



HAL
open science

HSP90 une cible thérapeutique dynamique : modulation des conformations peuplées par la liaison de ligands

Faustine Henot

► To cite this version:

Faustine Henot. HSP90 une cible thérapeutique dynamique : modulation des conformations peuplées par la liaison de ligands. Biologie structurale [q-bio.BM]. Université Grenoble Alpes [2020-..], 2022. Français. NNT : 2022GRALV044 . tel-03917557

HAL Id: tel-03917557

<https://theses.hal.science/tel-03917557>

Submitted on 2 Jan 2023

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : Biologie Structurale et Nanobiologie

Arrêté ministériel : 25 mai 2016

Présentée par

Faustine HENOT

Thèse dirigée par **Jérôme BOISBOUVIER, Directeur de recherche, CNRS**

préparée au sein du **groupe de RMN biomoléculaire, Institut de Biologie Structurale**
dans l'**École Doctorale Chimie et Sciences du Vivant**

HSP90 une cible thérapeutique dynamique : modulation des conformations peuplées par la liaison de ligands

Thèse soutenue publiquement le **1^{er} Juillet 2022**,
devant le jury composé de :

Dr. Carine VAN HEIJENOORT

Directrice de Recherche, CNRS (Rapporteur)

Dr. Robert SCHNEIDER

Maître de Conférence, Université de Lille (Rapporteur)

Dr. Oriane FRANCES

Chercheuse, Sanofi (Examinatrice)

Dr. Matthias FRECH

Directeur de recherche, Merck KGaA (Examinateur)

Dr. Winfried WEISSEHORN

Professeur, Université Grenoble Alpes (Président)

Dr. Jérôme BOISBOUVIER

Directeur de Recherche, CNRS (Directeur de thèse)



Remerciements

Je tiens tout d'abord à remercier Jérôme Boisbouvier, mon directeur de thèse, pour tout ce qu'il m'a appris ! Un grand merci pour les conseils, de m'avoir poussée à sortir de ma zone de confort de nombreuses fois et pour les « Il n'y a que ceux qui ne font rien qui ne font pas d'erreurs », parfois bienvenus.

Merci également à Pavel Macek pour son encadrement au commencement de ma thèse et à Elisa Rioual, Adrien Favier, Pierre Gans et Claire Loison pour leur contribution dans le projet HSP90, les discussions scientifiques et conseils.

Un grand merci à Rida Awad et Elodie Crublet de m'avoir formée au « wet lab » pour la production / purification de protéines, de m'avoir partagé tous les secrets d'un marquage spécifique méthyle réussi et d'avoir répondu à mes innombrables questions ! Merci également à Lionel Imbert, Isabel Ayala et Rime Kerfah pour leur aide lors de mes passages au labo.

Je tiens également à remercier Alicia Vallet et Adrien Favier pour leurs nombreux coups de main sur le spectromètre RMN comme les jours où il s'est pris la foudre ! Cédric Laguri, merci pour tes conseils, informatiques comme scientifiques, et ta bonne humeur quand tu passais dans le bureau. Merci à Paul Schanda et Dominique Marion pour les précieux cours théoriques de RMN partagés avec les autres étudiants du groupe.

Merci à Ricarda, Elisa, Massilia, Meriem et Irène pour tous les chouettes moments partagés dans le bureau, l'entraide et les fous rires ! Et particulièrement à Ricarda et Massilia de vous être creusé la tête avec moi sur diverses questions théoriques de RMN. Également merci à Arthur, Iva et tous les autres doctorants/stagiaires/post-doc du groupe pour la bonne ambiance au labo, les sorties randos, apéros... et à tous les membres permanents du groupe de RMN biomoléculaire que j'ai pu côtoyer durant ma thèse pour votre disponibilité et vos conseils.

Je tiens aussi à remercier Carine Van Heijenoort, Robert Schneider, Oriane Frances, Matthias Frech et Winfried Weissenhorn d'avoir accepté de faire partie de mon jury de thèse et d'évaluer mon travail. Ainsi que les membres de mon comité de suivi de thèse annuel : Claire Loison, Malene Ringkjøbing Jensen et Cédric Laguri pour leurs conseils et les discussions scientifiques enrichissantes.

Sur une note un peu plus personnelle, je souhaite remercier Eva, mon ancienne coloc, pour cette année et demie passée ensemble, les bons moments partagés (petite dédicace à Riri, Fifi et Loulou) et nos fameux entraînements de volley !

Un grand merci aussi à mes parents pour leur chasse aux typos dans ce manuscrit, d'avoir essayé de comprendre mon sujet de thèse et de m'avoir encouragée ! Et enfin un immense merci à Alex de m'avoir soutenue (et supportée !) tout au long de cette thèse et particulièrement pendant la rédaction du manuscrit.

Résumé en Français

HSP90 est une protéine chaperonne fondamentale, engagée dans le repliement, la stabilisation et la prévention de l'agrégation d'une multitude de protéines clientes. Parmi les protéines clientes de HSP90, on retrouve un nombre important d'oncoprotéines impliquées dans les marqueurs caractéristiques du cancer. De plus, il a également été reporté qu'un niveau élevé de protéine HSP90 était observé dans un grand nombre de cancers suggérant un rôle primordial de HSP90 dans la survie et la croissance de cellules cancéreuses. Ainsi, la chaperonne moléculaire a été reportée comme cible pour le développement de ligands à visée thérapeutique. HSP90 est une protéine très flexible, donc une cible complexe, montrant une variabilité structurale et ce, particulièrement au niveau de l'ATP-lid, segment à proximité du site de fixation des nucléotides/candidats médicaments, comme révélé par plus de 300 structures cristallographiques. Pour les études de conception de médicaments, il est crucial de caractériser la structure exacte de l'ATP-lid mais également d'étudier sa dynamique, cette dernière ayant un rôle clé à la fois sur la cinétique et la thermodynamique de la liaison d'un ligand.

Dans l'étude présentée ici j'ai utilisé la spectroscopie RMN en solution pour caractériser le domaine N-terminal de HSP90 α humaine de 26 kDa. En employant des techniques de marquages isotopiques avancées sur les groupements méthyles j'ai pu attribuer l'ensemble des fréquences du domaine N-terminal de HSP90 et démontrer que l'ATP-lid était en échange entre deux conformations dans la gamme de temps μ s-ms. Les structures de l'état majoritaire et de l'état minoritaire ont été élucidées. Tandis que l'état majoritaire peuple une conformation avec l'ATP-lid en position ouverte comme attendu d'après les structures cristallographiques disponibles du domaine N-terminal de HSP90, l'état minoritaire échantillonne une conformation fermée, distante de plus de 30 Å de l'état ouvert majoritaire. C'est la première fois que ce changement structural important, impliqué dans le cycle fonctionnel de HSP90, est observé pour le domaine N-terminal isolé de HSP90. En utilisant des expériences RMN de CPMG relaxation-dispersion, j'ai pu caractériser à la fois cinétiquement et thermodynamiquement l'échange entre les états ouvert et fermé du domaine N-terminal de HSP90 apo.

J'ai ensuite étudié l'influence de deux inhibiteurs, dérivés de la famille du résorcinol, sur le paysage énergétique de HSP90. J'ai pu montrer que bien que les deux ligands partagent une structure chimique comparable et une affinité similaire envers HSP90, ils modulent différemment la structure et la dynamique de leur protéine cible. L'un des ligands se lie à la conformation préexistante ouverte du domaine N-terminal de HSP90 et accélère l'échange avec l'état fermé par rapport à la protéine apo. A l'inverse, le second ligand, différant uniquement par un substituant, induit la formation d'une nouvelle hélice dans l'état majoritaire avec l'ATP-lid en position ouverte et possède un temps de résidence sur sa cible multiplié par 100. De plus, ce ligand induisant une nouvelle conformation fondamentale de la protéine décélère l'échange conformationnel. Ces résultats apportent une base structurale pour comprendre le rôle de la dynamique conformationnelle de la protéine cible dans la liaison d'un potentiel candidat médicament.

Mots clés : Structure / Dynamique / RMN / HSP90 / Liaison de ligand / Ajustement induit / Sélection conformationnelle / Marquage isotopique / Groupements méthyles / Attribution

Résumé en Anglais

HSP90 is a major chaperone engaged in the folding, stabilization and aggregation prevention of a plethora of client proteins. Among HSP90 clients are found numerous oncoproteins implicated in the hallmarks of cancer. Additionally, it has also been reported that a high level of HSP90 protein was observed in a large number of cancers suggesting a paramount role of HSP90 in survival and growth of cancer cells. Hence, HSP90 has been reported as a target for the development of ligands for therapeutical use. HSP90 is a very flexible protein, thus a challenging target, showing structural variability, especially on its ATP-lid covering the nucleotide/drug binding site, as pointed out by more than 300 crystallographic structures of the N-terminal domain of HSP90. For drug design studies, it is crucial to characterize the exact structure of the ATP-lid segment but also to investigate its dynamics as it was shown to play a role both in the kinetics and thermodynamics of the ligand binding.

In the here presented study, I used solution NMR spectroscopy to characterize the 26 kDa ATP-binding domain of human HSP90 α . Employing advanced isotopic labeling techniques on methyl groups I managed to assign all resonances of the N-terminal domain of HSP90 and demonstrated that the ATP-lid was in exchange between two conformations in the μ s-ms timescale. Both structures of the major and minor states were elucidated. While the major state is found with the ATP-lid in an open position as expected from the previously determined X-ray structures of HSP90 N-terminal domain, it was established that the ATP-lid also samples a closed conformation distant by up to 30 Å from the major state. This is the first time that this important structural change, involved in HSP90 functional cycle, is observed in the N-terminal domain of apo HSP90. Using NMR CPMG relaxation dispersion experiments, I characterized both kinetically and thermodynamically the exchange between the open and closed states of the N-terminal domain of apo HSP90.

I next investigated how two resorcinol derivatives modulate the energy landscape of HSP90. I could show that even though both ligands share a comparable chemical structure and a similar affinity towards HSP90, they differently modulate both structure and dynamics of their target protein. One of the two ligands binds to the pre-existing open conformation of the N-terminal domain of HSP90 and speeds up the exchange with the closed ATP-lid domain compared to the apo protein. On the contrary, the second ligand, differing only by one substituent, induces the formation of a new helix in the ATP-lid major state conformation and has a residence time on target increased by two orders of magnitude. While this ligand induces a new major state conformation, it also slows down the ATP-lid conformational exchange. These results bring a structural basis to understand target protein conformational dynamics' role in binding of a potential drug candidate.

Key words: Structure / Dynamics / NMR / HSP90 / Drug binding / Induced-fit / Conformational selection / Isotopic labeling / Methyl groups / Assignment

Glossaire

BEST : Band-selective Excitation Short-Transient

BLOSUM : BLOcks SUBstitution Matrix

CEST : Chemical Exchange Saturation Transfer

COSY : COrrrelated SpectroscopY

CPMG : Carr-Purcell-Meiboom-Gill

EXSY : EXchange SpectroscopY

FRET : Transfert d'énergie par résonance de type Förster

GM : Geldanamycine

HSP90 : Heat Shock Protein 90

HMQC : Heteronuclear Multiple Quantum Coherence

ITC : Isothermal Titration Calorimetry

L_{tR}^{Court} : Ligand à court temps de résidence

L_{tR}^{Long} : Ligand à long temps de résidence

NTD : Domaine N-terminal

NOE : Nuclear Overhauser Effect

NOESY : Nuclear Overhauser Effect SpectroscopY

PDB : Protein Data Bank

PCS : Pseudocontact Chemical Shift

PRE : Paramagnetic relaxation enhancement

RD : Radicicol

RMN : Résonance Magnétique Nucléaire

SAXS : Small Angle X-ray Scattering

SFX : Serial Femtosecond Cristallography

SOFAST : band Selective Optimized Flip Angle Short Transient

SPR : Surface Plasmon Resonance

SSX : Serial Synchrotron Cristallography

TOCSY : TOtal Correlation SpectroscopY

TROSY : Transverse Relaxation Optimized SpectroscopY

TR-SFX : Time Resolved - Serial Femtosecond Cristallography

TR-SSX : Time Resolved - Serial Synchrotron Cristallography

TRX : Time Resolved Cristallography

XFEL : X-ray Free Electron Laser

Table des matières

1. DYNAMIQUE DES PROTEINES	19
1.1 Importance de la dynamique des protéines	19
1.1.1 Contexte historique	19
1.1.2 Les protéines sont modulées par des changements structuraux d'amplitudes spatiale et temporelle variées	23
1.1.3 Analyse quantitative de la dynamique des protéines	27
1.1.3.1 Paysage énergétique	27
1.1.3.2 Apports potentiels d'une meilleure connaissance de la dynamique des protéines	30
1.2 Exemples de techniques expérimentales pour étudier la dynamique des protéines	31
1.2.1 Cryo-microscopie électronique	31
1.2.1.1 Principe de la microscopie électronique	31
1.2.1.2 Cryo-microscopie électronique (cryo-EM)	32
1.2.1.3 Etude de la dynamique par cryo-microscopie électronique	34
1.2.2 Cristallographie aux rayons X résolue en temps	38
1.2.2.1 Principe de la cristallographie aux rayons X	38
1.2.2.2 Cristallographie aux rayons X résolue en temps	40
1.2.2.3 Etude de la dynamique de la bactériorhodopsine par cristallographie aux rayons X résolue en temps	43
1.2.3 Transfert d'énergie par résonance de type Förster	47
1.2.3.1 Principe	47
1.2.3.2 FRET et dynamique des protéines	50
1.2.3.3 Exemple	51
1.2.4 Diffusion de rayons X aux petits angles (SAXS)	53
1.2.4.1 Principe	53
1.2.4.2 SAXS et dynamique	56
1.2.4.3 Exemple d'étude de dynamique par SAXS	57
1.3 Résonance Magnétique Nucléaire en solution (RMN)	59
1.3.1 RMN et relaxation	60
1.3.2 Méthodes pour quantifier la dynamique des protéines par RMN	62
1.3.2.1 RMN en temps réel ($\tau_{ex} > 1 \text{ s} / k_{ex} < 1 \text{ s}^{-1}$)	64
1.3.2.2 EXSY (Exchange Spectroscopy) ($\tau_{ex} \sim 10\text{-}5000 \text{ ms} / k_{ex} \sim 0.2\text{-}100 \text{ s}^{-1}$)	67
1.3.2.3 CEST (Chemical Exchange Saturation Transfert) ($\tau_{ex} \sim 2 - 50 \text{ ms} / k_{ex} \sim 20\text{-}500 \text{ s}^{-1}$)	71
1.3.2.4 CPMG-RD (Carr Purcell Meiboom Gill Relaxation – Dispersion) ($\tau_{ex} \sim 0.15\text{-}10 \text{ ms} / k_{ex} \sim 100\text{-}6000 \text{ s}^{-1}$)	74
1.3.2.5 Relaxation dispersion dans le référentiel tournant ($\tau_{ex} \sim 20\text{-}100 \mu\text{s} / k_{ex} \sim 10000\text{-}50000 \text{ s}^{-1}$)	78
1.3.2.6 Relaxation de spin nucléaire ($\tau_{ex} \sim \text{ps} - 10 \text{ ns}$)	81

2. DEVELOPPEMENTS METHODOLOGIQUES POUR L'ATTRIBUTION DU DOMAINE N-TERMINAL DE HSP90A HUMAINE	87
2.1 Heat Shock Protein 90	87
2.2 Données préalables sur l'attribution du squelette de HSP90 α -NTD	88
2.3 Intérêt des groupements méthyles	89
2.4 Marquage combinatoire des groupements méthyles de HSP90-NTD	91
2.5 Stratégies d'attribution des groupements méthyles d'une protéine	93
2.5.1 Transfert d'attribution du squelette aux groupements méthyles	93
2.5.1.1 Contexte	93
2.5.1.2 Développement des stratégies de marquage depuis les années 2010	97
2.5.1.3 Leucines et Valine pro-S : problématique	100
2.5.1.4 Développement d'un nouveau précurseur pour faciliter le transfert d'attribution du squelette aux groupements méthyles pro-S des leucines et valines	103
2.5.2 Attribution des groupements méthyles par mutagenèse	105
2.5.3 Attribution des groupements méthyles employant des méthodes basées sur la structure	108
2.5.3.1 Méthodes d'attribution automatique se basant sur l'effet paramagnétique	108
2.5.3.2 Méthodes d'attribution automatique employant l'effet NOE	113
2.6. Application des différentes stratégies d'attribution à HSP90-NTD Apo – Article I	115
2.7. Application des différentes stratégies d'attribution à HSP90-NTD en complexe avec deux ligands	129
2.7.1. HSP90-NTD en complexe avec un ligand conservant l'empreinte spectrale de la protéine	129
2.7.2 HSP90-NTD en complexe avec un ligand modifiant très largement l'empreinte spectrale de la protéine	132
3. ETUDE STRUCTURALE ET DYNAMIQUE DU DOMAINE N-TERMINAL DE HSP90 APO	147
3.1 Variabilité structurale de HSP90-NTD	148
3.1.1 Rôle et cycle fonctionnel d'HSP90	148
3.1.2 Etat de l'art de la PDB	149
3.1.3 Analyse des structures de la PDB	150
3.1.4 Structure de HSP90-NTD en solution à température ambiante	152
3.2 HSP90-NTD échantillonne plusieurs conformations distantes de plus de 30 Å	152
3.2.1 Analyse de NOEs N _H -N _H et CH ₃ -CH ₃	152
3.2.2 Incompatibilité des NOEs observés	153
3.2.3 Calcul de structure à un puis deux conformères	155
3.3 Mutants pour stabiliser individuellement chacune des conformations	156
3.3.1 Conception	156
3.3.2 Production des mutants	157
3.3.3 HSP90-NTD-R60A mutant stabilisant le segment [98-136] en position ouverte	157
3.3.4 HSP90-NTD-R46A mutant stabilisant le segment [98-136] en position fermée	158

3.4 Etude de la dynamique d'HSP90-NTD Apo	160
3.4.1 Acquisition des données de CPMG ¹⁵ N et ¹³ CH ₃	160
3.4.2 Traitement des données de relaxation-dispersion	162
3.4.2.1 Sélection des profils de relaxation-dispersion	162
3.4.2.2 Théorie de la relaxation dispersion et paramètres ajustés.....	164
3.4.2.3 Influence de chacun des paramètres sur l'ajustement	167
3.4.2.4 Vers un ajustement global des données de relaxation dispersion	168
3.4.2.5 Sélection des paramètres initiaux de l'ajustement global	169
3.4.2.6 Processus d'échange commun et ajustement global.....	170
3.4.2.7 Analyse critique des erreurs.....	173
3.4.3 Identification de l'état excité	174
3.4.3.1 Corrélation entre structures et observation des NOEs	174
3.4.3.2 Dépendance de la température	176
3.4.4 Caractérisation thermodynamique et cinétique de l'échange.....	178
3.4.5 Extraction des valeurs d'enthalpie et d'entropie entre l'état fondamental et excité.....	179
3.5 Article III.....	181
4. MODULATION DU PAYSAGE ENERGETIQUE DU DOMAINE N-TERMINAL DE HSP90 PAR LA LIAISON DE DIFFERENTS INHIBITEURS	205
4.1 HSP90 et cancer	205
4.1.1 HSP90 : une cible thérapeutique majeure	205
4.1.2 Différents types d'inhibiteurs de HSP90	207
4.1.2.1 Inhibiteurs de type ansamycine	207
4.1.2.2 Dérivés de résorcinol	208
4.1.2.3 Analogues de la purine.....	209
4.1.2.4 Stratégies de conception d'inhibiteurs de HSP90	210
4.2 HSP90 et inhibiteurs de type résorcinol	211
4.2.1 Deux familles de ligands.....	212
4.2.1.1 Une même affinité envers HSP90-NTD	212
4.2.1.2 Propriétés thermodynamiques et cinétiques différentes pour chaque catégorie de ligands	213
4.2.1.3 Influence des ligands reportée dans la littérature lors de la liaison à HSP90-NTD .	214
4.2.2 Sélection de ligands pour comparer les deux familles	215
4.3 Caractérisation de l'interaction du ligand à court temps de résidence avec HSP90	218
4.3.1 Etat fondamental	218
4.3.2 Etat excité	220
4.3.3 Influence du ligand « L _{TR} ^{Court} » sur la dynamique de HSP90-NTD à 293 K	221
4.3.4 Influence du ligand « L _{TR} ^{Court} » sur la dynamique de HSP90-NTD sur la plage de température [278-313 K].....	222
4.4 Caractérisation de l'interaction du ligand à long temps de résidence avec HSP90	224
4.4.1 Etat fondamental – une nouvelle structure	224

4.4.2 Etat excité	226
4.4.3 Influence du ligand « L_{TR}^{Long} » sur la dynamique de HSP90-NTD à 293 K	229
4.4.4 Influence du ligand « L_{TR}^{Long} » sur la dynamique de HSP90-NTD sur la plage de température [278-313 K].....	230
4.5 Discussion	231
4.5.1 HSP90-NTD Apo et en complexe avec le ligand « L_{TR}^{Court} » - un comportement similaire ...	231
4.5.2 HSP90-NTD complexe avec le ligand « L_{TR}^{Long} »	235
4.5.3 Extraction des valeurs de population de l'état excité et de différences de déplacements chimiques par CPMG-RD lorsque l'on tend vers un régime d'échange rapide	237
4.5.4 Vers un modèle à trois états pour HSP90-NTD	238
5. CONCLUSIONS ET PERSPECTIVES	241
5.1 Attribution de HSP90-NTD Apo	241
5.2 Structure et dynamique de HSP90-NTD Apo	242
5.3 Influence de la liaison de ligands à HSP90-NTD	242
5.3.1 Attribution de HSP90 en complexe avec deux ligands ayant des propriétés différentes ...	242
5.3.2 Structure et dynamique de HSP90-NTD en complexe	243
5.4 Modèle décrivant la différence de temps de résidence des deux ligands.....	245
6. REFERENCES BIBLIOGRAPHIQUES	249
7. ANNEXES	263
7.1. Informations Supplémentaires Article I	263
7.2. Informations Supplémentaires Article III	276

Liste des figures

1.1	Modèles « lock and key », « induced fit » et « conformational selection ».....	21
1.2	Echelle de temps des réarrangements conformationnels dans une protéine.....	23
1.3	Mouvement de large amplitude de l'enzyme SpeB.....	25
1.4	« Mouvements respiratoires » du domaine SH3 de la protéine JIP1.....	26
1.5	Coupe transversale d'un paysage énergétique représentant l'échange entre deux états.....	28
1.6	Paysage énergétique de la myoglobine.....	29
1.7	Paysage énergétique de la catalyse par la réductase dihydrofolate.....	30
1.8	Analyse à particule unique de cryo-microscopie électronique.....	33
1.9	Mécanisme de décodage de la traduction génétique.....	36
1.10	Différentes conformations du bactériophage pendant la phase d'infection.....	36
1.11	Paysage énergétique du ribosome de levure obtenu par cryo-microscopie électronique.....	37
1.12	Schéma de principe de la cristallographie aux rayons-X.....	38
1.13	Cartes de densités électroniques à différentes résolutions.....	39
1.14	Distinction de structures intermédiaires par Time Resolved Crystallography.....	40
1.15	Schéma de principe de la cristallographie aux rayons-X en série.....	41
1.16	Changements conformationnels de la bactériorhodopsine après photo-activation.....	44
1.17	Isomérisation du chromophore rétinale de la bactériorhodopsine.....	45
1.18	Changements structuraux permettant la re-protonation de la base de Schiff de la bactériorhodopsine.....	46
1.19	Echelle des temps du cycle de pompage de la bactériorhodopsine.....	46
1.20	Transfert d'énergie par résonance de type Förster.....	47
1.21	Instrument de sm-FRET à quatre voies d'acquisition.....	49
1.22	Modèle de l'enzyme F0-F1 ATP-synthase d' <i>E. coli</i>	51
1.23	Rotation de la F0-F1 ATP-synthase lors de l'hydrolyse et de la synthèse d'ATP.....	52
1.24	Enzyme F0-F1 ATP-synthase en présence d'AMPPNP.....	53
1.25	Schéma de principe du SAXS.....	54
1.26	Fonction de distribution des distances obtenue par SAXS.....	56
1.27	Représentation de Kratky sans dimension.....	56
1.28	Analyse par SAXS de la protéine kinase R.....	58
1.29	Modèle reproduisant les données expérimentales de la protéine kinase R.....	58
1.30	Représentation du retour à l'équilibre sur z de l'aimantation globale.....	60
1.31	Impact de l'échange chimique entre deux états sur les raies observées par RMN.....	63
1.32	Méthodes RMN pour étudier des phénomènes dynamiques à différentes échelles de temps.....	64
1.33	Réaction d'autoassemblage de la protéine TET2 suivie par RMN en temps réel.....	66
1.34	Modèle cinétique de l'autoassemblage de TET2 initié par un saut de pH.....	67
1.35	Principe de l'Exchange Spectroscopy.....	68
1.36	Evolution de l'intensité des résonances en fonction du temps de mélange EXSY.....	68
1.37	Attribution par mutagénèse des résonances I151 et I149 de la protéase ClpP.....	69
1.38	Echange conformationnel de ClpP suivi par EXSY et CPMG-RD.....	70
1.39	Potentielle route de relargage des substrats dégradés par ClpP.....	71
1.40	Principe du Chemical Exchange Saturation Transfert.....	72

1.41	Processus d'échange I de la protéine SOD1 immature suivi par CEST.....	73
1.42	Processus d'échange II de la protéine SOD1 immature suivi par CEST.....	74
1.43	Principe de la CPMG-RD avec un faible nombre d'impulsions refocalisantes.....	75
1.44	Principe de la CPMG-RD avec un nombre élevé d'impulsions refocalisantes.....	76
1.45	Evolution de la relaxation transversale en fonction de la fréquence CPMG appliquée....	76
1.46	Etude de GIRK par CPMG-RD.....	78
1.47	Principe de la relaxation dispersion dans le référentiel tournant.....	79
1.48	Etude de la dynamique de la protéine ubiquitine par relaxation dispersion dans le référentiel tournant.....	81
1.49	Représentation schématique des paramètres τ_e et S^2	83
1.50	Représentation séquentielle de R_2 et $[^1H]-^{15}N$ NOE pour les protéines ACTR et CBP dans leurs formes libres et liées.....	84
1.51	Représentation séquentielle de $J(0)$ et $J(0.87\omega_H)$ pour les protéines ACTR et CBP dans leurs formes libres et liées.....	85
1.52	Représentation séquentielle de τ_e et S^2 lorsque ACTR et CBP sont en complexe.....	86
2.1	Structure cristallographique d'HSP90 entière.....	88
2.2	Attribution partielle du squelette de HSP90-NTD Apo.....	89
2.3	Comparaison de sensibilité entre les expériences $^1H-^{15}N$ BEST-TROSY et méthyl $^1H-^{13}C$ SOFAST-HMQC.....	90
2.4	Déplacements chimiques spécifiques de chaque type de groupement méthyle.....	91
2.5	Identification de chaque type de groupement méthyle sur le spectre d'HSP90-NTD.....	92
2.6	Linéarisation de la leucine avec le précurseur α -ketoisovalérate non-stéréospécifique.....	94
2.7	Chemin biosynthétique simplifié de l'isoleucine.....	95
2.8	Etapes du transfert d'aimantation intra-résidu pour les séquences d'impulsions (HM)CM(CGBCA)NH et « out and back HCC-COSY ».....	97
2.9	Schéma réactionnel du protocole pour la production d'échantillons marqués U- $[^2H,^{15}N]$ -Leu/Val- $^{13}C^1H_3$ proS.....	97
2.10	Schéma réactionnel du protocole pour la production d'échantillons marqués U- $[^2H,^{13}C,^{15}N]$ -Leu/Val- $[^{13}C^1H_3]$ proR/ $[^{12}C^2H_3]$ proS.....	98
2.11	Schéma réactionnel du protocole pour la production d'échantillons marqués U- $[^2H,^{15}N]$, $[C', C\alpha, C\beta, C\gamma_2-^{13}C_4]$, $[\gamma_2-^1H_3]$ -Isoleucine.....	99
2.12	Schéma réactionnel du protocole pour la production d'échantillons marqués U- $[^2H,^{15}N]$, $[C', C\alpha, C\beta, C\gamma_1, C\delta_1-^{13}C_5]$, $[\delta_1-^1H_3]$ -Isoleucine.....	99
2.13	Techniques de transfert d'attribution du squelette aux groupements méthyles pro-S.....	101
2.14	Double marquage effectué par l'ajout simultané d'un précurseur marqué différemment..	102
2.15	Marquages d'acétoacétates disponibles commercialement et de celui souhaité.....	104
2.16	Etapes de la stratégie d'attribution par mutagénèse dirigée SeSAM améliorée.....	106
2.17	Spectres d'attribution de groupements méthyles avec la méthode de mutagénèse.....	107
2.18	Propriétés paramagnétiques de certains lanthanides.....	109
2.19	Influence d'un centre paramagnétique sur les groupements méthyles à proximité.....	110
2.20	Histogramme des performances de programmes d'attribution automatique des groupements méthyles.....	115
2.21	Superposition des spectres des groupements méthyles d'HSP90-NTD apo avec ceux de la protéine en complexe avec le ligand 5-(2,4-dihydroxy - phenyl) - 4 - (2 - fluoro - phenyl) - 2,4- dihydro -[1,2,4]triazol-3-one.....	129

2.22	¹⁵ N BEST-TROSY de HSP90-NTD en complexe avec le ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one.....	130
2.23	Spectre 2D ¹ H- ¹³ C de HSP90-NTD α en complexe avec le ligand 5 - (2,4 - dihydroxy - phenyl) -4-(2-fluoro-phenyl)-2,4-dihydro- [1,2,4] triazol-3-one.....	131
3.1	Cycle conformationnel d'HSP90 entière.....	148
3.2	Arbre illustrant le principe du regroupement des structures en clusters.....	150
3.3	Comparaison de 8 clusters de structures de HSP90-NTD.....	151
3.4	Matrice des contraintes de distance NOE observées.....	153
3.5	HSP90-NTD α échantillonne au minimum deux conformations en solution.....	154
3.6	Etat ouvert stabilisé par le mutant R60A.....	157
3.7	Etat fermé stabilisé par les mutants R46A et S129A.....	158
3.8	Ensembles conformationnels des états ouverts et fermés de HSP90 α -NTD.....	159
3.9	Séquence d'impulsions employée pour les expériences de CPMG-RD.....	161
3.10	Vitesse de relaxation transversale effective en fonction de la fréquence CPMG.....	163
3.11	Corrélation entre les paramètres obtenus par CPMG selon le nombre de champs magnétiques employés pour l'acquisition des expériences.....	166
3.12	Ajustements de données de relaxation dispersion montrant l'influence des différents paramètres k_{ex} , P^* , $ \overline{\Delta\omega_C} $ et $ \overline{\Delta\omega_H} $	168
3.13	Grilles de recherche pour des données de relaxation ¹³ C ¹ H et ¹⁵ N.....	170
3.14	Vérification du processus d'échange commun.....	171
3.15	Ajustement global des données de relaxation dispersion.....	173
3.16	Histogramme des paires de groupements méthyles proches dans l'espace pour uniquement l'« état fermé » et uniquement l' « état ouvert ».....	175
3.17	Corrélation des phénomènes observés par CPMG-RD et expériences NOESY.....	177
3.18	Diagramme schématique du paysage énergétique de HSP90-NTD α Apo.....	178
3.19	Extraction des contributions enthalpiques et entropiques entre l'état fondamental et l'état excité.....	180
4.1	Structures des différents inhibiteurs de type ansamycine développés.....	208
4.2	Structure de l'inhibiteur radicicol.....	208
4.3	Structures des différents inhibiteurs dérivés du résorcinol.....	209
4.4	Structure du premier inhibiteur synthétique d'HSP90.....	209
4.5	Structures des différents inhibiteurs analogues de la purine.....	210
4.6	Structure chimique commune des ligands disponibles pour cette étude.....	212
4.7	Profil thermodynamique des différents inhibiteurs de HSP90-NTD mesurés par ITC.....	213
4.8	Vitesse d'association des ligands (k_{on}) à HSP90-NTD en fonction de leur vitesse de dissociation (k_{off}).....	214
4.9	Spectres 2D ¹ H- ¹³ C de HSP90-NTD marquée méthyle en complexe avec chacun des quinze ligands.....	217
4.10	Récapitulatif des différentes caractéristiques des deux ligands sélectionnés.....	218
4.11	Superposition des spectres des groupements méthyles d'HSP90-NTD apo avec ceux de HSP90-NTD en complexe avec le ligand « L _{TR} ^{Court} ».....	219
4.12	Prédiction de structure secondaire pour HSP90-NTD Apo et en complexe avec le ligand « L _{TR} ^{Court} ».....	219

4.13	HSP90-NTD Apo et en complexe avec le ligand « L_{tr}^{Court} » : un état excité similaire.....	220
4.14	Diagramme schématique du paysage énergétique de HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} ».....	221
4.15	Evolution de la constante de vitesse d'échange conformationnel en fonction de la température pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} ».....	222
4.16	Evolution de la population de l'état excité de HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} » en fonction de la température.....	223
4.17	Superposition des spectres des groupements méthyles d'HSP90-NTD apo avec ceux de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».....	224
4.18	Prédictions de structure secondaire pour HSP90-NTD apo et en complexe avec le ligand « L_{tr}^{Long} ».....	225
4.19	Structure cristallographique de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».....	226
4.20	Segments de HSP90-NTD montrant une flexibilité/hétérogénéité structurale plus importante que la moyenne.....	228
4.21	Diagramme schématique du paysage énergétique de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».....	229
4.22	Evolution de la constante de vitesse d'échange conformationnel en fonction de la température pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».....	230
4.23	Evolution de la population de l'état excité de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} » en fonction de la température.....	231
4.24	Ajustement global des données de relaxation dispersion pour HSP90-NTD-R60A.....	232
4.25	Diagramme schématique du modèle proposé représentant le paysage énergétique de HSP90-NTD Apo et de HSP90-NTD-R60A.....	233
4.26	Différents modèles à trois états possibles.....	239
5.1	Diagramme schématique du paysage énergétique de l'échange de HSP90-NTD Apo et en complexe avec soit le ligand « L_{tr}^{Court} » soit le ligand « L_{tr}^{Long} ».....	245
5.2	Structures tridimensionnelles et surfaces des conformations possibles de HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} ».....	246
5.3	Structures tridimensionnelles et surfaces des conformations possibles de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».....	247

Liste des tableaux

Tableau 2.1 : Substitutions suggérées pour les résidus comportant des groupements méthyles.....	106
Tableau 3.1 : Tableau récapitulatif des paramètres employés dans les expériences CPMG $^{13}C^1H_3$ et ^{15}N	161

Préambule

L'objectif de ma thèse est d'étudier la dynamique de la chaperonne HSP90 et de voir comment son paysage énergétique est modulé par la liaison de différents inhibiteurs dérivés de la famille du résorcinol. L'importance de la dynamique des protéines pour effectuer leurs diverses fonctions ainsi que les informations que l'on peut obtenir en étudiant cette dynamique sont présentées dans le **chapitre 1**. Plusieurs techniques expérimentales permettant de mesurer la dynamique des protéines y sont introduites et un focus sur la résonance magnétique nucléaire, technique principale employée dans ma thèse, est effectué. J'ai choisi d'introduire les notions biologiques sur HSP90, nécessaires à la compréhension des différents chapitres, au début de chaque chapitre. Les développements méthodologiques mis en place pour l'attribution par RMN des résonances de ma protéine cible : le domaine N-terminal de HSP90, particulièrement dynamique, sont décrits dans le **chapitre 2**. L'étude de la structure et de la dynamique du domaine N-terminal de HSP90 dans son état Apo est reportée dans le **chapitre 3**. Le **chapitre 4**, quant à lui, décrit la modulation du paysage énergétique du domaine N-terminal de HSP90 par la liaison de deux inhibiteurs partageant une structure chimique très similaire mais différent à la fois cinétiquement et thermodynamiquement. Le **chapitre 5** clôture cette thèse en ouvrant des perspectives en vue d'une éventuelle poursuite de l'étude de la dynamique du domaine N-terminal de HSP90.

Chapitre 1

Dynamique des protéines

1.1 Importance de la dynamique des protéines

1.1.1 Contexte historique

Pendant plus de la moitié du 20^{ème} siècle, la grande majorité des protéines globulaires a été considérée comme adoptant une structure rigide et compacte. Dès 1900, grâce au travail de nombreux scientifiques tels que Mulder, Fischer, Hofmeister et de nombreux autres, il est reconnu que les protéines sont des macromolécules composées d'acides aminés joints entre eux par des liaisons peptidiques (Tanford and Reynolds 2001). A partir du milieu des années 1920, la plupart des chercheurs adoptent le point de vue selon lequel les protéines globulaires ont généralement une forme compacte dont le diamètre est nettement plus petit que celui obtenu pour la chaîne linéaire d'acides aminés composant cette même protéine. Plusieurs concepts physiques, étudiés notamment par Sutherland et Linderstrom-Lang, ont étayé cette idée en dévoilant la forte densité de charge à la surface de la protéine et la forme de cette dernière proche d'une sphère (Tanford and Reynolds 2001).

Les premiers motifs de diffraction obtenus avec des cristaux de pepsine en 1934 par Bernal et Crowfoot (Bernal and Crowfoot 1934) ont confirmé la forme compacte des protéines et montré une forme sphéroïde de la pepsine. La première structure tridimensionnelle de la myoglobine obtenue par cristallographie aux rayons X par Kendrew et al. en 1958 (Kendrew et al. 1958) a également conforté cette forme compacte et, additionnellement, la vision statique de la structure des protéines. Ces dernières possèderaient une structure ordonnée avec chaque atome ayant une position fixe dans l'espace. Toutefois, la représentation statique des protéines est apparue bien avant l'obtention de la première structure tridimensionnelle. En effet, le principe de « lock and key » décrit par Fischer en 1894 pour expliquer la spécificité des enzymes envers leurs substrats repose sur l'hypothèse d'une complémentarité de structures rigides entre l'enzyme (« lock ») et le substrat (« key »). Ainsi, les enzymes pourraient assurer leurs fonctions catalytiques uniquement si le substrat est structuralement complémentaire à l'enzyme (Fischer 1894) (Figure 1.1a).

Cependant, cette représentation statique apparaît insuffisante puisqu'elle ne permet pas d'expliquer certains phénomènes observés (Teilum et al. 2009). En effet, le modèle de « lock and key » est capable d'expliquer pourquoi un substrat analogue avec les mêmes sites réactifs que le substrat spécifique mais plus large que ce dernier ne montre pas d'activité catalytique. A l'inverse, le modèle ne permet pas de comprendre pourquoi un substrat analogue mais cette fois-ci plus petit que le substrat spécifique ne montre pas d'activité enzymatique (Koshland 1958, 1959; Neal 2021). Ainsi, en 1958, Koshland a proposé un nouveau modèle appelé « induced-fit » (modèle de l'ajustement induit) dans lequel la liaison d'un substrat/ligand à une enzyme/protéine cible entraîne un changement conformationnel de cette dernière pour s'adapter à la forme du substrat/ligand (Figure 1.1b), introduisant pour la première fois le rôle de la dynamique des protéines pour la fonction de ces dernières. L'étude par cristallographie aux rayons-X du lysozyme dans sa forme Apo (sans ligand) et Holo (en complexe avec un ligand) a permis de faire accepter par le plus grand nombre le modèle « induced-fit » et le changement conformationnel de la protéine associé – la protéine acceptant plusieurs ligands différents au même site de liaison et des différences structurales sont observées entre la forme Apo et Holo (Blake et al. 1965; Johnson and Phillips 1965).

En 1965, un troisième modèle appelé « conformational selection » (sélection conformationnelle) ou modèle de Monod-Wyman-Changeux a été proposé (Monod et al. 1965), prenant également en compte le rôle de la dynamique des protéines pour leur fonction. Dans ce modèle, la protéine échantillonnerait plusieurs états dans sa forme Apo et l'une d'entre elle serait sélectionnée lors de la liaison avec le ligand (Figure 1.1c).

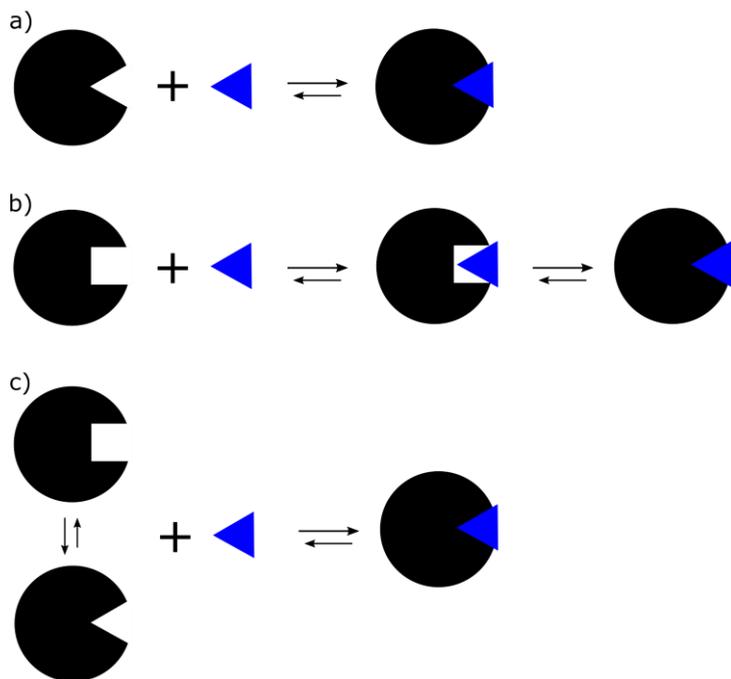


Figure 1.1 : a) Modèle « lock and key » introduit par Emil Fischer en 1894 (Fischer 1894) b) Modèle « induced fit » introduit par Koshland en 1958 (Koshland 1958) c) Modèle « conformational selection » introduit par Monod – Wyman – Changeux en 1965 (Monod et al. 1965). Pour chacun des trois modèles, la protéine est représentée en noir et le substrat en bleu.

Ce troisième modèle représente aussi en partie le point de vue de nombreux physiciens tels que Frauenfelder, Cooper et Careri puisqu'il répond aux principes de la thermodynamique statistique selon lesquels les protéines subissent en permanence des fluctuations aléatoires de structure. En effet, d'après Frauenfelder et Cooper (Cooper 1976; Neal 2021), les interactions électrostatiques maintenant la structure tridimensionnelle des protéines sont relativement faibles. Aussi, les protéines d'un échantillon seront en perpétuel mouvement, dû aux collisions entre les protéines et le solvant qui viendront déstabiliser les interactions de faible énergie, et ainsi échantillonneront un ensemble de sous-états dégénérés dans lesquels la conformation locale de parties de la protéine varie. Différentes méthodes expérimentales ont mis en évidence dans les années 1970-1980 les fluctuations conformationnelles des protéines globulaires. On notera par exemple, la phosphorescence, la résonance magnétique nucléaire et l'extinction de fluorescence (Allerhand et al. 1971; Lakowicz and Weber 1973; Saviotti and Galley 1974; Cooper 1976) ou encore la diffraction de rayons X à différentes températures utilisées comme méthodes pour étudier la dynamique des protéines (Frauenfelder et al. 1979).

Aussi, on peut se demander pourquoi cette dynamique conformationnelle des protéines, mise en évidence dès les années 1970 par différentes techniques expérimentales, n'a été une idée acceptée par le plus grand nombre qu'autour des années 2000. Plusieurs points développés par Neal (Neal 2021) permettent d'apporter

une partie de réponse à cette interrogation. Tout d'abord, bien que les deux modèles proposés (ajustement induit et sélection conformationnelle), mettent en évidence le rôle de la dynamique des protéines dans la fonction de ces dernières, il est possible de constater qu'ils ne s'opposent pas totalement à la représentation statique des protéines. En effet, bien qu'une dynamique importante soit nécessaire pour passer d'un état à l'autre dans le modèle de sélection conformationnelle, ce dernier ne met pas l'accent sur le passage de l'une à l'autre des conformations mais plutôt sur les états finaux échantillonnés. Aussi, les modèles proposés par Koshland et Monod-Wyman-Changeux présentent plusieurs conformations d'une même protéine mais les conformations présentes dans chaque modèle peuvent correspondre à une vue statique des protéines (Neal 2021). De plus, dans les années 1970, l'idée d'une dynamique conformationnelle des protéines provenait principalement des principes de la thermodynamique. Or, le rôle exact des fluctuations conformationnelles sur les fonctions des protéines n'était pas clairement défini. L'apparition du modèle de Cooper et Dryden (Cooper and Dryden 1984), définissant comment la dynamique de la protéine induite par la liaison d'un ligand/substrat pouvait entraîner des effets allostériques sur des sites de fixation autres sans entraîner de changements conformationnels macromoléculaires de la protéine d'intérêt, a permis de dessiner le premier lien entre dynamique et fonction de la protéine. Toutefois, comparé au précédent modèle décrivant l'effet allostérique par des changements structuraux (Monod et al. 1965), le modèle de Cooper et Dryden a eu peu d'impact au moment de sa publication et de fait pendant les années qui ont suivi. En effet, le modèle est parfois cité mais les changements structuraux pour expliquer l'allostérie, plus visuels, dominent par rapport aux changements dynamiques décrits par des équations de thermodynamique et mathématiques. Par exemple, on notera la publication de Petter et al. (Petter et al. 1990) où l'effet allostérique décrit est entièrement basé sur des changements structuraux pouvant correspondre à une vision statique des protéines. Petter et al., ne prennent pas en compte les changements de dynamique potentiels de leurs composés. Toutefois, dans une note de bas de page, ils reconnaissent que différents mécanismes allostériques n'impliquant pas de changements structuraux existent également pour expliquer l'allostérie et citent le travail de Cooper et Dryden.

Cependant, certains scientifiques tels que Akke, Braxton, Cooper ou Reinhart convaincus de l'importance de la dynamique des protéines ont poursuivi leurs travaux et dans les années 2000, ces derniers, reportant un rôle clé de la dynamique des protéines, ont fait changer la vision statique des protéines qui dominait depuis le début du 20^{ème} siècle en une vision dynamique qui ne pouvait plus être ignorée, où les protéines sont en perpétuel mouvement.

Aussi, il est maintenant communément admis que la dynamique des protéines se révèle cruciale pour que ces dernières puissent effectuer de nombreux processus biologiques complexes tels que la reconnaissance de partenaires physiologiques, la catalyse enzymatique, le repliement des protéines et la liaison à des ligands.

1.1.2 Les protéines sont modulées par des changements structuraux d'amplitudes spatiale et temporelle variées

Toutes les protéines sont dynamiques et cette propriété des protéines est encodée dans leur structure qui elle-même dépend de la séquence de ces dernières. La dynamique d'une protéine est caractérisée par un changement dans les coordonnées atomiques de la protéine en fonction du temps (Henzler-Wildman and Kern 2007). Lorsque la dynamique entraîne une large modification de la structure de la protéine on dit que cette dernière est alors flexible. Ainsi, les protéines peuvent être flexibles en conséquence de leur dynamique mais le fait qu'elles soient dynamiques n'implique pas forcément qu'elles soient flexibles (Teilum et al. 2009). En effet, toutes les protéines sont dynamiques mais des mouvements d'amplitude spatiale variée en résultent.

Ainsi, l'amplitude des mouvements d'une protéine peut varier du centième d'angström à la centaine d'angström mais l'échelle des temps caractéristiques des mouvements conformationnels d'une protéine s'étend également sur une multitude d'ordres de grandeur : de la femto seconde à plusieurs heures (Figure 1.2). Les mouvements d'une protéine fluctuent donc spatialement et temporellement sur une large amplitude.

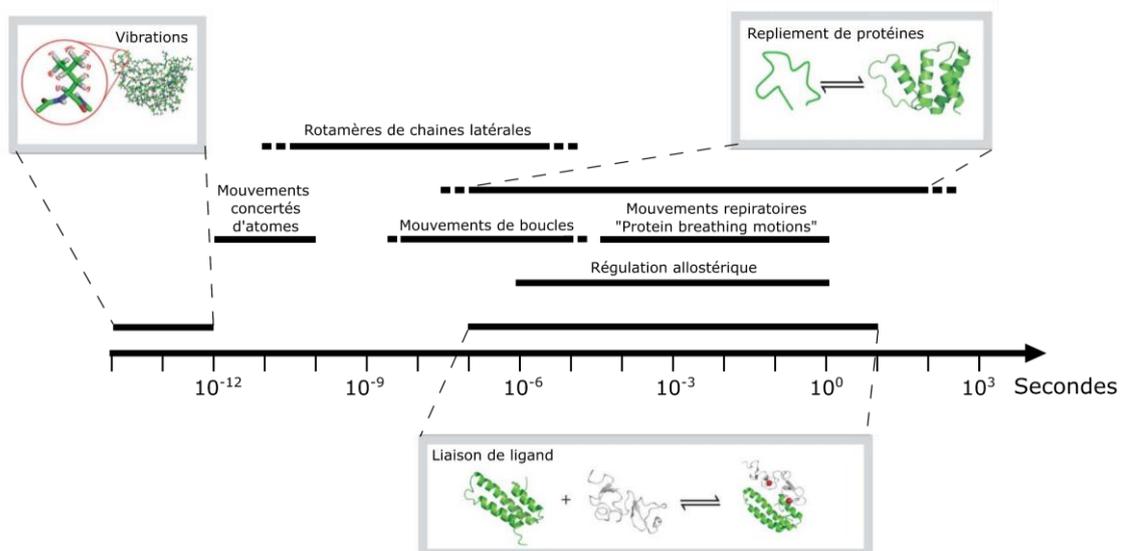


Figure 1.2 : Echelle de temps proposés pour les différents réarrangements conformationnels pouvant se produire dans une protéine. Figure adaptée de (Agarwal 2006; Henzler-Wildman and Kern 2007; Teilum et al. 2009; Skjærven et al. 2011).

Les vibrations atomiques, couvrant généralement une distance inférieure à 0.5 Å, interviennent dans la gamme de temps de la dizaine/centaine de femto secondes (10^{-14} - 10^{-13} s) et l'énergie de ces mouvements vibrationnels provient de l'énergie cinétique inhérente à la protéine (Petsko and Ringe 1984; Skjærven et al. 2011). En plus des fluctuations, deux autres catégories de mouvement peuvent être discernées. On note d'un côté les mouvements collectifs pouvant inclure des mouvements rapides mais peu fréquents tels que la rotation de 180° d'une tyrosine ou des mouvements lents tels que le mouvement d'une boucle ou d'un domaine et d'un autre côté les mouvements induits par une sollicitation externe telle que la liaison d'un ligand à la protéine. Les mouvements appartenant à la première catégorie peuvent échantillonner des distances de quelques dixièmes à plusieurs dizaines d'angströms (Petsko and Ringe 1984). Par exemple, la rotation d'une chaîne latérale pourra avoir une amplitude comprise entre 0.5 et 5 Å, les mouvements de domaines, plus larges, auront une amplitude de l'ordre de 5 à 10 Å et le repliement d'une protéine pourra atteindre plusieurs dizaines d'angströms (Samish et al. 2009). Les mouvements appartenant à la seconde catégorie peuvent également couvrir une large amplitude du dixième d'angström pour la liaison du ligand à quelques angströms pour les changements conformationnels du site actif ou les transitions allostériques. Toutefois il est à noter que ces mouvements sont protéines dépendant et ainsi les distances échantillonnées peuvent être plus disperses en amplitudes que les valeurs citées précédemment (Samish et al. 2009).

La dynamique des protéines peut, comme mentionné précédemment, impliquer des mouvements de diverses amplitudes. On notera par exemple le mouvement de large amplitude de l'exotoxine pyrogénique streptococcique B (SpeB) pour atteindre sa forme active. SpeB est une protéase à cystéine sécrétée et repliée sous la forme d'un précurseur inactif. Ce dernier est transformé en protéase active par clivage autolytique de 118 résidus du domaine N-terminal, région également appelé prodomaine (Kagawa et al. 2000). Les structures de SpeB sous sa forme active et celle inactive ont été résolues, respectivement par résonance magnétique nucléaire en solution (Wang et al. 2009) et cristallographie aux rayons X (Kagawa et al. 2000). L'ablation du prodomaine entraîne un changement conformationnel important : la boucle constituée des résidus 226 à 244 se déplace et ce, de plus de 25 Å (Figure 1.3). Le mouvement de cette boucle va permettre de libérer une seconde boucle contenant l'histidine 195 catalytique. His-195 pourra ainsi adopter sa conformation active dans laquelle une interaction avec la cystéine 47 est possible et SpeB sera fonctionnelle (Teillum et al. 2009).

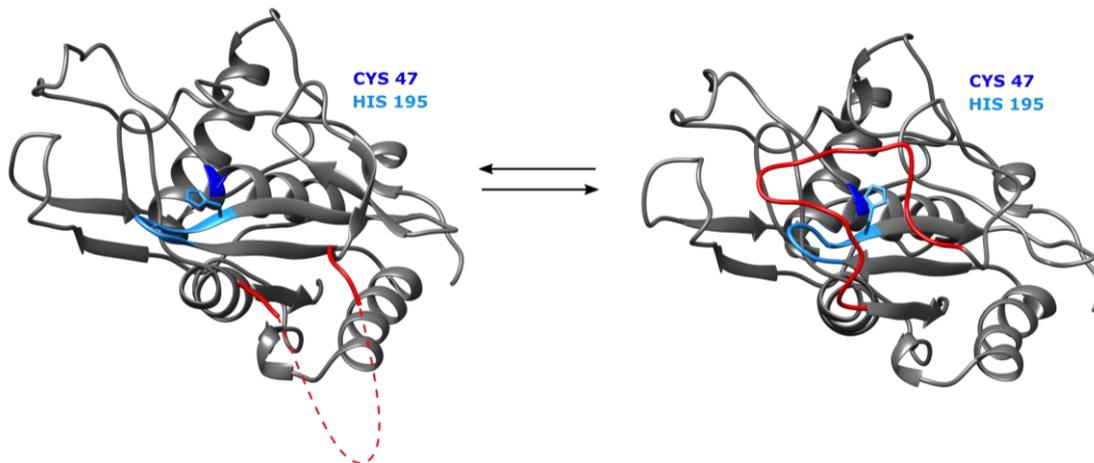


Figure 1.3 : A gauche structure de SpeB dans sa forme inactive (PDB : 1DKI), pour laquelle le prodomaine a été enlevé sur la figure, à droite, structure de SpeB dans sa forme active (PDB : 2JTC). En rouge est représentée la boucle constituée des résidus 226 à 244 se déplaçant de plus de 25 Å. En bleu clair est représentée la seconde boucle, contenant l'histidine 195, libérée par le mouvement de la première boucle. En bleu foncé est représentée la cystéine 47. Les pointillés rouges connectent la structure là où la densité électronique n'a pas permis de résoudre la structure. Figure inspirée de (Teilum et al. 2009).

De plus petits mouvements que le réarrangement structural effectué par SpeB, tel que notamment la rotation de 180° des chaînes latérales des acides aminés comportant un cycle aromatique, ont également été observés dans les protéines depuis les années 1970. Toutefois, les détails structuraux concernant le réarrangement des acides aminés localisés à proximité du cycle aromatique en rotation ainsi que les détails structuraux sur la rotation de ce dernier demeuraient peu connus. Une récente étude par Mariño Pérez et al. (Mariño Pérez et al. 2022) a analysé la dynamique du domaine replié SH3 de la protéine JIP1. JIP1 est une protéine de 711 acides aminés contenant des régions intrinsèquement désordonnées mais également deux domaines repliés : SH3 (Src Homology 3) et PID (Phosphotyrosine Interaction Domain). Dans le domaine SH3, la chaîne latérale d'une tyrosine enfouie dans le cœur hydrophobe du domaine est en rotation et les réarrangements structuraux de la protéine se produisant à proximité de la tyrosine ont été examinés. Cette étude révèle, en utilisant la résonance magnétique nucléaire (RMN) en solution, que SH3 est en échange entre deux conformations dans la gamme de temps μs - ms ($\sim 2600 \text{ s}^{-1}$). La structure cristallographique de SH3 a dévoilé une conformation éclipsée du cycle aromatique de la tyrosine dans l'état majoritaire ($\sim 97\%$). Mariño Pérez et al. ont conçu, en se basant sur une analyse de séquences de SH3 et de structures cristallographiques, des mutants stabilisant la conformation minoritaire peuplée par le domaine SH3. Ils ont montré que le segment 517-522, intrinsèquement dynamique, se réarrangeait pour stabiliser le cycle aromatique dans une conformation décalée. La RMN et la dynamique moléculaire ont permis de conclure que la rotation du

cycle aromatique de 180° ne contribuait que très peu au processus d'échange entre l'état majoritaire et minoritaire, cette dernière étant rapide, de l'ordre de la dizaine de nanosecondes. Ainsi, le domaine SH3 est en échange entre deux conformations, l'une majoritaire avec le cycle aromatique de la tyrosine en position éclipsée et une seconde minoritaire avec le cycle aromatique de la tyrosine en position décalée. Pendant la transition entre l'une et l'autre des conformations se crée un volume de vide autour de la tyrosine, permettant la rotation de 180° de la chaîne latérale de cette dernière. Ce volume vide, augmentant de plus de 65 \AA^3 le volume de la poche, est principalement créé par la réorientation de la chaîne latérale du résidu 520, localisé au sein du segment 517-522 intrinsèquement dynamique. Lorsque le cycle aromatique est stabilisé en position décalée on observe un réarrangement structural associé à une diminution du volume libre autour de la tyrosine, c'est un événement rare se produisant dans une échelle de temps lente de l'ordre de la milliseconde (Figure 1.4). Ces résultats récents montrent que la présence d'un acide aminé spécifique peut être à l'origine de dynamique et amènent à une compréhension structurale des mouvements associés à la rotation d'un cycle aromatique enfoui. Un changement subtil dans l'environnement d'une protéine dans une conformation énergétiquement défavorable peut conduire à des changements structuraux importants. Cette étude met en valeur l'importance de l'étude des phénomènes dynamiques en parallèle d'une étude structurale afin d'améliorer la compréhension du lien entre structures d'états parfois peu peuplés et fonctions de la protéine d'intérêt.

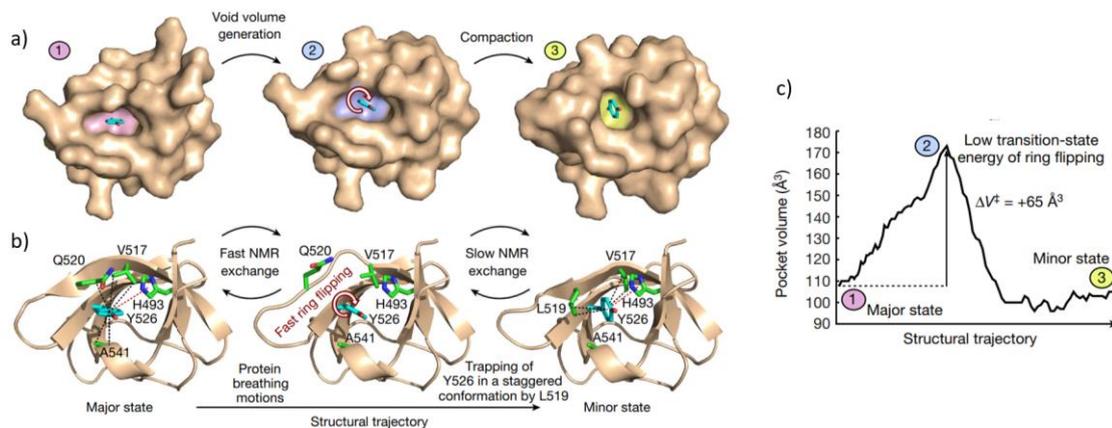


Figure 1.4 : a) Représentation des surfaces du domaine SH3 de JIP1. (1) l'état majoritaire, (2) un état intermédiaire et (3) l'état minoritaire. Le volume de la poche est coloré en rose, bleu et jaune respectivement pour les trois états. b) Structures tridimensionnelles du domaine SH3 de JIP1 représentées en fonction de la trajectoire structurale entre l'état majoritaire (1) et l'état minoritaire (3). Le cycle aromatique de la tyrosine 526 est stabilisé dans une conformation éclipsée en (1) par des interactions CH- π (avec les résidus 517, 520 et 541) et π - π (avec le résidu 493). La rotation du cycle aromatique de la tyrosine 526 est permis par la création d'un volume

vide autour de ce dernier lors du passage de (1) à (2). La rotation est occasionnellement interrompue lorsque le cycle est piégé en conformation décalée (3). Le cycle aromatique de la tyrosine 526 est stabilisé dans une conformation décalée en (3) par des interactions CH- π avec le résidu 519. c) Volume de la poche entourant la thréonine 526 en fonction de la trajectoire structurale entre l'état majoritaire (1), un état intermédiaire (2) et l'état minoritaire (3). Le passage de (1) à (2) entraîne une augmentation du volume de la poche et celui de (2) à (3) une diminution du volume de la poche. Figure extraite de (Mariño Pérez et al. 2022).

1.1.3 Analyse quantitative de la dynamique des protéines

1.1.3.1 Paysage énergétique

L'étude de la dynamique des protéines consiste généralement à caractériser les structures des différents états peuplés par une macromolécule ainsi qu'à déterminer leurs populations respectives et obtenir des informations sur la cinétique d'échange entre les différentes conformations peuplées. Ainsi, la description complète d'une protéine nécessite un paysage énergétique multidimensionnel pour prendre en compte à la fois la variation spatiale et temporelle des mouvements (Henzler-Wildman and Kern 2007).

Pour être complet le paysage énergétique doit comprendre tous les états échantillonnés par le système protéine-solvant et il est à noter que les conditions telles que le solvant utilisé, la pression ou la température du système sont des paramètres importants à définir lorsque l'on décrit un paysage énergétique. En effet, une variation de l'un ou plusieurs de ces paramètres peut avoir une influence sur la population des états ou, par exemple, la cinétique d'échange entre les différentes conformations. Plusieurs échelons constituent un paysage énergétique (Figure 1.5) tel que décrit par Ansari et al. (Ansari et al. 1985) en 1985 pour reporter les différents sous-états de la myoglobine.

On note l'échelon 1 constitué d'un faible nombre d'états séparés par une haute barrière énergétique. Généralement dans cet échelon les transitions sont peu fréquentes et reportent des mouvements collectifs, d'amplitude plus importante que ceux présents dans les échelons 2 et 3, se produisant dans la gamme de temps μ s-ms. Chaque état de l'échelon 1 est lui-même composé de sous-états constituant l'échelon 2. Dans cet échelon, les transitions entre les sous-états sont plus fréquentes et se produisent dans une gamme de temps plus rapide (nanosecondes). Les différents sous-états échantillonnés sont proches structurellement et, de même que l'échelon 1, les sous-états de l'échelon 2 sont composés de sous-états composant l'échelon 3. Dans ce dernier échelon, la dynamique d'échange entre les sous-états est encore plus rapide que dans l'échelon 2 (picosecondes). La rotation d'une chaîne latérale de surface peut

intervenir dans la gamme de temps picosecondes tandis que, par exemple, un mouvement de boucle peut se produire dans la gamme de temps nanosecondes (Henzler-Wildman and Kern 2007). Des sous-états peuvent également être présents dans l'échelon 3 impliquant, par exemple, des vibrations atomiques.

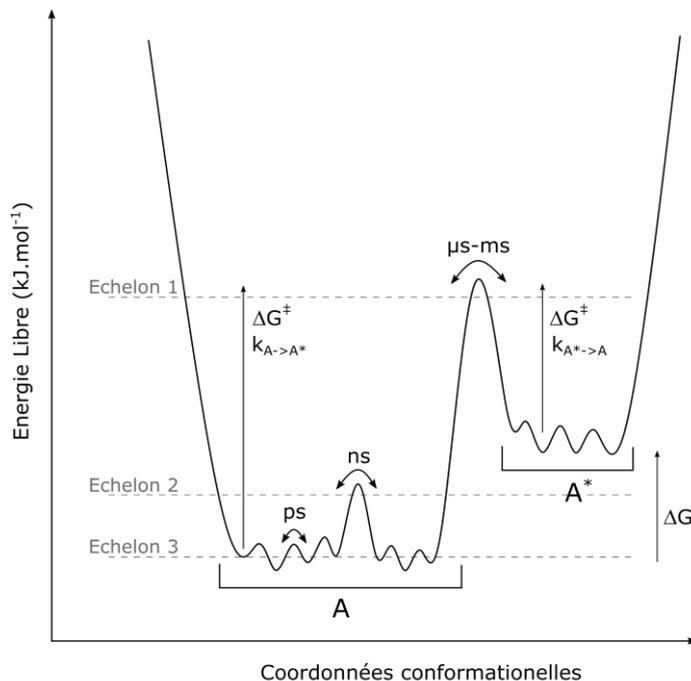


Figure 1.5 : Représentation d'une coupe transversale d'un paysage énergétique représentant l'échange entre deux états A et A^* dans l'échelon 1. Les sous-états sont les minimums de la surface d'énergie, les états de transition sont les maximums entre deux puits. Figure inspirée de (Henzler-Wildman and Kern 2007)

La première notion de paysage énergétique est apparue en 1985 pour la myoglobine. Ansari et al. (Ansari et al. 1985) ont observé que le comportement dynamique de la myoglobine variait en fonction de la température. En étudiant la relaxation de la carboxymyoglobine après dissociation entre 40 et 160 K ils ont pu définir l'existence de quatre échelons de sous-états (Figure 1.6).

Depuis 1985, Ansari et al. et Frauenfelder et al. (Ansari et al. 1987; Frauenfelder et al. 1991) ont poursuivi les études sur la myoglobine et son paysage énergétique. Ils ont montré que MbCO comporte au moins 3 sous-états dans l'échelon 1 puisque le spectre infrarouge de la protéine en complexe avec le monoxyde de carbone montre trois bandes à des fréquences différentes. Chacun des sous-états a la même séquence primaire de protéine mais lie le CO avec une géométrie différente et également une vitesse différente. Un second échelon de sous-états existe pour chaque sous-état de l'échelon 1 puisque la liaison du CO à la myoglobine pour chaque sous-état a une dépendance temporelle non exponentielle et que, de plus, la bande caractéristique de chaque sous-état de l'échelon 1 est non homogènement élargie. Enfin, l'existence de

sous-états dans les échelons 2 et 3 provient d'expériences optiques et de simulations sur ordinateur.

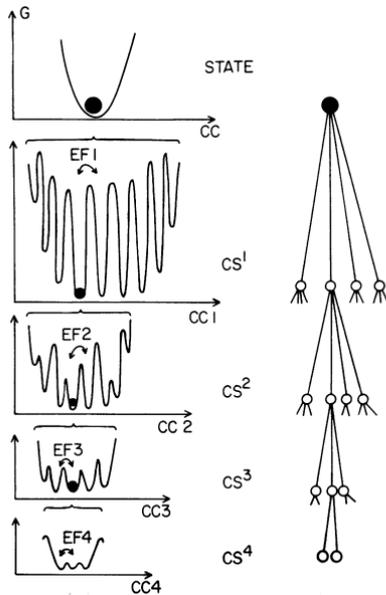


Figure 1.6: Description de l'arrangement hiérarchique des sous-états dans le paysage énergétique de la myoglobine. Figure extraite de (Ansari et al. 1985).

Depuis la première description du paysage énergétique de la myoglobine et l'adoption de la vision dynamique des protéines dans les années 2000, de nombreuses études ont reporté des informations sur le paysage énergétique et la dynamique de protéines variées. On notera par exemple l'étude par Boehr et al. (Boehr et al. 2006) du paysage énergétique dynamique lors de la catalyse par la réductase dihydrofolate. La réductase dihydrofolate catalyse la réduction de 7,8-dihydrofolate (DHF) en 5,6,7,8-tetrahydrofolate (THF) à l'aide du cofacteur nicotinamide adénine dinucléotide phosphate réduit (NADPH) selon un cycle à 5 étapes intermédiaires (Figure 1.7a). Les structures adoptées par l'enzyme aux différentes étapes ont été résolues par cristallographie aux rayons-X. En utilisant la résonance magnétique nucléaire, Boehr et al. ont pu montrer que chaque état intermédiaire du cycle était en échange entre un état majoritaire (dont la structure était déterminée par cristallographie) et un état minoritaire. Des informations sur la structure de l'état faiblement peuplé ont été obtenues et Boehr et al. ont déterminé que l'état excité de chacun des états intermédiaires avait une structure similaire à l'état majoritaire de l'état intermédiaire du cycle le précédant et/ou le suivant (Figure 1.7b). Les constantes de vitesse d'échange conformationnel entre l'état majoritaire et l'état minoritaire des étapes intermédiaires ont été déterminées par les expériences de RMN. La vitesse du cycle à cinq étapes est déterminée par les constantes de vitesse entre les états majoritaires et minoritaires de chaque étape.

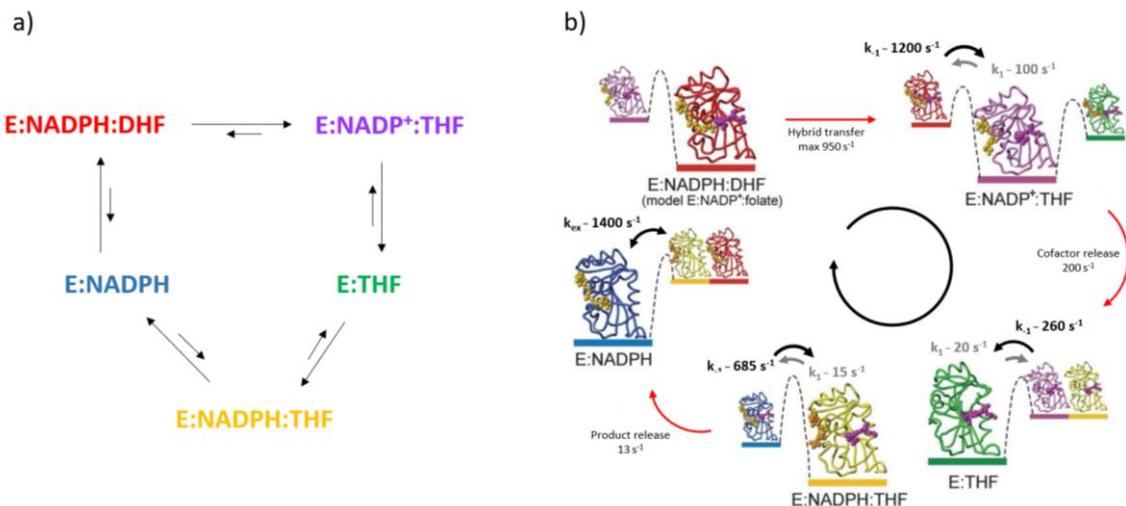


Figure 1.7 : a) Cycle des cinq étapes intermédiaires lors de la catalyse par la réductase dihydrofolate b) Paysage énergétique de la catalyse par la réductase dihydrofolate. Les états majoritaires/de plus basse énergie de chaque étape du cycle sont représentés avec les structures obtenues par cristallographie aux rayons-X et avec des couleurs correspondant à celle du a). Les états minoritaires/de plus haute énergie sont représentés en plus petite taille également avec les couleurs correspondant au a). La structure de l'état faiblement peuplé de chaque état intermédiaire a une structure similaire à l'état majoritaire de l'état intermédiaire le précédant et/ou le suivant. Les constantes de vitesse d'échange conformationnel déterminées par les expériences de RMN sont illustrées par des flèches noires et grises. Les flèches rouges indiquent les vitesses d'échange entre les complexes, mesurées par cinétique enzymatique dans des condition de pré-équilibre. Figure adaptée de (Boehr et al. 2006).

1.1.3.2 Apports potentiels d'une meilleure connaissance de la dynamique des protéines

Mieux appréhender la dynamique de chaque système et notamment avoir des connaissances sur son paysage énergétique apporte de multiples informations. En effet, avec ces données il est possible de visualiser comment la protéine se comporte et bouge d'un puit énergétique à un autre. Additionnellement, caractériser le chemin emprunté et déterminer les structures des états intermédiaires est également faisable. Potentiellement il est aussi possible d'obtenir des informations sur les contributions enthalpiques et entropiques lors du passage de l'une à l'autre des conformations. De plus, comprendre le lien entre structure et fonction protéique ne peut se faire avec une image statique des protéines. Obtenir des informations sur l'amplitude spatiale et temporelle des mouvements des protéines est donc d'un intérêt crucial. Il pourrait être également envisagé grâce à la caractérisation structurale et dynamique du système d'étude de concevoir de nouveaux ligands ciblant des conformations minoritaires de protéines.

1.2 Exemples de techniques expérimentales pour étudier la dynamique des protéines

De très nombreuses techniques existent pour étudier la dynamique des protéines, j'ai choisi de présenter dans cette partie une sélection de quatre techniques expérimentales, très utilisées en biologie structurale : la cryo-microscopie électronique, la cristallographie aux rayons X résolue en temps, le transfert d'énergie par résonance de type Förster et la diffusion de rayons X aux petits angles. Cette dernière méthode étant principalement utilisée pour l'étude de la dynamique des protéines intrinsèquement désordonnées. L'étude de la dynamique par résonance magnétique nucléaire, technique principale utilisée au cours de ma thèse pour étudier la dynamique de ma protéine cible sera détaillée dans la partie 1.3. Etudier la dynamique d'une protéine peut consister à obtenir des informations sur l'ensemble des conformations peuplées dans l'échantillon et à observer des variations structurales mais également à obtenir des informations sur les vitesses de transitions entre les différents états. Selon la technique employée, les informations qu'il sera possible d'obtenir vont différer.

1.2.1 Cryo-microscopie électronique

1.2.1.1 Principe de la microscopie électronique

En microscopie électronique, l'échantillon est irradié par un faisceau d'électrons accéléré à très haute vitesse et les électrons diffusés sont refocalisés par les lentilles du microscope pour obtenir une image de l'échantillon. La principale différence avec la microscopie optique réside dans le pouvoir résolutif des deux méthodes, lui-même lié à la longueur d'onde de la source de radiation. Ainsi les photons émis par la lumière auront une longueur d'onde comprise entre 400 et 750 nm tandis qu'un faisceau d'électron accéléré à 300 kV aura une longueur d'onde d'environ 0.002 nm permettant ainsi d'observer des macromolécules biologiques (Milne et al. 2013). La détermination de la structure de la bactériorhodopsine (petite protéine présente chez les archées fonctionnant comme pompe à protons) avec une résolution proche de 7 Å en 1975 (Henderson and Unwin 1975) a démontré l'utilité de la microscopie électronique en biologie structurale et pavé la voie d'études résolvant à une résolution presque atomique les cartes de densité de cette même protéine ainsi que d'autres protéines membranaires (Henderson et al. 1990; Gonen et al. 2006; Milne et al. 2013). Cependant, on notera que les échantillons biologiques sont sensibles à l'irradiation par un faisceau d'électrons et qu'avec cette technique des dommages tels que la rupture de certaines liaisons chimiques et la création de radicaux libres peuvent se produire, ainsi limiter la dose d'électrons appliquée sur les échantillons est primordiale. Un second désavantage de la technique est son faible rapport signal sur bruit. En effet, les principaux

composants des macromolécules biologiques sont le carbone l'azote et l'oxygène, des éléments avec un faible numéro atomique, qui diffusent plus difficilement les électrons car ils possèdent moins d'électrons autour de leur noyau.

Une technique développée pour pallier ce problème de rapport signal sur bruit pour l'observation des macromolécules biologiques est la microscopie électronique à coloration négative. Cette méthode permet d'obtenir des images basse résolution d'échantillons biologiques en ajoutant à ces derniers des sels d'atomes lourds tels que l'acide phosphotungstique, le molbydate d'ammonium ou l'acétate d'uranyl. Les colorants ajoutés vont permettre d'augmenter le contraste en se liant préférentiellement autour des particules biologiques, ainsi on obtiendra une empreinte négative de la protéine d'intérêt en se focalisant sur les zones de faible diffusion. En règle générale, cette méthode de faible résolution ($\sim 20 \text{ \AA}$) (Ohi et al. 2004) est utilisée pour évaluer les conditions employées et obtenir des informations sur l'homogénéité, la taille, la forme et la dispersion de l'échantillon. Mais comment obtenir avec la microscopie électronique à transmission des images à haute résolution d'un échantillon biologique sans entraîner de dommages sur ce dernier ?

1.2.1.2 Cryo-microscopie électronique (cryo-EM)

Dans les années 1980, Dubochet et al. (Adrian et al. 1984; Dubochet et al. 1988) ont introduit une nouvelle méthode de préparation d'échantillons pour la microscopie électronique consistant à vitrifier les macromolécules biologiques. Dans ce procédé, les protéines sont rapidement congelées dans l'éthane liquide dans le but d'obtenir une glace amorphe autour de l'échantillon afin de le préserver dans son environnement presque natif. Une maille de glace cristalline, cubique ou hexagonale, pourrait conduire à une distorsion de la structure du spécimen d'étude. Afin de préserver le caractère amorphe de la glace l'échantillon doit être conservé à une température inférieure à la température de transition cristalline, le plus souvent dans de l'azote liquide. Il a été reporté que les dommages entraînés par l'irradiation avec un faisceau d'électron était diminué par 6 lorsque la prise d'image s'effectuait à la température de l'azote liquide plutôt qu'à température ambiante (Dubochet et al. 1988; Milne et al. 2013). Ainsi avec la cryo-microscopie électronique, microscopie électronique effectuée à des températures cryogéniques, il serait possible d'augmenter la dose d'électrons appliquée pour parvenir à des dégâts similaires imposés sur l'échantillon (De Rosier and Klug 1968; Unwin and Henderson 1975; Milne et al. 2013). En pratique, en utilisant la cryo-microscopie électronique la dose d'électrons recommandée pour limiter les dommages sur les macromolécules biologiques est de l'ordre de $15 \text{ e}^- / \text{ \AA}^2$ à 200 kV pour une résolution quasi-atomique (Murata and Wolf 2018). Cependant, comme mentionné précédemment un des désavantages de la technique est le faible rapport signal sur bruit

obtenu, et ce également pour les échantillons vitrifiés. Pour résoudre ce problème une moyenne des images acquises avec une faible dose d'électron pour un grand nombre de particules identiques peut être effectué. Ainsi on peut obtenir des images haute résolution d'un échantillon biologique en combinant vitrification de l'échantillon, acquisition d'images à température cryogénique et en additionnant les images acquises sur un grand nombre d'unités analogues (Milne 2013).

La principale technique utilisée pour obtenir des images 3D d'un échantillon biologique en cryo-microscopie électronique est l'analyse de particule unique. Cette méthode repose sur le principe que les particules de l'échantillon sont aléatoirement orientées dans l'échantillon vitrifié. Ainsi, il va être possible de reconstruire l'image 3D de l'échantillon biologique en combinant un grand nombre de projections d'images 2D acquises représentant les différentes orientations des particules identiques (Milne et al. 2013). L'analyse de particule unique permet lors de la collection d'images 2D de limiter la dose d'électrons envoyée sur chaque particule puisque étant identiques et orientées différemment il sera possible de collecter des informations sur l'ensemble rotationnel de la particule sans avoir à utiliser une même particule irradiée de nombreuses fois avec à chaque fois une inclinaison différente. De plus cette technique augmente le rapport signal sur bruit puisqu'un ensemble de données d'analyse de particule unique consiste en milliers d'images et ainsi chaque orientation de la particule est elle-même représentée par plusieurs centaines d'images qui sommées améliorent la sensibilité (Murata and Wolf 2018).

Dans le processus d'analyse les images 2D sont tout d'abord regroupées en « classes » puis utilisées pour construire, avec un algorithme, une carte de densité 3D. Cette dernière est raffinée itérativement jusqu'à ce que sa résolution reste identique (Figure 1.8).

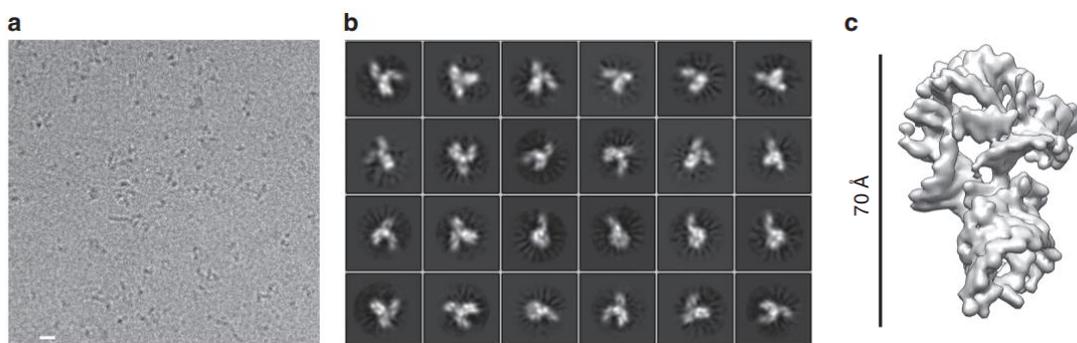


Figure 1.8 : Exemple d'analyse à particule unique de cryo-microscopie électronique du riboswitch de SAM-IV dans sa forme Apo (~40 kDa). a) Micrographe de cryo-EM. La barre d'échelle représente 100 Å. b) Moyenne des classes 2D calculées avec le logiciel RELION. c) Reconstruction finale en 3D de la protéine. Figure extraite de (Zhang et al. 2019).

Ainsi les structures tridimensionnelles de l'adénovirus humain D26 (150 MDa), du ribosome 70S (2.3 MDa), du dimère IDH1 (93 kDa) et de l'hémoglobine (64 kDa) ont été résolues par cryo-microscopie électronique avec des résolutions entre 3.4 Å et 3.8 Å (Merk et al. 2016; Agirrezabala et al. 2017; Yu et al. 2017; Khoshouei et al. 2017). Les récents progrès sur les caméras et notamment de celles à détection directe d'électrons ont contribué à l'obtention de structures tridimensionnelles par cryo-microscopie électronique à haute résolution (Murata and Wolf 2018).

Comme démontré par les exemples cités précédemment, la cryo-microscopie électronique peut être employée pour divers types d'échantillons de la protéine de plusieurs dizaines de kDa aux cellules entières. La taille maximum observable par cette technique est en pratique seulement déterminée par l'épaisseur maximale qui peut être pénétrée par le faisceau d'électron, soit d'environ 0.5 µm en utilisant un microscope électronique à transmission opérant à 300 kV (Murata and Wolf 2018). La taille minimum détectable est quant à elle actuellement située autour de 40 kDa (Zhang et al. 2019). Ainsi des protéines de petite taille (> 40 kDa, taille du record actuel, mais plus généralement > 100 kDa), des particules de virus, de larges assemblages moléculaires, des cellules ou même des bactéries peuvent être étudiées par cette technique.

1.2.1.3 Etude de la dynamique par cryo-microscopie électronique

La cryo-microscopie électronique permet, en plus de l'accès à des structures à résolution quasi atomique, de fournir des informations sur la dynamique des protéines/échantillons étudiés. En effet que ce soit en employant la méthode de l'analyse de particule unique ou celle de cryo-tomographie électronique (qui fonctionne sur un principe similaire mais où l'échantillon doit être incliné à différents angles lors de l'acquisition des images pour pallier le fait qu'il n'est pas possible d'observer les différentes orientations de la particule dans l'échantillon) plusieurs conformations hétérogènes de particules peuvent être identifiées dans les échantillons grâce à des programmes récents de classification 2D et 3D tels que RELION (Scheres 2012; Murata and Wolf 2018). Ces observations de petits changements conformationnels dans un ensemble d'images augmentent avec le développement de nouvelles innovations sur les détecteurs, les filtres d'énergie et les contrastes de phases qui permettent d'améliorer la résolution des structures obtenues. Ainsi, Zhao et al. ont reporté trois structures représentant trois états rotationnels différents de la V-ATPase de levure (Zhao et al. 2015) qui n'avaient jamais été observées auparavant. Les classifications 3D de tailles d'images différentes ont permis d'obtenir des cartes de résolution variant entre 6.9 Å pour la classe comportant le plus d'images (47 %) et 8.3 Å pour la classe comportant le moins d'images (17 %).

Dans une publication plus récente, Loveland et al. (Loveland et al. 2017) ont décrypté, en utilisant la cryo-microscopie électronique à analyse de particule unique, le mécanisme de décodage exact de la traduction génétique. Ils ont pour cela élucidé des structures haute résolution du ribosome procaryotique 70S lié à deux aminoacyl-ARN de transfert (aminoacyl-ARNt). Ces derniers servant à empêcher des erreurs de traduction du code génétique en s'assurant de la correspondance entre l'information portée par l'ARN messenger et l'acide aminé contenu dans la protéine. Dans cette étude les aminoacyl-ARNt, délivrés en complexe ternaire au ribosome avec le facteur d'élongation EF-Tu et la Guanosine TriPhosphate (GTP), différaient. L'un des aminoacyl-ARNt était porteur du même acide aminé que celui codé par l'ARN messenger (ARNm) tandis que le second aminoacyl-ARNt était porteur d'un acide aminé différent de celui codé par l'ARNm (e.g ARNt^{LYS} et ARNm codant pour une arginine).

Pour chacun des deux complexes ribosome-aminoacyl-ARNt, Loveland et al. ont déterminé trois structures à résolution quasi atomique (entre 3.2 et 4.0 Å) et ont ainsi pu révéler des différences dans les étapes de pré-accommodation lorsque le ribosome était lié à l'un ou à l'autre des aminoacyl-ARNt. Ils ont pu ainsi observer que le complexe ribosome avec l'aminoacyl-ARNt porteur du même acide aminé que celui codé par l'ARN messenger présentait différents changements conformationnels entre la première et la troisième structure. Notamment il a été observé que dans la première structure le complexe ternaire (aminoacyl-ARNt + EF-Tu + GTP) se liait à la petite sous unité 30S du ribosome 70S mais il n'est pas constaté d'appariement entre le codon de l'ARNm et l'anticodon de l'aminoacyl-ARNt de plus, le facteur d'élongation n'entre pas en contact avec la grande sous unité 50S du ribosome. Dans la seconde structure, l'appariement entre le codon et l'anticodon s'effectue et additionnellement, le résidu G530 appartenant au domaine de décodage adopte une conformation intermédiaire. Dans la troisième structure l'appariement reste en place, mais le résidu G530 change encore de conformation, entraînant une fermeture du sous domaine 30S ainsi EF-Tu entre en contact avec la sous unité 50S et est capable d'hydrolyser la GTP. A l'inverse, le complexe constitué du ribosome avec l'aminoacyl-ARNt porteur d'un acide aminé différent que celui codé par l'ARN messenger n'induit pas la fermeture du sous domaine 30S par le changement de conformation de G530 et ainsi le facteur d'élongation EF-Tu reste inactif (Figure 1.9).

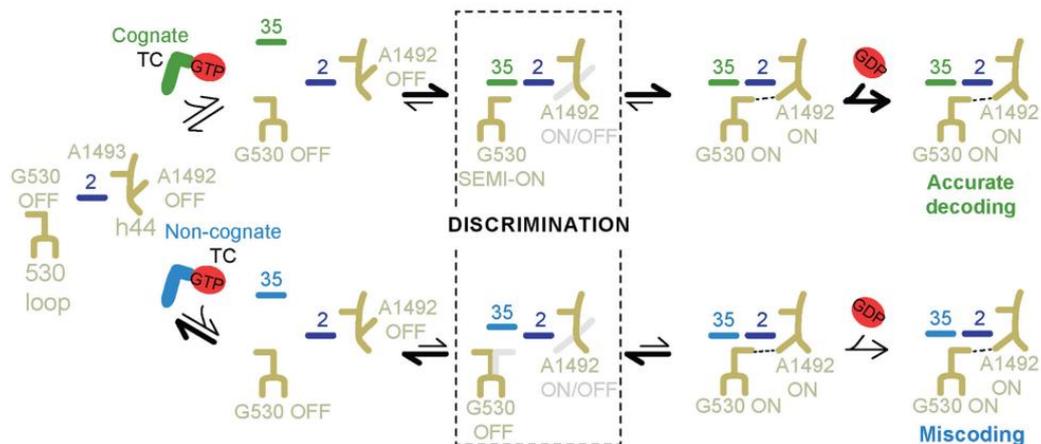


Figure 1.9 : Représentation schématique du mécanisme de décodage de la traduction génétique. Figure extraite de (Loveland et al. 2017).

La cryo-tomographie électronique a également été utilisée pour étudier la dynamique de l'infection par un bactériophage. Le travail de Hu et al. (Hu et al. 2015) dévoile une nouvelle vision du chemin mécanistique d'infection par un phage T4 en révélant des structures intermédiaires clés. Ces données ont été acquises sur des échantillons à différents temps t après infection par le phage. Ainsi, ils reportent qu'avant infection les longues fibres de la queue du phage n'interagissent pas toutes avec la cellule à infecter, la plupart sont repliées le long de la queue du phage. Cependant, au moins une doit être en interaction avec la cellule à infecter afin de promouvoir le réarrangement du socle. Lorsque le socle change de conformation et se lie de façon irréversible à la cellule à infecter, la contraction de la gaine entraîne l'entrée de la queue du phage dans le périplasm de la cellule et l'ADN du phage peut être injecté dans la cellule (Figure 1.10).

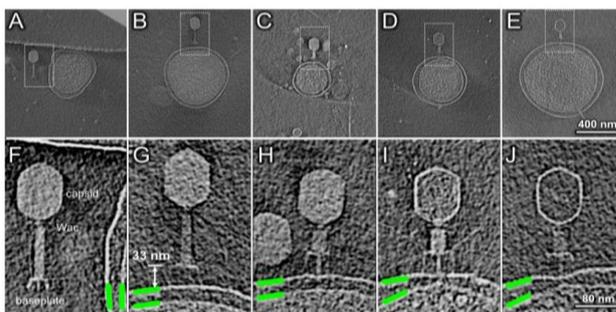


Figure 1.10 : Représentation de différentes conformations du bactériophage pendant la phase d'infection. Tomogrammes 3D capturés 30 s (A, F), 1 min (B, G), 3 min (C, H), 5 min (D, I) et 10 min (E, J) après l'infection. Les panels F à J sont des zooms de leurs panels respectifs A à E. Les membranes internes et externes sont en partie surlignées en vert. Figure extraite de (Hu et al. 2015).

Comme décrit par les études détaillées précédemment il est possible avec la cryo-microscopie électronique d'accéder à différentes conformations peuplées par un échantillon à un temps t . Cela peut se faire soit en piégeant différents états à différents temps t comme montré par Hu et al. (Hu et al. 2015) sur le bactériophage T4 ou en capturant de multiples conformations dans un même échantillon comme décrit par Zhao et al. et Loveland et al. respectivement sur la V-ATPase et le ribosome 70S (Zhao et al. 2015; Loveland et al. 2017).

Par ailleurs, il est également possible avec la cryo-microscopie électronique et l'utilisation de méthodes mathématiques avancées telle que par exemple le « manifold embedding » (Dashti et al. 2014) de suivre un continuum d'état structuraux. Ainsi, il serait possible d'observer le cycle entier d'une machine moléculaire et de représenter son paysage énergétique, si un nombre suffisant d'état conformationnel est présent dans l'échantillon hétérogène. Dashti et al. et Frank et al. ont démontré cela sur le ribosome de levure (Dashti et al. 2014; Frank and Ourmazd 2016). Ils ont utilisé la relation de Boltzmann pour déterminer la valeur de l'énergie libre (ΔG) entre deux conformations en se basant sur le nombre d'apparition de chacune de ces dernières et ont ainsi pu construire le paysage énergétique du ribosome (Figure 1.11).

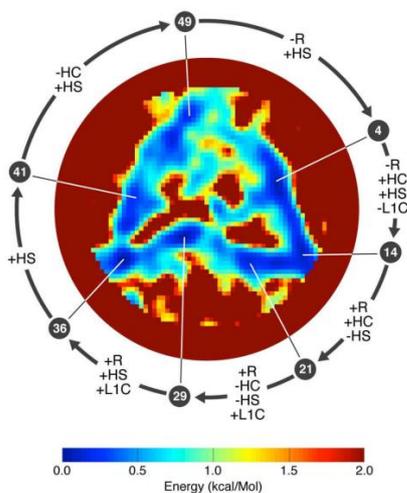


Figure 1.11 : Paysage énergétique du ribosome de levure obtenu par cryo-microscopie électronique. De bleu à marron l'énergie augmente et cette dernière a été tronquée à 2 kcal.mol^{-1} . La trajectoire de plus basse énergie (en bleu) est divisée en 50 états et les flèches indiquent les changements structuraux entre 7 d'entre eux, numérotés. Figure extraite de (Dashti et al. 2014).

Cette méthode également employée sur le récepteur de la ryanodine de type 1 a permis d'obtenir un film nous informant sur les mouvements fonctionnels de la protéine de 565 kDa (Dashti et al. 2020). Toutefois, il n'est pas possible d'obtenir des informations sur les vitesses de transition avec de telles expériences. Ainsi le chemin réactionnel suivi par le ribosome et présenté Figure 1.11 est basé sur une compilation de films montrant les changements conformationnels et reste putatif (Frank 2016, Miller 2021). La cryo-microscopie électronique ne permet pas l'accès à des données cinétiques caractérisant l'échange conformationnel.

Ainsi, la cryo-microscopie électronique permet d'obtenir des informations sur l'ensemble des conformations peuplées dans un échantillon et d'observer à haute résolution les différences structurales des différentes conformations, cependant, elle ne permet pas d'obtenir des informations sur la cinétique de l'échange conformationnel. Toutefois, la cinétique d'une réaction/d'un évènement (comme par exemple l'infection par le bactériophage) peut être suivie temporellement de façon indirecte en capturant à différents temps t la/les conformation(s) peuplée(s) par l'échantillon.

1.2.2 Cristallographie aux rayons X résolue en temps

1.2.2.1 Principe de la cristallographie aux rayons X

En cristallographie aux rayons X, un cristal de protéine purifiée est irradié par un faisceau de rayons X donnant lieu à des motifs de diffraction (Figure 1.12). Ces derniers vont ensuite être analysés afin d'accéder à de nombreuses informations sur le cristal tels que la taille de ses unités de répétition, le système cristallin, et permettre ensuite de calculer une carte de densité électronique qui, utilisée en même temps que la séquence de la protéine va conduire à la structure de la macromolécule cristallisée.

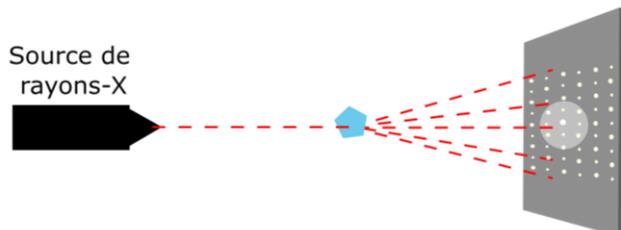


Figure 1.12 : Schéma de principe de la cristallographie aux rayons-X. Le cristal de protéine purifiée (bleu) est irradié par un faisceau focalisé de rayons-X donnant lieu à un motif de diffraction.

L'étape cruciale de la cristallographie aux rayons X réside dans l'obtention d'un cristal de l'échantillon souhaité. En effet, non seulement la protéine doit être soluble et homogène, mais elle doit pouvoir également être hautement concentrée sans précipiter, ni agréger, afin d'induire la nucléation de cristaux. Une quantité considérable de variables joue un rôle important dans cette phase de cristallisation tels que le tampon dans lequel se trouve la protéine, son pH, la nature des précipitants utilisés ainsi que leurs concentrations, la température et la technique de cristallisation employée. Cette phase de nucléation puis de croissance des cristaux est à l'heure actuelle l'étape limitante de la plupart des projets de cristallographie et est souvent la moins bien comprise (Smyth and Martin 2000).

Les rayons X irradiant le cristal peuvent être issus, entre autres, d'électrons accélérés provenant d'anneaux de stockage d'un synchrotron et doivent être focalisés et

collimatés pour générer un faisceau d'électrons parallèles. Plusieurs paramètres doivent être optimisés avant d'obtenir un motif de diffraction de qualité tels que par exemple la distance entre le cristal et le détecteur et, dans les cas idéaux, les motifs de diffractions obtenus sont acquis sur un seul et même cristal. En pratique, l'irradiation du cristal par le faisceau de rayons X entraîne des dommages irréversibles et plusieurs cristaux sont généralement nécessaires. Exposer les cristaux aux rayons X en descendant la température à des températures cryogéniques diminue les dommages et permet à la fois d'améliorer la qualité des données et de diminuer le nombre de cristaux nécessaires pour l'acquisition d'un jeu complet de données (Hope 1990).

L'analyse des motifs de diffraction est mathématiquement complexe mais dans ces derniers est encodée la structure de la protéine. Tout d'abord les dimensions de la maille élémentaire tridimensionnelle (plus petite unité de répétition dans le cristal) peuvent être déterminées à partir de l'espace entre deux tâches dans le motif de diffraction. Il est à noter que plus l'espace entre deux points sera grand plus la dimension de la maille sera petite. La forme de cette maille cristalline permet ensuite d'obtenir des informations sur le système cristallin. Sept systèmes cristallins existent : triclinique, monoclinique, orthorhombique, tétragonal, rhomboédrique, hexagonal et cubique. L'intensité des tâches dans le motif de diffraction apporte quant à elle des informations sur la localisation des atomes dans la maille cristalline (Wlodawer et al. 2008). Chaque tâche diffractée est caractérisée par une phase et une amplitude et tandis que l'amplitude peut être déduite de l'intensité de la tâche, obtenir des informations sur la phase se révèle plus difficile. Différentes méthodes ont été mises en place telles que le remplacement isomorphe et le remplacement moléculaire (Smyth 2000) pour caractériser la phase. La carte de densité électronique peut ensuite être construite (Figure 1.13) et après introduction de la séquence protéique et plusieurs étapes de raffinement une structure tridimensionnelle de la protéine cristallisée peut être obtenue.

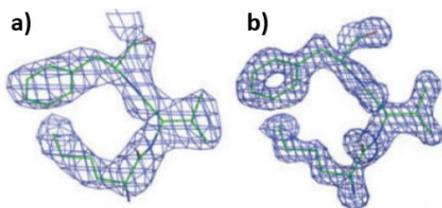


Figure 1.13 : Illustrations de cartes de densité électronique représentées avec une résolution de a) 3 Å et b) 2 Å. Figure extraite de (Wlodawer et al. 2008)

Depuis la première structure tridimensionnelle de myoglobine publiée en 1958 (Kendrew et al. 1958) avec une résolution de 6 Å et la première structure haute résolution (2 Å) du lysozyme par Blake et al. en 1965 (Blake et al. 1965), le domaine a

grandement évolué. Le but ultime des biologistes structuraux est d'observer non plus une structure statique à haute résolution mais une macromolécule en action et ce, toujours à résolution atomique. C'est ainsi que s'est développé la cristallographie aux rayons X résolue en temps.

1.2.2.2 Cristallographie aux rayons X résolue en temps

Les réactions impliquant des macromolécules biologiques induisent fréquemment des changements conformationnels. L'objectif de la cristallographie aux rayons X résolue en temps (TRX) et effectuée à température ambiante est de déterminer le mécanisme cinétique de réactions impliquant des macromolécules tout en déterminant simultanément les structures tridimensionnelles des états intermédiaires peuplés par les macromolécules. Avec cette méthode il est possible d'observer une macromolécule en action et d'obtenir des informations cinétiques sur la réaction et structurales sur les minimums locaux (états intermédiaires) peuplés, contrairement aux méthodes de « capture » chimique et cryogénique qui permettent de stabiliser un état intermédiaire en prolongeant sa durée de vie mais qui ne permettent pas d'avoir accès aux informations cinétiques de la réaction. Par TRX on distingue plusieurs structures intermédiaires indépendantes dans le temps mais dont la concentration est, elle, variable. L'accumulation des états intermédiaires à une concentration supérieure au seuil de détection est nécessaire pour la caractérisation de leur structure et cette augmentation de concentration sera dépendante des différentes constantes de vitesse du mécanisme cinétique de la réaction (Figure 1.14) (Šrajcar and Schmidt 2017).

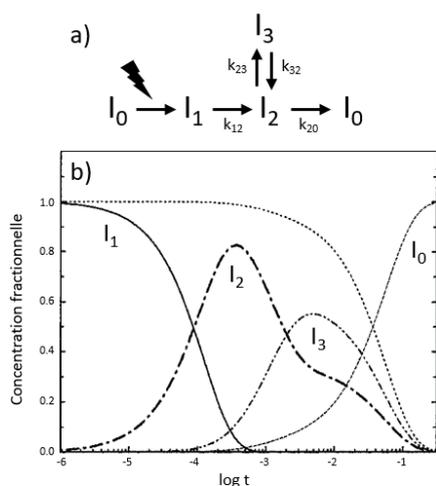


Figure 1.14 : a) Hypothétique réaction réversible photo-initiée. Trois intermédiaires (I_1 , I_2 et I_3) sont présents, I_0 est l'état initial et un mécanisme de cinétique chimique ramifié est décrit avec différentes vitesses d'échange k_{ij} . b) Diagramme de distribution des différents intermédiaires avec $I_0 + I_1 + I_2 + I_3 = 1$ déterminé à partir des constantes de vitesses k_{ij} . Plusieurs intermédiaires sont généralement présents à un temps t . Toutefois, pour qu'un intermédiaire soit observé expérimentalement, ses vitesses de formation et disparition doivent être telles qu'il ait le temps de s'accumuler à un niveau détectable. Adaptée de (Šrajcar and Schmidt 2017)

Afin d'être capable de suivre par TRX une réaction, cette dernière doit non seulement être déclenchée rapidement mais l'exposition aux rayons X pour sonder la réaction doit

également être de courte durée afin, dans les deux cas d'être plus bref que la durée de vie des états intermédiaires. De plus, il est primordial que le déclenchement de la réaction (initiation) se fasse de façon synchrone dans le plus grand nombre de molécules du cristal. Les changements structuraux de la macromolécule (ne devant pas briser le cristal) peuvent ainsi être suivis en enregistrant des motifs de diffractions acquis avec différents délais entre le déclenchement de la réaction et l'irradiation par les rayons X.

A l'origine la principale source de rayons X pour exécuter de la TRX était un synchrotron et la technique s'employait sur un cristal macroscopique. Cependant cette méthode présente plusieurs inconvénients. Tout d'abord, la résolution temporelle de l'expérience est limitée de par la durée de l'impulsion de rayon X (~ 100 ps), la taille des cristaux utilisés impose l'étude de réactions réversibles uniquement et à l'heure actuelle seules des réactions photoinduites sont généralement faisables. Bien que toutefois, la pénétration de la lumière ne soit que superficielle dans le cristal soulevant le problème de l'homogénéité de l'initiation de la réaction (Martin-Garcia 2021).

Le domaine a évolué et dans les années 2000 les lasers de rayons X à électrons libres (XFEL : X-ray Free Electron Laser) sont apparus produisant des rayons X très puissants et très courts (\sim femtoseconde). Ces lasers permettent d'obtenir une brillance d'un ordre de grandeur 9 fois plus important que le plus puissant des synchrotrons actuels (Šrajer and Schmidt 2017). L'utilisation de cristaux macroscopiques avec une source XFEL se révèle compliquée, en effet, la très forte intensité des rayons X délivrée par cette source cause des dommages non réversibles aux cristaux. Cette observation combinée à la large cohérence transversale des XFELs, permettant de focaliser le faisceau de rayon X sans perte de photons pour illuminer des cristaux de petite taille (nm), a pavé la voie de la « cristallographie femtoseconde en série » (SFX). Un flux de petits cristaux (μm -nm), à température ambiante, avec des orientations aléatoires est irradié par la source XFEL et un motif de diffraction est enregistré avant la destruction de chaque cristal (Figure 1.15). La SFX peut donc être employée pour la cristallographie « statique » mais également à la cristallographie résolue en temps (TR-SFX). La première expérience de TR-SFX avec une résolution quasi atomique date de 2014 (Tenboer et al. 2014).

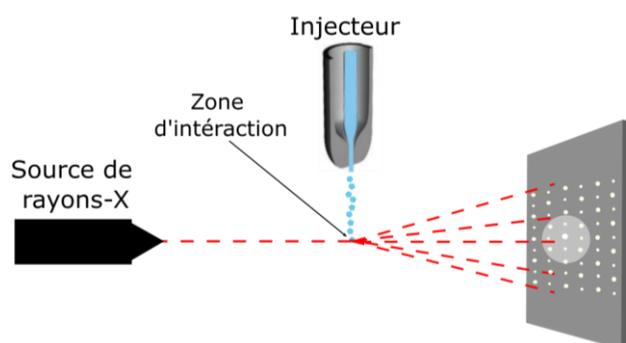


Figure 1.15 : Schéma de principe de la cristallographie aux rayons-X en série. Un flux de petits cristaux de protéine purifiée (bleu) est irradié par un faisceau focalisé de rayons-X donnant lieu à un motif de diffraction.

La TR-SFX est une technique qui présente de nombreux avantages. Tout d'abord, sa résolution temporelle (de l'ordre de la femtoseconde) est bien supérieure à celle que l'on peut obtenir en utilisant un synchrotron comme source de rayons X, de plus, les dommages infligés par les rayons X aux cristaux n'interviennent plus puisque l'on dépend, avec les micro-nano cristaux, du principe de diffraction avant destruction et enfin les cristaux sont, comme mentionné précédemment, de petite taille, ainsi ils ne diffractent qu'une seule fois et sont remplacés rapidement (Tenboer et al. 2014). Dans ce dernier point sont rassemblés plusieurs aspects positifs : le remplacement rapide des cristaux va permettre l'étude de réactions irréversibles par TR-SFX, la petite taille des cristaux va favoriser une illumination plus uniforme de ces derniers lors de réactions photoinduites et ainsi résoudre le problème d'homogénéité de l'initiation de la réaction, de plus une réaction pourra également être déclenchée par diffusion de petites molécules telles que des substrats ou des cofacteurs car le phénomène se produira plus vite dans un petit cristal (temps de diffusion plus petits) permettant de rester plus bref que la durée de vie des intermédiaires (Kupitz et al. 2017). Enfin, les petits cristaux peuvent accommoder de plus larges réarrangements structuraux ainsi que des changements importants dans la maille élémentaire sans détruire le cristal comparé aux cristaux macroscopiques (Kupitz et al. 2014).

Avec le TR-SFX l'étape cruciale ne sera plus d'obtenir un cristal macroscopique (de minimum 0.1 mm dans la dimension la plus longue) mais de faire croître un large nombre de micro/nano cristaux et de les conduire rapidement dans la zone d'intérêt où se situent les rayons X. Plusieurs méthodes qui diffèrent en termes de vitesse et de consommation de protéine ont été reportées. Ainsi, environ 200 mg de protéine seront nécessaires pour acquérir un jeu de données complet si l'on utilise un jet liquide, à l'inverse seulement 1 mg seront nécessaires avec un jet visqueux (Šrajer and Schmidt 2017).

De par la forte demande pour l'utilisation de sources XFEL, la faible disponibilité de ces dernières et les développements continus sur les synchrotrons, la cristallographie en série sur les synchrotrons est apparue (SSX « serial synchrotron crystallography »). De même que pour les sources XFEL, la TR-SSX s'est développée. Actuellement, presque tous les synchrotrons les plus puissants possèdent une ligne qui permet d'effectuer des expériences de SSX, toutefois la résolution temporelle de TR-SSX, comme mentionnée précédemment est limitée à ~100 ps.

1.2.2.3 Etude de la dynamique de la bactériorhodopsine par cristallographie aux rayons X résolue en temps

La bactériorhodopsine, protéine membranaire de 27 kDa présente chez les archées, est une pompe à protons se servant de l'énergie lumineuse pour transporter les protons à travers la membrane cellulaire (Figure 1.16a). Depuis les années 2000, des efforts considérables ont été mis en œuvre pour expliquer le mécanisme de cette pompe à protons, toutefois, des incohérences ont été reportées dans ces travaux (non corrélation des états intermédiaires « capturés » à température cryogénique ou par mutation avec les changements structuraux observés au cours du temps pour la protéine sauvage et la potentielle interprétation erronée des structures des intermédiaires « piégés » due aux dommages causés par le faisceau de rayon X sur les cristaux) (Nango et al. 2016). Depuis, plusieurs études employant la cristallographie aux rayons X en série et résolue en temps, s'affranchissant ainsi du problème des dommages dus aux radiations, ont cherché à élucider ce mécanisme qui restait incompris.

Nango et al. (Nango et al. 2016) ont utilisé la TR-SFX pour étudier les changements conformationnels adoptés par la bactériorhodopsine après photo-activation par un laser sur une gamme de temps allant de la nanoseconde à la milliseconde. Avec ces expériences, ils ont mis en lumière certaines étapes régulant le transport de protons du cytoplasme vers l'extérieur de la cellule. Ainsi, ils ont pu observer que la bactériorhodopsine, protéine membranaire constituée de sept hélices (numérotées de A à G), possède un chromophore rétinol lié de façon covalente par une base de Schiff protonée à la lysine 216 de l'hélice G et que ce dernier s'isomérisé après la photo-activation. Le chromophore rétinol situé au milieu des sept hélices change d'une configuration « trans » pour toutes ses doubles liaisons à une configuration « cis » pour le carbone 13 à l'absorption du photon de lumière visible par son système de double liaisons conjuguées (Figure 1.16b). Cette isomérisation, réorientant la base de Schiff vers le cytoplasme entraîne la rupture de la liaison hydrogène de cette dernière avec la molécule d'eau 402 (Figure 1.16c). Le mouvement du chromophore induit également le mouvement de deux résidus vers le cytoplasme (tryptophane 182 et leucine 93) permettant à une molécule d'eau (n° 452) de s'insérer entre la leucine 93 et la thréonine 89. La liaison hydrogène entre la base de Schiff et la molécule d'eau 452 (elle-même possédant une liaison hydrogène avec la thréonine 89) ou directement avec la thréonine 89 crée un chemin pour le transfert du proton de la base de Schiff vers l'aspartate 85. Différents réarrangements structuraux, notamment de l'hélice C vers la G, permettent d'augmenter le pKa de l'aspartate 85 qui accepte ainsi spontanément le proton de la base de Schiff. Une fois le proton accepté, la liaison hydrogène entre la thréonine 89 et l'aspartate 85 se rompt, la base de Schiff n'est ainsi plus connectée à l'aspartate 85

expliquant que le mécanisme de transfert de protons ne s'opère que dans un sens et ne soit pas réversible (Figure 1.16d).

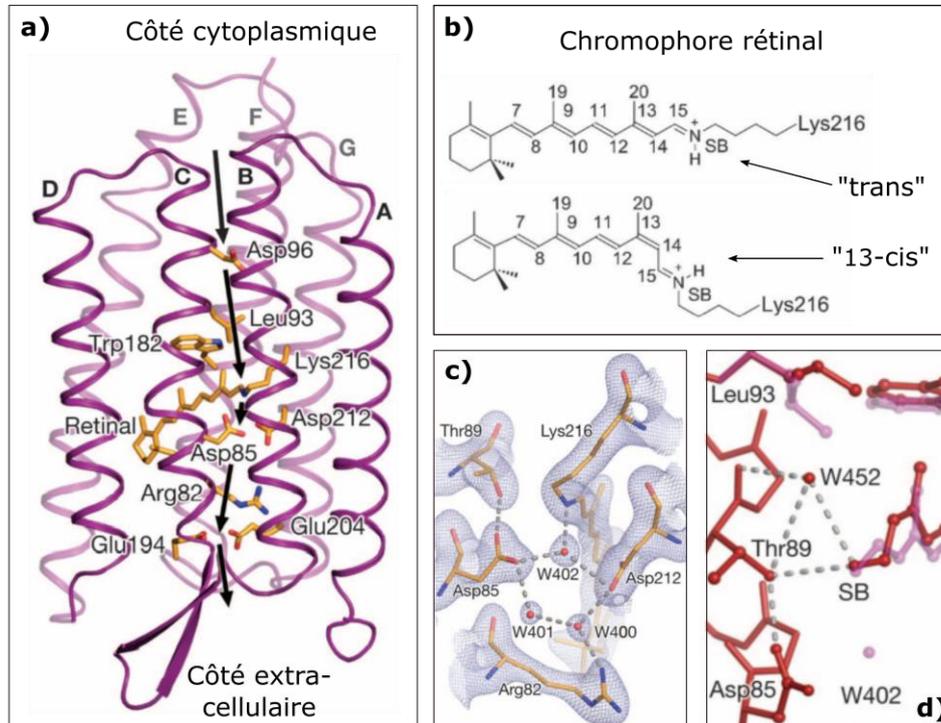


Figure 1.16 : a) Structure tridimensionnelle de la bactériorhodopsine. Les flèches représentent les étapes d'échange de proton effectuées par la pompe à protons. b) Schéma représentant le chromophore rétinol lié de façon covalente par une base de Schiff protonée à la lysine 216 en configuration « trans » ou « 13-cis ». c) Représentation de la densité électronique du site actif de la bactériorhodopsine au repos montrant la localisation de la molécule d'eau n°402. d) Modèle cristallographique montrant l'ensemble des liaisons hydrogènes potentiellement impliquées dans le transfert du proton de la base de Schiff vers l'aspartate 85. Adapté de (Nango et al. 2016).

La TR-SFX a également été utilisée par Nogly et al. (Nogly et al. 2018) pour étudier la bactériorhodopsine sur une gamme de temps plus rapide que Nango et al., allant de la femtoseconde à une dizaine de picosecondes après photo activation. L'objectif étant d'élucider les différentes conformations échantillonnées par le chromophore lors de son passage de sa configuration « trans » pour toutes les doubles liaisons à sa configuration « cis » pour le carbone numéro 13. Avec des structures cristallographiques ayant une résolution de 1.5 Å, ils montrent, étonnamment, que la base de Schiff et la molécule d'eau 402 s'éloignent l'une de l'autre de 0.6 Å après photo-activation et ce, avant même que l'isomérisation du chromophore se produise. Cette distance entre la molécule d'eau et la base de Schiff va à nouveau augmenter et passer de 3.25 Å à 3.65 Å en raison de la redistribution des charges très rapide sur le chromophore après absorption des photons,

phénomène qui est la force motrice de ce mouvement. De par cette polarisation, éloignant encore plus les deux molécules, la liaison hydrogène entre la base de Schiff et la molécule d'eau 402 est rompue. La barrière d'énergie pour passer de la configuration « trans » à « cis » est alors diminuée et l'isomérisation se produit. Cette étude complète celle de Nango et al., puisque ces deux études se focalisent sur des gammes de temps différentes, et permet d'apporter de nouvelles informations sur le début du mécanisme de la pompe à protons. Ainsi, Nango et al. décrivaient une isomérisation du chromophore qui conduisait à la rupture de la liaison hydrogène avec la molécule d'eau 402 et au mouvement de deux résidus, finalement, c'est la rupture de la liaison hydrogène qui conduit à l'isomérisation qui elle-même va conduire au mouvement des deux résidus qui entrainera les étapes suivantes du mécanisme de la pompe à protons (Figure 1.17).

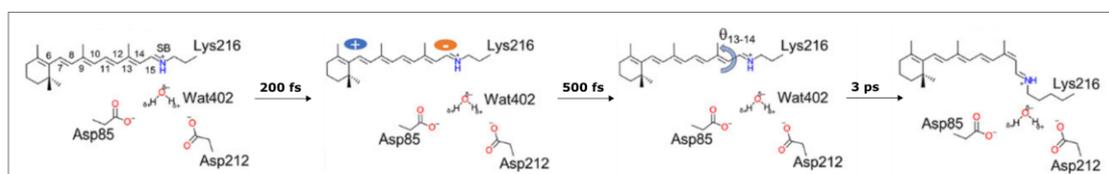
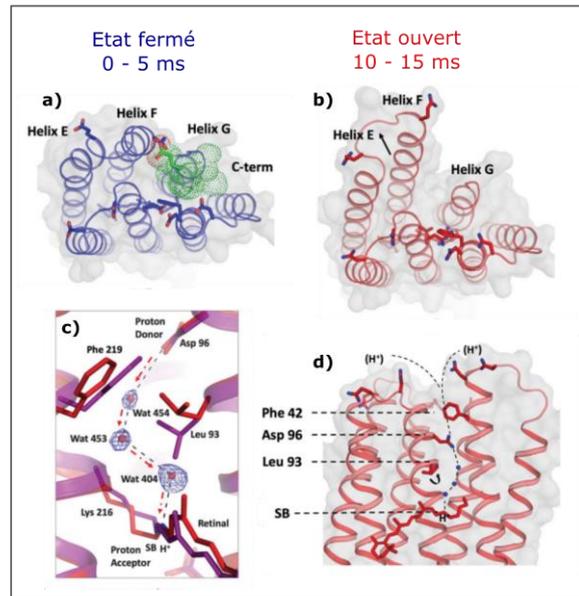


Figure 1.17 : Schéma décrivant les étapes de l'isomérisation du chromophore rétinol. Adapté de (Nogly et al. 2018).

