



**HAL**  
open science

**Spectrométrie de masse supramoléculaire :  
caractérisation de l'interaction non-covalente entre  
PEBP1/RKIP humaine et des analogues de nucléotides**

Lucie Jaquillard

► **To cite this version:**

Lucie Jaquillard. Spectrométrie de masse supramoléculaire : caractérisation de l'interaction non-covalente entre PEBP1/RKIP humaine et des analogues de nucléotides. Autre. Université d'Orléans, 2012. Français. NNT : 2012ORLE2020 . tel-00923153

**HAL Id: tel-00923153**

**<https://theses.hal.science/tel-00923153>**

Submitted on 2 Jan 2014

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

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



**UNIVERSITÉ D'ORLÉANS**



**ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES**

**CENTRE DE BIOPHYSIQUE MOLECULAIRE**

**THÈSE** présentée par :  
**Lucie JAQUILLARD**

soutenue le : **20 mars 2012**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline : Chimie analytique

**Spectrométrie de masse supramoléculaire:  
Caractérisation de l'interaction non-covalente  
entre PEBP1/RKIP humaine  
et des analogues de nucléotides**

**THÈSE dirigée par :**

**Martine CADENE**

Chargée de Recherche, CNRS Orléans

**RAPPORTEURS**

**Guillaume VAN DER REST**

Professeur, Université Paris Sud 11

**Olivier LAPREVOTE**

Professeur PU-PH, Université Paris 5

**JURY :**

**Luigi AGROFOGLIO**

Professeur, Université d'Orléans,  
Président du jury

**Françoise SCHOENTGEN**

Directeur de Recherche, Université Paris 6

**Pascale REVERDIAU**

Professeur, Université de Tours

**Guillaume VAN DER REST**

Professeur, Université Paris Sud 11

**Olivier LAPREVOTE**

Professeur PU-PH, Université Paris 5

**Martine CADENE**

Chargée de Recherche, CNRS Orléans

# Remerciements

*Le présent travail a été réalisé au sein du Centre de Biophysique Moléculaire à Orléans (CBM, UPR 4301 CNRS) sous la direction du Docteur Martine Cadène.*

*Je tiens à exprimer toute ma gratitude au Docteur Martine Cadène pour la confiance dont elle a fait preuve à mon égard en me proposant ce sujet de recherche. Je lui adresse également mes sincères remerciements pour sa disponibilité, sa rigueur scientifique, son enthousiasme et ses encouragements, que j'ai eu la chance de recevoir tout au long de ma thèse.*

*Je tiens à exprimer mes remerciements aux Professeurs Guillaume van der Rest et Olivier Laprévotte pour avoir accepté d'être rapporteurs de cette thèse et avoir assuré le travail d'évaluation qui en a découlé.*

*Mes remerciements s'adressent également au Professeur Luigi Agrofoglio, au Directeur de Recherche Françoise Schoentgen et au Professeur Pascale Reverdiau qui m'ont fait l'honneur de participer au jugement de ces travaux en tant qu'examineurs.*

*Je tiens à remercier tous les membres de mon équipe, Martine Beaufour, Guillaume Gabant, Kevin Mark, Emmanuelle Mebold, Cyril Colas et David Ginguené, pour leurs précieux conseils en spectrométrie de masse et leur gentillesse incontestable.*

*Je tiens à remercier tout particulièrement les personnes du projet METASUPP avec lesquelles nous avons travaillé en collaboration et sans qui ce travail n'aurait pas été possible :*

*Françoise Schoentgen et Ludivine Errien pour la production de la protéine, Françoise Schoentgen encore pour ses précieuses informations et idées sur la PEBP, Fabienne Saab pour la synthèse du peptide P3P, l'équipe de Luigi Agrofoglio, Vincent Roy, Aurélien Montagu, Zhili Zhang et Ozkan Sari, pour la synthèse des analogues de nucléotides, Laurette Tavel, Christian Damblon et Laurence Jouvensal, pour leurs études RMN, Pascale Reverdiau et Sophie Iochmann pour les tests d'inhibition de l'invasion, Christiane Guillouzo, Caroline Gaucher-Di Stasio et Myriam Ravache pour les tests de cytotoxicité et d'inhibition de la migration.*

*Je remercie également tous nos autres collaborateurs, Bertrand Castaing, Karine Loth et Rémy Lemeur pour le projet HU ; Bertarnd Castaing et Arthur Biela pour le projet Fpg ; Francis Gauthier, Brice Korkmaz et Christophe Epinette pour le projet Pr3.*

*Je remercie vivement le CNRS et la Région Centre pour le financement de cette thèse, et le Cancéropôle Grand Ouest, la Ligue contre le Cancer, FEDER et l'ANR pour le financement du projet.*

*Je souhaiterais remercier très chaleureusement tous mes collègues du CBM qui m'ont permis de réaliser ce travail dans une ambiance très sympathique.*

*Je tiens à remercier plus particulièrement,*

*Mon équipage d'aviron « Paquebot mais costauds » (Justo, Chrystelle, Guillaume, Ludivine).*

*Mon coloc de bureau, Guillaume, un inspirateur à tous niveaux et un ami, sincèrement merci.*

*Mon amie de la salsa, Hélène, toujours à l'écoute, remplie d'attentions, sincèrement merci.*

*Je voudrais remercier toute ma famille, tout particulièrement ma mère et mes deux sœurs, pour leur soutien et leur amour, qui m'ont ainsi offert toutes les conditions pour découvrir la recherche et trouver ma voie.*

*Enfin, un immense merci à celui qui a toujours cru en moi, qui est toujours auprès de moi, mon amour, JB.*

# Liste des abréviations

9AA	9 amino-acridine
4HCCA	Acide $\alpha$ -cyano-4-hydroxycinnamique
ABC	Bicarbonate d'ammonium
ACTH	acétylcholine
ADN	Acide désoxyribonucléique
ANP	nucléoside acyclique phosphonate
ARN	Acide ribonucléique
AMP	Adénosine monophosphate
AMPC	AMP cyclique
ADP	Adénosine diphosphate
ATP	Adénosine triphosphate
CE	Cap Exit : voltage à la sortie du capillaire de transfert
CID	Collision-induced dissociation
Da	Dalton (unité de mesure de masse moléculaire, équivalente au gramme/mole)
DC	Direct current (courant continu)
DMSO	diméthylsulfoxyde
DTT	dithiothréitol
Ec	énergie cinétique
Ep	énergie potentielle
ESI	Ionisation électrospray
ETD	electron transfer dissociation
Fpg	Formamidopyrimidine-DNA glycosylase
FMN	Flavine Mononucléotide
FWHM	full width at half maximum
FWI	Formic acid/Water/Isopropanol (3/1/2)
GPD	dissociation en phase gazeuse
GSH	glutathion

GTP	Guanosine triphosphate
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)ethanesulfonic Acid
HEWL	Hen Egg-White Lysozyme
IC50	concentration inhibitrice médiane
ISCID	In source CID- tension appliquée sur le maXis
IT	Trappe ionique
ITC	isothermal titration calorimetry
$k_{off}$	constante de dissociation cinétique
$k_{on}$	constante d'association cinétique
$K_D$	constante de dissociation à l'équilibre
MALDI	Désorption/ionisation laser assistée par matrice
MAP	Mitogen-activated protein
MES	Acide 2-(N-Morpholino)-éthanesulfonique
MS	Spectrométrie de masse
m/z	ratio masse sur charge
NADP	nicotinamide adenine dinucleotide phosphate
NAG3	<i>N-acetyl</i> glucosamine
NF- $\kappa$ B	Nuclear factor kappa B
NH <sub>4</sub> OAc	Acétate d'ammonium
NP-40	Tergitol-type NP-40 (nonyl phenoxyethoxyethanol)
PDE-5	Phosphodiesterase-5
PEBP	Phosphatidylethanolamine Binding Protein
Pi	Pression dans l'interface
PR3	Protéase à sérine 3
Q	Quadrupôle
RF	Radiofréquence
RKIP	Raf Kinase Inhibitory Protein
RMN	résonance magnétique nucléaire
SPR	Résonance des plasmons de surface

TCEP	Tris(2-carboxyethyl)phosphine
TFA	acide trifluoroacétique
TRIS	trihydroxyméthylaminométhane
TOF	Temps de vol
TWA	TFA, Water/Acetonitrile (0.1%, 66/34 v/v)
SK	Skimmer- tension appliquée sur ESI-IT
Vc	Tension d'accélération
V8	endoprotéinase-Glu-C
WAFA	Water/Acetonitrile, Formic Acid (50/50 v/v, 0.2%)

## Nomenclature des acides aminés :

A	Alanine	(Ala)
D	Acide aspartique	(Asp)
E	Acide glutamique	(Glu)
F	Phénylalanine	(Phe)
G	Glycine	(Gly)
H	Histidine	(His)
I	Isoleucine	(Ile)
K	Lysine	(Lys)
L	Leucine	(Leu)
P	Proline	(Pro)
R	Arginine	(Arg)
S	Sérine	(Ser)
T	Thréonine	(Thr)
V	Valine	(Val)
W	Tryptophane	(Trp)
Y	Tyrosine	(Tyr)





# Plan détaillé

<b>INTRODUCTION GENERALE .....</b>	<b>11</b>
<b>INTRODUCTION.....</b>	<b>15</b>
SECTION 1 : PEBP/RKIP, UNE PROTEINE SUPPRESSEUR DE METASTASES .....	16
I. Localisation et caractéristiques physico-chimiques .....	16
1. Localisation.....	16
2. Caractéristiques physico-chimiques.....	17
II. Structure .....	18
1. Structure primaire .....	18
1.1. Famille PEBP .....	18
1.2. PEBP/RKIP humaine .....	22
2. Structure secondaire.....	23
3. Structure tertiaire .....	24
3.1. hPEBP1 .....	24
3.2. Comparaison avec d'autres PEBPs .....	25
4. Alignements de structure .....	27
III. Ses fonctions .....	28
1. Inhibition de protéases .....	28
2. Inhibition de kinases .....	29
2.1. Inhibition de la voie des MAP kinases Raf/MEK/ERK .....	29
2.2. Inhibition de la voie NF- $\kappa$ B .....	31
2.3. Inhibition de la voie des récepteurs couplés aux protéines G.....	32
2.4. Régulation de l'activité de PEBP1 par phosphorylation .....	33
3. Au niveau de l'organisme .....	34
3.1. Suppresseur de métastases .....	34
3.2. Rôle dans la maladie d'Alzheimer .....	37
3.3. Rôle dans l'apoptose et la sensibilité aux chimiothérapies .....	37
3.4. Autres fonctions .....	38
IV. Interactions PEBP1 – ligand .....	39
1. Interaction covalente avec la locostatine.....	39
1.1. La locostatine .....	39
1.2. Protéine cible : PEBP1 .....	40
1.3. Mécanisme de liaison.....	41
2. Interaction non-covalente avec des nucléotides.....	41
2.1. Structure des nucléotides.....	41
2.2. Interactions avec PEBP1 .....	42
2.3. Analogues de nucléotides dans les traitements contre le cancer .....	45
3. Interactions protéine-protéine .....	49
4. Interactions avec des lipides membranaires ou des molécules hydrophobes .....	50
5. Interaction avec des molécules tampon.....	53
6. Autres interactions .....	54
6.1. Avec un inhibiteur de phosphodiesterase-5 (PDE5).....	54
6.2. Avec des dérivés morphiniques.....	54
6.3. Avec des composés portant un groupement 4-nitrophényl.....	55
7. Structure connue d'interaction avec des ligands : la poche à anion .....	56
7.1. Description de la poche .....	56
7.2. Rôles physiologiques de la poche chez les plantes et la levure.....	57
7.3. Plasticité conformationnelle et rôles possibles.....	57
SECTION 2 : LA SPECTROMETRIE DE MASSE SUPRAMOLECULAIRE .....	59
I. Les interactions non-covalentes impliquées dans les complexes supramoléculaires .....	59
1. Interactions ioniques .....	60
2. Forces de Van Der Waals .....	60

2.1.	Forces de Keesom .....	61
2.2.	Forces de Debye .....	62
2.3.	Forces de London .....	62
2.4.	Cas particulier des forces de van der Waals dans les protéines .....	63
3.	Liaisons hydrogènes .....	64
4.	Interactions résultant de l'effet hydrophobe .....	65
II.	Etude d'interactions non-covalentes par spectrométrie de masse .....	66
1.	Historique .....	66
2.	Spectromètre de masse à ionisation électrospray .....	68
2.1.	Schéma général d'un spectromètre de masse .....	68
2.2.	Principe de l'ionisation électrospray .....	68
2.3.	L'interface .....	72
2.4.	Principe de l'analyse des ions .....	74
3.	Compatibilité des tampons avec l'ESI .....	74
3.1.	Mécanisme de réduction des adduits de sodium sur la protéine avec un tampon acétate d'ammonium: .....	74
3.2.	Mécanisme de charge de la protéine en phase gazeuse avec un tampon ammonium, à pH neutre: .....	75
4.	Informations déduites du spectre de masse .....	76
4.1.	Conformation de la protéine .....	76
4.2.	Spécificité de l'interaction observée .....	76
4.3.	Stœchiométrie apparente du complexe .....	77
4.4.	Constantes cinétiques du complexe .....	77
5.	Reflète de l'équilibre en solution lors de l'analyse en phase gazeuse .....	79
5.1.	Conservation de l'équilibre chimique en phase gazeuse .....	79
5.2.	La conservation des interactions non-covalentes dépend du type d'interactions .....	80
III.	Cas particulier d'interactions non-covalentes: les systèmes à faible affinité .....	82
1.	Limites de l'étude par spectrométrie de masse de systèmes à faible affinité .....	82
1.1.	Problème de l'agrégation en phase gazeuse .....	82
1.2.	Problème de dissociation en phase gazeuse (GPD) .....	83
1.3.	Superposition des deux problèmes .....	84
2.	Stratégies développées pour pallier les limites des systèmes à faible affinité .....	85
2.1.	Stratégies pour s'affranchir de l'agrégation en phase gazeuse .....	85
2.2.	Stratégies pour s'affranchir de la dissociation en phase gazeuse .....	87
SECTION 3 : PRESENTATION DU TRAVAIL DE THESE .....		90
I.	Contexte : projet METASUPP .....	90
1.	Objectifs du projet .....	90
2.	Organisation .....	91
II.	Objectifs du travail de thèse .....	92
1.	Approfondir la relation structure-fonction de PEBP1 humaine .....	92
2.	Fournir des éléments pour la synthèse raisonnée d'activateurs des propriétés antimétastatiques naturelles de hPEBP1 .....	93
3.	Développer une méthode de suivi MS des complexes reflétant le comportement en solution ...	93
4.	Retombées potentielles .....	93
<b>RESULTATS ET DISCUSSION .....</b>		<b>95</b>
CHAPITRE I : OPTIMISATIONS EXPERIMENTALES, METHODE D'INJECTION ET PARAMETRES D'ANALYSE .....		97
I.	Mise au point d'une méthode d'injection directe « en sandwich » .....	99
1.	Choix du type d'infusion directe .....	99
2.	Débit d'infusion .....	100
3.	Méthode d'infusion directe ESI originale .....	101
4.	Effet du solvant du sandwich .....	102
4.1.	Sur hPEBP1/RKIP seule .....	102
4.2.	Sur le complexe hPEBP1 - BLA-C-04 .....	105
II.	Optimisation des paramètres instrumentaux .....	107
1.	Application à HU .....	108
1.1.	Optimisation de la tension d'accélération .....	108
1.2.	Optimisation de la température de désolvatation .....	112
2.	Application à hPEBP1 – GTP .....	113

CHAPITRE II: APPLICATION A DES SYSTEMES NON-COVALENTS .....	117
I. Etudes qualitatives.....	119
1. PR3-azapeptide .....	119
1.1. Contexte général.....	119
1.2. Etude du complexe.....	121
2. Fpg-Zn / NEIL2-Zn .....	125
2.1. Contexte .....	125
2.2. Objectifs.....	127
2.3. Optimisations pour les complexes protéine/métal.....	127
2.4. Confirmation d'édifices non-covalents protéine-zinc .....	128
2.5. Interaction avec la drogue 2-thioxanthine .....	132
2.6. Bilan .....	135
II. Etude cinétique : Cas de la formation de l'hétérodimère $\alpha\beta$ de la protéine HU .....	136
1. Contexte .....	136
2. Objectif .....	136
3. Résultats.....	137
3.1. Optimisation des conditions expérimentales pour observer les dimères HU .....	137
3.2. Formation d'hétérodimères à partir des deux homodimères .....	141
3.3. Bilan .....	147
 CHAPITRE III : DEVELOPPEMENT D'UNE NOUVELLE METHODOLOGIE POUR DETERMINER DES CONSTANTES D'EQUILIBRE DE DISSOCIATION DES COMPLEXES A FAIBLE AFFINITE .....	149
Contexte.....	151
Publication: <i>Improved accuracy of low affinity protein-ligand equilibrium dissociation constants directly determined by electrospray ionization mass spectrometry</i> .....	153
Supplementary materials.....	171
Annexes à la publication: .....	185
I. Validation du modèle .....	185
1. La myokinase ou adénylate kinase (AK) en complexe avec deux molécules d'ADP. ....	185
2. La concanavaleine A avec le méthyl- $\alpha$ -D-mannopyranoside: .....	188
II. Effet de la force ionique sur l'agrégation .....	188
III. Effet de la tension d'accélération sur $K_D$ et $f_{sat}$ .....	189
IV. Recommandation .....	191
 CHAPITRE IV: CARACTERISATION DU SITE NUCLEOTIDIQUE DE HPEBP1/RKIP .....	193
I. Recherche de ligands nucléotidiques par criblage.....	195
1. Principe de la synthèse raisonnée d'analogues nucléotidiques .....	195
2. Protocole de criblage.....	197
3. Criblage rapide avec les fractions de liaison.....	198
3.1. Analogues d'UMP.....	198
3.2. Nucléotides de référence .....	201
3.3. Variation de la longueur de chaîne de BLA-C-04.....	207
3.4. Deuxième série d'analogues d'UMP .....	208
3.5. Analogues de FMN .....	211
3.6. Analogues sans uracile .....	213
3.7. Analogues de GTP .....	214
3.8. Molécules du criblage de Shemon.....	216
4. Détermination de $K_D$ pour les meilleurs ligands .....	218
4.1. Analogues d'UMP .....	219
4.2. Dérivés de FMN .....	222
4.3. Dérivés de GTP .....	223
4.4. Classement général des $K_D$ .....	224
5. Bilan du criblage.....	225
6. Relation entre l'hydrophobicité des meilleurs ligands et les paramètres de liaison .....	228
II. Localisation du site de liaison nucléotidique.....	231
1. Indépendance des sites de liaison de nucléotide et de la locostatine.....	231
2. Parallèle avec la RMN : site nucléotidique et poche à anion confondues .....	232
3. Effet de la force ionique.....	234
III. Comparaison des liaisons moléculaires avec les effets cellulaires.....	239
1. Effet sur la migration cellulaire .....	239

2.	Effet sur l'invasion cellulaire.....	243
3.	Toxicité.....	247
4.	Comparaison et corrélation.....	248
4.1.	Des différents effets cellulaires.....	248
4.2.	Des résultats cellulaires et de MS.....	248
4.3.	Bilan.....	251
<b>CHAPITRE V : RECHERCHE D'UNE ACTIVITE ENZYMATIQUE DE hPEBP1/RKIP.....</b>		<b>253</b>
I.	Activité oxydase.....	255
1.	Observation de l'oxydation de hPEBP1 en présence de FMN.....	255
2.	Hypothèse sur le mécanisme d'oxygénation.....	257
2.1.	Monooxygénase à flavine.....	257
2.2.	Oxydation par échange direct avec une molécule substrat.....	258
3.	Carte peptidique de l'oxydation de hPEBP1 en présence de FMN.....	258
4.	Spécificité de la réaction d'oxygénation en présence de FMN.....	260
5.	Photo-oxydation en présence de FMN.....	262
6.	Absence d'oxydation avec les analogues de FMN.....	263
II.	Activité phosphodiesterase-5.....	265
1.	Contexte.....	265
2.	Résultats et discussion.....	266
III.	Activité GTPase ou ATPase.....	269
<b>MATERIELS ET METHODES.....</b>		<b>277</b>
I.	Matériels.....	277
II.	Production de hPEBP1 (100).....	278
III.	Protocole de dessalage.....	280
1.	Dessalage des nucléotides.....	280
2.	Dialyse des protéines.....	280
3.	Micro gel filtration.....	281
IV.	Conditions de digestion.....	281
V.	Conditions d'analyse.....	282
1.	Conditions ESI natives.....	282
2.	Conditions ESI dénaturantes.....	282
VI.	Paramètres instrumentaux.....	282
1.	ESI-IT.....	282
2.	MALDI-TOF.....	286
3.	ESI-UHR-Qq-TOF (maXis).....	287
<b>ANNEXES.....</b>		<b>289</b>
PRINCIPE DE L'ANALYSE DES IONS.....		290
DETERMINATION DU $K_D$ PAR MS.....		296
CARTES PEPTIDIQUES DES PROTEOLYSES DE hPEBP1.....		297
PROTOCOLE DE DESSALAGE DES NUCLEOTIDES.....		301
<b>BIBLIOGRAPHIE.....</b>		<b>303</b>
WEBOGRAPHIE.....		319

# Introduction générale

L'étude des interactions non-covalentes et des relations structure-fonction sont à la base de la compréhension des systèmes biologiques. Ainsi, l'ADN double brin, support de l'information génétique, est maintenu par des liaisons non-covalentes entre les bases complémentaires. Les protéines, qui participent au bon fonctionnement de tout l'organisme, sont maintenues dans une conformation active par des interactions non-covalentes. Les activités cellulaires sont assurées par des interactions non-covalentes temporaires entre anticorps et antigène, entre enzyme et substrat, ainsi que par d'autres complexes de type protéine-ligand.

L'étude d'édifices non-covalents est une application originale et plutôt inestimable de la MS. En effet, l'utilisation d'une technique qui détecte des espèces en phase gazeuse pour étudier des assemblages maintenus en solution par des interactions non-covalentes (interactions électrostatiques, liaisons hydrogène, interactions de type van der Waals, effet hydrophobe) peut sembler contre-intuitive en raison de la fragilité intrinsèque de certaines de ces interactions. Les travaux pionniers de deux équipes américaines (1, 2) ont permis de montrer clairement, au début des années 1990, que des liaisons non-covalentes spécifiques de type protéine/ligand peuvent être très bien préservées lors du processus de l'ionisation ESI.

La MS supramoléculaire apparaît aujourd'hui comme l'une des techniques de choix pour l'étude des interactions protéine/protéine ou protéine/ligand, en complément de techniques plus classiquement utilisées dans ce domaine telles que la spectrophotométrie, la fluorimétrie, la cristallographie, la RMN, ou la SPR, par exemple. Par rapport aux autres méthodes biophysiques, la spectrométrie de masse a l'avantage d'être rapide, précise, et peu consommatrice d'échantillon. La molécule à analyser ne doit remplir que deux conditions, être ionisable et volatile, car la MS analyse les composés en phase gazeuse selon leur rapport masse sur charge ( $m/z$ ). Il n'est pas nécessaire qu'elle possède des propriétés spectrales particulières, qu'elle soit cristallisable ou qu'elle soit fixée sur un support. Enfin, la MS mesure directement une dimension physique fondamentale des complexes, à savoir leur masse.

Récemment, la protéine Phosphatidylethanolamine Binding Protein 1 (PEBP1) / Raf Kinase Inhibitory Protein (RKIP) humaine a été découverte comme suppresseur naturel de métastases (3). Pour le cancer de la prostate, elle est d'ailleurs utilisable comme marqueur pronostique en

soutien aux décisions thérapeutiques (4). De plus, une petite molécule nommée locostatine, a été découverte pour inhiber la migration de certaines cellules, et sa cible cellulaire est justement la PEBP.

L'objectif de mon travail de thèse est de caractériser l'interaction non-covalente entre hPEBP1 et des nucléotides naturels ou analogues de nucléotides de synthèse. Plus particulièrement, il s'agit d'approfondir les relations structure-activité des complexes hPEBP1-nucléotide en les criblant par spectrométrie de masse supramoléculaire.

L'expression de la protéine hPEBP1 est ubiquitaire au sein des cellules d'un même organisme et sa séquence est très conservée entre les espèces. Ces caractéristiques suggèrent qu'elle joue un rôle essentiel dans la vie des cellules. Des activités d'inhibition de protéases, et de kinases régulant des voies de transduction du signal contrôlant la prolifération cellulaire, la survie ou la mort cellulaire, la réponse immunitaire ou la neurotransmission lui sont connues. Le rôle global de la PEBP dans la cellule n'est cependant pas encore clairement établi. Au démarrage de cette thèse, des structures cristallographiques de la protéine avaient déjà été obtenues mais il y avait très peu de données sur les relations structure-fonction.

Ce manuscrit s'articule en deux parties.

- L'Introduction fait l'objet d'un rappel bibliographique destiné à placer ce travail dans le contexte actuel des connaissances sur les protéines PEBP d'une part et de celles sur la spectrométrie de masse supramoléculaire d'autre part.
  
- La partie « Résultats et Discussion » est divisée en 5 chapitres.
  - Le chapitre I est consacré aux optimisations expérimentales qui ont été réalisées pour cribler des ligands rapidement et pour conserver des édifices non-covalents par MS.
  - Le chapitre II est la mise en application des optimisations du chapitre I pour réaliser, dans le cadre de collaborations, des études qualitatives ou quantitatives de systèmes non-covalents.
  - Le chapitre III concerne le développement d'une méthode de détermination de  $K_D$  plus précise, indispensable pour le criblage de ligands selon leur affinité.
  - Le chapitre IV est consacré à la caractérisation du site de liaison nucléotidique de la PEBP grâce au criblage d'analogues nucléotidiques réalisé par MS,

corrélé aux criblages en RMN et en biologie cellulaire par les équipes collaboratrices.

- Le chapitre V tente d'approfondir la question des fonctions de hPEBP1 par la recherche d'une activité enzymatique, en examinant tour à tour différents types de catalyse potentielle.





# Introduction

# Section 1 : PEBP/RKIP, une protéine suppresseur de métastases

La protéine Phosphatidylethanolamine Binding Protein (PEBP), également nommée Raf Kinase Inhibitory Protein (RKIP) chez certaines espèces, est le membre prototypique de la famille des protéines PEBPs. Les PEBPs ont conservé au cours de l'évolution la poche à anion (5). Initialement, en 1984, la protéine PEBP de mammifère fut purifiée à partir de cerveau de bœuf (6) et nommée "protéine cytosolique soluble de 23 kDa", puis caractérisée comme une protéine liant la phosphatidyléthanolamine (PEBP) (7). En 1999, à l'issue d'un criblage de molécules par double hybride chez le rat (8), PEBP fut nommée RKIP pour sa capacité à se lier et à inhiber la kinase Raf. La PEBP agit dans un grand nombre de processus cellulaires et physiologiques de l'organisme sain. Elle intervient par exemple au niveau de plusieurs voies de transduction des signaux cellulaires : elle contrôle les protéines G, inhibe les voies MAP-kinases et NFκB. Elle est impliquée dans différentes maladies comme le cancer et la maladie d'Alzheimer.

## I. Localisation et caractéristiques physico-chimiques

### 1. Localisation

Les PEBPs sont solubles dans le cytoplasme. Chez le rat, elles sont majoritairement associées à la membrane plasmique ou à celle du réticulum endoplasmique granuleux (9). L'étude de la PEBP1 bovine a montré que cette interaction se fait grâce à des liaisons non-covalentes, la protéine interagissant avec les groupements polaires chargés négativement des couches phospholipidiques (10). Une partie des PEBPs est également sécrétée dans le milieu extracellulaire. Ainsi, la protéine humaine hPEBP1 est présente dans le liquide céphalorachidien et dans le sérum (11).

Des études d'immunolocalisation ont montré que la PEBP1 est retrouvée au niveau de nombreux tissus. Dans le cerveau, rPEBP1 est présente dans les oligodendrocytes, les cellules de Schwann, les cellules de Purkinje dans le cervelet, ainsi qu'au niveau des six couches de cellules spécifiques du cortex et de l'hippocampe (12). On trouve aussi rPEBP1 dans des

cellules et tissus non neuronaux, comme les spermatides, les plaquettes sanguines, les cellules interstitielles de Leydig, les entérocytes, les mégacaryocytes de la rate, au niveau de l'épithélium de l'épididyme, de la prostate, du tube contourné proximal du rein, du petit intestin, des bronchioles, des oviductes et de l'utérus, ainsi qu'au niveau d'organes entiers comme le coeur, le foie, les ovaires, les glandes mammaires et la thyroïde (12). PEBP1 est observée dans la salive (13), mais pas dans le lait ni dans les sécrétions prostatiques (14). D'une manière générale, la PEBP1 est considérée comme une protéine ubiquitaire.

## 2. Caractéristiques physico-chimiques

La PEBP1 humaine est une protéine comportant 186 acides aminés. Son point de pH isoélectrique théorique est de 7.6, mais le calcul ne prend pas en compte la répartition des charges à la surface de la molécule. Il convient donc mieux d'utiliser le point isoélectrique mesuré expérimentalement, et qui est de 8.3.

Seule la PEBP1 bovine a été étudiée en détail par dichroïsme circulaire, par fluorescence statique et dynamique, et par microcalorimétrie (15). L'étude de sa stabilité thermique a montré un profil de dénaturation irréversible à pH neutre dû à sa précipitation, alors qu'à pH 4 la protéine est correctement repliée et sa dénaturation est réversible. Dans les deux cas, la transition se produit à 54°C (15).

La PEBP humaine recombinante exprimée chez *E.coli* et la PEBP bovine extraite de cerveau bœuf ne présentent pas de modifications post-traductionnelles visibles par spectrométrie de masse. Cependant, certaines modifications peuvent intervenir dans des situations bien précises, par exemple pour la régulation de l'activité de ces protéines. Ainsi, la phosphorylation de la serine 153 de la PEBP1 humaine par la protéine kinase C (PKC) permet de moduler ses capacités d'interaction avec d'autres protéines intervenant dans la voie des MAP-kinases et dans l'internalisation des récepteurs liés aux protéines G hétérotrimériques (16).

## II. Structure

### 1. Structure primaire

#### 1.1. Famille PEBP

La famille des PEBPs est un groupe de protéines très conservées qui compte plus de 400 membres ubiquitaires présents chez de nombreux organismes (mammifères, insectes, plantes, levures, bactéries). La séquence consensus qui définit l'appartenance à la famille des PEBP (6) est:

**DPD-x-P-x(11)-H-x(28-30)-G-x-HR**

Un alignement de séquences a été réalisé avec le logiciel en ligne Multalin [a] sur 12 PEBPs issues d'espèces différentes. Il met effectivement en évidence la séquence consensus précédemment décrite.

	< αA >	< βA >	< βB >	< βC >	< αB >	< βE >	< βF >		
Human	PVDLSKWSGPLSLQEVDEQPQHPLHVTYAGA	AVDELGKVL	TPTQVKNRPTSISWDGLD	SGKLY	TLVLT	DPD		71	
Monkey	PVDLSKWSGPLSLQEVDEQPQHPLHVTYAGA	ALDELGKVL	TPTQVKNRPTSISWDGLD	SGKLY	TLVLT	DPD		71	
Bovine	PVDLSKWSGPLSLQEVDERPQHPLQVKY	GGAEVDELGKVL	TPTQVKNRPTSITWDGLD	DPGKLY	TLVLT	DPD		71	
Rat	AADISQWAGPLSLQEVDEPPQHALLRVDY	GGVTVDELGKVL	TPTQVMNRPSSISWDGLD	PGKLY	TLVLT	DPD		71	
Mouse	AADISQWAGPLCLQEVDEPPQHALLRVDY	AGVTVDELGKVL	TPTQVMNRPSSISWDGLD	PGKLY	TLVLT	DPD		71	
Oncho	MLVDSAFKEHGIVPDVVSTAPTKLVNVS	YNLTVNLGNELTPTQVKNQPTKVS	WDRAE	PGALY	TLVMT	DPD		84	
Droso	NVRRIMKEMEVIPEILDEPPRELLRIKY	DNTIDIEEGKTY	TPTTELK	FQPRLD	WNA	DPESFY	TVLMIC	PD	74
Toxc	CAGCGFISSGIVPLVVTSA	PSRRVSVTFANNVQVNC	NTLTAQVANQPTVT	WEAQ	PNDRY	TLIMV	DPD	140	
Yeast	QMOKSVPQANAYVPQDD					DLFTLV	MTDPD	97	
Plasm	PPD						CYCFVLFM	VPD	78
Arab						VEIHGGDL	RSFFTL	IMTDPD	76
Antirr						VEVHGGDM	RSFFTL	IMTDPD	78
Consensus	..d.s.....vde.p...l.v.y.....elgk.l	tptqv.n.p....	wdg.dpg.ly	TLVmt	DPD				

	< βG >	< βH >	< βI >			
Human	APSRKDPKYREWHHFLVVMKGN	DIS	SGTVLSDYVGS	GPPKGTGLHRYVWLVY	124	
Monkey	APSRKDPKYREWHHFLVVMKGN	DIS	SGTVLSDYVGS	GPPKGTGLHRYVWLVY	124	
Bovine	APSRKDPKYREWHHFLVVMKGN	DIS	SGTVLSDYVGS	GPPKGTGLHRYVWLVY	124	
Rat	APSRKDPKFRWHHFLVVMKGN	DIS	SGTVLSEYVGS	GPPKGTGLHRYVWLVY	124	
Mouse	APSRKDPKFRWHHFLVVMKGN	DIS	SGTVLSDYVGS	GPPSGTGLHRYVWLVY	124	
Oncho	APSRKNPVFREWHHLLIINIS	GNVS	SGTVLSDYIGSG	PKRGTGLHRYVFLVY	137	
Droso	APNRENPMYRSWHLHVLVNV	PGLDIM	KGQPISEYFG	PLPKDSGIQRVILVY	127	
Toxc	FPSAANGQQGRLHWV	INIPGN	NIA	GGTTLAAFQ	PSTPAANTGVHRYVFLVY	193
Yeast	APSKTDHKWSEFCHLVECDL	KLLNEATH	ETSGATEFF	FAEFNTKGSNTLIEYMG	PAPPKGSGPHRYVFLLY	168
Plasm	YPSRLRPDGKEYIHWV	SGIKTKELIKGTQ	KNCVTILP	YVGSPIKGTGLHRIS	FIIS	136
Arab	VPGSPDPFLKEHLHWI	VNTNIPGTT	DATFGKEV	SYE	LPRPSIGIHRFVFLV	128
Antirr	VPGSPDPYLRHLHWI	VTDIPGTT	DSSFGKEV	SYE	MPRPNIGIHRFVFLV	130
Consensus	aPsrkdpk.rewhHvlvvn.kgn.is.....		.sgtvls.yvgs	GppkgtGIHRYvflvy		

	βI >	< αC >	< βJ >	< αD >		
Human	EQRPLKCD	EPILSNRSGDHRGKFKVASFRK	KYELRAPVAGTCYQAE	WDDYVPKLYEQLSGK	186	
Monkey	EQRPLKCD	EPILSNRSGDHRGKFKVASFRK	KYELGAPVAGACYQAE	WDDYVPKLYEQLSGK	186	
Bovine	EQEGPLKCD	EPILSNRSGDHRGKFKVASFRK	KYELGAPVAGTCYQAE	WDDYVPKLYEQLSGK	186	
Rat	EQEQLNCD	EPILSNKSGDNRGKFKVESFRK	KYHLGAPVAGTCFQAE	WDDSVPKLHDQLAGK	186	
Mouse	EQEQLSCD	EPILSNKSGDNRGKFKVETFRK	KYNLGA	PVAGTCYQAEWDDYVPKLYEQLSGK	186	
Oncho	KQPGS	ITDTQHG	GNRRNFKVMDF	FANKHHLGNPVAGNFFQAKHED	181	
Droso	QQSDKLD	DFDEK	MELSNADGHSNFDVMK	FTQKYEMGSPVAGNIFQSRWDEYVPEL	MKTLYGUSE	191
Toxc	RQPA	INSPLLN	LVQDSERPGFTTAF	ATQFNLGSPYAGNFYRSQA	241	
Yeast	KQPK	GVDSSK	FSKIKDRPNW	GYGTPATGVGK	WAKENNLQLVASNFFYAETK	219
Plasm	LIKEED	KDNITGLPHYK	GEKYITRVKFN	NYESVHNI	AQINNMKIVGYNWCQIEG	190
Arab	RQ	KQRRVIFPN	IPS	RDHFNTRKFAVEYDLGLP	VAAVFFNAQRETAARKR	177
Antirr	KQ	KKRQ	QAMLSPVWCRDGF	NTRKFTQENELGLP	VAAVFFNQCQRETAARRR	181
Consensus	.Q...lk.de.il.n..gd.r.kFkv..fr.ky.lgaPvag...qa.wdd.vpkI...l.g...					

Figure 1: Alignement de séquence pour une sélection de membres de la famille de PEBP : homme, singe, bovin, rat, souris, *Onchocerca volvulus* antigen 16 (OV16), *Drosophila*, *Toxocara canis* TES26 (Toxc), yeast (TFS1), *Plasmodium falciparum* (Plasf), *Arabidopsis thaliana* (TFL1) et *Antirrhinum* (CEN protein). La localisation des éléments de structures secondaires de la PEBP humaine est indiquée. Bleu : résidus conservés avec un pourcentage d'identité < 50%. Rouge : résidus conservés avec un pourcentage d'identité > 90%. L'alignement est basé sur la séquence de la PEBP humaine, les parties N-terminales d'*Onchocerca volvulus* antigen 16 (OV16), *Drosophila*, *Toxocara canis* TES26 (Toxc), yeast (TFS1), *Plasmodium falciparum* (Plasf), *Arabidopsis thaliana* (TFL1) et *Antirrhinum* (CEN protein) ne sont pas montrées.

La haute conservation de cette séquence au sein des membres de la famille PEBP suggère qu'elle joue un rôle clé sur le plan biologique.

D'après un autre alignement de séquences, Simister et ses collaborateurs ont distingué, en 2002, 4 sous-familles de PEBP de mammifères : PEBP-1, 2, 3 et 4 (17, 18). Cette classification réalisée à partir de 13 protéines de mammifères, ne permet cependant pas de classer correctement toutes les PEBPs, notamment celles des plantes. La majorité des différences de séquences entre les sous-familles est observée dans la région N-terminale, la plupart des variations se situant au niveau des dix premiers résidus. Les membres de la première famille, PEBP1, partagent des caractéristiques structurales et une séquence d'environ 190 résidus. La sous-famille PEBP4 est caractérisée par 2 insertions, entre les résidus 55-56 et 102-103 de hPEBP1, et d'une délétion, entre les résidus 131 et 134 de hPEBP1.

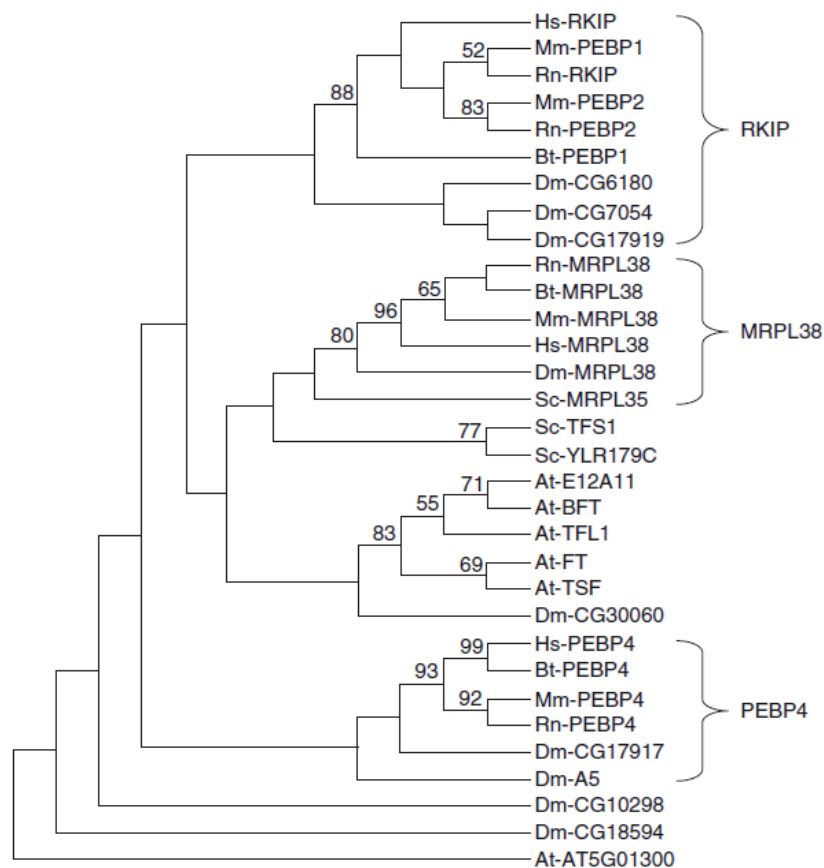


Figure 2: Arbre phylogénétique: Hs *Homo sapiens*, Rn *Rattus norvegicus*, Mm *Mus musculus*, Bt *Bos taurus*, Dm *Drosophila melanogaster*, Sc *Sacharomyces cerevisiae*, At : *Arabidopsis thaliana*, reproduit de la publication de Zeng (19)

En 2008 un arbre phylogénétique a été réalisé par l'équipe de Zeng, il indique que les protéines PEBPs se séparent en sous-familles distinctes au cours de l'évolution, incluant les familles PEBP1/2 (RKIP), PEBP4, la protéine L38 de ribosomes mitochondriaux (MRPL38) et d'autres groupes non-définis (19). Deux protéines sont homologues si elles ont divergé à partir d'une séquence ancêtre commune. Elles sont orthologues si leur divergence est due à la spéciation, à savoir le gène ancêtre commun se trouvait dans un organisme ancêtre commun avant la séparation des espèces. Elles sont paralogues si leur divergence est due à la duplication du gène ancêtre au sein d'un même organisme.

Les différentes sous-familles présentent des fonctions biologiques tantôt similaires et tantôt distinctes. Par exemple, les sous-familles reliées à PEBP (incluant RKIP, PEBP4 et PEBP2) inhibent toutes la voie de signalisation Raf/MEK/ERK (8, 16). Au contraire, l'expression de PEBP1 favorise l'apoptose (3) et des fonctions antimétastatiques (20) alors que PEBP4 inhibe l'apoptose (20).

Chez l'homme, les deux PEBPs caractérisées ont été nommées hPEBP1 (RKIP) et hPEBP4. Aucun homologue de PEBP2 n'a été identifié.

Des pourcentages d'identité ont été calculés entre différents homologues de PEBP1/2 et de PEBP4 pour mettre en évidence les différences entre les 2 sous-familles (Tableau 1).

Tableau 1: Pourcentage d'identité entre différents homologues de PEBP (calculs réalisés à partir des séquences entières avec le logiciel SIAS (Sequences Identities And Similarities) [b], h : human, m : mouse, r : rat, b : bovin

	hPEBP1	hPEBP4	mPEBP1	mPEBP2	mPEBP4	rPEBP1	bPEBP1	bPEBP4
hPEBP1								
hPEBP4	4							
mPEBP1	86	5						
mPEBP2	84	5	80					
mPEBP4	5	7	5	5				
rPEBP1	83	5	93	78	6			
bPEBP1	94	4	86	83	7	84		
bPEBP4	4	71	5	5	6	5	4	

bPEBP1 présente 94 % d'identité de séquence avec hPEBP1. Les PEBP1s de rat et de souris (rPEBP1 et mPEBP1) présentent 93% d'identité de séquence entre elles.

Plus généralement, hPEBP1, bPEBP1, rPEBP1, mPEBP1 et mPEBP2, ont entre elles au moins 78% d'identité de séquence alors qu'elles n'en ont pas plus de 7% avec hPEBP4, mPEBP4 ou de bPEBP4. Les deux sous-familles PEBP1/2 et PEBP4 sont clairement identifiés d'après les identités de séquence.

La PEBP de souris mPEBP2 est uniquement exprimée dans les testicules et joue un rôle spécifique dans la spermiogenèse (18). La PEBP2 de rat rPEBP2 est très peu décrite dans la littérature mais présente 91% d'identité de séquence avec mPEBP2 (18).

Par la suite, nous appellerons PEBP1/2, les PEBPs appartenant à la sous-famille RKIP définie par Zeng, à savoir la PEBP1 humaine (hPEBP1), la PEBP1 bovine (bPEBP1), les PEBPs 1 et 2 murines (mPEBP1 et 2), les PEBPs 1 et 2 du rat (rPEBP1 et 2).

## 1.2. PEBP/RKIP humaine

hPEBP1 est constituée de 187 acides aminés en comptant la méthionine initiale qui est par la suite coupée. Sa masse moléculaire moyenne est de 20 925.6 Da et la masse monoisotopique de 20 912.6 Da. La séquence consensus qui définit l'appartenance à la famille des PEBPs (6) se trouve entre les résidus 69 et 118 de la séquence primaire de la PEBP humaine sans méthionine (Figure 3).

```

      10           20           30           40           50           60
PVDLSKWSGP LSLQEVD EQP QHPLHV TYAG AAVDELGKVL TPTQVKNRPT SISWDGLDSG

      70           80           90           100          110          120
KLYTLVLT DP DAPSRKDPKY REWHHFLVVN MKGNDISSGT VLSDYVVGSGP PKGTGLHRYV

      130          140          150          160          170          180
WLVYEQDRPL KCDEPILSNR SGDRHGKFKV ASFRKKYELR APVAGTCYQA EWDDYVPKLY

EQLSGK

```

Figure 3 : Séquence en acides aminés de hPEBP1. En rouge, séquence consensus de la famille PEBP. En vert : les cystéines.



La protéine comporte deux cystéines libres en position 132 et 167. On utilise des agents réducteurs lors du stockage de la protéine pour éviter que les protéines ne polymérisent *via* des ponts disulfure intermoléculaires.

La protéine contient 24 résidus chargés négativement (D, E) et 30 résidus chargés positivement (K, R, H), elle devrait donc pouvoir s'ioniser facilement en mode positif pour les analyses en spectrométrie de masse.

## 2. Structure secondaire

Les structures cristallographiques de la PEBP1 humaine et de celle bovine ont été publiées simultanément et dans le même journal (*Structure*) (17, 21). Les deux structures sont extrêmement proches. Elles présentent toutes les deux 2 feuillets  $\beta$  antiparallèles et un élément C-terminal  $\alpha\beta\alpha$ . Pour la PEBP1 bovine, il est décrit que les deux feuillets  $\beta$  antiparallèles possèdent une topologie caractéristique de type « Greek key » (Figure 4).



Figure 4 : Motif « greek key »

Etant donné que mon travail de thèse a été réalisé avec la PEBP1 humaine, je m'appuierai sur la structure cristallographique de Banfield et ses collaborateurs pour décrire les différentes structures de la protéine.

D'après leurs travaux, hPEBP1 est composée de 4 hélices  $\alpha$  et de 9 brins  $\beta$  (Figure 5) (17). Un feuillet  $\beta$  composé de 6 brins ( $\beta C$ ,  $\beta J$ ,  $\beta I$ ,  $\beta F$ ,  $\beta G$  and  $\beta H$ ) compose la structure centrale de la protéine. Ce feuillet  $\beta$  est constitué de 4 brins antiparallèles bordés de chaque côté par deux petits brins, un parallèle ( $\beta C$ ) et un antiparallèle ( $\beta H$ ). Les trois brins restants dans la structure ( $\beta B$ ,  $\beta A$  and  $\beta E$ ) forment un petit feuillet  $\beta$  antiparallèle qui s'accole partiellement à un coin

du grand feuillet  $\beta$ . Sur les 4 hélices  $\alpha$ , une est formée à l'extrémité N-terminale ( $\alpha A$ ) et une autre à l'extrémité C-terminale. Les deux autres ( $\alpha B$  et  $\alpha C$ ) sont situés de part et d'autre du feuillet  $\beta$  principal. La protéine comporte deux grandes boucles en surface, l'une composée des résidus 104-118 (entre  $\beta H$  et  $\beta I$ ), et l'autre de 127-150 (entre  $\beta I$  et  $\alpha C$ ).

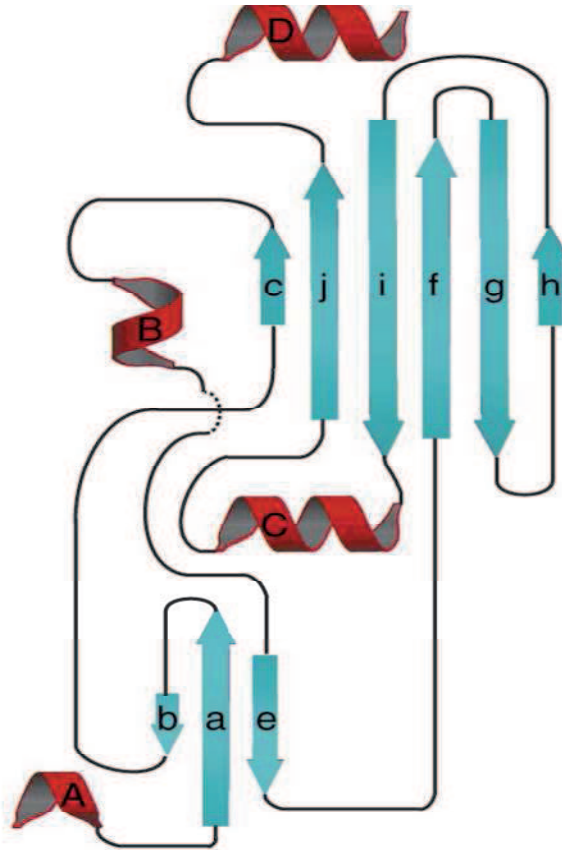


Figure 5: Représentation de la structure secondaire de hPEBP1 (17)

### 3. Structure tertiaire

#### 3.1. hPEBP1

Quatre grands types d'interactions interviennent dans le repliement final des protéines : les liaisons hydrogènes, les liaisons ioniques, les interactions dues à l'effet hydrophobe et les ponts disulfures.

hPEBP1 contient deux cystéines mais celles-ci ne forment pas de pont disulfure intramoléculaire. Dans la publication de Banfield et ses collaborateurs, la numérotation de la

séquence de hPEBP1 débute par la méthionine initiatrice, bien que celle-ci ne soit pas visible par cristallographie aux rayons X (17).

Il y a deux liaisons peptidiques en configuration *cis* formées par Pro74 et Glu83. La configuration *cis*, en principe instable, est probablement stabilisée par des interactions avec le reste de la protéine car il y a de nombreuses liaisons hydrogène dans cette zone. Les liaisons peptidiques *cis* autre que proline sont très rares dans des structures de protéines. La conservation stricte de cette orientation non favorable souligne un rôle structural important (17).

Dans la structure cristallographique, un cacodylate interagit avec la protéine *via* une liaison ionique, 4 liaisons hydrogène et des interactions de van der Waals. Une poche de liaison semble visible, localisée à la fin du feuillet  $\beta$  central. Les deux régions hautement conservées au sein de la famille PEBP sont intimement impliquées dans la détermination de la structure locale de ce site. Le glutamate *cis* est également localisé dans cette région. De plus, la poche se situe au sommet du moment dipolaire calculé pour la molécule, comme l'on pourrait s'y attendre pour une interaction avec la membrane chargée négativement. La conservation et la localisation de ces liaisons peptidiques en configuration *cis* à proximité de la zone putative de fixation des ligands anioniques laissent entrevoir un rôle majeur de ces liaisons dans les relations structure/fonction au sein de la famille des PEBP.

### 3.2. Comparaison avec d'autres PEBPs

Parmi les membres mammifères de la famille, la région de la poche située entre les résidus 63 et 126 (pour la PEBP1 bovine) montre une région de haute conservation (14) suggérant que ces éléments structuraux sous-tendent une fonction conservée dans cette région. Les structures de PEBP1 et ses homologues d'un certain nombre d'espèces ont été déterminées par cristallographie aux rayons X (5, 17, 18, 21–23). Ces structures cristallographiques révèlent une protéine globulaire avec un feuillet  $\beta$  central et une poche proéminente exposée au solvant. Les résidus qui constituent la poche appartiennent aux résidus hautement conservés cités précédemment. Sur 3 des 5 structures cristallographiques de PEBP de mammifères, un petit ligand chargé négativement est lié dans la poche (17, 21), confirmant que la poche est un site de liaison pour ligand anionique. Alors que deux des trois structures présentes dans la Protein Data Bank (1A44 et 1BEH) montrent une molécule de tampon dans le site à anion, la

troisième (1B7A) contient une O-phosphoryléthanolamine, la tête polaire soluble d'une phosphatidyléthanolamine (PE) phospholipidique. Le complexe bPEBP1 – O-phosphoryléthanolamine est la seule preuve structurale basée sur un cristal qui prouve que la poche accueille des ions d'origine physiologique. La liaison de la PE dans la poche est hypothétique et basée sur le fait que l'O-phosphoryléthanolamine, qui correspond à la tête du phospholipide, se lie à la poche. Récemment, Brady *et al.* ont trouvé par cristallographie aux rayons X que l'O-phosphotyrosine (pTyr) se liait à la poche à anion très conservée (PDB 2QYQ) (24). Quelle que soit l'espèce, les cristaux de PEBP1 ont été obtenus à des pH inférieurs à 5.5.

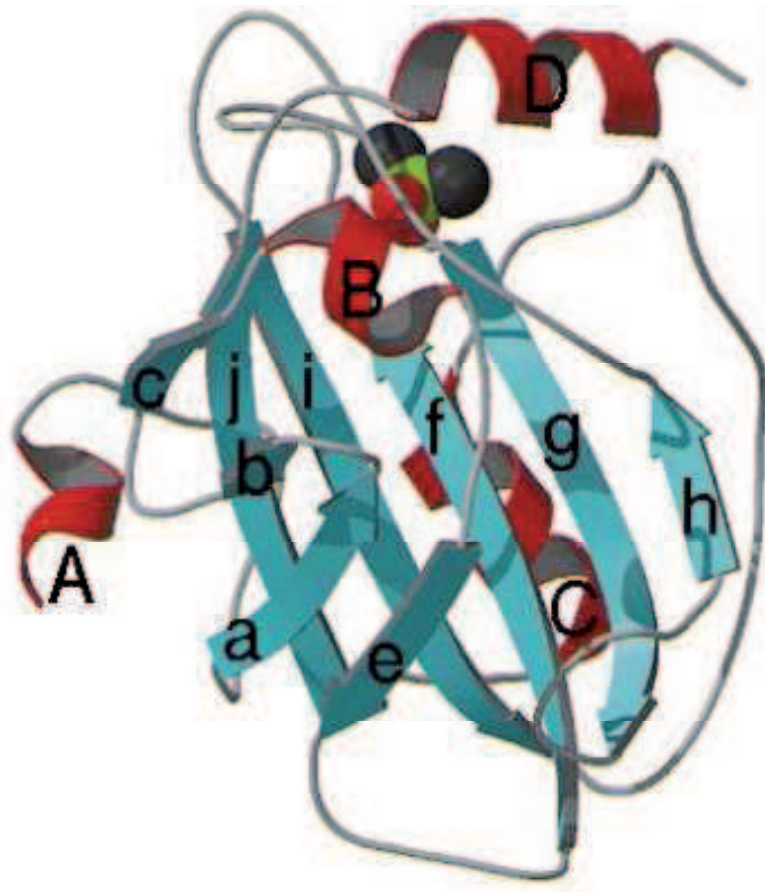


Figure 6: Structure tertiaire de la PEBP1 humaine avec un cacodylate dans la poche à anion (17)

## 4. Alignements de structure

Un alignement de structure a été réalisé en confrontant la structure de hPEBP1 (1BEH, chaîne A) à la base de donnée DALI (Tableau 2). La similarité entre deux structures est mesurée par le Z-score, qui compare des distances intramoléculaires. Le paramètre RMSD signifie Root Mean Square Deviation. Il indique la divergence de deux structures alignées l'une par rapport à l'autre.

Tableau 2 : Alignement de structure avec hPEBP1 (PDB 1BEH-A) réalisé avec la base de données DALI.

Molécule	Z-score	RMSD	Nombre d'acides aminés alignés	Nombre d'acides aminés total	Identité de séquence (%)
Famille PEBPs	39.1-8.9	0-3.2			100-23
Acyl coenzyme A thioestérase 4 humaine	6	3.1	106	405	10
Acyl coenzyme A thioestérase 2 mitochondrie humaine	5.7	3.2	109	408	10
Antigènes 4 humains des lymphocytes cytotoxiques T4	4.8	3.7	96	180	9
Facteur venin de cobra (CVF) homologue du complément C3b humain	4.7	2.9	80	613	9
Rho-GDI (GDP dissociation Inhibitor) 1 humain	4.6	3.4	98	138	9
Protéines sécrétrices de l'épididyme bovin (E1)	4.5	3.2	93	130	5
Protéines transformantes RhoA	4.6	3.6	99	136	8
ARF-like protein 2	4.4	3.1	91	141	5
Antigène tumoral cellulaire P53	4.4	3.6	105	182	4
Molécule du complément C5	4.4	3.2	89	1626	11
Thioester-containing protein 1	4.3	3.2	86	1283	12

Un tel travail avait déjà été réalisé en 1998 par Serre et ses collaborateurs avec la structure de bPEBP1 (21). En utilisant le même programme DALI, ils avaient trouvé que la structure la plus proche était celle de Rho-GDI (inhibiteur de la dissociation du GDP (GDI) des protéines

Rho), avec un Z-score de 4.3. Cette protéine possédait néanmoins un repliement beta distinct de bPEBP1.

Une similarité entre deux structures est significative si le Z-score est supérieur à 2. La similarité est forte si le pourcentage d'identité de séquence est supérieur à 20% ou le Z-score supérieur à  $n/10 - 4$  (n, le nombre de résidus de la protéine) et qu'il n'y a pas plus de 80 résidus sans aucune correspondance structurale. Dans notre cas,  $n/10 - 4 = 14$ .

Hormis les protéines de la famille PEBP, les protéines du tableau présentent toutes une similarité significative mais faible avec la structure de hPEBP1. Le Rho-GDI fait effectivement partie des protéines présentant le plus d'homologie structurale avec hPEBP1, comme l'acyl coenzyme A thioestérase humaine, l'antigène 4 humain des lymphocytes cytotoxiques T4 et le facteur de venin de cobra (CVF) homologue du complément C3b humain.

Parmi les protéines qui comptent à peu près le même nombre d'acides aminés que hPEBP1, soit entre 130 et 182 résidus, trois d'entre elles possèdent un site de liaison pour le GTP : Rho-GDI 1 humain, la protéine transformante RhoA et l'ARF-like protéine 2 (ADP-ribosylation factor-like protein 2).

### III. Ses fonctions

#### 1. Inhibition de protéases

mPEBP peut inhiber des protéases à sérine telles que la thrombine ( $K_i = 380$  nM), la neuropsine et la chymotrypsine ( $K_i = 1.8$   $\mu$ M), alors qu'elle ne partage aucune homologie avec les différentes familles d'inhibiteurs de protéases à sérine connues (25). Parmi les PEBPs de levures, Tfs1p est connue pour être un inhibiteur de la carboxypeptidase Y (26).

## 2. Inhibition de kinases

rPEBP1 est un inhibiteur efficace de plusieurs kinases intervenant dans diverses voies de transduction du signal : Raf-1, GRK2 (27), ainsi que les kinases TAK-1, NIK, IKK $\alpha$ , et IKK $\beta$  de la voie NF $\kappa$ B (28). La protéine porte d'ailleurs le nom de Raf Kinase Inhibitor Protein (RKIP) (8).

Des travaux publiés au cours des 12 dernières années ont montré que PEBP1 est impliquée dans 3 voies fondamentales de transduction du signal : la voie Raf/MEK/ERK, la voie NF $\kappa$ B et la voie de transduction médiée par les récepteurs couplés aux protéines G (GPCRs). Les voies Raf/MEK/ERK et NF $\kappa$ B sont des voies de transduction type MAP kinase (MAPK). Les voies MAP kinases sont constituées de 3 modules séquentiels de kinases. La MAPK kinase kinase est capable de phosphoryler et d'activer la MAPK kinase, qui à son tour phosphoryle et active la MAPK.

### 2.1. Inhibition de la voie des MAP kinases Raf/MEK/ERK

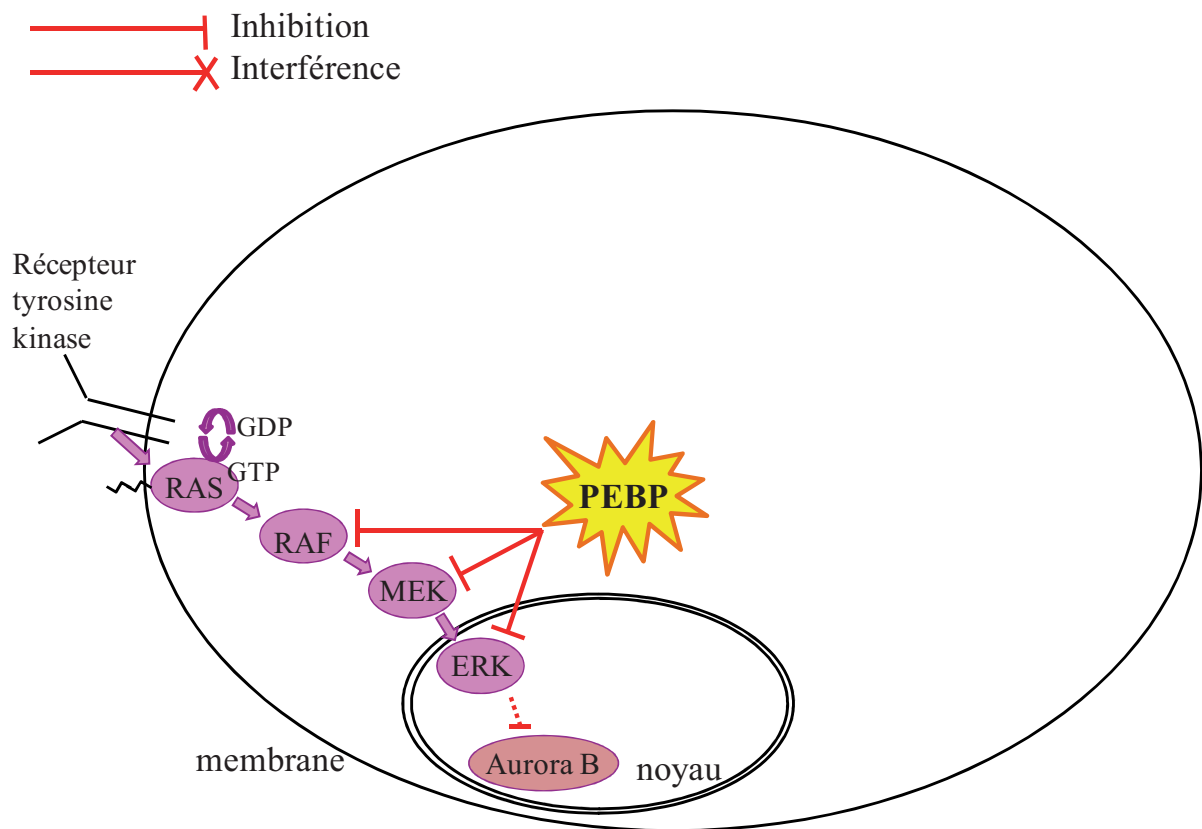


Figure 7 : Inhibition de la voie Raf/MEK/ERK par la PEBP

La voie Raf/MEK/ERK est essentielle pour la prolifération cellulaire, la différenciation, l'apoptose, la survie et la migration cellulaire (29). Sa dérégulation peut contribuer à de nombreuses maladies humaines comme le cancer (30, 31) et des problèmes lors du développement (32).

Des signaux extracellulaires tels que des facteurs de croissance sont capables d'activer une cascade de réactions en se fixant à des récepteurs transmembranaires à activité tyrosine kinase, les TRK. Après phosphorylation de la tyrosine, les récepteurs vont entraîner une activation de la protéine GTPase Ras qui va ensuite déclencher la cascade de phosphorylations de la voie MAPK. Ras va phosphoryler et activer la kinase Raf qui va elle-même phosphoryler et activer la kinase MEK qui va à son tour phosphoryler et activer la kinase ERK (Figure 7). Raf, MEK et ERK sont toutes des sérine/thréonine kinases, et sont de type MAPKKK, MAPKK et MAPK, respectivement. Sur toutes les protéines capables de se lier à Raf-1, seule la protéine PEBP (1, 2 ou 4) est connue pour inhiber la cascade de phosphorylation. Le premier rôle de PEBP1 dans la signalisation cellulaire a été identifié par double hybride chez le rat lors d'un criblage de molécules ayant la capacité de se lier à Raf-1 (8). La liaison de PEBP1 à Raf-1 inhibe directement l'activation de Raf-1 en bloquant la phosphorylation de ses 2 sites d'activation clés (Ser 338 et Tyr 341). Cela entraîne la dissociation du complexe Raf-1/MEK. La voie MAPK se retrouve bloquée (33). PEBP1 se lie aussi à MEK-1 et faiblement à ERK-2. PEBP1 ne peut se lier à Raf-1 et MEK simultanément et la liaison à l'un des deux partenaires est suffisante pour entraîner une inhibition de la cascade de phosphorylation. En inhibant différentes kinases de la voie, PEBP1 amplifie l'effet inhibiteur sur la cascade de kinases(34). PEBP1 ne se lie pas à Ras, et ne possède pas elle-même d'activité kinase.

Aurora B est une kinase localisée à l'intérieur des centromères en prométaphase, qui répare les erreurs de connection des chromosomes au fuseau dans le but d'assurer une bonne ségrégation durant l'anaphase. Plus récemment, Eves *et al.* (2006) ont démontré sur des cellules humaines et de rat que PEBP1 influence indirectement la kinase Aurora B et les contrôles du fuseau mitotique (35). La présence de PEBP1 phosphorylée au niveau des centrosomes et des kinétochores laisse penser que la protéine pourrait jouer un rôle durant la mitose. De plus, la déplétion de PEBP1 ou l'hyperactivation de Raf-1 augmente la vitesse du mouvement des cellules durant la mitose (35). La diminution de l'activité kinase d'Aurora B



en réponse à une diminution de PEBP1 engendre une hyperactivation de la voie MAPK. Le mécanisme d'action sur Aurora B n'est pas clair car Aurora B n'est pas un substrat direct. La régulation du fuseau mitotique et de la stabilité du génome par la protéine *via* la cascade Raf/MEK/ERK représente un mécanisme qui pourrait expliquer en partie son rôle dans la progression métastatique de tumeurs.

L'activité de PEBP1 n'est pas limitée à une seule voie de signalisation, elle régule également la voie NF- $\kappa$ B.

## 2.2. Inhibition de la voie NF- $\kappa$ B

PEBP1 inhibe chez le rat la voie de transduction du signal qui permet l'activation de la transcription du facteur NF- $\kappa$ B en réponse à une stimulation de TNF $\alpha$  ou d'interleukines 1 IL-1, des cytokines pro-inflammatoires (28).

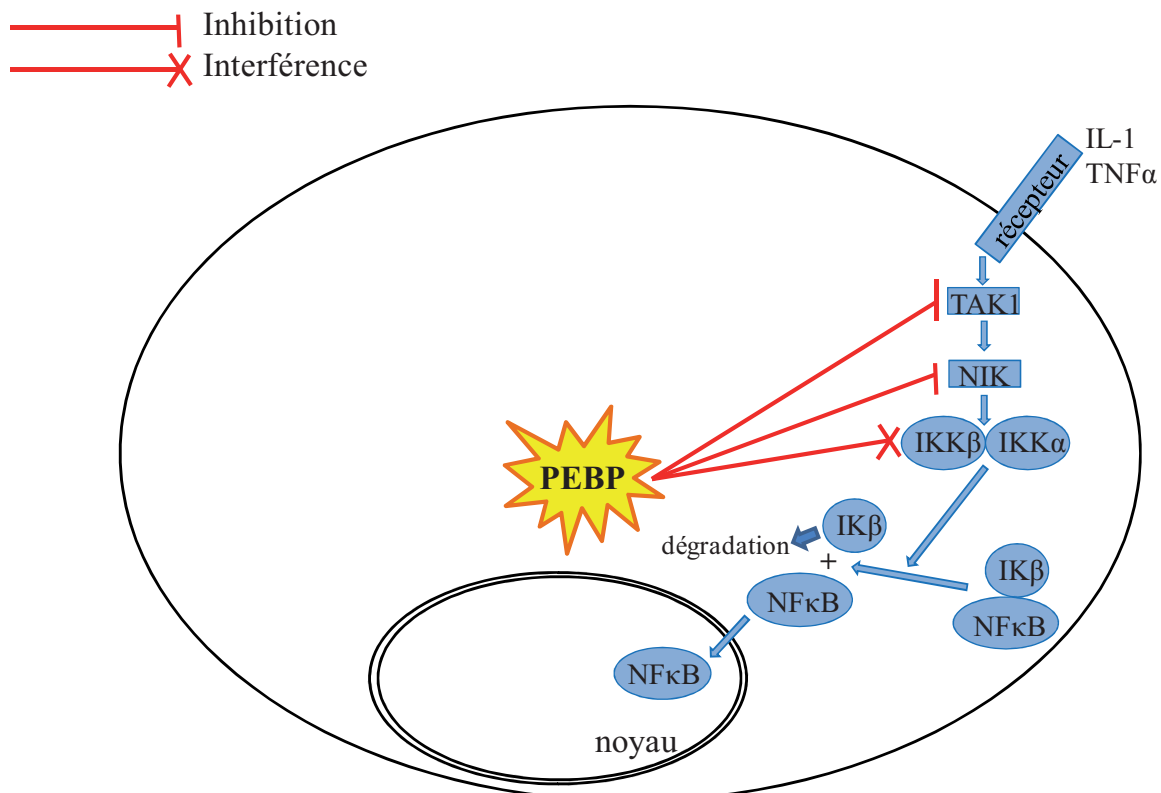


Figure 8: Inhibition de la voie NF- $\kappa$ B par la PEBP

Cette voie contrôle l'activation de récepteurs immunitaires et de cytokines (36). La famille nucléaire NFκB de facteurs de transcription joue un rôle critique dans la coordination de l'expression de gènes contrôlant la réponse immunitaire et les réponses au stress (36, 37). La modulation du niveau d'expression de PEBP1 affecte la voie NF-κB indépendamment de la voie Raf/MEK/ERK. Des analyses d'épistasie génétique impliquant l'expression ectopique de kinases jouant sur la voie NF-κB ont indiqué que PEBP1 agit en amont du complexe de kinases (IKKβ-IKKα), complexe induisant la phosphorylation et l'inactivation de l'inhibiteur de NF-κB (IKβ). Des expériences *in vitro* sur les kinases ont montré que PEBP1 inhibe l'activité de la kinase IKK induite par TNF-α. rPEBP1 interagit physiquement avec 4 kinases de la voie NF-κB : TGF β activated kinase (TAK1), NF-κB inducing kinase (NIK), IKKα and IKKβ (28). Cette voie a été proposée (mais non prouvée) comme médiant les effets anti-invasifs de hPEBP1.

### 2.3. Inhibition de la voie des récepteurs couplés aux protéines G

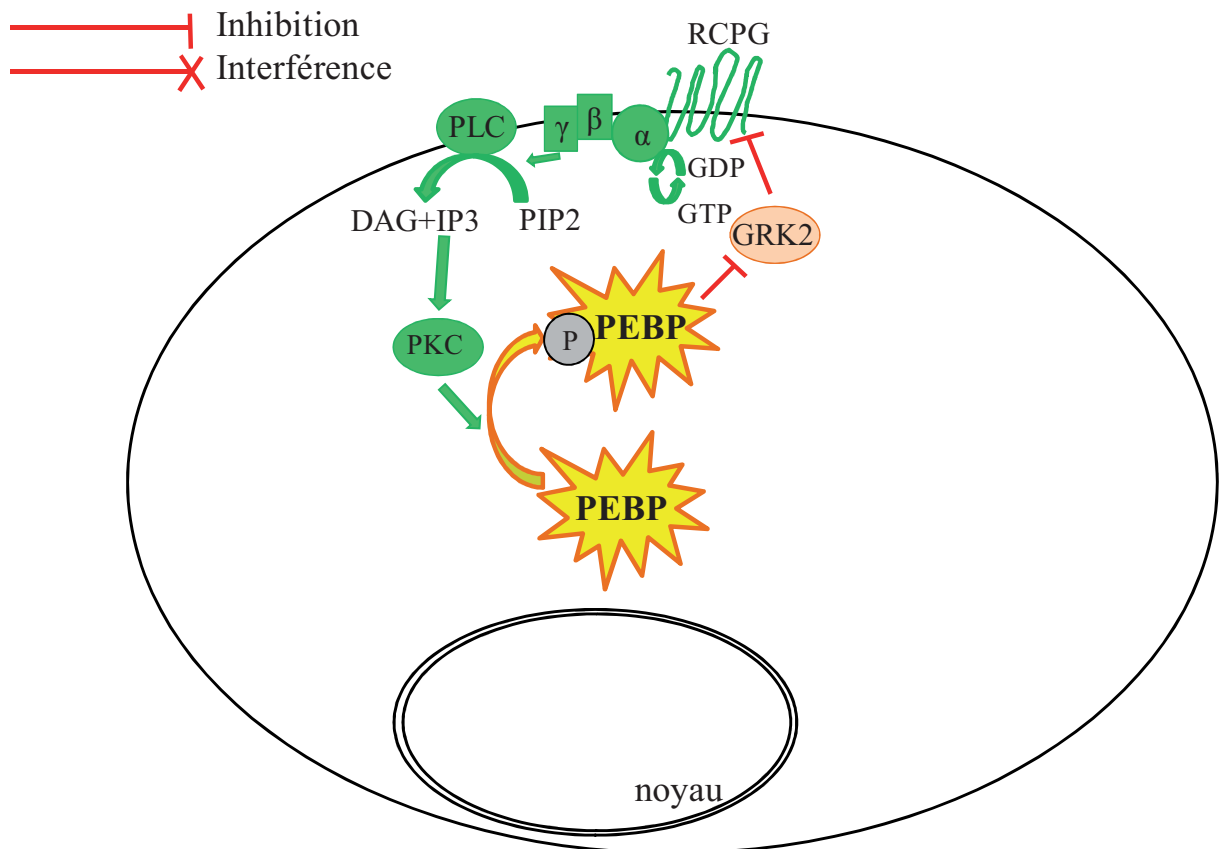


Figure 9: Inhibition de la voie des récepteurs couplés aux protéines G par la PEBP

PEBP1 peut moduler une autre voie de transduction du signal initié par les récepteurs couplés aux protéines G (RCPG). Ces récepteurs contrôlent des processus de régulation biologique essentiels tels que la régulation de la neurotransmission, de la libération d'hormones et d'enzymes, de l'inflammation et de la pression sanguine. À l'état inactif, la protéine G comprend les trois sous-unités  $G\alpha$ ,  $G\beta$  et  $G\gamma$ . L'activation du récepteur RCPG par son agoniste, se traduit par des modifications structurales qui favorisent le couplage du récepteur à la protéine G. Il en résulte une diminution de l'affinité de la sous-unité  $G\alpha$  pour le GDP, qu'elle expulse, et un GTP vient prendre sa place. La conséquence majeure de l'échange GDP/GTP semble être une dissociation du complexe hétérotrimérique :  $G\alpha$ -GTP et le complexe  $G\beta\gamma$  (38). Les sous-unités dissociées activent alors différents effecteurs spécifiques : adénylate cyclase (AC), phospholipase C (PLC)... (39). La PLC, une fois activée par la sous-unité  $G\alpha$ , clive le phosphatidyl-inositol-4,5-diphosphate (PIP2) en inositol-triphosphate (IP3) et en diacylglycérol (DAG). Ces deux second-messagers agissent à leur tour comme des effecteurs, le DAG active notamment la protéine kinase C (PKC).

La voie RCPG est contrôlée par une inhibition en boucle de rétrocontrôle (feedback). GRK2 est un inhibiteur qui phosphoryle les RCPGs actifs et initie leur internalisation (40) ce qui conduit au final à une inhibition de la voie RCPG. Il est intéressant de noter que GRK2 est une cible de PEBP1. Lorsque la rPEBP1 est phosphorylée sur la Sérine 153 par la PKC (16), elle se dissocie de Raf-1, se lie à GRK2 et bloque l'internalisation des récepteurs RCPG en inhibant GRK2, ce qui prolonge la durée de transduction du signal d'activation à travers la membrane (41, 42).

#### 2.4. Régulation de l'activité de PEBP1 par phosphorylation

La PKC phosphoryle des protéines cibles qui contrôlent la croissance, la différenciation et la transcription. La PKC phosphoryle la PEBP1 sur sa Sérine 153 (16) et régule ainsi finement l'activité de cette protéine, qui, sous forme non-phosphorylée inhibe la voie Raf/MEK/ERK et sous forme phosphorylée, lève l'inhibition de la voie RCPG. PEBP1 pourrait être un senseur environnemental qui bascule d'une régulation à une autre. Elle pourrait maintenir le système dans un état d'équilibre en conditions normales, être capable de répondre à un signal extérieur

et réguler finement négativement ou positivement les voies de signalisation en conditions anormales.

De plus, la forme phosphorylée de PEBP1 a été localisée dans des cellules humaines et de rat dans les régions du centrosome et des kinétochores de chromosomes prométaphasiques, et pourrait être impliquée dans la régulation de protéines du fuseau mitotique et du mouvement lors du cycle cellulaire (35).

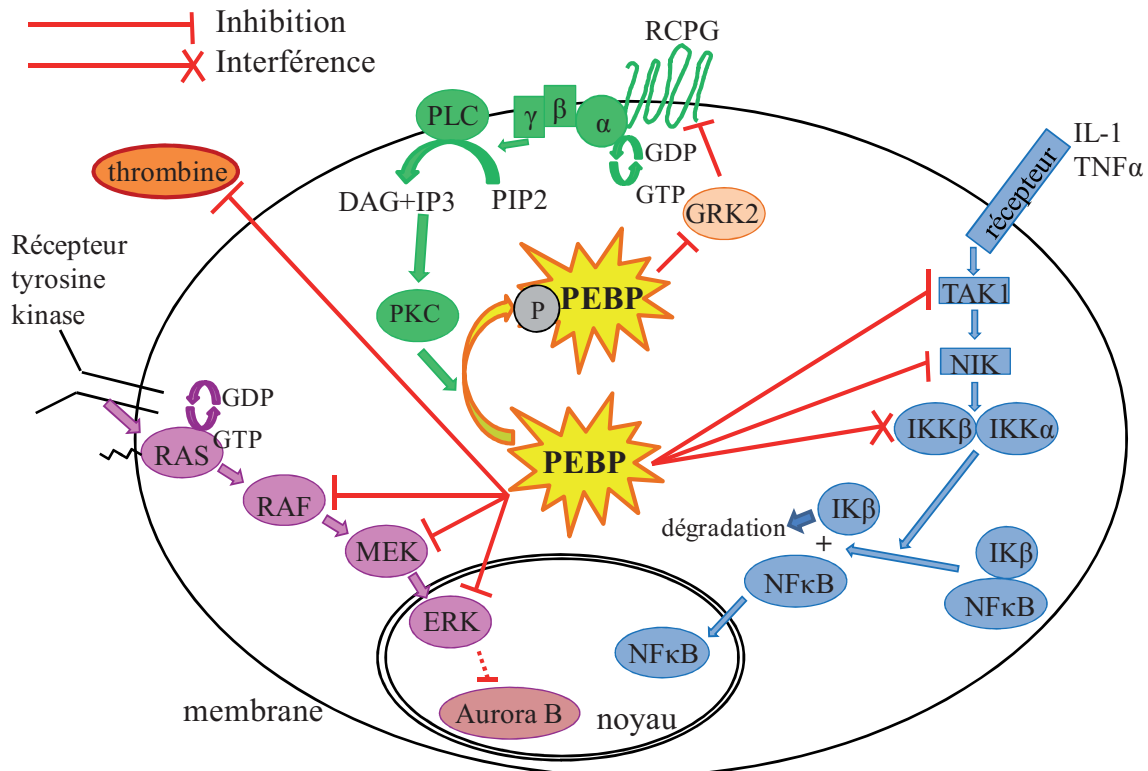


Figure 10: récapitulatif des activités cellulaires de PEBP1

### 3. Au niveau de l'organisme

#### 3.1. Suppresseur de métastases

Les métastases, qui correspondent à la formation de tumeurs secondaires dans des sites éloignés de celui de la tumeur primaire (43), se forment en plusieurs étapes. Dans un premier temps, une cellule cancéreuse acquiert un phénotype métastatique, quitte la tumeur primaire,

entre dans la circulation sanguine (invasion), envahit un nouveau tissu distant du premier site (migration), puis croît pour former une tumeur secondaire (Figure 11).

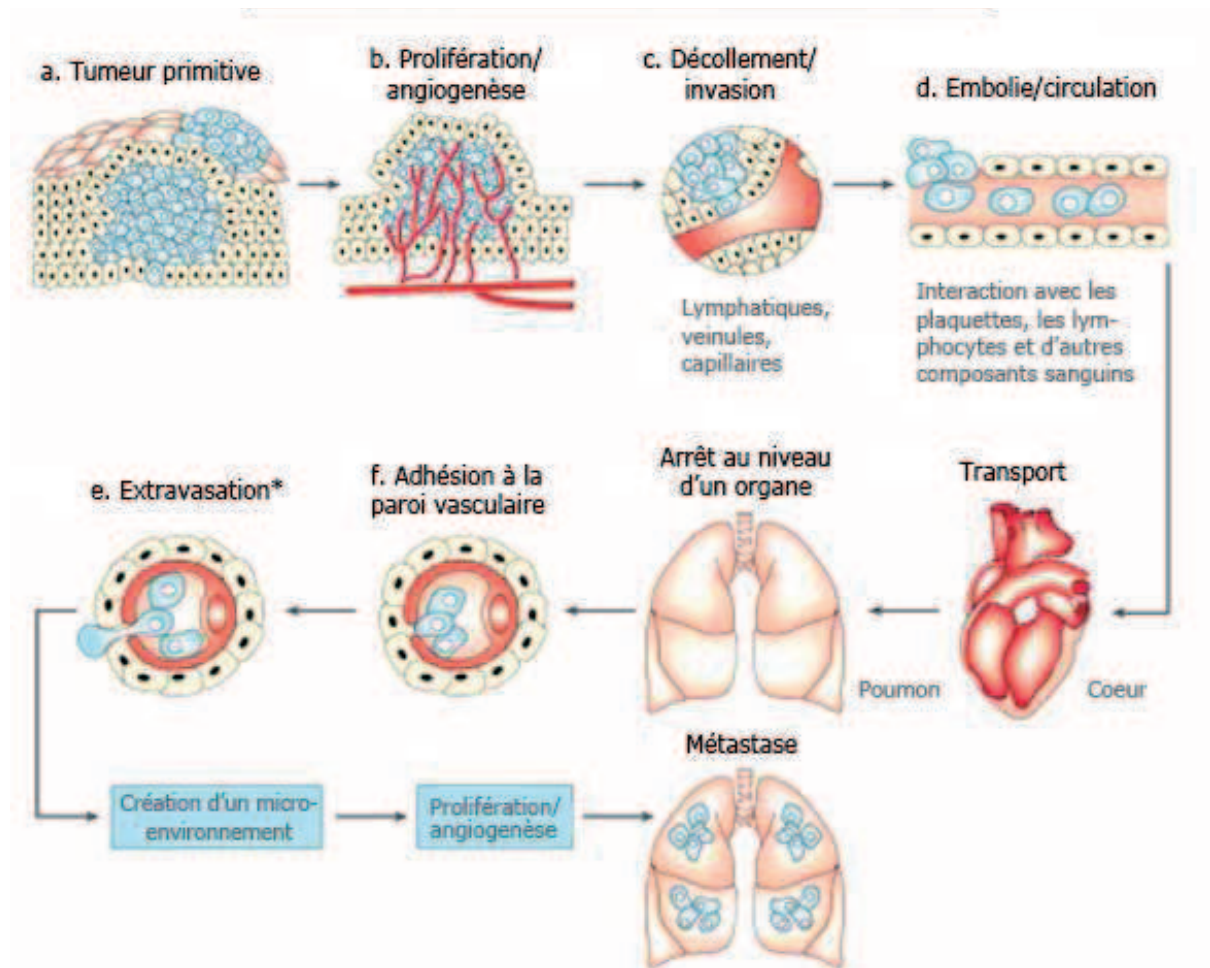


Figure 11 : Principales étapes de la formation des métastases (A. Bellahcène, V. Castronovo (Université de Liège)) [c]

Etant donné le processus multi-étapes, il est possible de bloquer la formation de métastases en inhibant au moins une des protéines indispensables à l'accomplissement d'une de ces étapes du processus (44). Les voies moléculaires impliquées dans le détachement, la migration de cellules malignes de leur site tumoral primaire et la colonisation invasive d'organes distants sont mal connues.

Les gènes qui favorisent ou suppriment les métastases sont d'un grand intérêt pour les avancées cliniques et diagnostiques, et pour la conception de traitements spécifiques des tumeurs à pouvoir métastatique. De nombreux travaux suggèrent que le gène de PEBP1 est le

produit d'un nouveau gène candidat sur la très courte liste des suppresseurs de métastases. Le premier élément de preuve de ce rôle vient de lignées cellulaires humaines dérivées de cancers métastatiques de la prostate, qui montrent des niveaux d'ARNm et de protéines hPEBP1 faibles comparés à ceux de cellules de lignées de tumeurs primaires non-invasives (3). Il a été suggéré qu'une expression faible ou nulle de hPEBP1 favorise l'apparition de caractéristiques invasives dans des cellules tumorales de la prostate (3). De plus, dans certaines études, la surexpression de hPEBP1 inhibe la migration et l'invasion (45, 46). Le traitement de cellules cancéreuses humaines de la prostate et du sein avec des agents chimiothérapeutiques entraîne une augmentation de l'expression de hPEBP1 et cette augmentation prédispose ces cellules cancéreuses à l'apoptose (47). Récemment, il a été montré que le niveau sérique de hPEBP1 a une valeur pronostique chez les patients atteints de cancer de la prostate, et un dosage sérique de hPEBP1 a été proposé comme marqueur pronostique en soutien aux décisions thérapeutiques (4). L'idée que hPEBP1 est un suppresseur naturel de métastases est confortée par l'observation d'une diminution significative de l'expression de hPEBP1 non seulement dans les cancers de la prostate invasifs, mais aussi dans les mélanomes malins (48), les métastases de cancer du sein localisées dans les ganglions lymphatiques (49), les insulinomes (50), les cancers colorectaux (51) et les hépatocarcinomes (46). En jouant sur les voies Raf-1 et NF- $\kappa$ B, hPEBP1 pourrait rendre certaines cellules partiellement sensibles aux signaux apoptotiques (47). Des travaux récents de Eves *et al.*, indiquent également que l'absence de hPEBP1 dans les cellules pourrait provoquer une augmentation de leur instabilité génétique (35). Considérant l'ensemble de ces données, l'altération de l'expression de hPEBP1 semble induire des caractéristiques qui sont généralement associées avec la croissance des cancers et leur propagation.

Une étude récente a montré que Snail, un médiateur de la transition épithélio-mésenchymateuse (EMT), pouvait inhiber la transcription de hPEBP1 et était corrélé négativement avec le niveau de hPEBP1 dans les tumeurs, ce qui est cohérent avec un rôle de hPEBP1 dans les métastases (52).

Il est possible qu'en plus d'inhiber l'invasion et la migration, hPEBP1 puisse empêcher l'acquisition d'un phénotype métastatique en inhibant la transition épithélio-mésenchymateuse.

### 3.2. Rôle dans la maladie d'Alzheimer

hPEBP1 est impliquée dans la maladie d'Alzheimer, maladie la plus connue dans les troubles de démence (53). La maladie d'Alzheimer se manifeste par un déclin progressif de multiples fonctions cognitives. Cela inclut des problèmes de mémoire, d'aphasie, d'apraxie, d'agnosie et/ou de perte de capacité à planifier et à organiser des activités de routine. Il semble y avoir deux formes basiques de la maladie d'Alzheimer : la forme précoce, qui affecte des individus de moins de 65 ans et la forme tardive, qui touche les plus de 65 ans. La forme précoce ou familiale est liée à des mutations pouvant avoir lieu sur 3 gènes différents (APP, PSEN1 and PSEN2) localisés sur 3 chromosomes différents (1, 14 et 21). Une diminution de l'expression des ARNm de PEBP1 a été relevée dans des cerveaux de patients atteints de la forme tardive de la maladie (54).

Les résidus 2 à 12 de hPEBP1 correspondent à l'Hippocampal Cholinergic Neurostimulating Peptide (HCNP), qui stimule la synthèse d'acétylcholine chez le rat. De plus, l'enzyme capable de cliver hPEBP1 en HCNP a été détectée dans l'hippocampe de rat (55). HCNP est impliqué dans le développement neuronal et la différenciation.

En tant que précurseur d'HCNP, hPEBP1 émerge comme un composant capable d'affecter l'action hautement régulée des cholines acétyltransférase neuronales. La formation de plaques amyloïdes  $\beta$  est corrélée avec la diminution d'expression de hPEBP1 dans l'hippocampe de souris (56). Ces observations motivent l'idée d'une implication possible de hPEBP1 dans les mécanismes pathogéniques de la maladie d'Alzheimer.

### 3.3. Rôle dans l'apoptose et la sensibilité aux chimiothérapies

Outre ses propriétés anti-métastatiques, hPEBP1 possède des propriétés pro-apoptotiques dans les cellules tumorales. L'exposition de cellules humaines de cancer du sein et de la prostate à des agents chimiothérapeutiques induit l'expression de PEBP1 et l'apoptose (47). De petits ARN interférents qui ciblent et réduisent l'expression de hPEBP1 diminuent l'apoptose induite par des agents chimiothérapeutiques. Au contraire, la restauration de l'expression de PEBP1 sensibilise les cellules à l'apoptose.

Ces résultats suggèrent que hPEBP1 joue un rôle essentiel dans l'apoptose induite par les agents chimiothérapeutiques. La perte de hPEBP1 pourrait donc conférer aux cellules

cancéreuses une résistance aux chimiothérapies et globalement favoriser la progression tumorale.

Plus récemment, il a été observé que la surexpression de hPEBP1 entraîne l'inhibition de l'expression de Yin Yang 1 (YY1) dans des cellules tumorales humaines de prostate et de mélanome, entraînant l'induction du récepteur de mort cellulaire DR5 et sensibilisant les cellules à l'apoptose médiée par TRAIL (57). Cette observation pourrait constituer une piste d'explication du rôle biologique de PEBP1 dans l'apoptose induite par les agents chimiothérapeutiques.

### 3.4. Autres fonctions

PEBP1 apparaît avoir une variété de rôles en fonction du tissu dans lequel elle se situe. PEBP1 est une PEBP ubiquitaire dans le sperme des mammifères (58), dans les testicules et les fluides épидидymaux (59, 60). PEBP1 pourrait jouer un rôle dans la biogenèse des membranes et dans la maintenance de la ségrégation des antigènes dans les spermatozoïdes (61). rPEBP1 est hautement exprimée dans les tubes séminifères, les cellules de Leydig qui produisent les stéroïdes et dans les élongations des spermatozoïde du rat (12). Elle est également fortement exprimée dans les tubes séminifères de souris (62). Le rôle spécifique de PEBP1 dans la reproduction reste encore peu clair. Des études sur des modèles de souris ont néanmoins fourni quelques indices intéressants sur le rôle de PEBP1 dans ces processus (63). En effet, la reproduction est fortement affectée chez des souris n'exprimant plus mPEBP1 (63).

Gibbons *et al.* ont également mis en évidence que mPEBP1 était impliquée dans la capacitation des spermatozoïdes (64). D'autres études sont nécessaires pour déterminer si mPEBP1 agit comme un récepteur de facteur de décapacitation ou comme un facteur de décapacitation. mPEBP1 joue néanmoins un rôle dans la fertilité des mâles.

Chez les plantes, le membre de la famille des PEBP, centroradialis (CEN), inhibe la floraison en maintenant l'indétermination des méristèmes apicaux et d'inflorescence (65). L'orthologue de PEBP/RKIP joue donc un rôle important dans la régulation de la différenciation cellulaire chez les plantes.



## IV. Interactions PEBP1 – ligand

### 1. Interaction covalente avec la locostatine

#### 1.1. La locostatine

##### 1.1.1. Structure

La locostatine est le nom trivial de la (4S)-3-[(E)-But-2-enoyl]-4-benzyl-2-oxazolidinone (66)

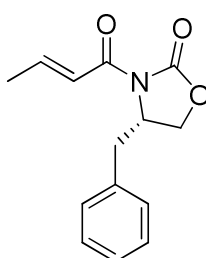


Figure 12: Structure chimique de la locostatine

##### 1.1.2. Fonction antimigratoire

La locostatine a été découverte lors d'un criblage d'agents cicatrisants sur des cellules épithéliales de rein de chien Madin-Darby (MDCK), au moyen d'un test de wound healing (Figure 13). Différents dérivés d'oxazolidinone ont été criblés (66). La (4S)-3-[(E)-But-2-enoyl]-4-benzyl-2-oxazolidinone a montré des propriétés inattendues d'inhibition de la cicatrisation, plus précisément de la migration cellulaire, ce qui lui a valu son nom de locostatine (Figure 13).

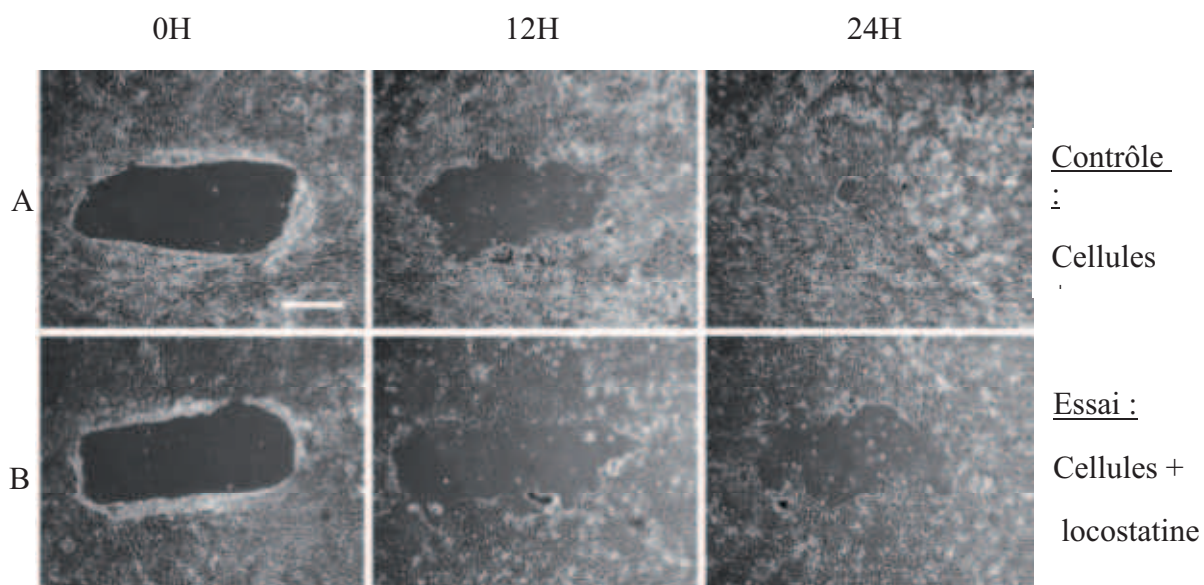


Figure 13: Images de microscopie à contraste de phase de cellules monocouches MDCK dans différentes conditions : A) Contrôle avec 0.1% de DMSO 0, 12 et 24 H après la blessure B) 50  $\mu$ M de locostatine, 0, 12 et 24 H après la blessure (66)

En 2008, Mc Henry et ses collaborateurs ont montré que la locostatine entraîne une diminution de l'adhésion cellule-substrat, que ce soit sur des cellules saines MDCK ou des cellules cancéreuses humaines d'adénocarcinome du sein (67).

## 1.2. Protéine cible : PEBP1

L'effet anti-migratoire de la locostatine est médié par la PEBP1 ; cet effet a été décrit à l'origine dans des cellules épithéliales de rein de chien Madin-Darby (MDCK) (68) qui sont des cellules normales et différenciées. La surexpression de PEBP1 dans ces cellules a conduit à des effets paradoxaux en termes de migration de cellules et d'adhésion cellule-cellule (67) ; en effet, la surexpression de PEBP1 a conduit à la dé-différenciation et à la perte du contact cellule-cellule. Ceci peut être dû au phénotype non-cancéreux des cellules qui ne présentent pas les mêmes caractéristiques ni le même comportement que les cellules métastatiques modèles.

Dans des cellules non-tumorales MEFs de rat, la locostatine a des effets délétères sur la structure du cytosquelette et sur la migration cellulaire même si les cellules n'expriment pas PEBP1 car de phénotype PEBP<sup>-/-</sup> ou ARNi, ce qui suggère que la locostatine pourrait avoir

d'autres cibles encore non-identifiées (69). Zhu *et al.*(2005) ont mis en évidence la liaison à trois enzymes : GSTO1-1, ALDH1A1 et POP (68). Les auteurs ont testé l'effet d'inhibiteurs spécifiques de chacune de ces enzymes sur la migration en wound healing et n'ont pas observé d'effet anti-migratoire. Cette observation les exclut donc comme cibles médiatrices d'un effet anti-migratoire.

### 1.3. Mécanisme de liaison

La locostatine se lie de manière covalente à rPEBP1 (68) et à hPEBP1. La liaison de la locostatine à la PEBP1 résulte d'une réaction de Michael où PEBP1 est le donneur de Michael et la locostatine l'accepteur de Michael. La PEBP1 agit comme un nucléophile et attaque l'acyl insaturé de la locostatine, qui constitue un centre électrophile. Sans cette liaison insaturée, la molécule ne possède plus de propriétés anti-migratoires (66).

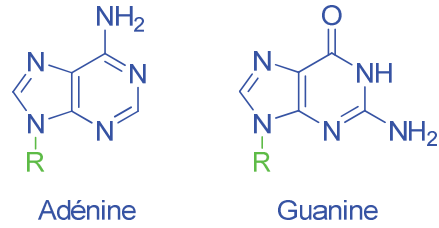
## 2. Interaction non-covalente avec des nucléotides

### 2.1. Structure des nucléotides

Un nucléotide est composé de trois parties :

- une base azotée variable en fonction du type de nucléotide (purine ou pyrimidine) fixée à l'atome de carbone du désoxyribose
- un sucre à cinq atomes de carbone
- un à trois groupements phosphates (ou acide phosphorique), fixés sur l'atome de carbone 5' du ribose

## Bases pyrimidiques



## Bases puriques

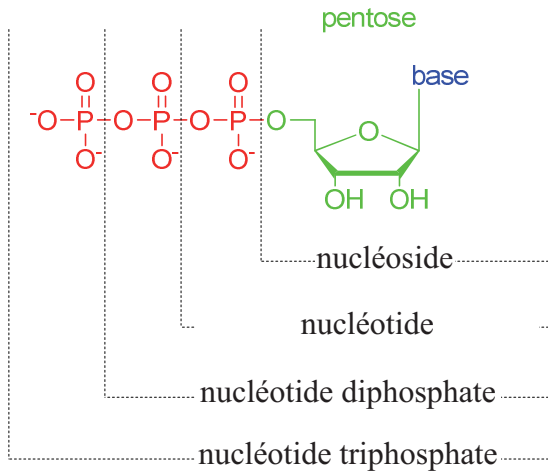
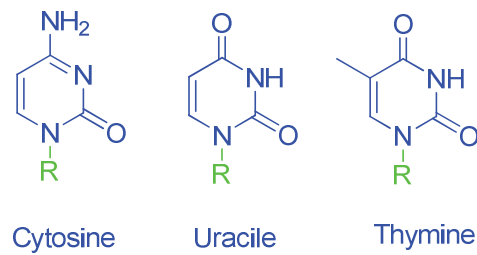


Figure 14: Composition chimique des nucléotides

## 2.2. Interactions avec PEBP1

bPEBP1 a été isolée de la fraction soluble d'extraits de cerveau bovin par chromatographie d'affinité avec un colorant immobilisé, le bleu Cibacron (6).

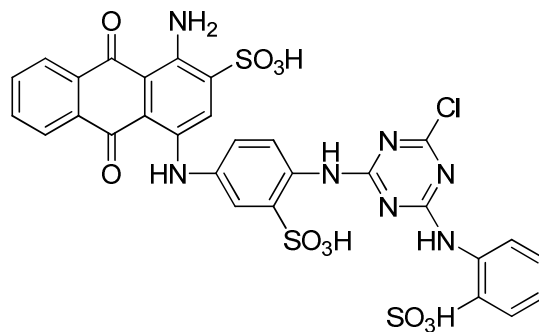
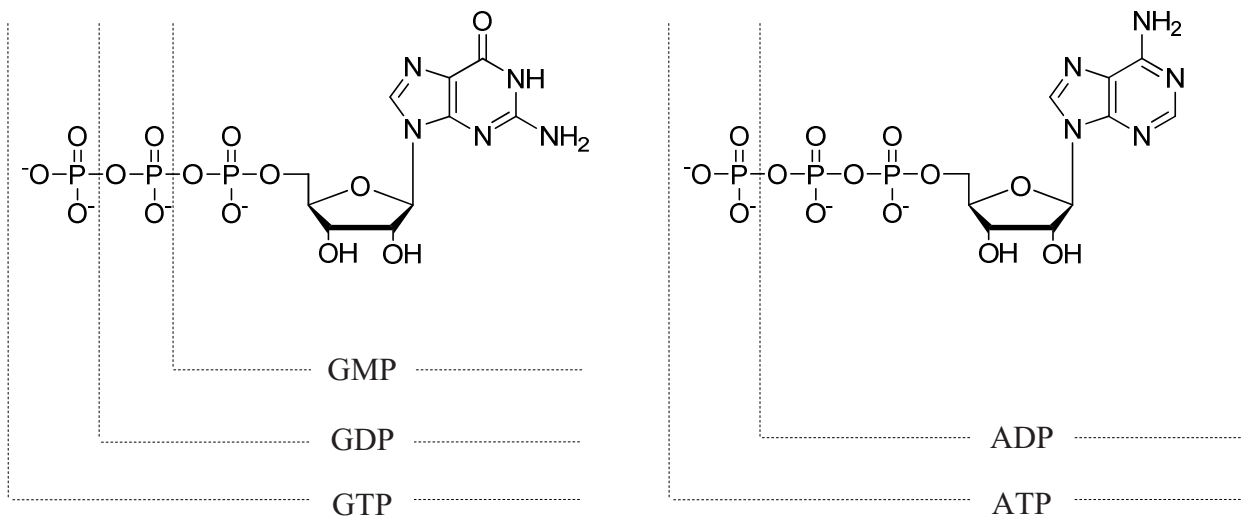


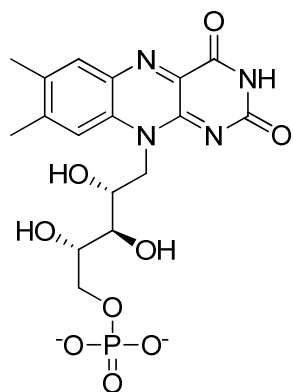
Figure 15: Structure du bleu Cibacron

Ce colorant est généralement considéré comme un analogue nucléotidique et ainsi l'affinité de bPEBP1 pour ce colorant a suggéré dès 1984 une fonction de liaison aux nucléotides pour la protéine.

