



HAL
open science

Etude par Résonance Magnétique Nucléaire de macromolécules biologiques: structure, dynamique et interactions. Application à une protéine de liaison aux odeurs

Ewen Lescop

► **To cite this version:**

Ewen Lescop. Etude par Résonance Magnétique Nucléaire de macromolécules biologiques: structure, dynamique et interactions. Application à une protéine de liaison aux odeurs. Biochimie [q-bio.BM]. Université Paris Sud - Paris XI, 2003. Français. NNT : . tel-00006777

HAL Id: tel-00006777

<https://theses.hal.science/tel-00006777>

Submitted on 31 Aug 2004

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 PARIS XI

Ecole doctorale: "Innovation Thérapeutique: du Fondamental à l'Appliqué"
Pôle: Ingénierie des protéines

Année:

Série Doctorat N°:

THÈSE

présentée et soutenue publiquement le 17 décembre 2003

pour l'obtention du grade de

Docteur de l'Université Paris XI - Orsay

spécialité "Cristallographie et RMN biologiques"

par

Mr LESCOP Ewen

Titre de la thèse:

Étude par Résonance Magnétique Nucléaire de Macromolécules Biologiques : Structure, Dynamique et Interactions. Application à une protéine de liaison aux odeurs.

Composition du jury

<i>Président :</i>	Pr. JANIN Joël	LEBS-CNRS (Gif-sur-Yvette)
<i>Rapporteurs :</i>	Pr. BONVIN Alexandre Dr. SIMORRE Jean-Pierre	Université Utrecht (Hollande) IBS-CNRS (Grenoble)
<i>Examineur :</i>	Pr. PERNOLLET Jean-Claude	INRA (Jouy-en-Josas)
<i>Directeur de thèse :</i>	Dr. GUITTET Eric	ICSN-CNRS (Gif-sur-Yvette)

Je dédie cette thèse à Sun Jie, dont l'amour me comble.

Remerciements

Je tiens tout d'abord à remercier Éric Guittet de m'avoir laissé autant de liberté dans mon travail. J'ai ainsi pu divaguer d'un sujet à un autre et étendre ainsi mes connaissances en RMN. Je lui suis également extrêmement reconnaissant pour les nombreuses discussions scientifiques très intéressantes que nous avons pu avoir.

Ce travail est le fruit d'une collaboration avec le laboratoire de Biochimie et Structure des Protéines de l'INRA (Jouy-en-Josas). Toute ma sympathie va à Loïc Briand, Valérie Bézirard, Claude Nespoulos et les autres membres de l'équipe de Jean-Claude Pernollet, pour les journées passées à la production de la protéine et pour les discussions autour de la table de la cafèt'.

Je remercie Alexandre Bonvin et Jean-Pierre Simorre d'avoir accepté la tâche ingrate d'être rapporteurs de ce travail ainsi que Joël Janin et Jean-Claude Pernollet pour avoir joué le rôle d'examineur.

Mais la science n'avance pas toute seule. J'aurai sans nul doute un souvenir très ému de l'ambiance du laboratoire et des personnes avec qui j'ai partagé ces quatre années ou qui ont été de passage : Nicolas Locker, Jérôme Coutant, Isabelle Lebars, Mounir Traïkia, Marie-Hélène Mazauric, Edith Gincel, Carine Van Heijenoort, Nicolas Birlirakis, Jean-François Gallard, Marie-Thérèse Martin, Alda Goncalves, Dominique Fourmy, Satoko Yoshizawa, Michael Domanski, Géraldine Goussard, Jonathan Farjon, Afaf Mikou, Matthieu Maesani, Stéphanie Cabal, Matthieu Gallopin, Nicolas Soler, Benoît Desmolaize, Muriel Canalia, Clotilde Husson, Emeric Miclet, Rafik Jrad, Stéphanie Juge, Van-Tai Nguyen (dans le désordre). Je leur suis reconnaissant de m'avoir supporté et j'espère que mes discussions sur le respect des cultures minoritaires ne leur ont pas été trop pénibles.

Il est nécessaire entre deux journées de labo de se reposer quelque peu. Il me vient à l'instant une pensée émue aux trois ans de collocation avec Gérald, Pedro, Lili, Axelle et Sun Jie.

La science, c'est bien mais réfléchir occupe l'esprit. La musique, c'est aussi bien et en plus, la pratiquer permet de réfléchir à autre chose. J'en profite donc pour remercier profondément les membres du groupe Deskomp (Gérald, Gaël, David, Remy, Manu et Freddy) pour tout le bonheur que j'ai eu lors des expériences musicales et humaines partagées avec eux. Je tiens à remercier tout spécialement mon « jumeau », Gérald Kergourlay, pour les aventures vécues en parallèle : j'ai énormément progressé à ton contact !

Et pour finir, mais j'aurais pu commencer par eux, je suis extrêmement reconnaissant envers mes parents, Paulette et Jean-Marie, pour m'avoir donné tous les éléments qui m'ont permis d'arriver jusqu'ici : trugarezh bras deoc'h. Un petit mot pour mon frère Morvan et ma soeur Soaz : kendalc'hit war ho hent !

Sommaire

Introduction

Constantes

Liste des abréviations

Partie I Olfaction et OBPs

7

Chapitre 1

L'olfaction : un moyen de communication chimique

- 1.1 La communication 9
 - 1.1.1 La communication est indispensable à la vie 9
 - 1.1.2 L'importance des petites molécules dans la communication 10
 - 1.1.3 La communication chimique et les sens 10
 - 1.1.4 L'abeille 10
- 1.2 Du stimulus odorant à la sensation olfactive ou à la réponse comportementale. 13
 - 1.2.1 Le stimulus chimique 13
 - 1.2.2 L'anatomie de l'organe olfactif 15
 - 1.2.3 Les récepteurs olfactifs (OR) 18
 - 1.2.4 Le codage cellulaire des stimuli 24
 - 1.2.5 Les événements périorécepteurs 27

Chapitre 2

Les OBPs

- 2.1 Les connaissances actuelles sur les OBPs 29
 - 2.1.1 Caractéristiques des OBPs 29
 - 2.1.2 Les OBPs de vertébrés 30
 - 2.1.3 Les OBPs d'insectes 32
 - 2.1.4 Aspects structuraux des OBPs et apparentés d'invertébrés. 35
 - 2.1.5 Des modèles fonctionnels des OBPs 38
- 2.2 ASP2 : une OBP de l'abeille domestique 40
 - 2.2.1 La mise en évidence des OBPs chez l'abeille domestique *Apis mellifera*. 40
 - 2.2.2 Les propriétés d'ASP2 42

2.3	Objectifs de la thèse	43
Bibliographie		45
Partie II Matériel et méthodes		53
Chapitre 1		
Production et purification d'ASP2 marquée en vue de l'étude par RMN		
1.1	Pourquoi utiliser un tel système de surexpression ?	55
1.2	Construction	56
1.3	Surexpression	56
1.4	Purification	56
1.5	Analyse	57
Chapitre 2		
Attribution de la protéine		
2.1	Attribution du squelette peptidique.	59
2.1.1	La stratégie d'attribution	60
2.1.2	Deux expériences hétéronucléaires tridimensionnelles : CBCANH et COCAH	61
2.2	Attribution des chaînes latérales	67
2.2.1	La chaîne aliphatique	67
2.2.2	Les groupes acide et amine des chaînes latérales de Arg, Asn, Asp, Gln et Glu.	68
2.2.3	Les groupes aromatiques	68
Chapitre 3		
Des données RMN à la résolution d'une structure		
3.1	Déplacements chimiques et structures secondaires	69
3.2	Contraintes diédrales	71
3.2.1	Définition des angles dièdres	71
3.2.2	Détermination expérimentale des angles ϕ : HNHA	72
3.2.3	Prédiction des angles ϕ et ψ du squelette : Talos	74
3.3	Les nOes caractéristiques des structures secondaires.	75
3.4	Les liaisons hydrogènes	77
3.4.1	Echange H/D	77
3.4.2	Mise en évidence directe : la HNC0 longue distance	77
3.5	Vers une reconstruction automatique : Aria	79
3.5.1	Reconstruction semi-manuelle d'une protéine	79
3.5.2	La philosophie d'Aria	80
3.5.3	La méthode d'attribution d'Aria	81
3.5.4	Les fichiers nécessaires à Aria	85
3.5.5	Implémentation d'Aria	85

Chapitre 4 Les couplages dipolaires résiduels
--

4.1	Le couplage dipolaire	87
4.2	Couplage dipolaire et orientation faiblement anisotrope	88
4.2.1	Cas du corps rigide	88
4.2.2	Cas d'une molécule flexible	90
4.2.3	Problème de dégénérescence	91
4.3	Les couplages dipolaires dans la pratique	91
4.3.1	Les milieux orientants	91
4.3.2	La mesure des couplages dipolaires résiduels.	93
4.3.3	Détermination du tenseur d'orientation.	94
4.4	Applications des couplages dipolaires résiduels	95
4.4.1	Affinement de structures avec les couplages dipolaires résiduels. . .	97
4.4.2	Orientation de sous-domaines de protéine	98
4.4.3	Identification du repliement de protéines.	98
4.4.4	Module : un programme pour l'orientation de domaines ou de frag- ments.	99
4.4.5	Couplages dipolaires résiduels et dynamique.	99
4.5	Les ondes dipolaires ("Dipolar Waves")	100
4.5.1	L'origine des ondes dipolaires	100
4.5.2	Calcul de l'orientation de l'hélice.	102

Chapitre 5 Analyse de la dynamique interne d'une protéine
--

5.1	Éléments de théorie	103
5.1.1	Mouvements moléculaires	103
5.1.2	La fonction de densité spectrale	104
5.1.3	Mouvement moléculaire et relaxation magnétique	105
5.1.4	Échantillonnage de la fonction de densité spectrale	106
5.1.5	L'effet nOe	109
5.1.6	Effet de l'échange chimique	109
5.1.7	Interprétation de la fonction de densité spectrale en terme de mou- vement	110
5.2	Mise en oeuvre expérimentale : mesures de R_1 , R_2 et de l'effet nOe	116
5.2.1	Les expériences	116

Chapitre 6 Méthodes d'investigation des propriétés d'une protéine
--

6.1	Les expériences filtrées ^{15}N en dimension F2	125
6.2	Titration en pH	126
6.3	Constante d'équilibre et déplacement chimique	126

Partie III Étude par RMN d'ASP2**129****Chapitre 1****Production et bioinformatique**

- | | | |
|-------|---|-----|
| 1.1 | Résultat de la production d'ASP2 simplement marquée (^{15}N) et doublement marquée (^{15}N - ^{13}C) | 131 |
| 1.2 | Analyse de la séquence d'ASP2 | 132 |
| 1.2.1 | Alignement de séquence | 132 |
| 1.2.2 | Structures secondaires | 132 |
| 1.2.3 | Prédiction de structure par TITO | 134 |

Chapitre 2**Attribution d'ASP2**

- | | | |
|-------|--|-----|
| 2.1 | Les premiers spectres d'ASP2 en solution | 137 |
| 2.2 | Les conditions expérimentales de l'attribution d'ASP2 | 137 |
| 2.3 | Attribution des résonances d'ASP2 en présence de TSP à pH 6 | 140 |
| 2.3.1 | Attribution du squelette à 308K et pH 6 | 140 |
| 2.3.2 | Attribution des chaînes latérales | 142 |
| 2.4 | Influence de la température, du TSP et du sel. | 144 |
| 2.4.1 | Effet du TSP | 144 |
| 2.4.2 | Effet de la température et du sel | 145 |
| 2.5 | Effet du pH : ASP2 libre est en équilibre entre deux formes au-dessus de pH 6,6. | 146 |
| 2.5.1 | Titration d'ASP2 libre en solution | 146 |
| 2.5.2 | Titration en pH du complexe ASP2-TSP | 151 |

Chapitre 3**La structure tridimensionnelle d'ASP2 en complexe avec le TSP.**

- | | | |
|-------|--|-----|
| 3.1 | La structure secondaire d'ASP2 | 155 |
| 3.1.1 | Les contraintes diédrales | 155 |
| 3.1.2 | Les liaisons hydrogènes | 159 |
| 3.1.3 | L'effet nOe | 163 |
| 3.2 | Le repliement tridimensionnel d'ASP2 en complexe avec le TSP | 164 |
| 3.2.1 | La modélisation par Aria. | 164 |
| 3.2.2 | La structure d'ASP2 en complexe avec le TSP. | 168 |
| 3.3 | Discussion | 176 |

Chapitre 4**La dynamique d'ASP2 à 308K et à 283K**

- | | | |
|-----|--|-----|
| 4.1 | Les données de relaxation R_1 , R_2 et l'effet nOe hétéronucléaire à 308K. . . | 177 |
| 4.2 | Temps de corrélation à 308K | 181 |
| 4.3 | L'analyse des données de relaxation par le modèle " <i>model-free</i> " à 308K . . | 181 |
| 4.4 | Influence de la température sur la dynamique. | 185 |
| 4.5 | Mise en évidence de l'échange chimique à 308K | 186 |

4.5.1	Comparaison de $R_{2,eff}$ et R_2	186
4.5.2	La région D30 à A38 présente de l'échange chimique	187
4.5.3	CPMG et Modelfree	188
4.6	Structure et dynamique	189
4.6.1	Analyse du temps de corrélation τ_c	190
4.6.2	La dynamique des structures secondaires	190
4.6.3	La dynamique des boucles	191
4.6.4	Ponts disulfure	191
4.6.5	La région D30 à S45	192
4.7	Conclusion	193

Chapitre 5

Étude des deux formes libres d'ASP2

5.1	Les couplages dipolaires et ASP2 libre	195
5.1.1	Mesure des couplages dipolaires résiduels.	195
5.1.2	Analyse des RDCs	196
5.2	Rafinement de la structure d'ASP2 par les RDCs.	198
5.2.1	Objectif de la modélisation	198
5.2.2	Méthode	199
5.2.3	Résultats	199
5.3	Les ondes dipolaires	202
5.4	La dynamique d'ASP2 libre en solution : monomère ou dimère?	204
5.5	Discussion sur les formes A et B	205

Chapitre 6

Étude des interactions d'ASP2 avec le TSP et l'IBMP

6.1	Une expérience "TSP contre ASP2- ^{15}N "	209
6.1.1	Objectif et protocole expérimental	209
6.1.2	Analyse des spectres	211
6.2	IBMP : un bon modèle de molécule odorante	211
6.2.1	Caractéristiques de l'IBMP	212
6.2.2	Attribution de l'IBMP dans l'eau	212
6.3	Une expérience "IBMP contre ASP2- ^{15}N "	214
6.3.1	Protocole expérimental	214
6.3.2	Analyse des spectres	214
6.3.3	Un modèle d'interaction à deux sites	217
6.3.4	Conclusion de l'expérience	219
6.4	Une expérience : "IBMP contre le complexe ASP2- ^{15}N /TSP"	219
6.4.1	Objectif et protocole expérimental	219
6.4.2	Analyse des spectres	219
6.4.3	Discussion	223
6.5	Une expérience "IBMP contre ASP2 ^{15}N - ^{13}C "	224
6.5.1	Objectif et protocole expérimental	224
6.5.2	Analyse des spectres.	224

6.5.3	Discussion	225
6.6	Discussion finale sur les interactions	225
6.6.1	Comment interpréter l'inversion de population entre A et B?	225
6.6.2	RMN versus microcalorimétrie	226
6.6.3	Les sites de fixation de l'IBMP	226
6.6.4	Un mécanisme d'entrée du ligand	227

Chapitre 7**Conclusion générale et perspectives**

7.1	Rappel des observations et des modèles proposés issus de l'étude d'ASP2 par RMN	229
7.2	Des données structurales des OBPs vers leur(s) fonction(s)	230
7.2.1	Comparaison des structures d'ASP2 et d'autres OBPs	230
7.2.2	Les propriétés d'interaction des complexes OBPs/ligands	232
7.3	Conclusion finale	235
7.4	Perspectives	236

Bibliographie **237****Annexe A**

Paramètres d'acquisition et de traitement des expériences 3D.	247
---	-----

Annexe B

Tableau des déplacements chimiques d'ASP2	249
---	-----

Annexe C

Expérience HNHA et Prédiction par Talos	253
---	-----

Annexe D

Protocole de préparation des Bicelles et Valeurs de RDC sur ASP2	257
--	-----

Annexe E

Programmes Octave d'analyse des Ondes Dipolaires	259
--	-----

Annexe F

Données de Relaxation et de Dynamique d'ASP2 à 308K et 283K	261
---	-----

Introduction

Dans nos sociétés post-modernes, le sens olfactif est ramené dans une grande mesure à un moyen de ressentir le beau, au même titre que la vision et l'ouïe. Le stimulus perceptif, qu'il soit visuel, auditif ou olfactif est souvent qualifié en termes esthétiques : un paysage est splendide, une musique divine, une fragrance exquise. La vision et l'ouïe ont été adoptées de tout temps par les artistes pour exciter nos émotions, en créant des oeuvres picturales et musicales. Les plus grands artistes sont reconnus pour savoir provoquer chez le spectateur-auditeur des émotions positives. Certes, des oeuvres telles que Guernica ou les pièces de violons de Philip Glass ont pour effet de créer le mal-être.

Le vingtième siècle a privilégié les sens visuel et auditif grâce au développement des techniques : la télévision s'est imposée dans la quasi-totalité des foyers, diffusant jusqu'à l'excès images et son. Depuis l'essor des supports sonores mobiles (cassette audio, disques compacts), la musique enregistrée est devenue un produit de base, au même titre que l'alimentation, excitant ainsi en permanence nos oreilles. Le sens olfactif est quant à lui sollicité par l'industrie du parfum, tirant profit des progrès importants de la chimie de synthèse au siècle précédent, ainsi que par la cuisine des grands chefs.

On le voit donc, l'Homme a dans son évolution, dévoyé considérablement les perceptions sensorielles de leur but initial : assurer la survie de l'espèce en lui permettant d'interagir avec son environnement immédiat. L'esthétique a pris le pas sur le vital.

Dans cette optique, l'odorat est resté le parent pauvre de la famille des sens, sans doute pour les raisons biomédicales : l'anosmie (privation d'odorat, total ou sélectif) est moins handicapante dans notre monde que la perte de la vision ou de l'ouïe. Les moyens mis en oeuvre pour la compréhension de l'olfaction ont été moindres. Ces études n'ont donc pas profité des développements des techniques industrielles ni des progrès en science fondamentale autant que ses cousins visuels et auditifs, bien que les progrès de la chimie de synthèse ait popularisé l'utilisation des parfums tout en conservant à la parfumerie son caractère artistique et heuristique. Longtemps, l'olfaction fut considérée comme le reliquat d'animalité chez l'Homme. Certains auteurs estiment en effet que la différence principale singularisant l'Homme de ses "voisins" animaux est l'appréhension de l'esthétique. Le fait de sentir une odeur rappelle donc à l'Homme qu'il a longtemps été une bête dénuée d'intelligence et de faculté à saisir le beau. Les codes sociaux occidentaux réprouvent autant la manifestation ostensible du flairage de ses congénères autant que celui des mets présentés à table. Cependant, nous faisons un usage important d'odorants naturels et artificiels pour masquer les odeurs nauséabondes de l'environnement. Cette utilisation de molécule de synthèse mène à la situation où la composition des productions industrielles ou d'un plat alimentaire, dont l'odeur est soit cachée, soit artificiellement exacerbée, reste transparente au consommateur, à moins de disposer d'outils

perfectionnés d'analyse.

Le peu d'intérêt marqué par les scientifiques à l'élucidation des mystères de l'olfaction se traduit par un faible nombre de publications : de la base de données SciSearch en 2000, il ressort que 1400 articles sont associés à l'adjectif « *olfactory* » pour 10 000 à « *visual* ». L'attrait peu enthousiaste porté par le public dans ce domaine est peut-être à l'origine de cette différence. A la décharge des chercheurs, on peut avancer sans risque que la complexité des mécanismes en jeu dans l'olfaction à la fois au niveau moléculaire, neurobiologique que psychologique, est plus profonde que dans le cas de la vision ou de l'ouïe. A partir de trois types de photorécepteurs différents, le cerveau arrive à reconstruire une image visuelle mentale de son environnement avec toutes ses nuances de couleur. La tâche dans le cas de l'olfaction semble plus difficile : les stratégies utilisant la continuité de la physique pour la lumière ou le son ne peuvent se transposer directement pour la détection de molécules chimiques. Or l'Homme arrive à discriminer au minimum 10000 odeurs différentes. La Nature a donc dû développé une méthode originale et complexe pour reconnaître et distinguer les molécules les unes des autres. La découverte des principes fondamentaux régissant la vision et l'ouïe est apparue relativement tôt dans l'Histoire des sciences, car s'ils sont relativement simples, ils ont surtout été accessibles par des moyens d'investigation modestes et facilement disponibles au chercheur du début du vingtième siècle. L'étude de l'olfaction requiert actuellement l'utilisation des techniques de pointe à la fois dans les domaines de l'imagerie, de la biologie moléculaire et de la biologie cellulaire et fait appel aux dernières avancées fondamentales effectuées dans la compréhension de la signalisation cellulaire, de la neurobiologie...

En 2003, beaucoup de zones d'ombre restent encore à explorer pour bien comprendre le mode de fonctionnement de l'olfaction à la fois du point de vue physiologique (intégration des signaux olfactifs au niveau du cortex olfactif) que moléculaire (étapes initiales de reconnaissance des molécules odorantes par les récepteurs olfactifs). L'objet de ce travail de thèse a été d'apporter quelques lumières sur le rôle d'une famille de protéines, appelée OBP (*Odorant-Binding Protein*, ou protéine de liaison aux odeurs) et impliquée sans doute dans la reconnaissance des odeurs par les récepteurs. Nous nous sommes attachés à l'étude d'une OBP d'abeille, dénommée ASP2 et découverte entre autres par Emmanuelle Danty en 1997. Afin de mieux comprendre son rôle dans l'olfaction chez l'abeille, nous avons mis en oeuvre les outils de la Résonance Magnétique Nucléaire afin d'obtenir les renseignements précieux que sont sa structure tridimensionnelle, sa dynamique et ses propriétés d'interaction avec quelques odeurs.

Ce mémoire va rassembler les informations que j'ai pu recueillir sur ASP2. Il s'articulera autour de trois points : présentation des connaissances actuelles sur l'olfaction et les OBPs, les méthodes spectroscopiques mises en oeuvre et les résultats sur ASP2.

Constantes

Perméabilité du vide	$\mu_0 = 4\pi * 10^{-7} TmA^{-1}$
Rapport gyromagnétique de 1H	$\gamma_H = 2,67522 * 10^8 s^{-1}T^{-1}$ $= 42,576375 MHz/T$
Rapport gyromagnétique de ^{15}N	$\gamma_N = -2,7126 * 10^7 s^{-1}T^{-1}$ $=$
Rapport gyromagnétique de ^{13}C	$\gamma_C = 6,7283 * 10^7 s^{-1}T^{-1}$ $=$
Dalton	$1Da = 1g/mol = 1,66.10^{-27} kg$
Constante de Boltzmann	$k_B = 1,38.10^{-23} J.K^{-1}$
Nombre d'Avogadro	$N_a = 6,023.10^{23}$
Permittivité du vide	$\mu_0 = 1,256 H.m^{-1}$
Constante de Planck	$h = 6,626.10^{-34} J.s^{-1}$

Liste des abréviations

- ADR** : Ambiguous Distance Restraint (Contrainte de distance ambiguë)
- ASP** : Antennal Specific Protein
- CPMG** : Séquence d'impulsions Carr-Purcell-Meiboom-Gill
- CSA** : Chemical Shift Anisotropy (Anisotropie de déplacement chimique)
- CSI** : Chemical Shift Index (Index de déplacement chimique)
- CSP** : Chemosensory Protéine (Protéine chimiosensorielle)
- HMQC** : Heteronuclear Multi Quanta Coherence
- HPLC** : High Performance Liquid Chromatography (Chromatographie liquide haute-performance)
- HSQC** : Heteronuclear Single Quantum Coherence
- IBMP** : 2-isobutyl-3-methoxy-pyrazine
- INEPT** : Insensitive Nucleus Enhancement by Polarization Transfert
- NOE** : Nuclear Overhauser Effect
- MUP** : Major Urinary Protein (Protéine majoritaire de l'urine)
- ODE** : Odorant Degrading Enzyme (Enzyme dégradant les odeurs)
- OBP** : Odorant Binding Protein (Protéine de liaison aux odeurs)
- OR** : Olfactory Receptor (Récepteur olfactif)
- OSN** : Olfactory Sensory Neuron (Neurone sensoriel olfactif)
- PBP** : Pheromone Binding Protein (Protéine de liaison aux phéromones)
- PCR** : Polymerase Chain Reaction (Réaction de polymérisation en chaîne)
- RDC** : Residual Dipolar Coupling (Couplage Dipolaire Résiduel)
- RMN** : Résonance Magnétique Nucléaire
- TSP** : sodium 3-Triméthylesilyle Propionate-2,2,3,3-d4

Première partie
Olfaction et OBPs

Chapitre 1

L'olfaction : un moyen de communication chimique

1.1 La communication

1.1.1 La communication est indispensable à la vie

La vie n'est possible que par un état stationnaire d'une population cellulaire mais également au sein même d'une cellule. Cet état est assuré par un ensemble de phénomènes et il suffit qu'un élément de la chaîne soit affaibli pour que l'édifice vacille voire s'effondre. Une mutation ponctuelle sur un récepteur membranaire est suffisante pour l'inactiver et déstabiliser une cellule et mettre en danger le fonctionnement d'un organisme : une mutation du gène du récepteur aquaporine II impliqué dans le transport d'eau dans les cellules rénales provoque par exemple un diabète insipide néphrogénique (dilution de l'urine). Les récepteurs sont les capteurs de l'environnement de la cellule. Sans information sur son environnement, la cellule dépérit rapidement car elle devient incapable de s'adapter et de modifier certains paramètres : pression osmotique, salinité,... De même, les organismes pluricellulaires existent car les milliers de cellules qui le composent sont capables de s'assembler, de rester soudées et d'évoluer. Elles se différencient afin de profiter d'une synergie profitable à l'organisme. Pour cela, elles doivent connaître l'état des autres cellules et la fonction qui lui sera dévolue. Enfin, à l'échelle d'une population d'organismes, une espèce survit et évolue grâce à la reproduction et au contact de ses congénères et de son environnement. A chacune de ces étapes, la communication est essentielle : c'est l'action ou le fait d'établir une relation entre deux ou plusieurs entités.

L'évolution a doté les organismes de plusieurs outils de communication plus ou moins complexes permettant cet échange : les sens. Ils exploitent au mieux les propriétés de la Nature. La physique a donné la vision (sensibilité aux ondes électromagnétiques) et l'ouïe (sensibilité aux ondes acoustiques), le toucher (sensibilité à la pression mécanique). On pourrait également ajouter la sensibilité à la température et au champ magnétique (chez les oiseaux). La chimie a donné l'olfaction et le goût (sensibilité aux molécules chimiques). A chaque sens sont associés des organes spécifiques (oeil, oreille, nez, papilles gustatives,...). En fonction de leurs besoins, les espèces disposent de tels ou tels organes ou les adaptent à leur spécificité. Ainsi, les oiseaux possèdent un organe qui leur permet de se situer dans le champ magnétique terrestre et de se diriger à partir des lignes de champ. La vision est inutile pour la chauve-souris : elle a

développé un organe d'émission et de réception des ultrasons.

1.1.2 L'importance des petites molécules dans la communication

Les petites molécules sont exploitées par la Nature dans de multiples phénomènes de communication à des niveaux très divers. Elles interviennent dans la régulation des cycles cellulaires (facteur de croissance), dans la communication au sein d'un même organisme : neurotransmetteurs (acétylcholine), hormones (oestrogène, testostérone) mais également entre les cellules d'organismes distincts : les spermatozoïdes semblent se diriger vers l'ovule en suivant un gradient de concentration en bourgeonal (Spehr et al. [2003]). A l'échelle de l'organisme supérieur, les molécules sont mises à contribution pour l'échange d'informations sociales intraspécifiques (marquage du territoire, reconnaissance entre la mère et ses petits,...) ou interspécifiques (répulsion de la proie vis-à-vis du prédateur, symbiose parasite-hôte). A cela, il faut ajouter la discrimination par l'odorat du danger (viande en état de décomposition) ou au contraire, d'une substance vitale (nectar de fleur pour l'abeille).

1.1.3 La communication chimique et les sens

Les sens associés à la communication chimique chez les vertébrés sont l'olfaction et le goût. Leur point commun est de faire appel aux molécules et à leurs propriétés chimiques (forme stérique, charge, polarité, groupe fonctionnel). Le rôle de ces sens est de renseigner l'organisme sur l'environnement chimique. Le système nerveux central interprète ensuite le stimulus par une sensation ou une propension à adopter un comportement spécifique. Pour le grand public, la distinction entre olfaction et goût est ténue. On peut cependant les définir chez les vertébrés en terme d'organes et de projections nerveuses dans le cerveau. L'organe nasal est dédié à l'olfaction tandis que les papilles gustatives sont spécialisées dans le goût. Du fait de la proximité et la sensibilité de ces organes, un stimulus peut les exciter de façon non exclusive. Plus généralement, le terme d'olfaction peut être appliqué à des systèmes chimiosensibles qui détectent des molécules chimiques émise par une source éloignée tandis que le goût requiert un contact "physique" avec la source (Eisthen [2002]).

1.1.4 L'abeille

1.1.4.1 Présentation de l'abeille

L'abeille appartient à l'ordre des hyménoptères, comme la guêpe et la fourmi. On la trouve un peu partout dans le monde sauf dans les régions où l'hiver est trop froid. L'abeille domestique, qui produit le miel est appelée *Apis mellifica* (ou *Apis mellifera*).

La plupart des espèces d'abeille sont sociales, la survie de l'individu est donc liée à la bonne marche de la colonie (ruche). Une colonie se compose d'une seule reine, d'ouvrières et de faux bourdons (mâles). La fonction des ouvrières diffère en fonction de leur âge (Biri [1997]) : l'abeille jeune s'occupe de l'entretien de la ruche tandis que les plus âgées récoltent le nectar et le pollen. Tous les "métiers" sont donc représentés : abeille-nourrice, ventileuse, dame d'honneur, cirière, architecte, bâtisseuse, sculpteuses, butineuses, operculatrices, balayeuses, fossoyeurs et gardiennes.

1.1.4.2 Petites molécules et abeilles

Si la vision est nécessaire à la communication entre les abeilles (la fameuse danse des abeilles), l'abeille utilise abondamment les petites molécules pour communiquer avec ses congénères mais aussi pour pouvoir s'orienter dans l'environnement : soit par l'olfaction, soit le toucher, associé au goût.

L'abeille butineuse doit pouvoir reconnaître les fleurs riches en nectar. Elle est capable de distinguer les espèces de fleurs sur la base de leur forme et de leur couleur mais également à partir de l'odorant qu'elles émettent (Wright et al. [2002]). Les plantes diffusent en effet un mélange de substances volatiles qui peuvent différer en nombre, en structures et en concentration. La fleur du muflier (*snapdragon*) contient par exemple un mélange complexe de méthylbenzoate, de β -ocimène, myrcène, linalool, méthyl-cinnamate, 3-hexényl-acétate, acétophénone, diméthoxytoluène, diphényléther et de nérolidol. Von Frisch montra en 1919 (Frisch [1969]) que l'abeille discriminait, après conditionnement, la plupart des paires d'odorants qui lui sont présentés. Laska (Laska et al. [1999]) aboutit à la même conclusion en utilisant des structures relativement variées (alcool, aldéhydes, cétones aliphatiques, terpènes, extrait de menthe, d'orange, de clou de girofle). L'abeille est donc capable de distinguer les fleurs à partir de leur bouquet floral et elle possède des facultés importantes d'apprentissage.

Toute une série de phéromones existe également au sein de la colonie (tableau 1.1). Leurs rôles sont bien définis et elles sont produites par des organes spécialisés. La cohésion sociale est assurée par des phéromones d'alarme, de défense, de recrutement, de reconnaissance de la colonie, de la caste, du sexe, de l'âge, de formation du couvain et de la régulation de la reproduction (Van der Meer et al. [1998], Briand et al. [2002]). La phéromone royale, exclusivement produite par la reine a pour effet de conforter la position sociale de la reine en évitant l'apparition d'une autre reine en inhibant le développement des ovaires des ouvrières. Les butineuses utilisent des phéromones pour le marquage des fleurs. L'effet du marquage peut être soit attractif, soit répulsif : après le passage d'une abeille sur une fleur, une seconde abeille sera soit attirée, soit repoussée. Le subtil jeu semble être lié à la qualité de la production de la fleur et à la durée du renouvellement en nectar dans le but d'optimiser les dépenses énergétiques de la récolte ¹.

1.1.4.3 Pourquoi étudier l'olfaction chez l'abeille

Outre l'intérêt purement fondamental de comprendre les mécanismes généraux de l'olfaction chez un insecte, la maîtrise du comportement des abeilles ainsi que l'optimisation de la production en miel, en cire et en gelée royale a un intérêt économique. En 1996, la production mondiale en miel a atteint 1,2 millions de tonnes. Avec une production d'environ 30 000 tonnes, la France est le premier producteur apicole européen. On comptait la même année 85 000 apiculteurs français dont 2000 à 3000 professionnels (Biri [1997]). D'autre part, le rôle de pollinisation joué par les abeilles est extrêmement important dans l'industrie agroalimentaire.

Grâce une explication moléculaire de l'odorat des abeilles, nous pouvons envisager par exemple d'orienter les abeilles vers un type particulier de plantes afin d'obtenir des miels d'une grande pureté en marquant volontairement les fleurs, ou au contraire en rendant les autres

¹d'après "Le marquage des fleurs visitées par les abeilles et les bourdons", J. Pierre, http://www.beekeeping.com/articles/fr/jacqueline-pierre/marquage_fleurs

Phéromone	Glande	Molécules	fonction
Nasonov	Nasonov	Géranol Acide géranique Acide nérolique (E)-citral, (Z)-citral (E-E)-farnasol nérol	<u>Attractive</u> orientation recrutement marquage d'un site marquage de la fleur ?
Footprint (empreinte de pas)	Anhart (Glande des tarsi)	?	<u>Attractive</u> Orientation marquage d'un site marquage de la fleur ?
Forage-making (marquage floral)	?	?	<u>Attractive</u> orientation marquage de la fleur
Alarme	Glande de Dufour	isoamyl-acétate N-butyl et hexyl acétate benzyl acétate isopentyl alcool N-octyl acétate (Z)-11-eicosen-1-ol	<u>Répulsive</u> alarme et défense marquage de la fleur
Royale	Glande mandibulaire de la reine	acide 9-céto-2(E)-décénoïque acide 9-hydroxy-2(E)-décénoïque méthyl p-hydroxy-benzoate 4-hydroxy-3-hydroxyphényléthanol	<u>Pouvoir de la reine</u>

Tab. 1.1: Récapitulatif des phéromones produites par les abeilles domestiques .

plantes répulsives. La maîtrise de l'agressivité des abeilles ainsi que le choix du lieu d'essaimage sont des applications potentielles à long terme qui faciliteront le travail de l'apiculteur. Et enfin, les abeilles sont essentielles à la pollinisation de nombreuses espèces végétales.

Plus généralement, des produits basés sur les phéromones d'insectes sont déjà utilisés pour contrôler le comportement d'insectes néfastes à l'homme. L'identification des récepteurs olfactifs de l'anophèle, moustique responsable de la malaria, peut mener d'ici à quelques années à un contrôle de la maladie par des composés olfactifs (Gilbert and Firestein [2002]). Cette stratégie a pour avantage de diminuer la diffusion des insecticides, responsables de dysfonctionnement du biotope et actuellement limités par l'apparition de souches résistantes.

D'autre part, nous le verrons, il existe dans le processus de reconnaissance des odorants, des protéines (OBPs) susceptibles d'avoir des applications biotechnologiques : construction de protéines de fusion pour la détection spécifique de petites molécules (par exemple : Schlehner and Skerra [2001]).

1.2 Du stimulus odorant à la sensation olfactive ou à la réponse comportementale.

Schématiquement, la reconnaissance de molécules odorantes se décompose en plusieurs étapes. La molécule odorante traverse le mucus olfactif (ou la lymphe sensillaire des insectes), un ou plusieurs récepteurs situés dans les membranes des neurones olfactifs et couplés à une protéine G est activé et les seconds messagers enclenchent la cascade de réactions menant à la génération d'un potentiel d'action. Le signal est ensuite adressé au bulbe olfactif (ou lobe antennaire) et intégré dans le cerveau central où la sensation prend alors naissance.

Nous allons décrire ces différentes étapes en commençant par quelques précisions sur le stimulus odorant, une description anatomique de l'organe olfactif et en nous intéressant plus longuement aux événements ayant lieu au cours de la conversion molécule-potentiel d'action. La distinction permanente entre les vertébrés et les invertébrés sera également discutée.

1.2.1 Le stimulus chimique

1.2.1.1 Définitions

Les molécules (ou médiateurs) chimiques odorantes sont des signaux émis par un être vivant (animal, végétal) et qui induisent chez un être vivant de la même espèce ou d'autres espèces, une modification du comportement et/ou de la physiologie. On classe les médiateurs chimiques odorants en fonction de leur action :

1. les médiateurs chimiques à action intraspécifique, ou **phéromones** :
 - (a) incitatrices , qui induisent une modification immédiate du comportement (ex. phéromone sexuelle)
 - (b) modificatrices, qui induisent une modification tardive de la physiologie et du comportement, (ex. phéromone royale de l'abeille qui bloque le développement des ouvrières)

Phylum	Classe	Genre et espèces	nom commun
Nématodes	Secernenteans	<i>Caenorhabditis elegans</i>	<i>C. elegans</i>
Arthropodes	Malacostracans	<i>Homarus americanus</i>	homard
	Insectes	<i>Apis mellifera</i>	abeille
		<i>Manduca sexta</i>	papillon
		<i>Drosophila melanogaster</i>	mouche
		<i>Locusta migratoria</i>	criquet migrateur
Vertébrés	Actinopterygians	<i>Brachydanio rerio</i>	poisson-zèbre
		<i>Ictalurus punctatus</i>	poisson-chat
	Amphibiens	<i>Ambystoma tigrinum</i>	salamandre
		<i>Xenopus laevis</i>	xénope
	Mammifères	<i>Homo sapiens</i>	homme
		<i>Mus musculus domesticus</i>	souris
		<i>Rattus norvegicus</i>	rat norvégien
	<i>Cavia porcellus</i>	porc	

Tab. 1.2: Organismes modèles étudiés dans le cadre de l'olfaction (adapté de Eisthen [2002]).

2. les médiateurs chimiques à action interspécifique (allélochimique), comme les signaux impliqués dans la relation abeille-planté et la pollinisation. On distingue :
 - (a) les **allomones**, bénéfiques pour l'espèce qui les émet (la plante)
 - (b) les **kairomones**, bénéfiques pour l'espèce qui les reçoit (l'abeille)

1.2.1.2 Caractéristiques des molécules odorantes

Le support du transport des médiateurs odorants dépend du milieu où vivent l'émetteur et le récepteur. Les animaux complètement aquatiques ne peuvent communiquer que par des molécules pouvant diffuser dans l'eau. Les poissons utilisent ainsi des acides aminés hydrophiles pour "dialoguer". Les espèces terrestres quant à elles, sont contraintes de passer par le biais de molécules diffusant dans l'air, et par conséquent plutôt hydrophobes et volatiles (faible poids moléculaire). En général, ces vecteurs d'information ou de communication sont de petites molécules organiques.

Il a été déterminé plusieurs milliers de molécules différentes possédant chez l'homme des propriétés olfactives. A partir de cette base de données, il a été possible d'établir quelques conditions indispensables pour d'une molécule soit capable d'exciter le système olfactif des espèces terrestres :

1. grande volatilité
2. hydrophobe
3. poids moléculaire inférieur à 310Da

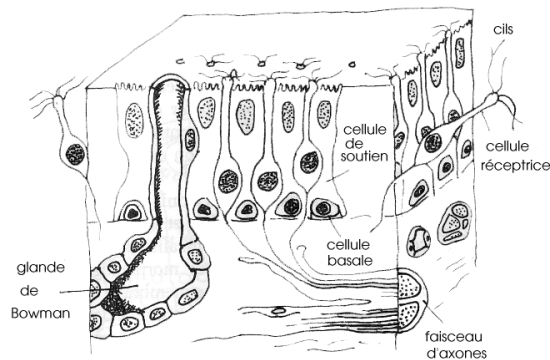


Fig. 1.1: L'épithélium olfactif de vertébré. Zone superficielle de la muqueuse olfactive, l'épithélium contient les neurones récepteurs, les cellules de soutien et les cellules basales. Plus profonde, la sous-muqueuse renferme les glandes à mucus (glande de Bowman), les faisceaux d'axones sensoriels et des vaisseaux sanguins (reproduction Holley [1999]).

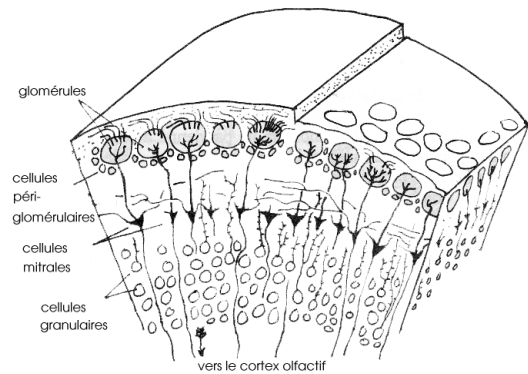


Fig. 1.2: Le bulbe olfactif. Dessin schématisé d'une portion du bulbe olfactif. Les glomérules et les différentes catégories de neurones sont disposés en couches concentriques. Les axones des cellules mitrales, neurones de deuxième ordre, rejoignent le cortex olfactif en formant le tractus olfactif latéral (reproduction Holley [1999]).

1.2.2 L'anatomie de l'organe olfactif

Les recherches sur l'olfaction dont il sera fait état dans la suite sont menées sur des organismes "balayant" une grande partie du phylum. Comme la communication chimique diffère en certains points entre les vertébrés et invertébrés, voire entre les phyla d'invertébrés (nématodes, arthropodes), il convient de faire un rappel sur le classement de ces organismes (tableau 1.2). Plusieurs articles de synthèse font le point sur les convergences et les divergences qui apparaissent au niveau du système olfactif des vertébrés et invertébrés : Hildebrand and Shepherd [1997], Strausfel and Hildebrand [1999], Eisthen [2002].

1.2.2.1 Les mammifères

La plupart des mammifères terrestres possèdent dans les chambres nasales, des systèmes neuronaux spécialisés, au nombre de cinq (Doty [2001]), définis par des projections différentes.

- ⇨ le système olfactif principal,
- ⇨ le système olfactif accessoire ou voméronasal (VNO),
- ⇨ le nerf trijumeau chimiosensible,
- ⇨ l'organe septal (organe de Maser),
- ⇨ le nerf terminal.

Le système olfactif principal est responsable des sensations olfactives (odeur de rose, de pomme, ...). Chez de nombreux mammifères, l'organe voméronasal est spécialisé dans la perception des phéromones (Monbaerts [1999b]). Les sensations somatosensorielles (irritation, brûlure, ...) sont relevées par le système chimiosensible. Il induit également des réactions de type reflex telles que la production de mucus ou l'arrêt de l'inhalation.

Les neurones sensoriels olfactifs (OSNs), au nombre de 1800 environ chez la souris, associés au système principal sont des cellules bipolaires (un axone et une dendrite) situées dans

l'épithélium olfactif (figure 1.1). Les dendrites sont vouées à la détection chimique tandis que les axones sont destinés à la transmission intercellulaire du signal induit. Les dendrites se terminent par des cils, qui s'étend dans le mucus pour couvrir une surface importante (jusqu'à $7m^2$ pour le berger allemand). Les axones sont projetés vers les glomérules, structures (neuropiles) formant avec les cellules gliales le bulbe olfactif (figure 1.2).

Outre des cellules sensorielles, l'épithélium olfactif contient des cellules de soutien qui isolent les cellules sensorielles les unes des autres et régulent l'environnement (ionique) des neurones olfactifs, et des cellules basales qui régénèrent la population neuronale de l'épithélium olfactif.

Le mucus se présente comme un milieu aqueux qui épouse la forme de l'épithélium olfactif. Les glandes de Bowman sont entre autres à l'origine de la sécrétion du mucus et de sa stabilité mécanique (glycoprotéines, olfacoméline). Sa fonction est multiple : défense immunitaire (immunoglobulines), détoxification (cytochrome p450), réservoir d'ions nécessaire à la neurogénèse des neurones olfactifs,...

1.2.2.2 Les invertébrés

Bien que les systèmes chimiosensibles des invertébrés montrent une énorme diversité à travers les genres, une grande similarité apparaît au niveau cellulaire. Dans tous les systèmes olfactifs, de la mouche, au homard, jusqu'au nématode, des neurones sensoriels bipolaires sont utilisés pour la détection des composés odorants. Les neurones sensoriels projettent des dendrites jusqu'à l'environnement interfacial et des axones, directement vers les centres de traitement supérieurs.

Le schéma 1.5 montre en coupe le système olfactif d'un insecte. Les sensilles sont des organules qui recouvrent les branches antennaires par milliers et qui fonctionnent comme des récepteurs : un système de pores tubulaires qui connecte le milieu extérieur à la lumière sensillaire renfermant les dendrites des neurones olfactifs. Chacune répond de façon spécifique à une classe de molécules. Il en existe au moins cinq types (Picimbon [2002]) :

- ↔ sensilles trichodées : impliquées dans la réception de phéromones sexuelles chez le mâle des papillons de nuit,
- ↔ sensilles basiconiques : présentes dans les deux sexes et impliquées dans la reconnaissance de molécules odorantes généralistes,
- ↔ sensilles placodées : plaques olfactives chez les abeilles et les scarabées,
- ↔ sensilles coeloconiques et chaétiques : réparties dans différentes parties du corps et sensibles aux molécules odorantes, au gaz carbonique, à la température et à l'humidité.

Chaque sensille contient quelques neurones et la lumen de chaque sensille est bien isolée de ses voisines. Les neurones olfactifs se prolongent par des dendrites jusqu'aux sensilles olfactives, où elles baignent dans la lymphe sensillaire. Les axones des neurones olfactifs convergent par le nerf antennaire vers les glomérules du lobe antennaire (schéma 1.4), équivalent du bulbe olfactif. Chaque antenne possède un lobe antennaire à sa base, où le traitement primaire du signal olfactif est effectué. Le lobe antennaire est une des trois ganglions formant le cerveau de l'insecte. Le ganglion le plus complexe (le protocerebrum) ainsi que le corps pédonculé reçoivent ensuite le signal du lobe antennaire (Voshall [2001]).

Le nombre de neurones chimiosensibles varie fortement : de 20 à 30 pour *C. elegans* à 1000 pour la drosophile (Tromel [1999]). Chez les insectes, les neurones olfactifs

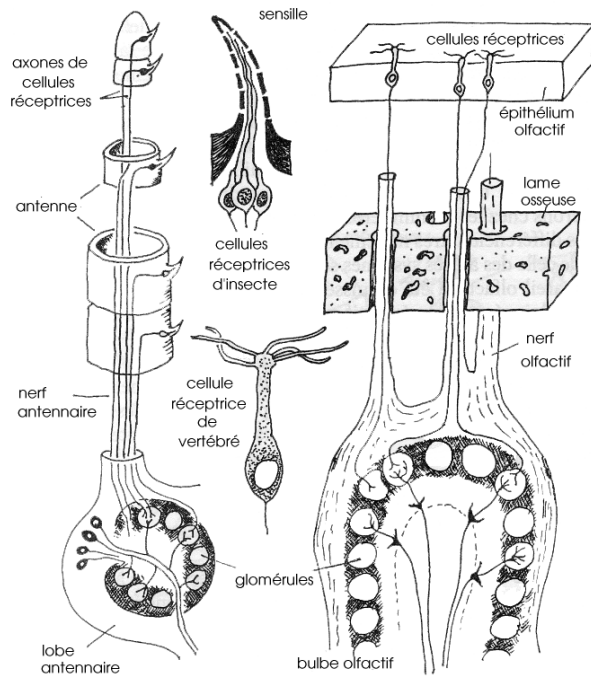


Fig. 1.3: Similitude d'organisation de l'entrée de la voie olfactive chez un insecte (à gauche) et un vertébré (à droite). Les axones des neurones sensoriels convergent dans les glomérules où ils font synapse avec des neurones de deuxième ordre (reproduction Holley [1999]).

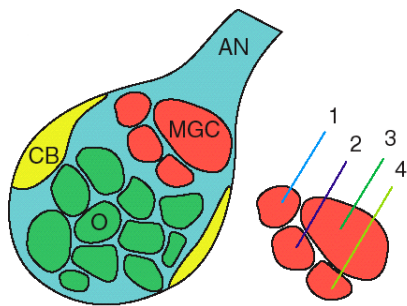


Fig. 1.4: Le lobe antennaire du papillon. Les axones provenant des antennes de l'insecte arrivent par le nerf antennaire (AN). Deux types de glomérules sont présents : les glomérules ordinaires (O), sexuellement isomorphiques et les complexes macroglomérulaires spécifiques des mâles (MGC), innervés par les neurones spécialisés dans la reconnaissance des phéromones sexuelles. Les glomérules sont flanquées de corps cellulaires d'interneurones (CB) (reproduction Hansson [2002]).

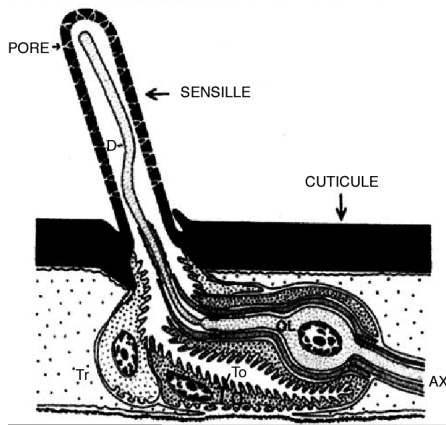


Fig. 1.5: Sensille olfactive d'un insecte. La lymphe sensillaire, dans laquelle baignent les dendrites (D), est sécrétée par cellules trichogènes (Tr) et tormogènes (To). Les cellules thécogènes ont le rôle de cellule gliale pour les neurones olfactifs. Chaque sensille contient les projections dendritiques de un à quatre neurones olfactifs (adapté de Kim and Smith [2001])

sont spécialisés : certains sont spécifiques des phéromones tandis que les autres détectent les odorants de plante (pour revue Hansson [2002]). Ces derniers sont peu nombreux.

1.2.3 Les récepteurs olfactifs (OR)

Cette mise au point bibliographique est fortement inspirée des revues de Monbaerts : Monbaerts [1999b,a, 2001].

1.2.3.1 La découverte des récepteurs olfactifs

L'élégance de la méthodologie mise en oeuvre dans la découverte des récepteurs olfactifs mérite qu'on s'y attarde. A la fin des années 1980, l'étude des mécanismes de la transduction des signaux olfactifs a débouché sur la mise en évidence d'une première cascade enzymatique impliquant l'AMP cyclique (cAMP) dans le contrôle du canal ionique responsable de la dépolarisation de la membrane des neurones olfactifs, qui conduit à la réponse électrique propagée. Puis une deuxième voie (IP3) fut découverte. Une protéine G (liant le GTP), appelée G_{olf} , participe au processus de transduction. Linda Buck et Richard Axel (Buck and Axel [1991]) en déduisirent que les récepteurs olfactifs devaient appartenir à la superfamille des récepteurs couplés aux protéines G. Ces récepteurs, implantés dans les membranes cellulaires, sont impliqués dans la communication cellulaire par leur faculté à lier les hormones, les neurotransmetteurs, les chimiokines et les ions calcium (pour revue, Pierce et al. [2002]) et contiennent sept hélices transmembranaires (superfamille 7TM) reliées par trois boucles extracellulaires et trois boucles intracellulaires de longueurs variables (figure 1.6). D'autre part, les récepteurs recherchés devaient appartenir à une famille multigénique pour expliquer la détection et la discrimination de structures chimiques si différentes et si nombreuses. Ils amplifièrent alors par PCR les ADNc préparés à partir des ARN messagers présents dans l'épithélium olfactif du rat. L'utilisation d'amorces dégénérées d'oligonucléotides correspondant aux deuxième et septième régions transmembranaires de récepteurs connus devait sélectionner les récepteurs 7TM. L'analyse des résultats montre l'existence d'une famille multigénique de récepteurs. Le fait que les récepteurs identifiés soient spécifiquement exprimés dans les cellules neuronales de l'épithélium olfactif a suggéré leur fonction de liaison aux odorants.

Après le rat, des familles de gènes de récepteurs olfactifs ont été découvertes dans une vingtaine d'espèces de vertébrés (mammifères, poissons, amphibiens, oiseaux). D'autre part, des

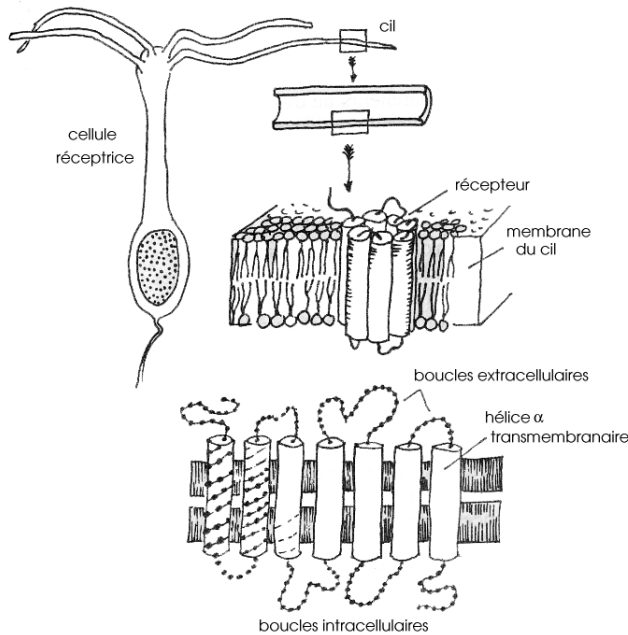


Fig. 1.6: Récepteurs 7TM : localisation et structure.

récepteurs olfactifs putatifs ont été mis en évidence par bioinformatique dans deux invertébrés (le nématode *C. elegans* et la drosophile). Des études d'expression fonctionnelle ont montré que certains répondaient à des odorants.

Les doutes sur la fonction olfactive de ces récepteurs ont été levés par des expériences menées par plusieurs équipes (Zhao et al. [1998], Zhang et al. [1997]) : en insérant les gènes codant pour ces récepteurs dans des systèmes hétérologues, ils ont observé une activité électrique spécifique de molécules odorantes.

1.2.3.2 Caractéristiques des ORs.

Gènes des ORs de vertébrés La complexité du répertoire des ORs est estimée à 1000 gènes chez la souris et le rat, à 350 chez l'homme (estimée après le séquençage du génome, compte tenu des pseudo-gènes) et à 100 gènes chez les poissons zèbre et chat (Axel [1995]). Ils représentent donc la famille la plus importante du génome des mammifères (1 – 2%). Si une partie des récepteurs 7TM s'expriment dans les neurones du système olfactif, on trouve deux familles également dans l'organe voméronasal (récepteurs voméronasaux, VR, Buck [2000]). Monbaerts met en garde contre une mauvaise estimation de ces nombres car d'autres familles de récepteurs non 7TM pourraient exister et une partie des récepteurs 7TM identifiés pourraient avoir une autre fonction que la perception olfactive (Monbaerts [1999b]).

Les ORs de vertébrés partagent la propriété de ne pas contenir d'intron, ce qui suggère que la diversité des récepteurs n'est pas obtenue par épissage alternatif. La présence de motifs conservés (de 2 à 10 acides aminés) dans les boucles intracellulaires IC1 et IC2 et les domaines transmembranaires TM3, TM5, TM6 et TM7 suffit pour classer un récepteur 7TM dans la famille des récepteurs olfactifs. Au contraire, les régions hypervariables sont situées dans les régions TM3, 4 et 5, susceptibles de former la poche d'interaction avec les ligands. D'autre part, on note que la seconde boucle extracellulaire est particulièrement longue et contient une

paires de cystéines fortement conservée. La superfamille des récepteurs se divise en familles et sous-familles d'après leur homologie de séquence, allant de 40% à 90% d'identité de séquence (Firestein [2001]).

D'autre part, les gènes des ORs sont disséminés dans le génome, sans doute une conséquence de l'évolution. Les ORs humains possèdent une originalité par rapport aux autres espèces. Le répertoire des ORs contient environ 50% de pseudogènes dus sans doute à un décalage du cadre de lecture, des mutations non-sens et/ou des délétions. La pression de la sélection est sans doute à l'origine de cette dégénérescence du répertoire, la sensibilité olfactive de l'homme étant très faible.

La découverte d'une classe de récepteurs chez le poisson chat a mené à la définition de deux classes :

- ↔ classe I : récepteurs susceptibles de reconnaître des ligands hydrophiles
- ↔ classe II : récepteurs susceptibles de reconnaître des ligands hydrophobes

La présence de récepteurs de classe I chez des vertébrés terrestres peut être interprétée comme un reliquat de l'évolution. De tels récepteurs existent chez l'homme et la souris. Cependant, des récepteurs de classe I répondent à des odorants aliphatiques volatiles et sont exprimés non spécifiquement dans le système olfactif. La question reste donc ouverte.

OR et OSN Les récepteurs olfactifs sont exprimés dans les terminaisons dendritiques des neurones olfactifs (OSN) au niveau des cils. Un récepteur s'exprime dans un nombre très faible de sous-ensemble de OSNs. Chez les vertébrés, il a été montré que chaque OSN exprime un type de récepteur, ou au plus quelques-uns (Dryer and Berghard [1999]). Par conséquent, l'hypothèse un récepteur-un neurone est de plus en plus admise (Malnic et al. [1999], Axel [1995]). Chez le rat et la souris, à l'exception d'une sous-famille, chaque récepteur est exprimé dans un nombre limité de neurones olfactifs (OSN) délimitant une zone de l'épithélium olfactif (il y en a quatre au total). A l'intérieur d'une zone, les OSNs exprimant des ORs distincts sont enchevêtrées pour former une mosaïque.

Comme un OSN donné répond à un spectre large d'odorants et qu'une molécule donnée active plusieurs glomérules, chaque récepteur doit donc reconnaître une multitude de molécules. Le code olfactif doit donc être de nature combinatoire.

La présence de transcrits d'ORs dans le bulbe olfactif, un site inaccessible aux molécules odorantes, est étonnante. Les ORs pourraient avoir une autre fonction : ils seraient utilisés pour le ciblage des axones des neurones olfactifs vers le bulbe olfactif durant l'ontogénèse (pour revue, St.John et al. [2002]) pour aboutir à la formation du système olfactif. Cette propriété est d'autant plus intéressante que les neurones olfactifs se renouvellent sans cesse durant la vie adulte (Doty [2001]) et que la carte odotopique du bulbe olfactif doit restée semblable au cours du temps.

1.2.3.3 Le cas des invertébrés

Les récepteurs olfactifs découverts chez *C. elegans* forment une superfamille indépendante de plus de 500 récepteurs (pour revue, Tromel [1999]). Les gènes ne présentent aucune homologie de séquence avec les récepteurs de vertébrés. D'autre part, il semble que plusieurs gènes s'expriment dans une même cellule sensorielle et que la fonction des récepteurs de *C. elegans* soit multiple : attraction, répulsion, détection des phéromones.

La drosophile adulte, quant à elle, possède au moins 60 types de récepteurs sans homologie avec les récepteurs de vertébrés, ni même entre eux (pour revue, Vosshall [2001]). Comme pour les vertébrés, il semble que chaque OSN exprime un seul type de récepteur à l'exception d'un récepteur qui s'exprime dans presque toutes les cellules sensorielles.

1.2.3.4 ORs et odorants

Les dernières études suggèrent que chaque récepteur est capable de fixer plusieurs odorants, et que chaque odorant se lie à plusieurs récepteurs.

Point de vue expérimental

Contrairement à la plupart des autres récepteurs couplés à des protéines G, chaque récepteur olfactif a un large spectre de fixation et est capable de lier plusieurs odorants (avec des spécificités et des affinités différentes). De plus, chaque molécule odorante active plusieurs récepteurs (Malnic et al. [1999] , Kajiyama et al. [2001]). Le codage des odorants est affecté par une légère différence dans la structure de la molécule aussi bien que par sa concentration. Bien entendu, la corrélation entre odorants et récepteurs n'est connue que pour un nombre limité de récepteurs et sur un spectre restreint de stimuli.

Les prérequis moléculaire à l'interaction récepteur-odorant ne sont pas clairement établis. On peut toutefois citer les études d'Araneda (Araneda et al. [2000]) qui rapporte la nécessité de la présence d'un groupement aldéhyde sur la petite molécule pour activer le récepteur I7 du rat et de Gaillard (Gaillard et al. [2002]), qui montre que le récepteur OR 912-93 murin ne se fixe qu'aux cétones aliphatiques. Sur la base de ces résultats, l'opinion consensuelle est que chaque récepteur reconnaît un motif structural propre à une famille d'odorants (odotopes) à la manière de l'interaction antigène-anticorps (voir paragraphe 1.2.3.6).

Site de fixation : une approche par la modélisation

L'étude de l'interaction au niveau moléculaire entre récepteurs et odorants est difficile à mettre en oeuvre du fait de l'absence de structure tridimensionnelle expérimentale disponible et de la difficulté d'expression des récepteurs en système hétérologue sous forme fonctionnelle en vue de mutagenèse dirigée. La structure des récepteurs est donc en général déduite de la structure cristallographique de la rhodopsine bovine, du profil d'hydrophobicité et d'alignement de séquence.

Malgré cela, les premiers travaux de Buck et Axel (Buck and Axel [1991]) proposant, sur la base de l'hypervariabilité, les domaines TM3, 4 et 5 comme site de fixation du récepteur de rat, se confirment au fur et à mesure des études de modélisation.

L'étude de Pilpel (Pilpel and Lancet [1999]) basée sur un alignement de séquences et la modélisation de la structure tridimensionnelle des récepteurs, a identifié 17 positions hypervariables situées sur les domaines TM3, 4 et 5. Ces résidus sont regroupés sur les faces des 3 hélices orientées vers l'intérieur du tonneau. Le site d'interaction défini au deux-tiers extracellulaire du tonneau est donc susceptible de pouvoir lier les odorants.

Grâce à la modélisation par arrimage moléculaire, Floriano (Floriano et al. [2000]) a identifié une série d'odorants potentiellement ligands d'un récepteur murin ainsi que le site de fixation (hélices TM3, 4 et 5). Par contre, dans un article plus récent (Vaidehi et al. [2002]), le

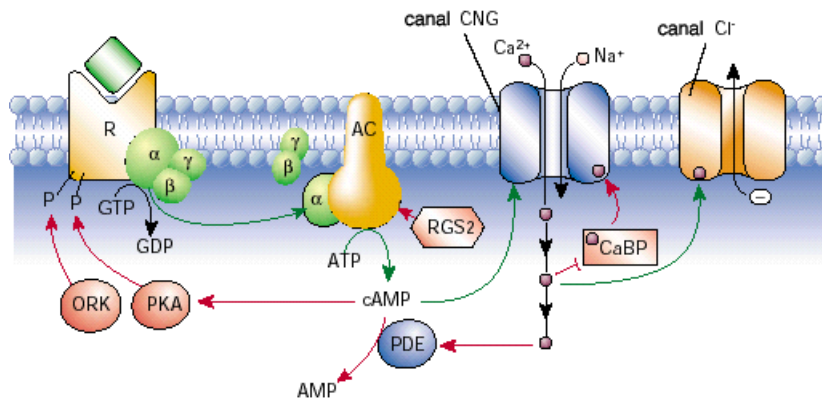


Fig. 1.7: Transduction sensorielle. Une cascade d'activités enzymatiques transduit la liaison d'une molécule à un récepteur en un signal électrique. C'est une transduction par un nucléotide cyclique dans laquelle toutes les protéines ont été identifiées, clonées, exprimées et caractérisées. AC : adénylyl cyclase ; canal CNG : canal contrôlé par un nucléotide cyclique ; PDE : phosphodiésterase ; PKA : protéine kinase A ; ORK : kinase du récepteur olfactif ; RGS : régulateur de la protéine G ; CaBP : protéine liant la calmoduline. Les flèches indiquent une activité stimulante (vert) ou inhibitrice (rouge). Reproduction de Firestein [2001].

site de fixation des récepteurs I7 de souris et de rat, déterminé par bioinformatique implique les hélices TM3,4 et 6.

1.2.3.5 La transduction du signal olfactif.

Une fois que le récepteur a lié une molécule, une cascade d'événements est initiée et transforme l'énergie chimique de liaison en un signal électrique en fonction de la qualité, la force et la durée des stimuli. Il existe plusieurs voies de signalisation à partir des protéines G. Dans le cadre de l'olfaction, les vertébrés semblent adopter deux voies préférentielles : la voie cAMP (AMP cyclique) et la voie IP₃ (inositol-triphosphate). Nous allons présenter la voie cAMP (figure 1.7).

Le récepteur active une protéine G (G_{olf}), qui active une adénylyl cyclase (AC). La cyclase convertit l'ATP cellulaire en AMP cyclique. Le cAMP se lie alors sur la face intracellulaire d'un canal ionique calcium/sodium (CNG). Ces ions peuvent alors traverser la membrane. Le potentiel électrique entre l'intérieur et l'extérieur du neurone olfactif au repos est de -65mV . Par le transfert ionique, le potentiel devient moins négatif. Si le potentiel atteint un seuil critique, un potentiel d'action se propage le long de l'axone jusqu'aux connections synaptiques dans le glomérule.

Cette cascade de réactions a pour effet d'amplifier et d'intégrer le stimulus. Il semble même qu'une seule molécule soit capable d'exciter un OSN, même si le signal n'est plus mesuré au niveau du cerveau. Le mécanisme de transduction du signal olfactif possède un second processus d'amplification : l'entrée de calcium provoque l'ouverture d'un canal chlore. La concentration en ion chlorure des OSNs est particulièrement élevée si bien que l'action de ce canal amplifie d'autant plus la dépolarisation.

La régulation de la sensibilité de ce mécanisme est assurée par l'entrée du calcium. Le cal-

cium active une protéine (sans doute la calmoduline) qui inhibe l'activité du canal CNG. Pour conserver l'activité électrique de la cellule, la "quantité" de stimulus doit alors constamment augmenter pour augmenter la concentration intracellulaire en cAMP.

Le cas des invertébrés est un peu plus complexe (Krieger and Breer [1999]) et moins bien compris actuellement. Les mécanismes de transduction ont été bien étudiés chez le homard, arthropode aquatique. Les deux voies cAMP et IP3 sont présentes dans une même cellule. Ils provoquent tous deux la création d'un potentiel d'action mais par des moyens opposés (hyperpolarisation et dépolarisation). Si, comme *C. elegans*, plusieurs récepteurs s'expriment dans un même neurone de homard, cette observation suggère que la cellule sensorielle agit comme une unité complexe d'intégration. Bien que peu d'information soit disponible sur la transduction des insectes (drosophile) ou de *C. elegans*, les deux voies précitées y ont apparemment un rôle également à jouer. Bien entendu, des différences importantes de mécanisme peuvent apparaître car les milieux dans lequel évoluent ces animaux sont très divers.

1.2.3.6 La relation structure-odorant : des approches théoriques.

La relation structure-odorant n'est pas véritablement établie : il s'agit de déterminer, voire de prédire la réponse d'un individu de telle espèce à un stimulus. On peut déjà préciser que la réponse est dépendante de l'espèce : l'isoamyl-acétate évoque pour l'humain une odeur de banane et représente pour l'abeille une phéromone d'alarme. Elle peut même être dépendante de l'individu au sein d'une espèce. Plusieurs théories s'affrontent pour discriminer les facteurs importants dans la reconnaissance, l'identification et l'effet de telle ou telle molécule mais sans pour autant satisfaire l'ensemble des observations expérimentales. Nous pouvons citer d'après la revue de L. Turin et F. Yoshii (Turin and Yoshii [2003]) :

la théorie stérique (Amoore [1964]) :

les caractéristiques olfactives des molécules sont définies par leur forme stérique et leur taille. Deux molécules ayant des encombrements stériques proches ont des propriétés olfactives identiques tandis que des molécules de volumes très différents seront bien discriminées. Les récepteurs suivent alors le modèle clé-serrure.

la théorie vibrationnelle (Wright [1982]) :

le récepteur olfactif est un spectromètre à infra-rouge moléculaire extrêmement fin, capable de distinguer des différences de vibrations moléculaires très faibles.

la théorie spectroscopique (Turin [1996]) :

basée sur la spectroscopie à effet tunnel électronique inélastique, cette théorie repose sur l'activation du récepteur olfactif par un transfert d'électron provenant de NADPH. En l'absence de molécule odorante au site de liaison, ce transfert ne s'effectue pas. En présence d'une molécule, de nouveaux niveaux d'énergie sont accessibles aux électrons, ce qui autorise leur transfert par effet tunnel du NADPH vers l'accepteur (le pont disulfure entre le récepteur et la protéine G). Le récepteur agit ainsi comme un spectromètre, n'étant activé que si la molécule contient un niveau d'énergie vibrationnelle égal à la différence d'énergie entre les donneurs et accepteurs d'électrons.

la théorie de l'odotope (Mori and Sheperd [1994]) :

de la même manière que les anticorps fixent spécifiquement l'épitope de l'antigène, un récepteur olfactif est capable de reconnaître un site, appelé odotope, sur une molécule.

Chaque molécule peut ainsi être reconnue par plusieurs récepteurs et inversement, chaque récepteur reconnaît plusieurs molécules présentant un fragment identique.

Ces modèles prennent en compte plus ou moins bien les caractéristiques suivantes des molécules :

- ↔ groupes fonctionnels (ex. molécule soufrée),
- ↔ remplacement isotopique (le cyclohexanone repousse fortement les blattes tandis que le cyclohexanone deutéré est inactif)
- ↔ spectre vibrationnel IR (deux molécules de même spectre peuvent avoir une odeur différente)
- ↔ énantiomérisation (la S-carvone a une odeur de cumin tandis que la R-carvone a une odeur de menthe verte)

A l'heure actuelle, la théorie de l'odotope reçoit des suffrages les plus nombreux. Cette interaction récepteur-ligand est plus générale que la simple olfaction car elle sert de paradigme pour la recherche pharmaceutique en vue de développer des inhibiteurs. Les théories spectroscopiques, bien que séduisante, doit encore faire ses preuves.