Maintenant que les étapes du transfert de protons par la bactériorhodopsine du cytoplasme vers l'extérieur de la cellule sont connues, la question de la re-protonation de la base de Schiff avec un proton provenant du côté du cytoplasme afin d'initier un nouveau cycle se pose. Obtenir la réponse à cette question permettrait de compléter le cycle du fonctionnement de la pompe à protons qu'est la bactériorhodopsine. Weinert et al. (Weinert et al. 2019) ont étudié ce phénomène en employant la TR-SSX avec une résolution temporelle de l'ordre de la milliseconde. De la même manière que dans les deux études précédentes, la pompe à protons est photo-activée par un laser opérant à une longueur d'onde d'environ 530 nm. La structure de l'état fermé observée de 0 à 5 ms est consistante avec les structures résolues par TR-SFX (Nango et al. 2016; Nogly et al. 2018). Pour les données collectées entre 10 et 15 ms, un état ouvert est caractérisé pour lequel un large réarrangement structural survient du côté cytoplasmique de la bactériorhodopsine comparé à la structure fermée (Figure 1.18a/b). La terminaison des hélices E, F et G est principalement affectée avec un réarrangement structural maximum de 9 Å observé pour le résidu glutamate 166 de l'hélice F. Ces changements ouvrent la voie pour une chaîne de trois molécules d'eau connectant la base de Schiff avec l'aspartate 96 (coté cytoplasmique), donneur de proton intracellulaire. La présence de ces molécules d'eau déstabilise le réseau d'interactions hydrophobiques permettant ainsi le transfert du proton de l'aspartate 96 à la base de Schiff entraînant le remplacement

de cette dernière au centre de la bactériorhodopsine. De petits changements induits par l'isomérisation du carbone 13 « cis » en configuration « trans » par relaxation sont observés. L'aspartate 96 ayant donné son proton est accessible pour une nouvelle protonation (Figure 1.18c/d).

Figure 1.18 : Réarrangement des hélices E, F et G entre l'état fermé (a) et l'état ouvert (b). c) Structure cristallographique de la bactériorhodopsine zoomée sur la zone montrant la chaîne de trois molécules d'eau connectant la base de Schiff avec l'aspartate 96, donneur de proton intracellulaire se formant dans l'état ouvert. d) Comme c) mais montrant une vue plus large de la bactériorhodopsine. Les résidus chargés sont représentés sous formes de tiges. Les pointillés montrent le chemin d'accessibilité pour un nouveau proton. Figure adaptée de (Weinert et al. 2019).



Ainsi la cristallographie en série résolue en temps a permis de caractériser en détail le cycle du transfert de protons de la bactériorhodopsine et ce sur une large gamme de temps de la femtosecondes à plusieurs centaines de millisecondes (Figure 1.19). Ces études montrent qu'effectuée à partir d'une source XFEL ou d'un synchrotron ces expériences sont de grande qualité et complémentaires.

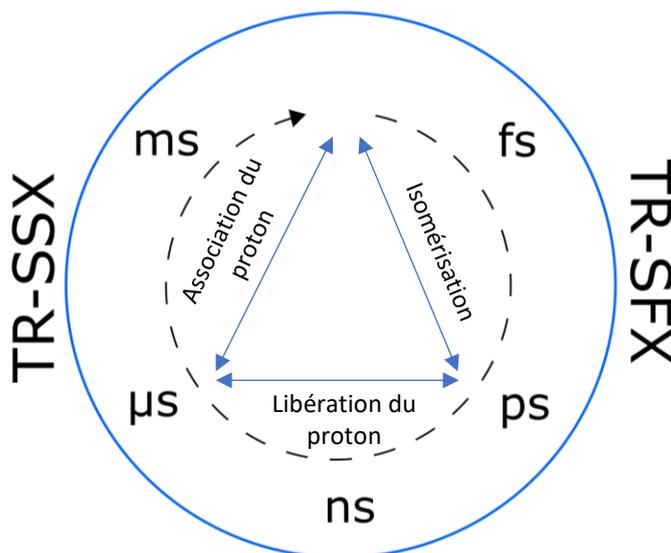


Figure 1.19 : Cycle du pompage de protons de la bactériorhodopsine avec les différents échelles de temps reportées pour les différentes étapes. La complémentarité entre TR-SSX et TR-SFX est également illustrée. Adapté de (Weinert et al. 2019).

Ainsi, la cristallographie résolue en temps donne accès à des informations sur l'ensemble des conformations peuplées dans un échantillon si ces dernières atteignent une population supérieure au seuil de détection. Des structures tridimensionnelles à haute résolution des différents intermédiaires peuvent être obtenues. La cinétique d'une réaction peut être suivie de la femtoseconde à la milliseconde, suivant la source de rayons-X employée, lorsqu'une synchronisation précise entre l'initiation de la réaction et l'irradiation par le faisceau de rayons-X est effectuée.

1.2.3 Transfert d'énergie par résonance de type Förster

1.2.3.1 Principe

Le transfert d'énergie par résonance de type Förster (FRET) est une méthode qui permet de caractériser la distance entre des sondes fluorescentes placées sur une macromolécule. Afin d'étudier une macromolécule avec cette méthode, il est primordial que cette dernière possède deux chromophores : un donneur et un accepteur. Un laser ayant une longueur d'onde spécifique excite le fluorophore donneur qui revient à son état fondamental soit en émettant un photon soit en transférant de façon non radiative son énergie au chromophore accepteur, s'ils sont suffisamment proches (en général < 10 nm). Ce dernier émettra alors également un photon pour retourner à son état initial (Figure 1.20). En comptant le nombre de photons émis par le donneur et par l'accepteur il est possible de remonter à l'efficacité du transfert d'énergie par résonance de type Förster et ainsi de mesurer la distance entre les deux sondes fluorescentes (Schuler 2013).

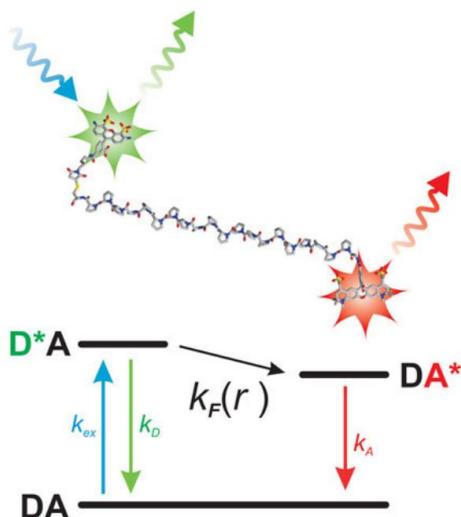


Figure 1.20: Molécule marquée avec un fluorophore donneur (vert) et un fluorophore accepteur (rouge). Le fluorophore donneur est excité spécifiquement et soit émet un photon pour retourner à son état fondamental soit transfère son énergie d'excitation au fluorophore accepteur qui émettra alors un photon pour revenir à son état fondamental. Figure extraite de (Schuler 2013).

Différentes propriétés caractérisent un fluorophore idéal pour une étude par FRET. Tout d'abord il doit posséder un haut rendement quantique (> 0.1) ainsi qu'un haut coefficient d'extinction, être photo stable et soluble en solution aqueuse (Roy et al. 2008). Parmi les fluorophores les plus employés on notera les colorants organiques fluorescents tels que, par exemple, les fluorophores des séries Alexa fluor ou cyanine, toutefois l'utilisation des nanocristaux semiconducteur luminescents ou points quantiques est en croissance. Les protéines fluorescentes quant à elle, bien que très utiles pour l'imagerie par FRET de cellules vivantes, sont généralement de tailles importantes limitant leur utilisation. En effet, un potentiel changement de fonction biologique de la protéine d'intérêt lorsque liée à ces dernières pourrait survenir. Enfin, les acides aminés naturels ne disposent pas, à l'exception du tryptophane, d'un rendement quantique suffisant mais ce dernier n'est pas photo stable. Concernant la paire de sondes fluorescentes donneur/accepteur il est souhaité que les deux sondes disposent d'un rendement quantique similaire mais surtout il est essentiel que les spectres d'absorptions des chromophores donneur et accepteur soient spectralement séparés (la longueur d'onde excitant le donneur ne doit pas exciter en même temps l'accepteur), que le spectre d'émission du fluorophore donneur recouvre en partie celui d'absorption du fluorophore accepteur (afin de permettre le transfert d'énergie) et que le spectre d'émission de l'accepteur ne recouvre pas celui du donneur (pour différencier l'émission par l'une ou l'autre des sources).

A l'origine, les expériences de FRET étaient réalisées sur un ensemble de molécules mais depuis les années 1990 il est possible d'étudier une molécule unique et à l'heure actuelle les expériences de FRET sur molécule unique sont les plus courantes. Deux types d'instrumentation existent : la fluorescence à réflexion interne totale employée pour les molécules immobilisées et les systèmes confocaux qui peuvent s'utiliser à la fois pour les molécules immobilisées et celles diffusant librement (Dimura et al. 2016). Dans les paragraphes suivant je m'attarderai sur ces derniers, plus exploités.

Pour effectuer une expérience de FRET sur particule unique, le signal émis par la molécule unique doit être supérieur à celui du fond contenant beaucoup de molécules de solvant. Une solution pour s'assurer de ce dernier point pour une molécule diffusant librement est de réduire drastiquement le volume de détection, par exemple en focalisant un faisceau laser très étroit sur la zone de détection, la probabilité de trouver deux molécules dans le volume de détection devenant négligeable. En pratique, après excitation de la molécule d'intérêt par le faisceau laser focalisé, un miroir dichroïque sélectionne les photons émis par cette dernière. Des filtres éliminent ensuite la lumière diffusée par les molécules du solvant et un sténopé (petit trou) placé en aval des filtres bloque les flux parasites. Dans un instrument à quatre voies d'acquisition, les photons sont tout d'abord séparés en deux selon leur polarisation (pour obtenir des informations

sur la mobilité rotationnelle des fluorophores) puis selon leur longueur d'onde (pour distinguer les photons émis par le fluorophore donneur de ceux émis par le fluorophore accepteur). Enfin, la lumière est focalisée sur des détecteurs très sensibles. Lorsqu'un photon arrive sur le détecteur il déclenche un signal électronique qui est enregistré (Figure 1.21).

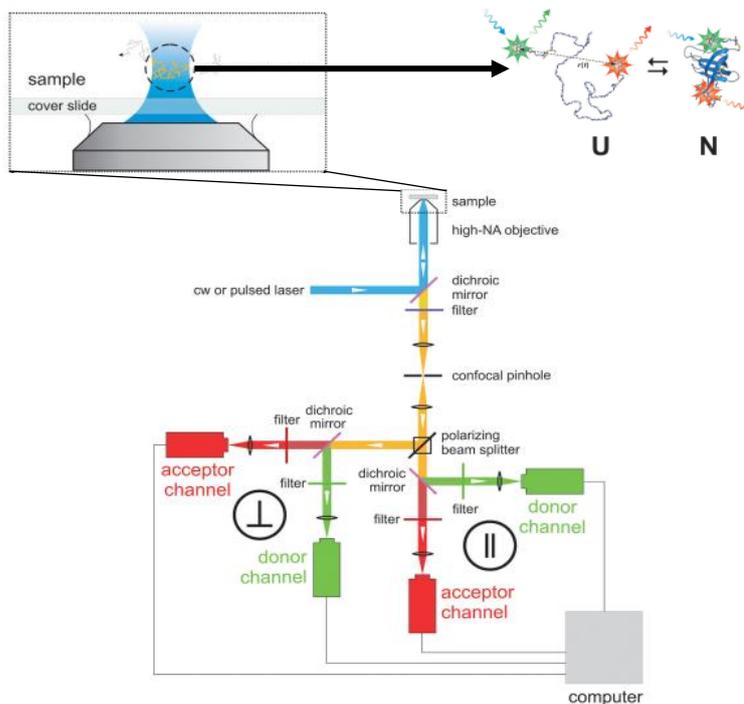


Figure 1.21 : Schéma d'un instrument de sm-FRET à quatre voies d'acquisition. La molécule unique diffuse librement dans le volume de détection. Si la protéine se diffuse dans sa forme native (N), l'efficacité de transfert entre le donneur et l'accepteur sera élevée tandis que si la protéine est dans sa forme non native (U) plus de photons émis par le fluorophore donneur seront détectés. Figure adaptée de (Schuler 2013).

L'analyse de données permet de tracer un graphique du nombre de photons donneurs et accepteurs arrivés dans un intervalle de temps t et d'en déduire l'efficacité du transfert d'énergie par résonance de type Förster ($E = n_A / (n_A + n_D)$) en fonction du nombre de photons accepteur (n_A) et donneur (n_D) détectés dans l'intervalle de temps. En ayant accès à l'efficacité de transfert E il est possible de remonter à la distance entre les deux fluorophores en utilisant l'équation suivante : $E = R_0^6 / (R_0^6 + r^6)$ avec r la distance entre les deux chromophores et R_0 le rayon de Förster (Equation 1.1), rayon définissant la distance donneur-accepteur donnant une efficacité de transfert de 50 % (Schuler 2013).

$$R_0 = (9000 \times \ln(10) \times \kappa^2 \times Q_D \times J) / (128 \times \pi^5 \times n^4 \times N_A) \quad (\text{Equation 1.1})$$

Avec κ une constante dépendant de l'orientation des fluorophores appartenant à l'intervalle [0-4], Q_D le rendement quantique du fluorophore donneur, J l'intégrale du recouvrement des spectre d'émission du donneur et d'absorption de l'accepteur, N_A la

constante d'Avogadro et n l'indice de réfraction du milieu dans lequel la macromolécule d'intérêt est en suspension.

1.2.3.2 FRET et dynamique des protéines

La méthode FRET sur particule unique (sm-FRET) est employée majoritairement pour deux applications. La première pour étudier la structure d'une protéine en solution en multipliant le nombre d'expériences avec des paires de fluorophores donneur/accepteur différentes pour obtenir des contraintes de distance. La seconde, qui nous intéresse particulièrement dans cette section, a pour but d'étudier la dynamique de la protéine d'intérêt en utilisant un nombre de paires de fluorophores donneur/accepteur restreint. Des interactions protéine-protéine, des clivages enzymatiques, le repliement d'une protéine ou des changements conformationnels peuvent être observés. En effet, en utilisant la technique de FRET à particule unique, les particules arrivent individuellement les unes après les autres dans la zone d'illumination du laser, ainsi des sous-états d'une protéine peuvent être détectés. Si l'expérience avait été effectuée sur un ensemble de particules, les sous-états auraient été moyennés et donc potentiellement non détectés. Additionnellement, une fois les sous-états identifiés, il est possible de suivre la cinétique de chaque état individuellement. Il est toutefois à noter qu'une des principales limitations de la technique employant le FRET pour mesurer la dynamique est l'unique mesure (triple mesure) de distance déterminée, due à l'emploi de deux (ou trois) chromophores. Cependant, lorsque des informations de structure à haute résolution sur la macromolécule d'intérêt sont disponibles, ce qui est très fréquemment le cas, il est possible d'interpréter la(es) distance(s) mesurée(s) par des changements conformationnels. (Henzler-Wildman and Kern 2007).

Avec les nombreuses configurations existantes d'appareils de sm-FRET il est possible de suivre la dynamique d'une protéine sur une vaste échelle de temps allant de la picoseconde à quelques secondes. Ces méthodes peuvent être séparées en deux catégories principalement en fonction de la vitesse de la cinétique qu'il leur est possible de couvrir. On notera d'un côté les méthodes qui sont basées sur des particules immobilisées sur une surface pouvant suivre une cinétique allant de la dizaine de millisecondes à la seconde. Ces configurations sont limitées en temps uniquement en raison de la perte de fluorescence des sondes fluorescentes au-delà de quelques secondes. Cependant, la vaste majorité des protéines exercent des mouvements caractérisés par une dynamique beaucoup plus rapide que la gamme de temps milliseconde-seconde. Les méthodes de sm-FRET sur les molécules diffusant librement en solution sont les principales utilisées pour suivre les changements conformationnels rapides allant de la picoseconde à la milliseconde (Mazal and Haran 2019).

1.2.3.3 Exemple d'étude de dynamique par sm-FRET

Diez et al (Diez et al. 2004) ont employé le sm-FRET pour étudier la rotation des sous unités de la FO-F1 ATP-synthase pendant la synthèse et l'hydrolyse de l'ATP. La FO-F1 ATP-synthase est une enzyme qui catalyse la formation d'ATP, chez les bactéries, les mitochondries et les chloroplastes, à partir d'ADP et de phosphate inorganique en convertissant l'énergie électrochimique d'un gradient de protons transmembranaire (Henzler-Wildman and Kern 2007). C'est également une enzyme qui fonctionne dans le sens inverse et peut hydrolyser l'ATP en ADP et phosphate inorganique dans le cas où le potentiel électrochimique est insuffisant (Okuno et al. 2011). La FO-F1 ATP-synthase est une protéine de haut poids moléculaire (~ 500 kDa) très étudiée et constituée de deux sous parties FO et F1. F1 est composée chez *E.coli* en sous unités : $\alpha_3\beta_3\gamma\delta\varepsilon$. La synthèse de l'ATP s'effectue dans la sous partie F1 et plus particulièrement au niveau des trois sous unités β qui peuvent adopter différentes conformations et qui constituent avec les sous unités α et δ le stator de cette machine moléculaire. La sous unité γ constitue le rotor de cette sous partie F1 et est liée à la sous partie FO grâce à la sous unité ε . FO possède également des sous unités ayant un rôle de rotor et est intégrée dans la membrane (Figure 1.22).

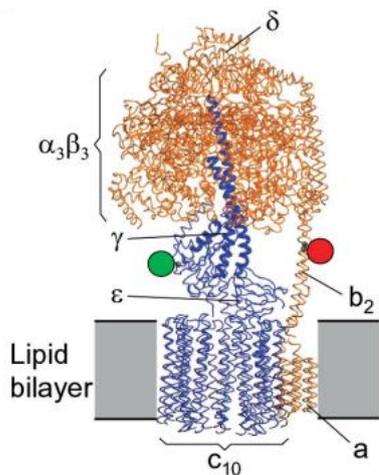


Figure 1.22 : Modèle de la FO-F1 ATP-synthase d'*E.coli*. Le fluorophore donneur (sphère verte) est liée à la sous unité γ de la sous partie F1. Le fluorophore accepteur (sphère rouge) est attaché à une sous unité (b) de la sous partie FO. Les sous unités de type rotor et stator sont représentées respectivement en bleu et en orange. Figure extraite de (Diez et al. 2004).

Il a été exposé qu'une rotation de la sous unité γ s'effectuait lors de l'hydrolyse de l'ATP (Noji et al. 1997), cependant la rotation de cette même sous unité lors de la synthèse d'ATP a seulement pu être démontré de façon indirecte (Zhou et al. 1997). Ainsi, de nombreuses questions restaient non élucidées. En utilisant le sm-FRET sur la FO-F1 ATP-synthase incorporée à un liposome et en attachant un chromophore donneur à la sous unité γ de la sous partie F1 et un chromophore accepteur à une sous unité de type stator de la sous partie FO, Diez et al. (Diez 2004) ont pu suivre en temps réel la rotation

progressive de la sous unité γ entraînée par le gradient de proton. Les photons émis par un liposome unique diffusant librement et traversant le volume confocal ont été enregistrés. Diez et al. ont tout d'abord montré que l'enzyme incorporée au liposome était fonctionnelle puis ils ont étudié la rotation de la sous unité γ . Lors de l'hydrolyse de l'ATP, le rapport de l'intensité du fluorophore donneur sur celle du fluorophore accepteur a permis de détecter la présence de trois états décrivant trois distances différentes entre les fluorophores donneurs et accepteurs. La rotation de la sous unité γ s'effectue en trois étapes de façon anti-horaire lorsque l'on regarde l'enzyme depuis la sous partie F0 (Figure 1.23a/c). A l'inverse, lors de la synthèse de l'ATP, une rotation en trois étapes a également été observée mais cette dernière s'effectuait dans le sens horaire lorsque la F0-F1 ATP-synthase était vue depuis la sous partie F0 (Figure 1.23b/c).

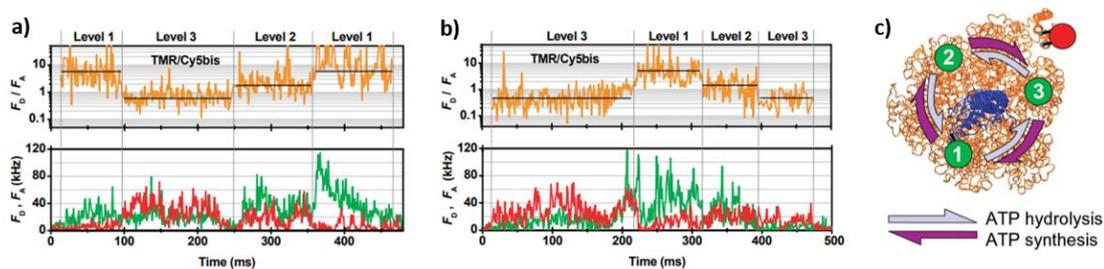


Figure 1.23 : Photons détectés pour un liposome contenant la F0-F1 ATP-synthase lors de (a) l'hydrolyse de l'ATP et (b) la synthèse de l'ATP. En vert sont reportées l'intensité des traces de fluorescence du donneur (FD) et en rouge celle de l'accepteur (FA). En jaune sont reportés les rapports d'intensité FD/FA. Trois niveaux distincts de FRET sont attribués (1, 2, 3), représentant trois conformations différentes. L'ordre de rotation des conformations est (1,3,2,1) dans le cas de l'hydrolyse de l'ATP et (3,1,2,3) dans le cas de la synthèse d'ATP. c) Coupe transversale au niveau du fluorophore donneur par la sous partie F0 de la F0-F1 ATP-synthase présentée Figure 1.22. Le fluorophore donneur (sphère verte) est en position la plus éloignée du fluorophore accepteur (sphère rouge) dans la conformation 1. La rotation est des conformations s'effectue dans le sens antihoraire pendant l'hydrolyse et dans le sens horaire pendant la synthèse d'ATP. Figure adaptée de (Diez et al. 2004).

Une expérience indépendante impliquant l'AMPPNP a également été effectuée et, en présence de cet analogue non hydrolysable de l'ATP, Diez et al. ont pu montrer que les trois conformations de la rotation étaient piégées individuellement et visibles par sm-FRET (Figure 1.24).

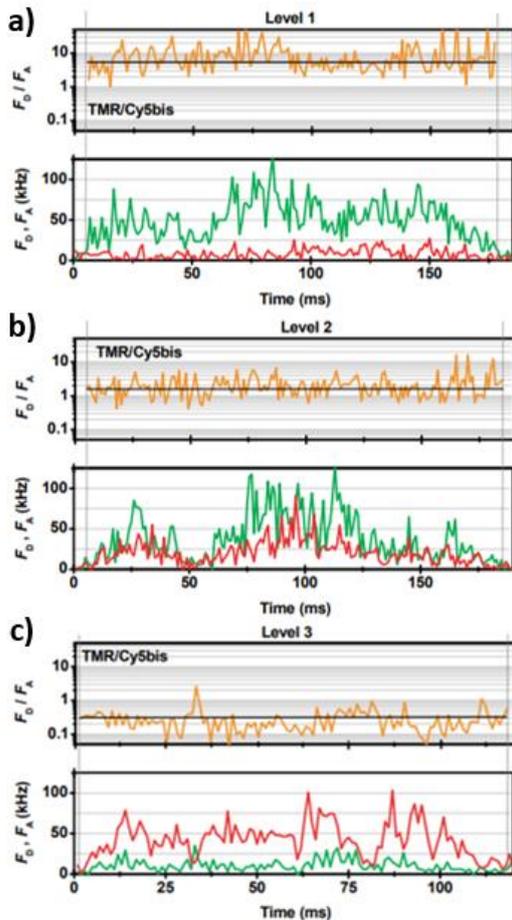


Figure 1.24 : Photons détectés pour trois liposomes diffusant librement individuellement et contenant la FO-F1 ATP-synthase en présence de 1mM d'AMPPNP (a,b,c). En vert est reportée l'intensité des traces de fluorescence du donneur (FD) et en rouge celle de l'accepteur (FA). En jaune est reporté le rapport d'intensité FD/FA montrant un niveau stable et rapportant qu'en présence d'AMPPNP l'enzyme peut être piégée dans chacune des trois conformations de la rotation (a,b,c). Figure extraite de (Diez et al. 2004).

Ainsi, le sm-FRET permet de détecter la présence de sous-états d'une protéine et également d'obtenir des informations cinétiques sur l'échange entre différentes conformations mais cette technique ne permet pas d'avoir accès à l'ensemble des conformations peuplées ni à des informations structurales à haute résolution. Lorsque les structures à haute résolution des molécules d'intérêt sont connues, il est possible d'interpréter la(es) distance(s) mesurée(s) par sm-FRET par des changements conformationnels.

1.2.4 Diffusion de rayons X aux petits angles (SAXS)

1.2.4.1 Principe

La diffusion de rayons X aux petits angles est une méthode permettant d'obtenir des informations globales à la fois sur la forme et la taille des protéines ou complexes étudiés mais il est également possible avec cette technique de caractériser quantitativement des systèmes flexibles tels que les protéines intrinsèquement désordonnées ou les protéines partiellement désordonnées (Kikhney and Svergun 2015).

Un faisceau de rayons X collimaté illumine l'échantillon homogène constitué généralement d'une protéine ou macromolécule en solution. Les intensités des rayons X diffusés aux petits angles par l'échantillon et également par le solvant seul sont enregistrées par un détecteur (Figure 1.25). Ainsi, en effectuant la différence il est possible d'obtenir uniquement le signal des protéines/macromolécules d'intérêt.

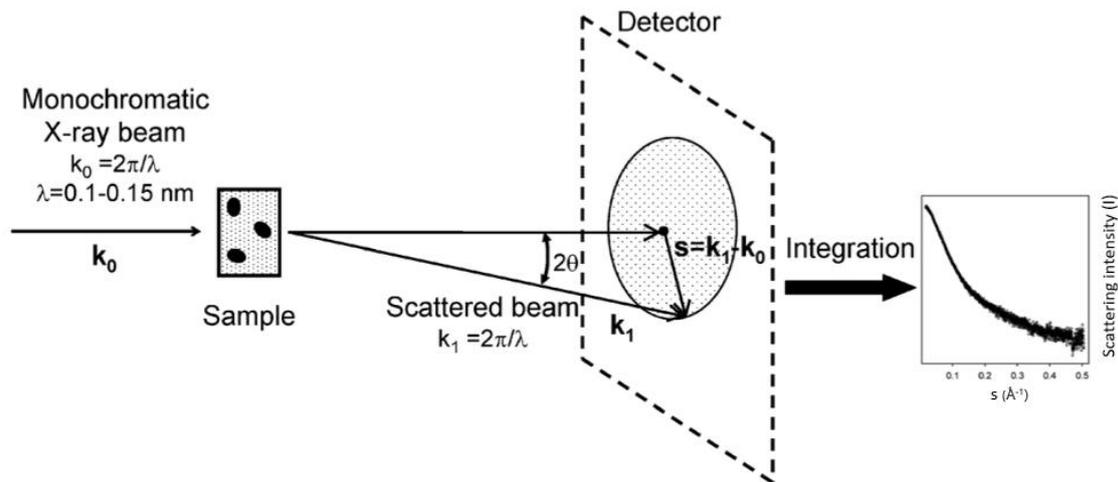


Figure 1.25 : Schéma de principe du SAXS. Figure adaptée de (Bernadó and Svergun 2012).

Le motif de diffusion de la protéine d'intérêt est décrit par l'intensité I en fonction de l'amplitude du vecteur de diffusion s avec $s = 4\pi\sin(\theta)/\lambda$. Avec λ la longueur d'onde de la radiation incidente et θ la moitié de l'angle entre la radiation incidente et celle diffusée (Jacques and Trehwella 2010). L'intensité diffusée par une particule peut être décrite par l'Equation 1.2 :

$$I(s) = \langle \int |(\rho(\vec{r}) - \bar{\rho}_s)e^{i\vec{s}\cdot\vec{r}} d\vec{r}|^2 \rangle \quad (\text{Equation 1.2})$$

Avec $\langle \rangle$ la moyenne sphérique, $\rho(\vec{r})$ la densité électronique de la particule et $\bar{\rho}_s$ la densité électronique du solvant.

Connaissant $I(s)$, il est possible d'obtenir des informations sur la taille globale de l'ensemble des particules en solution en calculant le rayon de gyration R_g . Ce dernier sera plus petit pour un échantillon constitué d'une protéine compacte par rapport à un échantillon composé d'une protéine étendue (les protéines des deux échantillons étant composées du même nombre d'acide aminés). En pratique, on obtient R_g avec l'approximation de Guinier, selon laquelle pour les petits angles l'intensité diffusée

dépend uniquement de deux paramètres (Equation 1.3) (Kikhney and Svergun 2015). Dans le cas de protéine intrinsèquement désordonnées le rayon de gyration correspond au rayon de gyration moyen des différentes conformations échantillonnées.

$$I(s) = I(0)e^{\frac{-s^2 R_g^2}{3}} \quad (\text{Equation 1.3})$$

De $I(s)$, il est également possible de connaître la fonction de distribution des distances $P(r)$ (Equation 1.4) qui est l'histogramme des distances entre toutes les paires d'atomes au sein d'une particule (Kikhney and Svergun 2015).

$$P(r) = \frac{r^2}{2\pi^2} \int_0^\infty \frac{s^2 I(s) \sin(sr)}{sr} ds \quad (\text{Equation 1.4})$$

La fonction de distribution des distances, toujours positive, débute à $r = 0$ pour $P(0)$. La plus grande distance r (nm) pour laquelle la fonction de distribution des distances s'annule est l'extension maximale (D_{\max}) de la particule diffusant les rayons X. Pour r supérieur à D_{\max} , $P(r)$ est égal à 0 (Figure 1.26). $P(r)$, symétrique pour les protéines compacte, est généralement asymétrique pour les protéines intrinsèquement désordonnées. Il est à noter que pour les protéines intrinsèquement désordonnées, la valeur D_{\max} obtenue n'est pas une moyenne pour les différentes conformations échantillonnées mais correspond à l'extension maximale de la conformation la plus étendue (Chan-yao-chong 2019).

Les données obtenues avec la technique de SAXS permettent aussi d'avoir des informations sur la flexibilité du système notamment en utilisant la représentation de Kratky sans dimension. Pour cela, il est nécessaire de tracer $(sR_g)^2 I(s)/I(0)$ en fonction de sR_g . Il est ainsi possible de comparer l'état de repliement de différentes protéines tout en s'affranchissant de leur taille. Ainsi une protéine globulaire sera représentée par une courbe de forme gaussienne avec un maximum atteint à la valeur de $sR_g = \sqrt{3}$. Une protéine désordonnée sera représentée par une courbe augmentant au-delà de la valeur $sR_g = \sqrt{3}$. Avec la représentation de Kratky il est possible de distinguer une protéine intrinsèquement désordonnée d'une protéine partiellement désordonnée constituée en partie de régions ordonnées (Figure 1.27).

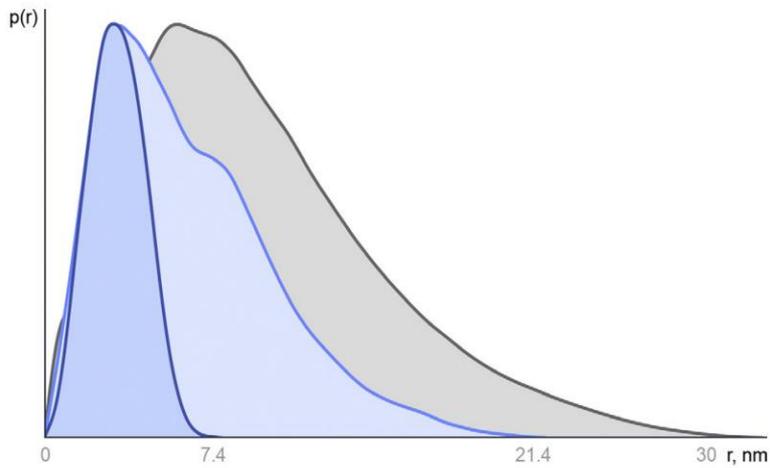


Figure 1.26 : Schéma représentant la fonction de distribution des distances $P(r)$ pour respectivement en bleu, bleu ciel et gris, une protéine globulaire, une protéine partiellement dépliée et une protéine non structurée. Figure extraite de (Kikhney and Svergun 2015).

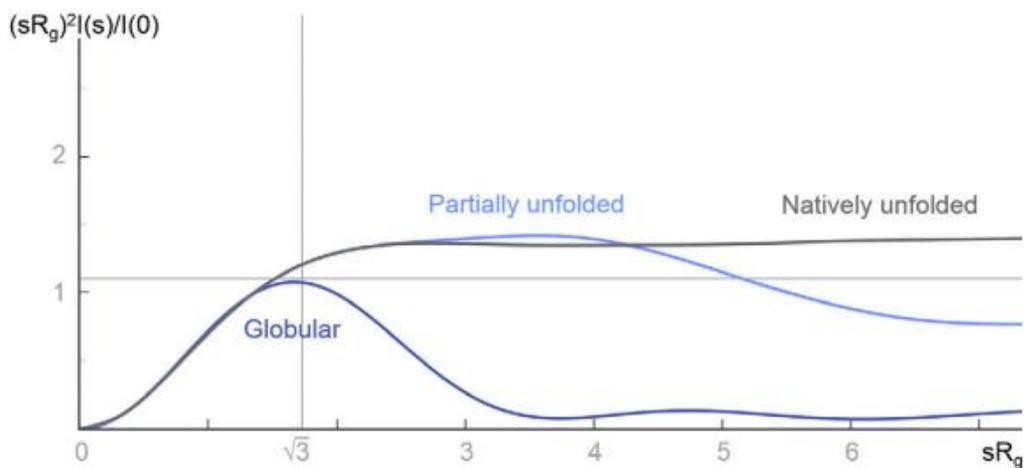


Figure 1.27 : Représentation de Kratky sans dimension pour respectivement en bleu, bleu ciel et gris, une protéine globulaire, une protéine partiellement dépliée et une protéine non structurée. Figure extraite de (Kikhney and Svergun 2015).

1.2.4.2 SAXS et dynamique

Lorsque l'échantillon en solution contient différents types de particules comme c'est le cas pour un échantillon d'une protéine intrinsèquement désordonnée, il peut être plus difficile d'analyser les résultats de SAXS. En effet, le SAXS ne permettra pas d'obtenir une conformation pour chacune des différentes conformations échantillonnées, cependant il pourra être utilisé pour la construction d'un modèle d'ensemble de structures décrivant la protéine intrinsèquement désordonnée étudiée.

Les protéines intrinsèquement désordonnées ou celles composées de multiples domaines liés entre eux par des parties flexibles peuvent être décrites comme des mélanges où chaque conformation de particule contribue au motif de diffusion observé par SAXS, avec l'intensité $I(s)$ définie d'après l'Equation 1.5 comme la somme linéaire de l'intensité de chaque particule (I_k) multipliée par sa fraction volumique (v_k), avec k variant de 1 à K , K étant le nombre de conformations échantillonnées.

$$I(s) = \sum_{k=1}^K v_k I_k(s) \quad (\text{Equation 1.5})$$

Plusieurs approches ont été développées pour caractériser la dynamique des protéines. On notera par exemple EOM en 2007 (Bernadó et al. 2007), rapidement suivi par MES (Pelikan et al. 2009), BSS-SAXS (Yang et al. 2010a) et EROS (Rózycki et al. 2011). Ces méthodes sont basées sur une stratégie commune composée de 2 étapes (Kikhney and Svergun 2015). La première consiste en la génération de milliers de conformères statiques différents de la protéine désordonnée et la seconde cherche à trouver un ensemble réduit parmi les conformères générés permettant de décrire les données expérimentales en employant l'Equation 1.5. La qualité des différents modèles construits est évaluée selon leur capacité à reproduire les données expérimentales obtenues avec l'échantillon de protéine désordonnée et ainsi il est possible d'obtenir des informations sur les différentes conformations échantillonnées par la protéine intrinsèquement désordonnée ou celle partiellement ordonnée.

1.2.4.3 Exemple d'étude de dynamique par SAXS

La protéine kinase R (PKR), kinase jouant un rôle clé dans le processus de défense immunitaire contre les infections virales, a été étudiée avec la technique de diffusion de rayons X (et neutrons) aux petits angles. Il apparaît que PKR est une protéine flexible avec des régions intrinsèquement non structurées qui peut peupler à la fois des conformations compactes comme des conformations étendues.

VanOudenhove et al. (VanOudenhove et al. 2009) ont acquis des expériences de SAXS sur PKR, une protéine constituée de trois domaines repliés, séparés par deux longs domaines flexibles. La fonction de distribution des distances obtenue à partir des rayons X diffusés, hautement asymétrique, avec un maximum atteint autour de 30 Å et un D_{\max} proche de 150 Å, indique que PKR échantillonne des conformations très étendues mais également des conformations plus compactes (Figure 1.28).

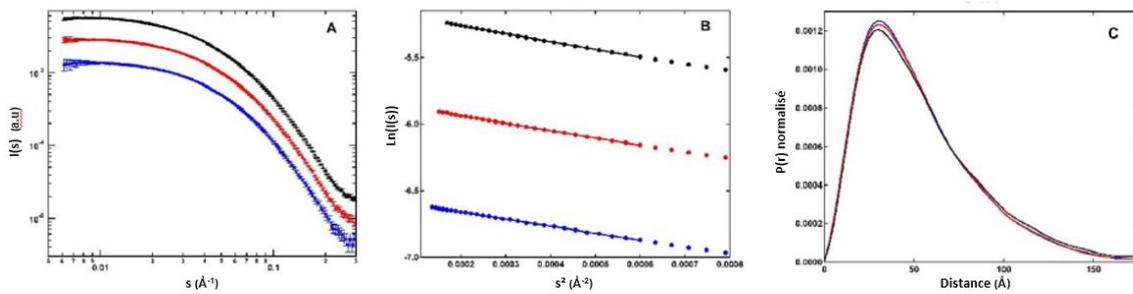


Figure 1.28 : Analyse des expériences de diffusion de rayons X aux petits angles de PKR respectivement à une concentration de 1 mg.ml^{-1} en bleu, 2 mg.ml^{-1} en rouge et 4 mg.ml^{-1} en noir. A) Motifs de diffusions B) Graphiques de Guinier C) Fonctions de distribution des distances. Figure adaptée de (VanOudenhove et al. 2009).

En employant l'approche EOM (Ensemble Optimized Method) (Bernadó et al. 2007) où les domaines repliés ont été considérés comme rigides et où les longs domaines flexibles échantillonnent différentes conformations, VanOudenhove et al. ont obtenu un modèle à 20 conformères qui reproduit presque à l'identique le motif de diffusion de la protéine mesurée expérimentalement par SAXS ($\chi^2 = 0.414$) (Figure 1.29A). La réduction du nombre de conformères dans le modèle entraîne une dégradation de la qualité de l'ajustement. Ainsi, le modèle montre que les domaines flexibles de PKR échantillonnent des conformations à la fois compactes et étendues (Figure 1.29B).

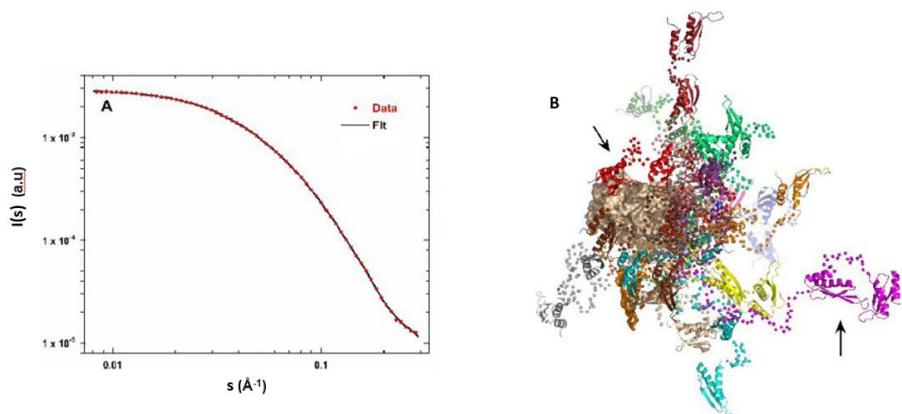


Figure 1.29 : A) Evaluation de la qualité du modèle obtenu avec l'approche EOM. En rouge les points expérimentaux et en noir l'ajustement obtenu avec les 20 conformères. B) Superposition des 20 conformères du modèle, alignés sur le domaine kinase de PKR représenté en beige. Les atomes Ca des domaines non structurés sont représentés par des sphères. Chaque conformation est représentée par une couleur différente. On note des conformations étendues comme montré par la flèche pour le conformère en magenta et des conformations compactes comme montré par la flèche pour le conformère rouge. Figure adaptée de (VanOudenhove et al. 2009).

Comme pour la cryo-microscopie électronique, la diffusion de rayons X aux petits angles ne permet pas d'obtenir des informations sur la cinétique de transition entre les différentes conformations peuplées. Le SAXS ne permettra pas non plus d'obtenir de structures hautement résolues pour l'ensemble des conformations peuplées par l'échantillon. C'est une technique qui, comme la résonance magnétique nucléaire apporte des informations moyennées. La diffusion de rayons X aux petits angles pourra être utilisée pour la construction d'un modèle d'ensemble de structures décrivant une protéine intrinsèquement désordonnée.

1.3 Résonance Magnétique Nucléaire en solution (RMN)

La RMN est une technique qui, contrairement à la cristallographie aux rayons X ou la cryo-microscopie électronique ne mesure pas la disposition des atomes directement dans l'espace mais, mesure la fréquence de précession des noyaux magnétiquement actifs (^1H , ^{13}C , ^{15}N , ^{31}P). Une irradiation à la fréquence de précession des noyaux (de l'ordre du MHz) induit des transitions entre les différents états d'énergie du système de spins, créant ainsi de l'aimantation dans le plan qui peut être enregistrée sous la forme d'un signal électrique. Le signal obtenu en fonction du temps peut ensuite être traité par transformée de Fourier pour obtenir un signal en fonction de la fréquence, un spectre RMN où les fréquences des noyaux actifs apparaissent sous forme de pics. Tous les noyaux appartenant au même élément chimique n'ont pas exactement le même environnement chimique et cela induit des résonances à différentes fréquences sur le spectre. On obtient donc des spectres avec des résonances à différentes fréquences caractérisant chacune le déplacement chimique d'un noyau d'intérêt. Grâce au transfert de l'aimantation à travers les liaisons (via le couplage scalaire) ou à travers l'espace (via l'effet NOE) en employant des séquences d'impulsions spécifiques, il est possible de connecter les atomes d'une molécule entre eux, de caractériser un même site par plusieurs résonances (par exemple avec un spectre 2D HSQC ou 3D HNC0) et d'obtenir des informations sur leur disposition dans l'espace. Ainsi, en utilisant la RMN, des informations sur la structure dans l'espace des molécules/protéines peuvent être obtenues. Un aspect également très important de la RMN est la possibilité d'étudier la dynamique des protéines sur une large gamme de temps en suivant la décroissance du signal mesuré. Avec cette technique il est possible d'obtenir des informations sur la dynamique locale, les vitesses de transition entre différents états peuplés mais également des informations structurales, à résolution atomique, sur l'ensemble des conformations peuplées.

J'ai choisi de présenter dans cette section les notions nécessaires à la compréhension de l'étude des phénomènes dynamiques par RMN et d'introduire différentes méthodes généralement employées pour caractériser la dynamique par RMN. Pour la compréhension des principes de base de la RMN, non détaillés dans cette partie, j'oriente le lecteur vers les livres de Prof. M. Levitt et de Dr. J. Keeler (Levitt 2001; Keeler 2002).

1.3.1 RMN et relaxation

Le processus de relaxation (retour à l'équilibre sur z du vecteur d'aimantation globale, \vec{M} , prenant en compte la contribution individuelle de chacun des spins), est décrit par deux termes appelés relaxation transversale et relaxation longitudinale. La première composante (R_2 (s^{-1})) correspond à la vitesse à laquelle les spins perdent leur cohérence après l'arrêt de l'impulsion de radiofréquence B_1 , autrement dit à la vitesse à laquelle la composante \vec{M}_{xy} de l'aimantation sera revenue à 0 (Equation 1.6). Le second terme, la relaxation longitudinale, correspond à la vitesse du retour des spins à l'équilibre de Boltzmann (R_1 (s^{-1})), soit la vitesse à laquelle la composante de l'aimantation \vec{M}_z sera revenue à son maximum (Equation 1.7) (Figure 1.30).

$$M_{xy}(t) = M_{xy}(0) \cdot e^{-t R_2} \quad (\text{Equation 1.6})$$

$$M_z(t) = M_z(0) \cdot (1 - e^{-t R_1}) \quad (\text{Equation 1.7})$$

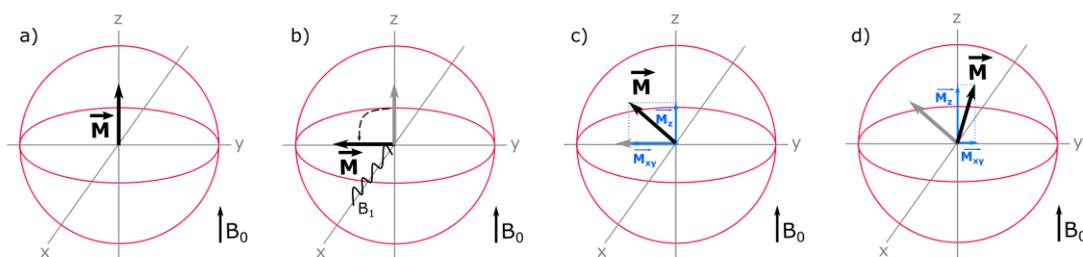


Figure 1.30 : Représentation du retour à l'équilibre sur z de l'aimantation globale \vec{M} . a) Aimantation \vec{M} à l'équilibre. b) Application de l'impulsion de radiofréquence B_1 selon x (impulsion 90°). En conséquence l'aimantation \vec{M} alignée selon z tourne autour de l'axe x et s'oriente dans le plan xy. La composante \vec{M}_{xy} de l'aimantation est alors maximale et la composante \vec{M}_z est nulle. c), d) Lorsque l'impulsion de radiofréquence s'arrête, l'aimantation qui se trouvait dans le plan xy retourne à l'équilibre en tournant autour de l'axe z. La composante \vec{M}_{xy} de l'aimantation diminue et celle \vec{M}_z augmente progressivement.

Pour une petite molécule R_1 sera approximativement égal à R_2 , à l'inverse, pour une protéine de haut poids moléculaire ($\tau_c > 10$ ns) R_2 sera très largement supérieur à R_1 et la relaxation transversale dominera.

Le phénomène de relaxation est causé non pas par les impulsions de radiofréquence qui généralement affectent les spins de la même manière mais par les interactions et les modulations aléatoires des champs magnétiques locaux. Chaque spin possède un moment magnétique qui agit tel un petit aimant générant son propre champ magnétique. Ce dernier aura une orientation et une magnitude qui sera variable dans le temps en fonction de l'agitation thermique. Toutefois, il sera toujours bien moins intense que le champ magnétique appliqué B_1 et décroîtra rapidement avec la distance, affectant possiblement uniquement les autres spins à proximité. Si la composante transverse d'un champ magnétique local oscille à, ou proche de, la fréquence de Larmor d'un spin à proximité, ce dernier sera affecté (i.e tourné) et la magnitude et l'orientation du vecteur d'aimantation globale, \vec{M} , seront affectées. Ainsi la modulation stochastique des champs magnétiques locaux, due aux mouvements de la protéine, est à l'origine de la relaxation longitudinale (impliquant un transfert d'énergie entre les spins et l'environnement autour qui est à l'équilibre). De plus, cette modulation aléatoire des champs magnétiques locaux est également à l'origine de la relaxation transversale avec une contribution non séculaire provenant des composantes transverses des champs locaux oscillant à la fréquence de Larmor d'un spin à proximité et une contribution séculaire due à la distribution des composantes z des champs locaux (Keeler 2002).

Pour un noyau de spin $\frac{1}{2}$, deux sources principales sont responsables de la génération des champs magnétiques locaux : l'interaction dipolaire et l'anisotropie de déplacement chimique. La première, également appelée mécanisme dipôle-dipôle, se produit à travers l'espace et implique deux spins : l'un exerçant le champ magnétique local, le second l'expérimentant. Le champ magnétique local créé par cette interaction dépend à la fois de la distance entre les deux spins, des moments gyromagnétiques des spins et de l'orientation du vecteur liant les deux spins par rapport à celle du champ magnétique B_0 .

L'anisotropie de déplacement chimique s'explique quant à elle par le fait qu'en présence d'un champ magnétique B_0 , les électrons autour des noyaux génèrent un champ magnétique appelé champ induit. Un noyau donné expérimentera donc un champ qui correspond à la somme du champ B_0 et de celui induit, ce dernier variant aléatoirement à tout instant t en fonction des mouvements de la protéine. Ainsi un noyau i donné d'une protéine aura un déplacement chimique anisotropique puisque les protéines constituant l'échantillon auront à un temps t une orientation différente. En pratique en RMN du liquide, la rotation de la molécule est rapide et donc le déplacement chimique

observé sur les spectres pour le noyau i correspond au déplacement chimique moyen (Keeler 2002).

Les mouvements de la protéine, tels que la rotation de cette dernière et les mouvements locaux, modulent aléatoirement les champs locaux, provenant, pour un noyau de spin $\frac{1}{2}$, de l'interaction dipolaire et de l'anisotropie de déplacement chimique, ainsi les mouvements de la protéine sont source de relaxation.

1.3.2 Méthodes pour quantifier la dynamique des protéines par RMN

La RMN est une technique clé pour étudier la dynamique puisqu'elle permet de quantifier cette dernière en utilisant de nombreuses sondes (les noyaux magnétiquement actifs) à résolution atomique, dans des conditions d'équilibre et sur une large gamme de temps allant de la picoseconde à plusieurs dizaines d'heures. Pour cela différents types d'expériences RMN sont utilisés pour suivre la cinétique de décroissance du signal en fonction de la gamme de temps dans laquelle la protéine est dynamique. Qu'il s'agisse de vibrations de liaisons, de mouvements de boucles, de réorientations de chaînes latérales, isomérisations cis-trans, mouvements de domaines ou repliements de protéines : tous les processus dynamiques peuvent affecter trois observables : ω , le déplacement chimique de la résonance, I et λ , respectivement l'intensité et la largeur de raie de cette dernière.

Si le phénomène dynamique expose la sonde RMN à, au minimum, deux environnements chimiques/états différents, on parle alors d'échange chimique pouvant donner lieu à différents spectres RMN en fonction de la vitesse d'échange entre les états. En effet, si l'on considère un échange entre deux états A et B pour un noyau donné, l'apparence du spectre RMN sera affectée par différents paramètres : P_A et P_B , respectivement les populations des deux états, la constante de vitesse d'échange conformationnel entre les deux états, $k_{ex} = k_A + k_B$, et la différence de déplacement chimique entre les deux états $|\Delta\omega| = |\omega_A - \omega_B|$. Trois principales gammes de régimes d'échange sont dénotées : le régime d'échange rapide où $k_{ex} \gg |\Delta\omega|$, celui intermédiaire où $k_{ex} \approx |\Delta\omega|$ et celui lent où $k_{ex} \ll |\Delta\omega|$ (Kleckner and Foster 2011). Dans le régime d'échange lent, les signaux des deux états sont observés (Figure 1.31). Ainsi pour un résidu en échange lent entre deux états A et B, on observera deux pics l'un à la fréquence ω_A et le second à la fréquence ω_B . L'intensité I de chacun des deux pics sera proportionnelle à la population de chaque état et la largeur de raie λ , définie par deux fois la valeur de la relaxation transversale observée (R_2^{eff}), de chaque pic sera influencée par la vitesse de départ de cet état (k_A) (Equation 1.8).

$$\lambda_A = 2 R_2^{eff}(A) = 2 (R_2^0 + k_A) \quad (\text{Equation 1.8})$$

Avec λ_A la largeur de raie de la raie au déplacement chimique ω_A , R_2^{eff} la vitesse de relaxation transversale observée, R_2^0 la vitesse de relaxation transversale dans le cas où il n'y a pas d'échange chimique (i.e influencée uniquement par l'interaction dipolaire et l'anisotropie de déplacement chimique) et k_A la constante de vitesse de l'état A.

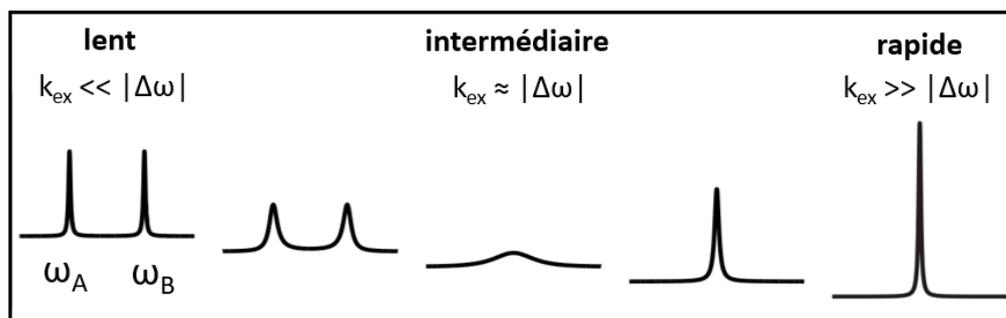


Figure 1.31 : Impact de l'échange chimique sur les raies observées par RMN en présence d'un échange entre deux états : A et B, avec dans cet exemple $P_A = P_B$. Dans le régime d'échange lent, à gauche, deux signaux sont observés aux fréquences des deux états ω_A et ω_B . Dans le régime intermédiaire, un unique pic est observé à un déplacement chimique entre ω_A et ω_B mais principalement la largeur de raie est grandement affectée par l'échange entre les deux états. Dans le cas du régime d'échange rapide, un seul signal est observé dont le déplacement chimique est proportionnel à la population et au déplacement chimique de chaque état.

Dans le cas du régime d'échange rapide, un seul signal est observé dont le déplacement chimique est proportionnel à la population et au déplacement chimique de chaque état d'après l'Equation 1.9 (Figure 1.31). La vitesse de relaxation transversale est définie par l'Equation 1.10.

$$\omega = \omega_A P_A + \omega_B P_B \quad (\text{Equation 1.9})$$

$$R_2^{eff} = P_A R_{2A} + P_B R_{2B} \quad (\text{Equation 1.10})$$

Avec R_2^{eff} la vitesse de relaxation transversale observée et R_{2A} et R_{2B} les vitesses de relaxation transversales de A et B respectivement.

Dans le cas du régime intermédiaire, un unique pic est observé à un déplacement chimique entre ω_A et ω_B mais, principalement, la largeur de raie est grandement affectée par l'échange entre les deux états (Equation 1.11) (Figure 1.31).

$$R_2^{eff} = R_2^0 + R_{ex} \quad (\text{Equation 1.11})$$

Avec R_2^{eff} la vitesse de relaxation transversale observée, R_2^0 la vitesse de relaxation transversale dans le cas où il n'y a pas d'échange chimique (i.e. influencée uniquement par l'interaction dipolaire et l'anisotropie de déplacement chimique) et R_{ex} la contribution de l'échange à la relaxation transversale.

Différentes méthodes de RMN peuvent ici être introduites pour obtenir des informations directes sur les phénomènes dynamiques se produisant dans les différentes gammes de temps. Les six méthodes détaillées ci-dessous couvrent quasiment toute l'échelle des temps allant de la picoseconde à plusieurs dizaines d'heures (Figure 1.32) et sont ordonnées en commençant par la méthode permettant d'accéder aux phénomènes les plus lents jusqu'à celle donnant accès à des informations sur les phénomènes les plus rapides. Pour chaque méthode un exemple d'utilisation, décrit dans la littérature, est reporté. Les couplages résiduels dipolaires ne permettant pas d'obtenir des informations sur le temps caractéristique du mouvement mais uniquement des informations sur le paramètre d'ordre ne seront pas détaillés dans la partie ci-dessous.

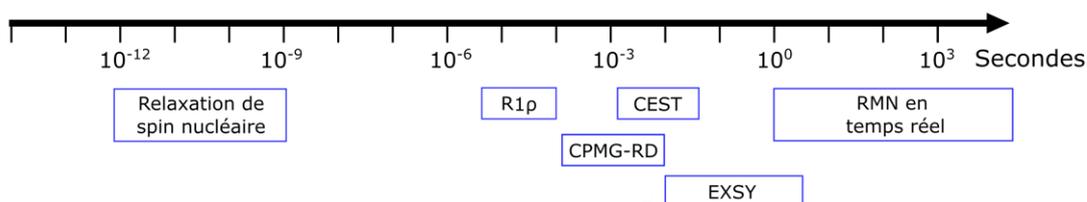


Figure 1.32 : Représentation graphique des méthodes de RMN permettant d'obtenir des informations directes sur les phénomènes dynamiques à différentes échelles de temps.

1.3.2.1 RMN en temps réel ($\tau_{ex} > 1 \text{ s} / k_{ex} < 1 \text{ s}^{-1}$)

La RMN en temps réel est employée pour des systèmes dont la constante de vitesse d'échange conformationnel est inférieure à 1 s^{-1} . Il s'agit d'une approche dont l'objectif est de suivre l'évolution de l'intensité des signaux RMN en fonction du temps dans une gamme de temps pouvant aller de la seconde à plusieurs dizaines d'heures. En pratique, le processus que l'on souhaite suivre est déclenché, par exemple par changement de tampon (pH, sels, ajout de cofacteur), saut de pression, induction lumineuse ou saut de température, puis des spectres RMN sont acquis rapidement pour suivre le processus.

A l'origine la RMN en temps réel était limitée à des expériences 1D et/ou à des processus dont la cinétique était relativement lente, de l'ordre de plusieurs minutes, puisqu'un spectre devait être acquis pour chaque point temporel (Kleckner and Foster 2011). Cependant, le développement de séquences d'impulsions telles que la SOFAST (band-Selective Optimized Flip Angle Short Transient) et la BEST (Band-selective Excitation Short-Transient) (Schanda and Brutscher 2005; Schanda et al. 2005; Amero et al. 2009), l'introduction de marquages isotopiques adaptés et l'amélioration matérielle telle que le développement de sondes cryogéniques et des aimants à haut champ ont permis de réduire drastiquement le temps d'acquisition des spectres 2D. Ainsi la RMN en temps réel peut, de nos jours, être suivie avec des expériences 2D qui permettent d'avoir accès à une bien meilleure résolution que les expériences 1D. Généralement les signaux d'une espèce présente à l'initiation de la réaction vont peu à peu disparaître et de nouvelles résonances caractéristiques d'une nouvelle espèce apparaissent au cours du temps. Il est possible de quantifier le changement de population et, en ajustant ces données, si l'on dispose de plusieurs sondes percevant cet échange, de voir si elles perçoivent toutes le même processus d'échange.

Exemple

Macek et al. (Macek et al. 2017) ont étudié par RMN en temps réel le processus d'oligomérisation de TET2 de *Pyrococcus horikoshii*. TET2 est un homo-dodécamère de 0.5 MDa, constitué de 12 sous-unités, qui participe au maintien de l'équilibre cellulaire en hydrolysant de courts peptides en acides aminés. En combinant haute température, marquage isotopique sélectif sur les groupements méthyles des alanines, dont le déplacement chimique fournit des informations sur la structure secondaire, et expériences de RMN optimisées afin d'acquérir en série des spectres RMN 2D, Macek et al. ont suivi en temps réel l'autoassemblage de TET2 avec une résolution temporelle de 15 secondes. TET2 a tout d'abord été stabilisé sous forme monomérique repliée à pH acide puis l'autoassemblage a été initié par un rapide saut de pH (de pH 4 à pH 8). Le premier spectre après initiation du processus d'assemblage a montré une résonance majoritaire très large, localisée au centre du spectre, caractéristique d'une protéine non structurée. L'intensité de la résonance au cours du temps est représentative d'un état intermédiaire, équivalente à zéro avant le saut de pH, maximum lors de l'acquisition du premier spectre après le saut de pH, puis diminuant au cours du temps. Un ensemble de résonances supplémentaires, en partie superposées avec le premier intermédiaire, est apparu dans la série de spectres RMN. Les résonances étaient cette fois-ci bien dispersées et localisées proches des signaux caractéristiques du dodécamère, indiquant la présence d'une (ou plusieurs) sous-espèce(s) repliée(s) en solution. Le suivi par RMN de l'intensité de ces résonances a montré une augmentation (jusqu'à 120 s) puis une diminution de leur intensité et a décelé la présence de plusieurs intermédiaires. Ces

derniers ont été observés par microscopie électronique à coloration négative et les clichés ont révélé des particules structurées mais avec une hétérogénéité en termes de taille et de forme de particules. Macek et al. ont ensuite montré que les oligomères s'assemblaient pour former le dodécamère (Figure 1.33).

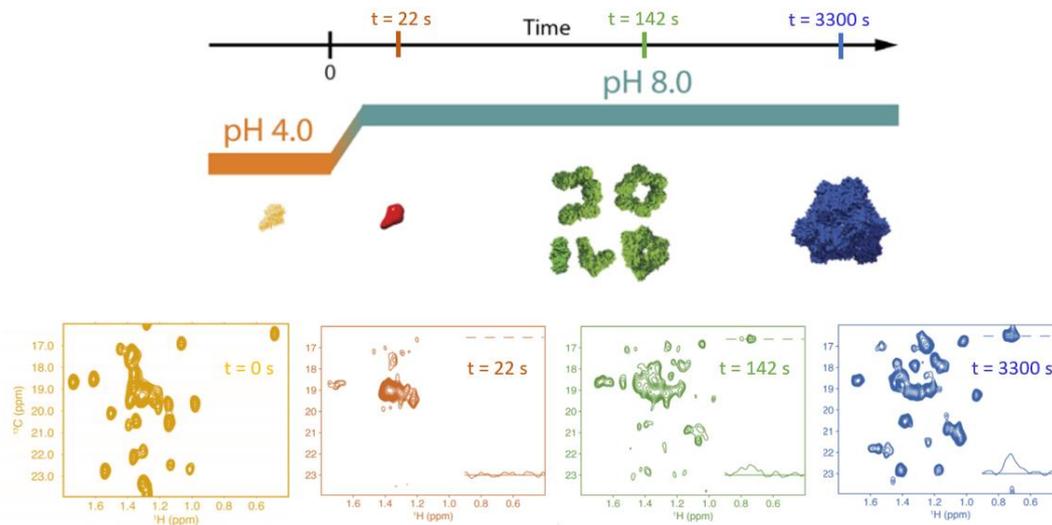


Figure 1.33 : Représentation schématique de la réaction d'autoassemblage de TET2 suivie par RMN en temps réel. Le spectre orange correspond au spectre des groupements méthyles des alanines du monomère replié de TET2 à pH acide. Le spectre rouge correspond au large et intense pic du monomère flexible obtenu pour TET2 juste après le saut de pH. Le spectre vert correspond au spectre des groupements méthyles des alanines des états intermédiaires de TET2 : les dimères/oligomères. On note toujours la présence mais moins intense du pic correspondant au monomère flexible. Le spectre bleu correspond au spectre des groupements méthyles des alanines de TET2 autoassemblé (acquis 55 minutes après initiation de l'oligomérisation). Figure adaptée de (Macek et al. 2017).

Des données cinétiques sur les différents intermédiaires éphémères ainsi que sur la formation de la particule native finale ont été obtenues par RMN en temps réel. L'étape limitante de l'autoassemblage semble être le passage du monomère non replié à un monomère replié, correspondant à la décroissance de l'intensité du signal du monomère déplié. Les monomères repliés à pH neutre exposent un grand nombre de résidus chargés au niveau de l'interface de dimérisation amenant à la formation de dimères. Très rapidement les dimères s'associent de façon stochastique par les interfaces de trimérisation et forment des oligomères de taille et de forme variables, jusqu'à ce qu'il n'y ait plus d'interfaces disponibles. Ainsi, de multiples chemins d'assemblages conduisent à la formation de TET2 dodécamérique (Figure 1.34).

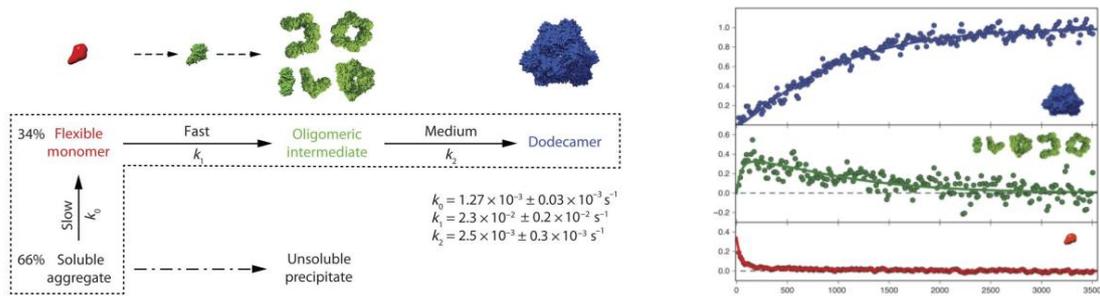


Figure 1.34 : A gauche, modèle cinétique de l'autoassemblage de TET2 initié par un saut de pH. Le saut de pH de 4 à 8 entraîne la conversion du monomère de TET2 replié en monomère flexible, non structuré. Le monomère replié à pH neutre est un intermédiaire nécessaire avant l'obtention d'oligomères éphémères de forme et de taille variables. Finalement, le dodécamère est formé. A droite, ajustement des courbes correspondant à l'intensité des pics mesurés et caractéristiques de chaque conformation. En rouge on retrouve les informations correspondant au monomère non replié, en vert, courbe caractéristique des différents oligomères et en bleu celle du dodécamère. Figure extraite de (Macek et al. 2017).

1.3.2.2 EXSY (Exchange Spectroscopy) ($\tau_{\text{ex}} \sim 10\text{-}5000 \text{ ms}$ / $k_{\text{ex}} \sim 0.2\text{-}100 \text{ s}^{-1}$)

L'EXSY est une méthode appropriée pour étudier un processus dynamique se produisant dans le régime d'échange lent avec des résonances observables pour les deux états : celui majoritaire (A) et celui minoritaire (B). Pour employer cette technique, la constante de vitesse d'échange conformationnel entre les deux états doit être comprise entre 0.2 et 100 s^{-1} . L'EXSY 2D repose sur une séquence d'impulsions semblable à celle de l'expérience NOESY 2D. Ainsi, un spectre 2D sans temps de mélange est acquis, montrant deux pics (l'un correspondant à l'état majoritaire AA et le second à l'état minoritaire BB) puis différents spectres 2D sont acquis avec différents temps de mélange. Dans chacune des expériences avec temps de mélange, à la fin de la période de préparation chaque sonde de la molécule peut être « étiquetée » avec son déplacement chimique pour identifier son état (AA ou BB). Lors de la période d'évolution l'aimantation est stockée sur z pendant le délai T (temps de mélange EXSY) au cours duquel l'interconversion entre les états A et B continue. Ainsi des sondes initialement dans l'état A pourront être passées, à la fin de la période d'évolution, dans l'état B et vice-versa. Après la période de détection et la transformation de Fourier il est possible d'observer cette interconversion entre les états sur le spectre 2D. En effet, si un échange s'est produit entre les deux états, on observera non plus deux pics correspondant aux états AA et BB en absence de temps de mélange mais quatre pics. 2 seront localisés aux mêmes endroits que AA et BB et correspondront aux sondes qui initialement étaient dans un état et sont restées dans cet état à la fin de la période d'évolution. Les deux

autres pics (AB et BA) correspondront aux sondes qui ont interchangeé d'état pendant la période d'évolution (Figure 1.35). Pour éviter la détection des pics de corrélation NOE, l'aimantation de l'hétéroatome (dans la figure suivante : ^{15}N) est à stocker sur z plutôt que l'aimantation du proton.

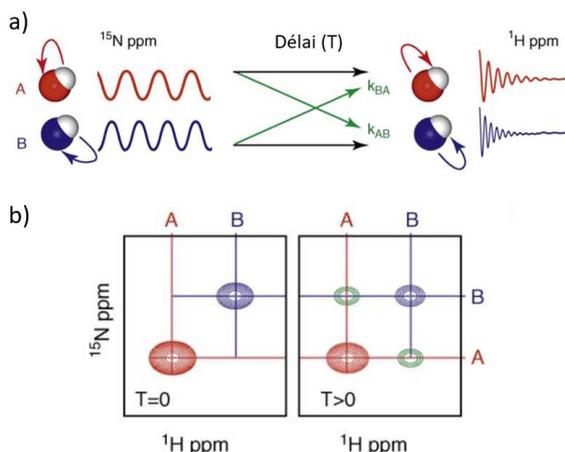


Figure 1.35 : a) Représentation schématique de la séquence d'impulsions RMN EXSY. b) Spectre EXSY simulé pour une paire de spin $^1\text{H}/^{15}\text{N}$ en échange chimique entre deux états, à gauche pour $T = 0$: on observe les pics des états AA et BB, à droite pour $T > 0$: on observe les pics des états AA et BB mais également deux pics supplémentaires correspondant aux sondes qui ont interchangeé d'état pendant la période d'évolution. Figure adaptée de (Mittermaier and Kay 2009).

En mesurant l'intensité des quatre différents pics aux différents temps de mélange puis en ajustant ces intensités il est possible d'extraire des informations sur les constantes d'échange k_A et k_B et donc sur la constante de vitesse d'échange conformationnel entre les deux états (Mittermaier and Kay 2009; Kleckner and Foster 2011; Charlier et al. 2016). Initialement, l'intensité des pics diagonaux augmente lorsque le temps de mélange augmente puis à partir d'un certain temps T , en raison de la relaxation des spins, l'intensité des pics diagonaux (AB et BA) diminue. Pour cette méthode il est important de respecter une période de récupération longue entre la détection et la période de préparation suivant la détection. En effet, l'EXSY quantifie l'intensité des pics pour en extraire la population il est donc fondamental que l'aimantation soit revenue à l'équilibre entre deux scans (Figure 1.36). Parmi les limitations possibles à la méthode EXSY on note : l'introduction de pics additionnels dans le spectre augmentant l'encombrement spectral et/ou la faible sensibilité due à l'apparition de pics parfois très peu intenses (Kleckner and Foster 2011).

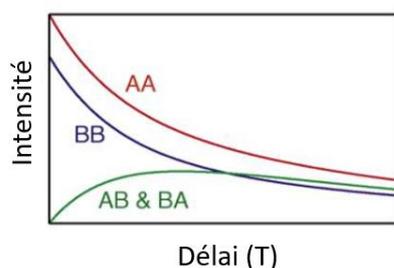


Figure 1.36 : Evolution de l'intensité des pics AA, BB, AB et BA en fonction du temps de mélange EXSY (T). Figure extraite de (Mittermaier and Kay 2009).

Lorsque l'élargissement des raies est tel qu'il empêche l'observation de l'un des deux états en échange chimique, l'EXSY ne peut plus être employé et différentes méthodes ont été développées pour obtenir des informations sur cet état invisible et notamment le CEST et la CPMG-RD, développées dans les parties suivantes.

Exemple

Sprangers et al. (Sprangers et al. 2005) ont étudié ClpP, une protéase oligomérique de 300 kDa composée de deux heptamères superposés l'un au-dessus de l'autre. Il est reconnu que l'enzyme protéolytique ClpP s'associe avec ClpA et ClpX qui lui délivrent les substrats à dégrader aux travers de ses pores axiaux, toutefois le mécanisme de relargage des produits dégradés restait hypothétique. Grâce à une étude de dynamique par RMN, Sprangers et al. ont proposé un chemin possible de sortie des produits après dégradation. En combinant perdeutération et marquage sélectif $^{13}\text{C}^1\text{H}_3$ sur les isoleucines δ_1 , ils ont pu obtenir des spectres RMN de qualité pour ClpP et ce, même à 0.5 °C. En attribuant les résonances des isoleucines δ_1 par mutagénèse, il est apparu que deux signaux correspondaient à l'isoleucine 151 et deux autres à l'isoleucine 149 avec dans les deux cas, une résonance large et une fine (Figure 1.37).

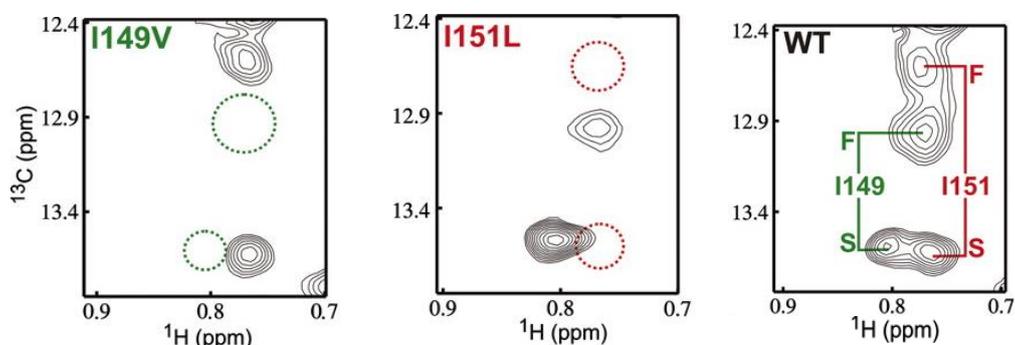


Figure 1.37 : Spectres 2D méthyle HMQC de ClpP marquée U- $^{15}\text{N},^2\text{H}$], Ile- δ_1 - $^{13}\text{C}^1\text{H}$] à 0.5 °C pour les mutants I149V, I151L et la protéine sauvage zoomés sur la région du spectre montrant les résonances des isoleucines 149 et 151. Deux raies, une large et une fine, correspondent à l'isoleucine 151 et deux à l'isoleucine 149. Figure extraite de (Sprangers et al. 2005).

Avec une expérience EXSY ils ont pu montrer que les résonances dédoublées étaient en échange conformationnel et ils ont pu quantifier la vitesse de cet échange ($k_{\text{ex}} \sim 61 \text{ s}^{-1}$) en effectuant un ajustement global après avoir montré que les deux isoleucines appartenait au même processus d'échange. Des expériences de CPMG relaxation-dispersion (voir partie 1.3.2.4) ont également été employée et confirment les données

mesurées avec les expériences EXSY ($k_{ex} \sim 68 \text{ s}^{-1}$). Les isoleucines 151 et 149, présentes dans chacun des monomères, sont localisées à la surface des régions équatoriales, à l'interface entre les deux heptamères (Figure 1.38). Aussi, la dynamique dans cette région pourrait être à l'origine de la sortie des substrats dégradés.

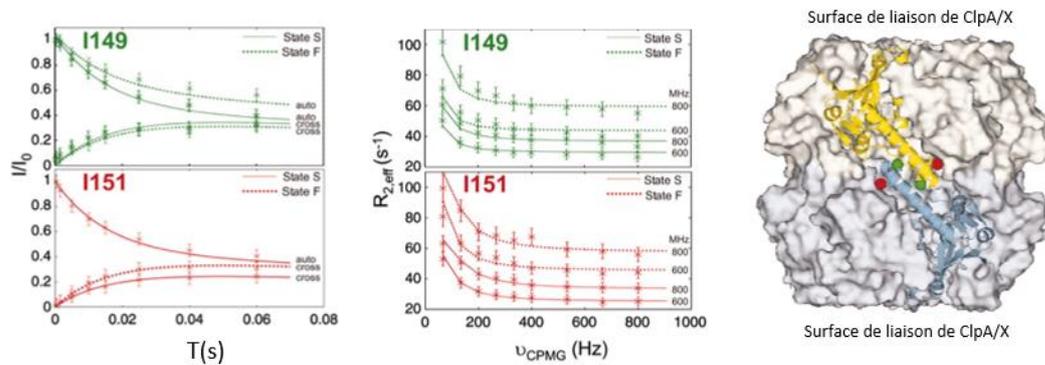


Figure 1.38 : A gauche, profils de décroissance des pics diagonaux et d'accumulation des pics de corrélation, en fonction du délai T , obtenus avec des expériences EXSY pour les isoleucines 149 et 151. Au centre, profils de relaxation-dispersion obtenus pour l'isoleucine 149 et l'isoleucine 151 avec des données acquises à deux champs magnétiques : 600 MHz et 800 MHz et à $0.5 \text{ }^\circ\text{C}$. A droite, surface tridimensionnelle de ClpP, les deux heptamères sont colorés respectivement en blanc et bleu ciel. La structure secondaire de deux monomères est représentée en bleu et en jaune. Les points verts et rouges représentent respectivement la localisation des isoleucines 149 et 151. Figure adaptée de (Sprangers et al. 2005).

Pour vérifier cette hypothèse, Sprangers et al. ont étudié l'activité protéolytique de plusieurs mutants et notamment du I151L et du A153C. Le mutant I151L ne montre aucune activité protéolytique bien que les déplacements chimiques des autres isoleucines restent identiques à ceux observés pour la protéine sauvage. Une orientation stricte des hélices semble donc importante pour l'activité protéolytique de ClpP. Le mutant A153C quant à lui, pouvant former des ponts disulfures a été étudié en conditions réduites et oxydées. Si la forme réduite permet d'observer les 4 pics des deux isoleucines 149 et 151, la forme oxydée ne permet pas l'observation de résonances pour ces deux isoleucines à basse température qui aurait pu permettre de montrer que l'échange était aboli lorsque les ponts disulfures étaient formés. Des expériences biochimiques employant la fluorescence ont été mises en place et ont montré que sous sa forme réduite le mutant A153C avait une activité protéolytique même si cette dernière était réduite par rapport à la protéine sauvage. A l'inverse, la protéine oxydée n'a pas montré d'activité protéolytique et le peptide est retenu plus de 10 h par ClpP. Ces résultats suggèrent fortement que la dynamique de la protéine au niveau des

isoleucines localisées à la surface des régions équatoriales est liée à la libération des substrats dégradés (Figure 1.39) (Sprangers et al. 2005; Mittermaier and Kay 2009).

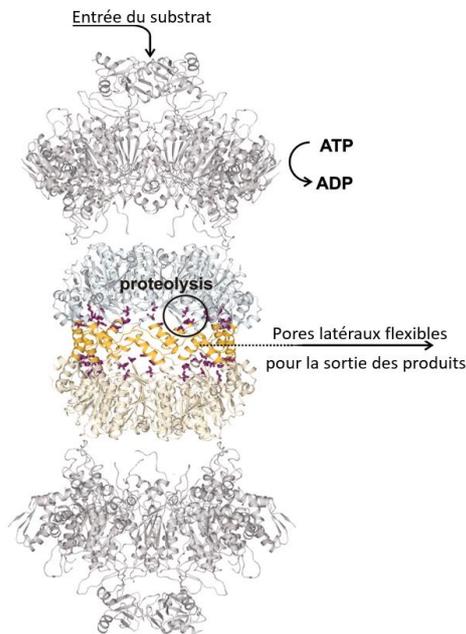


Figure 1.39 : Chemin possible pour la sortie du substrat dégradé. Le substrat est délivré à ClpP par ClpX, représentée en gris sur le schéma. Les hélices dynamiques de ClpP sont colorées en orange. La possible sortie des produits après dégradation est représentée par une flèche. Figure adaptée de (Sprangers et al. 2005).

1.3.2.3 CEST (Chemical Exchange Saturation Transfert) ($\tau_{\text{ex}} \sim 2 - 50 \text{ ms}$ / $k_{\text{ex}} \sim 20-500 \text{ s}^{-1}$)

Le CEST est une technique qui permet de « visualiser » les résonances d'un état invisible en se servant des résonances de l'état peuplé majoritairement avec lequel il est en échange. Pour pouvoir employer cette méthode l'échange entre l'état majoritaire et minoritaire doit être compris entre 20 et 500 s^{-1} . En pratique, un faible champ magnétique B_1 est appliqué selon x (ou y) à une fréquence ω_i pendant un temps T_{ex} avant l'acquisition du signal et cette étape est répétée pas à pas sur l'ensemble du spectre. L'objectif est de quantifier l'intensité du pic majoritaire visible en fonction de l'offset du champ magnétique B_1 par rapport à la position du pic majoritaire pour en déduire la position du pic de l'état invisible. Ainsi, si le champ magnétique B_1 est appliqué à la fréquence de résonance de l'état majoritaire alors ce dernier sera rapidement saturé et son aimantation sera nulle. Si B_1 est appliqué à une fréquence loin de la fréquence de résonance des états majoritaire et minoritaire cela n'aura pas d'influence sur l'intensité du pic majoritaire. Enfin, si B_1 est appliqué à la fréquence du pic minoritaire alors la saturation de ce dernier sera échangée avec l'état avec lequel il est en échange conformationnel et ainsi l'intensité du pic majoritaire sera plus faible (Figure 1.40). On pourra donc connaître la valeur du déplacement chimique entre l'état majoritaire et l'état minoritaire. Le CEST permet ainsi de retrouver des informations pour « voir » un état préalablement invisible (Vallurupalli et al. 2012, 2017; Charlier et al. 2016).

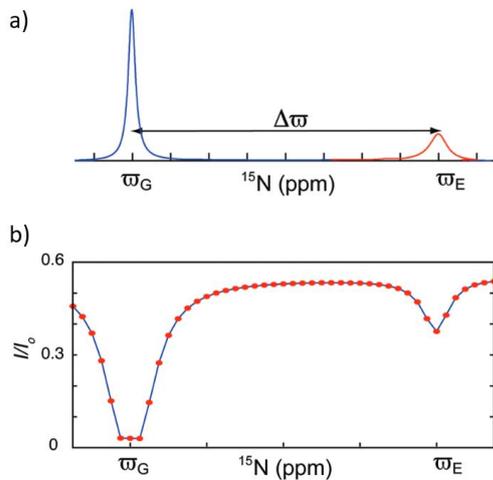


Figure 1.40 : a) Spectre d'un spin isolé en échange entre deux états G et E ayant des déplacements chimiques $\overline{\omega}_G$ et $\overline{\omega}_E$ et avec G l'état majoritaire. L'état minoritaire E observé à la fréquence $\overline{\omega}_E$ n'est en réalité pas observable sur le spectre en raison de l'élargissement des raies trop important et est présenté ici pour illustrer le propos. b) Graphique représentant le profil d'intensité I/I_0 du pic majoritaire en fonction de la position du champ magnétique B_1 . Avec I_0 l'intensité mesurée en absence de délai T_{ex} et I l'intensité mesurée après irradiation du champ B_1 . Le rapport d'intensité I/I_0 diminue quand B_1 est en résonance avec l'état majoritaire ou minoritaire. Figure adaptée de (Vallurupalli et al. 2012).

Exemple

La superoxyde dismutase SOD1 est une métalloenzyme dont la mutation du gène *sod1* l'encodant est reporté dans 20 % des cas de sclérose en plaque latérale amyotrophique familiales (SLAf). Malgré le lien entre SLAf et mutation du gène *sod1* et de nombreuses études, les événements moléculaires amenant à la maladie restent incompris. Dans sa forme mature, SOD1 est un homodimère dont chacun des monomères, constitué de 153 acides aminés, lie un atome de zinc et un atome de cuivre et comporte un pont disulfure. Dans sa forme la plus immature, SOD1 ne lie plus les ions métalliques et ne comporte plus de ponts disulfures. Comparé à la protéine mature, SOD1 immature a une plus grande probabilité d'être mal repliée et/ou d'agréger conduisant, potentiellement au développement de la SLAf. Sekhar et al. (Sekhar et al. 2015) ont étudié SOD1 immature et plus précisément sa dynamique par RMN en solution. En employant des expériences de CEST et de CPMG (voir partie 1.3.2.4) ils ont montré que SOD1 immature, sous forme monomérique dans les conditions RMN employées, était en échange entre sa forme fondamentale et au moins quatre états faiblement peuplés. Ces résultats apportent des informations précieuses pour comprendre au niveau atomique les potentiels mécanismes d'agrégation de SOD1 immature. Avec les expériences de CEST Sekhar et al. ont pu identifier deux des états excités de SOD1. Un premier processus d'échange a été observé où les expériences de CEST ont permis de visualiser les fréquences de résonance de l'état excité. Dans ce premier processus, la taille du pic de l'état minoritaire est proportionnelle à la concentration de la protéine, justifiant que l'état excité observé s'oligomérisse ou s'associe. Les expériences de CEST sur les atomes ^{15}N , ^{13}Ca et ^{13}C du squelette localisent l'échange dans la région de dimérisation du dimère natif mature de SOD1. De plus, les déplacements chimiques observés par CEST pour

l'état excité corrèlent avec les déplacements chimiques reportés dans la littérature pour l'homodimère de SOD1 mature (Figure 1.41). Aussi, l'état excité numéro 1 de SOD1 immature serait similaire à l'homodimère de SOD1 mature. Ce dernier aurait une demi vie de 3 ms et représenterait 3 % de la population (pour une concentration de SOD1 immature de 1.3 mM).

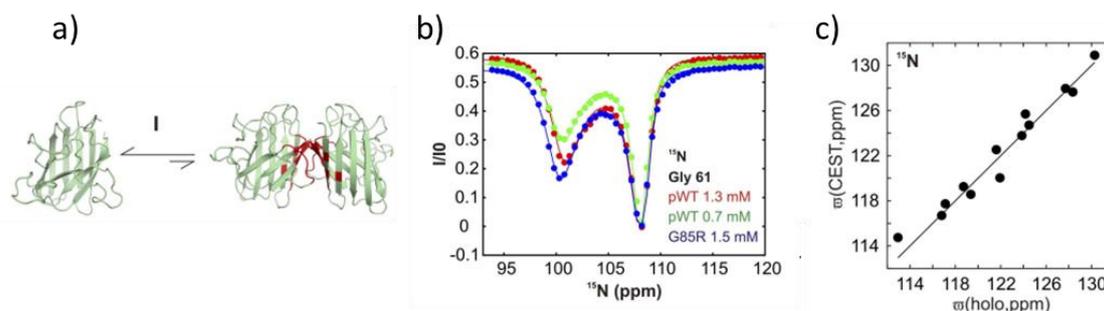


Figure 1.41 : a) Représentation schématique du processus d'échange I observé par CEST. b) Graphique représentant le profil d'intensité I/I_0 du pic majoritaire en fonction de la position du champ magnétique B_1 pour la glycine 61. Les courbes rouge et verte montrent l'influence de la concentration de SOD1 immature sur la taille du pic de l'état excité. En bleu : courbe pour le mutant G85R. c) Corrélation entre les déplacements chimiques de l'état excité observés par CEST et ceux reportés dans la littérature pour l'homodimère de SOD1 mature.

Un second processus d'échange a été observé et les expériences de CEST ont permis de visualiser les fréquences de résonance du deuxième état excité. Dans ce second processus, la taille du pic de l'état minoritaire n'est pas influencée par la concentration de la protéine. Les expériences de CEST sur les atomes ^{15}N , $^{13}\text{C}\alpha$ et $^{13}\text{C}'$ du squelette localisent l'échange au niveau d'une région désordonnée de SOD1 immature formant une hélice dans le cas de SOD1 mature. Les déplacements chimiques observés par CEST pour l'état excité du processus numéro 2 corrèlent avec les déplacements chimiques reportés dans la littérature pour l'hélice de SOD1 mature (Figure 1.42). Aussi, il apparaît que le second état excité de SOD1 immature correspond comme l'état fondamental à un monomère mais une hélice se forme de façon transitoire dans la région des résidus 130 à 140. Cet état excité aurait une demi vie de 13 ms et représenterait 2 % de la population.

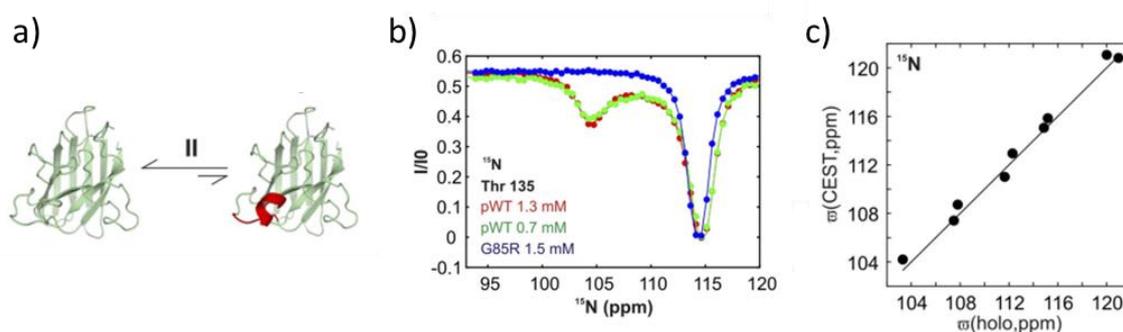


Figure 1.42 : a) Représentation schématique du processus d'échange II observé par CEST. b) Graphique représentant le profil d'intensité I/I_0 du pic majoritaire en fonction de la position du champ magnétique B_1 pour la thréonine 135. Les courbes rouge et verte montrent la non influence de la concentration de SOD1 immature sur la taille du pic de l'état excité. En bleu : courbe pour le mutant G85R. c) Corrélation entre les déplacements chimiques de l'état excité observés par CEST et ceux reportés dans la littérature pour l'homodimère de SOD1 mature.

Un des points clés pour l'identification de ces deux états excités a été la séparation des résidus appartenant aux différents processus d'échange. Les deux processus d'échange observés par CEST sont indépendants. En effet, la mutation G85R élimine le processus numéro 2 mais n'a que peu d'influence sur le processus numéro 1 (Figure 1.41b et Figure 1.42b). Les expériences de CPMG, sondant une dynamique plus rapide que le CEST, ont également apporté des informations sur deux autres états excités. Les modèles structuraux des quatre états excités identifiés et particulièrement ceux non-natifs, pourraient permettre l'éventuelle conception d'agents thérapeutiques permettant de prévenir l'agrégation de SOD1 immature conduisant potentiellement à la SLAf.

1.3.2.4 CPMG-RD (Carr Purcell Meiboom Gill Relaxation – Dispersion) ($\tau_{\text{ex}} \sim 0.15\text{-}10\text{ ms}$ / $k_{\text{ex}} \sim 100\text{-}6000\text{ s}^{-1}$)

La CPMG-RD est une des techniques principales pour quantifier l'échange chimique des macromolécules. Pour pouvoir employer cette méthode l'échange entre l'état majoritaire et minoritaire doit être compris entre 100 et 6000 s^{-1} (Kleckner and Foster 2011; Vallurupalli et al. 2011). Il faut également que les signaux observés montrent un élargissement des raies, soit une relaxation accrue $R_2^{\text{eff}} = R_2^0 + R_{\text{ex}}$ (avec R_2^0 la relaxation transverse intrinsèque et R_{ex} la contribution de l'échange à la relaxation). L'objectif de la CPMG-RD est de refocaliser l'élargissement des raies dû à l'échange (i.e de minimiser la contribution de l'échange à la relaxation transversale) en appliquant des impulsions 180° en série pendant un délai T (Vallurupalli et al. 2017). En pratique, on cherche à suivre l'évolution de la relaxation transversale (R_2^{eff}) en fonction de la fréquence à

laquelle sont appliquées les impulsions 180° . Avec cette technique on enregistre des spectres RMN 2D, acquis avec un nombre variable d'impulsions 180° séparées par des délais σ pendant un délai T constant (T_{CPMG}). Si le délai σ entre deux impulsions est long, le déphasage dû à l'échange entre les deux conformations se produira lors de chaque délai σ . Ainsi R_2^{eff} ne sera que très peu affecté par l'application des impulsions 180° (Figure 1.43).

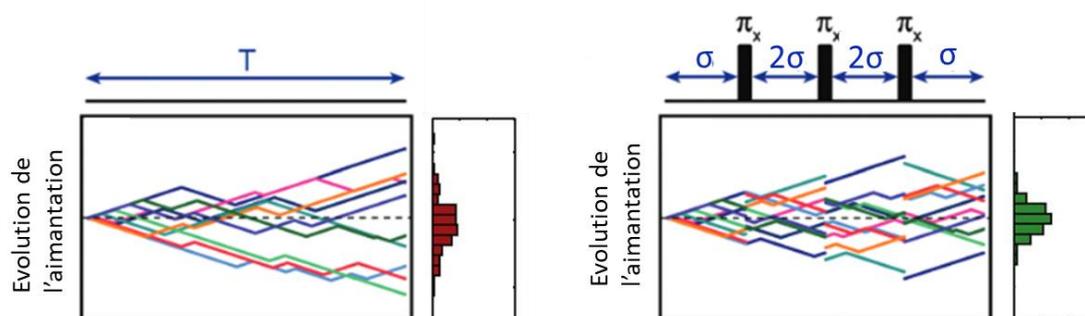


Figure 1.43 : Evolution de l'aimantation transverse en présence d'échange chimique entre deux états A et B. A gauche les fréquences de précessions échangent entre les fréquences de l'état A et de l'état B dû à l'échange chimique et ne sont pas refocalisées pendant le délai T . A droite, trois impulsions de 180° pendant le délai T refocalisent seulement en partie l'élargissement des raies dû à l'échange chimique. Le délai σ avant et après chaque impulsion est long, le déphasage dû à l'échange entre les deux conformations se produit. Figure adaptée de (Mittermaier and Kay 2009; Charlier et al. 2016).

A l'inverse, si le délai σ entre deux impulsions est court, le déphasage dû à l'échange entre les deux conformations n'aura pas le temps de se produire entre chaque délai σ . Ainsi R_2^{eff} sera fortement affecté par l'application des impulsions 180° puisque la contribution R_{ex} à la relaxation transverse aura en grande partie, voire totalement, été supprimée (Figure 1.44) (Charlier et al. 2016).

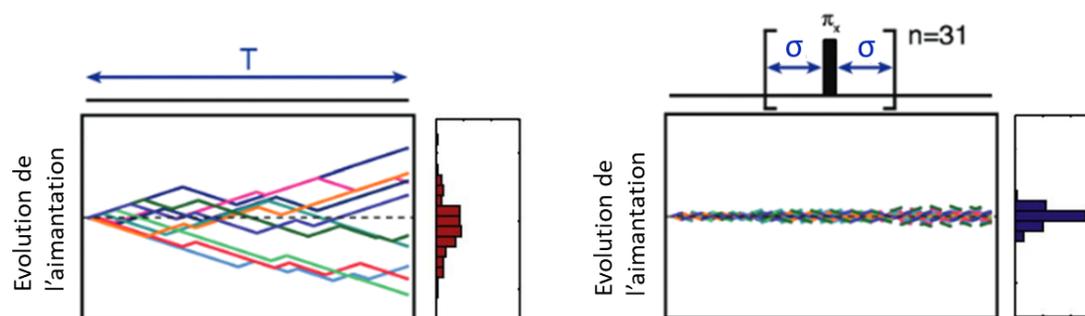


Figure 1.44 : Evolution de l'aimantation transverse en présence d'échange chimique entre deux états A et B. A gauche les fréquences de précessions échangent entre les fréquences de l'état A et de l'état B dû à l'échange chimique et ne sont pas refocalisées pendant le délai T. A droite, 31 impulsions de 180° pendant le délai T refocalisent presque totalement l'élargissement des raies dû à l'échange chimique. Le délai σ avant et après chaque impulsion est court, le déphasage dû à l'échange entre les deux conformations est en grande partie supprimé. Figure adaptée de (Mittermaier and Kay 2009; Charlier et al. 2016).

En effectuant un ajustement des courbes pour les différentes valeurs de R_2^{eff} obtenues aux différentes fréquences d'impulsions appliquées (ν_{CPMG}) (Figure 1.45) il est possible d'obtenir des informations sur la cinétique (k_{ex}) et la thermodynamique (P , P^*) de l'échange ainsi que sur la valeur de la différence de déplacement chimique entre l'état fondamental (majoritaire) et celui excité (minoritaire). Toutefois, il est à noter que la différence de déplacement chimique entre l'état fondamental et celui excité est obtenue en tant que valeur absolue, empêchant leur emploi pour, par exemple, calculer la structure de l'état excité (Bouvignies et al. 2011). La comparaison des déplacements chimiques dans la dimension indirecte pour une paire d'expériences HSQC acquises à différents champs magnétiques ou entre un spectre HSQC et un spectre HMQC acquis à champ magnétique identique permet de retrouver l'information sur le signe du déplacement chimique (Skrynnikov et al. 2002; Kovermann et al. 2016).

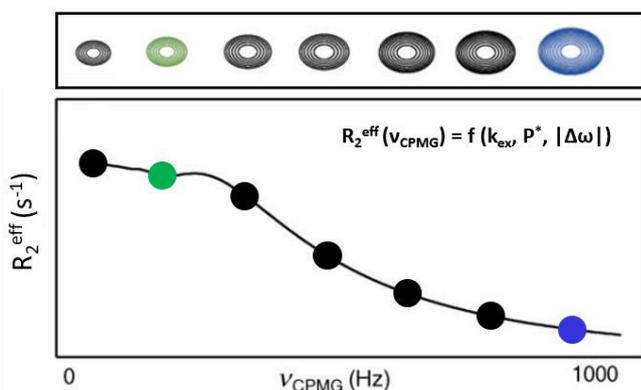


Figure 1.45 : Courbe montrant R_2^{eff} en fonction de la fréquence ν_{CPMG} . Les intensités des pics en vert et bleu correspondent qualitativement aux intensités obtenues figures 1.43 et 1.44 respectivement avec trois impulsions et 31 impulsions de 180°. Figure adaptée de (Mittermaier and Kay 2009).

Exemple

De par sa très faible constante d'affinité et la nature dynamique de l'interaction entre l'éthanol et ses protéines cibles, peu de connaissances sont disponibles sur les mécanismes moléculaires gouvernant l'action de l'éthanol. Pourtant, de telles informations pourraient être clés pour faciliter le développement de composés pharmaceutiques agissant sur le système nerveux puisqu'il a été reporté qu'éthanol et

anesthésiants avaient des effets similaires en se liant à des sites de fixations se recouvrant. Toyama et al. (Toyama et al. 2018) ont étudié GIRK, canal potassique à rectification entrante associé à la protéine G, et notamment ses changements structuraux et dynamiques lors de l'interaction avec l'éthanol. GIRK fonctionne comme un tétramère et est constitué d'une région cytoplasmique et d'une région transmembranaire avec un chemin pour le potassium au centre du tétramère. En se focalisant sur la partie cytoplasmique de GIRK, supposée lier l'éthanol, et en combinant marquage isotopique spécifique sur les groupements méthyles et acquisition d'expériences avec une relaxation transverse optimisée (expériences TROSY), Toyama et al. ont pallié l'absence de certaines résonances observée sur les groupements amides proches du site de fixation de l'éthanol et ont pu localiser ce dernier. Bien que des déplacements chimiques soient observés lors de la liaison de l'éthanol ces derniers étaient inférieurs à 0.1 ppm pour la majorité des groupements méthyles, suggérant que la structure du tétramère n'est pas substantiellement modifiée par la liaison d'éthanol. Toutefois, la dynamique de GIRK dans la gamme de temps μs -ms est impactée par la présence d'éthanol comme démontré par les expériences de CPMG-RD. GIRK dans son état Apo est en échange conformationnel. En effet, certains groupements méthyles montrent une contribution à l'échange supérieure à 20 s^{-1} . En faisant l'hypothèse d'un modèle à deux états la constante de vitesse d'échange conformationnel, la population de l'état excité ainsi que les différences de déplacement chimiques $|\Delta\omega_C|$ et $|\Delta\omega_H|$ ont été ajustées. Une constante de vitesse d'échange conformationnel de $3300 \pm 30 \text{ s}^{-1}$ et une population d'état excité de $11 \pm 0.4 \%$ ont été obtenues. Les groupements méthyles reportant de l'échange ont été cartographiés sur la structure de la partie cytoplasmique de GIRK. De cette cartographie, l'hypothèse a été faite que GIRK était en échange entre une conformation avec le G-loop ouvert (état fondamental) et une conformation avec le G-loop fermé (état excité). Le G-loop étant la porte de passage du potassium du côté transmembranaire de la région cytoplasmique de GIRK. Cette hypothèse a également été supportée par le mutant E304A stabilisant la forme fermée du G-loop. En présence d'éthanol à diverses concentrations, la dynamique de GIRK dans la gamme de temps μs -ms a également été étudiée. Il apparaît que la présence d'éthanol modifie le paysage énergétique et stabilise la conformation ouverte du G-loop. Plus la concentration d'éthanol augmente plus la conformation ouverte de GIRK est peuplée (Figure 1.46). En conclusion, la présence d'éthanol active GIRK et ainsi K^+ pourrait passer au travers de GIRK. Les auteurs suggèrent que ce mécanisme d'activation observé pour GIRK représente un mécanisme général de l'action de l'éthanol sur les protéines de signalisation du système nerveux.

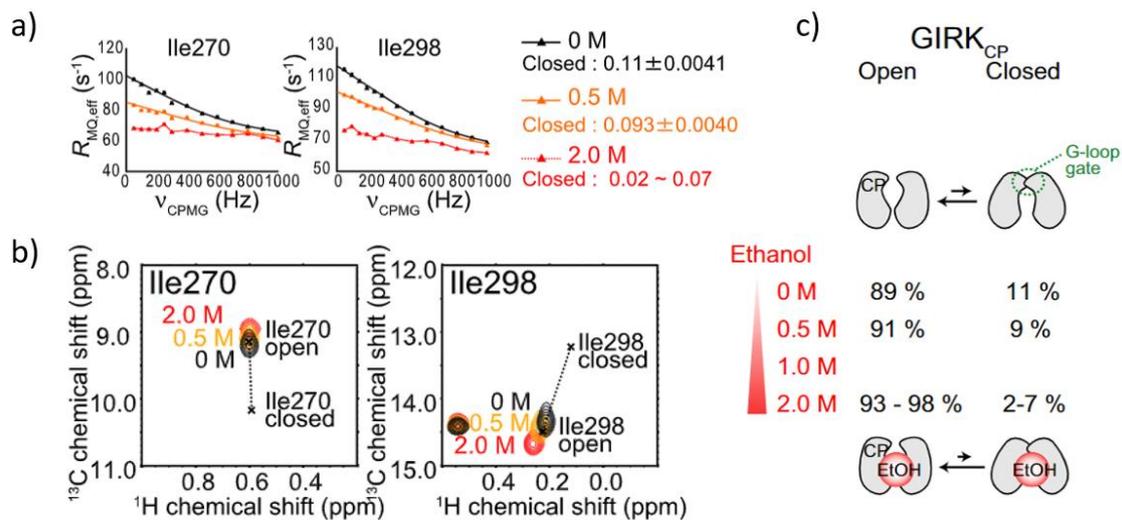


Figure 1.46 : a) Courbes de relaxation dispersion de la partie cytoplasmique de GyrA en l'absence (noir) et présence d'éthanol à 0.5 M (orange) ou 2M (rouge). b) Superposition des spectres 2D ¹H-¹³C zoomés sur l'isoleucine 270 et l'isoleucine 298 en absence (noir) et présence d'éthanol à 0.5 M (orange) et 2.0 M (rouge). c) Représentation schématique de l'influence de la concentration d'éthanol sur la conformation de la partie cytoplasmique de GyrA. Figure adaptée de (Toyama et al. 2018).

1.3.2.5 Relaxation dispersion dans le référentiel tournant ($\tau_{\text{ex}} \sim 20\text{-}100 \mu\text{s}$ / $k_{\text{ex}} \sim 10000\text{-}50000 \text{ s}^{-1}$)

La relaxation dispersion dans le référentiel tournant est une méthode adaptée pour étudier les phénomènes se produisant dans la gamme de temps 10000 - 50000 s⁻¹. Cette méthode fonctionne sur un principe similaire à la CPMG-RD. L'objectif est de réduire la relaxation de l'aimantation alignée parallèlement au champ effectif dans le référentiel tournant ($R_{1\rho}$) en appliquant des impulsions appelées « spin-lock » à différentes fréquences, en général comprises entre 1 et 50 kHz (Kleckner and Foster 2011). En pratique, on cherche à suivre l'évolution de $R_{1\rho}$ en fonction du champ effectif appliqué pour en déduire des informations sur l'échange se produisant. L'orientation du champ effectif dépend à la fois de l'amplitude de l'impulsion « spin lock » appliquée et de l'offset entre la fréquence à laquelle est appliquée l'impulsion et la résonance du pic de l'état majoritaire. Dans le référentiel tournant, la relaxation de l'aimantation alignée parallèlement au champ effectif ($R_{1\rho}$) est exprimée par : $R_{1\rho} = R_1 \cos^2(\theta) + R_2 \sin^2(\theta)$ avec R_2 et R_1 respectivement la relaxation transversale, et longitudinale et θ l'angle entre le champ effectif et l'axe z. Il est donc possible d'obtenir des informations sur R_2 en mesurant $R_{1\rho}$ et R_1 . Dans le cas d'un échange chimique, la contribution de l'échange à la relaxation R_{ex} , dépendant du champ effectif, peut être déterminée en représentant R_2

en fonction du champ effectif appliqué ($R_2 = R_2^0 + R_{ex}$) (Massi et al. 2005). Avec cette technique on enregistre des spectres RMN, acquis avec un champ effectif d'amplitude variable. En présence d'échange chimique $R_{1\rho}$ est augmenté puisque le noyau sera en échange entre deux conformations avec chacune un environnement chimique différent. Ainsi, si l'aimantation est initialement alignée avec un faible champ effectif lorsque l'échange se produit le noyau en échange passera une partie de la durée de l'impulsion « spin-lock » dans un environnement chimique différent, amenant à de larges variations de l'axe de rotation et à une trajectoire divergente de l'aimantation. L'aimantation alignée avec le champ effectif de l'état majoritaire décroîtra rapidement. Plus le champ effectif appliqué sera fort, plus la décroissance de l'aimantation alignée avec le champ effectif deviendra négligeable (Figure 1.47) (Charlier et al. 2016).

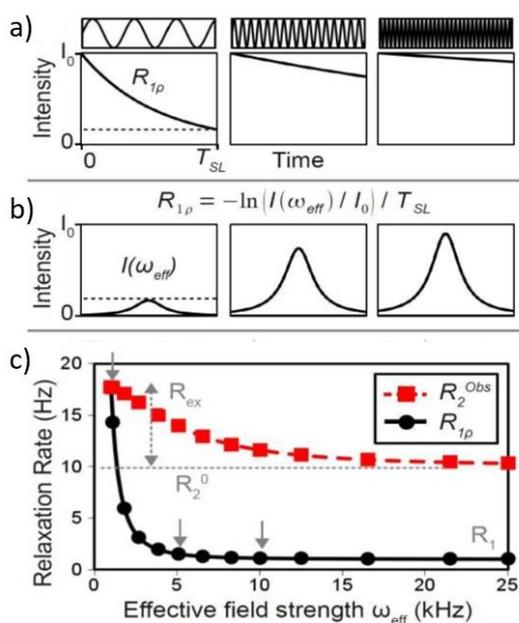


Figure 1.47: a) Plus le champ effectif appliqué est fort plus la relaxation $R_{1\rho}$ est réduite, l'échange est refocalisé. b) Signal observé sur les spectres RMN permettant d'en déduire : $R_{1\rho}$ en mesurant l'intensité des résonances. Avec I_0 , l'intensité de la raie en absence de délai de relaxation. c) Courbe montrant l'évolution de : $R_{1\rho}$ (en noir) et R_2 (en rouge) en fonction de la fréquence du champ effectif (ω_{eff}). Figure extraite de (Kleckner and Foster 2011).

En effectuant un ajustement des courbes pour les différentes valeurs de R_2 obtenues aux différentes valeurs d'amplitude du champ effectif il est possible d'obtenir des informations sur le processus d'échange. Toutefois il est à noter que la majorité des phénomènes se produiront dans le régime rapide ($k_{ex} \gg |\Delta\omega|$) aussi les populations des états et les valeurs de déplacements chimiques seront rassemblées en un unique paramètre $\phi_{ex} = P_A P_B |\Delta\omega|^2$, limitant l'interprétation des valeurs de populations et de déplacements chimiques.

Exemple

L'ubiquitine, protéine de 76 acides aminés très conservée, est une protéine clé du système ubiquitine-protéasome de dégradation de protéines. De façon très succincte, l'ubiquitine est activée par l'adénylation de la glycine C-terminale puis liée de façon covalente à E1, une enzyme activatrice. L'ubiquitine est ensuite transférée de E1 à E2, une enzyme de conjugaison. Finalement l'ubiquitine est liée de façon covalente au groupement ϵ d'un résidu lysine de la protéine cible à dégrader. Les protéines sur lesquelles plusieurs ubiquitines sont liées sont ensuite ciblées par le 26S protéasome pour dégradation (Massi et al. 2005; Kumar and Wiener 2019). Plusieurs acides aminés de l'ubiquitine ont été identifiés par mutagenèse et RMN comme ayant un rôle crucial dans les interactions de l'ubiquitine avec les enzymes E1 et E2. Notamment, un nombre important de résidus impliqués pour les interactions avec E1 est commun avec ceux impliqués pour les interactions avec E2. Massi et al (Massi et al. 2005) ont employé la méthode de relaxation dispersion dans le référentiel tournant sur les atomes ^{15}N pour quantifier les processus d'échanges affectant différents résidus de l'ubiquitine, la dynamique conformationnelle des surfaces d'interaction des protéines pouvant être une possible explication de la reconnaissance de plus d'un partenaire de liaison (James et al. 2003). A 280 K, quatre résidus : Ile23, Asn25, Thr55 et Val70 montrent une contribution de l'échange à la relaxation transversale significative. Les courbes de dispersion de chacun des quatre acides aminés en fonction du champ effectif (Figure 1.48a) ont été ajustées individuellement et une constante de vitesse d'échange conformationnel de l'ordre de 25000 s^{-1} a été obtenue pour chacun des quatre résidus suggérant qu'ils détectent tous le même phénomène d'échange (Figure 1.48b). Comme mentionné précédemment, dans le cas du régime d'échange rapide les populations ne peuvent pas être décorréliées de la différence de déplacement chimique. Aussi, un terme ϕ_{ex} , avec $\phi_{\text{ex}} = P_A P_B (\Delta\omega_N)^2$, a été obtenu pour chacun des quatre ajustements (Figure 1.48b). En faisant l'hypothèse que les quatre résidus voient le même phénomène d'échange avec des populations similaires alors la différence dans les valeurs de ϕ_{ex} dépend principalement de la valeur de $|\Delta\omega_N|$. Asn25 et Thr55 sont les résidus montrant respectivement la plus grande et la plus petite valeur de ϕ_{ex} . Avec ces informations, Massi et al. ont pu estimer, en considérant que les changements se produisant dans les liaisons hydrogènes et les angles dièdres contribuaient majoritairement à la dynamique conformationnelle (d'après des résultats provenant d'expériences supplémentaires) que la population de l'état excité était de l'ordre de 2 %.

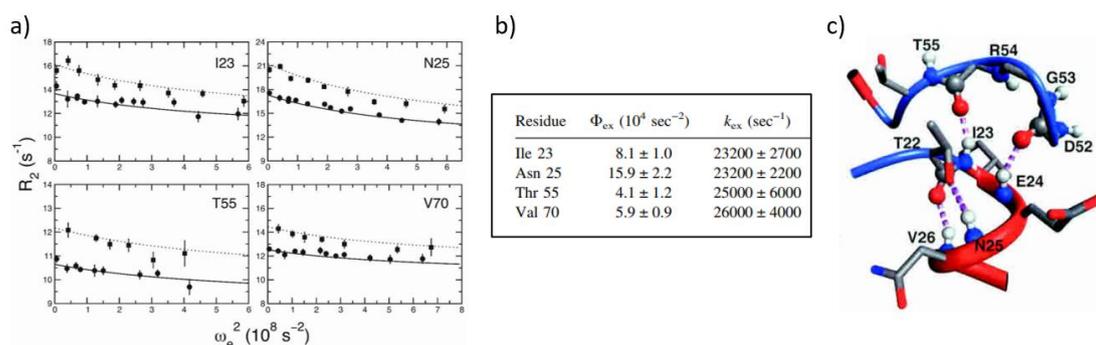


Figure 1.48 : a) Profils de relaxation dispersion en fonction du champ effectif obtenus pour les quatre résidus Ile23, Asn25, Thr55 et Val70 à $T = 280$ K. Les ronds et les carrés représentent respectivement les données collectées avec un champ magnétique statique de 11.7 et 14.1 Teslas. b) Paramètres obtenus sur l'échange chimique pour les résidus Ile23, Asn25, Thr55 et Val70 de l'ubiquitine. c) Zoom de la structure tridimensionnelle de l'ubiquitine montrant la connexion entre les résidus Ile23, Arg54 et Thr55. Les liaisons hydrogènes sont représentées en pointillées. Figure adaptée de (Massi et al. 2005).

Parmi les résidus de l'ubiquitine impliqués dans la liaison avec E1 on note, entre autres, Arg54, Ile44, Val70 et Arg72. Concernant les résidus impliqués dans l'interaction avec E2 on note les segments de résidus 40-49 et 70-76. Les résidus montrant de l'échange Ile23, Asn25, Thr55 et Val70 sont donc localisés dans les zones, ou proches des zones, interagissant avec les enzymes E1 et E2. Notamment, Ile23 et Arg54 sont tous les deux connectés à l'Arg54, soit de façon covalente ou par liaison hydrogène (Figure 1.48c) et Val70 est directement impliqué dans la région interagissant avec à la fois E1 et E2. La dynamique conformationnelle de l'ubiquitine autour de 40 μ s joue potentiellement un rôle important dans la reconnaissance des différents partenaires physiologiques.

1.3.2.6 Relaxation de spin nucléaire ($\tau_{ex} \sim$ ps - 10 ns)

Une technique privilégiée pour mesurer la dynamique des protéines dans la gamme de temps ps-ns est la relaxation de spins nucléaires. Les spins reviennent à l'équilibre après avoir été excités par l'impulsion de radiofréquence B_1 et cela est dû à la présence de champs magnétiques locaux oscillant à des fréquences de l'ordre du MHz. Ces derniers résultent principalement de deux phénomènes pour les spins $\frac{1}{2}$: l'anisotropie de déplacement chimique et l'interaction dipolaire, induits par la rotation moléculaire par rapport au champ magnétique fixe B_0 (ns) et par les mouvements internes de la protéine (ps) (Kleckner and Foster 2011).

Des informations sur la dynamique dans la gamme de temps ps-ns (jusqu'à maximum la valeur du τ_c de la protéine) peuvent être obtenues en acquérant des expériences RMN mesurant les vitesses R_1 et R_2 de retour à l'équilibre du système ainsi que des NOE hétéronucléaires (hetNOE).