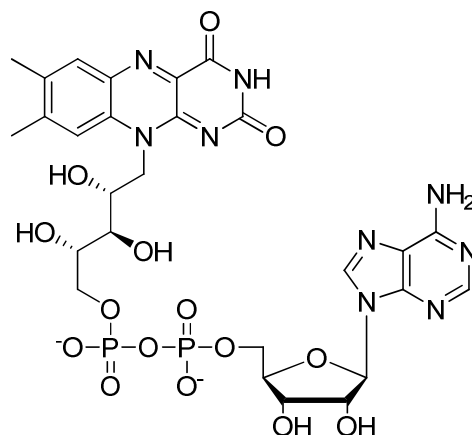
Un modèle de la PEBP bovine de cerveau a été construit sur la base d'une homologie structurale entre la bPEBP1 et le domaine N-terminal de la phosphoglycérate kinase (PGK) de levure (70). Ce modèle suggérait un possible site de liaison nucléotidique entre les résidus 112 et 125 (70) car ce site présentait des homologies avec le pli Rossman  $\beta\alpha\beta$ , repliement structural caractéristique de site nucléotidique (71, 72). Les structures cristallographiques humaine (17) et bovine (21) de 1998 ont néanmoins mis en évidence que la protéine ne comportait pas de motif  $\beta\alpha\beta$  complet.

La PEBP1 bovine de cerveau a été reportée pour se lier à des nucléotides car ceux-ci pouvaient éluer la protéine d'une colonne à affinité Blue Sepharose (73). Les nucléotides ont pu être classés selon capacité à éluer la protéine à pH 7.5 : FMN > GTP > GDP > GMP > FAD > ATP > NADP > CTP > UTP > ADP (73). L'AMP, l'AMPc, le GMPc et le NAD ont également été testés mais ils ne se liaient pas à la bPEBP1 (73).

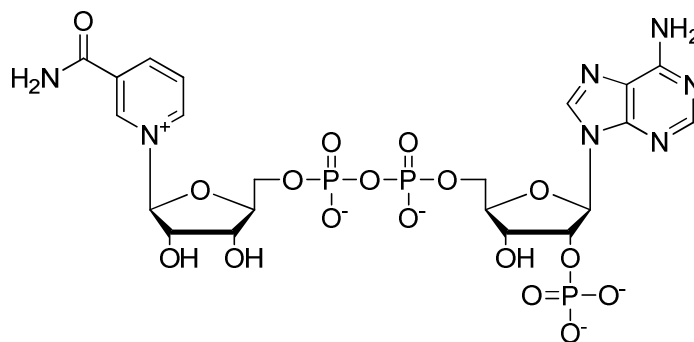




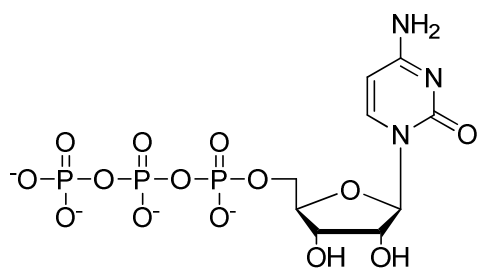
FMN



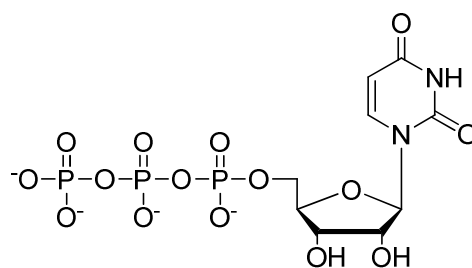
FAD



NADP<sup>+</sup>



CTP



UTP

Figure 16: Structures de nucléotides

La liaison directe du GTP à bPEBP1 a seulement été mesurée par autoradiographie lorsque la protéine était incubée en présence de petites protéines G monomériques (73), ce qui a mené à la conclusion que bPEBP1 devait spécifiquement interagir avec des petites protéines G.

## 2.3. Analogues de nucléotides dans les traitements contre le cancer

### 2.3.1. Analogues de nucléosides (sans phosphate)

Au début des années 60, des dérivés de nucléosides cytotoxiques furent identifiés comme des molécules potentiellement capables de jouer un rôle contre le cancer car elles pouvaient être considérées par l'organisme comme des antimétabolites interférant avec la croissance des cellules tumorales (74). Les analogues de nucléosides thérapeutiques cytotoxiques ressemblent beaucoup aux nucléosides naturels, ils subissent le même processus métabolique et sont incorporés dans le nouveau brin d'ADN ou d'ARN après avoir été phosphorylés, ce qui conduit à la terminaison de la chaîne et à l'apoptose (75). Malgré leur polarité, les analogues de nucléosides passent la membrane cellulaire grâce à des transporteurs de nucléosides (75).

Les analogues de nucléosides font partie des premiers agents chimiothérapeutiques à avoir été introduits pour le traitement du cancer. Les médicaments anti-cancéreux de type nucléosides les plus utilisés en clinique sont la capecitabine, la fludarabine, la cladribine la clofarabine, la gemcitabine, la troxacitabine et la cytarabine (Tableau 3 et Figure 17). La troxacitabine est la seule molécule à rentrer dans la cellule par diffusion passive (76).

Tableau 3 : Analogues nucléosidiques anticancéreux (77)

Analogues de nucléosides anticancéreux	Maladie
capecitabine, prodrogue de 5'-déoxy-5-fluorouridine	cancer colorectal métastatique (78), cancer du sein métastatique (79, 80)
cytarabine	leucémie myélogène aigüe, lymphome non-Hodgkinien (81)
fludarabine	lymphome de bas grade, leucémie lymphocytaire chronique (82)
cladribine	lymphome de bas grade (83), leucémie à tricholeucyte, leucémie lymphocytaire chronique (84)
clofarabine	leucémie myélogène aigüe infantile récidivants ou réfractaires (85)
gemcitabine	cancer pancréatique (86), cancer métastatique de la vessie (87, 88), cancer du poumon non à petites cellules (89)

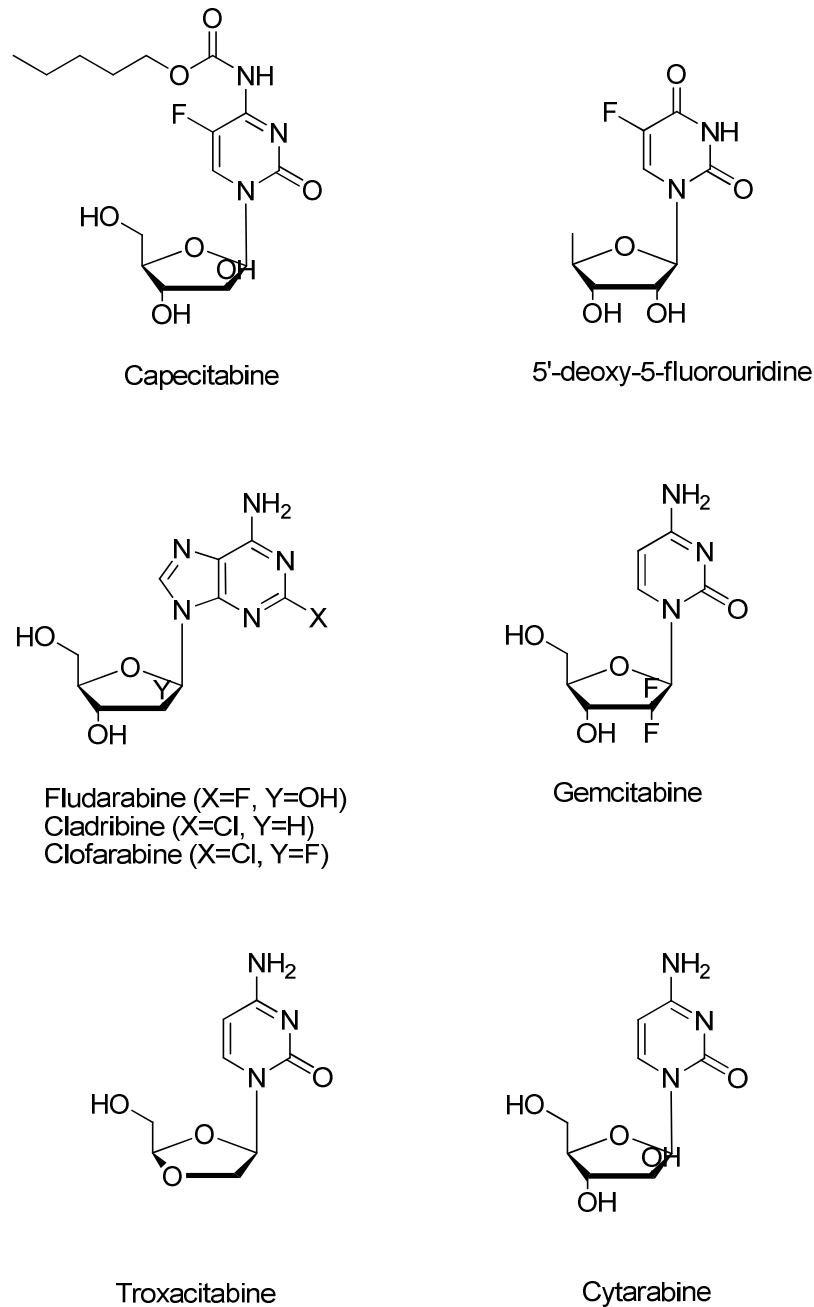


Figure 17: Structures des médicaments anti-cancéreux de type nucléosides

L'efficacité de ces médicaments n'est pas optimale car les analogues de nucléosides nécessitent d'être convertis en triphosphate pour présenter une activité thérapeutique. Cette réaction est cependant fortement limitée par la première étape de phosphorylation. Afin de régler ce problème de phosphorylation *in vivo* du médicament, les synthèses se sont rapidement tournées vers des molécules déjà phosphorylées : les nucléotides ou nucléosides phosphates.



### 2.3.2. Analogues de nucléosides phosphonates

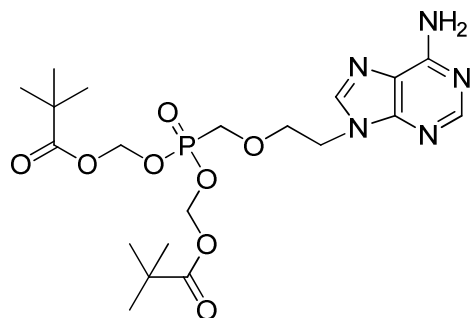
Bien que les nucléotides règlent le problème de la phosphorylation intracellulaire, un autre problème est apparu. Dans l'organisme, les phosphorylases hydrolysent les groupements phosphates des analogues de nucléotides qui deviennent à nouveau inactifs. Afin d'éviter cette hydrolyse, le groupement phosphate P-O-C- fût transformé en phosphonate P-C-O- non hydrolysable.

Des nucléosides acycliques phosphonates (ANPs) ont ensuite été développés et ont été principalement utilisés pour leurs propriétés antivirales. Les ANPs offrent une réponse antivirale plus longue que leurs équivalents phosphates car leur demi-vie intracellulaire est plus longue (74).

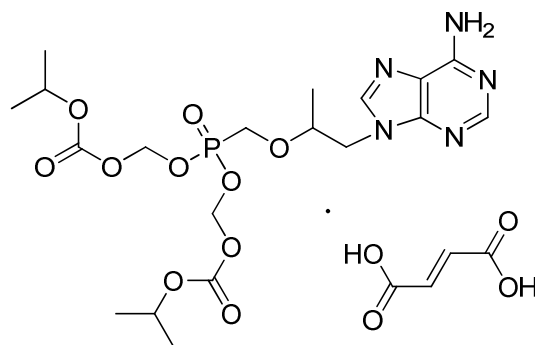
Les acides phosphoniques (phosphonates) sont rarement des molécules candidates pour être un médicament, malgré le potentiel de ce groupe fonctionnel à former des interactions de liaison unique avec une enzyme cible ou un récepteur. Le problème majeur d'un tel médicament est sa délivrance intracellulaire. Les acides phosphoniques portent une charge négative à pH physiologique, ce qui les rend très polaires. Cette forte polarité est sans doute à la base des difficultés de délivrance de ces molécules. Les espèces très polaires ne traversent pas la membrane plasmique par diffusion passive. Il existe néanmoins des transporteurs qui permettent de traverser la membrane de façon active: on connaît les transporteurs de nucléosides, et il pourrait aussi exister un transport actif d'analogues de nucléotides.

### 2.3.3. Prodrogues de nucléosides acycliques phosphonates

Pour neutraliser les groupements ionisables des phosphonates, des dérivations chimiques de type ester ont donc été mises en place. Les molécules dérivées forment alors des prodrogues. Pour deux médicaments, adefovir et tenofovir, le problème de biodisponibilité a donc été résolu en rajoutant des groupements protecteurs hydrophobes sur les fonctions hydroxyles du phosphonate et en formant ainsi des prodrogues : adefovir dipivoxil (Hepsera<sup>®</sup>) et tenofovir disoproxil fumarate (Viread<sup>®</sup>).



adefovir dipivoxil (Hepsera<sup>®</sup>)



tenofovir disoproxil fumarate (Viread<sup>®</sup>)

Figure 18 : Structures de l'Hepsera<sup>®</sup> et du Viread<sup>®</sup>

Les prodrogues libèrent facilement les médicaments par clivage des groupements protecteurs. Les travaux de Durand-Gasselin montrent que l'administration orale de la prodrogue tenofovir disoproxil fumarate (TDF) (Viread<sup>®</sup>) entraîne effectivement une meilleure distribution du tenofovir dans la circulation systémique (90).

En plus de propriétés antivirales et antibactériennes, adefovir dipivoxil (Hepsera<sup>®</sup>) présente une activité antitumorale. Il inhibe de façon dose-dépendante la croissance du chorio-carcinome chez le rat (91). Cidofovir (Vistide<sup>®</sup>), un ANP non-prodrogue, est également un agent antitumoral chez plusieurs modèles animaux : le carcinome nasopharyngé et le carcinome humain cervical exogreffé sur des souris nues athymiques (92–94). L'effet d'inhibition de la croissance tumorale de ce médicament pourrait être attribué à une induction de l'apoptose (74).

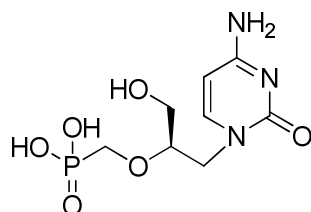


Figure 19 : Structure du Cidofovir (Vistide<sup>®</sup>)

### 3. Interactions protéine-protéine

Nous avons vu précédemment que PEBP1 interagit avec de nombreuses kinases. Nous allons détailler l'interaction de PEBP1 avec Raf-1. La liaison de la locostatine à PEBP1 inhibe l'interaction de PEBP1 avec Raf 1 (68). L'interaction de rPEBP1 avec Raf-1 a été étudiée par Park et ses collaborateurs (95). En combinant des approches mutationnelles, des études biochimiques, de la synthèse peptidique et de la résonance plasmonique de surface (SPR), ils ont mis en évidence que la région de Raf-1, qui comprend les résidus 331 à 349, était suffisante pour l'interaction avec rPEBP1 (95).

RPRGQRDpSpSYpYWEIEASEV

Figure 20: Séquence du peptide P3P phosphorylé

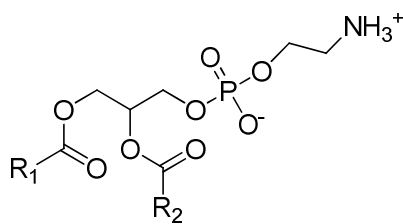
rPEBP1 se lie avec des constantes de dissociation de l'ordre du micromolaire au peptide 331-349 phosphorylé (Figure 20), alors que le peptide non-phosphorylé se lie avec une affinité plus faible, de l'ordre du millimolaire. La phosphorylation des sérines 338 et 339 et de la tyrosine 341 semble moduler l'affinité de l'interaction (95). Ces résultats peuvent paraître contradictoires avec ceux de Trakul et ses collaborateurs (96), qui montraient que la PEBP1 inhibait Raf-1 en empêchant la phosphorylation de la sérine 338 et de la tyrosine 341. Néanmoins, Park et ses collaborateurs soulignent que les deux résultats ne sont pas forcément contradictoires. Ils font l'hypothèse que la dissociation initiale de PEBP1 de Raf1 serait engendrée par la phosphorylation de PEBP1 sur la sérine 153, mais que la reliaison de PEBP1 à Raf1 nécessiterait ensuite la phosphorylation de Raf1. La reliaison de PEBP1 à Raf1 phosphorylée protégerait probablement ces sites de phosphorylation et permettrait ainsi de prolonger l'activation de Raf1 (95).

Récemment, Brady *et al.* ont trouvé par cristallographie aux rayons X que l'O-phosphotyrosine (pTyr) se liait à la poche à anion très conservée (PDB 2QYQ) (24). Cette structure est la première description d'une interaction qui fournit un modèle pour comprendre comment un ligand de type peptidique se lie dans la poche.

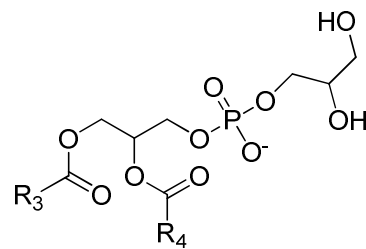
## 4. Interactions avec des lipides membranaires ou des molécules hydrophobes

La chromatographie d'affinité lipidique a été utilisée pour démontrer la liaison de stéroïdes ainsi que d'un composant de la membrane lipidique, la phosphatidyléthanolamine, à bPEBP1 (Figure 21) (6, 7). Des études mirent en évidence la présence d'un site saturable pour la phosphatidyléthanolamine, tandis que la phosphatidylcholine se liait de façon non spécifique (7).

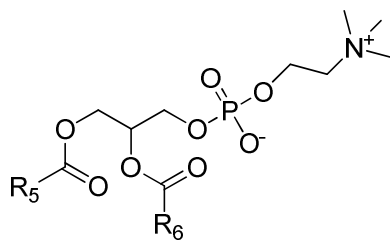
De plus, la poche de liaison anionique partage des homologues avec les domaines de liaison de phospholipides. bPEBP1 s'associe avec la membrane phospholipidique *via* des interactions électrostatiques avec des microdomaines chargés négativement à la surface de la membrane plutôt qu'en s'insérant dans la membrane (10).



Phosphatidyléthanolamine



Phosphatidylglycérol



Phosphatidylcholine

Figure 21: Structure de la phosphatidyléthanolamine, du phosphatidylglycérol et de la phosphatidylcholine

D'autres études ont montré que bPEBP1 se liait uniquement aux phosphatidylglycérol chargés négativement des liposomes et aux monocouches lipidiques (10) (Figure 21).

Les deux régions terminales de la PEBP1 bovine s'intègrent de la même façon à des membranes artificielles que la bPEBP1 entière, suggérant qu'elles peuvent participer à l'interaction de bPEBP1 avec la membrane (10). L'affinité entre la bPEBP1 et la PE ( $K_D = 10 \mu\text{M}$ ) est jugée plutôt faible par les auteurs (10). La région C-terminale de bPEBP1 est une hélice  $\alpha$  amphiphile localisée près de l'entrée de la poche à ligand. Cette hélice pourrait ancrer bPEBP1 à la membrane et protéger la poche à anion ce qui entrainerait l'exposition des résidus du site actif au solvant (10). Un tel changement structural serait consistant avec la flexibilité de la PEBP1 bovine décrite par Vallée *et al.* en 1999 (15).

D'après une étude RMN menée sur la PEBP1 de rat, la poche pourrait accueillir divers ligands tels que le 1,2-dihexanoyl-sn-glycéro-3-phosphoéthanolamine (DHPE), dihexanoylphosphatidylserine (DHPS), le dihexanoylphosphatidylglycerol (DHPG) et le dihexanoylphosphatidic acid (DHPA) (23) (Figure 22).

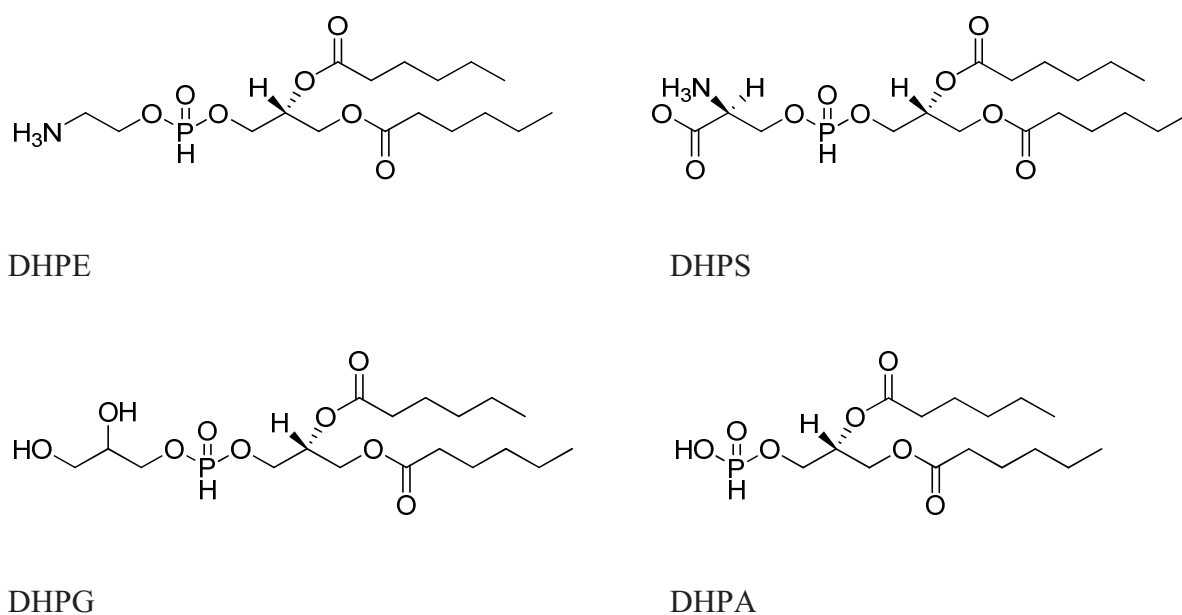
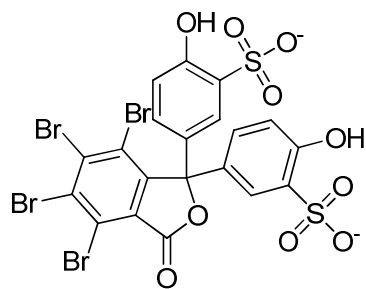
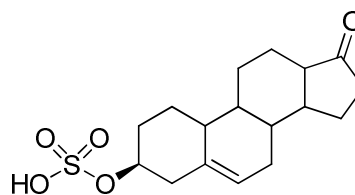


Figure 22: Structure du DHPE, DHPS, DHPG et DHPA

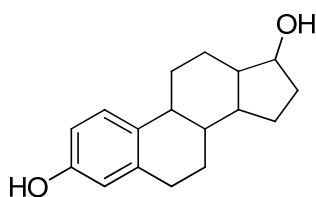
Durant la purification de bPEBP1, il apparut rapidement que la protéine restait associée à des lipides endogènes (triacylglycérides et esters de cholestérol) (6). L'interaction avec des molécules hydrophobes fut donc testée et un début de caractérisation montra qu'elle était capable de lier de manière non-covalente des anions organiques, la bromosulfophthaléine, le sulfate de déhydroépiandrostérone et une hormone, l'oestradiol-17 $\beta$  (Figure 23) avec des  $K_D$ s respectifs de 1, 100 et 333  $\mu\text{M}$  (7).



bromosulphthaléine



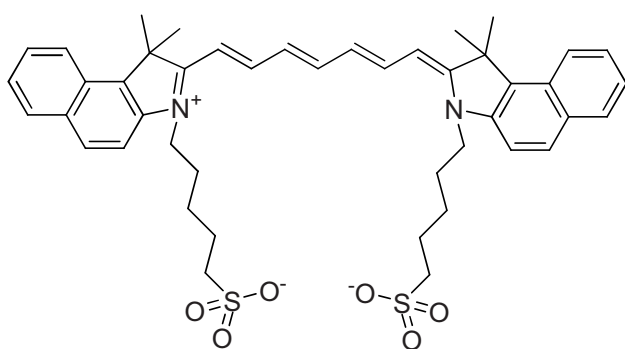
DHEA sulfate



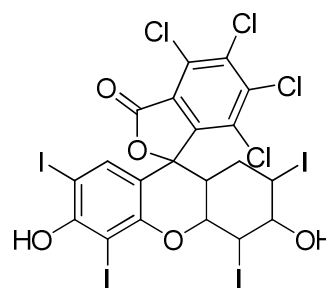
oestradiol-17β

Figure 23 : Structure du bromosulphthaléine, le sulfate de déhydroépiandrostérone (DHEA) et de l'oestradiol-17β

Au contraire, elle ne liait pas l'indocyanine, le rose du Bengal et le bleu Evans (Figure 24).



indocyanine



rose du Bengal

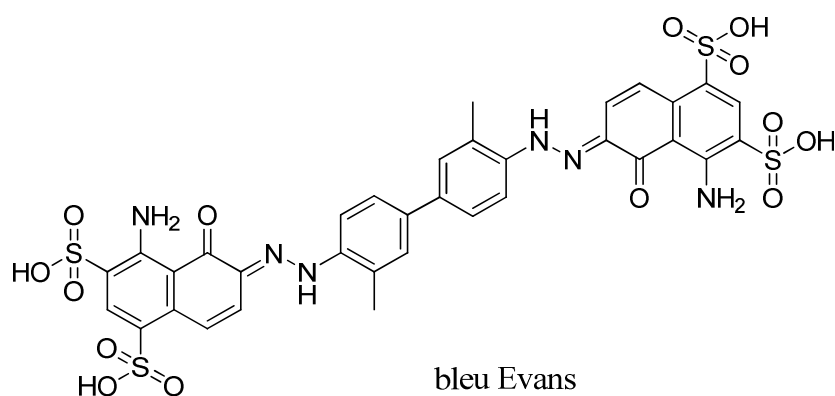


Figure 24: Structure de l'indocyanine, du rose du Bengal et du bleu Evans

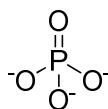
D'après les structures des différents ligands exposés précédemment, nous pouvons essayer de déterminer les fonctions qui semblent jouer un rôle dans l'interaction avec PEBP1. De nombreux ligands lipidiques portent un groupement phosphate, il est donc fort probable que cette fonction joue un rôle particulier dans l'interaction. Une molécule non chargée peut interagir avec bPEBP1, ce qui signifie qu'une interaction hydrophobe est suffisante pour permettre une liaison avec la protéine. De plus, en dehors des lipides membranaires, les molécules hydrophobes se liant à PEBP1 présentent toutes des multicycles accolés. Cette particularité n'est cependant pas suffisante pour la liaison.

## 5. Interaction avec des molécules tampon

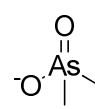
Les études de cristallographie aux rayons X ont montré une séquence hautement conservée dans la poche capable de lier des anions, tels que l'acétate, le phosphate (21) et le cacodylate (17) (Figure 25). Cette poche est la seule marque clairement identifiable d'un potentiel site de liaison au sein de l'architecture de la PEBP.



acétate



phosphate



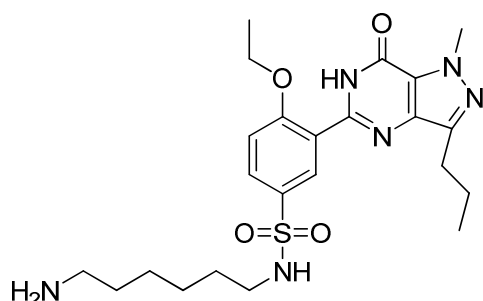
cacodylate

Figure 25: Structure de l'acétate, du phosphate et du cacodylate

## 6. Autres interactions

### 6.1. Avec un inhibiteur de phosphodiesterase-5 (PDE5)

L'interaction de PEBP1 (hPEBP1, mPEBP1 et 2, rPEBP1) avec un inhibiteur de phosphodiesterase-5 (PDE5) (Figure 26) a été mise en évidence en combinant un enrichissement basé sur l'affinité et de la spectrométrie de masse (97, 98). L'interaction a été confirmée par spectrophotométrie d'absorption UV-visible, par spectrofluorimétrie et par spectroscopie RMN.



inhibiteur de PDE5

Figure 26 : Structure de l'inhibiteur de PDE5 (97)

### 6.2. Avec des dérivés morphiniques

En 1990, Grandy *et al.* ont observé que rPEBP1 se liait à des molécules dérivées de morphine, suggérant son implication comme protéine liant la morphine (99). Dans des cellules chromaffines bovines de la matrice chromaffine intragranulaire, la morphine-6-glucuronide (M6G) (Figure 27) est un ligand endogène de bPEBP1 (11). Atmanène *et al.* ont observé que les PEBP1 humaines et bovines interagissaient avec des morphine-glucuronides mais pas avec la morphine (100).



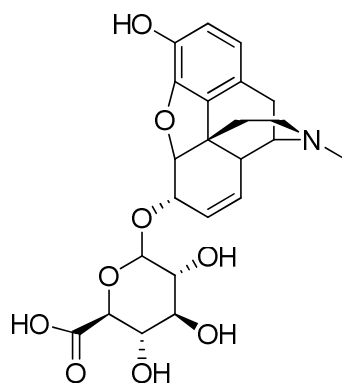
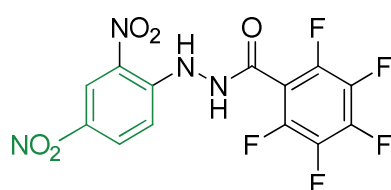


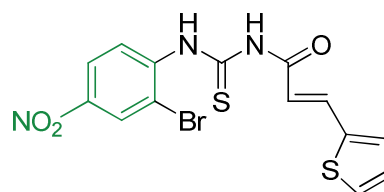
Figure 27: Structure de la morphine-6-glucuronide (M6G)

### 6.3. Avec des composés portant un groupement 4-nitrophényl

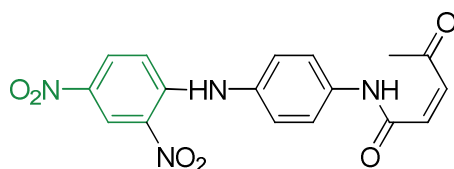
Un criblage par RMN de liaison à rPEBP1 mené sur environ 6000 composés (101) a permis de découvrir trois nouvelles molécules: N'-(2,4-dinitrophenyl)-2,3,4,5,6-pentafluorobenzohydrazide (composé n°26 dans la publication citée), N-{(2-bromo-4-nitrophenyl)amino]carbonothioyl}-3-(2-thienyl)acrylamide (composé 48) et 4-({4-[(2,4-dinitrophenyl)amino]phenyl} amino)-4-oxo-2-butenoic acid (composé 98) (Figure 28). Ces trois molécules portent un groupement 4-nitrophényl substitué en position 2.



Composé 26



Composé 48



Composé 98

Figure 28 : Structures des composés du criblage de Shemon *et al.* (101).

Contrairement au DHPE, les trois ligands n'affectaient ni la liaison de rPEBP1 à Raf-1, ni la phosphorylation de la sérine 153 de rPEBP1. Le composé 48 montre des effets cellulaires: il détériore la voie MAPK induite par EGF et l'activité métabolique (101).

## 7. Structure connue d'interaction avec des ligands : la poche à anion

### 7.1. Description de la poche

La poche à anion interagit avec la tyrosine phosphorylée (24), avec le cacodylate (17), l'acétate, le phosphate et l'O-phosphoryléthanolamine (21) (Figure 25 et 29).

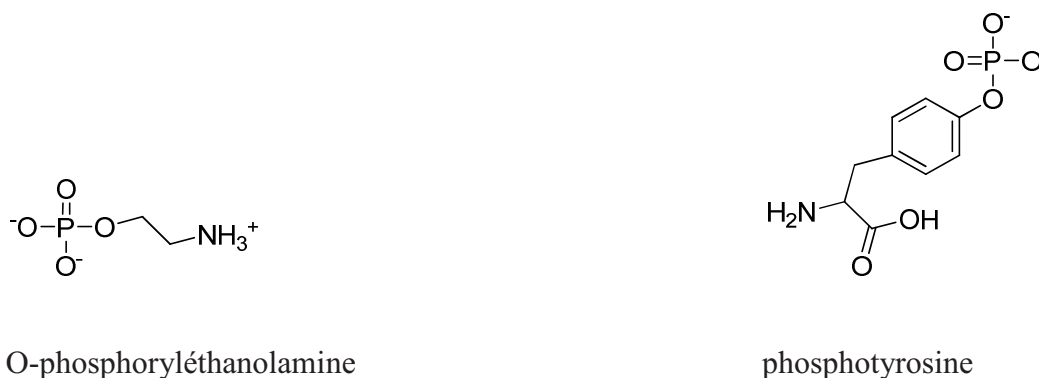


Figure 29 : Structures de l'O-phosphoryléthanolamine et de la phosphotyrosine

Le fond de la poche est composé de résidus capables de former des liaisons hydrogènes avec les anions. De plus, la poche est bordée de résidus aromatiques qui font que les parois de la poche sont souvent plus hydrophobes et peuvent s'engager dans des interactions de type van Der Waals avec les noyaux aromatiques des ligands amphiphiles comme la phosphotyrosine (95). Les interactions protéine-protéine n'impliquent vraisemblablement pas uniquement la poche à anion, mais la poche pourrait constituer une partie de l'interface d'interaction protéique (23, 102).

## 7.2. Rôles physiologiques de la poche chez les plantes et la levure

Les structures cristallographiques des protéines PEBPs depuis la bactérie jusqu'à l'*Homo sapiens* ont révélé une poche à anion remarquablement conservée (19). Le rôle physiologique de cette poche est illustré par des études sur les PEBPs de plante et de levure. Chez l'homologue de PEBP1, CEN, la mutation du motif conservé DPDxP au sein de la poche (l'équivalent de P74L) entraîne chez la tomate, lors du développement, une transition de la croissance racinaire vers la floraison (103). L'homologue chez *Saccharomyces cerevisiae*, Tfs1p, fonctionne comme une protéine inhibitrice de RasGAP, ce qui conduit à une hyperactivation de la Ras de levure, à l'activation de l'adenylyl cyclase, et à l'activation par augmentation d'AMPc, de la protéine kinase A. La mutation correspondant à P74L dans la poche annule l'interaction entre Tfs1p et RasGAP, et détruit ainsi la capacité de Tfs1p à réguler RasGAP (104).

La mutation P74L a également été étudiée sur la PEBP de rat. La mutation augmente la phosphorylation de la protéine et entraîne donc une activation de la voie Raf/MEK/ERK (23).

Dans l'étude de Bucquoy *et al.*, les résultats d'élution de colonne indiquent que les sites de liaison pour nucléotides et pour les lipides sont indépendants (73). En effet, en présence de lipides, le profil d'élution n'est pas affecté.

## 7.3. Plasticité conformationnelle et rôles possibles

D'après les résultats de Granovsky *et al.* (23), il semble que la structure et la dynamique de la poche lui permette d'intégrer deux signaux de régulation, la phosphorylation et la liaison d'un ligand.

Selon des analyses structurales réalisées par RMN, la poche est capable de changer de forme (23). En RMN, la forme des pics peut être utilisée pour déterminer la mobilité relative de certains résidus par rapport à d'autres. Granovsky et ses collaborateurs ont ainsi mis en évidence que l'hélice  $\alpha$  C-terminale et la boucle 127-150 étaient des régions flexibles sur une échelle de temps comprise entre la microseconde et la milliseconde (23). La mobilité de ces résidus est moins forte lorsqu'un ligand comme le DHPE est lié à la protéine.

Cette plasticité de conformation pourrait expliquer que la poche soit capable de lier des ligands chimiquement très différents, tels que le lipide DHPE et la protéine Raf-1. La

plasticité conformationnelle de la poche est sans doute nécessaire pour former différentes sortes d'interactions. La capacité de changer de conformation est la base générale de l'allostérie. DHPE inhibe non-compétitivement la phosphorylation de la tyrosine 42 de rPEBP1 catalysée par ERK2, ce qui suggère un couplage allostérique entre la liaison d'un ligand dans la poche et les interactions PEBP1 – kinases (23). Une hypothèse possible serait que les kinases-cibles de l'inhibition par PEBP1 se lient directement à la poche, ou alors que la ou les kinase(s) capable(s) de phosphoryler PEBP1 ne soi(en)t sensibles qu'à la conformation et à la flexibilité de la poche, *via* la liaison distante d'un ligand, sans se lier elle(s)-même(s) à la poche. La poche pourrait ainsi agir comme un senseur et/ou un régulateur des fonctions de PEBP1.

## Section 2 : La spectrométrie de masse supramoléculaire

*La spectrométrie de masse est une technique d'analyse qui permet d'étudier les interactions non-covalentes hPEBP1-ligand. Nous allons détailler les principes, les applications et les limites de la spectrométrie de masse supramoléculaire.*

### I. Les interactions non-covalentes impliquées dans les complexes supramoléculaires

Les interactions non-covalentes s'opposent clairement aux liaisons covalentes. La liaison covalente est une liaison chimique dans laquelle chacun des atomes liés met en commun au moins un électron d'une de ses couches externes afin de former un doublet d'électrons liant les deux atomes. Au contraire, l'interaction non-covalente n'implique pas de partage d'électrons mais des variations d'interactions électromagnétiques. Les interactions non-covalentes sont essentielles pour maintenir la structure tridimensionnelle des macromolécules. Elles maintiennent liés les deux brins d'ADN en une double hélice, elles stabilisent les structures secondaires et tertiaires des protéines, permettent les interactions de type enzyme-substrat, anticorps-antigène, protéine-ligand de manière plus générale. Elles sont impliquées dans de nombreux processus biologiques dans lesquels des macromolécules se lient spécifiquement mais transitoirement à d'autres molécules. Les interactions non-covalentes prises individuellement impliquent des énergies plutôt faibles. En effet, l'énergie libérée par la formation d'une interaction non-covalente est de l'ordre de 4 à 21 kJ par mole. Ces interactions travaillent de manière coopérative pour avoir globalement un effet significatif, et la combinaison de leurs forces est supérieure à la simple somme de ces forces. Il existe quatre types d'interactions non-covalentes communément admises:

- les interactions électrostatiques:
  - o interactions ioniques
  - o les forces de van der Waals
  - o les liaisons hydrogène
- les interactions issues de l'effet hydrophobe.

## 1. Interactions ioniques

L'interaction ionique (parfois dénommée liaison ionique ou pont salin) est une interaction entre deux groupements portant chacun une charge nette positive ou négative (soit au sein d'une molécule soit entre deux molécules). Cette interaction est la plus forte des interactions non-covalentes. Ces interactions peuvent conduire à un effet attractif (si charges opposées) ou répulsif (si charges de même signe). L'énergie d'interaction peut atteindre quelques centaines de kilojoules par mole, ce qui est assez proche de l'énergie d'une liaison covalente. L'origine de ces interactions est l'attraction ou la répulsion coulombienne, donnée par la loi de Coulomb :

$$F = \frac{q_1 \cdot q_2}{4 \pi \epsilon_0 \epsilon_r d^2}$$

où  $q_1$  et  $q_2$  sont les deux charges,  $d$  la distance entre les charges,  $\epsilon_0$  la permittivité du vide et  $\epsilon_r$  la permittivité relative du milieu. Dans l'eau,  $\epsilon_r$  est de 80 alors que dans le vide elle est de 1. Il est important de prendre en compte cette variation de permittivité suivant le milieu lors de l'analyse des interactions non-covalentes par spectrométrie de masse. En effet, dans l'instrument, les analytes passent d'un solvant aqueux ( $\epsilon_r = 80$ ) au vide ( $\epsilon_r = 1$ ). Les interactions ioniques sont donc fortement renforcées par cette désolvatation intrinsèque à l'analyse.



Figure 30: Illustration de l'interaction ionique

## 2. Forces de Van Der Waals

Une molécule n'a pas besoin d'une charge nette pour participer à des interactions électrostatiques : la présence d'un moment dipolaire permanent suffit à faire interagir les molécules polaires via des liaisons non-covalentes.

Les interactions entre molécules neutres se forment à partir d'interactions électrostatiques entre dipôles permanents et/ou dipôles induits. Les forces résultantes sont responsables d'interactions multiples entre atomes voisins non-liés et sont appelées communément forces de van der Waals. L'énergie potentielle typique de ce type d'interaction est de quelques kJ par mole. On englobe sous le nom de forces de van der Waals trois types d'interactions moléculaires [d, e] :

- forces entre deux dipôles permanents : forces de Keesom
- forces entre un dipôle permanent et un dipôle induit : forces de Debye
- forces entre deux dipôles induits : forces de dispersion de London

## 2.1. Forces de Keesom

Les forces de Keesom résultent d'une interaction intermoléculaire entre dipôles permanents. Lorsque, dans une molécule neutre, les barycentres des charges positives et des charges négatives ne coïncident pas, on dit que la molécule est polaire, donc qu'elle possède un moment dipolaire.

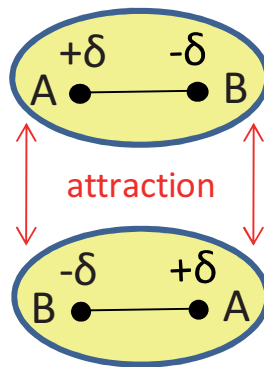


Figure 31: Illustration de la force de Keesom

Deux molécules polaires peuvent trouver des positions favorables pour maximiser l'attraction entre elles.

## 2.2. Forces de Debye

Les forces de Debye sont des forces intermoléculaires résultant de l'interaction entre un dipôle permanent et un dipôle induit. Une molécule polaire induit une déformation du nuage électronique des molécules apolaires voisines sous l'effet de son champ électrique. Cette déformation du nuage électronique est caractérisée par la polarisabilité de la molécule apolaire, qui augmente avec son nombre d'électrons.

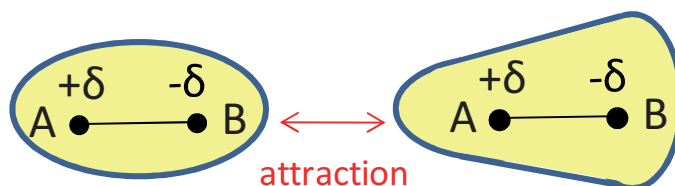


Figure 32: Illustration de la force de Debye

La molécule apolaire acquiert un moment dipolaire induit qui interagit avec le dipôle permanent de la première molécule.

## 2.3. Forces de London

Les forces de dispersion de London sont des forces faibles intermoléculaires créées par des dipôles induits. Les forces de London existent du fait que la densité électronique des molécules est probabiliste : il y a une forte chance à tout moment pour que celle-ci ne soit pas équitablement répartie à travers la molécule, ce qui crée un léger moment dipolaire. Les moments dipolaires vont varier très rapidement au cours du temps. Ainsi à chaque distribution inhomogène se crée un moment dipolaire induit qui peut interagir avec les moments dipolaires induits des molécules voisines et une force s'exerce entre les molécules. Naturellement, ces dipôles induits peuvent également réagir avec un dipôle permanent (molécule polaire) mais les forces qui en résultent sont alors les forces de Debye. L'effet apparaît également dans les molécules polaires, mais c'est alors les forces de Keesom qui prennent le dessus le plus souvent et les forces de London jouent alors un rôle souvent minime dans les interactions entre dipôles permanents.



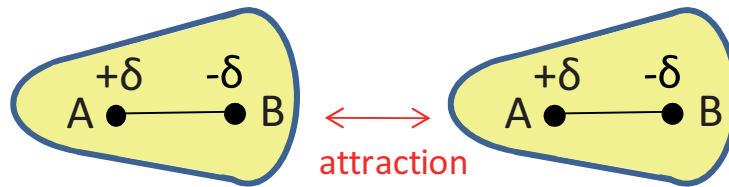


Figure 33: Illustration de la force de London

De manière générale, pour des molécules polaires, les forces de London sont très faibles devant les autres forces intermoléculaires, telles que les interactions ioniques, les liaisons hydrogène ou les forces de Keesom. Néanmoins, pour des molécules neutres, les forces de London sont la seule force attractive intermoléculaire.

#### 2.4. Cas particulier des forces de van der Waals dans les protéines

Au sein des protéines, les acides aminés portant un cycle aromatique (tryptophane, phénylalanine, tyrosine) peuvent interagir et former deux types d'arrangement:  $\pi$ -stacking et T-stacking.

Le  $\pi$ -stacking ou interaction  $\pi$ - $\pi$  est du à un recouvrement intermoléculaire d'orbitales p dans des systèmes  $\pi$ -conjugués (Figure 34).

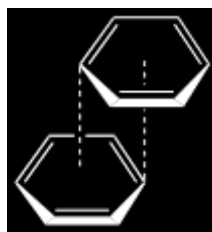


Figure 34: Interactions ( $\pi$ - $\pi$ )

Le T-stacking est souvent observé dans les protéines où un atome hydrogène partiellement chargé d'un système aromatique pointe perpendiculairement au centre d'un plan aromatique

d'un autre système aromatique (Figure 35). Cette interaction est aussi appelée "edge-face stacking".



Figure 35: Interactions T-stacking

### 3. Liaisons hydrogènes

L'origine de la liaison hydrogène est essentiellement électrostatique et de type dipôle-dipôle induit. Un hydrogène porté par un atome fortement électronégatif va porter une fraction de charge positive qui polarisera fortement une autre molécule possédant un doublet non-lié (Figure 36).

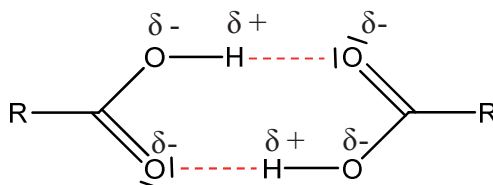


Figure 36: Illustration des liaisons hydrogènes

L'interaction forte entre le dipôle et le dipôle induit entraîne leur alignement et une grande proximité des atomes considérés par rapport aux interactions de type van der Waals. La liaison hydrogène est ainsi dirigée. Ces liaisons peuvent apparaître entre des molécules (intermoléculaires) ou avec différentes parties d'une même molécule (intramoléculaires). La liaison hydrogène est en partie responsable de la structure secondaire, tertiaire et quaternaire

des protéines et des acides nucléiques. En effet, pour la structure secondaire des protéines, les liaisons hydrogène sont à la base de la structuration en hélice alpha et en feuillet beta, par des interactions entre des groupes carbonyle (C=O, donneurs) et des N-H d'amide (accepteurs). De nombreux autres groupes chimiques des chaînes latérales sont également susceptibles de servir de donneurs (-OH, -NH, plus rarement -SH) ou d'accepteurs (O, N et de façon plus rare S). Bien que de la même famille que les forces de van der Waals, les liaisons hydrogène s'en distinguent par leur intensité : l'énergie de liaison de la liaison hydrogène est environ dix fois supérieure à celle des forces de van der Waals (20 à 30 kJ/mol).

#### 4. Interactions résultant de l'effet hydrophobe

L'effet hydrophobe est basé sur la tendance de l'eau à exclure les molécules non-polaires, ce qui conduit à une ségrégation de l'eau et des molécules non-polaires. L'eau pousse au regroupement des molécules apolaires entre elles, reliées par des liaisons de type van der Waals. Il est important de noter que la plus grande partie de l'énergie d'interaction est due non pas à ces interactions de van der Waals (qui sont faibles) mais aux contributions défavorables à la solvation de ces molécules apolaires dans l'eau. En absence d'eau, ce qui est le cas dans le vide du spectromètre de masse, seule la part d'énergie d'interaction due aux interactions de van der Waals subsiste. Les interactions hydrophobes se retrouvent donc très affaiblies par le passage en phase gazeuse. Les interactions hydrophobes sont par exemple impliquées dans le regroupement d'acides aminés apolaires conduisant à leur enfouissement au cœur de la protéine, ou à leur localisation stable au sein d'une bicouche lipidique membranaire.

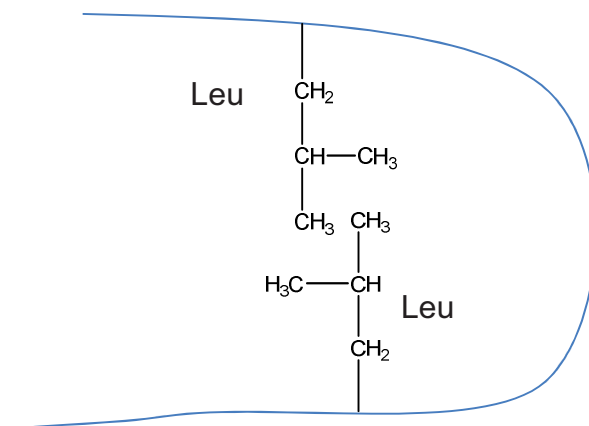


Figure 37: Illustration de l'effet hydrophobe : interaction entre les chaînes latérales non-polaires de leucines.

## II. Etude d'interactions non-covalentes par spectrométrie de masse

### 1. Historique

L'approche MS classique, appelée MS moléculaire, analyse en phase gazeuse les espèces individuelles qui étaient initialement présentes en solution, et cela après destruction de leurs éventuels assemblages non-covalents. La spectrométrie de masse supramoléculaire, ou MS native, s'effectue en conditions non dénaturantes et a pour objectif de garder intacts lors du transfert en phase gazeuse des complexes non-covalents qui préexistaient en solution.

L'étude de complexes non-covalents par MS peut apparaître comme une application originale et inattendue de la MS dans le domaine biologique. En effet, l'utilisation d'une technique qui détecte des espèces en phase gazeuse pour étudier des assemblages maintenus par des interactions non-covalentes décrites plus haut, pouvait sembler contre-intuitive en raison de la fragilité intrinsèque de certaines de ces interactions.

Dans les années 1990, deux groupes américains (122) ont mené des études pionnières sur un complexe récepteur/ligand et sur l'interaction globine/hème de la myoglobine, respectivement, et montré que des interactions non-covalentes spécifiques protéine/ligand peuvent survivre au processus de l'ionisation ESI. Bien que la MS non-dénaturante reste un domaine d'expertise d'un petit nombre de laboratoires, le nombre de publications relatant l'utilisation d'ESI-MS pour des complexes non-covalents (protéine/protéine, protéine/ligand, protéine/métal, protéine/ARN, protéine/ADN, ...) croît de façon exponentielle (105–110).

Comparée à des méthodes biophysiques plus classiques comme la spectrophotométrie, des techniques de fluorescence, la cristallographie, la RMN ou la SPR, la MS non-dénaturante est maintenant bien implantée comme une technique complémentaire pour caractériser des interactions protéine/ligand ou protéine/protéine. L'avantage le plus intéressant de la MS par rapport aux autres techniques biophysiques est sa capacité à fournir directement des renseignements sur chaque espèce individuelle présente en solution grâce à la mesure d'une masse précise. Au final, la MS non-dénaturante fournit des données informatives sur la stoechiométrie (en conditions de titrage), la spécificité et l'affinité d'un complexe formé en solution. Nous discuterons cependant de l'approche à utiliser pour obtenir avec exactitude ces

informations à partir de l'interprétation du spectre de masse ESI obtenu en conditions non dénaturantes.

Deux méthodes d'ionisation douce peuvent être utilisées pour l'analyse de macromolécules biologiques: le MALDI, ou matrix-assisted laser desorption/ionization (Tanaka *et al.*, 1988 et Karas *et al.*, 1988) (111, 112) et l'ESI (Fenn *et al.*, 1989) (113).

Le MALDI implique l'utilisation d'une matrice spécifique, c.-à-d. une petite molécule qui possède une forte absorption à la longueur d'onde du laser. Les matrices communément utilisées pour les protéines sont des dérivés d'acide cinnamique ou d'acide benzoïque, qui sont des acides faibles. Les solutions de matrice utilisées pour la préparation de l'échantillon contiennent généralement des solvants organiques ainsi que 0,1 à 1% d'acide fort, susceptible de rompre les interactions non-covalentes. La première étape de l'analyse MALDI consiste à diluer la solution d'analyte dans cette solution de matrice. Par conséquent, les interactions non-covalentes sont potentiellement détruites avant même l'étape de co-cristallisation matrice-analyte. De plus, la désorption laser de type MALDI, bien que considérée comme douce par rapport aux méthodes d'ionisation plus anciennes, crée une plume hautement énergisée peu favorable au maintien des interactions non-covalentes. Quelques études ont cependant rapporté la détection de complexes non-covalents avec une source MALDI sous des conditions très spécifiques : seul le spectre obtenu par le premier tir laser sur la couche supérieure de l'échantillon est enregistré. Ce phénomène a été appelé « first shot phenomenon » (114–117). Dans ces conditions spécifiques, les spectres accumulés ont montré des signaux correspondant à des complexes non-covalents. Cette méthode, outre qu'elle est plus laborieuse qu'une analyse MALDI classique, est également controversée, en raison de la possibilité reconnue d'observer des complexes non-spécifiques à forte fluence laser en analyse normale.

Dans le cas de l'analyse de complexes non-covalents, la meilleure méthode d'ionisation est l'ESI car elle permet de travailler avec des échantillons liquides. Sous réserve d'utilisation de tampons compatibles avec la MS et avec la conservation de la forme native de la protéine, tels que les tampons ammonium, les analytes peuvent donc être transférés de la solution à la phase gazeuse de la façon de la plus douce qui soit pour préserver les interactions non-covalentes. Nous allons détailler ce mode d'ionisation.

## 2. Spectromètre de masse à ionisation électrospray

### 2.1. Schéma général d'un spectromètre de masse

Le spectromètre de masse peut être divisé en quatre régions :

- la source, où s'effectue l'ionisation des molécules
- l'interface, qui complète la désolvatation et assure le transfert des ions de la pression atmosphérique au vide
- l'analyseur, où s'effectue la séparation des ions selon leur rapport  $m/z$
- le détecteur, où s'effectue la mesure du compte d'ions, qui constitue le signal
- le numériseur qui assure l'enregistrement des signaux en fonction du temps

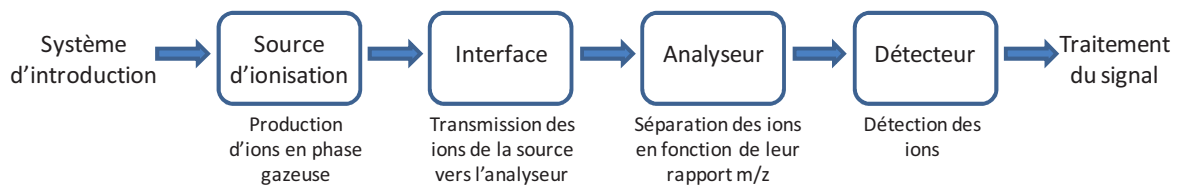


Figure 38: Représentation schématique d'un spectromètre de masse

### 2.2. Principe de l'ionisation électrospray

Le couplage ESI/MS a été développé par Fenn (118) et Aleksandrov (119) en 1984. En 1988, Fenn l'applique à l'étude de macromolécules chargées (113).

L'ionisation par électrospray consiste à générer des ions en phase gazeuse à partir d'un liquide soumis à un champ électrique d'environ 4000 V. On peut décrire ce processus par trois événements majeurs :

- la production de gouttelettes chargées à partir du liquide émergeant à l'extrémité de l'aiguille
- la réduction des gouttelettes par un processus réitéré d'évaporation et de fission des gouttelettes chargées par explosions coulombiennes
- le passage des ions en phase gazeuse à partir des gouttelettes chargées par un mécanisme encore débattu (deux modèles existants)

### 2.2.1. La production de gouttelettes chargées

Une solution d'électrolytes est introduite dans une aiguille sur laquelle est appliqué un champ électrique intense. Ce champ provoque la séparation électrophorétique des charges positives et négatives en solution afin d'induire un champ électrique contraire au champ imposé à la pointe de l'aiguille. En mode d'ionisation positive, l'aiguille métallique constitue l'anode (oxydation), à cet endroit des charges négatives sont neutralisées par oxydation de l'électrolyte, du solvant ou de l'aiguille. Les charges positives s'accumulent à la pointe de l'aiguille et à la surface du liquide. Pour un champ électrique critique, le liquide prend la forme d'un cône mieux connu sous le nom de « cône de Taylor » où il y a formation de gouttelettes enrichies en ions positifs.

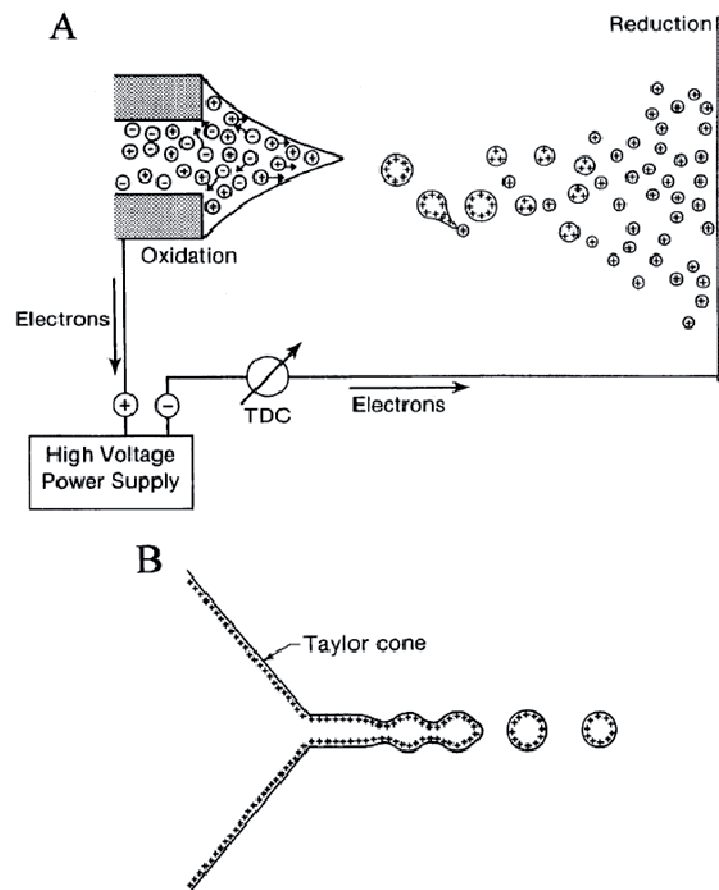


Figure 39 : Illustration de la production de gouttelettes chargées par ESI. Reproduction de la publication de Kebarle (120)

La technique d'ionisation électrospray a été miniaturisée en 1994 par Wilm et Mann (121, 122), on parle de « nanoélectrospray ». Ils ont utilisé des aiguilles avec un diamètre plus étroit. Les gouttelettes générées par nano-ESI sont environ dix fois plus petites que celles

obtenues par ESI classique. Ainsi la nano-ESI est plus efficace et par conséquent augmente la sensibilité. Cela permet de travailler avec des débits d'injection réduits qui autorisent des durées d'analyse plus longues et limitent les quantités d'échantillon consommées. En contrepartie, ces débits ne permettent pas de rincer la ligne d'infusion rapidement.

### 2.2.2. Réduction des gouttelettes: l'explosion coulombienne

Les gouttelettes chargées produites par électrospray se retrouvent dans une atmosphère chauffée. L'énergie thermique de l'air ambiant va entraîner l'évaporation progressive des molécules de solvant des gouttelettes qui subissent ainsi une diminution de leur volume. La diminution du rayon de la gouttelette a lieu jusqu'à ce que la limite de stabilité de Rayleigh soit atteinte, c'est-à-dire au moment où les forces de répulsion électrostatique sont égales aux forces de tension de surface (120, 123).

$$Q = 8\pi \sqrt{\epsilon_0 \gamma R_R^3}$$

$R_R$  : le rayon critique de Rayleigh

$Q$  : la charge de la gouttelette

$\epsilon_0$  : la permittivité du vide

$\gamma$  : la tension de surface

Équation 1 : Relation entre le rayon critique de Rayleigh et la charge  $Q$  de la gouttelette (124)

La gouttelette devient ensuite instable et forme par fissions coulombiennes des gouttelettes plus petites, qui contiennent environ 2% de la masse et 15% de la charge de la gouttelette d'origine. Les gouttelettes successives subiront le même processus et cela sur plusieurs cycles jusqu'à donner un ion fortement désolvaté.

### 2.2.3. La désolvatation des gouttelettes

Actuellement deux mécanismes sont proposés pour expliquer la production d'ions complètement désolvatés en phase gazeuse à partir de gouttelettes chargées : le mécanisme proposé par Dole appelé « Charged Residue Model (CRM) » (125) et le mécanisme proposé



par Iribarne et Thomson appelé « Ion Evaporation Mechanism (IEM)» (126–128) (Figure 40).

Le mécanisme de Dole (CRM) propose que la série de fissions coulombiennes de la gouttelette initiale aboutisse dans la source à des gouttelettes contenant une seule molécule chargée, le solvant s'évaporant ensuite dans l'interface jusqu'à obtenir un ion complètement désolvaté.

Le mécanisme d'Iribarne et Thomson (IEM) propose qu'à un stade intermédiaire de sa vie (avant le rayon critique de Rayleigh), le champ électrique à la surface de la gouttelette soit suffisant pour désorber des ions directement de la gouttelette (évaporation). Ce processus serait très important pour une gouttelette d'un rayon inférieur à 10 nm.

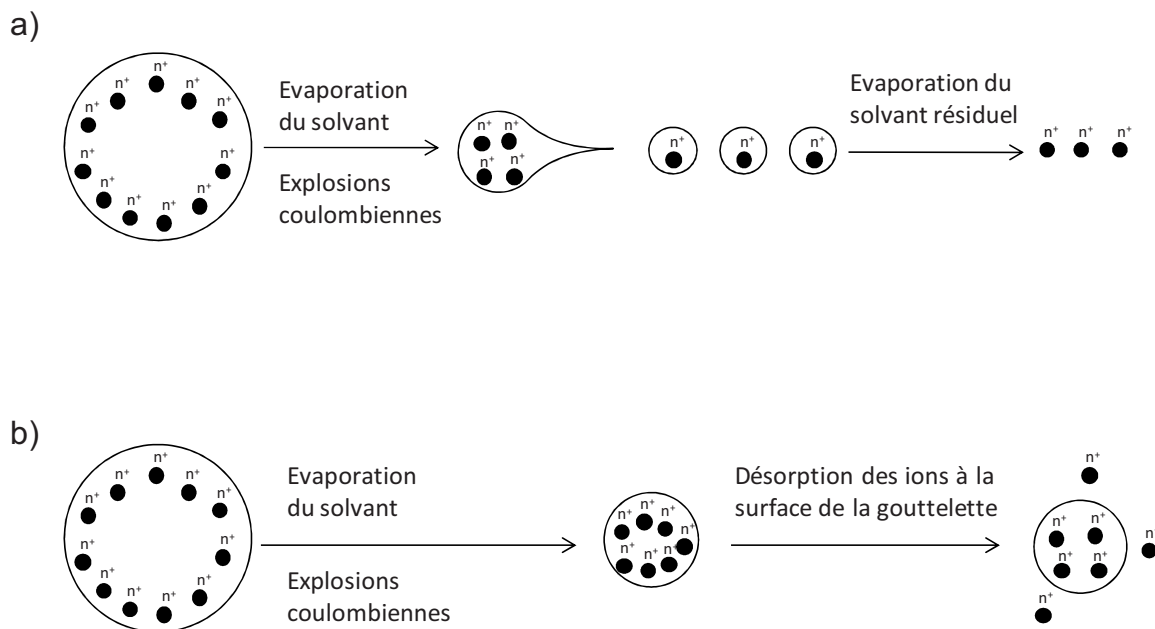


Figure 40 : Représentation schématique du modèle de Dole et coll. (a) et du modèle d'Iribarne et Thomson (b).

De nos jours, il est communément admis que les petits ions (sels) sont produits majoritairement par le mécanisme d'évaporation des ions (IEM) (129–132) alors que les protéines globulaires de taille plus importante sont produites par le mécanisme de Dole (132, 133).

## 2.3. L'interface

### 2.3.1. Principe de la transmission des ions

Après leur ionisation et leur désorption, les ions sont dirigés vers l'analyseur du spectromètre de masse sous l'action combinée de champs électriques et de gradients de pressions. La zone du spectromètre de masse appelée « interface » joue ce rôle très important.

En effet, l'interface sépare la zone de production des ions à pression atmosphérique de l'analyseur (vide poussé  $\approx 1.10^{-6}$  mbar). Il règne dans cette zone une pression intermédiaire de l'ordre du mbar. Cette pression intermédiaire est suffisamment faible pour que les ions acquièrent une énergie cinétique importante et suffisamment élevée pour entraîner des collisions entre les ions et les molécules de gaz ( $N_2$ ), convertissant ainsi l'énergie cinétique en énergie interne.

### 2.3.2. Rôle de l'interface dans la conservation des complexes non-covalents

Un point crucial pour conserver les interactions non-covalentes durant le processus d'ionisation/désolvation est l'optimisation des paramètres du spectromètre de masse qui contrôlent l'énergie communiquée aux ions dans cette interface. L'interface correspond physiquement à la zone entre le cône d'échantillonnage et le cône d'extraction. Pour l'ESI-IT, cela correspond donc à la zone entre le capillaire de transfert et le skimmer. Deux paramètres sont particulièrement importants et nécessitent d'être optimisés pour chaque nouveau système afin d'optimiser la sensibilité et la résolution des spectres de masses tout en préservant la dissociation des complexes. Le premier est la pression au niveau de cette interface, elle affecte l'efficacité des collisions. Le second est la tension d'accélération, elle contrôle l'énergie cinétique transmise aux ions de la source.

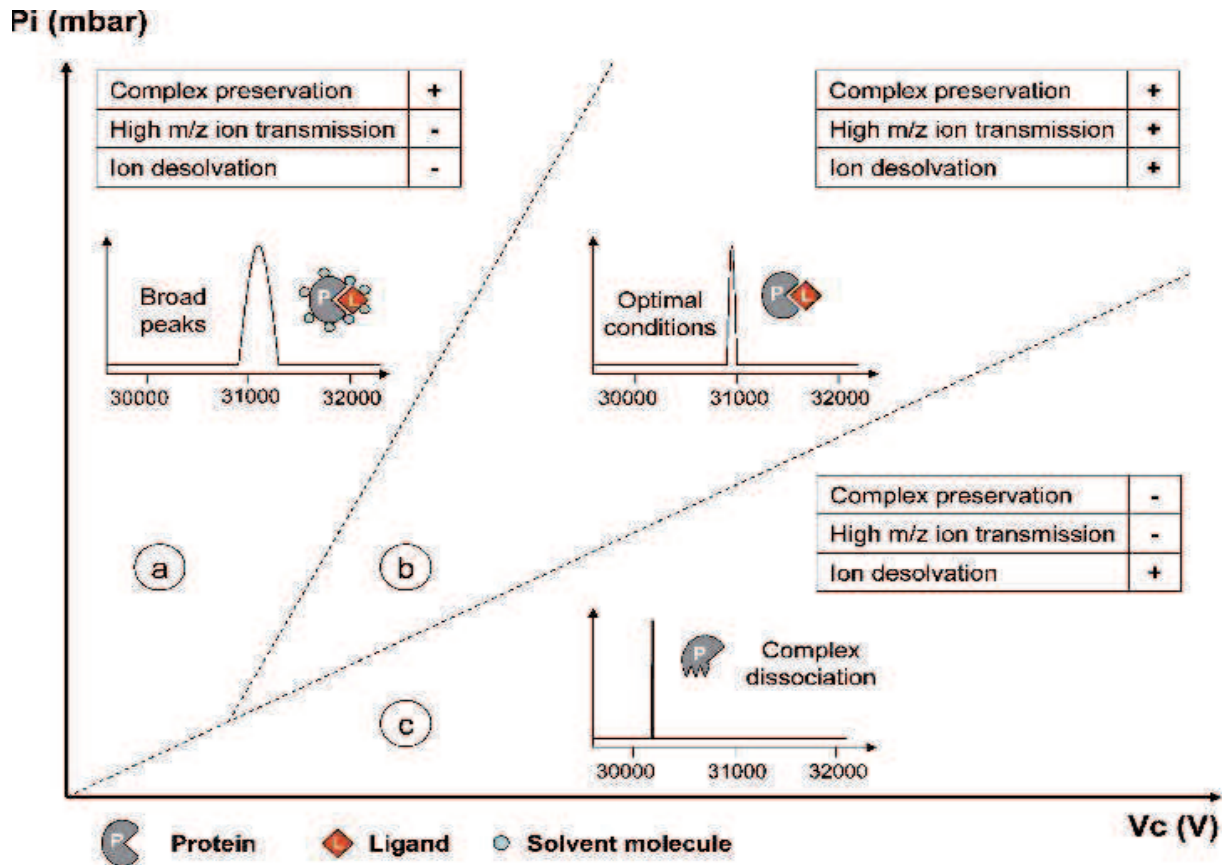


Figure 41: Représentation schématique de l'optimisation de la pression à l'interface ( $P_i$ ) et de la tension d'accélération ( $V_c$ ). **(a)** Région de désolvation incomplète, de faible transmission des ions de rapport  $m/z$  élevé et de préservation des complexes non-covalents ( $V_c$  faible et  $P_i$  élevée). **(b)** Région d'ajustements optimaux de  $V_c$  et  $P_i$ : région avec le meilleur compromis entre une désolvation efficace (pics étroits), une préservation des complexes non-covalents et une bonne transmission de ions de rapport  $m/z$  élevé. **(c)** Région de dissociation des complexes non-covalents, avec une faible transmission des ions de rapport  $m/z$  élevé ( $V_c$  élevée,  $P_i$  faible) alors que la désolvation est améliorée. Figure reproduite de la revue de Sanglier (110)

A des tensions d'accélération élevées, les ions ont des énergies cinétiques initiales plus élevées ce qui aboutit à des collisions de forte énergie et une possible dissociation des interactions faibles. Diminuer la tension d'accélération conduit à une perte de sensibilité considérable car la transmission des ions de forts  $m/z$  n'est plus correctement optimisée, la désolvation est moins efficace et la précision de masse s'en trouve fortement affectée. Il faut donc ajuster ces paramètres finement afin d'obtenir le meilleur compromis entre une désolvation suffisante, une transmission optimale des ions de  $m/z$  élevés et une préservation des édifices non-covalents en évitant les collisions en phase gazeuse destructives.

## 2.4. Principe de l'analyse des ions

Les études menées dans ce travail de thèse ont été principalement effectuées sur un appareil type ESI-IT, mais aussi sur un ESI-Q-TOF à haute résolution arrivé au cours de ma troisième année. Les principes de la trappe ionique et du TOF sont développés dans les annexes.

## 3. Compatibilité des tampons avec l'ESI

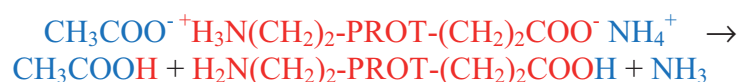
Habituellement, les tampons utilisés pour la purification ou l'extraction de protéine ou de complexes non-covalents (tampon phosphate, Tris, Hepes, MES, ...) sont des sels non-volatils qui ne sont donc pas compatibles avec l'analyse ESI-MS, même à l'état de traces. En conséquence, avant d'analyser un complexe non-covalent par MS, il est indispensable de changer le tampon de purification, procédure que nous appelons « étape de dessalage » et/ou « échange de tampon ». Le tampon d'analyse doit remplir deux conditions : (1), être compatible avec le processus d'ionisation électrospray, i.e. les tampons doivent être volatils et ne doivent pas provoquer la suppression d'ions, et (2), l'intégrité de l'assemblage non-covalent en solution doit être préservée. Les tampons ammonium sont ceux qui répondent le mieux à ces deux critères. Les tampons classiquement utilisés sont donc l'acétate d'ammonium, le bicarbonate d'ammonium, le bicarbonate de triéthylammonium (134) et l'eau (110). Ces tampons permettent de travailler à des pH compris entre 5 et 8,5. Des ajustements de pH vers des pH plus acides ou plus basiques peuvent être réalisés en ajoutant respectivement de faibles volumes d'acide formique ou d'hydroxyde d'ammonium. La force ionique d'un tampon peut aller de 10 à 500 mM suivant la stabilité du complexe (135). Dans la plupart des études, on utilise des tampons ammonium à des concentrations comprises entre 10 et 100 mM pour optimiser la qualité des spectres de masse.

### 3.1. Mécanisme de réduction des adduits de sodium sur la protéine avec un tampon acétate d'ammonium:

D'après Kebarle et Verkerk (136), un tampon comme l'acétate d'ammonium n'est pas utilisé que pour ses propriétés tampon, qui sont d'ailleurs plutôt faibles, ou ses propriétés non-dénaturantes. Il permet d'obtenir des spectres de masse très propres avec des protéines multiprotonées. En effet, utiliser de l'acétate d'ammonium permet de réduire la formation d'adduits de sodium sur la protéine. Le sodium est un contaminant très commun dans les

échantillons de protéines. Même quand la concentration en sodium de la solution nébulisée est faible, on observe des adduits de sodium sur la protéine qui peuvent perturber l'interprétation des spectres. Les adduits de sodium résultent de la formation en solution de paire d'ions entre les résidus acides de la protéine et le sodium. En présence d'acétate d'ammonium, les ions  $\text{NH}_4^+$  vont entrer en compétition avec les cations  $\text{Na}^+$  en solution, et les ions acétate avec les anions  $\text{Cl}^-$ . Plus la concentration relative en acétate d'ammonium en solution est large par rapport à celle de  $\text{NaCl}$ , plus l'équilibre se déplace vers la formation de paires d'ions n'impliquant au final que des ions acétate ou ammonium.

Lors du passage en phase gazeuse, l'activation collisionnelle que subit la protéine résultante conduit à une perte facile d'acide acétique et d'ammoniac, comme montré sur l'équation ci-dessous (136) :



La dissociation est facile car l'énergie nécessaire pour dissocier l'ammoniac ou l'acide acétique est seulement de 11 ou 15 kcal/mol respectivement, alors qu'elle est plus élevée pour le sodium ou le chlore (137).

En présence d'acétate d'ammonium, les adduits non spécifiques de sodium sont ainsi majoritairement éliminés par les mécanismes évoqués ci-dessus.

### 3.2. Mécanisme de charge de la protéine en phase gazeuse avec un tampon ammonium, à pH neutre:

Le phénomène de charge de la protéine en phase gazeuse est une réaction de transfert de proton,  $\text{AH}^+ + \text{B} \rightarrow \text{A} + \text{BH}^+$ , qui procède sans énergie d'activation lorsque la basicité en phase gazeuse du composé B est supérieure à celle de A. La notion de basicité en phase gazeuse est pertinente car le solvant a pratiquement complètement disparu lorsque ce phénomène de charge a lieu. Les basicités en phase gazeuse des résidus basiques ou des groupes amides à la surface de la protéine étant plus élevées que celle de l'ammoniac, les protéines se retrouvent facilement protonées avec des tampons ammonium (136). A pH neutre, la formation des ions mutichargés résulte en partie de la protonation des résidus basiques, mais est due essentiellement à la neutralisation des résidus acides.

## 4. Informations déduites du spectre de masse

### 4.1. Conformation de la protéine

Lorsqu'on analyse une biomolécule en conditions non dénaturantes (tampon ammonium avec pH et force ionique contrôlés), sa conformation native est préservée. Par conséquent, dans cet état replié, moins d'acides aminés sont accessibles pour être protonés que dans un état déplié. En conditions dénaturantes, comme par exemple dans une solution d'eau et d'acétonitrile acidifiée avec de l'acide formique, à pH 3, la protéine est dans un état déplié avec un nombre de résidus accessibles moyen plus élevé. Le nombre de charges effectives d'une protéine en conditions non dénaturantes est ainsi beaucoup plus faible que le nombre observé en conditions dénaturantes. Sur le spectre de masse, on détecte ainsi des ions à des  $m/z$  plus élevés qui portent moins de charges. Une distribution d'état de charge étroite avec des nombres de charge faibles sera donc caractéristique d'un spectre de protéine obtenu en conditions non dénaturantes. Au contraire, une distribution d'état de charge large avec des nombres de charge élevés sera caractéristique d'un spectre de protéine obtenu en conditions dénaturantes.

Les analyseurs traditionnellement utilisés pour travailler en conditions non dénaturantes seront donc ceux qui possèdent une gamme de masse étendue (supérieure à 4000  $m/z$ ), en particulier l'analyseur à temps de vol (TOF). De nombreux instruments avec source ESI sont néanmoins couplés à des quadropoles ou des trappes ioniques, mais il faut garder à l'esprit que leur gamme de  $m/z$  limitée peut constituer un obstacle technique à des applications MS non dénaturantes. Ainsi avec une trappe ionique limitée à des ions de  $m/z$  de 3000, on ne peut en général analyser des systèmes dont la masse dépasse 30 kDa.

### 4.2. Spécificité de l'interaction observée

Comme cela est détaillé par Smith *et al.*, en 1995, différents contrôles peuvent être réalisés pour affirmer qu'une interaction non-covalente est spécifique (138) :

- L'ajustement des conditions à l'interface ne doit pas modifier la stœchiométrie.
- Le complexe doit se dissocier suite à des modifications de conditions en solution (pH, température, tampon, ...).

- Le complexe doit se dissocier lorsqu'on varie les paramètres de l'interface, en d'autres termes, des conditions plus dissociantes doivent entraîner la rupture du complexe non covalent.
- La formation du complexe doit être sensible à des modifications de l'un des partenaires du complexe.

### 4.3. Stœchiométrie apparente du complexe

Le nombre de ligands liés à la protéine en phase gazeuse est obtenu facilement avec l'équation :

$$\frac{m(\text{complexe}) - m(\text{protéine})}{m(\text{ligand})}$$

Cette stœchiométrie est dite apparente car elle ne reflète pas forcément la saturation du complexe. Seul un titrage stœchiométrique donne accès à la stœchiométrie réelle.

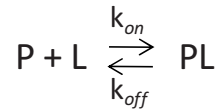
### 4.4. Constantes cinétiques du complexe

L'avantage unique de l'ESI-MS par rapport aux techniques biophysiques est de fournir directement des données sur toutes les espèces individuelles présentes en solution, grâce à la mesure précise de leurs masses. De plus, les intensités relatives des différentes espèces observées sur le spectre de masse peuvent servir à estimer l'abondance relative de chacun des composés, et ainsi donner une information sur les affinités relatives en solution. La combinaison de ces deux types d'information, la mesure de masse précise et les intensités relatives des pics, peut être utilisée rapidement pour mettre en évidence une interaction et estimer son affinité relative. De nombreuses interactions moléculaires avec des constantes de dissociation allant du nM au mM ont déjà été caractérisées par ESI-MS (139–143). Nous allons d'abord décrire les paramètres cinétiques usuels d'un complexe.

#### 4.4.1. Constante d'équilibre de dissociation ( $K_D$ )

##### 4.4.1.1. Définition

L'équation la plus simple décrivant la formation réversible d'un complexe protéine –ligand est la suivante :



L'équilibre est atteint lorsque les concentrations des espèces ne changent plus, ou quand la vitesse d'association du ligand à la protéine est égale à la vitesse de dissociation du complexe :

$$\underbrace{[P] \cdot [L] \cdot k_{on}}_{\text{Formation (2<sup>nd</sup> ordre)}} = \underbrace{[PL] \cdot k_{off}}_{\text{Dissociation (1<sup>er</sup> ordre)}}$$

Soit

$$\frac{k_{off}}{k_{on}} = \frac{[P] \cdot [L]}{[PL]} = K_D$$

$K_D$  est la constante d'équilibre de dissociation, elle s'exprime dans la dimension d'une concentration molaire, en unités M.  $K_D$  est inversement reliée à l'affinité d'une protéine pour un ligand, et quand la concentration en ligand est égale au  $K_D$ , la protéine est saturée à 50%.

#### 4.4.1.2. Détermination du $K_D$ par MS

La technique de détermination du  $K_D$  par ESI-MS est basée sur la détection et la quantification directe des ions PL (complexes protéine-ligand) et P (protéine libre) en phase gazeuse. Le principe de base est donc que la fraction de liaison ( $f_{\text{bound}}$ ) mesurée par MS est égale à la fraction de liaison à l'équilibre en solution.

L'équation quadratique classique établit une relation entre  $f_{\text{bound}}$ , le ratio de la concentration en protéine liée sur la concentration totale de protéine, qui constitue la variable mesurée,  $K_D$ , la constante de dissociation à l'équilibre du complexe protéine-ligand, qu'on cherche à déterminer, et  $[P]_0$  et  $[L]_0$ , respectivement les concentrations initiales en protéine et ligand :

$$f_{\text{bound}} = \frac{[PL]}{[P]_0} = \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4[P]_0 [L]_0}}{2[P]_0}$$



Le  $K_D$  peut être déterminé par une régression non-linéaire sur les données  $[L]_0$  et  $f_{\text{bound}}$  en utilisant l'équation quadratique classique, où  $[L]_0$  est la variable entrée en abscisse et  $f_{\text{bound}}$ , la variable entrée en ordonnée.

#### 4.4.2. Constante de vitesse de dissociation ( $k_{\text{off}}$ )

$k_{\text{off}}$  est la constante de 1<sup>er</sup> ordre de dissociation, elle s'exprime en  $s^{-1}$ .  $k_{\text{off}}$  est directement relié au temps de demi-vie du complexe :

$$t_{1/2} = \frac{\ln 2}{k_{\text{off}}}$$

Ce temps  $t_{1/2}$  correspond au temps au bout duquel 50% du complexe initialement présent est dissocié.

#### 4.4.3. Constante de vitesse d'association ( $k_{\text{on}}$ )

$k_{\text{on}}$  est la constante d'association bimoléculaire de 2<sup>ème</sup> ordre, elle s'exprime dans la dimension d'une concentration par unité de temps, en unités  $M^{-1} \cdot s^{-1}$ . Cette constante a une limite supérieure due au fait que deux molécules doivent d'abord diffuser et se rencontrer avant de pouvoir interagir. Pour deux molécules en diffusion libre, ç.à.d. se déplaçant librement en solution par mouvements browniens, la valeur maximale pour  $k_{\text{on}}$ , en d'autres termes la limite de diffusion, est d'environ  $1.10^8$  à  $1.10^9 M^{-1} \cdot s^{-1}$  (144).

## 5. Reflet de l'équilibre en solution lors de l'analyse en phase gazeuse

### 5.1. Conservation de l'équilibre chimique en phase gazeuse

Quand la masse du ligand est beaucoup plus petite que celle de la protéine, comme c'est souvent le cas, des résultats erronés de  $K_D$  peuvent être obtenus car la transmission du spectromètre de masse est masse-dépendante. Ainsi, il est plus avantageux d'utiliser seulement les intensités  $I_{PL}$  et  $I_P$  du complexe et de la protéine car les deux ont des masses similaires. Du fait du mécanisme de l'électrospray, on pourrait s'attendre à ce que l'augmentation de la concentration des solutés dans les gouttelettes chargées, engendrée par

l'évaporation du solvant, entraîne une mesure de  $K_D$  apparent sous-estimée par rapport au  $K_D$  en solution, correspondant à une affinité plus forte. Cependant, si les taux d'association et de réaction inverse sont lents comparés au temps d'évaporation de la gouttelette, on peut s'attendre à un bon accord entre les valeurs de  $K_D$  mesurées en phase gazeuse et en solution. En nanoélectrospray, il a été établi par Peschke et ses collaborateurs (2004) que la première génération de gouttelettes issues de l'électrospray était la source majeure d'ions (145). Même en prenant le taux d'association le plus rapide, à savoir la limite de diffusion dans l'eau pour la réaction  $P+L \rightarrow PL$ , il a été montré que ce temps est trop court pour qu'un nouvel équilibre correspondant à l'augmentation de la concentration des solutés provoquée par l'évaporation du solvant ait lieu. Comme les protéines sont de grande taille, la vitesse de diffusion dépend essentiellement de la diffusion du ligand, et la constante d'association à la limite de diffusion diminue lorsque la taille du ligand augmente. Des ligands de taille moyenne comme l'adénosine diphosphate/triphosphate, avec des constantes d'association à la limite de diffusion  $k_{on}$  compris entre  $10^6$  et  $10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$ , sont trop lents pour engendrer un écart d'équilibre qui donnerait une erreur significative sur la détermination de  $K_D$  (136). Ainsi, pour le nanospray, une déviation de l'équilibre n'est pas attendue pour des ligands qui ne sont pas trop petits (145).

En combinant des mesures de fluorescence et d'anémométrie de phase Doppler, le groupe de Zenobi a par la suite confirmé ces résultats dans le cas d'une source ESI en suivant l'évolution de la concentration en analyte et de la proportion de complexes à différents stades de la vie des gouttelettes chargées (146). D'après leurs travaux, il apparaît que dans des conditions usuelles, le processus ESI ne provoque pas de distorsion significative des équilibres chimiques en solution.

## 5.2. La conservation des interactions non-covalentes dépend du type d'interactions

La nature de l'interaction non-covalente a un effet sur sa résistance à la dissociation en phase gazeuse. Les interactions électrostatiques sont renforcées dans le vide. En 1995, Feng et ses collaborateurs ont étudié la liaison d'une spermine fortement basique à un peptide SBP (spermin-binding peptide) (147). Malgré une constante d'équilibre de dissociation élevée (100  $\mu\text{M}$  en solution), le complexe SBP/spermine a pu être détecté par ESI-MS. De plus, le complexe a été inhabituellement stable en phase gazeuse, ce qu'ils ont montré avec des expériences de CID. Alors que l'énergie de collision a été augmentée, la dissociation de

liaisons covalentes est observée avant celle du complexe non covalent car, en absence d'effet d'écrantage des charges par le solvant, un environnement sans solvant augmente la stabilisation coulombienne de charges opposées (147).

Les complexes avec des oligonucléotides impliquent usuellement des forces plutôt électrostatiques. Des études de complexes protéine/ADN (148) et de complexes ARN/aminoglycoside (140), sont des exemples où la dissociation du complexe en phase gazeuse est très difficile à obtenir, c'est-à-dire, les ions multichargés du complexe sont stables à de fortes énergies de collision dans l'interface.

À l'extrême opposé, le travail de Robinson *et al.* (149) avec la protéine liant l'acyl coenzyme A et des analogues d'acyl coenzyme A, a illustré un exemple où les données ESI-MS ne reflètent pas les résultats observés en solution. Les analogues d'acyl coenzyme A qui présentaient des  $K_D$  différents en solution ne présentent plus de différence dans les expériences ESI-MS (149). Une combinaison d'interactions hydrophobes, électrostatiques et de van der Waals maintiennent l'interaction protéine/ligand dans cet exemple. Des changements de la longueur de la chaîne carbonée de l'acyl, qui affectaient fortement la liaison en solution, ne semblent plus affecter la stabilité du complexe en phase gazeuse.

De la même manière, les stabilités en phase gazeuse de complexes non-covalents entre l'anhydrase carbonique bovine et des inhibiteurs para-substitués de benzenesulfonamide ne montrèrent aucune corrélation avec l'hydrophobicité des inhibiteurs (150). Pourtant les constantes de taux de dissociation en solution étaient bien corrélées avec l'hydrophobicité pour le même système. Robinson et ses collaborateurs notaient, en 1996, que pour les interactions qui relevaient principalement d'effet hydrophobe, la fraction de complexe observé en MS était faible ( $\leq 20\%$ ) (149). Il faut donc prendre ce phénomène en considération dans l'interprétation des résultats d'ESI-MS.

D'autres études ont suggéré que les complexes conservaient une stabilité similaire en passant de la solution à la phase gazeuse. Ainsi l'étude de Hunter *et al.*, en 1997, sur la liaison de l'hème à différentes formes d'apomyoglobine ou d'apocytochrome b5 a montré des résultats en MS comparables aux données cinétiques obtenues en solution (151). L'étude de Kraunsoe *et al.* sur la trypsine avec une variété d'inhibiteurs a montré par CAD que l'ordre des affinités de liaison en MS reflétait bien celui observé en solution (152).

### III. Cas particulier d'interactions non-covalentes: les systèmes à faible affinité

La spectrométrie de masse ESI est une méthode rapide et sensible pour déterminer des constantes de dissociation. Elle est cependant basée sur l'observation de l'interaction en phase gazeuse. L'équilibre en solution peut être significativement altéré lors du transfert en phase gazeuse. En effet, en phase gazeuse les interactions électrostatiques sont renforcées alors que les interactions hydrophobes sont fortement défavorisées à cause de l'absence de solvant. Ainsi, pour certains complexes non-covalents, en particulier pour ceux qui résultent au moins en partie d'interactions hydrophobes, le complexe en phase gazeuse ne reflète pas directement celui de la phase aqueuse. L'interaction peut être telle qu'il est difficile d'éviter une dissociation partielle du complexe spécifique dans l'interface de l'instrument (153–156). La conséquence en est que la fraction de liaison tend à être sous-estimée, d'autant plus que le phénomène de dissociation en phase gazeuse est étendu. De plus, pour observer en MS des complexes de faible affinité, il peut être nécessaire d'utiliser un large excès de ligand, ce qui conduit à des interactions non-spécifiques appelées agrégations en phase gazeuse. La stœchiométrie du complexe est alors modifiée comparée à celle qui est observable en solution. De plus, l'agrégation peut conduire à de faux positifs et a été montrée comme donnant des valeurs de  $K_D$  artificiellement faibles (157).

#### 1. Limites de l'étude par spectrométrie de masse de systèmes à faible affinité

##### 1.1. Problème de l'agrégation en phase gazeuse

L'agrégation en phase gazeuse, symptomatique d'une interaction non spécifique du ligand avec la protéine ou avec lui-même, se manifeste par des stœchiométries de liaison supérieures au nombre de sites de liaison dans la protéine. On peut réduire ces interactions non spécifiques en diminuant la concentration des analytes (138, 158). Dans le cas des systèmes à faible affinité, on est cependant limité par la faible fraction de liaison observable à ces concentrations.

Comme la masse d'un complexe non-spécifique est identique à celle d'un complexe spécifique à la même stœchiométrie, l'intensité du pic  $I_{PL}$  observé peut conduire à un  $K_D$  apparent qui serait trop faible. On pourrait penser que ces liaisons faibles non-spécifiques sont facilement dissociables durant les étapes de désolvatation dans l'interface, minimisant ainsi la contribution non-spécifique dans la mesure du  $K_D$  (145). Mais cette supposition sur les complexes non-spécifiques, qui peut paraître logique, ne reflète pas systématiquement les résultats obtenus. Il ne faut pas négliger que la liaison ion-neutre est forte en phase gazeuse et que la protéine est multichargée quand la gouttelette se désolvate. Une étude décrit exactement ces effets pour un complexe protéine/carbohydrate (159). Il a été observé non seulement des complexes non-spécifiques en phase gazeuse, mais aussi une liaison non-spécifique plus forte que celle spécifique. Dès 1991, Ganem *et al.* observaient déjà des complexes de stœchiométrie 1 : 2 et 1 : 3 de HEWL avec NAG3, qui pouvaient être réduits mais non éliminés en augmentant la tension d'accélération (160). Ils considéraient que ces complexes pouvaient être des artefacts d'adduits de NAG3 complexés à HEWL. Depuis, de nombreux travaux ont également traité d'agrégation en phase gazeuse, ont essayé de l'expliquer ou de la quantifier (145, 158, 159, 161–164).

L'historique de la formation d'une gouttelette montre qu'une protéine et un ligand n'ayant pas réagi ensemble peuvent se retrouver aléatoirement dans la même gouttelette. Après les étapes d'explosion coulombiques et d'évaporation totale d'une telle gouttelette, le ligand peut se retrouver agrégé à la protéine, ce qui va conduire à un complexe protéine-ligand artefactuel. Les agrégations peuvent être fortement renforcées si le ligand tend à l'autoagrégation.

## 1.2. Problème de dissociation en phase gazeuse (GPD)

Dans le travail d'Ogorzalek *et al.* (1993), l'affinité déterminée par MS pour le système protéine RNaseS/ peptide S a été trouvée plus faible que celle attendue d'après des expériences menées en phase aqueuse (165).

Des expériences MS qui ont échoué pour observer des complexes non-covalents alors qu'ils étaient pourtant attendus sont quelque fois rapportées (166, 167). Ces résultats trouvent aujourd'hui une explication avec la dissociation en phase gazeuse (GPD).

En 1997, Hunter *et al.* ont déterminé que le voltage nécessaire pour dissocier les ions complexes protéine-hème dans la région orifice-skimmer permet de mesurer la stabilité du

complexe et est corrélé à l'énergie d'activation nécessaire pour dissocier le complexe en solution (151).

Rogniaux *et al.* (1999) ont également utilisé la GPD, provoquée en augmentant la tension d'accélération dans l'interface source-analyseur, pour rapidement évaluer les stabilités relatives d'inhibiteurs non-covalents en phase gazeuse (168). Ils ont mesuré la stabilité du complexe en phase gazeuse avec le Vc50, voltage nécessaire pour dissocier 50 % du complexe initialement présent. De même qu'Hunter *et al.* (151), ils trouvent que le Vc50 est corrélé à l'énergie des interactions électrostatiques et des liaisons hydrogène impliquées dans le contact aldose réductase-inhibiteur (168).

En 2004, Clark et Konermann n'arrivaient pas à détecter le complexe trypsine pancréatique de bovin-inhibiteur benzamidine par ESI-MS (156). En 2007, Sun *et al.* ont pu le détecter en nanoESI-MS en travaillant dans des conditions de nébulisation très douces, mais l'abondance relative de complexes détectés reste plus faible que celle attendue (154).

La GPD a lieu lorsque l'énergie interne communiquée aux ions via des collisions avec des molécules gazeuses résiduelles dans la première partie du spectromètre de masse atteint l'énergie d'activation pour dissocier la molécule. Le nombre de collisions ainsi que l'énergie des collisions vont influencer la GPD.

### 1.3. Superposition des deux problèmes

Lors de l'étude de systèmes à faible affinité, il va être nécessaire de travailler avec de fortes concentrations en ligand pour pouvoir observer le complexe, ce qui pourra conduire à de l'agrégation en phase gazeuse. De plus, quelque soit le type d'interaction impliqué dans le complexe, l'affinité étant faible, il y aura plus de risque de dissociation en phase gazeuse.

Ces systèmes sont donc particulièrement délicats à étudier car les problèmes d'agrégation et de dissociation se superposeront lors de l'analyse en phase gazeuse.

La figure 42, ci-dessous, résume différents évènements possibles lors de l'analyse en phase gazeuse d'un complexe de faible affinité :

- le complexe se dissocie
- le complexe ne se dissocie pas
- le complexe se dissocie et un ligand s'agrège

- le complexe ne se dissocie pas et un ligand s'agrège

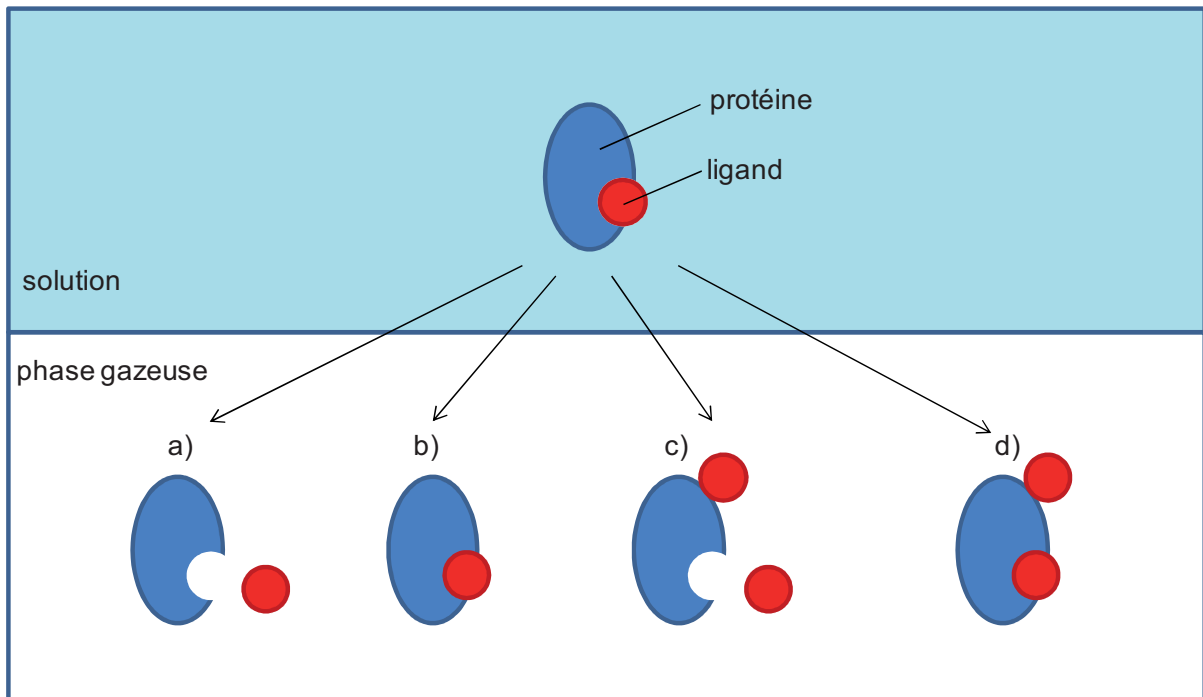


Figure 42: Représentation de différents évènements qui peuvent avoir lieu lors du passage en phase gazeuse d'un complexe protéine-ligand de faible affinité. a) l'interaction spécifique entre la protéine et le ligand (et une éventuelle interaction non-spécifique) est perdue; b) le complexe spécifique reste intact et il n'y a pas d'agrégation ; c) l'interaction spécifique entre la protéine et le ligand est perdue en présence d'une forte agrégation ; d) le complexe reste intact et une agrégation subsiste.

## 2. Stratégies développées pour pallier les limites des systèmes à faible affinité

### 2.1. Stratégies pour s'affranchir de l'agrégation en phase gazeuse

#### 2.1.1. Diminuer la concentration en ligand

Wang *et al.* (2003) ont testé l'influence de la concentration en ligand sur la mesure du  $K_A$  et ont montré qu'en travaillant avec des concentrations équimolaires en protéine et ligand, ils pouvaient minimiser la formation d'agrégation (157). Cette stratégie ne peut s'appliquer à des systèmes de faible affinité car il est nécessaire de travailler avec un large excès de ligand pour pouvoir obtenir une détermination fiable du  $K_D$ .

### 2.1.2. Quantifier l'agrégation avec des protéines de référence

Pour quantifier l'agrégation en phase gazeuse, le groupe de Sun et ses collaborateurs (2006) a décrit une méthode qui consistait à utiliser une protéine de référence connue pour ne pas se lier spécifiquement au ligand en solution (161). Ils faisaient l'hypothèse que l'agrégation était indépendante de la taille et de la structure de la protéine. Touboul *et al.* (2008) ont démontré que cette hypothèse était valide pour des interactions de type protéine-sucre neutre mais fautive dans le cas où les interactions étaient de type électrostatique (163). En 2010, Sun *et al.* ont approfondi le sujet et montrent que la liaison non-spécifique de ligands chargés à une protéine durant l'électrospray est un processus purement statistique, indépendant de la structure et de la taille de la protéine (162). Des expériences de dissociation réalisées sur les ions gazeux de complexes non-spécifiques suggèrent que les distributions non-équivalentes observées avec différentes protéines résultent de différents degrés de dissociation dans la source. Ils en concluent donc que la méthode avec protéine de référence peut aussi être utilisée avec des ligands ioniques à condition que la dissociation en source des interactions non-spécifiques soit minimisée. La variabilité de la stabilité de l'agrégation sur la protéine de référence dans le cas des ligands chargés rend cette méthode délicate à appliquer.

### 2.1.3. Distinguer qualitativement l'agrégation des interactions spécifiques

La méthode de Czuczy permet de discerner qualitativement les interactions spécifiques de celles non-spécifiques (164). La méthode est basée sur l'ionisation par spray électrosonique d'échantillons contenant des molécules de protéine et de ligand et sur la détection MS en utilisant la fonction scan de l'ion précurseur sur un instrument triple quadrupole. Les complexes intacts protéine-ligand sélectionnés par masse sont soumis à la fragmentation par CID dans la cellule de collision alors que le second analyseur est paramétré pour détecter le ligand protoné. Les complexes spécifiques et non-spécifiques montrent des fragmentations radicalement différentes et caractéristiques, ce qui constitue la base de la détection sélective des complexes spécifiques.

### 2.1.4. Quantifier l'agrégation par compétition avec un ligand de référence

Touboul *et al.* ont développé une méthode quantitative qui nécessite d'avoir un ligand de référence qui ne s'agrège pas sur la protéine et dont le  $K_D$  est connu (163). La méthode est basée sur la détermination de la concentration en ligand pour laquelle il y a 50 % d'inhibition, l'IC50, dans une expérience de compétition entre le ligand d'intérêt et celui de référence. Les



valeurs d'IC50 déterminées par MS sont ensuite converties en  $K_D$  en utilisant la valeur de  $K_D$  du ligand de référence. Dériver un  $K_D$  par calcul à partir d'une IC50 est une méthode moins fiable que de déterminer le  $K_D$  lui-même. De plus, une contrainte majeure de cette méthode est qu'elle nécessite l'existence au préalable d'un ligand de référence, connu pour se lier au même site de liaison que celui des ligands criblés, mais sans former d'agrégation en phase gazeuse.

#### 2.1.5. Quantifier l'agrégation et les interactions spécifiques par une méthode de déconvolution statistique

van der Rest et collaborateurs (169), comme Sun *et al.* (161, 162), proposent que l'agrégation en phase gazeuse est un phénomène statistique. Pour déterminer des constantes de liaison spécifiques, van der Rest et ses collaborateurs ont construit un modèle mathématique qui permet de déconvoluer les interactions spécifiques et non-spécifiques. Le modèle est basé sur une distribution binômiale des ligands sur le ou les sites de liaison spécifiques et sur une distribution de Poisson pour la liaison non-spécifique des ligands. Cette méthode nécessite de connaître la stœchiométrie de liaison car elle rentre en compte dans le développement mathématique (169).