1.2.4 Le codage cellulaire des stimuli

1.2.4.1 Le traitement de l'information au niveau du bulbe olfactif

Le mécanisme du codage olfactif commence à être compris dans le règne des vertébrés (pour revue, Mori et al. [1999], Buck [2000]). Le système olfactif central reçoit l'information des molécules odorantes par les axones des neurones olfactifs. L'information y est processée et intégrée pour former des images olfactives.

Le bulbe olfactif humain possède une structure corticale relativement simple contenant des milliers de modules de traitement du signal, appelés glomérules. Les axones des neurones olfactifs y forment des connections synaptiques avec les dendrites des cellules mitrales et des cellules à panache (figure 1.2). Ces cellules de second ordre sont les neurones de sortie du bulbe olfactif et forment le tractus olfactif qui rejoint le cortex olfactif.

D'un point de vue fonctionnel, on sait depuis une dizaine d'années (Freeman [1991]) qu'un odorant est codé dans l'ensemble du bulbe olfactif et non pas au niveau de quelques neurones. En analysant les électroencéphalogrammes, la présence d'odorant induit une impulsion électrique de 20 à 90Hz (onde γ) qui est distribuée dans tout le volume du bulbe.

Chaque glomérule (1800 chez la souris) peut être vu comme le point de convergence des OSNs associés au même type de récepteurs (1000 gènes environ) : c'est la convergence glomérulaire. En contrepartie, il est possible que chaque glomérule reçoive en entrée les axones de neurones exprimant différents récepteurs. Les récepteurs sont exprimés dans des OSNs répartis dans quatre zones et cette distribution semble être corrélée avec l'homologie de séquence entre récepteurs. Cette distribution spatiale apparaît conservée au niveau du bulbe olfactif (projection régiono-régionale).

De part la projection spécifique des dendrites des cellules mitrales et à panache vers un seul glomérule, la spécificité des glomérules se reflète sur celle de ces cellules qui l'innervent. Cela signifie que chaque cellule mitrale ou à panache présente une réponse en fonction des caractéristiques structurales des stimuli. Ces observations appuient la théorie de l'odotope : une

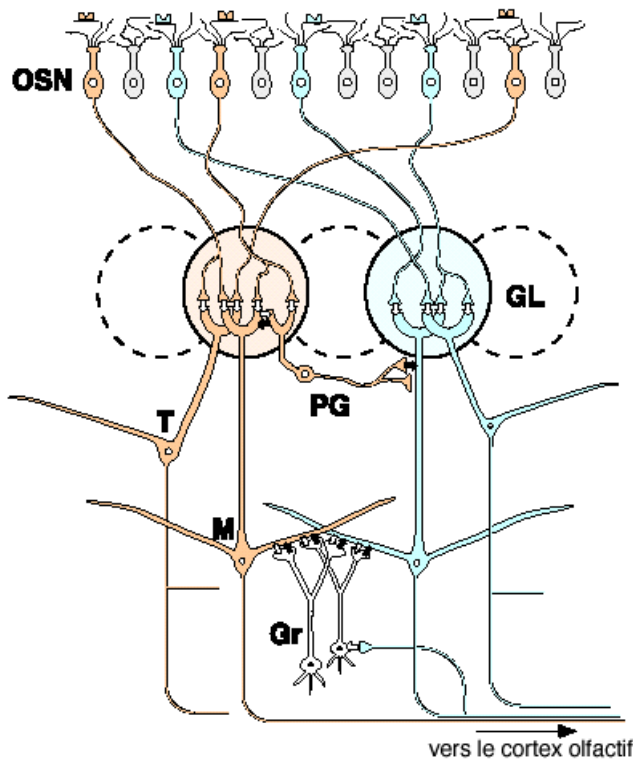


Fig. 1.8: Organisation synaptique du bulbe olfactif des mammifères. Les deux glomérules (en bleu et marron) expriment deux types de récepteurs. Les cellules mitrales (M) et à panache (T) sont les neurones de sortie et les cellules granulaire (Gr) et périglomérulaires (PG) sont des neurones de liaisons locales. OSN : neurone sensoriel olfactif, GL : glomérule. Les petites flèches dénotent des synapses d'excitation (blanc) et d'inhibition (noir). (Reproduction de Mori et al. [1999]).

molécule donnée possède plusieurs motifs structuraux reconnus par une série de récepteurs. Elle active par ce biais une série de glomérules dans le bulbe olfactif, donnant une représentation spatiale caractéristique d'un odorant. Un mélange d'odorants activerait alors une carte donnant une image "olfactive". La carte olfactive au niveau du bulbe olfactif suggère que les caractéristiques d'une molécule sont classées en deux parties : les caractéristiques primaires (groupe fonctionnel et position sur la molécule) et les caractéristiques secondaires (longueur de la chaîne). Si les premiers activent des domaines glomérulaires délimités du bulbe olfactif, les seconds activent des glomérules dispersés à travers ces domaines (Nagao et al. [2002]).

Bien entendu, cette représentation est simpliste. En effet, le bulbe est probablement le lieu d'un traitement initial de l'information par la présence de synapses inhibitrices entre les dendrites secondaires des cellules mitrales et à panache et les dendrites des interneurons de type granulaire et périglomérulaire. La figure 1.8 montre que l'activation d'une cellule mitrale par exemple mène à une inhibition en retour de la cellule par un effet rétroactif (médiée par les cellules granulaires) ainsi que l'inhibition de cellules voisines, y compris celles de glomérules voisins (médiée par les cellules périglomérulaires). Mori pense que cette inhibition latérale a pour effet de restreindre le spectre de reconnaissance des cellules mitrales et à panache par rapport aux neurones olfactifs avec lesquelles elles font synapses.

1.2.4.2 Intégration au niveau du cortex olfactif.

L'information olfactive qui traverse le bulbe olfactif atteint un ensemble d'aires paléocorticales regroupées sous le nom de cortex olfactif (les tubercules olfactifs, le cortex prépyriforme, le noyau cortical de l'amygdale et l'aire entorhinal, figure 1.9). Par la suite, l'information

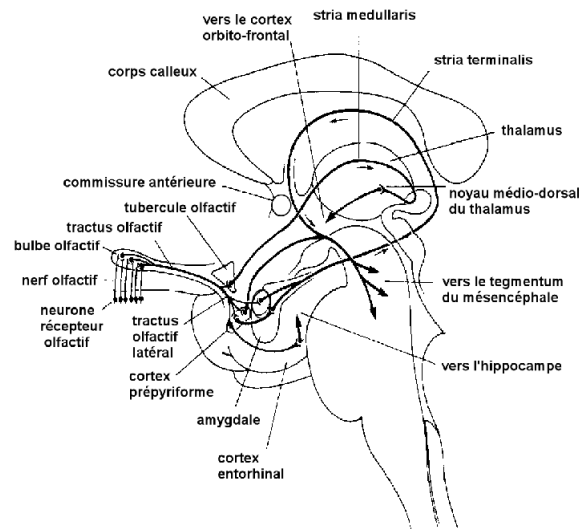


Fig. 1.9: Représentation schématique des voies olfactives chez l'Homme (Reproduction du site³).

olfactive accède au néocortex orbito-frontal qui est impliqué dans la perception consciente de l'odeur. En parallèle, les voies olfactives projettent en direction du système hypothalamo- limbique impliqué dans les processus affectifs et mnésiques (Buck [2000]). Les travaux de Zou et de ses collaborateurs (Zou et al. [2001]) ont permis de déterminer l'organisation du cortex olfactif. Nous pouvons tout de même faire état de quelques hypothèses formulées sur la base d'observations neuroanatomiques.

Les cellules mitrales et à panache se projettent directement vers le cortex olfactif primaire par le tractus olfactif latéral (Doty [2001]). L'imagerie fonctionnelle a récemment montré que la présence de molécules odorantes active chez l'Homme les cortex piriforme et orbitofrontal et le lobe frontal inférieur.

Les liens fonctionnels entre le bulbe olfactif et le cortex olfactif ne sont pas encore clairs. L'équipe de Mori a cependant mis en évidence un phénomène de décharges oscillatoires synchronisées (*synchronized oscillatory discharges*) entre les cellules mitrales et à panache, attribuées aux connexions synaptiques entre ces cellules et les cellules granulaires. Bien que ne reposant pas sur des observations directes, il émet l'hypothèse que les axones de sortie du bulbe olfactif convergent dans le cortex olfactif vers les neurones corticaux. Ainsi, deux glomérules représentant deux récepteurs différents sont reliés à un neurone cortical par les cellules mitrales et à panache, "synchronisées" ou non. L'excitation du neurone cortical cible serait fonction de la synchronisation des décharges en sortie des deux glomérules : si les signaux sont synchronisés, l'intégration temporelle en entrée du neurone cortical y induit un potentiel d'action tandis que des signaux asynchrones seraient sans effet. Dans ce modèle, le neurone cortical sert de détecteur combinatoire dont l'activité représente l'activité combinatoire des deux glomérules. Cette étape sera alors l'ultime étape de l'intégration du signal par la recombinaison de la décomposition du stimulus chimique en motifs structuraux.

1.2.4.3 Le cas de invertébrés

Le codage des odorants chez la drosophile est apparemment proche de celui utilisé par la souris : la règle de la convergence glomérulaire s'applique (pour revue Voshall [2001]). Le principe du codage spatial des odorants dans le lobe antennaire est donc une hypothèse séduisante, d'autant plus que chez l'abeille, le papillon, de même que chez le rat et la salamandre, une classe d'odorants donnée active une région bien délimitée du lobe antennaire, conservée à travers les individus, et que chez l'abeille, chaque glomérule d'un sous-ensemble est activé par un odorant de la classe (Galizia et al. [1998]). Cependant, Voshall fait état de divergence d'avis sur la façon dont le cerveau peut interpréter ces faits.

1.2.5 Les évènements périorécepteurs

Le processus discuté jusqu'à présent (des récepteurs au cortex) est susceptible de mener à une explication satisfaisante des observations. Cependant, la première étape de reconnaissance de la molécule est sans doute plus compliquée. En 1981, Vogt et Riddiford [1981] isolent dans l'antenne du papillon mâle *Antheraea polyphemus* deux familles de protéines. La première, appelée *pheromone-binding protein* (PBP), lie la composante principale de la phéromone sexuelle tandis que la seconde représente des estérases. Une sous-famille des estérases dégrade de plus la phéromone et est spécifiquement exprimée dans les éléments cuticulaires de l'insecte. Parallèlement, Pelosi isole une protéine liant le 2-isobutyl-3-méthoxypyrazine chez la vache (Pelosi et al. [1981, 1982]). La présence généralisée de telles protéines, à la fois chez les vertébrés et les invertébrés implique qu'elles ont un rôle important dans le processus de l'olfaction. Etant le sujet principal de notre travail, nous consacrerons une partie entière aux connaissances actuelles sur ces protéines.

Chapitre 2

Les OBPs

2.1 Les connaissances actuelles sur les OBPs

2.1.1 Caractéristiques des OBPs

A la suite de la découverte de Vogt en 1981 (Vogt and Riddiford [1981]), des dizaines protéines ayant la faculté de se lier réversiblement à des petites molécules volatiles ont été identifiées dans le milieu aqueux à proximité des neurones olfactifs, dans de nombreuses espèces appartenant à des classes très différentes (pour revue, Hansson [2002]) :

⇨ les insectes :

1. 17 espèces de la division Endopterygota (lépidoptères, diptères, coléoptères, hyménoptère)
2. 1 espèce de la division des hémiptéroïdes
3. 1 espèces des orthoptères (Ban et al. [2003])

⇨ les mammifères : porc-épic, souris, lapin, rat, homme, porc, boeuf

⇨ les amphibiens : le crapaud *Xenopus laevis*

Pour l'instant, la recherche de telles protéines est infructueuse pour les oiseaux, les tortues, les poissons ainsi que pour les vertébrés autre que mammifères.

Le terme de protéine de liaison aux molécules odorantes (OBP, *Odorant-Binding Protein*) est appliqué à toutes les protéines homologues exprimées dans l'épithélium olfactif. De part leur fonction, les OBPs d'insecte et de vertébrés sont classés sous le même nom mais elles appartiennent à des classes différentes. Les OBPs de vertébrés forment une sous-famille des lipocalines tandis que les OBPs d'insecte forment une classe distincte de part leur séquence primaire et leur repliement tridimensionnel. Malgré leurs caractéristiques structurales différentes et leur fonction toujours hypothétique, leurs propriétés biochimiques générales indiquent une fonction physiologique proche dans la transduction olfactive. Au fur et à mesure des avancées, l'un des pionniers de l'étude des OBPs, Pelosi, a consacré plusieurs revues aux OBPs (Pelosi and Maida [1990], Pelosi [1994], Pelosi and Maida [1995], Pelosi [1998]). Nous allons en tirer une grande partie des informations, en distinguant une nouvelle fois vertébrés et invertébrés.

2.1.2 Les OBPs de vertébrés

Les OBPs de mammifères appartiennent à la famille des lipocalines, contenant entre autres la β -lactoglobuline, la protéine de liaison au rétinol (RBP) et des chromophores comme l'insecticyanine. Cette famille est caractérisée par la liaison réversible à des ligands hydrophobes, par leur expression dans le milieu extracellulaire et par un repliement en tonneau β (pour revue, Flower [1996], Akerstrom et al. [2000]). Certaines lipocalines, y compris une OBP de boeuf (Boudjelal et al. [1996]) ont montré également qu'elles pouvaient se lier à des récepteurs de surface cellulaires (pour revue, Flower [2000a]). L'OBP de porc se fixe également avec une forte affinité (nM) et une grande sélectivité sur un récepteur olfactif humain (Matarazzo et al. [2002]). La définition d'OBP chez les vertébrés est rendue délicate par l'existence de lipocalines qui pourraient être classées parmi les OBPs par leurs propriétés d'interaction (pour revue, Pelosi [1995]), telles que l'aphrodisine (protéine abondante dans les décharges vaginales du hamster), les protéines urinaires de la souris (MUP) et du rat (α -microglobuline), les protéines salivaires de la souris et du porc, liant des molécules odorantes ou des composés endogènes, associés aux phéromones. La protéine MUP, par exemple, sécrétée dans les urines de souris mâles, fixe une phéromone et est capable d'agir elle-même comme phéromone en modifiant le cycle reproductif de ces rongeurs. Ces protéines se sont pas toutes exprimées spécifiquement dans les tissus olfactifs ce qui donne une définition plus ou moins artificielle des OBPs : une OBP est une protéine liant des molécules odorantes et exprimée spécifiquement dans l'épithélium olfactif. On les appelle alors aussi parfois *olfactory binding protein*. Cependant, on classe parfois les lipocalines liant odorants et/ou phéromones dans une sous-classe des OBPs, appelée PBP (*Pheromone-Binding Protein*), par homologie avec les PBPs d'insectes (Tegoni et al. [2000]).

2.1.2.1 Séquence

La plupart des OBPs de vertébrés totalement séquencées sont des OBPs de mammifères et contiennent entre 150 et 170 acides aminés (Tegoni et al. [2000]). L'identité de séquence entre OBPs est faible, la moyenne étant entre 20 et 30%. La plus grande identité rassemble les OBPs de porc et de boeuf (42%). Comme pour les autres lipocalines, la séquence -G-X-W- située à une distance de 15 à 20 résidus de l'acide aminé N-terminal est conservée (région conservée SCR1). On note également la conservation quasi systématique du motif YxxxYxG au milieu de la séquence et de quelques résidus disséminés. Le nombre de cystéines est lui-même très variable. On retrouve parfois plus d'identité de séquence entre deux OBPs d'espèces différentes qu'entre deux OBPs d'une même espèce, suggérant l'existence de sous-famille d'OBPs (trois au minimum) au sein des mammifères. D'autre part, une des sous-famille présente une plus grande similarité de séquence avec les protéines urinaires et salivaires.

2.1.2.2 Structure

Les OBPs de vertébrés adoptent un repliement tridimensionnel similaire à celui de toutes les lipocalines (pour revues Flower [2000b], Flower et al. [2000]). Le coeur de la structure des lipocalines est un tonneau β formé de huit brins β par une connectivité +1 (figure 2.1). Les boucles L1, L3, L5 et L7 d'un côté et les boucles L2, L4 et L6 sont orientées vers les deux ouvertures du tonneau. Les extrémités N- et C-terminales forment respectivement une

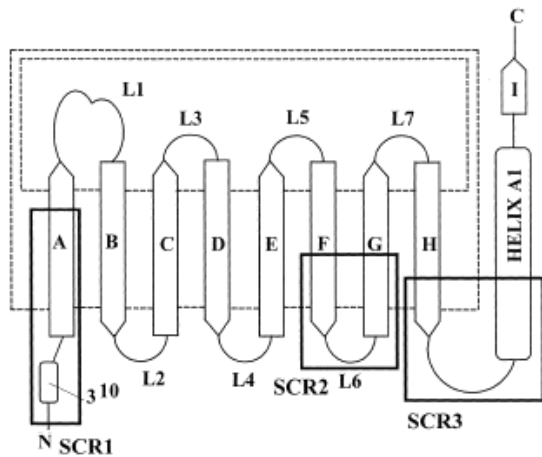


Fig. 2.1: Topologie de repliement des lipocalines. Les neuf brins β du feuillet β antiparallèle sont notés de A à I. Le tonneau β est formé par les 8 brins A à H. Les régions conservées sont marquées : SCR1 à SCR3. (reproduction de Flower et al. [2000])

hélice 3_{10} et une hélice α . L'orientation de l'hélice α est très variable suivant les protéines. Le repliement des lipocalines est caractérisé par trois régions conservées tant au sens de la séquence que de la structure tridimensionnelle : SCR1, SCR2 et SCR3. Le site de fixation des ligands de lipocalines est délimité par la cavité du tonneau β et la boucle externe L1. La diversité des résidus de la poche explique la reconnaissance des ligands par leur forme, leur taille, leurs caractéristiques chimiques.

Les OBPs diffèrent des protéines typiques des lipocalines (*kernel lipocalins*) par la non conservation des régions SCR2 et SCR3. Deux OBPs de vertébrés ont été cristallisées et leurs structures sont disponibles. Elles montrent des différences notables. Si l'OBP de porc (figure 2.2, Spinelli et al. [1998]) est monomérique, l'OBP bovine (figure 2.3, Bianchet et al. [1996], Tegoni et al. [1996]) est dimérique. Plus étonnant est le phénomène de “*domain swapping*” qui apparaît par l'échange au sein de l'OBP bovine des hélices α entre les deux monomères. Spinelli attribue ce phénomène à l'absence dans l'OBP bovine d'un pont disulfure entre les domaines β et α alors que la plupart des autres OBPs possèdent cette liaison.

Le ligand de l'OBP bovine, cristallisé avec la protéine et identifié par Ramoni (Ramoni et al. [2001]) au 1-octèn-3-ol racémique, se lie à l'OBP dans un rapport de une molécule par monomère. Les complexes d'OBP de porc obtenus par Vincent (Vincent et al. [2000]) sur une série de 7 molécules montrent que l'odorant se fixe à l'intérieur de la cavité de chaque monomère. Les interactions de type hydrophobes et apolaires sont majoritaires mais l'orientation de molécules structurellement proches (ex. aromatiques) est très différente dans chaque complexe.

2.1.2.3 OBPs et odorants

D'une manière générale, le spectre de fixation des OBPs de porc et de boeuf (les plus étudiées) est large (Pelosi [1998]) : une OBP fixe plusieurs molécules, de préférence hydrophobes et de taille moyenne (9-12 atomes de carbone) avec une forme plate ou susceptible d'adopter une conformation plate. Les composés aromatiques, tels que l'IBMP (2-isobutyl-3-méthoxy-pyrazine) ou le thymol sont représentatifs des bons ligands, au même titre que des molécules plates mais non totalement planaires telles que la menthone et la carvone et que des molécules flexibles comme le 3,7-diméthyl-octanol. Les constantes de dissociation sont dans la gamme du micromolaire, soit une interaction relativement faible.

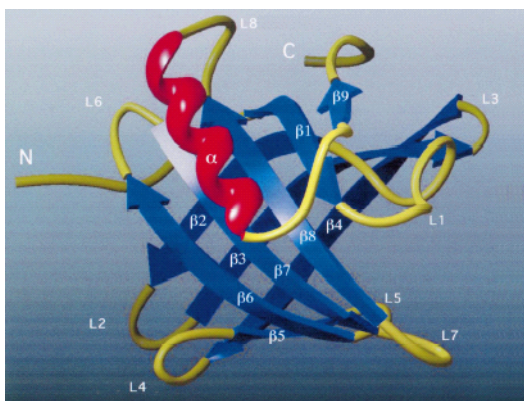


Fig. 2.2: Représentation de l'OBP de porc. Les brins β sont numérotés d'après leur ordre dans la séquence ($\beta 1=A, \beta 2=B, \dots$) (adapté de Tegoni et al. [2000])

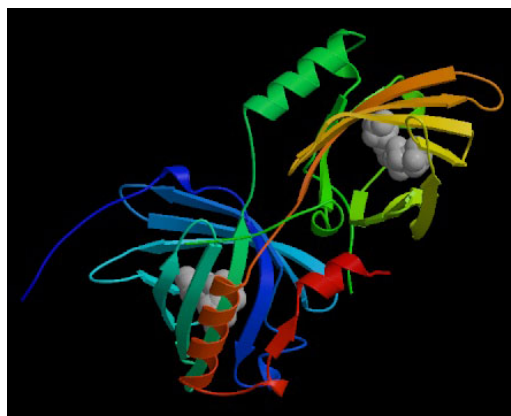


Fig. 2.3: Structure cristallographique de l'OBP bovine et son ligand (1-octen-3-ol).

2.1.3 Les OBPs d'insectes

Chez les insectes, les OBPs sont des protéines très solubles que l'on trouve en grande concentration dans la lymphe sensillaire (de l'ordre du millimolaire), équivalent du mucus des vertébrés. Les OBPs sont sécrétées par les cellules trichogènes et tormogènes (figure 1.5) avec un peptide signal de sécrétion (Steinbrecht et al. [1992, 1995]). D'autre part, les cellules théogènes (qui protègent les neurones olfactifs (cellules gliales)) sont probablement responsables de la dégradation des OBPs (par endocytose). En effet, la vitesse de renouvellement des OBPs dans la sensille est importante : $8 \cdot 10^7$ molécules/h/sensille (Vogt et al. [1989]).

Certaines OBPs ont une expression différentielle en fonction du sexe. Ainsi, ASP2, une OBP d'abeille, s'exprime en plus grande quantité chez l'ouvrière que chez le faux-bourdon (Danty et al. [1997]). Connaissant les rôles respectifs dévolus aux ouvrières et aux mâles, cette OBP est susceptible d'avoir un rôle général dans l'olfaction de l'abeille ouvrière. D'autre part, l'OBP découverte par Vogt chez le mâle du papillon *Antheraea polyphemus* (Vogt and Riddiford [1981]) lie la phéromone sexuelle émise par la femelle. Les OBPs présentent donc un éventail large de propriétés. Celles-ci définissent, chez les invertébrés, deux classes principales (Vogt et al. [1991]) :

- ⇨ La classe des OBPs liant spécifiquement les phéromones et exprimées dans des sensilles spécialisées de l'adulte et de préférence chez le mâle est appelée PBP (*Pheromone-Binding Protein*). Cependant, récemment, deux études vont à contre sens et rendent délicate la définition de PBP : Callahan et al. (Callahan et al. [2000]) rapporte la présence de PBP en quantité comparable chez le mâle et la femelle de trois lépidoptères ; plusieurs PBPs de papillon (Campanacci et al. [2001]) sont capables de fixer un spectre assez large d'odorants.
- ⇨ Par opposition, les OBPs liant un large spectre de molécules et non spécifiques du sexe et de l'âge sont appelées *general-OBP* (gOBP). Les gOBPs se répartissent en deux classes d'après leur séquence primaire : GOPB1 et GOPB2.

Plus récemment, une autre classe de protéines, apparentées aux OBPs a été découverte dans plusieurs parties du corps d'insectes. Ces protéines, appelées *Chemosensory Proteins* (CSP)

sont probablement impliquées dans la sensation chimique générale de contact, l'équivalent du goût des vertébrés.

Nous allons faire le point bibliographique sur les connaissances accumulées jusqu'à présent sur les deux classes d'OBPs et les CSP.

2.1.3.1 OBPs et séquence primaire : vers une classification

La séquence des OBPs (PBP et gOBP) est caractérisée en tout premier lieu par la présence et la topologie de trois ponts disulfure : CYS(I)-CYS(III), CYS(II)-CYS(V), CYS(IV)-CYS(VI).

En 1999, Vogt (Vogt et al. [1999]) établit une classification des OBPs d'insecte à partir d'une base de données de 43 protéines apparentées à des OBPs appartenant à 17 espèces de lépidoptères, 3 espèces de coléoptères, 2 espèces de diptère, une espèce d'hyménoptère et une espèce d'hémiptère (figure 2.4). L'origine des OBPs est détaillée dans la publication. Il résulte de cette étude que les OBPs s'organisent en plusieurs classes, certaines monophylétiques par rapport à l'ordre ou à l'espèce, certaines polyphétiques. Ainsi, la classe des OBPs de lépidoptères est isolée, aucune OBP d'espèce d'un autre ordre n'y est incluse. Des sous-classes correspondent aux PBPs et aux gOBPs (GOPB1 et GOPB2, déjà identifiés (Krieger et al. [1993, 1996], Vogt et al. [1991])). La famille des *Antennal Binding Protein X* (ABPX, Krieger et al. [1996]) représente une troisième sous-classe des OBPs. A contrario, une classe polyphétique contient des OBPs d'origines très diverses : hémiptère, coléoptère et lépidoptères (moitié inférieure de l'arbre). D'autre part, si certaines classes représentent un sous-ensemble d'OBPs appartenant à une même espèce, les OBPs d'une espèce peuvent être répartis sur l'ensemble de l'arbre, suggérant une fonction différente pour chaque sous-ensemble (ex : les protéines OS-E, OS-F, LUSH et PBPRPs de la drosophile).

Il faut noter que la littérature ne converge pas encore complètement vers la définition précédente des OBPs. En effet, l'auteur de la caractérisation de la famille de THP12 (Graham et al. [2003]) classe ces protéines dans les OBPs alors qu'ils ne contiennent que 2 ponts disulfure, qu'elles sont plus petites (12kDa) et qu'elles sont présentes dans l'hémolymphe du scarabée *Tenebrio molitor*. Elle utilise une classification des OBPs en fonction du nombre de ponts disulfure (isoformes 4-CYS et 6-CYS), tout en remarquant que les isoformes 6-CYS sont plus particulièrement exprimés dans les antennes.

De manière générale, nous pouvons garder à l'esprit que l'homologie de séquence entre des OBPs issus d'une même espèce est souvent faible alors qu'on retrouve des OBPs de forte homologie de séquence entre deux espèces différentes (Pelosi [1998]).

Malgré la grande diversité de séquence primaire des OBPs, elles partagent plusieurs caractéristiques biochimiques :

- ⇔ poids moléculaire compris entre 13 et 16kDa,
- ⇔ point isoélectrique acide (entre 4 et 6),
- ⇔ possibilité d'oligomérisation,
- ⇔ présence de trois ponts disulfure de topologie conservée,
- ⇔ profil d'hydrophobicité conservé,
- ⇔ structure en hélice α .

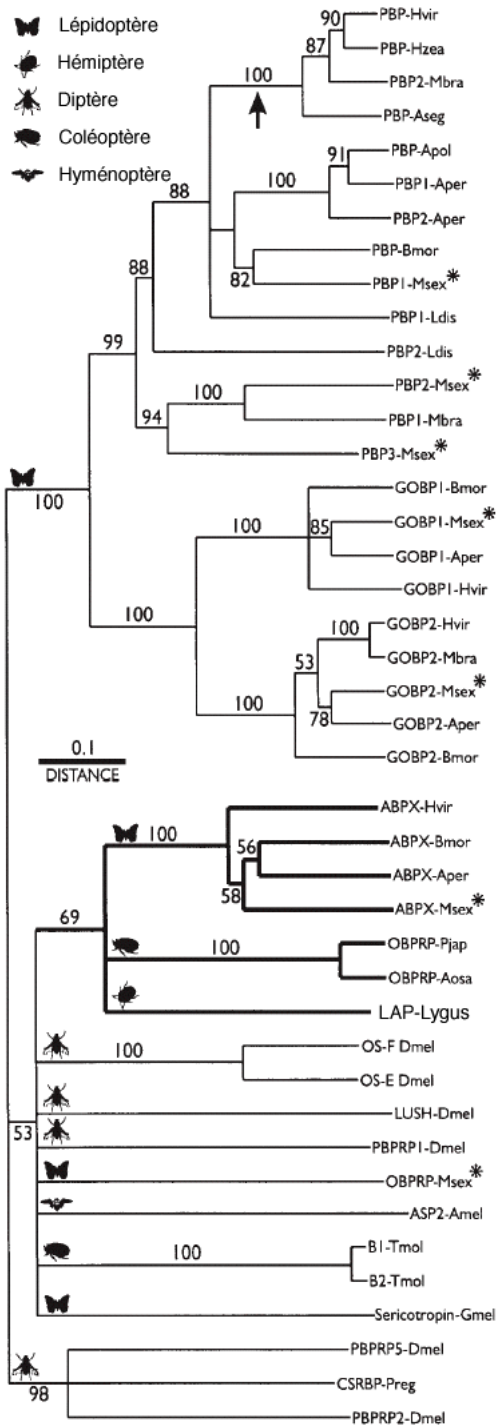


Fig. 2.4: Arbre de voisinage des OBPs d'insectes. La longueur des branches est proportionnelle et l'échelle de distance est indiquée. (reproduction de Vogt et al. [1999]). La première partie du nom correspond au type d'OBP (PBP ou GOBP) tandis que la seconde partie du nom désigne l'espèce d'après les correspondances du tableau suivant :

Ordre	Espèce	Abréviations
Lépidoptères (papillon)	<i>Antheraea polyphemus</i>	Apol
	<i>Antheraea pernyi</i>	Aper
	<i>Bombyx Mori</i>	Bmor
	<i>Manduca sexta</i>	Msex
	<i>Heliothis virescens</i>	Hvir
	<i>Helicoverpa zea</i>	Hzea
	<i>Mamestra brassicae</i>	Mbra
	<i>Lymantria dispar</i>	Ldis
	<i>Agrotis segetum</i>	Aseg
Coléoptères (scarabée)	<i>Popillia japonica</i>	Pjap
	<i>Anomala osakana</i>	Aosa
	<i>Tenebrio molitor</i>	Tmol
Diptère (mouches)	<i>Drosophila melanogaster</i>	Dmel
	<i>Phormia regina</i>	Preg
Hyménoptère	<i>Apis mellifera</i>	Amel
Hémiptère	<i>Lygus lineolaris</i>	Lygus

2.1.3.2 OBPs et odorants

La seule étude à ma connaissance de fixation d'odorants sur une GOPB concerne ASP2. Elle sera discutée ultérieurement. Par contre, plusieurs travaux sur les PBP montrent que les constantes d'association avec leurs phéromones respectives sont de l'ordre du micromolaire (Plettner et al. [2000], Du and Prestwich [1995], Bette et al. [2002], Rivière et al. [2003]). L'interaction au niveau moléculaire entre la PBP de *Bombyx mori* et le bombykol sera décrite au paragraphe 2.1.4.1.

2.1.3.3 Les CSP

Les CSPs (ou Chemosensory Proteins) sont des protéines apparentées aux OBPs de par leur localisation dans les tissus olfactifs, leur taille et leur fixation à de petites molécules. Ces protéines ont été identifiées chez la plupart des insectes : des papillons (Maleszka and Strange [1997], Bohbot et al. [1998], Picimbon et al. [2000], Meillour et al. [2000], Jacquin-Joly et al. [2001], Picimbon et al. [2001]), aux phasmes (Mameli et al. [1996], Marchese et al. [2000], Tuccini et al. [1996]) en passant par l'abeille (Danty et al. [1998a]). Elles diffèrent cependant des OBPs par une expression sur tout le corps (antennes, tarsi, labrum et proboscis). D'autre part, leur séquence est plus courte (100-110 acides aminés) et ne contient que 4 cystéines conservées arrangées en CYS(I)-CYS(II), CYS(III)-CYS(IV) (Angeli et al. [1999]). Les CSPs de plusieurs espèces d'insectes sont hautement conservées (Briand et al. [2002]) : de 40 à 60% d'identité de séquence. Il leur est attribué pour l'instant une fonction olfactive ou gustative (Nagnan-Le Meillour et al. [1996]) par leur existence dans les organes chimiosensibles (tels que les pattes) et leur faculté à lier des longues molécules très hydrophobes et des acides gras (Picone et al. [2001], Lartigue et al. [2002], Briand et al. [2002]), malgré l'absence de ligand naturel identifié pour la plupart des CSP. Cependant, une CSP de l'abeille (ASP3c) lie la phéromone du couvain (Briand et al. [2002]).

2.1.4 Aspects structuraux des OBPs et apparentés d'invertébrés.

A l'heure actuelle, la structure de trois protéines d'invertébrés de la famille des OBPs ou apparentées ont été élucidées.

2.1.4.1 La PBP de *Bombyx mori* : la première structure d'OBP d'insectes

Bombyx mori est le papillon aussi appelé bombyx du mûrier, ou papillon du ver à soie. Il s'agit du modèle d'invertébré le plus observé dans le cadre de la recherche sur l'olfaction. Ce papillon communique avec les autres membres de son espèce par le biais de phéromones. La première phéromone (bombykol) fut isolée et caractérisée chez le papillon du ver à soie par Butenandt en 1959. Comme mesuré par les études de dichroïsme circulaire (Wojtasek and Leal [1999]) et de RMN (Damberger et al. [2000]), BmPBP adopte un repliement différent en fonction du pH. Les deux conformations sont en équilibre lent pour un pH compris entre 4,9 et 6,0.

La structure cristallographique de la PBP de *Bombyx mori* à pH basique (pH>6,5) a été déterminée par Sandler (figure 2.5, Sandler et al. [2000]) en 2000, suivi par la résolution par RMN par Lee (Lee et al. [2002]). La BmPBP complexée avec le bombykol cristallise sous la

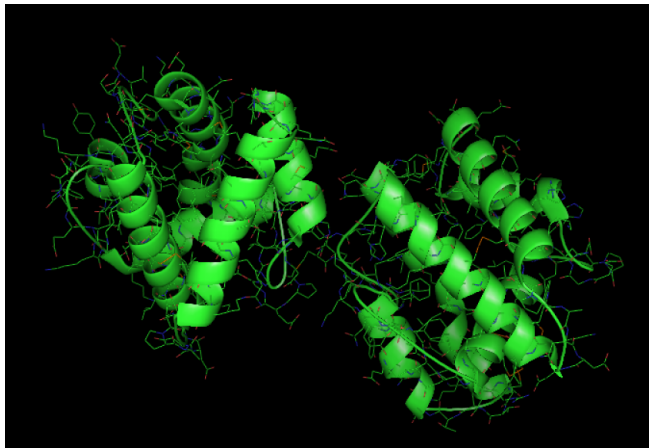


Fig. 2.5:
Structure cristallographique
de BmPBP en complexe
avec le bombykol.

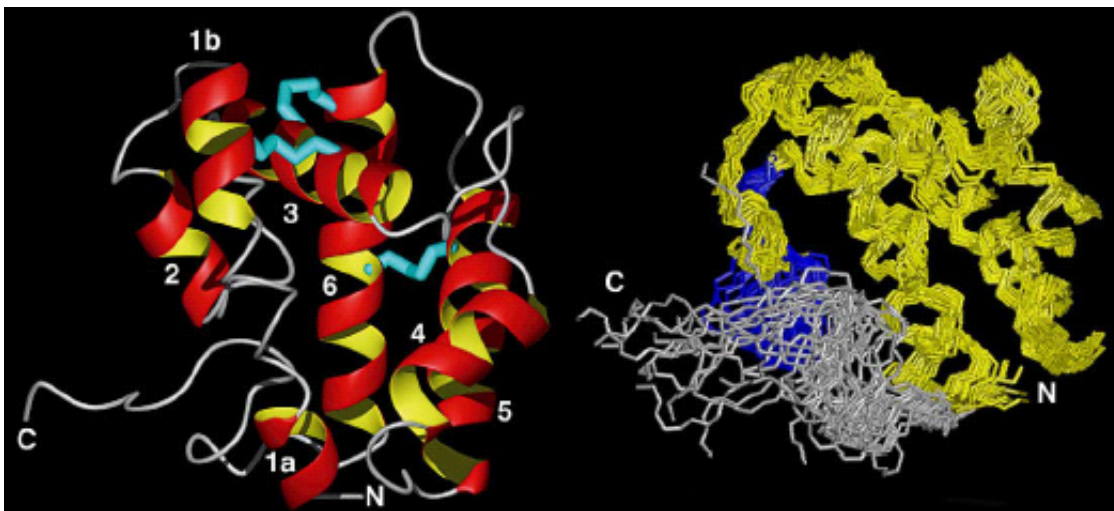


Fig. 2.6: Structure RMN de BmPBP. A droite, superposition des 20 structures de plus basse énergie. A gauche, représentation des structures secondaires d'une des structures. Les ponts disulfures sont montrés en couleur cyan (reproduction de Lee et al. [2002]).

forme d'un homodimère de 2x15,9kDa. La structure de chaque monomère est identique et est composée de six hélices, notées α_1 à α_6 . La première hélice est coudée (α_{1a} - α_{1b}). L'extrémité C-terminale (129-137) est en conformation étendue et flanquée entre l'hélice α_{1a} et l'hélice α_2 à la surface de la protéine. Le bombykol est liée à la BmPBP par chaque monomère, impliquant des résidus globalement hydrophobes de la poche délimitée par les quatre hélices antiparallèles α_1 , α_4 , α_5 et α_6 . Les phénylalanines 12 et 118 créent une surface hydrophobe, tandis que les chaînes latérales de M61, L62, I91 et V114, résidus plus variables, sont probablement à l'origine de la spécificité du ligand.

La structure obtenue en solution par RMN à pH basique (BmPBPB, figures 2.6) en absence de bombykol est très proche de la structure cristalline. L'hélice α_1 située en position N-terminale est séparée en deux α_{1a} (3-8) et α_{1b} (16-22). La superposition des 20 structures de plus basse énergie montre que l'extrémité C-terminale (129-142) est peu rigide. De même, la région comprise entre les résidus 33 et 45 montre une plus grande flexibilité que le reste de la

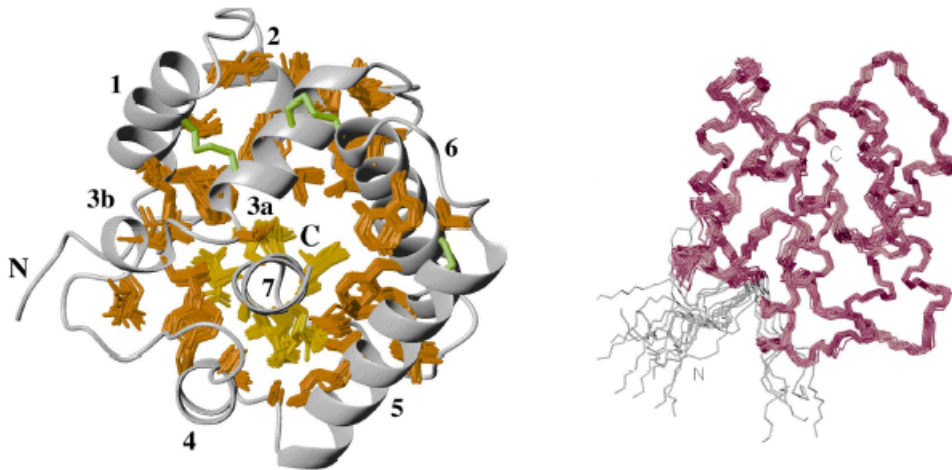


Fig. 2.7: Structure RMN de BmPBPA. A droite, superposition des 20 structures de plus basse énergie. A gauche, représentation des structures secondaires d'une des structures. Les chaînes latérales de plus faible exposition au solvant, qui forment le coeur hydrophobe, sont représentés : en jaune, les chaînes latérales de l'hélice α_7 , en orange, les autres. Les ponts disulfures sont représentés en vert (reproduction de Horst et al. [2001]).

protéine. La modélisation moléculaire reflète le peu de contraintes à moyenne distance (structuration secondaire) et l'absence de nOe à longue distance (repliement tertiaire) concernant les hydrogènes de ces deux régions. Ces observations de la modélisation furent confirmées par la mesure de l'effet Overhauser entre les noyaux ^{15}N et 1H , expérience qui met en évidence la flexibilité du squelette de la chaîne polypeptidique. L'absence de ligand n'induit donc pas de modification majeure de la structure.

A bas pH (4,5), la structure, BmPBPA, est considérablement modifiée. Les six hélices de la structure à haut pH ont une orientation identique, par contre, le segment C-terminal, étendu à haut pH, se structure en hélice α localisée dans le coeur de la structure (figures 2.7). L'hélice α_7 occupe alors l'espace qui correspond au site de fixation de la phéromone dans la structure cristalline du complexe. Le mécanisme probable de relarguage du ligand est donc identifié :

1. le ligand se fixe sur BmPBP au pH de la lymphe sensillaire,
2. Au contact des membranes, le pH diminue (van der Got et al. [1991]) et BmPBP change de conformation
3. le ligand est chassé par la formation de l'hélice α au coeur de la structure.

Ce modèle ne peut être généralisé car la fixation de la phéromone sur la PBP de *Lymantria dispar* est conservé à pH 4 (Kowcun et al. [2001]). De plus, ce mécanisme ne peut exister dans le cas d'OBP plus courtes, telles que la PBP de *Leucophaea maderae* (Rivière et al. [2003]).

L'état oligomérique de la PBP de *Bombyx mori* est à mettre en lumière des études de gel filtration (Leal [2000]) : à bas pH (pH<4,5), BmPBP est sous forme monomérique et à haut pH (pH>6), BmPBP est dimérique. BmPBP est un dimère à haut pH dans la forme cristalline alors que la structure en solution semble être monomérique. D'autre part, la structure à bas pH semble être monomérique.

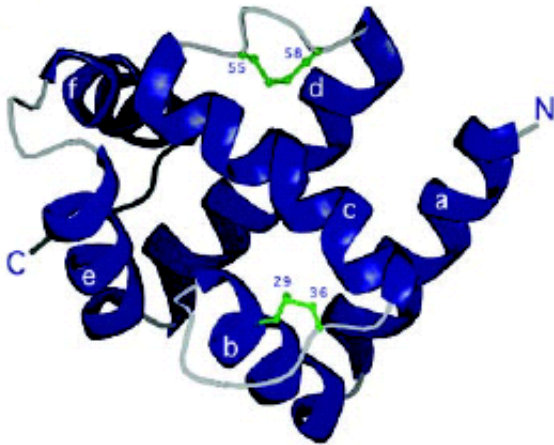


Fig. 2.8: Structure RMN de CSPMbra6

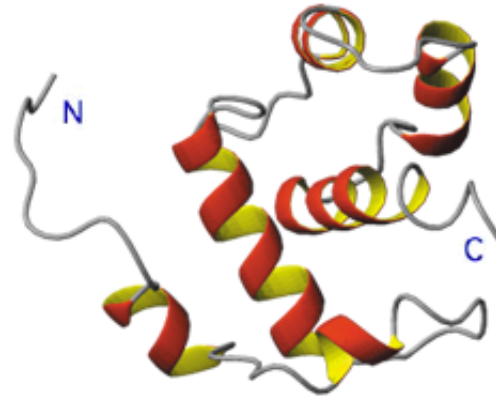


Fig. 2.9: Structure RMN de THP12

2.1.4.2 CSPMBra6 et THP12

Les deux autres structures de protéines apparentées aux OBPs sont CSPMbra6 et THP12. CSPMbra6 (Lartigue et al. [2002], Mosbah et al. [2002]) est une CSP produite dans les antennes et le proboscis du papillon *Mamestra brassicae* tandis que THP12 fait partie d'une famille de protéine de transport de l'hémolymphe du scarabée *Tenebrio molitor*. Elles sont plus courtes que la PBP *Bombyx mori* (112aa et 108aa respectivement) et ne partagent qu'une très faible homologie de séquence entre elles et avec l'ensemble des OBPs, y compris BmPBP. Par contre, leurs structures sont similaires et sont proches de celle de BmPBP à haut pH. Elles font toutes partie de la famille des structures hexahélicales.

La structure de THP12 (Rothmund et al. [1999]) est celle qui se rapproche le plus de BmPBP. Les cinq hélices $\alpha_2 - \alpha_6$ ont quasiment la même orientation relative dans les deux structures. L'hélice N-terminale α_1 est mal définie. Contrairement à BmPBP, le site de fixation des ligands (acides gras et phéromone 4-méthylnonanol) est un sillon hydrophobe .

2.1.5 Des modèles fonctionnels des OBPs

Malgré l'accumulation des observations expérimentales depuis ces dernières années, plusieurs modèles et leurs variantes restent en concurrence. Tous font bien entendu intervenir les OBPs, les récepteurs ainsi que les odorants. De manière générale, les enzymes de dégradation (ODE) participent à la terminaison de la molécule odorante par sa dégradation (Kaissling [1998]).

- ⇨ Les OBPs agissent comme un filtre avant l'activation des récepteurs (Vogt and Riddiford [1981]). Elles solubilisent (sélectivement ?) les molécules hydrophobes pour faciliter leur accès aux récepteurs ;
- ⇨ Les OBPs forment un complexe avec les molécules odorantes et le complexe interagit avec le récepteur (Prestwich et al. [1995]) ;

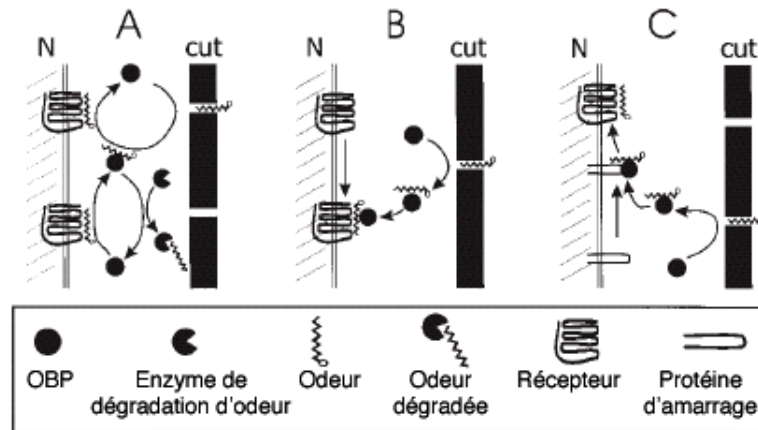


Fig. 2.10: Modèles d'action des OBPs d'insecte actuellement envisagés.

Dans les trois schémas, l'odorant entre dans la sensille olfactive à travers les pores de la cuticule (Cut). Les OBPs facilitent le transport de ces molécules hydrophobes à travers le milieu hydrophile de la lymphe sensillaire qui entoure les neurones (N). Pour la description des modèles, voir le texte. (adapté de Vogt et al. [1999])

- ⇨ Les OBPs inactivent les molécules pour éviter la désensibilisation des récepteurs ou pour permettre la détection d'un nouveau stimulus.

Différents mécanismes d'inactivation des stimulus ont été proposés :

- ⇨ inactivation directe de la molécule par liaison avec les OBPs (Vogt and Riddiford [1981]),
- ⇨ le complexe OBP-odorant est le substrat des enzymes de dégradation des odorants (ex. estérase) (Vogt et al. [1985]);
- ⇨ les OBPs éliminent les excès d'odorants qui entourent les neurones olfactifs (Kaissling [1998]).

Il est possible également qu'une protéine d'amarrage (*Sensory Neuron Membrane Protein*, ou SNMP, Rogers et al. [2001a,b]), de la famille des récepteurs CD-36 à deux domaines-transmembranaires et très abondante dans les membranes des neurones olfactifs, ait un rôle à jouer, en présélectionnant par exemple les complexes PBP-phéromone qui activeront ensuite le récepteur olfactif.

Ces modèles s'appuient sur diverses observations : les OBPs se lient aux odorants (par définition) et augmentent leur solubilité dans l'espace périneuronale, les OBPs se lient réversiblement à des récepteurs olfactifs (Boudjelal et al. [1996], Pophof [2002]), des enzymes de dégradation plus ou moins spécifiques sont présentes dans la lymphe sensillaire et le mucus (pour revue, Pelosi and Maida [1990]).

Le rôle des OBPs est prouvé d'un point de vue comportemental : l'absence de l'OBP LUSH de la drosophile se traduit par une modification du comportement vis-à-vis des alcools (Kim et al. [1998], Kim and Smith [2001]), le nombre de reines dans les colonies de fourmis de feu est régulé par un gène codant une PBP (Krieger and Ross [2002]).

Cependant, d'autres possibilités de fonction des OBPs ont été émises. Du fait de la non-spécificité de liaison de l'OBP bovine aux tissus olfactifs, les OBPs pourraient avoir un rôle

plus général dans la détoxification ou dans la signalisation cellulaire (Boudjelal et al. [1996]).

2.2 ASP2 : une OBP de l'abeille domestique

2.2.1 La mise en évidence des OBPs chez l'abeille domestique *Apis mellifera*.

Durant sa thèse au Laboratoire de Neurobiologie Comparée des Invertébrés sous la direction de Claudine Masson (Danty [1996]), Emmanuelle Danty a étudié entre autres les protéines solubles répondant aux caractéristiques attendues pour les OBPs (masse moléculaire entre 12 et 20kDa, hydrophile, point isoélectrique acide, localisation dans les antennes et les pattes).

Les travaux ont été menés sur des abeilles ouvrières et mâles élevées en conditions naturelles (i.e. dans les ruches) et hybrides des espèces *Apis mellifera ligustica*, *Apis mellifera mellifera* ainsi que sur les races pures. Les tissus (antennes, pattes, cerveau, thorax et hémolymphe) ont été broyés puis les protéines solubles et insolubles furent séparées par ultracentrifugation. L'analyse des protéines fut effectuée par électrophorèse en conditions non dénaturantes et la purification fut menée par chromatographie en phase inverse (RPLC).

Les protéines purifiées furent classées en fonction de leur tissu d'origine : les ASP (spécifiques des antennes) et les ALP (spécifiques des pattes ou des antennes et des pattes). La sous-classe ASP1 contient deux isoformes de même masse et de même séquence N-terminale mais de mobilité différente en gel d'électrophorèse. La sous-classe ASP2 contient une seule protéine. La sous-classe ASP3 contient deux isoformes d'une même protéine (ASP3a et ASP3b). Bien qu'ayant la même masse moléculaire et la même séquence N-terminale, les deux isoformes sont représentées de façon différente en fonction du sexe (tableau 2.1). Les ALPs sont des protéines solubles de masses comparables aux OBPs et contiennent une séquence consensus *GXXCXXE* trouvée fréquemment dans les OBPs. Cependant, leur localisation non spécifique dans les antennes supposent qu'elle ne sont pas des OBPs mais qu'elles partagent un certain nombre de propriétés.

La recherche d'homologie de séquence sur ASP1 a proposé l'insecticyanine, une protéine de l'hémolymphe de *Manduca sexta* responsable de la couleur de la chenille. Mais leurs distributions respectives dans l'organisme impliquent qu'ASP1 ne peut être une insecticyanine. L'abeille mâle est plus riche en ASP1 que l'ouvrière. De plus, ASP1 est exprimée spécifiquement dans les sensilles placodées, elle présente une identité de séquence significative avec les PBPs d'autres espèces et elle fixe l'acide 9-céto-2(E)-décénoïque et l'acide 9-hydroxy-2(E)-décénoïque, composants les plus actifs de la phéromone royale : cela suggère qu'ASP1 est une PBP (Danty et al. [1999, 1997]).

Le cas d'ASP2 sera discutée plus tard.

Briand et al. [2002] ont étudié les caractéristiques structurales et fonctionnelles d'ASP3c. ASP3c partage de 45% à 55% d'identité de séquence avec les CSPs d'autres espèces et comme elles, ne contient que deux ponts disulfure (CYS_I-CYS_{II} et CYS_{III}-CYS_{IV}). ASP3c appartient donc à la famille des CSP. L'étude par dichroïsme circulaire suggère que le repliement de ASP3c est analogue à celui de la CSP de *Mamestra brassicae* (6 hélices α et une boucle α - α).

Classe	$M_r(Da)$	%ASP ouvrières	%ASP mâles
ASP1a	13182	5	30
ASP1b	13182	12	30
ASP2	13695	45	4
ASP3a	11537	38	3
ASP3b	11537	0	33
ASP3c	12757	traces	traces
ALP	13000-13900		

Tab. 2.1: Protéines solubles dans les antennes et les pattes d'abeille. ASP : protéines spécifiques des antennes, ALP : protéines spécifiques des antennes et des pattes ou des pattes uniquement, %ASP : pourcentage d'ASP par rapport à la concentration totale d'ASP en fonction du sexe. (Danty et al. [1998b])

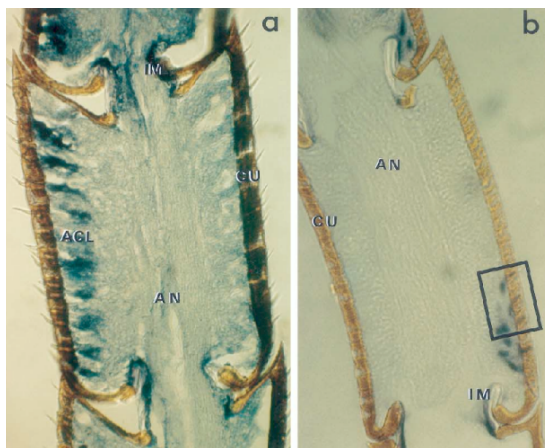


Fig. 2.11: Lieu d'expression de l'ARN messager d'ASP2 dans l'antenne d'une abeille adulte. a : chez l'ouvrière (X400), b : chez le mâle (X300). Abréviations : ac : cellule accessoire; acl : couche de cellules accessoires; an :nerf antennaire; cu : cuticule; in : membrane intersegmentaire; st : sensille trichodée. (reproduction de Danty et al. [1997])

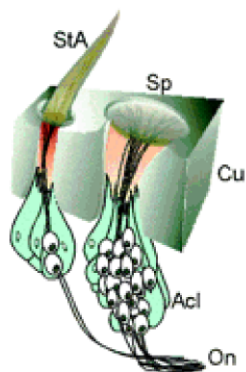


Fig. 2.12: Organisation cellulaire au niveau d'une sensille. Acl : couche de cellules auxiliaires; Cu : cuticule; On : nerf olfactif; Sp : sensille placodea; StA : sensilla trichodea A. (Reproduction de Danty et al. [1999])

2.2.2 Les propriétés d'ASP2

2.2.2.1 ASP2 : une OBP de l'abeille domestique.

L'hybridation in-situ d'une sonde à ARN (radiomarqué et complémentaire de l'ADNc codant la séquence d'ASP2) a révélé que l'ARN messenger cible est localisé essentiellement dans les régions olfactives des antennes : articles 3 à 10 de l'antenne pour les ouvrières, articles 3 à 11 pour les mâles. Les cellules marquées sont plus particulièrement les cellules auxiliaires chez les mâles. Chez l'ouvrière, l'ARNm est plus généralement produit dans la couche de cellules auxiliaires et épithéliales situé dans les invaginations de la cuticule au contact des sensilles sensorielles (voir figure 2.11). Cette protéine est exprimée exclusivement dans les antennes. Cette spécificité lui donna son nom : ASP2 (Antennal Specific Protein).

La masse moléculaire d'ASP2 mesurée en spectrométrie de masse est de $13695,2 \pm 1,6$ Da. Son point isoélectrique théorique est acide (4,8). Le profil d'hydrophobie montre qu'ASP2 est une protéine très hydrophile, comme le suggère sa forte solubilité dans l'eau.

La comparaison de la séquence d'ASP2 avec la séquence du polypeptide déduite du cadre de lecture du cDNA suggère qu'ASP2 subit un clivage de la partie N-terminale, correspondant au peptide signal de sécrétion au travers de la membrane des cellules auxiliaires sensillaires. La séquence d'ASP2 contient 6 cystéines qui sont appariées en ponts disulfure (CYS_I-CYS_{III}, CYS_{II}-CYS_V et CYS_{IV}-CYS_{VI}) lors des modifications post-traductionnelles (Briand et al. [2001]).

La séquence d'ASP2 ne partage qu'au maximum 8 à 15% d'identité avec les autres OBPs mais si les PBPs de lépidoptères présentent une forte identité de séquence (90%), les gOBPs n'ont que 40% d'identité entre elles. En revanche, les 6 cystéines restent parfaitement conservées et alignées (voir paragraphe 1.2.1).

L'ensemble des propriétés d'ASP2 (point isoélectrique acide, forte hydrophilie, faible masse moléculaire, alignement de séquence, absence de fixation des composés phéromonaux (Danty et al. [1999])) et sa localisation spécifique dans les antennes la range naturellement dans la famille des OBPs. Par ailleurs, le taux d'expression supérieur (5 :1) de la protéine chez les ouvrières par rapport aux mâles suggère qu'elle est moins impliquée dans la reconnaissance sélective des phéromones (PBP) que dans la détection des molécules odorantes (gOPB).

2.2.2.2 Les caractéristiques biochimiques d'ASP2

Une étude approfondie (Briand et al. [2001]) a été menée par l'équipe de J.-C. Pernollet sur les propriétés biochimiques d'ASP2 : arrangement des ponts disulfure, état oligomérique et structure secondaire. Les principaux résultats sont rappelés ici.

Arrangement des ponts disulfure Les expériences de digestion à la trypsine suivie d'un séquençage N-terminal et de spectrométrie de masse ont montré que les six cystéines d'ASP2 formaient trois ponts disulfure reliant Cys₂₁-Cys₅₃, Cys₄₉-Cys₁₀₇ et Cys₉₆-Cys₁₁₆.

Oligomérisation Les expériences de gel filtration suggèrent qu'à pH 7 et à 37 μ M, ASP2 adopte un état dimérique tandis qu'à pH 3, la protéine est en équilibre lent (par rapport à la durée d'éluion de la colonne) entre une forme monomérique et une forme dimérique.

Structure secondaire Le dichroïsme circulaire en UV lointain indique qu'ASP2 est composé essentiellement d'hélice α (50%) et très minoritairement de feuillet β (5%), à pH 7 et à pH 3. Par contre, l'UV proche indique qu'à pH 3, une transition apparaît, traduisant un changement de conformation locale des groupes aromatiques (augmentation de la flexibilité des chaînes latérales).

2.2.2.3 Les propriétés de fixation d'ASP2 avec des odorants

Par la technique VOBA, L. Briand et ses collègues (Briand et al. [2001]) ont testé la faculté d'ASP2 à lier plusieurs types de molécules choisies pour leurs structures différentes (figure 2.13) et leurs propriétés : la 2-heptanone et l'isoamylacétate sont des phéromones non-sexuelles et la 1,8-cinéol (odeur camphrée), la 2-isobutyl-3-méthoxy-pyrazine (IBMP, odeur de poivron) et la β -ionone (odeur de violette) sont des composés floraux. A pH 7,5 dans un tampon phosphate (500mM) et à une concentration de 1,5mM en protéine, ASP2 est capable de les fixer à l'exception de la β -ionone et dans une moindre mesure la 1,8-cinéol. Dans ces conditions, les rapports odorant :dimère d'ASP2 sont estimés à environ 3,6 pour la 2-heptanone et l'isoamylacétate et à environ 2,5 pour l'IBMP et la 1,8-cinéol.

Les composés se fixant sur ASP2 ont ensuite été utilisés pour une titration par microcalorimétrie à pH 7 dans une solution tamponnée par 50mM de phosphate de sodium. Il en ressort qu'il existe deux sites de fixation principaux (micromolaire) et un site plus faible (millimolaire) de l'IBMP et l'isoamylacétate sur la structure d'ASP2. Le tableau 2.2 rassemble les constantes d'association mesurées sur les quatre molécules testées. Dans les mêmes conditions à pH 3, l'absence de fixation est confirmée pour l'IBMP et l'isoamylacétate.

A partir du rapport 3 :2 entre les concentrations en odorant et en dimère d'ASP2, les auteurs émettent la possibilité de la présence d'un site de fixation micromolaire dans chaque monomère et d'un troisième site millimolaire, sans rôle biologique, situé à l'interface.

2.3 Objectifs de la thèse

L'étude d'ASP2 par Résonance Magnétique Nucléaire devait apporter des renseignements précieux sur la fonction d'ASP2. Cette technique s'avérait a priori particulièrement appropriée

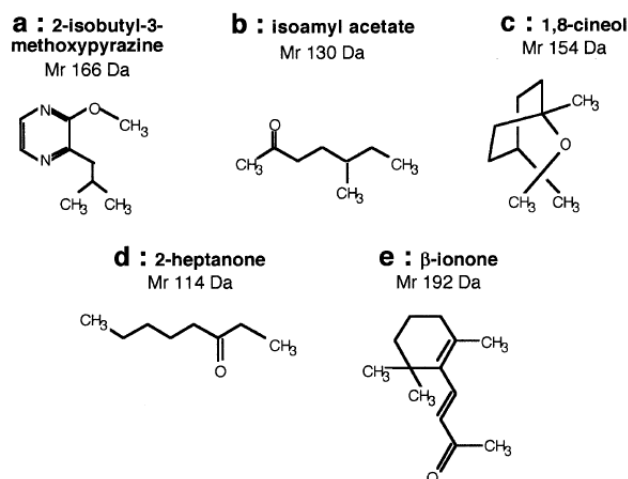


Fig. 2.13: Série de molécules étudiées pour les propriétés de fixation d'ASP2.

odorant	IBMP	isoamylacétate	1,8-cinéol	2-heptanone
Site 1, K_a	4,2±0,7	0,45±0,05	0,29±0,06	0,14±0,004
Site 2, K_a	4,8±0,7	0,28±0,03	0,32±0,06	0,12±0,005
Site 3, K_a	0,004±0,001	0,085±0,01	0,53±0,06	0,22±0,007

Tab. 2.2: Constantes d'association K_a ($M^{-1} \times 10^6$) des trois sites mesurées sur l'IBMP, l'isoamylacétate, la 1,8-cinéol et la 2-heptanone.

dans ce cas pour plusieurs raisons :

- ⇨ la forte solubilité de ce type de protéine permet d'utiliser des échantillons de concentration importante (environ millimolaire),
- ⇨ la possibilité du marquage isotopique rendait possible l'utilisation des techniques de RMN sur trois noyaux,
- ⇨ la taille relativement faible (13,6 kDa) du monomère laissait présager une relaxation des spins nucléaires suffisamment lente et donc favorable aux expériences hétéronucléaires tridimensionnelles classiques de la RMN,
- ⇨ la présence des trois ponts disulfure devait structurer fortement la chaîne polypeptidique,

Lors de début de cette étude en février 2000, la structure cristallographique de la PBP de *Bombyx Mori* n'était pas encore disponible et la prédiction de la structure secondaire d'ASP2 effectuée à cette époque laissait supposer que la structure d'ASP2 pouvait représenter un repliement nouveau. La détermination de structure d'ASP2 était donc d'un grand intérêt à la fois pour mieux appréhender son rôle biologique mais également pour ajouter un nouveau repliement dans les bases de données.

Nous nous sommes donc fixés comme objectif de déterminer la *structure* d'ASP2, sa *dynamique interne*, et ensuite les caractéristiques de ses *interactions avec quelques odorants* sélectionnés. Les chapitres suivants vont détailler les méthodes et les résultats obtenus.

Bibliographie

- B. Akerstrom, D.R. Flower, and J.P. Salier. Lipocalins 2000 : unity in diversity. *Biochim. Biophys. Acta*, 2000.
- J. Amoore. Current status of the steric theory of odour. *Ann. NY Acad. Sci.*, 116 :457–481, 1964.
- S. Angeli, F. Ceron, A. Scaloni, M. Monti, G. Monteforti, A. Minocci, R. Petacchi, and P. Pelosi. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from schistocerca gregaria. *Eur. J. Biochem.*, 262 :745–754, 1999.
- R.C. Araneda, A.D. Kini, and S. Firestein. The molecular receptive range of an odorant receptor. *Nature Neuroscience*, 3(12) :1248–1255, 2000.
- R. Axel. The molecular logic of smell. *Sci. Am.*, 273 :154–159, 1995.
- L. Ban, A. Scaloni, C. D'Ambrosio, L.Zhang, Y. Yan, and P. Pelosi. Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cell. Mol. Life Sci.*, 60 :390–400, 2003.
- S. Bette, H. Breer, and J. Krieger. Probing a pheromone binding protein of the silkworm *Antheraea polyphemus* by endogenous tryptophan fluorescence. *Insect Biochem. Mol. Biol.*, 32 :241–246, 2002.
- M.A. Bianchet, B. Bains, P. Pelosi, J. Pevsner, S.H. Snyder, H.L. Monaco, and L.M. Amzel. The three-dimensional structure of bovine odorant binding protein and its mechanisms of odor recognition. *Nat. Struct. Biol.*, 3(111) :934–939, 1996.
- M. Biri. *Le grand livre des abeilles*. Editions de Vecchi, 1997.
- J. Bohbot, F. Sobrio, P. Lucas, and P. Nagnan-Le Meillour. Functional characterization of a new class of odorant-binding proteins in the moth mamestra brassicae. *Biochem. Biophys. Res. Commun.*, 253 :489–494, 1998.
- M. Boudjelal, A. Sivaprasadarao, and J.B.C. Findlay. Membrane receptor for odour-binding proteins. *Biochem. J.*, 317 :23–27, 1996.
- L. Briand, C. Nespoulous, J. C. Huet, M. Takahashi, and J. C. Pernollet. Ligand binding and physico-chemical properties of ASP2, a recombinant odorant-binding protein from honeybee (*Apis mellifera L.*). *Eur. J. Biochem.*, 268(3) :752–60, 2001.
- L. Briand, N. Swasdipan, C. Nespoulos, V. Bézirard, F. Blon, J.-C. Huet, P. Ebert, and J.-C. Pernollet. Characterization of a chemosensory protein (ASP3C) from honeybee (*Apis Mellifera L.*) as a brood pheromone carrier. *Eur. J. Biochem.*, 269 :4586–4596, 2002.

- L. Buck and R. Axel. A novel multigen family may encode odorant receptors : a molecular basis for odor recognition. *Cell*, 65 :175–187, 1991.
- L.B. Buck. The molecular architecture of odor and pheromone sensing in mammals. *Cell*, 100 : 611–618, 2000.
- F.E. Callahan, R.G. Vogt, M.L. Tucker, J.C. Dickens, and A.K. Mattoo. High level expression of male specific pheromone binding proteins (PBPs) in the antennae of female noctuid moths. *Insect Biochem. Mol. Biol.*, 30 :507–514, 2000.
- V. Campanacci, J. Krieger, S. Bette, J.N. Sturgis, A. Lartigue, C. Cambillau, H. Breer, and M. Tegoni. Revisiting the specificity of mamestra brassicae and antherae polyphemus pheromone-binding proteins with a fluorescence binding essay. *J. Biol. Chem.*, 276 :20078–20084, 2001.
- F. Damberger, L. Nikonova, R. Horst, G. Peng, W.S. Leal, and K. Wüthrich. NMR characterization of a pH-dependent equilibrium between two folded solution conformations of the pheromone-binding protein from *Bombyx Mori*. *Prot. Sci.*, 9 :1038–1041, 2000.
- E. Danty. *Contribution à l'analyse des bases moléculaires de la détection des odeurs chez l'abeille domestique*. PhD thesis, Université de Montpellier II, 1996.
- E. Danty, G. Arnold, J.-C. Huet, D. Huet, C. Masson, and J.-C. Pernollet. Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera L.* (hymenoptera : Apidea). *Chem. Senses*, 23 :83–91, 1998a.
- E. Danty, G. Arnold, J. C. Huet, D. Huet, C. Masson, and J. C. Pernollet. Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera L.* (hymenoptera : Apidea). *Chem. Senses*, 23(1) :83–91, 1998b.
- E. Danty, L. Briand, C. Michard-Vanhee, V. Perez, G. Arnold, O. Gaudemer, D. Huet, J. C. Huet, C. Ouali, C. Masson, and J. C. Pernollet. Cloning and expression of a queen pheromone-binding protein in the honeybee : an olfactory-specific, developmentally regulated protein. *J Neurosci*, 19 : 7468–75, 1999.
- E. Danty, C. Michard-Vanhee, J. C. Huet, E. Genecque, J. C. Pernollet, and C. Masson. Biochemical characterization, molecular cloning and localization of a putative odorant-binding protein in the honey bee *Apis mellifera L.* (hymenoptera : Apidea). *FEBS Lett*, 414(3) :595–8, 1997.
- R.L. Doty. Olfaction. *Annu. Rev. Psychol*, 52 :423–452, 2001.
- L. Dryer and A. Berghard. Odorant receptors : a plethora of G-protein-coupled receptors. *TiPS*, 20 : 413–417, 1999.
- G. Du and G.D. Prestwich. Protein structure encodes the ligand binding specificity in pheromone binding proteins. *Biochemistry*, 34 :8726–8732, 1995.
- H.L. Eisthen. Why are olfactory systems of different animals so similar? *Brain Behav. Evol.*, 59 : 273–293, 2002.
- S. Firestein. How the olfactory system makes sense of scents. *Nature*, 413 :211–218, 2001.