Différents modèles peuvent être employés pour interpréter les vitesses de relaxation nucléaire et les NOE hétéronucléaires et essayer de comprendre la fonction protéique en étudiant la dynamique dans la gamme de temps ps-ns. On notera une première approche où les valeurs de R_1 , R_2 et hetNOE sont reportées en fonction de la séquence de la protéine pour identifier les régions avec une dynamique différente. Si la protéine n'a pas de dynamique interne et est en rotation isotropique alors les valeurs de R_1 , R_2 et hetNOE seront les mêmes tout au long de la séquence. Des variations dans les valeurs de ces trois observables indiquent des zones où une dynamique interne est présente. La mobilité interne peut également être caractérisée par la cartographie de la densité spectrale réduite en fonction de la séquence de la protéine. Avec les expériences mesurant des hetNOE et les vitesses R_1 et R_2 de retour à l'équilibre du système acquises à un seul champ magnétique il est possible, pour une paire N-H, d'en déduire la densité spectrale à différentes fréquences : $J(0.87\omega_H)$, $J(0)$ et $J(\omega_N)$ en fonction de la séquence de la protéine (en faisant l'hypothèse que la fonction de densité spectrale peut s'écrire $J(\omega) = a/\omega^2 + b$, avec a et b des constantes) (Farrow et al. 1995). $J(0.87\omega_H)$ est un indicateur clé de la dynamique dans la gamme de temps inférieure à la nanoseconde et des mouvements rapides dans la gamme de temps picoseconde entraînent une augmentation de la valeur de $J(0.87\omega_H)$ (Ebert et al. 2008; Kleckner and Foster 2011). Enfin, l'approche la plus utilisée est l'analyse « model-free ». Dans cette troisième approche les mouvements internes sont considérés indépendants et beaucoup plus rapides que la rotation de la molécule. La densité spectrale $J(\omega)$ pour une paire isolée H-N peut s'écrire avec l'Equation 1.12 (Lipari and Szabo 1982) :

$$J(\omega) = \frac{2}{5} \left(\frac{\tau_c S^2}{1 + \omega^2 \tau_c^2} + \frac{(1 - S^2) \tau_e}{1 + \omega^2 \tau_e^2} \right) \quad (\text{Equation 1.12})$$

Avec trois constantes S^2 , τ_c et τ_e . Avec S^2 le paramètre d'ordre, mesurant la restriction du mouvement, dont la valeur, comprise entre 0 et 1 est souvent tracée en fonction de la séquence de la protéine pour identifier les régions de rigidité (S^2 proche de 1) ou de flexibilité (S^2 proche de 0) dans la gamme de temps ps-ns. Avec τ_c le temps de corrélation, caractéristique de la protéine et τ_e le temps de corrélation interne renseignant sur la vitesse du mouvement (Figure 1.49).

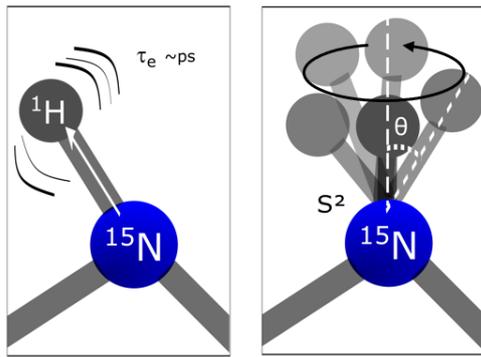


Figure 1.49 : A gauche représentation schématique de τ_e le temps de corrélation interne renseignant sur la vitesse du mouvement. A droite, représentation schématique de S^2 . Plus S^2 est proche de 1 plus l'angle θ sera petit. A l'inverse, plus l'angle θ sera grand plus S^2 sera proche de 0. Figure inspirée de (Mittermaier and Kay 2009).

Exemple

La relaxation de spin nucléaire a été employée pour étudier la dynamique d'un complexe formé par l'association de deux domaines de protéines désordonnées : ACTR et CBP (Ebert et al. 2008). ACTR et CBP sont deux coactivateurs de transcription dont le premier joue un rôle important dans la progression du cancer du sein et dont l'inhibition de la fonction pourrait amener à un traitement efficace. La structure du complexe formé entre le domaine de liaison de CBP et celui d'activation de ACTR a été déterminée par RMN. Si dans leurs formes libres les domaines de liaison de CBP et d'activation de ACTR sont, respectivement, partiellement déstructuré et non structuré, le complexe formé par l'association de ces deux domaines est, lui, structuré.

ACTR et CBP montrent une dispersion de déplacements chimiques limitée et si CBP montre deux segments ayant une structure secondaire en hélice définie, ACTR reporte une absence de structure secondaire dans son état libre. Dans leur état lié, les deux domaines comportent chacun trois segments montrant une structure secondaire en hélice.

Les données de relaxation de spin nucléaire montrant R_1 , R_2 et hetNOE en fonction de la séquence protéine pour chacun des deux domaines dans leurs états libres ou liés corroborent ces résultats (Figure 1.50). Toutes les valeurs de R_2 de ACTR dans sa forme libre sont situées en dessous de 5 s^{-1} et les valeurs de hetNOE $[^1\text{H}]-^{15}\text{N}$ sont toutes négatives. Les différences de R_2 et hetNOE $[^1\text{H}]-^{15}\text{N}$ pour les formes libre et liée de ACTR sont majeures.

Les différences entre les formes libres et liées du domaine de liaison de CBP pour R_2 et hetNOE $[^1\text{H}]-^{15}\text{N}$ sont moindres. En effet, CBP libre a été reporté comme molten globule et est donc moins désordonné que ACTR. On remarque toutefois, que certains résidus à l'extrémité de l'hélice 3 montrent une augmentation des valeurs de hetNOE $[^1\text{H}]-^{15}\text{N}$. Ces observations sont consistantes avec la formation d'une hélice supplémentaire pour le domaine de liaison de CBP lorsque le complexe avec ACTR est formé.

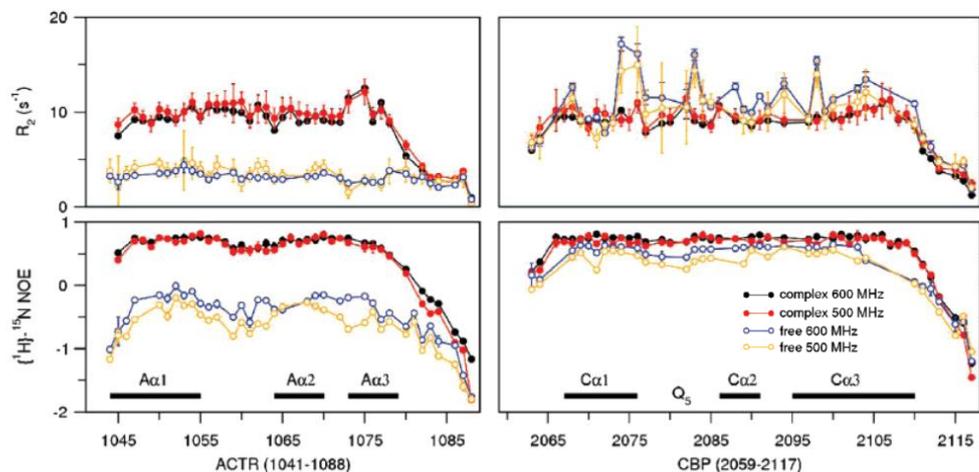


Figure 1.50: ^{15}N R_2 et ^1H - ^{15}N NOE des groupements amides du squelette du domaine d'activation d'ACTR et du domaine de liaison de CBP dans leurs formes libres à deux champs magnétiques (jaune et bleu) et dans leur forme liée à deux champs magnétiques (rouge et noir). Figure extraite de (Ebert et al. 2008).

La mobilité interne de ACTR et CBP a été étudiée avec la fonction de densité spectrale réduite. $J(0)$ et $J(0.87\omega_H)$ sont des indicateurs importants de la dynamique dans la gamme de temps inférieure à la nanoseconde. Des mouvements internes rapides dans l'échelle de temps de la picoseconde diminuent $J(0)$ et à l'inverse font augmenter la valeur de $J(0.87\omega_H)$. Comme pour l'analyse avec R_2 et hetNOE ^1H - ^{15}N , les différences entre $J(0)$ et $J(0.87\omega_H)$ sont moindres entre la forme libre et celle liée du domaine de liaison de CBP (Figure 1.51). De faibles valeurs de $J(0)$ sont observées pour les résidus terminaux de CBP montrant une plus grande dynamique interne rapide de cette région (déjà observée avec les valeurs décroissantes de hetNOE ^1H - ^{15}N) et cela est cohérent avec l'augmentation des valeurs de $J(0.87\omega_H)$ observée pour les mêmes résidus.

ACTR montre des valeurs de $J(0)$ et $J(0.87\omega_H)$ différentes dans sa forme libre ou liée. De hautes valeurs de $J(0)$ et de basses valeurs de $J(0.87\omega_H)$ dans la forme libre de ACTR mettent en évidence une dynamique interne rapide. En complexe, ACTR montre des valeurs de $J(0)$ identiques à CBP confirmant la formation d'une structure globulaire où les mouvements du squelette sont limités.

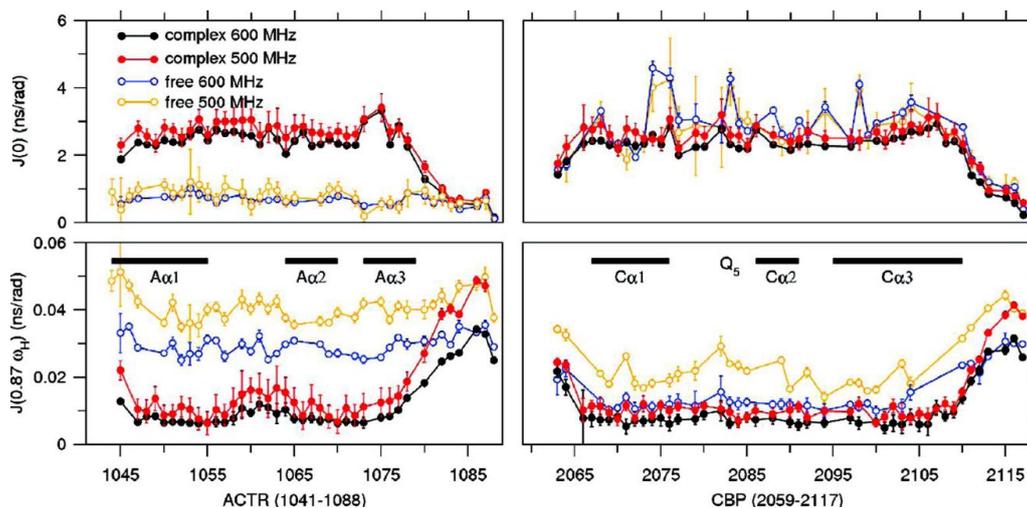


Figure 1.51 : Densités spectrales réduites $J(0)$ et $J(0.87\omega_H)$ des groupements amides du squelette du domaine d'activation d'ACTR et du domaine de liaison de CBP dans leurs formes libres à deux champs magnétiques (jaune et bleu) et dans leur forme liée à deux champs magnétiques (rouge et noir). Figure extraite de (Ebert et al. 2008).

L'analyse « model-free » a permis d'obtenir des informations sur les paramètres d'ordre (S^2) et les temps de corrélation interne (τ_e) pour le complexe (Figure 1.52). Il apparaît que la quasi-totalité des résidus ont un paramètre d'ordre entre 0.8 et 0.9, valeur typique d'une région structurée de protéine globulaire (Ebert et al. 2008) et les temps de corrélation interne sont en moyenne de 1 ns. Le temps de corrélation de la protéine (τ_c) a été estimé à 6.7 ns.

Ainsi, structure et dynamique de ACTR et CBP sont très distinctes dans leurs formes libres ou en complexe et un processus de repliement mutuel est observé. Habituellement la formation d'un complexe est favorisée entropiquement, les molécules de solvant sur les surfaces hydrophobes des domaines libres étant relâchées lors de la formation du complexe. Ici, la formation du complexe est très largement défavorisée entropiquement ($T\Delta S = -89.0 \text{ kJ.mol}^{-1}$) et cela est probablement dû à la perte de degré de liberté de ACTR et CBP et à leur repliement. Le coût entropique n'est pas compensé par une forte mobilité des atomes du squelette lorsque le complexe est formé comme l'ont montré les différentes valeurs élevées du paramètre d'ordre.

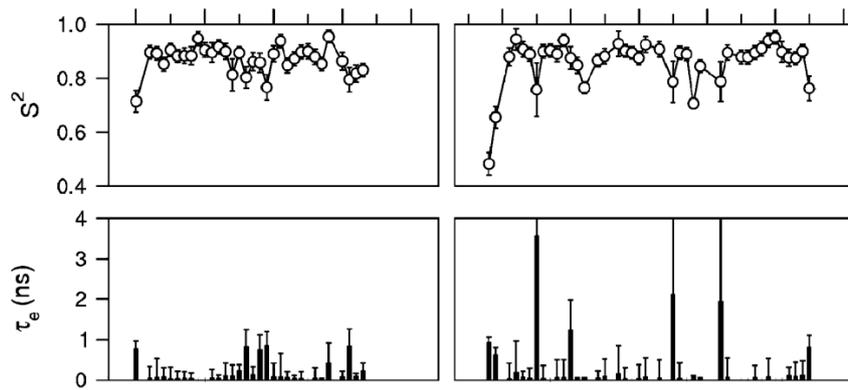


Figure 1.52 : Paramètres d'ordre et temps de corrélation interne en fonction de la séquence de la protéine pour le domaine d'activation d'ATCR et le domaine de liaison de CBP lorsque les deux protéines sont en complexe. Figure extraite de (Ebert et al. 2008).

Chapitre 2

Développements méthodologiques pour l'attribution du domaine N-terminal de HSP90 α humaine

2.1 Heat Shock Protein 90

HSP90 est une chaperonne moléculaire hautement conservée présente à la fois dans les bactéries et chez les eucaryotes (Farrelly and Finkelstein 1984; Bardwell and Craig 1987; Sorger and Pelham 1987; Song et al. 1995). Plusieurs homologues d'HSP90 existent dans les eucaryotes supérieurs avec des localisations différentes. On note TRAP1 présent dans les mitochondries, GRP94 dans le réticulum endoplasmique et HSP90 se trouvant dans le cytosol. HSP90, indispensable à la viabilité des eucaryotes (Borkovich et al. 1989), est la protéine la plus abondante du cytoplasme des cellules et peut représenter jusqu'à 2 % des protéines cytosoliques dans des conditions normales (Csermely et al. 1998). HSP90 se trouve sous deux isoformes majoritaires hautement homologues ayant 85 % de leur séquence identique : HSP90 α induit dans des conditions de stress et HSP90 β exprimé constitutivement (Millson et al. 2007; Johnson 2012).

HSP90 cytosolique est un homodimère d'environ 170 kDa dont chaque monomère est organisé en trois domaines distincts. Le domaine N-terminal (NTD), d'environ 25 kDa, possède le site de fixation de l'ATP (Prodromou and Roe 1997). Le domaine central (MD),

de 40 kDa, lié au domaine N-terminal par une longue chaîne d'acides aminés flexible, module l'hydrolyse de l'ATP en interagissant avec le γ -phosphate de ce dernier (Meyer et al. 2003). Enfin, le domaine C-terminal (CTD), d'environ 13 kDa, est indispensable à la dimérisation d'HSP90 (Nemoto et al. 1995; Meng et al. 1996; Hoter et al. 2018) (Figure 2.1).

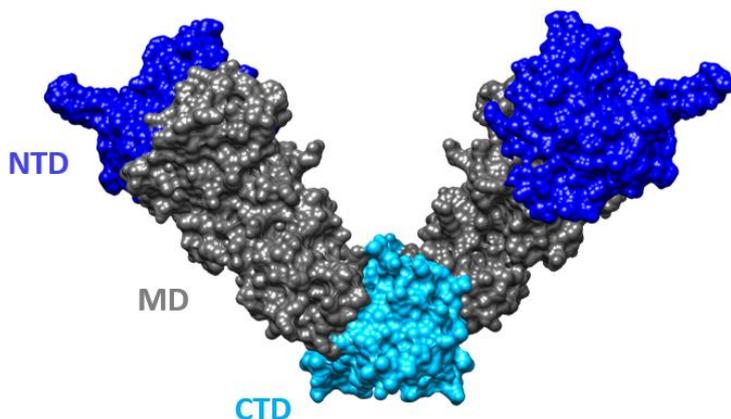


Figure 2.1 : Structure cristallographique d'HSP90 entière (PDB : 2IOQ). En bleu, gris et bleu ciel sont colorés respectivement le domaine N-terminal (NTD), le domaine central (MD) et le domaine C-terminal (CTD) de HSP90.

HSP90 est une chaperonne moléculaire particulièrement étudiée et notamment son domaine N-terminal, site de fixation de l'ATP. Durant ma thèse je me suis focalisée sur HSP90 α humaine et principalement sur son domaine N-terminal. Afin de pouvoir étudier par RMN la dynamique du domaine N-terminal d'HSP90 l'attribution de ses signaux est fondamentale. Les différentes méthodologies mises en place pour parvenir à une caractérisation des résonances de HSP90-NTD sont détaillées dans la suite de ce chapitre.

2.2 Données préalables sur l'attribution du squelette de HSP90 α -NTD

Le squelette du domaine N-terminal d'HSP90 α humaine Apo (sans ligand) a déjà été partiellement attribué par RMN (Jacobs et al. 2006; Park et al. 2011) à l'aide d'expériences de triple résonances classiques (HNCO/HNCACO, HNCOCA/HNCA et HNCOCACB/HNCACB). Deux attributions ont été déposées : BMRB 19560 et 7003 (Figure 2.2). Cependant, le squelette du segment L107-K116 à proximité du site de fixation de l'ATP n'est pas observé par RMN en raison d'une dynamique défavorable dans la gamme de temps μ s-ms (Figure 2.2). En effet, cette dynamique intermédiaire cause un élargissement important des signaux RMN des groupements amides les rendant invisibles/indistinguable du bruit du spectre (voir chapitre 1 partie 1.3.2).

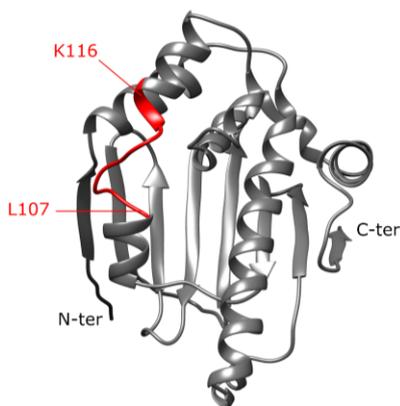
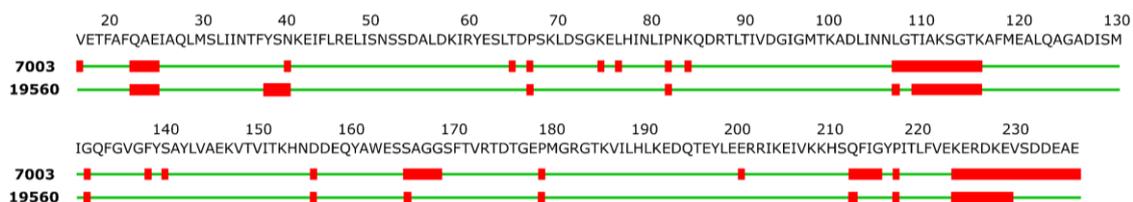


Figure 2.2 : Ci-dessus : séquence de HSP90 α -NTD. Les résidus pour lesquels les déplacements chimiques ^{15}N et H_N ont été attribués sont représentés par un trait vert pour chacune des deux attributions reportées (7003 et 19560). Ceux non attribués sont représentés par un segment rouge. A gauche : structure cristallographique de HSP90-NTD (PDB : 1YES). Le segment rouge correspond aux signaux du squelette invisibles par RMN en raison de la dynamique dans la gamme de temps μs -ms et appartenant au segment [L107-K116], localisé à proximité du site de fixation de l'ATP.

2.3 Intérêt des groupements méthyles

Du fait de la dynamique dans le régime intermédiaire la totalité des résidus amides du segment [L107-K116] est invisible. Les groupements méthyles ont, quant à eux, des propriétés de relaxation transversale et longitudinale favorables dues à leur rotation rapide autour de l'axe C-C et à leur flexibilité (Plevin and Boisbouvier 2012; Wiesner and Sprangers 2015). Cette dernière étant principalement observée lorsque les groupements méthyles sont situés au bout des chaînes latérales de long acides aminés. De plus, les groupements méthyles permettent d'obtenir des spectres sensibles grâce à leurs trois protons équivalents non échangeables. En comparant une expérience BEST-TROSY (^{15}NH), acquise à 20 °C sur un spectromètre opérant à une fréquence ^1H de 700 MHz avec un échantillon de HSP90-NTD à 0.5 mM marqué U- $[^2\text{H}, ^{15}\text{N}, ^{13}\text{C}]$ dont les protons amides ont été reprotonés, avec une HMQC-SOFAST ($^{13}\text{CH}_3$) acquise à 20 °C sur un spectromètre opérant à une fréquence ^1H de 700 MHz et avec un échantillon d'HSP90-NTD de concentration similaire marqué U- $[^2\text{H}, ^{15}\text{N}, ^{12}\text{C}]$, Ala- $[^{13}\text{C}^1\text{H}_3]^\beta$, Met- $[^{13}\text{C}^1\text{H}_3]^\epsilon$, Leu/Val- $[^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}$, Ile- $[^{13}\text{C}^1\text{H}_3]^{\delta 1}$, Thr- $[^{13}\text{C}^1\text{H}_3]^\gamma$, il a pu être constaté que l'expérience acquise sur les groupement méthyles était en moyenne 7 fois plus sensible que l'expérience BEST-TROSY (^{15}NH) (Figure 2.3). Il est à noter que chaque expérience a été acquise avec des conditions optimales et le même temps expérimental. Ainsi, les groupements méthyles se révèlent être des sondes idéales pour pallier le problème de dynamique dans la gamme de temps μs -ms et étudier HSP90-NTD.

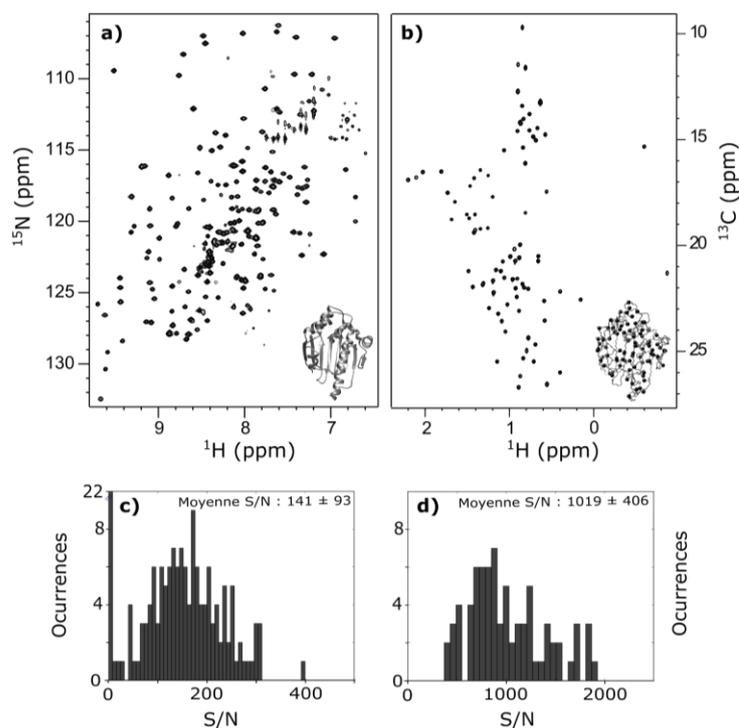


Figure 2.3 : a) Spectre 2D ^1H - ^{15}N BEST-TROSY de HSP90-NTD Apo. En bas à droite : structure cristallographique en rubans de HSP90-NTD (PDB : 1YES). b) Spectre méthyle ^1H - ^{13}C SOFAST-HMQC de HSP90-NTD. En bas à droite : structure cristallographique de HSP90-NTD, les sphères représentent les groupements méthyles. Les histogrammes c) et d) représentent le nombre d'occurrences en fonction d'une valeur de rapport signal/bruit (S/N), représentée avec un pas

respectivement de 10 et de 60 pour les signaux NH et CH_3 . Les spectres ont été acquis sur un spectromètre opérant à une fréquence ^1H de 700 MHz pendant une heure chacun en utilisant des échantillons de HSP90-NTD à 0.5 mM.

Les groupements méthyles sont également des sondes importantes qui apportent de nombreuses informations sur la protéine d'étude. En effet, lorsque utilisés dans une protéine perdeutérée, ils permettent de détecter des NOEs convertibles en contraintes de distance allant jusque 12 Å (Sounier et al. 2007; Ayala et al. 2020). Ainsi, en combinaison avec d'autres techniques de biologie structurale telles que la cryomicroscopie électronique (Gauto et al. 2019) ou la diffusion de rayons-X/neutrons aux petits angles (SAXS/SANS) (Lapinaite et al. 2013) il est possible d'obtenir des structures de complexes de plusieurs centaines de kDa. Mais grâce à ces sondes il est également possible d'obtenir à la fois des informations fonctionnelles (Rosenzweig et al. 2013; Mas et al. 2018) et des renseignements sur la dynamique moléculaire d'assemblages allant jusqu'à 1 MDa (Sprangers and Kay 2007; Mas et al. 2018).

Les protéines sont composées de six acides aminés comportant un (ou plusieurs) groupement(s) méthyle(s) : Alanine, Thréonine, Valine, Isoleucine, Méthionine et Leucine. Ces derniers représentent environ 40% des résidus de la protéine et sont localisés dans les cœurs hydrophobes et à la surface des sites d'interaction protéine/protéine (à l'exception des thréonines, possédant un groupement -OH

hydrophile, pouvant également se trouver accessibles au solvant et au niveau des sites d'interaction chargés) (Miller et al. 1987; Lo Conte et al. 1999). Dans le cas de HSP90-NTD, ces acides aminés méthylés font exactement 87/223 résidus soit 39% de la protéine et cinq d'entre eux (L107, T109, I110, A111, T115) se trouvent dans la boucle à proximité du site de fixation de l'ATP, pour laquelle les signaux des groupements amides du squelette sont invisibles par RMN.

2.4 Marquage combinatoire des groupements méthyles de HSP90-NTD

Afin de maximiser le nombre de sondes RMN disponibles pour étudier HSP90-NTD, un marquage combinatoire de différents types de méthyles a été mis en place. Cependant, il faut noter que les groupements méthyles résonnent dans une fenêtre spectrale étroite de l'ordre de 2 ppm en proton et 20 ppm en carbone sur un spectre 2D- $[^1\text{H}, ^{13}\text{C}]$ -HMQC (Figure 2.4).

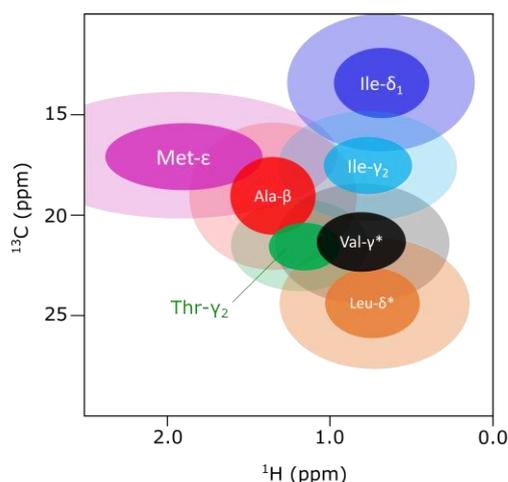


Figure 2.4: Fenêtre spectrale définissant les déplacements chimiques spécifiques de chaque type de groupement méthyle dans un spectre 2D $[^1\text{H}-^{13}\text{C}]$. Les diamètres des ellipses colorées et semi-transparentes représentent respectivement les déviations standard premières et secondes. Figure adaptée de (Plevin and Boisbouvier 2012).

Pour limiter le recouvrement des signaux RMN et obtenir des spectres bien résolus tout en gardant un nombre de sondes maximal, un choix de marquage a été effectué notamment pour les résidus comportant deux groupements méthyles (isoleucine, valine et leucine). Ainsi, un marquage stéréospécifique pro-S (Gans et al. 2010) a été mis en place pour les valines et leucines (également appelé respectivement γ_2 et δ_2 pour les valines et leucines) et un marquage régio-spécifique a été choisi pour les isoleucines (δ_1). Finalement la protéine étudiée est marquée : U- $[^2\text{H}, ^{15}\text{N}, ^{12}\text{C}]$, $^{13}\text{CH}_3$ – (Ile- δ_1 , Leu- δ_2 , Val- γ_2 , Ala- β , Met- ϵ , Thr- γ_2) en se basant sur le protocole décrit par Kerfah et al. (Kerfah et al. 2015c).

86 signaux sont observables sur un spectre 2D- ^1H - ^{13}C -HMQC de la protéine Apo, suggérant que seulement deux des groupements méthyles du domaine N-terminal d'HSP90, comportant 87 résidus portant un groupement méthyle marqué, sont potentiellement superposés (Figure 2.5a). On peut donc, grâce aux sondes méthyles, obtenir des informations sur la boucle préalablement invisible située à proximité du site de fixation de l'ATP. Afin d'identifier à quel type de groupement méthyle (Ala, Ile, Leu, Val, Met ou Thr) appartient chaque signal observé sur le spectre, HSP90-NTD a été produite avec différents marquages marquant spécifiquement un seul type de méthyle à la fois, sauf pour les leucines et valines : Ile- δ_1 , Ala- β , Met- ϵ , Thr- γ_2 , [Leu/Val] $^{\text{pro-S}}$ et Val $^{\text{pro-S}}$ (Figure 2.5b/c/d/e/f/g). Ces informations seront par la suite précieuses pour l'attribution des différents méthyles.

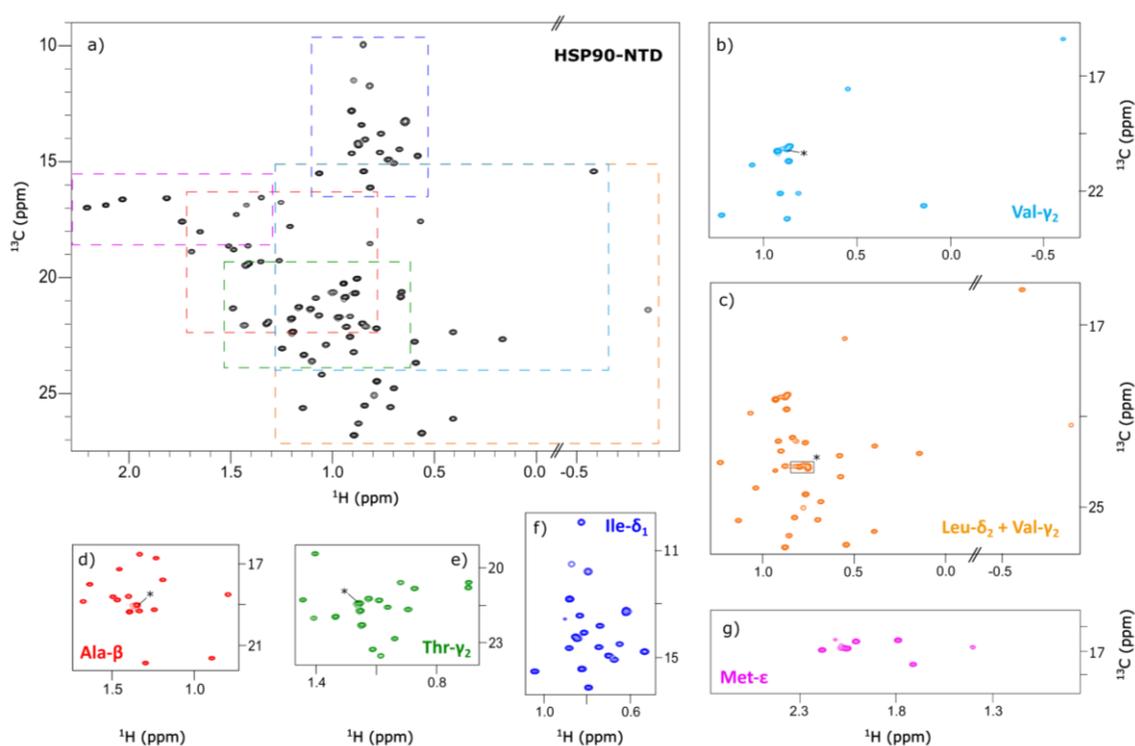


Figure 2.5 : a) Spectre 2D ^1H - ^{13}C -HMQC du domaine N-terminal d'HSP90. b-g) Spectres 2D ^1H - ^{13}C de HSP90-NTD marquée spécifiquement sur un(deux) type(s) de groupement(s) méthyle(s) respectivement les valines pro-S, les leucines et les valines pro-S simultanément, les alanines- β , les thréonines- γ_2 , les isoleucines- δ_1 et les méthionines- ϵ . Ceci permet d'identifier à quel type de résidu appartient chaque signal. Les pics annotés avec le symbole * sont des pics appartenant au tag poly histidine N-terminal non clivé pour les échantillons marqués spécifiquement sur un unique type de groupement méthyle (Henot et al. 2021).

2.5 Stratégies d'attribution des groupements méthyles d'une protéine

Comme évoqué précédemment, les groupements méthyles sont des sondes qui se révèlent cruciales pour des études à la fois structurales et dynamiques mais qui ne sont utiles que lorsqu'elles sont attribuées. Plusieurs méthodes ont été proposées dans la littérature pour résoudre cet obstacle majeur qu'est l'attribution des groupements méthyles. Différentes stratégies d'attribution peuvent être envisagées et les principales sont décrites ci-dessous.

2.5.1 Transfert d'attribution du squelette aux groupements méthyles

2.5.1.1 Contexte

A l'origine, des expériences RMN basées sur le couplage J_{H-H} telles que les expériences COSY (COrrélation SpectroscopY), COSY relayées et TOCSY (TOtal Correlation SpectroscopY) étaient utilisées pour attribuer les protons des chaînes latérales des protéines. Avec le développement du marquage isotopique et l'accès aux protéines marquées ^{13}C , les expériences RMN ont évoluées et les expériences ^1H - ^1H TOCSY/COSY/COSY-relayées ont été remplacées par des expériences ^{13}C - ^{13}C TOCSY/COSY/COSY-relayées où l'aimantation était transférée des chaînes latérales aux atomes déjà attribués du squelette. Les expériences ^{13}C - ^{13}C TOCSY/COSY/COSY-relayées présentent l'avantage d'un transfert d'aimantation plus efficace dû à une constante de couplage ($J_{C-C} \sim 35$ Hz) significativement plus grande que dans le cas des expériences ^1H - ^1H et qui, de plus, ne dépend pas de la conformation amenant donc à des fonctions de transfert de cohérences bien caractérisées (Eaton et al. 1990; Kerfah et al. 2015a). Ainsi, avec des expériences telles que la 3D-HCCH-TOCSY des chaînes latérales entières pouvaient être décrites, comme démontré par Bax et al. pour l'interleukin-1 β (17 kDa) (Bax et al. 1990). Toutefois, pour les protéines de poids moléculaire plus élevé (typiquement > 40 kDa), la sensibilité et la résolution deviennent des paramètres critiques.

Stratégies de marquage

Pour pallier ce problème de sensibilité et résolution, de nouveaux précurseurs protonant spécifiquement les groupements méthyles dans une protéine entièrement deutérée ont été proposés. Les premiers précurseurs introduits pour marquer les isoleucines (U- ^2H , ^{15}N , ^{13}C]-Ile- $^{13}\text{C}^1\text{H}_3$ $^{\delta 1}$) ou les leucines et valines (U- ^2H , ^{15}N , ^{13}C]-Leu/Val- $^{13}\text{C}^1\text{H}_3$ $^{\delta 1}$ / $^{13}\text{C}^1\text{H}_3$ $^{\delta 2}$) étaient deux keto-acides, respectivement : l' α -ketobutyrate (Gardner et al. 1997) et l' α -ketoisovalérate (Goto et al. 1999; Hajduk et al. 2000; Gross et al. 2003). Cependant, les chaînes latérales ramifiées marquées ^{13}C des

résidus valine, isoleucine et leucine compromettent, même avec ces nouveaux précurseurs, l'efficacité des approches TOCSY et COSY par la présence de deux chemins de cohérence.

Une solution présentée par Tugarinov et Kay (Tugarinov and Kay 2003a) a été de linéariser les acides aminés leucine et valine en employant un α -ketoisovalérate non-stéréospécifique ($^{13}\text{C}^1\text{H}_3/^{12}\text{C}^2\text{H}_3$) (Figure 2.6). Ainsi, le transfert d'aimantation dans les expériences TOCSY/COSY est plus efficace. Par contre, 50 % des groupements méthyles sont marqués pro-S ($[^{13}\text{C}^1\text{H}_3]^{\delta 2}/[^{12}\text{C}^2\text{H}_3]^{\delta 1}$) et l'autre moitié pro-R ($[^{13}\text{C}^1\text{H}_3]^{\delta 1}/[^{12}\text{C}^2\text{H}_3]^{\delta 2}$). La présence des groupements méthyles marqués pro-S et pro-R entraîne une forte augmentation de la probabilité de recouvrement des signaux et donc une baisse de la résolution de l'expérience. De plus, la sensibilité du signal est divisée par deux.

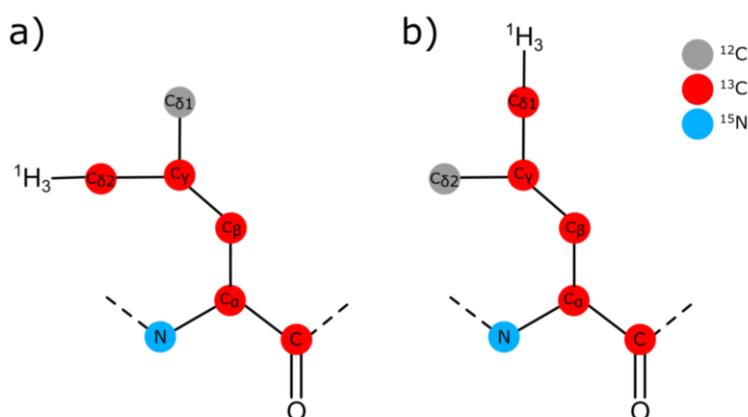


Figure 2.6 : Représentation schématique de la linéarisation de la leucine par un marquage utilisant l' α -ketoisovalérate non-stéréospécifique. 50 % des résidus leucine sont marqués U - $[^2\text{H},^{15}\text{N},^{13}\text{C}]$ - Leu - $[^{13}\text{C}^1\text{H}_3]^{\delta 2}/[^{12}\text{C}^2\text{H}_3]^{\delta 1}$ et les 50 % restants sont marqués U - $[^2\text{H},^{15}\text{N},^{13}\text{C}]$ - Leu - $[^{13}\text{C}^1\text{H}_3]^{\delta 1}/[^{12}\text{C}^2\text{H}_3]^{\delta 2}$. Les sphères rouges, grises et bleues représentent respectivement des atomes ^{13}C , ^{12}C et ^{15}N . Les atomes ^2H ne sont pas représentés.

Pour les isoleucines, le marquage linéaire des Ile- δ_1 (i.e. sans marquer les Ile- γ_2) n'est alors pas décrit (en 2003). En effet, le marquage linéaire des Ile- δ_1 implique un marquage ^{13}C de tous les carbones du C α au C δ de la chaîne or, le C β , contrairement aux trois autres atomes de carbone de la chaîne linéaire ne peut pas être marqué en utilisant le précurseur de l'isoleucine (l' α -ketobutyrate ^{13}C). Cela s'explique par le fait que les atomes C β et C γ_2 proviennent du pyruvate lui-même découlant de la glycolyse. Ainsi

pour marquer l'atome C β , l'addition de glucose [^2H , ^{13}C] est nécessaire et cela résulte par le marquage simultané de l'atome C γ_2 (Figure 2.7).

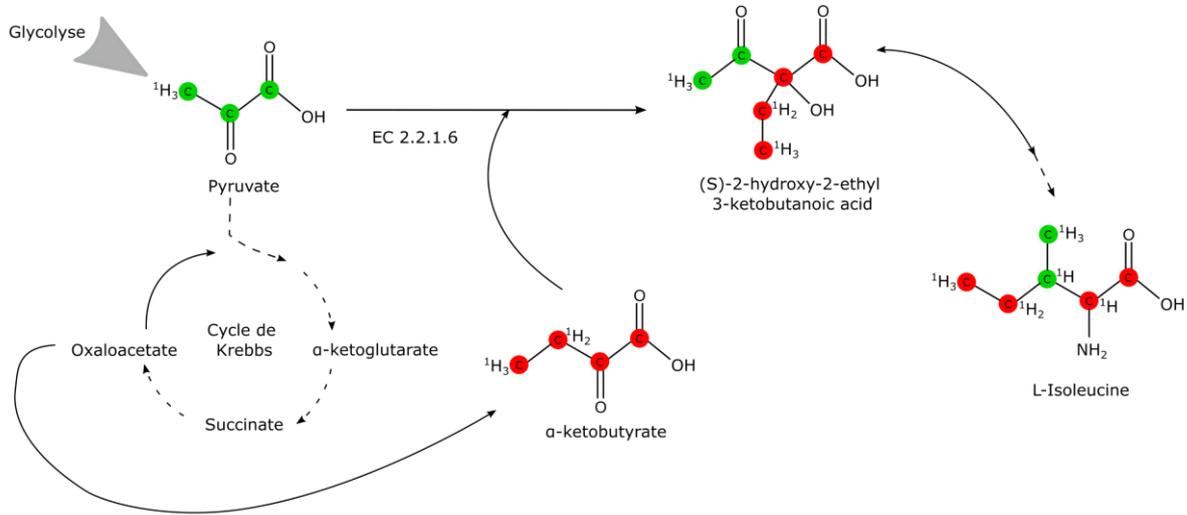


Figure 2.7 : Schéma représentant le chemin biosynthétique simplifié de l'isoleucine. Les sphères rouges et vertes représentent des atomes ^{13}C . Les carbones en vert proviennent directement du pyruvate et ceux en rouge de l' α -ketobutyrate. Les flèches en pointillés représentent des réactions en multi-étapes. Adapté de (Kerfah et al. 2015b).

Développement de séquences d'impulsions

Avec la disponibilité des marquages précédemment décrits, les avancées significatives dans l'attribution du squelette des assemblages moléculaires et, dans le but d'obtenir une attribution complète des groupements méthyles pour les protéines de haut poids moléculaires y compris celle des acides aminés leucine, valine et isoleucine, deux nouvelles séquences d'impulsions ont été développées par Tugarinov et Kay (Tugarinov and Kay 2003a). Pour employer ces méthodes afin d'attribuer les groupements méthyles de protéines ou d'assemblages de haut poids moléculaires, les résonances des signaux du squelette déjà attribuées sont utilisées. Pour cela, comme pour les expériences ^{13}C - ^{13}C TOCSY ou COSY, les groupements méthyles doivent impérativement être connectés aux atomes du squelette via une chaîne d'atomes de carbone ^{13}C .

La première séquence repose sur un transfert d'aimantation unidirectionnel des groupements méthyles aux groupements amides du squelette où l'aimantation sera détectée (HM)CM(CGBCA)NH (Figure 2.8a). Le transfert débute par les groupements méthyles car, grâce à leurs trois protons équivalents, l'aimantation de Boltzmann est plus importante. Cette méthode possède l'avantage de disposer d'une excellente

dispersion des déplacements chimiques dans les dimensions H_N et ^{15}N . Toutefois, il est à noter qu'une faible sensibilité de la corrélation entre les groupements méthyles et les NHs peut être observée dans le cas de protéines de haut poids moléculaire. La seconde méthode appelée « out and back - HCC-COSY » consiste à transférer l'aimantation depuis les groupements méthyles de la chaîne latérale jusqu'aux carbones du squelette et à revenir détecter l'aimantation sur les protons des groupements méthyles (Figure 2.8b).

Avec ces méthodes, en possédant une attribution du squelette de la protéine d'intérêt ainsi que d'un marquage spécifique permettant de connecter par une chaîne d'atomes ^{13}C les groupement méthyles au squelette, une attribution quasi complète des 276 groupements Ile- δ_1 , Val- γ^* et Leu- δ^* a pu être effectuée pour la Malate Synthase G (MSG) une protéine de 82 kDa (Tugarinov and Kay 2003a; Ayala et al. 2009). Il est à noter que dans le cas des isoleucines la non linéarité du système de spin compromet l'efficacité du transfert d'aimantation. Une impulsion sélective est alors employée dans la séquence d'impulsions afin d'empêcher l'aimantation d'évoluer vers l'atome de carbone Ile- γ_2 et de gagner en sensibilité en exploitant la dispersion des résonances carbone dans la chaîne de l'isoleucine (Tugarinov and Kay 2003b). Cependant malgré ces nouvelles séquences d'impulsions les spectres sont toujours très encombrés en raison du marquage utilisant l' α -ketoisovalérate non-stéréospécifique ($^{13}C^1H_3/^{12}C^2H_3$) pour le marquage des leucines et des valines.

Pour la suite de ma thèse, les expériences « out and back – HCC COSY » seront appelées « out and back – HCC » si elles permettent d'accéder au $C_{(i+1)}$ en partant du $H_{(i)}$, « out and back – HC(C)C » si elles permettent d'accéder au $C_{(i+2)}$ en partant du $H_{(i)}$ et « out and back – HC(CC)C » si elles permettent d'accéder au $C_{(i+3)}$ en partant du $H_{(i)}$. « Out and back – HC(C)C » et « out and back – HC(CC)C » sont des expériences COSY relayées. Les expériences « out and back » pouvant également être appliquées aux expériences ^{13}C - ^{13}C - TOCSY j'emploierai la notation « out and back » quand il pourra s'agir aussi bien de « out and back – HCC COSY » que de « out and back – ^{13}C - ^{13}C TOCSY ».

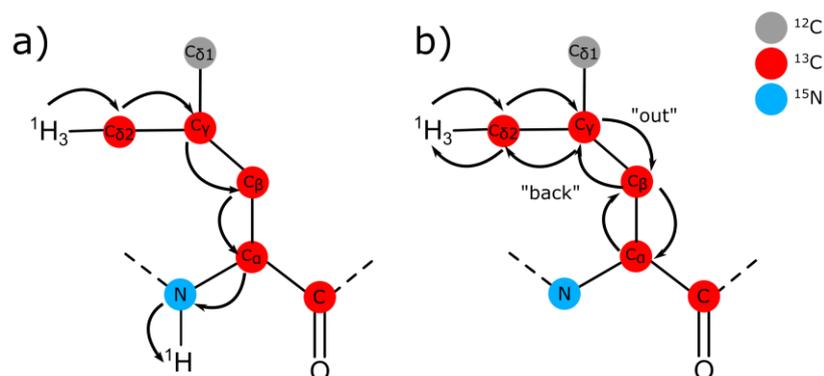


Figure 2.8 : Schémas représentant les différentes étapes du transfert d'aimantation intra-résidu, illustré pour une leucine marquée $U\text{-}[^2\text{H}, ^{15}\text{N}, ^{13}\text{C}]\text{-Leu-}[^{13}\text{C}^1\text{H}_3]^{62}/[^{12}\text{C}^2\text{H}_3]^{61}$ représentant 50 % des leucines de l'échantillon. a) pour la séquence d'impulsions (HM)CM(CGBCA)NH. b) pour la séquence d'impulsions « out and back HCC-COSY ». Figure adaptée de (Tugarinov and Kay 2003a). Les sphères rouges, grises et bleues représentent respectivement des atomes ^{13}C , ^{12}C et ^{15}N . Les atomes ^2H ne sont pas représentés.

2.5.1.2 Développement des stratégies de marquage depuis les années 2010

Leucines et Valines pro-R

Afin de limiter la superposition dans les spectres et faciliter l'étude des assemblages de haut poids moléculaire, un acétolactate marqué méthyle a été développé comme précurseur alternatif des leucines et valines (Gans et al. 2010). Ce dernier permet de marquer stéréo-spécifiquement les groupements méthyles $^{13}\text{C}^1\text{H}_3\text{-pro-S}$ (i.e. δ_2, γ_2) des leucines et valines (Figure 2.9). Un tel marquage divise par deux le nombre de pics observés dans le spectre tout en améliorant d'un facteur deux la sensibilité du spectre et en augmentant d'un facteur 4 l'intensité des pics de corrélation NOE observés lors de l'acquisition d'une expérience NOESY.

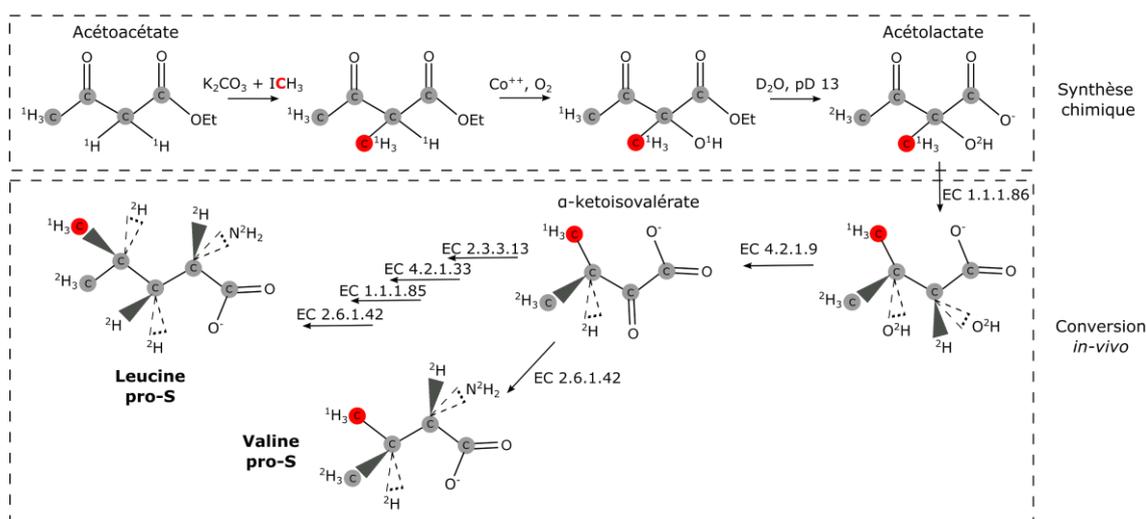


Figure 2.9 : Schéma réactionnel du protocole pour la production d'échantillons marqués $U\text{-}[^2\text{H}, ^{15}\text{N}]\text{-Leu/Val-}[^{13}\text{C}^1\text{H}_3]^{pro-S}$ adapté de (Gans et al. 2010). Les sphères grises et rouges représentent respectivement les atomes de carbone ^{12}C et ^{13}C . L'enzyme responsable de la catalyse de chaque réaction est indiquée par son numéro EC. EC 1.1.1.86 : ketol-acid-reductoisomerase, EC 4.2.1.9 : dihydroxy-acid hydratase, EC 2.3.3.13 : 2-isopropylmalate synthase, EC 4.2.1.33 3-isopropylmalate hydratase, EC 1.1.1.85 : 3-isopropylmalate dehydrogenase et EC 2.6.1.42 : branched-chain amino acid aminotransferase.

S'inspirant de ce nouveau précurseur un marquage linéaire ^{13}C reliant les groupements pro-R (i.e. δ_1 , γ_1) des leucines et valines au squelette a été mis en place (Figure 2.10) en marquant ^{13}C les atomes de carbone de l'acétoacétate (Mas et al. 2013).

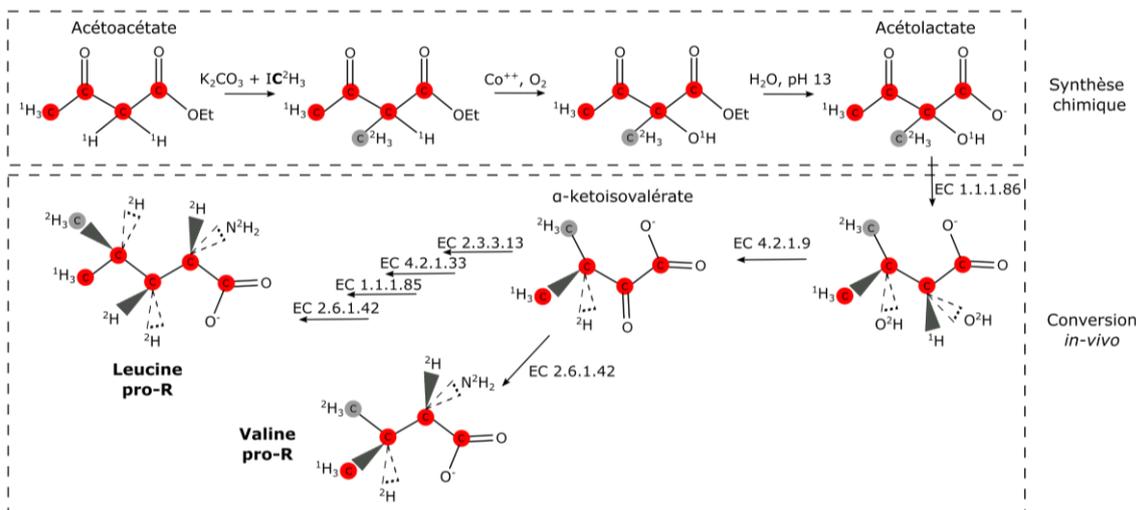


Figure 2.10 : Schéma réactionnel du protocole pour la production d'échantillons marqués U-[^2H , ^{13}C , ^{15}N]-Leu/Val-[[$^{13}\text{C}^1\text{H}_3$] $^{\text{proR}}$]/[[$^{12}\text{C}^2\text{H}_3$] $^{\text{proS}}$]. Les sphères grises et rouges représentent respectivement les atomes de carbone ^{12}C et ^{13}C . L'enzyme responsable de la catalyse de chaque réaction est indiquée par son numéro EC. Un milieu de culture en M9- D_2O est supposé.

Isoleucines γ_2 et δ_1

On notera également le développement d'une méthode pour le marquage des isoleucines, marquées ^{13}C sur le squelette ainsi que [$^{13}\text{C}^1\text{H}_3$] sur le groupement méthyle γ_2 mais [$^{12}\text{C}^2\text{H}_2$] et [$^{12}\text{C}^2\text{H}_3$] respectivement sur les groupements méthylène γ_1 et méthyle δ_1 (Figure 2.11). Ce marquage nécessite l'incorporation du précurseur 2-hydroxy-2-[[$^2\text{H}_5$]-ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate dans les cultures obtenu par synthèse chimique (Ayala et al. 2012).

Une méthode pour le marquage des isoleucines marquées ^{13}C sur le squelette ainsi que [$^{13}\text{C}^1\text{H}_3$] sur le groupement méthyle δ_1 mais [$^{12}\text{C}^2\text{H}_3$] sur le groupement méthyle γ_2 , résolvant ainsi le problème de linéarité du système de spin soulevé précédemment, a également été développée (Figure 2.12). Ce marquage nécessite l'introduction du précurseur : 2-hydroxy-2-(1',1'-[[$^2\text{H}_2$],1',2'-[[$^{13}\text{C}_2$]])-ethyl-1,2,3-[[$^{13}\text{C}_3$],4,4,4-[[$^2\text{H}_3$]]-3-oxobutanoate dans les cultures, ce dernier étant obtenu par synthèse enzymatique, par condensation du 3,3-[[$^2\text{H}_2$],1,2,3,4-[[$^{13}\text{C}_4$]]-2-ketobutyrate avec le 2-[[^{13}C]-3,3,3-[[$^2\text{H}_3$]]-pyruvate en utilisant l'aceto-hydroxy-acid synthase II (AHAS II) (Kerfah et al. 2015b).

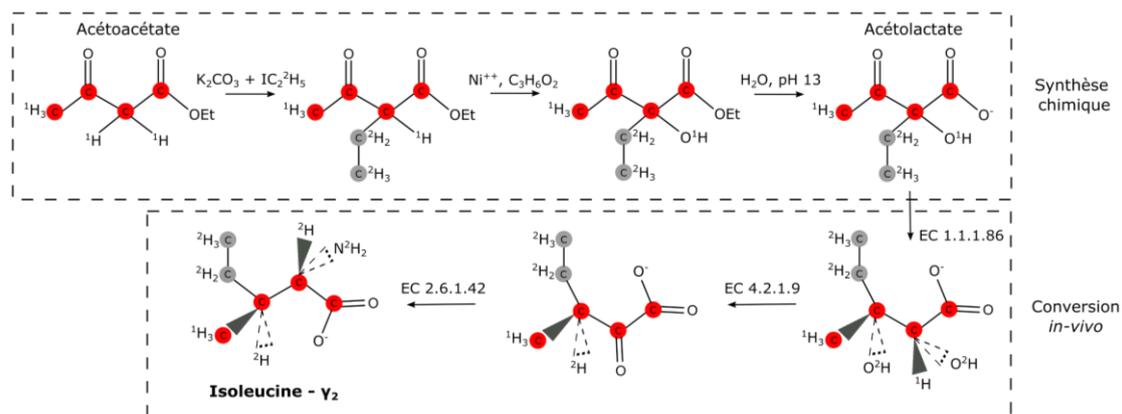


Figure 2.11 : Schéma réactionnel du protocole pour la production d'échantillons marqués U-[^2H , ^{15}N], [C' , $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma_2 - ^{13}\text{C}_4$], [$\gamma_2 - ^1\text{H}_3$]-Isoleucine. Les sphères grises et rouges représentent respectivement les atomes de carbone ^{12}C et ^{13}C . L'enzyme responsable de la catalyse de chaque réaction est indiquée par son numéro EC. Un milieu de culture en M9- D_2O est supposé.

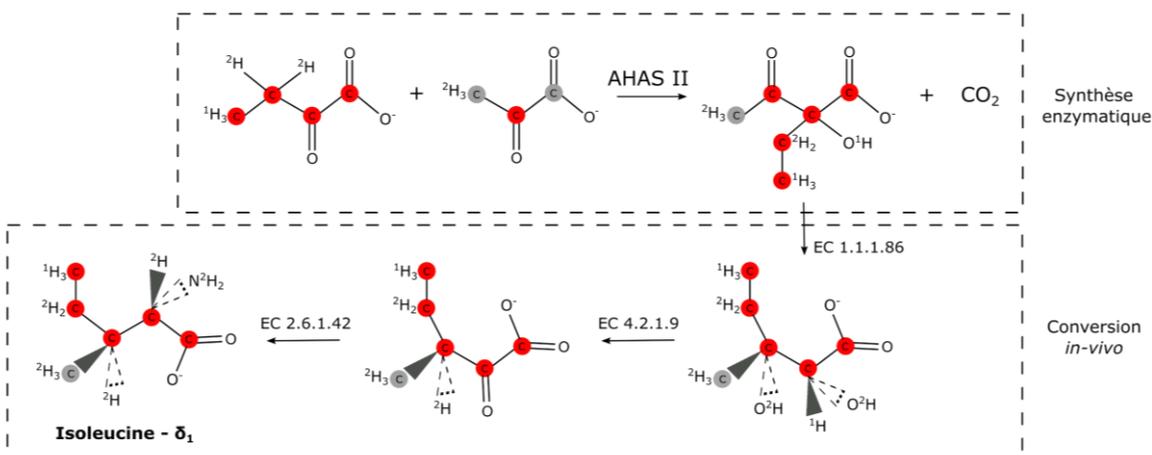


Figure 2.12 : Schéma réactionnel du protocole pour la production d'échantillons marqués U-[^2H , ^{15}N], [C' , $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma_1$, $\text{C}\delta_1 - ^{13}\text{C}_5$], [$\delta_1 - ^1\text{H}_3$]-Isoleucine. Les sphères grises et rouges représentent respectivement les atomes de carbone ^{12}C et ^{13}C . L'enzyme responsable de la catalyse de chaque réaction est indiquée par son numéro EC. Un milieu de culture en M9- D_2O est supposé.

Ainsi, un marquage linéaire ^{13}C liant les groupements méthyles aux atomes du squelette est maintenant disponible à la fois pour les leucines et valines pro-R et les isoleucines (γ_2 et δ_1). Additionnellement, ce marquage est stéréospécifique pour les leucines et valines connectant par des atomes ^{13}C les groupements méthyles pro-R au squelette et régiosélectif pour les isoleucines. Cela permet dans le cas des leucines et valines

d'obtenir des spectres plus sensibles et moins encombrés. Pour les isoleucines, il est maintenant possible d'attribuer les groupements méthyles δ_1 sans l'impulsion sélective employée précédemment pour empêcher l'aimantation d'évoluer vers l'atome de carbone Ile- γ_2 .

Ainsi, grâce à un marquage 100 % stéréospécifique ou 100 % régiosélectif avec une chaîne linéaire d'atomes ^{13}C liant les groupements méthyles au squelette, différentes stratégies ont été proposées et appliquées sur MSG (82 kDa) et TET2 (468 kDa) pour attribuer les groupements méthyles des isoleucines δ_1 et les groupements méthyles pro-*R* des leucines et valines (Mas et al. 2013; Kerfah et al. 2015a). On note qu'avec le marquage stéréospécifique des groupements méthyles des leucines et valines il est possible d'étendre l'attribution à des protéines de 500 kDa (Mas et al. 2013).

2.5.1.3 Leucines et Valine pro-S : problématique

Toutefois, le marquage des groupements méthyles pro-*S* des leucines et valines, à la fois moins cher et amenant à des résonances mieux dispersées (London et al. 2008), est souvent choisi à la place du marquage méthyle pro-*R* pour un échantillon marqué uniquement sur les groupements méthyles. Nous avons donc cherché à transférer l'attribution du squelette aux groupements méthyles pro-*S* des leucines et valines.

Avec les marquages disponibles, il est déjà possible d'attribuer les groupements méthyles pro-*S* des leucines et valines d'une protéine. Cependant, il faut utiliser deux marquages différents. Pour cela il est nécessaire soit d'utiliser deux échantillons : le premier avec un premier type de marquage afin de lier le méthyle pro-*R* linéairement aux atomes ^{13}C de la chaîne du squelette et le second avec un autre type de marquage pour lier le méthyle pro-*S* au méthyle pro-*R* soit à travers l'espace en utilisant une expérience NOESY (Sprangers and Kay 2007) ou au travers de liaisons covalentes en utilisant les expériences « out and back » (Mas et al. 2013) (Figure 2.13a). Autrement, il est possible d'utiliser un unique échantillon combinant les deux marquages à 50 % et d'employer une séquence d'impulsions « out and back » (Kerfah et al. 2015a) (Figure 2.13b). Le choix de l'emploi de l'une ou l'autre des techniques (avec un ou deux échantillons) s'effectue généralement suivant la protéine, l'emploi de la méthode avec un unique échantillon divisant par deux le rapport signal sur bruit.

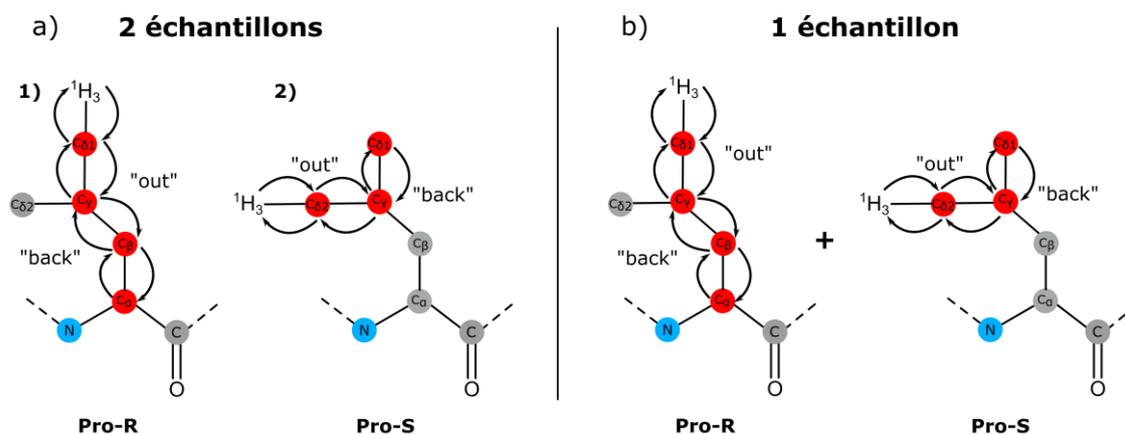


Figure 2.13 : Techniques de transfert d'attribution du squelette aux groupements méthyles pro-S (δ_2). Exemples schématiques pour une leucine a) 1) Sur le premier échantillon, un transfert employant une expérience « out-and-back HC(C)C » permet de lier le groupement méthyle pro-R au C_β du squelette déjà attribué et un transfert avec une expérience « out-and-back HC(CC)C » permet de lier le groupement pro-R au C_α du squelette également attribué. Sur le second échantillon, le groupement méthyle pro-R, attribué grâce aux expériences acquises avec le premier échantillon, va être lié au groupement pro-S avec une expérience « out and back HC(C)C » et permettre ainsi l'attribution de ce dernier. b) Un seul échantillon est nécessaire. Le transfert s'effectue avec une expérience « out-and-back HC(C)C » et une expérience « out-and-back HC(CC)C ». Ainsi il est possible de corrélérer le groupement pro-R (δ_1) de la leucine au squelette déjà attribué et dans le même temps le groupement pro-R au groupement pro-S. Les sphères rouges, grises et bleues représentent respectivement des atomes ^{13}C , ^{12}C et ^{15}N . Les atomes ^2H ne sont pas représentés. Sur cette figure la méthode d'attribution est montrée avec une expérience « out and back - COSY » mais de la même manière elle pourrait se faire avec une expérience « out and back – HCC TOCSY ».

L'inconvénient de la première méthode décrite réside dans la production de deux échantillons différents augmentant à la fois les coûts de production et le temps de préparation. La seconde méthode quant à elle ne nécessite qu'un unique échantillon. Ce dernier est constitué de 50% de protéine marquée ^{13}C sur le groupement méthyle pro-R ainsi que ^{13}C sur la chaîne de carbones reliant le méthyle pro-R au squelette et à 50 % de protéine marquée ^{13}C sur les deux groupements méthyles pro-R et pro-S ainsi que sur le carbone précédent les groupements méthyles. Le marquage ainsi décrit est possible par l'ajout simultané d'un même précurseur (l'acétolactate) marqué différemment lors de la production de la protéine. Ce dernier se trouve sous les formes : 2 hydroxy- 2-[$^2\text{H}_3$]-methyl-3-keto-1,2,3,4-[$^{13}\text{C}_4$] acide butanoïque (Figure 2.14 (1)) et 2-hydroxy- 2-[^{13}C]-methyl- 3-keto -3,4-[$^{13}\text{C}_2$]- 4,4,4-[$^2\text{H}_3$] acide butanoïque (Figure 2.14 (2)).

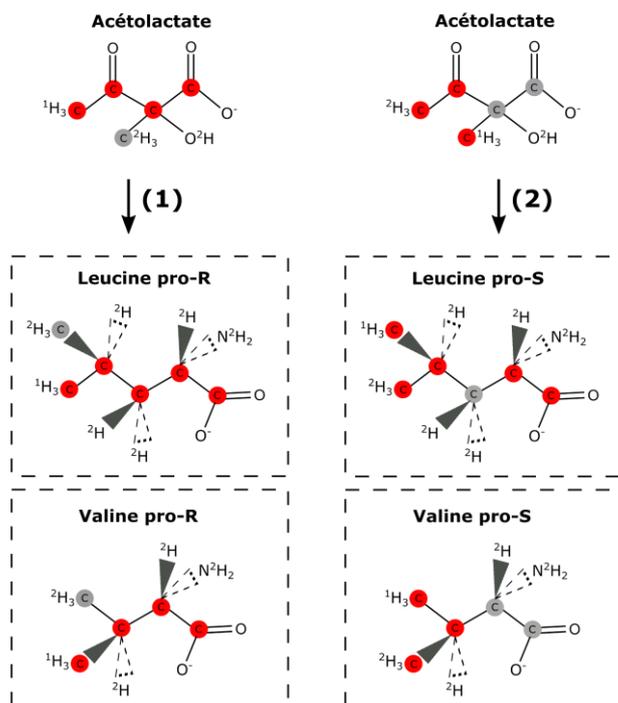


Figure 2.14 : Précurseurs et résidus marqués (résultant de l'ajout des précurseurs dans le milieu de culture) pour les leucines et les valines. Le précurseur (1) est le 2 hydroxy- 2- $[\text{}^2\text{H}_3]$ -methyl-3-keto-1,2,3,4- $[\text{}^{13}\text{C}_4]$ acide butanoïque. Le précurseur (2) est le 2-hydroxy- 2- $[\text{}^{13}\text{C}]$ -methyl- 3-keto -3,4- $[\text{}^{13}\text{C}_2]$ -4,4,4- $[\text{}^2\text{H}_3]$ acide butanoïque. Les sphères grises et rouges représentent respectivement les atomes de carbone ^{12}C et ^{13}C . Deux atomes ^{13}C des leucines (C' et Cα) proviennent du glucose U- $[\text{}^2\text{H}, \text{}^{13}\text{C}]$ ajouté dans la culture.

Toutefois, il est à noter que les groupements méthyles ne sont pas protonés de la même façon selon les deux types de marquages présents dans l'échantillon. Ainsi, dans le cas de la méthode employant un unique échantillon, des déplacements chimiques isotopiques de l'ordre de 0.8 ppm vont être présents lors du transfert d'attribution du groupement pro-R au groupement pro-S, provenant des différences magnétiques entre le ^1H et le ^2H , et compliquent l'analyse. Ces déplacements chimiques isotopiques peuvent être corrigés en re-référençant le second spectre dans la méthode employant deux échantillons pour transférer l'attribution des groupements méthyles pro-R aux groupements méthyles pro-S.

Le transfert de l'attribution des groupements méthyles pro-R aux groupements méthyles pro-S dans le cas de la méthode employant deux échantillons aurait également pu se faire à travers l'espace en utilisant une expérience NOESY (Sprangers and Kay 2007). Cela aurait nécessité le même premier échantillon que celui décrit précédemment (Figure 2.13a/1). Concernant le second échantillon le marquage aurait légèrement différé puisque le carbone précédant les deux groupements méthyles aurait été ^{12}C et les deux groupements méthyles auraient dû être protonés dans le but d'observer des pics de corrélation entre les groupements méthyles pro-R et pro-S d'un même résidu.

2.5.1.4 Développement d'un nouveau précurseur pour faciliter le transfert d'attribution du squelette aux groupements méthyles pro-S des leucines et valines

Afin de simplifier le transfert d'attribution du squelette aux groupements méthyles pro-S des leucines et valines, nous avons cherché à produire un échantillon connectant directement le groupement méthyle $^{13}\text{C}^1\text{H}_3$ -pro-S des leucines et des valines aux atomes du squelette par une chaîne linéaire de carbone ^{13}C , tout en marquant $^{12}\text{C}^2\text{H}_3$ le groupement méthyle pro-R afin d'optimiser la sensibilité de l'expérience. Pour cela, il est nécessaire d'employer un précurseur acétolactate marqué : le 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O ^2H]-3-oxo-4,4,4- $^2\text{H}_3$ - butanoate (Figure 2.15). Or, ce dernier ne peut pas être obtenu à partir du schéma de synthèse décrit précédemment (Figure 2.9) puisque le marquage ^{13}C des trois premiers carbones de l'acétoacétate n'est pas possible en utilisant des matériaux commercialement disponibles. L'acétoacétate étant obtenu par condensation de deux acétates, différents motifs de marquage ^{13}C peuvent être obtenus (Figure 2.15) mais pas celui souhaité. Grâce à une collaboration avec le Dr. Olivier Hamelin (LCBM, Grenoble) nous avons pu obtenir l'acétoacétate voulu. En effet, le Dr. Hamelin a mis en place une synthèse en 4 étapes d'un acétoacétate marqué de façon dissymétrique avec un rendement global de 27 %. La synthèse détaillée de ce nouveau précurseur est décrite dans l'article présenté en 2.6 et schématisée dans la figure 1 de ce dernier.

Avec ce nouveau précurseur, il est maintenant possible de corrélérer directement les groupements méthyles pro-S des leucines et des valines à leurs C α et C β respectifs en utilisant un unique échantillon et sans avoir à prendre en considération le déplacement chimique isotopique existant entre les atomes de ^{13}C , provenant des différences magnétiques entre le ^1H et le ^2H . Une application de ce nouveau précurseur est également décrite dans l'article présenté en 2.6.

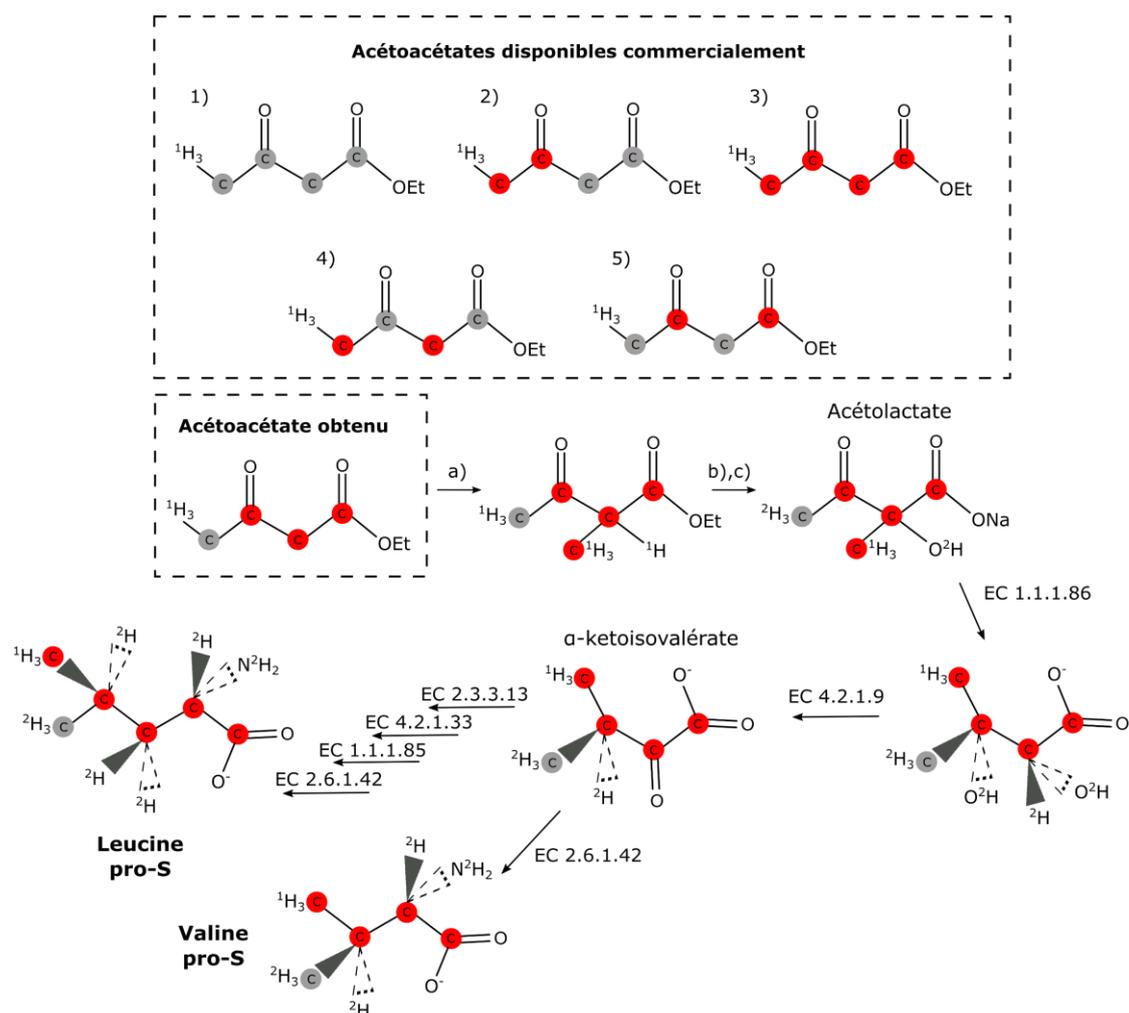


Figure 2.15 : Schéma représentant les marquages d'acétoacétate disponibles commercialement versus celui souhaité. Les sphères rouges et grises représentent les atomes de ^{13}C et ^{12}C , respectivement. Parmi les acétoacétates disponibles dans le commerce on distingue 1) l'éthyl-3-oxobutanoate, 2) l'éthyl-3,4- $^{13}\text{C}_2$ -3-oxo-butanoate, 3) l'éthyl - 1,2,3,4- $^{13}\text{C}_4$ -3-oxo-butanoate, 4) l'éthyl - 2,4- $^{13}\text{C}_2$ -3-oxo-butanoate et 5) l'éthyl - 1,3- $^{13}\text{C}_2$ -3-oxo-butanoate. L'acétoacétate souhaité est l'éthyl-1,2,3- $^{13}\text{C}_3$ -3-oxo-butanoate et est suivi du schéma réactionnel du protocole pour la production d'échantillons marqués U- ^{2}H , ^{15}N - Leu-[2,3,3,4- $^2\text{H}_4$ -1,2,3,4- $^{13}\text{C}_4$ - $^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$, Val-[2,3- $^2\text{H}_2$; 1,2,3- $^{13}\text{C}_3$; $^{13}\text{C}^1\text{H}_3$] pro-S/[$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$. a) K_2CO_3 , $^{13}\text{C}^1\text{H}_3$ -I, 0 °C, 18 h, EtO^1H , b) Cs_2CO_3 , $\text{P}(\text{OEt})_3$, O_2 , DMSO , 20 h, c) i - NaO^2H (2.5 M), $^2\text{H}_2\text{O}$, ^2H ; ii - ^2HCl 35 % ; tampon Tris pH 7.5.

2.5.2 Attribution des groupements méthyles par mutagénèse

L'attribution par mutagénèse dirigée consiste à muter un unique acide aminé possédant un groupement méthyle de la séquence de la protéine (par exemple l'alanine 24) et à le remplacer par un autre type d'acide aminé. Dans le même temps, tous les groupements méthyles du même type que le résidu muté vont être marqués (par exemple ici : les alanines). En comparant les spectres méthyles 2D-[^1H - ^{13}C] de la protéine de référence et de la protéine mutée il sera possible d'identifier la position du résidu muté dans la protéine de référence et ainsi de l'attribuer (dans notre exemple disparition du pic correspondant à l'alanine 24 dans le spectre du mutant).

Cette technique d'attribution est facile d'utilisation et applicable aussi bien pour les protéines que pour les gros assemblages moléculaires. En effet, plusieurs exemples présents dans la littérature décrivent l'attribution de groupements méthyles sur des protéines dont le poids moléculaire varie entre 50 kDa et 1 MDa. On notera, l'attribution des méthionines du poliovirus d'ARN polymérase (Yang et al. 2010b) et celles du 20S-CP-proteasome (Religa et al. 2010), l'attribution des isoleucines- δ_1 de ClpP (Sprangers et al. 2005) ainsi que la vérification de l'attribution des isoleucines- δ_1 de l'ATCase (Velyvis et al. 2009), l'attribution des alanines et isoleucines- δ_1 de TET2 (Amero et al. 2011) et celle des méthionines et valines- γ_2 de HSP60 (Mas et al. 2018).

Cependant, des inconvénients sont également présents. La mutation d'un acide aminé dans la séquence de la protéine peut entraîner une modification de l'environnement chimique des résidus situés à proximité de la mutation pouvant se traduire non pas uniquement par la disparition du pic attendu correspondant au résidu muté mais également par l'observation de déplacements chimiques secondaires. De plus, cette technique est particulièrement chronophage notamment dans les étapes de production et de purification des nombreux mutants nécessaires.

Afin de limiter l'impact de ces inconvénients une stratégie appelée SeSAM (Sequence-Specific Assignment of methyl groups by Mutagenesis) améliorée peut être employée. Initialement développée par Amero et al. (Amero et al. 2011) puis perfectionnée par Crublet et al. (Crublet et al. 2014) cette stratégie vise tout d'abord à minimiser le changement de structure de la protéine et les déplacements chimiques secondaires en sélectionnant un acide aminé de remplacement ayant une structure chimique similaire à celle du méthyle à substituer. Le choix des acides aminés de substitution pour remplacer le groupement méthyle voulu est basé sur les matrices de BLOSUM. Dans le tableau ci-dessous (Tableau 2.1) sont résumés les substitutions ainsi que les marquages suggérés (Henikoff and Henikoff 1992; Crublet et al. 2014).

Tableau 2.1 : Substitutions suggérées pour les résidus comportant des groupements méthyles.

Mutations	Marquage
Alanine → Serine	Alanine
Isoleucine → Valine	Isoleucine
Leucine → Isoleucine	Leucine
Valine → Isoleucine	Valine
Thréonine → Serine	Thréonine
Méthionine → Leucine	Méthionine

D'autre part, comme décrit par Amaro et al. (Amero et al. 2011), limiter le marquage à un seul type résidu contenant un groupement méthyl permet de réduire drastiquement les déplacements chimiques secondaires et ainsi approximativement 50% des méthyles peuvent être attribués sur la base de spectres pour lesquels seul le résidu marqué disparaît et aucun déplacement chimique secondaire n'est observé.

Finalement, pour pallier le temps important de la préparation des échantillons, la production et purification des mutants est parallélisée (Crublet et al. 2014). Dans la méthode SeSAM classique, les mutants sont produits avec des volumes de culture de l'ordre de 50mL (le volume étant à optimiser selon la protéine) dans des erlenmeyers individuels impliquant un grand investissement de temps, notamment pour les purifications, des coûts élevés de production et également de nombreuses manipulations augmentant les erreurs possibles d'inversion de mutants. Dans la stratégie améliorée, développée en milieu riche, non seulement la production et la purification sont parallélisées en plaques mais les volumes de culture sont également réduits (Figure 2.16). En utilisant culture et purification en plaques le risque d'inversion de mutants est réduit et le coût en isotopes diminué.

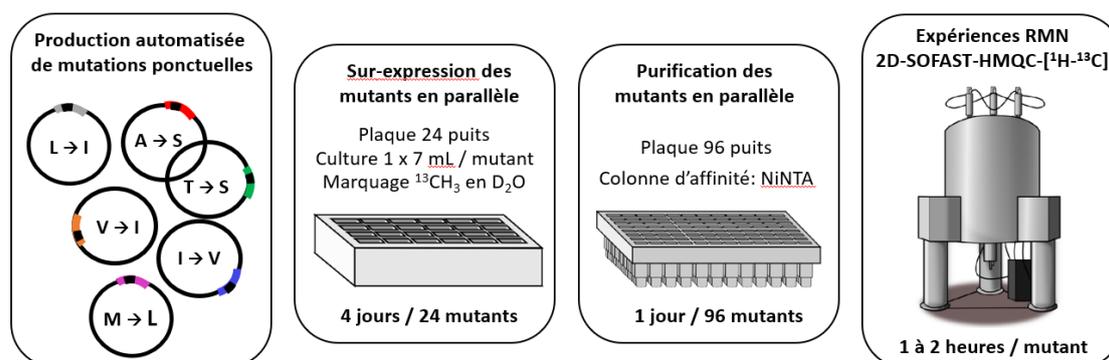


Figure 2.16 : Schéma représentant les différentes étapes de la stratégie d'attribution par mutagenèse dirigée SeSAM améliorée, adapté de (Crublet et al. 2014).

Pour l'analyse des mutants une micro-sonde RMN cryogénique 1.7 mm est utilisée, permettant de limiter le volume d'échantillon nécessaire (~40 μ L par tube). Cette sonde, combinée avec un passeur d'échantillons permet d'enregistrer un spectre 2D ^1H - ^{13}C HMQC en 1 à 2 heures.

En règle générale, les spectres des mutants obtenus peuvent être répartis en 2 catégories. Dans la première, la mutation d'un résidu entraîne la disparition d'un unique pic et n'induit que très peu ou pas de déplacement chimique secondaire pour les autres résidus. L'attribution est non ambiguë. Dans la seconde catégorie, la mutation induite affecte plusieurs signaux. Les spectres des mutants déjà attribués et la banque des spectres des autres mutants non attribués sont utilisés pour conclure sur l'attribution du mutant (Figure 2.17). La structure de la protéine, si disponible, permet de confirmer l'attribution du mutant dont le spectre appartient à la seconde catégorie en confirmant que les déplacements chimiques secondaires observés correspondent à des résidus proches dans l'espace du résidu muté. Si cela n'est pas suffisant, dans les cas les plus complexes, une expérience NOESY méthyle-méthyle pour obtenir des corrélations NOEs inter-méthyles à travers l'espace ainsi que la structure de la protéine sont employées pour attribuer ces résidus (voir partie 2.5.3).

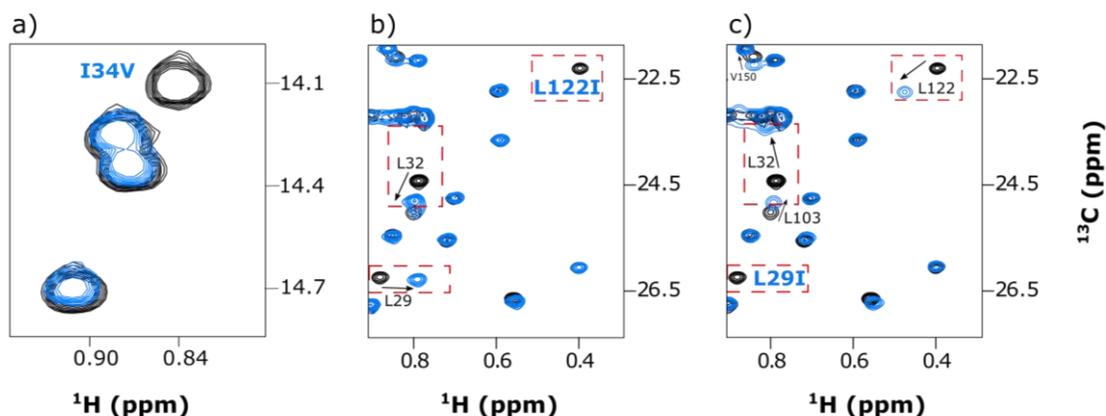


Figure 2.17 : Spectres d'attribution de groupements méthyles avec la méthode de mutagenèse. Exemples de spectres 2D- $[^1\text{H}$ - $^{13}\text{C}]$ -SOFAST méthyle TROSY de mutants d'HSP90-NTD a) pour lequel pas de déplacement chimique secondaire a été observé. b) et c) pour lesquels des déplacements chimiques secondaires ont été observés. Les spectres ont été acquis sur un spectromètre opérant à une fréquence ^1H de 850 MHz. Chaque spectre de mutant (bleu) est superposé avec le spectre correspondant de la protéine de référence (noir). Le signal correspondant à chaque mutant est annoté en couleur. Les déplacements chimiques secondaires sont indiqués par des flèches.

2.5.3 Attribution des groupements méthyles employant des méthodes basées sur la structure

Alternativement, une attribution des groupements méthyles peut être effectuée de façon manuelle ou automatique lorsque la structure 3D haute résolution de la protéine est connue. Pour cela, les données de RMN nécessaires peuvent être réparties en deux catégories. La première consiste à employer une méthode basée sur des données paramagnétiques et la seconde approche a pour but d'exploiter l'effet NOE (Nuclear Overhauser Effect) (Pritišanac et al. 2020).

2.5.3.1 Méthodes d'attribution automatique se basant sur l'effet paramagnétique

Différents logiciels permettent d'attribuer de façon automatique les groupements méthyles d'une protéine en se basant sur l'effet paramagnétique. Pour cela les effets paramagnétiques théoriques induits par le centre paramagnétique et calculés à partir de la structure tridimensionnelle haute résolution de la protéine ainsi que de la localisation de son centre paramagnétique sont comparés avec les effets observés dans les expériences acquises par RMN (Pritišanac et al. 2020).

Ces logiciels peuvent être répartis en deux catégories selon le mécanisme paramagnétique étudié. Toutefois, pour les deux types de méthodes l'introduction d'un centre paramagnétique (i.e. d'une entité chimique présentant un électron apparié) sur la protéine d'intérêt est indispensable. On notera d'un côté le logiciel « PRE assign » (Venditti et al. 2011) se basant sur l'augmentation de la relaxation due à la présence du centre paramagnétique (effet PRE) et d'un autre côté les logiciels tels que « POSSUM » (John et al. 2007) et « PARA assign » (Skinner et al. 2013; Lescanne et al. 2017) reposants sur les déplacements chimiques induits par la présence du centre paramagnétique (effet PCS). Les méthodes basées sur l'effet paramagnétique présentent l'avantage, grâce au moment magnétique élevé des électrons, de pouvoir avoir un effet à longue distance et ainsi par exemple permettre d'obtenir des informations sur un groupement méthyle isolé, contrairement à l'effet NOE.

Si tous les centres paramagnétiques donnent lieu à un effet PRE, d'intensité variable selon le centre introduit, à l'inverse, tous les centres paramagnétiques ne donnent pas lieu à un effet PCS. En effet, l'effet PCS induit ou non par le centre paramagnétique introduit dépend de la nature même de ce dernier. Le paramagnétisme d'un centre peut être décrit en fonction d'un tenseur de susceptibilité magnétique χ , variant sur les trois axes x, y et z. L'anisotropie de χ , définie par $\Delta\chi$ (avec $\chi = \chi_{\text{isotropique}} + \Delta\chi$), est responsable de l'effet paramagnétique PCS. Si $\Delta\chi$ est nul alors l'effet paramagnétique PCS sera nul et aucun déplacement chimique ne sera observé sur les spectres (Otting 2010; Softley et

al. 2020). Généralement, dans une protéine il est quasiment impossible d'avoir une sphère de coordination totalement symétrique, si bien que lorsqu'un cation paramagnétique ayant une relaxation électronique rapide par rapport à la rotation de la protéine se lie, il introduira un effet PCS dû à l'anisotropie du tenseur de susceptibilité magnétique.

Différents types de centre paramagnétiques existent. On note les composés organiques comportant un groupe nitroxyde, Mn^{2+} et Gd^{3+} , pour lesquels une relaxation électronique lente des spins domine et qui ne donnent pas lieu à des PCS, ils sont donc particulièrement intéressants pour mesurer des effets PRE sans avoir besoin de faire une attribution de la forme paramagnétique de la protéine étudiée. A l'inverse, tous les lanthanides, à l'exception de Gd^{3+} , possèdent un fort $\Delta\chi$ (Figure 2.18) et une relaxation électronique des spins plus rapide que la rotation de la molécule et ainsi induisent un effet PCS.

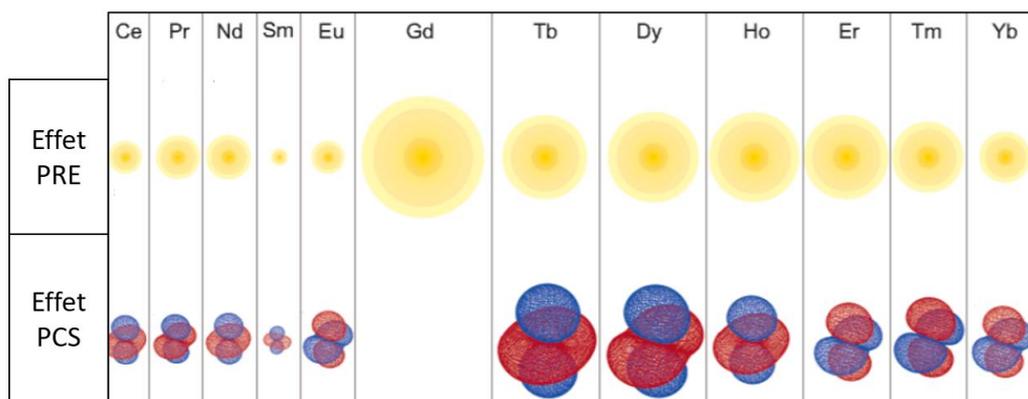


Figure 2.18 : Schéma représentant les propriétés paramagnétiques de certains lanthanides. Le rayon des sphères jaune indique l'amplitude de l'effet PRE. Les isosurfaces de l'effet PCS sont représentées telles que calculées avec le $\Delta\chi$ de la calbindin D_{9k} . Adaptée de (Pintacuda et al. 2007).

Quel que soit le centre paramagnétique introduit, proche de ce dernier les signaux des groupements méthyles seront élargis voire invisibles en raison de l'augmentation de la relaxation transversale induite par la présence du centre paramagnétique (effet PRE). La distance jusqu'à laquelle les signaux sont invisibles ainsi que celle jusqu'à laquelle l'intensité des signaux est affectée dépend de l'intensité de l'effet PRE dépendant lui-même de la nature du centre paramagnétique. L'effet PRE étant proportionnel à l'inverse de la puissance sixième de la distance entre le centre paramagnétique et le noyau d'intérêt, la sphère pour laquelle on observe un effet PRE est plus petite que celle pour laquelle on observe un effet PCS, qui elle sera proportionnelle à l'inverse de la

puissance cube de la distance entre le centre paramagnétique et le noyau d'intérêt. De la même manière, la distance jusqu'à laquelle les déplacements chimiques sont affectés par l'effet PCS va dépendre de la nature du centre paramagnétique introduit et de l'asymétrie de sa sphère de coordination. En règle générale la distance maximale pour laquelle un effet PRE peut être mesuré avec suffisamment de précision est de l'ordre de 25 Å dans le cas d'un groupement nitroxyde alors qu'un effet PCS peut toujours être mesuré jusqu'à 50 Å dans le cas d'un lanthanide à fort $\Delta\chi$ tel que le Dy³⁺ (Häussinger et al. 2009; Pritišanac et al. 2020) (Figure 2.19).

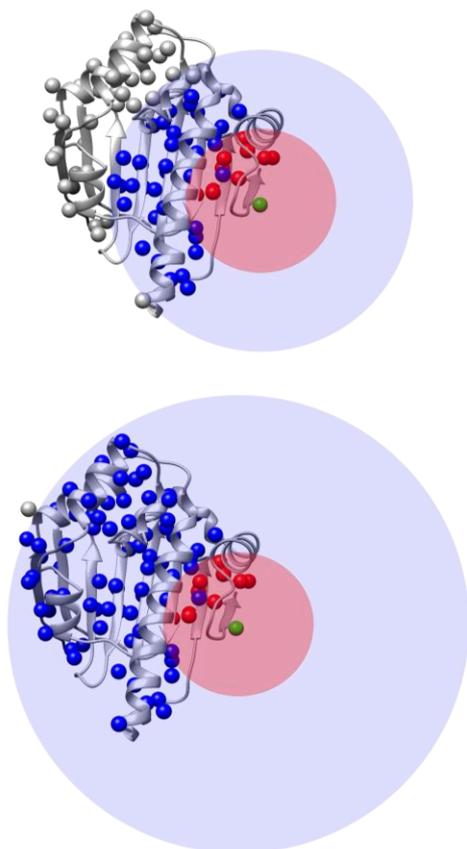


Figure 2.19 : Structures du domaine N-terminal d'HSP90 α . En haut : tous les groupements méthyles à une distance de 25 Å du centre paramagnétique (sphère verte – groupement nitroxyde) sont colorés en bleus. Ce sont les sondes pour lesquelles un effet PRE peut être détecté. Les groupements méthyles colorés en rouge sont ceux invisibles en raison de la trop grande proximité avec le centre paramagnétique qui conduit à une forte relaxation transversale. En bas : comme en haut, à l'exception de la nature du centre paramagnétique (lanthanide). Les sphères bleues représentent les groupements méthyles à 40 Å du centre paramagnétique et pour lesquels un effet PCS peut être observé. Les cercles semi transparents rouge et bleu représentent respectivement la zone où les groupements méthyles sont invisibles et celle où leur intensité / les déplacements chimiques sont affectés par la présence du centre paramagnétique. Adapté de (Pritišanac et al. 2020).

Pour l'attribution des groupements méthyles par PRE ou PCS, le groupement paramagnétique doit être introduit de façon spécifique sur la protéine, généralement sur une cystéine accessible ou sur une lysine. Et, ce dernier doit également avoir des degrés de liberté conformationnelle réduits afin de pouvoir mesurer avec précision les distances entre le centre paramagnétique et les groupements méthyles. Une paire de cystéines proches dans l'espace peuvent également être employées pour fixer le groupement paramagnétique et limiter ses degrés de liberté comme présenté par Lescanne et al. (Lescanne et al. 2017). Toutefois, contrairement à l'attribution

employant l'effet NOE, plusieurs échantillons avec un centre paramagnétique placé à différents endroits de la protéine sont généralement nécessaires pour assurer un fort taux d'attribution des groupements méthyles de la protéine. De plus pour mesurer l'effet paramagnétique, un spectre acquis avec la même protéine d'intérêt doit également être effectué lorsque cette dernière est liée à un centre diamagnétique. En effet, les déplacements chimiques d'une protéine sont sensibles à la liaison d'un ion métallique ou d'un groupement nitroxyde que celui-ci soit paramagnétique (Equation 2.1) ou diamagnétique (Equation 2.2).

Protéine d'intérêt + Centre paramagnétique = Déplacements chimiques dus à la présence du centre magnétique + Effet paramagnétique (Equation 2.1)

Protéine d'intérêt + Centre diamagnétique = Déplacements chimiques dus à la présence du centre magnétique (Équation 2.2)

Ainsi, la comparaison des deux spectres (Equation 2.1- Equation 2.2), en ayant choisi idéalement un centre diamagnétique similaire à celui paramagnétique, permet d'avoir accès uniquement à l'effet paramagnétique. Par exemple, deux lanthanides diamagnétiques existent La^{3+} et Lu^{3+} dont le rayon ionique est similaire au rayon ionique de différents lanthanides paramagnétiques.

Attribution basée sur les effets PRE : PRE-assign

En utilisant cinq mutants pour lesquels des groupements nitroxides paramagnétiques ont été introduits à différentes positions sur la protéine EIN (27 kDa), le logiciel « PRE-assign » (Venditti et al. 2011) a permis d'attribuer 47 % des 140 groupements méthyles alanine, isoleucine, leucine et valine correctement et sans ambiguïté (comparaison faite avec une attribution des groupements méthyles A, I, L et V effectuée avec une expérience NOESY). Afin de pouvoir mesurer l'effet PRE, un échantillon diamagnétique est préparé pour chacun des cinq mutants en ajoutant de l'acide ascorbique. Le programme fonctionne selon un protocole en trois étapes qui visent à simultanément chercher la meilleure attribution possible des groupements méthyles tout en en raffinant la position du groupement nitroxyde. Afin de limiter les attributions ambiguës, le type de chaque groupement méthyle (i.e. alanine, leucine/valine et isoleucine) est fourni. Il est reporté qu'approximativement un centre paramagnétique supplémentaire doit être ajouté pour chaque 6 kDa additionnel de protéine pour espérer obtenir une attribution complète et fiable des groupements méthyles. Cette technique est donc très chronophage en termes de préparation d'échantillons. De plus, 40 % des 140

groupements méthyles sont attribués correctement mais avec ambiguïté et 13% des attributions proposées par « PRE-assign » sont incorrectes pour la protéine EIN.

Attributions basées sur les effets PCS : POSSUM et PARA-assign

Le programme POSSUM (John et al. 2007) a été employé pour l'attribution des groupements méthyles d'un complexe de 30 kDa où les ions divalents du site actif ont été remplacés par un lanthanide. La³⁺ a été choisi comme centre diamagnétique et Dy³⁺ comme centre paramagnétique. L'emploi de Dy³⁺ ayant une forte anisotropie de tenseur de susceptibilité magnétique engendre d'importants déplacements chimiques par rapport au spectre de la protéine liée au centre diamagnétique. De plus ce lanthanide paramagnétique possède un fort effet PRE affectant grandement l'élargissement des lignes ¹H ainsi rendant invisible la quasi-totalité des groupements méthyles situés dans un rayon de 15 Å par rapport à Dy³⁺. Une solution proposée par John et al. a été d'introduire un second lanthanide : Yb³⁺ à la place de Dy³⁺. Ce second lanthanide possède à la fois des effets PCS et PRE réduits par rapport au Dy³⁺. En combinant l'analyse des spectres des différents échantillons marqués avec les différents lanthanides il a été possible d'attribuer 84 % des groupements alanine, isoleucine, méthionine, valine, thréonine et leucine du complexe étudié en comparant les effets PCS observés aux effets PCS théoriques calculés connaissant la structure du complexe. Ces chiffres sont augmentés lorsque la spécificité des groupements méthyles isoleucine, leucine et valine est définie. Cependant la connaissance de l'anisotropie du tenseur de susceptibilité magnétique, dont la détermination nécessite soit l'attribution d'un sous ensemble de résonances méthyles soit celle complète des groupements amides du squelette de l'échantillon diamagnétique, est indispensable et se révèle être une contrainte de taille (Pritišanac et al. 2020).

Le logiciel PARA-assign a également été utilisé pour attribuer des groupements méthyles, ceux isoleucine, leucine et valine de HSP90-NTD, ma protéine d'étude, en utilisant l'effet PCS. Différents mutants de HSP90-NTD ont été produits. En effet, cette dernière ne possède pas de cystéine naturellement présente dans sa séquence et les lanthanides ont été attachés par deux cystéines afin de limiter leurs degrés de liberté. D'après les auteurs, il a été reporté qu'avec deux mutants différents comportant chacun deux cystéines et un lanthanide (paramagnétique Yb³⁺ / diamagnétique Lu³⁺) une attribution fiable et non ambiguë de 60 % des 76 groupements méthyles (ile- δ_1 , leu δ_1/δ_2 et val γ_1/γ_2) d'HSP90-NTD était obtenue (Lescanne et al. 2017). Additionnellement, 36 % des résidus étaient attribués de façon ambiguë mais correcte. Toutefois il est à noter que la référence prise pour déterminer la véracité de l'attribution des groupements obtenue par PARA-assign est celle de Shah et al. (Shah et al. 2012) qui ne dévoile aucune attribution des groupements méthyles de HSP90-NTD. L'attribution des groupements

isoleucines par Lescanne et al, non divulguée, aurait pourtant pu être comparée avec celle de Park et al. (Park et al. 2011) disponible librement. Finalement en 2018, Lescanne et al. (Lescanne et al. 2018) ont publié l'attribution stéréospécifique des groupements valines et leucines de HSP90-NTD obtenue en employant les PCS. Cependant nous noterons que cette attribution, fiable d'après les auteurs, n'est en réalité juste qu'à 80 % (voir partie 2.7 – Article I).

Les méthodes d'attribution automatiques se basant sur les effets paramagnétiques sont très intéressantes pour aider à l'attribution des groupements méthyles. En effet, même si elles ne fournissent qu'une attribution partielle et qu'elles sont plus chronophages que les méthodes d'attribution automatiques se basant sur l'effet NOE puisqu'elles nécessitent la production de plusieurs échantillons, elles permettent d'apporter des informations à plus longue distance pouvant se révéler cruciales.

2.5.3.2 Méthodes d'attribution automatique employant l'effet NOE

Il existe de nombreux programmes permettant l'attribution automatique des résonances des groupements méthyles d'une protéine se basant sur l'effet NOE. Le premier en date MAP-XSII (Xu et al. 2009; Xu and Matthews 2013), développé en 2009, a été rapidement suivi par cinq autres programmes : FLAMEnGO 2.0 (Chao et al. 2014), MAGMA (Pritišanac et al. 2017), MAGIC (Monneau et al. 2017), MethylFLYA (Pritišanac et al. 2019) et MAUS (Nerli et al. 2021). Ces six derniers permettent de comparer le réseau de corrélations NOEs inter-méthyl à travers l'espace obtenu par des expériences NOESY 3D sur la protéine d'intérêt avec le réseau de corrélations NOEs extrait de la structure 3D haute résolution à disposition pour proposer une attribution des groupements méthyles la plus probable.

Ces logiciels utilisent différentes approches pour attribuer les groupements méthyles. Par exemple, MAP-XSII et FLAMEnGO 2.0 utilisent une permutation avec Monte Carlo où une attribution initiale est donnée aléatoirement à chaque groupement méthyle puis échangée itérativement. Chaque changement est évalué par une fonction qui est maximisée quand l'accord entre les données mesurées expérimentalement et prédites par la structure est optimal (Pritišanac et al. 2020). MAGIC quant à lui fonctionne avec une approche utilisant la densité du réseau. Ainsi, les possibilités d'attribution sont réduites en prenant en compte le fait que les contributions NOE sont limitées avec la distance et les échanges d'attributions se font uniquement localement (Monneau et al. 2017). Cependant, même si leurs approches sont variables il a été reporté que toutes les méthodes d'attribution automatique présentaient des performances accrues lorsque des informations sur le type d'acide aminé (Ala, Ile, Leu, Met, Thr, Val) étaient données

initialement (Pritišanac et al. 2020) et encore plus quand la stéréospécificité (leucine, valine) et régiosélectivité (isoleucine) du marquage étaient spécifiées.

L'utilisation de l'effet NOE pour attribuer les groupements méthyles d'une protéine présente plusieurs avantages, notamment le fait qu'aucune mutation ni modification n'est nécessaire sur la protéine d'intérêt. De plus, contrairement aux méthodes basées sur une approche paramagnétique, un unique échantillon est suffisant pour la collecte de données RMN. Toutefois certains inconvénients ont également été reportés. En effet, de par la nature « courte-distance » des NOEs et la possible diffusion de spin à temps de mélange élevés, l'interprétation des NOEs peut s'avérer compliquée et mener à des ambiguïtés dans l'attribution (Pritišanac et al. 2020). De plus, certains groupements méthyles peuvent être isolés, notamment lorsqu'ils se situent à la surface de la protéine et ainsi ne disposent pas de réseau de NOEs permettant leur attribution. Une solution proposée pour pallier, en partie, ce premier problème est de deutérer uniformément la protéine avec une protonation sélective uniquement sur les groupements méthyles comme proposé par Sounier et al. et Ayala et al. (Sounier et al. 2007; Ayala et al. 2020). Ainsi, des pics de corrélation NOEs peuvent être observés pour des sondes situées jusqu'à 10-12 Å l'une de l'autre. Enfin, une structure haute résolution est nécessaire pour définir avec exactitude les coordonnées des groupements méthyles. Il est à noter que cette dernière doit également être compatible avec les données RMN pour pouvoir en déduire des attributions. En effet, les données RMN en solution reportent des informations sur un ensemble d'états moléculaires alors que généralement à l'état cristallin une unique structure est obtenue (Pritišanac et al. 2020).

Le programme d'attribution automatique le plus récent (MAUS) a été comparé simultanément aux cinq autres logiciels d'attribution automatique en utilisant 8 protéines de référence (à l'exception de MAGMA qui a été comparé avec MAUS sur d'autres protéines de références) (Nerli et al. 2021). Il a été montré qu'en moyenne, les différents programmes d'attribution automatique permettent une attribution non-ambiguë et fiable (> 60 %) même si cette dernière est la majeure partie du temps uniquement partielle (Figure 2.20). Il est toutefois très utile d'avoir accès à une attribution fiable d'une partie des groupements méthyles d'une protéine faite de façon automatique et les programmes fournissent également des informations sur les attributions ambiguës qui sont précieuses pour terminer l'attribution manuellement.

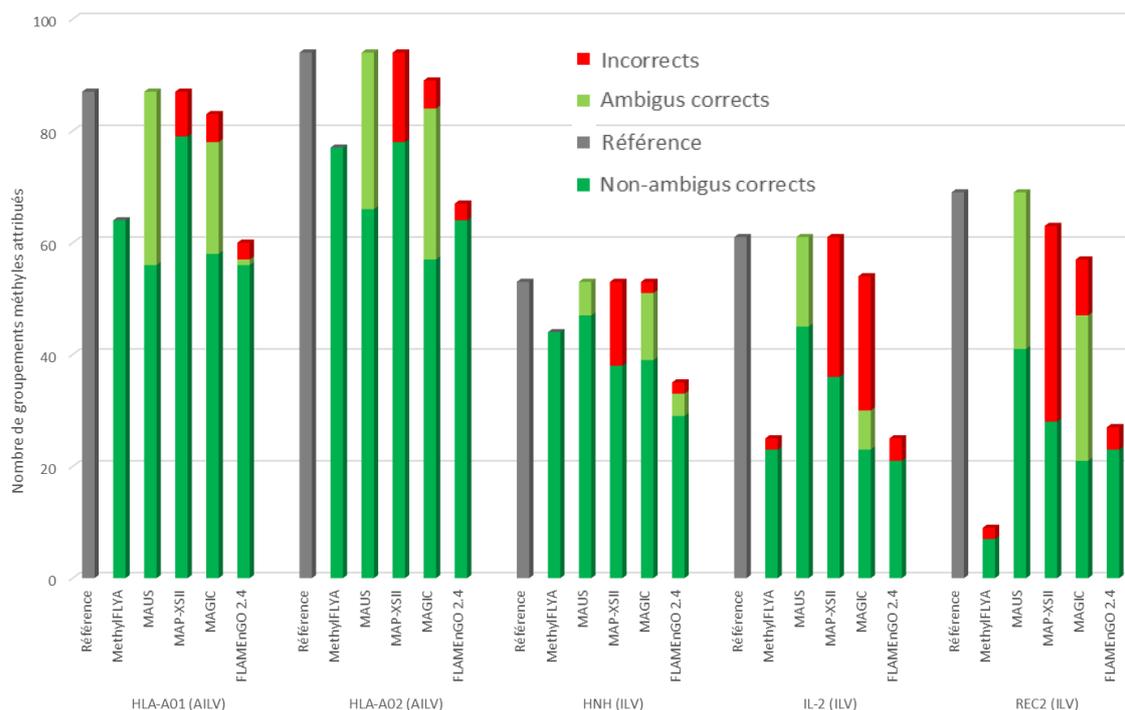


Figure 2.20 : Histogramme résumant les performances de différents programmes d'attribution automatique des groupements méthyles pour cinq protéines : HLA-A01, HLA-A02, HNH, IL-2 et REC2. Les données proviennent du tableau de Nerli et al. (Nerli et al. 2021).

2.6. Application des différentes stratégies d'attribution à HSP90-NTD Apo – Article I

Pour attribuer les groupements méthyles du domaine N-terminal de HSP90 Apo marquée U-[^2H , ^{15}N , ^{12}C], $^{13}\text{CH}_3$ – (Ile- δ_1 , Leu- δ_2 , Val- γ_2 , Ala- β , Met- ϵ , Thr- γ_2), les trois stratégies développées en 2.5.1, 2.5.2 et 2.5.3.2 ont été employées. En effet, même si HSP90-NTD est une protéine de taille modérée (26 kDa) dont l'attribution du squelette devrait être évidente et permettre d'attribuer par transfert les groupements méthyles Ala, Ile, Leu, Val de façon simple, il se trouve que la flexibilité de la protéine et sa dynamique dans la gamme de temps μs - ms rendent l'attribution du squelette et donc des groupements méthyles par transfert plus complexe qu'au premier abord.

De plus, HSP90-NTD est une protéine très flexible rendant cette cible particulièrement complexe pour les méthodes d'attribution automatique des groupements méthyles. Les taux de succès reportés par les différents programmes pour HSP90-NTD varient entre 24 et 69 % (Pritišanac et al. 2017, 2019, 2020; Monneau et al. 2017). Cependant, les programmes peuvent également attribuer de façon non ambiguë un résidu

incorrectement. En utilisant MAGIC je n'ai pu attribuer automatiquement que 34 % des groupements méthyles de la protéine avec un score de confiance élevé.

En employant les trois stratégies développées précédemment j'ai pu attribuer 100 % des groupements méthyles d'HSP90-NTD marquée U-[²H, ¹⁵N, ¹²C], ¹³CH₃ – (Ile-δ₁, Leu-ProS, Val-ProS, Ala-β, Met-ε, Thr-γ₂) et les résultats de ce travail d'attribution sont présentés dans l'article qui suit. La méthode SeSAM de mutagénèse dirigée a été utilisée pour attribuer la majorité des groupements méthyles des acides aminés : alanine, thréonine et méthionine. En effet, méthionines et thréonines ne peuvent être attribuées en utilisant le transfert d'attribution du squelette aux groupements méthyles dû, respectivement, à la présence d'un atome de soufre empêchant l'utilisation des expériences « out and back HCC/HC(C)C/HC(CC)C » et à l'absence de thréonine marquée Thr-[1,2,3-¹³C, 2,3-²H₂, ¹³C¹H₃-γ] disponible commercialement. Les corrélations NOEs inter-méthyles à travers l'espace ont également servi pour compléter l'attribution des alanines et thréonines. Concernant les isoleucines, leucines et valines la méthode du transfert d'attribution du squelette aux groupements méthyles a été employée en utilisant les précurseurs 2-hydroxy-2-(1',1'-[²H₂],1',2'-[¹³C₂])-ethyl-1,2,3-[¹³C₃],4,4,4-[²H₃]-3-oxobutanoate (Figure 2.12) ainsi que celui nouvellement développé 1,2,3-[¹³C₃]-2-[¹³C¹H₃]-2-[O²H]-3-oxo-4,4,4-[²H₃]-butanoate (Figure 1 – Article I) permettant de lier directement les groupements méthyles Ile- δ₁ et pro-S des leucines et valines au squelette par une chaîne de carbone marquée ¹³C. Ces derniers ont permis l'attribution de 100 % des isoleucines, leucines et valines pour lesquelles une attribution du squelette était disponible. La méthode SeSAM de mutagénèse dirigée a été utilisée pour compléter l'attribution des résidus isoleucine et leucine pour lesquels l'attribution du squelette n'était pas disponible.

Dans ce travail interdisciplinaire ma contribution a consisté en la production des différents mutants et des échantillons marqués de HSP90-NTD dans lesquels le nouveau précurseur développé par Dr. O. Hamelin a été incorporé. J'ai acquis les expériences 2D-SOFAST-[¹H-¹³C] sur les différents mutants et participé à l'acquisition des expériences de triples résonances et de « out and back HCC/HC(C)C/HC(CC)C » implémentées par Dr. J. Boisbouvier. J'ai analysé les différents types de spectres, attribué les résonances des groupements méthyles de HSP90-NTD et j'ai également rédigé l'article avec l'aide du Dr. J Boisbouvier.

Article I : « Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups : application to human Hsp90 N-terminal domain »



Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain

Faustine Henot¹ · Rime Kerfah² · Ricarda Törner¹ · Pavel Macek^{1,2} · Elodie Crublet² · Pierre Gans¹ · Matthias Frech³ · Olivier Hamelin⁴ · Jerome Boisbouvier¹

Received: 26 February 2021 / Accepted: 20 May 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Methyl moieties are highly valuable probes for quantitative NMR studies of large proteins. Hence, their assignment is of the utmost interest to obtain information on both interactions and dynamics of proteins in solution. Here, we present the synthesis of a new precursor that allows connection of leucine and valine pro-*S* methyl moieties to backbone atoms by linear ¹³C-chains. This optimized ²H/¹³C-labelled acetolactate precursor can be combined with existing ¹³C/²H-alanine and isoleucine precursors in order to directly transfer backbone assignment to the corresponding methyl groups. Using this simple approach leucine and valine pro-*S* methyl groups can be assigned using a single sample without requiring correction of ¹H/²H isotopic shifts on ¹³C resonances. The approach was demonstrated on the N-terminal domain of human HSP90, for which complete assignment of Ala-β, Ile-δ₁, Leu-δ₂, Met-ε, Thr-γ and Val-γ₂ methyl groups was obtained.

Keywords NMR · Assignment · Methyl groups · Acetolactate · HSP90

Introduction

Solution state NMR is the method of choice to characterize proteins at atomic level and to probe their dynamics over a wide range of biologically relevant timescales. However, for a long-time, study of high molecular weight proteins by NMR remained a challenge, notably due to the extensive line broadening of NMR signals in large proteins. Methyl groups have been widely studied and are extremely useful to overcome this issue that has hampered, in the past, quantitative NMR studies on large proteins. Indeed, due to the proton

multiplicity and their favorable relaxation properties, methyl groups allow the detection of NMR signals even for large proteins (Tugarinov et al. 2003). Nowadays, methyl groups are important probes to investigate molecular dynamics (Sprangers and Kay 2007) and to provide functional insight (Rosenzweig et al. 2013; Mas et al. 2018) on assemblies weighing up to 1 MDa. Specific labelling of methyl groups on perdeuterated large proteins allows the measurement of long-range distance restraints, up to 12 Å. (Sounier et al. 2007; Ayala et al. 2020) and enables, in combination with other structural biology techniques such as SANS/SAXS (Lapinaite et al. 2013) or Cryo-EM (Gauto et al. 2019), to solve the structure of complexes of several hundreds of kDa.