#### 2.1.6. Limites de ces stratégies

La principale limite de ces stratégies est qu'elles nécessitent certains pré-requis pour être applicables. Il faut en effet avoir une protéine ou un ligand de référence, ce qui suppose qu'il en existe. Le site de liaison ou la stœchiométrie doivent souvent être connus, et un minimum d'informations structurales est donc nécessaire.

La grande faiblesse de ces stratégies est qu'aucune ne traite le problème d'agrégation dans les cas relativement fréquents où le problème de la dissociation en phase gazeuse vient également compliquer l'interprétation des résultats obtenus par MS.

## 2.2. Stratégies pour s'affranchir de la dissociation en phase gazeuse

### 2.2.1. Augmenter la pression de l'interface

Schmidt *et al.* (2001) montrent qu'en augmentant uniquement la pression de l'interface, ils peuvent observer plus de complexes non-covalents intacts que de dissociés, à savoir plus

d'holomyoglobine que d'apomyoglobine ou de concanavoline tétramérique que de dimérique (170). Cependant cette stratégie n'est pas universellement suffisante.

### 2.2.2. Utiliser des additifs pour stabiliser le complexe

D'après Kitova et Klassen (171), la dissociation du complexe spécifique peut être réduite en utilisant des additifs tels que l'imidazole en solution. Il est suggéré que l'effet stabilisateur pourrait être dû à une formation en solution d'adduits d'imidazole sur le complexe. La dissociation de ces adduits par CID fournirait un effet thermalisant qui empêcherait que la liaison spécifique se dissocie aussi. Ils arrivent ainsi à préserver le complexe trypsine pancréatique bovine-inhibiteur benzamidine en ajoutant de l'imidazole et à obtenir dans ces conditions, des constantes d'association en accord avec les valeurs en solution de la littérature.

Bagal *et al.* ont montré que l'effet stabilisateur de l'ajout de vapeur d'imidazole est comparable à celui de l'ajout d'imidazole en solution (172). L'addition d'un autre composé, le sulfure d'hexafluorure, en phase gazeuse supprime également la dissociation du complexe en phase gazeuse (172). Ils proposent que le refroidissement évaporatif soit le premier mécanisme responsable des effets stabilisateurs de ces additifs en phase gazeuse. La dissociation des molécules d'imidazole liées non-spécifiquement (formées durant le processus l'électrospray) refroidit le complexe dans la source, et ainsi réduit les risques de GPD (172), aux dépendants cependant de la précision de température de mesure du  $K_D$ . Le risque associé à ces méthodes est que les additifs de "supercharging" altèrent la conformation de la protéine.

### 2.2.3. Corriger la fraction de liaison en utilisant un ligand de référence.

Klassen *et al.* proposent une nouvelle approche pour quantifier la GPD de complexes non-covalents (173). La méthode, appelée méthode avec ligand de référence, est basée sur l'analyse directe par ESI-MS et sur la compétition de ligand. Un ligand de référence, qui doit se lier spécifiquement à la protéine, au même site de liaison que le ligand d'intérêt, avec une affinité pour la protéine connue et qui forme un complexe protéine-ligand stable en phase gazeuse, est ajouté au mélange protéine-ligand en solution. A partir de la fraction de liaison du complexe protéine-ligand de référence observée directement par ESI-MS, l'affinité du ligand d'intérêt pour la protéine peut être déduite. Tjernberg et ses collaborateurs avaient déjà proposé une méthode semblable en 2004, mis à part le fait que le ligand de référence ne devait pas former avec la protéine un complexe stable en phase gazeuse (174).

#### 2.2.4. Limites de ces stratégies

Utiliser des additifs pour stabiliser la liaison non-covalente demande de faire des études préliminaires pour s'assurer que l'additif ne réagit avec aucun des réactifs, ce qui pourrait déplacer l'équilibre chimique. Les ligands de référence sont souvent employés pour des expériences de compétition, permettant ainsi de n'utiliser que le signal MS protéine-ligand de référence pour déterminer le  $K_D$  du ligand d'intérêt. La contrainte est qu'il faut déjà connaître un ligand spécifique au système.

Aucune méthode ne traite à la fois les problèmes d'agrégation en phase gazeuse et de GPD. Cependant, pour des systèmes de faible affinité, les risques d'agrégation sont accrus. Ces systèmes vont aussi avoir tendance à se dissocier plus facilement en phase gazeuse, car même si les interactions électrostatiques sont renforcées en phase gazeuse, elles le sont de manière proportionnelle à la force d'interaction en solution et restent donc globalement plus faibles que dans les systèmes à forte affinité.

Nous verrons que dans ce travail de thèse, les complexes hPEBP1-nucléotides étudiés sont sujets à l'agrégation en phase gazeuse ainsi qu'à la GPD. Nous avons donc dû mettre en place une nouvelle méthode qui puisse traiter les deux problèmes à la fois pour déterminer des  $K_D$  fiables.

# Section 3 : Présentation du travail de thèse

## I. Contexte : projet METASUPP

### 1. Objectifs du projet

Le but principal du projet METASUPP est de concevoir des petites molécules utilisables comme médicaments pour lutter contre la formation des métastases dans les cancers. En effet, le traitement des tumeurs primaires, basé sur des combinaisons de chirurgie, chimiothérapie, et irradiations, a fait beaucoup de progrès et c'est généralement le développement des métastases qui engage le pronostic vital des patients. Le projet est construit sur deux observations :

- La protéine hPEBP1 est impliquée dans un grand nombre de cancer où elle est reconnue pour être un suppresseur de métastases.
- La locostatine est une molécule qui inhibe la migration cellulaire et dont la cible cellulaire principale est précisément la protéine hPEBP1; l'interaction entre hPEBP1 et la locostatine est directe et supposée être covalente et non réversible au moment de l'initiation du projet.

L'hypothèse de travail centrale à ce projet est que des molécules de type locostatine, en se liant à hPEBP1, soient capables de renforcer le caractère anti-métastatique naturel de la protéine. Il devrait donc être possible de concevoir des molécules activatrices ("agonistes") de hPEBP1. Par ailleurs, hPEBP1 est capable de se lier à des nucléotides. Au démarrage de ce projet, il n'y avait pas d'information sur les sites de liaison de la locostatine et des nucléotides à hPEBP1. Par conséquent, le projet a aussi prévu de caractériser les sites de liaison à la locostatine et aux nucléotides, et de déterminer s'ils sont confondus. La conception rationnelle de molécules activatrices de hPEBP1 a donc été engagée en se basant d'une part sur la locostatine et d'autre part sur des structures de nucléotides. A terme, la synthèse et le criblage de composés à visée antimétastatique capables de se lier spécifiquement et réversiblement à la protéine, d'être stables et non toxiques, devrait produire des molécules candidates potentiellement utilisables en thérapeutique.

Le projet METASUPP est financé par l'Agence Nationale de la Recherche (ANR), le Cancéropole Grand Ouest, et la Ligue contre le Cancer.

## 2. Organisation

Pour caractériser les sites de liaison de la protéine et étudier le mécanisme moléculaire d'action des ligands possédant une activité potentielle d'activateur, nous avons combiné des techniques performantes permettant de synthétiser les ligands (chimie organique), d'étudier les interactions moléculaires (spectrométrie de masse et RMN) et de réaliser des tests d'activité sur cellules (biologie cellulaire). Ces trois volets sont nécessaires pour parvenir à la conception rationnelle de ligands de hPEBP1 forts, spécifiques et stables.

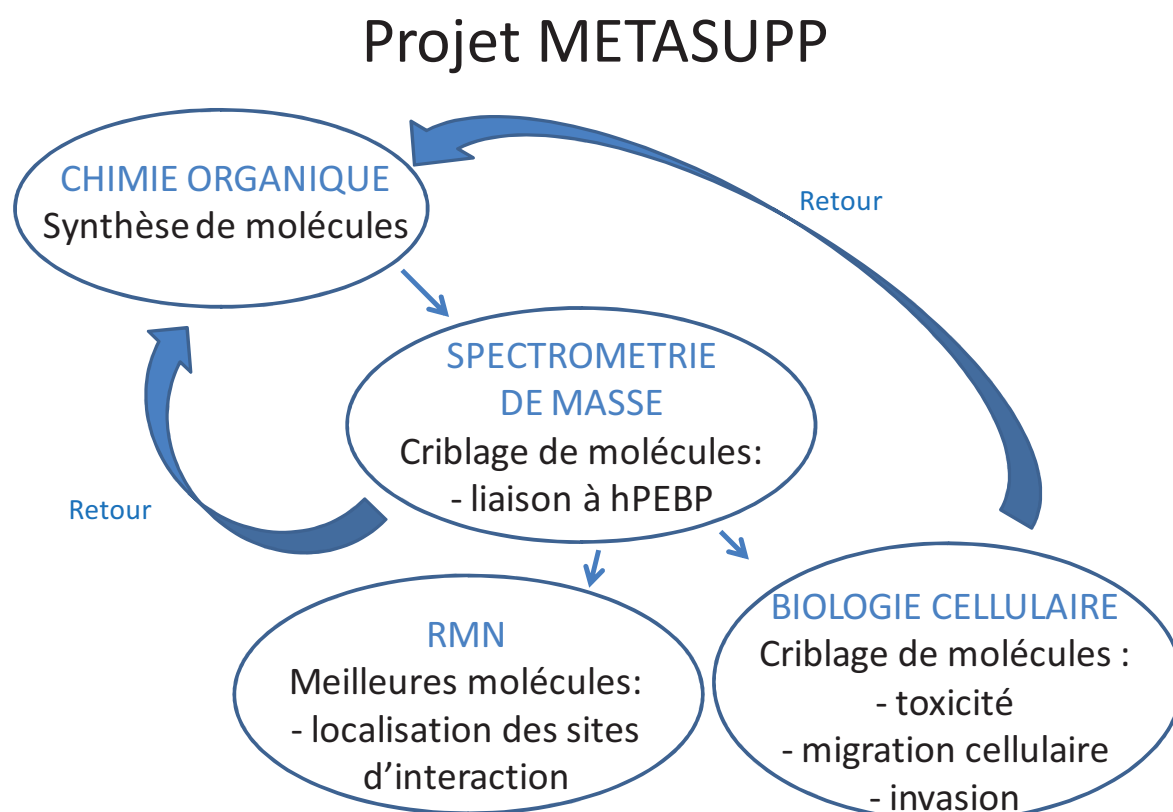


Figure 43: Représentation schématique de l'organisation du projet METASUPP

Les premières molécules sont criblées par MS pour leur capacité à se lier à hPEBP1. Les résultats de criblage sont ensuite pris en compte en chimie organique pour réorienter les synthèses vers de meilleurs ligands. En parallèle, les molécules de référence, en particulier les nucléotides, sont analysées en RMN pour déterminer leur site de liaison. Les meilleurs candidats issus du criblage moléculaire sont également analysés en RMN en complexe avec la protéine. Le criblage en biologie cellulaire est ciblé sur les propriétés antimigratoires et antimétastatiques des composés, et ces résultats de criblage sont eux aussi pris en compte pour les nouvelles synthèses de ligand. Les différentes équipes interagissent ainsi en boucle rétroactive.

Pour caractériser les sites de liaison de la protéine, des synthèses et des criblages sont réalisées pour les deux types de ligand (analogues de nucléotides et de locostatine), de façon indépendante.

Le criblage des analogues de nucléotides par spectrométrie de masse supramoléculaire et la contribution à la compréhension de leur mécanisme d'action constituent le point central de mon sujet de thèse.

## II. Objectifs du travail de thèse

### 1. Approfondir la relation structure-fonction de PEBP1 humaine

La structure et les fonctions d'une protéine sont intimement liées. La structure cristallographique de hPEBP1 est connue: on sait grâce à cette structure qu'elle possède une poche dite à anion. Au départ de ce projet les sites de liaison des nucléotides et de la locostatine n'étaient pas connus.

Une particularité de la protéine hPEBP1 par rapport à de nombreuses cibles classiques en cancérologie est qu'elle ne semble posséder aucune activité enzymatique. hPEBP1 intervient dans de nombreux phénomènes biologiques mais la fonction générale de cette protéine n'est pas élucidée. Pour ces raisons, nous avons souhaité approfondir la relation structure-fonction de la PEBP1 humaine non seulement en caractérisant le site de liaison nucléotidique de hPEBP1, mais également en recherchant une éventuelle activité enzymatique.

## 2. Fournir des éléments pour la synthèse raisonnée d'activateurs des propriétés antimétastatiques naturelles de hPEBP1

Dans le cadre du projet MétaSupp, la synthèse raisonnée d'analogues de nucléotides activateurs des propriétés antimétastatiques naturelles de hPEBP1 nécessite de cribler des ligands par MS pour leur affinité à PEBP1 humaine et de corrélérer les résultats de MS avec les résultats de criblage cellulaire.

## 3. Développer une méthode de suivi MS des complexes reflétant le comportement en solution

Le paramètre mesuré pour cribler des ligands selon leur affinité à une protéine est la constante d'équilibre de dissociation  $K_D$ . Notre étude bibliographique a mis en évidence qu'il n'existait pas de méthode capable de traiter à la fois l'agrégation en phase gazeuse et la GPD sans avoir à utiliser un ligand de référence. Or, pour hPEBP1, nous ne disposons pas de valeurs de  $K_D$  pour les ligands de référence. Par conséquent, une nouvelle méthode doit être développée pour pouvoir déterminer sans données préalables des  $K_D$  en phase gazeuse qui reflètent bien le  $K_D$  en solution.

## 4. Retombées potentielles

Le développement d'un savoir-faire pour l'étude de complexes non-covalents hPEBP1-nucléotides par spectrométrie de masse sur la plateforme du CBM, doit nous permettre d'étudier d'autres systèmes dans le cadre de collaborations visant à caractériser des liaisons protéine-protéine ou protéine ligand.





# Résultats et discussion

- Chapitre I : Optimisations expérimentales, Méthode d'injection et paramètres d'analyse
- Chapitre II : Application à des systèmes non-covalents
- Chapitre III : Développement d'une nouvelle méthodologie pour déterminer des constantes d'équilibre de dissociation des complexes à faible affinité
- Chapitre IV : Caractérisation du site nucléotidique de hPEBP1/RKIP
- Chapitre V : Recherche d'une activité enzymatique de hPEBP1/RKIP

# Chapitre I

Optimisations expérimentales  
Méthode d'injection et paramètres d'analyse

## *Résumé du chapitre I:*

*L'analyse de complexes non-covalents par MS nécessite des conditions expérimentales particulières, une adaptation du mode d'infusion et des optimisations fines de certains paramètres de l'ESI-IT. De plus, chaque nouveau système étudié nécessite d'ajuster spécifiquement une partie de ces paramètres. Les contraintes liées au criblage rapide de ligands selon leur affinité pour hPEBP1 nous ont amenées à choisir une infusion directe en ESI. Pour limiter la consommation d'échantillon, l'infusion a été miniaturisée en diminuant le débit d'infusion autant que possible et en développant une méthode originale d'« infusion en sandwich ». Il est montré que le choix du solvant utilisé pour constituer le sandwich est crucial pour préserver la conformation native de hPEBP1 et ses complexes non-covalents.*

*Afin de préserver au maximum les complexes lors du transfert en phase gazeuse, la tension d'accélération et la température de désolvatation ont été optimisées sur deux complexes, le dimère HU $\alpha_2$  et hPEBP1-GTP. Une diminution de la température de désolvatation, avec une valeur minimale de 120°C permet de limiter la dissociation du dimère HU $\alpha_2$  tout en gardant une désolvatation efficace. L'intégrité des complexes peut dépendre du voltage appliqué à la sortie du capillaire de transfert ( $V_{CE}$ ). Le dimère HU $\alpha_2$  est très sensible à la variation du voltage  $V_{CE}$ , la valeur de 200 V paraissant optimale. A l'inverse, le complexe hPEBP1-GTP est, lui, beaucoup moins sensible au réglage de la tension d'accélération. Dans certains cas, il est préférable d'appliquer un voltage de 40 V sur le skimmer pour préserver au maximum le complexe.*

La caractérisation des interactions PEBP-nucléotides et le criblage de composés analogues a imposé d'utiliser une méthode répondant à un certain nombre de critères:

- ✓ Miniaturisation de l'infusion par rapport à l'infusion classique pour obtenir:
  - \* une gestion économe des quantités d'échantillons de protéine cible (dont la préparation est coûteuse en temps et en moyens).
  - \* une infusion rapide et la possibilité de rincer facilement les lignes et l'aiguille d'ESI.

Pour ceci, une méthode d'infusion en sandwich a été développée.

- ✓ Conservation des liaisons non-covalentes

Pour cela des paramètres expérimentaux et instrumentaux ont été optimisés.

## I. Mise au point d'une méthode d'injection directe « en sandwich »

### 1. Choix du type d'infusion directe

Pour limiter la quantité de PEBP/RKIP humaine utilisée dans chaque expérience, nous avons dans un premier temps tenté de travailler en infusion directe avec une pointe émettrice nanoélectrospray, *a priori* moins consommatrice d'échantillon et donnant potentiellement une meilleure ionisation que l'électrospray. Cependant, il a été observé que le nettoyage après infusion directe nanoélectrospray nécessite jusqu'à 20 minutes par point expérimental, ce qui est incompatible avec le criblage rapide d'analogues de nucléotides. Nous avons donc choisi de travailler en infusion directe électrospray. De plus, la stabilité et la reproductibilité de l'électrospray est un avantage par rapport à l'infusion nanoélectrospray.

Pour minimiser autant que possible la quantité de matériel nécessaire à l'analyse, on peut jouer sur le débit d'injection ou sur le volume d'échantillon utilisé pour l'analyse.

## 2. Débit d'infusion

Le débit habituel en infusion ESI est de 3 à 4  $\mu\text{L}/\text{min}$ . Nous avons fixé la tension appliquée à l'aiguille à - 4000 V et étudié l'effet du débit d'infusion ESI sur l'intensité du signal de la myoglobine (Figure 44).

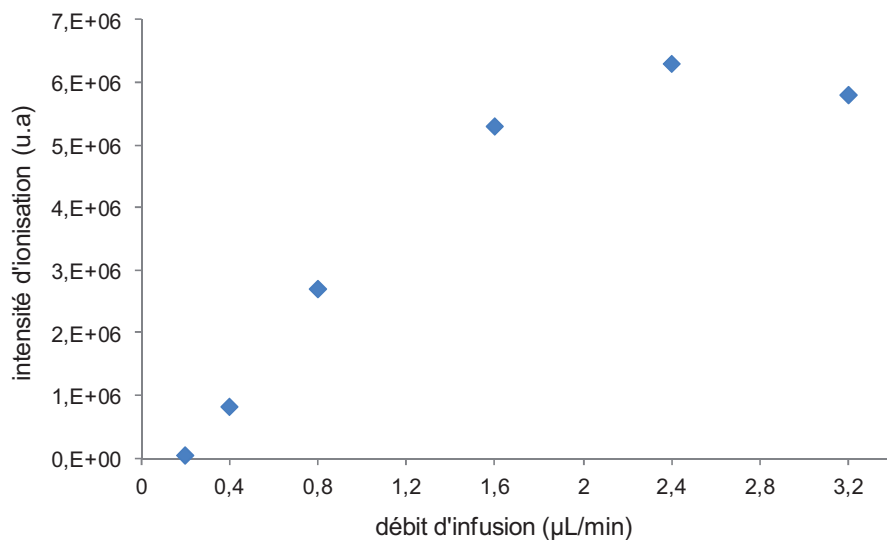


Figure 44: Effet du débit d'infusion sur l'intensité du pic majoritaire (21+) de la myoglobine.

L'intensité du signal est maximale pour un débit de 2,5  $\mu\text{L}$  par minute (Figure 44). Si l'on veut moyenner le signal sur au moins 2 minutes, un tel débit ne permet pas de travailler avec un volume d'échantillon suffisamment faible.

Afin d'approcher le débit nano (de 50 à 500 nL/min) sans perdre significativement de signal, une valeur compromise de 1,3  $\mu\text{L}/\text{min}$  a été choisie.

Pour obtenir le signal de PEBP native le plus intense possible à ce débit, le voltage appliqué à l'aiguille a ensuite été optimisé.

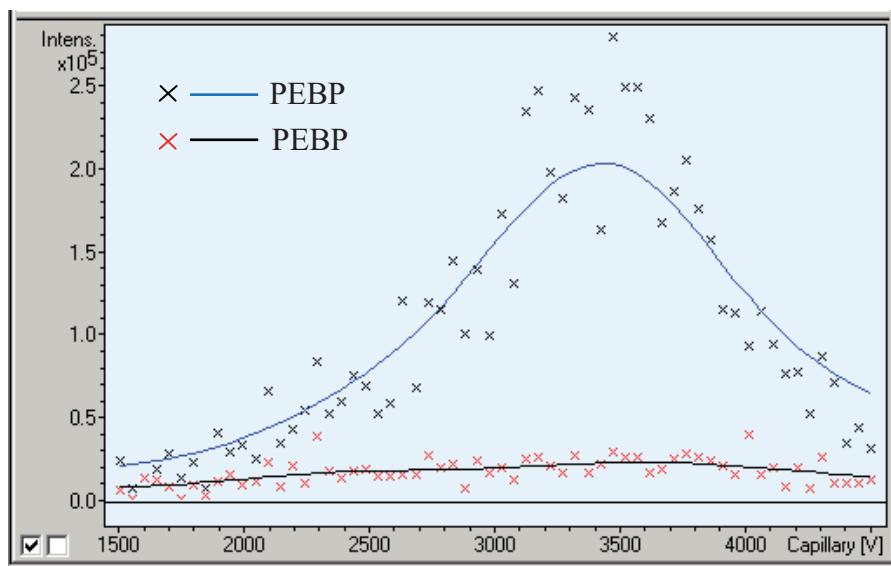


Figure 45: Intensités des pics 9+ et 10+ de la PEBP en fonction du voltage (négatif) appliqué à l'aiguille avec un débit d'infusion de 1.3  $\mu\text{L}/\text{min}$  (fonction Smartramp de EsquireControl).

D'après la Figure 45, pour un débit d'infusion de 1.3  $\mu\text{L}$  par minute, l'intensité des pics 9+ et 10+ de la PEBP est maximale lorsque la tension appliquée à l'aiguille est de - 3400 V.

### 3. Méthode d'infusion directe ESI originale

Dans la configuration pousse-seringue/source électrospray, le capillaire d'infusion possède un volume mort de 8  $\mu\text{L}$ . Au niveau du pousse-seringue, la butée du piston se situe à 6  $\mu\text{L}$  sur des seringues Hamilton de 50  $\mu\text{L}$ . Il faut donc injecter 14  $\mu\text{L}$  de solution d'analytes avant de voir apparaître le premier signal électrospray. Si l'on considère que l'on veut enregistrer le signal sur environ 3 minutes, l'échantillon ne devra pas faire moins de 5  $\mu\text{L}$ , ce qui impose de prélever au total environ 20  $\mu\text{L}$  d'échantillon pour une analyse. Seul un quart de l'échantillon introduit dans la seringue est réellement analysé par MS.

Pour éviter ce gaspillage de matériel, une méthode dans laquelle l'échantillon est pris en sandwich entre deux volumes de solvant dans la seringue d'injection a été développée (Figure 46). Cette méthode a été baptisée "infusion en sandwich".

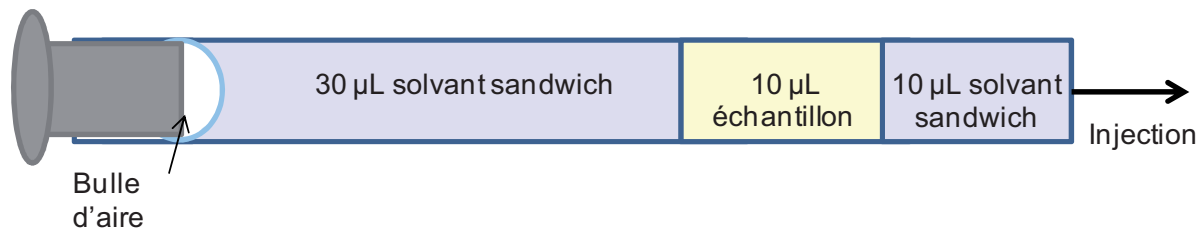


Figure 46: Représentation schématique de la méthode d'infusion en sandwich

Le « solvant sandwich » situé avant l'échantillon dans la seringue permet d'éliminer le volume mort du capillaire d'infusion et celui situé après l'échantillon pousse l'échantillon dans la source.

## 4. Effet du solvant du sandwich

### 4.1. Sur hPEBP1/RKIP seule

Le premier solvant utilisé pour faire « le sandwich » de l'échantillon de protéine native a été le WAFA, une solution eau/acétonitrile 50/50 v/v acidifié par 0.2% d'acide formique (Figure 48a). Ce solvant a été choisi pour faciliter le rinçage de la ligne d'infusion entre deux échantillons. Le spectre qui résulte de l'utilisation de ce solvant montre deux distributions d'états de charge. L'une, dans une gamme de  $m/z$  de 300 à 1000, est large avec des états de charge élevés, cette distribution étant caractéristique de celle d'une protéine dénaturée. L'autre, dans une gamme de  $m/z$  plus élevée, est étroite, avec des états de charge plus faibles. Cette distribution-là est caractéristique d'une protéine en conformation repliée, à l'état natif.

Pour aider à l'interprétation de la présence de ces deux distributions, une dénaturation thermique de hPEBP1 a été effectuée.

#### Dénaturation thermique de hPEBP1:

D'après les travaux de Vallée et ses collaborateurs (1999) (15), la PEBP bovine est dénaturée à partir de 54 °C. La protéine hPEBP1 a donc été chauffée à 60°C et la cinétique de dénaturation thermique a été suivie par MS (Figure 47).



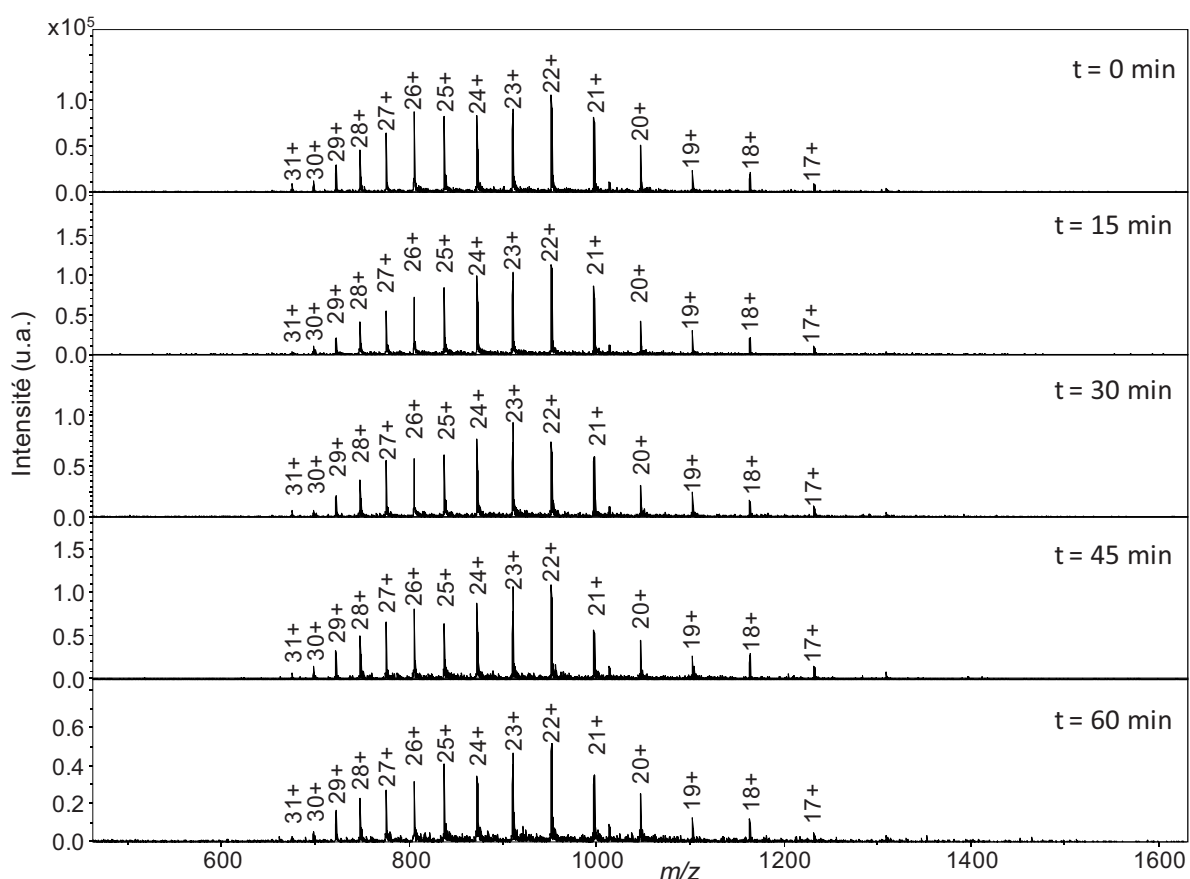


Figure 47: Spectres ESI-IT de la PEBCP à 1.8  $\mu\text{M}$  dans du WFA après 0, 15, 30, 45 ou 60 minutes de dénaturation thermique à 60°C

La dénaturation thermique ne modifie pas la distribution, déjà large avec des états de charge élevés au temps zéro, de la protéine. D'après cette cinétique, hPEBCP1 dénaturée est donc présente dès qu'elle est diluée dans le WFA.

Un échantillon de hPEBCP1 a été dénaturé par simple dilution de l'échantillon de protéine native dans du WFA puis analysé avec le même sandwich WFA et les mêmes conditions expérimentales que l'échantillon de protéine native de la Figure 48a. Le spectre est représenté sur la Figure 48b. Deux distributions sont observées, comme pour l'échantillon de hPEBCP1 native, mais les pics de la distribution correspondant à la protéine repliée sont d'intensité plus faible.

Ce contrôle suggère que le WFA utilisé pour faire le sandwich de l'échantillon de protéine native est capable de dénaturer celle-ci. Afin de conserver la configuration native de la PEBCP, le solvant du sandwich a ensuite été remplacé par un solvant non-dénaturant, le bicarbonate d'ammonium (ABC) (Figure 48c). Une seule distribution d'état de charge est observée, et

celle-ci est étroite avec des états de charge faibles. L'utilisation d'un solvant non-dénaturant pour faire le sandwich permet donc de préserver la conformation native de la PEBP.

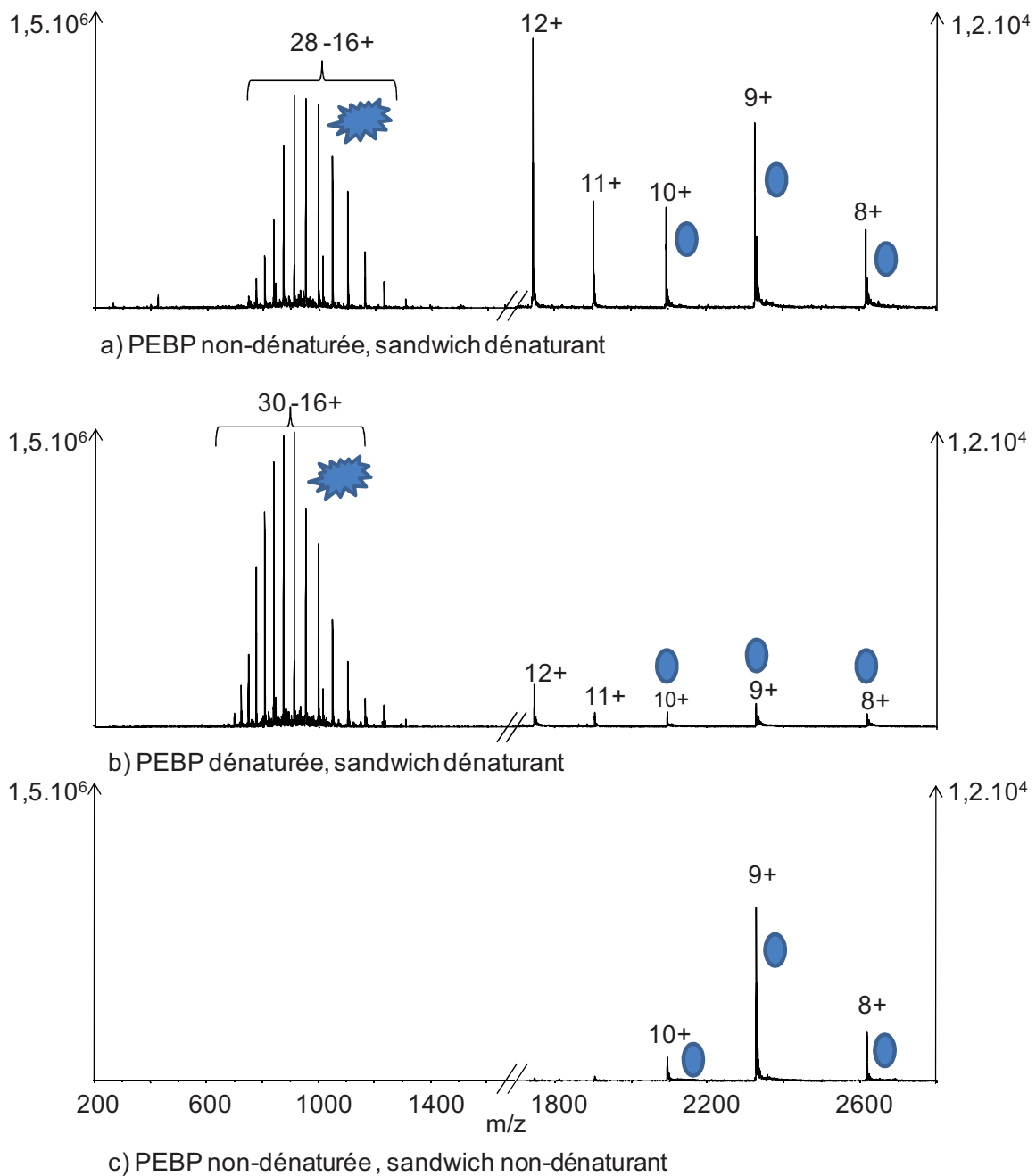


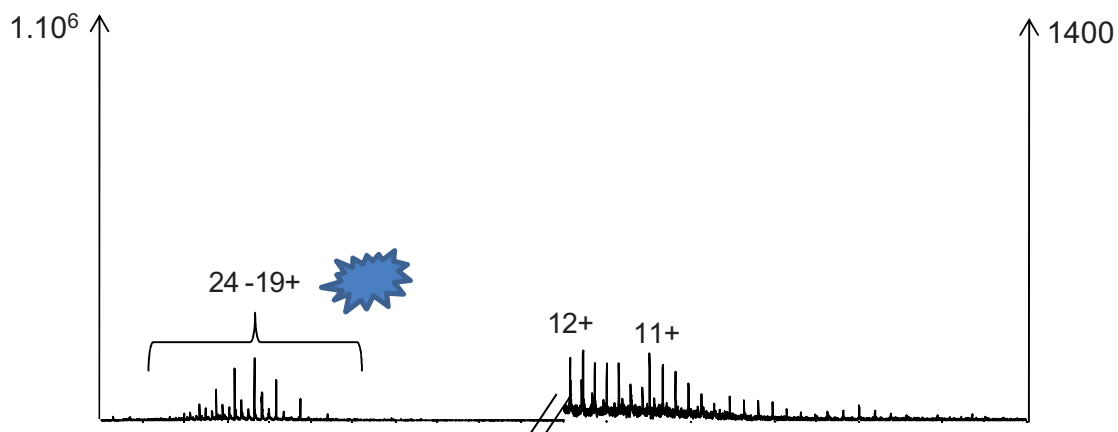
Figure 48: Effet de la composition du solvant sandwich sur l'état de PEBP et sur son ionisation. Deux spectres ESI-IT de la PEBP, enregistrés dans des gammes  $m/z$  différentes, ont été accolés. a) PEBP non-dénaturée avec un sandwich dénaturant, b) PEBP dénaturée avec un sandwich dénaturant, c) PEBP non-dénaturée avec un sandwich non-dénaturant. PEBP non-dénaturée = PEBP 18  $\mu\text{M}$  dans de l'ABC 20 mM. Sandwich non-dénaturant : ABC 20 mM. PEBP dénaturée : PEBP 1.8  $\mu\text{M}$  dans WAFA. Sandwich dénaturant : WAFA.

## 4.2. Sur le complexe hPEBP1 - BLA-C-04

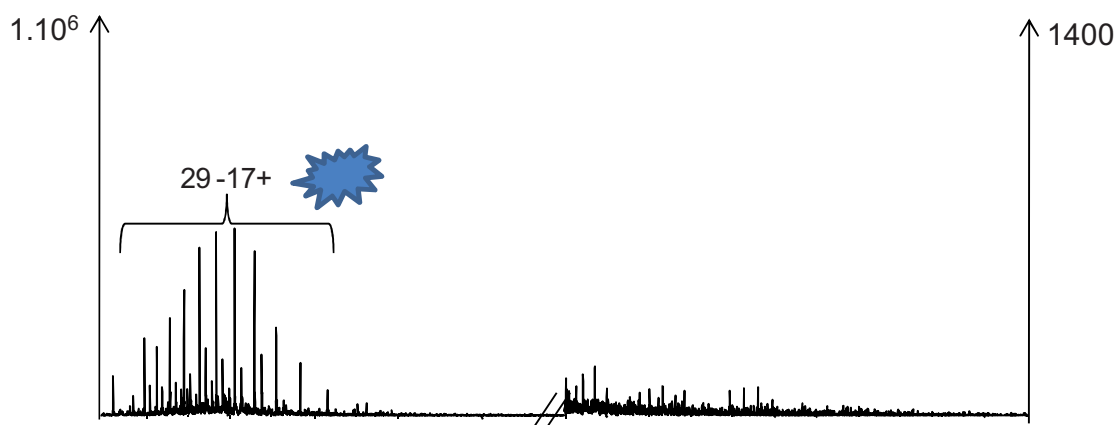
La même étude a ensuite été réalisée avec le complexe hPEBP1-BLA-C-04. La comparaison des spectres obtenus avec un solvant sandwich dénaturant ou non-dénaturant met également en évidence la nécessité de travailler avec un solvant sandwich non-dénaturant pour préserver la conformation native du complexe (Figure 49 au verso).

La méthode d'infusion en sandwich a l'avantage de miniaturiser l'infusion, c'est-à-dire de limiter la consommation d'échantillon et de faciliter le rinçage. Par ailleurs, l'utilisation d'un solvant sandwich non-dénaturant assure la préservation de la conformation native de la protéine.

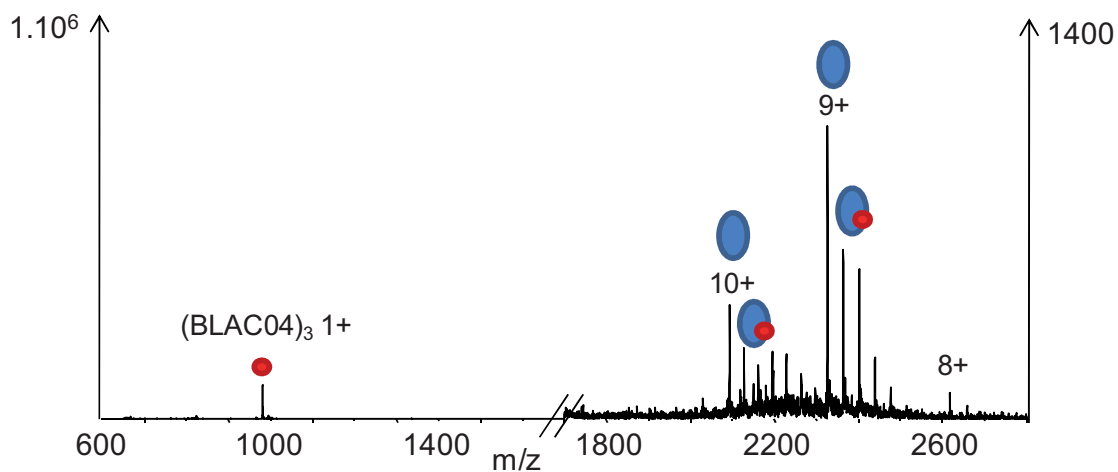
Toutes les analyses en MS native ont donc par la suite été réalisées avec cette méthode d'infusion en sandwich non-dénaturant, à un débit de 1.3  $\mu\text{L}/\text{min}$  et avec une tension d'aiguille de - 3400V.



a) PEBP-BLAC04 non-dénaturée, sandwich dénaturant



b) PEBP-BLAC04 dénaturée, sandwich dénaturant

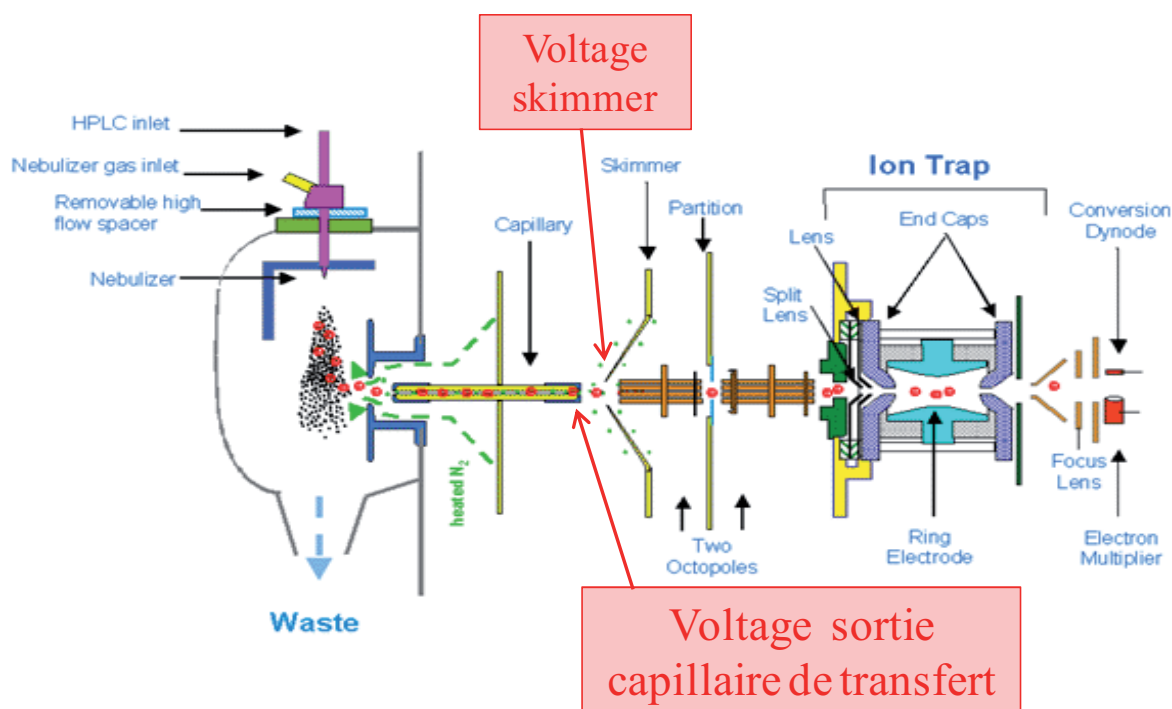


c) PEBP-BLAC04 non-dénaturée, sandwich non-dénaturant

Figure 49: Effet de la composition du solvant sandwich sur le complexe PEBP-BLA-C-04. Deux spectres ESI-IT du complexe PEBP-BLA-C-04, enregistrés dans des gammes  $m/z$  différentes, ont été accolés. a) complexe non-dénaturé avec un sandwich dénaturant, b) complexe dénaturé avec un sandwich dénaturant, c) complexe non-dénaturé avec un sandwich non-dénaturant.

## II. Optimisation des paramètres instrumentaux

Pour préserver les interactions non-covalentes, il est nécessaire d'optimiser la pression et la tension d'accélération à l'interface, paramètres qui permettent de contrôler l'énergie cinétique des ions. En ce qui concerne la trappe ionique Esquire HCT (Bruker, Bremen, Germany), il est difficile de faire varier la pression de façon simple. Nous n'avons donc pas changé sa valeur de 6.5 mbar, qui était déjà haute et donc optimale pour préserver les complexes non-covalents. Par contre, la tension d'accélération ( $V_c$ ), qui correspond à la différence de voltage entre la fin du capillaire de transfert ( $V_{CE}$ ) et le skimmer ( $V_{SK}$ ),  $V_c = V_{CE} - V_{SK}$ , a été optimisée.



Représentation schématique de l'instrument ESI-IT Esquire HCT de Bruker.

Des optimisations de tension d'accélération et de température de désolvation sont décrites pour deux systèmes, le dimère  $HU\alpha_2$  et complexe PEBP-GTP.

# 1. Application à HU

## 1.1. Optimisation de la tension d'accélération

L'étude du complexe HU est le sujet d'une collaboration avec Bertrand Castaing, responsable de l'équipe « Réparation de l'ADN : approches structurales et fonctionnelles », au Centre de Biophysique Moléculaire d'Orléans.

HU est une protéine présente à l'état de dimère en conditions physiologiques en solution de masse moyenne monomérique  $M_{\text{moy}}(\text{HU}\alpha) = 9535.0$  Da et dimérique  $M_{\text{moy}}(\text{HU}\alpha_2) = 19070$  Da. Vis et ses collaborateurs ont montré que pour observer HU sous forme dimérique, il est indispensable d'optimiser la pression à l'interface et la tension d'accélération (175).

Avant d'étudier l'effet de la variation de la tension d'accélération sur la préservation des interactions non-covalentes maintenant le dimère, des spectres du dimère  $\text{HU}\alpha_2$  dénaturé et du dimère non-dénaturé ont été obtenus avec le mode assisté d'EsquireControl (Figure 50). Le mode assisté permet d'utiliser la fonction « Smart Parameter Setting », qui optimise les voltages appliqués sur la cap exit, l'octopole 2 DC et l'octopole RF en fonction du rapport  $m/z$  ciblé. Les voltages appliqués sur le skimmer, l'octopole 1 DC, les lentilles 1 et 2, eux, restent constants et sont respectivement de 40, 8, - 5 et - 60 V. La fonction « Smart Parameter Setting » permet également d'ajuster l'optimisation en fonction de la stabilité du composé. Plus la stabilité est faible et plus le voltage de la cap exit est diminué.

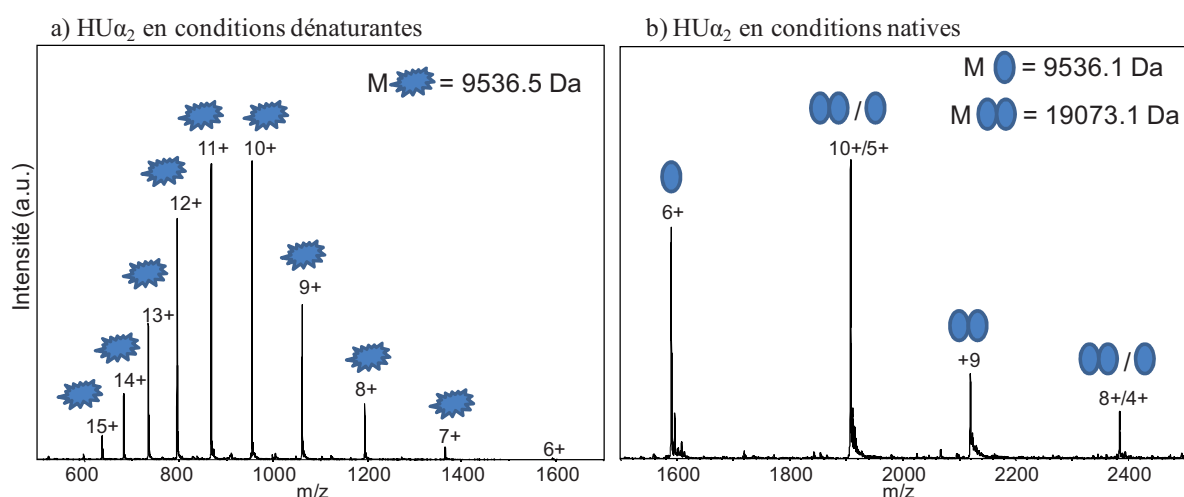


Figure 50: Spectres ESI-IT du dimère  $\text{HU}\alpha_2$  a) en conditions dénaturantes à 0.5  $\mu\text{M}$  dans WFA et b) en conditions natives à 10  $\mu\text{M}$  dans  $\text{NH}_4\text{OAc}$  50 mM pH 6.5.

En conditions dénaturantes, seul le monomère est visible, sa distribution d'états de charge est large (7+ à 15+) avec une intensité maximale pour le 10+. Le monomère est sous forme dépliée. En conditions natives, les pics les plus intenses sont dans la partie haute de la gamme de masse. Ils correspondent à du monomère ou du dimère en configuration native. Seul le pic à 2120  $m/z$  ne correspond qu'à du dimère, il correspond à l'état de charge 9+. Les pics à 1900 et 2350  $m/z$  peuvent correspondre à du dimère chargé 10+ et 8+ ou à du monomère chargé 5+ et 4+. Grâce à la présence d'adduits de potassium, on peut néanmoins déterminer qu'ils correspondent à un mélange de monomère et de dimère, le pic à 1900  $m/z$  correspondant à environ 50% de chaque entité et celui à 2350  $m/z$  ne correspondant pratiquement qu'à du monomère. Le pic à 1580  $m/z$  pourrait correspondre à du dimère 12+ ou du monomère 6+, cependant il n'y a pas de pic correspondant à du dimère 11+. Le pic à 1580  $m/z$  correspond donc forcément à du monomère chargé 6+.

Avec ces paramètres instrumentaux non-optimisés, l'intensité du pic caractéristique du dimère, le 9+, est inférieure à celle du pic caractéristique du monomère, le 6+.

Une fois les pics correspondant au dimère et au monomère ainsi identifiés, la tension d'accélération ( $V_c$ ) a été variée manuellement pour étudier son effet sur l'intensité du pic caractéristique du dimère, à 2120  $m/z$  (Figure 51).

Ces études ont été réalisées sur la gamme de masse 1800 – 2200  $m/z$ . L'analyse du spectre précédent (Figure 50) a mis en évidence que le dimère apparaît sous deux états de charge, 9+ et 10+. La variation de la tension d'accélération a pour but d'optimiser l'intensité d'un seul pic, celui qui ne correspond qu'à du dimère, à 2120  $m/z$ . Il est inutile d'optimiser celui à 1900  $m/z$  car la part de dimère dans ce pic ne peut pas être précisément connue. L'optimisation a donc été réalisée en cherchant une intensité relative du pic à 2120  $m/z$  par rapport à celui à 1900  $m/z$  la plus élevée possible.

La tension d'accélération peut être modulée en changeant la tension appliquée sur le skimmer ( $V_{SK}$ ) et/ou celle appliquée sur la fin du capillaire de transfert ( $V_{CE}$ ). De nombreux voltages ont été testés séparément sur le skimmer ou sur le cap exit. L'intensité relative du pic à 2120  $m/z$  par rapport à celui à 1900  $m/z$  est maximale pour deux couples  $V_{CE} / V_{SK}$  testés de 15/200 V et 100/200 V qui correspondent respectivement à deux tensions d'accélération différentes,  $V_c = 185$  V et  $V_c = 100$  V (Figure 51a et b)

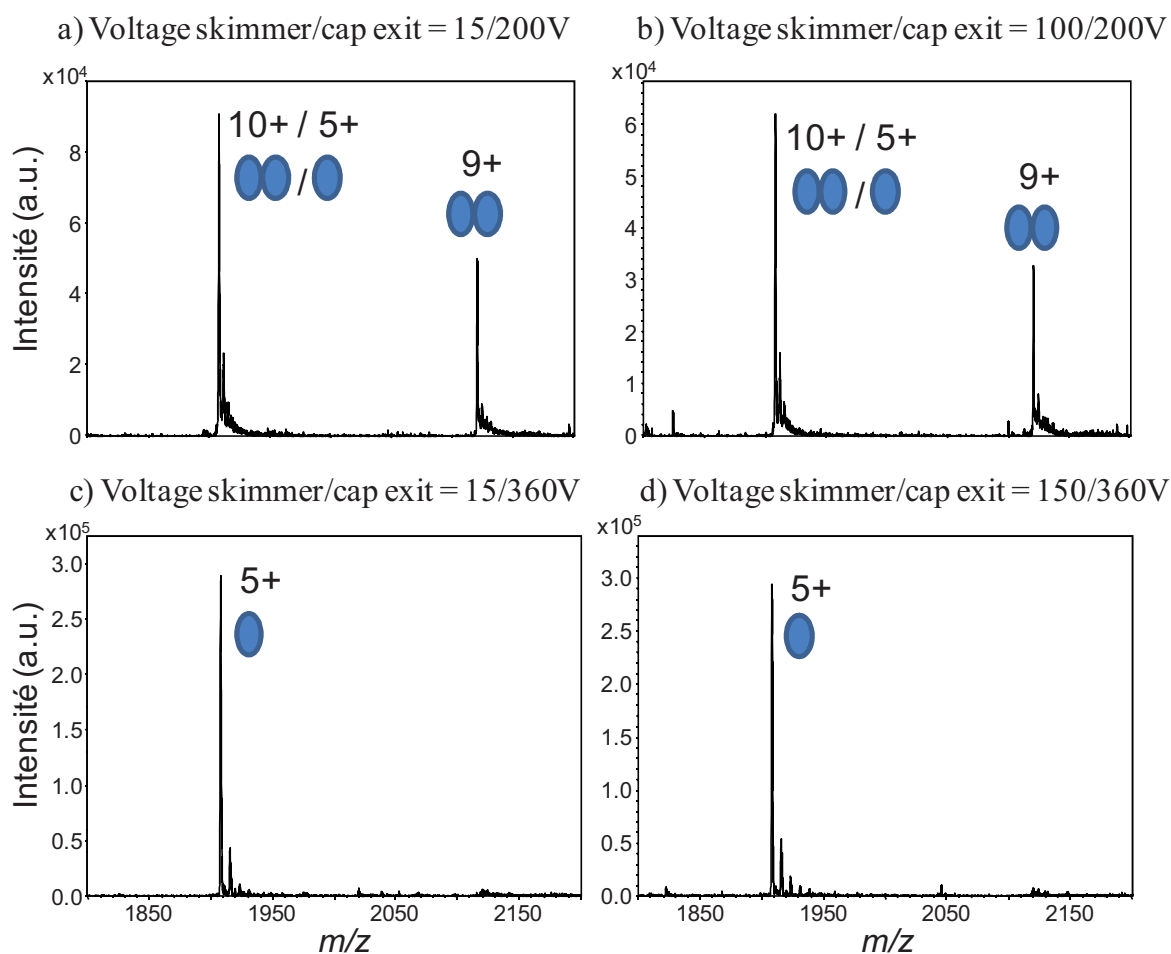


Figure 51 : Spectres ESI-IT du dimère  $HU\alpha_2$  en conditions natives à  $10 \mu\text{M}$  dans  $\text{NH}_4\text{OAc}$  50 mM pH 5 avec un couple  $V_{\text{SK}}/V_{\text{CE}}$  de a) 15/200V ( $V_C = 185 \text{ V}$ ), b) 100/200V ( $V_C = 100\text{V}$ ), c) 15/360V ( $V_C = 345\text{V}$ ) et d) 150/360V ( $V_C = 210\text{V}$ )

En revanche, avec les couples 15/360 V et 150/360 V, le dimère est pratiquement complètement dissocié (Figure 51 c et d). La variation du voltage du skimmer seul ne semble pas avoir d'effet sur la préservation du complexe. Seule la variation du voltage de la cap exit semble permettre d'optimiser cette préservation. Pour valider cette hypothèse, l'intensité relative du pic à 2120  $m/z$  par rapport à celui à 1900  $m/z$  a été suivie en faisant varier le voltage de la cap exit avec un voltage de skimmer fixe de 100V. La dissociation du dimère est minimale lorsque le voltage de la cap exit est d'environ 200V (Figure 52).



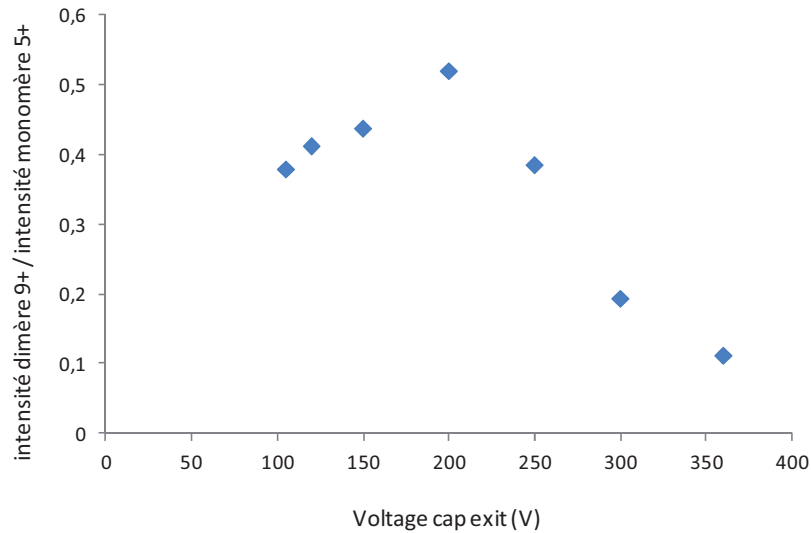


Figure 52: Effet du voltage de la cap exit sur la quantité relative de dimère par rapport au monomère (avec un  $V_{SK}$  fixe de 100 V)

Pour vérifier que le voltage appliqué sur le skimmer n'avait pas d'incidence sur la dissociation du dimère, la même étude a été réalisée en fixant le voltage du skimmer à 15V. (Figure 53)

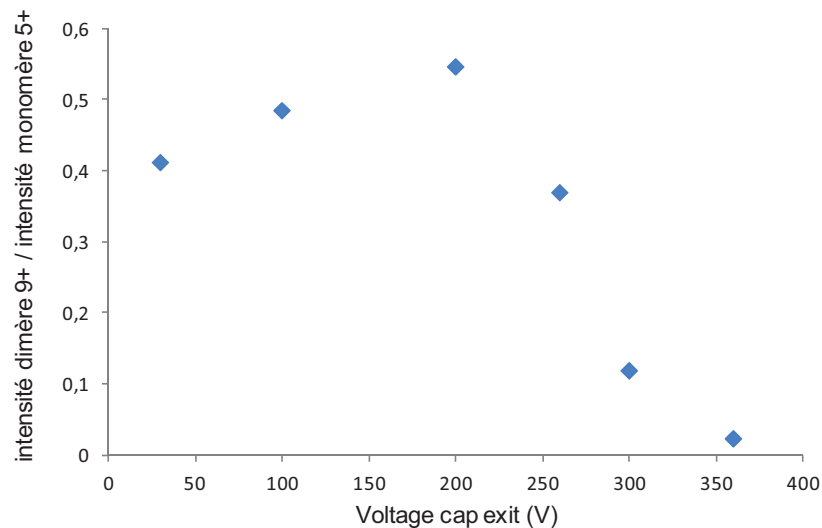


Figure 53 : Effet du voltage de la cap exit sur la quantité relative de dimère 9+ par rapport au monomère 5+ ( $V_{SK}$  fixé à 15V)

La dissociation est de nouveau minimale lorsque le voltage appliqué à la cap exit est de 200 V. Pour deux valeurs de voltage appliqué sur le skimmer, 15 et 100 V, le voltage optimal à appliquer sur le cap exit est de 200 V. D'autres voltages de skimmer, intermédiaires entre 15

et 100 V, ont été testés et ont confirmé que le voltage optimal à appliquer à la cap exit est de 200 V.

## 1.2. Optimisation de la température de désolvatation

Deux acquisitions ont été réalisées à des températures de désolvatation de 120°C et 365°C, avec un voltage appliqué à la cap exit fixe de 200V (Figure 54). A 365°C, l'intensité relative du pic caractéristique du dimère, à 2120  $m/z$ , par rapport à celle du pic à  $m/z$  1900, est moins élevée qu'à 120°C.

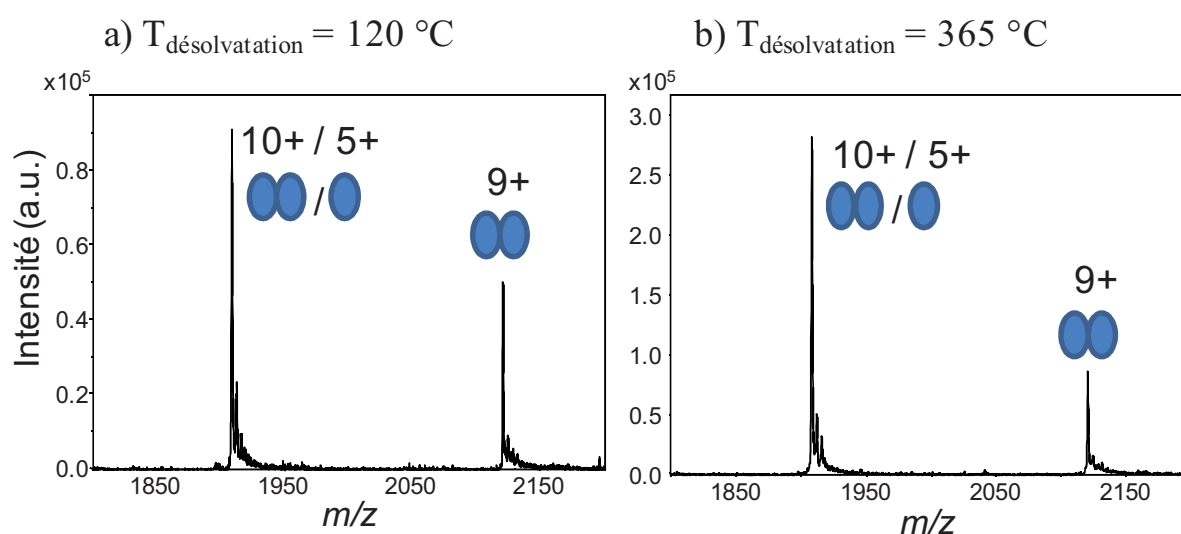


Figure 54: Spectres ESI-IT du dimère HU $\alpha_2$  en conditions natives à 10  $\mu$ M dans NH $_4$ OAc 50 mM pH 5 avec un couple  $V_{SK}/V_{CE}$  de 100/200V et une température de désolvatation de a) 120°C, b) 365°C

En conclusion, les conditions les plus douces pour étudier les dimères HU sont un voltage de la cap exit égal à 200V et une température de désolvatation de 120°C. Une température inférieure à 120°C n'a pas été testée pour éviter un encrassement trop rapide de la source.

L'optimisation des paramètres instrumentaux sur le système hPEBP1-GTP, archétype des complexes hPEBP1 - nucléotide qui composent le cœur de mon sujet de thèse, est présentée ci-après.

## 2. Application à hPEBP1 – GTP

Comme pour le système HU, le spectre du complexe a d'abord été enregistré sans optimisation particulière, avec le mode assisté d'Esquire Control, (Figure 55a) puis avec une forte tension d'accélération, pour étudier son effet sur la fraction de liaison hPEBP1-GTP/hPEBP1<sub>tot</sub> (Figure 55b).

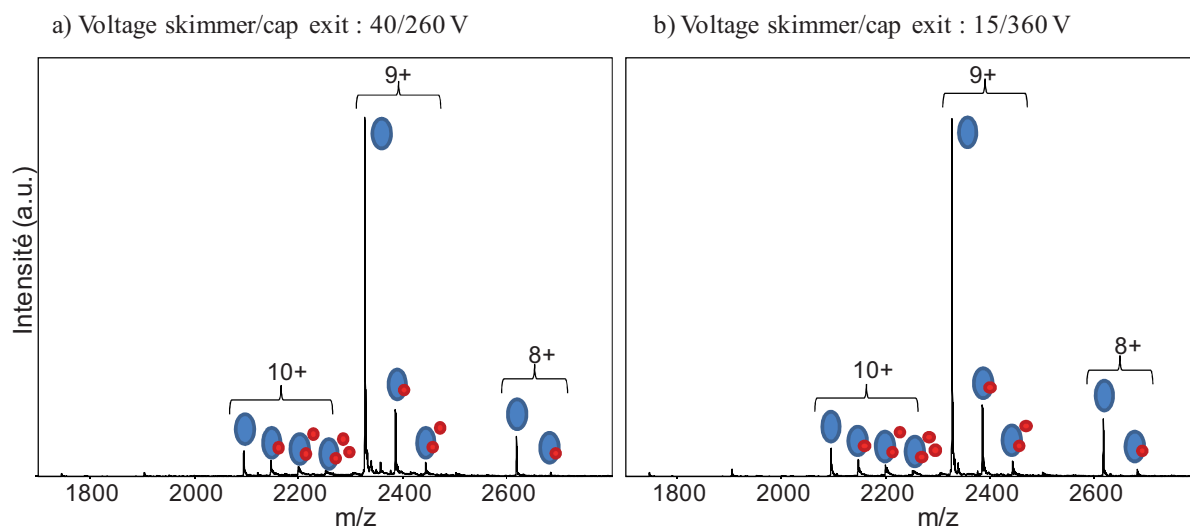


Figure 55: Spectres de masses ESI-IT de 12.6  $\mu\text{M}$  de hPEBP1 avec 40  $\mu\text{M}$  de GTP (delta-masse = + 523 Da) après 8 minutes d'incubation à 37°C dans l'ABC 20 mM, pH 7.4, a) avec  $V_{\text{SK/CE}}= 40/260$  V b) avec  $V_{\text{SK/CE}}= 15/360$  V

Des fractions de liaison ont été calculées à partir de la figure 55 et sont présentées dans le Tableau 4.

Tableau 4: Fraction de liaison du complexe hPEBP1-GTP en fonction de l'état de charge et de la tension d'accélération

	Int(hPEBP1-GTP) / Int(hPEBP1 <sub>tot</sub> )			
	10+	9+	8+	Moyenne pondérée
Voltage skimmer/cap exit : 40/260 V	0.27	0.16	0.10	0.17
Voltage skimmer/cap exit : 15/360 V	0.27	0.16	0.12	0.17

L'application d'une tension d'accélération la plus élevée possible entre la fin du capillaire et le skimmer ( $V_C = 345$  V) (Figure 55b) ne dissocie pas le complexe hPEBP1-GTP. Alors que

cette tension est très dissociante pour le complexe  $HU\alpha_2$ , elle ne l'est pas pour le complexe PEBP-GTP.

Pour valider ces premières observations, la fonction SmartRamp d'EsquireControl, qui permet d'étudier l'effet de la variation d'un paramètre sur des ions sélectionnés, a été utilisée.

L'objectif a été d'optimiser le signal du complexe hPEBP1-GTP, de stœchiométrie 1:1, par rapport à celui de hPEBP1 libre. Les ions les plus intenses, les états de charge 9+ du complexe et de hPEBP1 libre, respectivement de  $m/z$  2384.2 et 2326, ont été sélectionnés pour faire l'étude.

Dans un premier temps, le voltage du skimmer est resté fixé à 40V, valeur donnée par défaut par Esquirecontrol et le paramètre optimisé a été le voltage de la cap exit (Figure 56).

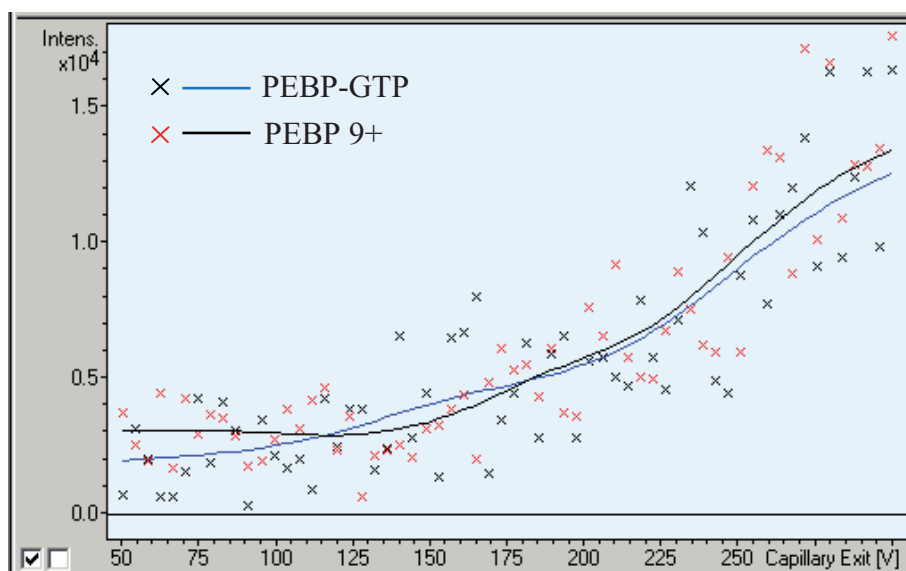


Figure 56: Intensités des états de charge 9+ de hPEBP1 et du complexe hPEBP1-GTP en fonction du voltage appliqué sur le cap exit avec un voltage skimmer fixé à 40 V.

Lorsque l'on cherche à détecter des ions autour de  $m/z$  2250, plus on augmente le voltage de la cap exit et plus on augmente l'intensité du signal du complexe et de hPEBP1 (Figure 56). L'intensité du signal augmente fortement dès que le voltage de la cap exit est supérieur à 200 V (Figure 56) et est maximum à  $V_{CE} = 360$  V

La même optimisation a été réalisée mais en fixant le voltage du skimmer à 100 V au lieu de 40 (Figure 57).

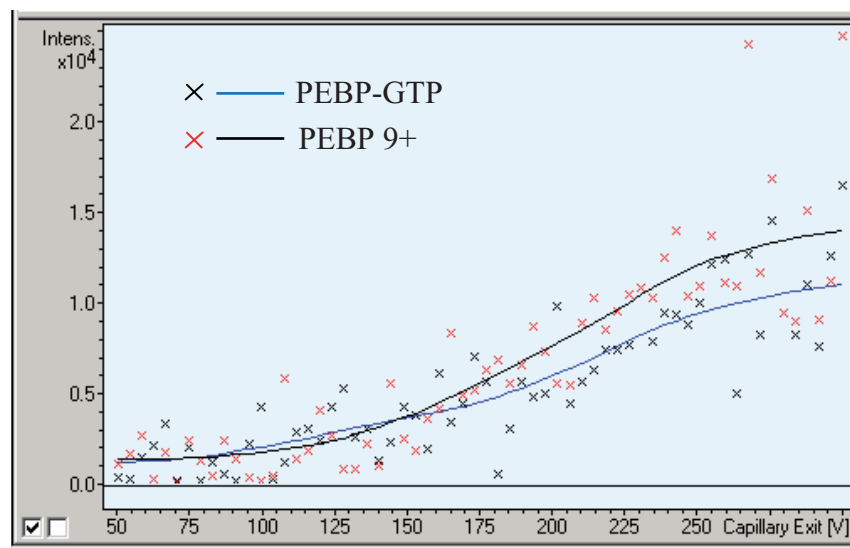


Figure 57: Intensités des états de charge 9+ de hPEBP1 et du complexe hPEBP1-GTP en fonction du voltage appliqué sur le cap exit avec un voltage skimmer fixé à 100 V.

Lorsque le voltage du skimmer est fixé à 100 V, l'intensité relative des deux composés varie significativement avec le voltage de la cap exit (Figure 57). A partir de  $V_{CE} = 150$  V, l'intensité du pic de hPEBP1 est plus élevée que celle du complexe. L'augmentation du voltage de la cap exit entraîne une dissociation du complexe, phénomène qui n'était presque pas observé avec un voltage skimmer de 40 V.

Pour éviter toute dissociation supplémentaire et préserver au maximum les complexes, les complexes de type hPEBP1-nucléotides ont été analysés avec un voltage de skimmer de 40 V et un voltage de cap exit de 260V. Ces voltages n'ont pas été optimisés pour chaque nouvel analogue de nucléotide afin que le criblage soit réalisé dans les mêmes conditions instrumentales pour tous les analogues.



# Chapitre II

Application à des systèmes non-covalents

*Résumé du chapitre II:*

*Outre les complexes hPEBP1-analogues de nucléotide, différents systèmes non-covalents d'intérêt thérapeutique en cancérologie ou dans les maladies pulmonaires ont été étudiés dans le cadre de collaborations. Deux systèmes font l'objet d'une étude purement qualitative pour valider par MS l'existence d'une interaction non-covalente. La MS ainsi a prouvé une interaction originale entre une protéine fortement hydrophobe, la protéinase PR3, et un azapeptide, ce qui a donné lieu à une publication. L'existence des complexes non-covalents entre le zinc et les protéines réparatrices de l'ADN, Fpg et NEIL2, initialement observée par des méthodes biochimiques, est également confirmée. De plus, un complexe ternaire Fpg-Zn-ADN a pu être observé entièrement intact lors de l'analyse MS.*

*La protéine HU est une protéine d'architecture de l'ADN qui contribue à l'action de protéines de réparation. Les dimères non-covalents  $HU\alpha_2$ ,  $HU\beta_2$  et  $HU\alpha\beta$  font l'objet d'une étude quantitative. Pour ces complexes, le pH et la concentration ionique du tampon ont été optimisés. La formation d'hétérodimère  $HU\alpha\beta$ , à partir des homodimères  $HU\alpha_2$  et  $HU\beta_2$ , est pratiquement instantanée à 37°C. A 27°C, la réaction peut être ralentie et le suivi des différents dimères au cours du temps semble valider un mécanisme de réaction original nécessitant des formes partiellement dénaturées des dimères. Il est également mis en évidence, que, contrairement aux hétérodimères inter-espèces, l'hétérodimère  $HU\alpha\beta$  d'*E.coli* est stable au cours du temps.*



La maîtrise des optimisations expérimentales et instrumentales pour l'étude de complexes non-covalents par spectrométrie de masse a été mise au service de l'analyse de complexes protéine-ligand différents des complexes PEBP-nucléotides, dans le cadre de plusieurs projets de recherche.

Trois collaborations ont ainsi été développées avec l'équipe « Réparation de l'ADN : approches structurales et fonctionnelles » de Bertrand Castaing, au Centre de Biophysique Moléculaire d'Orléans d'une part, et avec l'équipe « Protéases et pathologies pulmonaires » de Francis Gauthier à l'unité INSERM U618 - Faculté de Médecine d'autre part.

Les trois collaborations concernent les systèmes non-covalents suivants :

- protéase/azapeptide (équipe Francis Gauthier, collaborateur : Christophe Epinette)
- glycosylase/zinc/ADN/drogue (équipe Bertrand Castaing, collaborateur : Arthur Biela)
- dimères de la protéine HU, protéine ressemblant aux histones (équipe B. Castaing, collaborateur : B. Castaing).

Les deux premiers systèmes ont été étudiés d'un point de vue qualitatif pour valider une interaction non-covalente.

Le système HU a été exploré d'un point de vue cinétique et a nécessité des études qualitatives, quantitatives et cinétiques.

## I. Etudes qualitatives

### 1. PR3-azapeptide

#### 1.1. Contexte général

La protéinase 3 (PR3), connue aussi sous le nom de myéloblastine et d'antigène de Wegener, est une protéase à sérine. Elle est présente dans les granules primaires des granulocytes polymorphonucléaires neutrophiles et dans les monocytes, qui sont capables de protéolyser de

nombreux substrats biologiques. La PR3 dégrade une variété de protéines de la matrice extracellulaire, telles que l'élastine, la fibronectine, le collagène de type IV. La PR3 existe sous forme soluble ou associée à la membrane et constitue l'antigène principal de la granulomatose de Wegener, vascularite systémique nécrosante la plus connue chez les adultes. La structure de PR3 est stabilisée par quatre ponts disulfures et, dans une moindre proportion, par des glycosylations sur l'asparagine.

La PR3 de neutrophile ressemble structurellement et fonctionnellement à l'élastase de neutrophile (176, 177), néanmoins ses fonctions biologiques spécifiques ne sont pas clairement connues. Le rôle physiopathologique de la PR3 pourrait être mieux compris si des inhibiteurs spécifiques de PR3 étaient disponibles. Ceux-ci pourraient de plus être éventuellement utilisés en tant qu'agents thérapeutiques. Cependant, tous les inhibiteurs naturels de PR3 ciblent préférentiellement l'élastase de neutrophile (177, 178). Aucun inhibiteur connu, qu'il soit d'origine naturelle ou de synthèse, n'inhibe sélectivement la PR3.

Les azapeptides sont des molécules dans lesquelles l'atome de carbone d'une liaison peptidique est remplacé par un atome d'azote (Figure 58) (179, 180).

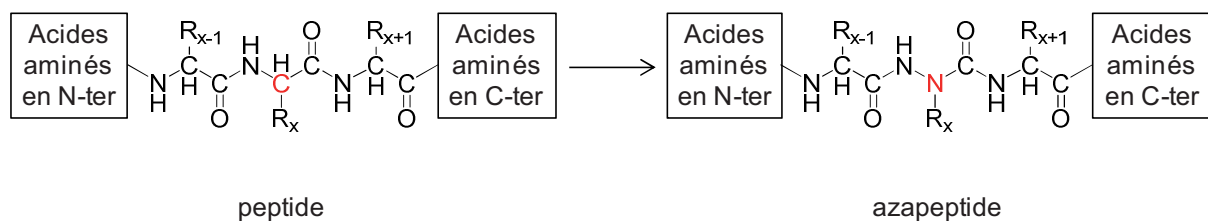


Figure 58: Structure du peptide comparée à celle de l'azapeptide

Ils sont plus stables que les peptides et ont déjà été utilisés en tant qu'inhibiteurs (180). Les azapeptides gardent les principales caractéristiques du peptide parent. L'équipe de F. Gauthier a utilisé ses connaissances sur la spécificité des substrats de PR3 pour modifier un substrat FRET spécifique de PR3. Les analyses FRET sur des substrats de séquence variable ont permis d'identifier les acides aminés interagissant le mieux avec la protéase des deux côtés du site de coupure, pour aboutir à un substrat se liant avec un  $K_m$  de l'ordre du micromolaire.

Nos collaborateurs ont émis l'hypothèse qu'en transformant un tel peptide-substrat spécifique de la PR3 en azapeptide, cela pourrait donner un inhibiteur compétitif qui inhiberait spécifiquement PR3 avec un  $K_i$  de l'ordre du micromolaire.

Les azapeptides contiennent généralement un groupe réactif labile leur permettant d'interagir avec la protéase-cible pour former un intermédiaire acyl-enzyme stable qui se dissocie très lentement (Schéma 1) (180).

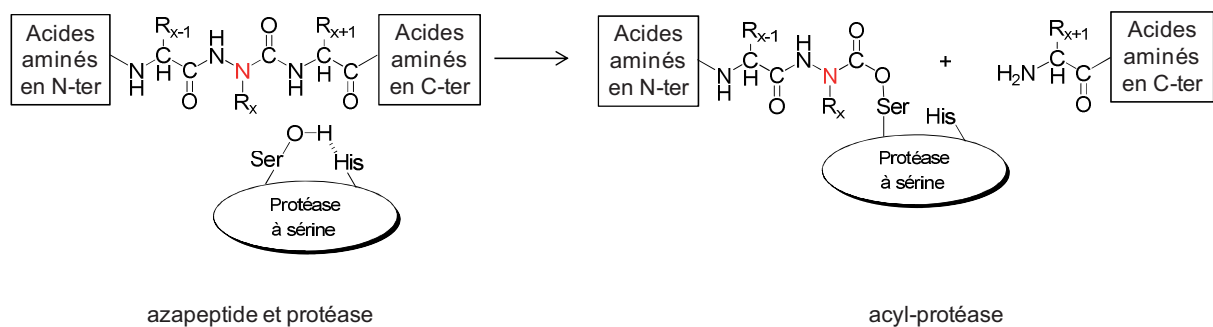


Schéma 1 : Génération d'acyl-protéase à partir d'azapeptide

L'interaction de cet azapeptide avec la PR3 a été étudiée dans l'objectif de déterminer le mécanisme d'inhibition. D'après les travaux de F. Gautier et ses collaborateurs, contrairement à la plupart des azapeptides, le mécanisme d'inhibition ne semble pas impliquer la formation d'un complexe acyl-enzyme stable (Schéma 1) mais relèverait plutôt d'un mécanisme de compétition réversible.

Les expériences qu'a effectuées l'équipe en cinétique enzymatique montrent que l'inhibition de l'enzyme apparaît rapidement, dès quelques minutes. Après une incubation prolongée, l'azapeptide semble cependant être dégradé par sa cible (181).

## 1.2. Etude du complexe

Dans ce contexte, notre travail a consisté à analyser en conditions natives le complexe PR3-azapeptide immédiatement après sa formation. Le premier objectif a été d'optimiser les paramètres instrumentaux afin d'observer la protéine PR3 par spectrométrie de masse. Cette étape est *a priori* compliquée par la présence d'un patch hydrophobe à la surface de PR3, qui impose qu'elle soit traitée d'une manière similaire à une protéine membranaire. Une fois cet

objectif atteint, le complexe a été formé en solution et analysé par MS native pour mettre en évidence la présence d'un édifice non-covalent et pour déterminer si l'azapeptide lié était dégradé ou non.

#### 1.2.1. Echange de tampon

La PR3 se trouvait initialement dans une solution d'HEPES 50 mM, pH 7.4, NaCl 150 mM contenant du détergent NP-40. Un échange de tampon contre du tampon ABC 20 mM a dû être effectué avant l'analyse en spectrométrie de masse. L'échange de tampon réalisé par HPLC ne permet cependant pas d'obtenir un signal par MS sur la trappe ionique Esquire HCT. La protéine PR3 n'est pas visible en conditions natives car le signal du détergent NP-40 résiduel domine le spectre.

Lors du rinçage avec de l'ABC 20 mM, un signal correspondant à la PR3 native est néanmoins apparu par simple effet de dilution de l'échantillon, conduisant à une concentration de NP-40 effective plus favorable. Le signal est proche du bruit de fond mais peut être amélioré en augmentant le débit du gaz séchant de 5 à 10 L/ min.

Afin d'obtenir un signal durant l'infusion de l'échantillon et non lors du rinçage, un autre instrument a été testé.

#### 1.2.2. Optimisations instrumentales

La PR3 a été analysée avec le maXis, ESI-Q-TOF à haute résolution. Contrairement à la trappe ionique Esquire HCT, un signal est visible dès l'injection de l'échantillon, et il n'est pas nécessaire d'attendre l'étape de rinçage. Le signal sur bruit est faible lorsque le débit du gaz séchant est de 4 L/min et fortement amélioré avec un débit de gaz séchant de 10 L/min (Figure 59). Des deltas masse de 162 Da sont observés sur la protéine. Il s'agit de glycosylations hétérogènes, modifications post-traductionnelles qui ont déjà été décrites dans la littérature (182).

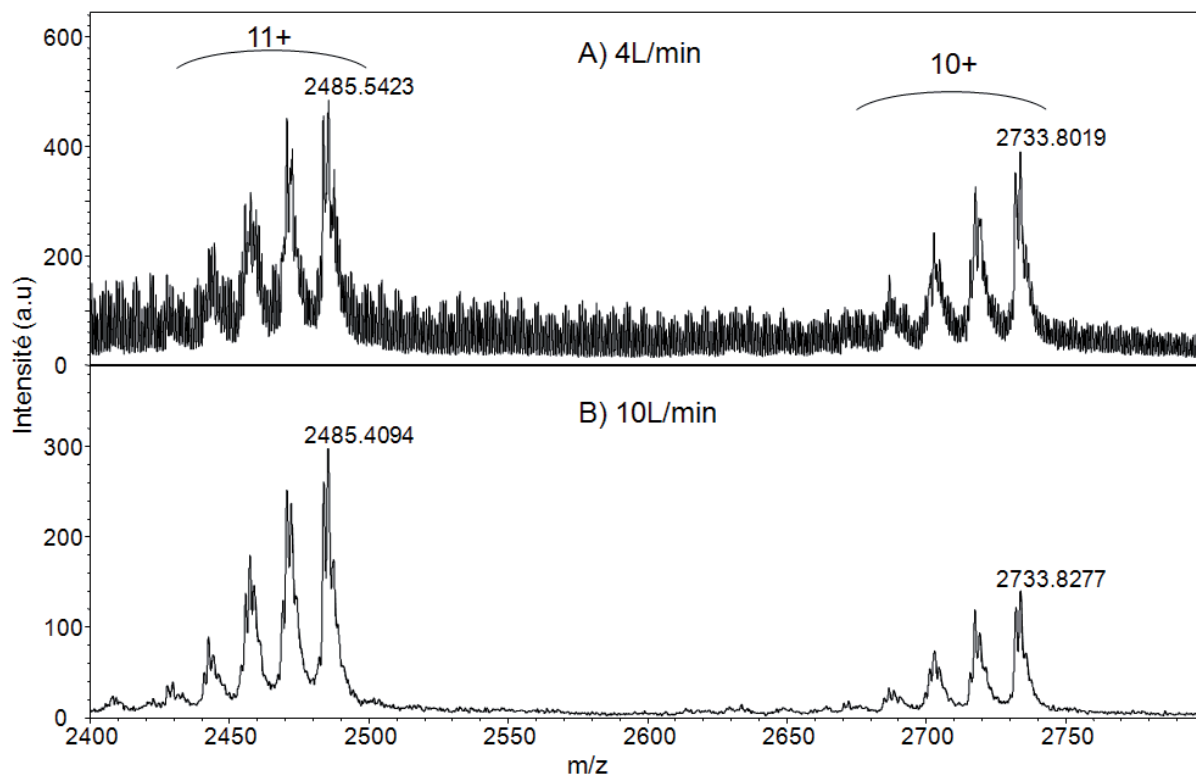


Figure 59: Spectres ESI-IT de 3 μM de PR3 dans l'ABC 20 mM, pH 8 avec un débit de gaz séchant de A) 4 L/min, B) 10 L/min.

### 1.2.3. Résultats

L'azapeptide a ensuite été incubé en solution avec la PR3 avant d'être analysé dans les conditions précédemment optimisées (Figure 60).

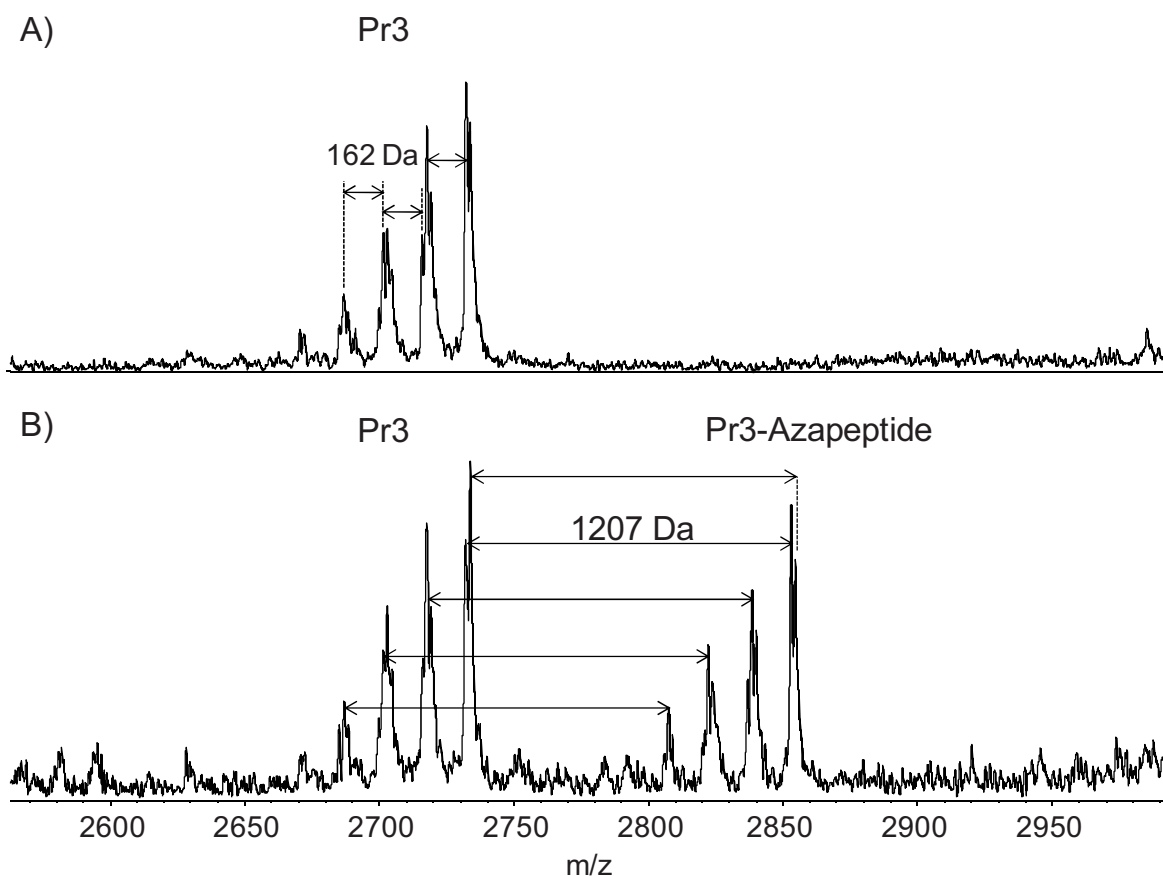


Figure 60 : Spectres ESI-IT dans l'ABC 20 mM, pH 8 de A) PR3 3  $\mu$ M, B) PR3 3  $\mu$ M incubée avec azapeptide 40  $\mu$ M. Seul l'état de charge 10+ est montré.

La PR3 seule ou en complexe apparaît sous trois états de charge, 11+, 10 et 9+. Seul l'état de charge 10+ est représenté ici (Figure 60). En présence d'azapeptide, il y a bien formation d'un complexe non-covalent avec la PR3. L'azapeptide est intact car le delta masse mesuré correspond exactement à la masse du peptide. Ici aussi des delta-masse de 162 Da correspondant à la glycosylation hétérogène de PR3 sont observés sur la protéine seule, et sur le complexe.

La technique de spectrométrie de masse a donc prouvé la formation d'un édifice non-covalent entre la PR3 et l'azapeptide et mis en évidence que l'azapeptide n'était pas dégradé dans ces conditions. Ce travail s'est intégré dans la publication intitulée « A selective reversible azapeptide inhibitor of human neutrophil proteinase 3 derived from a high affinity FRET substrate », qui est actuellement en presse pour *Biochemical Pharmacology* (181).

## 2. Fpg-Zn / NEIL2-Zn

### 2.1. Contexte

La protéine Fpg (Formamidopyrimidine-DNA glycosylase) de *Lactococcus lactis* et son orthologue humaine NEIL2 sont des glycosylases qui initient la première étape de réparation de l'ADN par excision de base.

La réparation par excision de base (ou BER pour Base excision repair en anglais) est un mécanisme de réparation d'un dommage au niveau d'une base individuelle de l'ADN. Un tel dommage est réparé par simple élimination de la base, suivi du clivage du désoxyribose, puis d'une nouvelle synthèse. Dans ce processus, une ADN glycosylase élimine la base endommagée et une endonucléase clive le désoxyribose. Une ADN polymérase remplit à nouveau l'espace libéré en utilisant la base opposée comme matrice.

La Fpg ou Formamidopyrimidine-DNA glycosylase est une enzyme particulière qui cumule les trois activités enzymatiques (183):

- Activité ADN-glycosylase : elle excise des bases endommagées par oxydation et produit ainsi un site apyrimidique/apuriques appelé site AP.
- Activité lyase AP 3' : elle clive la liaison phosphodiester 3'
- Activité lyase AP 5' : elle clive la liaison phosphodiester 5'

Fpg a plus d'affinité pour les purines oxydées (183). C'est un monomère constitué de deux domaines connectés entre eux par une zone charnière flexible. Elle possède deux motifs de liaison à l'ADN, le motif doigt de zinc et le motif hélice/tour/hélice/tour/hélice (Figure 61).

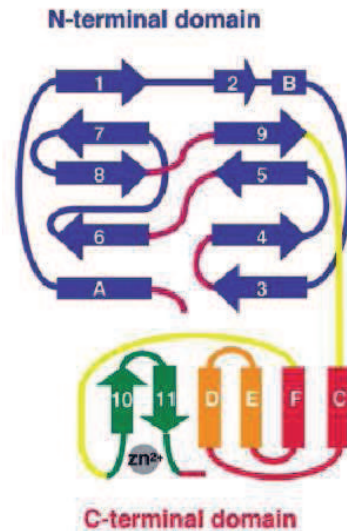


Figure 61 : Diagramme topologique de Fpg. Les brins beta sont représentés par des flèches. Les hélices alpha forment un motif hélice/tour/hélice/tour/hélice (figure reproduite de Sugahara *et al.* (184))

Le domaine à doigt de zinc est constitué de quatre cystéines conservées (deux avant le brin beta 10 et deux après le brin beta 11), du feuillet beta 10/11 et d'une molécule de zinc (Figure 61) (184).

L'équipe de B. Castaing travaille sur un inhibiteur de Fpg, la 2-thioxanthine. Les inhibiteurs d'enzyme de réparation de l'ADN peuvent être des outils utiles pour la potentialisation de la chimiothérapie du cancer ou être utilisés dans des études mécanistiques de la reconnaissance des dommages de l'ADN. L'inhibition provoquée par la 2-thioxanthine relève d'un mécanisme non-compétitif: la drogue est uniquement capable de se lier au complexe Fpg-ADN, et pas à la protéine libre (185). Cela suggère que la liaison de Fpg à l'ADN change significativement la conformation de la protéine, ce qui induit ensuite la reconnaissance d'un site par l'inhibiteur (185). Amara et ses collaborateurs ont effectivement mis en évidence que la Fpg subit de forts changements dynamiques suite à la liaison d'ADN endommagé (186).

D'après des premiers résultats obtenus par cristallographie, il semble que deux molécules de 2-thioxanthine se fixent sur deux cystéines de la zone conservée de Fpg. Des travaux sont réalisés actuellement pour savoir si la drogue peut également inhiber l'activité de NEIL-2, qui possède un motif à doigt de zinc légèrement différent de Fpg, avec 3 cystéines et une histidine. S'il n'y a pas d'inhibition, l'équipe réalisera des mutants de Fpg ou de NEIL-2 pour



identifier les résidus qui empêchent l'action de la drogue. Le criblage des mutants pourrait dans ce cas être réalisé par MS.

## 2.2. Objectifs

Dans ce contexte, l'équipe de B. Castaing a voulu confirmer les observations faites par cristallographie et biochimie avec une méthode complémentaire, la spectrométrie de masse supramoléculaire.

Pour disséquer et comparer par MS le mécanisme d'interaction de la drogue 2-thioxanthine avec soit Fpg, soit avec NEIL-2, trois objectifs ont été fixés:

- 1) confirmer l'existence d'édifices non-covalents Fpg/Zn et NEIL-2/Zn (187, 188): pour cela, les conditions expérimentales et les paramètres instrumentaux doivent être optimisés pour préserver les complexes protéine/métal.
- 2) vérifier que la drogue se fixe uniquement à Fpg complexée à l'ADN et non à Fpg libre
- 3) caractériser le mécanisme d'interaction de la drogue en déterminant s'il y a perte de l'atome de zinc ou modification de la conformation de la protéine.

Cet objectif nécessite d'optimiser les conditions expérimentales et les paramètres instrumentaux pour observer et préserver des complexes ternaires hydrides de type protéine/métal/ADN ou quaternaires de type protéine/métal/ADN/drogue.

## 2.3. Optimisations pour les complexes protéine/métal

Dans un premier temps, les paramètres du maXis ont été optimisés pour observer le complexe Fpg-Zn. Le paramètre de fragmentation dans la source isCID est fixé à 200 eV. L'augmentation du débit de gaz séchant, de 4 à 10 L/min permet d'obtenir une distribution d'états de charge plus étroite, avec des états de charge compris entre 11 et 14+ (Figure 62)

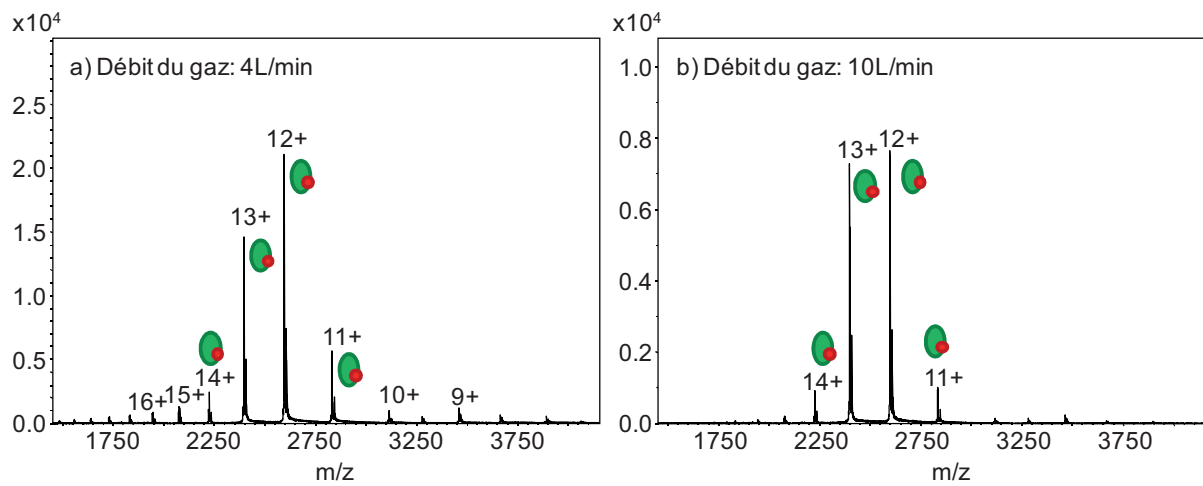


Figure 62: Spectres ESI-UHR-Q-TOF (maXis) de Fpg-Zn 10  $\mu$ M en conditions natives dans  $\text{NH}_4\text{OAc}$  50 mM pH 6.5 avec un débit de gaz a) 4 L/min et b) 10 L/min.

## 2.4. Confirmation d'édifices non-covalents protéine-zinc

### 2.4.1. Analyse de Fpg-Zn en conditions natives

Les espèces observées sont réparties sous quatre états de charge. Le zoom de l'état charge 12+ montre qu'il y a trois espèces (Figure 63).

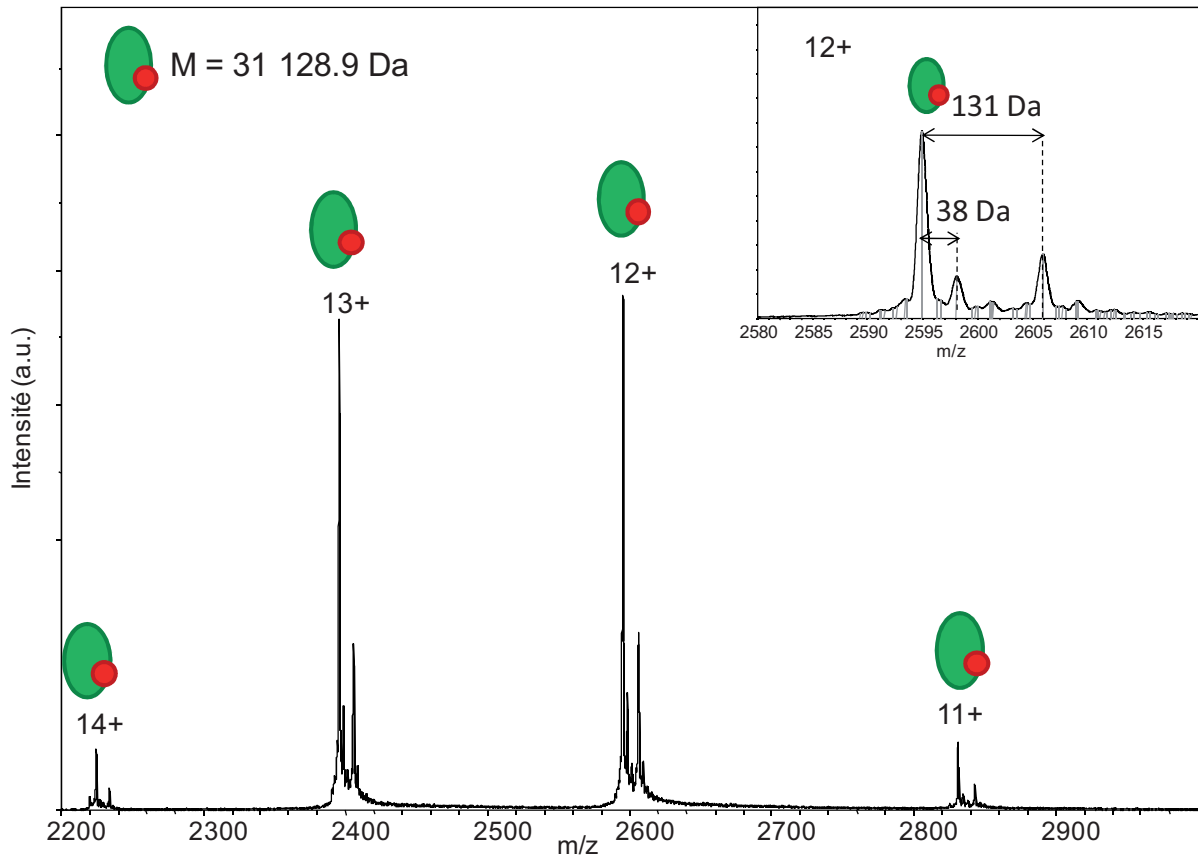


Figure 63: Spectre ESI-UHR-Q-TOF (maXis) de Fpg-Zn 10  $\mu$ M en conditions natives dans  $\text{NH}_4\text{OAc}$  50 mM pH 6.5 avec un débit de gaz de 10L/min.

D'après la masse théorique de la protéine ( $M = 31064$  Da) et du zinc ( $M = 65.4$  Da), le premier pic du massif chargé 12+ correspond au complexe Fpg-Zn, le deuxième, à ce complexe avec un adduit de potassium, et le troisième, au complexe du Zn avec la forme de Fpg qui comporte encore la méthionine initiale (elle-même sujette à l'adduction de potassium).

Les conditions ainsi mises au point permettent de mettre en évidence l'interaction Fpg-Zn en analyse MS native.

#### 2.4.2. Analyse de Fpg-Zn en conditions dénaturantes

Le complexe a ensuite été testé en conditions dénaturantes pour tester la nature, non-covalente ou non, de l'édifice.