-
- W.B. Floriano, N. Vaidehi, W.A. Goddard III, M.S. Singer, and G.M. Shepherd. Molecular mechanisms underlying differential odor responses of a mouse olfactory response. *Proc. Nat. Acad. Sci.*, 20 :10712–10716, 2000.
- D.R. Flower. The lipocalin protein family : structure and function. *Biochem.*, 318 :1–14, 1996.
- D.R. Flower. Beyond the superfamille : the lipocalin receptors. *Biochim. Biophys. Acta*, 1482 : 327–336, 2000a.
- D.R. Flower. Experimentally determined lipocalin structures. *Biochim. Biophys. Acta*, 1482 :46–56, 2000b.
- D.R. Flower, A.C.T. North, and C.E. Sansom. The lipocalin protein family : structural and sequence overview. *Biochim. Biophys. Acta*, 1482 :9–24, 2000.
- W.J. Freeman. The physiology of perception. *Scientific American*, 264 :78–85, 1991.
- K. Von Frisch. *Vie et moeurs des abeilles*. Editions Michel, Paris, 1969.
- I. Gaillard, S. Rouquier, J.-P. Pin, P. Mollard, S. Richard, C. Barnabé, J. Demaille, and D. Giorgi. A single olfactory receptor specifically binds a set of odorant molecules. *Eur. J. Neuroscience*, 15 : 409–418, 2002.
- C.G. Galizia, K. Nagler, B. Hölldobler, and R. Menzel. Odour coding is bilaterally symmetrical in the antennal lobes of honeybees (*Apis mellifera*). *Eur. J. Neurosci.*, 10 :2964–2974, 1998.
- A.N. Gilbert and S. Firestein. Dollars and scents : commercial opportunities in olfaction and taste. *Nature neuroscience*, 5 :1043–1045, 2002.
- L.A. Graham, D. Brewer, G. Lajoie, and P.L. Davies. Characterization of a subfamily of beetle odorant-binding proteins found in hemolymph. *Mol. Cell Proteomics*, 2 :541–549, 2003.
- B.S. Hansson. A bug’s smell - research into insect olfaction. *Trends Neurosci.*, 25(5) :270–274, 2002.
- R.K. Van der Meer, M.D. Breed, M.L. Winston, and K.E. Espelie. *Pheromone communication in social insects*. Westview, Boulder, Colo, 1998.
- J.G. Hildebrand and G.M. Shepherd. Mechanisms of olfactory discrimination : converging evidence for common principles across phyla. *Annu. Rev. Neurosci.*, 20 :595–631, 1997.
- A. Holley. *Eloge de l’odorat*. Editions Odile Jacob, 1999.
- R. Horst, F. Damberger, P. Luginbühl, P. Güntert, G. Peng, L. Nikonova, W.S. Leal, and K. Wüthrich. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *PNAS*, 98(25) :14375–14379, 2001.
- E. Jacquin-Joly, R.G. Vogt, and M.C. Francois P. Nagna-Le Meillour. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of mamestra brassicae. *Chem. Senses*, 26 :833–844, 2001.
- K. E. Kaissling. A quantitative model of odor deactivation based on the redox shift of the pheromone-binding protein in moth antennae. *Ann. NY Acad. Sci.*, 855 :320–2, 1998.

- K. Kajiya, K. Inaki, M. Tanaka, T. Haga, H. Kataoka, and K. Touhara. Molecular bases of odor discrimination : reconstitution of olfactory receptors that recognize sets of odorants. *J. Neuroscience*, 21(16) :6018–6025, 2001.
- M.S. Kim, A. Repp, and D.P. Smith. LUSH odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. *Genetics*, 150 :711–712, 1998.
- M.S. Kim and D.P. Smith. The invertebrate odorant-binding protein lush is required for normal olfactory behavior in *Drosophila*. *Chem. Senses*, 26 :195–199, 2001.
- A. Kowcun, N. Honson, and E. Plettner. Olfaction in the gipsy moth, *Lymantria dispar* : effect of pH, ionic strength and reductants on pheromone transport by pheromone binding proteins. *J. Biol. Chem.*, 276 :44770–44776, 2001.
- J. Krieger and H. Breer. Olfactory reception in invertebrates. *Science*, 286 :720–723, 1999.
- J. Krieger, H. Ganssle, K. Raming, and H. Breer. Odorant binding proteins of *Heliothis virescens*. *Insect Biochem. Mol. Biol.*, 23 :449–456, 1993.
- J. Krieger, E.V. von Nickisch-Roseneck, M. Mameli, P. Pelosi, and H. Breer. Binding proteins from the antennae of *Bombyx Mori*. *Insect Biochem. Mol. Biol.*, 26 :297–307, 1996.
- M.J.B. Krieger and K.G. Ross. Identification of a major gene regulating complex social behavior. *Science*, 295 :328–332, 2002.
- A. Lartigue, V. Campanacci, A. Roussel, A.M. Larsson, T.A. Jones, M. Tegoni, and C. Cambillau. X-ray structure and ligand binding study of a moth chemosensory protein. *J. Biol. Chem.*, 277 :32094–32098, 2002.
- M. Laska, C.G. Galizia, M. Giurfa, and R. Menzel. Olfactory discrimination ability and odor structure-activity relationships in honeybees. *Chem. Senses*, 24 :429–438, 1999.
- W.S. Leal. Duality monomer-dimer of the pheromone-binding protein from *Bombyx mori*. *Biochem. Biophys. Res. Comm.*, 268 :521–529, 2000.
- D. Lee, F. Damberger, G. Peng, R. Horst, P. Guntert, L. Nikonova, W. Leal, and K. Wüthrich. NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. *FEBS Letters*, 531 :314–318, 2002.
- R. Maleszka and G. Strange. Molecular cloning, by a novel approach, of a cDNA encoding a putative olfactory system in the labial palps of the moth *Cactoblastis cactorum*. *Gene*, 202 :39–43, 1997.
- B. Malnic, J. Hirono, T. Sato, and L.B. Buck. Combinatorial receptor codes for odors. *Cell*, 96 :713–723, 1999.
- M. Mameli, A. Tuccini, M. Mazza, R. Petacchi, and P. Pelosi. Soluble proteins in chemosensory organs of phasmids. *Insect Biochem. Molecul. Biol.*, 26 :875–882, 1996.
- S. Marchese, S. Angeli, A. Andolfo, A. Scaloni, A. Brandazza, M. Mazza, J.-F. Picimbon, W.S. Leal, and P. Pelosi. Soluble proteins from chemosensory organs of *Eurycantha calcarata* (insects, phasmatodea). *Insect Biochem. Molecul. Biol.*, 30 :1091–1098, 2000.

-
- V. Matarazzo, N. Zsürger, J.-C. Guillemot, O. Clot-Faybesse, J.-M. Botoo, C. Dal Farra, M. Crowe, J. Demaille, J.-P. Vincent, J. Mazella, and C. Ronin. Porcine odorant-binding protein selectively binds to a human olfactory receptor. *Chem. Senses*, 27 :691–701, 2002.
- P. Nagnan-Le Meillour, A.H. Cain, E. Jacquin-Joly, M.C. Francois, S. Ramachandran, R. Maida, and R.A. Steinbrecht. Chemosensory proteins from the proboscis of *Mamestra brassicae*. *Chem. Senses*, 25 :541–553, 2000.
- P. Monbaerts. Odorant receptor genes in humans. *Curr. Op. Gen. Dev.*, 9 :315–320, 1999a.
- P. Monbaerts. Seven-transmembrane proteins as odorant and chemosensory receptors. *Science*, 286 :707–711, 1999b.
- P. Monbaerts. The human repertoire of odorant receptor genes and pseudogenes. *Annu. Rev. Genomics Hum. Genet.*, 2 :493–510, 2001.
- K. Mori, H. Nagao, and Y. Yoshihara. The olfactory bulb : coding and processing of odor molecule information. *Science*, 286 :711–715, 1999.
- K. Mori and G.M. Sheperd. Emerging principles of molecular signal processing by mitral/tufted cells in the olfactory bulb. *Semin. Cell. Biol.*, 5(1) :65–74, 1994.
- A. Mosbah, V. Campanacci, A. Lartigue, M. Tegoni, C. Cambillau, and H. Darbon. solution structure of a chemosensory protein (CSP) from the moth *Mamestra brassicae*. *Biochem. J.*, 369 :39–44, 2002.
- H. Nagao, M. Yamaguchi, Yuji Takahashi, and K. Mori. Grouping and representation of odorant receptors in domains of the olfactory bulb sensory map. *Micros. Res. Tech.*, 58 :168–175, 2002.
- P. Nagnan-Le Meillour, J. C. Huet, M. Maibeche, J. C. Pernollet, and C. Descoins. Purification and characterization of multiple forms of odorant/pheromone binding proteins in the antennae of *Mamestra brassicae* (Noctuidae). *Insect. Biochem. Mol. Biol.*, 26(1) :59–67, 1996.
- P. Pelosi. Odorant-binding proteins. *Critic. Reviews Biochem. Molec. Biol.*, 29(3) :199–228, 1994.
- P. Pelosi. Perirecepteur events in olfaction. *J. Neurobiol.*, 30(1) :3–19, 1995.
- P. Pelosi. Odorant-binding proteins : structural aspects. *Ann. NY Acad. Sci.*, 855 :281–93, 1998.
- P. Pelosi, N. E. Baldaccini, and A. M. Pisanelli. Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine. *Biochem J*, 201(1) :245–8, Jan 1 1982.
- P. Pelosi and R. Maida. Odorant-binding proteins in vertebrates and insects : similarities and possible common function. *Chem. Senses*, 15(2) :205–215, 1990.
- P. Pelosi and R. Maida. Odorant-binding proteins in insects. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, 111(3) :503–14, 1995.
- P. Pelosi, A.M. Pisanelli, and N.E. Baldaccini A. Gagliardo. Binding of [3H]-2-isobutyl-3-methoxypyrazine to cow olfactory mucosa. *Chem. Senses*, 6 :77–85, 1981.
- J.-F. Picimbon. Les péri-récepteurs chimiosensoriels des insectes. *Médecine/Sciences*, 18 :1089–1094, 2002.

- J.-F. Picimbon, K. Dietrich, S. Angeli, A. Scaloni, J. Krieger, H. Breer, and P. Pelosi. Purification and molecular cloning of chemosensory proteins from *Bombyx mori*. *Arch. Insect Biochem. Physiol.*, 44 :120–129, 2000.
- J.-F. Picimbon, K. Dietrich, J. Krieger, and H. Breer. Identity and expression pattern of chemosensory proteins in *Heliothis virescens* (Lepidoptera, Noctuidae). *Insect Biochem. Mol. Biol.*, 31 :1173–1181, 2001.
- D. Picone, O. Crescenzi, S. Angeli, S. Marchese, A. Brandazza, L. Ferrara, P. Pelosi, and A. Scaloni. Bacterial expression and conformational analysis of a chemosensory protein from *Schistocerca gregaria*. *Eur. J. Biochem.*, 268 :4794–4801, 2001.
- K.L. Pierce, R.T. Premont, and R.J. Lefkowitz. Seven-transmembrane receptors. *Nature Reviews*, 3 :639–650, 2002.
- Y. Pilpel and D. Lancet. The variable and conserved interfaces of modeled olfactory receptor proteins. *Protein Science*, 8 :969–977, 1999.
- E. Plettner, J. Lazar, E.G. Prestwich, and G.D. Prestwich. Discrimination of pheromones enantiomers by two pheromone binding proteins from the gypsy moth *Lymantria dispar*. *Biochemistry*, 39 :8953–8962, 2000.
- B. Pophof. Moth pheromone binding proteins contribute to the excitation of olfactory receptor cells. *Naturwissenschaften*, 89 :515–518, 2002.
- G. D. Prestwich, G. Du, and S. LaForest. How is pheromone specificity encoded in proteins? *Chem. Senses*, 20(4) :461–9, 1995.
- R. Ramoni, F. Vincent, S. Grolli, V. Conti, C. Malosse, F. D. Boyer, P. Nagnan-Le Meillour, S. Spinelli, C. Cambillau, and M. Tegoni. The insect attractant 1-octen-3-ol is the natural ligand of bovine odorant-binding protein. *J. Biol. Chem.*, 276(10) :7150–5, Mar 9 2001.
- S. Rivière, A. Lartigue, B. Quenedey, V. Campanacci, J.-P. Farine, M. Tegoni, C. Cambillau, and R. Brossut. A pheromone binding protein from the cockroach *Leucophaea maderae* : cloning, expression and pheromone binding. *Biochem. J.*, 371 :573–579, 2003.
- M.E. Rogers, J. Krieger, and R.G. Vogt. Antennal SNMPs (sensory neuron membrane proteins) of Lepidoptera define a unique family of invertebrate CD36-like proteins. *J. Neurobiol.*, 49 :47–61, 2001a.
- M.E. Rogers, R.A. Steinbrecht, and R.G. Vogt. Expression of SNMP-1 in olfactory neurons and sensilla of male and female antennae of the silkworm *Antheraea polyphemus*. *Cell Tissue Res.*, 303 :433–446, 2001b.
- S. Rothmund, Y.-C. Liou, P.L. Davies, E. Krause, and F.D. Sönnichsen. A new class of hexahelical insect proteins revealed as putative carriers of small hydrophobic ligands. *Structure*, 7 :1323–1332, 1999.
- B.H. Sandler, L. Nikonova, W.S. Leal, and J. Clardy. Sexual attraction in the silkworm moth : structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.*, 7 :143–151, 2000.

-
- S. Schlehner and A. Skerra. Duocalins : engineered ligand-binding proteins with dual specificity derives from the lipocalin fold. *Biol. Chem.*, 382 :1335–1342, 2001.
- M. Spehr, J.A. Rifell, R.K. Zimmer, and H. Hatt. Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science*, 299 :2054–2058, 2003.
- S. Spinelli, R. Ramoni, S. Grolli, J. Bonicel, C. Cambillau, and M. Tegoni. The structure of the monomeric porcine odorant binding protein sheds light on the domain swapping mechanism. *Biochem.*, 37 :7913–7918, 1998.
- R.A. Steinbrecht, M. Laue, and G. Ziegelberger. Immunolocalization of pheromone binding protein and general odorant binding protein in olfactory sensilla of the silk moths *Antheraea* and *Bombyx*. *Cell tissue Res.*, 282 :203–217, 1995.
- R.A. Steinbrecht, M. Ozaki, and G. Ziegelberger. Immunocytochemical localization of pheromone-binding protein in moth antennae. *Cell. Tissue Res.*, 270 :287–302, 1992.
- J.A. St.John, H.J. Clarris, and B. Key. Multiple axon guidance cues establish the olfactory topographic map : how do these cues interact? *Int. J. Dev. Biol.*, 2002 :639–647, 2002.
- N.J. Strausfel and J.G. Hildebrand. Olfactory systems : common design, uncommon origins. *Curr. Op. Neurobiol.*, 9 :634–639, 1999.
- M. Tegoni, P. Pelosi, F. Vincent, S. Spinelli, V. Campanacci, S. Grolli, R. Ramoni, and C. Cambillau. Mammalian odorant binding proteins. *Biochim. Biophys. Acta*, 1482(1-2) :229–40, 2000.
- M. Tegoni, R. Ramoni, E. Bignetti, S. Spinelli, and C. Cambillau. Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nature Struct. Biol.*, 3 :863–867, 1996.
- E.R. Tromel. Chemosensory signaling in *C. elegans*. *BioEssays*, 21 :1011–1020, 1999.
- A. Tuccini, R. Maida, P. Rovero, , M. Mazza, and P. Pelosi. Putative odorant-binding proteins in antennae and legs of *Carausis morosus* (Insecta, Phasmatodea). *Insect Biochem. Molecul. Biol.*, 1 :19–24, 1996.
- L. Turin. A spectroscopic mechanism for primary olfactory preception. *Chem. Senses*, 21(6) :773–791, 1996.
- L. Turin and F. Yoshii. *Handbook of olfaction and gustation*, chapter Structure-odor relations : a modern perspective. Dekker, Marcel Incorporated, 2003. URL www.flexitral.com/research/review_final.pdf.
- N. Vaidehi, W.B. Floriano, R. Trabanino, S.E. Hall, P. Freddolino, E.J. Choi, G. Zamanakos, and W.A. Goddard III. Prediction of structure anf function of G protein-coupled receptors. *Proc. Nat. Acad. Sci.*, 99 :12622–12627, 2002.
- F.G. van der Got, J.M. Gonzalez-Manas, J.H. Lakey, and F. Pattus. A molten globule membrane-insertion at histidines of aerolysin from *Aeromonas hydrophila* : a lipid planar bilayer study. *Nature*, 354 :408–410, 1991.

- F. Vincent, S. Spinelli, R. Ramoni, S. Grolli, P. Pelosi, C. Cambillau, and M. Tegoni. Complexes of porcine odorant binding protein with odorant molecules belonging to different chemical classes. *J. Mol. Biol.*, 300(1) :127–39, 2000.
- R. G. Vogt, F. E. Callahan, M. E. Rogers, and J. C. Dickens. Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). *Chem Senses*, 24(5) :481–95, 1999.
- R. G. Vogt, A. C. Kohne, J. T. Dubnau, and G. D. Prestwich. Expression of pheromone binding proteins during antennal development in the gypsy moth *Lymantria dispar*. *J Neurosci*, 9(9) : 3332–46, 1989.
- R. G. Vogt, G. D. Prestwich, and M. R. Lerner. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. *J Neurobiol*, 22(1) :74–84, 1991.
- R. G. Vogt, L. M. Riddiford, and G. D. Prestwich. Kinetic properties of a sex pheromone-degrading enzyme : the sensillar esterase of *Antheraea polyphemus*. *Proc. Natl. Acad. Sci. USA*, 82(24) : 8827–8831, 1985.
- R.G. Vogt and L.M. Riddiford. Pheromone binding and inactivation by moth antennae. *Nature (Lond.)*, 293 :161–163, 1981.
- L.B. Vosshall. The molecular logic of olfaction in drosophila. *Chem. Senses*, 26 :207–213, 2001.
- H. Wojtasek and W. S. Leal. Conformational change in the pheromone-binding protein from *Bombyx mori* induced by pH and by interaction with membranes. *J Biol Chem*, 274(43) :30950–30956, 1999.
- G.A. Wright, B.D. Skinner, and B.H. Smith. Ability of honeybee, *Apis mellifera*, to detect and discriminate odors of varieties of canola (*Brassica rapa* and *Brassica napus*) and snapdragon flowers (*Antirrhinum majus*). *J. Chem. Ecology*, 28(4) :721–740, 2002.
- R.H. Wright. *The sense of smell*. CRCpress, Boca Raton, Florida, USA, 1982.
- Y. Zhang, J.H. Chou, J. Bradley, and C.I. Bargmann. The *Caenorhabditis elegans* seven-transmembrane protein ODR-10 functions as an odorant receptor in mammalian cells. *Proc. Nat. Acad. Sci. USA*, 1997 :12162–12167, 1997.
- H. Zhao, L. Ivic, J.M. Otaki, M. Hashimoto, and K. Mikoshiba. Functional expression of a mammalian odorant receptor. *Science*, 279 :237–242, 1998.
- Z. Zou, L.F. Horowitz, J.-P. Montmayeur, S. Snapper, and L.B. Buck. Genetic tracing reveals a stereotyped sensory map in the olfactory cortex. *Nature*, 414 :173–179, 2001.

Deuxième partie
Matériel et méthodes

Chapitre 1

Production et purification d'ASP2 marquée en vue de l'étude par RMN

Loïc Briand et ses collègues ont mis au point le protocole de surexpression et de purification d'ASP2 dans la levure *Pichia pastoris*. Ce travail a fait l'objet de deux publications : Briand et al. [1999] et Briand et al. [2001a]. La première décrit précisément le protocole tandis que la seconde montre que le système *Pichia pastoris* restitue un système d'expression hétérologue intéressant pour disposer d'une protéine marquée en isotopes ^{13}C et ^{15}N , prérequis pour une étude par RMN. Nous allons résumer ces deux articles.

1.1 Pourquoi utiliser un tel système de surexpression ?

Les systèmes bactériens du type *E. coli* ne disposant pas des outils nécessaires aux modifications post-traductionnelles, elles semblaient présenter peu d'avantages pour la production d'ASP2. D'autres OBPs ont été exprimées dans des corps d'inclusion (Plettner et al. [2000]) et repliées correctement, ou dans des cellules d'insecte (Krieger et al. [1992]) mais *Pichia pastoris* fournit les meilleurs éléments de surexpression de ce type de protéine : formation des ponts disulfure, pas de glycosylation, rendement important et marquage isotopique possible.

La particularité de cette levure méthylotrophe est de posséder un promoteur de l'alcool oxydase (AOX1), régulé fortement par le méthanol. En absence de méthanol, la cellule utilise la voie classique du métabolisme du glucose. Lorsque le méthanol est la seule source de carbone du milieu, la voie AOX1 est activée. L'alcool oxydase dégrade alors le méthanol en formaldéhyde à partir de l'oxygène moléculaire. L'alcool oxydase devient alors la protéine majoritairement exprimée ($\geq 30\%$ du total des protéines cellulaires) pour compenser sa faible affinité pour l'oxygène.

Les avantages de la levure *Pichia pastoris* pour le marquage isotopique d'une protéine ont été montrées par d'autres productions de protéines destinées à être étudiée dans un spectromètre de RMN (Wood and Komives [1999], Laroche et al. [1996]). Les rendements obtenus sont souvent très intéressants dans la mesure où le niveau de production de ses propres protéines est faible (4%, Briand et al. [1999]). Par ailleurs, elle arrive à croître en milieu minimum en conservant une densité cellulaire importante, indispensable à une production à prix raisonnable de protéine simplement et doublement marquée en isotopes stables. Il suffit de remplacer le

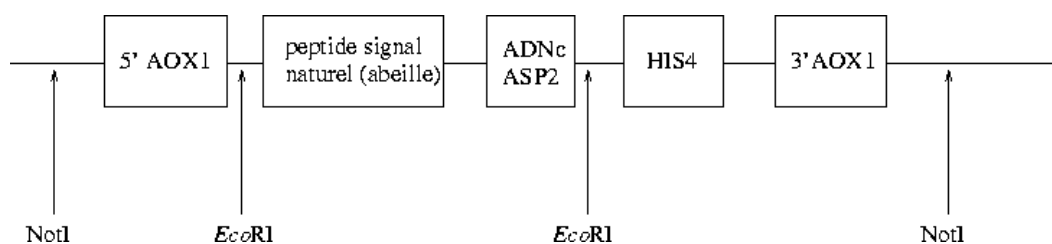


Fig. 1.1: Construction du vecteur d'expression pHIL-D2 pour ASP2.

sulfate d'ammonium par le sulfate d'ammonium marqué en ^{15}N pour marquer simplement et d'utiliser du ^{13}C -glucose et du ^{13}C -méthanol pour le double marquage.

1.2 Construction

Les souches *GS115* (*his4⁻*) de *Pichia pastoris* (Invitrogen) ont été transformées par la méthode des sphéroplastes avec le vecteur pHIL-D2 (pPIC9). La construction du vecteur est représenté sur la figure 1.1. Le clone codant pour ASP2 a été amplifié par PCR en utilisant un primer contenant la séquence du peptide signal naturel permettant la sécrétion de la protéine dans le milieu de culture.

1.3 Surexpression

Un clone producteur congelé à $-80^{\circ}C$ a été réisolé sur une boîte YPD incubée à $-30^{\circ}C$ pendant 48h. La croissance des cellules est initiée par une préculture de 5mL d'une colonie de levures transformées dans le milieu BMGlu (300 rpm, $29^{\circ}C$, 24h) dans un erlenmeyer de 500mL. 245mL de BMGlu sont ensuite ensemencés par les 5mL de préculture durant 24h (300 rpm, $29^{\circ}C$). Les cellules sont séparées du milieu BMGlu par centrifugation (600g, 10mn, $T_{ambiante}$) et le culot cellulaire est resuspendu dans 250 mL de BMM pour l'induction au méthanol (300 rpm, $29^{\circ}C$). La durée d'induction a varié de deux jours et demi à 6 jours. La concentration en méthanol fut maintenue constante à 0,5%, par ajout de 0,625mL de méthanol pur toutes les 12 heures environ. Après induction, le milieu est centrifugé (1500g, 20mn, $T_{ambiante}$) puis filtré à $0,2\mu m$. ASP2 se trouve sécrétée dans le surnageant.

La solution de BMM contenant ASP2 soluble est dialysée dans un tube à dialyse (seuil de coupure 8000Da, Servapor, France) à $4^{\circ}C$ pendant 3 jours contre de l'eau milliQ. Le bain de dialyse est renouvelé deux fois par jour. L'échantillon est ensuite congelé à $-80^{\circ}C$ puis lyophilisé durant 3 jours.

1.4 Purification

La purification d'ASP2 est effectuée en une seule étape par chromatographie HPLC en phase inverse sur une colonne C8 (Aquapore, 10 d.i. x 30mm, Perkin Elmer, France). La colonne est préalablement équilibrée avec l'éluant A (19% acétonitrile dans l'eau milliQ, 25mM

Milieu BMGlu	Milieu BMM
0,34% YNB (sans acide aminé) 4 μ g/ml D-biotine 100mM phosphate de potassium 1% glucose 1% sulfate d'ammonium pH = 6	0,34% YNB (sans acide aminé) 4 μ g/ml D-biotine 100mM phosphate de potassium 0.5% méthanol 1% sulfate d'ammonium pH = 6

Tab. 1.1: Composition des milieux BMGlu et BMM

acétate d'ammonium, pH 7). Après chargement, un gradient linéaire est appliqué de 19% à 24% en acétonitrile. Cette concentration est maintenue constante pendant 9min pour éliminer les impuretés hydrophiles, puis jusqu'à 100% de la solution B (50% d'acétonitrile dans l'eau milliQ, 25mM acétate d'ammonium, pH 7) en 50min. Le flux est ajusté à 2,5ml/mn. La mesure de densité optique à 215nm permet de suivre la concentration protéique en sortie de colonne. L'échantillon en sortie de colonne est de nouveau congelé à $-80^{\circ}C$ puis lyophilisé pendant 3 jours.

1.5 Analyse

L'analyse en spectrométrie de masse MALDI-TOF permet de contrôler la qualité de la protéine produite. 1 μ l de surnageant dialysé est mélangé à 24 μ l de matrice (solution saturée d'acide α -cyano-4-hydroxycinnamique dans 50% d'acétonitrile, 0,15% v/v d'acide trifluoroacétique). Un microlitre de solution est placé ensuite sur la cible et séché à l'air libre. L'analyse a été effectuée sur un PE Voyager-STR DE de Perkin-Elmer. La thiorédoxine de *E. coli* (11674, 49Da) et l'apomyoglobine du coeur de cheval (16952, 55Da) servent de calibration des masses.

Chapitre 2

Attribution de la protéine

L'étude par Résonance Magnétique Nucléaire de macromolécules biologiques ne devient en général exploitable qu'à partir du moment où on dispose de l'attribution des fréquences de résonance (déplacement chimique) des noyaux observables (1H , ^{13}C , ^{15}N) : c'est la première étape de la plupart des travaux sur les protéines. Au début de l'histoire de la RMN des protéines, les noyaux d'hydrogène étaient les seuls noyaux de spin 1/2 présents en abondance naturelle et exploitables dans ces macromolécules. La méthode d'attribution des hydrogènes est amplement discutée par Kurt Wüthrich dans son ouvrage "NMR of Proteins and Nucleic Acid" (John Wiley, 1986). Cette approche exploite les informations disponibles dans les expériences de type HOHAHA et NOESY en deux dimensions : la HOHAHA fournit les signatures de chaque type d'acide aminé tandis que la NOESY permet de déterminer l'enchaînement séquentiel par les corrélations $H_N^{i+1} \longleftrightarrow H_\alpha^i$ et $H_N^{i+1} \longleftrightarrow H_N^i$. Cette stratégie est applicable sur des protéines de taille limitée (inférieure à une centaine d'acides aminés). Au-delà, l'encombrement des spectres augmente le nombre de superpositions et donc l'ambiguïté de l'attribution séquentielle. De plus, cette méthode repose sur des corrélations spatiales et non scalaires, il en résulte que les erreurs d'attribution séquentielle sont aisées : une corrélation interprétée comme séquentielle peut en réalité, correspondre à une corrélation à longue distance (exemple : $H_N^i \longleftrightarrow H_\alpha^{i-3}$ dans une hélice α .)

L'avènement du marquage isotopique des protéines et des sondes trois noyaux a permis l'émergence d'une autre stratégie d'attribution non ambiguë. Elle est basée sur les transferts de cohérence par les couplages scalaires $^1J_{H-N}$ et $^1J_{H-C}$ et hétéronucléaires ($^1J_{C-N}$ et $^1J_{C-C}$) pour une protéine doublement marquée ^{13}C , ^{15}N . Dans un premier temps, il s'agit de travailler sur le squelette peptidique puis d'étendre l'attribution aux chaînes latérales.

2.1 Attribution du squelette peptidique.

L'attribution du squelette peut être obtenue par une série d'expériences tridimensionnelles hétéronucléaires manipulant les spins des quatre noyaux H_N , N , C_α et/ou C' , voire éventuellement le C_β : HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCANH et CBCA(CO)NH. Elles sont largement décrites dans plusieurs publications et livres (Cavanagh et al. [1996], Sattler et al. [1999]). Dans un premier temps, nous allons décrire la stratégie d'attribution puis nous allons détailler les expériences CBCANH et COCAH.

2.1.1 La stratégie d'attribution

Le nom des expériences provient du type de noyau édité, correspondant donc à une des trois dimensions et des chemins de cohérence suivis. Par exemple, la HNCO corrèle les noyaux $H_N^i - N^i - C'^{i-1}$. Les parenthèses autour d'un noyau signifient que le chemin de cohérence passe par ce noyau mais qu'il n'est pas édité : l'expérience HN(CA)CO corrèle les noyaux $H_N^i - N^i - C'^{i-1}$ et $H_N^i - N^i - C''^i$. Le tableau 2.2 relie les expériences aux corrélations observées.

Les expériences tridimensionnelles sont interprétées par paires : HNCO et HN(CA)CO, HNCA et HN(CO)CA, CBCANH et CBCA(CO)NH. En effet, l'expérience HNCA permet d'observer pour chaque $H_N - N$ du résidu i , le C_α intrarésiduel et le C_α^{i-1} interrésiduel. L'expérience HN(CO)CA (resp. HN(CO)CACB) permet de trancher car on n'y observe que les corrélations intrarésiduelles. En théorie, ces deux expériences sont suffisantes pour déterminer les relations de proximité séquentielle des tâches de corrélation de la ^{15}N -HSQC. Il faut également relier l'enchaînement des corrélations à la séquence protéique. Les déplacements chimiques des C_α de la plupart des acides aminés ne sont pas suffisamment caractéristiques pour les reconnaître rapidement. Excepté le C_α de la glycine qui va être observé autour de 40 ppm, les autres acides aminés se rangent par classe de déplacement chimique (voir figure 2.1). Par conséquent, connaître le déplacement chimique du C_α n'est pas suffisant pour placer un enchaînement de corrélations entre $H_N - N$ sur la séquence protéique. Pour résoudre ce problème, on utilise le couple CBCANH et CBCA(CO)NH qui porte les corrélations sur les C_α et C_β intrarésiduels et interrésiduels. La contrepartie à l'extension jusqu'aux C_β est une perte importante de la sensibilité. Par contre, son intérêt réside dans la forte dépendance du déplacement chimique de C_α et de C_β avec le type d'acide aminé. La combinaison des déplacements chimiques de ces deux carbones permet de lever l'ambiguïté sur la plupart des acides aminés. Le tableau 2.1 rassemble les classes de acides aminés distinguables par la simple connaissance des déplacements chimiques de ces deux carbones. Il faut cependant prendre des précautions car la structure secondaire peut influencer considérablement et faire sortir un carbone de la zone décrite dans la figure 2.3.

Parallèlement aux chemins par les carbones C_α et C_β , le couple HNCO - HN(CA)CO utilise le carbonyle de chaque acide aminé, en corrélant respectivement $H_N^i - N - C'^{i-1}$ et $(H_N^i - N - C'^{i-1} - H_N^i - N - C''^i)$. Il s'agit donc d'un moyen aisé de confirmer l'attribution basée sur la HNCA et de l'étendre au-delà des ambiguïtés (superpositions ou absence de signal) de la HNCA. Si la HNCO est l'expérience hétéronucléaire la plus sensible du fait des grandes valeurs de couplages scalaires impliqués (figure 2.1), elle n'est pas utilisée dans un premier temps pour l'attribution du squelette dans la mesure où la dispersion des carbonyle est relativement faible (environ 12ppm) par rapport aux carbones C_α (environ 30ppm) et la HN(CA)CO, contenant l'information de connectivité, est bien moins sensible (Sattler et al. [1999]).

L'utilisation des deux voies de connectivité est souvent indispensable pour établir l'attribution complète du squelette. Une expérience COCAH qui relie $C''-C_\alpha$ et H_α confirme l'attribution si on dispose de l'attribution des hydrogènes H_α . Si on ne connaît pas le déplacement chimique du H_α , cette expérience la fournit.

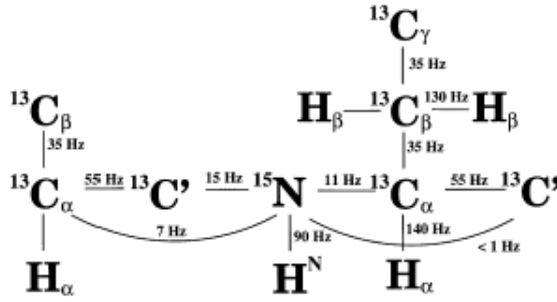


Fig. 2.1: Constantes de couplage entre les noyaux d'un peptide.

Ala	Gly	Pro, Val
Cys _{oxydée} , Asp, Phe, Leu, Asn, Tyr	Ile	Ser
Cys _{réduite} , Glu, His, Gln, Arg, Trp	Lys, Met	Thr

Tab. 2.1: Classes des acides aminés discernables sur la base des déplacements chimiques des C_α et C_β .

2.1.2 Deux expériences hétéronucléaires tridimensionnelles : CBCANH et COCAH

2.1.2.1 La CBCANH

Nous nous contenterons de détailler une seule séquence se terminant par l'acquisition de l'hydrogène amide : la CBCANH (Grzesiek and Bax [1992a]). Les autres du même type fonctionnent sur les mêmes principes spectroscopiques.

La séquence d'impulsions de la CBCANH que j'ai utilisée est représentée dans la figure 2.4. Les rectangles étroits et larges sont respectivement des impulsions 90° et 180° . Toutes les impulsions sur le canal carbone sont sélectives (gaussiennes). La fréquence du canal est centrée au milieu des carbones aliphatiques. Les impulsions sur le carbonyle sont hors résonance. Le second pulse 90° sur les carbones aliphatiques (marqué sp8) a un profil symétrique par rapport au premier. Nous pouvons isoler 5 étapes dans ce chemin de cohérence qui part des hydrogènes $H_\alpha - H_\beta$ pour finir à l'acquisition sur les hydrogènes amides :

a à b Transfert INEPT des hydrogènes aliphatiques vers leurs carbones par le couplage scalaire ${}^1J_{HC} \simeq 140\text{Hz}$. En *b*, l'aimantation se trouve suivant $H_z C_x$

b à c Temps constant T_C : édition du déplacement chimique des carbones C_α et C_β , découplage de ${}^1J_{C_{ali}C'}$ et ${}^1J_{C_{ali}N}$, transfert de l'aimantation $H_z C_x^\beta$ vers $H_z C_y^\beta C_z^\alpha$, refocalisation $H_z C_x \rightarrow C_y$ pendant le délai $2\tau_e = \frac{1}{2J_{HC_{ali}}}$. En *c*, les aimantations intéressantes sont suivant les cohérences $C_y^\beta C_z^\alpha$ et C_x^α

c à d Passage des états $C_y^\beta C_z^\alpha$ et C_x^α vers $C_z^\beta C_y^\alpha$ et C_x^α

d à e Transfert de l'aimantation vers les azotes des résidus *i* et *i*+1 avec découplage de ${}^1J_{HC}$ et refocalisation du terme $C_z^\beta C_y^\alpha$ vers un état pur C_x^α . En *e*, l'aimantation est suivant $C_y N_z^i$ et $C_y N_z^{i+1}$.

e à f Les états précédents sont transformés en : $C_z N_y^i$ et $C_z N_y^{i+1}$.

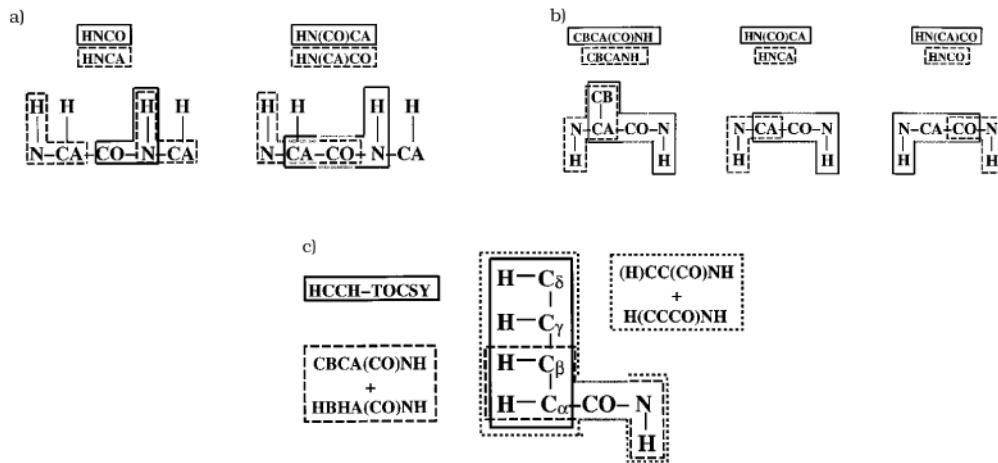


Fig. 2.2: Stratégie d'attribution d'une protéine marquée ^{15}N - ^{13}C basée sur expériences tridimensionnelles à base de transfert de cohérence. Les boîtes noires encadrant un système de spins indique la connectivité établie dans les expériences respectives. La partie grise de la boîte signifie que le spin est utilisé comme relais, sans édition du déplacement chimique. L'attribution séquentielle des résonances du squelette est obtenue par la combinaison des expériences **a**). L'attribution est améliorée par la combinaison des C_α et des C_β (**b**). Les systèmes de spin (H,C) identifiés dans une HCCH-TOCSY sont reliés aux résonances (H,N) du squelette en utilisant les déplacements chimiques des $H_{\alpha,\beta}$ et $C_{\alpha,\beta}$ des spectres HCCH-TOCSY et CBCA(CO)NH ou HBHA(CO)NH (reproduction de Sattler et al. [1999]).

f à g Temps constant T_N : édition du déplacement chimique de l'azote, découplage de $^1J_{NC'}$ et de $^1J_{NC_\alpha}$, création d'un terme $H_z N_y$ en antiphase. En *g*, l'aimantation est suivant $N_y^i H_z^i$ et $N_y^{i+1} H_z^{i+1}$.

g à h Passage des états $N_y^i H_z^i$ et $N_y^{i+1} H_z^{i+1}$ vers $N_z^i H_y^i$ et $N_z^{i+1} H_y^{i+1}$

h à i Refocalisation vers un état en phase H_y , suppression du solvant par watergate

Temps constant T_C Nous allons détailler le premier temps constant T_C . Entre les instants *b* et *c*, plusieurs termes d'interaction de couplage scalaire ne commutent pas avec la cohérence $H_z C_x$, ils peuvent donc a priori modifier cette cohérence. La construction du temps constant a pour but d'en éliminer les effets, ou au contraire de les conserver. Leur effet dépend de la durée totale pendant laquelle ils vont s'exercer. Il s'agit donc de calculer cette durée sachant qu'une impulsion 180° a pour effet d'inverser le signe de l'interaction scalaire si elle n'affecte qu'un seul des noyaux. Les lignes suivantes montrent la durée totale d'évolution de chacun des termes d'interaction ne commutant pas avec $H_z C_x$: $C_z^{ali} H_z$, $C_z^{ali} N_z$, $C_z^{ali} C'_z$, $C_z^\alpha C_z^\beta$ et C_z^{ali} .

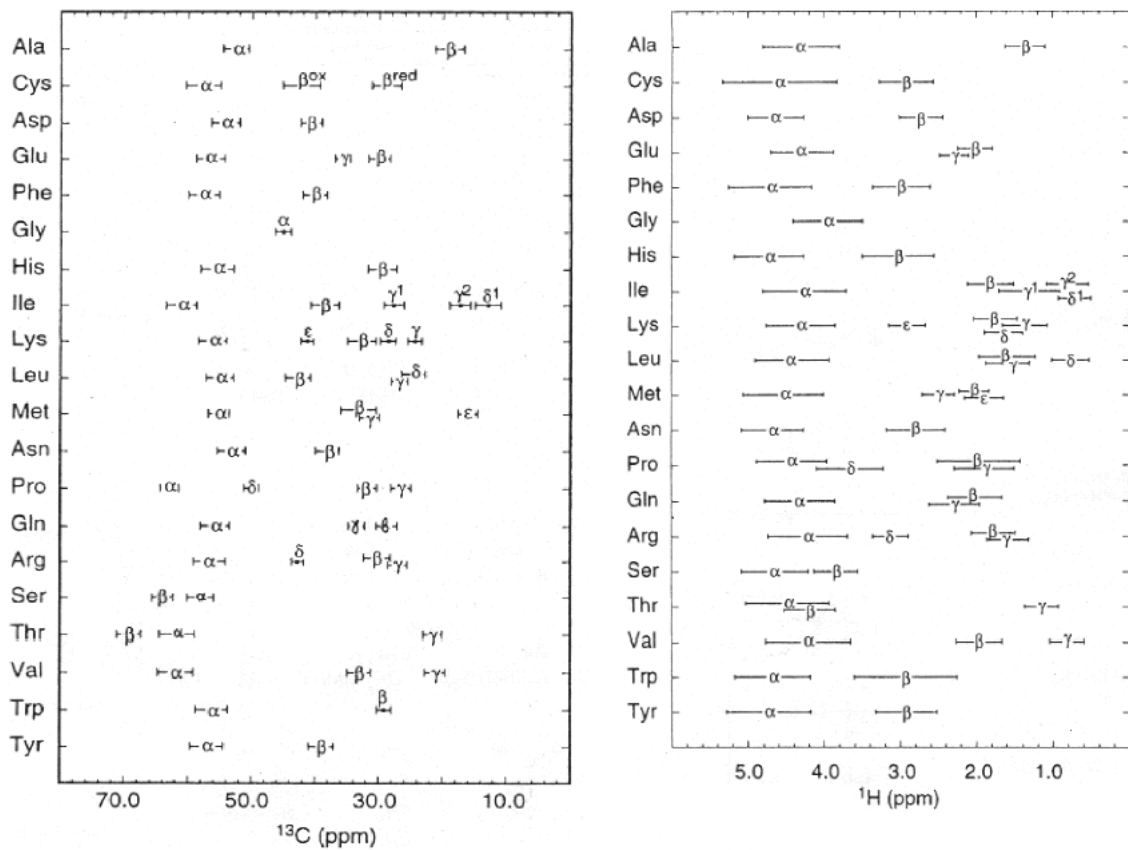


Fig. 2.3: Tableau des déplacements chimiques des noyaux de chaque type d'acide aminé, attendus dans une structuration en pelote statistique ou *random coil*. (Cavanagh et al. [1996])

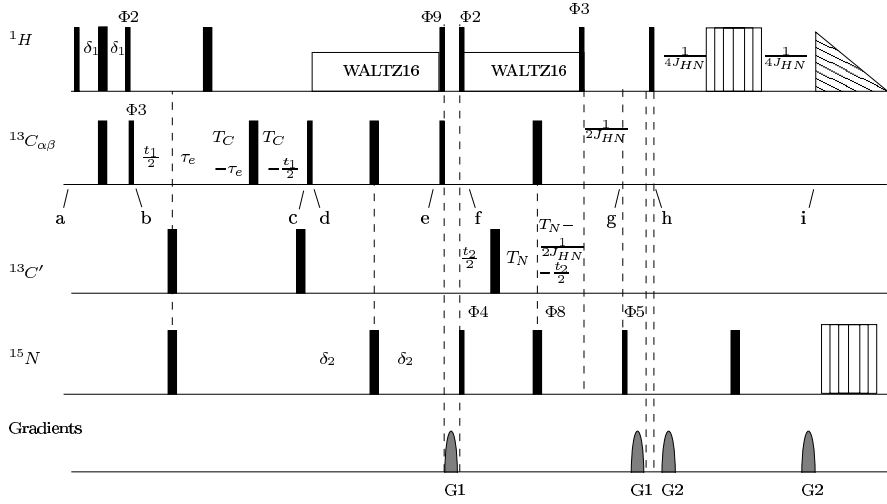


Fig. 2.4: Séquence d'impulsion de la 3D-CBCANH. La suppression de l'eau est assurée par une Watergate. Le cyclage de phase sur 16 phases est le suivant : $\Phi_2 = y$; $\Phi_3 = x, -x$; $\Phi_4 = 8x, -8x$; $\Phi_5 = 2x, -2x$; $\Phi_6 = -2y, 2y$; $\Phi_7 = -y$; $\Phi_8 = 4x, -4x$; $\Phi_9 = -x, -y$; $\Phi_{recepteur} = x, -x, -x, x, x, -x, -x, x, -x, x, x, -x, -, x, x, -x$. Les pulses non marqués sont suivant x. Les gradients G1 et G2 sont fixés à 10% et 20%, respectivement suivant les axes x et z . Les délais sont choisis à : $\delta_1 = \frac{1}{4J_{HC}} = 1,7ms$; $\delta_2 = 11ms$; $T_C = 3,6ms$; $\tau_e = 1,1ms$; $T_N = 12,4ms$; $\frac{1}{2J_{HN}} = 5,5ms$. Le mode d'acquisition est en TPPI dans les deux dimensions indirectes (phases Φ_3 et Φ_5)

$${}^1J_{C_{ali}H} : \frac{t_1}{2} + \tau_e - T_C + \tau_e + T_C - \frac{t_1}{2} = 2\tau_e \quad (2.1)$$

$${}^1J_{C_{ali}N} : \frac{t_1}{2} - \tau_e - T_C + \tau_e + T_C - \frac{t_1}{2} = 0 \quad (2.2)$$

$${}^1J_{C_{ali}C'} : \frac{t_1}{2} - \tau_e - T_C + \tau_e + T_C - \frac{t_1}{2} = 0 \quad (2.3)$$

$${}^1J_{C_{\alpha}C_{\beta}} : \frac{t_1}{2} + \tau_e + T_C - \tau_e + T_C - \frac{t_1}{2} = 2T_C \quad (2.4)$$

$$C_z^{ali} : \frac{t_1}{2} + \tau_e + T_C - \tau_e - T_C + \frac{t_1}{2} = t_1 \quad (2.5)$$

Les interactions scalaires du carbone aliphatique vers l'azote et le carbonyle ont une durée équivalente d'application nulle (lignes 2.2 et 2.3), ces termes seront refocalisés et ils n'auront pas d'effet sur l'évolution de la cohérence $H_z C_x$. Le terme d'interaction $C_z^{ali} H_z$ s'exerce pendant $2\tau_e$ (ligne 2.1). Le délai $2\tau_e$ est ajusté à $\frac{1}{2J_{CH}}$ pour transférer l'aimantation vers un état purement carbone. L'expérience n'étant pas de type aller-retour, ce transfert permet de s'éliminer l'hydrogène de la cohérence.

La ligne 2.4 met en évidence l'évolution de la cohérence de départ suivant le couplage entre les carbones aliphatiques. Ce transfert dû au couplage scalaire de $35Hz$ est indispensable pour observer au final un signal modulé par le déplacement chimique du C_{β} . Le lecteur attentif remarquera qu'il y a des pertes : de C_{α} vers $C_{\beta x}$ par le couplage entre les deux carbones et, éventuellement de $C_{\beta x}$ vers les C_{γ} par couplage scalaire. Les cohérences qui vont être conservées après l'instant c seront C_x^{α} et $C_y^{\beta} C_z^{\alpha}$. Ce second terme sera refocalisé par le module suivant.

Le dernier terme d'interaction (ligne 2.5) correspond au déplacement chimique des carbones, qui module la cohérence de départ par un facteur $\cos(2\pi\omega_C t_1)$, ω_C correspondant au déplacement chimique soit du C_α , soit du C_β .

La seconde impulsion sur le carbonyle a pour fonction d'éliminer l'erreur de phase induite par l'effet Bloch-Siegert (Sattler et al. [1999]) par la première impulsion sur l'aimantation transverse C_α/C_β .

Transfert carbone vers azote et refocalisation du terme $C_y^\beta C_z^\alpha$ Les constantes de couplage scalaire intra- et interrésiduels entre les carbones C_α et N ont des valeurs comprises respectivement dans les intervalles ${}^1J_{C_\alpha^i N^i} = 9 - 12.5 Hz$ et ${}^1J_{C_\alpha^i N^{i+1}} = 6 - 9 Hz$. D'autre part, ${}^1J_{C_\alpha C_\beta} = 35 - 40 Hz$. Avec $\delta_2 = 11 ms$, les transferts $C^i \rightarrow N^i$ et $C^i \rightarrow N^{i+1}$ sont optimisés et le terme $C_y^\beta C_z^\alpha$ est refocalisé vers un terme purement C_z^α . En e , l'aimantation est $C_z^\alpha N_z^i + C_z^\alpha N_z^{i+1}$, quel que soit l'hydrogène d'origine.

Temps constant T_N L'analyse du second temps constant T_N est analogue au précédent. Le couplage NC' est refocalisé sans effort (ligne 2.8) tandis que le terme en antiphase vis-à-vis du carbone est refocalisé grâce au couplage NC_α appliqué durant $2T_N$ (ligne 2.7) : la durée totale de ce temps constant est de $24.8 ms$. Le découplage en protons durant la première partie du temps constant détruit le terme de couplage scalaire HN . Il ne s'exprime que pendant $\frac{1}{2J_{HN}}$. Cela permet de créer en g la cohérence en anti-phase $H_z N_x$. Comme attendu, le déplacement chimique de l'azote s'exprime durant t_2 .

$${}^1J_{NH} : \quad 0 + \frac{1}{2J_{HN}} = \frac{1}{2J_{HN}} \quad (2.6)$$

$${}^1J_{NC_\alpha} : \quad \frac{t_2}{2} + T_N + T_N - \frac{1}{2J_{HN}} - \frac{t_2}{2} + \frac{1}{2J_{HN}} = 2T_N \quad (2.7)$$

$${}^1J_{NC'} : \quad \frac{t_2}{2} - T_N + T_N - \frac{1}{2J_{HN}} - \frac{t_2}{2} + \frac{1}{2J_{HN}} = 0 \quad (2.8)$$

$$N_z : \quad \frac{t_2}{2} + T_N - T_N + \frac{1}{2J_{HN}} + \frac{t_2}{2} + \frac{1}{2J_{HN}} = t_2 \quad (2.9)$$

Acquisition Le signal acquis est de la forme :

$$\begin{aligned} S(t) = & A_1 \cos(2\pi\delta_{C_\alpha^i} t_1) \cos(2\pi\delta_{N^i} t_2) \cos(2\pi\delta_{H_N^i} t_3) & H_\alpha^i - C_\alpha^i - N^i - H_N^i \\ & + A_2 \cos(2\pi\delta_{C_\alpha^i} t_1) \cos(2\pi\delta_{N^{i+1}} t_2) \cos(2\pi\delta_{H_N^{i+1}} t_3) & H_\alpha^i - C_\alpha^i - N^{i+1} - H_N^{i+1} \\ & + A_3 \cos(2\pi\delta_{C_\beta^i} t_1) \cos(2\pi\delta_{N^i} t_2) \cos(2\pi\delta_{H_N^i} t_3) & H_\beta^i - C_\beta^i - N^i - H_N^i \\ & + A_4 \cos(2\pi\delta_{C_\beta^i} t_1) \cos(2\pi\delta_{N^{i+1}} t_2) \cos(2\pi\delta_{H_N^{i+1}} t_3) & H_\beta^i - C_\beta^i - N^{i+1} - H_N^{i+1} \end{aligned}$$

Les coefficients A contiennent les fonctions trigonométriques associées au chemin de cohérence marqué à droite et différent en fonction du type d'acide aminé impliqué. Après transformée de Fourier tridimensionnelle, chaque bande extraite aux déplacements chimiques d'un couple $H_N - N$ contient quatre taches de corrélation correspondant aux carbones C_α et C_β intrarésiduels et interrésiduels (en $i-1$). Les résonances C_α de glycine et tous les C_β sont de signe opposé aux C_α des autres résidus. Cette expérience est relativement peu sensible (25 fois

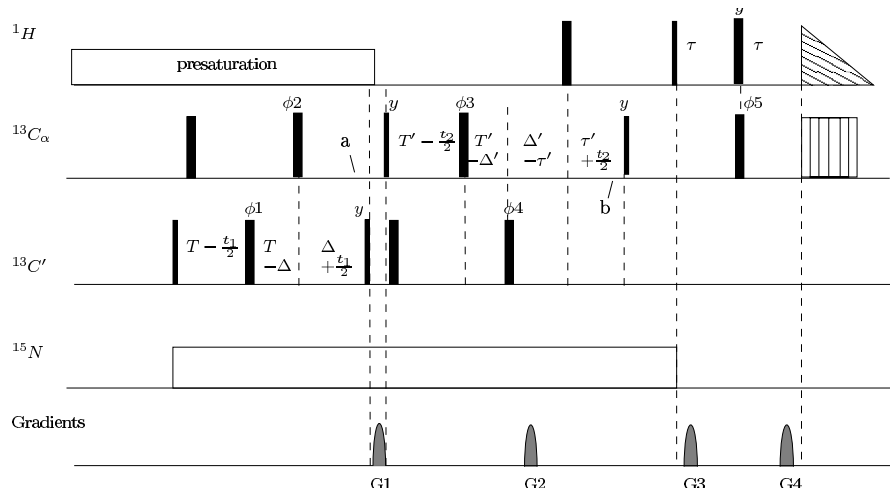


Fig. 2.5: Séquence d'impulsions COCAH. Les rectangles larges (resp. étroits) représentent des impulsions 180° (resp. 90°). Toutes les impulsions sur les canaux carbones sont sélectives (profil gaussien). Par défaut, les impulsions sont de phase x . Le cyclage de phases est le suivant : $\phi_1 = x, y; \phi_2 = 8x, 8(-x); \phi_3 = 2x, 2y; \phi_4 = 4y, 2(-y); \phi_5 = 2y, 2(-y), \phi_{acq} = x, 2(-x), x$. Les délais sont fixés à : $T = 4,6ms$, $\Delta = 4,5ms$, $T' = 3,5ms$, $\Delta' = 3,4ms$, $\tau = 1,6ms$, $\tau' = 1,1ms$. La sélection des cohérences est effectuée par la méthode echo-antiecho sur les gradients 2 et 4.

moins que la HNCO, Sattler et al. [1999]) du fait de la répartition de l'aimantation initiale de l'hydrogène sur les carbones de deux résidus.

2.1.2.2 Expérience COCAH

L'expérience COCAH permet de corréliser les résonances C', C_α et H_α d'un même résidu. La séquence d'impulsions est représenté sur la figure 2.5. Elle est basée sur la séquence publiée par Dijkstra (Dijkstra et al. [1994]). Je l'ai améliorée en utilisant des gradients pour nettoyer les cohérences non voulues et sélectionner le transfert carbone vers hydrogène tout en éliminant le signal du solvant.

Schématiquement, cette séquence commence sur le carbonyle. Le temps constant qui suit édite le déplacement chimique du carbonyle tout en découplant en proton et en azote et en laissant le couplage $C'-C_\alpha$ s'exprimer. Au point a, l'aimantation est sur $C'_x C_z^\alpha$. Après les deux impulsions $\pi/2$ sur le carbone, l'aimantation est sur $C'_z C_x^\alpha$. Durant le second temps constant (jusqu'en b), l'aimantation est modulée par le déplacement chimique de C_α et évolue sous les couplages $C' C_\alpha$ et $C_\alpha H_\alpha$. En b, l'aimantation est sur $C_x^\alpha H_z^\alpha$. Le reste de la séquence ramène l'aimantation sur l'hydrogène (INEPT retour) avant l'acquisition. Les deux temps constants commencent par des inversions pour compenser l'effet Bloch-Siegert.

Outre les corrélations du squelette, cette expérience corrèle également les noyaux $H_{ali}-C_{ali}-C'$ des résidus Asp, Asn, Glu et Gln.

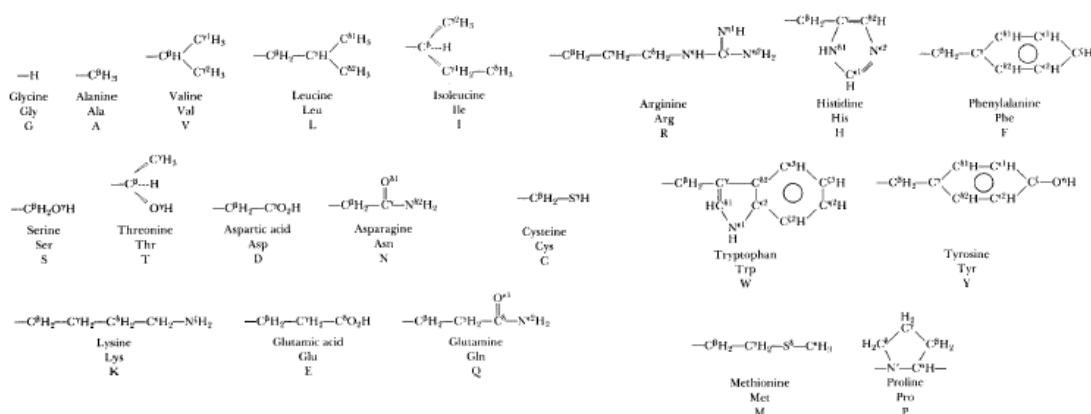


Fig. 2.6: Les chaînes latérales des 20 acides aminés naturels.

2.2 Attribution des chaînes latérales

Le repliement tertiaire d'une protéine étant obtenu essentiellement par des contraintes entre chaînes latérales, il convient de faire l'effort d'attribuer les protons des chaînes latérales. Les hydrogènes des chaînes latérales ne peuvent être tous traités de la même manière.

2.2.1 La chaîne aliphatique

L'attribution des chaînes aliphatiques sont accessibles par deux expériences tridimensionnelles fort utiles : la ^{15}N -TOCSY-HSQC et la ^{13}C -HCCH-TOCSY-HSQC.

La première corrèle par couplage scalaire isotrope dans les conditions de Hartmann-Hahn (Cavanagh et al. [1996]) chaque proton amide avec les hydrogènes de la chaîne latérale de chaque résidu. Le spectre TOCSY classique est éclaté dans une troisième dimension suivant le déplacement chimique de l'azote du squelette. Il convient éventuellement de faire varier le temps de mélange de 30ms à 100ms afin de faire apparaître tous les types d'hydrogène (H_α , H_β , H_γ , H_δ). Bien entendu, nous ne pourrions de cette manière atteindre les hydrogènes non couplés scalairement à ces derniers, à savoir les aromatiques, les S-méthyles de méthionine et les NH_2 des acides aminés glutamine et asparagine.

Disposant d'une protéine doublement marquée, nous pouvons mettre en oeuvre un autre type d'expérience : la ^{13}C -HCCH-TOCSY-HSQC. Le spin-lock est centré sur les carbones aliphatiques, et en suivant les conditions de Hartmann-Hahn, l'aimantation portée par un carbone va se transmettre sur toute la chaîne latérale. La première étape consiste à éditer les hydrogènes de la chaîne aliphatique, de transférer l'aimantation par INEPT vers les carbones aliphatiques, puis de répartir cette aimantation sur toute la chaîne latérale. Une INEPT retour fait revenir l'aimantation sur les hydrogènes et l'acquisition est effectuée. Le résultat est un spectre 3D contenant la même information qu'une $2\text{D}^{-1}\text{H}^{-1}\text{H}$ TOCSY éclatée suivant le déplacement chimique des carbones aliphatiques. Cette méthode présente l'avantage que le transfert HCCH-TOCSY est bien plus efficace qu'un transfert par les hydrogènes.

L'attribution des hydrogènes H_α est à confirmer par l'expérience HNHA décrite au paragraphe 3.2.2. Si cette séquence d'impulsions est utilisée pour déterminer la constante de

couplage ${}^3J_{H_N-H_\alpha}$, elle corrèle sans ambiguïté les protons amides et les protons H_α d'un même résidu.

2.2.2 Les groupes acide et amine des chaînes latérales de Arg, Asn, Asp, Gln et Glu.

Les protons amino des résidus asparagine et glutamine sont visibles en général lorsqu'ils ne sont pas en échange trop rapide avec l'eau. L'attribution de ces hydrogènes est particulièrement délicate pour une protéine simplement marquée ${}^{15}N$. En effet, ils ne sont couplés scalairement à aucun autre hydrogène et par conséquent, leur attribution ne peut se faire qu'à travers l'interprétation des spectres nOe. Ils donnent en effet des corrélations spatiales vers les hydrogènes intrarésiduels les plus proches : les H_β (Asn) et les H_γ (Gln). Si la protéine est doublement marquée ${}^{15}N$ - ${}^{13}C$, et que la chaîne latérale aliphatique est attribuée, on peut utiliser les couplages à travers les liaisons $C_{\text{ali}}-C'$ (expériences COCAH, paragraphe 2.1.2.2) et $N_{\text{ali}}-C'$ (HNCO optimisée pour les amines) pour obtenir une attribution non ambiguë.

L'attribution de l'arginine unique d'ASP2 fut aisée dans la mesure où le déplacement chimique de l'azote N_ϵ est caractéristique (entre 80 et 90ppm). Une carte HSQC avec une grande largeur spectrale suffit à attribuer l'hydrogène H_ϵ de l'arginine. Pour augmenter la résolution en dimension ${}^{15}N$, on diminue en général la largeur spectrale, le signal de la chaîne latérale de l'arginine apparaît alors replié dans le spectre. Les protons du groupe guanidinium sont en général invisible du fait de l'échange avec le solvant.

2.2.3 Les groupes aromatiques

Les cycles aromatiques des acides aminés Tryptophane, Histidine, Tyrosine et Phénylalanine sont particulièrement intéressants car ils sont souvent impliqués dans la formation d'un cœur hydrophobe. Ils représentent donc une source de contraintes particulièrement riche.

Afin d'établir l'attribution des hydrogènes des cycles aromatiques, j'ai utilisé un échantillon d'ASP2 non marqué dissout dans 100% de D_2O . Au bout de quelques heures, les protons amides sont échangés avec l'eau et seuls les protons aromatiques restent visibles à gauche de l'eau. A partir de la 2D-TOCSY, on reconnaît les hydrogènes couplés scalairement entre eux. La 2D-NOESY fait apparaître les corrélations entre les hydrogènes H_β des résidus aromatiques et les hydrogènes des cycles. Les hydrogènes H_δ et H_ϵ des tyrosines sont distinguables par l'intensité (respectivement forte et faible) de l'effet nOe vers les H_β . Bien qu'ambiguë, cette méthode reste suffisamment robuste lorsque le nombre de résidus aromatiques n'est pas trop important.

J'ai également modifié l'expérience 3D-NOESY-HSQC éditée ${}^{13}C$ pour éditer les déplacements chimiques des carbones des carbones aromatiques. Le délai autour des impulsions centrales de l'INEPT est alors de 1,3ms ($\langle {}^1J_{HC} \approx 195Hz \rangle$ sur les aromatiques). De cette séquence, j'ai obtenu l'attribution des carbones aromatiques.

Chapitre 3

Des données RMN à la résolution d'une structure

3.1 Déplacements chimiques et structures secondaires

De par sa grande dépendance vis-à-vis de l'environnement électronique, le déplacement chimique d'un noyau est lié à plusieurs facteurs structuraux : géométrie locale des liaisons, implication dans une liaison hydrogène, proximité de cycle aromatique, d'un noyau paramagnétique, ... Le déplacement chimique isotrope contient donc une quantité considérable d'informations sur l'environnement local de chaque spin et sur la structure de la protéine. Devant la difficulté rencontrée durant la récolte de contraintes de distances au fur et à mesure que la taille des protéines étudiées augmente, le déplacement chimique devient une contrainte qu'il faut savoir exploiter.

En l'absence de structure tridimensionnelle, il n'est pas encore possible de calculer *ab initio* le déplacement chimique de molécules aussi grosses que des protéines avec suffisamment de précision, on se contente donc d'établir des corrélations expérimentales entre déplacements chimiques et structure. Les premières recherches de corrélation structures secondaires-déplacements chimiques se sont portées sur les hydrogènes, seuls noyaux accessibles avant l'avènement du marquage isotopique. Du fait de la grande sensibilité du déplacement chimique des H^N à son implication dans une liaison hydrogène (Pardi et al. [1983]), il est difficile de séparer la contribution de la structuration secondaire de la contribution due à la liaison hydrogène. Par contre, Wishart et al. [1991] ont montré que le déplacement chimique moyen des H_α de résidus impliqués dans une structure secondaire s'éloignait des valeurs moyennes observées sur des acides aminés en pelote statistique (*random coil*). La déviation dépend du type de structure secondaire dans lesquelles les résidus sont impliqués : dans une hélice α , le déplacement chimique des H_α se trouve en moyenne décalé vers les hauts champs (-0,39 ppm par rapport au *random coil*), dans un brin β , il est décalé en moyenne vers les bas champs (+0,37 ppm par rapport au *random coil*). Afin d'automatiser et de rendre simple l'identification des structures secondaires d'une protéine à partir des déplacements chimiques des H_α , Wishart a développé un algorithme, appelé *Chemical Shift Index* (CSI) ou index de déplacement chimique. Le principe est d'associer à chaque résidu un nombre (-1, 0, +1) en fonction de l'écart de déplacement chimique de son H_α par rapport au déplacement chimique du même

Résidu	H_α	C_α	C_β	C'
Alanine	$4,35 \pm 0,10$	$52,5 \pm 0,7$	$19,0 \pm 0,7$	$177,1 \pm 0,5$
Cystéine (réduite)	$4,65 \pm 0,10$	$58,3 \pm 0,7$	$28,6 \pm 0,7$	$174,8 \pm 0,5$
Cystéine (oxydée)	$4,65 \pm 0,10$	$58,0 \pm 0,7$	$41,8 \pm 0,7$	$175,1 \pm 0,5$
Acide aspartique	$4,76 \pm 0,10$	$54,1 \pm 0,7$	$40,8 \pm 0,7$	$177,2 \pm 0,5$
Acide glutamique	$4,29 \pm 0,10$	$56,7 \pm 0,7$	$29,7 \pm 0,7$	$176,1 \pm 0,5$
Phénylalanine	$4,66 \pm 0,10$	$57,9 \pm 0,7$	$39,3 \pm 0,7$	$175,8 \pm 0,5$
Glycine	$3,97 \pm 0,10$	$45,0 \pm 0,7$	CSI de H_α ²	$173,6 \pm 0,5$
Histidine	$4,63 \pm 0,10$	$55,8 \pm 0,7$	$32,0 \pm 0,7$	$175,1 \pm 0,5$
Isoleucine	$3,95 \pm 0,10$	$62,6 \pm 0,7$	$37,5 \pm 0,7$	$176,9 \pm 0,5$
Lysine	$4,36 \pm 0,10$	$56,7 \pm 0,7$	$32,3 \pm 0,7$	$176,5 \pm 0,5$
Leucine	$4,17 \pm 0,10$	$55,7 \pm 0,7$	$41,9 \pm 0,7$	$177,1 \pm 0,5$
Méthionine	$4,52 \pm 0,10$	$56,6 \pm 0,7$	$32,8 \pm 0,7$	$175,8 \pm 0,5$
Asparagine	$4,75 \pm 0,10$	$53,6 \pm 0,7$	$39,0 \pm 0,7$	$175,1 \pm 0,5$
Proline	$4,44 \pm 0,10$	$62,9 \pm 4,0$	$31,7 \pm 0,7$	$176,0 \pm 4,0$
Glutamine	$4,37 \pm 0,10$	$56,2 \pm 0,7$	$30,1 \pm 0,7$	$176,3 \pm 0,5$
Arginine	$4,38 \pm 0,10$	$56,3 \pm 0,7$	$30,3 \pm 0,7$	$176,5 \pm 0,5$
Sérine	$4,50 \pm 0,10$	$58,3 \pm 0,7$	$62,7 \pm 0,7$	$173,7 \pm 0,5$
Thréonine	$4,35 \pm 0,10$	$63,1 \pm 0,7$	$68,1 \pm 0,7$	$175,2 \pm 0,5$
Valine	$3,95 \pm 0,10$	$63,0 \pm 0,7$	$31,7 \pm 0,7$	$177,1 \pm 0,5$
Tryptophane	$4,70 \pm 0,10$	$57,8 \pm 0,7$	$28,3 \pm 0,7$	$175,8 \pm 0,5$
Tyrosine	$4,60 \pm 0,10$	$58,6 \pm 0,7$	$38,7 \pm 0,7$	$175,7 \pm 0,5$

Tab. 3.1: Valeurs des déplacements chimiques *random coil* de H_α , C_α , C_β et C' et l'intervalle à prendre en compte pour la détermination des structures secondaires (Wishart et al. [1992], Wishart and Sykes [1994]). La référence utilisée est le TSP.

² : Les glycines n'ayant pas de C_β , le CSI du H_α est utilisé

résidu en *random coil* :

0 si le déplacement chimique observé est compris dans un intervalle autour du *random coil*,
+1 si le déplacement chimique observé est plus grand que la borne supérieure de l'intervalle,
-1 si le déplacement chimique observé est plus petit que la borne inférieure de l'intervalle.

La longueur de l'intervalle est fixée à 0,2 ppm ($\pm 0,1$ ppm) pour tous les résidus. Chaque index indique la propension d'un acide aminé à être structuré en hélice α (-1) ou en brin β (+1) ou autre (0).

Lorsque le déplacement chimique des carbones fut disponible, cette technique fut étendue aux C_α , C_β et C' (Wishart and Sykes [1994]), l'index étant défini de la même façon que pour les H_α . Il est à noter que le déplacement chimique du carbonyle est peu ou pas affecté par la liaison hydrogène et peut donc être inclus dans la prédiction des structures secondaires.

La procédure à suivre afin d'identifier les structures secondaires à partir des index du noyaux H_α est la suivante :

1. Une succession de 4 ou plus d'index "-1" non interrompue par un "+1" est une hélice

Noyau	H_α	C_α	C_β	C'
Hélice α	-1	+1		+1
Brin β	+1	-1	+1	-1

Tab. 3.2: Relation entre CSI et propension à former une structure secondaire. Un index +1 sur le C_β n'est pas corrélé à une propension à former une hélice α .

- α . Une succession de 3 ou plus de “+1” non interrompue par un “-1” est un brin β . Les autres régions sont en *random coil*.
2. Une densité locale supérieure à 70% est nécessaire pour définir une structuration secondaire. Cette densité locale doit être définie sur une fenêtre de 4 ou 5 résidus. Un minimum de 3 “+1” successifs est nécessaire pour en déduire un brin β et une hélice α ne peut être définie que par un minimum de 4 “-1” (non nécessairement consécutifs).
 3. Les extrémités des structures secondaires sont reconnues par un indice opposé ou par la succession d'indices nuls.

La procédure est identique pour les C_α et C' en inversant le signe des indices. Le cas des C_β est différent dans la mesure où l'indice associé à ce noyau ne peut être corrélé à l'identification des hélices α , la procédure précédente reste cependant valable pour l'identification des brins β , en utilisant le même signe d'indice que le H_α .

La valeur des déplacements chimiques *random coil* est tabulée dans le tableau 3.1. Le tableau 3.2 résume la propension des résidus à former une hélice α ou un brin β pour l'indice de chacun des noyaux.