For the past 20 years, a plethora of protocols overexpressing proteins in M9/²H₂O based *E. coli* growth medium and leading to specific protonation of methyl groups in perdeuterated proteins without scrambling of protons to other sites have been elaborated. On one hand, the direct incorporation of the methyl labelled amino acid in M9/²H₂O is employed for the specific labelling of ¹³C¹H₃-alanine (Isaacson et al. 2007; Ayala et al. 2009), ¹³C¹H₃-methionine (Gelis et al. 2007; Stoffregen et al. 2012) and ¹³C¹H₃-threonine (Velyvis

✉ Jerome Boisbouvier
jerome.boisbouvier@ibs.fr

¹ Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, Avenue des martyrs, 38044 Grenoble, France

² NMR-Bio, 5 place Robert Schuman, 38025 Grenoble, France

³ Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany

⁴ Univ. Grenoble Alpes, CEA, CNRS, IRIG, CBM, 38000 Grenoble, France

et al. 2012; Ayala et al. 2020). On the other hand, as leucine, valine and isoleucine residues are at the end of irreversible metabolic pathways in *E. coli*, precursors can be incorporated in the growth medium for their cost-effective labelling. The first precursors introduced to label isoleucine or leucine and valine residues were 2-keto acids: α -ketobutyrate (Gardner et al. 1997) and α -ketoisovalerate (Goto et al. 1999; Hajduk et al. 2000; Gross et al. 2003), respectively. However, α -ketoisovalerate used as a precursor for leucine and valine residues is leading to a non-stereospecific labelling of both pro-*S* and pro-*R* $^{13}\text{C}^1\text{H}_3$ groups resulting in overcrowded spectra for high molecular weight proteins, even when both sensitivity and resolution have been improved using a non-stereospecific $^{13}\text{C}^1\text{H}_3/^{12}\text{C}^2\text{H}_3$ α -ketoisovalerate (Tugarinov and Kay 2004b).

To prevent peak overlaps and to facilitate studies of large molecular weight assemblies, methyl labelled acetolactate has been used as an alternative precursor (Gans et al. 2010). This latter enables the stereospecific $^{13}\text{C}^1\text{H}_3$ -labelling of valine and leucine methyl groups and therefore halves the number of peaks observed whilst improving by two the sensitivity of the spectrum as compared to labelling at 50% pro-*S* and 50% pro-*R* methyl moieties using optimized $^{13}\text{C}^1\text{H}_3/^{12}\text{C}^2\text{H}_3$ α -ketoisovalerate (Tugarinov and Kay 2004b). This interesting precursor also enhances the intensity of NOE cross peaks and increases the distance threshold at which NOE cross peaks can be detected by 20% (Gans et al. 2010).

However, despite the tremendous progress made in protocols to selectively introduce protonated methyl groups in perdeuterated proteins, sequence specific assignment, essential for analyzing a variety of NMR data, remains an important challenge for large molecular assemblies. Several methods to solve the bottleneck of assignment of large proteins have been developed including parallel mutagenesis strategies (Amero et al. 2011), structure based approaches using the analysis of NOE cross peaks with various programs (Pritišanac et al. 2020), MAP-XSII (Xu and Matthews 2013), FLAMEnGO 2.0 (Chao et al. 2014), MAGMA (Pritišanac et al. 2017), MAGIC (Monneau et al. 2017), MethylFLYA (Pritišanac et al. 2019), MAUS (Nerli et al. 2021) or the “divide-and-conquer” approach (Gelis et al. 2007; Sprangers and Kay 2007) which is based on separation of large proteins into smaller fragments, assigning these and transferring the assignment back to the full-length protein. For proteins of moderate molecular weight or fragments of large assemblies for which backbone assignment is available, it is possible to connect methyl resonances to those of the backbone. This requires a sample with $^{13}\text{C}^1\text{H}_3$ labelled methyl groups connected to the backbone by a linear ^{13}C -chain. With such a sample, transfer from the assigned backbone to the methyl groups can be achieved using either unidirectional transfer from methyl groups to HN using (HM)CM(CGBCA)NH

experiments (Tugarinov and Kay 2003) or ‘out and back’ HCC relay triple resonance experiments (Tugarinov and Kay 2003; Ayala et al. 2012; Mas et al. 2013).

A combination of these techniques to assign methyl groups in addition with a stereospecific labelling, enhancing the sensitivity of the spectrum by a factor two and significantly reducing signal overlap, should lead to straightforward leucine and valine methyl group assignment. Labelling schemes, connecting non-stereospecifically both leucine and valine methyl groups (Tugarinov and Kay 2003), or only pro-*R* methyl moieties (Mas et al. 2013; Kerfah et al. 2015a), to the assigned backbone are already available. However, with such precursors additional samples are required either to stereospecifically assign the methyl group (Tugarinov and Kay 2004a; Gans et al. 2010) or to link the pro-*S* methyl to the pro-*R* one (Mas et al. 2013; Kerfah et al. 2015a). Here we introduce the synthesis of a new dissymmetric $^{13}\text{C}^2\text{H}$ -labelled acetolactate, a precursor that allows to directly connect the assigned leucine and valine backbone atoms to the pro-*S* methyl groups via a linear ^{13}C chain using only one sample. This new labelling scheme has been applied to the N-terminal domain of human HSP90 (HSP90-NTD) and we present here the full assignment of methyl moieties of this protein.

Materials and methods

Synthesis of 1,2,3- $^{13}\text{C}_3$ -4,4,4- $^2\text{H}_3$ -acetolactate

Synthesis of ethyl 1,2,3- $^{13}\text{C}_3$ -3-oxo-butanoate

A solution of LiHMDS (7.80 g, 46.6 mmol, 2.1 equiv.) in freshly dried THF (150 mL) was cooled to -78°C under argon. 1,2- $^{13}\text{C}_2$ -ethyl acetate (2.00 g, 22.2 mmol, Cambridge Isotope Laboratory, CIL) was added dropwise. The resulting solution was stirred at -78°C for 15 min, then 1- ^{13}C -acetyl chloride (1.60 mL, 22.2 mmol, 1 equiv., CIL) was added dropwise. The resulting mixture was stirred at -78°C for additional 30 min then quenched by addition of a 20% aqueous solution of ^1HCl (15 mL). After three extractions with Et_2O , the organics were combined, washed with saturated Na^1HCO_3 solution then dried over Na_2SO_4 . Concentration under vacuum affords the desired product (2.88 g) which was used in the next step without further purification).

^1H NMR: (C^2HCl_3), δ : 4.21 (dq, O-CH₂, $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, $^3\text{J}(\text{H}-^{13}\text{C})=3.2$ Hz, 2H); 3.46 (dt, $^{13}\text{C}-^{13}\text{CH}_2-^{13}\text{C}$, $^2\text{J}(\text{H}-^{13}\text{C})=6.5$ Hz, $^1\text{J}(\text{H}-^{13}\text{C})=130.1$ Hz, 2 H); 2.28 (dd, $\text{CH}_3-^{13}\text{C}$, $^3\text{J}(\text{H}-^{13}\text{C})=1.4$ Hz, $^2\text{J}(\text{H}-^{13}\text{C})=6.1$ Hz, 3H), 1.30 (t, OCH₂-CH₃, $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, 3H).

Synthesis of ethyl 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -3-oxo-butanoate $^{13}\text{C}^1\text{H}_3$ -I (752 mL, 11.99 μmoles , 1.1 equiv, CIL)

was slowly added to a solution of ethyl 1,2,3- $^{13}\text{C}_3$ -3-oxo butanoate (1.45 g, 10.90 mmol) in EtO^1H (50 mL) cooled to 0 °C before addition of K_2CO_3 (1.66 g, 11.99 mmol, 1.1 equiv.). The resulting suspension was warmed to room temperature then stirred for 18 h. The mixture was concentrated to the fifth before addition of a large volume of Et_2O . Excess of K_2CO_3 was filtered off and the filtrate concentrated under vacuum to the fifth before a further addition of Et_2O and a second filtration. Concentration under vacuum affords the desired product (1.03 g) as a colorless oil which was used in the next step without further purification.

^1H NMR: (C^2HCl_3), δ : 4.21 (dq, O- CH_2 , $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, $^3\text{J}(\text{H}-^{13}\text{C})=3.0$ Hz, 2H); 3.50 (dm, $^{13}\text{C}-^{13}\text{CH}-^{13}\text{C}$, $^1\text{J}(\text{H}-^{13}\text{C})=129.0$ Hz, 1H); 2.24 (dd, $\text{CH}_3-^{13}\text{C}$, $^3\text{J}(\text{H}-^{13}\text{C})=1.3$ Hz, $^2\text{J}(\text{H}-^{13}\text{C})=6.0$ Hz, 3H), 1.36 (dm, $^{13}\text{CH}_3$, $^1\text{J}(\text{H}-^{13}\text{C})=129.0$ Hz, 3H), 1.28 (t, O CH_2 - CH_3 , $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, 3H).

Synthesis of ethyl 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O ^1H]-3-oxo-butanoate To a solution of ethyl 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -3-oxo-butanoate (995 mg, 6.72 mmol) in DMSO (8 mL), Cs_2CO_3 was added (440 mg, 1.35 mmol, 0.2 equiv.). After O_2 bubbling for 15 min, P(OEt) $_3$ (233 mL, 1.35 mmol, 0.2 equiv.) was added. The resulting solution was stirred under O_2 atmosphere for 20 h. A large volume of Et_2O was then added followed by a saturated NaCl solution. The resulting phases were separated and the aqueous one was extracted one more time with Et_2O . The organics were combined, dried over Na_2SO_4 then concentrated under vacuum to obtain ethyl 1,2,3- $^{13}\text{C}_3$, 2- $^{13}\text{C}^1\text{H}_3$, 2-[O ^1H]-3-oxo-butanoate as a yellow oil (1.142 g) and pure enough to be used in the next step without further purification.

^1H NMR: (C^2HCl_3), δ : 4.25 (dq, O- CH_2 , $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, $^3\text{J}(\text{H}-^{13}\text{C})=3.2$ Hz, 2H); 4.17–4.24 (m, OH, 1H), 2.27 (dd, $\text{CH}_3-^{13}\text{C}$, $^3\text{J}(\text{H}-^{13}\text{C})=1.1$ Hz, $^2\text{J}(\text{H}-^{13}\text{C})=6.1$ Hz, 3H), 1.36 (dm, $^{13}\text{CH}_3$, $^1\text{J}(\text{H}-^{13}\text{C})=134.2$ Hz, 3H), 1.28 (t, O CH_2 - CH_3 , $^3\text{J}(\text{H}-^1\text{H})=7.1$ Hz, 3H).

Synthesis of sodium 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O ^2H]-3-oxo-4,4,4- $^2\text{H}_3$ -butanoate To a solution of ethyl 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O ^1H]-3-oxobutanoate (1.09 g) in $^2\text{H}_2\text{O}$ (4 mL), 0.4 equivalents of a solution of NaO^2H (2.5 M) in $^2\text{H}_2\text{O}$ were added dropwise in 40 min, using a syringe pump under argon. As soon as the addition was completed, ^1H NMR was carried out on a sample (few μL) in $^2\text{H}_2\text{O}$ in order to calculate the conversion percentage [*ratio* between the amount of hydrolyzed product (quadruplet at 1.60 ppm) and the amount of starting material (quadruplet at 1.7 ppm)]. 1.1 equivalents of a NaO^2H solution (2.5 M) was added over 30 min with the syringe pump. As soon as the addition was completed, an extraction with diethyl ether was carried out in order to remove the by-product coming from the previous step whose NMR signals prevent a good follow-up of the

hydrogen/deuterium ($^1\text{H}/^2\text{H}$) exchange on the 4- CH_3 (2.2 ppm). The $^1\text{H}/^2\text{H}$ exchange on 4- CH_3 was then monitored by ^1H NMR and carried out by successive addition of NaO^2H (2.5 M) until the integral of the doublet corresponding to the CH_3 reaches the value of 0.1 when the quadruplet at 1.6 ppm integrates for 1.5. The reaction was immediately neutralized with a concentrated ^2HCl solution to neutral pH and then buffered with Tris- ^1HCl , (1.0 M, pH 7.5 in $^2\text{H}_2\text{O}$). The concentration of the resulting solution was then determined by ^1H NMR using methanol or acetonitrile as internal reference. The final product (3.02 mmol) was stored at -80 °C.

^1H NMR: (C^2HCl_3), δ : 1.37 (dq, $^{13}\text{CH}_3$, $^1\text{J}(\text{H}-^{13}\text{C})=129.3$ Hz, $^2\text{J}(\text{H}-^{13}\text{C})=3.9$ Hz, 1 H).

Preparation of isotopically labelled HSP90-NTD samples

Escherichia coli BL21-DE3-RIL cells transformed with a pET-28 plasmid encoding the N-Terminal domain of HSP90 α from *Homo Sapiens* (HSP90-NTD) with a His-Tag and a TEV cleavage site were progressively adapted in three stages over 24 h to $\text{M9}/^2\text{H}_2\text{O}$. In the final culture, bacteria were grown at 37 °C in M9 medium with 99.85% $^2\text{H}_2\text{O}$ (Eurisotop), 1 g/L $^{15}\text{N}^1\text{H}_4\text{Cl}$ (Sigma Aldrich) and 2 g/L D-glucose- d_7 (for U- ^2H , ^{12}C , ^{15}N) HSP90-NTD samples) or D-glucose- $^{13}\text{C}_6$ - d_7 (CIL) (for U- ^2H , ^{13}C , ^{15}N) HSP90-NTD samples).

For methyl specifically labelled samples, the methyl labelled precursors or amino-acids were added to the media when the O.D at 600 nm reached 0.6 (Kerfah et al. 2015c):

- Labelling scheme (A): for production of the U- ^2H , ^{15}N , ^{13}C , Ile-[2,3,4,4- $^2\text{H}_4$; 1,2,3,4- $^{13}\text{C}_4$; $^{13}\text{C}^1\text{H}_3$] $^{\delta 1}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\gamma 2}$, Leu-[2,3,3,4- $^2\text{H}_4$; 1,2,3,4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$], Val-[2, 3- $^2\text{H}_2$; 1,2,3- $^{13}\text{C}_3$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$] HSP90-NTD, a solution containing the sodium 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O ^2H]-3-oxo-4,4,4- $^2\text{H}_3$ -butanoate precursor was added at a concentration of 172 mg/L 1 h before induction. 40 min later (20 minutes before induction) a solution containing 60 mg/L of sodium (S)-2-hydroxy-2-(1',1'- $^2\text{H}_2$), 1',2'- $^{13}\text{C}_2$)ethyl-3-oxo-1,2,3- $^{13}\text{C}_3$ -4,4,4- $^2\text{H}_3$ butanoate (Kerfah et al. 2015a) was added to the medium.
- Labelling scheme (B): for production of the U- ^2H , ^{15}N , ^{12}C , Ala- $^{13}\text{C}^1\text{H}_3$] $^{\beta}$, Met- $^{13}\text{C}^1\text{H}_3$] $^{\epsilon}$, Leu/Val- $^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$, Ile- $^{13}\text{C}^1\text{H}_3$] $^{\delta 1}$, Thr- $^{13}\text{C}^1\text{H}_3$] $^{\gamma}$ HSP90-NTD, a HLAM-A $^{\beta}$ I $^{\delta 1}$ M $^{\epsilon}$ LV $^{\text{pro-S}}$ T $^{\gamma}$ kit, purchased from NMR-Bio, was added before induction according to the manufacturer's protocol.
- Labelling scheme (C): U- ^2H , ^{15}N , ^{12}C] samples labelled on a single type of methyl group were produced in small scales (21 mL) to identify A $^{\beta}$, M $^{\epsilon}$, T $^{\gamma}$ methyl type (3 samples) or to complete assignment using single point mutants (Amero et al. 2011) (33 samples, list of mutants

presented in the legend of Fig. S4). Single point amino acid mutations were generated by GeneCust. For each of these samples a single type of methyl groups was labelled by addition of the corresponding NMR-Bio kit (SLAM-A^β, SLAM-M^ε, SLAM-I^{δ1}, SLAM-T^γ or DLAM-LV^{proS}) in M9²H₂O media 1 h before induction.

Protein production was induced by the addition of IPTG to a final concentration of 0.5 mM. The cultures were grown overnight at 20 °C before harvesting. Cells were collected by centrifugation at 5500 g for 20 min at 4 °C then lysed by sonication on ice in a buffer containing 20 mM phosphate sodium buffer at pH 7.4, 0.5 M NaCl, 0.05% β-ME, antiprotease (cOmplete® EDTA free, 1 tablet for 50 mL), 50 µg/mL DNase (Sigma Aldrich), 50 µg/mL RNase (Euromedex), and 0.25 mg/mL Lysozyme (Euromedex). After removal of cell debris by centrifugation (45,000×g, 30 min, 4 °C), the supernatant was purified using an affinity chromatography step (Ni-NTA, Superflow, QIAGEN) (labelling scheme A, B and C), followed by a size exclusion chromatography step (16/600 Superdex 75 PG, GE Healthcare) (labelling schemes A and B only). The gel filtration column was run with an isocratic step of the NMR buffer (20 mM Hepes, 150 mM NaCl, 1 mM TCEP, pH 7.5).

The HSP90-NTD proteins were concentrated, using an Amicon® 4 Centrifugal Filter Unit with a 10,000 MWCO (Merck), either in a 90%/10% ¹H₂O/²H₂O or in a 100% ²H₂O buffer containing 20 mM Hepes, 150 mM NaCl, 1 mM TCEP, pH 7.5. For labelling schemes A and B, samples were concentrated to 0.5 mM and 200 µL of each sample was loaded in 4 mm shigemi tube. The wild type and single point mutants of HSP90-NTD proteins labelled on only one methyl type (labelling scheme C) were concentrated at [0.1–0.4] mM and 40 µL of each sample was loaded in a 1.7 mm NMR tube.

NMR Spectroscopy

All NMR experiments acquired on HSP90-NTD samples were recorded at 298 K. 2D ¹H-¹³C SOFAST methyl TROSY (Amero et al. 2009) experiments to identify each methyl type as well as to assign individual methyl signals using single point mutants were recorded for an average duration of ~1.5 h each, on a spectrometer operating at a ¹H frequency of 850 MHz and equipped with a 1.7 mm cryogenically cooled, pulsed-field-gradient triple-resonance probe. All other NMR experiments were acquired using Bruker Avance III HD spectrometers equipped with 5 mm cryogenic probes (operating at a ¹H frequency of 600 or 950 MHz).

The 3D HCC, HC(C)C and HC(CC)C experiments (Tugarinov and Kay 2003; Ayala et al. 2009, 2012; Mas et al. 2013) were acquired on a spectrometer operating at a ¹H frequency of 600 MHz for a total duration of 4 days,

using a 0.5 mM sample of U-[²H, ¹⁵N, ¹³C], Ile-[2,3,4,4-²H₄; 1,2,3,4-¹³C₄; ¹³C¹H₃]^{δ1}/[¹²C²H₃]^{γ2}], Leu-[2,3,3,4-²H₄; 1,2,3,4-¹³C₄; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}], Val-[2,3-²H₂; 1,2,3-¹³C₃; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}] HSP90-NTD. The interscan delay was adjusted to 0.5–0.6 s, the heteronuclear ¹H → ¹³C transfer delay was set to 4 ms (1/(2 × ¹J_{HC})) and the homonuclear ¹³C → ¹³C transfer delay was fixed to 12.5 ms. The acquisition times were adjusted to 8–10.7 ms in the ¹³C indirect dimension, and to 70 ms in ¹H direct dimension.

For the sequential assignment of backbone resonances, a set of 6 BEST-TROSY 3D triple resonance experiments (HNCA, HN(CA)CB, HNCO, HN(CA)CO, HN(CO)CA and HN(COCA)CB (Favier and Brutscher 2019) were acquired on a Bruker Avance III HD spectrometer equipped with a cryogenic probe and operating at a ¹H frequency of 600 MHz for a total duration of 11 days using a 0.5 mM sample of the U-[²H, ¹⁵N, ¹³C] HSP90-NTD.

The 3D CCH HMQC-NOESY-HMQC NMR experiment (Tugarinov et al. 2005; Törner et al. 2020) was recorded over 3 days on a spectrometer operating at a ¹H frequency of 950 MHz using a 0.5 mM sample of U-[²H, ¹⁵N, ¹²C], Ala-[¹³C¹H₃]^β, Met-[¹³C¹H₃]^ε, Leu/Val-[¹³C¹H₃]^{pro-S}, Ile-[¹³C¹H₃]^{δ1}, Thr-[¹³C¹H₃]^γ HSP90-NTD. The interscan delay was set to 1.1 s. The heteronuclear ¹H → ¹³C transfer delay was set to 4 ms (1/(2 × ¹J_{HC})). The acquisition times in the ¹³C indirect dimension were set to 24.6 ms, (t_{1max}) and to 18.6 ms (t_{2max}). In the ¹H direct dimension t_{3max} was fixed to 80 ms. The NOE mixing period was set to 500 ms to detect a maximum number of long-range intermethyl NOEs.

Data processing and analysis

All data were processed and analyzed using nmrPipe/nmrDraw (Delaglio et al. 1995) and CcpNMR (Vranken et al. 2005). Automated methyl assignment was performed using MAGIC software (Monneau et al. 2017) using the reference structure of HSP90-NTD (PDB: 1YES). Input NOE lists for MAGIC were created with CcpNMR. MAGIC was run with a score threshold factor of 1 and distance thresholds of 7–10 Å using all inter methyl NOEs detected (S/N threshold of 5 was used) and given the assignment of isoleucines, leucines and valines previously obtained by the three 'out and back' HCC experiments as well as alanine, methionine and threonine methyl groups assigned by mutagenesis as additional input.

Results and discussion

To assign methyl groups of large perdeuterated proteins, previously assigned backbone resonances of these proteins can be used (Tugarinov and Kay 2003; Ayala et al. 2012; Mas et al. 2013). Nonetheless, to do so, methyl groups need to be

connected, via a linear chain of ^{13}C , to the backbone atoms in order to be able to apply optimized experiments to high molecular weight proteins. Strategies to label stereospecifically leucine and valine pro-*R* methyl groups and to connect them to backbone nuclei have already been proposed (Mas et al. 2013). However, it has to be noted that pro-*S* methyl groups are often chosen over pro-*R* methyl groups as they are both easier and cheaper to label stereospecifically (Gans et al. 2010). In order to simplify assignment of pro-*S* methyl groups using already assigned backbone resonances and to avoid the need of an additional sample to link the pro-*S* methyl to the pro-*R* one, a sample connecting the pro-*S* $^{13}\text{C}^1\text{H}_3$ -methyl groups to the backbone atoms by a linear ^{13}C -chain and labelled with $^{12}\text{C}^2\text{H}_3$ on pro-*R* methyl moieties to avoid signal loss would be optimal.

Synthesis of optimally labelled acetolactate precursors and proteins

Taking into account the specificity of leucine/valine metabolic pathway in *E. coli*, such an optimal labelling scheme can be achieved in $\text{M9}^2\text{H}_2\text{O}$ medium using $^{13}\text{C}^2\text{H}$ glucose (Kerfah et al. 2015c) as a carbon source together with 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O^2H]-3-oxo-4,4,4- $^{2}\text{H}_3$ -butanoate as suitably labelled acetolactate precursor. However, this latter cannot be synthesized from commercially available materials by the traditional route starting from acetoacetate (Gans et al. 2010) since the corresponding labelled starting material is not commercially available. Indeed, acetolactate chemical synthesis is achieved by reaction of iodomethane on acetoacetate (Gans et al. 2010). Whilst both $^{13}\text{C}^1\text{H}_3$ labelling of the methyl substituent in position 2 and

deuteration of the methyl group in position 4 can be obtained using $^{13}\text{C}^1\text{H}_3\text{I}$ as a starting synthesis material and hydrogen/deuterium exchange in controlled basic conditions (Gans et al. 2010), respectively, the ^{13}C labelling of only the first three carbons of the main chain using commercially available labelled acetoacetate materials is not achievable. Therefore, as acetoacetate can be obtained by condensation of two acetate moieties (Epstein et al. 1977), we decided to set up a synthesis of dissymmetrically labelled acetoacetate starting from commercially available ethyl 1, 2- $^{13}\text{C}_2$ -acetate with 1- ^{13}C -acetyl chloride. Based on reported procedures, we established a 4-step synthesis (Fig. 1) allowing to prepare the desired precursor with an overall yield of 27%. In brief, the dissymmetry is achieved by the Claisen condensation of ethyl-1,2, $^{13}\text{C}_2$ -acetate with 1- ^{13}C -acetyl chloride using an optimization of a reported procedure (Epstein et al. 1977) (a), followed by an alkylation in position 2 using $^{13}\text{C}^1\text{H}_3\text{I}$ (b) and a subsequent hydroxylation in position 2 (c). Finally, the last step combining both saponification of the ester and a hydrogen/deuterium exchange in position 4 is performed under controlled basic conditions (d). This last step is very delicate and requires a fine control of the basic condition as a methyl rearrangement above a pH of 13.5 can take place resulting in the interconversion of both methyl groups (Gans et al. 2010). Steps (b) and (c) are not stereoselective, hence, products of these latter steps were produced as racemic mixtures. The optimally labelled acetolactate, unstable at room temperature, was aliquoted and stored at $-80\text{ }^\circ\text{C}$.

The synthesized acetolactate precursor was incorporated in *E. coli*. $\text{M9}^2\text{H}_2\text{O}$ culture media without any further purification steps to label the overexpressed protein. Frozen acetolactate vials were thawed right before

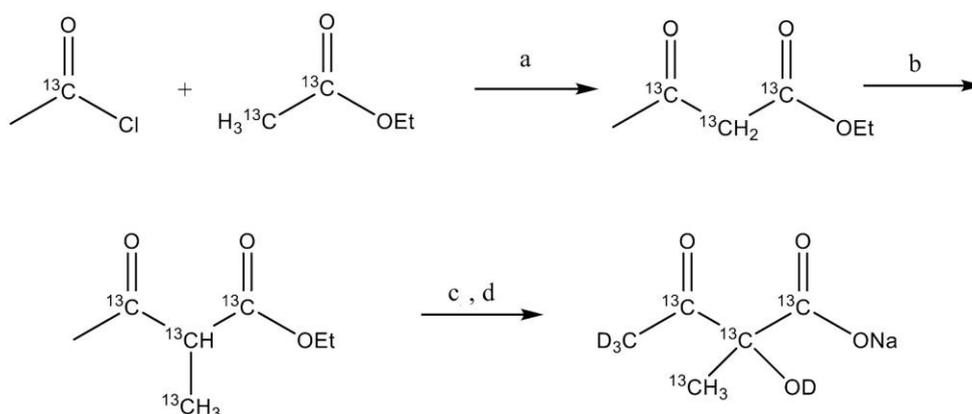


Fig. 1 Synthetic scheme for preparation of specifically labelled sodium 1,2,3- $^{13}\text{C}_3$ -2- $^{13}\text{C}^1\text{H}_3$ -2-[O^2H]-3-oxo-4,4,4- $^{2}\text{H}_3$ -butanoate. **a** i—LiHMDS (2.1 equiv.), dry THF, $-78\text{ }^\circ\text{C}$; ii—1,2- $^{13}\text{C}_2$ -ethyl acetate (1 equiv.), $-78\text{ }^\circ\text{C}$, 15 min.; iii—1- ^{13}C -acetyl chloride

(1.0 equiv.), $-78\text{ }^\circ\text{C}$, 30 min.; iv— ^1HCl 20%; **b** K_2CO_3 (1.1 equiv.), $^{13}\text{C}^1\text{H}_3\text{I}$ (1.1 equiv.), $0\text{ }^\circ\text{C}$, 18 h, EtOH ; **c** Cs_2CO_3 (0.2 equiv.), $\text{P}(\text{OEt})_3$ (0.2 equiv.), O_2 , DMSO, 20 h. **d** i— NaOH (2.5 M), $^2\text{H}_2\text{O}$; ii— ^2HCl 35%; Tris buffer pH 7.5; 27% overall yield

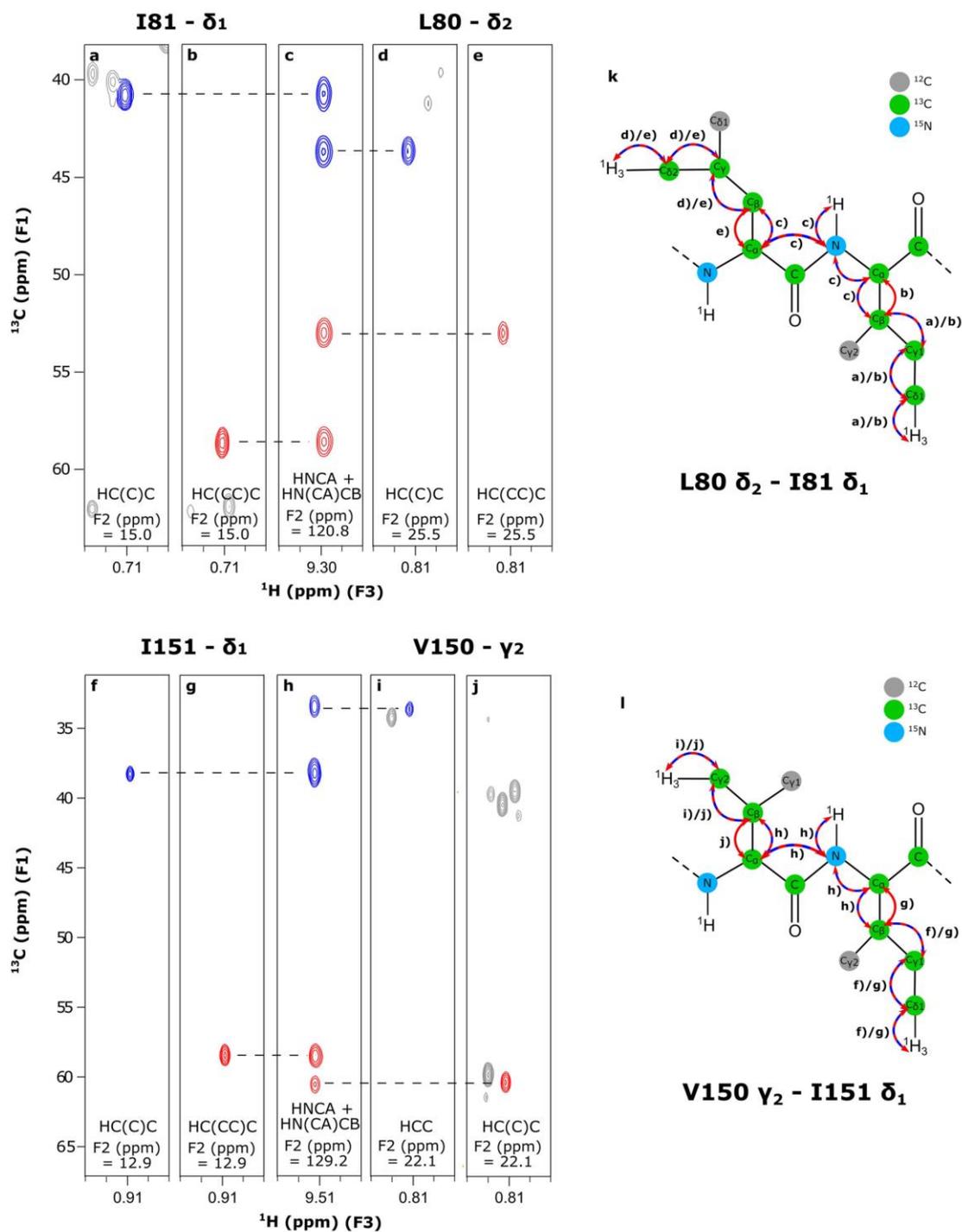


Fig. 2 Assignment transfer from the backbone to Leu^{pro-S}, Val^{pro-S} and Ile- δ_1 methyl groups of HSP90-NTD. Examples of 2D-extracts from 3D 'out and back' HCC (i), HC(C)C (a, d, f and j) and HC(CC)C (b, e and g) experiments correlating ^1H (F_3) and ^{13}C (F_2) methyl resonances with $^{13}\text{C}_{\beta}$ (blue) or $^{13}\text{C}_{\alpha}$ (red) in F_1 dimension. Panels c and h display the corresponding 2D HNCA and HN(CA)CB extracts for Ile-81 (c) and Ile-151 (h) allowing to connect Leu-80- δ_2 (pro-S), Ile-81- δ_1 , Val-150- γ_2 (pro-S), Ile-151- δ_1 methyl groups to previously assigned backbone atoms. 3D spectra were recorded on an NMR spectrometer operating at a proton frequency of 600 MHz using the U- ^{2}H , ^{15}N , ^{13}C], Ile-[2,3,4,4- $^2\text{H}_4$; 1,2,3, 4- $^{13}\text{C}_4$; $^{13}\text{C}^1\text{H}_3$] $^{\delta_1}/[^{12}\text{C}^2\text{H}_3]^{\gamma_2}$], Leu-[2,3,3,4- $^2\text{H}_4$; 1,2,3,4- $^{13}\text{C}_4$; $^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}/ [^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}$], Val-[2,3- $^2\text{H}_2$; 1,2,3- $^{13}\text{C}_3$; $^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}/ [^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}$] labelled sample or U- ^{2}H , ^{15}N , ^{13}C] labelled sample (3D HNCA and HN(CA)CB). k and l represent magnetization transfer schemes correlating with the strips (a, b, c, d, e) and (f, g, h, i, j), respectively. Except when specified, all hydrogen atoms are ^2H

addition in the culture medium to avoid degradation or methyl rearrangement. It has to be noted that only the 2-(S) stereoisomer of acetolactate is converted in vivo by ketol-acid reductoisomerase (EC1.1.1.86) and dihydroxy-acid dehydratase (EC 4.2.1.9) to form the stereospecifically labelled 2-keto-isovalerate. This latter is afterwards directly converted into valine or combined with $^{13}\text{C}/^2\text{H}$ -pyruvate, derived from the $^{13}\text{C}/^2\text{H}$ -glucose, to produce leucine with the desired labelling pattern. The 2-(R) stereoisomer of acetolactate is, itself, not a substrate of ketol-acid reductoisomerase and hence, induces no scrambling (Gans et al. 2010). The synthesized 1,2,3- $^{13}\text{C}_3$]-2- $^{13}\text{C}^1\text{H}_3$]-2-[O ^2H]-3-oxo-4,4,4- $^{2}\text{H}_3$]-butanoate can be mixed with other known precursors allowing also to connect methyl groups, such as Ile- δ_1 (Kerfah et al. 2015a, b; Törner et al. 2020), Ile- γ_2 (Ayala et al. 2012) or Ala- β (Ayala et al. 2009; Kerfah et al. 2015a; Törner et al. 2020), to backbone nuclei using a linear ^{13}C chain. In this study, we chose to label our sample on both leucine/valine pro-S and isoleucine- δ_1 methyl moieties, by adding Ile- δ_1 precursor (sodium (S)-2-hydroxy-2-(1',1'- $^{2}\text{H}_2$), 1', 2'- $^{13}\text{C}_2$)ethyl-3-oxo-1,2,3- $^{13}\text{C}_3$]-4,4,4- $^{2}\text{H}_3$]-butanoate – (Kerfah et al. 2015a, b) together with our new optimized acetolactate. The isoleucine precursor was added in the culture medium 40 min later than the new precursor in order to take into account co-incorporation incompatibilities between both the leucine/valine precursor and the isoleucine one. Indeed, enzymes from ILV-pathway have a tendency to process preferentially isoleucine precursor instead of leucine/valine precursors (Kerfah et al. 2015b, c). Incorporation of both precursors during the protein expression did not lead to a significantly different HSP90-NTD yield with regards to the yields obtained in standard M9/ $^2\text{H}_2\text{O}$ media. No scrambling was detected neither to the pro-R methyls groups of leucine and valine nor to the isoleucine- γ_2 site (Fig. S1a).

Connection of $^{\delta_1}\text{L}^{\delta_2}\text{V}^{\gamma_2}$ methyl groups to C_{α} and C_{β} atoms

Using this new precursor, it is possible to directly correlate pro-S methyl groups of leucine and valine residues to their respective C_{α} and C_{β} . We decided to apply this strategy to the N-terminal domain of human HSP90 (HSP90-NTD), an extensively studied protein, whose isoform assignment (α and β), including partial assignment of its methyl groups, is available (Jacobs et al. 2006; Elif Karagöz et al. 2011; Park et al. 2011; Lescanne et al. 2017, 2018). Here, we have focused on the N-terminal domain of the α isoform, a 29 kDa protein that contains 20 isoleucine, 18 leucine and 11 valine residues. This dynamic protein is particularly challenging to assign using automatic methyl assignment methods and reported success rates are ranging from 27 to 69% (Pritišanac et al. 2017, 2019, 2020; Monneau et al. 2017). In our hands, only 34% of the methyl groups (30/87) could be assigned automatically with (1) a single assignment, (2) a high NOE assignment completeness of the strip related to each peak (> 50%) and (3) a high total confidence score value (≥ 7) (Table S1) using the HSP90-NTD X-ray structure (PDB: 1YES) and experimentally detected NOE network. Therefore, this protein is a good candidate to assess our experimental strategy based on new precursors.

To do so, 200 μL at 0.5 mM of an optimally labelled sample was used to acquire three 'out and back' HCC, HC(C)C and HC(CC)C experiments (Tugarinov and Kay 2003; Ayala et al. 2009, 2012; Mas et al. 2013) connecting labelled $^{\delta_1}\text{L}^{\delta_2}\text{V}^{\gamma_2}$ methyls groups to $^{\text{I}}\text{L}^{\text{I}}\text{V}^{\text{I}}$, $^{\text{I}}\text{L}^{\text{I}}\text{V}^{\text{I}}$ and $^{\text{I}}\text{L}^{\text{I}}$ resonances. 100 and 94% of the expected C_{β} and C_{α} coherences, respectively, were observed for the 29 kDa HSP90-NTD at 298 K (τ_{C} ca. 20 ns) (Fig. 2). The three missing C_{α} resonances correspond to residues L56, I26 and I110, two of them being affected by extensive line broadening due to conformational exchange. With such a high percentage of observed C_{α} and C_{β} resonances we demonstrate the applicability of this method for medium size proteins. In order to validate the strategy for larger proteins, the labelling scheme was applied to the 87 kDa hetero hexameric protein prefoldin from *Pyrococcus horikoshii*, containing 2 α and 4 β -subunits. Only the β -subunits were labelled with the optimal labelling schemes described above, whilst the α -subunits remained perdeuterated (Fig. S1b). Three HCC experiments were collected on this 0.2 mM sample of prefoldin at 310 K (τ_{C} ca. 60 ns). 95 and 60% of the expected C_{β} and C_{α} coherences, respectively, were observed (Fig. S2) despite the high molecular weight of the protein and the presence of doubled peaks due to the presence of two inequivalent β -subunits in the $\alpha_2\beta_4$ hexameric prefoldin (Ohtaki et al. 2008). Such HCC experiments were also acquired at 343 K (τ_{C} ca. 30 ns) on this hyperthermophilic prefoldin sample, enabling transfer

of assignment between backbone atoms and methyl groups (Törner et al. 2021).

Application to the sequence specific assignment of I^{δ1}L^{δ2}V^{γ2} methyl groups of HSP90-NTD

Backbone sequential assignment was performed using 6 'BEST-TROSY' triple resonance experiments (Favier and Brutscher 2019). C_α and C_β resonances were assigned for 89% and 80% of the residues of HSP90-NTD respectively, excluding the loosely structured N-terminal [1–16] and C-terminal [225–236] regions. The segment L103-T115, that covers the ligand binding site, is invisible by NMR due to dynamics in the μs-ms timescale. Transfer of sequentially assigned backbone resonances to isoleucine-δ₁, leucine-δ₂ and valine-γ₂ methyl groups was achieved using 'out and back' HCC experiments acquired on a U-[²H, ¹⁵N, ¹³C], Ile-[2,3,4,4-²H₄; 1,2,3,4-¹³C₄; ¹³C¹H₃]^{δ1}/[¹²C²H₃]^{γ2}, Leu-[2,3,3,4-²H₄; 1,2,3,4-¹³C₄; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}], Val-[2,3-²H₂; 1,2,3-¹³C₃; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}] labelled HSP90-NTD sample. 2D extracts from the HCC experiments were compared with the corresponding ones from the 3D HNCA and HN(CA)CB experiments and using C_α and C_β resonances, all the I^{δ1}L^{δ2}V^{γ2} methyl groups could be unambiguously connected to previously assigned backbone atoms (Fig. 2). The assignment was transferred in one step, very simply and efficiently without having to correct for the isotopic shifts (Kerfah et al. 2015a). One must note that the methyl-δ₁ of Ile-33 and Ile-128 are superimposed in the 2D methyl-TROSY spectrum, but were unambiguously connected to the C_α resonances of both amino acids (Fig. S3). Remained unassigned only methyl groups of L103, I104, L107 and I110 for which backbone atoms are NMR-invisible due to extensive conformational exchange. Therefore, single

point mutagenesis was used to assign three of these last four I^{δ1} or L^{δ2} resonances (I104, L107 and I110) (Fig. 3). The remaining residue, L103, was assigned by a careful re-analysis of both HCC and backbone triple resonance experiments performed after the assignment of I104.

Sequence specific assignment of A^β M^ε T^γ methyl groups of HSP90-NTD

In the previous U-[²H, ¹⁵N, ¹³C], Ile-[2,3,4,4-²H₄; 1,2,3,4-¹³C₄; ¹³C¹H₃]^{δ1}/[¹²C²H₃]^{γ2}, Leu-[2,3,3,4-²H₄; 1,2,3,4-¹³C₄; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}], Val-[2,3-²H₂; 1,2,3-¹³C₃; [¹³C¹H₃]^{pro-S}/[¹²C²H₃]^{pro-R}] HSP90-NTD sample, only I^{δ1}L^{δ2}V^{γ2} methyl groups were connected to backbone by a linear ¹³C-chain. We did not incorporate labelled alanine in HSP90-NTD culture for the sample used to acquire the HCC experiments although it is commercially available with an optimal labelling pattern (1,2,3-¹³C₃, 2-²H-Ala). We recommend for future studies to incorporate labelled alanine in the culture medium to decrease the numbers of mutants required to complete the assignment. Regarding methionine and threonine residues, their assignment cannot be undertaken using HCC experiments. Indeed, the sulfur atom present on methionine residues prevents the use of HCC experiments to assign methionine methyl moieties from available backbone assignment and the Thr-[1,2,3-¹³C, 2,3-²H₂, ¹³C¹H₃-γ] labelled amino acid is not commercially available impeding the use of HCC 3D experiments to transfer assignment from backbone to threonine methyl groups in large proteins. Therefore, in order to assign the remaining 38 A^β M^ε and T^γ methyl groups, we decided to use a combination of both single point mutations and through space intermethyl NOE correlation peaks. To assign the remaining A^β, M^ε and T^γ methyl groups using inter-methyl NOE-connectivities a

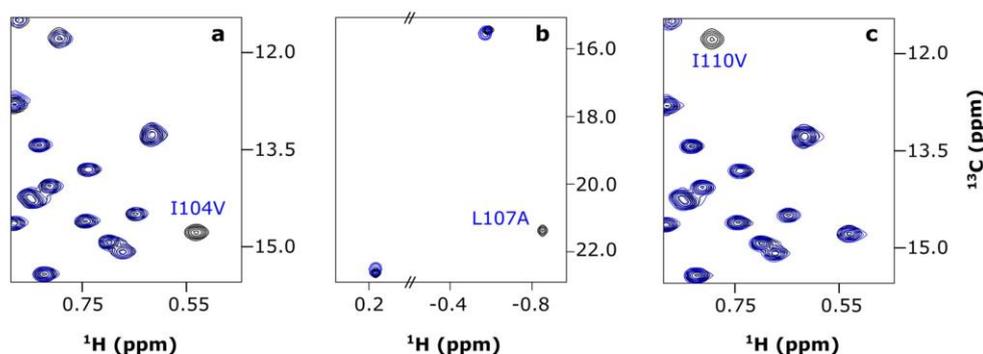


Fig. 3 Assignment of HSP90-NTD methyl groups belonging to the flexible loop covering ATP binding site. The 2D SOFAST methyl TROSY spectra were recorded using either the isoleucine to valine mutant samples using U-[²H, ¹²C, ¹⁵N]-Ile-[¹³C¹H₃]^{δ1} labelling scheme, or the leucine to alanine mutant sample using U-[²H, ¹²C,

¹⁵N]-Leu/Val-[¹³C¹H₃]^{pro-S} labelling scheme. Spectra were recorded at 298 K on a NMR spectrometer operating at a proton frequency of 850 MHz. **a** The HSP90-NTD mutant spectra of I104V. **b** L107A. **c** I110V. Each mutant spectrum extract (dark blue) was superimposed with the wild type protein extract (black)

U- ^2H , ^{15}N , ^{12}C], Ala- $^{13}\text{C}^1\text{H}_3]^\beta$, Met- $^{13}\text{C}^1\text{H}_3]^\epsilon$, Leu/Val- $^{13}\text{C}^1\text{H}_3]^\text{pro-S}$, Ile- $^{13}\text{C}^1\text{H}_3]^\delta1$, Thr- $^{13}\text{C}^1\text{H}_3]^\gamma$ HSP90-NTD sample was required. Indeed, intermethyl NOE connectivities between previously assigned isoleucine, leucine and valine residues and unassigned A $^\beta$, M $^\epsilon$ and T $^\gamma$ methyl groups simplify the assignment of methionine, threonine and alanine methyl moieties.

First small-scale samples with only one type of methyl group labelled A $^\beta$, M $^\epsilon$ or T $^\gamma$ were prepared to identify the amino acids type corresponding to each correlation remaining to assign in the 2D methyl-TROSY spectrum. Then, we overexpressed and purified in small scale 30 single point $^{13}\text{C}^1\text{H}_3$ -labelled mutants. In order to minimize secondary chemical shifts replacement amino acids that were structurally similar to the substituted amino acid were chosen (Crublet et al. 2014) (Fig. S4).

Each of these samples containing 100 to 500 μg of $^{13}\text{C}^1\text{H}_3$ -labelled single point mutant of HSP90-NTD were used to acquire a 2D SOFAST-methyl-TROSY spectrum (Amero et al. 2009, 2011) on a NMR spectrometer operating at a ^1H frequency of 850 MHz and equipped with a sample changer and a 1.7 mm cryogenic probe head. This mutant library allowed us to assign unambiguously 27 methyl groups. Three spectra of single point mutants were more complex to analyze due to chemical shift perturbations that result from the introduced mutation (Fig. S5). Out of the 38 A $^\beta$, M $^\epsilon$ and T $^\gamma$ methyl groups, eleven remained unassigned after the first analysis of the mutant library, among them eight methyls (one M $^\epsilon$, two A $^\beta$ and five T $^\gamma$ methyl groups) for which mutants were not available and three T $^\gamma$ methyl groups whose mutant spectra were challenging to analyze. The sole unassigned methionine methyl signal was unambiguously assigned as the last remaining methionine residue (M98).

To complete the assignment for the last 10 methyl groups, a 0.5 mM sample of U- ^2H , ^{15}N , ^{12}C], Ala- $^{13}\text{C}^1\text{H}_3]^\beta$, Met- $^{13}\text{C}^1\text{H}_3]^\epsilon$, Leu/Val- $^{13}\text{C}^1\text{H}_3]^\text{pro-S}$, Ile- $^{13}\text{C}^1\text{H}_3]^\delta1$, Thr- $^{13}\text{C}^1\text{H}_3]^\gamma$ labelled HSP90-NTD was prepared to acquire a 3D HMQC-NOESY-HMQC. A total of 344 intermethyl NOEs cross peaks with a $S/N \geq 5$ were detected between methyls distant by up to 10 Å. Examples of intermethyl NOE and a matrix presenting all the observed NOEs are displayed in Fig. 4. To assign the remaining 2 alanines and 8 threonines, the NOE connectivities and the previously assigned methyls (49 I $^\delta1$ L $^\delta2$ V $^\delta2$ and 28 A $^\beta$ M $^\epsilon$ and T $^\gamma$ methyl groups) were used as input for the program MAGIC. Six methyl groups [2 alanines (A141 and A145) and 4 threonines (T90, T88, T115 and T149)] were unambiguously assigned with both a high percentage of NOE correlation peaks assigned and a high confidence score. Four threonines (T94, T109, T176 and T184) were left either with multiple assignments or an assignment with a low confidence score. However, taking into account the information obtained from the NOE-based assignment, the mutant spectra displaying chemical shift perturbations (Fig. S5) were carefully reanalyzed allowing us to assign the four remaining threonine methyl signals.

Finally, all the 87 methyl groups of HSP90-NTD were assigned using both isoleucine precursor and the newly labelled acetolactate for isoleucine, leucine and valine methyl groups and this mixed NOE/mutants strategy for alanine, methionine and threonine methyl moieties (Fig. 5, Table S2). It can be noted that, there is, indeed, an overlap of two isoleucines (I33/I128) explaining only 86 visible peaks. Even though dynamics in the intermediate regime broaden the backbone resonances of the segment which is covering the ATP binding site [103–115] beyond the detection

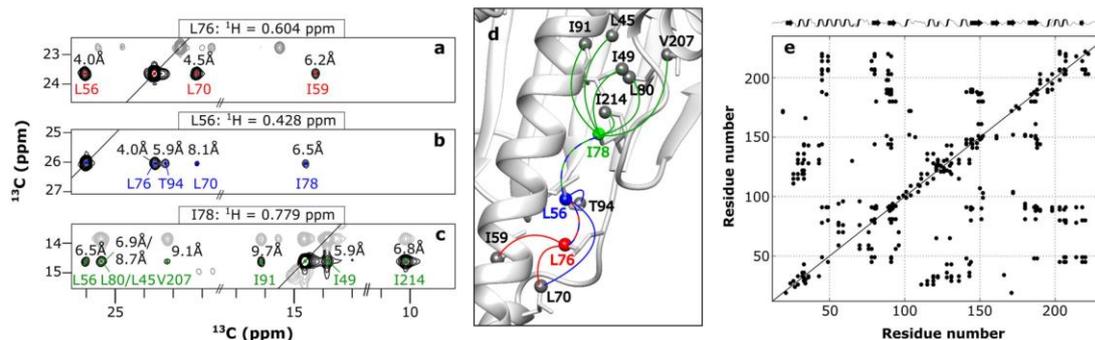


Fig. 4 Detected intermethyl NOEs in human HSP90-NTD. **a–c** Examples of 2D extracts of a 3D HMQC-NOESY-HMQC experiment recorded using U- ^2H , ^{12}C , ^{15}N]-Leu/Val- $^{13}\text{C}^1\text{H}_3]^\text{pro-S}$, Ile- $^{13}\text{C}^1\text{H}_3]^\delta1$, Met- $^{13}\text{C}^1\text{H}_3]^\epsilon$, Ala- $^{13}\text{C}^1\text{H}_3]^\beta$, Thr- $^{13}\text{C}^1\text{H}_3]^\gamma$ HSP90-NTD sample on a NMR spectrometer operating at a proton frequency of 950 MHz. The planes were extracted at the methyl proton frequencies of L76

(**a**), L56 (**b**) and I78 (**c**). The NOEs detected are colored in red, blue and green, respectively. **d** The NOEs detected in 2D extracts presented in panels a–c are displayed on the 3D structure of HSP90-NTD (PDB: 1YES) by lines (red for L76, blue for L56 and green for I78). **e** 2D matrix representing all the HSP90-NTD methyl residue pairs for which NOE cross-peaks have been detected

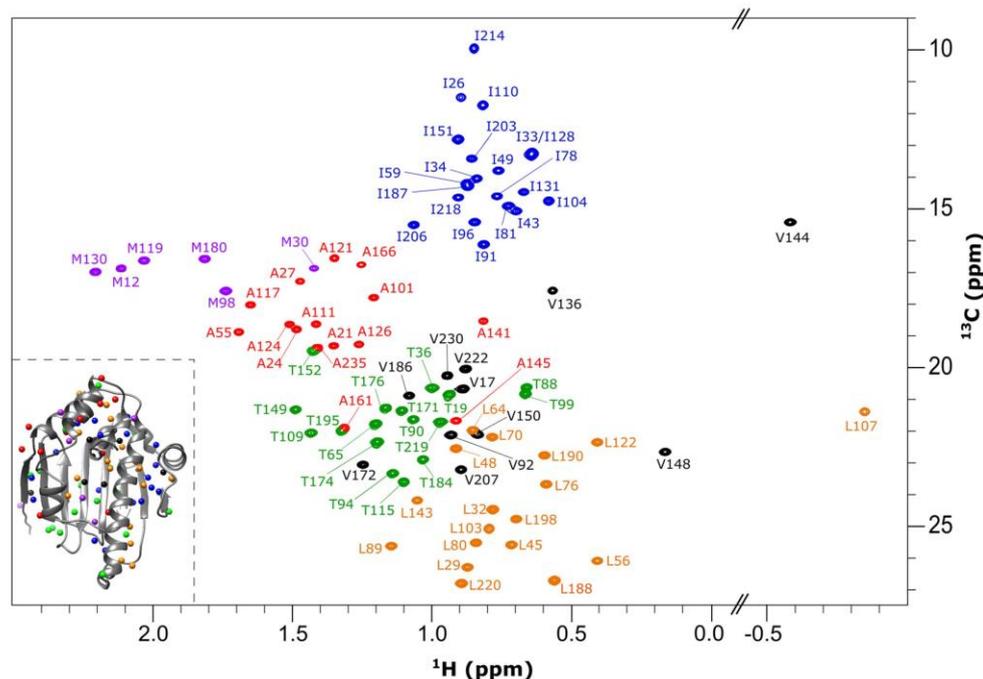


Fig. 5 Assigned 2D ^1H - ^{13}C SOFAST methyl TROSY spectrum of apo HSP90-NTD. Human HSP90-NTD was perdeuterated and specifically $^{13}\text{C}^1\text{H}_3$ -labelled on Leu/Val- $^{13}\text{C}^1\text{H}_3$ $^{\text{pro-S}}$, Ile- $^{13}\text{C}^1\text{H}_3$ $^{\delta 1}$, Met- $^{13}\text{C}^1\text{H}_3$ $^{\epsilon}$, Ala- $^{13}\text{C}^1\text{H}_3$ $^{\beta}$, Thr- $^{13}\text{C}^1\text{H}_3$ $^{\gamma}$ methyl groups. Each signal is annotated with the corresponding residue number. The spectrum was recorded on a NMR spectrometer operating at a proton frequency of

950 MHz. On the bottom left side an insert represents the 3D structure of human HSP90-NTD (PDB: 1YES). The methyl groups are represented by spheres. Alanines, isoleucines, valines, leucines, threonines and methionines are depicted in red, dark blue, black, orange, light green and purple, respectively

threshold, the according methyl probes are visible. The assignment of L103, I104, L107, T109, I110, A111 and T115 render this previously invisible region amenable to NMR studies.

Conclusion

This research reports the synthesis of a new dissymmetric $^{13}\text{C}/^2\text{H}$ -labelled acetolactate, a precursor that allows to connect directly, via linear ^{13}C chains, backbone atoms to the pro-*S* methyl groups of leucine and valine residues. This optimized precursor can be combined with isoleucine precursor and $^2\text{H}/^{13}\text{C}$ -alanine to enable the transfer of assignment from backbone to methyl groups with only one sample without requiring the correction of $^1\text{H}/^2\text{H}$ isotopic chemical shift for the ^{13}C resonances. We expect that this new precursor will ease the assignment of leucine and valine pro-*S* methyl groups of proteins using already assigned backbone resonances as it simplifies the analysis of the NMR experiments. This innovative labelling scheme was applied to the

29 kDa N-terminal domain of human HSP90 protein and to the 87 kDa hetero hexameric prefoldin complex. Using both isoleucine precursor and the newly labelled acetolactate, we managed to simply and efficiently transfer the backbone sequential assignment to all the isoleucine- δ_1 , leucine and valine pro-*S* methyl moieties of HSP90-NTD. This allowed us to confirm or correct the residue specific assignment of most isoleucine, leucine and valine methyl groups (2 assignments were corrected for the 49 ILV residues—Table S2) and the stereospecific assignment of prochiral methyl groups (5 stereospecific assignments were inverted for the 29 leucine and valine residues—Table S2). In addition to the full assignment of $\text{I}^{\delta 1}\text{L}^{\delta 2}\text{V}^{\gamma 2}$ methyl groups, we have used a mixed strategy based on mutagenesis and intermethyl NOEs to assign 38 new methyl resonances corresponding to the $\text{A}^{\beta}\text{M}^{\epsilon}\text{T}^{\gamma}$ methyl moieties of HSP90-NTD. Hence, we show that, despite extended conformational exchange that impedes the complete backbone assignment, we managed to detect and assign signals for all methyl probes including the ones belonging to the segment covering HSP90 ATP binding site.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10858-021-00370-0>.

Acknowledgements The authors thank Dr. R Awad and Mr. L. Imbert for advice and stimulating discussions. This work used the high field NMR and isotopic labelling facilities at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by grants from CEA/NMR-Bio (research program C24990) and by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02).

Data availability The FIDs acquired for this study are available in the biological magnetic resonance databank (bmrdb12) and the assignment have been deposited under the BMRB ID: 50786.

References

- Amero C, Schanda P, Asunción Durá M et al (2009) Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. *J Am Chem Soc* 131:3448–3449. <https://doi.org/10.1021/ja809880p>
- Amero C, Asunción Durá M, Noirclerc-Savoie M et al (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR* 50:229–236. <https://doi.org/10.1007/s10858-011-9513-5>
- Ayala I, Sounier R, Usé N et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* 43:111–119. <https://doi.org/10.1007/s10858-008-9294-7>
- Ayala I, Hamelin O, Amero C et al (2012) An optimized isotopic labelling strategy of isoleucine- γ 2 methyl groups for solution NMR studies of high molecular weight proteins. *Chem Commun* 48:1434–1436. <https://doi.org/10.1039/c1cc12932e>
- Ayala I, Chiari L, Kerfah R et al (2020) Asymmetric synthesis of methyl specifically labelled l-threonine and application to the NMR studies of high molecular weight proteins. *ChemistrySelect* 5:5092–5098. <https://doi.org/10.1002/slct.202000827>
- Chao F-A, Kim J, Xia Y et al (2014) FLAMeNGO 2.0: an enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances. *J Magn Reson* 245:17–23. <https://doi.org/10.1016/j.jmr.2014.04.012>
- Crublet E, Kerfah R, Mas G et al (2014) A cost-effective protocol for the parallel production of libraries of CH₃-specifically labeled mutants for NMR studies of high molecular weight proteins. *Methods Mol Biol* 1091:229–243. https://doi.org/10.1007/978-1-62703-691-7_17
- Delaglio F, Grzesiek S, Vuister GW et al (1995) NMRPipe: a multi-dimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293. <https://doi.org/10.1007/BF00197809>
- Elif Karagöz G, Duarte AMS, Ippel H et al (2011) N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. *Proc Natl Acad Sci USA* 108:580–585. <https://doi.org/10.1073/pnas.1011867108>
- Epstein J, Cannon P, Swidler R, Baraze A (1977) Amplification of cyanide ion production by the micellar reaction of keto oximes with phosphono- and phosphorofluoridates. *J Org Chem* 42:759–762. <https://doi.org/10.1021/jo00424a043>
- Favier A, Brutscher B (2019) NMRlib: user-friendly pulse sequence tools for Bruker NMR spectrometers. *J Biomol NMR* 73:199–211. <https://doi.org/10.1007/s10858-019-00249-1>
- Gans P, Hamelin O, Sounier R et al (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chemie - Int Ed* 49:1958–1962. <https://doi.org/10.1002/anie.200905660>
- Gardner KH, Kay LE, Chinchilla D, Fisher K (1997) Production and incorporation of ¹⁵N, ¹³C, ²H (1 H- δ 1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc* 119:7599–7600
- Gauto DF, Estrozi LF, Schwieters CD et al (2019) Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat Commun* 10:2697. <https://doi.org/10.1038/s41467-019-10490-9>
- Gelis I, Bonvin AMJJ, Keramisanou D et al (2007) Structural basis for signal sequence recognition by the 204-kDa translocase motor SecA determined by NMR. *Cell* 131:756–769. <https://doi.org/10.1016/j.cell.2007.09.039>
- Goto NK, Gardner KH, Mueller GA et al (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated ¹⁵N-, ¹³C-, ²H-labeled proteins. *J Biomol NMR* 13:369–374
- Gross JD, Gelev VM, Wagner G (2003) A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. *J Biomol NMR* 25:235–242
- Hajduk PJ, Augeri DJ, Mack J et al (2000) NMR-based screening of proteins containing ¹³C-labeled methyl groups. *J Am Chem Soc* 122:7898–7904. <https://doi.org/10.1021/ja0003501>
- Isaacson RL, Simpson PJ, Liu M et al (2007) A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. *J Am Chem Soc* 129:15428–15429. <https://doi.org/10.1021/ja0761784>
- Jacobs DM, Langer T, Elshorst B et al (2006) NMR backbone assignment of the N-terminal domain of human HSP90. *J Biomol NMR* 36:52
- Kerfah R, Hamelin O, Boisbouvier J, Marion D (2015a) CH₃-specific NMR assignment of alanine, isoleucine, leucine and valine methyl groups in high molecular weight proteins using a single sample. *J Biomol NMR* 63:389–402. <https://doi.org/10.1007/s10858-015-9998-4>
- Kerfah R, Plevin MJ, Pessey O et al (2015b) Scrambling free combinatorial labeling of alanine- β , isoleucine- δ 1, leucine-proS and valine-proS methyl groups for the detection of long range NOEs. *J Biomol NMR* 61:73–82. <https://doi.org/10.1007/s10858-014-9887-2>
- Kerfah R, Plevin MJ, Sounier R et al (2015c) Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* 32:113–122. <https://doi.org/10.1016/j.sbi.2015.03.009>
- Lapinaite A, Simon B, Skjaerven L et al (2013) The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature* 502:519–523. <https://doi.org/10.1038/nature12581>
- Lescanne M, Skinner SP, Blok A et al (2017) Methyl group assignment using pseudocontact shifts with PARAssign. *J Biomol NMR* 69:183–195. <https://doi.org/10.1007/s10858-017-0136-3>
- Lescanne M, Ahuja P, Blok A et al (2018) Methyl group reorientation under ligand binding probed by pseudocontact shifts. *J Biomol NMR* 71:275–285. <https://doi.org/10.1007/s10858-018-0190-5>
- Mas G, Crublet E, Hamelin O et al (2013) Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. *J Biomol NMR* 57:251–262. <https://doi.org/10.1007/s10858-013-9785-z>
- Mas G, Guan J-Y, Crublet E et al (2018) Structural investigation of a chaperonin in action reveals how nucleotide binding regulates

- the functional cycle. *Sci Adv* 4:eau:4196. <https://doi.org/10.1126/sciadv.aau4196>
- Monneau YR, Rossi P, Bhaumik A et al (2017) Automatic methyl assignment in large proteins by the MAGIC algorithm. *J Biomol NMR* 69:215–227. <https://doi.org/10.1007/s10858-017-0149-y>
- Nerli S, De Paula VS, McShan AC, Sgourakis NG (2021) Backbone-independent NMR resonance assignments of methyl probes in large proteins. *Nat Commun* 12:691. <https://doi.org/10.1038/s41467-021-20984-0>
- Ohtaki A, Kida H, Miyata Y et al (2008) Structure and molecular dynamics simulation of archaeal prefoldin: the molecular mechanism for binding and recognition of nonnative substrate proteins. *J Mol Biol* 376:1130–1141. <https://doi.org/10.1016/j.jmb.2007.12.010>
- Park SJ, Kostic M, Dyson HJ (2011) Dynamic interaction of Hsp90 with its client protein p53. *J Mol Biol* 411(1):158–173. <https://doi.org/10.1016/j.jmb.2011.05.030>
- Pritišanac I, Degiacomi MT, Alderson TR et al (2017) Automatic assignment of methyl-NMR spectra of supramolecular machines using graph theory. *J Am Chem Soc* 139:9523–9533. <https://doi.org/10.1021/jacs.6b11358>
- Pritišanac I, Würz JM, Alderson TR, Güntert P (2019) Automatic structure-based NMR methyl resonance assignment in large proteins. *Nat Commun* 10:4922. <https://doi.org/10.1038/s41467-019-12837-8>
- Pritišanac I, Alderson TR, Güntert P (2020) Automated assignment of methyl NMR spectra from large proteins. *Prog Nucl Magn Reson Spectrosc* 118–119:54–73. <https://doi.org/10.1016/j.pnmrs.2020.04.001>
- Rosenzweig R, Moradi S, Zarrine-Afsar A et al (2013) Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* 339:1080–1083. <https://doi.org/10.1126/science.1233066>
- Sounier R, Blanchard L, Wu Z, Boisbouvier J (2007) High-accuracy distance measurement between remote methyls in specifically protonated proteins. *J Am Chem Soc* 129:472–473. <https://doi.org/10.1021/ja067260m>
- Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445:618–622. <https://doi.org/10.1038/nature05512>
- Stoffregen MC, Schwer MM, Renschler FA, Wiesner S (2012) Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. *Structure* 20:573–581. <https://doi.org/10.1016/j.str.2012.02.012>
- Törner R, Awad R, Gans P et al (2020) Spectral editing of intra- and inter-chain methyl–methyl NOEs in protein complexes. *J Biomol NMR* 74:83–94. <https://doi.org/10.1007/s10858-019-00293-x>
- Törner R, Henot F, Awad R et al (2021) Backbone and methyl resonances assignment of the 87 kDa prefoldin from *Pyrococcus horikoshii*. *Biomol NMR Assignment*. <https://doi.org/10.1007/s12104-021-10029-4>
- Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc* 125:13868–13878. <https://doi.org/10.1021/ja030345s>
- Tugarinov V, Hwang PM, Ollershaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H-13 C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* 125:10420–10428. <https://doi.org/10.1021/ja030153x>
- Tugarinov V, Kay LE (2004a) Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G. *J Am Chem Soc* 126:9827–9836. <https://doi.org/10.1021/ja048738u>
- Tugarinov V, Kay LE (2004b) An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* 28:165–172
- Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional 1H-13 C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J Am Chem Soc* 127:2767–2775. <https://doi.org/10.1021/ja044032o>
- Velyvis A, Ruschak AM, Kay LE (2012) An economical method for production of 2H,13CH3-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PLoS ONE* 7:e43725. <https://doi.org/10.1371/journal.pone.0043725>
- Vranken WF, Boucher W, Stevens TJ et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins Struct Funct Genet* 59:687–696. <https://doi.org/10.1002/prot.20449>
- Xu Y, Matthews S (2013) MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins. *J Biomol NMR* 55:179–187. <https://doi.org/10.1007/s10858-012-9700-z>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2.7. Application des différentes stratégies d'attribution à HSP90-NTD en complexe avec deux ligands

Grâce à une collaboration avec le Dr. Matthias Frech (Merck KGaA), j'ai pu avoir accès pendant ma thèse à deux ligands pour lesquels l'attribution de HSP90-NTD en complexe avec ces derniers a été entreprise et est présentée ci-dessous. L'intérêt de ces deux inhibiteurs sera précisé au chapitre 4.

2.7.1. HSP90-NTD en complexe avec un ligand conservant l'empreinte spectrale de la protéine

HSP90-NTD en complexe avec le ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one présente un spectre des résonances du squelette et un spectre des résonances des groupements méthyles respectivement très similaires à ceux de la protéine dans son état Apo (Figure 2.21).

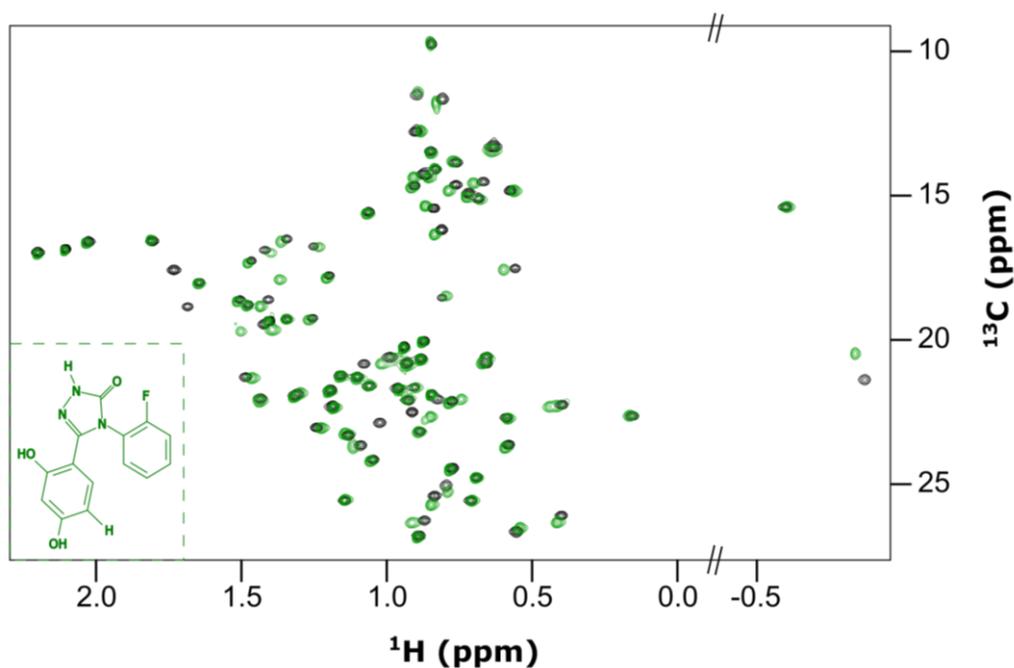


Figure 2.21 : Spectre ^1H - ^{13}C des groupements méthyles de la protéine HSP90-NTD apo en noir superposés avec ceux de la protéine HSP90-NTD liée au ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one en vert à la température de 293 K. La structure de ce dernier est représentée en bas à gauche du spectre.

Toutefois, il a été observé que lorsque la protéine était liée à cet inhibiteur, la stabilité du complexe était plus grande que dans le cas de la protéine apo. Ainsi la température d'acquisition des expériences RMN a pu être augmentée. D'une température maximale de 298 K pour la protéine apo à 0.6 mM il a été possible d'atteindre 308 K avec la protéine liée au 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one pour la même concentration de protéine. Cette augmentation de la température a probablement affecté la contribution de l'échange conformationnel (R_{ex}) à la relaxation transversale (R_2^{eff}). En acquérant de nouvelles expériences de triple résonances à haute température une résonance supplémentaire des groupements amides appartenant au segment [L107-K116] a pu être attribuée pour la protéine en complexe avec cet inhibiteur comparé à la protéine apo (Figure 2.22).

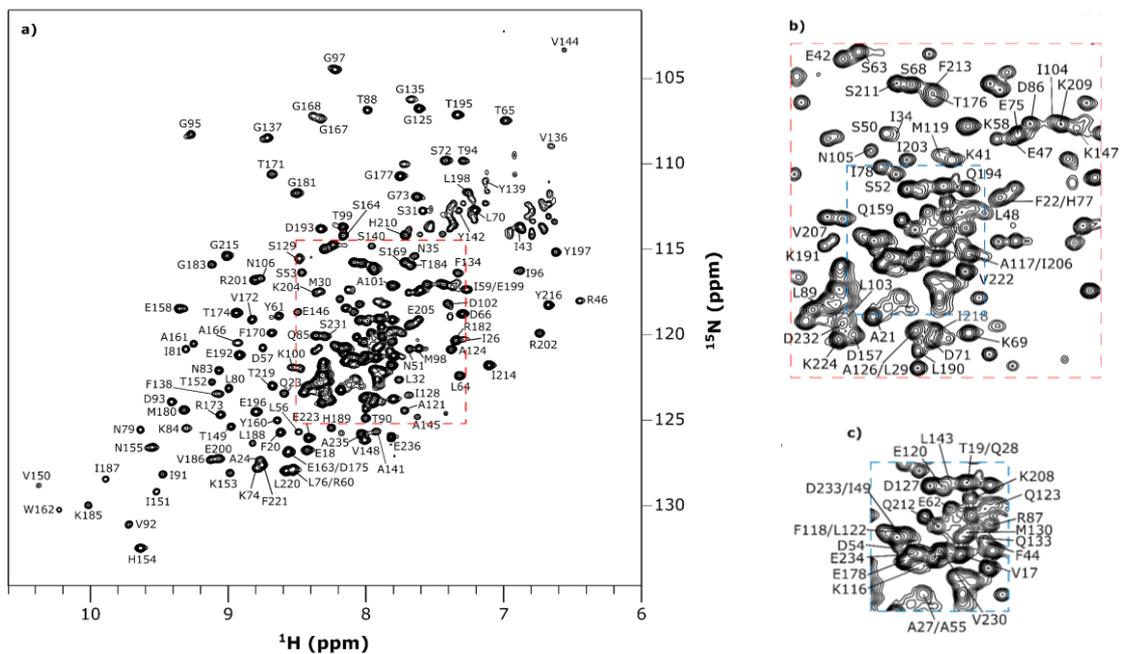


Figure 2.22 : a) Spectre ^{15}N BEST-TROSY de HSP90-NTD liée au ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one, acquis à 308 K et sur un spectromètre opérant à une fréquence proton de 600 MHz. b) et c) représentent des zooms du panneau a). Chaque résonance (a, b, c) est annotée avec le numéro de résidu correspondant.

Tous les groupements méthyles ont été attribués à cette même température pour HSP90-NTD en complexe (Figure 2.23). L'attribution s'est basée (1) sur un transfert de l'attribution des groupements méthyles de HSP90-NTD Apo à HSP90-NTD en complexe à 298 K puis à HSP90 liée à l'inhibiteur à 308 K en suivant l'évolution des pics en fonction de la température par paliers de 5 °C pour la protéine en complexe. L'attribution s'est poursuivie (2) en transférant l'attribution du squelette aux groupement méthyles pour les isoleucines, leucines et valines en employant la même stratégie que développée en 2.6 avec le précurseur optimisé et en utilisant le déplacement chimique de l'atome de C β pour les alanines. Finalement 6 résidus étaient non attribués après ces deux étapes parmi lesquels deux alanines, trois thréonines et une méthionine. HSP90-NTD a été produite avec deux différents marquages marquant spécifiquement un seul type de méthyle à la fois : Ala- β et Thr- γ_2 . En combinant les informations obtenues de ces spectres, l'attribution des groupements méthyles de HSP90-NTD Apo et la structure cristallographique d'HSP90-NTD en complexe (PDB : 5J64 (Amaral et al. 2017)) il a été possible d'attribuer chacun des six résidus restants.

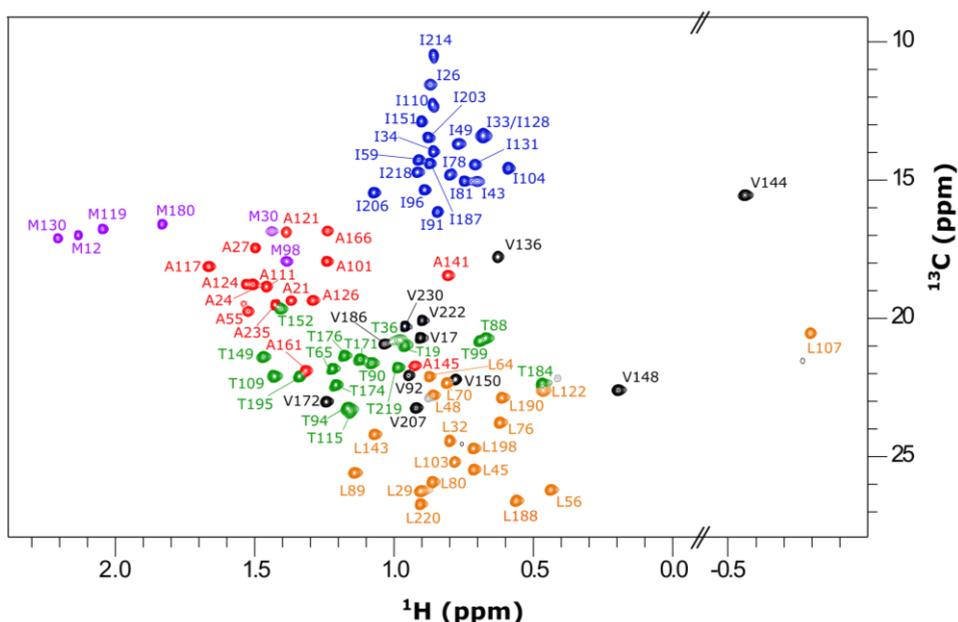


Figure 2.23 : Spectre SOFAST methyl TROSY 2D ^1H - ^{13}C de HSP90-NTD α en complexe avec le ligand 5 - (2,4 - dihydroxy - phenyl) -4-(2-fluoro-phenyl)-2,4-dihydro- [1,2,4] triazol-3-one et acquis à 308 K. Chaque signal est annoté avec l'attribution du résidu correspondant. Les alanines, isoleucines, thréonines, valines, leucines et méthionines sont colorées respectivement en rouge, bleu, vert, noir, orange et violet.

2.7.2 HSP90-NTD en complexe avec un ligand modifiant très largement l'empreinte spectrale de la protéine

HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide présente un spectre des résonances du squelette et un spectre des résonances des groupements méthyles respectivement très différents de ceux de la protéine dans son état Apo, suggérant une autre conformation de la structure de la protéine. En repartant de zéro, contrairement au ligand précédent, les différentes techniques employées pour attribuer HSP90-NTD apo ont été de nouveau utilisées pour attribuer la protéine en complexe. Ce ligand permet, comme le précédent, de stabiliser la protéine et d'acquérir des spectres à plus haute température (308 K). Ainsi, 100 % des groupements méthyles et 92 % des résidus du squelette hors prolines de HSP90-NTD en complexe avec le 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide ont été attribués. Par ailleurs, avec ce ligand induisant une tout autre conformation de l'empreinte spectrale de la protéine toutes sauf une des résonances du squelette appartenant au segment invisible [L107-K116] pour la protéine Apo ont été attribuées avec ce second inhibiteur.

La note d'attribution qui suit, acceptée (sous presse) dans le journal « Biomolecular NMR Assignment », présente les résultats ainsi que le processus d'attribution employé pour HSP90-NTD liée au 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. Dans cet article j'ai produit les différents mutants et les échantillons marqués de HSP90-NTD. J'ai acquis les expériences RMN et analysé ces dernières. J'ai attribué les résonances du squelette et des groupements méthyles et j'ai également rédigé l'article pour lequel je suis première et co-correspondante auteure.

Article II : "NMR Assignment of Human HSP90 N-terminal domain bound to a long residence time resorcinol ligand"

NMR Assignment of Human HSP90 N-terminal domain bound to a long residence time resorcinol ligand

Faustine Henot^{1,*}, Elodie Crublet², Matthias Frech³, Jerome Boisbouvier^{1,*}

1. Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS),
71, avenue des martyrs, F-38044 Grenoble, France.

2. NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.

3. Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany.

* correspondence to be addressed to: faustine.henot@ibs.fr or jerome.boisbouvier@ibs.fr

Abstract

HSP90 is a major molecular chaperone that helps both folding and stabilization of various client proteins often implicated in growth control and cell survival such as kinases and transcription factors. However, among HSP90 clients are also found numerous oncoproteins and, through its assistance to them, HSP90 has consequently been reported as a promising anticancer target. Several ligand chemotypes, including resorcinol type ligands, were found to inhibit HSP90, most of them in an ATP competitive manner. Binding of some of these ligands modify significantly the NMR spectrum of the HSP90 ATP binding domain compared to the apo protein spectrum, hampering assignment transfer from the previously assigned human HSP90 apo state. Here we report the assignment of the $^1\text{H}_\text{N}$, ^{15}N , $^{13}\text{C}'$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^1\text{H}_\text{methyl}$, and $^{13}\text{C}_\text{methyl}$ chemical shifts of the 29 kDa HSP90 N-terminal domain bound to a long residence time resorcinol type inhibitor: 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. 92 % of the backbone resonances and 100 % of the [^1H , ^{13}C]-resonances of A^β , M^ϵ , T^γ , $\text{L}^{\delta 2}$, $\text{V}^{\gamma 2}$ and $\text{I}^{\delta 1}$ methyl groups were successfully assigned, including for the first time the assignment of the segment covering the nucleotide/drug binding site. Secondary structure predictions based on the NMR assignment reveal a structural rearrangement of HSP90 N-terminal domain upon ligand binding. The long residence time ligand induces the formation of a continuous helix covering the ligand binding site of HSP90 N-terminal domain accounting for the large differences observed in the NMR spectra between the apo and bound proteins.

Key words: HSP90, resorcinol ligand, NMR, methyl groups, backbone, assignment

Biological Context

Heat shock protein 90 (HSP90) is a highly conserved molecular chaperone that is found both in bacteria and eukaryotes (Farrelly and Finkelstein 1984; Bardwell and Craig 1987; Sorger and Pelham 1987;

Song et al. 1995; Csermely et al. 1998) but which is, however, absent in archaea (Pearl and Prodromou 2006). Several homologues of HSP90 exist in higher eukaryotes with distinct locations: mitochondrial TRAP1, cytosolic HSP90,

Grp94 from the endoplasmic reticulum and HSP90C of the chloroplast (Johnson 2012). Cytosolic HSP90, which presence is essential for eukaryotes' viability (Borkovich et al. 1989), is one of the most bountiful proteins in the cytosol of eukaryotic cells (Röhl et al. 2013). It is a homodimeric protein that can be found under two major isoforms: HSP90 α induced upon stress conditions and HSP90 β constitutively expressed (Millson et al. 2007) that are highly homologous (Johnson 2012). Each monomer of both isoforms is organised in three distinct domains: the N-terminal domain (NTD) that contains the ATP binding site (Prodromou and Roe 1997), the middle domain (MD), attached to the NTD via a charged linker, interacting with the γ phosphate of ATP thus regulating HSP90 ATPase function (Meyer et al. 2003) and the C-terminal domain (CTD) essential for HSP90 dimerization (Nemoto et al. 1995; Meng et al. 1996; Hoter et al. 2018).

HSP90 is a dynamic and plastic molecular chaperone involved in both protein folding and prevention of protein aggregation (Miyata and Yahara 1992; Daturpalli et al. 2013; Zuehlke et al. 2018; Hoter et al. 2018) that processes its clients via an ATPase regulated cycle and the assistance of various co-chaperones. More than 20 of these latter are known to be interacting with the molecular chaperone in a sequential manner (Röhl et al. 2013). They bind, depending on the co-chaperone involved, either the NTD, the MD or the CTD of HSP90 and modulate HSP90's function under different ways (Taipale et al. 2010; Trepel et al. 2010; Li et al. 2012). Several hundreds of HSP90

clients have been identified and can be sorted into three different categories: protein kinases, transcription factors and a third class of unrelated proteins (Picard : <http://www.picard.ch/downloads/Hsp90interactors.pdf>). However, until now no common sequence nor evident structural homology have been found between the abundant client proteins (Zuehlke and Johnson 2010; Li et al. 2012). Among HSP90 clients can be found numerous oncoproteins participating to the hallmarks of cancers such as for instance: Raf-1, AKT, Her2, Bcr-Abl, Plk, CDK4, CDK6, Myt1, p53, c-MET, Apaf-1, Telomerase, HIF-1 α and MMP-2 (Whitesell and Lindquist 2005; Hanahan and Weinberg 2011; Miyata et al. 2012; Garg et al. 2016; Vartholomaïou et al. 2016). HSP90 not only helps both folding and stabilization of oncoproteins contributing to cancer features (Zuehlke and Johnson 2010), it has been reported that a high level of HSP90 molecular chaperone was found in a large number of malignant cells, implying a paramount importance of HSP90 in survival and growth of cancer cells (Ciocca and Calderwood 2005; Calderwood et al. 2006). Such an observation was noticed particularly for breast cancers (Jameel et al. 1992) as well as in acute leukemias (Yufu et al. 1992; Chant et al. 1995), pancreatic carcinomas (Ogata et al. 2000) and melanomas (McCarthy et al. 2008).

Through its implication in various types of cancer, HSP90 has appeared as a decisive target for inhibitors' development to tackle this complex disease. Several HSP90 inhibitors have been developed, most of them targeting the ATP binding site located in the NTD, inhibiting the

chaperone's ATPase activity (Whitesell and Lindquist 2005). Such inhibitors, blocking HSP90's functional cycle, were, thus, impeding the chaperone to assist its oncoprotein clients. A large number of HSP90 inhibitors, first natural compounds then followed by synthetic inhibitors, have been identified by pharmaceutical industries. Since 1998, 29 inhibitors binding to the NTD of HSP90 have entered clinical trials and, to date, 11 are still in trials. However, among these inhibitors, even though they are all different chemically, none has, as of yet, been approved (Yuno et al. 2018; Zuehlke et al. 2018; Wang et al. 2021).

In this study, we focused on the 29 kDa HSP90 α -NTD, previously assigned in its apo state (Jacobs et al. 2006; Park et al. 2011; Henot et al. 2021), bound to a long residence time resorcinol type inhibitor: 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4]triazol-3-yl] - N-furan - 2-ylmethyl-2,4-dihydroxy-N-methyl-benzamide. This ligand, when bound to the protein, induces conformational changes, notably on the segment covering the ATP/drug binding

site as shown by X-ray crystallography (Amaral et al. 2017), that correspond neither to the previously described open nor to the closed ATP-lid states sampled in solution by apo HSP90-NTD (Henot et al. Submitted). Such a new conformation has never been assigned by NMR. Spectra comparing the HSP90-NTD apo and bound to the long residence time resorcinol type inhibitor are highly different (Fig. 1) impeding a simple assignment transfer from the already assigned apo HSP90-NTD to the one in complex with the ligand. Additionally, a slow ligand exchange ($\tau_R \sim 50$ min) (Amaral et al. 2017) between its free and bound forms prevents both the use of EXSY experiments and titrations to transfer the assignment starting from the apo protein advocating for a new assignment of this particularly challenging and dynamic target.

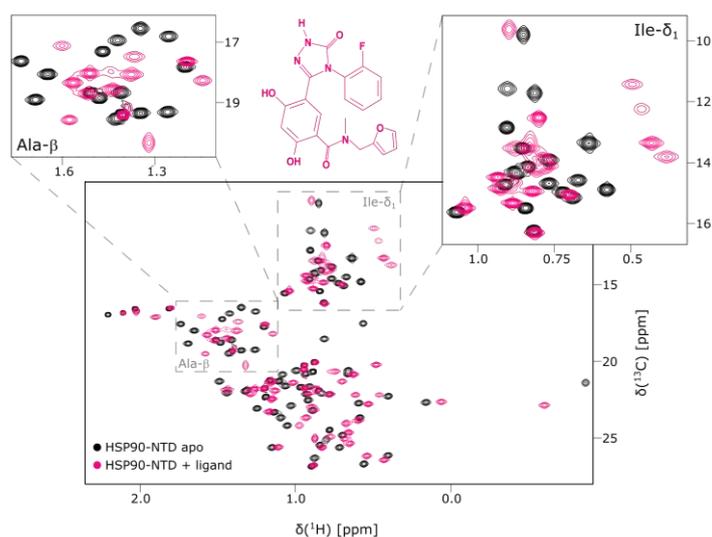


Figure 1: Spectra comparison between HSP90-NTD apo and in complex. Superimposition of ^1H - ^{13}C SOFAST methyl TROSY spectra of HSP90-NTD apo (black) and HSP90-NTD bound to the resorcinol type ligand 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2-ylmethyl-2,4-dihydroxy-N-methyl-benzamide (pink), both acquired at 293 K and at a ^1H frequency of 700 MHz. Zooms focusing on alanine and isoleucine regions are displayed.

Here, we present the resonance assignment of the $^1\text{H}_\text{N}$, ^{15}N , $^{13}\text{C}'$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^1\text{H}_\text{methyl}$, and $^{13}\text{C}_\text{methyl}$ chemical shifts of the 29 kDa HSP90-NTD bound to the inhibitor: 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4]triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. This assignment is a prerequisite to better understand HSP90-NTD conformational dynamics' role in binding inhibitors.

Methods and experiments

Protein expression and purification

E. coli BL21-DE3-RIL cells transformed with a pET-28 plasmid encoding the N-Terminal domain of HSP90 α from *Homo Sapiens* (HSP90-NTD) with a His-Tag and a TEV cleavage site were progressively adapted in three stages over 24 h to M9/ $^2\text{H}_2\text{O}$ (M9 in H_2O , M9 in 50% H_2O + 50% $^2\text{H}_2\text{O}$, M9 in $^2\text{H}_2\text{O}$). In the final culture, bacteria were grown at 37°C in M9 medium with 99.85 % $^2\text{H}_2\text{O}$ (Eurisotop). 2 g/L of either D-glucose-d7 (for U- ^2H , ^{15}N , ^{12}C) samples) or D-glucose- $^{13}\text{C}_6$ -d7 (for U- ^2H , ^{15}N , ^{13}C) samples) (CIL) and 1 g/L of $^{15}\text{N}^1\text{H}_4\text{Cl}$ (Merck) were used as carbon and nitrogen sources, respectively. In order to label methyl-containing amino-acids, isotopically labelled precursors or amino-acids were added to the media at specific time-intervals before induction.

- (Sample 1) Production of U- ^2H , ^{15}N , ^{13}C HSP90-NTD was undertaken.

- (Sample 2) For production of U- ^2H , ^{15}N , ^{13}C , Ile-[2,3,4,4- $^2\text{H}_4$; 1,2,3,4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\delta 1}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\gamma 2}$], Leu-[2,3,3,4- $^2\text{H}_4$; 1,2,3,4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$], Val-[2,3- $^2\text{H}_2$; 1,2,3- $^{13}\text{C}_3$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-}}$

$^{\text{S}}$ / [$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$] HSP90-NTD, a solution containing 172 mg/L of sodium 1, 2, 3- [$^{13}\text{C}_3$]-2- [$^{13}\text{C}^1\text{H}_3$]-2- [O ^2H]-3-oxo-4, 4, 4- [$^2\text{H}_3$]-butanoate precursor (Henot et al. 2021) was added to the culture medium when the O.D at 600 nm reached 0.6. 40 min later (20 min before induction) a solution containing 60 mg/L of sodium (S)-2-hydroxy-2-(1',1'- [$^2\text{H}_2$], 1', 2'- [$^{13}\text{C}_2$]) ethyl-3-oxo-1,2,3- [$^{13}\text{C}_3$]-4,4,4- [$^2\text{H}_3$] butanoate was added to the medium.

- (Sample 3) For production of U- ^2H , ^{15}N , ^{12}C], Ala- [$^{13}\text{C}^1\text{H}_3$] $^\beta$, Met- [$^{13}\text{C}^1\text{H}_3$] $^\epsilon$, Leu/Val- [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$, Ile- [$^{13}\text{C}^1\text{H}_3$] $^{\delta 1}$, Thr- [$^{13}\text{C}^1\text{H}_3$] $^\gamma$ HSP90-NTD, a HLAM-A $^\beta$ I $^{\delta 1}$ M $^\epsilon$ (LV) $^{\text{proS}}$ T $^\gamma$ kit, purchased from NMR-Bio, was added before induction according to the manufacturer's protocol.

- (Samples 4) U- ^2H , ^{15}N , ^{12}C] HSP90-NTD samples labelled on a single type of methyl probes were produced in small scales (21 mL) to identify A $^\beta$, M $^\epsilon$, T $^\gamma$ methyl type (3 samples) or to complete assignment using single point amino acids mutants (Amero et al. 2011) (29 samples: A21S, A24S, A27S, A55S, A101S, A111S, A117S, A121S, A124S, A126S, A161S, A166S, A235S, I34V, I128V, M12L, M30L, M119L, M130L, T19S, T36S, T65S, T99S, T109S, T115S, T171S, T174S, T195S, T219S). Single point amino acid mutations were generated by GeneCust. For each of these samples, a single type of methyl groups was labelled by addition of the corresponding NMR-Bio kit (SLAM-A $^\beta$, SLAM-M $^\epsilon$, SLAM-I $^{\delta 1}$, SLAM-T $^\gamma$) in M9/ $^2\text{H}_2\text{O}$ media when the O.D. at 600 nm reached 0.6, 1 h before induction.

Protein production was induced by addition of IPTG to a final concentration of 0.5 mM and protein expression was

performed at 20°C overnight. The purification protocol of HSP90-NTD is described in an earlier publication (Henot et al. 2021). The final NMR buffer conditions were 20 mM HEPES, 150 mM NaCl and 1 mM TCEP, pH 7.5 in either 100% $^2\text{H}_2\text{O}$ or 90% $\text{H}_2\text{O}/10\%$ $^2\text{H}_2\text{O}$. The resorcinol inhibitor: 5-[4-(2-Fluorophenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide was added in each sample in a 2.5 fold excess using a concentrated stock solution (provided by Merck KGaA). 200 μL of the final HSP90-NTD samples with the ligand were loaded in 4 mm shigemi tubes at concentrations of *c.a* 0.5 mM (labelling schemes 1, 3) and *c.a* 0.25 mM (labelling scheme 2). The wild type and single point mutants of HSP90-NTD proteins labelled on only one methyl type (labelling scheme 4), to which the ligand was added, were concentrated to [0.1-0.4] mM and 40 μL of each sample was loaded in a 1.7 mm NMR tube.

NMR Spectroscopy

All NMR experiments acquired on HSP90-NTD samples were recorded at 308 K, unless specified otherwise, with a ligand excess. Backbone resonances assignment was carried out by using both a $^1\text{H}-^{15}\text{N}$ TROSY experiment and full sets of BEST-TROSY 3D triple resonance experiments (HNCA, HN(CA)CB, HNCO, HN(CA)CO, HN(CO)CA and HN(COCA)CB (Favier and Brutscher 2019). All these latter experiments were acquired on a Bruker Avance III HD spectrometer equipped with a cryogenic probe and operating at a ^1H frequency of 600 MHz for a total duration of 15 days

using a 0.5 mM sample of U- $[\text{}^2\text{H}, \text{}^{15}\text{N}, \text{}^{13}\text{C}]$ HSP90-NTD.

Ile, Leu pro-S and Val pro-S methyl group assignment was undertaken using 3D “out and back” HCC, HC(C)C and HC(CC)C experiments (Tugarinov and Kay 2003; Ayala et al. 2009, 2012; Mas et al. 2013) acquired on a spectrometer operating at a ^1H frequency of 600 MHz for a total duration of 4 days, using a 0.25 mM sample of U- $[\text{}^2\text{H}, \text{}^{15}\text{N}, \text{}^{13}\text{C}]$, Ile- $[\text{}^2,3,4,4\text{-}^2\text{H}_4; \text{}^1,2,3,4\text{-}^{13}\text{C}_4; [\text{}^{13}\text{C}^1\text{H}_3]^{\delta 1}/[\text{}^{12}\text{C}^2\text{H}_3]^{\gamma 2}]$, Leu- $[\text{}^2,3,3,4\text{-}^2\text{H}_4; \text{}^1,2,3,4\text{-}^{13}\text{C}_4; [\text{}^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}/[\text{}^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}]$, Val- $[\text{}^2,3\text{-}^2\text{H}_2; \text{}^1,2,3\text{-}^{13}\text{C}_3; [\text{}^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}/[\text{}^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}]$ HSP90-NTD (sample 2).

2D $^1\text{H}-^{13}\text{C}$ SOFAST methyl TROSY (Amero et al. 2009) experiments, to identify the remaining methyl type (alanine, methionine, threonine), were recorded.

A 3D CCH HMQC-NOESY-HMQC NMR experiment (Tugarinov et al. 2005; Törner et al. 2020) was acquired over 2 days on a spectrometer operating at a ^1H frequency of 950 MHz using a 0.5 mM sample of U- $[\text{}^2\text{H}, \text{}^{15}\text{N}, \text{}^{12}\text{C}]$, Ala- $[\text{}^{13}\text{C}^1\text{H}_3]^{\beta}$, Met- $[\text{}^{13}\text{C}^1\text{H}_3]^{\epsilon}$, Leu/Val- $[\text{}^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}$, Ile- $[\text{}^{13}\text{C}^1\text{H}_3]^{\delta 1}$, Thr- $[\text{}^{13}\text{C}^1\text{H}_3]^{\gamma}$ HSP90-NTD (sample 3). The NOE mixing time was set to 750 ms to detect a maximum of long-range NOEs. Remaining methyl groups' assignment was performed using the automated methyl assignment software MAGIC (Monneau et al. 2017) and previously assigned isoleucine, leucine and valine methyl moieties together with the structure of N-HSP90 bound to the long residence time resorcinol inhibitor (PDB: 5J20).

These assignments were completed using 2D $^1\text{H}-^{13}\text{C}$ SOFAST methyl TROSY experiments on 29 single point

amino acids mutants recorded for an average duration of ~ 1.5 h each, on a spectrometer operating at a ^1H frequency of 850 MHz and equipped with a 1.7 mm cryogenically cooled, pulsed-field-gradient triple-resonance probe.

NMR data were processed using nmrPipe/nmrDraw (Delaglio et al. 1995) and analyzed using CcpNMR (Vranken et al. 2005)

Extent of assignments and data deposition

Backbone sequential assignment of HSP90-NTD bound to a long residence time resorcinol type ligand, first undertaken at 298 K, presented some challenges as 35 % of the non-proline residues were not detected in the 3D experiments. Dynamics in the μs -ms regime broadened these backbone resonances beyond detection (data not shown). In order to overcome this hurdle and to minimize conformational exchange's contribution to the transverse relaxation, both an increase of the temperature and a decrease of the magnetic field at which the spectra were recorded were used to increase the signal to noise ratio. As a result, 92 % of $^1\text{H}_\text{N}$, 92 % of $^{15}\text{N}_\text{H}$, 94 % of ^{13}CO , 96 % of $^{13}\text{C}_\alpha$ and 95 % of $^{13}\text{C}_\beta$ for the non-proline residues were detected after these new acquisition conditions and assigned (Fig. 2a, b, c) using 6 'BEST-TROSY' triple resonance experiments (Favier and Brutscher 2019). Stabilizing the protein with this resorcinol type ligand together with the optimization of NMR data acquisition conditions enabled us to assign for the first time almost all the backbone

resonances of the segment [L107-K116] covering the ligand binding site and undetectable in HSP90-NTD apo state (Jacobs et al. 2006; Park et al. 2011; Henot et al. 2021).

'Out and back' HCCs experiments, acquired on a U- $[\text{}^2\text{H}, \text{}^{15}\text{N}, \text{}^{13}\text{C}]$, Ile- $[\text{}^2,3,4,4\text{-}^2\text{H}_4; \text{}^1,2,3,4\text{-}^{13}\text{C}_4; [\text{}^{13}\text{C}^1\text{H}_3]^\delta]$, Leu- $[\text{}^2,3,3,4\text{-}^2\text{H}_4; \text{}^1,2,3,4\text{-}^{13}\text{C}_4; [\text{}^{13}\text{C}^1\text{H}_3]^{pro-S}/[\text{}^{12}\text{C}^2\text{H}_3]^{pro-R}]$, Val- $[\text{}^2,3\text{-}^2\text{H}_2; \text{}^1,2,3\text{-}^{13}\text{C}_3; [\text{}^{13}\text{C}^1\text{H}_3]^{pro-S}/[\text{}^{12}\text{C}^2\text{H}_3]^{pro-R}]$ labelled HSP90-NTD sample with a ligand excess, enabled to connect previously assigned backbone resonances to isoleucine- δ_1 , leucine- δ_2 and valine- γ_2 methyl groups. 75 % of the isoleucine- δ_1 , 100 % of the leucine- δ_2 and 73 % of the valine- γ_2 methyl groups could be unambiguously connected to the previously assigned C_α and C_β resonances (Fig. 3a). These first assigned methyl groups, in addition to the amino acid type corresponding to each methyl signal, the known 3D structure (PDB: 5J20) and the 369 intermethyl NOE connectivities recorded on the U- $[\text{}^2\text{H}, \text{}^{15}\text{N}, \text{}^{12}\text{C}]$, Ala- $[\text{}^{13}\text{C}^1\text{H}_3]^\beta$, Met- $[\text{}^{13}\text{C}^1\text{H}_3]^\epsilon$, Leu/Val- $[\text{}^{13}\text{C}^1\text{H}_3]^{pro-S}$, Ile- $[\text{}^{13}\text{C}^1\text{H}_3]^\delta$, Thr- $[\text{}^{13}\text{C}^1\text{H}_3]^\gamma$ HSP90-NTD sample, were given as input data for the automated methyl assigned performed by the MAGIC software (Monneau et al. 2017). 12 methyl groups (2 alanines (A141 and A145), 2 isoleucines (I203 and I131), 2 methionines (M98 and M180), 2 valines (V17 and V186) and 4 threonines (T90, T94, T152 and T184)) were unambiguously assigned with both a high percentage of NOE correlation peaks assigned and a high confidence score (Fig. 3b, c). The only unassigned valine methyl signal was unambiguously assigned as the last remaining valine residue (V230).

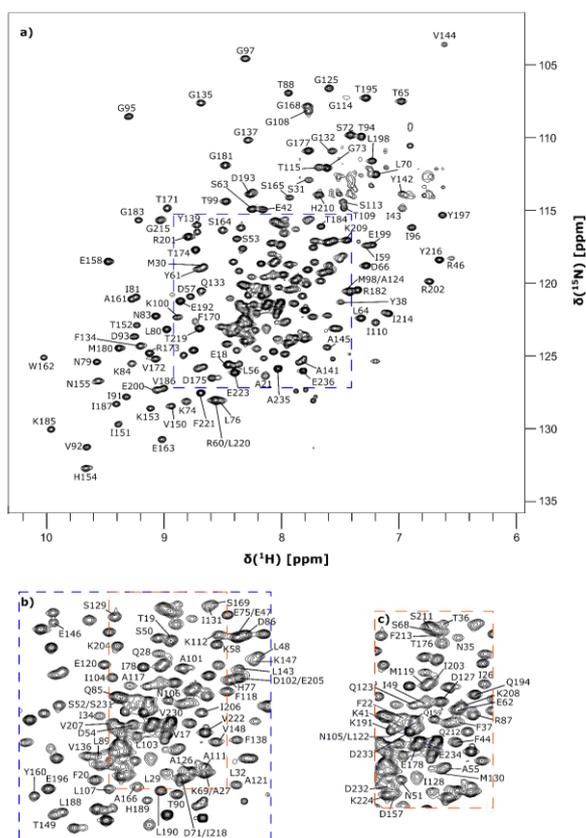
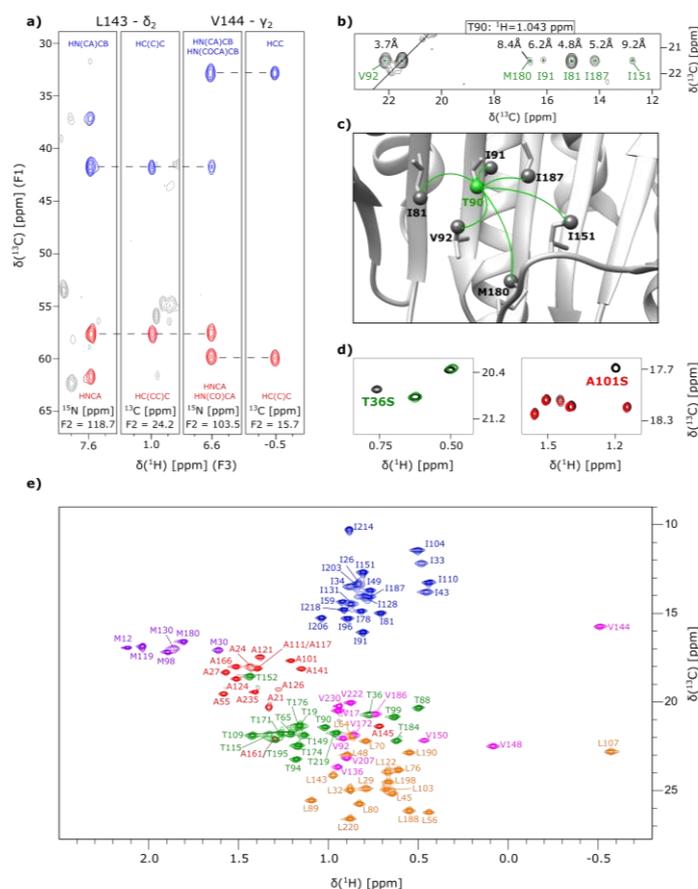


Figure 2: Backbone resonances assignment of HSP90-NTD bound to a long residence time resorcinol type ligand. a ^{15}N BEST-TROSY spectrum of HSP90-NTD bound to a resorcinol type ligand acquired at 308K and at a ^1H frequency of 600 MHz. On panels **b** and **c**, zooms from panel **a** are presented. Each assigned signal (**a**, **b**, **c**) is annotated with the corresponding residue number.

However, even though both leucine- δ_2 and valine- γ_2 methyl groups were fully assigned, 38 % of the methyl groups of HSP90-NTD bound to a long residence time resorcinol inhibitor remained unassigned (33 methyl groups). To complete the assignment, 29 single point amino acid mutants were over-expressed and purified in small scale cultures. Replacement amino acids which were structurally similar to the replaced amino acids were chosen to minimize secondary chemical shifts (Crublet et al. 2014). This mutant library allowed the unambiguous assignment of 22 methyl groups (Fig. 3d). Seven spectra were more challenging to analyze due to secondary chemical shifts. Eleven methyl groups remained unassigned, four for which mutants were not available and seven with complex mutant spectra to analyze due to chemical

shifts perturbations resulting from the introduced mutation. Among these 11 unassigned methyl groups were: three threonines (T88, T176 and T149), one isoleucine (I26) and seven alanines (A24, A27, A111, A117, A121, A124 and A126). By analyzing manually the NOE cross peaks' network, the three threonines and the isoleucine (partially superimposed with I203) could be assigned. Regarding the remaining seven alanines, their assignment was performed using a combination of the mutant spectra displaying secondary chemical shift perturbations, backbone C_β resonances, and the NOE cross peaks' network. Finally, 100 % of the methyl resonances of N-HSP90 bound to a long residence time resorcinol inhibitor could be assigned (Fig. 3e).

Figure 3: Methyl groups assignment of HSP90-NTD bound to a long residence time resorcinol type ligand. Panel a displays 2D-extracts from 3D 'out-and back' HCC, HC(C)C and HC(CC)C, HNCA, HN(CO)CA, HN(COCA)CB and HN(CA)CB experiments. In red C_{α} resonances and in blue C_{β} resonances. 3D spectra were recorded on an NMR spectrometer operating at a proton frequency of 600 MHz and using samples 1 and 2 to transfer the assignment from backbone resonances of L143 and V144 to their respective methyl groups. Panel b presents an extract of a 3D HMQC-NOESY-HMQC experiment recorded on a spectrometer operating at a ^1H frequency of 950 MHz. and using a U- ^2H , ^{12}C , ^{15}N -Leu/Val- $^{13}\text{C}^1\text{H}_3$ ^{pro-S}, Ile- $^{13}\text{C}^1\text{H}_3$ ^{δ^1} , Met- $^{13}\text{C}^1\text{H}_3$ ^{ϵ} , Ala- $^{13}\text{C}^1\text{H}_3$ ^{β} , Thr- $^{13}\text{C}^1\text{H}_3$ ^{γ} HSP90-NTD sample bound to a resorcinol type ligand. The plane was extracted at the methyl proton frequency of T90. Panel c depicts a zoom around T90- γ



(green sphere) of the 3D structure of HSP90-NTD bound to a resorcinol type ligand (PDB: 5J20). Neighboring methyl groups of T90- γ for which ^1H - ^1H NOEs are detected in the 2D extract presented in panels b are displayed with grey spheres and connected to T90- γ by green lines. **d** Examples of assignments performed by mutagenesis to complete the HCC/NOE based assignment. Each mutant 2D spectrum extract (T36S green, A101S red) is superimposed with the wild type protein 2D spectrum extract (black) for comparison. Panel **e** displays the fully assigned 2D ^1H - ^{13}C SOFAST methyl TROSY spectrum of HSP90-NTD bound to a resorcinol ligand acquired at 308 K. Each signal is annotated with the corresponding residue number. Alanines, isoleucines, valines, leucines, threonines and methionines are depicted in red, dark blue, pink, orange, light green and purple, respectively.

Secondary Structure Predictions

Assignment of the segment covering the drug binding site is achieved here for the first time on HSP90-NTD bound to a long residence time resorcinol type ligand, a flexible protein that has been shown to sample different conformations distant by more than 30 Å in its apo state (Henot et al. Submitted). Secondary structure predictions of HSP90-NTD in complex

with the long residence time resorcinol ligand, based on backbone chemical assignment using TALOS+ (Shen et al. 2009) displayed secondary structure elements in accordance with the crystal structure PDB: 5J20 (Amaral et al. 2017), notably when focusing on the segment [100-123] that corresponds to a long helical secondary structure (Fig. 4a, b, c, d, e).

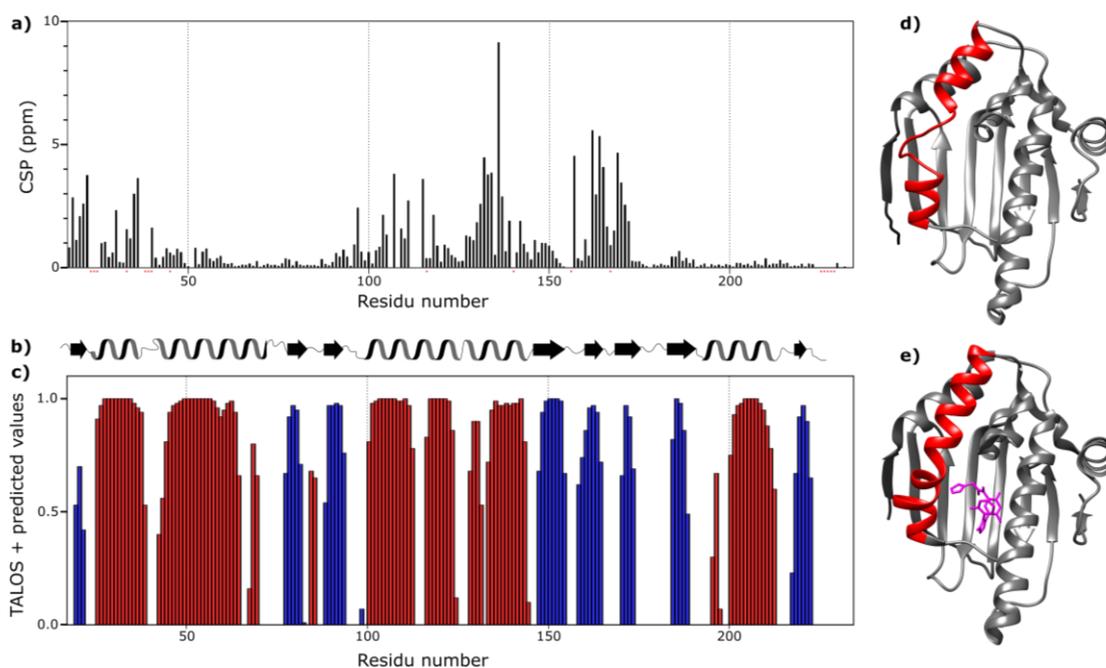


Figure 4: Secondary structure predictions – comparison to crystallographic data. Panel **a** shows the chemical shift perturbations (calculated using Ca, Cb and C' chemical shifts with the formula: $((\Delta\delta_{C\alpha})^2 + (\Delta\delta_{C\beta})^2 + (\Delta\delta_{C'})^2)^{\frac{1}{2}}$, with Δ being the difference between the chemical shifts of the apo protein and the ones of the protein in complex) induced by the binding of the ligand 5-[4-(2-Fluorophenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide (BMRB codes: 50786 and 51378) Residues for which the amide groups were not assigned for the protein in complex are represented by the symbol *. Panel **b** displays the secondary structure elements as reported on the crystallographic structure PDB: 5J20 (**c**). Panel **c** presents the secondary structure predictions based on backbone assignment of HSP90-NTD bound to the long residence time ligand using TALOS+. In red and in blue are predicted α -helix and β -sheets secondary structure elements, respectively. Panel **d** displays the structure of apo HSP90-NTD (PDB: 1YES) with the segment [100-123], constituted of two helices and a loop, colored in red. Panel **e** shows the structure of HSP90-NTD bound to the long residence time ligand (PDB: 5J20) with the helix [100-123] colored in red and the ligand 5-[4-(2-Fluorophenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide in magenta.

Such results highlight that a structural rearrangement is induced by the binding of the long residence time ligand. Both this latter structural rearrangement and the presence of the ligand account for the large chemical shift differences observed in the spectra.

The assignment of chemical shifts of backbone and methyl groups of HSP90-NTD, together with all the NMR data sets used have been deposited into the BMRB (<https://www.bmrwisc.edu/>) with the accession number 51378.

Acknowledgments

The authors thank Drs. P. Macek, P. Gans, R. Awad and R. Kerfah for advice and stimulating discussions. This work used the high field NMR and isotopic labelling facilities at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble

Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by grants from CEA/NMR-Bio (research program C24990), by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) and by a fellowship (to F.H) from "La Ligue contre le Cancer".

