



**HAL**  
open science

**Etude des interactions entre la peptidyl-prolyl cis/trans  
isomérase Pin1 et la protéine microtubulaire Tau.  
Recherche d'inhibiteurs ciblant la liaison de Pin1 à ses  
substrats phosphorylés**

Caroline Smet-Nocca

► **To cite this version:**

Caroline Smet-Nocca. Etude des interactions entre la peptidyl-prolyl cis/trans isomérase Pin1 et la protéine microtubulaire Tau. Recherche d'inhibiteurs ciblant la liaison de Pin1 à ses substrats phosphorylés. Biochimie [q-bio.BM]. Université du Droit et de la Santé - Lille II, 2004. Français. NNT: . tel-00354986

**HAL Id: tel-00354986**

**<https://theses.hal.science/tel-00354986>**

Submitted on 21 Jan 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Lille 2 – Droit et Santé  
U.F.R. de Médecine

## **THESE**

pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE DE LILLE 2**

Discipline *Chimie Organique*

Présentée et soutenue publiquement par

**Caroline SMET**

le 18 octobre 2004

**Etude des interactions entre la peptidyl-prolyl *cis/trans* isomérase Pin1 et la protéine microtubulaire tau. Recherche d'inhibiteurs ciblant la liaison de Pin1 à ses substrats phosphorylés.**

**Directeur de Thèse  
Guy LIPPENS**

### **JURY**

Docteur Luc BUEE  
Docteur Doris JACOBS  
Docteur Georges MASSIOT  
Docteur Ralf WARRASS  
Professeur Benoît DEPRez  
Docteur Guy LIPPENS

Président  
Rapporteur  
Rapporteur  
Examineur  
Examineur  
Directeur de Thèse



## Remerciements

J'aimerais adresser mes remerciements à l'ensemble des personnes qui ont contribué à ce travail qui a été aussi passionnant qu'enrichissant pour moi et qui a répondu à toutes mes attentes, et même au-delà.

Je remercie le Professeur Christian Sergheraert qui m'a accueillie à l'UMR 8525 et qui m'a accordé sa confiance.

Je tiens à remercier le Docteur Jacobs, le Docteur Massiot et le Docteur Warrass d'avoir accepté de s'investir dans cette thèse et de juger ce travail. Je remercie, à ce titre, l'ensemble des membres du jury.

Je remercie Guy Lippens, mon directeur de thèse, qui m'a fait découvrir la RMN des biomolécules et a contribué à enrichir cette thèse, en rendant ce travail toujours plus passionnant. Je me dois de souligner ici sa disponibilité et cet enthousiasme qui, semble-t-il, ne le quitte jamais. Je le remercie de m'avoir ouvert les portes du laboratoire de RMN, de sa confiance et de son soutien dans la recherche d'un post-doctorat.

Un merci tout particulier à Isabelle Landrieu, qui m'a transmis une partie de ses connaissances en matière de biochimie structurale, de la production de protéines recombinantes à la détermination des structures de protéines par RMN, sans oublier son investissement dans les diverses étapes de ce travail, ses conseils précieux et ses encouragements, à tous les moments qui constituent la vie d'un étudiant en thèse.

Je remercie Benoît Déprez, d'avoir accepté de faire partie du jury, de m'avoir accueillie dans son laboratoire à la Faculté de Pharmacie et d'avoir été présent pendant une grande partie de ma thèse. J'estime tout particulièrement son encadrement et ses conseils, notamment en matière de chimie médicinale et de *screening*.

Je remercie Luc Buée, ainsi que son équipe de l'Unité 422 de l'Inserm, pour leur contribution en matière de biologie cellulaire, avec une note particulière pour Luc Buée qui a participé à ce travail depuis le DEA, dont j'ai apprécié les conseils avisés sur tau et les cours improvisés de Bio. Cell., et qui sera présent, là encore, dans le jury.

Je remercie Jean-Michel Wieruszkeski et Gérard Montagne pour leur précieuse et incontournable expérience « RMNistique », mais aussi pour leur accueil chaleureux et leur sympathie. Je les remercie de tout cœur pour leur disponibilité à mon égard et le travail considérable qu'ils ont accompli et qui se cache derrière quasiment chaque expérience de cette thèse.

Je remercie l'incontournable Arnaud Leroy, qui m'a appris les bases de la fluorescence, pour son aide, ses conseils et sa bonne humeur communicative. Je le remercie également, ainsi qu'Alain Sillen, de m'avoir approvisionnée en protéine tau.

Merci à Oleg Melnyk qui m'a accueillie dans son équipe pour la synthèse des peptides.

Merci à Pierre Rousselot-Pailley qui m'a appris la synthèse peptidique.

Un grand merci à Eric Diesis et Hervé Drobecq, qui ont contribué au travail présenté ici, par leur soutien technique et leur disponibilité, sans oublier leur sympathie.

Merci à Jean-Frédéric Duckert pour la synthèse des composés à têtes polaires mimant le phosphate et son aide précieuse en synthèse organique.

Merci à Dragosh Horvath pour ses lumières en matière de modélisation moléculaire.

Merci à Gerhild Horvath de l'UMR8527 pour sa disponibilité et ses conseils pour les expériences d'ELISA, et à Marc Aumercier de l'UMR8526 pour m'avoir accueillie ponctuellement dans son laboratoire et pour son aide au cours des expériences de Biacore.

Merci à Xavier Hanouille, à Xavier Trivelli et à Benjamin Parent, mes compagnons de route du labo., pour leur gentillesse, leurs conseils et leur soutien psychologique.

Merci à Sylviane Piola et Amandine Menet, les deux stagiaires que j'ai encadrées au cours de ma thèse, pour leur contribution à la synthèse des peptides et les interrogations qu'elles ont soulevées qui m'ont permis d'avancer.

Un grand merci à l'ensemble des personnes constituant l'UMR8525, à la fois aux équipes de l'Institut de Biologie de Lille et aux équipes des laboratoires de Chimie Générale et de Chimie Organique de la Faculté de Pharmacie, pour leur accueil et leur gentillesse.

Et pour terminer, mais non les moindres, j'adresse mes remerciements les plus chaleureux à mes parents qui m'ont permis de réaliser mes projets et d'arriver là où j'en suis aujourd'hui, mais aussi à ma grand-mère et à ma tante qui ont toujours été présentes pour moi, mon oncle pour ses conseils et ses encouragements et évidemment Fabien, présent à chaque instant, qui m'a soutenue jusqu'au bout de cette thèse et sans doute encore au-delà.

*Je dédie cette thèse à mon papa...*

## Abstract

Phosphorylation is a major cellular process involved in regulation of protein function controlling inter-molecular associations, enzymatic activity, or ligand binding. Isomerisation of Ser/Thr-Pro peptidic bonds, phosphorylated by Pro-directed kinases, mostly involved in cell cycle control, stands as a novel regulation mode of protein function. Both signalling pathways are tightly related to through enzymes catalyzing the *cis/trans* isomerization process in phosphorylated Ser/Thr-Pro motifs as Pin1-type peptidyl-prolyl *cis/trans* isomerases (PPIases). Although Pin1 and homologous have an essential cellular role, their real molecular function remains still unclear. Molecular interactions between Pin1 and numerous phospho-proteins indicate a critical implication of Pin1 in cell cycle regulation and oncogenesis, and point out a relevant role as an emergent target in cancer treatment. Recently, interactions with the microtubule-associated Tau protein under its pathological hyperphosphorylated form, mediated by the unique phosphoThr231-Pro232 motif, could involve Pin1 in the regulation of Tau microtubule binding and in the Alzheimer's disease (AD) neuronal disorders.

We have targeted the interactions between Pin1 and Tau using small peptidic substrates to study the molecular processes that could explain the functional role of Pin1. Interactions with substrates through the phosphorylated Ser/Thr-Pro motifs are double : a WW binding domain is responsible for the substrate binding and the PPIase-type catalytic domain accelerates the *cis/trans* isomerization of proline bonds. A NMR screening of various phospho-Ser/Thr-Pro motifs encompassed in Tau protein points out a novel interacting site centered around the phosphoThr212-Pro213 motif, an AD-specific phosphorylation site.

We have extended investigations with Pin1 at the Tau full-length protein level. As for most of Pin1 substrates, Tau is characterized by a global absence of structuration that limits NMR studies. A regulatory role on enzymatic activity was shown for the WW domain through a peptidic fragment of 40 residues containing both Thr231 and Thr212 phosphorylated sites. Interactions with a mutant mimicking the phosphorylated form of Tau have been detected with the Pin1 catalytic domain involving an NMR assignment of the Tau protein based on peptide assignments so called "peptide mapping".

Phosphorylation of the Pin1 WW domain associated with inhibition of the binding function plays a regulatory role in Pin1 activity *in vivo*. The non-phosphorylated active form of the protein is mostly found in cancer cells and the phosphorylated inactive form in safe tissues.

We have targeted interactions between Pin1 and phospho-peptides based on the synthesis of small-molecules mimicking the phosphoThr-Pro dipeptide and performing high resolution NMR to investigate potential ligands targeting the Pin1 WW binding domain that could mimic the inactive form of the protein.

**Keywords** NMR spectroscopy, peptidyl-prolyl *cis/trans* isomerase, Pin1, Tau protein, biomolecular interactions, screening.

# SOMMAIRE

<b>Remerciements.....</b>	<b>3</b>
<b>Abstract .....</b>	<b>5</b>
<b>Abréviations .....</b>	<b>13</b>
<b>Introduction générale.....</b>	<b>15</b>
<b>Publications et communications.....</b>	<b>21</b>
<b>Chapitre I – Introduction. ....</b>	<b>23</b>
<b>Préambule - Introduction au cycle cellulaire et à sa régulation. ....</b>	<b>25</b>
<b>I – L’isomérisation des prolines, un nouveau mécanisme de régulation cellulaire de la fonction des protéines ?.....</b>	<b>27</b>
1.1 – Généralités autour de l’isomérisation <i>cis/trans</i> des prolines.....	27
1.2 – Catalyse de l’interconversion <i>cis/trans</i> . ....	31
1.3 – Trois classes de PPIases et des rôles biologiques différents. ....	34
<b>II – Pin1, une peptidyl-prolyl <i>cis/trans</i> isomérase essentielle dans la régulation du cycle cellulaire. ....</b>	<b>35</b>
2.1 – Une fonction biologique bien définie.....	35
2.1.1 – Le rôle de Pin1 dans le cycle cellulaire.....	37
2.1.2 – Le rôle de Pin1 dans la transcription.....	38
2.1.3 – Pin1, une cible thérapeutique émergente dans le traitement des cancers.....	40
2.2 – Un rôle moléculaire obscur. ....	42
2.2.1 – Structure de Pin1 : deux domaines indépendants ? .....	42
2.2.1.1 – Le domaine WW et la liaison du substrat, une fonction partagée par d’autres modules protéiques.....	43
2.2.1.2 – Le domaine catalytique : une structure et une spécificité de substrat différentes des autres PPIases. ....	50
2.2.1.3 – Indépendance structurale des deux domaines mais coopération fonctionnelle? .....	52
2.2.2 – Des relations entre structure et fonction de Pin1 mal définies. ....	54
2.2.2.1 – Pin1, une protéine phospho-spécifique?.....	54

2.2.2.2 – Reconnaissance des motifs phosphoSer/Thr-Pro par Pin1, un rôle moléculaire obscur. ....	56
<b>III – Les interactions de Pin1 avec la protéine microtubulaire tau. ....</b>	<b>61</b>
3.1 – La protéine tau : de la protéine microtubulaire au marqueur pathologique associé à la dégénération neuronale.....	61
3.1.1 – Fonction biologique de la protéine tau (Buee 2000). ....	61
3.1.2 – La protéine tau, marqueur moléculaire de la maladie d’Alzheimer. ....	66
3.2 – Des connaissances peu avancées sur les interactions entre les protéines Pin1 et tau..	72
3.2.1 – Rôle fonctionnel de Pin1 sur l’activité de tau phosphorylée.....	72
3.2.2 – Rôle moléculaire de Pin1 vis-à-vis de la protéine tau phosphorylée ?.....	75
3.2.3 – Pin1, une cible thérapeutique controversée dans le traitement de la maladie d’Alzheimer.....	76
<b>IV – Recherche d’inhibiteurs de la fonction de Pin1 .....</b>	<b>78</b>
4.1 – Rappels sur les inhibiteurs des parvulines.....	78
4.2 – Inhiber la liaison de Pin1 aux motifs phosphoSer/Thr-Pro. ....	80
<b>V – La RMN dans le processus de découverte de <i>hits/leads</i> pharmacologiques, une technique HTS parmi d’autres ?.....</b>	<b>81</b>
5.1 – Approche générale du criblage à haut débit par RMN.....	81
5.2 – La RMN dans les différentes étapes de la recherche d’un médicament.....	82
5.2.1 – Perturbations du signal de la cible.....	83
5.2.2 – Perturbations du signal du ligand. ....	84
5.2.3 – Suivi des interactions ligand-cible. ....	85
5.3 – Applications. Le criblage par RMN dans la recherche de composés actifs. ....	86
5.3.1 – L’approche par fragments. ....	86
5.3.2 – Expériences de compétition par RMN. ....	88
 <b>Chapitre II - Identification d’épitopes Ser/Thr-Pro phosphorylés de la protéine microtubulaire tau reconnus par la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1. ....</b>	 <b>91</b>
 <b>I – Criblage de phospho-épitopes de type Ser/Thr-Pro de la protéine tau reconnus par la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1.....</b>	 <b>93</b>

1.1 – L’outil peptidique pour le criblage d’épitopes phosphorylés de la protéine tau médiateurs des interactions avec la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1. ....	93
1.1.1 – Choix de phospho-épitopes de tau pertinents inclus dans des peptides. ....	93
1.1.2 – Synthèse de peptides phosphorylés par synthèse peptidique en phase solide.....	94
1.1.3 – Influence du pH sur l’état d’ionisation du phosphate. ....	98
1.2 – Bases thermodynamique et cinétique des interactions moléculaires.....	101
1.3 – Détermination des constantes d’affinité en milieu homogène. Utilisation de deux techniques spectroscopiques, la RMN et la fluorescence.....	105
1.3.1 – La RMN à haute résolution, un outil de détection des interactions de faible affinité. ....	105
1.3.1.1 – Calcul des constantes de dissociation et <i>mapping</i> des sites d’interaction par RMN à haute résolution. ....	109
1.3.1.2 – Etude par RMN du domaine WW isolé. Un modèle simplifié de Pin1 ?....	113
1.3.2 – Caractérisation de la cinétique de l’échange par RMN.....	116
1.3.3 – La fluorescence, une autre technique spectroscopique pour la détermination des constantes de dissociation en solution.....	132
1.3.3.1 – La fluorescence intrinsèque appliquée au domaine WW isolé, une information relative à l’environnement chimique.....	132
1.3.3.2 – Le FRET ( <i>Fluorescence Resonance Energy Transfer</i> ), une méthode plus sensible pour l’étude d’interactions biomoléculaires. ....	135
1.3.4 – Bilan des affinités des différents phospho-épitopes de la protéine tau dans la liaison à Pin1 obtenues en milieu homogène. ....	138
1.4 – Techniques de mesure de l’affinité des phospho-épitopes de tau pour Pin1 en milieu hétérogène. ....	144
1.4.1 – La résonance plasmonique de surface et la technologie Biacore®.....	145
1.4.2 – ELISA ( <i>Enzyme-Linked ImmunoSorbant Assay</i> ). ....	149
1.5 – Replacer les interactions Pin1-tau dans le contexte cellulaire. ....	154
<b>II – Le motif phospho-Thr212-Pro213, un nouveau site d’interaction entre la protéine tau et la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1 (Smet, Sambo et al. 2004).....</b>	<b>156</b>
2.1 – Le motif phospho-Thr212-Pro213 reconnu par le domaine WW de Pin1.....	156
2.2 - Etude conformationnelle autour du motif Thr212-Pro213 et isomérisation catalysée par la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1.....	156
2.2.1 – Détection de l’équilibre <i>cis/trans</i> des prolines par RMN à haute résolution et attribution de formes minoritaires. ....	157

2.2.2 – Influence de la phosphorylation sur l'équilibre <i>cis/trans</i> de la proline dans un motif Ser/Thr-Pro au niveau de substrats peptidiques. ....	160
2.2.3 – Catalyse de l'isomérisation <i>cis/trans</i> par Pin1 et son domaine catalytique isolé au niveau de substrats peptidiques. ....	162
2.3 – Validation au niveau de la protéine tau phosphorylée de l'épitope phosphoThr212-Pro213 comme site reconnu par Pin1.....	163
<b>III – Discussion. Méthodes de caractérisation des interactions biomoléculaires <i>in vitro</i> et découverte d'un nouvel épitope de tau impliqué dans l'interaction avec Pin1. ....</b>	<b>167</b>

<b>The Peptidyl Prolyl <i>cis/trans</i>-Isomerase Pin1 Recognizes the Phospho-Thr212-Pro213 Site on Tau (Smet, Sambo et al. 2004). ....</b>	<b>177</b>
---------------------------------------------------------------------------------------------------------------------------------------------	------------

## **Chapitre III – Tau, du peptide à la protéine. Etude des interactions avec la peptidyl-prolyl isomérase Pin1..... 187**

<b>I – Etude d'un fragment peptidique de 40 résidus comprenant deux sites d'interaction avec Pin1.....</b>	<b>189</b>
1.1 – Description du modèle peptidique.....	190
1.2 – Etude conformationnelle du fragment peptidique de 40 résidus de la région riche en prolines. ....	191
1.3 – Etude de <i>binding</i> avec la peptidyl-prolyl <i>cis/trans</i> isomérase Pin1. ....	194
1.3.1 – Etude avec la protéine Pin1 sauvage. ....	195
1.3.2 – Etude avec un mutant de Pin1 comportant une délétion de 6 résidus dans le <i>linker</i> inter-domaine. ....	198
1.4 – Modulation de l'activité catalytique <i>cis/trans</i> isomérase sur l'épitope Thr212/Ser214 par la présence d'un second phospho-épitope. ....	202
1.4.1 – Observation de l'interaction de Pin1 avec les sites phosphorylés du peptide....	202
1.4.2 – Spectroscopie d'échange appliquée à l'observation de l'interconversion <i>cis/trans</i> de la liaison phosphoThr212-Pro213. ....	204
1.4.3 – Inhibition de l'activité enzymatique autour de la liaison phosphoThr212-Pro213 en présence du site phosphoThr231-Pro232. ....	207
1.4.4 – Contrôles : rôle du domaine WW, rôle et position du second phospho-épitope.	209

1.5 – Bilan. Modulation de l'activité <i>cis/trans</i> isomérase de Pin1 autour de l'épitope Thr212/Ser214 dans des substrats multi-phosphorylés. ....	211
<b>II – Attribution de la protéine Tau par RMN et premières études d'interaction avec Pin1. ....</b>	<b>215</b>
2.1 – Attribution de la protéine Tau par « <i>mapping</i> peptidique ». ....	216
2.1.1 – Principe du « <i>mapping</i> peptidique ». ....	217
2.1.2 – Un outil pour la détection de structures locales. ....	222
2.1.3 – Confirmation de l'attribution par paires de résidus. ....	223
2.1.4 – Bilan. Avancées dans l'attribution de la protéine Tau avec le « <i>mapping</i> peptidique ». ....	226
2.2 – Etude des interactions entre Pin1 et un mutant de tau mimant l'état hyperphosphorylé. ....	228
2.2.1 – Le mutant Tau352E10 mime l'état hyperphosphorylé de la protéine tau (Eidenmuller, Fath et al. 2000). ....	228
2.2.2 – Investigations structurales du mutant Tau352E10 par RMN. ....	230
2.2.3 – Le domaine « catalytique » de Pin1 interagit avec le mutant Tau352E10. ....	234
<b>NMR peptide mapping of the neuronal Tau protein. ....</b>	<b>249</b>
<b>Chapitre IV - Conception, synthèse et criblage d'inhibiteurs ciblant la liaison de Pin1 aux motifs Ser/Thr-Pro phosphorylés. ....</b>	<b>265</b>
<b>I – Des inhibiteurs de Pin1 comme outil pour évaluer le rôle des interactions Pin1-tau dans la maladie d'Alzheimer. ....</b>	<b>267</b>
1.1 – Points de départ d'une stratégie visant à inhiber les interactions entre Pin1 et les motifs Ser/Thr-Pro phosphorylés. ....	267
1.1.1 – Cibler le domaine WW responsable de la liaison du substrat. ....	267
1.1.2 – Le dipeptide phosphoThr-Pro comme modèle pour la conception de ligands du domaine WW. ....	268
1.1.3 – Stratégie employée dans la recherche d'inhibiteurs de Pin1. ....	270
1.2 – Les outils mis en œuvre pour la synthèse et l'évaluation des ligands potentiels du domaine WW de Pin1. ....	272

1.2.1 – Synthèse en phase solide de ligands potentiels du domaine WW.....	272
1.2.2 – Choix d’une technique de criblage et de quantification de la liaison des composés au domaine WW de Pin1.....	273
1.2.2.1 – Compétition entre un phospho-peptide marqué et un ligand pour la liaison au domaine WW suivie par FRET.....	274
1.2.2.2 – La RMN : une technique mieux adaptée aux faibles affinités.....	281
<b>II – Synthèse de ligands du domaine WW basés sur la structure du motif phosphoThr-Pro et criblage par RMN.....</b>	<b>288</b>
2.1 – Mimes de proline.....	288
2.2 – Analogues de la phosphothréonine.....	297
2.2.1 – Etude préliminaire.....	297
2.2.2 – Analogues de type acide carboxylique : acide glutamique et dérivés.....	298
2.2.3 – Autres analogues de phosphate :diversité structurale de têtes polaires.....	301
2.3 – Introduction de motifs supplémentaires d’interaction avec le domaine WW.....	303
2.3.1 – Introduction d’un motif hydrophobe en position N-terminale.....	303
2.3.2 – Introduction d’un groupement accepteur de liaison hydrogène en position C-terminale.....	314
2.4 – Combinaison des éléments structuraux les plus favorables.....	316
2.5 – Tests cellulaires.....	318
<b>III – Discussion. La RMN comme outil de criblage pour des ligands de faible affinité. Application à la recherche d’inhibiteurs des interactions Pin1-tau.....</b>	<b>322</b>
<b>Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 and phosphopeptides (Smet, Duckert et al. soumis).....</b>	<b>329</b>
<b>Conclusion.....</b>	<b>341</b>
<b>Références.....</b>	<b>353</b>
<b>Matériels et Méthodes.....</b>	<b>361</b>
<b>Synthèse peptidique, biochimie des protéines, interactions biomoléculaires.....</b>	<b>367</b>
<b>Synthèse organique et caractérisation des composés. Criblage des composés par RMN et fluorescence.....</b>	<b>408</b>

## Abréviations

Ac	acétyl
Ac <sub>2</sub> O	anhydride acétique
APP	<i>amyloid precursor protein</i>
ATP	adénosine triphosphate
Boc	tertio-butoxycarbonyl
BSA	<i>bovin serum albumin</i>
Bn	benzyl
CD	<i>Circular dichroism</i>
CDK	<i>cyclin-dependent kinase</i>
CKS	<i>cyclin-dependent kinase subunit</i>
COSY	<i>correlation spectroscopy</i>
CPMG	Carr-Purcell-Meiboom-Gill
DCM	dichlorométhane
DIPEA	diisopropyléthylamine
DMF	diméthylformamide
DTT	dithiothreitol
DVB	divinylbenzène
EDT	ethanedithiol
EDTA	acide tétraacétiqueéthylènediamine
ELISA	<i>enzyme-linked immunosorbent assay</i>
FID	<i>free induction decay</i>
Fmoc	9-fluorénylméthoxycarbonyl
FT-IR	<i>Fourier transform infra-red spectroscopy</i>
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tétraméthyluronium hexafluorophosphate
HOBt	N-hydroxybenzotriazole
HSQC	<i>heteronuclear simple quantum correlation</i>
HTS	<i>High throughput screening</i>
IPTG	isopropyl-β-D-thiogalactose
K <sub>D</sub>	constante de dissociation
k <sub>on</sub>	constante de vitesse de la réaction de formation d'un complexe ( <i>on-rate</i> )
k <sub>off</sub>	constante de vitesse de la réaction de dissociation d'un complexe ( <i>off-rate</i> )
MALDI-TOF	<i>matrix assisted laser desorption ionisation - time of flight</i>
NFT	<i>neurofibrillary tangles</i>
NIMA	<i>never in mitosis A</i>
NOESY	<i>nuclear overhauser effect spectroscopy</i>
PBS	<i>phosphate buffer saline</i>
PEG	polyéthylène glycol
PHF	<i>paired helical filaments</i>
Pin1	<i>protein interacting with NIMA 1</i>
PPIase	peptidyl-prolyl <i>cis/trans</i> isomérase
ppm	partie par million
RMN	résonance magnétique nucléaire
RP-HPLC	<i>reverse phase - high pressure liquid chromatography</i>
RSA	relation structure/activité
SAXS	<i>Small-angle X-ray scattering</i>
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
STD	<i>saturation transfer difference</i>
TBS	<i>tris buffer saline</i>
TBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tétraméthyluronium tétrafluoroborate
TFA	acide trifluoroacétique
TIS	triisopropylsilane
TMSP	3-(triméthylsilyl)-2, 2, 3, 3-d <sub>4</sub> -propionate
TOCSY	<i>total correlation spectroscopy</i>
Tris	2-amino-2-hydroxyméthyl-1,3-propanediol

## Acides aminés usuels

<b>acide aminé</b>	<b>abréviation 3 lettres</b>	<b>abréviation 1 lettre</b>
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Acide aspartique	Asp	D
Acide glutamique	Glu	E
Cystéine	Cys	C
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Méthionine	Met	M
Phénylalanine	Phe	F
Proline	Pro	P
Sérine	Ser	S
Thréonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

## Introduction générale

### **Etude des interactions entre de la peptidyl-prolyl *cis/trans* isomérase Pin1 et la protéine microtubulaire Tau. Recherche d'inhibiteurs ciblant la liaison de Pin1 à ses substrats phosphorylés.**

Les mécanismes réactionnels mis en jeu dans le monde vivant sont très complexes et font appel à un grand nombre de processus différents au cours desquels de nombreux partenaires interagissent les uns avec les autres. Ces réactions, souvent essentielles au bon fonctionnement de la cellule, sont étroitement contrôlées par divers processus de régulation. La phosphorylation est un processus de régulation cellulaire majeur intervenant notamment dans le contrôle du cycle cellulaire. Elle fait partie des modifications dites post-traductionnelles qui interviennent au niveau des résidus Ser, Thr ou Tyr au sein des protéines et qui mettent en jeu l'ATP et un ensemble de kinases et phosphatases, enzymes coopérant pour assurer la régulation de l'état de phosphorylation des protéines substrats contrôlant ainsi leur fonction, leur localisation cellulaire.... Au niveau de certains motifs phosphorylés particuliers, les dipeptides Ser/Thr-Pro, ciblés notamment par les kinases contrôlant le cycle cellulaire (CDK/cycline...), une modification supplémentaire a été décrite dans certaines protéines. Un nouveau mécanisme de régulation émerge à partir de ces données récemment publiées concernant la reconnaissance de ces motifs par des enzymes appelées peptidyl-prolyl *cis/trans* isomérases (PPIases) qui isomérisent la liaison peptidique précédant la proline dans les motifs phosphoSer/Thr-Pro. Ce mécanisme ajouterait un point de régulation supplémentaire dans le contrôle du cycle cellulaire. D'ailleurs le prototype humain des PPIases, la protéine Pin1, apparaît comme essentiel. Contrairement à la phosphorylation qui se caractérise notamment par une modification covalente de la protéine, l'isomérisation de la proline est une modification conformationnelle où aucune liaison n'est formée ou rompue. L'isomérisation de la liaison peptidique précédant une proline représente un *switch* moléculaire intrinsèque qui pourrait être responsable du contrôle de l'activité enzymatique ou de la reconnaissance de ligands (Andreotti 2003). La régulation de ce processus serait assurée par les PPIases mais le mécanisme moléculaire et son impact fonctionnel dans la cellule ne sont pas encore connus à l'heure actuelle. Les PPIases sont généralement impliquées en tant que chaperonnes dans le processus de repliement des protéines natives lors de leur synthèse, processus cellulaire pour lequel l'isomérisation des prolines joue un rôle cinétiquement limitant. Elles sont regroupées en trois familles, dont celle des parvulines, décrites comme des

enzymes essentielles à la croissance cellulaire chez certains types de cellules à l'inverse des cyclophilines et des FKBP.

Pin1, une peptidyl-prolyl *cis/trans* isomérase humaine de la famille des parvulines, est notamment liée à la progression des cellules vers la mitose par sa spécificité pour les substrats phosphorylés. Pin1 a pour substrats de nombreuses phospho-protéines mitotiques, des facteurs de transcription phosphorylés et des protéines du cytosquelette comme la protéine tau, qui sera notre modèle dans l'étude du fonctionnement moléculaire Pin1 sur ses substrats. Pin1 se caractérise d'un point de vue structural par deux domaines assurant des fonctions différentes et reliés entre eux par un fragment peptidique flexible. Un petit module de liaison de type WW est responsable de la liaison des motifs phosphoThr/Ser-Pro et un domaine catalytique de 110 résidus isomériser les liaisons phosphoThr/Ser-Pro. Il a été proposé selon la théorie du « *tag and twist* » que les mécanismes de phosphorylation et d'isomérisation soient étroitement liés par l'intermédiaire des parvulines comme Pin1 et représentent un nouveau processus de signalisation cellulaire selon lequel les kinases phosphoryleraient les motifs Thr/Ser-Pro (*tag*) et ces mêmes motifs phosphorylés seraient ensuite reconnus par des peptidyl-prolyl *cis/trans* isomérases qui isomériseraient alors la liaison phospho-Thr/Ser-Pro (*twist*), conduisant à l'activation du substrat (Lu 2000).

Les interactions entre Pin1 et ses multiples phospho-épitopes restent obscures, en particulier parce que les motifs reconnus par Pin1 sont le plus souvent localisés dans des régions protéiques flexibles non structurées pour lesquelles l'intérêt d'un tel mécanisme de *switch* moléculaire n'est pas évident. C'est le cas de la partie régulatrice N-terminale de la phosphatase Cdc25 (Lu, Zhou et al. 1999; Zhou, Kops et al. 2000), de la protéine tau (Lu, Wulf et al. 1999), des facteurs de transcription p53 (Zacchi, Gostissa et al. 2002) et c-Jun (Wulf, Ryo et al. 2001) pour lesquels un ou plusieurs motifs d'interaction avec Pin1 de type phosphoSer/Thr-Pro ont été décrits mais toujours localisés dans ces régions flexibles qui jouent néanmoins un rôle essentiel de régulateur. Dans ce contexte, la résonance magnétique nucléaire est une technique particulièrement intéressante puisqu'elle permet l'observation directe de l'interconversion *cis/trans* des prolines et l'étude détaillée des mécanismes intervenant à l'échelle moléculaire, même dans des régions non structurées, impliquant l'activité isomérase de Pin1, et qui permettraient d'expliquer son rôle fonctionnel.

La protéine Pin1, outre son rôle dans le cycle cellulaire, serait impliquée dans la régulation de la liaison de la protéine neuronale tau aux microtubules. La phosphorylation de la protéine tau inhibe sa liaison aux microtubules et participe à la régulation physiologique des interactions entre tau et les microtubules. Une phosphorylation anormale de la protéine tau dans le contexte pathologique de la maladie d'Alzheimer, et d'autres maladies neurodégénératives appelées tauopathies, est impliquée dans les phénomènes de dégénérescence neuronale. La forme dite hyperphosphorylée de la protéine tau constitue un des marqueurs moléculaires de la maladie d'Alzheimer. L'interaction avec Pin1 permettrait de restaurer l'activité biologique de tau hyperphosphorylée (Lu, Wulf et al. 1999). L'implication de Pin1 dans les mécanismes de régulation de la liaison de tau aux microtubules et de la progression du cycle cellulaire pourrait ne pas être fortuite. Il est en effet proposé qu'une réactivation anormale du cycle cellulaire dans les neurones matures conduirait à la mort cellulaire et au processus de dégénérescence neurofibrillaire observés dans la maladie d'Alzheimer (Husseman, Nochlin et al. 2000; Delobel, Flament et al. 2002). A l'échelle moléculaire, les interactions entre Pin1 et tau sont peu avancées. Une étude a permis de mettre en évidence un site d'interaction unique centré autour de l'épitope phosphoThr231-Pro232 (Lu, Wulf et al. 1999) avec une affinité très élevée. Les connaissances structurales (Ranganathan, Lu et al. 1997; Bayer, Goettsch et al. 2003; Jacobs, Saxena et al. 2003) et les études de *binding* réalisées entre Pin1 et ses phospho-épitoques ont montré une interaction très limitée pour chacun des partenaires et des affinités modestes de l'ordre de quelques centaines de micromolaires (Verdecia 2000; Wintjens, Wieruszeski et al. 2001; Jacobs, Saxena et al. 2003) qui ne permettent pas d'expliquer la sélectivité et l'affinité de Pin1 pour l'épitope Thr231 de tau. Ces résultats structuraux ont permis d'établir un mode de liaison unique du domaine WW à ses multiples substrats, avec une spécificité pour les motifs phosphoSer/Thr-Pro en conformation *trans* (Wintjens, Wieruszeski et al. 2001), faisant intervenir une pince aromatique, constituée des chaînes latérales des résidus Trp34 et Tyr23, dans la liaison de la proline, et une boucle flexible constituée des résidus Ser16-Ser19 dans la liaison du groupement phosphate. Nous avons criblé plusieurs sites de phosphorylation de la protéine tau dont certains n'existent que dans les filaments hyperphosphorylés insolubles de la forme pathologique de la protéine, appelée PHF (*Paired Helicals Filaments*), caractéristiques de la maladie d'Alzheimer.

Les relations qui régissent les deux domaines de Pin1 sont encore mal définies et une éventuelle coopération inter-domaine, positive ou négative, pourrait être envisagée, en partie à cause de la flexibilité du fragment peptidique qui les relie. La dynamique qui régit le

comportement de Pin1 en solution montre une relative indépendance des deux domaines en l'absence de substrat (Bayer, Goettsch et al. 2003) et un rapprochement plus ou moins marqué selon le substrat, mais toujours très restreint (Jacobs, Saxena et al. 2003). Par ailleurs, les deux domaines sont essentiels à l'activité de Pin1 *in vivo* (Lu, Hanes et al. 1996). La phosphorylation du résidu Ser16 régule la fonction du domaine WW en inhibant la liaison des substrats et mène vers une délocalisation de Pin1 dans le cytosol alors que Pin1 est essentiellement nucléaire (Lu, Zhou et al. 2002). Une possible coopération fonctionnelle entre les deux domaines pourrait alors être envisageable pour expliquer le rôle moléculaire de Pin1 dans la régulation de la fonction biologique des phospho-protéines substrats. Nous avons donc tenté, à travers l'étude des interactions entre Pin1 et des peptides phosphorylés issus de la protéine tau, de comprendre les mécanismes qui ont lieu à l'échelle moléculaire.

Les investigations moléculaires à l'échelle peptidique, bien que très avantageuses, notamment dans le contexte de la reconnaissance de sites de phosphorylation, semblent limitées par la taille du substrat, le nombre de sites d'interaction potentiels et une totale absence de structure qui n'existe peut-être pas à l'échelle de la protéine entière. En particulier, l'isomérisation *cis/trans* catalysée par Pin1 n'a jamais été décrite à l'échelle d'une protéine entière. Des changements structuraux plus étendus peuvent également être envisagés. Ils pourraient expliquer les modulations fonctionnelles observées *in vivo* et être impliqués dans la régulation des interactions des substrats de Pin1 avec d'autres partenaires telle que l'activation de la déphosphorylation de tau et de Cdc25 par PP2A en présence de Pin1 (Zhou, Kops et al. 2000). Bien que la protéine tau ne semble présenter aucune structure globulaire stable, ce qui valide d'ailleurs en partie l'utilisation du modèle peptidique dans le criblage des phospho-épitopes de tau reconnus par Pin1 et l'étude de l'activité enzymatique de celle-ci sur de tels substrats, la présence de structures dites « résiduelles » n'est pas à écarter. Nous avons par conséquent étendu les études d'interaction entre Pin1 et tau à l'échelle de la protéine tau entière en utilisant la RMN. L'attribution de la protéine tau par RMN ne peut pas être réalisée avec les techniques conventionnelles d'attribution en raison d'une dégénérescence spectrale importante, inhérente aux protéines intrinsèquement non structurées. En combinant différentes approches fondées sur l'analyse des déplacements chimiques comparés aux valeurs « *random coil* », nous avons proposé une attribution de la protéine par identification de paires de résidus en exploitant la différence spectrale entre les isoformes de la protéine ainsi que des peptides courts. Les peptides, en particulier, dont l'absence de structure peut être facilement vérifiée, peuvent être utilisés comme sondes pour la détection de régions protéiques localement

structurées. La phosphorylation de la protéine tau *in vitro* par des kinases est problématique, notamment pour l'obtention de quantités suffisantes pour des études par RMN, et a justifié l'utilisation des phospho-peptides, dans un premier temps, pour déterminer les sites reconnus par Pin1. Nous avons contourné les problèmes de phosphorylation et commencé des études d'interaction avec Pin1 par RMN sur une protéine tau mutante contenant 10 mutations glutamates mimant l'état hyperphosphorylé de la protéine (Eidenmuller, Fath et al. 2000).

L'implication de Pin1 comme régulateur mitotique essentiel et l'aspect fondamental des interactions fonctionnelles entre Pin1 et les phospho-épitopes présents au sein de protéines étroitement impliquées dans le cycle cellulaire et dans l'oncogénèse, ont conduit à la validation de Pin1 comme agent de diagnostic de certains cancers et cible thérapeutique dans le traitement des cancers (Lu 2003; Ryo, Liou et al. 2003). Pin1 serait, en particulier, associée à l'expression de la cycline D1 (Liou, Ryo et al. 2002), surexprimée dans certains cancers. D'autre part, la forme active de Pin1, déphosphorylée au niveau du résidu Ser16 du domaine WW, est présente de manière prépondérante dans les cellules cancéreuses du sein alors que, dans les cellules saines, on retrouve majoritairement la forme phosphorylée inactive (Wulf, Ryo et al. 2001; Wulf, Ryo et al. 2003). Nous avons engagé une stratégie chimiothérapeutique avec la recherche d'antagonistes des interactions de Pin1 avec ses protéines substrats et fait le choix de cibler le domaine WW de Pin1 responsable de la liaison des phospho-épitopes, d'une part parce que la validation thérapeutique du domaine catalytique (chaperonne ou peptidyl-prolyl *cis/trans* isomérase) nécessiterait des études plus approfondies avant d'engager une stratégie et d'autre part, parce que, du point de vue biologique, l'inhibition du domaine WW par la liaison de petites molécules organiques s'apparenterait à la forme phosphorylée inactive retrouvée dans les tissus sains et mimerait la forme « normale » de Pin1. Le criblage de divers épitopes phosphorylés de la protéine tau par RMN et par fluorescence nous a permis de déterminer un motif structural de départ pour la conception et la synthèse d'inhibiteurs potentiels du domaine WW, sur la base du dipeptide phosphoThr-Pro, et d'adapter ces techniques au criblage rapide des molécules. La RMN, en particulier, a été mise en œuvre comme outil de criblage efficace pour la détection de faibles affinités et pour permettre une évaluation rapide du mode de liaison des ligands.



## Publications et communications

### **Articles.**

Smet, C., Duckert, J.-F., Wieruszeski, J.-M., Landrieu, I., Buée, L., Lippens, G., and Déprez, B. (2004) Control of protein-protein interactions: structure-based discovery of a low molecular weight inhibitors of the interactions between Pin1 and phosphopeptides. *J. Med. Chem. Lett.*, soumis.

Smet, C., Sillen, A., Leroy, A., Wieruszeski, J.-M., Landrieu, I., and Lippens, G. (2004) Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein. *Chembiochem*, accepté.

Smet, C., Sambo, A. V., Wieruszeski, J. M., Leroy, A., Landrieu, I., Buée, L., and Lippens, G. (2004) The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry*, **43**(7), 2032-40.

Lippens, G., Caron, L. and Smet, C. (2004) A microscopic view of chemical exchange : Monte Carlo simulation of molecular association. *Concepts Magn. Res. Part A*, **21A**(1), 1-9.

Lippens, G., Wieruszeski, J. M., Leroy, A., Smet, C., Sillen, A., Buée, L., and Landrieu, I. (2004) Proline-directed random-coil chemical shift values as a tool for the NMR assignment of the Tau phosphorylation sites. *Chembiochem*, **5**(1), 73-78.

Hamdane, M., Delobel, P., Sambo, A. V., Smet, C., Bégard, S., Violleau, A., Landrieu, I., Delacourte, A., Lippens, G., Flament, S., and Buée, L. (2003) Neurofibrillary degeneration of the Alzheimer-type: an alternate pathway to neuronal apoptosis? *Biochem. Pharmacol.*, **66**(8), 1619-1625.

Hamdane, M., Smet, C., Sambo, A. V., Leroy, A., Wieruszeski, J. M., Delobel, P., Maurage, C. A., Ghestem, A., Wintjens, R., Bégard, S., Sergeant, N., Delacourte, A., Horvath, D., Landrieu, I., Lippens, G., and Buée, L. (2002) Pin1: a therapeutic target in Alzheimer neurodegeneration. *J. Mol. Neurosc.*, **19**(3), 275-287.

### **Communications.**

#### Communications orales.

Smet, C., Wieruszeski, J. M., Landrieu, I., Buée, L., and Lippens, G. (2003) La RMN, un outil pour la détection des interactions de faible affinité. 3<sup>e</sup> Colloque annuel des doctorants, Lille, France.

Smet, C., Wieruszeski, J. M., Leroy, A., Landrieu, I., Buée, L., and Lippens, G. (2003) Etude des interactions entre la peptidyl-prolyl cis/trans isomerase Pin1 et la protéine microtubulaire Tau. Congrès RMN biologique et les enjeux du vivant, Société Française de Biochimie et de Biologie Moléculaire, 12-14 mars 2003, Clermont-Ferrand, France.

Lippens, G., Landrieu, I., Odaert, B., Wintjens, R., Smet, C., Wieruszeski, J. M., Buée, L. (2002) NMR study of CKS and Pin1, two proteins involved in the eukaryotic cell cycle.

Congrès La RMN, un outil pour la biologie V, Institut Pasteur de Paris, 4-6 février 2002, Paris, France.

Lippens, G., Smet, C., Wieruszeski, J. M., Leroy, A., Sillen, A., Sambo, A.-V., Bégard, S., Buée, L., and Landrieu, I. (2003) NMR interaction study of the prolyl *cis/trans* isomerase Pin1 and the microtubule-associated protein Tau. Obernai, France.

Communications par affiche.

Smet, C., Landrieu, I., Wieruszeski, J.-M., Leroy, A., Duckert, J.-F., Sambo, A.-V., Déprez, B., Buée, L. and Lippens, G. (2004) Screening of tau phospho-epitopes recognized by the human prolyl *cis/trans* isomerase Pin1. Inhibition of Pin1-phospho-epitope interactions with small-molecule ligands of the Pin1 WW binding module. EENC/Ampère Meeting on Magnetic Resonance, 6-11 septembre 2004, Lille, France.

Smet, C., Landrieu, I., Wieruszeski, J.-M., Leroy, A., Duckert, J.-F., Sambo, A.-V., Déprez, B., Buée, L. and Lippens, G. (2004) Targeting protein-protein interaction : Inhibition of Pin1–phospho-epitope interactions with small-molecule ligands of the Pin1 WW binding module. 40<sup>e</sup> Rencontres Internationales de Chimie Thérapeutique, Société de Chimie Thérapeutique, 30 juin-2 juillet 2004, Bordeaux, France.

Smet, C., Landrieu, I., Wieruszeski, J.-M., Leroy, A., Duckert, J.-F., Sambo, A.-V., Déprez, B., Buée, L. and Lippens, G. (2004) Peptide substrates for phospho-epitope screening and characterization of catalytical activity of the human prolyl *cis/trans* isomerase Pin1 by NMR. Design, synthesis and NMR screening of WW potential inhibitors. Workshop, Molecules as modulators: systems biology challenges chemistry, Aventis / Frankfurt University, 29-31 janvier 2004, Wiesbaden, Allemagne.

Landrieu, I., Leroy, A., Smet, C., Wieruszeski, J.-M., Wintjens, R., Sambo, A.-V., Inzé, D., Buée, L., and Lippens, G. (2003) Proline conformation as a regulatory element of molecular function, FEBS Meeting on Signal Transduction, 3-8 juillet 2003, Bruxelles, Belgique.

**CHAPITRE I**  
**INTRODUCTION**



## **Préambule - Introduction au cycle cellulaire et à sa régulation.**

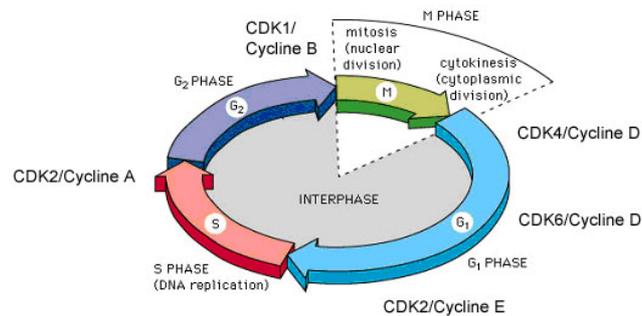
Un organisme eucaryote est formé de cellules. Chacune de ces cellules provient de la multiplication d'une cellule préexistante. Le cycle cellulaire est constitué principalement de deux phases : l'interphase et la mitose. Pendant l'interphase a lieu la réplication semi-conservative de l'ADN, tandis que la mitose est la division cellulaire. Elle permet de multiplier le nombre de cellules en conservant la quantité et la qualité des informations contenues dans les chromosomes. Le déroulement de la mitose dépend de mécanismes très précis. Ces mécanismes sont universels, c'est-à-dire communs, à quelques détails près, à tous les organismes eucaryotes, animaux et végétaux.

Le cycle cellulaire est divisé en 4 phases :  $G_1$ , S,  $G_2$  et M auxquelles on peut ajouter la phase  $G_0$  ou phase de quiescence. La régulation de la division cellulaire apparaît d'une très grande complexité car il existe des cascades de réactions interdépendantes dont il est difficile de situer le point de départ. Schématiquement la mitose, c'est-à-dire la division elle-même, est précédée des phases  $G_1$ , S et  $G_2$  au cours desquelles la cellule accumule, après mise en jeu de diverses transcriptions, le matériel nécessaire à la naissance de deux cellules-filles. On peut considérer que, pendant les phases  $G_1$  et S, il y a activation des gènes conduisant à la synthèse de protéines nécessaires à la synthèse d'ADN et pendant la phase  $G_2$ , synthèse de protéines nécessaires à la séparation et à la migration des chromosomes. Les phases sont définies de la manière suivante :

- phase  $G_1$  (G pour *gap*), phase précédant la phase S
- phase S ou phase de synthèse, au cours de laquelle la réplication de l'ADN s'effectue grâce à l'ADN polymérase de type III.
- phase  $G_2$  pré-mitotique
- phase M ou mitotique, c'est la phase de division cellulaire proprement dite.

Le cycle cellulaire est régulé par des protéines kinases cycline-dépendantes ou CDK (*Cyclin-Dependent Kinases*), c'est-à-dire que l'activité enzymatique de phosphorylation par ces kinases nécessite la présence d'une autre protéine appelée cycline. Chez l'Homme, on distingue plusieurs cyclines et plusieurs CDK. Les cyclines sont ainsi appelées parce qu'il y a une variation cyclique de leur concentration intracellulaire au cours des différentes phases du cycle cellulaire. Les cyclines D, liées aux CDK4 et CDK6, interviennent à la phase  $G_1$  et la cycline E intervient aux phases  $G_1$  et S. Les cyclines A, dont la concentration s'élève du début de la phase S jusqu'au début de la mitose, liées soit à CDK2 soit à Cdc2 (partie catalytique

d'une protéine kinase dépendant de la cycline A), interviennent aux phases S et G<sub>2</sub>, et les cyclines B, liées aux CDK1 et CDK2, aux phases G<sub>2</sub> et M.



**Figure 1.1** : Schéma du cycle cellulaire avec représentation des quatre phases : G<sub>1</sub>, S (phase de synthèse des acides nucléiques), G<sub>2</sub>, M (mitose) et des différents complexes CDK/cycline.

Il existe d'autres mécanismes de régulation des CDK :

- leur (dé)phosphorylation : la CDK1 est ainsi inhibée par la phosphorylation de la Thr14 et Tyr15, tandis que son activité maximale nécessite la phosphorylation de la Thr160. La protéine Cdc25, une phosphatase présentant une dualité pour les résidus Ser/Thr et Tyr phosphorylés, joue également un rôle essentiel en régulant la déphosphorylation des CDK. Cdc25 et le complexe CDK/cycline font partie d'un mécanisme de régulation positif par «*feed-back*», qui permet de déclencher le processus d'activation vers la mitose.
- la présence ou absence d'inhibiteurs spécifiques : les protéines p21 ou CIP1, qui inhibent les complexes cycline/CDK associés avec la phase G<sub>1</sub>, et dont l'expression elle-même est fortement régulée par la protéine p53 ; p27(Kip1) est un autre exemple, qui inhibe l'activation des complexes cycline E/CDK2 et cycline D/CDK4, et contrôle ainsi également la transition G<sub>1</sub>/S.
- la présence de petites sous-unités qui aident au transport des complexes CDK vers le substrat, telles que les CKS (*Cyclin-dependent Kinase Subunit*). Ces sous-unités agissent comme des modules d'interaction qui s'associent d'une façon non-covalente aux complexes CDK, en contraste avec d'autres modules tels que les domaines SH2 et SH3 qui font souvent partie intégrante de la séquence primaire des kinases multi-domaines.

Les substrats des complexes CDK sont majoritairement composés par le motif Ser/Thr-Pro. La structure cristalline d'un tel complexe montre, en outre, que la liaison peptidique doit être

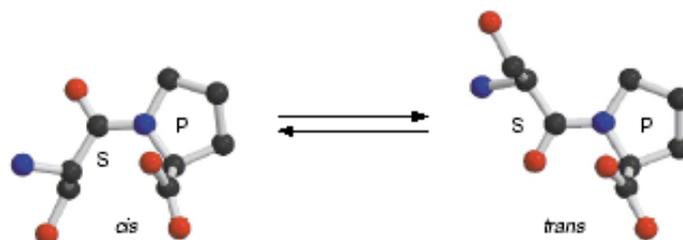
en *trans* pour une liaison optimale du peptide substrat. Ces motifs phosphorylés sont reconnus au sein des phospho-protéines mitotiques par l'anticorps monoclonal MPM-2, généré à partir d'extraits mitotiques totaux de cellules HeLa (Davis, Tsao et al. 1983). Récemment, un nouveau mécanisme de régulation a été proposé impliquant la conformation des prolines dans ces motifs Ser/Thr-Pro phosphorylés. De plus, ces épitopes sont également reconnus par des protéines appelées peptidyl-prolyl *cis/trans* isomérases (PPIases) dont le prototype humain, Pin1 (*Protein Interacting with NIMA 1*), a été découvert à l'issue d'un criblage génétique chez la levure recherchant des partenaires de la kinase NIMA, une protéine essentielle du cycle cellulaire (Lu, Hanes et al. 1996). La régulation de la conformation de la proline dans ces motifs par les PPIases a été proposée comme un nouveau mécanisme de régulation du cycle cellulaire, consécutif au processus de phosphorylation (Lu 2000). Sur le plan moléculaire, hormis pour les modèles peptidiques, peu de données structurales sur des protéines permettent de comprendre le rôle de ces enzymes dans la fonction biologique des protéines substrats.

## I – L'isomérisation des prolines, un nouveau mécanisme de régulation cellulaire de la fonction des protéines ?

### 1.1 – Généralités autour de l'isomérisation *cis/trans* des prolines.

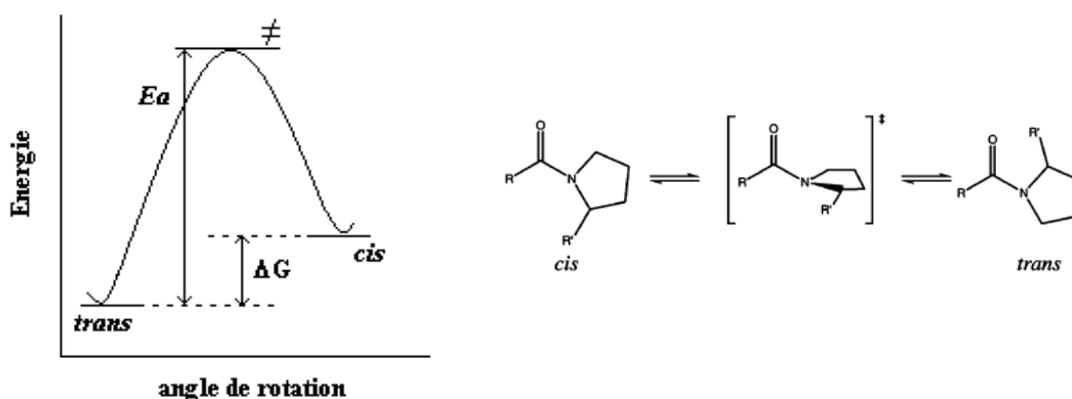
#### *Interconversion cis/trans et signalisation cellulaire.*

La proline est le seul acide aminé dont la chaîne latérale est comprise dans un cycle avec la chaîne principale. La fonction amine est alors incluse dans un cycle à 5 chaînons saturés (cycle pyrrolidine). Cette structure confère à la proline la remarquable propriété d'être représentée au sein des protéines sous deux conformations limites stables, les conformères *cis* et *trans*, par simple rotation autour de la liaison peptidique qui la précède (Figure 1.2).



**Figure 1.2** : Isomérisation de la liaison peptidique précédant la proline. Représentation de l'équilibre conformationnel *cis/trans* d'un motif Ser-Pro.

Alors que cette rotation de la liaison peptidique existe quels que soient les résidus qui la constituent et pour laquelle la seule forme stable est la forme *trans*, les liaisons X-Pro se caractérisent par une barrière énergétique faible concernant la transition de la forme *cis* vers la forme *trans* et *vice versa*, rendant l'isoforme *cis* accessible d'un point de vue énergétique. Le mécanisme de l'isomérisation est actuellement bien décrit dans le cas de petits peptides. En effet, pour des peptides courts contenant une proline, les données indiquent un passage par un état de transition via un mécanisme en une seule étape où le cycle pyrrolidinique de la proline est plan (Figure 1.3). L'interconversion *cis/trans* est donc régie par un équilibre thermodynamique défini par une enthalpie  $\Delta G$  ou une constante d'équilibre  $K$  ( $\Delta G = -RT \log K$ ) et une vitesse d'interconversion liée au coût énergétique, appelé énergie d'activation ( $E_a$ ), nécessaire pour franchir l'état de transition (Figure 1.3). La barrière énergétique d'interconversion entre les deux conformères limites est généralement comprise entre 10 et 30 kcal/mol, et la vitesse d'échange entre quelques secondes et quelques minutes (à 25°C), rendant ce processus cinétiquement limitant pour le repliement des protéines *in vitro*.



**Figure 1.3** : Mécanisme de l'isomérisation de la proline dans un motif X-Pro : interconversion *cis/trans* avec représentation de l'état de transition ( $\neq$ ) et diagramme d'énergie de l'isomérisation *cis/trans*.

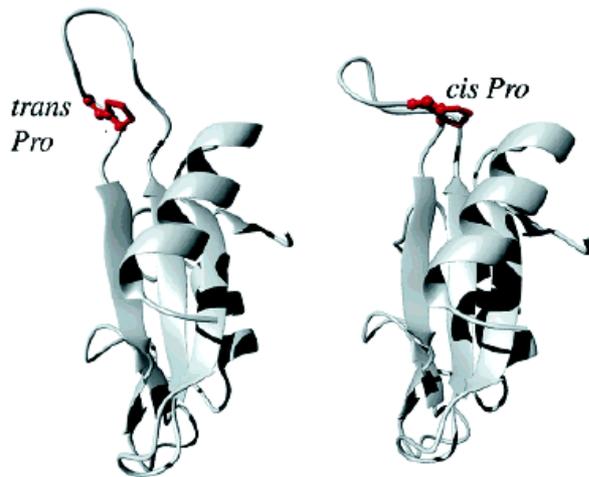
L'isomérisation aboutit à une hétérogénéité d'une ou plusieurs liaisons peptidiques impliquant une proline au sein de protéines natives structurées mais, dans la plupart des cas, il n'existe pas d'hétérogénéité conformationnelle (pour autant que l'on puisse la détecter) et les prolines adoptent soit une conformation *trans*, soit une conformation *cis*. Cette dernière est moins favorable dans le cas de petits peptides pour lesquels on trouvera 5 à 10% de formes *cis* mais peut être favorisée dans le contexte d'une protéine native, où elle sera stabilisée par la

structure globale de la protéine. Si l'isomérisation est bien caractérisée dans le cas de peptides courts au niveau moléculaire, peu de données structurales permettent de comprendre l'influence d'un tel mécanisme sur la fonction de protéines globulaires, à une exception près, celle de l'Itk, décrite dans le paragraphe suivant.

***Un exemple illustrant la modulation de la reconnaissance de ligands liée à l'interconversion cis/trans dans une protéine globulaire (Andreotti 2003).***

De manière générale, on note que la fréquence des prolines est élevée dans les boucles flexibles et que la conformation *trans* est préférentiellement exposée au solvant tandis qu'en conformation *cis*, la boucle forme des interactions intramoléculaires hydrophiles ou hydrophobes avec le corps de la protéine, stabilisant ainsi cette forme thermodynamiquement moins stable. Dans le cas du domaine SH2 de l'Itk (*Interleukine 2 tyrosine kinase*) par exemple, une boucle flexible intervenant dans la liaison du substrat comporte une liaison Asn286-Pro287 pour laquelle les deux conformations *cis* et *trans* coexistent à 40 et 60% respectivement. Les deux structures ont pu être résolues, montrant une structure étendue de la boucle dans la conformation *trans*, plus exposée au solvant (Figure 1.4). L'hétérogénéité conformationnelle concerne le plus souvent des régions hautement flexibles dans le contexte des protéines natives, par opposition aux régions hautement ordonnées. Cependant, dans certains cas, on a pu observer une proline hétérogène dans une structure secondaire stable.

L'hétérogénéité conformationnelle locale due à l'isomérisation d'une proline peut être plus ou moins étendue. Au niveau du domaine SH2 de l'Itk, des perturbations apparaissent dans la totalité de la boucle flexible mais concernent également 30% des résidus du corps replié de la protéine, modulant ainsi la reconnaissance de ligands par modification de la surface de liaison du ligand. Parfois, ces perturbations restent très localisées et ne concernent que quelques résidus entourant la proline. Dans l'exemple du domaine SH2 de l'Itk, cette hétérogénéité conformationnelle importante est impliquée dans un processus de signalisation et permet à la protéine en conformation *trans* de lier un phospho-peptide tandis qu'en conformation *cis*, elle interagit avec le domaine SH3 de l'Itk. L'isomérisation de la Pro287 est donc capable de modifier la surface moléculaire impliquée dans la liaison du phospho-peptide.



**Figure 1.4 :** Un exemple de l'effet de l'isomérisation de la proline dans la structure globulaire du domaine SH2 de l'Itk au niveau d'une boucle flexible. Dans la conformation *trans*, la boucle est plus exposée au solvant (d'après Andreotti, *Biochemistry* 2003).

Il faut également souligner que, dans certains cas, la proportion de formes *cis* étant faible (moins de 20%), la résolution de la structure dans la conformation *cis* ne peut être établie, ce qui freine la compréhension des phénomènes liés à la modulation de la fonction des protéines par le processus d'isomérisation. C'est le cas notamment des protéines non structurées ou présentant des fragments non repliés, dans lesquelles la proportion de formes *cis* est souvent inférieure à 10%, comme dans le cas de peptides courts. Si l'exemple décrit précédemment met en évidence au niveau moléculaire le rôle de l'isomérisation sur la fonction et la reconnaissance de ligands au niveau d'une protéine structurée, peu d'études permettent d'expliquer l'intérêt d'un tel processus sur des protéines hautement flexibles, totalement ou partiellement non structurées, pour lesquelles on s'attend à peu de modifications conformationnelles, en tous cas, à longue distance.

#### ***Pourquoi l'isomérisation des prolines est-elle si difficile à caractériser ?***

Dans la plupart des modifications post-traductionnelles, comme la phosphorylation ou la glycosylation, l'addition d'un ou plusieurs groupements chimiques module la fonction et/ou les interactions de la biomolécule. La plupart des techniques biochimiques (spectroscopie de masse, digestion enzymatique, reconnaissance par des anticorps, tests fonctionnels) sont parfaitement adaptées au suivi de telles modifications, en raison de leur sensibilité et de leur capacité à déterminer la localisation précise du site modifié. En revanche, des modifications post-traductionnelles qui sont de nature purement conformationnelle, avec conservation stricte de la composition chimique de la protéine, comme la conformation des prolines, ne sont pas détectables par ce type de techniques. L'absorption UV permet de détecter les deux conformères *cis* et *trans*. La mesure d'absorbance doit être précédée d'un brutal changement de pH pour altérer l'équilibre *cis/trans*, mais la gamme des pH pratiqués est souvent

incompatible avec les conditions permettant à la protéine de garder sa structure globulaire native. Les méthodes de choix pour détecter et caractériser les modifications conformationnelles sont les techniques structurales, comme la RMN et la cristallographie par rayons X. L'équilibre *cis/trans* des prolines étant relativement lent (de l'ordre de la minute) sur l'échelle de temps RMN, il est possible d'observer directement les deux conformations. Des contacts NOE caractéristiques entre les protons de la proline et de l'acide aminé précédent permettent d'attribuer sans ambiguïté les deux conformères, et l'intégration de leur signaux respectifs permet de quantifier les populations relatives (nous illustrerons ces aspects plus largement dans le chapitre II). La résolution de structures par cristallographie permet également de détecter des hétérogénéités conformationnelles. Cependant, les conditions de cristallisation peuvent sélectionner un conformère par rapport à l'autre, et l'observation des deux conformères dans les mêmes conditions de cristallisation est souvent fortuite. De manière générale, la structure des protéines obtenue par cristallisation est biaisée par les conditions mêmes de cristallisation et ne permet pas d'observer, dans la plupart des cas, une hétérogénéité conformationnelle, mais seulement l'un des deux conformères. De plus, dans le cas où la proline est localisée dans une région flexible, la résolution pauvre obtenue dans ces régions ne permet pas de détecter une isomérisation. Les techniques biophysiques qui permettent de déterminer la présence de structures secondaires (infra-rouge, dichroïsme circulaire) et de quantifier leur proportion dans la structure globale, ne sont pas suffisamment sensibles pour détecter une conformation locale autour d'une seule liaison peptidique.

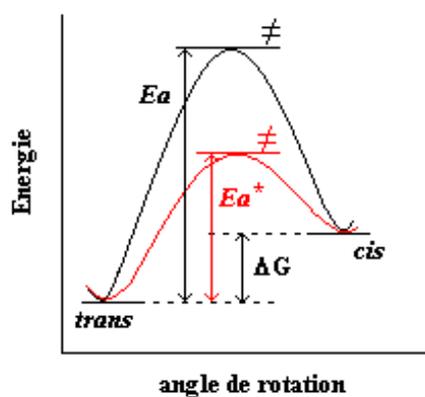
### **1.2 – Catalyse de l'interconversion *cis/trans*.**

Au sein de la cellule, les peptidyl-prolyl *cis/trans* isomérases sont impliquées dans le repliement des protéines après leur synthèse dans les ribosomes et jouent un rôle de chaperonnes moléculaires. Parmi ces protéines chaperonnes, on retrouve les protéines de choc thermique et les disulfures isomérases qui catalysent la formation et la rupture des ponts disulfures, permettant aux protéines d'acquérir leur structure tridimensionnelle correcte dont dépendra leur activité. Les PPIases font partie de ces protéines chaperonnes mais suivent un processus catalytique différent en accélérant l'isomérisation *cis/trans* des liaisons précédant une proline (liaisons X-Pro), processus intrinsèquement lent qui peut limiter la cinétique de repliement des protéines natives. L'acquisition d'une structure globulaire est essentielle (pour certaines protéines) non seulement pour assurer la fonction cellulaire requise, mais également pour ralentir le processus de dégradation métabolique des protéines par les protéases. Les PPIases sont des enzymes ubiquitaires, également impliquées dans la translocation des

protéines, ainsi que dans la transduction des signaux (Fischer, Bang et al. 1984; Schmid 1993; Schmid 2001).

### ***Description du rôle moléculaire des PPIases.***

Les peptidyl-prolyl *cis/trans* isomérases sont des protéines exprimées dans les cellules procaryotes et eucaryotes. Leur rôle premier est de lier et de stabiliser l'état de transition, permettant ainsi de compenser le coût énergétique de la distorsion de la liaison peptidique par une interaction favorable d'un point de vue énergétique dans le complexe PPIase/substrat. Cette interaction permet de diminuer l'énergie de l'état de transition ( $E_a^*$ ) et donc d'accélérer l'interconversion *cis/trans* (Figure 1.5).



**Figure 1.5 :** Diminution de l'énergie d'activation nécessaire à l'interconversion *cis/trans* par la stabilisation de l'état de transition ( $\ddagger$ ) par les peptidyl-prolyl *cis/trans* isomérases (courbe rouge).

Les PPIases ont été divisées en trois familles sur base de leur spécificité vis-à-vis d'un ligand : les cyclophilines, les FKBP et les parvulines. Leur fonction première, concernant l'isomérisation de la liaison peptidique précédant une proline, a surtout été montrée dans des modèles peptidiques courts. En revanche, l'action catalytique des PPIases sur des protéines structurées n'a pas été beaucoup étudiée à l'échelle moléculaire jusqu'à présent. D'autre part, si la fonction biologique des parvulines a été bien décrite en raison notamment de leur caractère essentiel, la fonction biologique des cyclophilines et des FKBP reste encore mal connue. Peu de données montrent l'activité *cis/trans* isomérases des PPIases, de manière directe, au niveau de protéines substrats structurées. L'activité *cis/trans* isomérase de la cyclophiline a été identifiée *in vitro* au niveau moléculaire dans le processus de repliement la ribonucléase T1 en cours de synthèse et ainsi associée au rôle de chaperonne (Schmid, Frech et al. 1996). Le plus souvent, dans le cas où on met en présence une PPIase et une protéine structurée présentant au moins une proline hétérogène du point de vue conformationnel, il n'est pas clairement mis en évidence si c'est l'isomérisation intrinsèque de la proline ou l'effet de l'isomérisation catalysée par les PPIases qui joue un rôle fonctionnel prépondérant, et si c'est au travers de cette proline particulière que se fait la reconnaissance avec la PPIase.

Le principal problème est de distinguer, au niveau cellulaire, la liaison de l'activité enzymatique.

### ***Caractérisation de la catalyse de l'isomérisation cis/trans des prolines.***

La première PPIase a été découverte en 1984 (Fischer, Bang et al. 1984). Un test permettant l'identification des PPIases potentielles est basé sur la haute sélectivité de la chymotrypsine pour les substrats du type X-Pro-Phe-pNA (X étant un acide aminé et pNA, la *para*-nitroaniline). La coupure du groupement pNA par la chymotrypsine se produit uniquement quand la liaison X-Pro du substrat est en conformation *trans*. Cette méthode est couramment utilisée pour détecter une activité PPIase ainsi que pour déterminer la spécificité d'une PPIase concernant le substrat par modification du résidu X. D'autres techniques peuvent être utilisées pour mettre en évidence une activité isomérase, comme une modification de la sensibilité aux protéases, de la reconnaissance par des anticorps, d'une activité enzymatique..., mais sont basées sur une détection indirecte de l'effet isomérase, ou plutôt, d'une détection d'une perturbation en présence de ladite PPIase qui pourrait provenir soit effectivement de la fonction isomérase, soit d'un autre effet, comme une simple liaison par exemple.

La catalyse de l'isomérisation des prolines et son rôle biologique restent flous. Bien que l'on puisse observer de manière directe l'isomérisation enzymatique des PPIases, cela reste limité aux substrats peptidiques uniquement. En parallèle, on a décrit de manière indirecte l'effet de l'isomérisation catalytique *in vivo* dans un certain nombre de processus biologiques, en raison de l'observation de modifications de cet effet biologique en présence de PPIases, caractérisées par des tests *in vitro* réalisés sur des peptides. Mais l'implication de la fonction isomérase ne serait-elle pas abusive, étant donné qu'aucune donnée structurale ne vient appuyer ces observations fonctionnelles ? Ainsi, le mode d'action des PPIases a été remis en question et on envisage un effet de *binding* plutôt qu'un effet catalytique (Fischer, Tradler et al. 1998; Ivery 2000), l'effet enzymatique sur l'isomérisation *cis/trans* pouvant être un artefact dû à l'utilisation de substrats peptidiques pour détecter cet effet. A partir des données biologiques *in vivo*, il est en effet, difficile de définir si les PPIases jouent le rôle d'enzymes ou de modules de liaison, intervenant comme médiateurs des interactions protéine-protéine.

La spectroscopie d'échange par RMN permet de détecter, de manière directe, l'accélération de l'isomérisation *cis/trans* d'une proline hétérogène en présence d'une PPIase mais l'effet est difficile à détecter notamment dans le cas de substrats protéiques. Bien qu'un tel effet de la cyclophiline A ait déjà été montré *in vitro* sur une protéine substrat, le domaine N-terminal de la protéine de capsid constituant le virus HIV-1 (*human immunodeficiency virus-type 1*),

l'importance de la catalyse de l'isomérisation de la Pro90 (86% en conformation *trans*) par la cyclophiline A sur la fonction biologique reste à déterminer. On peut déjà souligner à ce stade que de telles informations structurales sont encore pour le moment inexistantes avec les parvulines.

### 1.3 – Trois classes de PPIases et des rôles biologiques différents.

Les trois familles de PPIases ne partagent aucune similitude au niveau de leur structure tridimensionnelle et se distinguent par leur ligand respectif. Les cyclophilines tirent leur nom de l'affinité qu'elles possèdent pour la cyclosporine A, et les FKBP (*FK506-binding proteins*) de la molécule FK506. Enfin, les parvulines tirent leur nom du prototype de cette troisième famille de PPIases, la parvuline (du latin, *parvulus*, qui signifie très petit), isolée chez *E. coli* (Rahfeld, Rucknagel et al. 1994) et constituent une famille à part entière, car elles ne lient ni la cyclosporine A, ni le FK506, mais sont inhibées sélectivement par la juglone (Hennig, Christner et al. 1998), un composé naturel isolé des feuilles de noyer et des écorces de noix.

Les cyclophilines et les FKBP, que l'on regroupe sous le terme d'immunophilines, ont en commun un rôle dans le système immunitaire lors de la formation du complexe cyclophiline/CsA ou FKBP/FK506. Ces complexes ont permis d'expliquer certaines voies de transduction des signaux impliquant les cellules T mais la fonction cellulaire normale des immunophilines reste obscure. Il a en effet été montré que CsA et FKBP-12 agissent comme agents immunosuppresseurs en bloquant l'activation des cellules T. Cependant, cette activité ne provient pas de l'inhibition de l'activité isomérase mais de l'interruption de signaux au sein de mécanismes transductionnels par le complexe PPIase/ligand qui est le véritable inhibiteur. Des immunosuppresseurs tels que la CsA peuvent se lier aux cyclophilines, inhibent leur liaison à d'autres substrats et induisent une réponse immunitaire. Cependant, on sait aujourd'hui que les immunophilines agissent au niveau immunitaire en séquestrant la calcineurine, une phosphatase dépendant du complexe  $Ca^{2+}$ /calmoduline et n'ont pas de fonction directe spécifique dans le système immunitaire.

Malgré leur grande conservation au cours de l'évolution, les cyclophilines et les FKBP ne sont pas essentielles à la vie cellulaire. En revanche, les parvulines sont essentielles à la progression du cycle cellulaire chez la levure *Saccharomyces cerevisiae* (Ess1/Ptf1) et chez les mammifères (Pin1 et homologues) (Lu, Hanes et al. 1996). Elles ne sont pas essentielles par exemple, chez la souris en raison de redondances génétiques (Fujimori, Takahashi et al. 1999), ou la drosophile. Pin1 est étroitement impliquée dans de la régulation du cycle

cellulaire en raison de sa spécificité pour les épitopes MPM-2 (motifs phosphoSer/Thr-Pro) présents dans les substrats phosphorylés, eux-mêmes liés à différents processus cellulaires. Cependant il existe des parvulines qui n'ont pas de spécificité pour les substrats phosphorylés (Uchida, Fujimori et al. 1999; Sekerina, Rahfeld et al. 2000). Il faut noter à ce propos, que si Pin1 et ses homologues sont extrêmement spécifiques des motifs phosphoSer/Thr-Pro et ont une fonction bien définie, chez la bactérie et la souris, les parvulines n'ont pas d'activité sur ces mêmes motifs et leur fonction est inconnue. Les parvulines sont également impliquées dans la régulation de l'expression de gènes (Shaw 2002). De manière générale, aucune implication directe de l'isomérisation catalytique des liaisons précédant une proline n'est mise en évidence dans les différentes fonctions biologiques des PPIases et montrée à l'échelle moléculaire. C'est le cas notamment de Pin1, qui a été impliquée dans un grand nombre de processus de régulation de protéines mitotiques et de voies de signalisation, mais dont le rôle de PPIase n'est pas clairement mis en cause dans la fonction biologique.

## **II – Pin1, une peptidyl-prolyl *cis/trans* isomérase essentielle dans la régulation du cycle cellulaire.**

### **2.1 – Une fonction biologique bien définie.**

Pin1 a été isolée en 1996 lors d'un criblage en double hybride chez la levure (Lu, Hanes et al. 1996) dont le but était d'identifier des partenaires de NIMA (*Never in mitosis A*), une kinase essentielle dans la régulation du cycle cellulaire, stimulant l'entrée en mitose et présentant certaines analogies de fonction et de régulation avec Cdc2. Des voies de signalisation similaires à celle de NIMA ont par ailleurs été identifiées dans des cellules eucaryotes (Lu and Hunter 1995; Lu and Hunter 1995). D'un point de vue fonctionnel, Pin1 et NIMA ont des rôles assez similaires dans la cellule et sont co-localisées dans le noyau. La surexpression de Pin1 dans les cellules HeLa induit l'arrêt en mitose (les cellules sont bloquées à la transition G<sub>2</sub>/M). Inversement, la déplétion de Pin1 provoque l'apparition de phénotypes mitotiques précoces, accompagnée d'une activation précoce de Cdc25 et du complexe CDK/cycline mais empêche également les cellules de sortir de la mitose et les conduit ainsi tout droit vers l'apoptose (Hani, Stumpf et al. 1995; Lu, Hanes et al. 1996; Crenshaw, Yang et al. 1998; Shen, Stukenberg et al. 1998).

Pin1 est constituée de deux domaines : un domaine de liaison de type WW responsable de la liaison des substrats au niveau de motifs phosphoSer/Thr-Pro et un domaine catalytique assurant la fonction isomérase au niveau de ce même type de motifs (Ranganathan, Lu et al. 1997; Verdecia, Bowman et al. 2000). Les deux domaines sont essentiels à l'activité de Pin1 *in vivo*. La survie de la levure peut être assurée par la complémentation en Pin1 chez des levures mutantes porteuses de délétions du domaine WW ou du domaine catalytique. Dans des conditions de surexpression, le domaine catalytique seul peut assurer la fonction biologique mais ce n'est pas le cas du domaine WW (Lu, Zhou et al. 1999; Winkler, Swenson et al. 2000). Cependant, il faut souligner que les parvulines de plantes sont fonctionnelles sans ce domaine WW (Landrieu, De Veylder et al. 2000).

***Une grande diversité fonctionnelle des protéines-substrats.***

Les premiers substrats décrits pour Pin1 sont des phospho-protéines mitotiques, c'est-à-dire des protéines phosphorylées à la mitose, comme la phosphatase Cdc25, les kinases Myt1, Wee1, Plk1, NIMA. Ensuite il a été montré que Pin1 interagissait également avec des facteurs de transcription comme c-Jun,  $\beta$ -caténine, Pol II, p53, et des protéines de structure comme la protéine microtubulaire tau.

Substrat	Activité	Sites d'interaction de Pin1	Références
<b>Protéines mitotiques / cycle cellulaire</b>			
NIMA	Kinase	?	(Lu, Hanes et al. 1996)
Cdc25	Phosphatase	phosphoThr48/67	(Yaffe, Schutkowski et al. 1997; Lu, Zhou et al. 1999)
Plk1	Kinase	?	(Yaffe, Schutkowski et al. 1997)
Wee1	Kinase	?	(Shen, Stukenberg et al. 1998)
Myt1	Kinase	?	(Shen, Stukenberg et al. 1998)
Cycline D1	Cycline G <sub>1</sub> /S	phosphoThr286	(Liou, Ryo et al. 2002)
<b>Facteurs de transcription</b>			
Pol II	ARN Polymérase II	phosphoSer	(Albert, Lavoie et al. 1999; Verdecia, Bowman et al. 2000; Kops, Zhou et al. 2002)
p53	Suppresseur de tumeur	phosphoSer33/315, phosphoThr81	(Zacchi, Gostissa et al. 2002; Zheng, You et al. 2002)
c-Jun	Activateur transcriptionnel	phosphoSer63/73	(Wulf, Ryo et al. 2001)
$\beta$ -caténine	Activateur transcriptionnel	phosphoSer246	(Ryo, Nakamura et al. 2001)
<b>Protéine du cytosquelette</b>			
tau	MAP	phosphoThr231	(Lu, Wulf et al. 1999)

**Tableau 1.1 :** Quelques protéines-substrats de Pin1 classées selon leur rôle cellulaire dont les sites d'interaction avec Pin1 identifiés sont indiqués dans la troisième colonne.

### 2.1.1 – Le rôle de Pin1 dans le cycle cellulaire.

L'inhibition de Pin1 et de son activité PPIase bloque les cellules en mitose et conduit à l'apoptose chez la levure et les cellules HeLa, et induit une entrée prématurée en mitose chez *Xenopus laevis* (Crenshaw, Yang et al. 1998; Shen, Stukenberg et al. 1998; Winkler, Swenson et al. 2000). La surexpression de Pin1 dans des cellules HeLa ou dans des extraits de *Xenopus laevis* conduit à un arrêt en G<sub>2</sub> (Hani, Stumpf et al. 1995; Lu, Hanes et al. 1996). La déplétion de Pin1 chez la souris n'est pas létale mais conduit à des symptômes liés à une prolifération cellulaire (Liou, Ryo et al. 2002). Pin1 joue un rôle dans la progression de la mitose en raison de sa spécificité pour les substrats phosphorylés, pour la plupart, par Cdc2 au niveau d'épitopes Ser/Thr-Pro portés par diverses protéines régulatrices du cycle cellulaire comme Cdc25 (Shen, Stukenberg et al. 1998; Lu, Zhou et al. 1999), les kinases Myt1, Wee1, Plk1 et Cdc27 (Crenshaw, Yang et al. 1998; Shen, Stukenberg et al. 1998), qui sont également des antigènes de l'anticorps MPM-2.

Pin1 est impliquée dans un nouveau mode de régulation de la déphosphorylation. Il a été proposé que la catalyse de l'isomérisation *cis/trans* des liaisons phosphoSer/Thr-Pro pouvait être couplée à la *trans*-spécificité de PP2A dans le contrôle du processus de déphosphorylation : Pin1 stimule la déphosphorylation de Cdc25 par PP2A de manière catalytique (Zhou, Kops et al. 2000; Stukenberg and Kirschner 2001). La forme non phosphorylée de Cdc25 étant la forme inactive de la phosphatase, Pin1 régule par conséquent négativement l'activité de celle-ci. La protéine Cdc25 non phosphorylée n'est plus capable d'activer Cdc2, une kinase majeure impliquée dans la cascade de phosphorylations déclenchant l'entrée en mitose. Pin1 pourrait ainsi provoquer l'arrêt du cycle cellulaire en phase G<sub>2</sub>/M par régulation négative de la boucle d'activation. Quel est le rôle moléculaire de Pin1 dans le mécanisme d'activation de la déphosphorylation de Cdc25 par PP2A ? Est-il basé sur l'activité isomérase de Pin1 ? Deux réponses sont envisageables et sont quelques peu contradictoires. La première réponse possible envisage une activité enzymatique PPIase de Pin1. Il a été montré que Pin1 agit de manière catalytique dans ce processus (Stukenberg and Kirschner 2001). Les résultats concernant l'activité phosphatase de Cdc25 sont conflictuels et peu convaincants en raison des faibles variations d'activité mises en évidence en présence de Pin1. Concernant les possibles modifications de conformation des protéines, elles sont mises en évidence à l'aide d'outils dépendant probablement des conformations *cis/trans* locales (dont on ne doute pas de la perturbation par Pin1), par exemple des protéases *trans*-spécifiques ou de l'anticorps MPM-2. Une deuxième réponse peut être avancée suite aux

données mettant en évidence une compétition moléculaire entre Pin1 et CKS. Les CKS sont des petites sous-unités qui se lient au complexe CDK/cycline et sont essentielles au bon déroulement du cycle cellulaire. Elles joueraient un rôle d'ancrage pour les protéines substrats des kinases CDK/cycline lorsque le complexe est actif à la mitose. Le complexe SUC1/CKS stimule la phosphorylation de Cdc25 par Cdc2, activant ainsi la phosphatase Cdc25. La compétition entre le domaine WW de Pin1 et CKS pour un site phosphorylé de Cdc25 expliquerait le rôle antagoniste trouvé pour Pin1 dans l'entrée en mitose (Patra, Wang et al. 1999).

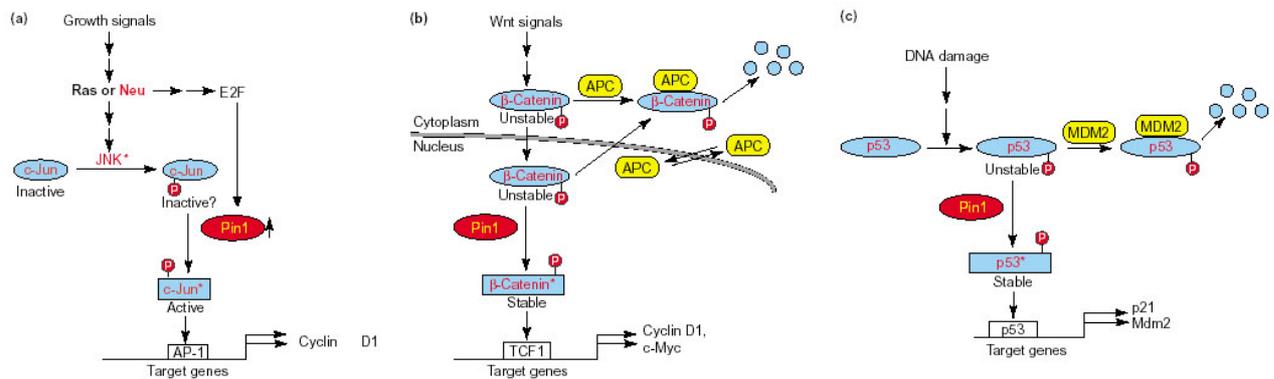
### **2.1.2 – Le rôle de Pin1 dans la transcription.**

Pin1 est capable de moduler la transcription par sa liaison à l'ARN Polymérase II via le domaine C-terminal répété (Verdecia, Bowman et al. 2000; Xu, Hirose et al. 2003) qui est multi-phosphorylé au niveau de nombreux sites Ser/Thr-Pro. En se liant au CTD, Pin1 pourrait inhiber sa déphosphorylation par FCP1, une phosphatase spécifique et contribuer à désactiver la transcription dans les cellules en mitose. De plus, Pin1 permet également d'inhiber la transcription des précurseurs d'ARNm et l'épissage des pré-ARNm stimulé par l'ARN Pol II (Xu, Hirose et al. 2003).

Pin1 n'a pas d'effet spectaculaire sur la transcription générale dans la cellule, que ce soit chez l'Homme ou la levure (Fujimori, Gunji et al. 2001). En revanche, elle joue un rôle important dans la transcription de quelques gènes bien spécifiques (Ryo, Liou et al. 2002). En effet, Pin1 régule l'activité de deux facteurs de transcription, c-Jun et  $\beta$ -caténine. C-Jun est phosphorylé sur deux motifs Ser63/Ser73-Pro par les kinases JNK, à la suite d'une série de signaux provenant de la voie de signalisation *Ras*. Cette phosphorylation, qui augmente l'activité transcriptionnelle de c-Jun pour ces gènes cibles dont celui de la cycline D1, est renforcée par l'action de Pin1 (Wulf, Ryo et al. 2001) (Figure 1.6a). En revanche, aucune donnée ne permet d'affirmer qu'il s'agisse bien de l'activité prolyl *cis/trans* isomérase de Pin1 qui soit impliquée ou de la liaison de Pin1 à c-Jun. L'hypothèse mécanistique de l'interaction entre Pin1 et  $\beta$ -caténine empêche une interaction de cette dernière avec la protéine APC (*Adenomatous Polyposis Coli*) en se liant à un site proche de celui de l'APC. Si l'on considère que cette interaction est cruciale pour la dégradation du facteur de transcription, la simple formation d'un complexe avec Pin1 pourrait réguler sa localisation subcellulaire et sa dégradation (Figure 1.6b).

***Pin1 dans la réponse aux altérations de l'ADN.***

Pin1 est capable de se lier au suppresseur de tumeurs p53 après modification post-traductionnelle par les kinases MAPK et CDK, au niveau de trois sites de phosphorylation (S33, T81 et S315), en réponse à des altérations de l'ADN permettant sa liaison à l'ubiquitine-ligase Mdm2 et sa dégradation consécutive (Zacchi, Gostissa et al. 2002; Zheng, You et al. 2002). Pin1 entre en compétition avec Mdm2 pour la liaison à p53 car les sites de liaison de Pin1 sont localisés à proximité des sites de liaison de Mdm2. La dissociation du complexe Mdm2-p53 conduit à la stabilisation de p53 tandis que Pin1 stimule également la liaison de p53 à l'ADN et son activité transcriptionnelle vis-à-vis du promoteur p21 (Wulf, Liou et al. 2002), conduisant à l'arrêt du cycle cellulaire et à l'apoptose en réponse à des dommages causés à l'ADN (Figure 1.6c). Dans le processus d'activation de p53, l'activité catalytique semble essentielle, mais le domaine WW est également nécessaire, et il a été proposé que l'isomérisation de p53 pourrait induire une modification structurale qui activerait sa fonction dans la transcription et qu'ainsi le rôle protecteur de p53, comme suppresseur de tumeurs, serait étroitement lié à Pin1.



**Figure 1.6 :** Trois mécanismes moléculaires impliquant Pin1 (a,b) dans la régulation de l'activité transcriptionnelle des protéines c-Jun et β-caténine sur l'expression de cycline D1, et (c) dans la stabilisation du suppresseur de tumeurs p53 et son activité transcriptionnelle (d'après Lu, *Trends Biochem Sci* 2004).

***Pin1 dans les voies de signalisation cellulaires.***

Les voies de signalisation *Ras/Neu* activant les kinases JNK sont impliquées dans la phosphorylation de c-Jun et permettent d'augmenter son activité transcriptionnelle (Figure 1.6a). L'expression de Pin1 est elle-même régulée par le facteur de transcription E2F et est nécessaire pour la transformation maligne de cellules épithéliales mammaires par la voie *Neu/Ras*, par l'activation de la cycline D1 (Ryo, Liou et al. 2002). D'autre part, dans le cancer

du colon, la  $\beta$ -caténine est activée par des mutations au niveau de son propre gène ou celui de l'APC. Dans le cancer du sein, elle est surexprimée. Pin1, par son interaction avec la  $\beta$ -caténine, contribue à la stabilisation et l'accumulation de  $\beta$ -caténine dans le noyau. Ainsi la surexpression de Pin1 dans les cancers contribue à activer la  $\beta$ -caténine, dont le niveau d'expression est corrélé à celui de Pin1.

### **2.1.3 – Pin1, une cible thérapeutique émergente dans le traitement des cancers.**

Pin1 est surexprimée dans les cellules cancéreuses et la surexpression de Pin1 contribue à la transformation des cellules et potentialise plusieurs processus impliqués dans l'oncogénèse, comme décrits dans les paragraphes précédents (Wulf, Ryo et al. 2001; Ryo, Liou et al. 2002; Ryo, Liou et al. 2003; Wulf, Ryo et al. 2003). La cycline D1, un des marqueurs du cycle cellulaire, en particulier, est impliquée directement dans les mécanismes de cancérisation car une surexpression de cycline D1 est observée dans de multiples tumeurs (elle est surexprimée dans 50% des cancers du sein) et il existe une nette corrélation entre la surexpression de Pin1 dans les cellules cancéreuses et le niveau de cycline D1. Pin1 semble jouer un rôle évident dans la régulation de la cycline D1, à deux niveaux :

- l'activation de la transcription du gène de la cycline D1 par les interactions que nous avons décrites précédemment avec c-Jun et  $\beta$ -caténine .
- Pin1 lie et stabilise la cycline D1 directement par une interaction directe au niveau du motif Thr286-Pro, ce qui empêcherait son interaction avec CRM1 qui la transporte vers le cytoplasme pour destruction (Alt, Cleveland et al. 2000). Cette interaction a des conséquences fonctionnelles puisque les souris Pin1-*knockout* possèdent un phénotype identique à celui des souris cycline D1-*knockout* (Liou, Ryo et al. 2002).

Ces preuves indiquent toutes un rôle crucial pour Pin1 dans la régulation de la cycline D1, qui est retrouvée dans de multiples tumeurs et contribue à la transformation de la cellule. L'inhibition de Pin1 paraît donc une bonne stratégie pour lutter contre le cancer.

#### ***Régulation de l'activité de Pin1.***

D'une façon tout à fait intéressante, l'interaction du domaine WW ne semble pas seulement être régulée par la phosphorylation de son substrat, mais également par la phosphorylation du domaine WW même (Lu, Zhou et al. 2002). La phosphorylation du domaine WW au niveau du résidu Ser16 inactive la protéine en inhibant sa capacité de liaison des substrats et joue

ainsi un rôle dans sa localisation subcellulaire. Alors que Pin1 est essentiellement localisée dans le noyau, Pin1 phosphorylée a une distribution cellulaire plus diffuse dans l'ensemble des compartiments, ressemblant à celle du domaine catalytique isolé (Lu, Zhou et al. 2002). Le domaine WW jouerait, par conséquent, un rôle dans la localisation cellulaire de Pin1 au niveau nucléaire au travers de la liaison des substrats. D'autre part, c'est la forme non phosphorylée active qui est retrouvée principalement dans les cellules mammaires cancéreuses alors que la forme phosphorylée inactive est présente majoritairement dans les cellules saines (Wulf, Ryo et al. 2001; Wulf, Ryo et al. 2003). Toutes ces données font de Pin1 une cible thérapeutique de choix dans le traitement des cancers. L'implication de Pin1 dans les mécanismes de cancérisation est liée à des interactions fonctionnelles avec les différentes protéines citées précédemment, mais le rôle moléculaire exact de Pin1, notamment en tant que peptidyl-prolyl *cis/trans* isomérase, n'est pas encore bien défini.

### ***Pin1 comme marqueur diagnostique et cible thérapeutique dans le cancer (Lu 2003).***

Pin1 est surexprimée dans la plupart des tissus cancéreux tels que le sein, la prostate, le colon, le poumon, .... La voie de signalisation Rb-E2F, impliquée dans la régulation de Pin1, dérégulée dans la plupart des cancers, pourrait jouer un rôle significatif dans la surexpression de Pin1, elle-même fortement corrélée avec la surexpression de protéines oncogéniques telles que la cycline D1 et la  $\beta$ -caténine. Pin1 peut également être utilisée comme marqueur moléculaire couvrant une large gamme de tumeurs. Le niveau de surexpression de Pin1 étant corrélé avec l'agressivité des tumeurs et l'issue clinique des cancers de la prostate, Pin1 peut être utilisée comme agent de diagnostic et de pronostic. Un haut niveau d'expression de Pin1 augmente le risque de rechute chez les patients atteints d'un cancer de la prostate (Ayala, Wang et al. 2003). La détection et la quantification du niveau d'expression de Pin1 permettrait de détecter précocément des cancers, d'identifier le stade atteint dans la maladie, d'établir une stratégie thérapeutique efficace et de suivre la réponse par rapport à la thérapie engagée.

### ***Pin1 et Pintex Pharmaceuticals, Inc. (Watertown, MA, USA).***

La société Pintex s'est engagée depuis 1999 dans le développement de molécules inhibant la fonction de Pin1 selon une stratégie de conception rationnelle basée sur la structure de l'enzyme et le développement en parallèle de diagnostics permettant d'établir le degré de sévérité de certains cancers et les risques de rechute après rémission. Toute la stratégie est fondée exclusivement sur Pin1 comme seule et unique cible thérapeutique couvrant un large

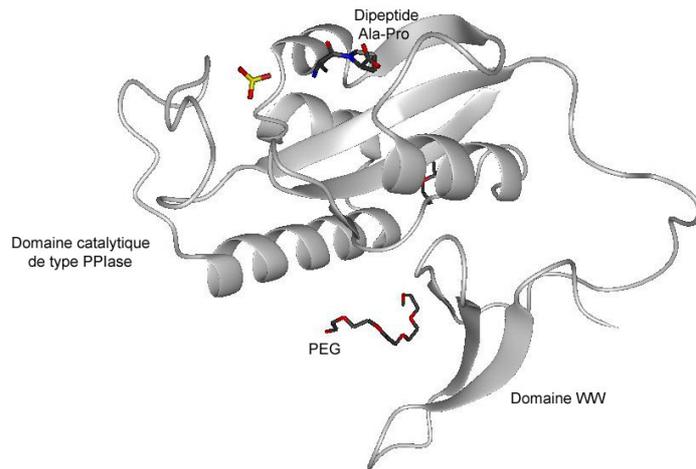
spectre de tumeurs, et a fait l'objet de plusieurs brevets, dont celui concernant la stratégie chimio-thérapeutique et le criblage des composés (PN/20030096387 ; « Peptidyl-prolyl cis-trans isomerase inhibitors and uses therefore », Noel, J. P. and Hunter, T. R.) couvrant la structure des inhibiteurs potentiels, les stratégies de synthèse, les modèles de criblage *in vitro*, sur base de l'inhibition de l'activité *cis/trans* isomérase sur des substrats peptidiques, et *in vivo*, sur la régulation de la mitose chez la levure, les applications....

### **2.2 – Un rôle moléculaire obscur.**

La structure de Pin1 est bien décrite et permet notamment d'expliquer sa spécificité pour les protéines phosphorylées à la fois au niveau du domaine de liaison N-terminal de type WW, impliqué dans la liaison du substrat, et au niveau du domaine catalytique C-terminal qui assure la fonction *cis/trans* isomérase. En effet, les deux domaines de Pin1 reconnaissent spécifiquement les motifs Ser/Thr-Pro phosphorylés au sein de phospho-protéines mitotiques ou de facteurs de transcription phosphorylés également reconnus par l'anticorps monoclonal mitotique MPM-2 (Winkler, Swenson et al. 2000). Pourtant, malgré un rôle cellulaire essentiel, la fonction moléculaire de Pin1 reste obscure car la plupart des protéines substrats de Pin1 présentent des parties non structurées dans lesquelles se trouvent les épitopes phosphoSer/Thr-Pro reconnus par Pin1. D'autre part, sa fonction isomérase n'a jamais été impliquée directement dans les différents mécanismes de régulation qu'on lui attribue.

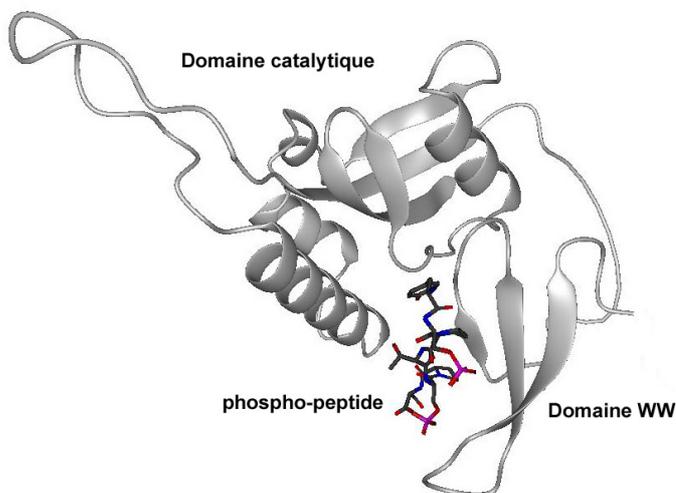
#### **2.2.1 – Structure de Pin1 : deux domaines indépendants ?**

Une première structure cristallographique de Pin1 en complexe avec le dipeptide Ala-Pro (Ranganathan, Lu et al. 1997) montre des interactions avec le substrat au niveau du domaine catalytique dans un site hydrophobe et un repliement de la boucle flexible contenant des résidus basiques autour d'un ion sulfate, tandis que le domaine WW interagit avec une molécule de PEG (Figure 1.7).



**Figure 1.7 :** Structure globale de Pin1 obtenue par diffraction des rayons X. La protéine a été co-cristallisée avec le dipeptide Ala-Pro et un ion sulfate que l'on retrouve logés dans le domaine catalytique. Une molécule de PEG se trouve à l'interface entre le domaine WW et le domaine catalytique.

Une seconde structure, où Pin1 forme un complexe avec un phospho-peptide issu du domaine C-terminal (CTD) de l'ARN Polymérase II, ne montre des interactions qu'avec le domaine WW uniquement (Verdecia, Bowman et al. 2000) (Figure 1.8). Dans les deux structures, même si les substrats sont localisés dans des sites différents de la protéine, les deux domaines sont rapprochés malgré la flexibilité que leur confère le *linker* qui les relie mais une coopération fonctionnelle, comme suggérée par la structure, n'a jamais été mise en évidence.



**Figure 1.8 :** Structure de Pin1 en complexe avec un peptide di-phosphorylé du domaine C-terminal de l'ARN Polymérase II (YpSPTpSPS) obtenue par diffraction des rayons X. La structure montre une interaction localisée au niveau du domaine WW de Pin1 avec le deuxième motif phosphoSer-Pro.

### 2.2.1.1 – Le domaine WW et la liaison du substrat, une fonction partagée par d'autres modules protéiques.

*Les modules de liaison à l'origine des interactions protéine-protéine, entre universalité et spécificité.*

La communication inter-cellulaire assure une fonction cruciale dans les organismes multicellulaires. Elle permet ainsi la croissance des cellules, leur différenciation et leur

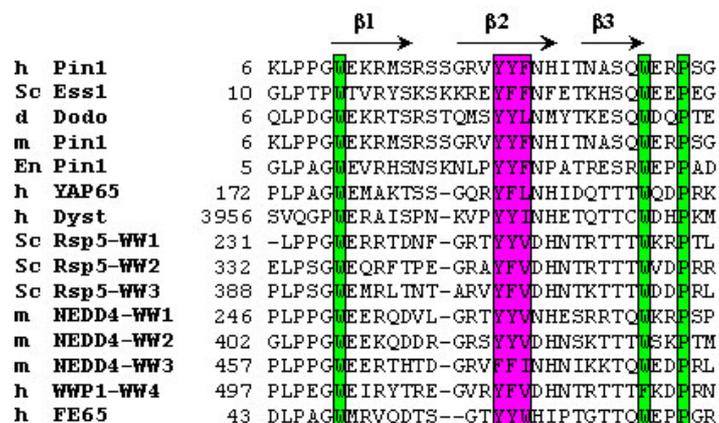
organisation en tissus spécifiques pouvant assurer une fonction précise dans un organisme. Ce processus dit de « signalisation » est responsable de phénomènes biologiques et pathologiques (Pawson, Raina et al. 2002). Parmi les conséquences de la rupture de ces processus de signalisation cellulaire se cachent des pathologies comme le cancer, le diabète, les désordres immunologiques... Les modules de liaison protéiques, qui se trouvent au cœur des interactions protéine-protéine, jouent un rôle essentiel dans la régulation de ces processus de signalisation et pourraient être impliqués dans un certains nombres de pathologies (Sudol 1996; Sudol, Sliwa et al. 2001). Les domaines de liaison ont pour caractéristique de lier leur substrat avec une affinité modérée mais avec une grande spécificité. Ils jouent un rôle important dans une large gamme de processus biologiques comme la régulation de l'activité enzymatique, la sous-localisation cellulaire des protéines, l'assemblage de complexes multi-protéiques.... Les domaines SH3 (*Src Homology 3*) par exemple, qui ont été les premiers modules de liaison identifiés, sont répertoriés dans plus de 1000 protéines incluant des protéines de levures jusqu'aux protéines humaines. Les protéines sont souvent constituées d'un assemblage de modules protéiques qui peuvent avoir une fonction catalytique (kinase, phosphatase, ubiquitine-ligase, hydrogénase...) ou une fonction de liaison qui permet les interactions avec d'autres protéines, des phospho-lipides, des acides nucléiques, des seconds messagers...

Il existe plusieurs types de modules médiateurs des interactions protéine-protéine hormis les domaines WW : les domaines SH2 qui lient les résidus Tyr phosphorylés, les domaines SH3 qui lient spécifiquement les régions riches en prolines, les domaines C2 qui régulent la fixation membranaire via le calcium de la protéine-hôte, et les domaines PH (*Pleckstrin Homology*) qui lient le phosphatidylinositol-4,5-bisphosphate et jouent un rôle dans les interactions protéine-membrane,... D'un point de vue fonctionnel, les domaines WW se rapprochent des domaines SH3 qui lient le même type de ligand que les domaines WW, c'est-à-dire des ligands riches en prolines ou contenant au moins une proline (Sudol 1996; Mayer 2001; Macias, Wiesner et al. 2002).

### ***Généralités sur les domaines WW.***

Les domaines WW ont été identifiés dans les protéines de divers organismes eucaryotes comme étant des domaines permettant les interactions physiques entre les protéines. Ce sont de petits modules de 38-40 acides aminés caractérisés par un pourcentage élevé de résidus hydrophobes et aromatiques, notamment deux tryptophanes conservés séparés de 20 ou 22

résidus, un ensemble de deux ou trois résidus aromatiques (en rose, Figure 1.9) localisés entre les deux tryptophanes et une proline conservée (en vert, Figure 1.9) (Verdecia, Bowman et al. 2000).



**Figure 1.9 :** Alignement de séquences de 15 domaines WW présents dans les protéines de différents organismes eucaryotes tels que l’homme (h), la souris (m), la drosophile (d) et la levure *Saccharomyces cerevisiae* (Sc), (d’après Verdecia, *Nat Struc Biol* 2000). Le numéro à gauche de la séquence primaire indique la position occupée par le premier résidu du domaine dans la séquence totale de la protéine dont il est extrait. Les régions organisées en une structure en triple feuillet  $\beta$  ( $\beta 1$ ,  $\beta 2$  et  $\beta 3$ ) sont indiquées par des flèches.

De manière générale, la partie N-terminale et la partie C-terminale sont globalement conservées, sans insertion ou délétion, et riches en résidus Pro. En revanche, la région centrale est moins conservée et présente un certain degré de variation en longueur de séquence. De plus, on note une répartition homogène de résidus polaires sur l’ensemble de la protéine (Figure 1.9). Les domaines WW peuvent être considérés comme les plus petites protéines monomériques qui se replient en l’absence de co-facteurs ou de ponts disulfures. D’un point de vue structural, ils se replient sous forme d’un triple feuillet  $\beta$  anti-parallèle. Il existe cinq classes de domaines WW selon la nature du ligand reconnu (Tableau 1.2). Dans tous les cas, au moins un résidu Pro est impliqué dans la séquence consensus reconnue. Certains groupes de domaines WW (groupes II, III et V) lient des séquences contenant plusieurs prolines, dites « séquences riches en prolines ». L’affinité des domaines WW pour leur ligand est de l’ordre de 1 à 100  $\mu$ M (Sudol 1996).

Groupe	Protéines représentatives	Séquences consensus du ligand
Groupe I	YAP65, Nedd4	PP <sub>x</sub> Y
Groupe II	FE65	PPPP <u>PP</u> LP
Groupe III	<i>Formin Binding Protein</i>	(P <sub>x</sub> xGM <sub>x</sub> PP)N
Groupe IV	Pin1, Ess1	PhosphoSer/Thr-Pro
Groupe V	PQBP-1	R <sub>x</sub> ( <sub>x</sub> )PPGPPP <sub>x</sub> R

**Tableau 1.2 :** Séquences consensus reconnues par les différentes classes de domaines WW où x représente un résidu quelconque.

Les domaines WW semblent être bien conservés au cours de l'évolution et se retrouvent dans divers organismes plus ou moins évolués comme les plantes, la levure, la mouche, les vers et les vertébrés. Ce sont des modules intracellulaires qui appartiennent à des protéines cytoplasmiques ou nucléaires. Différentes protéines porteuses d'un domaine WW ont été identifiées dans la co-activation de la transcription et dans la modulation de l'activité de la RNA Polymérase II. Cette liaison des domaines WW, notamment au niveau du domaine C-terminal (CTD), qui comporte plusieurs séquences répétées, est régulée par des mécanismes de phosphorylation/déphosphorylation. D'autre part, les domaines WW se retrouvent impliqués dans plusieurs pathologies notamment la maladie d'Alzheimer et le cancer (Sudol 1996).

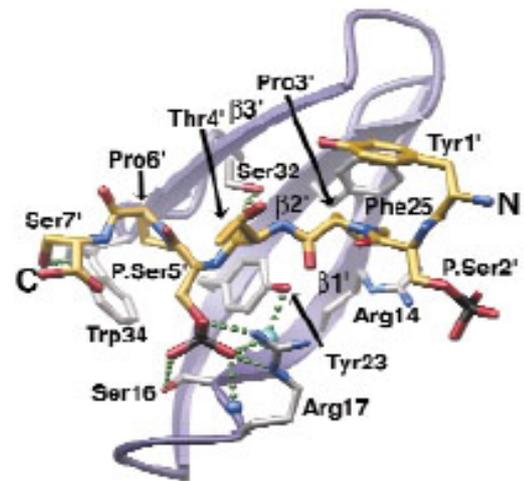
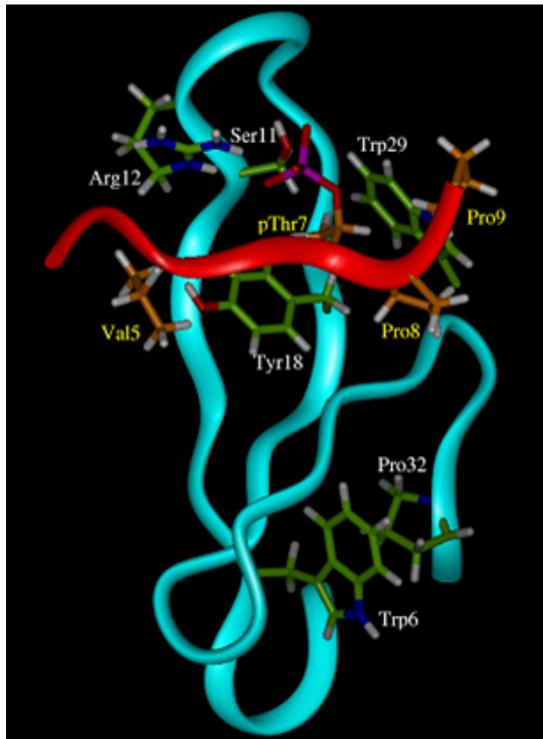
L'identification et la caractérisation des domaines WW proviennent de deux études indépendantes effectuées en 1994. Une première étude liée à l'identification d'une protéine possédant une séquence riche en prolines permettant la liaison au domaine SH3 de la protéine Yes, une tyrosine kinase de la famille Src. L'analyse séquentielle de cette protéine appelée YAP (*Yes-associated protein*), chez la souris, a révélé la présence d'une séquence répétée de manière imparfaite correspondant à un module protéique puisque des séquences hautement similaires ont été retrouvées dans diverses protéines non-homologues impliquées dans diverses fonctions cellulaires allant de l'architecture de la cellule aux mécanismes de régulation et de signalisation dans différentes espèces, de la levure aux eucaryotes supérieurs. Ce module a été nommé WW en raison des deux tryptophanes conservés qui le caractérisent (Sudol 1995; Sudol 1995). Un criblage fonctionnel a permis l'identification de deux ligands du domaine WW de la protéine YAP, les protéines WBP-1 et WBP-2, comportant chacune une région consensus riche en prolines. Des tests de liaison et de mutagenèse ont montré que le domaine WW de YAP liait spécifiquement la région riche en prolines avec une haute

affinité et que cette région se différenciait du motif consensus PXXP de liaison du domaine SH3 suggérant un rôle de médiateur des interactions protéine-protéine au niveau de régions riches en prolines de manière similaire mais distincte des domaines SH3 (Chen and Sudol 1995). Dans le même temps, une autre équipe identifiait un certain nombre de protéines à domaines Rsp5, du nom de la protéine dans laquelle ils ont été identifiés (Hofmann 1995), décrits comme de petits modules autonomes qui se replient sous forme d'unité indépendante sans l'aide de ponts disulfures et sans contact avec les résidus des autres parties de la protéine-hôte.

Les domaines médiateurs des interactions protéine-protéine sont présents dans une grande variété de protéines, sans fonction commune apparente car liées à des processus cellulaires différents, et peuvent être considérés comme le moyen le plus largement utilisés pour introduire une fonction supplémentaire dans une protéine existante. La fonction qu'ils remplissent n'est pas toujours clairement définie mais ils permettent d'établir des connexions entre les protéines d'un réseau de régulation complexe et de déterminer la fonction de la protéine hôte, d'amener un substrat à proximité de son site catalytique, de localiser les protéines dans leur compartiment cellulaire par le biais d'interactions du domaine avec ses ligands, ou encore les domaines peuvent être eux-mêmes impliqués directement dans des processus de régulation cellulaire.

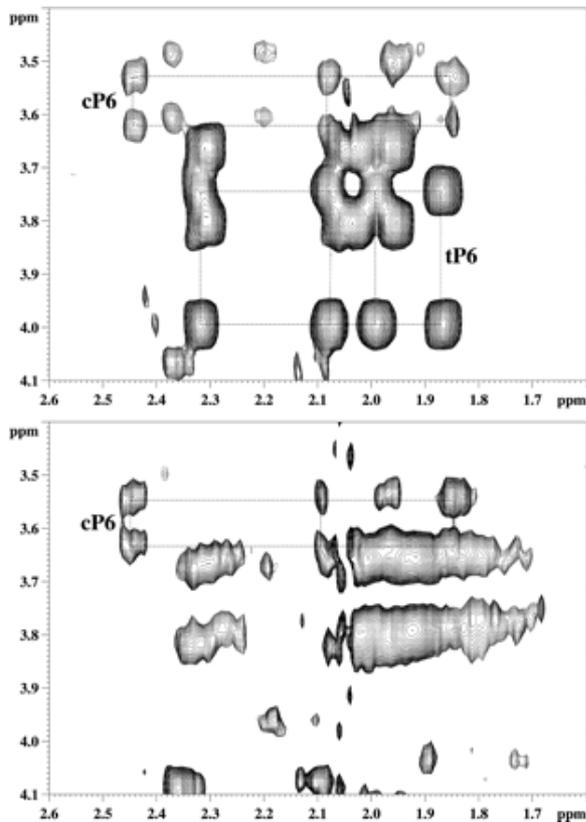
### ***Le domaine WW de Pin1 : son originalité, la phospho-spécificité.***

Le domaine WW de Pin1 est responsable de la liaison aux substrats contenant des sérines ou des thréonines phosphorylées suivies d'une proline, et en particulier à des phosphoprotéines impliquées dans la mitose (Lu, Zhou et al. 1999). Il lie les substrats phosphorylés par l'intermédiaire des résidus polaires Ser16-Arg17, localisés dans une boucle  $\beta$ -hairpin, qui sont en contact direct avec le résidu phosphorylé et d'une paire de résidus aromatiques Tyr23-Trp34 qui jouent un rôle de pince moléculaire qui accueille le résidu Pro (Verdecia, Bowman et al. 2000; Wintjens, Wieruszeski et al. 2001) (Figure 1.10).



**Figure 1.10** : Structure du domaine WW de Pin1 en complexe avec un phosphopeptide par RMN (à gauche, d'après Wintjens, *J Biol Chem* 2001) et par diffraction des rayons X (à droite, d'après Verdecia, *Nat Struc Biol* 2000)

L'étude par RMN de la liaison du domaine WW à deux peptides phosphorylés, l'un issu de la phosphatase Cdc25 de *Xenopus laevis* (EQPLpTPVTDL) et l'autre de la protéine tau humaine (KVSVVVRpTPPKSPS), a montré des interactions identiques à celles qui avaient été décrites dans la structure cristallographique de Pin1 entière avec un phospho-peptide. Des constantes de dissociation à l'équilibre de 117 et 230  $\mu\text{M}$  ont été définies pour les peptides Cdc25 et tau respectivement. Le domaine WW lie le substrat avec une orientation unique (Wintjens, Wieruszski et al. 2001) de l'extrémité N-terminale vers C-terminale (dite *N to C*), contrairement aux domaines SH3 où les deux directions du substrat par rapport au domaine sont possibles (*N to C* et *C to N*) (Zarrinpar and Lim 2000). D'autre part, le domaine WW de Pin1 possède une spécificité pour les liaisons phosphoSer/Thr-Pro en *trans* contrairement aux domaines 14-3-3 qui lient les motifs en *cis* ou en *trans* (Rittinger et al. 1999).



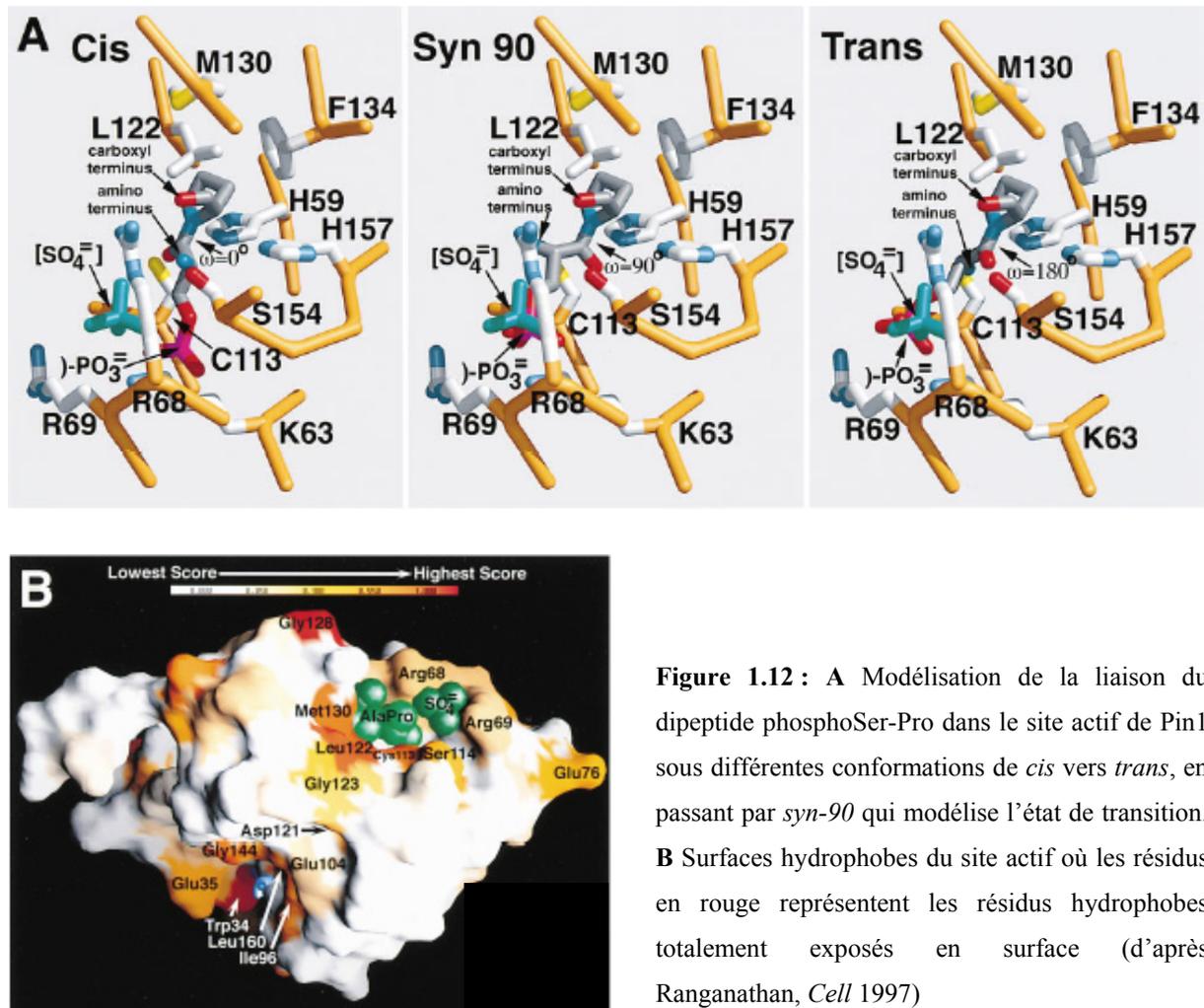
**Figure 1.11 :** *Trans*-spécificité de liaison du domaine WW au niveau de motifs phosphoThr-Pro. Le spectre TOCSY du phosphopeptide est représenté en détail au niveau des contacts  $^1\text{H}$ - $^1\text{H}$  avec les H $\delta$  des prolines. En haut, le peptide seul avec deux ensembles de signaux correspondant à la proline en *trans* (tP6) et à la proline en *cis* (cP6), en bas, le peptide en présence d'une quantité équimolaire de WW. On note une disparition des signaux correspondant à la forme *trans* par élargissement de ligne causé par l'interaction avec le domaine WW (d'après Wintjens, *J Biol Chem* 2001).

Seuls les domaines 14-3-3 avaient été décrits avant les domaines WW IV comme petits domaines indépendants permettant la liaison de résidus Ser ou Thr phosphorylés (Yaffe, Rittinger et al. 1997; Yaffe and Elia 2001), alors que les modules liant les phosphotyrosines, comme les domaines SH2, étaient bien caractérisés. Depuis, on a montré l'existence d'un autre module, appelé FHA, liant les motifs phosphoThr uniquement (Durocher, Taylor et al. 2000). L'importance de tels modules assurant les interactions avec des protéines phosphorylées vient du fait que la phosphorylation est étroitement impliquée dans les mécanismes de signalisation cellulaire (Yaffe and Elia 2001). Les motifs phosphoSer/Thr reconnus par les modules 14-3-3, WW et FHA recouvrent partiellement les motifs phosphorylés par différentes familles de kinases, permettant ainsi la formation de complexes de signalisation spécifiques qui peuvent être contrôlés par des combinaisons variées entre les kinases phosphorylant les résidus Ser/Thr, en amont, et les modules de liaison phospho-spécifiques, en aval (Yaffe and Elia 2001). Il a été montré que Pin1 est un régulateur mitotique essentiel qui interagit avec un grand nombre d'antigènes MPM-2, phosphorylés par des kinases dirigées contre des motifs Ser/Thr-Pro et que le domaine WW jouait un rôle essentiel dans l'activité de Pin1 en permettant probablement l'ancrage de Pin1 à ses substrats (Lu, Zhou et al. 1999).

### 2.2.1.2 – Le domaine catalytique : une structure et une spécificité de substrat différentes des autres PPIases.

*Le domaine catalytique de Pin1 présente une région basique impliquée dans la reconnaissance des protéines phosphorylées.*

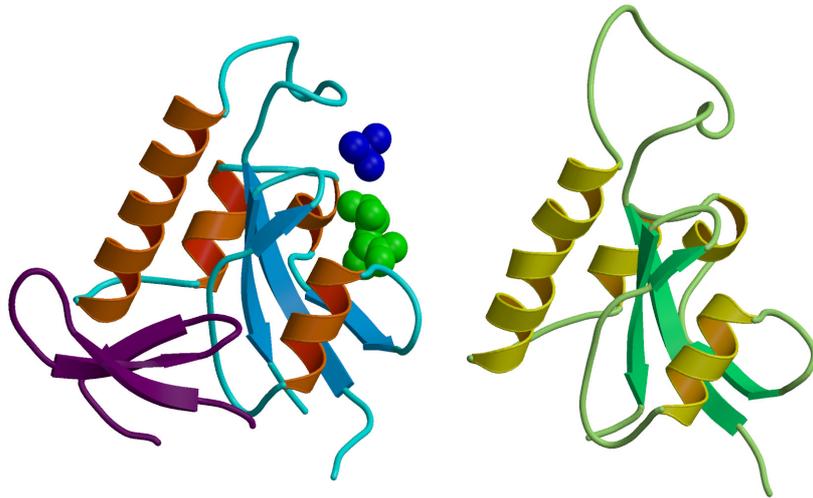
La première structure cristalline de Pin1, résolue à 1.35Å, montre la formation d'un complexe avec le di-peptide Ala-Pro en conformation *cis* (Ranganathan, Lu et al. 1997). La mutation de la cystéine du site actif a permis d'envisager un mécanisme catalytique incluant une catalyse acide-base et une liaison covalente pour l'isomérisation de la liaison peptidique. La mise en évidence d'un *cluster* basique constitué de trois résidus chargés positivement K63, R68 et R69, localisés dans la grande boucle flexible  $\beta_4/\alpha_1$ , a permis de mettre en évidence une interaction avec un résidu acide en position N-terminale de la proline, elle-même liée par un groupe de résidus hydrophobes (L122, M130, F134) (Figure 1.12B). C'est cette préférence pour les résidus acides, illustrée par la liaison d'un ion sulfate dans le *cluster* basique, qui a permis d'envisager pour Pin1, un rôle dans l'interaction avec des protéines phosphorylées impliquées dans la régulation du cycle cellulaire. Cependant, aucune structure n'a jamais montré la liaison d'un substrat phosphorylé, même de type peptidique, dans le site actif de Pin1. La seule preuve d'une interaction du domaine catalytique de Pin1 avec un substrat phosphorylé vient de l'étude de Jacobs et collaborateurs, montrant une interaction entre le domaine catalytique isolé de Pin1 et un peptide phosphorylé du domaine C-terminal de l'ARN Polymérase II et le peptide modèle Pintide phosphorylé, mais aucune interaction avec un peptide de Cdc25 (Jacobs, Saxena et al. 2003). La structure de Ranganathan montre également que, contrairement au domaine WW, le site actif peut adapter une liaison en *cis* et l'interaction avec le résidu en amont de la proline est optimale pour le dipeptide en conformation *syn-90*, mimant l'état de transition entre les formes *cis* et *trans*. Les résidus C113, H157, H59 et S154 entourent la liaison peptidique précédant la proline, et sont placés autour de l'axe de rotation de la liaison (Figure 1.12A). La fonction thiol du résidu C113, en particulier, interviendrait dans la formation d'une liaison covalente avec le substrat, par attaque nucléophile du groupement carbonyle de la liaison peptidique, qui présente un caractère partiel de double liaison, pour former un carbone tétravalent facilitant la rotation de la liaison peptidique. La mutation de ce résidu C113 induit une perte d'activité isomérase sur des substrats peptidiques. D'autre part, ce résidu est la cible précise de la juglone, inhibiteur sélectif des parvulines, qui se lie de manière covalente à la cystéine et perturbe la structure du site catalytique (Hennig, Christner et al. 1998).



### *Analogie structurale avec l'homologue de plante Pin1At*

La première parvuline de plante, Pin1At d'*Arabidopsis thaliana*, homologue de Pin1, est fonctionnelle sans domaine WW (Landrieu, De Veylder et al. 2000; Landrieu, Wieruszski et al. 2002) et présente une grande similarité de structure avec le domaine catalytique de Pin1, hormis l'insertion d'un motif IIL dans la boucle flexible et une lysine à la place d'une proline strictement conservée dans les autres eucaryotes (Figure 1.13). Depuis cette découverte, d'autres homologues dans trois autres espèces de plante ont confirmé cette absence de domaine WW (Metzner, Stoller et al. 2001). La structure tridimensionnelle de Pin1At permet de comprendre comment celle-ci peut être active sans ce domaine WW essentiel pour l'homologue humaine. La structure montre une région très électro négative, à la place de l'interface avec le domaine WW de Pin1, qui pourrait intervenir dans les contacts protéine-protéine par interaction électrostatique directe. D'autre part, d'un point de vue fonctionnel, la capacité du domaine catalytique Pin1At à compléter un mutant de levure *S. cerevisiae*

thermosensible pour la protéine Pin1, est supérieure au domaine catalytique Pin1 humain isolé, c'est à dire qu'un niveau d'expression inférieur est suffisant pour le même effet de complémentation (Metzner, Stoller et al. 2001; Yao, Kops et al. 2001). Ceci pourrait expliquer que, contrairement à l'homologue humaine, Pin1At n'a pas besoin de domaine WW pour être active.



**Figure 1.13** : Comparaison de la protéine Pin1 d'*Arabidopsis thaliana* (Pin1At) (à droite) constituée uniquement d'un domaine catalytique et de Pin1 (à gauche) (d'après Landrieu, *J Mol Biol* 2002). Le triple feuillet  $\beta$  du domaine WW de Pin1 est représenté en violet. Pin1 est ici représentée avec le dipeptide Ala-Pro et un ion sulfate (représentation « boules », en vert et bleu respectivement) dans le site catalytique. On note une grande similarité structurale du domaine catalytique de Pin1 avec Pin1At : un feuillet  $\beta$  central entouré de 4 hélices  $\alpha$ , et une boucle flexible  $\beta 1/\alpha 1$  (numérotation Pin1At) accueillant l'ion sulfate, qui est plus grande chez *Arabidopsis thaliana* que dans l'homologue humaine.

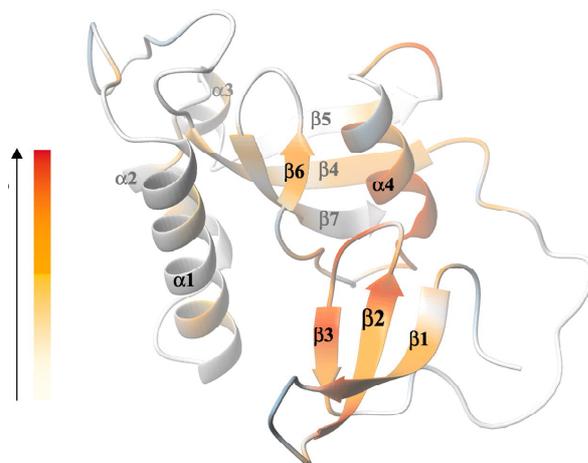
L'ensemble de ces données structurales ne permet pas de trancher de manière indubitable en faveur d'une fonction isomérase pour le domaine dit « catalytique » de Pin1, bien qu'une activité isomérase ait été mise en évidence sur des peptides, puisqu'au niveau moléculaire, une activité isomérase de Pin1 n'a jamais été montrée sur un substrat protéique.

### **2.2.1.3 – Indépendance structurale des deux domaines mais coopération fonctionnelle ?**

Le rôle joué par les deux domaines de Pin1 n'est pas clair. En effet, le domaine WW, en tant que module de liaison, est impliqué dans la liaison du substrat avec une affinité plus importante que le domaine catalytique (Verdecia, Bowman et al. 2000). Le mode d'interaction précis de Pin1 avec ses substrats n'est pas encore bien défini. Les structures se limitent aux substrats peptidiques, ce qui peut expliquer l'absence de coopération observée entre les deux

domaines au niveau moléculaire. En revanche, au niveau fonctionnel, l'importance des deux domaines semble indiquer un rôle coopératif, mais aucune information structurale ne permet de déterminer s'il s'agit d'une coopération autour d'un seul et même substrat ou de deux substrats distincts. La présence de plusieurs sites d'interaction pour Pin1 au sein d'un même substrat peut induire des interactions des deux domaines au niveau de ces différents sites. Mais on peut également envisager que le domaine WW et le domaine catalytique lient chacun un substrat, et que Pin1 jouerait un rôle de médiateur en favorisant les interactions entre deux protéines par un simple rapprochement conduisant à une augmentation de la concentration locale des deux partenaires l'un pour l'autre.

La recherche de données moléculaires permettant de définir le mode d'action de Pin1 est complexe, notamment parce que les interactions avec le domaine « catalytique » sont faibles et transitoires, et que les interactions avec le domaine WW l'emportent. En solution, les données indiquent une indépendance assez marquée des deux domaines (Bayer, Goettsch et al. 2003; Jacobs, Saxena et al. 2003) qui semble s'atténuer légèrement en présence de substrats de nature peptidique (Jacobs, Saxena et al. 2003). Mais ce rapprochement moléculaire, qui se traduit par des perturbations au niveau de l'hélice  $\alpha 4$  et le feuillet  $\beta 6$  du domaine catalytique constituant l'interface avec le domaine WW (Figure 1.14), n'indique pas pour autant une coopération fonctionnelle. On peut envisager que la présence de plusieurs sites d'interaction (par les phosphorylations multiples) au sein d'un même substrat de nature protéique, pourrait induire un rapprochement plus marqué et une coopération des deux domaines pour assurer la fonction biologique.



**Figure 1.14** : Variation des perturbations de déplacements chimiques des différents résidus de Pin1 lors de la liaison d'un phosphopeptide. Les perturbations sont graduées en fonction de leur intensité : en rouge, les résidus les plus fortement affectés par le substrat (d'après Jacobs, *J Biol Chem* 2003).

Nous allons décrire dans les paragraphes suivants le rôle moléculaire de Pin1 proposé dans l'interaction avec ses divers substrats phosphorylés.

### 2.2.2 – Des relations entre structure et fonction de Pin1 mal définies.

La structure globulaire des protéines est généralement associée directement à une fonction biologique (Thornton, Orengo et al. 1999) bien que des exemples de plus en plus nombreux de protéines sans structures globulaires ou comportant de longs fragments non repliés, assurant semble-t-il un rôle biologique prépondérant, viennent contredire cette théorie (Wright and Dyson 1999). Quoiqu'il en soit, de longue date, on a établi que structure et fonction étaient étroitement liées au sein des protéines et que la structure des protéines était impliquée dans divers mécanismes cellulaires comme la régulation de la reconnaissance de substrats et de l'activité enzymatique, le transport des protéines, le trafic des ions via les canaux ioniques, la transduction des signaux via les voies de signalisation intra- et inter-cellulaires, .... En ce qui concerne Pin1, bien que la structure tridimensionnelle soit connue et son rôle biologique bien établi, notamment par la connaissance des substrats reconnus par Pin1, permettant son implication dans un certain nombre de voies de signalisation, il manque des liens entre structure, effet moléculaire au niveau du substrat et fonction biologique.

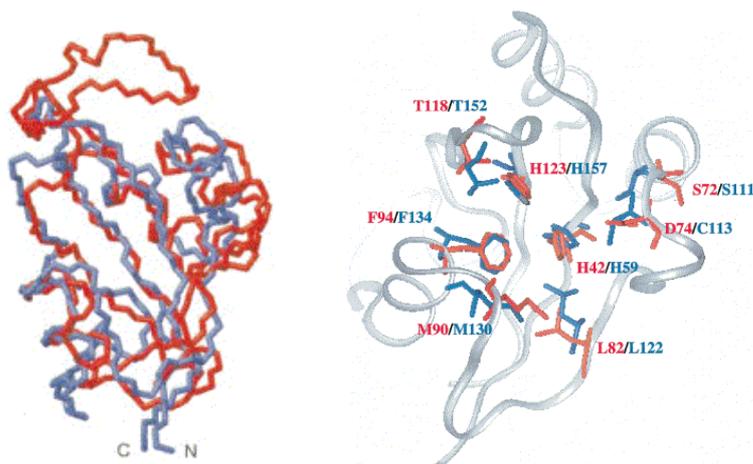
#### 2.2.2.1 – Pin1, une protéine phospho-spécifique?

Pin1 a été isolée à l'origine comme une protéine capable de se lier à la kinase NIMA. *In vitro*, la région de NIMA allant des résidus 280-396, localisée dans le domaine C-terminal de régulation, a été identifiée par *GST-pull down* dans l'interaction avec Pin1. Dans ces expériences, NIMA a été exprimée par voie recombinante et n'est donc pas phosphorylée. La phospho-spécificité de Pin1 a été proposée sur base des données cristallographiques mettant en évidence un *cluster* d'acides aminés basiques dans le domaine catalytique liant un ion sulfate (Ranganathan, Lu et al. 1997). La structure de Pin1 en complexe avec un phosphopeptide a permis de mettre en évidence des interactions phospho-spécifiques entre le domaine WW et le substrat (Verdecia, Bowman et al. 2000). Quel serait le mécanisme de reconnaissance moléculaire permettant à Pin1 de lier NIMA en l'absence de phosphorylation ? La présence d'un résidu acide en -1 de la proline, comme le glutamate ou l'aspartate, peut-il permettre la liaison de NIMA par Pin1 ?

#### ***Pin1, un facteur essentiel dans la régulation des substrats phosphorylés ?***

Pin1, en se liant de manière spécifique aux protéines phosphorylées au niveau de sites suivis d'un résidu Pro, catalyserait un changement conformationnel consécutif au processus de phosphorylation (Lu 2000). Cette fonction régulerait l'activité catalytique de la protéine substrat, sa déphosphorylation, sa localisation cellulaire, ou encore les interactions avec ses

autres partenaires. Ce mode d'interaction a été proposé comme un second facteur de régulation de la fonction des protéines en ajoutant un autre acteur au face à face kinases/phosphatases, incluant ainsi un degré de variabilité supplémentaire dans la fonction gouvernant la dynamique du cycle cellulaire. Pin1 pourrait, en quelque sorte, être nécessaire pour prévenir une activation mitotique non « planifiée ». Cependant, sa fonction biologique semble bien plus complexe puisque Pin1 est essentielle pour la croissance cellulaire et la différenciation (Lu, Hanes et al. 1996) alors que chez la levure, les cyclophilines et les FKBP peuvent être complètement inactivées sans induire la mort cellulaire. Il faut noter cependant que le gène de Pin1 est essentiel chez certains organismes comme la levure. On a également montré son importance dans les cellules humaines HeLa, alors que ce n'est pas le cas pour d'autres organismes. En général, les délétions de gènes codant pour des parvulines ne sont pas létales car des cyclophilines peuvent alors prendre en charge la fonction de Pin1, ce qui laisse supposer que certaines cyclophilines seraient capables de lier des motifs Ser/Thr-Pro phosphorylés (Fujimori, Gunji et al. 2001; Huang, Forsburg et al. 2001). Cependant, récemment, des parvulines ne possédant pas de spécificité pour les substrats phosphorylés, comme Par14, ont été identifiées et pourraient également remplacer la fonction de Pin1 (Sekerina, Rahfeld et al. 2000; Uchida, Takamiya et al. 2003). Par14 ne possède pas les caractéristiques qui confèrent à Pin1 sa phospho-spécificité. Dans sa partie N-terminale, un fragment de 35 résidus non replié et basique remplace le domaine WW de Pin1, gouvernant en grande partie la spécificité de Pin1 pour les phospho-protéines. La partie C-terminale de Par14 ne possède pas la grande boucle flexible  $\beta 4/\alpha 1$  responsable de la liaison du groupement phosphate des substrats dans le site actif de Pin1 (Figure 1.15). Il a été proposé que la partie N-terminale basique non structurée de Par14 pourrait, par analogie avec la structure des régulateurs transcriptionnels, être impliquée dans la liaison de l'ADN. D'ailleurs Par14 est localisée au niveau nucléaire, complexée à l'ADN (directement ou indirectement).



**Figure 1.15 :** Analogie structurale entre le domaine PPIase de Pin1 (rouge) et Par14 (bleu) montrant l'absence de la boucle responsable de la liaison du phosphate dans la structure de Par14, mais une grande conservation des résidus impliqués dans la liaison de la proline (L122, F134, M130 selon la numérotation de Pin1, en bleu) et de la liaison X-Pro (C113, H59, H157).

D'autre part, cette absence de spécificité pour le phosphate conférerait aux parvulines comme Par14, une plus large gamme de substrats. Etant donné leur spécificité distincte de substrats, leur structure et leur localisation cellulaire, Pin1 et Par14 pourraient avoir des fonctions différentes dans la cellule. Mais la grande spécificité de Pin1 pour les protéines phosphorylées restreint la gamme des substrats reconnus et serait, en quelque sorte, compensée par une spécificité moins importante de Par14 qui pourrait alors assurer, moins efficacement, la fonction de Pin1 quand celle-ci est déplétée dans la cellule.

#### 2.2.2.2 – Reconnaissance des motifs phosphoSer/Thr-Pro par Pin1, un rôle moléculaire obscur.

Dans la plupart des substrats, Pin1 reconnaît plusieurs motifs phosphoSer/Thr-Pro (Tableau 1.1) mais tous les motifs de ce type présents dans les substrats, ne sont pas systématiquement reconnus, suggérant un rôle de l'environnement de ces motifs dans le processus de reconnaissance par Pin1. Cependant ces observations sont en contradiction sur deux points avec les données structurales :

- Les structures de Pin1 ou de son domaine WW, en complexe avec des peptides phosphorylés (Verdecia, Bowman et al. 2000; Wintjens, Wieruszeski et al. 2001), ne montrent pas de contacts avec les résidus environnant les motifs phosphoSer/Thr-Pro, indiquant que la séquence ne joue pas un rôle prépondérant dans la liaison d'un motif phosphoSer/Thr-Pro particulier.
- On pourrait avancer l'hypothèse de l'accessibilité de ces sites, mais il faut souligner également que la plupart des motifs reconnus par Pin1 sont localisés dans des régions flexibles, non structurées, au sein du substrat, ce qui limite *a priori* les problèmes d'accessibilité.

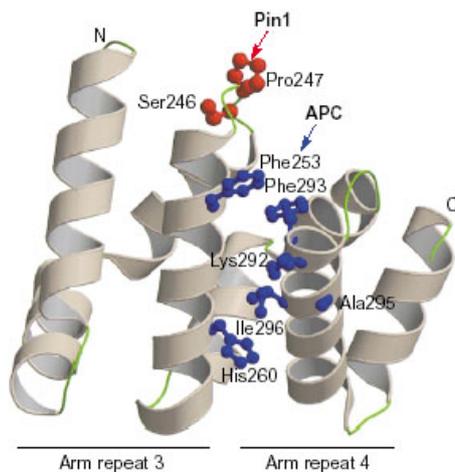
Comment, dans ce contexte, est dirigée la reconnaissance vers certains motifs phosphoSer/Thr-Pro particuliers ? Peut-on évoquer un processus de reconnaissance double impliquant l'intervention des deux domaines de Pin1 simultanément ?

### ***Effet catalytique sur l'isomérisation cis/trans ou simple liaison du domaine WW ?***

D'abord il faut rappeler que, dans des conditions normales d'expression, les deux domaines de Pin1 sont nécessaires à sa fonction essentielle *in vivo* (Lu, Zhou et al. 1999; Winkler, Swenson et al. 2000). Un changement structural étant difficile à mettre en évidence, il n'est pas certain que Pin1 exerce une activité isomérase sur tous ses substrats. Il a été montré que Pin1 catalysait une transition conformationnelle dans un peptide de Cdc25, un des régulateurs du complexe CDK2/Cycline, et serait ainsi essentielle au bon déroulement du cycle cellulaire (Crenshaw, Yang et al. 1998; Shen, Stukenberg et al. 1998). L'activité catalytique de Pin1 serait cruciale pour induire des changements conformationnels dans la partie N-terminale de Cdc25C. Cette partie, qui régule l'activité de la phosphatase, est phosphorylée lors de la transition G<sub>2</sub>/M mais n'a *a priori* pas de structure tertiaire stable. La phosphatase PP2A, qui régule négativement Cdc25C, ne serait capable de déphosphoryler le motif phosphoThr-Pro de Cdc25C que lorsque la proline est en *trans* (Zhou, Kops et al. 2000). Il a été montré que Pin1 stimulait la déphosphorylation de Cdc25 par PP2A de manière catalytique. Ces données proviennent d'une méthode de détection indirecte de l'activité isomérase par digestion différentielle de Cdc25 par une protéase sensible aux changements conformationnels en présence ou en absence de Pin1. La preuve irréfutable d'une activité PPIase ne pourra être obtenue que par résolution de la structure tridimensionnelle de cette protéine, qui peut s'avérer très compliquée, notamment parce que les motifs d'interaction entre Pin1 et Cdc25 sont localisés une région non structurée de la protéine. La présence d'une proline de Cdc25 en *cis* ainsi que des changements de conformation influençant son activité pourraient représenter deux éléments-clés potentiels de la régulation par Pin1.

Outre son activité isomérase, le rôle de Pin1 pourrait se limiter simplement à la liaison des cibles par le domaine WW sans catalyser l'isomérisation d'une liaison phosphoSer/Thr-Pro. Cette simple association contribuerait à masquer certains sites sur des protéines et à moduler leur activité ou leur liaison à d'autres partenaires. Il a en effet été observé que Pin1 est antagoniste de l'effet de stimulation de la phosphorylation du substrat Cdc25 par le complexe CDK/cycline obtenu en présence de p13<sup>SUC1</sup> (Patra, Wang et al. 1999), en entrant en compétition moléculaire directe avec SUC1/CKS, une sous-unité régulatrice des CDK, par l'intermédiaire de son domaine WW (Landrieu, Odaert et al. 2001). Cette effet de liaison est

également observé avec le facteur de transcription NFAT. La liaison de Pin1 à NFAT inhibe sa déphosphorylation par la calcineurine en bloquant son accès à NFAT (Liu, Youn et al. 2001). Pin1 se lie également à la  $\beta$ -caténine au niveau du motif phosphoSer246-Pro247 du domaine armadillo de celle-ci, localisé à proximité du site de liaison de l'APC, empêchant ainsi l'interaction entre ces deux partenaires (Ryo, Nakamura et al. 2001). Pin1 régule de même la stabilité des protéines par compétition avec des ligases d'ubiquitine pour la liaison à un site phosphorylé, par exemple la ligase d'ubiquitine Mdm2 pour p53 (Wulf, Liou et al. 2002; Zacchi, Gostissa et al. 2002; Zheng, You et al. 2002).



**Figure 1.16 :** Compétition pour la liaison de la  $\beta$ -caténine entre Pin1 et l'APC au niveau de deux sites d'interaction proches. Le site de liaison de Pin1 à la  $\beta$ -caténine est le seul site localisé dans une région organisée au sein d'une protéine substrat.

### ***Localisation des sites de reconnaissance de Pin1 dans des régions non structurées.***

Un mécanisme de changement conformationnel dans les motifs phosphoSer/Thr-Pro a été proposé comme nouveau mécanisme de régulation faisant intervenir des PPIases comme Pin1. Il faut noter que l'implication d'un tel changement de conformation au sein de protéines structurées et, par extension, au sein de protéines non structurées, est encore mal défini dans la mesure où les études ont porté jusqu'alors sur des peptides. On retrouve les kinases/phosphatases impliquées dans de nombreux processus cellulaires comme la régulation du cycle cellulaire, la transcription, les mécanismes de signalisation et dans de nombreuses pathologies (maladie d'Alzheimer, cancer...). La phosphorylation de résidus tels que les sérines et thréonines mais aussi les tyrosines constitue la formation de motifs de reconnaissance pour des modules de liaison chargés de lier le substrat phosphorylé et d'augmenter sa concentration locale au niveau du site catalytique de la protéine hôte (Yaffe 1999). La double reconnaissance des motifs Ser/Thr-Pro, dans un premier temps, par des kinases spécifiques dirigées contre des résidus précédant une proline, et dans un deuxième temps, après phosphorylation, par la peptidyl-prolyl isomérase Pin1, qui catalyserait l'interconversion *cis/trans* de ces motifs, s'impose comme un nouveau mécanisme de

régulation de la fonction des protéines avec des implications importantes comme la régulation du cycle cellulaire (Lu 2000; Lu, Liou et al. 2002). D'un point de vue moléculaire, les études de l'interaction entre Pin1 et ses multiples partenaires ne sont pas très avancées, entre autre, parce que les motifs phosphoSer/Thr-Pro reconnus par Pin1 se trouvent souvent dans des parties protéiques qui *a priori* ne sont pas organisées dans une structure globulaire. Tel est le cas pour la partie régulatrice de Cdc25, dont seule la structure de la partie catalytique est connue, ainsi que pour la partie N-terminale de p53 et la partie régulatrice de c-Jun, dont la zone d'interaction avec l'ADN s'organise en « *leucine zipper* » mais où le domaine régulateur n'a aucune structure tertiaire stable selon les serveurs de prédiction de structures secondaires. Pour tous ces exemples, les motifs reconnus se trouvent dans une partie non organisée qui joue néanmoins un rôle important de régulateur de la fonction de la protéine. D'autre part, l'absence d'organisation des régions protéiques impliquées dans l'interaction avec Pin1 soulève la question de l'intérêt d'une isomérisation catalytique dans la régulation de la fonction de ces protéines, dans la mesure où celle-ci ne provoquerait probablement pas un grand bouleversement structural, mais argumente plutôt en faveur d'un rôle de chaperonne, dont la fonction première est de se lier aux protéines non repliées, ou partiellement repliées, ainsi qu'aux protéines « mal repliées » pour prévenir les phénomènes d'agrégation. La fonction de stabilisation de Pin1 par contact moléculaire, en revanche, paraît plus classique, et permet de masquer, par simple liaison, un site de reconnaissance vis-à-vis d'un autre partenaire.

### ***Intérêt d'une catalyse de l'interconversion cis/trans dans des régions flexibles portant des prolines majoritairement en conformation trans ?***

Un paradoxe semble s'instaurer entre la capacité des PPIases à accélérer l'isomérisation *cis/trans* des prolines et la *trans*-spécificité de certaines protéines. En effet, en l'absence de protéines *cis*-spécifiques, quelle est l'utilité de catalyser l'interconversion *cis/trans* dans des régions flexibles caractérisées essentiellement par une majorité de conformations *trans* ? Par exemple, la phosphatase PP2A, qui est dirigée vers des résidus phosphoSer/Thr contenant une proline en +1, est dépendante de la conformation de cette liaison et déphosphoryle uniquement l'isomère en conformation *trans* (Zhou, Kops et al. 2000). L'isomérisation par Pin1 de ces liaisons spécifiques faciliterait leur déphosphorylation par PP2A (Zhou, Kops et al. 2000) et régulerait de cette manière la déphosphorylation de Cdc25. La structure du domaine WW a montré un élément très important quant à la fonction éventuelle de la protéine. En effet, non seulement une impossibilité stérique ne permet pas d'inclure une

proline en *cis* dans cette structure, mais le domaine ne lie pas du tout les motifs phosphoSer/Thr-Pro en *cis*. Ceci invalide l'hypothèse selon laquelle Pin1 se lierait grâce à son domaine WW à une proline en *cis*, qui serait dans un deuxième temps isomérisée vers sa forme *trans* par le domaine catalytique pour permettre l'action de PP2A. Dans tous les cas, on peut poser la question de l'utilité d'augmenter la population de formes *trans* dans des régions en contenant déjà au moins 90% (Wintjens, Wieruszeski et al. 2001).

### ***Pin1 et la stabilisation de l'architecture neuronale.***

La même absence de structure a été constatée pour la protéine microtubulaire tau, qui résiste à toute tentative de cristallisation, et pour laquelle un site de phosphorylation a été identifié dans la région riche en prolines (phosphoThr231-Pro232), impliquée dans la régulation de la liaison de tau aux microtubules (Lu, Wulf et al. 1999). La protéine Pin1, outre son rôle dans le cycle cellulaire, est un second facteur de régulation de la liaison de la protéine tau aux microtubules, le premier étant l'état de phosphorylation de tau. Les protéines tau hyperphosphorylées sont les constituants majeurs des neurones en dégénérescence neurofibrillaire (DNF) retrouvés dans la maladie d'Alzheimer. La liaison de Pin1 via les phospho-épitopes MPM-2 permettrait de restaurer l'activité biologique de tau (hyper)phosphorylée (Lu, Wulf et al. 1999). L'implication de Pin1 dans les mécanismes de régulation de la liaison de tau aux microtubules et de la progression du cycle cellulaire pourrait ne pas être fortuite. Il est en effet proposé qu'une réactivation anormale du cycle cellulaire dans les neurones matures conduirait à la mort cellulaire et au processus de dégénérescence neurofibrillaire observé dans la maladie d'Alzheimer (Husseman, Nochlin et al. 2000; Delobel, Flament et al. 2002).

\*\*\*\*\*

Il est maintenant assez clair que Pin1 fonctionne en partie par sa liaison à ses substrats. Le domaine WW entre alors en compétition avec d'autres protéines pour la liaison aux sites phosphorylés. Qu'en est-il alors de l'activité du domaine catalytique ? Il pourrait simplement s'agir d'un second site de liaison de plus faible affinité, dont l'occupation nécessiterait une augmentation de la concentration locale obtenue par liaison du domaine WW. De manière générale, les interactions de Pin1 avec ses substrats sont assez bien décrites d'un point de vue fonctionnel mais le rôle des deux domaines reste obscur, notamment celui du domaine

« catalytique », et mérite davantage d'investigations au niveau structural. Une étroite collaboration avec l'équipe du Dr Luc Buée (U422, Inserm, Lille) qui travaille depuis plusieurs années sur la caractérisation moléculaire des maladies liées à l'âge et, notamment, la détermination des relations entre la protéine tau et la maladie d'Alzheimer, nous a conduits à orienter nos recherches dans la compréhension des mécanismes régissant la reconnaissance par Pin1 des protéines phosphorylées, plus particulièrement vers la protéine tau. Leur expérience en biologie cellulaire, avec la mise au point de modèles cellulaires mimant un neurone en dégénérescence et l'état (hyper)phosphorylé de la protéine tau, nous donne la possibilité de confirmer nos données structurales. D'autre part, les interactions décrites entre Pin1 et tau sont assez contradictoires avec les données structurales exposées précédemment. Le paragraphe suivant est consacré à la description moléculaire de la protéine tau, à son rôle biologique, son implication pathologique dans la maladie d'Alzheimer de manière assez succincte, tant la littérature est prolifique sur ce sujet. Nous nous intéresserons plus particulièrement à un éventuel rôle de Pin1 dans le contexte pathologique de la maladie d'Alzheimer.

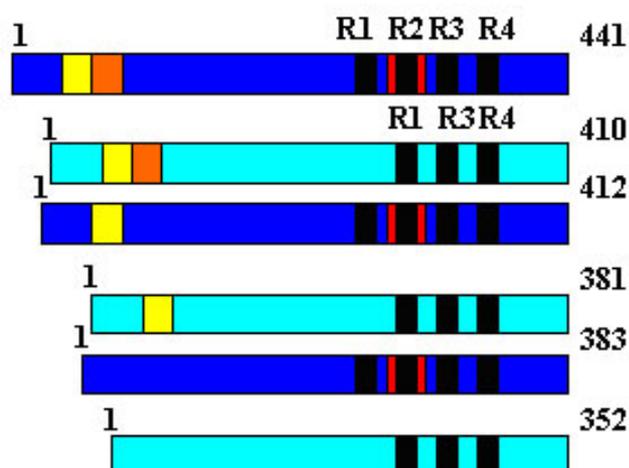
### **III – Les interactions de Pin1 avec la protéine microtubulaire tau.**

#### **3.1 – La protéine tau : de la protéine microtubulaire au marqueur pathologique associé à la dégénération neuronale.**

##### **3.1.1 – Fonction biologique de la protéine tau (Buee 2000).**

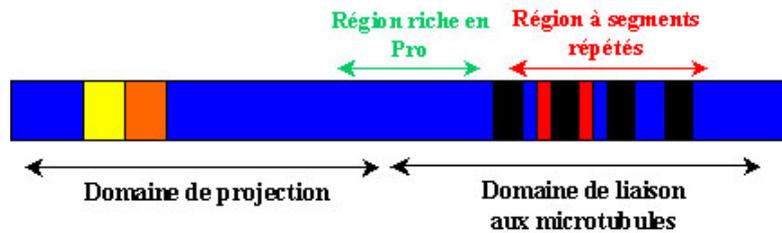
La protéine tau fait partie de la famille des MAP (*Microtubule-Associated Proteins* ou protéines associées aux microtubules). Les microtubules sont impliqués dans la stabilisation de la partie axonale du neurone favorisant les échanges avec le milieu extracellulaire (apport en glucose, oxygénation...). La fonction de la protéine tau, relative à la liaison et à la stabilisation des microtubules, permet aux monomères de tubulines de s'organiser en microtubules dans les neurones. La protéine tau existe sous six isoformes dans le cerveau humain adulte, obtenues par épissage alternatif à partir d'un même gène. Le gène des protéines tau est localisé sur le chromosome 17. Le transcrit primaire contient 16 exons. Les exons 2, 3 et 10 sont épissés de manière alternative et sont spécifiques du tissu cérébral adulte. L'épissage alternatif de ces 3 exons conduit à 6 combinaisons (2-3-10-; 2+3-10-; 2+3+10-; 2-3-10+; 2+3-10+; 2+3+10+ ). Chacune de ces isoformes a vraisemblablement un

rôle physiologique particulier puisqu'elles sont différenciellement exprimées au cours du développement et distribuées différemment dans les sous-populations neuronales. Une seule isoforme est présente chez le fœtus (isoforme fœtale) et ne comporte pas d'inserts codés par les exons 2, 3 et 10. Ensuite, les autres isoformes dont la longueur des séquences varie de 352 à 441 acides aminés vont apparaître au cours du développement (Figure 1.17).



**Figure 1.17 :** Représentation schématique des six isoformes de la protéine tau, produites par épissage alternatif de trois exons (fragment de séquence codé par l'exon 2 en jaune, exon 3 en orange et exon 10 en rouge). Les isoformes de tau diffèrent donc par la présence ou l'absence d'un segment répété dans la partie C-terminale R2 (31 AA,) codé par l'exon 10 et/ou d'une insertion de 28 (exon 2+) ou 58 (exons 2+, 3+) résidus dans la partie N-terminale.

La liaison aux microtubules a lieu par l'intermédiaire d'un domaine C-terminal qui comporte des segments de 18 acides aminés (*repeat domain*) répétés 3 ou 4 fois de façon imparfaite et séparés par des segments plus variables de 13 ou 14 acides aminés (*inter-repeat domain*). Les segments répétés contribuent différemment à la liaison aux microtubules : le segment répété R1 est le plus important. La région entre les segments R1 et R2 possède une capacité intrinsèque à lier les microtubules. Une région riche en prolines, localisée en amont du domaine à segments répétés, joue un rôle important dans la régulation de l'affinité de ce domaine pour les microtubules mais n'est pas capable de lier seule les microtubules. Un peptide de 7 acides aminés  $K^{224}KVAVVR^{230}$ , localisé dans la région riche en prolines, a une haute affinité pour les microtubules en combinaison avec le domaine à segments répétés (Goode 1997) (Figure 1.18). Une hypothèse proposée pour expliquer cette observation est que le domaine à segments répétés interagirait avec le domaine riche en prolines lors de la liaison aux microtubules (Goode 1997). Au niveau des tubulines, on sait uniquement que les 12 derniers acides aminés sont impliqués dans l'interaction avec tau (Chau 1998).



**Figure 1.18 :** Représentation schématique des domaines fonctionnels de la protéine tau. Les domaines protéiques de l'isoforme la plus longue de tau sont le domaine de projection dans la partie N-terminale, impliqué dans la liaison d'autres protéines, et le domaine de liaison aux microtubules constitué par les trois ou quatre segments répétés (indiqués par une double flèche rouge) et le domaine riche en prolines (indiqué par une double flèche verte).

La protéine tau subit des modifications post-traductionnelles au cours du développement, en particulier par des kinases/phosphatases. La phosphorylation de tau, en particulier dans le domaine à segments répétés, réduit l'affinité de tau pour les microtubules et provoque par conséquent leur dissociation. La phosphorylation de la Ser262 diminue très significativement la liaison de tau aux microtubules mais la seule phosphorylation de ce résidu ne permet pas d'abolir complètement la liaison (Biernat, Gustke et al. 1993). Il existe 80 sites de phosphorylation qui peuvent être regroupés en deux catégories : les sites inclus dans des motifs Ser/Thr-Pro localisés principalement dans la région riche en Pro (7 sites sur les 17 motifs Ser/Thr-Pro présents dans la protéine) et les sites Ser/Thr dispersés sur l'ensemble de la protéine.

***Aperçu de la « structure » moléculaire de la protéine tau. Le paradoxe avec le dogme structure/fonction.***

La protéine tau a été l'une des premières protéines pour laquelle on a parlé d'absence de structure intrinsèque, mais ce fût plusieurs années après la caractérisation de son rôle biologique. Paradoxalement, alors que son rôle biologique dans la liaison des microtubules et son rôle pathologique dans un certain nombre de maladies (regroupées sous le terme de tauopathies) sont largement décrits, peu de données structurales existent, et l'absence de structure elle-même en est la cause. Différentes études, impliquant des techniques macroscopiques de détermination structurale (SAXS, FT-IR, CD), ont permis de définir une absence de structure globale (Schweers, Schonbrunn-Hanebeck et al. 1994). De plus en plus de protéines « non structurées » sont mises à jour, présentant souvent un intérêt remarquable dans la compréhension des phénomènes biologiques et pathologiques (Plaxco and Gross

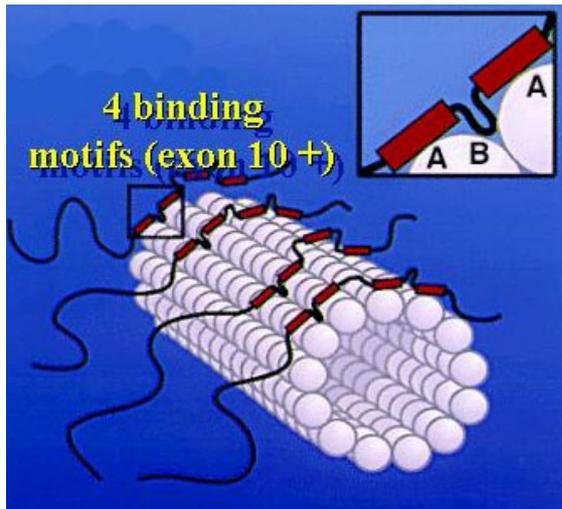
1997; Wright and Dyson 1999) et impliquant le développement de méthodes de détermination structurale pour ces protéines. La cristallographie et la RMN sont des techniques de choix pour la résolution de structures tridimensionnelles, mais, dans le cas des protéines non structurées, la flexibilité importante de celles-ci conduit à une absence de cristallisation et à une faible dispersion des déplacements chimiques. Des études par RMN ont pu être réalisées sur certaines protéines non structurées, comme l' $\alpha$ -synucléine, impliquée dans le processus neurodégénératif de la maladie de Parkinson (Bussel and Eliezer 2001). Pourtant aucune étude par RMN n'a encore été réalisée sur la protéine tau entière. En plus de sa nature non structurée, la protéine tau, dans son isoforme la plus longue, est une protéine de 441 résidus : deux paramètres qui font de cette protéine un mauvais candidat pour des études par RMN.

Depuis plus d'une dizaine d'années, de nombreuses études mettent en évidence la présence dans les cellules de protéines sans structure globulaire. Dès lors, l'existence de telles protéines vient contrarier la théorie, jusqu'alors bien établie, selon laquelle la fonction biologique des protéines était directement liée à leur structure tridimensionnelle, en se basant sur le principe de « la clé et la serrure » qui avait été introduit par Fischer en 1894 pour expliquer l'extrême spécificité des enzymes impliquées dans l'hydrolyse des sucres. Pourtant, les données actuelles indiquent la présence dans les cellules d'un grand nombre de régions protéiques, et même de protéines entières, ne possédant aucune structure tertiaire stable, et, dans certains cas, sans aucune structure secondaire, dans des conditions physiologiques. Chez les organismes eucaryotes, notamment, on dénombrerait plus de 30% de protéines ayant au moins une région « désordonnée » de plus de 50 résidus consécutifs, et proportionnellement, plus de protéines présentant de telles régions chez les eucaryotes que chez les bactéries par exemple. Il a été proposé, pour résoudre la question de ce paradoxe au dogme des relations structure/fonction, que la grande plasticité de ces protéines, liée à l'absence de structure globulaire ordonnée, serait plutôt un avantage du point de vue fonctionnel car elle leur permettrait d'interagir avec un plus grand nombre de partenaires (Kriwacki, Hengst et al. 1996), leur confèrerait de remarquables propriétés de mobilité (Daughdrill, Chadsey et al. 1997) et que le passage de l'état libre à l'état lié, s'accompagnant d'une transition d'un état désordonné à un état ordonné, serait un nouveau mécanisme de régulation *in vivo* (Kriwacki, Hengst et al. 1996). La plupart du temps, il s'agit de domaines protéiques intervenant dans la régulation de l'activité transcriptionnelle, qui adoptent une structure stable lors de la liaison à l'ADN ou l'ARN, parfois couplée à une structuration du substrat lui-même (Spolar and Record 1994). Les protéines structurées sont également capables d'adaptation, notamment au

niveau de régions exhibant un plus haut degré de flexibilité, permettant des fluctuations structurales en fréquence et en magnitude, et coïncidant généralement avec les régions impliquées dans les interactions avec d'autres protéines (Feher and Cavanagh 1999).