Cet algorithme est implémenté dans le programme CSI⁴

3.2 Contraintes diédrales

3.2.1 Définition des angles dièdres

Les angles dièdres ϕ et ψ du résidu i du squelette sont définis respectivement par les angles de torsion autour des liaisons $C'_{i-1} - N_i - C_{\alpha i} - C'_i$ et $N_i - C_{\alpha i} - C'_i - N_{i+1}$ (voir figure 3.1). Le troisième angle ω décrivant la liaison peptidique est en général plan (cis ou trans). La détermination des angles dièdres fournit donc des informations extrêmement précieuses sur la géométrie du squelette d'une protéine. La connaissance avec une grande précision des angles dièdres du squelette et des chaînes latérales est théoriquement susceptible de réaliser le repliement de la chaîne peptidique.

Les paramètres géométriques angulaires d'une molécule sont d'une manière générale accessibles en RMN par le biais des constantes de couplage 3J entre deux noyaux. L'équation semi-empirique de Karplus (Karplus [1959]) relie les constantes de couplage 3J entre deux noyaux quelconques séparés de trois liaisons formant un angle dièdre θ :

$$^3J = A \cos^2 \theta + B \cos \theta + C \quad (3.1)$$

Dans le cas des angles dièdres du squelette d'une protéine, Pardi (Pardi, Billeter, and Wüthrich [1984]) a précisé la valeur des constantes A, B et C pour le couplage $^3J_{HNH_\alpha}$ et

⁴disponible à l'adresse : <http://www.pence.ualberta.ca/software/csi/latest/csi.html>

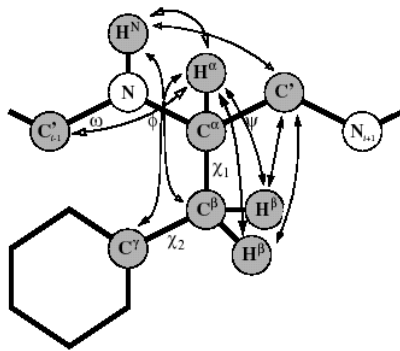


Fig. 3.1: Définition des angles dièdres du squelette peptidique (ϕ, ψ et ω) et de la chaîne latérale (χ_1, χ_2).

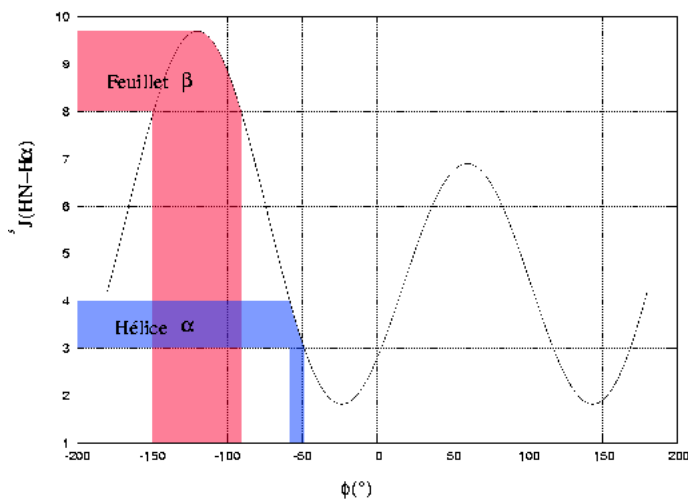


Fig. 3.2: Courbe de Karplus reliant la constante de couplage $^3J_{H_N H_\alpha}$ à l'angle dièdre ϕ . Les intervalles pour la constante de couplage des hélices α droites et des feuillets β canoniques sont indiquées en fonction des valeurs angulaires permises.

l'angle dièdre ϕ :

$$^3J(\theta) = 6,4 \cos^2 \theta - 1,4 \cos \theta + 1,9 \quad \text{avec } \theta = |\phi - 60^\circ| \quad (3.2)$$

3.2.2 Détermination expérimentale des angles ϕ : HNHA

L'expérience 3D-HNHA (Vuister and Bax [1993], Vuister and Bax [1994]) permet d'obtenir d'une manière quantitative des plages de valeurs pour les angles ϕ du squelette. Le coeur de cette séquence (figure 3.3) réside dans les deux modules HMQC où l'aimantation H_y est transférée sur l'azote (entre a et b), le déplacement chimique de l'azote édité entre b et c, et le couplage $^3J_{(H_N H_\alpha)}$ s'expriment. Le couplage $J_{H_N N}$ ne s'exprime pas entre b et c car il n'a aucun effet sur les termes à 0 et 2 quanta. Les déplacements chimiques des protons H_N et H_α sont édités pendant t_2 . Le deuxième module HMQC entre d et f permet de revenir sur le proton amide pour la détection. La séquence originelle a été modifiée par l'ajout d'une Watergate (Piotto et al. [1992]) avant l'acquisition pour améliorer la suppression du solvant. La dernière impulsion $90^\circ_{\Phi_y}$ sélectionne les états qui vont donner un signal purement en absorption mais en contrepartie, elle mène à un rapport Signal/Bruit plus faible. Le délai δ est ajusté à $1/(2J_{H_N N})$ pour optimiser le transfert $H_y^N \rightarrow H_x^N N_z$. Le délai ζ est fixé à $13,05ms$ pour

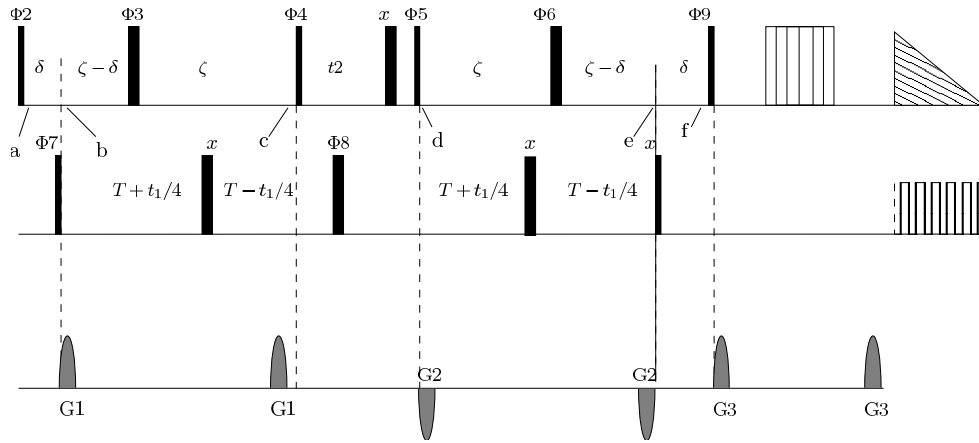


Fig. 3.3: Séquence d'impulsion de la 3D-HNHA. Les rectangles étroits et larges sont respectivement des impulsions 90° et 180° . La suppression de l'eau est assurée par une Watergate. Le cyclage de phase sur 16 phases est le suivant : $\Phi_2 = x$; $\Phi_3 = 4x, -4x$; $\Phi_4 = x, -x$; $\Phi_5 = 2x, -2x$; $\Phi_6 = 4x, -4x$; $\Phi_7 = x$; $\Phi_8 = 8x, 8y$; $\Phi_9 = y, -y$; $\Phi_{\text{récepteur}} = 8x, -8x$. Les délais ζ et δ valent respectivement 13,05ms et 4,5ms. Les gradients sont successivement d'intensité 2%, -2% et 15%.

optimiser le transfert $H_y^N \rightarrow H_x^N H_z^\alpha$. Il est légèrement inférieur à $1/(2J_{H^N H_\alpha})$ pour augmenter la sensibilité de l'expérience (relaxation T_2). Les deux dimensions ont été accumulées en mode States-TPPI (Marion et al. [1989]).

Le signal détecté pendant l'acquisition est décrit par :

$$S(t_1, t_2, t_3) = A \cos(\omega_N t_1) e^{i\omega_{H^N} t_3} \{ \cos^2(2\pi J_{H^N H_\alpha} \zeta) \cos(\omega_{H^N} t_2) - \sin^2(2\pi J_{H^N H_\alpha} \zeta) \cos(\omega_{H_\alpha} t_2) \} \quad (3.3)$$

Après transformée de Fourier, aux déplacements chimiques de l'azote et du proton amide du résidu i , il y aura deux pics : un pic sur la diagonale correspondant à l'aimantation restée sur H_N pendant les modules HMQC et un pic hors-diagonale au déplacement chimique du H_α du résidu i . Le rapport d'intensité des deux pics s'exprime alors par :

$$\frac{I_{\text{hors-diagonale}}}{I_{\text{diagonale}}} = -\tan^2(2\pi J_{H^N H_\alpha} \zeta) \quad (3.4)$$

L'estimation de $J_{H^N H_\alpha}$ s'obtient par inversion de cette relation. L'expression 3.3 n'est pas rigoureusement exacte dans la mesure où les états $2I_x^{H^N} I_z^{H_\alpha}$ et $I_y^{H^N}$ donnant lieu aux deux termes de la somme évoluent avec un temps de relaxation différent pendant le temps 2ζ . Dans la limite des mouvements lents pour les protéines, l'état antiphase $2I_x^{H^N} I_z^{H_\alpha}$ relaxe plus rapidement que la composante en phase $I_y^{H^N}$, entraînant une diminution de l'intensité du pic hors-diagonale par rapport au pic de la diagonale. Vuister et Bax ont proposé d'appliquer un facteur correctif global de 1,11 à l'ensemble des couplages mesurés. Cette correction appliquée à la nucléase du staphylocoque (SNase) est justifiée pour les résidus présentant peu de mobilité interne.

Le cas des glycines est particulier car ces résidus portent deux protons H_α qui vont chacun participer au déphasage de H_x^N pendant le délai 2ζ . L'interaction dipolaire forte entre les protons géminaux va induire une diminution de l'intensité de la tache la plus intense et une

feuillet β		hélice α		structure étendue	
$^3J_{H_N H_\alpha}$ (en Hz)	ϕ (en degré)	$^3J_{H_N H_\alpha}$ (en Hz)	ϕ (en degré)	$^3J_{H_N H_\alpha}$ (en Hz)	ϕ (en degré)
> 9,5	-120±25	> 6,4	-80±25	> 7,0	-120±55
entre 8,0 et 9,5	-120±35	entre 5,5 et 6,5	-70±25	< 7,0	indétermination
entre 7,0 et 8,0	-120±55	< 5,5	-60±25		
< 7,0	-115±60				

Tab. 3.3: Plage de valeurs de Φ en fonction de la structure secondaire et de la valeur de J

augmentation de l'intensité de la tache la plus faible. En l'absence d'attribution stéréospécifique, il n'est pas possible de conclure sur la valeur de ϕ .

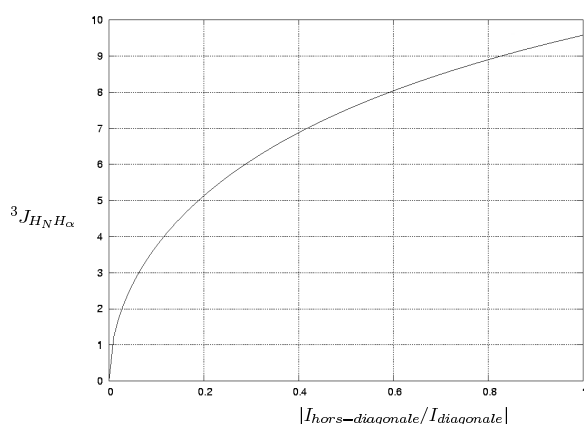


Fig. 3.4: Evolution de $^3J_{H_N H_\alpha}$ en fonction du rapport $I_{hors-diag}/I_{diag}$.

3.2.3 Prédiction des angles ϕ et ψ du squelette : Talos

Cornilescu (Cornilescu et al. [1999a]) a prolongé l'étude de Wishart sur la relation entre déplacements chimiques et structure secondaire d'une protéine. Il a mis au point un algorithme de prédiction des angles de torsion utilisant la séquence et les déplacements chimiques H_α , N , C_α , C_β et C' de la protéine. Le principe sous-jacent à ce travail est que si une succession d'acides aminés montre une similarité de déplacements chimiques avec une séquence d'une protéine de structure connue, alors les résidus centraux des deux chaînes ont probablement les mêmes angles de torsion ϕ et ψ . De plus, si les deux chaînes ont une grande similarité de séquence, alors la méthode devient encore plus robuste. Bien que le déplacement chimique de l'azote soit influencé par les liaisons hydrogènes, la mise au point de la procédure a montré qu'il pouvait être pris en compte mais avec un poids relativement faible.

Le logiciel Talos⁵ automatise la prédiction des angles de torsion en acceptant en entrée la séquence en acides aminés et le déplacement chimique des différents noyaux. Le logiciel dispose d'une base de données de 20 structures à haute résolution et de leur attribution. Il évalue ensuite la similarité S_{ij} entre le triplet (3 acides aminés successifs) centré sur le résidu i de la séquence entrée et tous les triplets j de la base de données. Ce facteur est calculé sur la base de la différence de déplacements chimiques des $3 \times 5 = 15$ noyaux et de la comparaison de

⁵disponible à l'adresse suivante : <http://spin.niddk.nih.gov/bax/software/TALOS/>

	β, β_p	α -Helix	3_{10} -Helix	Turn I	Turn II	Turn I'	Turn II'	Half-Turn
$d_{\alpha N}(i, i+4)$		=====						
$d_{\alpha\beta}(i, i+3)$		=====	=====					
$d_{\alpha N}(i, i+3)$		=====	=====	=====		=====	=====	
$d_{NN}(i, i+2)$		=====	=====	=====	=====	=====	=====	
$d_{\alpha N}(i, i+2)$			=====	=====	=====	=====	=====	=====
d_{NN}	=====	=====	=====	=====	=====	=====	=====	=====
$d_{\alpha N}$	=====				=====			=====
$^3J_{H\alpha}$ (Hz)	9 9 9 9 9 9 1 2 3 4 5 6	4 4 4 4 4 4 4 1 2 3 4 5 6 7	4 4 4 4 4 4 4 1 2 3 4 5 6	4 9 1 2 3 4	4 5 1 2 3 4	7 5 1 2 3 4	7 9 1 2 3 4	4 9 1 2 3 4

Fig. 3.5: Structure secondaire et nOe caractéristiques (Wüthrich [1986])

la composition en acides aminés. Les coefficients marquant le poids des $3 \times 6 = 18$ contributions furent déterminés heuristiquement. Talos rend en sortie pour chaque résidu, les triplets de la base de données présentant un facteur de similarité inférieur à un seuil fixé pour avoir au minimum 10 solutions. Chaque résidu (triplet) de la protéine analysée est donc finalement comparé à une dizaine de résidus (triplets) issus de la base de données.

Les angles de torsion des résidus centraux des triplets de plus haute similarité sont ensuite calculés. Si tous les angles (ϕ , ψ) sont situés dans une même zone (régions les plus favorables) de la carte de Ramachandran, Talos propose donc un encadrement du couple (ϕ , ψ) du résidu i autour de la moyenne des points des 10 triplets avec un indice de confiance élevé. Si un seul couple (ϕ , ψ) sort de la région définie par les 9 autres couples, celui-ci est donc éliminé, et la moyenne est calculée sur les 9 résidus. Si la déviation-standard est supérieure à 45° , la prédiction est notée comme ambiguë. Dans certains cas, la distribution des angles (ϕ , ψ) est trop éclatée, et aucune prédiction ne peut être effectuée.

L'intérêt de cette méthode est qu'elle utilise en elle-même la même information que la méthode du CSI (identification des hélices α et des brins β) tout en ajoutant le paramètre important que représente la séquence. Elle est plus puissante dans la mesure où le résultat est un encadrement assez fin des angles dièdres et où l'identification des hélices gauches est désormais possible.

3.3 Les nOes caractéristiques des structures secondaires.

Jusqu'à présent, l'observation de l'effet nOe entre deux noyaux d'hydrogènes représente la source principale de contraintes pour le repliement tertiaire des protéines. Bien entendu, la structure locale est maintenant accessible par des méthodes plus efficaces en terme de temps d'analyse (CSI, Talos) mais la présence de distances caractéristiques entre certains protons est une confirmation absolue de la prédiction. Les motifs de structure secondaire donnent lieu aux distances tabulées dans la figure 3.5. Dans le cas des structures en feuillet β , les proximités

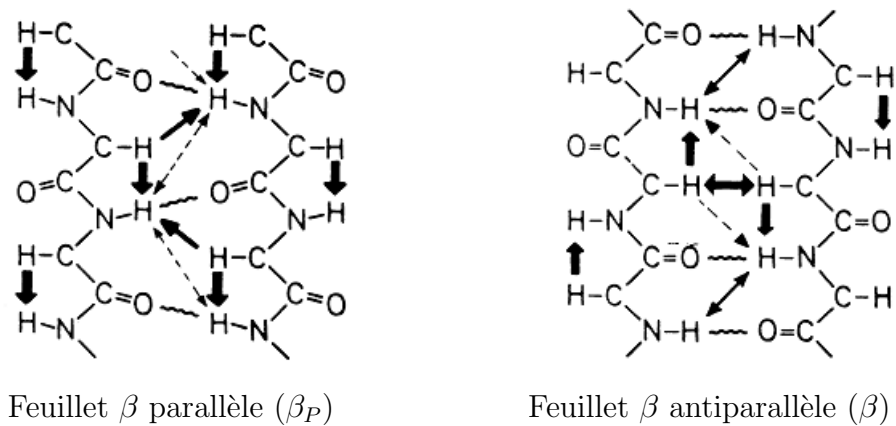


Fig. 3.6: Distances caractéristiques au sein des feuillets parallèle et antiparallèle

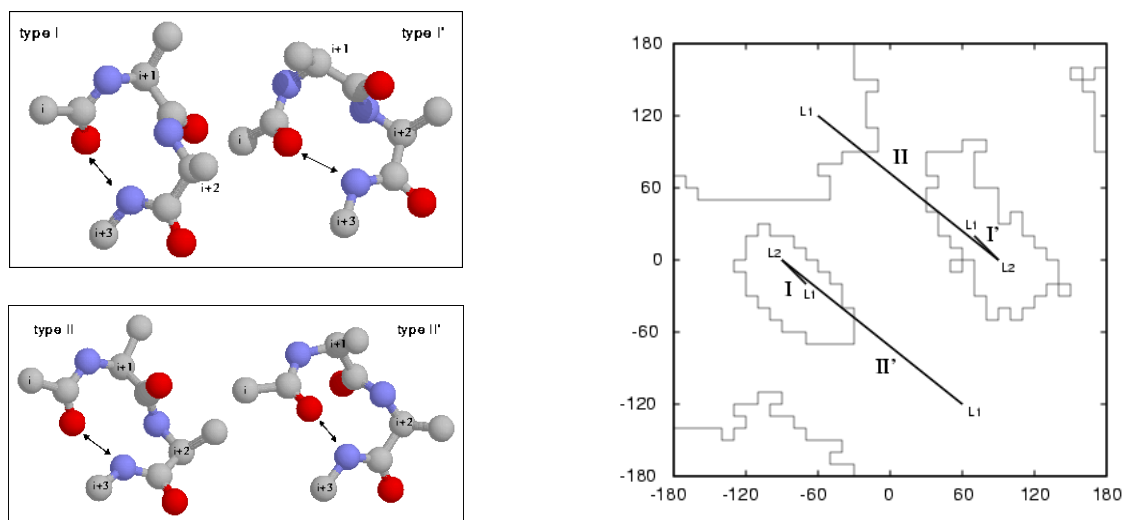


Fig. 3.7: Conformation du squelette peptidique en coude β et position dans la carte de Ramachandran des résidus L1 et L2 en fonction du type de coude.

d'hydrogènes entre les brins donnent l'orientation relative des brins (parallèle ou antiparallèle, voir figure 3.6).

Le troisième type de structure secondaire rencontré dans les protéines est le coude β (β -turn). Il permet à la chaîne peptidique de changer d'orientation. Il existe jusqu'à neuf familles de coudes β mais les coudes de type I et II sont les plus fréquemment rencontrés (figure 3.7). Chaque classe contient deux conformations, images l'une de l'autre dans un miroir (I et I', II et II'). Dans le cas d'un arrangement de deux brins β antiparallèles successifs dans la séquence primaire (β -hairpin), le raccord entre les deux brins se fait par l'intermédiaire d'un motif en coude β . On rencontre le plus souvent les coudes de type I' et II' dans les β -hairpin.

Le type de coude peut être reconnu aussi bien par les distances caractéristiques que par les angles dièdres de deux résidus centraux (appelés L1 et L2). Pour des raisons d'encombrement stérique, le résidu en L2 du coude de types I' et II' est généralement une glycine ainsi que le résidu L1 du coude de type II'⁶.

3.4 Les liaisons hydrogènes

3.4.1 Echange H/D

Un proton amide est *a priori* susceptible de s'échanger avec les protons du solvant. La solvatation d'une protéine non structurée dans D₂O se traduit donc par un décalage de la fréquence de résonance des hydrogènes amides (fréquences du spectromètre) vers la gamme de résonance du deutérium, tandis que les hydrogènes liés à un carbone résonnent toujours à la fréquence du proton. Dans le cas d'une protéine structurée, certains hydrogènes amides restent visibles car la vitesse d'échange proton-deutérium est très faible (de l'ordre de la durée d'une expérience de RMN). La vitesse d'échange peut être estimée sur une protéine fraîchement dissoute dans D₂O par l'accumulation d'une série de spectres (TOCSY, HSQC si la protéine est marquée ¹⁵N). Les protons s'échangeant rapidement avec le solvant disparaissent immédiatement, tandis que les protons peu labiles disparaissent progressivement.

On attribue en général cette protection vis-à-vis de l'échange H/D à l'implication de l'hydrogène dans une liaison hydrogène (Cavanagh et al. [1996], p548). Cependant, le partenaire de l'hydrogène amide n'est pas identifié par cette technique. Il peut être déterminé soit par la connaissance par un autre moyen de la structure secondaire (dans une hélice α , le partenaire de H_N^i est le carbonyle C^{i-4}), soit lorsque la convergence de la structure est suffisante pour éliminer les partenaires potentiels. La contrainte utilisée lors de la modélisation (par Aria, voir paragraphe 3.5) pour fixer une liaison hydrogène est la suivante :

$$\begin{aligned} 1, 3\text{\AA} < d_{HN^i-O^{i-4}} < 2, 5\text{\AA} \\ 2, 2\text{\AA} < d_{N^i-O^{i-4}} < 3, 4\text{\AA} \end{aligned} \quad (3.5)$$

3.4.2 Mise en évidence directe : la HNCO longue distance

Récemment, il a été mis en évidence à la fois sur les acides nucléiques (Dingley and Grzesiek [1998]) et les protéines (Cordier and Grzesiek [1999], Cornilescu et al. [1999b]) que les

⁶voir <http://www.cryst.bbk.ac.uk/PPS2/course/section3/hairpin.html>

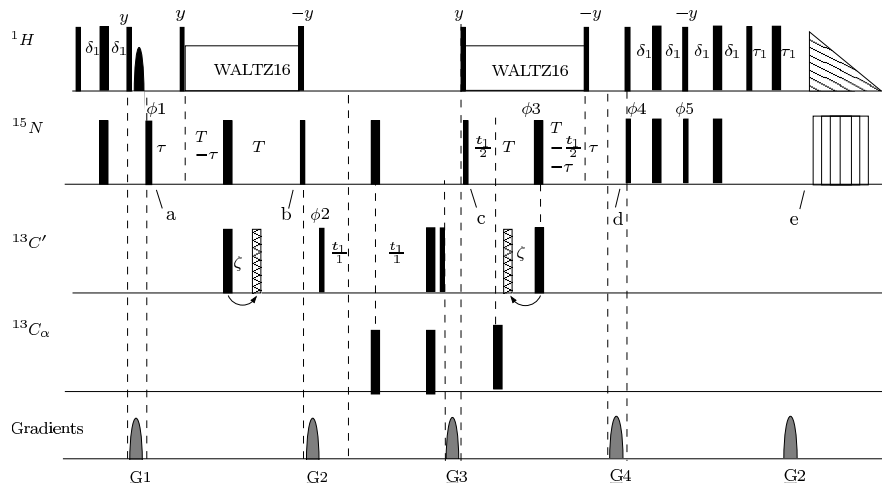


Fig. 3.8: Séquence d'impulsions HNC O longue distance (HNC O hb et HNC O hb_réf). Les rectangles larges (resp. étroits) représentent des impulsions 180° (resp. 90°). La quatrième impulsion proton est un 90° sélectif sur l'eau. Toutes les impulsions sur les canaux carbones sont sélectives (profil gaussien). Par défaut, les impulsions sont d'axe x . Le cyclage de phases est le suivant : $\phi_1 = 8x, 8(-x)$; $\phi_2 = x, (-x)$; $\phi_3 = 4x, 4(-x)$; $\phi_4 = 2x, 2(-x)$; $\phi_5 = 2(-y), 2y$, $\phi_{acq} = x, 2(-x), x, x, 2(-x), x, (-x), 2x, (-x), (-x), 2x, (-x)$. Les délais sont fixés à : $T = 66,6ms$, $\delta_1 = 2,3ms$, $\zeta = 16,6ms$, $\tau = 5,5ms$. La sélection des cohérences est effectuée par la méthode echo-antiecho sur les gradients 3 et 4. Pour la mesure de l'intensité de référence, deux impulsions carbonyles sont décalées de δ_2 .

constantes de couplage scalaire à travers une liaison hydrogène existaient et étaient mesurables par RMN sur les macromolécules (pour revue, Grzesiek et al. [2001]). La valeur des couplages scalaires à travers une liaison hydrogène d'une protéine est très faible (entre $-0,6$ et $1,4$ Hz entre un proton amide et un carbonyle ou un carboxyle) et est corrélée avec la longueur de la liaison hydrogène (Cornilescu et al. [1999c]).

La séquence d'impulsions qui permet de mettre en évidence une liaison hydrogène entre un proton amide (ou amino) et un carboxyle ou un carbonyle est basée sur l'expérience HNC O (Grzesiek and Bax [1992b], Kay et al. [1994]) utilisée pour l'attribution du squelette. La différence réside au niveau du délai d'évolution sous le couplage NC' (expérience HNC O hb, figure 3.8). Le couplage ${}^3hJ_{NC'}$ à la travers la liaison hydrogène est en effet beaucoup plus faible que le couplage ${}^1J_{NC'}$ (entre 13 et 17Hz). Après le transfert INEPT, l'aimantation est $H_z N_y$ en a. Entre a et b, l'aimantation est transférée vers le carbonyle tout en refocalisant le proton ($N_y C'_z$). Le déplacement chimique du carbonyle est ensuite édité. Entre c et d, un temps constant permet l'édition du déplacement chimique de l'azote, la défocalisation sous les couplages NH et NC' pour arriver en d à la cohérence $N_y H_z$. Le retour sur les protons se fait en utilisant le schéma PEP (Palmer III et al. [1991a]). Le délai T est optimisé à $66,6ms$ pour refocaliser le couplage ${}^1J_{NC'}$ et pour laisser le couplage ${}^3hJ_{NC'}$ s'exprimer. En principe, le spectre ne contient que les corrélations à travers les liaisons hydrogènes, cependant certains couplages intrarésiduels et séquentiels peuvent ne pas être suffisamment refocalisés et mènent alors à des taches supplémentaires.

En modifiant la séquence par un décalage des impulsions d'inversion du carbonyle durant

l'INEPT et l'INEPT retour, on peut quantifier la valeur de ce couplage. La durée de la séquence est identique mais le temps effectif de défocalisation et de refocalisation du couplage NC' est réduit à $2(T - \zeta)$. Le spectre HNCOb_{réf} fera alors apparaître préférentiellement les taches de corrélation $N^i - C'^{i-1}$. Le rapport d'intensité entre la tache de corrélation $N^i - C'^j$ (I_{hb}) et la tache de corrélation $N^i C'^{i-1}$ (I_{ref}) est lié à la valeur du couplage à travers la liaison hydrogène :

$$J_{N^i C'^j} = \frac{1}{2\pi T} \sqrt{\frac{I_{hb}}{I_{ref}}} \quad (3.6)$$

Cette expression fait l'approximation que $|2\pi J_{N^i C'^j} T| \ll 1$ (soit $J \ll 2,4Hz$) et $J_{N^i C'^{i-1}} \approx 15Hz$.

Une version TROSY de l'expérience HNCOb longue distance existe (Wang et al. [1999]). Nous ne la détaillerons pas.

3.5 Vers une reconstruction automatique : Aria

3.5.1 Reconstruction semi-manuelle d'une protéine

Les logiciels de modélisation moléculaire du type de CNS⁷ ou Xplor⁸, utilisent des contraintes de distances dérivées de l'intensité de la tache nOe entre deux hydrogènes. En considérant la dépendance de l'effet nOe avec la distance (Noggle and Shirmer [1971]), la distance peut être obtenue à partir de la relation suivante (Wüthrich [1986]) :

$$d_{ij} \propto NOE^{\frac{1}{6}} \quad (3.7)$$

où d_{ij} désigne la distance entre les noyaux i et j et NOE désigne l'intensité de la tache de corrélation.

Cette approche nécessite une calibration correcte des spectres à partir de certaines distances caractéristiques (protons géminaux, distance $HN^{i+3} - H_{\alpha}^i$ des hélices). Pour tenir compte des effets de diffusion de spin et de dynamique interne de la molécule, les contraintes sont définies sur une gamme de distance à partir de l'intensité (fort, moyen, faible) des taches de corrélations. De plus, une correction de distance est à apporter lorsqu'au moins un des partenaires est un méthyle car chacun des trois hydrogènes peut contribuer à l'effet nOe et la contrainte utilise le barycentre des hydrogènes des méthyles impliqués.

Lorsqu'un nombre suffisant de contraintes a été obtenu à partir des spectres, le travail de reconstruction peut commencer en utilisant les logiciels de modélisation. Les contraintes de distances sont introduites en général dans la fonction cible à minimiser par un terme de pseudo-énergie potentielle de forme harmonique :

$$E_{NOE} = k(d_{calc} - d_{exp})^2 \quad (3.8)$$

Les contraintes introduites sont respectées dès lors que ce terme d'énergie est nul. Dans la procédure de reconstruction, les contraintes sont affinées manuellement au fur et à mesure par

⁷disponible à l'adresse : <http://cns.csb.yale.edu/v1.1/>

⁸disponible à l'adresse : <http://xplor.csb.yale.edu/xplor/>

essai-erreur (figure 3.9). L'objectif final est d'arriver à une liste de distances compatibles et satisfaisant le terme géométrique du champ de force. Malheureusement, l'encombrement des spectres mène souvent à plusieurs attributions possibles pour certaines taches. L'étape d'aller-retour entre les calculs et les spectres est donc primordiale et peut être une source d'erreur catastrophique.

3.5.2 La philosophie d'Aria

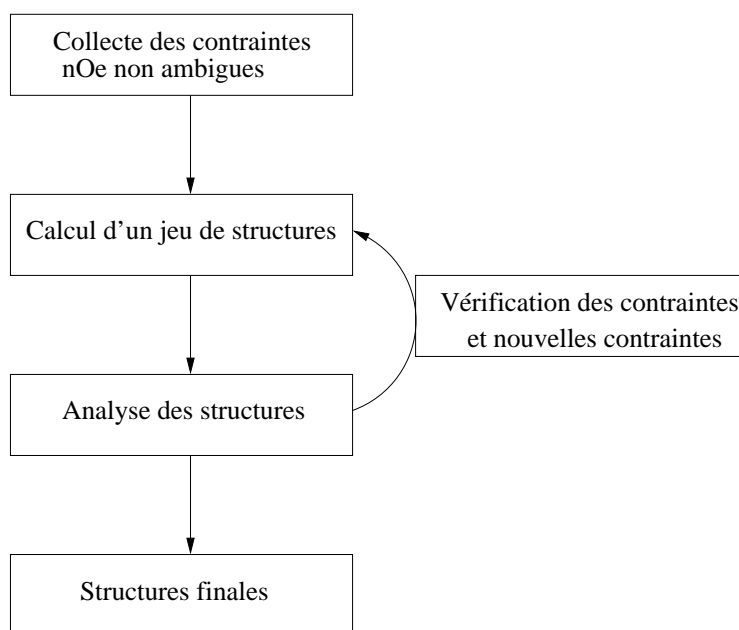


Fig. 3.9: Algorithme de reconstruction manuelle d'une macromolécule.

Durant mon travail sur ASP2, j'ai commencé par suivre la méthode précédente. En cours de reconstruction, je me suis rendu compte que certaines taches étaient *a priori* ambiguës et étaient capitales pour le repliement de la protéine. D'autre part, plusieurs équipes tentent de développer des logiciels d'attribution automatique des nOes dans le but d'accélérer cette étape. J'ai donc changé de stratégie en utilisant un algorithme d'attribution automatique des nOes ambiguës : ARIA.

Devant le nombre croissant de données structurales (nOe, couplages dipolaires résiduels, liaisons hydrogènes, centre paramagnétique, ...) que la RMN peut fournir au spectroscopiste, il devient indispensable de développer des outils informatiques permettant de réduire au maximum l'intervention humaine. En particulier, lors de la reconstruction d'une protéine de taille importante, les risques d'erreur d'attribution des taches de corrélation nOe sont grands.

Michael Nilges et ses collègues ont commencé à coder un algorithme de reconstruction de macromolécules proche de celui suivi par l'homme (figure 3.9). Ils ont introduit le concept de *contraintes de distance ambiguë* (ADR, *ambiguous distance restraint*) qui prend en compte toutes les attributions possibles pour une tache de corrélation. Ainsi, la contrainte de distance

\bar{D} introduite dans la modélisation devient :

$$\bar{D} \equiv \left(\sum_{a=1}^{N_\delta} d_a^{-6} \right)^{-1/6} \equiv NOE_{\delta_1, \delta_2}^{-\frac{1}{6}} \quad (3.9)$$

NOE_{δ_1, δ_2} représente l'intensité de la tache observée aux déplacements chimiques δ_1 et δ_2 . N_δ représente le nombre de paires d'hydrogènes aux déplacements chimiques δ_1 - δ_2 dans les deux dimensions suivant une tolérance définie par l'opérateur. d_a représente la distance dans une structure entre les hydrogènes d'une des N_δ paires.

Le logiciel Aria a été écrit dans le but de réduire de façon itérative le nombre de paires contribuant à chaque nOe observé, en éliminant automatiquement les paires ne contribuant pas physiquement à l'intensité. Le progrès apporté par Aria réside dans le choix des paramètres (tolérance sur les déplacements chimiques, poids affecté à chaque paires d'un ADR, . . .). L'utilisation de contraintes de distance ambiguë permet de s'affranchir des corrections de pseudo-atomes car les protons équivalents (méthyles, protons aromatiques) sont traités comme des nOe ambigus. L'utilisation du calcul de la matrice de relaxation assure en outre une calibration globale de chaque spectre en prenant en compte la diffusion de spin. Enfin, l'attribution stéréospécifique est laissée flottante durant la procédure par échange des déplacements chimiques des hydrogènes géminaux ou des méthyles de valine ou d'isoleucine. La sélection se fait sur la base du minimum d'énergie potentielle nOe.

3.5.3 La méthode d'attribution d'Aria

3.5.3.1 Schéma général

Aria utilise le schéma suivant pour attribuer les nOe observés à partir de l'attribution des résonances (non ambiguës) des protons de la protéine :

1. Pour chaque pic, les N_δ attributions possibles (basées sur les déplacements chimiques avec une tolérance δ_{ppm}) sont prises en compte.
2. Les 7 structures de plus basses énergies de l'itération i-1 sont sélectionnées.
3. Les distances caractéristiques \hat{d}_a moyennes de chaque paires de protons sont calculées sur ces 7 structures.
4. Les volumes nOe sont convertis en contraintes de distance en utilisant une calibration automatique des spectres.
5. Un critère de sélection est appliqué sur les différentes contributions. Les contributions les plus faibles sont éliminées.
6. Un ensemble de 20 structures est calculé à partir des nouvelles contraintes.

Le premier jeu de structures (itération 0) est évalué à partir du seul critère de compatibilité des déplacements chimiques ou à partir d'un modèle (issu d'un run précédent par exemple). Le schéma précédent est appliqué jusqu'à l'itération 8. Un nombre supérieur d'itérations ne semble pas apporter d'amélioration ni de l'attribution, ni de la convergence des calculs.

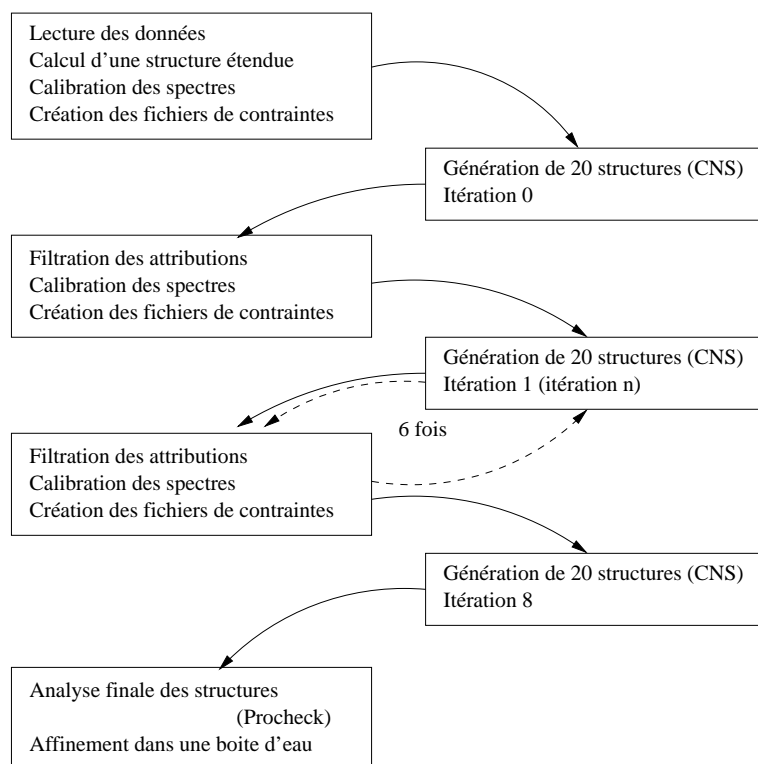


Fig. 3.10: Algorithme de reconstruction d'une macromolécule par Aria

3.5.3.2 Filtre d'attribution

A chaque itération, un pic peut correspondre à la somme des contributions de plusieurs paires de protons. Le volume total du pic sera la somme des contributions individuelles. Aria estime cette contribution C_n par l'expression suivante (\hat{d}_n représente la distance moyenne de la paire n dans les 7 structures de plus basses énergies de l'itération précédente) :

$$C_n = \frac{\hat{d}_n^{-6}}{\sum_{a=1}^{N_\delta} \hat{d}_a^{-6}} \quad (3.10)$$

Les paires sont ensuite réordonnées par contribution décroissante. Les attributions de plus faible probabilité sont éliminées de telle manière que :

$$\sum_{a=1}^{N_p} C_a > p \quad (3.11)$$

Le paramètre p détermine les N_p paires contribuant majoritairement au pic qui seront conservées par la suite. Ce paramètre vaut presque 1 dans la première itération et décroît jusqu'à 0,8 lors de la dernière itération. Plus la valeur finale de p est grande, plus l'attribution finale du pic restera ambiguë. La valeur de p décroît au fur et à mesure des itérations pour ne garder que les contributions qui ont physiquement un sens.

A chaque itération, deux fichiers de contraintes de distances sont générés. Lors de la première itération, le fichier `ambig.tbl` contient toutes les contraintes ADR définies sur la seule base des déplacements chimiques. Dans le fichier `unambig.tbl`, on trouve alors les contraintes entre hydrogènes dont les déplacements chimiques sont suffisamment caractéristiques pour qu'aucune ambiguïté sur l'attribution ne subsiste. Durant la première itération, la plupart des contraintes sont donc dans le fichier `ambig.tbl`. Au fur et à mesure des itérations, l'attribution des contraintes se précise par l'ajout d'informations sur les structures calculées et par le paramètre p . Les contraintes passent alors progressivement dans le fichier `unambig.tbl`.

3.5.3.3 Calibration des spectres

Dans l'approche classique, la simulation des spectres à partir d'un jeu de S structures utilise la relation $NOE = \frac{1}{C} d_{obs}^{-6}$ où d_{obs} peut être définie comme la moyenne arithmétique $d_{obs} = \frac{1}{S} \sum_{s=1}^S d_{ij,s}$ ou la moyenne $d_{obs} = \left(\sum_{s=1}^S d_{ij,s}^{-6} \right)^{-\frac{1}{6}}$.

En supposant que la dynamique est relativement homogène sur l'ensemble de la structure, un facteur de calibration globale C est obtenu par la relation :

$$C = \sum_{NOEs} \frac{d_{obs}^{-6}}{NOE} \quad (3.12)$$

Dans Aria, les spectres sont calibrés à partir d'un jeu de structures en simulant une carte nOe A . Pour éviter les biais introduits par la diffusion de spin, cette simulation ne repose pas sur les distances \hat{d}_a comme précédemment, mais prend en compte la matrice de relaxation R (τ_m est le temps de mélange nOe) avec un temps de corrélation de la protéine de τ_c :

$$R_{ij} = n_i \left(\frac{1}{\hat{d}_{ij}} \right)^6 \frac{\pi}{5} \gamma^4 \hbar^2 \left(\frac{6}{1 + 4\omega^2 \tau_c^2} - 1 \right) \quad (3.13)$$

$$A = e^{-R\tau_m} \quad (3.14)$$

L'approche par la matrice de relaxation ne donne pas les limites supérieures (comme dans l'approche classique) mais fournit directement une estimation de la distance. Les limites supérieures et inférieures sont données en estimant l'incertitude sur la distance à 12,5%.

$$\begin{cases} \text{limite inférieure} & = d_{obs} - \Delta^- \\ \text{limite supérieure} & = d_{obs} + \Delta^+ \\ \Delta^+ = \Delta^- = 0.125 d_{obs}^3 \end{cases} \quad (3.15)$$

Entre chaque itération, les spectres sont recalibrés après élimination des attributions les moins probables.

3.5.3.4 Traitement du bruit

Il est souvent difficile de distinguer dans les spectres les pics d'intensité faible et le bruit. La liste de pics d'une expérience contient donc en général des pics d'intensité très faible correspondant soit à un transfert entre protons éloignés, soit au bruit de l'expérience.

Par l'approche des contraintes ambiguës, un pic de bruit est interprété comme un signal physique et s'il existe une attribution possible à ce pic, les structures générées peuvent présenter des déformations très importantes. En effet, du fait de la forme harmonique de la fonction cible, l'introduction dans le calcul de contraintes fausses va mettre une pénalité trop importante sur ces violations et va mener à des violations de vraies contraintes. Il est donc primordial d'éliminer le bruit des spectres. Souvent, les contraintes issues du bruit sont violées dans les structures car celles-ci ne sont pas nécessairement compatibles entre elles.

Le critère utilisé par Aria pour éliminer le bruit de la liste de pics repose sur le nombre de violations de contraintes sur un jeu de structures. Si telle contrainte est violée dans plus de 50% des 7 meilleures structures d'une itération, cette contrainte peut être soit simplement notée, soit éliminée, soit ses limites sont élargies.

L'inconvénient de cette approche est que l'élimination d'une contrainte est sans effet sur l'ensemble du calcul. Aucun coût n'est imposé si bien que si une zone de la structure ne fait apparaître que peu de nOe, une seule tache n'est pas suffisante pour positionner correctement cette région. En revanche, lorsque l'attribution de la protéine n'est que partielle, l'élimination des taches provenant de la région non attribuée va éviter de trop contraindre la région déjà attribuée.

3.5.3.5 Choix des paramètres

L'utilisateur d'Aria peut régler à volonté tous les paramètres du calcul. Cependant, seuls deux paramètres semblent être réellement dépendants de la protéine : le seuil p et la tolérance Δ_{ppm} . Les autres paramètres dépendant de l'itération ont été ajustés empiriquement sur différentes protéines : dsRBD (domaine de liaison à l'ARN de *Escherichia Coli*), deux domaines de l'isomérase de ponts disulfure PDI et le domaine homologue à la pleckstrine de la β -spectrine. Ils n'ont pas à être modifiés lors d'une utilisation normale.

Le choix de la tolérance est un critère important pour la convergence et la qualité des structures. Ce paramètre est à ajuster au fur et à mesure des calculs. Le choix de Δ_{ppm} est guidé par plusieurs critères :

- ⇨ résolution des spectres,
- ⇨ estimation de l'incertitude sur l'attribution de la protéine,
- ⇨ la forme des pics

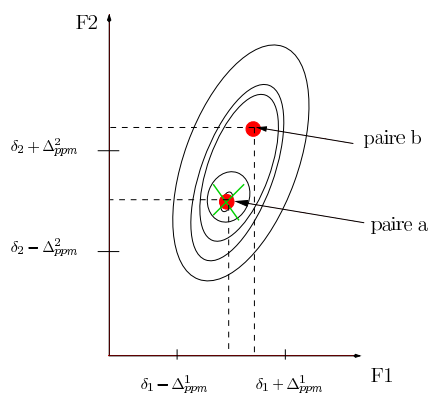


Fig. 3.11: Illustration d'une tolérance trop faible (voir le texte pour la légende)

Si Δ_{ppm} est trop petit, on prend le risque de rejeter la contribution faible d'une paire d'un pic large. Cette situation est illustrée par la figure 3.11. Les cercles rouges marquent l'emplacement du sommet des pics de corrélation des paires a et b. Les protons de la paire a sont plus proches que dans la paire b. L'application de peak-picking ne reconnaîtra qu'un seul maximum à (δ_1, δ_2) représenté par la croix verte. Avec les tolérances Δ_{ppm}^1 et Δ_{ppm}^2 définies comme sur la figure, la contrainte ambiguë correspondante ne contiendra pas la contribution de la paire b et la contribution de la paire a sera surestimée.

Si Δ_{ppm} est trop grand, la contrainte ambiguë pourra retenir des paires dont les déplacements chimiques ne sont pas compatibles avec les coordonnées du pic observé. De plus, les premières étapes du calcul seront beaucoup plus longues car le nombre de paires de protons à tester au départ sera plus grand.

La tolérance choisie doit donc être un compromis entre ces différents effets.

3.5.4 Les fichiers nécessaires à Aria

Outre les nOes, Aria gère également d'autres types de contraintes qui peuvent améliorer la convergence. La mise en oeuvre d'un calcul d'Aria nécessite de définir les fichiers suivants :

1. Séquence de la protéine, modèle de structure ou résultat du calcul précédent,
2. Attribution de la protéine pour chaque spectre,
3. Peak-picking pour chaque spectre,
4. Structure secondaire (CSI),
5. Prédiction des angles $\phi - \psi$ (Talos),
6. Constantes de couplages dipolaires résiduels,
7. Contraintes non ambiguës (ponts disulfure, liaison hydrogène,...).

3.5.5 Implémentation d'Aria

Aria est un ensemble de scripts écrits en langage Python s'interfaçant avec CNS (recompilé pour prendre en compte les ADRs). La version qui a été utilisée dans ce travail fut la 1.2⁹. L'installation d'Aria nécessite des ressources informatiques relativement importantes. En effet, 180 structures doivent être calculées. Le calcul d'une structure d'ASP2 dure 5mn sur un bi-processeur Athlon 2*1600MHz avec 1Go de mémoire vive. L'intérêt d'un cluster de PC devient évident mais le bi-processeur fut suffisant pour qu'un run d'Aria s'effectue en une nuit.

⁹disponible à l'adresse : <http://www.pasteur.fr/recherche/unites/Binfs/aria/>

Chapitre 4

Les couplages dipolaires résiduels

4.1 Le couplage dipolaire

Parmi les interactions reliant les spins, il en est une qui est omniprésente et qui fait que la résonance magnétique nucléaire est aussi utilisée pour l'étude des molécules : c'est le couplage dipolaire. Ce couplage s'établit entre deux spins A et B. Son origine physique provient de l'interdépendance de l'environnement magnétique des noyaux portant ces spins. Le moment magnétique dipolaire issu du spin I_A crée un champ magnétique dans l'espace. Le spin I_B situé à une distance r est sensible à ce champ. Il en résulte dans l'Hamiltonien du système un terme qui traduit cette interaction. Dans l'approche séculaire, il est à la fois proportionnel aux propriétés magnétiques des spins nucléaires (rapports gyromagnétiques), à la disposition réciproque des spins dans l'espace (vecteur \vec{r} entre les deux noyaux) et aux composantes z des spins I_A et I_B :

$$H^{dd}(t) = -\mu_0 \frac{\hbar\gamma_1\gamma_2}{4\pi^2} \cdot \left\langle \frac{3 \cos^2 \theta(t) - 1}{r^3(t)} \right\rangle \cdot I_{Az} \cdot I_{Bz} = D_{max}^{AB} \cdot \langle 3 \cos^2 \theta(t) - 1 \rangle \cdot I_{Az} \cdot I_{Bz} \quad (4.1)$$

$$\text{avec } D_{max}^{AB} = -\mu_0 \frac{\hbar\gamma_A\gamma_B}{4\pi^2 r^3} \text{ (Hz)}$$

L'angle θ est défini par l'angle entre le vecteur reliant les noyaux A et B et le champ magnétique \vec{B}_0 appliqué, qui définit l'axe z du repère du laboratoire et donc les composantes z des spins I_A et I_B . La moyenne est effectuée à la fois sur le temps et sur l'échantillon. On définit par constante de couplage dipolaire le facteur $D^{AB} = D_{max}^{AB} (3 \cos^2 \theta(t) - 1)$. Cette valeur sera maximale ($2D_{max}^{AB}$) pour une orientation du vecteur \vec{r} parallèle au champ \vec{B}_0 ($\theta = 0^\circ$) et minimale pour une orientation perpendiculaire ($\theta = 90^\circ$, $D_{min} = -D_{max}^{AB}$). Dans le cas d'un spin nucléaire proton amide et du spin de l'azote qui le porte ($r = 1,04$), la valeur maximale du couplage dipolaire vaut alors 21,6kHz. Il est positif car le produit des rapports gyromagnétiques du proton et de l'azote est négatif.

Dans un échantillon monocristallin, les vecteurs \vec{r} relatifs aux noyaux A et B de chaque molécule du cristal sont orientés de la même façon. L'opérateur densité évolue donc au cours d'une expérience RMN sous un terme de couplage dipolaire égal à D^{AB} . La plage de valeurs de déplacements chimiques isotropes des protons s'étale sur une fenêtre de fréquences d'environ 10000Hz. Le couplage dipolaire étant de l'ordre de 100kHz entre protons, l'éclatement des raies

est donc très important dans les spectres et les rend très difficilement interprétables lorsque la molécule contient un grand nombre d'hydrogène.

Dans le cas d'une solution isotrope, l'opérateur densité évolue sous l'Hamiltonien d'interaction moyenné sur le temps. On considère que durant le temps d'expression du déplacement chimique, le vecteur \vec{r} attaché à une molécule prend toutes les orientations possibles avec une probabilité équiprobable. La somme des $3 \cos^2 \theta - 1$ s'annule et par conséquent, le couplage dipolaire est sans effet direct sur les spectres RMN. Il n'agit qu'au second ordre dans la relaxation. En effet, les mécanismes de relaxation sont des processus essentiellement locaux. L'interaction dipolaire n'étant pas nulle à l'échelle atomique, elle procure une voie vers un retour à l'équilibre thermique du système par transfert de polarisation entre des spins couplés. Ce phénomène appelé effet nOe est constamment utilisé en RMN car il renseigne sur les proximités spatiales au sein d'une molécule.

L'échantillon analysé peut se trouver dans une situation intermédiaire entre les deux précédentes. Pour diverses raisons, l'hamiltonien d'interaction peut ne pas être annulé. Cela signifie que l'orientation du vecteur \vec{r} n'est plus totalement équiprobable dans toutes les directions de l'espace. La première application aux protéines de cet Hamiltonien non nul fut démontrée par Tolman (Tolman et al. [1995]) en utilisant l'anisotropie du tenseur de susceptibilité magnétique de la myoglobine paramagnétique. La protéine possédant un moment magnétique induit, elle s'oriente suivant une direction privilégiée. Deux années plus tard, Tjandra et Bax (Tjandra and Bax [1997]) ouvrirent la voie de l'étude de cet Hamiltonien sur d'autres protéines dont le moment magnétique induit est négligeable en utilisant des milieux faiblement orientants et en interprétant les données en termes structuraux.

4.2 Couplage dipolaire et orientation faiblement anisotrope

Pour les deux types d'orientation induites (soit directement, soit indirectement), le cadre théorique est identique : les articles de Prestegard (Prestegard et al. [2000]) et Bax (Bax et al. [2001]) font une description détaillée du couplage dipolaire résiduel et de son utilisation dans le cadre des macromolécules biologiques. Prestegard utilise le formalisme des harmoniques sphériques tandis que Bax choisit l'approche en coordonnées cartésiennes. C'est ce dernier que nous allons présenter. La littérature ne converge pas encore vers une définition unique de certaines grandeurs et en fonction des auteurs, il apparaît parfois un facteur 2 dans les formules. Nous ferons dans un premier temps l'hypothèse d'un corps rigide dans lequel le vecteur \vec{r} est constant (en norme et en direction) par rapport à un repère moléculaire, puis nous verrons les effets de la dynamique de ce vecteur sur le couplage dipolaire.

4.2.1 Cas du corps rigide

La moyenne à effectuer dans la formule 4.1 est à prendre uniquement sur la variation temporelle de θ étant donné que la distance r est constante. Le facteur $\langle 3 \cos^2 \theta - 1 \rangle$ ne se réduit plus à zéro du fait d'une répartition non équivalente pour tous les angles θ . Le développement de ce facteur doit tenir compte de cette anisotropie dans la distribution des angles θ . On définit un repère moléculaire arbitraire (x,y,z) et $c_i = \cos \alpha_i$, les cosinus directeurs du vecteur

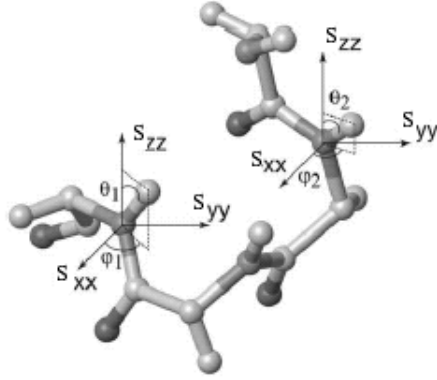


Fig. 4.1: Orientation de deux vecteurs de couplages dipolaires. Deux amides d'un fragment de protéines sont représentés. Le système d'axes du tenseur d'orientation est représenté par S_{xx} , S_{yy} et S_{zz} . Les angles ϕ et ψ définissent l'orientation des vecteurs \vec{r} par rapport au tenseur d'orientation

\vec{r} (supposé rigide) dans ce repère moléculaire. Les angles β_i définissent l'orientation des axes x, y et z par rapport au champ \vec{B}_0 . On obtient alors :

$$P_2(\cos \theta) = \frac{1}{2} \langle 3 \cos^2 \theta - 1 \rangle = \sum_{ij=\{x,y,z\}} S_{ij} \cos \alpha_i \cos \alpha_j \quad (4.2)$$

où $S_{ij} = \frac{3}{2} \langle C_i C_j \rangle - \frac{1}{2} \delta_{ij}$ et $C_k = \cos \beta_k$, δ : fonction de Kronecker

Cette représentation est un simple changement de repère moléculaire dans lequel le vecteur \vec{r} est supposé fixe. La sommation de P se ramène à sommer sur l'orientation du repère moléculaire vis-à-vis du champ, qui rentre alors dans les paramètres d'ordre S_{ij} , appelés paramètres d'ordre de Saupe et qui définissent la matrice d'ordre de Saupe. La diagonalisation de cette matrice mène aux vecteurs propres (axes x, y et z) et aux valeurs propres $\langle C_i \rangle^2$ associées qui sont physiquement interprétées comme la probabilité que la molécule soit orientée suivant l'axe i . Seule la différence relative entre les $\langle C_i \rangle^2$ contribue au couplage dipolaire. En définissant A_{ii} par $\langle C_i \rangle^2 = \frac{1}{3} + A_{ii}$, en écrivant $A_a = \frac{3}{2} A_{zz}$ et $A_r = A_{xx} - A_{yy}$ et en exprimant le vecteur \vec{r} dans le repère (x,y,z) en coordonnées polaires (θ , ϕ), le couplage dipolaire en milieu anisotrope devient :

$$D^{AB}(\theta, \phi) = D_{max} \left[\frac{1}{2} (3 \cos^2 \theta - 1) A_a + \frac{3}{4} A_r \sin^2 \theta \cos 2\phi \right] \quad (4.3)$$

La grandeur A_a est appelée composante axiale du tenseur d'orientation tandis que A_r est la composante rhombique du tenseur. Le lecteur prendra soin de ne pas confondre l'angle θ de cette relation qui est défini par l'angle entre le vecteur \vec{r} et l'axe z du repère moléculaire (axe d'orientation la plus probable de la molécule) et l'angle θ des formules précédentes qui est défini par l'angle entre \vec{r} et le champ \vec{B}_0 . Cette ambiguïté est omniprésente dans la littérature, par conséquent nous la conserverons par la suite. Ce développement appliqué à un système totalement orienté suivant le champ magnétique mène à $A_a = 1$. Dans un milieu partiellement orienté du type que l'on exploite pour les protéines, A_a traduit le degré d'orientation du milieu et est en général de l'ordre de $10^{-3} - 10^{-4}$. On trouve aussi parfois l'expression équivalente suivante avec $D_a^{AB} = \frac{1}{2} D_{max} A_a$ et $R = \frac{A_r}{A_a}$:

$$D^{AB}(\theta, \phi) = D_a^{AB} \left[(3 \cos^2 \theta - 1) + \frac{3}{2} R \sin^2 \theta \cos 2\phi \right] \quad (4.4)$$

D_a^{AB} est appelé magnitude du tenseur de couplage dipolaire résiduel et R la rhomboïté du tenseur. La valeur de R est comprise entre 0 et $2/3$. Lorsque l'orientation est axialement symétrique, $R = 0$ tandis qu'un tenseur totalement asymétrique donne une rhomboïté égale à $2/3$.

4.2.2 Cas d'une molécule flexible

L'Hamiltonien décrit par la formule 4.1 dépend du temps dans la mesure où le vecteur \vec{r} fluctue dans le temps en suivant les mouvements relatifs des atomes. L'effet sur l'hamiltonien se traduit par des variations temporelles à la fois sur la distance r et sur l'orientation θ par rapport au champ magnétique et rend l'interprétation des couplages dipolaires résiduels délicate. Les vibrations de la longueur de la liaison A-B (ex. N-H) ou de la distance entre les noyaux A et B (ex. H-H) est considérée comme étant décorrélée des autres mouvements, du fait de leur établissement sur des gammes de temps différents. Les moyennes sur r et sur θ peuvent donc être calculées séparément.

L'hamiltonien dipolaire dépend de r par $\langle 1/r^3 \rangle$. On introduit une distance effective r_{eff} définie par $1/r_{\text{eff}}^3 = \langle 1/r^3 \rangle$. Par souci de lisibilité, nous conserverons la notation r à la place de r_{eff} .

Dans la plupart des macromolécules biologiques, le vecteur \vec{r} dispose de degrés de libertés de mouvement autour de son axe moyen. Ces mouvements peuvent être de nature et d'amplitude différentes : rotation sur un cône, mouvements non axialement symétriques (ex. rotamères d'une chaîne latérale). En établissant un certain nombre d'hypothèses (tenseur d'orientation et mouvements internes décorrélés, mouvements internes restreints, mouvements internes axialement symétriques autour de l'axe moyen, pas d'échange entre plusieurs sites), la mobilité du vecteur a pour effet de réduire la valeur de la constante de couplage dipolaire résiduel d'un facteur S_{int} :

$$D^{AB}(\theta, \phi) = S_{\text{int}} D_{\text{max}} \left[\frac{1}{2} (3 \cos^2 \theta - 1) A_a + \frac{3}{4} A_r \sin^2 \theta \cos 2\phi \right] \quad (4.5)$$

Le paramètre d'ordre S_{int} introduit ici est formellement identique au paramètre d'ordre généralisé caractérisant la relaxation des spins. Cependant, il est sensible à une échelle de temps allant jusqu'à la milliseconde. Par conséquent, le paramètre d'ordre au sens couplage dipolaire est plus faible que le paramètre d'ordre au sens relaxation. En contrepartie, il n'intervient qu'à la puissance unité dans la valeur de constante de couplage dipolaire résiduel. En fonction de l'utilisation qui est faite des couplages dipolaires résiduels, il est par fois simplifié à 1, parfois à 0,8. On préconise en général de ne prendre en compte que les couplages dipolaires résiduels provenant de noyaux A-B dont le paramètre d'ordre généralisé au sens de la relaxation est supérieur à 0,6 (Clare et al. [1998b]). Dans ce cas, le paramètre d'ordre au sens du couplage dipolaire est constant sur le jeu de mesures et peut être inclus dans l'ordre du tenseur (D_a).

Dans le cas d'un saut entre deux ou plusieurs sites, le couplage dipolaire résiduel observée est la moyenne pondérée des couplages dipolaires résiduels correspondant à chaque site. L'information structurale contenue dans un tel type de couplage dipolaire résiduel est alors difficile à extraire du fait du manque de données expérimentales.

4.2.3 Problème de dégénérescence

Une des difficultés de l'interprétation des couplages dipolaires résiduels est la dégénérescence intrinsèque à l'expression mathématique 4.4. En effet, le couplage dipolaire résiduel d'un vecteur orienté dans la direction (θ, ϕ) sera égal aux couplages des vecteurs dans les directions $(\theta, \phi + \pi)$, $(\theta + \pi, \phi)$, et $(\theta, \pi - \phi)$. Les rotations de 180° correspondantes sont respectivement d'axes z , x et y . Le profil de couplages dipolaires d'une molécule sera donc identique après ces trois rotations axiales, il y a donc quatre orientations acceptables du repère moléculaire vis-à-vis du tenseur d'orientation. Inversement, il existe quatre orientations possibles du tenseur par rapport au repère moléculaire. Cette dégénérescence reste un obstacle lors de l'étude de l'orientation d'une molécule dans un seul milieu orientant. Une des solutions est l'utilisation de plusieurs milieux orientants de tenseurs non colinéaires.

4.3 Les couplages dipolaires dans la pratique

4.3.1 Les milieux orientants

L'origine physique de la création d'anisotropie orientationnelle dans un tube de RMN naît de l'interaction entre la molécule étudiée et des particules elles-mêmes orientées. L'orientation de ces particules peut être spontanée ou induite par un champ magnétique ou électrique (Hilbers and MacLean [1968]). Les interactions possibles sont les interactions stériques (dépendant de la forme de la molécule) et les interactions de charge. Le tenseur d'alignement dépend du type d'interaction et du milieu.

L'ordre induit doit impérativement être inférieur à 10^{-3} pour conserver l'aspect du spectre en solution isotrope. Bien entendu, il doit être supérieur à 10^{-5} pour que la constante de couplage résiduel puisse être observée. Ces valeurs dépendent évidemment du type de vecteur dipolaire. En théorie, tout type de vecteur ($^1H - ^1H$, $^{15}N - ^1H$, $^{13}C - ^1H$) peut être étudié. Dans la pratique, les constantes de couplage dipolaire entre protons sont difficilement interprétables du fait de la connaissance peu précise de la distance entre hydrogènes. Tjandra (Tjandra et al. [2000]) a cependant montré que l'utilisation de ce type de constante de couplage résiduel a amélioré la précision, l'exactitude et la qualité du calcul de la structure de l'ubiquitine.

Les milieux orientant les plus utilisés actuellement sont des phases nématiques (cristaux liquides), mais l'ordre peut également être induit par compression mécanique (gel de polyacrylamide contraint). Les phases nématiques sont obtenus par l'utilisation de bicelles (ex. mélange DMPC/DHPC), de bactériophages Pf1, de mélange ternaire chlorure ou bromure de cétylepyridium/hexanol/NaCl ou NaBr, de fragments de membrane pourpre, de cellulose, de mélange hexanol/alkyle-PEG. Le choix du milieu orientant est dicté par le diagramme de phase du cristal liquide. Les conditions expérimentales d'existence de la phase nématique (pH, salinité, concentration en nématogène, température) ne sont pas toujours adaptées à l'étude d'une protéine donnée. De plus, le type d'interaction qui gouverne l'orientation est un facteur important : la surface des bactériophages a une charge nette négative et est donc déconseillée pour l'orientation de molécules chargées globalement positivement.

Milieu orientant historique (Tjandra and Bax [1997]), la phase cristal liquide DMPC/DHPC (figure 4.2) forme des bicelles (figure 4.3) en solution (Sanders and Schwonek [1992]). Ce sont des particules micellaires planes constituées de phospholipides saturés. Le phospho-

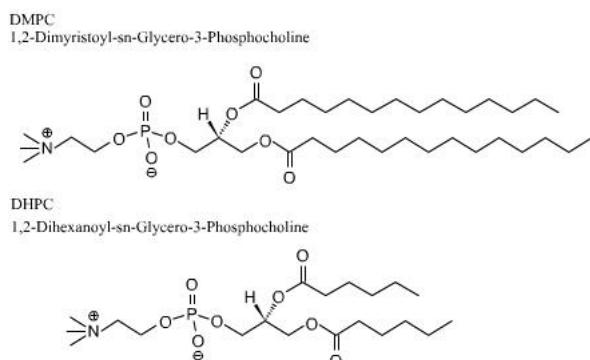


Fig. 4.2: Formules du DMPC et du DHPC.

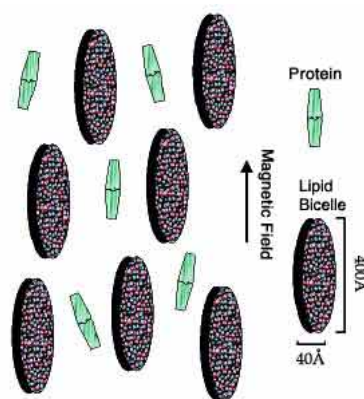


Fig. 4.3: Représentation schématique des bicelles. Le champ magnétique ordonne les disques bicellaires. La protéine en solution est alors contrainte stériquement à s'orienter préférentiellement parallèlement au champ.

lipide à chaîne longue (DMPC) constitue la bicouche tandis que les bords sont fermés par les phospholipides à chaîne courte (DHPC). En présence d'un champ magnétique, les disques s'orientent parallèlement au champ. La figure 4.3 est une représentation des bicelles. Le disque a un diamètre d'environ 200-250 Angstroems et une épaisseur de 40 Angstroems. Pour certains auteurs (Gaemers and Bax [2001]), il faut plutôt considérer l'image en négatif : les chaînes longues forment toujours la bicouche mais les chaînes courtes crée des pores au sein de la bicouche. La gamme de température d'existence de la phase nématique est comprise entre 25°C et 45°C . En dehors de cet intervalle, la phase est isotrope. La proportion $q = [\text{DMPC}] / [\text{DHPC}]$ ainsi que la concentration totale en nématogène sont les paramètres importants. Les bicelles sont stables sur une gamme assez large de pH, bien qu'un pH de 6 ralentit l'hydrolyse des phospholipides. L'ordre induit dépend à la fois de la concentration en bicelle et des concentrations relatives en nématogènes. Nous avons utilisé des échantillons de bicelle de rapport $q=3$ à des concentrations de 3% et 10%(w/v). A ces concentrations, la phase orientée existe pour une température de $35 - 42^{\circ}\text{C}$ (Tjandra and Bax [1997]).

Il existe également des nématogènes éther contenant une liaison carboxy-éther au lieu de la liaison carboxy-ester : ditétradécyle-PC (1,2-di-O-tétradécyle-*sn*-glycéro-3-phosphocholine) et dihexyle-PC (1,2-di-O-héxyle-*sn*-glycéro-3-phosphocholine). L'avantage de cette substitution est de protéger la phase bicellaire vis-à-vis de l'hydrolyse acide ou basique(Ottiger and Bax [1999]) et donc de pouvoir travailler sur une gamme de pH plus large (de 2,3 à 14). Enfin, Wang et ses collaborateurs (Wang et al. [1998]) ont développé des bicelles à partir d'un mélange DLPC (1,2-didodécanoyle-*sn*-glycéro-3-phosphocholine) et de CHAPSO (sulfonate de 3-(cholamidopropyle)diméthylammonio-2-hydroxyle-1-propane), qui permet de mesurer des couplages dipolaires résiduels sur une gamme de température allant de 12 à 50°C .

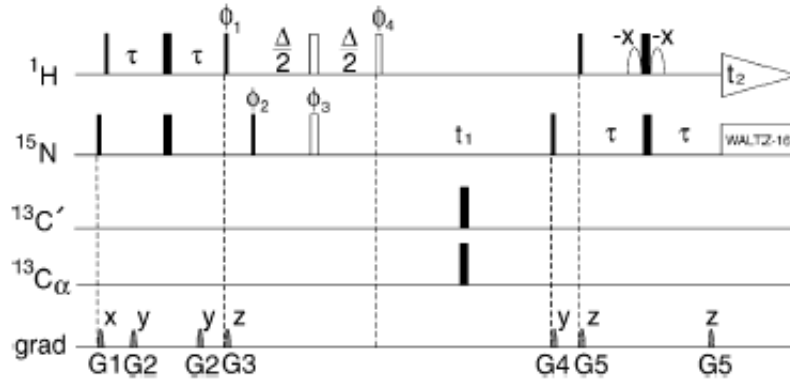


Fig. 4.4: Séquence d'impulsions de l'IPAP (^1H - ^{15}N). Les impulsions sont de phase x par défaut. Les impulsions creuses ne sont appliquées que pendant l'acquisition du spectre antiphase. L'acquisition des spectres en phase et en antiphase sont enchevêtrées. Les délais τ et Δ valent respectivement 2,5 et 5,3 ms. Les impulsions de découplage carbone n'ont pas été appliquées car la protéine est simplement marquée. Le cyclage de phase est le suivant : $\phi_1 = -y, y$; $\phi_2 = 2(x), 2(-x)$ pour la composante en phase ; $\phi_2 = 2(-y), 2(y)$ pour la composante en antiphase ; $\phi_3 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_4 = 8(x), 8(-x)$; $\phi_{rec} = x, 2(-x), x$. Une WaterGate (Piotto et al. [1992]) supprime le signal de l'eau. La détection en quadrature en t_1 est assurée en incrémentant ϕ_2 pour le spectre en phase et ϕ_2 et ϕ_3 simultanément pour le spectre en antiphase à la manière States-TPPI (Marion et al. [1989]).