References

- Amaral M, Kokh DB, Bomke J, et al (2017) Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat Commun* 8–2276. <https://doi.org/10.1038/s41467-017-02258-w>
- Amero C, Schanda P, Asunción Durá M, et al (2009) Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J Am Chem Soc* 131:3448–3449. <https://doi.org/10.1021/ja809880p>
- Ayala I, Hamelin O, Amero C, et al (2012) An optimized isotopic labelling strategy of isoleucine- γ 2 methyl groups for solution NMR studies of high molecular weight proteins. *Chem Commun* 48:1434–1436. <https://doi.org/10.1039/c1cc12932e>
- Ayala I, Sounier R, Usé N, et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* 43:111–119. <https://doi.org/10.1007/s10858-008-9294-7>
- Bardwell JC, Craig EA (1987) Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc Natl Acad Sci U S A* 84:5177–5181. <https://doi.org/10.1073/pnas.84.15.5177>
- Borkovich KA, Farrelly FW, Finkelstein DB, et al (1989) hsp82 Is an Essential Protein That Is Required in Higher Concentrations for Growth of Cells at Higher Temperatures. *Mol Cell Biol* 9:3919–3930. <https://doi.org/10.1128/mcb.9.9.3919>
- Calderwood SK, Khaleque MA, Sawyer DB, Ciocca DR (2006) Heat shock proteins in cancer: Chaperones of tumorigenesis. *Trends Biochem Sci* 31:164–172. <https://doi.org/10.1016/j.tibs.2006.01.006>
- Chant ID, Rose PE, Morris AG (1995) Analysis of heat-shock protein expression in myeloid leukaemia cells by flow cytometry. *Br J of Haematology* 90:163–168. <https://doi.org/https://doi.org/10.1111/j.1365-2141.1995.tb03395.x>
- Ciocca DR, Calderwood SK (2005) Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress Chaperones* 10:86–103. <https://doi.org/10.1379/csc-99r.1>
- Crublet E, Kerfah R, Mas G, et al (2014) A cost-effective protocol for the parallel production of libraries of CH₃-specifically labeled mutants for NMR studies of high molecular weight proteins. *Methods Mol Biol* 1091:229–243. https://doi.org/10.1007/978-1-62703-691-7_17
- Csermely P, Schnaider T, Soti C, et al (1998) The 90-kDa Molecular Chaperone Family: Structure, Function, and Clinical Applications. *A Comprehensive Review. Pharmacol Ther* 79:129–168. <https://doi.org/10.1016/s0163->

- 7258(98)00013-8
- Daturpalli S, Waudby CA, Meehan S, Jackson SE (2013) Hsp90 inhibits α -synuclein aggregation by interacting with soluble oligomers. *J Mol Biol* 425:4614–4628. <https://doi.org/10.1016/j.jmb.2013.08.006>
- Delaglio F, Grzesiek S, Vuister GW, et al (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293. <https://doi.org/10.1007/BF00197809>
- Farrelly FW, Finkelstein DB (1984) Complete sequence of the heat shock-inducible HSP90 gene of *Saccharomyces cerevisiae*. *J Biol Chem* 259:5745–5751. [https://doi.org/10.1016/s0021-9258\(18\)91077-x](https://doi.org/10.1016/s0021-9258(18)91077-x)
- Favier A, Brutscher B (2019) NMRLib: user-friendly pulse sequence tools for Bruker NMR spectrometers. *J Biomol NMR* 73:199–211. <https://doi.org/10.1007/s10858-019-00249-1>
- Garg G, Khandelwal A, Blagg BSJ (2016) Anticancer Inhibitors of Hsp90 Function: Beyond the Usual Suspects. *Adv Cancer Res* 129:51–88. <https://doi.org/10.1016/bs.acr.2015.12.001>
- Hanahan D, Weinberg RA (2011) Hallmarks of cancer: The next generation. *Cell* 144:646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Henot F, Kerfah R, Törner R, et al (2021) Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain. *J Biomol NMR* 75:221–232. <https://doi.org/10.1007/s10858-021-00370-0>
- Henot F, Rioual E, Favier A, et al Visualizing the Transiently Populated Closed-State of Human HSP90 ATP Binding Domain. Submitted. <https://doi.org/10.1101/2022.03.16.484593>
- Hoter A, El-Sabban ME, Naim HY (2018) The HSP90 family: Structure, regulation, function, and implications in health and disease. *Int J Mol Sci* 19:. <https://doi.org/10.3390/ijms19092560>
- Jacobs DM, Langer T, Elshorst B, et al (2006) NMR Backbone Assignment of the N-terminal Domain of Human HSP90. *J Biomol NMR* 36:52. <https://doi.org/10.1007/s10858-006-9030-0>
- Jameel A, Skilton RA, Campbell TA, et al (1992) Clinical and biological significance of HSP89 alpha in human breast cancer. *Int J Cancer* 50:409–415. <https://doi.org/10.1002/ijc.2910500315>
- Johnson JL (2012) Evolution and function of diverse Hsp90 homologs and cochaperone proteins. *Biochim Biophys Acta - Mol Cell Res* 1823:607–613. <https://doi.org/10.1016/j.bbamcr.2011.09.020>
- Li J, Soroka J, Buchner J (2012) The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. *Biochim Biophys Acta - Mol Cell Res* 1823:624–635. <https://doi.org/10.1016/j.bbamcr.2011.09.003>
- Mas G, Crublet E, Hamelin O, et al (2013) Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. *J Biomol NMR* 57:251–262. <https://doi.org/10.1007/s10858-013-9785-z>
- McCarthy MM, Pick E, Kluger Y, et al (2008) HSP90 as a marker of progression in melanoma. *Ann Oncol* 19:590–594. <https://doi.org/10.1093/annonc/mdm545>
- Meng X, Devin J, Sullivan WP, et al (1996) Mutational analysis of Hsp90 α

- dimerization and subcellular localization: dimer disruption does not impede 'in vivo' interaction with estrogen receptor. *J Cell Sci* 109:1677–1687
- Meyer P, Prodromou C, Hu B, et al (2003) Structural and Functional Analysis of the Middle Segment of Hsp90: Implications for ATP Hydrolysis and Client Protein and Cochaperone Interactions. *Mol Cell* 11:647–658. [https://doi.org/10.1016/s1097-2765\(03\)00065-0](https://doi.org/10.1016/s1097-2765(03)00065-0)
- Millson SH, Truman AW, Rácz A, et al (2007) Expressed as the sole Hsp90 of yeast, the α and β isoforms of human Hsp90 differ with regard to their capacities for activation of certain client proteins, whereas only Hsp90 β generates sensitivity to the Hsp90 inhibitor radicicol. *FEBS J* 274:4453–4463. <https://doi.org/10.1111/j.1742-4658.2007.05974.x>
- Miyata Y, Nakamoto H, Neckers L (2012) The Therapeutic Target Hsp90 and Cancer Hallmarks. *Curr Pharm Des* 19:347–365. <https://doi.org/10.2174/138161213804143725>
- Miyata Y, Yahara I (1992) The 90-kDa heat shock protein, HSP90, binds and protects casein kinase II from self-aggregation and enhances its kinase activity. *J Biol Chem* 267:7042–7047. [https://doi.org/10.1016/s0021-9258\(19\)50533-6](https://doi.org/10.1016/s0021-9258(19)50533-6)
- Monneau YR, Rossi P, Bhaumik A, et al (2017) Automatic methyl assignment in large proteins by the MAGIC algorithm. *J Biomol NMR* 69:215–227. <https://doi.org/10.1007/s10858-017-0149-y>
- Nemoto T, Ohara-Nemoto Y, Ota M, et al (1995) Mechanism of dimer formation of the 90-kDa heat-shock protein. *Eur J Biochem* 233:1–8. https://doi.org/10.1111/j.1432-1033.1995.001_1.x
- Ogata M, Naito Z, Tanaka S, et al (2000) Overexpression and Localization of Heat Shock Proteins mRNA in Pancreatic Carcinoma. *J Nippon Med Sch* 67:177–185. <https://doi.org/10.1272/jnms.67.177>
- Park SJ, Kostic M, Dyson HJ (2011) Dynamic interaction of Hsp90 with its client protein p53. *J Mol Biol* 411:158–173. <https://doi.org/10.1016/j.jmb.2011.05.030>
- Pearl LH, Prodromou C (2006) Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu Rev Biochem* 75:271–294. <https://doi.org/10.1146/annurev.biochem.75.103004.142738>
- Picard HSP90 INTERACTORS Chaperones and relatives. http://www.picard.ch/downloads/Hsp90_interactors.pdf. Accessed 13 Oct 2021
- Prodromou C, Roe SM (1997) Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone. *Cell* 90:65–75. [https://doi.org/10.1016/s0092-8674\(00\)80314-1](https://doi.org/10.1016/s0092-8674(00)80314-1)
- Röhl A, Rohrberg J, Buchner J (2013) The chaperone Hsp90: Changing partners for demanding clients. *Trends Biochem Sci* 38:253–262. <https://doi.org/10.1016/j.tibs.2013.02.003>
- Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223. <https://doi.org/10.1007/s10858-009-9333-z>
- Song HY, Dunbar JD, Yuan Xin Zhang, et al (1995) Identification of a protein with homology to hsp90 that binds the type 1 tumor necrosis factor receptor. *J Biol Chem* 270:3574–3581. <https://doi.org/10.1074/jbc.270.8.3574>

- Sorger PK, Pelham HR (1987) The Glucose-regulated Protein grp94 is Related to Heat Shock Protein hsp90. *J Mol Biol* 194:341–344.
[https://doi.org/10.1016/0022-2836\(87\)90380-9](https://doi.org/10.1016/0022-2836(87)90380-9)
- Taipale M, Jarosz DF, Lindquist S (2010) HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat Rev Mol Cell Biol* 11:515–528.
<https://doi.org/https://doi.org/10.1038/nrm2918>
- Törner R, Awad R, Gans P, et al (2020) Spectral editing of intra- and inter-chain methyl–methyl NOEs in protein complexes. *J Biomol NMR* 74:83–94.
<https://doi.org/10.1007/s10858-019-00293-x>
- Trepel J, Mollapour M, Giaccone G, Neckers L (2010) Targeting the dynamic HSP90 complex in cancer. *Nat Rev Cancer* 10:537–549.
<https://doi.org/10.1038/nrc2887>
- Tugarinov V, Kay LE (2003) Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J Am Chem Soc* 125:13868–13878.
<https://doi.org/10.1021/ja030345s>
- Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional ¹H-¹³C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J Am Chem Soc* 127:2767–2775.
<https://doi.org/10.1021/ja044032o>
- Vartholomaiou E, Echeverría P, Picard D (2016) Unusual Suspects in the Twilight Zone Between the Hsp90 Interactome and Carcinogenesis. *Adv Cancer Res* 129:1–30.
<https://doi.org/10.1016/bs.acr.2015.08.001>
- Vranken WF, Boucher W, Stevens TJ, et al (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct Funct Genet* 59:687–696.
<https://doi.org/10.1002/prot.20449>
- Wang L, Zhang Q, You Q (2021) Targeting the HSP90–CDC37–kinase chaperone cycle: A promising therapeutic strategy for cancer. *Med Res Rev* 1–27.
<https://doi.org/10.1002/med.21807>
- Whitesell L, Lindquist SL (2005) HSP90 and the chaperoning of cancer. *Nat Rev Cancer* 5:761–772.
<https://doi.org/10.1038/nrc1716>
- Yufu Y, Nishimura J, Nawata H (1992) High constitutive expression of heat shock protein 90 α in human acute leukemia cells. *Leuk Res* 16:597–605.
[https://doi.org/https://doi.org/10.1016/0145-2126\(92\)90008-U](https://doi.org/https://doi.org/10.1016/0145-2126(92)90008-U)
- Yuno A, Lee MJ, Lee S, et al (2018) Clinical evaluation and biomarker profiling of Hsp90 inhibitors. *Methods Mol Biol* 1709:423–441.
https://doi.org/10.1007/978-1-4939-7477-1_29
- Zuehlke A, Johnson JL (2010) Hsp90 and co-chaperones twist the functions of diverse client proteins. *Biopolymers* 93:211–217. <https://doi.org/10.1002/bip.21292>
- Zuehlke AD, Moses MA, Neckers L (2018) Heat shock protein 90: Its inhibition and function. *Philos Trans R Soc L B Biol Sci* 373:.
<https://doi.org/10.1098/rstb.2016.0527>

Chapitre 3

Etude structurale et dynamique du domaine N-terminal de HSP90 Apo

Ce chapitre de ma thèse est un travail interdisciplinaire portant sur l'étude à la fois de la structure et de la dynamique de HSP90-NTD et donnant lieu à un article présenté à la fin de ce chapitre. J'ai été impliquée dans la partie structurale où j'ai produit les différents échantillons marqués. J'ai également co-supervisé Elisa Rioual une stagiaire de Master 2, participé à l'acquisition et l'analyse des expériences NOESY, analysé les structures de la Protein Data Bank, conçu les mutants et produit ces derniers. La partie dynamique quant à elle représente le sujet principal de ma thèse. Tous les résultats présentés proviennent d'un travail personnel. J'ai choisi d'expliquer ici l'essentiel des résultats structuraux (même si pour cela des données déjà incluses dans l'article ont dû être réintroduites, toutefois, en une version plus compacte afin de limiter au maximum les répétitions) et d'approfondir la partie dynamique – une partie qui ne transparait pas dans le format plutôt réduit de la publication.

3.1 Variabilité structurale de HSP90-NTD

3.1.1 Rôle et cycle fonctionnel d'HSP90

HSP90 est une protéine chaperonne dynamique impliquée à la fois dans le repliement, la stabilisation et la prévention de l'agrégation de nombreuses protéines clientes. Pour exercer son rôle sur ses protéines clientes l'ATPase HSP90 adopte différentes conformations au cours d'un cycle où des changements structuraux complexes sont observés. Selon les modèles actuellement admis, HSP90 serait sous la forme d'un dimère ouvert. Lors de la liaison de l'ATP, le segment [G108-G135], appelé par la suite « ATP-lid », préalablement en position ouverte se stabilise en position fermée en recouvrant le site de fixation de l'ATP. Ce changement de conformation va exposer au solvant les résidus hydrophobes de l'hélice [A24-I34] conduisant à une dimérisation des deux domaines N-terminaux (NTD) d'HSP90 de façon transitoire. Cette dimérisation des NTDs est associée à un échange du premier brin β de chacun des NTDs qui stabilise davantage la forme fermée du dimère (McLaughlin et al. 2004; Hessling et al. 2009; Prodromou 2012; Schulze et al. 2016; Hoter et al. 2018). Une conformation encore plus compacte du dimère fermé d'HSP90 est obtenue lorsque les domaines MD et NTD se réarrangent pour assurer l'interaction du MD avec le γ -phosphate de l'ATP (Hessling et al. 2009; Prodromou and Pearl 2014). L'hydrolyse de l'ATP en ADP et phosphate inorganique conduit en fin de cycle à la réouverture du dimère au niveau des deux NTDs (Figure 3.1).

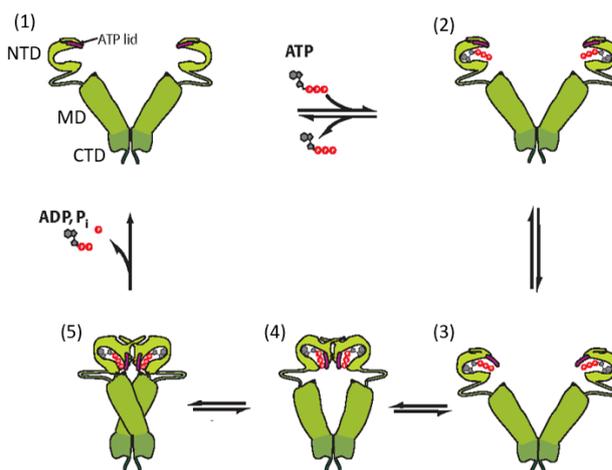


Figure 3.1: Cycle conformationnel d'HSP90 entière adapté de (Li and Buchner 2013). La liaison de l'ATP à la protéine entière Apo entraîne une fermeture de l'« ATP-lid » du NTD. Cette dernière induit une dimérisation transitoire des deux domaines NTD de la protéine qui adopte une conformation davantage compacte lorsque l'interaction entre le MD et le phosphate γ de l'ATP est assurée. L'hydrolyse de l'ATP entraîne la libération d'ADP et de phosphate inorganique et la chaperonne adopte à nouveau une conformation ouverte.

Lors de ce cycle, l'activité d'HSP90 est régulée par diverses modifications post-traductionnelles (Backe et al. 2020) ainsi que par de nombreuses co-chaperonnes. Plus d'une vingtaine de co-chaperonnes d'HSP90 ont été décrites, comme HOP, CDC37, p23, AHA1, CHIP et FKBP51/52, qui se lient à la chaperonne moléculaire de manière séquentielle au cours de son cycle fonctionnel (Zuehlke and Johnson 2010; Schopf et al. 2017). Le CTD possède un motif d'acides aminés MEEVD assurant la liaison de co-chaperonnes variées possédant des domaines tétratricopeptides (Taipale et al. 2010; Didenko et al. 2012; Biebl and Buchner 2019). Le MD et le NTD sont également impliqués dans la liaison avec d'autres co-chaperonnes telles que CDC37, AHA1 et p23.

Plusieurs centaines de protéines clientes d'HSP90 ont été identifiées, pouvant se classer en trois catégories : les facteurs de transcriptions, les protéines kinases ainsi qu'une troisième classe indépendante de protéines (Picard : http://www.picard.ch/downloads/Hsp90_interactors.pdf). Les protéines clientes se lient à la chaperonne moléculaire durant son cycle fonctionnel en fonction des différentes conformations qu'elle adopte. Toutefois, aucune séquence commune ni homologie structurale évidente n'ont été identifiées parmi la multitude de protéines clientes d'HSP90 (Zuehlke and Johnson 2010; Li et al. 2012).

HSP90 est une protéine essentielle qui présente une grande flexibilité. Les mouvements conformationnels de son domaine N-terminal initiés par la liaison de l'ATP puis son hydrolyse en ADP semblent être à l'origine de l'ensemble des réarrangements structuraux complexes suivis par la protéine entière. Je me suis donc intéressée en particulier durant ma thèse à la dynamique de ce domaine. HSP90 étant une chaperonne très étudiée, j'ai tout d'abord comparé les structures de HSP90-NTD α isolé déposées dans la Protein Data Bank (PDB) dans le but de m'en servir pour interpréter mes résultats de dynamique.

3.1.2 Etat de l'art de la PDB

HSP90 est non seulement une chaperonne moléculaire très étudiée mais c'est également une cible thérapeutique d'intérêt crucial. Depuis 1997, de nombreux inhibiteurs, ciblant le domaine NTD d'HSP90, ont été développés par les industries pharmaceutiques. Plusieurs centaines de structures du domaine N-terminal d'HSP90 ont donc été résolues et sont disponibles dans la PDB. La totalité de ces structures a été résolue par cristallographie aux rayons-X et, plus de 300 structures concernant uniquement le domaine N-terminal de la protéine humaine HSP90 α , sont déposées dans la banque de données. La première question que je me suis posée était : toutes ces structures sont-elles similaires ? Et si non, comment différent-elles ?

3.1.3 Analyse des structures de la PDB

Toutes les structures disponibles dans la PDB de HSP90 α -NTD humaine ont été classées en différents groupes (clusters) en fonction de leurs similarités structurales puis, les structures représentatives de chaque groupe ont été comparées. En effet, ainsi même les structures similaires très représentées n'induisent pas de biais dans la comparaison globale des structures. Pour les structures présentant plusieurs chaînes, toutes les chaînes ont été incluses dans l'analyse. Les structures dans lesquelles certains acides aminés étaient manquants ont été enlevées de l'analyse car empêchant une comparaison rigoureuse entre les structures. Au total, 334 structures de HSP90 α -NTD ont été sélectionnées et examinées en utilisant MaxCluster, un programme permettant de regrouper les structures en clusters (Herbert : <http://www.sbg.bio.ic.ac.uk/maxcluster>).

MaxCluster effectue un regroupement hiérarchique. Pour cela, l'ensemble des structures PDB fournies au programme sont alignées deux à deux en se basant sur les coordonnées des atomes C α de telle sorte que la racine quadratique de l'erreur moyenne (RQEM) calculée en utilisant la superposition entre toutes les paires de résidus donne le score le plus bas. Le regroupement commence ensuite en joignant les deux structures les plus proches (ayant le plus petit score RQEM) pour créer un nœud, et le processus est itéré en combinant à chaque fois soit les deux structures ou la structure et le nœud les plus proches, jusqu'à ce que toutes les structures/nœuds fassent partie du même groupe (Figure 3.2).

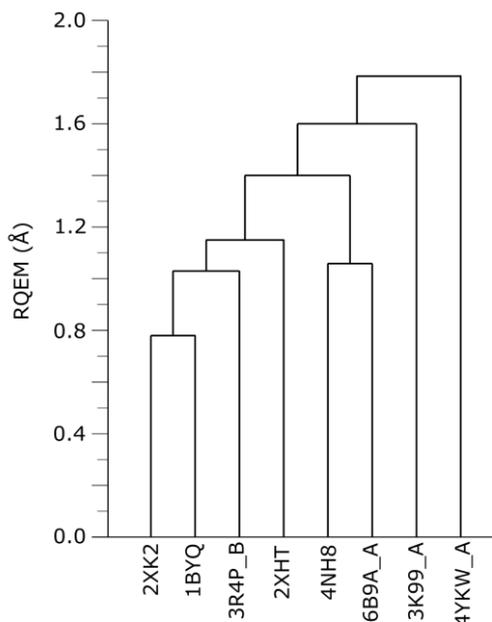


Figure 3.2 : Exemple d'arbre illustrant le principe du regroupement des structures en clusters. Les deux structures les plus proches se rejoignent pour créer un nœud, et le processus est itéré en combinant à chaque fois soit les deux structures ou la structure et le nœud les plus proches, jusqu'à ce que toutes les structures/nœuds fassent partie du même groupe.

Chaque arbre peut être divisé en n groupes avec n déterminé en fonction de la valeur seuil définie spécifiant la valeur RQEM maximum autorisée pour joindre deux structures/nœuds. Pour chaque groupe ainsi défini, MaxCluster donne la structure représentative. Cette dernière, également appelée le centroïde, est la structure qui présente le plus petit RQEM avec tous les autres membres du groupe (Herbert : <http://www.sbg.bio.ic.ac.uk/maxcluster>).

Ainsi de 334 structures initiales et en fixant une valeur seuil initiale de 1.05 Å, 8 groupes peuvent être constitués. La superposition deux à deux des 8 centroïdes de chaque groupe ainsi qu'une analyse de la racine quadratique de l'erreur moyenne (RQEM) en fonction de la séquence ont été effectuées sur les carbones C_{α} du squelette. Ces résultats montrent que les différences les plus importantes de structure entre les 8 centroïdes sont localisées au niveau de l'« ATP-lid » (segment [G108-G135]) - incluant le segment [L107-K116] pour lequel les groupements amides sont invisibles par RMN (Figure 3.3).

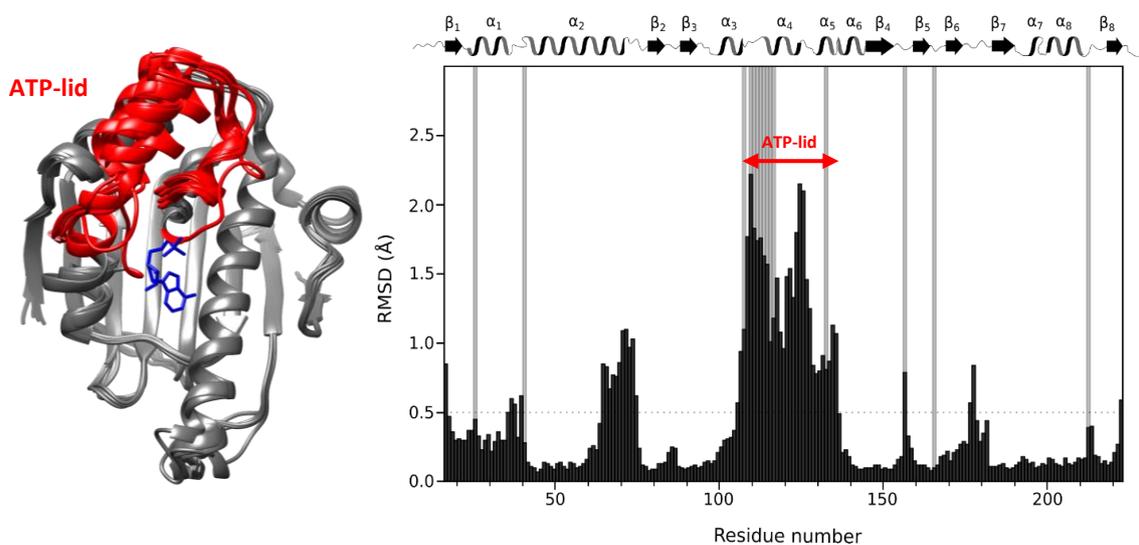


Figure 3.3 : A gauche : superposition des 8 centroïdes représentant les 8 clusters décrivant les 334 structures du domaine isolé de HSP90-NTD α disponibles dans la Protein Data Bank (le 5/01/2021). En rouge est représenté l'« ATP-lid » et en bleu un nucléotide se liant au domaine N-terminal. A droite : histogramme représentant la moyenne du RQEM calculée pour chaque paire de centroïde en fonction de la séquence (en noir). Les barres grises indiquent les résidus non attribués par (Jacobs et al. 2006; Park et al. 2011; Henot et al. 2021). Au-dessus de l'histogramme sont représentés les éléments de structure secondaire en fonction de la séquence, α pour les hélices et β pour les feuillets.

3.1.4 Structure de HSP90-NTD en solution à température ambiante

Ainsi, la question s'est posée de savoir quelles étaient les conformations échantillonnées par le domaine N-terminal de HSP90 α à température ambiante en solution puisqu'il n'a été observé qu'à l'état cristallisé. L'« ATP-lid » est en position ouverte (*i.e.* ne recouvrant pas le site de fixation de l'ATP) mais peut adopter différentes conformations (Figure 3.3) (hélice, boucle orientée vers le site de liaison de l'ATP ou boucle orientée vers l'extérieur de la protéine (Amaral et al. 2017)). Nous avons donc choisi de calculer l'ensemble conformationnel peuplé par le segment [98-136] (incluant l'hélice précédant l'« ATP-lid ») de HSP90-NTD Apo par RMN en acquérant des expériences NOESY afin d'en extraire des contraintes de distance. Les segments [17-97] et [137-223], inchangés dans les structures cristallographiques, ont été conservés fixes dans les calculs de structure décrits dans la suite de ce chapitre. Le calcul de structure a été effectué sur le segment [98-136], ce dernier ayant été défini dans notre article comme l'« ATP-lid », cependant, pour ma thèse j'ai préféré la définition de l'« ATP-lid » comme introduite par C. Prodromou (segment [G108-G135]) (Prodromou 2012).

3.2 HSP90-NTD échantillonne plusieurs conformations distantes de plus de 30 Å

Dans cette partie du projet, j'ai co-encadré Elisa Rioual, une étudiante en Master 2. Ensemble, nous avons analysé les spectres RMN obtenus à partir d'expériences NOESY dans le but d'obtenir les conformations échantillonnées par le segment [98-136] de HSP90-NTD. Elisa est ensuite devenue ma collaboratrice et a finalisé pendant sa première année de thèse les calculs de structures initiés pendant son M2.

3.2.1 Analyse de NOEs N_H-N_H et CH₃-CH₃

Les pics de corrélation NOEs obtenus avec des expériences de type NOESY peuvent être répartis en trois catégories. Il existe les NOEs séquentiels qui comprennent les interactions [i-1 ; i+1], utiles pour confirmer l'attribution de la protéine. Les NOEs locaux (moyenne distance) qui caractérisent des NOEs entre des sondes séparées par un maximum de 4 résidus, notamment intéressants pour déterminer la géométrie locale et les éléments de structure secondaire. Et enfin les NOEs longue distance entre des sondes proches dans l'espace mais éloignées dans la séquence de 5 résidus ou plus. Cette dernière catégorie de NOEs est particulièrement utile pour calculer la structure d'une protéine. Des pics de corrélation NOEs il est possible de dériver des contraintes de distance qui peuvent elles aussi être catégorisées de la même manière en contraintes de distance séquentielles, locales ou longues.

En acquérant une expérience 3D NOESY-BEST-TROSY (éditée HNH) sur un échantillon marqué [^2H , ^{15}N , ^{13}C] - HSP90-NTD, nous avons obtenu 41 contraintes de distance entre des paires de groupements H_N dont l'un, au moins, appartient au segment [98-136]. Parmi ces 41, seules trois sont des contraintes à longue distance (Figure 3.4). Toutefois, les NOEs obtenus avec l'expérience 3D NOESY-BTROSY (éditée HNH) nous ont permis de confirmer les éléments de structure secondaires prédits par TALOS+. En parallèle, nous avons cherché à détecter des NOEs inter-méthyles avec une expérience 3D HMQC-NOESY-HMQC (éditée CCH). Les groupements méthyles sont, en effet, des sondes idéales pour obtenir des informations à longue distance et ainsi calculer l'ensemble conformationnel peuplé par le segment [98-136] de HSP90-NTD, comprenant l'« ATP-lid ». Cette expérience a été effectuée sur un échantillon de HSP90-NTD marqué spécifiquement [^2H , ^{15}N , ^{12}C], $^{13}\text{C}^1\text{H}_3$ - (Ile- δ_1 , Leu- δ_2 , Val- γ_2 , Ala- β , Met- ϵ , Thr- γ_2). De cette expérience 3D, 111 contraintes de distance entre des paires de groupements méthyles dont l'un, au moins, appartient au segment [98-136] ont été obtenues. Parmi ces 111, plus de 82 % sont des contraintes longue distance (Figure 3.4).

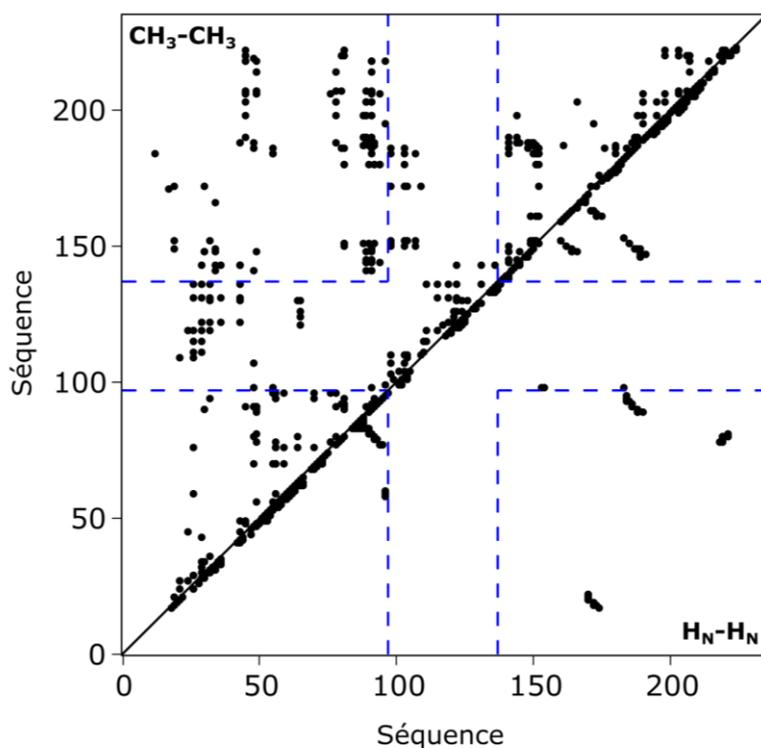


Figure 3.4 : Matrice résumant les contraintes de distance NOE observées. Dans le triangle en haut à gauche sont reportés les pics de corrélation NOEs $^{13}\text{C}^1\text{H}_3$ - $^{13}\text{C}^1\text{H}_3$ détectés pour HSP90-NTD α . Dans la région en bas à droite sont reportés les pics de corrélation NOEs H_N - H_N détectés pour HSP90-NTD α .

3.2.2 Incompatibilité des NOEs observés

En analysant les NOEs inter-méthyles détectés, des données incompatibles au regard des structures cristallographiques disponibles ont été observées. Par exemple

l'isoleucine 131, appartenant au segment [98-136] présente un pic de corrélation NOE avec la leucine 29 appartenant à la partie rigide de la protéine. Or l'isoleucine 131 est première voisine de la méthionine 130 et cette dernière possède des pics de corrélation NOE avec à la fois la thréonine 65 et la leucine 64, appartenant également à la partie fixe de la protéine et distantes de plus de 30 Å de la leucine 29 (Figure 3.5).

J'ai exclu la présence d'artéfacts en répétant l'expérience 3D HMQC-NOESY-HMQC deux fois, en utilisant deux échantillons de HSP90-NTD α marqués identiquement mais produits et purifiés de façon séparée. De plus, l'attribution présentée au chapitre 2 a été vérifiée. Pour cela les transferts d'attribution du squelette aux groupements méthyles pour les résidus isoleucine, leucine et valine ont été vérifiés et additionnellement un nouvel échantillon de HSP90-NTD marqué U-[^1H , ^{15}N , ^{13}C] a été produit pour transférer l'attribution du squelette aux groupements méthyles des alanines et thréonines avec des expériences « out and back HCC/HC(C)C ». Additionnellement, j'ai produit 23 mutants supplémentaires (soit un total de 56 mutants – voir chapitre 2, II.6) pour vérifier l'attribution des groupements méthyles des isoleucines, leucines et valines (Amero et al. 2011; Crublet et al. 2014). En conséquence, ces observations révélant des contraintes de distance ne pouvant être satisfaites simultanément ne peuvent s'expliquer qu'en prenant en compte plusieurs conformations très distinctes du segment [98-136] de HSP90-NTD.

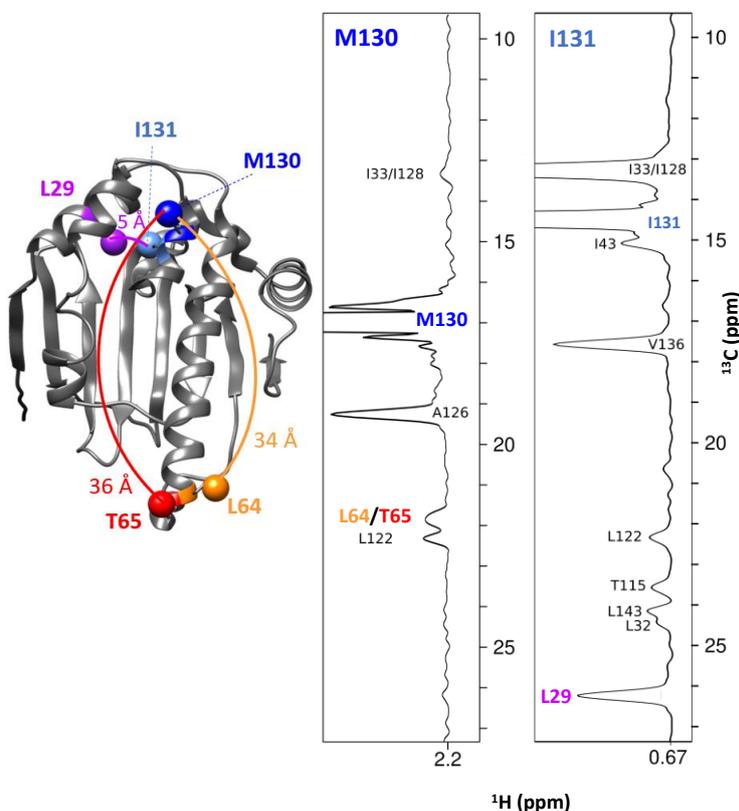


Figure 3.5: HSP90-NTD α échantillonne au minimum deux conformations en solution. A gauche : représentation, sur une structure 3D de HSP90-NTD α , d'exemples de NOEs inter-méthyles impliquant des résidus situés à plus de 30 Å. A droite : traces 1D extraites des expériences 3D CCH HMQC-NOESY-HMQC montrant M130 et I131 en pics diagonaux ainsi que leurs pics de corrélation NOEs associés. Le spectre NOESY a été acquis en un week-end sur un spectromètre opérant à une fréquence ^1H de 950 MHz et à 298K. Le temps de mélange NOE était de 400 ms.

3.2.3 Calcul de structure a un puis deux conformères

Nous avons effectué des calculs de structure en utilisant le logiciel CYANA (Combined assignment and dYnamics Algorithm for NMR Applications) (Güntert et al. 1997). La stratégie de calcul utilisée a été de fixer les parties conservées dans les structures tridimensionnelles : segment 11-97 et 137-223 et de déterminer la position du segment [98-136] d'après les contraintes expérimentales. Cela a été réalisé en extrayant les angles dièdres d'une structure préalablement résolue par cristallographie aux rayons X (PDB : 1YES) ce qui a permis de fixer la structure des domaines N- (11-97) et C- (137-223) terminaux. Leur position relative étant également fixée en utilisant des contraintes de distance C_{α} - C_{α} issues de la structure PDB : 1YES. L'ensemble conformationnel peuplé par le segment [98-136] a été déterminé en utilisant uniquement les données expérimentales constituées des contraintes de distance $^1H_N - ^1H_N$ and $^{13}CH_3 - ^{13}CH_3$ ainsi que les contraintes sur les angles dièdres (φ , ψ) dérivées de TALOS+ en utilisant l'attribution du squelette de HSP90-NTD (Henot et al. 2021). Un calcul avec un seul état a d'abord été effectué donnant un total de 25 violations associées à une valeur élevée de la fonction cible (74 kJ.mol^{-1}). Cette valeur est représentative des violations importantes des données de contraintes expérimentales par les structures calculées. Ainsi, ces résultats semblent confirmer notre hypothèse selon laquelle deux états ou plus représentent les conformations échantillonnées par le segment [98-136] en solution à température ambiante.

Nous avons donc choisi de mettre en place un protocole permettant le calcul simultané de deux états (Vögeli et al. 2014). Dans ce protocole, les contraintes de distance correspondent à des contraintes ambiguës et peuvent être satisfaites par l'un ou l'autre des deux états ou les deux états simultanément. Ce calcul à deux états nous a permis de conclure que toutes les contraintes pouvaient être satisfaites sans introduire de violation. Des tests supplémentaires incluant un 3^{ème} état ont également été faits mais aucune amélioration de la valeur de la fonction cible n'a été observée confirmant que deux états sont suffisants pour satisfaire toutes les contraintes structurales présentes en solution. Ainsi le segment [98-136] échantillonne deux ensembles conformationnels distincts de plus de 30 Å. Un premier état où ce dernier adopte une conformation ouverte qui est très comparable aux structures cristallographiques disponibles dans la PDB et un second où le segment adopte une conformation fermée – satisfaisant les contraintes de distance M130-T65 et M130-L64 utilisées en exemple précédemment (Figure 3.5). Cette seconde conformation où le segment [98-136] en position fermée recouvre complètement le site de fixation de l'ATP et où l'hélice-4 est dorénavant située à proximité de l'hélice-2 est totalement inédite pour le domaine N-terminal isolé de HSP90. Une analyse des contraintes de distance NOEs utilisées dans le calcul de structure a montré que 5 contraintes étaient spécifiques à la conformation avec le

segment [98-136] en position fermée et 22 contraintes de distance étaient caractéristiques du segment en position ouverte, les 125 autres étant communes aux deux ensembles conformationnels. Pour la suite, l'ensemble conformationnel ou le segment [98-136] adopte une conformation ouverte sera appelé « état ouvert » et celui ou le segment [98-136] adopte une conformation fermée, recouvrant le site de fixation de l'ATP sera appelé « état fermé ».

3.3 Mutants pour stabiliser individuellement chacune des conformations

3.3.1 Conception

Dans le but d'obtenir un unique jeu de contraintes NOE pour effectuer un calcul de structure avec un seul état mais également pour valider indépendamment l'existence des deux états en solution, j'ai cherché à concevoir des mutants permettant de stabiliser individuellement chacune des deux conformations. L'objectif étant de pouvoir calculer une structure tridimensionnelle avec une meilleure résolution et confirmer les structures des deux ensembles conformationnels obtenus.

Pour cela, à l'aide des structures disponibles, j'ai pu identifier des liaisons hydrogènes ou des ponts salins qui stabiliseraient l'une ou l'autre des conformations. Concernant l'« état ouvert », deux liaisons hydrogènes entre atomes des chaînes latérales (N40-D127 et S129-R46) ont été identifiées dans la structure cristallographique de HSP90 α -NTD (PDB : 1YES). Pour ce qui est de l'« état fermé », comme mentionné précédemment, il n'existait aucune structure du domaine N-terminal isolé avec le segment dans cette conformation. Par contre, la structure du dimère d'HSP90 β humain (PDB : 5FWM (Verba et al. 2016)) présente ses segments [98-136] en position fermée, recouvrant le site de fixation de l'ATP, dans les deux domaines N-terminaux. Cela nous a permis de construire par homologie une structure du NTD avec le segment [98-136] en position fermée pour HSP90 α -NTD. Notons qu'à l'époque où cette analyse a été faite la structure de cette conformation fermée d'HSP90 α humaine entière n'était pas encore disponible dans la PDB (PDB : 7L7J, (Lee et al. 2021)). Grâce à cette structure homologue, un pont salin (R60-D127) et une liaison hydrogène (D54-I131) ont été identifiés comme stabilisant le segment en position fermée. Cela m'a conduit à concevoir trois mutants qui potentiellement stabiliseraient complètement l'« état ouvert » : D54A, R60A et D127A ainsi que quatre mutants ayant l'effet inverse (i.e. stabiliser l'« état fermé ») : N40A, R46A, D127A et S129A. Le mutant I131A n'a pas été conçu car la liaison hydrogène entre D54 et I131 impliquait un atome de la chaîne latérale de D54 avec un atome du squelette de I131 ainsi le remplacement de I131 n'aurait pas modifié la liaison hydrogène.

3.3.2 Production des mutants

Les plasmides des six mutants ont été commandés chez Genecust puis les protéines ont été produites avec un marquage spécifique [^2H , ^{15}N , ^{12}C], $^{13}\text{C}^1\text{H}_3$ - (Ile- δ_1 , Leu-ProS, Val-ProS, Ala- β , Met- ϵ , Thr- γ_2). Le protocole de surexpression et purification employé est le même que pour la protéine HSP90-NTD sauvage et décrit dans l'article I présenté en 2.6. Les six mutants ont pu être produits et obtenus à des concentrations suffisantes pour une étude par RMN : c'est-à-dire pour les mutants D54A, R60A, D127A, N40A, R46A et S129A, respectivement : 0.54, 0.54, 0.31, 0.42, 0.52 et 0.41 mM. Les spectres 2D ^1H - ^{13}C HMQC acquis pour vérifier la qualité des échantillons ont montré que tous, à l'exception de D127A, étaient exploitables et ainsi, pour ces derniers, des expériences 3D HMQC-NOESY-HMQC ont été acquises. L'analyse du spectre du mutant D127A montre la présence de pics supplémentaires caractéristiques d'une hétérogénéité structurale. Ce résidu pourrait être impliqué dans des interactions stabilisant à la fois l'« état ouvert » et l'« état fermé » avec une constante de vitesse d'échange conformationnel entre les deux états très ralentie. Cependant, l'acquisition d'une expérience EXSY ne nous a pas permis de confirmer cette hypothèse. Aussi, nous nous sommes focalisés sur les autres mutants.

3.3.3 HSP90-NTD-R60A mutant stabilisant le segment [98-136] en position ouverte

L'analyse du spectre NOESY acquis sur le mutant HSP90-NTD-R60A révèle que les 5 NOEs longue distance caractéristiques de l'« état fermé » (M130-T65, M130-L64, A121-T65, A124-T65, A126-T65) sont absents pour cette construction (Figure 3.6). Ces résultats montrent que le mutant R60A déstabilise l'« état fermé » et diminue sa population en dessous du seuil de détection RMN.

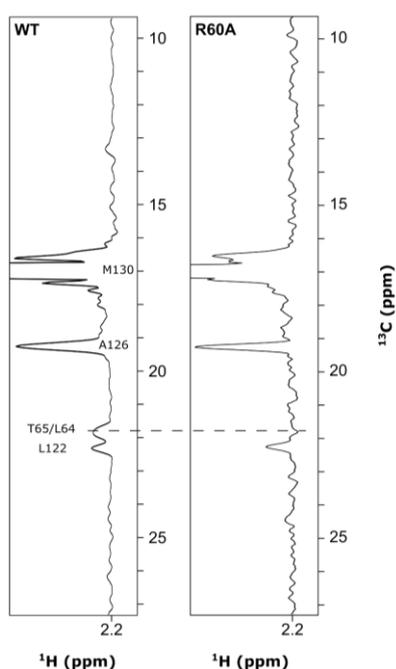


Figure 3.6 : Traces 1D extraites des expériences 3D CCH HMQC-NOESY-HMQC montrant M130 en pic diagonal ainsi que ses pics de corrélation NOEs associés pour la protéine sauvage HSP90-NTD (WT) et le mutant R60A-HSP90-NTD (R60A). Les spectres NOESY ont été acquis en 60 h sur des spectromètres opérant à des fréquences proton de 850 ou 950 MHz et à des températures de 298 K (WT) ou 293 K (R60A). Les temps de mélanges NOE étaient compris entre 300 et 400 ms. Les pics de corrélation M130-T65 et M130-L64 observés pour la trace 1D de HSP90-NTD WT sont absents pour la trace 1D du mutant R60A. Le pic de corrélation NOE M130-T65/L64 est un pic caractéristique de l'état fermé. Les pics de corrélation NOE M130-L122 et M130-A126 sont des pics intra segment [98-136].

Des calculs de structure ont été effectués avec le mutant HSP90-NTD-R60A, ce qui nous a permis de conclure qu'avec ce dernier toutes les contraintes NOE et angulaires étaient satisfaites avec un calcul employant un seul état. La structure obtenue pour ce mutant de HSP90-NTD est très similaire à celles obtenues par cristallographie aux rayons X (Figure 3.8).

3.3.4 HSP90-NTD-R46A mutant stabilisant le segment [98-136] en position fermée

A l'inverse du mutant HSP90-NTD-R60A, deux mutants R46A et S129A ont montré une augmentation de l'intensité des pics NOE caractéristiques de l'« état fermé » (Figure 3.7). De plus, pour ces deux mutants, une diminution d'environ 60 % de l'intensité des pics de corrélation NOE caractéristiques de l'« état ouvert » a été mesurée. Ces observations sont en accord avec le rôle stabilisant proposé de la liaison hydrogène R46-S129 pour l'état avec le segment [98-136] en position ouverte. Cependant, pour le mutant HSP90-NTD-N40A, supposé stabiliser également l'« état fermé », cette observation n'a pas été faite avec les expériences NOESY (Figure 3.7). Ainsi, le mutant N40A ne stabilise pas l'état fermé. Potentiellement la liaison N40-D127 n'est pas présente en solution à température ambiante ou cette liaison hydrogène existe mais n'est pas nécessaire pour la stabilisation de l'état ouvert.

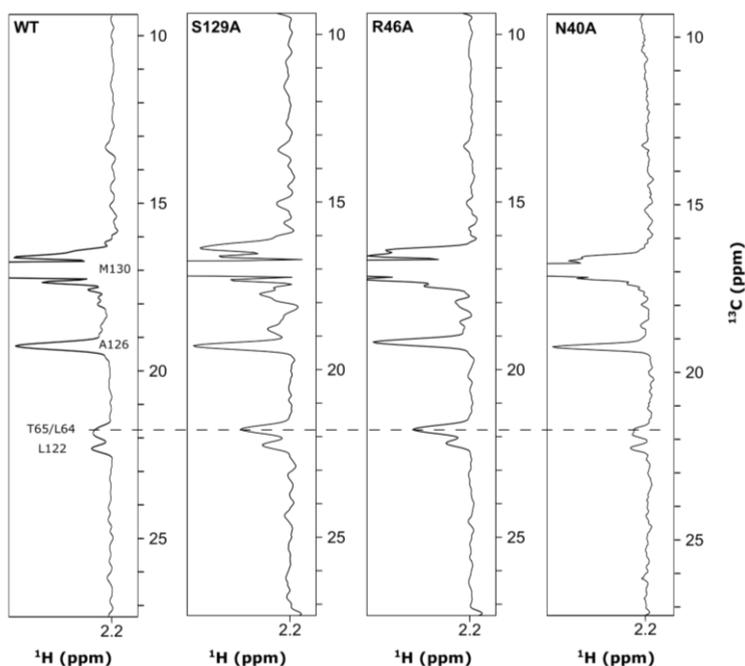


Figure 3.7 : Traces 1D extraites des expériences 3D CCH HMQC-NOESY-HMQC montrant M130 en pic diagonal ainsi que ses pics de corrélation NOEs associés pour la protéine sauvage HSP90-NTD (WT) et les mutants S129A, R46A et N40A. Les spectres NOESY ont été acquis en 2 à 3 jours sur des spectromètres opérant à des fréquences ^1H de 850 ou 950 MHz et à des températures de 298 K (WT) ou 288 K (R46A, S129A) et 293 K (N40A) pour prendre en compte la stabilité plus basse des mutants. Les temps de mélanges NOE étaient compris entre 0.3 et 0.4 s.

Les calculs de structure ont été poursuivis avec le mutant HSP90-NTD-R46A. En effet, bien que le mutant S129A montre des résultats similaires, la mutation du résidu arginine 46 en alanine se situe à une distance de plus de 20 Å du segment [98-136] et ainsi n'aurait un impact que mineur sur sa conformation contrairement à une mutation directe de la sérine 129 appartenant au segment [98-136].

Malgré la mutation stabilisant l'« état fermé », où le segment [98-136] recouvre le site de fixation de l'ATP, deux ensembles de NOEs, ceux caractéristiques de l'« état fermé » et ceux de l'« état ouvert » sont encore observés. Pour calculer une structure de l'« état fermé » uniquement, les pics de corrélation NOEs caractéristiques de l'« état ouvert » ont été identifiés dans les spectres du mutant HSP90-NTD-R60A (ne présentant qu'un seul état) et éliminés du calcul. La structure finale a été obtenue sans aucune violation des contraintes expérimentales. Dans cette structure les résidus formant les charnières permettant à l'« ATP-lid » de passer de la conformation ouverte à la conformation fermée sont les résidus [109-111] et [135-136] respectivement pour les 2 axes de pivot (Figure 3.8).

Comme mentionné précédemment, c'est la première fois qu'une structure avec le segment [98-136] recouvrant entièrement le site de fixation de l'ATP est reportée pour le domaine N-terminal isolé de HSP90 α Apo. Cela suggère que le domaine isolé de la protéine échantillonne déjà dans sa forme Apo cette conformation fermée du segment [98-136] observée uniquement dans le cycle conformationnel de la protéine entière après liaison de l'ATP.

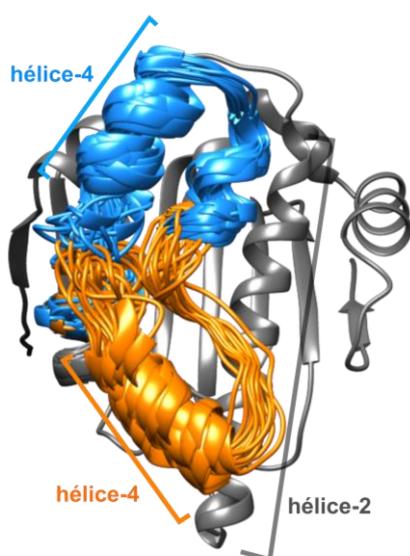


Figure 3.8 : Ensembles conformationnels des états ouverts et fermés du segment [98-136] de HSP90 α -NTD. Pour chaque état, les 20 conformères ayant la plus petite fonction cible sur CYANA ont été sélectionnés pour un raffinement supplémentaire employant la dynamique moléculaire en eau explicite. En bleu, ensemble conformationnel pour l'« état ouvert » calculé en utilisant les contraintes de distance RMN obtenues avec la construction R60A-HSP90 α -NTD. En orange, ensemble conformationnel pour l'« état fermé » calculé en utilisant les contraintes de distance RMN obtenues avec la construction R46A-HSP90 α -NTD. La partie invariable de la protéine est représentée en gris.

3.4 Etude de la dynamique d'HSP90-NTD Apo

Les résultats décrits précédemment montrent que HSP90 α -NTD est une protéine dynamique qui échantillonne deux ensembles conformationnels. J'ai cherché à caractériser de manière approfondie la cinétique et la thermodynamique de cet échange par RMN. La présence d'un seul jeu de signaux dans les expériences RMN 2D, que ce soit pour le spectre des méthyles ou des groupements amides, indique que les états sont en échange rapide par rapport à l'échelle des temps de la RMN ($\tau_{ex} < \sim 10$ ms). De plus, dans les spectres 2D, un élargissement des raies important pointe vers une dynamique dans la gamme de temps μ s-ms. J'ai donc utilisé des expériences de CPMG relaxation-dispersion pour caractériser l'échange entre les deux ensembles conformationnels observés. Cette technique, comme mentionné au chapitre 1, permet de quantifier l'augmentation de la relaxation transversale due à l'échange conformationnel et ainsi d'obtenir des informations sur la constante de vitesse d'échange conformationnel entre les deux conformations peuplées (k_{ex}), la population de l'état excité (P^*) ainsi que la valeur absolue de la différence de déplacement chimique entre l'état fondamental et celui excité. Les expériences CPMG peuvent être employées pour étudier la dynamique d'une protéine même quand l'état excité de cette dernière est très faiblement peuplé ($\sim 1\%$) et invisible sur les spectres (Kay 2016).

3.4.1 Acquisition des données de CPMG ^{15}N et $^{13}\text{CH}_3$

Des expériences de relaxation-dispersion CPMG méthyle TROSY $^{13}\text{C}^1\text{H}$ multiple quanta (Korzhnev et al. 2004a) modifiées avec l'angle d'excitation de Ernst (Figure 3.9), pour améliorer la sensibilité tout en diminuant le temps de recyclage (Amero et al. 2009), ont été acquises sur un échantillon de HSP90-NTD-WT marqué U- ^2H , ^{15}N , ^{12}C , Ala- $^{13}\text{C}^1\text{H}_3$ $^\beta$, Met- $^{13}\text{C}^1\text{H}_3$ $^\epsilon$, Leu/Val- $^{13}\text{C}^1\text{H}_3$ $^{\text{pro-S}}$, Ile- $^{13}\text{C}^1\text{H}_3$ $^{\delta 1}$, Thr- $^{13}\text{C}^1\text{H}_3$ $^\nu$ en utilisant des spectromètres opérant à des fréquences ^1H de 700 MHz et 850 MHz.

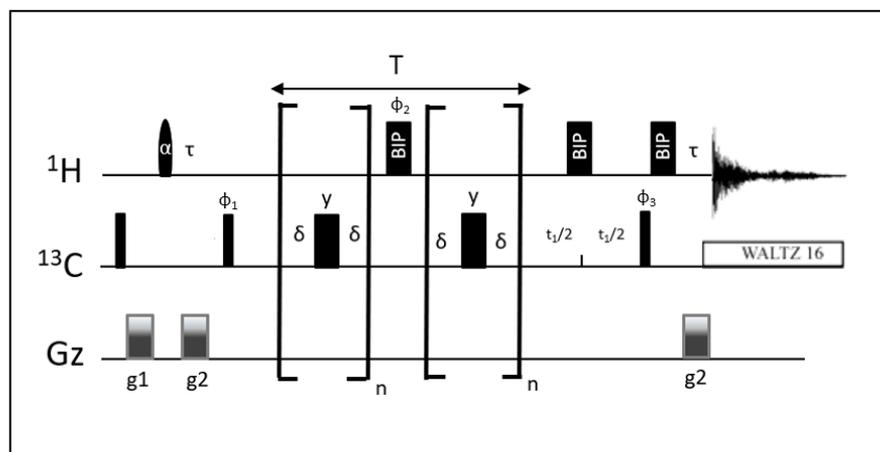


Figure 3.9 : Schéma de la séquence d'impulsions employée pour les expériences de relaxation-dispersion CPMG méthyle TROSY $^{13}\text{C}^1\text{H}$ multiple quanta (Korzhev et al. 2004a) modifiées avec l'angle d'excitation de Ernst. BIP = Broadband Inversion Pulse. $\alpha = 55^\circ$. Avec T le délai de relaxation CPMG, δ le délai situé avant et après chaque impulsion de 180° effectuée sur l'atome ^{13}C , n le nombre de fois où est répété le train d'impulsions CPMG avec $4n\delta = T$ et ν_{CPMG} la fréquence d'application des impulsions 180° refocalisantes avec $\nu_{\text{CPMG}} = 1/4\delta$. Avec $\phi_1 = x, -x, \phi_2 = 2(x), 2(y), 2(-x), 2(-y), \phi_3 = x$.

La puissance des impulsions refocalisantes (180°) ^{13}C a été fixée à 16 kHz. Le délai de relaxation CPMG (T_{CPMG}), le délai entre deux scans (d_1) et le temps d'acquisition dans la dimension indirecte ($t_{1\text{max}}$) étaient respectivement de 30 ms, 0.4 s et ca. 30 ms. Les expériences ont été effectuées à 293 K, en utilisant un échantillon de méthanol pour une calibration précise de la température de chaque spectromètre.

Pour les expériences de relaxation-dispersion, 16 fréquences de CPMG différentes allant de 66Hz à 2000 Hz (ν_{CPMG}) ont été acquises de manière intercalée. Des duplicatas ont été acquis pour les faibles fréquences de CPMG et additionnellement, un spectre de référence a été acquis en omettant le délai de relaxation. Le temps d'acquisition pour les expériences de relaxation-dispersion était de l'ordre de 25 h.

De la même manière, des expériences de relaxation-dispersion CPMG BEST-TROSY ^{15}N simple quantum (Franco et al. 2017) ont été acquises sur un échantillon de HSP90-NTD-WT marqué U- ^{2}H , ^{12}C , ^{15}N] en utilisant des spectromètres opérant à des fréquences ^1H de 600 MHz, 700 MHz et 850 MHz. Les paramètres employés pour ces expériences sont reportés dans le Tableau 3.1 ci-dessous.

Tableau 3.1 : Tableau récapitulatif des paramètres employés dans les expériences CPMG $^{13}\text{C}^1\text{H}_3$ et ^{15}N

Paramètres	Expériences $^{13}\text{C}^1\text{H}_3$	Expériences ^{15}N
Puissance impulsions (kHz)	16	4
T_{CPMG} (ms)	30	60
d_1 (s)	0.4 ^a	1.12
$t_{1\text{max}}$ (ms)	30	45
Température (K)	293	293
Durée (h)	25	70
Nombre de fréquences CPMG différentes	16 (66 Hz – 2000Hz)	11 (33 Hz – 1000 Hz)

^a Le délai plus court entre 2 scans pour les expériences de CPMG méthyle s'explique par l'utilisation d'une impulsion d'excitation faible (angle de Ernst équivalent à 55°).

Pour les sondes ^{15}N , comme les échantillons de HSP90-NTD sont perdeutérés et que nous souhaitons acquérir les expériences à haut voire très haut champ magnétique, la séquence d'impulsions CPMG a été implémentée dans une expérience BEST-TROSY simple quantum ce qui permet de bénéficier d'une meilleure résolution et d'une meilleure sensibilité.

Concernant les sondes méthyles, des expériences de relaxation-dispersion CPMG méthyle TROSY $^{13}\text{C}^1\text{H}$ multiple quanta ont été acquises car elles permettent d'obtenir, pour une macromolécule hautement deutérée, un spectre mieux résolu et jusqu'à trois fois plus sensible et que celui qui serait obtenu avec une expérience acquise en simple quantum (Tugarinov et al. 2003). Toutefois, avec la présence non seulement de l'hétéroatome mais également de celle du proton dans le plan, un paramètre supplémentaire sera à ajuster pour les courbes de relaxation-dispersion acquises en employant une séquence d'impulsions CPMG méthyle TROSY $^{13}\text{C}^1\text{H}$ multiple quanta (voir partie 3.4.2.2).

3.4.2 Traitement des données de relaxation-dispersion

3.4.2.1 Sélection des profils de relaxation-dispersion

L'intensité précise des pics des expériences de relaxation-dispersion a été déterminée en utilisant la routine d'ajustement non linéaire (nlinLS) du logiciel NMRPipe (Delaglio et al. 1995). Les profils de relaxation-dispersion ont été générés à partir des intensités de pics mesurées dans les expériences acquises avec les différentes fréquences CPMG divisées par les intensités des pics mesurées dans les spectres de référence (I_0) (Korzhnev et al. 2004a) et les vitesses de relaxation transversale effective (R_2^{eff}) ont été calculées avec l'Equation 3.1.

$$R_2^{\text{eff}}(v_{\text{CPMG}}) = - \frac{\ln\left(\frac{I_{\text{CPMG}}}{I_0}\right)}{T_{\text{CPMG}}} \quad (\text{Equation 3.1})$$

Tout d'abord une inspection visuelle de l'évolution de la vitesse de relaxation transversale effective en fonction de la fréquence CPMG a été effectuée afin d'exclure les courbes de relaxation dispersion trop bruitées. Les profils de relaxation-dispersion de pics non superposés pour lesquels la contribution de l'échange (R_{ex}) à la relaxation transversale, définie par la différence entre la valeur de R_2^{eff} extraite à v_{CPMG} minimum et celle extraite à v_{CPMG} maximum (respectivement, 66 Hz et 2000 Hz Figure 3.10), était supérieure à 2 s^{-1} ont été sélectionnés et ajustés avec le programme ChemEx. En effet, les profils de relaxation dispersion plats sont représentatifs de résidus ne percevant pas d'échange dans la gamme de temps μs -ms (ou de résidus impliqués dans l'échange mais

pour lesquels il n'y a pas de différence de déplacement chimique entre les deux états) (Figure 3.10). Pour l'ajustement avec le programme ChemEx, un modèle d'échange à deux états, comme suggéré par les expériences NOESY et le calcul de structure, a été utilisé (<https://github.com/gbouvignies/ChemEx>). Les erreurs sur les valeurs de R_2^{eff} ont été estimées avec deux fois le bruit observé dans le spectre. Cependant, quand ce dernier était inférieur à 2 % de la valeur de R_2^{eff} , une erreur planchée de 2 % a été admise (Korzhnev et al. 2004a).

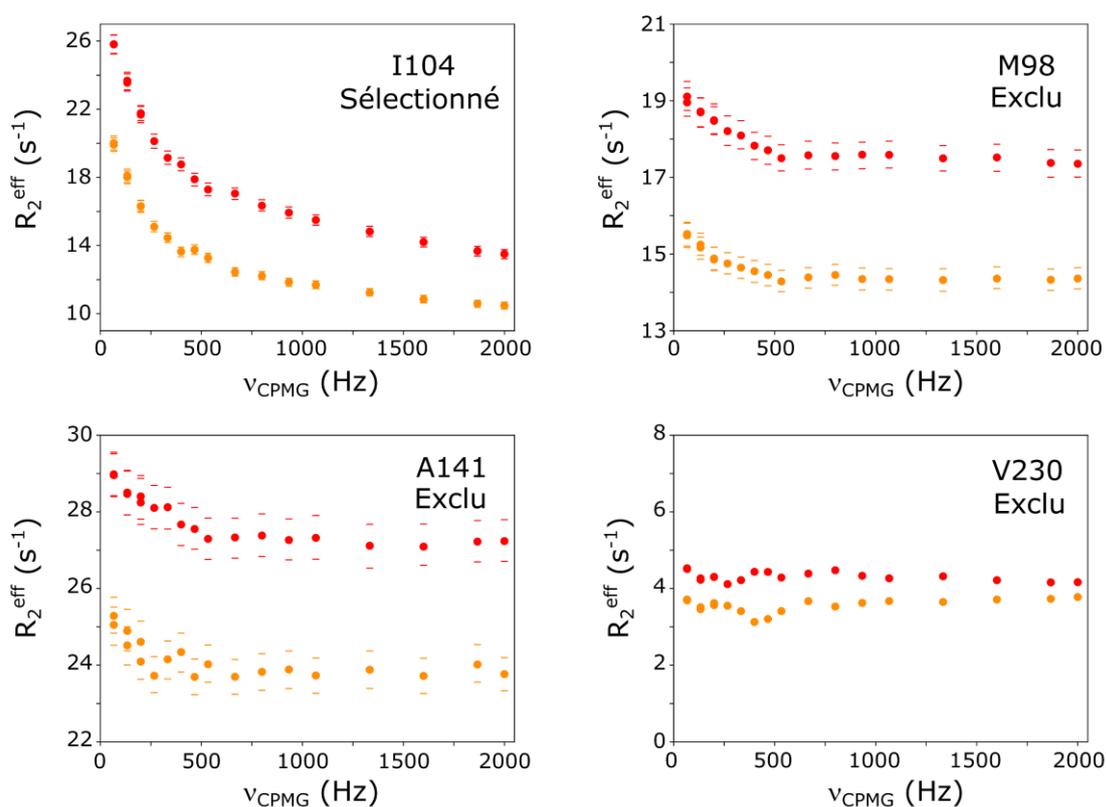


Figure 3.10 : Graphiques montrant la vitesse de relaxation transversale effective en fonction de la fréquence CPMG en rouge et orange pour respectivement des données acquises à 850 MHz et 700 MHz. En haut à gauche (I104), exemple d'un profil sélectionné pour les ajustements avec ChemEx. En haut à droite (M98), exemple d'un profil montrant une contribution de l'échange (R_{ex}) à la relaxation transversale inférieure à 2 s^{-1} et donc exclu de l'ajustement avec ChemEx. En bas à gauche (A141), exemple d'un profil montrant une contribution de l'échange (R_{ex}) à la relaxation transversale inférieure à 2 s^{-1} et bruité (notamment à 700 MHz) et donc non sélectionné pour l'ajustement avec ChemEx. En bas à droite (V230), exemple d'un profil plat signifiant l'absence d'échange dans la gamme de temps μs - ms pour ce résidu (ou qu'il s'agit d'un résidu impliqué dans l'échange mais pour lequel il n'y a pas de différence de déplacement chimique entre l'état fondamental et celui excité) et donc non gardé pour l'ajustement avec ChemEx.

3.4.2.2 Théorie de la relaxation dispersion et paramètres ajustés

Lorsque l'on fait un ajustement des courbes sur les données expérimentales pour un résidu on peut obtenir des informations, comme mentionné précédemment, sur la cinétique (k_{ex}), la thermodynamique (P^* , la population de l'état excité) et sur la valeur absolue de la différence de déplacement chimique entre l'état fondamental et celui excité.

En effet, si l'on a un échange entre deux états A et A*, se produisant dans la gamme de temps observable par CPMG-RD, où les populations des deux états sont très déséquilibrées ($P_A \gg P_{A^*}$), il est possible, pour une expérience de relaxation dispersion multiple quanta acquise avec un délai de CPMG constant, de décrire la relaxation transversale par l'équation 3.2 (Carver and Richards 1972; Korzhnev et al. 2004a).

$$R_2^{eff} = Re(\lambda_1) - \frac{1}{4n\delta} \ln(Q) \quad (\text{Equation 3.2})$$

$$\text{Avec } \lambda_1 = R_2^0 + \frac{1}{2} \left(k_{ex} - \frac{1}{2\delta} \cosh^{-1}(D_+ \cosh \eta_+ - D_- \cosh \eta_-) \right)$$

$$D_{\pm} = \frac{1}{2} \left(\frac{\psi + 2(\Delta\omega_X)^2}{\sqrt{\psi^2 + \zeta^2}} \pm 1 \right) \quad \eta_{\pm} = \sqrt{2}\delta \frac{1}{2} (\sqrt{\psi^2 + \zeta^2} \pm \psi)^{\frac{1}{2}}$$

$$\psi = (i\Delta\omega_H + (P_A - P_{A^*}) k_{ex})^2 - (\Delta\omega_X)^2 + 4 P_A P_{A^*} k_{ex}^2$$

$$\zeta = -2\Delta\omega_X (i\Delta\omega_H + (P_A - P_{A^*}) k_{ex})$$

$$\text{Et avec } Q = Re \left(1 - m_D^2 + m_D m_Z - m_Z^2 + \frac{1}{2} (m_D + m_Z) \sqrt{\frac{P_{A^*}}{P_A}} \right)$$

$$m_D = \frac{ik_{ex}\sqrt{P_A P_{A^*}}}{d_+ z_+} \left(z_+ + 2\Delta\omega_X \frac{\sin(z_+ \delta)}{\sin((d_+ + z_+) \delta)} \right)$$

$$m_Z = - \frac{ik_{ex}\sqrt{P_A P_{A^*}}}{d_- z_-} \left(d_- + 2\Delta\omega_X \frac{\sin(d_- \delta)}{\sin((d_- + z_-) \delta)} \right)$$

$$d_{\pm} = (\Delta\omega_H + \Delta\omega_X) \pm ik_{ex} \quad z_{\pm} = (\Delta\omega_H - \Delta\omega_X) \pm ik_{ex}$$

Avec R_2^{eff} la vitesse de relaxation transversale effective, R_2^0 la vitesse de relaxation transversale intrinsèque (en absence d'échange chimique) comprenant les contributions des interactions dipolaires et l'anisotropie de déplacement chimique. δ le délai situé avant et après chaque impulsion de 180° effectuée sur l'hétéroatome X et n le nombre de fois où est répété le train d'impulsions CPMG ($4n\delta$ étant égal à T le délai constant de relaxation CPMG). Avec P_{A^*} et P_A les populations respectives des états A^* et A , avec $P_A \gg P_{A^*}$ et $P_{A^*} = 1 - P_A$. $k_{ex} = k_A + k_{A^*}$ représente la constante de vitesse d'échange conformationnel entre les deux conformations A et A^* et $\Delta\omega$ est la valeur de la différence de déplacement chimique entre les états A et A^* (pour l'hydrogène $\Delta\omega_H$ et $\Delta\omega_X$ pour l'hétéroatome) en $\text{rad}\cdot\text{s}^{-1}$.

Dans le cas d'expériences simple quantum telles que les expériences acquises sur les sondes ^{15}N alors $\Delta\omega_H = 0$. Ainsi, l'équation 3.2 se simplifie en équation 3.3, également appelée expression de Carver-Richards, avec l'approximation que le terme $\ln(Q)$ est négligeable. En pratique, pour des expériences simple quantum acquises avec un délai de CPMG T constant et supérieur à 20 – 30 ms les erreurs faites en raison de l'omission du terme $\ln(Q)$ seront comprises dans le bruit (Carver and Richards 1972; Davis et al. 1994; Korzhnev et al. 2004a, b).

$$R_2^{eff} = R_2^0 + \frac{1}{2} \left(k_{ex} - \frac{1}{2\delta} \cosh^{-1}(D_+ \cosh\eta_+ - D_- \cosh\eta_-) \right) \quad (\text{Equation 3.3})$$

D'après ces équations on peut voir qu'avec un ajustement des courbes de relaxation dispersion dans le cadre d'expériences simple quantum ^{15}N on obtient des informations sur la cinétique (k_{ex}), la thermodynamique (P^* i.e. la population de l'état excité) et sur $|\Delta\omega_N|$ la valeur absolue de la différence de déplacement chimique de l'azote entre l'état fondamental et celui excité soit au total trois paramètres. Avec un ajustement des courbes de relaxation dispersion dans le cadre d'expériences multiple quanta $^{13}\text{C}^1\text{H}$ on obtient des informations sur k_{ex} , P^* et sur $|\Delta\omega_C|$ et $|\Delta\omega_H|$ les valeurs absolues de

déplacements chimiques entre l'état fondamental et celui excité pour le carbone et l'hydrogène soit au total quatre paramètres.

k_{ex} , P^* et la valeur absolue de la différence de déplacement chimique entre l'état fondamental et celui excité sont corrélées lorsque les données de relaxation dispersion sont acquises à un unique champ magnétique. En employant un set de données synthétiques pour lequel 1000 sets de données synthétiques incorporant du bruit ont été générés, il a été montré (Kovrigin et al. 2006), en ajustant les 1000 sets de données synthétiques avec l'équation 3.3, que les corrélations entre les trois paramètres (k_{ex} , $P^* = 1 - P$ et $|\Delta\omega_X|$) étaient conséquentes à un unique champ magnétique (Figure 3.11a). La corrélation entre la population et la différence de déplacement chimique ainsi que la corrélation entre la population et la constante de vitesse d'échange conformationnel sont les plus problématiques puisque pour une unique valeur de population, différentes valeurs de différence de déplacement chimique et/ou de constante de vitesse d'échange conformationnel peuvent être obtenues. Acquérir des expériences de relaxation-dispersion à deux champs magnétiques différents permet de casser cette corrélation (Figure 3.11b) (Kovrigin et al. 2006; Vallurupalli 2009) lorsque l'échange se produit dans la gamme de temps intermédiaire.

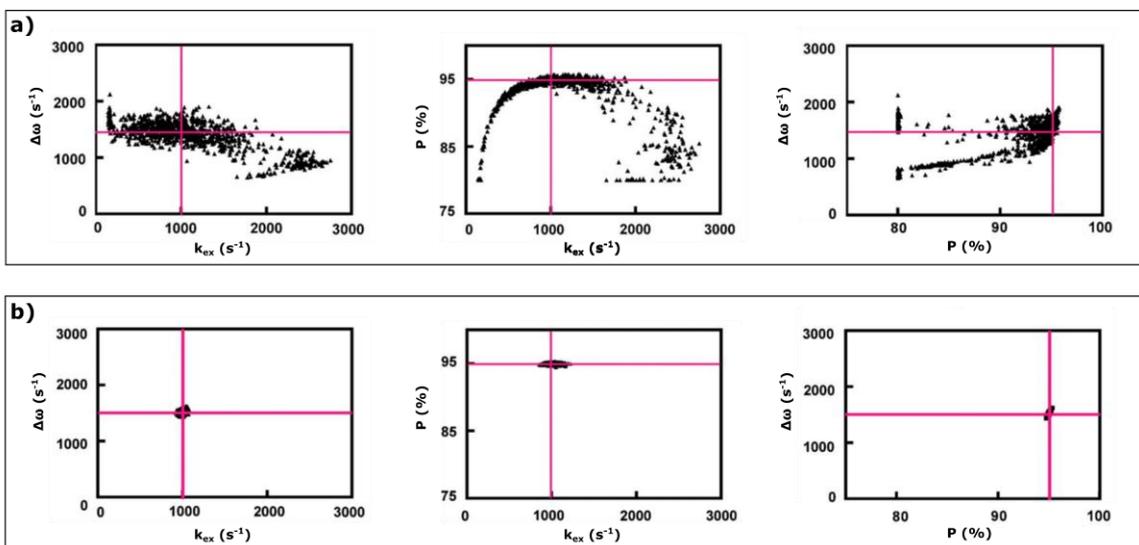


Figure 3.11 : a) Graphiques montrant la corrélation deux à deux entre les paramètres k_{ex} , P et $\Delta\omega$ lorsque les 1000 sets de données synthétiques incorporant du bruit sont ajustés avec l'équation 3.3 et qu'un unique champ magnétique est utilisé. Les lignes roses indiquent les paramètres initiaux utilisés pour générer le set de données synthétiques initial b) Graphiques montrant la corrélation deux à deux entre les paramètres k_{ex} , P et $\Delta\omega$ lorsque les 1000 sets de données synthétiques incorporant du bruit sont ajustés avec l'équation 3.3 et que deux champs magnétiques sont utilisés. Les lignes roses indiquent les paramètres initiaux utilisés pour générer le set de données synthétiques initial. Figure adaptée de (Kovrigin et al. 2006).

Ainsi, employer deux champs magnétiques différents est indispensable pour obtenir une étude fiable de la dynamique de notre protéine d'intérêt. L'ajout d'un troisième champ magnétique améliore légèrement l'estimation des paramètres obtenus mais cette faible amélioration ne justifie pas forcément le temps d'utilisation d'un spectromètre supplémentaire (Kovrigin et al. 2006).

Pour la suite de ma thèse j'emploierai la notation $|\overline{\Delta\omega}|$ pour décrire la variation de déplacement chimique entre l'état fondamental et l'état excité et non plus la notation $|\Delta\omega|$. En effet, j'ai utilisé deux champs magnétiques, aussi en employant la notation $|\Delta\omega|$ (s^{-1}) deux valeurs seraient nécessaires pour décrire les différences de déplacements chimiques pour chaque résidu en échange. En utilisant la notation $|\overline{\Delta\omega}|$, une unique valeur en ppm peut être employée pour décrire la différence de déplacement chimique entre l'état fondamental et l'état excité quel que soit le champ magnétique, simplifiant les notations.

3.4.2.3 Influence de chacun des paramètres sur l'ajustement

Dans le cas d'ajustements de données de relaxation dispersion obtenues avec des expériences CPMG multi quanta $^{13}C^1H$, 4 paramètres : k_{ex} , P^* et sur $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$ peuvent avoir une influence sur la courbe ajustée par rapport aux données initiales. J'ai donc cherché à comprendre l'influence de chacun de ces quatre paramètres sur l'ajustement. Pour cela j'ai ajusté à l'aide du programme ChemEx les données de relaxation dispersion obtenues à deux champs pour un résidu en faisant varier uniquement un paramètre et en conservant les trois autres constants et j'ai itéré cela pour tous les paramètres (Figure 3.12).

Les valeurs de $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$ choisies varient d'une valeur en ppm proche ($|\overline{\Delta\omega_C}|$) ou égale ($|\overline{\Delta\omega_H}|$) à 0 jusqu'à une valeur maximum correspondant à la déviation standard moyenne observée dans la Biological Magnetic Resonance Data Bank (BMRB) pour les déplacements chimiques carbone ou proton des groupements méthyles : Ala- β , Val- γ_2 , Thr- γ_2 , Ile- δ_1 , Leu- δ_2 et Met- ϵ .

Tandis que k_{ex} , la constante de vitesse d'échange conformationnel, a une influence sur la localisation du point d'inflexion de la courbe d'ajustement et sur la hauteur de l'ajustement entre sa valeur à $\nu_{CPMG} = 0$ Hz et $\nu_{CPMG} = 2000$ Hz, j'ai pu observer que P^* et $|\overline{\Delta\omega_C}|$ avaient tous les deux un effet uniquement sur la hauteur de l'ajustement entre sa valeur à $\nu_{CPMG} = 0$ Hz et $\nu_{CPMG} = 2000$ Hz. L'impact de $|\overline{\Delta\omega_H}|$ sur l'ajustement apparaît moindre. En effet, les courbes d'ajustement sont similaires même lorsque la valeur de $|\overline{\Delta\omega_H}|$ change de 0 à 0.1 ppm. Une influence sur la hauteur de l'ajustement pour les valeurs de fréquences CPMG faibles apparaît à partir de $|\overline{\Delta\omega_H}| = 0.2$ ppm. On note que

pour des fortes valeurs de $|\overline{\Delta\omega_H}|$, la contribution de l'échange (R_{ex}) à la relaxation transversale, définie par la différence entre la valeur de $R_{2,eff}$ extraite à ν_{CPMG} minimum et celle extraite à ν_{CPMG} maximum, peut être négative.

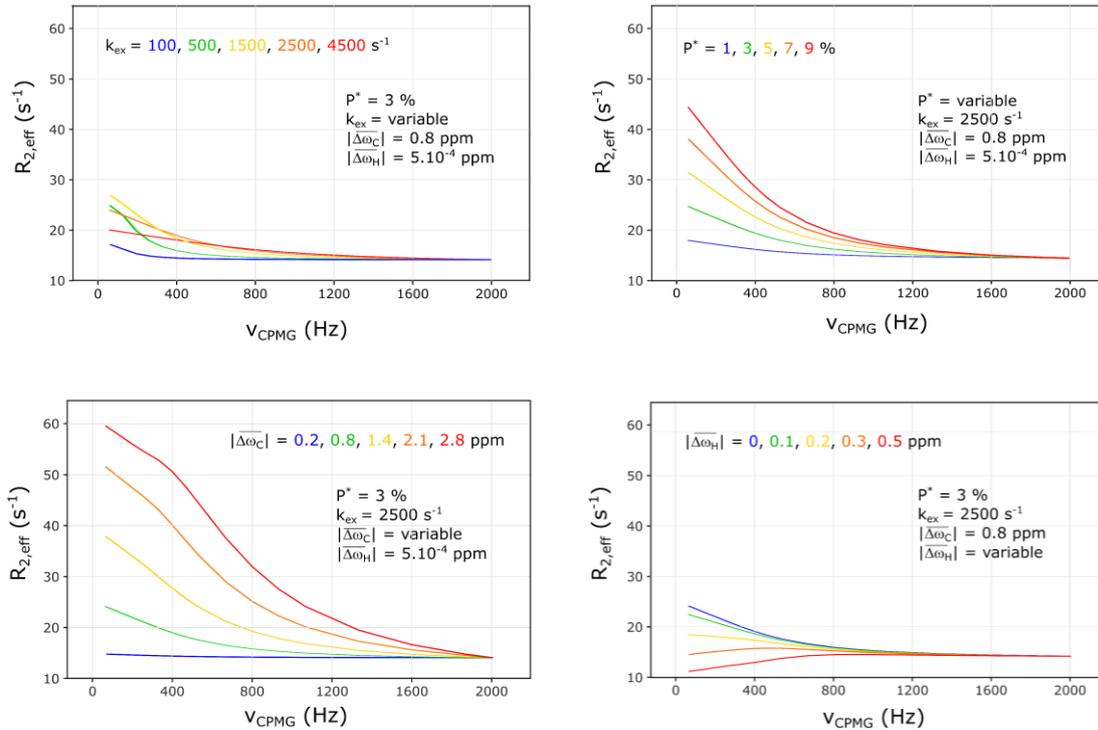


Figure 3.12 : Exemples d'ajustements des données de relaxation dispersion montrant l'influence des différents paramètres k_{ex} (en haut à gauche), P^* (en haut à droite), $|\overline{\Delta\omega_C}|$ (en bas à gauche) et $|\overline{\Delta\omega_H}|$ (en bas à droite). Les ajustements de relaxation dispersion représentés sont ceux obtenus à 850 MHz toutefois, chaque ajustement a été fait en employant deux champs 700 MHz et 850 MHz. Pour chaque panneau sont reportés les paramètres employés pour l'ajustement. De bleu, à vert, à jaune, à orange et à rouge les valeurs des paramètres variables, augmentent. Les valeurs de ces derniers sont également reportées sur chaque panneau et sont de 100, 500, 1500, 2500 et $4500 s^{-1}$ pour k_{ex} , de 1, 3, 5, 7, 9 % pour P^* , de 0.2, 0.8, 1.4, 2.1 et 2.8 ppm pour $|\overline{\Delta\omega_C}|$ et 0, 0.1, 0.2, 0.3 et 0.5 ppm pour $|\overline{\Delta\omega_H}|$.

3.4.2.4 Vers un ajustement global des données de relaxation dispersion

Avec l'ajustement des courbes de relaxation dispersion et l'obtention des paramètres k_{ex} , P^* , $|\overline{\Delta\omega_X}|$ et dans le cadre d'expériences multi quanta $|\overline{\Delta\omega_H}|$ des informations cruciales sur un état excité peuvent être obtenues. Toutefois, il est à noter que si l'on ajuste les courbes d'un unique résidu et ce, même avec des expériences acquises à deux champs on obtiendra quatre (trois) paramètres pour seulement 2 courbes ajustées dans

le cas d'expériences multi-quanta $^{13}\text{C}^1\text{H}$ (simple quantum ^{15}N), laissant apercevoir l'imprécision sur les données extraites de ces ajustements.

Une solution pour pallier ce problème d'imprécision sur les données extraites des ajustements est de procéder à des ajustements globaux. En effet, Kovermann et al. (Kovermann et al. 2016) décrivent que les mouvements lents d'une protéine tels que les changements conformationnels impliquent souvent le mouvement concerté d'un large groupe d'atomes et qu'il peut être attendu que ces résidus partagent une même constante de vitesse d'échange conformationnel et une même population d'état excité. Ainsi ajuster ensemble un large groupe d'atomes subissant le même échange conformationnel permettrait de réduire le rapport du nombre de paramètres à extraire des ajustements de courbes sur le nombre de courbes ajustées. Par exemple en faisant un ajustement global de N résidus avec des expériences acquises à 2 champs, je disposerai de $2N$ courbes pour trouver $2N + 2$ paramètres dans le cas d'expériences $^{13}\text{C}^1\text{H}_3$ et de $2N$ courbes pour trouver $N + 2$ paramètres dans le cas d'expériences ^{15}N , permettant ainsi d'augmenter la précision des valeurs obtenues de k_{ex} et P^* (Demers and Mittermaier 2009). Toutefois, la précision des valeurs de $|\overline{\Delta\omega_{\text{C}}}|$ et $|\overline{\Delta\omega_{\text{H}}}|$ ou de $|\overline{\Delta\omega_{\text{N}}}|$ qui sont obtenues à chaque fois pour un résidu ne peut que très légèrement être améliorée avec l'ajustement global.

3.4.2.5 Sélection des paramètres initiaux de l'ajustement global

Avant d'effectuer un ajustement global des résidus dynamiques il est important d'effectuer une grille de recherche en fixant à la fois la constante de vitesse d'échange conformationnel et la population de l'état excité afin d'identifier le(s) minimum(s) local(locaux) pour choisir ces paramètres comme paramètres initiaux dans l'ajustement global. Cette étape a été réalisée à la fois pour les sondes méthyles et celles ^{15}N sur la protéine HSP90-NTD-WT. Un unique et identique minimum local ($k_{\text{ex}} \sim 2600 \text{ s}^{-1}$ et $P^* \sim 3\%$) a été identifié pour les données de relaxation ^{15}N et $^{13}\text{CH}_3$, confirmant que les deux types de sondes observent le même phénomène (Figure 3.13). On note également que la grille de recherche effectuée sur les sondes méthyles semble être plus précise sur la valeur P^* par rapport à la valeur de k_{ex} et l'inverse est observé pour la grille de recherche effectuée sur les sondes ^{15}N .

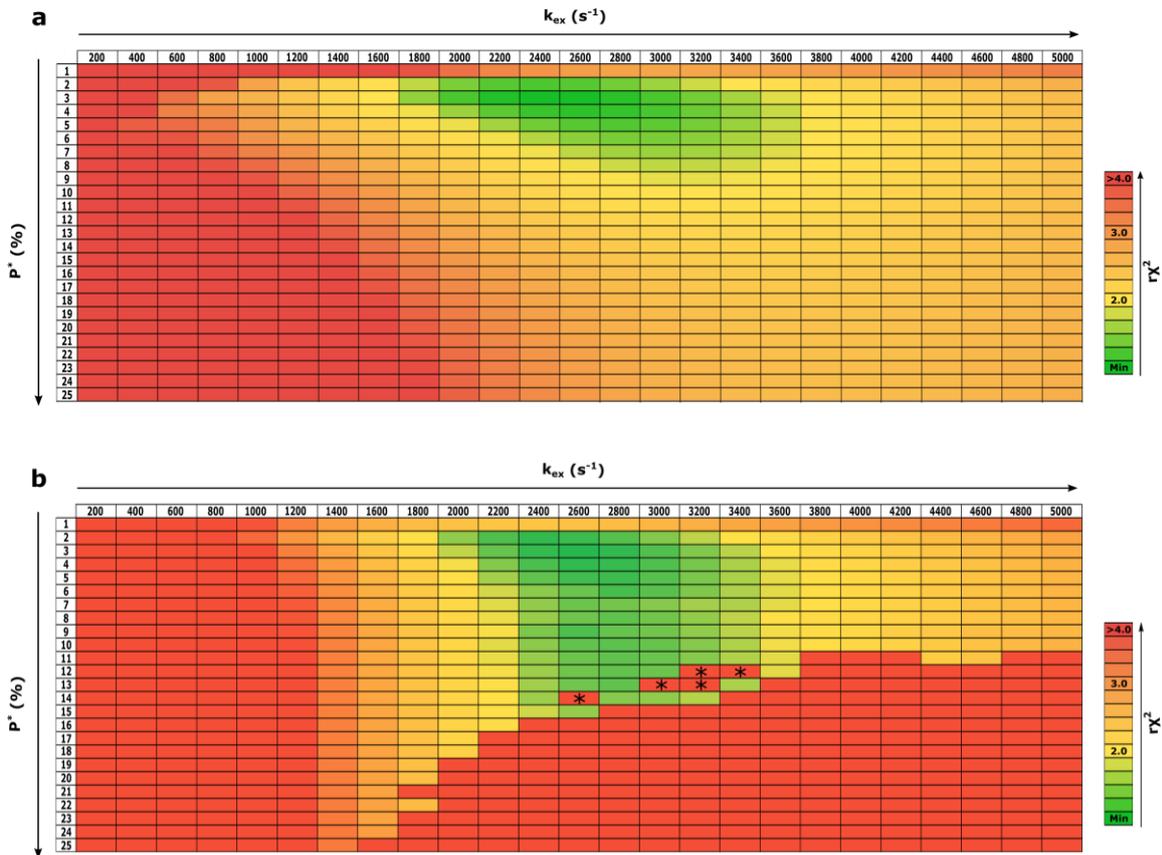


Figure 3.13 : Grilles de recherche représentant le χ^2 réduit obtenu pour un ajustement global lorsque la constante de vitesse d'échange conformationnelle et la population sont fixées. Pour des données de relaxation a) $^{13}\text{C}^1\text{H}$ et b) ^{15}N , acquises à 293 K. De vert à rouge les valeurs du χ^2 réduit augmentent. On note pour la grille de recherche des sondes ^{15}N des valeurs du paramètre χ^2 réduit, traduisant la qualité de l'ajustement, très élevées (représentés par le symbole *) au sein de la zone des valeurs très faibles de ce dernier. Une raison potentielle à cette observation est que le programme Chemex n'arrive pas à converger avec ces valeurs initiales de $k_{\text{ex}}\text{-}P^*$. Toutefois, même si la raison exacte d'une telle observation n'est pas définie avec précision cela n'influence pas le résultat de cette analyse montrant qu'un unique et identique minimum local existe pour les données de relaxation ^{15}N et $^{13}\text{CH}_3$. De plus, ces valeurs initiales de $k_{\text{ex}}\text{-}P^*$ donnant un χ^2 réduit élevé se situent tout de même relativement proches de la zone en bas à droite du graphique pour laquelle toutes les valeurs du paramètre χ^2 réduit sont très élevées.

3.4.2.6 Processus d'échange commun et ajustement global

Les données de relaxation dispersion CPMG effectuées sur les sondes méthyles donnant un rapport signal sur bruit 3 fois plus élevé que les données acquises sur les sondes ^{15}N , la suite de l'analyse, très chronophage, a été poursuivie uniquement sur les données

$^{13}\text{CH}_3$. Après cette étape de grille de recherche et avant d'effectuer un ajustement global des résidus il est primordial de s'assurer que ces résidus appartiennent au même processus d'échange. Pour cela un ajustement global des résidus préalablement sélectionnés est effectué en éliminant à tour de rôle un résidu de l'ajustement global à chaque fois. Ainsi si N résidus montrent un profil de relaxation dispersion intéressant, il faut effectuer N ajustements globaux avec N-1 résidus pour s'assurer qu'aucun d'entre eux ne tire la cinétique ou la thermodynamique vers des valeurs différentes que celles partagées par le plus grand nombre de résidus (*i.e.* pour appartenir au même processus d'échange, toutes les valeurs de k_{ex} obtenues pour les N ajustements doivent être comprises entre la moyenne des N valeurs de $k_{\text{ex}} \pm 3$ fois la déviation standard, la même règle s'applique pour P^*) (Figure 3.14).

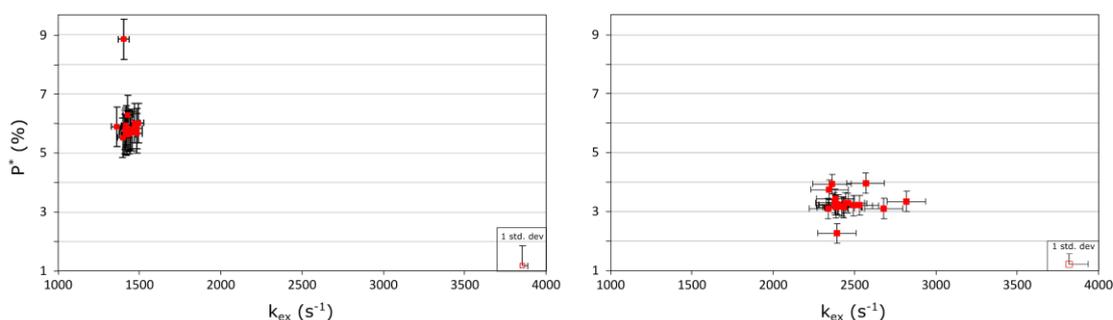
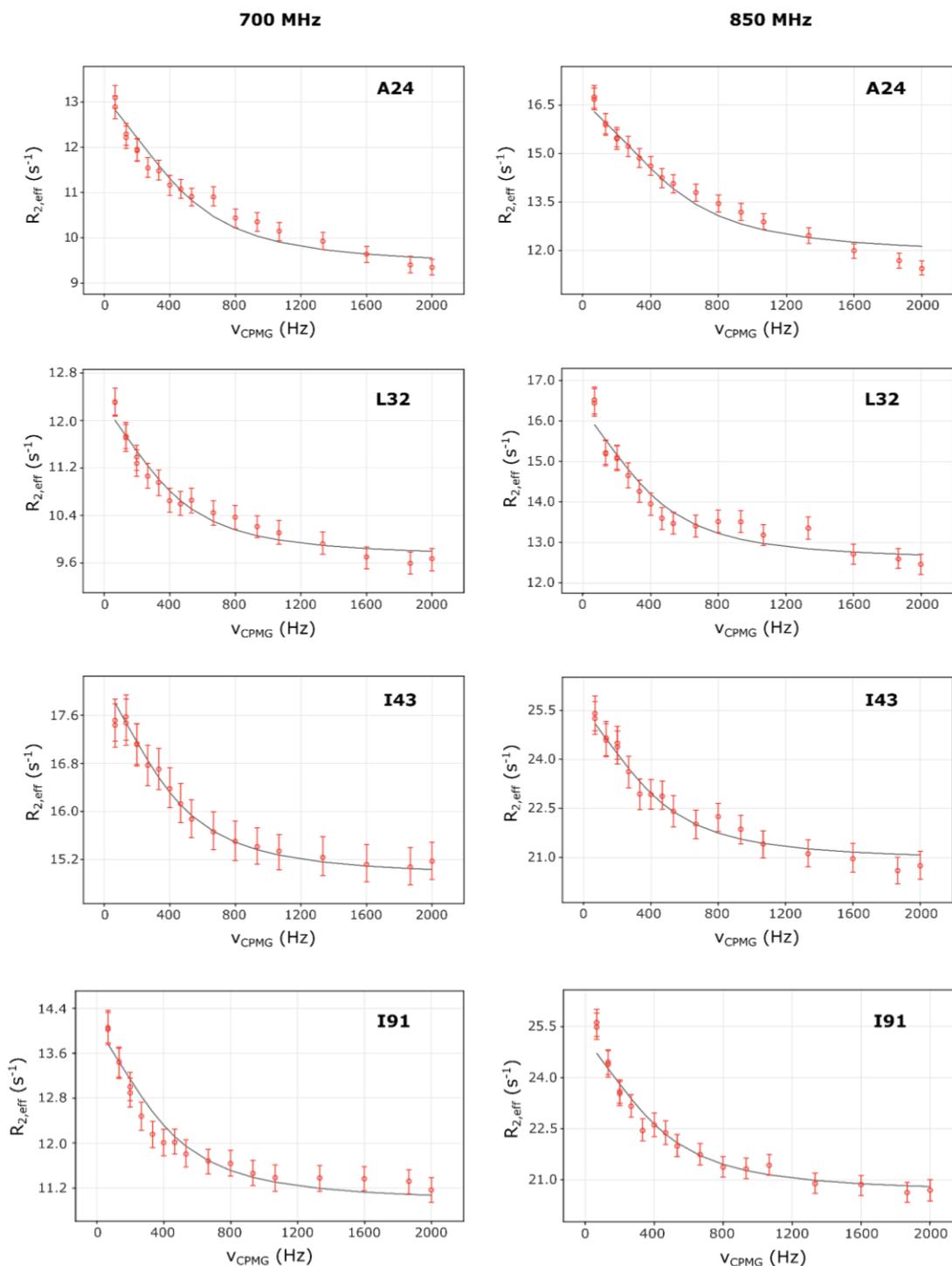


Figure 3.14 : Graphique représentant la population de l'état excité (P^*) en fonction de la constante de vitesse d'échange conformationnel (k_{ex}). Les N points rouges représentent chacun un ajustement global répété N-1 fois en éliminant à chaque fois un résidu différent de l'ajustement global. Les barres d'erreurs sur x et y sont, respectivement, les déviations standard sur k_{ex} et P^* . A gauche, la figure illustre que tous les résidus ajustés globalement n'appartiennent pas au même processus d'échange. Les données présentées montrent qu'un résidu tire la thermodynamique vers une valeur différente de celle partagées par le plus grand nombre de résidus. Ces données, employées ici uniquement à titre d'exemple pour illustrer des résidus n'appartenant pas au même processus d'échange proviennent d'un échantillon d'HSP90-NTD pour lequel les conditions sont différentes : présence d'un ligand et température différente, expliquant les différentes valeurs de k_{ex} et P^* comparé à la protéine Apo à 293 K. A droite, (données pour HSP90-NTD à 293 K, N=21) la figure illustre que tous les résidus ajustés globalement appartiennent au même processus d'échange. Les valeurs k_{ex} et P^* des ajustements globaux sont toutes comprises entre les valeurs moyennes de k_{ex} et $P^* \pm 3$ fois la déviation standard.

Finalement, tous les résidus appartenant au même processus d'échange peuvent être ajustés ensemble (Figure 3.15) et l'on peut ainsi obtenir des informations sur la cinétique, la thermodynamique et les valeurs absolues de différence de déplacement

chimique entre l'état fondamental et celui excité. Afin d'obtenir une estimation de la précision des résultats obtenus, 40 répétitions de Monte-Carlo ont été effectuées. Ainsi j'ai pu déterminer avec un ajustement global incluant 21 résidus que les conformations (ouverte et fermée) s'interconvertaient avec une constante de vitesse d'échange conformationnel $k_{ex} = k_1 + k_{-1} = 2490 \pm 61 \text{ s}^{-1}$ et que la population de l'espèce minoritaire était de $3.2 \pm 0.2 \%$ à 293 K.



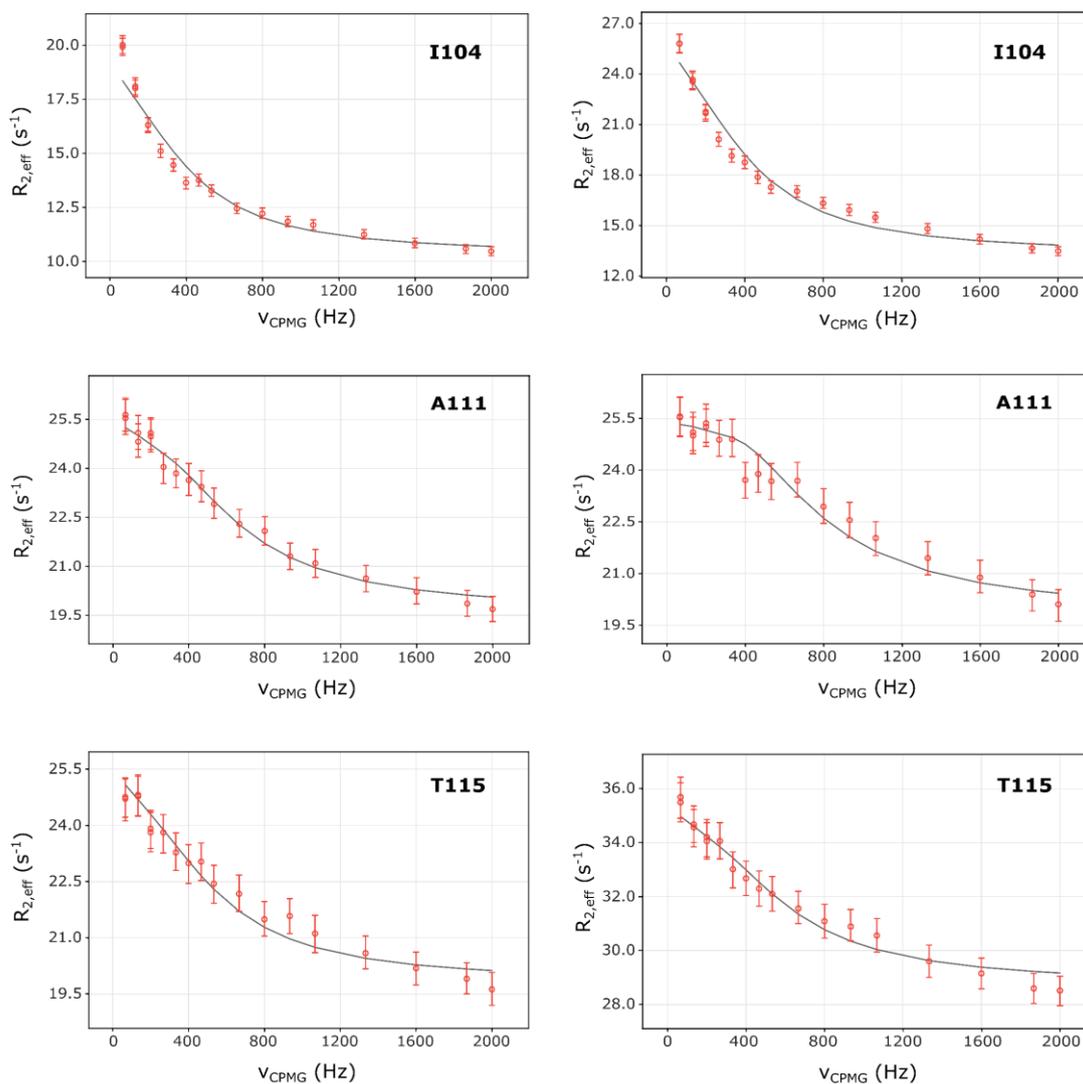


Figure 3.15 : Ajustement global des données de relaxation dispersion pour HSP90-NTD à 293 K comprenant 21 résidus. Les graphiques représentant R_2^{eff} en fonction de la fréquence CPMG (V_{CPMG}) montrent les ajustements globaux obtenus pour 7 des 21 résidus ajustés globalement. La colonne de gauche montre les ajustements obtenus pour les données acquises à 700 MHz et la colonne de droite pour les données acquises à 850 MHz.

3.4.2.7 Analyse critique des erreurs

De ces résultats, j'ai pu conclure que la précision sur les valeurs globales de k_{ex} et P^* était de l'ordre de 2.5 % et 6.3 % respectivement. Concernant les valeurs absolues des différences de déplacements chimiques entre l'état fondamental et celui excité, la précision que j'ai obtenue était en moyenne de 6 % pour les $|\overline{\Delta\omega_C}|$ et de 53 % pour les

$|\overline{\Delta\omega_H}|$. Cette observation, d'une très forte imprécision sur la valeur de $|\overline{\Delta\omega_H}|$, est cohérente avec les observations faites Figure 3.12 où $|\overline{\Delta\omega_H}|$ semblait avoir une influence moindre sur l'ajustement comparé à $|\overline{\Delta\omega_C}|$. Si les valeurs de k_{ex} et P^* sont des moyennes pour l'ensemble des résidus ajustés globalement (i.e. 2 valeurs : k_{ex} et P^* pour 21 résidus ajustés globalement soit pour 42 courbes dans le cas de HSP90-NTD à 293K), les valeurs absolues de différence de déplacements chimiques $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$ ne sont obtenues que lors de l'ajustement d'un seul résidu (i.e. 2 valeurs, $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$, pour un résidu donc pour 2 courbes dans les cas de HSP90-NTD à 293 K).

J'ai donc conclu de ces analyses que mon estimation de la précision pour les valeurs de k_{ex} et P^* pouvait être fiable tandis que celle obtenue pour $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$ devait être prise avec précaution. En faisant l'hypothèse que les conformations de l'état fondamental et de l'état excité seraient identiques sur une plage de température de 20°C, j'ai assumé que sur cette dernière les valeurs de $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$ seraient constantes, les déplacements chimiques des groupements méthyles variant peu avec la température, et que les valeurs de k_{ex} et P^* pourraient être influencées par la température. J'ai ensuite acquis des expériences de CPMG à quatre autres températures (278 K, 283 K, 288 K et 298 K). De la même manière que pour les données de CPMG-RD analysées à 293 K, un ajustement global des résidus appartenant au même processus d'échange a été réalisé pour chaque température. Les valeurs absolues des déplacements chimiques entre l'état fondamental et l'état excité ont pu être obtenues et la précision sur les valeurs absolues des déplacements chimiques était de l'ordre de 24 % pour $|\overline{\Delta\omega_C}|$ et 46 % pour $|\overline{\Delta\omega_H}|$. Avec de telles précisions pour les valeurs de $|\overline{\Delta\omega_C}|$ et $|\overline{\Delta\omega_H}|$, j'ai choisi pour la suite de ne considérer comme fiable uniquement les valeurs de k_{ex} et P^* pour mon analyse.

3.4.3 Identification de l'état excité

3.4.3.1 Corrélation entre structures et observation des NOEs

En utilisant uniquement les données des expériences de CPMG relaxation-dispersion il n'est pas possible d'identifier quel est l'état fondamental et quel est l'état excité entre les deux ensembles conformationnels obtenus précédemment (i.e. : l'« état ouvert » et l'« état fermé »). Pour tenter d'attribuer l'état excité à l'un ou l'autre des ensembles conformationnels j'ai utilisé les pics de corrélation NOEs en faisant l'hypothèse que les groupements méthyles étaient homogènement répartis dans la protéine et que si un état était peuplé pendant X % du temps alors l'intensité des NOEs observés pour cet état serait multipliée par ce même pourcentage. Une analyse attentive des ensembles de structures avec le segment [98-136] en position ouverte et fermée m'a permis d'identifier deux ensembles de paires de groupements méthyles proches dans l'espace

(< 11.5 Å) et présents dans un seul et unique ensemble de structures (ouvert ou fermé). Ainsi, 28 paires de groupements méthyles sont spécifiques de l' « état fermé » tandis que 95 sont caractéristiques de l' « état ouvert ». Ces paires de groupements méthyles correspondent aux NOEs interméthyles hypothétiques spécifiques à chacun des deux ensembles échantillonnés de façon transitoire par HSP90-NTD α . En étudiant minutieusement les spectres NOESY de chacun des deux mutants, R60A et R46A, stabilisant notre protéine d'intérêt soit avec le segment [98-136] respectivement en position ouverte ou en position fermée, j'ai constaté qu'expérimentalement on détectait 2 fois moins de NOEs attendus pour l' « état fermé » (mutant R46A) par rapport à l' « état ouvert » (mutant R60A) (Figure 3.16).

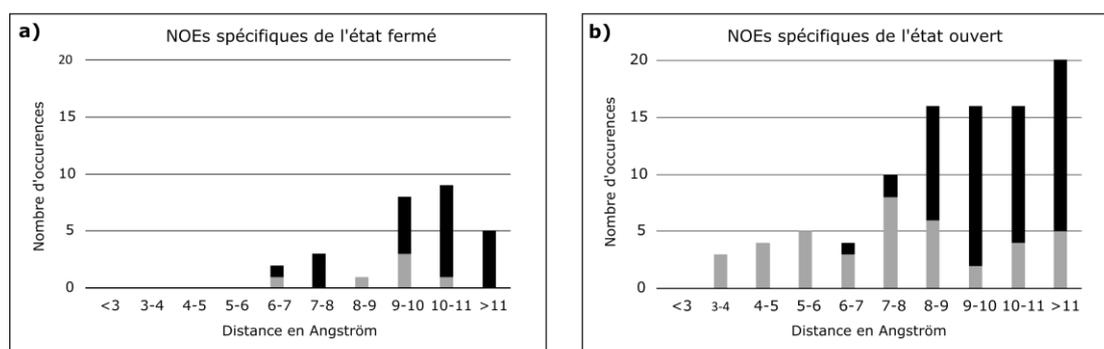


Figure 3.16 : Histogrammes illustrant, en fonction de la distance, le nombre de paires de groupements méthyles proches dans l'espace identifiées à partir des structures pour a) uniquement l' « état fermé » et b) uniquement l' « état ouvert ». En gris sont représentés les NOEs entre deux groupements méthyles observés sur les spectres des expériences NOESY pour respectivement a) R46A-HSP90-NTD et b) R60A-HSP90-NTD. En noir sont représentés les NOEs hypothétiques identifiés proches dans l'espace (< 11.5 Å) mais non détectés dans les spectres NOESY.

Un échantillonnage de la conformation fermée pour 20 % du temps donnerait un résultat similaire en termes de réduction du nombre de NOEs observés. En prenant en compte que la mutation R46A augmente d'environ 2.5 fois l'intensité moyenne des pics de corrélation NOE caractéristiques de l'état fermé par rapport à l'état ouvert (Figure 3.7) j'ai pu estimer à environ 8 % la population de l'état fermé de HSP90-NTD-WT. Ces analyses sont très qualitatives car elles comprennent de nombreuses approximations mais toutefois elles montrent clairement que l'ensemble conformationnel avec le segment [98-136] en position ouverte est l'état majoritaire tandis que la conformation avec le segment [98-136] en position fermée est l'état minoritaire.

3.4.3.2 Dépendance de la température

Additionnellement, j'ai acquis une expérience NOESY à basse température (278 K) en plus de celle enregistrée au préalable à 298 K pour HSP90-NTD-WT. Mon but en acquérant cette expérience était de confirmer, après une première analyse qualitative, que l'« état ouvert » était bien l'état majoritaire. De plus, j'ai cherché à prouver que ce qui était observé avec les spectres NOESY correspondait à l'échange conformationnel observé par CPMG. En effet, le réarrangement structural du segment [98-136] observé est un mouvement d'amplitude de plus de 30 Å et donc est manifestement plus complexe qu'une simple transition entre deux états comme utilisée pour quantifier la constante de vitesse d'échange conformationnel et la population de l'état excité par CPMG. Aussi, la question se posait de savoir si l'échange conformationnel observé par CPMG correspondait à ce qui était observé dans les spectres NOESY ou si un autre phénomène, totalement différent de celui observé par les expériences NOESY était étudié.

Avec les expériences de relaxation dispersion effectuées à cinq températures différentes j'ai pu observer que les populations de l'état excité varient en fonction de la température. Entre 293 K et 278 K, plus la température diminue plus la population de l'état excité augmente passant de 3.2 ± 0.1 % à 293 K à 5.1 ± 0.3 % à 278 K (Figure 3.17a/b). A 298 K, la population de l'état excité augmente à nouveau comparé à la valeur obtenue à 293 K, mais reste inférieure à celle observée à 278 et 283 K. Avec les expériences NOESY j'ai pu constater qu'à basse température (278 K) le rapport de l'intensité des pics caractéristiques de l'état fermé sur celle des pics spécifiques de l'état ouvert augmentait (Figure 3.17c/d) par rapport à celui observé à 298 K. Ainsi, les phénomènes observés avec les expériences de CPMG et NOESY corrélaient et confirment les résultats de l'analyse qualitative précédente révélant que l'« état ouvert » est l'état majoritaire.

Finalement, les deux analyses indépendantes décrites en 3.4.3.1 et 3.4.3.2 sont également en accord avec les données de cristallographie aux rayons X montrant que la totalité des structures du domaine N-terminal de HSP90-NTD reportées se trouvent avec le segment [98-136] en position ouverte. De plus, des simulations de dynamique moléculaire (détaillées dans l'article III présenté à la fin de ce chapitre) établissent que la conformation avec le segment [98-136] en position ouverte est plus stable que la conformation excitée métastable avec le segment [98-136] en position fermée. Ainsi, l'« état ouvert » est l'état majoritaire/fondamental et l'« état fermé » est l'état minoritaire/excité.

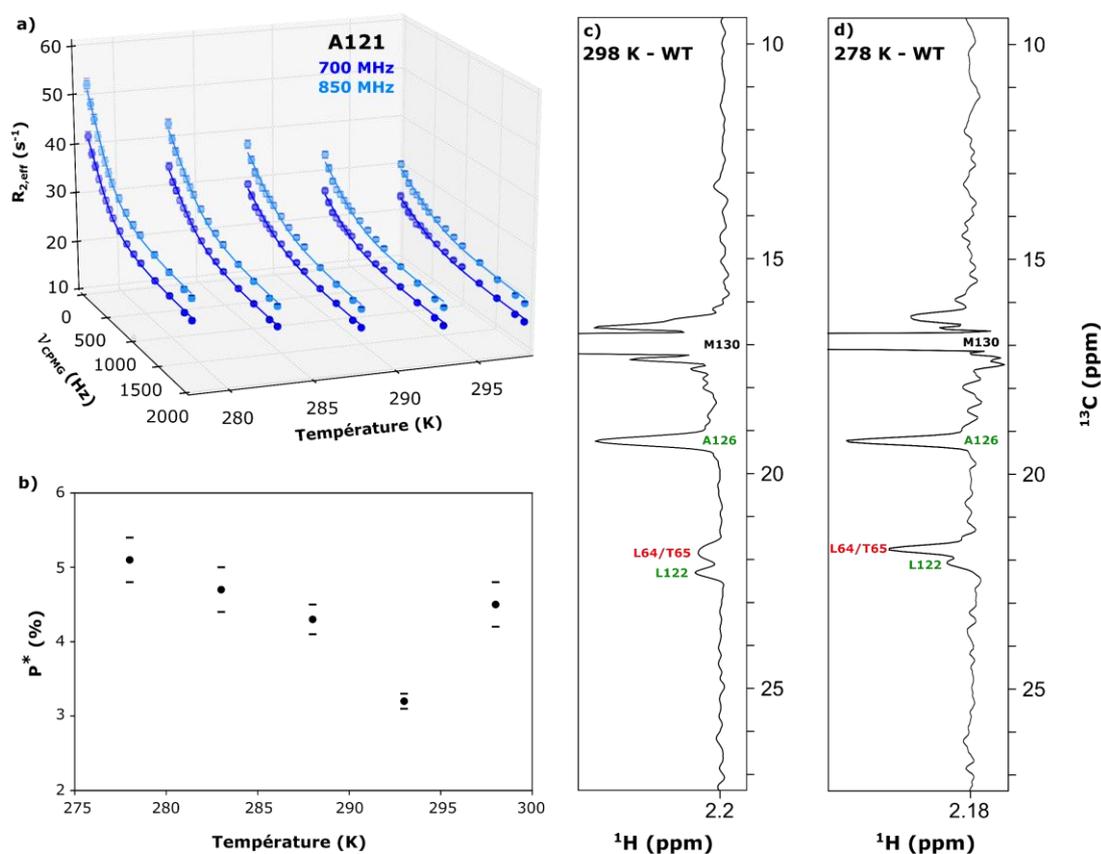


Figure 3.17 : a) Profils 3D de $^{13}\text{CH}_3$ -MQ CPMG relaxation-dispersion pour le groupement méthyle Ala-121 acquis à cinq différentes températures (de 278 K à 298 K). Les points expérimentaux sont indiqués par des cercles et ont été ajustés en employant un modèle d'échange à deux états. La précision sur les données a été estimée en utilisant une analyse Monte-Carlo. b) Graphique représentant l'évolution de la population de l'état excité en fonction de la température obtenue avec les expériences de CPMG-RD. Les barres d'erreur proviennent d'analyses Monte Carlo c)/d) Traces 1D extraites des expériences 3D CCH HMQC-NOESY-HMQC montrant M130 en pic diagonal ainsi que ses pics de corrélation NOEs associés pour la protéine sauvage HSP90-NTD (WT) à 298 K (c) et à 278 K (d). Les pics de corrélation NOEs en rouge représentent des pics caractéristiques de l'état fermé, ceux en vert correspondent à des pics intra segment [98-136] et donc non caractéristiques de l'état ouvert ni de l'état fermé. Les spectres NOESY ont été acquis en 2 à 3 jours sur des spectromètres opérant à des fréquences proton de 850 ou 950 MHz. Les temps de mélanges NOE sont respectivement de 400 et 240 ms pour prendre en compte l'influence de la température sur le τ_c .

3.4.4 Caractérisation thermodynamique et cinétique de l'échange

Les expériences de CPMG - relaxation dispersion, m'ont permis d'obtenir des informations sur la cinétique et la thermodynamique de l'échange entre l'état majoritaire (« état ouvert ») et l'état minoritaire (« état fermé »). Comme mentionné précédemment, avec un ajustement global incluant 21 résidus, les conformations ouvertes et fermées de HSP90-NTD Apo s'interconvertissent à une vitesse caractéristique $k_{ex} = k_1 + k_{-1} = 2490 \pm 61 \text{ s}^{-1}$ et la population de l'espèce minoritaire est de $3.2 \pm 0.2 \%$ à 293 K. L'énergie libre entre l'état fondamental et excité ainsi que l'énergie d'activation ont pu être déterminées en utilisant l'équation de Van't Hoff (Equation 3.4) et l'équation de Eyring (Equation 3.5) et sont respectivement de $8.3 \pm 0.2 \text{ kJ.mol}^{-1}$ et $61.0 \pm 0.2 \text{ kJ.mol}^{-1}$ à 293 K (Figure 3.18).

$$\Delta G = -RT \ln (K_{eq}) = -RT \ln \left(\frac{P^*}{1-P^*} \right) \quad (\text{Equation 3.4})$$

Avec ΔG l'énergie libre entre l'état fondamental et l'état excité en J.mol^{-1} , R la constante des gaz parfaits en $\text{J.mol}^{-1}.\text{K}^{-1}$, T la température en K, K_{eq} la constante d'équilibre et P^* la population de l'état excité en %.

$$k_1 = \frac{k_B T \kappa}{h} e^{\frac{-\Delta G^*}{RT}} \quad (\text{Equation 3.5})$$

Avec k_1 la constante de vitesse de l'état fondamental à l'état excité en s^{-1} , ΔG^* l'énergie d'activation entre l'état fondamental et l'état de transition en J.mol^{-1} , R la constante des gaz parfaits en $\text{J.mol}^{-1}.\text{K}^{-1}$, T la température en K, h la constante de Planck en J.s , k_B la constante de Boltzmann en J.K^{-1} et κ le coefficient de transmission, supposé égal à 1.

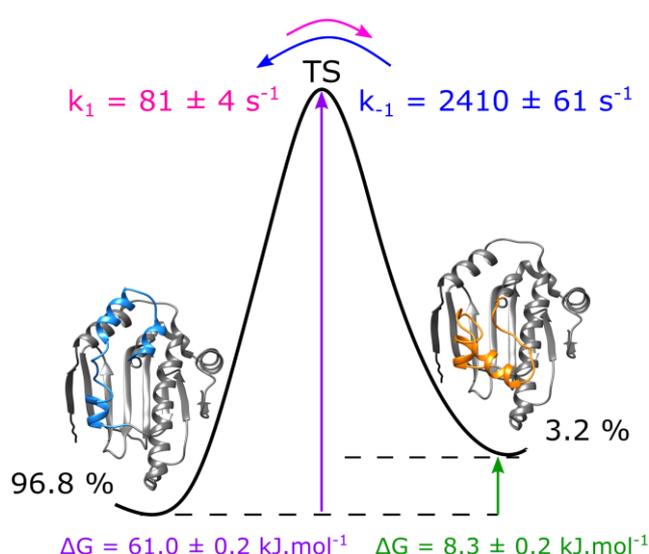


Figure 3.18 : Diagramme schématisé du paysage énergétique caractérisant l'échange entre l'état ouvert (fondamental) et l'état fermé (excité) de HSP90-NTD α . Les paramètres cinétiques et thermodynamiques de cet échange, extraits à 293 K, sont reportés.