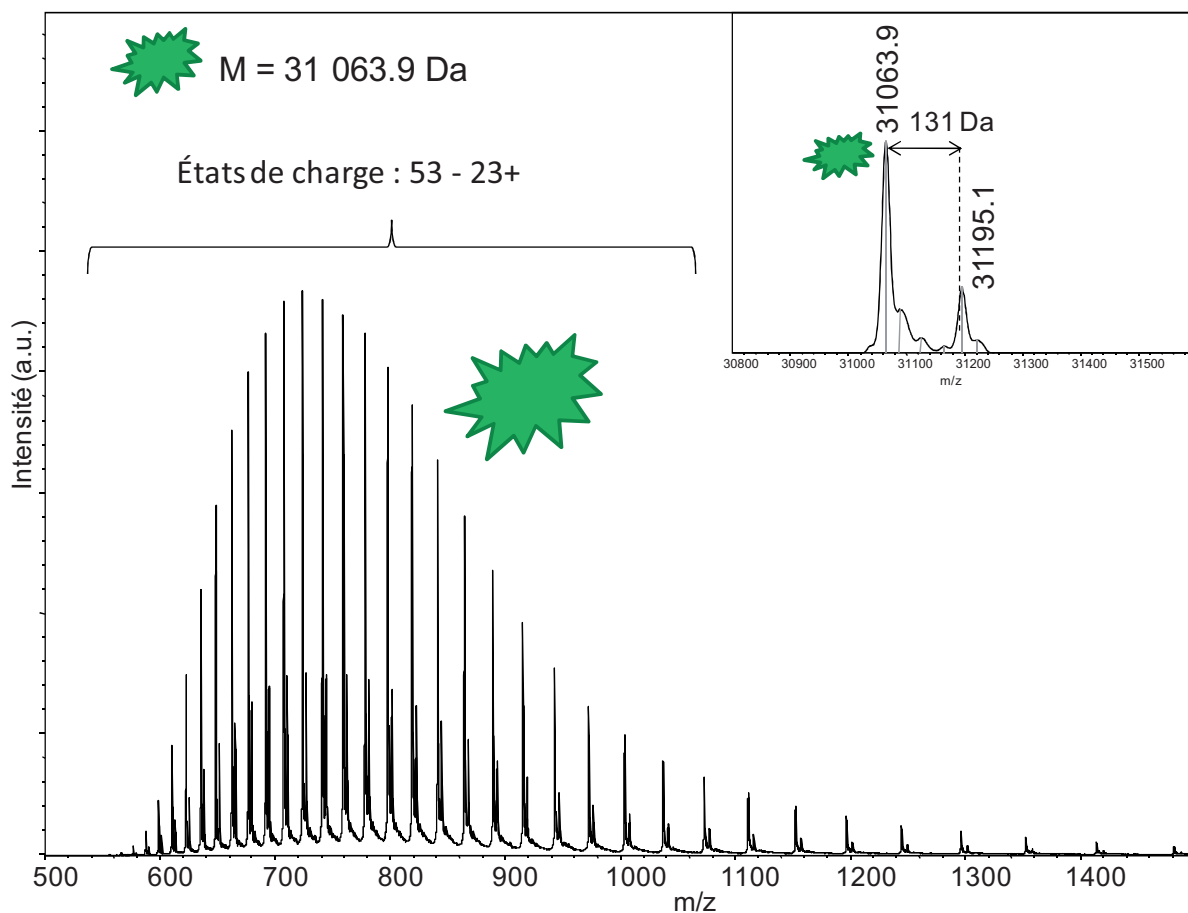


Figure 64: Spectres ESI-UHR-Q-TOF (maXis) de Fpg-Zn 0.5  $\mu$ M en conditions dénaturantes dans Wafa avec un débit de gaz de 10 L/min.

En conditions dénaturantes, la protéine n'est plus complexée avec le zinc (Figure 64). On observe toujours deux formes de Fpg, avec et sans méthionine initiale. La distribution des états de charge est très large, avec des états de charge beaucoup plus élevés qu'en conditions natives, ce qui confirme que la protéine est dénaturée.

L'existence d'un édifice non-covalent NEIL-2/Zn a aussi été vérifiée par MS.

#### 2.4.3. Analyse de Neil2-Zn en conditions natives

L'orthologue humain de Fpg, NEIL2 (M = 37 518 Da), a été étudié dans les mêmes conditions que le complexe Fpg-Zn.

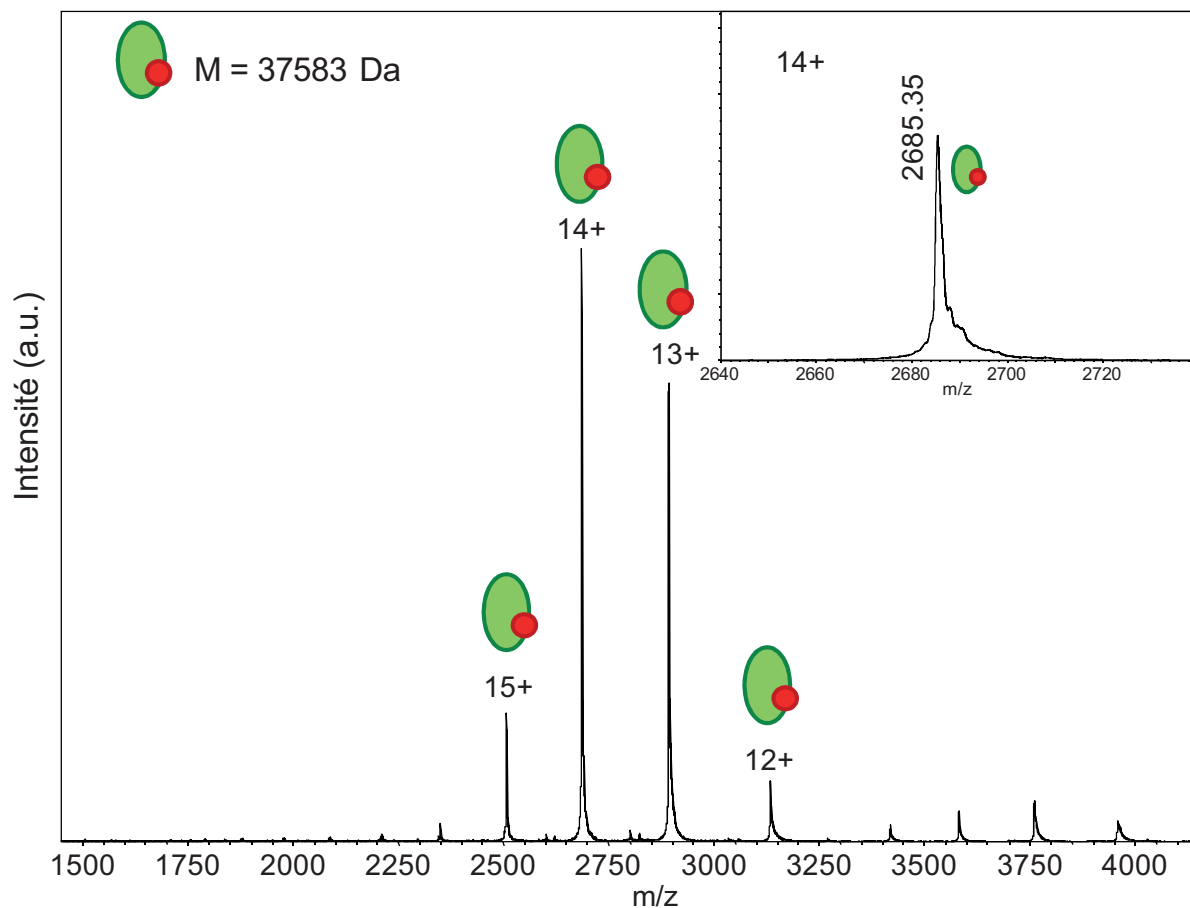


Figure 65: Spectre ESI-UHR-Q-TOF (maXis) de Neil2-Zn 10  $\mu$ M en conditions natives dans  $\text{NH}_4\text{OAc}$  50 mM pH 6.5 avec un débit de gaz de 10 L/min

En conditions natives, la masse observée ( $M = 37583$  Da) correspond au complexe NEIL2-Zn, il est visible avec des états de charge de 12+ à 15+ (Figure 65). Le complexe apparait sous une forme homogène, ce qui signifie que la méthionine initiale de la protéine est clivée à 100%.

#### 2.4.4. Analyse de Neil2-Zn en conditions dénaturantes

En conditions dénaturantes, la protéine est dénaturée et n'est plus liée au zinc car la masse observée est de 37 518.2 Da (Figure 66).

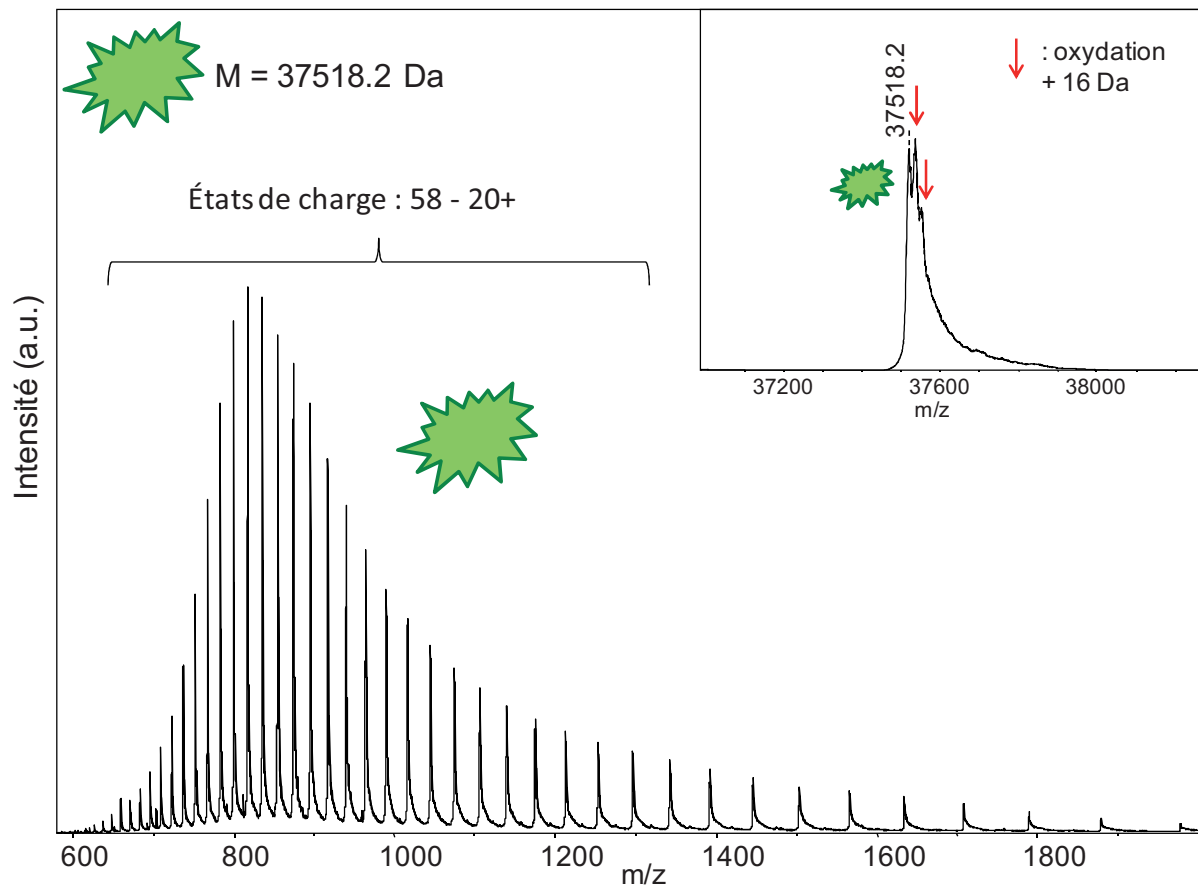


Figure 66: Spectre ESI-UHR-Q-TOF (maXis) de Neil2-Zn 0.5  $\mu$ M en conditions dénaturantes dans Wafa avec un débit de gaz de 10 L/min

Ces premiers résultats montrent qu'en optimisant les paramètres instrumentaux, les interactions Fpg-Zn et NEIL-2-Zn peuvent être complètement préservées. La nature ionique de ces interactions participe vraisemblablement fortement à sa préservation en phase gazeuse.

## 2.5. Interaction avec la drogue 2-thioxanthine

### 2.5.1. Interaction de la 2-thioxanthine avec Fpg-Zn en MS native

L'effet de l'inhibiteur 2-thioxanthine sur le complexe Fpg-Zn a ensuite été étudié. L'inhibiteur a été ajouté en solution au complexe Fpg-Zn. Malheureusement, aucun signal n'a pu être détecté pour le complexe. Seul l'inhibiteur libre apparaît dans le spectre. Il semble que l'ajout de la drogue diminue la solubilité de la protéine et entraîne son agrégation. Ce phénomène avait déjà été observé par l'équipe de B. Castaing. D'après leur expérience, le problème d'agrégation est moins fort lorsque la protéine Fpg-Zn est complexée avec de l'ADN.

L'étude a donc été répétée avec la protéine complexée à l'ADN : Fpg-Zn-ADN.

Les paramètres instrumentaux ont dû cependant être optimisés pour observer ce nouveau type de complexe hybride protéine-ADN.

## 2.5.2. Interaction de la 2-thioxanthine avec Fpg-Zn-ADN en MS native

### 2.5.2.1. Analyse du complexe Fpg-Zn-ADN en conditions natives

Un complexe de type protéine-ADN peut être difficile à observer par MS car la protéine et l'ADN sont des entités aux propriétés physicochimiques différentes voire opposées, ce qui a pour conséquence que l'on ne peut pas produire leur ionisation dans les mêmes conditions. Dans notre cas, le double brin d'ADN endommagé est de 8501Da, ce qui constitue environ un quart de la masse de la protéine (Figure 67).

CTC TTT **X**TT TCT CG :  $M_{\text{moy}} = 4162$   
GCG AGA AAC AAA GA :  $M_{\text{moy}} = 4339$

Figure 67: Séquence et masse du double brin d'ADN portant une base oxydée X, du 3-méthyl-3-hydroxy-hydantoin.

La contribution de la masse de l'ADN restant minoritaire dans la masse du complexe protéine-ADN, le complexe a d'abord été analysé dans les mêmes conditions que la protéine sans ADN. Avec un débit de gaz séchant de 4 L/min et en réduisant la valeur d'isCID à 100eV, 100% de complexe Fpg/Zn/ADN ont été observé (Figure 68)

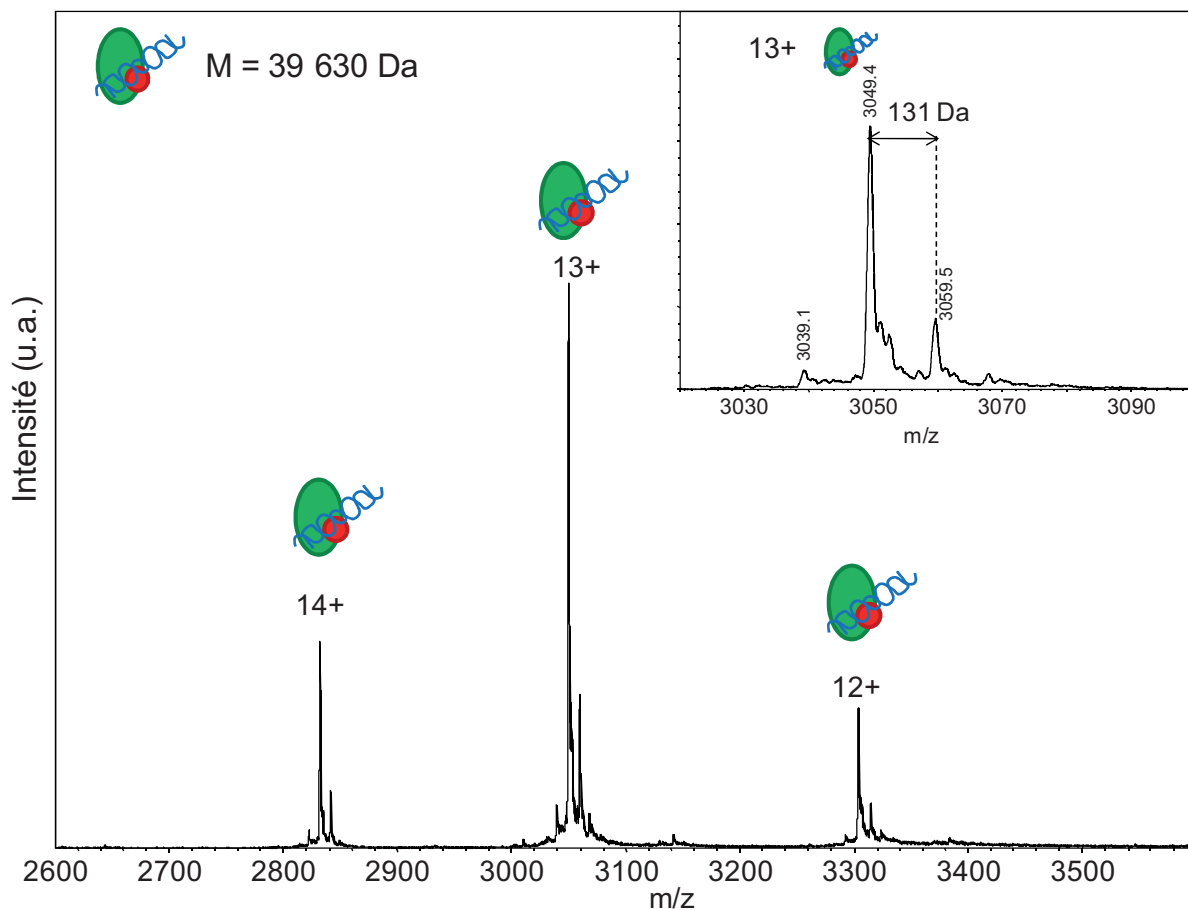


Figure 68: Spectre ESI-UHR-Q-TOF (maXis) de Fpg-Zn-ADN 10  $\mu$ M en conditions natives dans  $\text{NH}_4\text{OAc}$  50 mM pH 6.5 avec un débit de gaz de 4 L/min et isCID = 100eV.

La masse de l'édifice observé, à 39 630 Da, correspond exactement à la masse de la protéine complexée avec l'ADN et le zinc (8501 Da + 31129 Da = 39 630 Da).

#### 2.5.2.2. Analyse du complexe Fpg-Zn-ADN-2-thioxanthine en conditions natives

Comme il était possible d'observer le complexe Fpg-Zn-ADN en conditions natives, la drogue a ensuite été ajoutée. Aucun signal n'a pu être observé. Ceci suggère que, comme pour le complexe Fpg-Zn, l'ajout de la drogue entraîne aussi l'agrégation du complexe.

L'ajout de la drogue, avec la protéine seule, ou avec la protéine complexée avec l'ADN, semble entraîner une agrégation de l'échantillon et ne permet donc pas son étude par MS.



## 2.6. Bilan

L'étude de Fpg et NEIL-2 par MS native a confirmé leur complexation avec le zinc. Cette interaction subsiste lorsque que Fpg interagit avec de l'ADN. L'interaction de l'inhibiteur 2-thioxanthine avec la Fpg n'a pas pu être étudiée suite à l'agrégation induite par l'inhibiteur.

Toutes les optimisations instrumentales ont cependant été réalisées pour de futures études avec les systèmes Fpg-Zn et NEIL-2-Zn, qu'ils soient complexés ou non avec de l'ADN.

## II. Etude cinétique : Cas de la formation de l'hétérodimère $\alpha\beta$ de la protéine HU

### 1. Contexte

HU d'*Escherichia coli* est une protéine de type histone impliquée dans la compaction de l'ADN bactérien. Elle est capable d'induire des super-enroulements négatifs de l'ADN circulaire en présence de topoisomérase I. Les protéines HU se présentent sous la forme d'un homodimère dans la majeure partie des espèces bactériennes, à l'exception des entérobactéries. Chez *E. coli*, elle existe sous forme hétérodimérique  $\alpha\beta$  pour 90% et homodimérique  $\alpha_2$  ou  $\beta_2$  pour les 10% restant. Cependant, ces proportions varient au cours de la croissance cellulaire. HU se fixe sous forme de dimère et de manière non séquence-spécifique à l'ADN. Différentes études suggèrent fortement que chaque dimère HU d'*E. coli* joue un rôle biologique différent. L'équipe de Bertrand Castaing s'intéresse aux caractéristiques biophysiques et structurales des différents dimères de HU qui sous-tendent leurs disparités fonctionnelles. L'équipe a montré par microcalorimétrie que la dénaturation thermique des dimères était biphasique, le dimère natif passant par un état intermédiaire avant de former le monomère complètement dénaturé (189). Le dimère intermédiaire, partiellement dénaturé, n'est pas capable de se lier à l'ADN double brin avec une forte affinité (189). Cela suggère que l'équilibre état natif - état intermédiaire joue un rôle central dans la physiologie de la protéine HU d'*E. coli*. L'état intermédiaire est sans doute requis chez *E. coli* pour échanger les chaînes de HU, ce qui permet la formation de l'hétérodimère à partir des homodimères (189). Une étude de Vis *et al.* a montré qu'il était possible, en chauffant, de former un hétérodimère inter-spécifique à partir de deux homodimères HU de *Bacillus stearothermophilus* et *Bacillus subtilis*, cet hétérodimère n'était cependant pas stable dans le temps (135).

### 2. Objectif

L'objectif est d'utiliser la spectrométrie de masse non-covalente pour caractériser les différents dimères de HU ( $\alpha_2$ ,  $\beta_2$ ,  $\alpha\beta$ ) directement en solution, afin de pouvoir étudier la

formation de l'hétérodimère  $\alpha\beta$  à partir des deux homodimères, et ce, de façon qualitative et quantitative.

### 3. Résultats

#### 3.1. Optimisation des conditions expérimentales pour observer les dimères HU

L'optimisation des paramètres instrumentaux est décrite en exemple dans le chapitre I. Elle conduit à l'utilisation d'un voltage de la cap exit de 200V et d'une température de désolvation de 120°C.

Les paramètres expérimentaux de pH, de concentration du tampon et de température d'incubation ont également été optimisés pour favoriser l'observation de dimères.

##### 3.1.1. Effet du pH

D'après les études de spectrométrie de masse menées sur HU par Vis et ses collaborateurs (135, 175), il a été montré que la conformation native du dimère est préservée en présence de fortes concentrations en acétate d'ammonium (> 50 mM) et d'un pH légèrement acide de 5.

Deux pH d'une solution d'acétate d'ammonium à 50 mM ont été donc testés: pH 5 et pH 7 (Figure 69). L'hétérodimère préformé, purifié, et à forte concentration, est dilué dans l'un ou l'autre tampon pour atteindre le pH désiré, puis analysé en ESI-IT.

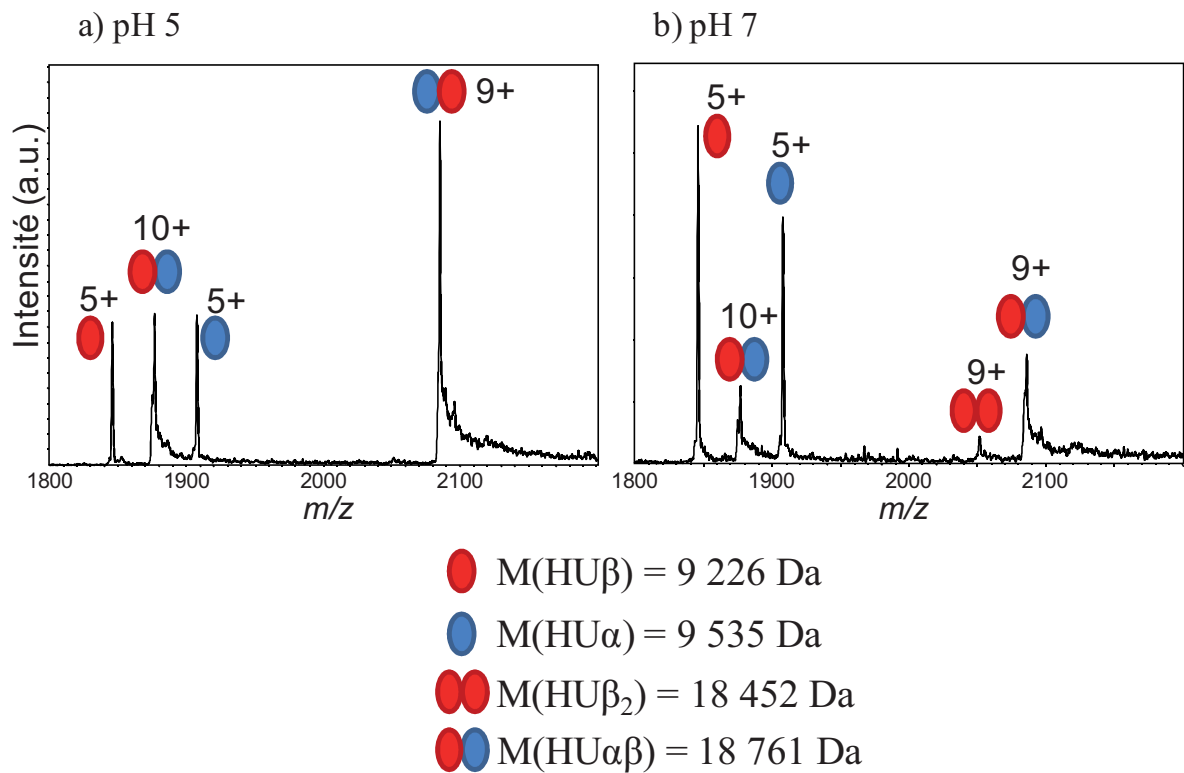


Figure 69: Spectres ESI-IT du dimère HU $\alpha\beta$  en conditions natives à une concentration de 10  $\mu$ M dans NH<sub>4</sub>OAc 50 mM, a) pH 5, b) pH 7, avec un couple V<sub>SK</sub>/V<sub>CE</sub> de 100/200V

Pour interpréter les spectres d'HU, on s'appuie sur le fait que:

- ✓ Les distributions d'états de charge des homodimères HU $\alpha_2$  et HU $\beta_2$ , et monomères HU $\alpha$  et HU $\beta$  étant superposés, seul les états de charge impairs du dimère correspondent exclusivement à du dimère, tandis que les états de charge pairs peuvent correspondre soit à du dimère, soit au monomère correspondant portant le double de charges.
- ✓ Pour l'hétérodimère HU $\alpha\beta$ , les distributions d'états de charge des dimères et monomères ne se superposent pas car les monomères HU $\alpha$  et HU $\beta$  ont des masses différentes: M<sub>moy</sub>(HU $\alpha$ ) = 9535.0 Da et M<sub>moy</sub>(HU $\beta$ ) = 9225.6 Da

L'analyse d'HU $\alpha\beta$  devrait montrer la présence de dimère, avec une dissociation en monomères plus ou moins forte en phase gazeuse.

A pH 5, l'hétérodimère HU $\alpha\beta$  apparaît clairement aux états de charge 9+ et 10+ (Figure 69a). Le spectre ne fait pas apparaître d'intensité significative d'homodimère HU $\beta_2$ . On peut donc en conclure qu'il n'y a pas d'homodimère à l'état de charge 10+ mais uniquement du

monomère à l'état de charge 5+. Le même raisonnement peut être appliqué à  $HU\alpha_2$ . On peut donc en conclure qu'il y a dissociation partielle de l'hétérodimère. Cette dissociation ne peut être quantifiée directement car on ne connaît pas les intensités relatives du dimère et du monomère à la même concentration.

A pH 7, il y a proportionnellement moins d'hétérodimère par rapport aux monomères qu'à pH 5 (Figure 69b). Un peu d'homodimère  $HU\beta_2$  est visible, ce qui suggère que le pic annoté comme  $HU\beta$  chargé 5+ peut également contenir une petite quantité de  $HU\beta_2$  chargé 10+. L'origine de l'espèce  $HU\beta_2$  dans cette expérience serait à clarifier. La dissociation accrue à pH 7 peut s'être produite soit en solution, soit dans l'instrument. Pour la suite des expériences, les dimères HU ont été incubés à pH 5 pour préserver au maximum le complexe.

### 3.1.2. Effet de la concentration en tampon

Une fois le pH fixé, l'effet de la force ionique a été étudié en faisant varier la concentration en acétate d'ammonium sur le système  $HU\alpha\beta$ .

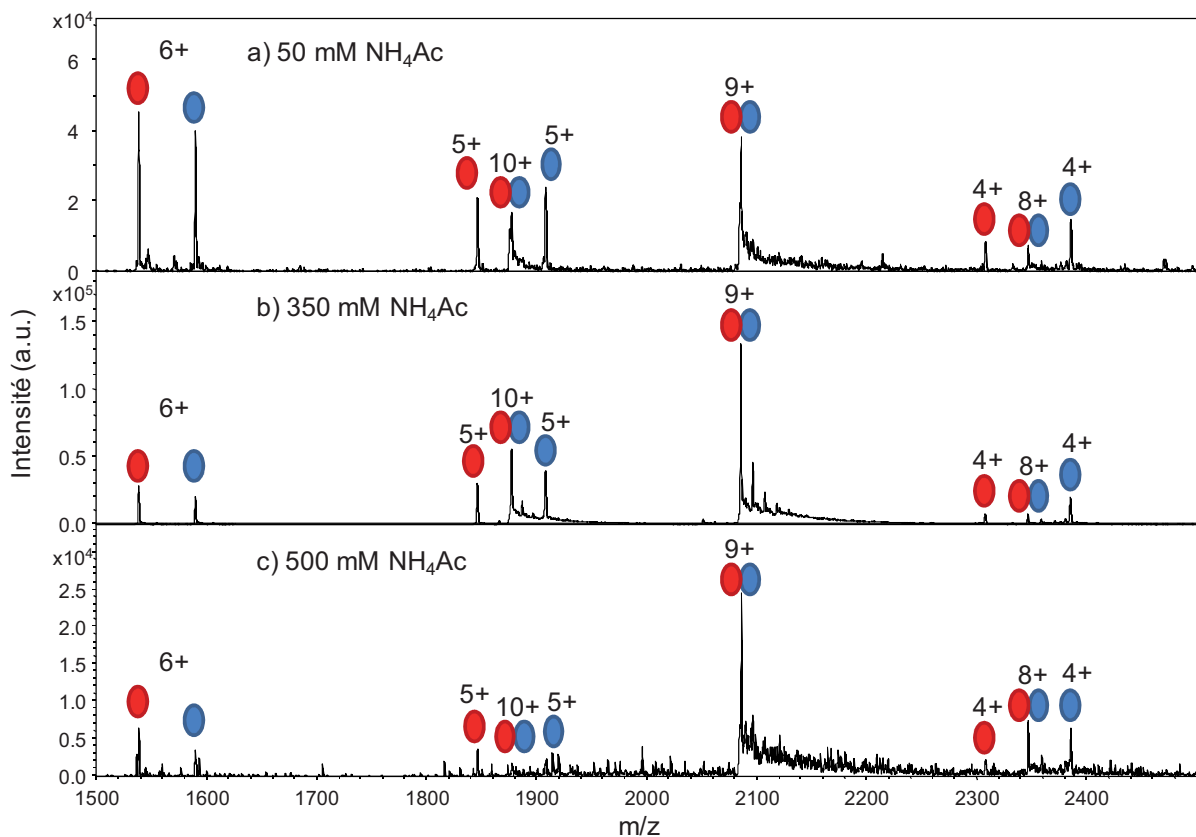


Figure 70: Spectres ESI-IT à pH 5 de 10  $\mu M$  de dimère  $HU\alpha\beta$  dans  $NH_4OAc$  a) 50 mM, b) 350 mM ou c) 500 mM.

Les spectres d'HU $\alpha\beta$  sont faciles à interpréter car les distributions des états de charge des monomères et des dimères ne se confondent pas. Avec 350 ou 500 mM d'acétate d'ammonium, on observe un pic majoritaire correspondant au dimère portant 9 charges. Le dimère est stabilisé par de fortes concentrations en acétate d'ammonium. A 500 mM d'acétate d'ammonium, le rapport signal/bruit se dégrade néanmoins.

Deux phénomènes peuvent être à l'origine de la dissociation du dimère. Soit la stabilité du complexe en solution nécessite une concentration ionique minimale, ce qui fait qu'à faible force ionique le complexe se dissocie en solution. Soit la préservation du complexe en phase gazeuse nécessite une forte concentration en acétate d'ammonium. Les contributions relatives de ces éventuels effets sont inconnues. Par la suite, les analyses de dimère HU ont été effectuées dans 350 mM d'ABC, pour mimer une forte concentration en sels et limiter les pertes de sensibilité.

### 3.1.3. Effet de la température d'incubation

D'après la littérature, l'hétérodimère HU $\alpha\beta$  se dénature à partir de 37°C et est complètement dénaturé à 80°C. Pour voir si la dénaturation du dimère était visible par MS et si les intensités relatives des dimères et monomères pouvaient ainsi être déterminées, l'effet de la température d'incubation sur la stabilité du complexe a été examiné.

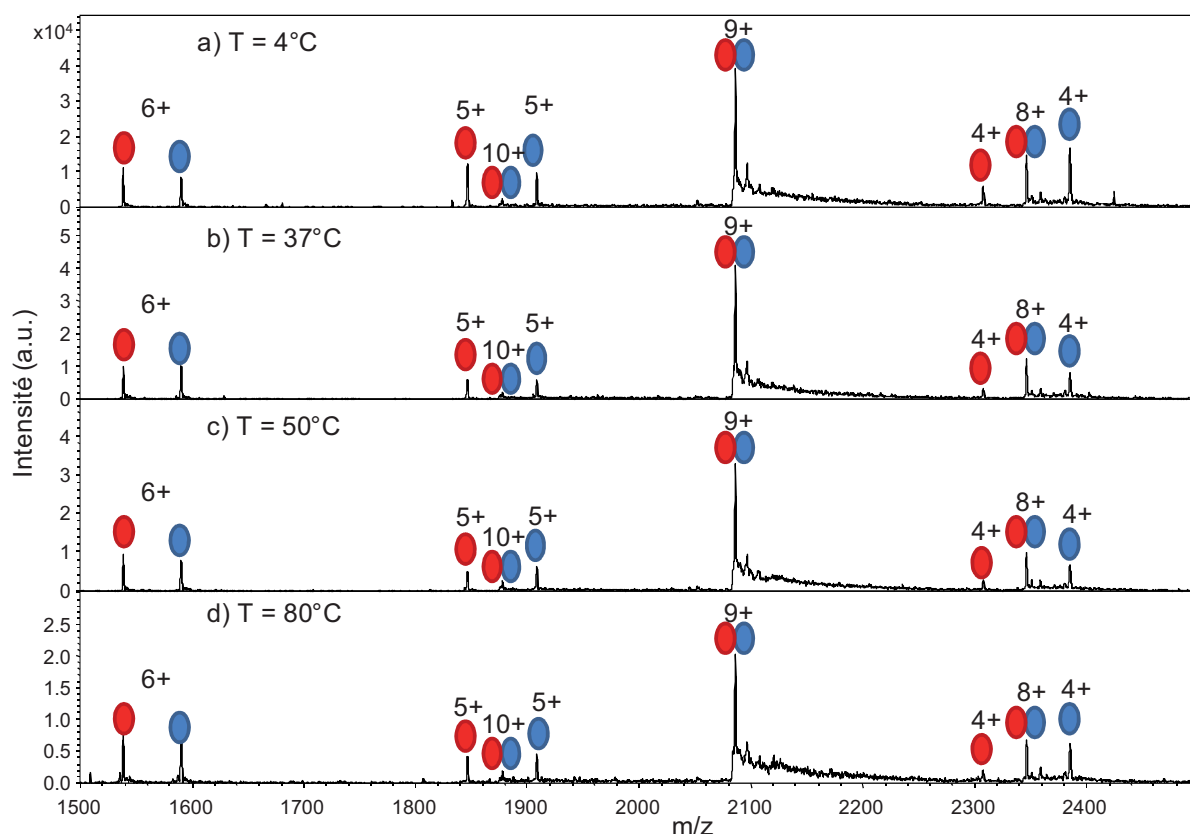


Figure 71: Spectres ESI-IT de 10  $\mu\text{M}$  HU $\alpha\beta$  en conditions natives dans  $\text{NH}_4\text{OAc}$  350 mM pH 5 après incubation à (a) 4°C, (b) 37°C, (c) 50°C, (d) 80°C.

L'hétérodimère HU $\alpha\beta$  a été incubé 30 minutes à quatre températures (4, 37, 50 et 80°C) puis analysé par MS native. Jusqu'à 50°C, la température d'incubation ne semble pas avoir d'effet sur la préservation du dimère. Une légère dissociation commence à être observée à partir de 80°C (Figure 71). Ce résultat n'est pas en accord avec les expériences de dénaturation thermique par microcalorimétrie (189). Il est possible que la dénaturation soit réversible dans ces conditions d'analyse, la ligne d'infusion n'étant pas thermostatée.

Cette expérience de dénaturation n'ayant pas abouti, les intensités relatives d'ionisation des dimères et monomères n'ont pas pu être déterminées.

### 3.2. Formation d'hétérodimères à partir des deux homodimères

Lors de l'étude de la formation d'hétérodimère à partir des deux homodimères, seuls les états de charge 9+ des différents dimères sont pris en compte pour la quantification relative.

### 3.2.1. Etude cinétique à différentes températures

Pour étudier la formation d'hétérodimère à partir d'homodimères, 5  $\mu\text{M}$  d' $\text{HU}\alpha_2$  ont été incubés avec 5  $\mu\text{M}$  d' $\text{HU}\beta_2$  dans un volume final de 100  $\mu\text{L}$  à la température voulue. 10  $\mu\text{L}$  d'échantillon ont ensuite été prélevés régulièrement pour suivre la cinétique de formation de l'hétérodimère.

Le rapport  $R = \text{intHU}\alpha\beta^+ / (\text{intHU}\alpha_2^+ + \text{intHU}\beta_2^+ + \text{intHU}\alpha\beta^+)$  est calculé pour suivre quantitativement la formation d'hétérodimère. Les facteurs d'ionisation des trois dimères et le pourcentage de préservation de chaque dimère n'étant pas connus, le paramètre R ne reflète pas directement le pourcentage d'hétérodimère en solution. Il donne cependant une idée de l'évolution de l'hétérodimère par rapport aux homodimères. Etant un rapport d'intensités, il élimine au moins les variabilités d'ionisation inter-analyse.

La cinétique de formation de l'hétérodimère a été étudiée à quatre températures : 2, 10, 27 et 37 °C.

Une première incubation a été réalisée à 37 °C (Figure 72).



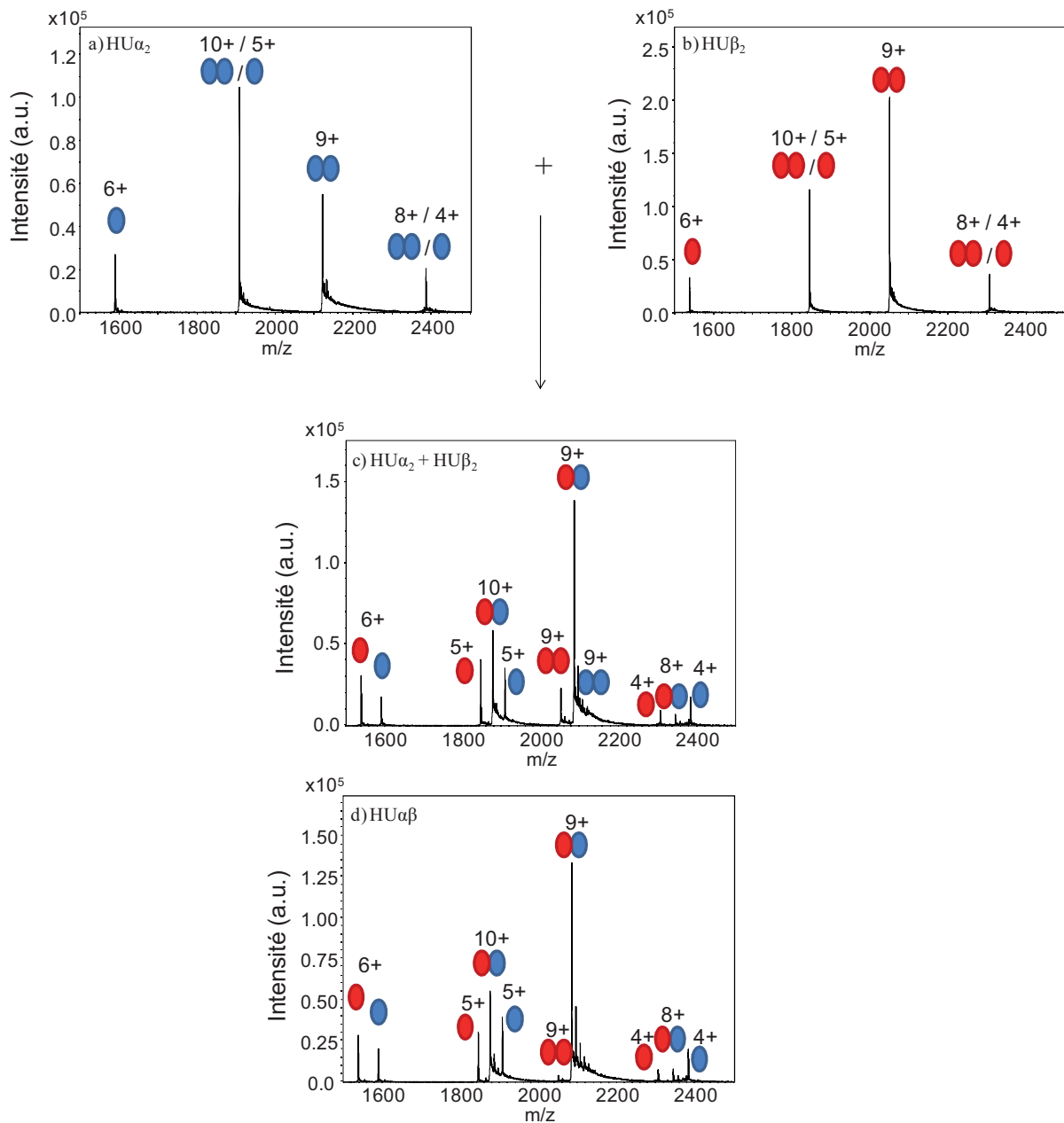


Figure 72: Spectres ESI-IT de (a) HU $\alpha_2$  5  $\mu$ M et (b) HU $\beta_2$  5  $\mu$ M **avant incubation**, puis (c) de HU $\alpha_2$  5  $\mu$ M et HU $\beta_2$  5  $\mu$ M **incubés** 6 minutes à 37 °C, (d) spectre **contrôle** de HU $\alpha\beta$  10  $\mu$ M dans NH<sub>4</sub>OAc 350 mM, pH 5.

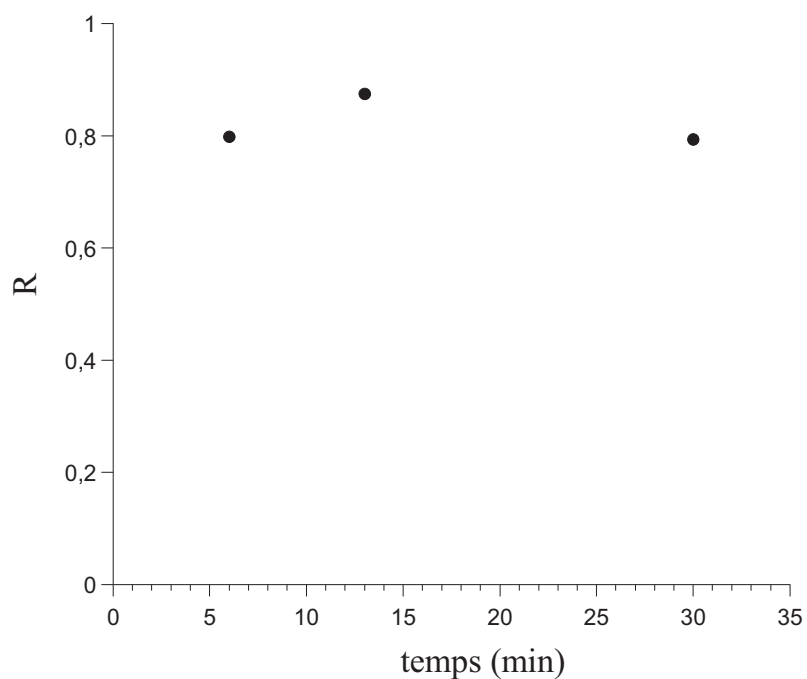


Figure 73: Cinétique de formation d'hétérodimère HU $\alpha\beta$  à partir de 5  $\mu$ M d'HU $\alpha_2$  et de 5  $\mu$ M d'HU $\beta_2$  incubés à 37°C dans NH<sub>4</sub>OAc 350 mM, à pH 5.  $R = \text{intHU}\alpha\beta 9^+ / (\text{intHU}\alpha_2 9^+ + \text{intHU}\beta_2 9^+ + \text{intHU}\alpha\beta 9^+)$

A 37°C, l'équilibre est atteint en moins de 6 minutes. Le paramètre R est supérieur à 80 %. Cette valeur de 80 % traduit sans doute une réaction totale. En effet, d'après les intensités d'ionisation de chaque dimère (Figure 72), l'intensité du dimère HU $\beta_2$  portant 9 charges est quatre fois plus élevée que les intensités correspondantes d'HU $\alpha_2$  ou d'HU $\alpha\beta$ . Le paramètre R ne peut donc atteindre une valeur de 100 % si l'intensité du dimère HU $\beta_2$  9+ est prise en compte dans le calcul. La valeur maximale que peut prendre R est donc *a priori* de 80%.

Deux autres incubations ont ensuite été réalisées à 2 et 10 °C (Figure 74).

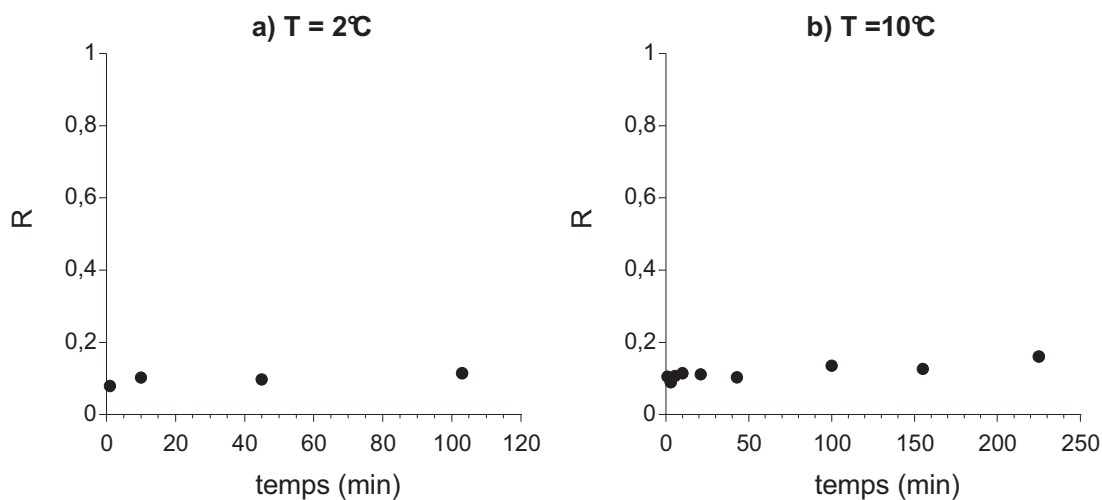


Figure 74: Cinétique de formation d'hétérodimère HU $\alpha\beta$  à partir de 5  $\mu\text{M}$  d'HU $\alpha_2$  et de 5  $\mu\text{M}$  d'HU $\beta_2$  incubés à a) 2°C ou b) 10°C dans NH<sub>4</sub>OAc 350 mM, à pH 5.  $R = \frac{\text{intHU}\alpha\beta 9^+}{(\text{intHU}\alpha_2 9^+ + \text{intHU}\beta_2 9^+ + \text{intHU}\alpha\beta 9^+)}$

A 2°C et 10 °C, le paramètre R reste inférieur à 15 % même après 4 heures d'incubation. La réaction est donc très lente lorsque la température est basse.

A 27°C, la formation d'hétérodimère a été étudiée sur 205 minutes. Le paramètre R semble approcher un équilibre au bout de ces 205 minutes avec une valeur d'équilibre d'environ 45% (Figure 75).

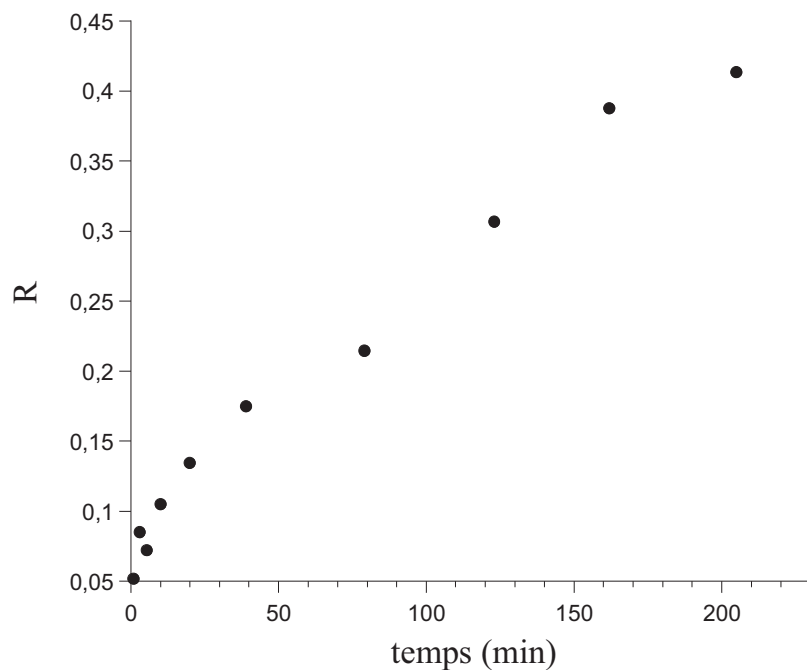


Figure 75: Cinétique de formation d'hétérodimère HU $\alpha\beta$  à partir de 5  $\mu\text{M}$  d'HU $\alpha_2$  et de 5  $\mu\text{M}$  d'HU $\beta_2$  incubés à 27°C dans NH<sub>4</sub>OAc 350 mM, à pH 5.  $R = \text{intHU}\alpha\beta^{9+} / (\text{intHU}\alpha_2^{9+} + \text{intHU}\beta_2^{9+} + \text{intHU}\alpha\beta^{9+})$

Seule la cinétique réalisée à 27°C peut être exploitée pour tenter d'expliquer le mécanisme de formation d'hétérodimère. La courbe paraît biphasique, avec une première phase entre 0 et 80 minutes et une autre entre 80 et 200 min.

D'après l'étude de Castaing *et al.* (189), la formation d'hétérodimère nécessite que les deux homodimères soient dans une conformation intermédiaire. Il existe en permanence un équilibre entre la forme native et la forme intermédiaire des dimères. Il se pourrait que cet équilibre soit une étape limitante de la réaction, avec formation rapide d'hétérodimère à partir des homodimères en conformation intermédiaire. La bimodalité observée va dans ce sens, car elle suggère qu'il ne s'agit pas d'une simple réaction  $\text{HU}\alpha_2 + \text{HU}\beta_2 \rightarrow 2\text{HU}\alpha\beta$ , mais qu'il existe au moins une étape supplémentaire à vitesse limitante. Cette bimodalité pourrait être la preuve en MS qu'il y a nécessairement besoin d'un état intermédiaire pour former l'hétérodimère. L'exploitation de ces résultats va nécessiter de modéliser le comportement de HU sur le plan mathématique afin de produire une équation adéquate pour extraire des données les paramètres cinétiques de la réaction.

### 3.2.2. Stabilité au cours du temps

La formation d'hétérodimères inter-espèce donne lieu à des dimères qui ne sont pas stables dans le temps (135). Il est donc intéressant de savoir si, en revanche, l'hétérodimère formé à partir de monomères d'une même espèce est stable.

Pour étudier la stabilité de l'hétérodimère au cours du temps, celui-ci a été formé à 37°C avec un temps d'incubation de 10 minutes puis sa stabilité a été suivie à 20 °C, selon un protocole identique à celui de la publication de Vis (135) (Figure 76).

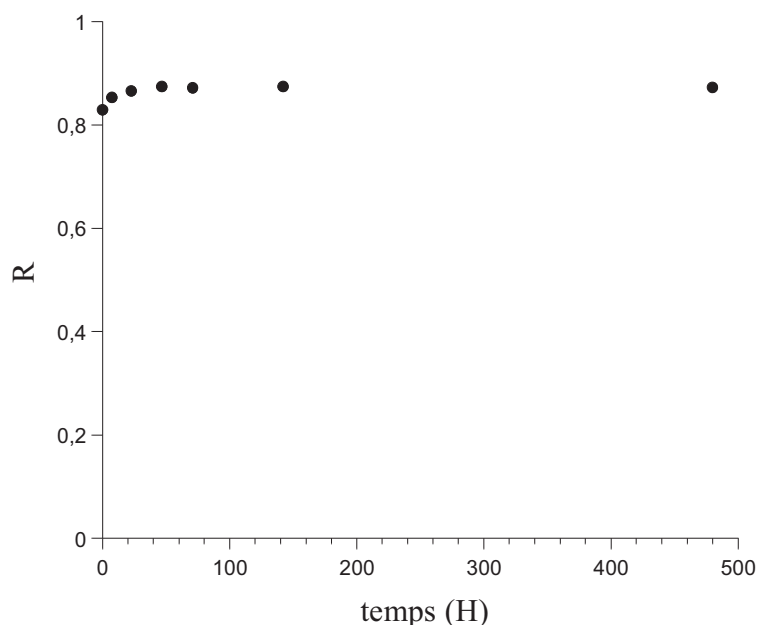


Figure 76: Stabilité à 20°C de l'hétérodimère HU $\alpha\beta$  formé dans une incubation de 10 minutes à 37°C par mélange des deux homodimères HU $\alpha_2$  et HU $\beta_2$ .  $R = \frac{\text{int HU}\alpha\beta 9^+}{(\text{int HU}\alpha_2 9^+ + \text{int HU}\beta_2 9^+ + \text{int HU}\alpha\beta 9^+)}$

Après 20 jours d'incubation à 20°C, l'hétérodimère HU $\alpha\beta$  est resté stable (Figure 76), contrairement à l'hétérodimère inter-espèce étudié par Vis.

Chez *E. coli*, le phénomène de formation de l'hétérodimère à partir des deux homodimères se passe à 37°C, en conditions physiologiques, alors que dans le cas de l'expérience inter-espèces, ce phénomène nécessite de faire chauffer les homodimères à 80°C pour les dénaturer et la formation du complexe est réversible. Cette observation est une preuve que la conformation intermédiaire permet la formation d'hétérodimère stable.

### 3.3. Bilan

Les 3 dimères HU ont été partiellement conservés en MS native. Les interactions qui maintiennent les dimères étant essentiellement hydrophobes, ce résultat est très encourageant.

L'étude de la formation d'hétérodimère HU $\alpha\beta$  à partir des deux homodimères HU $\alpha_2$  et HU $\beta_2$  en MS confirme que le phénomène se passe instantanément à 37°C chez *E. coli*. L'étude de la cinétique à 27°C a mis en évidence un comportement particulier, la formation de l'hétérodimère ayant un comportement biphasique.

Pour la suite, il serait très intéressant d'approfondir ce comportement cinétique. La technique de MS native est maintenant optimisée pour caractériser les différents dimères HU. Elle pourrait servir au criblage de mutants de HU pour localiser les résidus impliqués dans la formation d'hétérodimère.

# Chapitre III

Développement d'une nouvelle méthodologie pour déterminer des constantes d'équilibre de dissociation des complexes à faible affinité

### *Résumé du chapitre III:*

*La détermination par ESI-MS de constantes d'équilibre de dissociation ( $K_D$ ) qui reflètent fidèlement l'affinité d'un complexe protéine-ligand en solution continue de susciter un vif intérêt. Les difficultés à mesurer un  $K_D$  sont accrues dans le cas des complexes de faible affinité. Nous présentons une méthode de mesure de  $K_D$  accompagnée d'un modèle mathématique qui permet de traiter à la fois la dissociation en phase gazeuse (GPD) et l'agrégation. Pour cela, une correction mathématique rationnelle de la GPD ( $f_{sat}$ ) est combinée au développement d'un nouveau protocole expérimental qui traite, lui, l'agrégation en phase gazeuse. Un guide est fourni pour appliquer la méthode à des systèmes non-covalents protéine-ligand selon leur comportement cinétique.*

*L'approche est validée en comparant des valeurs de  $K_D$  déterminées par notre méthode avec des valeurs de  $K_D$  de la littérature déterminées en solution. L'influence du type d'interaction moléculaire et de paramètres instrumentaux est examinée dans le but d'une dissection fine des facteurs affectant la GPD. La méthode peut être appliquée de manière fiable à un large éventail de systèmes sans nécessiter l'utilisation d'un ligand ou d'une protéine de référence.*

*Ce développement méthodologique a fait l'objet d'une publication dans Journal of the American Society for Mass Spectrometry.*



## Contexte

Dans l'optique de cribler par MS des analogues de nucléotides pour leur capacité à interagir avec la PEBP, notre première approche a été de comparer des fractions de liaison. Les premières analyses ont été effectuées avec les deux nucléotides qui possédaient les meilleures capacités à éluer la PEBP d'une colonne d'affinité Blue-Sepharose : FMN et GTP (73). La protéine a été incubée avec du ligand en excès (500  $\mu$ M pour 18  $\mu$ M) et la fraction de liaison analysée par MS. Des stœchiométries allant jusqu'à 1:6 ont été observées sur l'état de charge 10+. Ces stœchiométries apparentes élevées correspondent à des interactions non-spécifiques entre la protéine et la petite molécule. Ces interactions non-spécifiques sont souvent décrites dans la littérature (149, 159–161, 163–165, 190). Elles résultent d'un phénomène d'agrégation en phase gazeuse qui a lieu lors de l'analyse MS. Pour différencier les interactions non-spécifiques qui peuvent avoir lieu en solution de celles résultant d'un transfert en phase gazeuse, le terme d'agrégation en phase gazeuse est préféré.

Une autre observation a pu être faite lors de ces premières analyses. En ayant 28 fois plus de ligand que de protéine, un pic correspondant à la protéine libre est toujours visible sur le spectre. Cela semble indiquer que l'affinité est faible, avec un  $K_D$  supérieur au micromolaire.

Le critère le plus approprié pour cribler des molécules selon leur affinité pour la PEBP est la constante d'équilibre de dissociation ( $K_D$ ). Le  $K_D$  est inversement proportionnel à l'affinité.

Pour déterminer des  $K_D$  sur des systèmes à faible affinité sensibles à l'agrégation en phase gazeuse, sans ligand de référence, les auteurs Daubenfeld et van der Rest ont décrit une méthode basée sur une simulation statistique de distributions de complexes (169). Elle consiste à distinguer les interactions spécifiques de l'agrégation par une méthode de déconvolution des données de masse (169). Lorsque nous avons appliqué cette méthode au complexe PEBP-GTP, un  $K_D$  de 2.5 mM a été trouvé.

Cependant, la représentation des données de fraction de liaison mesurée en fonction de la concentration en ligand a mis en évidence un phénomène de dissociation en phase gazeuse (GPD). Ceci contribue à expliquer la présence de protéine libre même en présence d'un large excès de ligand.

Une méthode de MS permettant de déterminer des  $K_D$  pour des systèmes de faible affinité, sensibles à l'agrégation et à la dissociation en phase gazeuse, était donc nécessaire pour réaliser le criblage des analogues de nucléotides. Des méthodes existent dans la littérature pour régler l'un ou l'autre de ces problèmes, mais pas les deux simultanément. De plus, elles utilisent généralement un ligand de référence, ayant un  $K_D$  connu et procèdent ensuite par des expériences de compétition. Le système PEBP-nucléotides étant encore peu étudié, il n'existait pas au début de ce travail pas de  $K_D$  pour les ligands de référence.

Nous avons ainsi entrepris de développer une méthode MS, qui ne nécessiterait pas de ligand de référence pour déterminer de façon fiable des constantes de dissociation pour des systèmes de faible affinité enclins simultanément à l'agrégation et à la dissociation en phase gazeuse.

## Publication:

Improved accuracy of low affinity  
protein-ligand equilibrium dissociation constants  
directly determined by  
electrospray ionization mass spectrometry





## RESEARCH ARTICLE

# Improved Accuracy of Low Affinity Protein–Ligand Equilibrium Dissociation Constants Directly Determined by Electrospray Ionization Mass Spectrometry

Lucie Jaquillard,<sup>1</sup> Fabienne Saab,<sup>1</sup> Françoise Schoentgen,<sup>2</sup> Martine Cadene<sup>1</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS UPR 4301, affiliated with Université d'Orléans, rue Charles-Sadron, 45071 Orléans, Cedex 2, France

<sup>2</sup>Institut de Minéralogie et de Physique des Milieux Condensés, CNRS UMR 7590, Université de Paris 6, 75252 Paris, Cedex 5, France

### Abstract

There is continued interest in the determination by ESI-MS of equilibrium dissociation constants ( $K_D$ ) that accurately reflect the affinity of a protein–ligand complex in solution. Issues in the measurement of  $K_D$  are compounded in the case of low affinity complexes. Here we present a  $K_D$  measurement method and corresponding mathematical model dealing with both gas-phase dissociation (GPD) and aggregation. To this end, a rational mathematical correction of GPD ( $f_{\text{sat}}$ ) is combined with the development of an experimental protocol to deal with gas-phase aggregation. A guide to apply the method to noncovalent protein–ligand systems according to their kinetic behavior is provided. The approach is validated by comparing the  $K_D$  values determined by this method with in-solution  $K_D$  literature values. The influence of the type of molecular interactions and instrumental setup on  $f_{\text{sat}}$  is examined as a first step towards a fine dissection of factors affecting GPD. The method can be reliably applied to a wide array of low affinity systems without the need for a reference ligand or protein.

**Key words:** Accuracy, Equilibrium dissociation constant,  $K_D$ , Gas phase dissociation, Aggregation, ESI, Model, Low affinity, Interaction, Protein-ligand, Mass spectrometry

## Introduction

Most of the cell functions that endow organisms with sustainable life rely on the interaction of proteins with each other, nucleic acids, glycans, lipids, or small molecules. The formation of a protein–ligand complex is generally mediated by noncovalent contacts, including hydrogen

bonds, ionic bonds, van der Waals interactions, and interactions based on the hydrophobic effect. The protein–ligand equilibrium dissociation constant  $K_D$  is an essential parameter used to describe a binding system, and to provide relative affinity information between ligands or between target proteins. The  $K_D$  can generally be determined by a biophysical method such as isothermal titration calorimetry, spectrophoto- or spectrofluorimetry, or surface plasmon resonance. In these techniques,  $K_D$  is determined by measurement of an incident property of the complex resulting from binding. Measuring the  $m/z$  of complex ions by MS is the only method allowing for the direct measurement of a fundamental physical dimension of the

**Electronic supplementary material** The online version of this article (doi:10.1007/s13361-011-0305-7) contains supplementary material, which is available to authorized users.

Correspondence to: Martine Cadene; e-mail: cadene@cnr-orleans.fr

Received: 5 August 2011  
Revised: 7 November 2011  
Accepted: 23 November 2011  
Published online: 20 March 2012

complex, in this case its mass. Following seminal work by Ganem and coworkers [1, 2] and Katta and Chait [3] for noncovalent complex detection, direct infusion electrospray ionization mass spectrometry (ESI-MS) has emerged as a powerful technique for quantifying the equilibrium dissociation constant  $K_D$  for protein–ligand interactions [4–6]. The search for accurate methods for  $K_D$  determination has also led to the development of indirect methods based on HDX and mass spectrometry (MS), such as PLIMSTEX [7] and SUPREX [8].

The direct ESI-MS  $K_D$  determination technique is based on the direct detection and quantification of PL (protein–ligand complex) and P (unbound protein) ions in the gas phase. A key assumption for this type of experiment is that the bound fraction ( $f_{\text{bound}}$ ) measured by MS is equivalent to the  $f_{\text{bound}}$  at equilibrium in solution. The classical quadratic Equation [9] establishes the relationship between  $f_{\text{bound}}$ , the ratio of bound protein concentration over total protein concentration, which is the measured variable,  $K_D$ , the protein–ligand equilibrium dissociation constant to be determined, and  $[P]_0$  and  $[L]_0$ , the initial protein and ligand concentrations, respectively:

$$f_{\text{bound}} = \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4[P]_0[L]_0}}{2[P]_0} \quad (1)$$

The equilibrium dissociation constant  $K_D$  can be determined by nonlinear regression analysis on a  $[L]_0$ ,  $f_{\text{bound}}$  dataset using Equation (1), where  $[L]_0$  is the x input variable and  $f_{\text{bound}}$  is the y variable. However, the  $f_{\text{bound}}$  measured in the gas phase reflects not only the solution-binding fraction, but also phenomena related to the transfer of the complex into the gas phase. These phenomena generate issues in the measurement of  $K_D$ , which are compounded in the case of low affinity complexes. First, the protein–ligand complex can be susceptible to partial dissociation in the interface of the instrument. This phenomenon is widely known as gas-phase dissociation (GPD). Avoiding GPD can be quite challenging [10–12]. As a consequence of GPD, the bound fraction tends to be underestimated. Second, the presence of free protein and ligand molecules in the electrospray droplet results in gas-phase aggregation, i.e., the formation of nonspecific protein–ligand complexes upon evaporation of the droplet. This is of concern particularly for low affinity systems which, in contrast to high-affinity systems, necessitate the addition of a significant excess of ligand to produce a measurable bound fraction. Protein–ligand complexes can be considered as low affinity if  $K_D \geq 1 \mu\text{M}$ . As a consequence of gas-phase aggregation, the observed complex stoichiometry is altered compared with the solution-phase stoichiometry [2]. Moreover, aggregation can lead to false positives and has been shown to result in artificially low  $K_D$  values [13]. Different strategies have been developed to determine the  $K_D$  of low

affinity complexes by ESI-MS. One strategy consists in disrupting aggregation using black-body infrared radiation dissociation (BIRD) [14]. However, in cases where gas-phase aggregation is stronger than the specific interaction, this strategy leads to the destruction of the complex of interest [15]. The Daubenfeld/van der Rest method [16] is based on the hypothesis that specific ligand binding follows a statistical distribution, which is different from the distribution of nonspecific, aggregation-related binding modes. By deconvoluting the raw binding fraction data into separate distributions, the authors can calculate the contribution of specific binding to the interaction. Strategies based on the use of a reference protein, which is known to lack specific binding for the ligand of interest, have been used to quantify aggregation [17]. Alternatively, competitive assays with a reference ligand have been developed as a quantitative method to deal with aggregation [18]. The main limit of all these strategies is that they correct for aggregation without taking into account potential GPD phenomena.

To deal with gas phase dissociation, thermal heating, cone voltage and pressure in the first pumping stage are usually optimized to preserve the noncovalent complex [19–21]. Certain solvent additives promote evaporative cooling, thereby enhancing the survival probability of some weakly bound species [12, 22]. Competition experiments with a reference ligand can be used to compensate for variable response factors [23]. None of the above strategies to correct potential GPD deals explicitly with the aggregation issue.

$K_D$  determination methods simultaneously dealing with both the GPD and aggregation issues have not been previously described. Here we present a complete MS-based approach with the aim to determine a  $K_D$ , which accurately reflects the affinity in solution. To this end, a mathematical correction of GPD based on a rational examination of solution-to-gas transfer events is combined with the development of a new experimental protocol to deal with gas-phase aggregation. Moreover, we examine in detail the consequences of the introduction of a micropurification step on the behavior of the complex prior to analysis, and develop appropriate equations for the  $K_D$  determination. A guide to apply the method to noncovalent protein–ligand systems according to their kinetic behavior is provided. The approach is validated by comparing the  $K_D$  values determined by this method with literature values determined in solution on reference protein–ligand systems. Furthermore, the influence of the type of molecular interactions and of the instrumental setup on GPD is examined as a first step towards a fine dissection of factors affecting GPD.

## Experimental

### Materials

Bovine pancreatic ribonuclease A (RNase A), hen egg-white lysozyme (HEWL), guanosine triphosphate (GTP), flavin mononucleotide (FMN), adenosine diphosphate (ADP),

cytosine triphosphate (CTP), tri-*N*-acetylchitotriose (NAG3), ammonium bicarbonate (ABC), and tris-(2-carboxyethyl) phosphine (TCEP) were purchased from Sigma (St. Louis, MO, USA). Ammonium acetate (NH<sub>4</sub>OAc) was purchased from Merck (Darmstadt, Germany).  $\beta$ -Mercaptoethanol was obtained from Amersham Biosciences (Uppsala, Sweden) and rabbit muscle creatine kinase (CK) from Roche (Mannheim, Germany). Tuning Mix, a mixture of fluorinated phosphazenes reference G2421A was from Agilent Technologies (Santa Clara, CA, USA). Acetonitrile was procured from Carlo Erba (Milan, Italy) and formic acid 90% (FA) from Fisher (Loughborough, UK). The P3P triphosphorylated peptide from Raf-1 with the sequence RPRGQRDpSpSYpY-WEIEASEV was a product of Fmoc synthesis. Details of the synthesis strategy will be described elsewhere. All solvents and buffers were prepared using 18 M $\Omega$  purified water (MilliQ reagent grade system, Millipore, Billerica, MA, USA).

### *hPEBP and Nucleotides Purification*

Recombinant human phosphatidyl ethanolamine binding protein (hPEBP), with an average  $M_r$  of 20,925.6 Da, was cloned and overexpressed in BL21DE3 *E. coli* as described by Schoentgen and coworkers [24]. The tagless protein was purified by QAE Sephadex A-50 chromatography, isoelectrofocusing, and Blue-Sepharose chromatography, as previously described [25]. Between purification steps, the protein was concentrated by centrifugation at 4000g for 20 min using a Vivaspin centrifugal separation vial with a 3 kDa cut-off filter (Sartorius Stedim, Aubagne, France). The purified protein was initially stored in a solution of 20 mM MES, 1 mM  $\beta$ -mercaptoethanol. As  $\beta$ -mercaptoethanol was found by MS to form covalent adducts with hPEBP by targeting the protein's free cysteines [26], a procedure for the removal of  $\beta$ -mercaptoethanol by various reducing agents was developed. The addition of TCEP to hPEBP stock solutions at a final concentration of 1 mM for one hour at 37 °C efficiently reduced the  $\beta$ -mercaptoethanol–cysteine bond [27]. The reduced protein solution was then dialyzed against water using Micro Dispodialyzer (Harvard Apparatus, Holliston, MA, USA) with a 10 kDa cut-off to eliminate excess TCEP. Purified hPEBP was stored in milliQ water with 2 equivalents of TCEP at –80 °C.

Commercial GTP, FMN, ADP and CTP nucleotides used in this study were desalted to prevent Na<sup>+</sup> adduct formation. For this purpose, a protocol derived from the RNA-desalting procedure of Limbach et al. was set up [28]. Desalting of nucleotide water solutions was achieved by precipitation in the presence of 2.5 M ammonium acetate and 70% ethanol at room temperature. The supernatant was discarded and the pellet dissolved in water for MS monitoring of adduct formation. After one to three precipitation cycles, the pellet was dried without heat in a Speed-Vac apparatus (Thermo Scientific, Waltham, MA, USA). Stock solutions at 5 mM in milliQ water were stored at –80 °C.

### *Micro Gel Filtration (MGF)*

Micro gel filtration spin columns (Zeba, Thermo Scientific) with a 7 kDa cut-off were used to desalt CK and RNase A and, where applicable, to eliminate excess ligand after incubation with the protein. Spin columns with a bed volume of 75 or 500  $\mu$ L were used. Efficient elimination of the conditioning buffer from the gel filtration phase required five cycles of equilibration with volumes of 50 or 300  $\mu$ L of incubation buffer, respectively. The sample was loaded onto the spin column and the collection performed immediately with a 30 s centrifugation at 1000 $\times$ g or 1500 $\times$ g, respectively.

### *Sandwich Infusion Method*

To limit protein consumption, a “sandwich” infusion method was developed. The syringe was filled with a volume of 10  $\mu$ L of sample trapped between volumes of solvent. The front solvent is used to clear the dead volume of the system, while the back solvent is used to push the sample into the source. Samples were infused at a flow rate of 1.3  $\mu$ L/min. Proteins in non-denaturing buffer analyzed with H<sub>2</sub>O:ACN:FA 49.8:50:0.2 (vol/vol) as a sandwich solvent exhibited a wide bell-shaped charge state distribution while proteins analyzed with non-denaturing buffer as sandwich solvent exhibited the high  $m/z$ , very narrow charge state distribution which is typical of native species [29]. Therefore, to preserve the native structure of the protein, non-denaturing buffer was systematically used for sandwich infusion.

### *Mass Spectrometry*

All MS measurements were performed in an ESI-ion trap model Ultra HCT PMT Discovery or Esquire HCT (Bruker, Bremen, Germany), or in a 4 GHz maXis ESI-UHR-Qq-TOF (Bruker).  $K_D$  measurement and kinetics experiments were performed by preparing solutions containing a fixed concentration of protein with a gamut of ligand concentrations. After incubation, samples were treated by MGF prior to MS measurement, or analyzed directly in MS. Samples were introduced using the above “sandwich” infusion method with a nebulizer gas pressure of 10 psi and a heated glass-capillary temperature of 120 °C.

On the ESI-ion traps the spray voltage was set to –3.4 kV, the dry gas flow to 5 L/min. For native protein ions detection, the capillary-skimmer voltage difference was adjusted to optimize detection of high  $m/z$  peaks and to preserve the noncovalent interactions between the protein and the ligand. Data was acquired at 8100  $m/z.s^{-1}$  for a minimum of 2 min using EsquireControl ver. 6.2 software (Bruker), and processed using DataAnalysis ver. 4.0 SP2 (Bruker). The maXis was run using a spray voltage of –4 kV and a dry gas flow of 4 L/min. The iSCID parameter was adjusted to preserve noncovalent interactions between the protein and the ligand and, in the case of CK, between dimer

subunits. Data was acquired at 0.2 or 1 Hz for a minimum of 2 min using micrOTOFcontrol ver. 3.1 (Bruker) and processed using DataAnalysis ver. 4.0 SP4 (Bruker). Peak intensities were used for the calculation of bound fractions as follows:

$$f_{\text{bound}} = \frac{\sum_n I(\text{PL}^{n+})}{\sum_n I(\text{PL}^{n+}) + I(\text{P}^{n+})} \quad (2)$$

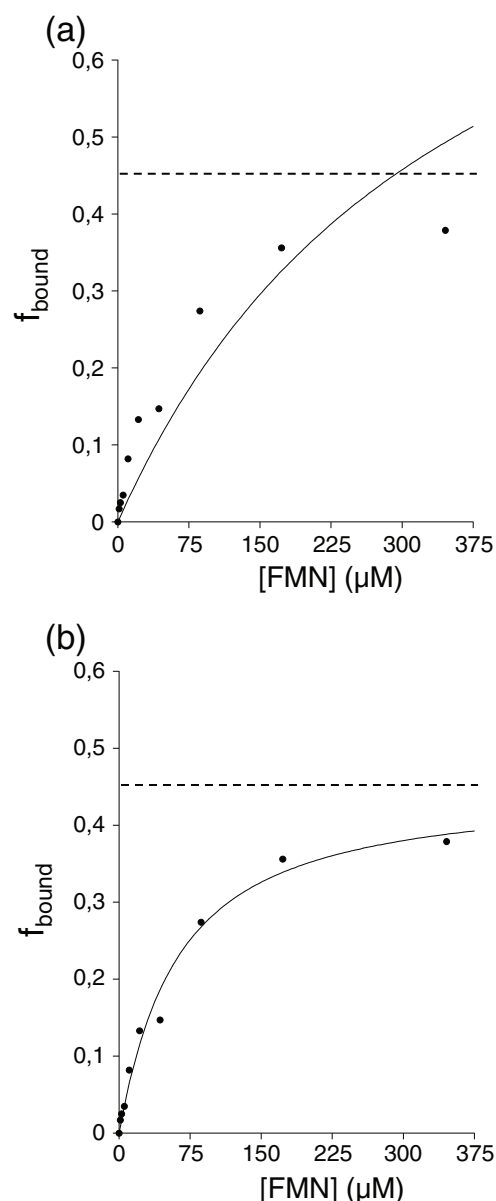
where  $I$  represents the intensity on the mass spectrum, PL and P, the protein-ligand complex and the free protein, respectively, and  $n$ , the ion charge state. Nonlinear regression analyses were performed using Kaleidagraph ver. 4.1.0. Errors are given as residual standard errors from the least-square regression process. To determine the error on  $K_D$  independently of the error introduced by floating a second parameter ( $f_{\text{sat}}$ ) in Equation (3) or (13) for the sake of direct comparison with the error on  $K_D$  determined using Equation (1), the  $f_{\text{sat}}$  value found by regression was re-injected as a constant and a second round of regression performed. This produces the same  $K_D$  value with the same  $R^2$  as the first round of regression.

## Results and Discussion

### Correcting for Gas-Phase Dissociation

The present study was initiated with the examination of the binding of hPEBP to nucleotides, for which relative affinities had been evaluated but  $K_D$  values were unknown [30]. The cone voltage parameter was optimized to improve the transmission of ions of high mass-to-charge ratio and to preserve noncovalent interactions by limiting concomitant dissociation. To investigate the binding of hPEBP to FMN, the equilibrium dissociation constant was determined by increasing FMN concentration while hPEBP concentration was kept constant. The bound fraction plotted as function of ligand concentration (Figure 1) does not tend to 1 as expected, appearing instead to plateau below 0.5 (dashed line). This observation indicates that only a fraction of complexes in solution is preserved during analysis in the gas phase and that gas-phase dissociation (GPD) occurs in spite of the cone voltage optimization. Even in the gentlest conditions, some amount of dissociation could not be prevented. As a consequence, the image of the equilibrium observed in the gas phase does not directly reflect the equilibrium in solution. Accordingly, when parameters of the classical equation (Equation (1)), which does not take into account the apparent partial dissociation in the gas phase, are fitted by nonlinear regression analysis to FMN-hPEBP binding data points, the fit is poor ( $R^2=0.83$ ) (Figure 1a).

To account for this phenomenon, we decided to attempt to correct for the partial GDP. For this purpose, we empirically introduced a floating parameter, which we called  $f_{\text{sat}}$ , which would correspond to the experimental bound fraction at saturation. If we hypothesize that partial GDP is constant throughout the experiment, we



**Figure 1.** MS-measured bound hPEBP fraction as a function of FMN concentrations following incubation of 12.6  $\mu\text{M}$  hPEBP with 0–346  $\mu\text{M}$  FMN at 37 °C in 20 mM ammonium bicarbonate, pH 8.4. Equation parameters were fitted to the data by nonlinear regression analysis using (a) Equation (1) to determine  $K_D$  without GPD correction ( $K_D=349\pm 55$   $\mu\text{M}$ ;  $R^2=0.83$ ); and (b) Equation (3) to determine  $K_D$  while applying GPD correction ( $K_D=54\pm 10$   $\mu\text{M}$ ,  $f_{\text{sat}}=0.45$ ;  $R^2=0.99$ )

can introduce the  $f_{\text{sat}}$  parameter as a multiplier in Equation (1), thus yielding Equation (3):

$$f_{\text{bound}} = f_{\text{sat}} \cdot \frac{[\text{P}]_0 + [\text{L}]_0 + K_D - \sqrt{([\text{P}]_0 + [\text{L}]_0 + K_D)^2 - 4[\text{P}]_0 [\text{L}]_0}}{2[\text{P}]_0} \quad (3)$$

The above data points were fitted with this corrected equation. Nonlinear regression analysis with Equation (3)



gave a much better fit, with  $R^2=0.99$ . In this case, a value of  $f_{\text{sat}}=0.45$  is found, in keeping with our empirical estimate (Figure 1b). The correlation coefficient is much higher with Equation (3) than with Equation (1), confirming our hypothesis that GPD is uniform throughout the experiment. Even in optimized transmission conditions, there is still 55% of dissociation in the gas phase in our ion trap instrument. Thus, it is crucial to take into account GPD in MS to estimate a  $K_D$ , which best reflects the  $K_D$  in solution. For the hPEBP-FMN system, correcting the GPD changes  $K_D$  by a factor of six.

It should be noted that the ratio of bound protein over free protein is often reported as the measured quantity, as it can be conveniently computed directly from the ratio of spectral PL and P intensities. However, this does not allow for the direct detection of the bound fraction at saturation. In contrast, as shown above, the measurement of the ratio of the bound protein over total protein makes the evaluation of the presence of GPD straightforward.

Dissociation in the gas phase can be considered as an irreversible step undergone by a fraction of the complex in the course of solution-to-gas transfer (Scheme 1):



Scheme 1. Solution to gas-phase transfer reaction

where the complex in solution ( $PL_{\text{sol}}$ ) partially dissociates in the gas phase to give free protein ( $P_{\text{GPD}}$ ) and free ligand ( $L_{\text{GPD}}$ ). If we define  $t_{\text{GPD}}$  as the amount of time in which the complex can undergo dissociation in the instrument interface, the  $[PL]_{\text{sol}}(t_{\text{GPD}})$  expression corresponds to the complex surviving the solution-to-gas transfer. We call the intact gas-phase complex  $PL_G$ :

$$[PL]_{\text{sol}}(t_{\text{GPD}}) = [PL]_G \quad (4)$$

To keep  $t_{\text{GPD}}$  constant, it is essential to keep the instrument voltages and spray conditions unchanged over the course of  $K_D$  measurements. Provided that this condition is respected, we have shown that dissociation in the instrument is constant throughout the experiment and that the decrease in complex ions population is proportional to the complex population in solution.

This translates mathematically as:

$$[PL]_G = [PL]_{\text{sol}} \cdot f_{\text{sat}} \quad (5)$$

Dividing both sides of the equation by the total protein concentration, Equation (5) combined with Equation (1) gives Equation (3). Namely, the binding fraction measured in the gas-phase reflects both GPD and the binding fraction in solution.

### Dealing with Gas-Phase Aggregation

Gas-phase aggregation is a well-known phenomenon observed for a number of protein–small molecule binding systems. As exposed in the preceding section, the  $K_D$  of the hPEBP-FMN complex (Figure 1b) was first determined by applying nonlinear regression analysis on the parameters of Equation (3) to fit to the bound fraction data at 1:1 stoichiometry. A value of  $K_D=54 \mu\text{M}$  was found for the complex. However, the measured bound fraction is skewed by aggregation. Even if GPD is corrected for by using Equation (3), aggregation is not taken into consideration thus far and consequently, an error in the estimation of  $K_D$  potentially remains.

In the course of studying the interaction of hPEBP with another nucleotide, GTP, apparent stoichiometries up to 1:4 were observed when reacting  $18 \mu\text{M}$  hPEBP with  $500 \mu\text{M}$

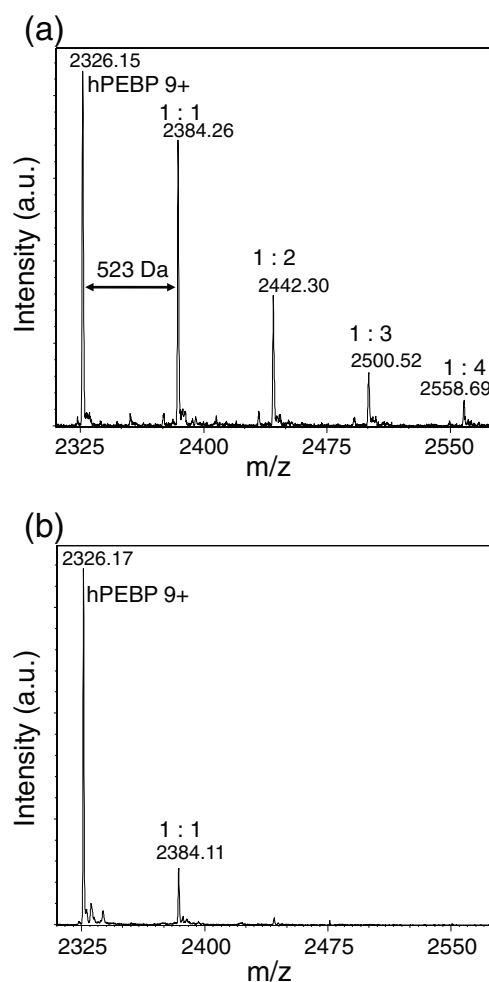
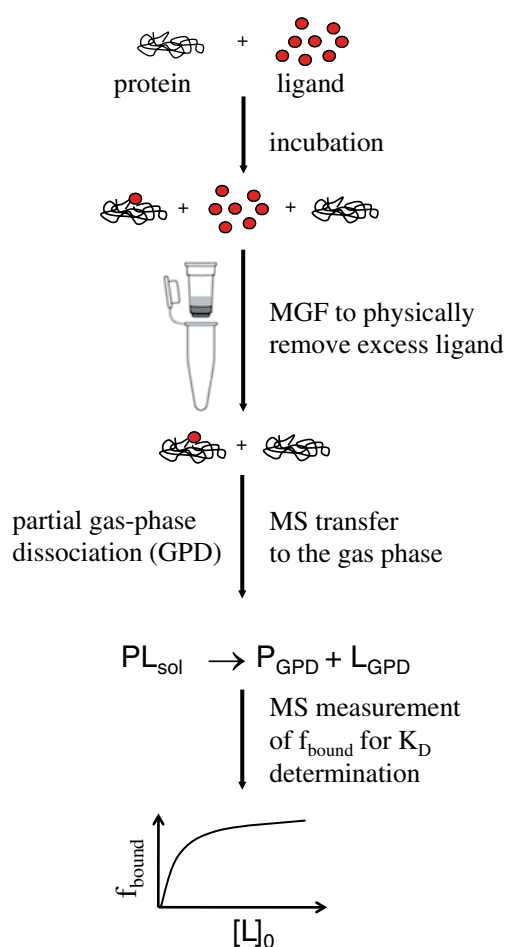


Figure 2. Positive ion native ESI mass spectra (charge state 9+) of  $18 \mu\text{M}$  hPEBP with  $500 \mu\text{M}$  GTP (delta-mass= $+523 \text{ Da}$ ) following incubation at  $37 \text{ }^\circ\text{C}$  in  $20 \text{ mM}$  ammonium bicarbonate, pH 8.4, (a) without micro gel filtration, (b) with micro gel filtration. Peaks are marked with their apparent stoichiometry

GTP (Figure 2a). The protein peaks distribution is typical of the presence of nonspecific interactions. Nonspecific interactions overlapped with specific interactions make data interpretation more complex, leading to the common strategy of decreasing the ligand concentration to avoid aggregation. This strategy however has an important drawback, in that low affinity complexes may not be observable at such low concentrations. Consequently, we decided to devise an alternative strategy aimed at preventing aggregation.

To this end, while the range of concentrations used was kept high enough to allow for complex measurement, excess free ligand present after incubation of the partners was eliminated by adding a step of micro gel filtration (MGF). After MGF, the apparent binding stoichiometry is down to 1:1 (Figure 2b). This step thus seems efficient at free ligand removal and prevention of aggregation. Adding the MGF step makes it possible to work with a large excess of ligand in the case of low affinity complexes ( $K_D \geq 1 \mu\text{M}$ ), while eliminating the risk of aggregation due to excess free ligand. This leads to a new experimental flow (Figure 3).



**Figure 3.** New experimental flow to determine low affinity  $K_D$  by ESI-MS

The main object of this work is to produce a  $K_D$  determination accurately reflecting the solution equilibrium. Both aggregation and GPD can affect  $K_D$  accuracy. Therefore our approach was designed to deal with both GPD and aggregation in the same experimental workflow and to integrate modeling of the system's behavior according to its kinetics into the approach.

### *An Integrated Approach and Model to Deal with GPD and Aggregation*

The new experimental flow for the determination of low affinity dissociation constants is presented in Figure 3. Scheme 2 is the general model corresponding to this new experimental flow, where  $[P]_1$ ,  $[L]_1$  and  $[PL]_1$  are the protein, ligand, and complex concentrations at equilibrium 1 in solution starting from  $[P]_0$  and  $[L]_0$  concentrations.

$[P]_F$ ,  $[L]_F$ , and  $[PL]_F$  are the protein, ligand, and complex concentrations immediately after MGF,  $[P]_2$ ,  $[L]_2$ , and  $[PL]_2$  are the protein, ligand, and complex concentrations reached when equilibrium is attained by liquid-phase dissociation (LPD) after MGF,  $[P]_{\text{GPD}}$ ,  $[L]_{\text{GPD}}$  are the protein and ligand concentrations resulting from gas-phase dissociation (GPD), and  $[P]_M$ ,  $[PL]_M$  are the relative protein and complex concentrations derived from the relative protein and complex peak intensities on the mass spectrum.

Equilibrium 1 is reached in solution after incubation of the protein with the ligand. MGF filtration is then used to eliminate free ligand with recovery of the protein in free or bound form. Equilibrium 2 in solution can be reached if the system follows fast kinetics or if one waits long enough. During MS analysis, the sample is transferred to the gas phase with the risks of dissociation previously exposed. The measurement is performed on the peak intensities for  $P_M$  and  $PL_M$  on the mass spectrum.

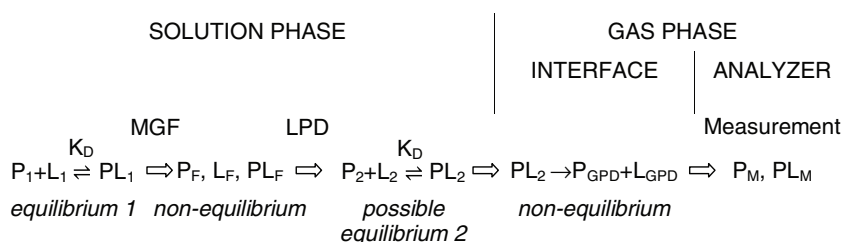
The dissociation constant  $K_D$  is the ratio of the dissociation rate constant  $k_{\text{off}}$  to the association rate constant  $k_{\text{on}}$ :

$$K_D = \frac{k_{\text{off}}}{k_{\text{on}}} \quad (6)$$

For low affinity systems, different combinations of  $k_{\text{off}}$  and  $k_{\text{on}}$  values can produce a high  $K_D$  value. The limit cases are a high  $k_{\text{off}}$  with a medium  $k_{\text{on}}$ , and a low  $k_{\text{off}}$  with a very low  $k_{\text{on}}$ . In the first case, the dissociation to a second equilibrium is very fast because  $k_{\text{off}}$  is high. In contrast, in the second case, the dissociation after MGF is very slow. The following model (Scheme 3) presents the limit cases subsequent to the MGF step.

In Scheme 3,  $[P]_i(t)$ ,  $[L]_i(t)$ , and  $[PL]_i(t)$  are the protein, ligand and complex concentrations in the intermediary step, respectively, at any given time  $t$  after MGF.

The “fast dissociation” case corresponds to a situation where the complex dissociates very quickly, so that the



Scheme 2. General model describing system behavior from the solution to the instrument analyzer

system reaches a new equilibrium in the analysis dead-time. In the “no dissociation” case on the other hand, the system has a very low dissociation rate constant, so that the complex does not have enough time to dissociate in any significant way in the analysis dead-time. Mathematical developments for the limit cases are presented below.

“Fast Dissociation” Case

The equation for  $K_D$  determination with a floating  $f_{sat}$  is derived from kinetic equations, the equations for mass action law and the law of conservation of mass as well as equations associated with the particular conditions of the experiment, such as MGF as shown in Table 1.

$f_{bound}$  is the bound fraction as measured by ESI-MS. The binding of a small ligand with a protein is assumed not to significantly alter the protein ionization and detection efficiencies during MS measurement.

Under this assumption, the intensities in the mass spectrum reflect the relative gas-phase concentrations of the complex and the free protein, so that the  $I(PL)_M / I(P)_M$  ratio corresponding to  $[PL]_M / [P]_{M tot}$  can be calculated by

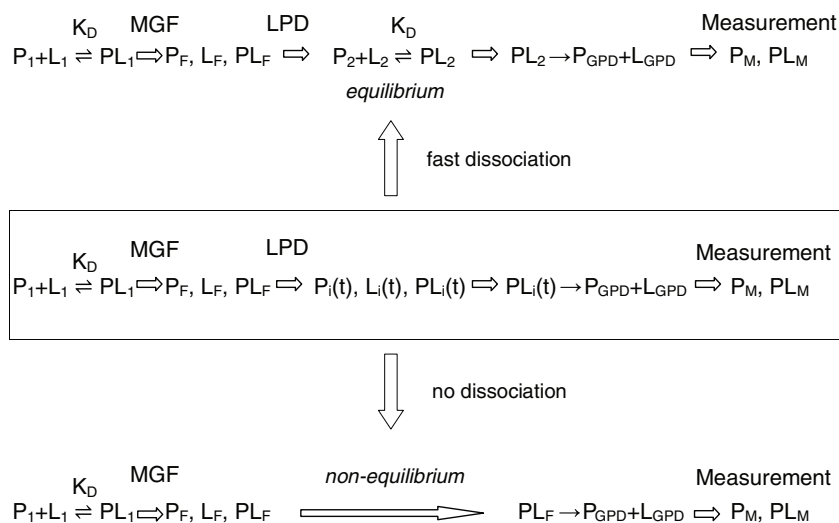
adding the protein peak intensities for PL or for P at different charge states. The  $a$  and  $b$  constants represent the fraction of recovery after the MGF step of protein and ligand, respectively. Since the MGF is optimized to recover most of the protein, constant  $a$  is close to 1, whereas constant  $b$  is close to 0 if the removal of excess ligand is efficient. During this step, the bound and free protein can be expected to have sufficiently close molecular masses and shapes to behave similarly, so that recovery of the material is identical for both.

Below is the mathematical development used to obtain a direct expression for the determination of the solution  $K_D$  from the mass-measured  $f_{bound}$  ratio:

$$f_{bound} = \frac{[PL]_M}{[P]_{M tot}} \tag{7}$$

Using mass conservation equations, we can demonstrate that:

$$[P]_{M tot} = a[P]_0 \tag{8}$$



Scheme 3. Limit cases to the general model

**Table 1.** Scheme and Equations Relating to the “Fast Dissociation” Case

	SOLUTION PHASE			GAS PHASE	
				INTERFACE	ANALYZER
	$P_1+L_1 \xrightleftharpoons{K_D} PL_1 \xrightarrow{\text{MGF}} P_F, L_F, PL_F \xrightarrow{\text{LPD}} P_2+L_2 \xrightleftharpoons{K_D} PL_2 \xrightarrow{\text{non-equilibrium}} PL_2 \rightarrow P_{GPD}+L_{GPD} \xrightarrow{\text{Measurement}} P_M, PL_M$				
	<i>equilibrium 1</i>			<i>non-equilibrium</i>	
Kinetic or equilibrium constant	$K_D = \frac{[P]_1[L]_1}{[PL]_1}$		$K_D = \frac{[P]_2[L]_2}{[PL]_2}$	$[PL]_2(t_{GPD}) = [PL]_2 \cdot f_{sat}$	
Mass conservation	$[P]_0 = [P]_1 + [PL]_1$ $[L]_0 = [L]_1 + [PL]_1$	$[P]_{F\text{ tot}} = [P]_F + [PL]_F$ $[L]_{F\text{ tot}} = [L]_F + [PL]_F$	$[P]_{2\text{ tot}} = [P]_2 + [PL]_2$ $[L]_{2\text{ tot}} = [L]_2 + [PL]_2$	$[P]_{G\text{ tot}} = [P]_G + [PL]_G$ $[L]_{G\text{ tot}} = [L]_G + [PL]_G$	$[P]_{M\text{ tot}} = [P]_M + [PL]_M$ $[L]_{M\text{ tot}} = [L]_M + [PL]_M$
Particular conditions		$[P]_F = a[P]_1$ $[L]_F = b[L]_1$ $[PL]_F = a[PL]_1$	$[P]_{2\text{ tot}} = [P]_{F\text{ tot}}$ $[L]_{2\text{ tot}} = [L]_{F\text{ tot}}$	$[P]_G = [P]_{GPD} + [P]_2$ $[L]_G = [L]_{GPD} + [L]_2$ $[PL]_G = [PL]_2(t_{GPD})$	$[P]_{M\text{ tot}} = [P]_{G\text{ tot}}$ $[PL]_M = [PL]_G$ $f_{bound} = \frac{[PL]_M}{[PL]_M + [P]_M}$

According to equations associated with the particular conditions of the experiment, we know that:

$$[P]_M = [PL]_2 \cdot f_{sat} \quad (9)$$

The combination of Equations (7), (8), and (9) gives:

$$f_{bound} = \frac{[PL]_2 \cdot f_{sat}}{a[P]_0} \quad (10)$$

By combining kinetic equations, the equations for the law of conservation of mass as well as equations associated

with the particular conditions of the experiment,  $[PL]_2$  can be expressed as a function of  $a$ ,  $b$ ,  $[P]_0$ ,  $[L]_0$ ,  $K_D$  and  $[PL]_1$  (Equation 11).

If  $[PL]_1$  can be expressed as a function of  $[P]_0$ ,  $[L]_0$ , and  $K_D$ , it will lead to an expression of  $[PL]_2$ , which only depends on these fixed values. An equation equivalent to the classical quadratic Equation (1) can be used to express  $[PL]_1$  (Equation 12).

Combining and reducing Equations (10), (11), and (12) produces the final expression of  $f_{bound}$  as a function of the fixed  $a$ ,  $b$ ,  $[P]_0$ , and  $[L]_0$  values and of the  $K_D$  and  $f_{sat}$  parameters (Equation 13).

$$[PL]_2 = \frac{a[P]_0 + b[L]_0 + (a-b)[PL]_1 + K_D - \sqrt{(a[P]_0 - b[L]_0 - (a-b)[PL]_1 - K_D)^2 + 4a[P]_0 K_D}}{2} \quad (11)$$

$$[PL]_1 = \frac{[PL]_0 + [L]_0 + K_D - \sqrt{([P]_0 - [L]_0 - K_D)^2 + 4[P]_0 K_D}}{2} \quad (12)$$

$$f_{bound} = \frac{f_{sat}}{2a[P]_0} \cdot \left( \frac{3a-b}{2}[P]_0 + \frac{a+b}{2}[L]_0 + \frac{2+a-b}{2}K_D - \frac{a-b}{2} \sqrt{([P]_0 - [L]_0 - K_D)^2 + 4[P]_0 K_D} - \sqrt{\left( \frac{a+b}{2}[P]_0 - \frac{a+b}{2}[L]_0 - \frac{2+a-b}{2}K_D + \frac{a-b}{2} \sqrt{([P]_0 - [L]_0 - K_D)^2 + 4[P]_0 K_D} \right)^2 + 4a[P]_0 K_D} \right) \quad (13)$$

In the “fast dissociation” case, nonlinear regression with Equation (13) gives direct access to the  $K_D$  in solution and to the  $f_{\text{sat}}$  linked to MS measurement of the studied system.

### “No Dissociation” Case

The mathematical development for the “no dissociation” case (Table S1) is simplified by the fact that there is no re-equilibration in solution after MGF. The beginning of the mathematical demonstration is similar to the “fast dissociation” case.

In the “no dissociation” case, the equivalent to Equation (10) is Equation (14). We know that  $f_{\text{bound}}$  can be expressed as a function of  $a$ ,  $[P]_0$ ,  $[L]_0$ , and  $f_{\text{sat}}$  by inserting Equation (12) into Equation (15), and inserting the final expression in Equation (14) to give Equation (16):

$$f_{\text{bound}} = \frac{[PL]_{\text{F}} \cdot f_{\text{sat}}}{a[P]_0} \quad (14)$$

$$\text{We know that } [PL]_{\text{F}} = a[PL]_1 \quad (15)$$

$$f_{\text{bound}} = f_{\text{sat}} \cdot \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4[P]_0[L]_0}}{2[P]_0} \quad (16)$$

Equation (16) is identical to Equation (3) found in the first section, meaning that the introduction of a MGF step for a non-dissociating complex does not alter the final mathematical expression for solution  $K_D$  measurement.

In the “no dissociation” case, the complex in solution stays intact after the MGF step. This means that this type of system is ideal for working with the large excess of ligand necessary for low affinity systems, and takes full advantage of the MGF step to prevent aggregation.

On the other hand, in the “fast dissociation” case, a second equilibrium in solution is reached following the MGF, so that the elimination of aggregation afforded by MGF does not translate into an effective increase in the measured bound fraction. In other words, this limit case corresponds to a situation where the measured species are equivalent to what would be observed directly with a lower gamut of ligand concentrations for which there is no aggregation. There are however practical advantages to applying MGF and Equation (13) to such a system. First, this approach does not necessitate a priori knowledge on the system and the mathematical treatment relies on the occurrence of two instances of the solution  $K_D$  corresponding to equilibriums 1 and 2, which gives the determination internal consistency. Second, the protein–ligand complex is usually formed

in a physiologically relevant buffer which may need to be buffer-exchanged before the MS analysis. Finally, the initial complex formation step will proceed faster at the high ligand concentrations allowed with an MGF step. This may be valuable for proteins and/or ligands with limited stability. In other cases with fast dissociating complexes where it is experimentally feasible to skip the MGF step and to decrease the ligand concentration, Equation (3) can be applied directly for  $K_D$  determination.

### “Partial Dissociation” Case

Between the two limit cases presented above, “partial dissociation,” corresponding to an intermediary case, occurs. Mathematical development for this case was attempted and is presented in [Supplementary Materials](#). This leads to the fairly complex Equation S14, which contains four remaining unknowns. Attempts at fitting this increased number of parameters would lead to a greater freedom for nonlinear regression to find compatible sets of parameter solutions, so as to render the result potentially meaningless.