4.3.2 La mesure des couplages dipolaires résiduels.

En phase anisotrope, l'hamiltonien sous lequel évolue l'opérateur-densité contient le terme de couplage dipolaire $D \cdot I_{Az} \cdot I_{Bz}$, formellement identique à l'hamiltonien de couplage scalaire $J \cdot I_{Az} \cdot I_{Bz}$. L'éclatement des raies dans une expérience de type HSQC non découplé est donc égal à $J+D$. Dans l'étude d'ASP2, nous nous sommes contentés de mesurer les constantes de couplage dipolaire résiduel entre les noyaux proton et amide du squelette, sur un échantillon marqué ^{15}N . En suivant la méthode de Tjandra et Bax (Tjandra and Bax [1997]), on mesure l'écartement des doublets vers 35°C ($J+D$, phase anisotrope) et à une température inférieure à 25°C (J , phase isotrope). Le couplage dipolaire résiduel est la différence entre les écartements aux deux températures. Il faut faire également attention au signe de J . La constante de couplage scalaire $^1J_{H_N N}$ vaut environ -94 Hz (Annala et al. [1999]). Un écart entre doublets supérieur en phase anisotrope par rapport à la phase isotrope signifie que le couplage dipolaire résiduel est négatif. Cependant, la littérature ne converge pas entièrement sur ce point. Si Annala mesure les couplages dipolaires résiduels en tenant compte du signe de J , Bax (Bax and Tjandra [1997]) considère que le couplage dipolaire résiduel est la différence nette d'éclatement. Le changement de signe intervient alors dans la magnitude du tenseur (D_a) et n'est pas problématique si on ne travaille que sur un seul type de vecteur.

Dans la pratique, la mesure de l'écart entre les doublets ne s'effectue pas sur une expérience de type HSQC non découplée dans une des dimensions. En effet, l'absence de découplage double le nombre de raies dans le spectre et rend par conséquent plus difficile l'analyse. La séquence d'impulsions que j'ai utilisée a été mise au point par Ottiger (figure 4.4, Ottiger et al. [1998]). Le but de cette expérience est d'enregistrer simultanément un spectre en phase, où les

composantes du doublet sont de même signe, et un spectre en antiphase, où les composantes sont de signe opposé. L'éclatement des raies est mesurée en dimension indirecte azote en raison d'une plus grande finesse de raie.

Le spectre en phase est accumulé de la même façon qu'une HSQC non découplée durant l'édition de l'azote : transfert INEPT de I_y vers $I_x S_z$, passage vers $I_z S_y$, évolution sous le déplacement chimique azote et le couplage $I_z S_z$, retour sur $I_y S_z$, INEPT retour pour revenir vers un état pur proton I_x , acquisition. L'opérateur-densité est proportionnel durant l'acquisition à $\cos(\pi J t_1) e^{-i\omega_N t_1} \cos(\omega_H t_2)$. Le spectre en antiphase est obtenu en insérant dans la séquence un module de refocalisation $\Delta/2 - 180^\circ - \Delta/2 - 90^\circ$ avant l'évolution pendant t_1 . L'état $I_z S_y$ devient avant le délai d'édition : $-\sin(\pi J \Delta) S_y$. La polarisation S_y évolue pendant t_1 vers un état en antiphase :

$$-\sin(\pi J \Delta) S_y \xrightarrow{t_1} -2 \cos(\omega_N t_1) \sin(\pi J t_1) \sin(\pi J \Delta) I_z S_x + 2 \sin(\omega_N t_1) \sin(\pi J t_1) \sin(\pi J \Delta) I_z S_y + IP \quad (4.6)$$

Les termes en phase IP (états S_x et S_y) ne seront pas convertis par la suite vers une polarisation I observable. L'élément INEPT retour transfère la polarisation vers une polarisation I observable modulée par $-\sin(\pi J t_1) e^{-i\omega_N t_1}$. La transformée de Fourier de ce signal donne lieu à un doublet dont les pics sont de signes opposés. Le délai Δ est ajusté à 5,3 ms pour optimiser la refocalisation vers l'état S_y . La dispersion des couplages scalaires N-H est suffisamment faible pour que le transfert soit efficace pour tous les résidus mais du fait de la composante dipolaire, ce transfert n'est pas toujours complet, en particulier lorsque l'ordre induit est trop important.

La somme et la soustraction de ces deux spectres mènent à des spectres ne contenant qu'une seule des composantes du doublet. Le mélange des expériences doit tenir compte des intensités relatives entre les deux expériences. L'intensité des pics en antiphase est légèrement plus faible du fait de la durée plus longue de l'expérience. Par cette méthode, l'encombrement des spectres n'est jamais supérieur à celui de l'HSQC.

4.3.3 Détermination du tenseur d'orientation.

La principale difficulté de l'interprétation des couplages dipolaires résiduels est la détermination du tenseur d'orientation. Lorsque l'interaction stérique est responsable de l'anisotropie d'orientation, ce tenseur peut être calculé si on connaît déjà le repliement global de la protéine. En effet, le tenseur d'orientation coïncide alors quasiment avec le tenseur de diffusion rotationnelle. Zweckstetter (Zweckstetter and Bax [2000]) modélise les bicelles par un mur infini parallèle au champ. Les protéines en solution, éloignées de ce mur ont un comportement isotrope, tandis que la fraction de protéines stériquement au contact des bicelles ont des orientations limitées. Le tenseur d'orientation est calculé à partir de la restriction des orientations de la protéine, fonction de sa distance au mur.

Quand on dispose d'un modèle structural, la rhomboïté du tenseur peut être dérivée d'une recherche par grille sur R afin de minimiser l'écart entre la mesure et le calcul des couplages

dipolaires résiduels (Clare et al. [1998b]). Lorsqu'aucune information structurale n'est disponible, la magnitude et la rhomboïté du tenseur peuvent être estimées à partir du spectre de poudre. Lorsque les vecteurs d'une molécule échantillonnent l'ensemble des orientations possibles vis-à-vis du tenseur, la distribution des valeurs de couplages dipolaires suit la loi du spectre de poudre classiquement retrouvé en RMN du solide. Les valeurs maximales D_{zz} et minimales D_{yy} et de plus forte probabilité D_{xx} du couplage dipolaire résiduel sont directement reliées aux paramètres D_a et R du tenseur. La figure 4.5 montre la valeur du couplage dipolaire résiduel en fonction de l'orientation du vecteur dipolaire \vec{r} par rapport aux axes propres du tenseur. La valeur est maximale ($D_{zz} = 2D_a$) quand \vec{r} est parallèle à l'axe z ($\theta = 0^\circ$), minimale ($D_{yy} = -D_a(1 + 1,5R)$) suivant l'axe y ($\theta = 90^\circ$, $\phi = 90^\circ$). La valeur de couplage la plus probable (zone blanche de la figure, $D_{xx} = -D_a(1 - 1,5R)$) est proche de celle obtenue pour une orientation suivant l'axe x ($\theta = 90^\circ$, $\phi = 0^\circ$).

On extrait donc la magnitude et la rhomboïté par :

$$D_a = \frac{D_{zz}}{2} \quad (4.7)$$

$$R = \frac{2(D_{xx} - D_{yy})}{3D_{zz}} \quad (4.8)$$

L'article de Clare (Clare et al. [1998a]) montre que cette méthode peut être utilisée pour estimer la magnitude et la rhomboïté du tenseur à partir de la distribution des couplages dipolaires résiduels. Bien entendu, cela implique que les vecteurs d'interaction dipolaire soient uniformément répartis dans l'espace. Ce n'est pas généralement le cas pour une protéine essentiellement constituée d'hélices. Les vecteurs N-H dans une hélice α sont orientés plus ou moins tous dans la même direction (voir section 4.5). La solution est de mesurer l'interaction d'autres vecteurs (ex. $C_\alpha - H_\alpha$) qui explorent une plus grande partie de l'espace. L'autre faiblesse de cette approche est de sous-estimer D_{zz} car la probabilité pour un vecteur de faire un angle θ avec l'axe z est égale à $\sin \theta$, soit nulle pour $\vec{r} \parallel \vec{z}$. La probabilité est donc faible que dans une protéine d'une centaine de résidus, la valeur théorique maximale de couplage soit atteinte. Clare (Clare et al. [1998b]) estime à 15 – 20% la sous-estimation de D_a . Par contre, la valeur de D_{yy} est plus fiable et est utilisée dans cet article pour calculer $D_a = -D_{zz}(1 + 1,5R)$ et R par une recherche par grille du R qui minimise l'énergie totale de la structure. Cette méthode fournit donc une valeur minimale pour la magnitude.

Lorsque le repliement est connu ou partiellement connu (protéine multidomaine), le tenseur d'orientation de la structure ou d'un fragment de structure peut être calculé soit par minimisation des moindres carrés de la fonction $\chi^2 = \sum_n (D_{ij}^{obs} - D_{ij}^{calc})^2 / \sigma_{ij}^2$ où σ_{ij} est l'incertitude sur la mesure de D_{ij}^{obs} , en variant aléatoirement les paramètres D_a , R et l'orientation du tenseur vis-à-vis du repère moléculaire (Logiciel Module, Dosset et al. [2001]), soit par décomposition en valeurs singulières (Losonczi et al. [1999]).

4.4 Applications des couplages dipolaires résiduels

Avant l'avènement des couplages dipolaires résiduels, le calcul de structure de macromolécules était basé sur des contraintes principalement locales. Les contraintes de distance issues

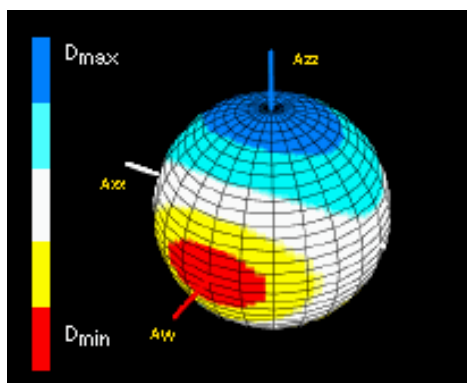


Fig. 4.5: Valeur du couplage dipolaire résiduel en fonction de l'orientation du vecteur \vec{r} dans le tenseur d'orientation défini par A_{xx} , A_{yy} et A_{zz} . Par définition, $|A_{zz}| > |A_{yy}| > |A_{xx}|$. (Image extraite du manuel de Module)

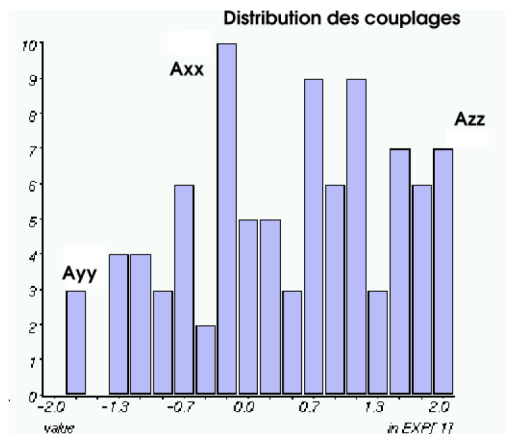


Fig. 4.6: Distribution des couplages dipolaires résiduels et extraction des paramètres du tenseur d'orientation

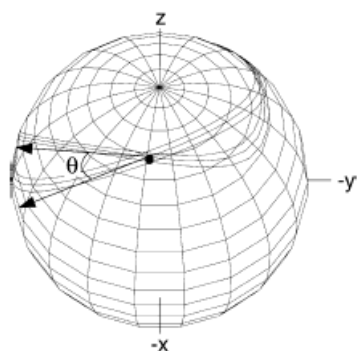


Fig. 4.7: Résolution du problème de dégénérescence par l'utilisation de deux milieux orientant en l'absence de données structurales. Les deux courbes sur la sphère sont définies par les couples (ϕ, ψ) d'un vecteur NH satisfaisant la relation 4.4 dans les deux milieux. Le repère (x, y, z) est un repère moléculaire arbitraire. La dégénérescence est réduite à au plus deux possibilités à l'intersection des courbes (Ramirez and Bax [1998]).

de l'observation de l'effet nOe entre protons sont limités à quelques Angstroems. La détermination de la constante de couplage scalaire 3J mène à des contraintes diédrales. Dans certains cas, la présence d'ion paramagnétique permet d'introduire des contraintes de distance allant jusqu'à quelques dizaines d'Angstroems. La puissance du couplage dipolaire résiduel est de contenir intrinsèquement des contraintes d'orientation globale de vecteurs au sein de la molécule. En effet, le tenseur d'orientation anisotrope est une donnée par essence globale à la molécule, car il naît de l'interaction de la molécule dans son ensemble avec le milieu ou avec le champ magnétique. Le couplage dipolaire résiduel issu de chaque vecteur intramoléculaire est donc lié fortement à ce tenseur. Lorsque le chercheur dispose d'un jeu de constantes de couplage dipolaire résiduel, il peut les utiliser de plusieurs façons en fonction de ses besoins. Nous allons décrire quelques-unes de ces applications. Les revues de De Alba (de Alba and Tjandra [2002]), Prestegard (Prestegard et al. [2000]) et Bax (Bax et al. [2001]) rassemblent la plupart des applications existantes à ce jour.

4.4.1 Affinement de structures avec les couplages dipolaires résiduels.

Les couplages dipolaires peuvent être utilisés pour améliorer la convergence d'un calcul de structures par l'incorporation dans la fonction cible de recuit simulé d'une pseudo-énergie potentielle traduisant la nécessité pour chaque vecteur \vec{r} de satisfaire le couplage dipolaire observé pour chaque mesure (Tjandra et al. [1997]) :

$$E_{dd} = W \sum_i (D_{i,th} - D_{i,obs})^2 \quad (4.9)$$

$D_{i,th}$ et $D_{i,obs}$ sont respectivement la constante de couplage dipolaire résiduel calculée à partir de la structure en cours de minimisation et la constante mesurée expérimentalement. La somme est prise sur toutes les constantes de couplage mesurées et W donne le poids attribué à ce type de contraintes par rapport aux autres contraintes (nOe, diédrales, champ de force moléculaire). Les logiciels de reconstruction (CNS, Aria, Xplor) proposent maintenant d'inclure ce terme. La difficulté de cette approche est la détermination du tenseur d'orientation (D_a , R et orientation du tenseur par rapport au repère moléculaire). L'orientation du tenseur est définie par un repère (O,X,Y,Z). Si on dispose de la magnitude et de la rhomboïcité du tenseur (voir paragraphe 4.3.3), l'orientation relative du tenseur d'orientation est laissée libre durant le protocole de recuit simulé jusqu'à obtention du minimum d'énergie.

Jens Meiler (Meiler et al. [2000a]) propose d'exclure le tenseur d'orientation du calcul en traduisant les constantes de couplage dipolaire en contraintes diédrales entre les vecteurs \vec{r} . Le principe de cette méthode est de comparer deux à deux les valeurs de couplages dipolaires résiduels. Si le tenseur est axialement symétrique ($R = 0$) et $D_a = 5Hz$, le fait que deux vecteurs amides aient une constantes de couplage résiduel égal à 9Hz, signifie qu'ils ont tous les deux une orientation de moins de 15° par rapport à l'axe z du tenseur, par conséquent leur orientation relative doit être comprise dans un intervalle (moins de 30° ou plus de 150°). L'absence de contraintes pour certains couples de vecteurs est un inconvénient majeure de cette méthode mais il est en partie contrebalancé par une nombre supérieur de contraintes (N contraintes pour l'approche précédente contre au plus $N(N - 1)/2$ contraintes diédrales). La

génération des contraintes diédrales est automatiquement assurée par le logiciel Dipocoup¹⁰ (Meiler et al. [2000b]) et le fichier de sortie est au format CNS.

On peut également citer la méthode développée par Moltke et Grzesiek (Moltke and Grzesiek [1999]). Ils construisent à partir du problème de minimisation des moindres carrés de la formule 4.9 qui dépend explicitement des paramètres d'ordre d'orientation par $D_{i,th}$ (non-linéaire en (θ, ϕ) et linéaire en S_{ij}), un nouveau problème de minimisation non-linéaire en (θ, ϕ) mais indépendant de S_{ij} . La contrepartie est un aplatissement du profil d'énergie potentielle.

4.4.2 Orientation de sous-domaines de protéine

Plusieurs articles font état de l'utilité des couplages dipolaires résiduels pour la détermination de l'orientation relative de sous-domaines. Les contraintes locales (nOe, angles dièdres) sont dans ces systèmes insuffisantes pour positionner correctement les deux sous-domaines l'un par rapport à l'autre. Disposant de la structure des deux domaines et un jeu de couplages dipolaires résiduels sur les deux domaines dans deux milieux orientant, on définit un tenseur unique pour les deux domaines (dont la structure a été conservée durant le recuit simulé) ce qui détermine une orientation unique des deux sous-domaines. Skrynnikov (Skrynnikov et al. [2000]) et Fischer (Fischer et al. [1999]) adoptèrent une telle solution pour déterminer l'orientation relative des deux domaines de leurs protéines, respectivement la MBP (protéine liant la maltodextrine) et la protéine ribosomale S4Δ41.

Cependant, l'approche d'un tenseur unique suggère que les mouvements entre les différentes sous-unité soient restreints. Lorsque les deux domaines présentent une dynamique relative fortement décorrélée, un seul tenseur d'orientation ne peut être défini. L'impossibilité de déterminer un tenseur unique est donc signe d'une forte dynamique.

4.4.3 Identification du repliement de protéines.

La détermination de la structure d'une protéine sur la base unique des constantes de couplage dipolaire est malheureusement encore inaccessible du fait de la présence de plusieurs minima d'énergie potentielle. Dans le cadre de la génomique structurale, l'utilisation des bases de données de structures devient indispensable pour accélérer la détermination de la structure d'une nouvelle protéine. Plusieurs équipes tente de relier un jeu de couplages dipolaires mesurés sur une protéine pour reconnaître dans la base de données, des structures ou des fragments de structures qui, dans un même tenseur d'orientation, donneraient le même motif de couplages dipolaires. En effet, Annala (Annala et al. [1999]) a montré que la ressemblance structurale entre protéines homologues se traduisait par une ressemblance du profil des couplages dipolaires résiduels.

Dans l'article de présentation de Dipocoup (Meiler et al. [2000b]), l'équipe de Griesinger présente leur idée. Il s'agit de rechercher dans la base de données les séquences dont le repliement est au moins partiellement homologue à la protéine d'intérêt. Cette recherche s'effectue en alignant les structures secondaires (obtenus par CSI) et en les orientant les unes par rapport aux autres en utilisant les couplages dipolaires. Pour cela, Dipocoup calcule les couplages

¹⁰disponible à l'adresse : <http://www.jensmeiler.de>

dipolaires puis les compare avec ceux obtenus sur la protéine d'intérêt. Les structures secondaires sont ensuite alignées progressivement et la qualité de l'alignement est estimée par la déviation standard entre la valeur mesurée et la valeur calculée sur la structure comparée. Le logiciel fournit donc un modèle structural satisfaisant les premières données de RMN. Bien entendu, il faut que la base de données contienne une structure homologue et par conséquent, un repliement entièrement nouveau ne donnerait pas de résultat fiable.

Delaglio (Delaglio et al. [2000]) adopte une méthode basée sur la reconstruction progressive de la structure à partir de fragments de 7 acides aminés. Il a extrait 350000 fragments de la banque de structure PDB. Pour chaque fragment, le meilleur ajustement entre les couplages mesurés et les couplages calculés sur ces fragments est estimé par un χ^2 . En tenant compte également les déplacements chimiques lorsqu'ils sont connus, les 20 fragments de χ^2 les plus faibles sont sélectionnés. Les angles (ϕ , ψ) de ces 20 fragments sont calculés et lorsque la majorité d'entre eux se trouve dans la même région de la carte de Ramachandran, ils sont utilisés pour la reconstruction. La pertinence de cette approche a été démontrée par les auteurs sur l'ubiquitine.

L'équipe de Baker a également mis au point un algorithme pour obtenir un modèle structural satisfaisant à la fois les couplages dipolaires et l'homologie de séquence (Rohl and Baker [2002]).

4.4.4 Module¹¹ : un programme pour l'orientation de domaines ou de fragments.

L'équipe de Blackledge à Grenoble met à la disposition de la communauté scientifique une application (Module, (Dosset et al. [2001]) qui permet d'orienter des modules (sous-domaines ou structures secondaires) les uns par rapport aux autres. Pour chaque module, il détermine le tenseur d'orientation sur la base des constantes de couplage dipolaire. Dans l'étape suivante, il aligne les tenseurs les uns par rapport aux autres. Du fait de la dégénérescence, les tenseurs et donc les modules peuvent être tournés de 180° autour des axes du tenseur. L'opérateur peut donc construire visuellement un modèle. Quelques contraintes de distances peuvent être utilisées pour arriver plus vite au modèle. La difficulté majeure est de gérer les dégénérescences dès lors que le nombre de modules est supérieur à trois.

4.4.5 Couplages dipolaires résiduels et dynamique.

Griesinger et Bruschweiler (Peti et al. [2002]) ont utilisé les RDCs pour sonder la dynamique des vecteurs observés. L'idée principale de leurs travaux est que les couplages dipolaires résiduels échantillonnent la dynamique du vecteur sur une gamme de temps plus large que celle accessible par relaxation. Les RDCs sont sensibles aux variations angulaires par rapport à l'axe z du tenseur d'orientation par le terme S_{int} . En utilisant plusieurs milieux orientants, l'axe de référence z prend différentes orientations et on peut alors étudier l'anisotropie de réorientation du vecteur dipolaire.

¹¹disponible à l'adresse : <http://www.ibs.fr/ext/labos/LRMN/softs/>

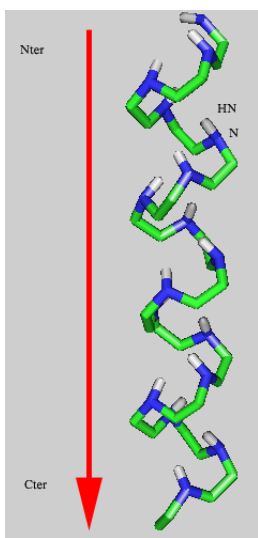


Fig. 4.8: Les vecteurs N-H sont distribués sur un cône d’angle 15° autour de l’axe de l’hélice.

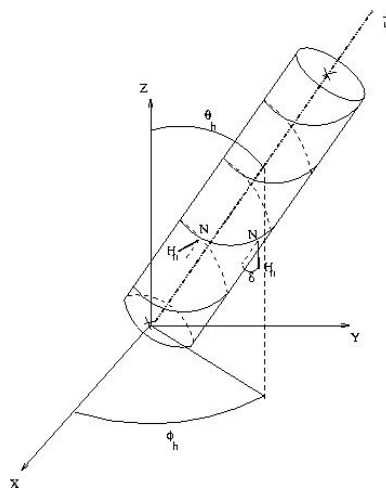


Fig. 4.9: Définition des angles θ_h et ϕ_h de l’hélice par rapport aux axes propres du tenseur d’orientation.

4.5 Les ondes dipolaires (“Dipolar Waves”)

4.5.1 L’origine des ondes dipolaires

Comme le montre la figure 4.8, les vecteurs N-H d’une hélice α parfaite ont une orientation restreinte sur un cône d’angle $\delta = 15,8^\circ$ de même axe que l’hélice. De plus, la périodicité de 3,6 résidus par tour dans l’hélice se retrouve sur l’orientation des vecteurs N-H : tous les 3,6 résidus, ce vecteur est dirigé dans la même direction. L’expression du couplage dipolaire est indépendante de la position dans l’espace et n’est liée qu’à l’angle entre ce vecteur et un axe. Le couplage dipolaire résiduel doit donc être périodique de période 3,6 résidus.

Cette propriété est déjà exploitée en RMN de l’état solide. Les expériences de type PISEMA (*polarization inversion spin exchange at magic angle*, Wu et al. [1994]) corrélient les interactions dipolaires et d’anisotropie de déplacement chimique (CSA). La non-colinéarité des deux Hamiltoniens associés mène à des courbes (PISA, *Polar Index Slant Angle*, Denny et al. [2001]) dans les spectres PISEMA paramétrées par l’angle τ de l’hélice par rapport à un axe de référence (normale à la bicouche, champ magnétique \vec{B}_0). La position sur la courbe du résidu n de l’hélice est lié à sa “polarité” PISA. Par cette méthode, on a ainsi accès par exemple à l’orientation d’une hélice par rapport à une bicouche sans disposer de l’attribution des résonances (Wang et al. [2000]).

Mesleh (Mesleh et al. [2002]) a introduit la réduction unidimensionnelle du PISA (qu’il appelle onde dipolaire, ou *dipolar wave*) en formulant la fonction périodique du couplage dipolaire résiduel dans une hélice en solution faiblement anisotrope :

$$D_{helice}(n) = D_a \left\{ (3 \cos^2(\theta_h - \delta \cos \rho) - 1) + \frac{3}{2} R \sin^2(\theta_h - \delta \cos \rho) \cos 2(\phi_h + \delta \sin \rho) \right\} \quad (4.10)$$

L’hélice α est orientée par les angles (θ_h, ϕ_h) par rapport au tenseur d’orientation. Les angles définissant la direction du vecteur N-H dans ce tenseur sont alors $(\theta_h - \delta \cos \rho, \phi_h + \delta \sin \rho)$

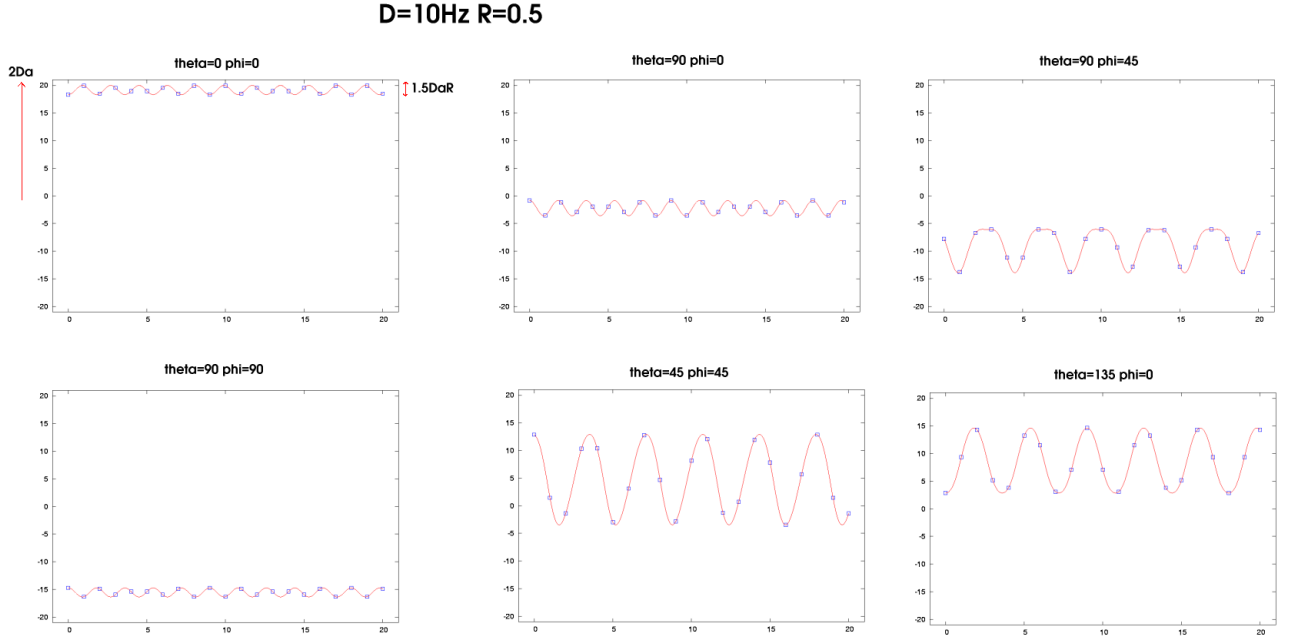


Fig. 4.10: Valeurs de couplages dipolaires résiduels simulés sur une hélice α parfaite pour un tenseur d’orientation de magnitude 10Hz et de rhomboïté 0,5. Le profil dépend de l’orientation (θ, ϕ) de l’hélice dans le repère du tenseur.

où ρ est la phase dans l’hélice (ρ est l’angle entre la projection du vecteur N-H dans le plan perpendiculaire à l’axe de l’hélice et un axe arbitraire de ce plan). La périodicité de l’orientation du vecteur N-H est contenue dans cette phase : $\rho = \frac{n}{3,6}2\pi + ph$ où n est le numéro du résidu et ph la phase initiale dans l’hélice¹². Faire varier ph revient à faire tourner l’hélice autour de son axe.

La relation 4.10 fait apparaître plusieurs dégénérescences par rapport à θ, ϕ et ph . On retrouve bien entendu les mêmes symétries par rapport aux axes du tenseur d’orientation que la formule habituelle des couplages dipolaires résiduels (formule 4.4, partie Matériel et Méthodes). La phase ph a pour effet de rajouter deux autres types de dégénérescence. Le signe des angles θ et ϕ ne peut être obtenu simultanément car un déphasage de ph se traduit par une inversion du signe de ces angles (relation ??). Cette opération est tout à fait possible car la phase ph n’a pas de sens physique et dépend de la façon dont on numérote les résidus au sein d’une hélice.

$$(\theta, \phi, ph) \leftrightarrow (-\theta, -\phi, ph + \pi) \quad (4.11)$$

La figure 4.10 rassemble les différentes allures de l’onde dipolaire simulée sur des hélices α parfaites dont l’orientation par rapport au tenseur varie. Bien entendu, la dégénérescence d’ordre quatre est conservée par la périodicité. L’orientation de l’hélice ne pourra donc être

¹²Remarque : l’article de Mesleh omet la phase ph dans son expression.

obtenue qu'aux rotations près.

Cette méthode permet a priori outre l'identification des hélices α , la distinction entre les hélices α et 3_{10} et également la présence d'un coude dans un hélice.

Mesleh publia son article quelques semaines après que nous ayons commencé la validation de cette hypothèse sur ASP2.

4.5.2 Calcul de l'orientation de l'hélice.

En l'absence de la connaissance du tenseur d'orientation, il y a donc 5 paramètres ajustables (D_a , R , θ_h , ϕ_h et ρ). Le nombre de données est limité par la longueur de l'hélice et par le nombre de mesures de couplages dipolaires résiduels sur cette hélice. Mesleh dans son article ne détaille pas la procédure d'ajustement qu'il a utilisée pour obtenir ces paramètres. Compte-tenu du grand nombre de paramètres à ajuster vis-à-vis du nombre de données, j'ai choisi de déterminer les paramètres du tenseur par un autre méthode (Module) et de procéder par une recherche par grille sur θ_h , ϕ_h et ρ (D_a et R restent constants ou peuvent évoluer dans un intervalle d'incertitude choisi par l'opérateur). La méthode de recherche par grille consiste à minimiser une fonction cible en partant de valeurs initiales de paramètres à ajuster qui varient. L'avantage est d'éviter de retenir un minimum local et la chance d'obtenir le minimum global est supérieure. Plus l'incrément des paramètres est faible, plus la probabilité est importante mais plus le calcul est long.

Pour accélérer le calcul, l'espace des (θ_h, ϕ_h) échantillonné durant la recherche par grille est restreint à $([0; 90^\circ], [0; 90^\circ])$. Les paramètres initiaux dans les autres quadrants dérivent des symétries et sont par conséquent inutiles dans la mesure où ils donneront les mêmes résultats aux symétries près. Une fonction χ^2 évalue la distance entre les couplages dipolaires calculés D_{calc}^n et mesurés D_{obs}^n sommés sur l'hélice :

$$\chi^2 = \frac{1}{N-p} \sum_{n=1}^N (D_{calc}^n - D_{obs}^n)^2 \quad (4.12)$$

L'indice n (numéro du résidu dans l'hélice) porte sur les N mesures de couplages dipolaires. N est inférieur à la longueur de l'hélice et les indices successifs ne correspondent par nécessairement à des résidus séquentiels. p est le nombre de paramètres à ajuster. Pour chaque jeu de paramètres initiaux (θ_h , ϕ_h et ρ), un algorithme d'optimisation non-linéaire cherche le minimum local le plus proche. En fin de recherche par grille, les paramètres calculés après minimisation de χ^2 le plus faible sont sélectionnés.

J'ai automatisé l'algorithme et les calculs par un programme écrit dans le langage Octave¹³, version libre de Matlab. Le programme d'optimisation non-linéaire que j'ai choisi s'appelle SOLNP¹⁴. Il présente l'avantage de pouvoir minimiser la fonction coût χ^2 sous contrainte, à savoir l'intervalle de variation de chacun des paramètres. Les paramètres du tenseur d'orientation sont par exemple libres d'évoluer dans un intervalle (1Hz pour D_a et 0,1 pour R) autour des valeurs calculées par Module. Le calcul de l'incertitude sur l'angle n'est pas pris en compte dans les programmes. Les programmes ainsi que leur utilisation sont disponibles à l'annexe D.

¹³disponible à l'adresse : <http://www.octave.org>

¹⁴disponible à l'adresse : <http://dollar.biz.uiowa.edu/col/ye/matlab.html>

Chapitre 5

Analyse de la dynamique interne d'une protéine

5.1 Éléments de théorie

Outre les mouvements de diffusion rotationnelle et translationnelle, une protéine en solution possède des degrés de liberté internes. On distingue donc les mouvements globaux de la protéine (diffusion d'un corps rigide) et les mouvements des vecteurs internes (liaisons atomiques).

La RMN est un outil puissant et très fin pour analyser l'ensemble de ces mouvements sur plusieurs ordres de grandeur et à l'échelle atomique. Chaque spin nucléaire agit comme une sonde extrêmement sensible aux modifications temporelles de son environnement. Nous allons dans un premier temps appréhender les différents types de mouvements présents au sein d'une macromolécule. Nous verrons ensuite les modèles qui permettent d'interpréter les mesures de la relaxation des spins nucléaires en terme de mouvements. Nous décrirons enfin l'effet de l'échange chimique sur les observables RMN et la façon dont on peut déterminer les acides aminés présentant de l'échange chimique. Le développement théorique qui suit est essentiellement une synthèse des documents suivants :

- ⇒ Thèse de Géraldine Goussard (Goussard [2000])
- ⇒ Cours Dynamique et Relaxation du DEA "Cristallographie et RMN biologiques" (Carine Van Heijenoort¹⁵)
- ⇒ Revue sur la RMN des mouvements internes dans les protéines (Korzhnev et al. [2001])
- ⇒ Manuel du Logiciel TENSOR2 : Aspects théoriques¹⁶

5.1.1 Mouvements moléculaires

Une macromolécule en solution présente divers degrés de liberté de mouvements internes qui sont liés à la forme de l'énergie potentielle du système. Il peut exister plusieurs conformations de basse énergie avec des barrières d'énergie suffisamment faibles pour qu'une molécule passe d'une forme à l'autre facilement (ex : rotation d'un cycle aromatique), certains noyaux peuvent alors passer d'un site à un autre et d'un déplacement chimique à un autre avec une

¹⁵<http://www.icsn.cnrs-gif.fr/guittet>

¹⁶http://www.ibs.fr/ext/labos/LRMN/softs/tensor/TENSOR2_DOC/theory.html

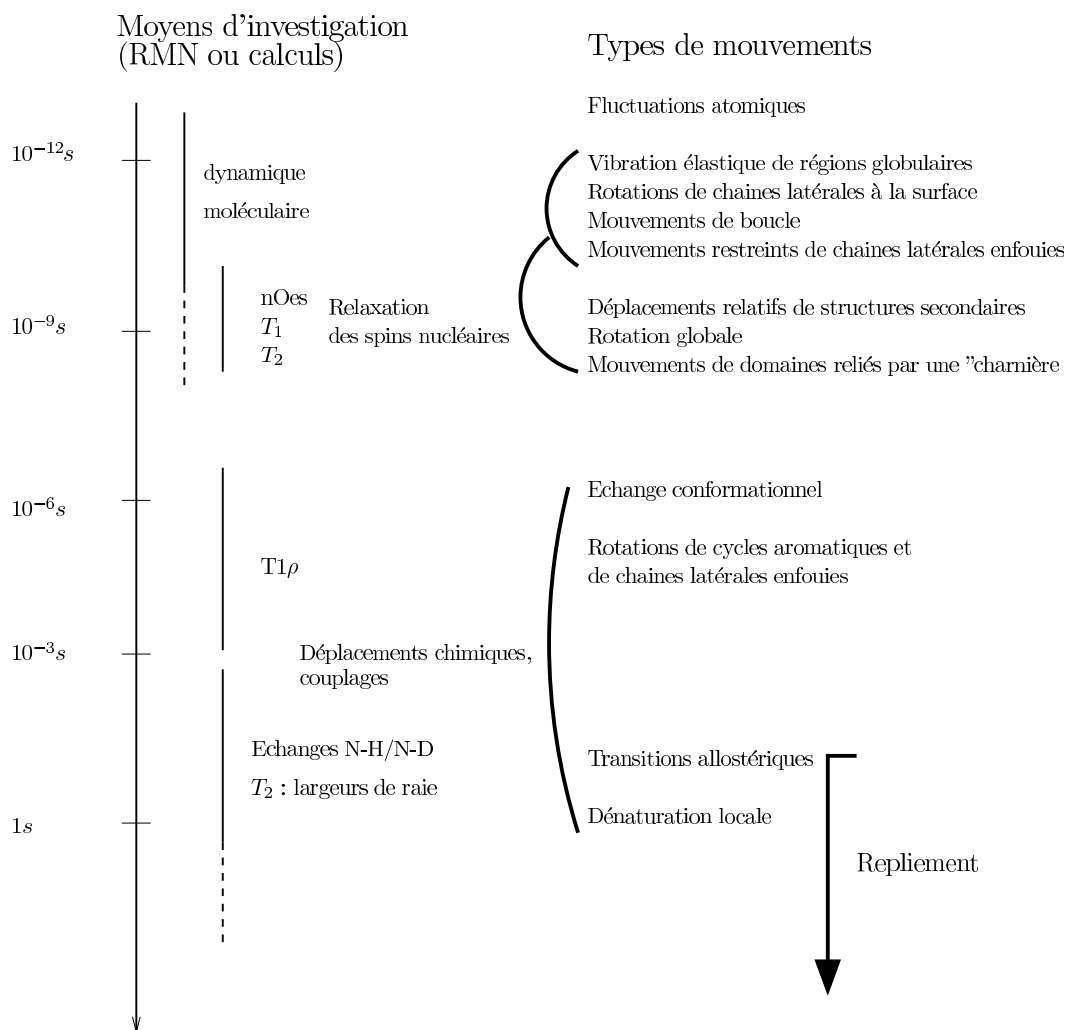


Fig. 5.1: Schéma représentatif de l'échelle des temps des mouvements accessibles par RMN.

vitesse donnée : c'est l'échange chimique. L'échange chimique met en jeu en général des mouvements de l'ordre de $\mu s - ms$. Sur une échelle de temps plus petite ($ps - ns$), les liaisons atomiques disposent d'une liberté de mouvement autour de leurs orientations moyennes. Les motifs structuraux tels que les boucles ont souvent une dynamique plus importante que les structures en hélice α et en feuillet β . La figure 5.1 rassemble les différents types de mouvements et leur échelle de temps. De nombreuses études portent sur l'analyse des corrélations entre la dynamique de la protéine et sa fonction.

5.1.2 La fonction de densité spectrale

Les mouvements atomiques induisent des fluctuations locales du champ magnétique. Par conséquent, les liens entre les mouvements spatiaux et le comportement des spins sont forts. La distribution des mouvements d'une protéine se réorientant en solution se traduit ainsi par une distribution du champ magnétique. Cette répartition est exprimée en terme de densité $J(\omega)$ de

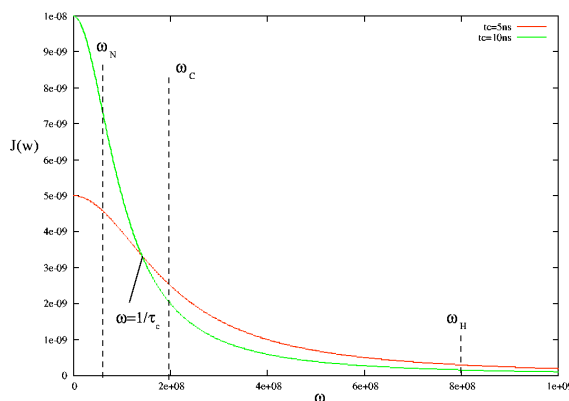


Fig. 5.2: Fonction de densité spectrale d’une molécule en diffusion rotationnelle isotrope sans mouvement interne : c’est une Lorentzienne unimodale. A un champ magnétique donné, plus la protéine étudiée a un poids moléculaire faible, plus le temps de corrélation associé est court et plus la courbe de densité spectrale est étalée.

puissance magnétique disponible à chaque fréquence. Cette fonction est mathématiquement la transformée de Fourier de la fonction d’autocorrélation $G(\tau) = f(t)f(t + \tau)$, qui décrit la ressemblance de la protéine avec elle-même au cours du temps. La fonction de densité spectrale est calculable pour une molécule sphérique rigide en rotation globale isotrope et aléatoire (mouvement brownien) :

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (5.1)$$

Le temps de corrélation τ_c est défini comme la durée que met un vecteur fixe par rapport à la sphère pour balayer un angle de 1 radian.

5.1.3 Mouvement moléculaire et relaxation magnétique

Il est bien connu que le signal détecté en solution par RMN (*free induced decay*) est décroissant. Il correspond au retour du système vers un équilibre thermodynamique après excitation : c’est la relaxation. On distingue la relaxation longitudinale T_1 (parallèle à \vec{B}_0) liée au retour à l’équilibre thermique de la population entre les spins α et β et la relaxation transversale T_2 (perpendiculaire à \vec{B}_0) liée à la perte de la cohérence des états à un quantum entre les différentes molécules de la solution. L’efficacité de la relaxation dépend de la répartition en fréquence des fluctuations d’interaction entre les spins. La relaxation est rapide quand il existe une proportion importante de mouvement à la fréquence de transition des spins (de l’ordre de 100MHz). Par ailleurs, les mouvements lents peuvent induire des déphasages de cohérences pour les composantes de spins situées dans le plan transversale (T_2). La mesure des temps de relaxation en RMN renseigne donc sur la répartition des énergies en fonction des fréquences de mouvements moléculaires, c’est-à-dire sur la fonction de densité spectrale J , aux fréquences de transition du système.

La tendance générale de l’évolution de la relaxation en fonction du poids moléculaire de l’objet étudié est néfaste pour le rapport signal sur bruit : plus l’objet est de taille importante, plus le temps de corrélation est long et plus la relaxation transversale T_2 est efficace. D’après la figure 5.3, la vitesse de relaxation transversale augmente avec le poids moléculaire tandis que la vitesse de relaxation longitudinale commence par baisser puis augmente. La taille critique de l’objet associée au changement de régime dépend du champ magnétique B_0 . Dans les conditions

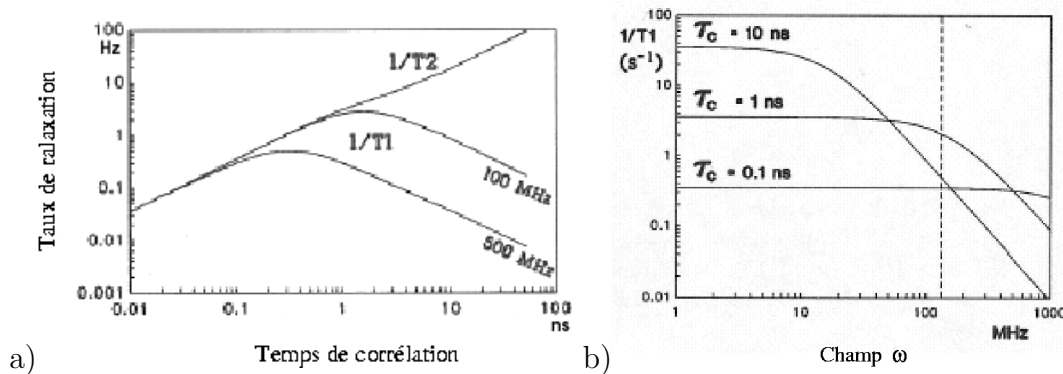


Fig. 5.3: Taux de relaxation $R = 1/T$ en fonction : a) du temps de corrélation rotationnel τ_c à deux champs, b) de la fréquence de Larmor ω_0 pour trois temps de corrélation. (Van deVen)

habituelles d'étude des protéines de poids supérieure à 10kDa, on se situe dans le régime où la taille de l'objet est favorable à T_1 mais défavorable à T_2 . Dans la majorité des expériences de RMN, les spins sont dans le plan transverse (ne serait-ce qu'à l'acquisition) et l'augmentation du poids moléculaire est de toute façon un handicap important.

Effet TROSY Face à cette limitation de taille, ce sont les mécanismes de relaxation eux-mêmes qui apportent une solution. La relaxation cause une perte importante du signal durant les expériences de type HSQC, essentiellement durant la période d'évolution du déplacement chimique. L'opérateur densité est à ce moment proportionnel à $I_z S_x$ (S : spin hétéronucléaire). La relaxation de cet état est essentiellement gouvernée par les termes d'autocorrélation dipôle-dipôle et CSA-CSA. Cependant, l'interférence de ces mécanismes de relaxation (corrélation croisée DD-CSA_{hétéronucléaire}) n'est pas négligeable. On peut profiter des propriétés de ce terme d'interférence. En effet, il contribue à la relaxation des cohérences $I_z S^+$ et $I_z S^-$ par des termes de signes opposés. La relaxation croisée diminue (resp. augmente) la valeur de relaxation transversale des deux cohérences. L'observation de la forme de raies des deux composantes du doublet dans une HSQC non découplée confirme l'importance de la relaxation croisée : un pic fin et intense et un pic large et peu intense. L'effet TROSY (pour revue Pervushin [2000]) utilise cette propriété en sélectionnant la cohérence de relaxation la plus lente. Malgré l'élimination de la moitié de l'aimantation initiale, le gain est important. Les termes d'autorelaxation et d'interférence dépendent de \vec{B}_0 et pour ce qui concerne le spin de l'azote lié à un proton, le terme d'interférence peut compenser le terme d'autocorrélation pour la cohérence $I_z S^+$ à un champ de 900MHz. A la limite, la composante $I_z S^-$ est très large, voire disparaît tandis que l'autre composante mène à une raie très fine et intense. Pour un système carbone aromatique-proton, le CSA du carbone est plus important et le champ optimal pour l'effet TROSY est de 600MHz.

5.1.4 Échantillonnage de la fonction de densité spectrale

Dans l'étude de la dynamique d'une protéine, on se limite en général à l'étude des hétéronoyaux azote et carbone. En effet, l'interprétation de la relaxation des protons est délicate

dans la mesure où la distance entre protons, qui, bien entendu, intervient dans les mécanismes de relaxation, est en général dépendante du temps et où chaque proton d'une protéine a dans son proche environnement plusieurs autres protons, rendant difficile l'extraction des différentes contributions. Par contre, dans les systèmes XH, la relaxation du noyau X est gouvernée en premier lieu par les mouvements du vecteur XH. Si des termes autres que les interactions directes entre X et H (dipôle-dipôle et CSA) interviennent, en particulier les interactions dipolaires entre hydrogènes (ρ_{HH}), ils sont négligeables pour la relaxation de l'azote (1 – 2%). On prend toutefois soin d'étudier la relaxation de l'azote sur une protéine simplement marquée afin d'éliminer les corrélations croisées entre les noyaux azote et carbone. Dans le cas du carbone, il faut également tenir compte de la relaxation croisée entre les carbones et azotes voisins ainsi que de l'effet du couplage scalaire ^{13}C - ^{13}C .

Dans la suite du développement, le spin des hétéronoyaux (^{15}N et ^{13}C) sera noté S et les temps de relaxation T_1 et T_2 sont relatifs au spin S .

Les constantes de relaxation T_1 et T_2 d'un spin S isolé sont respectivement issues des mesures du retour à l'équilibre des aimantations M_z et M_{xy} . En créant les états S_z et S_{xy} , on mesure aisément (voir paragraphes 5.2.1.1 et 5.2.1.2) la décroissance de l'intensité des aimantations $\langle S_z \rangle$ et $\langle S_{xy} \rangle$ par les relations de Bloch ($\langle S_z \rangle_{\text{éq}}$ est l'aimantation suivant l'axe z à l'équilibre thermique) :

$$\begin{aligned} \frac{d}{dt} \langle S_z \rangle &= -R_1 (\langle S_z \rangle - \langle S_z \rangle_{\text{éq}}) & \text{avec } R_1 &= 1/T_1 \\ \frac{d}{dt} \langle S_{xy} \rangle &= -R_2 (\langle S_{xy} \rangle - \langle S_{xy} \rangle_{\text{éq}}) & \text{avec } R_2 &= 1/T_2 \end{aligned} \quad (5.2)$$

soit

$$\begin{aligned} M_z(t) &= M_{z, \text{éq}} + (M_z(0) - M_{z, \text{éq}}) e^{-R_1 t} \\ M_{xy}(t) &= M_{xy, \text{éq}}(0) e^{-R_2 t} \end{aligned} \quad (5.3)$$

Ce modèle monoexponentiel n'est valable que si seuls les mécanismes de relaxation d'auto-corrélation sont actifs. La prise en compte des termes d'interférence de relaxation mènerait à une multi-exponentielle. Les séquences d'impulsions de mesure des temps de relaxation sont optimisées pour limiter l'effet de la relaxation croisée (paragraphe 5.2.1).

Dans un système à deux spins (I et S), chaque état propre relaxe avec une constante de vitesse de relaxation propre à chaque mécanisme de relaxation actif. En supposant que les mécanismes de relaxation n'interfèrent pas entre eux, ces états relaxent de manière indépendante, c'est-à-dire qu'il n'existe pas de mécanisme de relaxation autorisant un transfert de polarisation d'un état vers un autre (ex : $I_+ S_-$ vers $I_+ S_+$). Le transfert I_z vers S_z représente une exception : deux spins interagissant par couplage dipolaire ne relaxent pas indépendamment, la constante de vitesse de cette relaxation croisée est notée σ_{IS} , c'est l'effet Overhauser.

L'évolution phénoménologique de l'aimantation longitudinale d'un spin S en interaction dipolaire avec un spin I est alors gouvernée par l'équation de Solomon ($\langle S_z \rangle_{\text{éq}}$ et $\langle I_z \rangle_{\text{éq}}$ sont les magnétisations des spins S et I à l'équilibre thermodynamique, et $R_{1,I}$ la vitesse de relaxation longitudinale du spin I) :

$$\begin{aligned} \frac{d}{dt} \langle S_z \rangle &= -R_1 (\langle S_z \rangle - \langle S_z \rangle_{\text{éq}}) - \sigma_{IS} (\langle I_z \rangle - \langle I_z \rangle_{\text{éq}}) \\ \frac{d}{dt} \langle I_z \rangle &= -R_{1,I} (\langle I_z \rangle - \langle I_z \rangle_{\text{éq}}) - \sigma_{IS} (\langle S_z \rangle - \langle S_z \rangle_{\text{éq}}) \end{aligned} \quad (5.4)$$

On notera que la relaxation croisée n'intervient pas dans la relation 5.2, car dans l'expérience de mesure de R_1 , le cyclage de phase (mesure alternativement de $+\langle S_z \rangle$ et de $-\langle S_z \rangle$) élimine sa contribution à la décroissance.

Les outils de mécanique quantique appliqués aux mécanismes de relaxation dipôle-dipôle (DD) et d'anisotropie de déplacement chimique (CSA) montrent que les vitesses de relaxation R_1 et R_2 du noyau S et σ_{IS} ainsi que les états mixtes S et I sont reliées à la valeur de la fonction de densité spectrale $J(\omega)$ à cinq fréquences : $0, \omega_S, \omega_I - \omega_S, \omega_H, \omega_H + \omega_S$:

$$\begin{bmatrix} R_1 \\ R_2 \\ \sigma_{IS} \end{bmatrix} = \begin{bmatrix} 0 & 3d+c & d & 0 & 6d \\ 2\frac{3d+c}{3} & \frac{3d+c}{2} & \frac{d}{2} & 3d & 3d \\ 0 & 0 & -d & 0 & 6d \end{bmatrix} \begin{bmatrix} J(0) \\ J(\omega_S) \\ J(\omega_H - \omega_S) \\ J(\omega_H) \\ J(\omega_H + \omega_S) \end{bmatrix} \quad (5.5)$$

$$\text{avec } d = \frac{1}{10} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2 \gamma_S^2 \gamma_H^2}{\langle r_{SH}^3 \rangle^2} \text{ et } c = \frac{2\omega_S^2 (\sigma_{\parallel} - \sigma_{\perp})^2}{15}$$

($\sigma_{\parallel}, \sigma_{\perp}$ composantes parallèles et perpendiculaires du tenseur de déplacement chimique du noyau S , d et c sont des constantes liées respectivement aux interactions dipôle-dipôle et d'anisotropie de déplacement chimique, $d = 1,0389 \cdot 10^8 s^{-2}$ et $c = 1,6613 \cdot 10^8 .s^{-2}$ pour le vecteur NH à 600MHz)

Les différences entre les composantes parallèles et perpendiculaires ($\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp}$) du tenseur de déplacement chimique azote et carbone (de méthyles) valent respectivement $-160ppm$ et $25ppm$. Dans le cas d'une étude de la dynamique des vecteurs CH de méthyles, l'effet de l'anisotropie de déplacement chimique peut être négligé ($\leq 5\%$, Nicholson et al. [1992]).

Les constantes de vitesse de relaxation reliées aux autres états (états à zéro et deux quanta : $2I_z S_z, 2I_z S_+, 2I_- S_+$ et $2I_+ S_+$) sont d'une part difficiles à mesurer expérimentalement, et d'autre part, elles contiennent un terme de corrélation croisée entre le proton amide et les protons voisins, ce qui complique l'interprétation de ces relaxations en terme de mouvements.

Le système 5.5 est clairement indéterminé. La méthode de la "matrice réduite" permet de contourner ce problème. L'observation de la fonction de densité spectrale d'une protéine montre que dans le cas d'un système de spin $^{15}N-^1H$, la courbe est relativement "plate" autour de la fréquence ω_H : $J(\omega_H) \simeq J(\omega_H \pm \omega_S)$ et $J(\omega_H) \ll J(0)$. Dans cette approche, on remplace les trois valeurs de la densité spectrale de cette région par une valeur moyenne $\langle J(\omega_H) \rangle$ (Peng and Wagner [1992]). Le système 5.5 se simplifie :

$$\begin{bmatrix} J(0) \\ J(\omega_S) \\ \langle J(\omega_H) \rangle \end{bmatrix} = \begin{bmatrix} -\frac{3}{4(3d+c)} & \frac{3}{2(3d+c)} & -\frac{9}{10(3d+c)} \\ \frac{1}{3d+c} & 0 & -\frac{7}{5(3d+c)} \\ 0 & 0 & \frac{1}{5d} \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ \sigma_{IS} \end{bmatrix} \quad (5.6)$$

La mesure des constantes de relaxation permet alors d'explorer la fonction de densité spectrale sur trois gammes de fréquences : nulle, moyenne ($\omega_N = 50 - 80MHz$) et haute ($\omega_H = 500 - 800MHz$). La méthode de la matrice réduite n'est valable en général que pour le système de spin $^{15}N-^1H$ car pour le système $^{13}C-^1H$, la différence entre la valeur de la densité spectrale aux fréquences $\omega_H \pm \omega_C$ et à ω_H n'est plus négligeable par rapport à $J(0)$.

5.1.5 L'effet nOe

De manière général, lorsque deux spins sont suffisamment proches dans l'espace (ex. moins de 5Å entre deux protons), la perturbation de l'un influe sur l'évolution de l'autre (effet nOe). Dans le cas des protéine, la mesure de l'effet nOe η consiste à comparer l'intensité du signal issu du spin S en présence et en absence d'irradiation du spin I avant la mesure de d'intensité de $\langle S_z \rangle$. L'irradiation a pour effet d'équilibrer les populations du spin.

A l'état stationnaire, lorsque le spin I est saturé, $\langle I_z \rangle = 0$ et $\frac{d}{dt} \langle S_z \rangle = 0$. Les relations 5.4 donnent alors :

$$\langle S_z \rangle_{sat} = \langle S_z \rangle_{\acute{e}q} + \frac{\sigma_{IS}}{R_1} \langle I_z \rangle_{\acute{e}q} \quad (5.7)$$

où $\langle S_z \rangle_{\acute{e}q}$ et $\langle I_z \rangle_{\acute{e}q}$ représentent l'aimantation des spins S et I à l'équilibre, et $\langle S_z \rangle_{sat}$ l'aimantation du spin S à l'équilibre en présence de l'irradiation de I . L'effet nOe est défini par :

$$\eta = \frac{\langle S_z \rangle_{sat}}{\langle S_z \rangle_{eq}} = 1 + \frac{\gamma_H \sigma_{IS}}{\gamma_S R_1} \quad (5.8)$$

L'effet nOe hétéronucléaire $^1H \rightarrow ^{15}N$ correspondant à la liaison amide est alors limité théoriquement entre 0,834 (τ_c très grand) et $-3,71$ ($\tau_c \rightarrow 0$). L'analyse de l'effet nOe le long de la chaîne peptidique donne alors une estimation qualitative de la dynamique interne de la protéine. Si l'effet nOe est proche de 0,8, le vecteur \overrightarrow{NH} est relativement rigide tandis qu'une valeur de l'effet nOe plus faible indique que ce vecteur est plus mobile.

5.1.6 Effet de l'échange chimique

Vitesse d'échange et déplacement chimique. Les mécanismes d'échange chimique détectables par RMN ont lieu sur l'échelle de temps supérieures à la microseconde. Si la vitesse d'échange entre les différentes formes est inférieure à la différence de déplacements chimiques (échange lent), chaque site donnera lieu à des signaux de déplacements chimiques distincts. Si cette vitesse est beaucoup plus grande (échange rapide), un seul signal sera observé à la moyenne des déplacements chimiques, pondérée par la durée de vie dans chaque forme. Les effets de relaxation sont alors également moyennés et l'échange induit un mécanisme supplémentaire de relaxation. Un équilibre intermédiaire correspond à une vitesse d'échange de l'ordre de grandeur de la différence des déplacements chimiques.

Echange chimique et relaxation Considérons le cas d'un échange chimique rapide d'un noyau S entre deux sites A et B, en équilibre avec des vitesses d'interconversion k_1 et k_{-1} :



Si $\delta\omega_{ex}$ dénote la différence de déplacements chimiques entre les deux formes ($\delta\omega_{ex} = \omega_A - \omega_B$), l'échange intervient dans la fonction de densité spectrale par $J_{ex}(\omega)$ (Deverell et al.

[1970]) :

$$J_{ex}(\omega) = \frac{(k_1 + k_{-1})^2}{k_1 k_{-1}} (\delta\omega_{ex})^2 \frac{\tau_{ex}}{1 + \omega^2 \tau_{ex}^2} \quad (5.10)$$

avec $\tau_{ex} = \frac{1}{k_1 + k_{-1}}$, durée de vie réduite de l'échange.

On note k_{ex} la vitesse de l'échange : $k_{ex} = k_1 + k_{-1} = \frac{k_1}{p_B} = \frac{k_{-1}}{p_A} \cdot p_A$ et p_B sont les populations des formes A et B. La densité spectrale associée à l'échange chimique est donc une lorentzienne dont le point d'inflexion est à $1/\tau_{ex}$.

La durée de vie réduite minimale que l'on peut mesurer par RMN est de l'ordre de la microseconde, soit bien supérieure à $1/\omega_0$. La contribution de l'échange à la fonction de densité spectrale à fréquence élevée (de l'ordre de la fréquence de Larmor) est donc négligeable. Par contre, à fréquence nulle, la contribution n'est plus négligeable : $J_{ex} = J_{ex}(0) = \frac{k_1 + k_{-1}}{k_1 k_{-1}} (\delta\omega_{ex})^2$. L'effet de l'échange chimique n'est visible dans la densité spectrale qu'à la fréquence nulle. Il s'ajoute au terme de relaxation par les mouvements rapides. Parmi les constantes de vitesse de relaxation, $J(0)$ n'apparaît que dans R_2 . La présence d'échange chimique sur un résidu se traduit par un terme supplémentaire dans R_2 ($R_2 = R_2 + R_{ex}$, cas de la précession libre) avec :

$$R_{ex} = \frac{4\pi (\delta\omega_{ex})^2 p_A p_B \tau_{ex}}{1 + \omega^2 \tau_{ex}^2} \quad (5.11)$$

Ceci traduit également le fait que l'échange n'induit pas de transition α - β mais un déphasage des cohérences dans le plan transverse. Il est donc sans effet sur R_1 et ne joue que sur R_2 . L'identification de la présence d'échange chimique passe donc par la mesure de R_2 . Une valeur de R_2 anormalement élevée par rapport à R_1 (figure 5.3) peut être alors attribuée à de l'échange chimique.

Il existe d'autres méthodes utilisant la résonance magnétique nucléaire pour démontrer l'existence d'échange chimique. Elles ont été mises au point dans l'objectif de balayer les différents ordres de grandeur de vitesse d'échange :

- ↔ $\mu s - ms$: échange ZZ hétéronucléaire, relaxation $R_{1\rho}$,
- ↔ ms : séquence d'écho de spin CPMG,
- ↔ $s - h$: échange $^1H/{}^2H$ des protons amides.

Nous ne détaillerons que la séquence CPMG que nous avons mis en oeuvre pour mettre en évidence la présence d'échange chimique au sein de certains résidus d'ASP2 (5.2.1.4).

5.1.7 Interprétation de la fonction de densité spectrale en terme de mouvement

La mesure de R_1 , R_2 et η est utilisée pour déterminer la valeur de la fonction de densité spectrale aux fréquences 0, ω_S et ω_H et donc la répartition en fréquences des mouvements d'un vecteur. Si le vecteur étudié appartient à une protéine, l'évolution des trois valeurs de densité spectrale le long de la séquence peptidique nous renseigne déjà sur les résidus ayant un comportement dynamique original. Cependant, il est plus confortable de travailler sur des mouvements et pour cela, il est indispensable de proposer des modèles de mouvements pour recalculer les fonctions de densité spectrale. Nous nous contenterons de citer le modèle *model-free*.

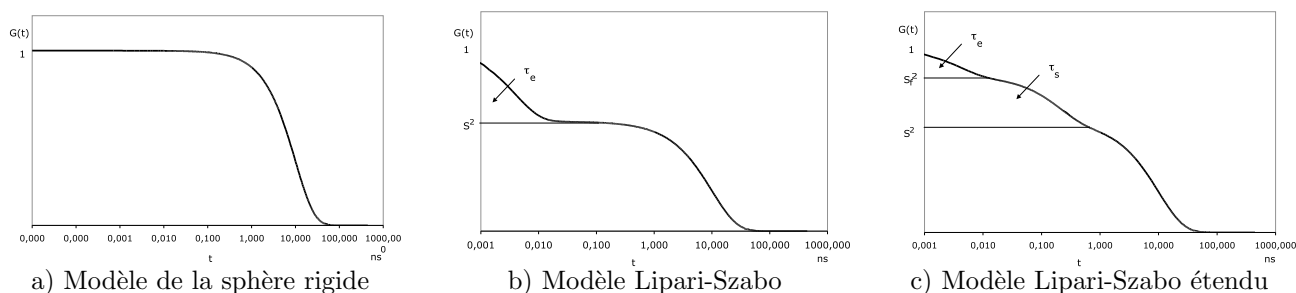


Fig. 5.4: La fonction d'autocorrélation G des trois modèles (l'échelle en abscisse est logarithmique).

5.1.7.1 L'analyse des données par *model-free*

Le modèle le plus couramment employé à l'heure actuelle est le modèle de Lipari-Szabo, aussi appelé *model-free* (Lipari and Szabo [1982a,b]). G. M. Clore et ses collègues l'ont amélioré en introduisant les mouvements internes lents (*model-free* étendu, Clore et al. [1990a]). Ces modèles se basent en première approche sur le mouvement d'une molécule sphérique rigide en rotation globale isotrope (temps de corrélation τ_c), puis les mouvements internes sont pris en compte par un paramètre d'ordre généralisé S^2 et du temps de corrélation τ_e (*model-free*). Un troisième temps de corrélation et un second paramètre d'ordre associés aux mouvements internes intermédiaires sont introduits quand leurs effets ne sont plus négligeables en terme de fonction de d'autocorrélation $G(\tau)$ (*model-free* étendu).

Rotation globale isotrope. Pour une molécule rigide animée d'un mouvement brownien isotrope, les vecteurs intramoléculaires $\overrightarrow{X\hat{H}}$ d'une protéine se réorientent de la même façon en suivant le temps de corrélation global τ_c de la protéine. La fonction d'auto-corrélation de la protéine est une monoexponentielle de paramètre τ_c (figure 5.4a) et la fonction de densité spectrale s'exprime simplement :

$$J_{rigide}(\omega) = \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (5.12)$$

Model-free (Lipari and Szabo [1982b,a]) Pour tenir compte des mouvements internes et donc de la dynamique du squelette peptidique ou des chaînes latérales, on considère que la décorrélation s'effectue en deux étapes. La première étape permet au vecteur d'explorer l'espace de ses orientations possibles. Cela se traduit par une décorrélation rapide (figure (5.4b)) jusqu'à un plateau à S^2 . Le temps de corrélation τ_e est la surface balayée jusqu'à S^2 . S^2 est le paramètre d'ordre qui reflète l'amplitude des mouvements. L'interprétation du temps de corrélation τ_e est délicate et en général, il n'est pas directement traduit par une représentation des mouvements rapides. Cette durée dépend du type de mouvement considéré et peut varier de quelques picosecondes pour une rotation d'un groupe méthyle, à quelques nanosecondes pour les mouvements des vecteurs $\overrightarrow{N\hat{H}}$ dans les boucles et les coudes. Le paramètre d'ordre S^2 traduit l'amplitude des mouvements internes.

Ces mouvements restent suffisamment restreints pour que la relaxation soit gouvernée en premier lieu par la réorientation globale de la protéine. Aucune hypothèse n'est nécessaire a

priori sur le type de mouvement (diffusion dans un cône, sur un cône, ...). Les mouvements internes des vecteurs \overrightarrow{XH} provoquent une décorrélation rapide de la fonction d'auto-corrélation d'un facteur d'ordre S^2 . Ce paramètre d'ordre généralisé peut être interprété comme étant l'amplitude de ces mouvements. Un paramètre d'ordre proche de 1 signifie que le vecteur \overrightarrow{XH} est fortement restreint autour de son orientation moyenne dans la molécule, elle-même en rotation. Par contre, un paramètre d'ordre très inférieur à 1 indique que le vecteur balaie un angle solide relativement grand. Si $S^2 = 0$, le vecteur \overrightarrow{XH} peut alors balayer toutes les orientations de manière équiprobable. La fonction de densité spectrale construite par Lipari et Szabo prend en compte les deux types de mouvements :

$$J_{LP}(\omega) = S^2 \frac{\tau_c}{1 + (\omega\tau_c)^2} + (1 - S^2) \frac{\tau'_e}{1 + (\omega\tau'_e)^2} \quad (5.13)$$

avec $\frac{1}{\tau'_e} = \frac{1}{\tau_c} + \frac{1}{\tau_e}$

Ils sont caractérisés par des temps de corrélation très différents. τ_e et τ_c valent typiquement 50ps et 10ns pour une protéine de poids moléculaire de l'ordre de la dizaine de kDa :

Sous la condition $\tau_e \ll \tau_c$ (mouvements internes très rapides), la fonction de densité spectrale peut être simplifiée en perdant un paramètre ajustable :

$$J_{LP}(\omega) = S^2 \frac{\tau_c}{1 + (\omega\tau_c)^2} \quad (5.14)$$

Model-free étendu (Clare et al. [1990a,b]) Pour certains résidus en boucle, dans un coude β , à la surface de la protéine ou à l'extrémité de brins β , le modèle de Lipari-Szabo n'est plus valable. La complexité des mouvements internes ne peut plus être décrite par un seul palier à S^2 . Clare et ses collègues ont ajouté un second palier associé à un second corrélation interne afin de tenir compte des mouvements intermédiaires (τ_s de 100ps à quelques nanosecondes, paramètre d'ordre S_s^2 , $S^2 = S_s^2 S_f^2$) afin d'expliquer les mesures expérimentales. Les mouvements sur l'échelle de temps de la picoseconde sont décrits par τ_f et S_f^2 . Les mouvements internes rapides et intermédiaires sont considérés comme décorrélés. Les mouvements intermédiaires peuvent être par exemple associés à un saut entre deux sites et les mouvements rapides à la diffusion du vecteur autour de ces deux sites. Cette hypothèse de travail mène à une nouvelle fonction de densité spectrale à cinq paramètres (τ_c , S^2 , τ_s , S_f^2 et τ_f) :

$$J_{\text{étendu}}(\omega) = S^2 \frac{\tau_c}{1 + (\omega\tau_c)^2} + (S_f^2 - S^2) \frac{\tau'_s}{1 + (\omega\tau'_s)^2} + (1 - S_f^2) \frac{\tau'_f}{1 + (\omega\tau'_f)^2} \quad (5.15)$$

avec $\frac{1}{\tau'_s} = \frac{1}{\tau_c} + \frac{1}{\tau_s}$, $\frac{1}{\tau'_f} = \frac{1}{\tau_c} + \frac{1}{\tau_f}$ et $S^2 = S_f^2 S_s^2$

Sous l'hypothèse $\tau_f \rightarrow 0$, cette fonction se simplifie et ne dépend plus que quatre paramètres (τ_c , τ_s , S^2 , S_f^2) :

$$J_{\text{étendu}}(\omega) = S^2 \frac{\tau_c}{1 + (\omega\tau_c)^2} + (S_f^2 - S^2) \frac{\tau'_s}{1 + (\omega\tau'_s)^2} \quad (5.16)$$

Le modèle de Lipari-Szabo est inclus dans celui-ci dans le cas où $S_f^2 = 1$ et $\tau_f = 0$ (mouvements internes rapides sans effet sur la relaxation) ou $S^2 = S_f^2$. On identifie alors $\tau_s = \tau_e$.

Les données expérimentales peuvent donc répondre à plusieurs modèles contenant un nombre différents de paramètres ajustables. Les logiciels Tensor2 et Modelfree utilisent les 5 modèles suivants (par ordre croissant du nombre de paramètres ajustables, y compris le terme d'échange R_{ex}) avec la même numérotation :

Modèle	Hypothèse(s)	Paramètres
1. <i>Lipari-Szabo simplifié</i>	τ_e est rapide ($\tau_e < 20ps$)	S^2
2. <i>Lipari-Szabo classique</i>	τ_e est actif sur la relaxation	S^2 et τ_e
3. <i>Lipari-Szabo simplifié avec échange</i>	Comme au 1. et présence d'échange	S^2 et R_{ex}
4. <i>Lipari-Szabo classique avec échange</i>	Comme au 2. et présence d'échange	S^2, τ_e et R_{ex}
5. <i>Lipari-Szabo étendu</i>	2 types de mouvements : S_s^2, τ_s (lent) et S_f^2, τ_f (rapide) et τ_f est négligeable	S_f^2, S_s^2 et $\tau_e = \tau_s$

Tab. 5.1: Définition des modèles utilisés par Tensor2 et model-free.