L'ordre de grandeur de nos valeurs d'énergie libre et d'énergie d'activation semble admissible. J'ai toutefois cherché à le comparer avec ce qui avait été reporté dans la littérature. Vallurupalli et al. (Vallurupalli et al. 2016) ont reporté une étude sur la dynamique du lysozyme T4 L99A effectuée avec des expériences de CPMG-RD et Xie et al. (Xie et al. 2020) ont reporté une étude sur la dynamique de la kinase ABL effectuée avec des expériences de CEST. Dans la première étude, Vallurupalli et al. ont une constante de vitesse d'échange conformationnel de 39000 s^{-1} et une population d'état excité de 2.3 % à 310 K. Ces valeurs ont été obtenues par extrapolation, en utilisant l'équation d'Arrhenius, des paramètres (k_{-1} , k_1) extraits des ajustements des courbes de relaxation dispersion à différentes températures [274 K – 293.4 K]. Ils montrent que l'énergie libre entre l'état fondamental et l'état excité est de $9.6 \pm 0.2 \text{ kJ.mol}^{-1}$ et que l'énergie d'activation entre l'état fondamental et l'état de transition est de $25.3 \pm 1.3 \text{ kJ.mol}^{-1}$ à la même température. Dans la seconde étude Xie et al. ont une constante de vitesse d'échange conformationnel de 46.8 s^{-1} entre l'état fondamental et le premier état excité et une population de 6 % pour le premier état excité. Ils décrivent une énergie libre d'environ 5.0 kJ.mol^{-1} entre l'état fondamental et le premier état excité et une énergie d'activation d'environ 152 kJ.mol^{-1} entre l'état fondamental et l'état de transition. Les valeurs de constante de vitesse d'échange conformationnel et de population d'état excité obtenues pour HSP90-NTD se situent entre celles obtenues par Vallurupalli et al. et Xie et al. Il en est de même pour la valeur d'énergie libre entre l'état fondamental et l'état excité et la valeur d'énergie d'activation entre l'état fondamental et l'état de transition de HSP90-NTD. Aussi, ces deux études semblent confirmer la crédibilité des énergies mesurées pour HSP90-NTD.

3.4.5 Extraction des valeurs d'enthalpie et d'entropie entre l'état fondamental et excité

En utilisant les populations des états peuplés de façon transitoire aux différentes températures il est possible, en faisant l'hypothèse que la contribution enthalpique (ΔH) et la contribution entropique (ΔS) à l'énergie libre de Gibbs (ΔG , avec $\Delta G = \Delta H - T\Delta S$) ne dépendent pas de la température, d'obtenir des informations sur les valeurs de ΔH et $-T\Delta S$ entre l'état fondamental (« état ouvert » de HSP90-NTD) et l'état excité (« état fermé » de HSP90-NTD) en utilisant l'équation de Van't Hoff (Equation 3.6). En effet, en effectuant un ajustement linéaire de la dépendance en $1/T$ de $\ln(K_{eq})$ on peut obtenir $-\Delta H/R$ avec la pente de la droite et $\Delta S/R$ avec l'ordonnée à l'origine et ainsi en déduire ΔH et ΔS .

$$\ln(K_{eq}) = \ln\left(\frac{P^*}{1-P^*}\right) = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \quad (\text{Equation 3.6})$$

Avec ΔH la contribution enthalpique en $J.mol^{-1}$, ΔS la contribution entropique en $J.mol^{-1}.K^{-1}$ entre l'état fondamental et l'état excité, R la constante des gaz parfaits en $J.mol^{-1}.K^{-1}$, T la température en K, K_{eq} la constante d'équilibre et P^* la population de l'état excité en %.

Ainsi l'« état fermé » est à la fois stabilisé enthalpiquement ($\Delta H = -7.9 \pm 2.6 \text{ kJ.mol}^{-1}$) et déstabilisé entropiquement ($-T\Delta S = 15.6 \pm 2.6 \text{ kJ.mol}^{-1}$ à 293 K) par rapport à l'état ouvert (Figure 3.19). De tels résultats pour ΔH indiquent que le système n'incluant pas uniquement la protéine mais également le solvant autour d'elle forme plus d'interactions stabilisantes dans l'état excité. Par exemple la création de liaisons hydrogène ou d'interactions Van der Waals et la solvation de groupements polaires participent à une contribution enthalpique favorable. La composante entropique quant à elle décrit le changement de désordre dans le système, incluant toujours la protéine et le solvant. Ainsi, la composante entropique ($-T\Delta S$) positive montre que le système est moins dynamique dans l'état excité. Il est à noter que la stabilisation enthalpique n'est pas assez importante pour compenser la pénalité entropique globale.

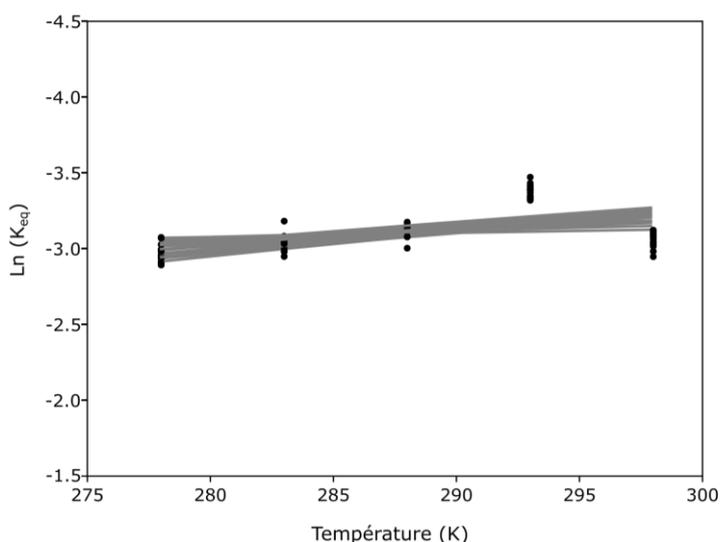


Figure 3.19 : Représentation graphique de 20 ajustements Monte Carlo de $\ln(K_{eq}) = \ln\left(\frac{P^*}{1-P^*}\right)$ en fonction de la température. Le but étant d'extraire les contributions enthalpiques et entropiques entre l'état ouvert et l'état fermé. P^* correspond à la population ajustée de l'état fermé obtenue en utilisant les données de relaxation-dispersion acquises à chacune des cinq différentes températures.

L'ordre de grandeur des valeurs obtenues a été comparé à ceux trouvés dans la littérature. Il apparaît que lysozyme T4 L99A est stabilisé enthalpiquement ($\Delta H = -15 \pm 3 \text{ kJ.mol}^{-1}$) et déstabilisé entropiquement ($-T\Delta S = 24.5 \pm 3.4 \text{ kJ.mol}^{-1}$ à 310 K) et les données obtenues par Vallurupalli et al. (Vallurupalli et al. 2016) sont du même ordre de grandeur que celles obtenues dans notre étude. Cependant, ayant formulé deux hypothèses pour arriver à ces conclusions (*i.e.* que l'échange entre l'« état ouvert » et

l'« état fermé » pouvait être représenté par un modèle à deux états et que la contribution enthalpique (ΔH) et la contribution entropique (ΔS) à l'énergie libre de Gibbs (ΔG) ne dépendent pas de la température) je pense qu'interpréter les valeurs chiffrées obtenues pour ΔH et $-\Delta S$ n'est pas pertinent, et uniquement le signe de ces dernières peut être qualitativement interprété, comme décrit ci-dessus. Il est facile de penser que l'erreur obtenue sur ΔH devrait être plus petite que celle obtenue pour $-\Delta S$, en effet ΔH est obtenue avec la pente de l'ajustement tandis que la valeur de $-\Delta S$ est obtenue avec l'ordonnée à l'origine de l'ajustement. Ainsi, une propagation de l'erreur importante pour arriver à 0 K et à l'ordonnée à l'origine serait responsable de la plus grande erreur sur la valeur de $-\Delta S$. En réalité lorsque l'on regarde les erreurs obtenues pour ΔH et $-\Delta S$, ces dernières sont similaires. D'après Lente et al. (Lente et al. 2005) cela peut s'expliquer par le fait que l'on peut également tracer l'équation 3.7 et obtenir les valeurs de ΔH avec l'ordonnée à l'origine et de $-\Delta S$ avec la pente.

$$T \ln (K_{eq}) = \ln \left(\frac{P^*}{1-P^*} \right) = \frac{\Delta S}{R} T - \frac{\Delta H}{R} \quad (\text{Equation 3.7})$$

Toutefois, les valeurs d'enthalpie et d'entropie entre l'état fondamental et excité n'étant pas extraites avec une précision suffisante, le modèle reposant sur des hypothèses conséquentes, ces dernières ne sont pas reportées dans l'article ci-dessous décrivant les principaux résultats de ce chapitre.

3.5 Article III

L'article III présenté ci-dessous (en révision pour Nature Communications), où je suis co-première auteure avec Elisa Rioual, intègre les résultats des expériences et analyses introduites dans ce chapitre. Additionnement, des simulations de dynamique moléculaire ont été effectuées pour corroborer les résultats obtenus par RMN et sont également présentées dans cet article. Dans ce travail, intégrant diverses méthodes d'analyses, j'ai préparé tous les échantillons, acquis une partie des spectres RMN et effectué l'analyse des structures de la PDB. De plus, j'ai analysé les expériences de CPMG-RD et en partie les expériences NOESY, principalement lors du co-encadrement d'Elisa Rioual pendant son stage de M2. J'ai également participé à la rédaction du manuscrit.

Article III « Visualizing the Transiently Populated Closed-State of Human HSP90 ATP Binding Domain »

Visualizing the Transiently Populated Closed-State of Human HSP90 ATP Binding Domain

Faustine Henot^{1,#}, Elisa Rioual^{1,2,#}, Adrien Favier¹, Pavel Macek^{1,3}, Elodie Crublet³, Pierre Josso², Bernhard Brutscher¹, Matthias Frech⁴, Pierre Gans¹, Claire Loison^{2,*}, Jerome Boisbouvier^{1,*}

Affiliation:

1. Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, avenue des martyrs, F-38044 Grenoble, France.

2. Institut Lumière Matière, University of Lyon, Université Claude Bernard Lyon 1, CNRS, F-69622, Villeurbanne, France.

3. NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.

4. Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany.

These authors contributed equally to this work

* correspondence to be addressed to: claire.loison@univ-lyon1.fr and jerome.boisbouvier@ibs.fr

Abstract:

HSP90 are abundant molecular chaperones, assisting the folding of several hundred client proteins, including substrates involved in tumor growth or neurodegenerative diseases. A complex set of large ATP-driven structural changes occurs during HSP90 functional cycle. However, the existence of such structural rearrangements in apo HSP90 has remained unclear. Here, we identified a metastable excited state in the isolated HSP90 ATP binding domain. We used solution NMR and mutagenesis to characterize structures of both ground and excited states. We demonstrated that in solution the HSP90 ATP binding domain transiently samples a functionally relevant ATP-lid closed state, distant by more than 30 Å from the ground state. NMR relaxation and molecular dynamics were combined to characterize the energy landscape corresponding to the transition between these interconverting states. The precise description of the dynamics and structures sampled by human HSP90 ATP binding domain is a paramount piece of information for the future design of new therapeutic ligands.

Keywords: HSP90, ATP-lid, Structure, Dynamics, NMR, Molecular Dynamics, Excited state

Introduction

Heat Shock Protein 90 (HSP90) is a ubiquitous ATP-dependant molecular chaperone involved in the folding of a plethora of client proteins, also modulating their cellular activities. HSP90's activity is itself regulated by post-translational modifications such as phosphorylation, nitration, methylation or acetylation¹ as well as a large number of different co-chaperones²⁻⁷, including Hop, PP5, p23, Sgt1, FKBP51/52, Cyp40, NudC and Cdc37. During its functional cycle, in order to fold client proteins, the chaperone HSP90 undergoes complex structural rearrangements driven by both ATP binding and hydrolysis^{8,9}. Among the identified HSP90-client proteins, a large fraction is related to cancer such as steroid hormone receptors, tumor suppressor p53, telomerase, hypoxia-inducible factor 1 α and kinases¹⁰. Furthermore, high cellular HSP90 expression levels are often associated with poor prognoses in many cancer types¹¹. Therefore, human HSP90 has been identified as a major anti-cancer drug target. There are four HSP90 homologues in human cells: GRP94, TRAP1, HSP90 α and HSP90 β ¹². While GRP94 and TRAP1 are respectively found in endoplasmic reticulum¹³ or mitochondria¹⁴, respectively, both HSP90 α and HSP90 β isoforms are highly abundant in the cytoplasm, representing 1 to 2 % of cellular proteins¹⁵⁻¹⁷. Under stress conditions or in tumor cells, the level of expression of HSP90 increases up to 7% of the total amount of expressed proteins¹⁰.

HSP90 is a homodimeric chaperone of *ca.* 170 kDa. Each monomer is composed of 3 domains: C-terminal

domain (CTD), middle domain, and N-terminal domain (NTD). The 13 kDa CTD is responsible for HSP90 dimerization, and harbours a MEEVD motif mediating the binding of various co-chaperones with tetratricopeptide repeat (TPR) domains¹⁸⁻²⁰. The middle domain (*ca.* 40 kDa) is involved in binding of both client proteins and co-chaperones and, additionally, modulates hydrolysis of ATP bound to the NTD^{3,21,22}. A long charged linker connects the middle domain and the NTD and modulates molecular interactions with client proteins²³. Finally, the NTD (*ca.* 25 kDa) is implicated in client protein and co-chaperone binding³. HSP90-NTD possesses an ATP binding site with an unusual structure named Bergerat fold²⁴. The particular environment of the ATP binding site offers the possibility to develop HSP90-specific inhibitors that are not affecting the activity of most other ATP-binding proteins. ATP binding and hydrolysis in the NTD domain drive the structural rearrangements of HSP90 during its functional cycle²⁵⁻²⁸. Therefore, the HSP90 ATP binding site is the target of most therapeutic ligands developed so far, and more than 300 crystal structures of HSP90-NTD, bound to different ligands, are available in the PDB. A particular challenge for the design of new inhibitors against this important cancer target is the presence of a highly flexible ATP-lid segment that can cover the nucleotide/drug binding site in HSP90^{29,30}.

In this article, we report on an atomic-resolution structural and dynamics investigation of HSP90-NTD conformations sampled in solution. We demonstrate that the ATP-lid segment of this crucial human chaperone, in addition

to the well characterized open state, also transiently populates a closed conformation that requires structural rearrangement of peptide segments over a distance of up to 30 Å. Using solution NMR spectroscopy and molecular dynamics (MD) simulation, we could determine atomic-resolution models of the open- and closed-state conformations, as well as derive kinetic and thermodynamic information for this structural rearrangement occurring in solution. This is the first time that such an excited closed-state structure is reported for isolated apo HSP90-NTD. Our study reveals that the closure of the ATP-lid, observed during the ATP-driven functional cycle of HSP90, is already sampled in the human chaperone before binding of ATP. We anticipate that our results will be important for the future drug design of new inhibitors against this challenging drug target.

Results

Conformational variability of HSP90 α -NTD. HSP90 has been identified as a major therapeutic target, especially against cancer^{2,31,32}. Since the late 90's a large number of competitive inhibitors, targeting the ATP binding site of the HSP90 α -NTD has entered clinical trials³³⁻³⁵. In the context of these drug development efforts, more than 300 atomic resolution structures of isolated HSP90 α -NTD in presence of various ligands have been determined, using X-ray crystallography. These structures can be separated into 8 main groups (Fig. 1a-c) showing distinct conformational properties. Pairwise superimposition of the HSP90 α -NTD centroid structures representing each cluster reveals that the

main structural differences are located on the segment covering the nucleotide/drug binding site. While large parts of these 8 centroids superimpose with an average root-mean-square deviation (RMSD) of *c.a.* 0.2 Å, the ATP-lid²⁹ (defined here from residues M98 to V136 including three helical segments) shows higher heterogeneity with an average pairwise RMSD of 1.4 Å (Fig 1d). In all the apo and nucleotide- or ligand-bound structures of the isolated HSP90 α -NTD, the ATP-lid is in the so-called open state²⁷ and does not cover the ATPase site and corresponding drug binding site. The observed structural variability of HSP90 α -NTD structures indicates that the ATP-lid, in the open state, can adopt various conformations in a crystalline environment depending on the crystallization conditions, and the stabilization by ligands or additives. However, a detailed picture of the conformational properties of the ATP-lid segment in solution is of utmost importance for future drug design, as this segment is located next to the major drug binding site of human HSP90. Therefore, we initiated a structural and dynamic investigation of HSP90-NTD using solution NMR spectroscopy to characterize at atomic resolution the conformations sampled by the ATP-lid.

The ATP-lid of HSP90-NTD populates two distinct conformations in solution.

Previous NMR investigations reported that human HSP90 α -NTD amide backbone NMR signals from N105 to K116 in the ATP-lid segment were undetectable³⁶⁻³⁸ (Fig 1d).

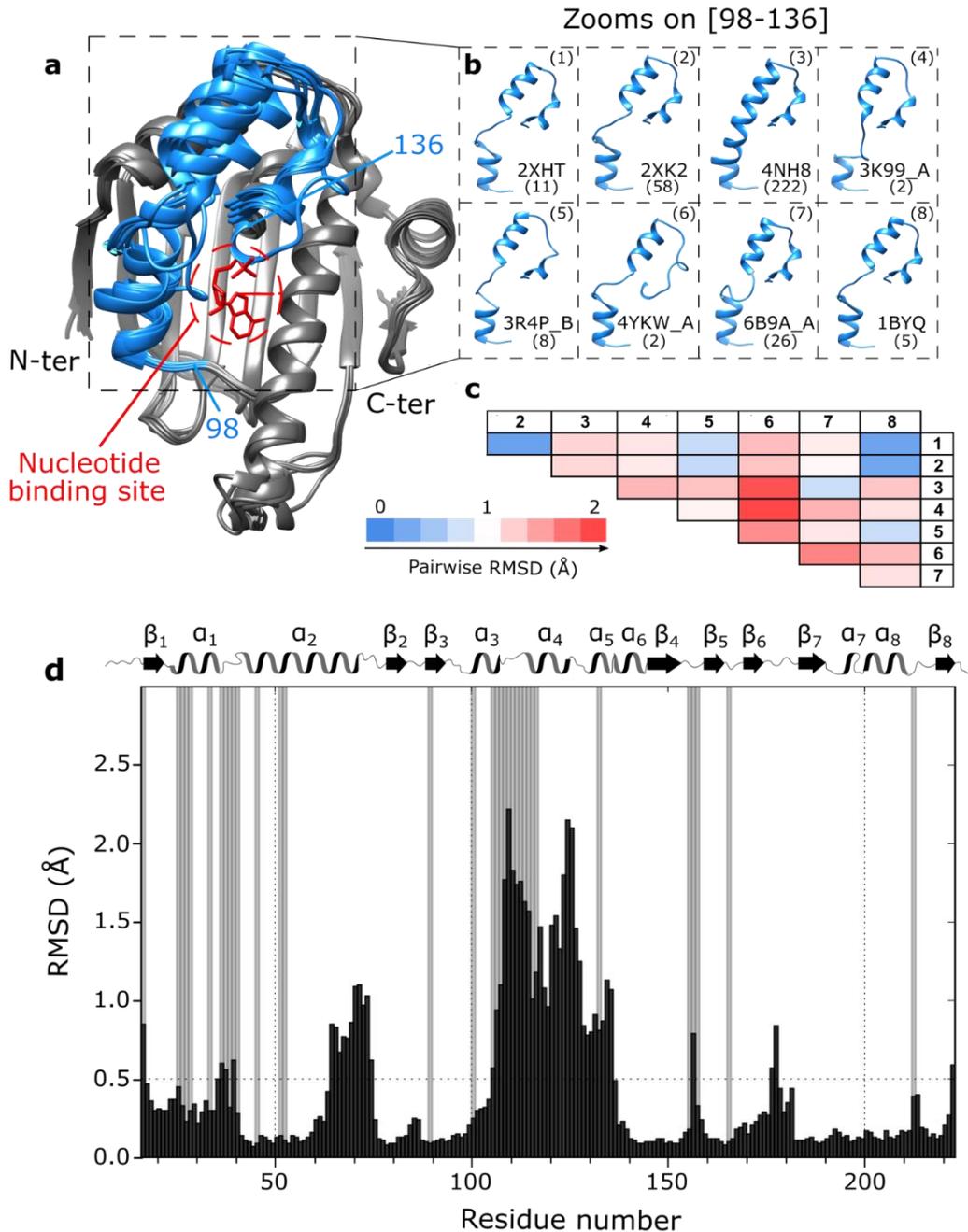


Fig. 1: Analysis of available structures of human HSP90 α -NTD. **a)** Superimposition of 8 centroids representing the 8 clusters describing the 334 structures of isolated HSP90 α -NTD available in the Protein Data Bank (on January the 5th of 2021). Clustering was performed using the MaxCluster program (<http://www.sbg.bio.ic.ac.uk/maxcluster>) with an Average Linkage type of hierarchical clustering and a threshold value of 1.05. In blue is depicted the segment [98-136] and in red the nucleotide. **b)** Zooms on the segment [98-136] for all the 8 centroids superimposed in a). PDB ID and number of structures present in each cluster are

disclosed next to each centroid. **c)** Table representing pairwise RMSD in Å between C α backbone atoms of each pair of the 8 centroids. (1: 2XHT, 2: 2XK2, 3: 4NH8, 4: 3K99_A, 5: 3R4P_B, 6: 4YKW_A, 7: 6B9A_A, 8: 1BYQ). Going from blue to red the RMSD values increase. **d)** Histogram of the averaged pairwise RMSD between C α backbone atoms of the 8 centroids in Å as a function of the residue number (black). Grey bars represent non-assigned backbone residues³⁸. On top of the histogram: secondary structure elements as a function of the residue number (α for helices and β for sheets).

This observation suggests that the ATP-lid is undergoing conformational exchange on the micro- to millisecond time scale. The broadening of backbone NMR signals in the ATP-lid is also particularly detrimental for NMR-distance restraint based structure determination using uniformly ^{13}C , ^{15}N -labeled HSP90 α -NTD sample. In order to overcome this problem, we decided to produce perdeuterated HSP90 α -NTD, specifically $^{13}\text{CH}_3$ -labeled on the A $^{\beta}$, I $^{\delta 1}$, L $^{\delta 2}$, M $^{\epsilon}$, T $^{\gamma}$ and V $^{\gamma 2}$ methyl positions³⁹. The resulting 87 ^1H , ^{13}C -labeled methyl groups all gave rise to detectable NMR signals. Most interestingly, 5 of these methyl groups (L107, T109, I110, A111 and T115) belong to the previously unassigned stretch of residues in the ATP-lid (Supporting Fig.1). Unambiguous sequence-specific ^1H and ^{13}C resonance assignments of all detected methyl groups was obtained from a set of through-bond correlation experiments³⁸, aided by an extensive mutagenesis-driven assignment of 33 methyl probes⁴⁰ (Supporting Fig.2).

Next, we recorded methyl NOESY spectra to derive a set of inter-methyl distance restraints. The applied labeling scheme that yields protonation only on the methyl moieties at the extremity of hydrophobic side chains is particularly useful to extract long-range distance restraints between methyl groups that are separated by up to 10 Å⁴¹. For HSP90 α -NTD, 597 inter-methyl NOEs (Supporting Fig 3A and Table 1) could be detected, from which 111 distance restraints were derived that involve at least one methyl group belonging to the ATP-lid segment. All these structurally meaningful restraints were assigned exclusively based

on chemical shifts. The inter-methyl distance restraints were complemented by 41 short-range backbone H $_N$ -H $_N$ distance restraints (Supporting Fig 2A and Table 1) derived from a 3D ^{15}N -edited NOESY spectrum, and 54 backbone ϕ, ψ -dihedral restraints derived from characteristic backbone chemical shifts^{38,42}.

Based on this set of experimentally derived structural restraints, we have performed structure calculations using a simulated annealing protocol in torsion angle space⁴³ to determine the structure and position of the ATP-lid segment relative the rigid HSP90 α -NTD scaffold comprised by the two segments [11-97] and [137-223]). To our surprise, the calculated structures showed a large number of distance violations (25) and a high final target energy (74 kJ.mol⁻¹). Careful analysis of the reported distance violations revealed a number of NMR-derived restraints that are incompatible with a single structure (Fig. 1a). As an example, a pair of inter-methyl NOEs is detected between the methyl group of I131- δ_1 located in the ATP-lid and the methyl moiety of L29- δ_2 in helix-1, while another set of NOE correlation peaks indicate that the preceding residue M130- ϵ in the ATP-lid is spatially close to L64- δ_2 and T65- γ , two methyl groups located at the end of the long helix-2 (Fig. 2ab). However, according to the previously determined X-ray structures of HSP90-NTD (Fig. 1), L29 is more than 30 Å away from L64 and T65. The specific assignment of these methyl groups, the analysis of the through-bond sequence specific assignments of backbone nuclei and the transfer to methyl-moieties were double checked³⁸ and completed by an

extensive mutagenesis driven assignment of 23 additional key methyl probes⁴⁰ (Supporting Fig.2) allowing us to exclude misassignment of these NOEs. As the NMR-derived distance restraints could not be satisfied with a single structure, we made the hypothesis that the ATP-lid populates several conformations.

Therefore, an additional round of structure calculation was performed⁴⁴ assuming two co-existing conformations of the ATP-lid segment (Fig. 2c), and allowing each experimental distance restraint to be satisfied by either one of the two structures, or both. The introduction of two distinct conformational states in our calculation protocol allowed to satisfy all experimental restraints, and also decreased the final target energy (Fig. 2d) by a factor of 16. Additional calculations including a third state did not result in further improvement of the target energy (Fig. 2d), indicating that a model including two states is sufficient to satisfy all experimental data .

When the two final average structures are superimposed on the common peptide segments [11-97] and [137-223], the backbone RMSD of the ATP-lid segment is *ca.* 20 Å (Fig. 1c). The first state corresponds to the structure with the ATP-lid in an open position, similar to the bundle of representative structures previously resolved for the isolated HSP90 α -NTD (Fig. 1a). The second state shows the ATP-lid in a closed conformation, covering the entire ATP-binding site, and helix-4 located near the end of long helix-2. A careful analysis of these structures showed that 22 NOE-derived distance restraints are satisfied only by the ATP-lid open state, while 5 distance restraints are specific to the ATP-

lid closed state (Fig. 2c, Supporting Table 2), and the remaining 125 restraints are satisfied in both states.

Stabilization of ATP-lid closed and open states by mutagenesis. Our structural model of the ATP-lid closed state points toward a salt bridge involving residues R60 and D127 as a potential contributor to the enthalpic stabilization of this so far unknown conformation of HSP90 α -NTD (Fig 2c). Therefore, we speculated that HSP90 α -NTD mutants suppressing the R60/D127 electrostatic interaction would push the population equilibrium toward the ATP-lid open conformation. Two ¹³CH₃-labeled samples of single-point HSP90 α -NTD mutants (D127A and R60A) were prepared and analyzed by NMR. While the D127A mutant showed NMR spectral signatures characteristic of a structurally heterogeneous sample, the R60A mutant yielded a single set of ¹H-¹³C correlation peaks -similar to the WT (Supporting Fig.4a). Analysis of a ¹³CH₃-edited NOESY spectrum recorded for this R60A mutant revealed that all expected NOE correlation peaks specific of the ATP-lid closed state were no longer detected for this mutant (Supporting Fig.5). This result indicates that disruption of the R60/D127 salt bridge decreases the population of the ATP-lid closed state to a level that is no longer detected by NOESY-type NMR experiments.

In a similar way, our structural model of the ATP-lid open state allowed us to identify a stabilizing hydrogen bond between the side chains of R46 and S129 (Fig 2c).

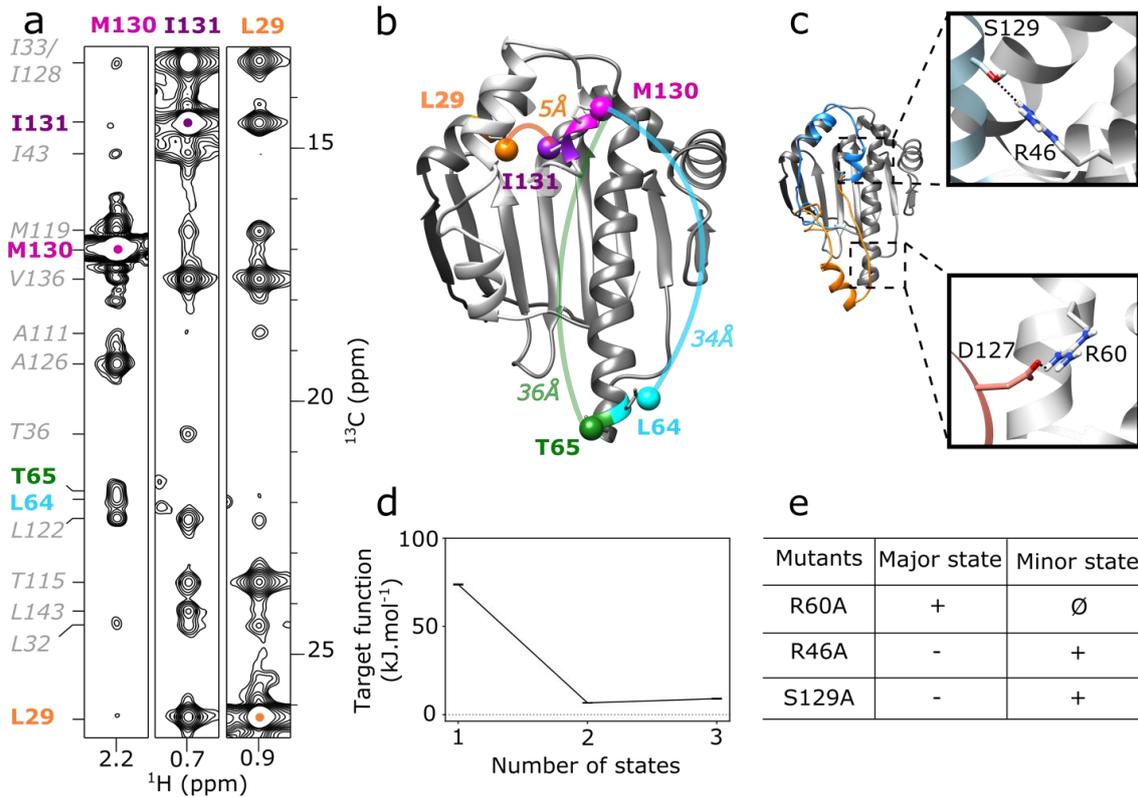


Figure 2: HSP90 α NTD samples two states in solution. **a**) 2D-strips extracted from $^{13}\text{CH}_3$ -edited 3D NOESY, at ^{13}C frequencies of M130, I131 and L29 methyl groups of WT-HSP90 α -NTD. The 3D HMQC-NOESY-HMQC experiment was acquired at 25°C on a spectrometer operating at a ^1H frequency of 950 MHz. **b**) Representation on a 3D structure of HSP90 α -NTD (PDB code: 1YES) of examples of intermethyl NOEs involving M130 and I131 with methyl groups distant by more than 30 Å. **c**) Representatives two states of HSP90 α -NTD calculated simultaneously⁵⁴ to satisfy all NMR restraints acquired on WT-HSP90 α -NTD. Boxes correspond to zooms on structure displaying interactions stabilizing the conformation of the ATP-lid either in closed (orange) or open conformation (blue) **d**) Variation of average CYANA target function for the 20 best structures according to the number of states used to calculate each conformer. **e**) Table summarizing intensity increases (+), decreases (-) or disappearance (∅) of characteristic NOEs corresponding to HSP90 α -NTD open and closed states (Supporting Table 2).

Again, we produced $^{13}\text{CH}_3$ -labeled samples of single-point HSP90 α -NTD mutants (R46A and S129A), in order to destabilize the open state and increase the population of the closed state. For both mutants, homogeneous ^1H - ^{13}C correlation spectra were obtained (Supporting Fig.4b), and methyl NOEs characteristic of both the open and closed states were detected. Still, these mutations resulted in a shift of population from the open to closed state, as the peak intensities of NOE correlation peaks specific for the closed state were enhanced by about a factor *ca.* 2.8 with respect to those characteristic of the open conformation (Supporting Fig.5). The observation of population shifts induced by single-point mutations derived from our structural models of the ATP-lid closed and open states also strongly supports our 2-states

model and the existence of an ATP-lid closed state in WT-HSP90 α -NTD in solution.

NMR refinement of structural models for the ATP-lid open and closed states.

Since the HSP90 α -NTD mutants possess populations of the two ATP-lid conformations strongly skewed toward either an open or a closed state, they provide an opportunity to further refine structural models. For structure calculation of the ATP-lid open state, a set of 114 distance restraints, involving at least one methyl group of the ATP-lid segment, was derived from the 3D $^{13}\text{CH}_3$ -edited NOESY spectrum acquired using $^{13}\text{CH}_3$ -labeled sample of R60A-HSP90 α -NTD. As the experimental set of distance restraints is self-consistent, calculations were performed assuming a single ATP-lid open state for each calculated conformer. No violation of experimental restraints was observed in the 20 best calculated structures. After a final refinement step using molecular dynamics in explicit water and in presence of experimental restraints⁴⁵, the structure of the ATP-lid segment can be superimposed on backbone atoms with a RMSD to the average structure of 0.9 Å (Fig. 3a, Supporting Table 3). The calculated ATP-lid open conformation forms three helices ([100-104], [115-123] and [128-135]) separated by short loop segments. As the backbone signals were not detectable for the segment [105-115], this part of the ATP-lid is less well defined due to a limited number of restraints.

For refinement of the ATP-lid closed state we based our structure calculation on NOE-derived distance restraints from the R46A mutant, as this mutation is located

far away from ATP-lid in closed state contrary to S129A mutant which may affect ATP-lid structure. For R46A mutant NOEs that are specific to the ATP-lid open state can still be detected, although at reduced intensity, we used the refined structure of the ATP-lid open state to identify and filter-out restraints specific to this conformation (Supporting Table 2). The remaining 81 inter-methyl NOEs, completed with 36 dihedral- and 41 backbone distance restraints (Supporting Table 1), were used to refine the structure of the ATP-lid segment in the closed state on the rigid HSP90 α -NTD scaffold. The final structural ensemble, superimposable to the average structure with a backbone RMSD of 1.6 Å (Fig. 3b), did not show any violation of experimental restraints. The hinge residues allowing the ATP-lid to switch from an open to a closed conformation are residues [109-111] and [135-136]. While the helix-3 and helix-4 are conserved in this structure, the segment corresponding to helix-5 in the ATP-lid open state is in an extended conformation in the closed state, thus enabling formation of a salt-bridge between R60 and D127 as well as methyl-methyl contacts between ATP-lid A121, A124, A126, M130 side-chains and methyl moieties L64 and T65.

The ATP-lid closed state is a metastable excited state. To study the stability of both HSP90 ATP-lid open and closed states, we have explored the conformational landscape around these two families of conformers using molecular dynamics simulations (without experimental restraints).

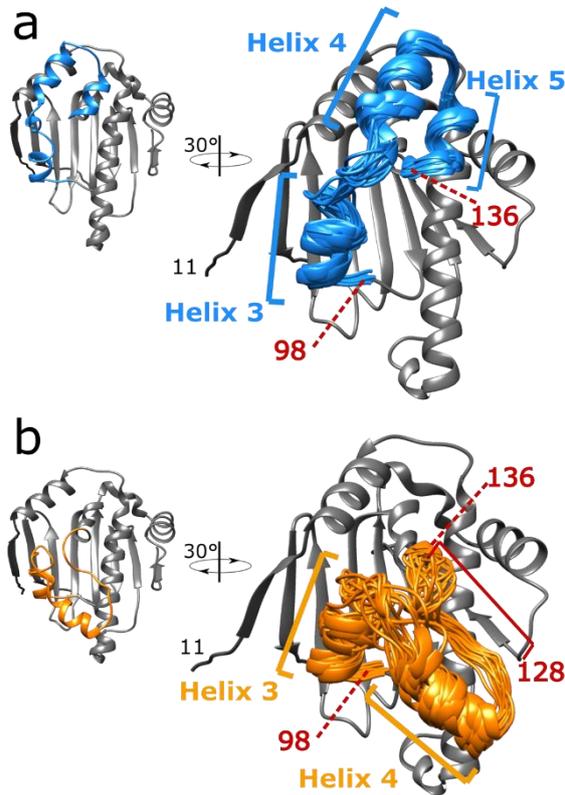


Figure 3: Solution structure ensembles of HSP90 α -NTD ATP-lid open and closed states. For each state, the 20 best CYANA conformers were selected for further restrained molecular dynamics refinement in explicit water. For each panel, the centroid representative conformer of the ensemble is presented on the left in an orientation similar to Fig 2b,c. On the right the structure ensemble is tilted by 30° and was superimposed on the coordinate of the centroid conformers. In **a**) structure ensemble for ATP-lid open/ground-state calculated using NMR structural distance restraints obtained using R60A-HSP90 α -NTD sample. In **b**) Structure ensemble for ATP-lid closed/excited-state calculated using NMR structural distance restraints obtained using R46A-HSP90 α -NTD sample. Helices 3, 4 and 5 correspond to ATP-lid helices. The positions of residues 98 and 136 are also indicated. The location of the unfolded α_5 helix in the closed state is also indicated (between 128 and 136).

Representative conformers of the structural ensembles obtained for the open and closed states were used as starting models for forty 1 μ s-molecular dynamics

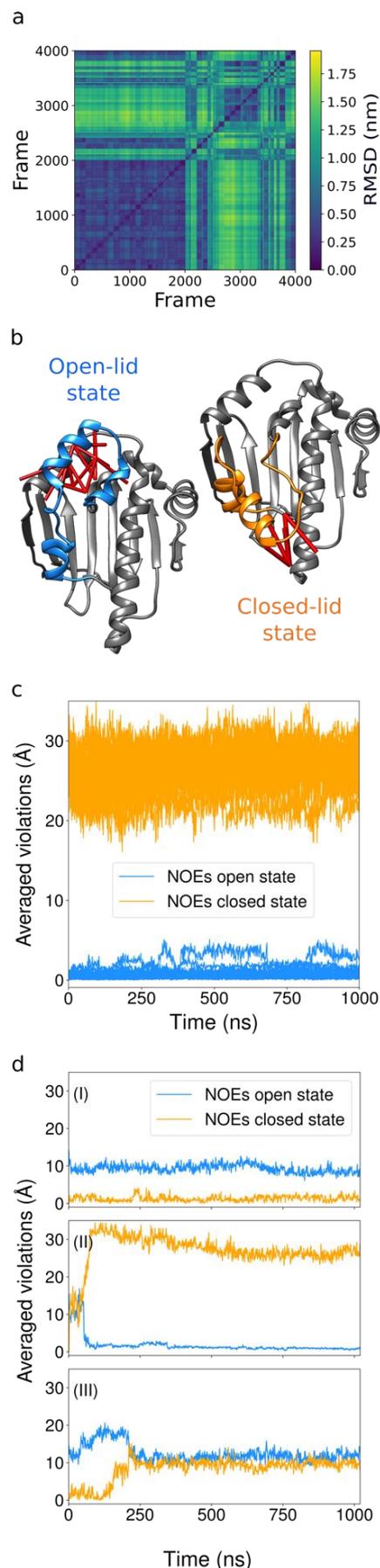
simulations. Simulations starting from an ATP-lid in the open state remained stable over the whole 1- μ s trajectory. More quantitatively, the pair-wise C α -RMSD over 2000 conformations extracted every 10 ns from the twenty 1- μ s long trajectories remains low (Fig 4.a bottom left corner), with an average of 4.0 Å. Furthermore, these structures satisfy the NOE contacts characteristic of the ATP-lid in the open state (Fig. 4.b), even if no restraints were applied during the molecular dynamics simulations. In contrast, the 2000 conformations extracted from the 20 trajectories starting with the ATP-lid in the closed state showed a larger averaged pairwise RMSD (9.3 Å, see the top right corner of Fig 4.a). This structural diversity may originate from a complex energy landscape around an excited state, and/or from the uncertainties of the initial experimental structures. Roughly, the trajectories show three typical behaviors. The first case scenario is represented by 7 trajectories (among 20) that were considered as stable, *i.e.*, they explore conformational space around the starting conformation without large violations of the NOE contacts that are characteristic of ATP-lid in the closed state during the whole 1 μ s-simulations (Fig. 4.d-I and Supporting Figure 6). This indicates that the refinement procedure has provided a state that is metastable.

In contrast, among the unstable trajectories, about 9 show a transition toward conformations almost compatible with the NOE restraints characteristic of the open state (Fig. 4.d-II). They are characterized by RMSD values, in the anti-diagonal corners of Fig 4.a, as low as 1.5 Å.

Figure 4: Molecular dynamic investigation of HSP90 α -NTD ATP-lid open and closed states.

a) Pairwise RMSD (in nm) among the ATP-lid structures obtained in the forty Molecular Dynamics simulations of 1 μ s duration each. The initial models are the ensemble of experimentally refined structures (Fig. 3). Hundred frames, separated by 10 ns from each other, are extracted from each simulation. The first 2000 frames are extracted from simulations starting with ATP-lid open-state conformers. The last ones are extracted from simulations starting with ATP-lid closed-state conformers. Going from blue to yellow, the RMSD values increase. **b)** Representation of NOE distance restraints characteristics of ATP-lid open and closed states on the corresponding structure. **c)** Violations of characteristics NOEs monitored during the molecular dynamics simulations performed without restraints, for the 20 simulations starting with the ATP-lid in the open state (See methods for definition of computed violations). The values are averaged over all the ATP-lid open (blue) or closed (orange) states specific distances restraints (Supporting Table 2). **d)** As (c), for a typical simulation starting with the ATP-lid in the closed state that remains stable (I), undergoes a transition towards the ATP-lid in the open-state (II), or derives towards region of conformational space that is neither the closed, nor the open ATP-lid state (III).

Finally, in 20 % of the trajectories, the ATP-lid populates conformations that are neither clearly the closed, nor the open state (Fig. 4.d-III). The observation of several spontaneous transitions of the ATP-lid from the closed to the open state, and the stability of the MD trajectories starting from the open state, indicate that the ATP-lid open state (the only state observed by X-ray crystallography - Fig. 1), is the ground state of HSP90-NTD, while the ATP-lid closed state is a metastable, less populated excited state. This conclusion is also supported by the observation of a small number (5) of characteristic NOEs for the closed ATP-lid state compared to a larger number (22) detected for the open ATP lid ground state (Supporting Table 2).



ATP-lid open and closed states exchange on the millisecond time scale.

The fact that a single set of NMR signals is detected for HSP90 α -NTD (WT, R60A and R46A mutants) indicates that the ATP-lid open and closed states are in fast exchange with respect to the NMR time scale ($\tau_{\text{ex}} < \sim 10$ ms). The observed severe line broadening of amide backbone resonances in the ATP-lid segment also points toward exchange dynamics on the micro- to millisecond time scale. In order to further quantify the kinetics (and thermodynamics) of the ATP-lid structural rearrangement, NMR relaxation-dispersion experiments were performed at 2 different magnetic field strengths using ^{13}C - ^1H methyl, as well as ^{15}N backbone amide probes. These data reveal that conformational exchange is mainly sensed by nuclei in the ATP-lid segment, as well as protein regions that are in close contact with the ATP-lid (Fig. 5a). Particularly strong exchange-contributions were detected in the hinge regions of the ATP-lid segment: residues [T109-T115] and V136. Assuming a simple two states model to describe the structural rearrangement of the ATP-lid, a systematic grid search of the model parameters fitting the relaxation dispersion data detected on backbone and methyl probes was performed. A single minimum was identified, indicating that the experimental data can be interpreted with the presence of a single low populated minor state (Supporting Fig.7). A global fit of the $^{13}\text{CH}_3$ -CPMG relaxation dispersion, characterized by a higher signal to noise ratio, enabled us to determine the global exchange rate $k_{\text{ex}}=k_1+k_{-1}= 2490 \pm 61 \text{ s}^{-1}$ and relative populations of the 2 states of $96.8 \pm 0.1 \%$

(ground state: ATP-lid open state) and $3.2 \pm 0.1 \%$ (excited state: ATP-lid closed state). This corresponds to half-life times of 8.6 ± 0.4 ms for the ATP-lid open state, and 0.29 ± 0.01 ms for the ATP-lid closed state. These populations and exchange rates correspond to a difference of Gibbs free energy of $8.3 \pm 0.2 \text{ kJ.mol}^{-1}$ between the 2 states, with an activation energy (from the ground to the transition state) of $61.0 \pm 0.2 \text{ kJ.mol}^{-1}$ ⁴⁶.

Discussion

In this work, we have demonstrated that the ATP-lid segment of human HSP90 α -NTD samples two distant conformations in solution. We have solved the solution structures of this important anti-cancer drug target with the ATP-lid either in an open or closed conformation. Such a closed state with the ATP-lid covering the nucleotide/drug binding site has never been observed before for isolated HSP90 α -NTD, neither in the apo state nor in complex with nucleotides or drugs (Fig. 1). In addition to our NMR data, we have used mutagenesis and molecular dynamics simulation to independently validate the existence of this meta-stable, low populated closed state in human HSP90 α -NTD. The calculation of Evolutionary Couplings (ECs) from sequence alignment of multiple homologous proteins is an emerging approach to identify conserved contacts and complement conventional approaches of protein structure determination⁴⁷. Starting from an alignment of 4414 homologous sequences of HSP90-NTD, 71 highly probable ECs ($p > 0.85$) involving at least one residue of the ATP-lid segment (Supporting Fig.8) were computed⁴⁷. Interestingly, we found

that both, the open and closed conformations are necessary to account for the structural contacts revealed by these ECs (Supporting Fig.8), further supporting the existence of an ATP-lid closed state.

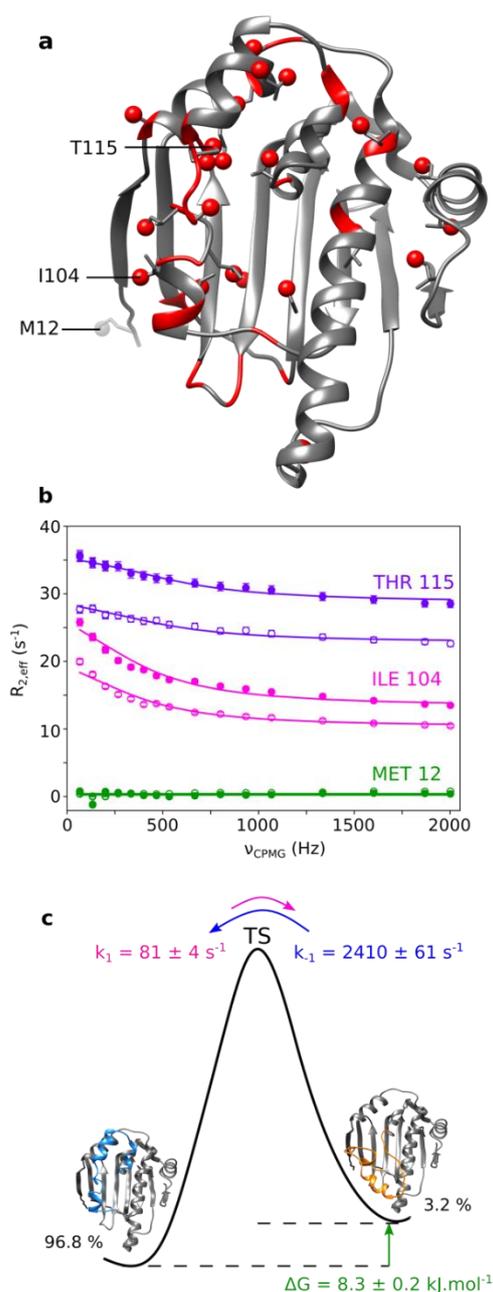


Figure 5: Relaxation dispersion study of HSP90 α -NTD. **a)** 3D structure of human HSP90 α -NTD displaying in red methyl- and backbone ^{15}N - probes for which conformational exchange in the μs -ms time scale was detected. **b)** Examples of $^{13}\text{CH}_3$ MQ CPMG relaxation dispersion profiles of Thr-115, Ile-104 and Met-12 plotted in purple, red and green, respectively. The filled circles for each color represent data acquired at 850 MHz and the empty circles represent data acquired at 700 MHz. Data displayed were acquired at 293 K using a U- ^2H , ^{15}N , ^{12}C , Ala- $^{13}\text{C}^1\text{H}_3$ $^\beta$, Ile- $^{13}\text{C}^1\text{H}_3$ $^{\delta 1}$, Leu- $^{13}\text{C}^1\text{H}_3$ $^{\delta 2}$, Met- $^{13}\text{C}^1\text{H}_3$ e , Thr- $^{13}\text{C}^1\text{H}_3$ g , Val- $^{13}\text{C}^1\text{H}_3$ $^{\gamma 2}$ HSP90-NTD sample. Experimental data were fitted to a two-sites exchange model (global fit of 21 relaxation dispersion curves). Precision on the data was estimated using Monte Carlo. **c)** Schematic diagram of the energy landscape for the exchange between the ATP-lid open (ground) and closed (excited) states of HSP90 α -NTD. Both thermodynamics and kinetics parameters of the exchange, extracted at 293 K, are displayed. Activation energies can be estimated using Eyring equation (assuming $\kappa = 1$). From the ground state to the transition state: $\Delta G^\ddagger = 61.0 \pm 0.2 \text{ kJ.mol}^{-1}$ and from the excited state to the transition state: $\Delta G^\ddagger = 52.7 \pm 0.1 \text{ kJ.mol}^{-1}$.

Using Molecular Dynamics simulations and NMR relaxation-dispersion experiments, we could establish that the ATP-lid in the open conformation is thermodynamically more stable than the ATP-lid in closed state, and that the population of the excited (closed) state is only 3-4 % at room temperature (Fig. 5). The open and closed states interconvert at a rate ($k_{\text{ex}} = k_1 + k_{-1}$) of 2.5 kHz, which seems to be very fast for such an extensive structural rearrangement. However, structural rearrangements at fast time scales were already reported for other conformational exchange events in proteins. An example is EIN, a 128 kDa homodimeric protein that switches from an open to a closed state at a speed faster than 10 kHz⁴⁸. Furthermore, PET fluorescence quenching experiments by Schulze *et al.*³⁰ monitoring yeast HSP90-

NTD ATP-lid structural rearrangement in isolated monomeric constructs reveals a motion of the ATP-lid with a characteristic rate of 1.5 kHz, even though a full closure of the lid over the nucleotide/drug binding site has not been observed. This frequency range is similar to our results obtained by NMR relaxation dispersion experiments for human HSP90 α -NTD. The transition pathway between open and closed ATP-lid states, distant by up to 30 Å, is obviously more complicated than a simple two-state model used here for the quantitative analysis of our NMR data. A more detailed analysis of the computed MD trajectories allows to obtain further information on the kinetics of the structural changes observed for the open-closed transition. The most pronounced rearrangements between the open and closed states are the reorientation of helix-3 and helix-4, as well as the folding/unfolding of helix-5. The MD trajectories describing a full transition from the closed to the open state show that the ATP-lid can tilt very rapidly, typically in a few hundreds of ns, while the folding of helix-5 and the associated local restructuring were not systematically observed during the 1- μ s simulations (Supporting Figure 9). Therefore, we may conclude that the transition pathway involves several structural rearrangements occurring at different timescales, with the unwinding/folding of the helix-5 being one of the rate limiting steps of the transition between ATP-lid open and closed states.

To assist the folding of client proteins, the ATPase HSP90 interacts with a large number of co-chaperones and undergoes a conformational cycle in which a complex set of structural changes

occurs. In currently proposed models of the HSP90 functional cycle, the two N-terminal domains of HSP90 transiently dimerize upon binding of ATP, client proteins and co-chaperones⁴⁹. Dimerization is associated with an exchange of the first β -strand (β 1) between both N-terminal domains, and a switch of the ATP-lid from an open to a closed state^{29,30,50,51}. Cryo-EM structures of full-length HSP90 co-vitrified with ATP, a co-chaperone and a client protein^{52,53} reveal an ATP-lid segment in a closed conformation associated to a dimerization of the N-terminal domains and a swap of the β 1-strands between both N-terminal domains (Fig. 6a-c). To investigate whether the closed ATP-lid conformation is responsible for dimerization of HSP90-NTD, we have identified 16 methyl-methyl proximities in the N-terminal domain of full-length human HSP90 α (PDB: 7L7J) that are characteristic of the dimeric HSP90 with a closed ATP-lid and swapped β 1-strands (Supporting Table 4). However, investigation of the inter-methyl NOEs observed for the HSP90-NTD mutant R46A, showing an enhanced population of the ATP-lid closed-state ensemble, revealed the absence of all the putative NOEs characteristic of a dimerization of HSP90 α -NTDs, confirming that the closed state of HSP90 α -NTD stays monomeric in solution.

The NMR-derived ATP-lid structure in closed state of HSP90 α -NTD was compared to the NTD extracted from full-length HSP90 α (FL-HSP90-NTD) focusing on the ATP-lid segment that is stabilized in the closed conformation by the binding of ATP, co-chaperone and client protein⁵³.

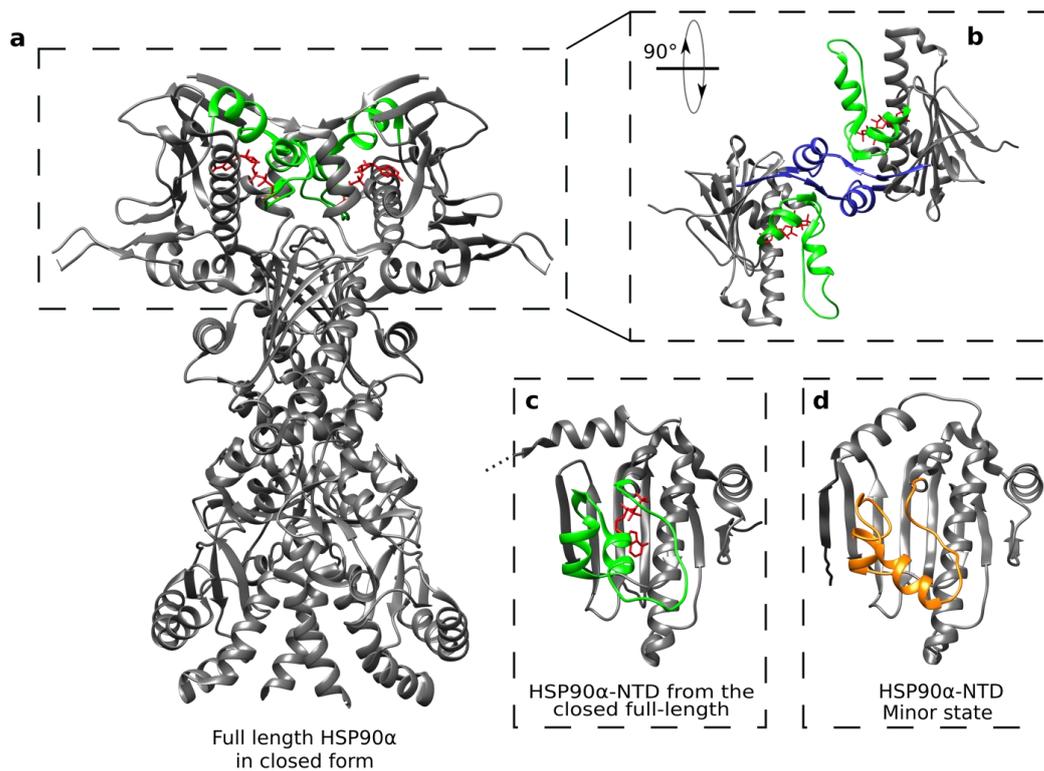


Figure 6: Structure comparison of HSP90 α -NTD excited state with full-length dimeric HSP90 α functional cycle intermediate. a) Structure of the homodimer full-length HSP90 α in closed form stabilized by p23, FKBP51 and ATP (PDB: 7L7J⁵³). The segment covering the nucleotide binding site and ATP are colored in green and red, respectively. b) Zoom on the two N-terminal domains of the full-length HSP90 α , turned by 90°. The blue segments represent the β strands exchanging between the two chains of the homodimer. c) N-terminal domain of HSP90 α (PDB: 7L7J). d) Average structure of the calculated ensemble for ATP-lid closed (excited) state of apo HSP90 α -NTD. The segment 98-136 is represented in orange.

The closed state of apo-HSP90 α -NTD can be superimposed on the ATP-lid of dimeric full length HSP90 α (PDB: 7L7J) with a RMSD of 4.1 Å (C α backbone atoms). This RMSD value highlights the resemblance of the NMR-observed metastable excited state to the structure of the ATP-lid in FL-HSP90-NTD (Fig. 6c-d). In particular, the secondary structural motifs describing the ATP-lid segment are

comparable: helix-3 and helix-4 are present, while the segment forming helix-5 in the open conformation is in an extended conformation in both FL-HSP90-NTD and the excited state of isolated HSP90-NTD. The presence of ATP in the structure of the full-length protein as well as the β -strand exchange leading to additional interactions may explain the remaining differences. Although our study was performed on isolated HSP90 α -NTD, our results suggest that the functionally relevant closed state of HSP90 ATP-lid pre-exists in apo HSP90 before dimerization driven by ATP binding. Therefore, ATP binding selects and stabilizes a preexisting conformation before inducing further structural rearrangements. These results highlight the importance of investigating protein dynamics and structure at a physiologically relevant temperature. This is particularly relevant for structure-based

design of new drugs against the ATP binding domain of human HSP90.

Methods

Preparation of isotopically labeled HSP90 α -NTD samples.

E. coli BL21-DE3-RIL cells transformed with a pET-28 plasmid encoding the N-Terminal domain of HSP90 α from Homo Sapiens (HSP90 α -NTD) with a His-Tag and a TEV cleavage site were progressively adapted in three stages over 24 h to M9/²H₂O. In the final culture, bacteria were grown at 37°C in M9 medium with 99.85 % ²H₂O (Eurisotop), 1 g/L ¹⁵N¹H₄Cl (Sigma Aldrich) and 2 g/L D-glucose-d₇ or D-glucose-¹³C₆-d₇ for U-[²H, ¹²C, ¹⁵N] and U-[²H, ¹³C, ¹⁵N] HSP90 α -NTD samples, respectively. For production of U-[²H, ¹⁵N, ¹²C], Ala-[¹³C¹H₃] ^{β} , Met-[¹³C¹H₃] ^{ϵ} , Leu-[¹³C¹H₃] ^{δ^2} , Ile-[¹³C¹H₃] ^{δ^1} , Thr-[¹³C¹H₃] ^{γ} , Val-[¹³C¹H₃] ^{γ^2} wild type HSP90 α -NTD or mutants (R46A, R60A, D127A, S129A), HLAM-A ^{β} I ^{δ^1} M ^{ϵ} (LV)^{proST γ} labeling kits (NMR-Bio) were added to the M9/²H₂O media³⁹. To cross-validate sequence specific assignment of methyl probes, 56 single point mutant samples were produced⁴⁰ (see supporting Fig.2 for list of mutants).

For all samples, protein production was induced by 0.5 mM of IPTG when the O.D at 600 nm reached *ca.* 0.8, overnight at 20 °C. HSP90 α -NTD samples were purified in two steps using a Ni-NTA affinity chromatography step followed by a size exclusion chromatography step as previously described³⁸. For relaxation dispersion experiments, an additional tag cleavage step was performed for the preparation of the wild type HSP90 α -NTD samples, by addition of TEV protease at room temperature for 2 h (in a ratio 1:10) before removal of His-tagged-TEV protease and cleaved His-Tag using reverse Ni-NTA affinity chromatography.

NMR Spectroscopy.

The final NMR buffer conditions were 20 mM

HEPES (pH 7.5), 150 mM NaCl and 1 mM TCEP in either 100% ²H₂O or 90% ¹H₂O/10% ²H₂O. Single point mutant samples to cross-validate the assignment were concentrated at [0.1-0.4] mM. 40 μ L of each sample was loaded in a 1.7 mm NMR tube and 2D ¹H-¹³C SOFAST methyl TROSY⁵⁵ experiments were recorded for an average duration of ~ 1.5 h each, on a spectrometer equipped with a 1.7 mm cryogenically cooled, pulsed-field-gradient triple-resonance probe and operating at a ¹H frequency of 850 MHz. All other NMR experiments were recorded on isotopically labeled HSP90 α -NTD samples (200 μ L of 0.5 mM labeled protein in 4 mm Shigemi tube) using a Bruker Avance III HD spectrometers equipped with 5 mm cryogenic probes.

3D HNH- and NNH-edited NOESY experiments were recorded at 293 K on a spectrometer operating at a ¹H frequency of 850 MHz using a U-[²H, ¹³C, ¹⁵N] wild type HSP90 α -NTD sample in 90% ¹H₂O/10% ²H₂O buffer. 3D CCH-edited HMQC-NOESY-HMQC NMR experiments^{56,57} were recorded on high field spectrometer, operating at a ¹H frequency of either 850 MHz or 950 MHz, using U-[²H, ¹⁵N, ¹²C], Ala-[¹³C¹H₃] ^{β} , Ile-[¹³C¹H₃] ^{δ^1} , Leu-[¹³C¹H₃] ^{δ^2} , Met-[¹³C¹H₃] ^{ϵ} , Thr-[¹³C¹H₃] ^{γ} , Val-[¹³C¹H₃] ^{γ^2} samples of wild type, R46A-, R60A-, or S129A-HSP90 α -NTD constructs in ²H₂O buffer. Each 3D CCH-edited NOESY experiment was recorded over *ca.* 2-3 days, at 298 K for wild type HSP90 α -NTD, 288 K for R46A- and S129A mutants or 293 K for R60A-HSP90 α -NTD.

Methyl TROSY ¹³C¹H-multiple quantum CPMG relaxation dispersion experiments⁵⁸ modified with Ernst angle excitation⁵⁵ experiments and ¹⁵N single quantum CPMG relaxation dispersion experiments were recorded at 293 K using U-[²H, ¹⁵N, ¹²C], Ala-[¹³C¹H₃] ^{β} , Ile-[¹³C¹H₃] ^{δ^1} , Leu-[¹³C¹H₃] ^{δ^2} , Met-[¹³C¹H₃] ^{ϵ} , Thr-[¹³C¹H₃] ^{γ} , Val-[¹³C¹H₃] ^{γ^2} WT- HSP90-NTD sample, on spectrometers operating at ¹H frequencies of 700 MHz and 850 MHz (¹³C¹H₃ experiments)

or of 600 MHz, 700 MHz and 850 MHz (^{15}N experiments). Methanol was used as a reference for precise temperature calibration on each spectrometer. The power of the ^{13}C refocusing pulse in CPMG was set to 16 kHz (or 4 kHz for ^{15}N), the CPMG relaxation period was set to 30 ms for ^{13}C - (60 ms for ^{15}N -) experiments. Dispersion profiles, recorded in an interleave manner, comprised 16 different CPMG frequencies (ν_{cpmg}) and were ranging from 66 Hz to a maximum of 2000 Hz, whilst dispersion profiles comprised 11 different ν_{cpmg} and were ranging from 33 Hz to a maximum of 1000 Hz for ^{15}N experiments. In addition, a reference spectrum was acquired for both types of experiments by omitting the relaxation period. Total experimental acquisition time for the relaxation dispersion experiment was 1 day for ^{13}C -experiments and 2 days for ^{15}N experiments.

Structural Restraints.

The following types of experimental distance restraints were used for structure calculations of HSP90-NTD variable segment from M98 to V136: (i) H_N — H_N restraints derived from ^{15}N -edited NOESY spectra acquired using a NOE mixing time of 240 ms; and (ii) methyl—methyl restraints from ^{13}C -edited NOESY spectra. A qualitative approach was used in the derivation of distance restraints. In particular, intermethyl distance restraints were fixed between the barycenters of each proton triplets with distance bound of 2-4 Å and 4-6 Å for strong and weak cross-peaks detected in 3D-NOESY acquired with a short mixing time (*i.e.* 100 ms) or with distance bound of 4-8 Å for cross-peaks observed only in experiment acquired with longer NOE mixing time (*ca.* 300 to 400 ms). In the case of the mutant HSP90-NTD-R46A, intermethyl distance restraints were fixed between the barycenters of each proton triplets with distance bound of 2-8 Å as no short mixing 3D-NOESY experiment was acquired. H_N — H_N distance bounds for

observed NOE cross-peaks were fixed to 2–6 Å. For R60-D127 and R46-S129 stabilizing interaction identified using mutagenesis, supplementary distances restraints were fixed between oxygen acceptor atoms and arginine side chain nitrogen/hydrogen donor atoms with distance upper limits of 2/3 Å, respectively. Backbone chemical shifts of WT-HSP90-NTD constructs were used to determine ϕ and ψ dihedral angles using TALOS+ software^{42,59,60} and angular restraints with a tolerance of $\pm 10^\circ$ were applied. For the HSP90-NTD segment from M98 to V136, a total of 54 dihedral angles restraints were used for the calculation of WT- and R60A-HSP90-NTD structures. Due to widely broadened missing peaks, no angular restraints were experimentally derived from residues 105 to 115.

For the structure refinement of R46A-HSP90-NTD excited state, the upper limit of the 5 distance restraints specific to the ATP-lid closed state (Supporting Table 2) were reduced to 6.5 Å to take into account the lower population of this transiently sampled state, while the 22 structural restraints specific of the ATP-lid in open conformation (Supporting Table 2) were excluded of distance restraints set. The angular restraints for helix-5 (from I128 to V136) were excluded for the calculation R46A-HSP90-NTD structures, as we had evidences from initial structure calculation that helix-3 and helix-4 are preserved while helix-5 unfolds in the excited state (Supporting Table 5). Total number and distribution of structural restraints can be found in supporting Fig.2 and Table 1.

Structure calculation.

Structural calculations were performed using CYANA 3.98.13 simulated annealing protocol⁴³. The invariable segments from residues 11 to 97 and from 137 to 223 (Fig. 1) were treated as a rigid body core using a set of 13434 C_α - C_α distance restraints and 477 dihedral angles (ϕ , ψ and χ_1) extracted from

X-ray crystal structure of apo HSP90-NTD (PDB code: 1YES⁶¹). The solution structures sampled by the variable segment of HSP90-NTD (Fig. 1) from residue M98 to V136 which harbors the specific loop covering the ATP-binding site was determined using NMR structural information obtained on 3 different samples (WT and R46A-, R60A-mutants). Initial structure calculation of WT- and R46A-HSP90-NTD were performed using a two-state structure calculation protocol enabling each distance restraints to be satisfied in either ATP-lid open- or closed- state or both states simultaneously⁵⁴.

Structural data acquired for the R60A-construct were used with a single state calculation protocol to refine the ATP-lid open structural ensemble (Fig. 3A). Conversely, the intermethyl NOE data set acquired for R46A-HSP90-NTD construct was used to refine the ATP-lid closed structural ensemble. For each HSP90-NTD structural state, a CYANA simulated annealing protocol using 25000 torsion angle dynamic steps⁴³ was used to generate one thousand conformers starting from random coordinates and using NMR experimental restraints. The twenty conformers characterized by the lower target function values were selected for further refinement using a CNS restrained molecular dynamics protocol using a full Lenard-Jones potential and explicit water molecules⁶². In brief, the explicit solvent refinement consisted of the five following steps: (i) immersion in a 7.0 Å shell of water molecules and energy minimization; (ii) slow heating from 100 to 500 K in 100 K temperature steps with 200 MD steps per temperature step (time/step 3 fs), with harmonic position restraints on the protein heavy atoms that were slowly phased out during the heating stage; (iii) refinement at 500 K with 2,000 MD steps (time step 4 fs); (iv) slow cooling from 500 K to 25 K in 25 K temperature steps with 200 MD steps per temperature step (time step 4 fs); (v) final energy minimization (200 steps).

Molecular Dynamics Simulations.

The experimentally refined ensembles were used as starting points for MD simulations. Forty separate simulations were run for 1 μ s, for the 20 different conformations observed for both the ATP-lid in the closed and open states.

1.1 Setup and parameters

The all-atom Amber ff14SB force field⁶³ was used with explicit water molecules described using TIP3P model⁶⁴. The Gromacs software⁶⁵ (single precision, 2020 versions) was used to perform preparation, equilibration and production steps. Each protein model was solvated in an octahedral box of about 13000 water molecules extending at least 13 Å away from the protein. Standard protonation states were assumed. The His189 was protonated on the N δ 1 atom, whereas the other three His residues were protonated on the N ϵ 2 atom, in accordance with the pKa calculations performed by H++ server⁶⁶ tested on 20 typical structures with relative permittivity of the solvent of 78, relative permittivity of the protein of 4 at a salt concentration of 0.17 M. This resulted in a net charge of -11 for the protein. NaCl salt and Na+ counterions were added to obtain an electrically neutral system, with a salt concentration of about 170 mM. For each simulation, a steepest-descent energy minimization was used until the maximum force was lower than 100 kJ.mol⁻¹.nm⁻¹, followed by 100 ps constant-volume equilibration and 100 ps constant-pressure equilibration, both performed with heavy nonwater atoms restrained toward the starting structure with a force constant of 1000 kJ.mol⁻¹.nm⁻². Finally, 1020 ns were performed with constant pressure, constant temperature, without restraints. The first 20 ns were considered as an initial equilibration, and the last 1000 ns were used for the analysis presented in the article. All bonds involving hydrogen atoms were constrained to the equilibrium value using 1 iteration to the 4th order of the Linear Constraint Solver⁶⁷, allowing for a time step of 2 ps for the leap-

frog integrator. The temperature was kept constant at 300 K using two stochastic velocity rescaling thermostats⁶⁸ for the solvent and for the protein, with a time constant of 2 ps. The pressure calculated with long-range correction was kept constant at 1 atm using a Parrinello-Rahman isotropic barostat⁶⁹ with a relaxation time of 2 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$. Long-range electrostatics were handled by smooth particle-mesh Ewald (PME) summation⁷⁰ with a fourth-order B spline interpolation and a grid spacing of 0.16 nm. The cutoff radius for Lennard-Jones interactions was initially set to 10 Å, but the neighbor list parameters were automatically optimized with an energy drift tolerance of $0.005 \text{ kJ.mol}^{-1}.\text{ps}^{-1}$ per particle⁷¹. The snapshots taken every 1 ns were analyzed with Gromacs and Mdanalysis⁷².

1.2 Analysis of the Molecular Dynamics Simulations

1.2.1 RMSD For each of the 40 trajectories, 100 conformations were extracted (every 10 ns), yielding 4000 structures. The pairwise RMSD was calculated using the TTclust tool⁷³. For each pair, the protein backbone atoms were superimposed on the rigid core of the protein (backbone of segments [11:97] and [137:223]), and the RMSD was calculated for the flexible ATP-lid, i.e., the backbone atoms for segment [98:136].

1.2.2 Violations Either for the closed or for the open state, the specific NOEs are defined by an ensemble of N couples of residues $\{ij\}$ (see Supporting Table 2). At each time step, the distances d_{ij} between the methyles are measured, and the violation V relative to a given state is the average over the N violations $v_{(ij)}$, i.e. $V = 1/N \sum_{(ij)} v_{(ij)}$, where $v_{ij} = \max(0, d_{ij} - d_{ij}^{\text{viol}})$. The distance d_{ij} between two protonated methyles is the distance between the two centers of masses of the three respective protons. We have chosen to use $d_{ij}^{\text{viol}} = \langle d_{ij} \rangle + 2 \sigma(d_{ij})$, where $\langle d_{ij} \rangle$ and $\sigma(d_{ij})$ are the average value and the standard deviation from the bundle of the 20 structures

obtained after refinement under NMR restraints respectively. The Supporting Tables S2A and S2B provides the details on the couples $\{ij\}$, the three protons defining the center of masses, and their respective d_{ij}^{viol} .

1.2.3 Secondary structure Analysis of the secondary structure per residue was done using DSSP. The analysis was done on 120 frames, extracted from the last 120 ns of the simulation⁷⁴. The 3 helical types (α , 3_{10} and π) were considered as helices.

Relaxation dispersion Data analysis.

Peak intensities for relaxation-dispersion CPMG experiments were determined using the nonlinear fitting routine nlinLS from the NMRPipe software suite⁷⁵. Relaxation dispersion profiles were generated from peak intensities, measured from pseudo 3D experiments using different CPMG frequencies and divided by intensities extracted from the reference spectrum⁵⁸. After visual inspection to exclude too noisy relaxation dispersion curves, relaxation dispersion data of the non-overlapping peaks for which exchange contribution (R_{ex}) to the transverse relaxation was $\geq 2 \text{ s}^{-1}$ were analyzed and fitted with the software ChemEx⁷⁶ using a two-state exchange model. Errors for $R_{2,\text{eff}}$ rate values were estimated from twice the noise measured in the spectra. However, when errors were less than 2% of the $R_{2,\text{eff}}$ rate value, an error of 2% was assumed⁵⁸.

First, a grid search using relaxation data was carried out where both conformational exchange rate and population of the excited state were fixed (Supporting Fig 7) to identify the minimum in order to select these parameters as starting parameters for the global numerical fitting (a single and identical local minimum was identified for both ^{15}N and $^{13}\text{CH}_3$ relaxation datasets). For each experimental temperature, a global numerical fit of $^{13}\text{CH}_3$ -CPMG relaxation dispersion profiles was performed to obtain global exchange parameters (exchange rate (k_{ex}) and

the population of the excited state (P^*) as well as residue specific parameters (absolute chemical shift difference between the ground and the excited state of the system under chemical exchange, $|\Delta\omega_C|$ and $|\Delta\omega_H|$). A jackknife resampling was run to ensure that all residues globally fitted were part of the same exchange process and finally 40 Monte-Carlo repeats were used to estimate precision on global fit results. Experimental data for 21 different methyl probes were used for the global fitting of the methyl ^{13}C - ^1H CPMG-RD experiments acquired at 293 K.

Data Availabilities.

All the experimental NMR data used for the structural and dynamics characterization of human HSP90-NTD as well as structures ensembles together with corresponding structural restraints used for the calculation of HSP90-NTD in ATP-Lip open (R60A mutant) and closed (R46A mutant) will be deposited in BMRB and PDB databases and accession codes will be provided before publication.

Acknowledgements.

The authors thank Drs. R. Awad and C. Laguri for help and advices. This work used the high field NMR and isotopic labeling facilities at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by grants from CEA/NMR-Bio (research program C24990), by the *Region Auvergne-Rhône-Alpes* (pack ambition recherche 2019 - P089), by the French National Research Agency in the framework of the "*Investissements d'avenir*" program

(ANR-15-IDEX-02) and by a fellowship (to F.H.) from "La Ligue contre le Cancer".

Author contributions.

C.L., E.R., F.H., J.B., M.F. designed the experiments; E.C. and F.H. prepared the samples; F.H., J.B., B.B. and P.M. set up and collected the NMR experiments; E.R., F.H., J.B., and P.G. analyzed the NOESY experiments; A.F. and E.R. performed structural calculations; F.H., J.B. and P.M. analyzed the CPMG experiments; C.L., E.R., and P.J. performed and analyzed the MD simulations; C.L., E.R., F.H. and J.B. wrote the manuscript. All authors discussed the results, corrected the manuscripts and approved the final version.

Competing interests:

The authors declare no competing interests.

References:

1. Backe, S. J., Sager, R. A., Woodford, M. R., Makedon, A. M. & Mollapour, M. Post-translational modifications of Hsp90 and translating the chaperone code. *Journal of Biological Chemistry* **295**, 11099–11117 (2020).
2. Zuehlke, A. & Johnson, J. L. Hsp90 and co-chaperones twist the functions of diverse client proteins. *Biopolymers* **93**, 211–217 (2010).
3. Schopf, F. H., Biebl, M. M. & Buchner, J. The HSP90 chaperone machinery. *Nature Reviews Molecular Cell Biology* **18**, 345–360 (2017).
4. Picard. HSP90 INTERACTORS Chaperones and relatives. Available at: <http://www.picard.ch/downloads/Hsp90interactors.pdf>. (Accessed: 13th October 2021)
5. Biebl, M. *et al.* NudC guides client transfer between the Hsp40/70 and Hsp90 chaperone systems. *Mol Cell* **82**, 555–569 (2022).

6. Biebl, M. M. *et al.* Structural elements in the flexible tail of the co-chaperone p23 coordinate client binding and progression of the Hsp90 chaperone cycle. *Nat. Commun.* **12**, (2021).
7. Lopez, A. *et al.* Client binding shifts the populations of dynamic Hsp90 conformations through an allosteric network. *Sci. Adv* **7**, 7295 (2021).
8. D'Annessa, I., Moroni, E. & Colombo, G. Visualizing the Dynamics of a Protein Folding Machinery: The Mechanism of Asymmetric ATP Processing in Hsp90 and its Implications for Client Remodelling. *J. Mol. Biol.* **433**, (2021).
9. Mader, S. L. *et al.* Conformational dynamics modulate the catalytic activity of the molecular chaperone Hsp90. *Nat. Commun.* **11**, (2020).
10. Sanchez-Martin, C., Serapian, S. A., Colombo, G. & Rasola, A. Dynamically Shaping Chaperones. Allosteric Modulators of HSP90 Family as Regulatory Tools of Cell Metabolism in Neoplastic Progression. *Front. Oncol.* **10**, (2020).
11. Wu, J. *et al.* Heat Shock Proteins and Cancer. *Trends in Pharmacological Sciences* **38**, 226–256 (2017).
12. Stechmann, A. & Cavalier-Smith, T. Evolutionary origins of Hsp90 chaperones and a deep paralogy in their bacterial ancestors. *J. Eukaryot. Microbiol.* **51**, 364–373 (2004).
13. Sorger, P. K. & Pelham, H. R. The Glucose-regulated Protein grp94 is Related to Heat Shock Protein hsp90. *J Mol Biol.* **194**, 341–344 (1987).
14. Felts, S. J. *et al.* The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties. *J. Biol. Chem.* **275**, 3305–3312 (2000).
15. Lai, B.-T., Chin, N. W., Stanek, A. E., Keh, W. & Lanks, K. W. Quantitation and Intracellular Localization of the 85K Heat Shock Protein by Using Monoclonal and Polyclonal Antibodies. *Mol. Cell. Biol.* **4**, 2802–2810 (1984).
16. Gupta, R. S. Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. *Mol. Biol. Evol.* **12**, 1063–1073 (1995).
17. Sreedhar, A. S., Kalmár, É., Csermely, P. & Shen, Y. F. Hsp90 isoforms: Functions, expression and clinical importance. *FEBS Lett.* **562**, 11–15 (2004).
18. Taipale, M., Jarosz, D. F. & Lindquist, S. HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* **11**, 515–528 (2010).
19. Didenko, T., Duarte, A. M. S., Karagöz, G. E. & Rüdiger, S. G. D. Hsp90 structure and function studied by NMR spectroscopy. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823**, 636–647 (2012).
20. Biebl, M. M. & Buchner, J. Structure, Function, and Regulation of the Hsp90 Machinery. *Cold Spring Harb. Perspect. Biol.* a034017 (2019). doi:10.1101/cshperspect.a034017
21. Meyer, P. *et al.* Structural and Functional Analysis of the Middle Segment of Hsp90: Implications for ATP Hydrolysis and Client Protein and Cochaperone Interactions. *Mol. Cell* **11**, 647–658 (2003).
22. Meyer, P. *et al.* Structural basis for recruitment of the ATPase activator Aha1 to the Hsp90 chaperone machinery. *EMBO J.* **23**, 1402–1410 (2004).
23. López, A., Elimelech, A. R., Klimm, K. & Sattler, M. The Charged Linker Modulates the Conformations and Molecular Interactions of Hsp90. *ChemBioChem* **22**, 1084–1092 (2021).

24. Dutta R & Inouye M. GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci* **25**, 24–28 (2000).
25. Obermann, W. M. J., Sondermann, H., Russo, A. A., Pavletich, N. P. & Hartl, F. U. In Vivo Function of Hsp90 Is Dependent on ATP Binding and ATP Hydrolysis. *J. Cell Biol.* **143**, 901–910 (1998).
26. Panaretou, B. *et al.* ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo. *EMBO J.* **17**, 4829–4836 (1998).
27. Prodromou, C. *et al.* Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone. *Cell* **90**, 65–75 (1997).
28. Schubert, J., Schulze, A., Prodromou, C. & Neuweiler, H. Two-colour single-molecule photoinduced electron transfer fluorescence imaging microscopy of chaperone dynamics. *Nat. Commun.* **12**, 6964 (2021).
29. Prodromou, C. The ‘active life’ of Hsp90 complexes. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823**, 614–623 (2012).
30. Schulze, A. *et al.* Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism. *Nat. Chem. Biol.* **12**, 628–635 (2016).
31. Ciocca, D. R. & Calderwood, S. K. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress Chaperones* **10**, 86–103 (2005).
32. Calderwood, S. K., Khaleque, M. A., Sawyer, D. B. & Ciocca, D. R. Heat shock proteins in cancer: Chaperones of tumorigenesis. *Trends Biochem. Sci.* **31**, 164–172 (2006).
33. Yuno, A. *et al.* Clinical Evaluation and Biomarker Profiling of Hsp90 Inhibitors. *Methods Mol. Biol.* **1709**, 423–441 (2018).
34. Zuehlke, A. D., Moses, M. A. & Neckers, L. Heat shock protein 90: Its inhibition and function. *Philos Trans R Soc L. B Biol Sci* **373**, (2018).
35. Wang, L., Zhang, Q. & You, Q. Targeting the HSP90–CDC37–kinase chaperone cycle: A promising therapeutic strategy for cancer. *Med. Res. Rev.* 1–27 (2021). doi:10.1002/med.21807
36. Jacobs, D. M. *et al.* NMR Backbone Assignment of the N-terminal Domain of Human HSP90. *J. Biomol. NMR* **36**, 52 (2006).
37. Park, S. J., Kostic, M. & Dyson, H. J. Dynamic interaction of Hsp90 with its client protein p53. *J. Mol. Biol.* **411**, 158–173 (2011).
38. Henot, F. *et al.* Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain. *J. Biomol. NMR* **75**, 221–232 (2021).
39. Kerfah, R., Plevin, M. J., Sounier, R., Gans, P. & Boisbouvier, J. Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr. Opin. Struct. Biol.* **32**, 113–122 (2015).
40. Amero, C. *et al.* A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J. Biomol. NMR* **50**, 229–236 (2011).
41. Sounier, R., Blanchard, L., Wu, Z. & Boisbouvier, J. High-accuracy distance measurement between remote methyls in specifically protonated proteins. *J. Am. Chem. Soc.* **129**, 472–473 (2007).
42. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).

43. Güntert, P., Mumenthaler, C. & Wüthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298 (1997).
44. Nichols, P. J. *et al.* The Exact Nuclear Overhauser Enhancement: Recent Advances. *Molecules* **22**, (2017).
45. Brünger, A. T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Cryst* **54**, 905–921 (1998).
46. Lente, G., Fábrián, I. & Poë, A. J. A common misconception about the Eyring equation. *New J. Chem.* **29**, 759–760 (2005).
47. Hopf, T. A. *et al.* The EVcouplings Python framework for coevolutionary sequence analysis. *Bioinformatics* **35**, 1582–1584 (2019).
48. Venditti, V., Tugarinov, V., Schwieters, C. D., Grishaev, A. & Clore, G. M. Large interdomain rearrangement triggered by suppression of micro-to millisecond dynamics in bacterial Enzyme I. *Nat. Commun.* **6**, (2015).
49. Hoter, A., El-Sabban, M. E. & Naim, H. Y. The HSP90 family: Structure, regulation, function, and implications in health and disease. *Int. J. Mol. Sci.* **19**, (2018).
50. Hessling, M., Richter, K. & Buchner, J. Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90. *Nat. Struct. Mol. Biol.* **16**, 287–293 (2009).
51. McLaughlin, S. H., Ventouras, L. A., Lobbezoo, B. & Jackson, S. E. Independent ATPase activity of Hsp90 subunits creates a flexible assembly platform. *J. Mol. Biol.* **344**, 813–826 (2004).
52. Verba, K. A. *et al.* Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* (80-.). **352**, 1542–1547 (2016).
53. Lee, K. *et al.* The structure of an Hsp90-immunophilin complex reveals cochaperone recognition of the client maturation state. *Mol. Cell* **81**, 3496–3508.e5 (2021).
54. Vögeli, B. *et al.* Towards a true protein movie: A perspective on the potential impact of the ensemble-based structure determination using exact NOEs. *J. Magn. Reson.* **241**, 53–59 (2014).
55. Amero, C. *et al.* Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J. Am. Chem. Soc.* **131**, 3448–3449 (2009).
56. Tugarinov, V., Kay, L. E., Ibraghimov, I. & Orekhov, V. Y. High-resolution four-dimensional ¹H-¹³C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J. Am. Chem. Soc.* **127**, 2767–2775 (2005).
57. Törner, R., Awad, R., Gans, P., Brutscher, B. & Boisbouvier, J. Spectral editing of intra- and inter-chain methyl–methyl NOEs in protein complexes. *J. Biomol. NMR* **74**, 83–94 (2020).
58. Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V. & Kay, L. E. Probing Slow Dynamics in High Molecular Weight Proteins by Methyl-TROSY NMR Spectroscopy: Application to a 723-Residue Enzyme. *J. Am. Chem. Soc.* **126**, 3964–3973 (2004).
59. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
60. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted

- from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
61. Stebbins, C. E. *et al.* Crystal structure of an Hsp90-geldanamycin complex: Targeting of a protein chaperone by an antitumor agent. *Cell* **89**, 239–250 (1997).
62. Nederveen, A. J. *et al.* RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins Struct. Funct. Genet.* **59**, 662–672 (2005).
63. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
64. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
65. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
66. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, 537–541 (2012).
67. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
68. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, (2007).
69. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182 (1981).
70. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
71. Páll, S. & Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650 (2013).
72. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
73. Tubiana, T., Carvaille, J. C., Boulard, Y. & Bressanelli, S. TTClust: A Versatile Molecular Simulation Trajectory Clustering Program with Graphical Summaries. *J. Chem. Inf. Model.* **58**, 2178–2182 (2018).
74. Kabsch, W. & Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **22**, 2577–2637 (1983).
75. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
76. Bouvignies, G. ChemEx - GitHub. Available at: <https://github.com/gbouvignies/ChemEx>. (Accessed: 2nd September 2021)

Chapitre 4

Modulation du paysage énergétique du domaine N-terminal de HSP90 par la liaison de différents inhibiteurs

4.1 HSP90 et cancer

4.1.1 HSP90 : une cible thérapeutique majeure

Comme mentionné dans le chapitre 3, HSP90 est une protéine chaperonne dynamique engagée à la fois dans le repliement, la stabilisation et la prévention de l'agrégation d'une multitude de protéines clientes. Parmi les protéines clientes de HSP90, on retrouve également un nombre important d'oncoprotéines impliquées dans les marqueurs caractéristiques du cancer – notamment dans ceux reportant sur les signaux de croissance (ErbB2, Src, Yes), l'apoptose de la cellule (p53) et la limite de prolifération (Miyata et al. 2012). Non seulement HSP90 assiste au repliement et à la stabilisation d'oncoprotéines clientes impliquées dans les marqueurs caractéristiques du cancer, mais il a également été reporté qu'un niveau élevé de protéine HSP90 était observé dans un grand nombre de cancers suggérant un rôle primordial de HSP90 dans la survie et la croissance de cellules cancéreuses (Whitesell and Lindquist 2005; Calderwood et al. 2006). Cette observation a été faite en particulier pour les cancers du sein (Pick et al.

2007), de la vessie (Lebret et al. 2003), des poumons (Shimamura and Shapiro 2008) et pour les mélanomes (McCarthy et al. 2008) et souvent de mauvais pronostics sont indiqués dans les cancers où le niveau de HSP90 augmente.

Concevoir qu'HSP90 puisse être une cible thérapeutique majeure pour lutter contre le cancer a été difficile puisqu'en effet il a été montré qu'une délétion du gène d'HSP90 était létal pour les eucaryotes (Borkovich et al. 1989; Aligue et al. 1994) cependant, la découverte du premier inhibiteur naturel d'HSP90 en 1994 : la geldanamycine a remédié à ce problème et depuis HSP90 est identifiée comme cible cruciale pour lutter contre cette maladie complexe (Whitesell et al. 1994; Chiosis and Neckers 2006). Dès lors, différents inhibiteurs d'HSP90 se liant au site de fixation de l'ATP ont été optimisés puis développés. Ces molécules, qu'elles soient naturelles ou synthétiques, inhibent l'activité ATPase de la chaperonne (Whitesell and Lindquist 2005) bloquant ainsi son cycle fonctionnel l'empêchant donc de replier et stabiliser les oncoprotéines. Toutefois, HSP90 est une protéine essentielle pour l'homme et les cellules saines requièrent elles aussi le soutien de la protéine chaperonne. On peut donc se demander pourquoi les inhibiteurs d'HSP90 sont plus sélectifs pour les cellules cancéreuses que pour les cellules saines puisque si cela n'était pas le cas le développement d'inhibiteurs d'HSP90 aurait été arrêté. Or, depuis 1998 un large nombre de ces derniers est entré en essai clinique et même si à l'heure actuelle aucun d'entre eux n'est approuvé, 11 sont toujours présentement en essai (Yuno et al. 2018; Zuehlke et al. 2018; Wang et al. 2021). Apporter une réponse à cette question soulevée est compliquée, toutefois plusieurs hypothèses pour répondre à cette interrogation sur la sélectivité tumorale des inhibiteurs d'HSP90 ont émergé. La première réside dans le fait que les conditions environnementales présentes dans les cellules cancéreuses sont plus dures que dans les cellules saines. Ainsi, un pH plus bas ou encore une hypoxie les rendraient plus dépendantes à HSP90 (Solit and Chiosis 2008; Li et al. 2009). Additionnellement, il a été montré que plus les cellules prolifèrent rapidement, plus le rôle d'HSP90 leur est indispensable (Miyata et al. 2012). Les cellules cancéreuses seraient davantage dépendantes du soutien de HSP90 alors que les cellules saines nécessiteraient sa présence de façon plus modérée. Cette dépendance accrue des cellules tumorales est confirmée par le niveau d'expression d'HSP90 qui atteint entre 4 et 7 % du total des protéines cellulaires contre 1 à 2 % dans des conditions normales (Sanchez-Martin et al. 2020), justifiant possiblement une accumulation des inhibiteurs d'HSP90 dans les cellules tumorales plutôt que dans les cellules saines. Une seconde hypothèse soulevée par Kamal et al. (Kamal et al. 2003; Chiosis and Neckers 2006) est la présence d'HSP90 uniquement sous forme d'un complexe multimérique dans les cellules cancéreuses ayant une forte affinité pour les inhibiteurs et pour l'ATP tandis qu'HSP90 dans les cellules normales serait présente uniquement dans un état latent et non pas sous forme de complexe multimérique avec plusieurs autres co-chaperones. De ces deux postulats

et avec l'observation faite que diverses doses d'inhibiteurs n'occupant qu'une faible fraction des sites de fixation d'HSP90 permettaient d'inhiber la croissance des cellules tumorales, Chiosis et Neckers (Chiosis and Neckers 2006) ont suggéré que les inhibiteurs d'HSP90 ciblaient une conformation d'HSP90 faiblement peuplée et potentiellement uniquement présente lorsqu'HSP90 est dans un complexe multimérique.

4.1.2 Différents types d'inhibiteurs de HSP90

Trois familles majeures d'inhibiteurs, se liant au domaine NTD de HSP90 se distinguent : les inhibiteurs de type ansamycine à benzoquinone, les dérivés de résorcinol et les analogues de la purine même si toutefois d'autres chimiotypes ont également été reportés et sont entrés, pour certains, en essai clinique.

4.1.2.1 Inhibiteurs de type ansamycine

Les composés de la famille des ansamycines à benzoquinone ont été les premiers reportés comme inhibiteurs de HSP90 et à être entrés en essais cliniques. La geldanamycine (GM) (Whitesell et al. 1994) est le premier composé de cette famille à avoir montré des résultats probants à la fois in-vivo et in-vitro dans des essais précliniques. Toutefois, en raison d'une forte hépatotoxicité, d'une instabilité chimique et d'une solubilité médiocre la phase clinique I n'a jamais été atteinte pour ce composé. Le 17-AAG, un analogue de la GM différant uniquement par un substituant a permis de diminuer la toxicité du composé chimique et d'atteindre les phases I et II des essais cliniques. Cependant sa faible solubilité, obligeant une administration par intraveineuse employant à la fois le DMSO et un phospholipide d'œuf comme vecteurs a entraîné une toxicité, stoppant la progression de cet inhibiteur dans les différentes phases cliniques. Une nouvelle tentative d'inhibiteur naturel de type ansamycine a été essayée avec le développement du 17-DMAG. Ce dernier, plus soluble que le 17-AAG, délivrable oralement et présentant une activité antitumorale similaire a toutefois été éliminé des essais cliniques par Kosan pour une toxicité inacceptable (Kim et al. 2009). Le composé le plus prometteur de cette classe d'inhibiteur est le IPI-504 qui semble plus performant que le 17-AAG, probablement grâce au nombre accru de liaisons hydrogènes qu'il peut former avec le dérivé d'hydroquinone formé (Sydor et al. 2006; Ge et al. 2006) même si toutefois une toxicité de ce dernier a tout de même été relevée (Ochiana et al. 2014) (Figure 4.1).

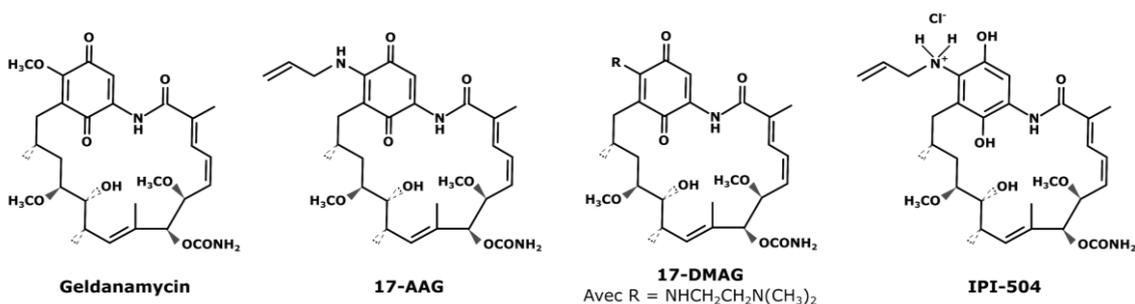


Figure 4.1 : Structures des différents inhibiteurs de type ansamycine développés. A gauche la geldanamycine, le premier inhibiteur reporté d’HSP90. Suivent dans l’ordre le 17-AAG, le DMAG et le IPI-504.

4.1.2.2 Dérivés de résorcinol

Une seconde famille importante d’inhibiteurs d’HSP90 est celle des dérivés de résorcinol, des composés chimiques contenant un groupement résorcinol. Parmi les premiers dérivés de résorcinol reportés on note le radicol (RD). Ce dernier, bien que montrant une forte activité in vitro n’a pas démontré d’activité in vivo. Cette observation est attribuée d’après Yang et Soga (Soga et al. 2003; Yang et al. 2004) à la forte réactivité des liaisons insaturées du carbonyle présentes dans sa structure qui sont métabolisées rapidement (Figure 4.2).

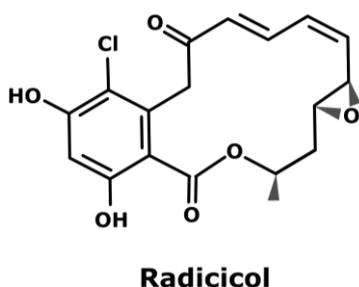


Figure 4.2 : Structure du radicol, inhibiteur d’HSP90 appartenant à la famille des dérivés de résorcinol.

La structure de base du RD permettant de se lier dans la poche de fixation de l’ATP avec des interactions clés reste un marqueur important des dérivés de la famille des résorcinol qui ont suivi. A l’heure actuelle au moins quatre dérivés de la famille des résorcinol ont atteint les phases I (KW-2478 et AT-13387) et II (AUY-922 et STA-9090) des essais cliniques (Figure 4.3) (Ochiana et al. 2014).

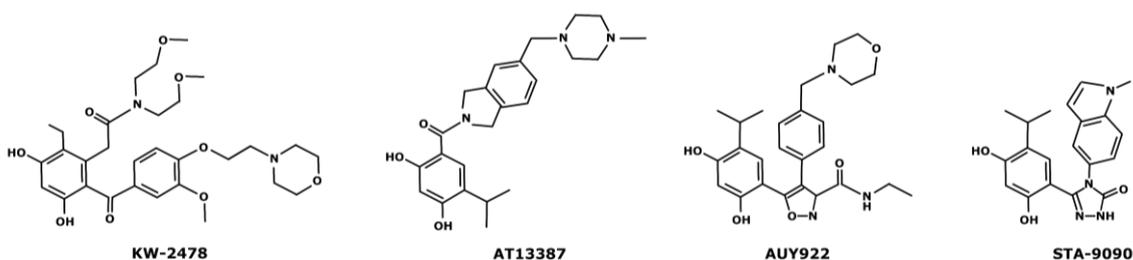


Figure 4.3 : Structures des différents inhibiteurs dérivés du résorcinol : KW-2478, AT13387, AUY922 et STA-9090.

4.1.2.3 Analogues de la purine

Le premier inhibiteur développé d'HSP90 appartenant à la famille des analogues de la purine est le PU3 en 2001. Ce dernier est le premier inhibiteur synthétique à montrer des propriétés phénotypiques similaires à la geldanamycine telles qu'une dégradation de HER2 et une inhibition de la croissance des cellules cancéreuses (Chiosis et al. 2001).

De nombreux groupes pharmaceutiques se sont inspirés de la structure du PU3 (Figure 4.4) pour développer de nouveaux inhibiteurs avec des propriétés pharmacocinétiques améliorées par rapports aux inhibiteurs de type ansamycine mais également reportant de plus faibles toxicités.

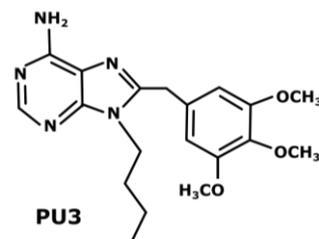


Figure 4.4 : Structure du premier inhibiteur synthétique d'HSP90 : PU3

Différentes modifications du squelette de PU3 ont été envisagées que ce soit sur le cycle purine, le cycle 8-aryl, le linker entre ces deux cycles ou le substituant alkyle en position 9. Les résultats montrent que les meilleurs candidats sont ceux pour lesquels des modifications du cycle 8-aryl et du substituant alkyle en position 9 ont été réalisées. Grâce à ces nombreux efforts quatre inhibiteurs dérivés de la purine sont entrés en essai cliniques avant 2014 : PU-H71, MPC-3100, BIIB021 et le Debio 0932 (Figure 4.5). Parmi ces quatre candidats médicaments, trois ont atteint la phase I des essais cliniques (PU-

H71, MPC-3100 et le Debio 0932), cependant seul un reste actif en phase I en 2022 (PU-H71). Concernant le BIIB021, la phase II des essais clinique a été atteinte pour deux études, dont l'une portant sur des patients atteints de tumeurs gastrointestinales montrant des résultats prometteurs toutefois la société Biogen Idec a décidé de s'éloigner des programmes de développements oncologiques (Ochiana et al. 2014).

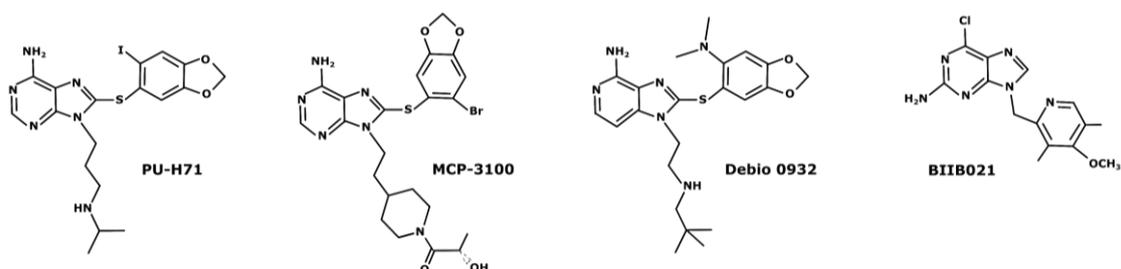


Figure 4.5 : Structures des différents inhibiteurs analogues de la purine : PU-H71, MCP-3100, Debio 0932 et BIIB021.

4.1.2.4 Stratégies de conception d'inhibiteurs de HSP90

Les différents types d'inhibiteurs ont été obtenus soit par criblage à haut débit (par exemple c'est le cas du dérivé de résorcinol AUY-922 (Figure 4.3)) soit par une approche rationnelle (notamment pour le dérivé de résorcinol AT-13387 (Figure 4.3) et les analogues de la purine : PU3 (Figure 4.4), PU-H71 (Figure 4.5) et MPC-3100 (Figure 4.5)). Concernant, l'optimisation du composé AT-13387, une conception employant la méthode des fragments a été utilisée. La RMN et la cristallographie aux rayons X à haut débit ont été employées pour, tout d'abord, trouver un fragment « point de départ ». Ce dernier a ensuite été optimisé pour donner un potentiel inhibiteur d'HSP90. Le passage du fragment initial en candidat idéal s'appuie sur un processus itératif nécessitant des données structurales. Pour ce qui est du développement des inhibiteurs PU3, PU-H71 et MPC-3100, les structures cristallographiques de HSP90-NTD liée à la GM, au RD et à l'ADP ont été utilisées. Ces dernières ont permis de révéler la présence d'un fold de Bergerat – une forme atypique de site de liaison de l'ATP présent uniquement pour la famille de protéines GHKL (Dutta and Inouye 2000). La disposition coudée de l'ATP dans ce fold de Bergerat a été employée pour développer l'inhibiteur PU3. De ce dernier se sont ensuite développés PU-H71 et MPC-3100.

Concernant le criblage à haut débit pour les inhibiteurs d'HSP90 différents criblages ont été reportés pouvant sonder un nombre considérable de candidats inhibiteurs (~56000 Rowlands et al. 2004). On note six criblages à haut débit reportant sur l'interaction

entre des ligands et une protéine d'HSP90 recombinante que ce soit humaine ou de levure et un criblage à haut débit employant des lysats de cellules cancéreuses humaines pour identifier de potentiels candidats médicaments plutôt que d'employer des protéines recombinantes (Du et al. 2007). Parmi les criblages utilisant des protéines recombinantes différentes techniques ont été utilisées. On note par exemple celui reportant le déplacement compétitif de l'inhibiteur [^3H]17AAG se liant au NTD de HSP90 α (Carreras et al. 2003), celui basé sur l'analyse de l'intensité de fluorescence d'un inhibiteur radicicol marqué lié à HSP90 β (Schilb et al. 2004), un essai colorimétrique ou un essai combinant une enzyme et la fluorescence pour mesurer l'activité ATPase de HSP90 de levure (Rowlands et al. 2004; Avila et al. 2006) ou encore des tests de polarisation de fluorescence (Kim et al. 2004; Howes et al. 2006) par exemple sondant la compétition de GM marquée à la protéine recombinante HSP90 α . Concernant le criblage employant des lysats de cellules cancéreuses un test de polarisation de fluorescence haut débit a été mis en place. Plus de 15000 composés ont été sondés résultant en 44 possibles « candidats médicaments » (Du et al. 2007).

4.2 HSP90 et inhibiteurs de type résorcinol

HSP90 est une protéine flexible capable d'adopter différentes conformations comme observé dans le chapitre 3 et la plasticité de cette dernière et notamment celle de ses poches de liaison, joue un rôle important dans sa capacité à se lier à ses partenaires, incluant les ligands à visée thérapeutique. Cependant, les méthodes actuelles de conception rationnelle de médicaments considèrent rarement ces différentes conformations et souvent une unique structure rigide de protéine est employée. Prendre en compte la dynamique des protéines, paramètre pour l'instant inexploité, pourrait permettre d'améliorer sensiblement le succès des méthodes rationnelles de conception de médicament. Un changement dans la dynamique de la protéine lors de la liaison du candidat médicament peut affecter à la fois la contribution enthalpique et entropique à l'énergie libre de Gibbs. Ainsi les contributions enthalpiques et entropiques peuvent participer à améliorer l'affinité de liaison entre un ligand et sa protéine cible mais de plus, la dynamique de la protéine peut également jouer un rôle central pour la cinétique. L'ajustement de la protéine aux candidats médicaments pourrait amener à un changement de vitesse de dissociation et ainsi impacter le temps de résidence des ligands un paramètre très important dans la conception de médicaments.

Grâce à une collaboration avec le Dr. Matthias Frech (Merck KGaA), j'ai pu avoir accès à un ensemble d'une quinzaine d'inhibiteurs de HSP90, appartenant à la famille des dérivés de résorcinol, pour mon étude. Ces ligands partagent tous un chimiotype commun (Figure 4.6) et sont des compétiteurs de l'ATP se liant au site de fixation de ce dernier.

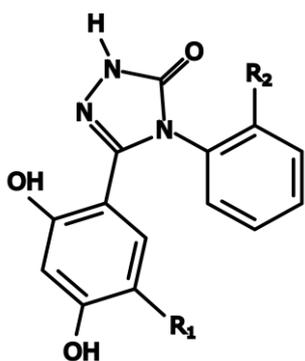


Figure 4.6 : Structure chimique commune de tous les ligands dérivés de la famille des résorcinols disponibles pour cette étude. R_1 et R_2 sont les deux substituants pouvant varier entre les différents inhibiteurs.

4.2.1 Deux familles de ligands

4.2.1.1 Une même affinité envers HSP90-NTD

Tous les inhibiteurs sélectionnés ont une affinité nanomolaire pour HSP90-NTD et ont une constante de dissociation de HSP90-NTD (K_D) similaire. K_D peut être décrite comme une fonction de l'énergie libre de Gibbs (ΔG) (Equation 4.1).

$$K_D = e^{\frac{-\Delta G}{RT}} \quad (\text{Equation 4.1})$$

ΔG peut être exprimé comme la somme de deux termes : la contribution enthalpique (ΔH) et la contribution entropique (ΔS) comme mentionné dans le chapitre 3. Dans le cas de la liaison d'un ligand à une protéine, et non plus de l'échange entre un état fondamental et un état excité d'une protéine Apo, ΔH mesure la chaleur associée à la liaison de l'inhibiteur. ΔS rend compte du changement dans le désordre de tout le système, incluant cette fois non seulement la protéine et le solvant mais également le ligand. La désolvatation des groupements polaires est généralement associée à une entropie favorable tandis qu'un changement entropique conformationnel du ligand est presque toujours considéré comme défavorable à la liaison (Freire 2008). Ainsi $\Delta G = \Delta H - T\Delta S$ et les ligands localisés sur la diagonale en pointillée représentant un ΔG constant ont la même affinité à HSP90-NTD (Figure 4.7).

4.2.1.2 Propriétés thermodynamiques et cinétiques différentes pour chaque catégorie de ligands

Toutefois, ces ligands diffèrent cinétiquement et thermodynamiquement et peuvent être répartis en deux catégories. D'un côté on observe des ligands pour lesquels la contribution enthalpique domine fortement la liaison et pour lesquels une cinétique de dissociation rapide de la protéine est observée, impliquant des temps de résidence de l'ordre de la seconde sur cette dernière (Figure 4.7, Figure 4.8, points verts). Dans une seconde catégorie, on observe des ligands pour lesquels la contribution entropique domine grandement la liaison et ces derniers montrent une cinétique de dissociation lente de leur cible, induisant des temps de résidence de l'ordre de la minute/heure sur le domaine N-terminal de HSP90 (Figure 4.7, Figure 4.8, points roses). Ces ligands seront par la suite reportés comme ligands à court temps de résidence « L_{tr}^{court} » et ligands à long temps de résidence « L_{tr}^{long} ».

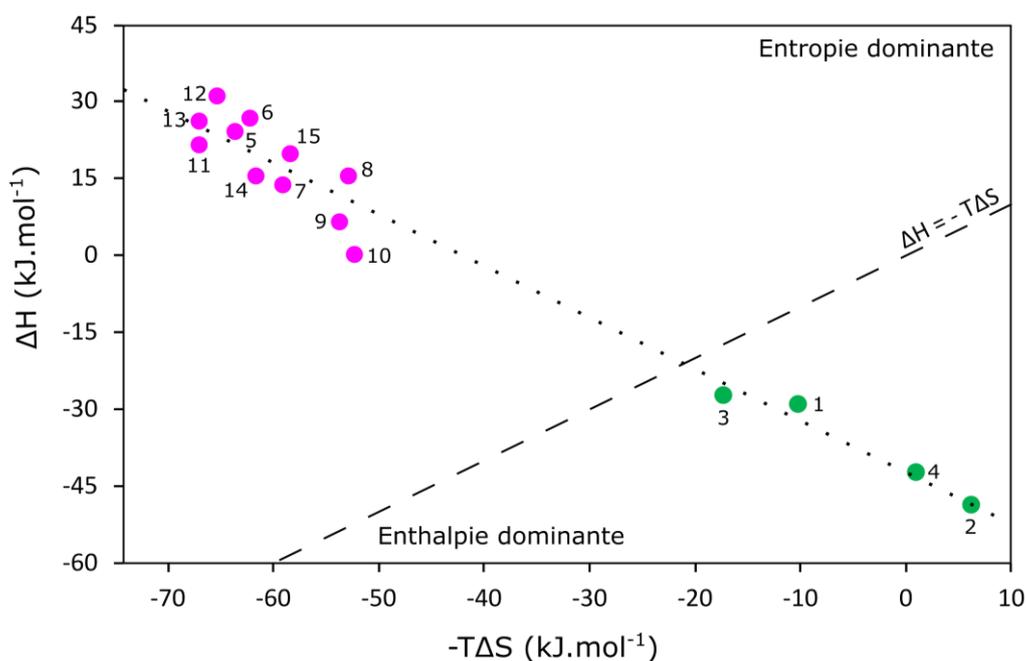


Figure 4.7 : Profil thermodynamique des différents inhibiteurs de HSP90-NTD mesurés par ITC. Les composantes enthalpiques et entropiques de l'énergie libre de Gibbs sont représentées pour chacun des inhibiteurs. En vert et rose respectivement les ligands dont la liaison est favorisée enthalpiquement et entropiquement. La droite en pointillée représente une énergie libre de Gibbs constante. Les tirets délimitent le graphique en deux zones. L'une où la contribution entropique domine la liaison et la seconde où la contribution enthalpique domine la liaison. Adapté de (Amaral et al. 2017).

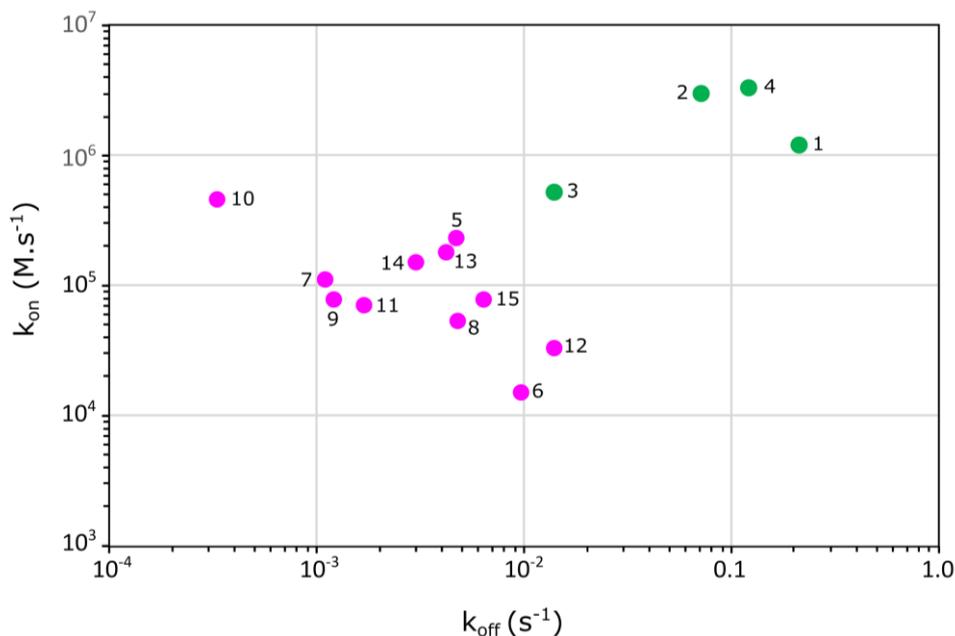


Figure 4.8 : Graphique logarithmique montrant la vitesse d'association des ligands (k_{on}) à HSP90-NTD en fonction de leur vitesse de dissociation (k_{off}) à la protéine cible. Les points verts et roses correspondent aux ligands dont la liaison est respectivement favorisée enthalpiquement et entropiquement.

Il est à noter, lorsque l'on compare la structure chimique des ligands des deux catégories, que le substituent R_1 (Figure 4.6) est décisif pour la répartition dans l'une ou l'autre des catégories. En effet, pour tous les ligands dont la liaison est favorisée enthalpiquement, le substituent R_1 est composé d'un unique atome. A l'inverse, pour tous les ligands dont la liaison est favorisée entropiquement le substituent R_1 est constitué, au minimum, d'une dizaine d'atomes. Le substituent R_2 , quant à lui, semble avoir peu d'influence sur la répartition entre les deux catégories. Pour tous les inhibiteurs R_2 est soit un atome de fluor ou de chlore soit un groupement méthyle ou éthyle.

4.2.1.3 Influence des ligands reportée dans la littérature lors de la liaison à HSP90-NTD

Une augmentation de la contribution entropique de 60 kJ.mol^{-1} a été observée dans le cas de la liaison de ligands de type « L_{TR}^{long} » par rapport aux ligands de type « L_{TR}^{court} ». Plusieurs mécanismes pourraient expliquer un tel changement de la composante entropique. En effet, une entropie favorable peut provenir d'un changement dans le nombre de molécules d'eau libérées lors de la liaison de l'inhibiteur à sa protéine cible, que ce soit provenant de la poche de liaison de la protéine ou de la couche d'hydratation du ligand mais un changement de cette composante peut également provenir d'un

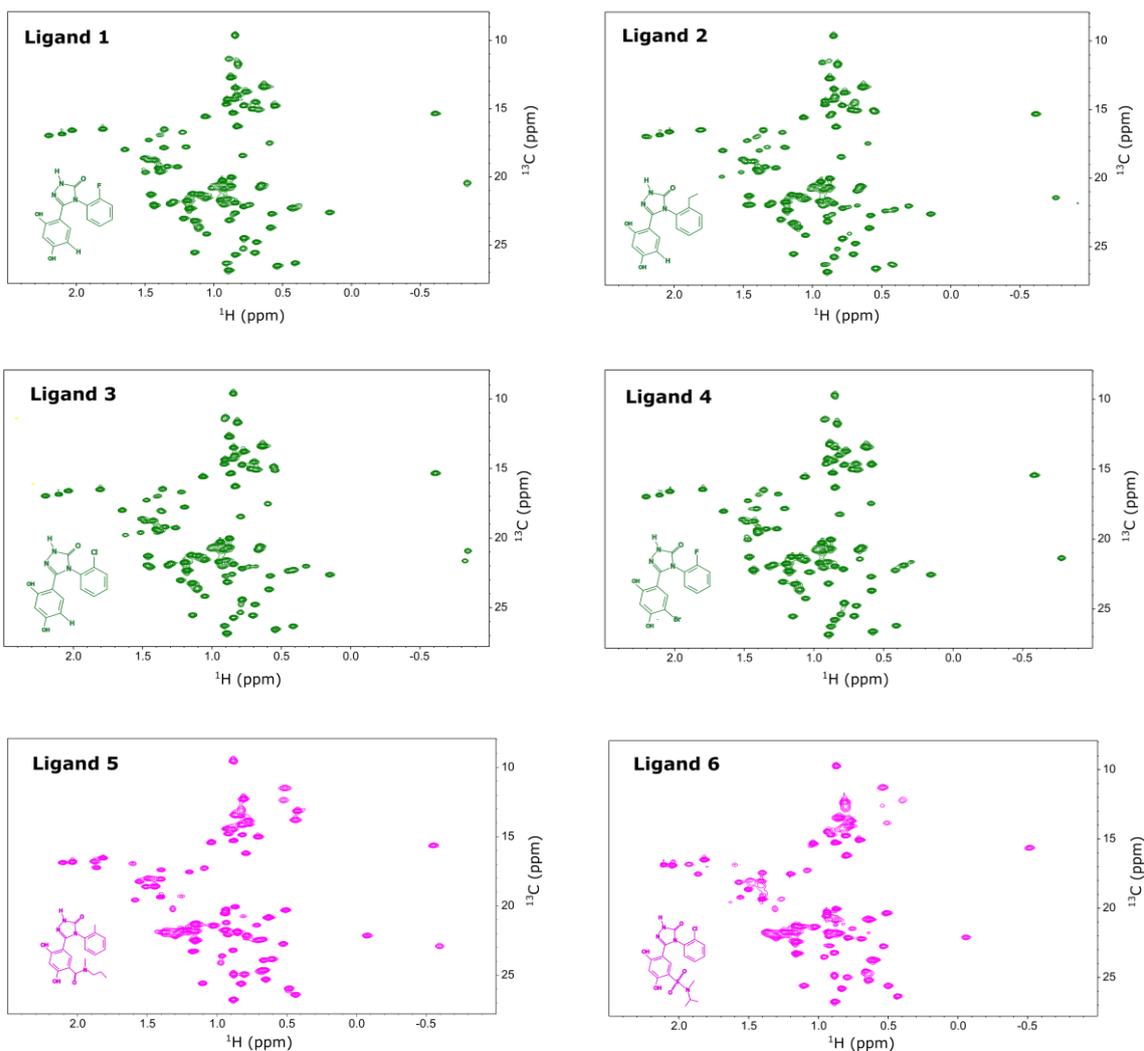
changement de la dynamique du ligand ou de la dynamique de la protéine cible elle-même.