Il a été montré que ces protéines flexibles ont d'importantes fonctions de régulation dans la cellule avec, comme preuve, une grande conservation au cours de l'évolution, et dont l'avantage fonctionnel résiderait dans une grande capacité d'adaptation, leur permettant d'interagir avec des substrats divers (Kriwacki, Hengst et al. 1996). Ce qu'il faut souligner, c'est que la flexibilité de telles protéines n'est pas synonyme d'absence de structure, autrement dit de structures totalement aléatoires. En réalité, ces protéines ont une structure qui leur est propre et l'échantillonnage conformationnel de leur squelette n'est pas si différent des protéines structurées (Smith, Fiebig et al. 1996), ni particulièrement astronomique (correspondant, théoriquement, à environ  $10^{100}$  conformations pour une protéine de 100 résidus). Elles se caractérisent pas une structure intrinsèque, appelée parfois structure résiduelle, qui est définie par un degré important de liberté mais tout de même limité, indiquant un repliement qui n'est pas aléatoire mais, d'une certaine manière, organisé (Plaxco and Gross 2001). Un des paramètres limitant est l'impossibilité stérique d'adopter certaines conformations, ce qui implique que, même dans de hautes concentrations d'urée, pour lesquelles la plupart des interactions hydrophobes maintenant la structure native de la protéine sont rompues, un certain degré d'organisation persiste. Une faible proportion de structure est difficilement décelable par les techniques classiques, notamment les techniques macroscopiques. Mais les techniques spectroscopiques, comme la RMN, sont capables d'identifier précisément des régions partiellement organisées.

Ainsi, on constate que les protéines dites « non structurées », adoptent une structure stable lors de la liaison de leur substrat. La protéine tau, par exemple, acquiert une structure particulière lors de la liaison aux microtubules par l'intermédiaire d'une interaction du domaine à segments répétés avec les monomères de tubuline (Panda, Goode et al. 1995). Il a été montré que le segment répété R1 et le domaine R1-R2 entre les deux segments répétés intervenait de manière prépondérante dans la stabilisation des microtubules. Un mécanisme moléculaire a été proposé selon lequel la liaison de tau à un monomère individuel de tubuline dans les microtubules induisait un changement conformationnel de la protéine qui renforce le pontage entre les monomères de tubuline par le fragment localisé entre les segments R1-R2.



**Figure 1.19 :** Modèle de liaison de la protéine tau aux microtubules au niveau de la région des domaines à segments répétés impliqués directement dans la liaison (d'après Panda, *Biochemistry* 1995).

Par ailleurs, l'implication au niveau pathologique de protéines sans structure stable ou (partiellement) dénaturées, augmente l'intérêt concernant la caractérisation des protéines natives non structurées et la compréhension des phénomènes menant à la perte de structure et parfois, à l'agrégation, notamment dans les processus neurodégénératifs pour lesquels de nombreuses corrélations ont été établies entre absence de structure et agrégation. On peut citer par exemple la maladie de Parkinson, liée à la formation d'agrégats d' $\alpha$ -synucléine appelés corps de Lewy et la maladie d'Alzheimer pour laquelle on trouve deux types d'agrégats caractéristiques : les dépôts amyloïdes formés de peptides  $A\beta$ , issus de la protéolyse de l'APP (*Amyloid Precursor Protein*), agrégés sous forme de feuillets  $\beta$  et les PHFs de protéines tau hyperphosphorylées. Cependant, la flexibilité naturelle de ces protéines n'implique pas nécessairement une propension à la perte de structure et à l'agrégation.

### 3.1.2 – La protéine tau, marqueur moléculaire de la maladie d'Alzheimer.

La protéine tau a fait l'objet de nombreuses recherches notamment ces dernières années depuis la découverte de son implication dans de nombreuses maladies neurodégénératives regroupées sous le terme générique de « tauopathies » dont fait partie la maladie d'Alzheimer. Les données les plus récentes sur la protéine tau concernent son implication dans les processus neurodégénératifs et les différents facteurs qui y sont liés. En revanche, très peu de données structurales sont disponibles à l'heure actuelle, pour les raisons que nous avons invoquées précédemment.

L'apoptose, ou mort programmée de la cellule, est un processus moléculaire et morphologique normal, menant à la destruction cellulaire contrôlée par la cellule elle-même.

Cette forme de mort cellulaire (l'autre forme étant la nécrose) est un phénomène naturel jouant un rôle essentiel dans la régulation des populations cellulaires des tissus et se trouve sous contrôle physiologique ou pathologique. En effet, des dérégulations dans le processus d'apoptose sont associées à un certain nombre de pathologies telles que l'immunodéficience, les maladies auto-immunes, les maladies neurodégénératives et le cancer.

Dans les désordres neurologiques à l'origine de certaines maladies neurodégénératives, il a été proposé que la dégénérescence neurofibrillaire survienne à la suite d'une réactivation du cycle cellulaire dans les neurones matures du cerveau adulte (Vincent 1996; Illenberger, Zheng-Fischhofer et al. 1998; Delobel, Flament et al. 2002), qui conduit à l'apoptose pour des cellules post-mitotiques différenciées. La maladie d'Alzheimer, en particulier, provient de désordres importants et progressifs d'ordre neurologique caractérisés par une perte neuronale sévère et deux types de lésions cérébrales, les plaques séniles et les fibres (ou NFT pour *neurofibrillary tangles*). Les plaques séniles résultent de l'accumulation extracellulaire sous forme de dépôts amyloïdes de peptide A $\beta$ , neurotoxique, provenant de l'APP ( *$\beta$ -amyloid precursor protein*). Les NFT correspondent, dans la maladie d'Alzheimer, à l'agrégation de protéine tau anormalement phosphorylée associée aux microtubules présente sous forme de filaments dits PHF (*paired helical filaments*). Contrairement aux plaques séniles dont la distribution touche un grand nombre d'aires du cerveau de manière variable (cortex cérébral et aires sous-corticales), les PHF sont représentés de manière très localisée dans certaines aires cérébrales bien précises. Etablir un diagnostic en faveur de la maladie d'Alzheimer nécessite de mettre en évidence la présence de ces deux formes de lésions cérébrales dans certaines régions spécifiques du cortex, nécessairement par examens *post-mortem*.

Il faut noter que la présence de PHF ou de protéine tau agrégée en général, n'est pas le propre de la maladie d'Alzheimer mais de nombreuses autres tauopathies (on en dénombre actuellement plus d'une vingtaine) impliquant un dysfonctionnement de la protéine tau. Dans la maladie d'Alzheimer, les six isoformes de protéine tau sont hyperphosphorylées et agrègent en formant les PHF. Dans d'autres maladies neurodégénératives, c'est soit la forme 3R, soit la forme 4R phosphorylée qui agrège. La régulation des isoformes joue un rôle important dans la morphologie cellulaire et le profil de phosphorylation de tau (Mailliot, Sergeant et al. 1998; Mailliot, Bussiere et al. 2000). La compréhension des mécanismes de formation de protéines tau pathologiques apportera de nouveaux diagnostics et traitements pour plusieurs maladies neurodégénératives.

***Structure moléculaire des PHF.***

Des études structurales ont été réalisées sur des fragments de la protéine tau, en particulier au niveau de la région d'intérêt, celle impliquée dans l'interaction fonctionnelle avec les microtubules (von Bergen, Friedhoff et al. 2000). De petits peptides ou la totalité du domaine à segments répétés ont permis de définir que cette région adoptait préférentiellement une structure en feuillets  $\beta$  dans les PHF consécutivement aux premières étapes de l'agrégation (Barghorn, Davies et al. 2004). Il faut noter cependant, que cette région isolée est intrinsèquement insoluble, favorisant ainsi le processus d'agrégation, notamment en l'absence du reste de la protéine. La partie C-terminale, en particulier, apparaît comme essentielle à la solubilisation de tau, car, lorsque la partie C-terminale de tau est digérée par les caspases, le fragment résultant de 421 résidus montre une cinétique d'agrégation plus rapide (Berry, Abraha et al. 2003; Gamblin, Chen et al. 2003). La détermination de la structure des fibres de tau par diffraction électronique a conduit à un modèle moléculaire présentant des feuillets  $\beta$  « empilés » entre les différentes molécules constituant les fibres, localisés dans certaines régions de la protéine, essentiellement dans le domaine à segments répétés. D'autres régions de la protéine, en revanche, restent accessibles à différentes protéases même dans les fibres (Goux 2002; Goux, Kopplin et al. 2004).

***Corrélation entre le niveau de PHF et les pertes cognitives. Le lien entre dépôts amyloïdes et les PHF (Dickson 2004).***

Deux caractéristiques moléculaires signent la physiopathologie de la maladie d'Alzheimer. Pendant longtemps on a cru que le peptide  $\beta$ -amyloïde, formant les plaques « séniles » extra-neuronales, jouait un rôle majeur dans la maladie mais, force est de constater qu'il n'y avait aucune corrélation entre la quantité de plaques extra-cellulaires et les symptômes tels que les pertes cognitives, il a fallu envisager un mécanisme impliquant les NFT, puisque la quantité de NFT, en revanche, est bien corrélée avec les pertes cognitives. Récemment, on a pu établir le lien manquant entre ces deux lésions neuropathologiques.

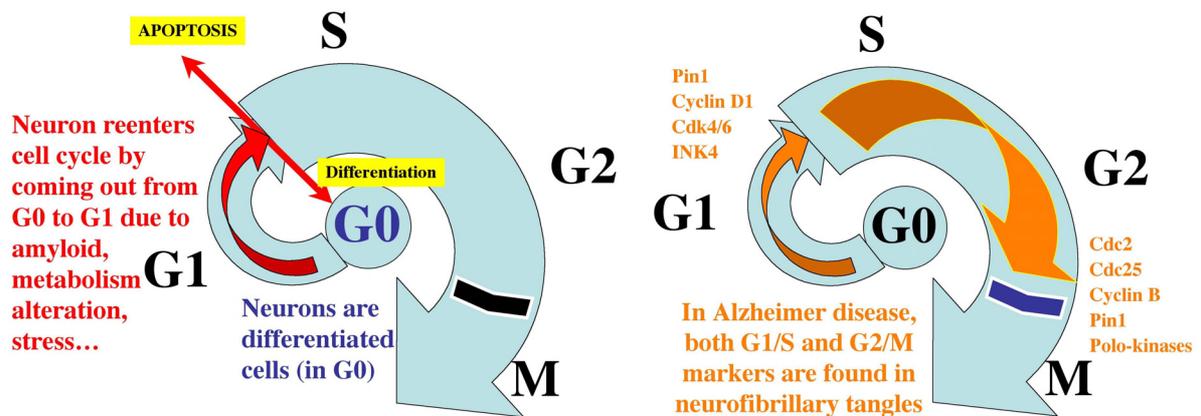
Bien que le rôle biologique précis de la protéine APP normale reste obscur, nous savons maintenant que plusieurs types de cellules et tissus produisent l'APP. La protéine fonctionne à travers la membrane externe des cellules par l'intermédiaire d'une courte partie saillante dans le milieu intracellulaire et d'une longue partie dans l'espace extracellulaire. Le peptide  $\beta$ -amyloïde, quant à lui, est coupé à partir de la section de l'APP qui « enjambe » la membrane cellulaire. Des travaux ont montré que l'APP est coupée d'une ou 2 façons. Dans un des

processus, l'APP est clivée en deux étapes : d'abord, une  $\beta$ -sécrétase coupe la protéine, puis un des fragments résultant est ensuite coupé par une  $\gamma$ -sécrétase, aboutissant ainsi au peptide  $\beta$ -amyloïde, fragment de 40-42 acides aminés, toxique. La thèse de l'apoptose comme mécanisme conduisant à la mort neuronale est quelque peu controversée mais de plus en plus de preuves indiquent qu'un tel mécanisme serait en cause dans le processus neuropathologique. Une approche pour étudier l'apoptose consiste à examiner les enzymes impliquées dans cette mort cellulaire programmée telles que les caspases (*cysteine aspartyl proteases*) participant notamment à une cascade de protéolyses intracellulaires, dont celle des protéines du cytosquelette et des protéines membranaires. L'activation des caspases est déclenchée par un double mécanisme apoptotique : un mécanisme intra-cellulaire menant à l'altération du réticulum endoplasmique et des mitochondries, provoqué par le peptide  $A\beta$  intra-cellulaire et un mécanisme extra-cellulaire impliquant des voies de signalisation par l'activation de récepteurs membranaires par le peptide  $A\beta$  extra-cellulaire. Une des conséquences de l'activation des caspases est le clivage de la protéine tau dans sa partie C-terminale. Cette forme tronquée de 20 résidus C-terminaux de la protéine se caractérise par une cinétique d'agrégation plus rapide et pourrait être à l'origine des NFT (Gamblin, Chen et al. 2003).

### ***Un nouveau mécanisme d'apoptose du neurone ? La dégénérescence neurofibrillaire (Hamdane, Delobel et al. 2003).***

Une réactivation aberrante du cycle cellulaire dans les neurones a été proposée comme mécanisme conduisant à la neurodégénérescence observée dans la maladie d'Alzheimer et se caractérise par la sortie de la phase de quiescence  $G_0$  et l'entrée dans le cycle cellulaire en phase  $G_1$  avec expression de cycline D1. Un grand nombre de cellules apoptotiques sont retrouvées dans le cerveau de patients atteints de la maladie d'Alzheimer, ainsi que des marqueurs précoces de l'apoptose. D'autre part, des marqueurs des transitions  $G_1/S$  (cycline D1, Cdk4/6 ...) et  $G_2/M$  (Cdc2, Cdc25, cycline B, kinases Polo) sont également détectés alors que seuls les marqueurs de la transition  $G_1/S$  sont détectés dans des pathologies liées à une réactivation du cycle cellulaire, suggérant un mécanisme apoptotique anormal. En  $G_0$ , les neurones ne se divisent pas et sont différenciés. En général, les cellules différenciées arrivant au *checkpoint*  $G_1/S$ , sont redirigées soit vers la différenciation, soit vers l'apoptose. Ici, dans la dégénération neuronale de type Alzheimer, la présence de marqueurs  $G_2/M$  indique un

cycle cellulaire anormal dans lequel le processus apoptotique serait avorté, et les neurones « agoniseraient » alors dans un processus dit de « dégénérescence neurofibrillaire ».



**Figure 1.20** : A gauche, mécanisme d'apoptose classique et à droite, mécanisme proposé pour la dégénérescence neurofibrillaire (d'après Hamdane, *Biochem Pharmacol* 2003).

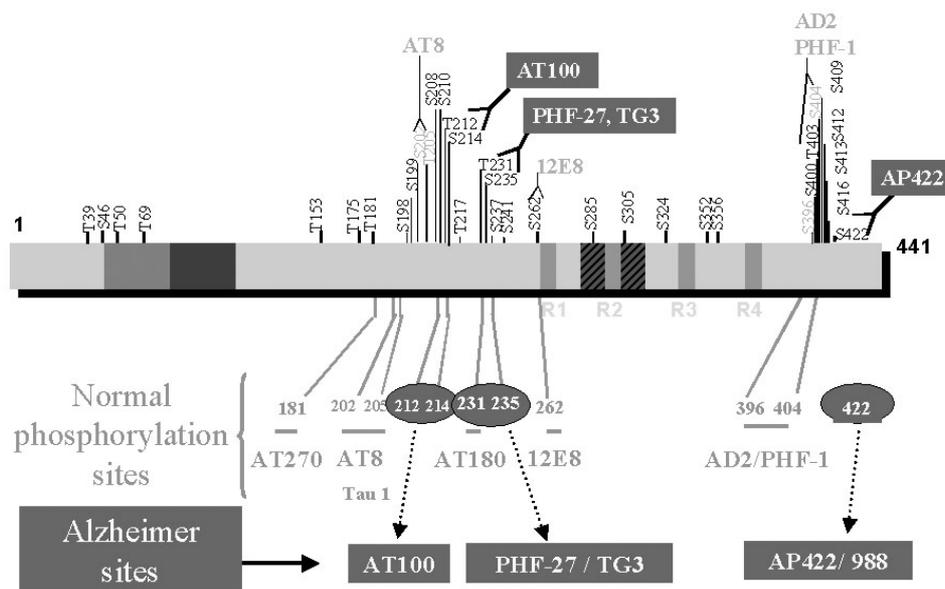
La phosphorylation anormale est la caractéristique majeure des PHF de protéines tau agrégées, observées dans la maladie d'Alzheimer. Trois phospho-épitopes sont clairement identifiés et détectés par les anticorps AT100, TG-3 et AP422. De plus, la présence de phospho-épitopes mitotiques, tel que Thr231, indiquent une implication de kinases mitotiques dans la phosphorylation de tau pouvant conduire à d'éventuels changements conformationnels et l'agrégation sous forme de filaments. L'hypothèse mitotique est également renforcée par la forte réactivité des PHF vis-à-vis de l'anticorps monoclonal MPM-2 spécifique des phospho-épitopes mitotiques (Vincent 1996). Les kinases de la famille des CDK, GSK3 $\beta$ , et MAPK pourraient phosphoryler les sites inclus dans des motifs Ser/Thr-Pro. La kinase CDK5 par exemple, est active principalement dans les neurones et est capable de phosphoryler tau. Elle est activée par une protéine régulatrice, p25, générée par clivage de p35 à la calpaïne, au cours du processus d'apoptose. La surexpression de CDK5/p25 dans des cultures de cellules neuronales conduit à l'hyperphosphorylation de tau et à la mort cellulaire (Patrick, Zukerberg et al. 1999). Parmi les phospho-épitopes anormaux de tau, l'épitope TG-3, associé au premier stade de la maladie, est retrouvé uniquement dans les cellules mitotiques et non dans les cellules en quiescence. La kinase Cdc2 génère l'épitope TG-3 dans les cellules mitotiques mais les kinases responsables de cette phosphorylation dans les neurones différenciés ne sont pas encore identifiées. Un des candidats possibles serait la kinase CDK5, qui phosphoryle le site Thr231 *in vitro* (Hamdane, Sambo et al. 2003).

***Profil de phosphorylation pathologique de la protéine tau.***

Les phénomènes liés à la phosphorylation de la protéine tau sont connus depuis de nombreuses années et bien décrits dans la littérature, notamment les différents sites de phosphorylation et les kinases/phosphatases associées, ainsi que le rôle de la phosphorylation dans la régulation de la liaison de tau aux microtubules au niveau physiologique d'une part, et d'autre part au niveau pathologique, dans le processus d'agrégation de la protéine tau associé aux tauopathies (Buee, Bussiere et al. 2000). Concernant le lien entre hyperphosphorylation de la protéine tau et agrégation sous forme de PHF, il faut signaler que la séquence d'événements n'est pas encore élucidée, c'est-à-dire que l'on n'a pas encore déterminé si l'hyperphosphorylation était la cause ou la conséquence de l'agrégation. En effet, l'hyperphosphorylation pourrait être due à une réactivation du cycle cellulaire dans des neurones matures, en accord avec un profil de phosphorylation des PHF semblable aux phosphorylations mitotiques, et serait de fait à l'origine de l'agrégation de la protéine tau et de la neurodégénérescence observée chez les patients atteints de la maladie d'Alzheimer. Ou bien, l'hyperphosphorylation serait consécutive à l'agrégation dans la mesure où certains sites potentiels de phosphorylation seraient moins accessibles aux phosphatases dans les fibres de tau par rapport aux kinases, de taille plus modeste. D'autre part, les relations entre phosphorylation et agrégation ne sont pas bien établies, entre autre parce que la phosphorylation de deux sites, la Ser214 dans la région riche en prolines et la Ser262 dans le domaine à segments répétés, provoque la dissociation de la liaison de tau aux microtubules et, en même temps, protège de l'agrégation (Schneider, Biernat et al. 1999).

Des anticorps avec des spécificités différentes ont été développés et permettent de détecter des épitopes au sein de la protéine tau dans différents états de phosphorylation : des sites non phosphorylés de la forme soluble normale de la protéine tau, des épitopes phosphorylés mitotiques et des épitopes phosphorylés spécifiques des PHF, caractérisés par les sites suivants : Thr212/Ser214 ou épitope AT100 (Hoffmann, Lee et al. 1997), Thr231/Ser235 ou épitope TG-3 (Jicha, Lane et al. 1997) et Ser422 ou épitope AP422. Ces derniers sont dits conformation-dépendants car ils ne reconnaissent pas les phospho-épitopes présents dans la forme phosphorylée normale de la protéine tau mais uniquement ces mêmes phospho-épitopes présents dans les PHF pathologiques (Buee, Bussiere et al. 2000). La conformation de la proline dans ces épitopes est un facteur qui peut être invoqué pour différencier les épitopes mitotiques des épitopes pathologiques. L'implication des prolines dans les motifs Ser/Thr-Pro sujets à la phosphorylation, suggère une possible modification de l'équilibre conformationnel

autour de la proline avec la phosphorylation et/ou l'agrégation mais ces hypothèses ne sont pas encore vérifiées au niveau moléculaire.



**Figure 1.21** : Sites de phosphorylation de la protéine tau normaux, reconnus par les anticorps mitotiques, et pathologiques reconnus par les anticorps dirigés contre les PHF dit « conformation-dépendants ».

### 3.2 – Des connaissances peu avancées sur les interactions entre les protéines Pin1 et tau.

C'est en 1995 qu'a été identifiée la nouvelle famille de peptidyl-prolyl *cis/trans* isomérases, les parvulines et leur homologue humaine Pin1. Mais c'est seulement en 1999 qu'ont été décrites pour la première fois les interactions entre Pin1 et la protéine microtubulaire tau (Lu, Wulf et al. 1999), publication qui fait référence dans ce domaine. De manière très surprenante, malgré 17 sites potentiels d'interaction au sein de la protéine tau correspondant aux motifs phosphoThr/Ser-Pro reconnus par Pin1, un seul et unique site, la Thr231, a été décrit comme un site de haute affinité (40 nM), nécessaire et suffisant pour permettre les interactions de tau avec le domaine de liaison de Pin1.

#### 3.2.1 – Rôle fonctionnel de Pin1 sur l'activité de tau phosphorylée.

Dans le cas des interactions entre Pin1 et la protéine tau, Lu et collaborateurs ont établi au niveau moléculaire, grâce à un criblage par ELISA de plusieurs peptides phosphorylés, que seul le peptide portant la phosphoThr231 était reconnu par Pin1. Pin1 interagit *in vitro* avec la protéine tau phosphorylée par la kinase Cdc2 purifiée et avec un extrait mitotique surexprimant la protéine tau, mais ne se lie pas à la protéine mutante T231A. *In vivo*, Pin1 est co-purifiée avec les PHF issus des neurones de patients atteints de la maladie d'Alzheimer et

co-localisée dans le cytoplasme des neurones avec les PHFs alors que dans les tissus normaux, Pin1 est essentiellement nucléaire (Lu, Wulf et al. 1999). La phosphorylation de tau par Cdc2 inhibe la liaison de tau aux microtubules *in vitro*, alors que Pin1 restaure cette capacité de liaison. Pin1 est détectée dans la fraction de protéine tau liée aux microtubules, mais n'interagit pas directement avec les microtubules. Il a donc été proposé que Pin1 pouvait se lier à tau phosphorylée par Cdc2 et restaurer son activité de liaison aux microtubules. D'autre part, la protéine phosphorylée n'est plus capable de promouvoir la polymérisation des microtubules, mais cette fonction est restaurée par Pin1 de manière dose-dépendante, sans effet sur la protéine tau non phosphorylée. Ni la cyclophiline, ni Pin1 mutée en Y23A ne sont capables de produire les mêmes effets sur la protéine tau phosphorylée. De même, la mutation K63A, concernant un résidu essentiel du *cluster* basique du site actif, ne permet pas de restaurer totalement la capacité de liaison de tau phosphorylée aux microtubules, indiquant que les deux domaines de Pin1 jouent un rôle dans la restauration de l'activité de tau phosphorylée.

Comme la déplétion de Pin1 induit l'arrêt en mitose et l'apoptose, Pin1 pourrait être nécessaire dans le neurone pour prévenir l'activation anormale d'événements mitotiques et contrôler la fonction des phospho-protéines comme tau. Il a été proposé que la réactivation du cycle cellulaire dans les neurones matures pourrait conduire à l'hyperphosphorylation de la protéine tau, dépassant les capacités de Pin1. La protéine tau phosphorylée ne pourrait plus se lier aux microtubules et l'accumulation de tau libre pourrait favoriser la formation des PHF, affectant la fonction neuronale. De plus, Pin1 serait ainsi « séquestrée » dans ces structures insolubles, entraînant une diminution de la concentration de Pin1 soluble dans les neurones et interférant ainsi avec la fonction biologique de Pin1.

Un certain nombre de points abordés dans l'article de Lu et collaborateurs paraissent obscurs, notamment le choix des peptides utilisés pour le criblage des différents épitopes phosphorylés. En effet, tous les peptides testés ne contiennent pas les motifs phospho-Ser/Thr-Pro reconnus par Pin1 et ne sont donc pas des phospho-épitopes pertinents pour déterminer les sites d'interactions avec Pin1. D'autre part, d'après la structure de Pin1 en complexe avec le peptide di-phosphorylé CTD de l'ARN Polymérase II obtenu par cristallographie (Verdecia 2000), les interactions entre le domaine WW de Pin1 et son substrat sont extrêmement limitées et se restreignent au seul motif phosphoSer-Pro. Ces données ont été par ailleurs confirmées par la structure du domaine WW isolé en solution

avec des peptides phosphorylés de tau et Cdc25 par RMN (Wintjens, Wieruszeski et al. 2001). Il est d'autant plus difficile de comprendre la grande spécificité autour du site Thr231 qui régit les interactions entre Pin1 et la protéine tau dans ce contexte. De plus, l'interaction a été « quantifiée » par un  $K_D$  relativement élevé de 40 nM, à partir des tests ELISA sur des peptides phosphorylés, ce qui est contradictoire avec les affinités généralement mesurées pour les substrats peptidiques de Pin1 (Verdecia, Bowman et al. 2000; Wintjens, Wieruszeski et al. 2001; Jacobs, Saxena et al. 2003). D'autre part, on peut douter de la fiabilité des valeurs de  $K_D$  obtenues par une telle méthode. Dans cette étude, Lu et collaborateurs ont montré des interactions entre Pin1 et la protéine tau phosphorylée soit par des extraits mitotiques, soit par la kinase Cdc2/cyclineB purifiée. La Thr231 étant décrite comme un site important de phosphorylation par Cdc2, il est possible que dans ce modèle, la phosphorylation de tau soit orientée majoritairement vers ce site et que d'autres sites ne soient pas phosphorylés dans ces conditions. L'anticorps TG-3, spécifique de la phosphorylation anormale de tau dans la maladie d'Alzheimer en T231/S235, est le seul anticorps utilisé dans les études décrites précédemment. D'autres anticorps comme AT100 et AP422 (Mailliot, Bussiere et al. 1998) n'ont pas été utilisés alors qu'ils sont aussi caractéristiques des phosphorylations anormales de la maladie. Il peut donc exister d'autres sites de reconnaissance entre Pin1 et tau anormalement phosphorylée dans les PHFs qui pourraient ne pas avoir été détectés, en particulier l'épitope AT100, qui n'est probablement pas (ou très peu) phosphorylé par Cdc2 et n'a pas été criblé parmi les phospho-peptides en ELISA.

### ***Stimulation de la déphosphorylation de tau par PP2A (Zhou, Kops et al. 2000).***

La liaison du domaine WW et l'activité catalytique de Pin1 sont toutes deux nécessaires pour retrouver l'activité de tau phosphorylée (Lu, Wulf et al. 1999). Pin1 serait capable de se lier à la protéine tau uniquement au niveau du site Thr231 phosphorylé et d'induire un changement conformationnel qui permettrait de restaurer la capacité de tau à lier les microtubules et à promouvoir leur assemblage (Zhou, Lu et al. 1999). Dans une autre étude, Lu et collaborateurs ont montré que Pin1 stimulait la déphosphorylation de tau par PP2A (Zhou, Kops et al. 2000), une phosphatase majeure dans la déphosphorylation de tau qui est également associée aux microtubules (Wang, Gong et al. 1995; Sontag, Nunbhakdi-Craig et al. 1999; Gong, Lidsky et al. 2000). Ainsi Pin1 peut restaurer l'activité de tau phosphorylée en favorisant sa déphosphorylation. D'autre part, Pin1 régule également la déphosphorylation d'autres antigènes MPM-2 impliqués dans la maladie d'Alzheimer, en stimulant, en particulier, la déphosphorylation de la protéine CTD de l'ARN Pol II par sa phosphatase

FCP1 (Kops, Zhou et al. 2002). Une fois encore, il est probable que l'activité catalytique de Pin1 sur l'interconversion *cis/trans* ait été proposée de manière abusive pour expliquer les effets biologiques observés (Lu, Liou et al. 2003).

***Modèle de souris Pin1 knockout développant des PHFs (Liou, Sun et al. 2003).***

Il a été proposé que Pin1 pourrait protéger de la dégénérescence neuronale observée dans la maladie d'Alzheimer. Un modèle de souris Pin1<sup>-/-</sup> a été utilisé pour rechercher des preuves génétiques liant Pin1 aux lésions neuropathologiques impliquant la protéine tau phosphorylée. Dans cette étude, Lu et collaborateurs montrent une corrélation inverse entre l'expression de Pin1 et la dégénérescence neurofibrillaire dans la maladie d'Alzheimer (Liou, Sun et al. 2003). En effet, chez la souris, la délétion de Pin1 n'est pas létale, mais le phénotype associé au *knockout* de Pin1 ressemble au phénotype des souris déficientes en cycline D1, avec des anomalies diverses liées à des problèmes de prolifération cellulaire, dont une diminution du poids corporel et des atrophies testiculaires et rétiniennes (Liou, Ryo et al. 2002). Ici, il est montré que la délétion de Pin1 provoque également des phénotypes neuropathologiques liés à l'âge, comme des pertes motrices et comportementales, ainsi que l'hyperphosphorylation de la protéine tau, la formation de filaments de tau et la dégénérescence neuronale. Cependant les souris développent des fibres sans présenter de dépôts amyloïdes, ce qui est quelque peu contradictoire avec les mécanismes proposés liant les plaques séniles et les fibres de tau, seules ces dernières étant corrélées avec les pertes cognitives dans la maladie d'Alzheimer. Il a été proposé que Pin1 pourrait être le lien manquant entre les dépôts amyloïdes et les NFT (Pilcher, H. R., Nature 424, 2003, commentaires).

**3.2.2 – Rôle moléculaire de Pin1 vis-à-vis de la protéine tau phosphorylée ?**

Pin1 restaurerait l'activité biologique de tau hyperphosphorylée soit par interaction directe, soit indirectement, en favorisant la déphosphorylation des PHF par la phosphatase PP2A. Les seules données structurales concernent la liaison du domaine WW à un peptide de tau phosphorylée en Thr231 (Wintjens, Wieruszeski et al. 2001) et indique une affinité relativement faible de 230 µM. La faible proportion de formes *cis* (3%) n'est pas le facteur limitant la liaison du domaine WW, qui lie spécifiquement les liaisons en *trans* (Wintjens, Wieruszeski et al. 2001). Cependant, cette faible proportion de forme *cis* n'est pas non plus limitante pour permettre l'action de PP2A, également *trans*-spécifique. Quel est dans ce contexte le rôle de Pin1 en tant que PPIase dans la stimulation de l'activité de déphosphorylation de PP2A, puisque l'épitope phosphoThr231-Pro232 est majoritairement en

conformation *trans* ? Deux scénarios sont envisageables : soit Pin1 se lie à une proline en *trans* par son domaine WW, et le domaine catalytique isomérisé une autre proline, soit le domaine catalytique agit d'abord, et le domaine WW stabilise la conformation *trans* jusqu'à l'arrivée d'une phosphatase PP2A. On peut envisager une autre hypothèse qui expliquerait le rôle fonctionnel de Pin1 sur la restauration de l'activité de tau phosphorylée : la reconnaissance d'un site autre que la Thr231 pour lequel il existerait une conformation en *cis* plus importante. On peut également considérer que les modèles peptidiques permettant une caractérisation aisée des populations *cis* et *trans* ne sont pas représentatifs des conformations qui s'établissent au sein de la protéine entière, même si celle-ci est caractérisée par un haut degré de flexibilité, supposant une grande similarité avec les peptides non structurés.

### **3.2.3 – Pin1, une cible thérapeutique controversée dans le traitement de la maladie d'Alzheimer.**

Si l'implication de Pin1 dans divers mécanismes de cancérisation semble claire, son rôle dans les mécanismes de neurodégénération observée dans la maladie d'Alzheimer reste incertain et soumis à controverse. De nombreuses preuves concernant une réactivation anormale du cycle cellulaire, et en particulier, des événements mitotiques, pourrait contribuer à l'hyperphosphorylation de la protéine tau et jouer un rôle important dans le développement de la maladie d'Alzheimer (Nagy, Esiri et al. 1998; Raina, Monteiro et al. 1999; Vincent 2000). Cependant, la séquence d'événements qui se déclenchent dans les neurones post-mitotiques, contribuant à la neurodégénérescence, reste à déterminer. La protéine Pin1 a été impliquée dans ce processus par sa capacité de liaison à des nombreuses protéines mitotiques phosphorylées, au niveau d'épitopes de l'anticorps MPM-2, par son activité sur la déphosphorylation et la fonction de ces protéines jouant un rôle dans la régulation de la mitose et dans les processus conduisant à la maladie d'Alzheimer. Par exemple, Pin1 est capable de restaurer l'activité de tau phosphorylée, directement par liaison aux épitopes Ser/Thr-Pro phosphorylés et restauration de la conformation de la protéine, ou indirectement par activation de la déphosphorylation de tau par la phosphatase PP2A (Lu, Wulf et al. 1999; Zhou, Kops et al. 2000). D'autre part, dans un modèle de souris Pin1<sup>-/-</sup>, il a été montré que l'absence de Pin1 était corrélée avec le développement de certaines caractéristiques moléculaires de la maladie d'Alzheimer, en l'occurrence la formation des NFT (Liou, Sun et al. 2003). Mais l'apparition des NFT sans la présence des plaques séniles est contradictoire avec la théorie émergente selon laquelle les dépôts amyloïdes seraient les précurseurs des

NFT. Dans cette logique, une déplétion en Pin1 déclencherait différents mécanismes menant à la neurodégénération et nécessiterait des activateurs de Pin1 pour y remédier.

Une autre théorie a été proposée pour expliquer le rôle de Pin1 dans la neurodégénération observée dans la maladie d'Alzheimer, expliquée par la réactivation du cycle cellulaire dans les neurones matures conduisant à la phosphorylation de tau avec l'implication de Pin1 en tant que régulateur mitotique dans ce phénomène (Hamdane, Delobel et al. 2003). Pin1 pourrait avoir un rôle dans le développement de la maladie, à la fois par la régulation de facteurs de transcription et par son rôle de régulation post-traductionnelle. Pin1 pourrait intervenir dans le mécanisme même de réactivation anormale du cycle cellulaire puisqu'il a été montré par Buée et collaborateurs que Pin1 augmente le niveau de cycline D1 dans des neuroblastomes, ce qui pourrait faciliter la transition de la phase G<sub>0</sub> à la phase G<sub>1</sub>, la dédifférentiation neuronale conduisant alors à l'apoptose (Hamdane, Smet et al. 2002; Hamdane, Delobel et al. 2003). D'autre part, l'hyperphosphorylation de tau et d'autres antigènes MPM-2 pourrait dépasser la capacité de Pin1 et la maintenir dans le cytoplasme (Lu, Wulf et al. 1999), diminuant la quantité de Pin1 active dans le noyau, provoquant ainsi l'entrée en mitose et l'apoptose (Lu, Hanes et al. 1996). La perturbation de l'équilibre entre Pin1 et ses substrats phosphorylés conduirait à une boucle d'amplification du phénomène d'hyperphosphorylation. L'insuffisance de Pin1 conduirait à un défaut de déphosphorylation par PP2A et augmenterait le phénomène d'hyperphosphorylation. Le phénomène conduirait (mais ça reste très hypothétique) à la formation de PHF. En parallèle, l'insuffisance de Pin1 conduirait à l'activation supplémentaire de Cdc2 en jouant sur l'activité de Cdc25 dont Pin1 stimule la déphosphorylation et donc l'inhibition. Dans ce cas précis, l'inhibition de Pin1 serait la voie thérapeutique dans le traitement de la maladie (Hamdane, Smet et al. 2002; Hamdane, Delobel et al. 2003).

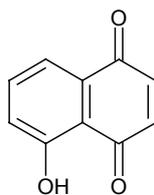
En résumé, le rôle de Pin1 comme cible thérapeutique dans la maladie d'Alzheimer est très controversé à l'heure actuelle, et deux théories s'opposent sur ce sujet. Le fait que la neurodégénérescence observée dans la maladie d'Alzheimer pourrait être liée à une réactivation du cycle cellulaire concorderait avec un possible rôle fonctionnel de Pin1 comme régulateur mitotique. La recherche de molécules pouvant interférer positivement ou négativement sur la fonction de Pin1 permettrait de comprendre le rôle de Pin1 dans la maladie d'Alzheimer, en suivant par exemple l'effet sur la phosphorylation de tau en Thr231 (Hampel, Buerger et al. 2001; Hampel, Buerger et al. 2004).

Nous avons présenté un ensemble de données concernant à la fois la fonction biologique de Pin1, impliquant la reconnaissance des divers substrats et les conséquences des interactions Pin1-substrats au niveau cellulaire, ainsi que des données moléculaires, beaucoup moins bien définies, ce qui explique peut-être que le mode de fonctionnement de Pin1 est encore, à l'heure actuelle, mal compris. C'est dans cette optique que s'insère notre travail : apporter des éléments structuraux, à l'échelle atomique, concernant le mode de reconnaissance des phospho-protéines par Pin1, et la protéine tau en particulier, et l'impact d'une éventuelle catalyse de l'isomérisation *cis/trans* des motifs phosphoSer/Thr-Pro dans des processus biologiques complexes, peut-être comme second élément régulateur de fonctions moléculaires, consécutif à la phosphorylation. Il s'agit là d'un premier aspect du travail qui sera présenté dans ce manuscrit. L'autre volet concerne la recherche d'inhibiteurs de Pin1 avec la synthèse de composés ciblant le domaine de liaison et le criblage de ces ligands potentiels par RMN. Les chapitres suivants vont présenter un bilan de la recherche sur les inhibiteurs de Pin1 et l'emploi de la RMN comme stratégie de criblage.

## IV – Recherche d'inhibiteurs de la fonction de Pin1

### 4.1 – Rappels sur les inhibiteurs des parvulines.

Contrairement aux cyclophilines et aux FKBP, peu d'inhibiteurs de parvulines sont connus. Récemment, un composé naturel extrait des coquilles de noix et feuilles de noyer, la juglone (5-hydroxy-1,4-naphtoquinone), a été décrit comme inhibiteur irréversible, spécifique des parvulines, qui agit par liaison covalente d'une molécule de juglone aux fonctions thiol de la protéine cible. Dans le cas de la parvuline d'*E. coli*, la juglone se lie à deux résidus Cys du domaine catalytique (Cys41 et Cys69) et agirait par modification partielle de la structure de celui-ci (Hennig, Christner et al. 1998; Chao, Greenleaf et al. 2001). Le mode d'action de cette molécule, basé sur sa réactivité comme accepteur de Michael vis-à-vis de fonctions thiols au niveau de résidus Cys exposés, induit une faible sélectivité. Cependant, la juglone est un inhibiteur spécifique des parvulines vis-à-vis des cyclophilines et des FKBP, car elle inactive sélectivement diverses parvulines telles que la parvuline chez *E. coli*, Ess1/Ptf1 de la levure et Pin1 humaine en présence d'autres PPIases.

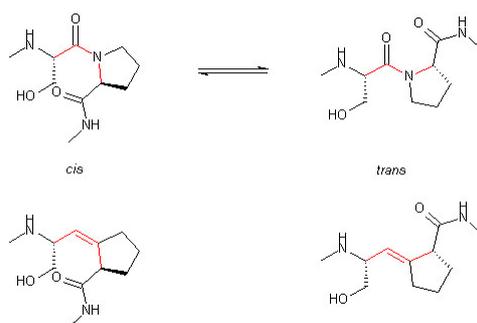


juglone

Des molécules basées sur la structure moléculaire de la juglone sont capables d'inhiber l'activité PPIase de Pin1 avec des  $IC_{50}$  dans la gamme de 1-10  $\mu\text{M}$  *in vitro* vis-à-vis d'un substrat peptidique, dont la séquence AEPF ne contient pas de résidus phosphoSer/Thr, mais pour laquelle une activité *cis/trans* isomérase est mesurable (Uchida, Takamiya et al. 2003). Les molécules, synthétisées initialement comme inhibiteurs potentiels des Topoisomérases I et II, n'ont montré aucune activité sur ces deux cibles mais une analogie structurale avec la juglone, notamment au niveau du double cycle aromatique, leur permet probablement de se lier au site actif de Pin1 via des interactions hydrophobes et de masquer de manière compétitive le site actif aux molécules de substrat. Les tests réalisés sur différentes lignées cellulaires cancéreuses ont permis de mettre en évidence une activité antiproliférative dans une gamme de concentrations de quelques micromolaires. Il a été proposé que l'addition de mimes de phosphate sur ces molécules pourrait augmenter la potentialité de ces composés comme agents anticancéreux.

Des mimes du dipeptide phosphoSer-Pro ont été envisagés notamment par remplacement de la liaison peptidique précédant la proline par une liaison thioxoamide dans des inhibiteurs réversibles de l'activité catalytique de Pin1, possédant des  $IC_{50}$  de quelques micromolaires (Zhang, Fussel et al. 2002).

En raison de la délocalisation électronique sur les trois sites constituant la fonction amide (O, C et N), la liaison peptidique possède un caractère partiel de double liaison aboutissant à un équilibre entre les conformères *cis* et *trans* qui s'apparentent aux configurations Z et E des alcènes (Figure 1.22). Des molécules isostères des dipeptides Ser-*cis*-Pro et Ser-*trans*-Pro basées sur des alcènes en configuration Z et E mimant les conformation *cis* et *trans* respectivement ont été proposés comme inhibiteurs potentiels de la fonction de Pin1 (Wang, Hart et al. 2003).



**Figure 1.22 :** Analogie entre les conformères *cis* et *trans* de la proline et les configurations (Z) et (E) d'une liaison insaturée comme l'indiquent les liaisons en rouge.

De manière générale, l'inhibition vise toujours la fonction enzymatique de Pin1 de manière assez classique, avec des tests de criblage des composés basés sur l'inhibition de l'activité PPIase vis-à-vis de substrats peptidiques non naturels et des méthodes de détection indirecte de la catalyse de l'isomérisation *cis/trans*. Aucune donnée structurale ne montre le site de liaison de l'inhibiteur dans le site actif de Pin1, qui n'est peut-être pas le site principal d'interaction, notamment pour les molécules contenant un phosphate, qui pourraient se loger dans le site de liaison du substrat au sein du domaine WW de Pin1.

#### 4.2 – Inhiber la liaison de Pin1 aux motifs phosphoSer/Thr-Pro.

Le domaine WW a été impliqué à différents niveaux dans la fonction de Pin1 : la liaison des substrats phosphorylés, dans un rôle compétiteur vis-à-vis d'autres partenaires dans le cas de certains substrats de Pin1 et, enfin dans la régulation de Pin1 à ses substrats et sa localisation cellulaire. Des données structurales ont permis d'identifier le dipeptide phosphoSer/Thr-Pro comme ligand minimal avec des affinités de liaison relativement faibles de l'ordre de 100  $\mu$ M. Nous avons envisagé une stratégie chimio-thérapeutique basée sur l'inhibition du domaine WW qui pourrait mimer la forme inactive de Pin1 telle qu'elle est trouvée majoritairement dans les cellules cancéreuses du sein (Wulf, Ryo et al. 2003), et des modifications pharmacophoriques du dipeptide phosphoThr-Pro, avec le criblage des composés synthétisés par RMN.

Le problème posé dans la première partie de cette introduction concernant l'absence d'informations structurales permettant de corréler le rôle d'isomérase joué par Pin1 (qui reste à démontrer) et la fonction biologique de cette dernière dans la machinerie cellulaire, soulève la nécessité d'établir un lien entre le rôle moléculaire de Pin1 sur ses substrats et cette

fonction. Cette issue inévitable ne peut être résolue que par des techniques d'investigations structurales dont la plus utilisée sans doute, est la cristallographie par rayons X. Cependant, dans le contexte de l'isomérisation *cis/trans* des liaisons précédant la proline, l'aspect dynamique de la catalyse de cette interconversion et l'implication de régions protéiques hautement flexibles peut donner l'avantage à la RMN. Cette technique, que nous maîtrisons bien au laboratoire, a été également mise en œuvre et optimisée pour le criblage de nos composés. Dans le paragraphe suivant, nous allons présenter l'état de la question concernant la RMN comme outil de criblage.

### **V – La RMN dans le processus de découverte de *hits/leads* pharmacologiques, une technique HTS parmi d'autres ?**

#### **5.1 – Approche générale du criblage à haut débit par RMN.**

Les avancées récentes dans le séquençage du génome humain et les méthodes développées en bio-informatique ont alimenté l'arsenal de cibles thérapeutiques pour le traitement des maladies infectieuses, dégénératives et chroniques... La caractérisation de telles cibles par les techniques biochimiques et physico-chimiques classiques est bien établie, avec l'implication déjà à ce stade de techniques à haut débit. Le développement de méthodes pour l'identification de petites molécules capables de se lier à une cible donnée doit encore être approfondi, de même que les techniques de caractérisation des interactions. Des méthodes comme le *docking in silico* et le criblage virtuel répondent en partie à cette demande, mais il faut encore augmenter l'effort en matière de caractérisation moléculaire des interactions entre les ligands et leurs cibles.

La RMN peut jouer un rôle clé dans ce domaine par sa capacité à analyser une large gamme de composés, très diversifiés d'un point de vue structural ainsi qu'un grand nombre de cibles, avec des fonction biologiques très diverses. De plus, l'intérêt principal de la RMN réside essentiellement dans la capacité de détecter une large gamme d'affinité entre les composés et leur cible (mM-nM), ce qui en fait une technique de choix tout particulièrement dans les premières étapes de la recherche d'un composé actif. Avec les efforts fournis en matière de décryptage du génome, on connaît actuellement une situation inédite avec les récepteurs orphelins pour lesquels aucun ligand n'est connu. La RMN peut dans ce contexte apporter un grand nombre de pistes en détectant des ligands de faibles affinités qui fournissent un point de

départ structural pour l'élaboration de ligands de plus haute affinité. Le point fort de la RMN par rapport aux techniques de criblage classiques, est l'apport d'informations structurales sur la cible ainsi que sur le mode de liaison du ligand en termes quantitatif et dynamique, qui permet de décortiquer finement les interactions entre le composé et sa cible en vue d'une éventuelle évolution structurale du ligand. Des banques de 200 000 composés peuvent être criblées dans un temps tout à fait raisonnable mais ce protocole reste encore limité à l'heure actuelle à l'industrie pharmaceutique.

Un des avantages majeurs de la RMN en terme de criblage est son caractère universel. La connaissance de la fonction biologique de la protéine-cible n'est pas nécessaire pour mettre en œuvre un test de criblage par RMN, qui peut être appliqué à une large gamme de protéines couvrant des fonctions très diverses, ce qui n'est pas le cas des techniques HTS basées sur des tests fonctionnels. La taille de la protéine n'est pas vraiment une limite contrairement à ce qui est couramment admis, tout dépend de l'information recherchée et de la technique RMN mise en œuvre. Nous en détaillerons quelques unes dans le paragraphe suivant. De plus, la connaissance d'un ligand de départ n'est pas un pré-requis comme cela peut être le cas en fluorescence. La mise en œuvre d'un test de criblage par compétition avec un ligand fluorescent nécessite de connaître un tel ligand pour une cible donnée, ainsi que son site de liaison. Lors du criblage, des composés qui se lieraient au niveau d'un autre site de la protéine-cible ne seraient pas détectés par absence de compétition avec le ligand fluorescent et alimenteraient le *pool* des faux négatifs. La RMN peut ainsi être utilisée dans la validation d'autres techniques de criblage HTS, en définissant le site d'interaction d'un ligand fluorescent par exemple.

### **5.2 – La RMN dans les différentes étapes de la recherche d'un médicament.**

La RMN peut être utilisée à diverses étapes du processus de conception d'un médicament : la validation d'une cible pharmacologique, la recherche et la validation de *hits*, la recherche de *leads* et leur optimisation, dans un approche basée sur la conception rationnelle guidée par les données structurales (Homans 2004).

Dans la validation de cibles, la RMN peut compléter l'approche par cristallographie dans la résolution de la structure tridimensionnelle, en fournissant des données dynamiques en solution, plus proches du contexte biologique de l'interaction ligand/récepteur que l'état cristallin. La résolution de structure peut être envisagée par RMN mais reste limitée à des protéines monomériques de 40 kDa avec les équipements les plus performants. Un point

important est la détection de protéines sans structures tridimensionnelles stables possédant un faible degré d'organisation. Ces protéines ne sont pas des cibles idéales, même si l'on envisage qu'elles se structurent au contact de leur ligands naturels *in vivo*. La protéine tau, que nous avons décrite dans les paragraphes précédents, sera étudiée dans un contexte structural dans notre approche des interactions avec Pin1. Nous aurons l'occasion d'aborder les problèmes de caractérisation de telles protéines par RMN, en particulier pour l'attribution séquentielle.

La RMN intervient dans la recherche de *hits*, c'est-à-dire de composés de faibles affinités (10  $\mu$ M-1 mM) dont la structure peut cependant être utilisée comme point de départ et dérivatisée pour la découverte d'un chef de file. Elle peut être utilisée dans la validation de *hits* issus d'autres techniques HTS. Les informations obtenues à ce niveau concernent : la validation de l'interaction, la définition du site de liaison, du mode de liaison du ligand, c'est-à-dire l'identification de l'« épitope » de liaison. Des expériences de compétition peuvent être mises en œuvre avec des ligands connus lorsqu'ils existent. Elle apporte également un grand nombre d'informations dans le processus dit *hit to lead*, pour la découverte d'un chef de file. L'apport des données structurales peut, à ce stade, donner un avantage considérable, en particulier dans l'investigation de sites de liaison additionnels qui devraient améliorer la liaison des ligands.

Dans ce contexte de criblage par RMN, trois types d'expériences peuvent être envisagées pour cribler et quantifier les interactions entre un composé et une cible donnée. Ces expériences se définissent par :

- le suivi de l'évolution du signal de la cible
- le suivi de l'évolution du signal du ligand
- le suivi des interactions entre le ligand et la cible

Les informations collectées sont aussi bien qualitatives (le composé se lie-t-il ou pas à la cible ? Quel est le mode de liaison ?) que quantitatives (quelle affinité ? Etablissement des relations structure/activité). Nous allons présenter succinctement dans les paragraphes suivants les différentes techniques envisageables (voir revue (Meyer and Peters 2003)).

### **5.2.1 – Perturbations du signal de la cible.**

Le suivi d'une interaction éventuelle avec un ligand peut s'effectuer en observant les résonances d'une protéine-cible. Les variations des résonances de la protéine uniformément

ou sélectivement enrichie en  $^{15}\text{N}$  sont détectées en fonction de la liaison du ligand au niveau de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Cette technique met à profit l'extrême sensibilité du déplacement chimique par rapport à l'environnement chimique de chacun des résidus de la protéine. Cette méthode permet de détecter des affinités très faibles de liaison et de cribler jusqu'à 1000 composés par jour par machine pour un système bien optimisé. Par rapport aux autres techniques HTS, la perturbation des déplacements chimiques fournit une identification du site de liaison au niveau de la cible de manière non ambiguë. Quand une série de composés est testée, l'information résultante permet de construire une image fidèle du complexe ligand-récepteur. Actuellement, les avancées technologiques que connaît la RMN depuis ces dernières années, avec l'arrivée des aimants à haut champ (800-900 MHz) et des nouvelles séquences impulsionsnelles (TROSY), ont repoussé les limites concernant la taille des protéines étudiées à environ 100 kDa. Cette approche consistant à observer les variations de signal de la protéine-cible sera largement illustrée dans notre travail concernant l'interaction de Pin1 avec des substrats peptidiques ou de petites molécules.

### 5.2.2 – Perturbations du signal du ligand.

L'observation des signaux du ligand est préférentiellement utilisée quand la quantité de protéine isotopiquement enrichie ou la taille de la protéine est un facteur limitant. Dans ce cas, de faibles quantités de protéines non marquées sont requises. La limite de ces techniques se situe au niveau de la dynamique du système. En réalité, l'observation est réalisée sur le *pool* de ligands libres en solution mais elle est capable de retracer l'« histoire » du ligand au cours du temps à condition que l'échange procède sur une échelle de temps comparable à celle de l'expérience. Cela concerne en l'occurrence des échanges relativement rapides. Si un échange entre l'état libre et l'état lié du ligand se produit dans un intervalle de temps inférieur à 100 ms-1 s, le « passé » du ligand est transféré à la fraction de ligand libre en solution. Le principe de ce « transfert » repose sur le fait que lorsque le ligand est lié à la protéine, il perd ses propriétés de petite molécule et adopte, pour un temps, le comportement de la protéine (relaxation plus rapide, diffusion plus lente, effets NOE plus importants et négatifs).

Ainsi, l'élargissement des signaux du ligand peut être le témoin d'une interaction avec un objet macromoléculaire, dont la mobilité et la flexibilité sont bien moins importantes et entraînent une augmentation de la vitesse de relaxation, inversement proportionnelle à la largeur de raie. Cette technique autorise le criblage de composés en mélange et permet de distinguer dans ce mélange les ligands des non-ligands. La limite, dans ce cas, est imposée par le recouvrement des signaux des différents composés du mélange, qui doivent être choisis

avec attention pour limiter ces problèmes de recouvrement de spectres. Le débit de ce type d'expérience est de 50 à 500 composés/jour/machine.

L'évolution technique avec les gradients de champ a mené vers les expériences de diffusion, plus sophistiquées mais basées sur le même principe que l'expérience décrite précédemment, c'est-à-dire, une variation de mobilité pour le ligand quand il est lié à la cible. Cette expérience est d'autant plus efficace que la cible est de taille importante, ce qui mène, en général, vers une différence de mobilité d'autant plus importante pour le ligand en présence et en l'absence de la protéine-cible. Le débit de ce type de méthode est relativement faible avec 10 à 100 composés testés par jour par machine.

Une autre technique, basée sur le transfert de saturation de la protéine vers le ligand, communément appelée STD (*Saturation Transfer Difference*), permet non seulement d'identifier un ligand d'un non ligand, éventuellement sur des mélange de composés, mais permet également de déterminer le site de liaison au niveau du ligand, appelé aussi « épitope ». Cette expérience consiste en l'irradiation du signal de la protéine dans une zone où aucun signal du ligand n'est observé (souvent, à  $\delta < 0$  ppm). Si le composé échange entre un état libre et un état lié, le transfert de saturation a lieu et, par différence avec la saturation de la même zone en l'absence de protéine, on peut définir le site d'interaction au niveau du ligand. L'avantage de cette technique concerne l'étendue de la gamme des affinités détectables et le débit d'environ 1000 composés/jour/machine.

Ces méthodes présentent toutes l'inconvénient de ne pas fournir d'information structurale sur le(s) site(s) de liaison au niveau de la cible.

### 5.2.3 – Suivi des interactions ligand-cible.

Le transfert de NOE permet d'étudier des interactions de faibles affinités (mM- $\mu$ M). L'avantage de cette technique de bas débits réside en l'évaluation de la structure du ligand ancré dans le site de liaison de la cible. Les expériences en milieux orientés apportent également des informations structurales sur l'interaction du ligand avec sa cible, basées sur les couplages dipolaires  $^1\text{H}$ - $^1\text{H}$  et  $^1\text{H}$ - $^{13}\text{C}$ , avec des contraintes d'orientation imposées par l'orientation de la cible, et permettent d'établir la conformation du ligand dans la poche de liaison. La résolution structurale du complexe avec les contraintes orientationnelles est plus rapide qu'avec les contraintes NOE, aussi en terme de temps machine pour la collecte des informations. Les expériences NOE avec filtrage isotopique permettent la détermination des contacts au niveau de l'interface du complexe et aboutissent à un *docking* précis du ligand dans le site de liaison.

### **5.3 – Applications. Le criblage par RMN dans la recherche de composés actifs.**

La RMN est devenue un outil majeur intervenant à différents stades dans le processus de découverte de médicaments. L'attrait de la RMN réside en la nature des informations fournies par rapport aux techniques HTS classiques, notamment en termes structuraux. Mais la RMN ne serait jamais devenue une technique de criblage sans les avancées technologiques réalisées ces dernières années, même si le débit de composés testés reste plus faible qu'avec d'autres techniques HTS. Au début des années 1990, l'effort fourni au niveau des techniques de criblage consistait essentiellement à augmenter le débit des composés testés jusqu'à cribler 100 000 composés par jour pour certaines techniques dite d'ultra-HTS, en réponse à une augmentation importante du nombre de cibles liée au décodage du génome humain et les avancées en génomiques et protéomiques, ainsi qu'au développement de la chimie combinatoire générant une grande quantité de composés de structures diversifiées. Actuellement la tendance est plutôt de privilégier la qualité de l'information plutôt que la quantité (Walters and Namchuk 2003). C'est dans ce contexte que la RMN s'est imposée comme technique de choix pour le criblage de banques de taille moyenne (de l'ordre de 10000 composés) avec la possibilité de tester soit 100 molécules individuelles par jour soit 1000 molécules en mélanges de 10 composés, de manière automatisée avec un passeur d'échantillons. L'inconvénient majeur de la RMN est lié à la faible sensibilité de la technique nécessitant de grandes quantités de protéines (plus de 200 mg). De plus l'apport d'informations structurales n'est applicable qu'aux petites protéines (de masse moléculaire inférieure à 40 kDa). Cependant ces limites évoluent avec les avancées technologiques, en particulier vers l'utilisation de quantités plus faibles de protéines avec l'utilisation de cryosondes (Hajduk, Gerfin et al. 1999).

#### **5.3.1 – L'approche par fragments.**

L'avantage majeur de la RMN en terme de criblage, hormis l'accessibilité aux informations de type relations structure-activité, est de détecter de faibles affinités de liaison (de l'ordre du millimolaire) qui seraient indétectables avec les techniques classiques de HTS. Ce point est particulièrement important lors de la recherche de *hits* pharmacologiques qui sont les composés précurseurs des *leads*. Lors de la découverte d'une cible pharmacologique, les premiers composés criblés ont pour la plupart peu, voire aucune, affinité pour la cible. Le criblage initial est une étape importante car il permet l'identification de points de départ chimiquement intéressants dans la recherche d'un futur médicament et doit par conséquent permettre de détecter des affinités faibles et différencier des ligands de faible affinité des non-

ligands. Il faut également souligner que la RMN permet de détecter des interactions entre un ligand et une protéine sans nécessairement passer par l'attribution de la protéine, au moins dans un premier temps, par simple considération des perturbations de déplacements chimiques de la protéine, ni même identifier la fonction biologique de ladite protéine.

C'est dans ce cadre qu'a été mise au point l'approche par fragments moléculaires de S. Fesik et P. Hajduk, respectivement vice-président et directeur de la plate-forme de criblage par RMN des Laboratoires Abbott (Shuker, Hajduk et al. 1996). Cette approche, également appelée « *SAR by NMR* » permet l'obtention de ligands d'affinité sub-micromolaire à partir d'un criblage par RMN de fragments. Ces fragments sont de très petites molécules structurellement peu évoluées qui présentent en conséquence, généralement, de faibles affinités pour la cible, comprises entre 10  $\mu$ M et 1 mM. Le but ultime étant de combiner ces fragments de faible affinité pour obtenir dans le cas optimal un ligand plus affiné, il faut qu'ils puissent se lier au niveau de sites différents dans la protéine-cible. De plus, ces sous-sites doivent être proches dans la structure de la protéine (Shuker, Hajduk et al. 1996; Hajduk, Meadows et al. 1997). Cette technique met en œuvre les étapes suivantes :

- le criblage des fragments
- l'identification des sites de liaison des fragments qui se lient à la cible et leurs affinités
- la détermination de la structure du complexe formé par deux fragments (ou plus) se liant à deux sous-sites distincts par cristallographie
- la synthèse « combinatoire » des fragments les plus affins pour chaque sous-site avec différents *linker*, faisant éventuellement intervenir la modélisation moléculaire *in silico*
- la détermination des affinités de ces nouveaux composés
- la détermination de leur conformation dans la cible

Cette stratégie permet de réduire à la fois la quantité de composés synthétisés ainsi que le temps nécessaire à l'acquisition des données. En effet, pour une protéine possédant deux sites de liaison distincts potentiels, relativement proches dans la structure globulaire de la protéine-cible, la synthèse et le criblage de  $10^4$  fragments équivalent à une banque virtuelle de  $10^8$  composés par liaison des fragments deux à deux. L'effort à fournir dans ce cas est de synthétiser et de cribler  $10^4$  fragments, de ne retenir et de n'optimiser que les plus efficaces d'entre eux, pas seulement en terme d'affinité mais également en tenant compte des possibilités d'optimisation à partir des données structurales et des données de modélisation

moléculaire. Cette approche nécessite l'implication d'une technique comme la RMN ou la cristallographie car elle requiert des informations structurales à différents stades : l'identification des sous-sites de la protéine, l'identification du sous-site occupé par un ligand donné, la caractérisation des complexes pour l'optimisation des combinaisons de fragments et la structure des ligands finaux dans la protéine.

Cette approche des fragments moléculaires a été modifiée et rendue accessible à d'autres techniques HTS. En effet, les fragments initiaux étant faiblement affins, une stratégie visant à augmenter « artificiellement » leur affinité permet d'utiliser des techniques HTS classiques qui ne permettent pas de détecter de si faibles affinités. Cette méthodologie a été développée récemment par la société Sunesis, sous le nom de *tethering*. Cette méthode consiste à cribler des fragments par liaison et ancrage covalent de ceux-ci à la protéine-cible (Erlanson, Braisted et al. 2000). L'ancrage utilisé est la fonction thiol de la chaîne latérale d'une Cys présente dans la protéine ou obtenue par mutagenèse dirigée, au niveau d'un site localisé à proximité d'un site de liaison au sein de la protéine-cible. Les fragments sont constitués de fonctions sulfures permettant ainsi, dans des conditions réductrices, un échange rapide entre thiol et sulfure consécutif à la liaison du fragment à la protéine. Les fragments peuvent ainsi être testés en mélange et les complexes, très stables dans des conditions réductrices avec des ligands peu affins, peuvent être identifiés par spectrométrie de masse.

### 5.3.2 – Expériences de compétition par RMN.

L'observation des signaux du ligand permet de contourner les problèmes concernant la taille de la protéine ou sa solubilité aux hautes concentrations requises pour l'observation des signaux de la protéine. L'inconvénient principal des techniques basées sur le suivi du signal du ligand est l'incapacité à détecter des ligands de haute affinité, qui apparaissent comme faux-négatifs. En effet, le ligand étant largement majoritaire par rapport à la protéine, les événements qui se produisent au niveau du ligand sont détectés à travers la fraction de ligand libre en solution. Mais le transfert de l'information entre la fraction de ligands liés et les ligands libres s'opère uniquement lorsque l'échange est rapide donc les constantes de dissociation relativement faibles. Lorsque l'échange est lent, peu de ligands sur la totalité des molécules de l'échantillon passent par un état lié à la protéine et ces événements ne se ressentent pas sur la fraction de ligands libres. Ainsi les ligands se liant fortement apparaissent comme faux-négatifs. Des méthodes ont été développées pour pallier ces inconvénients et permettre la détection de ligands hautement affins (avec des  $K_D$  sub-micromolaires).

Le criblage de composés peut s'effectuer en mélange par compétition avec un ligand connu de faible affinité, aussi appelé « espion » ou « sonde » ou encore « *reporter ligand* », d'où le nom de cette méthode appelée « *reporter screening* » (Jahnke, Floersheim et al. 2002). Dans cette technique, ce n'est pas la liaison du composé testé qui est observée mais la liaison du composé « espion ». Si on détecte une liaison entre la protéine et le ligand « espion », c'est qu'aucun des composés du mélange n'est capable de se lier à la cible avec une affinité plus élevée. A l'inverse, si le ligand « espion » est déplacé de la cible, c'est qu'un ou plusieurs composés présentent une affinité supérieure. Pour évaluer un comportement de ce type, on se place dans des conditions où le signal du ligand est élargi en raison de l'interaction avec la protéine. Lorsque le ligand est déplacé de son site de liaison par un composé, il retrouve la largeur de raie qu'il adopte dans l'état libre. Cette méthode se limite à la détection de ligands au niveau d'un seul et unique site, celui ciblé par le ligand « espion », contrairement à la technique évoquée précédemment. De plus, cette méthode permet de détecter des ligands peu solubles pourvu qu'ils aient une haute affinité.

Une autre approche qui permet d'augmenter la sensibilité de la détection, consiste en l'utilisation de marqueurs de spins qui sont introduits sur la protéine à proximité de son site de liaison (Jahnke, Rudisser et al. 2001). Le marqueur de spin est un composé paramagnétique dont la propriété est de provoquer un élargissement des signaux des protons localisés dans un environnement proche (à une distance inférieure à 15-20Å). Cette méthode permet de détecter, au travers de l'élargissement des signaux des composés, des ligands qui s'approchent du marqueur de spin lors de leur liaison au site actif de la protéine. Cette méthode permet de détecter des ligands avec une meilleure sensibilité nécessitant de plus faibles quantités de protéine. Cependant, elle présente les mêmes inconvénients concernant l'absence de détection des ligands de haute affinité.

Cette méthode a été dérivée pour la recherche de ligands ciblant un second site de liaison dans la protéine-cible. En effet, dans l'approche des fragments moléculaires, la technique repose sur l'identification de fragments se liant à deux sous-sites distincts de la protéine mais proches structurellement, pour lesquels les fragments ne doivent pas entrer en compétition. Une fois que le premier site a été identifié ainsi qu'un premier ligand, la recherche d'un second ligand pour un second site nécessite la saturation du premier site avec des concentrations élevées de ligand, basées sur des concentrations de protéine de 50 µM requises pour l'observation des signaux de la protéine et des constantes de dissociation relativement faibles. L'approche de l'observation des signaux du ligand est, à nouveau, une alternative à ces problèmes de concentration et aux questions de solubilité sous-jacentes. Le marquage de spin du premier

ligand permet de détecter l'élargissement des signaux du second ligand lorsque celui-ci se lie au second site, étant donnée la proximité spatiale des deux sites, nécessaire pour la liaison ultérieure des deux fragments (Jahnke, Florsheimer et al. 2003).

La RMN joue aujourd'hui un rôle significatif, si l'on en croit le nombre de revues sur le sujet parues ces cinq dernières années, dans le processus de recherche de candidats médicaments. Elle se présente comme une technique sensible et robuste pour la détection de ligands de faible affinité dans la découverte de *hits* pour une cible pharmacologique donnée. L'arsenal de techniques RMN permet de fournir en général la solution qui convient en fonction du problème posé. Dans le cas où la quantité de protéine enrichie n'est pas un facteur limitant, ni la taille, ni la solubilité de la protéine, la méthode de choix est d'étudier les signaux de la protéine au contact d'un ligand individuel ou de plusieurs ligands en mélange. Cette approche fournit en plus du criblage de base, une information sur le mode de liaison du ligand. Si ces conditions ne sont pas rassemblées, il est possible d'utiliser des techniques alternatives consistant à observer les signaux du ligand, moins versatiles en termes structuraux. Le criblage par RMN est universel et n'impose ni contraintes sur la fonction de la protéine, ni sur la structure et les propriétés physico-chimiques des composés. La RMN peut ainsi faire gagner un temps appréciable au début du programme de criblage par la rapidité de la mise au point du test, si les données structurales sont planifiées tôt dans la stratégie adoptée.

## **CHAPITRE II**

**Identification d'épitopes Ser/Thr-Pro phosphorylés de la protéine microtubulaire tau reconnus par la peptidyl-prolyl *cis/trans* isomérase Pin1.**

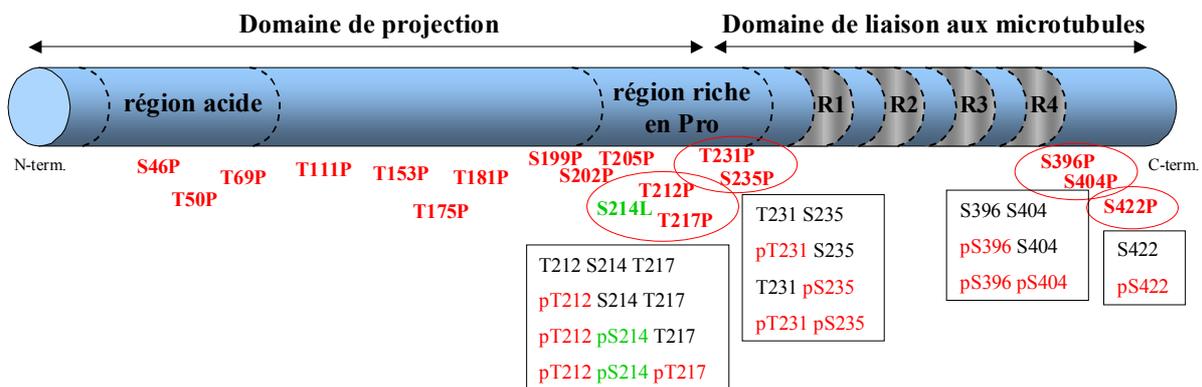


## I – Criblage de phospho-épitopes de type Ser/Thr-Pro de la protéine tau reconnus par la peptidyl-prolyl *cis/trans* isomérase Pin1.

### 1.1 – L'outil peptidique pour le criblage d'épitopes phosphorylés de la protéine tau médiateurs des interactions avec la peptidyl-prolyl *cis/trans* isomérase Pin1.

#### 1.1.1 – Choix de phospho-épitopes de tau pertinents inclus dans des peptides.

Le choix des phospho-épitopes pour le criblage s'est porté sur des sites de phosphorylation qui sont phosphorylés uniquement dans la forme « normale » soluble de la protéine tau (S396-S404), des sites qui sont phosphorylés dans les formes normales et pathologiques (T231/S235, S422) et un site spécifique des PHF (Thr212/S214). En effet, la Thr212 est un des rares sites de phosphorylation faisant partie d'un motif Thr-Pro phosphorylé uniquement dans les PHF. Ces épitopes sont localisés dans des régions protéiques impliquées dans la liaison ou dans la régulation de la liaison aux microtubules (Figure 2.1). Nous nous sommes intéressés tout particulièrement à la région riche en prolines qui concentre de nombreux motifs Ser/Thr-Pro reconnus par Pin1, notamment les épitopes TG3 et AT100 reconnaissant respectivement les sites Thr231/Ser235 et Thr212/Ser214 di-phosphorylés.



**Figure 2.1** : Représentation schématique de la protéine tau avec ses différents domaines incluant les 17 motifs Ser/Thr-Pro (annotés en rouge). Les sites encadrés font partie du criblage de phospho-épitopes reconnus par Pin1 inclus dans des séquences peptidiques regroupant un ou plusieurs sites dans différents états de phosphorylation (encadrés). La Ser214 est annotée en vert car il ne s'agit pas d'un site Ser-Pro mais le couple Thr212/Ser214 est un épitope spécifique des PHF détectés par l'anticorps AT100.

<i>Nom du peptide</i>	<i>Séquence</i>
Peptide T231S235	K K V A V V R T P P K S P S S A K
Peptide pT231	K K V A V V R pT P P K S P S S A K
Peptide pS235	K K V A V V R T P P K pS P S S A K
Peptide pT231pS235	K K V A V V R pT P P K pS P S S A K
Peptide T212S214T217	S R S R T P S L P T P P T R
Peptide pT212	S R S R pT P S L P T P P T R
Peptide pT212pS214	S R S R pT P pS L P T P P T R
Peptide pT212pS214pT217	S R S R pT P pS L P pT P P T R
Peptide S396	V Y K S P V V S G D T S P R H L
Peptide pS396	V Y K pS P V V S G D T S P R H L
Peptide pS396 pS404	V Y K pS P V V S G D T pS P R H L
Peptide S422	S I D M V D S P Q L A T L
Peptide pS422	S I D M V D pS P Q L A T L

**Tableau 2.1** : Séquences primaires des peptides de la protéine tau synthétisés pour le criblage des sites potentiels d'interaction avec Pin1.

La Thr212 est par ailleurs un site particulièrement intéressant car un même nombre de résidus la sépare de la Thr231 par rapport à la paire Thr47/Thr68 de la partie régulatrice de la phosphatase humaine Cdc25c, également reconnue par Pin1. En effet, dans les deux protéines, les régions protéiques contenant ces paires de résidus phosphorylés ne sont pas repliées en une structure globulaire, ce qui signifie très probablement que ces résidus sont séparés par la même distance spatiale. Il est alors envisageable que cette distance soit un facteur structural significatif dans les interactions entre Pin1 et ses partenaires, hypothèse qui peut être appuyée, d'un point de vue structural, par la grande flexibilité des deux domaines de Pin1 en solution (Bayer, Goettsch et al. 2003; Jacobs, Saxena et al. 2003) et qui suggère une possible interaction de chacun des sites avec les deux domaines de Pin1 simultanément.

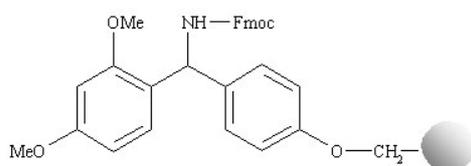
Le criblage des phospho-épitopes de la protéine tau a été effectué par l'intermédiaire de substrats peptidiques courts ne contenant qu'un ou deux sites potentiels d'interaction avec Pin1 (Figure 2.1, Tableau 2.1).

### **1.1.2 – Synthèse de peptides phosphorylés par synthèse peptidique en phase solide.**

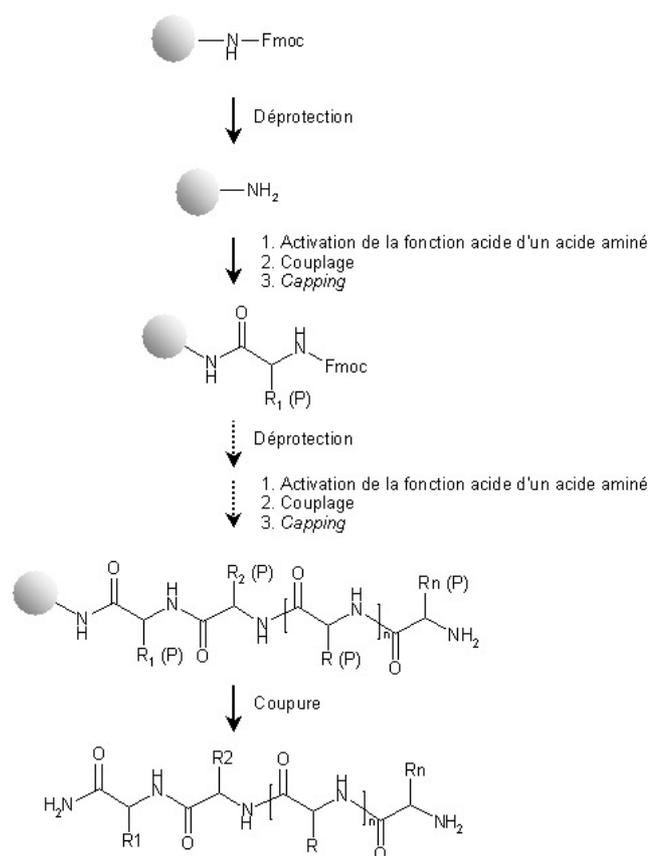
La synthèse mise en œuvre pour l'obtention des peptides phosphorylés est basée sur la technique décrite par Merrifield en 1963 de synthèse de peptides sur support solide (Merrifield 1963).

Les différents peptides phosphorylés et leurs homologues non phosphorylés ont été synthétisés selon une stratégie Fmoc classique. Cette stratégie est parfaitement adaptée à la synthèse des phosphopeptides à la fois par la disponibilité des acides aminés phosphorylés protégés (Novabiochem) utilisés lors de la synthèse et par la compatibilité des réactifs utilisés pour les réactions de couplage, de déprotection du groupement Fmoc, de coupure de la peptidyl-résine et de déprotection des chaînes latérales, en particulier avec le groupement phosphate que l'on choisit mono-protégé (commercial, Novabiochem) afin de réduire les réactions secondaires comme la réaction de  $\beta$ -élimination en milieu basique. De plus, la synthèse en stratégie Fmoc est automatisable, ce qui présente, dans le cas de synthèses multiples ou lors de la synthèse de peptides comportant un grand nombre de résidus, un grand intérêt en terme de temps de réalisation de la synthèse et de reproductibilité.

La synthèse des peptides phosphorylés a été effectuée sur synthétiseur automatique à flux continu de marque Pioneer (Applied Biosystems) sur des échelles allant 0.05 à 0.2 mmol avec l'utilisation de résines de type rink amide (Figure 2.2). Le schéma de synthèse est reproduit comme suit (Figure 2.3) : (i) déprotection de la fonction amine protégée par un groupement Fmoc par réaction avec une solution de pipéridine à 20% dans le DMF ; (ii) couplage d'un acide aminé activé par un mélange équimolaire HOBt/HBTU et 3 éq. de DIPEA dans le DMF ; (iii) réaction dite de *capping*, c'est-à-dire réaction d'acétylation par l'anhydride acétique en présence de DIPEA des fonctions amines non couplées. Le peptide est ensuite coupé de la résine à la fin de la synthèse par incubation de la résine dans une solution concentrée de TFA (concentration supérieure à 94% en volume) avec adjonction de *scavengers* tels que le triisopropylsilane ou l'éthanedithiol. Le mélange brut de peptides est isolé par précipitation dans un mélange éther : heptane. La purification et le contrôle de la pureté des fractions sont réalisés par HPLC en phase inverse. La colonne est équilibrée dans une solution aqueuse contenant 0.05% de TFA (v/v). Un gradient d'acétonitrile est mis en œuvre pour l'élution des peptides.



**Figure 2.2** : Structure de la résine Rink amide (Novabiochem) utilisée comme support solide dans la synthèse des peptides phosphorylés. Le *linker* de type Rink amide est ici greffé sur une résine polystyrène renforcée par 1% de divinylbenzène.



**Figure 2.3 :** Schéma de synthèse supportée employée pour la synthèse des phospho-peptides selon une stratégie Fmoc. La protection temporaire Fmoc des fonctions amine est clivée par incubation de la résine avec une solution de pipéridine 20% dans le DMF. Le couplage des acides aminés est réalisé par activation de la fonction amine par un mélange HOBt/HBTU en milieu basique dans le DMF et suivi d'une étape de *capping* par réaction avec l'anhydride acétique en milieu basique dans le DMF. Les chaînes latérales des acides aminés sont représentées par les groupements R1, R2, ... Rn protégés, quand cela est nécessaire, par des protections symbolisées par la notation (P). La réaction de coupure réalisée dans le TFA concentré en fin de synthèse permet de couper le peptide de la résine et de déprotéger simultanément les fonctions des chaînes latérales.

***Avantage de l'utilisation de synthons phosphorylés par rapport à la phosphorylation globale.***

L'avantage de travailler avec des substrats peptidiques pour l'étude des sites d'interaction de Pin1 réside en la sélectivité de la phosphorylation et en un taux de phosphorylation par site de 100%, accessible grâce à l'utilisation de synthons de type N- $\alpha$ -Fmoc-(L)-Thr/Ser-(PO(OBn)OH)-OH commerciaux (Novabiochem). Il est également possible de phosphoryler les résidus Ser, Thr ou Tyr présents dans les protéines ou les peptides selon une technique dite de « phosphorylation globale », soit par voie chimique soit par l'action de kinases plus ou moins sélectives. Par voie chimique, la sélectivité peut être obtenue dans un peptide par le

choix de protections orthogonales sur les différents résidus qui doivent subir une réaction de phosphorylation, de manière à conserver les protections sur les résidus qui ne doivent pas être phosphorylés après coupure du peptide de son support solide en fin de synthèse. Cette technique présente l'avantage de fournir en une seule synthèse le peptide phosphorylé et son homologue non phosphorylé. La phosphorylation peut également être envisagée sur le peptide fixé sur son support solide. On peut alors travailler en chimie Fmoc et utiliser des synthons de Ser ou Thr sans protection sur les chaînes latérales à conditions d'utiliser des conditions de couplages non acylantes. Une autre alternative est de mettre en œuvre des protections très labiles en milieu acide sur les chaînes latérales des résidus Ser ou Thr, comme le groupement trityle, qui peuvent être clivées par réaction avec 1% de TFA, conditions dans lesquelles les autres protections subsistent ainsi que la liaison entre le peptide et le *linker* du support solide. On procède alors à la réaction de phosphorylation en deux étapes : une réaction de phosphitylation et une réaction d'oxydation. La réaction de phosphitylation utilise des composés de type phosphochloridates, H-phosphonates ou phosphoramidites. Les plus couramment utilisés en chimie Fmoc sont les dibenzyl- et diterbutyl-N,N-dialkyl phosphoramidites, pour lesquels les groupements benzyle et terbutyle sont clivés lors de la coupure finale au TFA. L'étape d'oxydation peut être réalisée par action de l'acide m-chloroperbenzoïque, l'iode aqueux ou le terbutyl-hydroperoxyde. On choisira de préférence ce dernier réactif dans le cas de peptides contenant des résidus Cys, Met et Trp, sensibles aux conditions oxydantes car il est le moins favorable aux réactions secondaires d'oxydation sur ces résidus. Cette méthode cependant ne permet pas toujours d'obtenir des rendements de 100%, entre autre, parce qu'elle met en jeu plusieurs étapes réactionnelles. De plus, ces réactions utilisent des composés organophosphorés, soit très réactifs, facilement hydrolysables en milieux aqueux ou par la présence de traces d'eau dans les solvants organiques, ou encore au contact de l'air, comme les phosphochloridites et les phosphoramidites, soit beaucoup moins réactifs comme les phosphochloridates. Ces composés organophosphorés ont été exploités surtout dans la synthèse des oligonucléotides où ils jouent le rôle d'agents activateurs et ont connu un essor important avec le développement de la chimie supportée.

La phosphorylation par les kinases présente l'avantage d'être réalisée dans des conditions plus douces, favorables à la phosphorylation des protéines. La sélectivité est moins bonne que celle obtenue par la voie des réactifs organophosphorés dans le cas de peptide de synthèse, et elle est variable en fonction de la kinase employée. Le taux de phosphorylation est également plus

faible. Cependant, la phosphorylation mettant en œuvre des kinases présente un intérêt non négligeable, celui d'être plus proche des conditions cellulaires.

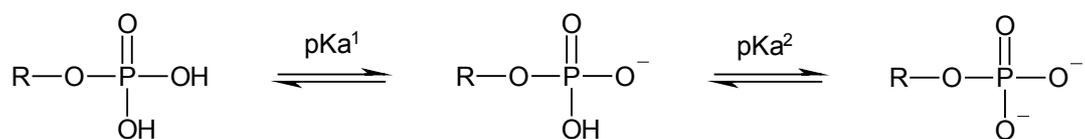
Pour la synthèse des phosphopeptides employés dans le criblage des phospho-épitopes de la protéine tau impliqués dans l'interaction avec Pin1, nous avons choisi d'utiliser les synthons de Thr ou Ser phosphorylés commerciaux, malgré le coût de ces composés, pour faciliter l'obtention des phospho-peptides garantissant ainsi une synthèse rapide, totalement automatisable car les conditions de couplage sont les mêmes que pour les autres acides aminés. Le seul inconvénient est la déprotection du groupement benzyle du phosphate qui nécessite des temps de réaction plus long dans le TFA comparé aux peptides non phosphorylés mais qui peut être achevée en 6 heures avec un rendement de déprotection du phosphate de 90%.

Les conditions expérimentales mises en œuvre pour la détermination de l'affinité des phospho-peptides jouent un rôle essentiel qui mérite une attention particulière. Les conditions de tampon utilisées, notamment le pH de la solution tampon, joue un rôle sur l'état d'ionisation du phosphate.

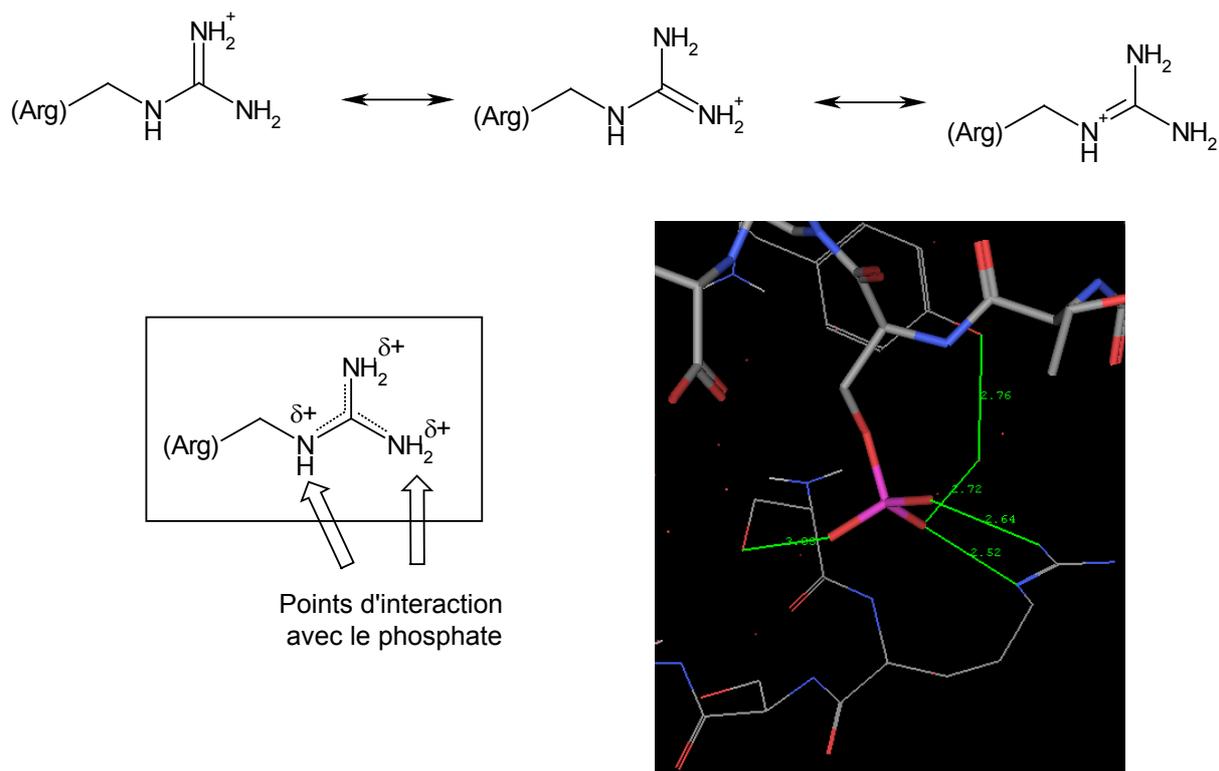
### 1.1.3 – Influence du pH sur l'état d'ionisation du phosphate.

En RMN, les conditions de pH sont limitées afin d'observer les protons échangeables tels que les H<sub>N</sub> du squelette protéique. A des pH trop élevés, les protons sont en échange rapide avec l'eau et leur signal est élargi. La zone de pH accessible se limite donc à des pH inférieurs à 8.0, voire 7.0, pour l'acquisition de spectres de protéines. Les spectres de Pin1 et du domaine WW isolé ont été réalisés à pH 6.80 afin de prendre en compte ces limites expérimentales mais également d'optimiser les conditions d'interaction en favorisant l'ionisation du groupement phosphate.

Le phosphate porté par les résidus Ser, Thr ou Tyr peut exister selon le pH sous trois formes différentes : une forme non ionisée à pH acide, une forme di-anionique à pH basique et une forme mono-anionique à pH intermédiaire.



Dans l'interaction avec Pin1, le phosphate porté par le substrat forme une liaison de type électrostatique avec le résidu Arg17 situé dans une boucle flexible  $\beta$ -hairpin entre les feuillet  $\beta 1$  et  $\beta 2$  du domaine WW. L'interaction est donc plus efficace quand le phosphate se présente sous la forme di-anionique, comme indiquée sur le **Schéma 2.1**, où sont représentés les distances entre le groupement phosphate du phospho-peptide et les différents résidus du domaine WW impliqués dans la liaison du phosphate Ser16 et Arg17.



**Schéma 2.1 :** Représentation schématique de la délocalisation électronique au niveau du groupement guanidinium de l'arginine et du mode de liaison du phosphate par le groupement guanidinium du résidu Arg17 du domaine WW d'après la structure cristallographique de Verdecia (Verdecia, Bowman et al. 2000).

Le pH des expériences de titration se situant autour de 6.80, il convient de vérifier si le phosphate est présent sous forme di-anionique dans ces conditions. Nous avons donc titré deux peptides phosphorylés par un simple dosage acide-base en suivant par RMN<sup>31</sup>P les modifications de déplacement chimique du phosphore. On peut ainsi déterminer sélectivement le pKa du groupement phosphate dans le contexte peptidique plutôt que le pI global du peptide.

Le calcul des valeurs de pKa par RMN peut être réalisé à partir de la formule de Henderson-Hasselbach (Fischer 2000) (Équation 2.1).

$$\delta = \delta_1 + (\delta_2 - \delta_1) \times \frac{10^{(pH-pK_1)}}{1 + 10^{(pH-pK_1)}} + (\delta_3 - \delta_2) \times \frac{10^{(pH-pK_2)}}{1 + 10^{(pH-pK_2)}}$$

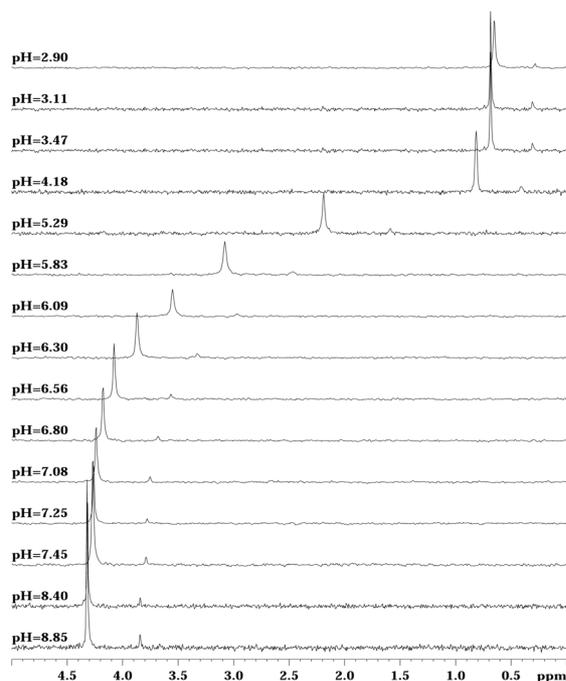
**Équation 2.1 :** Calcul du déplacement chimique d'un atome dans une espèce existant sous plusieurs formes acide/base en équilibre selon Henderson-Hasselbach en fonction du pH, des valeurs des différents pKa correspondant aux différents équilibre acido-basiques et des valeurs de déplacement chimiques correspondant aux plateaux ; pK<sub>1</sub> est le premier pKa (pKa acide), pK<sub>2</sub>, le deuxième pKa (pKa basique), δ<sub>1</sub>, la valeur de déplacement chimique correspondant au premier plateau (pH acide), δ<sub>3</sub>, la valeur de déplacement chimique correspondant au dernier plateau (pH basique), δ<sub>2</sub>, la valeur de déplacement chimique correspondant au plateau intermédiaire (pH intermédiaire, premier équilibre protonation/déprotonation).

Dans le cas d'un seul équilibre acide-base, que l'on utilisera pour déterminer le pKa<sub>2</sub> du phosphate présent dans les phospho-peptides, on peut simplifier la formule pour obtenir l'**Équation 2.2**.

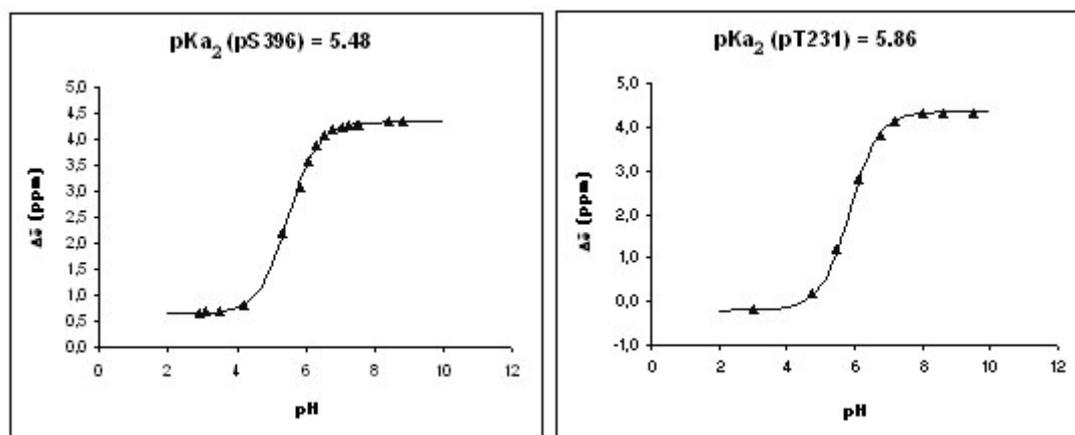
$$\delta = \delta_1 + (\delta_2 - \delta_1) \times \frac{10^{(pH-pK_1)}}{1 + 10^{(pH-pK_1)}}$$

**Équation 2.2 :** Équation 2.1 simplifiée dans le cas d'un seul équilibre acido-basique.

La mesure de la seconde acidité du phosphate porté par les résidus Ser et Thr dans les peptides pS396 et pT231 respectivement a été réalisée à partir de la détermination des variations de déplacement chimique du phosphore en fonction du pH sur une gamme allant de 2.5 à 9.0 environ (Figure 2.4).



**Figure 2.4 :** Spectres <sup>31</sup>P d'un phospho-peptide à différentes valeurs de pH. Le déplacement chimique du phosphore du groupement phosphate est utilisé pour calculer le pKa correspondant à l'équilibre R-PO<sub>4</sub>H/R-PO<sub>4</sub><sup>2-</sup> dans le contexte peptidique selon l'**Équation 2.2**.



**Graphique 2.1 :** Représentation des variations de déplacement chimique du  $^{31}\text{P}$  du groupement phosphate dans le peptide pS396 (à gauche) et pT231 (à droite) donnant respectivement des valeurs de  $\text{pK}_{\text{a}2}$  de 5.48 et 5.86.

Les  $\text{pK}_{\text{a}}$  déterminés pour la phosphoSer396 et la phosphoThr231 dans leur contexte peptidique aboutissent à des valeurs relativement proches de 5.48 et 5.86 (Graphique 2.1). Nous n'avons pas déterminé systématiquement le  $\text{pK}_{\text{a}}$  de chaque résidu phosphorylé dans leur contexte peptidique mais il semble que la séquence primaire n'influence pas de manière notable la valeur du  $\text{pK}_{\text{a}}$  du phospho-résidu.

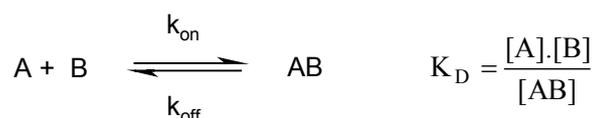
Nous avons également effectué une titration de Pin1 avec le peptide pT231 à pH 8.20 et comparé l'affinité trouvée avec celle de ce même peptide dans une titration réalisée à pH 6.80. Il faut noter qu'un certain nombre de résonances du domaine WW disparaissent à pH 8.20, mais aucune différence significative de l'affinité n'a été observée entre les deux conditions de pH mises en œuvre. Il n'y a donc pas de modification de l'équilibre entre les formes mono- et di-anioniques du phosphate entre les deux valeurs de pH.

## 1.2 – Bases thermodynamique et cinétique des interactions moléculaires.

Les interactions moléculaires non covalentes sont des événements essentiels dans les cellules et consistent en la liaison d'une ou plusieurs molécules appelées ligand à une macromolécule. Ces ligands peuvent être de petites molécules telles que les co-facteurs ou les substrats d'enzymes, ou des macromolécules dans le cas des interactions protéine-protéine, protéine-ADN .... La description des phénomènes de liaisons non-covalentes représente un aspect de la compréhension des mécanismes d'action enzymatique, de régulation et de signalisation cellulaire..., qui nécessite entre autres la connaissance du nombre de sites de liaison et la force de la liaison, la cinétique de formation et de dissociation des complexes. Il existe

plusieurs types de liaisons : la liaison à un site unique ou à plusieurs sites équivalents ou non et permettent d'adapter un seul type de molécule ou des molécules diverses. Dans le cas de sites multiples, les sites peuvent être inter-dépendants avec régulation de la liaison d'une molécule au niveau d'un site par la liaison d'une autre molécule à un autre site (site allostérique). Dans cette étude, nous nous limiterons à l'étude d'un site de liaison unique.

Les interactions moléculaires sont décrites par la formation de complexes entre deux partenaires ou plus. Dans le cas le plus simple d'une interaction bi-moléculaire entre deux composés A et B, deux types d'informations peuvent être extraites : l'information thermodynamique liée à la force de l'interaction, autrement dit l'affinité, régie par une constante thermodynamique, la constante de dissociation à l'équilibre  $K_D$  ou la constante d'affinité  $K_A$  ( $K_A = 1/K_D$ ), et l'information cinétique liée à la vitesse d'association et de dissociation, décrite par les constantes cinétiques d'association  $k_{on}$  ( $M^{-1}.s^{-1}$ ) et de dissociation  $k_{off}$  ( $s^{-1}$ ).



Les concentrations des composants à l'équilibre sont décrites par :

$$[A]_0 = [A] + [AB]$$

$$[B]_0 = [B] + [AB]$$

où  $[A]_0$  et  $[B]_0$  sont les concentrations initiales de A et B.

La concentration de complexe à l'équilibre s'exprime de la manière suivante à partir de l'expression du  $K_D$  :

$$[AB] = \frac{1}{2}([A]_0 + [B]_0 + K_D - \sqrt{([A]_0 + [B]_0 + K_D)^2 - 4[A]_0[B]_0})$$

Les paramètres cinétiques d'association et de dissociation sont décrits comme suit :

$$v_{\text{association}} = k_{on}[A][B] = +\frac{d[AB]}{dt} \quad \text{et} \quad v_{\text{dissociation}} = k_{off}[AB] = -\frac{d[AB]}{dt},$$

$$\text{donc, à l'équilibre, } \frac{d[AB]}{dt} = k_{on}[A][B] - k_{off}[AB] = 0,$$

$$\text{et } K_D = \frac{[A].[B]}{[AB]} = \frac{k_{off}}{k_{on}}.$$

En conséquence, cinétique d'association/dissociation et équilibre thermodynamique ne peuvent être dissociés pour décrire et mesurer un phénomène biologique d'interaction. En particulier, l'affinité est étroitement liée à la vitesse de dissociation des complexes. En effet, on peut estimer que la vitesse d'association est contrôlée principalement par la diffusion moléculaire, c'est-à-dire qu'on ne tient pas compte d'éventuels réarrangements structuraux ou de forces inter-moléculaires (interactions électrostatiques...) qui mènerait vers une association. Les valeurs communément utilisées pour décrire les associations sont de l'ordre de  $10^7 - 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$ . La constante de diffusion a été évaluée pour l'ubiquitine à  $1.7 \cdot 10^{-10} \text{ m}^2/\text{s}$  (la diffusion de l'eau étant de  $2.2 \cdot 10^{-9} \text{ m}^2/\text{s}$ ) et la distance moyenne entre deux molécules à  $70 \text{ \AA}$  pour une solution à  $1 \text{ mM}$  (Dingley, Mackay et al. 1995). Par conséquent, plus l'affinité est élevée, donc le  $K_D$  faible, plus le  $k_{\text{off}}$  sera faible, traduisant un échange lent entre la forme libre d'une protéine et la forme en complexe. Avec certaines techniques de mesure des interactions biologiques, comme le *pull-down* et la co-immunoprécipitation, des dissociations rapides sont défavorables à l'isolation des complexes qui peuvent être dissociés au cours des étapes de lavage par exemple. Ces techniques, contrairement aux techniques spectroscopiques comme la RMN ou la fluorescence, ne sont pas adaptées à la détection de complexes faiblement affins ( $K_D < 1 \text{ }\mu\text{M}$ ).

#### ***Caractérisation de la dynamique de l'échange - Temps de résidence.***

Dans l'hypothèse où la vitesse d'association dépend uniquement de la diffusion, soit un  $k_{\text{on}}$  de  $10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$ , la durée pendant laquelle deux partenaires restent en complexe, ou temps de résidence, peut être définie par la valeur du  $K_D$  comme le décrit le Tableau 2.2.

$K_D$	$k_{\text{off}} (\text{s}^{-1})$	Temps de résidence
1 nM	0.01	100 s
1 $\mu\text{M}$	10	100 ms
10 $\mu\text{M}$	100	10 ms
1 mM	10 000	100 $\mu\text{s}$
1 M	10 000 000	100 ns

**Tableau 2.2 :** Relations entre le  $K_D$ , la vitesse de dissociation et le temps de résidence.

D'un point de vue énergétique, la force d'une interaction est directement proportionnelle à la variation d'énergie libre  $\Delta G$ , décrite par l'équation suivante, où T est la température (en K),

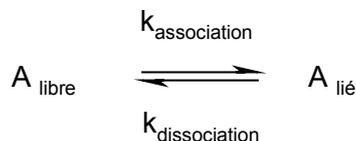
$\Delta H$ , la variation d'enthalpie (en kcal.mol<sup>-1</sup>) et  $\Delta S$ , la variation d'entropie du système (en kcal.mol<sup>-1</sup>.K<sup>-1</sup>) quand l'interaction se produit :

$$\Delta G \text{ (kcal.mol}^{-1}\text{)} = \Delta H - T\Delta S$$

$$\text{et } \Delta G = -RT \ln(K_D)$$

Comme le deuxième terme  $T\Delta S$  varie avec la température, une diminution de la température est favorable aux réactions dirigées par l'enthalpie et défavorable aux réactions gouvernées par l'entropie. La majorité des interactions biologiques sont gouvernées par l'enthalpie. Les variations d'enthalpies, contrairement aux variations d'entropie, sont faciles à prédire. Par exemple, dans les conditions standards (25°C, 1 atm), la formation d'une liaison hydrogène libère environ 1 kcal.mol<sup>-1</sup>, soit une diminution du  $K_D$  d'un facteur 5.4 environ, alors qu'une liaison ionique libère environ 2 kcal.mol<sup>-1</sup>, soit une diminution du  $K_D$  d'un facteur 30, et les deux simultanément, s'il y a additivité des contributions, représentent alors une diminution du  $K_D$  d'un facteur 150.

L'équilibre association/dissociation entre deux partenaires A et B qui forment un complexe peut être décrit de manière simplifiée par un équilibre entre la forme libre et la forme liée (en complexe) de l'un des deux partenaires comme suit :



où  $A_{\text{libre}}$  est la forme libre de A et  $A_{\text{lié}}$  la forme en complexe de A avec B et

$$k_{\text{association}} = k_{\text{on}} \cdot [B] \text{ (s}^{-1}\text{)} \quad \text{et} \quad k_{\text{dissociation}} = k_{\text{off}} \text{ (s}^{-1}\text{)}.$$

On peut alors calculer la constante de dissociation par mesure d'un paramètre macroscopique P, qui varie en fonction des quantités de  $A_{\text{libre}}$  et  $A_{\text{lié}}$  à l'équilibre. La variation de ce paramètre macroscopique  $\Delta P$  peut se définir comme suit :

$$[A]_0 P = [A]_0 (P_{\text{libre}} \% A_{\text{libre}} + P_{\text{lié}} \% A_{\text{lié}}),$$

où  $P_{\text{libre}}$  est la valeur de P en l'absence de B et  $P_{\text{lié}}$ , la valeur de P correspondant à l'état de saturation.

En simplifiant l'équation avec  $\Delta P = P - P_{\text{libre}}$  et  $\Delta P_{\text{max}} = P_{\text{lié}} - P_{\text{libre}}$ , on obtient :

$$\Delta P = \Delta P_{\text{max}} \frac{[A_{\text{lié}}]}{[A]_0}$$

Comme  $[A_{\text{lié}}]$  s'exprime en fonction du  $K_D$ , on peut alors calculer le  $K_D$  en fonction de  $[A]_0$ ,  $[B]_0$ ,  $\Delta P$  et  $\Delta P_{\text{max}}$  :

$$\Delta P = \frac{\Delta P_{\text{max}}}{2} \left( 1 + X + \frac{K_D}{[A]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[A]_0} \right)^2 - 4X} \right) \quad \text{où} \quad X = \frac{[B]_0}{[A]_0}$$

Dans le cas particulier des interactions protéine-protéine (ou protéine-peptide) qui nous intéressent, l'interaction est traduite par un équilibre entre les formes libre et liée (en complexe) de la protéine. Dans les paragraphes suivants, les paramètres macroscopiques qui permettront de déterminer la valeur des constantes de dissociation seront le déplacement chimique ( $\Delta\delta$ ) en RMN et l'intensité de fluorescence ( $\Delta I$ ) en fluorescence intrinsèque ou en FRET. En ce qui concerne les paramètres cinétiques qui gouvernent la formation et la dissociation des complexes, la RMN et la résonance plasmonique de surface pourront apporter les informations nécessaires à la détermination des constantes  $k_{\text{on}}$  et  $k_{\text{off}}$ . En RMN, par exemple, dans les conditions mises en œuvre pour la reconnaissance de phospho-épitopes de la protéine tau par Pin1, la vitesse d'association est essentiellement gouvernée par la diffusion. En revanche, la vitesse de dissociation aura un impact remarquable sur l'élargissement des signaux de la protéine qui interviennent dans l'interaction avec le substrat. Une méthode de simulation de l'échange chimique et son impact sur les signaux RMN sera présentée dans le **paragraphe 1.3.2.**

### **1.3 – Détermination des constantes d'affinité en milieu homogène. Utilisation de deux techniques spectroscopiques, la RMN et la fluorescence.**

#### **1.3.1 – La RMN à haute résolution, un outil de détection des interactions de faible affinité.**

La majorité des applications modernes de la RMN pour l'étude des macromolécules d'intérêt biologique implique l'analyse d'un spectre RMN à deux dimensions ou plus. Les développements méthodologiques ont été accompagnés par une utilisation croissante de la RMN pour l'étude des structures macromoléculaires et des interactions en solution. L'attribution des résonances des différents noyaux d'une molécule est la première étape d'une étude par RMN, et l'avènement des expériences multidimensionnelles a rendu cette tâche de plus en plus facile. Les informations structurales sont obtenues principalement par la mesure du NOE, qui fournit les contraintes concernant les distances internucléaires. En plus de la

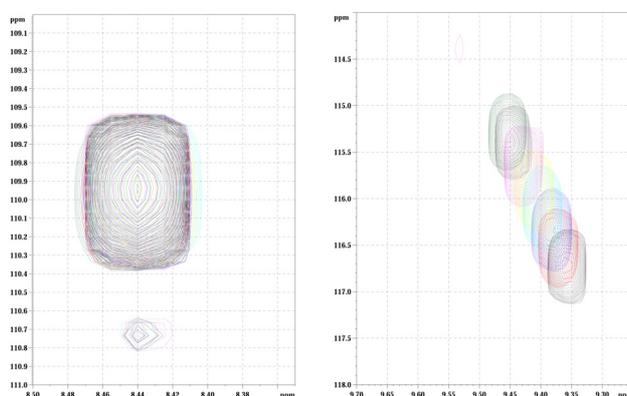
détermination des structures tridimensionnelles complètes de macromolécules en solution, la RMN fournit des informations très précieuses sur les structures locales, la dynamique conformationnelle et sur les aspects à la fois structuraux et cinétiques concernant les interactions avec de petites molécules. La RMN peut être utilisée pour caractériser l'affinité d'un ligand pour une protéine donnée, sa conformation et la vitesse de dissociation, mais elle peut également identifier les contacts entre les atomes du ligand et ceux de la protéine. Ce type d'informations peut être obtenu grâce au marquage isotopique. La capacité de combiner les informations structurales et dynamiques est peut-être le plus grand intérêt de la RMN dans le contexte de la biologie moléculaire. La RMN permet de détecter et de quantifier des interactions avec une grande sensibilité sans forcément connaître la fonction biologique de la protéine d'intérêt. On peut également avoir accès aux informations structurales à la fois de la protéine et de son substrat. Un des intérêts majeurs est la caractérisation des complexes macromoléculaires de faible affinité, pour lesquels peu de techniques d'investigation existent, mais également des complexes de haute affinité, accessibles par d'autres techniques (co-immunoprécipitation, *western blot*, test en double-hybride...).

Les études de *binding* entre Pin1 et des substrats peptidiques se situent dans le domaine des interactions de faibles affinités, comme c'est souvent le cas pour les interactions impliquant des enzymes. La RMN peut alors détecter de telles interactions au niveau de complexes qui sont qualifiés de « transitoires », non décelables par les techniques courantes d'investigation des interactions protéine-protéine, et distinguer des modulations d'affinité dans cette gamme inférieure au seuil détectable par ces techniques.

***Le déplacement chimique, une sonde sensible pour la caractérisation des constantes de dissociation.***

Le déplacement chimique d'un noyau est, par définition, une mesure de la dépendance de la fréquence de résonance de ce noyau en fonction de son environnement chimique. Le déplacement chimique est ainsi un paramètre extrêmement sensible à l'environnement chimique. Chaque atome de la molécule sera influencé par les atomes voisins à travers les liaisons chimiques mais également voisins dans l'espace. Cette propriété peut être utilisée pour identifier des interactions entre deux partenaires moléculaires, quantifier ces interactions en termes d'affinité de liaison et déterminer le site de l'interaction par *mapping* des déplacements chimiques.

Dans l'étude visant à identifier des phospho-épitopes de la protéine tau reconnus par Pin1, deux méthodes sont envisageables : le suivi des variations des déplacements chimiques au niveau du phospho-peptide ou le suivi des variations des déplacements chimiques au niveau de la protéine. Dans le premier cas, l'avantage est que la méthode requiert de faibles quantités de protéine non enrichie. Dans le deuxième cas, la technique nécessite l'enrichissement isotopique de la protéine, au moins en  $^{15}\text{N}$ , et des concentrations relativement élevées de protéine de l'ordre de 100  $\mu\text{M}$  et plus. Les concentrations de protéines sont variables et dépendent notamment de la taille de la protéine, dans la mesure où la taille est souvent corrélée avec la largeur de raie dont dépend la sensibilité (plus les raies sont fines, plus le signal est intense pour une même valeur de l'intégrale du pic, donc meilleure est la sensibilité). Dans le cas de détection des phénomènes de liaison au travers des signaux de la protéine enrichie en  $^{15}\text{N}$ , la quantité d'informations disponibles est plus importante. Elle concerne l'identification du(des) site(s) d'interaction du ligand dans la protéine et la nature de l'interaction. De plus, dans le cas de faibles affinités, les concentrations de ligand ajoutées sont largement supérieures à la concentration de protéine. Le transfert d'aimantation vers le noyau  $^{15}\text{N}$  permet de s'affranchir du signal du ligand, plus intense que celui de la protéine, qui pourrait masquer les signaux de celle-ci. Les résidus de la protéine impliqués dans l'interaction voient leur déplacement chimique modifié par la présence du ligand. De même, les résidus qui ne sont pas directement impliqués dans l'interaction mais proches du site d'interaction peuvent également subir des perturbations. La variation observée dépend de la nature de l'interaction, c'est-à-dire des fonctions chimiques impliquées. La RMN permet ainsi de détecter des interactions même très faibles mais spécifiques (Figure 2.5).



**Figure 2.5 :** Détail des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC superposés de la protéine Pin1 enrichie en  $^{15}\text{N}$  en présence de différentes concentration en peptide substrat. A gauche, le signal correspondant au résidu Trp11 qui ne varie pas en présence du peptide ; à droite, le signal correspondant au résidu Gln33, dont le déplacement chimique varie en fonction des concentrations de substrat ajouté.

L'affinité de liaison peut également être évaluée en mesurant la variation de déplacement chimique d'un ou plusieurs résidus de la protéine en fonction de quantités croissantes de ligand ajoutés (Équation 2.3). La variation de déplacement chimique peut prendre en compte la seule variation en proton mais également une combinaison des variations dans les dimensions proton et azote (Équation 2.4).

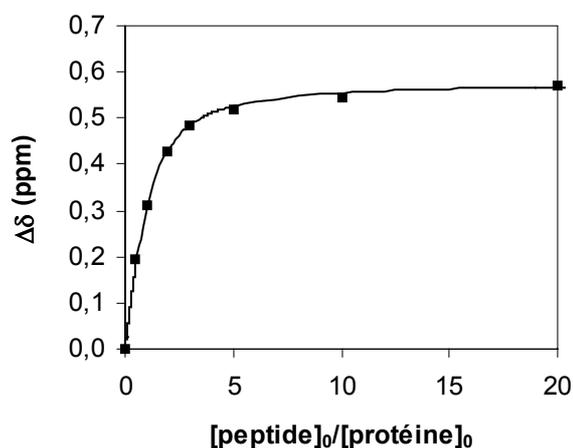
$$\Delta\delta(\text{ppm}) = 0.5\Delta\delta_{\text{max}} \left( 1 + X + \frac{K_D}{[\text{protéine}]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[\text{protéine}]_0} \right)^2 - 4X} \right)$$

**Équation 2.3 :** Equation utilisée pour la détermination des constantes de dissociation des complexes Pin1-phospho-peptides.

$$\Delta\delta (\text{ppm}) = [\Delta\delta^2(^1\text{H}) + 0.2 \Delta\delta^2(^{15}\text{N})]^{1/2}$$

**Équation 2.4 :** Variation pondérée du déplacement chimique d'un résidu en fonction de la variation en proton ( $^1\text{H}$ ) et de la variation en azote ( $^{15}\text{N}$ ).

On peut alors déterminer la constante de dissociation du complexe ( $K_D$ ) en faisant varier ce paramètre de telle sorte que la courbe théorique obtenue à partir de l'Équation 2.3 définisse au mieux les valeurs expérimentales (Graphique 2.2).



**Graphique 2.2 :** Représentation de la variation de déplacement chimique d'un résidu en fonction du rapport des concentrations initiales peptide sur protéine. La valeur du  $K_D$  est calculée de telle sorte que la courbe en trait plein sur le graphique, caractérisée par l'Équation 2.3, s'ajuste au mieux aux valeurs expérimentales correspondant aux points.

### 1.3.1.1 – Calcul des constantes de dissociation et *mapping* des sites d'interaction par RMN à haute résolution.

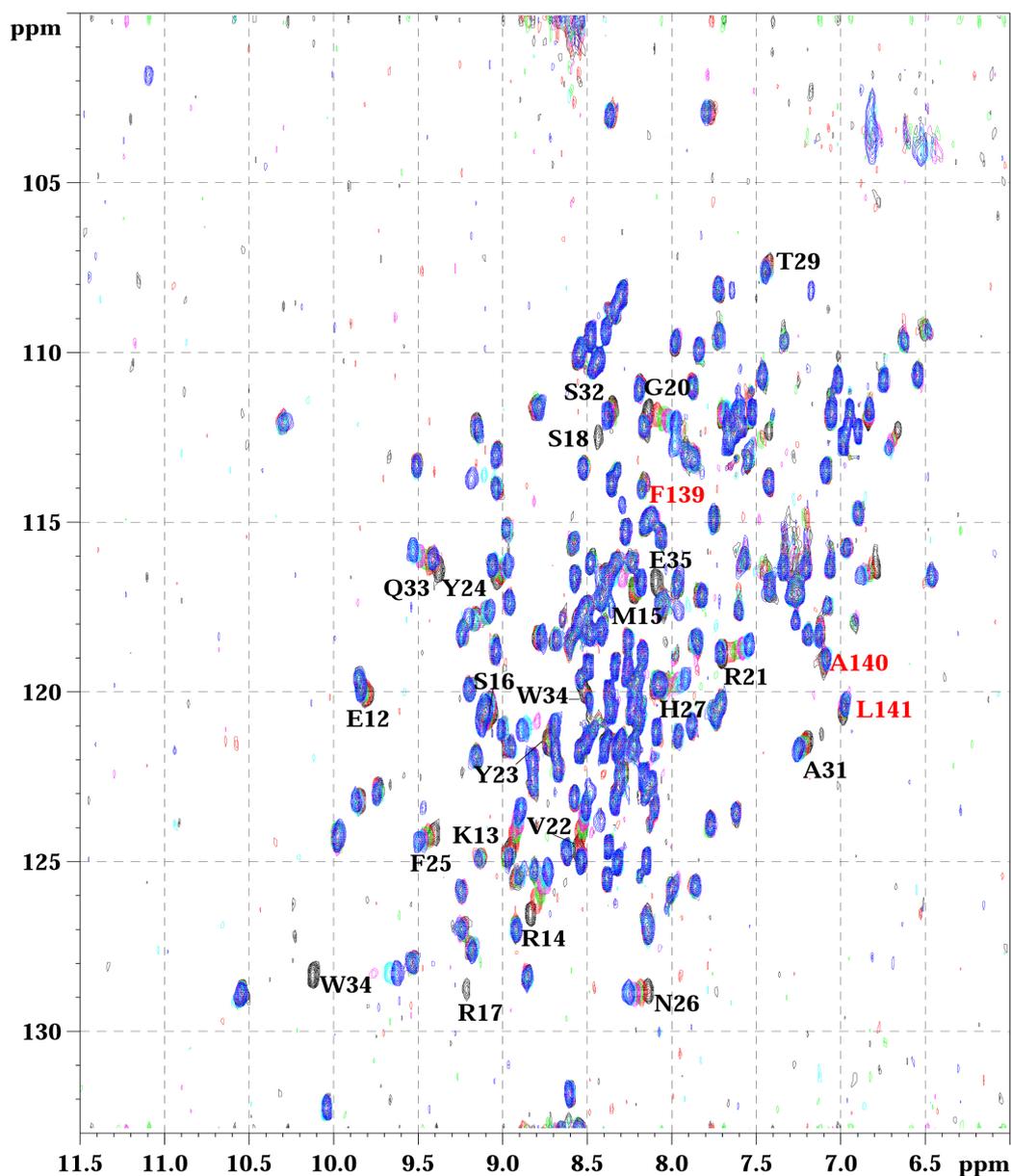
La protéine Pin1, produite par voie recombinante et uniformément enrichie en  $^{15}\text{N}$ , est solubilisée à une concentration de 200  $\mu\text{M}$  environ, dans un tampon Tris deutéré à 50 mM à pH 6.80 contenant 100 mM NaCl, 5 mM DTT, 5 mM EDTA et 10%  $\text{D}_2\text{O}$ . La solution de protéine est titrée par des quantités croissantes de peptides dont les ajouts lyophilisés, préparés à partir d'une solution-mère à 10 mM dans l'eau à pH 6.80, sont repris successivement par la solution de protéine afin de maintenir une concentration constante en protéine. Pour chaque points de titration, l'acquisition d'un spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC est effectuée avec 1024 x 128 points dans les dimensions  $^1\text{H}$  et  $^{15}\text{N}$  respectivement, et 64 *scans* par incrément.

Les résultats des études de *binding* ne montrent aucune sélectivité pour un phospho-épitope particulier de la protéine tau mais une graduation entre les différents sites avec une affinité plus importante pour les motifs phosphoThr-Pro par rapport aux motifs phosphoSer-Pro. L'épitope Thr212 en particulier présente une affinité significativement supérieure par rapport à l'épitope Thr231 (Tableau 2.3). De plus, le dipeptide minimal pThr-Pro se lie au domaine WW avec une affinité de 150  $\mu\text{M}$  confirmant que les interactions observées sur base de la structure cristallographique étaient bien limitées au motif phosphoSer/Thr-Pro dans la séquence peptidique du substrat.

<i>Peptide</i>	$K_D^{\text{PIN1}}$ (mM)
pT231	$0.38 \pm 0.08$
pS235	$0.90 \pm 0.2$
pT231 pS235	$0.18 \pm 0.05$
pT212	$0.10 \pm 0.03$
pT212 pS214	$0.16 \pm 0.02$
pT212 pS214 pT217	$0.10 \pm 0.01$
pS396	$1.0 \pm 0.2$
pS422	$1.0 \pm 0.2$
pThr-Pro (dipeptide)	$0.15 \pm 0.03$

**Tableau 2.3 :** Valeurs des constantes de dissociation et écart-types mesurés par RMN pour les différents complexes formés par les phospho-peptides de tau et la protéine Pin1. Les valeurs de  $K_D$  sont des valeurs moyennes calculées à partir des variations de déplacements chimiques de différents résidus de la protéine.

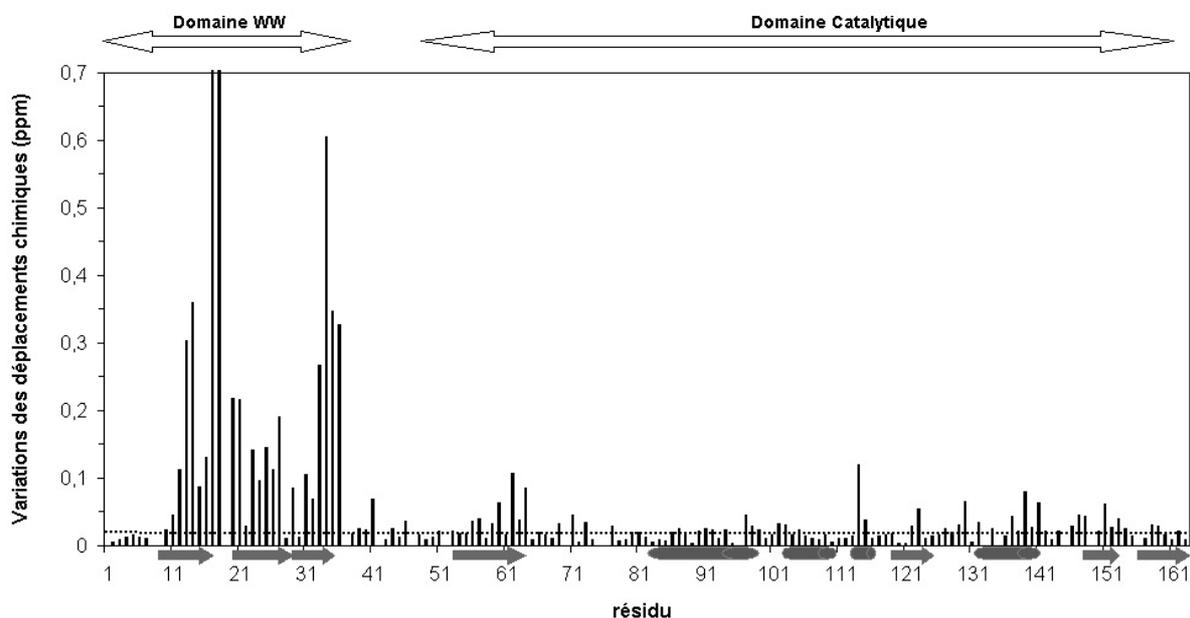
Les résonances affectées par les différents substrats sont principalement celles des résidus du domaine WW d'après l'attribution de Pin1 (Jacobs, Saxena et al. 2002). On note toutefois une très faible variation au niveau des résidus du site catalytique et un comportement intermédiaire pour les résidus du domaine catalytique qui se trouvent à l'interface avec le domaine WW. La structure cristalline de Pin1 en complexe montrait un rapprochement des deux domaines autour du peptide substrat di-phosphorylé malgré l'apparente flexibilité du fragment d'environ 10 résidus qui les sépare (Verdecia, Bowman et al. 2000). Cependant, les contacts majeurs étaient observés avec les résidus du domaine WW. La structure de la protéine en solution établie par RMN montre peu de contacts NOE inter-domaines (Bayer, Goettsch et al. 2003). De même une étude détaillée de la conformation de la protéine en présence de trois peptides substrats (Pintide, Cdc25 et CTD) montre un rapprochement plus ou moins marqué des deux domaines en fonction du peptide, mais toujours relativement faible (Jacobs, Saxena et al. 2003), notamment au niveau de résidus localisés dans l'hélice  $\alpha 4$  et le feuillet  $\beta 6$  du domaine catalytique qui constituent l'interface avec le domaine WW, indiquant une forte indépendance des deux domaines en solution autorisée par la présence d'un *linker* flexible entre les domaines. Dans le cas des peptides que nous avons étudiés, il semble que le comportement soit le même comme le montre la **Figure 2.6**. Les modifications majeures sont observées avec les résidus du domaine WW (annotés en noir). De faibles interactions ont lieu avec le domaine catalytique, essentiellement au niveau des résidus formant l'interface avec le domaine WW (annotés en rouge).



**Figure 2.6 :** Titration de la protéine Pin1 à 200 $\mu$ M par le peptide pT212 de la protéine tau. La figure représente une superposition des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC du domaine enrichi en  $^{15}\text{N}$  pour les rapports  $[\text{protéine}]_0/[\text{peptide}]_0$  suivants : en noir, 1 : 0 ; en rouge, 1 : 0.5 ; en vert, 1 : 1 ; en rose 1 : 2 ; en bleu clair, 1 : 5 ; en bleu foncé, 1 : 20. Les résidus du domaine WW sont annotés en noir. Les résidus annotés en rouge sont ceux du domaine catalytique de l'hélice  $\alpha 4$  située à l'interface avec le domaine WW.

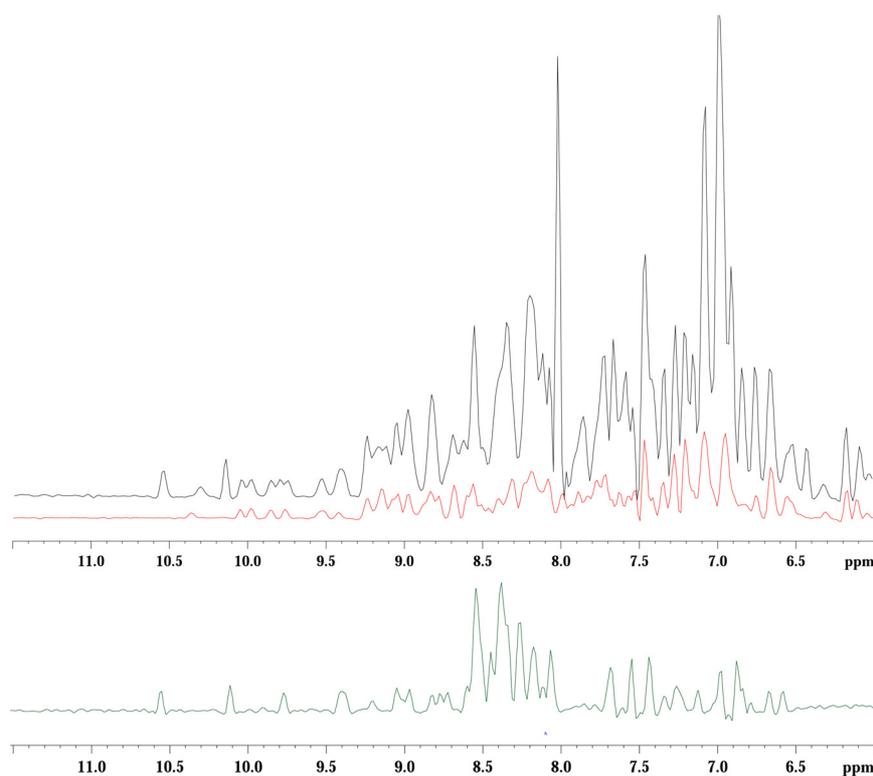
Un *mapping* des déplacements chimiques de la protéine en présence du peptide pT212 permet d'une part, de confirmer la prépondérance des interactions entre le peptide et le domaine de liaison de Pin1, et d'autre part, de déceler un léger rapprochement des deux domaines en présence du substrat, d'après les données de Jacobs et collaborateurs (Jacobs, Saxena et al. 2003). Les principales variations observées au niveau du domaine WW concernent la boucle

de liaison du phosphate comprenant les résidus Ser16, Arg17, Ser18 et Ser19, ainsi que les résidus entourant le Trp34 impliqué dans la liaison de la proline (Graphique 2.3).



**Graphique 2.3 :** Variations des déplacements chimiques des résidus de Pin1 induites par le peptide pT212 à une concentration de 4 mM par rapport à la protéine sans substrat. Les variations les plus importantes ont lieu au niveau du domaine WW de Pin1, et impliquent essentiellement les résidus de la boucle  $\beta$ -hairpin de liaison du phosphate (Ser16 à Ser19) et la zone autour du Trp34 qui lie le résidu Pro du peptide (Gln33 à Glu35).

L'indépendance relative des deux domaines de Pin1 est également remarquable au niveau de leurs largeurs de raie respectives. Les raies du domaine WW sont fines du fait de sa petite taille (moins de 40 résidus) et de sa structure compacte et bien repliée. Cette largeur est conservée dans le contexte protéique de Pin1 alors que les raies du domaine catalytique sont plus larges (Figure 2.7).

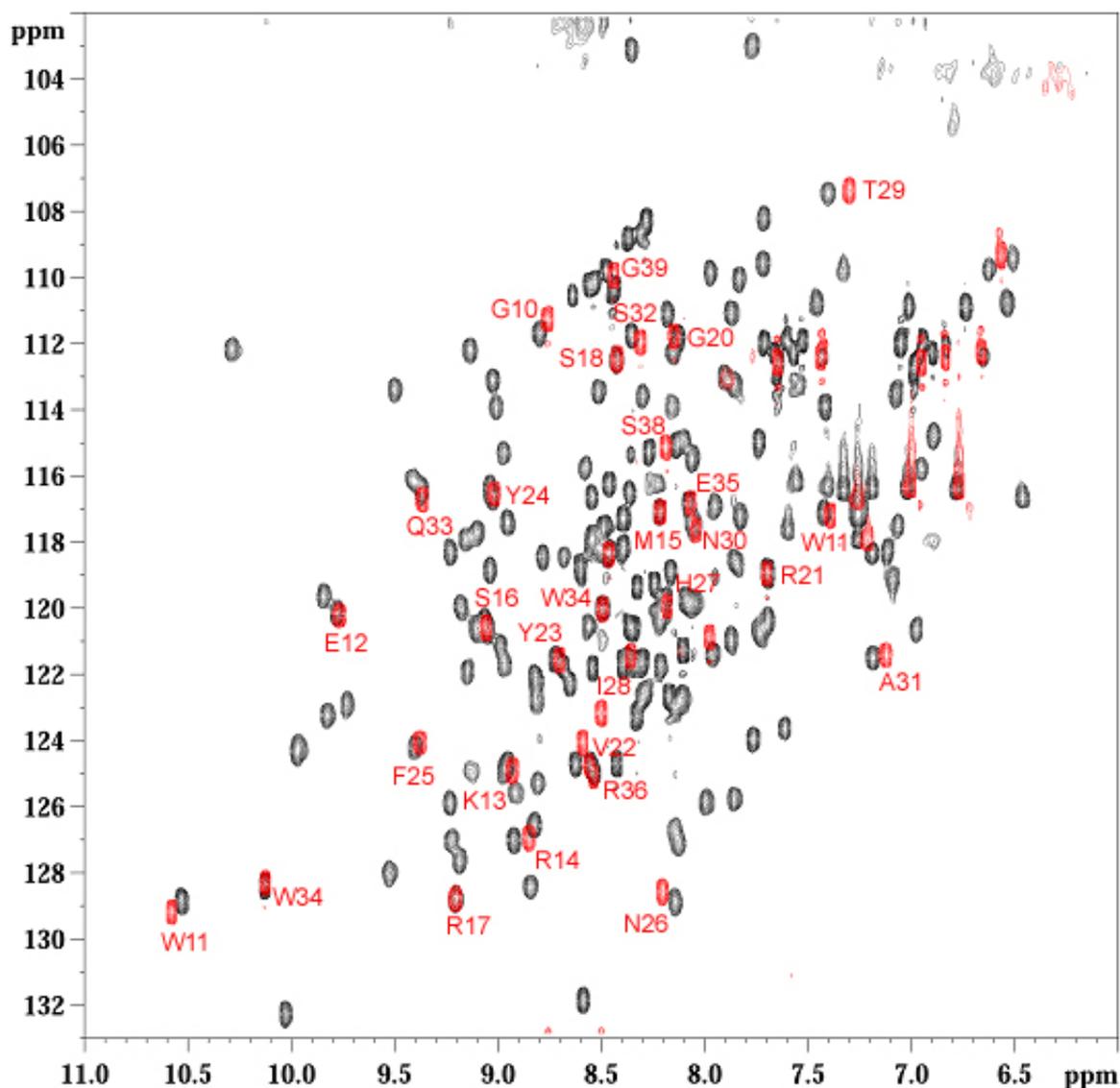


**Figure 2.7 :** Spectres  $^1\text{H}$  de Pin1 (en noir), de son domaine catalytique isolé (en rouge), et du domaine WW isolé (en vert). La largeur de raie du domaine WW isolé et dans son contexte protéique reste la même.

La relative indépendance du domaine WW de Pin1 et la prépondérance des interactions des phospho-peptides avec ce même domaine, gouvernant la liaison du substrat, nous a conduits à étudier ces interactions avec le domaine WW isolé.

### 1.3.1.2 – Etude par RMN du domaine WW isolé. Un modèle simplifié de Pin1 ?

Le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du domaine WW isolé de Pin1 présente une grande ressemblance avec le domaine WW dans son contexte protéique. Nous avons attribué les résonances du domaine isolé grâce aux données de RMN 3D (HNCO, HNCA, HNCOCA) et comparé ces attributions avec celles de Pin1 (Jacobs, Saxena et al. 2002). Peu de variations sont observées, notamment au niveau des résidus impliqués dans la liaison du substrat Ser16-Ser18, Trp34, Tyr23, Tyr24 (Figure 2.8). Les principales perturbations ont lieu au niveau de résidus situés à l'interface avec le domaine catalytique (Thr29-Ser32) ainsi que les résidus avoisinants dans le feuillet (Phe25-Ile28). Les perturbations concernent également les résidus situés à la fin de la séquence (Ser38, Gly39), en amont du *linker*.

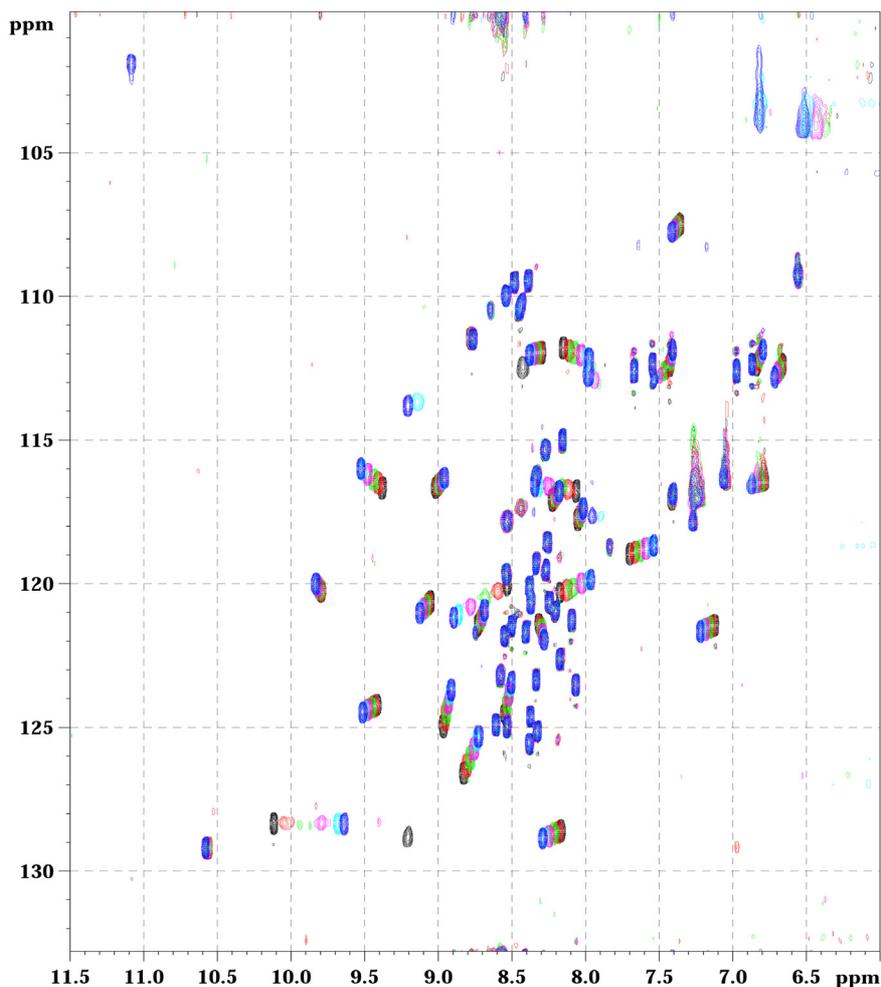


**Figure 2.8 :** Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC superposés de la protéine Pin1 entière (en noir) et du domaine WW isolé (en rouge) avec annotation des résidus du domaine WW en rouge.

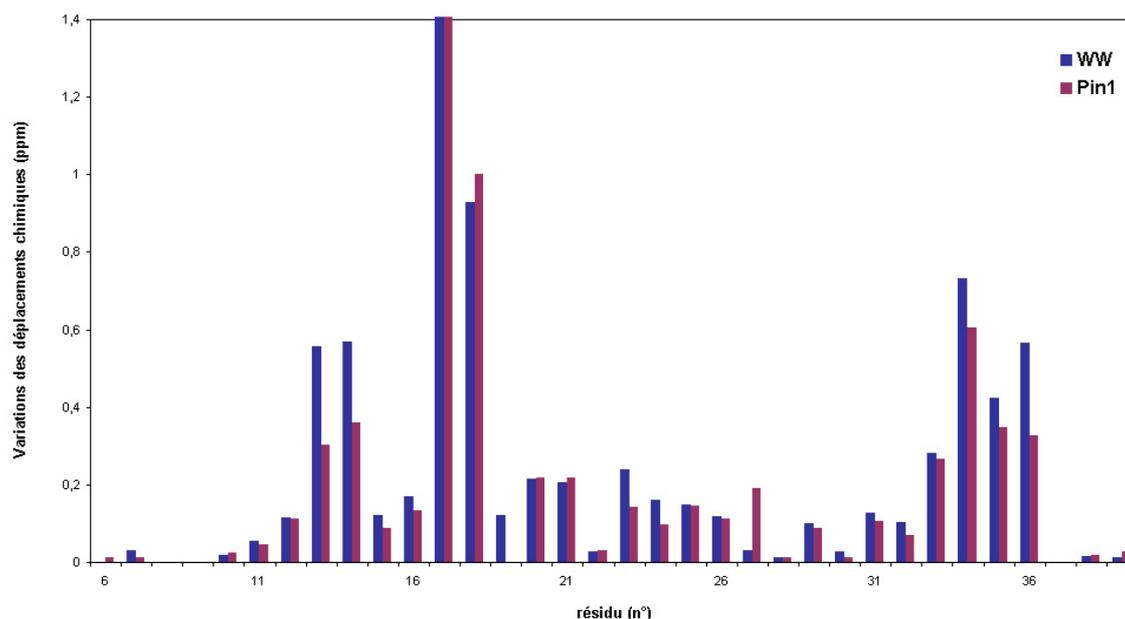
Les constantes de dissociation avec le domaine WW isolé ont été mesurées avec les peptides mono-phosphorylés de tau, pT231 et pT212, et tri-phosphorylé pT212 pS214 pT217, à titre comparatif (Tableau 2.4) dans les mêmes conditions qu'avec Pin1. Les valeurs obtenues sont proches des valeurs obtenues avec Pin1 entière pour les peptides mono-phosphorylés, indiquant que le domaine catalytique ne participe pas à la liaison du substrat. Le mode de liaison du substrat n'est pas affecté puisque les résonances perturbées sont les mêmes que celles du domaine WW dans la protéine Pin1 (Figure 2.9, Graphique 2.4). La liaison du peptide tri-phosphorylé est caractérisée par un  $K_D$  inférieur au  $K_D$  mesuré avec Pin1, avec une différence d'un facteur 5, indiquant une participation négative du domaine catalytique dans la liaison de ce substrat.

<i>Peptide</i>	$K_D^{WW}$ (mM)
pT231	$0.400 \pm 0.070$
pT212	$0.170 \pm 0.030$
pT212 pS214 pT217	$0.020 \pm 0.007$

**Tableau 2.4 :** Valeurs des constantes de dissociation et écart-types mesurés par RMN pour les complexes formés par les phospho-peptides de tau et le domaine WW isolé de Pin1. Les valeurs de  $K_D$  sont des valeurs moyennes calculées à partir des variations de déplacements chimiques de différents résidus de la protéine.



**Figure 2.9 :** Titration du domaine WW isolé à 200µM par le peptide pT212 de la protéine tau. La figure représente une superposition des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC du domaine enrichi en  $^{15}\text{N}$  pour les rapports  $[\text{protéine}]_0/[\text{peptide}]_0$  suivants : en noir, 1 : 0 ; en rouge, 1 : 0.5 ; en vert, 1 : 1 ; en rose 1 : 2 ; en bleu clair, 1 : 5 ; en bleu foncé, 1 : 20.



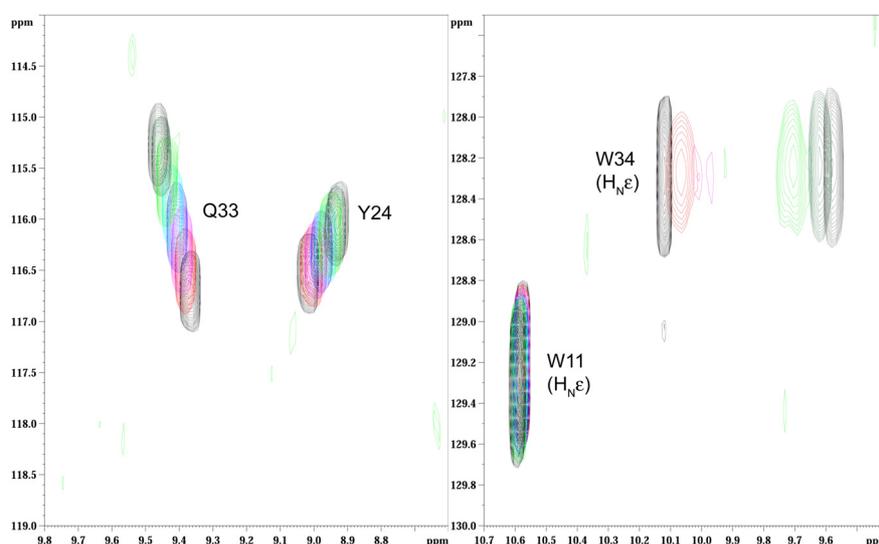
**Graphique 2.4 :** Mapping des déplacements chimiques des résidus du domaine WW isolé (bleu) et du domaine WW dans Pin1 (violet), en présence du peptide pT212 à une concentration de 4 mM par rapport à la protéine sans substrat. Les variations les plus importantes impliquent les résidus de la boucle  $\beta$ -hairpin de liaison du phosphate (Ser16 à Ser19) et la zone autour du Trp34 qui lie le résidu Pro du peptide (Gln33 à Glu35).

Le domaine WW de Pin1 peut être étudié isolé de son contexte protéique pour le criblage des phospho-épitopes de la protéine tau puisque le domaine catalytique ne semble pas intervenir de manière significative dans la liaison du substrat. Ces données ne permettent pas de déterminer si les deux domaines lient la même molécule-substrat ou deux molécules distinctes. Le premier cas est le moins probable dans le cas d'un substrat mono-phosphorylé comme le peptide pT212, compte-tenu du faible rapprochement des deux domaines. Le domaine catalytique interagit avec le peptide comme le montrent les perturbations au niveau des résonances du site catalytique mais on ne peut pas véritablement parler d'une liaison puisque le  $K_D$  est supérieur à 5 mM. La proportion de substrat interagissant avec le domaine catalytique ne diminue pas significativement la quantité de substrat accessible au domaine WW, de sorte que l'affinité du peptide n'est pas affectée par la présence ou l'absence du domaine catalytique.

### 1.3.2 – Caractérisation de la cinétique de l'échange par RMN.

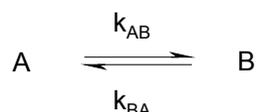
Le déplacement chimique est un paramètre extrêmement sensible, par définition, à l'environnement chimique. Pour un proton donné dans une molécule, par exemple, le déplacement chimique est un paramètre qui sera influencé par l'environnement direct du

proton à travers les liaisons covalentes mais également au travers des protons proches dans l'espace appartenant soit à la même molécule et relatif à l'échantillonnage conformationnel, soit à une autre molécule, sous l'effet d'une association inter-moléculaire. C'est l'effet des protons distants au travers des liaisons covalentes mais proches dans l'espace qui nous intéresse, en particulier l'effet de l'échange chimique entre une forme libre et une forme liée lors de l'association de deux molécules sous forme de complexe comme nous l'avons décrit pour les interactions entre Pin1 et des phospho-épitopes de la protéine tau. Dans les diverses expériences de titration par RMN que nous avons décrites dans le paragraphe précédent, nous montrions que l'ajout de ligand avait pour effet une perturbation des résonances des résidus de Pin1, ou du domaine WW isolé, qui se déplaçait d'une forme libre, en l'absence de ligand, vers une forme liée, à saturation de ligand. Pourtant, l'effet observé n'était pas le même pour les différentes résonances, traduisant un comportement différent des résidus de la protéine vis-à-vis du ligand (Figure 2.10). L'aboutissement de ces expériences de titration a été la détermination des constantes de dissociation des complexes formés entre Pin1 et les phospho-épitopes de tau, c'est-à-dire la détermination d'une constante thermodynamique dans les conditions données.



**Figure 2.10** : Comportement différent des résidus Q33/Y24 et W34 dans l'interaction du domaine WW avec un phospho-peptide de la protéine tau. Les ajouts consécutifs de substrat entraînent une variation des déplacements chimiques graduelle pour les résidus Q33 et Y24 (à gauche) et une disparition du signal de W34 par élargissement de raie puis une réapparition à l'approche de la saturation (à droite). Si on considère la dimension proton, dans le premier cas la variation entre les déplacements chimiques des états libre et lié est 0.1 ppm et dans le deuxième cas supérieure à 0.3 ppm.

L'effet de l'interaction sur la largeur de raie a été décrit d'un point de vue théorique par les équations de Bloch modifiées qui introduisent deux termes relatifs à l'échange chimique, en plus des termes relatifs à la fréquence de rotation de Larmor ( $\Omega$ ) et de relaxation  $T_2$ . Dans le cas le plus simple du passage d'un état A à un état B, représentant les formes libre et liée d'une protéine ou deux conformations d'une même molécule, les équations de Bloch modifiées s'écriront pour la magnétisation macroscopique M de la manière suivante :



$$\begin{aligned} \frac{d}{dt} M_x^A &= -\frac{1}{T_2^A} M_x^A + \Omega^A M_y^A - k_{AB} M_x^A + k_{BA} M_x^B & \frac{d}{dt} M_x^B &= -\frac{1}{T_2^B} M_x^B + \Omega^B M_y^B + k_{AB} M_x^A - k_{BA} M_x^B \\ \frac{d}{dt} M_y^A &= -\frac{1}{T_2^A} M_y^A - \Omega^A M_x^A - k_{AB} M_y^A + k_{BA} M_y^B & \frac{d}{dt} M_y^B &= -\frac{1}{T_2^B} M_y^B - \Omega^B M_x^B + k_{AB} M_y^A - k_{BA} M_y^B \end{aligned}$$

où les termes concernant l'échange chimique dérivent de la définition des concentrations de A et B au cours du temps :

$$\frac{d}{dt} [A] = -k_{AB}[A] + k_{BA}[B] \quad \frac{d}{dt} [B] = +k_{AB}[A] - k_{BA}[B]$$

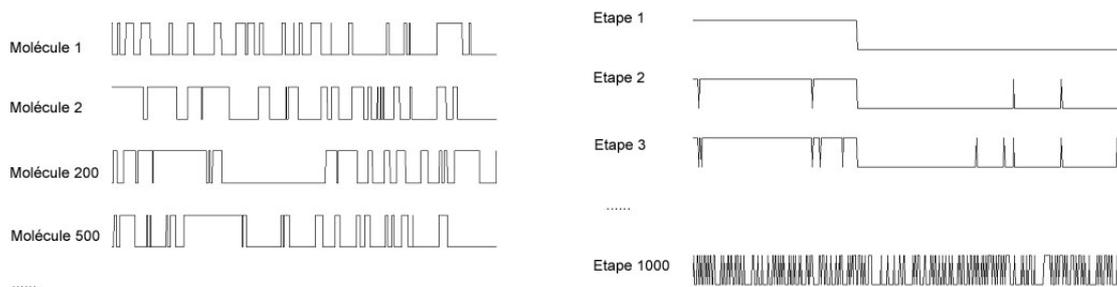
Les constantes cinétiques  $k_{AB}$  et  $k_{BA}$ , qui sont les constantes  $k_{on}$  et  $k_{off}$  dans le cas de figure d'un échange association/dissociation entre deux molécules, sont introduites dans ce système d'équations différentielles et la résolution analytique de celles-ci peut conduire à l'expression des constantes cinétiques qui pourront alors être appliquées à la caractérisation dynamique de différents systèmes. Mais le terme de magnétisation introduit ici est un terme global concernant l'ensemble moyen des molécules d'un échantillon, autrement dit il s'agit d'un paramètre macroscopique, qui ne reflète pas la situation de chaque molécule prise individuellement dans un échantillon. Dans le cas particulier de notre étude impliquant la liaison d'un ligand par une protéine, les données RMN traduisent l'état d'équilibre pour lequel, macroscopiquement, les proportions des formes libres et liées de la protéine sont invariantes au cours du temps. En revanche, à l'échelle des molécules individuelles, une molécule isolée n'est pas toujours dans un état libre ou lié mais peut passer d'un état à l'autre, dans la mesure où une autre molécule fait le chemin inverse de manière à assurer l'équilibre

macroscopique. C'est la définition du micro-équilibre ou équilibre microscopique. La simulation du comportement des molécules individuelles au cours de l'échange, sur une petite population de molécules, permettra de retrouver la largeur de raie expérimentale observée sur les spectres RMN. La simulation de l'échange chimique repose sur une décomposition du processus en deux temps :

- La simulation du comportement des molécules dans le temps, au cours d'un processus d'échange, sous forme de « trajectoires » avec l'utilisation de l'algorithme de Monte Carlo.
- La traduction des trajectoires en fréquence et la simulation d'un spectre RMN.

### ***Génération des trajectoires - Algorithme de Monte Carlo.***

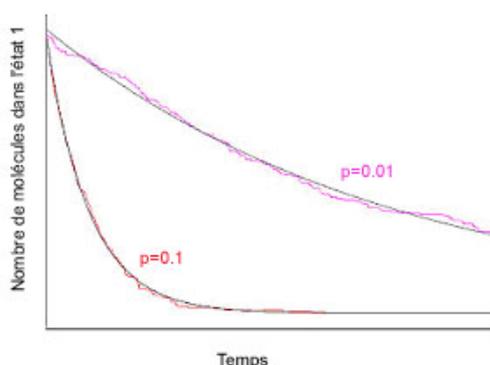
L'échange entre deux états, libre et lié, sera simulé par un saut entre un état « 0 » (état libre) et un état « 1 » (état lié) pour chacune des molécules de l'échantillon. Pour des raisons de capacités de calcul, nous limiterons le nombre de molécules  $m$  de l'échantillon à 1000 molécules. La simulation des « trajectoires » (représentées par une matrice) pour chacune des molécules dans le temps nécessite une décomposition en étapes  $n$ , qui seront limitées pour les mêmes raisons à 1000 étapes. La transition d'un état à l'autre au cours d'une étape sera régie par une probabilité de transition  $p$  fixée qui sera comparée à une valeur aléatoire  $r$ . On considérera que si  $r$  est inférieure à  $p$ , les molécules passeront l'état 1 à l'état 0 et que si  $r$  est supérieure à  $p$ , les molécules restent dans l'état 1. Initialement, on définit un nombre  $f$  de molécules dans l'état 1 et  $(m-f)$  molécules dans l'état 0,  $f$  étant calculé à partir de la valeur expérimentale de la constante de dissociation qui permet de définir, à l'équilibre, le nombre de molécules en complexe et le nombre de molécules libres en fonction des concentrations initiales des deux partenaires mis en présence. Pour assurer le micro-équilibre, il faut s'assurer que, pour  $x$  molécules qui passent de l'état 1 à l'état 0 au cours d'une étape,  $x$  molécules passent de l'état 0 à l'état 1, en comparant, à l'étape 2, la somme des  $f$  premières colonnes de la matrice par rapport à la valeur  $f$ , la différence donnant le nombre de molécules qui seront choisies aléatoirement parmi les  $(m-f)$  molécules dans l'état 0 qui feront la transition vers l'état 1. De cette manière, il aura toujours  $f$  molécules dans l'état 1 et  $(m-f)$  molécules dans l'état 0, assurant l'équilibre. La décision de faire ou non une transition est effectuée de manière itérative et aboutira au remplissage d'une matrice « trajectoires » de taille  $[m,n]$  avec  $m$ , le nombre total de molécules et  $n$ , le nombre d'étapes, chaque valeur (0 ou 1) indiquant l'état de chaque molécule à chaque étape (Figure 2.11).



**Figure 2.11** : Représentation schématique du passage de l'état 1 à l'état 0 et inversement (à gauche) pour une molécule donnée au cours de différentes étapes de la simulation et (à droite) à une étape donnée pour l'ensemble des molécules de l'échantillon.

### Définition de la probabilité de transition $p$ - introduction d'une échelle de temps.

La probabilité de transition  $p$  va gouverner le nombre de molécules qui vont faire la transition d'un état à l'autre à chaque étape de la simulation. Par conséquent, cette valeur va être directement liée à la vitesse de dissociation du complexe. Plus la probabilité de transition  $p$  est élevée, plus le nombre de molécules faisant la transition de l'état 1 à l'état 0 sera important donc plus rapide sera l'échange. Si, initialement,  $f$  molécules sur les  $m$  molécules totales de l'échantillon sont dans l'état 1 et si on passe outre la condition du micro-équilibre, au bout d'un certain laps de temps toutes les molécules vont irrémédiablement se retrouver dans l'état 0 selon une courbe qui décroît de manière mono-exponentielle. Si on simule cette situation, on s'aperçoit que le temps que vont mettre les molécules pour se trouver toutes dans l'état 0 dépend de la probabilité de transition  $p$  selon l'équation  $f_1(t) = f e^{-(1/p)t}$  où  $f_1$  est le nombre de molécules dans l'état 1. Plus  $p$  sera élevée, plus vite l'ensemble des molécules atteindra l'état 0 (Figure 2.12).

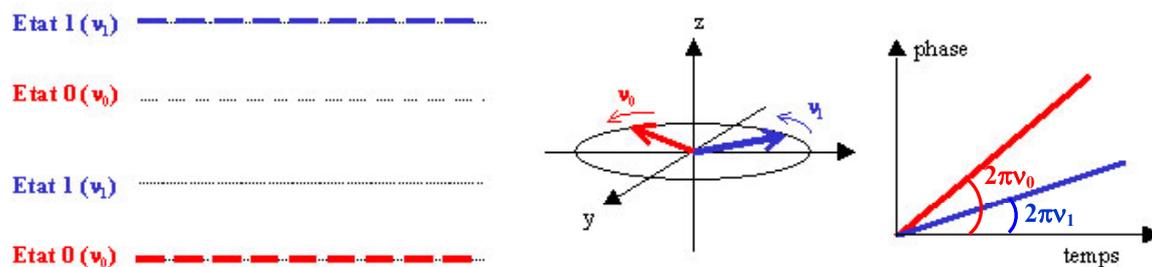


**Figure 2.12** : Variation de la probabilité de saut d'un état à l'autre théorique (courbes noires) et expérimentales en rose pour  $p=0,01$  et en rouge pour  $p=0,1$ .

### **Simulation du signal RMN.**

A chaque état, on associe une fréquence de résonance :  $\nu_0$  pour l'état 0 et  $\nu_1$  pour l'état 1. La relation entre la phase accumulée par le vecteur de magnétisation, qui est le paramètre directement impliqué dans la FID, et la fréquence est donnée par la relation linéaire (Figure 2.13) :

$$\Phi = 2\pi \nu \Delta t.$$



**Figure 2.13 :** Evolution de la phase en fonction du temps pour deux molécules données qui adopteraient deux comportements limites : soit toujours dans l'état 1, associé à la fréquence  $\nu_1$  (en bleu), soit toujours dans l'état 0, associé à la fréquence  $\nu_0$  (en rouge).

La Figure 2.13 illustre deux comportements limites de deux molécules, l'une qui serait toujours dans l'état 1, l'autre toujours dans l'état 0. La plupart des molécules alternent d'un état à l'autre au cours des différentes étapes et leur phase sera intermédiaire entre ces deux phases limites mais ces deux cas extrêmes ne sont pas moins probables.

Le calcul des spectres des molécules individuelles nécessite l'introduction de deux paramètres :

- l'intervalle de temps entre deux étapes
- l'intervalle d'échantillonnage du signal RMN.

L'intervalle entre deux événements à l'échelle moléculaire a été fixé à  $5 \mu\text{s}$ , soit pour une probabilité de 0.01, une transition toutes les  $500 \mu\text{s}$ . Typiquement, l'intervalle d'échantillonnage  $\Delta T$  est calculé à partir de l'inverse de la largeur de la fenêtre spectrale, soit  $200 \mu\text{s}$  pour une fenêtre spectrale de 5000 Hz (environ 8,3 ppm sur un spectromètre à 600 MHz). La phase est par conséquent calculée tous les 40 événements moléculaires. La FID est obtenue à partir du cosinus (partie réelle) et du sinus (partie imaginaire) de la phase et traduite

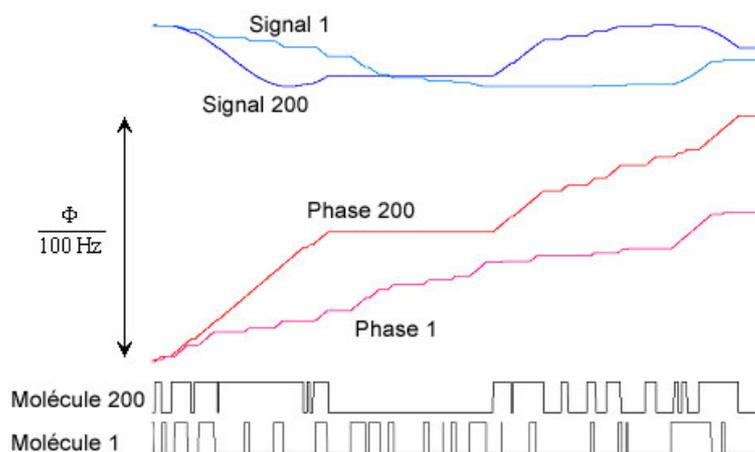
en signal RMN par application de la transformée de Fourier. Comme la transformée de Fourier est linéaire, on peut l'appliquer sur la somme des FID de chacune des molécules ou faire la somme des transformées de Fourier des FID de chacune des molécules.

### **Illustration pour une trajectoire donnée de l'impact de la différence de fréquence $\Delta\nu$ sur la largeur du signal RMN.**

Pour simplifier, nous avons pris, comme valeurs de fréquence, la valeur 0 pour l'état 0 et  $\Delta\nu$  pour l'état 1,  $\Delta\nu$  étant la différence de fréquence ( $\nu_0 - \nu_1$ ) observée expérimentalement entre l'état libre et l'état lié. Nous avons observé le comportement de deux molécules de l'échantillon. Leurs « trajectoires » ont été transformées en accumulation de phase, échantillonnée tous les 40 événements. Nous avons pris l'exemple de deux molécules : la molécule 200 qui effectue peu de transitions au cours du temps et reste par conséquent dans l'état 0 ou l'état 1 pendant de longues périodes, et la molécule 1, qui effectue des transitions entre les deux états plus fréquemment. L'origine de l'élargissement de raie provient de la différence de comportement des différentes molécules de l'échantillon. Nous avons illustré l'élargissement de raie à partir de deux exemples correspondant à deux valeurs de  $\Delta\nu$  : 100 Hz et 600 Hz.

#### Exemple 1 : $\Delta\nu = 100$ Hz

La molécule 200 changeant moins fréquemment d'état, la phase de la molécule 200 sera caractérisée par de longues périodes à la fréquence  $\Delta\nu$ , pour lesquelles la phase évolue au cours du temps avec une vitesse angulaire de  $2\pi\Delta\nu$  pour le vecteur de magnétisation, et de longues périodes à la fréquence 0, pour lesquelles la phase n'évolue pas. La molécule 1 changeant plus fréquemment d'état, la phase aura un comportement plus homogène, correspondant à une fréquence intermédiaire entre les fréquences 0 et  $\Delta\nu$  (Figure 2.14). L'ensemble des signaux de chacune des molécules de l'échantillon représenté en rouge dans la **Figure 2.15** et la somme des FID normalisée en noir illustrent la notion d'élargissement de raie des signaux RMN obtenus après transformée de Fourier (Figure 2.15). La largeur de raie correspond à l'hétérogénéité du comportement des molécules représentant l'échantillon.



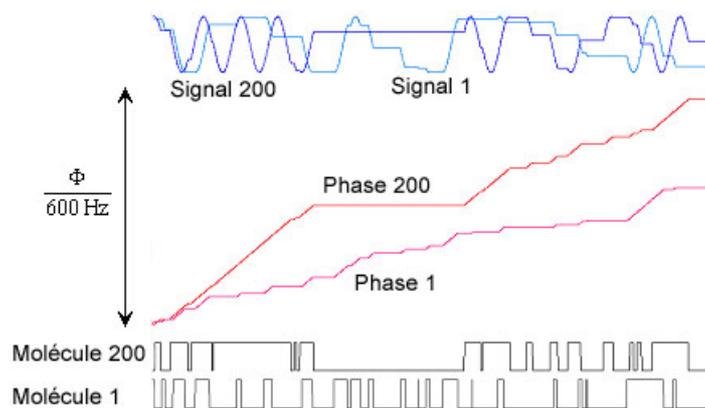
**Figure 2.14** : Correspondance entre la « trajectoire » de deux molécules, la phase et le signal FID dans le cas où  $\Delta\nu=100$  Hz.