La prise en compte du terme d'échange dans le modèle étendu amène le système à quatre inconnues pour trois mesures. La mesure à deux champs double le nombre de mesures. La dépendance en champ de R_1 , R_2 et de R_{ex} étant différente, il devient alors possible de reconnaître les résidus présentant à la fois une dynamique interne à deux temps de corrélation et de l'échange chimique. Dans le cas d'ASP2, nous nous sommes satisfaits des mesures à un seul champ. Seuls trois paramètres restent simultanément accessibles.

5.1.7.2 Quel modèle est choisi par le protocole ?

J'ai utilisé deux logiciels Tensor2¹⁷ (Dosset et al. [2000]) et ModelFree¹⁸ associé à FAST-Modelfree¹⁹ (Cole and Loria [2003]) pour modéliser les mouvements.. Ces deux approches s'appuient sur le protocole publié par Mandel (Mandel et al. [1996]), brièvement décrit au paragraphe suivant. D'une utilisation simple (un minimum de paramètres de calculs à ajuster), ils sont limités dans la mesure où ils restent des "boîtes noires". Pour valider les résultats, j'ai choisi de comparer l'analyse de la dynamique en utilisant à la fois les deux logiciels. Il leur suffit de connaître les valeurs de R_1 , R_2 et η avec leurs incertitudes expérimentales pour déterminer les paramètres de mobilité de chaque résidu.

Algorithme du choix du modèle. Les modèles définis au paragraphe précédent comportent à chaque fois des degrés de complexité supplémentaires. Il est possible d'affiner le modèle en ajoutant autant de temps de corrélation que souhaité. Cependant, les données expérimentales et leurs incertitudes limitent le nombre de paramètres de mouvement accessibles. En effectuant un jeu de mesures des trois constantes R_1 , R_2 et σ_{IS} , on calcule directement le

¹⁷<http://www.ibs.fr/ext/labs/LRMN/softs/>

¹⁸<http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/modelfree.html>

¹⁹<http://xbeams.chem.yale.edu/~loria/software.htm>

triplet $J(0)$, $J(\omega_N)$ et $J(\omega_H)$. Le modèle étendu, qui fournit le plus de renseignements, pourrait intuitivement être appliqué automatiquement mais il faut valider la signification physique des paramètres extraits. Il faut donc trouver le modèle décrivant suffisamment les données mais sans leur faire dire plus de choses qu'elles ne contiennent,

En général, lorsqu'un phénomène physique peut répondre de plusieurs modèles théoriques, on utilise les statistiques pour déterminer quel modèle est valide et lequel permet d'expliquer le phénomène en introduisant un minimum de paramètres. On peut toujours ajouter un paramètre ajustable mais le nouveau paramètre n'a plus statistiquement de sens. Pour cela, on estime la probabilité qu'un modèle satisfasse à l'expérience par la détermination d'une fonction cible χ^2 :

$$\chi^2 = \sum_{i=1}^N \frac{(V_{i,calc} - V_{i,exp})^2}{\sigma_{i,exp}^2} \quad (5.17)$$

avec N : nombre de points, V_i : valeur expérimentale ou calculée, $\sigma_{i,exp}$: incertitude sur la mesure

La grandeur χ^2 , ne dépendant pas du nombre de paramètres, décrit la distance entre les valeurs expérimentales et les valeurs calculées à partir du modèle. Plus il est faible, plus l'expérience est décrite de façon correcte par le modèle. L'adéquation d'un modèle est alors évaluée soit par un test du χ^2 où le minimum de χ^2 est évalué en fonction du nombre de paramètres du modèle, soit par Monte-Carlo. L'ajout d'un paramètre diminue automatiquement le χ^2 au-delà d'une valeur statistiquement significative, compte tenu des erreurs expérimentales. La détermination de cette limite, et donc du modèle qui sera choisi en fin d'algorithme, est gouvernée par la statistique du F, qui valide l'ajout statistiquement significatif d'un paramètre par rapport au nombre de variables et aux erreurs expérimentales.

La première étape de la procédure consiste à partir du modèle le plus simple. On détermine les paramètres ajustables en minimisant la fonction χ^2 par l'algorithme Levenberg-Marquardt. L'erreur sur les paramètres est obtenue par une simulation Monte-Carlo : on génère un jeu de données pseudo-expérimentale à partir des données mesurées et de l'incertitude de la mesure (on fait l'hypothèse d'une distribution normale), on calcule ensuite les paramètres correspondants à ce jeu de données. L'erreur sur les paramètres ajustés est calculée sur la base de la distribution des paramètres ainsi recalculés. Chaque jeu de données aléatoires correspond alors à un χ_{th}^2 . Le modèle est jugé acceptable dans un intervalle de confiance de 95% si 5% des χ_{th}^2 est supérieur à χ_{exp}^2 .

On applique ensuite cette méthode successivement aux autres modèles jusqu'à ce que le modèle soit acceptable dans un intervalle de confiance de 95% et que la statistique des F indique que l'ajout d'un paramètre n'est pas statistiquement justifié.

Il faut noter que les modèles 4 et 5 ont trois paramètres ajustables, le système n'a plus de degré de liberté interne et le test-F ne peut être utilisé. Ces modèles sont acceptés dans la mesure où le χ^2 est suffisamment proche de zéro. Dans le cas contraire, Tensor2 (ou modelfree) rejette les cinq modèles mais propose tout de même les valeurs de mobilité du meilleur des cinq modèles.

Il faut noter que les modèles 2 et 3 sont mathématiquement très proches et le choix entre l'introduction du terme d'échange ou d'un temps de corrélation τ_e peut s'avérer discutable. Si l'un des deux modèles est choisi, cela traduit simplement que l'ajustement des données nécessite

un terme supplémentaire, qui peut correspondre physiquement soit à de la dynamique interne rapide importante, soit à de la présence d'échange chimique. Les expériences de type CPMG permettent de lever en partie l'ambiguïté.

Les logiciels donnent en sortie après analyse un fichier contenant le modèle choisi, la valeur de χ^2 minimal, S_f^2 , S_s^2 , τ_e (τ_s pour le modèle 5), R_{ex} ainsi que les incertitudes associées pour chaque résidu. La colonne PO du fichier de sortie de Tensor2 correspond à S_s^2 et la colonne P1 à S_f^2 . S^2 se déduit par la relation : $S^2 = S_f^2 S_s^2$.

Détermination de τ_c Le lecteur aura remarqué que le temps de corrélation global τ_c n'est pas utilisé ici comme un paramètre ajustable. Il est extrait préalablement du rapport R_2/R_1 sur les résidus présentant le moins de flexibilité (situés *a priori* dans les structures secondaires) et non impliqués dans un phénomène d'échange chimique (Cloue et al. [1990a], Lee et al. [1997]). En effet, en considérant que pour ces résidus, le temps de corrélation des mouvements rapides (τ_e) est négligeable, la fonction de densité spectrale se réduit à :

$$J(\omega) = \frac{S^2 \tau_c}{1 + \omega^2 \tau_c^2} \quad (5.18)$$

Le rapport R_2/R_1 ne dépend alors plus de τ_e , ni de S^2 .

Une simulation Monte-Carlo évalue l'incertitude sur τ_c et valide également si le modèle de rotation globale isotrope est acceptable dans l'intervalle de confiance de 95%. Si le modèle de rotation isotrope est valable, alors le temps de corrélation τ_c est supposé identique pour tous les résidus, mêmes ceux présentant une forte mobilité ou/et de l'échange. On notera que si la réorientation de la protéine n'est pas isotrope, il faut prendre en compte 3 temps de corrélation pour définir le tenseur de diffusion orientationnelle. La valeur de τ_c calculée pour un résidu par le rapport R_2/R_1 dépend alors de l'orientation du vecteur. Ainsi, hors échange chimique et dynamique interne importante, de grandes variations de ce rapport au sein de la séquence indiquent une réorientation anisotrope. Ceci est particulièrement visible pour une structure dont les hélices α pointent dans des directions très différentes.

La valeur de τ_c calculée est à mettre en lumière avec celle dérivée de l'équation d'hydrodynamique de Stoke-Einstein (5.19) associée à une molécule sphérique de volume V (Palmer III [1996]).

$$\tau_c = \frac{\eta^* V}{k_B T} \quad (5.19)$$

avec $V = \frac{M}{\rho} (l.mol^{-1}) = \frac{10^{21}}{6,02 \cdot 10^{23}} \frac{M}{\rho} (nm^3)$
 η^* : viscosité du milieu, k_B : constante de Boltzmann, T : température

On considère en général qu'une protéine dispose à sa surface d'une monocouche d'hydratation de 0,32nm. Le volume spécifique de la protéine est 0,73cm³/g. De la masse de la protéine, on calcule son volume. Le volume est ensuite corrigé pour tenir compte de la couche d'hydratation (Palmer III [1996]).

Pour estimer la viscosité de l'eau à différentes températures, j'ai utilisé la relation suivante (Atkins [1982]) :

$$\log \left(\frac{\eta_{20}}{\eta_T} \right) = \frac{1,37023(T_{\circ C} - 20) + 8,36 \cdot 10^{-4}(T_{\circ C} - 20)^2}{109 + T_{\circ C}} \quad (5.20)$$

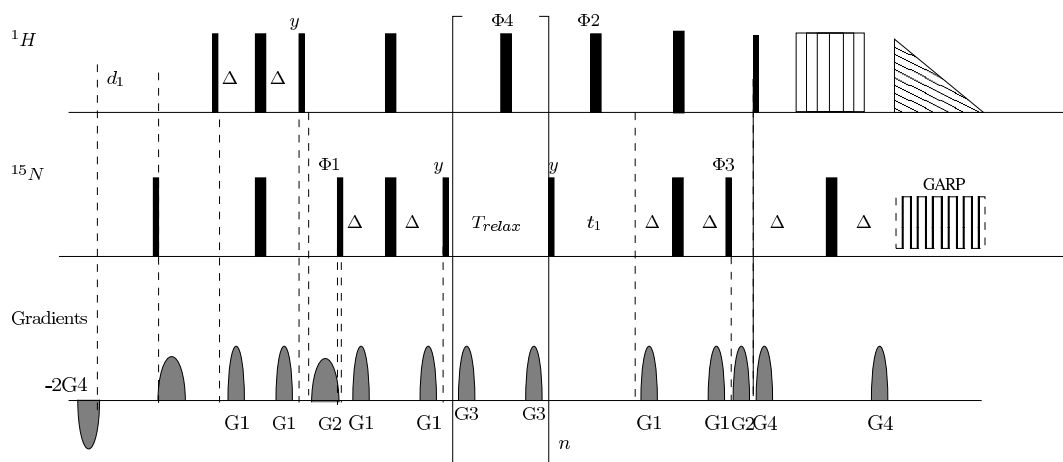


Fig. 5.5: Séquence d'impulsions pour la mesure de R_1 . Les impulsions sont par défaut suivant l'axe x . Le cyclage de phases est le suivant : $\phi_1 = x, -x$; $\phi_2 = x, x, -x, -x$; $\phi_3 = x, x, x, x, -x, -x, -x, -x$; $\phi_{acq} = x, -x, x, -x, -x, x, -x, x$. La détection en quadrature dans la dimension indirecte est assurée par le States-TPPI.

avec $\eta_{20} = 10^{-3} \text{kg.m}^{-1}.\text{s}^{-1}$ la viscosité de l'eau à 20°C

Ceci nous donne donc à 308K ($T = 34,85^\circ\text{C}$) une viscosité de l'eau pure de $\eta_{308\text{K}} = 0,867 * 10^{-3} \text{kg.m}^{-1}.\text{s}^{-1}$ et à 283K , $\eta_{283\text{K}} = 1,12 * 10^{-3} \text{kg.m}^{-1}.\text{s}^{-1}$.

5.2 Mise en oeuvre expérimentale : mesures de R_1 , R_2 et de l'effet nOe

5.2.1 Les expériences

5.2.1.1 Mesure de R_1

Séquence d'impulsions L'idée centrale de la séquence d'impulsions (figure 5.5) utilisée pour mesurer la relaxation longitudinale de l'azote, est de créer un état S_z hors équilibre puis de le laisser relaxer pendant T_{relax} . En variant le délai T_{relax} , on échantillonne la courbe de décroissance de l'intensité de la résonance étudiée.

On obtient l'état S_z en partant du spin I_z de l'hydrogène amide par transfert INEPT refocalisé : le premier transfert INEPT ($\Delta=2.6\text{ms}$) crée l'état antiphase $I_x S_z$. Le second INEPT refocalise vers l'état en phase S_z . On laisse ensuite cet état relaxer tout en inversant régulièrement l'aimantation proton par un 180° pour détruire le mécanisme de relaxation croisée DD-CSA. Le délai entre deux inversions est fixé à 6ms . Le déplacement chimique azote est édité puis on retourne vers l'hydrogène par la succession des deux INEPT retour. La suppression du signal de l'eau est effectuée par une watergate. Le délai entre deux FID doit être suffisamment grand ($d_1 = 4\text{s}$) pour que l'aimantation initiale proton soit à l'équilibre au début de chaque passage. La première impulsion azote suivie d'un gradient détruit l'aimantation ini-

tiale portée par l'azote pour que l'état S_z dont on mesure la relaxation provienne uniquement de l'hydrogène.

Le cyclage de phase est tel que l'effet nOe est compensé et que l'aimantation pour T_{relax} infini soit nulle.

Extraction des R_1 Le module *Rate Analysis* du logiciel NmrView a été développé afin d'extraire les paramètres d'une équation à partir d'un jeu de données expérimentales. L'équation de la relaxation que j'ai utilisée est une simple monoexponentielle à deux paramètres ajustables T et $I(0)$:

$$I(\delta) = I(0)e^{-\frac{T_{relax}}{T}}$$

avec $I(\delta)$: intensité du pic considéré au délai de relaxation δ .

On peut également choisir d'ajouter une constante C qui peut refléter une ligne de base imparfaite. La possibilité d'ajuster les données à une monoexponentielle justifie *a posteriori* l'élimination des mécanismes de relaxation croisée dans la séquence.

L'avantage principal du module *Rate Analysis* est d'analyser très rapidement un grand nombre de spectres à différents délais de relaxation δ . Il suffit du peak-picking et de l'attribution d'un seul spectre (délai le plus petit) pour extraire l'intensité des pics des autres spectres. L'intensité des pics est calculée sur la base du point d'intensité le plus élevée situé dans une limite de $\pm 25\%$ de la taille de la boîte autour du centre de la boîte (*Jitter*). Cette procédure est très efficace car elle autorise de faibles déplacements de fréquences d'un spectre à l'autre. L'incertitude sur l'intensité est déterminée par NmrView par calcul de l'écart-type du bruit sur une boîte ne contenant pas de signal. D'après Palmer III et al. [1991b], les taux de relaxation sont estimés avec une plus grande précision en utilisant l'intensité des pics au lieu du volume.

Pour chaque résidu, la valeur de l'intensité est représentée en fonction du délai δ . On définit ensuite l'équation à utiliser dans la procédure de minimisation. J'ai choisi de déterminer l'incertitude sur les paramètres extraits en utilisant la simulation Monte-Carlo décrite au paragraphe 5.1.7.2. Le nombre de simulations a été fixé à 200. Les 200 jeux de paramètres ainsi calculés suivent théoriquement une distribution gaussienne. J'ai choisi les extrémités de l'intervalle de confiance à 65%, c'est-à-dire les valeurs extrémales des 2*65 jeux de paramètres les plus proches des paramètres calculés à partir des intensités mesurées.

La procédure de minimisation est automatisée sur l'ensemble des résidus. Il convient alors de vérifier graphiquement les résultats, en particulier, l'extraction des intensités et la forme de l'exponentielle.

Il convient de noter que NmrView travaille par défaut sur T_1 et non R_1 . Le transfert vers Tensor2 nécessite donc de transformer les données en taux de relaxation. L'incertitude sur R_1 est alors estimée par :

$$\Delta R_1 = \frac{\Delta T_1}{T_1^2}$$

5.2.1.2 Mesure de R_2

Séquence d'impulsions La séquence d'impulsions (figure 5.6) construite pour mesurer la relaxation transversale est quasiment identique à celle étudiée précédemment. La différence se

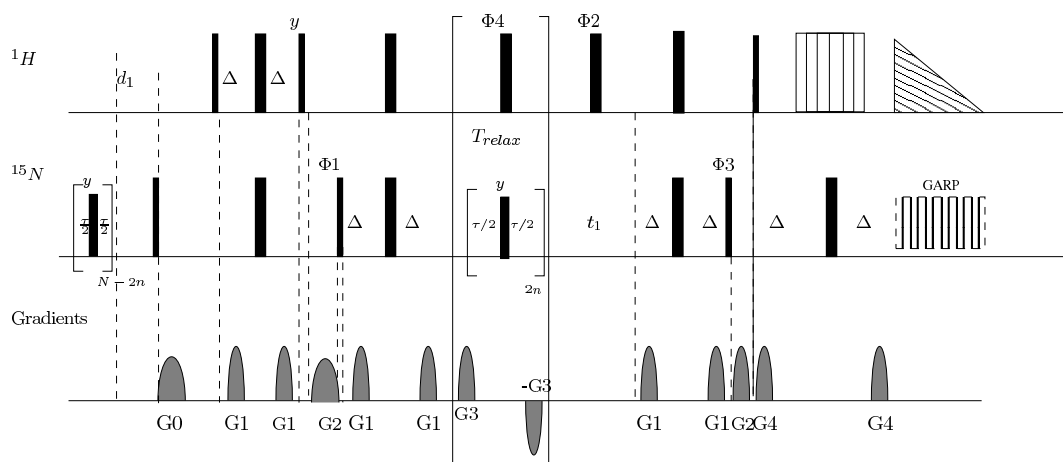


Fig. 5.6: Séquence d'impulsions pour la mesure de R_2 . Les impulsions sont par défaut suivant l'axe x . Le cyclage de phases est le suivant : $\phi_1 = x, -x$; $\phi_2 = x, x, -x, -x$; $\phi_3 = x, x, x, x, -x, -x, -x, -x$; $\phi_{acq} = x, -x, x, -x, -x, x, -x, x$. La détection en quadrature dans la dimension indirecte est assurée par le States-TPPI.

situé au milieu de la séquence. Au lieu de basculer l'aimantation S_x vers S_z , nous restons sur cet état transverse de phase S_x . Nous appliquons ensuite un train d'impulsion CPMG (voir paragraphe 5.2.1.4) sur l'azote pour minimiser les effets dus à l'échange chimique, l'inhomogénéité du champ et l'échange d'aimantation à zéro-quantum. L'intervalle entre les impulsions de refocalisation du train CPMG est 1ms. Ce délai est suffisant pour conserver l'aimantation transversale azote suivant l'axe x . Le CPMG détruit ainsi la composante de relaxation due à l'échange chimique entre des états de durée de vie supérieure à environ 10ms (Clare et al. [1990a]). La première série d'impulsions azote est hors-résonance et a pour effet l'homogénéisation de la température sur les différentes durées de T_{relax} . Les impulsions d'inversion de l'hydrogène réduisent l'aimantation de l'eau, refocalisent le couplage scalaire HN et limitent la corrélation croisée DD-CSA. Le retour sur l'hydrogène se fait comme pour la mesure de R_1 .

Extraction des R_2 La procédure que j'ai suivie pour extraire la constante R_2 de la série de spectres est exactement la même que pour extraire la valeur et l'incertitude sur R_1 .

5.2.1.3 Mesure de l'effet nOe η

Séquence d'impulsions L'introduction de l'effet nOe au paragraphe 5.1.5 laisse clairement entendre que la mesure de η impose deux expériences : une expérience de référence et une expérience pendant laquelle les protons amides sont préalablement saturés. La figure 5.7 détaille la séquence d'impulsions de ces deux expériences. L'expérience appelée *sat* commence par la saturation pendant le délai d'équilibration d_1 tandis que l'expérience *nonsat* laisse le système s'équilibrer naturellement. Avant la première impulsion azote, l'aimantation S_z est à l'équilibre : $S_{z, \text{éq}}$ dans l'expérience *nonsat* et $S_{z, \text{sat}}$ dans l'expérience *sat*. Le rapport des deux intensités détermine la valeur de η (relation 5.8). La suite des deux expériences est identique : édition du déplacement chimique de l'azote et refocalisation du couplage HN, module INEPT

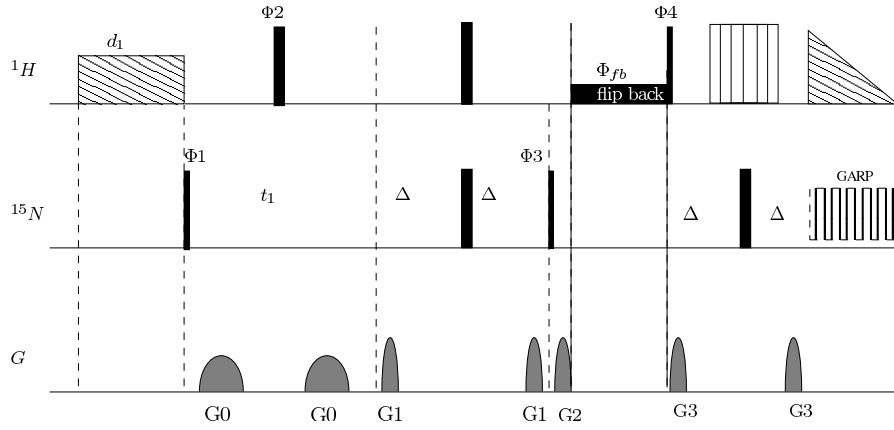


Fig. 5.7: Séquence d'impulsions pour la mesure de l'effet nOe hétéronucléaire 1H - ^{15}N . Les impulsions sont par défaut suivant l'axe x . Le cyclage de phases est le suivant : $\phi_1 = x, -x$; $\phi_2 = 8x, 8(-x)$; $\phi_3 = 2y, 2(-y)$; $\phi_4 = 4x, 4(-x)$; $\phi_{fb} = 4(-x), 4x$; $\phi_{acq} = x, -x, -x, x, -x, x, x, -x$. La détection en quadrature dans la dimension indirecte est assurée par le States-TPPI. Le délai Δ est fixé à 2,6ms. L'impulsion flip-back avant l'impulsion de lecture est calibrée pour une durée de 1ms. La saturation des protons pendant d_1 n'est effectuée que pour l'expérience *sat*, elle est absente pour l'expérience *nonsat*.

pour transférer la polarisation sur le proton amide (état antiphase $I_z S_y$), transfert vers l'état $I_y S_z$ et impulsion flip-back sélectif sur l'eau (durée de 1ms), refocalisation vers un état en phase I_x associée à une watergate puis acquisition du proton.

Le délai d_1 est suffisamment long pour l'établissement de l'équilibre avec et sans saturation (au minimum 4s). Il est gouverné par les relaxations longitudinales de l'azote et de l'hydrogène.

Les deux expériences sont enregistrées de manière enchevêtrée : chaque passage correspond à un passage des deux expériences. Les conditions des deux expériences (température, dérive du shim,...) sont ainsi aussi proches que possibles.

La sensibilité de la mesure de l'effet nOe hétéronucléaire est faible par rapport à la mesure de R_1 et de R_2 car on ne peut bénéficier en début de séquence d'un transfert de polarisation $^1H \rightarrow ^{15}N$. De plus, le délai d_1 allonge l'enregistrement du spectre de manière significative.

Extraction de η Le module HetNOE de NmrView a été développé afin d'automatiser la mesure de l'effet nOe η . Pour cela, il faut disposer des expériences saturée et non saturée. L'expérience non saturée contenant les intensités I_{eq} les plus importantes (car $\eta \leq 0,83$) est "peak-pickée" et attribuée. Le module mesure ensuite l'intensité I_{sat} des pics dans l'expérience saturée en mode *Jitter* (paragraphe 5.2.1.1). L'incertitude sur l'effet nOe est calculée (relation 5.21) sur la base de la déviation standard $sdev$ du bruit déterminée dans une région vide de signal (commande *getval*) et de l'intensité du pic non saturé. Cette approche fait l'approximation que le bruit est identique dans les deux expériences et que les incertitudes sur I_{sat} et I_{eq} sont identiques.

$$\Delta\eta = \Delta \left(\frac{I_{sat}}{I_{eq}} \right) = \frac{sdev}{I_{eq}} \quad (5.21)$$

5.2.1.4 La séquence CPMG : mise en évidence de l'échange chimique dans la gamme des millisecondes

Quelques mots sur le train d'impulsions CPMG Le train d'impulsions CPMG (Carr and Purcell [1954], Meiboom and Gill [1958]) est composé d'une série d'échos de spin, soit une série d'impulsions 180° de champ \vec{B}_1 séparées d'un délai $\tau = \tau_{CP}/2$. La fréquence CPMG est définie par : $\nu_{CPMG} = 1/2\tau_{CP}$. Cette séquence permet de figer l'aimantation transversale sur un axe (spin-lock) à condition que les impulsions soient large-bande afin d'éviter les effets hors-résonances. Outre la refocalisation de certains Hamiltoniens (couplages scalaire et dipolaire, relaxation croisée, CSA, ...), son intérêt réside également dans l'étude de systèmes en échange. Considérons en effet la situation présentée au paragraphe 5.1.6 d'un échange rapide ($k_{ex} > \delta\omega_{ex}$) sur l'échelle des déplacements chimiques entre les deux formes A et B : un seul déplacement chimique sera observé.

Durant le délai τ_{CP} , le spin situé initialement sur le site A peut basculer vers l'environnement B et donc changer de fréquence de précession. La refocalisation de l'Hamiltonien de déplacement chimique en fin d'écho ne sera pas totale ce qui se traduit par un terme de relaxation apparent appelé relaxation d'échange. Les contraintes expérimentales forcent le délai entre deux impulsions à être supérieur à $100\mu s$ soit une fréquence d'impulsions CPMG inférieure à 5000Hz. Nous disposons donc d'une technique pour étudier les échanges s'effectuant sur la gamme des fréquences d'échange inférieures à 5kHz.

Que se passe-t-il alors durant le délai τ_{CP} ? Si la fréquence d'interconversion entre les deux formes est suffisamment grande par rapport à la fréquence CPMG, le spin initialement sur A basculera plusieurs fois sur B entre deux inversions. On se trouve dans un échange rapide vis-à-vis de l'échelle définie par le CPMG. Les termes d'interaction et l'opérateur densité sont moyennés sur les deux formes, le train CPMG est transparent en terme de relaxation : il n'affecte pas le terme d'échange. Si la fréquence d'interconversion est faible par rapport à la fréquence CPMG, alors durant le délai τ_{CP} , le spin restera sur A ou sur B et l'Hamiltonien de déplacement chimique sera refocalisé au terme de l'écho de spin : c'est un échange lent vis-à-vis du CPMG. Le train CPMG détruit alors le terme de relaxation d'échange. Dans le cas où la fréquence d'interconversion est de l'ordre de la fréquence CPMG, l'échange peut être considéré comme intermédiaire : le terme de relaxation d'échange n'est pas entièrement détruit et sa valeur résiduelle dépend à la fois de la durée de vie des états et de la fréquence du CPMG :

$$R_{ex}(\tau_{CP}) = \Phi_{ex}\tau_{ex} \left[1 - \left(\frac{2\tau_{ex}}{\tau_{CP}} \right) \tanh \left(\frac{\tau_{CP}}{2\tau_{ex}} \right) \right] \quad (5.22)$$

$$= \Phi_{ex}\tau_{ex} \left[1 - (4\tau_{ex}\nu_{CPMG}) \tanh \left(\frac{1}{4\tau_{ex}\nu_{CPMG}} \right) \right] \quad (5.23)$$

$$\text{où } \Phi_{ex} = (\omega_A - \omega_B)^2 p_{APB}$$

La figure 5.8 rassemble les trois cas de figures. Le terme d'échange R_{ex} normalisé signifie que le coefficient devant le crochet de l'expression 5.23 a été fixé à 1.

D'un point de vue expérimental, la fréquence CPMG ne peut atteindre la valeur nulle car le délai τ_{CP} ne peut être infini. Le délai maximal est *a priori* gouverné par l'évolution sous le couplage scalaire hétéronucléaire 1J et par la relaxation T_2 . La séquence que nous avons

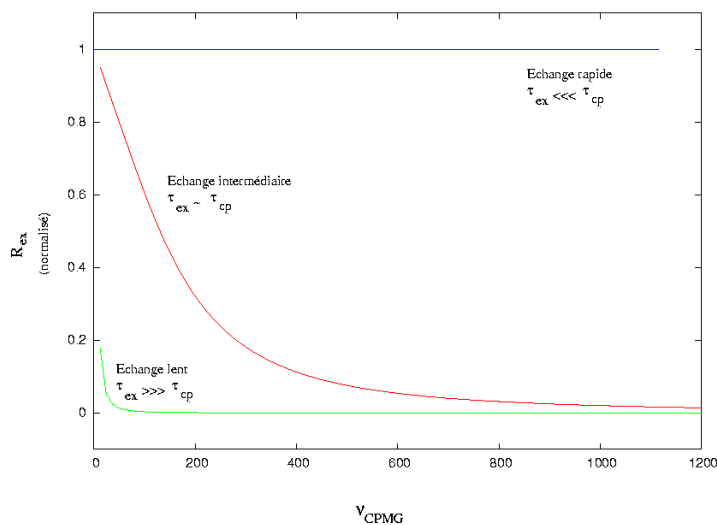


Fig. 5.8: Evolution du terme de relaxation d'échange normalisé en fonction de la durée de vie réduite τ_{ex} et de la fréquence du train d'impulsion CPMG.

utilisée permet de s'affranchir du couplage scalaire car le système évolue sous les termes en phase et en antiphase durant des durées égales. La dépendance de la relaxation transversale vis-à-vis de la fréquence CPMG trahit donc la présence d'échange chimique sur la gamme de temps du délai τ_{CP} .

Séquence d'impulsions La séquence qui m'a permis de montrer la présence d'échange chimique entre des états de durée de vie comprise entre 0,5ms et 5ms a été mise au point par Loria (Loria et al. [1999]). Elle est représentée sur la figure 5.9. Le principe de cette séquence est de comparer l'intensité du signal de chaque pic d'un spectre de type HSQC en variant la durée du délai $\tau_{CP} = 2\tau$ au sein du train d'impulsions CPMG (inversion de l'azote tous les 2τ). La variation de τ_{CP} mène en outre à une estimation de la composante d'échange R_{ex} .

Comme pour toutes les expériences de relaxation, la séquence commence par la destruction de l'aimantation naturelle (à l'équilibre) de l'azote. L'INEPT transfère la polarisation de l'hydrogène amide sur son hétéronoyau. Après le gradient G_2 de défocalisation de l'aimantation résiduelle non transférée sur l'azote, l'opérateur densité est proportionnel à $I_z S_y$ en b . Durant le délai T , on applique la séquence CPMG sur l'azote. Au milieu du délai T , on applique une impulsion 180° sur le proton pour supprimer l'interférence de relaxation entre les Hamiltoniens de CSA et d'interaction dipolaire. En c , l'opérateur densité est identique à celui en b multiplié par un facteur exponentiel traduisant la relaxation $R_{2,anti}$ de l'état en antiphase $I_z S_y$ durant T . Cependant, du fait du couplage scalaire HN pendant T , il se crée temporairement des états en phase S_x dont la relaxation transversale $R_{2,in}$ va être moyennée avec $R_{2,anti}$. Entre c et d , la polarisation est transférée vers l'état S_x . Le second temps constant T est identique au premier (CPMG sur l'azote et impulsion 180° sur le proton). Comme précédemment, les relaxations transversales des états en phase et en antiphase vont être moyennées sur la durée de présence dans chaque état. En e , l'opérateur densité est proportionnel à S_x et à une exponentielle de taux de décroissance effectif $R_{2,eff}$ durant les deux temps constants.

Entre e et f (délai $\Delta = \frac{1}{4J}$), le déplacement chimique de l'azote est édité et le premier transfert INEPT retour est effectué. La suite de la séquence sert à revenir sur le proton pour l'acquisition, par la méthode PEP (*Preservation of Equivalent Pathway*, Palmer III et al.

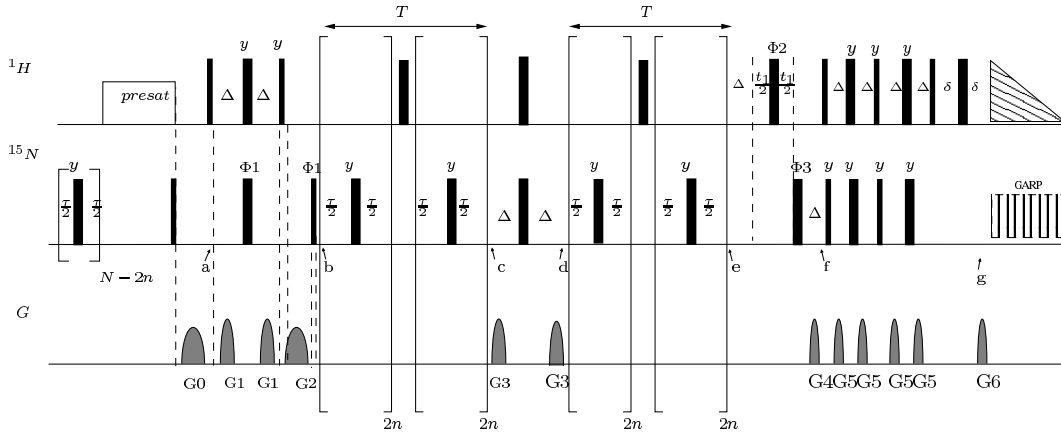


Fig. 5.9: Séquence CPMG compensée par la relaxation. Les impulsions sont suivant l'axe x par défaut. Le cyclage de phase est : $\phi_1 = x, -x$; $\phi_2 = 4x, 4(-x)$; $\phi_3 = 2x, 2y, 2(-x), 2(-y)$, $\phi_{acq} = x, -x, -x, x$. Le délai Δ vaut 2,6ms. La détection est effectuée en mode Echo-AntiEcho sur les gradients $G4$ et $G6$.

[1991a]) : au lieu de ne conserver que le chemin de cohérence issu de l'état S_y modulé par $\cos(\omega_N t_1)$, on sélectionne également le chemin issu de l'état $I_z S_y$ modulé par $\sin(\omega_N t_1)$ en inversant la phase ϕ_3 . Le rapport signal sur bruit est amélioré d'un facteur $\sqrt{2}$ par rapport à l'INEPT retour.

Carine Van Heijenoort observa durant la mise au point de l'expérience que le chauffage dû au CPMG était néfaste. Dépendant du nombre d'impulsions, il était plus important pour les hautes fréquences CPMG. Le rajout en début de séquence d'une série d'impulsions azote hors-résonance est suffisant pour équilibrer le chauffage entre les expériences de fréquences différentes.

Discussion La constante de vitesse de relaxation $R_{2,eff}$ est *a priori* une moyenne pondérée des vitesses de relaxation transversale des états en phase et en antiphase. Si le système est en échange chimique, il s'y ajoute le terme d'échange $R_{ex}(\tau_{CP})$. $R_{2,eff}$ s'exprime alors de la façon suivante :

$$R_{2,eff} = \epsilon R_{2,in} + (1 - \epsilon) R_{2,anti} + R_{ex}(\tau_{CP}) \quad (5.24)$$

Dans cette séquence d'impulsions, du fait de la symétrie de la période d'expression des relaxations R_{in} et R_{anti} , le système passe un temps identique sur les états en phase et en antiphase. Le coefficient ϵ vaut donc 0,5. L'intensité du signal mesuré dépend du délai τ_{CP} :

$$I_{(\tau_{CP})} = A e^{-2R_{2,eff}(\tau_{CP})T} \quad (5.25)$$

où le coefficient A rassemble les autres termes de relaxation, l'efficacité des transferts de polarisation et les propriétés magnétiques des spins nucléaires et T est la durée des CPMG.

Le terme d'échange est le seul dépendant de ν_{CPMG} (expression 5.23). En variant la fréquence du CPMG (limitée par le nombre minimal et maximal d'impulsions pendant T), on obtient la courbe $R_{2,eff} = f(\nu_{CPMG})$ dont on peut extraire les paramètres suivants :

$R_{2,\infty}$, τ_{ex} et Φ_{ex} ($R_{2,\infty}$ est la relaxation transversale moyenne des états en phase et en antiphase hors échange). En théorie, le terme d'échange peut être obtenu en deux spectres : $R_{ex} = R_{2,eff}(\tau_{CP} = +\infty) - R_{2,eff}(\tau_{CP} = 0)$. Mais le délai τ_{CP} est ici limité par le temps constant T (50ms). De même, le délai nul n'a pas de sens car il doit être significatif vis-à-vis de la durée des impulsions (plusieurs dizaines de microsecondes).

La mesure à deux fréquences (80Hz ($\tau_{cp} = 6, 2ms$) et 756Hz ($\tau_{CP} = 0, 66ms$) est suffisante pour déterminer de façon qualitative la présence ou non d'échange chimique sur tel ou tel résidu. En effet, en comparant les intensités I_{80Hz} et I_{756Hz} des pics de ces deux expériences pour un même résidu, on peut distinguer un résidu impliqué dans un échange chimique rapide (dans la gamme milliseconde) d'un autre. Si le rapport $I_{756Hz}/I_{80Hz} \simeq 1$, il n'y a pas d'échange dans cette gamme de temps. Par contre, si ce résidu présente de l'échange, le rapport sera inférieur à 1. Dans ce cas, le terme d'échange apparent $R_{ex,app}$ est défini par :

$$R_{ex,app} = -\frac{1}{T} \ln \left(\frac{I_{756Hz}}{I_{80Hz}} \right) \quad (5.26)$$

Il faut noter que le terme de relaxation d'échange apparent est inférieur au vrai terme d'échange du fait des approximations $R_{ex}(80Hz) = R_{ex}(0Hz)$ et $R_{ex}(756Hz) = R_{ex}(\infty)$.

Nous avons également mesuré la relaxation $R_{2,eff}$ aux deux fréquences. Pour une estimation précise de $R_{2,eff}$, il faudrait varier la durée du temps T tout en conservant la fréquence CPMG et extraire de la courbe exponentielle le taux de décroissance (Mulder et al. [2001]). Nous avons simplifié la procédure car notre objectif était d'obtenir des résultats qualitatifs plus que quantitatifs. Pour cela, la séquence précédente a été amputée des deux temps constants contenant les trains CPMG. La relaxation effective a été calculée par la relation :

$$R_{2,eff}(\tau_{CP}) = -\frac{1}{T} \ln \left(\frac{I_{\tau_{cp}}}{I_0} \right) \quad (5.27)$$

où I_0 et $I_{\tau_{CP}}$ sont respectivement l'intensité du signal sans et avec le temps constant.

L'incertitude sur les vitesses de relaxation calculées par les deux relations précédentes sont obtenues à partir de l'incertitude sur les intensités :

$$\Delta R_{ex,app} = \frac{1}{T} \frac{\Delta \left(\frac{I_{756Hz}}{I_{80Hz}} \right)}{\frac{I_{756Hz}}{I_{80Hz}}} \quad (5.28)$$

Le rapport d'intensités et l'estimation de l'incertitude ont été obtenus en utilisant le module HetNOE de NmrView.

Chapitre 6

Méthodes d'investigation des propriétés d'une protéine

6.1 Les expériences filtrées ^{15}N en dimension F2

Il est intéressant de sélectionner les résonances des protons aromatiques en éliminant les protons amides des spectres 1d et 2d dans la cas d'une protéine marquée ^{15}N . Pour ce faire, on modifie les séquences d'impulsions classiques (1d-Watergate, 2d-NOESY, 2d-TOCSY) basées sur une suppression de l'eau par Watergate.

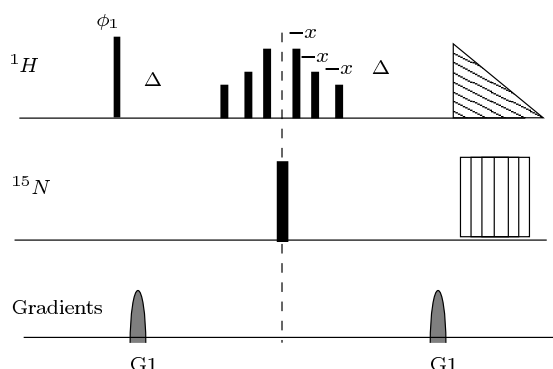


Fig. 6.1: Séquence d'impulsion 1d Watergate filtrée ^{15}N . Les impulsions par défaut sont suivant l'axe x . Le cyclage de phase est le suivant : $\phi_1 = x, -x, \phi_{acq} = x, -x$. Le délai Δ correspond à la moitié du délai entre la première impulsion et le début de l'acquisition.

La figure 6.1 montre la modification apportée à la Watergate. Dans la séquence classique découplée azote, l'inversion de l'azote à mi-parcours est absente. En présence de cette impulsion, l'aimantation H_y des amides évolue sous le couplage scalaire avec l'azote. Si le délai Δ est ajusté à $1/(4^1 J_{HN})$, cette aimantation se retrouve dans un état antiphase $H_x N_z$ qui sera détruit lors de l'acquisition par le découplage azote. Par contre, l'aimantation issue des noyaux non liés scalairement à des azotes sera refocalisée avant l'acquisition dans un état H_y . Le spectre ainsi accumulé ne laissera donc apparaître dans la région entre 6 et 9ppm que les résonances des aromatiques. Pour éliminer les protons amides du spectre, un délai théorique est d'environ 2,6ms. Cependant, il a été optimisé expérimentalement à 3ms dans toutes les expériences.

Le module Watergate avant acquisition des spectres bidimensionnels NOESY et TOCSY ont été modifiés de la même façon.

Acide aminé	pH < pKa	pH > pKa	pKa (modèle)	pKa (protéine)	Acide aminé	pH < pKa	pH > pKa	pKa (modèle)
Asp	COOH	COO ⁻	3,65	2-5,5	Ser	OH	O ⁻	13,0
Glu	COOH	COO ⁻	4,25	2-5,5	Thr	OH	O ⁻	13,0
His	N	NH ⁺	6,00	5-8	Tyr	OH	O ⁻	10,1
Lys	NH ₃ ⁺	NH ₂	10,53	9-12	Cys	SH	S ⁻	8,3
Arg	NH	NH ₂ ⁺	12,48	9-12				

Tab. 6.1: Etat de protonation des chaînes latérales acides et basiques en fonction du pKa. Valeurs des pKa de ces acides aminés et plage de valeurs rencontrées dans les protéines.

6.2 Titration en pH

L'évolution d'une carte ¹⁵N-HSQC en fonction du pH dépend de plusieurs facteurs. Si l'environnement magnétique local des noyaux amides (¹H et ¹⁵N) n'est pas modifié par un changement de pH, alors la résonance restera immobile dans le spectre. Par contre, si le proton amide est en échange avec le solvant, la vitesse de l'échange peut être augmentée par catalyse soit acide, soit basique. Ce phénomène se traduit dans les spectres par une baisse en intensité de la résonance observée. Une variation de l'intensité d'une tache peut alors être interprétée par la non implication de l'amide dans une liaison hydrogène.

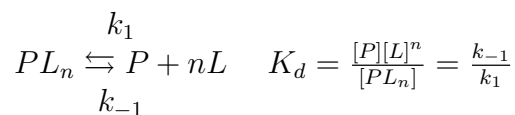
Bien entendu, la proximité spatiale d'un groupe acido-basique modifie l'environnement de l'amide en fonction de son degré de protonation. En général, la vitesse de protonation est rapide vis-à-vis de l'échelle des déplacements chimiques, si bien que les déplacements chimiques de l'amide sont une moyenne des déplacements chimiques entre les états où le groupe acido-basique est protoné ou non. Le suivi en pH d'un groupe amide qui se déplace dans la carte HSQC est donc un moyen de détecter et de mesurer le pKa des groupes acido-basiques. Le pKa des chaînes latérales acido-basiques d'acide aminé en solution (modèle) est rapporté dans le tableau 6.1. Ces pKa peuvent être modifiés dans une protéine.

6.3 Constante d'équilibre et déplacement chimique

Dans la situation où un noyau est en échange chimique rapide (vis-à-vis des déplacements chimiques) entre deux états A et B et où le déplacement chimique des deux états est différent, le déplacement chimique observé est une moyenne des deux déplacements chimiques pondérée par les populations relatives des deux états (relation 6.1).

$$\delta_{obs} = p_A \delta_A + p_B \delta_B \quad (6.1)$$

D'une manière générale, si le coefficient stoechiométrique de liaison d'un ligand L sur une protéine P est n , si les constantes de dissociation des sites de fixation sont identiques et en absence de coopérativité, cette réaction s'écrit :



Si les états P et PL_n sont respectivement les états libre et lié d'une protéine P, à concentration fixe en protéine, le déplacement chimique observé sur P dépend de la concentration en ligand L libre par la relation suivante :

$$\delta_{obs} = \frac{K_d}{K_d + [L]^n} \delta_P + \frac{[L]^n}{K_d + [L]^n} \delta_{PL_n} \quad (6.2)$$

En général, seule la concentration totale en ligand mis en solution est connue. La relation précédente devient alors (avec $\Delta_{max} = \delta_{PL_n} - \delta_P$) :

$$\delta_{obs} = \delta_P + \frac{\Delta_{max}}{2} \left\{ 1 + \frac{K_d}{n[P]} + \frac{[L_t]}{n[P]} - \left(\left(1 + \frac{K_d}{n[P]} + \frac{[L_t]}{n[P]} \right)^2 - 4 \frac{[L_t]}{n[P]} \right)^{\frac{1}{2}} \right\} \quad (6.3)$$

En mesurant la variation de déplacement chimique lors d'une titration en ligand, on peut alors estimer la constante de dissociation et le coefficient stoechiométrique n de l'équilibre. En particulier, lorsqu'à une concentration $[L]$, $\delta_{obs} = (\delta_A + \delta_B)/2$, alors $K_d = [L]^n$.

Si on mesure la variation de déplacement chimique du ligand lorsque sa concentration augmente à concentration en protéine constante, l'expression est plus délicate. Si on connaît le rapport des concentrations totale en protéine ($[P_t]$) et en ligand ($[L_t]$), et si le rapport stoechiométrique est de 1 :1 alors la concentration du complexe PL vaut :

$$[PL] = \frac{1}{2} \left\{ [P_t] + [L_t] + K_d - \sqrt{([P_t] + [L_t] + K_d)^2 - 4[P_t][L_t]} \right\}$$

On en déduit la concentration en ligand libre $[L]$ par la conservation de la matière et les populations relatives $p_L = \frac{[L]}{[L_t]}$ et $p_{PL} = \frac{[PL]}{[L_t]}$.

Troisième partie
Étude par RMN d'ASP2

Chapitre 1

Production et bioinformatique

1.1 Résultat de la production d'ASP2 simplement marquée (^{15}N) et doublement marquée (^{15}N - ^{13}C)

Nous avons suivi le protocole de production de protéine à quatre reprises afin de disposer de quantités suffisantes d'ASP2 pour les diverses expériences de RMN : une préparation doublement marquée ^{15}N - ^{13}C (février 2000) et trois préparations simplement marquées en ^{15}N (septembre 2000, mai 2002 et mars 2003). La production de la protéine marquée a fait l'objet d'une publication dans le journal *Protein Expression and Purification* (Briand et al. [2001b]).

Au cours des nombreuses préparations non marquées, il s'était avéré qu'ASP2 subissait une oxydation des méthionines (détectée en spectrométrie de masse). Afin de limiter cette altération, nous avons réduit la durée d'induction de 6 jours à 2 jours et demi. La quantité de protéine obtenue fut légèrement inférieure mais la qualité, suivie en spectrométrie de masse et en RMN fut supérieure. Le rendement moyen de la production dépend de la durée de l'induction. La quantité de poudre d'ASP2 disponible à partir de 250ml de culture et après purification a varié de 30 à 40mg. La poudre contient divers sels et la quantité d'ASP2 contenue dans la poudre est estimée par Loic Briand à environ 70%. J'ai utilisé par la suite ce facteur pour calculer la concentration de protéine en solution.

La figure 1.1 montre le profil d'éluion du milieu de culture de la levure *Pichia pasto-*

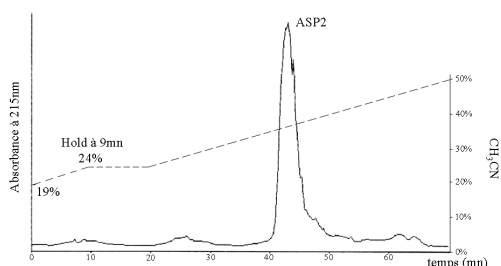


Fig. 1.1: Profil d'éluion d'ASP2 en chromatographie en phase inverse.

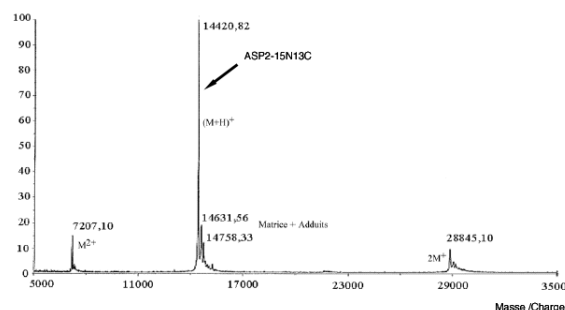


Fig. 1.2: Spectre de masse MALDI-TOF d'ASP2 doublement marquée ^{15}N - ^{13}C .

ris en fonction du gradient en acétonitrile. ASP2 sort pour une concentration d'acétonitrile comprise entre 32% et 37%. Le spectre de masse (figure 1.2) accumulé sur l'échantillon doublement marqué ^{15}N - ^{13}C fait apparaître un pic prédominant à 14420,82Da. La masse de la protéine non marquée étant de 13695Da, on peut estimer l'enrichissement isotopique à 95%. La possibilité d'effectuer ultérieurement des expériences de résonance magnétique nucléaire trois noyaux exclut donc l'hypothèse d'autres modifications traductionnelles (glycosylation) dans l'accroissement de la masse apparente. Les pics à 14632 et 14758Da sont attribués aux complexes protéine-matrice. La protéine doublement ionisée explique le pic à 7207Da. Le pic à 28845,1Da est sans doute la forme dimérique monoionisée, et minoritaire, d'ASP2. Cependant, il est connu que la technique MALDI-TOF fait apparaître ce type d'artefact. L'oxydation des méthionines est contrôlée par la finesse du pic majoritaire. Cet échantillon est particulièrement homogène du point de vue de cette dégradation. Les trois productions de protéine simplement marquée ont montré des résultats analogues, tant en profil d'élution qu'en spectrométrie de masse.

1.2 Analyse de la séquence d'ASP2

Les progrès récents de la bioinformatique permettent d'obtenir des prédictions des structures secondaires d'une protéine à partir de la séquence primaire. Dans certains cas, il est même possible de disposer d'un modèle tridimensionnel. Nous allons voir que les caractéristiques structurales globales d'ASP2 étaient disponibles dès le séquençage : structure secondaire et repliement.

1.2.1 Alignement de séquence

La recherche de séquences homologues par FASTA (tableau 1.1, Pearson and Lipman [1988]) montre qu'il n'y a pas de protéine à plus de 30% d'identité de séquence d'ASP2. Les séquences les plus proches sont toutes des OBPs. La présence des six cystéines conservées est la grande caractéristique des OBPs.

1.2.2 Structures secondaires

J'ai utilisé trois logiciels de prédiction de structure secondaire basés sur l'alignement de séquence : PHDsec²⁰, Psi-Pred²¹ et 3d-PSSM²². La prédiction de structure secondaire par PHDsec indique qu'ASP2 est composée de six hélices. Les résultats de la prédiction sont relatés dans le tableau 1.2.

La prédiction de structure secondaire est compatible avec les études de dichroïsme circulaire : la structure d'ASP2 est essentiellement constituée d'hélices α . Les méthodes Psi-Pred et 3d-PSSM indiquent la présence possible d'un brin β entre les résidus E65 et E69.

²⁰<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>

²¹<http://bioinf.cs.ucl.ac.uk/psipred/>

²²<http://www.sbg.bio.ic.ac.uk/~3dpssm/>

ASP2	--IDQDTVVAKYMEYLPDIMP CADELHISED IATNIQ---AA-KNGA---D---M---S--QL GCLKACVMKRIEMLK -GT-ELY
3D-PSSM	---SQE- VMKNLSL NFGKALDE CKKEM TLTDA I NEDFY---NFWKEGY---E---I---KNRET GCAIM CLSTKLN MLD -PEGNLH
mGen Threader	----SQ VMKNLSL NFGKALDE CKKEM TLTDA I NEDFY---NF-WKEG---Y---EIKNR--ET GCAIM CLSTKLN MLD -PE-GNL
SAM-T99	SQ---- EVMKNLSL NFGKALDE CKKEM TLTDA I NEDFY---NF-WKEG---YEIKN---R--ET GCAIM CLSTKLN MLD PEG-NLH
SAM-T99	SQ---- EVMKNLSL NFGKALDE CKKEM TLTDA I NEDFY---NF-WKEG---YEIKN---R--ET GCAIM CLSTKLN MLD PEG-NLH
FUGUE	----SQ VMKNLSL NFGKALDE CKKEM TLTDA I NEDFY NFWKE -GYEI---K---N---R--ET GCAIM CLSTKLN MLD -PEGNLH
FUGUE	----SQ VMKNLSL NFGKALDE CKKEM TLTDA I NEDFY---NF-WKEGYEIK---N---R--ET GCAIM CLSTKLN MLD -PE-GNL
ASP2	VEPVYKMIEVHAGNADDI QLVKGIANE CIENAKGET DEC NIGNKYTDCY---IEKL-----FS-----
3D-PSSM	HGNAMEFAK KHGADE TMA QL LIDIVHG- CEK STPAND DK CIWTLGVATCFKAEI HKLNWAP SMDVAVGE-----
mGen Threader	HHGNAMEFAK KHGADE TMA QL LIDIVHG CEK STPAND DK CIWTLGVATCF---KAEI-----H KLNWAP SMDVAVGE-
SAM-T99	HGNAMEFAK KHGADE -ETMA QL LIDIVHG CEK STPAND DK CIWTLGVATCF---KAEI-----H KLNWAP SMDVAVGE-
SAM-T99	HGNAMEFAK KHGADE -ETMA QL LIDIVHG CEK STPAND DK CIWTLGVATCF---KAEI-----H KLNWAP SMDVAVGE-
FUGUE	HGNAMEFAK KHGA -DETMA QL LIDIVHG CEK STPAND DK CIWTLGVATCF---KAEI-----H KLNWAP SMDVAVGE-
FUGUE	HHGNAMEFAK KHGADE TMA QL LIDIVHG CEK STPAND DK CIWTLGVATCF---KAEI-----H KLNWAP SMDVAVGE-

Tab. 1.3: Alignement de ASP2 et BmPBP.

En rouge, les résidus identiques

1.2.3 Prédiction de structure par TITO

Le serveur bioserv.cbs.cnrs-fr basé à Montpellier est dédié à la prédiction du repliement en se basant sur l'homologie de séquence (algorithmes 3D-PSSM, FUGUE, SAM-T99, mGen Threader, PDB-BLAST et Jpred2) de protéines dont la structure tridimensionnelle est connue et référencée dans la base de données Protein DataBank (www.rcsb.org/pdb). La séquence présentant la grande homologie de séquence calculée par ces algorithmes est ensuite sélectionnée par TITO. Le résultat final est un modèle tridimensionnel au format PDB de la séquence analysée à partir de la structure du meilleur candidat.

Dans le cas d'ASP2, la séquence de la PBP de *Bombyx Mori* (BmPBP) arrive dans le trio de tête dans tous les calculs d'alignement excepté PDB-BLAST, et représente le meilleur ajustement pour trois des six algorithmes (3D-PSSM, FUGUE, mGen Threader). Cependant, nous l'avons déjà vu, ASP2 ne partage que très peu d'identité de séquence (tableau 1.1) avec les autres protéines connues. En fonction des algorithmes choisis, le résultat de l'alignement avec BmPBP vaut de 13% à 19%. Le tableau 1.3 montrent le résultat des différents algorithmes d'alignement de BmPBP avec ASP2. Les alignements obtenus sont quasiment identiques entre eux à la différence de la région A37 à S45 où les algorithmes proposent chacun un alignement différent. La présence des trois cystéines bien conservées dans la famille des OBPs impose des contraintes importantes dans l'alignement. Outre les cystéines, sept résidus sont identiques : V6, E24, I31, G48, M60, L61, Q87 et D105. A l'exception du résidu L61, aucun de ceux-ci ne sont impliqués dans la liaison du bombykol.

Le modèle d'ASP2 obtenu par homologie de structure avec la structure cristallographique (référence 1DQE) de la BmPBP m'a servi de structure de départ pour l'identification de distances caractéristiques. Cependant, la séquence d'ASP2 est nettement plus courte. En se fiant à l'alignement des structures secondaires, c'est la partie C-terminale de BmPBP qui est absente, or celle-ci est importante dans sa fonction. Il faut donc prendre ce modèle avec précaution.

Chapitre 2

Attribution d'ASP2

2.1 Les premiers spectres d'ASP2 en solution

Le premier spectre ^{15}N -HSQC d'ASP2 enregistré sur l'échantillon doublement marqué dans l'eau (figure 2.1) nous a confortés dans notre projet de mener l'étude structurale d'ASP2. L'intensité des raies de ce spectre (800MHz et 308K) traduit un marquage réussi en ^{15}N tandis que le spectre ^{13}C -HSQC montre que l'isotope 13 du carbone a été incorporé correctement dans la protéine, corroborant ainsi l'analyse en spectroscopie de masse.

La finesse des raies nous informe qu'ASP2 est sans doute dans une forme monomérique. La forte dispersion des raies est en outre une indication de la structuration d'ASP2 en solution. Le choix de la température pour l'attribution a été fixé à 308K pour une finesse de raies favorable tout en évitant une déstructuration fatale de la protéine.

2.2 Les conditions expérimentales de l'attribution d'ASP2

Les conditions dans lesquelles a été menée l'attribution ne reflète pas tout à fait la complexité du comportement d'ASP2 en solution. En effet, nous verrons par la suite que le pH et le TSP affectent les spectres et donc l'état d'ASP2. L'attribution a été menée sur plusieurs échantillons. Toutes les expériences trois noyaux ont été enregistrées sur l'échantillon 1 (voir tableau 2.1). L'attribution des chaînes latérales fut effectuée à partir des tubes 2 et 3. Le remplacement isotopique de l'eau légère en eau lourde entre les tubes 1 et 2 fut possible par

Échantillon	1	2	3	4
Marquage de la protéine	^{15}N - ^{13}C	^{15}N - ^{13}C	^{15}N	non marqué
Concentration en protéine	1mM	1mM	0,7mM	0,5mM
Concentration en tampon phosphate de potassium	100mM	100mM	100mM	100mM
Rapport [TSP]/[ASP2]	30 :1	30 :1	17 :1	34 :1
pH	5,8	6,2	6,1	6,2
Concentration en D_2O	10%	100%	10%	100%

Tab. 2.1: Les conditions expérimentales de l'attribution et de la reconstruction d'ASP2.

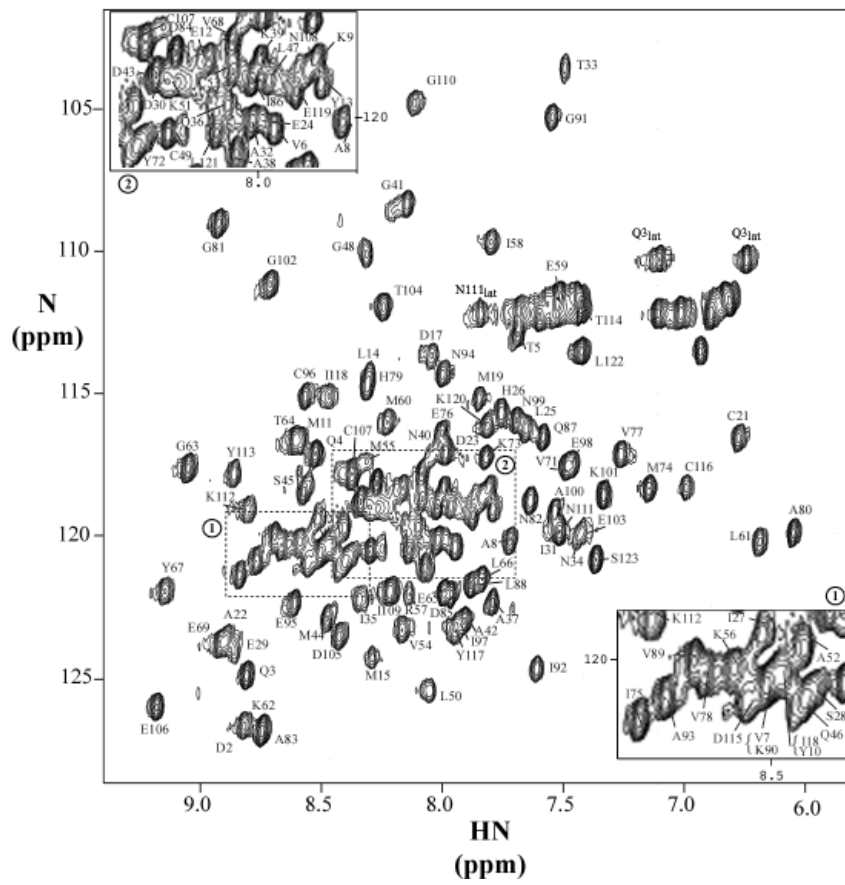


Fig. 2.1: Attribution de l'HSQC d'ASP2 à 308K (800MHz) (tube 1).

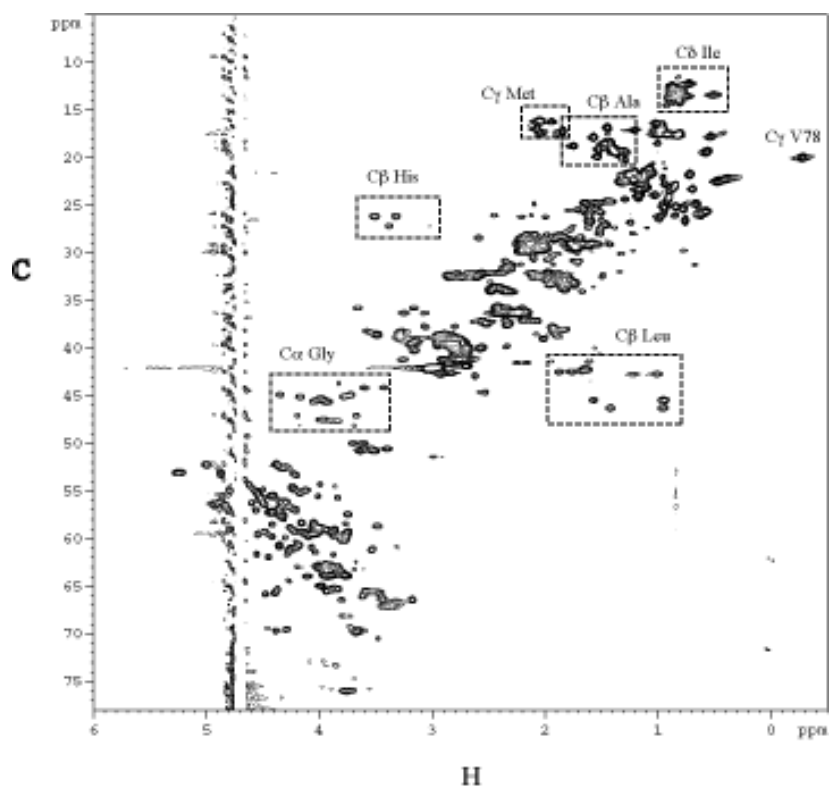


Fig. 2.2: Spectre HMQC ^{13}C à 308K (tube 1). La sélection des cohérences est assurée par des impulsions de gradient.

lyophilisation puis redissolution dans D₂O. Les chaînes latérales aromatiques ont été attribuées dans les conditions du tube 4.

Le choix du pH est un compromis entre le pH physiologique de la lymphe sensillaire (compris entre 6 et 7) et un pH proche du point isoélectrique (pI théorique=4,8). Le fait que le TSP se lie à ASP2 pourrait a priori être un problème lors de la calibration des spectres mais l'excès en TSP par rapport à la protéine est suffisant pour que l'on puisse considérer le déplacement chimique du TSP comme étant celui du libre. Les expériences d'interaction avec le TSP montrent que la différence de déplacement chimique de la protéine entre la forme libre et la forme liée du TSP est faible ce qui rendra aisée l'attribution de la protéine libre. Tous les échantillons contiennent de l'azide pour éviter la prolifération bactérienne.

Toutes les expériences ont été menées sur le spectromètre 800MHz équipé d'une sonde trois noyaux à gradients trois axes. L'expérience HNHA est la seule enregistrée sur le spectromètre 600MHz équipé d'une sonde trois noyaux à gradients trois axes. Les paramètres d'acquisition et de traitement des expériences sont récapitulés dans l'annexe A. Le traitement a été effectué par les logiciels XWINNMR et nmrPipe. Les logiciels Aurelia et NmrView m'ont permis d'analyser les spectres.

2.3 Attribution des résonances d'ASP2 en présence de TSP à pH 6

2.3.1 Attribution du squelette à 308K et pH 6

L'ensemble des expériences trois noyaux ont été menées sur l'échantillon doublement marqué ¹⁵N-¹³C à 308K (tube 1). Un extrait de l'expérience HNCA montre la qualité des spectres (figure 2.3). Ce spectre a été analysé en parallèle avec le spectre HN(CO)CA, qui permet de distinguer les taches intra- et inter-résiduelles. La connectivité par les C_α des deux tiers de la séquence a pu être ainsi déterminée. Le dernier tiers fut difficile à obtenir du fait de l'absence de signal ou de l'ambiguïté due à la superposition des déplacements chimiques, soit des couples N-H, soit des C_α. L'absence de proton amide des prolines est bien entendu une difficulté pour la prolongation de l'enchaînement. Les connectivités trouvées ont été positionnées sur la séquence à la fois par les déplacements chimiques particuliers des C_α des glycines (entre 40 et 48ppm) et des carbones β (expériences CBCA(CO)NH et HNCACB). L'utilisation du transfert par les carbonyles (expériences HNCO et HN(CA)CO, figure 2.4), indépendant des transferts par les carbones aliphatiques, a confirmé les connectivités précédemment déterminées et a étendu l'attribution à toute la séquence.

Les figures 2.3 et 2.4 illustrent l'utilité de travailler sur les deux types de connectivité. Le couple HNCO/HN(CA)CO permet de relier sans ambiguïté les résidus Q3 à V7. Par contre, entre V7 et A8, la superposition des résidus V7 et K90, ainsi que l'absence de taches dans la HN(CA)CO rendent la connexion délicate. Le passage par le carbone α lève toute ambiguïté et permet d'étendre la connectivité.

La carte ¹⁵N-HSQC (figure 2.1) attribuée montre que toutes les résonances des amides du squelette ont été attribuées. Les résonances des hydrogènes NH₂ de chaîne latérale d'asparagine et de glutamine (déplacement chimique de l'azote autour de 112ppm) sont attribuées (voir paragraphe 2.3.2.2). Seuls l'amine du résidu N-terminal et les prolines sont absents de

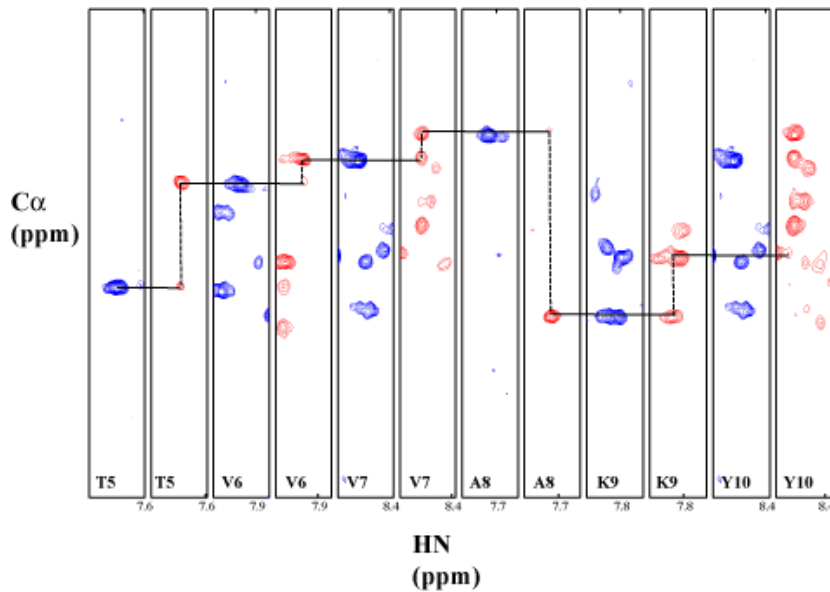


Fig. 2.3: Bandes extraites des plans ^{15}N de la HN(CO)CA (bleu) et de la HNCA (rouge) aux déplacements chimiques du proton amide des résidus T5 et Y10.

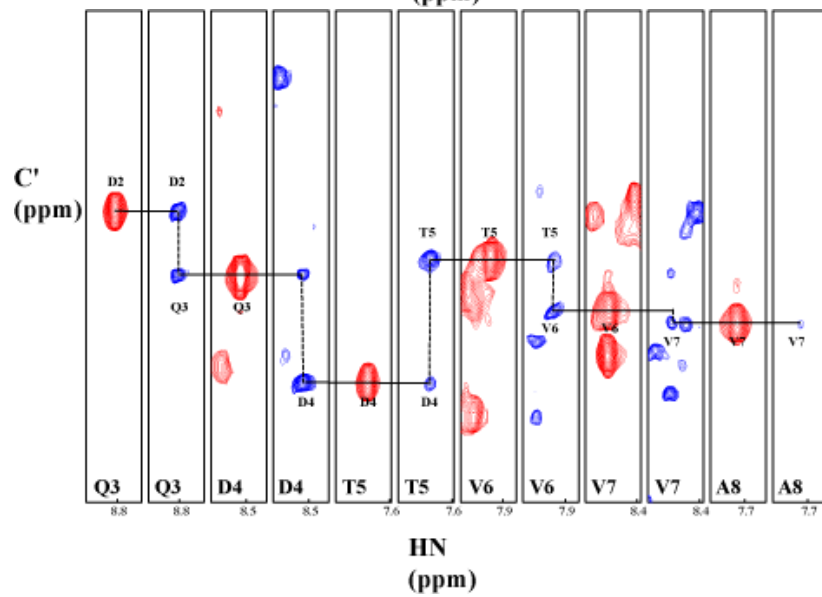


Fig. 2.4: Bandes extraites des plans ^{15}N de la HNCO (rouge) et de la HN(CA)CO (bleu) aux déplacements chimiques du proton amide des résidus Q3 et A8.