Des études réalisées par Amaral et al (Amaral et al. 2017), employant la cristallographie aux rayons X et la dynamique moléculaire suggèrent que le nombre de molécules d'eau libérées de la poche de liaison de la protéine lors de la liaison de l'inhibiteur ne peut à lui seul expliquer la large différence d'entropie observée pour les deux types de candidats médicaments. En effet, 4 molécules d'eau supplémentaires sont piégées dans la poche de liaison lorsqu'un inhibiteur de type « L_{tr}^{Court} » se lie à HSP90-NTD comparé à un inhibiteur de type « L_{tr}^{Long} ». Or, une molécule d'eau piégée entraîne une pénalité entropique d'environ 8.5 kJ.mol^{-1} , ainsi uniquement un peu plus de la moitié de l'augmentation de la contribution entropique observée lors de la liaison avec un ligand de type « L_{tr}^{Long} » peut être expliqué par la libération des molécules d'eau provenant de la poche de la protéine. Les propriétés de désolvatation des inhibiteurs « L_{tr}^{Long} » et « L_{tr}^{Court} » ont également été évaluées. Pour des composés non-chargés (tels que les ligands étudiés), une mesure de l'énergie libre de désolvatation donne une estimation de la valeur du changement entropique due au solvant (Amaral et al. 2017). Il a été mesuré que pour les ligands « L_{tr}^{Long} » leur énergie libre de désolvatation est plus défavorable ($\sim 98 \text{ kJ.mol}^{-1}$) que pour les ligands « L_{tr}^{Court} » ($\sim 76 \text{ kJ.mol}^{-1}$) et ce résultat va à l'encontre des entropies dérivées des mesures par ITC. Ces différentes observations impliquent que la liaison des ligands à la protéine a une influence sur la dynamique de cette dernière. Ainsi, j'ai souhaité caractériser à la fois structurellement et dynamiquement HSP90-NTD liée à ces ligands « L_{tr}^{Court} » et « L_{tr}^{Long} ».

4.2.2 Sélection de ligands pour comparer les deux familles

Pour cela, j'ai sélectionné un ligand appartenant à chacune des deux catégories : un « L_{tr}^{Court} » et un « L_{tr}^{Long} ». Pour choisir ces ligands parmi la quinzaine dont je disposais je souhaitais que ceux choisis ne diffèrent qu'uniquement par un groupement et que le spectre 2D ^1H - ^{13}C HMQC de la protéine marquée méthyle liée au ligand donne un spectre de bonne qualité. Ainsi le choix s'est réduit de 15 ligands possibles (Figure 4.9) à 5. En effet, en observant les spectres 2D ^1H - ^{13}C HMQC de la protéine marquée méthyle liée aux ligands j'ai éliminé les ligands n° 8, 13 et 15 induisant soit des spectres avec des pics dédoublés ou avec des intensités de pics très variables. Je me suis ensuite focalisée sur les spectres de la protéine en complexe avec les ligands « L_{tr}^{Court} » et j'ai pu constater que lorsqu'elle est liée aux ligands n° 2 et 3 la protéine a des spectres montrant plus de résonances que le nombre de groupements méthyles marqués. J'ai donc choisi de sélectionner uniquement les ligands 1 et 4 pour HSP90-NTD en complexe avec un ligand « L_{tr}^{Court} ». Pour ces deux ligands le substituent R_2 était un atome de fluor. Ne souhaitant

comparer que des ligands différant par un groupement j'ai donc gardé uniquement les ligands n°10, 11 et 12 pour HSP90-NTD en complexe avec un ligand « L_{tr}^{Long} ». Parmi les 5 ligands restants (n° 1, 4, 10, 11 et 12) il y avait 2 « L_{tr}^{Court} » et 3 « L_{tr}^{Long} ». Je souhaitais que les deux ligands choisis aient des temps de résidence différents et grâce aux mesures de SPR reportées par Amaral et al. (Amaral et al. 2017) j'ai pu y avoir accès. J'ai donc sélectionné un ligand de type « L_{tr}^{Court} » ayant un temps de résidence sur HSP90-NTD de 5 secondes (ligand 1 Figure 4.9) et un inhibiteur de type « L_{tr}^{Long} » ayant un temps de résidence de 50 minutes sur sa protéine cible (ligand 10 Figure 4.9). Plus précisément ces deux ligands étaient le 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one (« L_{tr}^{Court} ») et le 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methylbenzamide (« L_{tr}^{Long} ») pour lesquels l'attribution a été présentée au chapitre 2 (Figure 4.10).



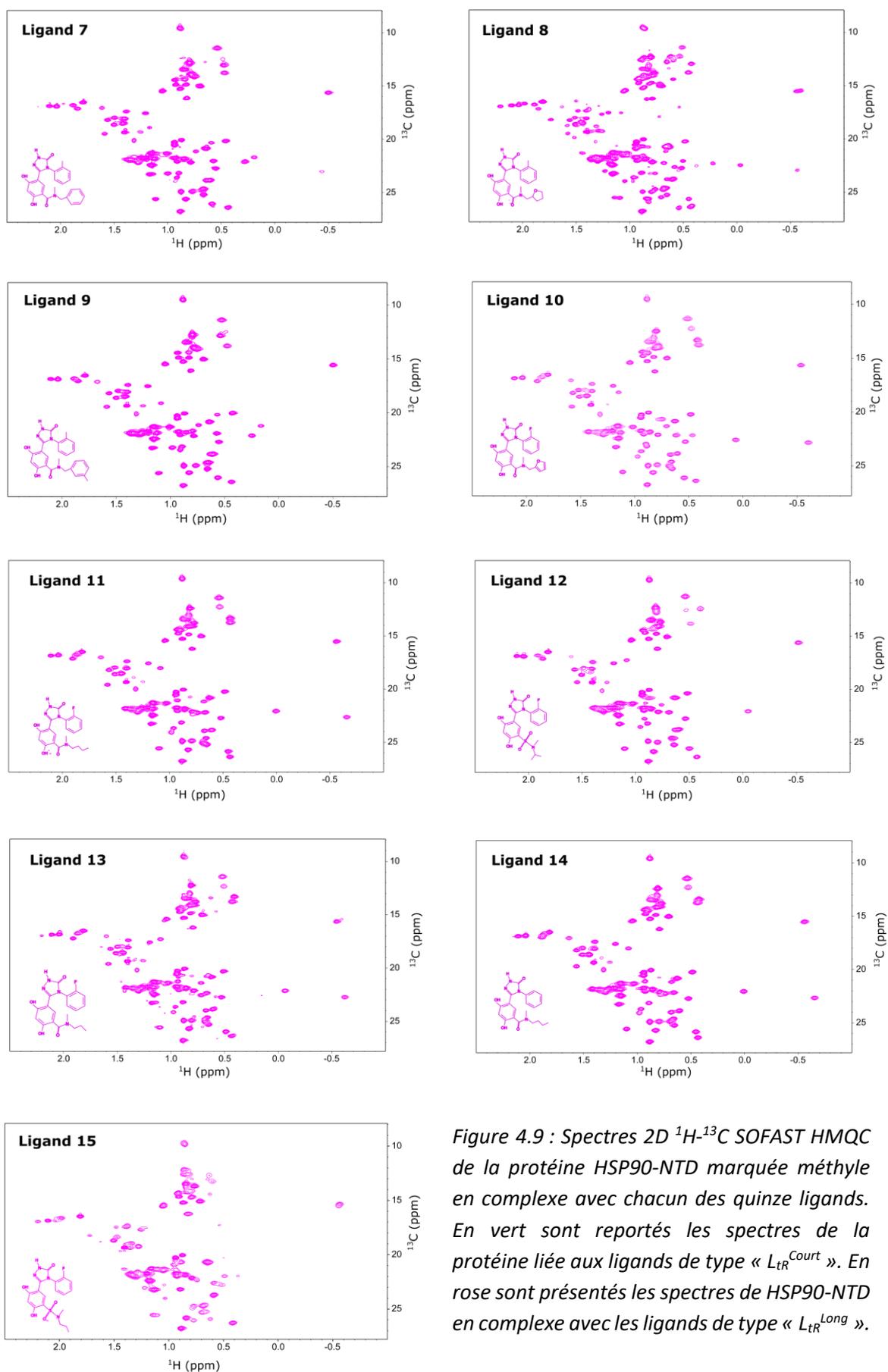


Figure 4.9 : Spectres 2D ^1H - ^{13}C SOFAST HMQC de la protéine HSP90-NTD marquée méthyle en complexe avec chacun des quinze ligands. En vert sont reportés les spectres de la protéine liée aux ligands de type « $L_{\text{tr}}^{\text{Court}}$ ». En rose sont présentés les spectres de HSP90-NTD en complexe avec les ligands de type « $L_{\text{tr}}^{\text{Long}}$ ».

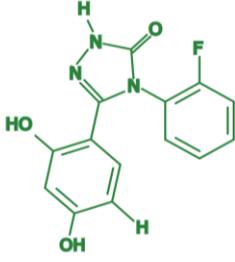
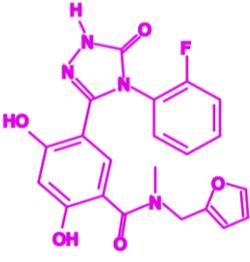
« L _{TR} ^{Court} »	« L _{TR} ^{long} »
Ligand 1	Ligand 10
	
5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-1,2,4-triazol-3-one	5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4]-triazol-3-yl]-N-furan-2-ylmethyl-2,4-dihydroxy-N-methyl-benzamide
$k_{on} = (1.2 \pm 0.2) \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$	$k_{on} = (4.6 \pm 0.5) \times 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$
$k_{off} = (2.1 \pm 0.3) \times 10^{-1} \text{ s}^{-1}$	$k_{off} = (3.3 \pm 0.2) \times 10^{-4} \text{ s}^{-1}$
$k_D = (1.8 \pm 0.1) \times 10^{-7} \text{ M}$	$k_D = (7.2 \pm 0.4) \times 10^{-10} \text{ M}$
$\Delta G = -39.1 \pm 0.3 \text{ kJ} \cdot \text{M}^{-1}$	$\Delta G = -52.2 \text{ kJ} \cdot \text{M}^{-1}$
$\Delta H = -29.1 \pm 1.1 \text{ kJ} \cdot \text{M}^{-1}$	$\Delta H = 0.0 \text{ kJ} \cdot \text{M}^{-1}$
$-T\Delta S = -10.1 \pm 1.3 \text{ kJ} \cdot \text{M}^{-1}$	$-T\Delta S = -52.2 \text{ kJ} \cdot \text{M}^{-1}$

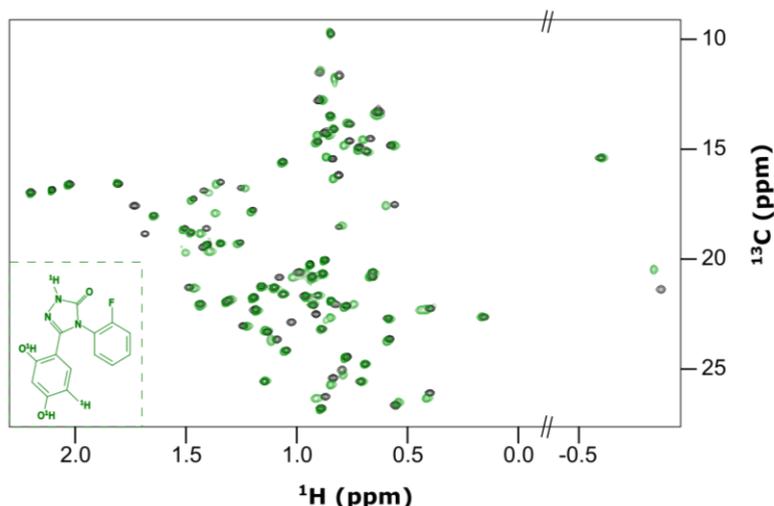
Figure 4.10 : Récapitulatif des différentes caractéristiques des deux ligands choisis. Pour le ligand 10, la chaleur associée à la liaison du ligand mesurée par ITC est proche de 0 à 298 K. Aussi, la valeur d'entropie a été déterminée en se basant sur l'affinité de liaison mesurée par SPR (Amaral et al. 2017).

4.3 Caractérisation de l'interaction du ligand à court temps de résidence avec HSP90

4.3.1 Etat fondamental

Comme mentionné au chapitre 2, HSP90-NTD a, lorsque liée au ligand « L_{TR}^{Court} », une empreinte spectrale similaire à celle de la protéine dans son état Apo (Figure 4.11) suggérant qu'elles ont des états fondamentaux similaires (*i.e.* avec l'« ATP-lid » en position ouverte). La valeur moyenne de la différence de déplacement chimique ($\Delta\delta_{\text{moyen}}$) entre la protéine Apo et la protéine en complexe avec le ligand 1, très faible (0.18 ppm) suggère également que les spectres de la protéine Apo et en complexe sont semblables. Avec $\Delta\delta_{\text{moyen}} = \sum_{i=1}^N ((\Delta\delta_{C\alpha(i)})^2 + (\Delta\delta_{C\beta(i)})^2 + (\Delta\delta_{C\gamma(i)})^2)^{\frac{1}{2}}$, N étant le nombre de résidus de la séquence de la protéine.

Figure 4.11 : Spectre ^1H - ^{13}C des groupements méthyles de la protéine HSP90-NTD Apo en noir superposés avec ceux de la protéine HSP90-NTD liée au ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluorophenyl)-2,4-dihydro-[1,2,4]triazol-3-one en vert à la température de 293 K. La structure de ce dernier est représentée en bas à gauche du spectre.



Pour valider cette hypothèse, une prédiction des structures secondaires de la protéine sous sa forme Apo et en complexe avec le ligand 1 a été effectuée avec TALOS + après l'attribution de leurs squelettes. Ces prédictions de structure secondaires montrent que ces dernières sont très semblables dans le cas de la protéine Apo ou en complexe confortant l'idée que HSP90-NTD Apo et liée au ligand « $L_{\text{tr}}^{\text{Court}}$ » partagent une structure similaire d'état fondamental (Figure 4.12).

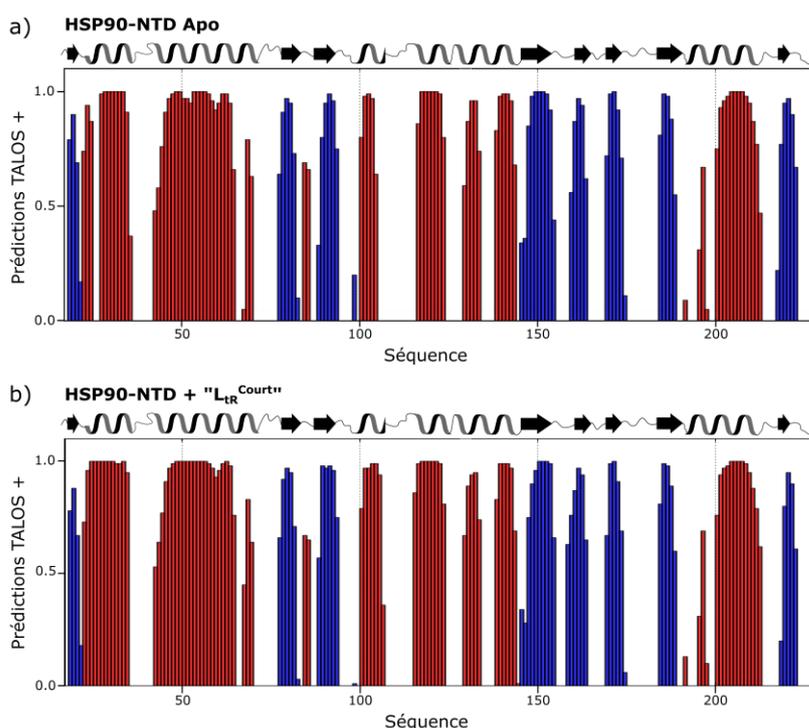


Figure 4.12 : Prédiction de structure secondaire données par le programme TALOS +. Les éléments de structure secondaire de types hélice α et feuillet β sont représentés respectivement en rouge et bleu pour a) HSP90-NTD Apo et b) HSP90-NTD en complexe avec le ligand « $L_{\text{tr}}^{\text{Court}}$ ».

4.3.2 Etat excité

Les expériences de relaxation dispersion montrant un échange conformationnel pour HSP90-NTD liée au ligand « L_{TR}^{Court} », un état excité existe pour HSP90-NTD en complexe. Une expérience 3D HMQC-NOESY-HMQC (éditée CCH) a été acquise sur un échantillon de HSP90-NTD perdeutérée, marquée spécifiquement sur six types de méthyles (Alanine- β , Méthionine- ϵ , Leucine-pro-S, Valine-pro-S, Isoleucine- δ_1 et Threonine- γ_2) et pour lequel un excès de ligand 1 (5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one) a été ajouté. De cette expérience NOESY 3D, 83 contraintes de distance ont été obtenues pour le segment [98-136] (segment pour lequel un ensemble structural a été calculé pour la protéine Apo au chapitre 2). Les NOEs caractéristiques de l'« état fermé » ont été observés (Figure 4.13), suggérant très fortement qu'HSP90-NTD Apo et en complexe avec le ligand de type « L_{TR}^{Court} » ne partagent pas uniquement un état fondamental similaire mais également un état excité semblable (l'« état fermé »).

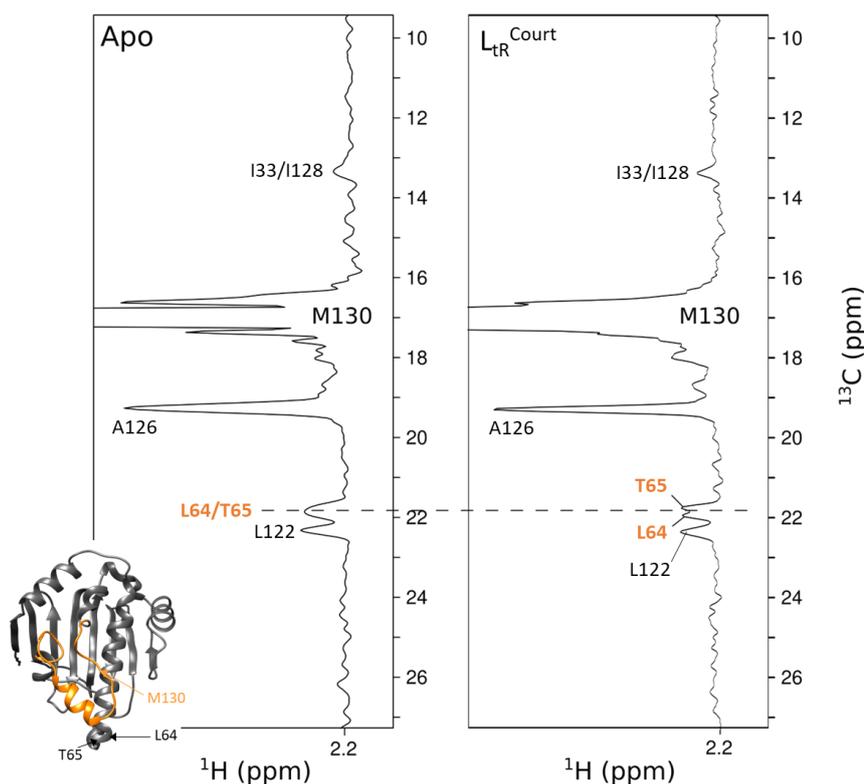


Figure 4.13 : Traces 1D extraites des expériences 3D CCH HMQC-NOESY-HMQC montrant M130 en pic diagonal ainsi que ses pics de corrélation NOEs associés pour la protéine sauvage HSP90-NTD Apo à gauche et en complexe avec le ligand 1 à droite. Les spectres NOESY ont été acquis en 2 à 3 jours sur des spectromètres opérant à des fréquences proton de 850 ou 950 MHz et à des températures de 298 K. Les temps de mélanges NOE étaient compris entre 300 et 400 ms. Les pics de corrélation M130-T65 et M130-L64 caractéristiques de l'état avec l'« ATP-lid » en position fermée sont également observés pour la trace 1D de HSP90-NTD en complexe.

4.3.3 Influence du ligand « L_{tr}^{Court} » sur la dynamique de HSP90-NTD à 293 K

En acquérant des expériences de CPMG-RD et en analysant les données avec un modèle à deux états en employant le même protocole que pour la protéine Apo, la constante de vitesse d'échange conformationnel ainsi que la population de l'état excité d'HSP90-NTD en complexe avec le ligand 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one ont pu être déterminées à 293 K. Il apparaît que le ligand a une influence sur la constante de vitesse d'échange conformationnel observée. En effet, la présence du ligand accélère l'échange entre l'« état ouvert » et l'« état fermé » comparée à la protéine Apo. Une constante de vitesse d'échange conformationnel de $3345 \pm 87 \text{ s}^{-1}$ est observée pour la protéine en complexe (Figure 4.14) contre une constante de vitesse d'échange conformationnel de $2490 \pm 61 \text{ s}^{-1}$ pour la protéine dans son état Apo. Concernant la population de l'état excité, il apparaît que la présence du ligand « L_{tr}^{Court} » entraîne une légère baisse de la population de l'état fermé comparé à la protéine Apo. La population passe de 3.2 % à 2.1 % en présence du ligand (Figure 4.14). On note que la valeur de k_1 , obtenue avec les valeurs de k_{ex} et P^* est, dans le cas de la protéine en complexe plus lente que celle obtenue pour la protéine Apo signifiant que l'énergie d'activation entre l'état fondamental et l'état de transition est plus basse dans le cas de la protéine Apo comparé à la protéine en complexe (61.0 kJ.mol^{-1} (Apo) versus 61.4 kJ.mol^{-1} (en complexe)). La valeur de k_{-1} quant à elle est accélérée dans le cas de la protéine en complexe par rapport à la protéine Apo, signifiant que l'énergie entre l'état excité et l'état de transition est plus faible dans le cas de la protéine en complexe (52.7 kJ.mol^{-1} (Apo) versus 52.0 kJ.mol^{-1} (en complexe)). Ce résultat est cohérent avec l'observation d'une plus faible population de l'état excité dans le cas de la protéine en complexe qui entraîne une diminution de la hauteur de la barrière énergétique entre l'état excité et l'état de transition.

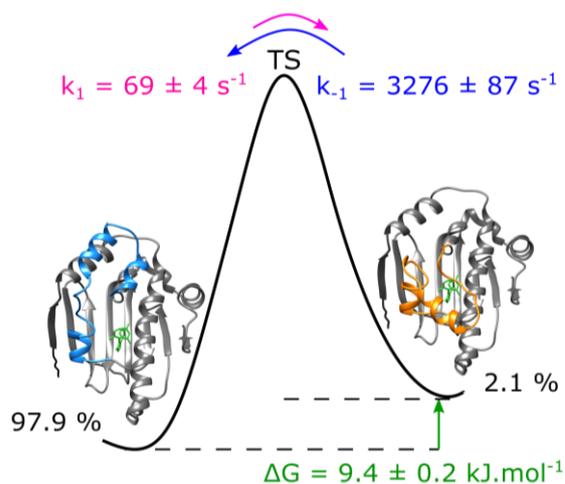


Figure 4.14 : Diagramme schématisique du paysage énergétique caractérisant l'échange entre l'état fondamental (ouvert) et l'état excité (fermé) de HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} ». Les paramètres cinétiques et thermodynamiques de cet échange, extraits à 293 K, sont reportés. Le ligand « L_{tr}^{Court} » est représenté en vert dans les structures tridimensionnelles.

4.3.4 Influence du ligand « L_{tR}^{Court} » sur la dynamique de HSP90-NTD sur la plage de température [278-313 K]

De la même manière que pour la protéine Apo j'ai acquis et analysé des expériences de relaxation dispersion pour la protéine en complexe pour 8 températures différentes allant de 278 K à 313 K avec un pas de 5 K. Le but final de ces expériences était d'extraire des informations sur la thermodynamique entre l'état fondamental et l'état excité.

En premier lieu, j'ai pu constater que l'évolution de la constante de vitesse d'échange conformationnel pour la protéine en complexe avec le ligand « L_{tR}^{Court} » augmentait avec la température – une évolution qui apparait cohérente avec les principes de la thermodynamique où généralement l'augmentation de la température augmente l'agitation thermique pour les molécules présentes dans un liquide et ainsi la constante de vitesse d'échange conformationnel (Figure 4.15). Il semble également qu'un plateau est atteint aux températures de 308 et 313 K, les deux derniers points, cependant cela n'est pas suffisant pour pouvoir en tirer une conclusion.

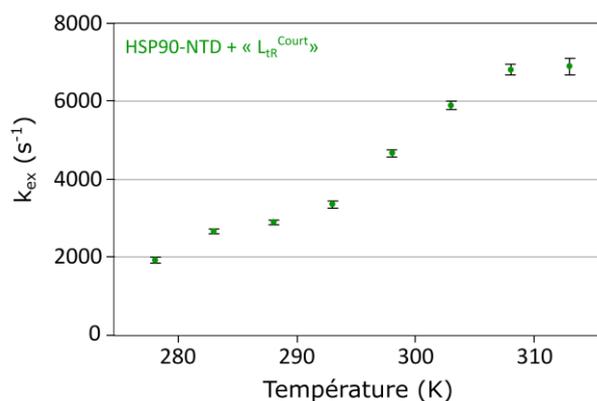
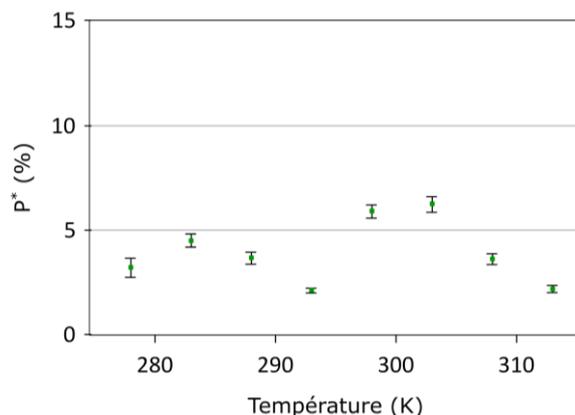


Figure 4.15 : Graphique représentant l'évolution de la constante de vitesse d'échange conformationnel (k_{ex}) en fonction de la température pour HSP90-NTD en complexe avec le ligand « L_{tR}^{Court} ». La précision sur les données a été estimée en utilisant une analyse Monte-Carlo.

Concernant les populations des états excités j'ai pu voir qu'en présence du ligand « L_{tR}^{Court} » la population de l'état excité en fonction de la température évoluait de façon complexe. En effet, la population de l'état excité augmente de 278 K à 283 K puis diminue de 283 K à 293 K avant d'augmenter de nouveau de 293 K à 303 K puis de diminuer de 303 K à 313 K (Figure 4.16). En observant ces données j'ai eu l'impression que deux phénomènes se produisaient, un à basse température entre 278 K et 293 K et un second à plus haute température entre 298 K et 313 K.

Figure 4.16 : Graphique montrant l'évolution de la population de l'état excité de HSP90-NTD en complexe avec le ligand « L_{tR}^{Court} » en fonction de la température. La précision sur les données a été estimée en utilisant une analyse Monte-Carlo.



A haute température, sur la gamme [298-313 K], les valeurs de k_{ex} dépassent 4500 s^{-1} . Plusieurs études suggèrent que lorsque l'on tend vers un régime d'échange rapide (*i.e.* que l'on tend vers $k_{ex}/\Delta\omega \gg 1$), même si deux champs magnétiques sont employés, décorréler la population de l'état excité des valeurs absolues de différences de déplacements chimiques devient complexe (Kleckner and Foster 2011; Vallurupalli et al. 2011). Ne pouvant extraire des ajustements globaux des valeurs précises de $|\Delta\omega|$ (voir chapitre 3) il est difficile d'obtenir une estimation de la valeur de $k_{ex}/\Delta\omega$ pour chaque sonde percevant l'échange et ce à chaque température. Cependant, en comparant qualitativement les moyennes des valeurs de $k_{ex}/\Delta\omega$ obtenues à 293 K et 298 K j'ai pu observer que cette moyenne était multipliée par plus de 2 à 298 K comparée à la moyenne obtenue à 293 K. Aussi, avec des valeurs de k_{ex} au-delà de $\sim 3500/4000 \text{ s}^{-1}$ pour HSP90-NTD en complexe, il semblerait que l'on tend vers le régime d'échange rapide et qu'en employant uniquement les courbes de relaxation dispersion il ne soit pas possible d'obtenir des valeurs de population d'état excité et de différences de déplacements chimiques décorréliées. Des approches pour pallier ce problème seront présentées en 4.5, mais notamment obtenir de façon externe les valeurs de différences de déplacement chimique permettrait d'obtenir les valeurs de la population d'état excité à hautes températures. Ainsi, seules les valeurs de population d'état excité entre 278 K et 293 K seraient exploitables mais déjà sur cette gamme de température, la population de l'état excité ne varie pas linéairement et éventuellement différents phénomènes pourraient être présents.

4.4 Caractérisation de l'interaction du ligand à long temps de résidence avec HSP90

4.4.1 Etat fondamental – une nouvelle structure

Contrairement au ligand « L_{TR}^{Court} », HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) - 5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide (« L_{TR}^{Long} ») montre des empreintes spectrales des groupements amides et des méthyles très différentes de la protéine Apo (Figure 4.17). Suggérant fortement que le ligand induit une conformation différente du domaine N-terminal de la protéine.

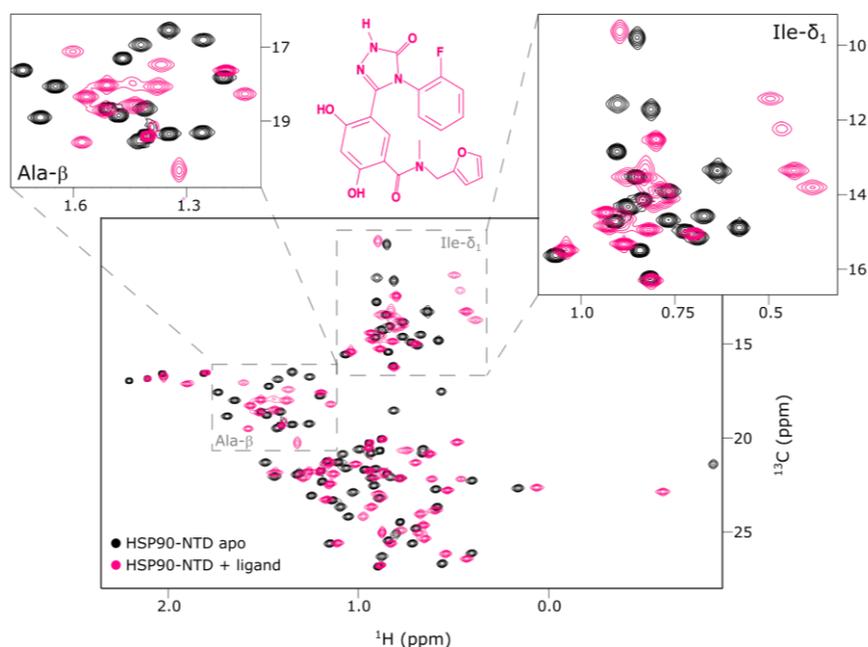
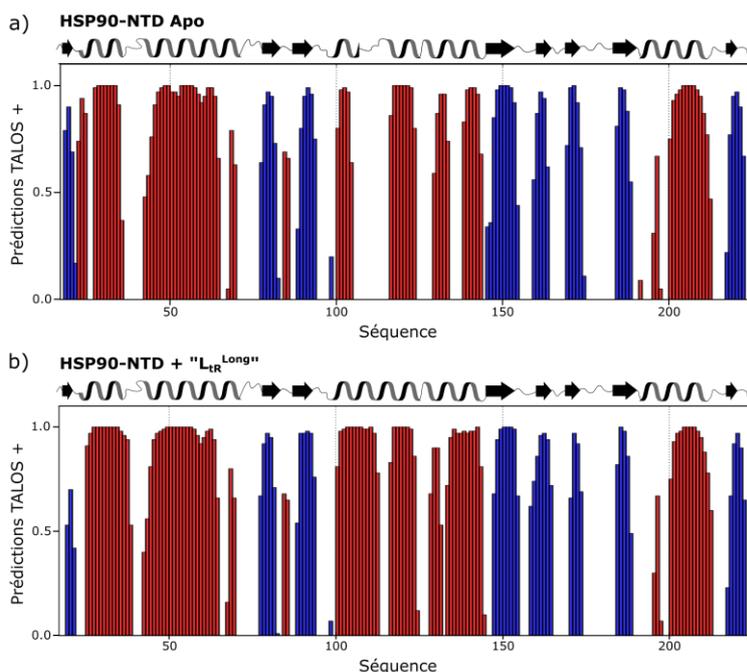


Figure 4.17 : Spectre 1H - ^{13}C des groupements méthyles de la protéine HSP90-NTD Apo en noir superposés avec ceux de la protéine HSP90-NTD liée au ligand 5-[4-(2-Fluoro-phenyl) - 5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide en rose à la température de 293 K.

La valeur moyenne de la différence de déplacement chimique ($\Delta\delta_{moyen}$) entre la protéine Apo et la protéine en complexe avec le ligand 10 est de 0.84 ppm, une valeur quatre fois plus élevée que lorsque la protéine Apo est comparée à la protéine en complexe avec le ligand 1. Cette valeur élevée suggère que les spectres de la protéine Apo et en complexe avec le ligand 10 sont effectivement différents. Avec $\Delta\delta_{moyen} = \sqrt{\sum_{i=1}^N ((\Delta\delta_{C\alpha(i)})^2 + (\Delta\delta_{C\beta(i)})^2 + (\Delta\delta_{C\gamma(i)})^2)}^{\frac{1}{2}}$, N étant le nombre de résidus de la séquence de la protéine.

J'ai attribué le squelette et les groupements méthyles d'HSP90-NTD en complexe à 308K (Chapitre 2 – Article II). Grâce à l'attribution de 92 % des résidus du squelette et notamment du segment [L107-K116], auparavant invisible pour la protéine Apo, j'ai pu conclure, avec une prédiction de la structure secondaire de la protéine en complexe avec le ligand 10 effectuée avec TALOS +, que le segment à proximité du site de fixation du ligand avait une structure secondaire en hélice (Figure 4.18).

Figure 4.18 : Prédictions de structure secondaire données par le programme TALOS +. Les éléments de structure secondaire de types hélice α et feuillet β sont représentés respectivement en rouge et bleu pour a) HSP90-NTD apo et b) HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ».



Ainsi, l'addition d'un ligand de type « L_{tr}^{Long} » induit de larges réarrangements structuraux dans la région [98-136]. Une structure de HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide est disponible dans la PDB (PDB : 5J20). Cette dernière a été obtenue par cristallographie aux rayons X et possède une structure secondaire en hélice pour le segment [100-123] (Figure 4.19). J'ai pu en conclure que cette structure est bien représentative de la structure de HSP90-NTD en complexe en solution à température de 308 K. Ainsi l'état fondamental (majoritaire) de HSP90-NTD Apo et en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide sont, comme présumé en observant les spectres RMN, différents.

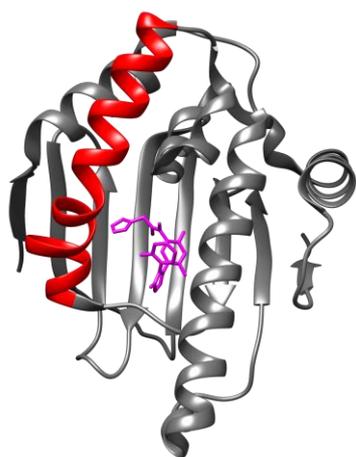


Figure 4.19 : Structure cristallographique tridimensionnelle de HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide (PDB : 5J20). En rose et rouge sont représentés respectivement, le ligand et la structure secondaire en hélice du segment [100-123]

4.4.2 Etat excité

De la même manière que pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} », les expériences de relaxation dispersion montrent un échange conformationnel pour HSP90-NTD liée au ligand « L_{tr}^{Long} ». Ainsi, un état excité existe également pour HSP90-NTD liée au ligand « L_{tr}^{Long} ». Pour obtenir des informations sur ce dernier, deux expériences 3D HMQC-NOESY-HMQC (éditées CCH) avec deux temps de mélange différents ont été acquises à 308 K sur un échantillon de HSP90-NTD perdeutérée, marquée spécifiquement sur six types de méthyles (Alanine- β , Méthionine- ϵ , Leucine-pro-S, Valine-pro-S, Isoleucine- δ_1 et Threonine- γ_2) et pour lequel un excès de ligand 10 : 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide a été ajouté.

Aucun, des 5 NOEs caractéristiques de l' « état fermé » n'est détecté pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} », quel que soit le temps de mélange de l'expérience. Or les expériences de relaxation dispersion montrant un échange conformationnel pour HSP90-NTD en complexe reportaient une population élevée d'état excité ($\sim 7\%$ à 308 K pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} » (dans le régime d'échange intermédiaire avec $k_{ex} \sim 1625\text{ s}^{-1}$), contre 3% pour la protéine Apo à 293 K). Aussi, de cette observation j'ai pu conclure que HSP90-NTD Apo et en complexe avec le ligand « L_{tr}^{Long} » avaient des états excités (minoritaires) différents.

J'ai choisi d'étudier plus en détails les pics de corrélation NOEs dans le but d'obtenir des informations sur l'état excité de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} » et pour cela j'ai utilisé le logiciel CYANA. Les distances $C_\alpha-C_\alpha$ ainsi que les angles dièdres correspondant à la structure cristallographique avec le ligand 10 (PDB : 5J20) ont été extraits de cette dernière puis utilisés en contraintes fixes. Les contraintes de distance obtenues à partir des pics de corrélation NOEs observés dans les spectres ont ensuite

été ajoutées. Une approche qualitative a été employée pour la dérivation des contraintes de distance à partir des NOE observés. Les contraintes de distance interméthyles ont été fixées entre les barycentres de chaque triplet de protons avec des distances comprises entre 2 et 6 Å pour les pics de corrélation détectés dans l'expérience 3D-NOESY avec un temps de mélange court (120 ms) et des distances comprises entre 4 et 8 Å pour les pics de corrélation observés uniquement dans l'expérience 3D-NOESY avec un temps de mélange plus long (360 ms). 113 contraintes de distance non-ambigües et 10 ambigües ont été dérivées des NOEs détectés dans l'expérience acquise avec un temps de mélange de 120 ms. 67 contraintes de distance supplémentaires non-ambigües et 21 ambigües ont été dérivées des pics de corrélation NOEs détectés avec l'expérience acquise avec un temps de mélange de 360 ms.

Parmi ces 211 contraintes de distance interméthyles, toutes sauf 2 sont en adéquation avec la structure cristallographique (PDB : 5J20). Ces 2 contraintes de distance violées proviennent de pics de corrélation NOEs observés dans le spectre NOESY acquis avec un temps de mélange long. Les NOEs observés ont été revérifiés ainsi que leur attribution. Les deux contraintes de distance violées (109-131 et 115-124), le sont respectivement de 2.2 et 0.9 Å. La structure cristallographique fixant l'orientation des chaînes latérales, des tests enlevant la fixation des angles χ_1 pour les résidus concernés par les violations ont été effectués, cependant, même ainsi, les 2 contraintes de distances restent violées de 2.1 et 0.9 Å.

Plusieurs hypothèses pourraient expliquer la présence de ces 2 contraintes de distance. La première est que ces deux pics de corrélation NOEs sont caractéristiques de l'état excité de HSP90-NTD en complexe avec le ligand « L_{TR}^{LONG} ». Mon hypothèse, décrivant le modèle le plus simple qui puisse être évoqué actuellement, est que l'état excité/minoritaire de HSP90-NTD en complexe avec le ligand « L_{TR}^{LONG} » est l'« état ouvert » (l'état majoritaire peuplé par HSP90-NTD Apo). Les structures de l'état fondamental de HSP90-NTD en complexe et de celle de son état excité hypothétique, structurellement très proches, justifieraient l'absence d'un nombre important de pics caractéristiques de l'état excité. En effet, la plupart de ces derniers ne seraient pas distinguables de ceux caractéristiques de l'état fondamental, aussi un ensemble majoritaire de pics satisfaisant les deux conformations à la fois serait obtenu. Pour tester cette hypothèse, les distances $C_\alpha-C_\alpha$ ainsi que les angles dièdres correspondant à la structure moyenne des 20 meilleurs conformères de l'état ouvert d'HSP90-NTD Apo ont été extraits de cette dernière puis utilisés en contraintes fixes. J'ai pu constater que pour les deux contraintes de distance 109-131 et 115-124 la violation diminuait. Passant de 2.2 et 0.9 Å à, respectivement 1.2 et 0.9 Å lorsque les angles χ_1 étaient fixés et à 0.6 et 0.8 Å lorsque ces derniers étaient laissés libres en utilisant la structure de l'état ouvert de HSP90-NTD Apo. Ces résultats montrent que l'état excité de HSP90-NTD en complexe

avec le ligand « L_{TR}^{Long} » est possiblement proche structurellement de l'état ouvert de HSP90-NTD apo. Toutefois, des violations persistent. La diffusion de spin pourrait être en partie responsable des violations. En effet, ces 2 contraintes de distance violées sont observées uniquement pour le NOESY acquis avec un long temps de mélange. Par exemple possiblement l'isoleucine 110, ayant à la fois un pic de corrélation NOE avec la thréonine 109 et l'isoleucine 131 servirait de relais pour expliquer que la thréonine 109 ait un pic de corrélation NOE avec l'isoleucine 131.

D'autres données, allant dans le même sens que les hypothèses précédentes et pouvant expliquer ces contraintes de distance est la présence de dynamique. L'analyse du B-facteur (ou facteur de Debye-Waller) reporté pour la structure cristallographique de HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide montre que les segments [119-127], [176-180] et [208-213] ont une flexibilité/hétérogénéité structurale plus importante que 1.5 fois la moyenne de la protéine (Figure 4.20). Parmi ces régions seul le segment [119-127] ne correspond pas à une boucle dynamique mais à un segment avec une structure secondaire en hélice et comprend le résidu 124 montrant des pics de corrélation NOEs convertibles en une contrainte de distance violée. Ce segment [119-127] est également à proximité des résidus 115 et 131 impliqués dans des pics de corrélation NOEs convertibles en contraintes de distance violées.

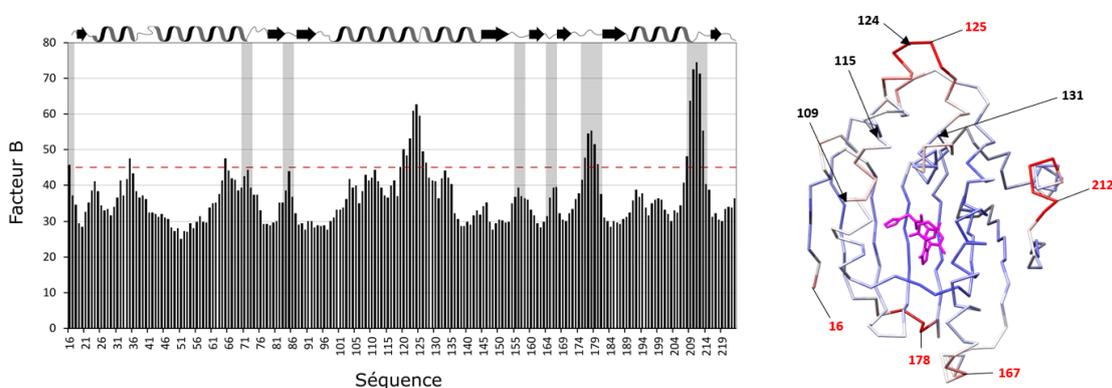


Figure 4.20 : A gauche, histogramme reportant le facteur B en fonction de la séquence de HSP90-NTD extrait de la structure cristallographique PDB : 5J20. Le trait pointillé rouge représente 1.5 fois le B-facteur moyen de la protéine. En gris sont représentées les zones de boucles flexibles où le B-facteur est au-dessus de la moyenne. A droite : représentation tridimensionnelle du squelette de HSP90-NTD (PDB : 5J20) coloré en fonction de la valeur du facteur B. De bleu à rouge le facteur B augmente. En rose est reporté le ligand 5-[4-(2-Fluoro-phenyl) -5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. En rouge sont reportés les résidus appartenant à des zones où le facteur B est élevé. En noir sont reportés les résidus montrant des pics de corrélation NOEs convertibles en contraintes de distance violées.

Pour la suite de ma thèse, le modèle le plus simple dans lequel l'état excité de la protéine en complexe avec le ligand « L_{tr}^{Long} » est l'état ouvert de la protéine Apo sera employé. D'après les résultats précédents avec les pics de corrélation NOEs cette hypothèse semble plausible mais n'est toutefois pas complètement vérifiée.

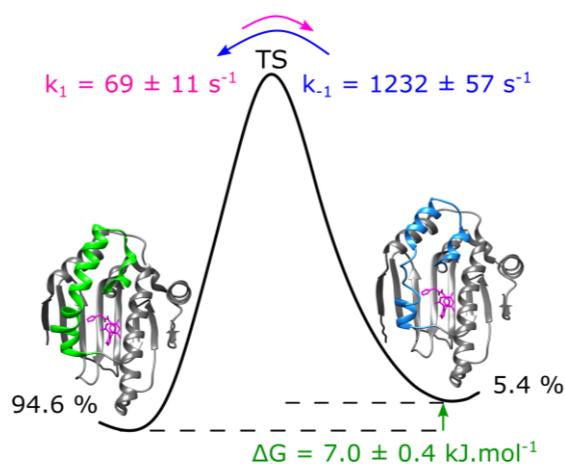
4.4.3 Influence du ligand « L_{tr}^{Long} » sur la dynamique de HSP90-NTD à 293 K

En acquérant des expériences de CPMG-RD et en analysant les données avec un modèle à deux états en employant le même protocole que pour la protéine Apo la constante de vitesse d'échange conformationnel ainsi que la population de l'état excité d'HSP90-NTD en complexe avec le ligand 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl]-N-furan-2-ylmethyl-2,4-dihydroxy-N-methyl-benzamide ont pu être déterminées à 293 K.

Il apparaît que la nature du ligand a une influence différente sur la constante de vitesse d'échange conformationnel observée. En effet, dans le cas du ligand « L_{tr}^{Court} » présence du ligand accélérerait l'échange entre l'« état ouvert » et l'« état fermé » comparée à la protéine Apo. A l'inverse, en présence du ligand de type « L_{tr}^{Long} » l'échange entre l'état majoritaire avec une hélice à proximité du site de liaison de l'ATP et celui minoritaire est ralenti. On note une constante de vitesse d'échange conformationnel de $1301 \pm 56 \text{ s}^{-1}$ en présence du ligand « L_{tr}^{Long} » (Figure 4.21).

Concernant la population de l'état excité, la présence du ligand « L_{tr}^{Long} », induisant une tout autre conformation de l'état fondamental et un état excité différent de la protéine Apo, semble stabiliser la conformation excitée – selon notre hypothèse, l'état ouvert. La population de cet état en présence du ligand est de 5.4 % (Figure 4.21).

Figure 4.21 : Diagramme schématique du paysage énergétique caractérisant l'échange entre l'état fondamental (avec une structure secondaire en hélice sur le segment [100-123]) et l'état excité (état ouvert) de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ». Les paramètres cinétiques et thermodynamiques de cet échange, extraits à 293 K, sont reportés. Le ligand « L_{tr}^{Long} » est représenté en rose dans les structures tridimensionnelles.



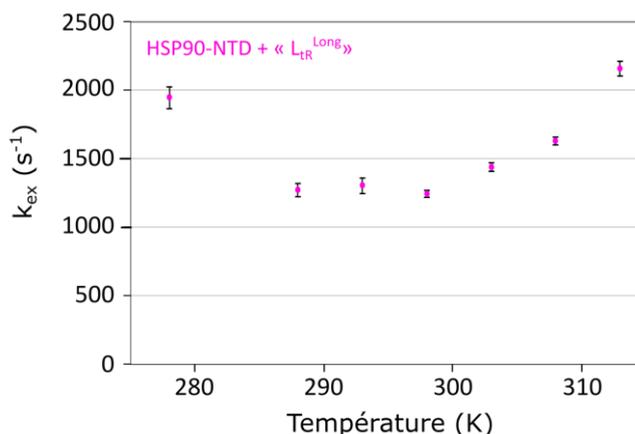
4.4.4 Influence du ligand « L_{tr}^{Long} » sur la dynamique de HSP90-NTD sur la plage de température [278-313 K]

De la même manière que pour la protéine en complexe avec le ligand « L_{tr}^{Court} » j'ai acquis et analysé des expériences de relaxation dispersion pour la protéine en complexe avec le ligand « L_{tr}^{Long} » pour 8 températures différentes allant de 278 K à 313 K avec un pas de 5 K.

On note ici l'absence de données à 283 K pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ». En effet, les données de relaxation-dispersion n'ont pas pu être ajustées à cette température. A basse température la qualité des spectres RMN pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} » se révèle être moindre comparée aux données acquises aux même températures pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} », que ce soit en termes de sensibilité ou de résolution sur le spectre. Les données à 278 K ayant été acquises sur un week-end, il m'a été possible de les ajuster, la sensibilité étant suffisante. Pour les données acquises à 283 K, les barres d'erreurs sur les vitesses de relaxation effectives à 700 et 850 MHz sont trop élevées. Lors de l'analyse pour la sélection des paramètres initiaux de l'ajustement global j'ai pu constater que le minimum local était très élargi, particulièrement pour la population, empêchant tout ajustement précis. Aussi, j'ai décidé de ne pas utiliser cette température lorsque la protéine était en complexe avec le ligand « L_{tr}^{Long} » pour la suite de l'analyse.

En premier lieu, j'ai pu constater que la constante de vitesse d'échange conformationnel pour la protéine en complexe, inférieure à 2160 s^{-1} quelle que soit la température entre 278 K et 313 K, n'augmentait pas de façon constante avec la température (Figure 4.22). En effet, sur la plage de température 278 K – 298 K la constante de vitesse d'échange conformationnel diminue puis stagne avec la température et cette observation va dans le sens contraire de ce que l'on aurait pu penser avec les principes de la thermodynamique.

Figure 4.22 : Graphiques représentant l'évolution de la constante de vitesse d'échange conformationnel (k_{ex}) en fonction de la température pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} ». La précision sur les données a été estimée en utilisant une analyse Monte-Carlo.



Concernant la population de l'état excité j'ai pu voir qu'en présence du ligand « L_{tr}^{Long} » l'évolution de la population de l'état excité en fonction de la température évoluait également de façon complexe (Figure 4.23), différents phénomènes semblent être présents. J'ai remarqué qu'en moyenne, sur la gamme de température [278 – 293 K], pour laquelle les données de population d'état excité pourraient être exploitables pour la protéine en complexe avec le ligand « L_{tr}^{Court} », la population d'état excité était plus élevée pour la protéine en complexe avec le ligand « L_{tr}^{Long} » (8.2 %) comparé à la protéine en complexe avec le ligand « L_{tr}^{Court} » (3.0 %).

La possibilité de la présence de différents états excités dont les populations sont pondérées différemment en fonction de la température existe. Aussi, l'interprétation de l'évolution de k_{ex} et de P^* en fonction de la température est très complexe.

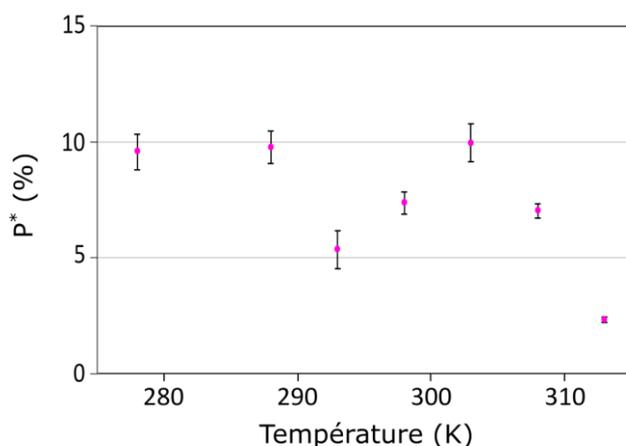


Figure 4.23 : Graphique montrant l'évolution de la population de l'état excité de HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} » en fonction de la température.

4.5 Discussion

4.5.1 HSP90-NTD Apo et en complexe avec le ligand « L_{tr}^{Court} » - un comportement similaire

Les empreintes spectrales très similaires de la protéine Apo ou en complexe avec le ligand « L_{tr}^{Court} » et les similitudes des spectres NOESY pour les deux protéines montrant les pics caractéristiques de l'état fermé suggèrent qu'un phénomène plus complexe qu'un simple modèle à deux états pourrait également se produire pour la protéine dans sa forme Apo, comme suggéré au 4.3.4 pour la protéine en complexe avec le ligand « L_{tr}^{Court} » en observant l'évolution de la population sur la gamme de température 278 K – 293 K.

Pour tenter d'élucider les phénomènes se produisant j'ai acquis des expériences de relaxation dispersion sur le mutant HSP90-NTD-R60A, présenté au chapitre 3. Ce mutant déstabilise la forme fermée de HSP90-NTD la rendant indétectable par RMN à 293 K. J'ai pu montrer qu'un échange entre un état fondamental et un état excité persistait pour le mutant R60A (Figure 4.24). Les valeurs de constante de vitesse d'échange conformationnel et de population d'état excité obtenues avec un ajustement global de 25 résidus pour le mutant R60A sont respectivement de : $3541 \pm 111 \text{ s}^{-1}$ et $2.7 \pm 0.2 \%$, en faisant l'hypothèse que l'on est toujours dans le régime d'échange intermédiaire et que l'on peut donc extraire la population de l'état excité.

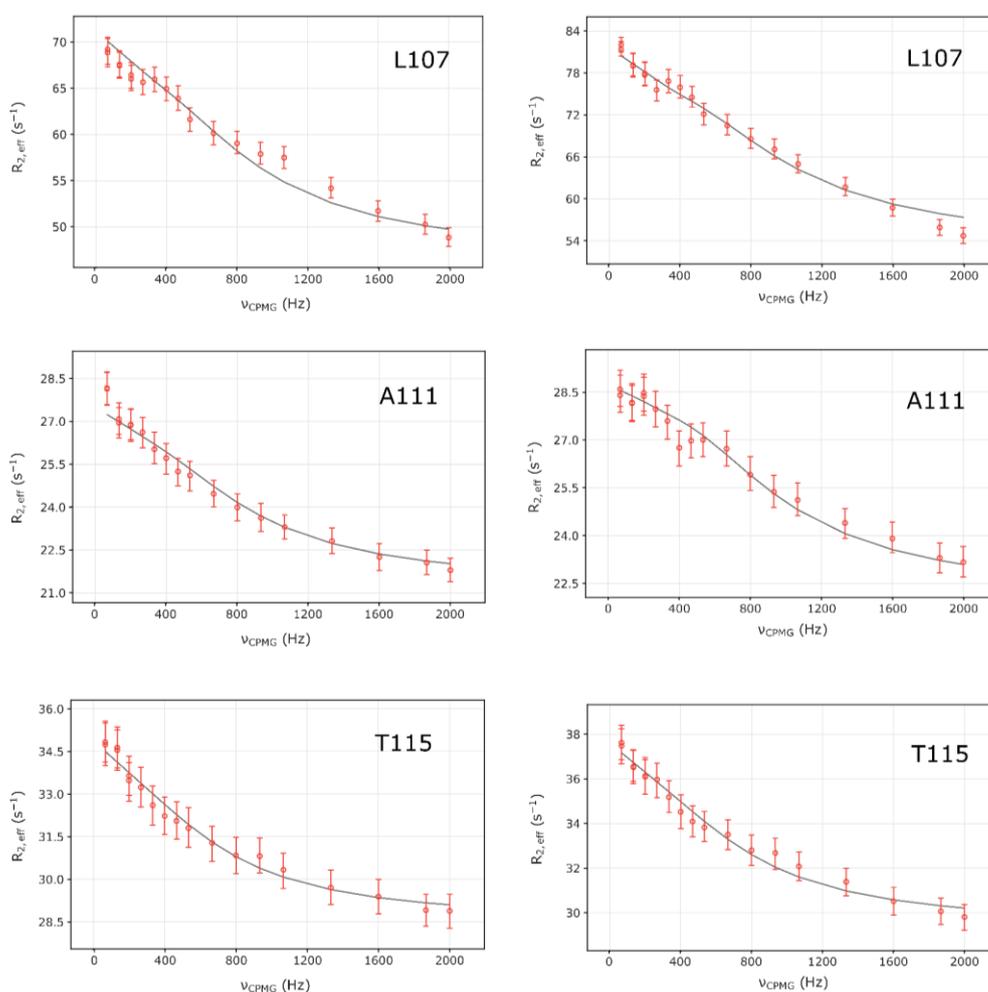


Figure 4.24 : Ajustement global des données de relaxation dispersion pour le mutant HSP90-NTD-R60A à 293 K comprenant 25 résidus. Les graphiques représentant $R_{2,eff}$ en fonction de la fréquence CPMG (ν_{CPMG}) montrent les ajustements globaux obtenus pour 3 des 25 résidus ajustés globalement. La colonne de gauche montre les ajustements obtenus pour les données acquises à 700 MHz et la colonne de droite pour les données acquises à 850 MHz.

De ces résultats, un modèle plus complexe qu'un modèle à deux états peut être proposé où un (ou plusieurs) état(s) intermédiaire(s) existe(nt) (Figure 4.25). Dans ce modèle, la protéine sauvage Apo et, par homologie, en complexe avec le ligand « L_{TR}^{Court} » échange entre un état majoritaire/fondamental ouvert et un état minoritaire/excité fermé et la transition de l'une à l'autre de ces conformations nécessite le passage par un (ou plusieurs) état(s) intermédiaire(s). Par exemple, on peut faire l'hypothèse qu'un des états intermédiaires pourrait être la structure ouverte mais avec l'hélice 5 dépliée comme observée dans l'état fermé (voir article III chapitre 3) ou une conformation avec le segment [98-136] qui descend mais n'atteint pas complètement la forme fermée.

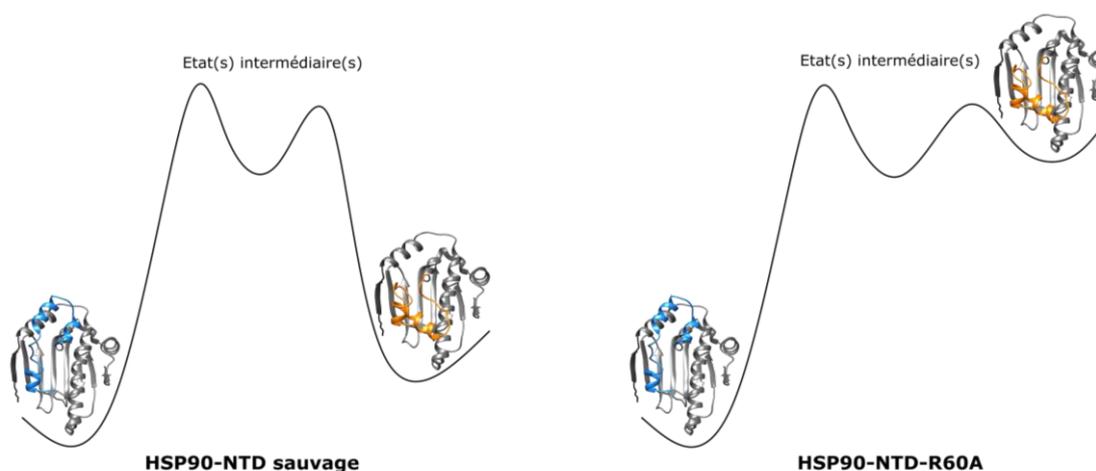


Figure 4.25 : Diagramme schématisant le paysage énergétique caractérisant à gauche l'échange entre l'état majoritaire/fondamental ouvert et l'état minoritaire/excité fermé de HSP90-NTD Apo et à droite l'échange entre l'état majoritaire/fondamental ouvert et un état intermédiaire de HSP90-NTD Apo pour la protéine HSP90-NTD-R60A.

Concernant le mutant HSP90-NTD-R60A, ce dernier déstabilise la forme fermée de HSP90-NTD. Cela a été observé avec la disparition des pics de corrélation NOEs caractéristiques de l'état fermé pour ce mutant (voir chapitre 3). Or, on peut noter que, comparé à la protéine sauvage Apo ($2490 \pm 61 \text{ s}^{-1}$ et $3.2 \pm 0.2 \%$), le mutant R60A entraîne une augmentation de la constante de vitesse d'échange conformationnel et, en faisant l'hypothèse que l'on est toujours dans le régime d'échange intermédiaire et que l'on peut obtenir des informations sur la population de l'état excité, une diminution de la population de l'état excité. Ainsi, l'état excité avec lequel l'état fondamental de HSP90-NTD-R60A est en échange serait plus haut en énergie que l'état excité avec lequel la protéine sauvage Apo est en échange. Possiblement, l'échange détecté pour le mutant R60A est l'échange entre la forme ouverte de HSP90-NTD et un état intermédiaire de HSP90-NTD Apo (Figure 4.25, à droite). Nous pouvons également faire

l'hypothèse que la protéine peuple deux états excités pour le mutant R60A, par exemple la conformation de HSP90-NTD ouverte mais avec l'hélice 5 dépliée ou une conformation avec le segment [98-136] qui descend mais n'atteint pas complètement la forme fermée et celle de HSP90-NTD avec l'ATP-lid en position fermée. La présence de la conformation HSP90-NTD avec l'ATP-lid en position fermée comme second état excité pourrait expliquer la non-observation des pics caractéristiques de l'état fermé, ce dernier étant possiblement trop peu peuplé pour être détecté sur le spectre NOESY.

Plusieurs états excités pourraient être présents, cependant, une analyse par CEST (voir chapitre 1 partie 1.3.2.3) pour déterminer le nombre d'états excités n'est pas envisageable puisque la constante de vitesse d'échange conformationnel est, dans notre cas, trop élevée.

Un modèle possible caractérisant l'échange de HSP90-NTD est que HSP90-NTD Apo dans son état fondamental adopte une conformation ouverte avec une boucle flexible à proximité du site de fixation des ligands/nucléotides. La rupture d'une ou plusieurs liaisons hydrogènes/ponts salins, maintenant cette conformation ouverte, induite par la dynamique de la protéine va entraîner la migration du segment [98-136]. Ce dernier sera stabilisé en position fermé lorsque de nouvelles interactions (liaisons hydrogènes/ponts salins) se formeront avec la partie basse de la protéine. Le mutant R60A, conçu pour rompre le pont salin entre R60 et D127 (chapitre 3 partie 3.3) et déstabiliser l'état fermé de la protéine, remplit son rôle puisqu'avec les expériences NOESY acquises à 293 K les pics de corrélation NOEs caractéristiques de l'état fermé ne sont plus observés. Toutefois, les pics de corrélation NOEs ne traduisent que de courtes distances et possiblement le mutant R60A est en échange entre la conformation ouverte et une conformation où le segment [98-136] s'approche de la position qu'il adopte dans l'état fermé mais ne l'atteint pas complètement, n'étant pas stabilisé par assez de liaisons hydrogènes ou ponts salins.

Aussi, pour vérifier si le mutant R60A supprime complètement l'échange avec une conformation où le segment [98-136] descend et se rapproche de l'état fermé, nous pourrions mettre en place des expériences employant l'effet PRE comme décrit au chapitre 2. Pour cela, il faudrait concevoir un échantillon de HSP90-NTD-R60A muté avec une cystéine localisée à un endroit tel qu'elle n'affecterait pas la structure de la protéine. Deux expériences 2D ^1H - ^{13}C HMQC acquises avec la protéine R60A mutée, en complexe soit avec un centre diamagnétique soit pour la seconde expérience avec un centre paramagnétique permettraient de répondre à notre interrogation en exploitant l'effet à plus longue portée des expériences PRE par rapport aux expériences NOE dans le but de détecter un état transitoire. L'expérience avec l'addition du centre diamagnétique servirait de référence et celle avec le centre paramagnétique permettrait d'observer le

mouvement vers le bas de la protéine du segment [98-136]. En effet, si l'état partiellement fermé est échantillonné par le mutant R60A, alors en présence d'un centre paramagnétique, les pics localisés dans l'état ouvert loin du centre paramagnétique mais situés, dans l'état partiellement fermé, proche de ce dernier seront élargis. A l'inverse si l'état partiellement fermé n'est pas échantillonné par le mutant R60A alors l'intensité des pics éloignés du centre paramagnétique dans l'état ouvert et proche de ce centre dans l'état partiellement fermé auront la même intensité dans le spectre de référence que dans celui acquis avec le centre paramagnétique puisque l'état partiellement fermé sera inexistant.

J'ai noté que le modèle à deux états actuel ne permet pas d'expliquer toutes les variations de population d'état excité observées, toutefois aucune preuve concrète pour un modèle à trois états n'a été révélée. Il est difficile de proposer un modèle à trois états qui pourrait représenter ce qui se passe en solution pour HSP90-NTD Apo et en complexe avec le ligand « L_{tr}^{Court} ». Possiblement, trois états sont en échange, l'état fondamental est l'état ouvert et deux états excités existeraient. L'un est l'état fermé et potentiellement, le second état excité pourrait être un état ouvert où l'hélice-5, dépliée dans l'état fermé, reste dépliée dans l'état ouvert ou alors le second état excité pourrait être un état où le segment [98-136] s'approche de la position qu'il adopte dans l'état fermé mais ne l'atteint pas.

La présence de deux états excités dont la population varie en fonction de la température pourrait permettre d'expliquer la forte imprécision obtenue pour $|\overline{\Delta\omega_c}|$, calculée pour la protéine Apo au chapitre 3 en ayant fait l'hypothèse que les conformations de l'état fondamental et de l'état excité seraient identiques sur une plage de température de 20 °C et que donc sur cette dernière les valeurs de $|\overline{\Delta\omega_c}|$ seraient constantes.

4.5.2 HSP90-NTD complexe avec le ligand « L_{tr}^{Long} »

Concernant HSP90-NTD en complexe avec le ligand « L_{tr}^{Long} », j'ai pu observer avec les expériences de relaxation dispersion que la protéine était en échange conformationnel sur toute la gamme de température allant de 278 K à 313 K. J'ai observé une variation atypique de la constante de vitesse d'échange conformationnel en fonction de la température. Sur la plage de température 278 K – 298 K la constante de vitesse d'échange conformationnel diminue lorsque la température augmente et sur la plage de température 298 K – 313 K, la constante de vitesse d'échange conformationnel augmente lorsque la température augmente. De plus, plusieurs phénomènes semblent

également se produire lorsque l'on observe l'évolution de la population en fonction de la température.

J'ai pu montrer (partie 4.4.1) que la présence de l'inhibiteur stabilisait à haute température un état fondamental avec une structure secondaire en hélice pour le segment [100-123] localisé à proximité du site de fixation des ligands. Avec une expérience NOESY, acquise à haute température également, j'ai pu montrer que les pics de corrélation NOEs caractéristiques de l'état fermé n'étaient pas détectés pour la protéine en complexe avec le ligand « L_{TR}^{Long} » et j'ai noté la présence de 2 pics de corrélation NOEs potentiellement caractéristiques d'un état excité proche structurellement de l'« état ouvert » de la protéine Apo.

Comme l'évolution de la constante de vitesse d'échange conformationnel en fonction de la température apparaissait atypique j'ai cherché à savoir si je pouvais confirmer la conformation fondamentale de mon modèle (i.e: où une hélice décrit la structure secondaire du segment [100-123]) à basse température. Avec des expériences de triple résonances à 293 K je n'ai pas pu conclure quant à la conformation de mon état fondamental à basse température. En effet, le segment [108-132] était invisible et je n'ai donc pas pu en déduire une prédiction de structure secondaire.

J'ai également acquis une expérience NOESY à basse température (286 K) pour voir si les pics de corrélation NOEs caractéristiques de l'état fermé ou ceux 109-131, 115-124 potentiellement caractéristiques d'un état excité proche structurellement de l'état ouvert étaient observés à basse température. Aucun pic caractéristique de l'état fermé ni les NOEs 109-131, 115-124 n'ont été observés dans les spectres à 286 K.

Aussi, il semble que les structures employées dans le modèle présenté en 4.4.3 pour décrire la cinétique et la thermodynamique de l'échange pour HSP90-NTD en complexe avec le ligand « L_{TR}^{Long} » à 293 K ne puissent pas être validées sur la gamme basse des températures. Par contre, rien n'indique que l'état fondamental avec une structure secondaire en hélice pour le segment [100-123] ne décrit pas l'état fondamental à basse température. Ainsi, les structures utilisées dans ce modèle semblent plus adaptées pour décrire les phénomènes se produisant à haute température. Toutefois, on note que même sur la plage [298 K – 313 K] l'évolution de la population d'état excité varie de façon complexe et potentiellement un modèle à deux états n'est pas suffisant pour décrire les phénomènes complexes se produisant.

Au vu de l'évolution de la constante de vitesse d'échange conformationnel et de la population d'état excité sur la gamme de température [278 K – 313 K], il peut être

supposé qu'un second état excité, potentiellement l'état fermé ou un autre état est présent à un faible pourcentage et que sa population varie en fonction de la température. Parmi les deux états excités l'un induirait une constante de vitesse d'échange conformationnel plus rapide et serait majoritairement présent à basse température.

4.5.3 Extraction des valeurs de population de l'état excité et de différences de déplacements chimiques par CPMG-RD lorsque l'on tend vers un régime d'échange rapide

Avec l'observation d'une constante de vitesse d'échange conformationnel (k_{ex}) élevée et donc d'un échange tendant vers le régime d'échange rapide (pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} » à haute température), il est difficile d'obtenir une population de l'état excité fiable, décorrélée des valeurs absolues de différences de déplacements chimiques, en employant uniquement les courbes de relaxation dispersion.

Différentes méthodes peuvent être employées pour être en mesure d'extraire séparément les valeurs de la population d'état excité de celles des différences de déplacements chimiques. La première consiste à acquérir de nouvelles expériences de CPMG-RD à plus haut champ magnétique, augmentant ainsi $\Delta\omega$ et par conséquent faisant diminuer le rapport $k_{ex}/\Delta\omega$. Pour HSP90-NTD en complexe avec le ligand « L_{tr}^{Court} », les expériences de CPMG-RD aux huit températures ont été acquises à deux champs magnétiques et le plus élevé était de 850 MHz. Avec les spectromètres disponibles à l'IBS, je pourrai acquérir des expériences de CPMG-RD à 950 MHz. Cependant, cela ne me permettrait de réduire que d'un facteur 1.12 le rapport $k_{ex}/\Delta\omega$. Ainsi, même avec cette réduction, je ne serai pas en mesure, pour les températures au-delà de 298 K, de retrouver les valeurs de $k_{ex}/\Delta\omega$ obtenues pour des températures inférieures ou égale à 293 K, plus de deux fois moins élevées que celle obtenue à 298 K. Des spectromètres avec un champ magnétique de 1.2 GHz existent actuellement toutefois, je ne serai, là non plus, pas en mesure de retrouver les valeurs de $k_{ex}/\Delta\omega$ obtenues pour des températures inférieures ou égale à 293 K puisque cela me permettrait de réduire d'un facteur 1.41 le rapport $k_{ex}/\Delta\omega$. De plus, le coût des expériences de CPMG-RD, relativement longues, serait très élevé sur de tels spectromètres et ces derniers étant très récents il est actuellement difficile de dire si ils sont suffisamment stables pour acquérir des expériences de CPMG-RD.

Une seconde méthode consisterait à manipuler k_{ex} en changeant la pression du système. En effet, nous avons pu observer qu'en changeant la température les valeurs de k_{ex}

varient considérablement pour HSP90-NTD en complexe avec le ligand « L_{tR}^{Court} ». Potentiellement, un changement de pression pourrait avoir la même fonction. Toutefois, il est à noter que lorsque l'on change la pression du système il est difficile d'anticiper les résultats, potentiellement la cinétique de l'échange serait ralentie et l'on pourrait obtenir deux ensembles de pics sur le spectre, l'un correspondant à l'état majoritaire et le second à l'état minoritaire et ainsi on pourrait extraire les valeurs de $\Delta\omega$, mais la protéine pourrait également ne pas être stable à plus haute pression dans son tampon actuel à pH 7.5.

Une troisième méthode, cette fois ne visant pas à moduler les valeurs de k_{ex} et $|\Delta\omega|$ pour tendre à nouveau vers le régime d'échange intermédiaire mais à rompre la corrélation entre la population de l'état excité et la différence de déplacement chimique existe. Vallurupalli et al. (Vallurupalli et al. 2011) montrent qu'en combinant analyse quantitative des profils de relaxation dispersion avec la comparaison des déplacements chimiques des pics de l'état majoritaire enregistrés avec des expériences HMQC et HSQC, l'extraction des valeurs de P^* et $\Delta\omega$ pour des systèmes tendant vers le régime d'échange rapide est possible. En effet, déplacements chimiques et vitesses de relaxation sont influencées différemment par le processus d'échange. Ainsi, les différences de déplacements chimiques $\overline{\omega_{SQ}} - \overline{\omega_{MQ}}$ introduites comme paramètres d'entrée additionnels dans le programme pour l'ajustement global des données de relaxation-dispersion permettraient d'obtenir des informations décorrélées sur P^* et $\Delta\omega$. De plus, nous aurions également accès au signe de $\Delta\omega$, paramètre très important pour obtenir plus d'informations sur l'état excité.

4.5.4 Vers un modèle à trois états pour HSP90-NTD

Une idée à développer serait de mettre en place une analyse des données de relaxation dispersion avec un modèle employant trois états pour la protéine Apo ou en complexe avec les différents inhibiteurs. Cependant, une telle analyse entraîne l'introduction de nouveaux paramètres et cette fois non plus 4 paramètres (k_{ex} , P^* , $|\Delta\omega_C|$ et $|\Delta\omega_H|$) seraient à extraire des ajustements des courbes de relaxation dispersion pour un résidu (modèle à deux états) mais huit paramètres : k_{ex-1} , k_{ex-2} , P^*_1 , P^*_2 , $|\Delta\omega_{C1}|$, $|\Delta\omega_{C2}|$, $|\Delta\omega_{H1}|$ et $|\Delta\omega_{H2}|$. Pour un ajustement global de N résidus dont les expériences de CPMG-RD ont été acquises à deux champs magnétiques cela implique $2N + 2$ paramètres dans le cas d'un modèle à deux états mais $4N + 4$ paramètres dans le cas d'un modèle à trois états pour $2N$ courbes $R_2 = f(v_{CPMG})$. Le nombre de paramètres à extraire des ajustements étant très conséquent par rapport au nombre de courbes de CPMG-RD disponibles il serait possible de réduire ce ratio en ajustant simultanément les données obtenues à différentes températures en faisant l'hypothèse que les différences de déplacements

chimiques ne varient pas en fonction de la température. Ainsi, en acquérant les données à, par exemple, trois températures, nous aurions $3 \times 2N$ courbes $R_2 = f(\nu_{\text{CPMG}})$ à ajuster pour extraire $4N + 4 \times 3$ paramètres.

Cette analyse, dont la méthode serait à développer, effectuée sur toute la gamme de température pour la protéine Apo (278K - 298K) ou en complexe avec les ligands (278 K – 313 K, uniquement en utilisant une des méthodes décrites en 4.5.3 pour HSP90-NTD en complexe avec le ligand « $L_{\text{TR}}^{\text{Court}}$ » à haute température) pourrait permettre de mieux comprendre ce qui se passe en solution. Peut-être que lorsque trois états seront employés dans le modèle, une évolution de la population linéaire en fonction de la température sera détectée pour chacun des états excités du modèle. Toutefois, il faudra réussir à distinguer le modèle correct parmi trois possibilités pour un modèle à trois états (Figure 4.26) et les résultats d'un tel modèle seront à analyser attentivement car l'un des principaux inconvénients d'un modèle à trois états est la présence possible de plusieurs minimums locaux en raison du nombre plus important de paramètres à ajuster (Xie et al. 2020).

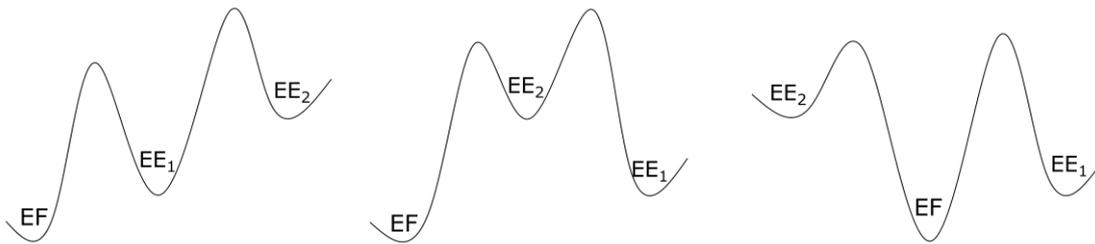


Figure 4.26 : Trois possibilités différentes de modèles à trois états, avec EF : état fondamental, EE_1 état excité 1 et EE_2 état excité 2.

Chapitre 5

Conclusions et perspectives

5.1 Attribution de HSP90-NTD Apo

Au cours de ma thèse, j'ai étudié le domaine N-terminal de HSP90, une protéine dynamique dont l'attribution du squelette a déjà été partiellement effectuée par RMN à l'aide d'expériences de triple résonances classiques (HNCO/HNCACO, HNCOCA/HNCA et HNCOCACB/HNCACB). Cependant, le squelette du segment L107-K116 de la protéine Apo, localisé à proximité du site de fixation de l'ATP, n'est pas observé par RMN en raison d'une dynamique défavorable dans la gamme de temps μ s-ms. Pour pallier ce problème et obtenir des informations sur ce segment invisible et pourtant crucial pour une étude structurale et dynamique de HSP90-NTD, j'ai choisi d'attribuer les groupements méthyles. Ces derniers ont des propriétés de relaxation transversale et longitudinale favorables grâce à leur rotation rapide autour de l'axe C-C et à leur flexibilité (pour les groupements méthyles situés au bout des chaînes latérales de longs acides aminés). De plus, ils permettent d'obtenir des spectres sensibles grâce à leurs trois protons équivalents non échangeables. En employant trois stratégies : le transfert d'attribution du squelette aux groupements méthyles (rendu possible grâce à l'emploi d'un nouveau précurseur développé en collaboration avec le Dr. O. Hamelin et à des expériences « out and back HCC/HC(C)C/HC(CC)C »), la méthode de mutagénèse dirigée améliorée et les corrélations NOEs inter-méthyles à travers l'espace, j'ai pu attribuer 100 % des groupements méthyles d'HSP90-NTD marquée U- ^{2}H , ^{15}N , ^{12}C], $^{13}\text{CH}_3$ – (Ile- δ_1 ,

Leu- δ_1 , Val- γ_2 , Ala- β , Met- ϵ , Thr- γ_2) dans son état Apo et ainsi obtenir des informations sur le segment L107-K116 puisque 5 résidus ayant des groupements méthyles appartiennent à ce segment : L107, T109, I110, A111 et T115.

5.2 Structure et dynamique de HSP90-NTD Apo

Grâce à cette attribution j'ai pu étudier la structure (en collaboration avec E. Rioual, une étudiante de Master 2 que j'ai co-encadrée, qui poursuit actuellement en thèse) et la dynamique de HSP90-NTD. Nous avons montré qu'en solution, à température ambiante le domaine N-terminal d'HSP90 échantillonnait deux ensembles conformationnels distants de plus de 30 Å. Le premier état, avec l'ATP-lid ouvert correspond à l'ensemble des conformations reportées dans la Protein Data Bank pour les plus de 300 structures cristallographiques du domaine N-terminal de HSP90 α . A l'inverse, le second état échantillonné, avec l'ATP-lid en position fermée, recouvrant complètement le site de fixation de l'ATP/des ligands, n'avait jamais été observé au préalable pour le domaine N-terminal de HSP90 isolé. J'ai ensuite employé des expériences de relaxation-dispersion pour caractériser l'échange se produisant dans la gamme de temps intermédiaire entre les deux états distants de plus de 30 Å et obtenir des informations sur la cinétique et la thermodynamique de l'échange. Les résultats suggèrent que l'état fermé de l'ATP-lid de HSP90 entière, fonctionnellement pertinent, pré existe dans la protéine Apo avant la liaison de l'ATP. La description précise des structures échantillonnées par le domaine N-terminal de HSP90 ainsi que la dynamique de ces dernières est une information clé pour la conception de nouveaux ligands à visée thérapeutique contre cette cible majeure pour lutter contre le cancer.

5.3 Influence de la liaison de ligands à HSP90-NTD

5.3.1 Attribution de HSP90 en complexe avec deux ligands ayant des propriétés différentes

J'ai ensuite étudié HSP90-NTD en complexe avec différents inhibiteurs appartenant à la famille des résorcinols : le 5-(2,4-dihydroxy-phenyl)-4-(2-fluoro-phenyl)-2,4-dihydro-[1,2,4]triazol-3-one et le 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. Le premier ligand a un temps de résidence court, de l'ordre de quelques secondes sur sa cible (« L_{tr}^{Court} ») et le second ligand a un temps de résidence de plusieurs dizaines de minutes sur sa cible (« L_{tr}^{Long} »).

Si la liaison du premier inhibiteur sur HSP90-NTD conserve l'empreinte spectrale de la protéine, que ce soit pour le spectre des groupements méthyles ou des groupements

amides, permettant un transfert d'attribution de la protéine Apo à la protéine en complexe, la liaison du second inhibiteur entraîne une très large modification des spectres de la protéine empêchant le transfert d'attribution. Additionnellement, l'échange lent du second ligand entre sa forme libre et liée (temps de résidence ~50 min) prévient l'emploi d'expériences EXSY et de titrations pour transférer l'attribution depuis le spectre de la protéine Apo promouvant une nouvelle attribution de la protéine cible, particulièrement dynamique, en complexe.

Ce ligand « L_{TR}^{Long} » m'a permis de stabiliser la protéine et d'acquérir des spectres à plus haute température. Ainsi, j'ai pu acquérir des expériences de triple résonances classiques à 308 K et attribuer 92 % des résidus du squelette hors prolines de HSP90-NTD en complexe avec le 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl] - N-furan - 2 - ylmethyl-2,4-dihydroxy-N-methyl-benzamide. Avec ce ligand induisant une tout autre conformation de l'empreinte spectrale de la protéine j'ai attribué, pour la première fois, toutes sauf une des résonances du squelette appartenant au segment qui était resté jusqu'alors invisible pour la protéine HSP90-NTD ([L107-K116]). Concernant les groupements méthyles, j'ai employé les trois mêmes stratégies qu'énoncées précédemment et ainsi, j'ai attribué 100 % des groupements méthyles de HSP90-NTD en complexe avec le ligand « L_{TR}^{Long} ».

5.3.2 Structure et dynamique de HSP90-NTD en complexe

La dynamique et la structure de HSP90-NTD en complexe avec les deux ligands a été étudiée. Tout d'abord, j'ai pu montrer que le ligand « L_{TR}^{Court} » stabilisait la même conformation fondamentale que la protéine Apo et que les deux protéines peuplaient également toutes les deux une conformation excitée, fermée où le segment [98-136] recouvrait le site de fixation de l'ATP/ligand. J'ai démontré qu'en présence du ligand « L_{TR}^{Court} », la constante de vitesse d'échange conformationnel entre l'état fondamental et l'état excité était accélérée par rapport au cas de la protéine Apo (Figure 5.1, en bas à gauche) lorsqu'un modèle à deux états était employé pour décrire la dynamique de l'échange. Toutefois, l'étude de la dynamique sur une large gamme de température montre qu'un phénomène plus complexe qu'un simple modèle à deux états pourrait se produire dans le cas de la protéine en complexe avec le ligand « L_{TR}^{Court} » et par homologie dans le cas de la protéine Apo. En effet, le réarrangement structural du segment [98-136] observé est un mouvement d'amplitude de plus de 30 Å et donc est très probablement plus complexe qu'une simple transition entre deux états comme utilisé pour quantifier la constante de vitesse d'échange conformationnel et la population de l'état excité par CPMG. Potentiellement plusieurs états intermédiaires pourraient exister, voire même, un second état excité.

Concernant le ligand « L_{tr}^{Long} », j'ai pu montrer que ce dernier induit une nouvelle conformation fondamentale à haute température pour la protéine en complexe pour laquelle une structure secondaire en hélice constitue le segment [100-123]. Comme pour la protéine en complexe avec le ligand « L_{tr}^{Court} », la protéine en complexe avec le ligand « L_{tr}^{Long} » est également en échange dans la gamme de temps μ s-ms. À l'aide d'expériences NOESY, j'ai pu faire l'hypothèse la plus simple que l'état excité de HSP90-NTD en complexe avec ce ligand était l'état ouvert de HSP90-NTD Apo ou une structure similaire à ce dernier. Contrairement au ligand « L_{tr}^{Court} », la présence du ligand diminue la constante de vitesse d'échange entre l'état fondamental et l'état excité comparé à la protéine Apo (Figure 5.1, en bas à droite) lorsqu'un modèle à deux états est employé. Cependant, l'étude de la dynamique, principalement des constantes de vitesse d'échange conformationnel, sur la gamme de température [278 K – 313 K] montre que potentiellement plusieurs phénomènes se passent, l'un à basse température entre 278 K et 298 K et un second à haute température entre 298 K et 313 K. De plus, une analyse de la population sur les deux gammes de températures semble montrer qu'un modèle à deux états n'est pas suffisant pour décrire totalement les phénomènes complexes se produisant.

La présence d'un troisième état n'a pas pu être confirmée au cours de ma thèse. Dans le modèle impliquant la protéine Apo/en complexe avec le ligand « L_{tr}^{Court} », notre hypothèse actuelle est que la structure du troisième état pourrait être la structure de HSP90-NTD ouverte mais avec l'hélice 5 dépliée comme observée dans l'état fermé (voir article III chapitre 3) ou une conformation avec le segment [98-136] qui descend mais n'atteint pas complètement la forme fermée. Des expériences avec un centre paramagnétique pourraient être envisagées sur le mutant R60A pour vérifier si l'échange avec une conformation où le segment [98-136] descend et se rapproche de l'état fermé (voir chapitre 4, partie 4.5.1) existe et ainsi valider ou invalider l'une de nos hypothèses. Concernant le ligand « L_{tr}^{Long} », proposer un modèle à trois états se révèle plus complexe. Toutefois j'ai pu noter qu'il s'agisse d'un modèle à deux états ou d'un modèle plus complexe, la présence d'un ligand qu'il soit « L_{tr}^{Court} » ou « L_{tr}^{Long} » module le paysage énergétique de HSP90-NTD et la nature du ligand a également une influence sur le paysage énergétique.

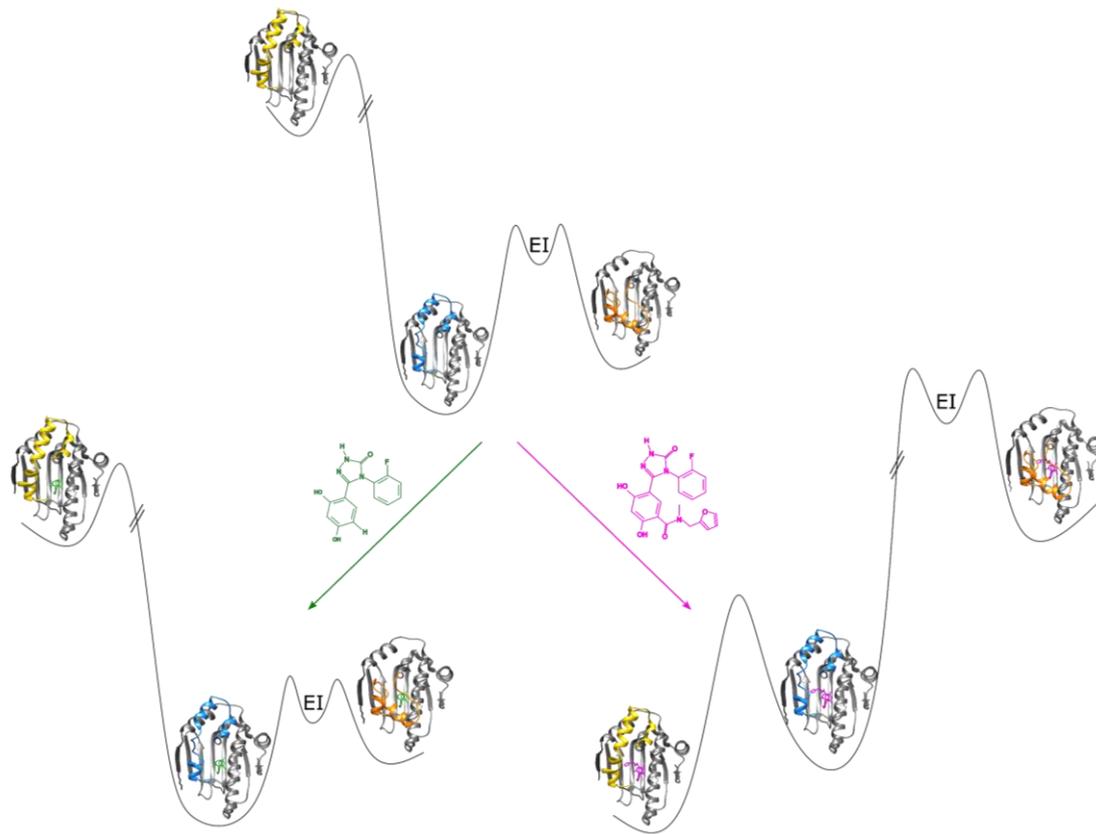


Figure 5.1 : Diagramme schématique du paysage énergétique caractérisant l'échange de HSP90-NTD Apo (en haut) et en complexe avec le ligand « L_{tr}^{Court} » (en bas à gauche) et le ligand « L_{tr}^{Long} » (en bas à droite). Les ligands « L_{tr}^{Court} » et « L_{tr}^{Long} » sont représentés respectivement en vert et rose dans les structures tridimensionnelles. Les structures tridimensionnelles sont représentées avec le segment [98-136] en jaune, bleu et orange pour, respectivement, les conformations ATP-lid ouvert avec une structure secondaire en hélice à proximité du site de fixation du ligand, ATP-lid ouvert avec une boucle hyperflexible à proximité du site de fixation du ligand et ATP-lid fermé. Les traits obliques indiquent une suite possible hypothétique du paysage énergétique non visible par les expériences de relaxation-dispersion. EI correspond à un(des) état(s) intermédiaire(s).

5.4 Modèle décrivant la différence de temps de résidence des deux ligands

Des modèles précédents présentés, je peux proposer une possible explication pour la différence de temps de résidence sur la protéine cible observée à 298 K entre le ligand pour lequel la contribution enthalpique domine fortement la liaison et celui pour lequel la contribution entropique domine fortement la liaison. Le ligand « L_{tr}^{Court} », pour lequel la contribution enthalpique domine la liaison, peut se dissocier de la protéine lorsque celle-ci adopte une conformation ouverte avec une boucle flexible (Figure 5.2a). Par contre, le ligand, même s'il est de petite taille, ne peut pas se dissocier de la

conformation fermée où le segment [98-136] recouvre totalement le site de fixation du ligand (Figure 5.2b). Aussi, le ligand se dissocie de la protéine pour une conformation représentant plus de 90 % de la population, l'état ouvert étant l'état fondamental dans le modèle avec le ligand « $L_{\text{tr}}^{\text{Court}}$ », expliquant son court temps de résidence sur le domaine N-terminal de HSP90.

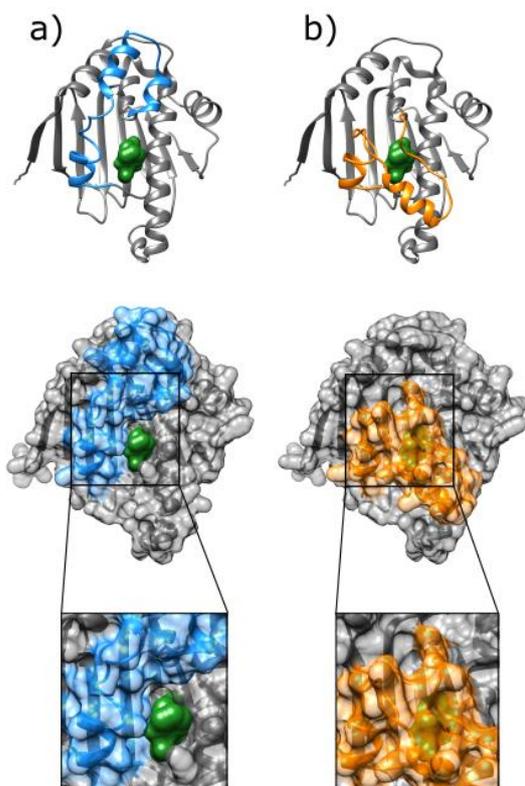
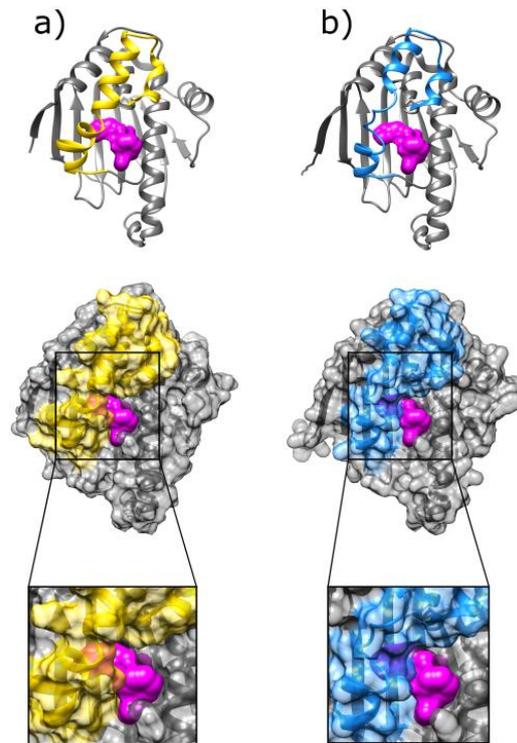


Figure 5.2 : Structures tridimensionnelles et surfaces des conformations possibles de HSP90-NTD en complexe avec le ligand « $L_{\text{tr}}^{\text{Court}}$ ». a) ouverte avec une boucle flexible constituante en partie le segment localisé à proximité du site de liaison du ligand, le ligand peut se dissocier de la protéine b) fermée, le ligand ne peut pas se dissocier de la protéine. Le ligand est représenté en vert.

A l'inverse, le ligand « $L_{\text{tr}}^{\text{Long}}$ », pour lequel la contribution entropique domine fortement la liaison, ne peut se dissocier de la protéine que lorsque celle-ci adopte une conformation ouverte avec une boucle hyperflexible, qui représente selon le modèle proposé son état excité. En effet, lorsque HSP90-NTD adopte une structure secondaire en hélice pour le segment [100-123], le ligand, de taille plus importante que le ligand « $L_{\text{tr}}^{\text{Court}}$ », sera piégé dans cette conformation rigide (Figure 5.3a), il peut se dissocier de la protéine uniquement lorsque celle-ci adopte une conformation ouverte où une boucle flexible constitue en partie le segment localisé à proximité du site de liaison du ligand (Figure 5.3b). Or, d'après les expériences de relaxation dispersion cette population ouverte représente au maximum 10 % de la population expliquant ainsi le long temps de résidence du ligand 5-[4-(2-Fluoro-phenyl)-5-oxo-4,5-dihydro-1H-[1,2,4] triazol-3-yl]-N-furan-2-ylmethyl-2,4-dihydroxy-N-methyl-benzamide sur HSP90-NTD.

Figure 5.3 : Structures tridimensionnelles et surfaces des conformations possibles de HSP90-NTD en complexe avec le ligand « $L_{\text{tr}}^{\text{Long}}$ ». a) structure secondaire en hélice pour le segment [100-123], le ligand ne peut pas se dissocier de la protéine b) ouverte avec une boucle hyperflexible constituant en partie le segment localisé à proximité du site de liaison du ligand, le ligand peut se dissocier de la protéine. Le ligand est représenté en rose.



Ainsi, la base de l'explication des différents temps de résidence provient des chemins de dissociation des ligands qui sont différents. Toutefois, la raison pour laquelle la liaison du ligand « $L_{\text{tr}}^{\text{Long}}$ » est stabilisée entropiquement et celle du ligand « $L_{\text{tr}}^{\text{Court}}$ » est stabilisée enthalpiquement reste non résolue. De plus, d'après les modèles proposés une conformation avec une hélice rigide à proximité du site de fixation du ligand est l'état fondamental pour HSP90-NTD en complexe avec le ligand « $L_{\text{tr}}^{\text{Long}}$ » alors qu'une conformation avec une boucle flexible à proximité du site de fixation du ligand est l'état fondamental pour HSP90-NTD en complexe avec le ligand « $L_{\text{tr}}^{\text{Court}}$ ». De telles données peuvent sembler contradictoires avec le fait que le ligand « $L_{\text{tr}}^{\text{Long}}$ » induit du désordre (i.e. ait une contribution entropique qui domine fortement la liaison). L'étude de la dynamique dans la gamme de temps μs - ms n'a pas permis d'obtenir de piste pour répondre à ce paradoxe observé, cependant, il est à noter que la part la plus importante de la contribution entropique est observée avec l'étude des mouvements plus rapides (Wand 2013) justifiant en partie pourquoi l'étude de la dynamique dans la gamme de temps μs - ms n'est pas en mesure d'expliquer le changement entropique observé par Amaral et al. (Amaral et al. 2017). La relaxation de spin nucléaire est une des techniques qui pourrait être employée pour étudier la dynamique rapide et notamment les mouvements se produisant dans la gamme de temps ps-ns, toutefois, la contribution importante de l'eau dans cette dynamique sera difficilement évaluée en employant la RMN.

Une autre technique permettant de comprendre à la fois le rôle de la dynamique de la protéine et de celle des molécules d'eau est la dynamique moléculaire. Employer cette méthode était compliqué dans le passé car les structures cristallographiques de HSP90-NTD ne reflétaient pas la réalité de la protéine en solution. Les résultats obtenus au cours de cette thèse ouvrent de nouveaux horizons pour évaluer la dynamique rapide via des simulations et obtenir des informations simultanément sur le solvant et la protéine.