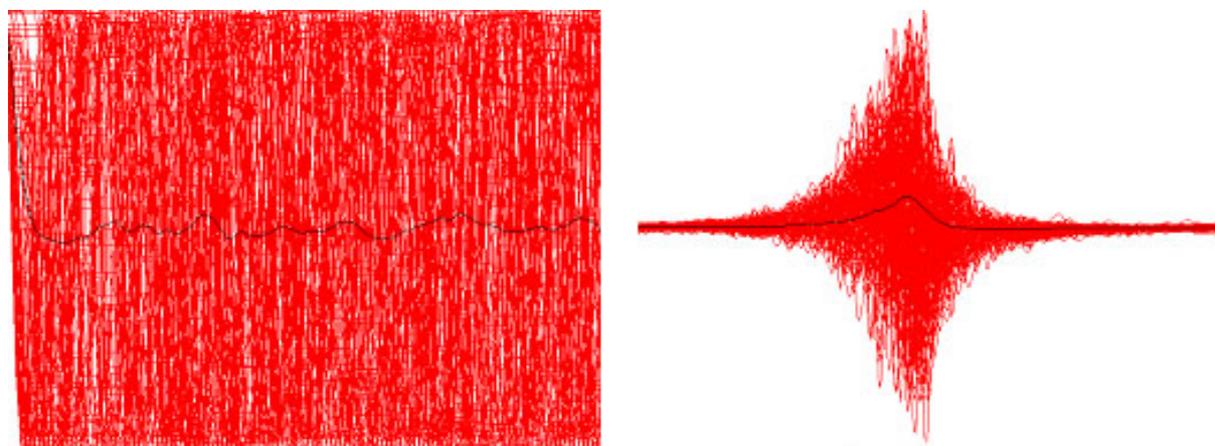
**Figure 2.15** : Somme des FID de toutes les molécules de l'échantillon (à gauche) et spectre RMN après transformée de Fourier (à droite) dans le cas où  $\Delta\nu=100$  Hz.

### Exemple 2 : $\Delta\nu = 600$ Hz

L'élargissement des signaux RMN est d'autant plus marqué que la différence de fréquence  $\Delta\nu$  est grande. La trajectoire des molécules 1 et 200 sont les mêmes que dans l'exemple précédent, mais lorsque les molécules sont dans l'état 1, la phase évolue plus rapidement que dans l'exemple précédent tandis que dans l'état 0, la phase n'évolue pas (Figure 2.16). La différence de fréquence introduit ici une divergence beaucoup plus importants entre les deux signaux au cours du temps. Les molécules représentant l'échantillon auront par conséquent des signaux très disparates qui conduiront à un élargissement de raie très important (Figure 2.17).



**Figure 2.16** : Correspondance entre la « trajectoire » de deux molécules, la phase et le signal FID dans le cas où  $\Delta\nu=600$  Hz.



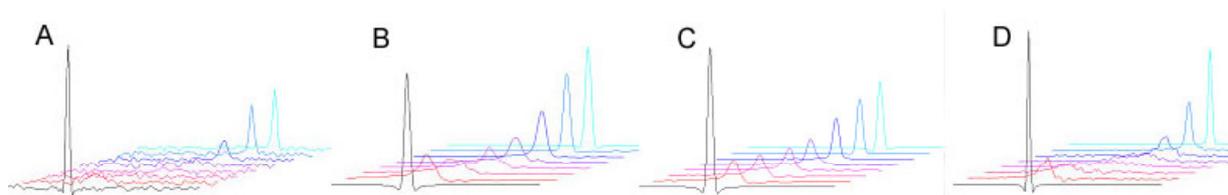
**Figure 2.17** : Somme des FID de toutes les molécules de l'échantillon (à gauche) et spectre RMN après transformée de Fourier (à droite) dans le cas où  $\Delta\nu=600$  Hz.

Cet algorithme peut être appliqué dans le cas d'échanges rapides mais aussi dans le cas d'échanges lents tant que la durée d'un événement moléculaire ne descend pas en-dessous de  $200 \mu\text{s}$  (soit une dissociation toutes les 20 ms), c'est-à-dire qu'il faut toujours qu'il y ait au moins un événement moléculaire dans un intervalle d'échantillonnage.

### ***Comparaison des spectres simulés avec les expériences RMN.***

La simulation a été illustrée sur l'exemple de la formation d'un complexe entre le domaine WW de Pin1 et le peptide de tau portant l'épitope phosphoThr231-Pro232 pour lequel une constante de dissociation d'environ  $300 \mu\text{M}$  a été déterminée, ce qui correspond à 3000 dissociations par seconde si on prend pour la constante de vitesse d'association  $k_{on}$  la valeur de  $10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$ , associée à la diffusion moléculaire. Pour une probabilité de transition de 0.01, cette valeur de  $k_{off}$  correspond à 300 000 événements moléculaires par seconde soit un

événement toutes les 3,33  $\mu\text{s}$  ou encore 60 événements par intervalle d'échantillonnage de 200  $\mu\text{s}$ . Nous avons choisi le signal du Trp34 inclus dans la poche aromatique responsable de la liaison du résidu Pro. La variation de déplacements chimiques pour ce résidu entre l'état libre et l'état lié est de 600 Hz. Nous avons effectué une simulation à partir de ces paramètres et comparé, pour différents ratios  $[\text{protéine}]_0 : [\text{peptide}]_0$ , les spectres protons avec les traces  $^1\text{H}$  obtenues à partir des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC effectués au cours de l'expérience de titration (Figure 2.18). La simulation n'est pas tout à fait satisfaisante dans la mesure où l'élargissement de raie observé expérimentalement n'est pas reproduit à l'identique dans la simulation, nécessitant une optimisation des paramètres cinétiques (Figure 2.18B et C). Une diminution de la vitesse d'association avec un  $k_{on}$  de  $0.5 \cdot 10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$ , et par conséquent de la vitesse de dissociation avec un  $k_{off}$  de 1500 Hz, pour maintenir une constante de dissociation de 300  $\mu\text{M}$ , permet d'obtenir une meilleure corrélation des spectres simulés avec les données expérimentales (Figure 2.18, A et D).



**Figure 2.18 :** Spectre expérimental (A) et spectres simulés (B,C,D) correspondant au proton  $\text{H}_N \epsilon$  de la chaîne latérale du Trp34 impliqué dans la liaison de la proline. Le spectre proton expérimental est issu du spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC dont on a extrait la trace  $^1\text{H}$  pour différents ratios  $[\text{protéine}]_0 : [\text{peptide}]_0$ . Les signaux sont reproduits à partir de la simulation avec un  $k_{on}$  de  $10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$  et un  $k_{off}$  de 3000 Hz dans le cas (B), avec un  $k_{on}$  de  $2 \cdot 10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$  et un  $k_{off}$  de 6000 Hz dans le cas (C), avec un  $k_{on}$  de  $0.5 \cdot 10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$  et un  $k_{off}$  de 1500 Hz dans le cas (D).

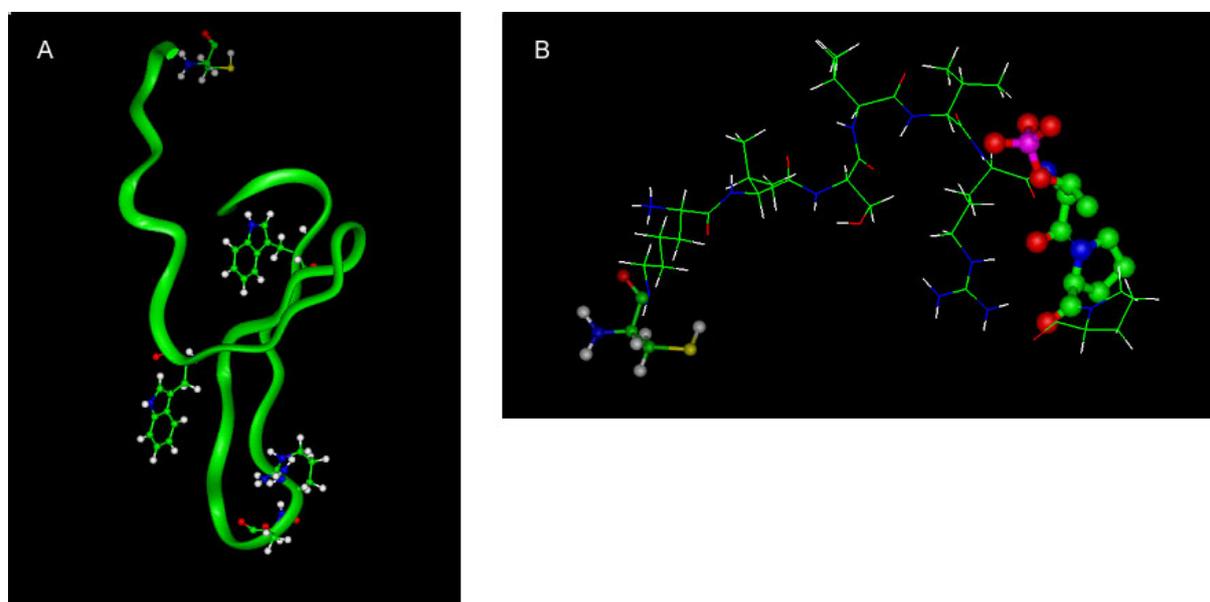
***Un modèle permettant d'évaluer l'influence de la concentration en substrat sur la vitesse d'association. Effet de concentration locale.***

Dans l'étude précédente, nous avons évalué les cinétiques d'association et de dissociation par RMN dans le cas de la formation et de la dissociation d'un complexe hétéro-moléculaire pour lequel la vitesse d'association ne tient compte que de la diffusion moléculaire, processus selon lequel deux molécules se rencontrent « par hasard » au cours du chemin parcouru en solution. La distance moyenne entre deux molécules en solution est définie dans ce cas par la relation suivante, où  $n$  est le nombre de molécules en solution par unité de volume ( $\text{m}^3$ ) :

$$\langle d \rangle = 0.6 n^{-1/3}$$

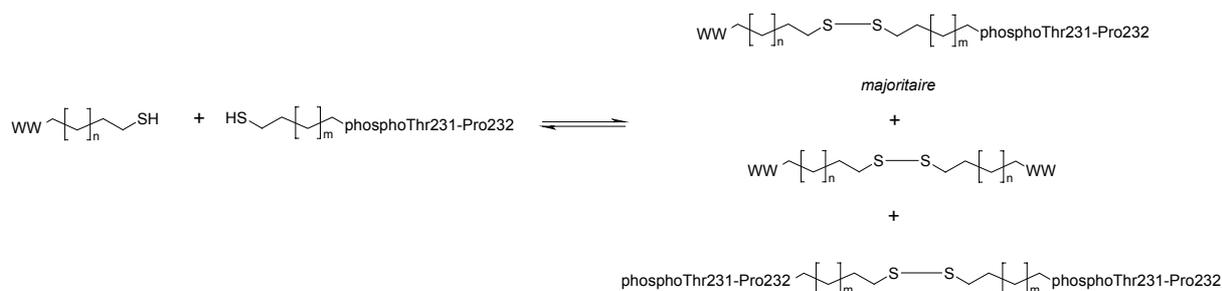
Cette distance moyenne est d'environ 70 Å pour une solution à une concentration de 1 mM. La RMN nous a permis, dans ce contexte, d'étudier la vitesse de dissociation des complexes formés en solution par interprétation des largeurs de raie.

Nous avons voulu montrer, dans ce paragraphe, l'influence de la concentration en substrat sur la vitesse d'association pour la formation de complexes avec un modèle limite où les deux partenaires sont associés de manière covalente et pour lesquels la vitesse de dissociation est la même que dans le cas d'un complexe hétéro-moléculaire « classique ». Dans ce modèle, le ligand de nature peptidique est lié covalamment à la protéine par le biais d'un pont disulfure, ce qui conduit à une augmentation de la concentration locale en substrat autour du site de liaison de la protéine. Cette molécule artificielle a été réalisée à partir du domaine WW isolé qui porte dans sa partie N-terminale, en plus d'une queue fusionnée de six histidines, une séquence additionnelle comportant une cystéine (Figure 2.19A) et d'un peptide de tau centré autour du phospho-épitope Thr231-Pro232 auquel on a ajouté par synthèse en phase solide une cystéine en position N-terminale (Figure 2.19B). La modélisation des deux molécules séparément valide la position des résidus Cys dans les deux molécules qui impose, après formation du pont disulfure, une distance suffisante entre l'épitope phosphorylé et le site de liaison du domaine WW pour ne pas gêner la formation du complexe.

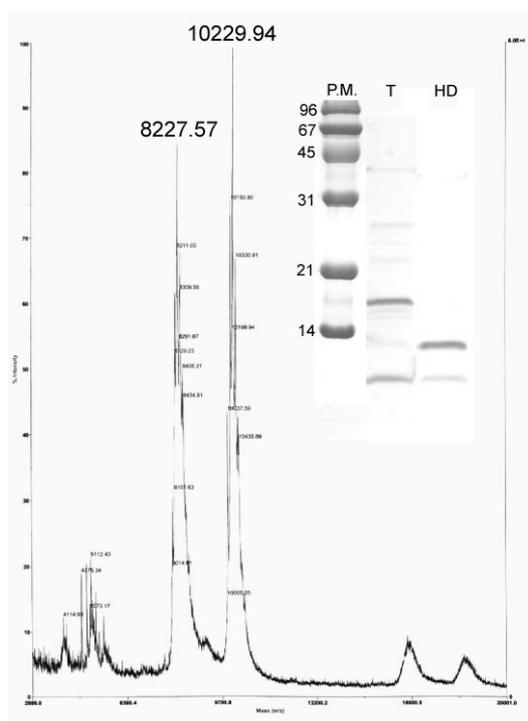


**Figure 2.19 :** (A) structure moléculaire du domaine WW avec représentation « *ball and stick* » des résidus Trp11, Trp34, Ser16 et Arg17 dans le domaine et la Cys en amont du domaine introduite entre la queue de 6 histidines et la protéine. (B) Structure du peptide de Tau centré autour du phospho-épitope Thr231-Pro232 en représentation « *ball and stick* » ainsi que le résidu Cys introduit en position N-terminale.

La formation du pont disulfure est effectuée dans le Tris 50 mM pH=8.0 sans ajout d'agent oxydant. Les conditions basiques favorisant la forme anionique de la fonction thiol des résidus Cys, plusieurs dimères peuvent être obtenus : l'hétéro-dimère WW-S-S-pT231 et les homo-dimères WW-S-S-WW et pT231-S-S-pT231. La formation de l'hétéro-dimère semble être favorisée par rapport à la formation des homo-dimères (Figure 2.20).



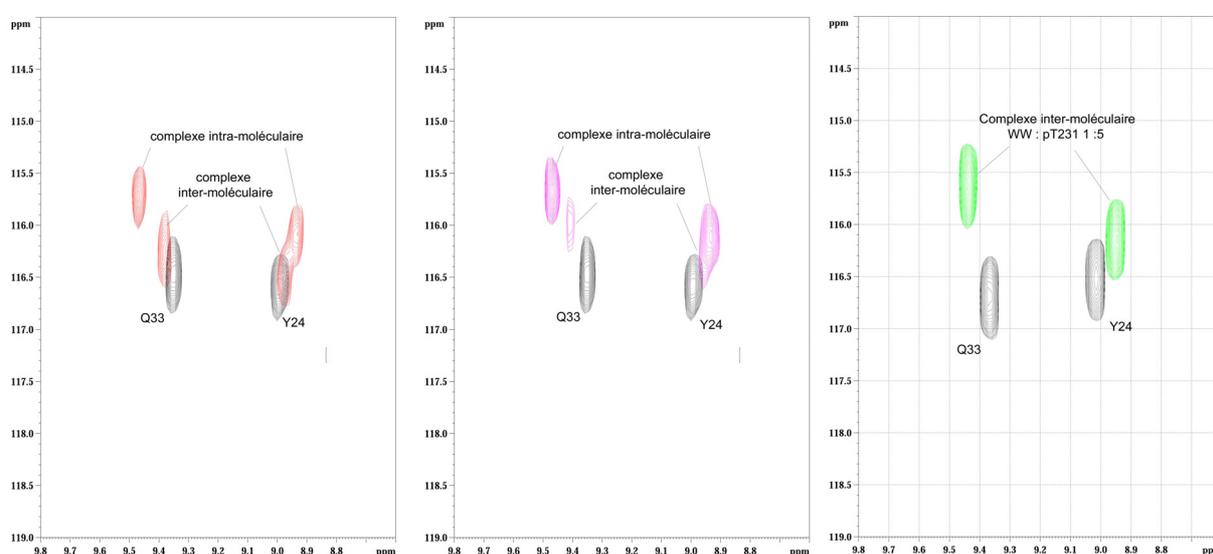
L'association sous forme de complexe inter-moléculaire, dans un premier temps, entre le domaine WW et le peptide, favorise le rapprochement des deux Cys et l'équilibre peut ainsi être déplacé dans le sens de la formation de l'hétéro-dimère.



**Figure 2.20 :** Caractérisation de la réaction d'oxydation en gel dénaturant SDS sans agent réducteur et en spectrométrie de masse MALDI-TOF. P.M., poids moléculaire, T, WW témoin, HD, hétéro-dimère WW-S-S-pT231.

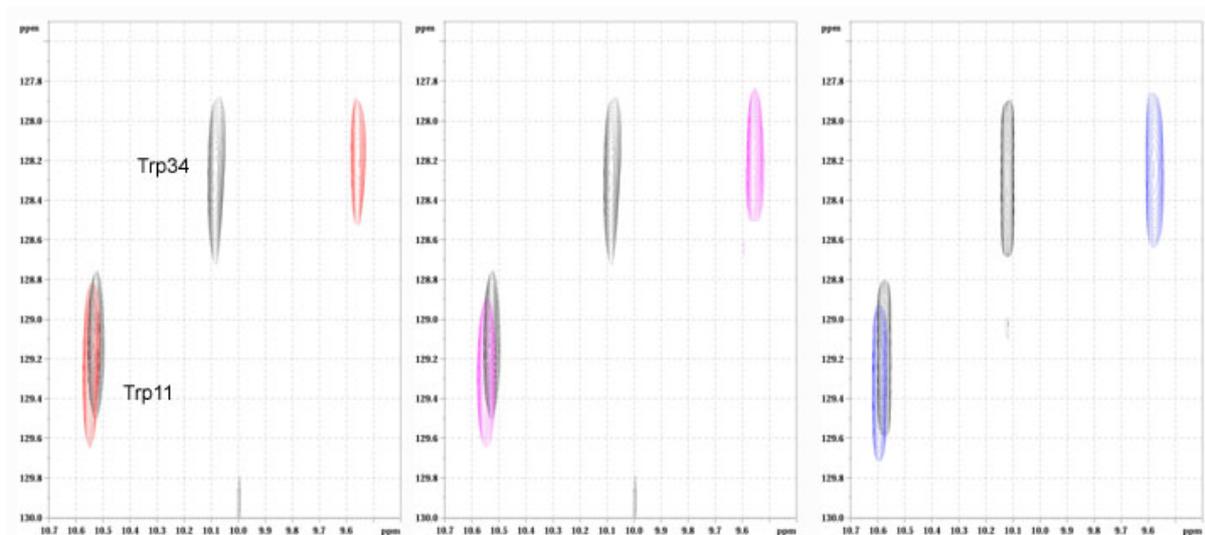
L'acquisition des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC est réalisée sur le domaine WW à 120  $\mu\text{M}$  en présence de 0,5 équiv. et 1 équiv. de peptide  $\text{Cys}_1(\text{pT231})_9$  dans un tampon Tris- $\text{d}_{11}$  50 mM à  $\text{pH}=8.0$ . On observe deux populations sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Figure 2.21) :

1. une population correspondant au(x) complexe(s) intermoléculaire(s) WW-S-S-pT231/pT231 et/ou WW/pT231, dans des ratios  $[\text{WW}]_0 : [\text{peptide}]_0$  de 1 :0.5 et 1 :1 respectivement,
2. une population majoritaire correspondant à l'équivalent d'un complexe intermoléculaire dans un ratio  $[\text{WW}]_0 : [\text{peptide}]_0$  de 1 :5, attribuable au complexe intramoléculaire WW-S-S-pT231.



**Figure 2.21** : Comparaison du complexe WW/pT231 et du complexe WW-S-S-pT231 à travers les spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC au niveau de deux résidus, Q33 et Y24, du domaine WW enrichi en  $^{15}\text{N}$ . Le domaine WW et le peptide pT231 sont mis en présence dans des rapports molaires 1 :0.5 (en rouge) et 1 :1 (en rose). Le spectre du mélange réactionnel d'oxydation brut à  $\text{pH}$  8.0 montre deux formes de complexes : le complexe intramoléculaire et le complexe intermoléculaire entre la protéine et le peptide pT231. Le complexe intermoléculaire où la protéine et le peptide sont mis en présence dans un rapport 1 :5 (en vert) à  $\text{pH}$  6.5 est indiqué à titre comparatif.

Dans le cas du Trp34, on observe pour un ratio 1 :0.5 un complexe semblable au complexe intermoléculaire où la protéine et le peptide sont dans un rapport 1 :20 (Figure 2.22). Des variations dues à un effet de  $\text{pH}$  sont observées entre le mélange réactionnel brut à  $\text{pH}$  8.0 et les spectres effectués lors des titrations classiques à  $\text{pH}$  6.5.



**Figure 2.22 :** Comparaison du complexe WW-S-S-pT231 dans des ratios molaires 1 :0.5 (rouge) et 1 :1 (rose)  $[WW]_0 : [peptide]_0$  à pH 8.0 et du complexe WW/pT231 avec un ratio molaire 1 :20 (bleu) pour le résidu Trp34 à pH 6.5.

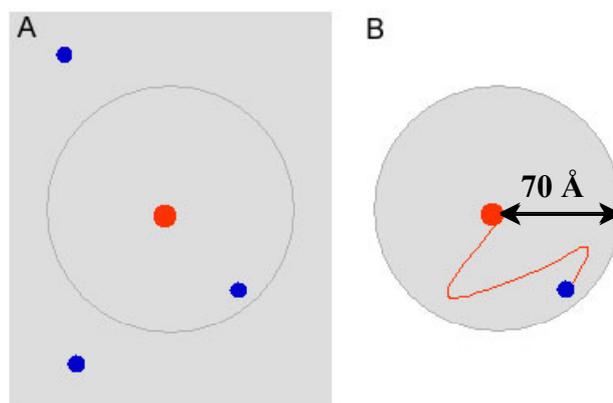
L'hétéro-dimère est un complexe dans lequel la protéine est proche de la saturation (65%) malgré les rapports de concentration  $[WW]_0 : [pT231]_0$  de 1 : 0.5 et 1 : 1 pour lesquels, dans les conditions d'un complexe inter-moléculaire WW/pT231, la protéine est saturée à environ 13% et 24% respectivement. Dans le cas du complexe intra-moléculaire, la proximité spatiale due à la contrainte du linker (21 résidus) maintient une concentration locale en peptide élevée à proximité du site de liaison du domaine WW. Le complexe paraît plus affiné mais le paramètre qui varie ici est la constante cinétique d'association, car la concentration du phospho-épitope phosphoThr231-Pro232 ne correspond plus à la concentration introduite initialement en solution (60 ou 120  $\mu\text{M}$ ) mais à une concentration locale qui s'apparente à une concentration de 1 mM environ d'après les données expérimentales.

$$k_{\text{association}} = k_{\text{on}} [\text{phosphoThr231 - Pr o232}]_{\text{locale}}$$

### ***Calcul de la distance moyenne entre le peptide et le domaine WW.***

- Cas de deux molécules liées covalamment par un fragment peptidique de 21 résidus (Schéma 2.2B).

Si on considère une distance moyenne de 3,4 Å entre deux résidus, la longueur totale du linker de 21 résidus est d'environ 20 fois 3,4 Å soit 70 Å. Le phospho-épitope se situe donc dans une sphère de 70 Å de rayon par rapport au domaine WW, ce qui correspond à un volume de  $1,34 \cdot 10^{-21}$  L et à une concentration de 1 molécule dans un volume de  $1,34 \cdot 10^{-21}$  L.



**Schéma 2.2 :** Comportement d'un peptide en solution (point bleu) par rapport au domaine WW (point rouge) dans le cas (A) de molécules indépendantes et dans le cas (B) de molécules liées de manière covalente.

- Cas de deux molécules libres en solution (Schéma 2.2A).

Pour un ratio 1 :1, soit une concentration en peptide de 120  $\mu\text{M}$ , il y a  $7,2 \cdot 10^{19}$  molécules dans 1L ou une molécule dans un volume de  $1,38 \cdot 10^{-20}$  L, soit un volume 10 fois plus grand que le volume du cas précédent. Cette concentration est 10 fois inférieure à la concentration calculée dans le cas précédent d'un complexe intra-moléculaire entre le domaine WW et le phospho-épitope liée par l'intermédiaire d'un fragment peptidique de 21 résidus. Ce ratio 1 :10 entre la concentration du domaine WW et la concentration en peptide dans la sphère de rayon 70 Å correspond approximativement au ratio observé expérimentalement pour le complexe intra-moléculaire.

Cette technique, où le substrat est fixé covalamment à la cible, est utilisée pour cribler des composés dans la recherche de ligands d'une protéine donnée, notamment dans l'industrie pharmaceutique (Sunesis Pharmaceuticals, San Francisco, USA). Connue sous le nom « *tethering* », cette stratégie permet de cribler rapidement des composés de faible affinité pour une cible pharmacologique donnée en liant par pont disulfure le composé à proximité du site de liaison pour lequel il existe un résidu Cys natif ou obtenu par mutation. La formation de ponts disulfures réversibles entre la protéine-cible et un ou plusieurs composés de bas poids moléculaire en mélange (jusqu'à 1200 composés) permet l'identification de complexes stables par spectrométrie de masse (Erlanson, Braisted et al. 2000). Cette méthode a été développée et utilisée dans la recherche de ligands pour des cibles dont aucun ligand n'est connu, pour permettre l'isolation de composés de faible affinité par une première étape de criblage et ensuite l'obtention, par pharmacomodulation, de composés plus affins (Braisted, Oslob et al. 2003; Hyde, Braisted et al. 2003; Arkin and Wells 2004).

Par ailleurs, notre intérêt pour les problèmes de concentration locale n'est pas anodin. Les mécanismes que nous observons dans nos tubes RMN, où les macromolécules sont solubilisées dans des solutions diluées, seules ou en présence d'un, voire deux partenaires, et que nous décrivons en détail à l'échelle atomique, sont probablement très éloignés de la réalité cellulaire (Ellis and Minton 2003). Le cytoplasme contient de nombreuses macromolécules présentes à hautes concentrations et qui occupent entre 5 et 40% du volume (soit plus de 400g/L). Par conséquent, le comportement d'une protéine dans des conditions expérimentales classiques, en quelques sortes isolée, peut être modifié de manière significative par rapport à un tel environnement (*crowding*) susceptible d'influencer de manière importante nombre de processus biochimiques tels que la stabilité des protéines et leur repliement, la régulation des interactions protéine-protéine ou protéine-ADN, la régulation des assemblages supramoléculaires, ainsi que les interactions avec des drogues potentielles. Les concentrations cellulaires en protéines peuvent également favoriser des interactions protéine-protéine qui mènent vers des processus pathologiques d'agrégation, que l'on retrouve souvent associés aux maladies liées à l'âge. Une explication possible au phénomène observé est la diminution du volume d'eau libre disponible dans la cellule, qui entraîne une augmentation de la concentration effective des molécules et un changement de la vitesse de diffusion due à une augmentation de la viscosité du milieu. Avec l'âge, on constate effectivement une diminution du volume cellulaire et un ralentissement des mécanismes de dégradation. Ces deux facteurs contribuent à une augmentation des concentrations protéiques cellulaires qui peuvent favoriser les phénomènes d'agrégation tels qu'ils sont observés dans certaines maladies neurodégénératives liées à l'âge, comme la formation de fibres d' $\alpha$ -synucléine associée à la maladie de Parkinson, la formation des NFT (*neurofibrillary tangles*) constitués de protéines tau et des dépôts amyloïdes constitués de peptides A $\beta$ , associés à la maladie d'Alzheimer. Des données expérimentales ont montré que le phénomène de *crowding* pouvait modifier significativement les constantes de dissociation des complexes et accélérer les processus d'agrégation (Uversky, Cooper et al. 2001; Shtilerman, Ding et al. 2002). D'autre part, la distribution des protéines dans les différents compartiments intracellulaires peut engendrer des concentrations locales élevées à des endroits précis de la cellule, dont la mesure des concentrations cellulaires globales ne rend pas compte (Ovadi and Saks 2004). La formation des complexes macromoléculaires, caractérisés par l'association étroite de plusieurs macromolécules, est également synonyme de concentrations locales élevées et permet à une molécule de passer directement d'un partenaire à l'autre dans le complexe sans passer par l'état libre en solution (Sontag, Nunbhakdi-Craig et al. 1999; Ovadi and Saks 2004).

### 1.3.3 – La fluorescence, une autre technique spectroscopique pour la détermination des constantes de dissociation en solution.

Les constantes de dissociation déterminées par RMN des différents phospho-épitopes de la protéine tau en complexe avec la peptidyl-prolyl *cis/trans* isomérase Pin1 sont en contradiction sur plusieurs points avec les résultats publiés par Lu et al. (Lu, Wulf et al. 1999) : la spécificité pour l'épitope T231 phosphorylé n'existe pas à l'échelle de substrats peptidiques et l'affinité mesurée est différente de plusieurs ordres de grandeur. Afin de confirmer nos résultats obtenus par RMN, nous avons reproduit les titrations du domaine WW isolé par les différents phospho-peptides en utilisant la fluorescence intrinsèque de la protéine ainsi que le FRET (*Fluorescence Resonance Energy Transfer*) qui est une méthode plus largement utilisée dans la quantification des interactions biomoléculaires.

#### 1.3.3.1 – La fluorescence intrinsèque appliquée au domaine WW isolé, une information relative à l'environnement chimique.

Les résidus aromatiques présents dans les protéines comme le tryptophane, la tyrosine et la phénylalanine sont responsables de la fluorescence intrinsèque des protéines. La fluorescence intrinsèque est un paramètre qui peut être utilisé pour détecter des changements structuraux au sein des protéines. La fluorescence d'une protéine repliée est fonction de la fluorescence des résidus aromatiques individuels. La majorité de l'émission de fluorescence est due aux résidus Trp, et à une très faible contribution des résidus Tyr et Phe. De plus, il peut y avoir transfert de fluorescence des résidus Tyr et Phe aux Trp quand ils sont suffisamment proches dans la structure.

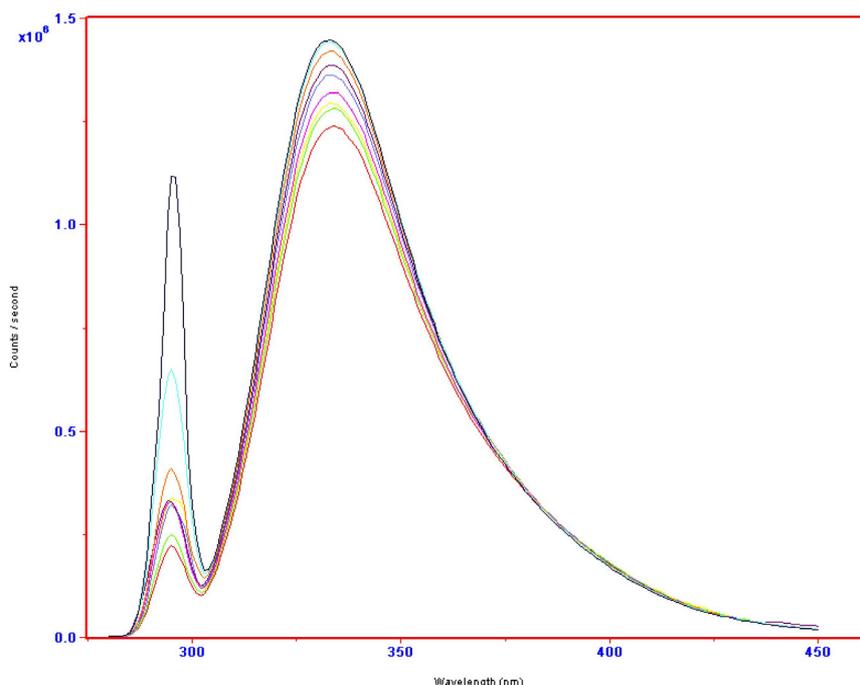
	<i>Durée de vie</i> ( <i>ns</i> )	<i>Absorption</i>		<i>Fluorescence</i>	
		Longueur d'onde (nm)	Absorption molaire	Longueur d'onde (nm)	Rendement quantique
Trp	2.6	280	5,600	348	0.20
Tyr	3.6	274	1,400	303	0.14
Phe	6.4	257	200	282	0.04

**Tableau 2.5** : Caractéristiques d'absorption et de fluorescence des trois résidus aromatiques Trp, Tyr et Phe.

La fluorescence des résidus aromatiques varie d'une protéine à l'autre en fonction de l'environnement chimique des résidus fluorescents comme l'exposition au solvant par exemple. En effet, l'émission de fluorescence (intensité, rendement quantique et longueur d'onde) du tryptophane est très sensible à la polarité du solvant. Par conséquent, l'intensité

d'émission de fluorescence varie mais également le maximum d'émission, en fonction de l'environnement plus ou moins hydrophobe des Trp dans la protéine. Par exemple, un tryptophane exposé à la surface d'une protéine aura un maximum de fluorescence entre 340 et 350 nm, alors que lorsqu'il est totalement enfoui dans le cœur hydrophobe, il émet à 330 nm. De manière générale, ce n'est pas l'intensité et la longueur d'onde du signal de fluorescence qui importe mais les variations de ce même signal en fonction des conditions, en l'occurrence, dans notre cas, la liaison d'un substrat.

La présence d'un phospho-peptide a pour effet d'augmenter l'intensité de fluorescence du domaine WW de Pin1 et de déplacer le maximum de fluorescence vers les longueurs d'onde plus faibles (« *blue-shift* »). En l'absence de substrat, le domaine WW donne un signal de fluorescence à 334 nm après excitation à 295 nm. En présence de 20 excès de substrat, le maximum est déplacé à 330-331 nm. D'autre part, l'augmentation de l'intensité de fluorescence est proportionnelle à la quantité de peptide ajoutée. Une constante de dissociation du complexe formé entre le domaine WW et le peptide substrat peut donc être calculée à partir d'une titration réalisée dans des conditions similaires à celle de la RMN avec 20  $\mu$ M de protéine non enrichie (Figure 2.23). L'augmentation de l'intensité de fluorescence et le « *blue-shift* » peuvent s'expliquer par un accroissement de l'hydrophobicité de l'environnement du Trp34 impliqué dans la liaison du résidu Pro du substrat. L'augmentation de l'intensité de fluorescence mesurée en présence du substrat est de l'ordre de 20% à saturation par rapport au signal de fluorescence du domaine WW en l'absence de substrat.



**Figure 2.23 :** Variation de l'intensité de fluorescence intrinsèque du domaine WW de Pin1 en fonction des concentrations croissantes ajoutées de phospho-peptide pT231. Représentation des spectres d'émission du domaine WW isolé après excitation sélective des résidus Trp à 295 nm pour différentes concentrations initiales en peptide substrat pT231. On observe à la fois une augmentation de l'intensité du maximum d'émission à 334 nm avec l'augmentation de la concentration en peptide et une légère variation de la longueur d'onde du maximum d'émission.

Les constantes de dissociation ont été calculées pour les différents phospho-peptides de tau décrits dans le **Tableau 2.1** et les valeurs de  $K_D$ , résumées dans le **Tableau 2.6**, sont du même ordre de grandeur que celles déterminées par RMN. Les valeurs de l'ordre du millimolaire qui ont été déterminées par RMN, n'ont pu être mesurées par fluorescence. Il semble donc que la détection de la liaison du substrat au domaine WW par fluorescence intrinsèque soit limitée à des valeurs de constantes de dissociation inférieures à 0.5-1 mM. En fluorescence, le signal mesuré est un signal global pour l'ensemble de la protéine alors que la RMN permet d'observer des phénomènes à l'échelle du résidu. Par conséquent, il est possible que des phénomènes concurrents à l'augmentation du signal de fluorescence viennent perturber la mesure, comme la concentration élevée en substrat qui est d'autant plus importante que la constante de dissociation est élevée et qui peut faire « écran » à l'énergie émise par la lampe et reçue par l'échantillon, ainsi qu'à l'énergie émise par l'échantillon et détectée par le photomultiplicateur.

<i>Peptide</i>	$K_D^{WW}$ (mM)
pT231	0.170
pT231pS235	0.100
pS235	N. d.
pT212	0.120
pT212pS214	0.040
pT212pS214pT217	0.030
pS396	N. d.
pS396 pS404	N. d.
pS422	N. d.

**Tableau 2.6 :** Valeurs des constantes de dissociation déterminées par mesure de la fluorescence intrinsèque pour les complexes formés par les phospho-peptides de tau et le domaine WW isolé de Pin1. N. d., non détectable.

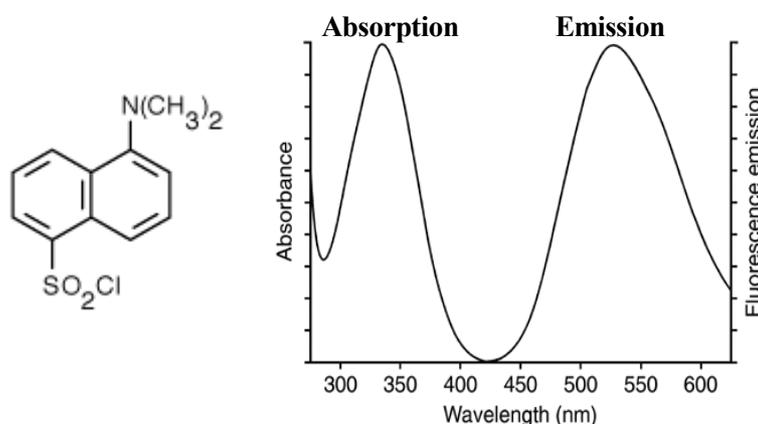
La titration de la protéine Pin1 entière par mesure des variations de l'intensité de fluorescence intrinsèque en présence de substrats peptidiques telle qu'elle est décrite pour le domaine WW n'a pas abouti à des résultats exploitables. La présence d'un troisième tryptophane en position 73 dans la séquence de Pin1 induit un signal d'émission plus intense que celui du domaine WW. Ce résidu, localisé dans le domaine catalytique à proximité des résidus Arg68 et Arg69 du site catalytique, impliqués dans la liaison du phosphate, peut induire des perturbations antagonistes au phénomène relatif à la liaison du substrat, mesuré au niveau du domaine WW. Nous avons donc mis en œuvre une expérience de FRET, qui est une technique couramment employée dans le domaine des investigations des interactions biomoléculaires.

### 1.3.3.2 – Le FRET (*Fluorescence Resonance Energy Transfer*), une méthode plus sensible pour l'étude d'interactions biomoléculaires.

Le transfert de fluorescence en FRET est une interaction fortement dépendante de la distance entre les états excités d'une paire donneur-accepteur pour laquelle l'excitation est transférée du donneur vers l'accepteur sans émission de photons, mais par résonance des orbitales de l'état excité de chacun. L'efficacité du FRET est proportionnelle à  $1/r^6$  où  $r$  est la distance intermoléculaire entre le donneur et l'accepteur. Cette technique est compatible, du point de vue des distances, avec les dimensions des macromolécules d'intérêt biologique et peut donc être mise en œuvre pour la détection de phénomènes impliquant des variations de distances intermoléculaires. Le transfert d'énergie peut avoir lieu sous trois conditions : (i) le donneur et l'accepteur sont suffisamment proches dans l'espace (de 10 à 100Å), (ii) le spectre d'émission de fluorescence du donneur recouvre au moins en partie le spectre d'absorption de l'accepteur

et (iii) les moments dipolaires de la paire donneur-accepteur adoptent une orientation parallèle.

Hormis la présence de fluorophores naturels dans les protéines conduisant à la fluorescence intrinsèque, il est possible de marquer les protéines avec des fluorophores synthétiques produisant une fluorescence dite extrinsèque. Ces marqueurs fluorescents sont d'excellents outils pour les études structurales et d'interactions par FRET. Différentes classes de molécules existent selon la nature de la fonction chimique visée dans la protéine : (1) les marqueurs des fonctions amines primaires, tels que les halogénures d'arylsulfonyle, les iso- et isothiocyanates, les N-succinimides et les anhydrides ; (2) les marqueurs de thiols comme les maléimides, les aziridines et les haloacétamides. Nous avons utilisé un dérivé de type chlorure de sulfonyle couramment employé pour étudier les interactions impliquant des protéines, le chlorure de dansyle, dont le spectre d'absorption recouvre bien le spectre d'émission de fluorescence des tryptophanes dans les protéines (Figure 2.24). Les chlorures de sulfonyle sont des réactifs très instables en milieu aqueux à des pH basiques, qui sont les conditions utilisées pour les réactions de couplage avec les fonctions amines des protéines. Cependant, le sulfonamide obtenu après conjugaison est extrêmement stable, même dans les conditions d'hydrolyse des protéines. Des réactions secondaires avec des fonctions phénol, alcool, thiol ou imidazole sont possibles mais peu probables avec les protéines et en particulier en milieu aqueux. De plus, les dérivés conjugués du chlorure de dansyle ne sont pas sensibles au *photobleaching*, processus au cours duquel la destruction irréversible du fluorophore entraîne une diminution de l'intensité de fluorescence dans le temps.



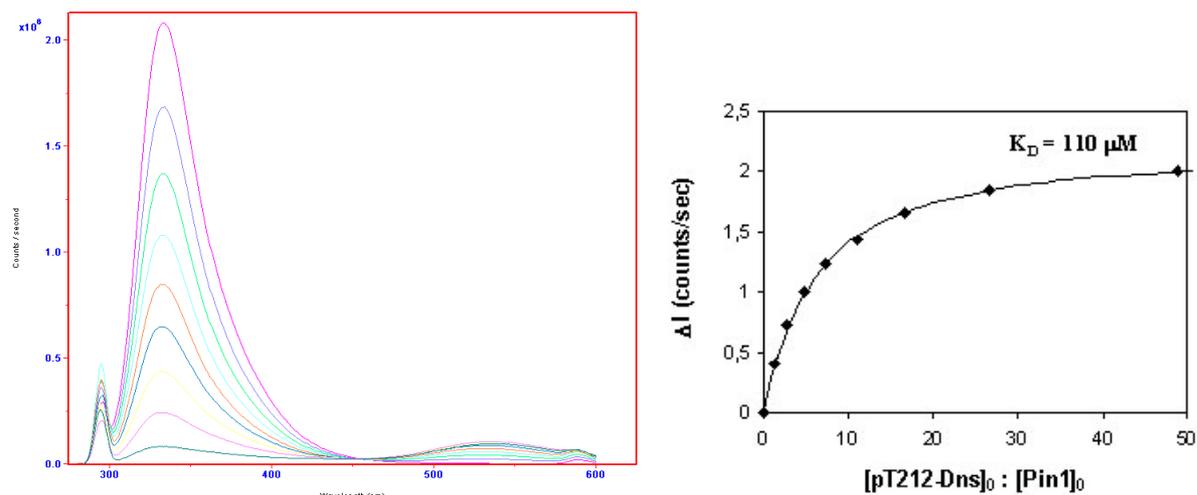
**Figure 2.24 :** Chlorure de 5-(diméthylamino)-naphthalène-1-sulfonyle ou chlorure de dansyle. Structure moléculaire et spectres d'absorption et d'émission dans le méthanol.

Dans notre cas, le marqueur fluorescent est porté par le peptide substrat synthétisé en phase solide. Le couplage du fluorophore peut alors être introduit aisément et sélectivement sur le peptide au cours de la synthèse. La réaction de couplage du chlorure de dansyle est effectuée sur l'extrémité N-terminale du peptide par réaction avec 5 équiv. de chlorure de dansyle et 5 équiv. de DIPEA dans le DMF, et suivie par une étape de *capping*. Le couplage du chlorure de dansyle est quantitatif, en particulier parce que les conditions de réaction en milieu organique limitent l'hydrolyse de la fonction chlorure de sulfonyle.

Les titrations de la protéine Pin1 entière et du domaine WW par le phospho-peptide pT212 portant un groupement dansyle ont été réalisées dans les mêmes conditions expérimentales qui ont été décrites précédemment pour la fluorescence, avec 20  $\mu\text{M}$  de protéine non enrichie. Un phénomène de FRET correspondant au transfert d'énergie entre les Trp du domaine WW, ou de Pin1, et le groupement dansyle porté par le peptide substrat, conduit à deux maxima d'émission de fluorescence à 334 nm et 530 nm correspondant respectivement aux Trp et au dansyle après excitation à 295 nm : on observe à la fois une diminution de l'intensité du maximum d'émission à 334 nm avec l'augmentation de la concentration en peptide et une légère augmentation du maximum d'émission à 530 nm traduisant le phénomène de FRET. Le calcul de la constante de dissociation est réalisé à partir de la courbe de titration représentant la variation de l'intensité d'émission à 334 nm en fonction du rapport des concentrations initiales en peptide et en protéine. Il faut noter que le rendement quantique du dansyle dans l'eau est très faible, ce qui rend difficilement exploitables les variations d'intensité de fluorescence du signal à 530 nm en fonction des quantités de substrat ajoutées (Figure 2.25). La valeur de la constante de dissociation obtenue pour le complexe formé par Pin1 et le peptide pT212 déterminée par FRET est sensiblement la même que celle déterminée par RMN. De même, les valeurs déterminées avec le domaine WW sont homogènes entre le peptide acétylé en N-term. et le peptide dansylé (Tableau 2.7), indiquant que le dansyle n'interfère pas dans la liaison du substrat et ne se lie pas à un autre site de la protéine.

<i>Peptide</i>	$K_D^{\text{Pin1}}$ (mM)	$K_D^{\text{WW}}$ (mM)
pT212-Ac	-	0.120
pT212-Dns	0.110	0.100

**Tableau 2.7 :** Comparaison des constantes de dissociation des complexes formés par le phospho-peptide de tau pT212 et le domaine WW isolé de Pin1 par mesure des variations de la fluorescence intrinsèque (pT212-Ac) et par FRET (pT212-Dns). Comparaison des valeurs de  $K_D$  déterminées par FRET pour la protéine Pin1 entière et le domaine WW isolé.



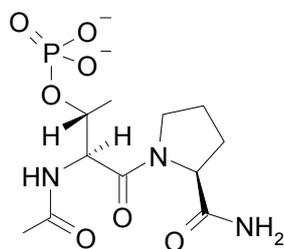
**Figure 2.25 :** Titration de la protéine Pin1 entière par le phospho-peptide pT212 portant une sonde fluorescente de type dansyle en N-term. et observation d'un phénomène de FRET correspondant au transfert d'énergie entre les Trp du domaine WW et le groupement dansyle porté par le peptide substrat. A gauche, spectres d'émission Pin1 après excitation sélective des résidus Trp à 295nm pour différentes concentrations initiales en peptide substrat pT212 avec deux maxima d'émission à 334 et 530 nm correspondant respectivement aux Trp et au dansyle. A droite, la courbe de titration correspondante, représentant la variation de l'intensité d'émission à 334 nm en fonction du ratio des concentrations initiales de peptide et de protéine, et le calcul de la constante de dissociation.

### 1.3.4 – Bilan des affinités des différents phospho-épitopes de la protéine tau dans la liaison à Pin1 obtenues en milieu homogène.

Différents sites de phosphorylation de type Ser/Thr-Pro de la protéine tau, incluant des sites de phosphorylation mitotique et des sites spécifiques des PHF pathologiques, sont des phospho-épitopes reconnus par la peptidyl-prolyl *cis/trans* isomérase Pin1, et plus particulièrement par son domaine WW, responsable de la liaison du substrat. Nous avons évalué l'affinité des phospho-peptides pour la protéine Pin1 entière ou son domaine WW isolé

- par RMN à haute résolution avec l'utilisation de protéines enrichies en  $^{15}\text{N}$  à partir des perturbations des déplacements chimiques des résidus de la protéine en fonction de la concentration en peptide,
- par mesure de la fluorescence intrinsèque du domaine WW isolé et par FRET avec un peptide portant une sonde fluorescente de type dansyle, à partir des variations de l'intensité d'émission de fluorescence à 334 nm, en fonction de la concentration en peptide.

Nous avons observé une gradation dans l'affinité des différents phospho-épitopes pour la protéine Pin1, infirmant la spécificité pour la Thr231 décrite précédemment. Nous avons également synthétisé et testé le dipeptide minimal phospho-Thr-Pro, et déterminé une affinité de 150  $\mu\text{M}$  pour la liaison au domaine WW.



$$K_D^{WW} = 150 \mu\text{M}$$

Ce résultat renforce la non-spécificité des interactions entre Pin1 et tau autour d'un épitope unique. La conception de ligands synthétiques du domaine WW de Pin1, basée sur la structure du motif primaire phospho-Thr-Pro, sera décrite dans le chapitre IV.

#### ***Effet des phosphorylations multiples sur l'interaction avec Pin1.***

La RMN a permis, outre le calcul des constantes de dissociation, d'identifier les sites d'interaction du peptide avec Pin1 par *mapping* des déplacements chimiques. L'interaction majeure avec les différents substrats est assurée par le domaine WW, décrit comme domaine de liaison au même titre que les domaines SH2, SH3,... et médiateur des interactions protéine-protéine (Lu, Zhou et al. 1999). Deux sites sont principalement affectés par la présence des phospho-peptides : les résidus localisés dans la boucle  $\beta$ -hairpin incluant les résidus Ser16-Ser19, responsables de la liaison du résidu phosphorylé et les résidus autour du Trp34 et Tyr23 impliqués dans l'interaction avec la proline en +1 du résidu phosphorylé. Les résidus du domaine catalytique sont peu affectés par le peptide. Cependant, on note des perturbations plus ou moins prononcées selon le substrat, pour les résidus situés à l'interface du domaine WW au niveau des éléments de structure secondaire  $\alpha$ 4 et  $\beta$ 6, traduisant un léger rapprochement des deux domaines de Pin1 lors de la liaison du substrat. D'autre part, les études de *binding* réalisées avec le domaine WW isolé ne montrent pas une différence significative de l'affinité par rapport à la protéine entière pour la liaison des peptides mono-phosphorylés, ce qui semble indiquer une faible participation du domaine catalytique, peut-être même inexistante, dans la liaison des substrats mono-phosphorylés.

Une graduation dans les affinités déterminées pour les différents phospho-épitopes, avec des valeurs allant de 100  $\mu\text{M}$  à 1  $\text{mM}$ , est contradictoire avec la spécificité pour la phospho-Thr231 décrite par Lu et al. L'épitope Thr212, en particulier, est significativement plus affiné que l'épitope Thr231 sur base du modèle peptidique. Dans les PHF, les phospho-épitopes Thr212 et Thr231 étant accompagnés des phospho-épitopes Ser214 et Ser235 respectivement (reconnus par les anticorps AT100 et TG3, respectivement), nous avons déterminé les constantes de dissociation des peptides di-phosphorylés contenant les paires Thr212/Ser214 et Thr231/Ser235. D'après les données obtenues par RMN, l'introduction d'un second site de phosphorylation n'a pas une influence claire sur l'affinité : la phosphoSer214 augmente le  $K_D$  de 60% alors que la phosphoSer235 diminue le  $K_D$  d'un facteur deux environ. Cependant, il faut noter que, dans le premier cas, il s'agit d'un site Ser-Leu qui n'a pas les spécificités requises pour interagir avec Pin1. D'autre part, dans la structure de Pin1 avec le peptide di-phosphorylé CTD de l'ARN Polymérase II (Verdecia, Bowman et al. 2000), un résidu Ser en C-term. du motif phosphoThr-Pro lié par le domaine WW établit une liaison hydrogène avec l'indole du Trp34. Nous avons dans un premier temps, sur base des affinités obtenues par RMN sur la protéine Pin1 entière, émis l'hypothèse que, dans le cas du peptide pT212, la Ser214 pourrait expliquer une affinité plus élevée qu'avec le peptide pT231 pour lequel la Pro233 n'offre pas de possibilité de liaison hydrogène. Dans le cas du peptide pT212 pS214, la phosphorylation de la Ser214 pourrait empêcher que la liaison hydrogène s'établisse, provoquant ainsi une diminution de l'affinité de 100 à 160  $\mu\text{M}$ . On retrouve ensuite l'affinité de 100  $\mu\text{M}$  avec une phosphorylation supplémentaire en Thr217. Dans le cas de la phosphorylation en Ser235, incluse dans un motif Ser-Pro, la présence de deux motifs d'interaction entraîne une augmentation de la concentration effective en substrat d'un facteur deux vis-à-vis du domaine WW. C'est-à-dire que, si on considère une simple addition des sites d'interaction, 1 mole de peptide pT231 pS235 correspondrait à 1 mole de pT231 et 1 mole de pS235 (en concentration initiale) se liant au domaine WW dans des proportions relatives à leurs affinités respectives. Le site phosphoSer235 représente à lui seul un  $K_D$  de 900  $\mu\text{M}$ , soit 2,4 fois celui de la phosphoThr231 seule. Donc l'effet d'une mole de pS235 correspondrait à l'effet de 0,42 mole de pT231 sur la liaison au domaine WW, ce qui est équivalent à 1,42 mole de pT231 au total pour 1 mole de pT231 pS235, donc une constante de dissociation théorique pour ce dernier de 270  $\mu\text{M}$  environ. La valeur déterminée expérimentalement étant de 180  $\mu\text{M}$ , on peut estimer qu'une simple augmentation de la concentration effective en sites de liaison est responsable de l'augmentation d'affinité pour le

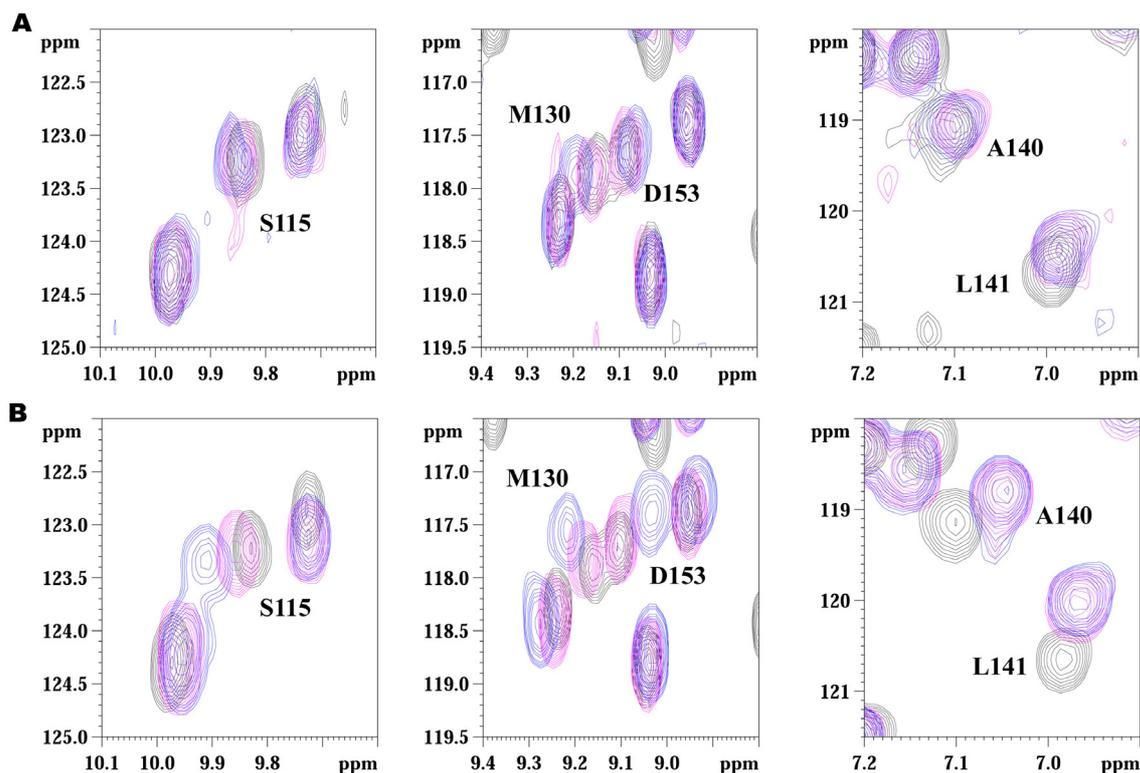
peptide di-phosphorylé pT231 pS235, et qu'aucune coopération n'intervient dans la liaison du substrat en présence d'un deuxième site de phosphorylation dans le contexte d'un substrat peptidique.

La première hypothèse concernant l'effet de la phosphorylation en Ser214 a été remise en cause par les données obtenues avec le domaine WW sorti de son contexte protéique. En effet, dans le cas du domaine WW isolé, la phosphorylation de la Ser214 induit une augmentation de l'affinité par rapport au peptide phosphorylé uniquement en Thr212. Cette augmentation de l'affinité pourrait être expliquée par un renforcement de la liaison hydrogène entre la Ser214 et l'indole du Trp34 quand celle-ci est phosphorylée, ce qui contredirait l'hypothèse préalablement énoncée à partir des seules données RMN. En ce qui concerne la diminution de l'affinité observée avec Pin1 entière, elle pourrait être expliquée par des interactions défavorables avec le domaine catalytique, ou une interaction plus forte au niveau du site catalytique diminuant la disponibilité de substrat pour le domaine WW (Figure 2.26).

#### ***Détail des interactions avec le peptide di-phosphorylé pT212 pS214.***

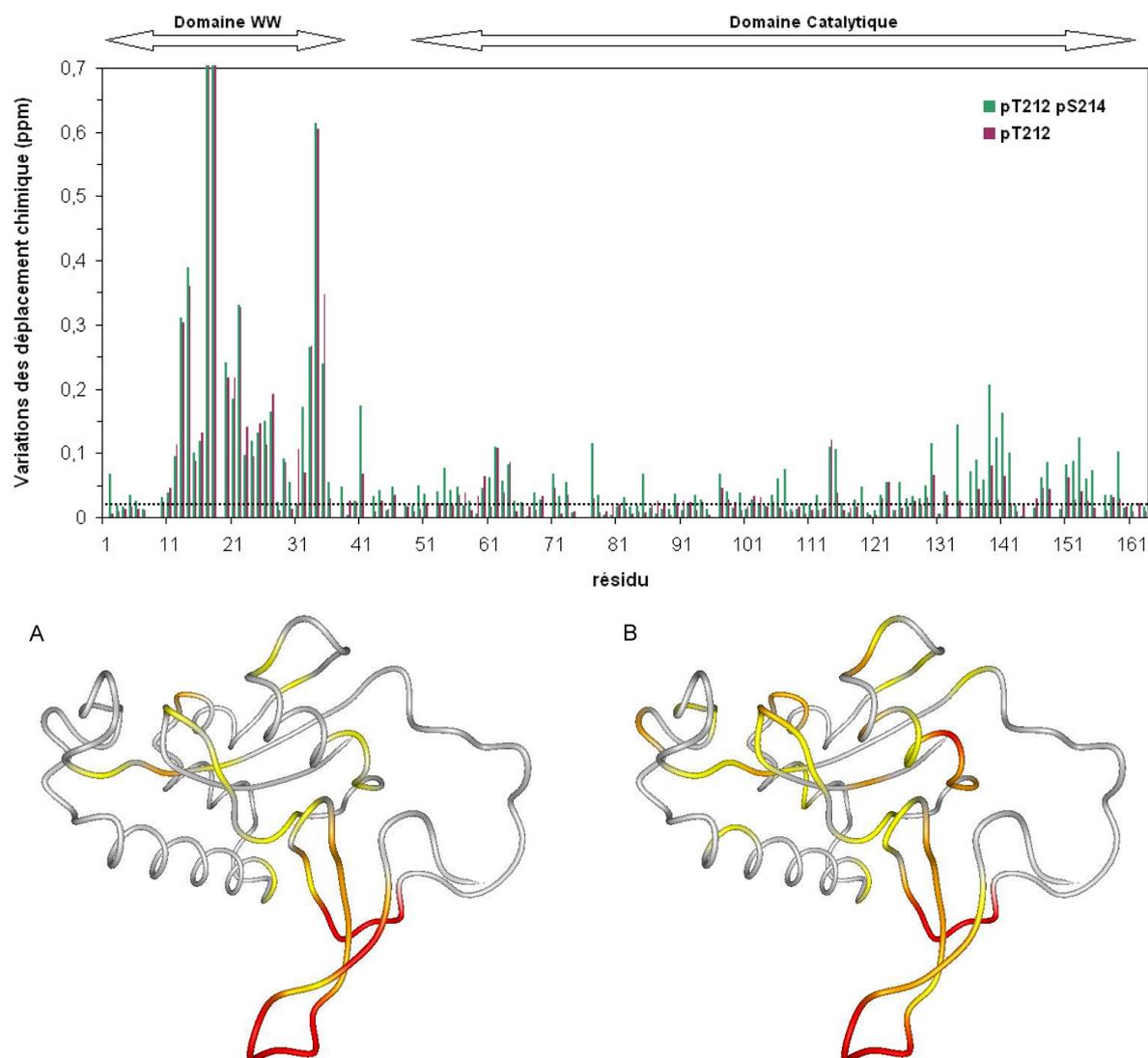
Globalement, les variations d'affinité observées entre les peptides mono- et di-phosphorylé, contradictoires entre Pin1 entière et le domaine WW, indiquent un rôle négatif du domaine catalytique sur la liaison du peptide di-phosphorylé. Les perturbations observées au niveau des résonances du domaine catalytique en présence du peptide di-phosphorylé en Thr212/Ser214 vont également dans ce sens. Même si ces observations sont difficilement interprétables, elles sont suffisamment significatives pour être relevées, en particulier si on compare ces perturbations avec celles des peptides mono-phosphorylé en Thr231 et di-phosphorylé en Thr231/Ser235.

Les résidus du site catalytique sont plus ou moins affectés en fonction du substrat peptidique. Le peptide di-phosphorylé pT212 pS214 induit des variations de déplacements chimiques plus marquées au niveau des résidus du site catalytique comme les résidus Ser114/115 et Met130 ainsi que des perturbations plus importantes des résidus à l'interface du domaine WW (Phe139-Leu141) par rapport au peptide mono-phosphorylé pT212 (Figure 2.26).



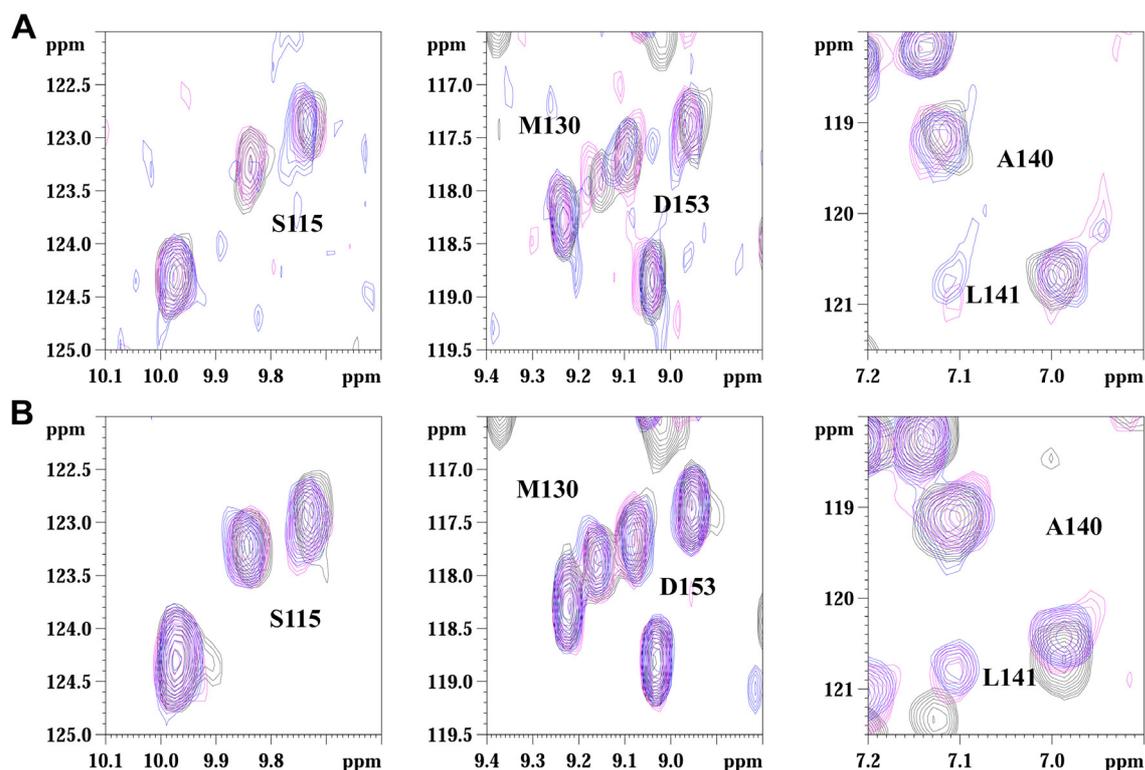
**Figure 2.26 :** Détails des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 enrichie en  $^{15}\text{N}$  avec des concentrations croissantes en peptides pT212 (A) et pT212 pS214 (B). En noir, la protéine seule à 200  $\mu\text{M}$  environ, en rose, en présence de 5 équiv. de peptide, en bleu, en présence de 20 équiv. de peptide. Les régions représentent les résidus S115 (gauche) et M130/D153 (milieu) localisés dans le site catalytique de Pin1 et les résidus A140/L141 du site catalytique de l'interface avec le domaine WW (droite).

Plus généralement, les perturbations engendrées par le peptide di-phosphorylé sont plus importantes que celles engendrées par le mono-phosphorylé et concernent essentiellement l'interface entre les deux domaines (Figure 2.27). Ces modifications infimes sont difficiles à expliquer mais elles suggèrent un rapprochement plus marqué des deux domaines en présence du peptide pT212 pS214, d'après l'étude structurale de Jacobs et collaborateurs (Jacobs, Saxena et al. 2003).



**Figure 2.27 :** *Mapping* comparés des peptides mono- et di-phosphorylés au niveau des sites Thr212 et Ser214. Le peptide di-phosphorylé pT212 pS214 (B) induit davantage de perturbations de déplacements chimiques par rapport au peptide mono-phosphorylé pT212 (A), notamment au niveau de l'hélice  $\alpha 4$  et du feuillet  $\beta 6$  traduisant un rapprochement plus marqué entre les deux domaines de Pin1.

Les effets observés avec la phosphorylation de la Ser214 dans le peptide pT212 pS214 ne sont pas reproduits avec la phosphorylation de la Ser235 dans le peptide di-phosphorylé pT231 pS235 que ce soit au niveau du site catalytique ou de l'interface inter-domaine, ce qui suggère que la phosphoSer235 ne joue pas le même rôle que la phosphoSer214 dans l'interaction avec Pin1. La deuxième phosphorylation n'est pas à la même position par rapport à la première phosphorylation dans les deux peptides ce qui pourrait expliquer la différence observée.



**Figure 2.28 :** Détails des spectres <sup>1</sup>H-<sup>15</sup>N HSQC de Pin1 enrichie en <sup>15</sup>N avec des concentrations croissantes en peptides pT231 (A) et pT231 pS235 (B). En noir, la protéine seule à 200 μM environ, en rose, en présence de 5 équiv. de peptide, en bleu, en présence de 20 équiv. de peptide. Les régions représentent les résidus S115 (gauche) et M130/D153 (milieu) localisés dans le site catalytique de Pin1 et les résidus A140/L141 du site catalytique de l'interface avec le domaine WW (droite).

#### 1.4 – Techniques de mesure de l'affinité des phospho-épitopes de tau pour Pin1 en milieu hétérogène.

L'avantage des techniques en phase hétérogène réside en une séparation plus facile des deux partenaires en présence, l'un étant immobilisé sur une surface tandis que l'autre est en solution. L'ELISA a été développé pour la détection d'anticorps dans des sérums. Le Biacore<sup>®</sup> dérive d'une technique appelée résonance plasmonique de surface que l'on a adaptée au criblage à haut débit notamment dans l'industrie pharmaceutique. Ces deux techniques sont largement utilisées dans le domaine des interactions biomoléculaires, la première essentiellement en tant que test qualitatif permettant de détecter des interactions entre deux biomolécules, la seconde fournissant des résultats quantitatifs concernant les constantes de dissociation de complexe et les cinétiques d'association et de dissociation. Ces deux méthodes sont fondées sur un principe commun qui repose sur l'immobilisation d'un des

deux partenaires sur une surface « fonctionnalisée », les différences concernent essentiellement l'immobilisation des molécules sur la surface et le système de détection.

#### **1.4.1 – La résonance plasmonique de surface et la technologie Biacore®.**

Le Biacore® correspond au développement technologique de la résonance plasmonique de surface, vers une spécialisation dans la détection de processus de *binding*, en particulier pour les biomolécules. Cette technologie trouve essentiellement ses applications dans l'industrie pharmaceutique comme technique HTS pour le *screening* de composés sur des cibles pharmacologiques et le développement « *hit to lead* », et également pour des études de *binding* entre macromolécules d'intérêt biologique dans les domaines de la neurobiologie, la biologie cellulaire et structurale, l'immunologie, ... ainsi que dans l'industrie agroalimentaire.

##### ***Principe de la méthode.***

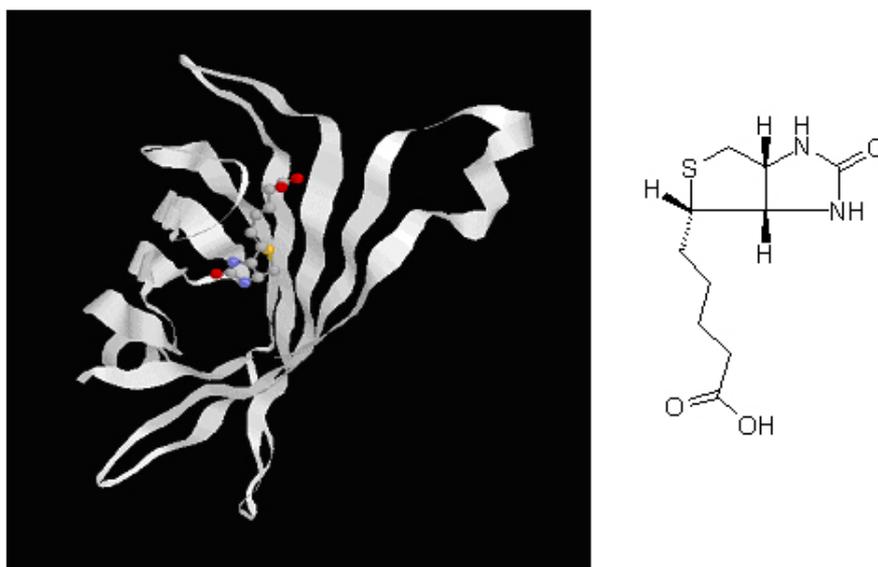
La surface sur laquelle un des partenaires de l'interaction est immobilisé, est une puce (*sensor chip*) constituée d'une fine pellicule d'or recouverte de dextran. Cette surface est ensuite fonctionnalisée en fonction de la nature du ligand qui doit y être immobilisé (peptide ou protéine, ADN, cellules, virus, ligands, ...). Un grand nombre de puces et une large gamme de techniques de fonctionnalisation et d'ancrage de substrats sont disponibles en fonction des applications. La maîtrise des conditions d'immobilisation aux surfaces conduit à des résultats reproductibles liés à une stabilité chimique permettant plusieurs utilisations et à la minimisation des phénomènes de liaison non-spécifique. Un des partenaires de l'interaction est fixé à la surface de la puce, l'autre partenaire est amené en solution au contact de la surface par un flux contrôlé précisément. Le suivi de l'association sous forme de complexe puis de la dissociation après passage d'une solution de lavage est réalisé instantanément par l'intermédiaire d'un sensorgramme. Le couplage d'une molécule sur la surface ou l'association en complexe résulte en une augmentation du signal et peut être ainsi quantifiée. La surface d'or recrée les conditions physiques de la résonance plasmonique de surface, qui permet de détecter d'infimes changements de masse de la phase aqueuse entourant la surface de la puce en mesurant des changements de l'indice de réfraction de celle-ci. Quand il y a association entre deux partenaires, la masse augmente, quand il y a dissociation, la masse diminue. Les deux processus peuvent donc être suivis en continu et en temps réel.

Les principaux avantages de cette technique sont la faible consommation de matériel biologique, la régénération possible des surfaces pour des utilisations répétées, la rapidité de

la mesure (quelques minutes) et une large gamme d'analytes compatibles avec la méthodologie allant des macromolécules d'intérêt biologique aux petites molécules de moins de 100 Da.

***Détection de la liaison du domaine WW isolé de Pin1 aux phospho-peptides.***

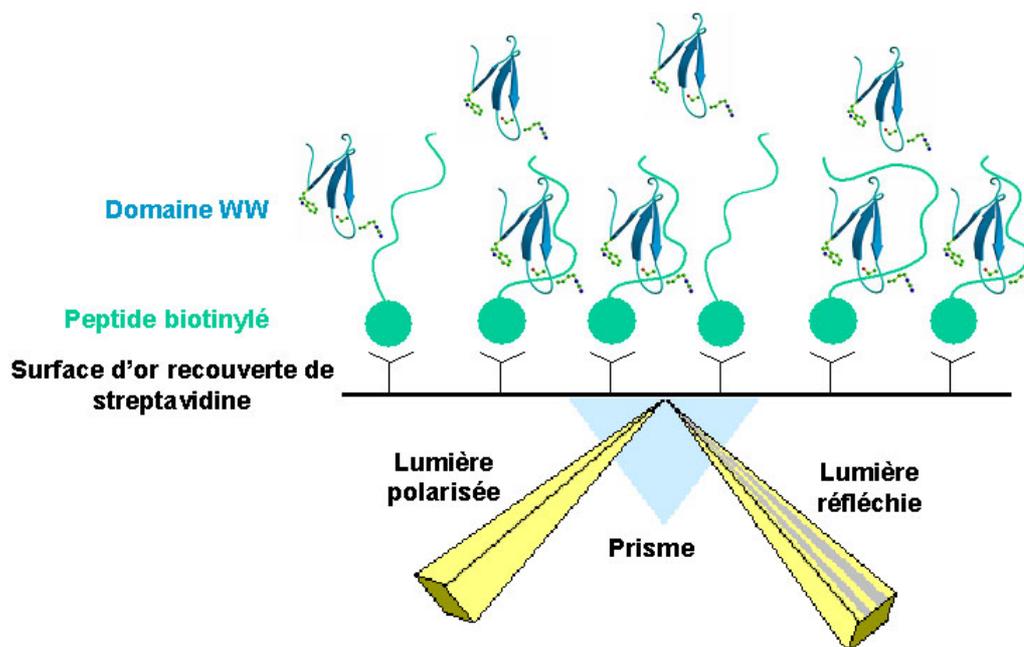
Dans un souci d'homogénéité avec les conditions mises en œuvre pour l'ELISA, nous avons fait le choix d'immobiliser les peptides phosphorylés sur la puce par un système mettant à profit la grande spécificité et la forte affinité ( $K_D=10^{-14}$  M) du couple biotine/streptavidine (Weber, Ohlendorf et al. 1989) (Figure 2.29). La synthèse des peptides en phase solide a permis l'introduction de manière quantitative et sélective de la biotine en position N-terminale sur le peptide par une étape de couplage similaire à celle utilisée pour les acides aminés (avec utilisation d'un mélange HOBt/HBTU comme agents d'activation en présence de DIPEA dans le DMF).



**Figure 2.29 :** Structure tridimensionnelle d'un monomère de streptavidine (à gauche) montrant la cavité où se loge la (D)-biotine dont la structure est donnée à droite. La streptavidine se présente sous forme d'un tétramère liant 4 molécules de biotine.

Une fonctionnalisation du dextran permet le couplage de la streptavidine (ou de la neutravidine) par réaction des fonctions amines primaires nucléophiles des résidus lysine exposés de la streptavidine sur les fonctions acides du dextran obtenues sous forme d'esters activés par réaction préalable avec le couple EDC/NHS en solution aqueuse à 200 mM et 50 mM respectivement. La solution est injectée à un débit de 10  $\mu$ L/min pour un volume effectif

de 70  $\mu\text{L}$  (le volume total est de 100  $\mu\text{L}$  avec 30  $\mu\text{L}$  volume mort). Une solution de streptavidine à 500  $\mu\text{g}/\text{mL}$  dans un tampon d'acétate de sodium (pH 3.5) est amenée en flux continu à 10  $\mu\text{L}/\text{min}$  avec un volume total de 120  $\mu\text{L}$ , puis on passe 60  $\mu\text{L}$  d'une solution d'éthanolamine 1M à 10  $\mu\text{L}/\text{min}$  pour désactiver les fonctions acides n'ayant pas réagi. On lave deux fois avec 10  $\mu\text{L}$  d'une solution de SDS à 0.03% puis on injecte 10  $\mu\text{L}$  d'une solution de peptide à 10  $\mu\text{L}/\text{min}$  suivi d'un lavage de 10  $\mu\text{L}$  au SDS à 0.03%.



**Schéma 2.3 :** Représentation du système mis en œuvre en Biacore<sup>®</sup> pour la détection d'interactions entre des peptides phosphorylés et le domaine WW de Pin1.

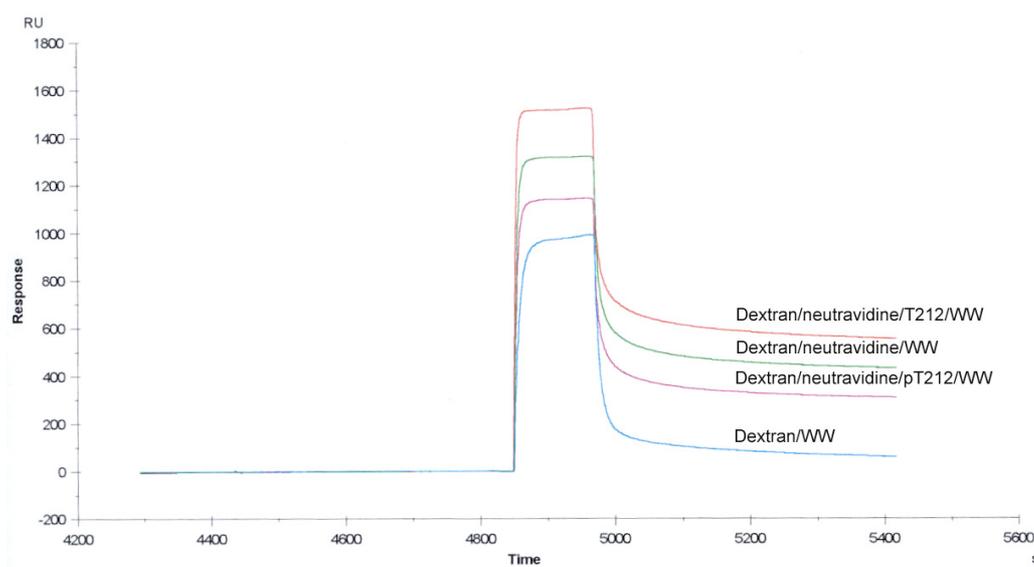
Les solutions de peptide et de protéine sont réalisées dans un tampon Tris 50 mM (pH 6.4), NaCl 100 mM, DTT 5 mM, EDTA 5 mM contenant 10% de BSA. Le système possédant 4 canaux utilisés en parallèle, les contrôles sont effectués simultanément à la mesure :

- Canal 1 : peptide T212 non phosphorylé biotinylé à 2  $\mu\text{g}/\text{mL}$  (contrôle négatif)
- Canal 2 : phospho-peptide pT212 biotinylé à 0.5  $\mu\text{g}/\text{mL}$
- Canal 3 : peptide T212 non phosphorylé biotinylé à 40 ng/mL (contrôle négatif)
- Canal 4 : phospho-peptide pT212 biotinylé 10 ng/mL.

Les résultats ont montré des signaux équivalents en terme de charge de la surface pour les peptides phosphorylés et les contrôles non phosphorylés (d'après l'unité de mesure en RU ou unité de résonance) dans les deux cas (Canaux 1/2 et 3/4).

Quatre cycles d'injection (de 40  $\mu$ L par cycle) d'une solution de WW à différentes concentrations (10 nM, 100 nM, 1  $\mu$ M et 200  $\mu$ M) sont effectués à 20  $\mu$ L/min sur les quatre canaux simultanément suivis d'une injection de 20  $\mu$ L d'une solution de SDS à 0.03%. On note pour chaque canal que l'intensité du signal augmente avec la concentration en protéine en solution et que les signaux des canaux 1 et 3 correspondant aux contrôles négatifs sont plus élevés que les signaux des canaux 2 et 4 respectivement, traduisant une association plus importante avec le peptide non phosphorylé qu'avec l'homologue phosphorylé en Thr212. Un contrôle a été effectué avec le domaine WW sans His-tag afin d'éviter d'éventuelles interférences entre la queue de 6 histidines chargée positivement et la surface de dextran chargée négativement. Des contrôles supplémentaires ont été effectués par interaction du domaine WW avec la surface sans peptide (Dextran/neutravidine/WW) et avec la surface de dextran sans streptavidine ou neutravidine (Dextran/WW). Les résultats obtenus avec le domaine WW sans His-tag, similaires à ceux obtenus avec le domaine WW contenant un His-tag, ne montrent pas d'interactions significatives entre la queue poly-histidine et la surface. En revanche, l'absence de peptide ne diminue pas les interactions du domaine WW avec la surface. Les intensités des signaux obtenus pour les différents contrôles par rapport au peptide phosphorylé sont les suivants en fonction du dernier élément immobilisé sur la surface (Figure 2.30) :

Peptide non phosphorylé > neutravidine > peptide phosphorylé > dextran



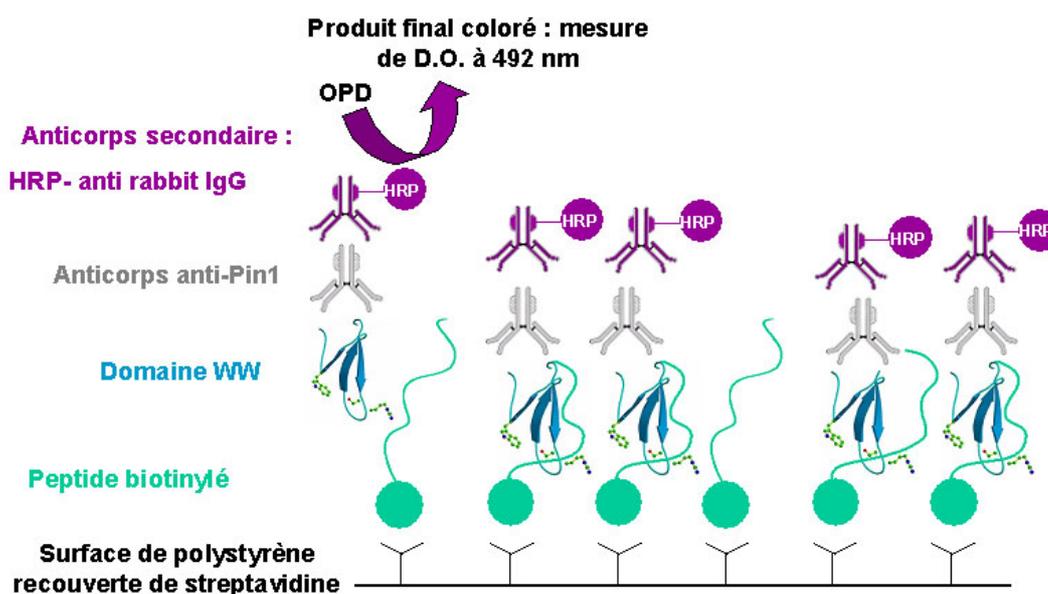
**Figure 2.30** : Courbes d'association et de dissociation du domaine WW sur les différentes surfaces de la puce correspondant aux 4 canaux.

Ces données sont révélatrices d'un phénomène de *binding* non-spécifique de la protéine avec la surface de dextran, avec la streptavidine et le peptide non phosphorylé malgré la présence de BSA dans le tampon qui assure en partie la saturation des sites de liaison non-spécifiques. Nous n'avons pas été capables de discerner l'interaction spécifique des interactions non-spécifiques avec cette méthode, probablement parce que l'interaction spécifique est trop faible et ne se distingue pas suffisamment des interactions non-spécifiques ou parce que la dissociation du complexe formé par le domaine WW et le phospho-peptide est trop rapide.

#### 1.4.2 – ELISA (*Enzyme-Linked ImmunoSorbant Assay*).

La technique ELISA est utilisée ici de manière assez originale par rapport à l'usage habituel qui consiste à recouvrir la surface du puits avec un anticorps. Dans la stratégie employée ici pour détecter des interactions entre Pin1 et des phospho-peptides, le protocole mis en œuvre est celui qui a été utilisé dans l'étude de Lu et collaborateurs (Jicha, Lane et al. 1997; Lu, Wulf et al. 1999). Comme en Biacore, le peptide phosphorylé est immobilisé sur la surface via la (D)-biotine par complexation à la streptavidine. En revanche, la streptavidine n'est pas fixée par couplage covalent à la surface, mais par simple adsorption sur des surfaces de polystyrène. Ces supports favorisent l'adsorption des macromolécules par formation de liaisons de Van der Waals, et sont optimisés en terme de quantités de molécules immobilisées tout en minimisant la proportion de macromolécules liées biologiquement inactives (plaques 96 puits, Nunc<sup>TM</sup> MaxiSorp, Nalge Nunc Int.). La neutravidine est utilisée à la place de la streptavidine (Pierce, ImmunoPure<sup>®</sup>), car il s'agit d'une forme modifiée, déglycosylée, de l'avidine et de la streptavidine, avec un point isoélectrique de 6.3, qui permet de réduire les liaisons et adsorptions non-spécifiques. La neutravidine, en solution à 25 µg/mL dans un tampon constitué de 30mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> (pH 7.2), 140 mM NaCl, 1 mM EDTA, est incubée pendant 3 heures à 37°C. Cette étape est suivie de période d'incubation d'une heure avec 2% de BSA dans le tampon TBS à température ambiante. Ensuite, le peptide biotinylé est immobilisé sur la surface via la formation d'un complexe de haute affinité entre la streptavidine (ou neutravidine) et la biotine, par incubation d'une solution de peptide de concentration variable en solution dans un tampon TBS (pH 8.0) contenant 2% de BSA pendant deux heures à température ambiante sous agitation. Les puits sont ensuite incubés dans les mêmes conditions successivement avec (i) une solution de protéine dans le TBS (pH 8.0) contenant 2% de BSA, de concentration variable, (ii) une solution d'anticorps anti-His-tag (ou anti-Pin1) diluée au 1000<sup>e</sup> dans le Tris 25 mM (pH 6.8) contenant 5% de lait et (iii) une solution d'anticorps secondaire couplé à l'enzyme HRP (*Horseradish Peroxidase*) diluée

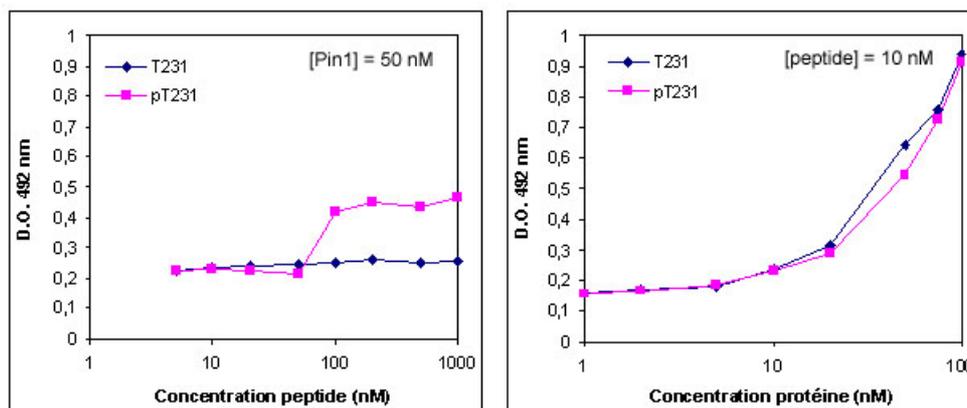
au 2000° dans le même tampon que l'anticorps précédent. Chaque étape d'incubation est suivie de trois lavages au TBS contenant 0.1 % Tween-20 et deux lavages au TBS. Après incubation de l'anticorps secondaire, la détection est réalisée par réaction catalytique de l'HRP sur l'OPD en présence de perborate de sodium dans un tampon citrate. La réaction conduit à la formation d'un produit coloré dont la concentration peut être estimée par lecture de l'absorption à 450 nm, ou à 492 nm après arrêt de la réaction par ajout d'acide sulfurique (Schéma 2.4).



**Schéma 2.4 :** Représentation du système mis en œuvre en ELISA pour la détection des interactions entre Pin1 et le peptide phosphorylé centré autour de l'épitope phospho-Thr231.

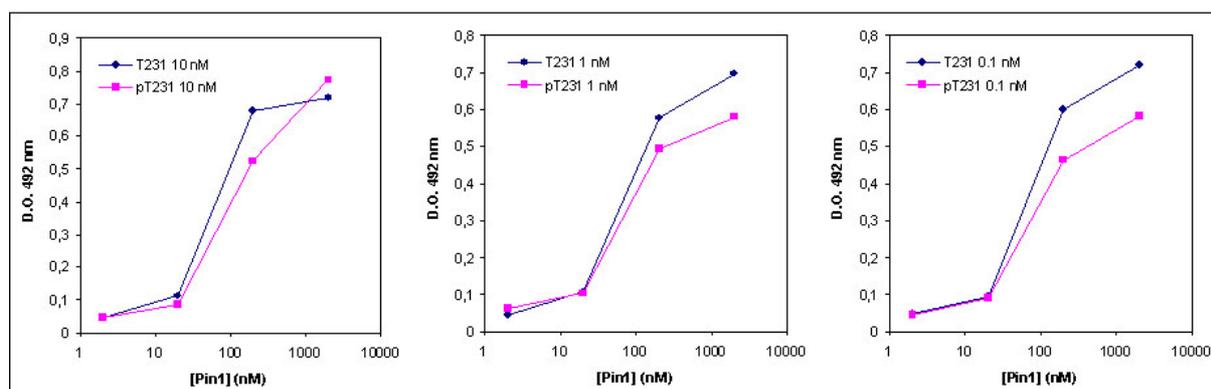
### **Résultats.**

Les résultats ont montré un comportement du système semblable aux résultats du Biacore. L'absorption du produit coloré issu de la réaction enzymatique catalysée par l'HRP sur l'OPD est identique (voire plus intense) avec le témoin non phosphorylé que l'absorption résultant de l'interaction avec le peptide phosphorylé (Graphique 2.5, Graphique 2.6). Les différents contrôles : sans peptide et/ou sans streptavidine ont montré des résultats positifs, l'intensité de l'absorbance étant toujours proportionnelle à la concentration de Pin1 en solution incubée dans les puits suggérant une adsorption de la protéine elle-même malgré l'utilisation de BSA (Calbiochem, BSA fractionV, RIA and ELISA Grade) dans les tampons.



**Graphique 2.5 :** Variation de l'intensité d'absorption à 492 nm pour des concentrations variables de peptide (T231, témoin non phosphorylé en bleu, pT231 phosphorylé, en rose) et une concentration de Pin1 constante de 50 nM (à gauche), pour une concentration de peptide constante de 10 nM et une concentration en Pin1 variable (à droite).

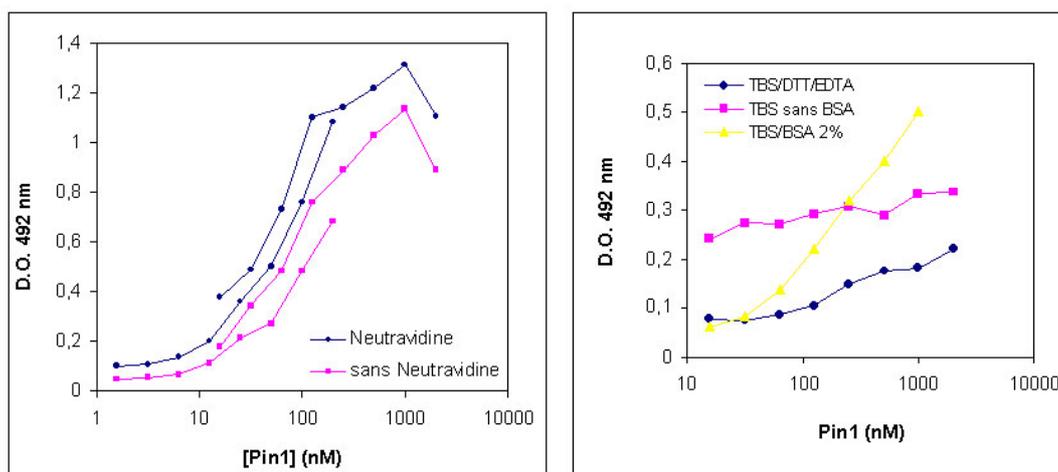
Un grand nombre de paramètres intervenant dans l'expérience ont été étudiés séparément : la dilution des anticorps, la concentration de la BSA, la concentration de détergent dans les lavages, le substrat de l'HRP (OPD ou ABTS). Deux anticorps ont été testés : anti-Pin1 polyclonal (avec Pin1 sans His-tag) et anti-His-tag monoclonal. La fixation des peptides a été vérifiée par détection avec l'anticorps AT180 dirigé contre l'épitope phospho-Thr231 avec ou sans neutravidine.



**Graphique 2.6 :** Variation de l'intensité d'absorption à 492 nm en fonction de la concentration en Pin1 pour différentes concentrations de peptide (témoin non phosphorylé T231 en bleu et phosphorylé pT231 en rose) : 10, 1, et 0.1 nM.

Les différents contrôles ont permis de vérifier la fixation des peptides T231/pT231 biotinylés via la formation d'un complexe avec la neutravidine (Graphique 2.8), même si la présence de

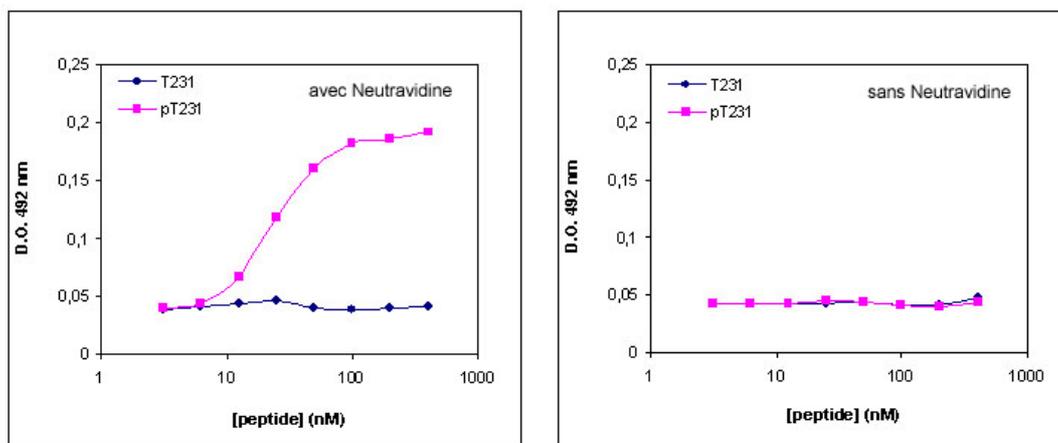
Pin1 est détectée en concentration relativement identique, en présence du peptide phosphorylé, du témoin non phosphorylé et en l'absence de peptide (Graphique 2.6, Graphique 2.7). En fait, les données montrent que Pin1 se fixe par adsorption dans les puits et que la quantité adsorbée est légèrement inférieure en l'absence de neutravidine, ce qui signifie la présence d'interactions non-spécifiques entre les deux protéines, mais la majorité du signal est due à l'adsorption directe de Pin1 sur la surface du puits. Cette adsorption est d'ailleurs minimisée par la présence de la neutravidine car la différence de signal en présence et en absence de neutravidine diminue au fur et à mesure que la concentration de Pin1 augmente (Graphique 2.7). D'autre part, l'ajout de BSA a une influence notable sur les interactions non-spécifiques puisque la BSA à 2% masque une partie des sites d'interaction accessibles à Pin1 mais n'est plus suffisante à partir d'une certaine concentration de Pin1 située entre 200 et 300 nM (Graphique 2.7). L'ajout de DTT et d'EDTA réduit également l'adsorption de Pin1 de manière significative (Graphique 2.7).



**Graphique 2.7 :** A gauche, variation de l'intensité d'absorption à 492 nm en absence de peptide en fonction de la concentration en Pin1 en présence de neutravidine (en bleu), en absence de neutravidine (en rose). A droite, variation de l'absorption à 492 nm en absence de peptide et de neutravidine en fonction de la concentration de Pin1 en solution incubée dans les puits : en présence de BSA à 2% (poids/volume) (en jaune), en présence de DTT et d'EDTA à 5 mM (en bleu) et en l'absence de BSA (en rose).

Les données de détection de Pin1 en présence de différentes concentrations de peptide phosphorylé ou témoin non phosphorylé, ou même en l'absence de peptide, suggèrent une possible absence de fixation du peptide via la neutravidine. La détection de l'épitope phospho-Thr231 avec l'anticorps monoclonal AT180 dirigé spécifiquement contre cet épitope a permis de vérifier la fixation du peptide : le signal est proportionnel à la concentration de

peptide phosphorylé et on note une absence de signal pour le peptide non phosphorylé ainsi que pour les deux peptides en l'absence de neutravidine (Graphique 2.8).



**Graphique 2.8 :** Détection par l'anticorps AT180 de l'immobilisation des peptides phosphorylé pT231 (en rose) et non phosphorylé (T231) en présence (à gauche) et en absence de Neutravidine (à droite) pour différentes concentrations.

D'après les données des différentes expériences réalisées par ELISA, les interactions non-spécifiques semblent masquer l'interaction spécifique concernant la liaison de Pin1 aux peptides phosphorylés, qui avait été observée par RMN et fluorescence, et qui est vraisemblablement trop faible pour être mesurable dans ce type d'expérience où plusieurs partenaires sont impliqués. En ELISA, comme en Biacore, les interactions non-spécifiques sont localisées au niveau de la surface sur laquelle sont immobilisés les peptides, qui permet également l'adsorption de Pin1 et dont les affinités sont du même ordre que celles qui existent entre Pin1 et les phospho-peptides (ou moins élevées mais plus nombreuses). De manière générale, ces deux techniques sont utilisées pour détecter des interactions plus fortes que celle qui est impliquée dans cette étude. En ELISA, on détecte généralement, des interactions anticorps/antigène dont les ordres de grandeurs sont donnés dans le Tableau 2.8.

<b>Interaction</b>	<b><math>K_D</math></b>
Streptavidine-biotine	$10^{-14}$ M
Anticorps-antigène avec un bon anticorps	$10^{-8}$ M- $10^{-10}$ M
Protéine liant l'ADN au niveau de sites spécifiques	$10^{-8}$ M- $10^{-10}$ M
Anticorps-antigène avec un anticorps faible	$10^{-6}$ M
Enzyme-substrat	$10^{-4}$ M- $10^{-10}$ M

**Tableau 2.8 :** Ordres de grandeurs des principales interactions entre biomolécules.

Ces méthodes de détermination de l'affinité en milieu hétérogène ne sont pas adaptées à la mesure d'affinités faibles qui ont pu être mesurées par RMN et fluorescence dans le cas de la liaison de Pin1 à des phospho-peptides, car elles introduisent des éléments supplémentaires qui n'interviennent pas directement dans l'interaction ciblée mais qui sont susceptibles d'ajouter des interférences dans la mesure, qui sont négligeables dans le cas de détection de fortes affinités mais qui masquent les interactions faibles dont le seuil n'est pas clairement défini.

### **1.5 – Replacer les interactions Pin1-tau dans le contexte cellulaire.**

Les conditions expérimentales de mesure des interactions entre deux ou plusieurs molécules sont soumises aux contraintes propres à la technique mise en œuvre, et elles sont nombreuses. Mais dans l'ensemble, ces conditions sont très éloignées des conditions physiologiques, c'est-à-dire celles de la cellule. Nous avons mis en œuvre dans les paragraphes précédents différentes techniques et discuté de l'apport de chacune d'entre elles dans le cas particulier des interactions entre Pin1 et des phospho-épitopes de la protéine tau. La tendance actuelle, avec les efforts fournis dans la connaissance du génome humain, des mécanismes régissant le métabolisme, des structures protéiques et des réseaux d'interaction, est d'aller vers l'étude des protéines sur le plan structural et fonctionnel mais aussi des interactions protéine-protéine ou protéine-ligand dans des conditions approchant l'environnement physiologique. La liaison des phospho-épitopes par le domaine WW de Pin1 se situant dans une gamme d'affinités allant de 0.1 à 1 mM mesurées *in vitro*, on peut poser la question de la pertinence de telles interactions de faible affinité dans un contexte cellulaire. D'ailleurs, les techniques de mesure d'affinité telles que la résonance plasmonique de surface et l'ELISA ont montré que des interactions non spécifiques de faible affinité mais plus nombreuses empêchaient l'observation de ces interactions pourtant spécifiques mais faibles associées à une dissociation rapide des complexes formés. Les interactions entre les protéines Pin1 et tau dans la cellule se produisent dans un contexte où sont présents différents constituants cellulaires, parfois en quantités importantes comme l'ATP, ainsi que d'autres substrats de Pin1.

#### ***La phospho-spécificité de Pin1 lui permet-elle de lier l'ATP ?***

L'ATP est présent en concentration importante dans la cellule de l'ordre du millimolaire car c'est une molécule essentielle qui est impliquée dans les transferts énergétiques (l'hydrolyse de l'ATP en ADP, c'est-à-dire la cassure de la liaison pyrophosphate, libère environ 7 kcal/mol) et intervient dans le processus de phosphorylation post-traductionnelle par les

kinases. La phospho-spécificité du domaine WW de Pin1 pourrait être impliquée dans la liaison de l'ATP et cette interaction serait susceptible d'altérer la liaison des phospho-épitopes car, même si l'interaction avec l'ATP est faible, les quantités importantes d'ATP dans les cellules pourraient masquer la liaison de Pin1 à ses substrats.

Nous n'avons pas observé d'interactions entre le domaine WW et l'ATP mais une interaction faible au niveau de la queue poly-histidine ainsi qu'une diminution de la concentration de la protéine d'un facteur deux en présence de 20 équiv. d'ATP (soit 2 mM en présence de 100  $\mu$ M de WW enrichi en  $^{15}\text{N}$ ), qui pourrait provenir d'une interaction de type « *stacking* » entre l'adénosine et les cycles aromatiques imidazoles de la queue poly-histidine.

Nous avons choisi d'évaluer les interactions entre Pin1 et l'ATP en raison des quantités cellulaires importantes de ce composé. Il est évident qu'un grand nombre de partenaires potentiels de Pin1 existent dans le contexte cellulaire. Une identification plus systématique de tels constituants, notamment des protéines phosphorylées, peut être effectuée par *pull-down*. Cette technique permet de trouver des complexes relativement stables à partir d'une protéine immobilisée sur un support solide et des partenaires potentiels dans un extrait cellulaire, par exemple. L'identification des protéines retenues par *pull-down* peut être effectuée par *western blot* en détectant les protéines avec des anticorps spécifiques ou plus largement par digestion enzymatique et spectrométrie de masse.

## **II – Le motif phosphoThr212-Pro213, un nouveau site d'interaction entre la protéine tau et la peptidyl-prolyl *cis/trans* isomérase Pin1 (Smet, Sambo et al. 2004).**

### **2.1 – Le motif phosphoThr212-Pro213 reconnu par le domaine WW de Pin1.**

Nous avons montré dans le paragraphe précédent que le domaine WW était responsable de la liaison du substrat et pouvait fonctionner même isolé du contexte de Pin1. Le *mapping* des phospho-épitopes de type phosphoSer/Thr-Pro de la protéine tau a montré que les interactions entre Pin1 et tau *in vitro* n'étaient pas limitées à un site unique avec une forte affinité mais pouvaient être médiées par plusieurs sites de faibles affinités, dont le site Thr212, pour lequel l'affinité la plus élevée avait été mesurée. D'autre part, il a été montré que le domaine WW liait les phospho-épitopes de manière *trans*-spécifique (Wintjens, Wieruszeski et al. 2001). Nous avons envisagé que l'affinité pour les différents phospho-épitopes de la protéine tau pouvait être modulée, non pas uniquement par la séquence peptidique environnante mais également par la conformation locale de ces motifs notamment autour des résidus Pro. Nous avons donc réalisé une étude conformationnelle de l'épitope Thr212-Pro213 au niveau de substrats peptidiques dans différents états de phosphorylation et étudié l'activité enzymatique de Pin1 pour valider cet épitope comme nouveau site d'interaction entre tau et Pin1.

### **2.2 - Etude conformationnelle autour du motif Thr212-Pro213 et isomérisation catalysée par la peptidyl-prolyl *cis/trans* isomérase Pin1.**

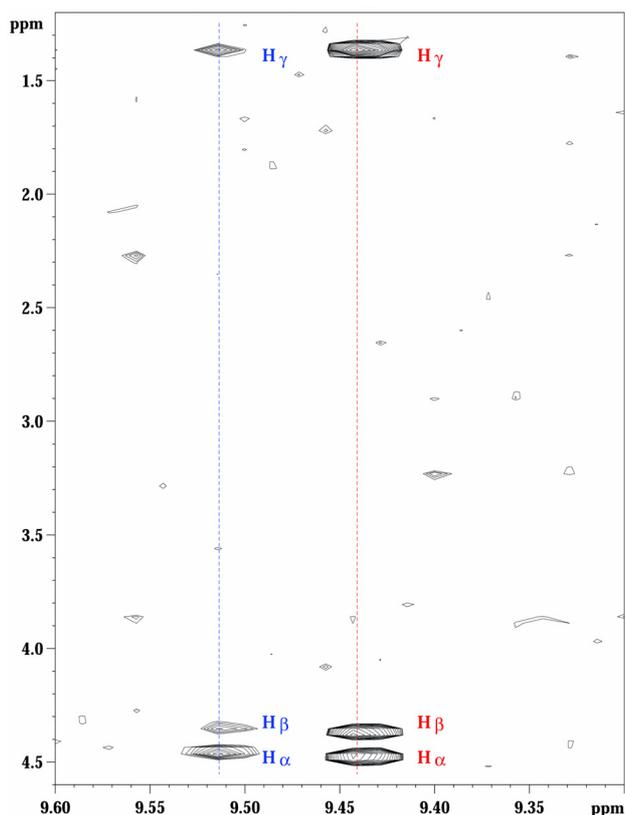
Le criblage des phospho-épitopes de la protéine tau, inclus dans des motifs Ser/Thr-Pro, a permis l'identification de la Thr212 comme site reconnu par le domaine de liaison de la peptidyl-prolyl *cis/trans* isomérase Pin1. Ce site, phosphorylé spécifiquement dans les PHF, pourrait contribuer, au niveau neuronal, à favoriser les interactions entre les deux protéines. D'autres sites à proximité de la Thr212 dans la séquence primaire de tau peuvent également être phosphorylés (Ser214 et Thr217), et la paire Thr212/Ser214 en particulier, est reconnue par l'anticorps AT100 conformationnellement-dépendant et spécifique des PHF. Nous avons donc réalisé une étude conformationnelle par RMN des peptides diversement phosphorylés autour de ces trois épitopes, et étudié l'influence de la phosphorylation sur l'équilibre conformationnel des prolines dans un contexte peptidique. L'identification de formes *cis* en proportion non négligeable, a permis d'étudier, au niveau d'un substrat peptidique, l'effet

catalytique de la peptidyl-prolyl isomérase Pin1 sur la vitesse d'échange entre les formes *cis* et *trans*.

### **2.2.1 – Détection de l'équilibre *cis/trans* des prolines par RMN à haute résolution et attribution de formes minoritaires.**

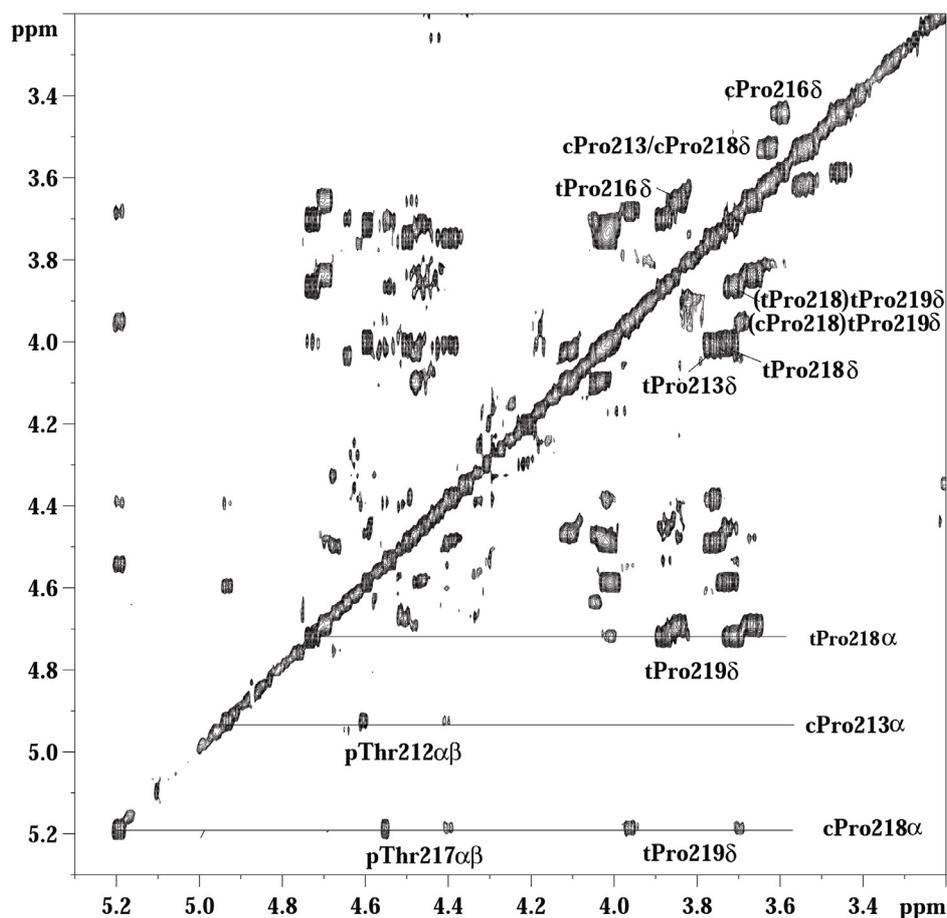
Deux acides aminés ont un rôle particulier au niveau conformationnel et structural dans les séquences polypeptidiques : la glycine et la proline. Si la glycine, avec l'absence de chaîne latérale induit plus de flexibilité dans la chaîne polypeptidique, la proline, dont la chaîne latérale est comprise dans un cycle à 5 chaînons, induit une rigidité plus importante par rapport aux autres acides aminés naturels. La présence de prolines dans les séquences peptidiques conduit à des équilibres conformationnels entre les conformères limites *cis* et *trans*, que l'on peut distinguer dans la mesure où l'échange entre les deux formes est relativement lent sur l'échelle de temps des événements biologiques. D'ailleurs, cet équilibre *cis/trans* est considéré comme une des étapes cinétiquement limitantes dans le repliement des protéines. Dans des protéines structurées, la présence de formes *cis*, qui est la forme la moins favorable d'un point de vue thermodynamique, peut être relativement importante. La présence de prolines peut ainsi jouer un rôle conformationnel clé dans la régulation d'une activité enzymatique ou dans la reconnaissance de ligands. Dans des peptides courts, contenant 10 à 20 résidus, les formes *cis* atteignent rarement les 10%.

La RMN est une technique particulièrement bien adaptée à l'observation de cet équilibre entre les conformères *cis* et *trans* des prolines. L'échange conformationnel se faisant sur une échelle de temps plus lente que l'échelle de temps RMN, les deux conformères donnent des résonances distinctes au niveau des résidus Pro mêmes, mais également au niveau des résidus voisins (Figure 2.31). Cependant, dans les peptides, la faible proportion des formes *cis*, ainsi que la faible sensibilité de la RMN, peuvent limiter la détection de telles conformations. Dans le cas des substrats peptidiques utilisés pour le criblage des phospho-épitopes de tau, certaines conformations *cis* ont pu être détectées grâce à l'utilisation d'une cryosonde, qui a grandement facilité l'acquisition des spectres dans des temps corrects et l'attribution des formes minoritaires.



**Figure 2.31** : Observation de l'équilibre *cis/trans* de la Pro213 au niveau des contacts  $^1\text{H}$ - $^1\text{H}$  du résidu phospho-Thr212 précédant sur un spectre TOCSY dont le détail représenté ici correspond aux corrélations entre les protons  $\text{H}_\text{N}$  et les protons  $\text{H}_\alpha$ ,  $\text{H}_\beta$  et  $\text{H}_\gamma$  de la chaîne latérale de la Thr phosphorylée. Les signaux les plus intenses, annotés en rouge, correspondent à la forme majoritaire liée à la Pro213 en *trans* (90% environ). Les signaux annotés en bleu correspondent à la forme minoritaire liée à la Pro213 en *cis* (10% environ).

L'attribution des peptides a été réalisée à partir d'expériences TOCSY et NOESY standards à 20°C, avec des temps de mélange respectifs de 60 et 400 ms, permettant l'attribution des résidus au travers des contacts scalaires  $^1\text{H}$ - $^1\text{H}$  avec le TOCSY ainsi que l'attribution séquentielle au travers des contacts dipolaires  $^1\text{H}$ - $^1\text{H}$  avec le NOESY. Les peptides sont solubilisés à une concentration de 2 mM dans un tampon aqueux contenant 25mM de Tris deutéré à pH 6.40 et 50 mM NaCl. L'attribution des prolines a fait l'objet d'une attention particulière et a été réalisée à partir de spectres acquis dans le  $\text{D}_2\text{O}$ , afin de minimiser le signal résiduel de l'eau. Les formes *trans* de prolines sont identifiées grâce à des contacts NOE entre les protons  $\text{H}_\delta$  de la proline et le proton  $\text{H}_\alpha$  du résidu précédent dans la séquence primaire, les formes *cis* par des contacts NOE, distincts des formes *trans*, entre les protons  $\text{H}_\alpha$  de la proline et du résidu précédent (Figure 2.32).



**Figure 2.32 :** Attribution du peptide pT212 pS214 pT217 triphosphorylé sur un spectre NOESY (MT=400ms) dans le D<sub>2</sub>O. Les formes *trans* sont attribuées à partir de contacts NOE H $\alpha$ -H $\delta$  entre le proton  $\delta$  de la Pro et le proton  $\alpha$  du résidu qui la précède (Pro219). Les formes *cis* sont attribuées à partir de contacts H $\alpha$ -H $\alpha$  entre le proton  $\alpha$  de la Pro et la proton  $\alpha$  du résidu qui la précède (Pro213, Pro218). Les contacts H $\delta$ 1-H $\delta$ 2 sont annotés pour les Pro213, Pro216 et Pro218 en *trans* et en *cis*, ainsi que la Pro219 en *trans* précédée soit par la Pro218 en *trans* (tPro218)tPro219, soit par la Pro 218 en *cis* (cPro218)tPro219.

L'attribution complète en proton du peptide pT212 pS214 pT217 tri-phosphorylé est donnée dans le **Tableau 2.9**. Les peptides non, mono- et di-phosphorylés ont été attribués de la même manière. Les peptides utilisés comme substrat pour l'identification des sites d'interaction de Pin1 au sein de la protéine tau ne sont pas structurés et adoptent une conformation étendue dite *random coil*. Aucun couplage NOE H<sub>N</sub>-H<sub>N</sub> n'a été observé sauf dans le cas des peptides de la série T212 pour lesquels il existe un faible contact H<sub>N</sub>-H<sub>N</sub> entre les résidus Ser214 et Leu215, plus intense quand la Ser214 est phosphorylée.

		$H_N$	$H\alpha$	$H\beta$		$H\gamma$		$H\delta$	
		$H_N$	$H\alpha$	$H\beta1$	$H\beta2$	$H\gamma1$	$H\gamma2$	$H\delta1$	$H\delta2$
208	S	8.38	4.43	3.87					
209	R	8.5	4.44	1.91	1.79	1.67		3.21	
210	S	8.49	4.45	3.89					
211	R	8.48	4.37	1.91	1.79	1.67		3.23	
212 (t)	pT	9.45	4.50	4.40		1.38			
212 (c)	pT	9.52	4.61	4.41		1.27			
213 (t)	P	-	4.43	2.34	2.09	2.04	1.93	4.01	3.78
213 (c)	P	-	4.94	2.41	2.12	1.95	1.84	3.61	3.48
214	pS	9.10	4.44	4.11	3.98				
215 (t)	L	8.40	4.69	1.71	1.63	1.59		0.97	0.94
215 (c)	L	7.80	4.51	1.60		1.43		0.91	
216 (t)	P	-	4.49	2.31	2.03	1.89		3.87	3.68
216 (c)	P	-	4.60	2.41	2.13	1.95	1.85	3.62	3.48
217 (t)	pT	9.03	4.60	4.47		1.37			
217 (c)	pT	9.20	4.56	4.40		1.30			
218 (t)	P	-	4.72	2.4	2.09	2.05	1.9	4.01	3.75
218 (c)	P	-	5.20	2.52	2.17	2.00	1.86	3.62	3.54
219	P	-	4.55	2.36	2.09	2.05	1.93	3.89	3.71
220	T	8.29	4.32	4.22		1.22			

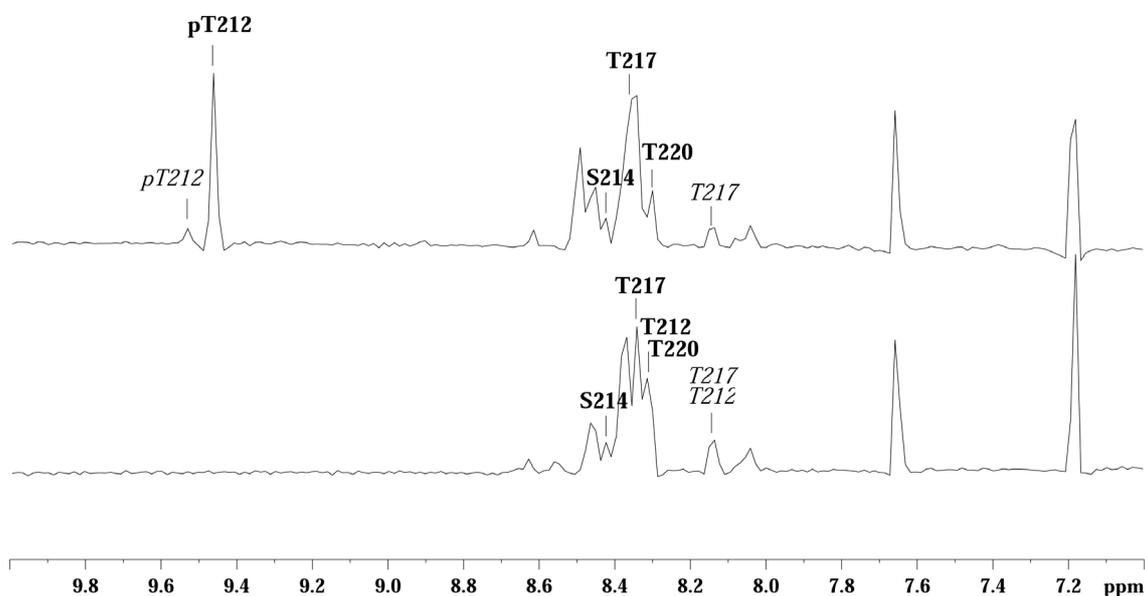
**Tableau 2.9 :** Attribution  $^1\text{H}$  des résidus du peptide pT212 pS214 pT217 tri-phosphorylé dont les formes majoritaires de Pro en *trans* ou liées à une Pro en *trans* sont annotées par (t), les formes minoritaires de Pro en *cis* ou liées à une Pro en *cis* sont annotées par (c).

### 2.2.2 – Influence de la phosphorylation sur l'équilibre *cis/trans* de la proline dans un motif Ser/Thr-Pro au niveau de substrats peptidiques.

Dans la série des peptides centrés autour des épitopes Thr231 et Ser235, le résidu Pro232 présente une population de conformères *cis* inférieure à 3%, confirmant les résultats précédemment publiés (Daly, Hoffmann et al. 2000). Dans le cas des peptides centrés autour des épitopes Thr212, Ser214 et Thr217, les formes *cis* sont plus importantes pour les Pro213, 216 et 218. Nous avons étudié l'influence de la phosphorylation sur la conformation des peptides non phosphorylé et tri-phosphorylé au niveau des 4 résidus Pro les constituant.

La phosphorylation des résidus Thr et Ser a une influence notable sur le déplacement chimique du proton  $H_N$  du résidu phosphorylé, qui se déplace vers les bas champs et qui peut ainsi être détecté facilement au niveau d'un spectre à une dimension (Figure 2.33). Les formes

minoritaires dues à la présence d'un résidu Pro en +1, peuvent également être détectées facilement et sont représentées par un signal de plus faible intensité.



**Figure 2.33** : Spectres  $^1\text{H}$  des peptides T212 S214 T217 (non phosphorylé) et pT212 S214 T217 (mono-phosphorylé). Les formes majoritaires des résidus T212/pT212, T217 sont annotées en caractère gras, les formes minoritaires en italique.

La quantification des populations relatives des formes *cis* et *trans* a été déterminée par l'intégration des signaux au niveau des spectres à une dimension et des corrélations H $\delta$ 1/H $\delta$ 2 sur les spectres NOESY (Tableau 2.10). Une proportion identique de formes *cis* dans les peptides mono-et tri-phosphorylés, indique que la phosphorylation n'a pas un impact significatif sur l'équilibre conformationnel des prolines dans le contexte peptidique.

<i>Peptide</i>	% <i>cis</i>			
	<i>P213</i>	<i>P216</i>	<i>P218</i>	<i>P219</i>
T212 S214 T217	6%	8%	12%	< 1%
pT212 pS214 pT217	7%	8%	13%	< 1%

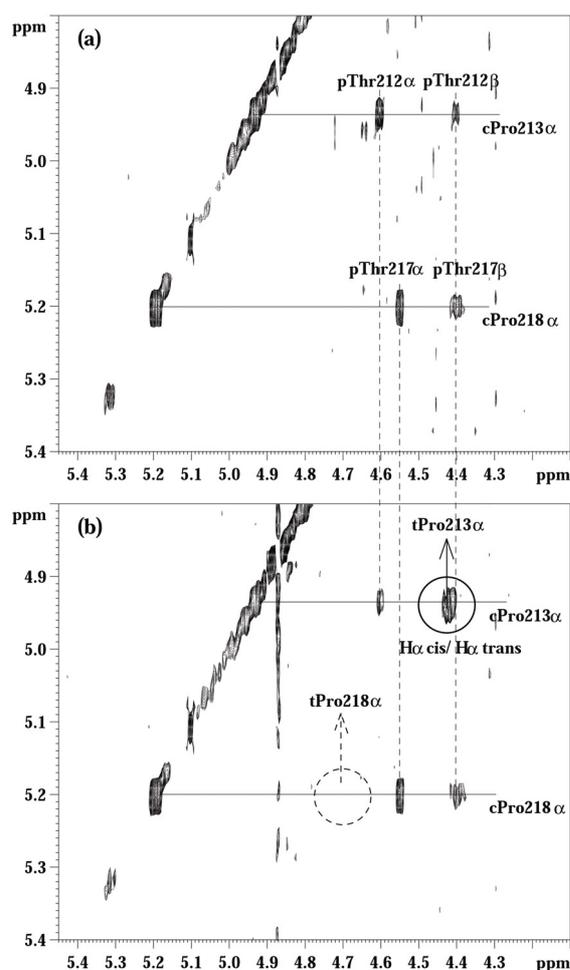
**Tableau 2.10** : Proportion de formes *cis* des différentes prolines dans les peptides T212 non phosphorylé et pT212 pS214 pT217 tri-phosphorylé.

La présence de prolines en conformation *cis* dans des proportions relativement importantes et détectables par RMN, notamment pour les Pro213 et 218 précédées d'un résidu phosphorylé,

autorise une étude de l'activité catalytique d'isomérisation de Pin1 et de son domaine catalytique isolé.

### **2.2.3 – Catalyse de l'isomérisation *cis/trans* par Pin1 et son domaine catalytique isolé au niveau de substrats peptidiques.**

L'isomérisation catalytique de Pin1 a été étudiée au niveau des peptides centrés autour des phospho-épitopes Thr212/Ser214 et Thr217 inclus dans des motifs Thr-Pro pour lesquels les formes *cis* sont détectables par RMN, alors que l'absence de formes *cis* détectables dans le cas de l'épitope Thr231, rend moins aisée l'observation de l'activité catalytique d'isomérisation. Nous avons donc utilisé la spectroscopie d'échange EXSY avec un temps de mélange de 400 ms pour détecter l'isomérisation catalysée par Pin1. L'observation d'un pic d'échange reste très délicate et doit être réalisée dans des conditions expérimentales adéquates, c'est-à-dire dans le D<sub>2</sub>O et à des concentrations de 2 mM en peptide. Des signaux supplémentaires apparaissent en présence d'une quantité catalytique de protéine Pin1 (environ 50 μM) traduisant l'accélération de l'échange entre les formes *cis* et *trans* en présence de l'enzyme, échange trop lent pour être détecté en l'absence d'enzyme. Un pic d'échange entre les protons H $\alpha$  des formes *cis* et *trans* de la Pro213 peut être facilement identifié grâce à l'attribution des formes minoritaires, ainsi que des contacts entre les protons H $\delta$  des formes *cis* et *trans*. En revanche, aucun pic d'échange n'est observé pour la Pro218, alors que la proportion de forme *cis* est plus importante que pour la Pro213 (Figure 2.34). La catalyse d'isomérisation est donc sélective au niveau de l'épitope phosphoThr212-Pro213 dans le contexte du peptide tri-phosphorylé car l'isomérisation de la liaison phosphoThr217-Pro218 n'est pas catalysée.



**Figure 2.34 :** Détail d'un spectre EXSY  $^1\text{H}$ - $^1\text{H}$  (MT=400ms) autour de la zone des protons  $\alpha$  des formes *cis* des Pro213 et Pro218 du peptide pT212 pS214 pT217 triphosphorylé montrant des contacts avec les protons  $\alpha$  et  $\beta$  des pThr212 et pThr217 respectivement. (a) Peptide seul à 2 mM en l'absence de Pin1 ; (b) Peptide à 1.5 mM en présence de 50  $\mu\text{M}$  de Pin1. Un signal supplémentaire correspondant à un contact entre le  $\text{H}\alpha$  de la forme *cis* de la Pro213 et le  $\text{H}\alpha$  de la forme *trans* est encerclé en trait plein. Le même type de contact pour la Pro218 n'est pas observé (cercle en trait pointillé).

La même expérience a été réalisée avec le domaine catalytique isolé sur le peptide triphosphorylé et a conduit aux mêmes résultats. L'isomérisation catalytique a également été observée avec Pin1 entière sur le peptide mono-phosphorylé en Thr212.

### 2.3 – Validation au niveau de la protéine tau phosphorylée de l'épitope phosphoThr212-Pro213 comme site reconnu par Pin1.

Dans le cadre d'une collaboration avec le groupe du Dr Luc Buée (Inserm, U422), nous avons validé l'épitope phosphoThr212-Pro213 comme motif intervenant dans la reconnaissance de tau par Pin1. Des cellules COS ont été transfectées de manière stable avec des plasmides codant pour la protéine tau sauvage et les mutants Tau T231A, T212A et T212A/T231A. Les cellules ont été traitées avec l'acide okadaïque, un inhibiteur de la phosphatase PP2A, permettant d'obtenir un taux élevé de phosphorylation des protéines tau après 42 heures. La phosphorylation des résidus T212 et T231 a été vérifiée dans les lysats cellulaires par immuno-réaction avec des anti-corps spécifiques des sites T212 et T231 phosphorylés

respectivement. L'absence de phosphorylation a également été vérifiée pour les mutants contenant un site de phosphorylation muté en alanine (Figure 2.35A).

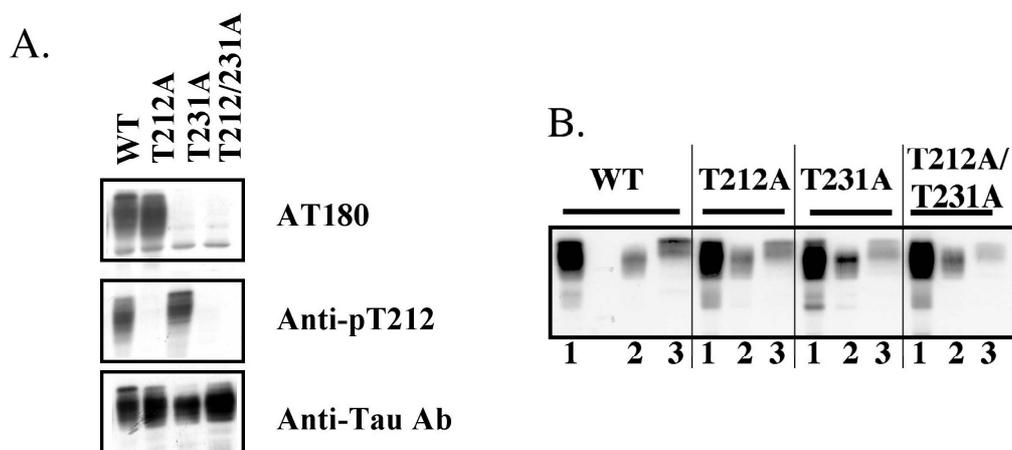
Les expériences de *pull-down* menées avec la protéine Pin1, le domaine WW isolé et le domaine catalytique isolé, fusionnés en position N-terminale à une GST, ont été réalisées à partir des extraits solubles obtenus dans les cellules COS surexprimant tau sauvage et les mutants. La présence de protéine tau dans les différentes fractions a été détectée avec l'anticorps polyclonal M19G qui reconnaît tau indépendamment de son degré de phosphorylation et des sites de phosphorylation. L'analyse des fractions correspondant au lysat total (Figure 2.35, ligne 1), à la fraction de protéines non retenues (Figure 2.35, ligne 2) et à la fraction de protéines retenues (Figure 2.35, ligne 3) a permis de déterminer la présence ou non de protéine tau dans les différentes fractions. Un contrôle a été effectué avec la GST seule et n'a révélé aucune interaction avec les protéines tau phosphorylées. De même, le domaine catalytique n'a retenu aucune des protéines tau phosphorylées. La protéine tau sauvage est retenue par Pin1, de même que les différents mutants. On notera que les protéines retenues, dans les 4 cas, sont les protéines de plus haut poids moléculaire parmi les différentes populations de protéines phosphorylées présentes dans l'extrait total, c'est-à-dire les protéines de plus haut degré de phosphorylation. Cependant, d'après l'intensité de signaux, on remarque que les quantités retenues pour les mutants sont moins importantes que pour la protéine sauvage. De même que le mutant T212A (phosphorylé en T231, entre autres sites de phosphorylation) est davantage retenu que le mutant T231A (phosphorylé en T212, entre autres sites de phosphorylation), lui-même davantage retenu que le double mutant T212A/T231A, indiquant la présence d'autres sites de phosphorylation que les T212 et T231. En ce qui concerne la différence de signal entre les mutants T212A et T231A, on ne peut pas conclure entre les trois propositions suivantes :

- soit la T231 présente un taux de phosphorylation plus élevé que la T212, donc le mutant T212A se lie davantage même si les deux sites sont d'affinités comparables,
- soit la T231 présente un taux de phosphorylation identique ou moindre par rapport à la T212 mais une affinité beaucoup plus élevée,
- soit la T231 présente à la fois un taux de phosphorylation plus important et une affinité plus élevée que la T212.

Même si ces expériences ne permettent pas de conclure quant aux affinités relatives de deux sites T212 et T231 contrairement aux expériences effectuées par RMN et par fluorescence,

ces données sont complémentaires car elles indiquent que les deux sites sont reconnus par Pin1 au sein de la protéine tau totale présentant des taux de phosphorylation plus proches des taux de phosphorylation cellulaires. Elles indiquent également la présence d'autres sites d'interaction qui ne sont pas nécessairement moins affins mais peut-être « moins phosphorylés » par rapport aux T231 et T212 par deux interprétations :

- la liaison du double mutant T212A/T231A
- la liaison préférentielle des protéines tau de haut poids moléculaire, donc de haut degré de phosphorylation.



**Figure 2.35 :** (A) Détection des sites de phosphorylation dans les lysats cellulaires totaux de cellules COS au sein des protéines tau sauvage (WT) et mutantes (T212A, T231A, T212A/T231A) par réaction immunologique avec des anticorps spécifiques dépendant de la phosphorylation : AT180 reconnaît la phospho-T231 et l'anticorps anti-pT212 reconnaît la phosphoThr212, un anticorps anti-Tau est utilisé comme contrôle, car indépendant du site de phosphorylation. (B) *Pull-down* réalisés sur des extraits de cellules COS surexprimant la protéine tau sauvage (WT) ou les mutants T212A, T231A et T231A/T212A. Les fractions de lysat total (1), de protéines non retenues (2) et de protéines retenues (3), par interaction avec Pin1 ou son domaine WW isolé fusionnés à une GST, sont analysées par immuno-détection. Les protéines tau sont détectées avec l'anticorps polyclonal M19G (Sambo et al., soumis).

Ces données montrent que, comme cela avait été suggéré par les données de RMN et de fluorescence, la reconnaissance de la protéine tau par Pin1 passe par la liaison de plusieurs phospho-épitopes et non un site unique. Cependant, il faut noter que les constantes de dissociation déterminées par RMN et confirmées par fluorescence, de l'ordre de 100  $\mu\text{M}$  à 1 mM, ne sont pas compatibles avec les expériences de *pull-down*, car elles impliquent des dissociations rapides, de l'ordre de 1500  $\text{s}^{-1}$  (d'après la simulation de l'échange chimique exposée dans le paragraphe 1.3.2). On peut considérer que la cinétique de dissociation est la

même par rapport aux expériences en milieu homogène mais que la cinétique d'association dans un tel système peut être modifiée et ne plus être uniquement dépendante de la diffusion moléculaire. Si la dissociation est rapide, le complexe pourra être déplacé lors des étapes de lavage. Les résultats obtenus par *pull-down* suggèrent une affinité élevée de Pin1 pour les protéines tau hautement phosphorylées qui permettrait l'isolation des complexes.

### **III – Discussion. Méthodes de caractérisation des interactions biomoléculaires *in vitro* et découverte d'un nouvel épitope de tau impliqué dans l'interaction avec Pin1.**

#### ***Intérêt de caractériser les interactions entre Pin1 et tau dans le contexte de la maladie d'Alzheimer.***

La phosphorylation de la protéine tau a été au centre de nombreuses questions concernant tant son rôle biologique lié à la régulation de la liaison de tau aux microtubules, que son rôle pathologique lié à l'agrégation de protéines tau hyperphosphorylées sous forme de PHF dans les neurones de patients atteints de la maladie d'Alzheimer. Des études immunologiques ont montré que la protéine tau chez l'adulte était moins phosphorylée que l'isoforme fœtale (tau 352, isoforme la plus courte), mais que les PHF l'étaient davantage, en termes de nombre de sites de phosphorylation et niveau de phosphorylation. En effet, les PHF possèdent un grand nombre de sites de phosphorylation et la plupart de ces sites ont un taux de phosphorylation plus élevé que la isoforme fœtale. La Thr231, en particulier, est quasiment totalement phosphorylée dans les PHF alors que ce site n'est phosphorylé qu'à 30% dans la protéine tau fœtale. Au niveau physiologique, la phosphorylation de tau provoque sa dissociation des microtubules et par suite, leur dépolymérisation sous forme de monomères de tubuline. La phosphorylation des épitopes Ser262 et Ser214 est directement impliquée dans la dissociation du complexe formé entre la protéine tau et les microtubules. Cependant, même si les PHF sont une forme hyperphosphorylée de la protéine tau, on ne peut pas affirmer de façon univoque que phosphorylation et agrégation soient liées. La même Ser262 qui, lorsqu'elle est phosphorylée provoque la dissociation de tau des microtubules, protègerait de l'agrégation. De plus, il est possible que l'hyperphosphorylation soit postérieure au phénomène d'agrégation qui empêcherait l'action des phosphatases mais n'empêcherait pas l'action des kinases, plus petites.

Outre la phosphorylation, la présence de résidus prolines dans des motifs Ser/Thr-Pro sujets à la phosphorylation par les kinases dirigées contre les sites contenant des prolines (CDK5/p25, MAPK, GSK3 $\beta$ ...), pourrait représenter un second facteur impliqué dans l'agrégation de tau. Deux types de données suggèrent un rôle particulier de ces motifs phosphoSer/Thr-Pro dans l'agrégation de tau. Premièrement, des anticorps spécifiques des PHF reconnaissent

sélectivement des motifs Ser/Thr-Pro phosphorylés dans les PHF et non lorsque ces mêmes motifs phosphorylés sont présents au sein de la protéine tau normale, et inversement, des épitopes spécifiques de la forme phosphorylée normale ne reconnaissent pas ces motifs dans les PHF. C'est le cas par exemple des anticorps AT180 et TG-3, le premier reconnaît l'épitope Thr231 dans la forme normale de tau, l'autre dans les PHF. Deuxièmement, la peptidyl-prolyl *cis/trans* isomérase humaine Pin1, se lierait aux PHF par l'intermédiaire d'un site unique centré autour de la phosphoThr231-Pro232 et restaurerait l'activité altérée des PHF dans la liaison aux microtubules (Lu, Wulf et al. 1999).

Au niveau moléculaire, Pin1 est spécifique des motifs Ser/Thr-Pro phosphorylés. Pin1 est fortement impliquée dans le cycle cellulaire par sa spécificité pour des substrats phosphorylés et de nombreuses phospho-protéines reconnues par Pin1 sont des antigènes de l'anticorps mitotique monoclonal MPM-2 dirigé contre les phospho-protéines mitotiques présentes dans la plupart des cellules eucaryotes (Vandré, Davis et al. 1986). Cette spécificité concerne ses deux domaines : le domaine WW impliqué dans la liaison du substrat et le domaine catalytique qui isomérisent spécifiquement les liaisons phospho-Ser/Thr-Pro, d'où la spécificité de Pin1 pour les phospho-protéines mitotiques. L'implication d'un mécanisme de *switch* moléculaire au niveau des liaisons précédant une proline, et notamment dans les parties non structurées de protéines comme c'est le cas de la plupart des substrats de Pin1, n'est pas encore bien établie. Les interactions entre Pin1 et ses substrats phosphorylés ont été décrites pour quelques protéines au niveau moléculaire avec, parfois, la présence de sites multiples de reconnaissance : les sites Thr48/67 de la phosphatase Cdc25, les Ser63/73 du facteur de transcription c-Jun. En revanche, pour la protéine tau, un seul et unique site d'interaction, centré autour de la Thr231, serait nécessaire et suffisant pour permettre les interactions avec Pin1 (Lu, Wulf et al. 1999). Cependant, les données structurales concernant la liaison de Pin1, à travers son domaine de liaison WW, à des phospho-peptides montrent que les interactions sont limitées au dipeptide phosphoSer/Thr-Pro au niveau du substrat. Ces résultats ne concordent pas avec la sélectivité pour l'épitope Thr231 au sein de la protéine tau. Nous avons donc remis en cause la spécificité de la liaison de Pin1 au niveau du seul épitope Thr231 et étudié les interactions entre Pin1 et d'autres phospho-épitopes de tau, notamment les phospho-épitopes impliqués dans des motifs Ser/Thr-Pro, présents dans les PHF (Thr212/Ser214, Thr231/Ser235 et Ser422) mais également des phospho-épitopes mitotiques (Ser396 et Ser404).

***Méthodes de criblage des phospho-épitopes de tau impliqués dans l'interaction avec Pin1.***

Nos études de *binding* par RMN et fluorescence entre Pin1, ou son domaine WW isolé, et des phospho-peptides issus de la séquence de la protéine tau, ont permis de mettre en évidence une gradation des affinités des différents phospho-épitopes dans une gamme d'affinités faibles allant de 0.1 à 1 mM. Nous avons également tenté de caractériser ces interactions en milieu hétérogène avec des techniques comme le Biacore et l'ELISA mais ces expériences ont échoué en raison de la prépondérance des interactions non spécifiques nombreuses qui masquent les interactions faibles entre Pin1 et ses phospho-épitopes.

La RMN a permis de déterminer les sites d'interaction des peptides phosphorylés au niveau du domaine WW et du domaine catalytique de Pin1, et de distinguer des interactions différentes d'un peptide à l'autre notamment au niveau du site catalytique et de l'interface entre les deux domaines de Pin1. Si les variations d'affinité observées entre les peptides mono-phosphorylé en Thr212 et di-phosphorylé en Thr212/Ser214 sont faibles et peu significatives, l'étude plus détaillée des sites d'interaction permet de suspecter un effet moléculaire à plus large échelle au niveau de la protéine Pin1 induite par une phosphorylation supplémentaire en Ser214. La RMN a également fourni des informations sur la cinétique concernant l'association en complexe de Pin1 et ses phospho-épitopes et la dissociation des complexes. De plus, la RMN permet l'observation directe de l'isomérisation *cis/trans* des liaisons phosphoSer/Thr-Pro, intrinsèque et catalysée par le domaine enzymatique de Pin1. La fluorescence a permis de confirmer les données concernant l'affinité des phospho-épitopes pour Pin1 obtenues par RMN mais plus rapidement et avec l'utilisation de moins de matériel. Cependant, les données sont dépourvues d'informations structurales.

Le criblage de différents sites de phosphorylation de tau par ELISA avait montré une spécificité pour le site phosphoThr231 et une affinité élevée caractérisée par un  $K_D$  de 40 nM (Lu, Wulf et al. 1999). Nous ne sommes pas parvenu à reproduire ces résultats par ELISA, ni par résonance plasmonique surface, technique qui possède de nombreuses similitudes avec l'ELISA en ce qui concerne le contexte moléculaire de l'interaction. Nous avons dans un premier temps émis l'hypothèse d'une possible concentration locale en peptide à proximité du site de liaison du domaine WW ou du domaine catalytique, favorisée par la fixation du peptide sur une surface. La forte affinité obtenue par cette méthode dans l'étude de Lu et al. pourrait s'expliquer par l'additivité des interactions au niveau de chacun des domaines et, par conséquent, multiplication des constantes de dissociation. Nous avons envisagé de reproduire

les expériences par ELISA avec Pin1 entière et son domaine WW isolé avec différents peptides dans le but de vérifier ces hypothèses. Cependant, les interactions non spécifiques engendrées par la méthode elle-même ont masqué les faibles interactions entre Pin1 et les phospho-épitopes de tau. Ces résultats ont amené une autre question concernant la pertinence de telles interactions dans un contexte cellulaire, où de nombreux composés sont susceptibles d'interférer dans l'interaction. Nous avons en particulier étudié les interactions entre Pin1 et l'ATP, en raison de la phospho-spécificité du domaine WW et des concentrations importantes d'ATP dans les cellules et montré une absence de liaison de l'ATP par le domaine WW.

Le bilan que nous pouvons dresser concernant les différentes méthodes utilisées dans l'investigation des interactions entre Pin1 et des phospho-épitopes de tau est résumé dans le **Tableau 2.11** :

	<i>Données thermodynamiques (<math>K_D</math>)</i>	<i>Données cinétiques (<math>k_{on}</math>, <math>k_{off}</math>)</i>	<i>Gamme d'affinités accessibles</i>	<i>Remarques générales</i>
<b>RMN</b>	+	+	Faibles ( $K_D > 0.1$ mM)	Informations structurales (site d'interaction, modifications conformationnelles)
<b>Fluorescence/FRET</b>	+	-	Moyennes à élevées ( $K_D < 0.5$ mM)	Rapidité, sensibilité
<b>Biacore</b>	+	+	Elevées ( $K_D < 1$ $\mu$ M)	Interactions non spécifiques non négligeables
<b>ELISA</b>	+	-	Elevées ( $K_D < 1$ $\mu$ M)	Influence de la surface liant l'un des partenaires (effet de concentration locale, modification conformationnelle)

**Tableau 2.11** : Apport et limites des différentes techniques d'investigation des interactions biomoléculaires *in vitro*.

***Graduation dans l'affinité des différents sites phosphoSer/Thr-Pro de la protéine tau dans la liaison au domaine WW de Pin1.***

Nous n'avons pas observé une haute affinité pour le site unique phosphoThr231 (Lu, Wulf et al. 1999) mais de faibles affinités pour plusieurs sites de phosphorylation. L'importance de la dynamique de la phosphorylation a été discutée dans le contexte du cycle cellulaire (Deshaies and Ferrell 2001; Nash, Tang et al. 2001) où l'équilibre entre phosphorylation et déphosphorylation joue un rôle primordial dans la transition d'une phase à l'autre. Il a été proposé que six phosphorylations de la protéine Sic1 par le complexe Cln/CDK étaient

nécessaires pour activer le processus d'ubiquitination de Sic1 et sa dégradation consécutive, enclenchant ainsi l'entrée en phase S du cycle cellulaire. Cette étude suggère que les phosphorylations multiples pourraient jouer un rôle dans un processus de *switch* qui régulerait le passage d'une phase à une autre du cycle cellulaire. Cet effet pourrait être obtenu plus aisément par multiplication de sites de faibles affinités (« *velcro* ») plutôt qu'avec un seul site de haute affinité (« *glue* ») (Deshaies and Ferrell 2001). Ce processus de phosphorylations multiples peut être soulevé dans le contexte de la reconnaissance d'épitopes phosphorylés par Pin1, dont tous les substrats possèdent une région non structurée contenant plusieurs sites Ser/Thr-Pro phosphorylés correspondant au site reconnu par Pin1.

### ***Identification d'un nouvel épitope phosphorylé de tau impliqué dans l'interaction avec Pin1.***

L'épitope Thr212, avec une affinité de 100  $\mu\text{M}$ , s'est révélé plus affin que l'épitope Thr231 dont l'affinité a été déterminée à 380  $\mu\text{M}$ . Nous avons montré le rôle prépondérant du domaine WW dans la liaison du substrat et l'importance de la Thr212 dans la médiation des interactions Pin1-tau par deux techniques spectroscopiques distinctes. De plus, comme Pin1 active *in vivo* la déphosphorylation de la phosphoThr231 par PP2A (Zhou, Kops et al. 2000), et pas de la phosphoThr212, ce dernier pourrait représenter un point d'ancrage pour Pin1 localisé à proximité du site phosphoThr231. D'autre part, la graduation observée pour les différents phospho-épitopes confirme la non spécificité des interactions autour de la Thr231 en accord avec les données structurales montrant peu d'interaction avec les résidus environnant le motif phospho-Ser/Thr-Pro dans la séquence peptidique. L'affinité de 150  $\mu\text{M}$  du dipeptide phospho-Thr-Pro renforce le caractère limité des interactions au niveau du substrat à ce seul motif. Le fait que des sites d'interaction avec Pin1, comme la Thr212, puissent ne pas avoir été détectés lors de l'étude proposée par Lu et al., peut s'expliquer par la spécificité de phosphorylation de certaines kinases pour des sites particuliers. Dans cette étude, la protéine tau avait été phosphorylée *in vitro* par l'action du complexe Cdc2/cycline B capable de phosphoryler l'épitope Thr231, mais phosphorylant la Thr212 avec un faible taux de conversion de 0.1 unité de phosphate par site (Drewes, Lichtenberg-Kraag et al. 1992), ce qui expliquerait que l'épitope Thr212 n'ait pas été détecté initialement. Un tel problème ne se pose pas avec la synthèse de substrats peptidiques par synthèse en phase solide où le taux de phosphorylation par site est de 100%.

L'effet des phosphorylations multiples sur la liaison à Pin1 peut également être discuté, étant donné les résultats contradictoires que nous avons obtenus. La phosphorylation des épitopes Thr231/Ser235, reconnus par l'anticorps TG3 spécifique des PHF, induit une augmentation de l'affinité pour le domaine WW d'un facteur deux. En revanche, la phosphorylation de la Ser214 à proximité de phosphoThr212 n'a pas le même effet selon si l'on considère les interactions avec Pin1 entière ou son domaine WW isolé. La phosphorylation de la Ser214 diminue l'affinité du peptide di-phosphorylé pour Pin1 par rapport au peptide mono-phosphorylé en Thr212. Mais cette diminution n'est pas observée avec le domaine WW isolé. Des interactions plus importantes avec le site catalytique de Pin1 observées avec le peptide di-phosphorylé par rapport au peptide mono-phosphorylé pourraient être à l'origine de la perte de l'affinité observée. Avec une phosphorylation supplémentaire en Thr217, on récupère une affinité de 100  $\mu$ M, indiquant que le site Thr217 est également un site reconnu par Pin1.

Il a été montré que la phosphorylation pouvait, dans certains cas, modifier l'équilibre *cis/trans* et ralentir la vitesse d'interconversion dans les motifs Ser/Thr-Pro phosphorylés, et serait ainsi impliquée dans la régulation de la dynamique de ces motifs (Schutkowski, Bernhardt et al. 1998). Une précédente étude avait établi que, dans des substrats peptidiques, la phosphorylation des résidus Thr231 et Ser235 ne modifiait pas l'équilibre *cis/trans* des prolines en position +1 (Daly, Hoffmann et al. 2000), mais que la phosphorylation de la Thr231 pourrait affecter la vitesse d'isomérisation et inhiber la liaison aux microtubules (Lu, Wulf et al. 1999). Nous avons réalisé une étude conformationnelle des substrats peptidiques incluant le site Thr212 dans différents états de phosphorylation, avec une attention particulière pour la détermination des proportions de formes *cis* pour les prolines présentes dans la séquence, entre autre pour exclure la possibilité d'une proline totalement en *cis*, interdisant les interactions avec le domaine WW qui lie ses substrats de manière *trans*-spécifique (Wintjens, Wieruszski et al. 2001). L'attribution des formes *cis* de Pro ainsi que leur quantification a été réalisée sur un spectromètre à 600 MHz, sur des échantillons à une concentration de 2 mM et facilitée par l'utilisation d'une cryosonde. Malgré l'absence de structure des différents peptides, une hétérogénéité conformationnelle des Pro213, Pro216 et Pro218 a pu être détectée. En revanche, la proportion de Pro219 en conformation *cis* est trop faible pour être détectée et peut être estimée à moins de 1%. En comparant les populations relatives des peptides diversement phosphorylés, nous avons pu conclure que la phosphorylation n'induisait pas de modifications significatives dans l'équilibre *cis/trans* des prolines. Cependant, nos données ne nous permettent pas d'observer une éventuelle modification de la

vitesse d'interconversion *cis/trans* de ces prolines. De plus, la phosphorylation pourrait, à l'échelle de la protéine tau entière, induire des modifications conformationnelles locales, notamment autour des sites Ser/Thr-Pro, qui seraient à l'origine de l'agrégation sous forme de PHF observée pour la protéine hyperphosphorylée. Les données obtenues à l'échelle peptidique, ne permettent pas d'impliquer la phosphorylation dans ces mécanismes d'agrégation car elle ne modifie pas, de manière notable, les équilibres *cis/trans*. Des études devront être réalisées à l'échelle de la protéine entière pour permettre d'établir un éventuel lien entre phosphorylation, perturbations conformationnelles locales et agrégation.

Si les domaines WW et catalytique sont essentiels à l'activité biologique de Pin1 *in vivo* (Lu, Zhou et al. 1999; Winkler, Swenson et al. 2000), *in vitro* les deux domaines semblent pourtant relativement indépendants tant au niveau structural que fonctionnel. Un *mapping* des déplacements chimiques des résidus de Pin1 a montré que les résidus principalement affectés par le substrat sont ceux du domaine WW, mais on peut également observer des perturbations plus ou moins marquées au niveau de résidus du domaine catalytique, les plus affectés étant ceux localisés à l'interface avec le domaine WW. Les faibles perturbations au niveau du site actif, indiquent que le domaine catalytique interfère peu dans la liaison du peptide au domaine WW. Ces données concordent avec deux études menées par RMN, combinant un *mapping* des déplacements chimiques, l'évaluation des contacts NOE inter-domaines et des couplages dipolaires résiduels en milieu orienté, qui démontrent une faible « communication » entre les deux domaines dont l'ampleur dépend du substrat (Bayer, Goettsch et al. 2003; Jacobs, Saxena et al. 2003). Les variations des déplacements chimiques des résidus du domaine catalytique dénotent d'une interaction supérieure à 10 mM, difficilement quantifiable, même par RMN. Cependant, si on ne peut pas raisonnablement parler de liaison, une activité catalytique est envisageable. Comme nous avons pu quantifier des populations de conformères *cis* avoisinant 10% pour les Pro213 et Pro218, nous avons étudié l'activité catalytique de Pin1 sur un peptide tri-phosphorylé au niveau des sites Thr212-Ser214-Thr217 en ajoutant 50  $\mu$ M d'enzyme à une solution de 1.5 mM de peptide dans le D<sub>2</sub>O. De manière tout à fait inattendue, l'activité catalytique a été détectée au niveau du seul motif phospho-Thr212-Pro213 par l'apparition d'un pic d'échange entre les formes *cis* et *trans* du proton H $\alpha$  de la Pro213. Un pic d'échange similaire n'est pas observé au niveau du motif phospho-Thr217-Pro218, pour lequel pourtant, la population de conformères *cis* est deux fois plus importante. Ce dernier motif étant inclus dans un motif PTPP, une contrainte structurale pourrait empêcher l'action catalytique de Pin1. Le même phénomène est observé avec le

domaine catalytique isolé indiquant que le domaine WW n'est pas nécessaire à l'activité enzymatique *in vitro*. Cependant, le rôle d'un tel *switch* moléculaire catalysé par Pin1 au niveau d'une protéine non structurée et sur de faibles populations de formes *cis*, reste mal résolu et les connaissances actuelles ne permettent pas encore d'expliquer l'interaction fonctionnelle qui lie Pin1 et ses substrats, et tau en particulier. Dans les études de *binding* qui ont été décrites dans ce chapitre, la protéine tau se trouve dans un état libre, c'est-à-dire non liée aux microtubules. Dans la mesure où l'interaction avec les microtubules induit une structuration de la protéine tau, il est envisageable que des prolines avec de fortes proportions de formes *cis* apparaissent au niveau des motifs Ser/Thr-Pro phosphorylés par la biais de contraintes structurales imposées par l'état lié, et que ceux-ci soient les véritables substrats de Pin1 *in vivo*.

Nous avons validé la pertinence de l'épitope phosphoThr212-Pro213 à l'échelle de la protéine tau entière. Nous avons employé une technique de phosphorylation de la protéine tau qui a permis de phosphoryler entre autres les sites Thr231 et Thr212 par blocage de la phosphatase PP2A par action de l'acide okadaïque, dans des cellules COS transfectées de manière stable avec la protéine tau. Des interactions de la protéine tau phosphorylée *in vivo* avec Pin1 et son domaine WW isolé ont été identifiées par *pull-down* et comparées à la liaison des mutants T231A, T212A et T231A/T212A. Dans les quatre cas, des interactions avec Pin1 sont observées indiquant que les sites Thr231 et Thr212 sont impliqués dans la reconnaissance par Pin1, mais qu'il existe encore d'autres sites. L'intensité relative des signaux est difficilement interprétable car il est difficile de faire la distinction à ce niveau entre affinité et taux de phosphorylation de chaque site. Cependant, ces résultats appuient la théorie des phosphorylations multiples qui seraient nécessaires à Pin1 pour lier ses substrats, d'une part parce qu'il existe plusieurs sites de phosphorylation, dont les Thr212/231, et d'autre part parce que, dans les expériences de retard sur gel, Pin1 lie les protéines tau hautement phosphorylées. Ces résultats suggèrent également que l'affinité du complexe formé par les protéines Pin1 et tau isolé lors de l'expérience de *pull-down* serait plus élevée que les affinités déterminées pour les peptides phosphorylés. En effet, l'isolation de complexes par *pull-down* nécessite des affinités relativement élevées avec des  $K_D$  que l'on peut estimer inférieurs à 1  $\mu\text{M}$ , ce qui semble être contradictoire avec les affinités mesurées par RMN et par fluorescence, et suggère la possibilité d'une interaction avec une forme de haute affinité induite par les multiples phosphorylations.

Nous avons essayé d'aller plus loin dans la caractérisation des interactions entre Pin1 et la protéine tau au niveau moléculaire. Dans le chapitre suivant, nous nous sommes intéressés à des substrats plus complexes. Un fragment peptidique de 40 résidus comportant plusieurs sites de phosphorylation a été utilisé comme modèle pour déterminer l'influence de sites de phosphorylation distants impliquant les résidus Thr212 et Thr231. Dans la paire Thr212/Thr231, un même nombre de résidus sépare les deux sites que dans la paire Thr48/Thr67 de Cdc25. Dans les deux protéines, ces paires sont localisées dans des parties non structurées, impliquant une distance similaire entre les résidus. Ce paramètre structural pourrait se révéler particulièrement pertinent dans la régulation des interactions entre Pin1 et ses partenaires. Nous avons ensuite étendu l'investigation des interactions de Pin1 avec la protéine tau entière au travers d'un mutant mimant l'état hyperphosphorylé de la protéine.



## The Peptidyl Prolyl *cis/trans*-Isomerase Pin1 Recognizes the Phospho-Thr212-Pro213 Site on Tau<sup>†</sup>

Caroline Smet,<sup>‡</sup> Anne-Véronique Sambo,<sup>§</sup> Jean-Michel Wieruszeski,<sup>‡</sup> Arnaud Leroy,<sup>‡</sup> Isabelle Landrieu,<sup>‡</sup> Luc Buée,<sup>§</sup> and Guy Lippens<sup>\*,‡</sup>

*Institut de Biologie de Lille, Institut Pasteur de Lille, UMR CNRS 8525, BP 245, F-59019 Lille Cedex, France, and Institut de Médecine prédictive et recherche thérapeutique, INSERM U422, Place de Verdun, F-59045 Lille Cedex, France*

*Received August 19, 2003; Revised Manuscript Received December 12, 2003*

**ABSTRACT:** The interaction between the neuronal Tau protein and the Pin1 prolyl *cis/trans*-isomerase is dependent on the phosphorylation state of the former. The interaction site was mapped to the unique phospho-Thr231-Pro232 motif, despite the presence of many other Thr/Ser-Pro phosphorylation sites in Tau and structural evidence that the interaction site does not significantly extend beyond those very two residues. We demonstrate here by NMR and fluorescence mapping that the Alzheimer's disease specific epitope centered around the phospho-Thr212-Pro213 motif is also an interaction site, and that the sole phospho-Thr-Pro motif is already sufficient for interaction. Because a detectable fraction of the Pro213 amide bond in the peptide centered around the phospho-Thr212-Pro213 motif is in the *cis* conformation, catalysis of the isomerization by the catalytic domain of Pin1 could be investigated via NMR spectroscopy.

Tau is a microtubule-associated protein that occurs mainly in neurons (1). Although derived from a single gene, it can exist through alternative splicing as six isoforms that are developmentally regulated (2). Further variability is introduced by posttranslational phosphorylation, with a fetal isoform of Tau carrying significantly more phosphate groups in embryonic stages than the six Tau isoforms in adult brain (3). A decade of research has led to the discovery of a large number of kinases that can phosphorylate Tau on specific threonine (Thr) or serine (Ser) residues. Mapping of the phosphorylation sites was done with specific antibodies (4–6) or with protein chemistry techniques, involving mass spectroscopy and/or Edman degradation (3, 7, 8). It was found that many kinases belong to the proline (Pro)-directed protein kinase family, with an enzymatic activity for one of the 17 Ser/Thr-Pro motifs in full-length Tau. Other kinases, however, such as PKA<sup>1</sup> were found to phosphorylate Ser or Thr residues that are not followed by a Pro (9), whereas

GSK-3 $\beta$  can modify both Ser and Thr residues that are or are not followed by a Pro (10–12).

The main motivation for this extensive research was and remains twofold. First, phosphorylation was found to modulate the binding of Tau to the microtubules, with a lower affinity for the phosphorylated Tau, and might as such influence the dynamic equilibrium of the tubulin/microtubule system (13). Second, in the paired helical filaments (PHF), one of the hallmarks of Alzheimer's disease of which Tau is the main component, the protein is heavily phosphorylated (1). Hyperphosphorylation in the later life stages might reflect an erroneous reactivation of the cell cycle in the aging neurons, proposed on the basis of the agreement between AD and mitosis-related phosphorylation patterns (14–16). Still, the relationship between phosphorylation and aggregation is not clear, and certain phosphorylation sites (Ser214 and Ser262) were recently described as being protective against aggregation (17).

The phospho-epitope mapping by immunochemistry heavily relies on a number of specific antibodies that recognize (i) the protein in its unphosphorylated soluble form, (ii) certain phospho-epitopes, or (iii) the phosphorylated protein aggregated into PHF tangles as found in the diseased brain. Examples of the latter are the AT100 and PHF-27/TG3 antibodies that recognize the AD specific phosphorylation sites composed of the Thr212–Ser214 and Thr231–Ser235 epitopes, respectively (14, 18). Whereas those antibodies are commonly classified as being “conformation-dependent”, the precise conformational changes that they detect have still not been fully elucidated (19, 20). Because many of the phosphorylation sites are Pro-directed, one possible hypothesis is that the Pro conformation would change upon phosphorylation and/or aggregation. The Pro residue is indeed unique in the sense that the energetic barrier between its *trans* and *cis* conformation is significantly lower than for

<sup>†</sup> Part of this work was funded by a grant from the Génopole of Lille, by the INSERM, the Région Guadeloupe, and by the Institute for the Study of Aging (ISOA). The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais, the CNRS, and the Institut Pasteur de Lille.