Thus, it is more useful to examine conditions for which any low affinity system can be brought to behave according to one of the two limit cases. The next section provides general guidelines for setting up the  $K_D$  determination experiments. For a fine tuning of experimental conditions,

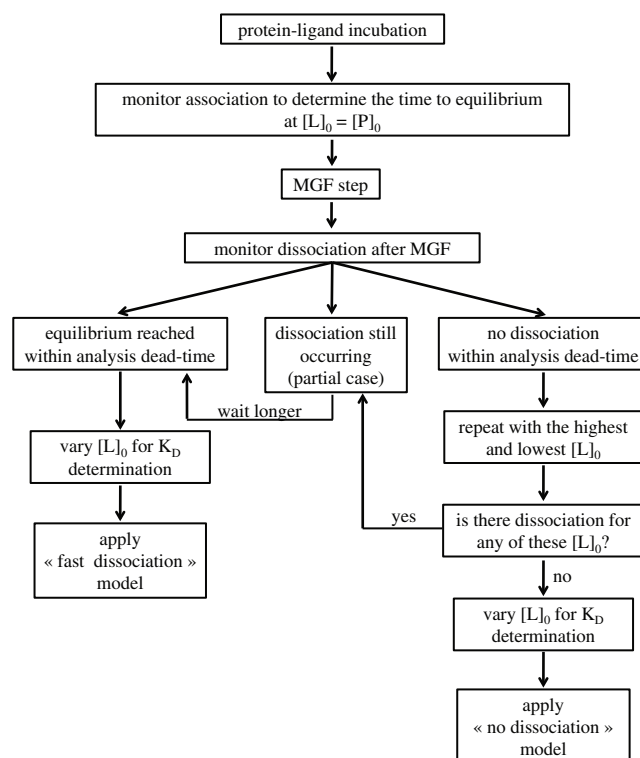


Figure 4. Experimental decision tree to aid in  $K_D$  determination by ESI-MS

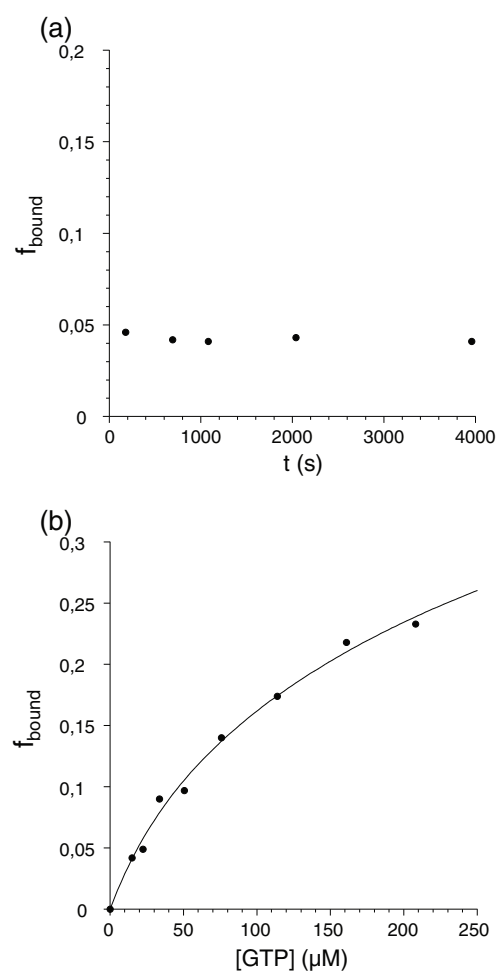
graphs simulating the system's kinetic behavior according to theoretical rate and equilibrium constants (Figures S1 and S2) can be used. The equations for building such simulations for a given set of initial conditions are provided in [Supplementary Materials](#).

### Guidelines to Determine which Model to Apply

In practical situations, the  $K_D$  is usually unknown, with at best an order of magnitude guesstimate for its value. The same can be said for the  $k_{on}$  and  $k_{off}$  values. Some preliminary information is thus necessary to ensure that low affinity  $K_D$  constants are measured in a safe way with the right model applied. The experimental decision tree (Figure 4) shows a strategy to gather preliminary information and to select the model to apply. The first step consists in determining a safe incubation time to ensure that complexes in solution are studied at equilibrium. This should be done at the 1:1 protein:ligand stoichiometric point where the longest association times occur, as shown by a plot of the association behavior of binding systems (Figure S1). The next step consists in applying MGF and determining which case the studied system falls into. This is done by simply monitoring dissociation post-MGF over a few time points. If dissociation is still occurring after the analysis dead-time, the studied system falls into the “partial dissociation” case. The system can be then brought back to the “fast dissociation” case by allowing it to re-equilibrate for a longer period of time after MGF. Alternatively, a system's behavior can be manipulated to reach one of the limit cases by playing on the range of initial ligand concentrations. This approach requires a finer understanding of the effect of ligand concentrations on kinetic behavior, and is described in [Supplementary Materials](#) along with simulation tools to assist in experimental choices. Once it is ensured that the studied system falls either into the “fast dissociation” or the “no dissociation” case,  $K_D$  determination can be performed by varying the ligand concentration and applying the appropriate model.

### Application to hPEBP-GTP

As shown above, this system is sensitive to aggregation. Preliminary monitoring of association kinetics of 18  $\mu\text{M}$  hPEBP with a ligand concentration close to the protein concentration and with a high GTP concentration (36 and 500  $\mu\text{M}$ , respectively) was conducted to determine a safe incubation time. The equilibrium is reached in less than 2 min (data not shown). To figure out which model applies to this system, dissociation kinetics was monitored after a MGF step following 8 min of incubation in the above conditions (Figure 5a). The complex dissociates very fast in solution, so that in less than 180 s, a new equilibrium has already been reached. Consequently, the hPEBP-GTP complex follows the “fast dissociation” model. Applying nonlinear regression analysis with Equation (13) on hPEBP-



**Figure 5.** (a) MS-measured bound hPEBP fraction as a function of time after micro gel filtration. The complex formed by incubating 18  $\mu\text{M}$  hPEBP with 36  $\mu\text{M}$  GTP at 37 °C in 20 mM ABC, pH 8.4, was micro-purified using Zeba microspin columns. The time to reach 95% dissociation is below 180 s. (b) MS-measured hPEBP bound fraction as a function of ligand concentration after micro gel filtration. The complex was formed by incubating 18  $\mu\text{M}$  PEBP with 0–208  $\mu\text{M}$  GTP at 37 °C in 20 mM ABC, pH 8.4. The  $K_D$  and  $f_{\text{sat}}$  parameters of Equation (13) were fitted by nonlinear regression analysis to the data

GTP  $f_{\text{bound}}$  data gives a very good fit (Figure 5b) with  $R^2=0.99$ ,  $K_D=40\pm 1.0$   $\mu\text{M}$  and  $f_{\text{sat}}=0.52$ .

Since hPEBP and GTP are stable over time and the incubation buffer in this instance is compatible with native mass spectrometry, one can choose to skip the MGF step to determine  $K_D$ . In this case, the ligand concentration must be decreased to avoid aggregation. To this end, a range of ligand concentrations was screened to determine the highest point for which, in the absence of MGF, no stoichiometries above 1:1 are observed. Since the longest association times occur around the protein:ligand stoichiometric point, the association time in these conditions was confirmed to be below 2 min. In the case of the hPEBP-GTP complex, the fit using Equation (3) is very good with  $R^2=0.99$  and gives

**Table 2.**  $K_D$  Constants Determined Through MS Measurement with and without  $f_{sat}$  Correction versus  $K_D$  Constants Determined by In-solution Methods

P-L complex	Experimental conditions	$K_D$ measurement without GPD correction (Equation (1))		$K_D$ measurement with GPD correction (Equations (3) or (13))			Literature $K_D$ ( $\mu\text{M}$ )
		$K_D$ ( $\mu\text{M}$ )	$R^2$	$K_D$ ( $\mu\text{M}$ )	$f_{sat}$	$R^2$	
RNase-CTP	10 mM $\text{NH}_4\text{OAc}$ , pH 6.8; 25°C	17 $\pm$ 1.4	0.911	1.6 $\pm$ 0.24	0.47 $\pm$ 0.07	0.971	0.7 [31]; 0.5 - 2 <sup>a</sup> [32–34]
PEBP-P3P	20 mM ABC, pH 7.4; 25°C	160 $\pm$ 15	0.927	11 $\pm$ 0.74	0.26 $\pm$ 0.02	0.995	20 [35]
HEWL-NAG3	50 mM $\text{NH}_4\text{OAc}$ , pH 6.6; 37°C	19 $\pm$ 1.2	0.993	20 $\pm$ 1.2	1.02 $\pm$ 0.08	0.993	8.6 [36]
(CK) <sub>2</sub> -2ADP	10 mM ABC, pH 7.9; 37°C <i>ESI-UHR Q-TOF</i>	15 $\pm$ 1.2	0.951	4.5 $\pm$ 0.54	0.61 $\pm$ 0.17	0.964	24 [16] – 33 <sup>b</sup> ; 40 [37]
PEBP-GTP	20 mM ABC, pH 8.4; 37°C	140 $\pm$ 6.7	0.984	36 $\pm$ 0.62	0.42 $\pm$ 0.03	0.999	n.d.
PEBP-GTP	20 mM ABC, pH 8.4; 37°C + MGF	230 $\pm$ 26	0.896	55 $\pm$ 5.0	0.40 $\pm$ 0.07	0.970	n.d.
PEBP-GTP	20 mM ABC, pH 8.4; 37°C + MGF, <i>ESI-UHR Q-TOF</i>	98 $\pm$ 4.6	0.962	40 $\pm$ 1.0	0.52 $\pm$ 0.05	0.993	n.d.
PEBP-FMN	20 mM ABC, pH 8.4; 37°C + MGF	240 $\pm$ 28	0.854	35 $\pm$ 3.1	0.25 $\pm$ 0.05	0.967	n.d.

Unless otherwise stated, an ESI-IT instrument was used.

[31] MS, Zhang et al. 2003; [32] ITC, Straume et al. 1992; [33] DSC, Wiseman et al. 1989; [34] CD, Jones et al. 2002; [35] SPR, Park et al. 2006; [36] fluorescence, Schindler et al. 1977; [16] MS, Daubenfeld et al., 2006, [37] ligand radio-labeling, Burbaum et al. 1989.

For PEBP-nucleotide complexes, there is no reference  $K_D$  in solution. However, the measured  $K_D$  values rank the same as previously determined ranking of affinities.

Errors are standard errors of fitting. For the sake of comparison with errors on  $K_D$  determined using Equation (1), errors on  $K_D$  with  $f_{sat}$  correction (Equation (3) or (13)) were calculated by a second round of regression analysis using the determined  $f_{sat}$  as a fixed parameter.

MS=mass spectrometry; ITC=isothermal titration calorimetry; DSC=differential scanning calorimetry; CD=circular dichroism; SPR=surface plasmon resonance; NMR=nuclear magnetic resonance; n.d. = not determined

<sup>a</sup>Value for 2'-CMP.

<sup>b</sup>Recomputed  $K_D$ , taking into account free protein in peaks of apparent stoichiometry >1, as modeled by the authors.

$K_D=36\pm0.62$   $\mu\text{M}$  with  $f_{sat}=0.42$  (data not shown). The similarity of these values with values obtained at higher ligand concentrations with a MGF step confirms that the new experimental protocol is efficient for preventing aggregation.

With the right tools in hand to deal with aggregation and the effect of its remedy on kinetic behavior, the model can be validated from the point of view of the GPD correction by applying this approach to other systems.

### Comparison of $K_D$ Values Determined Using GPD Correction with $K_D$ Values Determined in Solution

To validate the model, the accuracy of the GPD correction and the correlation of the MS-measured equilibrium dissociation constants  $K_D$  with literature  $K_D$  values were evaluated. The degree of data correlation when performing regression analysis using theoretical Equation (1), which does not take GPD into account, was contrasted with the result obtained using Equation (3), which corrects for GPD (Table 2).

The comparison was made for six systems: RNase-CTP, PEBP-P3P, HEWL-NAG3, (CK)<sub>2</sub>-2ADP, PEBP-FMN, and PEBP-GTP. The correlation coefficient is generally greater with Equation (3) than with Equation (1). Moreover, with Equation (3) the correlation coefficient is improved to a higher extent for complexes which present a low stability in the gas phase, as measured by a low  $f_{sat}$ . For these complexes, the dissociation constant can vary by as much as a factor of 10 as a consequence of substituting Equation (3) for Equation (1) in the regression analysis. For complexes with a higher resistance to GPD such as HEWL-NAG3, a very minor correction for GDP is required, and accordingly the difference between a determination with and without GPD correction is significantly smaller. The improvement in correlation coefficients when

GPD correction is applied suggests that the MS bound fraction indeed reflects both solution binding and a phenomenon of GPD, thus confirming the necessity to include this correction in model building.

The MS-measured  $K_D$  computed with Equation (3) was then compared with the literature values found for  $K_D$  determined by various methods based on the analysis of complexes in solution (Table 2). The comparison was made for four reference systems: RNase-CTP, PEBP-P3P, HEWL-NAG3, and (CK)<sub>2</sub>-2ADP. In the literature,  $K_D$  values have been determined by various methods in solution (ITC, DSC, CD, SPR, fluorescence or ligand radio-labeling). Solution-measured  $K_D$  values of 0.5 to 1.6  $\mu\text{M}$  can be found in the literature for the RNase-CMP system [32–34]. These were found to be close to that of RNase-CTP as determined by another MS study [31]. As 2'CMP was unobtainable worldwide, our method was applied to the RNase-CTP system. The MS measured  $K_D$  value computed with Equation (3) is again closer to the literature values than the one computed with Equation (1). The same observation can be made for the PEBP-P3P system. For the HEWL-NAG3 complex, the MS-measured  $K_D$  values computed with Equation (3) or (1) are nearly identical and both are close to the literature value. We can thus conclude that HEWL-NAG3 is resistant to GPD and can be considered as a reference for GPD. Last, (CK)<sub>2</sub>-2ADP, is a dimer with one binding site per monomer. Again in this case, the  $K_D$  value found with Equation (3) is lower than that found with Equation (1). Both are however lower than the  $K_D$  measured by in-solution method. Van der Rest and coauthors studied this complex by MS and succeeded in discriminating the aggregation phenomenon from specific interactions using a statistical approach [16]. An average  $K_D$  value of 24  $\mu\text{M}$  for the two

independent binding sites was reported, in keeping with the literature [16]. A  $K_D$  value of 33  $\mu\text{M}$  can be found by recomputing the data to take free protein in peaks of higher apparent stoichiometry, as modeled by the authors, into account. For this system a  $K_D$  value of 5  $\mu\text{M}$  was found with Equations (3) and 15  $\mu\text{M}$  with Equation (1). A  $f_{\text{sat}}$  value of 0.61 suggests a significant level of GPD. If GPD also occurred during the Daubenfeld/van der Rest MS analysis, it could lead to an overestimation of the  $K_D$  value, which could explain the observed difference. On the whole, the comparison of the MS measured  $K_D$  values with literature data validates the use of our GPD model to determine a  $K_D$  closer to the reality of solution equilibria.

We chose to calculate a weighted average over all charge states to determine  $K_D$ . However, to study the effect of the charge state on the  $K_D$ , we also determined  $K_D$  based on individual charge state data (Table S3, Supplementary Materials). In a majority of cases, the best fit is obtained with the major charge state, while being close to the fit obtained with the weighted average. Surprisingly, the weighted average sometimes gives a better fit than any of the individual charge states. The six studied systems do not seem to show a unique trend of  $K_D$ -charge state correlation. Until a better understanding of charge state effects is gained, it seems that the use of weighted average data is a reasonable choice.

To calculate binding fractions, we assumed that binding of a small ligand with the target protein does not significantly alter the protein ionization and detection efficiencies during MS measurement. This seems to be the case for the ligands studied in this work. The delta mass observed upon ligand binding corresponds to the mass of a globally neutral molecule. Nucleotides possess two ionic groups of opposite charge and behave as zwitterions. At the near-neutral pH used in these MS measurements, they may have an overall net charge of zero, so that binding to the protein does not directly alter its charge state. If not, it would have led to a significant shift in the charge state distribution, which was not observed. Moreover, these ligands are relatively small compared to the target protein.

For massive ligands, or for ligands that behave as strictly ionic molecules at the pH of measurement, such as ionic detergents or metal ions, this assumption may not hold, potentially leading to effects on the complex ionization and detection efficiency. In such cases, the  $f_{\text{sat}}$  correcting factor would compensate for the effects.

Several methods have been developed in recent years to deal with aggregation or GPD. They have in common the use of a reference protein or ligand. One such strategy is based on half-maximal inhibitory concentration ( $\text{IC}_{50}$ ) measurements in a competitive assay with a non-aggregating ligand [18].  $\text{IC}_{50}$  values determined by MS are then converted into  $K_D$ . The main disadvantage of this method is the need for a reference ligand. In addition, the reference ligand has to meet the double condition of not forming aggregates with the protein, and having a known solution  $K_D$  [18]. Nonspecific binding of charged biomolecules is a

statistical process, independent of protein structure and size [38]. This observation has led to the development of a method using a reference protein to quantify aggregation [16, 38]. However, the stability of nonspecific complexes is sensitive to the nature of the protein ions. Consequently, the reference protein method for correcting aggregation can only be applied provided that GPD of nonspecific complexes is minimized. In the presence of GPD, this method cannot be applied.

There are two existing methods using a reference ligand for quantifying the affinity of protein-ligand complex that are prone to GPD [23, 39]. The reference ligand has to bind with a known affinity and at the same binding site as the ligand of interest. The reference ligand method is based on the MS measurement of the reference complex in a competitive assay with the ligand of interest. In addition, one of the methods requires that the reference ligand is not prone to GPD [39].

The necessity to have a reference ligand or protein is very restrictive, particularly for exploratory projects where little may be known beforehand about the protein binding site. In addition, none of the above methods takes into account both GPD and aggregation. In contrast, the approach described in this work has the advantage not to require a reference protein or ligand. The presence of GPD can easily be checked and resolved, regardless of the type of complex and aggregation controlled upstream of MS analysis.

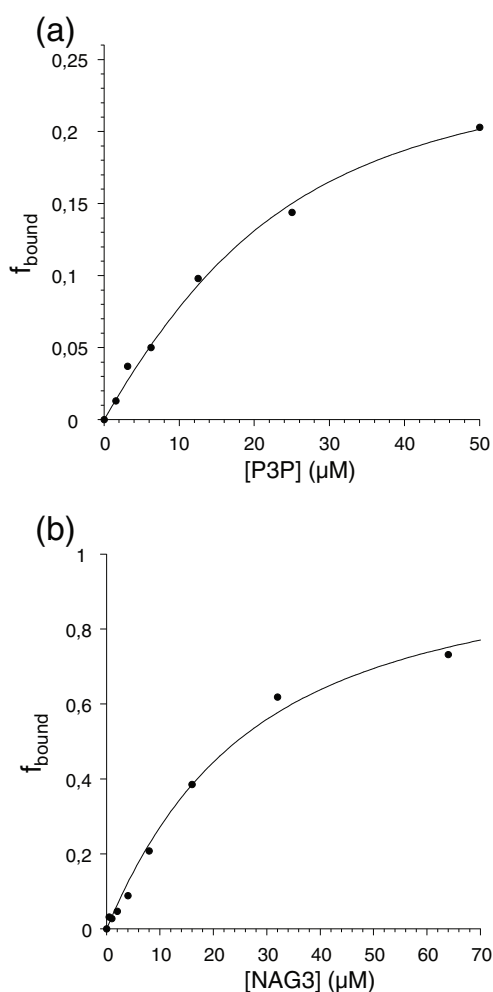
### *Effect of the Type of Noncovalent Interaction on Gas-Phase Dissociation*

Considerations on the effect of the type of molecular interaction in the complex provide a first step in the dissection of the  $f_{\text{sat}}$  parameter. Going from solution to a solventless gas phase environment can alter the strength of noncovalent interactions. Ionic interactions and hydrogen bonds can be strengthened, while the hydrophobic effect is for the most part not conserved [40, 41]. If we consider that the bound fraction measured by MS is a vestige of the complex in solution, that is to say that we measure the ratio of complex that has survived the transfer into the analyzer, then it is reasonable to consider the influence of the type of interactions on the GPD for a given system. To explore this aspect, we can compare the  $f_{\text{sat}}$  results obtained for the HEWL-NAG3 and the hPEBP-P3P systems.

HEWL-NAG3 is an interaction based on hydrogen bonds with hydrophobic contacts [42, 43]. P3P is a triphosphorylated peptide from Raf1, which has been synthesized and used to obtain information on the hPEBP-Raf1 interaction by studying the hPEBP-P3P complex [35]. The hPEBP-P3P interaction can be surmised to be at least partially based on ionic interactions between amino acid residues and the phosphate groups of the triphosphorylated peptide.

While HEWL-NAG3 (Figure 6a) and hPEBP-P3P (Figure 6b) have similar equilibrium dissociation constants, they have radically different  $f_{\text{sat}}$  (Table 2). The bound





**Figure 6.** GPD behavior comparison: MS-measured bound protein fraction as a function of ligand concentration. **(a)** hPEBP (18  $\mu\text{M}$ ) incubated with 0–50  $\mu\text{M}$  P3P in 20 mM ABC, pH 7.4, at 25 °C. **(b)** HEWL (10  $\mu\text{M}$ ) incubated with 0–64  $\mu\text{M}$  NAG3 in 50 mM  $\text{NH}_4\text{OAc}$ , pH 6.6, at 37 °C. The  $K_D$  and  $f_{\text{sat}}$  parameters of Equation (3) were fitted to the data by nonlinear regression analysis

fraction at saturation for HEWL-NAG3 was determined by nonlinear regression analysis to be 1.02. This means that there is apparently no discernible dissociation in the gas phase in the course of the analysis. If so, our model successfully handles this system by retrieving an  $f_{\text{sat}}$  value that conforms to reality. In cases where an  $f_{\text{sat}}$  value close to 1 is found, regression can be repeated with a fixed parameter value  $f_{\text{sat}}=1$ , in other words the classical equation (Equation (1)) can be applied. The hPEBP-P3P complex, on the other hand, reaches a bound fraction of 0.26 at saturation, which means only 26% of the solution complex survives during the MS analysis. The corrective equation (Equation (3)) must be applied here to take into account the GPD of the complex and thus avoid an overestimation of  $K_D$ .

In view of the  $f_{\text{sat}}$  value found for HEWL-NAG3, the hypothesis that hydrogen bonds are sufficient to maintain the complex intact during the solution-gas phase transition is

plausible. The hPEBP-P3P complex on the other hand largely dissociates in the gas phase, implying that ionic interactions and hydrogen bonds are not major contributors to the binding and that positional sequence-dependent interactions relying on the hydrophobic effect play an important part in complex formation. This aspect could be further explored from a statistical point of view by examining a large number of known systems for which information on the type of interactions is available.

### *Effect of the Instrument on the Gas-Phase Dissociation*

A consideration of the parameters affecting GPD also calls attention to the influence of the instrumental setup on the dissociation. In particular, the interface pressure and the cone voltage are likely contributors. To test the effect of the instrument on gas-phase dissociation, we determined the equilibrium dissociation constant of hPEBP-GTP using an ESI-IT (HCT Ultra PTM Discovery, Bruker) and an ESI-UHR-Qq-TOF (maXis, Bruker). Equilibrium dissociation constants determined by ESI-IT MS and ESI-UHR-Qq-TOF MS are within each other's margin of measurement error (Table 2). However, the complex is somewhat more labile in the ESI-IT mass spectrometer, with  $f_{\text{sat}}=0.42$ , than in the ESI-UHR-Qq-TOF ( $f_{\text{sat}}=0.52$ ). The interface pressure could be a factor in the dissociation in the interface as a higher pressure favors complex stability [19, 20]. However, interface pressure is actually higher for ESI-IT than for ESI-UHR-Qq-TOF with values of 6.5 versus 3.5 mbar. Consequently, this parameter cannot explain a higher dissociation in ESI-IT. More likely, differences in energetics linked to the transmission time and voltages may explain the small difference in the percentage of GPD. The equilibrium dissociation constants with GDP correction are close, suggesting that the  $f_{\text{sat}}$  parameter compensates for the level of GPD regardless of the instrument and that consequently, the methodology can be applied independently of instrument choice.

The dissection of the  $f_{\text{sat}}$  parameter with respect to the instrumental setup should be further explored. The effect of various instrument parameters on ion energetics and their subsequent impact on dissociation events opens up the possibility of perfecting our understanding of the GPD phenomenon. The knowledge gained would provide a basis for the design of mass spectrometers geared towards preserving noncovalent interactions. Moreover, a precise physical description of GPD could lead to a predictive model for binding system behavior.

## Conclusion

The present work describes a comprehensive approach dealing with both GPD and aggregation in order to determine a  $K_D$  constant that accurately reflects solution binding. The method gives access to a reliable  $K_D$  by ESI-

MS without requiring reference ligands or proteins, or preliminary knowledge on the tendency of the system to form aggregates or to dissociate in the gas phase. We introduce the GPD correction factor  $f_{\text{sat}}$ , which corresponds to the experimental bound fraction at saturation. The system is freed from aggregation by adding a micro-gel filtration step (MGF) to remove excess ligand after incubation of the ligand-protein. The behavior of the complex is modeled according to its kinetics, allowing for the mathematical development of the corresponding binding equations. The introduction of a MGF step has consequences for the accurate measurement of  $K_D$ , which leads to specific recommendations for the application of this type of methods.

The  $f_{\text{sat}}$  constant reflects the type of interaction in the protein-ligand complex, bringing insight into the nature of the complex.  $f_{\text{sat}}$  also provides a way to quantitatively measure the ability of an instrument to preserve noncovalent interactions. This work opens up perspectives for a fine dissection of  $f_{\text{sat}}$  to help in the design of mass spectrometers specialized in native MS analysis or to gain insight into the nature of noncovalent complexes.

## Acknowledgments

The authors thank Ludivine Errien for technical assistance with the expression and purification of human PEBP, as well as Drs. Laurette Tavel and Christian Damblon for the kind gift of  $^{15}\text{N}$  labeled human PEBP. Support by the ANR through the award of a "Programme Blanc" grant, and funding by the Ligue Nationale contre le Cancer and the Cancéropôle Grand-Ouest are gratefully acknowledged. L.J. benefited from Ph.D. funding by CNRS and Région Centre.

## References

- Ganem, B., Li, Y.T., Henion, J.D.: Detection of noncovalent receptor-ligand complexes by mass spectrometry. *J. Am. Chem. Soc.* **113**, 6294–6296 (1991)
- Ganem, B., Li, Y.T., Henion, J.D.: Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry. *J. Am. Chem. Soc.* **113**, 7818–7819 (1991)
- Katta, V., Chait, B.T.: Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **113**, 8534–8535 (1991)
- Greig, M.J., Gaus, H., Cummins, L.L., Sasmor, H., Griffey, R.H.: Measurement of macromolecular binding using electrospray mass spectrometry. Determination of dissociation constants for oligonucleotide: serum albumin complexes. *J. Am. Chem. Soc.* **117**, 10765–10766 (1995)
- Lim, H., Hsieh, Y.L., Ganem, B., Henion, J.: Recognition of cell-wall peptide ligands by vancomycin group antibiotics: Studies using ion spray mass spectrometry. *J. Mass Spectrom.* **30**, 708–714 (1995)
- Loo, J.A., Peifeng, H., McConnell, P., Tom Mueller, W., Sawyer, T.K., Thanabal, V.: A study of Src SH2 domain protein-phosphopeptide binding interactions by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **8**, 234–243 (1997)
- Zhu, M.M., Rempel, D.L., Du, Z., Gross, M.L.: Quantification of protein-ligand interactions by mass spectrometry, titration, and H/D exchange: PLIMSTEX. *J. Am. Chem. Soc.* **125**, 5252–5253 (2003)
- Roulhac, P.L., Powell, K.D., Dhungana, S., Weaver, K.D., Mietzner, T. A., Crumbliss, A.L., Fitzgerald, M.C.: SUPREX (Stability of unpurified proteins from rates of H/D exchange) analysis of the thermodynamics of synergistic anion binding by Ferric-Binding Protein (FbpA), a bacterial transferrin. *Biochemistry* **43**, 15767–15774 (2004)
- Bieth, J.: In: Fritz, H., Tschesche, H., Greene, L.J., Truscheit, E. (eds.) Bayer Symposium V, Proteinase Inhibitors, pp. 463–469. Springer Verlag, New York (1974)
- van Dongen, W.D., Heck, A.J.R.: Binding of selected carbohydrates to apo-concanavalin A studied by electrospray ionization mass spectrometry. *Analyst* **125**, 583–589 (2000)
- Clark, S.M., Konermann, L.: Determination of ligand-protein dissociation constants by electrospray mass spectrometry-based diffusion measurements. *Anal. Chem.* **76**, 7077–7083 (2004)
- Sun, J., Kitova, E.N., Klassen, J.S.: Method for stabilizing protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **79**, 416–425 (2007)
- Wang, W., Kitova, E.N., Klassen, J.S.: Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **75**, 4945–4955 (2003)
- Wang, W., Kitova, E.N., Sun, J., Klassen, J.S.: Blackbody infrared radiative dissociation of nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization: The nature of the noncovalent interactions. *J. Am. Soc. Mass Spectrom.* **16**, 1583–1594 (2005)
- Peschke, M., Verkerk, U.H., Kebarle, P.: Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method. *J. Am. Soc. Mass Spectrom.* **15**, 1424–1434 (2004)
- Daubenfeld, T., Bouin, A.-P., van der Rest, G.: A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 1239–1248 (2006)
- Sun, J., Kitova, E.N., Wang, W., Klassen, J.S.: Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **78**, 3010–3018 (2006)
- Touboul, D., Maillard, L., Grässlin, A., Moumne, R., Seitz, M., Robinson, J., Zenobi, R.: How to deal with weak interactions in noncovalent complexes analyzed by electrospray mass spectrometry: Cyclopeptidic inhibitors of the nuclear receptor coactivator 1-STAT6. *J. Am. Soc. Mass Spectrom.* **20**, 303–311 (2009)
- Schmidt, A., Bahr, U., Karas, M.: Influence of pressure in the first pumping stage on analyte desolvation and fragmentation in nano-ESI MS. *Anal. Chem.* **73**, 6040–6046 (2001)
- Tahallah, N., Pinkse, M., Maier, C.S., Heck, A.J.: The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid Commun. Mass Spectrom.* **15**, 596–601 (2001)
- Sanglier, S., Atmanene, C., Chevreux, G., Dorsselaer, A.V.: Non-denaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries. *Methods Mol. Biol.* **484**, 217–243 (2008)
- Bagal, D., Kitova, E.N., Liu, L., El-Hawiet, A., Schmier, P.D., Klassen, J.S.: Gas phase stabilization of noncovalent protein complexes formed by electrospray ionization. *Anal. Chem.* **81**, 7801–7806 (2009)
- Tjernberg, A., Carnö, S., Oliv, F., Benkestock, K., Edlund, P.-O., Griffiths, W.J., Hallén, D.: Determination of dissociation constants for protein-ligand complexes by electrospray ionization mass spectrometry. *Anal. Chem.* **76**, 4325–4331 (2004)
- Atmanene, C., Laux, A., Glattard, E., Muller, A., Schoentgen, F., Metz-Boutigue, M.-H., Aunis, D., Van Dorsselaer, A., Stefano, G. B., Sanglier-Cianfèrari, S., Goumon, Y.: Characterization of human and bovine phosphatidylethanolamine-binding protein (PEBP/RKIP) interactions with morphine and morphine-glucuronides determined by noncovalent mass spectrometry. *Med. Sci. Monit.* **15**, BR178–187 (2009)
- Bernier, I., Jollès, P.: Purification and characterization of a basic 23 kDa cytosolic protein from bovine brain. *Biochim. Biophys. (BBA)-Prot. Struct. Mol. Enzymol.* **790**, 174–181 (1984)
- Beaufour, M., Vaillant, A.T., Schoentgen, F., Cadene, M.: Mass Spectrometric Characterization of the Interaction Between PEBP/RKIP and Locostatin. *24th Annual Conference of the French Society for Mass Spectrometry*. Pau, France (2007)
- Gabant, G., Beaufour, M., Schoentgen, F., Cadene, M.: Towards a Dissection of the Molecular Mechanism of Binding of Locostatin to the

- PEBP/RKIP Protein, a Metastasis Suppressor, *25th Annual Conference of the French Society for Mass Spectrometry*, Grenoble, France (2008)
28. Limbach, P.A., Crain, P.F., McCloskey, J.A.: Molecular mass measurement of intact ribonucleic acids via electrospray ionization quadrupole mass spectrometry. *J. Am. Soc. Mass Spectrom.* **6**, 27–39 (1995)
  29. Jaquillard, L., Gabant, G., Schoentgen, F., Agrofoglio, L., Cadene, M.: Noncovalent Interactions Between the PEBP/RKIP Protein and Nucleotide analogs: *57th ASMS Conference on Mass Spectrometry and Allied Topics*. Philadelphia, PA (2009)
  30. Bucquoy, S., Jollès, P., Schoentgen, F.: Relationships between molecular interactions (nucleotides, lipids and proteins) and structural features of the bovine brain 21-kDa protein. *Eur. J. Biochem.* **225**, 1203–1210 (1994)
  31. Zhang, S., Van Pelt, C.K., Wilson, D.B.: Quantitative determination of noncovalent binding interactions using automated nanoelectrospray mass spectrometry. *Anal. Chem.* **75**, 3010–3018 (2003)
  32. Straume, M., Freire, E.: Two-dimensional differential scanning calorimetry: simultaneous resolution of intrinsic protein structural energetics and ligand binding interactions by global linkage analysis. *Anal. Biochem.* **203**, 259–268 (1992)
  33. Wiseman, T., Williston, S., Brandts, J.F., Lin, L.N.: Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137 (1989)
  34. Jones, C.L., Fish, F., Muccio, D.D.: Determination of RNase A/2'-cytidine monophosphate binding affinity and enthalpy by a global fit of thermal unfolding curves. *Anal. Biochem.* **302**, 184–190 (2002)
  35. Park, S., Rath, O., Beach, S., Xiang, X., Kelly, S.M., Luo, Z., Kolch, W., Yeung, K.C.: Regulation of RKIP binding to the N-region of the Raf-1 kinase. *FEBS Lett.* **580**, 6405–6412 (2006)
  36. Schindler, M., Assaf, Y., Sharon, N., Chipman, D.M.: Mechanism of lysozyme catalysis: Role of ground-state strain in subsite D in hen egg-white and human lysozymes. *Biochemistry* **16**, 423–431 (1977)
  37. Burbaum, J.J., Knowles, J.R.: Internal thermodynamics of enzymes determined by equilibrium quench: Values of  $K_{int}$  for enolase and creatine kinase. *Biochemistry* **28**, 9306–9317 (1989)
  38. Sun, N., Soya, N., Kitova, E.N., Klassen, J.S.: Nonspecific interactions between proteins and charged biomolecules in electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **21**, 472–481 (2010)
  39. El-Hawiet, A., Kitova, E.N., Liu, L., Klassen, J.S.: Quantifying labile protein-ligand interactions using electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **21**, 1893–1899 (2010)
  40. Robinson, C.V., Chung, E.W., Kragelund, B.B., Knudsen, J., Aplin, R.T., Poulsen, F.M., Dobson, C.M.: Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-CoA ligand binding and assembly. *J. Am. Chem. Soc.* **118**, 8646–8653 (1996)
  41. Wu, Q., Gao, J., Joseph-McCarthy, D., Sigal, G.B., Bruce, J.E., Whitesides, G.M., Smith, R.D.: Carbonic anhydrase-inhibitor binding: From solution to the gas phase. *J. Am. Chem. Soc.* **119**, 1157–1158 (1997)
  42. Cheetham, J.C., Artymiuk, P.J., Phillips, D.C.: Refinement of an enzyme complex with inhibitor bound at partial occupancy: Hen egg-white lysozyme and tri-*N*-acetylchitotriose at 1.75 Å resolution. *J. Mol. Biol.* **224**, 613–628 (1992)
  43. Jecklin, M.C., Touboul, D., Bovet, C., Wortmann, A., Zenobi, R.: Which electrospray-based ionization method best reflects protein-ligand interactions found in solution? A comparison of ESI, nanoESI, and ESSI for the determination of dissociation constants with mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 332–343 (2008)

# SUPPLEMENTARY MATERIALS

FOR

## IMPROVED ACCURACY OF LOW AFFINITY PROTEIN-LIGAND EQUILIBRIUM DISSOCIATION CONSTANTS DIRECTLY DETERMINED BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

Lucie Jaquillard<sup>1</sup>, Fabienne Saab<sup>1</sup>, Françoise Schoentgen<sup>2</sup>, Martine Cadene<sup>1</sup>

1. Centre de Biophysique Moléculaire, CNRS UPR 4301, affiliated with Université d'Orléans, rue Charles-Sadron, F-45071 Orléans Cedex 2, France.

2. Institut de Minéralogie et de Physique des Milieux Condensés, CNRS UMR 7590, Université de Paris 6, 75252 PARIS cedex 05, France

Table S1. Schemes and equations relating to the “no dissociation” case

	SOLUTION PHASE		GAS PHASE	
			INTERFACE	ANALYZER
	$P_1 + L_1 \xrightleftharpoons{K_D} PL_1 \xrightarrow{MG} P_F, L_F, PL_F$ <i>equilibrium 1</i> <i>non-equilibrium</i>		$PL_F \rightarrow P_{GPD} + L_{GPD} \rightarrow P_M, PL_M$ <i>non-equilibrium</i>	
Kinetic or equilibrium constant	$K_D = \frac{[P]_1[L]_1}{[PL]_1}$		$[PL]_F(t_{GPD}) = [PL]_F \cdot f_{sat}(t_{GPD})$	
Mass conservation	$[P]_0 = [P]_1 + [PL]_1$ $[L]_0 = [L]_1 + [PL]_1$	$[P]_{F\ tot} = [P]_F + [PL]_F$ $[L]_{F\ tot} = [L]_F + [PL]_F$	$[P]_{G\ tot} = [P]_G + [PL]_G$ $[L]_{G\ tot} = [L]_G + [PL]_G$	$[P]_{M\ tot} = [P]_M + [PL]_M$ $[L]_{M\ tot} = [L]_M + [PL]_M$
Particular conditions		$[P]_F = a[P]_1$ $[L]_F = b[L]_1$ $[PL]_F = a[PL]_1$	$[P]_G = [P]_{GPD} + [P]_F$ $[L]_G = [L]_{GPD} + [L]_F$ $[PL]_G = [PL]_F(t_{GPD})$	$[P]_{M\ tot} = [P]_{G\ tot}$ $[PL]_M = [PL]_G$ $f_{bound} = \frac{[PL]_M}{[PL]_M + [P]_M}$

*Mathematical development for the “partial dissociation” case*

In the “partial dissociation” case, the complex dissociates after MGF without reaching a new equilibrium. Between the MGF step and the transfer to the gas phase, the species are annotated “i” for this intermediary state. In this intermediary state, at every time t, the species concentrations change.

Table S2. Schemes and equations relating to the “partial dissociation” case

	SOLUTION PHASE			GAS PHASE	
				INTERFACE	ANALYZER
	$P_1 + L_1 \xrightleftharpoons{K_D} PL_1 \xrightarrow{\text{MGF}} P_F, L_F, PL_F \xrightarrow{\text{LPD}} P_i(t), L_i(t), PL_i(t) \xrightarrow{\text{non-equilibrium}} PL_{LPD} \rightarrow P_{GPD} + L_{GPD} \xrightarrow{\text{non-equilibrium}} P_M, PL_M$				Measurement
	<i>equilibrium 1</i>	<i>non-equilibrium</i>	<i>non-equilibrium</i>	<i>non-equilibrium</i>	
Kinetic or equilibrium constant	$K_D = \frac{[P]_1[L]_1}{[PL]_1}$			$[PL]_{LPD}(t_{GPD}) = [PL]_{LPD} \cdot f_{sat}(t_{GPD})$	
Mass conservation	$[P]_0 = [P]_1 + [PL]_1$ $[L]_0 = [L]_1 + [PL]_1$	$[P]_{F\ tot} = [P]_F + [PL]_F$ $[L]_{F\ tot} = [L]_F + [PL]_F$	$[P]_{i\ tot} = [P]_i(t) + [PL]_i(t)$ $[L]_{i\ tot} = [L]_i(t) + [PL]_i(t)$	$[P]_{G\ tot} = [P]_G + [PL]_G$ $[L]_{G\ tot} = [L]_G + [PL]_G$	$[P]_{M\ tot} = [P]_M + [PL]_M$ $[L]_{M\ tot} = [L]_M + [PL]_M$
Particular conditions		$[P]_F = a[P]_1$ $[L]_F = b[L]_1$ $[PL]_F = a[PL]_1$	$[PL]_i(t_{LPD}) = [PL]_{LPD}$ $[P]_i(t_{LPD}) = [P]_{LPD}$ $[L]_i(t_{LPD}) = [L]_{LPD}$	$[P]_G = [P]_{GPD} + [P]_{LPD}$ $[L]_G = [L]_{GPD} + [L]_{LPD}$ $[PL]_G = [PL]_{LPD}(t_{GPD})$	$[P]_{M\ tot} = [P]_{G\ tot}$ $[PL]_M = [PL]_G$ $f_{bound} = \frac{[PL]_M}{[PL]_M + [P]_M}$

The aim is to express  $f_{bound}$  as function of the known parameters  $a$ ,  $b$ ,  $[P]_0$ ,  $[L]_0$  and of the unknown factors  $K_D$  and  $f_{sat}$  to be determined by regression.

The equivalent of Equation 10 for the “partial dissociation” case is:

$$f_{bound} = \frac{[PL]_{LPD} \cdot f_{sat}}{a [P]_0} = \frac{[PL]_i(t_{LPD}) \cdot f_{sat}}{a [P]_0} \quad (S1)$$

First  $[PL]_i(t)$  can be expressed as a function of known parameters and of t.

By definition:

$$\frac{d[PL]_i(t)}{dt} = k_{on} \cdot [P]_i(t) \cdot [L]_i(t) - k_{off} \cdot [PL]_i(t) \quad (S2)$$

At every time t,

$$[P]_i(t) = [P]_{Ftot} - [PL]_i(t) \quad (S3)$$

$$[L]_i(t) = [L]_{Ftot} - [PL]_i(t) \quad (S4)$$

Inserting Equations S3 and S4 into Equation S2 gives:

$$\frac{d[PL]_i(t)}{dt} = k_{on} \cdot ([P]_{Ftot} - [PL]_i(t)) \cdot ([L]_{Ftot} - [PL]_i(t)) - k_{off} \cdot [PL]_i(t) \quad (S5)$$

which can be developed into:

$$\frac{d[PL]_i(t)}{dt} = k_{on} ([PL]_i(t))^2 - [PL]_i(t) (k_{on} [L]_{Ftot} + k_{on} [P]_{Ftot} + k_{off}) + k_{on} \cdot [P]_{Ftot} \cdot [L]_{Ftot} \quad (S6)$$

Note that Equation S6 can be described as the polynomial expression:

$$\frac{d[PL]_i(t)}{dt} = a' ([PL]_i(t))^2 + b' [PL]_i(t) + c' \quad (S7)$$

with the following constants:

$$a' = k_{on}$$

$$b' = -(k_{on} [L]_{Ftot} + k_{on} [P]_{Ftot} + k_{off})$$

$$c' = k_{on} \cdot [P]_{Ftot} \cdot [L]_{Ftot}$$

It can be shown that the discriminant  $\Delta = b'^2 - 4a'c'$  is always positive.

The integral of Equation S7 is:

$$\int_0^t dt = \int_0^t \frac{d[PL]_i(t)}{a' ([PL]_i(t))^2 + b' [PL]_i(t) + c'}$$

(S8)

which gives:

$$t = \left[ \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \ln \left| \frac{2a'[PL]_i(t) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[PL]_i(t) + b' + \sqrt{b'^2 - 4a'c'}} \right| \right]_0^t \quad (S9)$$

which develops as:

$$t = \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \left( \ln \left| \frac{2a'[\text{PL}]_i(t) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t) + b' + \sqrt{b'^2 - 4a'c'}} \right| - \ln \left| \frac{2a'[\text{PL}]_F + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \right| \right) \quad (\text{S10})$$

The expression dependant on  $[\text{PL}]_i(t)$  is isolated and an exponential is applied to both sides of the equality:

$$e^{\sqrt{b'^2 - 4a'c'} \cdot t + \ln \left| \frac{2a'[\text{PL}]_F + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \right|} = \left| \frac{2a'[\text{PL}]_i(t) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t) + b' + \sqrt{b'^2 - 4a'c'}} \right| \quad (\text{S11})$$

Equation S11 is simplified and both sides of the equality multiplied by  $(2a'[\text{PL}]_i(t) + b' + \sqrt{b'^2 - 4a'c'})$ :

$$\left( 2a'[\text{PL}]_i(t) + b' + \sqrt{b'^2 - 4a'c'} \right) \cdot \frac{-2a'[\text{PL}]_F - b' + \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \cdot e^{\sqrt{b'^2 - 4a'c'} \cdot t} = -2a'[\text{PL}]_i(t) - b' + \sqrt{b'^2 - 4a'c'} \quad (\text{S12})$$

Equation S12 is rearranged into:

$$\begin{aligned} [\text{PL}]_i(t) = & \\ & \left( -b' + \sqrt{b'^2 - 4a'c'} - (b' + \sqrt{b'^2 - 4a'c'}) * \frac{-2a'[\text{PL}]_F - b' + \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \cdot e^{\sqrt{b'^2 - 4a'c'} \cdot t} \right) / \\ & \left( 2a' + 2a' \frac{-2a'[\text{PL}]_F - b' + \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \cdot e^{\sqrt{b'^2 - 4a'c'} \cdot t} \right) \end{aligned} \quad (\text{S13})$$

Equations S1 and S13 combine into:

$$f_{\text{bound}} = \frac{\left( -b' + \sqrt{b'^2 - 4a'c'} - (b' + \sqrt{b'^2 - 4a'c'}) * \frac{-2a'[\text{PL}]_F - b' + \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \cdot e^{\sqrt{b'^2 - 4a'c'} \cdot t} \right) \cdot f_{\text{sat}}}{\left( 2a' + 2a' \frac{-2a'[\text{PL}]_F - b' + \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \cdot e^{\sqrt{b'^2 - 4a'c'} \cdot t} \right) a [\text{P}]_0} \quad (\text{S14})$$

with:

$$a' = k_{\text{on}}$$

$$b' = -(k_{\text{on}} [\text{L}]_{\text{Ftot}} + k_{\text{on}} [\text{P}]_{\text{Ftot}} + k_{\text{off}})$$

$$c' = k_{on} \cdot [P]_{Ftot} \cdot [L]_{Ftot}$$

and

$$[PL]_F = a \cdot \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4 [P]_0 [L]_0}}{2} \quad (\text{from Equations 12 and 15})$$

$$[L]_{Ftot} = (a - b) \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4 [P]_0 [L]_0}}{2} + b [L]_0$$

$$[P]_{Ftot} = a[P]_0$$

According to Equation S14,  $f_{bound}$  depends on four known parameters  $a$ ,  $b$ ,  $[P]_0$ , and  $[L]_0$ , and on five unknown parameters  $K_D$ ,  $k_{on}$ ,  $k_{off}$ ,  $t$ , and  $f_{sat}$ , which could be brought back to four unknown parameters based on the relation between  $K_D$ ,  $k_{on}$ , and  $k_{off}$  (Equation 6).

However, attempts at fitting this increased number of parameters by non-linear regression analysis would lead to a greater freedom for the regression to find compatible sets of parameter solutions, so as to render the result potentially meaningless.

### *Complement to "Guidelines to determine which model to apply"*

We have developed and presented in the main article an explicit model for system behavior in a new workflow, which offers the opportunity to determine the solution  $K_D$  for low affinity systems while addressing both aggregation and gas phase dissociation issues. For a given system, conscious choices can be made in adjusting the conditions for  $K_D$  determination to facilitate the choice and application of the model.

Each binding system can be described in kinetic terms by its equilibrium and kinetic constants ( $K_D$ ,  $k_{on}$ ,  $k_{off}$ ), and its behavior can be predicted as a function of the ligand concentration. The predicted behavior can be then used to assist in adjusting conditions.



*Generating a graph to simulate the association kinetics of the complex (Figure S1)*

This section develops an equation to simulate the progress of complex formation given  $K_D$ ,  $k_{on}$ ,  $k_{off}$  and initial concentrations and shows a simulation graph as an example of application. We define  $t_{x\%}$ , the time to reach  $x$  % equilibrium, as the time necessary for the complex concentration to reach  $x$  % of the equilibrium concentration (Equilibrium 1 in this work). For the sake of accuracy, we have developed this equation for a second-order reaction, even though some ligand concentrations in the experiment may meet pseudo-first order conditions.

When  $x = 100\%$ ,  $\frac{[PL]_i(t_{100\%})}{[PL]_1} = 1$ , which gives  $[PL]_i(t_{100\%}) = [PL]_1$

When  $x = 0\%$ ,  $\frac{[PL]_i(t_{0\%})}{[PL]_1} = 0$ , which gives  $[PL]_i(t_{0\%}) = 0$  (S15)

The time to reach  $x$  % equilibrium is the time for which this equality is true:

$$x \% = \frac{[PL]_i(t_{x\%})}{[PL]_1} \quad (S16)$$

Equation S16 is equivalent to:

$$[PL]_i(t_{x\%}) = x [PL]_1 \quad (S17)$$

with:

- $[PL]_i(t_{x\%})$ : the complex concentration at  $x\%$  of equilibration
- $[PL]_1$ : the complex concentration reached at equilibrium 1

In the above section *Mathematical development for the “partial dissociation” case*, we found the relation between  $t_{x\%}$  and  $[PL]_i(t_{x\%})$ :

$$t_{x\%} = \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \left( \ln \left| \frac{2a'[\text{PL}]_i(t_{x\%}) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t_{x\%}) + b' + \sqrt{b'^2 - 4a'c'}} \right| - \ln \left| \frac{2a'[\text{PL}]_i(t_{0\%}) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t_{0\%}) + b' + \sqrt{b'^2 - 4a'c'}} \right| \right) \quad (\text{S10})$$

Equations S15, S17 and S10 combine into:

$$t_{x\%} = \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \left( \ln \left| \frac{2a'x[\text{PL}]_1 + b' - \sqrt{b'^2 - 4a'c'}}{2a'x[\text{PL}]_1 + b' + \sqrt{b'^2 - 4a'c'}} \right| - \ln \left| \frac{b' - \sqrt{b'^2 - 4a'c'}}{b' + \sqrt{b'^2 - 4a'c'}} \right| \right) \quad (\text{S18})$$

with:

$$a' = k_{\text{on}}$$

$$b' = -(k_{\text{on}}[\text{L}]_0 + k_{\text{on}}[\text{P}]_0 + k_{\text{off}})$$

$$c' = k_{\text{on}} \cdot [\text{P}]_0 \cdot [\text{L}]_0$$

and

$$[\text{PL}]_1 = \frac{[\text{P}]_0 + [\text{L}]_0 + K_D - \sqrt{([\text{P}]_0 + [\text{L}]_0 + K_D)^2 - 4[\text{P}]_0[\text{L}]_0}}{2} \quad (12)$$

Equation S18 can be used to calculate the time to reach complete equilibrium.

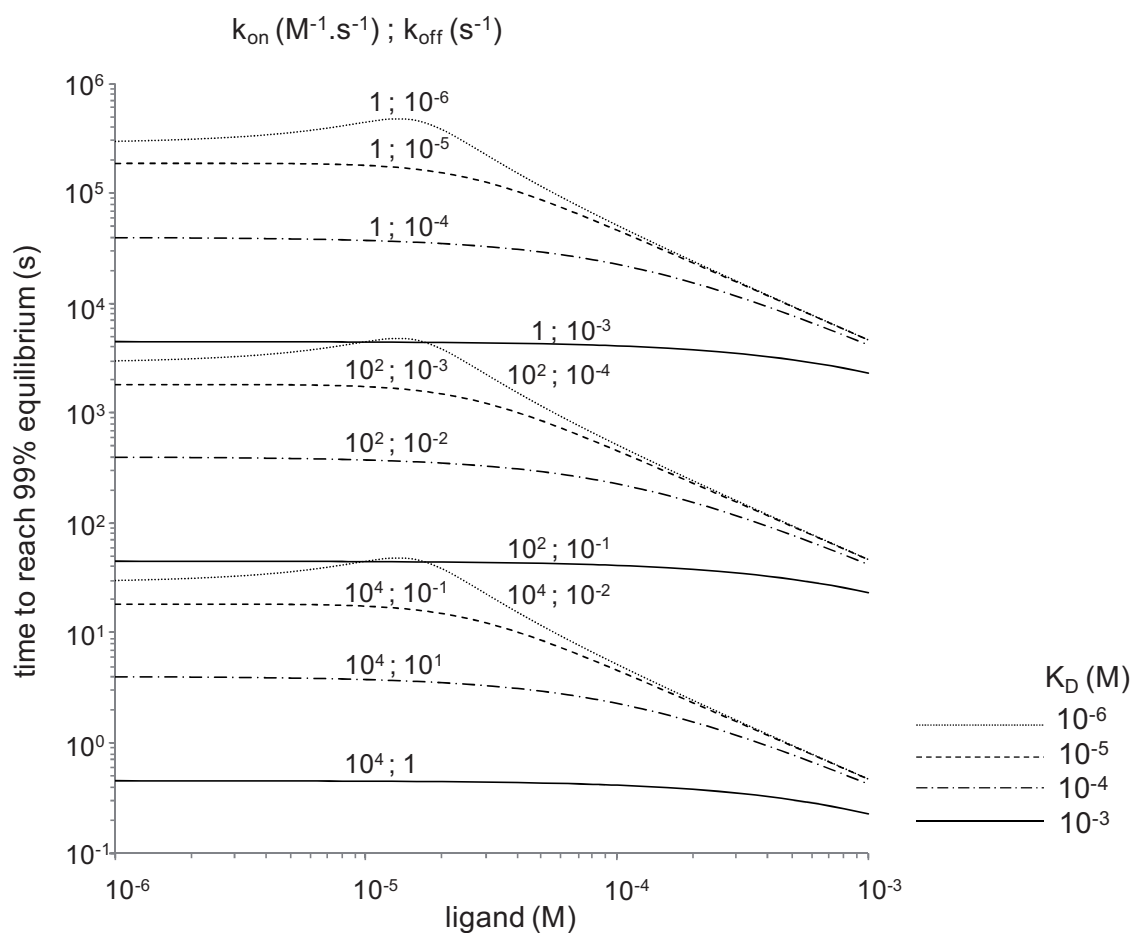


Figure S1. Effect of equilibrium and kinetic constants on association behavior. Theoretical curves were plotted for time to reach 99% equilibrium as a function of ligand concentration for different sets of  $K_D$ ,  $k_{on}$  and  $k_{off}$  values (dotted lines:  $K_D = 10^{-6}$  M; dashed lines:  $K_D = 10^{-5}$  M; dash-dot lines:  $K_D = 10^{-4}$  M; solid lines:  $K_D = 10^{-3}$  M). This numerical application is based on  $[P]_0 = 15 \mu\text{M}$ .

As an example, the theoretical graph in Figure S1 was built by calculating  $t_{99\%}$ , the time to reach 99% equilibrium for complexes with low affinity  $K_D$  values of  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ , or  $10^{-3}$  M, association rate constants  $k_{on}$  of 1,  $10^2$ , or  $10^4$   $\text{M}^{-1}\cdot\text{s}^{-1}$  and thirty-one ligand concentrations between  $10^{-6}$  and  $10^{-3}$  M. Complexes can be grouped by their  $K_D$  as shown with dotted, solid or dashed lines. For  $k_{on}$  values of  $10^6$  and  $10^8$   $\text{M}^{-1}\cdot\text{s}^{-1}$ , association is even faster and these curves are not represented on the graph. No assumption regarding the order of the reaction was made. The bump observed with some of the curves results from second-order conditions.

*Generating of a graph simulating dissociation behavior of a complex after MGF (Figure S2)*

This section develops an equation to simulate the transition of a complex to a new equilibrium after the MGF step, given  $K_D$ ,  $k_{on}$ ,  $k_{off}$  and initial concentrations. Simulation graphs used to define the “partial dissociation” zone are shown.

The time to reach  $x$  % equilibrium after MGF ( $t_{x\%}$ ) is the time necessary for the complex concentration to undergo  $x$  % of the decrease necessary to reach the complex concentration at the new equilibrium point (corresponding to Equilibrium 2).

$$\text{When } x = 100\%, [\text{PL}]_F - [\text{PL}]_i(t_{100\%}) = [\text{PL}]_F - [\text{PL}]_2$$

$$\text{When } x = 0\%, [\text{PL}]_F - [\text{PL}]_i(t_{0\%}) = 0, \text{ which gives } [\text{PL}]_F = [\text{PL}]_i(t_{0\%}) \quad (\text{S19})$$

The time to reach  $x$  % equilibrium after MGF is the time for which this equality is true:

$$x \% = \frac{[\text{PL}]_F - [\text{PL}]_i(t_{x\%})}{[\text{PL}]_F - [\text{PL}]_2} \quad (\text{S20})$$

Equation S20 is equivalent to:

$$[\text{PL}]_i(t_{x\%}) = (1 - x)[\text{PL}]_F + x [\text{PL}]_2 \quad (\text{S21})$$

with:

- $[\text{PL}]_F$ : the complex concentration just after MGF. If  $a$  is the yield in protein after MGF,  $[\text{PL}]_F = a[\text{PL}]_1$
- $[\text{PL}]_i(t_{x\%})$ : the complex concentration at  $x\%$  of re-equilibration after MGF
- $[\text{PL}]_2$ : the complex concentration reached at equilibrium 2

Again, we can use Equation S10 relating  $t_{x\%}$  to  $[\text{PL}]_i(t_{x\%})$ :

$$t_x \% = \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \left( \ln \left| \frac{2a'[\text{PL}]_i(t_x\%) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t_x\%) + b' + \sqrt{b'^2 - 4a'c'}} \right| - \ln \left| \frac{2a'[\text{PL}]_i(t_0\%) + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_i(t_0\%) + b' + \sqrt{b'^2 - 4a'c'}} \right| \right) \quad (\text{S10})$$

Equations S19, S21 and S10 combine into:

$$t_x \% = \frac{1}{\sqrt{b'^2 - 4a'c'}} \times \left( \ln \left| \frac{2a'((1-x)[\text{PL}]_F + x[\text{PL}]_2) + b' - \sqrt{b'^2 - 4a'c'}}{2a'((1-x)[\text{PL}]_F + x[\text{PL}]_2) + b' + \sqrt{b'^2 - 4a'c'}} \right| - \ln \left| \frac{2a'[\text{PL}]_F + b' - \sqrt{b'^2 - 4a'c'}}{2a'[\text{PL}]_F + b' + \sqrt{b'^2 - 4a'c'}} \right| \right) \quad (\text{S22})$$

with:

$$a' = k_{\text{on}}$$

$$b' = -(k_{\text{on}} [\text{L}]_{\text{Ftot}} + k_{\text{on}} [\text{P}]_{\text{Ftot}} + k_{\text{off}})$$

$$c' = k_{\text{on}} \cdot [\text{P}]_{\text{Ftot}} \cdot [\text{L}]_{\text{Ftot}}$$

and

$$[\text{PL}]_F = a \cdot [\text{PL}]_1$$

$$[\text{PL}]_1 = \frac{[\text{P}]_0 + [\text{L}]_0 + K_D - \sqrt{([\text{P}]_0 + [\text{L}]_0 + K_D)^2 - 4[\text{P}]_0[\text{L}]_0}}{2} \quad (12)$$

$$[\text{PL}]_2 = \frac{a[\text{P}]_0 + b[\text{L}]_0 + (a-b)[\text{PL}]_1 + K_D - \sqrt{(a[\text{P}]_0 - b[\text{L}]_0 - (a-b)[\text{PL}]_1 - K_D)^2 + 4a[\text{P}]_0K_D}}{2} \quad (11)$$

$$[\text{L}]_{\text{Ftot}} = (a - b) [\text{PL}]_1 + b [\text{L}]_0$$

$$[\text{P}]_{\text{Ftot}} = a[\text{P}]_0$$

Equation S22 can be used to calculate  $t_{95\%}$ . and  $t_{5\%}$ .

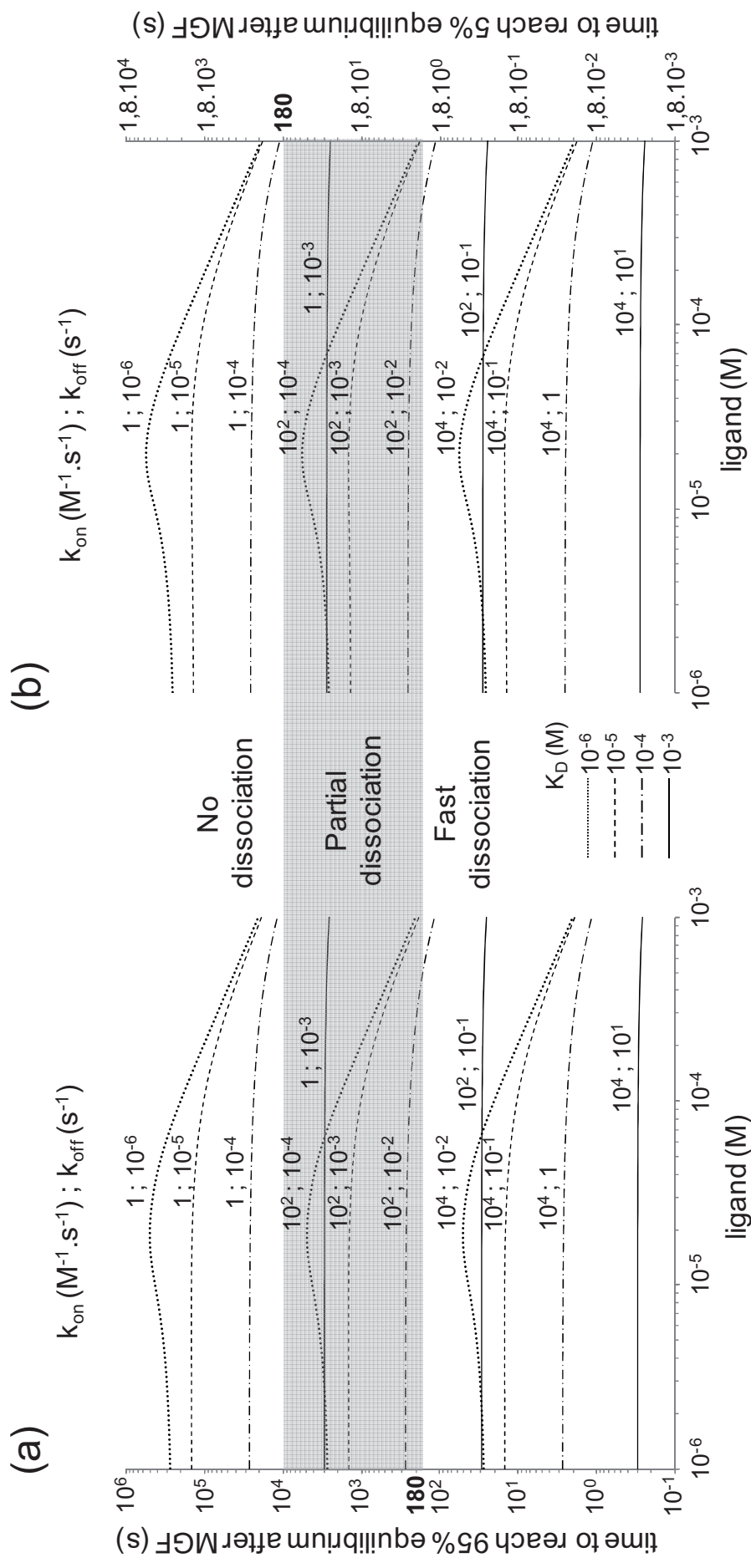


Figure S2. Effect of equilibrium and kinetic constants on dissociation behavior. Theoretical curves were plotted for (a) time to reach 95% equilibrium after MGF or (b) time to reach 5% equilibrium after MGF as a function of ligand concentration for different sets of  $K_D$ ,  $k_{\text{on}}$  and  $k_{\text{off}}$  values (dotted lines:  $K_D = 10^{-6} \text{ M}$ ; dashed lines:  $K_D = 10^{-5} \text{ M}$ ; dash-dot lines:  $K_D = 10^{-4} \text{ M}$ ; solid lines:  $K_D = 10^{-3} \text{ M}$ ). The greyed area (partial dissociation) corresponds to dissociation between 5 and 95% in the analysis dead-time (in this example, 180 s). The upper and lower areas correspond to less than 5% and over 95% dissociation, respectively. This numerical application is based on  $[P]_0 = 15 \text{ }\mu\text{M}$ ,  $a = 0.7$  and  $b = 0.15$ .

The theoretical graphs in Figure S2 were built by calculating the time to reach 95% or 5% equilibrium after MGF for complexes with low affinity  $K_D$  values of  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$  or  $10^{-3}$  M, association rate constants  $k_{on}$  of 1,  $10^2$ ,  $10^4$ ,  $10^6$  or  $10^8$   $M^{-1}.s^{-1}$  and thirty-one ligand concentrations between  $10^{-6}$  and  $10^{-3}$  M. Complexes can be grouped by their  $K_D$  as shown with dotted, solid or dashed lines. For the sake of simplification, the curves corresponding to  $k_{on}$  of  $10^6$  and  $10^8$   $M^{-1}.s^{-1}$  are not represented on these graphs. The amount of time required to reach 95% and 5% are represented on Figure S2a and S2b, respectively. Again, we made no assumption regarding the order of the reaction and the bump observed with some of the curves results from second-order conditions.

The dead-time of analysis is the elapsed time between the withdrawal of a reaction aliquot and the measurement in the instrument. If we consider that a system is fast dissociating when over 95% complex dissociation has occurred in the analysis dead-time (in this case 180 s), then all systems located in the white area below the grayed area correspond to the "fast dissociation" model. Conversely, considering that dissociation below 5% within the analysis dead-time is negligible, the "no dissociation" model applies when the system is described by a point located in the upper white area. The grayed area corresponds to partial dissociation conditions.

When  $k_{off}$  is sufficiently high ( $k_{off} \geq 10^{-1} s^{-1}$ ), the complex re-equilibration occurs within 180 s, regardless of  $k_{on}$ . These systems all fall into the lower white area and match the "fast dissociation" model. For complexes with slower dissociation kinetics ( $k_{off} < 10^{-1} s^{-1}$ ), the degree of dissociation will depend on the equilibrium dissociation constant and the ligand concentration.

When  $k_{on}$  and  $K_D$  are sufficiently low ( $k_{on} \leq 1 \text{ M}^{-1} \cdot \text{s}^{-1}$  and  $K_D \leq 10^{-4} \text{ M}$ ), the complex will not dissociate within 180 s. For complexes with  $1 \leq k_{on} \leq 10^2 \text{ s}^{-1}$ , the degree of dissociation will depend on the equilibrium dissociation constant and the ligand concentration.

As explained in Results and Discussion, it is possible to increase the time between the MGF and MS analysis to bring a partially dissociating system into the “fast dissociation” case. Alternatively, it is possible to have the whole concentration range fall into the “no dissociation” or the “fast dissociation” zone by playing on the range of ligand concentrations used. The flat-looking curves in Figure S2 are the exception, for which one can only play on time to reach full dissociation.

If the 1:1 protein:ligand stoichiometric point falls into the “no dissociation” case and the lowest and/or highest ligand concentrations into the “partial dissociation” case, the range of ligand concentrations can be narrowed and/or displaced so that they all fall into the “no dissociation” zone. The “no dissociation” model can then be applied to this system.

Similarly, if the 1:1 protein:ligand point falls into the “partial dissociation” case and the lowest and/or highest ligand concentrations into the “fast dissociation” case, one should reduce the concentration range around the lowest or the highest ligand concentration, excluding the 1:1 protein:ligand point, to bring back the system into the “fast dissociation” zone. The “fast dissociation” model can then be applied to this system.

By applying one of these strategies, i.e. playing on the analysis dead-time or on the ligand concentration, all the studied systems can fall either into the “fast dissociation” or into the “no dissociation” case, and thus allowing for the  $K_D$  determination to be easily performed.



Table S3. Relation between charge state and  $K_D$

P-L complex	Experimental conditions	Charge state	$K_D$ measurement with GPD correction (Equation 3 or 13)		
			$K_D$ ( $\mu M$ )	$f_{sat}$	$R^2$
RNase-CTP	10 mM $NH_4OAc$ , pH 6.8; 25°C	Weighted mean	$1.6 \pm 0.24$	$0.47 \pm 0.07$	0.971
		8+	$0.62 \pm 0.31$	$0.37 \pm 0.07$	0.891
		7+ *	$2.1 \pm 0.25$	$0.51 \pm 0.07$	0.976
PEBP-P3P	20 mM ABC, pH 7.4; 25°C	Weighted mean	$11 \pm 0.74$	$0.26 \pm 0.02$	0.995
		10+	$1.7 \pm 1.6$	$0.60 \pm 0.16$	0.818
		9+ *	$24 \pm 1.0$	$0.31 \pm 0.03$	0.996
HEWL-NAG3	50 mM $NH_4OAc$ , pH 6.6; 37°C	8+	$21 \pm 1.8$	$0.14 \pm 0.03$	0.987
		Weighted mean	$20 \pm 1.2$	$1.02 \pm 0.08$	0.993
		8+ *	$19 \pm 1.1$	$1.01 \pm 0.07$	0.997
(CK) <sub>2</sub> -2ADP	10 mM ABC, pH 7.9; 37°C <i>ESI-UHR Q-TOF</i>	7+	$38 \pm 5.1$	$1.03 \pm 0.31$	0.973
		Weighted mean	$4.5 \pm 0.54$	$0.61 \pm 0.17$	0.964
		21+	$1.7 \pm 0.38$	$0.57 \pm 0.14$	0.947
		20+ *	$5.2 \pm 0.55$	$0.68 \pm 0.19$	0.968
PEBP-GTP	20 mM ABC, pH 8.4; 37°C	19+ *	$9.7 \pm 0.85$	$0.73 \pm 0.27$	0.965
		18+	$2.6 \pm 0.43$	$0.47 \pm 0.12$	0.957
		Weighted mean	$36 \pm 0.62$	$0.42 \pm 0.03$	0.999
		10+	$8.7 \pm 1.4$	$0.41 \pm 0.07$	0.969
PEBP-GTP	20 mM ABC, pH 8.4; 37°C + MGF	9+ *	$26 \pm 1.9$	$0.37 \pm 0.07$	0.984
		8+	$36 \pm 1.9$	$0.34 \pm 0.07$	0.911
		Weighted mean	$55 \pm 5.0$	$0.40 \pm 0.07$	0.970
		10+	$18 \pm 4.0$	$0.42 \pm 0.08$	0.899
PEBP-GTP	20 mM ABC, pH 8.4; 37°C + MGF, <i>ESI-UHR Q-TOF</i>	9+ *	$60 \pm 5.1$	$0.43 \pm 0.08$	0.972
		8+	$170 \pm 7.3$	$0.65 \pm 0.15$	0.989
		Weighted mean	$40 \pm 1.0$	$0.52 \pm 0.05$	0.993
		10+ *	$29 \pm 0.65$	$0.46 \pm 0.03$	0.995
PEBP-FMN	20 mM ABC, pH 8.4; 37°C + MGF	9+ *	$42 \pm 0.90$	$0.55 \pm 0.04$	0.995
		8+	$64 \pm 1.8$	$0.66 \pm 0.10$	0.990
		Weighted mean	$35 \pm 3.1$	$0.25 \pm 0.05$	0.967
		10+	$27 \pm 3.2$	$0.41 \pm 0.09$	0.948
		9+ *	$36 \pm 5.4$	$0.24 \pm 0.09$	0.911
		8+	$52 \pm 12$	$0.16 \pm 0.12$	0.795

\* major peak

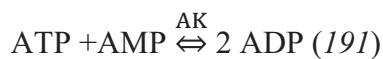
# Annexes à la publication:

## I. Validation du modèle

Dans le cadre des expériences destinées à valider notre modèle, trois systèmes non décrits dans la publication ont été étudiés. L'interaction  $\beta$ -lactoglobuline-piperine n'étant pas conservée lors de l'analyse MS, aucune donnée n'est décrite ici.

### 1. La myokinase ou adénylate kinase (AK) en complexe avec deux molécules d'ADP.

L'AK, en présence de magnésium, catalyse la réaction :



D'après la littérature, l'AK comporte un site de liaison pour l'AMP et un site pour l'ATP, et l'ADP peut se loger dans les deux sites (191). D'après les résultats de Zenobi et ses collaborateurs, le complexe AK-2ADP peut-être observé intact par MS après une incubation rapide (192).

Avant d'appliquer notre modèle à ce système, la protéine seule et un mélange AK /ADP (1 : 5) ont été analysés en conditions natives par ESI-IT (Figure 77).

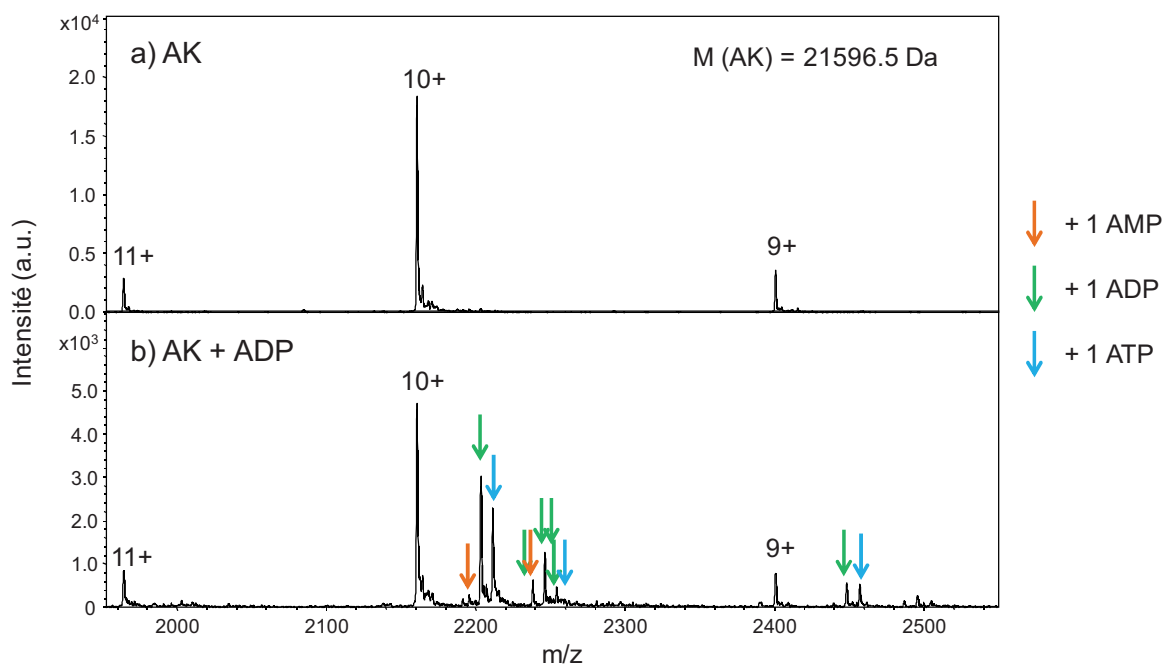


Figure 77: Spectres ESI-IT de 10  $\mu\text{M}$  d'adénylate kinase (AK) dans  $\text{NH}_4\text{OAc}$  50 mM, pH 6.6, a) seule, b) avec 50  $\mu\text{M}$  d'ADP. Les flèches orange, vertes et bleues indiquent respectivement la liaison d'un ligand AMP, ADP et ATP.

La protéine apparaissant sous trois états de charge, avec une distribution étroite, la conformation native de la protéine semble bien préservée. La masse moléculaire mesurée de la protéine est de 21 596.5 Da. Or, la masse théorique de la protéine sans modification est de 21 683 Da. La masse mesurée correspond à la protéine sans la méthionine initiale et avec une acétylation. L'ajout d'ADP pur entraîne l'apparition de plusieurs pics correspondant aux complexes AK-AMP, AK-ADP, AK-ATP, AK-AMP-ADP, AK-2ADP, AK-ADP-ATP. Le pic correspondant au complexe AK-AMP est très peu intense. Les deux pics les plus intenses sont ceux correspondant aux complexes AK-ADP et AK-ATP.

Contrairement aux résultats de MS publiés par Jecklin (192), malgré une analyse rapide de l'échantillon, l'AK semble avoir catalysé la formation d'ATP et d'AMP à partir d'ADP. L'AK que nous avons analysée est la même que celle décrite dans la section matériel de la publication de Jecklin : l'adénylate kinase de muscle de poulet achetée chez Sigma (192). Nous n'observons pourtant pas la même masse. Jecklin et ses collaborateurs observent un composé à 24.4 kDa alors que la masse théorique de l'AK est de 21.7 kDa. Le composé analysé dans cette publication ne semble donc pas être de l'adénylate kinase de muscle de poulet. Ceci expliquerait aussi l'absence de réaction enzymatique observable pour cette protéine à 24.4 kDa.

D'après nos résultats, la catalyse de l'ADP semble avoir lieu sans ajout de magnésium. D'après la littérature, le magnésium est pourtant indispensable à l'activité phospho-transférase de l'AK (193, 194).

Afin de vérifier que l'AMP et l'ATP ont été formés par l'ajout d'AK et qu'ils ne proviennent pas de l'échantillon d'ADP lui-même, l'ADP a été analysé par ESI-IT, en absence et en présence d'AK (Figure 78).

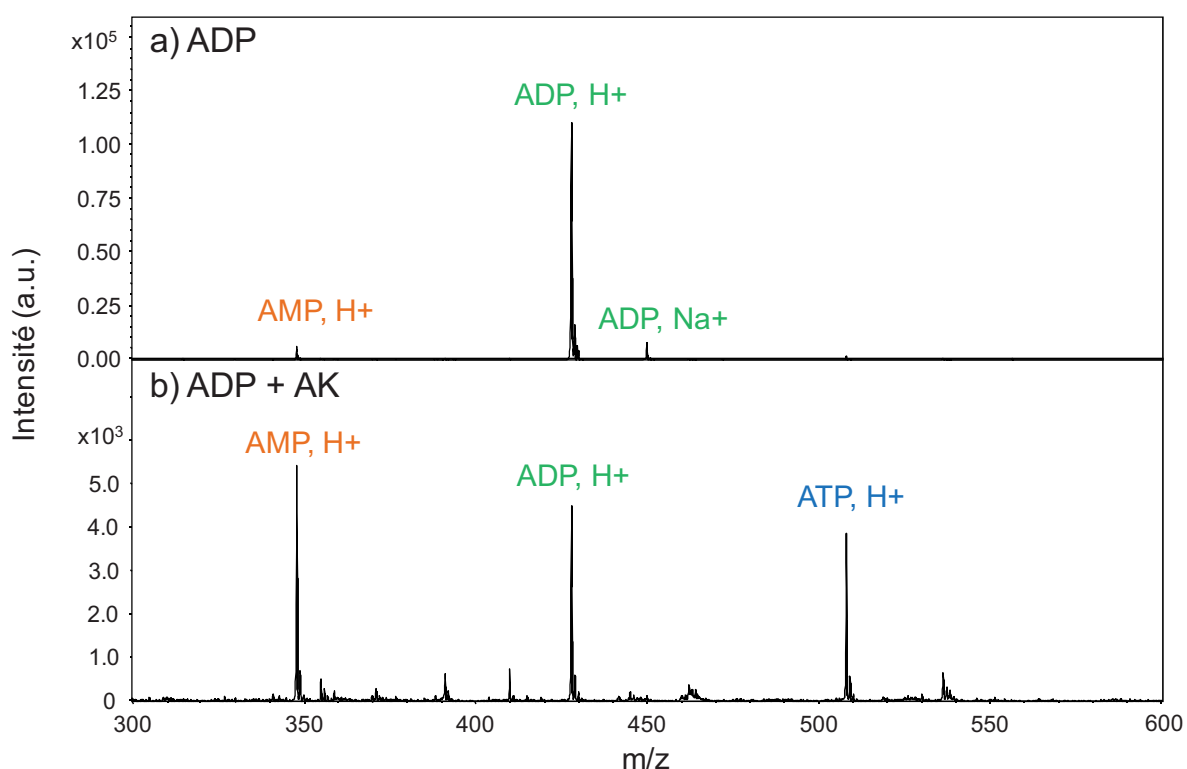


Figure 78 : Spectres ESI-IT de, a) 30  $\mu\text{M}$  d'ADP, b) 50  $\mu\text{M}$  d'ADP avec 10  $\mu\text{M}$  d'adénylate kinase (AK), dans  $\text{NH}_4\text{OAc}$  50 mM, pH 6.6.

L'analyse de l'ADP seul donne un pic majoritaire correspondant à l'ADP protoné, et deux pics très minoritaires correspondant à un adduit de sodium sur l'ADP et une molécule d'AMP protonée. L'AMP observé correspond vraisemblablement à une dégradation chimique en solution ou à un contaminant dans le stock d'ADP. En présence d'AK, un pic correspondant à de l'ATP protoné apparaît et l'intensité relative du pic d'AMP protoné par rapport à celle de l'ADP protoné augmente fortement.

Ces spectres confirment que l'AK entraîne la formation d'AMP et d'ATP. Le  $K_D$  du système AK-ADP ne pouvant être déterminé à cause de cette réaction enzymatique, ce système n'a pas pu être utilisé pour valider notre modèle. Il serait intéressant de considérer si cette méthode peut être utilisée à des fins de déterminations des constantes de Michaelis d'un substrat.

## 2. La concanavoline A avec le méthyl- $\alpha$ -D-mannopyranoside:

La concanavoline A a été observée à pH 5 sous forme dimérique (MW = 51.2 kDa) et en complexe avec deux atomes de manganèse ( $2 * 54.9$  Da), sous trois états de charge: 17+, 16+ et 15+. La masse du méthyl- $\alpha$ -D-mannopyranoside est de 194 Da. Des micro gel filtrations ont été réalisées pour éliminer le manganèse mais sans succès. La présence de manganèse élargissant les pics de la concanavoline A, il n'était pas possible de voir précisément la liaison du méthyl- $\alpha$ -D-mannopyranoside sur la protéine. De plus, le produit concanavoline A n'était pas pur. Les deux peptides précurseurs étaient encore visibles. Ce système n'a donc pas été utilisé pour valider notre modèle.

## II. Effet de la force ionique sur l'agrégation

D'après la littérature, l'agrégation de biomolécules chargées est un processus uniquement statistique (162, 169). Cependant, la stabilité des complexes non-spécifiques est, elle, sensible à la nature des ions protéiques (162). La stabilité des complexes non-spécifiques, comme celle des complexes spécifiques, peut être renforcée par l'ajout d'imidazole (162, 171, 172). A l'inverse, il serait intéressant de diminuer la stabilité des complexes non-spécifiques pour éliminer les problèmes d'agrégation en phase gazeuse. L'augmentation de la force ionique pourrait peut-être diminuer la stabilité des complexes non-spécifiques et ainsi les éliminer de l'échantillon gazeux. L'effet de l'augmentation de la force ionique a donc été testé sur l'agrégation du GTP sur hPEBP1 (Figure 79).

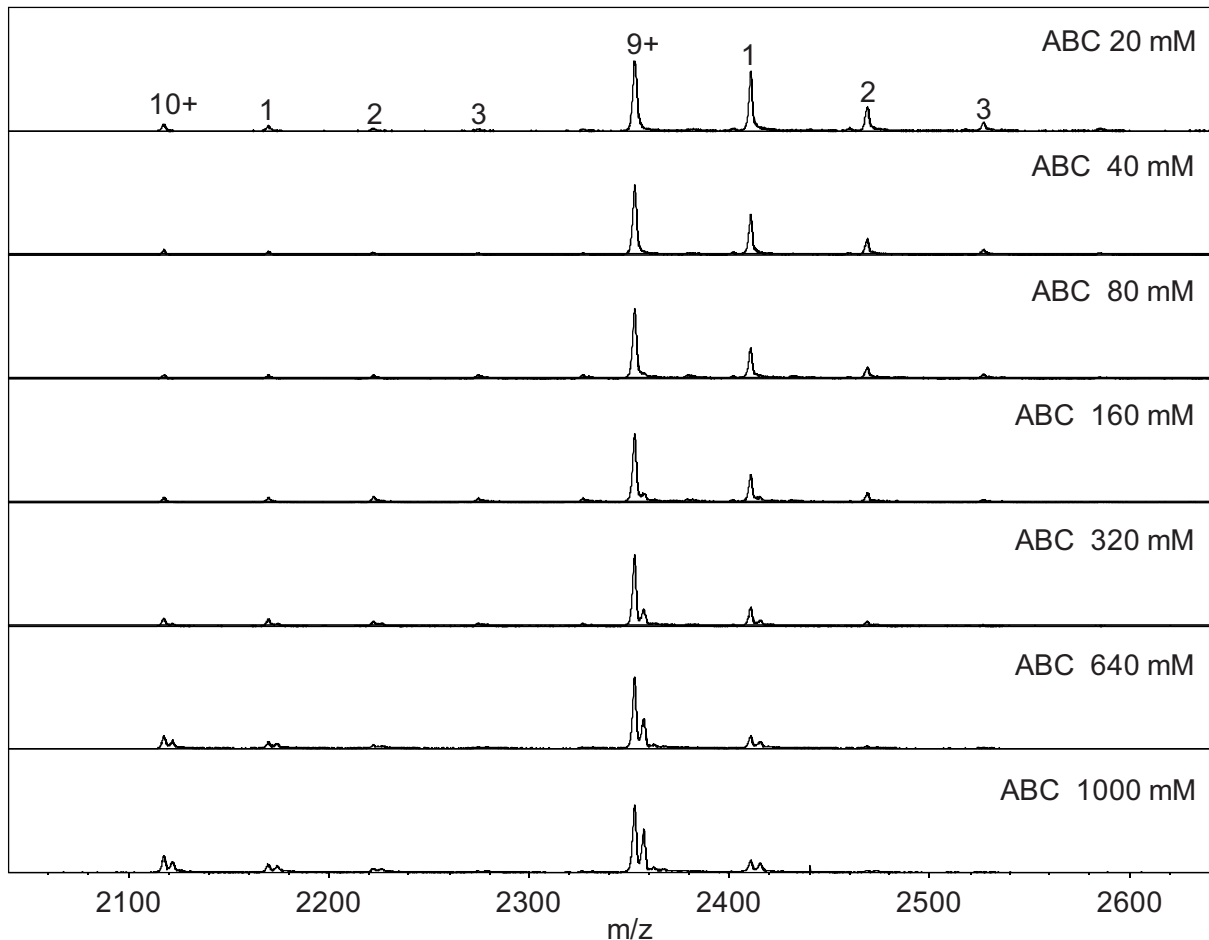


Figure 79: Spectres ESI-IT de la  $^{15}\text{N}$  PEBP à  $12,6 \mu\text{M}$  incubée 8 minutes à  $37^\circ\text{C}$  avec  $126 \mu\text{M}$  de GTP dans 20, 40, 80, 160, 320, 640 ou 1000 mM d'ABC à pH 7.4. Seuls les états de charge 9+ et 10+ sont montrés. La stœchiométrie apparente est indiquée sur les pics.

Lors de l'augmentation de la force ionique, une diminution de la stœchiométrie apparente est observée, ce qui correspond au moins à une diminution de l'agrégation. Ainsi même si l'agrégation de biomolécules chargées est un processus purement statistique, donc ne dépendant pas de la nature des biomolécules, ce processus est dépendant de la force ionique.

Il est possible que les complexes spécifiques soient également déstabilisés par l'augmentation de la force ionique. Ceci sera approfondi dans le chapitre IV.

### III. Effet de la tension d'accélération sur $K_D$ et $f_{\text{sat}}$

Les résultats de criblage PEBP-analogues de nucléotides font l'objet du chapitre IV. Cependant, nous présentons ici un résultat sur ce type de complexe qui paraît intéressant du point de vue de la méthodologie développée et de la possibilité de disséquer le paramètre  $f_{\text{sat}}$ .

Des  $K_D$  ont été déterminés pour 7 systèmes dans des conditions dissociantes, avec une tension d'accélération élevée ( $V_c = 345$  V) ou dans des conditions peu dissociantes, avec une tension d'accélération standard ( $V_c = 220$  V) (Tableau 5).

Tableau 5:  $K_D$  déterminés par des mesures de MS avec deux tensions d'accélération différentes :  $V_C = 220$  V, conditions *a priori* non-dissociantes ou  $V_C = 345$  V, conditions *a priori* dissociantes

complexe P-L	Conditions expérimentales	Tension d'accélération ( $V_C$ )	Mesure de $K_D$ en corrigeant la GPD (Equation 3 ou 13)	
			$K_D$ ( $\mu M$ )	$f_{sat}$
RNase-CTP	10 mM $NH_4OAc$ , pH 6.8; 25°C	220	$1.6 \pm 1.3$	$0.47 \pm 0.07$
		345	$1.3 \pm 1.6$	$0.51 \pm 0.10$
PEBP-FMN	20 mM ABC, pH 7.4; 37°C	220	$20 \pm 4.4$	$0.55 \pm 0.06$
		345	$30 \pm 11$	$0.65 \pm 0.13$
PEBP-BLA-C-051	20 mM ABC, pH 7.4; 37°C	220	$1.4 \pm 3.6$	$0.18 \pm 0.07$
		345	$1.5 \pm 3.7$	$0.11 \pm 0.04$
PEBP-BLA-C-054	20 mM ABC, pH 7.4; 37°C	220	$2.8 \pm 2.3$	$0.25 \pm 0.04$
		345	$2.5 \pm 1.6$	$0.17 \pm 0.02$
PEBP-UMP	20 mM ABC, pH 7.4; 37°C	220	$23 \pm 8$	$0.34 \pm 0.06$
		345	$14 \pm 7$	$0.18 \pm 0.04$
PEBP-BLA-C-055	20 mM ABC, pH 7.4; 37°C	220	$51 \pm 15$	$0.33 \pm 0.07$
		345	$43 \pm 12$	$0.24 \pm 0.03$
PEBP-BLA-C-056	20 mM ABC, pH 7.4; 37°C	220	$52 \pm 6$	$0.33 \pm 0.02$
		345	$64 \pm 21$	$0.22 \pm 0.04$

Ces expériences permettent de tester l'influence d'un paramètre instrumental, la tension d'accélération, sur la dissociation en phase gazeuse de complexes.

Les résultats montrent que les valeurs de  $K_D$  déterminées avec une tension d'accélération de 220 ou de 345 V sont très proches. Au contraire, le paramètre  $f_{sat}$  varie entre les deux conditions, il absorbe ainsi les variations instrumentales sans que cela n'impacte la valeur de  $K_D$ . La valeur de  $f_{sat}$  tend à diminuer pour 5 des complexes et augmente légèrement pour les complexes RNase-CTP et PEBP-FMN. Pour ces deux derniers systèmes, la dissociation en

phase gazeuse semble moins sensible au réglage de la tension d'accélération. Il serait intéressant de savoir si la sensibilité de  $f_{\text{sat}}$  à la tension d'accélération est corrélée avec le type de liaison protéine-nucléotide.

## IV. Recommandation

Pour s'assurer que le paramètre  $f_{\text{sat}}$  ne reflète que le phénomène de dissociation en phase gazeuse, il est important de titrer le site actif de la protéine d'intérêt en conditions stœchiométriques.





# Chapitre IV

Caractérisation du site nucléotidique de  
hPEBP1/RKIP

#### *Résumé du chapitre IV :*

*Pour caractériser le site nucléotidique de hPEBP1 et contribuer à la découverte de molécules anti-métastases, une recherche de ligand par criblage a été réalisée. L'équipe de Luigi Agrofoglio a réalisé la synthèse raisonnée des analogues de nucléotides en s'appuyant sur la structure des meilleurs nucléotides du criblage de Bucquoy et al. : GTP et FMN (73). Le criblage rapide est effectué par une mesure par MS d'une fraction de liaison avec hPEBP1. Les criblages d'analogues nucléotidiques et de nucléotides de référence mettent en évidence des critères structuraux indispensables pour qu'une molécule se lie à hPEBP1 : la présence d'un groupement chargé ou type NO<sub>2</sub>, d'une structure apparentée à une base azotée et d'un cycle. Des déterminations de K<sub>D</sub> montrent que les fractions de liaison seules ne sont pas représentatives des affinités relatives des ligands pour la protéine.*

*Le criblage par K<sub>D</sub> permet de classer les ligands par affinité pour hPEBP1. Les molécules analogues de FMN ou d'uracile sont plus affines que les analogues de GTP. Les résultats de criblage par l'affinité et leur corrélation avec le logP des ligands calculé montrent qu'une part importante de l'affinité est liée au caractère hydrophobe du ligand.*

*D'après une expérience de compétition, les sites d'interaction du GTP et de la locostatine sont indépendants. Les études de RMN faites par Laurette Tavel, Christian Damblon et Laurence Jouvensal, montrent que le GTP et le FMN se lient, entre autre, dans la poche à anion de hPEBP1. Les résultats de la MS combinés à ceux de la RMN prouvent que contrairement à rPEBP1, la PEBP1 humaine est capable de lier le GTP, et mettent en évidence une influence de la force ionique sur l'affinité du GTP et du FMN pour hPEBP1. Ces résultats sont en cours de publication.*

*Les meilleurs ligands issus du criblage par MS ont été testés en biologie cellulaire, sous forme de pro-drogues, pour leur capacité à inhiber la migration et l'invasion et pour leur cytotoxicité. Le crible sur la migration a permis de sélectionner les molécules à cribler dans le test de l'invasion. Il en ressort plusieurs molécules dont l'intérêt est de montrer une activité inhibitrice de l'invasion sur les lignées tumorales, tout en étant peu toxiques pour les cellules saines.*

# I. Recherche de ligands nucléotidiques par criblage

## 1. Principe de la synthèse raisonnée d'analogues nucléotidiques

Des molécules qui miment la structure des nucléotides FMN ou GTP ont été synthétisées à l'ICOA (Institut de Chimie Organique et Analytique) par l'équipe de Luigi Agrofoglio (Vincent Roy, Aurélien Montagu, Zhili Zhang et Ozkan Sari).

La synthèse raisonnée est basée sur le découpage du FMN ou du GTP en motifs structuraux caractéristiques (Figure 80). Elle consiste ensuite à réaliser différentes associations de ces motifs structuraux et à moduler les motifs. Cela permet ainsi de tester le rôle de ces motifs structuraux dans la liaison à hPEBP1.

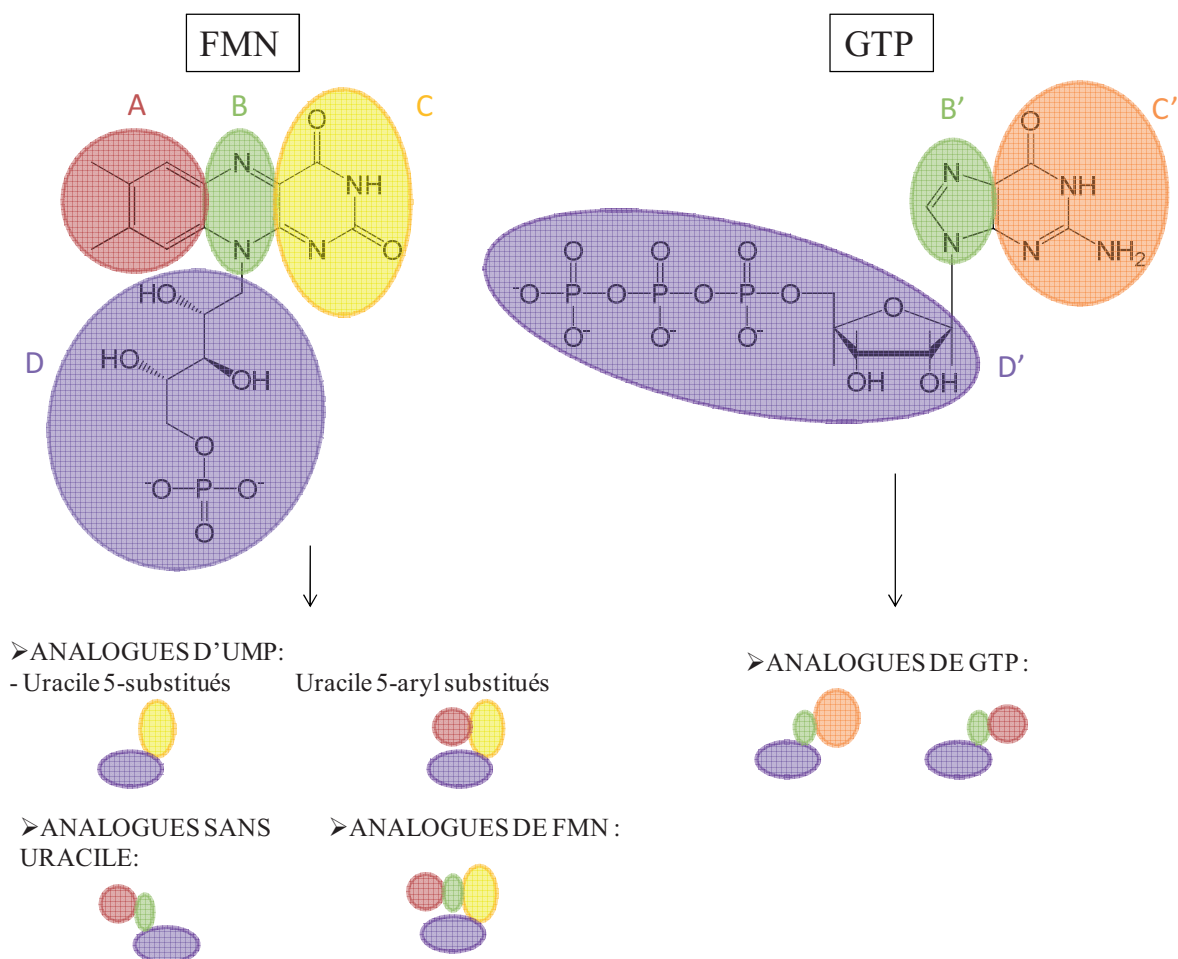



Figure 80 : Découpage du FMN ou du GTP en motifs structuraux caractéristiques et séries d'analogues associées. A : cycle hydrophobe, B : pyrimidine 1,4, C : uracile, pyrimidine 1, 3


D : sucre et phosphate. B' : imidazole, C' : dérivé de pyrimidine 1, 3, D' : sucre et triphosphate. \* : motif changé

La série analogue d'uracile conserve de fait la partie pyrimidine du FMN. Pour les motifs D ou D', la modulation consiste systématiquement à remplacer le phosphate ou le triphosphate par un phosphonate pour éviter l'hydrolyse du groupement (on parle donc de phosphonucléoside) et la partie sucre par une chaîne carbonée portant une insaturation (excepté pour un analogue où le sucre acyclique est conservé). Seuls des monophosphonucléosides ont été synthétisés. D'après des études d'homologie structurale, les chaînes de 4 carbones portant une insaturation miment la structure du ribose du GTP alors que celles de 5 carbones miment la structure du ribitol, sucre acyclique du FMN

A partir du FMN, trois séries d'analogues ont été synthétisés :

- ANALOGUES D'UMP :

- Phosphonates uracile 5-substitués :   
Les analogues de cette série comportent uniquement les motifs C et D du FMN. Le motif C est éventuellement modulé par des substitutions en position 5. Le motif D est constitué d'une chaîne de 3 carbones portant une insaturation.

- Phosphonates uracile 5-aryl substitués :   
Les analogues de cette série comportent les motifs A, C et D du FMN. Le motif C n'est pas changé. Le motif A est éventuellement modulé par des substitutions en position 4. Pour le motif D, la longueur de la chaîne carbonée portant une insaturation est variée (entre 3 et 5 carbones). Pour un des analogues, deux fonctions hydroxyles sont conservées sur la chaîne carbonée.

- ANALOGUE SANS URACILE : 

Un analogue comporte les motifs A, B et D du FMN. Le motif A n'est pas changé. Le motif B est remplacé par un imidazole. Le motif D est constitué d'une chaîne de 5 carbones portant une insaturation.



- ANALOGUES DE FMN :

Seul le motif D du FMN a été modulé. La chaîne carbonée n'est plus rattachée au cycle central mais à l'uracile. La longueur de la chaîne a également été variée.

A partir du GTP, une seule série d'analogues a été synthétisée :



- ANALOGUES DE GTP :

Par rapport au GTP, les motifs C'et D' sont modifiés, le motif B' ne change pas. Le motif D'est une chaîne de 5 carbones portant une insaturation. Le motif C' varie d'un analogue à l'autre.

## 2. Protocole de criblage

Dans le projet METASUPP, il a été défini que tous les analogues de nucléotides devaient d'abord être criblés par MS pour leur capacité à se lier à la PEBP. Ensuite les meilleurs ligands sont analysés plus finement en spectrométrie de masse pour mesurer leur  $K_D$ , et être testés par RMN pour déterminer leurs sites d'interaction et si possible un  $K_D$ . En parallèle, ils doivent être testés par biologie cellulaire sous forme de prodrogues pour tester leur capacité à inhiber la migration ou l'invasion sur des cellules tumorales à potentiel métastatique.

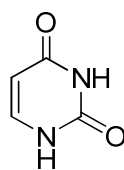
Le criblage rapide des analogues de nucléotides pour leur capacité à se lier à la PEBP consiste à comparer les fractions de liaison, de complexes sur protéine totale, déterminées par MS. Les analogues de nucléotides sont mis à incuber durant 8 minutes à 37 °C avec la hPEBP1 dans un tampon ABC d'une concentration de 20 mM. Les échantillons sont ensuite analysés par ESI-IT en utilisant la méthode d'infusion en sandwich et les optimisations instrumentales développés dans le chapitre I.

### 3. Criblage rapide avec les fractions de liaison

#### 3.1. Analogues d'UMP

##### 3.1.1. Description

Les phosphononucleosides acycliques de cette série contiennent tous la base uracile qui est au sein de la structure du FMN.

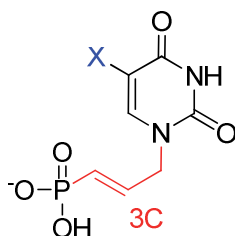
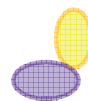


uracile

Figure 81: Structure de l'uracile

Les analogues d'UMP se répartissent en deux sous-groupes :

- Les phosphonates uracile 5-substitués (Tableau 6)

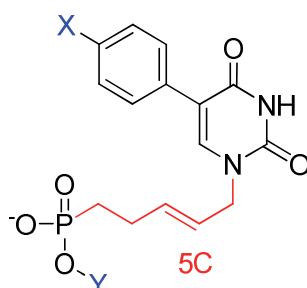
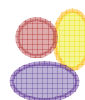


Ils comportent une base uracile substituée en position 5 par un halogène ou un groupement non-benzénique et une chaîne carbonée de 3 carbones portant une insaturation et se terminant par un phosphonate.