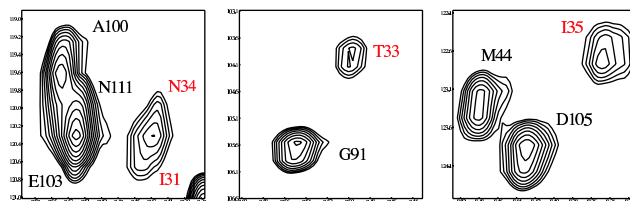


Fig. 2.5: Extraits de ^{15}N -HSQC. Les taches de corrélation des résidus I31, T33, N34 et I35 sont moins intenses que celles des autres résidus.

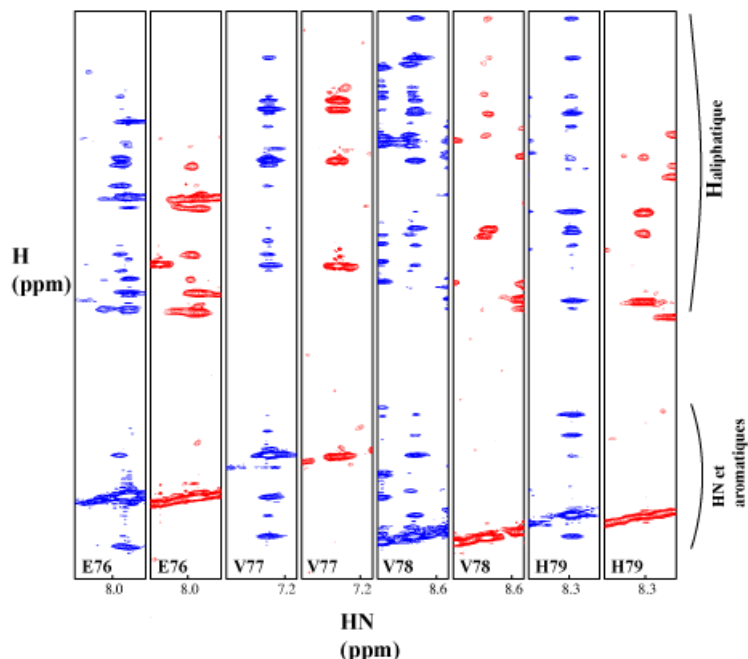


Fig. 2.6: Bandes extraites des plans ^{15}N de la ^{15}N -NOESY-HSQC (bleu) et de la ^{15}N -TOCSY-HSQC (rouge) aux déplacements chimiques du proton amide des résidus E76 à H79 (800MHz et 308K).

la carte. On note dès à présent la faible intensité des signaux provenant des résidus I31 à I35 (figure 2.5).

Le déplacement chimique des carbones β des cystéines (entre 36 et 41ppm) confirme l'état oxydé et donc l'implication des six cystéines dans des ponts disulfure. Le tableau des déplacements chimiques d'ASP2 à 308K est en annexe B.

2.3.2 Attribution des chaînes latérales

2.3.2.1 Attribution des chaînes aliphatiques

L'attribution des chaînes latérales a été menée en deux étapes. Tout d'abord, j'ai déterminé l'attribution des hydrogènes de chaînes latérales aliphatiques en utilisant simultanément les spectres ^{15}N -TOCSY-HSQC ($\tau_m=70\text{ms}$), ^{15}N -NOESY-HSQC ($\tau_m=150\text{ms}$) accumulés sur le tube d'ASP2 marqué ^{15}N à 308K (tube 3). Malgré la faible constante de couplage $^3J_{H^N H_\alpha}$ attendue pour les résidus en hélice α , le transfert TOCSY est suffisant pour observer des taches de corrélation entre le proton amide et les chaînes latérales d'un grand nombre de résidus. Pour exemple, l'attribution de la chaîne latérale de la valine 77 et de l'histidine 79 se lit directement (figure 2.6). Par contre, pour les résidus E76 et V78, situées dans des régions encombrées du spectre HSQC, l'attribution des hydrogènes aliphatiques est plus délicate : il faut utiliser le spectre NOESY qui contient en général les corrélations spatiales entre un proton amide et les H_α et H_β du résidu précédent dans la séquence. Bien entendu, l'attribution à partir du spectre NOESY doit être confirmée par une méthode non ambiguë.

Pour une partie des résidus, l'ambiguïté due à l'encombrement des spectres et/ou à un mauvais transfert TOCSY a été levée en utilisant la ^{13}C -HCCH-TOCSY-HSQC et la ^{13}C -NOESY-HSQC sur le tube 2. En partant des déplacements chimiques connus des carbones C_α et C_β , l'attribution des carbones de fin de chaînes latérales a été possible, conjointement

aux hydrogènes qu'ils portent. Ces expériences sont aussi une confirmation des attributions d'hydrogènes obtenues à partir de l'amide du squelette.

2.3.2.2 Attribution des aminos de chaîne latérale

J'ai attribué les NH_2 des chaînes latérales en partie sur la base des interactions dipolaires qu'ils présentent avec les protons aliphatiques H_β (asparagine) ou H_γ (glutamine). Lors des premiers jeux de structures, aucune contrainte ne les impliquant n'a été introduite pour ne pas orienter la convergence. Lorsqu'un repliement global satisfaisant fut atteint, cette attribution fut confirmée par la présence de nOe entre les NH_2 et des hydrogènes d'autres résidus et cela dans la majorité des structures. Ainsi, la chaîne latérale de l'asparagine 94 s'orientait préférentiellement vers l'acide aminé V68, la présence de nOe entre les H_γ de ce résidu et un NH_2 a donné l'attribution des NH_2 de N94.

Lors des expériences de mise en évidence directe de liaison hydrogène (paragraphe 3.1.2.2), j'ai dû mettre en oeuvre une panoplie d'expériences complémentaires permettant une attribution des carbonyles (et carboxyles) des chaînes latérales. La COCAH m'a permis de déduire ces attributions à partir des corrélations entre $H_\beta-C_\beta-C'_\gamma$ des asparagines et aspartates et entre $H_\gamma-C_\gamma-C'_\delta$ de glutamate et glutamine. Par une version bidimensionnelle de la HNCO classique optimisée pour l'observation des aminos (corrélation aminos-carbonyles), l'attribution des aminos de glutamine et asparagine devenait totalement non-ambiguë. A l'exception des glutamates, dont l'attribution des carboxyles est inutile pour la reconstruction de la structure, tous les carbonyles et carboxyles de Asp, Asn et Gln ont été attribués. La superposition des carbonyles de Q36, Q46 et Q87 à 180ppm rend cette stratégie caduque. Par conséquent, les aminos de ces trois glutamines ne sont attribués que sur la base des nOes vers les protons H_γ de chaîne latérale.

2.3.2.3 Attribution des protons aromatiques

Les protons aromatiques d'ASP2 (deux histidines, une phénylalanine et six tyrosines) ont été attribués avec une démarche identique à celle utilisée pour l'attribution des aminos, en utilisant des spectres NOESY ($\tau_m = 100ms$) et TOCSY ($\tau_m = 70ms$) sur un échantillon d'ASP2 non marquée dissoute dans D_2O (tube 4, figure 2.7). Du fait de la faible concentration en eau légère, une présaturation au début de la séquence d'impulsion suffit à éliminer l'eau résiduelle. Au bout de quelques jours, la plupart des protons amides se sont échangés avec le solvant et se trouvent donc sous une forme deutérée. La région à gauche de l'eau dans le spectre ne contient donc que les protons aromatiques et quelques protons très lentement échangeables. Les six tyrosines ont été attribués à partir des corrélations des H_δ du noyau aromatique vers les H_β intrarésiduels. Les H_ϵ donnent parfois des corrélations (plus faibles) également vers les H_β mais sont de toute façon non ambiguës par la corrélation TOCSY vers les H_δ . Pour les six tyrosines, le déplacement chimique des protons H_δ est supérieur à celui de H_ϵ . J'ai attribué la seule phénylalanine de la séquence sur la base des trois résonances visibles en TOCSY dont deux sont très rapprochées mais distinguables vers 7,15ppm. Les nOes entre les H_δ et les H_β confirment l'appartenance de ces résonances à la phénylalanine 122. Le cas des histidines est plus délicat. La carte TOCSY montre deux taches de corrélations faibles ne correspondant ni aux tyrosines ni à la phénylalanine, et pouvant être des corrélations $H_{\delta 2} - H_{\epsilon 1}$ d'histidine.

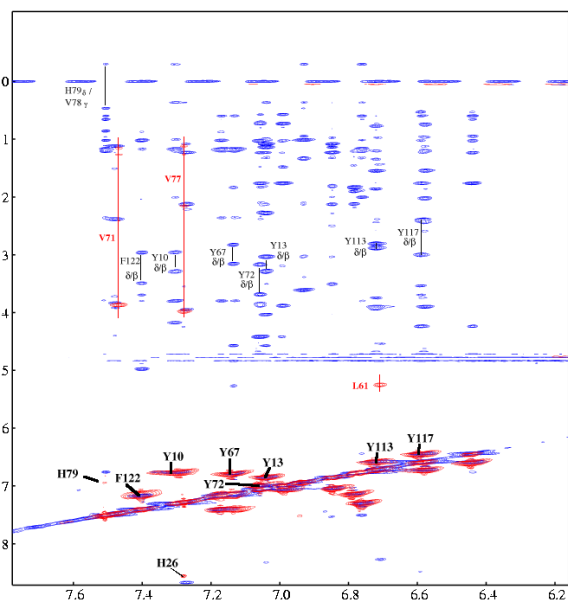


Fig. 2.7: Extrait de la superposition de la TOCSY ($\tau_m=70\text{ms}$, rouge) et la NOESY ($\tau_m=100\text{ms}$, bleu) d'ASP2 pour l'attribution des aromatiques. En bas du spectre, les taches de corrélations entre protons aromatiques intrarésiduels sont annotées. En haut du spectre, les verticales en rouge correspondent aux amides non échangés, en noir les corrélations entre les protons H_δ et H_β intrarésiduels.

La faible d'intensité de ces deux pics est justifiée à la fois par un transfert plus long dans le cas des histidines (${}^4J_{H_{\delta 2}-H_{\epsilon 1}}$) que pour les tyrosines (${}^3J_{H_\delta-H_\epsilon}$) et par un nombre deux fois inférieur de protons. Pour distinguer entre les deux histidines H26 et H79, je n'ai pu utiliser les corrélations vers les protons aliphatiques intrarésiduels. Je me suis basé sur les nOes vers les autres résidus au fur et à mesure de la reconstruction. La présence en particulier de nOes entre une des histidines et les méthyles de la valine 78 a orienté l'attribution de cette histidine à H79. L'autre aromatique d'histidine ne donnant pas de nOes, je l'ai attribuée par défaut à H26. La suite de la reconstruction a donné raison à cette attribution. Il faut noter que les tyrosines et la phénylalanine doivent être en rotation rapide car on n'observe que la moyenne des déplacements chimiques entre hydrogènes situés symétriquement par rapport à l'axe du noyau aromatique.

2.3.2.4 Remarques

Il faut noter qu'aucune attribution stéréospécifique (attribution spécifique des protons des méthyles de valine par exemple) n'a été déterminée. Le protocole de reconstruction d'Aria prend en compte cet aspect en laissant flotter l'attribution des hydrogènes géminaux.

L'attribution de la protéine a fait l'objet d'une publication dans le journal *Journal of Biomolecular NMR* (Lescop et al. [2001]).

2.4 Influence de la température, du TSP et du sel.

2.4.1 Effet du TSP

Le TSP (Sodium 3-(triméthylesilyl)propionate-2,2,3,3-d4, figure 2.8) est utilisé en RMN comme étalon pour déplacements chimiques. Le pic correspondant à la résonances des méthyles est à 0ppm.

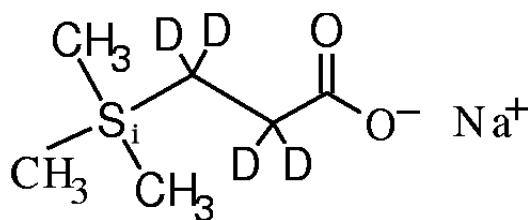


Fig. 2.8: Structure du TSP.

En général, cette petite molécule n'interagit pas avec les protéines étudiées. Dans notre cas, il s'est avéré que dans une expérience NOESY, il apparaissait des taches de corrélation entre les aromatiques (Y10, Y113, Y117 et F122) de la protéine et le TSP (à 0 ppm), preuve indiscutable de l'interaction entre la protéine et la molécule. Dans le chapitre 5, nous détaillerons les interactions entre le TSP et la protéine. A ce point de la discussion, il convient de préciser que la variation du déplacement chimique de TSP entre la forme liée et la forme libre est très faible (environ 0,0124ppm). Par conséquent, la calibration des spectres et donc l'analyse des déplacements chimiques (CSI et Talos) reste valable.

2.4.2 Effet de la température et du sel

Effet de la température Une série de titrations en température a été menée sur un tube d'ASP2 marquée ^{15}N (tube 3) contenant le TSP en rapport 17 :1. Les spectres sont représentés sur la figure 2.9. Très peu de résonances se déplacent fortement (au-delà de 4-6Hz/K en proton) avec la température : elles sont attribuées et mises en exergue par des flèches. Les autres résonances se déplaçant légèrement sont simplement attribuées.

Le fait que le déplacement chimique d'un proton amide est modifié par la température implique que ce noyau est accessible au solvant. Dans un modèle simple, un proton amide donné est en échange entre deux sites : enfoui dans la protéine et au contact du solvant (*random coil*). En augmentant la température, la vitesse d'échange de cet équilibre est augmentée et le déplacement chimique observé se déplace vers le déplacement chimique de l'eau. C'est effectivement ce qu'on observe : les résonances sont déplacées en général vers les bas déplacements chimiques lorsque la température augmente.

On constate que les résonances les plus affectées sont localisées sur la structure primaire : partie N-terminale (D2 et Q3), résidus L50, K62 à Y64, résidus G81 à A83, segment K101 à E106. Les premiers résidus d'une protéine sont en général peu structurés et sont alors au contact du solvant. A l'exception de L50, les autres résidus d'ASP2 fortement affectés sont, d'après la prédiction de structure secondaire, impliqués dans des segments reliant deux hélices ou dans les extrémités de ces hélices. Cette expérience conforte donc la prédiction de structure secondaire. Cependant, certains segments de la séquence semblent être en échange avec le solvant alors qu'ils sont structurés : L50, V54, E95-C96. Il est donc probable que l'hélice α dans laquelle ils sont impliqués n'est pas aussi structurée que les autres hélices.

On peut noter que I31, T33 et N34 font partie des résidus qui ne se déplacent pas au-delà de 2Hz/K.

Une expérience NOESY à environ 60°C sur une protéine non marquée a montré que les motifs généraux de la structure étaient conservés. La structure de la protéine est donc très stable vis-à-vis des hautes températures. Il semble que ce soit un caractère général des OBPs (Ban et al. [2003], Picone et al. [2001], Rothmund et al. [1999]). Il est possible que la tem-

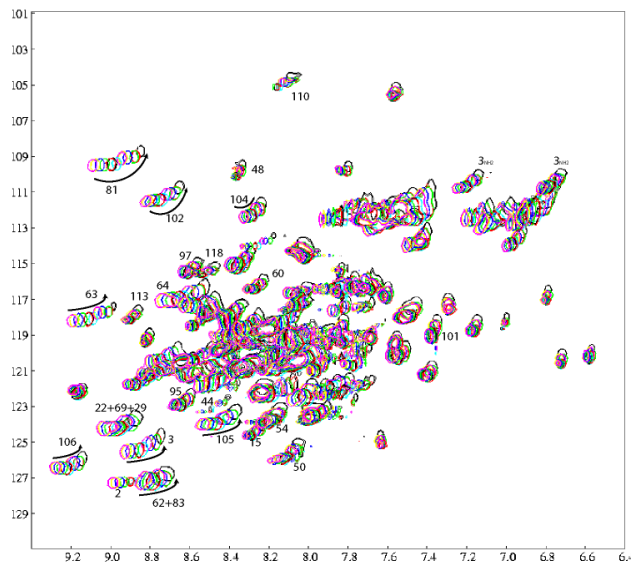


Fig. 2.9: Spectres ^{15}N -HSQC en fonction de la température (de 294K, en magenta, à 318K, en noir, par pas de 2K) à pH 6 en présence de TSP.

pérature atteinte dans les ruches (jusqu'à 45°C) explique cette stabilité, d'autant plus que la régulation de la température est moins efficace dans les antennes que dans le reste du corps de l'abeille.

Effet du sel Durant l'étude d'ASP2, une série de tubes a été préparée soit sans tampon salin, soit avec un tampon de 100mM de phosphate de potassium. Afin de contrôler l'effet potentiel de la salinité du milieu, j'ai enregistré un spectre ^{15}N -HSQC sur un échantillon d'ASP2- ^{15}N en absence de sel et sur le même échantillon dans lequel j'ai ajouté le tampon phosphate pour arriver à une concentration de 100mM (spectres non montrés). Je n'ai observé aucune modification notable de la carte. La présence de ce tampon n'influe donc pas sur l'état d'ASP2.

2.5 Effet du pH : ASP2 libre est en équilibre entre deux formes au-dessus de pH 6,6.

Dans cette section, nous allons étudier l'effet du pH sur les spectres d'ASP2. Bien que la discussion présente n'ait pas encore fait état de la structure d'ASP2, nous allons prendre un peu d'avance sur le chapitre 2. Nous utiliserons le résultat du calcul de structure pour visualiser et interpréter les résultats des titrations. J'ai effectué deux titrations en pH de la protéine marquée ^{15}N : en absence et en présence de TSP.

2.5.1 Titration d'ASP2 libre en solution

2.5.1.1 Protocole expérimental

Le tube qui a servi à la titration en pH contient $550\mu\text{l}$ d'une solution à 0,3 mM d'ASP2 marquée ^{15}N fraîchement dissoute, 50mM NaN_3 et 10% de D_2O . Dans ces conditions, le pH

résidu	δ_{HN} (ppm)	δ_N (ppm)
94	8,12	114,4
95	8,40	121,5
96	8,85	117,8
98	7,74	119,0
99	7,88	115,4

Tab. 2.2: Attribution des résidus N94 à N99 d'ASP2 dans la forme libre.

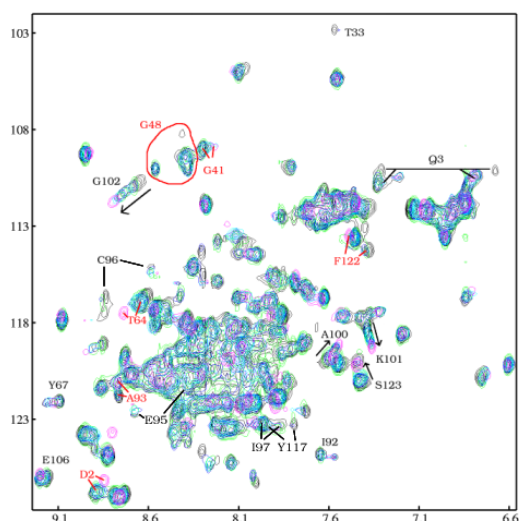


Fig. 2.10: Superposition des spectres ^{15}N -HSQC entre pH 8-8,5 et pH 6,1 à 308K (du noir vers le magenta). Les résidus dédoublés sont attribués en noir. Les déplacements importants sont suivis par une flèche. En rouge, les résidus dont la résonance est fortement déplacée entre pH 6,6 et 6,1.

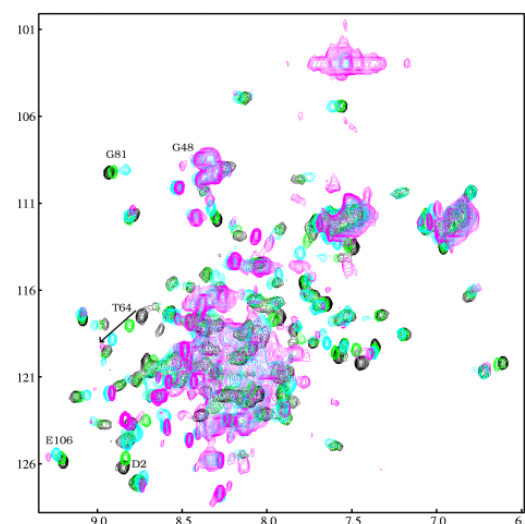


Fig. 2.11: Superposition des spectres ^{15}N -HSQC à pH 6,1 (noir), 5,7 (vert), 4,3 (bleu) et 3,5 (magenta).

est équilibré entre 8 et 8,5. J'ai ajusté le pH ensuite par ajout successif d'une solution environ molaire d'acide chlorhydrique. J'ai ainsi balayé le pH entre environ 8 et 3,5 : 8-8,5 ; 7-7,5⁺ ; 7-7,5⁻ ; 6,5-7⁺ ; 6,5-7⁻ ; 6,6 ; 6,1 ; 5,7 ; 4,3 et 3,5. Jusqu'à pH 6,6, le pH a été mesuré par papier pH, en-dessous, le pH a été mesuré par un pH-mètre calibré entre 4 et 7. Le papier pH est moins précis que le pH-mètre, ce qui explique les intervalles de valeurs pour les pH supérieurs à 6,6. L'exposant ⁺ (resp. ⁻) indique que le pH est dans la gamme supérieure (resp. inférieure) de l'intervalle. Pour chaque pH, j'ai accumulé un spectre ^{15}N -HSQC et un spectre Watergate filtré ^{15}N pour éliminer les résonances amides.

2.5.1.2 Analyse des spectres HSQC et 1d Watergate filtrée ^{15}N .

L'attribution de la protéine libre a été effectuée par une NOESY-HSQC éditée ^{15}N . La plupart des résonances peuvent être attribuées sur la base de l'attribution déjà menée sur le

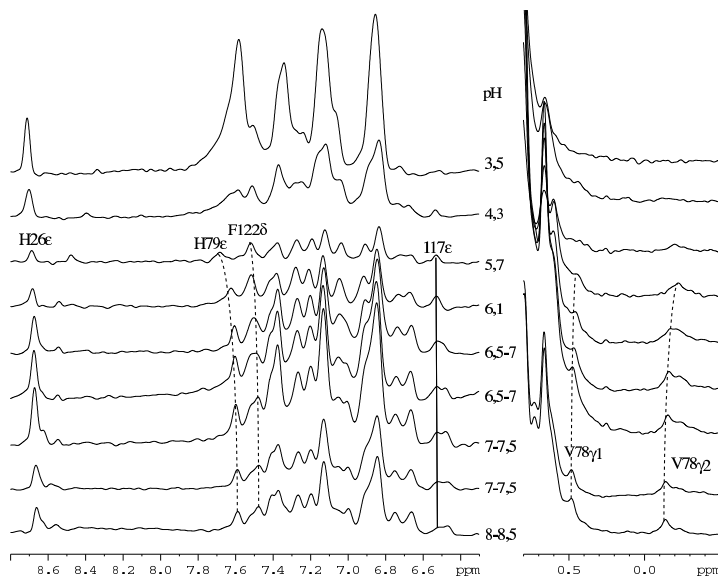


Fig. 2.12: Titration en pH suivie par des spectres 1d Watergate filtrée ^{15}N dans la région des aromatiques et des méthyles de la valine V78 (à droite).

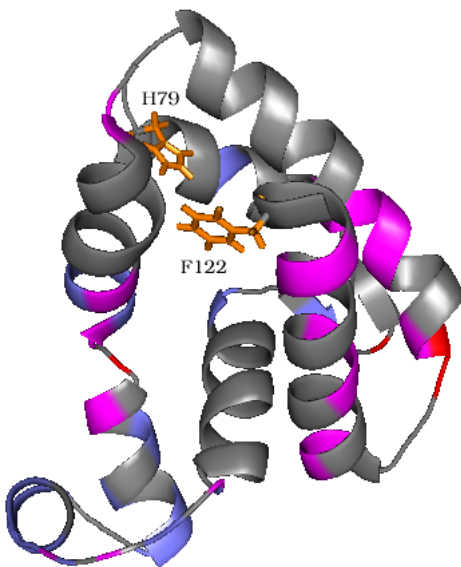


Fig. 2.13: Titration en pH au-dessus de 6,6. La variation de déplacement chimique du squelette est codée par la couleur : du bleu au rouge, les déplacements sont de plus en plus importants. Les chaînes latérales de H79 et F122 sont en orange.

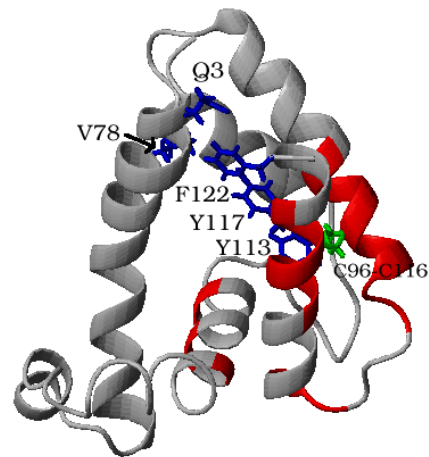


Fig. 2.14: Deux environnements en échange lent. Les résidus pour lesquels au minimum deux résonances amides ont été observées sont en rouge. En bleu, les chaînes latérales présentes sous deux formes (Q3, V78, Y113, Y117 et F122). En vert, le pont disulfure C96-C116.

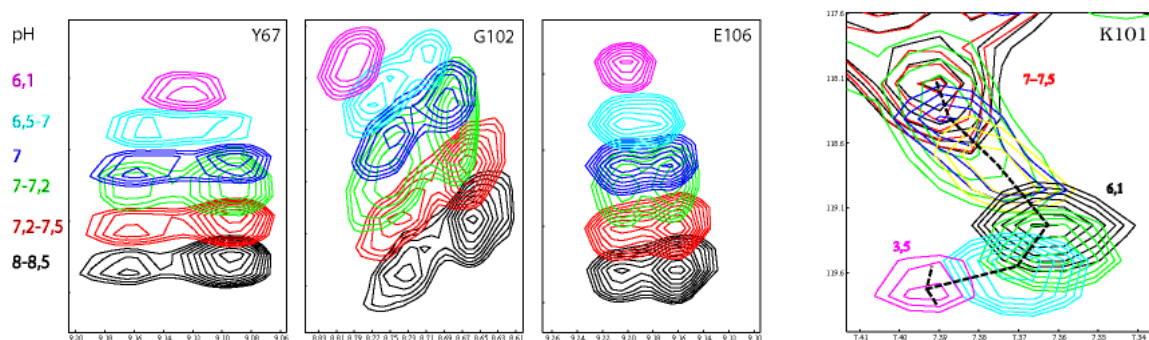


Fig. 2.15: A gauche, titration en pH des protons amides des résidus Y67, G102 et E106 entre pH 8-8,5 et 6,1. A droite, déplacement de la résonance de K101 entre pH 8-8,5 et 3,5. Les trois valeurs de pH marquent les points caractéristiques de la titration.

complexe ASP2/TSP. Il faut préciser que les segments 94-99 et 114-118 subissent une forte variation de déplacement chimique entre la forme libre à pH 7-8 et en complexe avec le TSP : les déplacements chimiques des résidus 94 à 99 de la forme libre sont tabulés (tableau 2.2). L'isoleucine I97 n'a pu être attribuée.

Au-dessus de pH 6,1 Lorsque la protéine est solubilisée (pH 8-8,5), la carte HSQC montre un dédoublement de certaines résonances (par exemple, Y67, G102, E106, voir figure 2.15) avec une forme majoritaire et une forme minoritaire. Les résidus G48 et L50 sont représentés par trois ou quatre résonances. Les résonances des deux formes des régions E95-C96 et C116-Y117-I118 de la protéine sont relativement éloignées les unes des autres (reliés par les traits sur le spectre HSQC). La différence de fréquences de résonance des deux composantes des autres résidus est inférieure à 20Hz. De même, les protons H_ϵ de la tyrosine Y117, les protons δ de la phénylalanine 122 et le méthyle γ_2 de la valine V78 sont en échange lent entre deux formes (figure 2.12). Du fait de la superposition au centre de la région des aromatiques, le comportement des autres aromatiques n'a pu être suivi. La distribution des sites en équilibre lent entre deux états est illustrée sur la structure d'ASP2 (figure 2.14).

En baissant le pH, deux phénomènes ont lieu. Une cinquantaine de résonances se déplacent plus ou moins fortement (figures 2.10 et 2.13) du fait d'une titration de pKas. Les raies des résidus K101 et E102 sont les plus fortement déplacées. Étonnamment, l'intensité des deux formes en échange s'inverse : la forme minoritaire devient majoritaire. Cela se voit sur les résonances des chaînes latérales de V78 et Y117, ainsi que sur les protons amides de Y67, G102, E106 (figure 2.15), M60, A100 et D105. A un pH d'environ 7, les deux formes sont présentes avec la même intensité tandis qu'à pH 6,1, une seule résonance subsiste pour tous les résidus, y compris la glycine G48 et L50. Cette transition est visible également sur les chaînes latérales.

Entre pH 7 et pH 6,1, un certain nombre de résonances sont singulièrement déplacées (en rouge sur la figure 2.10). La plupart sont insensibles au pH et, bien qu'il soit possible qu'à ce pH, nous titrions un pKa plus faible, il est plus probable que ce déplacement soit en réalité l'apparition de la forme minoritaire au détriment de la forme majoritaire.

Nous pouvons donc conclure qu'à un pH supérieur à 7, deux formes d'ASP2 existent en solution, en échange lent : une **forme B** (pour basique) et une **forme A** (pour acide). A pH

6,1, nous sommes en présence d'une forme unique A pour tous les résidus, forme minoritaire à haut pH. Les plus fortes différences de déplacements chimiques entre les formes sont localisées sur la structure : la zone contenant le plus de dédoublement concerne le segment peptidique reliant les deux cystéines C96 (hélice α_5) et C116 (hélice α_6) impliquée toutes deux dans le même pont disulfure.

Nous notons également le déplacement avec le pH de la résonance du proton H_ϵ de l'histidine H79. La gamme de pH dans laquelle nous nous situons correspond au pKa attendu d'une histidine (6-7). Si la fréquence de résonance de ce proton se déplace, c'est que nous observons vraisemblablement la protonation progressive d'un des azotes du cycle aromatique de cette histidine.

Au-dessous de pH 6 J'ai mesuré trois points de pH en dessous de pH 6,1 : 5,7, 4,73 et 3,5 (figure 2.11). A pH 3,5, la dispersion des résonances est plus faible. La protéine subit sans doute une transition vers un état "*molten-globule*" dans lequel la structure est partiellement dénaturée. Une partie des résidus peut être attribuée en suivant les résonances. Cependant, leur distribution non localisée sur la structure ne permet pas de définir une région de la structure se dépliant avant une autre.

Entre pH 6,1 et 4,5, il est fort probable que les groupes acides des aspartates et glutamates soient titrés, comme le montre entre autre le déplacement, à partir de pH 6,1, de G81, Y67, E106, K112 et Y113, proches spatialement respectivement de D85, E65 et E98, E103 et E106, D115. Le cas de la lysine K101 est intéressant. La figure 2.15 montre que le mouvement de la résonance associée à K101 n'est pas continu. Il semble donc que le point d'inflexion à pH 6,1 traduit la présence de deux groupes acido-basiques de pKas différents dans l'environnement de la lysine. Il s'agit probablement des lysines K112 ou K101 et des glutamates E65 ou E103.

La dénaturation partielle est réversible Durant cinq jours, la solution est restée à pH 3,5. J'ai ensuite vérifié la stabilité de la protéine en remontant le pH à 5,8. Le spectre ^{15}N -HSQC est alors complètement retrouvé. La dénaturation partielle est donc réversible. En ajustant à nouveau le pH à 7,5, le dédoublement des raies réapparaît. A pH 12,4, une partie des résonances disparaît à cause de l'augmentation de la vitesse d'échange entre les protons amides et les protons aqueux. Cependant, les deux formes A et B sont toujours en équilibre lent : les résonances associés aux acides aminés G48, I58, Y67, I92 entre autres, sont dédoublées.

Sur toute la gamme de pH, on constate que la résonance du proton H_{26_ϵ} est fixe avec le pH, même dans l'état de *molten-globule*. Cela suggère que le groupe imidazole de ce résidu n'est pas affecté par la dénaturation partielle de la structure et confirme donc l'absence de nOe et son orientation vers le solvant.

2.5.1.3 Discussion

Au-dessus de pH 6,6, le suivi de la titration en pH des résonances est délicat du fait de l'hétérogénéité d'ASP2 libre. Plusieurs raies sont dédoublées et par conséquent, la détermination de tout pKa est impossible. C'est le cas en particulier pour la lysine 101 et la glycine 102. Il aurait été intéressant de déterminer le pKa titré car il est fort probable que le groupe acido-basique à l'origine de ces titrations est sans doute le même pour ces deux résidus.

Le passage à une solution très acide (pH 3,5) dénature partiellement la structure (*molten-globule*). Certaines structures secondaires, voire tertiaires, sont conservées comme le suggère la dispersion des raies dans la carte ^{15}N -HSQC. Cependant, le repliement global de la protéine est affecté comme le montre le spectre des aromatiques, où les résonances sont très resserrées. Ce résultat confirme les spectres de dichroïsme circulaire effectués à pH 3 par l'équipe de l'INRA et qui concluaient à une augmentation de la flexibilité des chaînes latérales.

Le cycle en pH indique que la dénaturation partielle est réversible ainsi que la reformation de la forme B. Il ne peut donc s'agir dans notre cas d'un exemple d'hystérésis en pH.

Mais le résultat le plus important de cette expérience est la mise en évidence d'un équilibre lent d'ASP2 à pH basique entre deux formes A et B, dont les populations respectives dépendent du pH. De par les points de pH échantillonnés, le pH exact de transition n'a pu être déterminé. Il est compris entre 6,1 et 6,6. Les spectres très similaires de ces deux formes suggèrent que les modifications à l'origine de la séparation n'ont qu'un effet très faible sur la structure d'ASP2. Cependant, la région localisée autour du pont disulfure semble être importante. *A priori*, l'observation d'une telle hétérogénéité peut être dû à plusieurs facteurs : équilibre monomère-dimère où les hélices α_5 et α_6 constitueraient l'interface du dimère, équilibre entre deux conformations stables de la protéine (isomérisation d'un pont disulfure, changement d'orientation d'une chaîne latérale, ...). Cela m'a motivé à entreprendre une caractérisation des deux formes. Ce sera l'objet du chapitre 4.

Quelques soucis de reproductibilité Ces différents résultats sont à nuancer. Durant la durée de ma thèse, j'ai préparé plusieurs échantillons d'ASP2 sans TSP. La dépendance en pH telle que relatée ici n'a pu être reproduite. Quel que soit le pH, la forme B est restée majoritaire alors que la forme A est en général observée. Les hypothèses de l'existence d'un cycle d'hystérésis ou d'un effet de la concentration en protéine sont exclues. Une explication possible est une hétérogénéité entre les lots d'origine, due par exemple à des concentrations en sel différentes. En effet, certains lots avaient un aspect cristallin tandis que d'autres étaient sous la forme d'une poudre fine.

D'autre part, les résidus dont les résonances sont présentes en double ne sont pas toujours les mêmes d'un échantillon à un autre. Cependant, j'ai noté un dédoublement systématique des résidus marqués en rouge sur la figure 2.13. Dans certains échantillons, les résonances correspondant aux acides aminés L14, A22, I58, Y67 sont apparues dédoublées.

2.5.2 Titration en pH du complexe ASP2-TSP

2.5.2.1 Protocole expérimental

La titration en pH sur le complexe ASP2-TSP a été menée sur le tube d'ASP2 marquée ^{15}N - ^{13}C , utilisé pour l'attribution (tube 1). Le pH a été ajusté à partir d'une solution de soude environ molaire. La solution initiale était à pH 6,2. J'ai effectué quatre points à pH 6,5, 6,8, 7 et 7,3. La titration de chaque point de pH a été suivie par un spectre ^{15}N -HSQC.

Ces groupes ont des pKa éloignés de la zone titrée et ne sont sans doute pas à l'origine des déplacements observés.

Chapitre 3

La structure tridimensionnelle d'ASP2 en complexe avec le TSP.

ASP2 lie le TSP. Je n'ai observé ce phénomène que tard dans le processus de reconstruction. Je suis donc allé jusqu'au bout de la reconstruction d'un complexe ASP2/TSP. La caractérisation de la forme libre d'ASP2 sera discutée au chapitre 5.

3.1 La structure secondaire d'ASP2

3.1.1 Les contraintes diédrales

3.1.1.1 *Chemical Shift Index*

La figure 3.1 rassemble les résultats de l'analyse par la méthode CSI. Le CSI brut par type de noyau est affiché ainsi que la structure secondaire consensus (H : hélice α , B : structure étendue). Comme prédit par les algorithmes PSSM et PHDsec, ASP2 est composée essentiellement d'hélices α . La structure secondaire résultant du consensus entre les CSI des quatre noyaux place 8 fragments en hélices : 3-13, 21-24, 28-32, 36-41, 44-57, 70-78, 82-99, 106-121 et deux brins β : 60-62 et 65-68. Il faut noter que les résidus cystéine 96 et 116 ont tendance à briser l'enchaînement des deux dernières hélices. Le pont disulfure qu'ils forment est peut-être à l'origine d'une perturbation du squelette des hélices.

La propension des résidus à former une hélice (au sens du CSI) n'est pas homogène sur la séquence. La figure 3.2 représente la différence entre les déplacements chimiques observés sur ASP2 et les déplacements chimiques en pelote statistique (*random coil*). L'écart par rapport au *random coil* du déplacement chimique du C_α est en moyenne de 3ppm pour les régions en hélice α . Le segment 28-41 contient deux zones prédites en hélice, cependant, l'écart par rapport au *random coil* est en moyenne beaucoup plus faible. Cette tendance est visible également sur le noyau C' . Bien que la variation du H_α soit peu ou prou de même grandeur entre les différentes hélices, l'écart sensiblement plus faible du déplacement chimique de la région 28-41 par rapport aux autres hélices prédites est peut-être le reflet d'une instabilité plus grande des hélices 28-32 et 36-41 (Eliezer et al. [1998]).

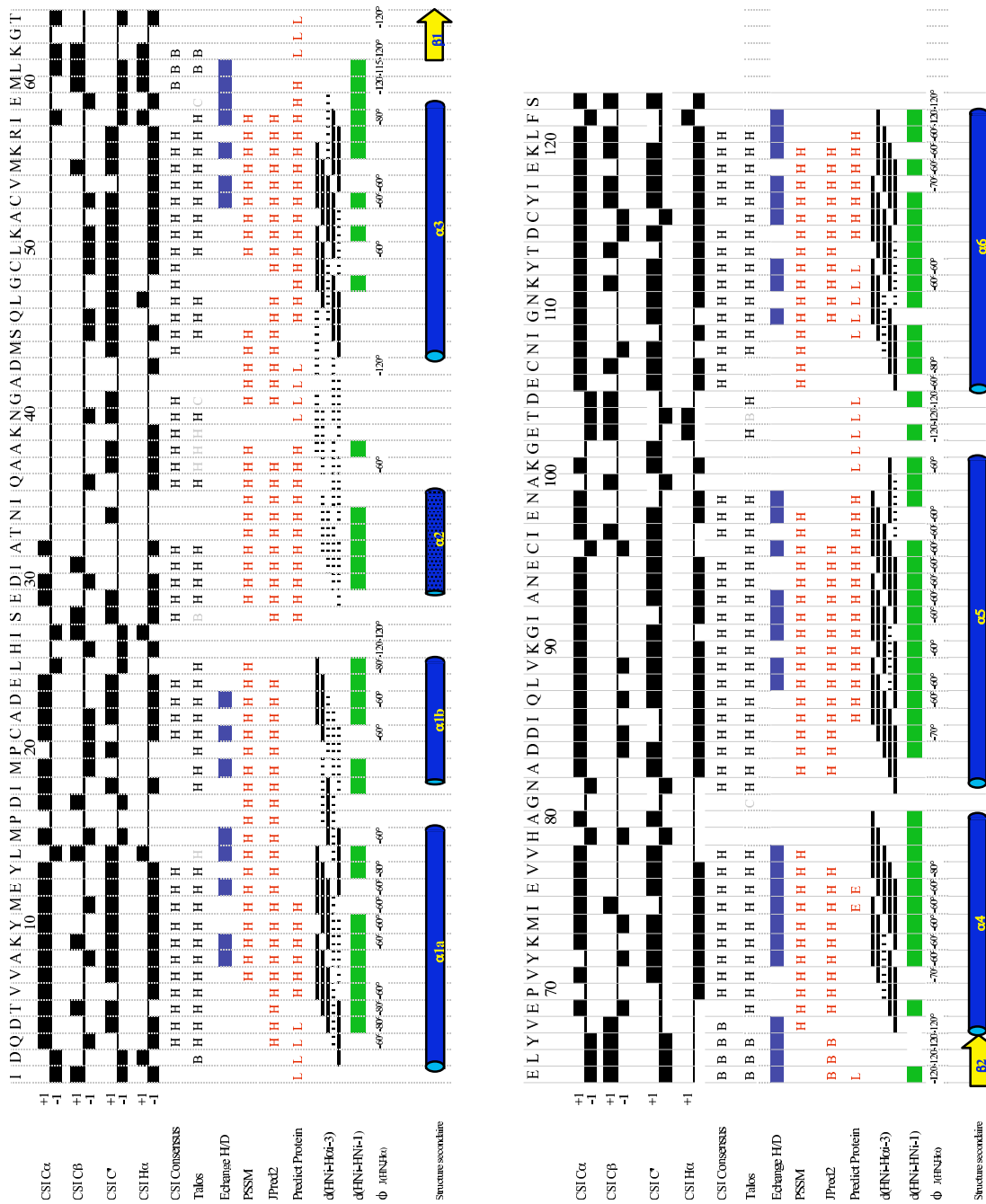


Fig. 3.1: Structure secondaire d'ASP2 à 308K : CSI, Talos, échange H/D, prédiction par PSSM et PredictProtein, nOes $H_N^i - H_\alpha^{i-3}$ (trait plein : non ambigu ; en pointillé : ambigu), intensité des nOes $H_N^i - H_N^{i-1}$, structure secondaire consensus.

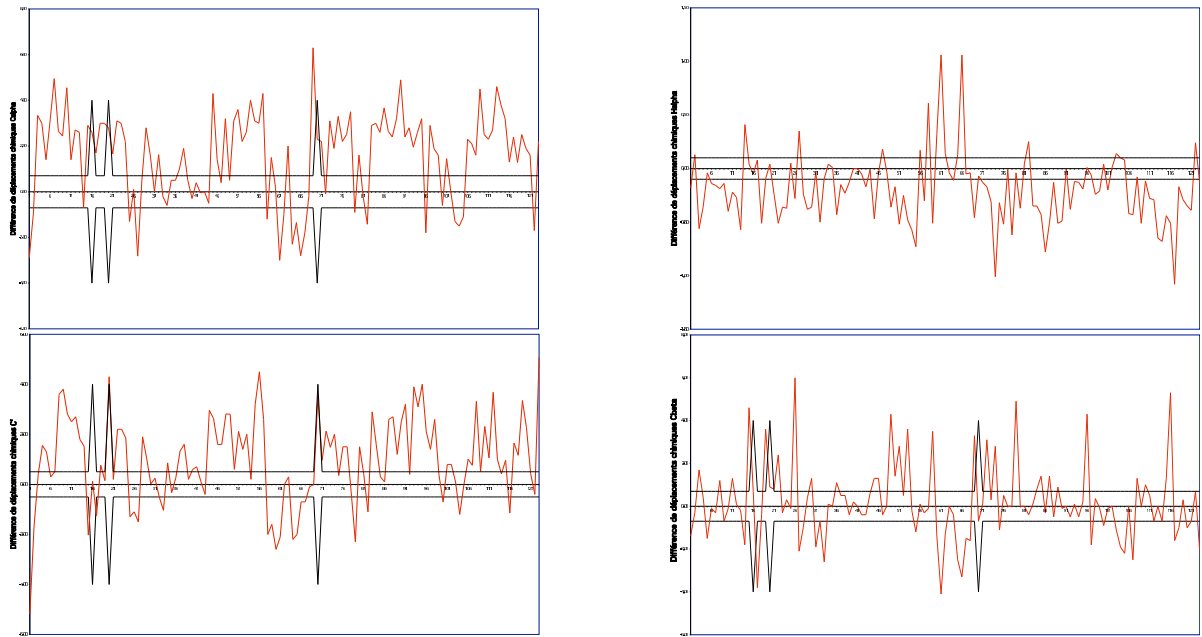


Fig. 3.2: Différence entre le déplacement chimique des noyaux C_α , C_β , C' et H_α d'ASP2 à 308K et le déplacement chimique en *random coil*. Si la différence est supérieure (resp. inférieure) à la limite, un indice de +1 (resp. -1) est affecté au résidu.

3.1.1.2 Talos

Les angles dièdres ϕ et ψ de 86 résidus ont pu être prédits par Talos (résultats en Annexe C). La figure 3.3 rassemble les résidus pour lesquels une prédiction fiable a pu être obtenue : en vert, les résidus dont Talos a conclu à une prédiction satisfaisante, en gris, les autres. Les résultats sont rassemblés dans la figure 3.1 en fonction de la position des angles dièdres du squelette dans la carte de Ramachandran. Les résidus situés dans la zone de la carte de Ramachandran réservée respectivement à l'hélice α , à la structure étendue et au coude α sont notés H, B et C. Les résidus ne montrant qu'une propension sont notés en gris. L'utilisation de Talos confirme en grande partie le résultat de CSI. Ainsi, Talos prédit l'existence de sept hélices α de plus de quatre résidus. Parmi les prédictions moins fiables (huit triplets dans la même zone de la carte de Ramachandran au lieu de neuf), neuf résidus montrent une propension à être en hélice, en structure étendue ou en coude (noté en gris dans le tableau récapitulatif). On peut noter que les résidus 37 à 40 ont une propension à former une hélice. Ceci suggère peut-être l'existence d'une courte hélice entre Q36 et N40.

3.1.1.3 Détermination des angles ϕ : l'expérience HNHA

L'expérience HNHA menée sur ASP2 a permis de déterminer 72 valeurs de $^3J_{H_N H_\alpha}$ sur les 123 résidus. J'ai appliqué la correction de 11% proposée par Vuister et Bax en faisant l'hypothèse qu'ASP2 et la nucléase du staphylocoque (Torchia et al. [1989]) ont un temps de corrélation proche. J'ai vérifié que ce facteur de correction était correct sur les résidus K62, T64, L66, E65 du feuillet β dont les valeurs de couplages sont suffisamment grandes par rapport à la largeur de raies pour les mesurer directement sur le spectre HSQC. Les deux



Fig. 3.3: Talos : un logiciel de prédiction des angles dièdres du squelette peptidique. Exemple du résidu E24 d'ASP2. A droite : les 10 triplets de la base de données de plus haute similarité avec le triplet centré sur E24. A gauche : carte de Ramachandran. En vert, les 9 résidus dans une zone favorable. En rouge, le résidu dans une zone défavorable. Au milieu : séquence d'ASP2. En vert, les résidus pour lesquels Talos a pu prédire les angles dièdres avec une grande confiance.

raisons essentielles pour lesquelles le couplage n'a pu être obtenu sont la superposition des raies dans la ^{15}N -HSQC et l'absence de signal en dehors de la diagonale. Pour certains résidus dont la tache hors-diagonale n'est pas visible (M11, M15), l'intensité $I_{\text{hors-diagonale}}$ est donc inférieure au bruit, j'ai alors accordé une borne supérieure pour $^3J_{H_N H_\alpha}$. Lorsqu'il y a superposition sur la diagonale, l'intensité de la tache sur la diagonale est un majorant de I_{diag} . Cela permet dans certaines situations de minorer $^3J_{H_N H_\alpha}$.

Les constantes de couplage ont ensuite été traduites en plages de valeurs pour l'angle ϕ . Pour certaines valeurs de J intermédiaires (entre 6 et 8 Hz), la connaissance de la structure secondaire est indispensable pour distinguer entre les plages de valeurs autour de -60° et autour de -120° . Ne disposant pas de l'attribution stéréospécifique des glycines, les valeurs de couplage mesurées ne peuvent être interprétées en termes géométriques. La figure 3.1 résume les angles ϕ qui ont pu être déterminés. L'ensemble des résultats est en annexe C.

3.1.1.4 Comparaison des différentes méthodes.

Nous avons donc utilisé trois méthodes pour déterminer les angles ϕ du squelette d'ASP2. Il est intéressant de comparer le résultat issu de chaque méthode. Le résultat est représenté sur la figure 3.4 et la figure 3.5 rapporte les valeurs de $^3J_{H_N H_\alpha}$ déterminées. On constate que les résultats issus de CSI et Talos donnent des résultats tout à fait cohérents, en particulier pour la prédiction des hélices α et des brins β , ainsi que la boucle entre les hélices α_5 et α_6 . Ces méthodes utilisant les mêmes données, il n'est donc pas surprenant qu'elles fournissent des informations proches.

D'autre part, la mesure du couplage scalaire doit être liée à la structure secondaire. On remarque en effet que la valeur de J est inférieure à 7Hz sur l'ensemble des hélices. De même, sur les deux brins β , cette valeur est supérieure à 9Hz. Les valeurs de J élevées sur la boucle reliant les hélices α_5 et α_6 indiquent que celle-ci adopte sans doute une conformation étendue.

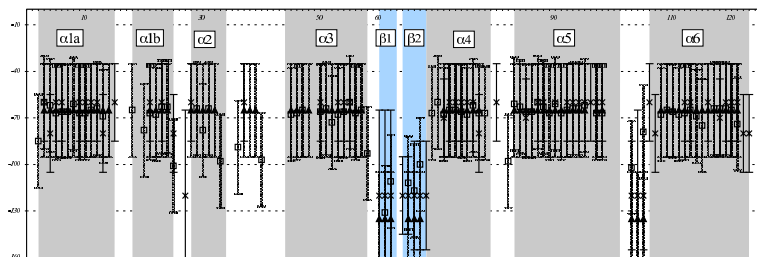


Fig. 3.4: Comparaison des trois méthodes de détermination des angles dièdres ϕ ($^{\circ}$). Les angles ϕ prédits ou déterminés par les méthodes CSI, Talos et HNHA sont respectivement représentés par un triangle, un carré et une croix. Les barres d'erreur de CSI et Talos sont en pointillés et de HNHA en trait plein.

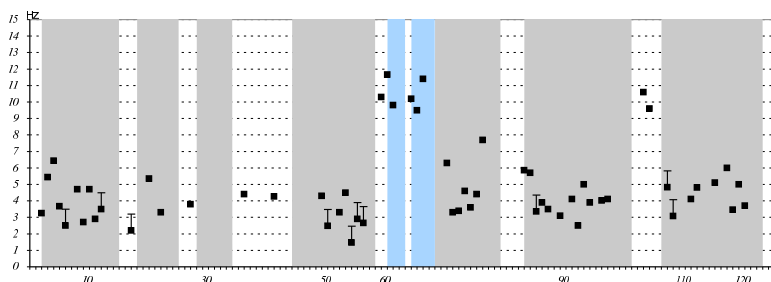


Fig. 3.5: Valeur de $^3J_{H_NH_\alpha}$ en fonction de la séquence d'ASP2. La barre d'erreur supérieure signifie que la valeur de J affichée est une valeur minimale.

La traduction en termes d'angles ϕ est aussi cohérente pour chaque résidu sur l'ensemble des trois méthodes.

3.1.2 Les liaisons hydrogènes

3.1.2.1 Echange H/D

L'échange H/D met en évidence l'implication d'un proton amide dans une liaison hydrogène. En dissolvant ASP2 doublement marquée ^{13}C - ^{15}N dans D_2O (tube 1 à tube 2), j'ai suivi par une série d'HSQC la disparition des signaux des protons amides ^1H (figure 3.6). Une partie des protons amides a disparu quasiment immédiatement (dans les premières minutes). Une autre partie est visible pendant quelques heures. Au bout de 12 heures, il reste environ une

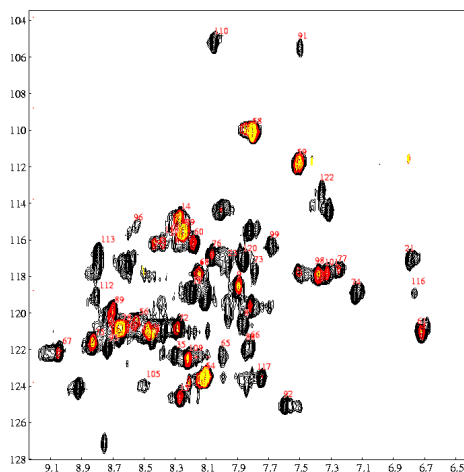


Fig. 3.6: Superposition de trois HSQC d'ASP2 fraîchement redissoute dans D_2O à 308K. En noir, après 20mn, en rouge, après 6h, en jaune, après 13h. Les taches attribuées de façon non ambiguë sont annotées.

Résidu (H^N)	Partenaire (O)	Résidu (H^N)	Partenaire (O)	Résidu (H^N)	Partenaire (O)	Résidu (H^N)	Partenaire (O)
A8	D4	K56	A52	M74	P70	E98	N94
K9	T5	I58	V54	I75	V71	N99	E95
E12	A8	E59	?	E76	Y72	K101	I97 ¹
L14	Y10	M60	M55 ¹	V77	K73	G110	E106
M15	M11	L61	M55 ¹	V78	M74	K112	N108
D17	L14 ¹	K62	E65 ¹	H79	I75 ²	Y113	I109
M19	M15 ¹	E65	K62 ²	L88	D84	C116	K112
C21	I18 ¹	L66	?	V89	D85	Y117	Y113
A22	M19 ¹	Y67	M60 ²	G91	Q87	I118	D115
D23	P20 ¹	V68	?	I92	L88	K120	C116
C53	C49	Y72	V68 ²	A93	V89	L121	Y117
V54	L50	K73	E69 ²	C96	I92	F122	I118

Tab. 3.1: Les protons amides en échange lent et leurs partenaires de liaison hydrogène. L'exposant ¹ signifie que le partenaire du proton amide a été déterminé à partir de la structure sans avoir été introduit ni comme contrainte au préalable, ni pour la structure finale. L'exposant ² signifie que le partenaire a été déterminé pendant la convergence et que la contrainte a été introduite après ce stade. Pour les autres résidus, le partenaire a été déterminé par l'existence d'une hélice α et la contrainte a été utilisée dès le début des calculs.

dizaine de taches. L'attribution d'un bon nombre des protons amides restant au bout de 20min est aisée, mais une dizaine de taches sont restées non attribuées du fait de superposition. De plus, quelques taches ont bougé, sans doute du fait d'une variation de pH lors des étapes de lyophilisation et de redissolution.

Pour les résidus en échange lent (disparition après 20min) et attribués sans ambiguïté, j'ai introduit dans le calcul de structure une liste de contraintes de liaison hydrogène. Pour les résidus au coeur des hélices définies par CSI et Talos, le partenaire du proton amide en échange lent est l'oxygène du carbonyle du résidu en i-4. Pour le feuillet β , les partenaires ont été déterminés après convergence suffisante des structures. Le partenaire des résidus aux extrémités des hélices n'a été déterminé qu'en toute fin de calcul. De la même façon, la première hélice semblait dès le départ de l'étude être coudée (analyse CSI et présence des deux prolines en position 16 et 20), par conséquent les partenaires des résidus 17 à 23 n'ont été déterminés que tard dans la convergence des structures.

J'ai obtenu l'attribution des taches ambiguës sur la base des structures préliminaires. Par exemple, les résonances des résidus A83 et K62 sont superposées dans la HSQC. Les structures préliminaires nous indiquaient que la lysine 62 était impliquée dans un β -hairpin alors que l'alanine 83 est dans une boucle. Par conséquent, j'ai attribué la tache à la lysine K62. Cependant, pour ne pas introduire de biais trop important, je n'ai pas utilisé ces contraintes dans les calculs de structure.

A l'exception de E59, L66 et V68, tous les partenaires des protons lentement échangeables ont pu être déterminés. Les résultats sont recensés dans le tableau 3.1.

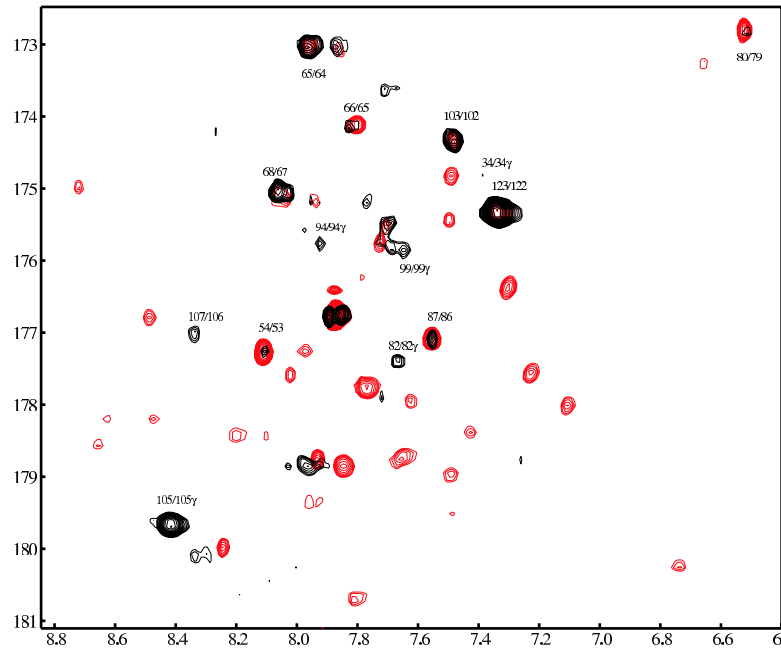


Fig. 3.7: Superposition du premier plan $^1\text{H}-^{13}\text{CO}$ des expériences HNCOhb (en noir) et HNCOhb_réf (en rouge) à 308K. Les taches attribuées correspondent aux corrélations N^i-C^{i-1} non refocalisées ($N^{65}-C^{64}$ par exemple) et aux corrélations $N^i-C'_\gamma$ ($N^{105}-C'_\gamma$ 105).

3.1.2.2 Expérience HNCO longue distance

Afin de lever les ambiguïtés sur les partenaires impliqués en liaison hydrogène et compléter l'étude de l'échange isotopique, j'ai mis en oeuvre la technique mise au point par F. Cordier et S. Grzesiek. Etant donné le faible nombre de signaux attendus dans les spectres, je n'ai accumulé que le premier plan de chacune des dimensions indirectes (plans HN et HCO). Les spectres sont représentés sur la figure 3.7. Les corrélations correspondant aux liaisons hydrogènes sont en principe présentes sur le spectre HNCOhb et absentes du spectre HNCOhb_réf.

Les spectres montrent en effet l'existence de quelques taches susceptibles d'être des liaisons hydrogènes. Si la plupart de ces taches sont de faible intensité et peuvent être considérées éventuellement comme du bruit, la résonance à (8,4ppm ; 179,6ppm) est très intense. Celle-ci doit donc correspondre à un transfert par couplage scalaire faible d'un azote à un carbonyle. Le plan HN de l'expérience HNCOhb identifie l'hydrogène à l'aspartate 105 mais la recherche du carbonyle couplé à l'azote de 105 resta infructueuse sous l'hypothèse que ce carbonyle était sur le squelette.

J'ai donc supposé que les taches visibles dans le spectre HNCOhb pouvaient correspondre à une corrélation entre l'azote et le carbonyle ou le carboxyle des chaînes latérales des aspartate, asparagine, glutamate et glutamine. Pour cela, il fallait disposer de l'attribution des carbones. En diminuant par deux le délai des INEPTs de la HNCO classique, le transfert des aminos vers l'azote qui les porte est optimisé. De cette manière, j'ai obtenu l'attribution des carbonyles C'_γ et C'_δ des asparagines 40, 82, 94, 99 et 108 et des glutamines 3 et 46. A partir ce moment, il s'est avéré que je pouvais attribuer une partie des taches de faible intensité à des corrélations

Résidu	${}^3J_{NC'_\gamma}$ (Hz)	χ_1 (°)
N82	0,43	-41 ± 28
N94	0,42	-103 ± 27
N99	0,32	-178 ± 7
D105	0,66	-176 ± 2

Tab. 3.2: Valeur des constantes de couplage ${}^3J_{NC'_\gamma}$ mesurées et l'angle dièdre χ_1 correspondant dans la structure.

$N^i-C'_\gamma$ intra-résiduelles au sein des résidus N34, N82, N94 et N99. L'absence de corrélations $N^i-C'_\delta$ pour les glutamines s'expliquent simplement par une constante de couplage plus faible.

Il restait encore à attribuer la tache la plus intense. J'ai alors considéré que ce signal provenait du couplage 3J entre l'azote et le carbone γ de la chaîne latérale de l'aspartate D105. Pour confirmer cette hypothèse, j'ai obtenu l'attribution du carbonyle au moyen de l'expérience COCAH optimisée pour observer les transferts $C'-C_\beta-H_\beta$. J'ai donc pu attribuer cette corrélation au couplage $N-C'_\gamma$ de l'aspartate 85. Les carbonyles des autres aspartates ont été aussi attribués mais ces aspartates ne donnent aucune corrélation $N^i-C'_\gamma$.

Pour expliquer la grande valeur du couplage $N^i-C'_\gamma$, on peut imaginer que la chaîne latérale de D105 est figée dans une conformation favorable au couplage. L'analyse des structures montre que D105 est situé dans la boucle entre les hélices α_5 et α_6 et que sa chaîne latérale possède une conformation très bien définie. L'angle dièdre $N-C_\alpha-C_\beta-C'_\gamma$ (χ_1) est proche de 180° . L'évolution de la relation de Karplus nous indique que le couplage scalaire est maximal pour cette valeur d'angle dièdre. La recherche de liaison hydrogène par des critères géométriques (distances, angles) par Molmol indique que dans plus de la moitié des structures, il existe une liaison hydrogène entre l'amide du résidu 108 et un des deux oxygènes de l'acide terminal de D105. L'existence d'une liaison hydrogène entre N^{108} et C_γ^{105} fige la conformation de la chaîne latérale de D105. Celle-ci est alors probablement à l'origine de la forte valeur de couplage observé. Peu après notre analyse, Banci (Banci et al. [2002]) a rapporté également l'observation de ces corrélations intrarésiduelles, alors que les premières publications n'en font pas état explicitement.

Nous pouvons estimer la valeur de la constante de couplage en utilisant la relation 3.6 (partie Matériel et Méthodes). J'ai pu ainsi obtenir les constantes de couplages intrarésiduelles pour les résidus N82, N94, N99 et D105 (tableau 3.2). Les angles χ_1 rapportés dans le tableau sont calculés par Molmol à partir des structures finales.

Ces valeurs sont à mettre en lumière avec les constantes de couplage 3J mesurées auparavant sur l'angle dièdre χ_1 . Après une recherche dans la littérature, je n'ai pas trouvé de référence donnant une estimation des couplages ${}^3J_{NC'_{ali}}$ sur les résidus acides ou amines. Cependant, Hu (Hu et al. [1997]) a mesuré les constantes de couplages ${}^3J_{NC'_\gamma}$ sur les résidus aromatiques. Il trouve une corrélation entre ce couplage et la valeur de χ_1 . Pour les résidus aromatiques bien définis dans la structure, l'angle dièdre χ_1 est soit trans ($\chi_1 = 180^\circ$) et ${}^3J_{NC'_\gamma} \approx 2,4Hz$, soit gauche ($\chi_1 = -60^\circ$) et ${}^3J_{NC'_\gamma} \leq 0,5Hz$. Cette corrélation n'est pas évidente pour nos mesures où les liaisons en trans (D105 et N99) ne sont pas associées à des valeurs de couplage aussi élevées que 2,4Hz (tableau 3.2).

Nous avons donc échoué dans l'objectif que nous nous étions fixé (déterminer les liaisons hydrogènes). Cette méthode semble avoir pourtant porté ses fruits sur plusieurs protéines, et de taille plus importante. Il est possible que le délai T ne soit pas correctement optimisé.

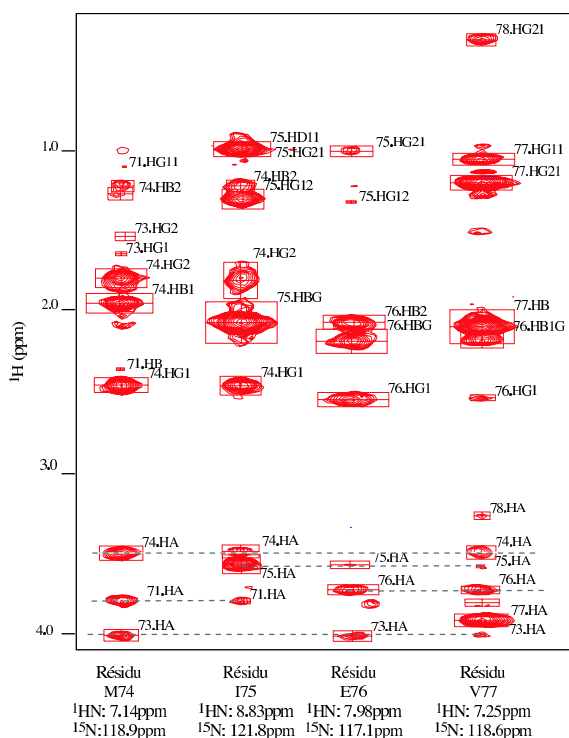


Fig. 3.8: Bandes extraites de l'expérience 3D-NOESY-HSQC aux déplacements chimiques H_N et N des acides aminés 74 à 77. Les taches de corrélation caractéristiques d'une hélice ($H_N^i \rightarrow H_\alpha^{i-3}$ et $H_N^i \rightarrow H_\alpha^{i-4}$) se retrouvent sur les traits en pointillés.

Nous avons également essayé la version TROSY (Wang et al. [1999]) de l'expérience mais sans succès. Il n'apparaît aucune nouvelle corrélation. De plus, la suppression d'un des deux doublets dans la dimension d'acquisition est imparfaite. Cela s'explique par la relative petite taille d'ASP2 : la séquence de Wang sélectionne les doublets par relaxation et est donc plus adaptée à des protéines de taille importante.

3.1.3 L'effet nOe

J'ai exploité trois spectres utilisant des transferts de type Overhauser pour extraire des contraintes de distances : une NOESY éditée ^{13}C (dans D_2O , $\tau_m = 120\text{ms}$, tube 2), une NOESY éditée ^{15}N ($\tau_m = 150\text{ms}$, tube 3) et une NOESY dans D_2O ($\tau_m = 100\text{ms}$, tube 4). Toutes les expériences ont été menées à 800MHz. La première donne des contraintes entre hydrogènes de chaînes aliphatiques, la seconde contient les corrélations entre protons amides et tous les protons, tandis que la dernière contient les corrélations concernant les aromatiques. Chacun des spectres a été analysé sous NmrView. Un fichier d'attribution a été créé par spectre pour tenir compte des légères différences de déplacements chimiques dues aux conditions expérimentales (température et pH) sensiblement différentes. Un "peak-picking" propre a été effectué pour mener à des fichiers contenant respectivement 2094, 1696 et 372 pics de corrélations. Bien entendu, une partie de ces pics sont redondants, comme par exemple les corrélations H^N-H^N qui se retrouvent en double dans la ^{15}N -NOESY-HSQC.

La figure 3.8 illustre la qualité du spectre ^{15}N -NOESY-HSQC dans les régions peu encombrées. Les corrélations caractéristiques de l'hélice α présentes entre les résidus V71 et V77 se lisent directement sur le spectre.

3.2 Le repliement tridimensionnel d'ASP2 en complexe avec le TSP

3.2.1 La modélisation par Aria.

3.2.1.1 Paramétrisation d'ARIA

Pour mettre en oeuvre un calcul d'ARIA, il faut tout d'abord créer un fichier "new.html" qui définit les liens vers les fichiers de contraintes, les listes de pics et les fichiers d'attribution de la protéine correspondant à chaque expérience, ainsi que la tolérance pour chaque dimension. Le lancement d'ARIA sur ce fichier crée l'arborescence, transforme les fichiers d'entrée au format ARIA et crée un fichier run.cns, qui contient toutes les informations sur la paramétrisation de la modélisation (recuit simulé) et sur la procédure de l'attribution automatique.

J'ai utilisé par défaut les paramètres contenus dans ce fichier. Cependant, j'ai modifié les ponts disulfure (C21-C53, C49-C107 et C96-C116). Le fichier de paramètres énergétiques est "parallhdg5.3".

J'ai utilisé les trois listes de pics obtenus par peak-picking automatique dans NmrView sur les expériences 3D de type NOESY éditées ^{15}N et ^{13}C et l'expérience bidimensionnelle dans D_2O (pour les aromatiques). Pour le calcul de la matrice de relaxation, j'ai défini le champ (800MHz) et le temps de mélange des expériences. Le temps de corrélation τ_c fut d'abord fixé à 9ns. Puis, quand j'ai obtenu la valeur expérimentale par les expériences de relaxation, j'ai utilisé la valeur 4,87ns.

Lors des premiers calculs, je n'ai imposé aucune attribution aux pics nOes. Les structures calculées alors convergeaient très mal. La raison principale était un problème de tolérance trop faible associé à une attribution trop peu précise. En effet, le déplacement chimique d'un noyau varie d'une expérience à l'autre en fonction des paramètres expérimentaux (température, pH, résolution), de traitement (prédiction linéaire) et d'analyse (calibration) et varie également d'un pic à l'autre dans une même expérience (quelques centièmes de ppm pour le proton). Pour obtenir une attribution plus précise pour chaque expérience, j'ai donc attribué dans NmrView la plupart des pics non ambigus. A partir des attributions de nOe associé à un hydrogène, NmrView fait la moyenne des déplacements chimiques de cet hydrogène. Dans une seconde étape, j'ai donc utilisé les listes de pics partiellement attribuées et les fichiers d'attribution modifiés par NmrView. Lorsque les fichiers d'attribution furent plus précis, j'ai diminué la tolérance et utilisé les listes sans attribution des nOes pré-imposées. La tolérance sur les hétéronoyaux ^{15}N et ^{13}C est alors de 0,5ppm. La tolérance sur les déplacements chimiques des protons dépend de la résolution : 0,02ppm pour la dimensions d'acquisition et 0,03ppm pour les dimensions indirectes. Lorsque la structure a commencé à converger, j'ai utilisé comme structure de départ, une structure de la série de calculs précédente (/begin/aria_template.pdb).