\* To whom correspondence should be addressed. Telephone: +33/(0) 320871229. Fax: +33/(0) 320871233. E-mail: guy.lippens@pasteur-lille.fr.

<sup>‡</sup> UMR CNRS 8525.

<sup>§</sup> INSERM U422.

<sup>1</sup> Abbreviations: PKA, protein kinase A; GSK-3 $\beta$ , glycogen synthase kinase 3 $\beta$ ; AD, Alzheimer's disease; PP2A, protein phosphatase 2A; PHF, paired helical filaments; CDK, cyclin-dependent kinase; MAPK, mitogen-activated protein kinase; DMF, dimethylformamide; RP-HPLC, reverse phase high-pressure liquid chromatography; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; IPTG, isopropyl thiogalactoside; PBS, phosphate-buffered saline; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser exchange spectroscopy; EXSY, exchange spectroscopy; FRET, fluorescence resonance energy transfer;  $K_D$ , dissociation constant; TFE, trifluoroethanol.

all other amino acids, leading to a higher population of the *cis* form. In small peptides, this population is typically on the order of a few percent, but in the context of a folded protein, Pro residues can be found that are completely in the *cis* conformation.

Indirect evidence of a structural role of the proline conformation comes from the recent finding that Pin1, a prolyl *cis/trans*-isomerase that is essential for the cell cycle and highly conserved within all eukaryotic species (21, 22), interacts with Tau (23). Pin1 was reported to restore the microtubule binding capacity of Cdc2/cyclin B-phosphorylated Tau (23). This might result from a direct physical interaction between both partners, or alternatively, Pin1 might assist proline-directed phosphatase PP2A in its catalytic activity on phosphorylated Tau. Indeed, at least at the level of short peptide substrates, PP2A only catalyzes the dephosphorylation of a threonine when it is followed by a proline residue in the *trans* conformation (24). Finally, the large reservoir of phosphorylated Tau in the PHF tangles might lead to depletion of soluble Pin1, and as such interfere with Pin1's other functions. Because Pin1 is a key regulator of the mitotic transition, the earlier finding that Alzheimer's disease might be related to a reactivation of the cell cycle (14–16) is in agreement with a functional role for Pin1 in AD.

Pin1 is constituted by two domains (25) both essential for its *in vivo* activity: a WW binding domain (26) which binds specifically pThr/pSer-Pro motifs and a peptidyl-prolyl *cis/trans*-isomerase domain which can act catalytically to promote conformational changes in pThr/pSer-Pro bonds by accelerating the intrinsic slow *cis/trans* isomerization of these bonds at least on small peptidic substrates (27). When the interaction between phosphorylated Tau and Pin1 was reported, the authors identified a unique interaction site at the phosphorylated Thr231 residue. Combined peptide mapping and mutagenesis studies indicated that phosphorylation of this proline-directed site, which is phosphorylated significantly in both fetal and adult rat brains (3) but is also one of the phosphorylation sites detected on PHF Tau, is both necessary and sufficient for the interaction with Pin1. Thr231 can be phosphorylated by GSK-3 $\beta$ , but only after phosphorylation of Ser235 by CDK5 or MAPK (6, 28, 29). It can be directly phosphorylated by CDK5 in the presence of heparin (30), and with a low efficiency by the p35cdc2-cyclin A complex (31).

Structural details of the interaction of Pin1 with its various substrates come from several high-resolution X-ray and NMR structures. Pin1 in complex with peptide substrates derived from the C-terminal domain of the polymerase II, the CDC25 phosphatase, or Tau itself indicates a dominant role for the WW domain in the interaction, with an aromatic clamp formed by one of the Trp residues and a Tyr holding onto the Pro side chain, whereas Arg and Ser in the loop region connecting the two  $\beta$ -strands make hydrogen bonds with the phosphorylated side chain of the Thr (32, 33). The interaction seems, however, mainly limited to this precise dipeptide, raising the question of the structural basis for the reported selectivity. In a recent NMR study, the only conformational selectivity we could detect was at the level of the proline isomerization state, with the pThr-Pro moiety interacting only when the Pro is in the *trans* conformation (33).

On the basis of the discrepancy between this structural information and the reported uniqueness of the Pin1 interaction site on Tau, we decided to further investigate this aspect by NMR spectroscopy. As the Thr212 residue is one of the few proline-directed and PHF specific phosphorylation sites, we centered our peptide on this residue. Another reason for the choice of this precise peptide is that the two reported interaction sites on human CDC25C, Thr47 and Thr68, are separated by exactly the same number of residues as Thr212 and Thr231 of Tau (34). Although the sequences of both proteins between their respective two threonine residues do not align, they both are considered random coil polymers without well-defined structure, and their separation might therefore be the relevant structural factor. Moreover, as other sites in the direct vicinity of the Thr212 site can equally be phosphorylated (Ser214 and Thr217), we have first investigated the influence of phosphorylation on the conformation of the peptide, by performing a detailed NMR analysis of the non-, mono-, di-, and triphosphorylated peptide. Their interaction with Pin1 was studied at the level of binding to the WW domain in the full-length protein, and the conformational changes of the proline conformation induced by the catalytic domain of Pin1 were followed by exchange spectroscopy. It was found that the different peptides do bind the Pin1 protein even better than the initially reported pThr231-Pro232 motif, and that at least the conformational exchange of the pThr212-Pro213 bond from *cis* to *trans* or *vice versa* can be catalyzed by Pin1. Because these results do not agree with the previously reported uniqueness of the pThr231-Pro231 site, we have looked into the possibility of other interaction sites on Tau, and confirmed that a minimal pThr-Pro dipeptide already can bind with a good affinity. These combined results show that the interaction of Pin1 and Tau is not limited to one unique site, and pose the intriguing question of a functional cooperativity between the WW and catalytic domain of Pin1 while it interacts with the hyperphosphorylated Tau.

## MATERIALS AND METHODS

**Peptide Synthesis and Purification.** Peptides were obtained by solid phase synthesis with introduction of selectively phosphorylated residues using appropriate building blocks of side chain-protected phosphothreonine or phosphoserine, as described previously (33). A dansyl fluorophore was added as a last step in the peptide synthesis of the pT212 peptide, by reaction of 5 equiv of dansyl chloride and 10 equiv of diisopropylethylamine in DMF. After TFA cleavage, peptides were purified by RP-HPLC on a C18 Nucleosil column (Machee-Nagel, Duren, Germany) and eluted with an acetonitrile gradient. The homogeneity of the fractions was verified by MALDI-TOF mass spectrometry.

**Expression and Purification of Recombinant Pin1 and the Catalytic Domain (Pin1<sub>CAT</sub>).** Pin1 with an N-terminal histidine tag fusion was produced in *Escherichia coli* strain BL21(DE3) (Novagen), carrying the recombinant PIN1-pET28 plasmid (generous gift from M. Yaffe, Harvard University, Cambridge, MA). The recombinant strain was grown at 37 °C in an LB medium containing kanamycin (20 mg/L) until the OD<sub>600</sub> reached ~0.6. The cells were then harvested by centrifugation, and the pellet was resuspended in half the volume in M9 medium with <sup>15</sup>NH<sub>4</sub>Cl (Cambridge Isotope Laboratories, Cambridge, MA) as the nitrogen source

[6 g/L Na<sub>2</sub>HPO<sub>4</sub>, 3 g/L KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/L NaCl, 4 g/L glucose, 1 g/L <sup>15</sup>NH<sub>4</sub>Cl, 0.12 g/L MgSO<sub>4</sub>, 20 mg/L kanamycin, and MEM vitamin cocktail (Sigma)]. The culture was incubated at 37 °C for 1 h before induction with 0.5 mM IPTG followed by growth for 3 h at 31 °C. The cell pellet was resuspended in a lysis buffer [50 mM Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> (pH 8.0), 300 mM NaCl (buffer A), 10 mM imidazole, 1 mM DTT, 0.1% NP40, and a protease inhibitor cocktail (Roche)], and cell lysis was performed by sonication. The soluble extract was loaded on a nickel affinity column (Chelating Sepharose Fast Flow, Amersham Pharmacia Biotech). Unbound proteins were extensively washed with 20 mM imidazole in buffer A, and the protein of interest was eluted with 250 mM imidazole in buffer A.

The GST–Pin1<sub>CAT</sub> fusion protein was produced in the BL21(DE3) star strain using the pGEX-6P plasmid (Amersham Pharmacia Biotech). The production was performed in LB medium with 0.5 mM IPTG induction followed by incubation for 3 h at 31 °C before the cells were harvested. The pellet was resuspended in phosphate-buffered saline (PBS) [130 mM NaCl, 3 mM KCl, 10 mM NaHPO<sub>4</sub>, and 2 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.4)] containing 1% Triton X-100, 20 mM β-mercaptoethanol, and a protease inhibitor cocktail. GST fusion proteins were purified from the soluble extract by use of a GSTrap Chelating Sepharose column (Amersham) equilibrated in PBS (pH 7.2). Unbound proteins were extensively washed with PBS, and proteins were eluted after incubation with Precision Protease (Amersham) in 50 mM Tris (pH 8.0), 100 mM NaCl, and 4 mM EDTA at 4 °C for 20 h.

**NMR Spectroscopy.** NMR experiments were performed on a Bruker DMX spectrometer (Bruker, Karlsruhe, Germany) operating at 14.1 T and equipped with a cryogenic triple-resonance probe head. Spectral parameters were as specified in ref 33. Standard TOCSY and NOESY experiments, with mixing times of 60 and 400 ms, respectively, were obtained with 1 or 2 mM peptide solutions in 50 mM deuterated Tris (pH 6.3) (Cambridge Isotope Laboratories) and 100 mM NaCl at 293 K. <sup>15</sup>N-labeled Pin1 HSQC spectra were recorded on a 200 μM protein sample in 50 mM deuterated Tris (pH 6.3), 100 mM NaCl, and a 5 mM DTT/EDTA mixture at the same temperature. For the titration experiments, a 200 μM solution of Pin1 was added to the appropriate amounts of the lyophilized peptide. The chemical shift perturbations of individual resonances calculated with the relationship  $\Delta\delta$  (parts per million) =  $[\Delta\delta^2(^1\text{H}) + 0.2\Delta\delta^2(^{15}\text{N})]^{1/2}$  were used to derive the dissociation constants (33). For the observation of catalytic activity, 50 μM full-length Pin1 or its catalytic domain alone, each solubilized in D<sub>2</sub>O buffer, was added to a 1.5 mM solution in D<sub>2</sub>O of a triply phosphorylated peptide, and EXSY spectra with a mixing time of 400 ms were recorded.

**Fluorescence Spectroscopy.** FRET experiments were carried out on a PTI (Lawrenceville, NJ) fluorescence spectrometer by exciting tryptophan residues of Pin1 at 295 nm. A 20 μM solution of Pin1 in 50 mM Tris (pH 6.3) containing 100 mM NaCl and a 5 mM DTT/EDTA mixture was added to lyophilized aliquots of the peptide to obtain a full titration curve. When the dansylated peptides were used, the fluorescence quenching of the tryptophan indole groups by the dansyl moiety was assessed at 334 nm. The fluorescence

Table 1: Peptide Sequences Synthesized during This Study<sup>a</sup>

peptide	sequence
T231	KKVAVVRT <sub>231</sub> PPKSPSSAK
pT231	KKVAVVR <b>pT</b> <sub>231</sub> PPKSPSSAK
pT231-pS235	KKVAVVR <b>pT</b> <sub>231</sub> PPK <b>pS</b> <sub>235</sub> PSSAK
T212	SRSRT <sub>212</sub> PSLPTPPTR
pT212	SRSR <b>pT</b> <sub>212</sub> PSLPTPPTR
pT212-pS214	SRSR <b>pT</b> <sub>212</sub> <b>pS</b> <sub>214</sub> LPTPPTR
pT212-pS214-pT217	SRSR <b>pT</b> <sub>212</sub> <b>pS</b> <sub>214</sub> <b>LpT</b> <sub>217</sub> PPTR
pThr-Pro dipeptide	pTP

<sup>a</sup> Phosphorylated residues are in bold, and numbered according to their position in the 441-residue isoform. The peptides are represented in the text with the one-letter code, whereas the individual residues are represented with the three-letter code.

variation was used to calculate the *K<sub>D</sub>* as described previously (33).

## RESULTS

**NMR Characterization of the Peptide Epitopes.** Four different peptides were synthesized, corresponding to a nonphosphorylated sequence, and peptides containing the single pThr212 residue, pThr212/pSer214, and finally pThr212/pSer214/pThr217. As a control, we also synthesized sequences centered around Thr231, with phosphorylation of this latter site alone or of both Thr231 and Ser235 (Table 1).

Whereas assignment of such short peptides can now be routinely carried out according to the sequential assignment procedure based on combined TOCSY and NOESY spectra (38), the main problem for detecting and characterizing the minor *cis* forms proved to be less straightforward because of sensitivity reasons. We used 2 mM samples dissolved in 270 μL of aqueous buffer in H<sub>2</sub>O or D<sub>2</sub>O together with a cryogenic probe on a 600 MHz instrument to obtain good quality spectra of even the minor forms. The resulting spectra on all peptides show that phosphorylation has a large effect on the amide proton of the carrying Thr or Ser residue (39), making its ready detection from a one-dimensional spectrum feasible (Figure 1). The same phosphorylated residues followed by a Pro residue in the *cis* conformation show up as low-intensity peaks close to the H<sub>N</sub> frequency of the major pThr/pSer signals followed by a Pro in the *trans* conformation.

When we first analyzed the peptides centered around the Thr231 residue, we confirmed the previously reported results for this epitope (20). The *cis* isomer of the pThr231-Pro232 peptide bond is populated at a level of <3% in water, and the conformation of the peptide as determined by NOE spectroscopy is essentially random. For the peptides centered around Thr212, confirmation of the Pro conformation came from the NOESY spectrum recorded in D<sub>2</sub>O on both the non- and triphosphorylated peptides, where we observed specific NOE contacts between the H<sub>α</sub> (and H<sub>β</sub>) protons of Thr212 and the H<sub>α</sub> proton of *cis*-Pro213, and the same contacts between the corresponding protons of Thr217 and Pro218 (Figure 2). Other minor forms correspond to Pro216, which is preceded by a Leu, whereas no *cis* conformation could be detected for Pro219, despite the fact that its conformation could even be more constrained due to the preceding Pro residue.

Integration of the different NMR signals in the one-dimensional spectra or on the H<sub>D</sub>–H<sub>D</sub> NOE cross-peaks,

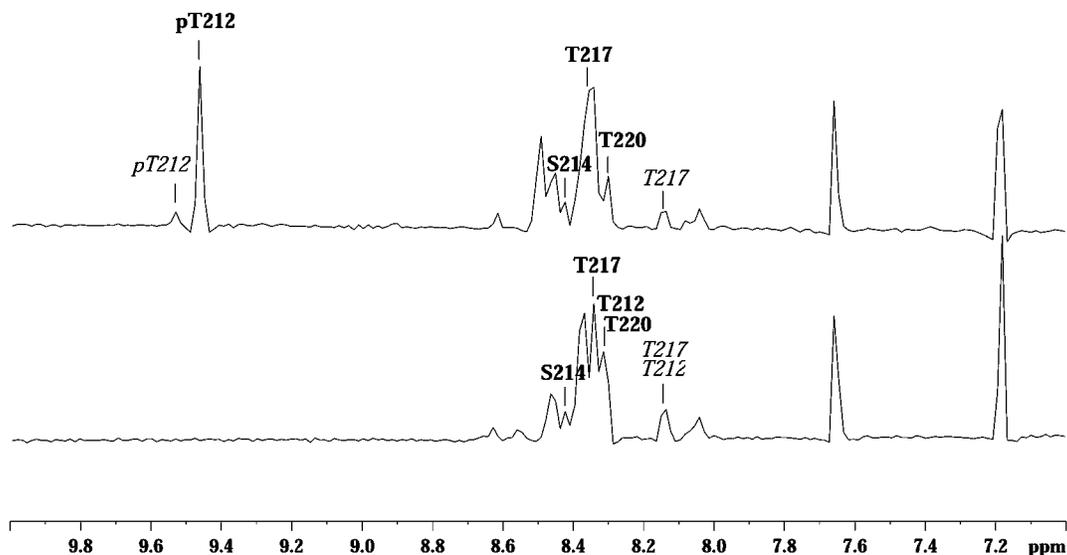


FIGURE 1:  $^1\text{H}$  NMR spectra (zoom of the NH region) of the T212 peptide phosphorylated at the Thr212 position or unphosphorylated. Only Thr212, Ser214, and Thr217 are annotated. Those in bold precede a *trans*-Pro form, and those in italics precede a *cis*-Pro form.

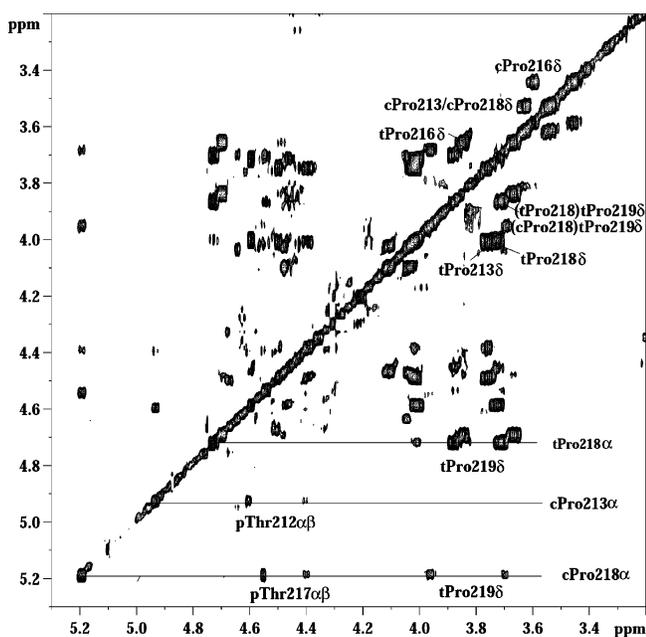


FIGURE 2: Assignment of the triphosphorylated T212 peptide on a 400 ms NOESY spectrum in  $\text{D}_2\text{O}$ .  $\text{H}\delta\text{-H}\alpha$  NOE contacts were assigned for *trans*-Pro isoforms and  $\text{H}\alpha\text{-H}\alpha$  for *cis*-Pro forms of Pro213 and Pro218.  $\text{H}\delta 1\text{-H}\delta 2$  NOE contacts are annotated for all *trans*-Pro forms, for *cis*-Pro213, *cis*-Pro216, and *cis*-Pro218 forms, and for *trans*-Pro219 preceded by *cis*-Pro218.

Table 2: Relative Populations of the *cis/trans* Ratio for the Different Pro Residues in the Non- or Triphosphorylated T212 Peptide

peptide	<i>cis/trans</i> ratio (%)			
	P213	P216	P218	P219
T212	6	8	12	<1
pT212-pS214-pT217	7	8	13	<1

where a reasonable dispersion exists, allowed the estimation of the relative populations of the *trans* and *cis* conformers (Table 2). Importantly, it was found that even a triple phosphorylation did not significantly change the relative populations.

**Interaction Mapping of the Peptides with Pin1.** Rather than using the isolated WW domain as we reported in a previous NMR study (33), we decided to work directly with the full-length  $^{15}\text{N}$ -labeled Pin1 protein. Near-complete assignments for this protein were previously reported (40), enabling a direct identification of the WW resonances in the full-length protein. When we titrated the phosphopeptide centered around the Thr231 epitope against a constant concentration of full-length Pin1, we found a nearly identical value for the dissociation constant. A similar titration experiment with a nonphosphorylated peptide did not yield detectable binding. When we repeated the experiment with the pT212 peptide, carrying a single phospho group on the Thr212 side chain, we detected a 3-fold stronger interaction, confirming that pThr231 is neither a unique nor even the best binding epitope. Resonances on the Pin1 protein that shifted upon addition of the pT212 peptide were mainly assigned to the WW domain, with the strongest shifts for the same residues (Ser16–Gly20 and Gln33–Glu35, with a very pronounced variation for the Trp34  $\text{HN}\epsilon 1$  side chain moiety) that were identified in the titration with the pT231 peptide (33). Small variations were equally detected for some resonances on the catalytic domain. A first category of those was assigned to residues at the interface of the WW and catalytic domain, including, for example, the Phe139 and Leu141 residues. On the basis of the chemical shift variation of their resonances as a function of the increasing peptide concentration, we derived a  $K_D$  constant similar to the one describing the interaction with the WW domain (Table 4). A second group of resonances that shift slightly upon addition of peptides are those of the catalytic site, including Met130, Leu122, Ser114, and Ser115. Because even at the 20-fold peptide excess used in our titration study, saturation of the chemical shift variations was not attained, we could not derive a reliable  $K_D$  value for the catalytic domain. This very weak interaction is in agreement with our findings for the interaction between phosphopeptides and the Pin1 protein of *Arabidopsis thaliana*, which consists only of a catalytic domain (41).

Table 3: Chemical Shift Values of the Triply Phosphorylated pT212-pS214-pT217 Peptide<sup>a</sup>

		H <sub>N</sub>	H <sub>α</sub>	H <sub>β</sub>		H <sub>γ</sub>		H <sub>δ</sub>	
				H <sub>β1</sub>	H <sub>β2</sub>	H <sub>γ1</sub>	H <sub>γ2</sub>	H <sub>δ1</sub>	H <sub>δ2</sub>
208	S	8.38	4.43	3.87					
209	R	8.5	4.44	1.91	1.79	1.67		3.21	
210	S	8.49	4.45	3.89					
211	R	8.48	4.37	1.91	1.79	1.67		3.23	
212 (t)	pT	9.45	4.50	4.40		1.38			
212 (c)	pT	9.52	4.61	4.41		1.27			
213 (t)	P	—	4.43	2.34	2.09	2.04	1.93	4.01	3.78
213 (c)	P	—	4.94	2.41	2.12	1.95	1.84	3.61	3.48
214	pS	9.10	4.44	4.11	3.98				
215 (t)	L	8.40	4.69	1.71	1.63	1.59		0.97	0.94
215 (c)	L	7.80	4.51	1.60		1.43		0.91	
216 (t)	P	—	4.49	2.31	2.03	1.89		3.87	3.68
216 (c)	P	—	4.60	2.41	2.13	1.95	1.85	3.62	3.48
217 (t)	pT	9.03	4.60	4.47		1.37			
217 (c)	pT	9.20	4.56	4.40		1.30			
218 (t)	P	—	4.72	2.4	2.09	2.05	1.9	4.01	3.75
218 (c)	P	—	5.20	2.52	2.17	2.00	1.86	3.62	3.54
219	P	—	4.55	2.36	2.09	2.05	1.93	3.89	3.71
220	T	8.29	4.32	4.22		1.22			

<sup>a</sup> Minor form assignments of residues in the peptide with a Pro in the *cis* conformation are given.

Table 4: Dissociation Constant Values Obtained by NMR for Various Phosphorylated Peptide Substrates against the Full-Length Pin1 Protein<sup>a</sup>

peptide	$K_D^{\text{PIN1}}$ (mM)	peptide	$K_D^{\text{PIN1}}$ (mM)
pT231	0.38 ± 0.1	pT212-pS214	0.16 ± 0.02
pT231-pS235	0.18 ± 0.05	pT212-pS214-pT217	0.10 ± 0.01
pT212	0.10 ± 0.03	pThr-Pro dipeptide	0.15 ± 0.03

<sup>a</sup> NMR dissociation constant values are based on chemical shift perturbations of several residues ( $N > 5$ ). The data for the dipeptide were obtained with the isolated WW domain.

To investigate the influence of multiple phosphorylations on the interaction, we performed similar titration experiments with three different peptides, including doubly phosphorylated Alzheimer-type epitopes pThr231/pSer235 and pThr212/pSer214. Amazingly, the effect on the  $K_D$  for Pin1 of a second phosphorylation is not the same for both peptides; the addition of a phosphate on Ser235 increases the affinity by a factor of 2, whereas phosphorylation of Ser214 decreases the affinity for Pin1. Addition of a third phosphorylation site to the latter peptide (Thr217) was necessary to restore the initial dissociation constant of 100  $\mu\text{M}$ .

The dissociation constant values for the pT212 peptide–Pin1 interaction were confirmed in an independent experiment based on the fluorescence variation of Pin1 upon titration in a dansylated pT212 peptide. Because of energy transfer from the tryptophan emitting at 334 nm to the dansyl group on the N-terminus of the pT212 peptide, the fluorescence of the tryptophans at 334 nm is efficiently quenched (Figure 4a). The resulting decrease in fluorescence intensity can more readily be quantified as the increase in the intrinsic tryptophan fluorescence that we had previously used with a nondansylated peptide as a substrate (42). The  $K_D$  value of 110  $\mu\text{M}$  confirms the data from the NMR titration experiment (Figure 4b), and demonstrates the absence of interference by the dansyl group.

**Catalytic Activity of Pin1.** The absence of a detectable amount of the *cis* form in the pThr231 peptide makes it a

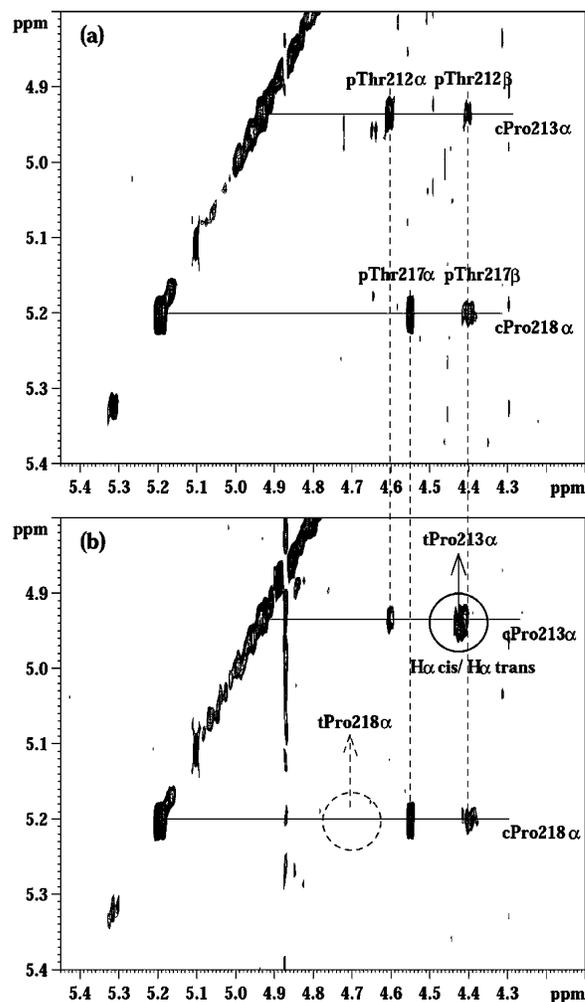


FIGURE 3: (a) Region around H<sub>α</sub> *cis*-Pro213 and Pro218 in the 400 ms EXSY spectrum of the triphosphorylated T212 peptide. Correlations involving pThr212 and pThr217 H<sub>α</sub> and H<sub>β</sub> minority forms are annotated. (b) Same region in the 400 ms EXSY spectrum of the triphosphorylated T212 peptide containing 50  $\mu\text{M}$  Pin1 showing an additional signal that correlated H<sub>α</sub> *cis* and H<sub>α</sub> *trans* of Pro213 (circled). The corresponding signal for Pro218 is not observed as indicated by the dashed circle.

less suitable substrate for the detection of a potential catalytic action of Pin1. Because the situation with the epitope centered around pThr212 is different, and because the use of the cryogenic probe had allowed the full assignment of all minor *cis* forms (Table 3), we used EXSY spectroscopy to probe the catalytic activity of Pin1. The experimental setup was similar to the one we used to demonstrate the catalytic activity of the *A. thaliana* Pin1 protein (35), with 400 ms EXSY spectra of the triply phosphorylated peptide in D<sub>2</sub>O with or without Pin1 added. Without Pin1, the isomerization proved to be too slow to give detectable cross-peaks connecting the signal of the same proton on a peptide in two conformations. When we added a catalytic amount of the protein, however, several additional cross-peaks appeared. The contact between the H<sub>α</sub> protons of Pro213 in the *cis* and *trans* conformations was easily identified (Figure 3), despite the presence of a weak contact between the H<sub>α</sub> proton of Pro213 in the *cis* conformation and the H<sub>β</sub> proton of pThr212. The structural transition of the Thr212-Pro213 amide bond was further confirmed by a contact between the H<sub>δ</sub> protons of Pro213 in the *cis* and *trans* conformations.

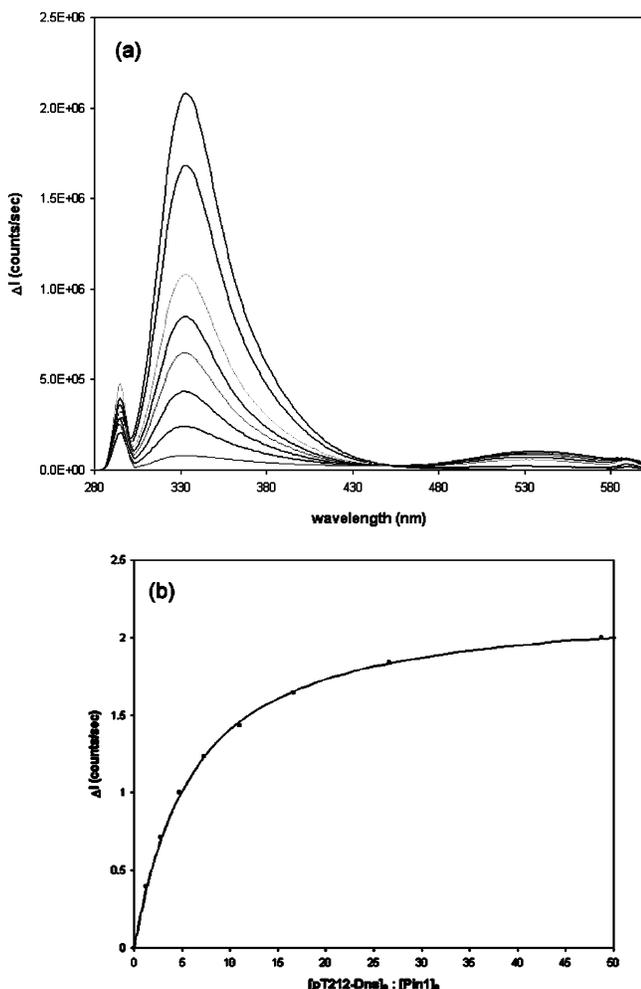


FIGURE 4: (a) Fluorescence titration of Pin1 by the pT212 Tau peptide coupled to a dansyl group at the N-terminus (pT212-Dns) involving FRET. Increasing amounts of the peptide were added to a 20  $\mu$ M solution of Pin1, resulting in a quenching of the fluorescence signal at 334 nm and a slight increase in the dansyl signal at 540 nm. (b) Graphical representation of the  $K_D$  calculation for the pT212 peptide based on the intensity variations of the fluorescence at 334 nm plotted against the molar ratio of the pT212 dansylated peptide to Pin1 concentration.

The isolated character of the H $\alpha$  proton of Pro218 in the *cis* conformation should make the observation of an exchange-mediated contact straightforward, but even with a long mixing time of 400 ms, we could not observe such a contact. This indicates that Pin1 selectively catalyzes the conformational change of the peptide bond between pThr212 and Pro213, but not the one between pThr217 and Pro218. When we performed the same experiment with the isolated catalytic domain of Pin1, very similar results were obtained.

## DISCUSSION

The phosphorylation of Tau has been the subject of intense studies. A first line of research was aimed at the mapping of the *in vivo* phosphorylated sites of fetal or adult soluble Tau (3) as well as the Tau found in the insoluble PHF aggregates (7, 8). Both immunochemical and protein chemistry techniques led to the conclusion that adult Tau carries significantly fewer phosphates than fetal Tau, and that phosphorylation on PHF Tau is even more pronounced than in fetal Tau. The dual character of the hyperphosphorylation stems from the fact that a larger number of phosphorylation

sites has been identified in PHF Tau than in fetal Tau, and second, the occupation level of the individual sites is higher in PHF Tau than in fetal Tau. An example of this latter factor is the Thr231 site, which is only phosphorylated to a level of 30% in fetal Tau, whereas no unphosphorylated Thr231 is found in PHF Tau (8).

Whereas phosphorylation of the Ser262 residue or Ser214 is known to promote detachment of Tau from the microtubules, and might hence lead to microtubule destabilization (13), a clear relationship between phosphorylation and aggregation has not been established. Recent results, for example, have demonstrated that phosphorylation of the same Ser262 site actually prevents aggregation (17), in apparent contradiction with the initial hypothesis that hyperphosphorylation is a causal factor for aggregation. The sole factor of phosphorylation might thus not be sufficient to explain the physicochemical changes that govern microtubule interaction or tangle formation. Additional evidence for this comes from the recent finding that the perturbed interaction between phosphorylated Tau and tubulin might be restored by the parvulin Pin1 (23). This latter enzyme belongs to the class of prolyl *cis/trans*-isomerases that catalyze the conformation of the prolyl bond. However, whereas the cyclophilins or FKBP binding proteins, the two main other classes of prolyl *cis/trans*-isomerases, are at least in yeast not essential for correct cell functioning (43), Pin1 or its yeast analogue Ess1 is crucial for correct progress through the cell cycle. Many of the recognition motifs of Pin1 coincide moreover with the epitopes of the monoclonal MPM2 antibody. This latter recognizes mitosis specific phosphoproteins, confirming further the role of Pin1 in the cell cycle.

Pin1 is characterized by a specificity at the substrate level, as it catalyzes the conformational change only when the proline is preceded by a phosphorylated Thr or Ser residue. The same motif is recognized by the second structural unit, the 40-amino acid WW domain, that seemingly promotes protein-protein interactions and is essential for the correct functioning of the protein (44). The interaction of Pin1 with a large number of phosphoproteins involved in the cell cycle has led to the hypothesis of a second conformational mechanism of posttranslational control after the covalent modification of phosphorylation. A proline *cis/trans* isomerization could be the structural switch, but direct evidence of such a mechanism with true protein targets has only very recently begun to emerge (45). Finally, it should be noted that the described interaction sites on several key mitotic proteins are in the unstructured regulatory parts of the protein. Examples are the Thr48 and Thr67 sites on the regulatory part of CDC25C (46, 47), or the Ser63 and Ser73 epitopes on c-JUN (48).

Because phosphorylated Tau extracted from the brains of people with AD is equally recognized by the mitotic MPM2 antibodies (49), Lu et al. (23) set out to study its interaction with Pin1. They discovered several intriguing aspects. First, Pin1 would be able to restore the microtubule binding capacity of Cdc2/cyclin B-phosphorylated Tau, and second, the interaction with phospho-Thr231 would be necessary and sufficient for this interaction. However, the observed specificity is not easy to explain at the structural level, as several studies have proven that the main interaction motif of the WW domain is the pThr-Pro dipeptide, with a minor contribution of the residue following the Pro (32, 33).

Moreover, although Thr231 is indeed phosphorylated to a higher level in PHF Tau, it can at least partially be phosphorylated in fetal Tau, raising the question of the conformational specificity of the Pin1–Tau interaction. For these reasons, we decided to further investigate the uniqueness of the interaction.

Of those sites that are unique to PHF Tau, two (Thr212 and Ser422) are proline-directed sites. We decided to first investigate the former site, because it is in the vicinity of the Thr231 residue, with the same distance from the latter as the two epitopes described on CDC25C (34). Moreover, because this Thr212 residue has at least two other phosphorylation sites in its immediate vicinity, we synthesized the peptides in different combinations of phosphorylation. NMR spectroscopy was used to investigate first the conformational aspects of those peptides, with specific attention being paid to the *cis/trans* ratios of the different prolines that are involved. Indeed, the absence of interaction between one of those epitopes and Pin1 could be due to the latter containing a Pro residue completely constrained in the *cis* conformation, where interaction with the WW domain is no longer possible (32, 33). A previous study with peptides centered around Thr231 had shown that the phospho-Thr–Pro bond in this peptide is predominantly in the *trans* conformation, with only 3% in the *cis* conformation. Addition of TFE, a solvent that is supposed to mimic the more hydrophobic environment of the inner protein, even further reduced this fraction (20). Therefore, it is not clear how a local structural change in such a small amount of the *cis* conformation containing chains could form the structural switch that explains the functional interaction between Pin1 and Tau. Of course, our results are obtained on a small peptide substrate and hence do not necessarily reflect the situation in the full-length protein. Still, recent NMR results of our group tend to confirm that the random coil chemical shift can be used efficiently to assign at least partially the full-length protein (51), confirming its unstructured nature and validating the peptide approach.

In our T212 peptide, the pool of peptide bonds in the *cis* conformation is somewhat larger, and this occurs not only at the level of the Thr212–Pro213 bond. Indeed, using a 2 mM sample and a measuring time of 12 h on a 600 MHz spectrometer equipped with a cryoprobe, we were able to distinguish the conformational heterogeneity at the level of the Pro213, Pro216, and Pro218 peptide bonds, and assign the different peptides in solution (Figure 2). On the basis of the integration of the one-dimensional proton spectrum and of the proline  $\delta 1$ – $\delta 2$  cross-peaks in the NOESY spectrum, we obtained the relative concentrations for the *trans* and *cis* peptide bonds preceding Pro213, Pro216, and Pro218 (Table 1). Only for Pro219 could we not detect a *cis* conformation, which limits its level to an estimated 1%. However, when comparing the relative populations in the differently phosphorylated peptides, we conclude that there is no clear modification of the *cis/trans* distribution with an increasing level of phosphorylation. As for the secondary structure of the peptides, the only noteworthy NOE contact that we observed was a medium-intensity contact between the amide protons of Ser214 and Leu215. This might indicate a kink in the peptide structure, but it remains to be seen whether this contact is an artifact of the short size of the peptide or whether it is truly present in the full-length protein.

The interaction of the different peptides with Pin1 was first investigated in a series of titration experiments against full-length Pin1. The  $^{15}\text{N}$ -labeled protein was used simultaneously to derive the dissociation constant from a fit of the chemical shift variations as a function of increased peptide concentration and to map those residues that directly interact with the peptide. A first titration experiment with the pT231 peptide resulted in a similar  $K_D$  value as previously determined against the isolated WW domain (33), confirming the primordial role of the WW domain as a protein–protein interaction module. When we performed a second titration series with a peptide phosphorylated at both Thr231 and Ser235, mimicking thereby the recognition site of the Alzheimer specific TG3 antibody, the affinity increased roughly 2-fold. We are at present using calorimetry to study whether this increased affinity results in a loss of entropy in the unbound state, or from a true increase in the enthalpy of binding.

The main result of our study came when we used the pT212 peptide in the same titration study. Indeed, the affinity of this peptide proved to be even better than that of the peptide centered on pThr231. Because this result casts doubt on the proposed uniqueness of this interaction site, we repeated the same measurement by fluorescence spectroscopy with a peptide modified with a dansyl group at its N-terminus. Energy transfer between Trp34 of the WW domain and the dansyl acceptor proved to be very efficient, and the resulting fluorescence decrease proved to be easier to interpret than the intrinsic fluorescence changes that we had previously used to confirm independently the NMR  $K_D$  value for the pT231–WW interaction (42). Despite the addition of a dansyl, the resulting  $K_D$  value confirmed our previous results derived by NMR.

Contrary to the case of the pT231 peptide, a second phosphorylation on the Ser214 site decreases the affinity by a factor of 2. We can rationalize this observation with the crystal structure of Pin1 in complex with a peptide derived from the C-terminal domain of polymerase II, where the hydroxyl Ser residue downstream of the phospho-Thr–Pro recognition motif also makes a hydrogen bond with the WW Trp34 side chain (32). In the pT212 peptide, we also find a Ser214 residue, and from the fact that the affinity of pT212 is higher than that of the pT231 peptide, where the Pro233 side chain does not offer the possibility of a hydrogen bond, we can safely assume that Ser214 indeed contributes to the increased affinity. Phosphorylation, however, would introduce a bulky phosphate group, which might be harder to fit than the simple hydroxyl group. Only when we create a second canonical phospho-Thr217–Pro218 interaction motif in the triply phosphorylated peptide did we find our initial dissociation constant of 100  $\mu\text{M}$ , indicating that both dipeptides are recognized with similar affinities. This latter result motivated us to measure the affinity of the isolated phospho-Thr–Pro dipeptide, devoid of any sequence context. The dissociation constant measured is on the same order of magnitude as that for the pT212 peptide, confirming that the main motif of interaction is indeed the dipeptide, and that the other amino acids only weakly affect this binding.

Because the interaction mapping studies were performed with the full-length Pin1 protein, we could also observe potential interaction sites with the catalytic domain. A first category of residues whose amide correlation shifted slightly

upon addition of the peptide includes residues at the interface between the WW and catalytic domain. Because the  $K_D$  derived from those chemical shift variations is identical to the one obtained for the WW domain, we conclude that it concerns the same binding event, and that binding of the peptide to the WW domain slightly interferes with the catalytic domain. This is in agreement with two recent NMR studies, where chemical shift mapping combined with interdomain NOEs and residual dipolar couplings led to the conclusion of possible cross-talk between the two domains, although the fine details seem to be substrate-dependent (50, 51). An independent interaction site, be it with a very weak affinity, was identified in the long loop that was found to contain the Ala-Pro dipeptide in the first crystal structure (25). However, as was the case for Pin1 from *A. thaliana*, only a lower bound of 10 mM could be determined for the affinity constant.

On the basis of the fact that both *cis* and *trans* conformations can be observed for Pro213 and Pro218 in this triphosphorylated peptide, we have investigated the Pin1 catalytic activity by adding 50  $\mu$ M enzyme to a 1.5 mM solution of the peptide in D<sub>2</sub>O. Unexpectedly, the only Pin1-dependent prolyl *cis/trans* isomerization detected through an *cis*-Ha-*trans*-Ha exchange cross-peak was for the Pro213 residue. A similar correlation peak was not observed for the Pro218 residue, although this motif is characterized by twice as much *cis* conformer as the peptide bond preceding Pro213. The presence of a Pro residue before and after the pThr217-Pro218 moiety possibly imposes a structural constraint such that the *cis/trans* interconversion can no longer be catalyzed by the Pin1 enzymatic domain.

Phosphorylation of the different Thr/Ser-Pro sites of the Tau protein is controlled by various well-known kinase-phosphatase complexes. Phosphorylation can be implemented among others *in vitro* with purified activated kinases such as the Cdc2-cyclin B complex used in the initial study identifying the Pin1-Tau interaction (23). However, this latter kinase was previously shown to incorporate only 0.1 phosphate unit per Tau molecule, even with an incubation time of 24 h (29). This latter factor might explain why the interaction with other sites such as the Thr212-Pro213 site described here initially went undetected, as the Cdc2-cyclin B complex probably only slightly catalyzes, if at all, the phosphorylation of this site (41). This is of course in sharp contrast to the peptide studies, where chemical synthesis allows for a 100% incorporation of a phosphate at any given position. As for the important question of whether T212 can interact with Pin1 in the full-length protein, preliminary pull-down results with recombinant Tau phosphorylated by a COS cell extract in the presence of the PP2A phosphatase blocking okadaic acid show that the T231A mutant and the T212A mutant still interact with Pin1 (A.-V. Sambo, M. Hamdane, and L. Buée, unpublished results). The double mutation (T212A/T231A) leads to a further decrease in the level of the retained protein, but still not a complete absence of interaction, suggesting the presence of even other interaction sites.

In conclusion, we have shown through careful peptide mapping that the phospho-Thr231-Pro232 dipeptide is not the only interaction motif between Tau and the cell cycle specific prolyl *cis/trans*-isomerase Pin1, and that the phospho-Thr212-Pro213 motif has an even higher affinity for the

Pin1 WW domain. The enzymatic activity of Pin1 on this latter site was confirmed through exchange NMR spectroscopy. Many questions about the molecular details of the Pin1-Tau interaction remain, however, and one of the most important is the possible cooperativity between both Thr212 and Thr231 sites, which could be a model for the action of Pin1 on other substrates such as Cdc25C. Finally, all our work is related to the interaction between Pin1 and Tau with the latter in its phosphorylated but free state. In the microtubule-bound state, however, we cannot exclude the possibility that one of several phosphorylated Thr/Ser-Pro motifs of Tau is forced into a *cis* conformation, and that those motifs are the true substrates for the prolyl *cis/trans*-isomerase activity of Pin1.

Future efforts will be directed to larger substrates encompassing both sites, in both their free and microtubule-bound state, and should be able to describe in further detail the intriguing functional role of the WW and prolyl isomerase domains of Pin1.

## REFERENCES

- Delacourte, A., and Buée, L. (1997) *Int. Rev. Cytol.* 171, 167–224.
- Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D., and Crowther, R. A. (1989) *Neuron* 3, 519–526.
- Watanabe, A., Hasegawa, M., Suzuki, M., Takio, K., Morishima-Kawashima, M., Titani, K., Arai, T., Kosik, K. S., and Ihara, K. Y. (1993) *J. Biol. Chem.* 268, 25712–25717.
- Biernat, J., Mandelkow, E.-M., Schröter, C., Lichtenberg-Kraag, B., Steiner, B., Berling, B., Meyer, H., Mercken, M., Vandermeeren, A., Goedert, M., and Mandelkow, E. (1992) *EMBO J.* 11, 1593–1597.
- Otvos, L., Feiner, L., Lang, E., Szendrei, G. I., Goedert, M., and Lee, V. M.-Y. (1994) *J. Neurosci. Res.* 39, 669–673.
- Goedert, M., Jakes, R., Crowther, R. A., Cohen, P., Vanmechelen, E., Vandermeeren, M., and Cras, P. (1994) *Biochem. J.* 301, 871–877.
- Hasegawa, M., Morishima-Kawashima, M., Takio, K., Suzuki, M., Titani, K., and Ihara, Y. (1992) *J. Biol. Chem.* 267, 17047–17054.
- Morishima-Kawashima, M., Hasegawa, M., Takio, K., Suzuki, M., Yoshida, H., Watanabe, A., Titani, K., and Ihara, Y. (1995) *Neurobiol. Aging* 16, 365–371.
- Brandt, R., Lee, G., Teplow, D. B., Shalloway, D., and Abdel-Ghany, M. (1994) *J. Biol. Chem.* 269, 11776–11782.
- Wang, J. Z., Wu, Q., Smith, A., Grundke-Iqbal, I., and Iqbal, K. (1998) *FEBS Lett.* 436, 28–34.
- Godemann, R., Biernat, J., Mandelkow, E., and Mandelkow, E.-M. (1993) *FEBS Lett.* 454, 157–164.
- Illenberger, S., Zheng-Fischhofer, Q., Preuss, U., Stamer, K., Baumann, K., Trinczek, B., Biernat, J., Godemann, R., Mandelkow, E. M., and Mandelkow, E. (1998) *Mol. Biol. Cell* 9, 1495–1512.
- Gustke, N., Steiner, B., Mandelkow, E. M., Biernat, J., Meyer, H. E., Goedert, M., and Mandelkow, E. (1992) *FEBS Lett.* 307, 199–205.
- Vincent, I., Rosado, M., and Davies, P. (1996) *J. Cell Biol.* 132, 413–425.
- Vincent, I., Zheng, J. H., Dickson, D. W., Kress, Y., and Davies, P. (1998) *Neurobiol. Aging* 19, 287–296.
- Delobel, P., Flament, S., Hamdane, M., Mailliot, C., Sambo, A.-V., Bégard, S., Sergeant, N., Delacourte, A., Vilain, J.-P., and Buée, L. (2002) *J. Neurochem.* 83, 1–9.
- Schneider, A., Biernat, J., von Bergen, M., Mandelkow, E., and Mandelkow, E. M. (1999) *Biochemistry* 38, 3549–3558.
- Hoffmann, R., Lee, V. M., Leight, S., Varga, I., and Otvos, L., Jr. (1997) *Biochemistry* 36, 8114–8124.
- Jicha, G. A., Lane, E., Vincent, I., Otvos, L., Jr., Hoffmann, R., and Davies, P. (1997) *J. Neurochem.* 69, 2087–2095.
- Daly, N. L., Hoffmann, R., Otvos, L., Jr., and Craik, D. J. (2000) *Biochemistry* 39, 9039–9046.
- Lu, K., Hanes, S., and Hunter, T. (1996) *Nature* 380, 544–547.

22. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) *Science* 278, 1957–1960.
23. Lu, P.-J., Wulf, G., Zhou, X. Z., Davies, P., and Lu, K. P. (1999) *Nature* 399, 784–788.
24. Stukenberg, P. T., and Kirschner, M. W. (2001) *Mol. Cell* 7, 1071–1083.
25. Ranganathan, R., Lu, K., Hunter, T., and Noel, J. (1997) *Cell* 89, 875–886.
26. Sudol, M., Bork, P., Einbond, A., Kastury, K., Druck, T., Negrini, M., Huebner, K., and Lehman, D. (1995) *J. Biol. Chem.* 270, 14733–14741.
27. Fischer, G., and Aumuller, T. (2003) *Rev. Physiol., Biochem. Pharmacol.* 148, 105–150.
28. Ishiguro, K., Shiratsuchi, A., Sato, S., Omori, A., Arioka, M., Kobayashi, S., Uchida, T., and Imahori, K. (1993) *FEBS Lett.* 325, 167–172.
29. Drewes, G., Lichtenberg-Kraag, B., Doring, F., Mandelkow, E. M., Biernat, J., Goris, J., Doree, M., and Mandelkow, E. (1992) *EMBO J.* 11, 2131–2138.
30. Paudel, H. K., and Li, W. (1999) *J. Biol. Chem.* 274, 8029–8038.
31. Vulliet, R., Halloran, S. M., Braun, R. K., Smith, A. J., and Lee, G. (1992) *J. Biol. Chem.* 267, 22570–22574.
32. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noël, J. P. (2000) *Nat. Struct. Biol.* 7, 639–643.
33. Wintjens, R., Wieruszeski, J.-M., Rousselot-Pailley, P., Drobecq, H., Lippens, G., and Landrieu, I. (2001) *J. Biol. Chem.* 276, 25150–25156.
34. Crenshaw, D., Yang, J., Means, A., and Kornbluth, S. (1998) *EMBO J.* 17, 1315–1327.
35. Landrieu, I., De Veylder, L., Fruchart, J. S., Odaert, B., Casteels, P., Portetelle, D., Van Montagu, M., Inze, D., and Lippens, G. (2000) *J. Biol. Chem.* 275, 10577–10581.
36. Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M. J., Delacourte, A., Michalski, J.-C., and Caillet-Boudin, M. L. (2003) *Biochim. Biophys. Acta* 1619, 167–176.
37. Mailliot, C., Bussiere, T., Caillet-Boudin, M. L., Delacourte, A., and Buée, L. (1998) *Neurosci. Lett.* 255, 13–16.
38. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
39. Bienkiewicz, E. A. (1999) *J. Biomol. NMR* 15, 203–206.
40. Jacobs, D. M., Saxena, K., Grimme, S., Vogtherr, M., Pescatore, B., Langer, T., Elshorst, B., and Fiebig, K. M. (2002) *J. Biomol. NMR* 23, 163–164.
41. Landrieu, I., Wieruszeski, J.-M., Wintjens, R., Inze, D., and Lippens, G. (2002) *J. Mol. Biol.* 320, 321–332.
42. Hamdane, M., Smet, C., Sambo, A.-V., Leroy, A., Wieruszeski, J. M., Delobel, P., Maurage, C. A., Ghestem, A., Wintjens, R., Bégard, S., Sergeant, N., Delacourte, A., Horvath, D., Landrieu, I., Lippens, G., and Buée, L. (2002) *J. Mol. Neurosci.* 19, 275–287.
43. Dolinski, K., Muir, S., Cardenas, M., and Heitman, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 13093–13098.
44. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science* 283, 1325–1328.
45. Andreotti, A. (2003) *Biochemistry* 42, 9515–9524.
46. Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) *Genes Dev.* 12, 706–720.
47. Wulf, G. M., Ryo, A., Wulf, G. G., Lee, S. W., Niu, T., Petkova, V., and Lu, K. P. (2001) *EMBO J.* 20, 3459–3472.
48. Kondratick, C. M., and Vandre, D. D. (1996) *J. Neurochem.* 67, 2405–2416.
49. Bayer, E., Goettsch, S., Mueller, J. W., Griewel, B., Guiberman, E., Mayr, L. M., and Bayer, P. (2003) *J. Biol. Chem.* 278, 26183–26193.
50. Jacobs, D. M., Saxena, K., Vogtherr, M., Bernado, P., Pons, M., and Fiebig, K. (2003) *J. Biol. Chem.* 278, 26174–26182.
51. Lippens, G., Wieruszeski, J.-M., Leroy, A., Smet, C., Sillen, A., Buée, L., and Landrieu, I. (2004) *ChemBioChem* 4, 100–105.

BI035479X



## **CHAPITRE III**

**Tau, du peptide à la protéine. Etude des interactions avec la peptidyl-prolyl  
*cis/trans* isomérase Pin1.**



## **I – Etude d'un fragment peptidique de 40 résidus comprenant deux sites d'interaction avec Pin1.**

Nous avons montré dans le chapitre précédent que Pin1 possédait plusieurs sites de reconnaissance d'affinités variables au sein de la protéine tau. Le *mapping* peptidique de plusieurs sites phosphoSer/Thr-Pro a déterminé que ces motifs étaient des substrats de faible affinité du domaine WW. Le *pull-down* réalisé sur la protéine tau phosphorylée a montré que Pin1 reconnaissait d'autres sites de phosphorylation que les sites Thr212 et Thr231, et qu'elle liait les fractions de protéines tau les plus hautement phosphorylées avec, vraisemblablement, une affinité élevée.

Nous sommes allés plus loin dans la compréhension des phénomènes qui régissent la reconnaissance des substrats par Pin1. Il a été proposé que Pin1 pourrait jouer un rôle régulateur dans le cycle cellulaire en modulant les processus de phosphorylation/déphosphorylation par isomérisation des liaisons Ser/Thr-Pro après leur phosphorylation, introduisant ainsi un degré de variabilité supplémentaire pour contrôler le passage d'une phase à l'autre. Le rôle du domaine WW n'est pas clair dans ce scénario. Sa présence est essentielle pour assurer la fonction de Pin1 *in vivo* (Lu, Zhou et al. 1999; Winkler, Swenson et al. 2000; Lu, Zhou et al. 2002), cependant, les homologues de Pin1 chez les plantes ne possèdent pas de domaine WW (Landrieu, De Veylder et al. 2000; Landrieu, Wieruszeski et al. 2002). D'ailleurs, Pin1 peut être active *in vivo* en l'absence du domaine WW en condition de surexpression du domaine catalytique (Lu, Zhou et al. 1999). D'autre part, la phosphorylation du domaine WW (en Ser16) est impliquée dans une régulation négative de la liaison de Pin1 à ses substrats et induit une localisation diffuse de Pin1 au niveau cellulaire, similaire à celle du domaine catalytique isolé, alors que Pin1 non phosphorylée est localisée principalement au niveau nucléaire. Le domaine WW pourrait intervenir dans l'ancrage de Pin1 au niveau d'un site phosphoThr/Ser-Pro dans une protéine et augmenter ainsi la concentration locale du domaine catalytique autour d'un motif phosphoThr/Ser-Pro (identique à celui lié par le domaine WW ou un autre) par analogie avec les conditions de surexpression du domaine enzymatique isolé. Cet effet de concentration locale a été envisagé au travers des phosphorylations multiples pour expliquer la reconnaissance de Sic1 par la protéine Cdc4 du complexe SCF après six étapes de phosphorylation (Deshaies and Ferrell 2001). Dans le cas de Pin1, la présence de deux sites de reconnaissance des substrats pourrait impliquer la liaison simultanée de deux motifs phosphoThr/Ser-Pro au sein d'une même protéine avec une haute

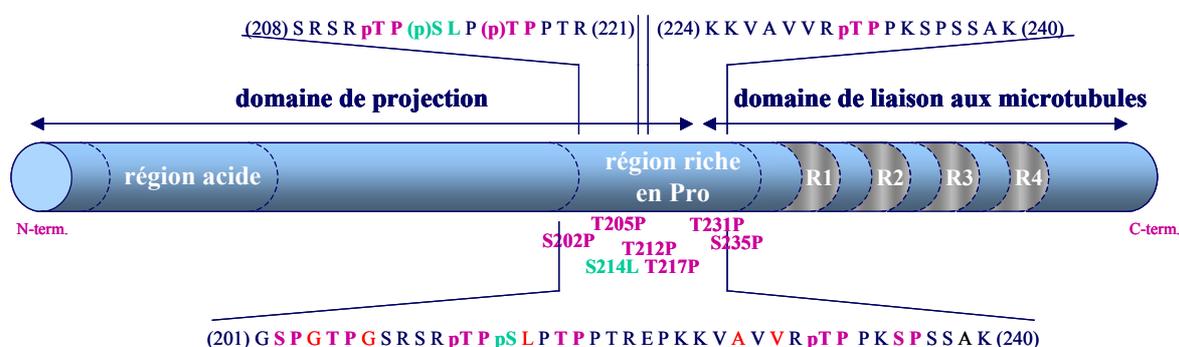
affinité par addition des énergies de liaison (la liaison du deuxième motif étant favorisée par la liaison du premier), bien que le domaine catalytique lie les substrats avec une très faible affinité. Nous avons envisagé que la flexibilité des deux domaines, conférée par le fragment peptidique inter-domaine de 10 résidus, puisse permettre à chacun d'entre eux de lier un motif phosphoThr/Ser-Pro de la protéine tau, à condition que les deux sites soient localisés sur la protéine à une distance qui autorise l'interaction avec les deux domaines. D'autre part, la distance de 19 résidus séparant les sites Thr212 et Thr231, est retrouvée dans les deux motifs phosphoThr-Pro (Thr48 et Thr67) de la phosphatase Cdc25, également reconnus par Pin1. Les deux sites étant localisés dans des régions non structurées, cette distance de 19 résidus devrait être équivalente dans les deux protéines et pourrait représenter un facteur structural au niveau du substrat permettant sa reconnaissance par Pin1.

La structure de la peptidyl-prolyl *cis/trans* isomérase Pin1 soulève de nombreuses questions concernant le mode de fonctionnement des deux domaines dans la reconnaissance des diverses phospho-protéines. Quelle est leur fonction intrinsèque ? Existe-t-il une association synergique de leur fonction respective pour assurer la liaison du substrat ou l'activité enzymatique ? Ont-ils un rôle antagoniste ? Quel est le rôle du fragment peptidique flexible reliant les deux domaines ? La flexibilité et la relative indépendance des deux domaines jouent-elles un rôle clé dans la fonction de Pin1 ? Les expériences décrites dans le chapitre II, avec l'utilisation de substrats peptidiques courts, n'ont pas permis de répondre à ces questions, en particulier parce que la taille des peptides ne rendait pas compte des phénomènes qui peuvent exister à l'échelle de la protéine tau entière. Dans ce chapitre, le but est de mettre en évidence une éventuelle coopération entre les deux domaines de Pin1 en présence d'un substrat peptidique plus long comprenant deux sites d'interaction.

### **1.1 – Description du modèle peptidique.**

Nous avons obtenu par synthèse peptidique en phase solide un peptide de 40 acides aminés comportant 3 sites de phosphorylation en position Thr212, Ser214 et Thr231, représentant une grande partie de la région riche en prolines de la protéine tau. Nous avons également introduit 5 résidus enrichis en <sup>15</sup>N à des endroits « stratégiques », en position Gly204/207, Leu215, Ala227 et Val229, pour faciliter la détection de prolines en conformation *cis* (Schéma 3.1). Le peptide a été synthétisé selon une stratégie classique, décrite pour les peptides courts du chapitre II, avec un simple couplage avec 10 équivalents d'acides aminés pour le fragment [K224-K240], et en double couplage pour le fragment [G201-P223] avec deux fois 5

équivalents d'acides aminés, pour prévenir d'éventuels problèmes de couplage liés à la taille du peptide. Les acides aminés enrichis en  $^{15}\text{N}$  étant plus onéreux et disponibles en quantités restreintes, nous avons procédé systématiquement pour ceux-ci à un double couplage avec deux fois 1.1 équivalents.

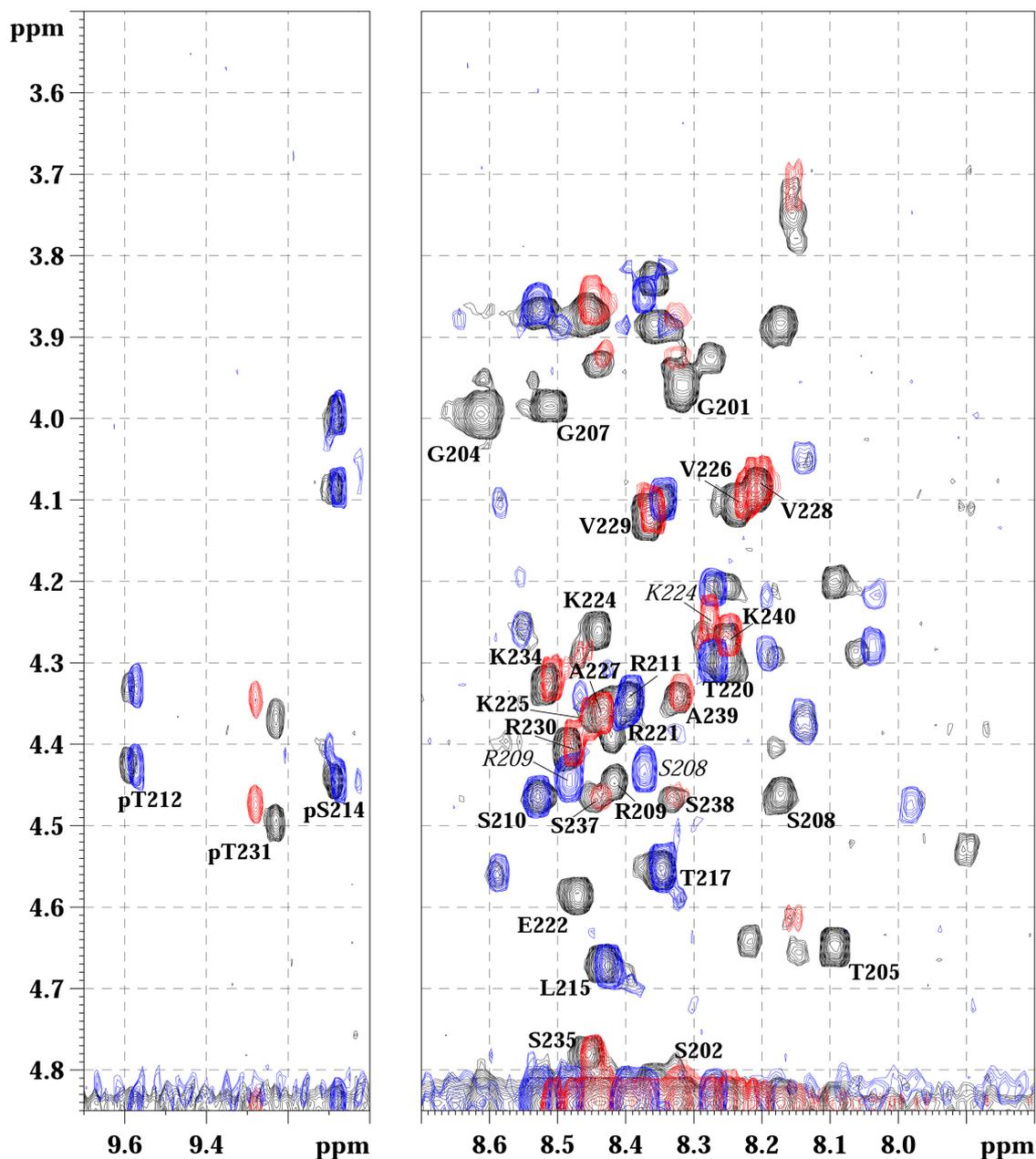


**Schéma 3.1** : Séquences primaires des peptides avec annotation des sites phosphoThr/Ser-Pro en caractères gras. La numérotation du premier et dernier résidu de la séquence peptidique, numérotés selon l'isoforme la plus longue, est indiquée en début et fin de séquence. Les résidus notés en rouge correspondent aux résidus enrichis en  $^{15}\text{N}$  incorporés dans la séquence du peptide long. Les peptides courts [208-221] et [224-240], diversement phosphorylés, ont été utilisés comme contrôle.

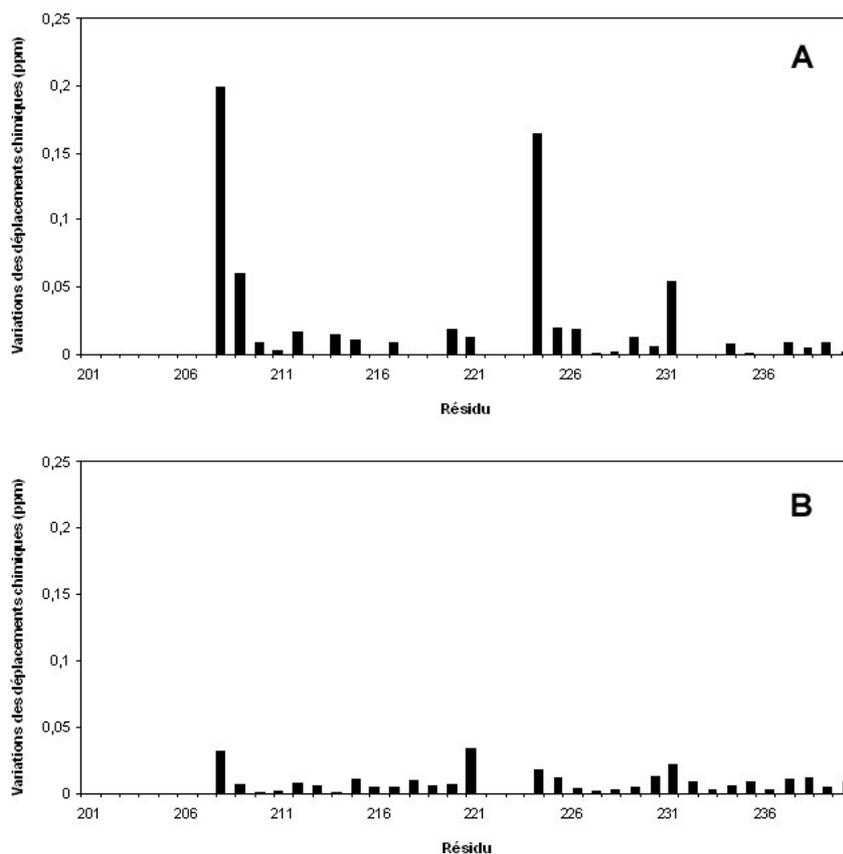
## 1.2 – Etude conformationnelle du fragment peptidique de 40 résidus de la région riche en prolines.

Nous nous sommes intéressés au préalable à la caractérisation conformationnelle du fragment peptidique de 40 résidus englobant les deux peptides courts, [S208-R221] di-phosphorylé en Thr212/Ser214 et [224-240] phosphorylé en Thr231, dont la taille permet la formation éventuelle de structures secondaires. Une séquence de 40 acides aminés représente également une taille limite pour une analyse structurale selon une stratégie RMN homonucléaire classique d'attribution des peptides. Les spectres homonucléaires TOCSY des deux peptides courts sont des sous-spectres du peptide long, notamment au niveau des corrélations qui s'établissent entre les protons  $\text{H}_\text{N}$  et ceux des chaînes latérales (Figure 3.1), indiquant un échantillonnage conformationnel similaire entre le peptide long et les deux peptides courts. Hormis pour les résidus S208, R209, d'une part, et pour le résidu K224 d'autre part, qui constituent le début de la séquence primaire (« effet de bord ») des peptides [208-221]pT212/pS214 et [224-240]pT231 respectivement, seules les résonances des résidus phosphorylés subissent de légères modifications de déplacements chimiques entre les peptides courts et le peptide long et ces perturbations concernent essentiellement les protons amide  $\text{H}_\text{N}$

(Graphique 3.1). Etant donné que seuls les résidus phosphorylés sont concernés, ces perturbations peuvent s'expliquer par de petites variations de pH entre les différents peptides car les spectres ont été effectués dans des conditions de pH proches de la seconde acidité ( $pK_a^2$ ) du phosphate (Tris deutéré 50 mM pH 6.4, NaCl 100 mM). Nous n'avons pas observé de modifications significatives des équilibres *cis/trans* pour chaque résidu Pro commun dans les différentes séquences (Tableau 3.1).



**Figure 3.1 :** Superposition des spectres TOCSY des peptides [201-240]pT212/pS214/pT231 (noir), [208-221]pT212/pS214 (bleu) et [224-240]pT231 (rouge) : détails des corrélations  $^1H_N$ - $^1H_\alpha$ . Les résidus S208, R209, K224 annotés en italique correspondent aux résonances dans les peptides courts [208-221]pT212/pS214 et [224-240]pT231. Seules les formes majoritaires correspondant aux résidus Pro en conformation *trans* sont annotées.



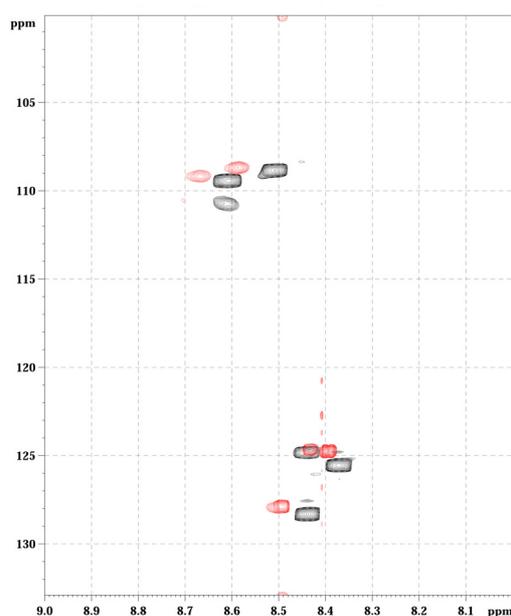
**Graphique 3.1 :** Variations de déplacements chimiques des résonances des peptides courts [208-221] pT212/pS214 et [224-240]pT231 par rapport au peptide [201-240] pT212/pS214/pT231 englobant les deux peptides précédents : (A) variations des résonances des protons H<sub>N</sub>, (B) variations des résonances des protons H<sub>α</sub>.

	Contacts NOE avec la (phospho)Thr (-1)			% conformère <i>cis</i>	
	H <sub>α</sub> (ppm)	H <sub>α</sub> (ppm)	H <sub>β</sub> (ppm)	long	court
<b>Pro213<i>cis</i></b>	4.941	4.604 (α/α)	4.413 (α/β)	9.6%	5%
<b>Pro218<i>cis</i></b>	5.106	4.372 (α/α)	4.049 (α/β)	9.8%	10.4%
<b>Pro232<i>cis</i></b>	5.155	4.494 (α/α)	4.409 (α/β)	4%	3%

**Tableau 3.1 :** Attribution des résidus Pro213/218/232 en *cis* par les contacts NOE avec les formes minoritaires des thréonines Thr212/217/231 (phosphorylées ou non). Proportions de conformation *cis* pour chaque proline dans le peptide long [201-240]pT212/pS214/pT231 et dans les peptides courts [208-221]pT212/pS214 et [224-240]pT231.

L'acquisition d'un spectre <sup>1</sup>H-<sup>15</sup>N HSQC à 20°C sur une solution de peptide à 1 mM a permis de détecter six corrélations <sup>1</sup>H-<sup>15</sup>N alors que seulement cinq résidus sont enrichis en <sup>15</sup>N. A 37°C, on n'observe plus que cinq résonances (Figure 3.2), suggérant un équilibre

conformationnel *cis/trans* dans la partie N-terminale du peptide. Le résidu Gly204 donne deux signaux à 20°C correspondant à des formes *cis* (30%) et *trans* (70%) de la Pro203. Le dédoublement des résonances de la Gly204 n'est pas observable sur le spectre homonucléaire car il ne concerne que la résonance de l'azote, plus sensible à l'environnement chimique, les résonances du proton H<sub>N</sub> dans les conformères *cis* et *trans* étant dégénérées. Cette importante proportion de conformère *cis* pour la Pro203 peut être favorisée par une flexibilité plus importante des extrémités peptidiques ainsi que par la présence de 3 résidus Gly parmi les 7 résidus N-terminaux.



**Figure 3.2 :** <sup>1</sup>H-<sup>15</sup>N HSQC du peptide pT212 pS214 pT231 à 20°C (noir) et à 37°C (rouge).

En résumé, aucune différence significative n'est observée entre les peptides courts et le peptide long car les déplacements chimiques sont sensiblement les mêmes indiquant un même échantillonnage conformationnel. Les formes minoritaires des peptides courts sont retrouvées dans le peptide long avec globalement les mêmes intensités indiquant que l'équilibre *cis/trans* des différents résidus Pro n'est pas modifié. On note une proportion de Pro203*cis* de 30% favorisée probablement par sa localisation à l'extrémité N-terminale du peptide.

### 1.3 – Etude de *binding* avec la peptidyl-prolyl *cis/trans* isomérase Pin1.

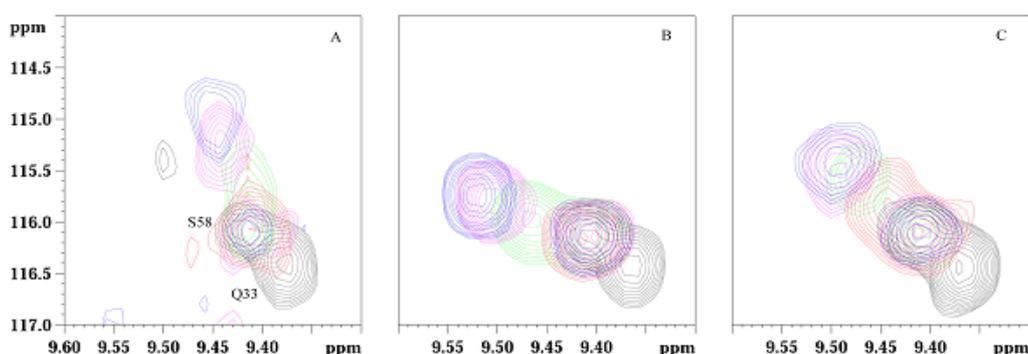
Nous avons évalué l'affinité du fragment peptidique de 40 résidus contenant les deux épitopes Thr212/Ser214 et Thr231 phosphorylés dont les constantes de dissociation des complexes formés avec Pin1 ont été déterminées pour les peptides courts à 160 et 380 μM respectivement (chapitre II).

***Hypothèse d'une coopération positive des deux domaines de Pin1 dans la liaison d'un fragment peptidique comprenant deux phospho-épitopes .***

Si les deux domaines lient chacun un épitope, les énergies de liaison pourraient s'additionner dans le cas le plus favorable, c'est-à-dire si les sites de chacun des partenaires sont placés de manière optimale. Avec des constantes de dissociation de l'ordre de 100  $\mu\text{M}$  pour le domaine WW et 10 mM pour le domaine catalytique, la liaison des deux épitopes simultanément devrait conduire à des constantes de dissociation de l'ordre de 1  $\mu\text{M}$  environ. Dans un cas moins favorable, des réarrangements structuraux seraient nécessaires avant l'interaction entre les deux partenaires, ce qui conduirait à un gain d'énergie moins important.

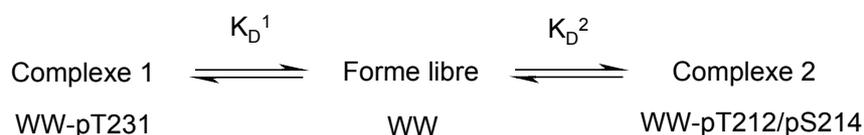
**1.3.1 – Etude avec la protéine Pin1 sauvage.**

La constante de dissociation du complexe formé entre Pin1 et le peptide [201-240]pT212/pS214/pT231 a été évaluée à 70  $\mu\text{M}$ . Cette valeur ne correspond pas à une liaison coopérative des deux épitopes par les deux domaines de Pin1 mais plutôt à une simple combinaison des affinités des deux épitopes isolés, avec la liaison au domaine WW d'un épitope à la fois. L'augmentation de l'affinité par rapport à celle des deux épitopes isolés provient en grande partie, d'une augmentation de la concentration effective en substrat qui correspond à 2 moles de dipeptide phosphoThr-Pro par mole de peptide. On note par ailleurs un comportement intermédiaire concernant la variation des résonances des résidus Q33 et N26 du domaine WW, induite par le peptide long par rapport aux deux peptides courts (Figure 3.3) qui confirme la liaison des épitopes pT212/pS214 et pT231 alternativement.

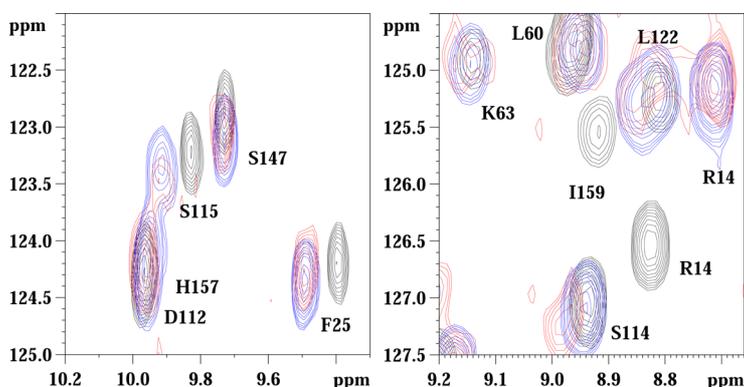


**Figure 3.3 :** Détail des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 en présence de concentrations croissantes de peptides [224-240]pT231 (A), [208-221]pT212/pS214 (B) et [201-240]pT212/pS214/pT231 (C). Les corrélations  $^1\text{H}$ - $^{15}\text{N}$  correspondent à la protéine Pin1 seule (noir) autour des résonances de S58 et Q33, en présence de 1 équiv. (rouge), 2 équiv. (vert), 5 équiv. (rose) et 20 équiv. (bleu) de peptide.

La dissociation des complexes étant rapide, le domaine WW échange très vite (de l'ordre de 1500 dissociations par seconde) entre la forme libre et les deux formes liées. On observe alors un signal moyen entre la forme libre du domaine WW, la forme liée à l'épitope pT212/pS214 et la forme liée à l'épitope pT231.



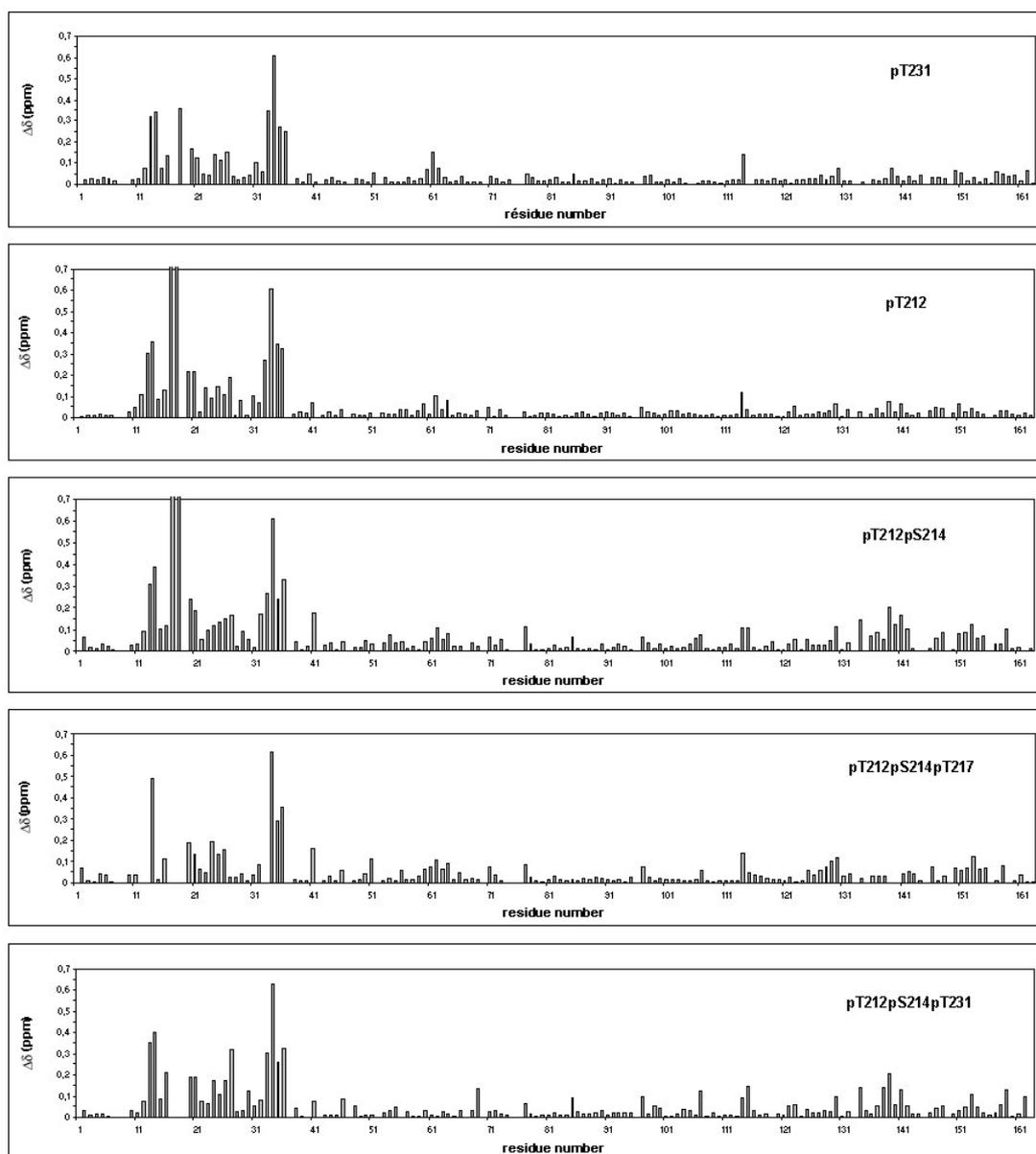
Le détail des perturbations induites par le peptide long établit un mode de liaison similaire aux peptides courts au niveau des résonances du domaine WW. Au niveau du domaine catalytique, on note des perturbations des résidus de l'interface inter-domaine assez comparables à celles induites par le peptide court [208-221]pT212/pS214 (Figure 3.5), c'est-à-dire un rapprochement plus marqué des deux domaines autour du substrat. On note également un élargissement plus important des résidus du site catalytique (Figure 3.4) que l'on retrouve plutôt avec le peptide [224-240]pT231.



**Figure 3.4** : Détails des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 seule (noir), en présence de 20 équiv. du peptide [208-221]pT212/pS214 (bleu) et en présence de 20 équiv. de peptide [201-240]pT212/pS214/pT231 (rouge) montrant les résidus du site catalytique (S114, S115, K63, L122) et les résidus du domaine catalytique localisés à l'interface avec le domaine WW (S147, I159) ainsi que des résidus du domaine WW (F25, R14).

Les variations du *mapping* des déplacements chimiques d'un substrat à l'autre sont faibles. On constate que les variations les plus importantes du domaine catalytique se situent au niveau des résidus de l'interface qui sont très peu perturbés en présence des peptides mono-phosphorylés mais qui sont plus marqués avec le peptide di-phosphorylé en T212/S214 et le peptide long. En revanche, avec le peptide tri-phosphorylé en T212/S214/T217 ont observe

des élargissements de raies plus importants, en particulier pour les résidus F139-L140 dans l'hélice  $\alpha_4$  et S16-S19 dans la boucle de liaison du phosphate du domaine WW.



**Figure 3.5 :** Mapping des déplacements chimiques des résidus de Pin1 en présence des peptides mono-phosphorylés [224-240]pT231, [208-221]pT212, du peptide di-phosphorylé [208-221]pT212/pS214 et des peptides tri-phosphorylés [208-221]pT212/pS214/pT217 et [201-240]pT212/pS214/pT231.

Les variations de déplacements chimiques induites par le peptide long ne permettent pas de détecter un comportement particulier des deux domaines de Pin1 en présence d'un substrat contenant deux phospho-épitoques. On note cependant un élargissement des résonances du site catalytique plus important qu'avec le peptide court [208-221]pT212/S214, semblable au

peptide [224-240]pT231. Un tel élargissement de signal peut être dû à une dissociation plus lente entre le substrat et le site actif ou à une dynamique différente du site catalytique de Pin1, en particulier de la boucle flexible qui adapte le phosphate.

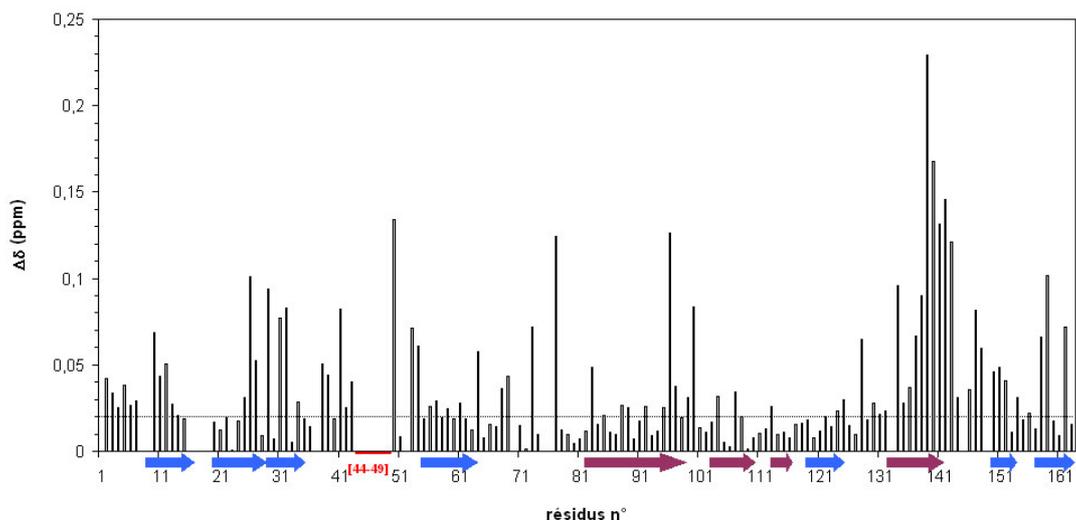
Le gain en affinité pour la liaison du peptide long comportant les deux phospho-épitopes Thr212/Ser214 et Thr231 est légèrement plus important qu'attendu si on considère une simple augmentation de la concentration en substrat due à la présence de deux motifs phosphoThr-Pro par molécule de peptide. On pourrait envisager que l'introduction de la phosphoThr231 restaurerait, pour le peptide di-phosphorylé, l'affinité de 100  $\mu$ M observée pour le peptide mono-phosphorylé en Thr212. En effet, dans le chapitre II, nous avons observé une diminution de l'affinité du peptide contenant la phosphoThr212 avec l'ajout d'une phosphorylation en Ser214 et suggéré une possible interaction avec le domaine catalytique qui serait à l'origine de cette diminution. De plus, avec le domaine WW isolé, nous avons remarqué plutôt une augmentation de l'affinité avec la phosphorylation de la Ser214, renforçant l'idée d'une perturbation de la liaison causée par le domaine catalytique. On pourrait envisager qu'une interaction du domaine catalytique préférentiellement en Thr231 restaurerait l'affinité de l'épitope Thr212/Ser214 di-phosphorylé pour le domaine WW malgré la phosphorylation en Ser214.

### **1.3.2 – Etude avec un mutant de Pin1 comportant une délétion de 6 résidus dans le *linker* inter-domaine.**

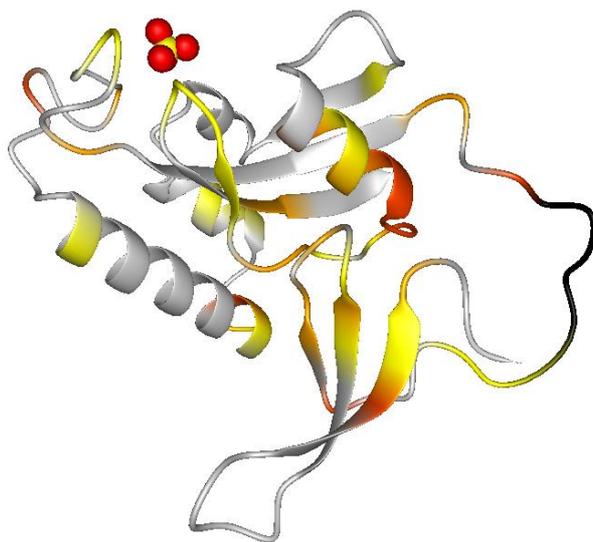
Comme les différences d'affinité et de mode de liaison sont peu significatives, nous avons tenté de mettre en évidence un effet structural avec un mutant de Pin1, **Pin1 $\Delta$ [44-49]**, comportant une délétion de 6 résidus dans le fragment peptidique qui relie les deux domaines. Si il existe effectivement un effet structural entre les deux épitopes Thr212/Ser214 et Thr231 qui permet d'adapter les deux domaines de Pin1 sur chacun des épitopes, la délétion dans le *linker* pourrait avoir un effet négatif en réduisant la mobilité des deux domaines l'un par rapport à l'autre et en raccourcissant la distance entre les deux sites d'interaction de Pin1.

#### ***Etude structurale du mutant Pin1 $\Delta$ [44-49].***

Nous avons déterminé l'impact d'une délétion dans le *linker* sur la structure du mutant à partir du *mapping* des déplacements chimiques de Pin1 $\Delta$ [44-49] (Figure 3.6).



**Graphique 3.2 :** Mapping des déplacements chimiques des résidus du mutant Pin1 $\Delta$ [44-49] comparés à la protéine sauvage.



**Figure 3.6 :** Mapping des déplacements chimiques des résidus de Pin1 représenté schématiquement sur la structure cristallographique de Ranganathan et al. (Ranganathan, Lu et al. 1997). Les variations de déplacements chimiques supérieures à 0.1 ppm sont représentées en rouge et les variations comprises entre 0.04 et 0.1 ppm en orange, les délétions en noir. Le sulfate de la structure cristalline indique la poche du site catalytique liant le phosphate.

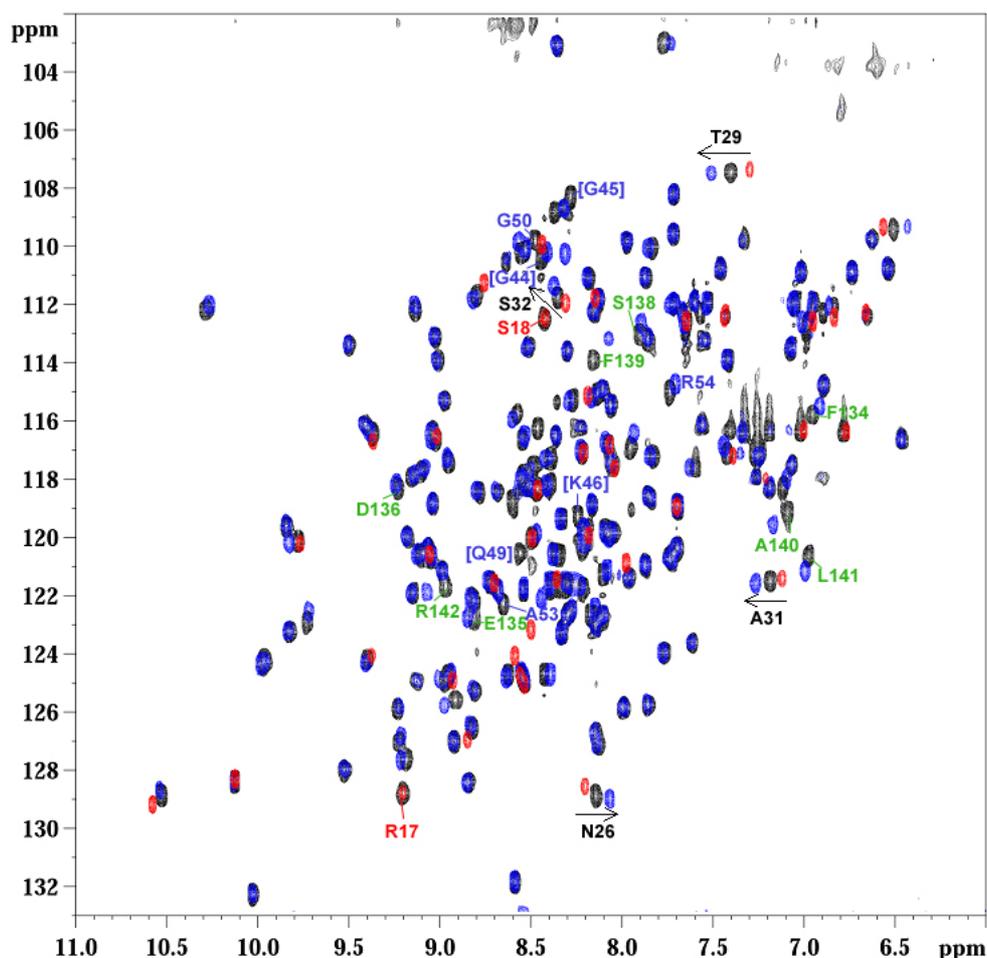
Les résidus de la boucle  $\beta$ -hairpin impliqués dans la liaison du phosphate du domaine WW (R17, S18) sont élargis dans le mutant indiquant une dynamique différente par rapport à la protéine sauvage ou le domaine WW isolé. La plupart des résidus du domaine WW sont affectés, notamment les résidus N26 à S32, localisés à l'interface du domaine catalytique. Le domaine WW étant un petit domaine constitué d'un triple feuillet  $\beta$  compact, les perturbations

des résonances des résidus à l'interface se propagent à l'ensemble du feuillet (Figure 3.6). Dans le domaine catalytique, la région la plus fortement affectée par la délétion concerne les résidus de l'hélice  $\alpha_4$  et du feuillet  $\beta_6$  constituant l'interface avec le domaine WW (Figure 3.6). L'amplitude de ces perturbations est équivalente à celles du domaine WW (Graphique 3.2) contrairement aux variations induites par les substrats peptidiques, pour lesquelles on observe toujours un forte prédominance des perturbations du domaine WW. On observe également un élargissement de certains résidus du site catalytique dans le mutant Pin1 $\Delta$ [44-49] qui sont impliqués dans le liaison du phosphate, en particulier les Arg68/69 et le Trp73.

Les perturbations touchant principalement l'interface entre les deux domaines, on peut envisager, concernant les modifications structurales qu'entraîne la délétion dans le *linker*, soit un éloignement des deux domaines ou, au contraire, un rapprochement. La comparaison des résonances des résidus N26, T29, A31 et S32 du domaine WW dans la protéine Pin1 sauvage, dans le mutant Pin1 $\Delta$ [44-49] ou dans le domaine WW isolé (Figure 3.7), indique que les déplacements chimiques relatifs, en valeur absolue, suivent l'ordre suivant (indiqué par les flèches dans la Figure 3.7) :

$$WW \text{ isolé} < Pin1 \text{ sauvage} < Pin1\Delta[44-49].$$

Si on considère que, dans le cas du domaine WW isolé, la distance entre les deux domaines de Pin1 est infinie, cet ordre indiquerait que, dans le mutant, les deux domaines soient plus proches que dans la protéine sauvage. Cependant, les perturbations restreintes des déplacements chimiques indiquent des modifications structurales faibles, du même ordre que celles induites par la délétion complète du domaine catalytique, suggérant que les deux domaines restent relativement indépendants l'un de l'autre.



**Figure 3.7 :** Superposition des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 sauvage (noir), du mutant Pin1 $\Delta$ [44-49] (bleu) et du domaine WW isolé (rouge). Les résidus annotés en rouge correspondent aux résidus de la boucle  $\beta$ -hairpin impliqués dans la liaison du phosphate du domaine WW. Les résidus annotés en bleu correspondent aux résidus localisés dans la zone de délétion du mutant quand ils sont notés entre crochets, ou proche du site de mutation. Les résidus de l'interface sont annotés en vert, pour les résidus du domaine catalytique, et, en noir, pour les résidus du domaine WW.

L'interprétation des variations de déplacements chimiques des résidus du domaine catalytique constituant l'interface est plus complexe. Si on compare la variation des résidus S138-R142 en présence d'un substrat de nature peptidique, on constate que les variations vont dans le sens opposé à celles du mutant. Mais, dans le cas des peptides, deux contributions conduisent aux variations des résidus de l'interface, l'interaction directe du peptide avec le domaine catalytique et l'influence de la liaison du peptide sur les interactions domaine-domaine. Ces deux effets sont indissociables sans le *mapping* du domaine catalytique isolé en présence du peptide. On note cependant, que les perturbations des résidus N26 et S32 induites par les peptides vont également dans le sens opposé à celles du mutant.

### ***Etude de binding avec la protéine mutante Pin1 $\Delta$ [44-49].***

L'impact de la délétion de 6 résidus dans le *linker* a été évaluée sur l'affinité du peptide de 40 résidus contenant les deux épitopes, Thr212/Ser214 et Thr231, reconnus par Pin1. La constante de dissociation a été déterminée à 70  $\mu$ M comme pour la protéine Pin1 sauvage. On note cependant, un élargissement plus marqué des signaux du mutant pour des concentrations élevées de peptide (10 et 20 équivalents).

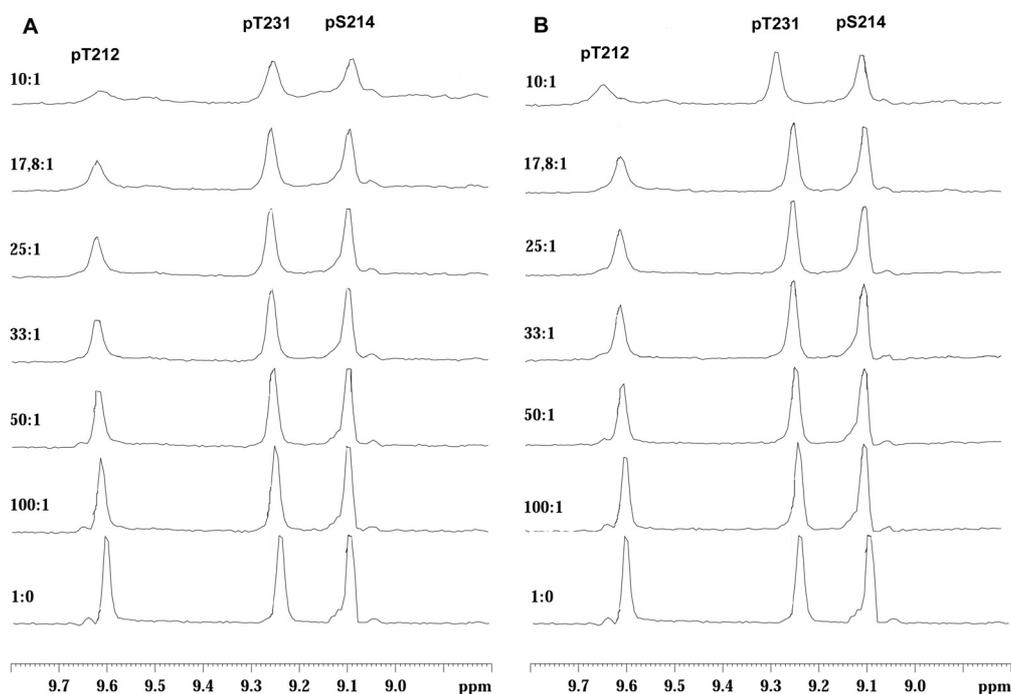
En résumé, les études de *binding* menées sur le fragment peptidique de la région riche en prolines de la protéine tau n'ont pas montré un effet clair sur le comportement de Pin1 face à la présence de deux sites de reconnaissance au sein du même peptide si on considère uniquement la liaison du substrat ( $K_D$ ). Une coopération entre les deux domaines de Pin1 peut être envisagée dont l'effet sur la liaison serait difficilement décelable, car on ne peut pas réellement parler d'une liaison du domaine catalytique mais plutôt d'une interaction transitoire. En regardant les perturbations de Pin1, on observe un effet global rassemblant l'effet de l'interaction avec le site Thr212/Ser214 et l'effet de l'interaction avec le site Thr231 alternativement. L'échange entre les diverses formes de la protéine (libre, en complexe avec le site Thr212/Ser214, en complexe avec le site Thr231) étant rapide par rapport à l'échelle de temps RMN, nous observons un comportement moyen qui peut conduire à l'élargissement des signaux que nous avons remarqué. Nous avons donc poursuivi l'investigation des interactions entre Pin1 et ce fragment de 40 acides aminés de la protéine tau en observant le comportement du substrat.

## **1.4 – Modulation de l'activité catalytique *cis/trans* isomérase sur l'épitope Thr212/Ser214 par la présence d'un second phospho-épitope.**

### **1.4.1 – Observation de l'interaction de Pin1 avec les sites phosphorylés du peptide.**

Nous avons tenté de caractériser l'interaction en observant le comportement du substrat en présence de différentes concentrations de Pin1 (Figure 3.8A) et du domaine WW isolé (Figure 3.8B). On observe un élargissement (et un déplacement) des signaux des résidus phosphoThr212, phosphoThr231 et phosphoSer214 en présence de Pin1 ou du domaine WW isolé. Ces perturbations indiquent une interaction de Pin1, et en particulier de son domaine WW, avec les motifs phosphoThr212-Pro213 et phosphoThr231-Pro232. En ce qui concerne les perturbations de la phosphoSer214, on peut envisager deux explications : soit Pin1

interagit directement avec la phosphoSer214 (comme le suggéraient les  $K_D$  mesurés pour le peptide di-phosphorylé en Thr212/Ser214), soit les perturbations sont liées à l'interaction de Pin1 avec la phosphoThr212 localisée à proximité. On remarque des perturbations plus importantes avec le site phosphoThr212 par rapport au site phosphoThr231 qui peuvent provenir d'une interaction de plus haute affinité. De manière générale, un élargissement plus important est détecté en présence de Pin1 par rapport au domaine WW et peut traduire plusieurs phénomènes : une modification de l'affinité, une modification du  $\Delta\delta_{\max}$  entre l'état libre et l'état lié ou une interaction avec un objet de taille plus importante.



**Figure 3.8 :** Interactions entre le peptide [201-240]pT212/pS214/pT231 et Pin1 entière (A) ou le domaine WW isolé (B). Le peptide est solubilisé à une concentration de 1 mM. La protéine est ajoutée à partir d'une solution concentrée à 500  $\mu$ M pour obtenir les concentrations relatives peptide : protéine indiquées à gauche des spectres  $^1\text{H}$ .

Là encore, la nature des interactions des deux domaines de Pin1 avec un substrat peptidique complexe comportant deux phospho-épitopes reste floue. Nous avons opté pour une autre stratégie, toujours basée sur l'observation du substrat, pour tenter d'éclaircir le mode d'action de Pin1. Nous avons mis en évidence une activité enzymatique sélective au niveau de l'épitope Thr212/Ser214 (chapitre II) en présence d'un autre phospho-épitope en Thr217. Nous avons envisagé que le mode d'action de Pin1 en présence de substrats multi-phosphorylés pouvait être décelé au travers de l'activité catalytique en Thr212/Ser214 et que

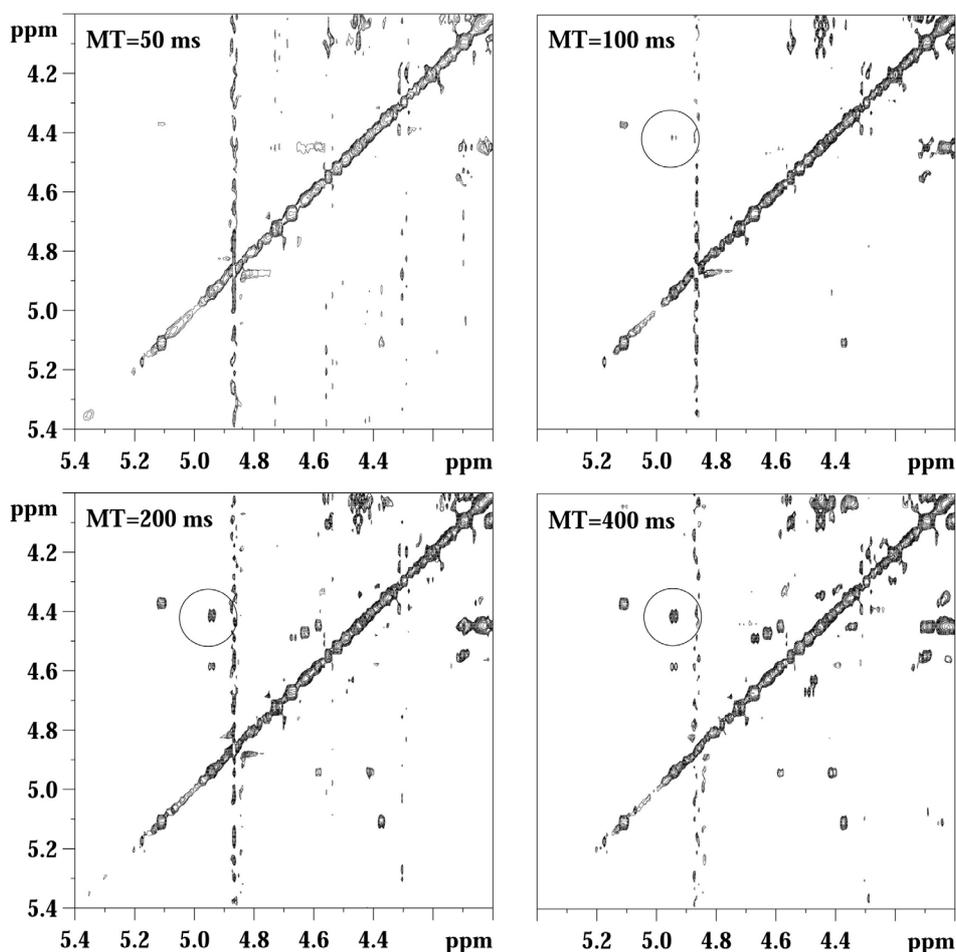
la présence d'un second site de phosphorylation en Thr231 au sein du même substrat, pouvait perturber l'activité enzymatique en Thr212/Ser214, de manière positive ou négative.

#### 1.4.2 – Spectroscopie d'échange appliquée à l'observation de l'interconversion *cis/trans* de la liaison phosphoThr212-Pro213.

La spectroscopie de type EXSY mesure un échange chimique, ici entre un proton présent sur une proline en conformation *trans*, associée à une fréquence de résonance  $\delta_{trans}$ , pendant un temps  $t_1$ , ensuite la proline change de conformation pendant le temps de mélange (ou *mixing time*, MT) et le proton se retrouve sur une proline en conformation *cis*, associée à une fréquence de résonance  $\delta_{cis}$ , pendant un temps  $t_2$ . En absence de catalyse, les signaux correspondant aux conformations *cis* et *trans* sont détectés distinctement, traduisant un échange lent entre les deux populations de l'ordre de la seconde. En présence de Pin1 ou du domaine catalytique isolé, on constate une augmentation de la vitesse d'interconversion *cis/trans* qui se traduit par un pic d'échange. Mais la vitesse d'échange reste toujours très lente par rapport à l'échelle de temps RMN car on observe toujours distinctement les signaux de la forme *trans* et les signaux de la forme *cis*. Nous avons mis en œuvre la technique dite du « *build up* » qui permet de mesurer la variation de l'intensité du pic d'échange en fonction du temps de mélange qui correspond au transfert de magnétisation entre l'impulsion et la détection. Plus le temps de mélange est long, plus il y aura de « sauts » entre la forme *cis* et la forme *trans* (et inversement), donc plus le pic d'échange sera intense (il en va de même pour les couplages dipolaires  $^1\text{H}$ - $^1\text{H}$  NOE), mais cet effet est contrebalancé par une perte d'aimantation par relaxation  $T_1$ . L'intensité du pic d'échange est donc proportionnelle à la vitesse d'échange. Nous avons typiquement utilisé des temps de mélange de 50, 100 200, 300, 400 et 500 ms, sur des échantillons à 3.5 mM dans un tampon Tris deutéré 50 mM pH 6.4, NaCl 100 mM dans le  $\text{D}_2\text{O}$  100%, ce qui correspond à des expériences d'environ 48 heures (Figure 3.9).

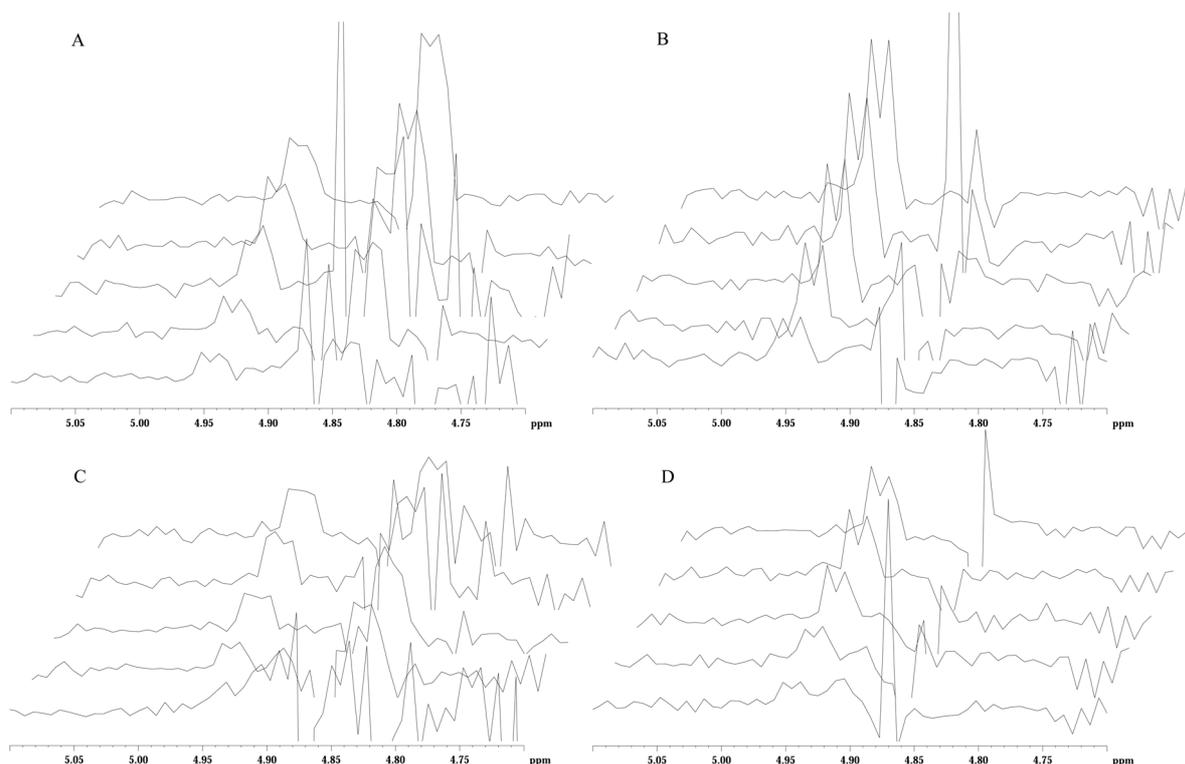
Une attribution des formes minoritaires a été réalisée sur le peptide long [201-240]pT212/pS214/pT231 pour la plupart des prolines de la séquence. Les équilibres *cis/trans* ont été déterminés pour les Pro213/218/232 en particulier (Tableau 3.1). On constate une proportion équivalente de conformations *cis* pour les Pro213 et Pro218 (10%) et une proportion plus faible pour la Pro232 (4%). Des proportions similaires de conformères *cis* ont été déterminées pour les Pro218 et Pro232 dans leurs peptides courts respectifs par rapport au peptide long, mais elle est deux fois plus faible pour la Pro213 (5%).

L'acquisition de spectres d'échange EXSY a été réalisée sur les peptides long [201-240]pT212/pS214/pT231 et court [208-221]pT212/pS214 en solution à 3.33 mM en présence de 0.033 mM Pin1 ou du domaine catalytique isolé. Pour les deux peptides, seul le pic d'échange entre les formes *cis* et *trans* de la Pro213 est observé traduisant une activité catalytique sélective en Thr212/Ser214 et aucune activité en Thr231.



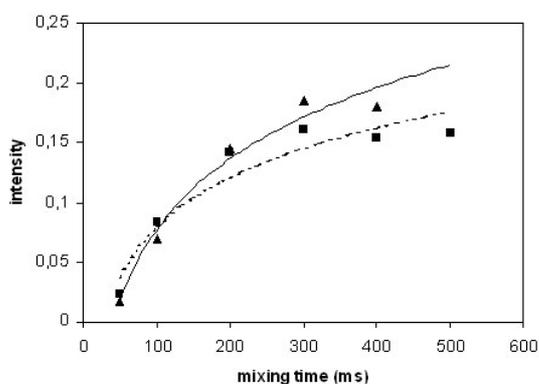
**Figure 3.9 :**Détail de spectres  $^1\text{H}$ - $^1\text{H}$  EXSY du peptide pT212 pS214 à 3.33 mM en présence de 33.3  $\mu\text{M}$  de protéine Pin1 obtenus à différents temps de mélange. Le signal encerclé correspond au pic NOE entre le proton  $\text{H}\alpha$  de la *cis*Pro213 et le proton  $\text{H}\beta$  de la phospho-Thr212 et au pic d'échange entre les protons  $\text{H}\alpha$  de la *cis*Pro213 et la *trans*Pro213.

L'intégration du pic d'échange  $\text{H}\alpha$  *cis* /  $\text{H}\alpha$  *trans* de la Pro213 est réalisée à partir des projections 1D (Figure 3.10) et permet de définir l'évolution de l'échange en fonction du temps de mélange (Graphique 3.4). Toutes les intégrales sont calibrées sur le signal du pic d'échange à 400 ms (intégrale de 1) obtenu sur le peptide court [208-221]pT212/pS214 en présence de Pin1.



**Figure 3.10 :** Projections 1D du pic d'échange  $H\alpha$  *cis*/  $H\alpha$  *trans* Pro213 (représenté sur la **Figure 3.9**) pour différents temps de mélange : 50, 100, 200, 300 et 400 ms réalisés sur le peptide [201-240]pT212/pS214/pT231 (3.33 mM) avec 33.3  $\mu$ M Pin1 (A) ou 33.3  $\mu$ M du domaine catalytique isolé (C) et sur le peptide [208-221]pT212/pS214 (3.33 mM) avec 33.3  $\mu$ M Pin1 (B) ou 33.3  $\mu$ M du domaine catalytique isolé (D).

Le pic d'échange  $H\alpha$  *cis* /  $H\alpha$  *trans* de la Pro213 se superpose au contact NOE entre les protons  $H\alpha$  de la Pro213*cis* et  $H\beta$  de la forme minoritaire de la phosphoThr212 comme décrit dans le chapitre II. La variation du signal NOE  $H\alpha$  Pro213*cis* /  $H\alpha$  phosphoThr212 en fonction du temps de mélange est mesurée comme contrôle (courbe rouge Graphique 3.4) et les résultats sont similaires d'un peptide à l'autre (Graphique 3.3).



**Graphique 3.3 :** Variation de l'intensité des pics NOE entre les protons  $H\alpha$  Pro213*cis* et  $H\alpha$  phosphoThr212 dans les peptides [208-221]pT212/pS214 (en pointillés) et [201-240]pT212/pS214/pT231 (en trait plein).

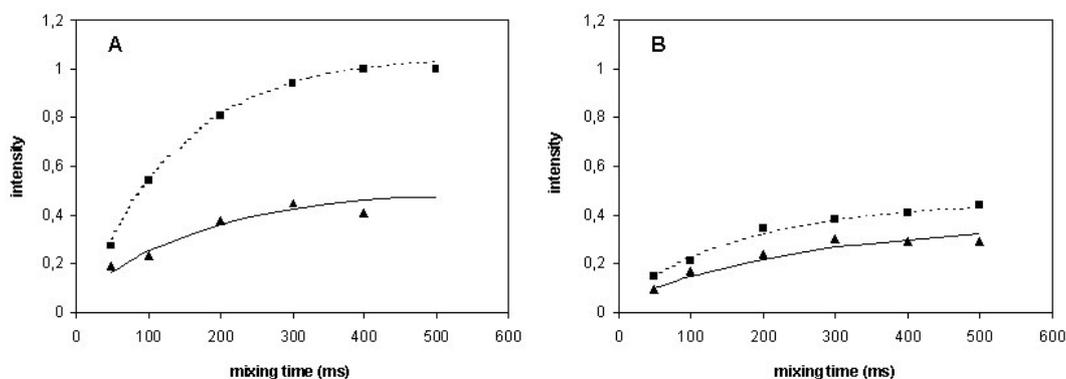
### 1.4.3 – Inhibition de l'activité enzymatique autour de la liaison phosphoThr212-Pro213 en présence du site phosphoThr231-Pro232.

Les *build up* permettent de suivre l'évolution de l'interconversion en Thr212/Ser214 catalysée par Pin1 ou son domaine enzymatique isolé et de calculer une vitesse d'échange ( $k_{\text{exch}}$ ) à partir de la formule suivante :

$$\% [cis \rightarrow trans] = \frac{1 - \exp(-2k_{\text{exch}} \times MT)}{1 + \exp(-2k_{\text{exch}} \times MT)}$$

La comparaison des courbes de *build up* en fonction du substrat permet d'établir que :

- la vitesse de l'échange catalysé par Pin1 entière est plus importante pour le peptide court ( $k_{\text{exch}} = 1 \text{ s}^{-1}$ ) que pour le peptide long ( $k_{\text{exch}} = 0.3 \text{ s}^{-1}$ ) avec une perte de 70% environ (Graphique 3.4A),
- la vitesse de l'échange catalysé par le domaine catalytique isolé est plus faible qu'avec la protéine entière dans les deux cas,
- la vitesse de l'échange catalysé par le domaine catalytique isolé est plus importante pour le peptide court que pour le peptide long mais la différence est moins significative qu'avec l'interconversion catalysée par Pin1 (Graphique 3.4B).



**Graphique 3.4 :** Courbes représentatives des *build up* représentés **Figure 3.10**. En trait plein, le peptide [201-240]pT212/pS214/pT231 et en pointillés, le peptide [208-221]pT212/pS214, (A) avec Pin1, (B) avec le domaine catalytique isolé. En rouge, intégration du signal NOE  $H\alpha$  Pro213<sub>cis</sub> /  $H\beta$  phosphoThr212 sur le peptide [208-221]pT212/pS214 seul.

A partir de ces observations, plusieurs conclusions peuvent être établies :

- Une perte de 60% de l'activité en Thr212/Ser214 au niveau du peptide court avec le domaine catalytique isolé par rapport à Pin1 indique que le domaine WW stimule l'activité enzymatique. Ces résultats semblent indiquer que le domaine WW servirait

de point d'ancrage pour augmenter la concentration locale du domaine catalytique au niveau du substrat. En présence d'un deuxième épitope de reconnaissance, l'effet est moins significatif.

- La présence d'un second phospho-épitope reconnu par Pin1 perturbe de manière significative la catalyse en Thr212/Ser214, et bien que nous n'ayons pas observé d'activité enzymatique en Thr231, des interactions de cet épitope avec le site catalytique pourraient détourner ce dernier de l'épitope phosphoThr212/Ser214, provoquant une diminution de l'activité enzymatique.
- La différence d'activité du domaine catalytique isolé entre le peptide long et le peptide court semble indiquer que le site Thr231 est effectivement reconnu par celui-ci, réduisant la fraction d'enzyme disponible pour l'épitope Thr212/Ser214.

D'après les affinités relatives des deux épitopes présents dans le peptide long, on peut envisager qu'ils sont tous deux reconnus par le domaine WW et que le domaine WW est occupé environ 70% du temps par le site Thr212/Ser214 et 30% du temps par le site Thr231. Les deux épitopes peuvent donc servir alternativement de point d'ancrage pour le domaine WW de Pin1. Si on considère que le domaine enzymatique interagit avec l'épitope lié par le domaine WW, on peut estimer que le domaine enzymatique n'occupe le site Thr231 que 30% du temps, ne conduisant qu'à une diminution de l'activité en Thr212/Ser214 du même ordre. Cependant, la perte d'activité en Thr212/Ser214 est plus importante et permet d'envisager une contribution supplémentaire apportée par un effet « croisé », avec la liaison du domaine WW en Thr231 et la catalyse en Thr212/Ser214, et inversement. En résumé, la catalyse de l'échange en Thr212/Ser214 proviendrait de plusieurs contributions :

- des contributions « positives » : liaison / catalyse en Thr212/Ser214, liaison en Thr231 et catalyse en Thr212/Ser214
- des contributions « négatives » : liaison / catalyse en Thr231, liaison en Thr212/Ser214 et catalyse en Thr231.

Les proportions relatives de ces différentes contributions sont difficiles à évaluer, mais globalement l'effet sur la catalyse en Thr212/Ser214 est négatif. Les données obtenues à partir du domaine enzymatique isolé indiquent que l'épitope phosphoThr231 est reconnu par celui-ci. Avec des affinités de 160 et 380  $\mu$ M pour les phospho-épitopes Thr212/Ser214 et Thr231 respectivement, si on néglige les contributions croisées, on peut estimer que 70% du domaine WW se lie en Thr212/Ser214 et 30% en Thr231, ce qui aurait pour conséquence une

diminution d'environ un tiers de l'activité catalytique en Thr212/Ser214 dans le peptide long par rapport au peptide court. La diminution observée étant de 60% environ, on ne peut pas négliger l'effet « croisé » et estimer une contribution de l'effet « croisé » s'élevant à 80%, ce qui aurait pour conséquence une présence du domaine catalytique autour du site Thr212/Ser214 que 40% du temps.

#### **1.4.4 – Contrôles : rôle du domaine WW, rôle et position du second phospho-épitope.**

Des contrôles ont été effectués pour valider les hypothèses émises précédemment, en particulier pour confirmer le rôle du domaine WW dans l'ancrage au substrat et le rôle de l'épitope Thr231 dans l'inhibition de l'activité catalytique en Thr212/Ser214.

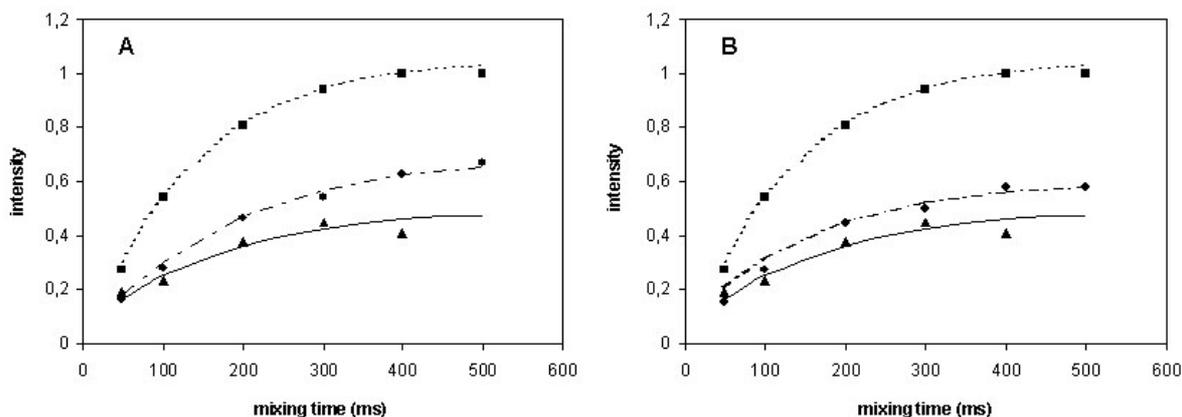
##### ***Rôle du domaine WW dans l'ancrage de Pin1 au substrat.***

Afin de vérifier l'importance du domaine WW dans l'ancrage de Pin1 au substrat et son effet sur la concentration locale du domaine catalytique autour d'un épitope, nous avons évalué l'activité du domaine enzymatique isolé en présence d'une quantité équimolaire du domaine WW isolé, c'est-à-dire en l'absence du *linker* qui relie les deux domaines. Nous n'avons pas retrouvé l'activité isomérase qui avait été mesurée avec Pin1 au niveau du peptide [208-221]pT212/pS214, ce qui permet de souligner l'importance du *linker* qui permet de connecter les deux domaines et confirme un effet d'ancrage de Pin1 au substrat assuré par le domaine WW pour augmenter la concentration de l'enzyme autour du substrat.

##### ***Rôle de la phosphoThr231 dans l'activité catalytique en Thr212/Ser214.***

Afin de confirmer si l'épitope Thr212/Ser214 influe sur la concentration locale du domaine catalytique autour de l'épitope Thr231, nous avons évalué l'activité de Pin1 sur un mélange équimolaire (3,33mM) des deux peptides courts [208-221]pT212/pS214 et [224-240]pT231, c'est-à-dire lorsque les deux phospho-épitopes sont présents mais localisés sur des molécules différentes. Nous avons mesuré une activité isomérase correspondant aux deux tiers environ de l'activité mesurée au niveau du seul peptide [208-221]pT212/pS214, donc plus importante qu'avec le peptide long où les deux épitopes sont présents sur la même molécule de substrat (Graphique 3.5A). Ce résultat traduit une liaison du domaine WW en Thr231 et la localisation du domaine catalytique sur ce site provoquant ainsi une diminution de la fraction de domaine catalytique lié en Thr212/Ser214. La diminution de l'activité enzymatique d'un tiers concorde avec les affinités relatives des deux épitopes pour le domaine WW de Pin1. L'absence d'effet

« croisé » explique la différence d'activité observée avec le peptide long comprenant les deux épitopes. Ces résultats confirment l'hypothèse du rôle inhibiteur de la phosphoThr231 dans l'activité catalytique en Thr212/Ser214 dû à l'effet « croisé » qui permet l'ancrage du domaine WW à un phospho-épitope et l'interaction du domaine enzymatique avec l'autre phospho-épitope.



**Graphique 3.5 :** Activité de Pin1 en Thr212/Ser214 sur le peptide [208-221]pT212/pS214 à 3.33mM (en petits pointillés), sur le peptide [201-240]pT212/pS214/pT231 à 3.33 mM (en trait plein) et (A) sur les peptides [208-221]pT212/pS214 et [224-240]pT231 en mélange équimolaire à 3.33 mM chacun (en pointillés larges) ou (B) sur le peptide [208-221]pT212/pS214/pT217 à 3.33 mM (en pointillés larges).

#### ***Influence de la localisation du second phospho-épitope sur l'activité en Thr212/Ser214.***

Afin d'évaluer si la position de l'épitope Thr231 a une importance particulière dans l'inhibition de l'activité en Thr212/Ser214, nous avons effectué les mêmes mesures avec un peptide tri-phosphorylé court contenant un second phospho-épitope proche, localisé en Thr217.

Nous avons observé une diminution de l'activité par rapport au peptide court ne contenant que l'épitope Thr212/Ser214 mais une activité plus importante qu'avec le peptide long contenant l'épitope Thr231 (Graphique 3.5B). Si ces variations observées sont significatives, l'activité de Pin1 en Thr212/Ser214 mesurée en présence d'une phosphorylation en Thr217, portée par le même substrat, ressemble à l'activité mesurée sur le peptide court di-phosphorylé en Thr212/Ser214 en présence du peptide mono-phosphorylé en Thr231, pour lequel l'effet « croisé » n'était pas possible. Ces données permettraient de conclure à un effet « croisé » moins efficace qui pourrait être dû à une position moins favorable du second phospho-épitope

en Thr217 par rapport à la Thr231. Cependant, une diminution de l'activité indique que le site Thr217 perturbe quand même l'activité enzymatique en Thr212/Ser214 et que la position du second site de reconnaissance n'est pas le seul facteur qui influence celle-ci, mais la présence d'un second motif de reconnaissance même est suffisante.

### **1.5 – Bilan. Modulation de l'activité *cis/trans* isomérase de Pin1 autour de l'épitope Thr212/Ser214 dans des substrats multi-phosphorylés.**

Nous avons précédemment émis l'hypothèse que la séparation dans la séquence primaire des résidus phosphoThr212/231 de la protéine Tau et phosphoThr48/67 de Cdc25, deux partenaires cellulaires de Pin1, pourrait être un paramètre structural déterminant à l'origine d'une possible régulation des interactions entre Pin1 et ses partenaires. Les deux sites étant localisés dans des régions non structurées, cette distance de 19 acides aminés dans la séquence primaire représente vraisemblablement une distance équivalente entre les deux sites dans l'espace. Par ailleurs, la présence d'un fragment peptidique non structuré d'une dizaine de résidus, conférant une flexibilité importante aux deux domaines de Pin1, pourrait favoriser l'accessibilité de chaque site à chacun des domaines de Pin1 simultanément.

#### ***L'introduction d'un second phospho-épitope inhibe l'activité enzymatique en Thr212/Ser214.***

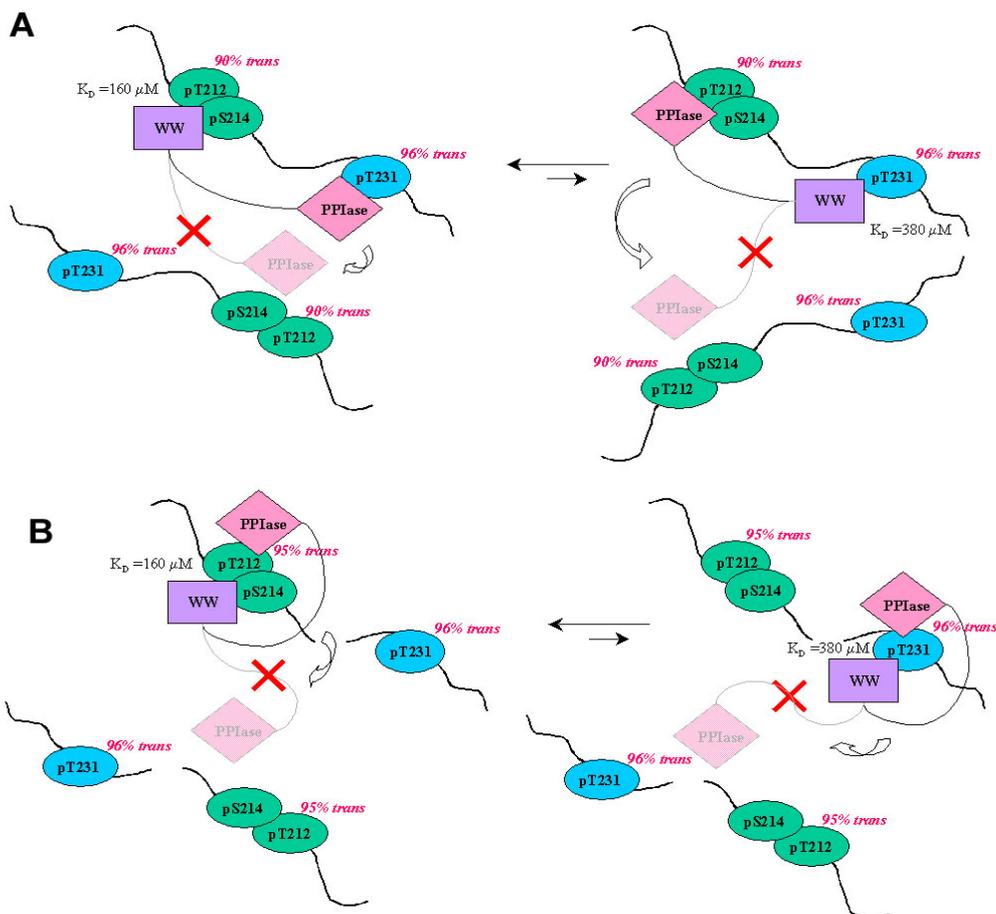
Nous avons observé une diminution de l'activité catalytique de Pin1 autour du phospho-épitope Thr212/Ser214 en présence d'un second épitope reconnu par Pin1 (phosphoThr231-Pro232 ou phosphoThr217-Pro218). L'effet principal est dû à la présence du deuxième site de reconnaissance lui-même, en tant que site de liaison pour le domaine WW, d'après le *build up* réalisé sur les deux peptides courts en mélange avec Pin1 (Graphique 3.5A), mais la position de ce second épitope semble avoir une influence sur l'activité enzymatique en Thr212/Ser214 (Graphique 3.5B).

#### ***Rôle d'ancrage de Pin1 au substrat assuré par le domaine WW.***

Le domaine WW joue un rôle prépondérant dans l'activité catalytique car il permet un ancrage de Pin1 au substrat et, vraisemblablement, une augmentation de la concentration locale en enzyme autour du substrat d'après les données montrant une diminution importante de l'activité isomérase en absence du domaine WW (Graphique 3.4).

**Proposition d'un mode d'interaction de Pin1 aux substrats multi-phosphorylés.**

L'effet moléculaire d'un second épitope de reconnaissance au sein du même substrat reste obscur. On sait que la présence même d'un deuxième site « détourne » Pin1 de l'épitope Thr212/Ser214 mais n'explique pas totalement la perte d'activité mesurée. Le fait que les deux épitopes soient sur la même molécule-substrat joue également un rôle non négligeable (Graphique 3.5A). Nous avons envisagé un effet « croisé » mettant en œuvre la liaison à l'épitope Thr212/Ser214 et la « catalyse » en Thr231 et inversement, qui apporterait une contribution supplémentaire, plutôt négative, sur l'isomérisation en Thr212/Ser214. La position du second phospho-épitope (en 217 ou en 231) influence également l'activité enzymatique en Thr212/Ser214, puisque l'activité en Thr212/Ser214 en présence de l'épitope Thr217 phosphorylé est similaire à l'activité mesurée en présence de deux phospho-épitopes localisés sur deux molécules différentes ne permettant pas la contribution due à l'effet « croisé » et suggérant l'impossibilité, avec le site Thr217, d'avoir cet effet « croisé ». Ces données semblent indiquer l'existence d'un facteur structural reliant les deux phospho-épitopes pour moduler l'activité enzymatique de Pin1.



### ***Rôle des phosphorylations multiples dans la régulation de l'activité isomérase ?***

Bien que nous ayons déterminé des populations de conformères *cis* pour les Pro218 et Pro232 de 10% et 4% respectivement, nous n'avons pas observé la catalyse de l'interconversion *cis/trans* par Pin1 pour les deux motifs phosphoThr217-Pro218 et phosphoThr231-Pro232. On pourrait envisager une catalyse moins efficace au niveau de ces épitopes, qui pourrait provenir d'une rigidité plus importante du squelette peptidique, en raison de la présence d'une autre proline en position 219 et 233, respectivement (avec, dans le cas de la Thr217, une proline en 216 également), et qui expliquerait que le pic d'échange ne soit pas observable. Si on poursuit le raisonnement à l'extrême, on peut envisager qu'il s'agisse de simples sites de liaison pour le domaine catalytique, qui occuperaient le site actif avec une faible affinité. Ces deux épitopes pourraient jouer un rôle régulateur sur l'activité en Thr212/Ser214 en provoquant une diminution de l'activité isomérase en Thr212/Ser214 par liaison du site enzymatique.

Des expériences supplémentaires sont nécessaires pour étayer ces hypothèses. La comparaison de l'orientation relative des deux domaines de Pin1 en présence des substrats, court ou long, dans des milieux orientés pourrait permettre de confirmer une interaction de chacun des domaines avec chacun des épitopes. Nous avons commencé des études dans ce sens, mais nous n'avons pas encore réussi à déterminer les conditions qui permettraient d'orienter Pin1 dans un milieu contenant des phages pour déterminer les constantes dipolaires relatives des deux domaines. En effet, des interactions non spécifiques avec les phages conduisent à un élargissement des signaux de la protéine. On pourrait également envisager, pour vérifier l'importance de la position du second épitope sur l'activité enzymatique en Thr212/Ser214, de refaire un *build up* avec la protéine mutante Pin1 $\Delta$ [44-49]. Le raccourcissement du *linker* inter-domaine ne modifie pas l'affinité pour le peptide long par rapport à la protéine sauvage mais pourrait avoir un impact plus significatif sur l'activité catalytique.

### ***Contradiction entre l'effet moléculaire observé et la fonction biologique sur PP2A.***

*In vivo*, l'effet fonctionnel de Pin1 se traduit par une augmentation de l'activité de PP2A, une phosphatase *trans*-spécifique, pour déphosphoryler la phosphoThr231. Aucune activité de déphosphorylation n'est détectée au niveau du site Thr212 (Zhou, Kops et al. 2000). Il faut cependant noter que la Thr212 a probablement un niveau de phosphorylation plus bas que la Thr231 *in vivo*. Quel est le mode d'action de Pin1 sur la stimulation de la déphosphorylation

de la Thr231 par PP2A ? Nos données sur les interactions de Pin1 avec les peptides phosphorylés à l'échelle moléculaire indiquent que le site phosphoThr231 est majoritairement en conformation *trans*, alors quel pourrait être le rôle d'une prolyl *cis/trans* isomérase comme Pin1 dans ce contexte ? Nos données moléculaires sont difficilement réconciliables avec les données cellulaires, d'autant plus qu'aucune catalyse n'est observée en Thr231. Les fondements même des hypothèses avancées sur le mode de fonctionnement de Pin1, concernant un éventuel effet catalytique d'isomérisation *cis/trans* de la liaison phosphoThr231-Pro232 par Pin1 aidant la déphosphorylation par PP2A, sont peut-être à revoir (Zhou, Kops et al. 2000; Stukenberg and Kirschner 2001). Il n'existe probablement pas d'isomérisation enzymatique en Thr231, et quel en serait l'intérêt puisque PP2A déphosphoryle les liaisons en *trans* et que celle-ci ne présente que 4% en conformation *cis* ? On peut envisager un effet conformationnel plus global concernant la protéine tau entière pour expliquer le rôle de Pin1, effet que l'on ne pourrait pas détecter avec des substrats peptidiques. On peut également imaginer que, dans les PHF, la situation soit différente et que la proportion de formes *cis* soit plus importante, nécessitant l'intervention de Pin1 pour faciliter la déphosphorylation par PP2A et restaurer la liaison de tau aux microtubules. Dans ce cas, le site Thr231 n'est peut-être pas le site d'action principal pour Pin1, mais il peut s'agir de l'épitope Thr212/Ser214, phosphorylé uniquement dans les PHF (Hoffmann, Lee et al. 1997).

#### ***Vers des études avec la protéine tau entière.***

Les études que nous avons menées au travers de substrats peptidiques, pour tenter de percer le mode de fonctionnement de Pin1 à l'échelle moléculaire, ont été limitées par la taille restreinte des peptides, n'autorisant l'investigation que de quelques motifs phosphoSer/Thr-Pro de la protéine tau qui compte 17 motifs de ce type, susceptibles d'être des substrats de Pin1. Nous avons concentré notre étude sur des épitopes de la région riche en prolines, mais un phospho-épitope de plus haute affinité pour Pin1 peut exister dans une autre région de la protéine. D'autre part, dans le contexte de la protéine entière, des modulations conformationnelles peuvent exister autour de ces motifs Ser/Thr-Pro, concernant localement l'équilibre conformationnel des prolines, ou un réarrangement conformationnel plus global de la protéine, dont les peptides ne rendraient pas compte. Ces modulations structurales pourraient avoir un impact significatif sur la reconnaissance de ces phospho-épitopes par Pin1. On peut par exemple envisager une proline majoritairement en conformation *cis* dans le contexte de la protéine tau entière, qui ne serait pas reconnue par le domaine WW. Nous

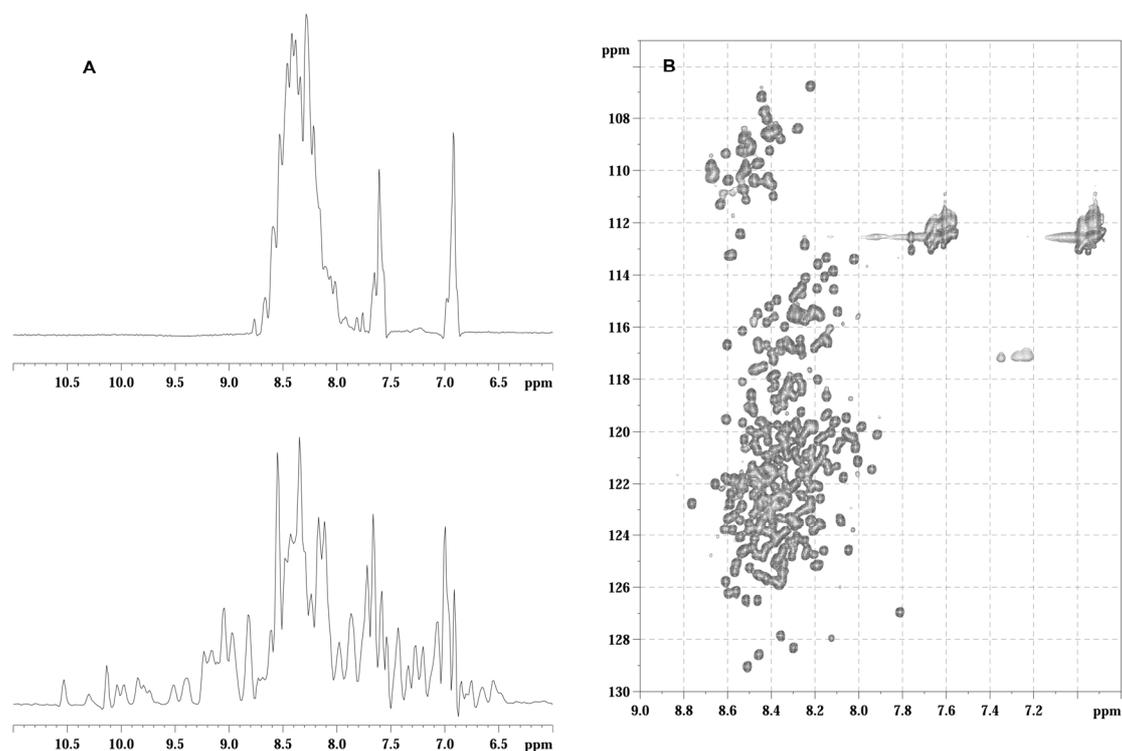
avons donc franchi l'obstacle de la taille du substrat pour étudier les interactions entre Pin1 et la protéine tau entière, qui implique une difficulté supplémentaire concernant la phosphorylation. Nous avons, dans un premier temps, proposé une méthode d'attribution de la protéine par RMN et contourné l'obstacle de la phosphorylation avec une protéine mutante mimant l'état hyperphosphorylé de tau par mutation de 10 résidus sérine ou thréonine en glutamate (Eidenmuller, Fath et al. 2000).

## **II – Attribution de la protéine Tau par RMN et premières études d'interaction avec Pin1.**

Nous avons entamé au laboratoire un vaste projet autour de la protéine tau, avec un intérêt particulier pour le processus d'agrégation de la protéine à l'échelle atomique avec le suivi de la cinétique d'agrégation par RMN en présence de poly-anions (héparine, acide arachidonique), dont les résultats ne seront pas discutés ici (Sillen, Leroy et al. soumis). L'implication de la phosphorylation dans ce processus d'agrégation sera également envisagée, et dans ce contexte, nous voulons déterminer le rôle de Pin1. L'ensemble de ces investigations nécessite l'attribution, au moins partielle, de la protéine tau et fera l'objet de ce début de paragraphe.

Les caractéristiques intrinsèques de la protéine tau, qui sont la taille (441 résidus pour l'isoforme la plus longue et 352 pour l'isoforme la plus courte), une distribution déséquilibrée en acides aminés (5 résidus acide aminé représentent 50% de la séquence), une grande flexibilité et peu ou pas d'éléments de structures secondaires, mènent vers une faible dispersion des signaux protons en RMN et ne permettent pas d'envisager une attribution séquentielle classique de la protéine par RMN.

L'aspect global du spectre proton de la protéine tau ressemble davantage au spectre d'un grand peptide non structuré qu'à celui d'une protéine globulaire, avec une dispersion des protons  $H_N$  sur moins de 1 ppm, alors que, pour une protéine comme Pin1 par exemple, la dispersion des protons  $H_N$  s'étend sur un peu plus de 3 ppm (Figure 3.11A). En revanche, une dispersion correcte en azote sur 25 ppm est plus engageante pour une étude de la protéine tau entière par RMN (Figure 3.11B). De plus, une faible largeur de raie pour une protéine de 50 kDa indique un haut degré de flexibilité, et constitue un atout non négligeable.



**Figure 3.11 :** (A) Comparaison des spectres protons des protéines tau (en haut ) et Pin1 (en bas). (B) Spectre <sup>1</sup>H-<sup>15</sup>N HSQC de la protéine tau indiquant une dispersion « normale » dans la dimension <sup>15</sup>N.

Une attribution partielle de la protéine tau a été réalisée selon deux stratégies étroitement liées, mettant en œuvre l'utilisation de peptides courts issus de la protéine (« *mapping* peptidique ») et l'utilisation des expériences RMN tridimensionnelles standards. Les résultats qui seront présentés ici concernent essentiellement l'approche par « *mapping* peptidique ».

## 2.1 – Attribution de la protéine Tau par « *mapping* peptidique ».

### *Introduction au concept de random coil en RMN.*

Le déplacement chimique d'un atome au sein d'une protéine est défini par deux contributions : celle de la séquence primaire, c'est-à-dire l'influence des résidus environnants dans la séquence, et celle de l'échantillonnage conformationnel induit par les éléments de structure secondaire/tertiaire (ou l'absence d'éléments de structure), c'est-à-dire l'influence des atomes proches dans l'espace. Les déplacements chimiques *random coil* ont été définis selon deux techniques distinctes qui n'apportent chacune qu'une des deux contributions à la fois :

- Une technique basée sur la détermination des déplacements chimiques <sup>13</sup>C et <sup>15</sup>N d'un résidu X dans un petit peptide de 5 résidus GGXGG (Wishart, Bigam et al. 1995). Cette méthode ne tient pas compte de l'influence de la séquence primaire mais définit

assez bien le concept « non structuré », car le pentapeptide n'adopte aucune structure secondaire stable et échantillonne un vaste espace conformationnel. Si on veut déterminer l'influence des résidus en position -2, -1, +1 et +2 dans la séquence primaire, il faudrait synthétiser  $20^5$  pentapeptides (en supposant qu'ils n'adoptent pas une conformation privilégiée).

- Une technique basée sur une moyenne des déplacements chimiques de protéines, dont la structure est connue, issus des bases de données. Cette méthode interprète l'absence de structure comme étant la moyenne d'un ensemble de structures diverses, ce qui n'est pas tout à fait exact par rapport au concept *random coil*, mais tient compte de l'influence de la séquence primaire et permet d'échantillonner l'ensemble des contributions des résidus en positions -2, -1, +1, +2 sur le déplacement chimique du résidu central.

Si on considère que la protéine tau n'est pas structurée, on peut supposer que les spectres de peptides courts issus de la séquence primaire de la protéine sont des sous-spectres du spectre de la protéine totale. Une technique basée sur l'attribution de peptides de 15 résidus et sur la superposition graphique des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC des peptides à celui de la protéine, nous a permis d'avancer dans l'attribution de tau. Cette méthode présente l'avantage de combiner les deux techniques décrites précédemment : elle permet de s'affranchir de l'influence de la séquence primaire car, les peptides étant issus de la protéine, la séquence primaire est la même (« effets de bord » mis à part), et l'absence de structure secondaire peut être aisément vérifiée dans le contexte de peptides courts.

### 2.1.1 – Principe du « *mapping* peptidique ».

Le point central du « *mapping* peptidique » est que le même acide aminé présent dans le peptide ou dans la protéine entière doit donner le même signal de corrélation  $^1\text{H}$ - $^{15}\text{N}$  si il échantillonne un espace conformationnel identique dans les deux contextes. Même si l'idée paraît simple, la correspondance des résonances est loin d'être évidente quand on compare la taille des deux objets (1,5-2 kDa pour les peptides et 50 kDa pour la protéine).

Les peptides représentent souvent un modèle pour l'étude des protéines, permettant une attribution relativement aisée malgré une faible dispersion spectrale, grâce aux valeurs des tables de déplacements chimiques « *random coil* »  $^1\text{H}$  décrits pour de petits peptides modèles de 5 résidus (Bundi and Wuthrich 1977). L'utilisation de séquences impulsionsnelles standards de type TOCSY et NOESY permet l'attribution des résidus par type et l'attribution séquentielle à la fois. En combinant les deux informations, on peut attribuer ainsi des peptides

allant jusqu'à 20 résidus et même parfois davantage. Ainsi, le fait que le déplacement chimique, un paramètre très sensible à l'environnement chimique, soit proche des valeurs « *random coil* » indique une tendance non structurée. D'autre part, l'analyse des informations sur les contacts NOE dans le contexte peptidique est relativement aisée et renseigne immédiatement sur la présence (ou l'absence) de structures secondaires. On peut donc envisager l'utilisation de fragments peptidiques courts pour décrire des régions protéiques flexibles sans structures apparentes et, par extension, des protéines totalement non structurées. Les peptides doivent être suffisamment petits pour permettre leur synthèse par voie chimique et leur purification, ainsi que leur caractérisation par RMN, c'est-à-dire l'attribution séquentielle et la détermination de l'état de « structuration ». Une taille restreinte limite également la possibilité pour le peptide d'adopter une structure stable. Dans ce contexte, le peptide peut également être utilisé comme sonde pour déterminer les régions protéiques présentant une structure secondaire résiduelle. La RMN est une technique privilégiée car elle peut fournir de nombreuses indications de l'état de « structuration » à l'échelle atomique, résidu par résidu. Trois paramètres peuvent être utilisés pour caractériser des éléments de structures secondaires : le déplacement chimique, les constantes de couplage  $^3J[{}^1\text{H}-{}^1\text{H}]$ , les contacts NOE  ${}^1\text{H}-{}^1\text{H}$ .

Le déplacement chimique est, de loin, le paramètre le plus facile à exploiter, car l'interprétation nécessite uniquement l'acquisition de simples spectres  ${}^1\text{H}-{}^{15}\text{N}$  HSQC. De plus, le déplacement chimique est une valeur moyenne prenant en compte l'ensemble des déplacements chimiques correspondant chacun à une des conformations adoptées par la protéine. L'interprétation des constantes de couplage nécessite l'acquisition et l'exploitation de spectres DQF-COSY, nécessitant le calcul des constantes de couplage pour environ 400 couplages  $\text{H}_\text{N}-\text{H}_\alpha$ . Quant à l'interprétation des couplages dipolaires NOE, elle est très délicate car, outre le nombre considérable de signaux à traiter, l'intensité du contact NOE est biaisée. En effet, contrairement au déplacement chimique, le pic NOE est proportionnel à l'inverse de la puissance 6 de la distance inter-atomique, de sorte que l'intensité du pic NOE rendra compte préférentiellement de la conformation la plus « repliée ». Par exemple, si deux protons sont séparés de 2Å dans une conformation repliée, représentant 10% de l'échantillonnage conformationnel, et 5Å dans une conformation étendue, représentant 90% de l'échantillonnage conformationnel, la distance moyenne pondérée entre les deux protons sera de 4,7Å tandis que le pic NOE moyen sera équivalent à une distance de 2,9Å. En distance moyenne, cette distance de 2,9Å correspond à 70% de l'échantillonnage conformationnel en conformation repliée et 30% en conformation étendue. L'avantage du déplacement chimique

est que le déplacement chimique de la conformation repliée n'aura pas plus de poids que le déplacement chimique de la conformation étendue sur le déplacement chimique moyen. D'autre part, le déplacement chimique est, par définition, un paramètre extrêmement sensible à l'environnement chimique, en particulier pour les signaux  $^{15}\text{N}$ , ce qui rend ces derniers très difficilement prédictibles avec les méthodes de prédiction des déplacements chimiques basées sur les peptides ou les bases de données (Wang and Jardetzky 2002; Meiler 2003).

### *Découpage de la protéine tau en peptides courts localisés dans diverses régions.*

Différentes régions de l'isoforme la plus longue de la protéine Tau ont été couvertes par le « mapping » peptidique : les peptides [12-24], [36-48] et [42-56] décrivent la partie N-terminale, les peptides [208-221] et [224-240] la région riche en prolines, le peptide [303-317] le deuxième segment répété, et les peptides [393-408] et [416-428] la partie C-terminale (Figure 3.12).

1	MAEPRQEFEV	<u>MEDHAGTYGL</u>	<u>GDRK</u> DQGGYT	MHQDQ <u>EGDTD</u>	<u>AGLKE</u> <u>SPLQ</u> T
51	<u>PTEDGSEEPG</u>	<u>SETSDAKSTP</u>	<u>TAEDVTAPLV</u>	<u>DEGAPGKQAA</u>	<u>AQP</u> <u>HT</u> E <u>I</u> PE <u>G</u>
101	<u>TTAEEAGIGD</u>	TPSLEDEAAG	HVTQARMVSK	SKDGTGSDDK	KAKGADGKTK
151	IATPRGAAPP	GQKGQANATR	I PAKTPPAPK	TPPSSGEPPK	SGDRSGYSSP
201	GSPGTPG <u>SRS</u>	<u>RTPSLTPPT</u>	<u>REP</u> <u>KKVAVVR</u>	<u>TPPKSPSSAK</u>	SRLQTAPVPM
251	PDLKNVSKI	GSTENLKHQP	<u>GGGKVQI</u> INK	<u>KLDLSNVQSK</u>	<u>CGSKDNIKHV</u>
301	PG <u>GGSVQIVY</u>	<u>KPVDLSK</u> VTS	KCGSLGNIHH	KPGGGQVEVK	SEKLDLDFKDRV
351	QSKIGSLDNI	THVPGGGNKK	IETHKLTFRE	NAKAKTDHGA	E <u>I</u> <u>VYKSPVVS</u>
401	<u>GDTSPRHL</u> SN	VSSTG <u>SIDMV</u>	<u>DSPQLATL</u> AD	EVSASLAKQG	L

**Figure 3.12 :** Régions de la protéine tau (dans son isoforme la plus longue de 441 résidus) couvertes par le « mapping » peptidique indiquées en caractères gras soulignés. La séquence de l'isoforme la plus courte (fœtale) de 352 résidus est définie par les caractères noirs, les zones de délétions correspondant aux exons 2(-), 3(-) et 10(-) sont notées en rouge.

Le choix de la longueur de la séquence peptidique est déterminé par des critères prenant en compte :

- Un « effet de bord » et par conséquent le nombre utile de résidus dans la séquence aboutissant à une attribution efficace des signaux de la protéine entière,
- la taille critique d'un peptide permettant son attribution malgré une faible dispersion des signaux causée par l'absence de structure et sa caractérisation structurale.

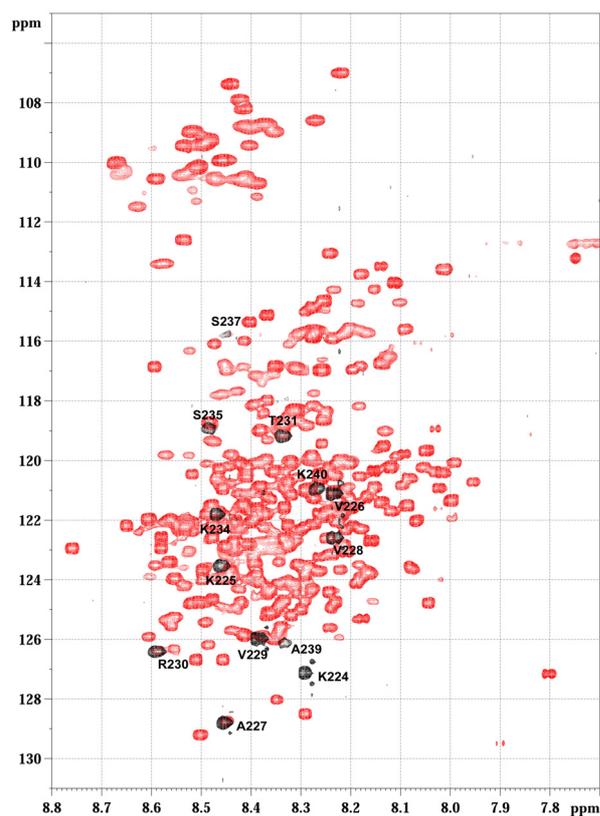
Les peptides ont été synthétisés selon une procédure classique de synthèse peptidique en phase solide en stratégie Fmoc sur synthétiseur automatique en flux continu suivi d'une coupure du peptide de son support solide par réaction avec une solution concentrée de TFA.

Les peptides bruts sont purifiés en phase inverse et caractérisés par spectrométrie de masse MALDI-TOF et attribués par RMN homonucléaire à l'aide des séquences impulsives TOCSY et NOESY (temps de mélange de 69 ms et 400 ms respectivement). Les valeurs des déplacements chimiques des protons  $H_{\alpha}$ , ainsi que les protons des chaînes latérales des résidus peptidiques, ont des valeurs proches des valeurs « *random coil* ». De plus, seuls des contacts NOE sont observables entre les protons  $H_{\alpha}$  du résidu (i) et le proton  $H_N$  du résidu (i+1) permettant l'attribution séquentielle des peptides mais n'indiquant aucun élément de structure secondaire stable. D'autre part, l'acquisition de spectres  $^1H$ - $^{15}N$  HSQC en abondance naturelle en  $^{15}N$ , nécessite une solubilité des peptides à une concentration relativement élevée de 4 mM dans l'eau à pH neutre. L'acquisition des données sur un spectromètre à 600 MHz équipé d'une cryosonde nécessite 296 scans soit 8 heures de temps machine. Ainsi la limite d'attribution de la protéine par l'utilisation de peptides courts réside d'une part, en la synthèse et la purification d'un grand nombre de peptides, ainsi qu'en l'acquisition des données par RMN et l'attribution des peptides, sans oublier la limite de solubilité des peptides, notamment ceux localisés dans la région à segments répétés. Le principe d'attribution par *mapping* peptidique fait ensuite intervenir la superposition graphique des spectres  $^1H$ - $^{15}N$  HSQC des peptides à celui de la protéine tau entière.

#### ***Correspondance entre les signaux du peptide et ceux de la protéine.***

Dans certains cas, nous avons trouvé une bonne correspondance entre les signaux des peptides et ceux de la protéine entière, nous permettant d'avancer dans l'attribution de la protéine. Par exemple, dans la région riche en prolines, on trouve une bonne correspondance entre les résonances des résidus V226 à S235 (Figure 3.13). Les résidus S237 et S238 sont élargis dans le contexte peptidique et ne sont pas visibles sur le spectre  $^1H$ - $^{15}N$  HSQC. Les résonances des résidus en début et fin de séquence dans le peptide ne correspondent pas avec ceux de la protéine en raison d'une modification de leur environnement immédiat. Ce point sera discuté dans le paragraphe suivant.

Une description des informations fournies par le « *mapping* peptidique » des différentes régions de la protéine tau est présentée plus en détails dans la première version d'un article soumis à la revue Chembiochem, donné en fin de chapitre.



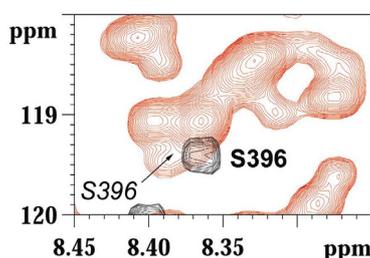
**Figure 3.13 :** Superposition des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine tau entière (rouge) et d'un peptide [K224-K240] de la région riche en prolines (noir).

### *Caractérisation de l'« effet de bord ».*

Les résidus localisés dans les extrémités N- et C-terminales du peptide isolé de la protéine subissent une modification de leur environnement séquentiel et leur déplacement chimique n'est plus représentatif de leur environnement protéique. C'est ce que nous avons appelé « effet de bord ». Cet effet a déjà été observé dans la première partie de ce chapitre avec la comparaison des spectres des peptides courts [208-221] et [224-240] par rapport à celui du peptide long, pour lesquels un « effet de bord » de deux résidus avait été mis en évidence. Une autre preuve provient de la séquence de la protéine même et de la comparaison des spectres isoformes de 441 et 352 résidus. Les résidus Y310 et Y394 dans la protéine ont les mêmes voisins en position -2, -1 et +1 alors que le résidu en position +2 varie (P312 et S396), conduisant à de légères modifications de leur pic de corrélation  $^1\text{H}$ - $^{15}\text{N}$  sur le spectre. De même les résidus V309 et V393, qui ont les mêmes voisins en -1, +1 et +2 mais un voisin différent en position -2 (Q307 et E391), n'ont pas exactement le même pic de corrélation sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Le résidu G42, en revanche, n'a pas le même voisin en position +3 dans l'isoforme longue (E45) et dans l'isoforme courte (A103, d'après la numérotation de l'isoforme longue) mais donne le même signal de corrélation  $^1\text{H}$ - $^{15}\text{N}$  dans les spectres de deux isoformes.

### 2.1.2 – Un outil pour la détection de structures locales.

Le déplacement chimique apporte ici des informations structurales importantes concernant la protéine. D'autant plus que le déplacement chimique de l'azote est extrêmement sensible, ce qui le rend, en général, difficilement prédictible par les méthodes de prédiction des déplacements chimiques. Le fait que les pics de corrélation  $^1\text{H}$ - $^{15}\text{N}$  du peptide et de la protéine se superposent bien rend d'autant plus fiable les conclusions quant à un échantillonnage conformationnel identique de la séquence d'acides aminés dans les contextes peptidique et protéique. En revanche, une différence dans les pics de corrélation permet de détecter des régions protéiques qui adoptent une structure différente que la même séquence dans le peptide et qui devraient être examinées plus en détail par d'autres techniques RMN (Figure 3.14).



**Figure 3.14 :** Superposition des spectres de la protéine tau entière (rouge) et du peptide [393-408] illustrant, pour le fragment  $\text{V}^{393}\text{YKSPV}^{398}$ , une absence de correspondance entre les signaux du peptide et ceux de la protéine avec un agrandissement de la région autour du signal de la S396.

Ainsi, dans le peptide [V393-L408], nous observons des différences dans la partie N-terminale du peptide qui vont au-delà des deux résidus impliqués dans l'« effet de bord » et s'étendent jusqu'à la Val398 (résidu n°6 dans la séquence peptidique). Les signaux des autres résidus du peptide se superposent avec les signaux de la protéine, mais n'adoptent pas la même largeur de raie, plus élargie dans cette région protéique. Ce phénomène que nous observons pour la partie N-terminale du peptide peut provenir du fait que nous ayons rompu un *cluster* hydrophobe qui s'établit au sein de la protéine au niveau de la séquence « IVYK ». L'absence du résidu Ile392 dans le peptide empêcherait la formation de ce *cluster* hydrophobe et aboutirait, par conséquent, à une différence entre les signaux du peptide et de la protéine. Cette région pourra par la suite être examinée plus en détail.