Tableau 6: Modulations chimiques du groupe de ligands « phosphonates uracile 5-substitués »

Ligand	Nombre de carbones entre la base et le phosphonate	X
BLA-C-03	3	H
BLA-C-09	3	Cl
BLA-C-17	3	F

- Les phosphonates uracile 5-aryl substitués (Tableau 7)



Ils comportent une base uracile substituée en position 5 par un benzène, une chaîne carbonée de 3 ou 4 carbones portant une insaturation qui se termine par un phosphonate. Le benzène peut également être substitué en position 4.

Tableau 7: Modulations chimiques du groupe de ligands « phosphonates uracile 5-aryl substitués »

Ligand	Nombre de carbone entre la base et le phosphonate	X	Y
BLA-C-02	3	H	CH <sub>2</sub> -CH <sub>3</sub>
BLA-C-04	3	O-CH <sub>3</sub>	H
BLA-C-06	4	H	H
BLA-C-10	4	* S entre uracile et benzène	H
BLA-C-11	3	F	H
BLA-C-12	4	F	H

Ces premiers analogues d'UMP permettent donc de tester l'effet de 5 modulations chimiques sur la fraction de liaison avec hPEBP1 :

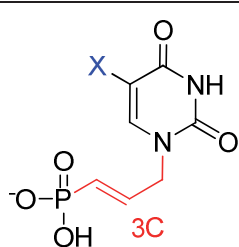
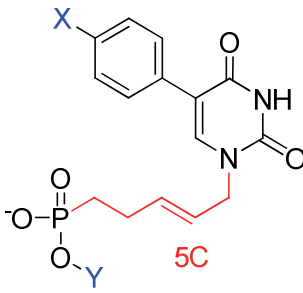
- ✓ substitution halogénée en position 5 de l'uracile
- ✓ ajout d'un benzène (motif A) en position 5 de l'uracile
- ✓ substitution en position 4 du benzène (X)
- ✓ protection partielle du phosphonate (Y)
- ✓ variation de la longueur de la chaîne carbonée entre la base et le phosphonate pour les phosphonates uracile 5-aryl substitués



### 3.1.2. Résultats et discussion

D'après le criblage de ce premier lot, aucune fraction de liaison n'est visible par MS pour les phosphononucléosides acycliques qui ne sont pas aryl-substitués (Tableau 8).

Tableau 8: Criblage des phosphononucléosides acycliques par la mesure de fraction de liaison à PEBP à pH 8.4 : 18  $\mu$ M de PEBP avec 500  $\mu$ M de ligand.

Famille chimique	Structure	Nom	Modulation chimique		Nombre de carbones	Fraction de liaison (%)
			X	Y		
phosphonates uracile 5-substitués		BLA-C-03	H	H	3	0
		BLA-C-09	Cl	H	3	0
		BLA-C-017	F	H	3	0
phosphonates uracile 5-aryl substitués		BLA-C-02	H	C <sub>2</sub> H <sub>5</sub>	3	0
		BLA-C-04	O-CH <sub>3</sub>	H	3	16
		BLA-C-06	H	H	4	5
		BLA-C-010	* S entre uracile et benzène	H	4	4
		BLA-C-011	F	H	3	7
		BLA-C-012	F	H	4	4

La présence d'un benzène en position 5 de l'uracile sur ces analogues semble nécessaire pour observer une fraction de liaison. De plus, le seul analogue aryl-substitué qui ne se lie pas à la PEBP est celui qui possède encore une protection sur le phosphonate. Pour l'anecdote, la protection partielle de l'analogue n'était pas intentionnelle mais a résulté d'une étape de déprotection incomplète durant la synthèse. Cet analogue a permis de mettre en évidence que, pour se lier à la PEBP, l'analogue doit non seulement être aryl-substitué mais également contenir un phosphonate non protégé.

La substitution, en position 4 du benzène, d'un atome d'hydrogène (BLA-C-06) par un atome de fluor (BLA-C-12) n'a aucun effet significatif sur la fraction de liaison. Par contre, la substitution par un groupement méthoxy entraîne une augmentation de la fraction de liaison.

Ainsi, la meilleure fraction de liaison est obtenue avec BLA-C-04, qui possède une chaîne carbonée de 3 carbones et une fonction méthoxy sur le benzène. La variation de la longueur de la chaîne carbonée a été testée avec les composés BLA-C-011 et BLA-C-012 qui contiennent respectivement des chaînes de 4 et 3 carbones. La différence de fraction de liaison mesurée n'est pas assez significative pour en tirer des conclusions.

## 3.2. Nucléotides de référence

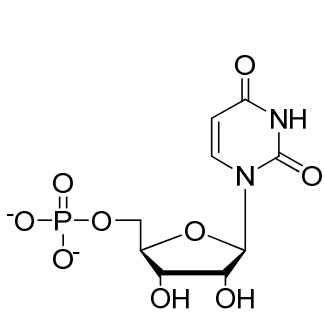
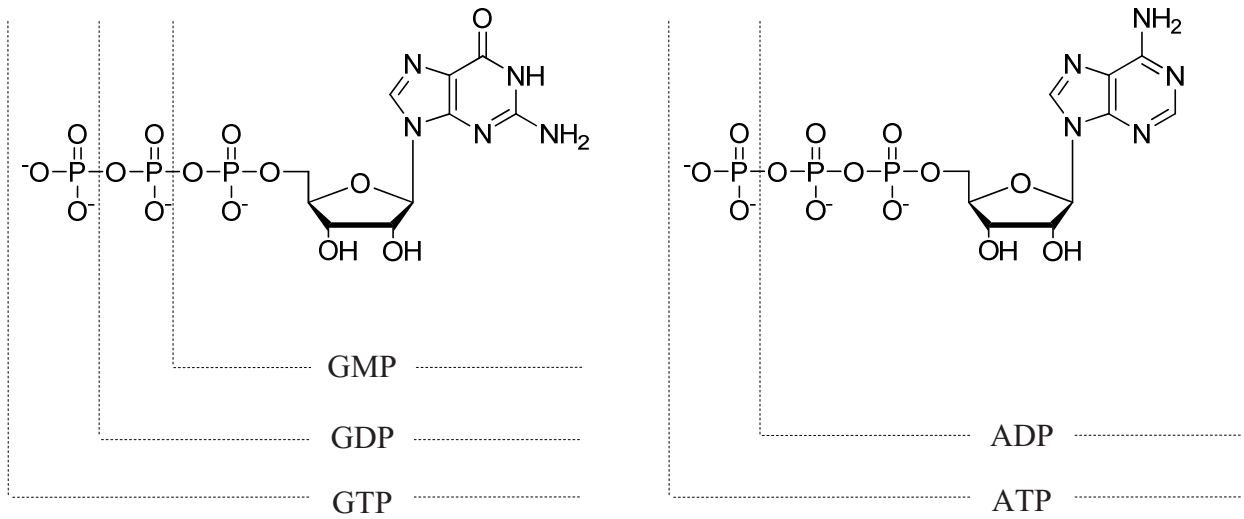
### 3.2.1. Description

L'étude de Bucquoy et ses collaborateurs a permis de classer les nucléotides selon capacité à éluer la protéine à pH 7.5 : FMN > GTP > GDP > GMP > FAD > ATP > NADP > CTP > UTP > ADP (Bucquoy 1994). L'AMP, l'AMPc, le GMPc et le NAD avaient également été testés mais ils ne se liaient pas à la bPEBP1 (73).

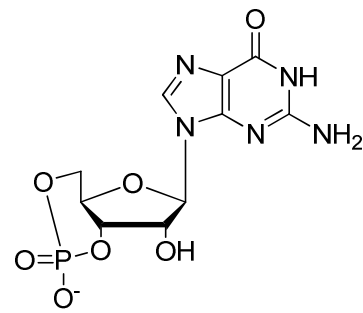
Des mononucléotides, GTP, GDP, GMP, ATP, ADP, UMP, GMPc et FMN et des dinucléotides, FAD, NAD<sup>+</sup>, NADH,H<sup>+</sup>, NADP<sup>+</sup> et NADPH,H<sup>+</sup> ont été sélectionnés pour tester leur liaison à la PEBP humaine.

#### ✓ Mononucléotide :

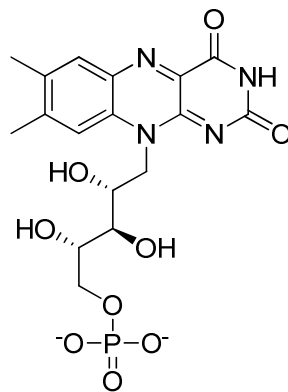
Les mononucléotides sélectionnés sont mono-, di- ou tri-phosphorylés, ou cycliques, ce qui permet d'étudier l'effet du nombre de groupements phosphates sur la fraction de liaison. Certains nucléotides comportent le même nombre de phosphate mais ont des bases différentes. L'effet seul de la base sur la fraction de liaison peut ainsi également être étudié.



UMP



GMPc

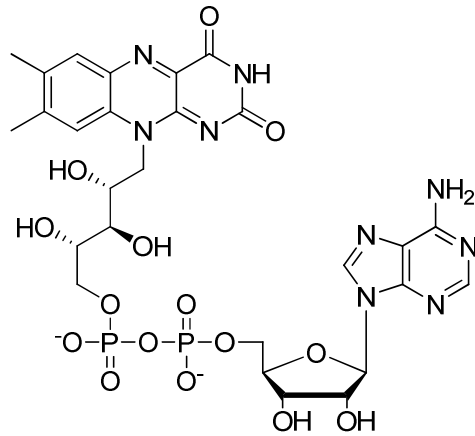


FMN

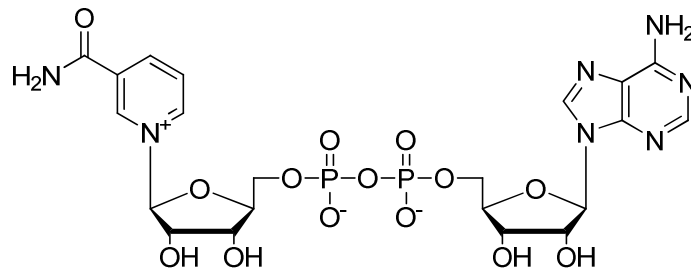
Figure 82: Structure des mononucléotides

✓ Dinucléotide :

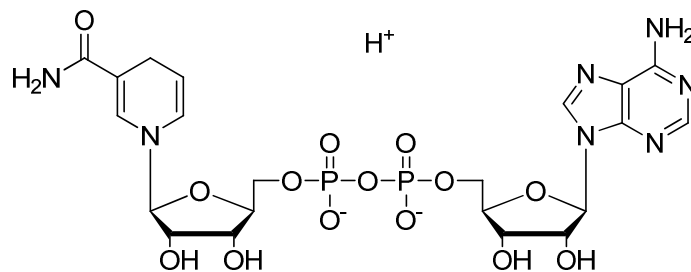
Les dinucléotides sélectionnés sont di- ou tri-phosphorylés, et oxydés ou réduits. L'effet de ces propriétés peut donc être évalué sur la fraction de liaison.



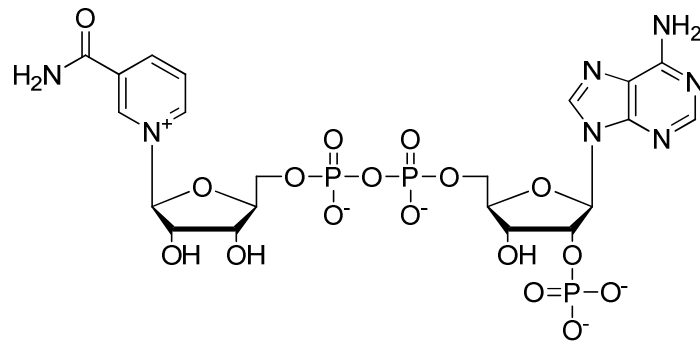
FAD



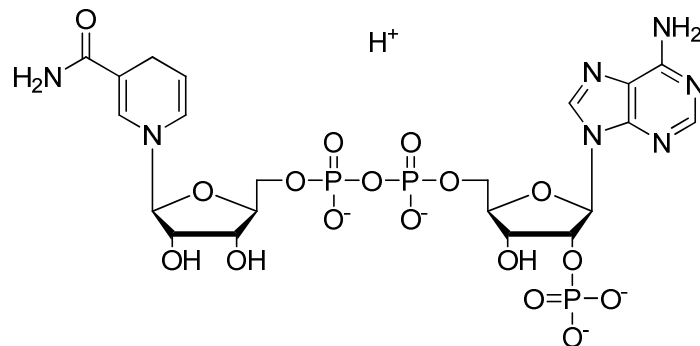
NAD<sup>+</sup>



NADH, H<sup>+</sup>



NADP<sup>+</sup>

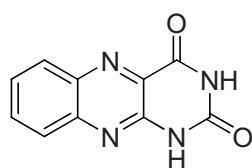


NADPH, H<sup>+</sup>

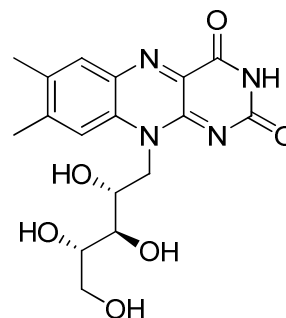
Figure 83 : Structure des dinucléotides

- ✓ Base azotée et nucléoside

Pour voir si le phosphate et le sucre sont indispensables à la liaison, la riboflavine, qui correspond au FMN sans le phosphate, et l'alloxazine, qui correspond au FMN sans la partie sucre ni le phosphate et sans les méthylations sur le benzène ont été criblées. Le lumichrome est la molécule qui correspond exactement au FMN sans la partie sucre et le phosphate, mais celle-ci n'a pas été testée.



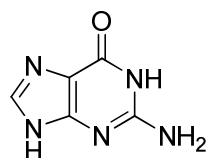
alloxazine



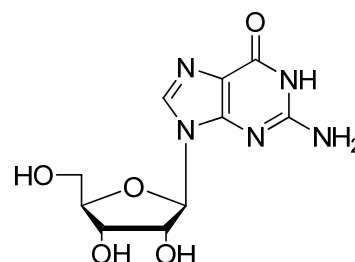
riboflavine

Figure 84: Structure de l'alloxazine et de la riboflavine

La guanine et la guanosine qui correspondent respectivement au GTP sans le sucre et les triphosphates, et au GTP sans les triphosphates ont également été testées.



guanine



guanosine

Figure 85: Structure de la guanine et de la guanosine

Lors du premier criblage, les phosphonucléosides acycliques ont été ajoutés en fort excès : 500 pour 18  $\mu\text{M}$ . Cela a conduit à des stœchiométries apparentes très élevées, signe d'agrégation en phase gazeuse. Afin de limiter ce phénomène, une étape de MGF a été ajoutée avant l'analyse MS pour le criblage des nucléotides de référence.

### 3.2.2. Résultats et discussion

Les résultats du criblage mettent en évidence que les formes sans phosphate, la guanine, la guanosine, la riboflavine et l'alloxazine, ne présentent pas de fraction de liaison visible avec la PEBP (Tableau 9).

Tableau 9: Criblage de nucléotides de référence par la mesure de fraction de liaison à PEBP à pH 8.4 et 37°C: 18  $\mu$ M de PEBP avec 500  $\mu$ M de ligand en utilisant la MGF.

Ligand	Fraction de liaison (%)	Ligand	Fraction de liaison (%)
Guanine	0	FMN	18.5
Guanosine	0	FAD	21
GMPc	0	ADP	12
GMP	11	ATP	18
GDP	13	NAD <sup>+</sup>	14
GTP	12-16	NADP <sup>+</sup>	25
alloxazine	0	NADH, H <sup>+</sup>	12
riboflavine	0	NADPH, H <sup>+</sup>	24

Le GMP cyclique qui ne possède pas de phosphate libre mais un sous forme cyclisé ne présente pas non plus de fraction de liaison visible avec PEBP, ce qui est en accord avec les résultats de Bucquoy (73).

Le NAD semble se lier à la hPEBP1, alors qu'il ne semble pas se lier à bPEBP1 d'après les résultats de Bucquoy et coauteurs (73).

Les fractions de liaison des séries GMP/GDP/GTP, FMN/FAD, ADP/ATP, NAD<sup>+</sup>/NADP<sup>+</sup>, NADH,H<sup>+</sup>/NADPH,H<sup>+</sup> laissent penser que plus il a de phosphates sur la molécule et plus la fraction de liaison est élevée. Il est probable que les phosphates renforcent la liaison avec la PEBP via des interactions électrostatiques.

La nature oxydée ou réduite des cofacteurs NAD et NADP n'a pas d'effet sur la fraction de liaison.

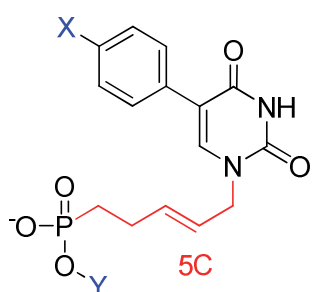
A nombre de phosphore égal (monophosphonate/monophosphate et diphosphate), les composés à base de flavine présentent de meilleures fractions de liaison. Pour les

triphosphates, ce sont les molécules NADP<sup>+</sup> et NADPH,H<sup>+</sup> qui présentent les meilleures fractions de liaison, ce sont aussi les meilleures fractions de liaison de l'ensemble des molécules testées dans ce criblage. Afin d'approfondir ces résultats, la recherche d'une éventuelle activité enzymatique détenue par la protéine PEBP a été entreprise. Des tests d'activité enzymatique d'oxydo-réduction ou de phosphorylation ont été réalisés et seront présentés dans le chapitre V. La suite du criblage est, elle, restée centrée sur les mononucléotides, presque aussi affins pour hPEBP1 et plus abordables du point de vue de la synthèse organique.

### 3.3. Variation de la longueur de chaîne de BLA-C-04

Suite au premier criblage d'analogues d'UMP, nous avons voulu examiner l'effet de la longueur de la chaîne carbonée sur le meilleur ligand, BLA-C-04, dans l'optique d'optimiser cette longueur pour les nouvelles synthèses d'analogues. Ainsi, les composés BLA-C-04, BLA-C-026 et BLA-C-025 qui comportent respectivement 3, 4 et 5 carbones ont constitué le deuxième lot d'analogues à être criblé. Lors de ce criblage, une étape de MGF a également été ajoutée avant l'analyse MS pour limiter les phénomènes d'agrégation.

Tableau 10: Effet de la variation de la longueur de la chaîne carbonée du composé BLA-C-04 sur la fraction de liaison à PEBP pH 8.4 et 37°C: 18 μM de PEBP avec 500 μM de ligand en utilisant la MGF

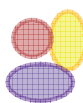
Famille chimique	Structure	Nom	Modulation chimique		Nombre de carbones	Fraction de liaison (%)
			X	Y		
phosphonates uracile 5-aryl substitués		BLA-C-04	O-CH <sub>3</sub>	H	3	13
		BLA-C-026	O-CH <sub>3</sub>	H	4	15
		BLA-C-025	O-CH <sub>3</sub>	H	5	12



Le composé BLA-C-026 n'est pur qu'à 38%. Pour avoir une concentration finale de 500  $\mu\text{M}$  de ligand, la pureté apparente a été prise en compte, mais le résultat doit être interprété avec précaution.

Les fractions de liaison observées pour les molécules BLA-C-04, BLA-C-025, BLA-C-026 suggèrent qu'il n'y a pas d'effet de la variation de la longueur de la chaîne carbonée sur la fraction de liaison. Les synthèses suivantes ont été orientées vers des analogues de BLA-C-025, avec une chaîne à cinq carbones, car le meilleur mononucléotide testé, le FMN, possède une chaîne à 5 carbones.

### 3.4. Deuxième série d'analogues d'UMP



#### 3.4.1. Description

Un nouveau lot de phosphononucléosides acycliques dérivés de BLA-C-025 a été criblé. L'UMP a également été testé car il correspond à la molécule de référence à partir de laquelle tous ces phosphononucléosides dérivent étant donné qu'ils comportent tous une base uracile. La pharmaco-modulation du composé BLA-C-025 a consisté à:

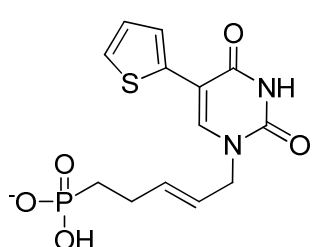
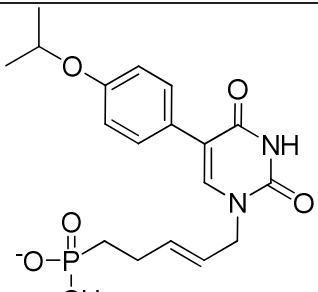
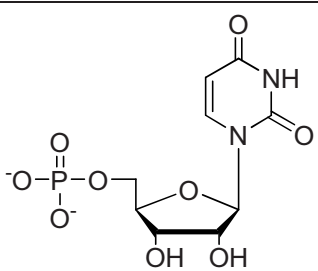
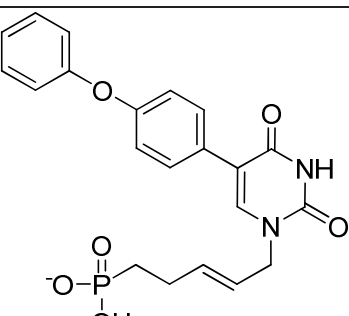
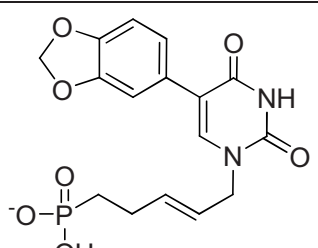
- ✓ modifier le méthoxy substitué en position 4 du benzène : trifluorométhoxy (BLA-C-030), isopropyloxy (BLA-C-032), ethoxy (BLA-C-029), benzyloxy (BLA-C-028)
- ✓ ajouter une substitution en position 3 du benzène : BLA-C-027 et BLA-C-031
- ✓ introduire un diol à la place de l'insaturation : BLA-C-033
- ✓ remplacer le benzène par un thiophène : BLA-C-034

#### 3.4.2. Résultats et discussion

Lors de ce criblage, le ligand a été ajouté en faible excès, 88 pour 12, ce qui permet d'éviter l'étape de MGF tout en limitant les phénomènes d'agrégation. D'autre part, l'incubation a été réalisée à pH 7.4 pour se rapprocher des conditions physiologiques et faciliter la comparaison avec les résultats du criblage cellulaire. Dans ces conditions, la fraction de liaison avec BLA-C-025 est de 25 %, par conséquent plus élevée qu'à pH 8.4: où elle n'est que de 12 %. En solution à pH 7.4, la PEBP ayant plus de chances d'avoir davantage de charges positives qu'à 8.4 et BLA-C-025 étant chargé négativement, il est très probable que l'interaction avec la PEBP soit donc renforcée à pH 7.4.

Malgré les différentes modulations réalisées, tous les analogues de ce criblage se lient à la PEBP. Il y a un facteur de 42% entre les deux valeurs extrêmes de fractions de liaison.

Tableau 11: Criblage des analogues d'UMP dérivés de BLA-C-025 par la mesure de fraction de liaison à PEBP à pH 7.4 et 37°C: 12  $\mu$ M de PEBP avec 88  $\mu$ M de ligand.

Ligand	Structure	Fraction de liaison (%)
BLA-C-034		33
BLA-C-032		28
UMP		27
BLA-C-028		25
BLA-C-027		25

BLA-C-025		25
BLA-C-031		24
BLA-C-04		23
BLA-C-033		21
BLA-C-030		20
BLA-C-029		19

D'après ces données (Tableau 11), le remplacement du benzène par un thiophène n'est pas néfaste pour la liaison. Au contraire, cet analogue présente la plus grande fraction de liaison

du criblage. L'introduction d'un diol à place de l'insaturation (BLA-C-033) ne montre pas d'effet significatif sur la fraction de liaison. La variation de l'alkyloxy substitué en position 4 du benzène entraîne des différences de fraction de liaison significatives. Le composé portant le groupement éthoxy (BLA-C-029) présente la fraction de liaison la plus faible, 19%, et le composé portant le groupement isopropyloxy (BLA-C-032) une des plus élevées, 28%. Ces deux composés restant très semblables, il est difficile d'interpréter cette différence de fraction de liaison. Finalement, l'ajout d'une substitution en position 3 du benzène n'a pas d'effet sur la fraction de liaison.



### 3.5. Analogues de FMN

#### 3.5.1. Description

Deux analogues de FMN ont été synthétisés. Ils ont la particularité de porter la chaîne carbonée sur l'uracile et non sur le cycle central. La conséquence majeure de ce déplacement est que les électrons sont beaucoup moins délocalisés ce qui peut conduire à une perte des propriétés oxydo-réductrices de la molécule. Les activités oxydo-réductrices des analogues de FMN seront étudiées dans le chapitre V. Tout comme les analogues de mononucléotides, ils portent un groupe phosphonate et non phosphate. Ils diffèrent l'un de l'autre par la longueur de la chaîne carbonée (4 carbones pour BLA-C-051 et 5 pour BLA-C-054).

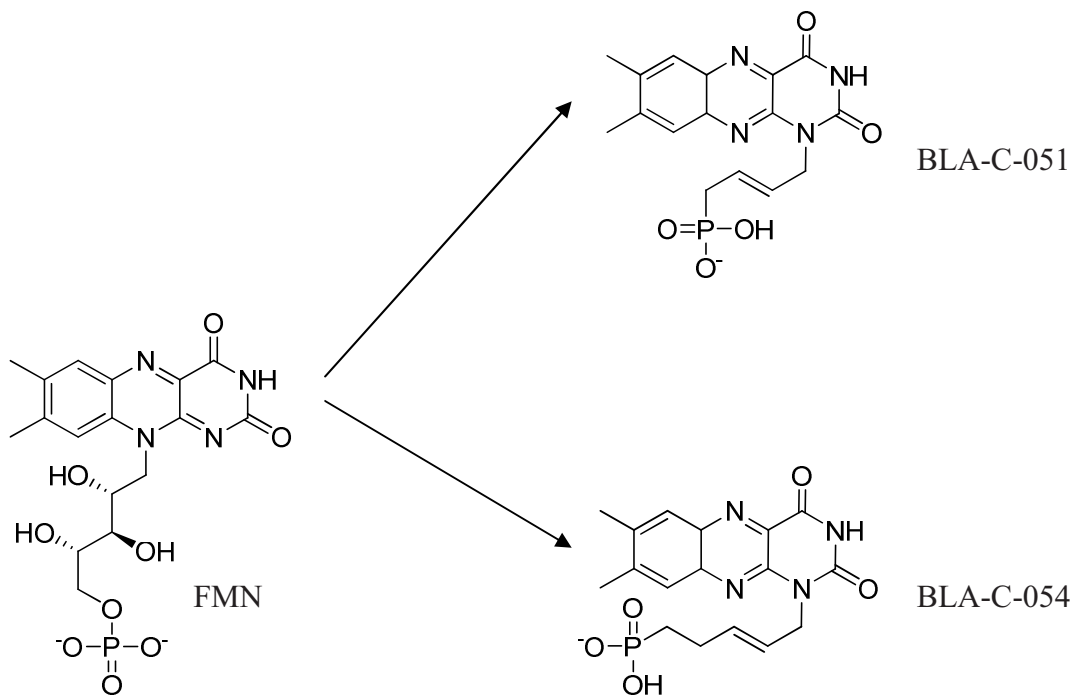
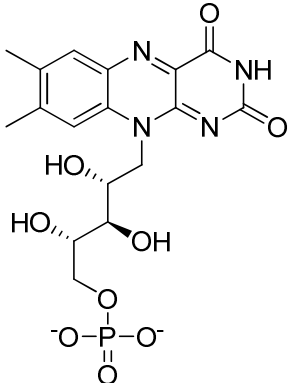
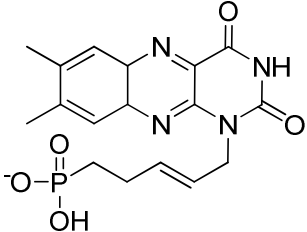
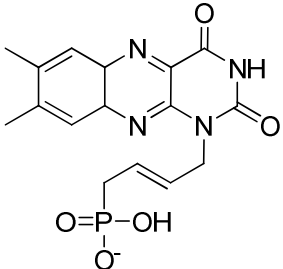


Figure 86: Structure du FMN et des analogues de FMN synthétisés.

### 3.5.2. Résultats et discussion

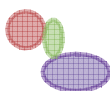
Le criblage a été réalisé à pH 7.4. La fraction de liaison mesurée avec le FMN à ce pH est plus élevée que celle obtenue à pH 8.4 (criblage des nucléotides de référence).

Tableau 12: Criblage d'analogues de FMN par la mesure de fraction de liaison à PEBP à pH 7.4 et 37°C: 12  $\mu$ M de PEBP avec 88  $\mu$ M de ligand.

Ligand	Structure	Fraction de liaison (%)
FMN		37
BLA-C-054		33
BLA-C-051		28

L'analogue de FMN avec une chaîne de 5 carbones, BLA-C-054, se lie à la PEBP avec une fraction de liaison comparable à celle du FMN. L'absence de sucre combiné à la ramification de la chaîne carbonée sur l'uracile ne semble donc pas avoir d'effet significatif sur la liaison. Par contre, l'analogue de FMN avec une chaîne de 4 carbones, BLA-C-051, présente une fraction de liaison plus faible que celle du FMN. La diminution de la longueur de la chaîne carbonée, de 5 à 4 carbones, pourrait donc avoir un léger effet négatif sur la liaison à PEBP.

### 3.6. Analogues sans uracile



#### 3.6.1. Description

Un analogue de FMN sans base uracile a été synthétisé : BLA-C-049. Le motif A du FMN, le diméthylbenzène, est conservé et le motif B, une pyrazine, est transformé en imidazole. Le motif C, l'uracile, est absent. Le rôle de l'uracile dans la liaison avec hPEBP1 peut ainsi être testé.

#### 3.6.2. Résultats et discussion

BLA-C-049 ne se lie pas à hPEBP1. Comme les analogues de FMN ont montré que le remplacement du sucre par une chaîne carbonée n'est pas délétère pour la liaison, seule la perte de l'uracile peut expliquer la perte de la liaison.

Tableau 13: Criblage de l'analogue de FMN sans uracile par la mesure de fraction de liaison à PEBP à pH 7.4 et 37°C: 12  $\mu$ M de PEBP avec 88  $\mu$ M de ligand.

Ligand	Structure	Fraction de liaison (%)
FMN		37
BLA-C-049		0

L'uracile qui constitue le motif C du FMN est donc indispensable dans la liaison avec hPEBP1.

### 3.7. Analogues de GTP



#### 3.7.1. Description

Deux analogues de GTP ont été synthétisés. Ils portent un phosphonate à la place du triphosphate et ne contiennent pas de ribose. Les composés BLA-C-055 et BLA-C-056 sont très proches. BLA-C-055 diffère de BLA-C-056 par la perte de l'amine de la base.

L'analogue de FMN sans uracile, BLA-C-049, peut également être classé parmi les analogues de GTP. En effet, il est identique aux analogues de GTP, BLA-C-055 et BLA-C-056, à part pour le motif C' du GTP, le dérivé pyrimidique, qui est remplacé par un diméthylbenzène. Cet analogue permet de tester l'importance du motif C' du GTP dans la liaison à hPEBP1.

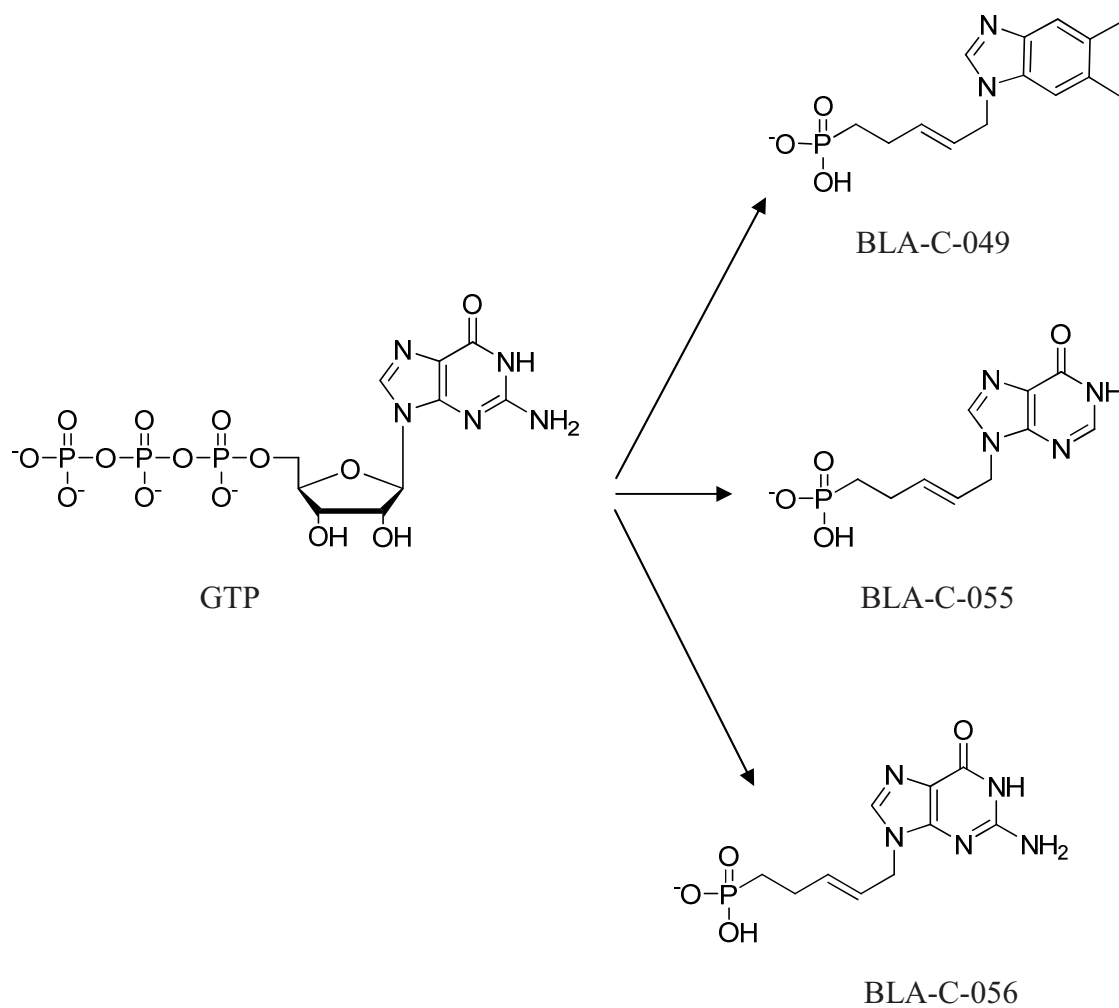
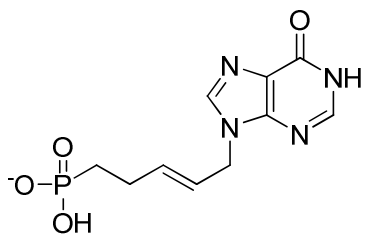
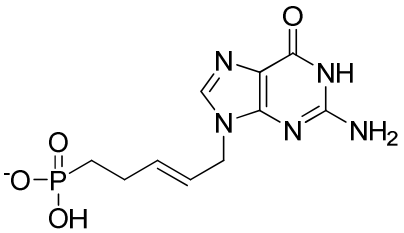
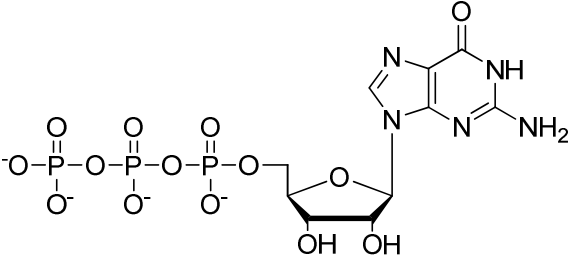
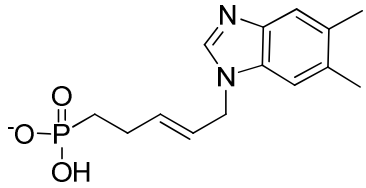


Figure 87: Structure du GTP et des analogues de GTP synthétisés.

### 3.7.2. Résultats et discussion

BLA-C-049 ne se lie pas à hPEBP1 (Tableau 14). Les deux autres analogues de GTP (BLA-C-055 et BLA-C-056) ont conservé le dérivé pyrimidique et se lient à PEBP. On peut donc en déduire que la conservation du dérivé pyrimidique (motif C') est essentielle pour observer une fraction de liaison entre la hPEBP1 et l'analogue. Sa structure, ou au minimum une partie de sa structure, doit donc jouer un rôle clé dans l'interaction PEBP-nucléotides.

Tableau 14: Criblage d'analogues de GTP par la mesure de fraction de liaison à hPEBP1 à pH 7.4 et 37°C: hPEBP1 12 µM avec 88 µM de ligand.

Ligand	Structure	Fraction de liaison (%)
BLA-C-055		25
BLA-C-056		21
GTP		21
BLA-C-049		0



Les analogues BLA-C-055 et BLA-C-056 présentent des fractions de liaisons très proches de celle du GTP. Ainsi, l'absence combinée de sucre cyclique et de triphosphate n'a pas d'effet significatif sur la fraction de liaison.

### 3.8. Molécules du criblage de Shemon

#### 3.8.1. Description

Un criblage par RMN de liaison à rPEBP1 mené sur environ 6000 composés par Shemon et ses collaborateurs (101) a permis de découvrir 3 nouvelles molécules possédant une affinité sur la protéine de rat: le N'-(2,4-dinitrophenyl)-2,3,4,5,6-pentafluorobenzohydrazide (composé n°26 dans la publication citée), le N-[(2-bromo-4-nitrophenyl)amino]carbonothioyl}-3-(2-thienyl)acrylamide (composé 48) et le 4-({4-[(2,4-dinitrophenyl)amino]phenyl} amino)-4-oxo-2-butenoic acid (composé 98). Les trois molécules portent un groupement 4-nitrophényl substitué en position 2. Seuls les composés 26 et 98 ont pu être solubilisés et ont ensuite été analysés en conditions natives.

Ces composés peuvent présenter des analogies avec le FMN suivant leur représentation. Ils pourraient notamment mimer les trois cycles accolés du FMN.

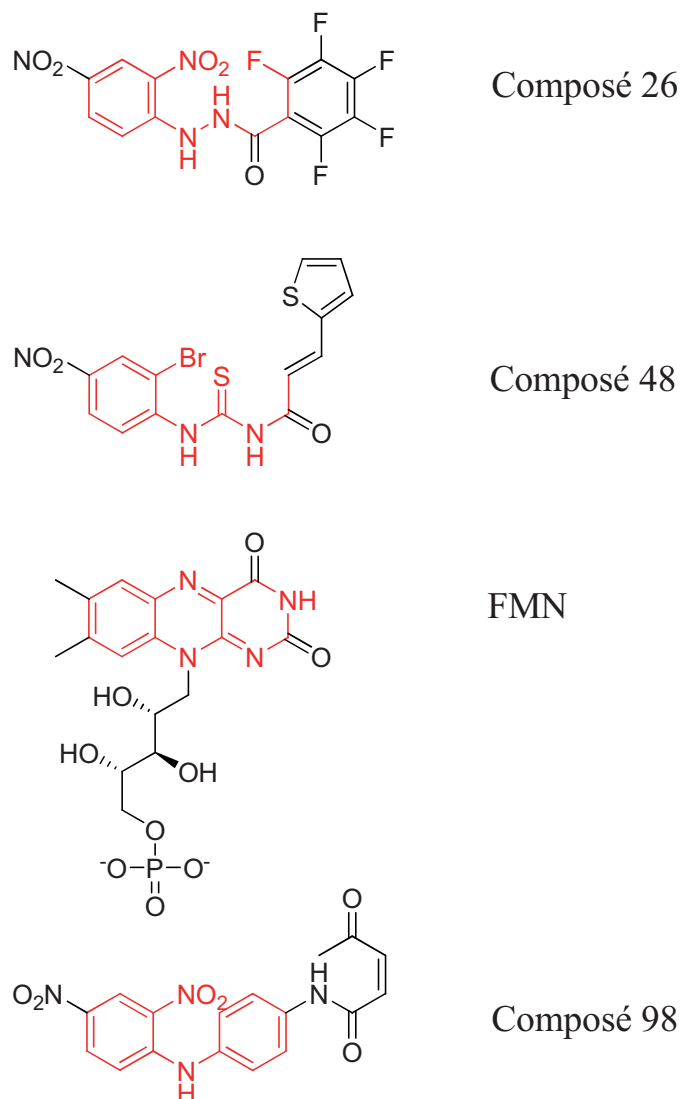
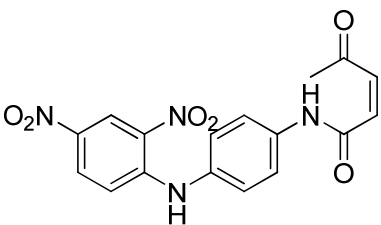
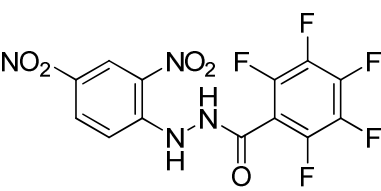


Figure 88: Alignement des structures des composés 26, 48 et 98 de Shemon (101) avec celle du FMN.

### 3.8.2. Résultats et discussion

Nous avons pu montrer que la molécule « Shemon 98 », qui ne possède ni phosphate ni phosphonate, se lie à la PEBP humaine. Ainsi la présence d'un groupe phosphate/phosphonate n'est pas indispensable pour la liaison avec hPEBP1 (Tableau 15).

Tableau 15: Criblage des molécules Shemon par la mesure de fraction de liaison à hPEBP1 à pH 7.4 et 37°C: hPEBP1 12  $\mu$ M avec 88  $\mu$ M de ligand.

Ligand	Structure	Fraction de liaison (%)
Shemon 98		23
Shemon 26		0

La molécule « Shemon 98 » porte deux nitrations sur un benzène, ce qui pourrait constituer des équivalents au phosphate ou au phosphonate et expliquer la liaison. Cependant, la molécule « Shemon 26 » possède aussi un benzène avec deux nitrations, et aucune liaison n'est observée. Si le dinitrobenzène est nécessaire à la liaison, il n'est pas suffisant. La molécule « Shemon 98 » possède une partie structurale qui présente des similitudes avec la base uracile, deux fonctions dioxo et un atome d'azote. C'est peut-être cette partie-là de la molécule qui mime une base, et qui permet, en combinaison avec le benzène portant deux nitrations, à la molécule de se lier à hPEBP1.

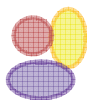
#### 4. Détermination de $K_D$ pour les meilleurs ligands

D'après un calcul simple effectué avec l'équation quadratique classique, connaissant les concentrations initiales en ligand et en protéine ainsi que les fractions de liaison obtenues lors du criblage rapide, les ligands présentent des affinités faibles pour la PEBP, avec un  $K_D$  *a priori* supérieur au micromolaire. De plus, étant donné les stœchiométries apparentes élevées

observées pour de fortes concentrations en ligand, ils ont tendance à s'agréger de manière non-spécifique à la PEBP.

La détermination de  $K_D$  a donc été réalisée avec la méthode développée dans le chapitre III qui permet de traiter des systèmes de faible affinité, sensibles à l'agrégation et sans  $K_D$  de référence. Les incubations ont été réalisées à 37 °C dans de l'ABC 20 mM, à un pH de 7.4, pour se rapprocher des conditions physiologiques.

#### 4.1. Analogues d'UMP

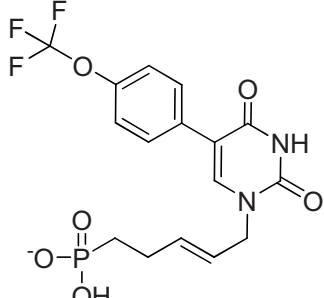
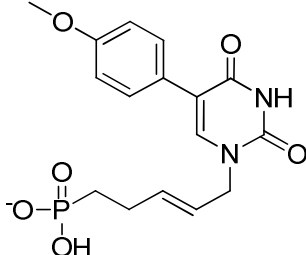
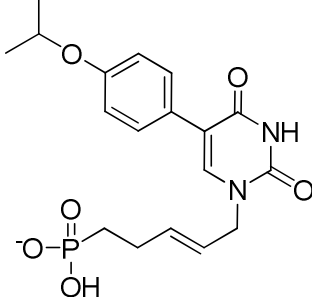
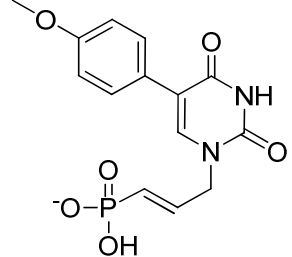
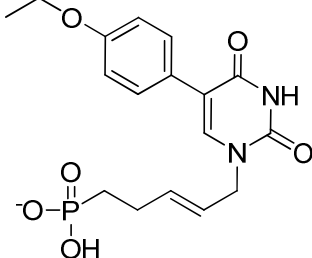


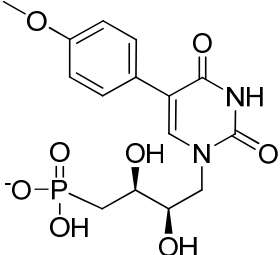
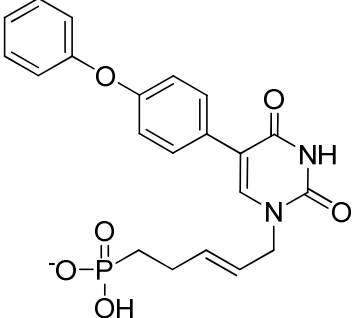
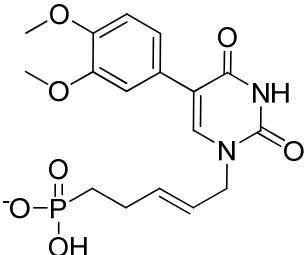
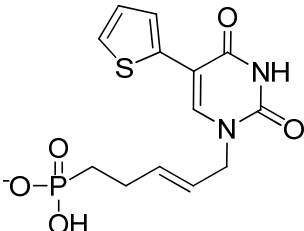
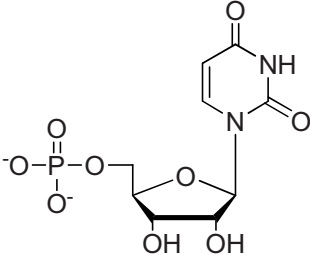
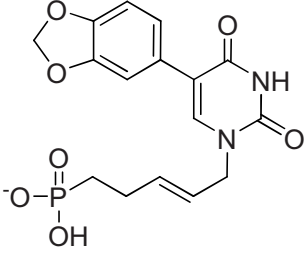
Les analogues d'UMP développés à partir de la structure de BLA-C-025 sont classés du plus affin au moins affin (Tableau 16). Le classement obtenu avec le  $K_D$  est assez différent de celui obtenu avec la fraction de liaison. Cela suggère que la sensibilité à la dissociation en phase gazeuse n'est pas égale pour tous les ligands. Ceci a effectivement été mis en évidence par la variabilité des valeurs de  $f_{sat}$  obtenues lors des déterminations de  $K_D$ . Par conséquent, la fraction de liaison seule n'est pas suffisamment représentative de l'affinité du ligand pour la PEBP pour permettre un classement des affinités relatives.

De plus, les fractions de liaison ont été obtenues dans des conditions où il y avait encore de l'agrégation en phase gazeuse et ce phénomène d'agrégation en phase gazeuse n'est pas identique pour tous les ligands. En effet, pour les trois quarts des analogues testés, on a vu une stœchiométrie apparente entre 3 et 5, et pour le quart restant, entre 1 et 2. Ainsi, la fraction de liaison déterminée par MS n'est pas représentative de celle en solution car elle est altérée par les phénomènes d'agrégation et de dissociation en phase gazeuse. Les déterminations de  $K_D$  ont, elles, été réalisées en prenant soin d'éviter les conditions d'agrégation.

La fraction de liaison doit donc être rétrospectivement considérée comme une manière d'évaluer qualitativement, au mieux semi-quantitativement, s'il y a liaison ou non avec la PEBP. L'affinité, quant à elle, doit être déterminée par une mesure de  $K_D$ .

Tableau 16: Mesure de  $K_D$  à pH 7.4 et 37 °C des analogues d'UMP

Ligand	Structure	$K_D$ ( $\mu\text{M}$ )
BLA-C-030		$2 \pm 4$
BLA-C-025		$3 \pm 5$
BLA-C-032		$3 \pm 4$
BLA-C-04		$7 \pm 4$
BLA-C-029		$8 \pm 3$

Ligand	Structure	$K_D$ ( $\mu\text{M}$ )
BLA-C-033		$10 \pm 4$
BLA-C-028		$17 \pm 5$
BLA-C-031		$19 \pm 4$
BLA-C-034		$20 \pm 4$
UMP		$23 \pm 8$
BLA-C-027		$48 \pm 13$

Le classement des dérivés de BLA-C-025 selon le  $K_D$  met en évidence une plus grande amplitude de variation que celui obtenu avec la fraction de liaison. Le remplacement du benzène par un thiophène (BLA-C-034) n'améliore pas l'affinité. Au contraire, il tend à la diminuer. L'introduction d'un diol à place de l'insaturation (BLA-C-033) ne montre pas d'effet significatif sur l'affinité. La longueur de la chaîne carbonée, 3 carbones au lieu de 5, pour le composé BLA-C-04, n'a toujours pas d'effet significatif sur l'affinité.

Les ligands les plus affins, avec un  $K_D$  compris entre 2 et 10  $\mu\text{M}$ , sont ceux qui sont structurellement proches de BLA-C-025. Ils ne comportent qu'une seule substitution, de type alkyloxy, en position 4 du benzène. Les ligands les moins affins, avec un  $K_D$  compris entre 17 et 48  $\mu\text{M}$ , sont les plus éloignés structurellement de BLA-C-025. Soit ils comportent des substitutions lourdes au niveau du benzène, soit ils ne contiennent pas de benzène (BLA-C-034 et UMP).

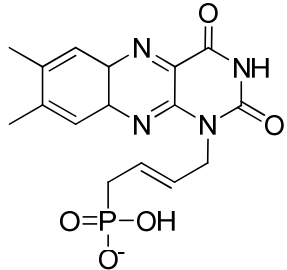
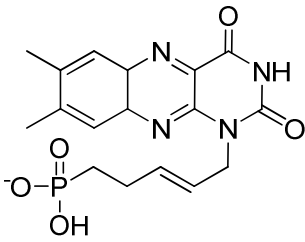
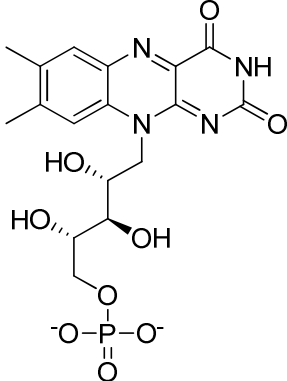
En conclusion, la longueur de la chaîne carbonée et l'introduction d'un diol ne semble pas avoir d'effet sur l'affinité. Par contre, l'ajout d'un benzène peu substitué dans la structure d'un ligand analogue d'UMP permet d'augmenter d'un facteur de 10 son affinité pour la PEBP.

#### 4.2. Dérivés de FMN



D'après le  $K_D$  des analogues de FMN, le déplacement de la chaîne carbonée du cycle central à l'uracile augmente l'affinité du ligand pour la PEBP. La longueur de la chaîne carbonée (4 carbones pour BLA-C-051 et 5 pour BLA-C-054) n'a pas d'effet sur l'affinité.

Tableau 17: Mesure de  $K_D$  à pH 7.4 et 37 °C des dérivés de FMN.\*moyenne sur deux déterminations de  $K_D$

Ligand	Structure	$K_D$ ( $\mu\text{M}$ )
BLA-C-051		$1 \pm 4$
BLA-C-054		$3 \pm 2$
FMN		$18 \pm 5^*$

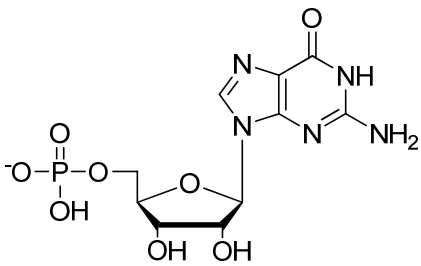
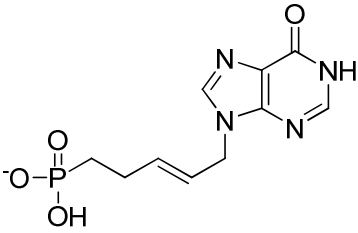
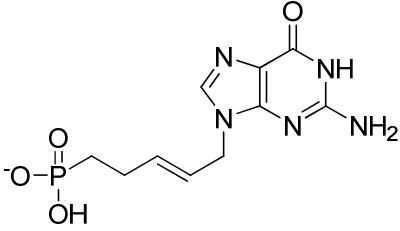
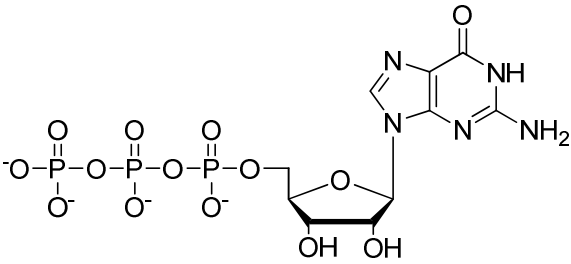
### 4.3. Dérivés de GTP



BLA-C-055 et BLA-C-056 sont plutôt des dérivés de GMP que de GTP. La perte de l'amine primaire de la guanine (BLA-C-055) n'a aucun effet sur l'affinité pour la PEBP (Tableau 18). Les deux analogues de GMP sont presque deux fois moins affins que le GMP lui-même. Cette différence est peut-être due au fait que le GMP contient un cycle de plus que BLA-C-055 ou BLA-C-056.



Tableau 18: Mesure de  $K_D$  à pH 7.4 et 37 °C des dérivés de GTP. Sauf indication contraire, les  $K_D$  ont été mesurés par ESI-IT. \* mesure en UHR-Q-TOF

Ligand	Structure	$K_D$ ( $\mu\text{M}$ )
GMP		$27 \pm 12$
BLA-C-055		$51 \pm 15$
BLA-C-056		$52 \pm 6$
GTP		$89 \pm 48^*$ $120 \pm 20$

#### 4.4. Classement général des $K_D$

Le FMN et ses analogues sont plus affins pour la PEBP que le GMP et ses analogues. Les analogues d'UMP sont également plus affins que le GMP et ses analogues. L'UMP et le GMP ont néanmoins des affinités très proches, ce qui suggère qu'ils possèdent tous les deux le squelette « minimal » pour se lier à la PEBP.

Tableau 19: Classement des  $K_D$  obtenus à pH 7.4 et 37 °C pour les meilleurs ligands. Rouge : dérivés de FMN, noir : dérivés de l'uracile, bleu : dérivés de GTP. Sauf indication contraire, les  $K_D$  ont été mesurés par ESI-IT. \* mesure en UHR-Q-TOF

Ligand	$K_D$ ( $\mu$ M)
BLA-C-051	1 $\pm$ 4
BLA-C-030	2 $\pm$ 4
BLA-C-054	3 $\pm$ 2
BLA-C-025	3 $\pm$ 5
BLA-C-032	3 $\pm$ 4
BLA-C-004	7 $\pm$ 4
BLA-C-029	8 $\pm$ 3
BLA-C-033	10 $\pm$ 4
BLA-C-028	17 $\pm$ 5
FMN	18 $\pm$ 5
BLA-C-031	19 $\pm$ 4
BLA-C-034	20 $\pm$ 4
UMP	23 $\pm$ 8
GMP	27 $\pm$ 12
BLA-C-027	48 $\pm$ 13
BLA-C-055	51 $\pm$ 15
BLA-C-056	52 $\pm$ 6
GTP	89 $\pm$ 48 * 120 $\pm$ 20

Parmi les dérivés d'UMP, il est intéressant de noter que les molécules comportant un benzène substitué en position 4 ont une affinité du même ordre de grandeur que les dérivés de FMN alors que la molécule (BLA-C-027) comportant un benzène modifié, avec une structure proche de la guanine, a une affinité du même ordre de grandeur que les dérivés du GMP.

## 5. Bilan du criblage

Suite à ces criblages, on peut déterminer qu'un analogue nucléotidique doit répondre à trois critères pour présenter une fraction de liaison avec la PEBP, visible par MS :

- Porter un groupement chargé négativement (non protégé) type phosphate, phosphonate ou un groupement type NO<sub>2</sub>.
- Comporter une base azotée ou une structure apparentée (par exemple, avec un atome d'azote et deux fonctions oxo).
- Contenir au moins deux cycles, en comptant la base

Parmi les molécules qui répondent à ces trois critères, les analogues d'UMP qui présentent les meilleures affinités sont celles qui possèdent un benzène, substitué uniquement en position 4 par un groupement peu volumineux. Les analogues de FMN qui présentent 3 cycles de 6 carbones accolés mais avec la chaîne carbonée attachée à l'uracile sont également parmi les meilleurs ligands du criblage

Il est intéressant de déterminer le squelette commun entre les meilleurs ligands (Figure 89). Les meilleurs ligands possèdent tous une base uracile simple ou accolée à un cycle, un benzène substitué par un groupement peu volumineux en position 4 et, une chaîne E-but-2-ène ou E-pent-2-ène phosphonate, rattachée à l'uracile.

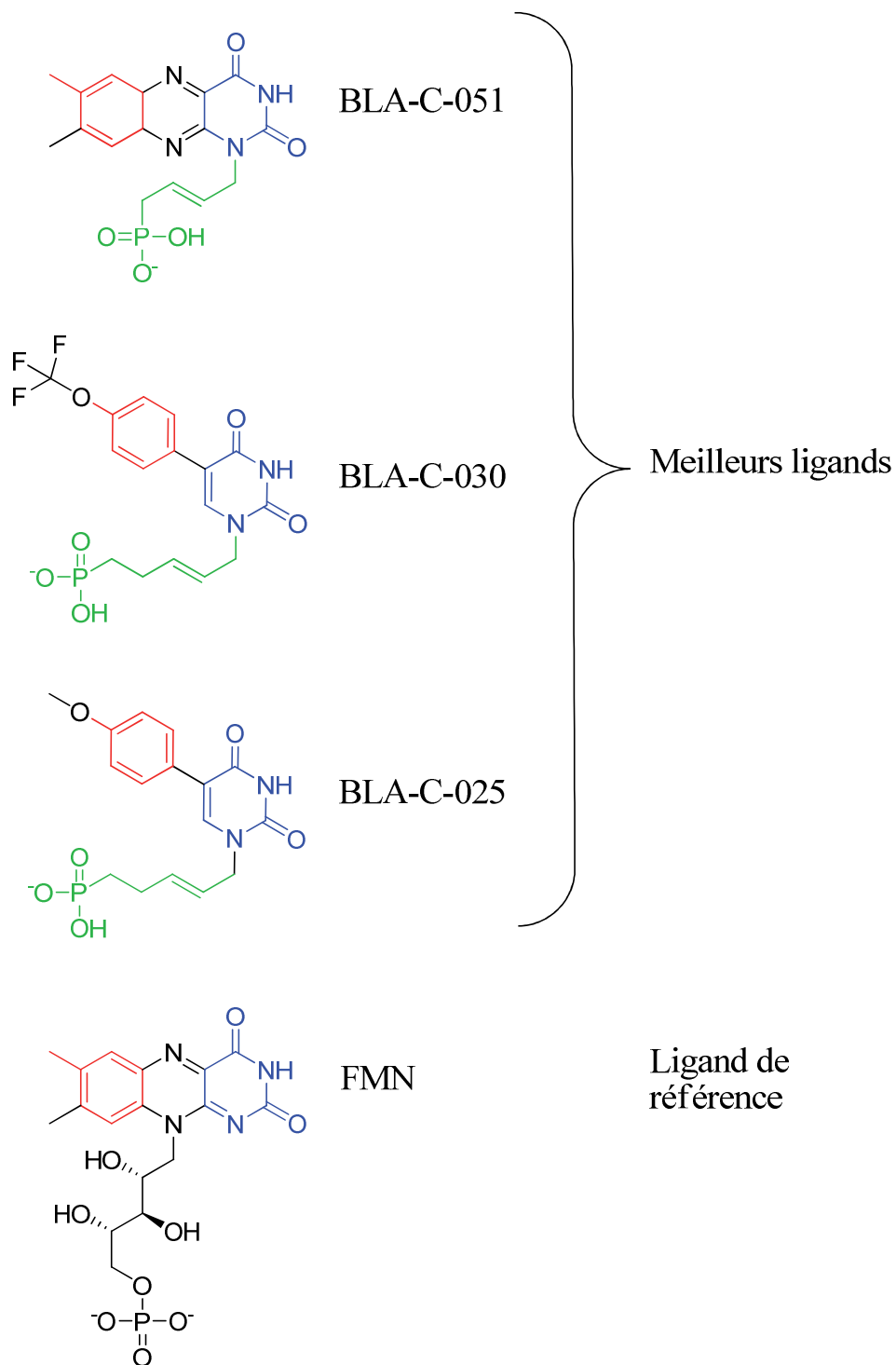


Figure 89: Squelette commun entre les meilleurs ligands BLA-C-051, BLA-C-030 et BLA-C-025 comparé au nucléotide de référence : FMN. En bleu: uracile, en rouge: méthyl benzène, en vert : E-but-2-ène ou E-pent-2-ène avec phosphonate.

## 6. Relation entre l'hydrophobicité des meilleurs ligands et les paramètres de liaison

Dans l'article de méthodologie (chapitre III), le paramètre  $f_{\text{sat}}$  a été interprété comme un indicateur du type d'interactions qui maintiennent la liaison protéine-ligand. Ainsi, un  $f_{\text{sat}}$  proche de 1 signifie que la liaison est principalement maintenue par des interactions résistantes ou peu sensibles à la dissociation en phase gazeuse telles que les interactions électrostatiques/ioniques et les liaisons hydrogène, alors qu'un  $f_{\text{sat}}$  proche de 0 signifie qu'elle est principalement maintenue par des interactions de type hydrophobe.

Le GTP, le GMP et le FMN présentent des  $f_{\text{sat}}$  relativement élevés ( $> 0.40$ ), ce qui suggère que la liaison de ces trois molécules à la PEBP repose en grande partie sur des interactions électrostatiques (Tableau 20). BLA-C-030, -032, -051, -025 et -033 présentent des  $f_{\text{sat}}$  faibles ( $< 0.20$ ). Ces molécules semblent donc se lier à la PEBP principalement par des interactions hydrophobes (Tableau 20).

Tableau 20: Calcul de logP et valeurs de  $K_D$  et  $f_{\text{sat}}$  déterminées à pH 7.4 et 37 °C pour les meilleurs ligands. Rouge : dérivés de FMN, noir : dérivés de l'uracile, bleu : dérivés de GTP

Ligand	$f_{\text{sat}}$	$K_D$ ( $\mu\text{M}$ )	logP
BLA-C-030	0.13	$2 \pm 4$	1.6
BLA-C-032	0.17	$3 \pm 4$	1.4
BLA-C-051	0.18	$1 \pm 4$	0.9
BLA-C-025	0.19	$3 \pm 5$	0.7
BLA-C-033	0.19	$10 \pm 4$	-1.4
BLA-C-031	0.22	$19 \pm 4$	0.3
BLA-C-029	0.24	$8 \pm 3$	1.1
BLA-C-054	0.25	$3 \pm 2$	1.1
BLA-C-004	0.25	$7 \pm 4$	0.1
BLA-C-028	0.31	$17 \pm 5$	2.4
BLA-C-056	0.33	$52 \pm 6$	-2.2
UMP	0.34	$23 \pm 8$	-2.6
BLA-C-027	0.41	$48 \pm 13$	0.5
GTP	0.42	$89 \pm 48$ * $120 \pm 20$	-4.5
BLA-C-055	0.43	$51 \pm 15$	-1.6

GMP	0.47	27 ± 12	-3.5
FMN	0.51	18 ± 5	-1.4
BLA-C-034	0.65	20 ± 4	0.4

Le logP est une mesure de solubilité différentielle d'un composé chimique dans deux solvants (coefficient de partage octanol/eau). Il est égal au logarithme du rapport des concentrations de la substance quand elle se répartit entre une phase d'octanol et une phase aqueuse :  $\text{LogP} = \text{Log}(C_{\text{oct}}/C_{\text{eau}})$ . Cette valeur permet d'appréhender le caractère hydrophile ou hydrophobe (lipophile) d'une molécule. Le logP est inversement corrélé à l'hydrophobicité. Les logP de chaque molécule ont donc été calculés avec le logiciel en ligne Molinspiration (Tableau 20) [f]. La méthode de prédiction de logP développée par Molinspiration est basée sur les contributions de groupements chimiques. Les concepteurs du logiciel ont montré que le logP d'un composé peut être déterminé par la somme des logP des fragments moléculaires non-chevauchants, ces fragments moléculaires étant définis comme un ou plusieurs atomes liés covalamment l'un à l'autre au sein de la molécule.

Le paramètre  $f_{\text{sat}}$  reflétant *a priori* plusieurs composantes dépendantes du type d'interactions, une éventuelle corrélation entre l'hydrophobicité des ligands et le paramètre  $f_{\text{sat}}$  a été recherchée (Figure 90).

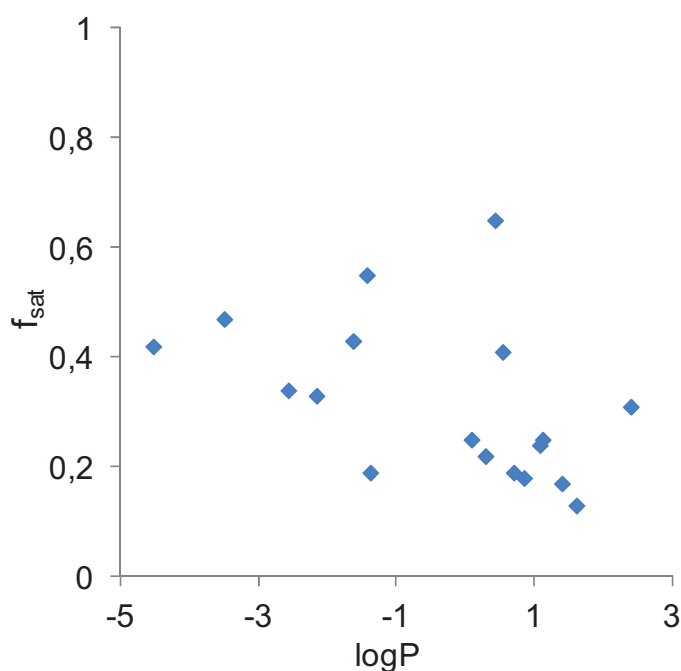


Figure 90: Relation entre le logP et le  $f_{\text{sat}}$  des meilleurs analogues de nucléotides

Une régression linéaire sur ces données conduit à un  $R^2$  de 0.12. Il n'y a donc aucune corrélation entre le  $\text{fsat}$  et le  $\log P$  (Figure 90). La relation entre l'hydrophobicité de la molécule et le comportement du complexe en phase gazeuse n'est donc pas si directe.

Le  $K_D$ , en revanche, peut avoir une relation plus directe avec  $\log P$ . Il est en effet possible que l'affinité soit liée au caractère hydrophobe du ligand, si celui-ci est engagé dans des liaisons hydrophobes avec la protéine.

Une corrélation entre le  $\log P$  et le  $K_D$  a donc également été recherchée (Figure 91).

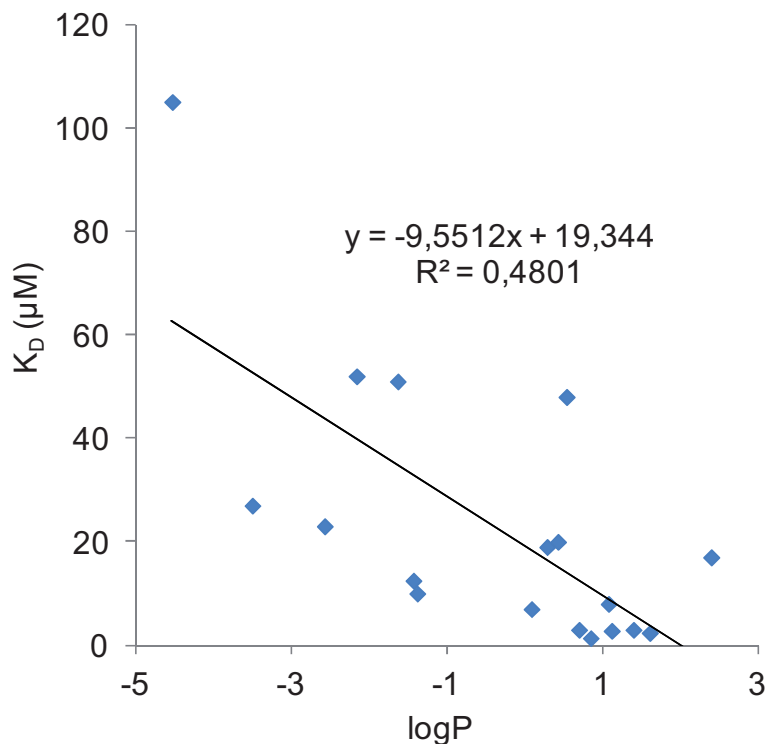


Figure 91: Relation entre le  $K_D$  et le  $\log P$  des meilleurs analogues de nucléotides

La régression linéaire donne un  $R^2$  de 0.48. Il n'y a donc qu'une très faible corrélation entre le  $\log P$  et le  $K_D$  (Figure 91). Le  $K_D$  tend à diminuer quand le  $\log P$  augmente. Autrement dit, l'affinité tend à augmenter avec l'hydrophobicité du ligand (Figure 91). Cependant l'affinité résulte probablement d'un jeu complexe d'interactions, de différents types, et de leurs relations spatiales.

Cela suggère qu'une partie de l'affinité est liée au caractère hydrophobe du ligand. Ceci est en accord avec le fait que les analogues d'UMP qui se lient à hPEBP1 sont ceux possédant un benzène en plus.

## II. Localisation du site de liaison nucléotidique

Au démarrage du projet Métasupp, les sites d'interaction des nucléotides et de la locostatine avec la PEBP n'étaient pas connus. La relation de dépendance entre les deux sites doit cependant pouvoir être déterminée par MS.

### 1. Indépendance des sites de liaison de nucléotide et de la locostatine

Afin de déterminer si les deux sites de liaison sont indépendants, la protéine a été pré-incubée avec différentes concentrations de GTP. La locostatine à une concentration finale de 1 mM a ensuite été ajoutée à l'incubation et le complexe PEBP-locostatine a été analysé par MALDI. Cette expérience a été réalisée par Emmanuelle Mebold en se basant sur les données d'affinité du GTP que j'ai obtenues (Figure 92).

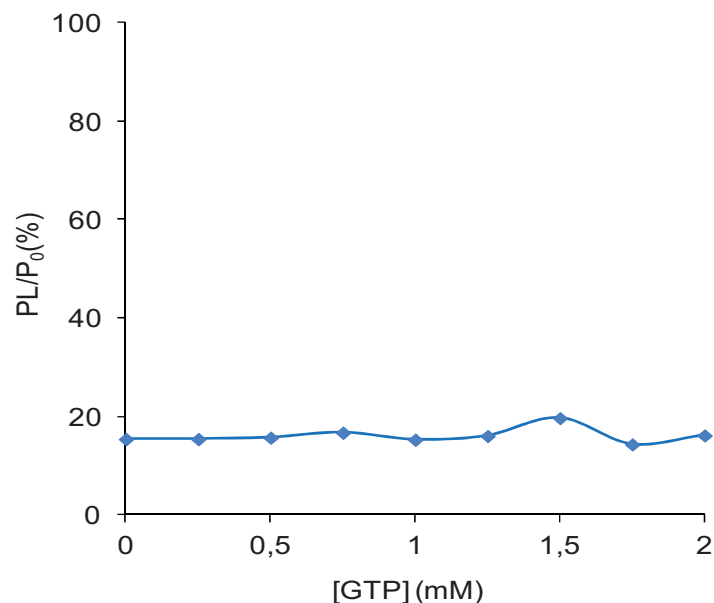


Figure 92: Fraction de liaison PEBP-locostatine obtenue en compétition avec des concentrations croissantes de GTP. hPEBP1 6.7  $\mu$ M dans HEPES-NaOH 25 mM, pH 7.7, a été pré-incubé 10 minutes à 37°C avec du GTP, puis incubée 3h à 37°C avec 1 mM final de locostatine. Les échantillons sont analysés après MGF.



Les résultats montrent que la liaison PEBP-locostatine ne varie pas quelle que soit la concentration de GTP pré-incubée avec la protéine (Figure 92). Ces résultats suggèrent que les deux sites de liaison sont indépendants.

## 2. Parallèle avec la RMN : site nucléotidique et poche à anion confondues

De la PEBP  $^{15}\text{N}$  a été produite dans le but d'étudier les complexes en RMN (195). D'après les résultats de RMN (Laurette Tavel et Christian Damblon, Département de Chimie, Université de Liège, Liège, Belgique et Laurence Jouvensal, Equipe RMN, Centre de Biophysique Moléculaire, Orléans), le GTP et le FMN se lient dans la poche à anion de hPEBP1 (195). En effet, les hydrogènes du squelette peptidique de tous les résidus de la poche à anion sont perturbés par l'ajout de FMN ou de GTP (Figure 93). Les deux molécules perturbent aussi d'autres résidus en dehors de la poche, ce qui suggère qu'une partie seulement des nucléotides se loge dans la poche (195).

La partie anionique du GTP et du FMN, le phosphate, se loge donc probablement dans cette poche. Le fait de ne pas observer d'interaction entre les ligands neutres et la PEBP n'est donc pas dû à une limite de la méthode de MS mais est une conséquence logique du mode d'interaction.

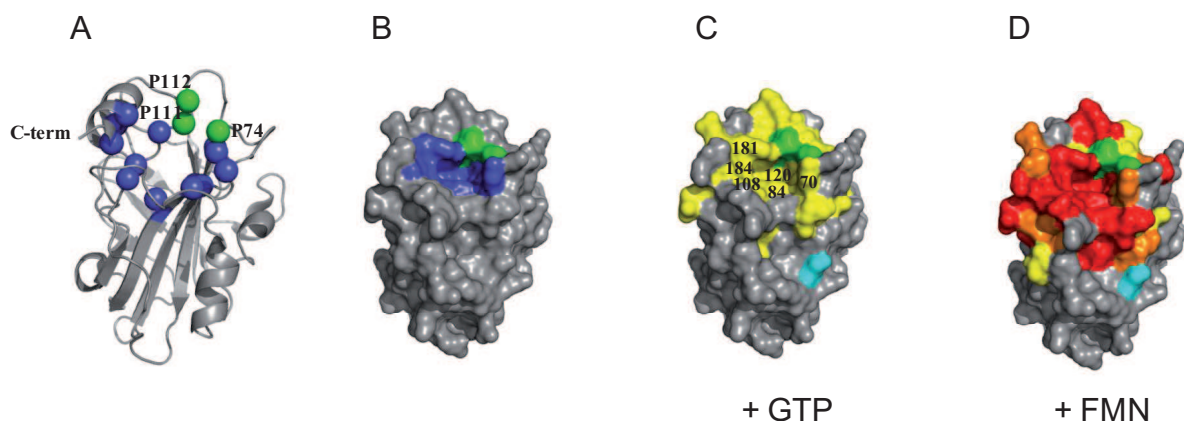


Figure 93: Poche à anion de la PEBP (PDB 2QYQ), A) sur la structure en ruban, B) sur la représentation de surface. Les résidus indiqués en bleu sont D70, A73, Y81, W84, H86, V107, G108, G110, H118, Y120, L180, Y181, and L184. Les prolines 74, 111 and 112, qui appartiennent à la poche de hPEBP, ne sont pas détectées par spectres HSQC, et sont indiquées en vert. Les perturbations chimiques des H du squelette peptidique induites par la liaison C) du GTP ou D) du FMN à pH 6.5 sont indiquées en rouge : résidus en échange lent,

en orange : résidus en échange intermédiaire, jaune : résidus en échange rapide. La sérine 153 est indiquée en bleu cyan comme un point de référence.

Les perturbations induites par la liaison du GTP sur la PEBP résultent d'échanges rapides alors que ceux induits par le FMN d'échanges lents. Ces différences suggèrent que même si les deux nucléotides se lient dans la poche à anion, leurs mécanismes d'interaction diffèrent.

Les deux méthodes de chimie analytique (MS et RMN) montrent néanmoins que les sites de liaison du GTP et du FMN sont fortement chevauchants. Les  $K_D$ s des complexes PEBP-GTP et PEBP-FMN ont pu être déterminés sur la base d'une stœchiométrie 1:1 par RMN dans deux jeux de conditions différentes, qui sont distincts de ceux utilisés en spectrométrie de masse. Les valeurs de  $K_D$  déterminées par RMN sont rassemblées dans le tableau 21 avec celles obtenues par MS à pH 7.4 et 8.4.

Tableau 21: Valeurs de  $K_D$  de nucléotides déterminées par RMN et MS.

Composé	$K_D$ ( $\mu$ M) par RMN			$K_D$ ( $\mu$ M) par MS	
	à pH 6.5/ T = 20°C MES 10 mM	à pH 7.5/ T = 20°C HEPES 10mM	à pH 7.5/ T = 30°C HEPES 10mM, NaCl 100mM	à pH 7.4/ T = 37°C ABC 20 mM	à pH 8.4/ T = 37°C ABC 20 mM
GTP	669 ± 140	n.d.	3425 ± 1967	89 ± 48	40 ± 15
FMN	14 ± 9	14 ± 11	252 ± 84	5.3 ± 4.2	35 ± 13

Note: les erreurs indiquées pour les données MS reflètent l'erreur de mesure à la fois sur  $K_D$  et sur  $f_{sat}$ .

Les valeurs de  $K_D$  déterminées par MS, avec une force ionique constante de 20 mM et une température de 37°C, montrent que l'affinité du GTP augmente lorsque le pH varie de 7.4 à 8.4 alors que celle du FMN diminue.

Une expérience a également été réalisée en RMN sur le complexe PEBP-FMN pour faire varier le pH, de 6.5 à 7.5, en gardant la température à 20°C et la force ionique constante à 10 mM. Dans ces conditions le  $K_D$  ne change pas, ce qui signifie que l'augmentation du pH de 6.5 à 7.5 n'a aucune influence sur l'interaction PEBP-FMN (Tableau 21).

Pour FMN, les valeurs de  $K_D$  déterminées en MS et en RMN, à pH physiologique avec une force ionique peu élevée, sont proches, bien que les températures soient différentes. Une variation de température de 20 à 37°C ne semble donc pas avoir de conséquence sur l'affinité de FMN pour la PEBP. La température n'ayant pas d'effet sur le  $K_D$  de PEBP-FMN, les valeurs de  $K_D$  déterminées par RMN, à pH physiologique, à deux forces ioniques différentes, 10 et 110 mM, et deux températures différentes, 20 et 30°C, montrent donc que l'augmentation du  $K_D$  est due à l'augmentation de la force ionique (Tableau 21). L'affinité du FMN pour la PEBP semble donc diminuer lorsque la force ionique augmente.

Ces déterminations de  $K_D$  de molécules de référence obtenues par une méthode RMN combinées à celles obtenues en MS, dans des conditions se rapprochant des conditions physiologiques, montrent que contrairement à rPEBP1 (101), la PEBP1 humaine est capable de lier le GTP. Une publication sur ce sujet est en cours de révision.

### 3. Effet de la force ionique

La RMN ayant montré que la force ionique influençait l'affinité d'un nucléotide pour la PEBP, cet effet a été approfondi par une étude en MS.

La stabilité en solution d'édifices supramoléculaires principalement maintenus par des interactions électrostatiques est très fortement dépendante de la force ionique du tampon utilisé. En effet, en présence de fortes concentrations de sels, les charges se trouvent davantage écrantées, réduisant ainsi la force des interactions électrostatiques.(196–198).

L'effet de l'augmentation de la force ionique a été testé sur le complexe hPEBP1-GTP dans des conditions où il n'y a pas d'agrégation visible. La fraction de liaison a ensuite été déterminée en fonction de la force ionique (Figure 94).

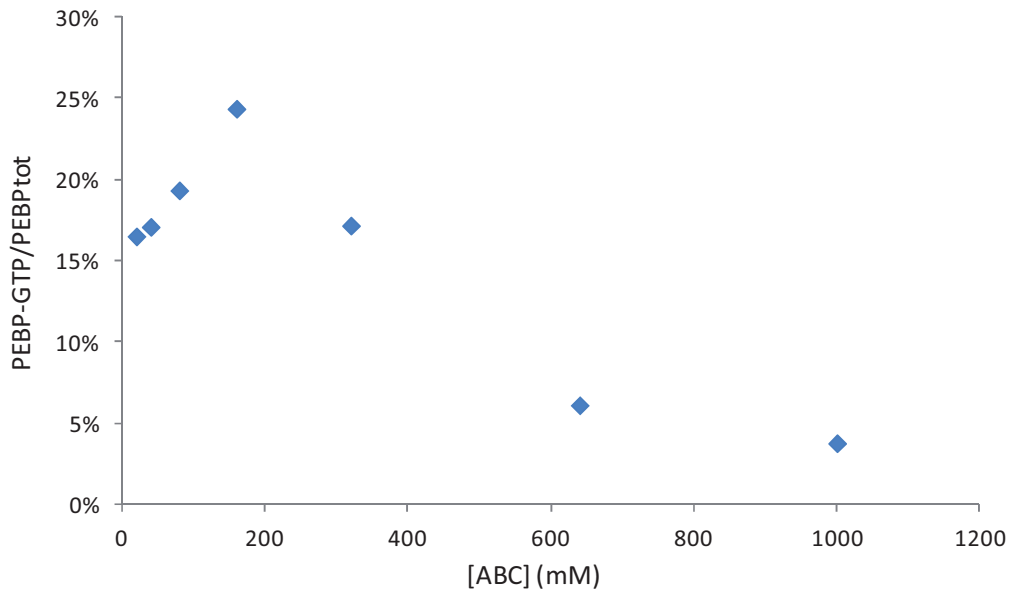


Figure 94: Effet de la concentration en ABC sur la fraction de liaison d'une incubation de 12.6  $\mu\text{M}$  de PEBP avec 25.2  $\mu\text{M}$  de GTP

La fraction de liaison est maximale pour une concentration en ABC de 160 mM, puis diminue fortement jusqu'à 1000 mM d'ABC (Figure 94). Ce résultat prouve que la force ionique affecte les interactions hPEBP1-GTP.

Afin de confirmer l'effet de la force ionique sur la liaison, le  $K_D$  a été déterminé sur le complexe PEBP-GTP à des concentrations en ABC de 20, 110, 250 et 500 mM (Figure 95). Une concentration de 110 mM se rapproche de la force ionique physiologique d'une cellule, qui est d'environ 140 mEq.

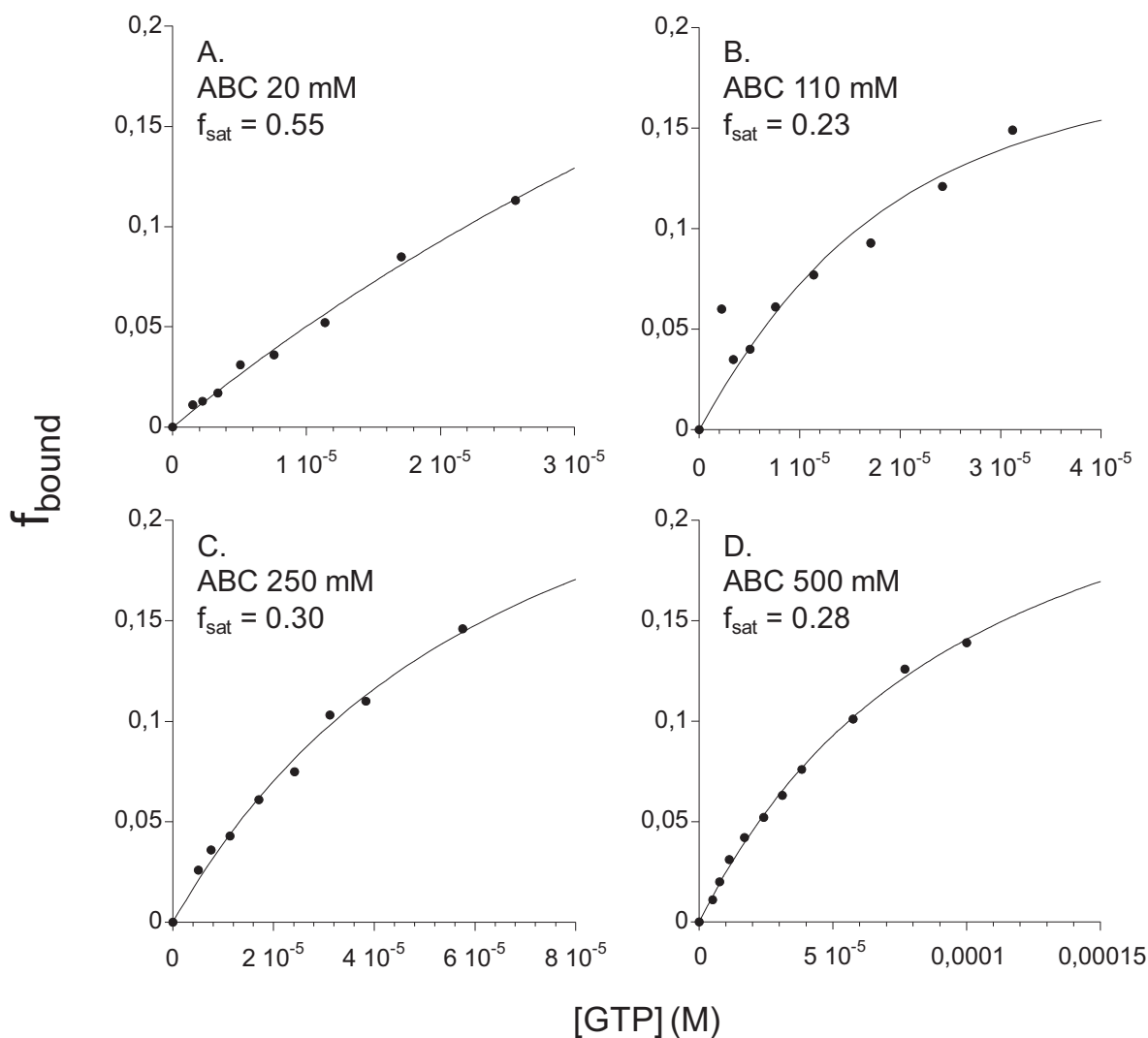


Figure 95: Fraction de liaison de PEBP mesurée par MS en fonction de la concentration en GTP. Le complexe est formé en incubant  $12.6 \mu\text{M}$  PEBP du GTP, à  $37^\circ\text{C}$ , pH 7.4 et à différentes concentrations en ABC A) 20, B) 110, C) 250 et D) 500 mM. Le  $K_D$  and le  $f_{\text{sat}}$  sont obtenus par régression non-linéaire avec l'Equation 13 de la publication Jaquillard *et al.*

Les résultats de la régression sont présentés dans le tableau 22.

Tableau 22: Effet de la concentration en ABC sur les valeurs de  $K_D$  et de  $f_{\text{sat}}$  déterminées sur le complexe PEBP-GTP à pH 7.4 et  $37^\circ\text{C}$ .

[ABC] (mM)	$K_D$ ( $\mu\text{M}$ )	$f_{\text{sat}}$
20	$89 \pm 48$	$0.55 \pm 0.23$
110	$14 \pm 6.7$	$0.23 \pm 0.04$
250	$57 \pm 13$	$0.30 \pm 0.04$
500	$92 \pm 9.0$	$0.28 \pm 0.02$

La variation du  $K_D$  en fonction de la force ionique (Tableau 22) est en accord avec la variation de la fraction de liaison en fonction de la force ionique (Figure 95). L'affinité du GTP pour la PEBP est maximale à une concentration en ABC de 110 mM. De façon étonnante, c'est aussi à cette concentration en ABC que le paramètre  $f_{\text{sat}}$  est le plus bas, ce qui peut signifier que c'est dans ces conditions que la part d'interaction hydrophobe est la plus grande. Il semble donc que l'affinité du GTP pour la PEBP et la part d'interaction hydrophobe soient corrélées. Les interactions hydrophobes forment ainsi une part importante de l'interaction PEBP-GTP.

Les changements d'affinité observés avec la variation de la force ionique sont peut-être dus à des changements de conformation de la protéine avec la force ionique. La PEBP nécessite peut-être une certaine concentration en sels pour adopter la conformation la plus propice à l'interaction avec le GTP. En d'autres termes, la PEBP semble adopter, à cette force ionique proche des conditions physiologiques, une conformation proche de la conformation idéale du point de vue fonctionnel. Par contre, une trop forte concentration en sels peut déstabiliser les interactions de type électrostatique.

Afin de voir si la concentration ionique a un effet visible sur la conformation de la PEBP, son effet sur la distribution des états de charge a été analysé sur la protéine seule (Figure 96).

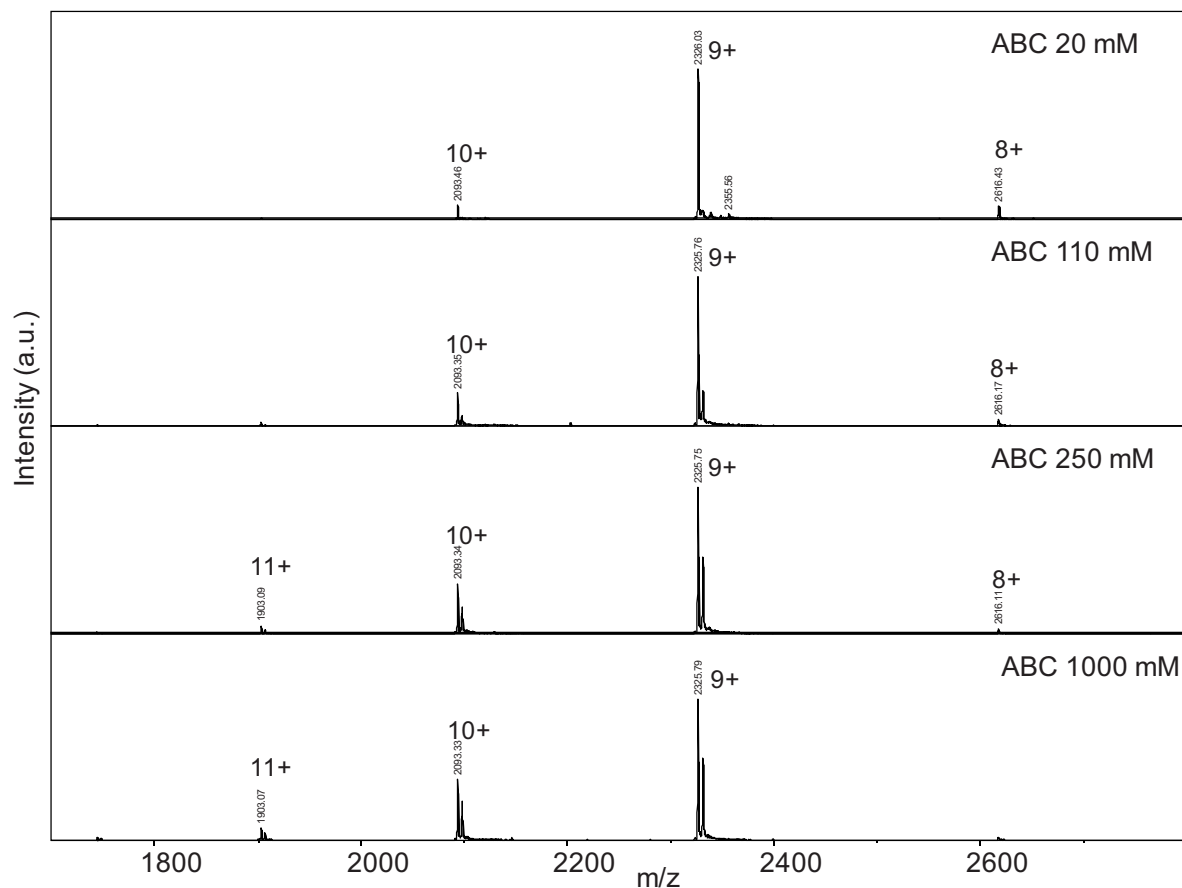


Figure 96: Spectres ESI-IT de la PEBP à 12,6  $\mu\text{M}$  dans 20, 110, 250 ou 1000 mM d'ABC et pH 7.4

L'augmentation de la concentration en bicarbonate d'ammonium entraîne une augmentation des nombres de charges et nombre d'états de charge de la PEBP et une augmentation du pourcentage d'adduits d'acide formique. A pH constant, l'apparition d'un état de charge plus élevé, 11+, suggère que de fortes concentrations en ABC conduisent à une structure de PEBP un peu moins compacte (Figure 96). La PEBP semble cependant conserver globalement sa conformation native même à 1000 mM d'ABC.

Il semble important de ne pas travailler à de fortes concentrations en ABC en routine pour éviter de surcharger l'instrument en sels et limiter les adduits d'acide formique. Par contre, il paraît intéressant de travailler avec 110 mM d'ABC au lieu de 20 mM pour se rapprocher davantage des conditions physiologiques.

Le  $K_D$  de PEBP-GTP déterminé à pH 7.4 ou 7.5 par MS dans 110 mM d'ABC ( $K_D = 14 \pm 7 \mu\text{M}$ ) n'est pas du tout en accord avec celui déterminé par RMN dans 10 mM d'HEPES avec

100 mM de NaCl ( $K_D = 3400 \pm 2000 \mu\text{M}$ ). Même si la technique de MS peut sous-estimer le  $K_D$  et que la RMN la surestime, la différence entre les deux valeurs reste très importante. Il se peut qu'à concentration ionique égale, le NaCl perturbe plus fortement les interactions électrostatiques que l'ABC. Il serait intéressant de faire ce type d'étude sur le complexe PEBP-FMN, pour lequel il a été observé en RMN un effet opposé, ç.à.d. une baisse d'affinité lors du passage d'une force ionique de 10 à 110 mEq.

### III. Comparaison des liaisons moléculaires avec les effets cellulaires

La migration cellulaire est un processus fondamental dans les mécanismes d'invasion cellulaire et de métastase. L'invasion cellulaire implique à la fois l'altération des propriétés d'adhésion des cellules et l'augmentation de leur motilité associée à des propriétés accrues de contractilité et la capacité à modifier l'environnement matriciel (199–201). Le potentiel anti-métastatique de la protéine RKIP/PEBP avec de petites molécules ayant un effet agoniste (ou activateur) de la protéine, est donc évalué *in vitro* par la mesure de leur effet sur l'invasion et la migration cellulaire par deux tests complémentaires. L'effet sur la migration cellulaire est évalué en mesurant la vitesse de réparation d'une lésion introduite expérimentalement dans une monocouche de cellules en culture (méthode du « wound-healing »). L'effet sur l'invasion est testé en chambre de Boyden en mesurant la capacité des cellules à franchir un mime de la matrice extracellulaire puis les pores d'une membrane de polycarbonate. Ces deux tests ont été réalisés sur des lignées tumorales dérivées de tumeurs humaines invasives et connues pour être très invasives dans les modèles animaux.

#### 1. Effet sur la migration cellulaire

L'analyse du pouvoir antimigratoire des analogues nucléotidiques est effectuée par le test du « wound healing » automatisé sur la plateforme de Rennes. Le test consiste à rayer des monocouches cellulaires confluentes de façon à détruire une bande de cellules. Plusieurs lavages permettent d'éliminer ensuite les cellules détachées ainsi que les débris. Puis, par la prise de photographies au microscope à intervalles programmés et réguliers, on mesure la vitesse de comblement de la déchirure par les cellules, proportionnelle à leur motilité (Figure



97). Le pouvoir antimigratoire de 8 molécules a été analysé en comparaison à celui de la locostatine.

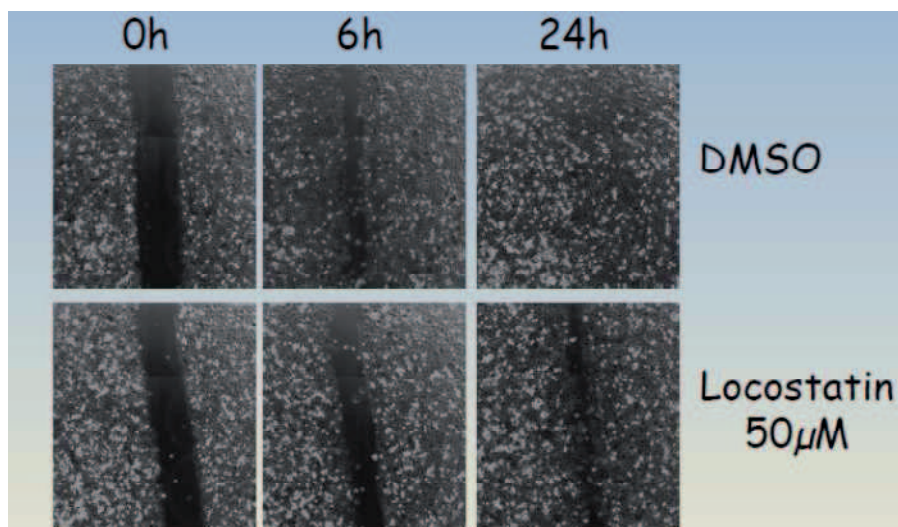
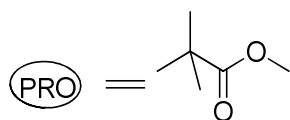


Figure 97: Effet de la locostatine par rapport à un contrôle DMSO dans le test du wound-healing sur cellules PC3 (cancer de la prostate).

Deux lignées cellulaires MDA-MB-231 (cancer du sein) et PC3 (cancer de la prostate), sélectionnées pour leur grande capacité migratoire, ont été utilisées. Il est à noter que, du fait que le GTP et le FMN sont abondants dans la cellule et impliqués dans de nombreux processus, nous n'avons pas voulu les utiliser comme molécules de référence dans ce crible. Pour ces tests, on a donc utilisé la locostatine, molécule de référence d'une autre série d'analogues synthétisés dans le projet Metasupp, comme contrôle positif d'un effet anti-migratoire. Les analogues de nucléotides se liant à hPEBP1 ont tous été testés à la même concentration (25  $\mu\text{M}$ ), concentration à laquelle la locostatine inhibe de façon significative la migration des deux types cellulaires.

Tableau 23 : Pourcentage d'inhibition de la migration des cellules en présence des différentes molécules à une concentration de 25µM.



Nom du composé forme prodrogue	Structure du composé testé	Nom du composé	Cellules PC3	Cellules MDA-MB- 231
contrôle		-	0 ± 6.1	0 ± 1.6
locostatine		-	33 ± 4.0	40 ± 5.8
BLA-C-050		BLA-C-032	-9 ± 3.1	18 ± 4.1
BLA-C-050		BLA-C-032	6 ± 5.3	15 ± 9.7
BLA-C-036		équivalent de BLA-C-032 avec 4 carbones Non- synthétisé	15 ± 4.8	16 ± 4.4

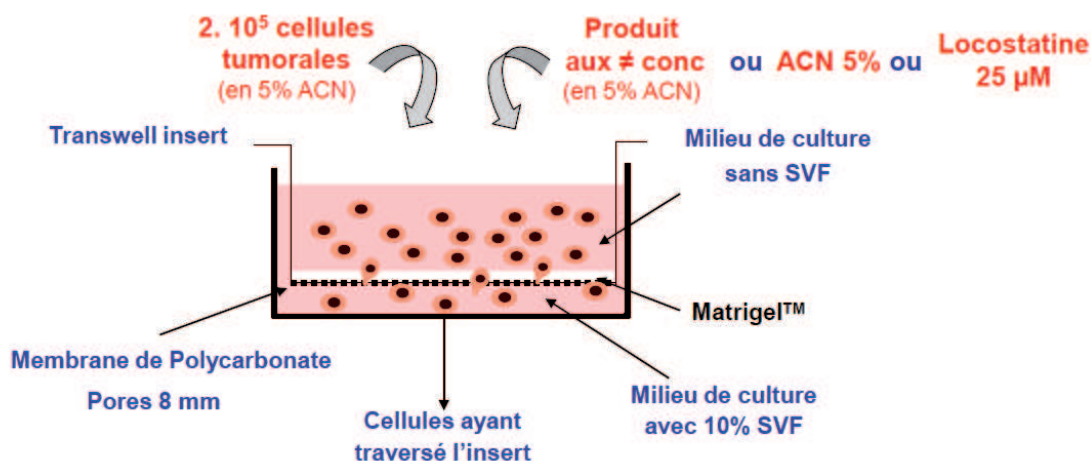
BLA-C-044		BLA-C-028	$11 \pm 3.7$	$1 \pm 1.6$
BLA-C-045		BLA-C-031	$-5 \pm 4.6$	$7 \pm 1.3$
BLA-C-046		BLA-C-025	$7 \pm 4.6$	$5 \pm 2.7$
BLA-C-047		BLA-C-027	$9 \pm 5.9$	$17 \pm 3.7$
BLA-C-048		BLA-C-034	$-8 \pm 6.8$	$12 \pm 6.9$

Les résultats de migration cellulaire pour un analogue donné ne sont pas forcément parallèles entre les lignées PC3 et MDA-MB-231 (Tableau 23). Si l'on s'intéresse prioritairement aux ligands susceptibles d'agir sur plusieurs types de cancers, on peut sélectionner d'emblée un sous-groupe de composés prodrogues pour lesquels le composé drogue se lie à PEBP. De plus, les ligands présentant le même comportement entre les deux lignées facilitent la comparaison avec les données moléculaires. BLA-C-050, BLA-C-036, BLA-C-044, BLA-C-046 et BLA-C-047 entraînent une légère inhibition de la migration cellulaire, qui reste dans tous les cas inférieure à celle induite par la locostatine. BLA-C-050 et BLA-C-036 ne diffèrent structurellement que d'un carbone au niveau leur chaîne carbonée, et leurs effets sur la migration sont proches. Les résultats sont donc cohérents. Le composé BLA-C-032 qui est la forme drogue de la prodrogue BLA-C-050 a également été testée. Alors qu'elle ne semble pas pénétrer les cellules PC3, elle semble avoir été aussi efficace que la prodrogue dans les cellules MDA.