Références bibliographiques

- Adrian M, Dubochet J, Lepault J, McDowell A (1984) Cryo_electron microscopy of viruses. *Nature* 308:32–36
- Agarwal PK (2006) Enzymes: An integrated view of structure, dynamics and function. *Microb Cell Fact* 5:2 <https://doi.org/10.1186/1475-2859-5-2>
- Agirrezabala X, Samatova E, Klimova M, et al (2017) Ribosome rearrangements at the onset of translational bypassing. *Sci Adv* 3:e1700147
- Aligue R, Akhavan-Niak H, Russell P (1994) A role for Hsp90 in cell cycle control: Weel tyrosine kinase activity requires interaction with Hsp90. *EMBO J* 13:6099–6105
- Allerhand A, Doddrell D, Glushko V, et al (1971) Conformation and Segmental Motion of Native and Denatured Ribonuclease A in Solution. Application of Natural-Abundance Carbon-13 Partially Relaxed Fourier Transform Nuclear Magnetic Resonance. *J Am Chem Soc* 93:544–546
- Amaral M, Kokh DB, Bomke J, et al (2017) Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat Commun* 8:2276 <https://doi.org/10.1038/s41467-017-02258-w>
- Amero C, Asunción Durá M, Noirclerc-Savoye M, et al (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR* 50:229–236. <https://doi.org/10.1007/s10858-011-9513-5>
- Amero C, Schanda P, Asunción Durá M, et al (2009) Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J Am Chem Soc* 131:3448–3449. <https://doi.org/10.1021/ja809880p>
- Ansari A, Berendzen J, Bownet SF, et al (1985) Protein states and proteinquakes. *Prot Natl Acad Sci* 82:5000–5004
- Ansari A, Berendzen J, Braunstein D, et al (1987) Rebinding and relaxation in the myoglobin pocket. *Biophys Chem* 26:337–355
- Avila C, Hadden MK, Ma Z, et al (2006) High-throughput screening for Hsp90 ATPase inhibitors. *Bioorganic Med Chem Lett* 16:3005–3008 <https://doi.org/10.1016/j.bmcl.2006.02.063>.
- Ayala I, Chiari L, Kerfah R, et al (2020) Asymmetric Synthesis of Methyl Specifically Labelled L-Threonine and Application to the NMR Studies of High Molecular Weight Proteins. *Chemistry-Select* 5:5092–5098. <https://doi.org/10.1002/slct.202000827>
- Ayala I, Hamelin O, Amero C, et al (2012) An optimized isotopic labelling strategy of isoleucine- γ 2 methyl groups for solution NMR studies of high molecular weight proteins. *Chem Commun* 48:1434–1436. <https://doi.org/10.1039/c1cc12932e>
- Ayala I, Sounier R, Usé N, et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* 43:111–119. <https://doi.org/10.1007/s10858-008-9294-7>
- Backe SJ, Sager RA, Woodford MR, et al (2020) Post-translational modifications of Hsp90 and translating the chaperone code. *J. Biol. Chem.* 295:11099–11117
- Bardwell JC, Craig EA (1987) Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc Natl Acad Sci U S A* 84:5177–5181. <https://doi.org/10.1073/pnas.84.15.5177>

- Bax A, Clore M, Gronenborn A (1990) 1H-1H Correlation via Isotropic Mixing of ¹³C Magnetization, a New Three-Dimensional Approach for Assigning 1H and ¹³C Spectra of ¹³C-Enriched Proteins. *J Magn Reson* 88:425–431
- Bernadó P, Mylonas E, Petoukhov M V., et al (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* 129:5656–5664. <https://doi.org/10.1021/ja069124n>
- Bernadó P, Svergun DI (2012) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol Biosyst* 8:151–167. <https://doi.org/10.1039/c1mb05275f>
- Bernal J, Crowfoot D (1934) X-Ray Photographs of Crystalline Pepsin. *Nature* 133:794–795
- Biebl MM, Buchner J (2019) Structure, Function, and Regulation of the Hsp90 Machinery. *Cold Spring Harb Perspect Biol* a034017. <https://doi.org/10.1101/cshperspect.a034017>
- Blake C, Koenig D, Mair G, et al (1965) Structure of hen egg-white lysozyme. *Nature* 4986:757–761
- Boehr DD, McElheny D, Dyson HJ, Wrightt PE (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313:1638–1642. <https://doi.org/10.1126/science.1130258>
- Borkovich KA, Farrelly FW, Finkelstein DB, et al (1989) hsp82 Is an Essential Protein That Is Required in Higher Concentrations for Growth of Cells at Higher Temperatures. *Mol Cell Biol* 9:3919–3930. <https://doi.org/10.1128/mcb.9.9.3919>
- Bouvignies G, Vallurupalli P, Hansen DF, et al (2011) Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* 477:111–117. <https://doi.org/10.1038/nature10349>
- Calderwood SK, Khaleque MA, Sawyer DB, Ciocca DR (2006) Heat shock proteins in cancer: Chaperones of tumorigenesis. *Trends Biochem Sci* 31:164–172. <https://doi.org/10.1016/j.tibs.2006.01.006>
- Carreras CW, Schirmer A, Zhong Z, Santi D V. (2003) Filter binding assay for the geldanamycin-heat shock protein 90 interaction. *Anal Biochem* 317:40–46. [https://doi.org/10.1016/S00032697\(03\)0060-5](https://doi.org/10.1016/S00032697(03)0060-5)
- Carver J., Richards R. (1972) A General Two-Site Solution for the Chemical Exchange Produced Dependence of T2 Upon the Carr-Purcell Pulse Separation. *J Magn Reson* 6:89–105
- Chan-yao-chong M (2019) Structure et dynamique de protéines intrinsèquement désordonnées : Caractérisation par une approche combinant dynamique moléculaire avancée et SAXS
- Chao F-A, Kim J, Xia Y, et al (2014) FLAMEnGO 2.0: An enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances. *J Magn Reson* 245:17–23. <https://doi.org/10.1016/j.jmr.2014.04.012>
- Charlier C, Cousin SF, Ferrage F (2016) Protein dynamics from nuclear magnetic relaxation. *Chem Soc Rev* 45:2410–2422. <https://doi.org/10.1039/c5cs00832h>
- Chiosis G, Neckers L (2006) Tumor selectivity of Hsp90 inhibitors: the explanation remains elusive. *ACS Chem. Biol.* 1:279–284
- Chiosis G, Timaul MN, Lucas B, et al (2001) A small molecule designed to bind to the adenine nucleotide pocket of Hsp90 causes Her2 degradation and the growth arrest and differentiation of breast cancer cells. *Chem Biol* 8:289–299
- Cooper A (1976) Thermodynamic fluctuations in protein molecules (stochastic and relaxation processes/"breathing"). *Biochemistry* 73:2740–2741

- Cooper A, Dryden DTF (1984) European Biophysics Journal Allostery without conformational change A plausible model. *Eur Biophys J* 11:103–109
- Crublet E, Kerfah R, Mas G, et al (2014) A cost-effective protocol for the parallel production of libraries of CH₃-specifically labeled mutants for NMR studies of high molecular weight proteins. *Methods Mol Biol* 1091:229–243 https://doi.org/10.1007/978-1-62703-691-7_17
- Csermely P, Schnaider T, Soti C, et al (1998) The 90-kDa Molecular Chaperone Family: Structure, Function, and Clinical Applications. A Comprehensive Review. *Pharmacol Ther* 79:129–168. [https://doi.org/10.1016/s0163-7258\(98\)00013-8](https://doi.org/10.1016/s0163-7258(98)00013-8)
- Dashti A, Mashayekhi G, Shekhar M, et al (2020) Retrieving functional pathways of biomolecules from single-particle snapshots. *Nat Commun* 11:4734. <https://doi.org/10.1038/s41467-020-18403-x>
- Dashti A, Schwander P, Langlois R, et al (2014) Trajectories of the ribosome as a Brownian nanomachine. *Proc Natl Acad Sci U S A* 111:17492–17497. <https://doi.org/10.1073/pnas.1419276111>
- Davis D, Perlman M, London R (1994) Direct measurements of the Dissociation-Rate Constant for Inhibitor Enzyme Complexes via the T₁ρ and T₂ (CPMG) Methods. *J Magn Reson* 104:266–275
- De Rosier D, Klug A (1968) Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature* 217:130–134
- Delaglio F, Grzesiek S, Vuister GW, et al (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293. <https://doi.org/10.1007/BF00197809>
- Demers JP, Mittermaier A (2009) Binding mechanism of an SH3 domain studied by NMR and ITC. *J Am Chem Soc* 131:4355–4367. <https://doi.org/10.1021/ja808255d>
- Didenko T, Duarte AMS, Karagöz GE, Rüdiger SGD (2012) Hsp90 structure and function studied by NMR spectroscopy. *Biochim Biophys Acta - Mol Cell Res* 1823:636–647. <https://doi.org/10.1016/j.bbamcr.2011.11.009>
- Diez M, Zimmermann B, Börsch M, et al (2004) Proton-powered subunit rotation in single membrane-bound F₁OF₁-ATP synthase. *Nat Struct Mol Biol* 11:135–141. <https://doi.org/10.1038/nsmb718>
- Dimura M, Peulen TO, Hanke CA, et al (2016) Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr Opin Struct Biol* 40:163–185
- Du Y, Moulick K, Rodina A, et al (2007) High-throughput screening fluorescence polarization assay for tumor-specific Hsp90. *J Biomol Screen* 12:915–924. <https://doi.org/10.1177/1087057107306067>
- Dubochet J, Adrian M, Chang J, et al (1988) Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* 21:129–228. <https://doi.org/10.1017/S0033583500004297>
- Dutta R, Inouye M (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci* 25:24–28. [https://doi.org/10.1016/s0968-0004\(99\)01503-0](https://doi.org/10.1016/s0968-0004(99)01503-0)
- Eaton H, Fesik S, Glaserandgaryp S, Drobny (1990) Time Dependence of ¹³C-¹³C Magnetization Transfer in Isotropic Mixing Experiments Involving Amino Acid Spin Systems. *J Magn Reson* 90:452–463
- Ebert MO, Bae SH, Dyson HJ, Wright PE (2008) NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* 47:1299–1308. <https://doi.org/10.1021/bi701767j>

- Farrelly FW, Finkelstein DB (1984) Complete sequence of the heat shock-inducible HSP90 gene of *Saccharomyces cerevisiae*. *J Biol Chem* 259:5745–5751. [https://doi.org/10.1016/s0021-9258\(18\)91077-x](https://doi.org/10.1016/s0021-9258(18)91077-x)
- Farrow NA, Zhang O, Szabo A, et al (1995) Spectral density function mapping using ¹⁵N relaxation data exclusively. *J Biomol NMR* 6:153–162
- Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch Chem Gesellschaft* 27:2985–2993
- Franco R, Gil-Caballero S, Ayala I, et al (2017) Probing Conformational Exchange Dynamics in a Short-Lived Protein Folding Intermediate by Real-Time Relaxation–Dispersion NMR. *J Am Chem Soc* 139:1065–1068. <https://doi.org/10.1021/jacs.6b12089>
- Frank J, Ourmazd A (2016) Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods* 100:61–67. <https://doi.org/10.1016/j.ymeth.2016.02.007>
- Frauenfelder H, Petsko G, Tsernoglou D (1979) Temperature-dependant X-ray diffraction as a probe of protein structural dynamics. *Nature* 280:558–563
- Frauenfelder H, Sligar SG, Wolynes PG (1991) The Energy Landscapes and Motions of Proteins. *Science* 254:1598–1603
- Freire E (2008) Do Enthalpy and Entropy Distinguish First in Class From Best in Class? *Drug Discov Today* 13:869–874. <https://doi.org/10.1016/j.drudis.2008.07.005>
- Gans P, Hamelin O, Sounier R, et al (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chemie - Int Ed* 49:1958–1962. <https://doi.org/10.1002/anie.200905660>
- Gardner KH, Kay LE, Chinchilla D, Fisher K (1997) Production and Incorporation of ¹⁵N, ¹³C, ²H (1H- δ 1 Methyl) Isoleucine into Proteins for Multidimensional NMR Studies. *J Am Chem Soc* 119:7599–7600
- Gauto DF, Estrozi LF, Schwieters CD, et al (2019) Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat Commun* 10:2697 <https://doi.org/10.1038/s41467-019-10490-9>
- Ge J, Normant E, Porter JR, et al (2006) Design, synthesis, and biological evaluation of hydroquinone derivatives of 17-amino-17-demethoxygeldanamycin as potent, water-soluble inhibitors of Hsp90. *J Med Chem* 49:4606–4615. <https://doi.org/10.1021/jm0603116>
- Gonen T, Cheng Y, Sliz P, et al (2006) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438(7068):633–638
- Goto NK, Gardner KH, Mueller GA, et al (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated ¹⁵N-, ¹³C-, ²H-labeled proteins. *J Biomol NMR* 13:369–374
- Gross JD, Gelev VM, Wagner G (2003) A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. *J Biomol NMR* 25:235–242
- Güntert P, Mumenthaler C, Wüthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 273:283–298. <https://doi.org/10.1006/jmbi.1997.1284>
- Hajduk PJ, Augeri DJ, Mack J, et al (2000) NMR-based screening of proteins containing ¹³C-labeled methyl groups. *J Am Chem Soc* 122:7898–7904. <https://doi.org/10.1021/ja000350l>
- Häussinger D, Huang JR, Grzesiek S (2009) DOTA-M8: An extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J Am Chem Soc* 131:14761–14767. <https://doi.org/10.1021/ja903233w>
- Henderson R, Baldwin JM, Ceskat TA, et al (1990) Model for the Structure of Bacteriorhodopsin Based on High-resolution Electron Cryo-microscopy. *J Mol Biol* 213:899–929

- Henderson R, Unwin P (1975) Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* 257:28–32
- Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Biochemistry* 89:10915–10919
- Henot F, Kerfah R, Törner R, et al (2021) Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain. *J Biomol NMR* 75:221–232. <https://doi.org/10.1007/s10858-021-00370-0>
- Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450:964–972. <https://doi.org/10.1038/nature06522>
- Herbert A MaxCluster. <http://www.sbg.bio.ic.ac.uk/~maxcluster/index.html>. Accessed 17 Dec 2020
- Hessling M, Richter K, Buchner J (2009) Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90. *Nat Struct Mol Biol* 16:287–293. <https://doi.org/10.1038/nsmb.1565>
- Hope H (1990) CRYSTALLOGRAPHY OF BIOLOGICAL MACROMOLECULES AT ULTRA LOW TEMPERATURE. *Biophys Chem* 19:107–126
- Hoter A, El-Sabban ME, Naim HY (2018) The HSP90 family: Structure, regulation, function, and implications in health and disease. *Int J Mol Sci* 19,2560. <https://doi.org/10.3390/ijms19092560>
- Howes R, Barril X, Dymock BW, et al (2006) A fluorescence polarization assay for inhibitors of Hsp90. *Anal Biochem* 350:202–213. <https://doi.org/10.1016/j.ab.2005.12.023>
- Hu B, Margolin W, Molineux IJ, Liu J (2015) Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc Natl Acad Sci U S A* 112:E4919–E4928. <https://doi.org/10.1073/pnas.1501064112>
- Jacobs DM, Langer T, Elshorst B, et al (2006) NMR Backbone Assignment of the N-terminal Domain of Human HSP90. *J Biomol NMR* 36:52. <https://doi.org/10.1007/s10858-006-9030-0>
- Jacques DA, Trehwella J (2010) Small-angle scattering for structural biology - Expanding the frontier while avoiding the pitfalls. *Protein Sci* 19:642–657. <https://doi.org/10.1002/pro.351>
- James LC, Roversi P, Tawfik DS (2003) Antibody multispecificity mediated by conformational diversity. *Science* 299:1362–1367. <https://doi.org/10.1126/science.1079731>
- John M, Schmitz C, Ah YP, et al (2007) Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc* 129:13749–13757. <https://doi.org/10.1021/ja0744753>
- Johnson JL (2012) Evolution and function of diverse Hsp90 homologs and cochaperone proteins. *Biochim Biophys Acta - Mol Cell Res* 1823:607–613. <https://doi.org/10.1016/j.bbamcr.2011.09.020>
- Johnson L, Phillips D (1965) Structure of some crystalline lysozyme-inhibitor complexes determined by X-Ray analysis at 6 Å resolution. *Nature* 208:761–763
- Kagawa TF, Cooney JC, Baker HM, et al (2000) Crystal structure of the zymogen form of the group A Streptococcus virulence factor SpeB: An integrin-binding cysteine protease. *PNAS* 97:2235–2240
- Kamal A, Thao L, Sensintaffar J, et al (2003) A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* 425:407–410

- Kay LE (2016) New Views of Functionally Dynamic Proteins by Solution NMR Spectroscopy. *J Mol Biol* 428:323–331. <https://doi.org/10.1016/j.jmb.2015.11.028>
- Keeler J (2002) *Understanding NMR Spectroscopy*. Wiley
- Kendrew J, Bodo G, Dintzis H, et al (1958) A three-dimensional model of the myoglobin molecule obtained by X-Ray analysis. *Nature* 181:662–666
- Kerfah R, Hamelin O, Boisbouvier J, Marion D (2015a) CH₃-specific NMR assignment of alanine, isoleucine, leucine and valine methyl groups in high molecular weight proteins using a single sample. *J Biomol NMR* 63:389–402. <https://doi.org/10.1007/s10858-015-9998-4>
- Kerfah R, Plevin MJ, Pessey O, et al (2015b) Scrambling free combinatorial labeling of alanine-β, isoleucine-δ₁, leucine-proS and valine-proS methyl groups for the detection of long range NOEs. *J Biomol NMR* 61:73–82. <https://doi.org/10.1007/s10858-014-9887-2>
- Kerfah R, Plevin MJ, Sounier R, et al (2015c) Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* 32:113–122. <https://doi.org/10.1016/j.sbi.2015.03.009>
- Khoshouei M, Radjainia M, Baumeister W, Danev R (2017) Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. *Nat Commun* 8:16099. <https://doi.org/10.1038/ncomms16099>
- Kikhney AG, Svergun DI (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett*. 589:2570–2577
- Kim J, Felts S, Llauger L, et al (2004) Development of a fluorescence polarization assay for the molecular chaperone Hsp90. *J Biomol Screen* 9:375–381. <https://doi.org/10.1177/1087057104265995>
- Kim Y, Alarcon S, Lee S, et al (2009) Update on Hsp90 Inhibitors in Clinical Trial. *Curr Top Med Chem* 9:1479–1492. <https://doi.org/10.2174/156802609789895728>
- Kleckner IR, Foster MP (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta - Proteins Proteomics* 1814:942–968
- Korzhnev DM, Kloiber K, Kanelis V, et al (2004a) Probing Slow Dynamics in High Molecular Weight Proteins by Methyl-TROSY NMR Spectroscopy: Application to a 723-Residue Enzyme. *J Am Chem Soc* 126:3964–3973. <https://doi.org/10.1021/ja039587i>
- Korzhnev DM, Kloiber K, Kay LE (2004b) Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: Theory and application. *J Am Chem Soc* 126:7320–7329. <https://doi.org/10.1021/ja049968b>
- Koshland DE (1958) APPLICATION OF A THEORY OF ENZYME SPECIFICITY TO PROTEIN SYNTHESIS. *Proc Natl Acad Sci* 44:98–104
- Koshland DE (1959) Enzyme Flexibility and Enzyme Action. *J Cell Physiol* 54:245–258
- Kovermann M, Rogne P, Wolf-Watz M (2016) Protein dynamics and function from solution state NMR spectroscopy. *Q Rev Biophys* 49,e6. <https://doi.org/10.1017/S0033583516000019>
- Kovrigin EL, Kempf JG, Grey MJ, Loria JP (2006) Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. *J Magn Reson* 180:93–104. <https://doi.org/10.1016/j.jmr.2006.01.010>
- Kumar M, Wiener R (2019) Closing the gap of ubiquitin activation. *Proc. Natl. Acad. Sci. U. S. A.* 116:15319–15321
- Kupitz C, Basu S, Grotjohann I, et al (2014) Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* 513:261–265

- Kupitz C, Olmos JL, Holl M, et al (2017) Structural enzymology using X-ray free electron lasers. *Struct Dyn* 4(4):044003. <https://doi.org/10.1063/1.4972069>
- Lakowicz JR, Weber G (1973) Quenching of Fluorescence by Oxygen. a Probe for Structural Fluctuations in Macromolecules. *Biochemistry* 12:4161–4170. <https://doi.org/10.1021/bi00745a020>
- Lapinaite A, Simon B, Skjaerven L, et al (2013) The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature* 502:519–523. <https://doi.org/10.1038/nature12581>
- Lebret T, Watson RWG, Molinié V, et al (2003) Heat shock proteins HSP27, HSP60, HSP70, and HSP90: Expression in bladder carcinoma. *Cancer* 98:970–977. <https://doi.org/10.1002/cncr.11594>
- Lee K, Thwin AC, Nadel CM, et al (2021) The structure of an Hsp90-immunophilin complex reveals cochaperone recognition of the client maturation state. *Mol Cell* 81:3496–3508.e5. <https://doi.org/10.1016/j.molcel.2021.07.023>
- Lente G, Fábrián I, Poë AJ (2005) A common misconception about the Eyring equation. *New J Chem* 29:759–760. <https://doi.org/10.1039/b501687h>
- Lescanne M, Ahuja P, Blok A, et al (2018) Methyl group reorientation under ligand binding probed by pseudocontact shifts. *J Biomol NMR* 71:275–285. <https://doi.org/10.1007/s10858-018-0190-5>
- Lescanne M, Skinner SP, Blok A, et al (2017) Methyl group assignment using pseudocontact shifts with PARAssign. *J Biomol NMR* 69:183–195. <https://doi.org/10.1007/s10858-017-0136-3>
- Levitt M. (2001) Spin dynamics: Basics of Nuclear Magnetic Resonance. Wiley
- Li J, Buchner J (2013) Structure, function and regulation of the Hsp90 machinery. *Biomed. J.* 36:106–117
- Li J, Soroka J, Buchner J (2012) The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. *Biochim Biophys Acta - Mol Cell Res* 1823:624–635. <https://doi.org/10.1016/j.bbamcr.2011.09.003>
- Li Y, Zhang T, Schwartz SJ, Sun D (2009) New developments in Hsp90 inhibitors as anti-cancer therapeutics: Mechanisms, clinical perspective and more potential. *Drug Resist Updat* 12:17–27. <https://doi.org/10.1016/j.drug.2008.12.002>
- Lipari G, Szabo A (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J Am Chem Soc* 104:4546–4559
- Lo Conte L, Chothia C, Janin J (1999) The Atomic Structure of Protein-Protein Recognition Sites. *J Mol Biol* 285:2177–2198
- London RE, Wingad BD, Mueller GA (2008) Dependence of amino acid side chain ¹³C shifts on dihedral angle: Application to conformational analysis. *J Am Chem Soc* 130:11097–11105. <https://doi.org/10.1021/ja802729t>
- Loveland AB, Demo G, Grigorieff N, Korostelev AA (2017) Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* 546:113–117. <https://doi.org/10.1038/nature22397>
- Macek P, Kerfah R, Erba EB, et al (2017) Unraveling self-assembly pathways of the 468-kDa proteolytic machine TET2. *Sci Adv* 3:e1601601
- Mariño Pérez L, Ielasi FS, Bessa LM, et al (2022) Visualizing protein breathing motions associated with aromatic ring flipping. *Nature* 602:695–700. <https://doi.org/10.1038/s41586-022-04417-6>
- Martin-Garcia JM (2021) Protein dynamics and time resolved protein crystallography at synchrotron radiation sources: Past, present and future. *Crystals* 11:521–543. <https://doi.org/10.3390/cryst11050521>

- Mas G, Crublet E, Hamelin O, et al (2013) Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. *J Biomol NMR* 57:251–262. <https://doi.org/10.1007/s10858-013-9785-z>
- Mas G, Guan J-Y, Crublet E, et al (2018) Structural investigation of a chaperonin in action reveals how nucleotide binding regulates the functional cycle. *Sci Adv* 4:eau4196. <https://doi.org/doi:10.1126/sciadv.aau4196>
- Massi F, Grey M, Palmer III A. (2005) Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1 ρ relaxation experiments. *Protein Sci* 14:735–742. <https://doi.org/10.1110/ps.041139505>
- Mazal H, Haran G (2019) Single-molecule FRET methods to study the dynamics of proteins at work. *Curr. Opin. Biomed. Eng.* 12:8–17
- McCarthy MM, Pick E, Kluger Y, et al (2008) HSP90 as a marker of progression in melanoma. *Ann Oncol* 19:590–594. <https://doi.org/10.1093/annonc/mdm545>
- McLaughlin SH, Ventouras LA, Lobbezoo B, Jackson SE (2004) Independent ATPase activity of Hsp90 subunits creates a flexible assembly platform. *J Mol Biol* 344:813–826. <https://doi.org/10.1016/j.jmb.2004.09.055>
- Meng X, Devin J, Sullivan WP, et al (1996) Mutational analysis of Hsp90 α dimerization and subcellular localization: dimer disruption does not impede 'in vivo' interaction with estrogen receptor. *J Cell Sci* 109:1677–1687
- Merk A, Bartesaghi A, Banerjee S, et al (2016) Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* 165:1698–1707. <https://doi.org/10.1016/j.cell.2016.05.040>
- Meyer P, Prodromou C, Hu B, et al (2003) Structural and Functional Analysis of the Middle Segment of Hsp90: Implications for ATP Hydrolysis and Client Protein and Cochaperone Interactions. *Mol Cell* 11:647–658. [https://doi.org/10.1016/s1097-2765\(03\)00065-0](https://doi.org/10.1016/s1097-2765(03)00065-0)
- Miller S, Janin J, Lesk AM, Chothia C (1987) Interior and Surface of Monomeric Proteins. *J Mol Biol* 196:641–656
- Millson SH, Truman AW, Rácz A, et al (2007) Expressed as the sole Hsp90 of yeast, the α and β isoforms of human Hsp90 differ with regard to their capacities for activation of certain client proteins, whereas only Hsp90 β generates sensitivity to the Hsp90 inhibitor radicicol. *FEBS J* 274:4453–4463. <https://doi.org/10.1111/j.1742-4658.2007.05974.x>
- Milne JLS, Borgnia MJ, Bartesaghi A, et al (2013) Cryo-electron microscopy - A primer for the non-microscopist. *FEBS J* 280:28–45. <https://doi.org/10.1111/febs.12078>
- Mittermaier AK, Kay LE (2009) Observing biological dynamics at atomic resolution using NMR. *Trends Biochem Sci* 34:601–611. <https://doi.org/10.1016/j.tibs.2009.07.004>
- Miyata Y, Nakamoto H, Neckers L (2012) The Therapeutic Target Hsp90 and Cancer Hallmarks. *Curr Pharm Des* 19:347–365. <https://doi.org/10.2174/138161213804143725>
- Monneau YR, Rossi P, Bhaumik A, et al (2017) Automatic methyl assignment in large proteins by the MAGIC algorithm. *J Biomol NMR* 69:215–227. <https://doi.org/10.1007/s10858-017-0149-y>
- Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12:88–118. [https://doi.org/10.1016/S0022-2836\(65\)80285-6](https://doi.org/10.1016/S0022-2836(65)80285-6)

- Murata K, Wolf M (2018) Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta - Gen Subj* 1862:324–334. <https://doi.org/10.1016/j.bbagen.2017.07.020>
- Nango E, Royant A, Kubo M, et al (2016) A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* 354:1552–1557
- Neal J (2021) Protein Structure, Dynamics, and Function: A Philosophical Account of Representation and Explanation in Structural Biology
- Nemoto T, Ohara-Nemoto Y, Ota M, et al (1995) Mechanism of dimer formation of the 90-kDa heat-shock protein. *Eur J Biochem* 233:1–8. https://doi.org/10.1111/j.1432-1033.1995.001_1.x
- Nerli S, De Paula VS, McShan AC, Sgourakis NG (2021) Backbone-independent NMR resonance assignments of methyl probes in large proteins. *Nat Commun* 12:691. <https://doi.org/10.1038/s41467-021-20984-0>
- Nogly P, Weinert T, James D, et al (2018) Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science* 361(6398):eaat0094. <https://doi.org/10.1126/science.aat0094>
- Noji H, Yasuda R, Yoshida M, Kinosita Jr K (1997) Direct observation of the rotation of F1-ATPase. *Nature* 386:299–302
- Ochiana S, Taldone T, Gabriela Chiosis G (2014) Designing Drugs Against Hsp90 for Cancer Therapy. In: Houry W (ed) *The Molecular Chaperones Interaction Networks in Protein Folding and Degradation*
- Ohi M, Li Y, Cheng Y, Walz T (2004) Negative Staining and Image Classification-Powerful Tools in Modern Electron Microscopy. *Biol Proced Online* 6:23–34
- Okuno D, Iino R, Noji H (2011) Rotation and structure of FoF1-ATP synthase. *J Biochem*. 149:655–664
- Otting G (2010) Protein NMR using paramagnetic ions. *Annu Rev Biophys* 39:387–405. <https://doi.org/10.1146/annurev.biophys.093008.131321>
- Park SJ, Kostic M, Dyson HJ (2011) Dynamic interaction of Hsp90 with its client protein p53. *J Mol Biol* 411:158–173. <https://doi.org/10.1016/j.jmb.2011.05.030>
- Pelikan M, Hura GL, Hammel M (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen Physiol Biophys* 28:174–189
- Petsko GA, Ringe D (1984) FLUCTUATIONS IN PROTEIN STRUCTURE FROM X-RAY DIFFRACTION. *Ann Rev Biophys Bioeng* 13:331–371
- Petter R, Salek J, Sikorski C, et al (1990) Cooperative Binding by Aggregated Mono-6-(alkylamino)- β -cyclodextrins. *J Am Chem Soc* 112:3860–3868
- Picard HSP90 INTERACTORS Chaperones and relatives. <http://www.picard.ch/downloads/Hsp90interactors.pdf>. Accessed 13 Oct 2021
- Pick E, Kluger Y, Giltner JM, et al (2007) High HSP90 expression is associated with decreased survival in breast cancer. *Cancer Res* 67:2932–2937. <https://doi.org/10.1158/0008-5472.CAN-06-4511>
- Pintacuda G, John M, Su XC, Otting G (2007) NMR structure determination of protein-Ligand complexes by lanthanide labeling. *Acc Chem Res* 40:206–212. <https://doi.org/10.1021/ar050087z>
- Plevin M, Boisbouvier J (2012) Isotope-Labeling of Methyl Groups for NMR Studies of Large Proteins. In: Clore M, Potts J (eds) *Recent Developments in Biomolecular NMR*. Royal Society Chemistry
- Pritišanac I, Alderson TR, Güntert P (2020) Automated assignment of methyl NMR spectra from large proteins. *Prog Nucl Magn Reson Spectrosc* 118–119:54–73. <https://doi.org/10.1016/j.pnmrs.2020.04.001>

- Pritišanac I, Degiacomi MT, Alderson TR, et al (2017) Automatic Assignment of Methyl-NMR Spectra of Supramolecular Machines Using Graph Theory. *J Am Chem Soc* 139:9523–9533. <https://doi.org/10.1021/jacs.6b11358>
- Pritišanac I, Würz JM, Alderson TR, Güntert P (2019) Automatic structure-based NMR methyl resonance assignment in large proteins. *Nat Commun* 10:4922. <https://doi.org/10.1038/s41467-019-12837-8>
- Prodromou C (2012) The “active life” of Hsp90 complexes. *Biochim Biophys Acta - Mol Cell Res* 1823:614–623. <https://doi.org/10.1016/j.bbamcr.2011.07.020>
- Prodromou C, Pearl L (2014) Structural Basis of Hsp90 Function. In: Machajewski T, Gao Z (eds) *Inhibitors of Molecular Chaperones as Therapeutic Agents*. Royal Society Chemistry
- Prodromou C, Roe SM (1997) Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone. *Cell* 90:65–75. [https://doi.org/10.1016/s0092-8674\(00\)80314-1](https://doi.org/10.1016/s0092-8674(00)80314-1)
- Religa T, Sprangers R, Kay LE (2010) Dynamic Regulation of Archaeal Proteasome Gate Opening As Studied by TROSY NMR. *Science* 328:98–102. <https://doi.org/10.1126/science.1178994>
- Rosenzweig R, Moradi S, Zarrine-Afsar A, et al (2013) Unraveling the Mechanism of Protein Disaggregation Through a ClpB-DnaK Interaction. *Science* 339:1080–1083. <https://doi.org/10.1126/science.1233066>
- Rowlands MG, Newbatt YM, Prodromou C, et al (2004) High-throughput screening assay for inhibitors of heat-shock protein 90 ATPase activity. *Anal Biochem* 327:176–183. <https://doi.org/10.1016/j.ab.2003.10.038>
- Roy R, Hohng S, Ha T (2008) A practical guide to single-molecule FRET. *Nat. Methods* 5:507–516
- Rózycki B, Kim YC, Hummer G (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure* 19:109–116. <https://doi.org/10.1016/j.str.2010.10.006>
- Samish I, Gu J, Klein ML (2009) PROTEIN MOTION: SIMULATION. pp 909–939
- Sanchez-Martin C, Serapian SA, Colombo G, Rasola A (2020) Dynamically Shaping Chaperones. Allosteric Modulators of HSP90 Family as Regulatory Tools of Cell Metabolism in Neoplastic Progression. *Front Oncol* 10:1177. <https://doi.org/10.3389/fonc.2020.01177>
- Saviotti ML, Galley WC (1974) Room Temperature Phosphorescence and the Dynamic Aspects of Protein Structure. *Proc Nat Acad Sci* 71:4154–4158
- Schanda P, Brutscher B (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* 127:8014–8015. <https://doi.org/10.1021/ja051306e>
- Schanda P, Kupče E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33:199–211. <https://doi.org/10.1007/s10858-005-4425-x>
- Scheres SHW (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–530. <https://doi.org/10.1016/j.jsb.2012.09.006>
- Schilb A, Riou V, Schoepfer J, et al (2004) Development and implementation of a highly miniaturized confocal 2D-FIDA-based high-throughput screening assay to search for active site modulators of the human heat shock protein 90 β . *J Biomol Screen* 9:569–577. <https://doi.org/10.1177/1087057104265538>
- Schopf FH, Biebl MM, Buchner J (2017) The HSP90 chaperone machinery. *Nat. Rev. Mol. Cell Biol.* 18:345–360

- Schuler B (2013) Single-molecule FRET of protein structure and dynamics-a primer
- Schulze A, Beliu G, Helmerich DA, et al (2016) Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism. *Nat Chem Biol* 12:628–635. <https://doi.org/10.1038/nchembio.2111>
- Sekhar A, Rumpfheldt JA, Broom HR, et al (2015) Thermal fluctuations of immature SOD1 lead to separate folding and misfolding pathways. *Elife* 4:e07296 <https://doi.org/10.7554/eLife.07296.001>
- Shah DM, Ab E, Diercks T, et al (2012) Rapid protein-ligand costructures from sparse NOE data. *J Med Chem* 55:10786–10790. <https://doi.org/10.1021/jm301396d>
- Shimamura T, Shapiro GI (2008) Heat shock protein 90 inhibition in lung cancer. *J Thorac Oncol* 3:. <https://doi.org/10.1097/JTO.0b013e318174ea3a>
- Skinner SP, Moshev M, Hass MAS, Ubbink M (2013) PARAssign - Paramagnetic NMR assignments of protein nuclei on the basis of pseudocontact shifts. *J Biomol NMR* 55:379–389. <https://doi.org/10.1007/s10858-013-9722-1>
- Skjærven L, Reuter N, Martinez A (2011) Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules: A computational approach. *Future Med Chem* 3:2079–2100. <https://doi.org/10.4155/fmc.11.159>
- Skrynnikov NR, Dahlquist FW, Kay LE (2002) Reconstructing NMR spectra of “invisible” excited protein states using HSQC and HMQC experiments. *J Am Chem Soc* 124:12352–12360. <https://doi.org/10.1021/ja0207089>
- Smyth MS, Martin JHJ (2000) Review x Ray crystallography. *J Clin Pathol Mol Pathol* 53:8–14
- Softley CA, Bostock MJ, Popowicz GM, Sattler M (2020) Paramagnetic NMR in drug discovery. *J Biomol NMR* 74:287–309. <https://doi.org/10.1007/s10858-020-00322-0>
- Soga S, Shiotsu Y, Akinaga S, Sharma S V (2003) Development of Radicol Analogues. *Curr Cancer Drug Targets* 3:359–369
- Solit DB, Chiosis G (2008) Development and application of Hsp90 inhibitors. *Drug Discov. Today* 13:38–43
- Song HY, Dunbar JD, Yuan Xin Zhang, et al (1995) Identification of a protein with homology to hsp90 that binds the type 1 tumor necrosis factor receptor. *J Biol Chem* 270:3574–3581. <https://doi.org/10.1074/jbc.270.8.3574>
- Sorger PK, Pelham HR (1987) The Glucose-regulated Protein grp94 is Related to Heat Shock Protein hsp90. *J Mol Biol* 194:341–344. [https://doi.org/10.1016/0022-2836\(87\)90380-9](https://doi.org/10.1016/0022-2836(87)90380-9)
- Sounier R, Blanchard L, Wu Z, Boisbouvier J (2007) High-accuracy distance measurement between remote methyls in specifically protonated proteins. *J Am Chem Soc* 129:472–473. <https://doi.org/10.1021/ja067260m>
- Sprangers R, Gribun A, Hwang PM, et al (2005) Quantitative NMR spectroscopy of supramolecular complexes: Dynamic side pores in ClpP are important for product release. *Proc Natl Acad Sci* 102:16678–16683
- Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445:618–622. <https://doi.org/10.1038/nature05512>
- Šrajter V, Schmidt M (2017) Watching proteins function with time-resolved x-ray crystallography. *J Phys D Appl Phys* 50(37):373001. <https://doi.org/10.1088/1361-6463/aa7d32>
- Sydor JR, Normant E, Pien CS, et al (2006) Development of 17-allylamino-17-demethoxygeldanamycin hydroquinone hydrochloride (IPI-504), an anti-cancer agent directed against Hsp90. *Proc Natl Acad Sci* 103:17408–17413

- Taipale M, Jarosz DF, Lindquist S (2010) HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat Rev Mol Cell Biol* 11:515–528. <https://doi.org/10.1038/nrm2918>
- Tanford C, Reynolds J (2001) *Nature's robots : A history of proteins*. Oxford University Press
- Teilum K, Olsen JG, Kragelund BB (2009) Functional aspects of protein flexibility. *Cell Mol Life Sci* 66:2231–2247. <https://doi.org/10.1007/s00018-009-0014-6>
- Tenboer J, Basu S, Zatsepin N, et al (2014) Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* 346:1242–1246. <https://doi.org/10.1126/science.1259357>
- Toyama Y, Kano H, Mase Y, et al (2018) Structural basis for the ethanol action on G-protein-activated inwardly rectifying potassium channel 1 revealed by NMR spectroscopy. *Proc Natl Acad Sci U S A* 115:3858–3863. <https://doi.org/10.1073/pnas.1722257115>
- Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H-13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* 125:10420–10428. <https://doi.org/10.1021/ja030153x>
- Tugarinov V, Kay LE (2003a) Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J Am Chem Soc* 125:13868–13878. <https://doi.org/10.1021/ja030345s>
- Tugarinov V, Kay LE (2003b) Side Chain Assignments of Ile δ 1 Methyl Groups in High Molecular Weight Proteins: an Application to a 46 ns Tumbling Molecule. *J Am Chem Soc* 125:5701–5706. <https://doi.org/10.1021/ja021452>
- Unwin PNT, Henderson R (1975) Molecular Structure Determination by Electron Microscopy of Unstained Crystalline Specimens. *J Mol Biol* 94:425–440
- Vallurupalli P (2009) *Chemical Exchange Workshop Series on NMR and Related Topics*
- Vallurupalli P, Bouvignies G, Kay LE (2012) Studying “invisible” excited protein states in slow exchange with a major state conformation. *J Am Chem Soc* 134:8148–8161. <https://doi.org/10.1021/ja3001419>
- Vallurupalli P, Bouvignies G, Kay LE (2011) Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR. *J Phys Chem B* 115:14891–14900. <https://doi.org/10.1021/jp209610v>
- Vallurupalli P, Chakrabarti N, Pomès R, Kay LE (2016) Atomistic picture of conformational exchange in a t4 lysozyme cavity mutant: An experiment-guided molecular dynamics study. *Chem Sci* 7:3602–3613. <https://doi.org/10.1039/c5sc03886c>
- Vallurupalli P, Sekhar A, Yuwen T, Kay LE (2017) Probing conformational dynamics in biomolecules via chemical exchange saturation transfer: a primer. *J Biomol NMR* 67:243–271. <https://doi.org/10.1007/s10858-017-0099-4>
- VanOudenhove J, Anderson E, Krueger S, Cole JL (2009) Analysis of PKR Structure by Small-Angle Scattering. *J Mol Biol* 387:910–920. <https://doi.org/10.1016/j.jmb.2009.02.019>
- Velyvis A, Schachman HK, Kay LE (2009) Assignment of Ile, Leu, and Val methyl correlations in supra-molecular systems: An application to aspartate transcarbamoylase. *J Am Chem Soc* 131:16534–16543. <https://doi.org/10.1021/ja906978r>

- Venditti V, Fawzi NL, Clore GM (2011) Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear overhauser enhancement spectroscopy. *J Biomol NMR* 51:319–328. <https://doi.org/10.1007/s10858-011-9559-4>
- Verba KA, Wang RYR, Arakawa A, et al (2016) Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* 352:1542–1547. <https://doi.org/10.1126/science.aaf5023>
- Vögeli B, Orts J, Strotz D, et al (2014) Towards a true protein movie: A perspective on the potential impact of the ensemble-based structure determination using exact NOEs. *J Magn Reson* 241:53–59. <https://doi.org/10.1016/j.jmr.2013.11.016>
- Wand AJ (2013) The dark energy of proteins comes to light: Conformational entropy and its role in protein function revealed by NMR relaxation. *Curr Opin Struct Biol* 23:75–81. <https://doi.org/10.1016/j.sbi.2012.11.005>
- Wang CC, Houg HC, Chen CL, et al (2009) Solution structure and backbone dynamics of streptopain: Insight into diverse substrate specificity. *J Biol Chem* 284:10957–10967. <https://doi.org/10.1074/jbc.M807624200>
- Wang L, Zhang Q, You Q (2021) Targeting the HSP90–CDC37–kinase chaperone cycle: A promising therapeutic strategy for cancer. *Med Res Rev* 42:156–182. <https://doi.org/10.1002/med.21807>
- Weinert T, Skopintsev P, James D, et al (2019) Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science* 365:61–65
- Whitesell L, Lindquist SL (2005) HSP90 and the chaperoning of cancer. *Nat Rev Cancer* 5:761–772. <https://doi.org/10.1038/nrc1716>
- Whitesell L, Mimnaugh EG, De Costat B, et al (1994) Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: Essential role for stress proteins in oncogenic transformation. *Cell Biol* 91:8324–8328
- Wiesner S, Sprangers R (2015) Methyl groups as NMR probes for biomolecular interactions. *Curr Opin Struct Biol* 35:60–67. <https://doi.org/10.1016/j.sbi.2015.08.010>
- Wlodawer A, Minor W, Dauter Z, Jaskolski M (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J* 275:1–21. <https://doi.org/10.1111/j.1742-4658.2007.06178.x>
- Xie T, Saleh T, Rossi P, Kalodimos CG (2020) Conformational states dynamically populated by a kinase determine its function. *Science* 370(6513):eabc2754. <https://doi.org/10.1126/science.abc2754>
- Xu Y, Liu M, Simpson PJ, et al (2009) Automated assignment in selectively methyl-labeled proteins. *J Am Chem Soc* 131:9480–9481. <https://doi.org/10.1021/ja9020233>
- Xu Y, Matthews S (2013) MAP-XSII: An improved program for the automatic assignment of methyl resonances in large proteins. *J Biomol NMR* 55:179–187. <https://doi.org/10.1007/s10858-012-9700-z>
- Yang S, Blachowicz L, Makowski L, et al (2010a) Multidomain assembled states of Hck tyrosine kinase in solution. *Proc Natl Acad Sci* 107:15757–15762. <https://doi.org/10.1073/pnas.1004569107//DCSupplemental>
- Yang X, Welch JL, Arnold JJ, Boehr DD (2010b) Long-range interaction networks in the function and fidelity of poliovirus RNA-dependent RNA polymerase studied by nuclear magnetic resonance. *Biochemistry* 49:9361–9371. <https://doi.org/10.1021/bi100833r>

- Yang ZQ, Geng X, Solit D, et al (2004) New efficient synthesis of resorcinylic macrolides via ynolides: Establishment of cycloproparadicicol as synthetically feasible preclinical anticancer agent based on Hsp90 as the target. *J Am Chem Soc* 126:7881–7889. <https://doi.org/10.1021/ja0484348>
- Yu X, Veesler D, Campbell MG, et al (2017) Cryo-EM structure of human adenovirus D26 reveals the conservation of structural organization among human adenoviruses. *Sci Adv* 3:e1602670
- Yuno A, Lee M-J, Lee S, et al (2018) Clinical Evaluation and Biomarker Profiling of Hsp90 Inhibitors. *Methods Mol Biol* 1709:423–441. https://doi.org/10.1007/978-1-4939-7477-1_29
- Zhang K, Li S, Kappel K, et al (2019) Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat Commun* 10,5511 <https://doi.org/10.1038/s41467-019-13494-7>
- Zhao J, Benlekbir S, Rubinstein JL (2015) Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. *Nature* 521:241–245. <https://doi.org/10.1038/nature14365>
- Zhou Y, Duncan TM, Cross RL (1997) Subunit rotation in Escherichia coli F₁oF₁-ATP synthase during oxidative phosphorylation. *Proc Natl Acad Sci* 94:10583–10587
- Zuehlke A, Johnson JL (2010) Hsp90 and co-chaperones twist the functions of diverse client proteins. *Biopolymers* 93:211–217. <https://doi.org/10.1002/bip.21292>
- Zuehlke AD, Moses MA, Neckers L (2018) Heat shock protein 90: its inhibition and function. *Philos Trans R Soc Lond B Biol Sci* 373:(1738):20160527. <https://doi.org/10.1098/rstb.2016.0527>

Annexes

7.1. Informations Supplémentaires Article I

Optimized Precursor to Simplify Assignment Transfer between Backbone Resonances and Stereospecifically labelled Valine and Leucine Methyl Groups: Application to Human Hsp90 N-Terminal Domain

Faustine Henot¹, Rime Kerfah², Ricarda Törner¹, Pavel Macek^{1,2}, Elodie Crublet², Pierre Gans¹, Matthias Frech³, Olivier Hamelin⁴, Jerome Boisbouvier^{1, *}

1. Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, avenue des martyrs, F-38044 Grenoble, France.

2. NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.

3. Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany.

4. Univ. Grenoble Alpes, CEA, CNRS, IRIG, CBM-F-38000 Grenoble. France.

* correspondence to be addressed to: jerome.boisbouvier@ibs.fr

SUPPLEMENTARY INFORMATION

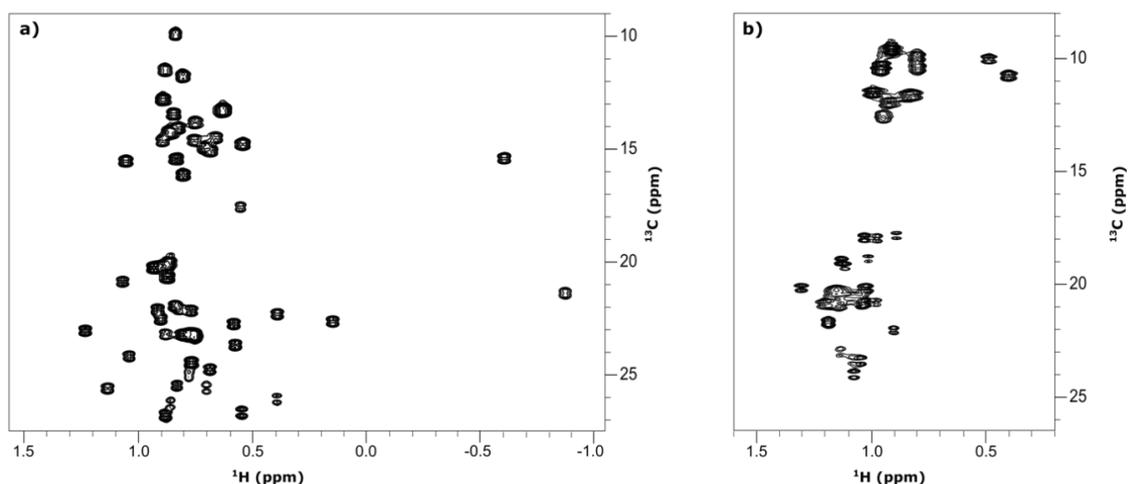


Fig. S1 2D ^1H - ^{13}C SOFAST methyl TROSY spectrum of U- $[^2\text{H}$, ^{15}N , ^{13}C], Ile-[2, 3, 4, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; $[^{13}\text{C}^1\text{H}_3]^{\delta_1}/[^{12}\text{C}^2\text{H}_3]^{\gamma_2}$], Leu-[2, 3, 3, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; $[^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}/[^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}$], Val-[2, 3- $^2\text{H}_2$; 1, 2, 3- $^{13}\text{C}_3$; $[^{13}\text{C}^1\text{H}_3]^{\text{pro-S}}/[^{12}\text{C}^2\text{H}_3]^{\text{pro-R}}$] HSP90-NTD **a**) and prefoldin **b**) samples (labelling scheme A applied to prefoldin β subunit only), recorded before the HCC experiments to check the sample quality. The labelling strategy used Ile- δ_1 precursor: sodium (S)-2-hydroxy-2-(1',1'- $[^2\text{H}_2]$, 1', 2'- $[^{13}\text{C}_2]$) ethyl-3-oxo-1,2,3- $[^{13}\text{C}_3]$ -4,4,4- $[^2\text{H}_3]$ -butanoate) and the suitably labelled acetolactate precursor: 1, 2, 3- $[^{13}\text{C}_3]$ -2- $[^{13}\text{C}^1\text{H}_3]$ -2-[O ^2H]-3-oxo-4, 4, 4- $[^2\text{H}_3]$ -butanoate for the labelling of Leu and valine pro-S methyl groups. No additional peaks were detected neither to the pro-R methyls groups of leucine and valine nor to the isoleucine- γ_2 site, confirming the absence of isotopic scrambling. The peaks are splitted in the carbon dimension due to the presence of $^1J_{\text{CC}}$ coupling as expected from the labelling scheme of the produced samples.

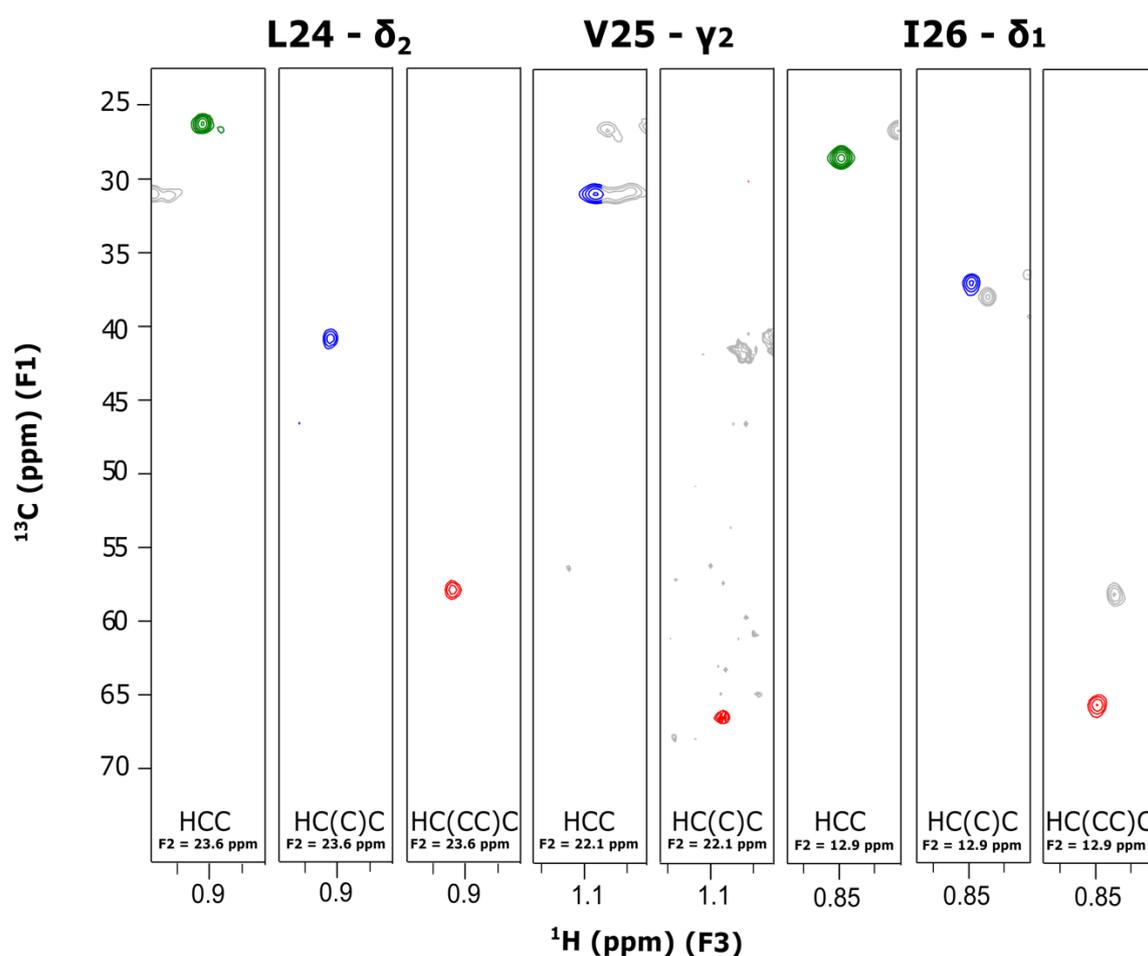


Fig. S2 Application to the 87 kDa prefoldin from *Pyrococcus horikoshii*. Examples of 2D-extracts from 3D ‘out and back’ HCC , HC(C)C and HC(CC)C experiments correlating ^1H (F₃) and ^{13}C (F₂) methyl resonances with $^{13}\text{C}_\gamma$ (green) $^{13}\text{C}_\beta$ (blue) or $^{13}\text{C}_\alpha$ (red) in F₁ dimension of Leu-24, Val-25 and Ile-26 of *Pyrococcus horikoshii* prefoldin subunit β . Spectra have been recorded at 310 K using 0.2 mM U- ^{2}H , ^{15}N , ^{13}C], Ile-[2, 3, 4, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\delta 1}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\gamma 2}$], Leu-[2, 3, 3, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$], Val-[2, 3- $^2\text{H}_2$; 1, 2, 3- $^{13}\text{C}_3$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$] sample of prefoldin on an NMR spectrometer operating at a proton frequency of 950 MHz. The acquisition parameters were similar to that used for HSP90-NTD, described in Materials and Methods section. Prefoldin was expressed and purified according to the protocol described in Törner et al. (2020) and labelled on the β subunit only using labelling scheme A, described in Materials and Methods section of the present manuscript.

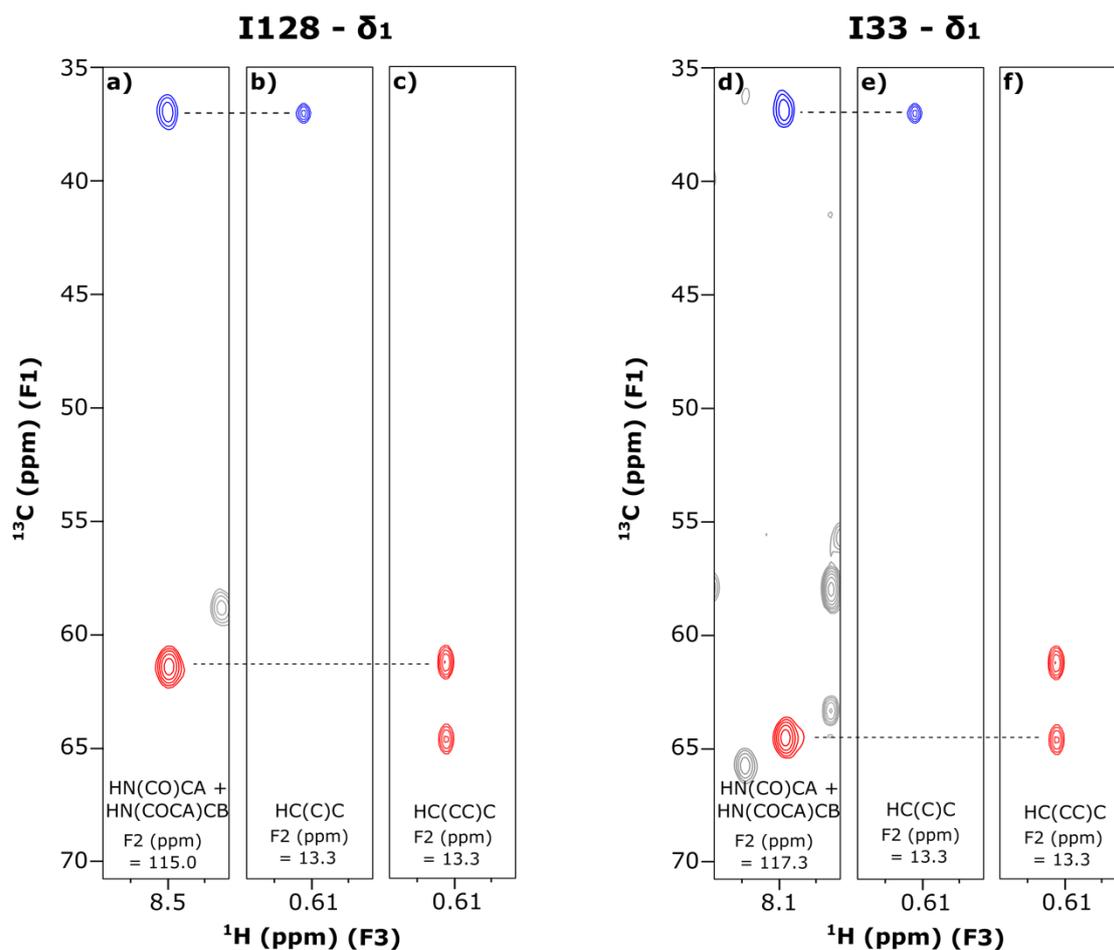


Fig. S3 Assignment transfer from the backbone to the Ile- δ_1 methyl groups of Ile-33 and Ile-128 of HSP90-NTD. Examples of 2D-extracts from 3D 'out and back' HC(C)C (b and e) and HC(CC)C (c and f) experiments correlating ^1H (F₃) and ^{13}C (F₂) methyl resonances with $^{13}\text{C}_\beta$ (blue) or $^{13}\text{C}_\alpha$ (red) in F₁ dimension. Panels a and d display the corresponding 2D HN(CO)CA and HN(COCA)CB extracts for Ile-128 (a) and Ile-33 (d) allowing to connect Ile-33- δ_1 and Ile-128- δ_1 methyl groups to previously assigned backbone atoms. 3D spectra were recorded on an NMR spectrometer operating at a proton frequency of 600 MHz using the U- ^2H , ^{15}N , ^{13}C], Ile-[2, 3, 4, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\delta_1}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\text{V}2}$], Leu-[2, 3, 3, 4- $^2\text{H}_4$; 1, 2, 3, 4- $^{13}\text{C}_4$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$], Val-[2, 3- $^2\text{H}_2$; 1, 2, 3- $^{13}\text{C}_3$; [$^{13}\text{C}^1\text{H}_3$] $^{\text{pro-S}}$ /[$^{12}\text{C}^2\text{H}_3$] $^{\text{pro-R}}$] labelled sample or U- ^2H , ^{15}N , ^{13}C] labelled sample (3D HN(CO)CA and HN(COCA)CB).

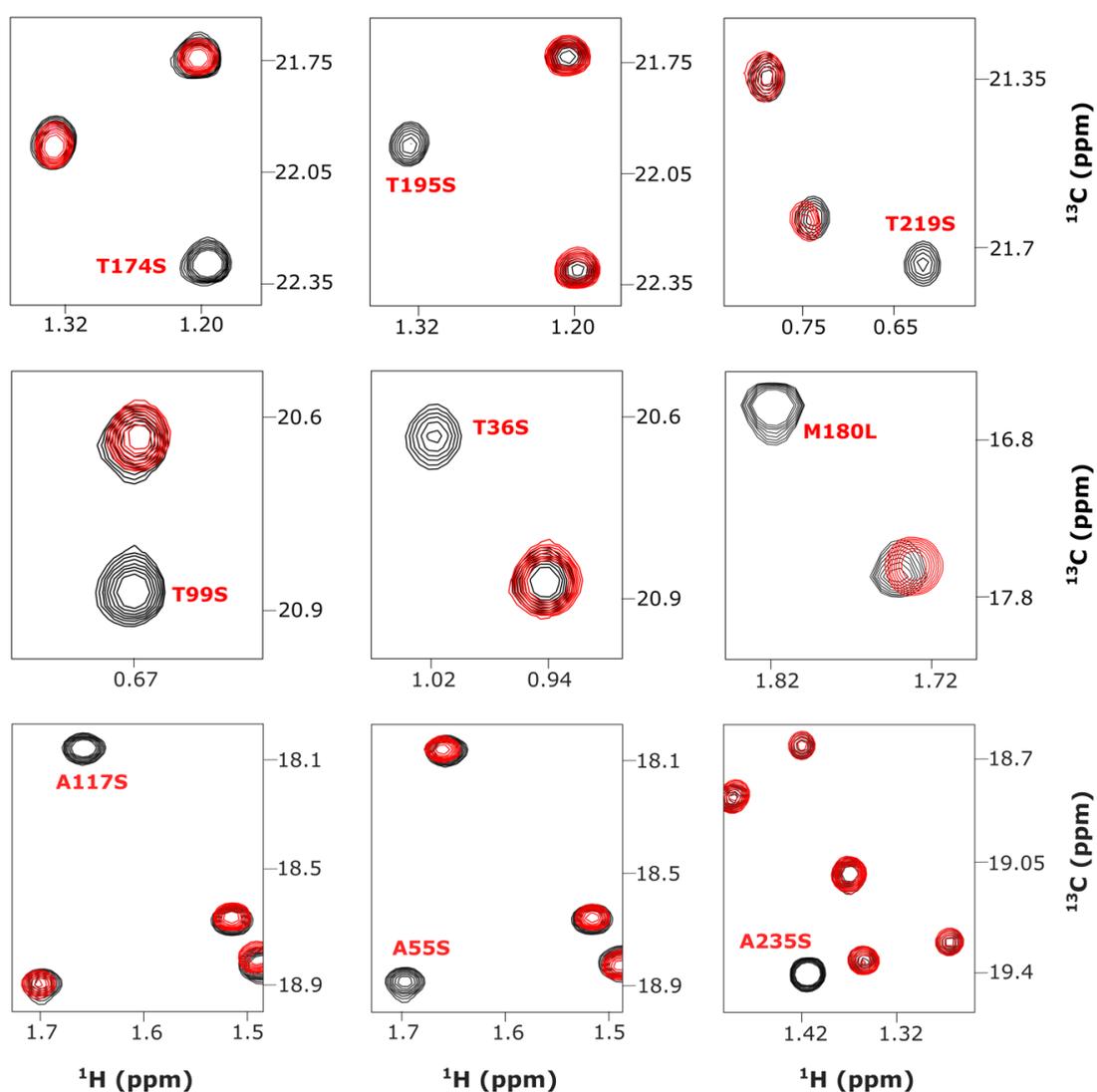


Fig. S4 Assignment of methyl groups using mutagenesis. Examples of 2D ^1H - ^{13}C SOFAST methyl TROSY spectra of HSP90-NTD mutants, for which no or minor chemical shift perturbations are observed. Spectra were acquired on an NMR spectrometer operating at a proton frequency of 850 MHz. Each mutant spectrum extract (colored) is superimposed with the corresponding spectra acquired using wild type HSP90-NTD methyl labelled spectra (black). The signal corresponding to each mutant is annotated in color. List of the mutants produced for this study: M12L, M30L, M119L, M130L, M180L, A21S, A24S, A27S, A55S, A101S, A111S, A117S, A121S, A124S, A126S, A161S, A166S, A235S, T19S, T36S, T65S, T99S, T109S, T115S, T152S, T171S, T174S, T184S, T195S, T219S, I104V, L107A, I110V.

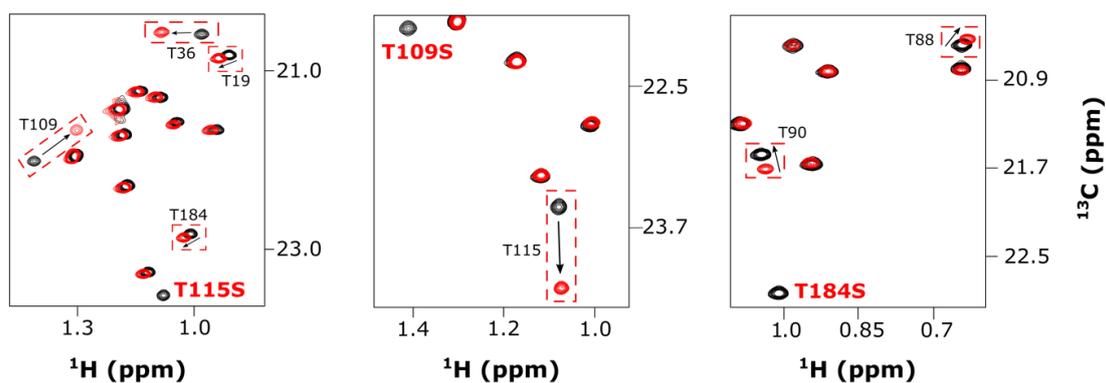


Fig. S5 Assignment of methyl groups using mutagenesis. Examples of 2D ^1H - ^{13}C SOFAST methyl TROSY spectra of HSP90-NTD mutants, for which chemical shift perturbations. Spectra were acquired on an NMR spectrometer operating at a proton frequency of 850 MHz. Each mutant spectrum extract (colored) is superimposed with the corresponding spectra acquired using wild type HSP90-NTD methyl labelled spectra (black). The signal corresponding to each mutant is annotated in color. Black arrows represent secondary chemical shifts.

Table S1: Output file from MAGIC software for the assignment of HSP90-NTD methyl groups.

The 1st column displays the best assignment. The 2nd and 3rd columns give the ¹³C and ¹H chemical shifts, respectively. The 4th column represents the sum of all peak-peak connection confident scores for each peak. The 5th column shows the NOE assignment completeness of the strip related to each peak and the 6th column displays the list of all possible assignment with their scores (Monneau et al. 2017).

The automated methyl assignment was performed using the reference structure of HSP90-NTD (PDB: 1YES). MAGIC was run with a score threshold factor of 1, distance thresholds of 7–10 Å and using the 344 inter methyl NOE cross peaks (S/N ≥ 5) detected (Fig. 4e). In addition, MAGIC was given the methyl type for each methyl group, previously identified using specific methyl labelling. In blue are the methyl groups assigned automatically by MAGIC that have 1) a single assignment, 2) a high NOE assignment completeness of the strip related to each peak (> 50 %) and 3) a high total confidence score value (≥ 7).

Predicted Assignment	¹³ C ppm	¹ H ppm	Score	% NOE	Possible Assignment
V144CG*-HG*	15.487	-0.566	31.48	0.86	{'V144CG': 215.504}
L56CD*-HD*	26.036	0.428	9.0	1.0	{'L56CD': 215.504}
L89CD*-HD*	25.59	1.159	26.22	1.0	{'L89CD': 215.504}
L45CD*-HD*	25.544	0.731	36.08	1.0	{'L80CD': 207.896 'L45CD': 215.504}
L80CD*-HD*	25.474	0.86	33.17	1.0	{'L80CD': 215.504 'L45CD': 207.896}
L198CD*-HD*	24.747	0.717	15.0	1.0	{'L198CD': 215.504}
L32CD*-HD*	24.451	0.802	13.67	0.8	{'L32CD': 215.504}
L143CD*-HD*	24.158	1.061	13.0	1.0	{'L143CD': 215.504 'L122CD': 206.867}
L76CD*-HD*	23.647	0.605	9.0	1.0	{'L70CD': 207.413 'L76CD': 215.504}
T94CG2-HG2	23.292	1.157	7.0	1.0	{'T184CG2': 206.767 'T152CG2': 207.402 'T94CG2': 215.504}
T115CG2-HG2	23.566	1.118	24.6	1.0	{'T115CG2': 215.504}
M130CE-HE	17.026	2.213	5.0	0.25	{'M130CE': 215.504}
V207CG*-HG*	23.223	0.91	27.69	1.0	{'V207CG': 215.504}
V172CG*-HG*	23.024	1.257	3.0	0.5	{'V172CG': 215.504}
L190CD*-HD*	22.76	0.615	33.53	1.0	{'L190CD': 215.504}

V148CG*-HG*	22.646	0.178	24.0	0.83	{'V148CG': 215.504}	
L48CD*-HD*	22.527	0.93	19.5	0.67	{'L48CD': 215.504}	
T19CG2-HG2	22.331	1.204	2.0	0.0	{'T19CG2': 215.504 'T174CG2': 215.504}	{'T176CG2': 215.504 'T171CG2': 215.504}
V92CG*-HG*	22.125	0.946	27.35	1.0	{'V92CG': 215.504}	
L70CD*-HD*	22.186	0.801	10.39	1.0	{'L70CD': 215.504}	{'L76CD': 207.413}
I214CD1-HD1	10.135	0.86	13.38	1.0	{'I78CD1': 207.021}	{'I214CD1': 215.504}
I26CD1-HD1	11.564	0.908	14.0	1.0	{'I26CD1': 215.504}	
I110CD1-HD1	11.885	0.831	2.0	1.0	{'I110CD1': 215.504}	
I151CD1-HD1	12.894	0.916	15.0	0.57	{'I151CD1': 215.504}	
I33CD1-HD1	13.355	0.661	36.6	1.0	{'I33CD1': 215.504}	
I203CD1-HD1	13.467	0.874	33.35	1.0	{'I203CD1': 215.504}	
I49CD1-HD1	13.848	0.775	34.45	0.86	{'I49CD1': 215.504}	
I34CD1-HD1	14.091	0.859	2.0	1.0	{'I34CD1': 215.504}	
I187CD1-HD1	14.32	0.88	12.37	1.0	{'I187CD1': 215.504}	
I59CD1-HD1	14.25	0.89	8.37	0.75	{'I96CD1': 207.538}	{'I59CD1': 215.504}
I131CD1-HD1	14.458	0.688	23.0	0.86	{'I131CD1': 215.504}	
I78CD1-HD1	14.622	0.779	24.0	1.0	{'I218CD1': 207.374 'I214CD1': 207.021}	{'I78CD1': 215.504}
M119CE-HE	16.667	2.047	26.67	0.71	{'M119CE': 215.504}	
I218CD1-HD1	14.66	0.916	22.0	0.8	{'I218CD1': 215.504 'I214CD1': 206.692}	{'I78CD1': 207.374}
I104CD1-HD1	14.808	0.565	4.0	1.0	{'I104CD1': 215.504}	
I81CD1-HD1	14.949	0.733	42.89	0.45	{'I81CD1': 215.504}	
I43CD1-HD1	15.1	0.714	10.0	1.0	{'I43CD1': 215.504}	{'I28CD1': 208.077}
I96CD1-HD1	15.467	0.865	5.92	0.67	{'I96CD1': 215.504}	{'I59CD1': 207.538}
I206CD1-HD1	15.535	1.076	28.33	0.83	{'I206CD1': 215.504}	
I91CD1-HD1	16.128	0.832	77.51	0.81	{'I91CD1': 215.504}	
A124CB-HB	16.617	1.357	6.0	0.67	{'A121CB': 215.504 'A124CB': 215.504}	{'A126CB': 207.223}
A166CB-HB	16.771	1.262	2.0	1.0	{'A166CB': 215.504}	
A27CB-HB	17.326	1.486	2.0	1.0	{'A27CB': 215.504}	{'A24CB': 215.504}
V136CG*-HG*	17.675	0.576	19.67	1.0	{'V136CG': 215.504}	

A101CB-HB	17.824	1.225	3.0	1.0	{'A101CB': 215.504}	
A141CB-HB	18.538	0.827	15.0	1.0	{'A145CB': 207.062	'A141CB': 215.504}
A121CB-HB	18.655	1.528	4.33	1.0	{'A121CB': 215.504	'A117CB': 207.058 'A126CB': 207.223 'A124CB': 215.504}
A111CB-HB	18.69	1.429	5.0	1.0	{'A111CB': 215.504	'A27CB': 207.455 'A24CB': 207.059}
A24CB-HB	18.812	1.5	2.0	1.0	{'A27CB': 215.504	'A24CB': 215.504}
A126CB-HB	19.272	1.272	9.83	1.0	{'A121CB': 207.223	'A126CB': 215.504}
M180CE-HE	16.604	1.829	29.55	0.64	{'M180CE': 215.504}	
T152CG2-HG2	19.402	1.427	3.23	0.4	{'T152CG2': 215.504}	
V222CG*-HG*	20.053	0.894	5.0	1.0	{'V222CG': 215.504}	
T36CG2-HG2	20.646	1.008	8.0	1.0	{'T36CG2': 215.504}	
T88CG2-HG2	20.629	0.675	8.0	0.75	{'T88CG2': 215.504	'T195CG2': 207.358}
T99CG2-HG2	20.83	0.676	2.0	0.33	{'T99CG2': 215.504}	
L220CD*-HD*	26.75	0.906	12.31	1.0	{'L220CD': 215.504}	
V186CG*-HG*	20.862	1.1	19.26	0.57	{'V150CG': 207.024	'V186CG': 215.504}
T176CG2-HG2	21.263	1.172	2.0	0.0	{'T171CG2': 215.504	'T176CG2': 215.504 'T174CG2': 215.504 'T19CG2': 215.504}
T149CG2-HG2	21.331	1.498	9.0	1.0	{'T149CG2': 215.504}	
A145CB-HB	21.67	0.932	16.27	1.0	{'A145CB': 215.504	'A141CB': 207.062}
T90CG2-HG2	21.638	1.078	29.0	0.71	{'T90CG2': 215.504}	
T195CG2-HG2	21.988	1.34	5.0	1.0	{'T195CG2': 215.504	'T88CG2': 207.358}
L188CD*-HD*	26.645	0.575	38.1	0.78	{'L188CD': 215.504}	
T109CG2-HG2	22.02	1.444	3.0	1.0	{'T109CG2': 215.504}	
L122CD*-HD*	22.419	0.431	8.1	0.75	{'L143CD': 206.867	'L122CD': 215.504}
L103CD*-HD*	24.974	0.797	2.73	1.0	{'L103CD': 215.504	'L107CD': 207.733 'L64CD': 207.687}
A161CB-HB	21.9	1.327	7.0	0.67	{'A161CB': 215.504}	
L29CD*-HD*	26.252	0.89	27.0	1.0	{'L29CD': 215.504}	
A55CB-HB	19.402	1.427	3.23	0.0	{'A24CB': 208.147	'A166CB': 208.147 'A21CB': 215.504 'A111CB': 207.455 'A27CB': 208.147 'A117CB': 215.504 'A235CB': 215.504 'A55CB': 215.504}

					'A124CB': 207.058}
T65CG2-HG2	20.867	0.944	2.0	0.0	{'T171CG2': 215.504 'T176CG2': 215.504 'T174CG2': 215.504 'T65CG2': 215.504 'T19CG2': 215.504 'T219CG2': 215.504}
T184CG2-HG2	21.357	1.117	2.0	1.0	{'T171CG2': 207.047 'T152CG2': 207.267 'T94CG2': 206.547 'T176CG2': 207.047 'T174CG2': 207.047 'T65CG2': 207.047 'T19CG2': 207.047 'T219CG2': 207.047 'T184CG2': 215.504}
T219CG2-HG2	21.756	1.216	3.0	0.0	{'T171CG2': 215.504 'T152CG2': 208.147 'T94CG2': 207.402 'T176CG2': 215.504 'T174CG2': 215.504 'T65CG2': 215.504 'T19CG2': 215.504 'T219CG2': 215.504 'T184CG2': 207.267}

Not Assigned	¹³ C ppm	¹ H ppm	
3	16.949	2.087	NotAss ['M30CE' 'M98CE' 'M12CE']
5	17.58	1.732	NotAss ['M30CE' 'M98CE' 'M12CE']
49	16.898	1.43	NotAss ['M30CE' 'M98CE' 'M12CE']
53	18.034	1.659	NotAss ['A21CB' 'A117CB' 'A235CB']
58	18.938	1.694	NotAss ['A21CB' 'A117CB' 'A235CB']
60	19.325	1.362	NotAss ['A21CB' 'A117CB' 'A235CB']
64	20.266	0.959	NotAss ['V17CG' 'V150CG' 'V230CG']
67	20.699	0.9	NotAss ['V17CG' 'V150CG' 'V230CG']
73	21.392	-0.838	NotAss ['L64CD' 'L107CD']
75	21.709	0.974	NotAss ['T171CG2' 'T174CG2']
80	21.965	0.865	NotAss ['L64CD' 'L107CD']
82	22.208	0.844	NotAss ['V17CG' 'V150CG' 'V230CG']
86	22.86	1.022	NotAss ['T171CG2' 'T174CG2']

Table S2: Assignment of HSP90-NTD methyl groups

Met	H(ϵ)	C(ϵ)
M12	2.117	16.889
M30	1.426	16.886
M98	1.741	17.594
M119	2.035	16.644
M130	2.210	17.000
M180	1.818	16.593

Thr	H(γ_2)	C(γ_2)
T19	0.942	20.832
T36	1.002	20.622
T65	1.205	21.746
T88	0.664	20.605
T90	1.07	21.611
T94	1.144	23.293
T99	0.667	20.819
T109	1.437	22.033
T115	1.104	23.567
T149	1.492	21.304
T152	1.430	19.470
T171	1.111	21.333
T174	1.199	22.314
T176	1.169	21.250
T184	1.036	22.863
T195	1.326	21.960
T219	0.973	21.686

Ala	H(β)	C(β)
A21	1.356	19.310
A24	1.489	18.799
A27	1.476	17.297
A55	1.696	18.870

A101	1.211	17.806
A111	1.419	18.635
A117	1.654	18.031
A121	1.353	16.570
A124	1.514	18.640
A126	1.265	19.266
A141	0.819	18.528
A145	0.917	21.654
A161	1.318	21.910
A166	1.257	16.777
A235	1.415	19.384

Ile	H(δ_1)	C(δ_1)
I26 ^b	0.899	11.547
I33/I128 ^a	0.647	13.317
I34 ^a	0.842	14.078
I43 ^a	0.702	15.097
I49 ^a	0.765	13.830
I59 ^a	0.876	14.287
I78 ^a	0.771	14.631
I81 ^a	0.729	14.935
I91 ^a	0.818	16.139
I96 ^a	0.850	15.440
I104 ^a	0.585	14.778
I110 ^b	0.821	11.783
I131 ^a	0.675	14.498
I151 ^a	0.910	12.853
I187 ^a	0.876	14.287
I203 ^a	0.861	13.455
I206 ^a	1.069	15.528
I214 ^a	0.853	10.015
I218 ^a	0.909	14.670

Leu	H(δ_2)	C(δ_2)
-----	-----------------	-----------------

L29 ^c	0.876	26.240
L32 ^c	0.786	24.429
L45 ^d	0.719	25.537
L48 ^c	0.918	22.526
L56 ^c	0.411	26.037
L64 ^c	0.853	21.951
L70 ^c	0.787	22.168
L76 ^c	0.594	23.642
L80 ^c	0.846	25.472
L89 ^c	1.149	25.579
L103 ^c	0.798	25.037
L107 ^d	-0.851	21.374
L122 ^c	0.411	22.329
L143 ^d	1.056	24.148
L188 ^d	0.565	26.652
L190 ^c	0.601	22.739
L198 ^c	0.702	24.73
L220 ^c	0.898	26.742

Val	H(γ^2)	C(γ^2)
V17 ^d	0.892	20.658
V92 ^c	0.936	22.102
V136 ^c	0.570	17.586
V144 ^c	-0.582	15.446
V148 ^c	0.168	22.630
V150 ^c	0.839	22.084
V172 ^c	1.251	23.028
V186 ^c	1.085	20.870
V207 ^c	0.899	23.183
V222 ^c	0.884	20.037
V230	0.949	20.244

^a Residues assigned by (Park et al. 2011) and confirmed by this study.

^b Assignment was inverted for these two residues in our study compared to (Park et al. 2011).

^c Residues stereospecifically assigned by (Lescanne et al. 2018) and confirmed by this study.

^d Stereospecific assignment was inverted compared to (Lescanne et al. 2018).

7.2. Informations Supplémentaires Article III

Visualizing the Transiently Populated Closed-State of Human HSP90 ATP Binding Domain

Faustine Henot^{1,#}, Elisa Rioual^{1,2,#}, Adrien Favier¹, Pavel Macek^{1,3}, Elodie Crublet³, Pierre Josso², Bernhard Brutscher¹, Matthias Frech⁴, Pierre Gans¹, Claire Loison^{2,*}, Jerome Boisbouvier^{1,*}

1. Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, avenue des martyrs, F-38044 Grenoble, France.

2. Institut Lumière Matière, University of Lyon, Université Claude Bernard Lyon 1, CNRS, F-69622, Villeurbanne, France.

3. NMR-Bio, 5 place Robert Schuman, F-38025 Grenoble, France.

4. Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany.

These authors contributed equally to this work

* correspondence to be addressed to: claire.loison@univ-lyon1.fr and jerome.boisbouvier@ibs.fr

SUPPLEMENTARY INFORMATION

Supporting Table 1: Summary of structural restraints.	p. 2
Supporting Table 2: Definition of NOE-derived distance restraints specific of a single state.	p. 3
Supporting Table 3: Summary of NMR refined structure statistics.	p. 4
Supporting Table 4: Identification of characteristic distances of dimeric forms of HSP90 α -NTD.	p. 5
Supporting Table 5: NMR-based structural calculations supporting the absence of helix-5 in ATP-lid closed state.	p. 6
Supporting Figure 1: Assigned 2D NMR spectra of HSP90 α -NTD.	p. 7
Supporting Figure 2: HSP90 α -NTD assignment.	p. 8
Supporting Figure 3: Summary of NOEs distance restraints.	p. 9
Supporting Figure 4: Assigned 2D ¹ H- ¹³ C SOFAST methyl TROSY of HSP90 α -NTD mutants	p. 10
Supporting Figure 5: Intensity variation of characteristic NOEs.	p. 11
Supporting Figure 6: Time evolution of the violations of the NOE-derived distance restraints specific of either the open or the closed conformation for the 20 non-restrained molecular dynamics simulations starting from ATP-lid closed state conformers.	p. 12
Supporting Figure 7: Analysis of HSP90-NTD dynamics using relaxation dispersion experiments.	p. 14
Supporting Figure 8: Computed Evolutionary Couplings (ECs) for HSP90-NTD.	p. 15
Supporting Figure 9: Probability of helical structures along the sequence of the ATP lid, in 4 MD simulations depicting a transition of the ATP-lid from a closed to an open state.	p. 16

Supporting Table 1: Summary of structural restraints used for structure calculation of WT-, R60A- and R46A-HSP90 α -NTD.

		WT ^(a)	R60A	R46A
Experimental distance restraints				
CH₃-CH₃		111	114	81^(b)
Sequential / local ($\leq i+4$) / long range / ambiguous		4 / 12 / 75 / 20	5 / 13 / 78 / 18	4 / 15 / 57 / 5
H_N-H_N		41	41	41
Sequential / local ($\leq i+4$) / long range / ambiguous		18 / 19 / 3 / 1	18 / 19 / 3 / 1	18 / 19 / 3 / 1
H-bonds		0	2	2
Experimental angular restraints				
Dihedrals	φ, ψ	36+54^(a)	54	36^(c)
Restraints for rigid body				
Dihedrals	φ, ψ	475 + 477 ^(a)	477	475
Distances	C α -C α	13434 + 13434 ^(a)	13434	13434

(a) Structure calculations using NMR restraints derived from WT-HSP90-NTD sample were performed using a two-state structure calculation protocol enabling each experimental distance restraint to be satisfied in either ATP-lid open- or closed-state or both states simultaneously⁵⁴. The angular restraints, and distance restraints for the rigid body core (segments 11-97 and 137-223) were applied on both states simultaneously. Dihedral restraints, corresponding to helix-5 in the ATP-lid open state (from residues 128 to 136), were not applied for the closed ATP-lid state (See note (c) for further information).

(b) A total of 103 inter-methyl distance restraints were experimentally extracted using R46A-HSP90 α -NTD sample to characterize the ATP-lid segment (from residue M98 to V136). Removal of R46/S129 hydrogen bond by mutagenesis leads to an increase of the population of the closed state, but NOEs corresponding to an ATP-lid in both open and closed conformations were detected simultaneously. Therefore, we used the previously refined solution structure with the ATP-lid in an open position (obtained using R60A-HSP90 α -NTD sample) to identify and exclude 22 NOEs specifically assigned to the ATP-lid in an open state (see supporting Table 2). The remaining 81 NOEs were completed with backbone restraints previously obtained on WT-HSP90 α -NTD sample for the final structure calculation using a single state structure calculation protocol.

(c) Structure calculation allowed us to identify large violations due to incompatibilities between helix-5 dihedral restraints and NOEs specific of the ATP-lid closed state (Supporting Table 5). Therefore, the angular restraints from residues 128 to 136 were omitted for the final CYANA structure calculation of ATP-lid closed state using restraints derived from the R46A-HSP90 α -NTD sample.

Supporting Table 2: Definition of NOE-derived distance restraints specific of a single state.

A) NOE-derived distance restraints specific of the ATP-lid closed state. NOEs observed only for the WT- and R46A-HSP90 α -NTD mutants and not for R60A-HSP90 α -NTD sample. In the structure ensembles calculated with a two-state protocol using distance restraints derived from NOEs detected on WT- and R46A-HSP90-NTD, those restraints are satisfied only in the ATP-lid closed state and violated by more than 5 Å in the ATP-lid open state.

Restraint	Residue i	Pseudoatom i	Residue j	Pseudoatom j	d_{ij}^{\max} (Å) ^a	d_{ij}^{viol} (Å) ^b
1	LEU64	HD2*	MET130	HE*	6.5	10.08
2	THR65	HG2*	ALA121	HB*	6.5	10.43
3	THR65	HG2*	ALA124	HB*	6.5	9.85
4	THR65	HG2*	ALA126	HB*	6.5	10.19
5	THR65	HG2*	MET130	HE*	6.5	10.19

B) NOE-derived distance restraints specific of the ATP-lid open state. NOEs observed for WT-, R46A-, or R60A-HSP90 α -NTD samples. In the structure ensembles calculated with a two-state protocol using distance restraints derived from NOEs detected on WT- and R46A-HSP90-NTD, those distance restraints are satisfied only in the ATP-lid open state and violated by more than 5 Å in the ATP-lid closed state.

Restraint	Residue i	Pseudoatom Pi	Residue j	Pseudoatom Pj	d_{ij}^{\max} (Å) ^a	d_{ij}^{viol} (Å) ^b
1	ALA24	HB*	MET119	HE*	8	11.87
2	ILE26	HD1*	THR115	HG2*	8	5.54
3	ILE26	HD1*	MET119	HE*	8	8.59
4	ILE26 ^c	HD1*	ILE131 ^c	HD1*	8	f
5	LEU29	HD2*	THR115	HG2*	8	6.63
6	LEU29	HD2*	MET119	HE*	8	6.16
7	LEU29	HD2*	LEU122	HD2*	8	8.89
8	LEU32	HD2*	MET119	HE*	8	4.52
9	LEU32	HD2*	LEU122	HD2*	8	6.24
10	LEU32 ^c	HD2*	MET130 ^c	HE*	8	f
11	LEU32	HD2*	ILE131	HD1*	8	9.65
12	THR36	HG2*	LEU122	HD2*	8	8.53
13	THR36	HG2*	ILE131	HD1*	8	11.65
14	ILE43	HD1*	LEU122	HD2*	8	9.42
15	ILE43 ^d	HD1*	MET130 ^d	HE*	8	f
16	ALA111	HB*	MET119	HE*	8	10.55
17	THR115	HG2*	VAL136	HG2*	8	6.64
18	MET119	HE*	VAL136	HG2*	8	8.76
19	LEU122	HD2*	VAL136	HG2*	8	8.31
20	LEU122	HD2*	LEU143	HD2*	8	7.98
21	LEU32	HD2*	THR115	HG2*	8	10.72
22	LEU32	HD2*	ALA126	HB*	8	8.53
23	MET119 ^e	HE*	ILE131 ^e	HD1*	8	f
24	LEU122 ^e	HD2*	ILE131 ^e	HD1*	8	f

^a For each NOE cross-peak observed associated to a pair of residue {i,j}, d_{ij}^{\max} defines the upper limit for the distance restraints applied for structure calculations.

^b In the MD simulations, the violation relative to a given NOE-derived distance restraint differs from zero if the distance between the pseudoatoms Pi and Pj is larger than d_{ij}^{viol} (see more details Section 1.2.2 of Methods).

^c NOEs not observed in the construct HSP90-R46A

^d NOEs not observed in the construct HSP90-R60A

^e NOEs not observed in the construct HSP90-WT

^f Distance not used to calculate violation in the analysis of the MD trajectories.

Supporting Table 3: Summary of NMR refined structure statistics for HSP90-NTD in ATP-lid open and closed states.

	Closed state (R46A-HSP90-NTD)	Open state (R60A-HSP90-NTD)
RMSD to mean structure^a		
<i>Backbone atoms (Å)</i>	1.52 ± 0.28	0.86 ± 0.18
<i>Heavy atoms (Å)</i>	2.08 ± 0.29	1.30 ± 0.21
Distance restraints violations		
<i>Average number of violations (>0.4Å)</i>	0	0
<i>Highest violation (Å)</i>	0.37	0.24
Ramachandran analysis^b		
<i>Most favored (%)</i>	87.8	90.0
<i>Additionally allowed region (%)</i>	10.7	9.0
<i>Generously allowed region (%)</i>	1.3	0.9
<i>Disallowed region (%)</i>	0.2	0.1

^aRMSD to the mean structure for the ATP-lid (from the residue 98 to 136) calculated for the 20 structures with the lower CYANA target functions refined using CNS restrained molecular dynamics.

^bRamachandran statistics calculated with PROCHECK for both states of HSP90-NTD.

Supporting Table 4: Identification of characteristic distances of dimeric forms of HSP90 α -NTD.

Close by pairs of methyl groups representative of a possible dimerization of HSP90-NTD with the ATP-lid closed state and intermolecular exchange of the two β_1 strands are shown column 1. Such pair of methyl groups correspond to inter-chain putative NOEs involving one methyl group from residue 13 to 41 and any other methyl group of the NTD for which the measured distance was ≤ 11 Å in the full-length dimer HSP90 α in closed form (PDB: 7L7J) but > 11 Å for both the solution structure ensemble representative of the ground/open and excited/closed ATP-lid states (Fig. 3). In the second column, presence/absence of putative specific NOEs of HSP90 α -NTD dimer have been investigated. Their presence/absence has been verified in the 3D $^{13}\text{C}_3$ -edited NOESY spectrum acquired using the methyl-labeled R46A-HSP90 α -NTD sample.

Contacts ($d_{\text{HH}} \leq 11$ Å) specific of dimeric full-length HSP90α	Present in NOESY spectra of R46A-HSP90-NTD ?
17 - 149	No
17 - 174	No
17 - 176	No
21 - 26	No
21 - 110	No
24 - 29	No
24 - 30	No
24 - 33	No
24 - 110	No
24 - 136	No
29 - 109	No
32 - 111	No
33 - 24	No
33 - 111	No
36 - 111	No
36 - 136	No

Supporting Table 5: NMR-based structural calculations supporting the absence of helix-5 in ATP-lid closed state.

All CYANA structural calculations performed without angular restraints for helix-5, but including all experimental NOE-derived distance restraints have a target function below 5 kJ.mol⁻¹ and no violation of NMR experimental restraints. In contrast, when dihedral restraints are introduced to force the formation of helix-5, the target function increases above 6 kJ.mol⁻¹ and 3 to 4 violations of NMR structural restraints are observed.

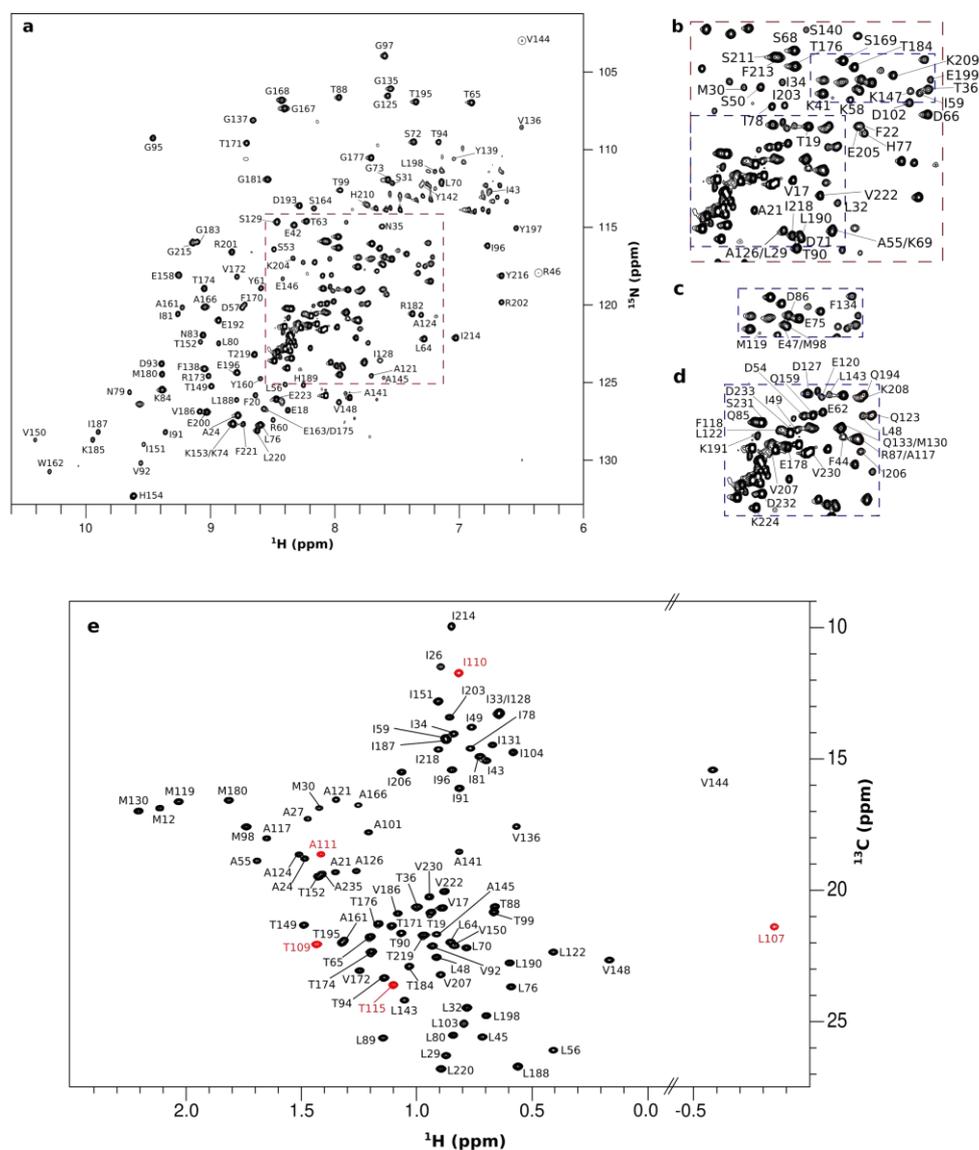
Helices present in the structural calculation	Energy (kJ.mol ⁻¹)	Distance restraint violation	Angular restraint violation
None	4.51 ± 0.04	∅	∅
α3	4.54 ± 0.05	∅	∅
α4	4.60 ± 0.04	∅	∅
α5	6.13 ± 0.06	2 ^a	1 ^b
α3α4	4.56 ± 0.02	∅	∅
α3α5	6.13 ± 0.08	2 ^a	1 ^b
α4α5	6.23 ± 0.09	2 ^a	2 ^c
α3α4α5	6.14 ± 0.07	2 ^a	2 ^c

^a Violated distance restraints: THR65 HG2* - MET130 HE* and ILE131 HD1* - LEU143 HD2*

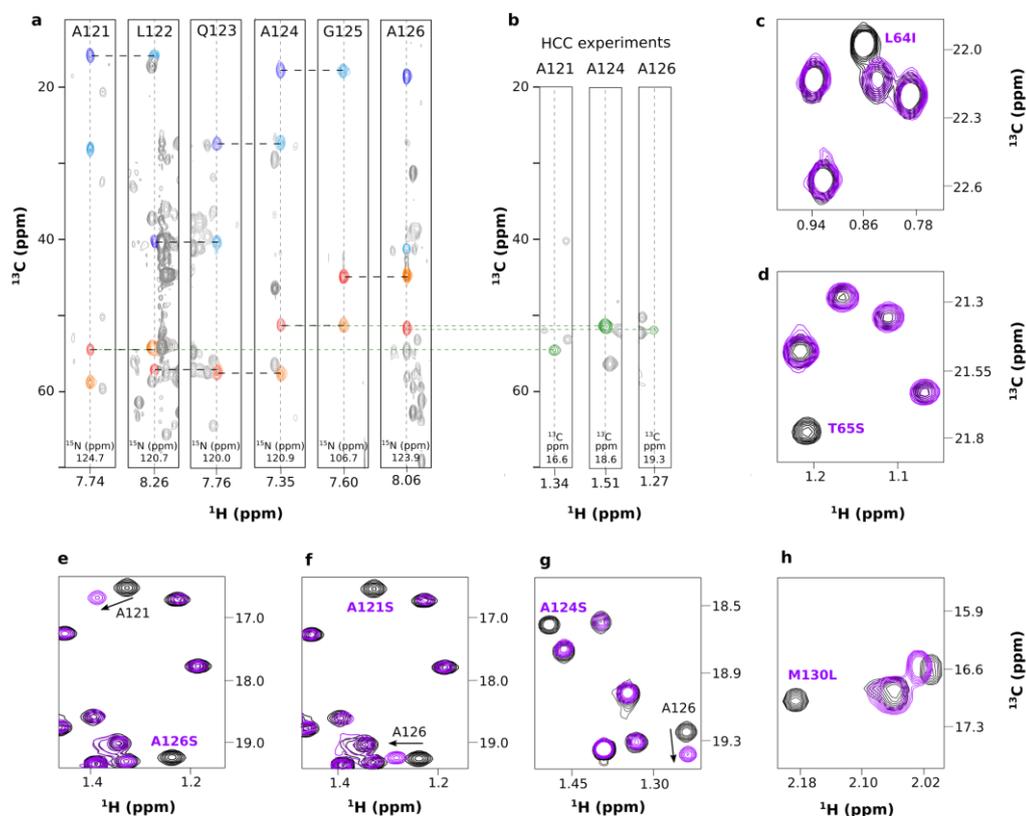
^b Violated angular restraints: PHI SER211

^c Violated angular restraints: PHI SER211 and CHI1 ILE214

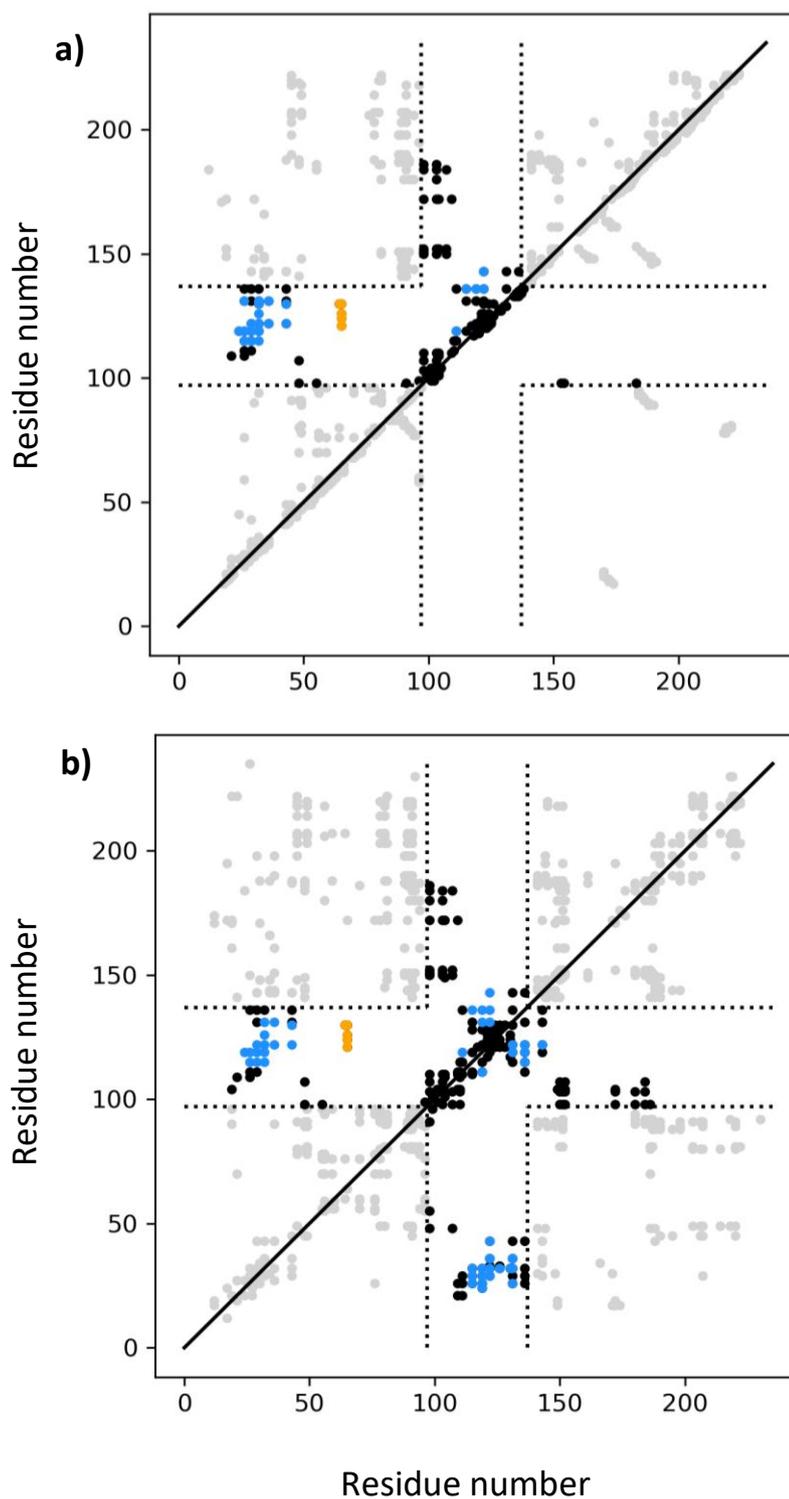
∅ : no violation



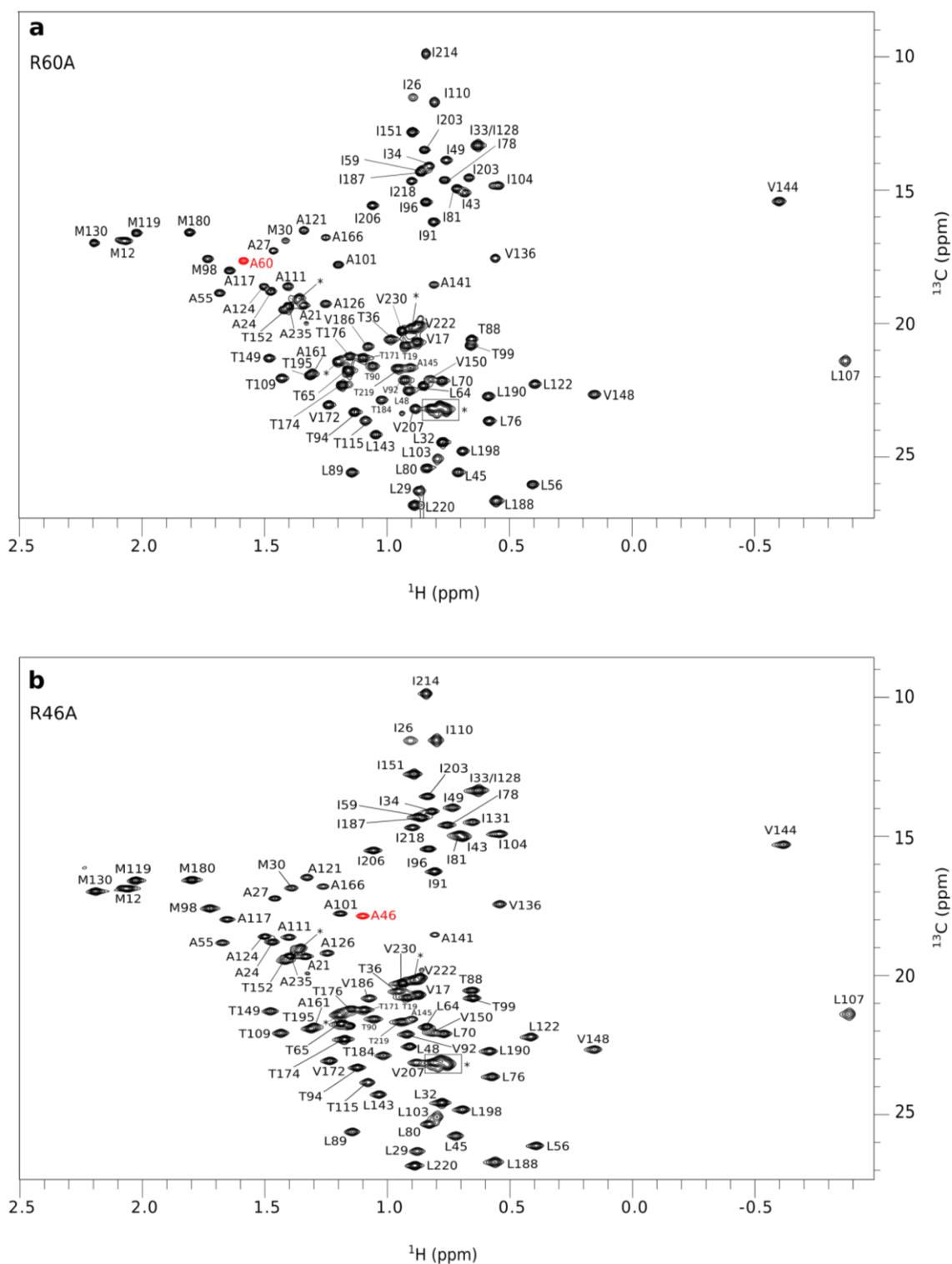
Supporting Figure 1: Assigned 2D NMR spectra of HSP90 α -NTD. **a)** ^{15}N -BEST-TROSY spectrum of HSP90 α -NTD in its apo state, acquired at 293 K and at a ^1H frequency of 600 MHz. Panels **b-d**) present zooms of 2D ^{15}N -BEST-TROSY spectrum. Each assigned signal (a, b, c, d) is annotated with the corresponding residue number³⁶⁻³⁸. Panel **e**) displays the fully assigned 2D ^1H - ^{13}C SOFAST methyl TROSY spectrum of HSP90 α -NTD apo, recorded at 298 K. Each signal is annotated with the corresponding residue number³⁸. In red are displayed residues belonging to the stretch [105-116] invisible on ^{15}N -BEST-TROSY due to conformational exchange in the μs -ms timescale.



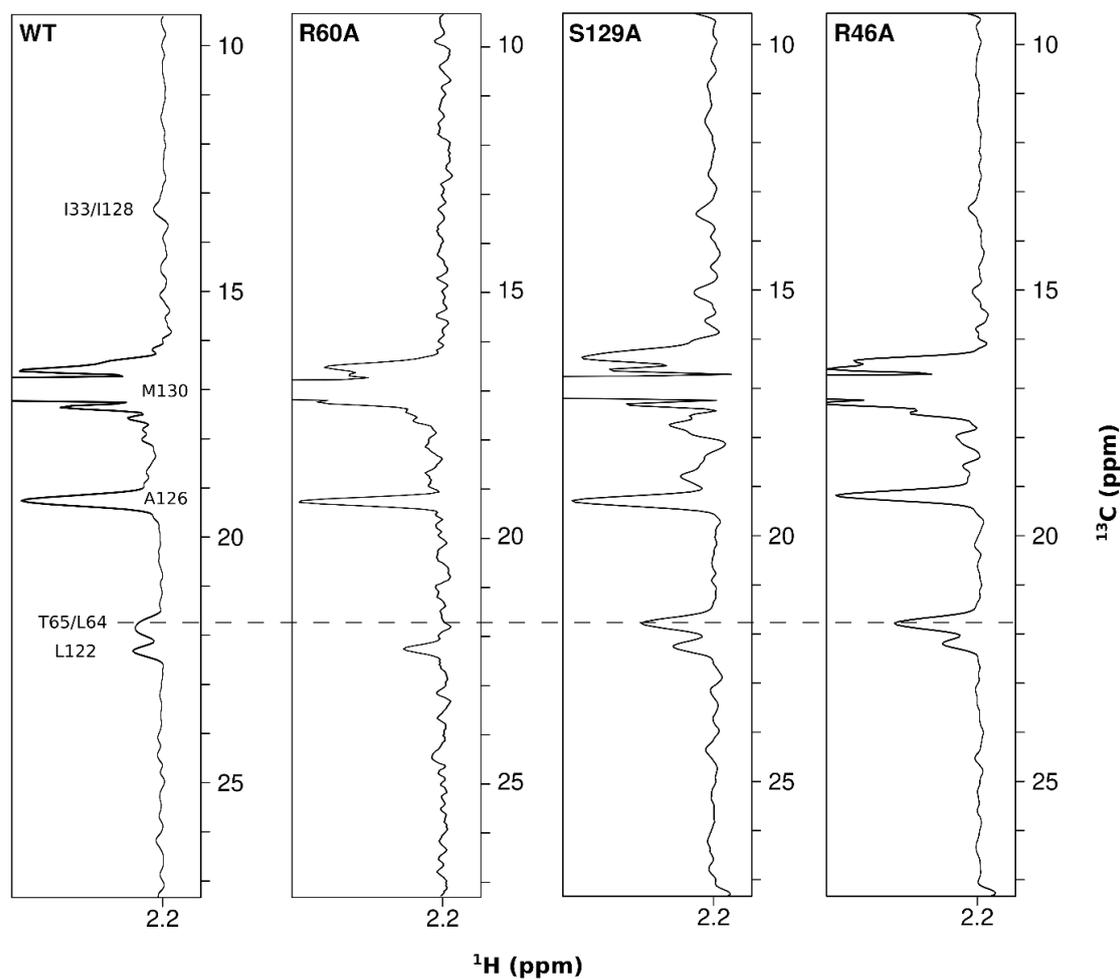
Supporting Figure 2: HSP90 α -NTD assignment. **a)** Examples of 2D-extracts from 3D HNCA, HN(CO)CA, HNCACB, HN(CO)CACB showing backbone correlations from A121 to A126. Spectra were acquired at 298 K on a NMR spectrometer operating at a ^1H frequency of 600 MHz using a U- $[\text{2H}$, ^{13}C , $^{15}\text{N}]$ -HSP90 α -NTD sample. $\text{C}\alpha_{(i)}$, $\text{C}\alpha_{(i-1)}$, $\text{C}\beta_{(i)}$, $\text{C}\beta_{(i-1)}$, are colored in red, orange, dark blue and light blue, respectively. **b)** Examples of 2D-extracts from ‘out and back’ HCC experiment to transfer assignment from backbone atoms to the methyl groups. **c-h)** Cross-validation of assignment of methyl groups using mutagenesis. Examples of 2D ^1H - ^{13}C SOFAST methyl TROSY spectra of 6 HSP90 α -NTD mutants. Each mutant spectrum extract (purple) is superimposed with the wild type protein (black). A total of 56 single point amino acid mutations were generated by GeneCust and for each of these samples a single type of methyl group was labeled by addition of the corresponding NMR-Bio kit (SLAM-A $^{\beta}$, SLAM-M $^{\epsilon}$, SLAM-I $^{\delta 1}$, SLAM-T $^{\gamma}$, SLAM-V $^{\text{proS}}$ or DLAM-LV $^{\text{proS}}$) in M9/ $^2\text{H}_2\text{O}$ media. The following mutants were produced and analyzed by NMR: M12L, M30L M119L, M130L, M180L, A21S, A24S, A27S, A55S, A101S, A111S, A117S, A121S, A124S, A126S, A161S, A166S, A235S, T19S, T36S, T65S, T99S, T109S, T115S, T152S, T171S, T174S, T184S, T195S, T219S, I26V, I33V, I34V, I81V, I104V, I110V, I128V, I131V, I214V, V17I, V136I, V150I, V172I, V186I, V222I, V230I, L29I, L32I, L45I, L48I, L56I, L64I, L70I, L107A, L122I, L143I).



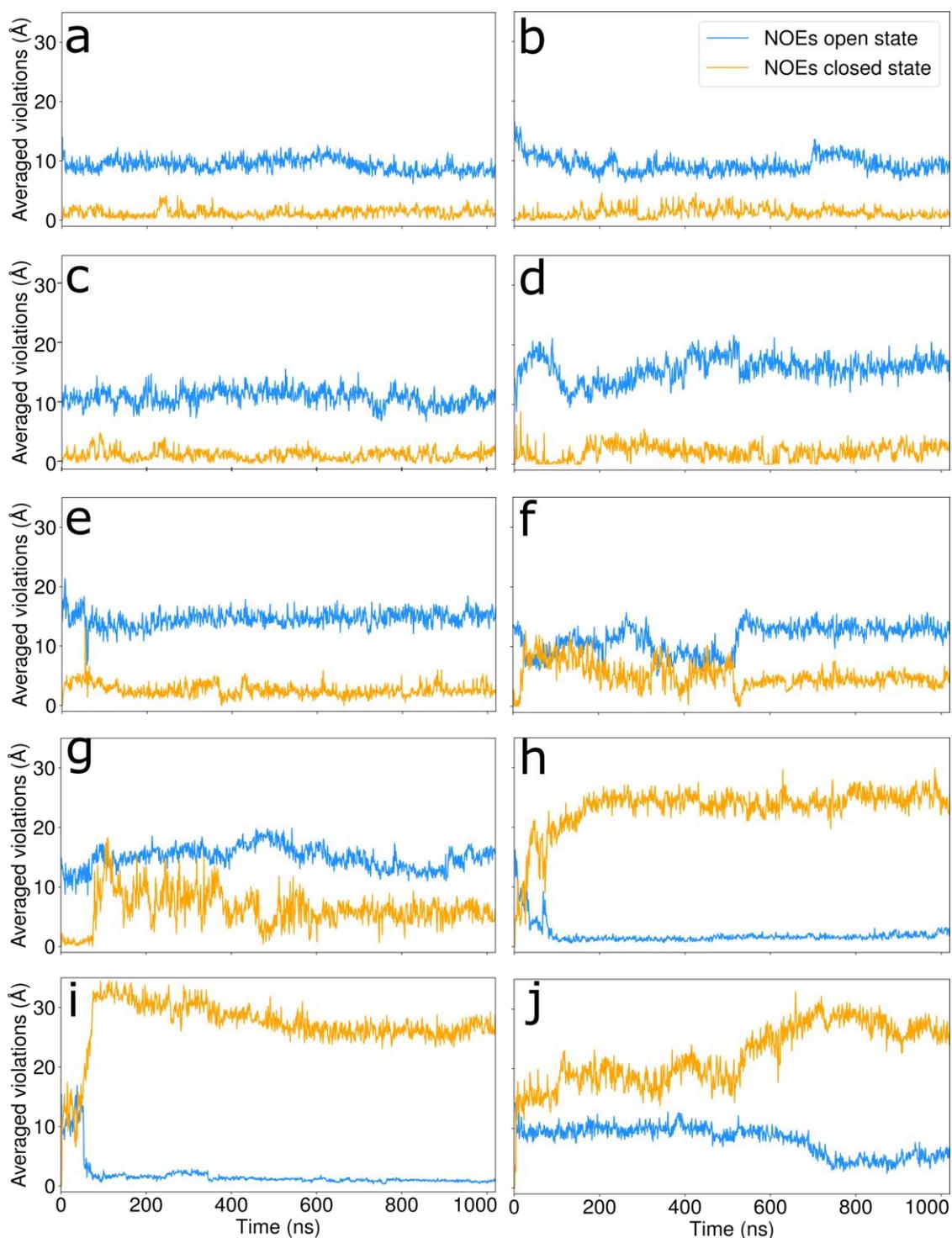
Supporting Figure 3: Summary of NOEs distance restraints. a) Experimentally detected CH₃-CH₃ (top left) and H_N-H_N (bottom right) NOEs in WT-HSP90α-NTD. b) Experimentally detected CH₃-CH₃ NOEs for R46A- (top left) and R60A- (bottom right) WT-HSP90α-NTD. Black points represent NOEs of the ATP-lid present in both states. Blue/orange points represent NOEs specific of ATP-lid in open/closed states, respectively (Supporting Table 2).



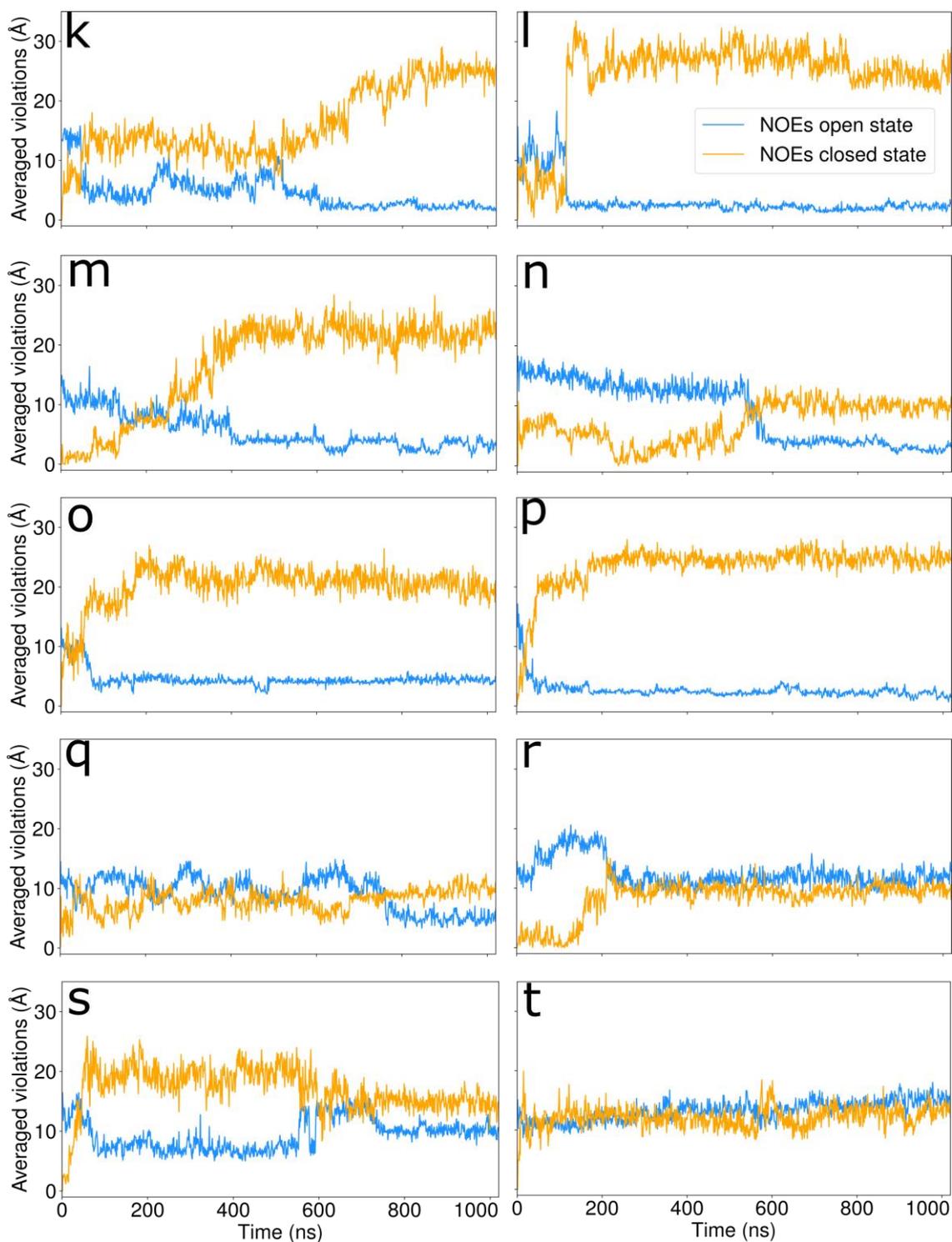
Supporting Figure 4: Assigned 2D ^1H - ^{13}C SOFAST methyl TROSY of HSP90 α -NTD mutants. a) R60A- and **(b)** R46A-HSP90 α -NTD mutants, recorded at 293 K. Each signal is annotated with the corresponding residue number. Peaks annotated with the symbol * are peaks belonging to the uncleaved N-terminal tag. Extra alanine methyl is colored in red.



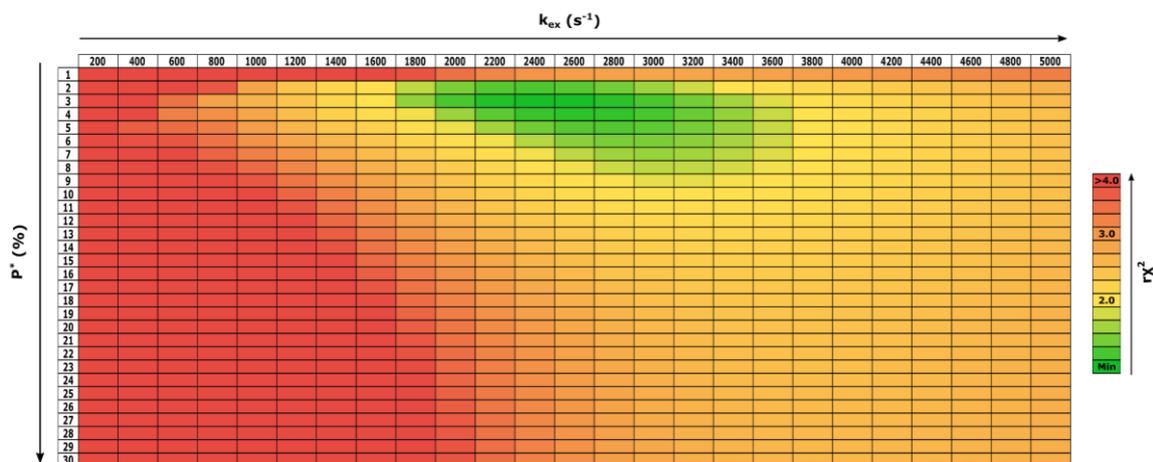
Supporting Figure 5: Intensity variation of characteristic NOEs. 1D traces extracted from 3D CCH HMQC-NOESY-HMQC experiments showing M130 as diagonal peak and associated NOEs cross peaks for the wild type protein HSP90 α -NTD and three mutants: R60A, S129A and R46A. 3D ^{13}C -edited NOESY spectra were acquired during *ca.* 2 days each using high-field NMR spectrometers operating at 950 or 850 MHz and at a temperature of 25°C (WT), 20°C (R60A) or 15°C (R46A), taking into account the lower stability of mutants. NOE mixing time were optimized to delays ranging from 0.3 (15°C) to 0.4 s (25°C), corresponding to the estimated maximum intensities of NOEs buildups for each experimental condition.



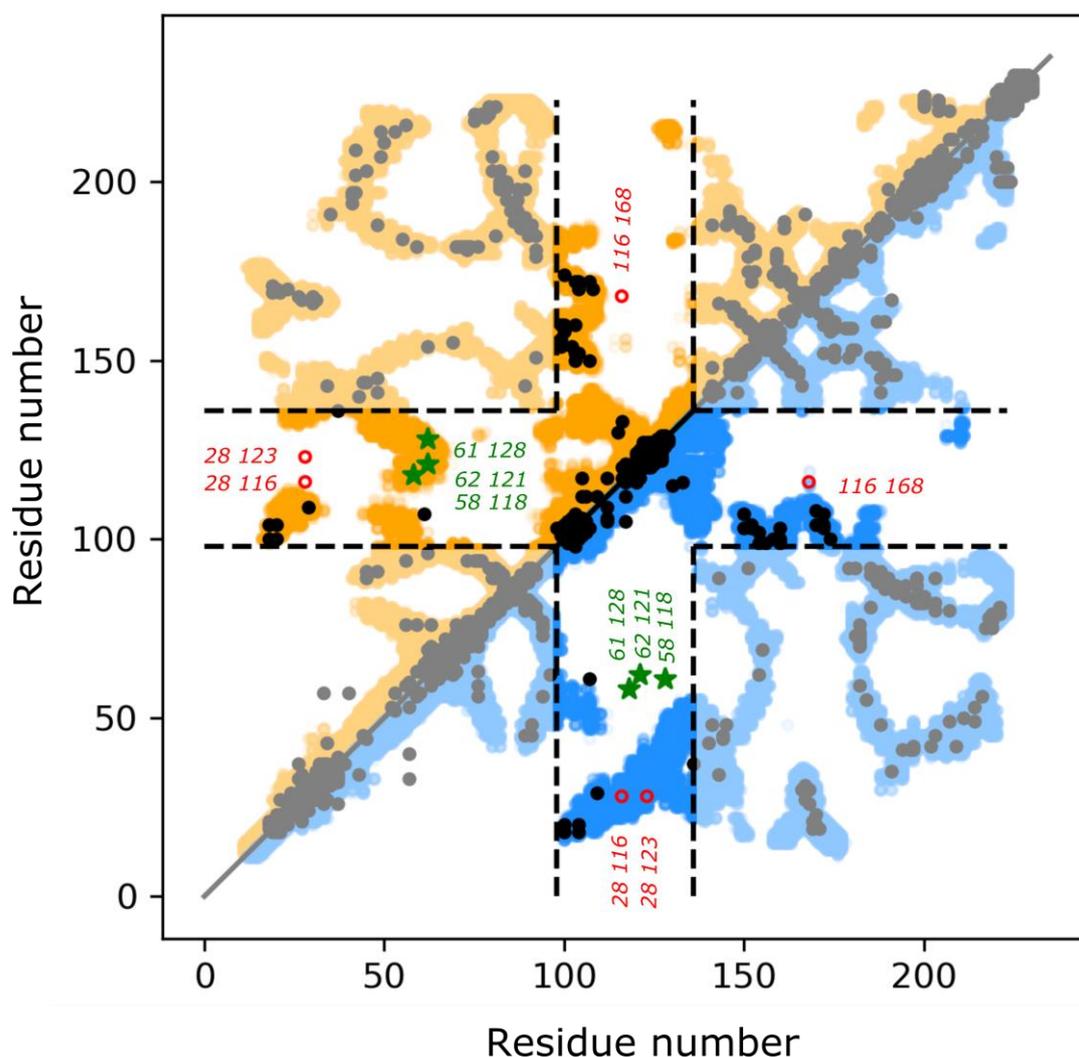
Supporting Figure 6a. Time evolution of the violations of the NOE-derived distance restraints specific of either the open or the closed conformation (orange and blue respectively), for the 20 non-restrained molecular dynamics simulations starting from ATP-lid closed state conformers. The violations are defined in section 1.2.2 of Methods, and the simulations were classified according to the violations. **a-c:** relatively stable trajectories, compatible with closed-state-NOE distances during several hundreds of nanoseconds. **d-g:** relatively stable trajectories, remaining nearby the closed state during several hundreds of nanoseconds. **h-j:** unstable trajectories with transition toward a conformation close to the open-state.



Supporting Figure 6b: same as Supporting Figure S6a. **k-p:** unstable trajectories with transition toward a conformation close to the open-state. **q-t:** unstable trajectories with transition toward a conformation neither closes nor open-state.

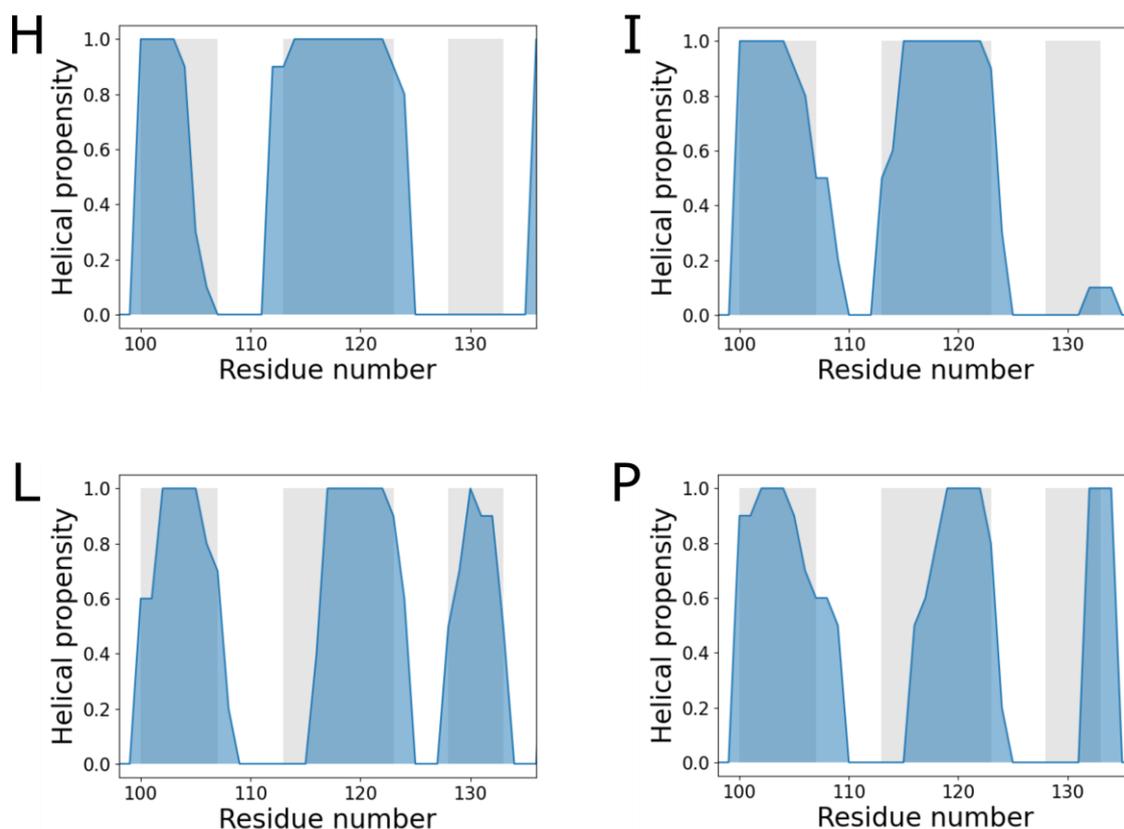


Supporting Figure 7: Analysis of HSP90-NTD dynamics using relaxation dispersion experiments. Grid search table displaying the reduced χ^2 of the fit obtained from ^{13}C -relaxation data acquired at 293 K, when exchange rate and population of the excited state were fixed. Going from green to red the reduced χ^2 values increase.



Supporting Figure 8: Computed Evolutionary Couplings (ECs) for HSP90-NTD.

Contact map of top-ranked predicted ECs (black dots) overlaid on HSP90-NTD contact maps of ground/open (blue) and excited/closed (orange) ATP-lid states. 453 highly probable ECs ($p > 0.85$) predicted by EVcouplings (<https://v2.evcouplings.org>) were computed⁴⁷ starting from an alignment of 4414 homologue sequences of HSP90-NTD coming from the data base UniRef90. 71 ECs involving at least one residue of the HSP90 protein segment [98-136] are highlighted in the middle of the matrix. The contact maps have been plotted using both ground/open and excited/closed structures ensembles (Fig. 3). A colored dot was plotted for each C_{α} - C_{α} distance ≤ 15 Å extracted from the 20 best structures obtained for both states. The red circles (green stars) correspond to ECs only satisfied in ground/open (respectively excited/closed) ATP-lid state.



Supporting Figure 9: Probability of helical structures along the sequence of the ATP lid, in 4 MD simulations depicting a transition of the ATP-lid from a closed to an open state. The helical propensity are shown in blue. For comparison, the helical regions of the ground state are represented in gray (helices 3, 4 and 5). Each panel is annotated with the letter corresponding to panel of Supporting Fig. 6, presenting the 20 different MD simulations starting from closed-state conformers. These four trajectories were selected on the basis of their violations of the specific NOE of the open state, which were lower than *ca.* 3.5 Å. On average over these four selected simulations, the helical propensity of the helix-5 (residues [128-134]) is $28 \pm 18\%$. In comparison, on average over the 20 simulations on the ground state, the helical propensity of the helix-5 is much higher: $63 \pm 5\%$ on the last 120 ns. Therefore, the formation of the helix-5 is not systematically observed after a swap of the ATP-lid from the closed to the open state.