L'attribution par *mapping* peptidique a été confirmée avec l'utilisation des expériences de RMN tridimensionnelles standards, dont l'exploitation des déplacements chimiques  $^{13}\text{C}$ , sur base des valeurs *random coil*, peut permettre d'avancer dans l'attribution de la protéine tau,

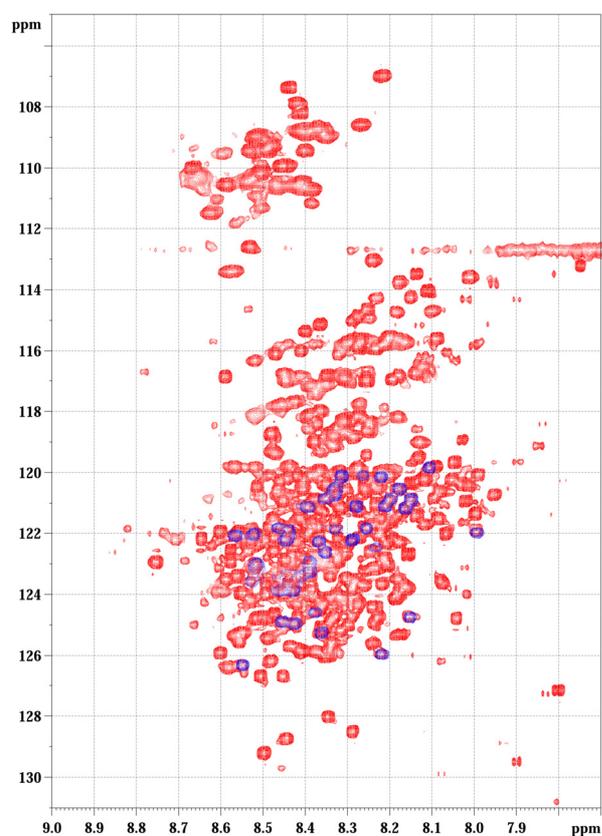
mais est utilisée ici pour confirmer les attributions obtenues par les peptides, plutôt que pour une attribution *ab initio*.

### 2.1.3 – Confirmation de l'attribution par paires de résidus.

Le principe d'attribution des protéines par RMN est basé sur l'acquisition d'une collection de spectres tridimensionnels de type HNCO, HNcaCO, HNCA, HNcoCA, CBCANH, CBCAcoNH dont le point de départ commun est le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. L'exploitation de l'HNCA, par exemple, permet d'extraire les résonances  $^{13}\text{C}\alpha$  des résidus (i) et (i-1) pour chaque pic de corrélation  $^1\text{H}$ - $^{15}\text{N}$ . A partir d'un pic de corrélation  $^1\text{H}$ - $^{15}\text{N}$  pour un résidu (i) donné, on peut donc retrouver le pic de corrélation du résidu précédent (i-1) en passant par son  $\text{C}\alpha$ , et ainsi de suite... Ce processus fonctionne bien dans la mesure où les résonances sont suffisamment dispersées. En cas de dégénérescence de plusieurs valeurs  $^{13}\text{C}\alpha$  par exemple, on fait appel aux autres expériences tridimensionnelles et les ambiguïtés sont levées à partir des résonances  $^{13}\text{CO}$  ou  $^{13}\text{C}\beta$ . Cette technique est couramment utilisée pour l'attribution séquentielle de protéines adoptant une structure globulaire stable qui donnent une bonne dispersion des signaux. En revanche, dans le cas de protéines de structure étendue, comme la protéine tau, une dégénérescence importante des déplacements chimiques  $^{13}\text{C}$  ne permet pas d'appliquer cette méthode d'attribution. Cependant, on peut tirer avantage de cette situation, en considérant que les déplacements chimiques  $^{13}\text{C}\alpha/\beta$  sont très proches des valeurs *random coil* données par les tables (Wishart, Bigam et al. 1995), à  $\pm 0.5$  ppm près. On peut ainsi exploiter les données des expériences RMN 3D et attribuer un type de résidu à un pic de corrélation  $^1\text{H}$ - $^{15}\text{N}$ , pour un résidu (i), à partir des valeurs  $^{13}\text{C}\alpha$  et  $^{13}\text{C}\beta$  données par la comparaison des spectres HNCA/HNcoCA et CBCANH/CBCAcoNH respectivement, et identifier le type de résidu qui le précède, résidu (i-1), de la même manière. En revanche, on ne peut pas déterminer quel est le pic de corrélation  $^1\text{H}$ - $^{15}\text{N}$  de ce résidu (i-1) en raison de la dégénérescence spectrale. On peut de cette manière identifier des paires de résidus, et si, par chance, il n'existe qu'une seule paire de ce type dans la séquence primaire de la protéine, on peut l'attribuer sans ambiguïté. Sachant que l'isoforme la plus longue de la protéine tau contient 79 paires uniques dans sa séquence, et 49 paires de résidus, qui existent en double exemplaire, il est possible d'envisager une attribution partielle de la protéine. Si on ne prend en compte que, sur les 49 couples de paires de la séquence de Tau441, les 19 couples qui possèdent une paire absente dans la protéine Tau352 et deviennent par conséquent unique dans l'isoforme la plus longue, on peut non seulement lever l'ambiguïté pour ces 19 paires,

mais également pour les 19 paires complémentaires présentes dans les deux isoformes. Si on poursuit ce raisonnement à l'extrême, on peut considérer les peptides comme des isoformes ultra-courtes de la protéine, et lever ainsi des ambiguïtés. L'attribution de la protéine par les peptides peut être vérifiée paire par paire avec cette technique et, dans certains cas, nous avons pu lever des ambiguïtés concernant des paires existant en plusieurs exemplaires. On peut également souligner qu'en combinant les deux stratégies, on se rapproche d'une technique classique d'attribution des protéines, c'est-à-dire que, avec les résonances  $^{13}\text{C}\alpha/\beta$  des expériences RMN 3D, on peut identifier et attribuer (sans ambiguïté) une paire de résidus à partir du pic de corrélation  $^1\text{H}-^{15}\text{N}$  du résidu (i) et, grâce au peptide, on peut retrouver le pic de corrélation  $^1\text{H}-^{15}\text{N}$  du résidu (i-1) sur le spectre de la protéine entière. Ces résultats ont fait l'objet d'une publication (Smet, Sillen et al., *Chembiochem*, sous presse).

Le marquage sélectif de la protéine en  $^{15}\text{N}$  au niveau des résidus Lys, Leu ou Val a permis de faciliter l'attribution de ces résidus, dont les pics de corrélation  $^1\text{H}-^{15}\text{N}$  sont situés dans la zone la plus encombrée du spectre.



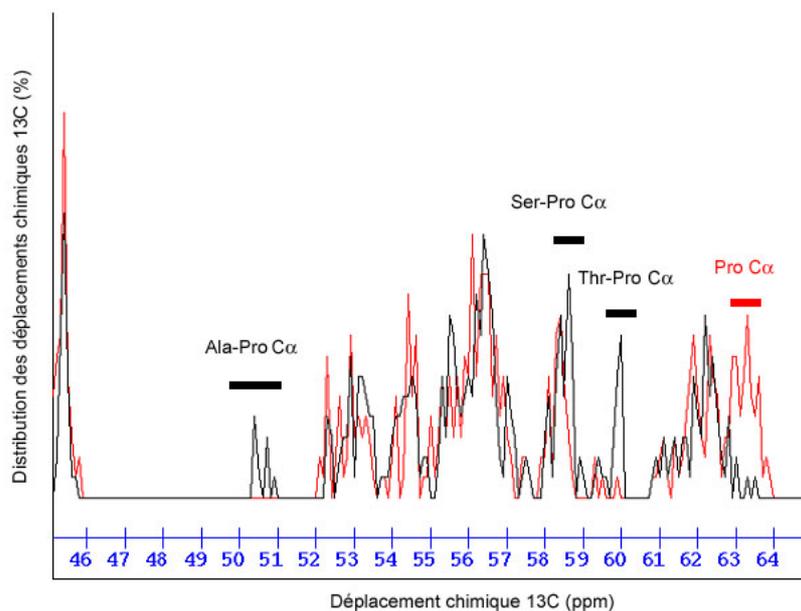
**Figure 3.15 :** Superposition des spectres  $^1\text{H}-^{15}\text{N}$  HSQC de Tau uniformément enrichie en  $^{15}\text{N}$  en rouge, et sélectivement enrichie en  $^{15}\text{N}$ Lys en bleu.

***Particularité de la proline et des résidus précédant une proline pour l'attribution des sites de phosphorylation inclus dans des motifs Ser/Thr-Pro (Lippens, Wieruszeski et al. 2004).***

L'attribution de paires de résidus de la protéine tau a été mise en œuvre au départ avec l'identification des résidus environnant une proline dans des paires Pro-X et X-Pro, où X est n'importe quel acide aminé. La proline possède en effet les particularités suivantes :

- La proline ne possédant pas de proton  $H_N$ , donc pas de pic de corrélation  $^1H-^{15}N$ , les valeurs  $^{13}C\alpha$  des prolines ne peuvent être obtenues qu'à partir des valeurs  $^{13}C\alpha$  (i-1) sur le pic de corrélation du résidu suivant la proline par comparaison des HNCA et HNcoCA. (courbe rouge, Figure 3.16)
- La proline induit une diminution de 2 ppm pour la résonance du  $^{13}C\alpha$  du résidu qui la précède (courbe noire, Figure 3.16) par rapport aux valeurs  $^{13}C\alpha$  du même résidu dans un motif X-Ala (Wishart, Bigam et al. 1995).

Cette particularité a pu être exploitée notamment pour l'attribution des résidus Ser ou Thr dans des motifs Ser/Thr-Pro, qui nous intéressent tout particulièrement en tant que sites de phosphorylation par les kinases dirigées contre des résidus précédant une proline (CDK, PKA, GSK3 $\beta$ ) et en tant que sites permettant la reconnaissance des protéines, après leur phosphorylation, par Pin1.



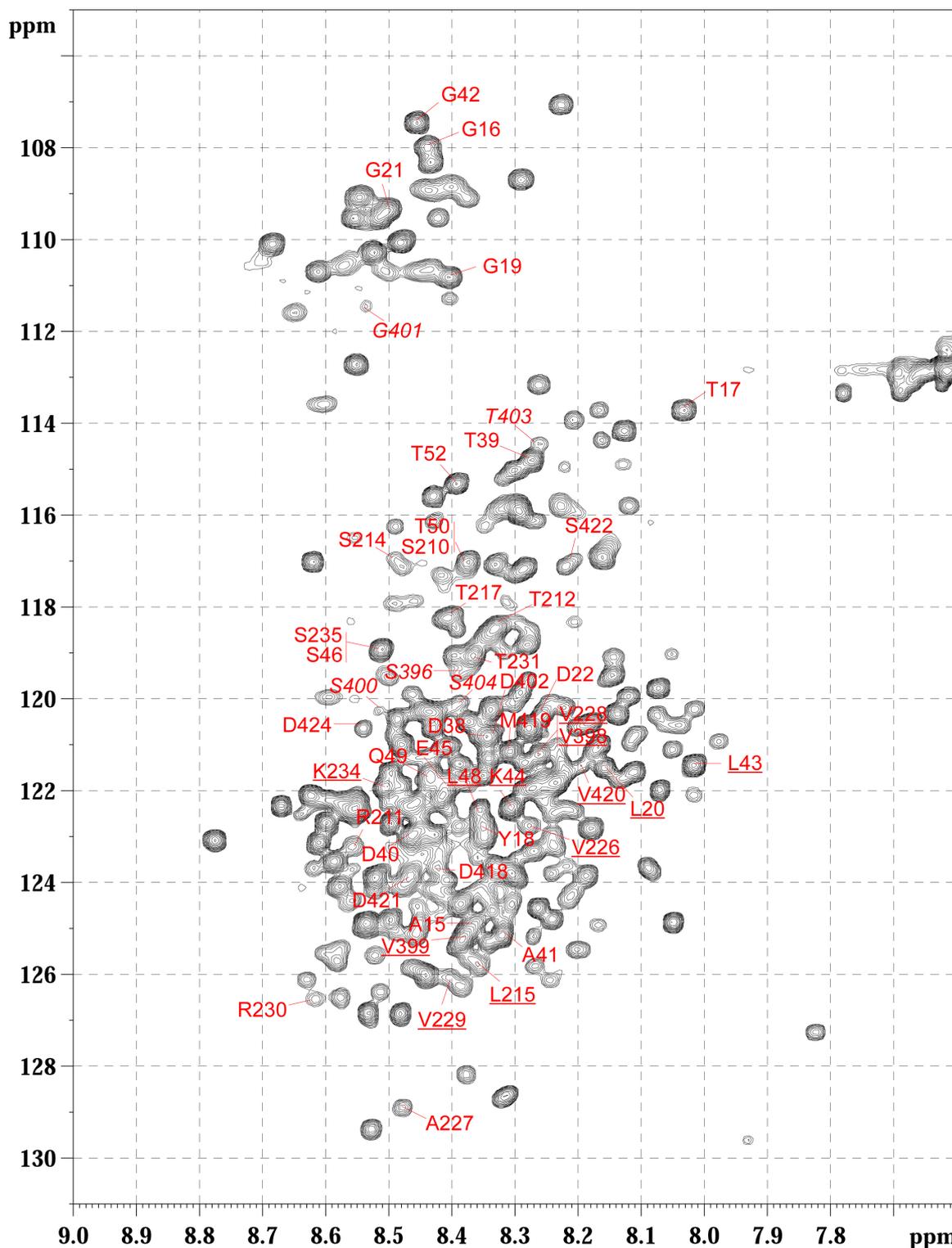
**Figure 3.16 :** Distribution des déplacements chimiques en  $^{13}C$  de la protéine tau des  $C\alpha$  (i) (en noir) obtenus à partir de l'expérience HNCA et des  $C\alpha$  (i-1) (en rouge) obtenus à partir de l'expérience HNcoCA. Les résidus Ala, Ser et Thr précédant une Pro sont identifiés à partir des valeurs de déplacements chimiques « *random coil* » données par les tables (Wishart, Bigam et al. 1995).

#### 2.1.4 – Bilan. Avancées dans l'attribution de la protéine Tau avec le « *mapping* peptidique ».

Le « *mapping* peptidique » a permis d'avancer dans l'attribution de la protéine tau mais également de lever des ambiguïtés liées à la méthode d'attribution par paires de résidus. Cette technique a mis en évidence certaines régions de la protéine qui n'adoptaient pas un même échantillonnage conformationnel que les peptides correspondants, notamment dans la partie N-terminale et dans les segments répétés. Il a été proposé que la partie N-terminale de tau puisse interagir avec la région de liaison des microtubules notamment parce que les deux régions sont reconnues simultanément par l'anticorps Alz50 et que des mutations dans l'une ou l'autre de ces régions induisaient une perte d'affinité pour l'anticorps (Carmel, Mager et al. 1996). La reconnaissance de la région N-terminale et de la région de liaison des microtubules par Alz50 semble indiquer qu'il existe une conformation « repliée » de la protéine. Cependant, l'anticorps ne permet pas d'obtenir des informations sur la dynamique car il ne permet pas de distinguer si la protéine est le plus souvent en conformation « repliée » ou en conformation « étendue ». En revanche, la RMN peut apporter ce type d'information. La comparaison des pics de corrélation  $^1\text{H}$ - $^{15}\text{N}$  des résidus N-terminaux dans le spectre du peptide et celui de la protéine montre des différences assez faibles (par exemple, la Thr17), indiquant que cette conformation « repliée » est probablement transitoire, toutes précautions gardées quant à l'effet du « repliement » sur la perturbation des déplacements chimiques.

Un autre point mérite d'être souligné sur les informations structurales fournies par le « *mapping* peptidique » et concerne les formes *cis* des prolines. Nous avons pu établir, pour le peptide [S208-R221], central dans nos études d'interaction avec Pin1, une proportion de conformères *cis* pour les résidus Pro213, Pro216 et Pro218 de l'ordre de 10%. De manière très inattendue, nous avons pu détecter et attribuer des formes minoritaires dans le spectre de la protéine tau entière, parce que ces signaux minoritaires de faibles intensités se superposaient aux signaux correspondant à la Thr212 et à la Thr217 liées à des prolines en *cis* dans le peptide et, de manière tout à fait surprenante, les proportions relatives des conformères *cis* des Pro213 et Pro218 du peptide sont conservées dans la protéine totale.

L'ensemble de ces données sur l'attribution de la protéine tau par « *mapping* peptidique » a été soumis à publication dans la revue Chembiochem. La version de soumission, contenant le détail des attributions et les informations apportées par les peptides, est fournie en fin de chapitre. Un spectre récapitulant les résonances attribuées à l'aide de cette méthode est donné **Figure 3.17**.



**Figure 3.17** : Attribution de l'isoforme la plus longue de la protéine Tau (441 résidus) par *mapping* peptidique. Les résidus soulignés correspondent à des résonances confirmées par marquage sélectif en  $^{15}\text{NLeu}$ ,  $^{15}\text{NLys}$  et  $^{15}\text{Nval}$ . Les résidus annotés en italique correspondent à la zone S396-S404 dans la partie C-terminale pour lesquels les résonances sont élargies sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine entière.

## **2.2 – Etude des interactions entre Pin1 et un mutant de tau mimant l'état hyperphosphorylé.**

La phosphorylation de tau peut être réalisée *in vitro* avec des kinases purifiées ou *in vivo* dans des cellules mitotiques. Ces méthodes permettent d'obtenir un échantillon hétérogène de protéines phosphorylées. L'hétérogénéité concerne à la fois les sites de phosphorylation et les taux de phosphorylation. Un mélange hétérogène de protéines phosphorylées, bien que difficile à caractériser, représente un modèle plus réaliste des processus de phosphorylation cellulaires par rapport aux peptides phosphorylés, pour lesquels le taux de phosphorylation par site est de 100%. La caractérisation des interactions avec Pin1 à l'échelle atomique s'avère plus complexe encore.

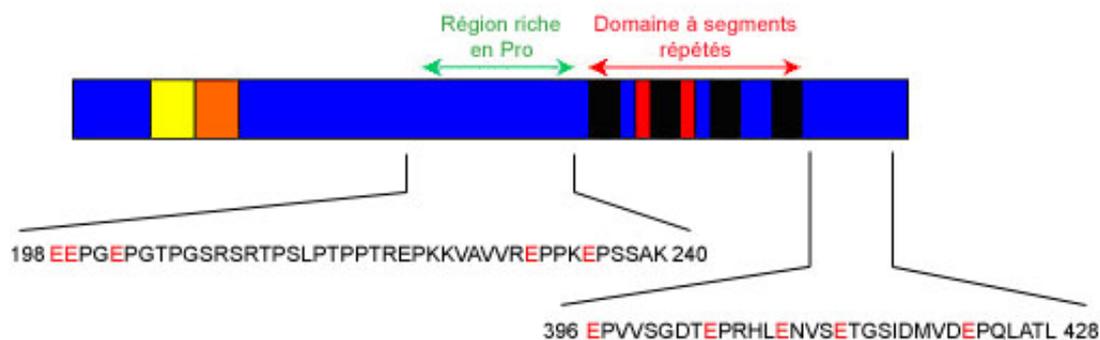
Nous avons étudié les interactions entre Pin1 et un mutant de la protéine tau mimant l'état hyperphosphorylé, pour lequel des sites de phosphorylation majeurs de tau ont été mutés en glutamate. Bien que les peptides phosphorylés aient permis de déterminer plusieurs sites d'interaction de Pin1 au sein de la protéine tau, la taille de ces mêmes peptides semble être un facteur limitant pour définir une éventuel coopération entre les deux domaines. Les mutations glutamate représentent un compromis entre les peptides et la protéine phosphorylés, dans le sens où les mutations sont introduites à des endroits précis de la protéine et miment une phosphorylation totale du site muté, comme les peptides, tout en ayant la taille de la protéine tau entière.

Nous présenterons, dans ce paragraphe, des données préliminaires concernant l'étude structurale du mutant mimant l'état hyperphosphorylé de la protéine tau par RMN et l'investigation des interactions avec Pin1. Bien que l'exploitation de ces données, limitée par l'attribution des différentes protéines, mérite d'être approfondie davantage, nous proposerons, sur base des résultats obtenus, un mécanisme fonctionnel possible permettant la restauration de la liaison de tau phosphorylée aux microtubules par Pin1.

### **2.2.1 – Le mutant Tau352E10 mime l'état hyperphosphorylé de la protéine tau (Eidenmuller, Fath et al. 2000).**

Les mutations de résidus sérine et thréonine en aspartate ou glutamate sont couramment employées pour mimer un état totalement phosphorylé. Dans la protéine tau, en particulier, les mutations des résidus Ser214 et Ser416 ont été utilisées pour étudier l'effet de la phosphorylation sur la structure de tau et sa fonction de liaison aux microtubules (Leger, Kempf et al. 1997). Il a été montré que la phosphorylation régulaient les propriétés structurales de tau et l'interaction avec les microtubules. Le processus d'agrégation pathologique semble

lié à la phosphorylation mais l’hypothèse d’une déficience de la déphosphorylation dans les filaments de protéines tau ne peut être exclue. De plus, la phosphorylation des résidus Ser262 et Ser214 empêche la liaison de tau aux microtubules mais protège également de l’agrégation. Les sites de phosphorylation caractéristiques des PHF sont localisés dans les régions encadrant le domaine de liaison aux microtubules, hormis pour la Ser262 du premier segment répété, et correspondent pour la plupart à des sites Ser/Thr-Pro (Morishima-Kawashima, Hasegawa et al. 1995). La moitié de ces sites ont été mutés en glutamate dans le mutant Tau352E10 (Figure 3.18). Les dix glutamates miment un état permanent d’hyperphosphorylation, qui se traduit par un effet conformationnel dû à l’accumulation de charges négatives, détecté par une mobilité électrophorétique réduite similaire à la protéine tau hyperphosphorylée sauvage. Pourtant, les charges négatives des glutamates ne miment pas l’effet des polyanions tels que l’héparine, nécessaires à la formation de filaments *in vitro*, car en l’absence d’héparine, les protéines sauvages phosphorylée et mutante ne forment pas de filaments (Eidenmuller, Fath et al. 2000).



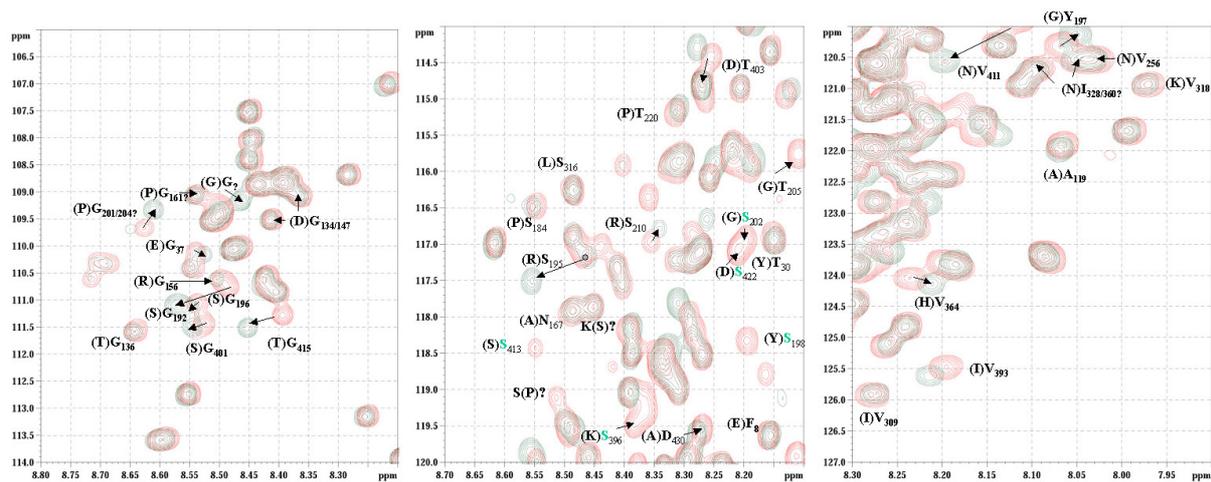
**Figure 3.18** : Représentation schématique de la protéine tau dans son isoforme la plus longue avec, en rouge, les sites de mutations glutamate dans les deux régions encadrant le domaine fonctionnel de liaison aux microtubules, l’une en N-terminal du domaine à segments répétés (région riche en prolines), l’autre en C-terminal. Toutes les mutations sont de type Ser/Glu sauf la Thr231/Glu.

Le mutant ne présente pas toutes les caractéristiques de la protéine hyperphosphorylée sauvage. La protéine Tau352E10 supprime la promotion de l’assemblage des microtubules, comme la protéine phosphorylée sauvage, mais affecte davantage la nucléation que la croissance des microtubules (Eidenmuller, Fath et al. 2000). Les mutations dans la région riche en prolines suffisent pour bloquer l’assemblage des microtubules (Eidenmuller, Fath et al. 2001). La protéine Tau352E10 agrège sous forme de filaments avec une cinétique plus lente que la protéine sauvage et présente un niveau d’agrégation plus faible. De même, la

phosphorylation de la protéine sauvage par certaines kinases, comme la PKA, réduit le niveau d'agrégation et indiquerait que l'hyperphosphorylation serait consécutive à la formation des filaments et, par là même, une conséquence du processus d'agrégation. L'absence d'interactions de la protéine mutée en glutamates avec la phosphatase PP2A indique que, probablement de manière identique, la protéine hyperphosphorylée, agrégée sous forme de filaments, n'est plus déphosphorylée efficacement. Un des mécanismes proposés pour la formation des filaments passe par une oxydation de la protéine par formation d'un pont disulfure inter-moléculaire via le résidu Cys322. Cependant, la formation de filaments avec le mutant Tau352E10 peut également être réalisée *in vitro* dans des conditions réductrices, en présence de DTT. L'ensemble de ces données montre que l'hyperphosphorylation mimée par 10 mutations glutamates contribue à augmenter le *pool* de protéines tau libres dans la cellule, en déplaçant la protéine (1) de la surface des microtubules, (2) du complexe formé avec PP2A, ainsi que (3) de la surface de la membrane corticale, compensant ainsi la plus faible propension du mutant à former des filaments (Eidenmuller, Fath et al. 2000).

### **2.2.2 – Investigations structurales du mutant Tau352E10 par RMN.**

Le mutant Tau352E10 présente des perturbations des résonances par rapport à la protéine sauvage Tau352, autres que celles correspondant au résidu muté et à ses proches voisins dans la séquence. On observe des perturbations à longue distance concernant des résidus localisés dans la partie N-terminale de la protéine (G37) et dans le domaine à segments répétés (V364), ainsi que des élargissements de raies. Ces données peuvent indiquer un réarrangement conformationnel global de la protéine causé par les dix mutations glutamates. Les variations de déplacements chimiques ne sont pas toujours corrélées avec la position du résidu correspondant par rapport au site de mutation. Par exemple, le résidu Y197, précédant la mutation de la S198, est proportionnellement beaucoup moins perturbé que les résidus G196 et S195 (Figure 3.19). Comment le remplacement des résidus Ser ou Thr par des Glu peut provoquer un changement conformationnel ? On peut avancer l'hypothèse d'un effet de charge car on remplace des résidus neutres par des résidus chargés négativement. La concentration de charges négatives localisées dans deux régions de la protéine peut entraîner une répulsion et l'« ouverture » consécutive de la protéine.

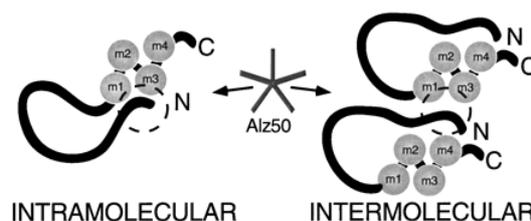


**Figure 3.19 :** Attribution partielle des perturbations des déplacements chimiques induites par les mutations glutamates dans l'isoforme courte de la protéine tau. On observe des variations importantes qui ne sont pas toujours corrélées avec la localisation du résidu correspondant, par rapport au site de mutation (par exemple, G37, S195 et V364).

Bien que la protéine tau n'adopte pas de structure stable, la flexibilité importante du squelette, dont témoigne la largeur de raie de la plupart des résonances, pourrait induire un repliement de la protéine sauvage de manière transitoire, comme nous l'avons déjà suggéré avec l'observation des peptides couvrant une partie de la région N-terminale.

### ***Hypothèse d'un repliement de la protéine tau.***

L'éventualité d'un repliement de la protéine a été proposée suite à l'isolation de l'anticorps Alz50 à partir des PHF, qui présente une sélectivité conformationnelle mais n'est pas phospho-dépendant (Carmel, Mager et al. 1996). Il a été montré que Alz50 reconnaissait un épitope discontinu de tau, c'est-à-dire un épitope se présentant sous forme de deux sites distants dans la séquence primaire de la protéine et tous deux nécessaires à la liaison de l'anticorps avec une haute affinité. Les deux sites identifiés couvrent les 18 résidus N-terminaux (reconnus également par l'anticorps M19G) et la région de liaison aux microtubules pour laquelle la localisation précise du site de liaison n'est pas bien définie (Figure 3.20). Alz50 reconnaît préférentiellement les PHF plutôt que la protéine tau recombinante, et détecte sélectivement une conformation particulière de tau dans les filaments, indiquant l'implication d'un effet conformationnel précurseur dans la formation des filaments.

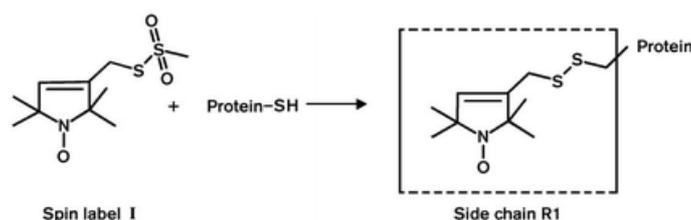


**Figure 3.20 :** Modèle proposé pour expliquer la sélectivité de l’anticorps Alz50 (Carmel, Mager et al. 1996). Alz50 reconnaît un épitope discontinu englobant à la fois la partie N-terminale et la région des domaines répétés de la protéine tau, préférentiellement de manière intra-moléculaire.

Nous avons envisagé que l’hyperphosphorylation mimée par les mutations glutamates pouvait modifier l’échantillonnage conformationnel de la protéine tau et conduire vers une conformation plus étendue qui serait contradictoire avec ces données. On peut envisager que la protéine non phosphorylée adopte également une conformation repliée différente de celle des PHF, qui ne serait par conséquent pas reconnue par Alz50. L’utilisation de marqueurs paramagnétiques, introduits sélectivement sur les résidus Cys dans les protéines, permet de détecter des résidus proches du site marqué par élargissement des résonances pour les résidus se trouvant à proximité de l’élément paramagnétique (Hubbell, Mchaourab et al. 1996). Cette technique a été utilisée, en particulier, pour détecter des contacts à longue distance dans des protéines hautement flexibles (Lietzow, Jamin et al. 2002).

### ***Marquage de spins et détection d’un effet à longue distance.***

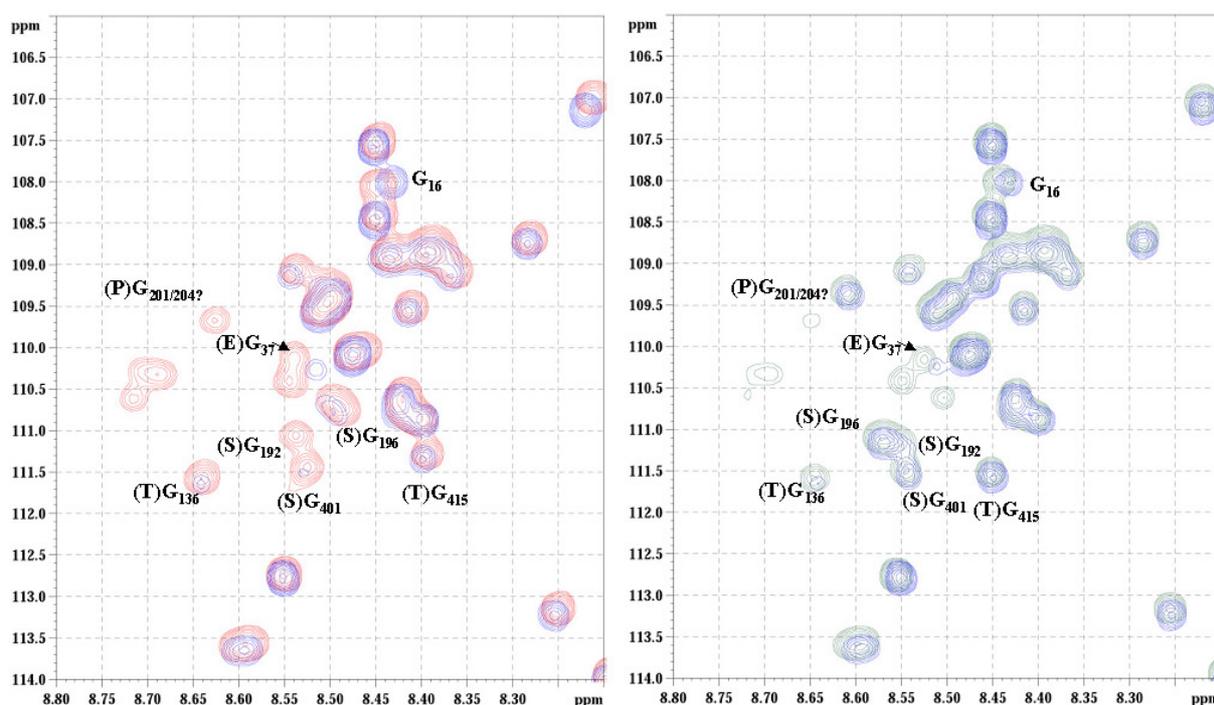
Nous avons appliqué cette technique pour détecter d’éventuels contacts à longue distance entre des résidus proches du site de marquage et des résidus localisés plus loin dans la séquence primaire de la protéine mais qui pourraient être affectés par l’effet paramagnétique en raison d’une proximité spatiale. Nous avons introduit un marqueur de type nitroxyde, le (1-oxyl-2,2,5,5-tétraméthyl-3-pyrroline-3-méthyl) méthanesulfonate, sélectivement sur l’unique cystéine de la protéine Tau352 sauvage ou Tau352E10 mutante en position 322 (Figure 3.21).



**Figure 3.21 :** Réaction de couplage du marqueur paramagnétique portant un groupement nitroxyde sur la fonction thiol d’une cystéine (Berliner, Grunwald et al. 1982; Hubbell, Mchaourab et al. 1996).

Nous avons observé deux types de comportement pour les résidus affectés par la présence du groupement paramagnétique :

- un élargissement de raies, qui est l'effet typiquement observé et exploité dans ce type d'expérience,
- un *shift* de certaines résonances. Ces *shifts* disparaissent lors de la réduction du groupement nitroxyde, effectuée comme contrôle, suggérant une influence directe du groupement paramagnétique sur la structure de la protéine. Le marqueur utilisé présente pourtant l'avantage d'une liberté rotationnelle limitée et d'une distance réduite avec la liaison disulfure (Berliner, Grunwald et al. 1982).



**Figure 3.22** : Détection de contacts à longue distance par marquage de spin avec un marqueur paramagnétique couplé à la Cys322 dans la protéine Tau352 sauvage (à gauche) et mutante Tau352E10 (à droite). Comparaison des spectres, dans la zone des glycines, de la protéine sauvage Tau352 avec (bleu) et sans marqueur (rouge) et de la protéine mutante avec (bleu) et sans marqueur (vert).

On observe un élargissement des résonances correspondant à des résidus de la région N-terminale (G37) à la fois pour la protéine sauvage et pour le mutant ainsi que des *shifts* des résonances (G16). La région riche en prolines semble affectée par l'effet paramagnétique dans la protéine sauvage mais pas dans la protéine mutée (G192, G196, G201 ou G204). La région C-terminale semble également être perturbée uniquement dans la protéine sauvage (Figure 3.22). Ces données sont préliminaires et doivent faire l'objet d'une attribution plus

systematique des résonances affectées dans les deux protéines, pour pouvoir établir un *mapping* des régions proches de la Cys322 localisée dans le domaine à segments répétés. Cependant, on peut déjà proposer que des perturbations plus étendues dans la protéine sauvage pourraient traduire un état « plus replié » par rapport au mutant, indiquant que l'effet de charge, induit par les mutations glutamates, aurait pour conséquence un effet conformationnel sur la protéine.

Nous avons examiné si des interactions entre la protéine Tau352E10 et Pin1 étaient possibles étant donné que le mutant mime l'état hyperphosphorylé de la protéine tau (Eidenmuller, Fath et al. 2000; Eidenmuller, Fath et al. 2001) et qu'il a été montré que Pin1 était capable de lier et restaurer l'activité de tau phosphorylée (Lu, Wulf et al. 1999). La spécificité de Pin1 pour les substrats phosphorylés est due à la présence de régions basiques, notamment dans le domaine catalytique (avec les résidus K63/R68/R69), permettant la liaison d'un résidu acide en position N-terminale d'une proline (Ranganathan, Lu et al. 1997). Le remplacement des résidus sérines ou thréonines phosphorylés par des glutamates ou des aspartates mime essentiellement la forme mono-anionique du phosphate dans les protéines, et pourrait permettre la liaison de Pin1 au niveau des motifs Glu-Pro de la protéine mutante. Nous pourrions ainsi mesurer l'effet des phosphorylations multiples à l'échelle de la protéine tau entière dans le processus de reconnaissance moléculaire par Pin1.

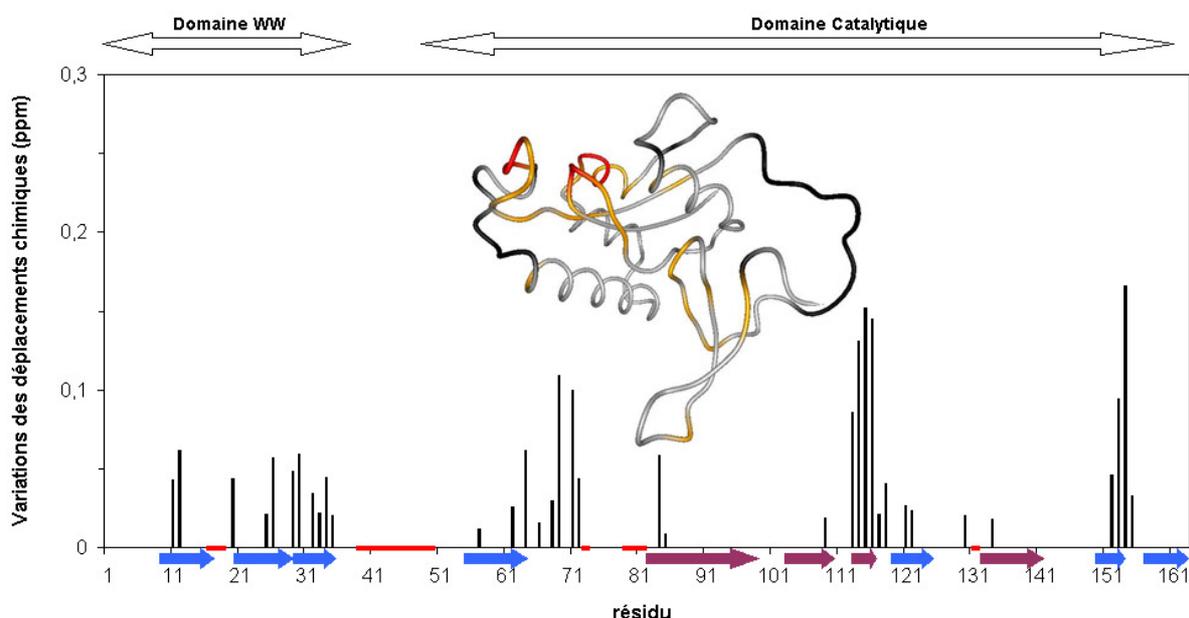
### **2.2.3 – Le domaine « catalytique » de Pin1 interagit avec le mutant Tau352E10.**

L'avantage de l'utilisation de protéine isotopiquement enrichie en RMN est qu'il est possible d'observer des interactions entre deux protéines en ne regardant que l'une d'entre elles. On peut ainsi regarder les perturbations des signaux de Pin1 sans être gêné par les signaux de la protéine Tau352E10, et inversement. Les conditions mises en œuvre pour l'ensemble des expériences décrites dans les paragraphes suivants utilisent des solutions de protéines non enrichies ou uniformément enrichies en  $^{15}\text{N}$ , à des concentrations de 100  $\mu\text{M}$  en tampon Tris deutéré à 25 mM pH 6.80, contenant 50 mM NaCl, 2.5 mM DTT/EDTA.

Nous avons observé des perturbations assez importantes des signaux de Pin1 en présence d'une quantité stœchiométrique de Tau352E10 et, de manière surprenante, avec les résidus du site catalytique essentiellement. Les effets observés proviennent plutôt d'un *binding* que d'une catalyse, bien que l'interaction concerne le site actif de Pin1, car les quantités utilisées ne sont pas catalytiques.

### **Mode de liaison de Pin1 à la protéine Tau352E10.**

Un *mapping* du site d'interaction de Pin1 a permis de définir plus précisément les régions impliquées dans la liaison du substrat (Figure 3.23) et montre de faibles interactions avec le domaine WW. Les perturbations concernent essentiellement la région autour du site de liaison de la proline et s'étendent à une partie du feuillet  $\beta$ . Le site catalytique de Pin1 semble le plus affecté par le mutant, notamment au niveau de la zone Cys113-Ala118, intervenant dans la liaison du lien peptidique du motif X-Pro (Ranganathan, Lu et al. 1997), ainsi que la région Thr152-Gly155. Un élargissement des résonances du *linker* est observé et peut soit traduire une interaction directe avec le substrat, qui n'a cependant jamais été montrée avec les substrats peptidiques (Jacobs, Saxena et al. 2003), soit être une conséquence indirecte de l'interaction avec le site catalytique par modification de la dynamique de ce fragment flexible.



**Figure 3.23 :** *Mapping* des déplacements chimiques de Pin1 uniformément enrichie en  $^{15}\text{N}$  en présence d'une quantité équimolaire de Tau352E10. Les interactions se situent essentiellement dans le site catalytique au niveau des résidus C113-A118 et T152-G155. On note de faibles interactions avec le domaine WW et un élargissement important des résidus constituant le *linker* qui lie les deux domaines. Les traits rouges indiquent les signaux élargis qui disparaissent du spectre. Les perturbations supérieures à 0,1 ppm sont annotées en rouge sur la structure de Pin1, celles comprises entre 0,05 et 0,1 ppm, en orange, et les signaux élargis, en noir.

On note, en particulier, l'apparition d'un signal à 5,8 ppm sur le spectre de Pin1. Ce signal a été attribué au résidu Lys77, homologue à la Lys22 de Pin1 d'*Arabidopsis thaliana*, dont la structure a été déterminée dans notre laboratoire. En effet, le résidu Lys22, localisé dans une

boucle flexible du site actif, responsable de la liaison du phosphate, n'est observable qu'en présence de sulfate qui permet une stabilisation de la boucle (Landrieu, Wieruszeski et al. 2002). Ce signal, qui n'avait jamais été observé avec les peptides phosphorylés, semble indiquer une meilleure stabilisation de la boucle du site actif par le mutant, en l'absence de phosphorylation.

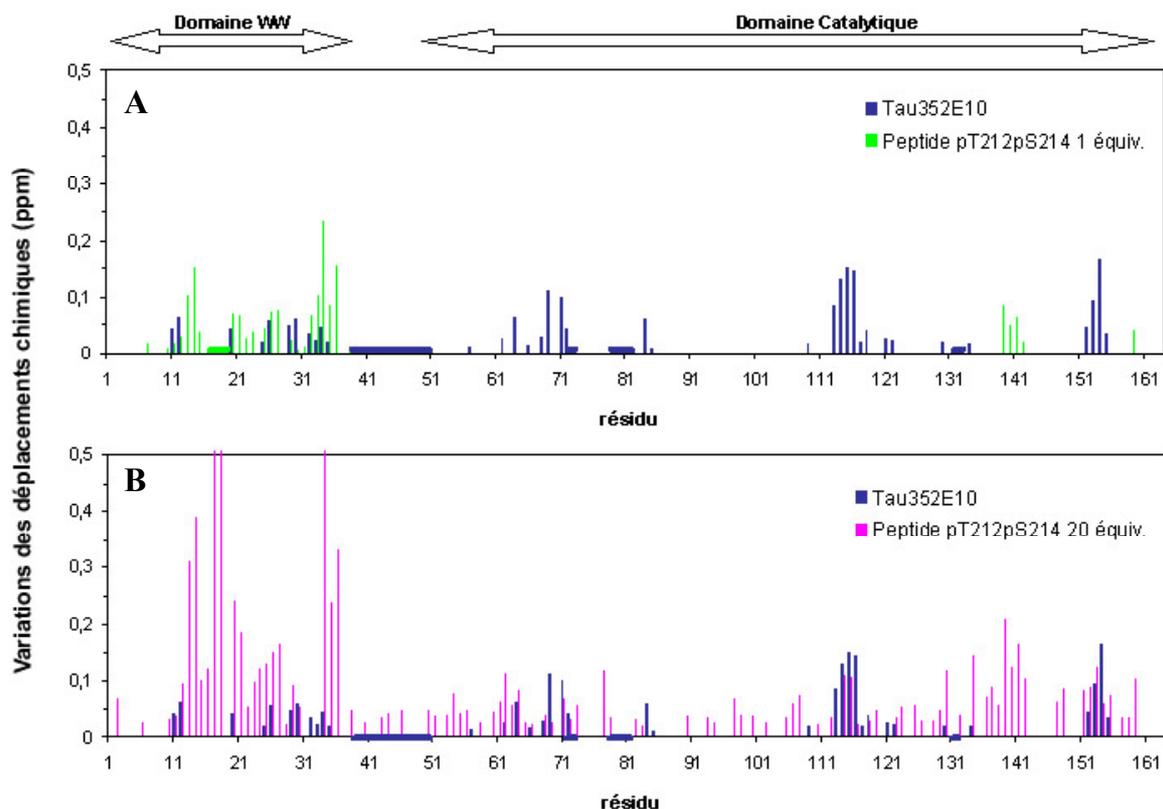
***Une spécificité d'interaction différente des peptides phosphorylés.***

Les perturbations induites par le mutant Tau352E10 au niveau du domaine WW, sont faibles comparées à celles induites par les peptides phosphorylés dans les mêmes proportions (ratio molaire 1 : 1), conditions dans lesquelles, avec les substrats peptidiques, aucune perturbation du site actif n'était observée. En revanche, les *shifts* induits par Tau352E10 sur le site actif sont reproduits avec 20 équivalents de peptides phosphorylés (Figure 3.24).

Ces données indiquent clairement que la protéine tau mutante ne reproduit pas le même mode de liaison que les peptides phosphorylés en ne faisant pas intervenir le module de liaison de Pin1, qui était le principal acteur dans la liaison des substrats peptidiques. Ce phénomène peut provenir de l'absence de phosphate dans le mutant Tau352E10 et indiquer une phosphospécificité plus accentuée du domaine WW par rapport au domaine catalytique. Le site actif de Pin1 serait capable de lier, comme cela avait déjà été montré (Ranganathan, Lu et al. 1997), divers motifs X-Pro où X serait un résidu acide, alors que le domaine WW ne pourrait lier que les motifs phosphoSer/Thr-Pro.

D'autre part, les amplitudes identiques des signaux du site actif induites par 20 équivalents de peptide phosphorylé et un seul équivalent de protéine mutante indiquent :

- soit la liaison de plusieurs sites distincts du mutant avec des affinités relativement faibles, similaires à celle des peptide,
- soit la liaison d'un site unique avec une plus haute affinité.



**Figure 3.24 :** Mapping des déplacements chimiques des résidus de Pin1 en interaction avec 1 équiv. de protéine mutante Tau352E10 (en bleu) comparé avec le *mapping* induit par (A) 1 équiv. (en vert) et (B) 20 équiv. de peptide di-phosphorylé pT212 pS214 (en rose).

Comme il existe quatre motifs Glu-Pro dans la protéine tau sauvage (aux positions 3, 58, 187 et 222 dans Tau441), nous avons vérifié que l'effet observé avec Tau352E10 était spécifique à la protéine mutante et qu'il n'était pas dû à la liaison des motifs Glu-Pro présents dans la protéine sauvage. Les contrôles ont été réalisés dans les mêmes conditions (ratio molaire 1 :1) avec les isoformes longue (Tau441) et courte (Tau352) de la protéine sauvage et ont montré qu'il n'existait aucune interaction notable avec Pin1. Les faibles perturbations qui ont été observées au niveau du domaine WW uniquement, peuvent provenir d'une liaison de faible affinité des motifs Glu-Pro et excluent l'implication de ces motifs dans la liaison du mutant Tau352E10 par le site actif.

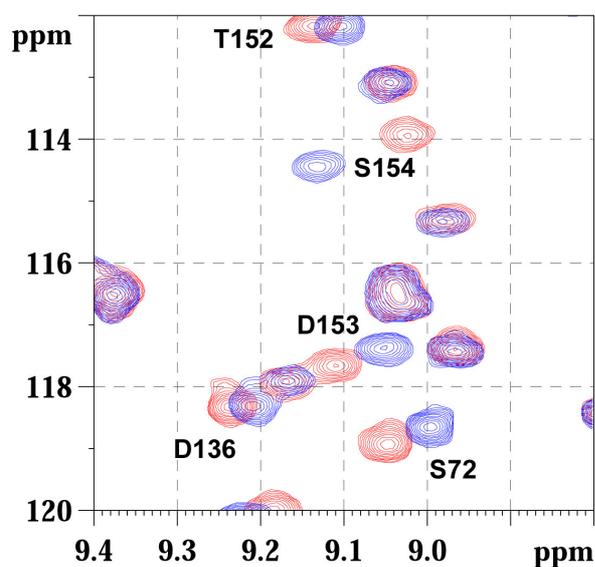
***Mode de liaison de Pin1 au mutant Tau352E10 : une simple liaison des motifs Glu-Pro ?***

Afin de déterminer si Pin1 reconnaît le mutant au travers d'une simple liaison des motifs Glu-Pro, nous avons utilisé un substrat peptidique, provenant de la région riche en prolines (fragment [224-240]), centré autour du site de phosphorylation Thr231, que nous avons muté en glutamate et en aspartate. Aucune interaction entre Pin1 et le peptide [224-240]pT231E n'a

été observée, même en présence de 20 équivalents de peptide, hormis de faibles perturbations des résidus du domaine WW impliqués dans la liaison de la proline (S32-E35) et du phosphate (R17-S18), qui pourraient lier le motif Glu-Pro avec une très faible affinité ( $K_D > 5$  mM). La mutation glutamate, mimant l'état phosphorylé de la Thr231, inhibe totalement les interactions du peptide avec Pin1 et montre qu'aucun des domaines n'est capable de lier les motifs Glu-Pro. Nous pouvons donc établir que Pin1 ne reconnaît pas le mutant Tau352E10 en liant simplement les motifs Glu-Pro mais utilise un mécanisme plus complexe, qui reste à définir. On peut envisager par exemple, la reconnaissance d'un motif Glu-Pro particulier faisant intervenir les résidus environnants, soit dans la séquence primaire directe entourant le motif, soit à proximité dans l'espace environnant celui-ci.

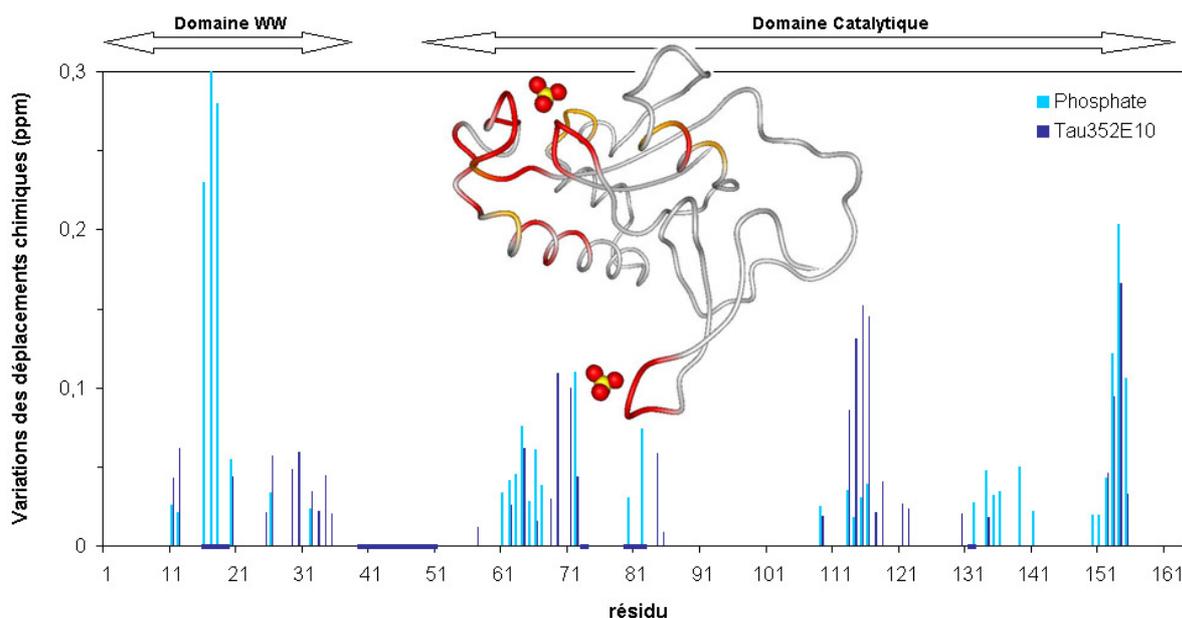
### *Compétition avec le phosphate*

Afin de préciser le mode de liaison du mutant Tau352E10, nous avons réalisé une compétition avec le phosphate inorganique, dont l'affinité pour Pin1 a été évaluée à 2 mM environ (Bayer, Goettsch et al. 2003). Une étude préliminaire de l'interaction du phosphate avec Pin1 a été réalisée dans les conditions habituelles, en présence de 25 mM de phosphate (tampon  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  pH 6.80), soit 250 équivalents. On note une perturbation importante des déplacements chimiques des résidus Leu61-Ser72 de la boucle  $\alpha 1/\beta 1$  du domaine catalytique impliquée dans la liaison du sulfate, d'après la structure cristallographique de Pin1 en complexe avec le dipeptide Ala-Pro (Ranganathan, Lu et al. 1997), et de la zone Thr152-Gly155 (Figure 3.25).



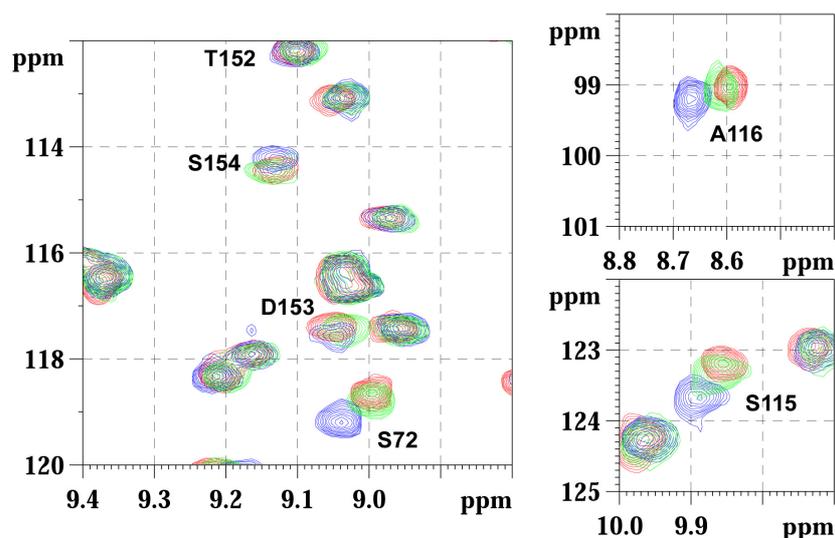
**Figure 3.25 :** Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 100  $\mu\text{M}$  (rouge) et en présence de 25 mM phosphate (bleu).

Les perturbations induites par le mutant Tau352E10 concernent principalement les résidus impliqués dans la liaison du phosphate dans le site actif de l'enzyme (Figure 3.26), mais ne recouvrent pas l'ensemble des perturbations induites par le phosphate, notamment au niveau du domaine WW (Ser16-Ser19). Inversement, les perturbations induites par le phosphate ne recouvrent pas la totalité des perturbations provoquées par le mutant, en particulier dans la région Cys113-Ala118 et pour le *linker* inter-domaine Ser38-Glu51.



**Figure 3.26 :** Mapping des déplacements chimiques de Pin1 en présence d'une quantité équimolaire de protéine mutante Tau352E10 (en bleu foncé) ou en présence de 25 mM de phosphate (en bleu clair). Les interactions avec le mutant se situent essentiellement dans le site catalytique, au niveau des résidus C113-A118 et T152-G155, la seconde région étant également fortement impliquée dans la liaison du phosphate. On note un élargissement important de tous les résidus constituant le *linker* qui lie les deux domaines (indiqué par les traits bleus). Les régions colorées sur la structure de Pin1 correspondent aux résidus perturbés par le phosphate inorganique, présentant des variations supérieures à 0.1 ppm (rouge) et comprises entre 0.05 et 0.1 ppm (orange).

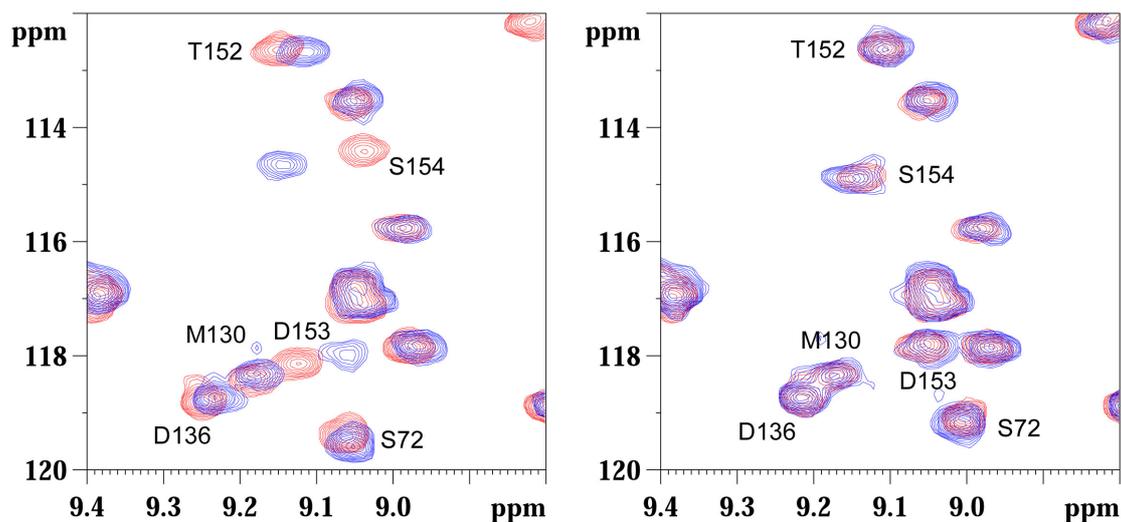
La compétition avec le phosphate permet de déterminer plus finement la nature des interactions entre Pin1 et le mutant Tau352E10. L'ajout de 250 équivalents de phosphate à la solution de Pin1 contenant une quantité stœchiométrique de Tau352E10, n'inhibe pas totalement l'interaction entre les deux protéines, puisque des perturbations subsistent encore pour certaines résonances (S115, A116 Figure 3.27). D'autre part, on ne retrouve pas les résonances des résidus du *linker*, indiquant également que la liaison du mutant n'est pas totalement déplacée. On constate également que certaines résonances retrouvent la fréquence qu'elles adoptent en présence du phosphate uniquement (S72, T152-S154 Figure 3.27).



**Figure 3.27** : Comparaison des perturbations induites par le phosphate sur l'interaction Pin1/ Tau352E10. En rouge Pin1 et 25 mM phosphate, Pin1 : Tau352E10 1 : 1 sans phosphate (en bleu) et avec phosphate (en vert).

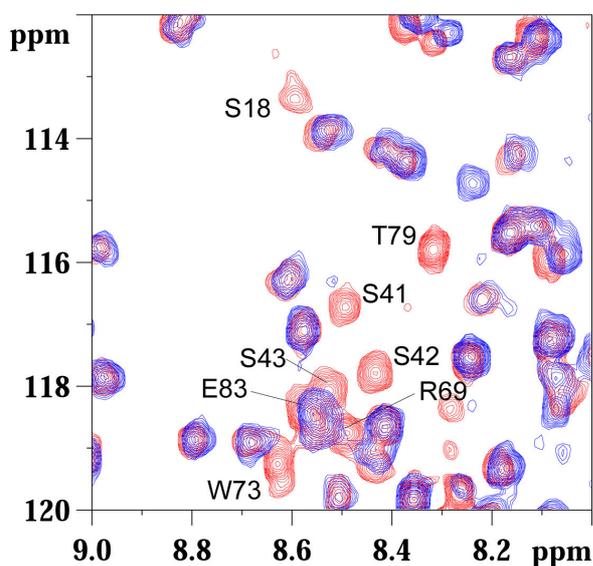
Si on regarde plus précisément les variations des résonances du site actif, on remarque que l'ajout de 25 mM de phosphate dans l'échantillon contenant les deux protéines en quantité équimolaire a deux effets sur les résonances de Pin1 :

- Certains résidus Thr152-Ser154 sont peu ou pas perturbés par la présence du mutant dans ces conditions (Figure 3.28), ce qui n'est pas surprenant puisqu'il s'agit d'un des sites les plus fortement affectés par le phosphate. Il est important de noter que les perturbations des résonances induites par 25 mM phosphate concernant ce site de liaison sont identiques à celles induites par 100  $\mu$ M de mutant (comparaison Figure 3.28/ Figure 3.26).
- Certains résidus retournent vers la fréquence de résonance qu'ils adoptent en présence du phosphate seul, suggérant que le phosphate déplace le mutant du site de liaison de Pin1. Certains résidus du site catalytique impliqués directement dans la liaison du phosphate (R68, R69) ainsi que les résidus proches (W73, T79, C113-A116) voient leur résonances perturbées par la protéine Tau352E10 (100  $\mu$ M), même en présence de 25 mM de phosphate (Figure 3.29), mais de manière plus restreinte. Ces données indiquent que le phosphate n'inhibe pas la totalité de l'interaction pour déplacer complètement le mutant.



**Figure 3.28 :** Interaction entre Tau352E10 (100  $\mu$ M) et Pin1 (100  $\mu$ M) avec ou sans phosphate (25 mM), indiquant les sites de liaison du phosphate occupés par la protéine mutante Tau352E10. Spectre de Pin1 seule (rouge) et en présence du mutant Tau352E10 (bleu) sans phosphate (à gauche), avec 25 mM de phosphate (à droite).

D'autre part, aucune perturbation n'est observée avec le mutant pour les résidus de l'hélice  $\alpha 4$  et du feuillet  $\beta 6$  formant l'interface avec le domaine WW, comme nous l'avons montré pour certains peptides phosphorylés. L'absence de rapprochement des deux domaines peut s'expliquer, dans le cas présent, par l'absence d'interaction avec le domaine WW. L'élargissement des résonances du *linker* inter-domaine montre une modification de la dynamique de celui-ci. Une dynamique différente des deux domaines en présence du substrat est probablement à l'origine de ce phénomène, et l'absence d'interaction avec le domaine WW, en particulier, peut en être la cause.



**Figure 3.29 :** Interaction entre Pin1 et Tau352E10 en présence de 25 mM de phosphate. En rouge, spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 à 100  $\mu$ M en présence de 25 mM de phosphate, en bleu, spectre de Pin1 dans les mêmes conditions avec 100  $\mu$ M de Tau352E10.

En résumé, on peut définir un mode de liaison du mutant Tau352E10 basé sur la liaison du site actif de Pin1 uniquement, et concernant principalement la région impliquée dans la liaison du phosphate, pour laquelle on observe une compétition partielle avec le phosphate inorganique. Mais il existe aussi un autre site d'interaction entre Pin1 et le mutant, caractérisé par les résidus Cys113-Ala118, localisé à proximité de la région liant le phosphate. Ce second site ne fait pas l'objet d'une compétition directe avec le phosphate, mais la liaison de ce site par le mutant est conditionnée par la liaison du site du phosphate. Ces données montrent que la liaison du site du phosphate de Pin1 est essentielle à la liaison du substrat.

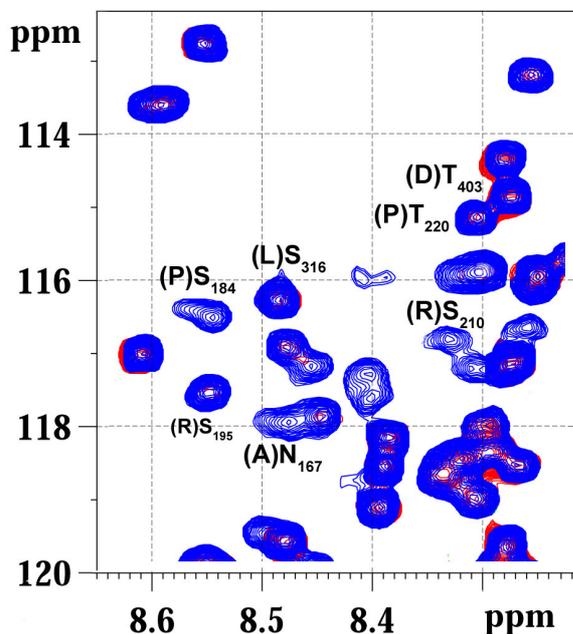
Le site défini par la région Cys113-Ala118, impliqué à l'origine dans la liaison des motifs X-Pro (Ranganathan, Lu et al. 1997), représente un point d'ancrage supplémentaire du mutant pour lequel la liaison n'est pas directement déplacée par le phosphate, mais indirectement par inhibition de la liaison du site du phosphate. De plus, on remarque que, contrairement aux phospho-peptides, le site de liaison de la proline (Leu122, Met130, F134) n'est pas affecté, indiquant que ce site n'est pas occupé par le mutant. L'ensemble de ces données montrent que 250 équivalents de phosphate, dont la constante de dissociation avec Pin1 est relativement faible (2 mM), ne suffisent pas à déplacer totalement le complexe formé par Pin1 et le mutant Tau352E10 (ratio 1 : 1), en particulier parce que l'affinité de Pin1 pour le mutant est probablement plus élevée que pour le phosphate.

Nous avons regardé les manifestations de l'interaction de Pin1 avec le mutant Tau352E10 en observant les perturbations des résonances de ce dernier. L'absence de structure et la limite imposée par l'attribution des signaux, représentent autant d'obstacles pour l'interprétation des données, mais semblent incontournables pour déterminer le site d'interaction au niveau du mutant et évaluer l'effet de la liaison de Pin1. Nous allons présenter des données préliminaires, qui suggèrent, pour la première fois, un mode d'interaction fonctionnel entre Pin1 et un substrat protéique à l'échelle moléculaire.

#### ***Investigations des perturbations induites par Pin1 au niveau du mutant Tau352E10.***

Au premier abord, l'effet de Pin1 sur la protéine mutante Tau352E10 semble induire de très faibles perturbations de cette dernière. On observe quelques *shifts* concernant un nombre limité de résonances. L'effet le plus inattendu se traduit par l'apparition de nouvelles résonances sur le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Figure 3.30). Quand on descend le seuil de détection des signaux, on s'aperçoit que ces résonances, qui apparaissent en présence de Pin1, existent

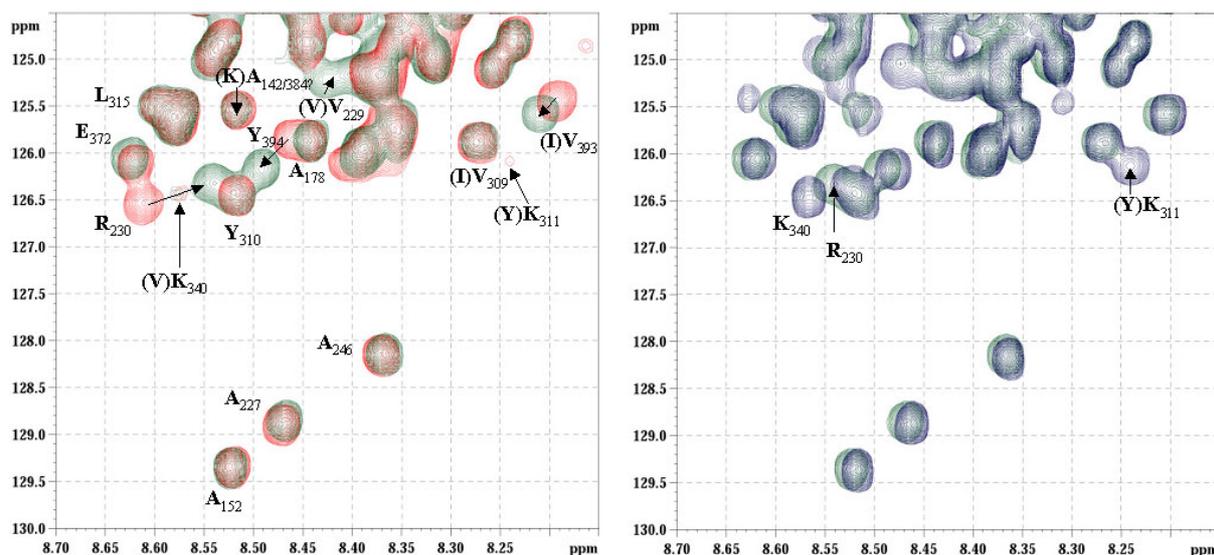
dans le mutant Tau352E10 mais sont très élargies, donc à peine détectables. Donc Pin1 a un effet d'« affinement » sur la largeur de certaines résonances ce qui traduit davantage un effet dynamique qu'un effet de liaison.



**Figure 3.30** : En rouge, spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine mutante Tau352E10 seule à 100  $\mu\text{M}$  et, en bleu, avec une quantité stœchiométrique de Pin1 (non enrichie).

Là où cela devient très intéressant, c'est que la plupart de ces résonances, élargies dans le mutant, sont observables dans la protéine Tau352 sauvage avec le même niveau de signal que l'ensemble des résonances, bien qu'elles ne concernent pas des résidus proches du site de mutation. Nous avons suggéré un effet conformationnel au niveau du mutant Tau352E10 induit par la présence des glutamates (paragraphe 2.2.2), qui pourrait être à l'origine de ces perturbations à longue distance. Des perturbations plus étendues pour la protéine sauvage par rapport à la protéine mutante, en présence d'un groupement paramagnétique introduit sur la Cys322 sélectivement, suggèrait une conformation plus « repliée » de la protéine sauvage. Il se pourrait alors, que Pin1 inverse la tendance, en « repliant » la protéine mutante dans une conformation proche de celle de la protéine sauvage.

En présence de Pin1, on observe l'apparition de signaux qui étaient élargis à la fois dans le mutant Tau352E10 et dans la protéine sauvage Tau352 (K311, K340). Ce phénomène pourrait traduire une stabilisation de ces résidus par la liaison de Pin1. On observe également un *shift* pour la résonance du résidu R230 proche du site de mutation T231E qui proviendrait d'une liaison classique de Pin1 (Figure 3.31).



**Figure 3.31** : Interactions entre  $^{15}\text{N}$ Tau352E10 et Pin1. Spectres de Tau352 sauvage (en rouge), de Tau352E10 (en vert), et de Tau352E10 : Pin1 1 : 1 (en bleu). L'interaction avec Pin1 se traduit par un effet sur la largeur de raie de certains résidus (K311, K340), qui n'apparaissent ni dans le spectre de Tau352, ni dans celui de Tau352E10.

L'interaction de Pin1 avec la protéine Tau352E10 aboutit donc à deux effets : un effet de liaison classique et un effet conformationnel, qui sont difficiles à discerner par simple *mapping* des déplacements chimiques. L'interprétation de la variation des signaux est donc plus délicate car elle nécessite la déconvolution des perturbations dues au *binding* et dues au changement conformationnel.

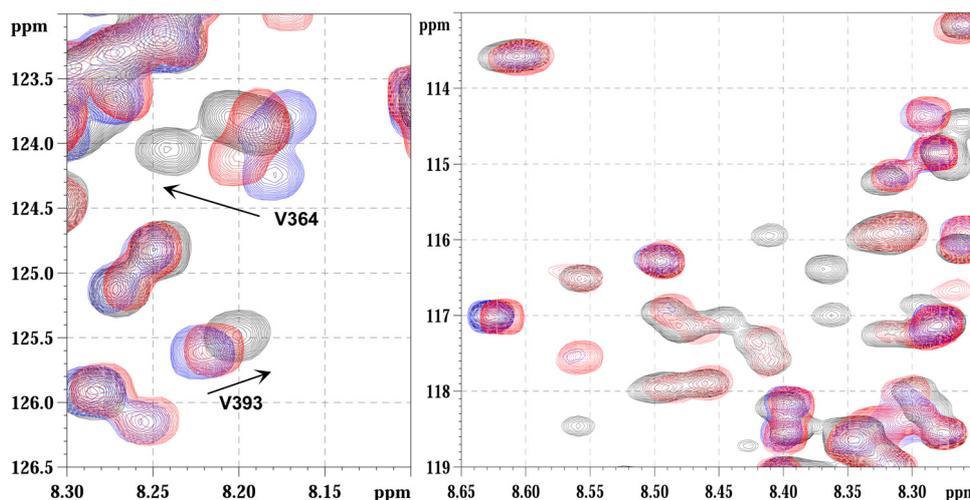
***Pin1 restaure-t-elle la structure de la protéine tau sauvage pour le mutant mimant l'état hyperphosphorylé ?***

On peut supposer que l'accumulation de charges négatives, due aux glutamates, au sein du mutant Tau352E10 induit des modifications conformationnelles, par simple répulsion électrostatique. On peut par ailleurs prouver un certain « dépliement » de la protéine mutante par rapport à la protéine sauvage sur base (1) des perturbations des résonances entre les deux protéines, accompagnées d'un élargissement important de certains signaux, qui vont au-delà des zones de mutation, et (2) des perturbations induites par un groupement paramagnétique localisé dans la région des segments répétés, avec un effet à longue distance plus prononcé dans la protéine sauvage que dans la protéine mutante. L'interaction de Pin1, sélective de la protéine mutante Tau352E10, produit deux effets distincts sur les résonances de Tau352E10 : (1) un *shift* concernant un nombre restreint de résonances et (2) l'affinement de certains signaux, qui retrouvent une largeur de ligne comparable à celle qu'ils avaient dans la protéine

sauvage Tau352. Quel est l'effet produit au niveau moléculaire qui aboutit à un tel élargissement ou affinement des signaux ? On peut envisager une modification de la dynamique de l'échantillonnage conformationnel entre les protéines sauvage et mutante, ou un échantillonnage conformationnel différent. L'interaction avec Pin1 pourrait stabiliser certaines régions du mutant par simple liaison et Pin1 pourrait être, dans ce contexte, une protéine chaperonne. Quels sont les sites de liaison de Pin1 ou est-ce qu'il s'agit d'un site unique ? Il se peut que Pin1 ne lie qu'un seul site au sein du mutant Tau352E10 mais qu'un réarrangement conformationnel à longue distance puisse induire une perturbation des signaux dans diverses régions de la protéine, comme nous l'observons. A l'heure actuelle, nous ne sommes pas capables de dissocier l'effet de *binding* de l'effet conformationnel pour répondre à cette question. L'effet de Pin1 est multiple sur les résonances de la protéine tau (Figure 3.32) :

- Un affinement de certaines résonances du mutant, qui étaient élargies par rapport aux mêmes résonances dans la protéine sauvage,
- Un affinement des résonances du mutant, qui étaient également élargies dans la protéine sauvage,
- Un *shift* des résonances du mutant vers les résonances de la protéine sauvage (représenté par les flèches sur la Figure 3.32 pour les résidus V364, V393),

L'ensemble de ces comportements suggère une restauration de la structure de la protéine sauvage au sein du mutant Tau352E10.

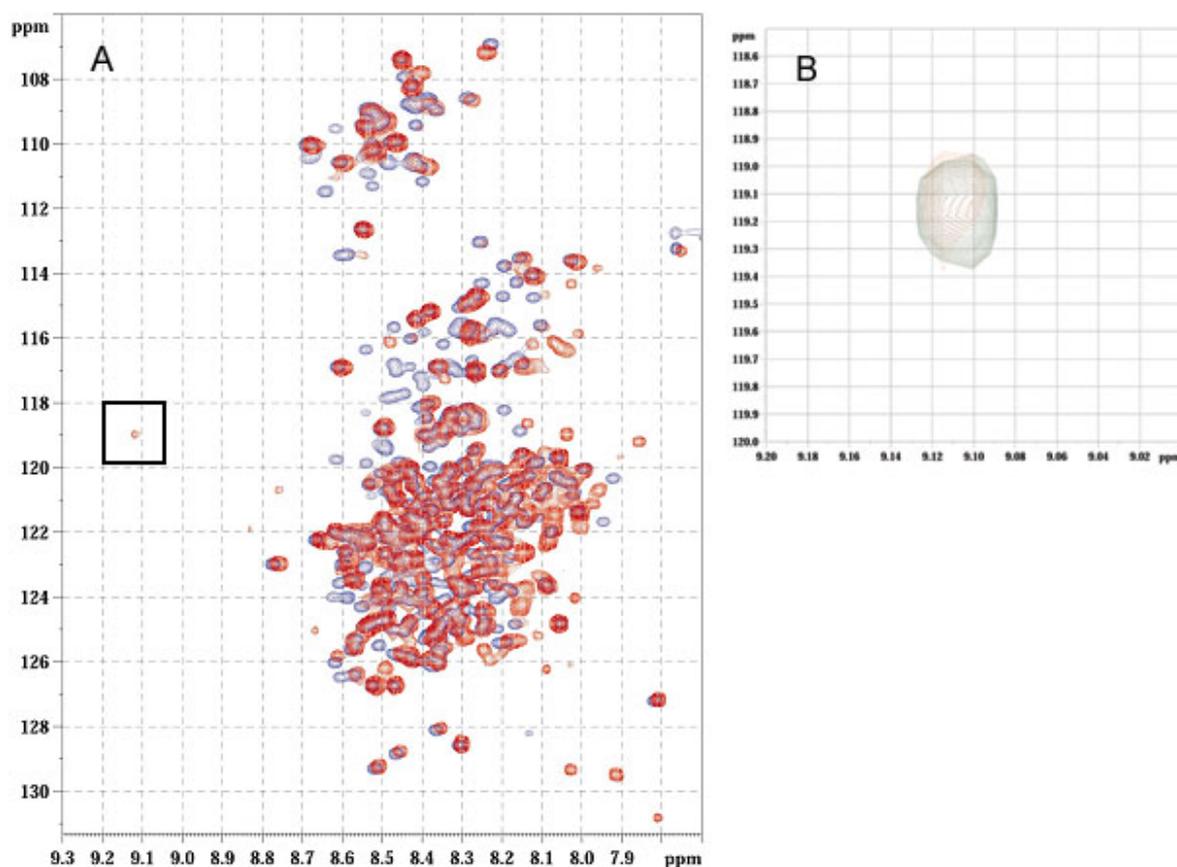


**Figure 3.32** : Spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC de la protéine tau : en noir, Tau352 100  $\mu\text{M}$ , en bleu, Tau352E10 100  $\mu\text{M}$ , en rouge Tau352E10 100  $\mu\text{M}$  : Pin1 100  $\mu\text{M}$ .

***Perspectives concernant l'interaction de Pin1 avec la protéine tau phosphorylée.***

On peut envisager un effet plus drastique de Pin1 en présence d'un substrat phosphorylé, en particulier grâce à l'intervention du domaine WW, qui ne reconnaît pas le mutant Tau352E10. Il est probable que, dans l'étude que nous avons décrite précédemment, l'utilisation de quantités stœchiométriques de Pin1 par rapport au substrat Tau352E10 soient nécessaires parce que le domaine WW ne participe pas à la liaison du substrat. Nous avons montré dans la première partie de ce chapitre que le domaine WW pouvait jouer un rôle dans l'ancrage de Pin1 au substrat et induire une augmentation de la concentration locale en enzyme autour de la protéine substrat. L'utilisation de quantités catalytiques de Pin1 (1/100<sup>e</sup>) permettait de détecter un effet sur l'isomérisation *cis/trans* au niveau du motif phosphoThr212-Pro213. On peut envisager le même mécanisme, à l'échelle de la protéine tau entière, avec ancrage de Pin1 via le domaine WW, à condition que la protéine tau soit phosphorylée. On peut donc supposer que, le mutant n'étant pas reconnu par le domaine WW, il faille augmenter la concentration de l'enzyme, jusqu'à un ratio molaire 1 :1 (voire même davantage) entre Pin1 et Tau352E10, pour pouvoir détecter un effet. Nous avons commencé à évaluer des conditions de phosphorylation de la protéine tau avec un suivi de la cinétique et la détermination des sites de phosphorylation par RMN. Nous envisageons d'utiliser des kinases purifiées GSK3 $\beta$  et PKA, qui sont capables de phosphoryler l'épitope Thr212/Ser214 reconnu par l'anticorps AT100 (Zheng-Fischhofer, Biernat et al. 1998), ou des extraits cellulaires issus de cellules mitotiques (de souches neuroblastomes SY5Y surexprimant l'activation p25 de la CDK5, grâce à une collaboration avec l'équipe du Dr Luc Buée). Les premiers essais de phosphorylation, avec les extraits mitotiques, réalisés directement dans le tube RMN, ont montré une perturbation importante du spectre de la protéine tau sauvage (Tau441), notamment dans la région des sérines et thréonines, qu'il est difficile d'interpréter. Cependant, l'apparition d'une résonance de faible intensité à bas champ, que l'on observe typiquement pour les résidus phosphorylés dans les peptides, permet d'envisager une phosphorylation de la protéine (Figure 3.33A). La comparaison avec le spectre <sup>1</sup>H-<sup>15</sup>N HSQC du peptide long de la région riche en prolines, utilisé dans la première partie de ce chapitre, comportant trois résidus phosphorylés aux positions Thr212, Ser214 et Thr231, montre de manière très surprenante, une parfaite superposition de la résonance à 9.11 ppm de la protéine tau avec la résonance de la phosphoSer214 du peptide (Figure 3.33B). D'autres sites Ser/Thr de la protéine tau ont pu être phosphorylés, mais le niveau de phosphorylation est probablement trop bas pour être détectable. La caractérisation des sites de phosphorylation par immunodétection a été réalisée

par l'équipe du Dr Luc Buée, et a mis en évidence un profil de phosphorylation identique aux données publiées (Hamdane, Sambo et al. 2003).



**Figure 3.33 :** (A) Phosphorylation de la protéine tau par un extrait mitotique de cellules SY5Y de type neuroblastomes par activation de p25 (rouge) avec le contrôle (bleu). (B) Comparaison du signal apparaissant à 9.11 ppm sur le spectre (A) avec le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du peptide de 40 résidus [201-240] comportant trois sites de phosphorylation en Thr212, Ser214 et Thr231 (noir), avec un agrandissement de la région autour du signal de la phosphoSer214.

Les modifications observées sur le spectre de la protéine lors de l'incubation avec les extraits mitotiques de cellules SY5Y, peuvent provenir de la phosphorylation de plusieurs sites Ser/Thr et éventuellement d'un réarrangement conformationnel consécutif à la phosphorylation, comme nous l'avons observé pour le mutant Tau352E10. A l'heure actuelle, nous poursuivons les essais de phosphorylation de tau *in vitro*, dans le but d'étudier le rôle structural de la phosphorylation sur la protéine tau soluble et dans le processus d'agrégation (par exemple, pour la détection et l'attribution des formes minoritaires liées à des prolines en conformation *cis*, et leur rôle dans l'agrégation). Nous voulons également déterminer, comme nous l'avons fait pour les phospho-peptides, le rôle de Pin1 dans la liaison de substrats multi-phosphorylés, et les conséquences structurales et fonctionnelles de cette interaction.



## **NMR peptide mapping of the neuronal Tau protein.**

Caroline Smet<sup>[a]</sup>, Alain Sillen<sup>[a]</sup>, Arnaud Leroy<sup>[a,b]</sup>, Jean-Michel Wieruszeski<sup>[a]</sup>, Isabelle Landrieu<sup>[a]</sup> and Guy Lippens\*<sup>[a]</sup>,

<sup>[a]</sup> C. Smet, Dr. A. Sillen, Dr. A. Leroy, Dr. J.-M. Wieruszeski, Dr. I.

Landrieu, Dr. G. Lippens

CNRS – Université de Lille 2 UMR 8525, Institut Pasteur de Lille,

BP 245 – F-59019 Lille Cedex France.

Fax: (+33) 3-20-87-12-33

e-mail: [guy.lippens@ibl.fr](mailto:guy.lippens@ibl.fr).

<sup>[b]</sup> Dr. A. Leroy, Laboratoire de biochimie appliquée, Faculté de Pharmacie à Châtenay-Malabry (Paris XI), Tour D4 2ème étage, 5 rue Jean-Baptiste Clément, 92296 Chatenay-Malabry Cedex, France.

## **Abstract**

Extracting residue specific structural information from the NMR spectra of the neuronal Tau protein has proved a real challenge. We propose a new strategy based on “peptide NMR mapping”, where the <sup>1</sup>H-<sup>15</sup>N HSQC spectra of small peptides derived from the native sequence are compared with that of the full-length protein. NMR verification of the random coil nature of the peptides can be done in the classical way, and the good agreement between correlation peaks of the same residue in the isolated peptide or embedded in the full protein confirm the random coil structure of the amino acid even in the full protein.

**Key words** NMR spectroscopy – peptide – random coil – secondary structure – Tau protein.

## Introduction

In our previous contribution, we have paved the way for at least a partial NMR assignment of the neuronal Tau protein<sup>1</sup>. The random coil carbon chemical shifts as derived from small peptide models<sup>2</sup> were at the basis of our strategy, and the relative success confirms the previous description of Tau as a random polymer in solution<sup>3,4</sup>. However, a detailed interpretation of all chemical shifts including the amide nitrogen shift in terms of conformation is far from straightforward, as sequence effects equally play a role<sup>2,5,6</sup>. Because the nitrogen chemical shift is the only NMR parameter that leads to a reasonable dispersion in Tau, extracting maximal conformational information from it will be crucial to further analyse the conformation of Tau in solution.

The strategy we propose here is based on the comparison of the <sup>1</sup>H-<sup>15</sup>N HSQC of full-length Tau with the same spectra of short peptides derived from the Tau sequence. We hereby explicitly eliminate the local primary sequence as an influence on the chemical shift. Preliminary NMR analysis of the latter peptides allows to establish their random coil conformation, defined as a dynamical ensemble of conformations without any stable secondary structure. Agreement of the corresponding protein and peptide HSQC spectra indicates a

similar sampling of conformational space of the amino acids in the isolated peptide or embedded in the full protein. Disagreement, on the other hand, indicates some structural features that justify further investigations.

## Results

### *Peptide synthesis and NMR analysis*

Several peptides of 13-17 amino acids (Figure 1) were synthesized by solid phase synthesis using a classical Fmoc strategy. After cleavage from the resin, they were purified by reverse phase high-pressure liquid chromatography (HPLC), and their purity was further assessed by MALDI-TOF mass spectrometry. None of the sequences gave particular difficulties for synthesis or purification. The peptides were solubilised at a concentration of 4 mM in the same buffer as used for the protein spectra, and standard DQF-COSY, TOCSY and NOESY spectra were recorded for the assignment. The absence of secondary structure elements was verified by the measurement of coupling constants and analysis of the NOE pattern. Coupling constants were consistently found in the range of 6-7 Hz, and except for the sequential H<sub>α</sub>(i) – H<sub>N</sub> (i+1) contacts, no other NOE cross peaks of significant intensity were detected.

### ***Comparison of peptide and protein spectra***

Heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra were recorded overnight at natural abundance using a cryoprobe on a 600MHz instrument. The signal to noise ratio of the resulting spectra was satisfactory, partially due to the sharp line width of the peptide resonances. Spectra were overlaid with the HSQC spectrum of  $^{15}\text{N}$  labelled Tau441 or with the same spectrum of the foetal isoform Tau352. Verification of the assignments was performed on the basis of the triple resonance spectra that were described in our preceding work. In the following paragraphs, we will briefly comment the different peptide spectra.

### ***Peptide [Lys<sub>224</sub>-Lys<sub>240</sub>]***

This peptide is centered around Thr<sub>231</sub>, one of the critical proline directed phosphorylation sites of Tau<sup>7</sup>, and equally contains a second potential phosphorylation site Ser<sub>235</sub>. Phosphorylated in mitotic cells<sup>8</sup>, both sites together constitute the Alzheimer specific PHF-27 epitope<sup>9</sup>, whereas the first Thr-Pro motif is after phosphorylation equally recognized by the Pin1 prolyl *cis/trans* isomerase<sup>10</sup>. Because both (Val)Arg<sub>230</sub> and (Val)Ala<sub>227</sub> motifs are unique in the full-length Tau sequence, our previously described pairwise assignment strategy<sup>1</sup> has already allowed the unambiguous assignment of both

residues. The peptide resonances were found to coincide perfectly with both peaks (Figure 2a). The correlation peaks of the first two residues of the peptide, Lys<sub>224</sub> and Lys<sub>225</sub>, did not coincide with the equivalent signals in the protein, as verified on a HSQC spectrum of a Lys selectively labelled sample (Figure 2b), leading to the conclusion that border effects extend to at least two residues. The resonance of the Thr<sub>231</sub> as previously determined based on the Pro-induced random coil chemical shift variation<sup>11</sup> coincides with the peak found for the same residue in the peptide. Further verification of the other resonances in the crowded region of the full-length protein spectrum was done on the basis of strips extracted from the CBCANH and CBCA(CO)NH spectra on the doubly labelled Tau441 (Figure 3). Whereas assignment on the basis of those same 3D spectra turned out to be very complex if not impossible, verification of the peptide derived assignment is straightforward (Figure 3). The same 3D spectra allow the separation of the two Val residues at positions 226 and 228, degenerate in their amide proton frequencies (Figure 2a). The C-terminal part of the peptide following Pro<sub>236</sub> was characterized by resonance broadening of the amide signals, hindering the assignment of the two Ser residues.

### ***Peptides [Glu<sub>12</sub>-Lys<sub>24</sub>] and [Glu<sub>36</sub>-Leu<sub>48</sub>]***

At first, we noted slight shifts between several correlation peaks of the [Glu<sub>12</sub>-Lys<sub>24</sub>] peptide and the HSQC spectrum of Tau441. Because we tentatively assigned these shifts to the presence of a His-tag introduced in Tau441 for purification purposes, we compare in Figure 4 the peptide spectra with that of <sup>15</sup>N-labelled Tau352 sample without His-tag. Each peptide contains one (Ala)Gly motif among the four present in full-length Tau. Whereas previously we could not further distinguish them, the superposition of the peptide HSQC spectra allows to assign the peaks at [8.45, 107.5] ppm and [8.44, 108.0] ppm to Gly<sub>42</sub> and Gly<sub>16</sub>, respectively. The two correlation peaks corresponding to the remaining (Ala)Gly<sub>107</sub> and (Ala)Gly<sub>120</sub> residues are found at [8.44, 108.4] ppm and at [8.22, 107.1] ppm, but without complementary information, we cannot yet distinguish them. Interestingly, Thr<sub>17</sub> is close to a (Gly)Thr peak in the Tau352 spectrum, but we note a slight shift in both proton and nitrogen dimensions. At this moment, we cannot distinguish subtle differences in pH or differential sampling of the conformational space by this residue as the origin of the imperfect match.

### ***Peptide [Ser<sub>208</sub>-Arg<sub>221</sub>]***

This peptide is centered around the Thr<sub>212</sub> residue, that together with Ser<sub>214</sub> upon phosphorylation of both generates the AT10 binding site<sup>9</sup>. Although the [Ser<sub>208</sub>-Arg<sub>221</sub>] and the [Lys<sub>224</sub>-Lys<sub>240</sub>] peptides both belong to the proline rich region of Tau, we detected a significant amount of *cis* conformation only for the Thr<sub>212</sub>-Pro<sub>213</sub> (6%) and Thr<sub>217</sub>-Pro<sub>218</sub> (12%) peptide bonds<sup>12</sup>. The peptide when phosphorylated on Thr<sub>212</sub> and/or Ser<sub>214</sub> not only interacts with the Pin1 WW domain, but is equally a substrate for the catalytic domain of Pin1<sup>12</sup>. Overlap of the two spectra is again very good, and for example the Thr<sub>217</sub> correlation peak, previously assigned without ambiguity on the basis of its unique (Pro)Thr(Pro) motif<sup>1</sup>, exactly coincides with the same resonance in the peptide. When we compare the Ser/Thr regions of the peptide and Tau352 HSQC spectra, we note next to the well-overlapping intense peaks in both spectra, the presence of some minor forms (Figure 5). Interestingly, the peptide resonance that corresponds to the Thr<sub>217</sub> minor *cis* form (assigned on the basis of its H<sub>N</sub> frequency), overlaps well with a minor peak in the Tau352 HSQC spectrum. Whereas its intensity was too low to allow verification on the basis of the corresponding triple resonance spectra, we are reasonably confident that the assignment to the Thr<sub>217</sub> is correct. The other weak correlation peaks, assigned on the

basis of the peptide homonuclear spectra to the minor forms of Thr<sub>212</sub> and Ser<sub>214</sub> surrounding Pro<sub>213</sub> in its *cis* form, are equally present in the Tau352 spectrum but at a lower level (data not shown). This validates the peptide model not only at the level of its major conformation, but equally shows that the energetic barriers separating *trans* and *cis* conformers are comparable in the peptide and the full-length protein.

#### ***Peptide [Gly<sub>303</sub>-Lys<sub>317</sub>]***

This peptide contains the PHF6 motif previously defined to be the nucleus of the aggregation because of its strong interaction with the proteolytic PHF43 fragment that itself did form fibers in solution<sup>13</sup>. The synthesis of the peptide had not given particular problems, but not unexpectedly, we did not succeed to solubilize it in our standard buffer to the 4mM concentration that we had used for all other peptides. Based on the signal intensity of the NMR signals in the supernatant, we estimate that only 200μM peptide remains in solution, whereas the rest readily precipitates. This concentration, though not sufficient for a <sup>1</sup>H-<sup>15</sup>N HSQC at natural abundance, did allow the recording of the classical homonuclear spectra. Based on the TOCSY and NOESY spectra, we could assign the proton frequencies of the amide protons, and in Figure 6, we show the

assignment of the V<sub>309</sub> and Y<sub>310</sub> amide protons on top of the HSQC spectrum of <sup>15</sup>N labelled Tau441. Our previous assignment strategy with pairwise motif search did not succeed to unambiguously assign both residues, as exactly the same motif is found later in the sequence (I<sub>393</sub>VYK<sub>396</sub>). However, for both residues, two candidate correlation peaks had been identified in the HSQC spectrum (Figure 6). The H<sub>N</sub> chemical shift values of Val<sub>309</sub> and Tyr<sub>310</sub> in the peptide spectrum indicate that the peaks at [8.24, 126.1] ppm and [8.51, 126.4] ppm correspond to the two residues.

#### ***Peptides [Val<sub>393</sub>-Leu<sub>408</sub>] and [Ser<sub>416</sub>-Leu<sub>428</sub>]***

The [Val<sub>393</sub>-Leu<sub>408</sub>] peptide was initially chosen because it encompasses two additional phosphorylation sites, Ser<sub>396</sub> and Ser<sub>404</sub>, that both can be phosphorylated in mitotic cells.<sup>14</sup> Solubility of this peptide was not a problem, and we could record a natural abundance HSQC spectrum. When superimposed upon the Tau441 HSQC (Figure 7), the two N-terminal residues (Val<sub>393</sub> and Tyr<sub>394</sub>) expectedly do not coincide with the corresponding peaks in the full-length protein, but this discrepancy extends to Lys<sub>395</sub> (that was unambiguously assigned in the full-length Tau spectrum as the sole (Tyr)Lys motif not followed by a Pro) and even Ser<sub>396</sub>. For the latter resonance, the chemical shift change was however more limited

(Figure 7). After the Pro in position 397, however, the peptide resonances again overlap correctly with the corresponding protein resonances, and a good agreement was equally found for the [Ser<sub>416</sub>-Leu<sub>428</sub>] peptide (data not shown).

## Discussion

Because the absence of stable secondary structure can be readily established in peptide models based on average J coupling values and the absence of long range NOE effects, they have been successfully used to define the so-called “random coil” chemical shift values for the different amino acids. Proposed first for the H $\alpha$  and aliphatic protons<sup>15</sup> and later on for the C $\alpha$ , C $\beta$  and carbonyl carbons<sup>2</sup>, it has been recognized that the direct environment of the residue, in the form of its next neighbours, equally determines the chemical shift value<sup>5</sup>. A different strategy of establishing random coil chemical shifts starts from a statistical analysis of the combined chemical shift assignments and structural data as deposited in the Biomagresbank and Protein Data Base<sup>6,16</sup>. The advantage of this latter approach is that the large body of data should contain at least some examples of every possible di- or tripeptide, allowing the deconvolution of secondary structure effects from the electron withdrawing effects caused by the neighbouring residue side chain.

The definition of “random coil”, however, in these latter studies potentially deviates from the notion of “ensemble of conformations” that characterizes small peptides, because it is defined as “not in the  $\alpha$ -helix or  $\beta$ -sheet zones of the Ramachandran chart”. Attempts to define a “random coil” conformation in a natively unfolded protein such as the neuronal protein Tau are even more challenging. Even though NMR spectroscopy remains the method of choice, we face methodological problems – how to extract conformational parameters from the overcrowded spectra? – but equally a conceptual problem. Indeed, both definitions as an ensemble of conformations (similar as in a peptide), or as the absence of well-defined  $\alpha$ -helix or  $\beta$ -sheet conformation can be adopted, but are not easily distinguishable in the context of a full-length protein.

In order to solve simultaneously both problems, we have here adopted a strategy of “peptide NMR mapping”. It consists of the synthesis and spectral recording of relatively small peptides, that can be easily analysed by homonuclear NMR spectroscopy, leading not only to their full resonance assignment but equally to their easy classification of “random coil” peptides. Peptides should be sufficiently small allowing for their straightforward chemical synthesis, purification, and NMR analysis. An additional reason for a limited length is that they should not adopt a stable structure. They

should however not be too short, because we expect some spectral border effects for those residues that are at the N- or C-terminus of the peptide, and hence do not have the same chemical environment as their equivalents in the full length protein. Because we previously had synthesized a number of peptides of 13 to 17 amino acids containing potential phosphorylation sites<sup>12</sup>, we first used these and extended them with some additional peptides from the N-terminus or third microtubule binding repeat (Figure 1). NMR analysis of all those peptides led to the conclusion of absence of stable secondary structure elements in all of them. The second stage of the procedure consists in the recording of <sup>1</sup>H-<sup>15</sup>N HSQC spectra on these peptides in the same conditions of buffer, pH and temperature as for the full length protein. Present cryogenic probe heads<sup>17,18</sup> installed on a 600MHz spectrometer enable such spectral recordings within one night, if the peptide concentration exceeds 4mM. The graphical superposition of the peptide spectra with the full-length Tau HSQC completes the procedure.

Central to our approach is the statement that the same amino acid stretch in an isolated peptide or in a full protein sequence should give an identical HSQC spectrum when the individual amino acids sample the same conformational space in both contexts. That this will be actually the case, of course, is not obvious at all, as

especially the nitrogen chemical shift is extremely sensitive to the spatial environment of the amide group. Evidence for this is clear from the Tau HSQC spectrum: despite the macroscopically unfolded nature of the protein<sup>3,4</sup> and the unfavourable amino acid composition, reflecting in a mere 0.8 ppm chemical shift dispersion for the amide protons, the nitrogen shifts span the full range from 105 to 130 ppm. Strikingly, empirical prediction methods using the peptide or database derived chemical shift parameters still have the most problems with nitrogen chemical shifts as well<sup>6,16</sup>. The consequence of this extreme sensitivity is that in the case of a good agreement, we can safely say that the amino acids indeed sample conformational space in a very similar way, be they embedded in a short peptide or in the full protein. Disagreements, on the other hand, indicate structural differences between the isolated peptide or its equivalent in the full-length protein, and therefore point to regions in the protein that should be examined more in detail by other NMR methods.

Because the mere superposition of peaks is not sufficient to conclude the identical resonance assignment, we have used the same triple resonance spectra exploited in our previous work<sup>1,11</sup> to verify rather than determine the assignment. The superposition of the [Lys<sub>224</sub>-Lys<sub>240</sub>] peptide and Tau441 spectra (Figure 2a) indeed shows a good

agreement with the obviously good agreement for the isolated Ala<sub>227</sub> and Arg<sub>230</sub> correlation peaks. The strips extracted for the Thr<sub>231</sub>-Ala<sub>227</sub> stretch at the exact positions of the peptide resonances (Figure 3), moreover show that the sequential walk indeed is possible, and thereby confirm that the superposition is valid. However, the CBCA(CO)NH strip extracted for Thr<sub>231</sub> shows the difficulty that we would face if we had to use the same information for the ab initio assignment of full length Tau. We indeed observe three C $\alpha$  and two C $\beta$  peaks, but have no indication how to combine them into meaningful pairs. This situation is even worse for the strip of Val<sub>228</sub>, whose <sup>1</sup>H-<sup>15</sup>N correlation peak is in the crowded central region of the HSQC spectrum.

In order to ascertain the border effects, and hence the fraction of every peptide that is useful to assist the assignment, we compared the same peptide spectrum with that of a <sup>15</sup>N Lys labelled sample. It can be seen that both the peptide N- and C-terminal Lys correlations (Lys<sub>224</sub> and Lys<sub>240</sub>) do not overlap with any peak in the Lys labelled Tau441, but we equally did not manage to connect the Lys<sub>225</sub> correlation with the corresponding Val<sub>226</sub>, indicating that border effects span at least two residues. Independent confirmation of this comes from the aforementioned Gly<sub>42</sub> or Val<sub>309</sub>/Val<sub>393</sub> and Tyr<sub>310</sub>/Tyr<sub>394</sub> residues. Gly<sub>42</sub> within the peptide derived from the

Tau441 sequence equally well agrees with the Tau352 spectrum, although they differ starting from the residue in position 45 (Glu for Tau441, Ala for Tau352). The two Val/Tyr motifs share in the full-length protein the same neighbours, but they do differ in the residues at the ( $i \pm 2$ ) position, leading to slight differences in their amide chemical shift values in the protein spectra. In conclusion, these limited border effects mean that every peptide contributes to some ten assignments. These same assignments, of course, can equally lift some of the ambiguities that we had previously reported<sup>1</sup>. An example is the (Gly)Thr motif, with two candidate peaks at [8.03, 113.7] and [8.27, 113.0] ppm for two possible residues at positions 17 or 135. The peptide assisted assignment of Thr<sub>17</sub> to the former peak immediately removes the ambiguity for both. Those residues for whom the peptide NMR mapping was helpful in the assignment were underlined in Figure 2 of the preceding paper<sup>1</sup>.

The peptide derived from the beginning of the third microtubule binding repeat contains the local sequence motif (Val<sub>306</sub>-Lys<sub>311</sub>) that was previously shown to form  $\beta$  structure upon aggregation<sup>13</sup>. In agreement with this, the peptide precipitated at the millimolar concentrations that we used for the natural abundance HSQC recording. This clearly shows that the full protein context plays a solubilizing role for this element. Preliminary data in our laboratory

show that poly-anions such as heparin, known to lead to the formation of Paired Helical Fragments of Tau<sup>19</sup>, interfere with this role and allow for interaction of this fragment even in the full protein context. Long range NOE effects in the homonuclear NOE spectrum were not observed for this peptide, and the agreement of the amide proton chemical shifts with those of the cross peaks that we had tentatively assigned to Val<sub>309</sub> and Tyr<sub>310</sub> indicates that even in the full-length Tau protein free in solution, this PHF6 fragment still does not adopt a stable secondary structure. However, because of the poor quality of the NOE spectrum, we cannot exclude that the central residues [Ile308-Lys311] do form a hydrophobic cluster through their side chains, and this as well in the protein as in the peptide. Evidence for this comes from the analysis of the [Val<sub>393</sub>-Leu<sub>408</sub>] peptide, where we did observe discrepancies beyond the first two N-terminal residues. Lack of the N-terminal Ile residue in this peptide might break the cluster, and this structural difference might extend further on and explain the spectral discrepancies that we observe (Figure 7).

Of particular interest in the framework of the proposed<sup>10</sup> interaction between Tau and Pin1 is the [Ser<sub>208</sub>-Arg<sub>221</sub>] peptide. We showed that this peptide in its phosphorylated form can interact with the Pin1 WW domain, and that its two proline directed phosphorylation

motifs (Thr<sub>212</sub>-Pro and Thr<sub>217</sub>-Pro) in the small peptide are present to some extent in the *cis* conformation. NMR spectroscopy led to a quantitative measure of 6 respectively 12% of *cis* form for both peptide bonds, but equally showed a Pin1 catalytic specificity for the isomerization of the first peptide bond<sup>12</sup>. Because of the excellent signal to noise in our peptide HSQC spectra, we were able to observe the amide correlations of those minor forms as well. Interestingly, they did not only coincide with the location of correlation peaks in the Tau352 HSQC spectrum (where the latter was chosen because it was less crowded in this region), but the relative intensities indicate that even the populations of *trans* and *cis* isomers are conserved. Our current efforts to phosphorylate Tau in vitro should allow us to extend this observation to the phosphorylated protein, as <sup>1</sup>H-<sup>15</sup>N HSQC spectra of the phosphorylated peptides can readily be acquired and assigned.

In conclusion, we propose here a method dubbed “peptide NMR mapping” that should be of general use to study natively unfolded proteins. It consists of the synthesis and NMR analysis of a series of short peptides, and the subsequent superposition of their <sup>1</sup>H-<sup>15</sup>N HSQC spectrum with that of the full-length protein. Triple resonance spectra on the doubly labelled protein can be used to verify the

assignment, where verification is substantially easier than ab initio assignment on the basis of the same spectra. Besides the methodological advancement that our method allows for in the assignment procedure, we directly obtain important structural data. Indeed, because of the extreme sensitivity of nitrogen chemical shift to the 3D environment, we conclude that in the case of good agreement, the same amino acids in the isolated peptide or embedded in the full protein sample conformational space in a very similar way. In our case of the Tau protein, we have found in general a very good agreement.

## **Experimental Section**

### ***Expression and purification of the recombinant Tau protein.***

Expression and purification of the uniformly labelled samples was as previously described.<sup>1, 7</sup> Purification of the foetal construct Tau352 without His-tag was performed by a first heating step followed by cationic exchange chromatography with a Mono-S column (Amersham Pharmacia).

### ***Peptide synthesis and purification***

Peptides were synthesized using a classical Fmoc strategy on a Pioneer<sup>TM</sup> peptide continuous flow synthesizer with intermediate capping after each coupling step. Cleavage of peptide from the resin and side chain protecting groups were performed using TFA mixtures containing either triisopropylsilane 2.5% or triisopropylsilane 1% : ethanedithiol 2.5% for peptide containing Cys or Met residues. Crude peptides were then purified by RP-HPLC on a C18 Nucleosil column performing a linear gradient of acetonitrile. Fractions were checked both by RP-HPLC and MALDI-TOF mass spectrometry. NMR samples of peptides were prepared at a concentration of 4 mM in the same buffer that was used for the recombinant Tau protein.

### ***NMR Spectroscopy.***

All NMR spectra were recorded at 20°C on a 600MHz Bruker DMX spectrometer (Bruker, Karlsruhe, Germany) equipped with a cryogenic triple resonance probe head with actively shielded z-gradient. Triple resonance spectra on full-length Tau441 or Tau352 were performed as described previously.<sup>7</sup> The homonuclear peptide spectra for peptide assignment and structural verification were recorded with DQFCOSY or Water flipback NOESY and TOCSY sequences, with mixing times of 500ms and 69ms for the latter. <sup>1</sup>H-<sup>15</sup>N HSQC were recorded with 256 scans per increment, and 256 nitrogen increments.

### ***Acknowledgments***

We thank E. Diesis and A. Menet for help with the peptide synthesis. Dr. L. Buée (Lille, France) gave us many valuable insights about Tau and Alzheimer's disease. The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, and the Institut Pasteur de Lille. A.S. is funded by a European Training and Mobility Grant (HPRN-CT-2002-00241).

### **References**

1. G. Lippens, A. Leroy, A. Sillen, J.-M. Wieruszeski, C. Smet, and I. Landrieu, *ChemBioChem*, submitted.
2. D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR*, **1995**, 5, 67-81.
3. D. W. Cleveland, S. Y. Hwo, M. W. Kirschner, *J Mol Biol.* **1977**, 116, 227-247
4. O. Schweers, E. Schonbrunn-Hanebeck, A. Marx, and E. Mandelkow, *J. Biol. Chem.* **1994**, 269, 24290-24297.
5. S. Schwarzingler, G.J. Kroon, T.R. Foss, J. Chung, P.E. Wright, H. J. Dyson, *J. Am. Chem. Soc.* **2001**, 123, 2970-2978.
6. Y. Wang and O. Jardetzky, *J Am Chem Soc.* **2002**, 124, 14075-14084.
7. M. Morishima-Kawashima, M. Hasegawa, K. Takio, M. Suzuki, H. Yoshida, K. Titani and Y. Ihara, *J Biol Chem.* **1995**, 270, 823-829.
8. I. Vincent, M. Rosado and P. Davies, *J Cell Biol.* **1996**, 132, 413-425.
9. R. Hoffmann, V.M. Lee, S. Leight, I. Varga, and L. Otvos, *Biochemistry* **1997**, 36, 8114-8124.

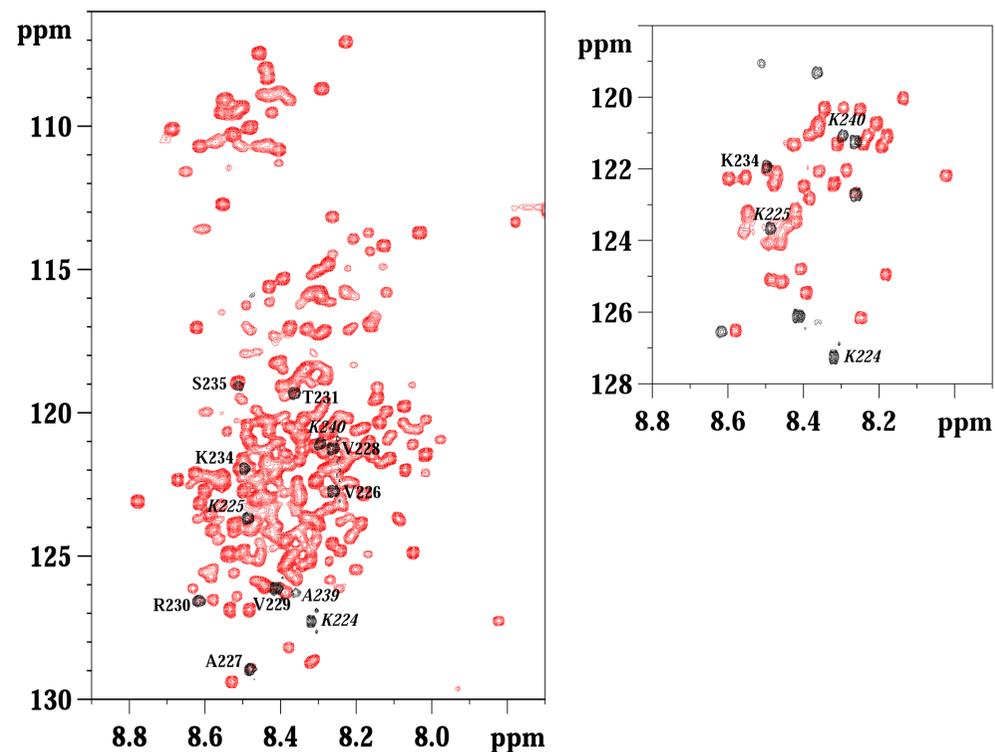
10. P.-J. Lu, G. Wulf, X Z. Zhou, P. Davies, K. P. Lu, K. P. *Nature* **1999**, 399, 784-788.
11. G. Lippens, J.-M. Wieruszeski, A. Leroy, C. Smet, A. Sillen, L. Buée and I. Landrieu, *ChemBioChem* **2004**, 5, 73-78.
12. C. Smet, A.-V. Sambo, J.-M. Wieruszeski, A. Leroy, I. Landrieu, L. Buée, G. Lippens, *Biochemistry*. **2004**, 43, 2032-2040.
13. M. von Bergen, P. Friedhoff, J. Biernat, J. Heberle, E.-M. Mandelkow and E. Mandelkow, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 5129-5134.
14. U. Preuss and E.M. Mandelkow, *Eur J Cell Biol.* **1998**, 76, 176-184.
15. A. Bundi and K. Wüthrich, *Biopolymers* **1979**, 18, 285-298.
16. J. Meiler, *J Biomol NMR.* **2003**, 26, 25-37.
17. P. Styles, N.F. Soffe, C.A. Scott, D.A. Cragg, F. Row, D.J. White and P.C.J. White *J. Magn. Res.* **1984**, 60, 397-404.
18. P. Styles, N. Soffe and C. A. Scott, *J. Magn. Res.* **1989**, 84, 376-378.
19. M. Goedert, R. Jakes, M.G. Spillantini, M. Hasegawa, M.J. Smith, and R.A. Crowther *Nature* **1996**, 383, 550-553.

## Figures

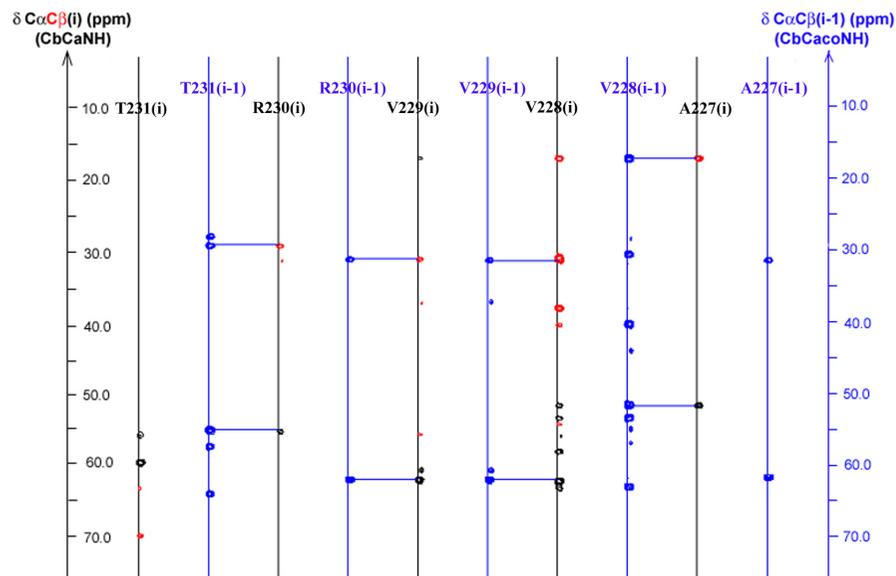
```

MAEPRQEFEV MEDHAGTYGL GDRKQDGGYT MHQDQEGDTD 40
AGLKESPLQT PTEGSEEPG SETSDAKSTP TAEDVTAPLV 80
DEGAPGKQAA AQPHTIPEG TTAEEAGIGD TPSLEDEAAG 120
HVTQARMVSK SKDGTGSDDK KAKGADGKTK IATPRGAAPP 160
GQKGQANATR IPAKTTPAPK TTPSSGEPK SGDRSGYSSP 200
GSPGTPGSR SRTPLSLTPPT REPKKVAVVR TPKSPSSAK 240
SRLQTAPVPM PDLKNVSKI GSTENLKHQP GGGKVQIINK 280
KLDLSNVQSK CGSKDNIKHV PGGGSVQIVY KPVDLSKVT S 320
KCGSLGNIHH KPGGGQVEVK SEKLDKDRV QSKIGSLDNI 360
THVPGGNKK IETHKLFRE NAKAKTDHGA EIVYKSEVVS 400
GDTSPRHLSN VSSTGSIDMV DSPQLATLAD EVSASLAKQG 440
L 441
  
```

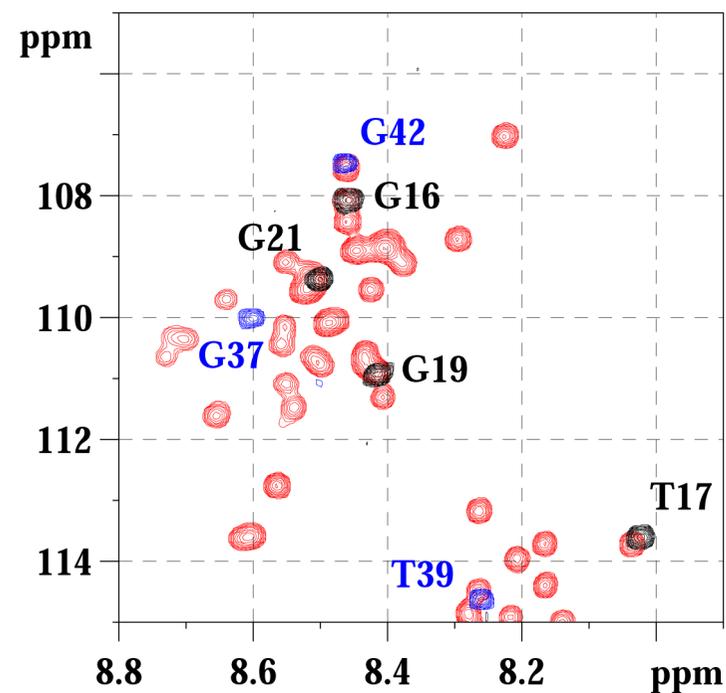
**Figure 1** Primary sequence of the longest isoform of human Tau441, indicating the inserts that distinguish it from the foetal form (box). In yellow are indicated the different peptides used for this study.



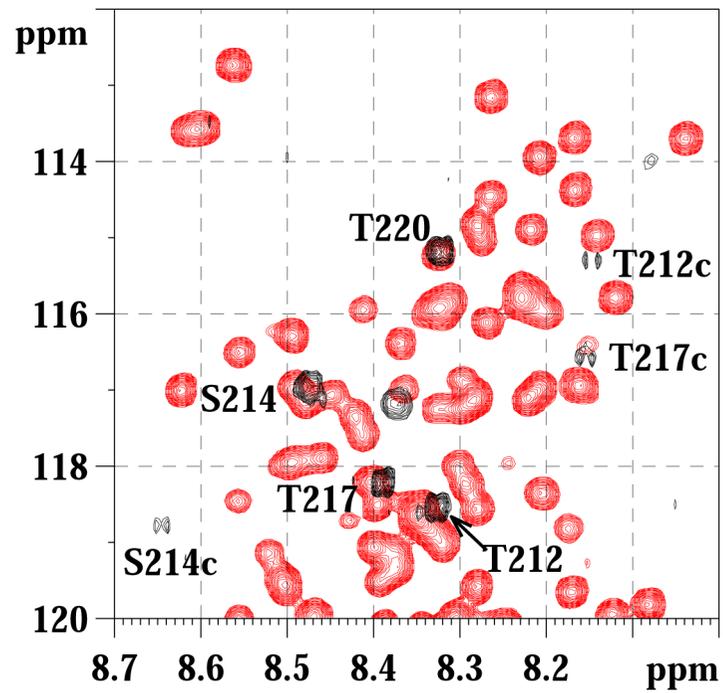
**Figure 2** Superposition of the uniformly (left) or  $^{15}\text{N}$ -Lys labelled (right) Tau441 HSQC spectrum (red) with the same spectrum on the [Lys<sub>224</sub>-Lys<sub>240</sub>] peptide (black).



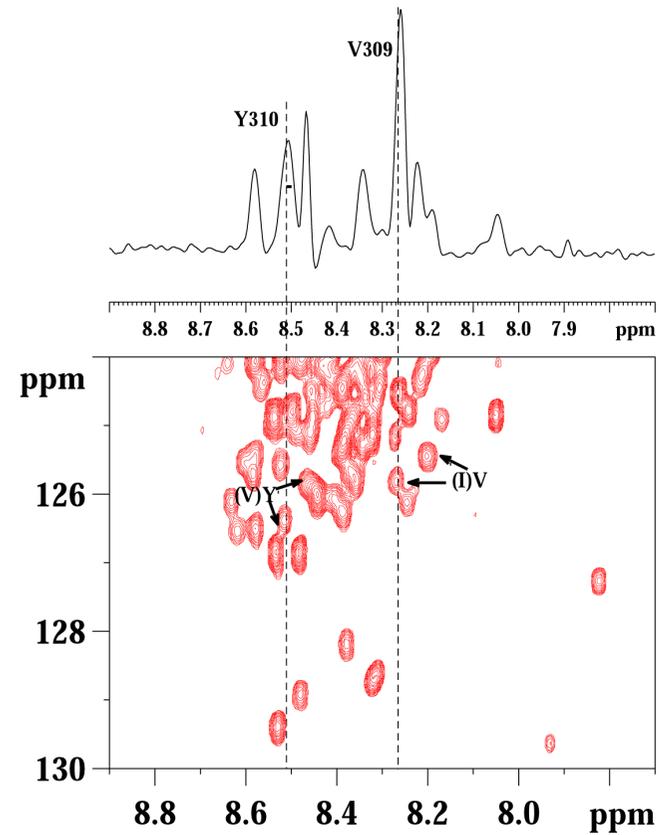
**Figure 3** Strips extracted from the CBCANH (black for positive, red for negative contours) and CBCA(CO)NH (blue) spectrum of  $^{15}\text{N}/^{13}\text{C}$  labelled Tau441 confirming the assignment of the [Thr<sub>231</sub>-Ala<sub>227</sub>] stretch.



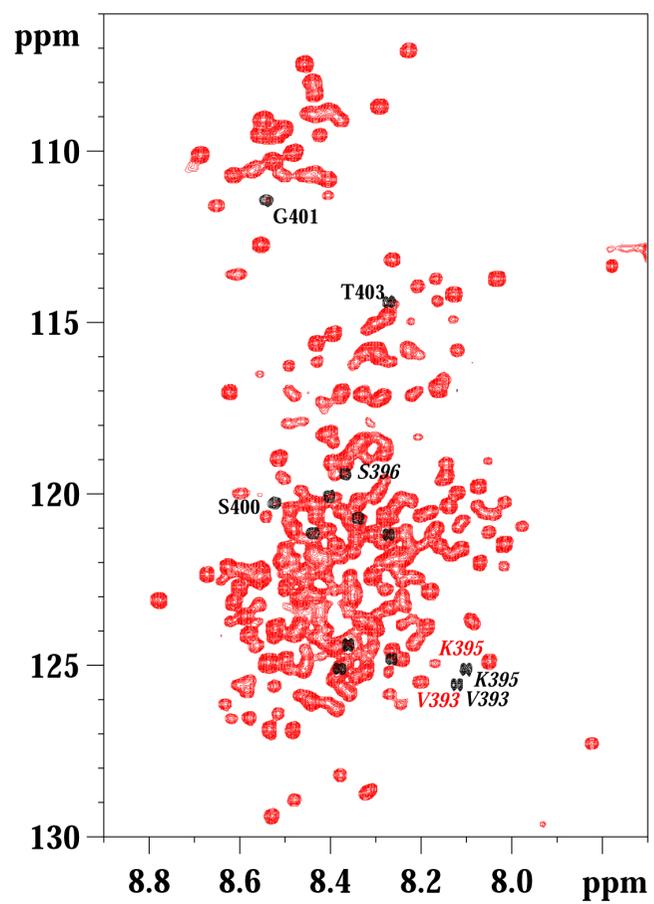
**Figure 4** Overlay of the uniformly  $^{15}\text{N}$  labelled Tau352 spectrum (red) with the HSQC spectra of [Glu<sub>12</sub>-Lys<sub>24</sub>] (black) and [Glu<sub>36</sub>-Leu<sub>48</sub>] (blue) peptides.



**Figure 5** Superposition of the uniformly  $^{15}\text{N}$  labelled Tau352 spectrum (red) with the HSQC spectra of [Ser<sub>208</sub>-Arg<sub>221</sub>] (black). Resonances corresponding to the minor *cis* forms are indicated.



**Figure 6** Proton 1D spectrum of the [Gly<sub>303</sub>-Lys<sub>317</sub>] peptide (top) and corresponding zone of the Tau441 HSQC (bottom). Candidate peaks for both (Ile)Val and (Val)Tyr as previously assigned<sup>1</sup> are indicated.



**Figure 7** Overlay of the Tau441 and [Val<sub>393</sub>-Leu<sub>408</sub>] HSQC spectra. Residues for which discrepancies were observed between both spectra are indicated in italic.

## **CHAPITRE IV**

**Conception, synthèse et criblage d'inhibiteurs ciblant la liaison de Pin1 aux motifs Ser/Thr-Pro phosphorylés.**



## **I – Des inhibiteurs de Pin1 comme outil pour évaluer le rôle des interactions Pin1-tau dans la maladie d’Alzheimer.**

### **1.1 – Points de départ d’une stratégie visant à inhiber les interactions entre Pin1 et les motifs Ser/Thr-Pro phosphorylés.**

#### **1.1.1 – Cibler le domaine WW responsable de la liaison du substrat.**

Deux stratégies sont envisageables pour inhiber les fonctions de Pin1 : cibler le domaine catalytique, directement responsable de la fonction peptidyl-prolyl *cis/trans* isomérase, ou le domaine WW, impliqué dans la liaison du substrat et dont l’état de phosphorylation (en Ser16) régule l’activité et la localisation cellulaire de Pin1 dans les cellules HeLa (Lu, Zhou et al. 2002). Le fait que le domaine WW soit tout aussi essentiel que le domaine catalytique pour assurer la fonction biologique de Pin1 (Lu, Zhou et al. 1999) et son rôle dans la régulation de l’activité (Lu, Zhou et al. 2002) ont motivé le choix du domaine WW comme cible dans la recherche de molécules inhibitrices de la fonction de Pin1.

L’activité du domaine WW de Pin1 serait régulée par l’état de phosphorylation de la Ser16, résidu localisé dans la boucle de liaison du phosphate du domaine WW, qui diminuerait sa capacité à lier les substrats *in vitro* et influencerait sur la localisation cellulaire de Pin1. En effet, la phosphorylation de Pin1 *in vitro* par la kinase PKA inhibe totalement les interactions entre Pin1 et divers antigènes de l’anticorps MPM-2 (Lu, Zhou et al. 2002). D’autre part, dans les cellules HeLa surexprimant Pin1 ou ses domaines isolés, Pin1 et le domaine WW sont localisés très clairement au niveau nucléaire dans la cellule alors que le domaine catalytique isolé se distribue de manière plus diffuse au niveau nucléaire comme au niveau cytoplasmique. Ces données indiquent un rôle déterminant du domaine WW dans la localisation cellulaire de Pin1 (Rippmann, Hobbie et al. 2000; Lu, Zhou et al. 2002). Ainsi la phosphorylation de la Ser16, en inhibant la liaison du domaine WW à ses substrats, entraîne une localisation diffuse de Pin1 au niveau cellulaire, c’est-à-dire une distribution dans tous les compartiments cellulaires telle qu’elle est décrite pour le domaine catalytique (Lu, Zhou et al. 2002). Par conséquent, la liaison du domaine WW à ses substrats influe sur l’adressage de Pin1 au niveau cellulaire. Du point de vue biologique, la forme majoritairement retrouvée dans les tissus sains est la forme phosphorylée inactive du domaine WW. La forme déphosphorylée est détectée principalement dans des cellules cancéreuses du sein

surexprimant Pin1 (Wulf, Ryo et al. 2001) ce qui signifie que l’inhibition du domaine WW par de petites molécules mimerait la forme « normale » de la protéine Pin1.

Une stratégie plus classique visant à inhiber la fonction enzymatique de Pin1 a été entreprise avec des molécules mimant la juglone, un inhibiteur des parvulines, dont le mode d’action repose sur la liaison covalente d’une molécule de juglone à la cystéine du site actif de Pin1 conduisant à une dénaturation partielle et une perte d’activité (Hennig, Christner et al. 1998; Chao, Greenleaf et al. 2001; Uchida, Takamiya et al. 2003). Il faut souligner que la RMN permet de détecter directement l’activité isomérase, mais un test des inhibiteurs potentiels basé sur la perte d’activité catalytique serait beaucoup plus délicat à mettre en œuvre, notamment dans un contexte de *screening*, compte-tenu des contraintes exposées dans les chapitres précédents. Un test de *binding* basé sur la capacité de liaison des composés au domaine WW est plus simple et plus robuste.

Travailler avec le domaine WW isolé plutôt qu’avec la protéine Pin1 entière présente également certains avantages pratiques car l’expression par voie recombinante donne de meilleurs rendements pour le domaine WW isolé et celui-ci présente une plus grande stabilité alors que Pin1 a tendance à s’agréger en solution par oxydation de la Cys113, résidu essentiel à l’activité catalytique mais qui favorise l’agrégation de la protéine dans des conditions non réductrices (Bayer, Goettsch et al. 2003). La stabilité du domaine WW autorise un conditionnement sous forme de lyophilisat et un stockage plus long. D’autre part, les signaux RMN du domaine WW ont une largeur de raie fine du fait de sa petite taille ce qui induit une meilleure qualité de spectres, favorable au criblage des inhibiteurs par RMN.

### **1.1.2 – Le dipeptide phosphoThr-Pro comme modèle pour la conception de ligands du domaine WW**

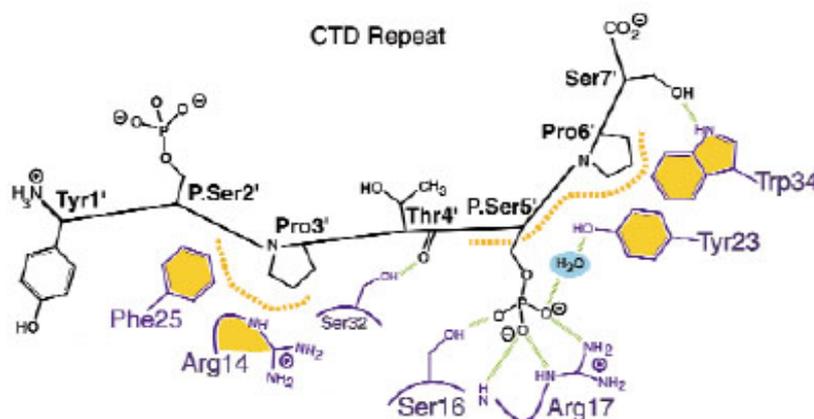
Les structures de Pin1 et de son domaine WW en complexe avec un phospho-peptide montrent une spécificité de l’interaction au niveau du motif phosphoThr/Ser-Pro dans les séquences peptidiques. Notamment, la structure cristallographique de Pin1 en complexe avec un peptide di-phosphorylé issu du domaine C-terminal de la ARN polymérase II met en évidence plusieurs types d’interactions entre le peptide et le domaine WW de Pin1 (Verdecia, Bowman et al. 2000) :

1. Des liaisons hydrogènes entre le groupement phosphate de la phosphosérine et les fonctions alcool et guanidinium des chaînes latérales des résidus Ser16 et Arg17

respectivement ainsi qu’avec le H<sub>N</sub> de l’Arg17 et la fonction phénol de la Tyr23 via une molécule d’eau (Figure 4.1).

2. Des liaisons hydrophobes entre le cycle pyrrolidine de la proline et les noyaux aromatiques phénol et indole des chaînes latérales des résidus Tyr23 et Trp34 (Figure 4.1).

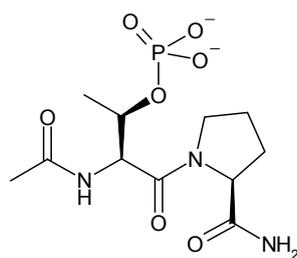
La présence d’une liaison hydrogène entre la fonction alcool de la sérine suivant la proline dans la séquence du peptide et l’hydrogène du noyau indole du résidu Trp34 apporte un point d’interaction supplémentaire moins essentiel. Les études de liaison de substrats peptidiques par le domaine WW de Pin1 réalisées par RMN homonucléaire ont montré un mode de liaison unique du substrat et la *trans*-spécificité du domaine WW (Wintjens, Wieruszeski et al. 2001).



**Figure 4.1** : Représentation des interactions entre le domaine WW et le peptide CTD de l’ARN Polymérase II indiquant les éléments structuraux intervenant dans l’interaction (d’après Verdecia, *Nat Struct Biol* 2000).

Les affinités de différents peptides de la protéine tau déterminées en RMN varient d’un peptide à l’autre mais restent toujours dans une gamme allant de 100µM à 1mM pour des peptides mono-phosphorylés, les affinités les plus faibles étant obtenues pour des peptides comportant le motif phosphoSer-Pro par rapport au motif phosphoThr-Pro. L’affinité du dipeptide phosphoThr-Pro montre que ce motif est le point central de l’interaction avec le domaine de liaison de Pin1. Cependant, il semble que les résidus environnants jouent un rôle non négligeable dans la liaison. En effet, pour le peptide centré autour du motif phosphoThr212-Pro213 ( $K_D=100 \mu M$ ), on retrouve le motif pTPS identique à celui du peptide du domaine C-terminal de l’ARN polymérase II. En revanche, le peptide centré autour du motif phosphoThr231-Pro232 ( $K_D=380 \mu M$ ) ne possède pas sérine en +1 du motif phosphoThr-Pro mais une proline, ce qui expliquerait l’affinité plus faible observée.

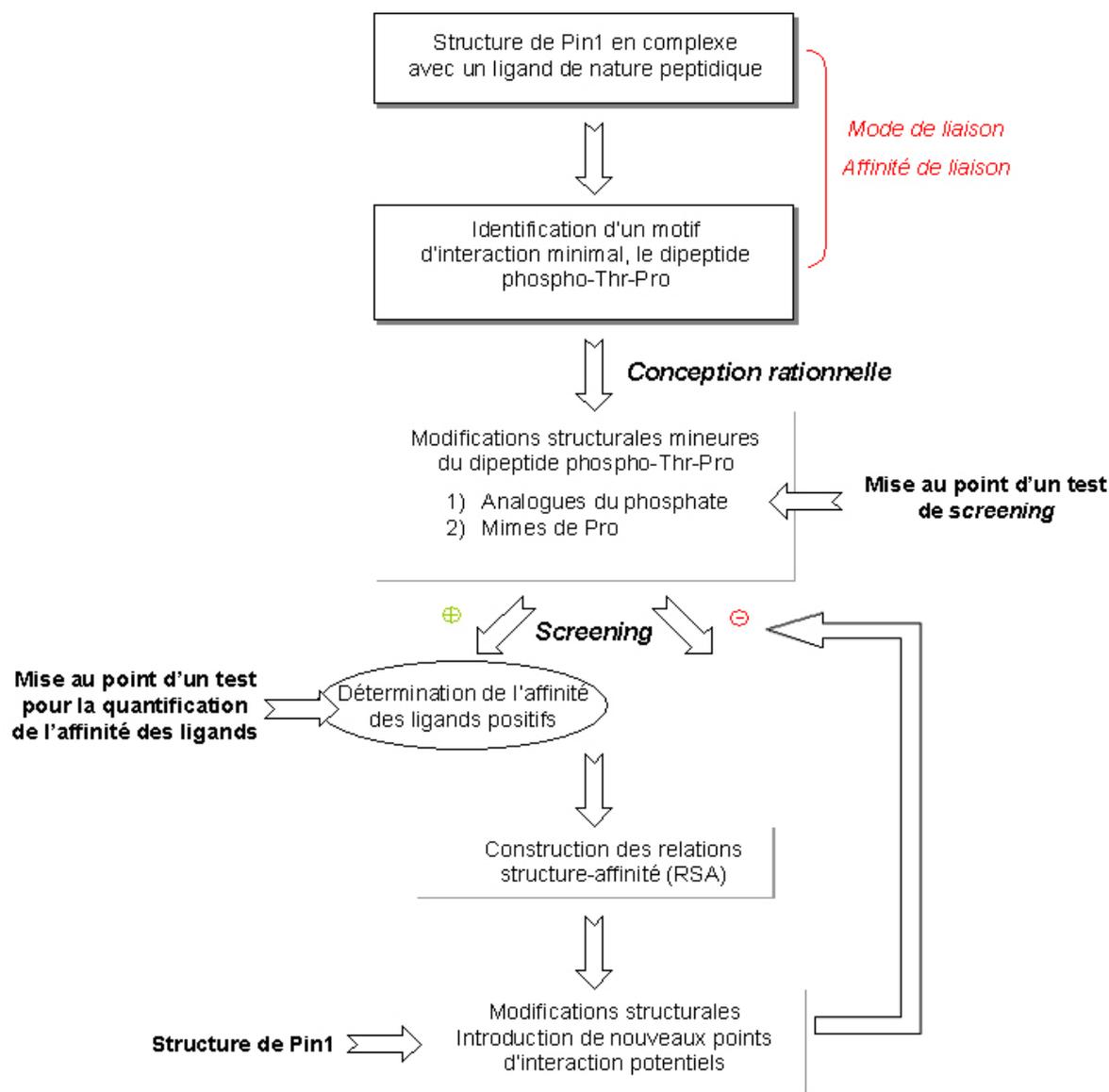
Le dipeptide phosphoThr-Pro (**1**) a été choisi comme point de départ pour la conception de ligands du domaine WW. La métabolisation importante des groupements phosphate *in vivo* et leur impact négatif sur la biodisponibilité des composés impliquent leur remplacement dans la structure des composés synthétisés, de même pour la liaison peptidique. Dans un premier temps, ces éléments seront conservés afin de minimiser les changements structuraux et pouvoir évaluer leur importance dans la liaison au domaine WW de Pin1. Le remplacement du groupement phosphate sera également envisagé dans un deuxième temps.



(1)

### 1.1.3 – Stratégie employée dans la recherche d’inhibiteurs de Pin1.

La stratégie employée pour la conception de ligands du domaine WW est basée sur une stratégie classique de conception rationnelle basée sur la structure de la cible définie, qui est ici le domaine de liaison WW de la peptidyl-prolyl *cis/trans* isomérase humaine Pin1, et sur la connaissance d’un motif primaire permettant une liaison au domaine WW avec une faible affinité, qui est dans notre cas le dipeptide phosphoThr-Pro (**1**) (Schéma 4.1). La stratégie employée met en œuvre la synthèse de mimes du dipeptide phosphoThr-Pro, le criblage et l’évaluation de l’affinité *in vitro* pour le domaine WW de Pin1. L’établissement des relations structure-affinité au niveau des ligands et la recherche de nouveaux points d’interaction d’après la structure du domaine WW permettront de définir de nouvelles molécules pour lesquelles on peut espérer une meilleure affinité. Un test cellulaire évaluera finalement, pour les composés les plus affins, leur effet *in vivo*.



**Schéma 4.1 :** Description de la stratégie employée pour la synthèse, le criblage et l’évaluation des ligands potentiels du domaine WW de Pin1.

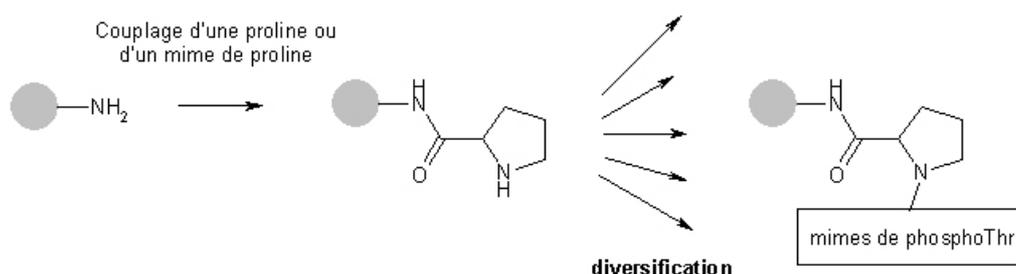
La conception de mimes du dipeptide phosphoThr-Pro (composé-référence **1**) a été décomposée en deux parties consistant en la recherche, d’une part, d’analogues de la partie polaire constituée du groupement phosphate (têtes polaires) et d’autre part, de mimes de la partie apolaire constituée du cycle hydrophobe de type pyrrolidinique. Afin de minimiser les changements structuraux pour que l’affinité des composés reste mesurable, les analogues de phosphothréonine seront associés à la proline, et inversement, les mimes de proline seront associés à la phosphothréonine.

## 1.2 – Les outils mis en œuvre pour la synthèse et l'évaluation des ligands potentiels du domaine WW de Pin1.

Dans un premier temps, l'élaboration de molécules capables de lier le domaine WW nécessite, pour être efficace, un test permettant de discriminer rapidement les composés « actifs » des composés « non actifs ». Le choix du domaine WW comme modèle simplifié dans la recherche d'inhibiteurs de Pin1 permet ainsi l'utilisation d'un test relativement simple basé uniquement sur la capacité de liaison des composés au domaine WW, beaucoup plus facile à mettre en œuvre qu'un test d'inhibition d'une activité *cis/trans* isomérase. D'autre part, la stratégie de synthèse utilisant la chimie supportée permet de mettre au point des conditions de réaction qui pourront être étendues à l'obtention de petites banques de composés par synthèse combinatoire.

### 1.2.1 – Synthèse en phase solide de ligands potentiels du domaine WW.

Le principe de la synthèse peptidique en phase solide a été repris et adapté à la synthèse des mimes du dipeptide phosphoThr-Pro (**1**) pour deux raisons. La première est que les composés de type mimes de proline couplés à une phosphothréonine sont très proches d'un point de vue structural de la molécule de base. On conserve notamment la liaison amide présente dans le dipeptide (**1**). La deuxième est que la stratégie de synthèse en phase solide permet de synthétiser un lot d'analogues de proline couplés à la résine et ensuite de diversifier les têtes polaires de manière simple et rapide (Schéma 4.2).



**Schéma 4.2 :** Stratégie de synthèse par couplage de la proline sur une phase solide et diversification ultérieure des têtes polaires.

La mise au point d'une stratégie de synthèse supportée pourra ensuite être employée en synthèse combinatoire pour assembler, à l'issue d'un premier criblage, les mimes de proline et les analogues de phosphothréonine pertinents.

La synthèse supportée présente de nombreux avantages par rapport à la synthèse en phase homogène. L’utilisation d’excès de réactifs permet d’obtenir de meilleurs rendements, les excès de réactifs ainsi que les produits secondaires formés en solution peuvent être aisément éliminés par lavage. Les synthèses comportant très souvent plusieurs étapes réactionnelles ne nécessitent pas l’isolement et la purification des intermédiaires réactionnels. Cependant, l’inconvénient majeur de la synthèse supportée est le coût élevé dû à l’utilisation d’excès de réactif, de grandes quantités de solvant et du coût du support solide. Les échelles de synthèse doivent par conséquent être optimisées pour obtenir la quantité minimale nécessaire au criblage et aux études de *binding*, c’est-à-dire quelques dizaines de milligrammes de composé, ce qui revient à l’utilisation de quantités de résine de l’ordre de quelques grammes selon la charge de celle-ci.

### **1.2.2 – Choix d’une technique de criblage et de quantification de la liaison des composés au domaine WW de Pin1.**

Le choix d’une technique de criblage des molécules synthétisées prend en compte les critères tels que la sensibilité de la méthode, la reproductibilité, le coût... Le criblage des différents phospho-épitopes Ser/Thr-Pro au sein de la protéine tau a été réalisé par RMN dans le but de déterminer le mode d’interaction des différents peptides avec Pin1, à savoir le mode liaison au domaine WW et les résidus impliqués dans la liaison, la caractérisation de l’interaction avec le domaine catalytique et la caractérisation de l’activité enzymatique, la mise en évidence d’une éventuelle coopération entre les deux domaines pour la liaison ou pour l’isomérisation catalytique. Les efforts à fournir pour l’obtention de telles données sont considérables en termes de coût et de temps, notamment pour la production de quantités relativement importantes de protéines recombinantes enrichies en  $^{15}\text{N}$ , de peptides sélectivement phosphorylés obtenus par synthèse en phase solide et de temps machine sur un spectromètre RMN à 600MHz. La mise au point d’un test de criblage doit également prendre en considération le nombre de molécules à tester.

La RMN peut être optimisée dans le but de cribler de petites banques constituées de quelques dizaines de composés et même des collections plus importantes en testant les composés en mélange (Hajduk, Gerfin et al. 1999). L’étape de criblage ne nécessite pas obligatoirement, dans le cas d’un petit domaine tel que le domaine WW de Pin1, un lot de protéine enrichie en  $^{15}\text{N}$ . La discrimination des ligands des non-ligands se fait par l’intermédiaire de perturbations des résidus de la protéine et peut tout à fait utiliser les résonances  $^1\text{H}$  uniquement. De manière

évidente, l’utilisation d’un lot de protéine enrichie uniformément en  $^{15}\text{N}$  facilite la détection des signaux  $\text{H}_\text{N}$  de la protéine puisqu’en présence d’excès de ligand, on peut ainsi s’affranchir, par filtrage  $^{15}\text{N}$ , des signaux de celui-ci ainsi que de la majeure partie du signal de l’eau. L’avantage de la RMN comme technique de criblage est de pouvoir déterminer, en même temps que l’affinité, le mode de liaison du ligand, ce qui nécessite alors l’enrichissement isotopique  $^{15}\text{N}$  de la protéine-cible. De plus, les composés qui vont être synthétisés et testés initialement, à l’instar du dipeptide (**1**), seront très peu affins vis-à-vis du domaine WW car très proches du dipeptide d’un point de vue structural, et la RMN est une technique particulièrement bien adaptée à la détection d’affinités faibles. Cependant, l’un des handicaps les plus importants de la RMN est la faible sensibilité et le fait que le temps d’acquisition des données soit inversement proportionnel à la concentration de la protéine en solution.

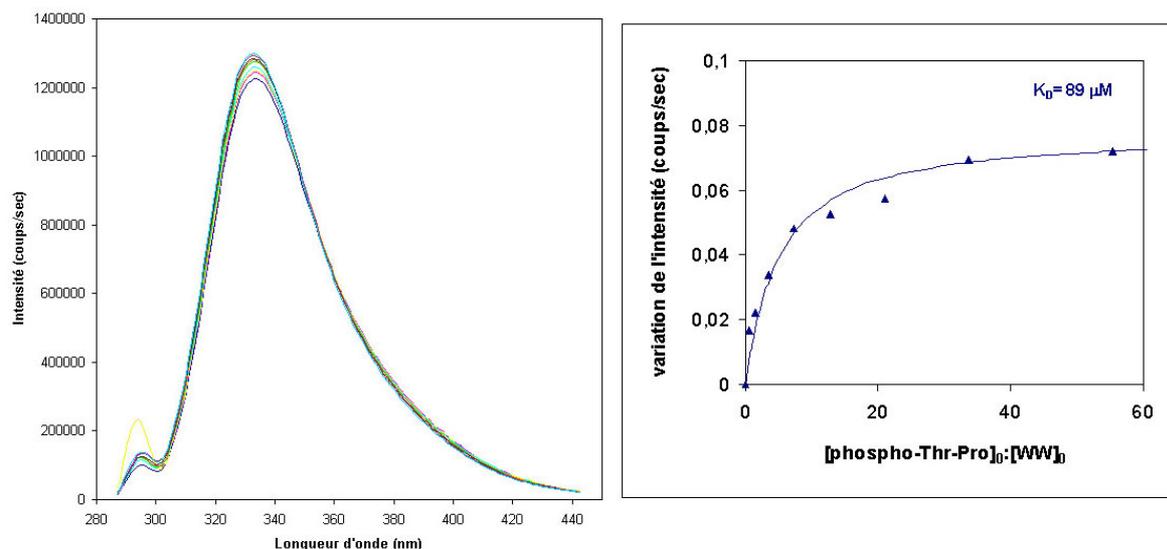
La fluorescence (FRET ou polarisation) est une technique plus largement utilisée pour le criblage de composés et présente un double avantage par rapport à la RMN : la réduction des quantités de protéine mise en œuvre pour le criblage, et par conséquent de composé à tester, et un temps considérablement réduit pour l’acquisition des données. Travailler avec une concentration en protéine de 20  $\mu\text{M}$  nécessite une acquisition d’au moins 2 heures en RMN sans cryosonde et 20 min avec cryosonde alors qu’elle ne nécessite que quelques minutes avec un fluorimètre à balayage pour l’acquisition d’un spectre d’émission. La sensibilité et la rapidité font de la fluorescence une technique intéressante pour le criblage des molécules dans la recherche d’inhibiteurs du domaine WW de Pin1. Mais la faible affinité du composé-référence et celles attendues pour les premières molécules mimant celui-ci peuvent en limiter l’usage. Nous avons évalué, dans un premier temps, l’utilisation de la fluorescence comme méthode de criblage.

#### **1.2.2.1 – Compétition entre un phospho-peptide marqué et un ligand pour la liaison au domaine WW suivie par FRET.**

##### ***Fluorescence intrinsèque du domaine WW et détermination de $K_D$ .***

Comme nous l’avons décrit dans le chapitre II pour les peptides phosphorylés, il est possible de mettre en œuvre la fluorescence intrinsèque du domaine WW pour détecter des phénomènes de liaison qui modifient l’hydrophobicité de l’environnement du tryptophane. Nous avons envisagé d’exploiter la fluorescence intrinsèque du domaine WW pour le criblage

des composés et pour la détermination des affinités. Cette technique s’est avérée limitée par les faibles variations des deux paramètres mesurés, l’intensité d’émission de fluorescence et la position du maximum d’émission (Figure 4.2). D’autre part, l’utilisation de concentrations importantes en ligand due à de faibles affinités conduit à un phénomène d’« écran » qui se caractérise par une forte diminution de l’énergie émise par le tryptophane et qui avait déjà été observé avec les phospho-peptides de tau de plus faibles affinités (chapitre II).



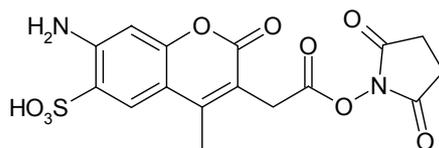
**Figure 4.2 :** Titration du domaine WW en concentration constante à 20  $\mu\text{M}$  par le dipeptide (**1**). La procédure utilisée est celle décrite dans le chapitre II pour la titration du domaine WW par des peptides phosphorylés. On notera que la variation maximale de l’intensité de fluorescence à 334 nm est plus faible qu’avec les peptides. Elle représente ici environ 6% de l’intensité maximale.

La constante de dissociation d’un complexe est une valeur universelle que permet de définir la quantité de complexe dans des conditions données. La RMN et la fluorescence intrinsèque permettent d’accéder à des données quantitatives de dissociation. Mais il n’est pas toujours possible de calculer un  $K_D$ . Dans certains cas, il n’est pas possible d’isoler et de quantifier la protéine-cible (récepteurs membranaires et nucléaires, canaux ioniques...). Le criblage de composés peut alors faire intervenir une autre stratégie basée sur la compétition avec un ligand ou un substrat connu pour quantifier l’interaction. Les paramètres mesurés sont l’ $\text{IC}_{50}$  ou la constante d’inhibition  $K_i$  qui définissent les concentrations de composés nécessaires pour déplacer un ligand ou inhiber une activité. Le FRET est classiquement utilisé dans cette optique à la fois en test biochimique et en test cellulaire. Il faut noter cependant que, contrairement au  $K_D$  qui est une valeur absolue, l’ $\text{IC}_{50}$  ou le  $K_i$  sont des valeurs qui dépendent

du système mis en œuvre et ne sont valables que pour comparer des composés évalués dans les mêmes conditions. La connaissance d’un ligand « naturel » du domaine WW, le peptide centré autour de l’épitope Thr212 phosphorylé, a permis d’utiliser le FRET en mettant en compétition les molécules synthétisées avec le phospho-peptide.

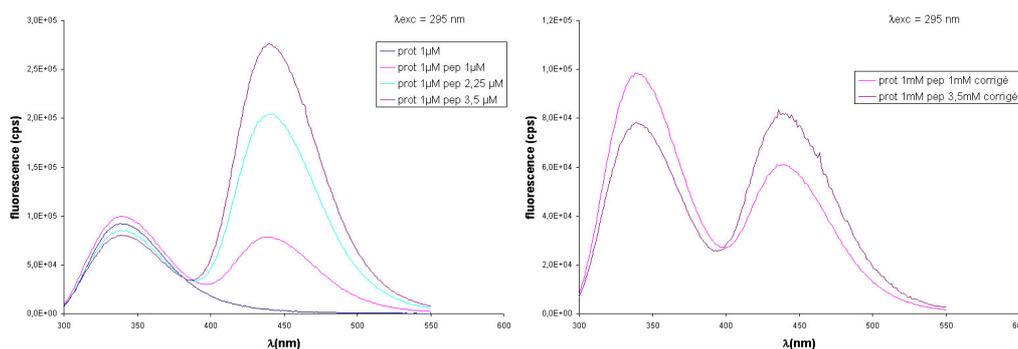
#### ***Compétition avec un phospho-peptide suivie par FRET et détermination d’IC<sub>50</sub>.***

Le choix de la sonde fluorescente est important pour l’élaboration d’un test fiable, rapide et sensible. L’émission du dansyle à 520 nm est faible dans l’eau, par conséquent seul le processus de *quenching* du Trp permet de quantifier l’affinité du domaine WW pour le peptide. L’AlexaFluor350<sup>®</sup> (Molecular Probes) possède les qualités requises qui sont : le maximum de longueur d’onde d’excitation (350 nm) - coïncidant avec le maximum d’émission du Trp dans le domaine WW, un bon rendement quantique, une bonne brillance, la photostabilité et la solubilité en solution aqueuse.



**Figure 4.3** : Structure de l’AlexaFluor350<sup>®</sup> sous forme d’ester de succinimide, qui permet un couplage sur les amines primaires.

La sonde fluorescente (Figure 4.3) est couplée sélectivement à la fonction amine située sur la chaîne latérale du résidu lysine placé en position N-terminale sur le peptide N-acétylé centré autour du phospho-épitope Thr212 (voir séquence dans le chapitre II). La constante de dissociation du complexe formé par le domaine WW et le peptide fluorescent est évaluée à 20  $\mu$ M selon la méthode utilisée pour la détermination de l’affinité des phospho-épitopes de la protéine tau décrite dans le chapitre II. La constante de dissociation est calculée à partir des variations de l’émission de fluorescence des résidus Trp du domaine WW à 334 nm. L’émission de fluorescence à 450 nm de la sonde AlexaFluor<sup>®</sup>350 excitée à 295 nm n’est pas négligeable. Par conséquent l’émission à 450 nm en présence du domaine WW correspond au FRET dû à la liaison du domaine WW au peptide porteur de la sonde fluorescente et à l’excitation directe de la sonde. Les variations d’émission de fluorescence à 450 nm ne peuvent être exploitées que si elles sont corrigées par l’émission de la sonde seule excitée à 295 nm (Figure 4.4).



**Figure 4.4 :** Expérience de FRET entre le domaine WW et le peptide centré autour de l’épitope Thr212 couplé covalamment à l’AlexaFluor 350® (Molecular Probes) après excitation à 295 nm. A gauche, spectre d’émission de fluorescence du domaine WW seul à une concentration de 1 μM excité à 295 nm (en bleu foncé) et spectres d’émission du complexe WW-peptide pour différentes concentrations en peptide (en rose, ratio 1 :1, en bleu clair, ratio 1 :2.25, en violet, ratio 1 :3.5). A droite, spectres d’émission corrigés par l’émission du peptide seul excité à 295 nm pour les ratios 1 :1 (en rose) et 1 :3.5 (en violet).

La compétition des molécules avec un phospho-peptide marqué pour la liaison au domaine WW peut être exploitée pour le criblage des différents composés synthétisés mais également dans la détermination de la concentration inhibitrice à 50% ou  $IC_{50}$ , à partir de la courbe représentant le pourcentage d’inhibition dont l’expression est décrite par l’Équation 4.1, en fonction de la concentration de composé ajouté à un mélange WW/phospho-peptide de concentration constante.

$$\% \text{ Inhibition} = \frac{I - I_0}{I_0 - I_{100}}$$

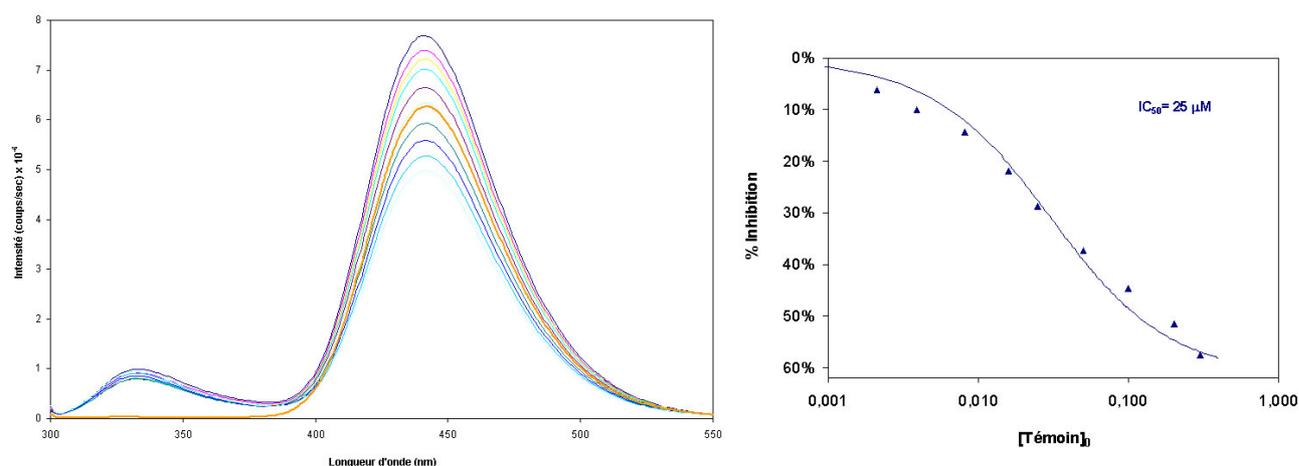
**Équation 4.1 :** Calcul du pourcentage d’inhibition d’un composé pour un mélange WW : phospho-peptide fixé où  $I$  est la valeur de l’émission de fluorescence à 334 nm du domaine WW pour une quantité donnée de ligand,  $I_0$  et  $I_{100}$ , les valeurs de l’émission à 334 nm à 0% et 100% d’inhibition respectivement.

Par exemple, quand deux composés entrent en compétition pour la liaison d’un même site, si on connaît le  $K_D$  du ligand ( $K_D^L$ ), on peut calculer le  $K_D$  du compétiteur ( $K_D^I$ ) à partir de l’ $IC_{50}$  par la formule de Cheng et Prusoff (Équation 4.2).

$$K_D^L = \frac{[L] K_D^I}{IC_{50} - K_D^I}$$

**Équation 4.2 :** Formule reliant les constantes de dissociation d’un ligand  $L$  par rapport à une protéine-cible, la concentration fixe  $[L]$  du ligand et l’ $IC_{50}$  d’un inhibiteur  $I$  à la constante de dissociation de celui-ci.

La validité du système mis en oeuvre est vérifiée par une expérience mettant en compétition pour la liaison du domaine WW un phospho-peptide marqué et le même phospho-peptide non marqué. L’expérience est réalisée avec une solution de domaine WW à 10  $\mu\text{M}$  dans un tampon Tris 50 mM (pH 6.8), NaCl 100 mM, DTT 5 mM en présence de 25  $\mu\text{M}$  de phospho-peptide (pT212) marqué avec la sonde AlexaFluor350. Des quantités croissantes de peptide témoin pT212 non marqué sont ajoutées successivement à la solution et des spectres d’émission de fluorescence sont acquis après excitation à 295 nm (Figure 4.5).

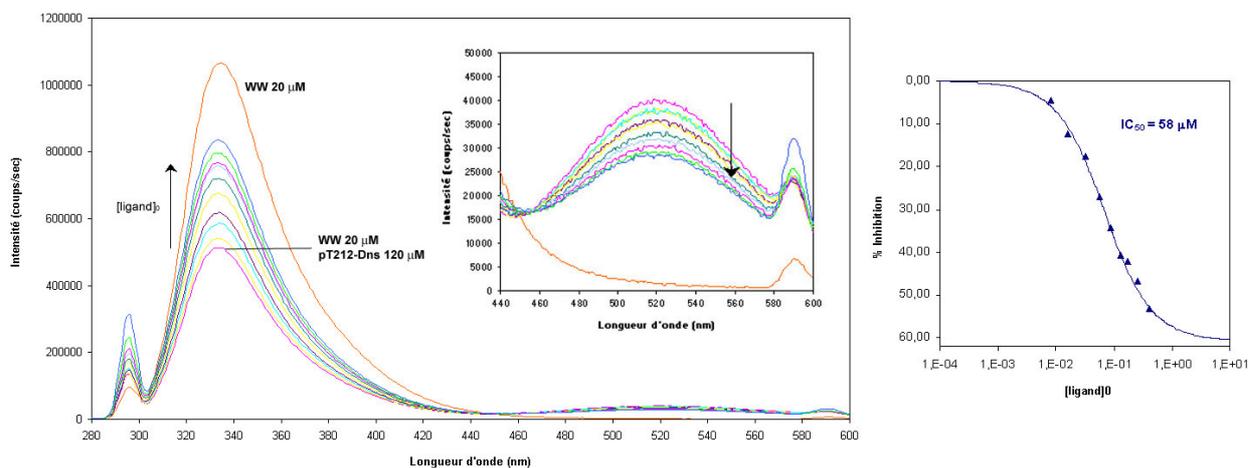


**Figure 4.5 :** Compétition de liaison entre le domaine WW à 10  $\mu\text{M}$  et le peptide pT212 porteur d’une sonde AlexaFluor350<sup>®</sup> à 25  $\mu\text{M}$  avec un peptide témoin (pT212) sans sonde fluorescente. La concentration inhibitrice à 50% est évaluée à 25  $\mu\text{M}$ .