## 2. Effet sur l'invasion cellulaire

Le principe de la chambre de Boyden est de mesurer la capacité des cellules à franchir l'obstacle constitué d'une membrane de polycarbonate dotée de pores de 8  $\mu\text{m}$  (Figure 98). Pour que l'on puisse tester les effets spécifiquement sur l'invasion et non plus seulement sur la migration cellulaire, il est capital que les cellules soient capables de progresser d'abord à travers un mime de la matrice extracellulaire. Pour cela, la membrane est prétraitée avec 0,8 mg/mL de Matrigel<sup>TM</sup> (B&D Biosciences) qui forme un film à sa surface. La chambre inférieure du dispositif contient du milieu de culture enrichi de 10% de sérum de veau foetal (SVF) pour agir comme chimio-attractant. 48 heures après ensemencement, le nombre de cellules ayant migré dans la chambre inférieure est défini (Figure 98). L'invasion cellulaire au travers du Matrigel et au travers de pores de taille inférieure à la cellule nécessite une grande plasticité des cellules et une capacité à dégrader les fibres matricielles par des protéases telles que les métalloprotéinases (202).

## Chambre de Boyden



**Comptage après 48h d'invasion**

**MDA en ACN : 40-50% d'invasion**

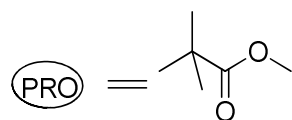
**A549 en ACN : 20-25% d'invasion**

Figure 98 : Principe de la chambre de Boyden

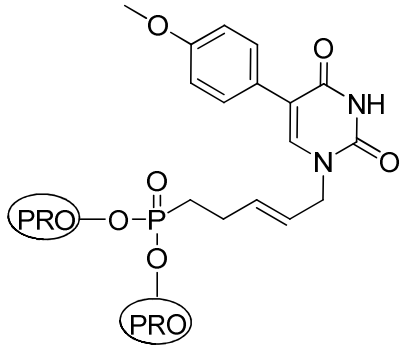
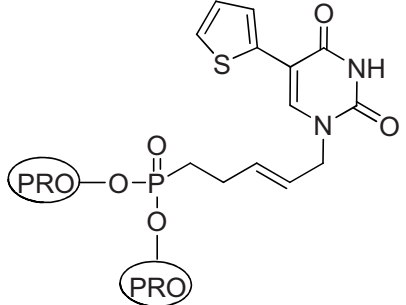
Les premiers résultats obtenus avec différentes lignées cellulaires ont montré que les propriétés invasives pouvaient varier d'une lignée à l'autre. Au total six lignées ont été testées avec insert nu et les lignées MDA-MB-231 (cancer du sein) et A549 (cancer du poumon) ont été retenues pour leur pouvoir migratoire.

Le test d'invasion cellulaire a été appliqué sur la locostatine (contrôle positif dans ce test) et sur les meilleurs analogues nucléotidiques sortis du criblage par MS (Tableau 24).

Tableau 24: Pourcentage d'inhibition maximale de l'invasion des cellules en présence des différentes molécules à la concentration maximale non-toxique.



Nom du composé forme prodrogue	Structure du composé testé	Nom du composé	Concen- -tration testée ( $\mu$ M)	Cellules A549 (poumon)	Cellules MDA- MB-231 (sein)
locostatine		-	25	48	48
testé sous forme pas prodrogue		BLA-C-032	800	n.d.	0
BLA-C-050		BLA-C-032	50	21	32
BLA-C-036		équivalent de BLA-C-032 avec 4 carbones Non- synthétisé	800	0	0

BLA-C-046		BLA-C-025	50	47	30
BLA-C-048		BLA-C-034	800	≈ 10	≈ 20

Quatre analogues de nucléotide sous forme prodrogue ont été testés pour l'invasion cellulaire, BLA-C-050, BLA-C-036, BLA-C-046 et BLA-C-048. La forme drogue de BLA-C-050, BLA-C-032, a également été testée et n'a montré aucun effet sur l'invasion des cellules MDA, alors qu'elle avait un effet anti-migratoire. BLA-C-050 présente, elle, un effet inhibiteur sur les deux types de cellules.

BLA-C-036 est la seule prodrogue qui ne présente aucun effet inhibiteur. Parmi les quatre molécules, c'est également la seule qui présente une chaîne carbonée de quatre carbones. Son homologue à 5 carbones (BLA-C-050) présente une activité inhibitrice, ce qui laisse penser qu'une chaîne de cinq carbones est indispensable pour l'activité inhibitrice de l'invasion.

BLA-C-046 est la prodrogue nucléotidique qui présente le plus fort pourcentage d'inhibition de l'invasion, cela sur les cellules A549. Le pouvoir inhibiteur de cette prodrogue à la concentration de 50  $\mu\text{M}$  est équivalent à celui de la locostatine à une concentration de 25  $\mu\text{M}$ .

BLA-C-048 montre aussi un effet anti-invasif plus léger mais ceci à des concentrations élevées de l'ordre de 800  $\mu\text{M}$  sur les cellules MDA-MB 231 et A549.

Les résultats présentés dans le tableau viennent pour la plupart d'une seule expérience. Dans un souci de fiabilité, il serait nécessaire de les répéter avant d'en tirer de plus amples conclusions.

### 3. Toxicité

Les analogues nucléotidiques ont aussi été criblés par rapport à leur capacité à induire une toxicité, c'est-à-dire leur capacité à bloquer la prolifération par cytotoxicité anti-tumorale et/ou à induire la mort de cellules normales fibroblastiques (Tableau 25). La toxicité sur cellules normales est particulièrement intéressante car elle augure de la marge thérapeutique que l'on peut espérer d'un analogue actif sur les métastases. Huit molécules ont été testées sur six lignées cellulaires représentatives des tumeurs les plus fréquentes chez l'homme et sur fibroblastes. En utilisant une gamme de concentration, un  $CC_{50}$  a été défini pour chacune d'elles. Le  $CC_{50}$  (50% Cytotoxic Concentration) correspond à la concentration en prodrogue à laquelle 50% de cellules sont tuées. Ici encore, la locostatine sert de contrôle positif dans le test. Sur l'ensemble des lignées, les analogues nucléotidiques testés n'induisent une toxicité qu'à des concentrations élevées, supérieures à celle observée pour la locostatine (Tableau 25).

Tableau 25 :  $CC_{50}$  , en  $\mu M$ , de prodrogues nucléotidiques sur différentes lignées cellulaires.

Nom du composé	Nom du composé forme prodrogue	HuH7 (foie)	Caco-2 (colon)	HCT116 (colon)	MDA-MB 231 (sein)	PC3 (prostate)	NCI (poumon)	<b>fibroblaste</b>
	locostatine	7	6	8	11	8	n.d.	<b>7</b>
BLA-C-032	-	>25	>25	>25	>25	>25	>25	<b>&gt;25</b>
BLA-C-032	BLA-C-050	43	47	40	36	39	29	<b>&gt;70</b>
BLA-C-032 *4C	BLA-C-036	>70	>70	>70	48	>50	>50	<b>&gt;50</b>
BLA-C-028	BLA-C-044	37	48	>50	41	28	27	<b>&gt;70</b>
BLA-C-031	BLA-C-045	>50	>50	47	48	>50	33	<b>&gt;70</b>
BLA-C-025	BLA-C-046	>50	44	41	42	43	34	<b>40</b>
BLA-C-027	BLA-C-047	>50	45	45	33	>50	33	<b>&gt;50</b>
BLA-C-034	BLA-C-048	>50	40	37	32	38	35	<b>&gt;50</b>

Un des buts majeurs du projet Métasupp est de trouver des molécules actives contre l'invasion par les cellules cancéreuses mais non toxiques sur les cellules saines, dans notre cas les fibroblastes. Les prodrogues BLA-C-050, 46 et 48 sont actives contre l'invasion et peu ou pas



toxiques sur cellules saines. De plus, elles sont antiprolifératives avec un large spectre vis-à-vis du type du cancer d'origine des cellules tumorales. Les molécules BLA-C-044 et BLA-C-047 n'ont pas encore été testées au niveau de l'invasion, mais elles ont des propriétés anti-migratoires et antiprolifératives avec une faible toxicité.

## 4. Comparaison et corrélation

### 4.1. Des différents effets cellulaires

On peut noter que les pourcentages d'inhibition de l'invasion sont plus forts que ceux de la migration, ce qui suggère que ces molécules n'inhibent pas seulement la migration mais doivent également nuire à la plasticité des cellules et/ou à leur capacité à dégrader les fibres matricielles par des protéases.

La locostatine est aussi toxique sur les cellules cancéreuses que sur les cellules saines. Le criblage d'analogues de nucléotides a mis en évidence des molécules qui présentent l'avantage, en plus d'avoir des propriétés anti-migratoires ou anti-invasives, d'être peu ou pas toxiques sur les cellules saines. C'est le cas de BLA-C-046 et de BLA-C-50, qui possèdent un pouvoir clair anti-invasif sur les cellules MDA-MB231 et A549.

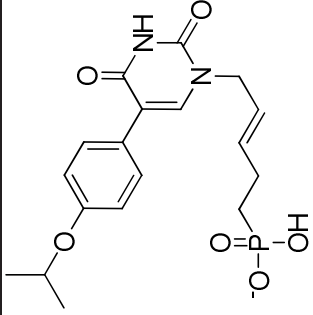
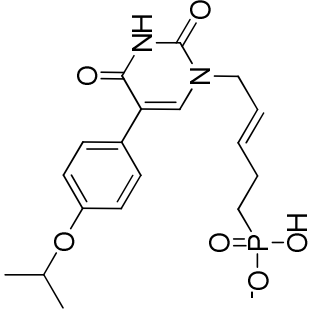
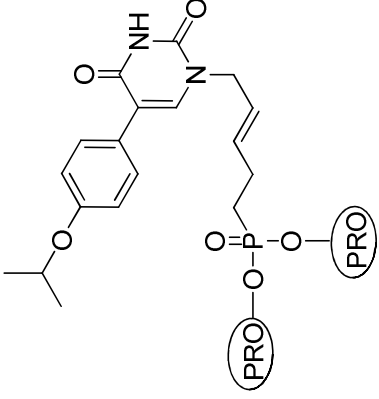
### 4.2. Des résultats cellulaires et de MS

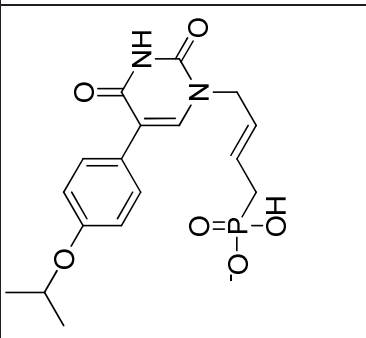
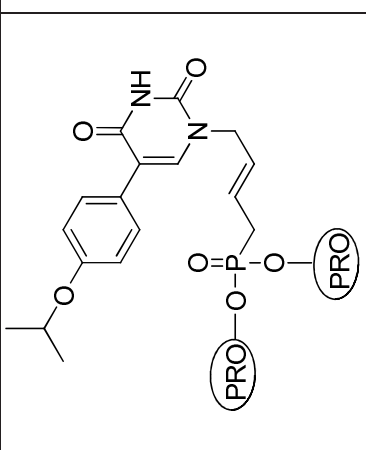
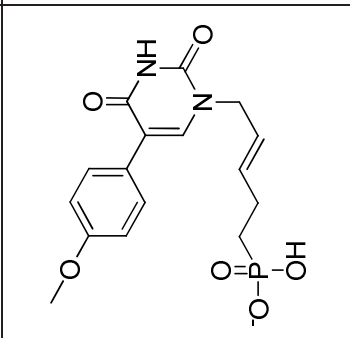
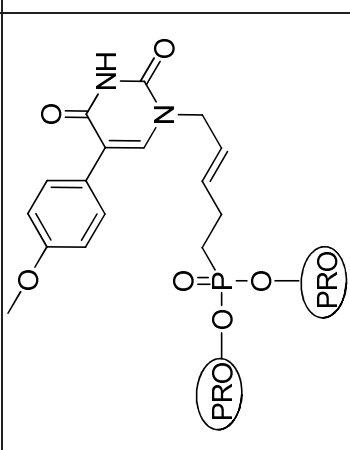
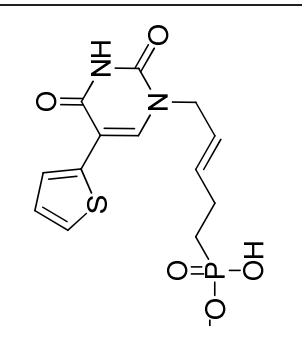
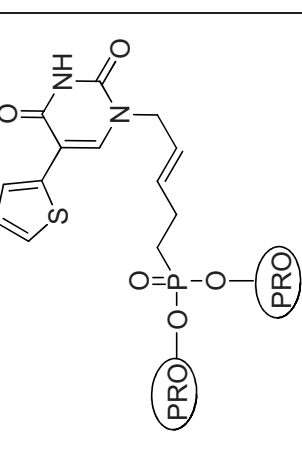
Il est difficile de tirer des conclusions entre les résultats obtenus en spectrométrie de masse et ceux obtenus en biologie cellulaire, notamment pour l'invasion cellulaire, car peu d'analogues ont été testés et les résultats proviennent pour la plupart d'une unique expérience. La comparaison a tout de même été réalisée sur les 4 ligands criblés en commun par les deux méthodes (Tableau 26).

En confrontant les valeurs d'inhibition de l'invasion à celles de  $K_D$  et de  $f_{sat}$ , sur ce petit nombre de composés, on observe une bonne corrélation entre l'activité anti-invasive et l'affinité pour hPEBP1.

Pour confirmer la corrélation entre le  $K_D$  et la capacité d'un ligand à inhiber l'invasion, il serait important de faire un contrôle négatif c.-à-d. de tester en cellulaire un ligand qui, d'après la MS, ne se lie pas à hPEBP1. La prodrogue de BLA-C-049 serait un bon contrôle.

Tableau 26 : Résultats obtenus en spectrométrie de masse ( $K_D$ ,  $f_{sat}$ ) sur les drogues avec ceux d'inhibition de l'invasion cellulaire obtenues sur les prodrogues correspondantes.

Nom du composé	Structure du composé	$K_D$ ( $\mu\text{M}$ )	$f_{sat}$	Nom du composé forme prodrogue	Structure de la prodrogue	Concentration testée ( $\mu\text{M}$ )	% inhibition invasion Cellules A549 (poumon)	% inhibition invasion Cellules MDA-MB-231 (sein)
BLA-C-032		3	0.17	testé sous forme pas prodrogue	testé sous forme pas prodrogue	800	n.d.	0
BLA-C-032		3	0.17	BLA-C-050		50	21	32

équivalent de BLA-C-032 avec 4 carbones Non-synthétisé		n.d.	n.d.	BLA-C-036		800	0	0
BLA-C-025		3	0.19	BLA-C-046		50	47	30
BLA-C-034		20	0.65	BLA-C-048		800	≈ 10	≈ 20

### 4.3. Bilan

Au départ de ce projet, on n'anticipait pas que des analogues de nucléotides puissent présenter des propriétés anti-métastatiques. Cependant, les composés BLA-C-046 et BLA-C-50 inhibent l'invasion des cellules tumorales mammaires et pulmonaires, et BLA-C-50 présente l'avantage d'être non toxique sur les cellules saines.



# Chapitre V

Recherche d'une activité enzymatique de  
hPEBP1/RKIP

### *Résumé du chapitre V:*

*Sur la base de nos observations (Chapitre IV) et de données de la littérature, nous avons recherché une activité enzymatique chez hPEBP1, pour laquelle une telle activité n'a jamais été décrite.*

*hPEBP1 est oxydée en présence de FMN, et le phénomène est fortement accentué en présence de GTP. Dans une première approche pour tenter d'expliquer ce phénomène, une activité oxydase FMN-dépendante a été recherchée. Les peptides oxydés de hPEBP1 ne coïncident pas avec les résidus de la PEBP en liaison avec FMN, ce qui met en doute la spécificité de la réaction d'oxydation.*

*Pour évaluer la spécificité de la réaction d'oxydation, trois peptides ont été mis en contact avec le FMN +/- GTP en absence de hPEBP1. Les peptides ont été oxydés malgré l'absence de protéine. Le phénomène observé ne dépend donc pas d'une activité oxydase de hPEBP1, qui n'a ainsi pas pu être mise en évidence. Le rôle du GTP dans une oxydation accrue des peptides en présence de FMN (indépendante de hPEBP1) reste à expliquer.*

*hPEBP1 se liant à un inhibiteur de PDE-5, une activité phosphodiesterase (PDE) a été recherchée. Dans nos conditions expérimentales, une activité PDE sur le GMPc ou l'AMPc n'a pas pu être mise en évidence pour hPEBP1.*

*Lors du criblage, des déphosphorylations ont été observées sur certains nucléotides, la PEBP a donc été testée pour une activité de phosphatase. La MS montre qu'elle n'est ni une GTPase, ni une ATPase. Au vu de la variété des conditions nécessaires à l'activité des différentes classes d'enzymes capables de se lier aux ligands nucléotidiques oxydo-réducteurs ou cycliques de hPEBP1, la recherche d'une activité enzymatique ne peut cependant pas être exhaustive.*

Une particularité de la protéine hPEBP1 par rapport à de nombreuses cibles classiques en cancérologie est qu'elle ne semble posséder aucune activité enzymatique. Partant d'une observation sur le complexe hPEBP1-FMN détaillée plus loin, nous avons souhaité approfondir la relation structure-fonction de la PEBP humaine non seulement en caractérisant le site de liaison nucléotidique de hPEBP1, mais également en recherchant une éventuelle activité enzymatique. De plus, hPEBP1 intervenant dans de nombreux phénomènes biologiques, la découverte d'une activité enzymatique pourrait contribuer à éclaircir son rôle général dans la cellule.

- Avant la mise au point des protocoles de criblage, des expériences très préliminaires de compétition entre FMN et GTP pour la liaison à hPEBP1 ont mis en évidence un phénomène d'oxydation de la protéine en présence de FMN, fortement amplifié par l'addition de GTP. Ce résultat nous a amené à rechercher une éventuelle activité oxydase de la PEBP dépendante du FMN.
- La liaison d'un inhibiteur de phosphodiesterase 5 (PDE5) à rPEBP1, découverte par Heck et ses collaborateurs, nous a conduits à tester une éventuelle activité PDE-5 de hPEBP1 (101).
- Des pertes apparentes de 80 Da ont été observées lors de l'étude de complexes hPEBP1-nucléotides. Celles-ci pourraient correspondre à des déphosphorylations. La présence d'une activité phosphatase chez hPEBP1, de type GTPase ou ATPase, a donc été testée.

## I. Activité oxydase

### 1. Observation de l'oxydation de hPEBP1 en présence de FMN

Lors de l'étude des complexes PEBP-FMN et PEBP-GTP par MS, une légère oxydation de la PEBP a été observée en présence de FMN. Une expérience de compétition entre FMN et GTP, dans des conditions d'agrégation non-maîtrisée, a également mis en évidence que ce phénomène d'oxydation était amplifié par l'addition de GTP.

La Figure 99 présente les spectres issus d'incubations de hPEBP1 avec le GTP, le FMN ou les 2 nucléotides à 4°C durant 5 minutes dans ABC 20 mM, pH 8.4. On observe l'apparition d'espèces avec des delta-masse multiples de 16 Da par rapport à hPEBP1. Ces delta-masse



sont typiques d'additions d'oxygène, c.à.d. d'une oxydation de la protéine. Le phénomène d'oxydation est beaucoup moins fort à 4°C qu'à 37°C, mais il est plus facile d'interpréter les spectres effectués sur les incubations à 4°C.

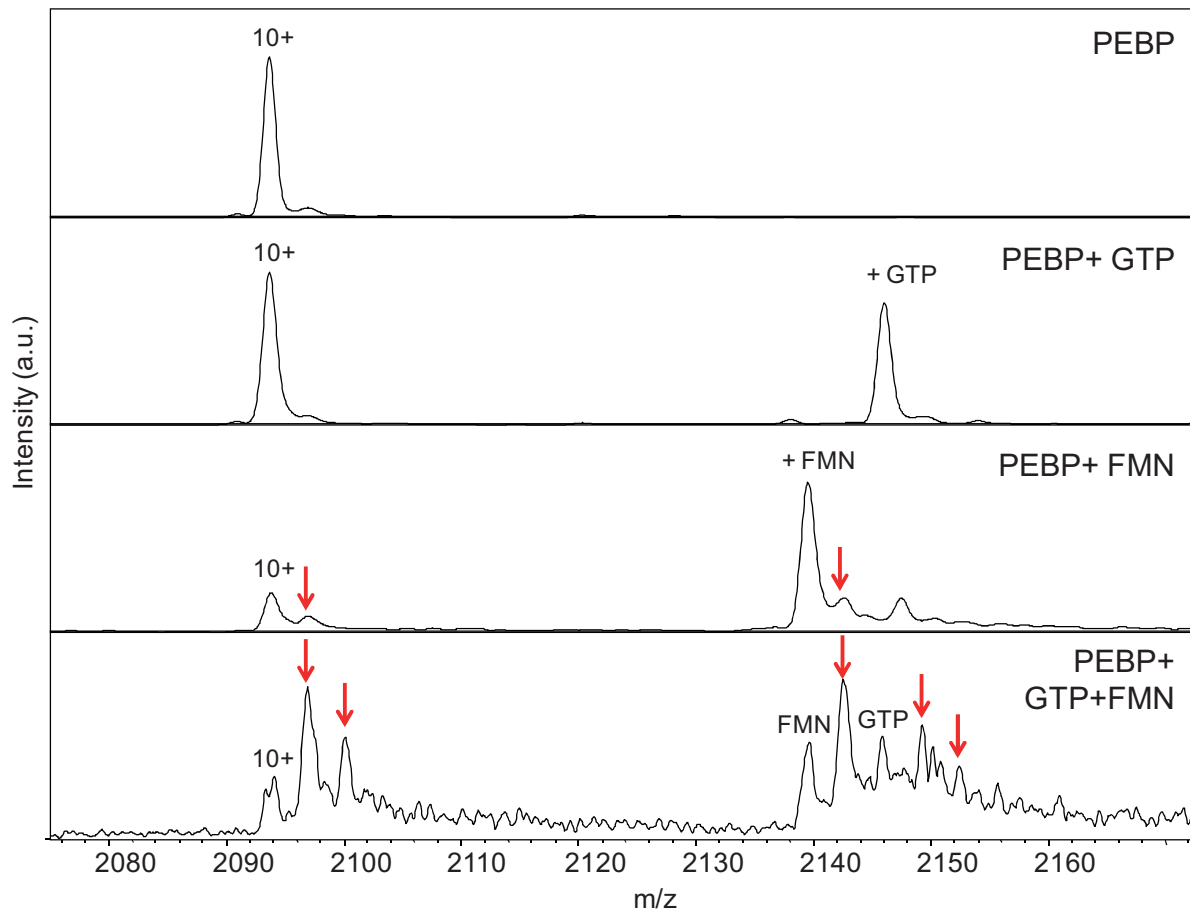


Figure 99 : Spectres ESI-IT de PEBP 18  $\mu$ M dans l'ABC 20 mM, pH 8.4, seule ou en présence de 500  $\mu$ M de GTP, avec 500  $\mu$ M de FMN ou les 2, incubée 5 minutes à 4°C. Seul l'état de charge 10 + est montré, les flèches rouges indiquent une oxydation entraînant un delta masse de 16 Da.

Cette observation conduit à penser que le FMN, co-substrat naturel de réactions d'oxydo-réduction, est responsable d'une oxydation de la protéine. L'intérêt de cette observation est qu'elle suggère que la PEBP elle-même utilise le FMN comme co-substrat. Nous avons donc fait l'hypothèse que hPEBP1 possédait une activité enzymatique monooxygénase, qui n'avait encore jamais été décrite, les PEBPs n'ayant pas d'activité catalytique connue.

## 2. Hypothèse sur le mécanisme d'oxygénation

### 2.1. Monooxygénase à flavine

La famille des monooxygénases à flavine consiste en un groupe d'enzymes qui catalyse des réactions chimiques *via* une flavine liée, qu'on appelle également groupement prosthétique ou cofacteur flavine (203–205). Les flavines sont des molécules dont la structure est basée sur un composé hétérocyclique à trois cycles, l'isoalloxazine. Les réactions de monooxygénation impliquent une oxydation des hétéroatomes nucléophiles (203–205).

L'oxygénation observée sur hPEBP1 pourrait être produite par une activité enzymatique monooxygénase avec le cofacteur FMN:

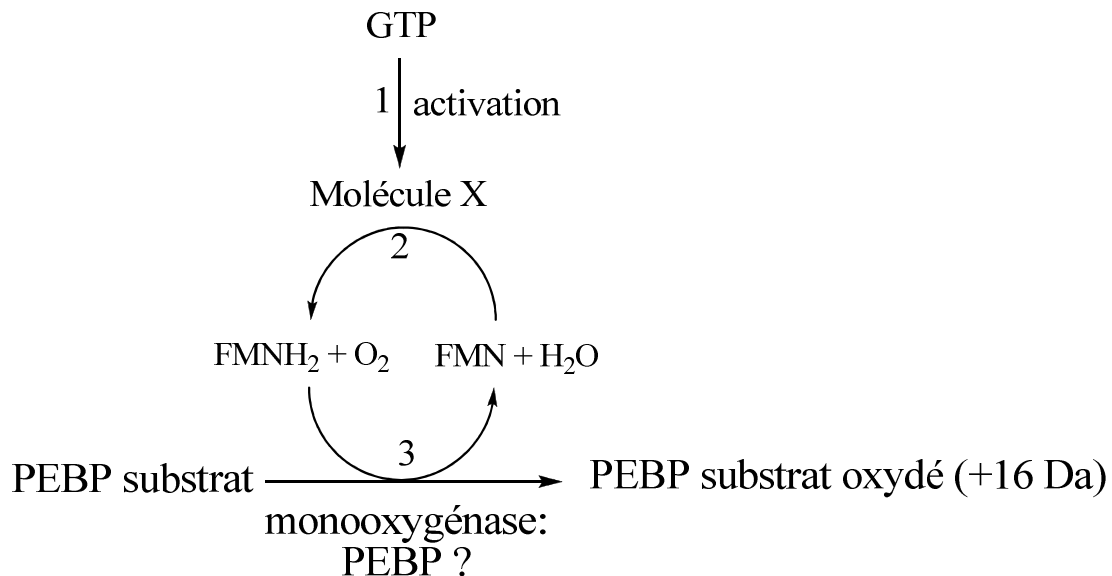


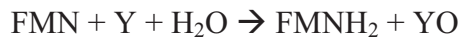
Schéma 2 : Mécanisme hypothétique d'oxydation de la PEBP en présence de FMN, fortement amplifié par l'addition de GTP. (1) Activation d'une molécule X par le GTP. (2) Réduction du FMN par la molécule X activée. (3) Ajout d'un seul atome du dioxygène sur la PEBP substrat, catalysée par la PEBP monooxygénase et son cosubstrat FMN réduit.

D'après le mécanisme hypothétique proposé (Schéma 2), hPEBP1 s'autooxyderait. L'oxygénation nécessiterait un cofacteur FMN sous forme réduite, FMNH<sub>2</sub>. Le nucléotide FMN acheté chez Sigma est principalement sous forme oxydé. Ainsi, en présence uniquement de FMN, l'oxygénation serait très modérée car le cosubstrat sous forme réduite serait limitant. Par contre, en présence de GTP, la réduction de FMN serait accélérée et entraînerait une forte oxydation de la PEBP. Il existe des mécanismes de catalyse de l'oxydation du NADPH en

NADP<sup>+</sup> avec réduction du FMN, par contre le GTP n'est pas a priori connu pour pouvoir remplir ce rôle. Il faudrait donc faire l'hypothèse d'une molécule X présente dans l'incubation et capable, en présence de GTP, de réduire le FMN.

## 2.2. Oxydation par échange direct avec une molécule substrat

Alternativement, l'oxydation pourrait procéder selon le schéma hypothétique suivant :



où Y est une molécule cible de l'oxydation.

On peut noter qu'il s'agit de la réaction inverse de la triméthylamine-N-oxide réductase bactérienne sur la triméthylamine N-oxide (206):



## 3. Carte peptidique de l'oxydation de hPEBP1 en présence de FMN

Pour aider à l'élucidation du mécanisme d'oxydation en présence de FMN, et pour déterminer s'il découle réellement d'une activité catalytique de la protéine, nous avons d'abord cherché à identifier les résidus touchés par l'oxydation et à déterminer s'ils étaient localisés dans une zone particulière, proche du site de liaison du FMN.

Pour connaître les résidus oxydés, des cartes peptidiques des échantillons ont été réalisées. Ces cartes sont réalisées par protéolyse spécifique à la fois sur la protéine intacte et sur sa forme oxydée. En obtenant deux cartes avec deux protéinases différentes, la trypsine et l'endoprotéinase-Glu-C, on s'assure d'avoir une couverture de séquence maximale. La trypsine coupe en C-terminal des résidus lysine ou arginine, et l'endoprotéinase-Glu-C (V8) en C-terminal des acides glutamiques ou aspartiques. La carte peptidique finale montre les régions de la protéine qui portent les oxydations (Figure 100). Les cartes correspondant aux expériences de protéolyse prises individuellement sont présentées dans les annexes.

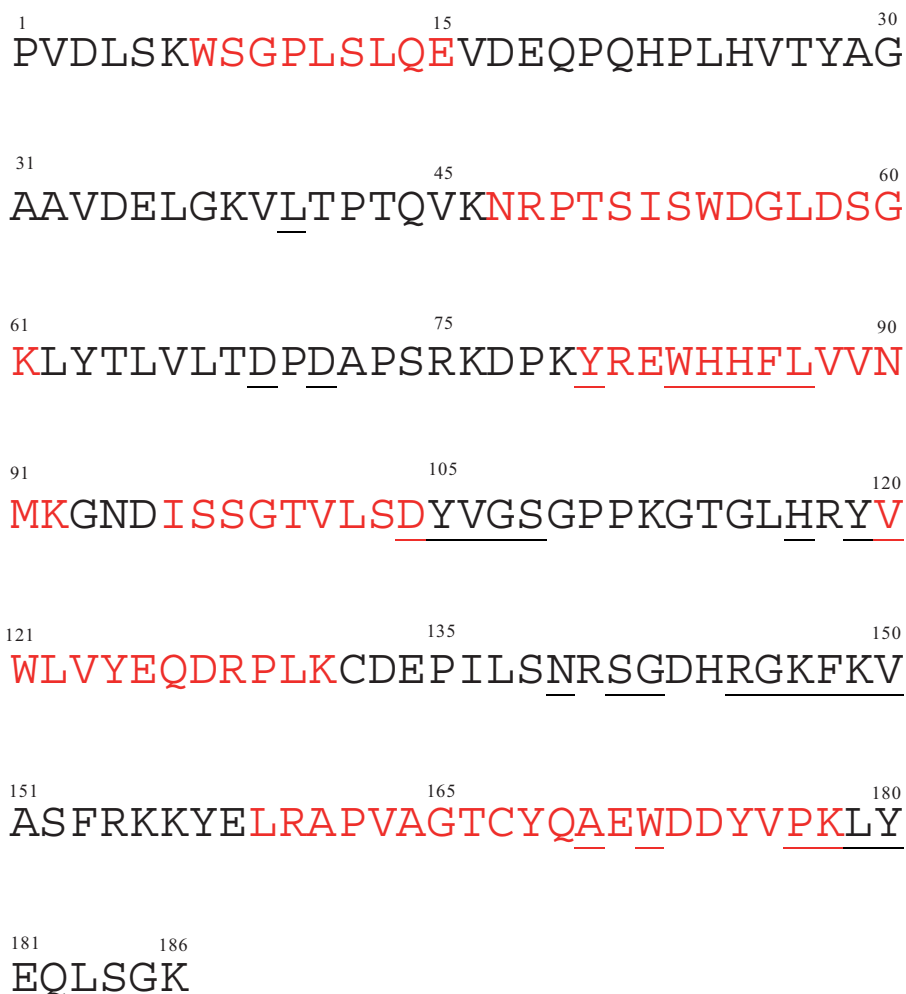


Figure 100: Carte des peptides oxydés (en rouge) de la PEBP en présence de FMN. Les résidus de la PEBP, en affinité forte avec le FMN à pH 7.5 d'après la RMN, sont soulignés.

Contrairement à ce que l'on pourrait attendre d'une réaction localisée à proximité d'un substrat ou d'un cosubstrat tel que le FMN dans un site enzymatique, l'oxydation touche plusieurs peptides, répartis sur l'ensemble de la séquence et non pas colocalisés avec les résidus en interaction avec FMN. Cette observation a soulevé la question de la spécificité de la réaction d'oxydation de hPEBP1 en présence de FMN. D'autre part, il est intéressant de noter que 6 des 7 peptides oxydés contiennent un tryptophane dans leur séquence (Figure 100). L'oxydation de certains résidus comme les tryptophanes pourrait être provoquée soit par le FMN par lui-même, soit par des produits peroxydes contaminants du réactif. La spécificité de la réaction d'oxygénation a donc été étudiée.

#### 4. Spécificité de la réaction d'oxygénation en présence de FMN

Afin de déterminer si le phénomène d'oxydation observé en présence de FMN, et fortement amplifié par l'addition de GTP, est spécifique à la PEBP et si celle-ci possède une activité monooxygénase FMN-dépendante, l'oxydation de la PEBP et d'un mélange de peptides (« pepmix ») a été étudiée en présence de FMN et/ou de GTP. Trois peptides, la bombésine (1619 m/z), l'adrénocorticotrophine (ACTH) clip 1-17 (2093 m/z) et la somatostatine (3147 m/z) sont particulièrement intéressants dans le mélange pepmix car ils contiennent tous les trois un résidu tryptophane.

La Figure 101 présente les résultats obtenus sur la somatostatine du Pepmix.

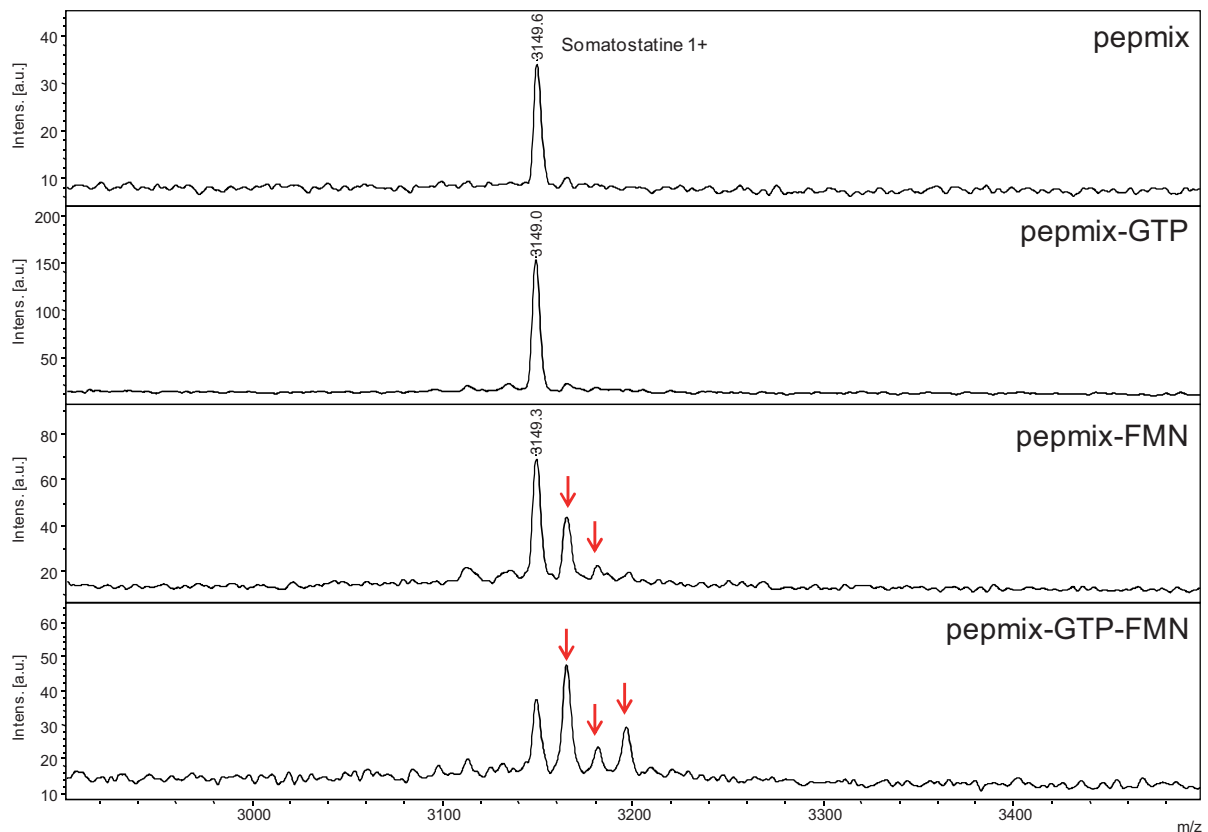


Figure 101 : Spectres MALDI-TOF de Pepmix 6.7  $\mu$ M dans l'ABC 20 mM, pH 8.4, seule, avec 200  $\mu$ M de GTP, avec 200  $\mu$ M de FMN ou les 2, incubée 3 heures à 37°C. L'échantillon analysé est dilué au 1/100<sup>ème</sup> dans 4HCCA/FWI La gamme de masse est centrée sur la somatostatine. Les flèches rouges indiquent une oxydation entraînant un delta masse de 16 Da.

Au bout de 3 heures d'incubation à 37°C, la somatostatine seule ou en présence de GTP n'est pas oxydée. Par contre l'ajout de FMN induit de l'oxydation, et le phénomène est accentué en présence de GTP.

Les effets du FMN ont également été étudiés sur le peptide [1-17] de l'adrénocorticotrophine (ACTH) (Figure 102).

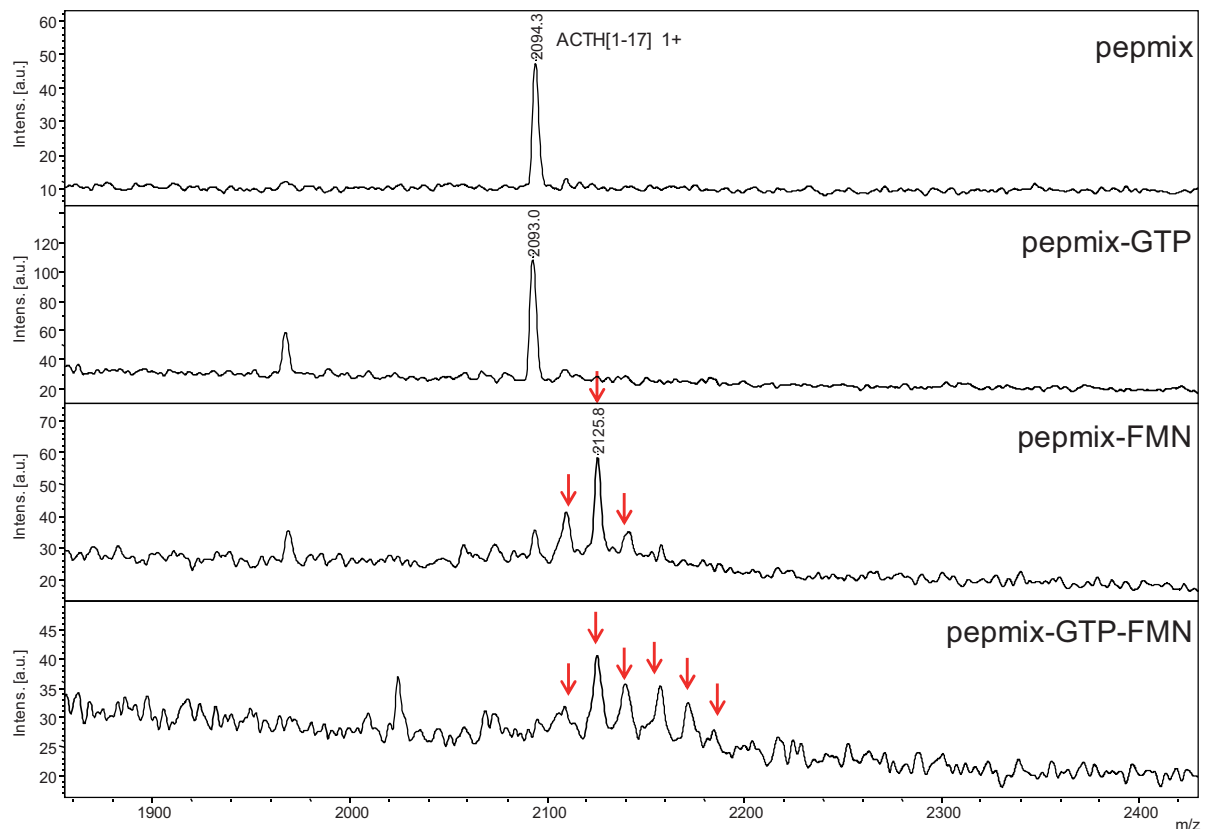


Figure 102 : Spectres MALDI-TOF de Pepmix 6.7  $\mu\text{M}$  dans l'ABC 20 mM, pH 8.4, seule, avec 200  $\mu\text{M}$  de GTP, avec 200  $\mu\text{M}$  de FMN ou les 2, incubée 3 heures à 37°C. L'échantillon analysé est dilué au 1/100<sup>ème</sup> dans 4HCCA/FWI. La gamme de masse est centrée sur l'ACTH clip 1-17. Les flèches rouges indiquent une oxydation entraînant un delta mass de 16 Da.

L'oxydation induite par le FMN est encore plus forte sur l'ACTH clip 1-17. La forme non-oxydée n'est presque pas visible.

Les peptides du pepmix sont oxydés au même degré que hPEBP1 en présence de FMN ou FMN/GTP. La hPEBP1 ne possède donc pas d'activité catalytique de l'oxydation. Le FMN semble être un oxydant naturel, qui ne nécessite pas de monooxygénase.

## 5. Photo-oxydation en présence de FMN

D'après la littérature, les flavines (riboflavine, FMN), sont photoactivables avec passage de l'état singulet de base à un état triplet activé sous l'action de la lumière. Ces propriétés photodynamiques sont d'ailleurs impliquées dans des phénomènes de photosensibilisation d'origine accidentelle ou génétique. La photoactivation de la flavine provoque la formation d'oxygène singulet ( $^1\text{O}_2$ ) à partir de l'état triplet de base de l'oxygène moléculaire. L'oxygène singulet est une espèce très réactive vis-à-vis des biomolécules et peut provoquer leur dégradation. Le mécanisme de photodégradation par oxydation est décrit sur le Schéma 3 (207).

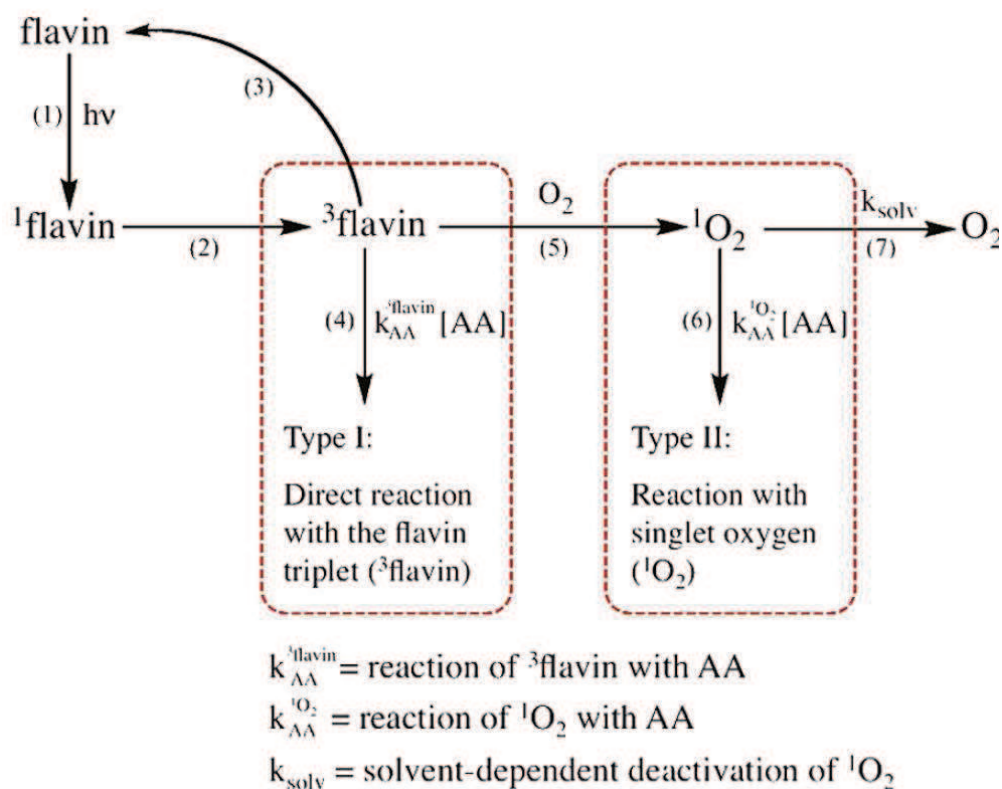


Schéma 3 : Voie de photodégradation des acides aminés médiée par les flavines. Reproduit de la publication de Remucal et McNeill (207)

Les flavines sont excitées par la lumière (réaction (1)), puis réagissent directement avec des substrats (type I), ou produisent l'oxygène singulet réactif vis-à-vis par exemple d'acides

aminés (type II) (Schéma 3). Les dérivés réactifs de l'oxygène qui sont formés par la riboflavine photoactivée peuvent induire l'oxydation des acides aminés d'une protéine (207).

D'après la littérature, 5 acides aminés sont particulièrement sensibles à l'oxydation induite par photo-activation : tyrosine, tryptophane, méthionine, histidine et cystéine (208–210).

Ces propriétés oxydantes de FMN peuvent expliquer les phénomènes d'oxydation observés précédemment. Il reste à expliquer pourquoi l'oxydation est amplifiée en présence de GTP.

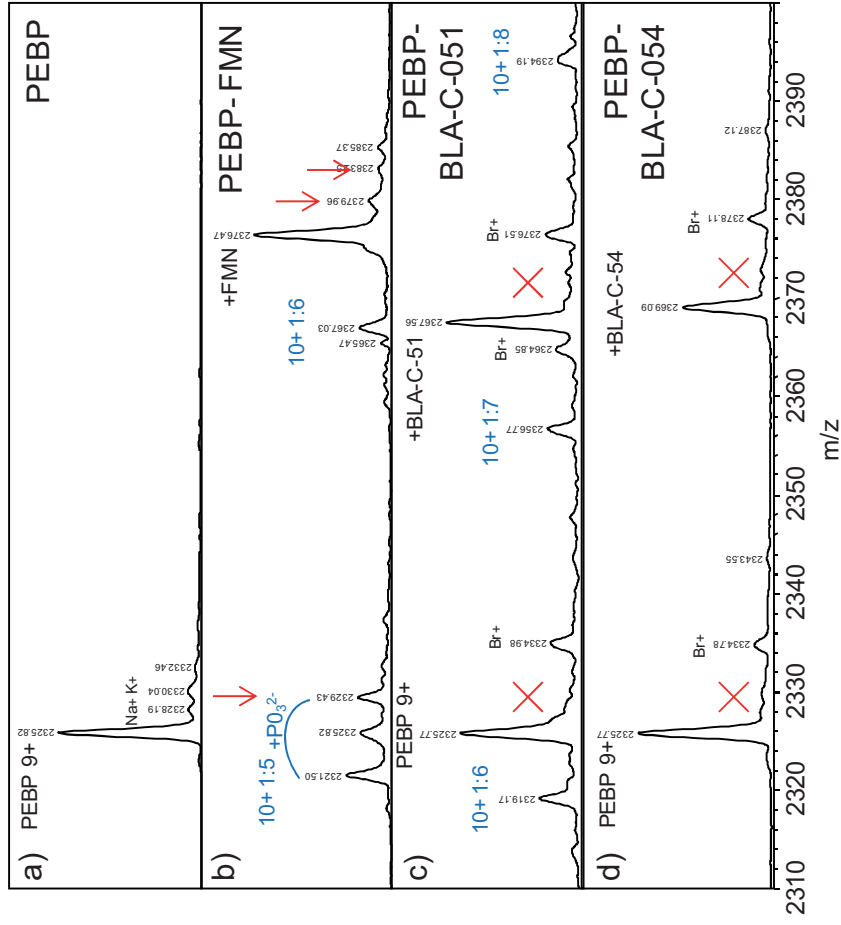
## 6. Absence d'oxydation avec les analogues de FMN

Le criblage d'analogues de FMN a mis en évidence que deux molécules, BLA-C-051 et BLA-C-054, se lient à la PEBP avec une bonne affinité ( $K_D$  de 1-3  $\mu$ M). Ces molécules ont la particularité de porter la chaîne carbonée sur l'uracile et non sur le cycle central. La conséquence majeure de ce déplacement est une perte potentielle des propriétés oxydo-réductrices de l'analogue par diminution de la délocalisation des électrons. Le potentiel oxydo-réducteur du FMN étant associé à de nombreux phénomènes physiologique d'une part, et à une photo-oxydation accidentelle capable d'endommager les cellules d'autre part, il est souhaitable de synthétiser des analogues de nucléotide qui n'induisent pas d'oxydation cellulaire. Le pouvoir oxydant des deux analogues de FMN a donc été testé sur la PEBP (Figure 103). Les expériences ont été réalisées sur les deux analogues de FMN et le FMN lui-même. L'amplification de l'oxydation en présence de GTP a également été testée

La Figure 103 met en évidence que les deux analogues n'induisent pas d'oxydation de la PEBP, que ce soit seul ou en présence de GTP. Ils présentent ainsi une caractéristique très intéressante pour des molécules thérapeutiques antimétastatiques potentielles.



## 1) SANS GTP



## 2) AVEC GTP

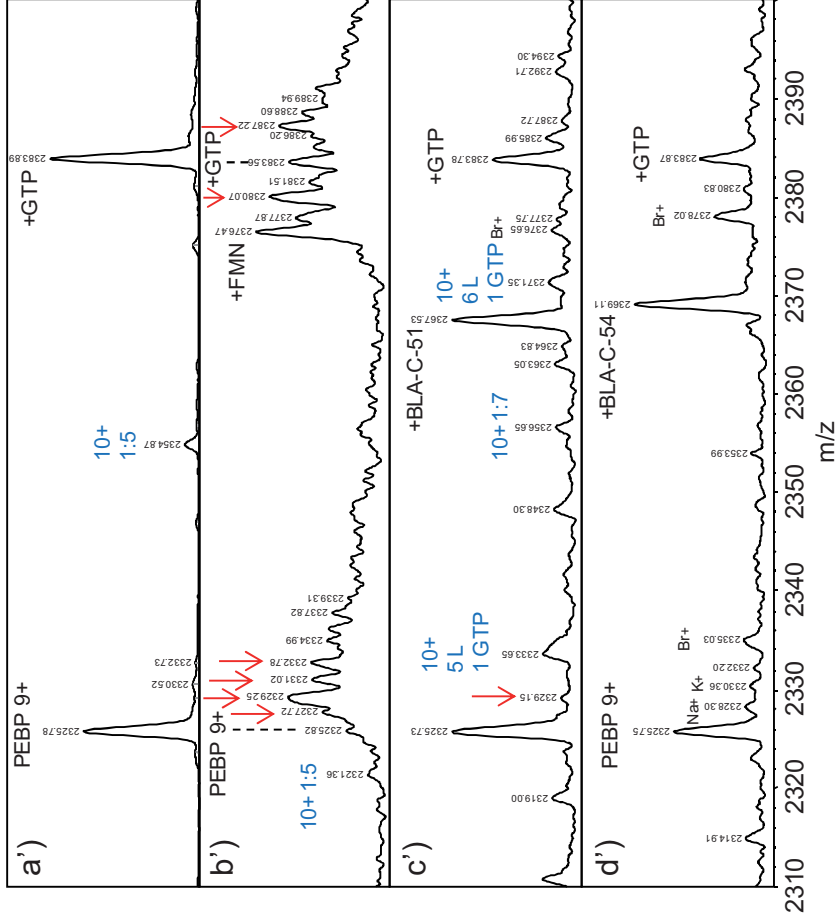


Figure 103: Spectres ESI-IT de PEBP 12.6  $\mu\text{M}$  dans l'ABC 20 mM, pH 7.4, 1) SANS GTP a) PEBP seule, b) avec 400  $\mu\text{M}$  de FMN, c) avec 400  $\mu\text{M}$  de BLA-C-051, d) avec 400  $\mu\text{M}$  de BLA-C-054 ou 2) AVEC GTP a') uniquement, b') plus 400  $\mu\text{M}$  de FMN, c') plus 400  $\mu\text{M}$  de BLA-C-051, d') plus 400  $\mu\text{M}$  de BLA-C-054. Incubation 1 heure à 37°C. Seul l'état de charge 9+ est montré. Des pics correspondant à des stoechiométries élevées de l'état de charge 10+ sont néanmoins visibles et annotés en bleu. Les flèches rouges indiquent une oxydation entraînant un delta mass de 16 Da. Les croix rouges indiquent une absence d'oxydation.

## II. Activité phosphodiesterase-5

### 1. Contexte

D'après une publication de Heck et ses collaborateurs, un inhibiteur de la phosphodiesterase-5 (PDE-5) est capable de se lier à la PEBP (9797, 98). Les cibles préférentielles de cet inhibiteur étant les PDE-5, la PEBP pourrait présenter une activité PDE-5. L'activité PDE-5 consiste à hydrolyser le GMPc en GMP (Figure 104). Le GMPc joue entre autre un rôle prépondérant dans la fonction érectile et les inhibiteurs de PDE-5 sont ainsi classiquement utilisés comme agents thérapeutiques pour les troubles de cette fonction.

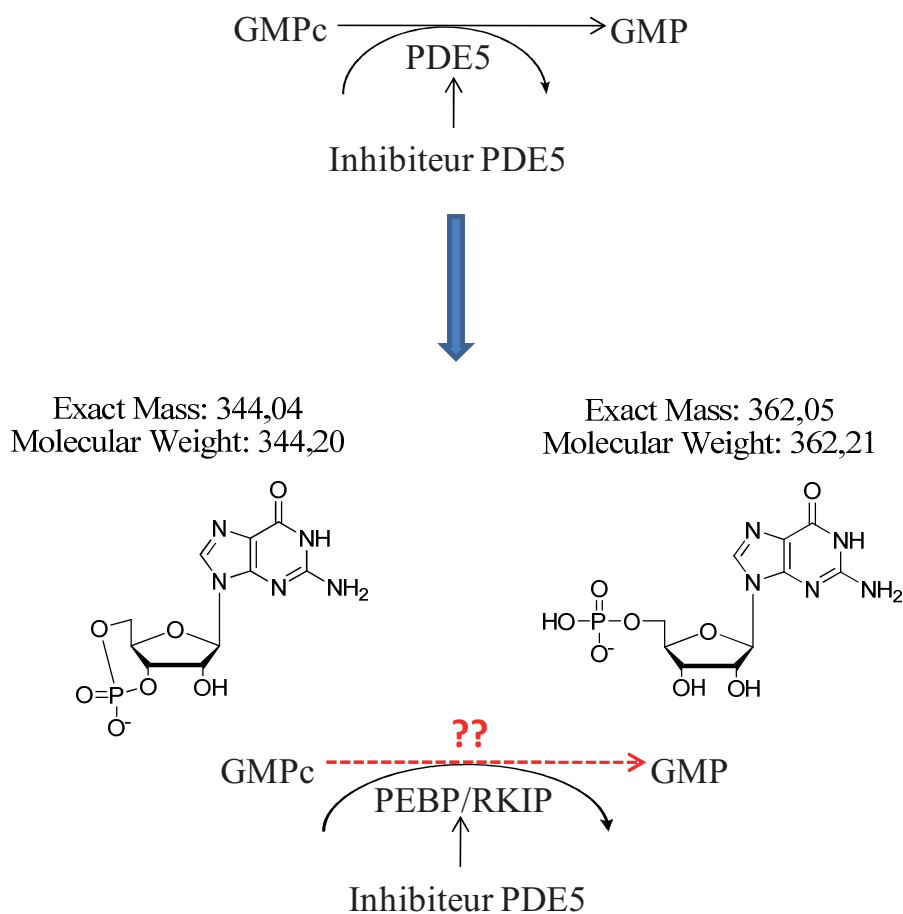


Figure 104: Description de l'activité PDE5

Si la PEBP présente une activité PDE, alors elle doit être capable d'hydrolyser le GMPc en GMP (Figure 104). L'hydrolyse du GMPc s'accompagnerait alors d'un delta-masse de +18 Da. Les PDE nécessitent classiquement la présence de magnésium pour être actives (211, 212).

## 2. Résultats et discussion

Une incubation de la PEBP avec le GMPc en présence de  $MgCl_2$  à 10 mM a donc été réalisée, et l'état du GMPc a été suivi par analyse MALDI-TOF en mode négatif dans une gamme de  $m/z$  correspondant à l'anion monochargé du nucléotide (Figure 105).

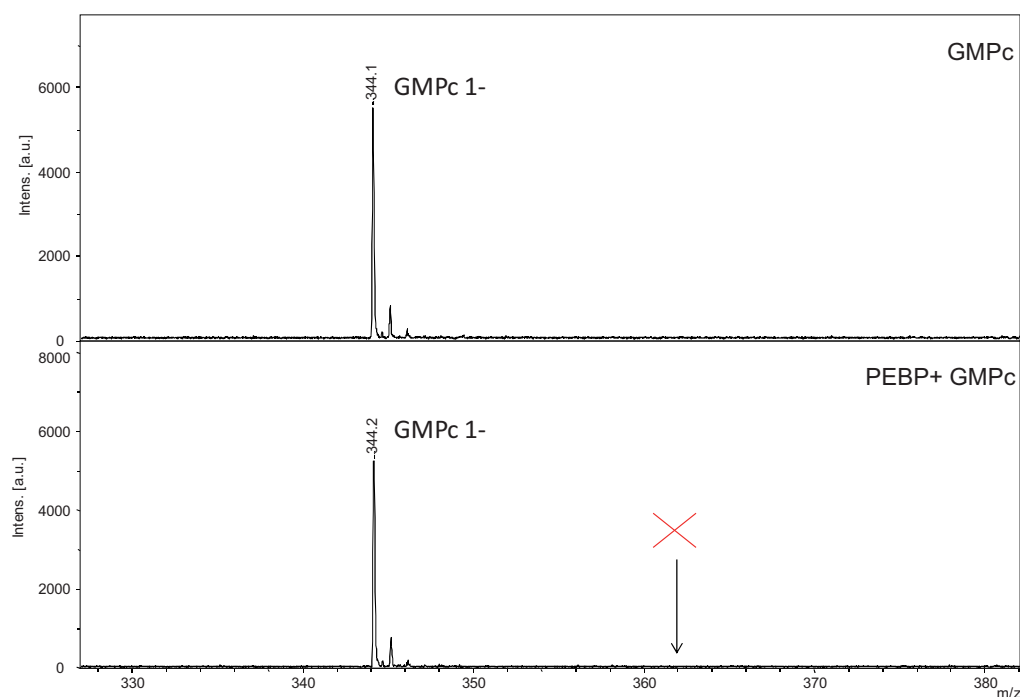


Figure 105: Spectres MALDI-TOF de GMPc 1 mM avec ou sans PEBP 12.6  $\mu M$  en présence de 10 mM  $MgCl_2$  incubés 2H à 37 °C dans ABC 20mM, pH 7.4. Dépôt : mélange 1:1 avec 9AA/EtOH, calibration externe.

Aucune hydrolyse n'est observée. La même expérience a été réalisé avec de l'AMPc pour tester une éventuelle activité PDE sur l'AMPc.

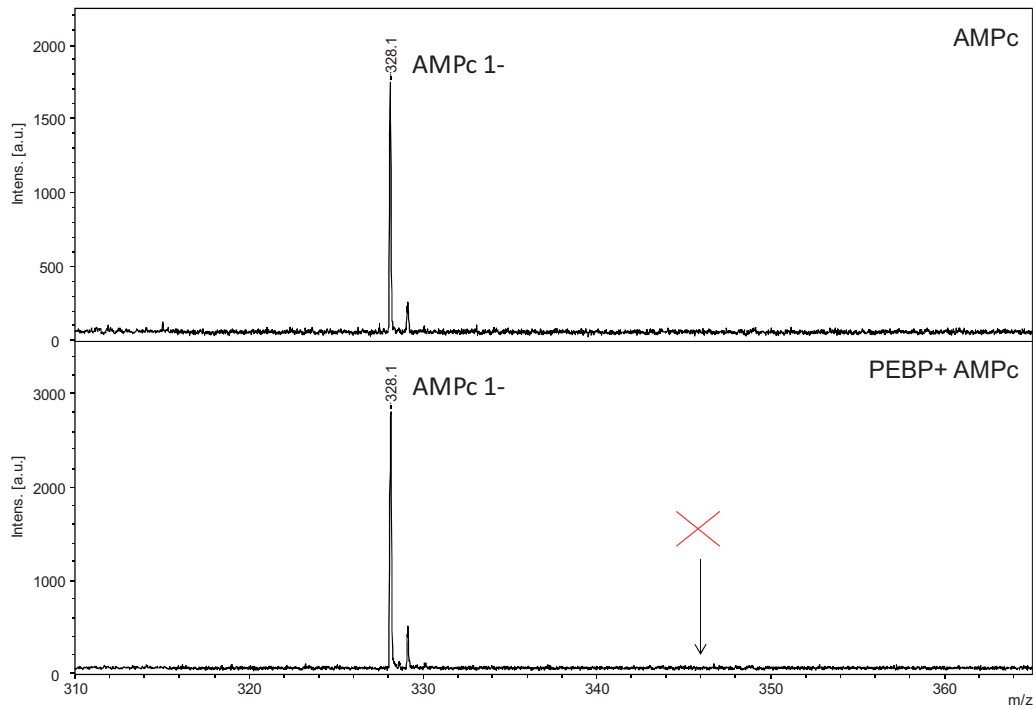


Figure 106 : Spectres MALDI-TOF d'AMPc 1mM avec ou sans PEBP 12.6  $\mu$ M en présence de 10 mM MgCl<sub>2</sub> incubés 2H à 37 °C dans ABC 20mM, pH 7.4. Dépôt : mélange 1:1 avec 9AA/EtOH, calibration externe.

Comme pour le GMPc, aucune activité PDE n'est observée sur l'AMPc.

Les deux nucléotides cycliques ont également été ajoutés simultanément car des études ont établi que certaines PDE nécessitent leur présence conjointe pour être actives (213).

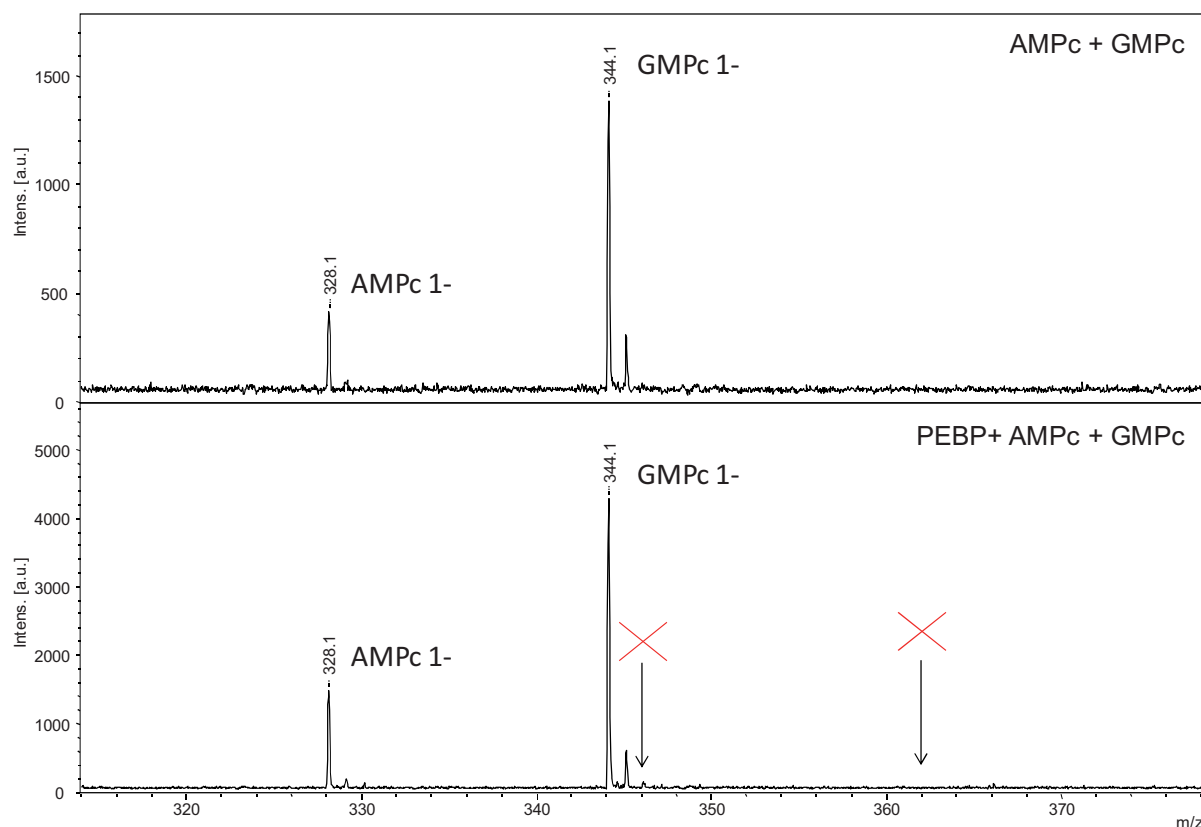


Figure 107: Spectres MALDI-TOF d'AMPc et de GMPc 1mM avec ou sans PEBP 12.6  $\mu$ M en présence de 10 mM MgCl<sub>2</sub> dans ABC 20mM pH 7.4 2H d'incubation à 37 °C. Mélange 1:1 avec 9AA/EtOH, calibration externe.

Dans un deuxième temps, nous nous sommes interrogés sur la concentration de substrat dans l'incubation. En effet, avec 1 mM de substrat et seulement 12.6  $\mu$ M de PEBP, il se peut que la formation du produit hydrolysé soit trop faible par rapport à la quantité initiale de substrat. Dans ce cas, la gamme dynamique de l'instrument peut être insuffisante pour que le produit soit détecté en spectrométrie de masse. Pour obtenir un meilleur ratio produit/substrat, on peut tenter de diminuer la concentration en substrat en conservant la concentration de protéine.

La même expérience a été réalisée avec 63  $\mu$ M de nucléotide et analysée par ESI-IT après MGF pour éliminer le MgCl<sub>2</sub>. Le même résultat est obtenu. Les ions correspondant aux complexes hPEBP1-AMPc et hPEBP1-GMPc ont également été recherchés dans les hautes masses, aucun delta-masse correspondant à de l'AMP ou du GMP n'est visible sur la PEBP.

Bien qu'un inhibiteur de PDE-5 se lie à hPEBP1, nous n'avons pas pu mettre en évidence d'activité phosphodiesterase de la protéine.

### III. Activité GTPase ou ATPase

Des pertes de 80 Da correspondant à des déphosphorylations apparentes ont été observées lors de l'étude de complexes PEBP-nucléotides. Nous avons donc cherché à mettre en évidence pour hPEBP1 une fonction GTPase ou ATPase. Les GTPases ou ATPases nécessitent aussi du  $MgCl_2$  pour être fonctionnelles (214–217). La PEBP a donc été incubée avec le GTP ou l'ATP, en présence de  $MgCl_2$ , et l'état des nucléotides suivi par analyse MALDI-TOF en mode négatif dans une gamme de  $m/z$  correspondant à leur anion monochargé.

L'activité GTPase a d'abord été testée (Figure 108)

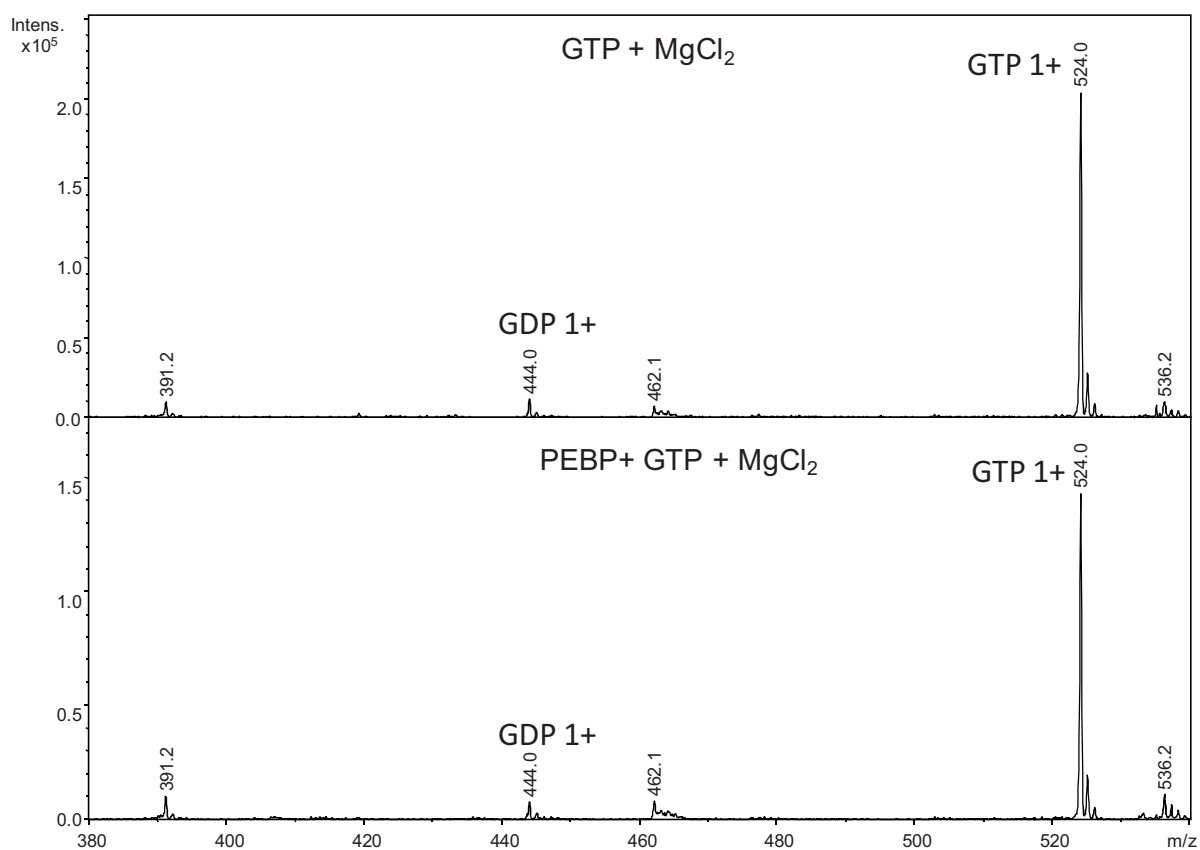


Figure 108 : Spectres ESI-IT de GTP 288  $\mu$ M avec ou sans 18  $\mu$ M de PEBP en présence de 10 mM  $MgCl_2$  incubés 4H à 37 °C dans ABC 20mM, pH 7.4. MGF avant analyse pour éliminer le  $MgCl_2$ .

Un pic à  $-80$  Da, correspondant au GDP, est visible dans le contrôle. Ce GDP observé correspond vraisemblablement à une dégradation chimique en solution ou à un contaminant de synthèse ou de purification dans le stock de GTP. En présence de hPEBP1, l'intensité du pic correspondant au GDP n'augmente pas, et le pic correspondant ne diminue pas par rapport à l'ion de background  $m/z = 391$ . La protéine ne possède donc pas d'activité GTPase dans ces conditions expérimentales.

La même expérience a été réalisée avec de l'ATP au lieu de GTP (Figure 109).

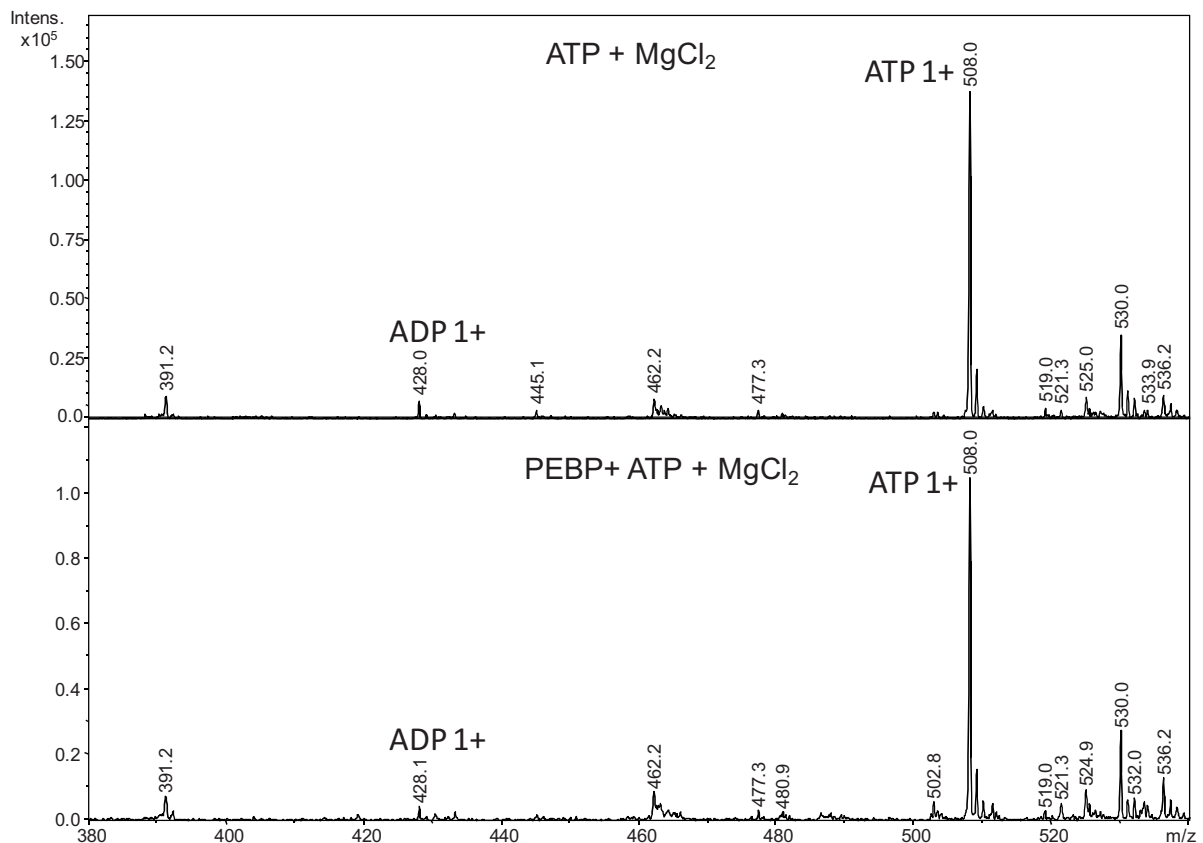


Figure 109: Spectres ESI-IT de 288  $\mu$ M d'ATP avec ou sans 18  $\mu$ M de PEBP en présence de 10 mM  $MgCl_2$  incubés 4H à 37  $^{\circ}C$  dans ABC 20mM, pH 7.4. MGF avant analyse pour éliminer le  $MgCl_2$ .

De nouveau, un pic à  $-80$  Da, correspondant à de l'ADP, est visible dans le contrôle, et aucun changement n'est observé avec l'ajout de PEBP. Cet ADP observé correspond vraisemblablement à une dégradation chimique en solution.

En conclusion, dans les conditions testées, nous n'avons pas pu mettre en évidence pour hPEBP1 humaine d'activité ATPase ou GTPase.

La recherche d'une éventuelle activité enzymatique de la PEBP1 humaine montre que la protéine n'est ni une monooxygénase, ni une GMPC ou AMPc phosphodiesterase, ni une GTPase ou ATPase. Le champ des activités enzymatiques possibles pour cette protéine reste encore très ouvert.





## Conclusion générale et perspectives

Les grands objectifs de ma thèse s'inscrivent dans le cadre du projet MétaSupp. Ils ont consisté à utiliser la spectrométrie de masse pour approfondir la relation structure- activité des ligands nucléotidiques de hPEBP1, et pour fournir des éléments pour la synthèse raisonnée d'activateurs des propriétés antimétastatiques naturelles de cette protéine, dans l'optique d'obtenir des candidats thérapeutiques.

L'analyse de complexes non-covalents par MS a nécessité des mises au point de la préparation et du mode d'infusion directe ESI de l'échantillon, ainsi que des optimisations fines de paramètres de l'ESI-IT. Pour limiter la consommation d'échantillon, l'infusion a été miniaturisée en développant une méthode originale d'« infusion en sandwich ». Afin de préserver au maximum les complexes lors du transfert en phase gazeuse, la tension d'accélération et la température de désolvatation ont été optimisées pour chaque nouveau système, deux d'entre eux étant donnés en exemple.

La détermination rapide et peu consommatrice par ESI-MS de constantes d'équilibre de dissociation ( $K_D$ ) qui reflètent fidèlement l'affinité d'un complexe protéine-ligand en solution suscite un vif intérêt. Il n'existait pas dans la littérature de méthode capable de traiter à la fois l'agrégation en phase gazeuse et la dissociation en phase gazeuse (GPD) sans avoir à utiliser un ligand de référence. Pour hPEBP1, les  $K_D$  pour les ligands de référence étaient inconnus. De plus, les difficultés à mesurer un  $K_D$  sont accrues dans le cas des complexes à faible affinité. Nous présentons ici une méthode de mesure de  $K_D$  accompagnée d'un modèle mathématique qui permet de traiter à la fois la GPD et l'agrégation. Pour cela, une correction mathématique rationnelle de la GPD ( $f_{sat}$ ) est combinée au développement d'un nouveau protocole expérimental qui traite l'agrégation. Un guide est fourni pour appliquer la méthode selon leur comportement cinétique des complexes. L'approche a été validée en comparant des valeurs de  $K_D$  déterminées par notre méthode avec des valeurs de  $K_D$  en solution de la littérature. La méthode, publiée dans *JASMS*, peut être appliquée de manière fiable à un large éventail de systèmes sans nécessiter l'utilisation d'un ligand ou d'une protéine de référence.

Pour caractériser le site nucléotidique de hPEBP1 et contribuer à la découverte de molécules anti-métastases, une recherche des déterminants structuraux optimaux du ligand a été réalisée par criblage d'analogues de synthèse. L'équipe de Luigi Agrofoglio a réalisé la synthèse raisonnée des analogues de nucléotides en s'appuyant sur la structure des meilleurs nucléotides issus du criblage sur colonne de Bucquoy *et al.* : GTP et FMN. Le criblage rapide a été effectué pour tous les ligands nucléotidiques par une mesure par MS d'une fraction de liaison avec hPEBP1, et par la détermination du  $K_D$  pour les molécules de référence et des meilleurs ligands de synthèse. Les molécules analogues de FMN ou d'uracile sont plus affines que les analogues de GTP. Les criblages mettent en évidence des critères structuraux indispensables pour qu'une molécule se lie à hPEBP1 : la présence d'un groupement chargé ou type  $NO_2$ , d'une structure apparentée à une base azotée et d'un cycle. Les résultats de  $K_D$  des ligands et leur corrélation avec le logP de ceux-ci montrent qu'une part importante de l'affinité est liée au caractère hydrophobe du ligand. Les résultats de la MS combinés à ceux de la RMN prouvent que contrairement à rPEBP1, la PEBP1 humaine est capable de lier le GTP, et mettent en évidence une influence de la force ionique sur l'affinité du GTP et du FMN pour hPEBP1. Ces résultats sont en cours de publication.

En termes de structure, la RMN a montré que les sites de liaison du GTP et du FMN sont fortement confondus et ancrés dans la poche à anion, dans laquelle se loge probablement le groupement phosphate. Le fait de ne pas observer d'interaction entre les ligands neutres et la PEBP n'est donc pas dû à une limite de la méthode MS mais est une conséquence logique du mode d'interaction. Le site de liaison du GTP semble indépendant de celui de la locostatine, molécule sans fonction anionique et dont le site ne contiendrait donc pas la poche à anion. Les meilleurs ligands issus du criblage par MS ont été testés en biologie cellulaire, sous forme de pro-drogues, pour leur capacité à inhiber la migration et l'invasion et pour leur cytotoxicité. Le crible sur la migration a permis de sélectionner les molécules à cribler dans le test de l'invasion. Il en ressort quelques molécules dont l'intérêt est de montrer une activité inhibitrice de l'invasion sur les lignées tumorales, tout en étant peu toxiques pour les cellules saines.

Sur la base de nos observations et de données de la littérature, nous avons recherché une activité enzymatique chez hPEBP1, pour laquelle une telle activité n'a jamais été décrite. Aucune activité d'oxydase, de phosphodiesterase-5 sur le GMPc et l'AMPc, ou de déphosphorylase sur l'ATP ou le GTP n'a pu être mise en évidence.

Le savoir-faire pour l'étude de complexes non-covalents développé dans le cadre du projet principal a été mis à profit pour la caractérisation d'autres systèmes dans le cadre de collaborations. Certains de ces systèmes (protéine/protéine, protéine/peptide, protéine/métal et protéine/métal/ADN) ont présenté des difficultés pour se solubiliser et/ou s'ioniser, ou pour préserver les complexes formés en MS. Malgré ces challenges analytiques, un hybride ternaire protéine/métal/ADN, une protéine avec un patch hydrophobe en surface, et des dimères protéiques maintenus essentiellement par des interactions hydrophobes ont pu être analysés par MS supramoléculaire. Une de ces collaborations a permis de confirmer le mécanisme d'action d'un inhibiteur de synthèse sur une protéinase du neutrophile en prouvant la présence d'une liaison non-covalente entre ces partenaires, ce qui a donné lieu à une publication.

Pour aller plus loin dans la caractérisation du site de liaison nucléotidique et la recherche de ligands à forte affinité de hPEBP1, il serait utile de cribler des composés d'origine naturelle afin de collecter davantage d'informations sur les déterminants structuraux du ligand "idéal".

Sachant que les meilleurs ligands analogues de FMN identifiés par criblage par MS sont non-oxydants, il serait très intéressant de les tester pour leur activité anti-invasive. Ceci permettrait notamment de savoir si la longueur de la chaîne carbonée joue réellement un rôle crucial dans effet anti-invasif des analogues nucléotidiques.

Les études par RMN nous apportent des renseignements sur les résidus de la protéine dont la fonction peptidique est perturbée par la liaison d'un ligand. L'attribution des chaînes latérales de hPEBP1 en RMN et les études de perturbation de ligands marqués au  $^{13}\text{C}$ , prévue dans ce projet, permettraient d'affiner les données sur les sites de liaison de hPEBP1. Des structures cristallographiques de hPEBP1 en interaction avec des molécules ont été obtenues par le passé. Si l'on pouvait cristalliser hPEBP1 en complexe avec un analogue nucléotidique, des études par diffraction aux rayons X nous donneraient les résidus de hPEBP1 en interaction directe avec l'analogue nucléotidique. Il serait alors plus facile de faire de la synthèse raisonnée pour trouver des ligands plus affins.

Nous avons mis en évidence que le paramètre  $f_{\text{sat}}$ , la sensibilité de  $f_{\text{sat}}$  à la tension d'accélération et la sensibilité de  $f_{\text{sat}}$  à la force ionique sont influencés par la part de l'interaction hydrophobe d'une liaison. Il serait intéressant d'aller plus loin et de pouvoir quantifier la contribution de l'interaction hydrophobe, des autres types d'interaction, et de leurs combinaisons à la liaison, en combinant d'autres éléments avec ceux que nous avons déjà identifiés.

Pour continuer la recherche d'une éventuelle activité enzymatique de hPEBP1, on pourrait s'inspirer des protéines qui présentent de fortes homologues de structure avec elle. De plus, même si nous n'avons pas mis en évidence d'activité d'oxydo-réduction d'hPEBP1, la protéine pourrait jouer un rôle dans le stress oxydatif. hPEBP1 présente en effet une affinité de l'ordre du micromolaire pour FMN qui est un agent oxydant. Plusieurs publications montrent que hPEBP1 ou des protéines homologues interviennent dans le stress oxydatif (104, 218–220). hPEBP1 pourrait capter les agents oxydants comme FMN et « détoxifier » ainsi la cellule, ce que pourrait corroborer notre observation d'une oxydation de hPEBP1 en présence de FMN et de GTP. Dans cette hypothèse, il faudrait considérer l'effet de l'oxydation sur les fonctions de la protéine.

Inversement, il serait intéressant de tester les analogues de FMN déjà synthétisés ou à venir sur des oxydoréductases, en particulier sur celles impliquées dans les phénomènes tumoraux ou invasifs, en comparant des résultats d'études au niveau moléculaire avec des tests sur l'invasion et d'autres phénomènes cellulaires importants pour la cancérologie.

# Matériels et méthodes

## I. Matériels

Sauf indication contraire, tous les solvants sont de qualité HPLC. Tout le matériel utilisé pour la publication « Improved accuracy of low affinity protein-ligand equilibrium dissociation constants directly determined by electrospray ionization mass spectrometry » a déjà été décrit dans la section *Matériel et Méthode* de celle-ci.

Les analogues de nucléosides phosphonates acycliques ont été synthétisés par l'équipe de Luigi Agrofoglio : Vincent Roy, Aurélien Montagu, Zhili Zhang et Ozkan Sari par cross-métathèse (221).

Les nucléotides GTP, GDP, GMP, ATP, ADP, UMP, GMPc, FMN, FAD, NAD<sup>+</sup>, NADH,H<sup>+</sup>, NADP<sup>+</sup> et NADPH,H<sup>+</sup>, la guanine, la guanosine, la riboflavine et l'alloxazine proviennent de chez Sigma (St. Louis, MO, USA). Les trois molécules issues du criblage de Shemon et ses collaborateurs (« Shemon 26, 48 et 98 ») ont été synthétisées par Hit2lead (San Diego, CA, USA).

Les protéines HU sous forme de HU $\alpha$ 2, HU $\beta$ 2 et HU $\alpha\beta$  ont été produites et purifiées par l'équipe de Bertrand Castaing (189, 222). Les protéines Fpg et NEIL2 ont également été produites et purifiées par l'équipe de Bertrand Castaing (183).

La production et la purification de la protéase PR3 et de l'azapeptide sont décrites dans la section *Matériel et Méthode* de la publication « A selective reversible azapeptide inhibitor of human neutrophil proteinase 3 derived from a high affinity FRET substrate ».

## II. Production de hPEBP1 (100)

Des ARNm poly-A<sup>+</sup> ont été extraits d'oligodendrocytes humains en utilisant le mini kit « Oligotex direct mRNA » de QIAGEN. L'ARNm poly-A<sup>+</sup> purifié a ensuite été soumis à la RT-PCR (Reverse Transcriptase-Polymerase Chain Reaction) en utilisant le système « TITAN one tube RT-PCR system » (Roche, Penzberg, Allemagne).

La PEBP humaine (hPEBP) a été produite sans tag pour être comparable à celle bovine. L'amorce sens est : 5'-ATATCATATGCCGGTGGACCTCAGCAAGTGGT-3', avec un site de restriction *Nde* I (souligné) à l'extrémité 3' de l'ORF. L'amorce antisens est: 5'-TATACTCGAGCTACTTCCCAGACAGCTGCTCATACAGTTT-3' avec un site de restriction *Xho* I à l'extrémité 3' de l'ORF.

Le protocole suivant a été utilisé pour la PCR: dénaturation initiale, 30 min à 60°C, 2 min à 94°C; hybridation and élongation, 30 s à 94°C, 30 s à 64°C, 2 min à 68°C, 30 cycles; élongation prolongée, 68°C 7 min. Le produit de PCR a été cloné dans le plasmide pET 31b(+) (Novagen) en utilisant les endonucléases de restriction *Nde*I and *Xho*I (New England Biolabs) and la ligase T4 DNA (Invitrogen). Des bactéries compétentes *E. coli* BL21(DE3) (Novagen) ont été transformées avec le plasmide pET31b(+) contenant l'insert de hPEBP. Pour la sur-expression de hPEBP, 1 L de milieu LB a été inoculé avec 15 mL de cellules *E. coli* BL21 (pET31bhPEBP) and incubé avec agitation sur la nuit à 30°C jusqu'à atteindre une densité optique de 0.7 à 600 nm; ensuite l'IPTG a été ajouté à une concentration finale de 1 mM pour induire l'expression de hPEBP. Après une incubation de 150 min à 30°C, les cellules ont été agrégées par centrifugation à 4000 g pendant 5 min à 4°C puis congelées à -20°C. Le culot de cellules gelées a été décongelé dans 30 mL de tampon QAE (Tris HCl 50 mM pH 8, NaCl 50 mM, EDTA 1 mM, NaN<sub>3</sub> 0.05%) maintenu au froid dans la glace. La suspension a été soniquée et centrifugée. Le surnageant a été directement chargé sur une colonne QAE Sephadex A-50. La purification de hPEBP a ensuite été réalisée, selon le protocole décrit pour la bPEBP naturelle, c.-à-d., trois étapes de purification ont été réalisées: chromatographie QAE Sephadex A-50, isoélectrofocalisation et chromatographie Blue-Sepharose (6). Entre les étapes de purification, la protéine a été concentrée par centrifugation à 4000g pendant 20 minutes en utilisant des tubes de séparation par centrifugation, Vivaspin,

avec un filtre de 3 kDa de cut-off (Sartorius Stedim, Aubagne, France). La protéine purifiée a été initialement conservée dans une solution de MES 20 mM, avec 1 mM  $\beta$ -mercaptoéthanol et de l'azide de sodium. L'action principale du  $\beta$ -mercaptoéthanol est de réduire les ponts disulfures de la protéine. Cependant, le  $\beta$ -mercaptoéthanol a été détecté par MS sous forme d'adduits covalents avec la hPEBP, par réaction avec les cystéines libres (223).

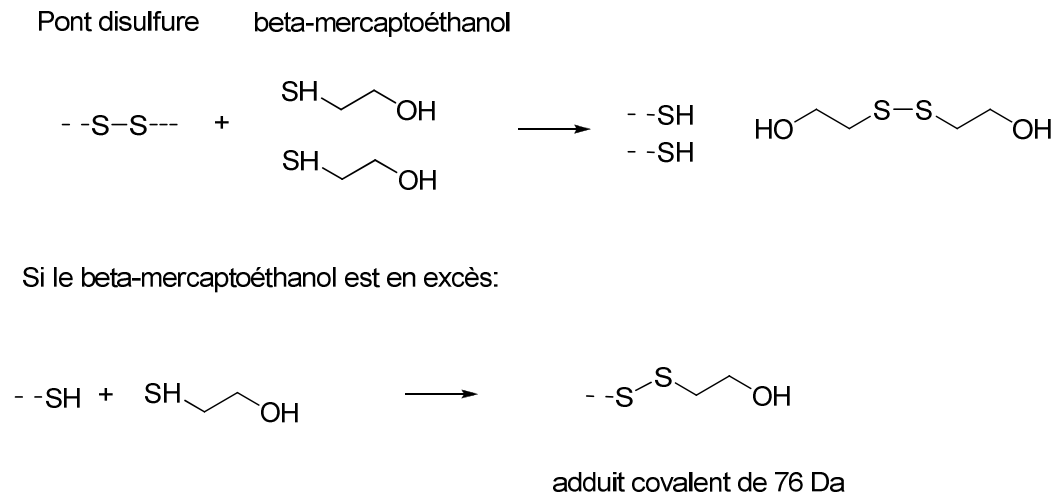


Figure 110: Mécanismes de réduction des ponts disulfures et d'adduction des cystéines libres par le beta-mercaptoéthanol

Un protocole d'élimination du  $\beta$ -mercaptoéthanol par différents agents réducteurs a donc été utilisé. L'addition de TCEP 1 mM final à la solution stock d'hPEBP pendant 1 heure à 37 °C réduit efficacement les ponts  $\beta$ -mercaptoéthanol-cystéine (224).

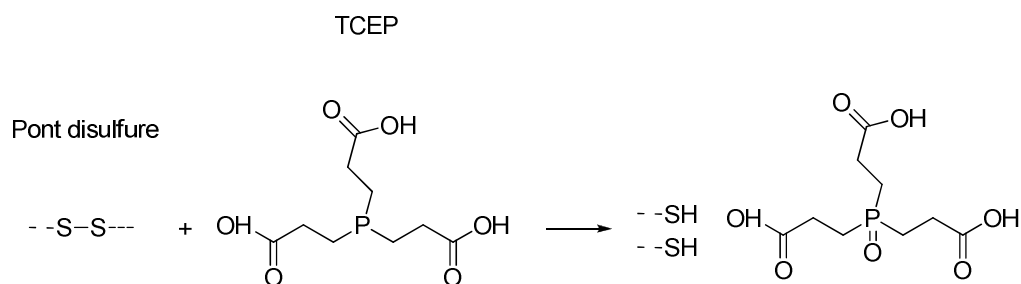


Figure 111 : Mécanisme de réduction des cystéines libres par le TCEP



La solution de protéine réduite a ensuite été dialysée contre de l'eau mQ en utilisant des Micro Dispodialyzer (Harvard Apparatus, Holliston, MA) avec un cut-off de 10 kDa pour éliminer le MES et l'excès de TCEP. La protéine purifiée a été conservée à -80°C dans de l'eau mQ avec 2 équivalents de TCEP.

La séquence de la protéine purifiée a été vérifiée par MS. Environ 10 mg de protéine purifiée ont été obtenu avec 1 L de milieu LB inoculé avec des cellules *E. coli* BL21 (pET31b-hPEBP).

### III. Protocole de dessalage

#### 1. Dessalage des nucléotides

Le protocole présenté en annexe a été adapté de la section « Precipitation of RNA » de la publication de Limbach et ses collaborateurs (225).

#### 2. Dialyse des protéines

Des Micro Dispodialyzer (Harvard Apparatus, Holliston, MA) avec un cut-off de 10 kDa ont été utilisés pour purifier différentes protéines. La face extérieure de la membrane du microdispodialyzer est conditionnée 5 fois avec 100 µL de tampon de dialyse. Un volume de 100 µL d'échantillon à dialyser est ensuite introduit dans le cône et déposé sur la membrane. Le microdispodialyzer est mis à flotter dans le bain de dialyse constitué d'eau mQ, avec agitation magnétique. Le bain est changé à intervalles réguliers pour accélérer la dialyse.

L'échantillon est récupéré par centrifugation et analysé par MS pour vérifier l'efficacité de dialyse.

### 3. Micro gel filtration

Les filtrations sur gel sont réalisées avec des micro spin colonnes « Zeba™ Micro Spin Desalting Columns » (Thermo, Waltham, MA, USA). Le cut-off est de 7 000 Da. Des volumes de résine de 75, 500 et 2000 µL permettent de dessaler respectivement trois gammes de volumes d'échantillons, entre 2 et 12 µL, entre 30 et 130 µL, entre 200 et 700 µL. Les « Zebas » ont été utilisés pour dessaler des protéines ou pour éliminer de petits ligands, de masse inférieure à 7 000 Da.

La solution de conservation de la résine est de l'azide de sodium 0.05%. Avant de dessaler l'échantillon, il est indispensable de conditionner la résine avec le tampon d'incubation. Les protocoles d'utilisation des « Zebas » de 75 ou 500 µL ont été optimisés. Une élimination efficace du tampon de conditionnement de la résine nécessite 5 cycles d'équilibration avec 50 ou 300 µL de tampon d'incubation respectivement. L'échantillon est ensuite chargé sur la colonne et collecté immédiatement après 30 secondes de centrifugation, respectivement à 1000 x g or 1500 x. La protéine seule ou en complexe avec un ligand est alors prête pour l'analyse en spectrométrie de masse.

## IV. Conditions de digestion

La digestion avec la trypsine (Proméga) est réalisée avec un ratio enzyme : substrat de 1 : 2 (masse : masse), durant 2 heures à 37 °C dans 50 mM d'ABC avec 20% d'acétonitile.

La digestion avec l'endoprotéinase GluC (V8) (Roche) est réalisée avec un ratio enzyme : substrat de 1 : 25 (masse : masse), durant 3 heures à 25 °C dans 50 mM de NH<sub>4</sub>OAc, à pH 8, avec 10% d'acétonitile.

## V. Conditions d'analyse

### 1. Conditions ESI natives

En conditions natives, les protéines ont été analysées à des concentrations comprises entre 4 et 18  $\mu\text{M}$ . Le tampon d'incubation est soit du bicarbonate d'ammonium (ABC), soit de l'acétate d'ammonium ( $\text{NH}_4\text{OAc}$ ). Le pH des tampons d'incubation a été ajusté avec de l'acide formique, de l'acide acétique ou de l'hydroxyde d'ammonium. Lorsqu'un ligand est très hydrophobe, 5% de solvant organique sont tolérés dans l'incubation finale sans altération de la distribution des états de charge de la protéine.

### 2. Conditions ESI dénaturantes

En conditions dénaturantes, les protéines ont été analysées à des concentrations comprises entre 0.1 et 2  $\mu\text{M}$ . Le tampon d'incubation est un mélange 50/50 v/v d'acétonitrile et d'eau mQ avec 0.2% final d'acide formique. Ce mélange est appelé Wafa pour Water Acetonitrile Formic Acid. Les analyses en conditions dénaturantes ont été réalisées en diluant au 1/10<sup>ème</sup> l'échantillon de protéine native dans du solvant Wafa.

## VI. Paramètres instrumentaux

### 1. ESI-IT

Deux trappes ioniques ont été utilisées au cours de ma thèse, le modèle Esquire HCT (Bruker, Bremen, Germany), puis le modèle Ultra HCT PMT Discovery (Bruker).

Le schéma de l'instrument est représenté sur la Figure 112.

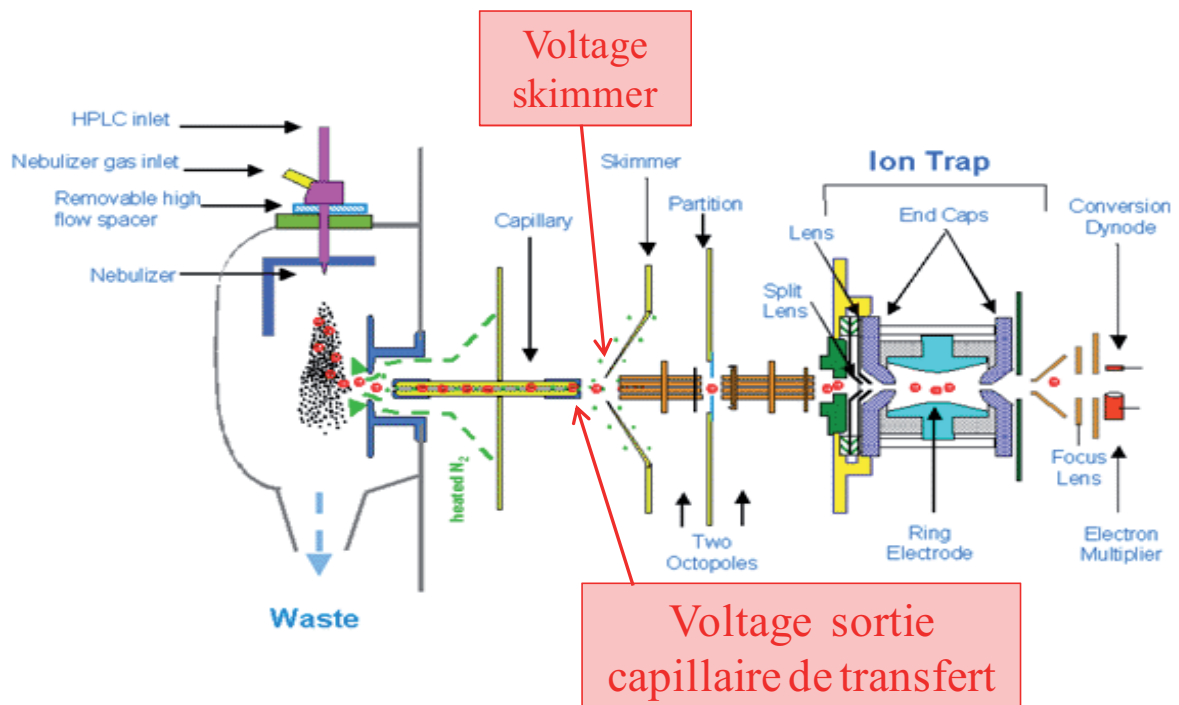


Figure 112 : Représentation schématique de l'instrument ESI-IT Esquire HCT de Bruker

La calibration de la trappe est réalisée avec du Tuning Mix, de référence G2421A chez Agilent Technologies (Santa Clara, CA)

Pour toutes les analyses en mode positif, les paramètres instrumentaux étaient fixés ainsi :

Polarity	Trap
<input checked="" type="radio"/> Positive	<input checked="" type="checkbox"/> ICC
<input type="radio"/> Negative	SmartTarget <input type="text" value="90000"/>
<input type="checkbox"/> Alternating	Max. Accu Time <input type="text" value="200.00"/> ms
<input checked="" type="radio"/> Link Edit	Scan <input type="text" value="200"/> to <input type="text" value="2000"/> m/z
<input type="radio"/> Unlink Edit	Averages <input type="text" value="10"/>
for Parameter	<b>Rolling Averaging</b>
<input type="checkbox"/> Apply to all Segments	<input checked="" type="checkbox"/> On No. <input type="text" value="2"/>
	<input type="button" value="Apply"/>

Environ 90000 ions sont accumulés dans la trappe, en une durée maximale de 200 millisecondes. La moyenne de 10 microscans est réalisée pour donner un spectre. La moyenne glissante du dernier spectre et de la moyenne précédente au cours du temps est ensuite effectuée. La gamme de  $m/z$  à scanner est adaptée au système que l'on analyse.

Le débit du gaz séchant est de 5 L/min et la pression du gaz de nébulisation de 10 psi. L'offset appliqué sur la partie « end plate » est de - 500 V.