La concentration pour laquelle l’inhibition de la liaison du peptide pT212 marqué au domaine WW est de 50% est évaluée à 25  $\mu\text{M}$  (Figure 4.5). Cette valeur n’est pas en accord avec l’affinité déterminée pour les deux peptides, qui sont de 100  $\mu\text{M}$  pour le peptide sans sonde et de 20  $\mu\text{M}$  pour le peptide marqué. D’autre part, l’émission de fluorescence du domaine WW en l’absence de FRET ( $I_{100}$ ), c’est-à-dire en l’absence de peptide marqué, n’est pas récupérée avec un large excès de peptide témoin et sature à 60% du signal de fluorescence du domaine WW seul. L’émission de fluorescence à 450 nm en présence d’un large excès de peptide témoin est inférieure à l’émission de fluorescence du peptide marqué seul, excité à 295 nm (courbe orange sur la Figure 4.5). L’émission de fluorescence de la sonde par excitation à 295 nm n’est pas négligeable et cette intensité de fluorescence est atteinte pour une quantité équivalente de peptide marqué et de peptide témoin. Ces résultats indiquent un comportement anormal du système mis en oeuvre pour la détermination des  $\text{IC}_{50}$ , suggérant une interaction

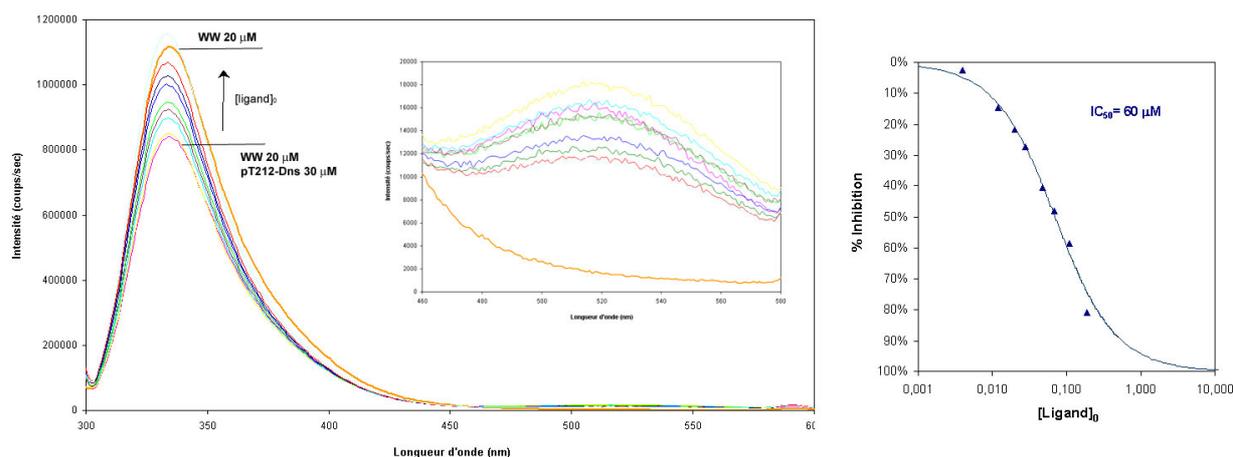
supplémentaire entre le domaine WW et le peptide marqué au niveau d’un site de liaison différent de celui qui intervient dans la liaison du dipeptide phosphoThr-Pro. On peut donc supposer une interaction entre le domaine WW et la sonde fluorescente elle-même puisqu’il s’agit du seul élément qui différencie les deux peptides, ce qui expliquerait par ailleurs une affinité pour le peptide fluorescent cinq fois supérieure à l’affinité du peptide témoin.

Le choix d’une autre sonde fluorescente s’est porté sur le dansyle qui a été utilisé dans le chapitre II pour la détermination de l’affinité de phospho-épitopes de la protéine tau reconnus par Pin1. Le rapport des concentrations entre le domaine WW et le peptide phosphorylé marqué est fixé à 1 : 6 avec une concentration de 20  $\mu\text{M}$  de protéine en solution, ce qui correspond d’après la courbe de *binding* à environ 50% de saturation de la protéine. Malgré une apparente cohérence des résultats, c’est-à-dire une augmentation du signal de fluorescence à 334 nm et une diminution du signal à 520 nm en fonction de concentrations croissantes en ligand, les problèmes décrits précédemment avec la sonde AlexaFluor sont également rencontrés avec le dansyle à savoir une saturation du signal à 334 nm à 60% du signal de fluorescence du domaine WW seul. L’ajout de concentrations importantes de ligand nécessaires au déplacement du peptide phosphorylé provoque ensuite une diminution du signal à 334 nm (Figure 4.6).



**Figure 4.6 :** Détermination de l’ $IC_{50}$  d’un composé par FRET. Compétition d’un ligand avec le peptide pT212 à 120  $\mu\text{M}$  porteur d’une sonde dansyle pour la liaison au domaine WW (20  $\mu\text{M}$ ). A gauche, augmentation de l’intensité de fluorescence à 334 nm avec l’augmentation de la concentration en ligand avec, en encadré, un agrandissement du spectre d’émission dans la zone 440-600 nm, montrant une diminution de l’émission de fluorescence du groupement dansyle à 520 nm avec l’augmentation de la concentration en ligand. A droite, calcul de l’ $IC_{50}$  par mesure de la variation de l’intensité de fluorescence à 334 nm en fonction de la quantité de ligand.

Une autre expérience a été réalisée avec cette fois un rapport de concentration de 1 : 1.5 entre le domaine WW et le peptide phosphorylé marqué, correspondant à 20% de saturation de la protéine (Figure 4.7). Dans ce cas de figure, la totalité du signal du domaine WW est récupérée avec un excès de ligand d’environ 4 équivalents par rapport au peptide phosphorylé. Avec des concentrations supérieures en ligand, le signal à 334 nm dépasse le signal de fluorescence du domaine WW seul car la liaison du ligand modifie la fluorescence intrinsèque du domaine WW (Figure 4.2). La valeur de l’intensité de fluorescence à 100% d’inhibition ( $I_{100}$ ) est par conséquent déterminée par mesure de la fluorescence intrinsèque du domaine WW en l’absence de peptide marqué et en présence d’une concentration de ligand pour laquelle l’émission à 334 nm est maximale. On obtient alors un signal à 334 nm de 10% supérieur à celui du domaine WW en l’absence de ligand. On fait donc les calculs d’ $IC_{50}$  avec une valeur d’ $I_{100}$  correspondant à 110% de la valeur de fluorescence déterminée pour le domaine WW seul.



**Figure 4.7 :** Détermination de l’ $IC_{50}$  d’un composé par FRET. Compétition d’un ligand avec le peptide pT212 à 30  $\mu$ M porteur d’une sonde dansyle pour la liaison au domaine WW (20  $\mu$ M). A gauche, augmentation de l’intensité de fluorescence à 334 nm avec l’augmentation de la concentration en ligand avec, en encadré, un agrandissement du spectre d’émission dans la zone 440-600 nm, montrant une diminution de l’émission de fluorescence du groupement dansyle à 520 nm avec l’augmentation de la concentration en ligand. A droite, calcul de l’ $IC_{50}$  par mesure de la variation de l’intensité de fluorescence à 334 nm en fonction de la quantité de ligand.

En résumé, la détermination de l’affinité des ligands par mesure de la fluorescence intrinsèque du domaine WW reste limitée aux composés non fluorescents et les variations du maximum d’émission étant relativement faibles, la détermination du  $K_D$  des différents ligands potentiels est peu fiable dans ce contexte. La compétition avec un ligand de nature peptidique suivie par

FRET ne permet pas d’accéder directement au  $K_D$  mais à l’ $IC_{50}$  des composés. Les problèmes rencontrés avec cette technique sont similaires à ceux décrits précédemment comme les concentrations élevées de ligand et de phospho-peptide marqué qui vont de paire avec les faibles affinités et qui mènent notamment à une diminution des signaux de fluorescence par effet d’« écran ».

Nous avons mis au point des conditions de criblage par RMN pour déterminer les constantes de dissociation des composés et différencier des ligands dans la gamme des faibles affinités. Nous avons largement utilisé cette technique dans la recherche de phospho-épitopes de la protéine tau reconnus par Pin1 dans le chapitre II et nous avons pu l’optimiser facilement pour le criblage de ligands potentiels et la détermination des constantes de dissociation des composés issus du criblage.

#### **1.2.2.2 – La RMN : une technique mieux adaptée aux faibles affinités.**

La RMN peut être utilisée efficacement dans la recherche de ligands du domaine WW de Pin1 pour plusieurs raisons :

- Le domaine WW est une petite protéine d’environ 4,4 kDa pour laquelle le système d’expression permet la production de plusieurs milligrammes de protéine enrichie en  $^{15}N$  (environ 10 mg/L en milieu minimum M9) et dont le nombre restreint de résidus s’accompagne d’une bonne résolution des signaux RMN au niveau de spectres  $^1H$ - $^{15}N$  HSQC.
- Une largeur de raie fine due à un temps de relaxation long en solution est également un atout pour mettre en œuvre un test de criblage relativement rapide car il est alors possible d’utiliser de faibles quantités de protéine pour des temps d’acquisition de quelques dizaines de minutes.

#### ***Utilisation de la RMN à haute résolution pour le criblage des composés.***

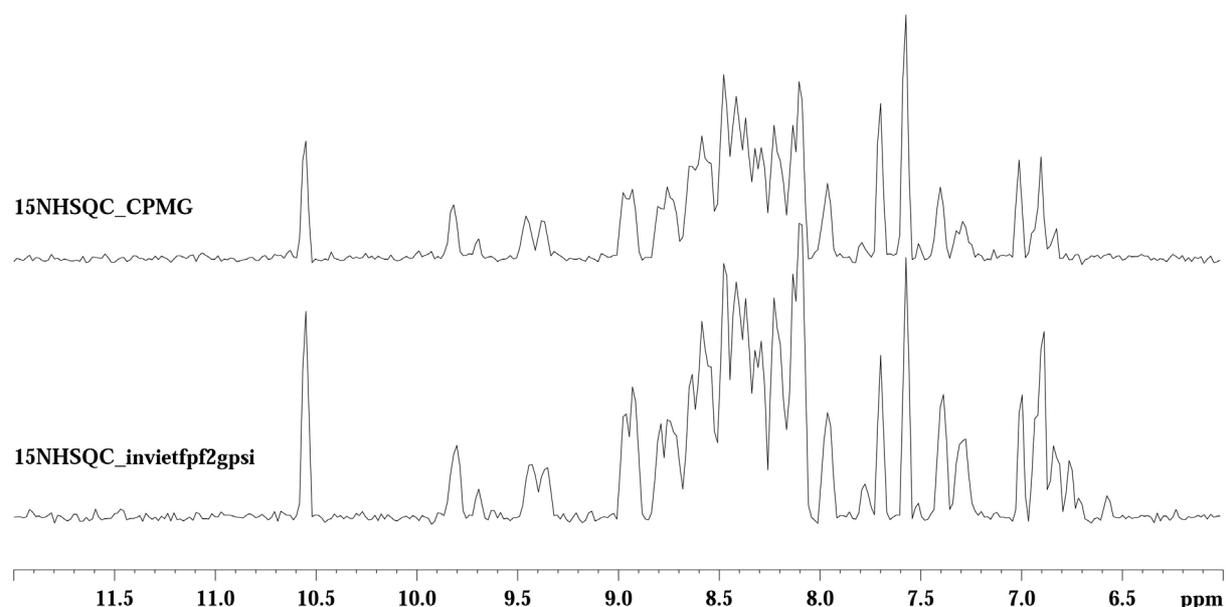
L’utilisation de la RMN $^1H$  comme technique de criblage ne nécessite pas obligatoirement l’attribution des signaux de la protéine ni la production de protéine enrichie en  $^{15}N$ , plus onéreuse et donnant de moins bons rendements. Cependant, l’acquisition des spectres protons ne permet pas de s’affranchir des résonances des ligands présents en large excès qui peuvent masquer les signaux de la protéine (voir un exemple Figure 4.14). Afin d’exploiter au maximum toutes les données accessibles grâce à la RMN à haute résolution, le criblage a été

mis en œuvre avec l’utilisation de protéine enrichie en  $^{15}\text{N}$  et l’acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC permettant l’identification des résonances affectées par le ligand dès l’étape de criblage et de détecter de manière précoce un mode de liaison du ligand différent du dipeptide phosphoThr-Pro, privilégiant ainsi la qualité de l’information par rapport à la quantité de composés testés.

Différents paramètres doivent être pris en considération pour effectuer le criblage des composés synthétisés par RMN dans la recherche de ligands du domaine WW. La validation de la RMN mono-dimensionnelle passe par l’évaluation de la concentration de protéine en solution conjointement avec l’optimisation du temps d’acquisition et le choix d’une séquence impulsionnelle adéquate, ainsi que l’évaluation de la concentration en ligand à mettre en œuvre pour détecter une liaison sur une gamme d’affinités aussi large que possible.

#### ***Optimisation de la séquence impulsionnelle.***

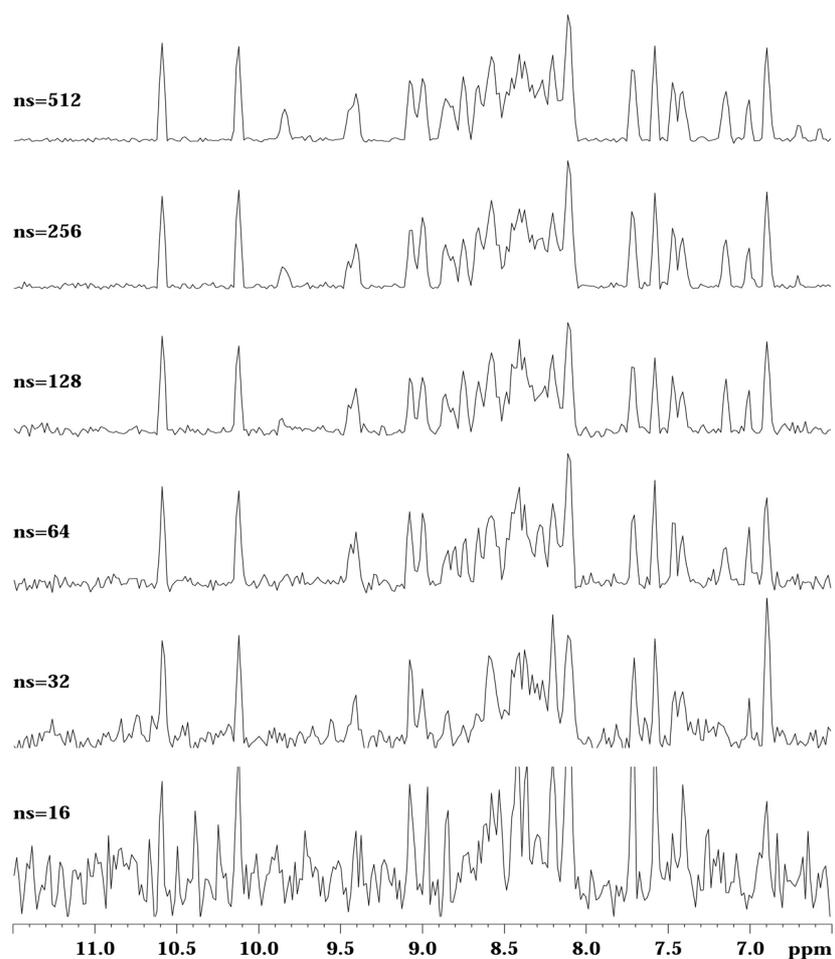
L’efficacité d’une technique de criblage réside en la quantité de composés qui peuvent être testés en minimisant les artefacts conduisant à la détection de faux positifs et faux négatifs, et en minimisant le matériel et le temps nécessaires à l’acquisition de l’information. L’utilisation d’une séquence impulsionnelle plus sensible pour l’acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC, que la séquence classique CPMG, permet d’augmenter le rapport signal : bruit pour une même concentration en protéine comparativement à une séquence impulsionnelle classique avec un simple transfert d’INEPT (Cavanagh and Rance 1993).



**Figure 4.8 :** Comparaison des séquences impulsionnelles  $^1\text{H}$ - $^{15}\text{N}$  HSQC sur le spectre  $^1\text{H}$  du domaine WW. L’augmentation de la sensibilité de la détection est estimée à un gain de 40%.

**Optimisation de la concentration en protéine et du temps d’acquisition.**

Le temps d’acquisition nécessaire étant inversement proportionnel à la concentration de protéine en solution, il est nécessaire de trouver un compromis entre ces deux facteurs. La RMN étant une technique peu sensible, une concentration relativement importante en protéine est requise pour l’obtention de données de bonne qualité. L’effort à fournir pour exprimer et purifier la protéine enrichie en  $^{15}\text{N}$  n’est pas proportionnel à la quantité de protéine produite. En effet, exprimer et purifier 20 mg de WW enrichi en  $^{15}\text{N}$ , soit 2L de culture en milieu minimum M9, est identique à la production de 10 mg soit 1L de culture. L’accumulation des spectres permet d’augmenter le rapport signal : bruit de sorte que le signal propre à la protéine devienne exploitable. Le rapport signal : bruit (S/N) est proportionnel à la racine carrée du nombre de *scans* (illustré avec une sonde classique BBI sur 20  $\mu\text{M}$  de protéine Figure 4.9). Diminuer la concentration de protéine en solution d’un facteur 2 revient à multiplier par 4 le nombre de *scans* nécessaires pour l’acquisition d’un spectre avec le même rapport S/N.



**Figure 4.9 :** Diminution du nombre de *scans* sur un échantillon de protéine de concentration constante (20  $\mu\text{M}$ ) sur un spectromètre à 300 MHz équipé d’un sonde BBI. En normalisant les spectres sur l’intensité des signaux à 10.6 et 10.2 ppm, on peut observer l’augmentation du bruit pour une intensité du signal équivalente.

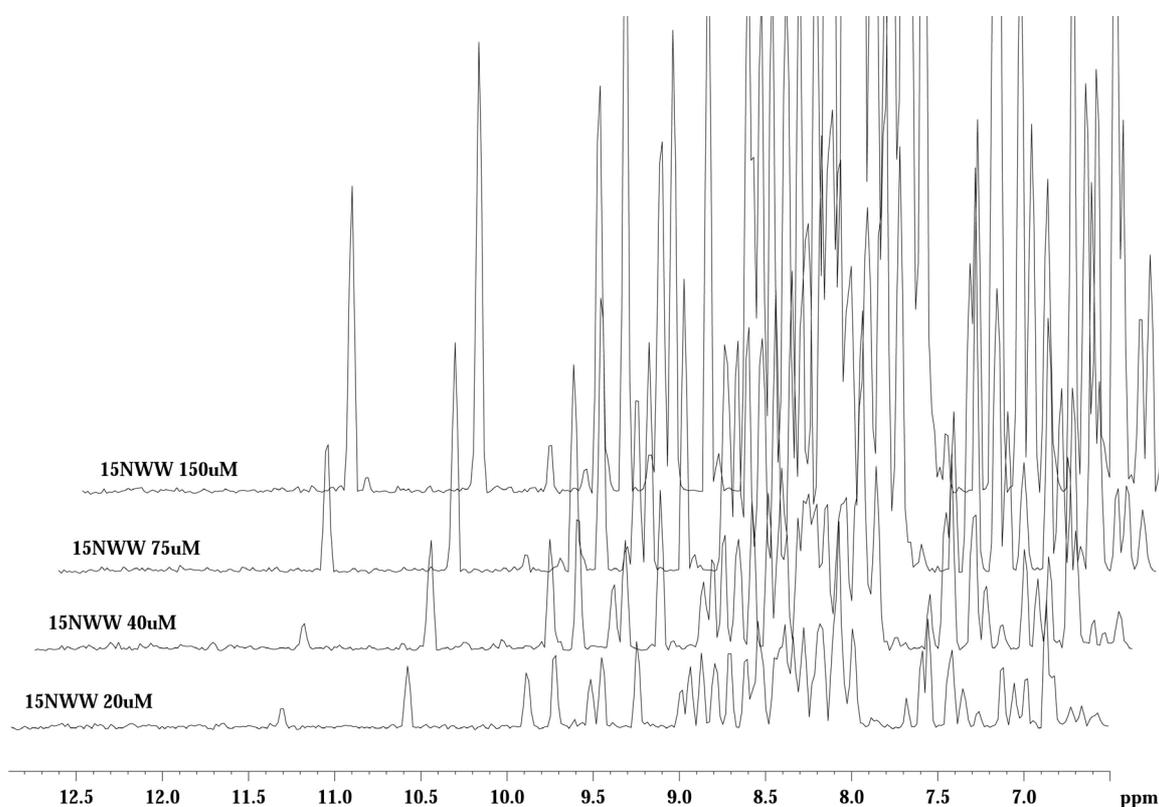
Sur un spectromètre équipé d’une cryosonde, qui permet d’augmenter la sensibilité de la détection de manière significative par rapport à une sonde classique, il est possible de faire l’acquisition d’un spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC avec 20  $\mu\text{M}$  de protéine en 20 minutes, soit 8 *scans*. Il s’agit d’un temps acceptable qui permet de tester individuellement jusqu’à 16 ligands par jour, ou d’effectuer deux titrations par jour avec 8 concentrations de ligand par titration pour déterminer les constantes de dissociation.

L’optimisation de la concentration de protéine utilisée pour la détermination de la liaison ou non des composés au domaine WW a été réalisée sur un spectromètre à 600MHz équipé d’une cryosonde. L’acquisition d’un spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC est réalisée sur un échantillon de protéine contenant 20 équivalents de ligand, conditions dans lesquelles la liaison du ligand est en particulier détectée au travers d’une résonance à 11.35 ppm correspondant à l’Arg17 impliquée dans la liaison du phosphate. Ce signal est relativement élargi comparativement aux autres signaux de la protéine notamment ceux qui ne sont pas impliqués dans la liaison du ligand. Il faut également souligner que dans le cas de ligands ayant une constante de dissociation de l’ordre de 100  $\mu\text{M}$ , le signal de l’Arg17 dans sa forme liée est visible, ce qui n’est pas le cas pour des ligands moins affins avec ce même rapport molaire ligand : protéine. Pour des affinités plus faibles, certains signaux des résidus impliqués dans la liaison du composé peuvent être fortement élargis et difficilement détectables. Il faut donc tenir compte d’un tel comportement lorsque l’on détermine la concentration de protéine mise en jeu pour le criblage des composés.

L’acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC avec 8 *scans* (soit 20 minutes d’acquisition) a été effectuée sur des solutions de protéine de différentes concentrations, obtenues par dilutions successives à partir d’une solution de concentration 150  $\mu\text{M}$  contenant 3 mM de ligand (Figure 4.10). On note que, pour une concentration de 20  $\mu\text{M}$ , le signal à 11,35 ppm (Arg17) est encore observable mais devient plus difficile à détecter alors que le signal à 10,6 ppm correspondant au Trp11 qui n’intervient pas dans la liaison du ligand, est beaucoup plus fin. De plus, en diluant la protéine, on diminue la quantité de complexe même si le ratio ligand : protéine reste constant. Ainsi, les signaux comme le Trp11 qui ne sont pas impliqués dans la liaison du ligand vont diminuer linéairement avec la dilution alors que les signaux des résidus impliqués dans l’interaction comme l’Arg17 vont diminuer de manière non linéaire. La

concentration de 20  $\mu\text{M}$  constitue par conséquent une limite inférieure pour laquelle le rapport signal : bruit donne des données exploitables avec 8 *scans*.

Pour résumer, sur un spectromètre équipé d’une cryosonde qui permet d’augmenter de manière notable la sensibilité, avec des acquisitions de 8 *scans* et une concentration de 20  $\mu\text{M}$  de domaine WW enrichi en  $^{15}\text{N}$ , il est possible de tester 16 molécules ou de faire 2 titrations par jour. Avec un spectromètre équipé d’une sonde BBI (à 300MHz), le même rapport signal : bruit est obtenu avec une concentration de protéine multipliée par un facteur 8 ou 64 fois plus de *scans*, soit 160  $\mu\text{M}$  avec 8 *scans* ou 20  $\mu\text{M}$  avec 512 *scans*.



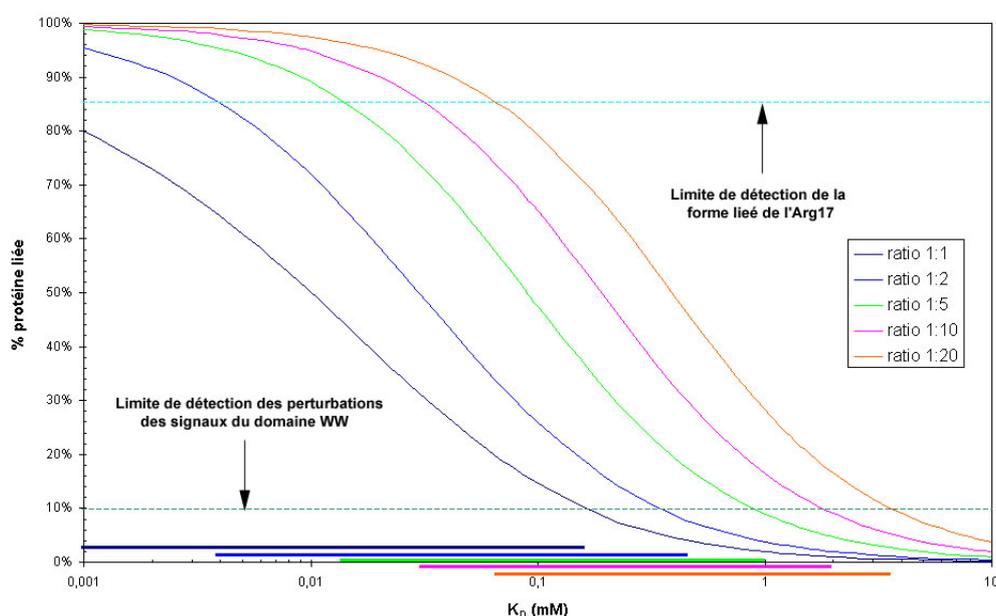
**Figure 4.10 :** Diminution de la concentration en protéine de 150 à 20  $\mu\text{M}$  en présence de 20 équivalents de ligand, et effet sur le rapport signal : bruit lors de l’acquisition de spectres protons découplés  $^{15}\text{N}$  avec 8 *scans* sur un spectromètre à 600MHz équipé d’une cryosonde.

#### ***Choix d’une concentration de ligand.***

Pour le criblage des ligands potentiels du domaine WW, 20 équivalents de ligand sont ajoutés à une solution contenant 20  $\mu\text{M}$  de domaine WW enrichi en  $^{15}\text{N}$  dans un tampon Tris deutéré à 25 mM (pH 6.8) contenant 50 mM NaCl et 2.5 mM DTT. Le choix d’un ratio molaire ligand : protéine est déterminé en fonction de la concentration en protéine et de l’ordre de grandeur de l’affinité qu’il faut détecter.

Avec le système que nous étudions, on peut définir trois plages d’affinité à partir de deux paramètres :

- le seuil de détection des perturbations des signaux de la protéine, qui est estimé à 10% de protéine complexée,
- le signal de la forme liée de l’Arg17 (le signal de la forme libre n’est pas observable en l’absence de ligand à cause de la dynamique de la boucle) qui apparaît pour une proportion de protéine complexée estimée à 80-85% (Graphique 4.1, Figure 4.11). Il faut cependant souligner que la forme liée de l’Arg17 peut résonner à une autre fréquence que celle indiquée sur la figure en présence d’un ligand contenant un analogue du phosphate.

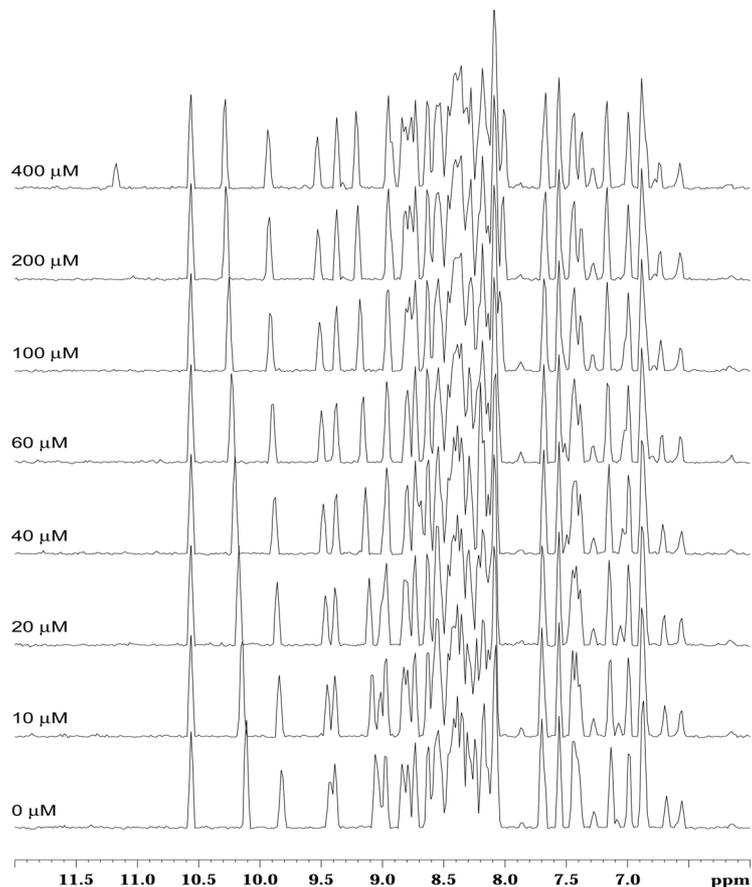


**Graphique 4.1 :** Gammes des constantes de dissociation détectables en fonction de la concentration en ligand utilisée pour le criblage pour une concentration de protéine de 20  $\mu\text{M}$  avec un seuil de détection des variations de déplacements chimiques fixé à 10% de protéine liée. Un plafond de 85% a été fixé par rapport au signal de l’Arg17 à partir duquel ce signal devient détectable. Cette valeur donne un ordre de grandeur de la limite inférieure du  $K_D$  d’un ligand dès l’étape de criblage. Avec une concentration de ligand de 400  $\mu\text{M}$ , soit un ratio ligand : protéine de 20, la gamme des constantes de dissociation se situe entre 70  $\mu\text{M}$  et 3 mM.

Ces valeurs ont été déterminées à partir des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC du composé de référence (**1**). Un ratio de 20 entre la concentration du ligand et la concentration de la protéine (pour une concentration de WW de 20  $\mu\text{M}$ ) permet de distinguer :

- les composés ayant une constante de dissociation supérieure à 3 mM si aucune perturbation des signaux du domaine WW n’est détectée,

- les composés ayant une constante de dissociation inférieure à 70  $\mu\text{M}$  si la forme liée de l’Arg17 est visible,
- les composés ayant une constante de dissociation comprise entre 70  $\mu\text{M}$  et 3 mM, si des perturbations sont détectées mais si la forme liée de l’Arg17 n’est pas présente.



**Figure 4.11** : Suivi des interactions entre Pin1 et le dipeptide phosphoThr-Pro (projections 1D des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC). La liaison du ligand est suivie par ajouts successifs de quantités croissantes de ligand, dont les concentrations initiales sont indiquées à gauche des spectres, sur une solution de WW uniformément enrichi en  $^{15}\text{N}$  à une concentration constante de 20  $\mu\text{M}$ . L’apparition d’un pic à 11.2 ppm correspondant à l’Arg17 indique une stabilisation de la boucle  $\beta$ -hairpin par liaison au substrat.

#### ***Détermination d’une constante de dissociation par RMN.***

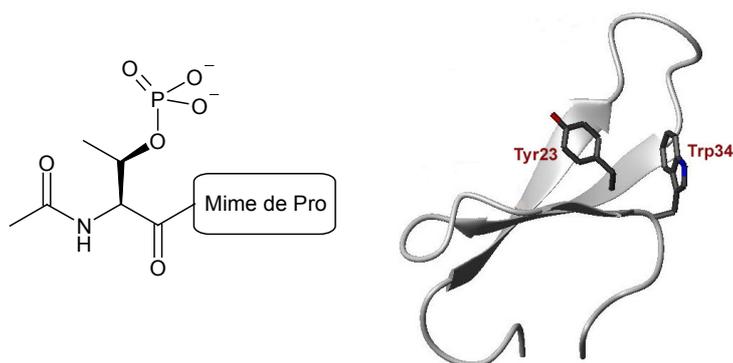
La constante de dissociation d’un complexe formé entre le domaine WW et un ligand peut être évaluée par RMN de la manière décrite dans le chapitre II avec l’acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC pour différents rapports molaires ligand : protéine et faisant intervenir la variation de déplacement chimique des résonances de résidus impliqués dans la liaison du substrat en fonction de la concentration de ligand ajouté. Cette technique de détermination de la constante de dissociation des complexes par RMN présente, en outre, l’avantage de définir le mode de liaison des ligands par *mapping* des déplacements chimiques.

## II – Synthèse de ligands du domaine WW basés sur la structure du motif phosphoThr-Pro et criblage par RMN.

La connaissance d’un ligand de faible affinité représenté par le dipeptide phosphoThr-Pro ainsi que la structure du domaine WW en complexe avec un peptide phosphorylé, nous a permis de débiter la recherche de ligands du domaine WW possédant une meilleure affinité (*hits* pharmacologiques) par modification de la structure du ligand de départ. Nous avons dans un premier temps recherché des mimes du dipeptide phosphoThr-Pro en mimant la proline et le phosphate séparément. Dans un deuxième temps, nous avons, grâce à la structure du domaine WW de Pin1, recherché de nouveaux points d’interaction permettant, par ajout de groupements supplémentaires, d’augmenter l’affinité des ligands.

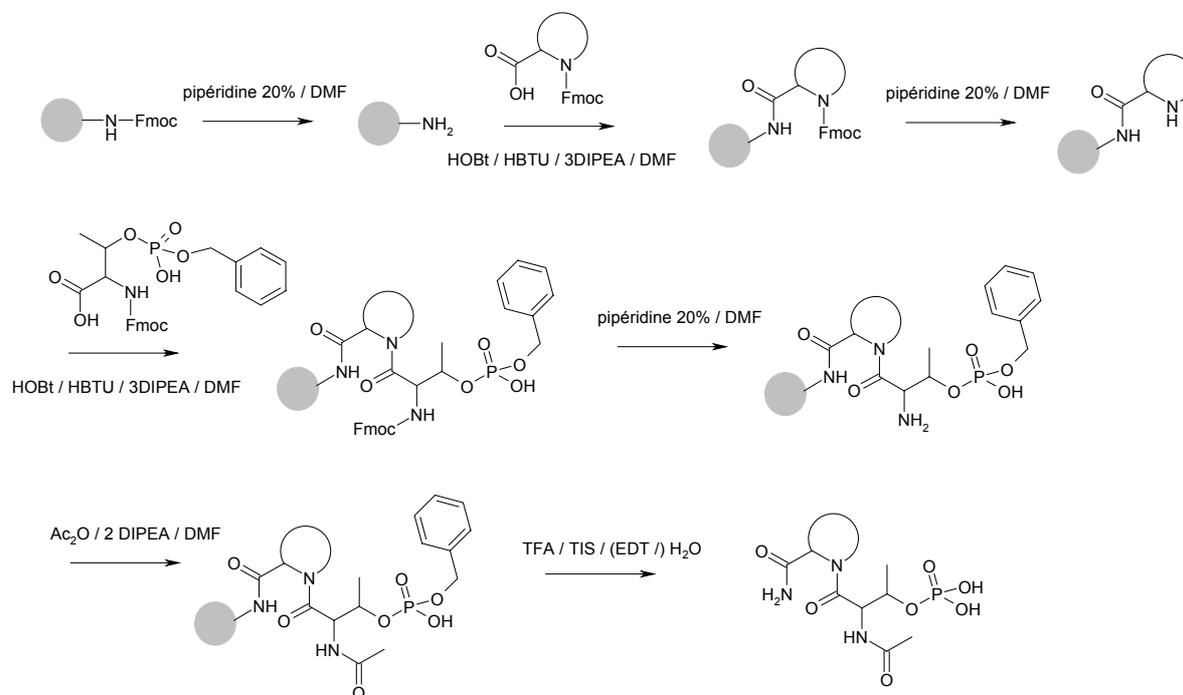
### 2.1 – Mimes de proline.

La proline lie le domaine WW au niveau d’une poche hydrophobe constituée des chaînes latérales des résidus Tyr23 et Trp34 qui forment une pince aromatique (Figure 4.12). Cette poche est relativement petite et exposée au solvant. Nous avons donc exploré la taille de la poche en faisant varier la taille du cycle pyrrolidine de la proline, en introduisant des substituants endocycliques et exocycliques, ou en utilisant des cycles plus rigides. Les composés synthétisés pour l’évaluation des mimes de proline ont pour base structurale le motif décrit **Figure 4.12**, correspondant au couplage de la phosphothréonine via la fonction acide sur la fonction amine du mime de proline et acétylation de la fonction amine de la phosphothréonine.



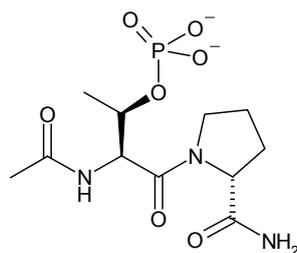
**Figure 4.12 :** Structure de base des molécules impliquées dans la recherche de mimes de proline et structure tridimensionnelle du domaine WW montrant les résidus Trp34 et Tyr23 de la poche aromatique impliquée dans la liaison du résidu proline.

La synthèse supportée mise en œuvre pour l'obtention de ces composés (Figure 4.13) permettra ensuite une rapide diversification du (des) meilleur(s) analogue(s) de proline (*hit explosion*).



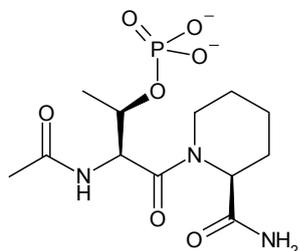
**Figure 4.13 :** Schéma général de synthèse supportée pour l'obtention des dérivés (2) à (9) excepté le composé (6) pour lequel, le couplage de la phosphothréonine met en œuvre le chlorure d'oxalyle pour l'activation de la fonction acide de la phosphothréonine.

Le choix des mimes de proline s'est porté dans un premier temps sur des molécules simples, proches du point de vue structural de la proline. On peut ainsi évaluer les éléments structuraux qui influencent la liaison au domaine WW de Pin1. Pour conserver une affinité de liaison mesurable, le second élément entrant dans l'interaction, le groupement phosphate, est conservé dans cette série. Une première modification de la stéréochimie du cycle pyrrolidinique conduit à la molécule (2).

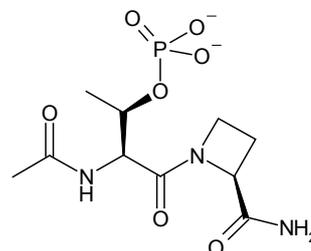


(2)

Nous avons également sélectionné dans une deuxième série d’analogues de la proline, l’homologue supérieur (pipéridine) et l’homologue inférieur (azétidine) de la pyrrolidine aboutissant, après couplage à la phosphoThr, aux molécules **(3)** et **(4)** respectivement.

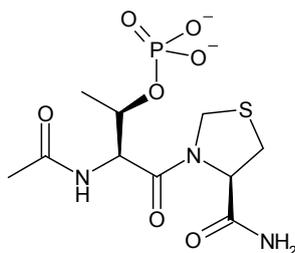


**(3)**



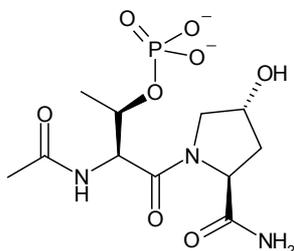
**(4)**

Dans une troisième série d’analogues, nous explorons les possibilités de substitution du cycle pyrrolidine. Le composé **(5)** est obtenu par remplacement isostérique du méthylène 4 par un atome de soufre.

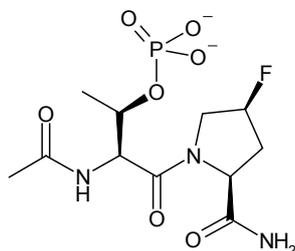


**(5)**

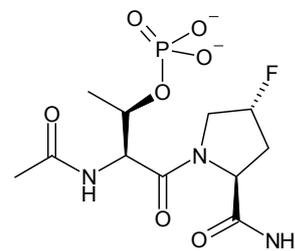
Le composé **(6)** est substitué en position 4 du cycle pyrrolidinique par un groupement hydroxyle en *trans* et les 4-*cis*- et 4-*trans*-fluoro-(L)-proline sont introduites en remplacement de la (L)-proline, conduisant par couplage à la phosphothréonine aux composés **(7)** et **(8)** respectivement.



**(6)**

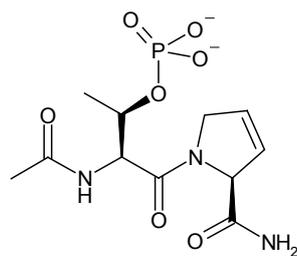


**(7)**



**(8)**

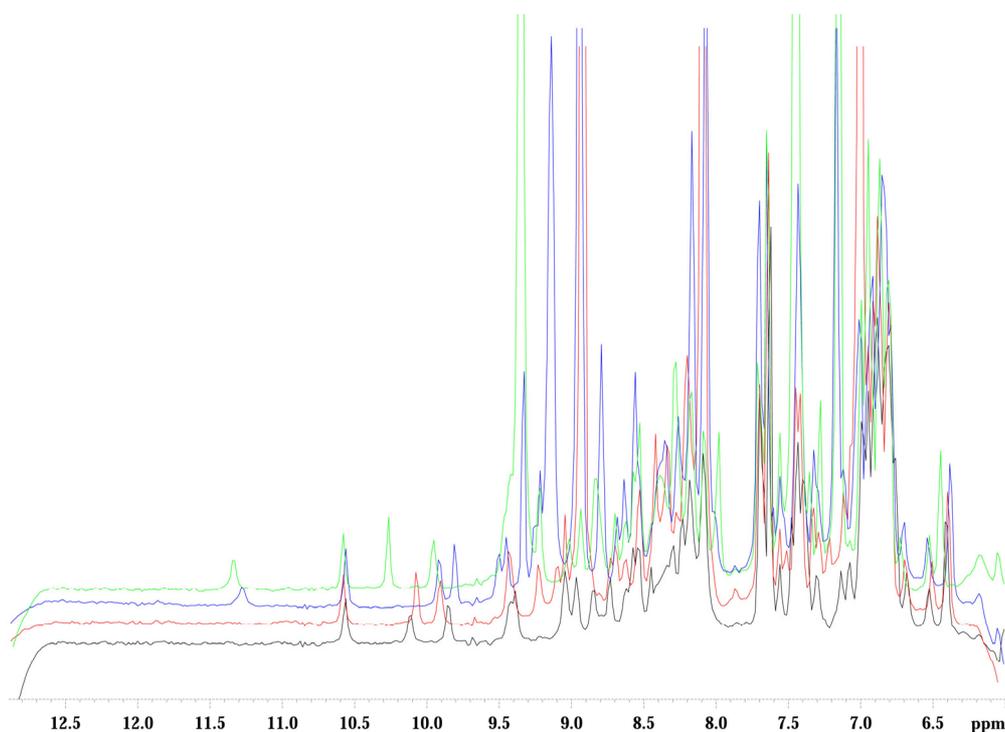
Dans une quatrième série, le couplage de la 3,4-dehydro-(L)-proline à la phosphothréonine, conduit au composé **(9)**.



(9)

Les composés (2) à (9) ont été testés par RMN à haute résolution et les constantes de dissociation des complexes formés avec le domaine WW ont été calculées comme décrit pour le composé-référence (1).

La **Figure 4.14** montre un exemple de criblage des composés (2), (4) et (6) par RMN  $^1\text{H}$  réalisé sur des solutions contenant 60  $\mu\text{M}$  de WW non enrichi et 1.2 mM de ligand. On observe pour les composés (2) et (4) un signal à 11.3 ppm correspondant à l'Arg17 indiquant que l'état de saturation de la protéine est presque atteint avec 20 équiv. de ligand. Dans le cas du composé (6), la saturation n'est pas atteinte mais les perturbations des déplacements chimiques de la protéine indiquent une liaison du composé au domaine WW.



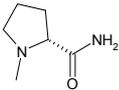
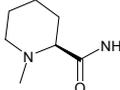
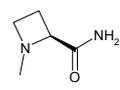
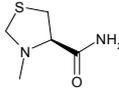
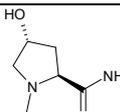
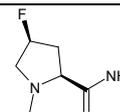
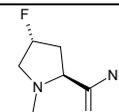
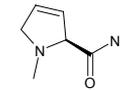
**Figure 4.14** : Détection de la liaison de trois composés constitués des mimes de Pro de type azétidine (bleu), 4-hydroxy-Pro (rouge) et (D)-Pro (vert) par RMN par comparaison avec le spectre du domaine WW seul (noir) à une concentration de 60  $\mu\text{M}$ . Les ligands sont présents à une concentration de 1.2 mM soit 20 équiv. par rapport à la protéine. Les spectres  $^1\text{H}$  sont réalisés sur un spectromètre à 600 MHz équipé d'une cryosonde.

La RMN à haute résolution a permis, outre la détermination des constantes de dissociation, de déterminer le mode de liaison des ligands par *mapping* des déplacements chimiques des résidus du domaine WW en l’absence de ligand et en présence de 20 équiv. de ligand. On peut ainsi déterminer les régions du domaine affectées par le ligand et définir le site de liaison. Les expériences de titration permettent donc :

- d’obtenir des informations quantitatives sur les ligands avec la constante de dissociation et la détermination des groupements pharmacophoriques influençant la liaison au domaine WW, c’est-à-dire les relations structure/affinité (SAR).
- d’obtenir des informations qualitatives sur le mode de liaison du ligand et le site de liaison du domaine WW.

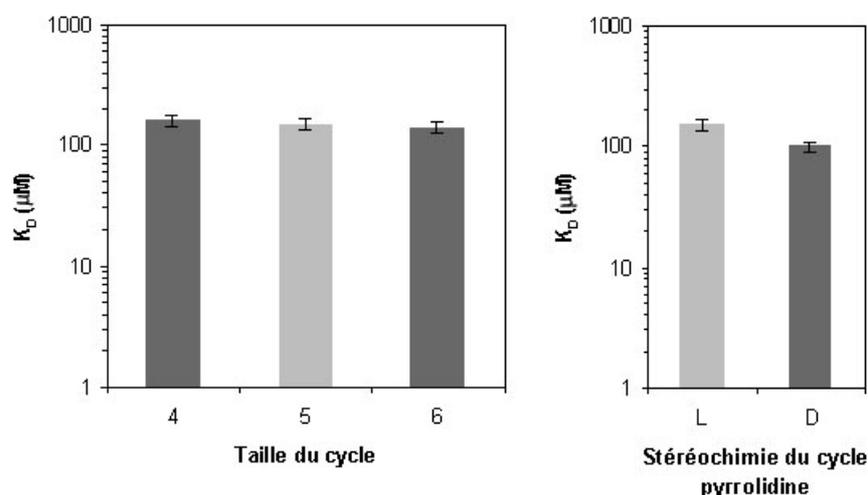
### ***Bilan des constantes de dissociation et relations structure-affinité (SAR) pour les mimes de proline.***

Les cycles homologues de la proline constituant les composés **(3)** et **(4)**, ainsi que le stéréoisomère de la proline dans le composé **(2)** ont une affinité similaire à celle du composé référence **(1)** dont la constante de dissociation a été évaluée à 150  $\mu\text{M}$ . Un remplacement isostérique réalisé avec l’analogue de type (L)-thiaproline (composé **(5)**), dont le  $K_D$  a été déterminé à 400  $\mu\text{M}$ , influence davantage la liaison au domaine WW mais avec un effet négatif. L’effet d’une substitution peut avoir une influence plus ou moins marquée selon l’encombrement stérique du substituant et selon la position de ce substituant par rapport au plan moyen du cycle. Dans le cas de la 4-*trans*-hydroxy-(L)-proline (composé **(6)**), la diminution de l’affinité est relativement importante avec un  $K_D$  de 800  $\mu\text{M}$ . La substitution d’un atome d’hydrogène par un atome de fluor montre un impact différent sur la liaison au domaine WW en fonction de la position du fluor par rapport au cycle. Dans le cas de la 4-*cis*-fluoro-proline (composé **(7)**), on observe une légère diminution de l’affinité (400  $\mu\text{M}$ ) du même ordre que le remplacement isostérique dans le composé **(5)**. En revanche, pour la 4-*trans*-fluoro-proline (composé **(8)**), la constante de dissociation de 40  $\mu\text{M}$  représente une augmentation d’affinité significative par rapport au composé-référence **(1)** et constitue un gain d’un facteur 10 par rapport au diastéréoisomère **(7)**. Le mime le plus remarquable en terme d’affinité est la 3,4-dehydro-proline (composé **(9)**). La gain en affinité pour la liaison au domaine WW s’élève à un facteur 7 par rapport au composé-référence **(1)** d’après la constante de dissociation évaluée à 20  $\mu\text{M}$  (Tableau 4.1).

	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>Mimes de Pro</b>								
<b>K<sub>D</sub> (μM)</b>	100	140	160	400	800	400	40	20

**Tableau 4.1 :** Constantes de dissociation des mimes de proline couplés à la phosphothrénine basés sur une analogie structurale avec le composé-référence (1).

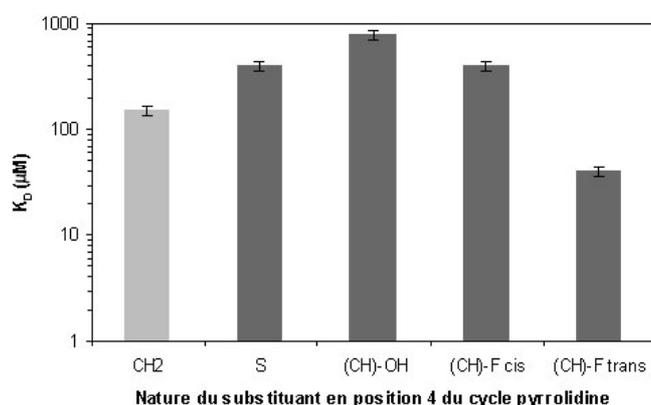
Cycles homologues de la pyrrolidine et énantiomère : Une variation de la taille du cycle et de la stéréochimie, dans la mesure où il n'y a pas de substituant sur le carbonyle de la proline, ne perturbe pas de façon significative l'interaction avec le domaine WW (Graphique 4.2).



**Graphique 4.2 :** Représentation de l'influence de la taille du cycle mimant la proline et de la stéréochimie du cycle pyrrolidinique dans la liaison de la pince aromatique Tyr23-Trp34 du domaine WW. Le composé de référence est indiqué en gris clair

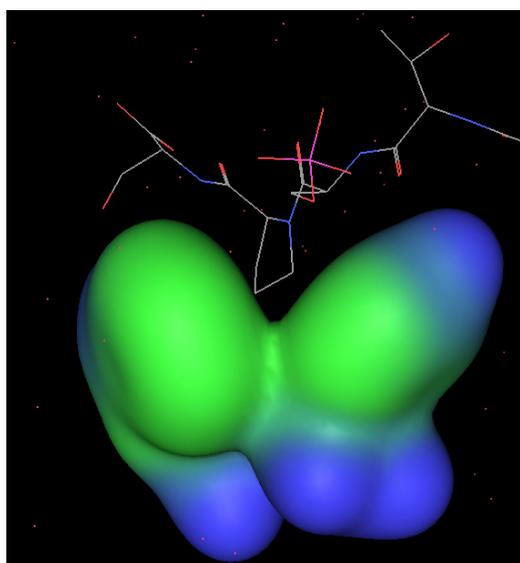
Substitution du cycle pyrrolidine et isostérie en position 4 : L'effet d'une substitution sur le cycle pyrrolidinique au niveau du méthylène 4 est plus complexe (Graphique 4.3). Son influence dépend de plusieurs facteurs : le caractère endo- ou exo-cyclique, les hétéroatomes impliqués et la stéréochimie. Les modifications étudiées sont l'introduction d'un soufre en position 4, la substitution d'un atome d'hydrogène du méthylène 4 par un groupement hydroxyle en position *trans* par rapport au carboxyle de la proline ou un fluor en position *cis* ou *trans*. Deux paramètres atomiques peuvent expliquer les différences observées en terme d'affinité pour le domaine WW : la taille des atomes et leur position par rapport au plan

moyen du cycle pyrrolidinique. Les atomes d’hydrogène et de fluor ont des tailles comparables (leurs rayons de Van der Waals étant de 1,2 et 1,47 Å respectivement) alors que les atomes de soufre et d’oxygène sont plus gros. L’effet d’encombrement stérique est beaucoup plus important pour le substituant hydroxyle que pour le fluor, de même le remplacement d’un méthylène par un soufre induit une gêne stérique plus importante par rapport à la poche aromatique du domaine WW qui lie la proline. D’autre part, la stéréochimie d’un substituant exo-cyclique joue un rôle non négligeable car le fluor en *cis* est très défavorable par rapport au fluor en position *trans*.



**Graphique 4.3 :** Représentation de l’influence d’une substitution du méthylène 4 du cycle pyrrolidinique de la (L)-proline.

Il faut noter à ce propos que le méthylène en position 4 est celui qui est dans le « creux » de la pince aromatique constituée des chaînes latérales des résidus Trp34 et Tyr23 donc probablement le plus sensible à l’encombrement stérique.

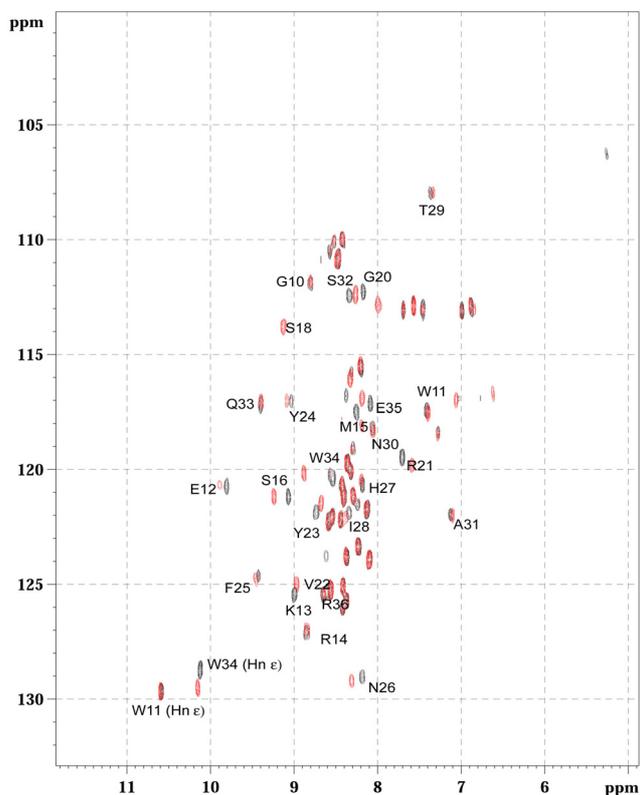


**Figure 4.15 :** Surface hydrophobe de la poche aromatique du domaine WW de Pin1 impliquée dans la liaison d’un résidu Pro.

La 3,4-dehydro-proline augmente l’affinité d’un facteur 7 par rapport à la proline qui pourrait s’expliquer par un encombrement stérique moins important au niveau des atomes situés au creux de la pince aromatique.

**Mode de liaison des ligands par mapping des déplacements chimiques.**

Un *mapping* des déplacements chimiques des résonances du domaine WW est réalisé pour les composés **(8)** et **(9)**, afin de déterminer si l’augmentation de l’affinité par rapport au composé-référence **(1)** est due à un mode de liaison différent du ligand.

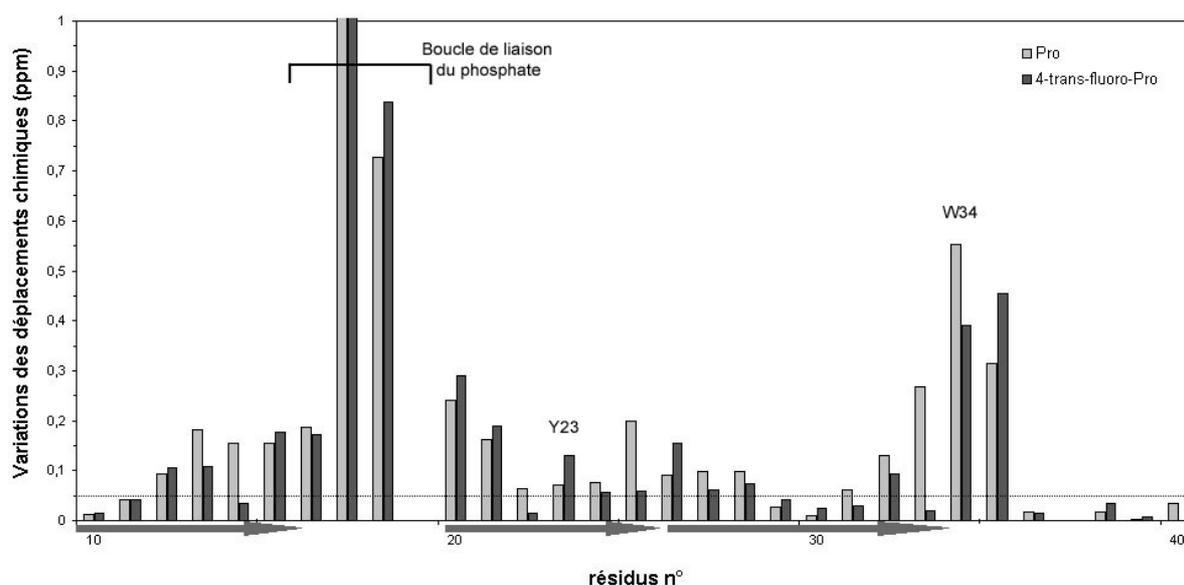


**Figure 4.16 :** Perturbations des résidus du domaine WW en présence de 20 équiv. du composé **(8)** (spectre rouge) par rapport à l’apo-WW (spectre noir).

Globalement, le spectre  $^1\text{H}-^{15}\text{N}$  HSQC du domaine WW semble affecté de manière quasi-identique pour les ligands **(2)** à **(9)** par rapport au dipeptide **(1)**. Les variations induites par les différents ligands sont faibles et souvent difficiles à expliquer en terme de modification des interactions avec le domaine WW en fonction de la nature du ligand. La nature des groupements chimiques impliqués peut influencer le déplacement chimique de la forme liée du domaine WW sans pour autant expliquer les variations d’affinité et définir clairement le mode de liaison du ligand. De plus, même si le seul élément structural qui varie d’un ligand à l’autre dans cette série est le mime de proline, la nature même de ce motif et son positionnement dans le site de liaison peut influencer le positionnement du phosphate. Ainsi,

on observe pour les composés **(8)** et **(9)** des variations plus marquées des résidus de la boucle de liaison du phosphate qui peuvent traduire un meilleur positionnement de ce dernier par rapport à la boucle  $\beta$ -hairpin du domaine WW. D’autre part, un réarrangement conformationnel du domaine WW peut être à l’origine des variations de déplacement chimique observées d’un ligand à l’autre. Le mapping des déplacements chimiques du domaine WW ne permet donc pas de définir clairement si les perturbations observées sont relatives à la nature du ligand lui-même ou à un réarrangement structural induit par la liaison du ligand ou les deux.

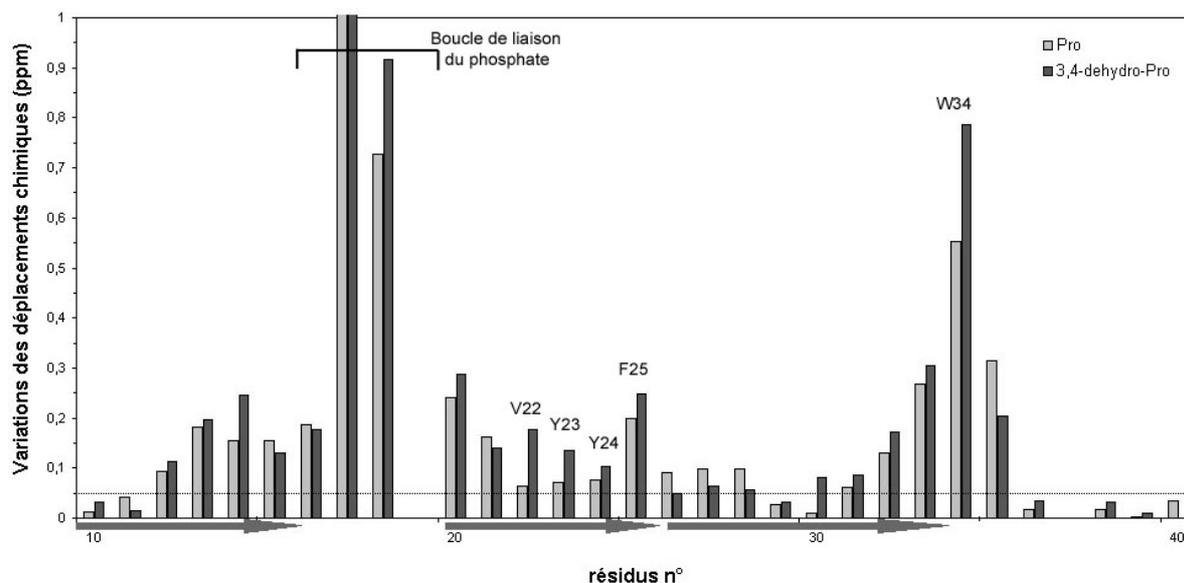
La 4-*trans*-fluoro-proline ne modifie pas de manière notable les déplacements chimiques des résidus Phe25 et Gln33 alors que le dipeptide **(1)** induisait des perturbations significatives. L’atome de fluor pourrait être mis en cause directement dans ce phénomène, notamment par son caractère électro attracteur, ou par un positionnement différent du cycle pyrrolidinique dans la pince aromatique du domaine.



**Figure 4.17 :** Mapping des déplacements chimiques des résidus du domaine WW en présence du composé **(8)** comparé avec le mapping du composé-référence **(1)**. Les flèches indiquent les feuilletts  $\beta$ .

La 3,4-dehydro-proline induit des perturbations plus marquées pour les résidus Val22, Tyr23, Tyr24 et Trp34 qui peuvent être induites par le cycle insaturé lui-même ou traduire un réarrangement structural de la poche aromatique autour de la 3,4-dehydro-proline.

Enfin, il faut noter que le *mapping* ne rend pas compte de l’ensemble des phénomènes impliqués dans la liaison du ligand, en particulier de la dynamique du processus de liaison. Par exemple, on a noté pour le composé **(9)** un élargissement plus important des résonances du domaine WW, en particulier pour les résidus Trp34 et Arg17, témoignant d’une dynamique différente de l’échange entre le domaine WW et le ligand comme une vitesse de dissociation plus lente du complexe.



**Figure 4.18 :** *Mapping* des déplacements chimiques des résidus du domaine WW en présence du composé **(9)** comparé avec le *mapping* du composé-référence **(1)**.

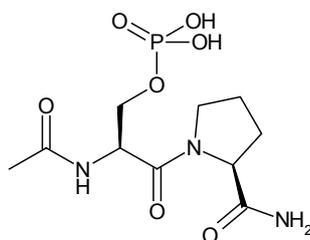
## 2.2 – Analogues de la phosphothréonine.

Les analogues de la phosphothréonine concernent des analogues du phosphate sous forme de têtes polaires comportant un hydrogène acide, mais introduisent également des variations de la chaîne carbonée qui porte la tête polaire.

### 2.2.1 – Etude préliminaire.

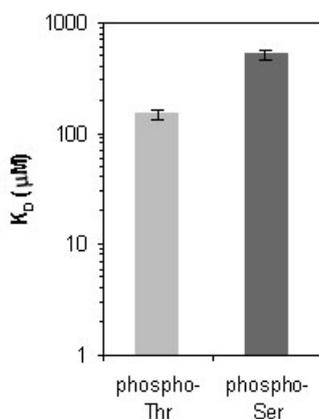
Nous avons exploré les conditions d’interaction du groupement phosphate avec la boucle  $\beta$ -hairpin du domaine WW en envisageant des têtes polaires mimant le groupement phosphate et des modifications structurales de la thréonine. Une étude préliminaire a été réalisée pour évaluer l’affinité du dipeptide phosphoSer-Pro **(10)**. D’après les études de *binding* réalisées par RMN sur des peptides phosphorylés de la protéine tau dans le chapitre II, les peptides contenant un résidu phosphoSer présentaient une affinité plus faible que les peptides

contenant un résidu phosphoThr. Afin de déterminer si cette différence d'affinité est due à la sérine elle-même ou à la séquence primaire des peptides, nous avons déterminé l'affinité du composé **(10)** pour le domaine WW de Pin1. De plus, les analogues de phosphothréonine qui sont présentés dans ce paragraphe sont obtenus par formation d'une tête polaire sur une chaîne carbonée non ramifiée, dérivée de l'acide propionique, relevant plutôt d'une analogie avec la sérine.



**(10)**

Une constante de dissociation de 510  $\mu\text{M}$  a été mesurée pour le composé **(10)**. On observe par conséquent une variation significative de l'affinité entre les résidus phosphoThr et phosphoSer indiquant un rôle positif du méthyle du résidu Thr dans la liaison du domaine WW soit par liaison hydrophobe avec un résidu du domaine WW, soit par expulsion d'une molécule d'eau dans un site hydrophobe ou par rigidification de la chaîne latérale de l'acide aminé.

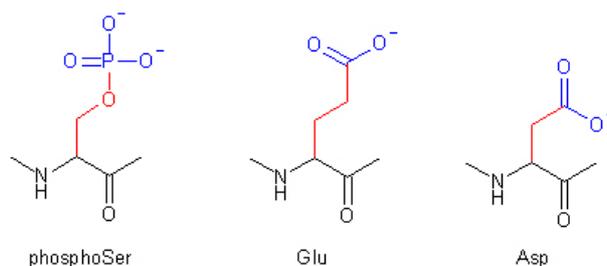


**Figure 4.19 :** Représentation des constantes de dissociation comparées des dipeptides phosphoThr/Ser-Pro (composés **(1)** et **(10)**).

### 2.2.2 – Analogues de type acide carboxylique : acide glutamique et dérivés.

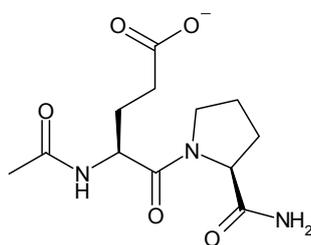
Des mutations des acides aminés sérine ou thréonine par des glutamates sont souvent employées dans les protéines pour mimer la phosphorylation de ces mêmes résidus (Maciejewski, Peterson et al. 1995; Leger, Kempf et al. 1997; Eidenmuller, Fath et al. 2000).

Ces mutations sont basées sur l’analogie structurale entre la sérine dans son état phosphorylé et l’acide glutamique (Figure 4.20). Bien que les charges et la délocalisation électronique soient différentes (deux charges négatives délocalisées sur trois atomes pour le phosphate, une charge négative délocalisée sur deux atomes pour le carboxylate), la longueur de chaîne et la présence d’une charge négative font de l’acide glutamique un mime facile à mettre en œuvre par simple mutation dans la séquence primaire de la protéine, stratégie moins coûteuse et plus simple que la phosphorylation.



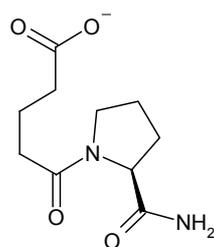
**Figure 4.20** : Analogie structurale entre la phosphosérine, l’acide glutamique et l’acide aspartique. La longueur de la chaîne carbonée portant la tête polaire est plus proche de la phosphosérine dans l’acide glutamique que dans l’acide aspartique.

Une série de molécules basées sur le couplage de l’acide glutamique ou d’analogues structuraux de l’acide glutamique à la proline a été obtenue par synthèse en phase solide. La première étape de couplage de la proline permet l’obtention d’un lot de résine qui après fractionnement peut subir une étape de diversification. Le couplage de l’acide glutamique protégé par un groupement tBu sur la fonction acide de la chaîne latérale conduit après acétylation de la fonction amine et coupure au TFA au composé référence **(11)**.

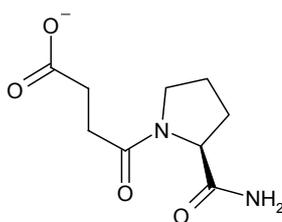


**(11)**

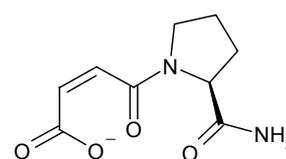
Le couplage à la proline de l’anhydride glutarique, de l’anhydride succinique et de l’anhydride maléique conduit respectivement aux composés **(12)**, **(13)** et **(14)** après une étape de coupure au TFA.



(12)

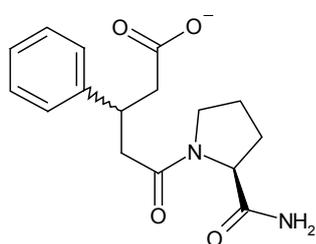


(13)

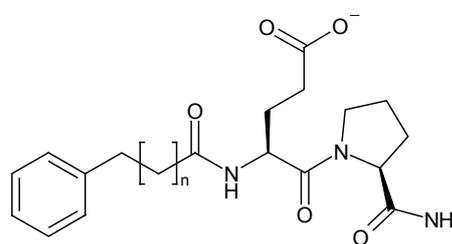


(14)

Le composé (15) possède comme précurseur l'anhydride 3-phénylglutarique pour lequel le cycle benzénique substitue un proton méthylénique en position 3 de l'acide glutarique.



(15)



(16) n = 1

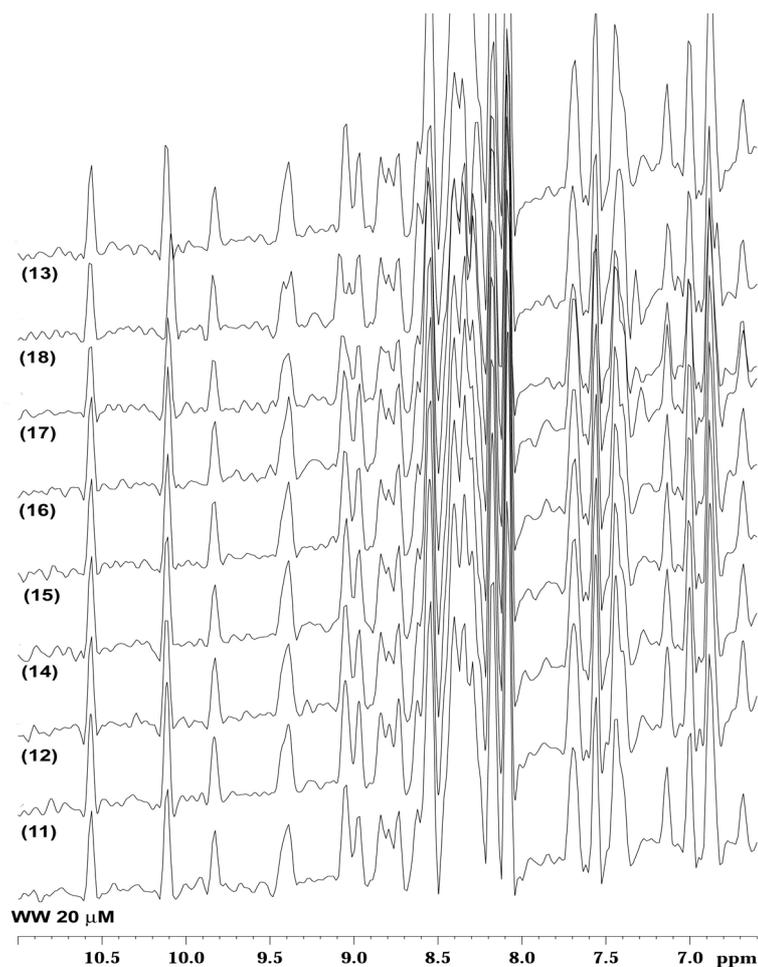
(17) n = 2

(18) n = 3

La structure du domaine WW en complexe avec un substrat peptidique montre une proline face au cycle benzénique du résidu Phe25 (Figure 4.1). Nous avons envisagé la possibilité d'une nouvelle interaction entre le ligand et le domaine WW au niveau du résidu Phe25 via un groupe aromatique. Un cycle benzénique est introduit sur le dipeptide Glu-Pro par l'intermédiaire d'une chaîne aliphatique de longueur variable, afin d'explorer le positionnement optimal du cycle benzénique par rapport au résidu Phe25 du domaine WW lors de la liaison. Pour les composés (16) à (18), les acides 3-phénylpropionique, 4-phénylbutyrique et 5-phénylvalérique sont couplés sur la fonction amine de l'acide glutamique du dipeptide Glu-Pro, après activation de l'acide par un mélange équimolaire HOBt/HBTU (1 équivalent par rapport à l'acide) et 3 équivalents de DIPEA.

Le criblage pour les ligands (11) à (16) n'a montré aucune interaction entre les ligands et le domaine WW mais on note une liaison faible pour les ligands (17) et (18), caractérisée par une constante de dissociation supérieure à 5 mM (Figure 4.21). Les affinités très faibles (supérieures à 3 mM) obtenues pour les composés (11) à (18) montrent qu'un acide carboxylique ne peut pas remplacer le phosphate dans la liaison au domaine WW, mais les

composés (17) et (18) offrent de nouvelles perspectives concernant l’ajout de ballasts hydrophobes permettant une interaction additionnelle avec le domaine WW. Nous avons envisagé qu’il puisse y avoir une interaction hydrophobe supplémentaire qu’à la condition que l’interaction principale centrée autour du dipeptide phosphoThr-Pro soit conservée. Les mêmes groupements hydrophobes seront impliqués par la suite dans l’étude d’une interaction hydrophobe additionnelle avec des composés où le motif d’interaction de base est conservé (voir paragraphe 2.3.1).

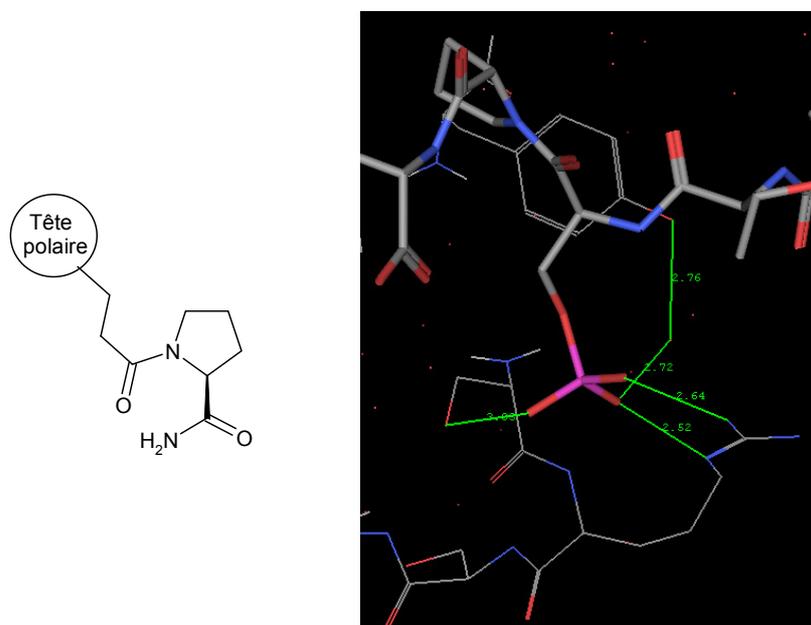


**Figure 4.21 :** Criblage des composés (11) à (18) par RMN sur un spectromètre à 600 MHz équipé d’une cryosonde. Le domaine WW enrichi en  $^{15}\text{N}$  est mis en solution à une concentration de 20  $\mu\text{M}$  dans un tampon Tris deutéré à 25 mM pH 6.80, NaCl 50 mM, DTT 2.5 mM. Les différents ligands sont présents à une concentration de 400  $\mu\text{M}$  (soit 20 équiv.).

### 2.2.3 – Autres analogues de phosphate : diversité structurale de têtes polaires.

Les composés synthétisés pour l’évaluation des analogues du phosphate ont pour base le motif décrit **Figure 4.22**, correspondant au couplage d’un acide dérivé de l’acide propionique

fonctionnalisé en position 3 pour permettre l’obtention d’une tête polaire et dont la fonction acide carboxylique est couplée à la fonction amine de la proline.



**Figure 4.22 :** Structure de base des composés impliqués dans la recherche de têtes polaires mimant le groupement phosphate et mode de liaison du groupement phosphate par interaction avec les résidus Ser16 et Arg17 du domaine WW.

Les têtes polaires sont des motifs comportant au moins un hydrogène acide, le plus proche du phosphate étant le phosphonate, présent dans le composé (19). L’acide carboxylique dans le composé (13) ou l’ $\alpha$ -diacide dans le composé (20) sont des analogues couramment utilisés comme nous l’avons décrit dans le paragraphe précédent.

	(13)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
<i>Analogue de phosphate</i>								
$K_D$ ( $\mu M$ )	>5000	>5000	>5000	>5000	>5000	>5000	>5000	>5000

**Tableau 4.2 :** Constantes de dissociation des têtes polaires mimant le groupement phosphate.

D’autres analogues plus distants structurellement sont basés sur des motifs comportant un hydrogène acide comme le tétrazole constituant le composé (21), le squaramide (22), le trifluorosulfonamide (23). La 3,4-difluoroaniline (24) et la 3-nitroaniline (25) quant à elles, pourraient interagir via leur doublets non liants.

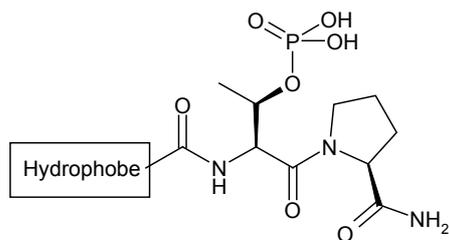
Tous les composés constitués de mimes de phosphate ont montré des affinités très faibles de plus de 5 mM, c’est-à-dire en dehors de la gamme mesurable par RMN, indiquant une extrême spécificité du domaine WW pour le phosphate et une adaptabilité très limitée de la boucle liant le phosphate.

### **2.3 – Introduction de motifs supplémentaires d’interaction avec le domaine WW.**

La perte drastique d’affinité pour les ligands dont la phosphothréonine a été modifiée à la fois au niveau de sa chaîne carbonée (longueur, ramification, substitution...) et de sa tête polaire, ainsi que l’influence des modifications apportées au niveau du cycle pyrrolidinique de la proline le plus souvent en défaveur de la liaison du ligand au domaine WW nous a conduits à envisager de nouveaux points d’interaction entre les ligands et le domaine WW. Une interaction supplémentaire de type hydrophobe avait été étudiée à travers les dipeptides Glu-Pro auxquels avaient été ajoutées des chaînes aliphatiques de longueurs variées portant un groupement phényle. Malgré l’ajout de ces ballasts hydrophobes, l’affinité n’avait pu être déterminée ( $K_D > 3$  mM). Cependant, il faut noter que le dipeptide Glu-Pro (composé **(11)**) avait une affinité vraisemblablement encore plus faible. Nous avons donc poursuivi l’étude de l’influence de groupements hydrophobes sur la liaison des composés au domaine WW en conservant le motif d’interaction de base, le dipeptide phosphoThr-Pro.

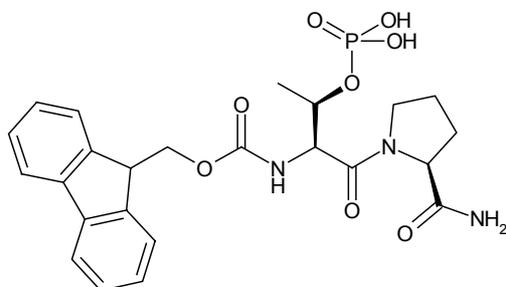
#### **2.3.1 – Introduction d’un motif hydrophobe en position N-terminale.**

Une nouvelle série de composés possédant un nouveau point d’interaction potentiel avec le domaine WW de Pin1 a été synthétisée. Le motif de base phosphoThr-Pro est conservé pour situer les composés dans une gamme d’affinité mesurable et permettre la détection de variations positives ou négatives des affinités. Un groupement hydrophobe est ajouté via une fonction acide par couplage à la fonction amine primaire de la phosphothréonine aboutissant à une série de composés dont la structure générique est décrite dans la **Figure 4.23**. Dans un premier temps, nous avons conservé le groupement protecteur Fmoc de l’amine primaire (composé **(26)**) ou couplé, après déprotection de l’amine, des synthons de type acide carboxylique à chaînes aliphatiques de longueurs variables portant un groupement phényle, sur base du dipeptide phosphoThr-Pro (composés **(28)** à **(31)**) et sur base du dipeptide phosphoSer-Pro (composé **(32)**).

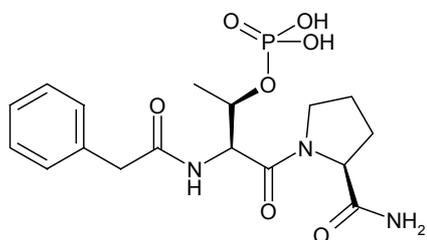


**Figure 4.23** : Structure de base des composés impliqués dans l'étude de l'effet d'un groupe hydrophobe sur la liaisons des composés au domaine WW de Pin1.

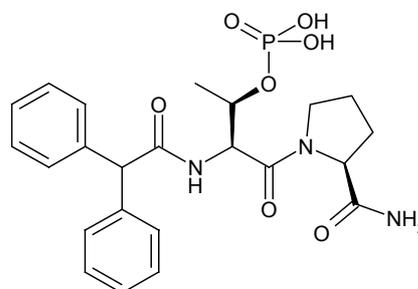
La stratégie de synthèse en phase solide est celle décrite pour l'obtention des composés **(1)** à **(9)**. Le couplage des acides précurseurs de type 2-phénylacétique, 2,2-bisphénylacétique, 3-phénylpropionique, 4-phénylbutyrique et 5-phénylvalérique est réalisé comme décrit pour les composés **(16)** à **(18)**.



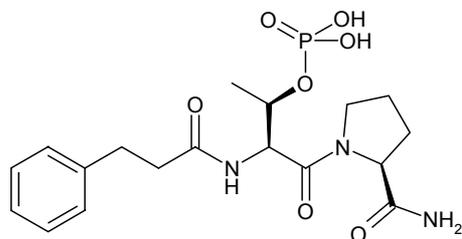
**(26)**



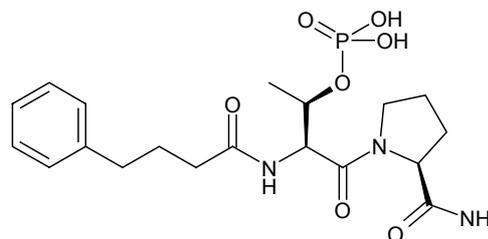
**(27)**



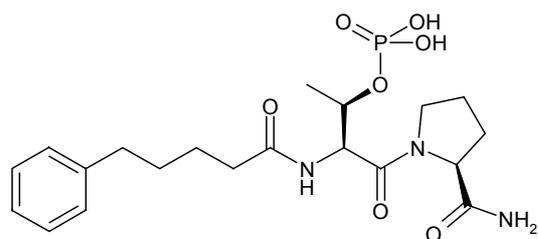
**(28)**



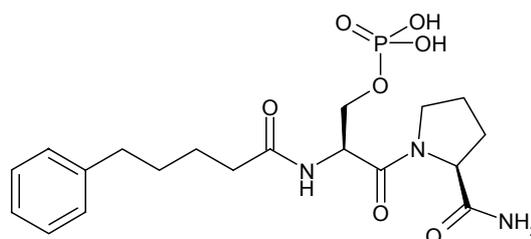
**(29)**



**(30)**



(31)



(32)

**Bilan des constantes de dissociation : amélioration de l'affinité des ligands par ajout d'un motif hydrophobe supplémentaire.**

L'addition d'un ballast hydrophobe sur le dipeptide phosphoThr-Pro augmente de manière significative l'affinité des composés résultants pour le domaine WW de Pin1. Le groupement Fmoc dans le composé (26) permet d'atteindre une constante de dissociation de 15  $\mu\text{M}$  soit un gain d'un facteur 10 en affinité par rapport au composé acétylé (1). Les chaînes aliphatiques portant un groupement phényle dans les composés (27) à (31) permettent d'aboutir à des constantes de dissociation comprises entre 30 et 41  $\mu\text{M}$  (Tableau 4.3).

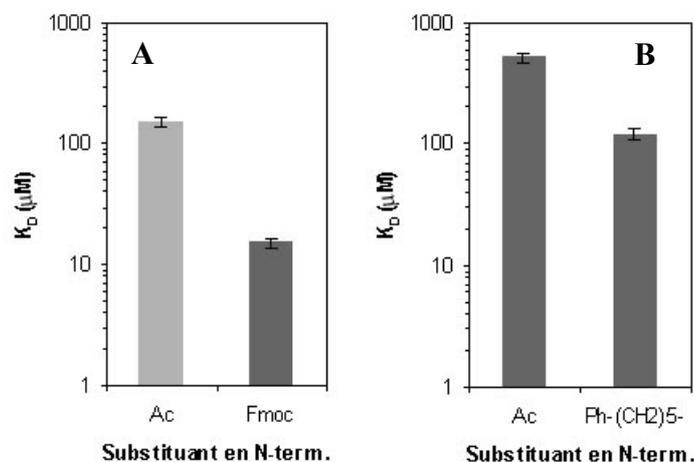
	(26)	(27)	(28)	(29)	(30)	(31)	(32)
<b>Groupement hydrophobe</b>							
<b><math>K_D</math> (<math>\mu\text{M}</math>)</b>	15	35	38	32	30	41	120

**Tableau 4.3 :** Constantes de dissociation des complexes entre le domaine WW et les composés (26) à (32).

**Relations structure-affinité concernant les groupements hydrophobes.**

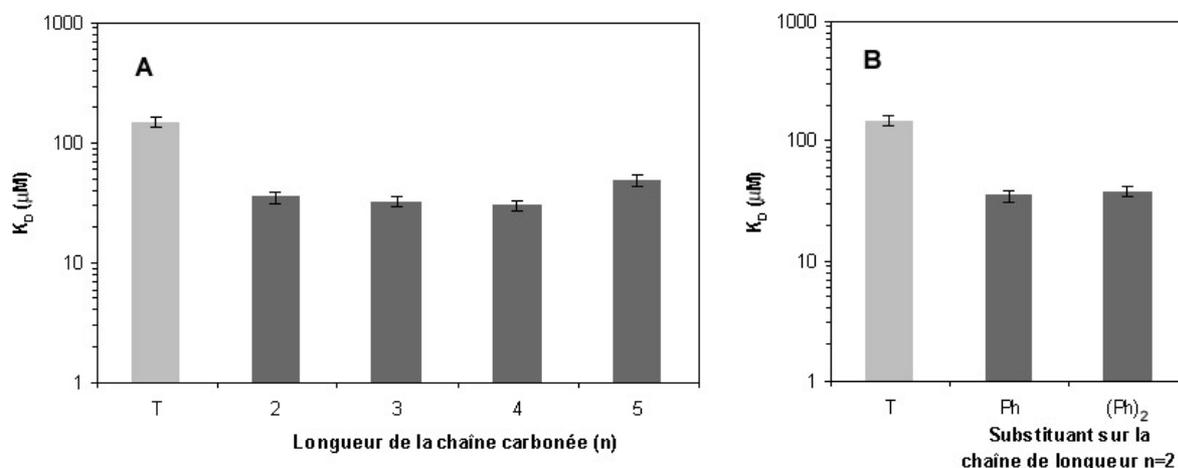
Le groupement Fmoc conduit au plus important gain d'affinité parmi les six ballasts hydrophobes testés (Graphique 4.4A). La longueur de la chaîne aliphatique portant le phényle influence peu l'affinité des composés pour le domaine WW. L'introduction d'un substituant 5-phénylvaléroyle en position N-terminale sur le dipeptide phosphoSer-Pro permet d'augmenter l'affinité d'un facteur 4 par rapport au dipeptide phosphoSer-Pro acétylé (10) (Graphique 4.4B). Les composés (27), (29)-(31) sont caractérisés par des constantes de dissociation comprises entre 30 et 40  $\mu\text{M}$  soit une perte d'un facteur 2 à 3 par rapport au groupement Fmoc. Même si les variations d'affinité en fonction de la longueur de la chaîne

aliphatique sont faibles, le composé le plus affiné est obtenu avec la chaîne dérivée de l’acide 4-phényl-butyrique (Graphique 4.5A), soit une chaîne à 4 carbones, analogue au groupement Fmoc.



**Graphique 4.4 :** Représentation des  $K_D$  en fonction de la nature du substituant en position N-terminale, pour les ligands contenant le dipeptide phosphoThr-Pro (A) ou le dipeptide phosphoSer-Pro (B).

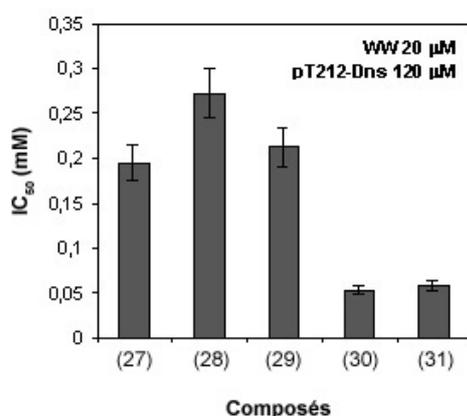
L’introduction d’un deuxième substituant phényle (noté Ph<sub>2</sub>) sur le dérivé de l’acide 2-phénylacétique (noté Ph) ne modifie pas significativement l’affinité (Graphique 4.5B).



**Graphique 4.5 :** Représentation des constantes de dissociation des ligands phosphoThr-Pro en fonction de la longueur de la chaîne carbonée en position N-terminale portant un groupement phényle (A) ou en fonction de la nature du groupement hydrophobe sur un dérivé de l’acide acétique (B). Le témoin (T) est le dipeptide phosphoThr-Pro (1) acétylé en position N-terminale.

En résumé, l’apport d’un groupement hydrophobe en terme d’affinité dépend principalement de la nature du groupement aromatique et, plus modérément, de la position de celui-ci par rapport au dipeptide phosphoThr-Pro.

Pour cette série de composés, les  $IC_{50}$  ont été calculées par compétition avec le peptide centré autour du motif phosphoThr212-Pro213 marqué avec une sonde dansyle pour la liaison au domaine WW. La compétition entre les inhibiteurs et le substrat peptidique est suivie par FRET. Le domaine WW est solubilisé dans un tampon Tris 50 mM pH 6.8, NaCl 100 mM, DTT 5 mM à 20  $\mu$ M. Le peptide est solubilisé à 120  $\mu$ M conduisant d'après le  $K_D$  du peptide de 100  $\mu$ M à 50% de protéine liée. Des concentrations croissantes de ligand sont ajoutées à la solution précédente et le calcul de l' $IC_{50}$  est basé sur l'augmentation de l'intensité d'émission de fluorescence à 334 nm. Les résultats pour les composés (27) à (31) sont résumés dans le Graphique 4.6.



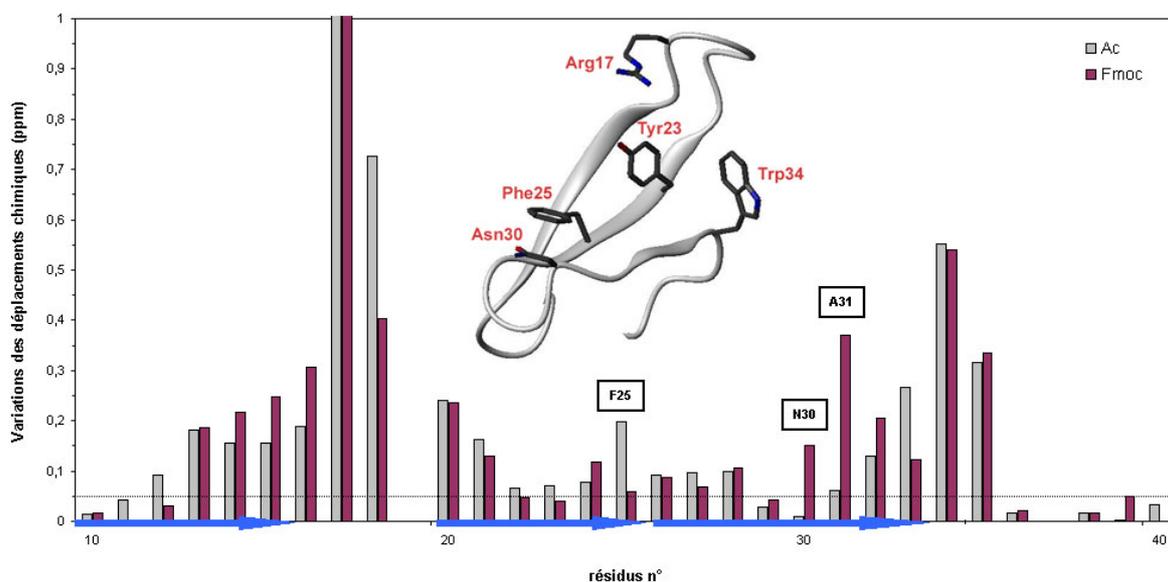
**Graphique 4.6 :**  $IC_{50}$  des composés (27) à (31) évalués en FRET par compétition avec 120  $\mu$ M de peptide substrat pour la liaison au domaine WW à 20  $\mu$ M, soit 50% (10  $\mu$ M) de protéine en complexe.

Les  $IC_{50}$  relatives pour les composés (27) à (31) ne sont pas homogènes avec les  $K_D$  et les différences sont plus marquées entre les composés à chaîne aliphatiques de deux et trois carbones et les composés à chaîne aliphatique de quatre et cinq carbones. L' $IC_{50}$  est reliée au  $K_D$  à la seule condition que les composés et le peptide substrat entrent en compétition pour le même site. D'après les  $IC_{50}$ , le  $K_D$  du peptide substrat ainsi que les concentrations initiales du domaine WW et du peptide, le calcul des constantes de dissociation des composés (27) à (29) aboutirait à des valeurs de l'ordre de 100  $\mu$ M et de l'ordre de 25  $\mu$ M pour les composés (30) et (31). Une telle différence avec les  $K_D$  déterminés par RMN peut s'expliquer par la présence d'un motif supplémentaire d'interaction au niveau des ligands pour lequel il n'existe pas de compétition avec le peptide dans la liaison au domaine WW. On observe deux comportements distincts en fonction de la longueur de la chaîne aliphatique. Dans le cas des molécules (27) à (29), seule la partie contenant le motif phosphoThr-Pro entre en compétition avec le peptide et la liaison de la queue hydrophobe est consécutive à la liaison du motif phosphoThr-Pro. On

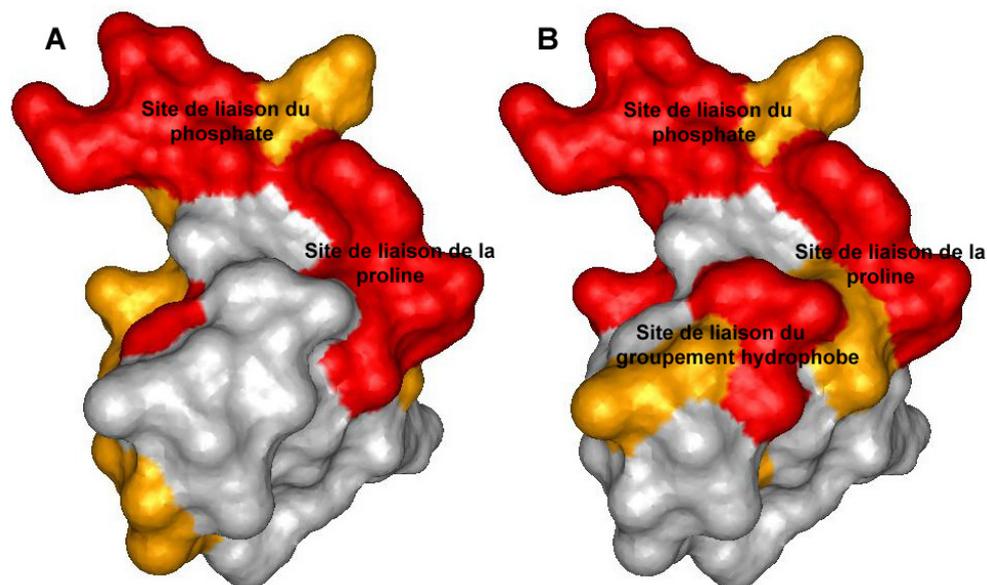
sait d’ailleurs que cette partie de la molécule possède une affinité similaire à celle du peptide substrat (de l’ordre de 100  $\mu$ M). Dans le cas des molécules (30) et (31), l’augmentation d’affinité pourrait provenir de la liaison préalable de la queue hydrophobe (même de très faible affinité), qui favoriserait la liaison par augmentation de la concentration locale, et donc la compétition, du motif phosphoThr-Pro avec le peptide substrat, aboutissant à une affinité apparente plus élevée d’un facteur 4.

**Identification d’un site de liaison additionnel au niveau du domaine WW par mapping des déplacements chimiques.**

Le *mapping* des déplacements chimiques des résidus du domaine WW a permis l’identification d’un nouveau site d’interaction avec les groupements hydrophobes introduits sur la fonction amine du dipeptide phosphoThr-Pro. On observe pour les composés (26) à (32) une perturbation plus importante des résidus Asn30 et Ala31 ainsi qu’une perturbation moins importante du résidu Phe25 qui peut correspondre à un effet antagoniste entre les perturbations dues à la liaison du dipeptide phosphoThr-Pro et la liaison du groupe hydrophobe sur le déplacement chimique de la Phe25. On remarque également des perturbations plus importantes de la Ser16 et moins importantes de la Ser18 pour les dérivés portant un groupe hydrophobe par rapport au dérivé acétylé (Graphique 4.7, Figure 4.24).



**Graphique 4.7 :** *Mapping* des déplacements chimiques comparés des composés (1) et (26) permettant l’identification d’un nouveau site d’interaction avec le groupement Fmoc au niveau des résidus Phe25, Asn30 et Ala31. Représentation de la poche constituée des résidus Phe25 et Asn30 sur la structure du domaine WW.

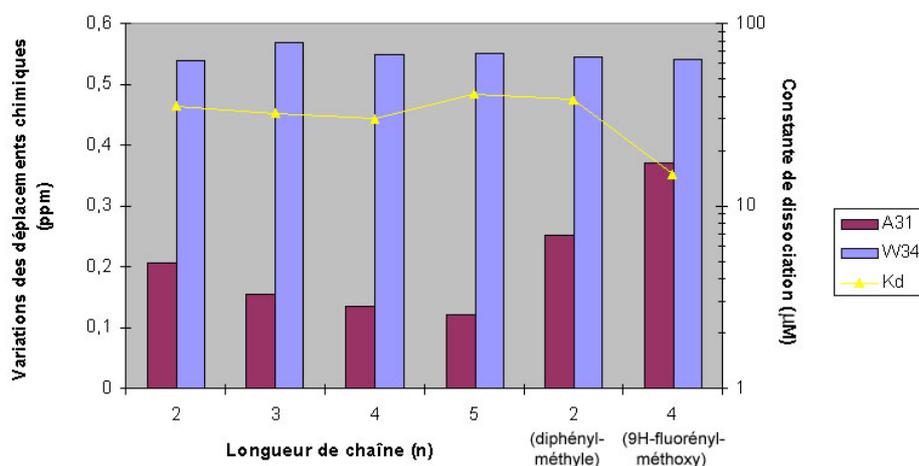


**Figure 4.24 :** Surfaces d'interaction des différents sites du domaine WW pour les dipeptides phosphoThr-Pro acétylé (A) et Fmoc (B). Les surfaces représentées en rouge correspondent à des variations de déplacements chimiques supérieures à 0.1 ppm, les surface représentées en orange correspondent à des variations de déplacements chimiques comprises entre 0.04 et 0.1 ppm.

#### ***Spécificité de l'interaction des groupements hydrophobes avec le domaine WW.***

Les perturbations de déplacements chimiques des résidus Ala31 et Asn30 varient avec la longueur de la chaîne aliphatique portant le substituant phényle, alors que les résidus impliqués dans la liaison de la proline comme les résidus Trp34 ne varient pas (Graphique 4.8). Les perturbations de l'Ala31 sont les plus importantes avec les groupements aromatiques les plus encombrants (9H-fluorényle >  $\alpha$ -diphénylméthyle > phényle) et diminuent quand la longueur de la chaîne aliphatique augmente mais ne sont pas corrélées avec les constantes de dissociation des ligands (Graphique 4.8). Ces données doivent quand même être relativisées compte-tenu de l'incertitude concernant la détermination des  $K_D$  et les variations de déplacements chimiques.

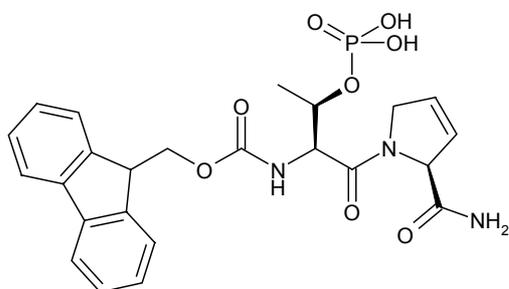
Les résidus Ser16 à Ser18 localisés dans la boucle qui lie le phosphate présentent des perturbations de déplacements chimiques très variables d'un ligand à l'autre indiquant plutôt des changements de positionnement de la boucle pour les différents ligands.



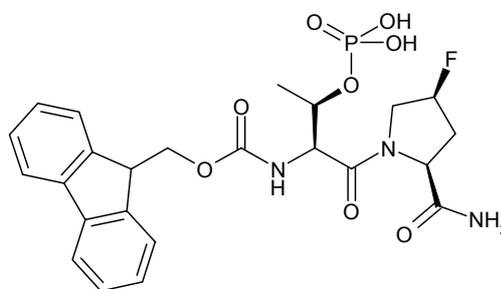
**Graphique 4.8 :** Comparaison des variations de déplacements chimiques des résidus W34 et A31 en fonction de la longueur de chaîne aliphatique (n), la nature du groupement aromatique (phényle, α-diphénylméthyle ou 9H-fluorényle) et la constante de dissociation des différents ligands.

***Association de mimes de proline couplés à la phosphothréonine et de groupements hydrophobes en position N-terminale.***

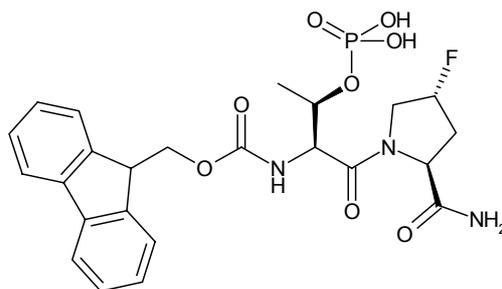
Les mimes de proline de type 3,4-dehydro-proline et 4-*cis/trans*-fluoro-proline ont été couplés à la Fmoc-phosphothréonine, dont le groupement Fmoc a été conservé comme groupe hydrophobe pour l'obtention des composés **(33)** à **(35)**. Nous nous attendions à obtenir avec les dérivés Fmoc de la 3,4-dehydro-proline et 4-*trans*-fluoro-proline des constantes de dissociation de l'ordre du micromolaire étant donné les constantes de dissociation des dérivés acétylés correspondants (20 et 40 μM respectivement). L'affinité des composés **(34)** et **(35)** a augmenté mais, de manière surprenante, l'affinité du composé **(33)** a diminué (Tableau 4.4). Avec la 4-*trans*-fluoro-proline, même si l'affinité est meilleure, le gain n'est pas aussi important qu'avec les dérivés de la proline (composés **(26)** / **(1)**).



**(33)**



**(34)**

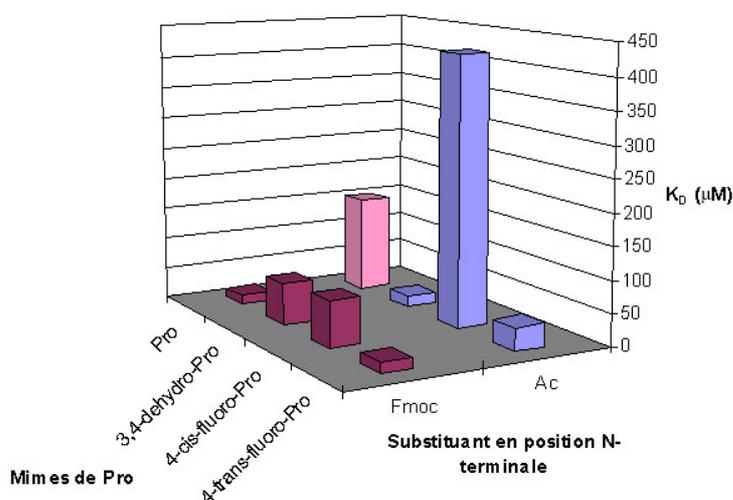


(35)

	(33)	(34)	(35)
<b>Mimes de Pro (dérivés Fmoc)</b>			
<b><math>K_D</math> (<math>\mu M</math>)</b>	65	70	15

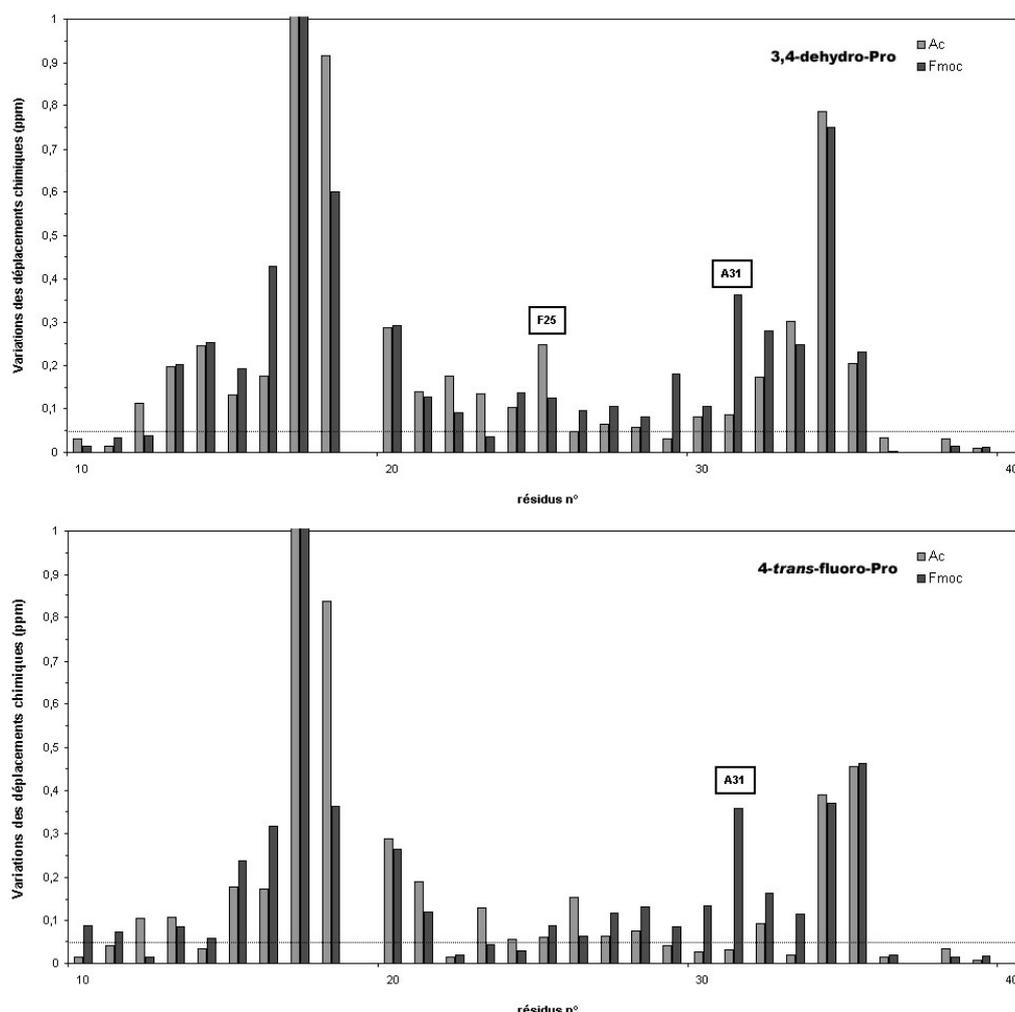
**Tableau 4.4 :** Constantes de dissociation des mimes de proline de type 3,4-dehydro-Pro, 4-*cis/trans*-fluoro-Pro couplés à la N-Fmoc-phosphoThr.

Dans le cas des dérivés de la 4-*cis*-fluoro-proline, on gagne un facteur 5 en affinité en passant de l'acétyle au Fmoc mais seulement un facteur 2 pour les dérivés de la 4-*trans*-fluoro-proline et on perd un facteur 3 pour les dérivés de la 3,4-dehydro-proline (Graphique 4.9). Une explication possible est le repliement de la molécule sur elle-même qui représente une conformation « fermée » défavorable à la liaison au domaine WW. Un apport d'énergie supplémentaire doit alors être fourni pour passer en conformation « ouverte » avant la liaison au domaine WW.



**Graphique 4.9 :** Relations structure-affinité des composés Fmoc par rapport aux composés acétylés contenant les mimes de proline de type 3,4-dehydro-, 4-*cis/trans*-fluoro-proline couplés à la N-Fmoc-phosphoThr.

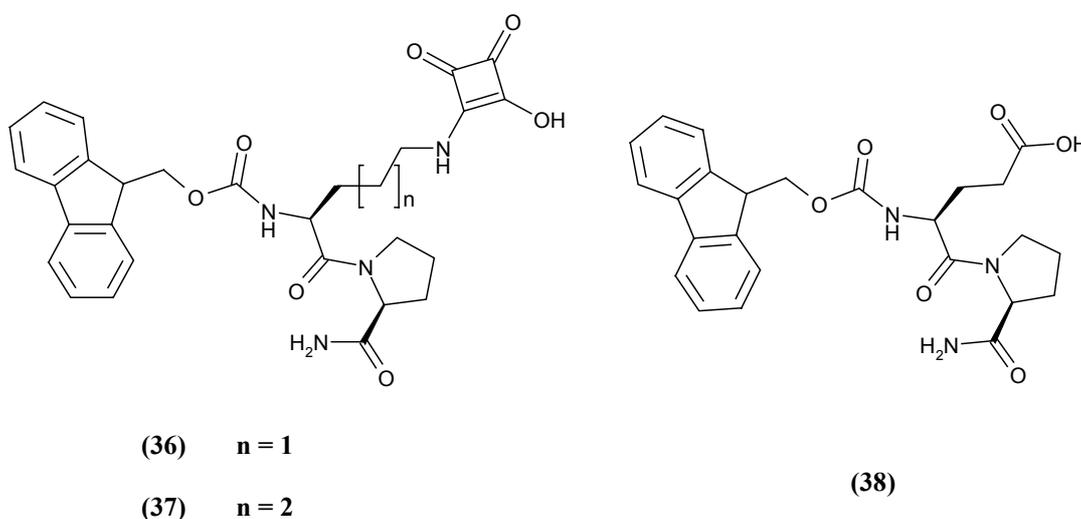
Le mode de liaison de cette série de composés est similaire au mode de liaison du dérivé Fmoc (**26**) du dipeptide phosphoThr-Pro (Graphique 4.10). Comme pour les dérivés de la proline, on remarque une perturbation plus importante du résidu Ala31 pour les dérivés Fmoc par rapport aux dérivés acétylés. Ces deux mimes de proline induisaient des perturbations plus importantes de la Tyr23 par rapport à la proline en version acétylée, traduisant un positionnement différent (peut-être meilleur) du cycle dans la poche aromatique du domaine WW. Les dérivés Fmoc, en revanche, induisent des perturbations de la Tyr23 équivalentes au dérivé Ac-phosphoThr-Pro (**1**). On remarque, comme pour les dérivés comportant une proline, des perturbations plus importantes de la Ser16 et moins importantes de la Ser18 pour les dérivés portant un groupe Fmoc par rapport aux dérivés acétylés.



**Graphique 4.10 :** Mapping des déplacements chimiques comparés des composés (**9**)/(**33**) et (**7**)/(**35**) indiquant l’influence d’un groupement Fmoc sur les perturbations des résidus du domaine WW pour les mimes de proline de type 3,4-dehydro-proline et 4-trans-fluoro-proline.

**Perspectives concernant le remplacement du phosphate par une tête polaire.**

L'apport énergétique des groupements hydrophobes pour la liaison des ligands au domaine WW peut être mis à profit pour gagner en affinité de manière suffisamment significative de sorte que les composés contenant un mime du phosphate aient une affinité mesurable. Nous avons synthétisé des dérivés squaramides de l'ornithine (**36**) et de la lysine (**37**) couplés à la proline pour lesquels la protection Fmoc de la fonction amine a été conservée. Un composé obtenu par couplage de l'acide N-Fmoc glutamique sur la proline (**38**) a été également synthétisé. La constante de dissociation des deux premiers composés a été évaluée à 970 et 540  $\mu\text{M}$  respectivement et celle du dérivé glutamate a été estimée à 500  $\mu\text{M}$  (Tableau 4.5).



	(36)	(37)	(38)
$K_D$ ( $\mu\text{M}$ )	970	540	$\sim 500$

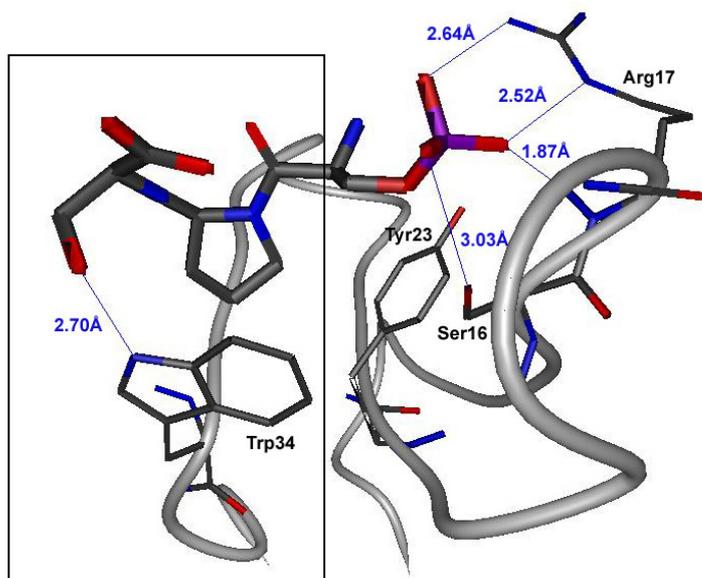
**Tableau 4.5 :** Constantes de dissociation des composés constitués d'analogues de phosphate de type squaramide ou carboxylate.

La présence du groupement Fmoc a permis d'accéder à une gamme d'affinités mesurables mais ces affinités restent faibles. Néanmoins, dans le cas des dérivés squaramides de l'ornithine et de la lysine, la chaîne aliphatique est plus longue que la chaîne de la thréonine et, dans les trois composés, la chaîne n'est pas ramifiée. On peut donc s'attendre à une augmentation d'affinité pour des dérivés squaramides ou carboxylates basés sur le squelette carboné de la thréonine, avec un gain d'un facteur 3 environ, d'après l'affinité du composé (**10**) contenant une phosphosérine, soit des affinités de l'ordre de 200 à 300  $\mu\text{M}$ . Ces affinités

(hypothétiques) sont du même ordre que le composé initial **(1)**, ce qui laisse à penser que le remplacement du groupement phosphate reste un point très délicat.

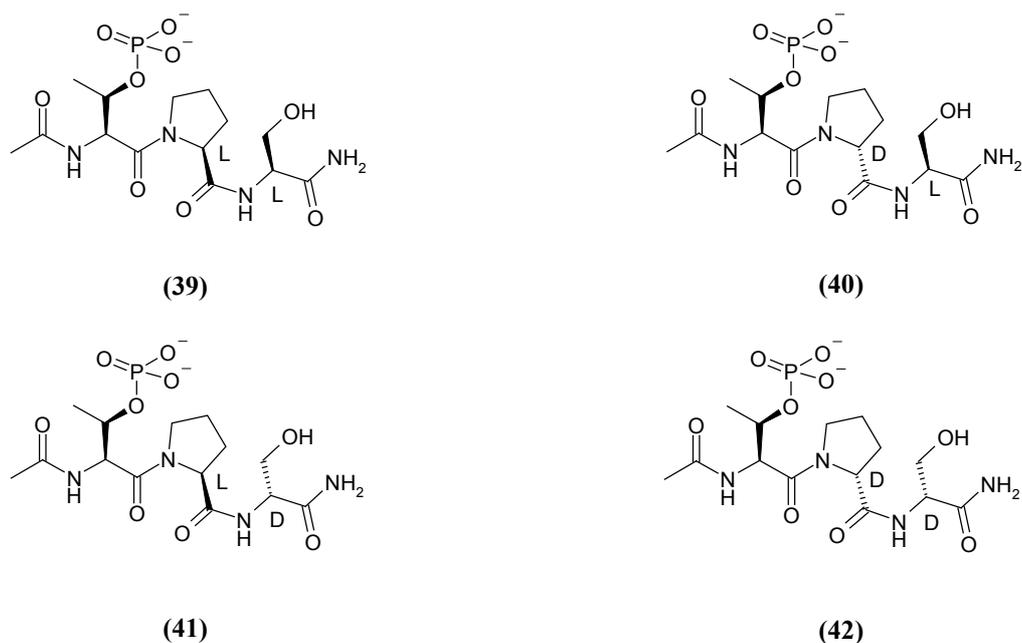
### 2.3.2 – Introduction d’un groupement accepteur de liaison hydrogène en position C-terminale

D’après la structure cristallographique de Pin1 en complexe avec un peptide phosphorylé du domaine C-terminal de l’ARN polymérase II, la fonction alcool du résidu sérine en +1 par rapport au motif phosphoThr-Pro établit une liaison hydrogène avec le H<sub>N</sub> de la chaîne latérale du Trp34 (Figure 4.25).



**Figure 4.25** : Eléments structuraux du domaine WW intervenant dans la liaison du dipeptide phosphoThr-Pro et d’un résidu Ser supplémentaire en position C-terminale.

Pour évaluer l’implication de cette liaison hydrogène dans la liaison à Pin1, nous avons synthétisé le tripeptide Ac-phosphoThr-(L)-Pro-(L)-Ser-NH<sub>2</sub> **(39)**. Comme le composé **(2)**, obtenu par couplage de la phosphoThr sur l’énantiomère de la proline, avait donné une affinité du même ordre de grandeur que le composé référence **(1)**, le tripeptide Ac-phosphoThr-(D)-Pro-(L)-Ser-NH<sub>2</sub> **(40)** ainsi que les épimères de ces deux tripeptides contenant la (D)-sérine (composés **(41)** et **(42)**) ont également été synthétisés pour déterminer l’importance de la position de la sérine dans la structure du motif d’interaction phosphoThr-Pro pour la liaison au domaine WW.



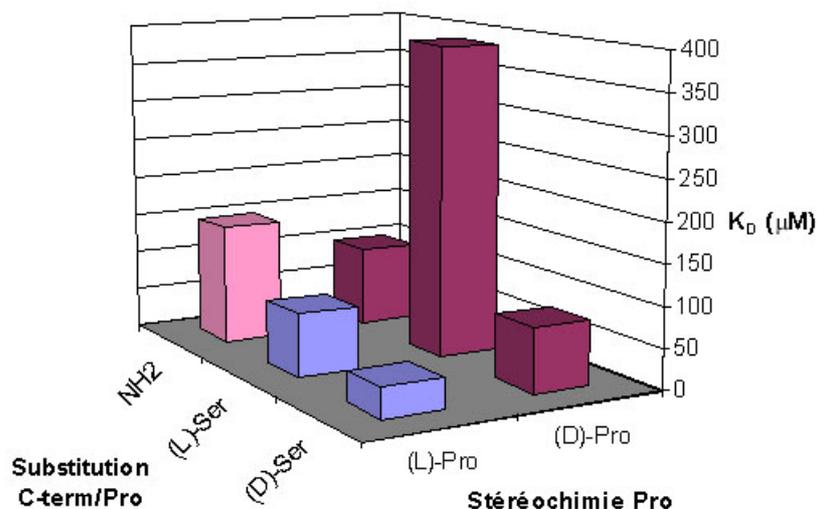
	(39)	(40)	(41)	(42)
$K_D$ ( $\mu M$ )	80	390	40	80

**Tableau 4.6 :** Constantes de dissociation des tripeptides combinant les énantiomères (L) ou (D) de la proline aux énantiomères (L) ou (D) de la sérine, associés à la (L)-phosphothréonine.

Les composés **(39)** et **(40)** se distinguent par l'épimérisation du carbone  $\alpha$  de la proline. Mais cette différence de stéréochimie se traduit par une diminution significative de l'affinité d'un facteur 5 environ pour le composé contenant la (D)-proline. L'épimérisation supplémentaire du carbone  $\alpha$  de la sérine dans le composé **(42)** permet de retrouver l'affinité initiale de 80  $\mu M$  du composé **(39)**. En revanche, la seule épimérisation du carbone  $\alpha$  de la sérine dans le composé **(41)** induit une augmentation de l'affinité d'un facteur 2 par rapport au composé **(39)**.

***Rôle positif d'un accepteur de liaison hydrogène sur la fonction carbonyle de la proline pour la liaison au domaine WW.***

Dans la série de composés contenant une (L)-proline, l'ajout d'une (L)-sérine en position C-terminale du dipeptide phosphoThr-Pro permet d'augmenter l'affinité d'un facteur 2 et l'ajout d'une (D)-sérine d'un facteur 4. Avec les composés contenant une (D)-proline, l'ajout d'une (L)-sérine provoque une diminution de l'affinité par rapport au dipeptide **(2)** et l'ajout d'une (D)-sérine n'a pas un effet significatif mais augmente l'affinité d'un facteur 2 par rapport au composé référence **(1)** (Graphique 4.11).

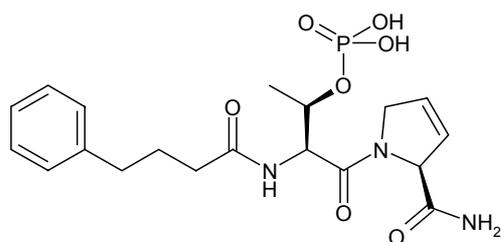


**Graphique 4.11 :** Relations structure-affinité pour les tripeptides (39) à (42) comparées aux dipeptides phosphoThr-Pro (1) et (2).

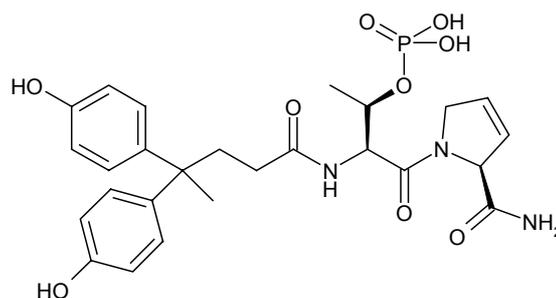
L’ajout d’un accepteur de liaison hydrogène sur la fonction acide de la proline permet d’augmenter l’affinité des ligands sauf dans le cas de la (D)-proline et peut fournir une voie supplémentaire de pharmaco-modulation pour optimiser la liaison au domaine WW. L’atome d’oxygène étant accepteur de liaison hydrogène, la fonction alcool peut être remplacée par une fonction éther (méthoxy, éthoxy) pour augmenter la biodisponibilité des composés.

#### 2.4 – Combinaison des éléments structuraux les plus favorables.

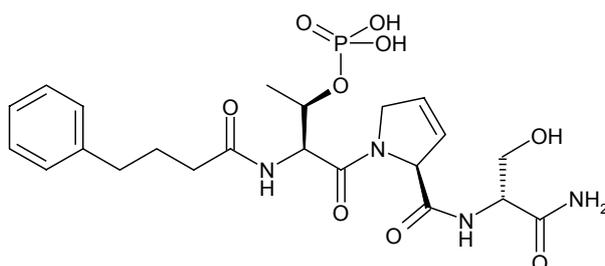
Malgré la perte d’affinité des composés basés sur la 3,4-dehydro-proline possédant un groupement Fmoc, nous avons conservé ce mime de proline dans l’investigation de ligands porteurs d’autres groupements aromatiques dans le but d’améliorer l’affinité pour le domaine WW. Nous avons choisi des groupements aromatiques plus flexibles portés par des chaînes aliphatiques à quatre carbones ayant comme précurseurs les acides 4-phénylbutyrique (43) et 4,4-bis(4-hydroxyphényl)valérique (44). Un composé basé sur la structure du composé (43) a été synthétisé avec couplage de la (D)-sérine sur la fonction acide de la 3,4-dehydro-(L)-proline (45).



(43)



(44)



(45)

	(43)	(44)	(45)
$K_D$ ( $\mu M$ )	28	5	7

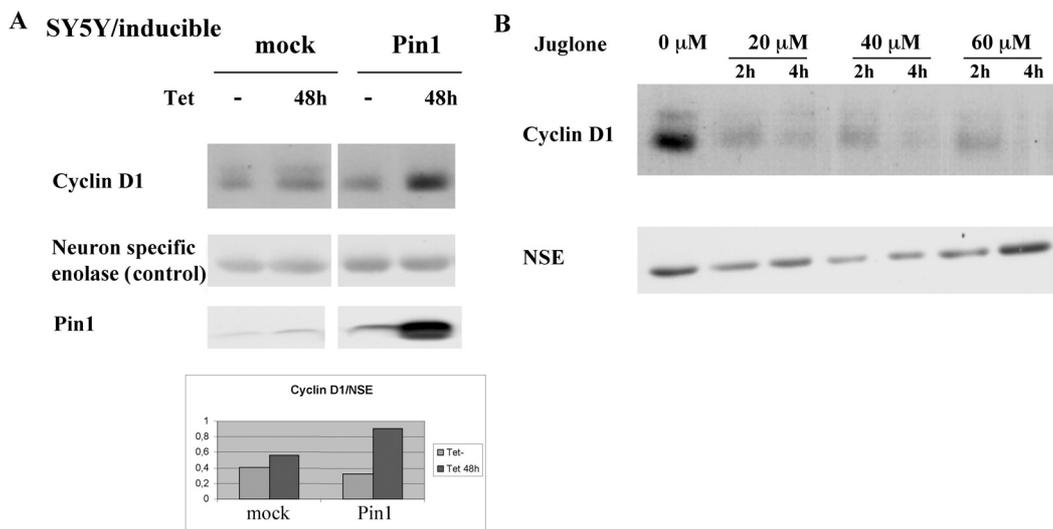
**Tableau 4.7 :** Constantes de dissociation des composés constitués de la 3,4-dihydro-(L)-proline couplée à la phosphothréonine avec comme groupements hydrophobes en position N-terminale le 4,4-bis(4-hydroxyphényl)valéroyle dans le composé (44) et le 4-phénylbutyroyle dans les composés (43) et (45), ce dernier comportant en plus une (D)-sérine en position C-terminale.

Le groupement 4-phénylbutyroyle du composé (43) induit une perte d'affinité moins importante que le Fmoc en comparaison au composé acétylé (9) mais n'améliore pas l'affinité, probablement parce que le repliement de la molécule par « collapse hydrophobe » n'est pas compensé par le gain d'affinité induit par le groupement hydrophobe, contrairement au groupement 4,4-bis(4-hydroxyphényl)valéroyle dans le composé (44) qui permet d'augmenter l'affinité d'un facteur 4 par rapport au composé acétylé (9). L'ajout d'une (D)-sérine dans la structure du composé (43) permet d'augmenter l'affinité d'un facteur 4 pour le composé (45). On retrouve ce même gain d'affinité quand on ajoute une (D)-sérine sur la position C-terminale du dipeptide Ac-phosphoThr-(L)-Pro dans le composé (41).

### 2.5 – Tests cellulaires.

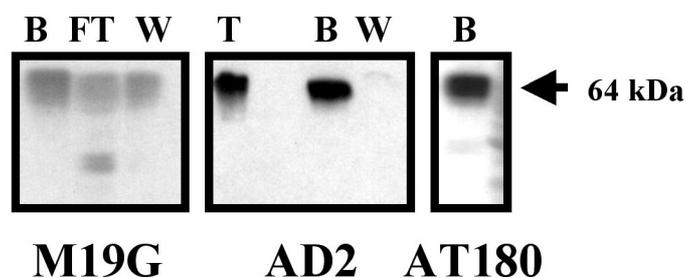
Les molécules ayant franchi l’étape de criblage sont soumises à un test cellulaire basé sur un modèle de neurones en différenciation (en collaboration avec l’équipe U422 Inserm du Dr Luc Buée). Les cellules SY5Y sont transfectées de manière stable avec la protéine tau humaine et peuvent exprimer Pin1 par induction. La différenciation de cellules de type neuroblastomes en cellules neuronales est réalisée par induction à l’acide rétinoïque 1  $\mu\text{M}$ . La surexpression de Pin1 est induite par action de la tétracycline sur le répresseur Tet contrôlant négativement le gène promoteur de l’expression de Pin1.

Il a été montré que Pin1 régulait positivement l’expression de cycline D1 par activation de c-Jun et stabilisation de la  $\beta$ -caténine (Ryo, Nakamura et al. 2001; Wulf, Ryo et al. 2001). Pin1 régule également positivement l’expression de cycline D1 dans les cellules neuronales. La surexpression de Pin1 induit une augmentation du niveau de cycline D1 (Figure 4.26A). L’inhibition de l’activité enzymatique de Pin1 avec la juglone, un inhibiteur de l’activité *cis/trans* isomérase sélectif des parvulines, induit une diminution du niveau de cycline D1 (Figure 4.25B). Après 48 heures d’induction, une augmentation de l’immunoréactivité de la cycline D1 a été observée. Une diminution de l’immunoréactivité de la cycline D1 est observée dans les cellules SY5Y induite par la tétracycline et traitées par la juglone à 20  $\mu\text{M}$ . Le niveau de la NSE (*neuron specific enolase*) est utilisé comme contrôle dans les deux cas.



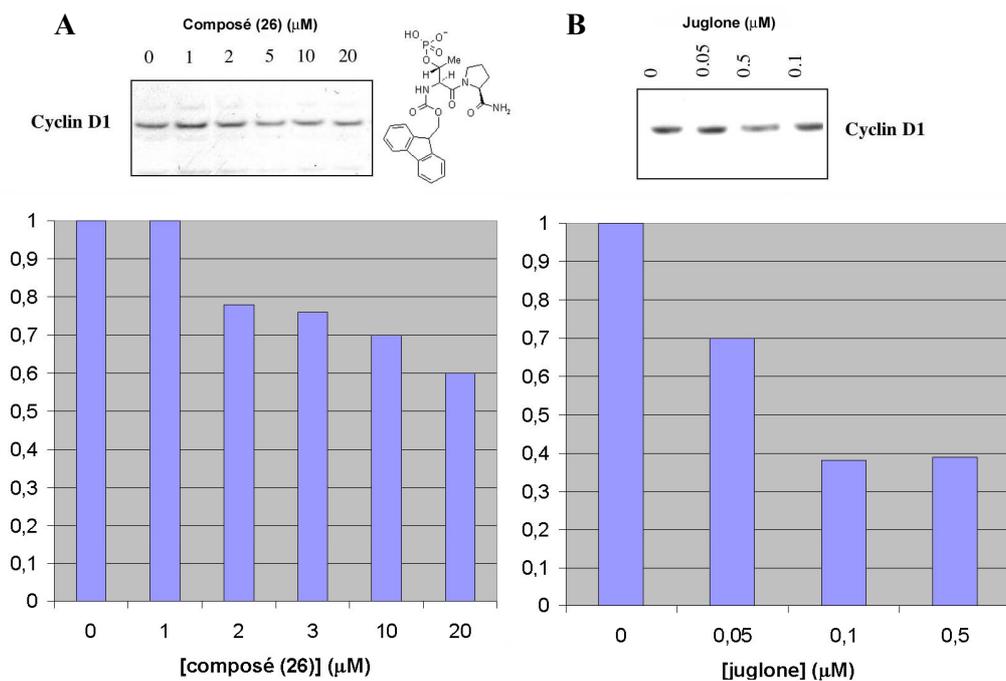
**Figure 4.26 :** Corrélation entre le niveau d’expression de Pin1 et le niveau d’expression de cycline D1 dans les neuroblastomes SY5Y différenciés. (A) L’induction de Pin1 par la tétracycline conduit à une surexpression de Pin1 après 48H corrélée avec une surexpression de cycline D1 dans les cellules SY5Y. (B) L’inhibition de l’activité catalytique de Pin1 surexprimées dans les cellules induites par la juglone entraîne une diminution du niveau d’expression de cycline D1.

Dans les cellules en différenciation SY5Y, il est possible de détecter la phosphorylation de tau par divers anticorps, dont l’AT180 dirigé spécifiquement contre le site T231. La balance phosphorylation/déphosphorylation est assurée par l’action de phosphatases comme PP2A, notamment au niveau de la T231. La liaison de tau phosphorylée dans les cellules SY5Y différenciées à Pin1 est suivie par *pull-down*. Les différentes fractions sont analysées par immunoréactivité vis-à-vis de divers anticorps : l’anticorps M19G qui reconnaît la protéine tau non phosphorylée (au niveau du fragment N-terminal), l’anticorps AD2 qui reconnaît les sites de phosphorylation mitotique S396/S404 et l’anticorps AT180 qui reconnaît le site T231 phosphorylé (Figure 4.27). Comme Pin1 stimule la déphosphorylation de tau par PP2A (Zhou, Kops et al. 2000), il est également envisageable de détecter l’activité ou l’inhibition de l’activité de Pin1 en suivant le niveau de phosphorylation du site reconnu par l’anticorps AT180 (Hamdane et al., non publié).



**Figure 4.27** : Détection de protéines liant Pin1 dans un lysat cellulaire SY5Y surexprimant tau. La protéine Pin1 est fixée, via une queue de 6 histidines, sur résine d’agarose chargée en ions  $\text{Ni}^{2+}$ . A gauche, détection de protéines tau (non phosphorylée) par l’anticorps anti-tau polyclonal M19G dans les fractions de protéines liées (B), non liées (FT) et après lavages (W). Au centre, détection de protéine tau phosphorylée en S396/S404 par l’anticorps AD2 dans l’extrait total (T), dans les fractions contenant les protéines liées (B) et dans la fraction de lavage (W). A droite, détection de protéine tau phosphorylée en T231 dans la fraction contenant les protéines liées (B).

Les composés les plus affins ont été évalués dans le test cellulaire décrit ci-dessus (Figure 4.28A). Les cellules SY5Y de type neuroblastomes sont différenciées en neurones par l’acide rétinoïque ou le NGF pendant 14 jours. Dans les cellules neuronales, Pin1 régule positivement l’expression de la cycline D1 (Hamdane, Smet et al. 2002) ainsi l’inhibition de Pin1 est déterminée par le niveau d’expression de la cycline D1. Des doses d’environ 10  $\mu\text{M}$  de composé induisent une diminution du niveau de cycline D1 similaire à 1  $\mu\text{M}$  de juglone (Figure 4.28B).



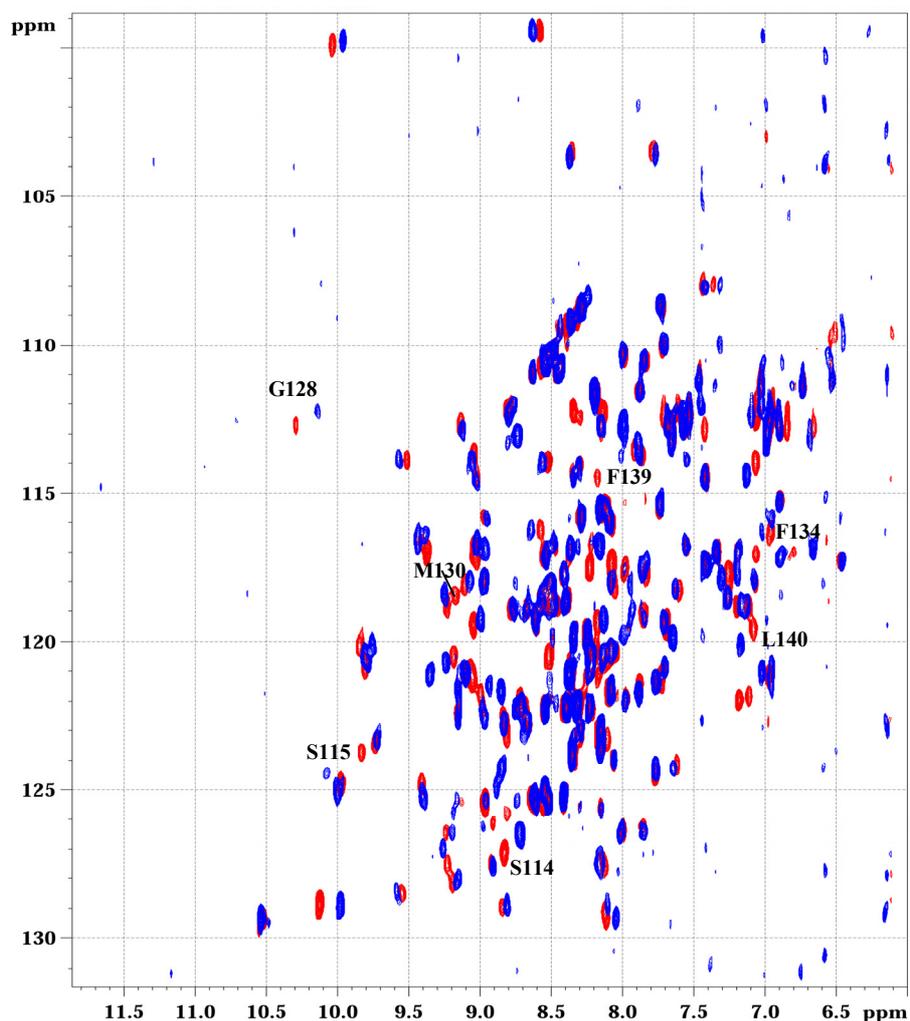
**Figure 4.28** : Tests cellulaires réalisés sur le composé (26) dont l’activité est mesurée par la diminution du niveau d’expression de cycline D1 (A) et comparée à la juglone (B).

Les tests cellulaires ont permis de montrer une corrélation entre l’inhibition fonctionnelle de Pin1 par la juglone et la diminution du niveau de cycline D1 dans le neurone. Ces effets ont été reproduits avec les composés synthétisés qui ont montré les meilleures affinités par RMN mais sont obtenus avec des doses plus importantes qu’avec la juglone. Il est impossible de déterminer pour le moment si ces effets sont dus à l’interaction des ligands avec le domaine WW ou avec le domaine catalytique et, par conséquent, si ces molécules agissent en inhibant la liaison des substrats de Pin1 au niveau du domaine WW ou en inhibant l’activité catalytique directement.

#### ***Interaction des ligands avec le domaine catalytique de Pin1.***

Les composés mimant le dipeptide phosphoThr-Pro ciblent le domaine WW qui est impliqué dans la régulation de l’activité de Pin1 *in vivo* et dans la localisation cellulaire de Pin1 (Lu, Zhou et al. 2002) et dont le rôle est aussi crucial que celui du domaine enzymatique *in vivo* (Lu, Zhou et al. 1999). Nous avons évalué le mode d’interaction des composés avec la protéine Pin1 entière et déterminé des interactions avec le domaine catalytique pour le composé (26) ainsi qu’une constante de dissociation de 12 µM, similaire à celle du complexe formé avec le domaine WW isolé. Des perturbations très nettes sont observées pour les

résidus du site catalytique de Pin1 ainsi que pour les résidus situés à l’interface avec le domaine WW (Figure 4.29).



**Figure 4.29** : Interaction entre le dipeptide Fmoc-phosphoThr-Pro et Pin1. En bleu, le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC de Pin1 seule à 125  $\mu\text{M}$  et en rouge, le rapport des concentrations 1 :20 entre la protéine et le ligand.

On peut également envisager que les molécules n’agissent pas sur Pin1 mais ciblent d’autres protéines, comme les phosphatases PP2A ou Cdc25. Un *pull-down* pourrait permettre de « piéger » la(es) protéine(s)-cible(s) dans les extraits cellulaires SY5Y en utilisant les composés synthétisés.

### **III – Discussion. La RMN comme outil de criblage pour des ligands de faible affinité. Application à la recherche d’inhibiteurs des interactions Pin1-tau.**

#### *La RMN comme technique de criblage.*

Dans le but de mieux comprendre l’implication moléculaire de Pin1 dans le processus neurodégénératif de la maladie d’Alzheimer et les relations entre Pin1 et la protéine tau constituant les fibres intraneuronales, nous avons privilégié la piste des petites molécules jouant le rôle d’inhibiteurs de Pin1. La RMN a été utilisée dans les premières étapes de la recherche de ligands du domaine WW de Pin1 selon une stratégie de conception rationnelle à partir d’un composé de départ de faible affinité, le dipeptide phosphoThr-Pro, et la connaissance de la structure tridimensionnelle de la cible. Nous avons évalué, dans un premier temps, l’affinité de composés très polaires constitués de mimes de phosphate liés à une proline ou de mimes de proline liés à une phosphothréonine. Les changements structuraux peu importants concernant les mimes de proline et le remplacement du phosphate laissent présager des affinités proches de celle du dipeptide (150  $\mu\text{M}$ ) et nécessitent une technique adaptée capable de faire la distinction entre des composés de faible affinité. D’un point de vue pratique, la RMN n’est pas une technique couramment utilisée en criblage mais l’industrie pharmaceutique porte un intérêt grandissant pour la RMN dans ce domaine car elle présente des caractéristiques que ne possèdent pas les méthodes de criblage classiques basées sur la fluorescence (FRET, polarisation) ou les tests fonctionnels. Les deux principaux attraits de la RMN en criblage, qui ont orienté le choix de cette technique dans le cas de la recherche d’inhibiteurs de Pin1, sont la détection d’affinités faibles et l’apport en informations structurales.

L’utilisation de quantités relativement faibles de protéine (20  $\mu\text{M}$ ) pour le criblage a été facilitée par :

- l’utilisation d’une cryosonde
- la petite taille du domaine WW associée à une faible largeur de raie et l’utilisation d’une séquence impulsionnelle améliorant la sensibilité
- l’enrichissement isotopique en  $^{15}\text{N}$  de la protéine qui permet de filtrer le signal du ligand et le signal de l’eau.

Il faut souligner qu’une concentration de 20  $\mu\text{M}$  est faible pour la RMN mais relativement élevée par rapport aux techniques HTS classiques. L’acquisition de spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC est réalisée pour chaque composé en 20 minutes. Il est clair que l’utilisation de la RMN ici n’est pas optimale pour le criblage de grandes chimiothèques (10 000 à 100 000 composés). On pourrait optimiser les conditions pour augmenter l’efficacité du criblage par RMN en testant des composés en mélange. Une telle approche est à associer avec une stratégie de synthèse combinatoire qui permet d’obtenir rapidement des banques de plusieurs centaines de composés. La RMN est généralement utilisée dans les premières étapes de criblage pour la recherche de *hits* lorsque aucune structure de départ n’est connue pour une cible donnée car elle permet de « pêcher » des composés d’affinité relativement faible (100  $\mu\text{M}$  - 1 mM), qui échapperaient aux techniques HTS classiques en limitant la proportion de faux positifs.

Nous avons mis au point des conditions de criblage avec 20  $\mu\text{M}$  de protéine uniformément enrichie en  $^{15}\text{N}$  qui permettent de déterminer rapidement dans quelle gamme d’affinité se situent les ligands. Avec 400  $\mu\text{M}$  de composé (soit 20 équivalents), il est possible de faire la distinction entre des composés dont le  $K_D$  est supérieur à 3 mM, inférieur à 70  $\mu\text{M}$  ou compris entre ces deux valeurs. La limite supérieure a été estimée à 3 mM en raison du seuil de détectabilité des perturbations des signaux du domaine WW correspondant à 10% de complexe. La limite inférieure a été établie à 70  $\mu\text{M}$  en fonction du seuil de détection de la forme liée de l’Arg17 estimé à 85% de complexe. Pour les composés induisant des perturbations des déplacements chimiques du domaine WW, le calcul de l’affinité des ligands par RMN, a permis de faire la distinction entre des composés présentant de faibles affinités pour le domaine WW (entre 0.01 et 1 mM) et a permis de définir des relations structure-affinité pour des composés faiblement affins en même temps qu’un mode de liaison.

La structure cristallographique du domaine WW (Verdecia, Bowman et al. 2000) a guidé l’investigation d’un nouveau site d’interaction supplémentaire avec le ligand. Une poche hydrophobe constituée des résidus Phe25 et Asn30 est explorée avec des ligands basés sur le motif phosphoThr-Pro contenant un groupement hydrophobe de type aromatique ancré sur une chaîne aliphatique de longueur variable pour déterminer si le groupement aromatique se lie au niveau du domaine et déterminer la distance optimale par rapport au motif primaire d’interaction. Les composés dérivés de cette structure ont révélé une meilleure affinité et leur

mode de liaison a pu également être déterminé. Nous avons confirmé la liaison avec le site d'interaction présumé grâce aux données structurales.

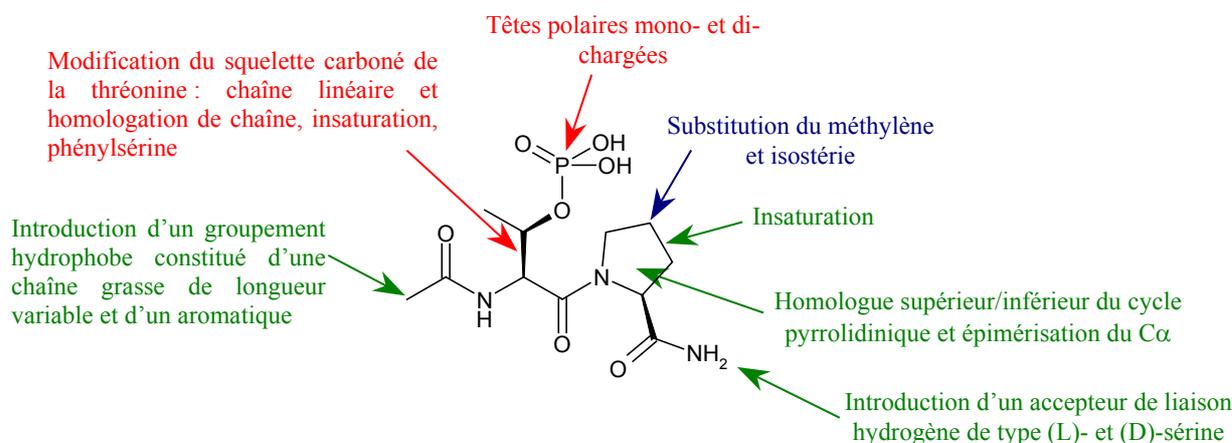
### ***Chimie en phase solide et diversification rapide des composés.***

L'utilisation de la chimie supportée a permis l'obtention rapide de molécules relativement complexes (poids moléculaire élevé, plusieurs centres d'asymétrie, présence d'une fonction acide), facilement diversifiables sans modification de la voie de synthèse, en quantité suffisante pour la réalisation des étapes de criblage et de *binding* par RMN.

### ***Modifications pharmacophoriques apportées au dipeptide phosphoThr-Pro.***

Les modifications structurales apportées au composé de référence (**1**) concernent (Figure 4.30) :

- le cycle pyrrolidinique de la proline (substitution, isostérie, homologation, insaturation, épimérisation),
- le groupement phosphate, remplacé par des têtes polaires (carboxylate,  $\alpha$ -dicarboxylate, phosphonate, squaramide, tétrazole, trifluorosulfonamide, 3,4-difluoroaniline, 3-nitroaniline) et des modifications du squelette carboné de la thréonine (simultanément),
- l'introduction de groupes hydrophobes constitués de chaînes aliphatiques de longueur variable portant un aromatique sur la fonction amine de la phosphoThr (ou analogues),
- l'introduction d'un groupe accepteur de liaison hydrogène sur la fonction carbonyle de la proline (ou analogues).



**Figure 4.30** : Récapitulatif des modifications pharmacophoriques apportées sur le dipeptide phosphoThr-Pro (**1**). Les modifications positives sont annotées en vert, négatives en rouge, et positives ou négatives en bleu selon les modifications effectuées.

La recherche d'analogues de la proline a permis d'avoir une idée de la profondeur et la flexibilité de la poche hydrophobe du domaine WW impliquée dans la liaison du cycle pyrrolidinique. Le remplacement du méthylène en position 4 du cycle par un isostère ou la substitution par un groupe hydroxyle indique que le volume est limité et ne peut accueillir que des cycles proches d'un point de vue stérique comme l'homologue inférieur (azétidine) et l'homologue supérieur (pipéridine). La substitution du même méthylène par un fluor est tolérée en position *trans* et induit une augmentation significative de l'affinité. Un tel composé pourrait être exploité efficacement comme substrat-test pour le criblage de composés par compétition (Jahnke, Floersheim et al. 2002) suivie par RMN<sup>19</sup>F (Dalvit, Flocco et al. 2002). La RMN<sup>19</sup>F possède de nombreux avantages par rapport à la RMN<sup>1</sup>H, tels que la sensibilité, une gamme de déplacements chimiques très larges (300 ppm), l'absence de présaturation du signal de l'eau et un seul signal car un seul atome de fluor. L'introduction d'une insaturation est également très favorable à la liaison du domaine WW avec un gain d'affinité d'un facteur 7 environ ce qui suggère que le remplacement de la pyrrolidine par un cycle aromatique de type pyrrole pourrait être également favorable. Nous avons tenté de synthétiser un composé pour lequel la proline est remplacée par l'acide 2-pyrrole carboxylique mais l'acylation de l'azote aromatique du pyrrole par la phosphothréonine dans des conditions basiques fortes (avec NaH dans le DCM anhydre) a conduit à la perte du groupement phosphate et à l'obtention d'un mélange de plusieurs composés. L'épimérisation du carbone  $\alpha$  de la proline n'induit pas de modification significative de l'affinité, suggérant que la fonction carboxamide portée par ce même carbone n'intervient pas dans la liaison. Il est envisageable de synthétiser des composés ne comportant pas ce groupement carboxamide mais il faudra adapter la voie de synthèse pour l'obtention de ces composés.

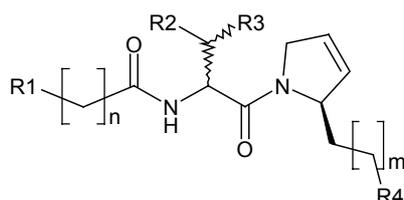
L'introduction de groupements additionnels sur le composé de base phosphoThr-Pro au niveau de la fonction amine de la phosphothréonine et sur la fonction carbonyle de la proline a conduit à une amélioration significative d'affinité. L'ajout de groupements hydrophobes sur la phosphothréonine a l'impact le plus significatif avec des augmentations d'affinités pour le domaine WW allant jusqu'à 10 fois celle du dipeptide phosphoThr-Pro avec le groupement Fmoc. La combinaison du groupement Fmoc et des mimes de proline de type 4-*trans*-fluoro-proline et 3,4-dehydro-proline couplés à la phosphothréonine ne provoque pas une augmentation d'affinité aussi drastique pour le mime fluoré par rapport à la proline, et induit même une perte d'affinité pour le mime insaturé. Cette perte d'affinité est probablement due à une conformation repliée en solution du ligand, peu favorable à la liaison au domaine WW.

Le remplacement du 9H-fluorényle par des aromatiques moins rigides permet d’augmenter l’affinité des ligands. L’introduction d’une sérine en configuration L ou D sur les dipeptides phosphoThr-Pro, au niveau du carbonyle de la proline, a permis d’évaluer l’impact d’un accepteur de liaison hydrogène sur la liaison au domaine WW. La position de l’hydroxyle de la sérine par rapport à la proline influence l’affinité des ligands suggérant un rôle positif (établissement d’une liaison hydrogène) ou négatif (encombrement stérique) de celui-ci. Les diastéréoisomères LLL et LDD (d’après la stéréochimie des carbones  $\alpha$  des trois résidus phosphoThr, Pro et Ser respectivement) sont équivalents en termes d’affinité. Le composé de stéréochimie LLD est plus affiné que l’épimère LLL (ou LDD), lui-même beaucoup plus affiné que l’épimère LDL. L’apport d’une (D)-sérine en terme d’affinité a été confirmé avec le composé portant un mime de type 3,4-dehydro-proline. D’autre part, s’il faut garder la liaison peptidique, le fait d’avoir des acides aminés (D) est un gage d’une meilleure stabilité métabolique.

Le remplacement du groupement phosphate a été moins fructueux car les composés constitués de têtes polaires mimant le phosphate dans le composé **(1)** n’ont pas d’affinité détectable ce qui correspond à une perte d’affinité d’un facteur 20 et plus par rapport au composé de départ. L’introduction d’un groupement hydrophobe Fmoc a permis d’obtenir des affinités mesurables pour les composés contenant un carboxylate ou un squaramide comme tête polaire, mais plus faibles que le composé de départ **(1)**.

### *Perspectives.*

Les données obtenues sur les différentes parties constituant le ligand permettent d’envisager des modifications structurales autour de la forme générique décrite ci-dessous. En particulier, l’influence de la chaîne carbonée de la thréonine n’a pas été étudiée et semble jouer un rôle non négligeable dans la mesure où le remplacement par une sérine diminue l’affinité des ligands significativement.



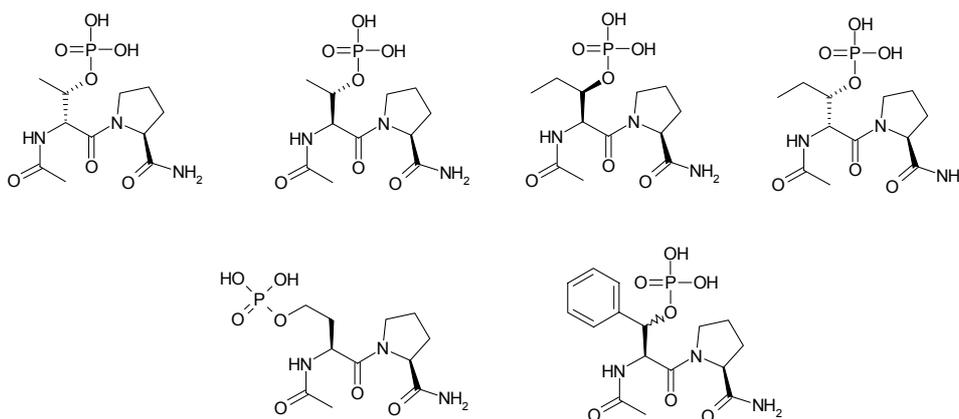
R1 : aromatique substitué, aromatique condensé...  
 R2 : Me, Et, iPr, Ph ....  
 R3 : squarate, carboxylate, dicarboxylate...  
 R4 : OH, SH, OMe, SMe...

L’investigation d’analogues de thréonine nécessite la mise au point d’une stratégie de synthèse incluant la phosphorylation d’une fonction alcool. Une stratégie de phosphorylation

en phase solide a été décrite pour la phosphorylation des sucres (Parang, Fournier et al. 2001) et permet un ancrage des molécules au support solide par la tête polaire.

On peut envisager comme analogues de thréonine des acides- $\alpha$ -aminés de configuration différente :

- la (D)-thréonine, un énantiomère pour lequel les configurations des deux carbones asymétriques sont inversées
- l'allo-thréonine, un diastéréoisomère pour lequel seule la configuration du carbone  $\alpha$  est inversée
- la DL-hydroxynorvaline, un homologue de chaîne, qui revient à remplacer le groupement méthyle de la thréonine par un éthyle
- l'homosérine, un isomère de position du groupement hydroxyle de la thréonine
- la 3-phénylsérine, qui revient à remplacer le groupement méthyle de la thréonine un phényle.



Une stratégie de synthèse en phase solide par ancrage du composé par la tête polaire permettrait de diversifier les autres fragments du ligand et, en particulier, remplacer la fonction carboxamide obtenue à partir de la fonction acide de la proline qui n'est pas nécessaire à la liaison du domaine WW. On peut également envisager la synthèse de mimes contraints du dipeptide phosphoThr-Pro pour éviter le collapse hydrophobe observé avec certaines molécules.



## Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 and phosphopeptides.

Caroline Smet,<sup>1</sup> Jean-Frédéric Duckert,<sup>1</sup> Jean-Michel Wieruszkeski,<sup>1</sup> Isabelle Landrieu,<sup>1</sup> Luc Buée,<sup>2</sup> Guy Lippens,<sup>1</sup> and Benoît Déprez.<sup>\*1</sup>

CNRS-University of Lille2 UMR 8525, 59000 Lille, France

RECEIVED DATE (will be automatically inserted after manuscript is accepted).

Interactions involving phosphorylated Ser/Thr-Pro motifs in proteins play a key role in numerous regulatory processes in the cell. Here, we investigate potential ligands of the WW binding domain of Pin1 in order to inhibit protein-protein interactions between Pin1 and phosphopeptides. Our structure-based strategy implies the synthesis of analogues of the Ac-Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro-NH<sub>2</sub> dipeptide and relies on high resolution NMR spectroscopy to accurately measure the affinity constants even in the high micromolar range.

Protein-protein interactions have an important role in most biological processes. A number of small domains present in various proteins with different biological functions are recurrent in protein-protein interactions and stand as such in the core of the cellular protein network.<sup>1</sup> Interactions mediated by phosphorylated moieties on their targets are of particular interest, since crucial regulation processes involve a balance between phosphorylation and dephosphorylation pathway allowed by kinase and phosphatase activity.<sup>2</sup>

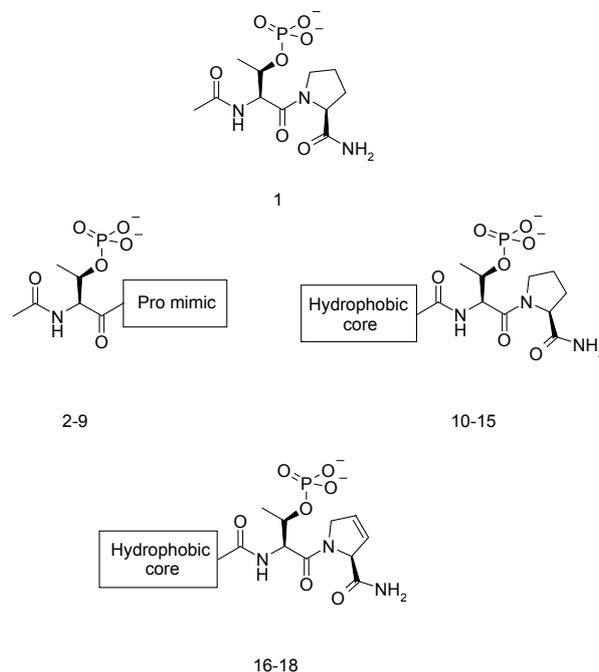
Pin1 is a peptidyl-prolyl *cis/trans* isomerase constituted by two domains, a catalytic prolyl *cis/trans* isomerase domain and a WW binding module. Both share a specificity for phospho-Thr/Ser-Pro motifs. At the cellular level, Pin1 is shown to be an essential protein involved in cell cycle regulation, since Pin1 inhibition prevents entry into mitosis and cell division.<sup>3,4</sup> Overexpression of Pin1 in numerous human cancers indicates its implication in several oncogenic pathways.<sup>5,6</sup> Therefore, Pin1 is proposed as a therapeutic target for cancer treatment.<sup>7</sup> Pin1 also interacts with the neuronal microtubule-associated tau protein when the latter is phosphorylated at specific motifs, and thereby restores its ability to polymerise the tubulin into microtubules.<sup>8</sup>

Inhibition of Pin1 activity with small-molecules targeting the enzymatic domain blocks cell proliferation in cancer cell lines. Juglone was the first of this category of irreversible ligands, and the synthesis and characterization of several derivatives has been described.<sup>9,10</sup> The WW domain of Pin1, however, is equally essential for its *in vivo* activity: in healthy tissue, phosphorylation of the WW Ser16 inhibits its binding role, and inactivates the whole protein.<sup>11</sup> In breast cancer cells, the WW domain is dephosphorylated, leading to the active form of the protein.

At the molecular level, the WW domain has a flexible loop encompassing Ser16-Ser19 residues that binds the phosphate moiety on Ser or Thr residues and an aromatic clamp

constituted by its Tyr23 and Trp34 side chains that binds the prolyl residue in the +1 position of the phosphoSer/Thr.<sup>12,13</sup>

The H-Thr(PO<sub>3</sub>H<sub>2</sub>)-Pro-NH<sub>2</sub> dipeptide **1** (Figure 1) was assessed by NMR as the minimal requirement for binding to the Pin1 WW module exhibiting a relatively high dissociation constant of 150 μM and a similar mode of binding as larger peptidic substrates.<sup>14</sup> Based on the molecular structures of the WW domain and the ligand **1**, we investigate here potential ligands for the Pin1 WW binding domain, by sequentially considering compounds that contain proline or phosphate mimics, and further extending these compounds with additional hydrophobic cores (Figure 1). High resolution NMR proved a sensitive assay even for very weak ligands,<sup>15,16</sup> with cryogenic probe technology limiting appropriately the amount of protein and ligand.<sup>17</sup> Moreover, structural information could be obtained at every step through a simple mapping of chemical shift variations of the WW domain.



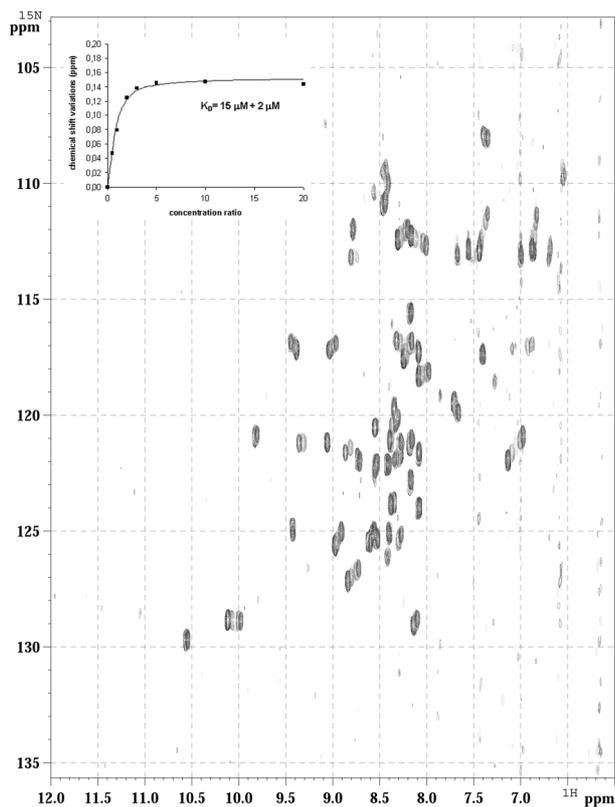
**Figure 1.** Structures of synthetic WW domain-binding analogs. Structure details of phosphate and proline mimics are given in Table 1 and hydrophobic cores in Table 2.

**Synthesis.** We have investigated proline mimics coupled to phosphothreonine through an amide bond and hydrophobic cores according to the template described in Figure 1 for the compounds **2-9** and **10-18** respectively. Compounds were synthesized by solid phase chemistry on polystyrene solid support functionalized by a “rink amide” linker. The chemistry involved is very similar to solid phase peptide synthesis. Amine functions were protected by a Fmoc group and deprotected by incubation of the resin with a 20% piperidine solution in DMF. Carboxylic acid activation was performed by a mixture of 1 equiv.HOBT/ 1 equiv. HBTU and 3 equiv. DIPEA in DMF. When needed, the terminal amine function of the dipeptide was acetylated with acetic anhydride and DIPEA in DMF.

\* To whom correspondence should be addressed. Phone : +33/(0) 320964024. Fax : +33/(0) 320964709. E-mail : bdeprez@pharma.univ-lille2.fr.

<sup>1</sup> CNRS-University of Lille2 UMR 8525. 3 rue du Professeur Laguesse, 59000 Lille, France.

<sup>2</sup> Inserm U422, place de Verdun, 59000 Lille, France.



**Figure 2.** NMR titration of the  $^{15}\text{N}$  labeled WW domain with the ligand **15**. Overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra acquired on a  $50\mu\text{M}$  WW domain  $^{15}\text{N}$ -labelled sample with increasing amounts of compound **15**. In the left top corner, a titration curve plotting the chemical shift variations against the concentration ratio **15**:WW.

**Discovery of novel reversible ligands for the Pin1 WW binding module.** In an initial phase, we screened compounds by a rapid 1D NMR experiment at a twenty-fold excess of ligand to WW domain. The WW domain at  $20\mu\text{M}$  was  $^{15}\text{N}$ -labelled, allowing the efficient filtering of the ligand resonances. The signal of the  $\text{H}_{\text{N}\epsilon}$  of the Trp34 (at 10.2 ppm) was used as a sensitive probe to detect perturbations of the Trp side chain chemical environment upon binding of the proline ring. When interaction was detected, we performed an additional titration study with a  $50\mu\text{M}$   $^{15}\text{N}$ -labelled WW

sample and increasing amounts of ligands (from 0.5 to 20 equiv.). Gradual chemical shift perturbations with increasing ligand concentrations were used to derive the dissociation constant (Figure 2).

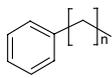
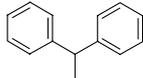
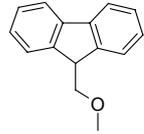
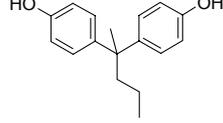
The aromatic clamp was found to accept Pro mimics with structural variations of the pyrrolidine ring. Despite the relatively low affinities, NMR titrations yielded reliable SAR results for analogues of Ac-Thr( $\text{PO}_3\text{H}_2$ )-Pro- $\text{NH}_2$  (**2-9**). Compounds where the pyrrolidine ring is replaced by an azetidine or a piperidine bind the WW domain without significant changes of affinity. The same observation is made for the epimer at the  $\text{C}_\alpha$  of proline. Substitutions have contradictory effects on binding since the replacement of proline by 4-hydroxyproline, 4-*cis*-fluoroproline or thiazolidine-4-carboxylic acid isoster decrease dramatically the affinity for the WW domain. However replacement of proline with a 4-*trans*-fluoroproline improves significantly the binding as well as the introduction of an insaturation at position 3-4 (Table 1) of the pyrrolidine ring. Contrary to the proline mimics, polar groups used to replace the phosphate moiety all exhibited dissociation constants as low as 5 mM, pointing out an extreme specificity of the flexible loop for phosphate binding.

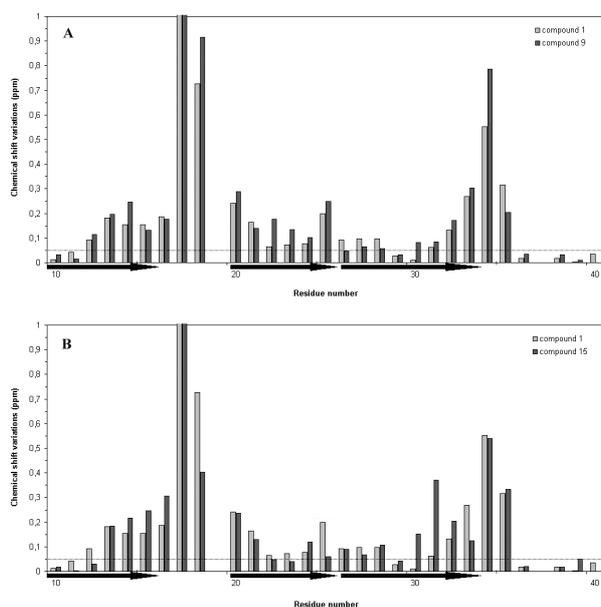
Faced with restricted possibilities for structural modifications around the two major points of interaction of peptide **1** with the WW domain, we have examined further WW hydrophobic patches in the near proximity of the main binding site, in order to gain in affinity. A potential hydrophobic surface constituted by Phe25 and Ala31 was targeted with a phenyl ring anchored to the Thr( $\text{PO}_3\text{H}_2$ )-Pro dipeptide via aliphatic chains of various length (compounds **10-14**) and with the Fmoc group (**15**). A significant increase of binding affinity was observed for these compounds with a maximum for the 4-phenylbutyric acid precursor and for the Fmoc group that confirms the agreement with the 4 carbon-length chain (Table 2). Nevertheless, interactions with the phosphate binding loop and the prolyl ring binding pocket remain essential and appear as a minimal requirement for ligands to adapt the domain since the replacement of phosphothreonine by a glutamic acid in dipeptides carrying the same hydrophobic cores did not significantly bind the WW domain (data not shown).

**Table 1.** Structures of proline mimics (**2-9**) and dissociation constant ( $K_D$ ) of resulting ligands.

	2	3	4	5	6	7	8	9
Pro mimic structure								
$K_D$ ( $\mu\text{M}$ )	100	160	140	400	800	36	420	18

**Table 2.** Structures of hydrophobic cores and dissociation constant of resulting ligands based on proline (**10-15**) and 3,4-dehydroproline (**16-18**) derivatives (underlined).

compound :	10	11	12	<u>16</u>	13	14	15	<u>17</u>	<u>18</u>
Hydrophobic core structure	 $n=1,2,3,4$								
$n =$	1	2	3	<u>3</u>	4			<u>66</u>	<u>5</u>
$K_D$ ( $\mu\text{M}$ )	35	32	30	<u>28</u>	48	38	15	<u>66</u>	<u>5</u>



**Figure 3.** NMR chemical shift mapping of the  $^{15}\text{N}$  labelled WW domain. Comparison of the mapping of WW domain in presence of 20 equiv. of compound **1** and **9** (A), in presence of 20 equiv. of compound **1** and **15** (B).

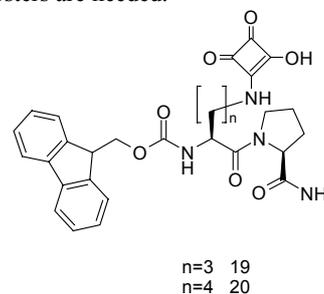
**Consistency of binding mode in the series.** Chemical shift interaction mapping was performed for each compounds of significantly improved affinity, to verify the consistency of the binding mode. Very similar modes of binding were observed for the various proline mimics in comparison with the dipeptide **1**. The 3,4-dehydroproline moiety induces larger chemical shift perturbations for both phosphate and proline binding sites, and especially a stronger effect on the Y23 amide peak (Figure 3A), together with a significant line broadening concerning most of interacting residues. The binding site of the hydrophobic core that replaces the acetyl group in the phospho-Thr-Pro dipeptide is given by the chemical shift mapping of compounds **10-15**. Larger variations of Ala31 amide shifts (Figure 3B) and a dependence of the Ala31 variations with the chain length carrying a phenyl group (**10-13**) that is not shared by Trp34 or Arg17 verify our initial structural hypothesis.

Combining the data on chain length and nature of the aromatic group with the previously determined optimal proline mimics provides a further perspective for improving the affinity. The combination of the three essential motifs

was further investigated in compounds **16** to **18** described in Figure 1 based on the 3,4-dehydroproline analog. In the proline series (compounds **10** to **15**), the replacement of the acetyl group by  $\omega$ -arylalkanoyl chains or a Fmoc group increased affinity as expected. Surprisingly, except for 4,4-bis(4-hydroxyphenyl)-pentanoyl (**18**), most of the hydrophobic cores induce a decrease of affinity (Table 2) in the dehydroproline series, suggesting different structural constraints in the free ligand of the two series.

Finally, to confirm that our small ligands can actually disrupt a protein-protein interaction, we have conducted competition experiments using fluorescence energy transfer. Our ligands were tested for their ability to displace a tetradecapeptide centered on the Thr ( $\text{PO}_3\text{H}_2$ )-Pro(213) of tau protein<sup>14</sup> from the WW domain. In this setup, our best ligands have shown to be able to displace the peptide with  $K_d$  consistent with their  $K_d$  (data not shown).

**Revisiting phosphate mimics on the basis of the additional hydrophobic interaction.** The latter ligands with affinity in the low micromolar range allowed us to revisit the possible phosphate mimics. Squaramide is singly charged at physiological pH and thank to delocalization of this charge can accept strong hydrogen bonds at each of its oxygen atoms. Compounds carrying this group instead of the phosphate were synthesized and their binding studied by NMR. In the Fmoc series (**19** and **20**), compounds displayed interesting affinities (respectively 910 and 540  $\mu\text{M}$ ), thereby providing new chemical perspectives to replace the phosphate moiety for which metabolically stable and cell-permeable isosters are needed.



**Conclusion.** Our NMR-based strategy has led to the conception, synthesis and testing of new ligands for the Pin1 WW domain. Whereas the proline in the initial phosphoThr-Pro dipeptide could be successfully replaced by a 3,4-dehydro- or 4-*trans*-fluoro-proline, the specificity for the phosphate moiety proved more challenging. Based on the

structure of the WW domain, we extended these ligands with an N-terminal aromatic group, resulting in a new series of ligands with low micromolar affinity. NMR proved crucial throughout this development, as a technique to bridge the gap between very low affinity hit compounds and conventional tools to measure affinity of small molecules to protein, such as fluorescence polarization or energy transfer. It allowed initial screening, and generated reliable SAR at low levels of affinity, as well as structural information on the binding mode of our ligands. The ligands developed here will be further tested in a cell-based assay.

**Acknowledgment.** The authors thank the Institute for the Study Of Aging (ISOA, the United States) for financial support. We thank Gérard Montagne, Hervé Drobecq and Eric Diesis for their technical contribution to this work. The 600 MHz facility used in this study was funded by the Région Nord-Pas de Calais (France), the CNRS, and the Institut Pasteur de Lille.

**Supporting information available** Synthetic procedures and spectral characterization of compounds **1** to **20** are available.

## References

- (1) Pawson, T.; Raina, M.; Nash, P. Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett* **2002**, *513*, 2-10.
- (2) Yaffe, M. B.; Smerdon, S. J. The Use of in Vitro Peptide-Library Screens in the Analysis of Phosphoserine/Threonine-Binding Domain Structure and Function. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 225-244.
- (3) Lu, K. P.; Hanes, S. D.; Hunter, T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **1996**, *380*, 544-547.
- (4) Winkler, K. E.; Swenson, K. I.; Kornbluth, S.; Means, A. R. Requirement of the prolyl isomerase Pin1 for the replication checkpoint. *Science* **2000**, *287*, 1644-1647.
- (5) Wulf, G. M.; Ryo, A.; Wulf, G. G.; Lee, S. W.; Niu, T. et al. Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *Embo J* **2001**, *20*, 3459-3472.
- (6) Wulf, G.; Ryo, A.; Liou, Y. C.; Lu, K. P. The prolyl isomerase Pin1 in breast development and cancer. *Breast Cancer Res* **2003**, *5*, 76-82.
- (7) Lu, K. P. Prolyl isomerase Pin1 as a molecular target for cancer diagnostics and therapeutics. *Cancer Cell* **2003**, *4*, 175-180.
- (8) Lu, P. J.; Wulf, G.; Zhou, X. Z.; Davies, P.; Lu, K. P. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* **1999**, *399*, 784-788.
- (9) Uchida, T.; Takamiya, M.; Takahashi, M.; Miyashita, H.; Ikeda, H. et al. Pin1 and Par14 peptidyl prolyl isomerase inhibitors block cell proliferation. *Chem Biol* **2003**, *10*, 15-24.
- (10) Hennig, L.; Christner, C.; Kipping, M.; Schelbert, B.; Rucknagel, K. P. et al. Selective inactivation of parvulin-like peptidyl-prolyl cis/trans isomerases by juglone. *Biochemistry* **1998**, *37*, 5953-5960.
- (11) Lu, P. J.; Zhou, X. Z.; Liou, Y. C.; Noel, J. P.; Lu, K. P. Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. *J Biol Chem* **2002**, *277*, 2381-2384.
- (12) Wintjens, R.; Wieruszkeski, J. M.; Drobecq, H.; Rousselot-Pailley, P.; Buee, L. et al. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J Biol Chem* **2001**, *276*, 25150-25156.
- (13) Verdecia, M. A.; Bowman, M. E.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **2000**, *7*, 639-643.
- (14) Smet, C.; Sambo, A. V.; Wieruszkeski, J. M.; Leroy, A.; Landrieu, I. et al. The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* **2004**, *43*, 2032-2040.
- (15) Widmer, H.; Jahnke, W. Protein NMR in biomedical research. *Cell Mol Life Sci* **2004**, *61*, 580-599.
- (16) Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins. *Science* **1997**, *278*, 497-499.
- (17) Hajduk, P. J.; Gerfin, T.; Boehlen, J. M.; Haberli, M.; Marek, D. et al. High-throughput nuclear magnetic resonance-based screening. *J Med Chem* **1999**, *42*, 2315-2317.

## Control of protein-protein interactions: structure-based discovery of low molecular weight inhibitors of the interactions between Pin1 and phosphopeptides.

Caroline Smet, Jean-Frédéric Duckert, Jean-Michel Wieruszkeski, Isabelle Landrieu, Luc Buée, Guy Lippens, and Benoît Déprez.

### Experimental section – solid phase synthesis

Abbreviations :

DCM : dichloromethane

DIPEA : N,N-diisopropylethylamine

DTT : dithiothreitol

DMF : dimethylformamide

HBTU : 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

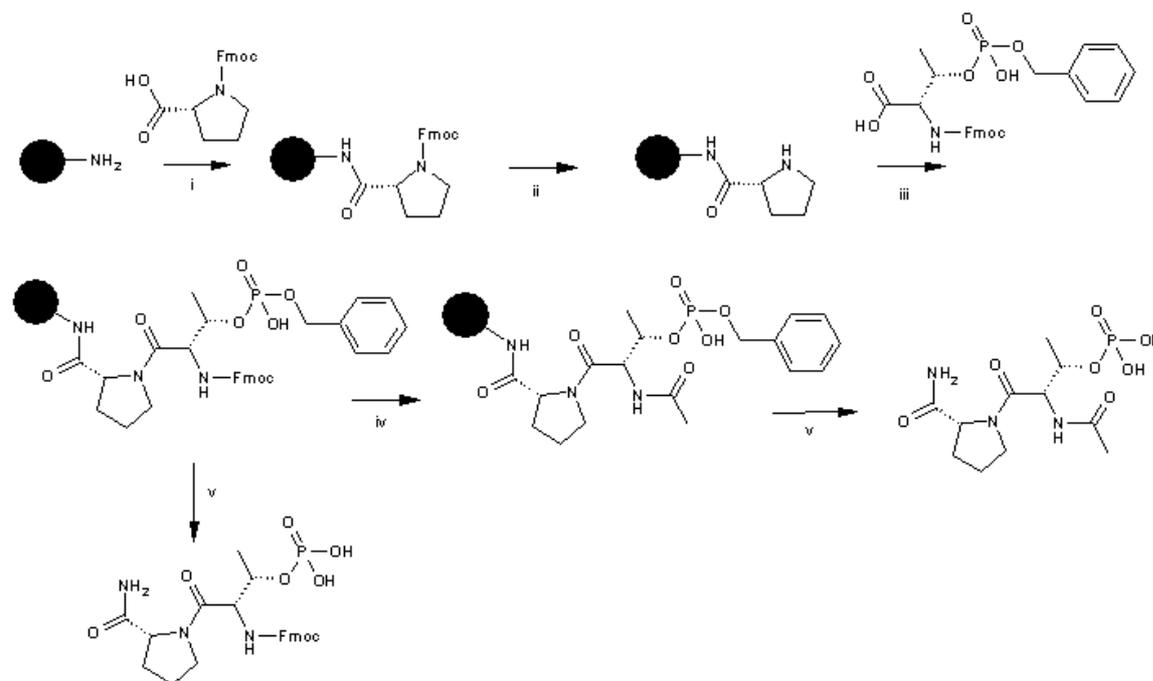
HOBt 1-hydroxybenzotriazole

THF : tetrahydrofuran

TFA : trifluoroacetic acid

Protected natural amino acid, coupling reagents (HOBt and HBTU) and resin were purchased from Novabiochem. The solid phase used is a polystyrene resin cross-linked with 2% divinylbenzene carrying a rink amide MBHA linker (loading 0.74 mmol/mg). Purity was verified by a LC-MS system. Data were recorded on a Varian 1200 MS system using a TSKgel OSH supelco column. Water (A) and acetonitrile (B) were used as eluents containing 0.1% and 0.12% formic acid respectively. After sample injection, the column is equilibrated in solution A for 2 min and a 0-100%B 7.30 min-gradient is performed at a flow rate of 1 mL/min to elute compounds that are detected at 215 nm and 254 nm. Characterisation of compounds solubilized in DMSO- $d_6$  was also done at 300K with  $^1\text{H}$  and  $^{13}\text{C}$  NMR recorded on a 300 MHz Bruker spectrometer equipped with a BBI probe.

Scheme 1<sup>a</sup>



General scheme of synthesis for potential ligands of the Pin1 WW domain. <sup>a</sup> Reagents and conditions : (i) HOBt, HBTU, 3 eq. DIPEA, DMF, room temp, 1 h; (ii) piperidine 20%, DMF, room temp, 25 min; (iii) HOBt, HBTU, 3 eq. DIPEA, DMF, room temp, 1 h; (iv) (a) piperidine 20%, DMF, room temp, 25 min; (b)  $(\text{CH}_3\text{CO})_2\text{O}$ , 2 eq. DIPEA, DMF, room temp, 15 min; (v) TFA 95%, TIS 2.5%, water 2.5%, room temp, 5 h.

## General procedure for compounds 1 to 18

### Coupling of amino acids.

A mixture of 3 eq. Fmoc protected amino acid, 3 eq. HOBt, 3 eq. HBTU were dissolved in DMF and 9 eq. DIEA were added. The resin beads were suspended in the solution at room temperature for 1 hour. The solution was removed by filtration under vacuum and the resin beads were extensively washed with DMF.

### Removal of Fmoc protecting group.

The resin was deprotected in 20% (v/v) piperidine in DMF (3 mL for 300 mg resin). The incubation of resin beads was performed at room temperature for 15 min. The solution was removed by filtration under vacuum and the resin beads were extensively washed with DMF.

### Amine acetylation.

5 eq. acetic anhydride were dissolved in DMF and 10 eq. DIEA was added. The solution was incubated with resin bead at room temperature for 15 min. The solution was removed by filtration under vacuum and the resin beads were extensively washed with DMF.

### Coupling of hydrophobic cores.

A mixture containing hydrophobic cores carrying a carboxylic acid (5 eq.), 1-hydroxybenzotriazole (5 eq.) and HBTU (5 eq.) were dissolved in DMF and 10 eq. DIPEA was added. The coupling reaction was performed at room temperature for 1 hour. The solution was removed by filtration under vacuum and the resin beads were extensively washed with DMF.

### Ninhydrine and chloranil tests.

Each deprotection and coupling step was monitored by a colorimetric test on a small fraction of resin bead. Fmoc removal of resin linker and phosphothreonine was performed with a positive ninhydrine test and coupling of Pro mimics and hydrophobic cores or acetylation with a negative ninhydrine test. Removal of Fmoc protecting group on Pro mimics and coupling of phosphothreonine were monitored with a chloranil test.

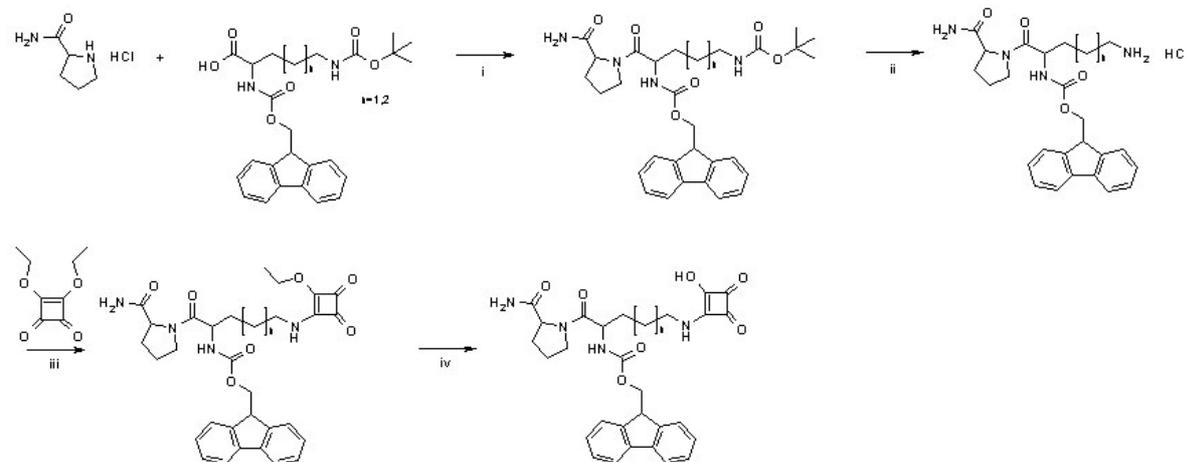
### TFA cleavage of peptidyl-resin and benzyl protecting group and compound isolation.

Resin beads were extensively washed with DMF, DCM and 10% TFA in DCM and incubated for 5 hours at room temperature with a TFA solution containing 2.5% Triisopropylsilane(v/v) and 2.5% water (v/v).

The solution was filtered under vacuum and the resin was washed with 2 times TFA. The combined filtrates were collected in a cold mixture of ether : pentane (v/v). A white precipitate was formed and the precipitation was achieved at 20°C overnight. The precipitate was isolated by filtration and washed with ether, then solubilized in water and lyophilised.

## General procedure for compounds 19 and 20

Scheme 2<sup>a</sup>



General scheme of synthesis for ligands 19-20 of the Pin1 WW domain. <sup>a</sup> Reagents and conditions : (i) DCC/DCM, 0°C, 2 h; (ii) HCl/dioxane, 0°C then room temperature, 3 h; (iii) (a) NEt<sub>3</sub>/EtOH (b) diethylsquarate/EtOH, room temperature, 12 h (iv) HCl/DCM/24H room temperature.

#### Coupling of ornithine and lysine protected derivatives on prolinamide.

3 eq. ornithine/lysine derivatives were solubilized in DCM (15.8 mmol in 8 mL) and added to a solution of 1.5 eq. DCC in DCM at 0°C (7.9 mmol in 2 mL). The precipitate was filtered and the filtrate was added to a solution of 1 eq. prolinamide in DCM (5.25 mmol in 3 mL). The reaction was stirred for 2 hours at 0°C. The crude mixture was evaporated and resolubilized in DCM. The precipitate was removed by filtration and the organic layer was washed twice with an saturated NaHCO<sub>3</sub> aqueous solution then dried over MgSO<sub>4</sub> and evaporated. After purification on silica column in DCM/EtOH 95/5.

#### Removal of Boc protecting group.

1 eq. of the Boc protected compound was solubilized in DCM (1.37 mmol in 14 mL) and cooled in a ice bath. A solution of HCl 4N in dioxane was prepared and added drop wise to the preceding solution at 0°C. The mixture was stirred at room temperature for 3 h then the crude mixture was evaporated. The deprotected product is obtained with quantitative yield.

#### Coupling of diethylsquarate on ornithine/lysine side chain amine function.

A solution of 1.2 eq. diethylsquarate in ethanol (1.64 mmol in 5 mL) was prepared. 1 eq. of ornithine/lysine N-Fmoc protected derivative coupled to prolinamide as solubilized in ethanol (1.37 mmol in 8 mL) and 1.5 eq. triethylamine was added. The mixture was added drop wise in the diethylsquarate solution and was stirred at room temperature for 12 hours. The crude mixture was evaporated and purified on silica column in DCM/MeOH 9/1.

#### Removal of Ethyl protecting group.

The ethylsquarate derivative was solubilized in THF (1.084 mmol in 6 mL) and stirred with a solution of HCl 4N in THF (12 mmol in 3 mL) at room temperature for 20 hours. The crude mixture was evaporated. The solid was solubilized in DCM and extracted four times with a NH<sub>4</sub>HCO<sub>3</sub> 0.3N aqueous solution. The aqueous solution was lyophilized.

#### Compound (1).

Yield 25%

LC-MS  $t_R$  1.38 min 98%  $m/z=336$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.03 (d, 1H, 7.60 Hz); 7.36 (s, 1H); 7.05 (s, 1H); 4.62 (t, 1H, 7.31 Hz) – 55.5; 4.45 (dd, 1H, 6.43 Hz, 2.92 Hz) – 72.4; 4.21 (dd, 1H, 8.18 Hz, 5.26 Hz) – 60.0; 3.77 (m, 1H) – 47.6; 3.63 (m, 1H) – 47.6; 2.04 (m, 1H) – 29.6; 1.88 (m, 1H) – 24.8; 1.85 (s, 3H) – 22.6; 1.79 (m, 1H) – 29.6; 1.26 (d, 3H, 5.85 Hz) – 18.1.

#### Compound (2).

Yield 30%

LC-MS  $t_R$  1.52 min 98%  $m/z=336$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.16 (d, 1H, 7.31); 7.07 (s, 1H); 6.94 (s, 1H); 4.59 (t, 1H, 6.43 Hz) – 55.9; 4.41 (m, 1H) – 73.2; 4.19 (dd, 1H, 8.18 Hz, 2.34 Hz) – 60.8; 3.67 (m, 2H) – 47.9; 1.99 (m, 1H) – 30.2; 1.87 (s, 3H) – 23.2; 1.77 (m, 3H) – 30.2, 24.9; 1.21 (d, 3H 6.43 Hz) – 18.5.

#### Compound (3).

Yield 23%

LC-MS  $t_R$  0.87 min 95%  $m/z=322$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.05 (d, 1H, 7.31 Hz); 7.43 (s, 1H); 7.21 (s, 1H); 4.51 (m, 1H) – 60.8; 4.42 (m, 1H) – 71.7; 4.39 (m, 1H) – 53.8; 4.19 (m, 2H) – 49.0; 2.46 (m, 1H) – 20.2; 2.08 (m, 1H) – 20.2; 1.89 (s, 3H) – 22.7; 1.24 (d, 3H, 5.85 Hz).

**Compound (4).**

Yield 15%

LC-MS  $t_R$  3.62 min 96%  $m/z=350$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 9.70-8.50 (s(b), 1H) ; 8.13-7.17 (s,1H) ; 7.61-7.00 (s,1H) ; 5.01-4.96 (t, 1H, 2.92 Hz) – 52.7-56.6; 4.70-4.45 (t, 1H, 6.72 Hz) – 55.0-55.4; 4.40-4.22 (m, 1H) – 70.7; 4.43-2.37 (m, 1H) – 39.9; 4.05-3.12 (td, 1H, 12.86 Hz, 2.92 Hz) – 43.9; 2.18-1.33 (m, 1H) – 26.6; 2.21-1.41 (m,1H) – 20.3; 1.58-1.40 (m, 1H) – 25.9; 1.54-1.19 (m, 1H) – 25.3; 1.86-1.80 (s, 3H) – 23.0-22.6; 1.55-1.32 (m, 2H) – 21.0; 1.13 (d, 3H, 6.12 Hz) – 19.2.

**Compound (5).**

Compound (5) is obtained according to the general procedure with modifications of coupling of phosphothréonines building block on amine function of thiaproline. Amine function was deprotonated by incubation of resin with 2 eq. DBU in DMF for 10 min. Activation of carboxylic acid function was performed by reaction with 1 eq. oxalyl chloride in DCM/DMF. After removal of the DBU solution, the coupling solution was added to the resin and incubation was spread over 1 hour at room temperature. The cleavage was realized using a solution containing 94% TFA, 2,5% 1,2-ethanedithiol, 1% triisopropylsilane and 2,5% water.

Yield 13%

LC-MS  $t_R$  1.45 min 90%  $m/z=354$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.32 (s, 1H); 7.48 (s, 1H); 7.19 (s, 1H); 5.12 (d, 1H, 9.06 Hz) – 49.9; 4.70 (m, 1H) – 63.1; 4.68 (m, 1H) – 56.5; 4.60 (d, 1H, 9.06 Hz) – 49.9; 4.43 (m, 1H) – 72.34; 4.13 (d, 1H, 9.06 Hz) – 54.4; 4.05 (d, 1H, 9.06 Hz) – 54.4; 3.25 (m, 1H) – 33.6; 3.04 (m, 1H) – 33.6; 1.87 (s, 3H) – 22.8; 1.26 (d, 3H, 5.85 Hz) – 18.3.

**Compound (6).**

Compound (6) is obtained according to the general procedure with, for the acetylation step, reaction of 1 eq. acetic anhydride and 2 eq. DIEA dissolved in DMF with resin beads at room temperature for 15 min.

Yield 22%

LC-MS  $t_R$  0.63 min 92%  $m/z=352$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.55 (s, 1H) ; 7,13 (s, 1H) ; 7.95 (d, 1H, 8.48 Hz) ; 4.67 (dd, 1H, 8.48 Hz, 6.14 Hz) – 55.6 ; 4.47 (m, 1H) – 72.5 ; 4.31 (m, 2H) – 59.5, 69.6 ; 3.67 (d, 2H, 3.22 Hz) – 56.3 ; 2.04 (m,1H) – 38.6 ; 1.88 (s, 3H) – 22.9 ; 1.80 (m, 1H) – 38.6 ; 1.24 (d, 3H, 6.43 Hz) – 18.4.

**Compound (7).**

Yield 57%

LC-MS  $t_R$  0.90 min 96%  $m/z=354$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.10 (d, 1H, 7.60 Hz); 7.53 (s, 1H); 7.19 (s, 1H); 5.43 (m, 0.5H) – 92.0; 5.25 (m, 0.5H) – 94.6; 4.63 (t, 1H, 6.72 Hz) – 55.1; 4.49 (m, 1H) – 72.1; 4.44 (m, 1H) – 56.7; 4.07-3.99 (m, 1H) – 54.1; 3.95-3.90 (m, 1H) – 53.7; 2.40-2.25 (m, 1H) – 35.8; 2.25-2.17 (m, 1H) – 35.8; 1.87 (s, 3H) – 22.4; 1.31 (d, 3H, 6.14 Hz) – 18.1.

**Compound (8).**

Yield 70%

LC-MS  $t_R$  0.83 min 92%  $m/z=354$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 8.12 (d, 1H, 7.60 Hz); 7.55 (, 1H); 7.13 (s, 1H); 5.45 (m, 0.5H) – 92.0; 5.27 (m, 0.5H) – 94.5; 4.55 (t, 1H, 7.31 Hz) – 56.1; 4.40 (m, 1H) – 72.2; 4.31 (m, 1H) – 58.6; 4.29-3.85 (m, 1H) – 54.5; 4.22-3.71 (m, 1H) – 53.9; 2.45-2.03 (m, 1H) – 36.5; 2.39-1.88 (m, 1H) – 36.1; 1.85 (s, 3H) – 22.6; 1.26 (d, 3H, 6.14 Hz) – 18.2.

**Compound (9).**

Yield 62%

LC-MS  $t_R$  1.05 min 92%  $m/z=335.3$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 8.16 (d, 1H, 7.02 Hz); 7.39 (s, 1H); 7.06 (s, 1H); 6.02 (dd, 1H, 6.43 Hz, 2.05 Hz) – 128.3; 5.79 (dd, 1H, 6.43 Hz, 2.05 Hz) – 126.6; 4.92 (m, 1H) – 67.8; 4.60 (m, 1H) – 54.4; 4.52 (m, 1H) – 55.9; 4.46 (m, 1H) – 54.4; 4.42 (m, 1H) - 72.6; 1.85 (s, 3H) – 22.7; 1.31 (d, 3H, 6.14 Hz) – 18.0.

**Compound (10).**

Yield 54%

LC-MS  $t_R$  4.24min 96%  $m/z= 412$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 8.27 (d, 1H, 7.60 Hz); 7.41 (s, 1H); 7.27 (m, 4H) – 129.8, 128.6; 7.19 (m,1H) – 126.6; 7.09 (s, 1H);4.64 (t, 1H, 7.31 Hz) – 56.0; 4.50 (td, 1H, 9.35 Hz, 6.43 Hz) – 72.4; 4.21 (dd, 1H, 8.18 Hz, 5.55 Hz) – 60.3; 3.73 (m, 1H) – 47.9; 3.6 (m, 1H) – 47.9; 3.51 (s, 2H) – 42.2; 2.05 (m, 1H) – 29.9; 1.87 (m, 1H) – 25.0; 1.79 (m, 1H) – 25.0; 1.74 (m, 1H) – 29.9; 1.26 (d; 3H, 6.43 Hz) – 18.1.

**Compound (11).**

Yield 58%

LC-MS  $t_R$  4.46 min 98%  $m/z= 426$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 8.08 (d, 1H, 7.60 Hz); 7.36 (s, 1H); 7.23 (m, 4H) – 128.5; 7.16 (m,1H) – 126.0; 7.05 (s, 1H);4.64 (t, 1H, 7.31 Hz) – 55.7; 4.46 (m, 1H) – 72.3; 4.20 (dd, 1H, 8.77 Hz, 6.14 Hz) – 60.0; 3.74 (m, 1H) – 47.6; 3.6 (m, 1H) – 47.6; 2.80 (t, 2H, 8.77 Hz) – 31.2; 2.45 (t, 2H, 8.48 Hz) – 36.9; 2.06 (m, 1H) – 29.7; 1.86 (m, 1H) – 24.9; 1.78 (m, 1H) – 24.9; 1.74 (m, 1H) – 29.7; 1.23 (d; 3H, 6.43 Hz) – 18.0.

**Compound (12).**

Yield 56%

LC-MS  $t_R$  4.74 min 98%  $m/z= 440$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 7.99 (d, 1H, 7.31 Hz); 7.37 (s, 1H); 7.27 (m, 2H) – 128.6; 7.18 (m,3H) – 128.6, 126.2; 7.05 (s, 1H);4.62 (t, 1H, 7.02 Hz) – 55.9; 4.46 (m, 1H) – 72.6; 4.21 (dd, 1H, 8.18 Hz, 5.55 Hz) – 60.3; 3.79 (m, 1H) – 47.9; 3.63 (m, 1H) – 47.9; 2.54 (t, 2H, 8.18 Hz) – 35.1; 2.17 (td, 2H, 7.60 Hz, 3.22 Hz) – 35.0; 2.06 (m, 1H) – 30.0; 1.87 (m, 2H) – 25.1; 1.76 (m, 2H) – 27.5; 1.74 (m, 1H) – 30.0; 1.23 (d; 3H, 6.14 Hz) – 18.0.

**Compound (13).**

Yield 62%

LC-MS  $t_R$  5.01 min 97%  $m/z= 454$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 8.04 (d, 1H, 7.60 Hz); 7.36 (s, 1H); 7.26 (m, 2H) – 128.5; 7.17 (m,3H) – 128.5, 126.1; 7.02 (s, 1H);4.58 (t, 1H, 6.72 Hz) – 55.8; 4.44 (m, 1H) – 72.2; 4.21 (dd, 1H, 8.48 Hz, 5.26 Hz) – 60.0; 3.78 (m, 1H) – 47.7; 3.62 (m, 1H) – 47.7; 2.56 (t, 2H, 7.02 Hz) – 35.0; 2.16 (t, 2H, 7.02 Hz) – 35.0; 2.05 (m, 1H) – 29.8; 1.87 (m, 1H) – 24.9; 1.79 (m, 1H) – 24.9; 1.74 (m, 1H) – 29.8; 1.52 (m, 4H) – 30.8, 25.0; 1.27 (d; 3H, 6.43 Hz) – 18.1.

**Compound (14).**

Yield 65%

LC-MS  $t_R$  5.18 min 96%  $m/z= 488$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.48 (d, 1H, 7.89 Hz); 7.45 (s, 1H); 7.30 (m, 8H) – 128.9; 7.28 (m,2H) – 126.9; 7.14 (s, 1H); 5.22 (s, 1H) – 56.1; 4.72 (t, 1H, 6.72 Hz) – 55.7; 4.55 (m, 1H) – 72.0; 4.20 (dd, 1H, 7.89 Hz, 5.85 Hz) – 60.3; 3.73 (m, 1H) – 47.8; 3.59 (m, 1H) – 47.8; 2.06 (m, 1H) – 29.8; 1.87 (m, 1H) – 25.0; 1.76 (m, 1H) – 25.0; 1.72 (m, 1H) – 29.8; 1.23 (d; 3H, 6.72 Hz) – 18.0.

**Compound (15).**

Yield %

LC-MS  $t_R$  5.52 min 99%  $m/z= 516$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.60 Hz) - 120.3; 7.72 (d, 2H, 7.60 Hz) – 125.8; 7.46 (d, 2H, 6.72 Hz); 7.41 (s, 1H); 7.40 (t, 2H, 7.31 Hz) – 128.2; 7.32 (t, 2H, 7.60 Hz) – 127.6; 7.04 (s, 1H); 4.46 (m, 2H) – 72.5, 57.5; 4.25 (m, 4H) – 66.6, 46.9, 60.1; 3.78 (m, 1H) – 47.7; 3.64 (m, 1H) – 47.7; 2.06 (m, 1H) – 29.7; 1.88 (m, 2H) – 24.9; 1.77 (m, 1H) – 29.7; 1.32 (d, 3H, 5.85 Hz) – 18.0.

**Compound (16).**

Yield 38%

LC-MS  $t_R$  4.68 min 92%  $m/z= 438$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.15 (d, 1H, 6.43 Hz); 7.41 (s, 1H); 7.26 (m, 2H) – 128.9; 7.16 (m, 3H) – 128.9, 126.4; 7.05 (s, 1H); 6.02 (dd, 1H, 6.43 Hz, 2.05 Hz) – 128.4; 5.79 (dd, 1H, 6.14 Hz, 2.05 Hz) – 126.6; 4.93 (m, 1H) – 68.0; 4.62 (d, 1H, 2.05 Hz) – 54.7; 4.45 (d, 1H, 2.05 Hz) – 54.7; 4.39 (m, 1H) – 56.5; 4.34 (m, 1H) – 72.7; 2.53 (t, 2H, 7.02 Hz) – 35.4; 2.15 (td, 2H, 7.31 Hz, 2.34 Hz) – 35.2; 1.77 (q, 2H, 7.31 Hz) – 27.6; 1.32 (d, 3H, 6.14 Hz) – 18.2.

**Compound (17).**

Yield 57%

LC-MS  $t_R$  5.48 min 90%  $m/z= 516$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.88 (d, 2H, 7.31 Hz) – 120.3; 7.70 (m, 3H) – 125.9; 7.41 (m, 3H) – 128.3; 7.32 (t, 2H, 7.31) – 127.7; 7.05 (s, 1H); 6.00 (dd, 1H, 6.72 Hz, 2.05 Hz) – 128.3; 5.80 (dd, 1H, 6.72 Hz, 2.05 Hz) – 126.5; 4.96 (dd, 1H, 4.97 Hz, 2.05 Hz) – 67.8; 4.58 (m, 1H) – 54.4; 4.45 (m, 2H) – 54.4, 72.3; 4.34 (m, 1H) – 57.9; 4.21 (m, 3H) – 47.1, 66.7; 1.36 (d, 3H, 6.14 Hz) – 18.1.

**Compound (18).**

Yield 35%

LC-MS  $t_R$  4.58 min 99%  $m/z= 560$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 9.17 (s(b), 1H); 8.07 (d, 1H, 6.43 Hz); 7.37 (s, 1H); 7.04 (s, 1H); 6.92 (t, 4H, 8.77 Hz) – 128.2; 6.64 (t, 4H, 8.77 Hz) – 115.0; 6.00 (dd, 1H, 6.72 Hz, 2.34 Hz) – 128.2; 5.78 (dd, 1H, 6.43 Hz, 2.34 Hz) – 126.4; 4.91 (dd, 1H, 4.97 Hz, 2.05 Hz) – 67.6; 4.57 (d, 1H, 2.05 Hz) – 54.0; 4.45 (m, 2H) – 54.0, 55.7; 4.38 (m, 1H) – 72.4; 2.18 (m, 2H) – 37.5; 1.89 (m, 2H) – 31.3; 1.44 (s, 3H) – 27.7; 1.30 (d, 3H, 6.14 Hz) – 17.8.

**Compound (19).**

Yield 30%

LC-MS  $t_R$  5.48 min 90%  $m/z= 546$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 7.88 (d, 2H, 7.60 Hz) – 120.2; 7.73 (d, 2H, 7.31 Hz) – 125.8; 7.64 (d, 1H, 7.89 Hz); 7.41 (t, 2H, 7.60 Hz) – 128.1; 7.32 (t, 2H, 7.31 Hz) – 127.6; 7.32 (s, 1H); 6.80 (m, 2H); 4.26 (m, 1H) – 52.7; 4.25 (m, 1H) – 66.1; 4.19 (m, 1H) – 60.1; 4.22 (m, 2H) – 47.0; 3.58 (m, 2H) – 47.0; 3.39 (m, 2H) – 42.4; 1.99-1.79 (m, 2H) – 29.4; 1.92-1.83 (m, 2H) – 24.6; 1.70-1.54 (m, 4H) – 27.6.

### Compound (20).

Yield 30%

LC-MS t<sub>R</sub> 5.53 min 99% m/z= 560 (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz): δ (ppm) = 7.89 (d, 2H, 7.60 Hz) – 120.3; 7.73 (d, 2H, 7.31 Hz) – 125.7; 7.59 (d, 1H, 7.89 Hz); 7.40 (t, 2H, 7.60 Hz) – 127.92; 7.32 (t, 2H, 7.31 Hz) – 127.5; 7.27 (s, 1H); 6.94 (s, 1H); 6.86 (t, 1H, 6.14 Hz); 4.28 (m, 1H) – 66.1; 4.21 (m, 1H) – 60.1; 4.20 (m, 1H) – 52.9; 4.22 (m, 2H) – 47.0; 3.59 (m, 2H) – 47.0; 3.43-3.36 (m, 2H) – 42.6; 1.99-1.79 (m, 2H) – 29.3; 1.96-1.82 (m, 2H) – 24.9; 1.70-1.50 (m, 2H) – 30.3; 1.45 (m, 2H) – 30.9; 1.37 (m, 2H) – 22.2.

### Peptide synthesis for IC50 determination.

The peptide sequence was isolated from Tau protein around the Thr212-Pro213 epitope as described<sup>1</sup>. Peptide was obtained by solid phase synthesis with introduction of selectively phosphorylated phosphothreonine using appropriated building block (Novabiochem). The dansyl fluorescent probe (Sigma) was introduced selectively at the N-terminal extremity by reaction of dansyl chloride (5 eq.) and DIEA (10 eq.) in DMF and followed by a TFA cleavage. Peptide was isolated by precipitation of the TFA cleavage solution in a mixture of ether: heptane (v/v) and centrifugation. The peptide was solubilized in water and purified by RP-HPLC on a C<sub>18</sub> nucleosil column equilibrated in a 0.05% TFA aqueous solution (solution A). Separation of peptides in the crude mixture was carried out at 50°C at a flow rate of 2 mL/min using an acetonitrile linear 60 min-gradient from 0 to 50 % solution B (80% CH<sub>3</sub>CN in solution A). Homogeneous fractions as checked by RP-HPLC on a C<sub>18</sub> nucleosil column and MALDI-TOF MS were pooled and lyophilized.

### Protein expression and purification.

The WW domain carrying a N-terminus His<sub>6</sub> tag was produced in *E. coli* BL21(DE3) strain, carrying the pET28a plasmid under control of T7 promoter with a sequence encoding the His<sub>6</sub> tag. Cells were grown at 37°C in a LB medium containing kanamycine (20 mg/L) until O.D.600 reached about 0.6, allowing rapid growth to high cell density. Then cells were harvested by centrifugation. The pellet was resuspended in M9 medium concentrated 2 times containing 4 g/L glucose, 1 g/L <sup>15</sup>NH<sub>4</sub>Cl, 1 mM MgSO<sub>4</sub>, 20 mg/L kanamycine and MEM vitamin cocktail (Sigma). The mixture was incubated at 37°C for 1 hour and induction was performed with 0.5 mM IPTG (final concentration) at 31°C for 3 hours. Cells were centrifuged and the pellet was resuspended in 50 mM Na<sub>2</sub>HPO<sub>4</sub> / NaH<sub>2</sub>PO<sub>4</sub> pH 8.0, 300 mM NaCl buffer (Buffer A) containing 10 mM imidazole, 1 mM DTT, 0.1 % NP40 and complemented with a protease inhibitor cocktail (Roche). The soluble lysate was obtained by sonication and centrifugation. The soluble extract was charged on a nickel-nitrilotriacetic affinity column (Chelating Sepharose Fast Flow, Amersham). Unbound proteins were washed with 20 mM imidazole in Buffer A and the WW domain was eluted with 250 mM imidazole in Buffer A (pH 7.8).

### NMR screening and dissociation constant calculation.

NMR experiments were performed on a Bruker DMX 600 MHz spectrometer at 293K. <sup>15</sup>N-labeled WW spectra were recorded on a 50μM sample in a 50 mM deuterated Tris-HCl (pH=6.8) buffer containing 50 mM NaCl, 5 mM DTT and 2.5 mM EDTA in H<sub>2</sub>O 90% : D<sub>2</sub>O 10%. The solution of WW was added on lyophilized aliquots of ligand prepared from a concentrated solution. Quantification of chemical shift perturbations was calculated with the Equation 1.

### Equation 1

$$\Delta\delta(\text{ppm}) = \sqrt{\Delta\delta(^1\text{H})^2 + 0.2 \Delta\delta(^{15}\text{N})^2}$$

Dissociation constant was calculated by fitting on a graphical representation the theoretical curve obtained from Equation 2 with experimental points.

*Equation 2*

$$\Delta\delta(\text{ppm}) = 0.5\Delta\delta_{\text{max}} \left( 1 + X + \frac{K_D}{[\text{protein}]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[\text{protein}]_0} \right)^2 - 4X} \right)$$

X is the ligand on protein molar ratio,  $[\text{protein}]_0$  the initial concentration of protein,  $\Delta\delta$  the chemical shift variation calculated as described above and  $\Delta\delta_{\text{max}}$  the  $\Delta\delta$  value at the protein saturation level.

#### **FRET competition experiments for $IC_{50}$ determination.**

FRET experiments were achieved on a PTI fluorescence spectrometer by exciting tryptophane residues of WW domain at 295 nm. As for NMR titrations, a 20  $\mu\text{M}$ -solution of WW in protonated Tris 50mM pH 6.8 buffer containing 50mM NaCl and 5mM DTT/EDTA and variable concentrations of dansylated peptide was added on lyophilized aliquots of ligand. The dissociation constant of the WW - peptide complex was evaluated to 100  $\mu\text{M}$  as described<sup>1</sup>.  $IC_{50}$  determination was calculated as described for dissociation constants determined by NMR with the Equation 2 by replacing  $\Delta\delta$  by the inhibition percentage given by Equation 3 and  $K_D$  by  $IC_{50}$ .

*Equation 3*

$$\% \text{ Inhibition} = \frac{I_0^{334} - I^{334}}{I_0^{334} - I_{100}^{334}}$$

where  $I^{334}$  is the fluorescence intensity at 334 nm upon ligand addition,  $I_0^{334}$ , the fluorescence intensity without inhibition and  $I_{100}^{334}$ , the fluorescence intensity at maximal inhibition.

- (1) Smet, C.; Sambo, A. V.; Wieruszkeski, J.-M.; Leroy, A.; Landrieu, I. et al. The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau. *Biochemistry* **2004**, *43*, 2032-2040.

## **CONCLUSION**



## Conclusion.

### **Pin1, une peptidyl-prolyl *cis/trans* isomérase phospho-spécifique ?**

Cette question en cache deux en réalité. Elle remet en cause la phospho-spécificité de Pin1 d'une part et d'autre part son activité enzymatique de peptidyl-prolyl *cis/trans* isomérase. Pin1 a été décrite par Lu et collaborateurs comme une protéine régulatrice capable de provoquer une réorganisation conformationnelle au sein des protéines substrats après leur phosphorylation par les kinases dirigées contre les motifs Ser/Thr-Pro. Ces mêmes motifs sont reconnus spécifiquement par Pin1 de manière phospho-dépendante. Pin1 a ainsi été impliquée dans un mécanisme de régulation de la fonction des protéines constitué d'une première étape de phosphorylation suivi d'un réarrangement conformationnel par isomérisation de la liaison phosphoSer/Thr-Pro. Il est en effet tentant de conclure à un rôle dans la réorganisation structurale des substrats car on n'a jamais détecté d'autres modifications covalentes, comme la déphosphorylation par exemple. De plus, *in vitro*, sur base de substrats peptidiques modèles, on a montré une activité enzymatique sur l'isomérisation des prolines, généralement de manière indirecte par la sensibilité de protéases telles que la chymotrypsine à hydrolyser des liaisons peptidiques comportant une proline en conformation *trans*, dont l'activité est stimulée par Pin1. Cependant, des méthodes d'observations directes par RMN de la catalyse de l'interconversion *cis/trans* des liaisons phosphoSer/Thr-Pro ont également permis de mettre en évidence une telle activité de Pin1 et de son domaine catalytique isolé (Landrieu, De Veylder et al. 2000; Smet, Sambo et al. 2004). D'autre part, les modèles peptidiques ont montré une phospho-spécificité de Pin1 au travers des structures cristallographiques (Ranganathan, Lu et al. 1997; Verdecia, Bowman et al. 2000) ou des structures par RMN (Wintjens, Wieruszeski et al. 2001), des études de *binding* (Landrieu, Wieruszeski et al. 2002; Jacobs, Saxena et al. 2003; Smet, Sambo et al. 2004) ou des études de compétition avec le phosphate ou le sulfate (Landrieu, Wieruszeski et al. 2002; Bayer, Goettsch et al. 2003).

Nos études sur les substrats peptidiques contribuent à préciser le rôle de Pin1 comme une PPIase phospho-spécifique. Nous avons utilisé comme substrat la protéine tau, dont les interactions avec Pin1 sont encore mal comprises, notamment au niveau de l'impact biologique de cette interaction dans la fonction cellulaire neuronale de la protéine tau. La protéine tau joue un rôle fondamental dans l'architecture neuronale avec la stabilisation et la

polymérisation des microtubules. Sa fonction est régulée négativement par la phosphorylation, notamment au niveau de motifs Ser/Thr-Pro reconnus par Pin1, qui régulerait positivement l'activité de tau. Il a été proposé que Pin1 pouvait se lier à la protéine tau (hyper)phosphorylée dans les PHF, une des caractéristiques moléculaires de la maladie d'Alzheimer, et restaurer la fonction biologique de tau phosphorylée, à la fois par la liaison du domaine WW et par l'intervention du domaine catalytique qui induirait un changement conformationnel dans la protéine. Le mode de reconnaissance de la protéine tau par Pin1 montre une sélectivité de reconnaissance autour du site unique phosphoThr231, reconnu par l'anticorps AT180. Cette observation nous a semblé biaisée par les données cellulaires montrant la Thr231 comme site majeur de phosphorylation de tau dans les PHF et l'absence d'investigations au niveau d'autres motifs phosphorylés spécifiques des PHF.

Nos études de *binding* réalisées par RMN et par fluorescence ont permis de montrer une spécificité pour les substrats phosphorylés et ont également mis en lumière le rôle essentiel joué par le domaine WW dans la liaison du substrat. Les affinités, relativement faibles (de l'ordre de 100  $\mu$ M à 1 mM), déterminées pour les différents peptides phosphorylés de tau sont gouvernées par l'interaction des substrats avec le domaine WW, le domaine catalytique ne participant pas clairement à la liaison, ou de manière très transitoire. Ces données indiquent qu'aucun phospho-épitope n'est reconnu sélectivement par Pin1. L'analyse des équilibres conformationnels intrinsèques de ces différentes liaisons phosphoSer/Thr-Pro nous a permis d'exclure une restriction conformationnelle due à une proline majoritairement en conformation *cis* qui limiterait la liaison du domaine WW *trans*-spécifique (Wintjens, Wieruszeski et al. 2001). D'autre part, le rôle moléculaire des deux domaines de Pin1 n'a pas été totalement élucidé, notamment quant à la participation du domaine catalytique dans la liaison. Au travers des substrats peptidiques, nous avons mis en évidence un rôle ambigu du domaine catalytique dans la liaison des sites de phosphorylation de la protéine tau. Ainsi nous avons montré que l'épitope AT100 (phosphoThr212/Ser214), un phospho-épitope spécifique de la protéine tau pathologique dans la maladie d'Alzheimer, était reconnu par Pin1 avec une affinité similaire à celle de l'épitope pathologique TG-3 incluant le site Thr231 phosphorylé. Mais la présence du domaine catalytique diminue l'affinité par rapport à la liaison du domaine WW seul, ce qui n'est pas le cas pour l'épitope mono-phosphorylé en Thr212. D'autre part, nous avons montré, toujours au niveau de substrats peptidiques, une activité catalytique sur l'isomérisation *cis/trans* des liaisons phosphoSer/Thr-Pro de manière directe par RMN. Contrairement à la liaison des phospho-épitopes, nous avons montré une sélectivité de

l'activité enzymatique pour l'épitope AT100. Nous avons effectivement mesuré une accélération de l'interconversion *cis/trans* de la liaison phosphoThr212-Pro213 en présence de Pin1 ou de son domaine catalytique isolé, alors qu'aucune activité n'a été détectée pour les liaisons phosphoThr217-Pro218 et phosphoThr231-Pro232. Ces liaisons étant majoritairement en *trans* dans ces fragments peptidiques flexibles non structurés, dont la taille est comprise entre 15 et 20 résidus, nous nous sommes demandés quel était l'intérêt de catalyser l'interconversion *cis/trans* dans ce contexte. Même si le domaine WW jouait un rôle dans la stabilisation des conformations *trans*, quel en serait l'intérêt puisque les motifs phosphoSer/Thr-Pro dans ces peptides sont déjà à plus de 90% en *trans* et qu'aucune protéine *cis*-spécifique n'a été décrite pour le moment. En effet, il a été montré que Pin1 stimulait l'activité phosphatase de PP2A, également *trans*-spécifique, sur la déphosphorylation de Cdc25 et tau de manière catalytique en agissant par une modification conformationnelle du substrat (Zhou, Kops et al. 2000; Stukenberg and Kirschner 2001). Nos données structurales ne sont pas en accord avec le rôle de l'isomérisation catalytique sur l'activité de PP2A dans la mesure où la conformation *trans* des liaisons n'est pas ici un facteur limitant. Nous avons envisagé que le modèle peptidique puisse présenter, en terme de taille du substrat, une limite significative dans les phénomènes que nous cherchons à déterminer.

En effet, les observations faites sur les peptides phosphorylés ne sont peut-être que des artefacts à double titre. D'abord, un peptide pourrait ne pas rendre compte des équilibres conformationnels qui ont lieu au sein de la protéine entière qui peut comporter des prolines majoritairement ou totalement en *cis*. Ce cas de figure est peu probable avec la protéine tau, qui est réputée pour être une protéine non structurée hautement flexible. On peut donc raisonnablement supposer que, dans ce contexte, le peptide est un bon modèle pour une protéine ne possédant pas de structure globulaire stable, voire même secondaire (ça reste à déterminer). Mais une structure résiduelle n'est pas à écarter, ni une organisation plus globale. Il faut noter que l'anticorps Alz50 reconnaissant un épitope discontinu de la protéine tau a été mis en évidence et permettrait de détecter une conformation particulière de la protéine tau avec la double reconnaissance de la partie N-terminale et de la région d'interaction avec les microtubules, localisée dans la partie C-terminale. Cette conformation ne pourrait être que transitoire mais nous devons prendre en compte une telle possibilité de structure, qui pourrait impliquer la présence éventuelle de prolines en conformation *cis*. Deuxièmement, l'activité isomérase mise en évidence au travers des substrats peptidiques pourrait également être un artefact. En effet, la structure de Pin1 en complexe avec le dipeptide Ala-Pro montre une

liaison du substrat en conformation *cis* dans le site actif et il a été proposé que la conformation la plus favorable est la conformation *syn-90* mimant l'état de transition plan. De plus, le mécanisme proposé fait intervenir la Cys113 qui, accompagnée des résidus Ser154, His59 et His157, lierait de manière covalente et réversible le substrat avec attaque nucléophile du carbonyle, lequel passerait par un état tétravalent, facilitant la rotation de la liaison peptidique (Ranganathan, Lu et al. 1997). Ces données indiquent que le site actif assurant la fonction « isomérase » ne pourrait intervenir que dans la stabilisation de la conformation *cis* et serait un simple site de liaison de faible affinité. De plus, le site actif de Pin1 serait capable d'adaptation plus large avec la reconnaissance des motifs contenant une proline précédée d'un résidu acide (Ranganathan, Lu et al. 1997).

Nous avons élargi notre étude à un fragment peptidique plus long (de 40 résidus acides aminés), comprenant presque la totalité de la région riche en proline de la protéine tau, incluant deux sites de reconnaissance majeurs de Pin1, l'épitope Thr212/Ser214 diphosphorylé (AT100) et l'épitope Thr231 mono-phosphorylé (AT180). Cette région n'est pas impliquée directement dans la liaison des microtubules mais joue un rôle régulateur dans cette fonction de la protéine tau. Nous envisageons la possibilité de trouver des modifications conformationnelles par rapport aux peptides courts qui pourraient impliquer une modulation supplémentaire (en plus de la séquence primaire autour des sites de phosphorylation) dans la reconnaissance des divers sites de phosphorylation par Pin1 ou qui expliquerait le rôle positif de Pin1 sur la déphosphorylation de la Thr231 par PP2A. Nous avons également émis l'hypothèse que plusieurs sites de reconnaissance au sein d'un même substrat pourraient réguler la liaison de Pin1. En effet, un facteur important dans la structure de Pin1 est la présence d'un fragment flexible d'une dizaine de résidus reliant les deux domaines, qui confère aux domaines WW et catalytique de Pin1 une relative autonomie. Cette indépendance moléculaire peut se traduire au niveau fonctionnel soit par l'interaction de chacun des domaines avec deux épitopes différents au sein d'un même substrat, que l'on peut espérer alors mettre en évidence avec le fragment de 40 résidus, soit par l'interaction de chacun des domaines avec des épitopes localisés dans deux substrats distincts, qui conférerait alors à Pin1 un rôle de médiateur. Dans le premier cas, le domaine WW servirait de point d'ancrage sur le substrat pour augmenter la concentration locale de l'enzyme autour d'un autre phospho-épitope. Cette interprétation est assez logique dans la mesure où, au niveau cellulaire, les deux domaines de Pin1 sont nécessaires à son activité, sauf dans le cas où le domaine catalytique est surexprimé, ce qui s'apparente au phénomène de concentration locale engendrée par la

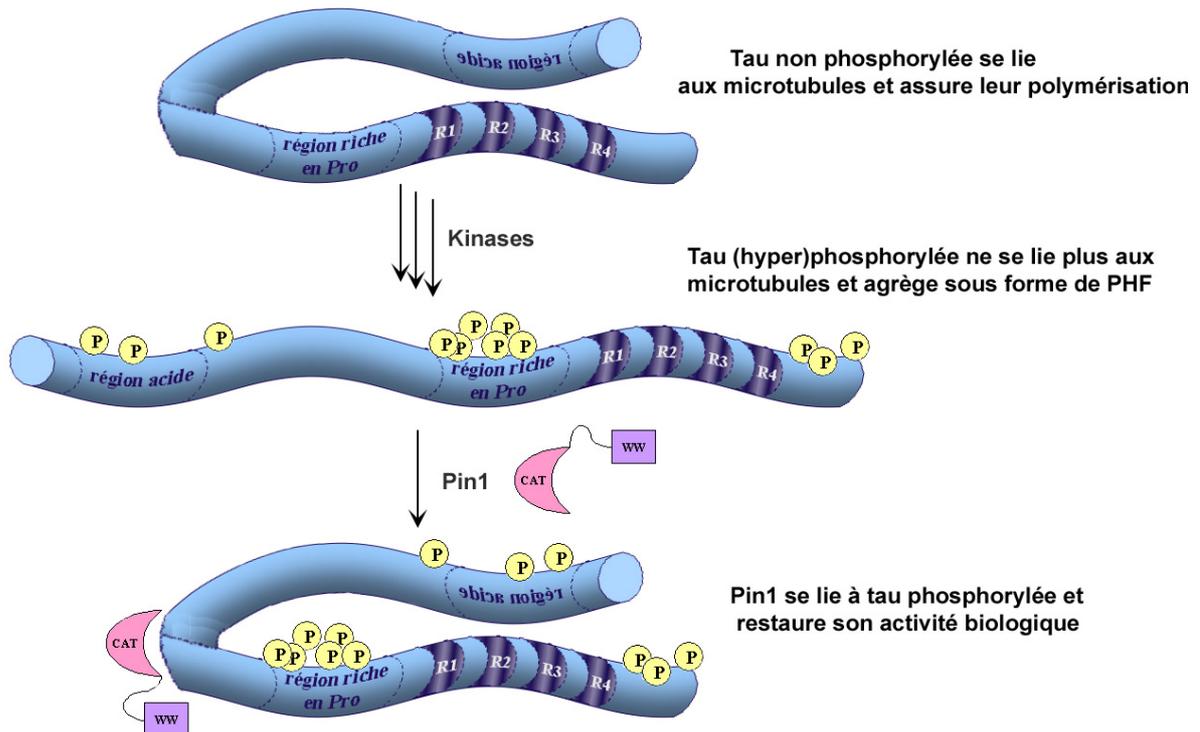
liaison du domaine WW au substrat. Dans le deuxième cas, le domaine « catalytique » serait en fait un deuxième site de liaison de plus faible affinité, et Pin1, en se liant à deux substrats distincts par l'intermédiaire de ses deux domaines, contribuerait à augmenter la concentration locale des deux partenaires l'un pour l'autre dans un complexe macromoléculaire où Pin1 exercerait un rôle d'adaptateur. Cette deuxième hypothèse est plus difficile à mettre en œuvre. Nous avons envisagé l'investigation de la première hypothèse avec notre peptide modèle de 40 résidus de la région riche en prolines. L'analyse conformationnelle n'a pas montré de différence dans les équilibres *cis/trans* des différentes prolines par rapport aux peptides courts. Si, au niveau de la liaison du substrat, les affinités mesurées ne permettent pas de trancher en faveur d'une liaison coopérative des deux domaines via les deux sites présents dans le peptide, l'investigation de l'activité catalytique a révélé des aspects très intéressants sur le mode de fonctionnement moléculaire de Pin1 vis-à-vis de son substrat. D'abord, le domaine WW contribue à augmenter significativement l'activité enzymatique de Pin1. Ensuite, la présence d'un deuxième site de reconnaissance pour Pin1 au sein du même substrat diminue clairement l'activité *cis/trans* isomérase mesurée au niveau du site Thr212/Ser214, en tout cas de manière plus marquée que lorsque les deux phospho-épitopes sont localisés sur deux molécules distinctes. Une explication possible réside en la liaison d'un site par le domaine WW et la liaison du deuxième site par le domaine catalytique sous l'effet d'une concentration locale, qui oriente la liaison du domaine catalytique vers un épitope présent sur le même substrat plutôt que sur une autre molécule. Nous avons ainsi montré que, même si aucune activité catalytique n'était mesurée en Thr231, la simple liaison de ce site par le domaine « catalytique » pourrait le détourner du site Thr212/Ser214 où il exerce une fonction, semble-t-il, enzymatique sur l'isomérisation de la liaison phosphoThr212-Pro213. Ces données montrent que le domaine catalytique peut jouer un rôle dans la liaison des substrats et que la présence d'un second site de reconnaissance pour Pin1 peut moduler l'activité isomérase. Ces résultats suggèrent l'importance des phosphorylations multiples dans un mécanisme qui gouvernerait la liaison de Pin1 à ses substrats. Il a en effet été montré, dans le contexte plus général de la régulation du cycle cellulaire au niveau de la transition G1/S, que la présence de plusieurs sites de phosphorylation d'affinités modérées pouvait moduler plus efficacement qu'un site unique de phosphorylation de haute affinité, la reconnaissance d'un substrat et permettait de contrôler plus finement la réponse consécutive au processus de phosphorylation (Deshaies and Ferrell 2001). Dans le cas de Pin1, des phosphorylations multiples pourraient être invoquées dans le processus de reconnaissance puisque la protéine tau possède 17 sites de type Ser/Thr-Pro localisés dans diverses régions de la protéine et essentiellement dans la

région riche en prolines. Nous avons dénombré plusieurs sites de phosphorylation qui pouvaient jouer un rôle médiateur dans l'interaction avec Pin1 comme les sites T212, T217, T231, S235, S396, S422, avec des modulations d'affinités allant de 100  $\mu$ M à 1 mM. On peut envisager qu'il existe encore d'autres sites de liaison pour Pin1 qui n'ont pas été pris en compte lors du *mapping* des sites par le biais des peptides phosphorylés. D'autre part, le fragment de 40 résidus n'est peut-être pas représentatif du même fragment compris dans la protéine totale, au même titre que les peptides courts. Il a été montré que la région riche en prolines pouvait moduler la liaison de tau aux microtubules et que la phosphorylation de cette région à elle seule suffirait à abolir cette liaison (Eidenmuller, Fath et al. 2000; Eidenmuller, Fath et al. 2001). Le mécanisme moléculaire qui se cache derrière cette régulation n'est pas connu. Nous avons constaté une limite du modèle peptidique dans l'exploration des mécanismes qui ont lieu à l'échelle de la protéine entière, notamment dans la recherche du mode d'action de Pin1 pour restaurer la fonction de tau phosphorylée. Bien que nous n'ayons pas mis en évidence de modifications conformationnelles lors du passage des peptides courts au peptide long de 40 résidus, ceci n'exclut pas que, dans le contexte protéique, il existe des prolines en conformation *cis*.

Qu'en est-il de l'interaction entre Pin1 et ses substrats protéiques ? Même si, pour bon nombre d'entre eux, la phosphorylation a été reportée comme un élément essentiel gouvernant la reconnaissance par Pin1, la découverte initiale de Pin1 par interaction avec la protéine NIMA n'impliquait aucune phosphorylation. En revanche, dans le cas de la protéine tau, la phosphorylation est une condition essentielle puisque la protéine non phosphorylée n'interagit pas. L'obstacle de la phosphorylation nous a conduits, dans un premier temps, à explorer les interactions de Pin1 avec un mutant de la protéine tau Tau352E10 où la plupart des sites de phosphorylation majeurs ont été mutés en glutamate (au nombre de 10 sites Ser/Thr dont 7 sites Ser/Thr-Pro), hormis les sites Ser214 et Ser262 interférant directement la liaison de tau aux microtubules. Ce mutant a montré certaines caractéristiques moléculaires et fonctionnelles de la protéine tau hyperphosphorylée comme une mobilité électrophorétique identique, une tendance à former des filaments ainsi qu'une incapacité à s'associer à PP2A et à promouvoir l'assemblage des microtubules (Eidenmuller, Fath et al. 2000). Nos résultats, pour le moins surprenants, ont montré que le site actif du domaine « catalytique » de Pin1 pouvait lier le mutant de manière très spécifique, notamment avec la région impliquée dans la liaison du phosphate, avec une relativement haute affinité. En revanche si le site de liaison de Pin1 est assez bien défini, le site reconnu au niveau de Tau352E10 n'est pas identifié. Il peut

s'agir d'un motif Glu-Pro particulier qui n'existe pas dans la protéine tau sauvage. Il faut également souligner que, dans le contexte d'un peptide mutant, le motif Glu231-Pro232 n'est pas reconnu par Pin1. Dans la protéine tau entière, le mode de reconnaissance est peut-être plus « sophistiqué », en faisant intervenir les résidus environnants dans la séquence ou à proximité dans l'espace si on considère que la protéine peut adopter une conformation particulière qui reste à définir. D'autre part, l'absence de phosphate semble indiquer que Pin1, ou plus précisément son domaine catalytique, présente des capacités à adapter un autre motif que les motifs phosphoSer/Thr-Pro. Cependant, le mutant se lie au site de liaison du phosphate dans le site actif de manière très similaire au phosphate lui-même. Les premières données que nous avons obtenues semblent indiquer une modification conformationnelle du mutant Tau352E10, qui adopterait une structure plus « étendue » que la protéine sauvage. La liaison du site actif de Pin1, perturbe Tau352E10 d'une manière assez inhabituelle car nous avons pu identifier des interactions au niveau de diverses régions de la protéine, qui ne correspondent pas toujours au site de mutation. Pin1 étant présente en quantités stœchiométriques, ces perturbations très étendues pourraient argumenter en faveur d'un réarrangement conformationnel plus global du substrat, plutôt qu'une simple liaison. Les perturbations des déplacements chimiques du mutant Tau352E10 semblent indiquer, par ailleurs, un retour vers une conformation semblable à celle de la protéine sauvage. Il est, pour le moment, difficile de dissocier l'effet de liaison de Pin1 à Tau352E10 d'un éventuel effet conformationnel. Une telle réorganisation structurale du substrat au niveau des régions protéiques flexibles supposerait un rôle de chaperonne pour Pin1, permettant la restauration de la conformation initiale perturbée par la phosphorylation, effet ne nécessitant pas une catalyse enzymatique de l'interconversion *cis/trans*. Dans ce cas, il ne s'agirait pas d'accélérer l'isomérisation *cis/trans* d'une proline, mais de stabiliser une conformation, le temps d'acquiescer la structure finale stable. On peut envisager qu'avec un substrat phosphorylé, le domaine WW pourrait être engagé dans une interaction supplémentaire et intervenir au niveau moléculaire pour lier une région distincte de celle impliquée dans l'interaction avec le site actif. On pourrait alors s'attendre à un effet plus marqué sur la structure de la protéine phosphorylée, dans le sens d'un « repliement » grâce à la liaison de deux régions distantes dans la protéine qui serait maintenue dans une proximité spatiale par l'intermédiaire de la contrainte imposée par le *linker* inter-domaine de Pin1. Toutes ces hypothèses devront faire l'objet d'une étude plus approfondie, qui passe nécessairement par la phosphorylation de la protéine tau et l'attribution de la protéine et des sites de phosphorylation par RMN pour définir les motifs de reconnaissance intervenant dans la

liaison de Pin1. D'autres expériences RMN, comme l'expérience de STD (*Saturation Transfer Difference*), permettraient, à condition que l'échange soit suffisamment rapide, de localiser le (les) site(s) de liaison de Pin1 et de se prémunir de l'effet « conformationnel » sur les résonances du substrat. Avec des expériences d'alignement en milieux orientés, on peut espérer observer des orientations distinctes pour différentes régions de la protéine tau phosphorylée ou non, en présence ou en l'absence de Pin1.



Pour conclure sur les données moléculaires concernant le mode d'action de Pin1 sur ses substrats, on peut désormais soulever deux faits contradictoires avec les hypothèses émises précédemment sur le rôle biologique de Pin1. Premièrement, il n'est pas certain que le site actif de Pin1 soit totalement phospho-spécifique car nous avons montré que ce site peut reconnaître et participer à la liaison de substrats non phosphorylés au travers d'un motif dont la séquence reste à déterminer. Deuxièmement, l'effet PPIase n'a toujours pas été montré sur un substrat de nature protéique mais nos données préliminaires sur la protéine tau nous permettent d'envisager la possibilité d'un rôle de chaperonne, qui reste également à confirmer, notamment avec un substrat phosphorylé car le domaine WW, contrairement au domaine « catalytique », semble lier de manière très spécifique les substrats phosphorylés.

## **Pin1, une cible thérapeutique dans le traitement de la maladie d'Alzheimer ?**

Il a été proposé que Pin1 pourrait jouer un rôle dans la réactivation du cycle cellulaire des neurones post-mitotiques. Nous avons commencé une stratégie visant à inhiber le domaine WW de Pin1 qui est indispensable à sa fonction biologique et qui a récemment été impliqué dans la régulation de la liaison de Pin1 à ses substrats par le processus de phosphorylation. L'inhibition de la liaison des substrats provoque une délocalisation de Pin1 du noyau vers le cytoplasme avec un profil de localisation cellulaire semblable à celui du domaine catalytique isolé. Le domaine WW jouerait donc un rôle important dans la régulation de l'activité de Pin1. Inhiber le domaine WW avec de petites molécules reviendrait à produire une forme inactive de Pin1. Sur base du dipeptide phosphoThr-Pro, qui est le motif minimal d'interaction avec le domaine WW de Pin1, présentant une affinité modeste de 100  $\mu$ M, nous avons introduit de faibles modifications structurales avec la conservation de la dualité partie apolaire/ partie polaire du composé référence. Nous avons synthétisé des molécules basées sur la structure du dipeptide avec, soit des mimes de la proline couplés à la phosphothréonine, soit des têtes polaires mimant le groupement phosphate en conservant la proline comme partie apolaire. Les modifications en terme d'affinité pour le domaine WW sont plutôt négatives sauf dans le cas de la 4-*trans*-fluoro-proline et de la 3,4-dehydro-proline, où les améliorations restent modérées. En ce qui concerne les têtes polaires, aucune n'a la potentialité du phosphate dans la liaison de la boucle flexible  $\beta$ 1/ $\beta$ 2 du domaine WW. Les pertes d'affinité pour les composés résultants sont même assez drastiques et limitent grandement le remplacement du phosphate. Cependant, le remplacement de cette tête polaire est crucial à double titre : le phosphate est naturellement très vite hydrolysé dans les cellules par l'action des phosphatases et le passage des membranes cellulaires est fortement compromis avec des composés très polaires tels que les mimes du dipeptide phosphoThr-Pro.

Les molécules comportant, en plus de la partie basée sur le dipeptide phosphoThr-Pro, un groupement hydrophobe, présentent toutes une amélioration globale de l'affinité. Dans certains cas, la probable formation d'un « collapse hydrophobe » au sein du composé même, contraint le ligand à un réarrangement structural préalable avant sa liaison au domaine WW et entraîne un gain d'affinité plus modéré par rapport à ce que l'on pouvait espérer. Des molécules plus rigides pourraient pallier ces limitations et ouvrir de nouvelles voies dans l'approche de l'inhibition du domaine WW de Pin1.

En revanche, si les composés sont capables de se lier au domaine WW et d'entrer en compétition avec un substrat peptidique issu de la protéine tau (dont la séquence est centrée autour de l'épitope phosphoThr212-Pro213), on remarque également *in vitro* que les composés se lient au site actif du domaine catalytique de Pin1, avec probablement une plus faible affinité. *In vivo*, dans un modèle cellulaire mimant un neurone post-mitotique surexprimant les protéines Pin1 et tau, les molécules semblent avoir une activité détectable, pour des concentrations d'environ 10  $\mu$ M et de manière dose-dépendante, et provoquer une diminution du niveau de cycline D1. Des données préliminaires, non publiées, montrent également que les composés entraînent une augmentation du niveau de phosphorylation de tau en Thr231. Mais, actuellement, le mode d'action de ces composés n'est pas encore élucidé. En effet, il a été montré par Buée et collaborateurs que le niveau d'expression de Pin1 était corrélé au niveau de cycline D1 également dans les neurones (Hamdane, Smet et al. 2002). Mais rien n'indique que les composés agissent par inhibition de la fonction de Pin1, en tout cas de la fonction du domaine WW uniquement. Nous avons d'ailleurs mis en évidence *in vitro* une interaction des composés avec le site actif de Pin1. Concernant le niveau de phosphorylation en Thr231, Lu et collaborateurs ont montré que Pin1 stimulait la déphosphorylation de tau et Cdc25 par PP2A de manière catalytique (Zhou, Kops et al. 2000; Stukenberg and Kirschner 2001). De nouveau, il est envisageable d'expliquer l'augmentation de la phosphorylation en Thr231 par une inhibition du site catalytique de Pin1 ou par l'inhibition directe de phosphatases telles que PP2A. Ces données préliminaires indiquent toutefois que les ligands sont capables de franchir les barrières cellulaires, même en présence d'un groupement phosphate. On peut également, dans ce contexte, envisager de tester ces molécules sur des lignées de cellules cancéreuses, puisque Pin1 est surexprimée dans le cancer et joue un rôle dans l'activation de divers processus d'oncogénèse. De plus, Pin1 est retrouvée majoritairement sous sa forme active non phosphorylée dans les cellules cancéreuses du sein.

## Références.

- Albert, A., S. Lavoie, et al. (1999). "A hyperphosphorylated form of RNA polymerase II is the major interphase antigen of the phosphoprotein antibody MPM-2 and interacts with the peptidyl-prolyl isomerase Pin1." *J Cell Sci* **112** (15): 2493-500.
- Alt, J. R., J. L. Cleveland, et al. (2000). "Phosphorylation-dependent regulation of cyclin D1-dependent cellular transformations." *Genes Dev* **14**: 3102.
- Andreotti, A. H. (2003). "Native state proline isomerization : an intrinsic molecular switch ?" *Biochemistry* **42**(32): 9515-9524.
- Arkin, M. R. and J. A. Wells (2004). "Small-molecule inhibitors of protein-protein interactions: progressing towards the dream." *Nat Rev Drug Discov* **3**(4): 301-17.
- Ayala, G., D. Wang, et al. (2003). "The prolyl isomerase Pin1 is a novel prognostic marker in human prostate cancer." *Cancer Res* **63**(19): 6244-51.
- Barghorn, S., P. Davies, et al. (2004). "Tau paired helical filaments from Alzheimer's disease brain and assembled in vitro are based on beta-structure in the core domain." *Biochemistry* **43**(6): 1694-703.
- Bayer, E., S. Goetsch, et al. (2003). "Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding." *J Biol Chem* **278**(28): 26183-93.
- Berliner, J. L., J. Grunwald, et al. (1982). "A novel reversible spin-label : papain active site labeling and inhibition." *Anal Biochem* **119**(2): 450-455.
- Berry, R. W., A. Abraha, et al. (2003). "Inhibition of tau polymerization by its carboxy-terminal caspase cleavage fragment." *Biochemistry* **42**(27): 8325-31.
- Biernat, J., N. Gustke, et al. (1993). "Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: distinction between PHF-like immunoreactivity and microtubule binding." *Neuron* **11**(1): 153-63.
- Braisted, A. C., J. D. Oslob, et al. (2003). "Discovery of a potent small molecule IL-2 inhibitor through fragment assembly." *J Am Chem Soc* **125**(13): 3714-5.
- Buee, L., T. Bussiere, et al. (2000). "Tau protein isoforms, phosphorylation and role in neurodegenerative disorders." *Brain Res Brain Res Rev* **33**(1): 95-130.
- Bundi, A. and K. Wuthrich (1977). "1H NMR titration shifts of amide proton resonances in polypeptide chains." *FEBS Lett* **77**(1): 11-4.
- Bussel, R., Jr. and D. Eliezer (2001). "Residual structure and dynamics in Parkinson's disease-associated mutants of  $\alpha$ -synuclein." *J. Biol. Chem.* **276**(49): 45996-46003.
- Carmel, G., E. M. Mager, et al. (1996). "The structural basis of monoclonal antibody Alz50's selectivity for Alzheimer's disease pathology." *J Biol Chem* **271**(51): 32789-32795.
- Cavanagh, J. and M. Rance (1993). *Annu Rep NMR Spec* **27**: 1-58.
- Chao, S. H., A. L. Greenleaf, et al. (2001). "Juglone, an inhibitor of the peptidyl-prolyl isomerase Pin1, also directly blocks transcription." *Nucleic Acids Res* **29**(3): 767-73.
- Chau, M.-F., Radeke, M., de Inés, C., Barasoain, I., Kohlstaedt, L., and Feinstein, S. (1998). "The microtubule-associated protein tau cross-link to two distinct sites on each  $\alpha$  and  $\beta$  tubulin monomer via separate domains." *Biochemistry* **37**: 17692-17703.
- Chen, H. I., and Sudol, M. (1995). "The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules." *Proc. Natl. Acad. Sci.* **92**: 7819-7823.
- Crenshaw, D. G., J. Yang, et al. (1998). "The mitotic peptidyl-prolyl isomerase, Pin1, interacts with Cdc25 and Plx1." *Embo J* **17**(5): 1315-27.

- Dalvit, C., M. Flocco, et al. (2002). "Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures." Comb Chem High Throughput Screen **5**(8): 605-11.
- Daly, N. L., R. Hoffmann, et al. (2000). "Role of phosphorylation in the conformation of tau peptides implicated in Alzheimer's disease." Biochemistry **39**(30): 9039-46.
- Daughdrill, G. W., M. S. Chadsey, et al. (1997). "The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28." Nat Struct Biol **4**(4): 285-91.
- Davis, F. M., T. Y. Tsao, et al. (1983). "Monoclonal antibodies to mitotic cells." Proc Natl Acad Sci U S A **80**: 2926-2930.
- Delobel, P., S. Flament, et al. (2002). "Abnormal Tau phosphorylation of the Alzheimer-type also occurs during mitosis." J Neurochem **83**(2): 412-20.
- Deshaies, R. J. and J. E. J. Ferrell (2001). "Multiple phosphorylation and the countdown to S phase." Cell **107**: 819-822.
- Dickson, D. W. (2004). "Apoptotic mechanisms in Alzheimer neurofibrillary degeneration: cause or effect?" J Clin Invest **114**(1): 23-7.
- Dingley, A. J., J. P. Mackay, et al. (1995). "Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2." J Biomol NMR **6**(3): 321-8.
- Drewes, G., B. Lichtenberg-Kraag, et al. (1992). "Mitogen activated protein (MAP) kinase transforms tau protein into an Alzheimer-like state." Embo J **11**(6): 2131-8.
- Durocher, D., I. A. Taylor, et al. (2000). "The molecular basis of FHA domain: phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms." Mol Cell **6**(5): 1169-82.
- Eidenmuller, J., T. Fath, et al. (2000). "Structural and functional implications of tau hyperphosphorylation: information from phosphorylation-mimicking mutated tau proteins." Biochemistry **39**(43): 13166-75.
- Eidenmuller, J., T. Fath, et al. (2001). "Phosphorylation-mimicking glutamate clusters in the proline-rich region are sufficient to simulate the functional deficiencies of hyperphosphorylated tau protein." Biochem J **357**(Pt 3): 759-67.
- Ellis, J. R. and A. P. Minton (2003). "Join the crowd." Nature **425**: 27-28.
- Erlanson, D. A., A. C. Braisted, et al. (2000). "Site-directed ligand discovery." Proc Natl Acad Sci U S A **97**(17): 9367-9372.
- Feher, V. A. and J. Cavanagh (1999). "Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F." Nature **400**(6741): 289-93.
- Fischer, G. (2000). "Chemical aspects of peptide bond isomerisation." Chem Soc Rev **29**: 119-127.
- Fischer, G., H. Bang, et al. (1984). "[Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides]." Biomed Biochim Acta **43**(10): 1101-11.
- Fischer, G., T. Tradler, et al. (1998). "The mode of action of peptidyl prolyl cis/trans isomerases in vivo: binding vs. catalysis." FEBS Lett **426**(1): 17-20.
- Fujimori, F., W. Gunji, et al. (2001). "Crosstalk of prolyl isomerases, Pin1/Ess1, and cyclophilin A." Biochem Biophys Res Commun **289**(1): 181-90.
- Fujimori, F., K. Takahashi, et al. (1999). "Mice lacking Pin1 develop normally, but are defective in entering cell cycle from G(0) arrest." Biochem Biophys Res Commun **265**(3): 658-63.
- Gamblin, T. C., F. Chen, et al. (2003). "Caspase cleavage of tau: linking amyloid and neurofibrillary tangles in Alzheimer's disease." Proc Natl Acad Sci U S A **100**(17): 10032-7.

- Gong, C. X., T. Lidsky, et al. (2000). "Phosphorylation of microtubule-associated protein tau is regulated by protein phosphatase 2A in mammalian brain. Implications for neurofibrillary degeneration in Alzheimer's disease." *J Biol Chem* **275**(8): 5535-44.
- Goode, B. L., P. E. Denis, et al. (1997). "Functional Interactions between the proline-rich and repeat regions of tau enhance microtubule binding and assembly region." *Mol Biol Cell* **8**: 353-365.
- Goux, W. J. (2002). "The conformations of filamentous and soluble tau associated with Alzheimer paired helical filaments." *Biochemistry* **41**(46): 13798-806.
- Goux, W. J., L. Kopplin, et al. (2004). "The formation of straight and twisted filaments from short tau peptides." *J Biol Chem* **279**(26): 26868-75.
- Hajduk, P. J., T. Gerfin, et al. (1999). "High-throughput nuclear magnetic resonance-based screening." *J Med Chem* **42**(13): 2315-7.
- Hajduk, P. J., R. P. Meadows, et al. (1997). "Discovering high-affinity ligands for proteins." *Science* **278**(5337): 497,499.
- Hamdane, M., P. Delobel, et al. (2003). "Neurofibrillary degeneration of the Alzheimer-type: an alternate pathway to neuronal apoptosis?" *Biochem Pharmacol* **66**(8): 1619-25.
- Hamdane, M., A. V. Sambo, et al. (2003). "Mitotic-like tau phosphorylation by p25-Cdk5 kinase complex." *J Biol Chem* **278**(36): 34026-34.
- Hamdane, M., C. Smet, et al. (2002). "Pin1: a therapeutic target in Alzheimer neurodegeneration." *J Mol Neurosci* **19**(3): 275-87.
- Hempel, H., K. Buerger, et al. (2001). "Tracking of Alzheimer's disease progression with cerebrospinal fluid tau protein phosphorylated at threonine 231." *Ann Neurol* **49**(4): 545-6.
- Hempel, H., K. Buerger, et al. (2004). "Measurement of phosphorylated tau epitopes in the differential diagnosis of Alzheimer disease: a comparative cerebrospinal fluid study." *Arch Gen Psychiatry* **61**(1): 95-102.
- Hani, J., G. Stumpf, et al. (1995). "PTF1 encodes an essential protein in *Saccharomyces cerevisiae*, which shows strong homology with a new putative family of PPIases." *FEBS Lett* **365**(2-3): 198-202.
- Hennig, L., C. Christner, et al. (1998). "Selective inactivation of parvulin-like peptidyl-prolyl cis/trans isomerases by juglone." *Biochemistry* **37**(17): 5953-60.
- Hoffmann, R., V. M. Lee, et al. (1997). "Unique Alzheimer's disease paired helical filament specific epitopes involve double phosphorylation at specific sites." *Biochemistry* **36**(26): 8114-24.
- Hofmann, K., Bucher, P. (1995). "The rsp5-domain is shared by proteins of diverse functions." *FEBS Letters* **358**: 153-157.
- Homans, S. W. (2004). "NMR spectroscopy tools for structure-aided drug design." *Angew Chem Int Ed Engl* **43**(3): 290-300.
- Huang, H. K., S. L. Forsburg, et al. (2001). "Isolation and characterization of the Pin1/Ess1p homologue in *Schizosaccharomyces pombe*." *J Cell Sci* **114**(Pt 20): 3779-88.
- Hubbell, W. L., H. S. Mchaourab, et al. (1996). "Watching proteins move using site-directed spin labeling." *Structure* **4**(7): 779-783.
- Husseman, J. W., D. Nochlin, et al. (2000). "Mitotic activation: a convergent mechanism for a cohort of neurodegenerative diseases." *Neurobiol. Aging* **21**(6): 815-828.
- Hyde, J., A. C. Braisted, et al. (2003). "Discovery and characterization of cooperative ligand binding in the adaptive region of interleukin-2." *Biochemistry* **42**(21): 6475-83.
- Illenberger, S., Q. Zheng-Fischhofer, et al. (1998). "The endogenous and cell cycle-dependent phosphorylation of tau protein in living cells: implications for Alzheimer's disease." *Mol Biol Cell* **9**(6): 1495-512.

- Ivery, M. T. G. (2000). "Immunophilins : switched on protein binding domains?" Med Res Rev **20**(6): 452-484.
- Jacobs, D. M., K. Saxena, et al. (2002). "1H, 13C and 15N backbone resonance assignment of the peptidyl-prolyl cis-trans isomerase Pin1." J Biomol NMR **23**(2): 163-4.
- Jacobs, D. M., K. Saxena, et al. (2003). "Peptide binding induces large scale changes in inter-domain mobility in human Pin1." J Biol Chem **278**(28): 26174-82.
- Jahnke, W., P. Floersheim, et al. (2002). "NMR reporter screening for the detection of high-affinity ligands." Angew Chem Int Ed Engl **41**(18): 3420-3.
- Jahnke, W., A. Florsheimer, et al. (2003). "Second-site NMR screening and linker design." Curr Top Med Chem **3**(1): 69-80.
- Jahnke, W., S. Rudisser, et al. (2001). "Spin label enhanced NMR screening." J Am Chem Soc **123**(13): 3149-50.
- Jicha, G. A., E. Lane, et al. (1997). "A conformation- and phosphorylation-dependent antibody recognizing the paired helical filaments of Alzheimer's disease." J Neurochem **69**(5): 2087-95.
- Kops, O., X. Z. Zhou, et al. (2002). "Pin1 modulates the dephosphorylation of the RNA polymerase II C-terminal domain by yeast Fcp1." FEBS Lett **513**(2-3): 305-11.
- Kriwacki, R. W., L. Hengst, et al. (1996). "Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity." Proc Natl Acad Sci U S A **93**(21): 11504-9.
- Landrieu, I., L. De Veylder, et al. (2000). "The Arabidopsis thaliana PIN1At gene encodes a single-domain phosphorylation-dependent peptidyl prolyl cis/trans isomerase." J Biol Chem **275**(14): 10577-81.
- Landrieu, I., B. Odaert, et al. (2001). "p13(SUC1) and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope." J Biol Chem **276**(2): 1434-8.
- Landrieu, I., J. M. Wieruszkeski, et al. (2002). "Solution structure of the single-domain prolyl cis/trans isomerase PIN1At from Arabidopsis thaliana." J Mol Biol **320**(2): 321-32.
- Leger, J., M. Kempf, et al. (1997). "Conversion of serine to aspartate imitates phosphorylation-induced changes in the structure and function of microtubule-associated protein tau." J Biol Chem **272**(13): 8441-8446.
- Lietzow, M. A., M. Jamin, et al. (2002). "Mapping long-range contacts in a highly unfolded protein." J Mol Biol **322**: 655-662.
- Liou, Y. C., A. Ryo, et al. (2002). "Loss of Pin1 function in the mouse causes phenotypes resembling cyclin D1-null phenotypes." Proc Natl Acad Sci U S A **99**(3): 1335-40.
- Liou, Y. C., A. Sun, et al. (2003). "Role of the prolyl isomerase Pin1 in protecting against age-dependent neurodegeneration." Nature **424**(6948): 556-61.
- Lippens, G., J. M. Wieruszkeski, et al. (2004). "Proline-Directed Random-Coil Chemical Shift Values as a Tool for the NMR Assignment of the Tau Phosphorylation Sites." Chembiochem **5**(3): 256.
- Liu, W., H. D. Youn, et al. (2001). "Binding and regulation of the transcription factor NFAT by the peptidyl prolyl cis-trans isomerase Pin1." FEBS Lett **496**(2-3): 105-8.
- Lu, K. P. (2000). "Phosphorylation-dependent prolyl isomerization: a novel cell cycle regulatory mechanism." Prog Cell Cycle Res **4**: 83-96.
- Lu, K. P. (2003). "Prolyl isomerase Pin1 as a molecular target for cancer diagnostics and therapeutics." Cancer Cell **4**(3): 175-180.
- Lu, K. P. (2004). "Pinning down cell signaling, cancer and Alzheimer's disease." Trends Biochem Sci **29**(4): 200-9.
- Lu, K. P., S. D. Hanes, et al. (1996). "A human peptidyl-prolyl isomerase essential for regulation of mitosis." Nature **380**(6574): 544-7.

- Lu, K. P. and T. Hunter (1995). "The NIMA kinase: a mitotic regulator in *Aspergillus nidulans* and vertebrate cells." *Prog Cell Cycle Res* **1**: 187-205.
- Lu, K. P. and T. Hunter (1995). "Evidence for a NIMA-like mitotic pathway in vertebrate cells." *Cell* **81**(3): 413-24.
- Lu, K. P., Y. C. Liou, et al. (2003). "Proline-directed phosphorylation and isomerization in mitotic regulation and in Alzheimer's Disease." *Bioessays* **25**(2): 174-81.
- Lu, K. P., Y. C. Liou, et al. (2002). "Pinning down proline-directed phosphorylation signaling." *Trends Cell Biol* **12**(4): 164-72.
- Lu, P. J., G. Wulf, et al. (1999). "The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein." *Nature* **399**(6738): 784-8.
- Lu, P. J., X. Z. Zhou, et al. (2002). "Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function." *J Biol Chem* **277**(4): 2381-4.
- Lu, P. J., X. Z. Zhou, et al. (1999). "Function of WW domains as phosphoserine- or phosphothreonine-binding modules." *Science* **283**(5406): 1325-8.
- Macias, M. J., S. Wiesner, et al. (2002). "WW and SH3 domains, two different scaffolds to recognize proline-rich ligands." *FEBS Lett* **513**(1): 30-7.
- Maciejewski, P. M., F. C. Peterson, et al. (1995). "Mutation of Ser90 to glutamic acid mimics phosphorylation of bovin prolactin." *J Biol Chem* **270**: 27661-27665.
- Mailliot, C., T. Bussiere, et al. (1998). "Alzheimer-specific epitope of AT100 in transfected cell lines with tau: toward an efficient cell model of tau abnormal phosphorylation." *Neurosci Lett* **255**(1): 13-6.
- Mailliot, C., T. Bussiere, et al. (2000). "Pathological tau phenotypes. The weight of mutations, polymorphisms, and differential neuronal vulnerabilities." *Ann N Y Acad Sci* **920**: 107-14.
- Mailliot, C., N. Sergeant, et al. (1998). "Phosphorylation of specific sets of tau isoforms reflects different neurofibrillary degeneration processes." *FEBS Lett* **433**(3): 201-4.
- Mayer, B. J. (2001). "SH3 domains: complexity in moderation." *J Cell Sci* **114**(Pt 7): 1253-63.
- Meiler, J. (2003). "PROSHIFT :protein chemical shift prediction using artificial neural networks." *J Biomol NMR* **26**(1): 25-37.
- Merrifield, R. B. (1963). "Solid phase peptide synthesis." *J Am Chem Soc* **85**: 2149-2154.
- Metzner, M., G. Stoller, et al. (2001). "Functional replacement of the essential ESS1 in yeast by the plant parvulin D1Par13." *J Biol Chem* **276**(17): 13524-9.
- Meyer, B. and T. Peters (2003). "NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors." *Angew Chem Int Ed Engl* **42**(8): 864-90.
- Morishima-Kawashima, M., M. Hasegawa, et al. (1995). "Proline-directed and non-proline-directed phosphorylation of PHF-tau." *J Biol Chem* **270**(2): 823-9.
- Nagy, Z., M. M. Esiri, et al. (1998). "The cell division cycle and the pathophysiology of Alzheimer's disease." *Neuroscience* **87**(4): 731-9.
- Nash, P., X. Tang, et al. (2001). "Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication." *Nature* **414**: 514-521.
- Ovadi, J. and V. Saks (2004). "On the origin of intracellular compartmentation and organized metabolic systems." *Mol Cell Biochem* **256-257**(1-2): 5-12.
- Panda, D., B. L. Goode, et al. (1995). "Kinetic stabilization of microtubule dynamics at steady state by tau and microtubule-binding domains of tau." *Biochemistry* **34**(35): 11117-27.
- Parang, K., E. J. Fournier, et al. (2001). "A solid phase reagent for the capture phosphorylation of carbohydrates and nucleosides." *Org Lett* **3**(2): 307-9.
- Patra, D., S. X. Wang, et al. (1999). "The xenopus Suc1/Cks protein promotes the phosphorylation of G(2)/M regulators." *J Biol Chem* **274**(52): 36839-42.

- Patrick, G. N., L. Zukerberg, et al. (1999). "Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration." *Nature* **402**(6762): 615-22.
- Pawson, T., M. Raina, et al. (2002). "Interaction domains: from simple binding events to complex cellular behavior." *FEBS Lett* **513**(1): 2-10.
- Plaxco, K. W. and M. Gross (1997). "Cell biology. The importance of being unfolded." *Nature* **386**(6626): 657, 659.
- Plaxco, K. W. and M. Gross (2001). "Unfolded, yes, but random? Never!" *Nat Struct Biol* **8**(8): 659-60.
- Rahfeld, J. U., K. P. Rucknagel, et al. (1994). "Confirmation of the existence of a third family among peptidyl-prolyl cis/trans isomerases. Amino acid sequence and recombinant production of parvulin." *FEBS Lett* **352**(2): 180-4.
- Raina, A. K., M. J. Monteiro, et al. (1999). "The role of cell cycle-mediated events in Alzheimer's disease." *Int J Exp Pathol* **80**(2): 71-6.
- Ranganathan, R., K. P. Lu, et al. (1997). "Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent." *Cell* **89**(6): 875-86.
- Rippmann, J. F., S. Hobbie, et al. (2000). "Phosphorylation-dependent proline isomerization catalyzed by Pin1 is essential for tumor cell survival and entry into mitosis." *Cell Growth Differ* **11**(7): 409-16.
- Ryo, A., Y. C. Liou, et al. (2003). "Prolyl isomerase Pin1: a catalyst for oncogenesis and a potential therapeutic target in cancer." *J Cell Sci* **116**(Pt 5): 773-83.
- Ryo, A., Y. C. Liou, et al. (2002). "PIN1 is an E2F target gene essential for Neu/Ras-induced transformation of mammary epithelial cells." *Mol Cell Biol* **22**(15): 5281-95.
- Ryo, A., M. Nakamura, et al. (2001). "Pin1 regulates turnover and subcellular localization of beta-catenin by inhibiting its interaction with APC." *Nat Cell Biol* **3**(9): 793-801.
- Schmid, F. X. (1993). "Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions." *Annu Rev Biophys Biomol Struct* **22**: 123-42.
- Schmid, F. X. (2001). "Prolyl isomerases." *Adv Protein Chem* **59**: 243-82.
- Schmid, F. X., C. Frech, et al. (1996). "Catalyzed and assisted protein folding of ribonuclease T1." *Biol Chem* **377**(7-8): 417-24.
- Schneider, A., J. Biernat, et al. (1999). "Phosphorylation that detaches tau protein from microtubules (Ser262, Ser214) also protects it against aggregation into Alzheimer paired helical filaments." *Biochemistry* **38**(12): 3549-58.
- Schutkowski, M., A. Bernhardt, et al. (1998). "Role of phosphorylation in determining the backbone dynamics of the serine/threonine-proline motif and Pin1 substrate recognition." *Biochemistry* **37**(16): 5566-75.
- Schweers, O., E. Schonbrunn-Hanebeck, et al. (1994). "Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure." *J Biol Chem* **269**(39): 24290-7.
- Sekerina, E., J. U. Rahfeld, et al. (2000). "NMR solution structure of hPar14 reveals similarity to the peptidyl prolyl cis/trans isomerase domain of the mitotic regulator hPin1 but indicates a different functionality of the protein." *J Mol Biol* **301**(4): 1021-1035.
- Shaw, P. E. (2002). "Peptidyl-prolyl isomerases: a new twist to transcription." *EMBO Rep* **3**(6): 521-6.
- Shen, M., P. T. Stukenberg, et al. (1998). "The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins." *Genes Dev* **12**(5): 706-20.
- Shtilerman, M. D., T. T. Ding, et al. (2002). "Molecular crowding accelerates fibrillization of  $\alpha$ -synuclein : could an increase in the cytoplasmic protein concentration induce Parkinson's disease?" *Biochemistry* **41**: 3855-3860.

- Shuker, S. B., P. J. Hajduk, et al. (1996). "Discovering high-affinity ligands for proteins: SAR by NMR." *Science* **274**(5292): 1531-4.
- Smet, C., A. V. Sambo, et al. (2004). "The peptidyl prolyl cis/trans-isomerase Pin1 recognizes the phospho-Thr212-Pro213 site on Tau." *Biochemistry* **43**(7): 2032-40.
- Smith, L. J., K. M. Fiebig, et al. (1996). "The concept of a random coil. Residual structure in peptides and denatured proteins." *Fold Des* **1**(5): R95-106.
- Sontag, E., V. Nunbhakdi-Craig, et al. (1999). "Molecular interactions among protein phosphatase 2A, tau, and microtubules. Implications for the regulation of tau phosphorylation and the development of tauopathies." *J Biol Chem* **274**(36): 25490-8.
- Spolar, R. S. and M. T. Record, Jr. (1994). "Coupling of local folding to site-specific binding of proteins to DNA." *Science* **263**(5148): 777-84.
- Stukenberg, P. T. and M. W. Kirschner (2001). "Pin1 acts catalytically to promote a conformational change in Cdc25." *Mol Cell* **7**(5): 1071-83.
- Sudol, M. (1996). "The WW module competes with the SH3 domain?" *Trends Biochem Sci* **21**(5): 161-3.
- Sudol, M. (1996). "Structure and function of the WW domain." *Prog Biophys Mol Biol* **65**(1-2): 113-32.
- Sudol, M., P. Bork, et al. (1995). "Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain." *J Biol Chem* **270**(24): 14733-41.
- Sudol, M., H. I. Chen, et al. (1995). "Characterization of a novel protein-binding module--the WW domain." *FEBS Lett* **369**(1): 67-71.
- Sudol, M., K. Sliwa, et al. (2001). "Functions of WW domains in the nucleus." *FEBS Lett* **490**(3): 190-5.
- Thornton, J. M., C. A. Orengo, et al. (1999). "Protein folds, functions and evolution." *J Mol Biol* **293**: 333-342.
- Uchida, T., F. Fujimori, et al. (1999). "Identification and characterization of a 14 kDa human protein as a novel parvulin-like peptidyl prolyl cis/trans isomerase." *FEBS Lett* **446**(2-3): 278-82.
- Uchida, T., M. Takamiya, et al. (2003). "Pin1 and Par14 peptidyl prolyl isomerase inhibitors block cell proliferation." *Chem Biol* **10**(1): 15-24.
- Uversky, V. N., E. M. Cooper, et al. (2001). "Accelerated  $\alpha$ -synuclein fibrillation in crowded milieu." *FEBS Lett* **515**: 99-103.
- Vandré, D. D., F. M. Davis, et al. (1986). "Distribution of the cytoskeletal proteins sharing a conserved phosphorylated epitope." *Eur J Cell Biol* **41**: 72-81.
- Verdecia, M. A., M. E. Bowman, et al. (2000). "Structural basis for phosphoserine-proline recognition by group IV WW domains." *Nat Struct Biol* **7**(8): 639-43.
- Vincent, I. (2000). "Cycling to the finish." *Neurobiol Aging* **21**(6): 757-60.
- Vincent, I., M. Rosado, et al. (1996). "Mitotic mechanisms in Alzheimer's disease?" *J Cell Biol* **132**: 413-425.
- von Bergen, M., P. Friedhoff, et al. (2000). "Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure." *Proc Natl Acad Sci U S A* **97**(10): 5129-34.
- Walters, W. P. and M. Namchuk (2003). "Designing screens: how to make your hits a hit." *Nat Rev Drug Discov* **2**(4): 259-66.
- Wang, J. Z., C. X. Gong, et al. (1995). "Dephosphorylation of Alzheimer paired helical filaments by protein phosphatase-2A and -2B." *J Biol Chem* **270**(9): 4854-60.
- Wang, X. J., S. A. Hart, et al. (2003). "Serine-cis-proline and serine-trans-proline isosteres: stereoselective synthesis of (Z)- and (E)-alkene mimics by Still-Wittig and Ireland-Claisen rearrangements." *J Org Chem* **68**(6): 2343-9.

- Wang, Y. and O. Jardetzky (2002). "Investigation of the neighboring residue effects on protein chemical shifts." *J Am Chem Soc* **124**(47): 14075-14084.
- Weber, P. C., D. H. Ohlendorf, et al. (1989). "Structural origins of high-affinity biotin binding to streptavidin." *Science* **243**: 85-88.
- Winkler, K. E., K. I. Swenson, et al. (2000). "Requirement of the prolyl isomerase Pin1 for the replication checkpoint." *Science* **287**(5458): 1644-7.
- Wintjens, R., J. M. Wieruszkeski, et al. (2001). "<sup>1</sup>H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides." *J Biol Chem* **276**(27): 25150-6.
- Wishart, D. S., C. G. Bigam, et al. (1995). "<sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects." *J Biomol NMR* **5**: 67-81.
- Wright, P. E. and H. J. Dyson (1999). "Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm." *J Mol Biol* **293**(2): 321-31.
- Wulf, G., A. Ryo, et al. (2003). "The prolyl isomerase Pin1 in breast development and cancer." *Breast Cancer Res* **5**(2): 76-82.
- Wulf, G. M., Y. C. Liou, et al. (2002). "Role of Pin1 in the regulation of p53 stability and p21 transactivation, and cell cycle checkpoints in response to DNA damage." *J Biol Chem* **277**(50): 47976-9.
- Wulf, G. M., A. Ryo, et al. (2001). "Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1." *Embo J* **20**(13): 3459-72.
- Xu, Y. X., Y. Hirose, et al. (2003). "Pin1 modulates the structure and function of human RNA polymerase II." *Genes Dev* **17**(22): 2765-76.
- Yaffe (1999). "Grabbing phosphoproteins." *Nature* **402**: 30-31.
- Yaffe, M. B. and A. E. H. Elia (2001). "Phosphoserine/threonine-binding domains." *Current Opinion in Cell Biology* **13**: 131-138.
- Yaffe, M. B., K. Rittinger, et al. (1997). "The structural basis for 14-3-3:phosphopeptide binding specificity." *Cell* **91**(7): 961-71.
- Yaffe, M. B., M. Schutkowski, et al. (1997). "Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism." *Science* **278**(5345): 1957-60.
- Yao, J. L., O. Kops, et al. (2001). "Functional conservation of phosphorylation-specific prolyl isomerases in plants." *J Biol Chem* **276**(17): 13517-23.
- Zacchi, P., M. Gostissa, et al. (2002). "The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults." *Nature* **419**(6909): 853-7.
- Zarrinpar, A. and W. A. Lim (2000). "Converging on proline: the mechanism of WW domain peptide recognition." *Nat Struct Biol* **7**(8): 611-3.
- Zhang, Y., S. Fussel, et al. (2002). "Substrate-based design of reversible Pin1 inhibitors." *Biochemistry* **41**(39): 11868-77.
- Zheng, H., H. You, et al. (2002). "The prolyl isomerase Pin1 is a regulator of p53 in genotoxic response." *Nature* **419**(6909): 849-53.
- Zheng-Fischhofer, Q., J. Biernat, et al. (1998). "Sequential phosphorylation of Tau by glycogen synthase kinase-3beta and protein kinase A at Thr212 and Ser214 generates the Alzheimer-specific epitope of antibody AT100 and requires a paired-helical-filament-like conformation." *Eur J Biochem* **252**(3): 542-52.
- Zhou, X. Z., O. Kops, et al. (2000). "Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins." *Mol Cell* **6**(4): 873-83.
- Zhou, X. Z., P. J. Lu, et al. (1999). "Phosphorylation-dependent prolyl isomerization: a novel signaling regulatory mechanism." *Cell Mol Life Sci* **56**(9-10): 788-806.

## **MATERIELS ET METHODES.**



**Première Partie – Synthèse Peptidique, Biochimie des Protéines, Interactions Biomoléculaires (techniques de mesure).**

<b>I - Synthèse peptidique en phase solide.....</b>	<b>367</b>
1.1 – Synthèse en phase solide de peptides phosphorylés.....	367
1.2 - Purification des peptides par RP-HPLC. ....	371
Un exemple de suivi des réactions de couplage pendant la synthèse et des différentes étapes de purification et de caractérisation d'un peptide en solution après coupure de la phase solide. ....	372
Principe d'attribution séquentielle d'un peptide par RMN combinant les expériences RMN 2D homonucléaires TOCSY et NOESY. ....	374
Récapitulatif des différents peptides synthétisés pour l'étude des interactions Pin1-tau....	376
Synthèse d'un fragment peptidique de 40 acides aminés.....	377
<b>II – Expression et purification de protéines recombinantes.....</b>	<b>381</b>
2.1 – Expression des protéines Pin1, WW et Pin1 <sub>CAT</sub> par technique recombinante. ....	382
2.1.1 – Expression de protéines non marquées en milieu riche. ....	382
2.1.2 – Expression de protéine enrichie <sup>15</sup> N ou <sup>15</sup> N/ <sup>13</sup> C en milieu minimum. ....	383
2.2 – Purification de Pin1, de Pin1 <sub>CAT</sub> et du domaine WW par chromatographie d'affinité (1 <sup>re</sup> étape de purification).....	385
2.2.1 – Purification par affinité d'une queue poly-histidine (His-tag) pour le nickel (Pin1, WW).....	385
2.2.2 – Purification par affinité de la GST pour le glutathion (WW, Pin1 et Pin1 <sub>CAT</sub> )...	386
2.3 – Purification de Pin1 par chromatographie d'exclusion stérique (2 <sup>e</sup> étape). ....	389
2.4 – Purification du domaine WW par RP-HPLC (2 <sup>e</sup> étape de purification). ....	390
2.5 – Dessalage et préparation des échantillons dans les tampons utilisés pour la RMN ou la fluorescence par chromatographie d'exclusion sur gel non résolutif.....	392
Détermination de la concentration de Pin1 et du domaine WW en solution. ....	393
Courbes d'étalonnage Bradford établies pour Pin1 et le domaine WW isolé.....	395
<b>III – Acquisition des données par RMN et par fluorescence. Procédure de titration.....</b>	<b>398</b>
3.1 – RMN.....	398
RMN des peptides. ....	398
RMN de la protéine Pin1 ou du domaine WW. ....	399
Procédure de titration par RMN. ....	400

Calcul des constantes de dissociations ( $K_D$ ).....	401
3.2 – Fluorescence intrinsèque et FRET. ....	401
<b>IV - Protocoles généraux des expériences en milieu hétérogène (Biacore<sup>®</sup>, ELISA). ....</b>	<b>402</b>
4.1 – Biacore <sup>®</sup> . ....	402
4.2 – ELISA.....	403
<b>Structures de Pin1 et des complexes dans la PDB (<i>Protein Data Bank</i>). ....</b>	<b>406</b>
<b>Données d’attribution par RMN de Pin1 et du domaine WW dans la BMRB (<i>BioMagResBank</i>). ....</b>	<b>407</b>

## Deuxième Partie – Synthèse organique et caractérisation des composés.

### Criblage des composés par RMN et fluorescence.

<b>I - Synthèse et caractérisation des ligands potentiels du domaine WW de Pin1. ....</b>	<b>408</b>
1.1 - Synthèse des composés .....	408
Procédure n°1 : Procédure générale pour les composés (1)-(10), (19)-(35), (38)-(45). ....	409
Procédure n°2 : Procédure générale pour la synthèse en parallèle du dipeptide Glu-Pro et des dérivés dans les composés (11) à (18). ....	412
Procédure n°3 : Procédure générale pour les composés (36) et (37). ....	415
1.2 - Caractérisation des composés. ....	416
Composé (1). ....	416
Composé (2). ....	417
Composé (3). ....	417
Composé (4). ....	418
Composé (5). ....	419
Composé (6). ....	420
Composé (7). ....	420
Composé (8). ....	421
Composé (9). ....	421
Composé (10). ....	422
Composé (11). ....	422
Composé (12). ....	423
Composé (13). ....	423
Composé (14). ....	424
Composé (15). ....	424

## Matériels et Méthodes

Composé (16).....	425
Composé (17).....	425
Composé (18).....	426
Composé (26).....	427
Composé (27).....	427
Composé (28).....	428
Composé (29).....	428
Composé (30).....	429
Composé (31).....	429
Composé (32).....	430
Composé (33).....	430
Composé (34).....	431
Composé (35).....	432
Composé (36).....	433
Composé (37).....	434
Composé (38).....	435
Composé (39).....	436
Composé (40).....	436
Composé (41).....	437
Composé (42).....	437
Composé (43).....	438
Composé (44).....	438
Composé (45).....	439
<b>II - Screening par RMN et calcul des constantes de dissociation.....</b>	<b>440</b>
2.1 - Expression et purification du domaine WW recombinant.....	440
2.2 - <i>Screening</i> par RMN.....	440
<b>III – Compétition suivie par FRET pour le calcul des IC<sub>50</sub>.....</b>	<b>441</b>
3.1 - Synthèse du peptide Ac-KSRSRpT212PSLPTPPTR-NH <sub>2</sub> intervenant dans les expériences de compétition pour le calcul des IC <sub>50</sub> par FRET.....	441
3.2 - Couplage de l'AlexaFluor350®.....	442
3.3 - Screening par FRET.....	443

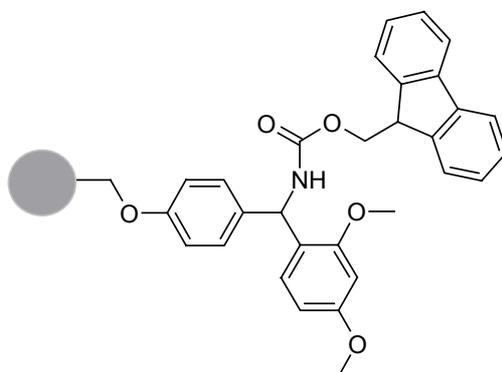


## Première Partie – Synthèse Peptidique, Biochimie des Protéines, Interactions Biomoléculaires (techniques de mesure).

### I - Synthèse peptidique en phase solide.

#### 1.1 – Synthèse en phase solide de peptides phosphorylés.

Les différents peptides phosphorylés et leurs homologues non phosphorylés sont synthétisés par synthèse en phase solide selon une stratégie Fmoc. Le support solide utilisé pour la synthèse est basé sur un support de type polystyrène ponté par 1% de divinylbenzène qui assure la résistance du support et le gonflement dans des solvants organiques très hydrophobes (DMF, DCM, NMP). La résine utilisée est choisie en fonction des conditions de déprotection qui implique une coupure en milieu acide concentré. L'obtention de peptide sous une forme carboxamide pour la partie C-terminale nécessite le choix d'une résine qui, après coupure, générera une fonction amide. Les résines de type Rink amide (Novabiochem) correspondent à ce critère et sont labiles en milieu acide fort (TFA > 50%).



**Schéma 1 :** Structure du *linker* d'une résine Rink amide.

Pour les synthèses automatiques sur synthétiseurs à flux continu (Pioneer™ Peptide Synthesizer), la pression du système nécessite une plus grande résistance du support solide. La résine utilisée est de type PAL-PEG-PS (Advanced ChemTech).

**Choix des dérivés d'acides aminés.** Les acides aminés protégés utilisés pour la synthèse sont commerciaux (Novabiochem) et sont décrits dans le Tableau 1. Les peptides sont linéaires et ne possèdent pas de dérivatisation de la chaîne latérale nécessitant une déprotection sélective

de celles-ci, par conséquent la stratégie de synthèse est classique. Les acides aminés utilisés comportent une protection Fmoc de la fonction amine et une protection orthogonale à la protection Fmoc sur la chaîne latérale quand celle-ci comporte une fonction réactive. On choisit des protections labiles en milieu acide qui pourront être déprotégées en même temps que l'étape de coupure du peptide de son support solide. La phosphorylation est réalisée par la couplage de synthons commerciaux d'acides aminés Ser ou Thr phosphorylés mono-protégés par un groupement benzyle (Tableau 1).

Acide aminé	Nom complet du dérivé Fmoc avec protection de la chaîne latérale
Ala	N- $\alpha$ -Fmoc-L-alanine
Arg	N- $\alpha$ -Fmoc-N <sup>ε</sup> -2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl-L-arginine
Asn	N- $\alpha$ -Fmoc-N- $\beta$ -trityl-L-asparagine
Asp	acide N- $\alpha$ -Fmoc-L-aspartique $\beta$ -terbutyl ester
Glu	acide N- $\alpha$ -Fmoc-L-glutamique $\gamma$ -terbutyl ester
Cys	N- $\alpha$ -Fmoc-L-cystéine
Gln	N- $\alpha$ -Fmoc-N- $\gamma$ -trityl-L-glutamine
Gly	N- $\alpha$ -Fmoc-L-glycine
His	N- $\alpha$ -Fmoc-N-im-trityl-L-histidine
Ile	N- $\alpha$ -Fmoc-L-isoleucine
Leu	N- $\alpha$ -Fmoc-L-leucine
Lys	N- $\alpha$ -Fmoc-N- $\epsilon$ -t.-Boc-L-lysine
Met	N- $\alpha$ -Fmoc-L-méthionine
Phe	N- $\alpha$ -Fmoc-L-phenylalanine
Pro	N- $\alpha$ -Fmoc-L-proline
Ser	N- $\alpha$ -Fmoc-O-terbutyl-L-serine
pSer	N- $\alpha$ -Fmoc-O-benzyl-L-phosphoserine
Thr	N- $\alpha$ -Fmoc-O-terbutyl-L-threonine
pThr	N- $\alpha$ -Fmoc-O-benzyl-L-phosphothreonine
Trp	N- $\alpha$ -Fmoc-N-in-t.-Boc-L-tryptophane
Tyr	N- $\alpha$ -Fmoc-O-terbutyl-L-tyrosine
Val	N- $\alpha$ -Fmoc-L-valine

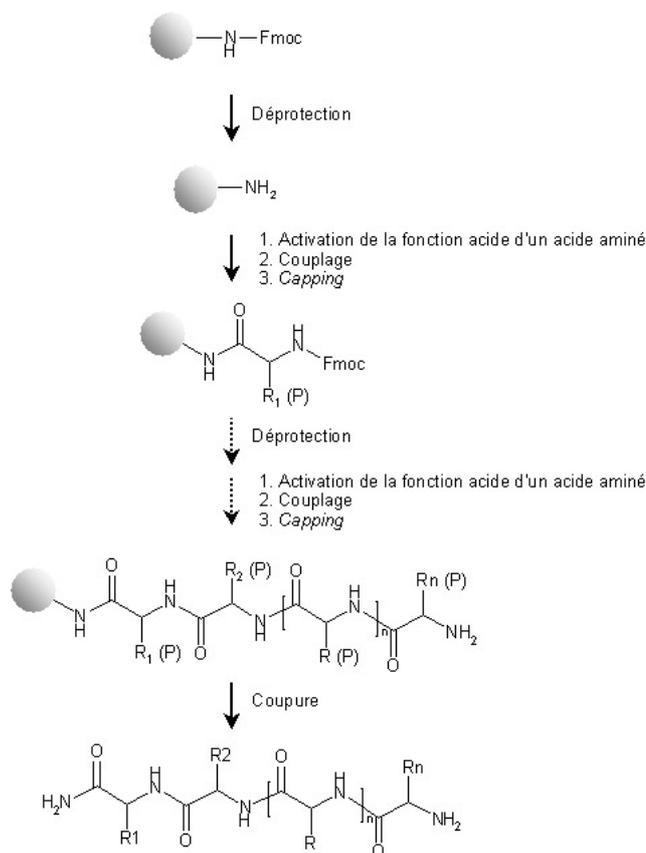
**Tableau 1 :** Dérivés acides aminés utilisés pour la synthèse en phase solide des peptides.

**Principe de la synthèse.** Le schéma de synthèse est répété de la façon suivante pour chaque cycle de couplage d'un acide aminé :

1. gonflement de la résine dans le solvant approprié
2. déprotection de l'amine par coupure du groupement Fmoc en milieu basique
3. activation de la fonction acide carboxylique et d'un acide aminé et couplage à la résine
4. réaction de *capping* correspondant à la « désactivation » des fonctions amines libres de la résine.

## Matériels et Méthodes

Les étapes 2, 3 et 4 sont répétées successivement pour chaque acide aminé de la séquence (ordre d'introduction des acides aminés au cours de la synthèse : du résidu C-terminal vers le résidu N-terminal). Une fois la synthèse terminée, on procède à la coupure du peptide de la résine et à la déprotection simultanée des fonctions portées par les chaînes latérales.



**Schéma 2 :** Schéma général de synthèse des peptides en phase solide en chimie Fmoc.

**Déprotection des fonctions amines protégées par un groupement Fmoc.** Les fonctions amines en  $\alpha$  de la fonction carboxylique des acides aminés sont protégées par le groupement protecteur Fmoc et déprotégées en milieu basique par réaction avec une solution de pipéridine à 20% dans le DMF. Les fonctions réactives des chaînes latérales (acide carboxylique, amine, alcool, phénol, cystéine, guanidine, thiol, imidazole) nécessitent une stratégie de protection orthogonale à la protection Fmoc. Les fonctions latérales sont par conséquent protégées par des groupements protecteurs labiles en milieu acide et seront déprotégées simultanément à la coupure du peptide de la résine.

**Couplage d'un acide aminé protégé.** La réaction de couplage met en œuvre les agents d'activation HOBt et TBTU en quantité équimolaire par rapport à l'acide aminé et 3 équiv. de DIPEA dans le DMF. Le couplage des acide aminé est effectué à température ambiante par simple couplage avec 10 équiv. d'acide aminé par équiv. de chaînes peptidiques en croissance sur la résine pendant 45 minutes. Le choix du simple ou double couplage (avec deux fois 5 équiv. d'acide aminé) tient compte de la longueur de la séquence peptidique et de sa difficulté, c'est-à-dire de la présence ou non d'une succession de résidus hydrophobes ou encombrés, qui peuvent conduire à l'agrégation du peptide sur résine.

**Réaction de *capping*.** Le *capping* est une réaction d'acétylation des fonctions amines qui n'ont pas réagi lors du couplage. Elle permet de « clôturer », par un groupement acétyle résistant aux conditions basiques et acides utilisées, les chaînes qui présenteraient une délétion dans leur séquence et permet ainsi de faciliter l'étape de purification ultérieure. On fait réagir une solution à 3% d'anhydride acétique (v/v) et 0.3 % de DIEA (v/v) dans le DMF.

**Introduction d'une sonde fluorescente en position N-terminale du peptide.** Le dansyle, utilisé dans les expériences de FRET, est introduit sélectivement sur la fonction amine terminale en fin de synthèse par incubation de la résine dans une solution contenant 5 équiv. de chlorure de dansyle et 10 équiv. de DIPEA dans le DMF pendant 45 min à température ambiante, sous agitation et à l'abri de la lumière. Ensuite la résine est lavée au DMF et on procède à une réaction de *capping*.

**Introduction de la (D)-Biotine en position N-terminale du peptide.** La (D)-biotine, couplée aux peptides, permet la fixation des peptides par liaisons fortes à la streptavidine, dans les expériences de Biacore<sup>®</sup> et d'ELISA. La (D)-biotine est introduite sélectivement sur la fonction amine terminale en fin de synthèse par incubation de la résine dans une solution contenant 5 équiv. (D)-Biotine, 5 équiv. HOBt, 5 équiv. HBTU et 10 équiv. de DIPEA dans le DMF pendant 45 min à température ambiante, sous agitation. Ensuite la résine est lavée au DMF et on procède à une réaction de *capping*.

**Réaction de coupure de la peptidyl-résine et de déprotection des chaînes latérales.** Le mélange de coupure est constitué de 95% TFA ; 2.5% TIS ; 2.5% H<sub>2</sub>O ou 94% TFA ; 1% TIS ; 2.5% EDT ; 2.5% H<sub>2</sub>O pour les peptides contenant un résidu Cys ou Met. La résine est incubée avec le mélange de coupure sous agitation à température ambiante pendant 2 heures

pour les peptides non phosphorylés ou 5 heures pour les peptides phosphorylés (10 mL de mélange de coupure pour 500 mg de résine). Après réaction, le mélange de coupure contenant le peptide est isolé par filtration de la résine et collecté dans un mélange éther : heptane (v/v) froid (200 mL pour 10 mL de mélange de coupure). On observe la formation d'un précipité blanc correspondant au mélange de peptides. La précipitation est achevée à  $-20^{\circ}\text{C}$  pendant la nuit. Les peptides sont isolés par centrifugation (3000 RPM,  $4^{\circ}\text{C}$ , 10 minutes). Le mélange brut est lavé 3 fois avec 40 mL d'éther froid puis dissout dans un minimum d'eau. Le résidu d'éther est évaporé par bullage d'azote dans la solution aqueuse. Le peptide est isolé par lyophilisation. Le mélange brut est contrôlé par spectrométrie de masse MALDI-TOF et RP-HPLC sur colonne C18 Nucleosil avant purification.

### 1.2 - Purification des peptides par RP-HPLC.

Le contrôle du mélange brut après coupure au TFA et lyophilisation, l'étape de purification et le contrôle de la pureté des fractions sont réalisés par HPLC en phase inverse sur colonne C18 Nucleosil. La colonne est équilibrée dans une solution aqueuse contenant 0.05% de TFA (v/v). Un gradient d'acétonitrile est mis en œuvre pour l'élution des peptides. La programmation du gradient pour les différentes étapes HPLC est donnée dans le Tableau 2 .

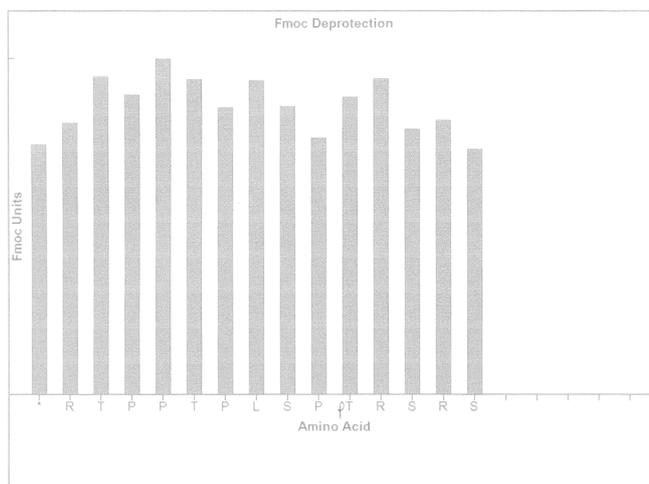
<i>Etape</i>	<i>gradient</i>	<i>temps</i>
Contrôle des peptides bruts	0 à 80 % CH <sub>3</sub> CN	30 min
Purification	0 à 50% CH <sub>3</sub> CN	60 min
Contrôle des fractions	0 à 50% CH <sub>3</sub> CN	10 min
Contrôle des peptides purifiés	0 à 80 % CH <sub>3</sub> CN	30 min

**Tableau 2 :** Programmation des gradients d'acétonitrile utilisé pour les étapes de purification et de contrôle de la pureté des peptides avant et après purification par HPLC en phase inverse sur colonne Nucleosil C18.

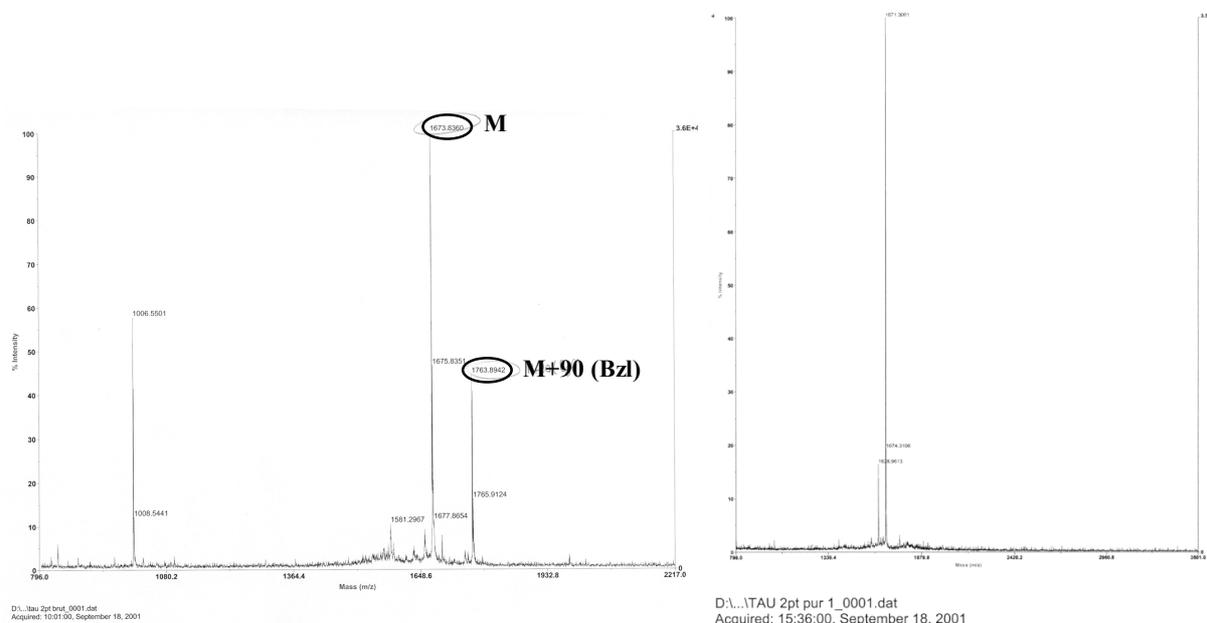
Les fractions contenant le peptide pur d'après les contrôles par spectrométrie de masse MALDI-TOF et RP-HPLC sont rassemblées et contrôlées. L'attribution séquentielle des peptides par RMN est réalisée avec l'utilisation des séquences impulsionsnelles homonucléaires TOCSY et NOESY.

*Un exemple de suivi des réactions de couplage pendant la synthèse et des différentes étapes de purification et de caractérisation d'un peptide en solution après coupure de la phase solide.*

Monitoring de la réaction de déprotection de la fonction amine libérant le fulvène par mesure de l'absorption UV de la solution de pipéridine 20% dans le DMF. Cette mesure permet un suivi indirect des réactions de couplage de manière qualitative.



Caractérisation des peptides par spectrométrie de masse avant et après purification.



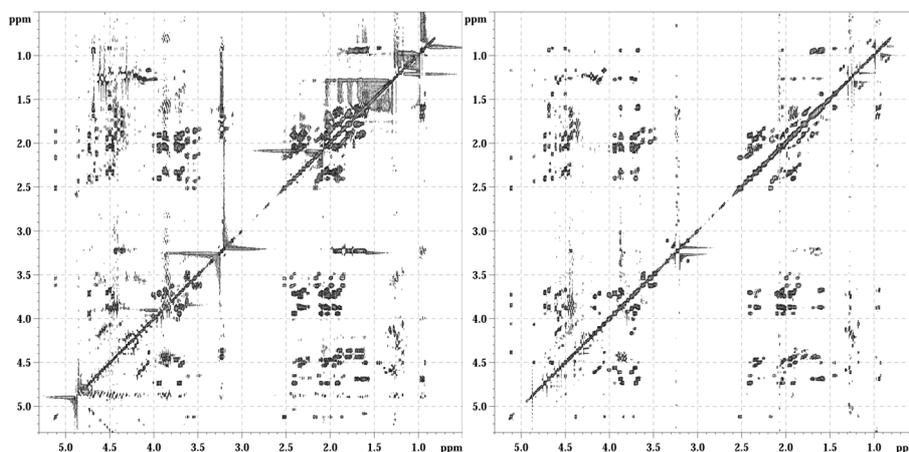
**Figure 1 :** Spectres de masse MALDI-TOF réalisés sur un peptide brut (à gauche) après coupure au TFA et isolement par précipitation dans un mélange éther : heptane (v/v) et sur le même peptide après purification (à droite).



**Principe d'attribution séquentielle d'un peptide par RMN combinant les expériences RMN 2D homonucléaires TOCSY et NOESY.**

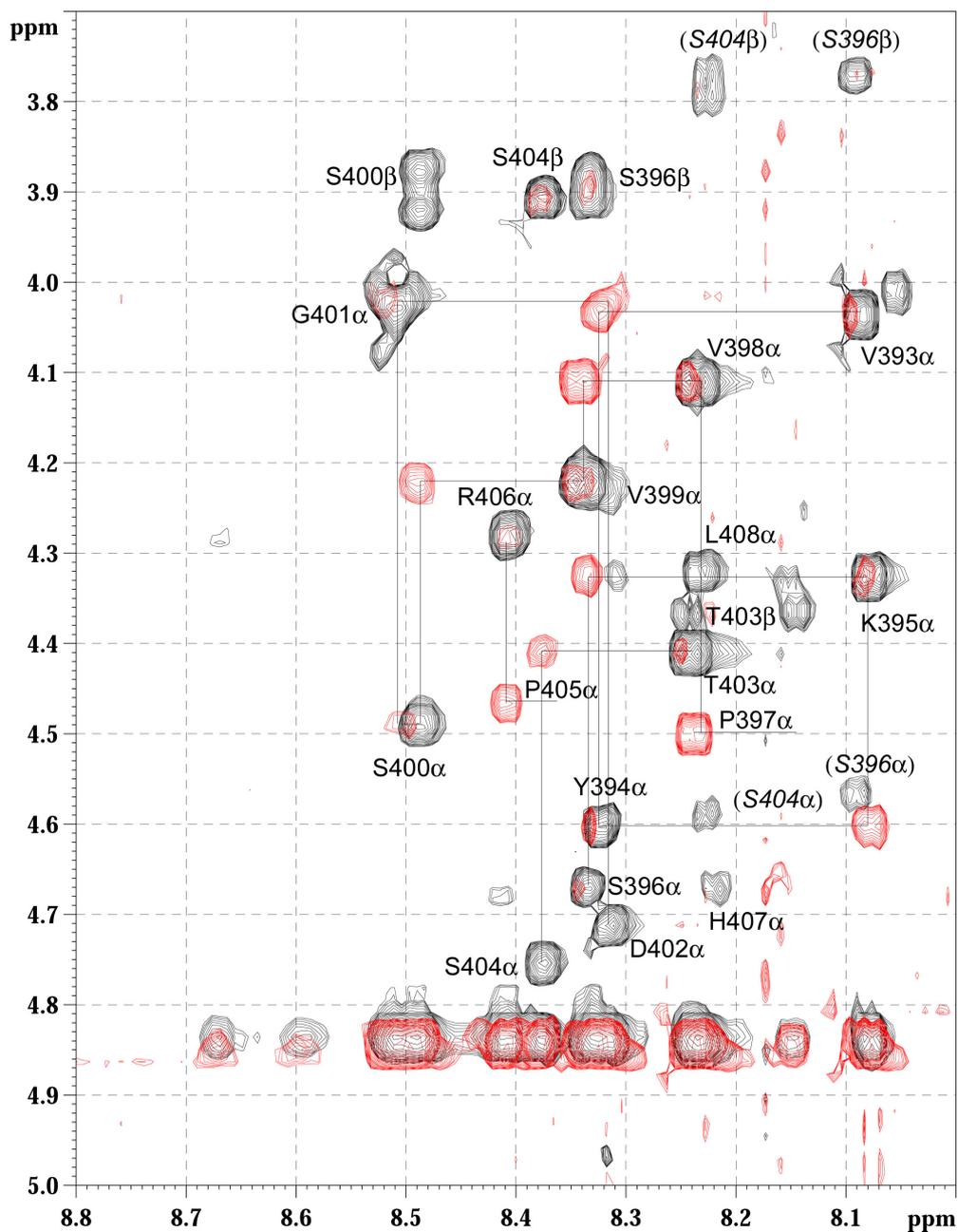
Les peptides sont solubilisés à une concentration de 2 mM dans un tampon Tris-d<sub>11</sub> 50 mM pH 6.40, NaCl 100 mM contenant 10% D<sub>2</sub>O et le TSPA-d<sub>4</sub> (ou 3-triméthylsilyl-d(3,3',2,2')-propionate de sodium) comme référence. L'attribution des prolines, en particulier des formes *cis*, est réalisée dans le D<sub>2</sub>O 100% (même tampon).

L'attribution séquentielle des peptides par RMN est réalisée avec l'utilisation des séquences impulsionnelles homonucléaires TOCSY (temps de mélange 69 ms) et NOESY (temps de mélange de 400 ms) (Figure 3).



**Figure 3 :** Spectres homonucléaires TOCSY (à gauche) et NOESY (à droite) réalisé sur un peptide dans le D<sub>2</sub>O. Les formes *cis* des résidus Pro sont visibles dans le D<sub>2</sub>O à 5.1 et 4.85 ppm et sont invisibles sur les spectres acquis dans l'eau.

L'identification du type d'acide aminé est réalisée par exploitation du spectre TOCSY qui permet d'obtenir des couplages <sup>1</sup>H-<sup>1</sup>H à longue distance entre les différents protons d'un résidu acide aminé (couplages <sup>3</sup>J, <sup>4</sup>J, <sup>5</sup>J). Le spectre NOESY permet de définir des couplages entre des protons proches dans l'espace. On peut ainsi détecter des éléments de structure secondaire ou plus simplement, dans le cas de peptides non structurés, des couplages dipolaires entre les protons d'acides aminés voisins dans la séquence. Dans le cas des peptides de tau qui ont été synthétisés pour cette étude, on ne trouvera que des couplages H<sub>N</sub>(i)-H<sub>α</sub>(i-1) qui permettront d'identifier les résidus dans l'ordre de la séquence et de réaliser l'attribution séquentielle du peptide.



**Figure 4 :** Attribution séquentielle d'un peptide avec l'utilisation des spectres TOCSY (en noir) et NOESY (en rouge). Les  $H_{\alpha}$  (et les  $H_{\beta}$  dans le cas des résidus Ser et Thr) sont annotés pour les formes majoritaires correspondant aux résidus Pro en *trans* d'après les contacts NOE entre les protons  $H_{\delta}$  de la Pro (i) et les protons  $H_{\alpha}$  du résidu (i-1). Les formes minoritaires des résidus Ser396 et Ser404 en position (i-1) d'un résidu Pro en *cis* sont annotés en italique entre parenthèses.

**Récapitulatif des différents peptides synthétisés pour l'étude des interactions Pin1-tau.**

Peptide	Séquence	N-term	C-term	P.M.	MS (MALDI- TOF)	Pureté (HPLC) - Temps de rétention	Rendement (%)
<b>FRAGMENT [224-240]</b>							
T231 S235	K K V A V V R T P P K S P S S A K	Ac	amide	1822.22	1822.14 (+)	97% - 10.58 min	20%
		Biotine	amide	2005.47	2006.29 (+)	96% - 10.57 min	58%
pT231	K K V A V V R pT P P K S P S S A K	Ac	amide	1901.19	1902.28 (+)	92% - 10.02 min	46%
		Biotine	amide	2083.44	2084.63 (+)	98% - 10.68 min	44%
pS235	K K V A V V R T P P K pS P S S A K	Ac	amide	1901.19	1902.29 (+)	95% - 10.42 min	28%
pT231 pS235	K K V A V V R pT P P K pS P S S A K	Ac	amide	1981.17	1981.02 (-)	96% - 10.05 min	15%
tau pT231E	K K V A V V R E P P K S P S S A K	Ac	amide	1849.34	1850.38 (+)	95% - 10.37 min	39%
tau pT231D	K K V A V V R D P P K S P S S A K	Ac	amide	1835.34	1835.39 (+)	98% - 10.0 min	44%
C-pT231	C K K V A V V R pT P P K S P S S A K	Ac	amide	2004.29	2004.25 (-)	98% - 10.12 min	37%
<b>FRAGMENT [208-221]</b>							
T212 S214 T217	S R S R T P S L P T P P T R	Ac	amide	1593.82	1593.87 (+)	98% - 12.69 min	28%
		Biotine	amide	1778.05	1778.18 (+)	91% - 10.93 min	42%
pT212	S R S R pT P S L P T P P T R	Ac	amide	1673.80	1673.84 (-)	99% - 9.85 min	50%
		Biotine	amide	1858.05	1859.18 (+)	95% - 12.33 min	39%
		Dns	amide	1866.76	1866.15 (-)	98% - 13.06 min	43%
KpT212	K S R S R pT P S L P T P P T R	Ac	amide	1801.98	1800.34 (-)	93% - 9.47 min	49%
(AlexaFluor)KpT212	(AlexaFluor)K S R S R pT P S L P T P P T R	Ac	amide	2097.25	2094.60 (-)	90% - 9.82 min	40%
pT212 pS214	S R S R pT P pS L P T P P T R	Ac	amide	1753.78	1751.90 (-)	99% - 11.58 min	45%
pT212 pS214 pT217	S R S R pT P pS L P pT P P T R	Ac	amide	1833.76	1832.44 (-)	99% - 11.26 min	25%
<b>FRAGMENT [393-408]</b>							
S396 S404	V Y K S P V V S G D T S P R H L	Ac	amide	1783.03	1784.13 (+)	98% - 13.54 min	53%
pS396	V Y K pS P V V S G D T S P R H L	Ac	amide	1863.01	1861.02 (-)	92% - 12.31 min	21%
pS396 pS404	V Y K pS P V V S G D T pS P R H L	Ac	amide	1942.99	1940.04 (-)	92% - 12.44 min	13%
<b>FRAGMENT [416-428]</b>							
S422	S I D M V D pS P Q L A T L	Ac	amide	1430.65	1431.83 (+)	94% - 15.59 min	53%
pS422	S I D M V D pS P Q L A T L	Ac	amide	1510.63	1508.46 (-)	94% - 13.81 min	19%
pT212 pS214 pT231	G S P G T P G S R S R pT P pS L P T P P T R E P K K V A V V R pT P P K S P S S A K	Ac	amide	4381.40	4380.18 (-)	95% - 11.33 min	20%

**Tableau 3 :** Résumé des séquences peptidiques, des caractéristiques des parties N- et C-terminales, de la pureté déterminée par HPLC et du rendement de synthèse des peptides après purification.

***Synthèse d'un fragment peptidique de 40 acides aminés.***

Description de la synthèse :

Le peptide comporte 5 résidus enrichis en  $^{15}\text{N}$  : G204, G207, L215, A227 et V229.

La synthèse a été décomposée en trois parties :

- la synthèse du fragment compris entre K240 et R230 (dans le sens de synthèse, de C-vers N-term.), ensuite l'introduction de la  $^{15}\text{N}$ -V229 est réalisé par couplage manuel.
- la synthèse du fragment compris entre V228 et P216, ensuite l'introduction de la  $^{15}\text{N}$ -L215 est réalisé par couplage manuel.
- la synthèse du fragment compris entre pS214 et G201

La synthèse a été réalisée en simple couplage avec 10 équiv. d'acides aminés par couplage, pour le fragment compris entre K240 et R221 et en double couplage avec deux fois 5 équiv. pour le fragment compris entre T220 et G201.

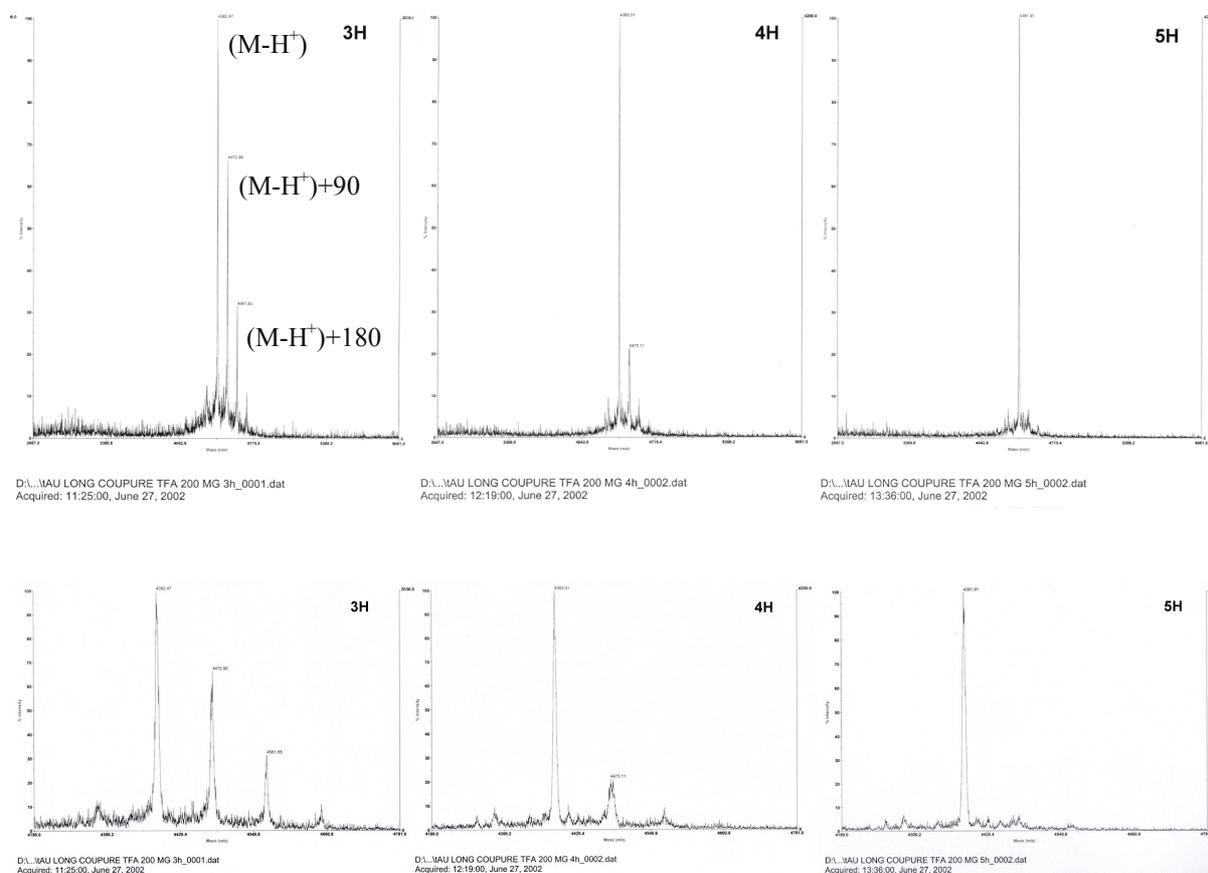
L'introduction des acides aminés Fmoc Gly et Ala enrichis en  $^{15}\text{N}$  a été effectuée en synthèse automatique comme pour les acides aminés non enrichis et les acides aminés phosphorylés. En revanche, nous disposions de peu de quantité de Fmoc Leu et Val enrichis en  $^{15}\text{N}$  qui ont nécessité une modification du nombre d'équivalents mis en jeu dans la réaction de couplage. Pour ces acides aminés, la réaction de couplage a été effectuée manuellement avec une fois 1.2 équiv. d'Ala et deux fois 1.2 équiv. de Leu.

Pour ce peptide, les feuilles récapitulatives de la synthèse, le suivi des réactions de déprotection des fonctions amines, les spectres de masse ainsi que les chromatogrammes avant et après purification sont donnés dans les pages suivantes à titre d'exemple.



## Matériels et Méthodes

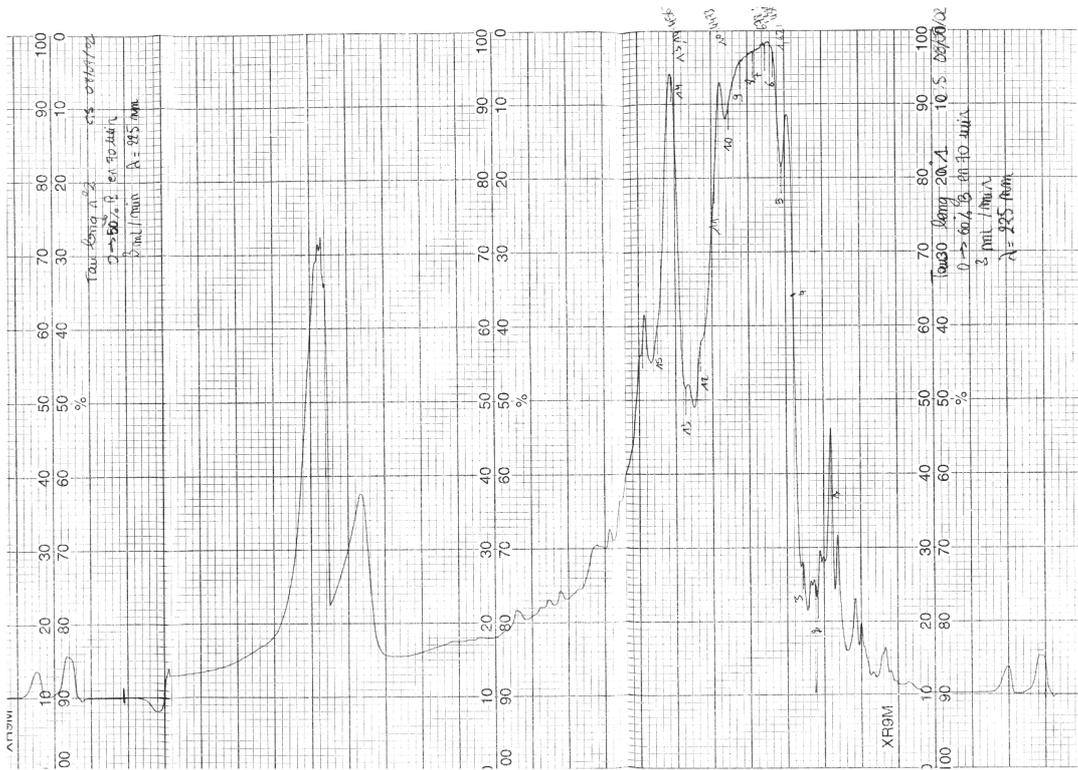
Suivi de la réaction de coupure du peptide dans le TFA par spectrométrie de masse MALDI-TOF.



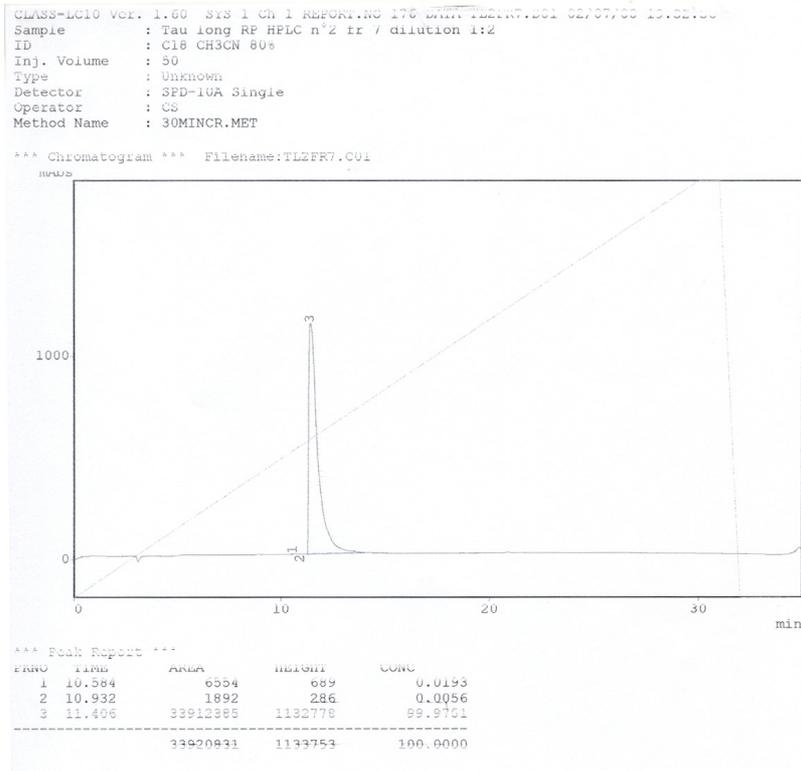
La réaction de coupure dans le TFA à 95% pour les peptides phosphorylés pose généralement des problèmes de déprotection du groupement phosphate portant un groupe protecteur de type benzyle. La coupure du peptide de la résine est généralement achevée en 2H (avec 50% de TFA) mais une augmentation des temps d'incubation de la résine dans le TFA jusqu'à 5H permet d'améliorer significativement la coupure du benzyle. La réaction peut être suivie par spectrométrie de masse MALDI-TOF en prélevant 1  $\mu$ L de solution de coupure dans le milieu réactionnel que l'on dilue au 1:1000<sup>e</sup> dans l'eau. En comparant l'intensité des pics correspondant aux masses  $(M-H^+)$ ,  $(M-H^+)+90$  et  $(M-H^+)+180$ , on peut obtenir une information quantitative (ces données ont été confirmées par HPLC) très rapidement sur l'avancée de la réaction.

# Matériels et Méthodes

## Profil chromatographique du peptide brut et purification à l'échelle semi-préparative.



## Profil chromatographique du peptide après purification.



## II – Expression et purification de protéines recombinantes.

La mise en oeuvre de la production de protéines par voie recombinante possède plusieurs intérêts comme la quantité de protéine produite, la possibilité d'introduire dans la partie N- ou C-terminale de la protéine d'intérêt une protéine (ou un peptide) de fusion pour faciliter la purification et le marquage isotopique de la protéine d'intérêt pour des études par RMN. Les quantités de protéine produites peuvent atteindre ainsi plusieurs dizaines de milligrammes selon le niveau d'expression qui dépend du vecteur d'expression, de l'organisme hôte, du milieu de culture et permettent de satisfaire aux exigences d'une étude par RMN notamment pour l'attribution séquentielle et la détermination de la structure tri-dimensionnelle. Les nouvelles technologies RMN comme le développement de cryosondes permettent une augmentation de la sensibilité et donc de diminuer les concentrations en protéine de l'échantillon tout en obtenant un rapport signal sur bruit correct.

**Une petite banque de plasmides.** Nous disposons de protéine Pin1 sous diverses formes pour les études des interactions : des plasmides contenant la protéine dans sa totalité, les domaines WW et catalytiques isolés, des mutants, avec deux systèmes de protéine de fusion (His-tag et GST) selon les besoins (Tableau 4). Nous disposons également des isoformes longues (441 résidus) et fœtale (352 résidus) de la protéine Tau humaine ainsi que d'un mutant contenant 10 résidus Glu remplaçant des résidus Ser ou Thr. Toutes les protéines sont exprimées chez *Escherichia coli* (souche BL21).

<i>Protéine</i>	<i>Vecteur</i>
Pin1	pET15b (His-tag) pGEX-4T
WW	pET28a (His-tag) pGEX-4T
Pin1 <sub>CAT</sub>	pGEX-6P
Pin1 $\Delta$ [44-49]	pET15b (His-tag)

**Tableau 4 :** Liste des vecteurs disponibles pour la production des différentes protéines recombinantes de la famille de Pin1 utilisées pour les études d'interaction.

Les clonages ont été réalisés par le Dr Isabelle Landrieu (UMR CNRS 8525) ainsi que par Séverine Bégard de l'équipe du Dr Luc Buée (U422 Inserm).

Les protéines tau ont été fournies par deux collaborateurs du laboratoire (UMR CNRS 8525) dont les thématiques de recherches sont centrées sur la protéine tau, le Dr. Alain Sillen (agrégation de la protéine tau *in vitro*) et le Dr. Arnaud Leroy (phosphorylation de la protéine tau).

### **2.1 – Expression des protéines Pin1, WW et Pin1<sub>CAT</sub> par technique recombinante.**

Les protocoles suivis pour l'expression et la purification des protéines recombinantes Pin1 et ses deux domaines isolés, ainsi que le mutant Pin1 $\Delta$ [44-49], sont des protocoles relativement classiques comprenant :

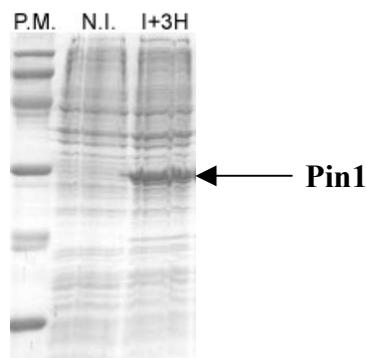
- une culture soit en milieu riche LB pour l'obtention de protéine non enrichies, soit en milieu minimal M9 pour l'obtention de protéines enrichies <sup>15</sup>N ou <sup>15</sup>N/<sup>13</sup>C.
- une première étape de purification par chromatographie d'affinité en fonction la protéine de fusion (HisTag ou GST)
- une seconde étape de purification soit par chromatographie d'exclusion stérique pour Pin1, soit par phase inverse pour le domaine WW.

Dans le cas où la protéine de fusion est la GST, on procède à une étape de digestion enzymatique entre les deux étapes de purification.

**Précultures.** La préculture est effectuée en milieu LB, avec ajout de l'antibiotique (20 mg/L pour la kanamycine ou 100 mg/mL pour l'ampicilline) et ensemencement soit à partir d'une colonie prélevée sur boîte, soit à partir d'une conservation dans le glycérol 50%. Le milieu est incubé à 37°C pendant la nuit.

#### **2.1.1 – Expression de protéines non marquées en milieu riche.**

Les protéines non enrichies en élément isotopique sont produites en milieu riche de type LB. Les milieux sont incubés à 37°C pendant 3 à 4 heures par dilution de la préculture au 1 :50<sup>e</sup> jusqu'à ce que la turbidité du milieu lue par mesure de l'absorbance à 600 nm atteigne la valeur 0.6. L'induction est effectuée avec 0.5mM d'IPTG en concentration finale à partir d'une solution à 1M et les milieux sont incubés à 31°C pendant 3 à 4 heures. Les cellules sont ensuite collectées par centrifugation (6000 RPM – 30 min – 4°C) et resuspendues dans 40 mL de tampon d'extraction pour 1 à 2 L de culture puis congelées.



**Figure 5** : Suivi de l'expression de Pin1 sur gel SDS PAGE 12.8%. P.M., poids moléculaires ; N.I., non induit ; I+3H, après 3H d'induction.

*Les poids moléculaires utilisés dans les différents gels sont les suivants :*

*14.4, 21.5, 31, 45, 66 et 97*

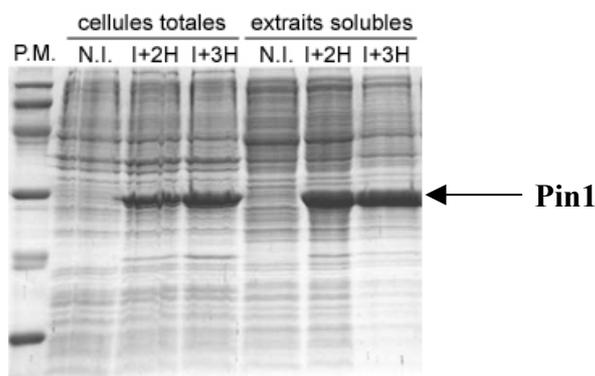
### **2.1.2 – Expression de protéine enrichie $^{15}\text{N}$ ou $^{15}\text{N}/^{13}\text{C}$ en milieu minimum.**

Dans un premier temps, l'incubation des bactéries à 37°C pendant 3 à 4 heures en milieu riche LB permet d'atteindre rapidement une haute densité de cellules contrôlée par mesure de l'absorbance à 600 nm (comme décrit dans le paragraphe précédent). Lorsque l'absorbance atteint une valeur de 0.6, les cellules sont collectées par centrifugation (3000 RPM – 30 min – 20°C) et resuspendues en milieu minimal M9 sélectivement enrichi en  $^{15}\text{N}$  par ajout de  $^{15}\text{NH}_4\text{Cl}$  à 1g/L comme unique source d'azote et/ou enrichi en  $^{13}\text{C}$  par ajout de [ $^{13}\text{C}$ ]-glucose à 4g/L comme seule source de carbone. Les milieux sont incubés à 37°C pendant 1 heure pour permettre le remplacement progressif de l'azote et du carbone en abondance naturelle par les isotopes  $^{15}\text{N}$  et/ou  $^{13}\text{C}$  au sein des acides aminés synthétisés par la bactérie. Lors du changement de milieu, les cellules peuvent être resuspendues dans le même volume ou concentrées de 2 à 4 fois pour réduire les quantités de  $^{15}\text{NH}_4\text{Cl}$  et de [ $^{13}\text{C}$ ]-glucose utilisées. L'induction est réalisée à 31°C par ajout de 0.5 mM d'IPTG en concentration finale. Les cellules sont incubées pendant 3 à 4 heures puis collectées par centrifugation (6000 RPM – 30 min – 4°C) et resuspendues dans un tampon d'extraction (40 à 45 mL de tampon d'extraction pour 1 à 2 L de culture). La suspension est ensuite congelée jusqu'à l'étape de purification. La production de Pin1 et du domaine WW isolé dans ces conditions est estimée à 10 mg/L environ.