En conditions natives, la température de désolvatation est de 120 °C, la tension appliquée à l'aiguille de - 3400 V (voir les optimisations dans le chapitre I), avec un débit d'infusion de 1.3 µL/min. En conditions dénaturantes, le débit est le même, la température de désolvatation est de 250°C, et la tension appliquée à l'aiguille de - 4000 V

Deux modes peuvent ensuite être choisis pour fixer les voltages : le mode « assisté » et le mode « expert »

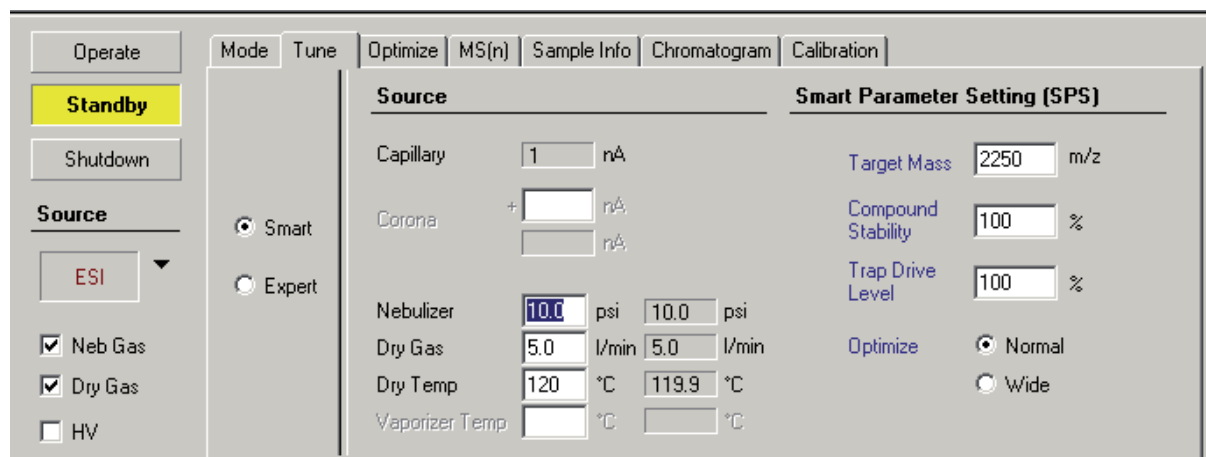


Figure 113: Configuration « assistée » pour fixer les différents voltages

Le mode « assisté » permet d'utiliser la fonction « Smart Parameter Setting », qui optimise les voltages appliqués sur le cap exit, l'octopole 2 DC et l'octopole RF en fonction du rapport  $m/z$  ciblé. Les voltages appliqués sur le skimmer, l'octopole 1 DC, les lentilles 1 et 2 ne changent pas et sont respectivement de 40, 8, -5 et -60 V. La fonction « Smart Parameter Setting » peut

optimiser les paramètres instrumentaux en fonction de la stabilité observée sur un composé de  $m/z$  donné. Plus la stabilité est faible et plus le voltage de cap exit est diminué.

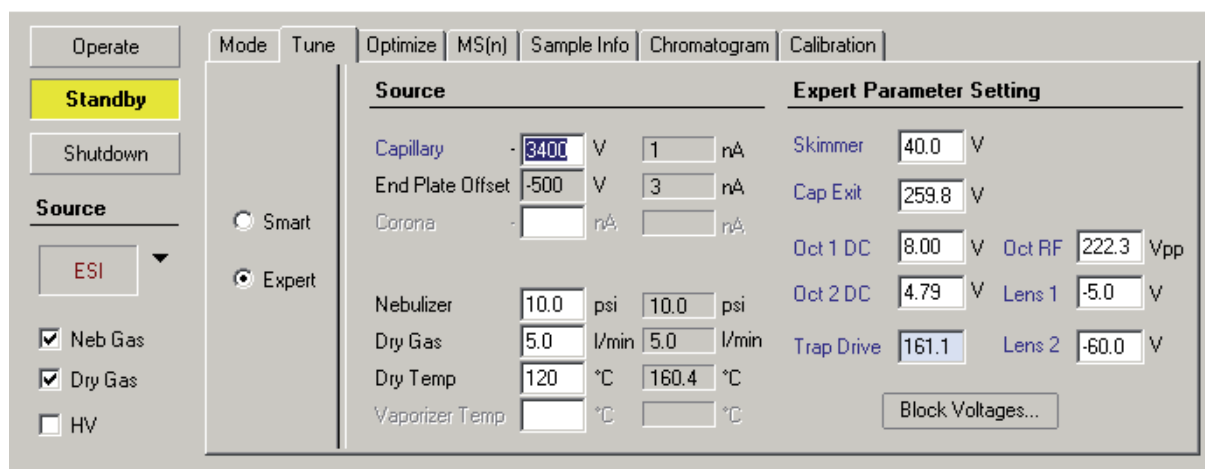


Figure 114 : Configuration « experte » pour fixer les différents voltages

Le mode « expert » permet d'ajuster librement tous les voltages, notamment celui appliqué sur le skimmer qui n'est pas accessible en mode « assisté ». La nécessité d'augmenter le voltage appliqué sur l'octopole RF pour obtenir un signal pour les ions de rapport  $m/z$  élevé est un signe d'encrassement de l'instrument.

De manière générale, les voltages sont optimisés afin de trouver un compromis entre l'énergie communiquée aux ions et la préservation des édifices non-covalents.

L'application « optimisation » peut être également utilisée pour étudier l'effet d'un paramètre sur au maximum deux ions d'intérêts (Figure 115).

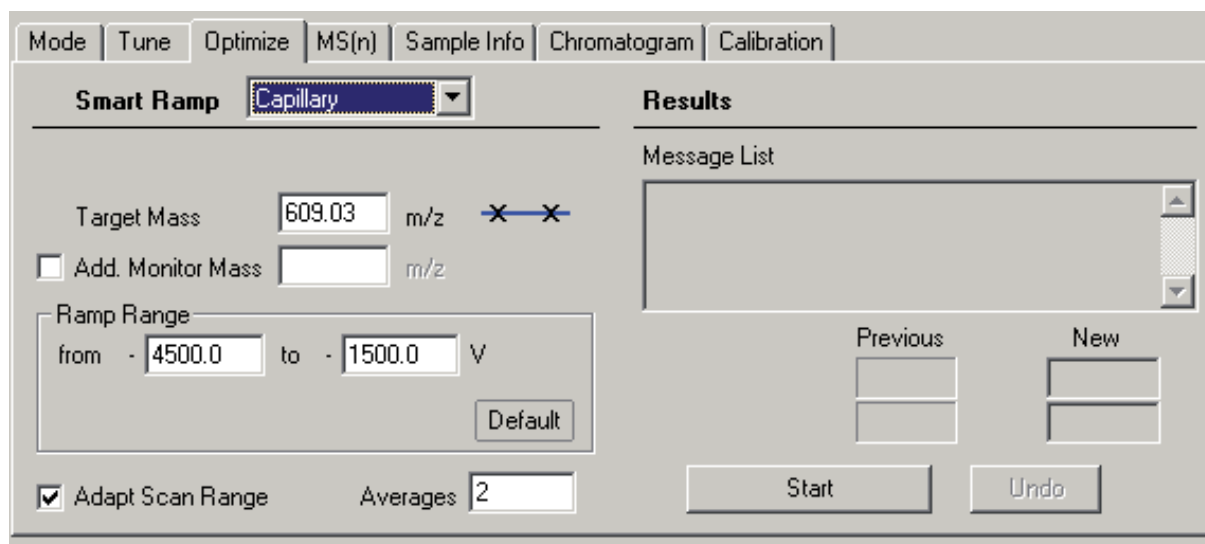


Figure 115: Configuration de l'application « optimisation »

## 2. MALDI-TOF

Deux MALDI-TOF ont été utilisés au cours de ma thèse, le modèle Autoflex I (Bruker, Bremen, Germany), puis le modèle Ultraflex II TOF/TOF (Bruker).

La calibration est externe sur dépôt adjacent et réalisée avec l'apomyoglobine (référence A8673 de Sigma) pour les analyses de protéine, avec le mélange Peptide Calibration Standart II (Bruker) pour les peptides issus de la protéolyse de hPEBP1, et finalement avec la matrice 9-aminoacridine (9AA) et le GTP pour les nucléotides.

Les échantillons de hPEBP1 ou de myoglobine ont été analysés après dilution de l'échantillon dans un rapport 1: 20 ou 1: 10 (pour hPEBP1) ou 1 : 50 (pour la myoglobine) avec une solution de 4HCCA/TWA ou de 4HCCA/FWI. Les échantillons de peptides sont analysés par dilution dans un rapport 1: 20 avec une solution de 4HCCA/TWA, et les échantillons de nucléotides par dilution dans un rapport 1: 1 avec une solution de 9AA/EtOH. Le mélange Peptide Calibration Standart II est analysé par dilution dans un rapport 1: 20 avec une solution dans 4HCCA/TWA pour la calibration externe mais par dilution dans un rapport 1: 100 avec une solution dans 4HCCA/FWI pour les tests d'oxydations en présence de FMN. Les dépôts ont été réalisés par la technique en couche ultra mince (UTL) (226, 227) sur une cible dorée à

l'or fin ou sur une cible en acier poli MPT 384 (Bruker). Si nécessaire, les dépôts sont lavés une ou deux fois avec une solution aqueuse de TFA 0.1%.

### 3. ESI-UHR-Qq-TOF (maXis)

Le maXis 4 Ghz UHR-QTOF (Bruker) est un instrument de pouvoir résolutif  $> 50\,000$ . Il a été installé sur la plateforme dans le cadre de la Fédération ICOA-CBM en décembre 2010.

La calibration du maXis est réalisée avec du Tuning Mix, de référence G1969-85000 chez Agilent Technologies (Santa Clara, CA).

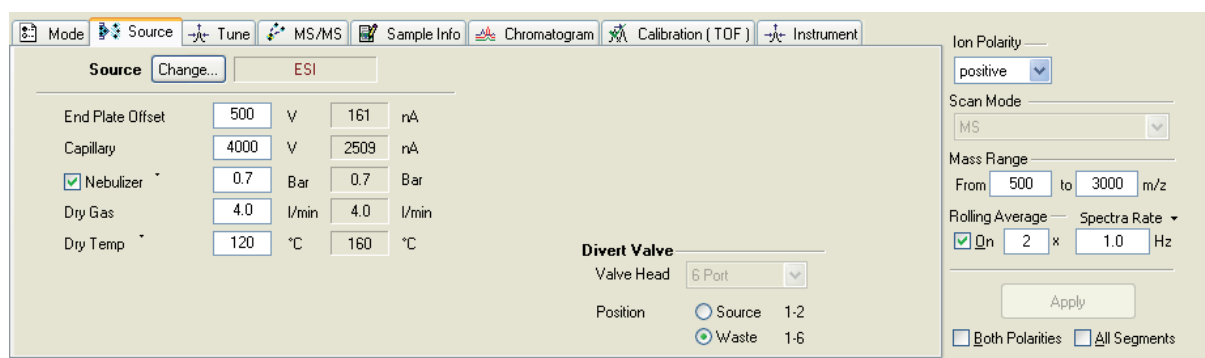


Figure 116: Onglet « Source » du logiciel d'acquisition est micrOTOFcontrol

Le voltage de l'aiguille du maXis est fixé à - 4000 V. L'offset appliqué sur la partie « end plate » est de - 500 V. La température de désolvatation est de 120 °C et la pression de 0.7 bar. Le débit de gaz séchant est par défaut de 4 L/min (Figure 116). Ce paramètre a été fixé à 10L/min pour analyser en conditions natives la PR3, Fpg-Zn et NEIL2-Zn. La gamme de masse varie en fonction du système étudié.

Le paramètre isCID (in source collision induced dissociation) est ajusté pour préserver les interactions non-covalentes entre la protéine et le ligand. Une valeur trop haute peut entraîner la dissociation du complexe. Les autres paramètres ne sont pas optimisés mais les valeurs correspondent à la méthode « Tune-high » de Bruker



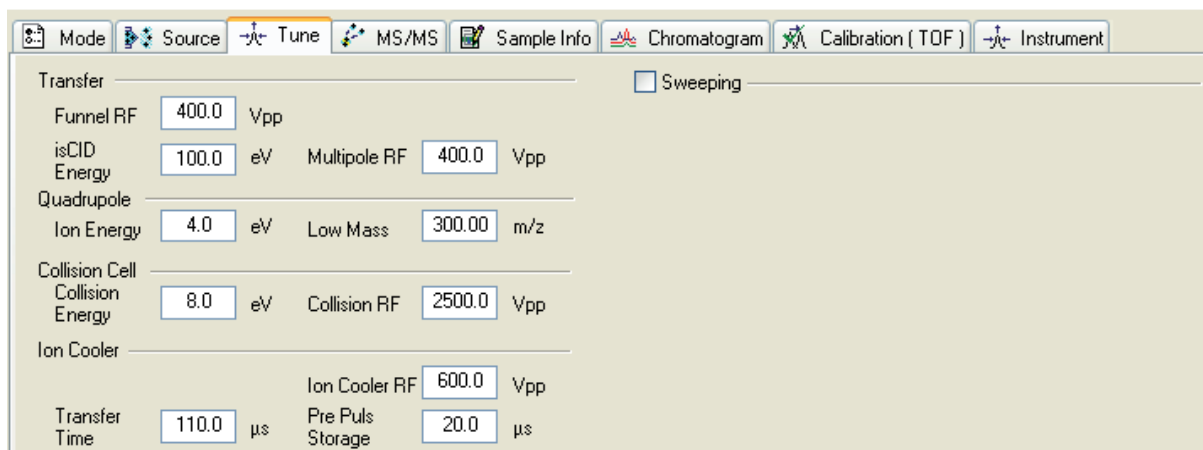


Figure 117 : Onglet « Tune » du logiciel d'acquisition est micrOTOFcontrol

Le logiciel d'acquisition est micrOTOFcontrol version 3.1 (Bruker) et les données sont retraitées avec DataAnalysis version 4.0 SP4 (Bruker).

# Annexes

# Principe de l'analyse des ions

Le principe physique sur lequel est basée la détermination du rapport  $m/z$  dépend du type d'analyseur du spectromètre de masse. On dénombre plusieurs types d'analyseurs tels que l'analyseur quadripolaire (Q), à temps de vol (TOF), à trappe ionique (IT), à résonance cyclotronique (ICR) et à secteur magnétique. Tous ces analyseurs peuvent être couplés à l'ionisation électrospray.

Pour comparer les performances des différents analyseurs, on s'intéresse principalement à deux paramètres :

- la gamme de masse
- la résolution

Tableau 27 : Performances des spectromètres de masse

Analyseurs	Gamme de masse (m/z)	Résolution FWHM
Quadripôle (Q)	8000	2000
Trappe ionique (IT)	4000	2000
Temps de vol (TOF)	> 500 000	20 000
Q-TOF haute résolution	8000	60 000
Orbitrap	4000	30 000 - 100 000
Cyclotron à résonance des ions (FTICR)	4000	100 000 - 1 000 000

La gamme de masse va définir la limite du rapport  $m/z$  des ions analysables. Le pouvoir résolusif va correspondre à la capacité de fournir des signaux que l'on peut distinguer pour deux ions de masses voisines ( $R=m/\delta m$ ). On parle maintenant couramment de résolution FWHM (full width at half maximum) et dans ces cas  $\delta m$  est prise à 50% de son maximum. La résolution est un paramètre déterminant dans l'analyse par spectrométrie de masse. Une bonne résolution permet de déterminer la masse de chaque pic isotopique d'un composé et plus

généralement de séparer des espèces différentes. Dans le cas des complexes non covalents, la résolution de l'appareil sera un paramètre déterminant pour distinguer l'interaction de petites molécules sur une protéine de haut poids moléculaire.

### 1. L'analyseur à temps de vol

L'analyseur à temps de vol est constitué de deux zones, une zone d'accélération des ions par application de potentiels électriques de plusieurs dizaines de kVolts et une zone libre de champ où règne un vide poussé.

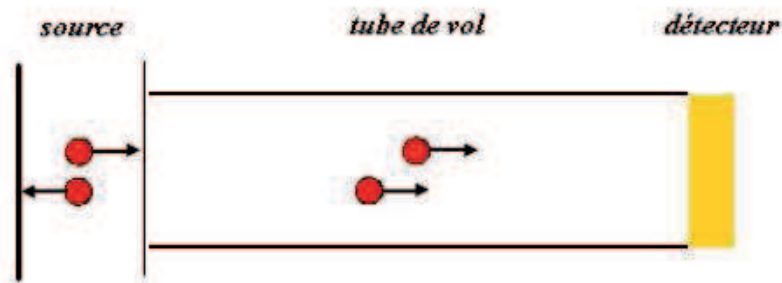


Figure 118 : Représentation schématique de l'analyseur TOF

L'énergie potentielle ( $E_p$ ) d'une particule chargée dans un champ électrique est donnée par la relation :  $E_p = z \cdot e \cdot U$  (1)

$z$ , la charge de l'ion

$e$ , la charge élémentaire

$U$ , la tension d'accélération

Quand la particule chargée est accélérée dans la première zone, son énergie potentielle est transformée en énergie cinétique ( $E_c$ ) :  $E_c = \frac{1}{2} \cdot m \cdot v^2$  (2)

m, la masse

v, la vitesse de l'ion

L'égalité des deux énergies rend compte du fait que l'énergie potentielle est convertie en énergie cinétique :  $E_p = E_c$  (3)

$$\text{D'où: } z.e.U = \frac{1}{2}.m.v^2 \quad (4)$$

La vitesse d'une particule chargée dans une zone libre de champ est constante. Elle peut donc être exprimée en fonction de L, la longueur du tube de vol, et t, le temps de vol :

$$v = \frac{L}{t} \quad (5)$$

$$\text{La combinaison des équations 5 et 6 donne : } t = L.\left(\frac{m}{2.z.e.U}\right)^{\frac{1}{2}} \quad (6)$$

Plus le rapport m/z d'un ion est élevé et plus il mettra de temps à traverser le tube de vol.

Depuis la fin des années 80, le TOF a été couplé au mode d'ionisation MALDI car l'injection pulsée des ions est compatible avec une mesure de temps. Le couplage entre le mode d'ionisation électrospray et l'analyseur TOF est apparu après car il nécessite lui une injection orthogonale pulsée des ions vers l'analyseur, afin de limiter la perte de résolution due la dispersion transversale en énergie cinétique. Une autre amélioration consiste à appliquer un miroir ou réflecteur électrostatique à l'extrémité de l'analyseur, afin d'augmenter la résolution (228).

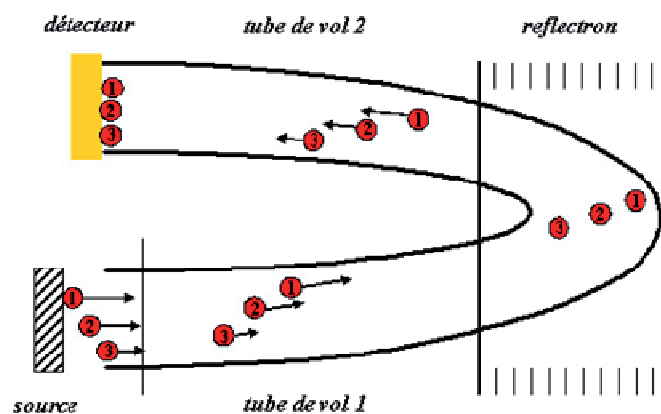


Figure 119 : Représentation schématique de l'analyseur TOF-TOF

La dispersion en énergie cinétique des ions diminue après réflexion des ions dans un miroir constitué de grilles cylindriques sur lesquelles sont appliqués des potentiels décroissants. Les analyseurs TOF sont désormais des analyseurs dont la résolution peut dépasser 60 000 et présentent l'avantage unique de pouvoir analyser en théorie des composés de rapport  $m/z$  infini. Le mode d'ionisation électrospray couplé au TOF semble donc particulièrement bien adapté à l'analyse de molécules biologiques de très hauts poids moléculaires.

## 2. La trappe à ions

Paul et Steinwedel, les inventeurs de l'analyseur quadripolaire, ont décrit la première trappe ionique en 1960 (229, 230). L'instrument a ensuite été modifié par Stafford *et al* (231). Le piège est constitué de trois électrodes à section hyperbolique : une électrode annulaire encadrée par des électrodes d'entrée et de sortie - qui forment les calottes supérieure et inférieure du dispositif (Figure 120).

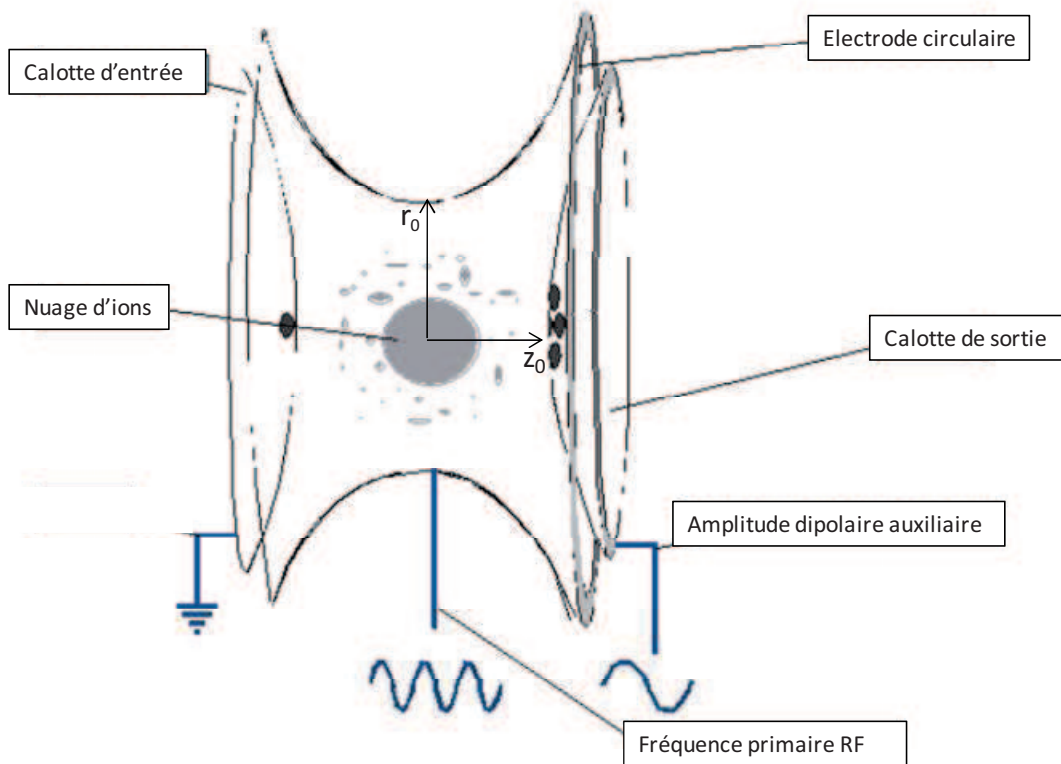


Figure 120: La géométrie de la trappe ionique

L'analyse et la détection des ions s'effectuent dans un même espace, suivant des séquences temporelles successives. Une tension en radiofréquence  $V \cdot \cos(2 \pi f t)$  combinée ou non à une tension continue  $U$  est appliquée entre l'électrode centrale et les deux électrodes calottes. Le champ résultant est alors tridimensionnel. Des ions de toutes masses sont piégés dans la trappe et ont une trajectoire dimensionnelle formant un huit. L'utilisation d'un gaz de refroidissement de type Argon ou Hélium favorise le piégeage et le confinement des ions dans la trappe. Les inventeurs ont proposé d'utiliser la trappe ionique comme un spectromètre de masse en appliquant dans l'axe  $z$  une fréquence de résonance pour éjecter les ions d'un rapport  $m/z$  donné. Dans le cas des instruments quadripolaires, les potentiels sont ajustés de façon à ce que seuls les ions d'un  $m/z$  sélectionné passent à travers les cylindres. Le principe est différent pour la trappe ionique. Des ions de différentes masses sont présents ensemble

dans la trappe, et sont éjectés séquentiellement selon leurs masses de telle façon à obtenir un spectre.

Une analyse mathématique basée sur les équations de Mathieu permet de localiser des domaines de stabilité dans lesquelles des ions d'une masse donnée ont une trajectoire stable. Ces aires peuvent être décrites dans un diagramme, avec en ordonnée,  $U$ , la tension continue et en abscisse  $V$ , l'amplitude de la tension alternative. Les aires dans lesquelles les ions sont stables sont celles pour lesquelles les trajectoires n'excèdent jamais les dimensions de la trappe,  $z_0$  et  $r_0$ . Un balayage de l'amplitude de la radiofréquence  $V$  entraînera donc l'expulsion des ions piégés selon cet axe, vers le détecteur.

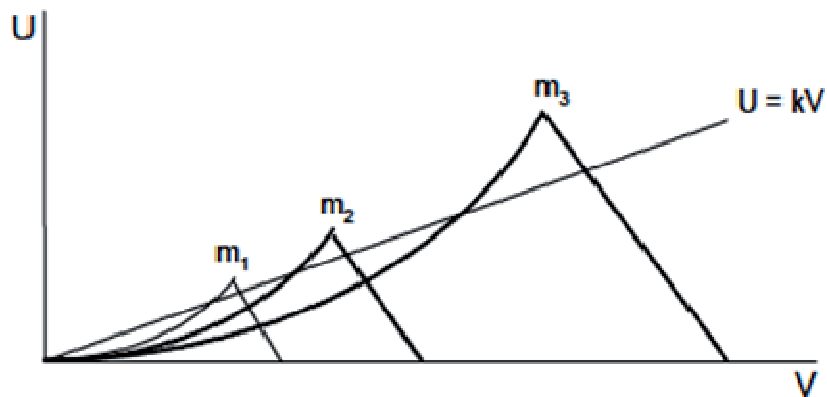


Figure 121: Diagramme de stabilité pour trois ions de masse  $m_1 < m_2 < m_3$  en fonction des valeurs  $U$  et  $V$  pour un quadropole.



## Détermination du $K_D$ par MS

Equation quadratique classique de détermination de  $K_D$  :

$$f_{\text{bound}} = \frac{[\text{PL}]}{[\text{P}]_0} = \frac{[\text{P}]_0 + [\text{L}]_0 + K_D - \sqrt{([\text{P}]_0 + [\text{L}]_0 + K_D)^2 - 4[\text{P}]_0 [\text{L}]_0}}{2[\text{P}]_0}$$

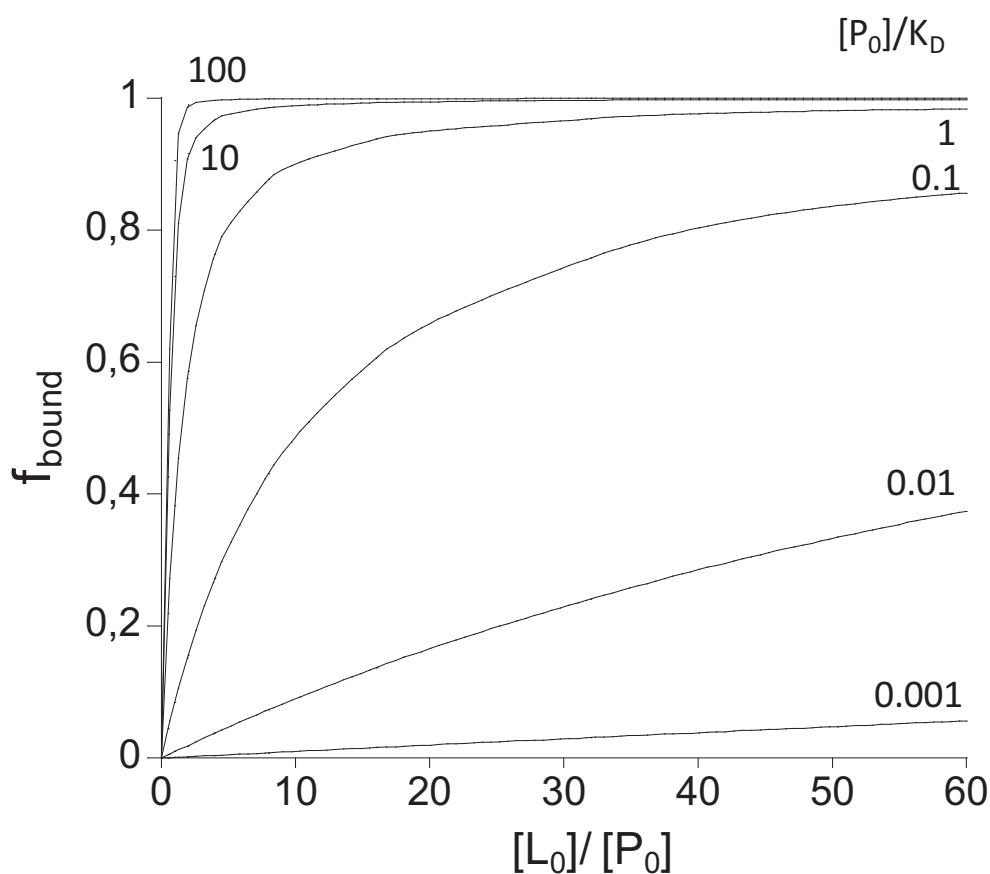


Figure 122 : Courbes théoriques des fractions de liaison ( $[\text{PL}] / [\text{P}]_0$ ) en fonction des ratios molaires de ligand sur protéine ( $[\text{L}]_0 / [\text{P}]_0$ ) pour les valeurs de ratio ( $[\text{P}]_0 / K_D$ ) indiquées (inspiré de la publication de Bieth (232)).

# Cartes peptidiques des protéolyses de hPEBP1

## Protéolyse de hPEBP1 par la trypsine :

Carte peptidique (trypsine) de la PEBP humaine seule

1 15 30  
PVDLSK<sup>W</sup>SGPLSLQEVD<sup>W</sup>EQPHLV<sup>W</sup>TYAG

31 45 60  
AAVDEL<sup>W</sup>GKVLTP<sup>W</sup>TQVKNRPT<sup>W</sup>SIS<sup>W</sup>DGLDSG

61 75 90  
KLYTLVLTDPDAPSRKDPKYRE<sup>W</sup>HHFLV<sup>W</sup>VN

91 105 120  
MKGNDI<sup>W</sup>SSGTVLSDYVGS<sup>W</sup>GPPKGTGLHRYV

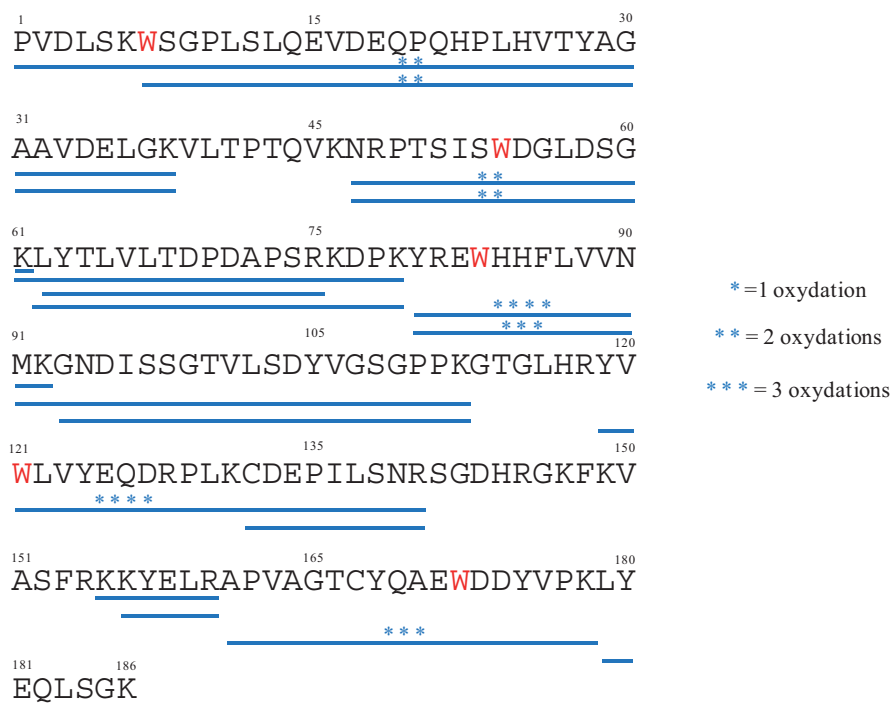
121 135 150  
<sup>W</sup>LVYEQDRPLKCDEPILSNRSGDHRGKFKV

151 165 180  
ASFRKKYELRAPVAGTCYQAE<sup>W</sup>DDYVPKLY

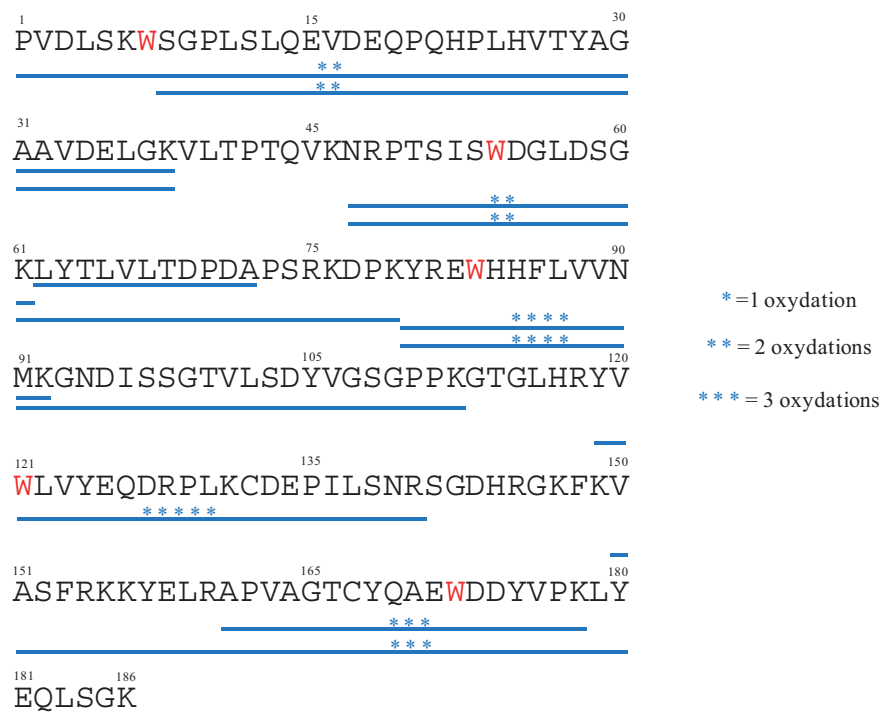
181 186  
EQLSGK

\* =1 oxydation

## Carte peptidique (trypsine) de la PEBP humaine en mélange avec le FMN

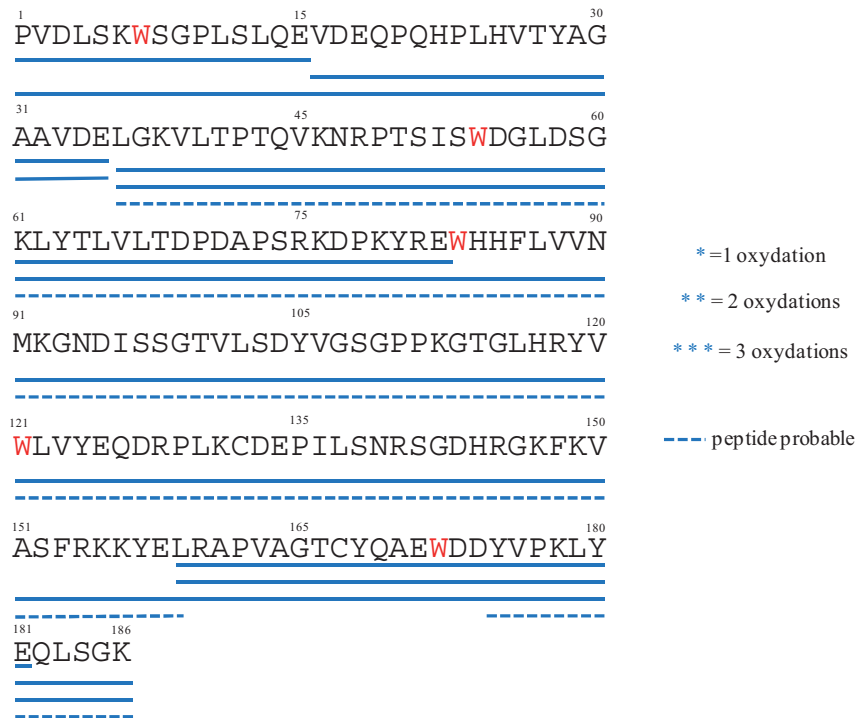


## Carte peptidique (trypsine) de la PEBP humaine en mélange avec le FMN et le GTP

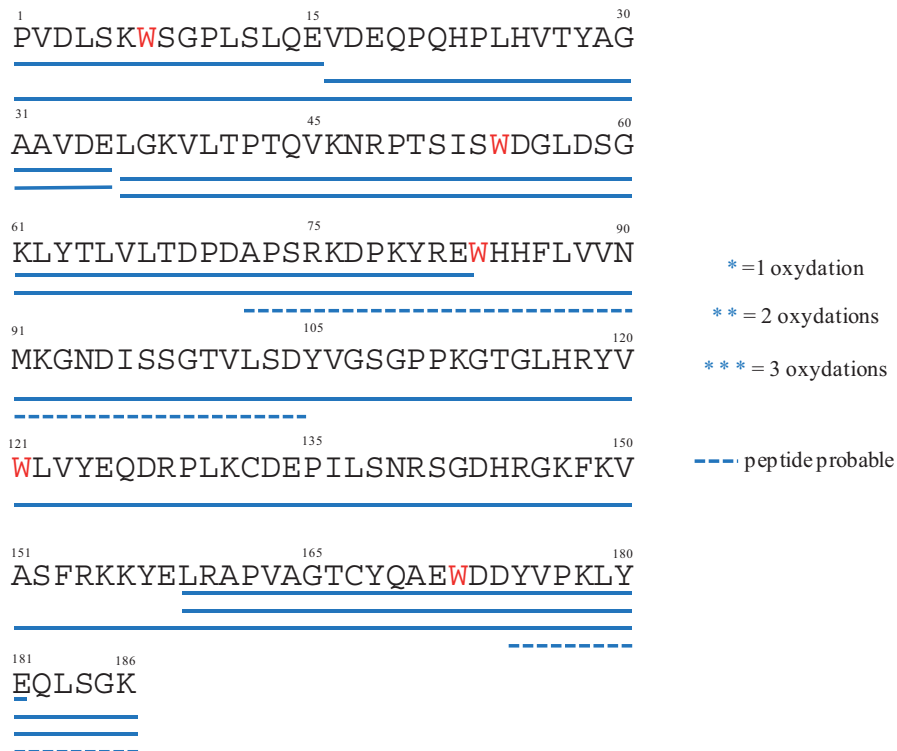


## Protéolyse de hPEBP1 par l'endoprotéinase GluC (V8):

Carte peptidique (V8) de la PEBP humaine



Carte peptidique (V8) de la PEBP humaine en mélange avec le GTP





# Protocole de dessalage des nucléotides

Adapté de la section « Precipitation of RNA » de la publication de Limbach et ses collaborateurs (225)

- Préparer entre 50 et 70  $\mu\text{L}$  de solution de nucléotide à 40mM (solvant : eau).
- Ajouter 1/3 volume d'acétate d'ammonium 10M tel que la concentration finale en acétate d'ammonium soit de 2.5M.
- Précipiter la solution en ajoutant 2.5 volumes d'éthanol à 95% froid (conservé à  $-80^{\circ}\text{C}$ ) tel que le pourcentage final d'éthanol soit de 70%.
- Centrifuger 15 minutes à 12 500 rpm.
- Oter le surnageant et rincer le culot avec environ 400  $\mu\text{L}$  d'éthanol à 70% froid ( $-80^{\circ}\text{C}$ ).
- Centrifuger 15 minutes à 12 500 rpm, ôter le surnageant.
- Répéter le cycle de dissolution dans l'eau, ajout acétate d'ammonium, précipitation et rinçage, 1 à 3 fois selon les cas. Par exemple, dans le cas du FMN, on perd beaucoup de matériel dans le surnageant, on ne peut donc pas faire plus de 2 cycles.
- Faire évaporer le solvant du culot au speed-vacpour créer un stock de nucléotide sous forme de poudre



# Bibliographie

- (1) Katta, V., and Chait, B. T. (1991) Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry, *J. Am. Chem. Soc.* *113*, 8534–8535.
- (2) Ganem, B., Li, Y. T., and Henion, J. D. (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry, *J. Am. Chem. Soc.* *113*, 6294–6296.
- (3) Fu, Z., Smith, P. C., Zhang, L., Rubin, M. A., Dunn, R. L., Yao, Z., and Keller, E. T. (2003) Effects of raf kinase inhibitor protein expression on suppression of prostate cancer metastasis, *J. Natl. Cancer Inst* *95*, 878–889.
- (4) Fu, Z., Kitagawa, Y., Shen, R., Shah, R., Mehra, R., Rhodes, D., Keller, P. J., Mizokami, A., Dunn, R., Chinnaiyan, A. M., Yao, Z., and Keller, E. T. (2006) Metastasis suppressor gene Raf kinase inhibitor protein (RKIP) is a novel prognostic marker in prostate cancer, *Prostate* *66*, 248–256.
- (5) Serre, L., Pereira de Jesus, K., Zelwer, C., Bureaud, N., Schoentgen, F., and Bénédetti, H. (2001) Crystal structures of YBHB and YBCL from *Escherichia coli*, two bacterial homologues to a Raf kinase inhibitor protein, *J. Mol. Biol.* *310*, 617–634.
- (6) Bernier, I., and Jollés, P. (1984) Purification and characterization of a basic 23 kDa cytosolic protein from bovine brain, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* *790*, 174–181.
- (7) Bernier, I., Tresca, J.-P., and Jollès, P. (1986) Ligand-binding studies with a 23 kDa protein purified from bovine brain cytosol, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* *871*, 19–23.
- (8) Yeung, K., Seitz, T., Li, S., Janosch, P., McFerran, B., Kaiser, C., Fee, F., Katsanakis, K. D., Rose, D. W., Mischak, H., Sedivy, J. M., and Kolch, W. (1999) Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP, *Nature* *401*, 173–177.
- (9) Roussel, G., Nussbaum, F., Schoentgen, F., Jollès, P., and Nussbaum, J. L. (1988) Immunological investigation of a 21-kilodalton cytosolic basic protein in rat brain, *Dev. Neurosci* *10*, 65–74.
- (10) Vallée, B. S., Tauc, P., Brochon, J. C., Maget-Dana, R., Lelièvre, D., Metz-Boutigue, M. H., Bureaud, N., and Schoentgen, F. (2001) Behaviour of bovine phosphatidylethanolamine-binding protein with model membranes. Evidence of affinity for negatively charged membranes, *Eur. J. Biochem* *268*, 5831–5841.
- (11) Goumon, Y., Muller, A., Glattard, E., Marban, C., Gasnier, C., Strub, J.-M., Chasserot-Golaz, S., Rohr, O., Stefano, G. B., Welters, I. D., Van Dorsselaer, A., Schoentgen, F., Aunis, D., and Metz-Boutigue, M.-H. (2006) Identification of morphine-6-glucuronide in chromaffin cell secretory granules, *J. Biol. Chem* *281*, 8082–8089.
- (12) Frayne, J., Ingram, C., Love, S., and Hall, L. (1999) Localisation of phosphatidylethanolamine-binding protein in the brain and other tissues of the rat, *Cell Tissue Res* *298*, 415–423.
- (13) Hirtz, C., Chevalier, F., Centeno, D., Egea, J. C., Rossignol, M., Sommerer, N., and de Périère, D. (2005) Complexity of the human whole saliva proteome, *J. Physiol. Biochem* *61*, 469–480.



- (14) Schoentgen, F., and Jollès, P. (1995) From structure to function: possible biological roles of a new widespread protein family binding hydrophobic ligands and displaying a nucleotide binding site, *FEBS Lett* 369, 22–26.
- (15) Vallée, B., Teyssier, C., Maget-Dana, R., Ramstein, J., Bureaud, N., and Schoentgen, F. (1999) Stability and physicochemical properties of the bovine brain phosphatidylethanolamine-binding protein, *Eur. J. Biochem* 266, 40–52.
- (16) Corbit, K. C., Trakul, N., Eves, E. M., Diaz, B., Marshall, M., and Rosner, M. R. (2003) Activation of Raf-1 Signaling by Protein Kinase C through a Mechanism Involving Raf Kinase Inhibitory Protein, *J. Biol. Chem.* 278, 13061–13068.
- (17) Banfield, M. J., Barker, J. J., Perry, A. C., and Brady, R. L. (1998) Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a role in membrane signal transduction, *Structure* 6, 1245–1254.
- (18) Simister, P. C., Banfield, M. J., and Brady, R. L. (2002) The crystal structure of PEBP-2, a homologue of the PEBP/RKIP family, *Acta Crystallogr. D Biol. Crystallogr* 58, 1077–1080.
- (19) Zeng, L., Imamoto, A., and Rosner, M. R. (2008) Raf kinase inhibitory protein (RKIP): a physiological regulator and future therapeutic target, *Expert Opin. Ther. Targets* 12, 1275–1287.
- (20) Wang, X., Li, N., Liu, B., Sun, H., Chen, T., Li, H., Qiu, J., Zhang, L., Wan, T., and Cao, X. (2004) A novel human phosphatidylethanolamine-binding protein resists tumor necrosis factor alpha-induced apoptosis by inhibiting mitogen-activated protein kinase pathway activation and phosphatidylethanolamine externalization, *J. Biol. Chem.* 279, 45855–45864.
- (21) Serre, L., Vallée, B., Bureaud, N., Schoentgen, F., and Zelwer, C. (1998) Crystal structure of the phosphatidylethanolamine-binding protein from bovine brain: a novel structural class of phospholipid-binding proteins, *Structure* 6, 1255–1265.
- (22) Banfield, M. J., and Brady, R. L. (2000) The structure of Antirrhinum centroradialis protein (CEN) suggests a role as a kinase regulator, *J. Mol. Biol.* 297, 1159–1170.
- (23) Granovsky, A. E., Clark, M. C., McElheny, D., Heil, G., Hong, J., Liu, X., Kim, Y., Joachimiak, G., Joachimiak, A., Koide, S., and Rosner, M. R. (2009) Raf Kinase Inhibitory Protein Function Is Regulated via a Flexible Pocket and Novel Phosphorylation-Dependent Mechanism, *Mol Cell Biol* 29, 1306–1320.
- (24) Simister, P. C., Burton, N. M., and Brady, R. L. (2011) Phosphotyrosine Recognition by the Raf Kinase Inhibitor Protein, *Forum Immun Dis Ther* 2, 59–70.
- (25) Hengst, U., Albrecht, H., Hess, D., and Monard, D. (2001) The Phosphatidylethanolamine-binding Protein Is the Prototype of a Novel Family of Serine Protease Inhibitors, *Journal of Biological Chemistry* 276, 535–540.
- (26) Caesar, R., and Blomberg, A. (2004) The stress-induced Tfs1p requires NatB-mediated acetylation to inhibit carboxypeptidase Y and to regulate the protein kinase A pathway, *J. Biol. Chem* 279, 38532–38543.
- (27) Keller, E. T., Fu, Z., and Brennan, M. (2005) The biology of a prostate cancer metastasis suppressor protein: Raf kinase inhibitor protein, *J. Cell. Biochem* 94, 273–278.
- (28) Yeung, K. C., Rose, D. W., Dhillon, A. S., Yaros, D., Gustafsson, M., Chatterjee, D., McFerran, B., Wyche, J., Kolch, W., and Sedivy, J. M. (2001) Raf Kinase Inhibitor Protein Interacts with NF- $\kappa$ B-Inducing Kinase and TAK1 and Inhibits NF- $\kappa$ B Activation, *Mol. Cell. Biol.* 21, 7207–7217.

- (29) Kolch, W. (2000) Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions, *Biochem. J* 351 Pt 2, 289–305.
- (30) Dhillon, A. S., Hagan, S., Rath, O., and Kolch, W. (2007) MAP kinase signalling pathways in cancer, *Oncogene* 26, 3279–3290.
- (31) Roberts, P. J., and Der, C. J. (2007) Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer, *Oncogene* 26, 3291–3310.
- (32) Schubbert, S., Shannon, K., and Bollag, G. (2007) Hyperactive Ras in developmental disorders and cancer, *Nat. Rev. Cancer* 7, 295–308.
- (33) Yeung, K., Janosch, P., McFerran, B., Rose, D. W., Mischak, H., Sedivy, J. M., and Kolch, W. (2000) Mechanism of Suppression of the Raf/MEK/Extracellular Signal-Regulated Kinase Pathway by the Raf Kinase Inhibitor Protein, *Mol. Cell. Biol.* 20, 3079–3085.
- (34) Saab, F., Routier, S., Méroux, J., Bénateau, V., and Schoentgen, F. (2010) Comparison of the efficiencies of two TR-FRET methods to detect in vitro natural and synthesized inhibitors of the Raf/MEK/ERK signaling pathway, *IJHTS* 81.
- (35) Eves, E. M., Shapiro, P., Naik, K., Klein, U. R., Trakul, N., and Rosner, M. R. (2006) Raf kinase inhibitory protein regulates aurora B kinase and the spindle checkpoint, *Mol. Cell* 23, 561–574.
- (36) Li, X., and Stark, G. R. (2002) NFkappaB-dependent signaling pathways, *Exp. Hematol* 30, 285–296.
- (37) Karin, M., and Greten, F. R. (2005) NF-kappaB: linking inflammation and immunity to cancer development and progression, *Nat. Rev. Immunol* 5, 749–759.
- (38) Neer, E. J. (1995) Heterotrimeric G proteins: organizers of transmembrane signals, *Cell* 80, 249–257.
- (39) Neves, S. R., Ram, P. T., and Iyengar, R. (2002) G Protein Pathways, *Science* 296, 1636–1639.
- (40) Krupnick, J. G., and Benovic, J. L. (1998) The role of receptor kinases and arrestins in G-coupled receptor regulation, *Annu. Rev. Pharmacol. Toxicol.* 38, 289–319.
- (41) Krosiak, T., Koch, T., Kahl, E., and Höllt, V. (2001) Human phosphatidylethanolamine-binding protein facilitates heterotrimeric G protein-dependent signaling, *J. Biol. Chem* 276, 39772–39778.
- (42) Lorenz, K., Lohse, M. J., and Quitterer, U. (2003) Protein kinase C switches the Raf kinase inhibitor from Raf-1 to GRK-2, *Nature* 426, 574–579.
- (43) Yoshida, B. A., Sokoloff, M. M., Welch, D. R., and Rinker-Schaeffer, C. W. (2000) Metastasis-suppressor genes: a review and perspective on an emerging field, *J. Natl. Cancer Inst* 92, 1717–1730.
- (44) Welch, D. R., and Rinker-Schaeffer, C. W. (1999) What defines a useful marker of metastasis in human cancer?, *J. Natl. Cancer Inst* 91, 1351–1353.
- (45) Schuierer, M. M., Bataille, F., Hagan, S., Kolch, W., and Bosserhoff, A.-K. (2004) Reduction in Raf kinase inhibitor protein expression is associated with increased Ras-extracellular signal-regulated kinase signaling in melanoma cell lines, *Cancer Res* 64, 5186–5192.
- (46) Lee, H. C., Tian, B., Sedivy, J. M., Wands, J. R., and Kim, M. (2006) Loss of Raf kinase inhibitor protein promotes cell proliferation and migration of human hepatoma cells, *Gastroenterology* 131, 1208–1217.
- (47) Chatterjee, D., Bai, Y., Wang, Z., Beach, S., Mott, S., Roy, R., Braastad, C., Sun, Y., Mukhopadhyay, A., Aggarwal, B. B., Darnowski, J., Pantazis, P., Wyche, J., Fu, Z.,

- Kitagwa, Y., Keller, E. T., Sedivy, J. M., and Yeung, K. C. (2004) RKIP sensitizes prostate and breast cancer cells to drug-induced apoptosis, *J. Biol. Chem* 279, 17515–17523.
- (48) Park, S., Yeung, M. L., Beach, S., Shields, J. M., and Yeung, K. C. (2005) RKIP downregulates B-Raf kinase activity in melanoma cancer cells, *Oncogene* 24, 3535–3540.
- (49) Hagan, S., Al-Mulla, F., Mallon, E., Oien, K., Ferrier, R., Gusterson, B., García, J. J. C., and Kolch, W. (2005) Reduction of Raf-1 kinase inhibitor protein expression correlates with breast cancer metastasis, *Clin. Cancer Res* 11, 7392–7397.
- (50) Zhang, L., Fu, Z., Binkley, C., Giordano, T., Burant, C. F., Logsdon, C. D., and Simeone, D. M. (2004) Raf kinase inhibitory protein inhibits beta-cell proliferation, *Surgery* 136, 708–715.
- (51) Minoo, P., Zlobec, I., Baker, K., Tornillo, L., Terracciano, L., Jass, J. R., and Lugli, A. (2007) Loss of raf-1 kinase inhibitor protein expression is associated with tumor progression and metastasis in colorectal cancer, *Am. J. Clin. Pathol* 127, 820–827.
- (52) Beach, S., Tang, H., Park, S., Dhillon, A. S., Keller, E. T., Kolch, W., and Yeung, K. C. (2008) Snail is a repressor of RKIP transcription in metastatic prostate cancer cells, *Oncogene* 27, 2243–2248.
- (53) Hebert, L. E., Scherr, P. A., Bienias, J. L., Bennett, D. A., and Evans, D. A. (2003) Alzheimer Disease in the US Population: Prevalence Estimates Using the 2000 Census, *Arch Neurol* 60, 1119–1122.
- (54) Maki, M., Matsukawa, N., Yuasa, H., Otsuka, Y., Yamamoto, T., Akatsu, H., Okamoto, T., Ueda, R., and Ojika, K. (2002) Decreased expression of hippocampal cholinergic neurostimulating peptide precursor protein mRNA in the hippocampus in Alzheimer disease, *J. Neuropathol. Exp. Neurol* 61, 176–185.
- (55) Ojika, K., Kojima, S., Ueki, Y., Fukushima, N., Hayashi, K., and Yamamoto, M. (1992) Purification and structural analysis of hippocampal cholinergic neurostimulating peptide, *Brain Research* 572, 164–171.
- (56) George, A. J., Holsinger, R. M. D., McLean, C. A., Tan, S.-S., Scott, H. S., Cardamone, T., Cappai, R., Masters, C. L., and Li, Q.-X. (2006) Decreased phosphatidylethanolamine binding protein expression correlates with Abeta accumulation in the Tg2576 mouse model of Alzheimer's disease, *Neurobiol. Aging* 27, 614–623.
- (57) Baritaki, S., Katsman, A., Chatterjee, D., Yeung, K. C., Spandidos, D. A., and Bonavida, B. (2007) Regulation of tumor cell sensitivity to TRAIL-induced apoptosis by the metastatic suppressor Raf kinase inhibitor protein via Yin Yang 1 inhibition and death receptor 5 up-regulation, *J. Immunol* 179, 5441–5453.
- (58) Frayne, J., McMillen, A., Love, S., and Hall, L. (1998) Expression of phosphatidylethanolamine-binding protein in the male reproductive tract: immunolocalisation and expression in prepubertal and adult rat testes and epididymides, *Mol. Reprod. Dev* 49, 454–460.
- (59) Jones, R., and Brown, C. R. (1987) Identification and characterization of the 2D6 and Mr 23,000 antigens on the plasma membrane of rat spermatozoa, *Biochem. J* 241, 353–360.
- (60) Jones, R., and Hall, L. (1991) A 23 kDa protein from rat sperm plasma membranes shows sequence similarity and phospholipid binding properties to a bovine brain cytosolic protein, *Biochimica et Biophysica Acta (BBA) - Protein Structure and*

- Molecular Enzymology* 1080, 78–82.
- (61) Perry, A., Wiley, E. L., and Albores-Saavedra, J. (1994) Pagetoid spread of intratubular germ cell neoplasia into rete testis: a morphologic and histochemical study of 100 orchiectomy specimens with invasive germ cell tumors, *Hum. Pathol* 25, 235–239.
  - (62) Hickox, D. M., Gibbs, G., Morrison, J. R., Sebire, K., Edgar, K., Keah, H.-H., Alter, K., Loveland, K. L., Hearn, M. T. W., de Kretser, D. M., and O'Bryan, M. K. (2002) Identification of a novel testis-specific member of the phosphatidylethanolamine binding protein family, *pebp-2*, *Biol. Reprod* 67, 917–927.
  - (63) Moffit, J. S., Boekelheide, K., Sedivy, J. M., and Klysik, J. (2007) Mice lacking Raf kinase inhibitor protein-1 (RKIP-1) have altered sperm capacitation and reduced reproduction rates with a normal response to testicular injury, *J. Androl* 28, 883–890.
  - (64) Gibbons, R., Adeoya-Osiguwa, S. A., and Fraser, L. R. (2005) A mouse sperm decapacitation factor receptor is phosphatidylethanolamine-binding protein 1, *Reproduction* 130, 497–508.
  - (65) Bradley, D., Carpenter, R., Copsey, L., Vincent, C., Rothstein, S., and Coen, E. (1996) Control of inflorescence architecture in *Antirrhinum*, *Nature* 379, 791–797.
  - (66) Mc Henry, K. T., Ankala, S. V., Ghosh, A. K., and Fenteany, G. (2002) A non-antibacterial oxazolidinone derivative that inhibits epithelial cell sheet migration, *Chembiochem* 3, 1105–1111.
  - (67) Mc Henry, K. T., Montesano, R., Zhu, S., Beshir, A. B., Tang, H.-H., Yeung, K. C., and Fenteany, G. (2008) Raf kinase inhibitor protein positively regulates cell-substratum adhesion while negatively regulating cell-cell adhesion, *J. Cell. Biochem* 103, 972–985.
  - (68) Zhu, S., Mc Henry, K. T., Lane, W. S., and Fenteany, G. (2005) A chemical inhibitor reveals the role of Raf kinase inhibitor protein in cell migration, *Chem. Biol* 12, 981–991.
  - (69) Shemon, A. N., Eves, E. M., Clark, M. C., Heil, G., Granovsky, A., Zeng, L., Imamoto, A., Koide, S., and Rosner, M. R. (2009) Raf Kinase Inhibitory Protein Protects Cells against Locostatin-Mediated Inhibition of Migration, *PLoS ONE* 4, e6028.
  - (70) Schoentgen, F., Seddiqi, N., Bucquoy, S., Jollès, P., Lemesle-Varloot, L., Provost, K., and Mornon, J.-P. (1992) Main structural and functional features of the basic cytosolic bovine 21 kDa protein delineated through Hydrophobic Cluster Analysis and molecular modelling, *Protein Eng.* 5, 295–303.
  - (71) Rao, S. T., and Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins, *J. Mol. Biol* 76, 241–256.
  - (72) Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) Chemical and biological evolution of a nucleotide-binding protein, *Nature* 250, 194–199.
  - (73) Bucquoy, S., Jollès, P., and Schoentgen, F. (1994) Relationships between molecular interactions (nucleotides, lipids and proteins) and structural features of the bovine brain 21-kDa protein, *Eur. J. Biochem* 225, 1203–1210.
  - (74) De Clercq, E., and Holý, A. (2005) Acyclic nucleoside phosphonates: a key class of antiviral drugs, *Nat Rev Drug Discov* 4, 928–940.
  - (75) Zhang, J., Visser, F., King, K. M., Baldwin, S. A., Young, J. D., and Cass, C. E. (2007) The role of nucleoside transporters in cancer chemotherapy with nucleoside drugs, *Cancer Metastasis Rev* 26, 85–110.

- (76) Gourdeau, H., Clarke, M. L., Ouellet, F., Mowles, D., Selner, M., Richard, A., Lee, N., Mackey, J. R., Young, J. D., Jolivet, J., Lafrenière, R. G., and Cass, C. E. (2001) Mechanisms of uptake and resistance to troxacitabine, a novel deoxycytidine nucleoside analogue, in human leukemic and solid tumor cell lines, *Cancer Res* 61, 7217–7224.
- (77) Damaraju, V. L., Damaraju, S., Young, J. D., Baldwin, S. A., Mackey, J., Sawyer, M. B., and Cass, C. E. Nucleoside anticancer drugs: the role of nucleoside transporters in resistance to cancer chemotherapy[dagger], *Oncogene* 22, 7524–7536.
- (78) Hoff, P. M., Ansari, R., Batist, G., Cox, J., Kocha, W., Kuperminc, M., Maroun, J., Walde, D., Weaver, C., Harrison, E., Burger, H. U., Osterwalder, B., Wong, A. O., and Wong, R. (2001) Comparison of oral capecitabine versus intravenous fluorouracil plus leucovorin as first-line treatment in 605 patients with metastatic colorectal cancer: results of a randomized phase III study, *J. Clin. Oncol* 19, 2282–2292.
- (79) Blum, J. L., Jones, S. E., Buzdar, A. U., LoRusso, P. M., Kuter, I., Vogel, C., Osterwalder, B., Burger, H. U., Brown, C. S., and Griffin, T. (1999) Multicenter phase II study of capecitabine in paclitaxel-refractory metastatic breast cancer, *J. Clin. Oncol* 17, 485–493.
- (80) Talbot, D. C., Moiseyenko, V., Van Belle, S., O'Reilly, S. M., Alba Conejo, E., Ackland, S., Eisenberg, P., Melnychuk, D., Pienkowski, T., Burger, H.-U., Laws, S., and Osterwalder, B. (2002) Randomised, phase II trial comparing oral capecitabine (Xeloda[reg]) with paclitaxel in patients with metastatic/advanced breast cancer pretreated with anthracyclines, *Br J Cancer* 86, 1367–1372.
- (81) Ellison, R. R., Holland, J. F., Weil, M., Jacquillat, C., Boiron, M., Bernard, J., Sawitsky, A., Rosner, F., Gussoff, B., Silver, R. T., Karanas, A., Cuttner, J., Spurr, C. L., Hayes, D. M., Blom, J., Leone, L. A., Haurani, F., Kyle, R., Hutchison, J. L., Forcier, R. J., and Moon, J. H. (1968) Arabinosyl cytosine: a useful agent in the treatment of acute leukemia in adults, *Blood* 32, 507–523.
- (82) Chun, H. G., Leyland-Jones, B., and Cheson, B. D. (1991) Fludarabine phosphate: a synthetic purine antimetabolite with significant activity against lymphoid malignancies, *J. Clin. Oncol* 9, 175–188.
- (83) Dimopoulos, M. A., Kantarjian, H., Weber, D., O'Brien, S., Estey, E., Delasalle, K., Rose, E., Cabanillas, F., Keating, M., and Alexanian, R. (1994) Primary therapy of Waldenström's macroglobulinemia with 2-chlorodeoxyadenosine, *J. Clin. Oncol* 12, 2694–2698.
- (84) Seymour, J. F., Kurzrock, R., Freireich, E. J., and Estey, E. H. (1994) 2-chlorodeoxyadenosine induces durable remissions and prolonged suppression of CD4+ lymphocyte counts in patients with hairy cell leukemia, *Blood* 83, 2906–2911.
- (85) Lindemalm, S., Liliemark, J., Gruber, A., Eriksson, S., Karlsson, M. O., Wang, Y., and Albertioni, F. (2003) Comparison of cytotoxicity of 2-chloro- 2'-arabino-fluoro-2'-deoxyadenosine (clofarabine) with cladribine in mononuclear cells from patients with acute myeloid and chronic lymphocytic leukemia, *Haematologica* 88, 324–332.
- (86) Burris, H. A., 3rd, Moore, M. J., Andersen, J., Green, M. R., Rothenberg, M. L., Modiano, M. R., Cripps, M. C., Portenoy, R. K., Storniolo, A. M., Tarassoff, P., Nelson, R., Dorr, F. A., Stephens, C. D., and Von Hoff, D. D. (1997) Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial, *J. Clin. Oncol* 15, 2403–2413.
- (87) Stadler, W. M., Kuzel, T., Roth, B., Raghavan, D., and Dorr, F. A. (1997) Phase II

- study of single-agent gemcitabine in previously untreated patients with metastatic urothelial cancer, *J. Clin. Oncol* 15, 3394–3398.
- (88) von der Maase, H., Hansen, S. W., Roberts, J. T., Dogliotti, L., Oliver, T., Moore, M. J., Bodrogi, I., Albers, P., Knuth, A., Lippert, C. M., Kerbrat, P., Sanchez Rovira, P., Wersall, P., Cleall, S. P., Roychowdhury, D. F., Tomlin, I., Visseren-Grul, C. M., and Conte, P. F. (2000) Gemcitabine and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study, *J. Clin. Oncol* 18, 3068–3077.
- (89) Schiller, J. H., Harrington, D., Belani, C. P., Langer, C., Sandler, A., Krook, J., Zhu, J., and Johnson, D. H. (2002) Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer, *N. Engl. J. Med* 346, 92–98.
- (90) Durand-Gasselin, L., Van Rompay, K. K. A., Vela, J. E., Henne, I. N., Lee, W. A., Rhodes, G. R., and Ray, A. S. (2009) Nucleotide Analog Prodrug, Tenofovir Disoproxil, Enhances Lymphoid Cell Loading Following Oral Administration in Monkeys, *Mol Pharm* 6, 1145–1151.
- (91) Hatse, S., Naesens, L., Degève, B., Segers, C., Vandeputte, M., Waer, M., De Clercq, E., and Balzarini, J. (1998) Potent antitumor activity of the acyclic nucleoside phosphonate 9-(2-phosphonylmethoxyethyl)adenine in choriocarcinoma-bearing rats, *Int. J. Cancer* 76, 595–600.
- (92) Neyts, J., Sadler, R., De Clercq, E., Raab-Traub, N., and Pagano, J. S. (1998) The antiviral agent cidofovir [(S)-1-(3-hydroxy-2-phosphonyl-methoxypropyl)cytosine] has pronounced activity against nasopharyngeal carcinoma grown in nude mice, *Cancer Res.* 58, 384–388.
- (93) Muroso, S., Raab-Traub, N., and Pagano, J. S. (2001) Prevention and inhibition of nasopharyngeal carcinoma growth by antiviral phosphonated nucleoside analogs, *Cancer Res.* 61, 7875–7877.
- (94) Andrei, G., Snoeck, R., Piette, J., Delvenne, P., and De Clercq, E. (1998) Inhibiting effects of cidofovir (HPMPC) on the growth of the human cervical carcinoma (SiHa) xenografts in athymic nude mice, *Oncol. Res.* 10, 533–539.
- (95) Park, S., Rath, O., Beach, S., Xiang, X., Kelly, S. M., Luo, Z., Kolch, W., and Yeung, K. C. (2006) Regulation of RKIP binding to the N-region of the Raf-1 kinase, *FEBS Letters* 580, 6405–6412.
- (96) TRAKUL, N., and ROSNER, M. R. (2005) Modulation of the MAP kinase signaling cascade by Raf kinase inhibitory protein, *Cell Res* 15, 19–23.
- (97) Dadvar, P., O’Flaherty, M., Scholten, A., Rumpel, K., and Heck, A. J. R. (2009) A chemical proteomics based enrichment technique targeting the interactome of the PDE5 inhibitor PF-4540124, *Mol. BioSyst.* 5, 472–482.
- (98) Dadvar, P., Kovanich, D., Folkers, G. E., Rumpel, K., Raijmakers, R., and Heck, A. J. R. (2009) Phosphatidylethanolamine-binding proteins, including RKIP, exhibit affinity for phosphodiesterase-5 inhibitors, *Chembiochem* 10, 2654–2662.
- (99) Grandy, D. K., Hanneman, E., Bunzow, J., Shih, M., Machida, C. A., Bidlack, J. M., and Civelli, O. (1990) Purification, cloning, and tissue distribution of a 23-kDa rat protein isolated by morphine affinity chromatography, *Mol. Endocrinol* 4, 1370–1376.
- (100) Atmanene, C., Laux, A., Glattard, E., Muller, A., Schoentgen, F., Metz-Boutigue, M.-H., Aunis, D., Van Dorsselaer, A., Stefano, G. B., Sanglier-Cianféroni, S., and Goumon, Y. (2009) Characterization of human and bovine phosphatidylethanolamine-binding protein (PEBP/RKIP) interactions with morphine and morphine-glucuronides

- determined by noncovalent mass spectrometry, *Med. Sci. Monit* 15, BR178–187.
- (101) Shemon, A. N., Heil, G. L., Granovsky, A. E., Clark, M. M., McElheny, D., Chimon, A., Rosner, M. R., and Koide, S. (2010) Characterization of the Raf kinase inhibitory protein (RKIP) binding pocket: NMR-based screening identifies small-molecule ligands, *PLoS ONE* 5, e10479.
  - (102) Rath, O., Park, S., Tang, H., Banfield, M. J., Brady, R. L., Lee, Y. C., Dignam, J. D., Sedivy, J. M., Kolch, W., and Yeung, K. C. (2008) The RKIP (Raf-1 Kinase Inhibitor Protein) conserved pocket binds to the phosphorylated N-region of Raf-1 and inhibits the Raf-1-mediated activated phosphorylation of MEK, *Cell. Signal* 20, 935–941.
  - (103) Pnueli, L., Carmel-Goren, L., Hareven, D., Gutfinger, T., Alvarez, J., Ganai, M., Zamir, D., and Lifschitz, E. (1998) The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1, *Development* 125, 1979–1989.
  - (104) Chautard, H., Jacquet, M., Schoentgen, F., Bureaud, N., and Bénédicti, H. (2004) Tfs1p, a member of the PEBP family, inhibits the Ira2p but not the Ira1p Ras GTPase-activating protein in *Saccharomyces cerevisiae*, *Eukaryotic Cell* 3, 459–470.
  - (105) Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrom Rev* 16, 1–23.
  - (106) Loo, J. A., Peifeng, H., McConnell, P., Tom Mueller, W., Sawyer, T. K., and Thanabal, V. (1997) A study of Src SH2 domain protein--phosphopeptide binding interactions by electrospray ionization mass spectrometry, *Journal of the American Society for Mass Spectrometry* 8, 234–243.
  - (107) Heck, A. J. R., and Heuvel, R. H. H. van den. (2004) Investigation of intact protein complexes by mass spectrometry, *Mass Spectrometry Reviews* 23, 368–389.
  - (108) van den Heuvel, R. H. H., and Heck, A. J. R. (2004) Native protein mass spectrometry: from intact oligomers to functional machineries, *Curr Opin Chem Biol* 8, 519–526.
  - (109) Potier, N., Rogniaux, H., Chevreux, G., and Van Dorsselaer, A. (2005) Ligand-metal ion binding to proteins: investigation by ESI mass spectrometry, *Meth. Enzymol* 402, 361–389.
  - (110) Sanglier, S., Atmanene, C., Chevreux, G., and Dorsselaer, A. V. (2008) Nondenaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries, *Methods Mol. Biol* 484, 217–243.
  - (111) Tanaka, Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Communications in Mass Spectrometry* 2, 151–153.
  - (112) Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Anal. Chem.* 60, 2299–2301.
  - (113) Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules, *Science* 246, 64–71.
  - (114) Kiselar, J. G., and Downard, K. M. (2000) Preservation and detection of specific antibody--peptide complexes by matrix-assisted laser desorption ionization mass spectrometry, *J. Am. Soc. Mass Spectrom.* 11, 746–750.
  - (115) Strupat, K., Rogniaux, H., Van Dorsselaer, A., Roth, J., and Vogl, T. (2000) Calcium-induced noncovalently linked tetramers of MRP8 and MRP14 are confirmed by

- electrospray ionization-mass analysis, *J. Am. Soc. Mass Spectrom.* 11, 780–788.
- (116) Wattenberg, A., Sobott, F., Barth, H.-D., and Brutschy, B. (2000) Studying noncovalent protein complexes in aqueous solution with laser desorption mass spectrometry, *International Journal of Mass Spectrometry* 203, 49–57.
- (117) Horneffer, V., Forsmann, A., Strupat, K., Hillenkamp, F., and Kubitscheck, U. (2001) Localization of Analyte Molecules in MALDI Preparations by Confocal Laser Scanning Microscopy, *Anal. Chem.* 73, 1016–1022.
- (118) Yamashita, M., and Fenn, J. B. (1984) Application of electrospray mass spectrometry in medicine and biochemistry, *Iyo Masu Kenkyukai Koenshu* 9, 203–206.
- (119) Alexandrov, M. L., Gall, L. N., Krasnov, N. V., Nikolaev, V. I., Pavlenko, V. A., and Shkurov, V. A. (1984) Extraction of ions from solutions at atmospheric pressure, mass spectrometric analysis of bioorganic substances, *Dokl. Akad. Nauk SSR, (Moscow)* 277, 379–383.
- (120) Kebarle, P. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry, *Journal of Mass Spectrometry* 35, 804–817.
- (121) Wilm, M. (1994) Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?, *International Journal of Mass Spectrometry and Ion Processes* 136, 167–180.
- (122) Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source, *Anal. Chem.* 68, 1–8.
- (123) Gomez, A., and Tang, K. (1994) Charge and fission of droplets in electrostatic sprays, *Physics of Fluids* 6, 404.
- (124) Cole, R. B. (1997) *Electrospray ionization mass spectrometry: fundamentals, instrumentation, and applications*, Wiley.
- (125) Dole, M. (1968) Molecular Beams of Macroions, *J. Chem. Phys.* 49, 2240.
- (126) Iribarne, J. V., Dziedzic, P. J., and Thomson, B. A. (1983) Atmospheric pressure ion evaporation-mass spectrometry, *International Journal of Mass Spectrometry and Ion Physics* 50, 331–347.
- (127) Iribarne, J. V. (1976) On the evaporation of small ions from charged droplets, *J. Chem. Phys.* 64, 2287.
- (128) Thomson, B. A., and Iribarne, J. V. (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure, *J. Chem. Phys.* 71, 4451.
- (129) Kebarle, P., and Tang, L. (1993) From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry, *Analytical Chemistry* 65, 972A–986A.
- (130) Wang, Y., Schubert, M., Ingendoh, A., and Franzen, J. (2000) Analysis of non-covalent protein complexes up to 290 kDa using electrospray ionization and ion trap mass spectrometry, *Rapid Communications in Mass Spectrometry* 14, 12–17.
- (131) Sakairi, M., Yergey, A. L., Siu, K. W., Le Blanc, J. C., Guevremont, R., and Berman, S. S. (1991) Electrospray mass spectrometry: application of ion evaporation theory to amino acids, *Anal. Chem.* 63, 1488–1490.
- (132) Gamero-Castano, and Mora. (2000) Kinetics of small ion evaporation from the charge and mass distribution of multiply charged clusters in electrosprays, *J Mass Spectrom* 35, 790–803.
- (133) Kebarle, P., and Peschke, M. (2000) On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions, *Analytica Chimica Acta* 406, 11–35.
- (134) Lemaire, D., Marie, G., Serani, L., and Laprevote, O. (2001) Stabilization of Gas-



- Phase Noncovalent Macromolecular Complexes in Electrospray Mass Spectrometry Using Aqueous Triethylammonium Bicarbonate Buffer, *Analytical Chemistry* 73, 1699–1706.
- (135) Vis, H., Dobson, C. M., and Robinson, C. V. (1999) Selective association of protein molecules followed by mass spectrometry., *Protein Sci* 8, 1368–1370.
- (136) Kebarle, P., and Verkerk, U. H. (2009) Electrospray: From ions in solution to ions in the gas phase, what we know now, *Mass Spectrometry Reviews* 28, 898–917.
- (137) Verkerk, U. H., and Kebarle, P. (2005) Ion-Ion and Ion-Molecule Reactions at the Surface of Proteins Produced by Nanospray. Information on the Number of Acidic Residues and Control of the Number of Ionized Acidic and Basic Residues, *Journal of the American Society for Mass Spectrometry* 16, 1325–1341.
- (138) Smith, R. D., and Light-Wahl, K. J. (1993) The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: Promise, pitfalls and prognosis, *Biological Mass Spectrometry* 22, 493–501.
- (139) Griffey, R. H., Hofstadler, S. A., Sannes-Lowery, K. A., Ecker, D. J., and Crooke, S. T. (1999) Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A* 96, 10129–10133.
- (140) Sannes-Lowery, K. A., Griffey, R. H., and Hofstadler, S. A. (2000) Measuring dissociation constants of RNA and aminoglycoside antibiotics by electrospray ionization mass spectrometry, *Anal. Biochem* 280, 264–271.
- (141) Greig, M. J., Gaus, H., Cummins, L. L., Sasmor, H., and Griffey, R. H. (1995) Measurement of Macromolecular Binding Using Electrospray Mass Spectrometry. Determination of Dissociation Constants for Oligonucleotide: Serum Albumin Complexes, *Journal of the American Chemical Society* 117, 10765–10766.
- (142) Daniel, J. M., McCombie, G., Wendt, S., and Zenobi, R. (2003) Mass spectrometric determination of association constants of adenylate kinase with two noncovalent inhibitors, *J. Am. Soc. Mass Spectrom* 14, 442–448.
- (143) Jørgensen, T. J. D., Roepstorff, P., and Heck, A. J. R. (1998) Direct Determination of Solution Binding Constants for Noncovalent Complexes between Bacterial Cell Wall Peptide Analogues and Vancomycin Group Antibiotics by Electrospray Ionization Mass Spectrometry, *Analytical Chemistry* 70, 4427–4432.
- (144) Metzler, D. E., and Metzler, C. M. (2001) *Biochemistry: the chemical reactions of living cells*, Academic Press.
- (145) Peschke, M., Verkerk, U. H., and Kebarle, P. (2004) Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method, *J. Am. Soc. Mass Spectrom* 15, 1424–1434.
- (146) Wortmann, A., Jecklin, M. C., Touboul, D., Badertscher, M., and Zenobi, R. (2008) Binding constant determination of high-affinity protein-ligand complexes by electrospray ionization mass spectrometry and ligand competition, *J Mass Spectrom* 43, 600–608.
- (147) Feng, R. (1995) Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics, p 1264, Atlanta, GA.
- (148) Cheng, X., Harms, A. C., Goudreau, P. N., Terwilliger, T. C., and Smith, R. D. (1996) Direct measurement of oligonucleotide binding stoichiometry of gene V protein by mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A* 93, 7022–7027.
- (149) Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen,

- F. M., and Dobson, C. M. (1996) Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein–CoA Ligand Binding and Assembly, *Journal of the American Chemical Society* 118, 8646–8653.
- (150) Wu, Q., Gao, J., Joseph-McCarthy, D., Sigal, G. B., Bruce, J. E., Whitesides, G. M., and Smith, R. D. (1997) Carbonic Anhydrase-Inhibitor Binding: From Solution to the Gas Phase, *Journal of the American Chemical Society* 119, 1157–1158.
- (151) Hunter, C. L., Mauk, A. G., and Douglas, D. J. (1997) Dissociation of heme from myoglobin and cytochrome b5: comparison of behavior in solution and the gas phase, *Biochemistry* 36, 1018–1025.
- (152) Kraunsoe, J. A., Aplin, R. T., Green, B., and Lowe, G. (1996) An investigation of the binding of protein proteinase inhibitors to trypsin by electrospray ionization mass spectrometry, *FEBS Lett.* 396, 108–112.
- (153) van Dongen, W. D., and Heck, A. J. R. (2000) Binding of selected carbohydrates to apo-concanavalin A studied by electrospray ionization mass spectrometry, *Analyst* 125, 583–589.
- (154) Sun, J., Kitova, E. N., Sun, N., and Klassen, J. S. (2007) Method for identifying nonspecific protein-protein interactions in nanoelectrospray ionization mass spectrometry, *Anal. Chem* 79, 8301–8311.
- (155) Clark, S. M., and Konermann, L. (2003) Diffusion measurements by electrospray mass spectrometry for studying solution-phase noncovalent interactions, *J Am Soc Mass Spectrom* 14, 430–441.
- (156) Clark, S. M., and Konermann, L. (2004) Determination of Ligand–Protein Dissociation Constants by Electrospray Mass Spectrometry-Based Diffusion Measurements, *Analytical Chemistry* 76, 7077–7083.
- (157) Wang, W., Kitova, E. N., and Klassen, J. S. (2003) Influence of Solution and Gas Phase Processes on Protein–Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Analytical Chemistry* 75, 4945–4955.
- (158) Smith, R. D., Light-Wahl, K. J., Winger, B. E., and Loo, J. A. (1992) Preservation of non-covalent associations in electrospray ionization mass spectrometry: Multiply charged polypeptide and protein dimers, *Organic Mass Spectrometry* 27, 811–821.
- (159) Wang, W., Kitova, E. N., and Klassen, J. S. (2005) Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability, *Anal. Chem* 77, 3060–3071.
- (160) Ganem, B., Li, Y. T., and Henion, J. D. (1991) Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry, *Journal of the American Chemical Society* 113, 7818–7819.
- (161) Sun, J., Kitova, E. N., Wang, W., and Klassen, J. S. (2006) Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry, *Anal. Chem* 78, 3010–3018.
- (162) Sun, N., Soya, N., Kitova, E. N., and Klassen, J. S. (2010) Nonspecific interactions between proteins and charged biomolecules in electrospray ionization mass spectrometry, *J. Am. Soc. Mass Spectrom* 21, 472–481.
- (163) Touboul, D., Maillard, L., Grässlin, A., Moumne, R., Seitz, M., Robinson, J., and Zenobi, R. (2009) How to Deal with Weak Interactions in Noncovalent Complexes Analyzed by Electrospray Mass Spectrometry: Cyclopeptidic Inhibitors of the Nuclear Receptor Coactivator 1-STAT6, *Journal of the American Society for Mass*

- Spectrometry* 20, 303–311.
- (164) Czuczy, N., Katona, M., and Takats, Z. (2009) Selective Detection of Specific Protein-Ligand Complexes by Electrosonic Spray-Precursor Ion Scan Tandem Mass Spectrometry, *Journal of the American Society for Mass Spectrometry* 20, 227–237.
- (165) Loo, R. R. O., Goodlett, D. R., Smith, R. D., and Loo, J. A. (1993) Observation of a noncovalent ribonuclease S-protein/S-peptide complex by electrospray ionization mass spectrometry, *J. Am. Chem. Soc.* 115, 4391–4392.
- (166) Knight, W. B., Swiderek, K. M., Sakuma, T., Calaycay, J., Shively, J. E., Lee, T. D., Covey, T. R., Shushan, B., Green, B. G., and Chabin, R. (1993) Electrospray ionization mass spectrometry as a mechanistic tool: mass of human leucocyte elastase and a beta-lactam-derived E-I complex, *Biochemistry* 32, 2031–2035.
- (167) Huang, H. W., and Wang, K. T. (1996) Structural characterization of the metal binding site in the cysteine-rich region of HIV-1 Tat protein, *Biochem. Biophys. Res. Commun.* 227, 615–621.
- (168) Rogniaux, H., Van Dorsselaer, A., Barth, P., Biellmann, J. F., Barbanton, J., van Zandt, M., Chevrier, B., Howard, E., Mitschler, A., Potier, N., Urzhumtseva, L., Moras, D., and Podjarny, A. (1999) Binding of aldose reductase inhibitors: correlation of crystallographic and mass spectrometric studies, *J. Am. Soc. Mass Spectrom* 10, 635–647.
- (169) Daubenfeld, T., Bouin, A.-P., and van der Rest, G. (2006) A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry, *J. Am. Soc. Mass Spectrom* 17, 1239–1248.
- (170) Schmidt, A., Bahr, U., and Karas, M. (2001) Influence of pressure in the first pumping stage on analyte desolvation and fragmentation in nano-ESI MS, *Anal. Chem* 73, 6040–6046.
- (171) Sun, J., Kitova, E. N., and Klassen, J. S. (2007) Method for stabilizing protein-ligand complexes in nanoelectrospray ionization mass spectrometry, *Anal. Chem.* 79, 416–425.
- (172) Bagal, D., Kitova, E. N., Liu, L., El-Hawiet, A., Schnier, P. D., and Klassen, J. S. (2009) Gas phase stabilization of noncovalent protein complexes formed by electrospray ionization, *Anal. Chem* 81, 7801–7806.
- (173) El-Hawiet, A., Kitova, E. N., Liu, L., and Klassen, J. S. (2010) Quantifying Labile Protein-Ligand Interactions Using Electrospray Ionization Mass Spectrometry, *Journal of the American Society for Mass Spectrometry* 21, 1893–1899.
- (174) Tjernberg, A., Carnö, S., Oliv, F., Benkestock, K., Edlund, P.-O., Griffiths, W. J., and Hallén, D. (2004) Determination of Dissociation Constants for Protein-Ligand Complexes by Electrospray Ionization Mass Spectrometry, *Analytical Chemistry* 76, 4325–4331.
- (175) Vis, H., Heinemann, U., Dobson, C. M., and Robinson, C. V. (1998) Detection of a Monomeric Intermediate Associated with Dimerization of Protein Hu by Mass Spectrometry, *Journal of the American Chemical Society* 120, 6427–6428.
- (176) Korkmaz, B., Moreau, T., and Gauthier, F. (2008) Neutrophil elastase, proteinase 3 and cathepsin G: physicochemical properties, activity and physiopathological functions, *Biochimie* 90, 227–242.
- (177) Korkmaz, B., Horwitz, M. S., Jenne, D. E., and Gauthier, F. (2010) Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases,

- Pharmacol. Rev.* 62, 726–759.
- (178) Moreau, T., Baranger, K., Dadé, S., Dallet-Choisy, S., Guyot, N., and Zani, M.-L. (2008) Multifaceted roles of human elafin and secretory leukocyte proteinase inhibitor (SLPI), two serine protease inhibitors of the chelonianin family, *Biochimie* 90, 284–295.
- (179) Adessi, C., and Soto, C. (2002) Converting a peptide into a drug: strategies to improve stability and bioavailability, *Curr. Med. Chem.* 9, 963–978.
- (180) Zega, A. (2005) Azapeptides as pharmacological agents, *Curr. Med. Chem.* 12, 589–597.
- (181) Epinette, C., Croix, C., Jaquillard, L., Marchand-Adam, S., Kellenberger, C., Lalmanach, G., Cadene, M., Viaud-Massuard, M.-C., Gauthier, F., and Korkmaz, B. (2011) A selective reversible azapeptide inhibitor of human neutrophil proteinase 3 derived from a high affinity FRET substrate, *Biochemical Pharmacology*.
- (182) Specks, U., Fass, D. N., Finkielman, J. D., Hummel, A. M., Viss, M. A., Litwiller, R. D., and McDonald, C. J. (2007) Functional Significance of Asn-linked Glycosylation of Proteinase 3 for Enzymatic Activity, Processing, Targeting, and Recognition by Anti-neutrophil Cytoplasmic Antibodies, *Journal of Biochemistry* 141, 101–112.
- (183) Le Bihan, Y.-V., Angeles Izquierdo, M., Coste, F., Aller, P., Culard, F., Gehrke, T. H., Essalhi, K., Carell, T., and Castaing, B. (2011) 5-Hydroxy-5-methylhydantoin DNA lesion, a molecular trap for DNA glycosylases, *Nucleic Acids Res.* 39, 6277–6290.
- (184) Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000) Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8, *EMBO J* 19, 3857–3869.
- (185) Speina, E., Cieśla, J. M., Graziewicz, M.-A., Laval, J., Kazimierczuk, Z., and Tudek, B. (2005) Inhibition of DNA repair glycosylases by base analogs and tryptophan pyrolysate, Trp-P-1, *Acta Biochim. Pol.* 52, 167–178.
- (186) Amara, P., Serre, L., Castaing, B., and Thomas, A. (2004) Insights into the DNA repair process by the formamidopyrimidine-DNA glycosylase investigated by molecular dynamics, *Protein Sci.* 13, 2009–2021.
- (187) Boiteux, S., O'Connor, T. R., Lederer, F., Gouyette, A., and Laval, J. (1990) Homogeneous *Escherichia coli* FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites, *J. Biol. Chem.* 265, 3916–3922.
- (188) Das, A., Rajagopalan, L., Mathura, V. S., Rigby, S. J., Mitra, S., and Hazra, T. K. (2004) Identification of a Zinc Finger Domain in the Human NEIL2 (Nei-like-2) Protein, *Journal of Biological Chemistry* 279, 47132–47138.
- (189) Ramstein, J., Hervouet, N., Coste, F., Zelwer, C., Oberto, J., and Castaing, B. (2003) Evidence of a Thermal Unfolding Dimeric Intermediate for the *Escherichia coli* Histone-like HU Proteins: Thermodynamics and Structure, *Journal of Molecular Biology* 331, 101–121.
- (190) Cunniff, J., and Vouros, P. (1995) False positives and the detection of cyclodextrin inclusion complexes by electrospray mass spectrometry, *Journal of the American Society for Mass Spectrometry* 6, 437–447.
- (191) Callaghan, O. H., and Weber, G. (1959) Kinetic studies on rabbit-muscle myokinase, *Biochem J* 73, 473–485.
- (192) Jecklin, M. C., Touboul, D., Bovet, C., Wortmann, A., and Zenobi, R. (2008) Which

- electrospray-based ionization method best reflects protein-ligand interactions found in solution? a comparison of ESI, nanoESI, and ESSI for the determination of dissociation constants with mass spectrometry, *J. Am. Soc. Mass Spectrom* 19, 332–343.
- (193) De Weer, P., and Lowe, A. G. (1973) Myokinase equilibrium. An enzymatic method for the determination of stability constants of magnesium complexes with adenosine triphosphate, adenosine diphosphate, and adenosine monophosphate in media of high ionic strength, *J. Biol. Chem* 248, 2829–2835.
- (194) Bowen, W. J., and Kerwin, T. D. (1954) The kinetics of myokinase. I. Studies of the effects of salts and ph and of the state of equilibrium, *Archives of Biochemistry and Biophysics* 49, 149–159.
- (195) Tavel, L., Jaquillard, L., Karsisiotis, A. I., Saab, F., Jouvensal, L., Brans, A., Delmas, A., Schoentgen, F., Cadene, M., and Damblon, C. Ligand binding study of human PEBP1/RKIP: interaction with nucleotides and Raf-1 peptides evidenced by NMR and mass spectrometry., *PlosOne*.
- (196) Zhou, M., Sandercock, A. M., Fraser, C. S., Ridlova, G., Stephens, E., Schenauer, M. R., Yokoi-Fong, T., Barsky, D., Leary, J. A., Hershey, J. W., Doudna, J. A., and Robinson, C. V. (2008) Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3, *Proceedings of the National Academy of Sciences* 105, 18139–18144.
- (197) Yamazaki, K., Iwura, T., Ishikawa, R., and Ozaki, Y. (2006) Effects of Ionic Strength on the Thermal Unfolding Process of Granulocyte-Colony Stimulating Factor, *Journal of Biochemistry* 139, 41–49.
- (198) Horovitz, A., Serrano, L., Avron, B., Bycroft, M., and Fersht, A. R. (1990) Strength and co-operativity of contributions of surface salt bridges to protein stability, *Journal of Molecular Biology* 216, 1031–1044.
- (199) Wolf, K., Wu, Y. I., Liu, Y., Geiger, J., Tam, E., Overall, C., Stack, M. S., and Friedl, P. (2007) Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion, *Nat Cell Biol* 9, 893–904.
- (200) Friedl, P., and Wolf, K. (2003) Tumour-cell invasion and migration: diversity and escape mechanisms, *Nat. Rev. Cancer* 3, 362–374.
- (201) Chambers, A. F., Groom, A. C., and MacDonald, I. C. (2002) Dissemination and growth of cancer cells in metastatic sites, *Nat. Rev. Cancer* 2, 563–572.
- (202) Chabot, V., Reverdiau, P., Iochmann, S., Rico, A., Sénécal, D., Goupille, C., Sizaret, P.-Y., and Sensebé, L. (2006) CCL5-enhanced human immature dendritic cell migration through the basement membrane in vitro depends on matrix metalloproteinase-9, *J. Leukoc. Biol.* 79, 767–778.
- (203) Joosten, V., and van Berkel, W. J. (2007) Flavoenzymes, *Current Opinion in Chemical Biology* 11, 195–202.
- (204) van Berkel, W. J. H., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts, *J. Biotechnol.* 124, 670–689.
- (205) Cashman, J. R. (1995) Structural and catalytic properties of the mammalian flavin-containing monooxygenase, *Chem. Res. Toxicol.* 8, 166–181.
- (206) Silvestro, A., Pommier, J., Pascal, M. C., and Giordano, G. (1989) The inducible trimethylamine N-oxide reductase of *Escherichia coli* K12: its localization and inducers, *Biochim. Biophys. Acta* 999, 208–216.

- (207) Remucal, C. K., and McNeill, K. (2011) Photosensitized Amino Acid Degradation in the Presence of Riboflavin and Its Derivatives, *Environ. Sci. Technol.* *45*, 5230–5237.
- (208) Montaña, M. P., Blasich, N., Haggi, E., and García, N. A. (2009) Oxygen uptake in the vitamin B-sensitized photo-oxidation of tyrosine and tryptophan in the presence of uracil: kinetics and mechanism, *Photochem. Photobiol.* *85*, 1097–1102.
- (209) Posadaz, A., Biasutti, A., Casale, C., Sanz, J., Amat-Guerri, F., and García, N. A. (2007) Rose Bengal-Sensitized Photooxidation of the Dipeptides l-Tryptophyl-l-Phenylalanine, l-Tryptophyl-l-Tyrosine and l-Tryptophyl-l-Tryptophan: Kinetics, Mechanism and Photoproducts¶, *Photochemistry and Photobiology* *80*, 132–138.
- (210) Bertolotti, S. G., García, N. A., and Argüello, G. A. (1991) Effect of the peptide bond on the singlet-molecular-oxygen-mediated sensitized photo-oxidation of tyrosine and tryptophan dipeptides. A kinetic study, *Journal of Photochemistry and Photobiology B: Biology* *10*, 57–70.
- (211) Srivastava, D., Fox, D. A., and Hurwitz, R. L. (1995) Effects of magnesium on cyclic GMP hydrolysis by the bovine retinal rod cyclic GMP phosphodiesterase, *Biochem. J* *308 ( Pt 2)*, 653–658.
- (212) Percival, M. D., Yeh, B., and Falgueyret, J.-P. (1997) Zinc Dependent Activation of cAMP-Specific Phosphodiesterase (PDE4A), ,, *Biochemical and Biophysical Research Communications* *241*, 175–180.
- (213) Zaccolo, M., and Movsesian, M. A. (2007) cAMP and cGMP Signaling Cross-Talk, *Circulation Research* *100*, 1569–1578.
- (214) Zotter, A., Bodor, A., Oláh, J., Hlavanda, E., Orosz, F., Perczel, A., and Ovádi, J. (2011) Disordered TPPP/p25 binds GTP and displays Mg<sup>2+</sup>-dependent GTPase activity, *FEBS Lett.* *585*, 803–808.
- (215) Lai, T. S., Slaughter, T. F., Koropchak, C. M., Haroon, Z. A., and Greenberg, C. S. (1996) C-terminal deletion of human tissue transglutaminase enhances magnesium-dependent GTP/ATPase activity., *The Journal of Biological Chemistry* *271*, 31191–31195.
- (216) Guymer, D., Maillard, J., Agacan, M. F., Brearley, C. A., and Sargent, F. (2010) Intrinsic GTPase activity of a bacterial twin-arginine translocation proofreading chaperone induced by domain swapping, *FEBS J.* *277*, 511–525.
- (217) Reid, J. D., and Hunter, C. N. (2004) Magnesium-dependent ATPase Activity and Cooperativity of Magnesium Chelatase from *Synechocystis* sp. PCC6803, *Journal of Biological Chemistry* *279*, 26893–26899.
- (218) Al-Mulla, F., Bitar, M. S., Feng, J., Park, S., and Yeung, K. C. (2012) A new model for raf kinase inhibitory protein induced chemotherapeutic resistance, *PLoS ONE* *7*, e29532.
- (219) Bradley, D., Carpenter, R., Copsey, L., Vincent, C., Rothstein, S., and Coen, E. (1996) Control of inflorescence architecture in *Antirrhinum*, *Nature* *379*, 791–797.
- (220) Al-Mulla, F., Bitar, M. S., Al-Maghrebi, M., Behbehani, A. I., Al-Ali, W., Rath, O., Doyle, B., Tan, K. Y., Pitt, A., and Kolch, W. (2011) Raf Kinase Inhibitor Protein RKIP Enhances Signaling by Glycogen Synthase Kinase-3{beta}, *Cancer Res.*
- (221) Kumamoto, H., Topalis, D., Broggi, J., Pradère, U., Roy, V., Berteina-Raboin, S., Nolan, S. P., Deville-Bonne, D., Andrei, G., Snoeck, R., Garin, D., Crance, J.-M., and Agrofoglio, L. A. (2008) Preparation of acyclo nucleoside phosphonate analogues based on cross-metathesis, *Tetrahedron* *64*, 3517–3526.
- (222) Coste, F., Hervouet, N., Oberto, J., Zelwer, C., and Castaing, B. (1999) Crystallization

- and preliminary X-ray diffraction analysis of the homodimeric form alpha2 of the HU protein from *Escherichia coli*, *Acta Crystallogr. D Biol. Crystallogr.* 55, 1952–1954.
- (223) Beaufour, M., Vaillant, A. T., Schoentgen, F., and Cadene, M. (2007) Mass spectrometric characterisation of the interaction between PEBP/RKIP and locostatin, Pau, France.
- (224) Gabant, G., Beaufour, M., Schoentgen, F., and Cadene, M. (2008) Towards a dissection of the molecular mechanism of binding of locostatin to the PEBP/RKIP protein, a metastasis suppressor, Grenoble, France.
- (225) Limbach, P. A., Crain, P. F., and McCloskey, J. A. (1995) Molecular mass measurement of intact ribonucleic acids via electrospray ionization quadrupole mass spectrometry, *Journal of the American Society for Mass Spectrometry* 6, 27–39.
- (226) Gabant, G., and Cadene, M. (2008) Mass spectrometry of full-length integral membrane proteins to define functionally relevant structural features, *Methods* 46, 54–61.
- (227) Cadene, M., and Chait, B. T. (2000) A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins, *Anal. Chem.* 72, 5655–5658.
- (228) Mamyrin, B. A., Karataev, V. I., Shmikk, D. V., and Zagulin, V. A. (1973) The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution, *Soviet Journal of Experimental and Theoretical Physics* 37, 45.
- (229) Paul, W., and Steinwedel, H. S. (1960) US Patent 2 939 952.
- (230) Todd, J. F. J. (1991) Ion trap mass spectrometer—past, present, and future (?), *Mass Spectrometry Reviews* 10, 3–52.
- (231) Stafford Jr., G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E., and Todd, J. F. J. (1984) Recent improvements in and analytical applications of advanced ion trap technology, *International Journal of Mass Spectrometry and Ion Processes* 60, 85–98.
- (232) Bieth, J. (1974) Some kinetic consequences of the tight binding of protein-proteinase-inhibitors to proteolytic enzymes and their application to the determination of dissociation constants, pp 463–469, Springer Verlag. Fritz, H., Tschesche, H., Greene, L.J., Truscheit, E., New York, NY.

# Webographie

[a] <http://multalin.toulouse.inra.fr/multalin/>

[b] <http://imed.med.ucm.es/Tools/sias.html>

[c] [http://reflexions.ulg.ac.be/cms/j\\_5740/glossaire?letter=M](http://reflexions.ulg.ac.be/cms/j_5740/glossaire?letter=M)

[d] [http://www.educnet.education.fr/rnchimie/phys/electros\\_m/vanderwaals/vdwaals.htm](http://www.educnet.education.fr/rnchimie/phys/electros_m/vanderwaals/vdwaals.htm)

[e] <http://ead.univ-angers.fr/~jaspard/Page2/COURS/3CoursdeBiochSTRUCT/1EAU/1H2O.htm>

[f] <http://www.molinspiration.com/cgi-bin/properties>



## Lucie JAQUILLARD

### Spectrométrie de masse supramoléculaire: Caractérisation de l'interaction non-covalente entre PEBP1/RKIP humaine et des analogues de nucléotides

#### Résumé :

L'étude des interactions non-covalentes et des relations structure-fonction est à la base de la compréhension des systèmes biologiques. La MS supramoléculaire est une technique de choix pour l'étude des interactions protéine/protéine ou protéine/ligand. Dans le cadre d'études qualitatives ou quantitatives, pour chaque système étudié, les conditions expérimentales et les paramètres instrumentaux ont été optimisés pour conserver le complexe en phase gazeuse (1).

L'objectif principal de ce travail est de caractériser le site nucléotidique de hPEBP1 et de contribuer à la découverte de molécules anti-métastases. Sur le plan fonctionnel, une activité enzymatique de hPEBP1 n'a pas pu être mise en évidence. Pour ce projet, une méthode MS de détermination de  $K_D$  de complexes à faible affinité, plus précise et ne nécessitant pas l'utilisation d'un ligand de référence a été développée (2). Une recherche des déterminants structuraux d'un ligand optimal de hPEBP1 a été réalisée par criblage de composés issus d'une synthèse raisonnée basée sur la structure des nucléotides FMN et GTP et par la détermination de leur  $K_D$  (3). Les criblages ont montré que les critères structuraux indispensables pour la liaison sont la présence d'un groupement chargé ou donneur d'électrons, d'une structure apparentée à une base azotée et d'un cycle additionnel. Une part importante de l'affinité est liée au caractère hydrophobe du ligand. Certains ligands de synthèse ont montré une activité inhibitrice de l'invasion des lignées tumorales.

(1) Epinette *et al.* *Biochem. Pharm.* 2011. (2) Jaquillard *et al.* *J. Am. Soc. Mass. Spectrom.* 2012. (3) Tavel, Jaquillard *et al.* en révision.

Mots clés : spectrométrie de masse, interaction protéine-ligand, non-covalent, hPEBP1, RKIP, antimétastatique, nucléotides,  $K_D$ , faible affinité, criblage, invasion

### Supramolecular mass spectrometry: Characterization of the noncovalent interaction between human PEBP1/RKIP and nucleotide analogs

#### Summary:

The study of noncovalent interactions and structure-function relationships provides the basis for the understanding of biological systems. Supramolecular MS is a favored technique to dissect protein/protein or protein/ligand interactions. In the context of qualitative or quantitative studies, experimental conditions and instrumental parameters have been optimized for each system to preserve the noncovalent complex in the gas-phase (1).

The main objective of this work is to characterize the nucleotide site of hPEBP1 and to contribute to the discovery of antimetastatic molecules. Functionally, a catalytic activity for hPEBP1 could not be detected. For this project, an original MS method to more accurately determine  $K_D$  for low-affinity complexes without a reference ligand was developed (2). Structural features of an optimal hPEBP1 ligand were determined by screening compounds based on FMN and GTP nucleotides in the context of a rational design approach, using  $K_D$  determination to rank affinities (3). Screening highlighted that the essential structural requirements for binding hPEBP1 consist in a charged group or an electron donor, a structure related to a cyclic nitrogenous base and an additional cycle. A significant part of the affinity depends on the hydrophobic nature of the ligand. Some of the synthesized nucleotide analogs are active as inhibitors of invasion in tumor cell lines.

(1) Epinette *et al.* *Biochem. Pharm.* 2011. (2) Jaquillard *et al.* *J. Am. Soc. Mass. Spectrom.* 2012. (3) Tavel, Jaquillard *et al.* in revision.

Keywords : mass spectrometry, noncovalent, hPEBP1, RKIP, antimetastatic, nucleotide,  $K_D$ , low affinity, screening, invasion