La solubilité de la protéine dans le tampon d'extraction est contrôlée sur gel SDS-PAGE après lyse des cellules (Figure 6) par comparaison avec la bande de protéine obtenue dans les

## Matériels et Méthodes

cellules totales (la quantité de cellules prélevées dans le milieu de culture correspond au volume équivalent à un volume de 1 mL dont la densité mesurée à 600 nm est de 0.1).



**Figure 6 :** Suivi de l'expression de Pin1 en milieu minimal M9 sur gel SDS PAGE 15% et contrôle de la solubilité de la protéine dans le tampon d'extraction. P.M., poids moléculaire ; N.I., non induit ; I+2H et I+3H, après 2H et 3H d'induction respectivement, dans les cellules et dans l'extrait soluble (la même quantité de cellule est prélevée dans le milieu de culture dans les deux cas).

### Caractéristiques des composés utilisés dans les cultures pour l'enrichissement isotopique des protéines en $^{15}\text{N}/^{13}\text{C}$ .

La concentration des volumes de culture de 2 ou 4 fois lors du passage du milieu LB au milieu M9 nécessite l'utilisation de quantités moins importantes de  $^{15}\text{NH}_4\text{Cl}$  et de  $[^{13}\text{C}]$ -glucose, permettant ainsi des économies substantielles si on considère le coût de ces composés (Tableau 5).

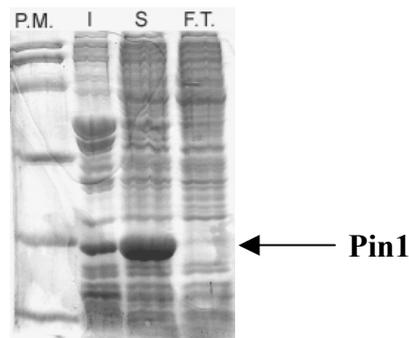
Nom	Chlorure d'ammonium	(D)-Glucose
Isotope	( $^{15}\text{N}$ , 98%+)	(U- $^{13}\text{C}$ 6, 99%)
Code produit	NLM-467-5	CLM-1396-1
Quantité	5 G	1 G
Prix	\$190.00	\$210.00
Formule	* $\text{NH}_4\text{Cl}$	$\text{HO}^*\text{CH}_2(^*\text{CHOH})_4^*\text{CHO}$

**Tableau 5 :** Caractéristiques du chlorure d'ammonium et du glucose (Cambridge Isotope Laboratories) utilisés pour l'enrichissement isotopique des protéines pour les études par RMN.

## 2.2 – Purification de Pin1, de Pin1<sub>CAT</sub> et du domaine WW par chromatographie d'affinité (1<sup>re</sup> étape de purification).

### 2.2.1 – Purification par affinité d'une queue poly-histidine (His-tag) pour le nickel (Pin1, WW).

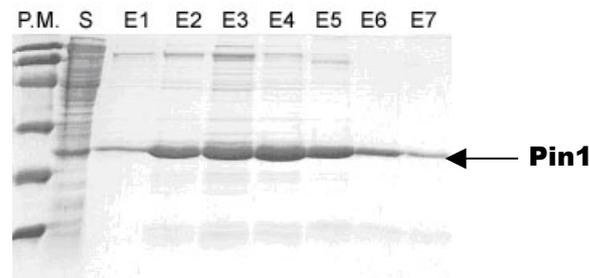
**Lyse des cellules et extraction dans la fraction soluble.** Les cellules issues de la culture sont resuspendues dans un tampon d'extraction constitué de 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.5 mM DTT, 1% NP40, cocktail d'inhibiteurs de protéases (Roche). Les cellules obtenues à partir d'un à deux litres de culture sont resuspendues dans 40-45 mL de tampon d'extraction. La suspension est congelée jusqu'à la purification puis décongelée au moment de la lyse. On ajoute 10 mg de lysozyme pour 40-45 mL de tampon d'extraction, puis les cellules sont lysées par sonication. L'extrait soluble est isolé par centrifugation (15000 RPM à 4°C pendant 30 min) puis filtré à 0.22µ.



**Figure 7 :** Lyse des cellules BL21 surexprimant Pin1 dans le vecteur pET15b par sonication. Les fractions insolubles (I) et solubles (S) sont séparées par centrifugation. L'extrait soluble contenant la majorité de la protéine produite est chargé sur la colonne de nickel. La fraction des protéines qui n'est pas retenue sur la colonne (F.T.) ne contient pas la protéine d'intérêt.

**Purification sur colonne HiTrap Chelating 5 mL chargée en nickel (Amersham).** La colonne est rincée dans l'eau avec 5 volumes de colonne, chargée en 1 volume de colonne avec une solution de NiSO<sub>4</sub> à 0.1M, et enfin rincée à l'eau avec 5 volumes de colonne. Une fois chargée en nickel, la colonne peut être réutilisée plusieurs fois pour la purification d'une même protéine. L'extrait est chargé en un seul passage à 1mL/min sur la colonne équilibrée dans un tampon contenant 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 8.0 et 300 mM NaCl (tampon A) auquel on ajoute 10 mM imidazole. La colonne est ensuite lavée à 2mL/min avec le tampon A

contenant 20 mM d'imidazole et la protéine est éluée à 2 mL/min avec le tampon A à pH 7.80 contenant 250 mM d'imidazole. La protéine est éluée sur 1.5 à 2 volumes de colonne.



**Figure 8 :** Suivi par gel SDS PAGE 15% de la purification de Pin1 (pET15b) par affinité de la queue His-tag pour le nickel. S, fraction soluble totale ; E1 à E7, fractions d'élution 1 à 7 (1 à 2 mL par fraction).

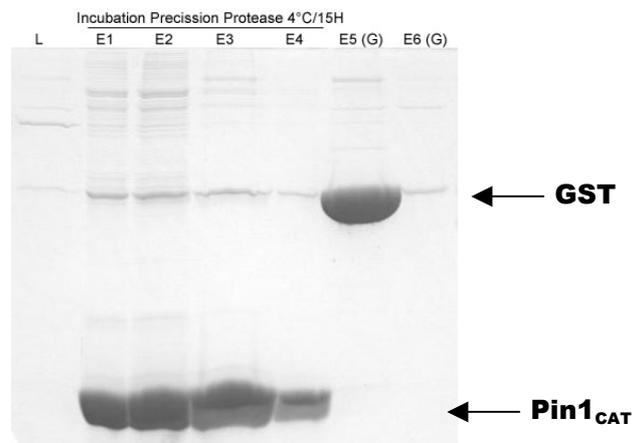
### 2.2.2 – Purification par affinité de la GST pour le glutathion (WW, Pin1 et Pin1<sub>CAT</sub>).

**Purification du domaine WW en fusion GST.** Après décongélation des culots resuspendus dans le tampon d'extraction (50 mM PBS pH=7.3; 2 mM DTT ; 10 mM EDTA; 0.1% Triton X-100 ; cocktail d'inhibiteurs de protéases), les cellules sont lysées par lyse chimique avec 15mg de lysozyme / 40mL d'extrait (40 mL d'extrait pour 1L de culture). La viscosité est réduite par ajout de RNase/DNase et l'extrait soluble est isolé par centrifugation puis filtré (0.22µ). L'extrait est chargé sur une colonne (GSTrap Glutathione Sepharose 4B Fast Flow 5mL, Amersham Pharmacia) équilibrée dans le PBS pH=7.3 à 0.5mL/min. La colonne est ensuite lavée avec 100mL de PBS pH=7.3; 5mM EDTA. On élue la protéine de fusion dans un tampon Tris 50 mM pH=8.0 ; NaCl 100 mM ; EDTA 5 mM ; Glutathion 10 mM avec 1 à 2 volumes de colonne.

**Purification de Pin1 et du domaine catalytique Pin1<sub>CAT</sub> en fusion GST.** Le protocole est identique à celui décrit précédemment pour le domaine WW pour la lyse des cellules, la charge et le lavage de la colonne. L'élution de la protéine d'intérêt est réalisée par digestion dite « on-colonne » avec la thrombine dans le cas de Pin1 ou la Precision Protease® (Amersham) dans le cas du domaine PPIase. La colonne est préalablement équilibrée avec 5 volumes de colonne de tampon d'élution sans protéase (PBS pH=7.3 ; EDTA 5mM pour Pin1 et Tris 50 mM pH=7.8 ; NaCl 150 mM ; EDTA 1 mM ; DTT 1 mM pour le domaine PPIase). La digestion est réalisée avec 200 unités de protéase pour une colonne de 5 mL

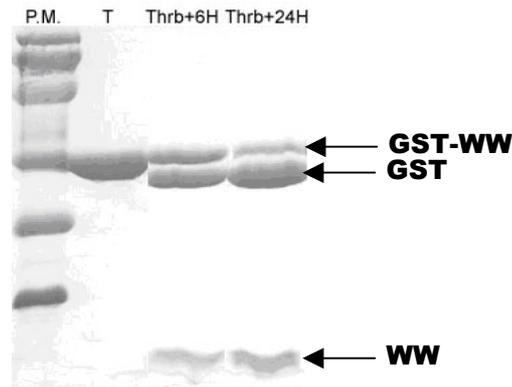
## Matériels et Méthodes

(correspondant à 8 mg de protéine environ) dans 1 volume de colonne de tampon d'élution. La colonne est incubée à 20°C pour la thrombine et à 4°C pour la Precision Protease® pendant 24H. La protéine d'intérêt est éluée dans 1 volume de colonne de PBS pH=7.3 (Figure 9). La colonne est ensuite régénérée par élution de la GST (et éventuellement de la protéine de fusion non digérée) avec le tampon utilisé pour l'élution de la GST-WW. Il est à noter que la Precision Protease®, contrairement à la thrombine, est fusionnée à une GST et se lie à la phase stationnaire. Dans le cas d'une digestion à la thrombine, 2 mM de PMSF en concentration finale sont ajoutés pour éviter une dégradation ultérieure de la protéine d'intérêt par protéolyse.



**Figure 9 :** Suivi de l'élution de Pin1<sub>CAT</sub> sur gel SDS PAGE 15% après digestion « on-colonne » de la GST-Pin1<sub>CAT</sub> par incubation avec la Precision Protease (Amersham) pendant 15H à 4°C. L, lavage ; E, fractions d'élution : E1 à E4, élution par passage du tampon sans protéase après incubation de la colonne avec la Precision Protease pendant 15H/4°C, E5-E6 (G), élution de la GST digérée avec ajout de 10 mM de glutathion dans le tampon.

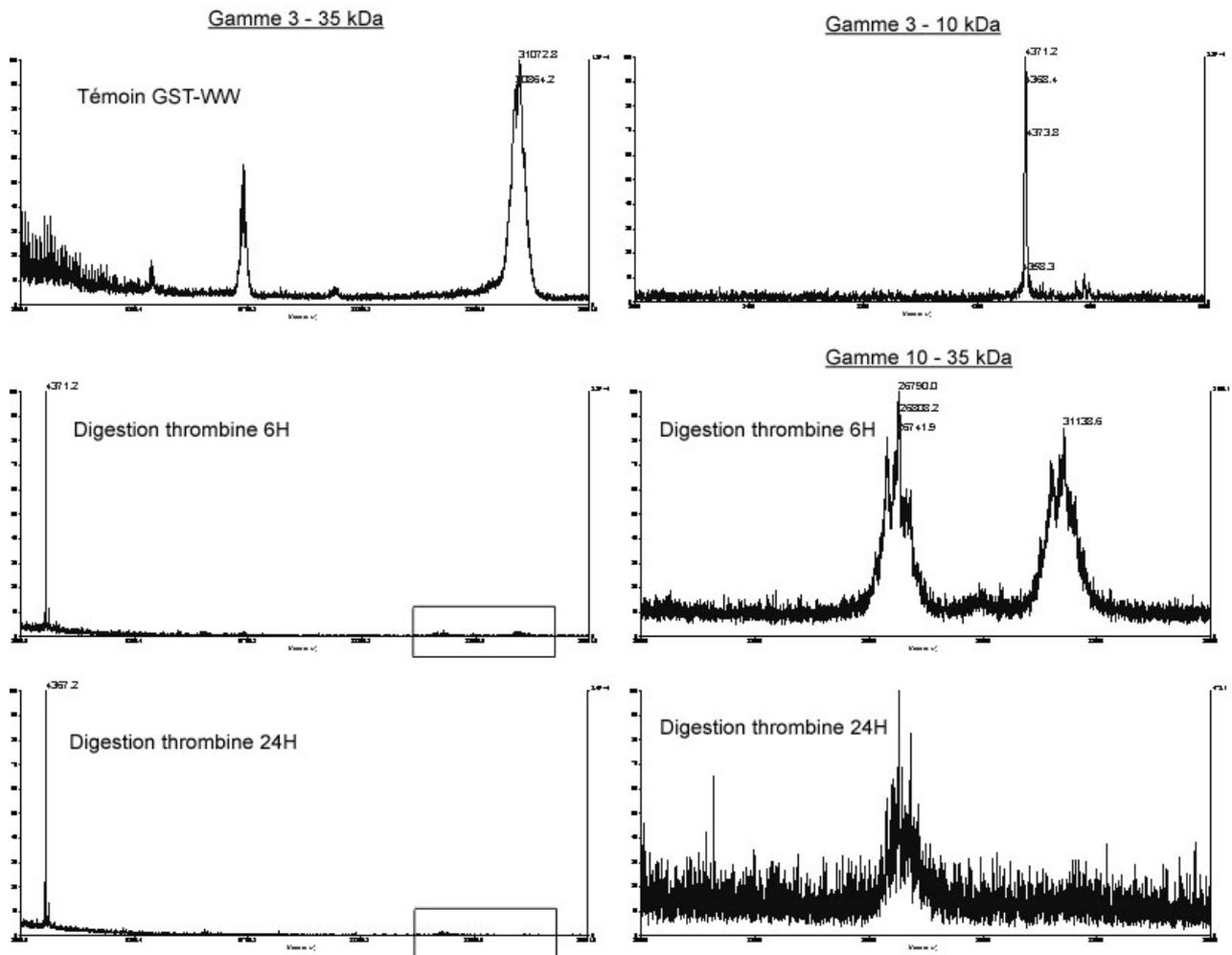
**Digestion du domaine WW fusionné à la GST.** Si la digestion de la queue d'histidines n'est pas obligatoire pour les études par RMN car il s'agit d'un petit fragment peptidique d'environ 1.5 kDa (pET15b), non structuré, qui n'interagit pas avec la protéine d'intérêt (cela a été vérifié) et ne modifie donc pas les résonances des résidus du domaine WW, la digestion de la GST est obligatoire car il s'agit d'une protéine de 26kDa. La digestion de la GST est réalisée par incubation de 0.1 unité de thrombine par mg de protéine directement dans le tampon d'élution à 20°C pendant 24H. La thrombine et la GST sont ensuite éliminées après digestion lors d'une seconde étape de purification par RP-HPLC (voir paragraphe 2.4).



**Figure 10 :** Suivi de la digestion de la protéine GST fusionnée au domaine WW par la thrombine dans le temps sur gel SDS PAGE 15%.

### **Suivi de la digestion de la GST fusionnée au domaine WW par spectrométrie de masse MALDI-TOF.**

Le suivi de la digestion peut être également réalisé par MALDI-TOF qui permet de suivre la réaction quasiment en temps réel par rapport au suivi par chromatographie en gel dénaturant SDS PAGE. Le domaine WW s'ionisant beaucoup mieux que les protéines GST ou GST-WW, la comparaison de l'aire des signaux n'est pas proportionnelle aux concentrations relatives des différentes protéines en solution. En revanche, en travaillant sur la gamme des masses allant de 10 à 35 kDa, il est possible de comparer les signaux de la GST et de la protéine de fusion GST-WW (Figure 11) dont l'aire des pics est proportionnelle aux bandes observées sur gel SDS PAGE (Figure 10).



**Figure 11 :** Suivi par spectrométrie de masse MALDI-TOF de la cinétique de digestion de la GST-WW par la thrombine. A gauche, spectres de masse réalisés sur le mélange brut de digestion sur une gamme allant de 3000 à 35 000 Da pour différents temps de réaction : 0, 6 et 24 heures. A droite, les spectres de masse correspondant à une gamme de détection entre 3 et 10 kDa (en haut) et 10 à 35 kDa pour 6 et 24 heures de réaction (en bas).

### 2.3 – Purification de Pin1 par chromatographie d’exclusion stérique (2<sup>e</sup> étape de purification).

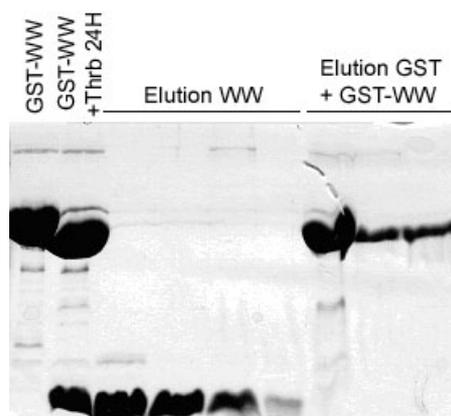
La colonne Sepharose PrepGrade (Amersham) de 100 mL est équilibrée dans un tampon contenant 50 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> pH 7.6, 250 mM NaCl, 5 mM β-mercaptoéthanol avec 1.5 volumes de colonne. Les fractions de la première étape de purification par affinité sont rassemblées (entre 3 et 5 mL au total) et injectées directement à partir du tampon d’élution de l’étape de purification sur colonne de nickel. La protéine est éluée dans le tampon d’équilibration décrit ci-dessus avec un volume d’élution de 75 mL environ.



**Figure 12 :** Evaluation de la pureté de Pin1 après une seconde étape de purification par chromatographie d'exclusion stérique.

#### 2.4 – Purification du domaine WW par RP-HPLC (2<sup>e</sup> étape de purification).

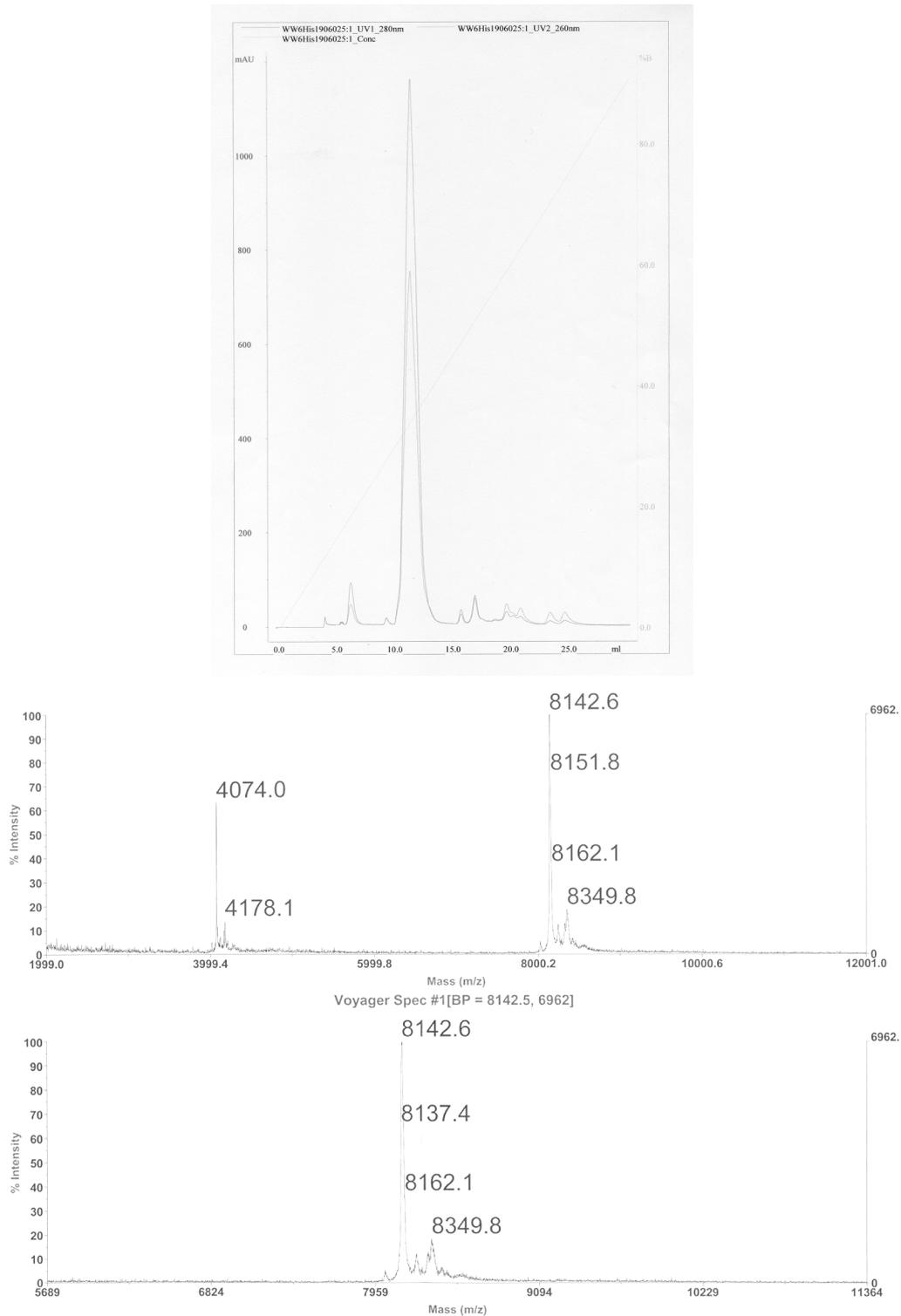
L'HPLC phase inverse est utilisée comme seconde étape de purification du domaine WW. Cette seconde étape est aussi utilisée après digestion de la GST ou de la queue d'histidines avec la thrombine et permet ainsi d'éliminer les fragments après digestion ainsi que la thrombine. L'obtention de protéine sous forme lyophilisée présente également l'avantage de faciliter et de prolonger la conservation. Les fractions d'éluion digérées ou non par la thrombine sont injectées sur une colonne SOURCE 15RPC ST4.6/100 (Amersham) de 1.667mL équilibrée au préalable dans une solution aqueuse contenant 0.05% TFA (v/v). L'éluion est réalisée avec un gradient d'acétonitrile à 1mL/min de 0 à 85% d'acétonitrile en 20 volumes de colonne. Les fractions collectées sont contrôlées par spectrométrie de masse MALDI-TOF et sur gel dénaturant SDS PAGE à 15% acrylamide. Les fractions 1 à 4 contiennent le domaine WW pur, les fractions suivantes contiennent en mélange la GST digérée et la GST fusionnée au WW.



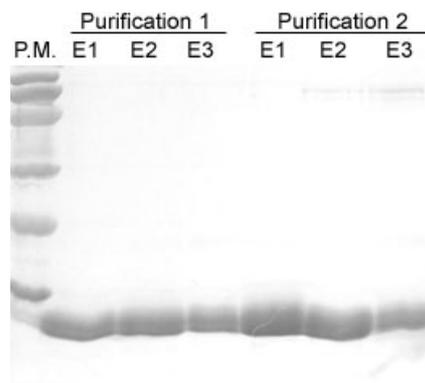
**Figure 13 :** Suivi de la purification du mélange GST-WW, GST et WW par RP-HPLC après coupure à la thrombine sur gel SDS PAGE 15%.

## Matériels et Méthodes

Le domaine WW comportant un queue poly-histidine est également purifié par phase inverse dans les mêmes conditions qui ont été décrites précédemment pour la purification du mélange issu de la digestion de la protéine de fusion GST-WW.



**Figure 14 :** Profil chromatographique de purification du domaine WW portant une queue de 6 His fusionnée en N-term. par RP-HPLC et spectre de masse ( $m/z = 8140$ ).



**Figure 15** : Seconde étape de purification du domaine WW portant une queue de 6 His fusionnée en N-term. par RP-HPLC sur les fractions d'élution de la première étape de purification par affinité. La purification a été réalisée en deux fois car la quantité de protéine est limitée à 5 mg. E1 à E3, fractions d'élution 1 à 3.

Après la purification par RP-HPLC, les fractions contenant le domaine WW pur sont rassemblées et lyophilisées.

### 2.5 – Dessalage et préparation des échantillons dans les tampons utilisés pour la RMN ou la fluorescence par chromatographie d'exclusion sur gel non résolutif.

Les colonnes utilisées sont de type Sephadex G-25, NAP-10 ou NAP-5 selon le volume de l'échantillon. Les spécifications pour les volumes de charge et d'élution sont données dans le **Tableau 6**.

Les colonnes sont équilibrées dans les tampons utilisés pour les expériences de RMN (avec 3-4 volumes de colonne) : Tris- $d_{11}$  25 mM pH 6.80, 50 mM NaCl, 5 mM EDTA, 5 mM DTT. Les spécifications des différentes colonnes en fonction du volume d'échantillon à échanger sont données dans le Tableau 6. On charge la colonne avec le volume d'échantillon spécifié puis on élue la protéine avec le même tampon. Les volumes d'élution de la protéine sont donnés dans le Tableau 6 en fonction de la colonne utilisée.

	<i>Volume de colonne (mL)</i>	<i>Volume de charge (mL)</i>	<i>Volume d'élution (mL)</i>
<b>G-25</b>	5	2.5	3
<b>NAP-10</b>	2	1.0	1.5
<b>NAP-5</b>	1	0.5	1.0

**Tableau 6** : Volumes de charges et d'élution pour les différentes colonnes de dessalage.

Cette méthode est utilisée préférentiellement à la dialyse qui nécessite plusieurs litres de tampon pour échanger quelques millilitres d'échantillon, pour des raisons de coûts du tampon Tris deutéré (Tableau 7).

## Matériels et Méthodes

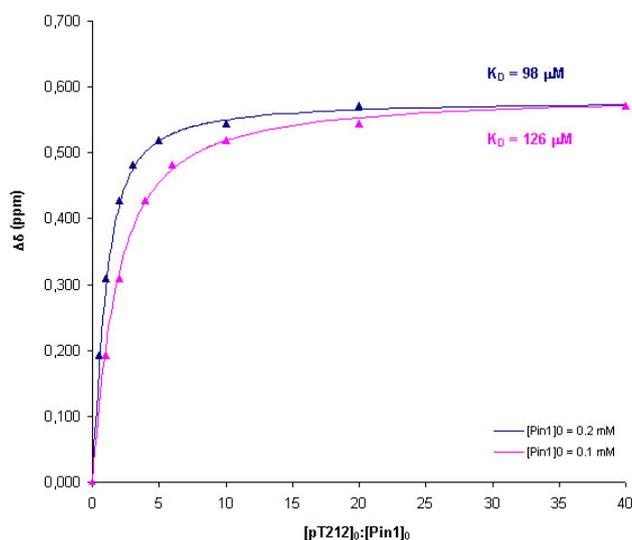
Nom	TRIS
Isotope	(D11, 98%)
Code produit	DLM-1814-5
Quantité	5 G
Prix	\$465.00
Formule	(DOCD2)3CND2

**Tableau 7 :** Caractéristique du Tris-d11 (Cambridge Isotope Laboratories) utilisé pour les études d'interactions par RMN.

La concentration de la protéine contenue dans les fractions d'éluion est estimée par dosage Bradford (voir ci-dessous). Les fractions contenant la protéine sont rassemblées et concentrées si nécessaire par ultrafiltration (Vivaspin 2). Pour Pin1 et Pin1<sub>CAT</sub>, on utilise une membrane avec un *cut-off* de 5 kDa, pour la protéine tau, un *cut-off* de 3 kDa.

### **Détermination de la concentration de Pin1 et du domaine WW en solution.**

La connaissance de la concentration initiale de protéine en solution est déterminante dans le calcul de la constante de dissociation d'autant plus quand elle se situe dans la même gamme que le  $K_D$ . On a ainsi remarqué qu'une erreur dans la détermination de la concentration en protéine conduisait à des variations significatives de la constante de dissociation (Graphique 1).



**Graphique 1 :** Calcul de la constante de dissociation d'un complexe avec deux valeurs différentes de concentration de protéine initiale, en bleu 200  $\mu\text{M}$  et en rose 100  $\mu\text{M}$ , à partir des mêmes valeurs expérimentales. Une diminution de la concentration de protéine d'un facteur 2 entraîne une variation du  $K_D$  de 28%.

De plus, nous avons remarqué une variation notable des signaux RMN des échantillons de protéines en solution mis en œuvre pour les différentes titrations.

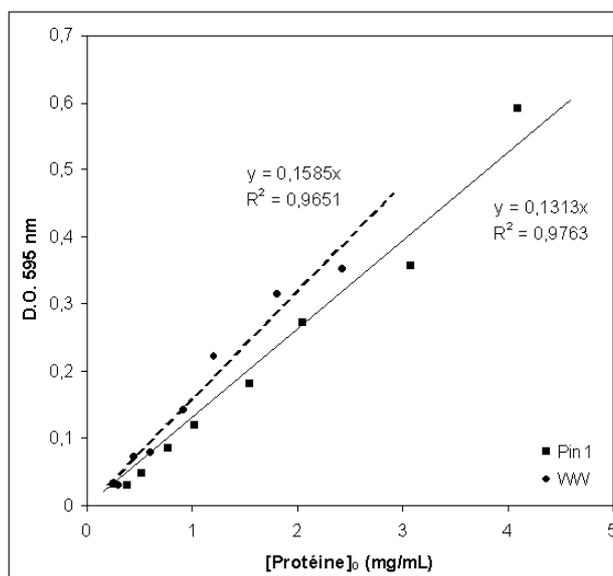
Notre but est de comparer les affinités de différents peptides de tau pour la protéine Pin1 et son domaine de liaison isolé par titration. Travailler sur un lot de protéine utilisable pour chacune des titrations est un moyen de pouvoir établir, sans ambiguïté sur la concentration de la protéine, des affinités relatives, à défaut de résoudre la question de l'affinité absolue. Cependant, l'acquisition de spectres RMN corrects d'un point de vue rapport signal sur bruit nécessite des concentrations de protéine élevées dans des conditions restrictives de tampon et de pH. L'effort fourni pour atteindre de telles concentrations, souvent très au-dessus des concentrations cellulaires, implique des phénomènes de dégradation et d'agrégation importants, défavorables à la production d'un lot de protéine.

Le choix de la technique de mesure utilisée pour déterminer la concentration de protéine en solution dépend de plusieurs facteurs comme la quantité de protéine disponible, les produits entrant dans la composition du tampon, la protéine elle-même. Les techniques les plus couramment utilisées sont les techniques colorimétriques (Bradford, Lowry, BCA) qui mettent à profit un changement de coloration d'une solution lors de la réaction avec la protéine. Le test Bradford (Bradford, 1976) met en jeu la variation du maximum d'absorption d'un colorant, le bleu de Coomassie G-250. La concentration de la protéine est obtenue à partir d'une courbe d'étalonnage et de l'application de la loi de Beer dans la mesure où le coefficient d'extinction molaire du complexe colorant-protéine est constant sur une gamme de concentrations données. Le test est basé sur un changement colorimétrique du colorant qui peut exister sous 3 formes : une forme cationique de couleur rouge, une forme anionique de couleur bleue et une forme neutre de couleur verte. Au pH du test, le colorant est sous forme cationique. La liaison avec la protéine, par l'intermédiaire des résidus basiques, en particulier l'arginine, et des résidus aromatiques, entraîne la stabilisation de la forme anionique qui peut être détectée à 595 nm. La limite de la méthode vient de la présence dans le tampon de la protéine de composés capables de modifier l'équilibre entre les 3 formes du colorant. Certains détergents ou tampons basiques créent des interférences dans la mesure en stabilisant la forme neutre du colorant. Le test de Lowry (Lowry et al., 1951) pallie à ce type de limites en adaptant le test Bradford aux protéines solubilisées dans un tampon contenant de hautes concentrations en détergent. Les concentrations de Pin1 et du domaine WW en solution sont déterminées par dosage des protéines en utilisant la méthode de Bradford pour laquelle une courbe d'étalonnage a été établie à partir de dilutions successives d'une solution-mère dont la concentration a été déterminée par hydrolyse acide totale.

## Matériels et Méthodes

Le dosage des acides aminés présents dans une protéine après hydrolyse totale de celle-ci est une technique qui permet, en connaissant la séquence de la protéine, de déterminer sa concentration. Elle repose sur 2 étapes : (i) l'hydrolyse de la protéine par une solution d'acide chlorhydrique 6N à 110°C pendant 24 à 48H et (ii) le dosage des acides aminés libérés par chromatographie. Avant analyse par HPLC, les acides aminés libres sont dérivatisés par réaction avec le phénylthiocyanate (PITC) à pH basique. Les dérivés phénylthiocarbamyls obtenus sont injectés sur une colonne C18, élués par un gradient d'acétonitrile et détectés à 254 nm. La norleucine est ajoutée dans l'échantillon comme étalon interne. Parallèlement, une solution standard d'acides aminés dérivatisés analysée sur un autre canal est utilisée comme calibration pour déterminer la quantité de chaque acide aminé présent dans l'échantillon. Pour traiter les données, il convient de connaître les artefacts dus à l'hydrolyse acide ou à la réaction de dérivatisation. En effet, certains acides aminés sont dégradés pendant l'hydrolyse acide, comme le tryptophane, la thréonine, la sérine et la cystéine ou partiellement oxydé comme la méthionine. L'asparagine et la glutamine sont converties en acide aspartique et glutamique respectivement. Certains tampons, dont le Tris, et détergents interfèrent dans le processus de dérivatisation.

### *Courbes d'étalonnage Bradford établies pour Pin1 et le domaine WW isolé.*

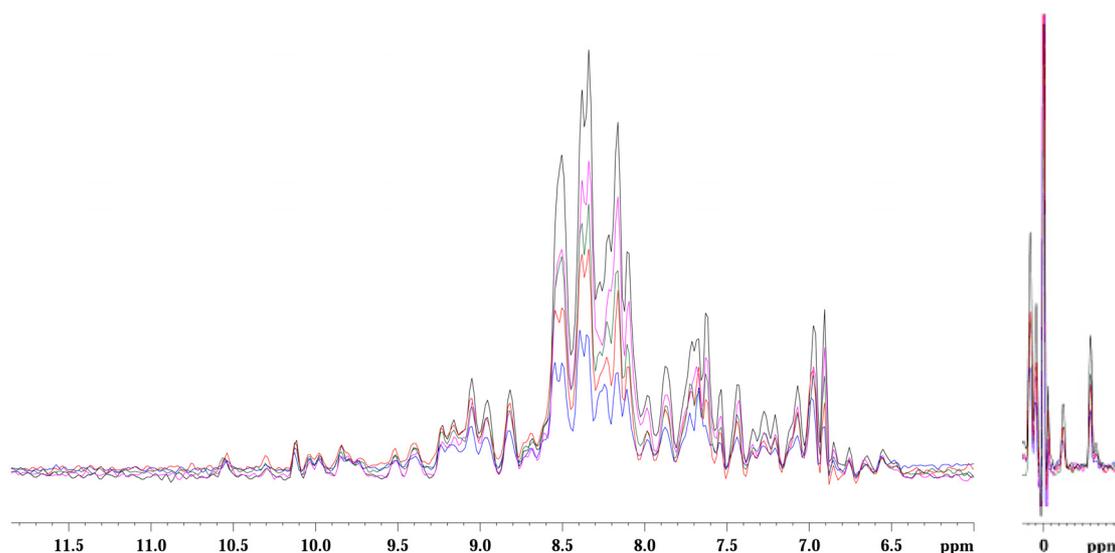


**Graphique 2 :** Courbes d'étalonnage basées sur le test Bradford pour les protéines WW et Pin1. Les valeurs de la densité optique (D.O.) à 595 nm sont mesurées pour des solutions de protéine de concentrations variables obtenues par dilutions successives à partir d'une solution-mère dont la concentration a été déterminée par hydrolyse acide totale.

## Matériels et Méthodes

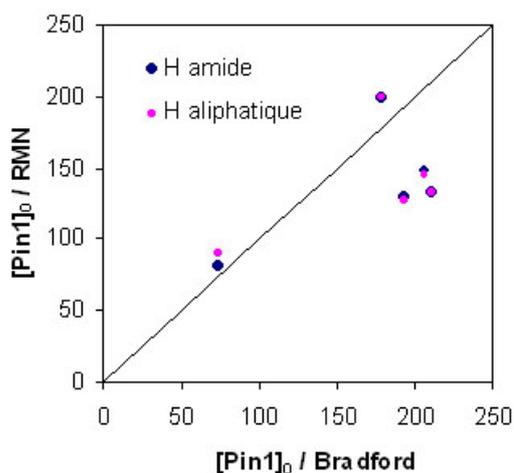
La détermination de la concentration de protéine en solution est réalisée à partir du test colorimétrique de Bradford par mesure de la D.O. à 595 nm pour 2  $\mu$ L de solution de protéine diluée dans 1 mL de solution de colorant commercial diluée au  $1/5^e$ . Une courbe d'étalonnage est établie de la même manière par dilution successive au  $1/2^e$  à partir d'une solution-mère concentrée dont la concentration est déterminée par dosage des acides aminés après hydrolyse acide totale. Trois échantillons sont préparés pour chaque concentration de protéine en solution. Chaque point de la droite d'étalonnage correspond à la moyenne des trois lectures de D.O. pour chaque concentration de protéine.

Les concentrations des échantillons de protéine utilisés pour les différentes titrations ont été déterminées à partir de la courbe d'étalonnage basée sur le test Bradford et concordent avec les variations observées au niveau l'intensité des signaux RMN entre les différents échantillons. Une calibration des concentrations de protéine dans les différents échantillons de titration RMN peut ainsi être mise en œuvre sur base d'une comparaison de l'intensité des résonances protons en 1D. Cette méthode nécessite que les spectres soient acquis dans les mêmes conditions, que les paramètres de traitement des données (*processing*) soient les mêmes et n'est valable que pour comparer deux spectres de la même protéine, dans la mesure où la largeur de raie varie d'une protéine à l'autre. La calibration nécessite un phasage correct des spectres et le choix d'une ou plusieurs résonances isolées et, de préférence, appartenant à des protons non échangeables car moins sensibles à l'environnement, en particulier au pH. On choisit en général les protons des groupements méthyle des résidus valine, leucine ou isoleucine dont certains résonnent entre 0 et  $-2$  ppm dans les protéines structurées en raison de la proximité de groupements aromatiques dans le cœur hydrophobe (Figure 16).



**Figure 16 :** Représentation du spectre proton pour différentes solutions de Pin1. On peut observer à travers l'intensité des signaux, une variation de la concentration de la solution de protéine, notamment au niveau des protons méthyles non échangeables (à droite) résonnant à des valeurs inférieures à 0 ppm.

Les valeurs des concentrations déterminées par le dosage Bradford ont été comparées avec les valeurs déterminées par RMN sur base de la comparaison des intensités des résonances des protons amides et des protons aliphatiques résonnant à moins de 0 ppm (Figure 17). On obtient une corrélation moyenne entre les valeurs obtenues par RMN avec, comme référence arbitraire, la courbe noire **Figure 16** pour laquelle la concentration de protéine a été fixée à 200  $\mu$ M, et les valeurs de concentration déterminées par Bradford.



**Figure 17 :** Corrélation entre les concentrations de Pin1 déterminées par le test de Bradford, par absorbance UV à 595 nm d'une solution de 1 mL de colorant additionnée de 2  $\mu$ L de solution de protéine, et les concentrations obtenues par RMN, par comparaison de l'intensité des signaux entre les différentes solutions de Pin1 au niveau des protons amide (bleu) et des protons aliphatiques (rose).

### **III – Acquisition des données par RMN et par fluorescence. Procédure de titration.**

#### **3.1 – RMN.**

##### ***RMN des peptides.***

Toutes les expériences ont été enregistrées sur un spectromètre Bruker DMX 600MHz, équipé d'une sonde cryogénique. La température nominale de l'échantillon est de 293K. Les échantillons contiennent 1 ou 2mM de peptide dans un tampon de 50mM Tris-d11 (Cambridge Isotope Laboratories, USA) et 100mM NaCl, à pH 6.3

Les expériences à une dimension ont été enregistrées avec une séquence Watergate pour la suppression de l'eau (Piotto et al., J. Biomol. NMR ).

Paramètres d'acquisition :

- Temps de répétition : 1s
- Durée de l'impulsion : 10 à 10.5 $\mu$ s (à -6dB)
- Largeur spectrale : 14ppm
- Nombre de points complexes : 2k
- Nombre d'accumulations : 32 ou 8

La Transformée de Fourier a été effectuée avec 2k points complexes, après multiplication du signal par une fonction sinusoïdale décalée de  $\pi/3$ .

L'attribution séquentielle du peptide est basée sur la stratégie classique décrite par K. Wüthrich, avec un spectre TOCSY pour identifier la nature des acides aminés à partir des connectivités avec le proton amide, et les spectres NOESY pour relier des acides aminés consécutifs.

Le spectre TOCSY a été enregistré comme un z-TOCSY avec une séquence de mélange isotrope, et une suppression de l'eau basée sur la séquence Watergate avec retour vers +z de l'aimantation de l'eau (Dhalluin et al.).

Paramètres d'acquisition :

- Temps de répétition : 1s

## Matériels et Méthodes

- Largeur spectrale : 14ppm (F<sub>1</sub>) x 14ppm (F<sub>2</sub>)
- Nombre de points complexes : 256 (F<sub>1</sub>) x 1024 (F<sub>2</sub>)
- Nombre d'accumulations : 8
- Séquence DIPSI : temps de mélange = 60ms,  $\gamma B_1 = 8.2\text{kHz}$

Le spectre a été transformé comme une matrice de 1k x 1k, après multiplication du signal par une fonction sinusoïdale décalée de  $\pi/2$  (en F<sub>1</sub>) et de  $\pi/3$  (en F<sub>2</sub>).

Le spectre NOESY a été enregistré avec une suppression de l'eau basée sur la séquence Watergate avec retour vers +z de l'aimantation de l'eau (Lippens et al.).

Paramètres d'acquisition :

- Temps de répétition : 1s
- Largeur spectrale : 14ppm (F<sub>1</sub>) x 14ppm (F<sub>2</sub>)
- Nombre de points complexes : 256 (F<sub>1</sub>) x 1024 (F<sub>2</sub>)
- Nombre d'accumulations : 8
- Temps de mélange = 400ms

Le spectre a été transformé comme une matrice de 1k x 1k, après multiplication du signal par une fonction sinusoïdale décalée de  $\pi/2$  (en F<sub>1</sub>) et de  $\pi/3$  (en F<sub>2</sub>).

### ***RMN de la protéine Pin1 ou du domaine WW.***

Les spectres homonucléaires de la protéine ont été enregistrés avec les mêmes paramètres que ceux des peptides.

Les spectres hétéronucléaires <sup>1</sup>H-<sup>15</sup>N HSQC ont été enregistrés sur deux spectromètres (600 et 300MHz), et avec plusieurs séquences :

- a) Séquence HSQC avec suppression par Watergate et retour de l'aimantation de l'eau vers l'axe +z
- b) Addition d'une séquence CPMG pour limiter la perte du signal due à la relaxation T<sub>2</sub> qui accompagne l'échange chimique
- c) Séquence de sensibilité augmentée, jouant sur un double transfert d'INEPT permettant de préserver les deux composantes x et y de l'aimantation.

Nous avons observé que pour la protéine Pin1 et son domaine WW, la séquence (b) avec CPMG permettait d'observer les fonctions amides correspondant à la boucle  $\beta 2$ - $\beta 3$  du

## Matériels et Méthodes

domaine WW ou la boucle du site catalytique, qui sont caractérisés par un échange chimique lent (Jacobs et al., 2003). Pour le domaine WW isolé, par contre, la séquence (c) dite « *sensitivity enhanced* » permet de gagner le facteur  $\sqrt{2}$  (voir texte principal Figure 8 p. 282).

Paramètres d'acquisition :

- Temps de répétition : 1s
- Largeur spectrale : 33ppm ( $F_1$ ,  $^{15}\text{N}$ ) x 14pm ( $F_2$ ,  $^1\text{H}$ )
- Nombre de points complexes : 32 ( $F_1$ ) x 1024 ( $F_2$ )
- Nombre d'accumulations : 8 ou 64 (dépendant de la concentration et du champs, voir Figure 10, 285)

Le spectre a été transformé comme une matrice de 512 x 1k, après multiplication du signal par une fonction sinusoïdale décalée de  $\pi/2$  (en  $F_1$ ) et de  $\pi/3$  (en  $F_2$ ).

Les spectres hétéronucléaires des peptides en abondance naturelle ont été enregistrés avec la même séquence (c), mais en utilisant 256 scans par incrément.

- Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. (1992) *J Biomol NMR*. **2**(6), 661-665.
- Dhalluin C, Wieruszski J, Lippens G. An Improved Homonuclear TOCSY Experiment with Minimal Water Saturation (1996) *J Magn Reson B*. **111**(2), 168-170.
- Lippens, G., Dhalluin, C., Wieruszski, J.-M. Use of a water flip-back pulse in the homonuclear NOESY experiment. (1995) *J Biomol NMR*. **5**(3), 327.

### ***Procédure de titration par RMN.***

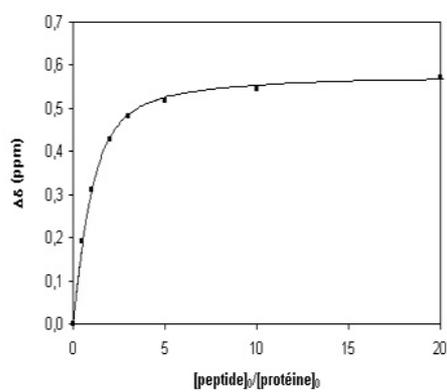
La protéine enrichie en  $^{15}\text{N}$  est solubilisée dans un tampon Tris-d11 50 mM pH 6.4, NaCl 100 mM, DTT 5 mM, EDTA 5 mM à une concentration de 200  $\mu\text{M}$  environ (voir détermination des concentrations de protéines en solution). Des quantités de peptide sont ajoutées à la solution de protéine dont le volume et la concentration restent constants au cours de la titration. Les quantités sont calculées pour obtenir les ratio [peptide] : [protéine] suivants : 0.5:1, 1:1, 2:1, 3:1, 5:1, 10:1 et 20:1. La solution de protéine est ajoutée successivement à des fractions de peptide lyophilisées obtenues à partir d'une solution concentrée de peptide à 10 mM dans l'eau ajustée à pH 6.4.

**Calcul des constantes de dissociations ( $K_D$ ).**

Le calcul des constantes de dissociation des complexes formés entre Pin1 (ou son domaine WW isolé) et les peptides phosphorylés de la protéine tau est effectué en reliant les données expérimentales à la courbe théorique obtenue à partir de l'Équation 1.

$$\Delta\delta(\text{ppm}) = 0.5\Delta\delta_{\text{max}} \left( 1 + X + \frac{K_D}{[\text{protéine}]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[\text{protéine}]_0} \right)^2 - 4X} \right) \text{ où } X = \frac{[\text{peptide}]_0}{[\text{protéine}]_0}$$

**Équation 1 :** Equation utilisée pour la détermination des constantes de dissociation des complexes Pin1-phospho-peptides.



On fait varier la valeur du  $K_D$  et on ajuste la valeur  $\Delta\delta_{\text{max}}$  pour que la courbe théorique (en trait plein, ci-dessus) correspondent au mieux aux points expérimentaux. Cette correspondance entre les valeurs théoriques et expérimentales est déterminée par la méthode des moindres carrés.

On obtient la valeur du déplacement chimique global  $\Delta\delta$  à partir des spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC en moyennant les variations du déplacement chimique en proton  $\Delta\delta(^1\text{H})$  et les variations du déplacement chimique en azote  $\Delta\delta(^{15}\text{N})$  selon l'Équation 2.

$$\Delta\delta(\text{ppm}) = [\Delta\delta^2(^1\text{H}) + 0.2 \Delta\delta^2(^{15}\text{N})]^{1/2}$$

**Équation 2 :** Variation pondérée du déplacement chimique d'un résidu en fonction de la variation en proton ( $^1\text{H}$ ) et de la variation en azote ( $^{15}\text{N}$ ).

**3.2 – Fluorescence intrinsèque et FRET.**

Les expériences de fluorescence intrinsèque et de FRET ont été réalisées sur un spectromètre PTI Fluorescence (Lawrenceville, New Jersey, USA) par excitation des tryptophanes de la

protéine à 295 nm. Les spectres d'émission couvrent la gamme 280-450 nm en fluorescence intrinsèque et 280-600 nm en FRET. La protéine est solubilisée à 20 mM dans un tampon Tris 50 mM pH 6.40, NaCl 100 mM, DTT 5 mM, EDTA 5 mM. La solution de protéine est ajoutée à des fractions de peptide lyophilisées obtenues à partir d'une solution concentrée de peptide à 10 mM dans l'eau ajustée à pH 6.4.

Pour les calculs de  $K_D$  à partir des expériences de fluorescence, on procède de la même façon que par RMN, en remplaçant le  $\Delta\delta$  et le  $\Delta\delta_{\max}$  en ppm par les valeurs  $\Delta I^{334}$  et  $\Delta I^{334}_{\max}$  représentant les variations de l'intensité d'émission de fluorescence à 334 nm.

## IV - Protocoles généraux des expériences en milieu hétérogène (Biacore<sup>®</sup>, ELISA).

### 4.1 – Biacore<sup>®</sup>.

#### **Préparation de la surface de la *sensorchip*. Couplage de la streptavidine/neutravidine.**

Une fonctionnalisation du dextran permet le couplage de la streptavidine (ou de la neutravidine) par réaction des fonctions amines primaires nucléophiles des résidus lysine exposés de la streptavidine sur les fonctions acides du dextran obtenues sous forme d'esters activés par réaction préalable avec le couple EDC/NHS en solution aqueuse à 200 mM et 50 mM respectivement. La solution est injectée à un débit de 10  $\mu\text{L}/\text{min}$  pour un volume effectif de 70  $\mu\text{L}$  (le volume total est de 100  $\mu\text{L}$  avec 30  $\mu\text{L}$  volume mort). Une solution de streptavidine à 500  $\mu\text{g}/\text{mL}$  dans un tampon d'acétate de sodium (pH 3.5) est amenée en flux continu à 10  $\mu\text{L}/\text{min}$  avec un volume total de 120  $\mu\text{L}$ , puis on passe 60  $\mu\text{L}$  d'une solution d'éthanolamine 1M à 10  $\mu\text{L}/\text{min}$  pour désactiver les fonctions acides n'ayant pas réagi. On lave deux fois avec 10  $\mu\text{L}$  d'une solution de SDS à 0.03%.

#### **Fixation des peptides biotinylés à la streptavidine/neutravidine.**

Les solutions de peptide et de protéine sont réalisées dans un tampon Tris 50 mM (pH 6.4), NaCl 100 mM, DTT 5 mM, EDTA 5 mM contenant 10% de BSA. Le système possédant 4 canaux utilisés en parallèle, les contrôles sont effectués simultanément à la mesure :

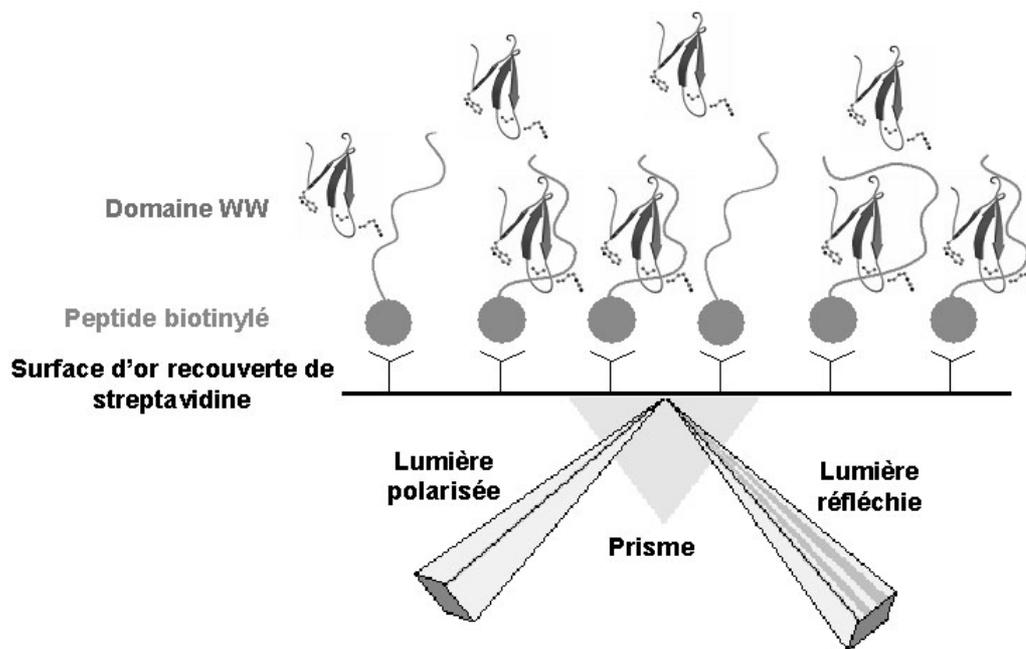
- Canal 1 : peptide T212 non phosphorylé biotinylé à 2  $\mu\text{g}/\text{mL}$  (contrôle négatif)
- Canal 2 : phospho-peptide pT212 biotinylé à 0.5  $\mu\text{g}/\text{mL}$
- Canal 3 : peptide T212 non phosphorylé biotinylé à 40 ng/mL (contrôle négatif)

- Canal 4 : phospho-peptide pT212 biotinylé 10 ng/mL.

Les solutions de peptides sont injectées à 10  $\mu\text{L}/\text{min}$  pour un volume de 10  $\mu\text{L}$ . La surface est ensuite lavée avec 10  $\mu\text{L}$  de SDS à 0.03% à 10  $\mu\text{L}/\text{min}$ . Les résultats ont montré des signaux équivalents en terme de charge de la surface pour les peptides phosphorylés et les contrôles non phosphorylés (d'après l'unité de mesure en RU ou unité de résonance) dans les deux cas (Canaux 1/2 et 3/4).

### Interactions avec le domaine WW de Pin1.

Quatre cycles d'injection (de 40  $\mu\text{L}$  par cycle) d'une solution de WW à différentes concentrations (10 nM, 100 nM, 1  $\mu\text{M}$  et 200  $\mu\text{M}$ ) sont effectués à 20  $\mu\text{L}/\text{min}$  sur les quatre canaux simultanément suivis d'une injection de 20  $\mu\text{L}$  d'une solution de SDS à 0.03%. Le signal mesuré en RU est proportionnel à la quantité de protéine injectée.



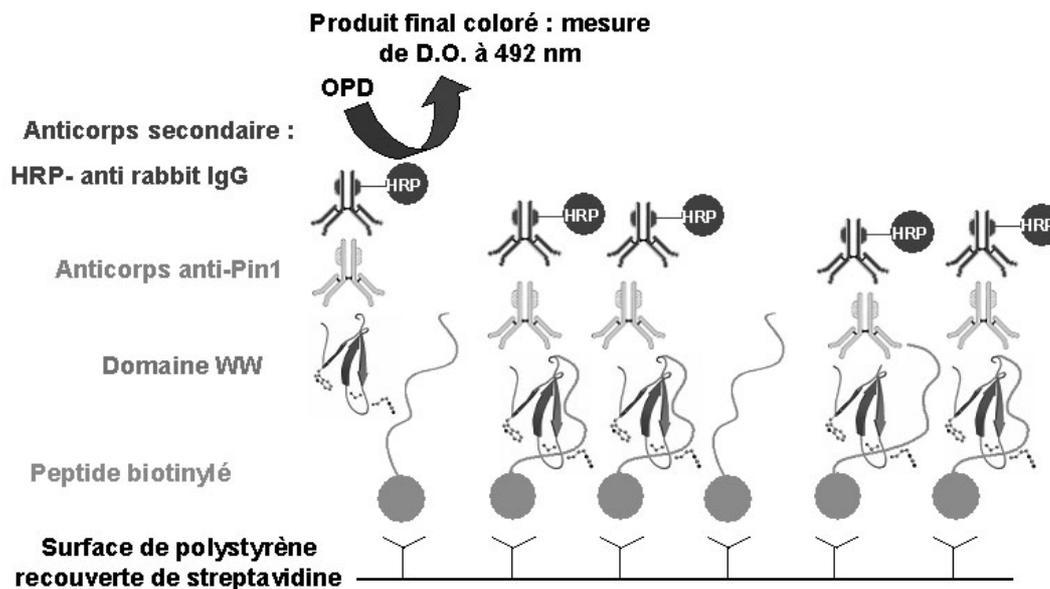
**Schéma 3 :** Représentation du système mis en œuvre en Biacore<sup>®</sup> pour la détection d'interactions entre des peptides phosphorylés et le domaine WW de Pin1.

### 4.2 – ELISA.

Les tests ELISA ont été effectués sur plaque (Nunc<sup>™</sup> MaxiSorp 96 puits, Nalge Nunc Int.) par fixation d'un des deux partenaires, ici le peptide, et par mesure qualitative de la quantité de protéine Pin1 « piégée » dans les puits par interaction avec les peptides.

Les différentes étapes du protocole sont les suivantes :

1. *coating* de la surface des puits par la Neutravidine® (ImmunoPure® Pierce)
2. fixation des peptides biotinylés phosphorylé et témoin non phosphorylé
3. fixation de la protéine Pin1
4. fixation de l'anticorps anti-His<sub>5</sub>
5. fixation de l'anticorps secondaire couplé à l'HRP (Horseradisch Peroxidase)
6. révélation par réaction avec l'OPD (orthophénylène diamine)



**Schéma 4** : Représentation du système mis en œuvre en ELISA pour la détection des interactions entre Pin1 et le peptide phosphorylé centré autour de l'épitope phospho-Thr231.

1. **Coating de la Neutravidine®.** La neutravidine est solubilisée à 10 µg/mL dans un tampon à KH<sub>2</sub>PO<sub>4</sub> 20 mM / K<sub>2</sub>HPO<sub>4</sub> 10 mM pH=7.2 ; NaCl 140 mM ; EDTA 1 mM, 50 µL sont introduits dans chaque puits et la plaque est incubée à température ambiante pendant 3 heures. On retire la solution de neutravidine puis on incube chaque puits avec 200 µL du même tampon contenant 2% de BSA (Calbiochem, BSA fractionV, RIA and ELISA Grade) pour limiter les interactions non spécifiques ultérieures. La plaque est incubée 30 min à température ambiante.
2. **Fixation du peptide biotinylé.** Les peptides sont solubilisés à 1 µM dans le tampon précédent sans BSA et dilués à la concentration initiale désirée dans le même tampon contenant 2% de BSA. Une gamme de dilution en peptide est réalisée si nécessaire par

## Matériels et Méthodes

dilution sériée sur plaque. 50  $\mu$ L de la solution de peptide à la concentration souhaitée sont introduits dans les puits et la plaque est incubée à température ambiante pendant 1 heure.

3. **Fixation de Pin1.** La protéine Pin1 solubilisée dans un tampon Tris 25 mM pH=6.80, NaCl 100 mM avec 2% de BSA à la concentration initiale désirée à partir d'une solution stock concentrée. Une gamme de dilution est réalisée si nécessaire par dilution sériée sur plaque. 50  $\mu$ L en volume final sont introduits dans chaque puits et la plaque est incubée à température ambiante pendant 1 heure.
4. **Fixation de l'anticorps anti-His<sub>5</sub>.** L'anticorps est solubilisé dans un tampon Tris 25 mM pH=6.8 ; 2% BSA à une dilution de 1/250<sup>e</sup>. 50  $\mu$ L sont introduits dans les puits et la plaque est incubée à température ambiante pendant 1 heure.
5. **Fixation de l'anticorps secondaire HRP.** L'anticorps (IgG de souris) est solubilisé dans un tampon Tris 25 mM pH=6.8 ; 2% BSA à une dilution de 1/500<sup>e</sup>. 50  $\mu$ L sont introduits dans les puits et la plaque est incubée à température ambiante pendant 1 heure.
6. **Révélation.** L'OPD (Sigma) est utilisé comme substrat pour la réaction enzymatique avec l'HRP lié à l'anticorps secondaire. L'OPD est solubilisé à 0.8 mg/mL dans un tampon phosphate/citrate à pH=5.0 (Sigma) contenant du perborate de sodium comme agent oxydant. 50  $\mu$ L sont introduits dans les puits. La lecture est effectuée à 450 nm ou à 492 nm après arrêt de la réaction avec H<sub>2</sub>SO<sub>4</sub> 0,1 M.

Des lavages sont effectués entre chaque étape du protocole avec un tampon TBS contenant 0.1% de Tween-20 (trois lavages) suivis par deux lavages au TBS.

## **Structures de Pin1 et des complexes dans la PDB (*Protein Data Bank*).**

### **1. Structure cristallographique de Pin1 en complexe avec un dipeptide Ala-Pro.**

Auteurs : J. P. Noel, R. Ranganathan, T. Hunter

Titre : Pin1 Peptidyl-Prolyl Cis-Trans Isomerase From Homo Sapiens

Référence : Ranganathan, R., Lu, K. P., Hunter, T., Noel, J. P.: Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 89 pp. 875 (1997)

Code PDB : 1PIN

### **2. Structure cristallographique de Pin1 en complexe avec un peptide diphosphorylé du domaine C-terminal de l'ARN polymérase II (YpSPTpSPS).**

Auteurs : M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter, J. P. Noel

Titre : Structural Basis For The Phosphoserine-Proline Recognition By Group IV WW Domains

Référence : Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., Noel, J. P.: Structural Basis for Phosphoserine-Proline Recognition by Group Iv Ww Domains *Nat.Struct.Biol.* 7 pp. 639 (2000)

Code PDB : 1F8A

### **3. Structure RMN en solution du domaine WW de Pin1.**

Auteurs : R. Wintjens, J.-M. Wieruszski, H. Drobecq, G. Lippens, I. Landrieu

Titre : Solution Structure Of Pin1 WW Domain

Références : Wintjens, R., Wieruszski, J.-M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., Landrieu, I.: 1H NMR Study on the Binding of Pin1 Trp-Trp Domain with Phosphothreonine Peptides *J.Biol.Chem.* 276 pp. 25150 (2001)

Code PDB : 1I6C

### **4. Structure RMN en solution du domaine WW de Pin1 en complexe avec un peptide de tau.**

Auteurs : R. Wintjens, J.-M. Wieruszski, H. Drobecq, G. Lippens, I. Landrieu

Titre : Solution Structure Of Pin1 WW Domain Complexed With Human  $\tau$ Phosphothreonine Peptide

Références : Wintjens, R., Wieruszski, J.-M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., Landrieu, I.: 1H NMR Study on the Binding of Pin1 Trp-Trp Domain with Phosphothreonine Peptides *J.Biol.Chem.* 276 pp. 25150 (2001)

Code PDB : 1I8H

### **5. Structure RMN en solution du domaine WW de Pin1 en complexe avec un peptide de Cdc25.**

Auteurs : R. Wintjens, J.-M. Wieruszeski, H. Drobecq, G. Lippens, I. Landrieu

Titre : Solution Structure Of Pin1 WW Domain Complexed With Cdc25 Phosphothreonine Peptide

Références : Wintjens, R., Wieruszeski, J.-M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., Landrieu, I.: 1H NMR Study on the Binding of Pin1 Trp-Trp Domain with Phosphothreonine Peptides *J.Biol.Chem.* 276 pp. 25150 (2001)

Code PDB : 1I8G

### **6. Structure RMN en solution de Pin1.**

Auteurs : E. Bayer, S. Goettsch, J. W. Mueller, B. Griewel, E. Guiberman, L. Mayr, P. Bayer

Titre : Solution Structure of human Pin1

Références : Bayer, E., Goettsch, S., Mueller, J. W., Griewel, B., Guiberman, E., Mayr, L., Bayer P.: Structural Analysis of the Mitotic Regulator Hpin1 in Solution : Insights into Domain Architecture and Substrate Binding. *J.Biol.Chem.* 278 pp. 26183 (2003)

Code PDB : 1NMV

## **Données d'attribution par RMN de Pin1 et du domaine WW dans la BMRB (*BioMagResBank*).**

### **1. Domaine WW de Pin1**

Références : Wintjens, R., Wieruszeski, J.-M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., Landrieu, I.: 1H NMR Study on the Binding of Pin1 Trp-Trp Domain with Phosphothreonine Peptides *J.Biol.Chem.* 276 pp. 25150 (2001)

Code BRMB : 4882

### **2. Pin1.**

Références : Jacobs, D.M., Saxena, K., Grimme, S., Vogtherr, M., Pescatore, B., Langer, T., Elshorst, B., Fiebig, K.M. : Backbone 1H, 13C, and 15N Chemical Shift Assignments for the peptidyl-prolyl cis-trans isomerase Pin1 *J. Biomol. NMR* 23 pp 163 (2002)

Code BRMB : 5305

## **Deuxième Partie – Synthèse organique et caractérisation des composés. Criblage des composés par RMN et fluorescence.**

### **I - Synthèse et caractérisation des ligands potentiels du domaine WW de Pin1.**

Le projet concernant la recherche d'inhibiteurs du domaine WW de Pin1, dont la partie expérimentale est présentée ici, a été financé par ISOA (Institute for the Study Of Aging, USA).

Les acides aminés naturels protégés par un groupement Fmoc, les agents de couplage (HOBT et HBTU), la résine proviennent de chez Novabiochem. La phase solide utilisée consiste en une résine de polystyrène pontée par 1% de divinylbenzène greffée par un fragment de type Rink amide MBHA dont la charge est de 0.74 mmol/g.

La pureté des composés a été vérifiée par un système LC-MS. Les données sont collectées par un système Varian 1200MS composé d'une colonne TSKgel OSH supelco. L'eau (A) et l'acétonitrile (B) sont utilisés comme éluants et contiennent respectivement 0.1% et 0.12% d'acide formique. Après l'injection de l'échantillon, la colonne est équilibrée dans la solution A pendant 2 min et on élue les constituants par un gradient linéaire allant de 0 à 100%B en 7.30 min avec un flux de 1 mL/min. Les composés élués sont détectés simultanément à 215 et 254 nm. La caractérisation des composés solubilisés dans le DMSO-d<sub>6</sub> est réalisée à 300K avec l'acquisition de spectres protons, COSY et <sup>13</sup>C-<sup>1</sup>H HSQC sur un spectromètre Bruker à 300MHz équipé d'une sonde BBI.

#### **1.1 - Synthèse des composés**

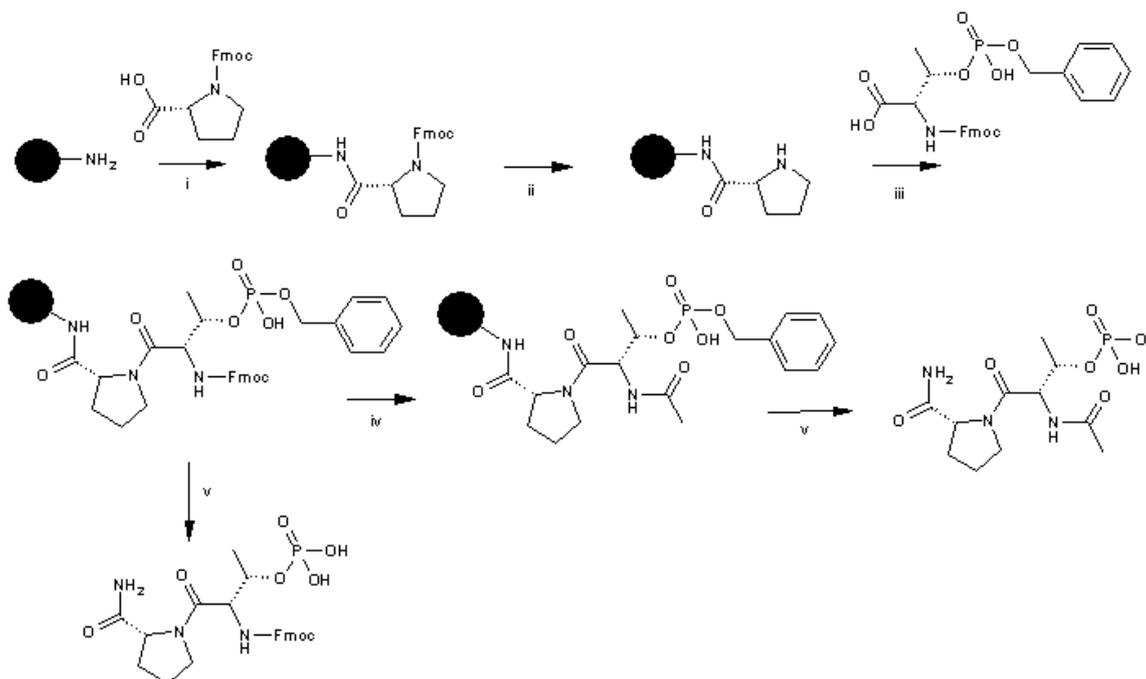
Les composés comportant une phosphothréonine ou une phosphosérine ont été synthétisés selon la procédure n°1, mettant en œuvre de faibles quantités (3 équiv.) pour les synthons commerciaux N- $\alpha$ -Fmoc-(L)-Thr/Ser(PO(OBn)OH)-OH (Novabiochem) pour des raisons de coûts des réactifs : 620€/5g pour la phosphothréonine et 510€ pour la phosphosérine.

Les composés comportant un analogue de phosphosérine comportant une tête polaire de type acide carboxylique (acide glutamique et dérivés) ont été synthétisés selon la procédure n°2.

Les composés comportant d'autres têtes polaires (phosphonates,  $\alpha$ -diacide, squarate, tétrazole, trifluorosulfonamide, 3,4-difluoroaniline, 3-nitroaniline) ont été synthétisés par le

Dr Jean-Frédéric Duckert en stage post-doctoral dans le laboratoire qui s'est investi dans le projet avec la synthèse d'une série de composés à têtes polaires conservant la partie proline. Seulement deux des molécules synthétisées dans cette série ont montré des affinités mesurables. Il s'agit de molécules comportant une tête polaire de type squarate ancrée sur la fonction amine de l'ornithine et la lysine, avec conservation de la protection Fmoc de l'amine. Pour ces deux molécules, la synthèse et la caractérisation des produits finis ont été décrites dans la procédure n°3.

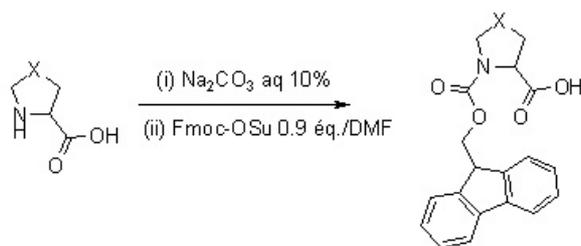
**Procédure n°1 : Procédure générale pour les composés (1)-(10), (19)-(35), (38)-(45).**



**Schéma 5 :** Schéma général de synthèse de ligands potentiels du domaine WW de Pin1. <sup>a</sup> Réactifs et conditions: (i) HOBt, HBTU, 3 éq. DIPEA, DMF, temp. amb., 1 h; (ii) pipéridine 20%, DMF, temp. amb., 25 min; (iii) HOBt, HBTU, 3 éq. DIPEA, DMF, temp. amb., 1 h; (iv) (a) pipéridine 20%, DMF, temp. amb., 25 min; (b)  $(\text{CH}_3\text{CO})_2\text{O}$ , 2 éq. DIPEA, DMF, temp. amb., 15 min; (v) TFA 95%, TIS 2.5%, eau 2.5%, temp. amb., 5 h.

**Protection de la fonction amine des mimes de proline par un groupement Fmoc.**

Les dérivés de type (D)-proline, acide 2-(L)-azétidine carboxylique, acide 2-(L)-pipécolique carboxylique (ou (L)-Homoproline), 4-*trans*-hydroxy-(L)-proline, 4-thia-(L)-proline, 4-*trans*-fluoro-(L)-proline, 4-*cis*-fluoro-(L)-proline, 3,4-dehydro-(L)-proline sont protégés préalablement (sur une échelle de 150 mg/dérivé) au niveau de leur fonction amine secondaire par un groupement Fmoc avant d'être engagés dans la synthèse en phase solide.



On pèse 1 éq. de proline ou dérivés (1.33 mmol dans 3.2 mL) et on solubilise le composé dans une solution aqueuse de  $\text{Na}_2\text{CO}_3$  10% que l'on place sous agitation et que l'on refroidit dans un bain de glace. On pèse 0.9 éq. de N-(9-fluorénylméthoxycarboxyloxy)-succinimide (Fmoc-OSu) que l'on dissout dans le DMF (1.19 mmol dans 2.72 mL). On additionne cette solution goutte à goutte à la solution préparée précédemment. Un précipité blanc se forme instantanément. Après l'addition du Fmoc-OSu, on retire le bain de glace et on laisse la réaction se poursuivre sous agitation pendant 2H à température ambiante. Le milieu réactionnel est ensuite dilué dans l'eau distillée (100 mL) jusqu'à obtention d'une solution claire. La phase aqueuse est extraite avec 10 mL d'éther et 2 fois 10 mL d'acétate d'éthyle. La phase aqueuse est refroidie dans un bain de glace et acidifiée jusqu'à pH 2.0 par ajout d'HCl concentré. On voit se former un précipité blanc ou une huile. L'ensemble est extrait avec 6 fois 10 mL d'acétate d'éthyle. Les phases organiques sont rassemblées et lavées avec 3 fois 10 mL d'une solution saturée de NaCl et 2 fois avec 10 mL d'eau distillée, puis séchées sur  $\text{MgSO}_4$  et évaporées. On obtient un produit visqueux que l'on peut traiter avec de l'éther de pétrole froid pour obtenir un produit sous forme solide. Les produits protégés sont obtenus avec un rendement supérieur à 90%, caractérisés par spectrométrie de masse et par RMN ( $^1\text{H}$ ,  $^{13}\text{C}$ ).

#### Couplage des acides aminés protégés.

On pèse 3 équiv. d'acide aminé N- $\alpha$ -Fmoc, 3 équiv. HOBt et 3 équiv. HBTU que l'on dissout dans un minimum de DMF. On ajoute 9 équiv. de DIPEA. La solution se colore légèrement en jaune. Cette solution est ajoutée à la résine et l'ensemble est incubé à température ambiante pendant 1H. La solution est ensuite retirée par filtration sous vide et la résine est lavée 3 ou 4 fois au DMF (5-6 mL par lavage).

#### Coupage du groupement protecteur Fmoc.

Les fonctions amines protégées par un groupement Fmoc sont déprotégées par incubation de la résine avec une solution de pipéridine à 20% dans le DMF (3 mL pour 300 mg de résine)

pendant 25 min à température ambiante. La solution est ensuite retirée par filtration sous vide et la résine est lavée 3 ou 4 fois au DMF.

### **Acétylation de la fonction amine.**

On dissout 5 équiv. d'anhydride acétique dans 3 mL de DMF et on ajoute 10 équiv. de DIPEA. La solution est ensuite incubée avec la résine pendant 15 min à température ambiante. Puis la résine est filtrée et lavée 3 ou 4 fois au DMF.

### **Couplage des acides hydrophobes.**

On pèse 5 équiv. d'acide, 5 équiv. HOBt et 5 équiv. HBTU et on les dissout dans un minimum de DMF, puis on ajoute 10 équiv. de DIPEA. La réaction de couplage est réalisée à température ambiante pendant 1H. La solution est ensuite retirée par filtration sous vide et la résine est lavée 3 ou 4 fois au DMF.

### **Tests à la ninhydrine ou au chloranil.**

Chaque réaction de couplage et de déprotection fait intervenir un test colorimétrique de détection des amines primaires (test à la ninhydrine) ou des amines secondaires dans le cas de la proline et des mimes de proline (test au chloranil). On prélève quelques grains de résine que l'on incube avec les différentes solutions mises en œuvre pour les différents tests. Le test à la ninhydrine est positif après coupure du groupement Fmoc du fragment Rink amide de la résine et de la fonction amine de la phosphothréonine. Le test à la ninhydrine est négatif après couplage de la proline ou d'un mime de proline sur la résine, après couplage d'un acide hydrophobe ou après acétylation d'une fonction amine. La coupure du groupement Fmoc de la proline ou d'un mime de proline aboutit à un test au chloranil positif tandis que le couplage de la phosphothréonine sur la proline ou un mime de proline aboutit à un test au chloranil négatif.

Test à la ninhydrine (ou test de Kaiser) : on prépare les trois solutions suivantes

1. on dissout 1g de ninhydrine dans 20 mL d'éthanol
2. on dissout 80g de phénol dans 20 mL d'éthanol
3. on mélange 2 mL d'une solution aqueuse de cyanure de potassium à 0.001 M et 98 mL de pyridine

## Matériels et Méthodes

On ajoute deux gouttes de chaque solution à la résine prélevée, on chauffe à 120°C pendant 5 min. Le test est positif (présence d'amines primaires) si la solution et la résine sont de couleur bleu foncé (presque noire). Si la résine et la solution sont incolores, le test est négatif.

Test au chloranil : on prépare une solution de p-chloranil à 2% dans le DMF et une solution d'acétaldéhyde à 2% dans le DMF. On ajoute une goutte de chaque solution à la résine prélevée. Le test est positif (présence d'amines secondaires) si la résine est de couleur bleu foncé. Le test est négatif si la résine reste incolore.

### **Coupage au TFA des composés de la phase solide et protection du groupement protecteur benzyle de la phosphothréonine. Isolation des composés.**

La résine est lavée abondamment 3 ou 4 fois au DMF, puis de la même façon au dichlorométhane et enfin on procède à un lavage avec une solution à 10% TFA dans le dichlorométhane (2 lavages d'un volume correspondant à celui de la résine). La résine est ensuite incubée avec une solution de TFA contenant 2.5% de triisopropylsilane (v/v) et 2.5% d'eau (v/v) pendant 5H à température ambiante sous agitation. La solution est ensuite filtrée sous vide et la résine est lavée deux fois avec quelques mL de TFA pur. Les filtrats sont rassemblés et collectés dans un mélange éther : pentane (v/v) froid (-20°C) dont le volume est d'au moins 10 fois celui de la solution de TFA. Un précipité blanc apparaît et la précipitation est ensuite achevée pendant la nuit à -20°C. Le précipité est ensuite isolé par filtration sur fritté n°4, lavé à l'éther froid et solubilisé dans l'eau. La solution aqueuse obtenue est congelée et lyophilisée.

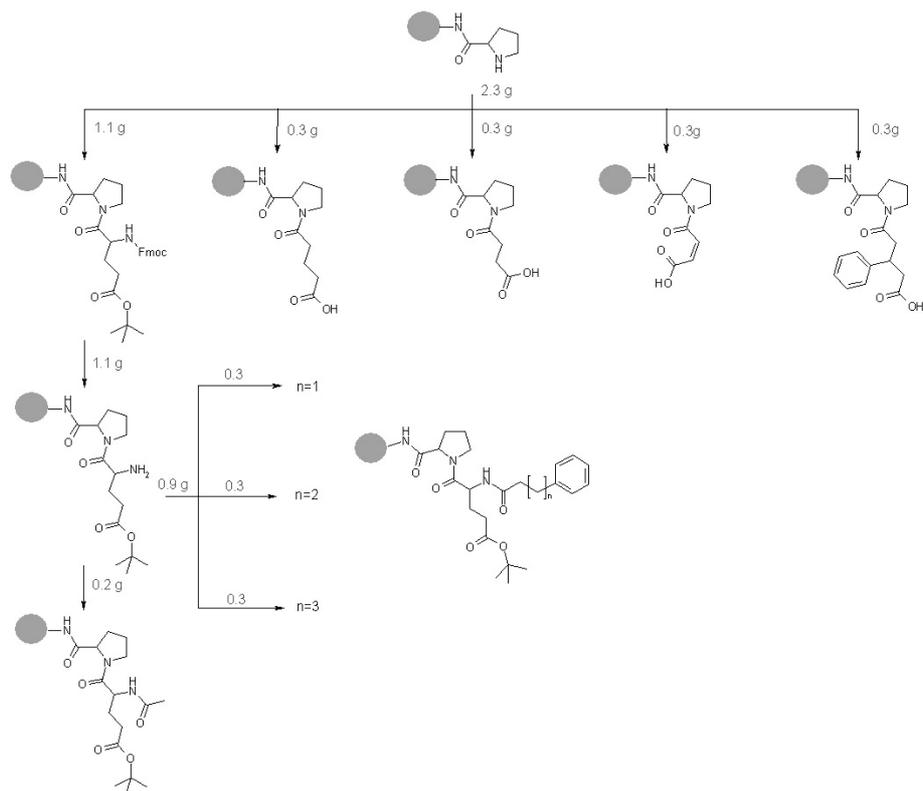
### ***Procédure n°2 : Procédure générale pour la synthèse en parallèle du dipeptide Glu-Pro et des dérivés dans les composés (11) à (18).***

Une série de molécules basées sur le couplage de l'acide glutamique ou d'analogues structuraux de l'acide glutamique à la proline a été obtenue par synthèse en phase solide. La première étape de couplage de la proline permet l'obtention d'un lot de résine qui, après fractionnement, peut subir une première étape de diversification. Une deuxième étape de diversification est réalisée sur le dipeptide Glu-Pro par introduction d'un groupement hydrophobe sur la fonction amine terminale.

La première étape de couplage de la proline sur 2.3 g de résine de type Rink amide MBHA (Novabiochem), comme décrit dans la procédure n°1, est suivie d'une étape diversifiante

## Matériels et Méthodes

après séparation du lot de résine. Les couplages des anhydrides glutarique, succinique, maléique et 3-phénylglutarique sont réalisés sur 0.3 g de résine chacun. Un couplage de l'acide glutamique est réalisé sur 1.1 g de résine laquelle subit une seconde étape de diversification. On réalise une étape d'acétylation de la fonction amine sur 0.2 g de résine. Les acides 3-phénylpropionique, 4-phénylbutyrique et 5-phénylvalérique sont couplés sur 0.3 g de phase solide chacun.



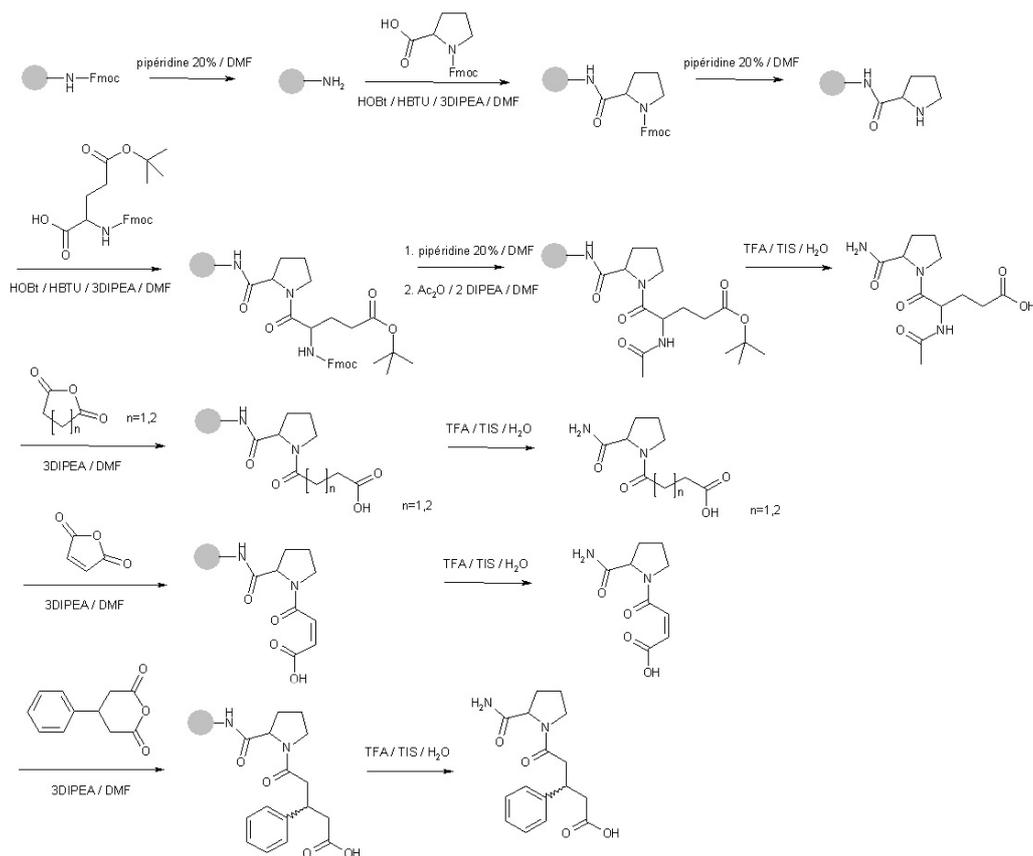
**Schéma 6 :** Schéma de synthèse parallèle en phase solide du dipeptide Glu-Pro et de ses dérivés.

### Couplage de l'acide glutamique et de ses dérivés.

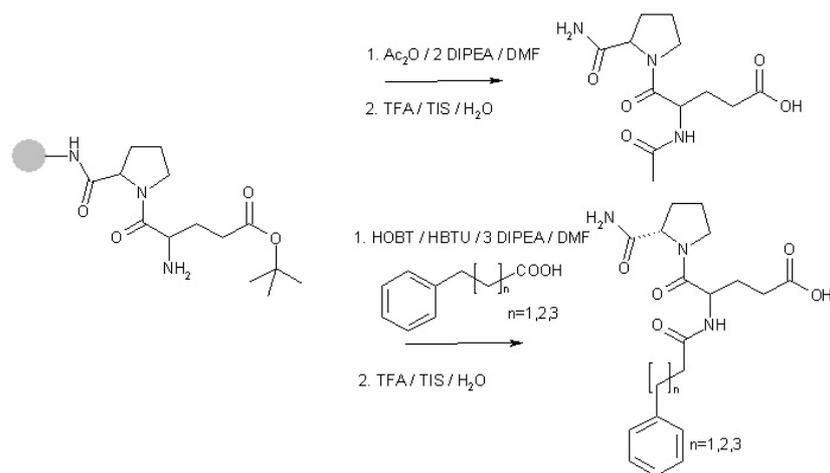
Le couplage de l'acide glutamique est réalisé avec 5 équiv. d'acide, 5 équiv. HOBt et 5 équiv. HBTU solubilisés dans le minimum de DMF. On ajoute 15 équiv. DIPEA et on incube la solution obtenue avec la résine pendant 1H à température ambiante.

Le couplage des anhydrides glutarique, succinique, maléique et 3-phénylglutarique est réalisé avec 5 équiv. d'anhydride solubilisé dans un minimum de DMF et 15 équiv. de DIPEA. La solution est incubée avec la résine pendant 1H à température ambiante.

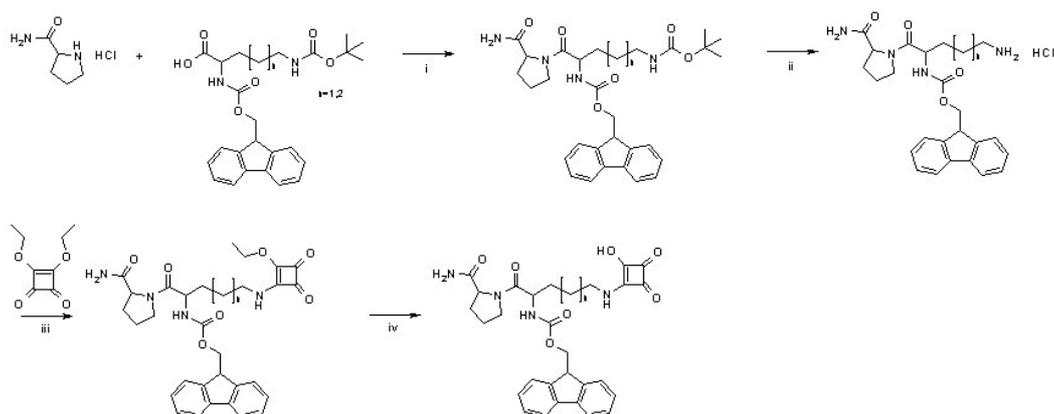
## Matériels et Méthodes



Les précurseurs acides 3-phénylpropionique, 4-phénylbutyrique et 5-phénylvalérique ont été couplés sur la fonction amine de l'acide glutamique du dipeptide Glu-Pro ancré sur résine, comme décrit dans la Procédure n°1.



La coupure des composés de la résine ainsi que leur isolement sont réalisés dans les mêmes conditions que décrites dans la procédure n°1 mais les temps d'incubation dans le TFA sont plus courts (environ 2H).

**Procédure n°3 : Procédure générale pour les composés (36) et (37).**

**Schéma 7 :** Schéma général pour la synthèse des ligands (36) et (37) du domaine WW de Pin1. <sup>a</sup> Réactifs et conditions : (i) DCC/DCM, 0°C, 2 h; (ii) HCl/dioxane, 0°C puis température ambiante, 3 h; (iii) (a) NEt<sub>3</sub>/EtOH (b) diéthylsquarate/EtOH, température ambiante, 12 h (iv) HCl/DCM/24H température ambiante.

### Couplage de l'ornithine et de la lysine protégés par un groupement Fmoc sur la prolinamide.

3 éq. des dérivés N- $\alpha$ -Fmoc-ornithine/lysine(N- $\epsilon$ -Boc) sont solubilisés dans le DCM (15.8 mmol dans 8 mL) et ajoutés à une solution contenant 1.5 éq. DCC dans le DCM à 0°C (7.9 mmol dans 2 mL). Le précipité est filtré et le filtrat est ajouté à une solution de 1 éq. de prolinamide dans le DCM (5.25 mmol dans 3 mL). Le mélange est placé sous agitation pendant 2H à 0°C. Le mélange brut est ensuite évaporé et re-solubilisé dans le DCM. Le précipité est retiré par filtration et la phase organique est lavée deux fois avec une solution aqueuse saturée de NaHCO<sub>3</sub> et ensuite séchée sur MgSO<sub>4</sub> puis évaporée. Le mélange est purifié sur silice et élué avec une solution DCM/EtOH 95/5.

### Coupage du groupement protecteur Boc.

1 éq. des composés préparés précédemment sont solubilisés dans le DCM (1.37 mmol dans 14 mL). La solution est refroidie dans un bain de glace. Une solution de HCl 4N dans le dioxane est préparée et ajoutée au goutte à goutte à la solution précédente à 0°C. Le mélange est placé sous agitation à température ambiante pendant 3H et le mélange brut est ensuite évaporé. Le produit déprotégé est obtenu quantitativement.

### Couplage du diéthylsquarate sur les fonctions amines des chaînes latérales des dérivés de l'ornithine et de la lysine.

Une solution de 1.2 éq. de diéthylsquarate est préparée dans l'éthanol (1.64 mmol in 5 mL). On pèse 1 éq. du dérivés N- $\alpha$ -Fmoc-ornithine/lysine couplé à la prolinamide, on le solubilise

## Matériels et Méthodes

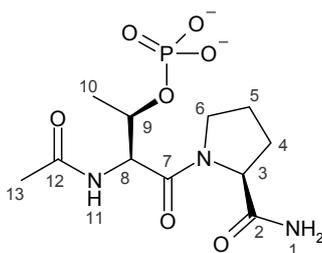
dans l'éthanol (1.37 mmol dans 8 mL) et on ajoute 1.5 éq. de triéthylamine. Cette solution est ajoutée au goutte à goutte à la solution de diéthylsquarate et l'ensemble est mis sous agitation à température ambiante pendant 12H. Le mélange brut est évaporé et purifié sur colonne de silice. Les composés sont élués avec une solution DCM/MeOH 9/1.

### Coupage des groupements éthyles du derives du diéthylsquarate.

Le dérivé diéthylsquarate est solubilisé dans le THF (1.084 mmol dans 6 mL). Cette solution est diluée dans une solution HCl 4N dans le THF (12 mmol dans 3 mL) et l'ensemble est placé sous agitation à température ambiante pendant 20H. Le mélange brut est évaporé, le solide obtenu est solubilisé dans le DCM et extrait 4 fois avec une solution aqueuse de  $\text{NH}_4\text{HCO}_3$  0.3N. Le produit est obtenu de manière quantitative après lyophilisation.

## 1.2 - Caractérisation des composés.

### Composé (1).



Rendement 25%

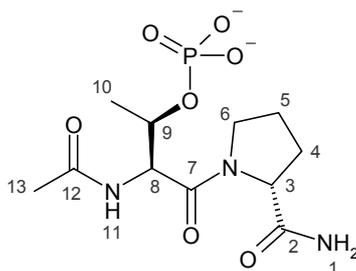
LC-MS  $t_R$  1.38 min 98%  $m/z=336$  ( $M-H^+$ )

$^1\text{RMN } ^1\text{H}$  (300MHz,  $\text{DMSO-d}_6$ ) –  $\text{RMN } ^{13}\text{C}$  (75MHz):  $\delta$  (ppm) = 8.03 (d, 1H, 7.60 Hz) [11]; 7.36 (s, 1H br) [1]; 7.05 (s, 1H br) [1]; 4.62 (t, 1H, 7.31 Hz) – 55.5 [8]; 4.45 (dd, 1H, 6.43 Hz, 2.92 Hz) – 72.4 [9]; 4.21 (dd, 1H, 8.18 Hz, 5.26 Hz) – 60.0 [3]; 3.77 (m, 1H) – 47.6 [6]; 3.63 (m, 1H) – 47.6 [6]; 2.04 (m, 1H) – 29.6 [4]; 1.88 (m, 1H) – 24.8 [5]; 1.85 (s, 3H) – 22.6 [13]; 1.79 (m, 1H) – 29.6 [4]; 1.26 (d, 3H, 5.85 Hz) – 18.1 [10].

---

<sup>1</sup> L'attribution des protons et des carbones associés (valeur entre crochets) est déterminée à partir des valeurs de déplacements chimiques proton avec, entre parenthèses, la multiplicité, l'intégration du signal et les constantes de couplage  $^3J[^1\text{H}-^1\text{H}]$ , et les valeurs de déplacements chimiques du carbone associé obtenu par un spectre  $^1\text{H}-^{13}\text{C}$  HSQC. br, broad (résonance élargie).

**Composé (2).**

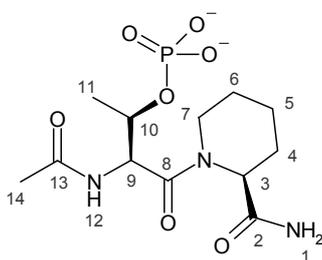


Rendement 30%

LC-MS  $t_R$  1.52 min 98%  $m/z=336$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.16 (d, 1H, 7.31) [11] ; 7.07 (s, 1H) [1] ; 6.94 (s, 1H) [1] ; 4.59 (t, 1H, 6.43 Hz) – 55.9 [8] ; 4.41 (m, 1H) – 73.2 [9] ; 4.19 (dd, 1H, 8.18 Hz, 2.34 Hz) – 60.8 [3] ; 3.67 (m, 2H) – 47.9 [6] ; 1.99 (m, 1H) – 30.2 [4] ; 1.87 (s, 3H) – 23.2 [13] ; 1.77 (m, 3H) – 30.2 [4] ; 1.77 (m, 3H) – 24.9 [5] ; 1.21 (d, 3H 6.43 Hz) – 18.5 [10].

**Composé (3).**

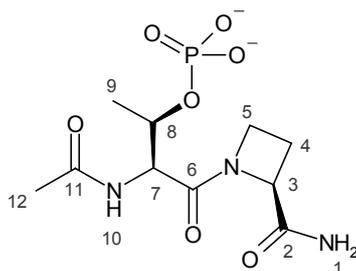


Rendement 15%

LC-MS  $t_R$  3.62 min 96%  $m/z=350$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 9.70-8.50 (s, 1H br) [12] ; 8.13-7.17 (s,1H) [1] ; 7.61-7.00 (s,1H) [1] ; 5.01-4.96 (t, 1H, 2.92 Hz) – 52.7-56.6 [3] ; 4.70-4.45 (t, 1H, 6.72 Hz) – 55.0-55.4 [9] ; 4.40-4.22 (m, 1H) – 70.7 [10] ; 4.43-2.37 (m, 1H) – 39.9 [8] ; 4.05-3.12 (td, 1H, 12.86 Hz, 2.92 Hz) – 43.9 [8] ; 2.18-1.33 (m, 1H) – 26.6 [4] ; 2.21-1.41 (m,1H) – 26.3 [4] ; 1.58-1.40 (m, 1H) – 25.9 [6] ; 1.54-1.19 (m, 1H) – 25.3 [6] ; 1.86-1.80 (s, 3H) – 23.0-22.6 [14] ; 1.55-1.32 (m, 2H) – 21.0 [5] ; 1.13 (d, 3H, 6.12 Hz) – 19.2 [11].

**Composé (4).**



Rendement 23%

LC-MS

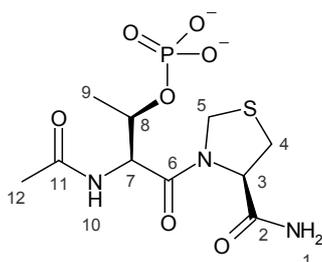
$t_R$  0.87 min 95%

$m/z=322$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.05 (d, 1H, 7.31 Hz) [1] ; 7.43 (s, 1H) [1] ; 7.21 (s, 1H) [10] ; 4.51 (m, 1H) – 60.8 [3] ; 4.42 (m, 1H) – 71.7 [8] ; 4.39 (m, 1H) – 53.8 [7] ; 4.19 (m, 2H) – 49.0 [5] ; 2.46 (m, 1H) - 20.2 [4] ; 2.08 (m, 1H) – 20.2 [4] ; 1.89 (s, 3H) – 22.7 [12] ; 1.24 (d, 3H, 5.85 Hz) – 19.4 [9].

**Composé (5).**

Pour ce dérivés, le couplage de la phosphothréonine est réalisé par activation de l'acide au chlorure d'oxalyle (3 équiv.) dans le DCM avec ajout de 50 mL de DMF. La fonction amine est préalablement déprotonnée par incubation de la résine avec 9 équiv. de DBU (1,8-diazabicyclo[5,4,0]undec-7-ène) dans le DCM. La solution de DBU est retirée par filtration et la résine est ensuite incubée avec la solution d'acide activé. La coupure du composé du support solide est réalisée avec une solution contenant 94% TFA, 2,5% d'éthanedithiol, 1% de triisopropylsilane et 2,5% d'eau.



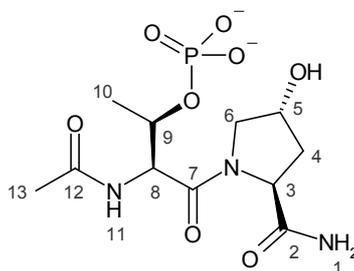
Rendement 13%

LC-MS  $t_R$  1.45 min 90%  $m/z= 354 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 8.32 (s, 1H) [10] ; 7.48 (s, 1H) [1] ; 7.19 (s, 1H) [1] ; 5.12 (d, 1H, 9.06 Hz) – 49.9 [5] ; 4.70 (m, 1H) – 63.1 [3] ; 4.68 (m, 1H) – 56.5 [7] ; 4.60 (d, 1H, 9.06 Hz) – 49.9 [5] ; 4.43 (m, 1H) - 72.34 [8] ; 4.13 (d, 1H, 9.06 Hz) – 54.4 [5] ; 4.05 (d, 1H, 9.06 Hz) – 54.4 [5] ; 3.25 (m, 1H) – 33.6 [3] ; 3.04 (m, 1H) – 33.6 [4] ; 2.97 (m, 1H) – 36.4 [4] ; 2.88 (m, 1H) – 36.4 [4] ; 1.87 (s, 3H) – 22.8 [12] ; 1.26 (d, 3H, 5.85 Hz) – 18.3 [9].

**Composé (6).**

L'acétylation de la fonction amine de la phosphothréonine est réalisée par incubation de la résine avec 1 équiv. d'anhydride acétique et 2 équiv. de DIPEA dans le DMF.

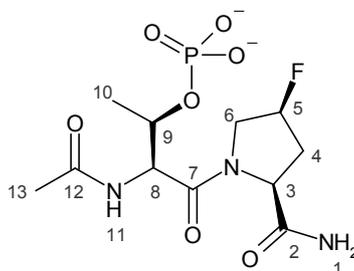


Rendement 22%

LC-MS  $t_R$  0.63 min 92%  $m/z=352$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.95 (d, 1H, 8.48 Hz) [11] ; 7.55 (s, 1H) [1] ; 7.13 (s, 1H) [1] ; 4.67 (dd, 1H, 8.48 Hz, 6.14 Hz) – 55.6 [8]; 4.47 (m, 1H) – 72.5 [9] ; 4.31 (m, 2H) – 59.5 [3] ; 4.31 (m, 2H) – 69.6 [5] ; 3.67 (d, 2H, 3.22 Hz) – 56.3 [6] ; 2.04 (m, 1H) – 38.6 [4] ; 1.88 (s, 3H) – 22.9 [13] ; 1.80 (m, 1H) – 38.6 [4] ; 1.24 (d, 3H, 6.43 Hz) – 18.4 [10].

**Composé (7).**

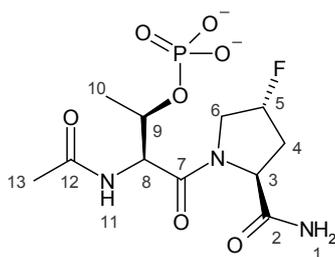


Rendement 70%

LC-MS  $t_R$  0.83 min 92%  $m/z=354$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.10 (d, 1H, 7.60 Hz) [11] ; 7.53 (s, 1H) [1] ; 7.19 (s, 1H) [1] ; 5.43 (m, 0.5H) – 92.0 [5] ; 5.25 (m, 0.5H) – 94.6 [5] ; 4.63 (t, 1H, 6.72 Hz) – 55.1 [8] ; 4.49 (m, 1H) – 72.1 [9] ; 4.44 (m, 1H) – 56.7 [3] ; 4.07-3.99 (m, 1H) – 54.1 [6] ; 3.95-3.90 (m, 1H) – 53.7 [6] ; 2.40-2.25 (m, 1H) – 35.8 [4] ; 2.25-2.17 (m, 1H) – 35.8 [4] ; 1.87 (s, 3H) – 22.4 [13] ; 1.31 (d, 3H, 6.14 Hz) – 18.1 [10].

**Composé (8).**



Rendement 57%

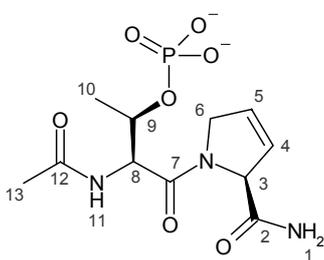
LC-MS

$t_R$  0.90 min 96%

$m/z=354$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.12 (d, 1H, 7.60 Hz) [11] ; 7.55 (s, 1H) [1] ; 7.13 (s, 1H) [1] ; 5.45 (m, 0.5H) – 92.0 [5] ; 5.27 (m, 0.5H) – 94.5 [5] ; 4.55 (t, 1H, 7.31 Hz) – 56.1 [8] ; 4.40 (m, 1H) – 72.2 [9] ; 4.31 (m, 1H) – 58.6 [3] ; 4.29-3.85 (m, 1H) – 54.5 [6] ; 4.22-3.71 (m, 1H) – 53.9 [6] ; 2.45-2.03 (m, 1H) – 36.5 [4] ; 2.39-1.88 (m, 1H) – 36.1 [4] ; 1.85 (s, 3H) – 22.6 [13] ; 1.26 (d, 3H, 6.14 Hz) – 18.2 [10].

**Composé (9).**



Rendement 62%

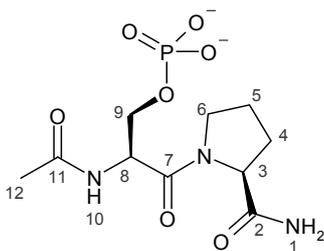
LC-MS

$t_R$  1.05 min 92%

$m/z=335.3$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.16 (d, 1H, 7.02 Hz) [11] ; 7.39 (s, 1H) [1] ; 7.06 (s, 1H) [1] ; 6.02 (dd, 1H, 6.43 Hz, 2.05 Hz) – 128.3 [4] ; 5.79 (dd, 1H, 6.43 Hz, 2.05 Hz) – 126.6 [5] ; 4.92 (m, 1H) – 67.8 [3] ; 4.60 (m, 1H) – 54.4 [6] ; 4.52 (m, 1H) – 55.9 [8] ; 4.46 (m, 1H) – 54.4 [6] ; 4.42 (m, 1H) - 72.6 [9] ; 1.85 (s, 3H) – 22.7 [13] ; 1.31 (d, 3H, 6.14 Hz) – 18.0 [10].

**Composé (10).**

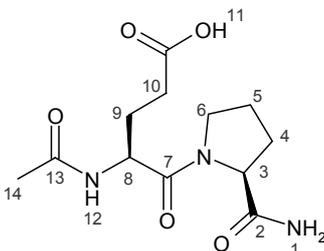


Rendement 28%

LC-MS  $t_R$  1.07 min 99 %  $m/z= 321$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz) :  $\delta$  (ppm) = 8.3 (d, 1H, 7.89 Hz) [10]; 7.26 (s, 1H br) [1]; 7.0 (s, 1H br) [1]; 4.79 (td, 1H, 7.89 Hz, 5.55 Hz) – 51.4 [8]; 4.19 (dd, 1H, 8.48 Hz, 3.80 Hz) – 59.9 [3]; 4.04 (m, 1H) – 64.8 [9]; 3.83 (m, 1H) – 64.8 [9]; 3.63 (m, 2H) – 47 [6]; 2,03 (m, 1H) – 29.5 [4]; 1.88 (m, 2H) – 24.4 [5]; 1.84 (s, 3H) – 22.3 [12]; 1.80 (m, 1H) – 29.5 [4].

**Composé (11).**

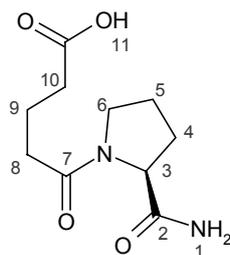


Rendement 34%

LC-MS  $t_R$  3.62 min 100%  $m/z= 283$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz) :  $\delta$  (ppm) = 12.08 (s, 1H) [11]; 8.10 (d, 1H, 7.89 Hz) [12]; 7,27 (s, 1H) [1]; 6.87 (s, 1H) [1]; 4.52 (m, 1H) – 49.9 [8]; 4.2 (dd, 1H, 3.80 Hz, 8.48 Hz) – 59.8 [3]; 3.61 (m, 2H) – 47 [6]; 2.29 (m, 2H) – 29.8 [10]; 2.01 (m, 1H) – 29.5 [4]; 1,90 (m, 3H) – 26.8 [9]; 1;84 (m, 3H) – 24.5 [5]; 1;82 (s, 3H) – 22.4 [14]; 1.77 (m, 1H) – 29.5 [4]; 1;64 (m, 1H) – 26.8 [9].

**Composé (12).**

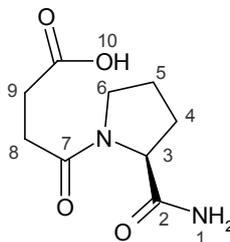


Rendement 43%

LC-MS  $t_R$  3.82 min 100 %  $m/z=$  227 (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.5 (s, 0.5H) [1] ; 7.24 (s, 1H) [1] ; 7.11 (s, 0.5) [1] ; 6.86 (s, 1H) [1] ; 4.23 (dd, 0.5H, 2.92 Hz, 8.48) – 60.1 [3] ; 4.18 (dd, 1H, 2.92 HZ, 8.77 Hz) – 59.6 [3] ; 3.51 (m, 1H) – 46.6 [6] ; 3.4 (m, 1H) – 46.6 [6] ; 2.3 (m, 2H) – 33.1 [8] ; 2.2 (m,2H) – 32 [10] ; 2.0 (m, 1H) – 29.6 [4] ; 1.85 (m, 2H) – 24.4 [5] ; 1.7 (m, 2H) – 19.8 [9] ; 1.81 (m, 1H) – 29.6 [4] ; 12,02 (s, 1H br) [11].

**Composé (13).**

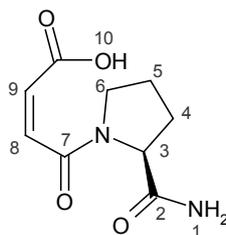


Rendement 38%

LC-MS  $t_R$  3.58 min 100 %  $m/z=$  212 (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.31 (s, 1H br) [10] ; 7.17 (s, 1H) [1] ; 6.90 (s, 1H) [1] ; 4.29 (dd, 1H) – 60.1 [3] ; 4.16 (dd, 1H) – 59.8 [3] ; 3.59 (t, 1H) – 47 [6] ; 3.44 (t, 1H) – 47 [6] ; 3.33 (m, 1H)– 40.5 [6] ; 2.49 (m, 1H) – 28.9 [9] ; 2.4 (m, 1H) – 28.9 [9] ; 2.19 (m, 1H) – 32 [4] ; 2.19 (m, 1H) – 28.7 [8] ; 2.0 (m, 1H) – 30 [4] ; 1.92 (m, 1H) – 32 [4] ; 1.84 (m) – 30 [4] ; 1.86 (m, 2H) – 24 [5] ; 1.81 (m, 1H) – 30 [4] ; 1.75 (m, 2H) – 22.6 [5].

**Composé (14).**

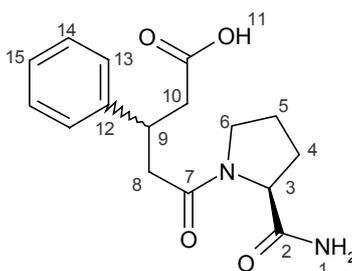


Rendement 30%

LC-MS  $t_R$  3.84 min 100%  $m/z= 210$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.87 (s, 1H br) [10] ; 7.93 (s, 1H) [1]; 7.66 (s, 1H) [1]; 6.7 (d, 1H, 12.11 Hz) – 137.3 [8]; 6.08 (d, 1H, 12.15 Hz) – 126.2 [9]; 4.11 (t, 1H) – 59 [3] ; 3.2 (m, 2H) – 46.5 [6] ; 2.26 (q, 1H) – 29.8 [4] ; 1.88 (m, 1H) – 29.8 [4] ; 1.88 (m, 2H) – 23.5 [5].

**Composé (15).**

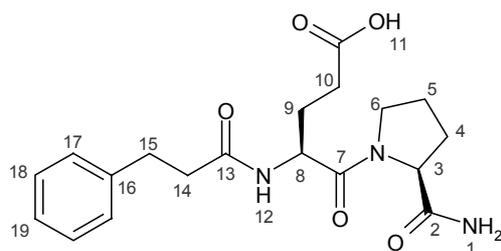


Rendement (total) 40%

LC-MS  $t_R$  4.73/4.87 min 45/45 %  $m/z= 302$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 12.0 (s, 1H br) [11] ; 7.27 (m, 3H) – 128.3 [14] ; 7.27 (m, 3H) – 128.9 [15] ; 7.24 (s, 1H) [1] ; 7.18 (m, 2H) – 127 [13] ; 6.86 (s, 1H) [1] ; 4.19 (m, 1H) – 60 [3]; 4.09 (m, 1H) – 60 [3]; 3.53 (m, 2H) – 48 [6]; 3.53 (m, 2H) – 38.5 [9]; 3.4 (m, 1H) – 47.5 [6]; 3.28 (m, 2H) – 48 [6]; 3.28 (m, 2H) – 47.5 [6]; 2.75 (m, 2H) – 40.5 [10]; 2.6 (m, 2H) – 41 [8] ; 1.95 (m, 1H) – 30 [4] ; 1.82 (m, 1H) – 25 [5] ; 1.79 (m, 1H) – 30 [4].

**Composé (16).**

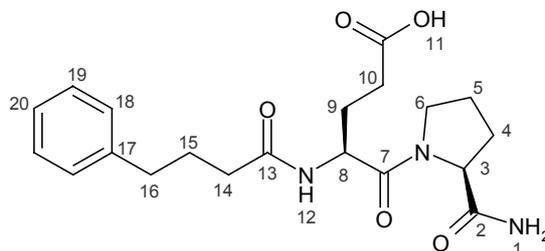


Rendement 54%

LC-MS  $t_R$  4.85min 98%  $m/z= 373$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.1 (d, 1H) [12] ; 7.27 (s, 1H br) [1] ; 6.89 (s, 1H br) [1] ; 7.23 (m, 2H) – 128.8 [18] ; 7.16 (m, 3H) – 128.8 [17] ; 7.16 (m, 3H) – 126.5 [19] ; 4.54 (q, 1H) – 50 [8] ; 4.19 (dd, 1H) – 60 [3] ; 3.56 (t, 2H) – 47.3 [6] ; 2.79 (t, 2H) – 31.6 [15] ; 2.41 (q, 2H) – 37.1 [14] ; 2.24 (t, 2H) – 30.2 [10] ; 2.0 (m, 1H) – 29.9 [4] ; 1.88 (m, 2H) – 25 [5] ; 1.88 (m, 2H) – 27.3 [9] ; 1.79 (m, 2H) – 29.9 [4] ; 1.79 (m, 2H) – 25 [5] ; 1.63 (m, 1H) – 27 [9].

**Composé (17).**

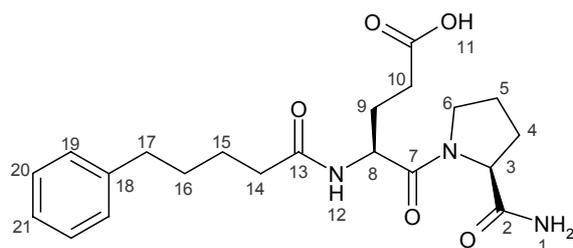


Rendement 61%

LC-MS  $t_R$  5.17min 100%  $m/z= 387$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 12.05 (s, 1H br) [11] ; 8.06 (d, 1H) [12] ; 7.28 (s, 1H br) [1] ; 7.28 (s, 2H) – 128.8 [19] ; 7.17 (m, 3H) – 128.8 [18] ; 7.17 (m, 3H) – 126.4 [20] ; 6.85 (s, 1H) [1] ; 4.53 (q, 1H) – 50 [8] ; 4.21 (dd, 1H) – 60 [3] ; 3.62 (dd, 2H) – 47.3 [6] ; 2.54 (t, 2H) – 35.1 [16] ; 2.29 (t, 2H) – 30.3 [10] ; 2.11 (m, 2H) – 34.9 [14] ; 2.0 (m, 1H) – 29.8 [4] ; 1.9 (m, 2H) – 27 [9] ; 1.9 (m, 2H) – 24.3 [5] ; 1.75 (m, 4H) – 27.5 [15] ; 1.75 (m, 4H) – 29.8 [4] ; 1.75 (m, 4H) – 24.9 [5] ; 1.68 (m, 1H) – 27 [9].

**Composé (18).**



Rendement 58%

LC-MS

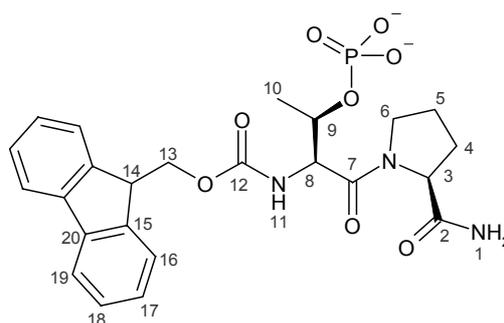
$t_R$  5.53min 100%

$m/z = 401$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.05 (d, 1H) [12] ; 7.27 (s, 1H br) [1] ; 7.27 (m, 2H) – 128.9 [20] ; 7.19 (m, 3H) – 128.9 [19] ; 7.17 (m, 3H) – 126.3 [21] ; 6.85 (s, 1H br) [1] ; 4.52 (q, 1H) – 50.2 [8] ; 4.20 (dd, 1H) – 60.1 [3] ; 3.60 (m, 2H) – 47.4 [6] ; 2.56 (t, 2H) – 35.5 [17] ; 2.28 (t, 2H) – 31 [10] ; 2.13 (m, 2H) – 35.2 [14] ; 1.98 (m, 1H) – 29.9 [4] ; 1.88 (m, 2H) – 24.5 [5] ; 1.88 (m, 2H) – 27.1 [9] ; 1.81 (m, 2H) – 24.5 [5] ; 1.75 (m, 2H) – 29.9 [4] ; 1.65 (m, 1H) – 27.1 [9] ; 1.5 (m, 2H) – 31 [16] ; 1.49 (m, 2H) – 25.3 [15].

*Les composés (19) à (25) contenant des têtes polaires de type phosphonate (19),  $\alpha$ -diacide (20), tétrazole (21), squaramide (22), trifluorosulfonamide (23), 3,4-difluoroaniline (24) et 3-nitroaniline (25), ont été synthétisés, purifiés et contrôlés par le Dr Jean-Frédéric Duckert.*

**Composé (26).**

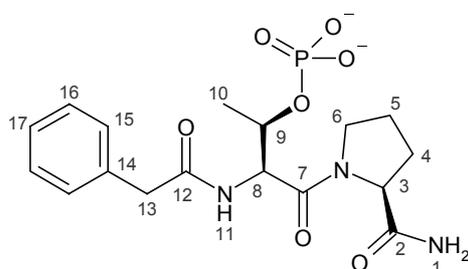


Rendement 52%

LC-MS  $t_R$  5.52 min 99%  $m/z= 515$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.60 Hz) - 120.3 [19] ; 7.72 (d, 2H, 7.60 Hz) – 125.8 [16] ; 7.46 (d, 2H, 6.72 Hz) [11] ; 7.41 (s, 1H) [1] ; 7.40 (t, 2H, 7.31 Hz) – 128.2 [17] ; 7.32 (t, 2H, 7.60 Hz) – 127.6 [18] ; 7.04 (s, 1H) [1] ; 4.46 (m, 2H) – 72.5[9]; 4.46 (m, 2H) – 57.5 [8]; 4.25 (m, 4H) – 66.6 [13]; 4.25 (m, 4H) – 46.9 [14]; 4.25 (m, 4H) – 60.1 [3]; 3.78 (m, 1H) – 47.7 [6]; 3.64 (m, 1H) – 47.7 [6]; 2.06 (m, 1H) – 29.7 [4]; 1.88 (m, 2H) – 24.9 [5]; 1.77 (m, 1H) – 29.7 [4]; 1.32 (d, 3H, 5.85 Hz) – 18.0 [10].

**Composé (27).**

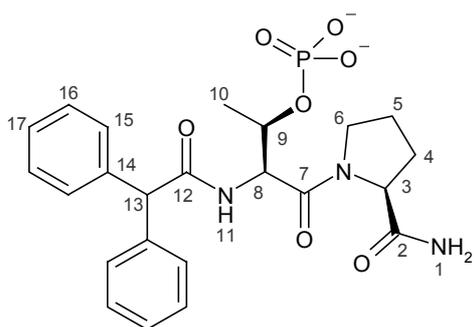


Rendement 54%

LC-MS  $t_R$  4.24min 96%  $m/z= 412$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.27 (d, 1H, 7.60 Hz) [11] ; 7.41 (s, 1H) [1] ; 7.27 (m, 4H) – 129.8 [15] ; 7.27 (m, 4H) – 128.6 [16] ; 7.19 (m,1H) – 126.6 [17] ; 7.09 (s, 1H) [1] ; 4.64 (t, 1H, 7.31 Hz) – 56.0 [8] ; 4.50 (td, 1H, 9.35 Hz, 6.43 Hz) – 72.4 [9] ; 4.21 (dd, 1H, 8.18 Hz, 5.55 Hz) – 60.3 [3] ; 3.73 (m, 1H) – 47.9 [6] ; 3.6 (m, 1H) – 47.9 [6] ; 3.51 (s, 2H) – 42.2 [13] ; 2.05 (m, 1H) – 29.9 [4] ; 1.87 (m, 1H) – 25.0 [5] ; 1.79 (m, 1H) – 25.0 [5] ; 1.74 (m, 1H) – 29.9 [4] ; 1.26 (d; 3H, 6.43 Hz) – 18.1 [10].

**Composé (28).**

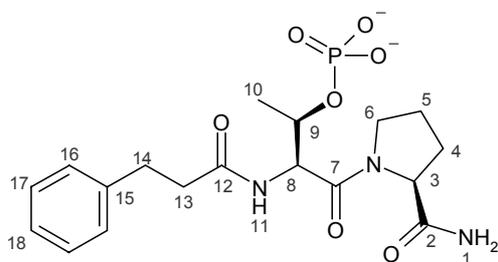


Rendement 65%

LC-MS  $t_R$  5.18 min 96%  $m/z= 488$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.48 (d, 1H, 7.89 Hz) [11] ; 7.45 (s, 1H) [1] ; 7.30 (m, 8H) – 128.9 [15,16] ; 7.28 (m,2H) – 126.9 [17] ; 7.14 (s, 1H) [1] ; 5.22 (s, 1H) – 56.1 [13] ; 4.72 (t, 1H, 6.72 Hz) – 55.7 [8] ; 4.55 (m, 1H) – 72.0 [9] ; 4.20 (dd, 1H, 7.89 Hz, 5.85 Hz) – 60.3 [3] ; 3.73 (m, 1H) – 47.8 [6] ; 3.59 (m, 1H) – 47.8 [6] ; 2.06 (m, 1H) – 29.8 [4] ; 1.87 (m, 1H) – 25.0 [5] ; 1.76 (m, 1H) – 25.0 [5] ; 1.72 (m, 1H) – 29.8 [4] ; 1.23 (d; 3H, 6.72 Hz) – 18.0 [10] .

**Composé (29).**

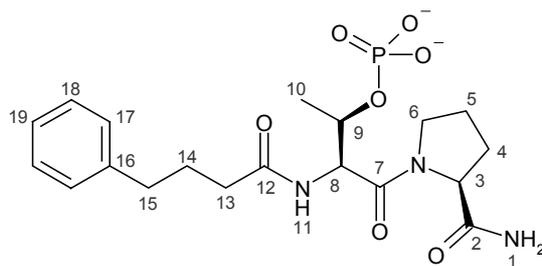


Rendement 58%

LC-MS  $t_R$  4.46 min 98%  $m/z= 426$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.08 (d, 1H, 7.60 Hz) [11] ; 7.36 (s, 1H) [1] ; 7.23 (m, 4H) – 128.5 [16,17] ; 7.16 (m,1H) – 126.0 [18] ; 7.05 (s, 1H) [1] ; 4.64 (t, 1H, 7.31 Hz) – 55.7 [8] ; 4.46 (m, 1H) – 72.3 [9] ; 4.20 (dd, 1H, 8.77 Hz, 6.14 Hz) – 60.0 [3] ; 3.74 (m, 1H) – 47.6 [6] ; 3.6 (m, 1H) – 47.6 [6] ; 2.80 (t, 2H, 8.77 Hz) – 31.2 [14] ; 2.45 (t, 2H, 8.48 Hz) – 36.9 [13] ; 2.06 (m, 1H) – 29.7 [4] ; 1.86 (m, 1H) – 24.9 [5] ; 1.78 (m, 1H) – 24.9 [5] ; 1.74 (m, 1H) – 29.7 [4] ; 1.23 (d; 3H, 6.43 Hz) – 18.0 [10].

**Composé (30).**

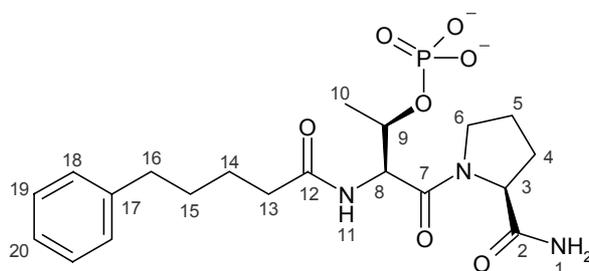


Rendement 56%

LC-MS  $t_R$  4.74 min 98%  $m/z= 440$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz) :  $\delta$  (ppm) = 7.99 (d, 1H, 7.31 Hz) [11] ; 7.37 (s, 1H) [1] ; 7.27 (m, 2H) – 128.6 [18] ; 7.18 (m, 3H) – 128.6 [17] ; 7.18 (m, 3H) – 126.2 [19] ; 7.05 (s, 1H) [1] ; 4.62 (t, 1H, 7.02 Hz) – 55.9 [8] ; 4.46 (m, 1H) – 72.6 [9] ; 4.21 (dd, 1H, 8.18 Hz, 5.55 Hz) – 60.3 [3] ; 3.79 (m, 1H) – 47.9 [6] ; 3.63 (m, 1H) – 47.9 [6] ; 2.54 (t, 2H, 8.18 Hz) – 35.1 [15] ; 2.17 (td, 2H, 7.60 Hz, 3.22 Hz) – 35.0 [13] ; 2.06 (m, 1H) – 30.0 [4] ; 1.87 (m, 2H) – 25.1 [5] ; 1.76 (m, 2H) – 27.5 [14] ; 1.74 (m, 1H) – 30.0 [4] ; 1.23 (d; 3H, 6.14 Hz) – 18.0 [10].

**Composé (31).**

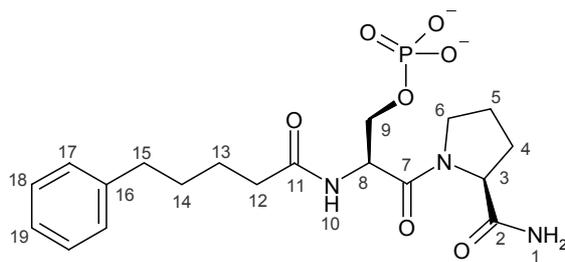


Rendement 62%

LC-MS  $t_R$  5.01 min 97%  $m/z= 454$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz) :  $\delta$  (ppm) = 8.04 (d, 1H, 7.60 Hz) [11] ; 7.36 (s, 1H) [1] ; 7.26 (m, 2H) – 128.5 [19] ; 7.17 (m, 3H) – 128.5 [18] ; 7.17 (m, 3H) – 126.1 [20] ; 7.02 (s, 1H) [1] ; 4.58 (t, 1H, 6.72 Hz) – 55.8 [8] ; 4.44 (m, 1H) – 72.2 [9] ; 4.21 (dd, 1H, 8.48 Hz, 5.26 Hz) – 60.0 [3] ; 3.78 (m, 1H) – 47.7 [6] ; 3.62 (m, 1H) – 47.7 [6] ; 2.56 (t, 2H, 7.02 Hz) – 35.0 [16] ; 2.16 (t, 2H, 7.02 Hz) – 35.0 [13] ; 2.05 (m, 1H) – 29.8 [4] ; 1.87 (m, 1H) – 24.9 [5] ; 1.79 (m, 1H) – 24.9 [5] ; 1.74 (m, 1H) – 29.8 [4] ; 1.52 (m, 4H) – 30.8 [15] ; 1.52 (m, 4H) – 25.0 [14] ; 1.27 (d; 3H, 6.43 Hz) – 18.1 [10].

**Composé (32).**

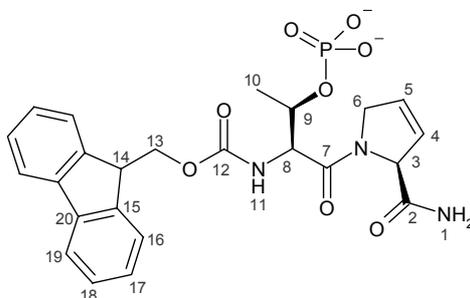


Rendement 58 %

LC-MS  $t_R$  4.91 min 96 %  $m/z= 438$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.24 (d, 1H, 7.89 Hz) [10] ; 7.29 (s, 1H br) [1] ; 7.29 (m, 2H, 7.31 Hz) – 128.6 [18] ; 7.0 (s, 1H br) [1] ; 7.20 (m, 3H) – 128.6 [17] ; 7.20 (m, 3H) – 126 [19] ; 4.79 (dd, 1H, 7.60 Hz) – 51.5 [8] ; 4.19 (dd, 1H, 3.80 Hz, 9.35 Hz) – 60.2 [3] ; 4.06 (m, 1H) – 64.9 [9] ; 3.84 (m, 1H) – 64.9 [9] ; 3.65 (m, 2H) – 47.2 [6] ; 2.55 (t, 2H, 7.02 Hz) – 35 [15] ; 2.15 (t, 2H, 6.72 Hz) – 34.9 [12] ; 2.0 (m, 1H) – 30.7 [4] ; 1.85 (m, 2H) – 24.5 [5] ; 1.79 (m, 1H) – 29.7 [4] ; 1.51 (m, 2H) – 30.6 [14] ; 1.49 (m, 2H) – 25 [13].

**Composé (33).**

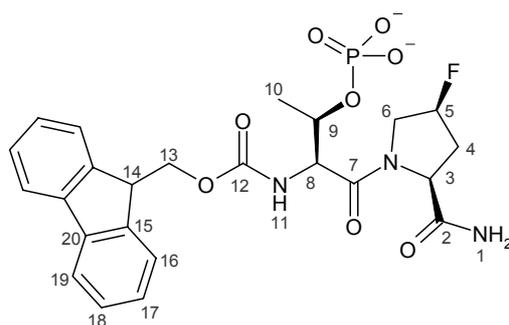


Rendement 57%

LC-MS  $t_R$  5.48 min 90%  $m/z= 514$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.88 (d, 2H, 7.31 Hz) – 120.3 [19] ; 7.70 (m, 3H) – 125.9 [16] ; 7.70 (m, 3H) [11] ; 7.41 (m, 3H) – 128.3 [17] ; 7.41 (m, 3H) [1] ; 7.32 (t, 2H, 7.31) – 127.7 [18] ; 7.05 (s, 1H) [1] ; 6.00 (dd, 1H, 6.72 Hz, 2.05 Hz) – 128.3 [4] ; 5.80 (dd, 1H, 6.72 Hz, 2.05 Hz) – 126.5 [5] ; 4.96 (dd, 1H, 4.97 Hz, 2.05 Hz) – 67.8 [3] ; 4.58 (m, 1H) – 54.4 [6] ; 4.45 (m, 2H) – 54.4 [6] ; 4.45 (m, 2H) – 72.3 [9] ; 4.34 (m, 1H) – 57.9 [8] ; 4.21 (m, 3H) – 47.1 [14] ; 4.21 (m, 3H) – 66.7 [13] ; 1.36 (d, 3H, 6.14 Hz) – 18.1 [10].

**Composé (34).**



Rendement 64%

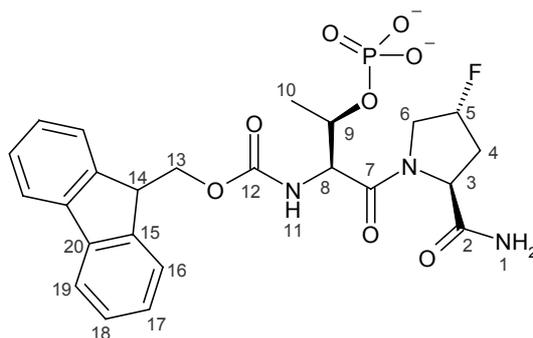
LC-MS

$t_R$  5.48min 93%

$m/z = 534$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.89 Hz) – 120.5 [19] ; 7.72 (d, 2H, 7.60 Hz) – 125.9 [16] ; 7.56 (d, 1H, 7.02 Hz) [11] ; 7.42 (t, 2H, 7.02 Hz) – 128.4 [17] ; 7.33 (t, 2H, 7.60 Hz) – 127.7 [18] ; 7.15 (s, 1H br) [1] ; 7.04 (s, 1H br) [1] ; 5.34 (m, 0.5H) – 92.2 [5] ; 5.34 (m, 0.5H) – 94.7 [5] ; 4.46 (m, 2H) – 72.4 [9] ; 4.46 (m, 2H) – 59.0 [3] ; 4.42 (m, 1H) – 57.4 [8] ; 4.28 (m, 1H) – 66.7 [13] ; 4.21 (m, 2H) – 66.7 [13] ; 4.21 (m, 2H) – 46.8 [14] ; 4.04 (m, 0.5H) – 54.3 [6] ; 3.95 (m, 0.5H) – 54.0 [6] ; 3.6 (m, 0.5H) – 54.3 [6] ; 3.6 (m, 0.5H) – 54.0 [6] ; 2.41-2.25 (m, 1H) – 36.0 [4] ; 2.27-2.17 (m, 1H) – 36.0 [4] ; 1.36 (d, 3H, 6.14 Hz) – 18.2 [10].

**Composé (35).**



Rendement 53%

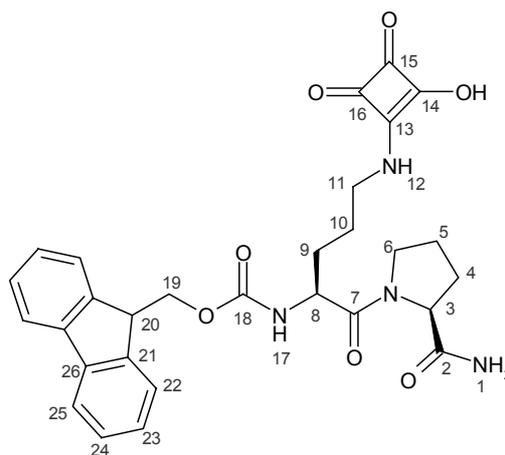
LC-MS

$t_R$  5.64 min 92%

$m/z = 534$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.60 Hz) – 120.5 [19]; 7.73 (d, 2H, 7.60 Hz) – 126.1 [16]; 7.63 (s, 1H br) [1]; 7.53 (d, 1H, 6.72 Hz) [11]; 7.42 (t, 2H, 7.60 Hz) – 128.4 [17]; 7.32 (t, 2H, 7.60 Hz) – 128.0 [18]; 7.16 (s, 1H br) [1]; 5.46 (m, 0.5H) – 92.0 [5]; 5.28 (m, 0.5H) – 95.2 [5]; 4.43 (m, 1H) – 72.7 [9]; 4.41 (m, 1H) – 58.5 [8]; 4.36 (m, 1H) – 59.0 [3]; 4.29-3.86 (m, 1H) – 55.0 [6]; 4.22-3.72 (m, 1H) – 54.6 [6]; 4.22 (m, 3H) – 47.1 [13]; 4.22 (m, 3H) – 67.0 [14]; 2.46-2.07 (m, 1H) – 36.6 [4]; 2.39-1.93 (m, 1H) – 36.1 [4]; 1.32 (d, 3H, 6.14 Hz) – 18.3 [10].

**Composé (36).**

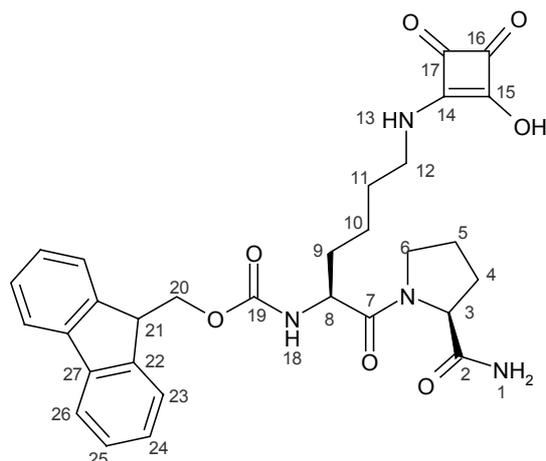


Rendement 30%

LC-MS  $t_R$  5.48 min 90%  $m/z= 546 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 7.88 (d, 2H, 7.60 Hz) – 120.2 [25] ; 7.73 (d, 2H, 7.31 Hz) – 125.8 [22] ; 7.64 (d, 1H, 7.89 Hz) [17] ; 7.41 (t, 2H, 7.60 Hz) – 128.1 [24] ; 7.32 (t, 2H, 7.31 Hz) – 127.6 [23] ; 7.32 (s, 1H) [1] ; 6.80 (m, 2H) [1]; 6.80 (m, 2H) [12]; 4.26 (m, 1H) – 52.7 [8] ; 4.25 (m, 1H) – 66.1 [20] ; 4.19 (m, 1H) – 60.1 [3] ; 4.22 (m, 2H) – 47.0 [19] ; 3.58 (m, 2H) – 47.0 [6] ; 3.39 (m, 2H) – 42.4 [11] ; 1.99 (m, 2H) – 29.4 [4] ; 1.79 (m, 2H) – 29.4 [4] ; 1.92 (m, 2H) – 24.6 [5] ; 1.83 (m, 2H) – 24.6 [5] ; 1.70 (m, 1H) – 27.6 [9] ; 1.54 (m, 3H) – 27.6 [9] ; 1.54 (m, 3H) – 27.6 [10].

**Composé (37).**

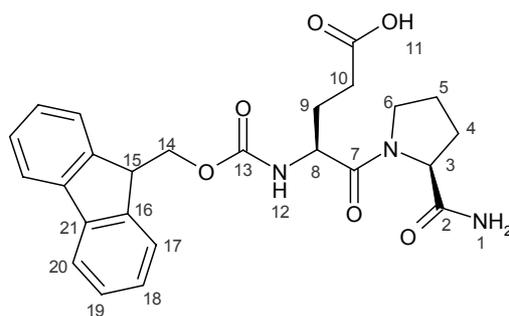


Rendement 30%

LC-MS  $t_R$  5.53 min 99%  $m/z= 560 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.60 Hz) – 120.3 [26] ; 7.73 (d, 2H, 7.31 Hz) – 125.7 [23] ; 7.59 (d, 1H, 7.89 Hz) [18] ; 7.40 (t, 2H, 7.60 Hz) – 127.92 [25] ; 7.32 (t, 2H, 7.31 Hz) – 127.5 [24] ; 7.27 (s, 1H) [1] ; 6.94 (s, 1H) [1] ; 6.86 (t, 1H, 6.14 Hz) [13] ; 4.28 (m, 1H) – 66.1 [21] ; 4.21 (m, 1H) – 60.1 [3] ; 4.20 (m, 1H) – 52.9 [8] ; 4.22 (m, 2H) – 47.0 [20] ; 3.59 (m, 2H) – 47.0 [6] ; 3.43 (m, 2H) – 42.6 [12]; 3.36 (m, 2H) – 42.6 [11]; 1.99 (m, 2H) – 29.3 [4] ; 1.79 (m, 2H) – 29.3 [4] ; 1.96 (m, 2H) – 24.9 [5] ; 1.82 (m, 2H) – 24.9 [5] ; 1.70 (m, 2H) – 30.3 [9] ; 1.50 (m, 2H) – 30.3 [9] ; 1.37 (m, 2H) – 22.2 [10].

**Composé (38).**

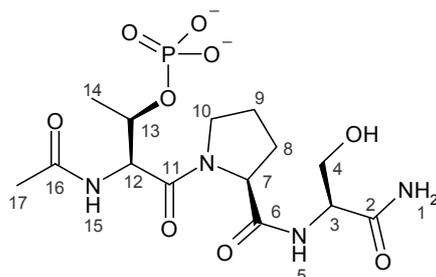


Rendement 60%

LC-MS  $t_R$  6.15min 99%  $m/z= 466 (M+H^+)$

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 7.89 (d, 2H, 7.60 Hz) – 120.7 [20] ; 7.72 (d, 2H, 7.31 Hz) – 126.1 [17] ; 7.64 (d, 1H, 8.18 Hz) [12] ; 7.42 (t, 2H, 7.60 Hz) – 128.5 [18] ; 7.32 (t, 2H, 7.31 Hz) – 127.7 [19] ; 7.27 (s, 1H br) [1] ; 6.87 (s, 1H br) [1] ; 4.29 (m, 1H) – 52.2 [8] ; 4.20 (m, 4H) – 66.5 [14] ; 4.20 (m, 4H) – 60.1 [3] ; 4.20 (m, 4H) – 47.3 [15] ; 3.58 (t, 2H, 6.72 Hz) – 47.3 [6] ; 2.32 (t, 2H, 7.60 Hz) – 30.4 [10] ; 2.0 (m, 1H) – 30 [4] ; 1.89 (m, 1H) – 26.8 [9] ; 1.84 (m, 2H) – 25.1 [5] ; 1.75 (m, 1H) – 30 [4] ; 1.69 (m, 1H) – 26.8 [9]

**Composé (39).**

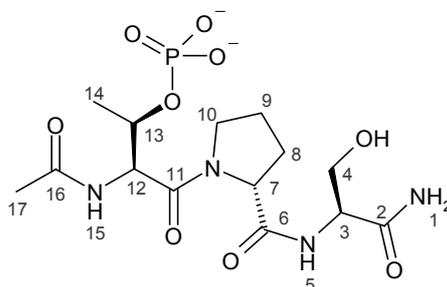


Rendement 68%

LC-MS  $t_R$  1.38 min 93%  $m/z= 422 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz) :  $\delta$  (ppm) = 8.10 (d, 1H) [15] ; 7.85 (d, 1H) [5] ; 7.17 (s, 1H br) [1] ; 7.12 (s, 1H br) [1], 4.62 (t, 1H) – 56 [12] ; 4.43 (m, 1H) – 72.7 [13] ; 4.34 (t, 1H) – 60.6 [7] ; 4.14 (q, 1H) – 56 [3] ; 3.80 (m, 1H) – 48.1 [10] ; 3.67 (m, 1H) – 48.1 [10] ; 3.58 (t, 2H) – 62 [4] ; 2.07 (m, 1H) – 29.5 [8] ; 1.85 (m, 6H) – 29.5 [8] ; 1.85 (m, 6H) – 24.8 [9] ; 1.85 (m, 6H) – 22.6 [17] ; 1.27 (d, 3H) – 18.2 [14].

**Composé (40).**

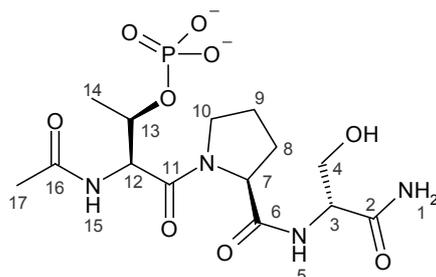


Rendement 45%

LC-MS  $t_R$  1.46 min 88%  $m/z= 422 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz) :  $\delta$  (ppm) = 8.06 (d, 1H) [15] ; 7.85 (d, 1H) [5] ; 7.17 (s, 1H br) [1] ; 7.11 (s, 1H br) [1] ; 4.61 (t, 1H) – 56.2 [12] ; 4.40 (m, 1H) – 72.1 [13] ; 4.34 (m, 1H) – 60.8 [7] ; 4.11 (q, 1H) – 56 [3] ; 3.71 (m, 2H) – 48.1 [10] ; 3.62 (m, 2H) – 62.3 [4] ; 2.0 (m, 2H) – 29.8 [8] ; 2.0 (m, 2H) – 24.7 [9] ; 1.88 (m, 2H) – 29.8 [8] ; 1.88 (m, 2H) – 24.7 [9] ; 1.87 (s, 3H) – 22.8 [17] ; 1.2 (d, 3H) – 18.2 [14].

**Composé (41).**

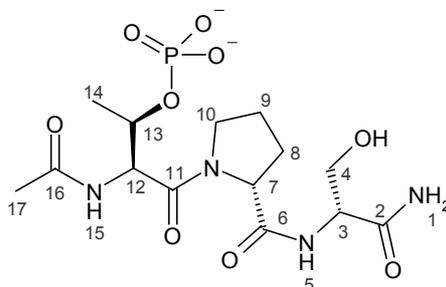


Rendement 72%

LC-MS  $t_R$  1.67 min 94%  $m/z= 422 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 8.28 (d, 1H, 8.48 Hz) [5] ; 8.17 (d, 1H, 8.48 Hz) [15] ; 7.18 (s, 2H br) [1] ; 4.48 (t, 1H ; 7.31 Hz) – 56.6 [12] ; 4.36 (m, 2H) – 72.6 [13] ; 4.36 (m, 2H) – 61.7 [7] ; 4.10 (m, 1H) – 55.9 [3] ; 3.63 (m, 3H) – 48.2 [10] ; 3.63 (m, 3H) – 62.7 [4] ; 2.06 (m, 1H) – 29.8 [8] ; 1.98 (m, 1H) – 25.4 [9] ; 1.80 (m, 5H) – 29.8 [8] ; 1.80 (m, 5H) – 25.4 [9] ; 1.80 (m, 5H) – 22.9 [17] ; 1.24 (d, 3H, 6.14 Hz) – 17.9 [14].

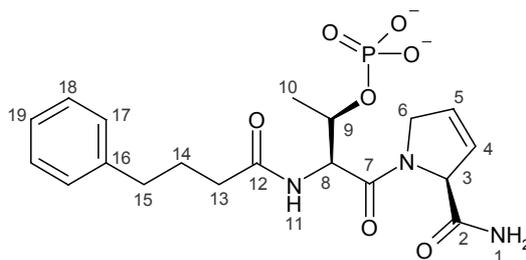
**Composé (42).**



Rendement 68%

LC-MS  $t_R$  3.01min 90%  $m/z= 422 (M-H^+)$

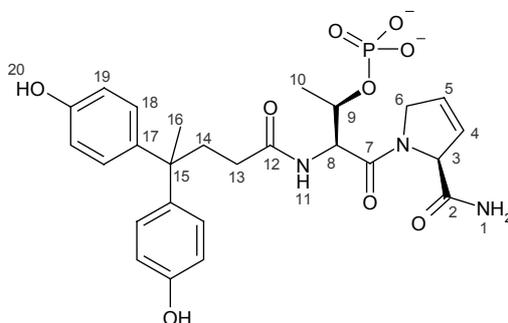
RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 8.68 (s, 1H br) [15] ; 7.54 (d, 1H, 8.48 Hz) [5] ; 7.11 (s, 1H br) [1] ; 6.75 (s, 1H br) [1] ; 4.48 (t, 1H ; 7.31 Hz) – 56.8 [12] ; 4.40 (m, 1H) – 71.6 [13] ; 4.31 (dd, 1H, 3.22 HZ , 8.48 Hz) – 61.2 [7] ; 4.10 (m, 1H) – 56.4 [3] ; 3.88 (m, 1H) – 47.9 [10] ; 3.70 (m, 1H) – 47.9 [10] ; 3.64 (m, 2H) – 61.8 [4] ; 2.11 (m, 1H) – 29.4 [8] ; 1.90 (m, 3H) – 29.4 [8] ; 1.90 (m, 3H) – 24.2 [9] ; 1.85 (s, 3H) – 22.3 [17] ; 1.24 (d, 3H, 6.14 Hz) – 17.9 [14].

**Composé (43).**

Rendement 38%

LC-MS  $t_R$  4.68 min 92%  $m/z= 438$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 8.15 (d, 1H, 6.43 Hz) [11] ; 7.41 (s, 1H) [1] ; 7.26 (m, 2H) – 128.9 [18] ; 7.16 (m, 3H) – 128.9 [17] ; 7.16 (m, 3H) – 126.4 [19] ; 7.05 (s, 1H) [1] ; 6.02 (dd, 1H, 6.43 Hz, 2.05 Hz) – 128.4 [4] ; 5.79 (dd, 1H, 6.14 Hz, 2.05 Hz) – 126.6 [5] ; 4.93 (m, 1H) – 68.0 [3] ; 4.62 (d, 1H, 2.05 Hz) – 54.7 [6] ; 4.45 (d, 1H, 2.05 Hz) – 54.7 [6] ; 4.39 (m, 1H) – 56.5 [8] ; 4.34 (m, 1H) – 72.7 [9] ; 2.53 (t, 2H, 7.02 Hz) – 35.4 [15] ; 2.15 (td, 2H, 7.31 Hz, 2.34 Hz) – 35.2 [13] ; 1.77 (q, 2H, 7.31 Hz) – 27.6 [14] ; 1.32 (d, 3H, 6.14 Hz) – 18.2 [10].

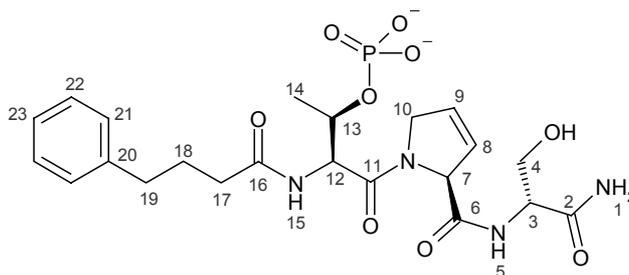
**Composé (44).**

Rendement 35%

LC-MS  $t_R$  4.58 min 99%  $m/z= 560$  (M-H<sup>+</sup>)

RMN <sup>1</sup>H (300MHz, DMSO-d<sub>6</sub>) – RMN <sup>13</sup>C (75MHz):  $\delta$  (ppm) = 9.17 (s, 1H br) [20] ; 8.07 (d, 1H, 6.43 Hz) [11] ; 7.37 (s, 1H) [1] ; 7.04 (s, 1H) [1] ; 6.92 (t, 4H, 8.77 Hz) – 128.2 [18] ; 6.64 (t, 4H, 8.77 Hz) – 115.0 [19] ; 6.00 (dd, 1H, 6.72 Hz, 2.34 Hz) – 128.2 [4] ; 5.78 (dd, 1H, 6.43 Hz, 2.34 Hz) – 126.4 [5] ; 4.91 (dd, 1H, 4.97 Hz, 2.05 Hz) – 67.6 [3] ; 4.57 (d, 1H, 2.05 Hz) – 54.0 [6] ; 4.45 (m, 2H) – 54.0 [6] ; 4.45 (m, 2H) – 55.7 [8] ; 4.38 (m, 1H) – 72.4 [9] ; 2.18 (m, 2H) – 37.5 [13] ; 1.89 (m, 2H) – 31.3 [14] ; 1.44 (s, 3H) – 27.7 [16] ; 1.30 (d, 3H, 6.14 Hz) – 17.8 [10].

**Composé (45).**



Rendement 43%

LC-MS

$t_R$  4.63min 89%

$m/z= 524 (M-H^+)$

RMN  $^1H$  (300MHz, DMSO- $d_6$ ) – RMN  $^{13}C$  (75MHz):  $\delta$  (ppm) = 8.38 (d, 1H, 8.18 Hz) [5]; 8.19 (d, 1H, 5.85 Hz) [12]; 7.26 (m, 2H) – 128.6 [18]; 7.16 (m, 3H) – 126.1 [19]; 7.16 (m, 3H) – 128.8 [17]; 6.05 (dd, 1H, 6.14 Hz, 1.75 Hz) – 128.1 [8]; 5.78 (dd, 1H, 6.72 Hz, 2.34 Hz) – 126.1 [9]; 5.15 (q, 1H, 2.34 Hz) – 67.6 [7]; 4.62 (m, 1H) – 54.2 [10]; 4.51 (m, 1H) – 54.2 [10]; 4.44 (m, 1H) – 56.1 [12]; 4.36 (m, 1H) – 72.8 [13]; 4.12 (td, 1H, 4.68 Hz, 8.18 Hz) – 55.4 [3]; 3.60 (qd, 2H, 10.82 Hz, 5.85Hz) – 61.7 [4]; 2.53 (t, 2H, 7.31 Hz) – 34.9 [17]; 2.13 (td, 2H, 7.60 Hz, 2.05 Hz) – 34.9 [19]; 1.77 (q, 2H, 7.89 Hz) – 27.3 [18]; 1.30 (d, 3H, 5.85 Hz) – 17.9 [14].

## II - Screening par RMN et calcul des constantes de dissociation.

### 2.1 - Expression et purification du domaine WW recombinant.

Le domaine WW porteur d'une queue polyhistidine fusionnée en position N-terminale est exprimé chez E.coli souche BL21(DE3) dotée du plasmide pET28b. Les bactéries sont incubées à 37°C pendant 3 à 4 heures dans 1L de LBensemencé avec les cellules provenant d'une préculture au 1/100<sup>e</sup>. La protéine recombinante est ensuite produite, soit en milieu LB pour la production de protéine non enrichie, soit en milieu minimum M9 avec <sup>15</sup>NH<sub>4</sub>Cl (1 g/L) comme seule source d'azote pour un enrichissement isotopique <sup>15</sup>N, à 31°C pendant 3 heures après induction avec 0.5mM d'isopropylthiogalactose (IPTG). Les cellules sont isolées par centrifugation et resuspendues dans un tampon phosphate 50 mM à pH=8.0 contenant 300 mM NaCl, 0.5 mM DTT, 10 mM imidazole, 0.1% NP40 et un cocktail d'inhibiteurs de protéases (Complete<sup>TM</sup> EDTA-free, Roche). Les cellules sont lysées par sonication et l'extrait soluble isolé par centrifugation est purifié sur une colonne d'affinité chargée en ions Ni<sup>2+</sup> (Chelating Sepharose Fast Flow, Amersham Pharmacia Biotech). La protéine est éluée dans un tampon phosphate 50 mM à pH=8.0 contenant 300 mM NaCl et 250 mM d'imidazole.

Une seconde étape de purification en phase inverse (Source 15RPC ST 4.6/100, Amersham) permet d'obtenir une protéine d'une pureté supérieure à 95% dans les conditions suivantes : la colonne est équilibrée dans une solution aqueuse contenant 0.05% de TFA et la protéine est éluée par un gradient linéaire d'acétonitrile (de 0 à 85% acétonitrile en 16 mL) contenant 0.05% de TFA. Les volumes injectés à partir des fractions d'éluion de la chromatographie d'affinité, dans lesquelles la concentration de la protéine est estimée par un dosage Bradford, correspondent à des quantités de protéines inférieures à 5 mg.

### 2.2 - Screening par RMN

Les expériences de criblage par RMN sont réalisées sur un spectromètre Bruker DMX 600 MHz équipé d'un sonde cryogénique, à 293K. Les spectres <sup>1</sup>H-<sup>15</sup>N HSQC sont réalisés sur une solution à 50 μM pour les titrations et 20 μM pour les tests de criblage de domaine WW enrichi en <sup>15</sup>N solubilisé dans un tampon Tris deutéré (d11) 25 mM pH 6.8, NaCl 50 mM, 5 mM DTT et 2.5 mM EDTA dans un mélange H<sub>2</sub>O 90% : D<sub>2</sub>O 10%. Pour le criblage des composés, on prépare 6mL d'une solution de <sup>15</sup>N-WW à 200 μM que l'on fragmente en 10 échantillons de 60 μL. Sur un des échantillons, on ajoute 540 μL de tampon (contrôle), sur les autres échantillons, on ajoute 120 μL d'une solution de ligand à 2 mM dans l'eau à pH 6.80

(ajusté par ajout de NaOH aq concentré), et on complète à 600  $\mu\text{L}$  avec 420  $\mu\text{L}$  de tampon pour obtenir un ratio molaire ligand:protéine de 20:1

La solution de WW est ajoutée sur des aliquots de ligand lyophilisés préparés à partir d'une solution concentrée dans l'eau. Le  $K_D$  des ligands est calculé à l'aide de l'Equation 1 comme décrit pour les peptides phosphorylés. La quantification des perturbations des déplacements chimiques de la protéine en présence du ligand est calculée à partir de l'Equation 2.

### Equation 1

$$\Delta\delta(\text{ppm}) = 0.5\Delta\delta_{\text{max}} \left( 1 + X + \frac{K_D}{[\text{protéine}]_0} - \sqrt{\left( 1 + X + \frac{K_D}{[\text{protéine}]_0} \right)^2 - 4X} \right)$$

### Equation 2

$$\Delta\delta(\text{ppm}) = \sqrt{\Delta\delta(^1\text{H})^2 + 0.2\Delta\delta(^{15}\text{N})^2}$$

X est le ratio molaire ligand sur protéine,  $[\text{protéine}]_0$  est la concentration initiale de protéine,  $\Delta\delta$  la variation de déplacement chimique calculée comme décrit dans l'Equation 1 et  $\Delta\delta_{\text{max}}$  la valeur de  $\Delta\delta$  à saturation de la protéine. La constante de dissociation est calculée en ajustant sur la représentation graphique la courbe théorique obtenue par l'Equation 2 aux points expérimentaux.

## III – Compétition suivie par FRET pour le calcul des $\text{IC}_{50}$ .

### 3.1 - Synthèse du peptide Ac-KRSRpT212PSLPTPPTR-NH2 intervenant dans les expériences de compétition pour le calcul des $\text{IC}_{50}$ par FRET.

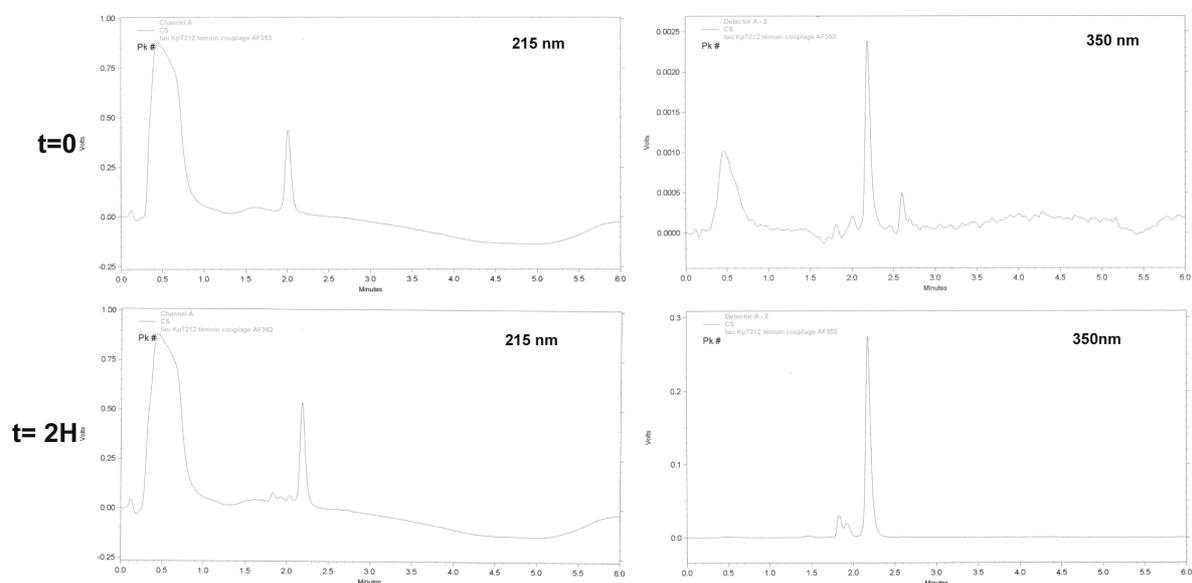
La synthèse du peptide est réalisée dans les conditions standards de synthèse peptidique en phase solide mettant en œuvre une stratégie Fmoc sur un synthétiseur automatique de marque Pioneer<sup>TM</sup> (Applied Biosystems). La déprotection des fonctions amines, en  $\alpha$  de la fonction acide, protégées par un groupement Fmoc est réalisée par une solution de pipéridine à 20% dans le DMF. Le couplage des acides aminés (10 équivalents) est effectué en simple couplage avec un mélange équimolaire d'HOBt/HBTU comme agents d'activation en présence de diisopropyléthylamine (DIPEA), suivi d'une étape de *capping* réalisée avec une solution d'anhydride acétique 0.3M et de DIPEA 0.6M. Les acides aminés N- $\alpha$ -Fmoc (Novabiochem)

utilisés pour la synthèse sont les suivants : Ser(tBu), Thr(tBu), phosphoThr(OBn), Arg(Pbf), Lys(Boc), Pro, Leu (Tableau 1). Après couplage du dernier acide aminé, un *capping* final suit la déprotection du groupement Fmoc. La coupure de la peptidyl-résine et la déprotection des chaînes latérales sont effectuées simultanément par réaction avec une solution constituée de TFA à 95%, triisopropylsilane à 2.5% et d'eau à 2.5%. La résine est incubée dans ce mélange pendant 6 heures à température ambiante. La solution de coupure est séparée de la résine par filtration et le mélange peptidique brut est isolé par précipitation dans un mélange éther-heptane (v/v). La purification est réalisée par RP-HPLC sur colonne C18 Nucleosil équilibrée dans une solution aqueuse contenant 0.05% TFA. Un gradient d'élution allant de 0 à 30% d'acétonitrile en 30 minutes permet de séparer les différents constituants du mélange. Les fractions collectées sont ensuite caractérisées par RP-HPLC et par spectrométrie de masse MALDI-TOF. On obtient ainsi un peptide de pureté supérieure à 95%.  $m/z=1801.9$  ( $M+H^+$ ).

### 3.2 - Couplage de l'AlexaFluor350<sup>®</sup>

On dissout le peptide à 10 mg/mL dans le DMF anhydre et on ajoute une quantité équimolaire d'AlexaFluor350<sup>®</sup> en solution stock à 10 mg/mL dans le DMF anhydre et 2 équivalents de DIPEA. Le mélange est incubé à température ambiante sous agitation et à l'abri de la lumière pendant 2 heures. Le suivi de la réaction est réalisé par RP-HPLC sur colonne C18 à 215 nm et 350 nm. La purification est réalisée dans les mêmes conditions décrites précédemment pour le peptide avec un gradient allant de 0 à 40% d'acétonitrile en 40 min. Le peptide est ensuite caractérisé par RP-HPLC et spectrométrie de masse MALDI-TOF.  $m/z=2093.9$  ( $M+H^+$ ).

Suivi de la réaction de couplage de l'AlexaFluo350<sup>®</sup> sur la fonction amine de la lysine en N-terminal du peptide.



### 3.3 - Screening par FRET.

Les expériences de FRET sont réalisées sur un spectromètre PTI par excitation des tryptophanes du domaine WW à 295 nm. Une solution de 20  $\mu\text{M}$  de WW dans le Tris 50 mM pH 6.8, NaCl 50 mM, DTT/EDTA 5mM contenant des concentrations variables (entre 30 et 120  $\mu\text{M}$ ) de peptide porteur d'une sonde fluorescente (tau pT212) est ajoutée à des aliquots lyophilisés de ligands. Les  $\text{IC}_{50}$  sont calculées à partir de l'Equation 1 où  $\Delta\delta$  est remplacé par le pourcentage d'inhibition donné par l'Equation 3, lui-même défini à partir des variations de l'intensité d'émission du domaine WW à 334 nm. Le  $\text{K}_D$  correspond alors à l' $\text{IC}_{50}$ .

#### Equation 3

Calcul du pourcentage d'inhibition d'un composé pour un mélange WW : phospho-peptide fixé où I est la valeur de l'émission de fluorescence à 334 nm du domaine WW pour une quantité donnée de ligand,  $I_0$  et  $I_{100}$ , les valeurs de l'émission à 334 nm à 0% et 100% d'inhibition respectivement.

$$\% \text{ Inhibition} = \frac{I - I_0}{I_0 - I_{100}}$$