En fin de processus, j'ai accumulé un spectre NOESY édité ^{15}N sur un tube d'ASP2 marqué ^{15}N de pH 6 à 308K et à 283K. Quelques taches de corrélation HN s'étant déplacées, j'ai pu ajouter quelques contraintes supplémentaires définies par un encadrement de la distance fonction de l'intensité de la tache observée.

En phase finale de calcul, j'ai utilisé les contraintes attribuées par Aria de façon non ambiguë (fichier unambig.tbl) pour relancer un calcul en mode données déjà calibrées (*already calibrated data*), en partant d'une structure étendue. Ce fichier contient des définitions de

Expérience	Total	Attribués	Attribution		Contraintes (unambig.tbl)		
			ambiguë	non ambiguë	Total	ambiguë	non ambiguë
3D-NOESY-HSQC ^{13}C	2094	1964 (94%)	687 (35%)	1277 (65%)	1006	197	809
3D-NOESY-HSQC ^{15}N	1696	1623 (96%)	490 (30%)	1133 (70%)	1007	114	893
NOESY dans D ₂ O	372	344 (93%)	105 (30%)	239 (70%)	163	9	154
Total	4162	3931 (95%)	1282 (33%)	2649 (67%)	2198	320	1878

Tab. 3.3: Statistiques de l'attribution des pics nOe par Aria. Les trois premières colonnes contiennent les nombres de pics (total par expérience, attribué et type d'attribution). Les pourcentages sont rapportés par rapport au nombre contenu dans la colonne précédente. Les trois dernières colonnes rassemblent les statistiques sur les contraintes finales de distances.

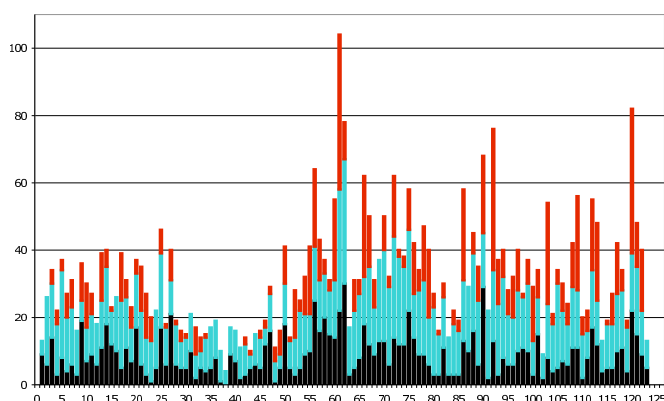


Fig. 3.9: Nombre de contraintes des distances (nOe) par résidu. En noir, les contraintes intrarésiduelles, en cyan, les contraintes entre deux résidus séparés par moins de 2 résidus, en rouge, les contraintes à longue distance.

distances ambiguës (ADR). Dans ces conditions, Aria n'est qu'une interface par rapport à CNS et n'interprète aucunement les données. Un jeu de 100 structures a été généré et j'ai conservé les 20 structures de plus basse énergie. Le module Analysis d'Aria permet d'analyser les 20 structures en termes de violation de contraintes géométriques et expérimentales et fournit une structure moyenne minimisée afin de réduire les déviations des paramètres géométriques (angles, longueur de liaisons) par rapport à une structure idéale.

3.2.1.2 Résultats de l'attribution des corrélations nOe et des contraintes de distances.

La grande majorité des pics obtenus par peak-picking sous NmrView ont pu être attribués par Aria, soit de manière ambiguë, soit de manière non ambiguë. Le tableau 3.3 rassemble les statistiques de l'attribution. Une moyenne de 95% des pics ont été attribués dans les fichiers .xpk (format NmrView). Les 5% restants non attribués correspondent à des pics de très faible intensité (qui peuvent alors être interprétés comme du bruit expérimental) ou à des pics associés à la superposition de plusieurs taches de corrélation (en particulier dans la région des chaînes latérales de Glu, Gln, Lys,...). Dans cette dernière situation, le centre du pic intense peut n'être corrélé à aucun couple de déplacement chimique d'hydrogènes potentiels. Plus des deux tiers des pics attribués l'ont été de manière non ambiguë (un seul couple d'hydrogènes en interaction).

Les trois dernières colonnes du tableau 3.3 rassemblent les statistiques sur le fichier unambig.tbl qui a servi pour le calcul final de la structure. Les contraintes définies dans ce fichier sont environ deux fois moins nombreuses que le nombre de pics attribués. A cela deux raisons, la première étant la redondance des corrélations nOe, en particulier dans l'expérience NOESY-HSQC ^{13}C , la seconde étant que la partie des contraintes ambiguës (ADR) dont Aria n'a pu déterminer avec suffisamment de confiance l'attribution, n'a pas été utilisée (709 contraintes). Près de 2200 contraintes ont donc servi à la reconstruction finale d'ASP2. J'ai ajouté à ce fichier des contraintes supplémentaires dont les corrélations nOe sont superposées ou invisibles dans les cartes analysées par Aria. Elles concernent le segment 28-42. Je me suis basé sur une expérience NOESY-HSQC- ^{15}N à 283K. A cette température, la séparation des corrélations de la ^{15}N -HSQC permet d'isoler des résidus (D30, I31 et I34, Q36) auparavant superposés avec d'autres corrélations et certaines corrélations apparaissent. Sur la base de la propension des résidus 29 à 34 à former une hélice (Talos, intensité des corrélations $H_N^i-H_N^{i-1}$), j'en ai extrait la contrainte ($H_N^{34}-H_\alpha^{31}$). D'autre part, j'ai ajouté les contraintes ($H_N^{34}-H_{\gamma12}^{35}$) et ($H_N^{36}-H_\beta^{37}$). Par ailleurs, des corrélations très ambiguës impliquaient les protons amides des résidus D40 et G41 et une résonance à 1,43ppm susceptible d'être l'un des méthyles des A37 et A38, et le proton amide de G41 vers une résonance à 4,17ppm, susceptible d'être le proton H_α des résidus A37, A38 ou K39. Compte tenu de la propension de ce segment à former une hélice α , j'ai attribué ces nOe à des corrélations (i, i+3).

La figure 3.9 montre la distribution du nombre de contraintes impliquant chaque résidu : contraintes intrarésiduelles, à moyenne distance (H^i-H^j , avec $|i-j| \leq 3$) et à longue distance ($|i-j| \geq 4$). Le nombre moyen de contraintes par résidu est de 32. La majorité des résidus sont impliqués dans des contraintes à moyenne et longue distance ce qui traduit la bonne structuration globale de la protéine. Par contre, sur le segment 28-45, le nombre total de contraintes est beaucoup plus faible et aucune contrainte à longue distance n'implique le segment 35-41. A l'opposé, les résidus L61, K62, I92 et L121 fournissent de très nombreuses contraintes (plus de 75). Ces résidus seront donc importants pour le repliement de la structure.

3.2.1.3 Les 20 structures de plus basse énergie : les violations de contraintes et carte de Ramachandran.

Parmi les 20 structures de plus basse énergie, seules 4 structures présentent des violations de nOe supérieures à 0,5Å (au maximum 4 violations par structure). 13 structures violent au moins une contrainte diédrale au-dessus de 5° (au maximum 5 contraintes de ce type sont violées par structure). Sur les 20 structures de plus basse énergie, analysées dans le répertoire *analysis*, il ressort que le nombre total de violations de contraintes de distances varie de 85 à 123. Ces contraintes se répartissent entre les taches de corrélation nOe (de 65 à 107 violations) et les liaisons hydrogènes (de 12 à 23 violations par structure). 92% des contraintes violées sur ces 20 structures correspondent à des distances supérieures à l'intervalle autorisé (distance effective supérieure à la borne supérieure de l'intervalle). Aucune de ces violations n'est supérieure à 0,3Å.

La méthode utilisée par Aria pour définir les intervalles de distances est plus restrictive que celle utilisant les intervalles fort-moyen-faible. Par conséquent, le nombre de violations, relativement important, n'est pas inquiétant dans la mesure où on pourrait sans trop de risque élargir légèrement les intervalles.

Type d'énergie	angles impropres	angles de liaison	angles dièdres	longueur de liaison	Interaction Van der Waals	distance (nOe+ liaison hydrogènes)	angles dièdres	Total par structure
	Contraintes géométriques					Contraintes expérimentales		
Minimum	6,1	48,4	604,5	3,45	65,35	8,2	2	744,8
Maximum	36,2	95,1	637,0	8,11	115,45	173,1	146,8	1045,9
Moyenne	11,8	61,8	616,7	5,13	83,8	37,1	17,9	834,1
rmsd moyen des violations par structure	0,28 (°)	0,3411 (°)	40,89 (°)	0,00164 (Å)		0,029 (Å)	0,93 (°)	

Tab. 3.4: Résumé des statistiques énergétiques des 20 structures d'ASP2 de plus basse énergie et des violations par type de contrainte.

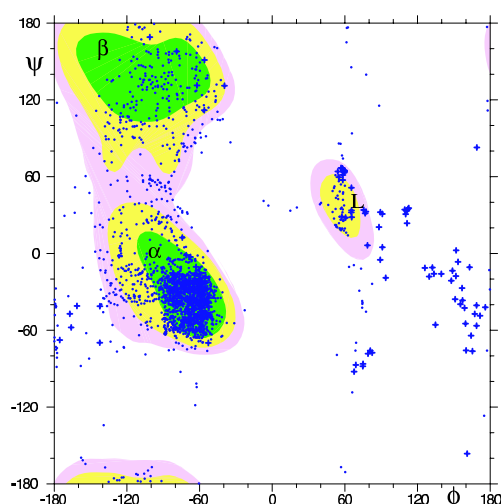


Fig. 3.10: Carte de Ramachandran des 20 meilleures structures. Chaque acide aminé des 20 structures est représenté par un point en fonction de ses angles ϕ et ψ . Les glycines sont représentés par une croix.

Le tableau 3.4 résume les statistiques sur l'énergie des 20 structures finales.

La carte de Ramachandran (figure 3.10) est une représentation des angles dièdres des acides aminés au sein d'une protéine. Trois zones se distinguent correspondant aux angles dièdres attendus dans les hélices α (zone α), les structures en brin β (zone β) et les boucles (zone L). Le logiciel Molmol a défini ces trois zones à partir de l'étude de 378 structures cristallographiques obtenues à haute résolution. La couleur des zones correspondent à la densité des points : 80% (respectivement 95% et 98%) des résidus sont dans la zone verte (respectivement jaune et violette). A l'exception des glycines, la grande majorité des résidus ($20 \times 123 = 2460$ résidus au total) sont dans les zones autorisées. On constate une très forte densité dans la zone α , corollaire de la structure en α d'ASP2. Les résidus situés dans les zones non autorisées (en blanc) sont localisés sur la structure dans le segment 26 à 45.

L'analyse par Procheck indique que 80% des résidus d'ASP2 sont dans la région centrale, 17,8% dans la région acceptable (jaune) et 1,67% dans la région acceptable étendue. 0,5% des résidus sont situés hors des zones acceptables.

Segment	Atomes	écart quadratique moyen
1-23, 49-123	bb	0,71Å
1-23, 49-123	heavy	1,09Å
24-48	bb	2,46Å

Tab. 3.5: Valeurs des écarts quadratiques moyens de différentes sélections d'atomes (bb : squelette, heavy : hétéroatomes) sur l'ensemble des 20 structures.

3.2.2 La structure d'ASP2 en complexe avec le TSP.

3.2.2.1 Un ensemble de 20 structures

Les 20 structures de plus basse énergie sont représentées sur la figure 3.11. A l'exception, du segment 24-48, la structure est très bien définie avec un écart quadratique moyen de 0,71Å sur les atomes du squelette. Compte tenu de l'absence de contraintes à longue distance engageant les hydrogènes du segment 34 à 45, celui-ci est très mal défini et son orientation par rapport au coeur structuré de la protéine ne peut être précisée. La figure 3.12 montre le squelette du segment 24-48 des 20 meilleures structures, superposés sur 24-48 (écart quadratique moyen de 2,46Å). Une autre façon de visualiser la mauvaise définition de ce segment est de représenter l'écart quadratique moyen de chaque résidu le long de la séquence par rapport à la structure moyenne (figure 3.13). Si pour les résidus T5 à A21 et de C49 à L121, l'écart quadratique moyen par résidu est inférieur à 1Å, il augmente considérablement à partir de la fin de l'hélice α_{1b} pour atteindre une valeur maximale pour le résidu K39 avant de revenir progressivement à 1Å. Ce profil est typique d'un mouvement autour d'une charnière formée par le pont disulfure C21-C53, par le peu de contraintes à longue distance impliquant les résidus compris entre les deux cystéines et par l'absence totale de telles contraintes impliquant les résidus centraux. On remarque également la dispersion conformationnelle des cinq résidus en position N-terminale, des deux derniers résidus de la séquence et du résidu A100.

3.2.2.2 Les structures secondaires

Les hélices α On peut définir six hélices, dont la première est coudée. Afin de conserver la nomenclature de la structure de la PBP de *Bombyx mori* à pH 6 élucidée par RMN (Lee et al. [2002]), ces hélices ont été dénommées α_{1a} (3-13), α_{1b} (19-24), α_2 (31-35), α_3 (44-57), α_4 (70-78), α_5 (82-99) et α_6 (106-121). La première hélice α est coudée du fait de la présence des deux prolines P16 et P20. Le coude forme un angle de 135°.

β -hairpin et coudes Les résidus M60 à K62 et E65 à Y67 impliqués dans des brins β forment un β -hairpin (figure 3.14), le coude β défini par les résidus 62 à 65 est de type I' comme le montrent les nOes et la position des résidus 63 et 64 dans la carte de Ramachandran, représentés sur la figure 3.15. Le segment peptidique reliant les hélices α_3 et α_4 est composé de quatre résidus (H79-A80-G81-N82). Une analyse de la conformation du squelette montre qu'ils forment un coude. Malgré la distance qui sépare le proton amide de N82 de l'oxygène du carbonyle de H79 (entre 3,2 et 4,3Å), importante pour une liaison hydrogène, les angles dièdres des résidus A80 et G81 ainsi que la présence de glycine en position L2 indiquent que ce coude est de type II.

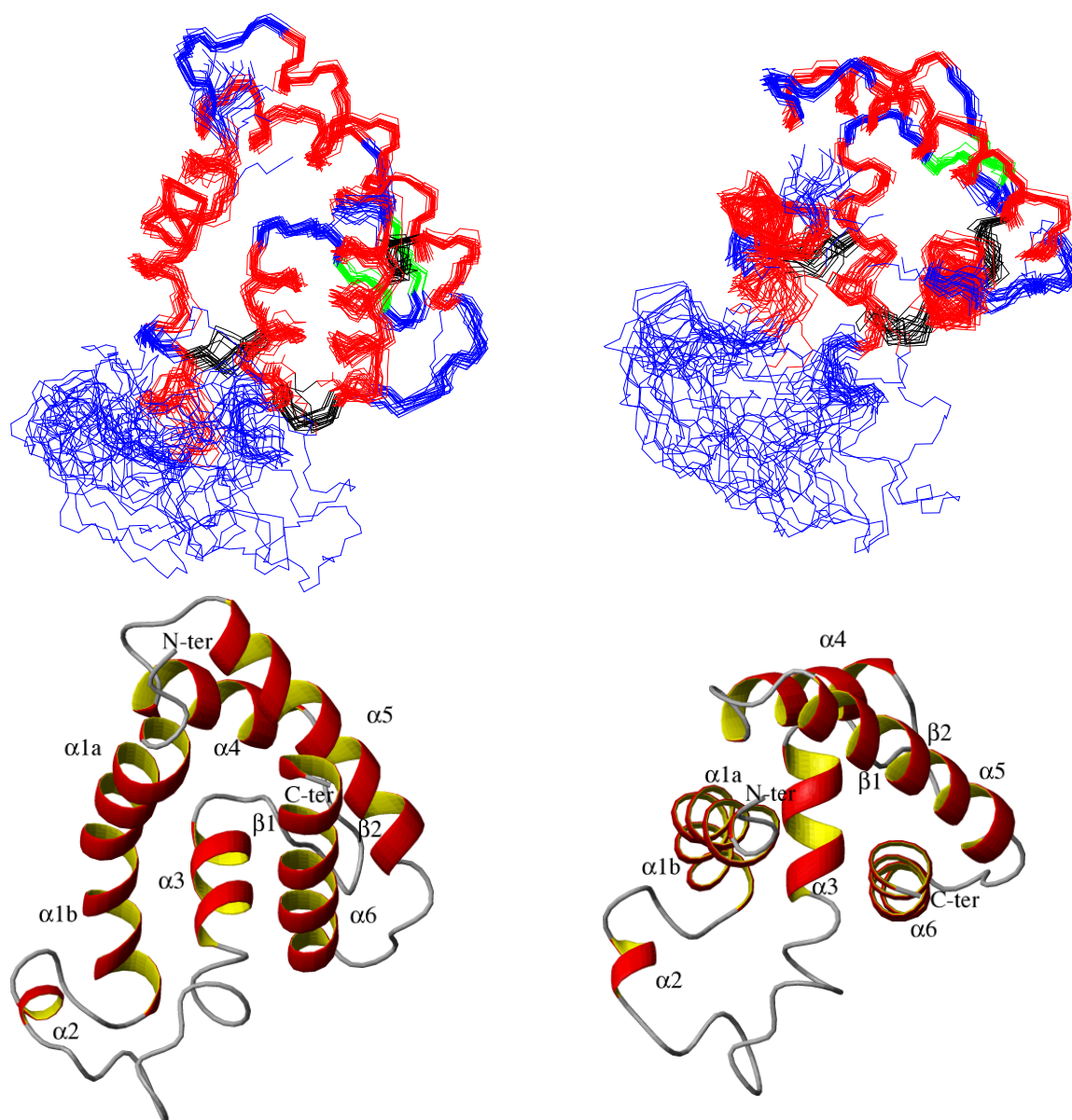


Fig. 3.11: Représentations de la structure d'ASP2 en complexe avec le TSP. En haut, superposition des 20 structures de plus basse énergie entre les résidus 1 à 23 et 49 à 123. Seuls le squelette peptidique et les ponts disulfure (en noir) sont représentés. Les hélices α sont en rouge et les acides aminés en structure β sont en vert; en bas, représentation des structures secondaires de la structure de plus basse énergie. Les figures de droite correspondent à une rotation d'environ 90° des figures de gauche.

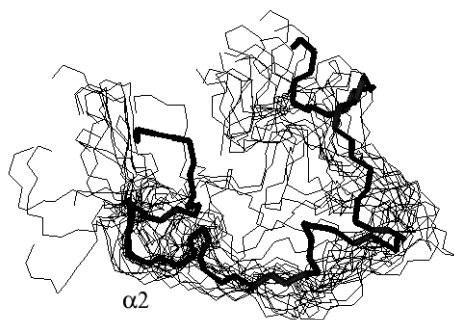


Fig. 3.12: Superposition du segment 24-48. Seul le squelette de cette région est illustré.

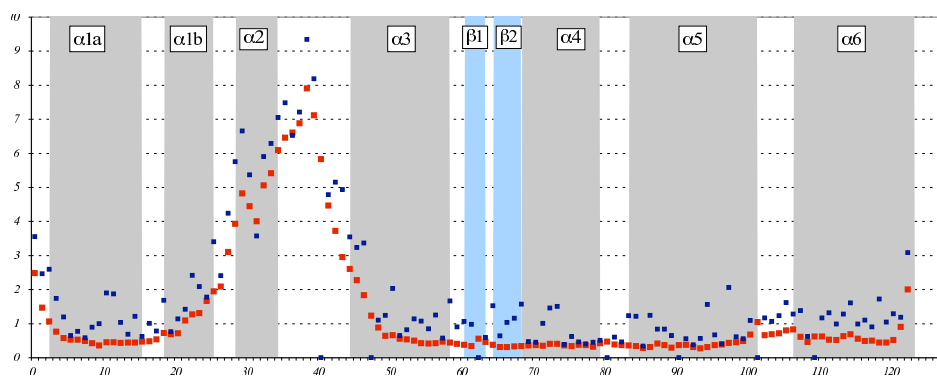


Fig. 3.13: Écart quadratique moyen par résidu (\AA) des atomes du squelette (en rouge) et des hétéroatomes de chaîne latérale (en bleu) en fonction de la séquence.

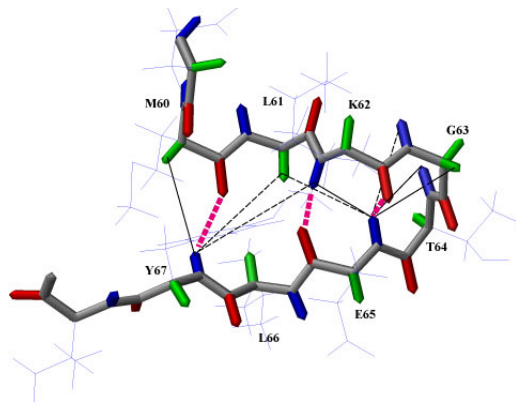


Fig. 3.14: Corrélations nOe observées au sein du β -hairpin entre les protons amides (en bleu), les hydrogènes H_α (en vert) et liaisons hydrogènes entre les oxygènes (en rouge) et les protons amides. En trait plein, les corrélations fortes, en pointillés, les corrélations faibles.

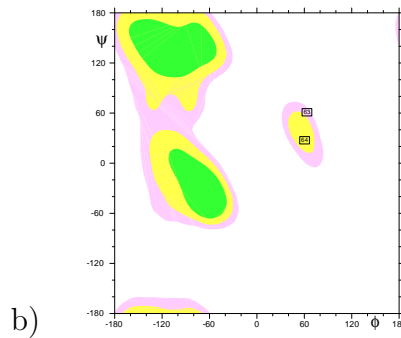
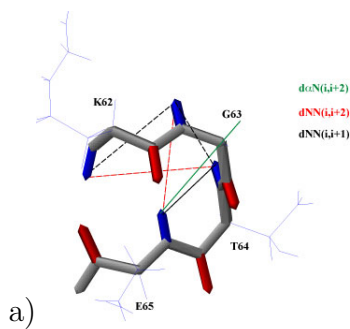


Fig. 3.15: Coude β défini par les résidus 62 à 65. En a), conformation des résidus 62 à 65 d'ASP2, en b), position des résidus 63 et 64 dans la carte de Ramachandran.

Isomérisation des prolines La superposition des prolines des 20 structures (figure 3.16) indique que les trois prolines P16, P20 et P70 sont en conformation trans.

3.2.2.3 Arrangement des hélices α .

La figure 3.17 rassemble les nOes observés en fonction de la séquence. La majorité des corrélations concerne bien entendu des résidus séquentiels ou proches dans la séquence. Outre les trois ponts disulfure, un grand nombre de nOes (points loin de la diagonale) assurent le repliement des structures secondaires d'ASP2.

Les nOes impliquant des résidus proches séquentiellement des cystéines permettent d'orienter les hélices reliées par les ponts disulfure les unes par rapport aux autres ($\alpha_{1b} - \alpha_3$, $\alpha_3 - \alpha_6$

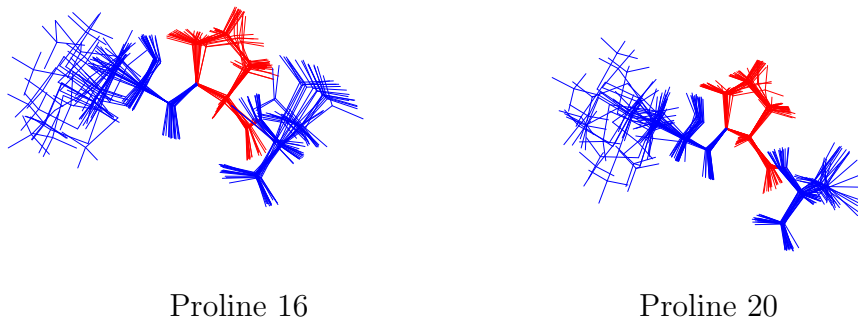


Fig. 3.16: Isomérisation des Prolines 16 et 20

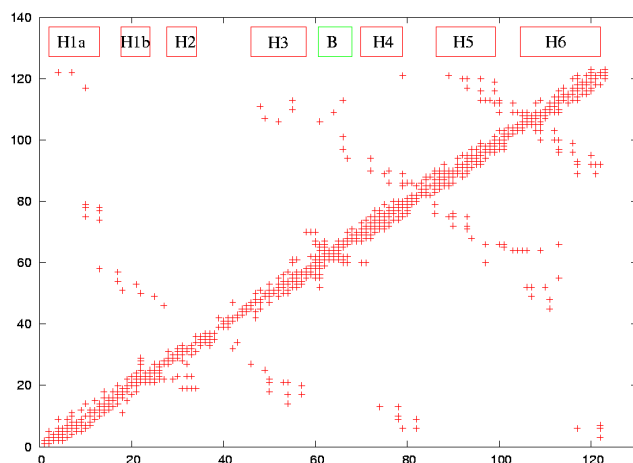


Fig. 3.17: Distribution des nOes non ambiguës en fonction de la séquence. Les nOes hors diagonales (longues distances) sont responsables de la structure tertiaire de la protéine.

	α_{1a}	α_3	α_4	α_5	α_6
α_{1a}		70,9°	102,5°	75,4°	148°
α_3	70,9°		95,7°	109,1°	92,7°
α_4	102,5°	95,7°		152,4°	50,6°
α_5	75,4°	109,1°	152,4°		136,6°
α_6	148°	92,7°	50,6°	136,6°	

Tab. 3.6: Angles formés par les hélices d'ASP2.

et $\alpha_5 - \alpha_6$). Les paires d'hélices non reliées par pont disulfure sont orientées par des nOes à longue distance. Ainsi, l'extrémité N-terminale de l'hélice α_1 montre des corrélations avec les résidus I75, V78 et H79 de l'extrémité C-terminale de l'hélice α_4 . En particulier, un des deux méthyles de la valine 78 a un déplacement chimique négatif qui s'explique par la proximité spatiale avec les deux cycles aromatiques des tyrosines 10 et 13. La fermeture de la structure dans cette région est montrée par les nOes présents entre les résidus Q3 à V7 et le cycle aromatique de la phénylalanine 122 de l'hélice α_6 C-terminale. Par ailleurs, les résidus de la région structurée en β -hairpin montrent plusieurs corrélations avec les chaînes latérales des résidus de la boucle liant les hélices α_5 et α_6 (T64-L66 vers K101-D105).

Les hélices α_{1a} et α_{4-6} forment un faisceau d'hélices antiparallèles convergeant vers les extrémités N- et C-terminales. L'autre extrémité du faisceau est fermée par l'hélice α_3 , fortement protégée du solvant d'un côté par l'hélice α_{1b} et de l'autre côté par le β -hairpin et l'hélice α_6 . Les hélices α_{1a} et α_{4-6} sont quasiment parallèles tandis que l'hélice α_3 a une orientation globalement perpendiculaire aux autres (tableau 3.6). Le logiciel Molmol a détecté deux liaisons hydrogènes rigidifiant localement la structure et impliquant des chaînes latérales : $NH_{\zeta}^{K62}-O_{\epsilon}^{E59}$ et $NH_{\delta}^{N82}-O_{\delta}^{D85}$.

La figure 3.18 décrit la position des résidus polaires autour de l'axe des différentes hélices. La partie des hélices exposée au solvant est relativement riche en amines et en acides. C'est particulièrement vrai pour l'hélice α_6 dont l'amphiphilie est presque parfaite.

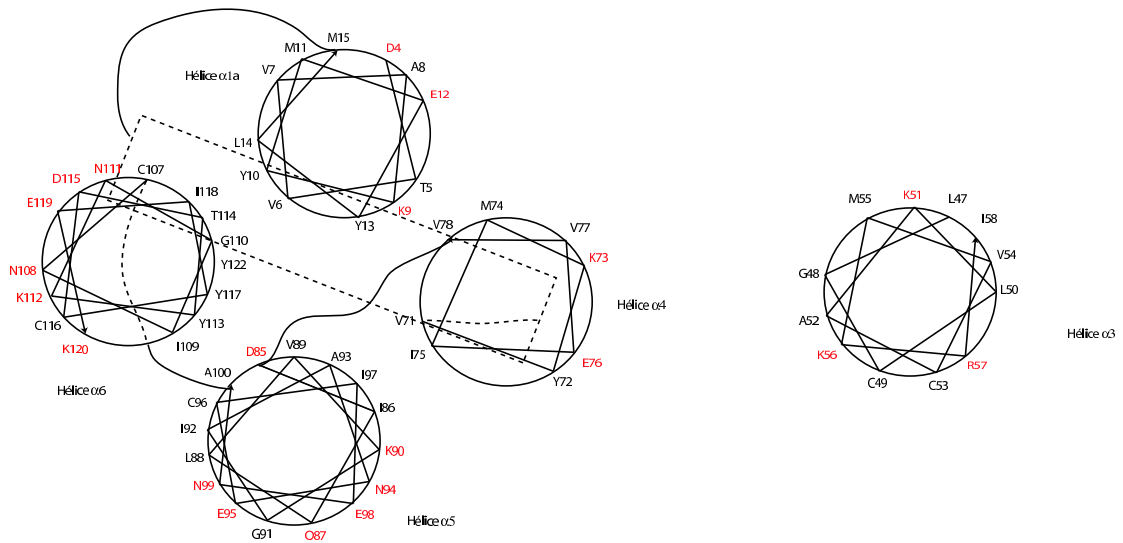


Fig. 3.18: Orientation des résidus autour des hélices. En rouge, les résidus polaires. Les hélices $\alpha_1 - \alpha_{4-6}$ sont vues suivant une vue de dessus. Le rectangle en pointillé représente grossièrement la position de l'hélice α_3 .

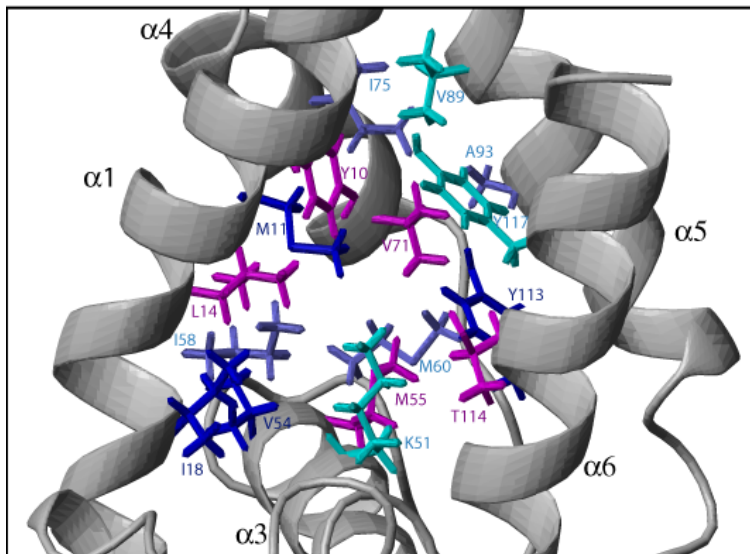


Fig. 3.19: Vue sur la poche hydrophobe. Les chaînes latérales formant la poche sont représentées.

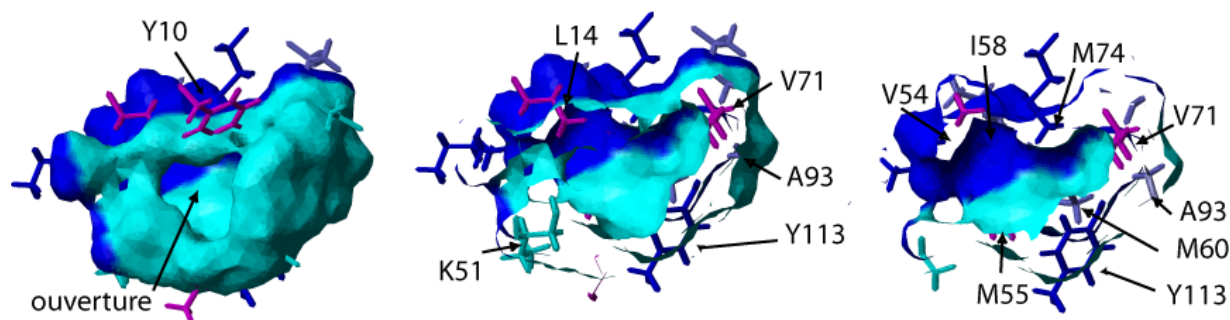


Fig. 3.20: Une plongée dans la poche hydrophobe. Le point de vue est identique à celui de la figure 3.19.

3.2.2.4 Un coeur hydrophobe

Le repliement d'ASP2 montre que l'arrangement des hélices crée une poche entourée de résidus essentiellement hydrophobes à l'exception de la lysine 51 (figure 3.19). Les atomes impliqués dans la formation de la surface interne sont les suivants : H_{ϵ}^{Y10} , H_{OH}^{Y10} , SMe^{M11} , Me_{δ}^{L14} , Me_{δ}^{I18} , $H_{\alpha,\delta,\epsilon,\zeta}^{K51}$, $H_{\alpha,\gamma}^{V54}$, SMe^{M55} , Me_{δ}^{I58} , H_{γ}^{M60} , SMe^{M60} , Me_{γ}^{V71} , SMe^{M74} , $Me_{\gamma,\delta}^{I75}$, Me_{γ}^{V89} , H_{α}^{A93} , Me^{A93} , $H_{\delta,\epsilon}^{Y113}$, H_{OH}^{Y113} , $H_{\alpha,\beta}^{T114}$, H_{OH}^{T114} , H_{β}^{Y117} , $H_{\delta,\epsilon}^{Y117}$, H_{OH}^{Y117} . Ces atomes sont suffisants pour couvrir au maximum la surface interne. La plupart de ces atomes sont apolaires, à l'exception des hydroxyles des tyrosines 10, 113 et 117 et de la thréonine 114, concentrées dans une moitié de la cavité. Le segment 29-45 semble être impliqué dans la poche car dans les structures dont nous disposons, on observe une ouverture de la poche entre les résidus Y10, L14, I18, K51, T114 et I117. Si la conformation du segment mal défini mène à une structure plus fermée, il est probable que des chaînes latérales des acides aminés Q36 à N40 viennent refermer la poche.

La forme de la poche est globalement un pavé de côtés $7\text{\AA} * 7\text{\AA} * 3\text{\AA}$. La figure 3.20 visualise les chaînes latérales formant la poche par trois vues successives en partant de l'ouverture. La surface a été générée à partir des volumes de Van der Waals ($1,4\text{\AA}$). L'ouverture principale a une forme triangulaire de côté 3\AA . On peut remarquer sur la vue de droite la présence d'une petite ouverture entre les acides aminés V54 et M55 mais son diamètre est très faible. Le squelette au niveau de ces résidus ferme cette petite cavité. Elle ne doit pas avoir de rôle particulier dans le processus d'interaction avec un ligand.

3.2.2.5 Tenseur d'inertie et surface de potentiel électrostatique

Le tenseur d'inertie de la structure d'ASP2 La forme globale d'une protéine en solution peut être approchée par l'analyse des axes principaux. L'axe principal de plus grand moment rejoint les atomes les plus éloignés de la molécule. Une molécule dont les moments principaux sont de même longueur est sphérique. Le calcul par Molmol des axes principaux de l'ensemble des 20 structures d'ASP2 indique que les moments moyennés sur l'ensemble sont dans un rapport de (1,00 ; 0,59 ; 0,78). ASP2 présente donc une structure quelque peu éloignée d'une sphère parfaite. La rhomboïcité du tenseur d'inertie de la protéine vaut 0,60.

L'orientation des axes est illustrée sur la figure 3.21. Bien entendu, la forte dispersion de

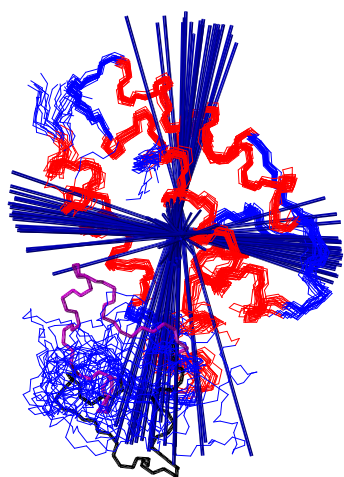


Fig. 3.21: Les axes principaux de la structure d'ASP2. L'axe principal le plus long est vertical.

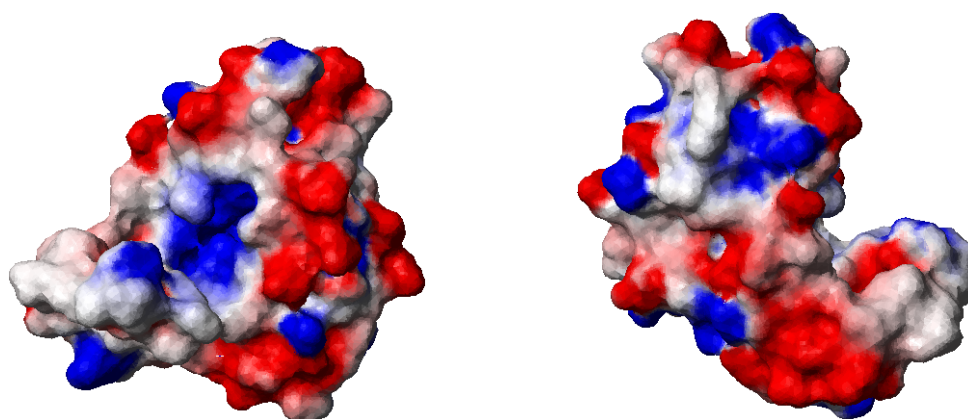


Fig. 3.22: Surface de potentiel électrostatique de la structure moyenne d'ASP2. Les deux vues sont prises de part et d'autre de la structure. Les potentiels positifs sont en bleu et les potentiels négatifs en rouge.

l'orientation des axes est due majoritairement à la conformation du segment 29-45. Pour la structure la plus fermée (segment coloré en magenta), les rapports de longueurs principales est (1,00 ; 0,63 ; 0,79) tandis que pour la structure la plus ouverte (en noir), ces rapports sont (1,00 ; 0,47 ; 0,68).

Les charges de surface de la structure d'ASP2. Le calcul du potentiel électrostatique de surface de la structure moyenne d'ASP2 (figure 3.22) indique que le contact entre le solvant et la protéine est assuré essentiellement par des chaînes latérales chargées (Glutamate, Aspartate et Lysine, Arginine). Lorsqu'on regarde la distribution de ces résidus sur la structure, à l'exception de la lysine K51, toutes les chaînes latérales chargées positivement sont orientés vers le solvant. Les zones de la surface au potentiel négatif (en rouge) sont dues aux chaînes latérales de Glu et Asp. A l'exception de la D85, dont le groupe acide n'est pas en contact avec le solvant, tous ces résidus sont orientés vers le solvant.

Les résidus chargés sont susceptibles d'être impliqués dans des ponts salins. D'après la structure, on peut envisager la présence de tels ponts entre les résidus K56 et E24, K62 et E59, K90 et E76, K112 et E103, NH_3^+ de I1 et D85.

3.3 Discussion

Nous venons de discuter des caractéristiques de la structure d'ASP2 en complexe avec le TSP. Le lecteur aura remarqué que le TSP n'a pas été introduit dans la modélisation. La modélisation avec le TSP n'aurait sans doute pas amélioré le calcul. En effet, le TSP est deutéré à l'exception des 3 groupes méthyles. Par conséquent, les contraintes possibles n'impliqueraient que ces méthyles et ne pourraient en aucun cas contraindre la chaîne aliphatique du TSP. Compte tenu de la résolution et du nombre de contraintes utilisées, il est raisonnable de considérer que la structure obtenue est bien celle d'ASP2 en complexe.

Le cas du segment 29-48 est à considérer plus avant. Les différentes cartes NOESY que j'ai accumulées contiennent des signaux provenant de cette région de la protéine, mais ceux-ci sont faibles, voire très faibles (résidus I31, T33, N34, I35). D'autre part, les déplacements chimiques de ces résidus sont tels que les quelques nOes engageant ces résidus ont mené à de très nombreuses d'ambiguïtés d'attribution, rendant le processus de reconstruction manuelle caduque sur cette région. L'utilisation d'Aria n'a pas fourni de nouvelles contraintes, et ce, sans doute à cause de l'algorithme d'attribution d'Aria : aucun poids n'est ajouté au terme d'énergie lorsqu'un pic est éliminé, et ce, pour pouvoir discriminer entre bruit et signal. Les quelques nOes pouvant a priori contraindre fortement l'orientation de ce segment par rapport au coeur de la protéine ont été éliminés ou attribués vers des contraintes à moyenne distance (structure secondaire). A ce moment de la reconstruction s'est donc posée la question : le segment 29-45 est-il sous plusieurs conformations ou en dynamique interne originale, menant à des signaux faibles, ou alors, serait-ce un exemple de limitation de la RMN à l'étude d'une protéine d'une taille respectable. Il se peut en effet que par malchance, les proximités spatiales n'ont pu être observées car superposées à d'autres. Afin de répondre à cette interrogation cruciale, le deuxième objectif initial de la thèse (étude de la dynamique interne d'ASP2) a pris tout son sens. Le chapitre suivant décrira les résultats des différentes expériences de relaxation menées sur le squelette d'ASP2 ainsi que leur interprétation et la signification de la structure au vu de ces résultats.

Chapitre 4

La dynamique d'ASP2 à 308K et à 283K

Une étude complète de la relaxation de l'azote ^{15}N du squelette a été menée à 308K, soit la température à laquelle la structure a été résolue. Pour analyser l'effet de la température, en particulier sur le segment mal défini de la structure, j'ai mené la même série d'expériences à 283K. Les résultats à 308K seront amplement détaillés (section 3.1 à 3.3) et la comparaison entre les deux températures sera discutée qualitativement dans la section 3.4. Nous verrons ensuite comment la dynamique du squelette peut être reliée à la structure obtenue.

4.1 Les données de relaxation R_1 , R_2 et l'effet nOe hétéronucléaire à 308K.

La mesure des constantes de relaxation du spin azote ^{15}N a été menée sur un tube d'ASP2 marqué ^{15}N , fraîchement dissoute dans $450\mu\text{l}$ d'eau (10% de D_2O) à $\text{pH}=6$ en l'absence de sel et avec le TSP dans un rapport de 5 :1.

Un échantillonnage correct de la courbe de relaxation T_1 impose de prendre des points répartis de façon homogène jusqu'à l'extinction du signal. Nous avons choisi de prendre les valeurs suivantes de δ : 12, 48, 72, 128, 300, 420, 600, 840 et 1200ms. Les spectres ont été accumulés dans le désordre afin d'évacuer toute erreur systématique. De plus, le point à 72ms a été doublé pour estimer l'erreur sur les intensités et pour vérifier qu'il y avait pas eu de dérive du champ ou de l'homogénéité du champ durant la durée de l'expérience. Le gain avant l'échantillonnage numérique doit être identique quel que soit le délai pour une comparaison possible des intensités relatives. L'échantillonnage de la courbe mono-exponentielle de relaxation transversale a consisté en 8 points : 8, 16, 32, 48, 64, 136, 256, 400ms. Les points à 48, 64 et 256ms ont été doublés. Comme $R_2 \geq R_1$, la décroissance est plus rapide dans le cas de la mesure de R_2 que de la mesure de R_1 , par conséquent le dernier point de R_2 est plus court (400ms) que pour R_1 (1200ms).

Les spectres ont été enregistrés sur le spectromètre Bruker 600MHz. Chaque expérience a été réalisée avec 128 points complexes en dimension indirecte et 1024 en dimension d'acquisition. Pour chaque t_1 , 16 passages ont été enregistrés. Un filtre sinus a été appliqué dans les deux dimensions. Chaque dimension a été complétée par zero-filling jusqu'à une taille de

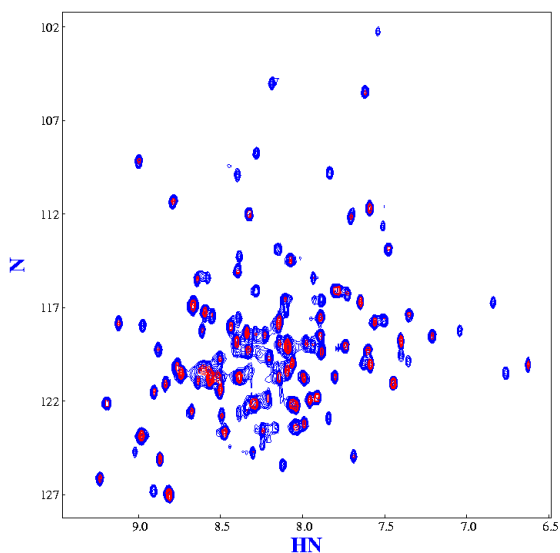


Fig. 4.1: Superposition de deux spectres de mesure de R_1 à 12ms (bleu) et à 600ms (rouge) à 308K. Les deux spectres sont représentés avec le même niveau d'affichage. L'intensité à long délai de relaxation est clairement plus faible qu'à court délai.

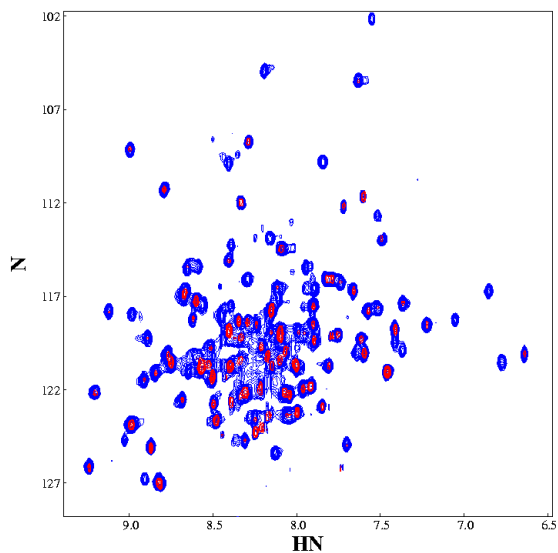


Fig. 4.2: Superposition de deux spectres de mesure de R_2 à 8ms (bleu) et à 256ms (rouge) à 308K. Les deux spectres sont représentés avec le même niveau d'affichage. L'intensité à long délai de relaxation est clairement plus faible qu'à court délai.

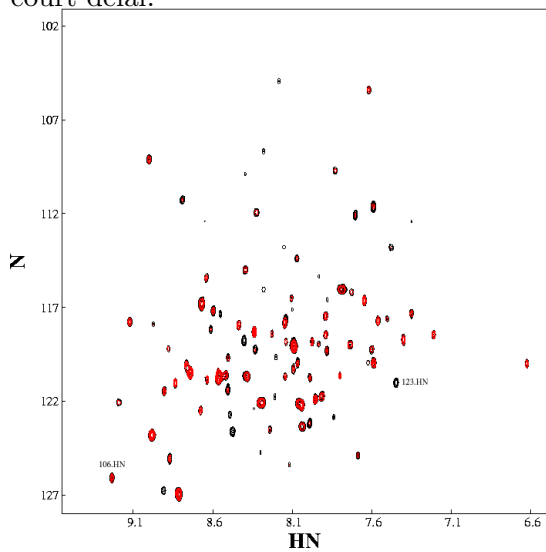


Fig. 4.3: Superposition de deux spectres de mesure de l'effet nOe hétéronucléaire à 308K représentés au même niveau. Seuls les niveaux positifs sont affichés. En noir, le spectre non saturé ; en rouge, le spectre saturé. La baisse d'intensité due à l'effet nOe est visible entre autres sur le résidu C-terminal (123) tandis que l'effet nOe de l'azote du résidu 106 est plus proche de la limite supérieure.

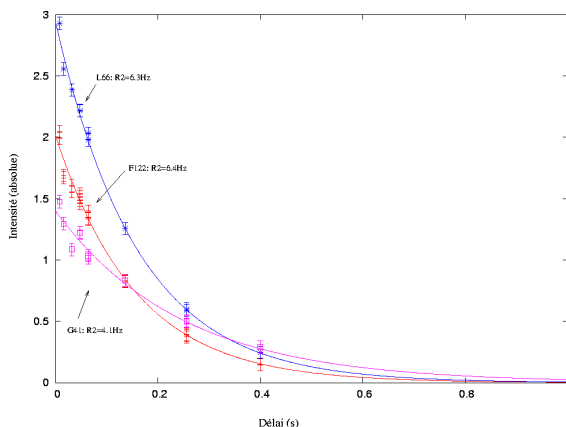


Fig. 4.4: Exemples d'ajustement de R_2 . La décroissance plus lente de l'intensité de la glycine 41 traduit un R_2 plus élevé que pour les résidus 66 et 122. La forte dispersion des intensités du résidu 41 conduit à une incertitude plus élevée sur le R_2 de ce résidu.

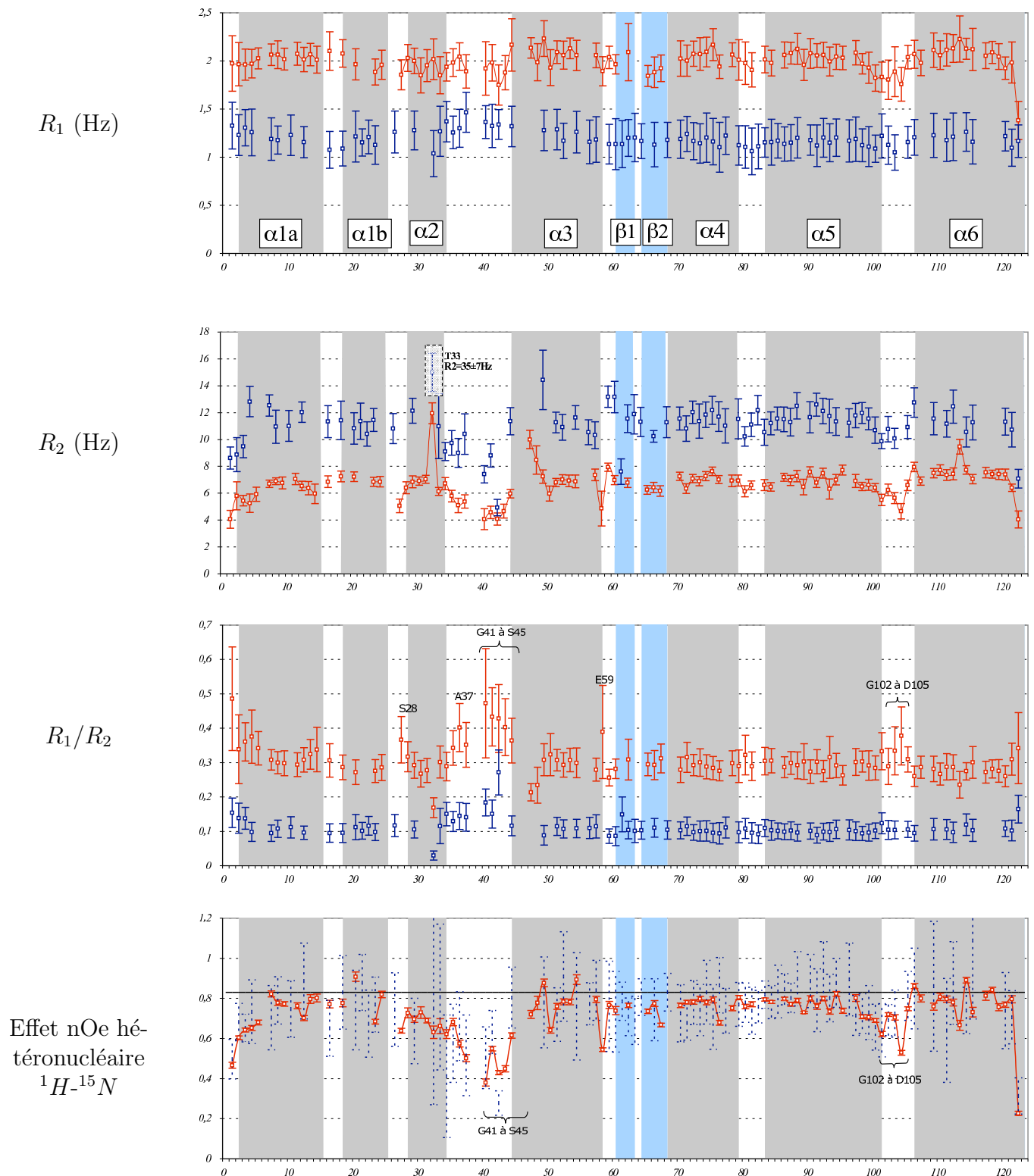


Fig. 4.5: Valeurs de R_1 , R_2 , R_1/R_2 et nOe hétéronucléaire d'ASP2 à 308K (en rouge) et à 283K (en bleu) à 600MHz. Valeurs en Annexe F.

256x2048 points avant transformée de Fourier et une correction de ligne de base a été effectuée. Le logiciel NmrPipe a servi pour le traitement.

Pour la mesure de l'effet nOe hétéronucléaire, j'ai fixé le délai d_1 d'établissement de l'équilibre sous irradiation des protons amides à 10s à 308K ($\sim 20T_1(N_z)$) et à 5s à 283K ($\sim 6T_1(N_z)$). Chaque spectre a été accumulé sur 32 passages pour obtenir une matrice complexe de 200*1024 points.

Les figures 4.1, 4.2 et 4.3 représentent quelques spectres accumulés pour la mesure des vitesses de relaxation et de l'effet nOe hétéronucléaire. La figure 4.4 donne trois exemples du fit des intensités mesurées sur les spectres de mesure de R_2 pour les résidus G41, L66 et F122.

En suivant les protocoles décrits dans la partie méthodologie, j'ai extrait les vitesses de relaxation longitudinale et transversale des spins azote 15 du squelette de 96 résidus. Un ajustement monoexponentiel a pu être obtenu pour tous les résidus, l'hypothèse qui néglige la relaxation croisée est donc justifiée. A pH 6, tous les résidus autres que N-ter et les prolines donnent une tache de corrélation, y compris les résidus de la région relativement flexible dans la structure obtenue. Pour des raisons de superpositions de déplacements chimiques, les paramètres de relaxation n'ont pu être obtenus pour les 22 acides aminés suivants : V7-I18-I27, M11-T64, A22-E69, D23-D40, K39-L47-I86, Q46-K90, K56-V78, R57-I109, K62-A83, I97-Y117. Par ailleurs, la tache de corrélation du résidu 26 était trop faible dans les spectres à long délai par rapport à son voisin N99 pour pouvoir donner des valeurs acceptables. Je l'ai donc exclu par la suite. Je disposais donc d'un jeu de paramètres de relaxation complet pour 96 résidus répartis de façon homogène le long de la séquence. Les valeurs de R_1 , R_2 et nOe hétéronucléaire sont rassemblés dans les figures 4.5.

La vitesse de relaxation longitudinale R_1 moyenne sur l'ensemble de la séquence est de 2Hz. L'écart-type moyen vaut 0,16Hz soit une incertitude moyenne de environ 8%. Excepté le résidu en position C-terminale, les valeurs de R_1 s'écartent peu de la valeur moyenne (entre 1,74 et 2,23Hz).

La vitesse de relaxation transversale R_2 moyenne sur l'ensemble de la séquence est 6,7Hz avec une incertitude moyenne de 0,45Hz, soit environ 7%. La dispersion des R_2 est beaucoup plus importante que les R_1 . On remarque qu'une majorité des résidus ont leur R_2 compris autour de la valeur moyenne de 7Hz. Les résidus en position N-ter (de 2 à 6) voient leur R_2 augmenter progressivement pour atteindre la valeur moyenne. Deux autres zones (de Q36 à K51 et de G102 à D105) comportent également des valeurs de R_2 significativement plus faibles que la moyenne. La première zone correspond à la seconde moitié de la région de la structure mal définie tandis que la seconde forme la boucle entre les deux dernières hélices α . Certains résidus ont des valeurs de R_2 se distinguant particulièrement de leurs voisins : ce sont les résidus T33, G48, C49 et T114 avec un R_2 très élevé (supérieure à 8Hz) et S28 et E59 (R_2 inférieur à 5Hz). La première série de résidus (R_2 élevés) est susceptible de présenter de l'échange chimique du fait d'un terme de relaxation supplémentaire tandis que les deux résidus de faible R_2 ont une dynamique interne importante dans la gamme pico- nanoseconde.

Les valeurs de nOe hétéronucléaire sont toutes inférieures à la limite maximale autorisée par la théorie excepté les résidus C21, L50, M55, C107 et E119 dont l'intervalle de confiance ne contient pas la valeur 0,83. On peut expliquer ces valeurs soit par une incertitude sous-estimée, soit par une saturation partielle de l'eau durant l'expérience de référence, menant à une sous-estimation de l'intensité de référence (Grzesiek and Bax [1993]). Pour les autres résidus, le nOe est compris entre 0,23 et 0,83. Il se dégage les mêmes tendances que pour les

R_2 : les extrémités N- et C-terminales, ainsi que les résidus de A32 à G48 et de N99 à D105 ont un nOe hétéronucléaire inférieur à 0,71. Ces zones de la protéine ont donc une dynamique interne importante sur la gamme de temps ns-ms.

4.2 Temps de corrélation à 308K

Il est intéressant de calculer le rapport R_1/R_2 . La grande majorité des résidus ont un tel rapport égal à environ 0,3. La région comprise entre les résidus A37 à L50 se démarque avec des valeurs globalement supérieures. Par contre, les valeurs particulièrement élevées du R_2 des résidus T33, G48 et C49 se traduisent par un rapport R_1/R_2 remarquablement faible. Il est à noter que la valeur de R_1/R_2 n'est pas significativement différente entre les résidus de l'hélice 3 et ceux des autres hélices. Or, cette hélice a une orientation quasiment orthogonale aux autres hélices qui ont des orientations assez colinéaires. Le rapport R_1/R_2 du noyau azote est sensible à l'anisotropie de réorientation globale de la protéine (Barbato et al. [1992]). Nous pouvons donc conclure que le tenseur de diffusion rotationnelle d'ASP2 est plutôt isotrope, correspondant à une protéine relativement globulaire. La valeur moyenne du rapport R_1/R_2 sur les hélices a été utilisée par Tensor2 pour estimer le temps de corrélation globale d'ASP2 à 308K : $\tau_c = 4,87 \pm 0,04ns$. Le modèle isotrope est également suffisant d'après ce logiciel pour satisfaire les données expérimentales. Cette valeur a été utilisée pour l'ajustement des modèles de mouvements. La discussion sur le temps de corrélation observé aura lieu au paragraphe 4.6.1.

4.3 L'analyse des données de relaxation par le modèle "model-free" à 308K

Détermination des meilleurs modèles.

Le modèle développé par Lipari et Szabo et étendu par Clore a été appliqué au jeu de données expérimentales. Les résultats obtenus par Tensor2 sont rassemblés dans la figure 4.6. Le protocole décrit dans le paragraphe 5.1.7.2 a convergé vers 95 jeux de paramètres de mouvement (tous les résidus sauf le résidu E95). L'analyse effectuée par le logiciel FAST-Modelfree est extrêmement proche et permet de discuter sur la fiabilité des modèles pour certains résidus. Il est à noter que les critères de convergence de FAST-Modelfree sont plus sévères car 7 résidus sur 96 (C21, G48, L50, M55, C107, T114 et D115) n'ont pu être décrits correctement alors que Tensor2 se satisfaisait de l'un ou l'autre des modèles et le résidu 95 a pu étonnamment être décrit par FAST-Modelfree par le modèle 1 ($S^2 = 0,89$). D'autre part, 5 résidus (D43, M44, M60, C96 et M119) sont décrits par des modèles différents dans les deux logiciels. Nous discuterons des résultats de Tensor2 en premier lieu puis lorsque nécessaire, nous verrons quel est l'avis de Modelfree sur quelques résidus particuliers. Malgré la rhomboïté non négligeable du tenseur d'inertie de la structure d'ASP2 (0,6), le calcul utilisant un tenseur de diffusion rotationnelle anisotrope n'a pas amélioré l'ajustement. Nous analyserons donc les paramètres de dynamique associés au tenseur isotrope.

La moitié des résidus (47 résidus sur 96) est correctement et suffisamment décrit par le modèle 1 (S^2). La répartition homogène de ces résidus sur la séquence confirme qu'ASP2 se

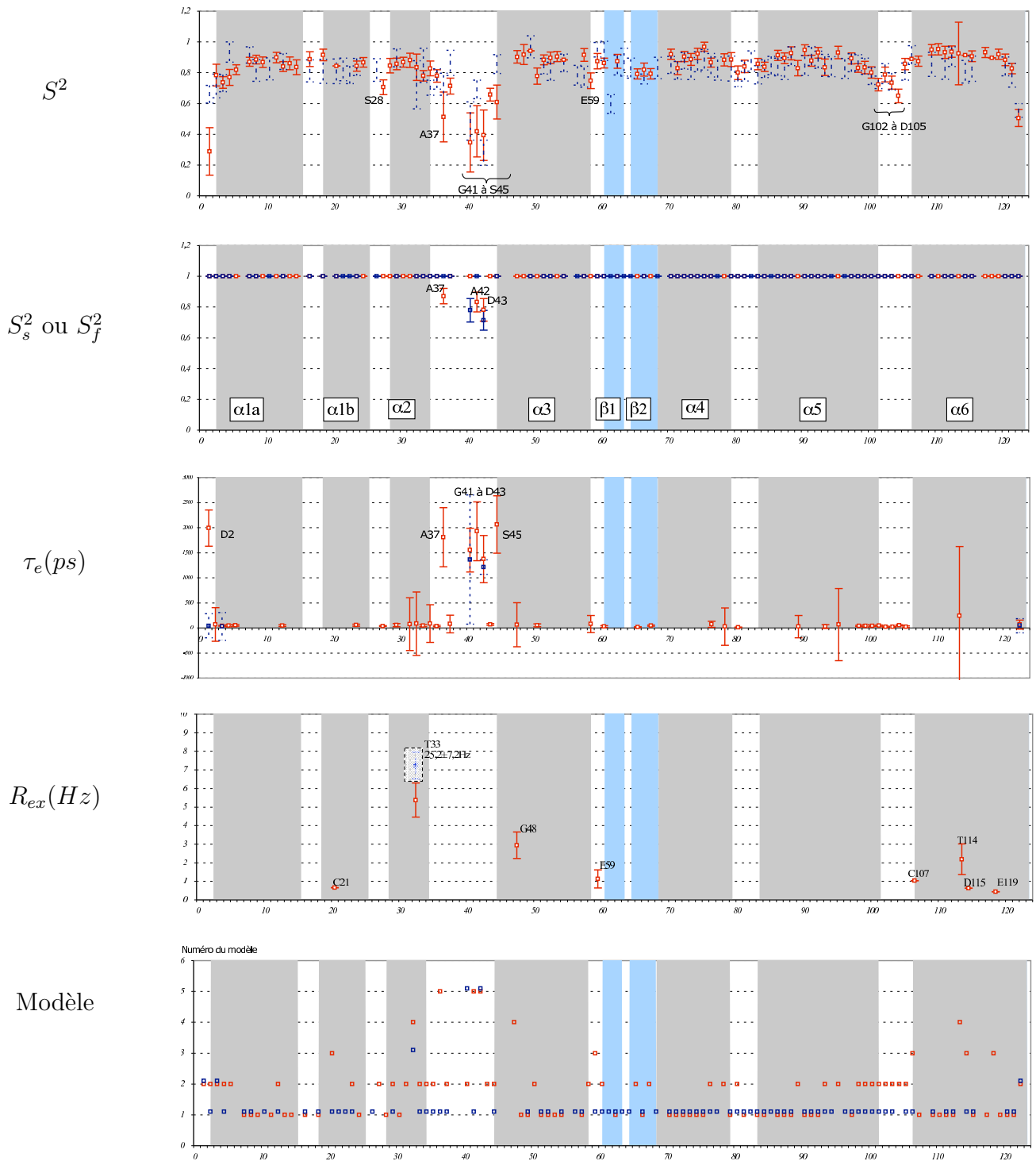


Fig. 4.6: Valeurs du numéro de modèle choisi par Tensor, de S^2 , τ_e et R_{ex} à 308K. Les résultats à 308K sont en rouge et à 283K en bleu. Les résidus présentant un comportement dynamique particulier à 308K sont annotés. Valeurs en Annexe F.

comporte de façon isotrope en solution car un temps de corrélation τ_c unique est suffisant pour décrire la réorientation de vecteurs de direction très différente. Les mouvements internes de ces résidus ont une faible amplitude ($S^2 \geq 0,7$) et ont lieu sur une gamme de temps inférieure à 100ps, si bien que la contribution du second terme de la fonction de densité spectrale donnée par l'équation 5.13 à R_1 et à R_2 est négligeable. La valeur du nOe hétéronucléaire de ces résidus étant supérieure à 0,7, la limite supérieure des mouvements internes peut être fixée à environ 20ps (Cloue et al. [1990a]).

Le modèle 2 (S^2 et τ_e) a été choisi pour la plupart des autres résidus (37 sur 96) : ils ne présentent ni dynamique interne complexe, ni échange chimique. Pour les autres résidus, la prise en compte de l'échange chimique (modèles 3 et 4, 8 résidus sur 96) ou d'un temps de corrélation supplémentaire traduisant la présence de dynamique interne de l'ordre de la nanoseconde (modèle 5, 3 résidus sur 96) a été nécessaire. Les résidus D2, G41 et S45 sont intéressants car Tensor2 et Modelfree ont tous deux considéré que le modèle 2 pouvait expliquer les données expérimentales avec des paramètres S^2 et τ_e quasiment identiques. Ce modèle fait l'hypothèse d'un temps de corrélation globale τ_c et d'un temps de corrélation des mouvements internes τ_e qui sont en principe sur deux gammes de temps distinctes. Le fit du modèle 2 a donné à τ_e une valeur de l'ordre de la nanoseconde pour ces trois résidus, soit de l'ordre de grandeur de la réorientation globale, alors qu'on attend plutôt une valeur de l'ordre de quelques centaines de picosecondes au plus. Ce résultat étonnant est explicable de trois façons : le jeu de mesures est erroné ou le modèle choisi cache des inconnues importantes. La première proposition ne semble pas raisonnable dans la mesure où les données issues de ces deux résidus semblent être cohérentes et dans la gamme de valeurs de leurs résidus voisins. Si on regarde de plus près le modèle 2 (équation 5.13), il est formellement identique au modèle 5 (équation 5.16) à la condition que $S_f^2 = 1$ (mouvements internes rapides très restreints dans l'espace) et $\tau_e' = \tau_s'$. Le temps de corrélation anormalement élevé obtenu par minimisation du modèle 2 pour les résidus 2, 41 et 45 est alors à interpréter dans les mêmes termes que le τ_s du modèle étendu. Cette explication est également confortée pour les résidus 41 et 45 car les résidus 42 et 43 ont un comportement dynamique similaire (modèle 5). Les mouvements très rapides des résidus 2, 41 et 45 ont donc une amplitude relativement faible et des mouvements internes lents importants.

Analyse des paramètres d'ordre S^2 .

Les paramètres d'ordre calculés montrent qu'ASP2 est globalement structurée en solution : la moyenne de S^2 sur la séquence vaut 0,82. La grande majorité des résidus en hélice ou en brin ont un S^2 compris entre 0,83 et 0,92 (tableau 4.1). Les mouvements internes de la dernière hélice sont encore plus restreints : $\langle S^2 \rangle = 0,92$. Comme observé en général sur les protéines, les premiers résidus en position N-terminale ont des mouvements moins restreints. Le S^2 croît jusqu'à l'alanine 8 puis reste constant au niveau de la valeur moyenne des S^2 sur les parties structurées en hélice ou en brin. De même, S^2 décroît en partie C-terminale mais beaucoup plus vite qu'il ne croît en N-ter : la chute de S^2 est brutale à partir du résidu F122. On note également que la boucle reliant les deux dernières hélices (entre 101 et 105) est plus mobile que les hélices qui l'entourent alors que la boucle située entre les hélices H4 et H5 semble être quasiment aussi rigide que les hélices. La zone comprise entre les résidus Q36 et S45 présente un profil original dans la mesure où ces résidus ont un paramètre d'ordre inférieur à

Résidu	Structure secondaire	Valeur de S^2
3-12	α_{1a}	0,83
15-25	α_{1b}	0,87
29-34	α_2	0,85
47-58	α_3	0,89
69-77	α_4	0,90
83-99	α_5	0,88
109-121	α_6	0,92
37-45		0,52

Tab. 4.1: Moyenne des S^2 sur les hélices α et la région A37-S45

0,7 signifiant la présence de mouvements internes de forte amplitude.

Modelfree et Tensor2 convergent tous deux vers les mêmes valeurs de S^2 pour la très grande majorité des résidus. 7 résidus ont des valeurs de S^2 significativement différentes mais avec une différence inférieure à 10% en fonction de l'application utilisée : Q3, M60, V68, K101, D105, M119 et S123. La différence de S^2 sur les résidus D43 et M44 est supérieure (environ 40%). Une explication simple est le choix d'un modèle différent pour les résidus D43, M44, M60 et M119. Le nombre de paramètres ajustables est alors différent ce qui affecte bien entendu les valeurs de S^2 . Pour les autres résidus, le modèle est identique (modèle 2). Le temps de corrélation τ_e calculés pour ces résidus par les deux logiciels étant également significativement distincts, il est donc probable que l'ajustement de S^2 et de τ_e est paramétré de façon différente dans les deux logiciels.

Analyse des mouvements internes.

Les temps de corrélation des mouvements internes sont plutôt inhomogènes le long de la séquence avec des valeurs allant de la dizaine de picosecondes à 2 nanosecondes.

Les mouvements internes de la plupart des vecteurs de liaison N-H ont lieu sur une gamme de temps inférieure à 80ps, voire même inférieure à 20ps (modèle 1). Dans ce cas, le paramètre d'ordre généralisé S^2 est associé directement à la restriction de ces mouvements.

Les vecteurs de liaison N-H présentant des mouvements sur la gamme de la nanoseconde sont concentrés essentiellement entre les résidus 41 à 45. L'acide aminé D2 en position N-terminale et l'alanine 37 ont également une dynamique de cet ordre. Les résidus 2, 41 et 45 sont décrits par le modèle (S^2, τ_e) avec τ_e de l'ordre de la nanoseconde, cela signifie que ces résidus ont des mouvements internes rapides très restreints ($S_f^2 \simeq 1$). Pour les autres, la valeur de S_f^2 étant proche de celle du S^2 des résidus ajustés par le modèle simple, en suivant l'intuition de Clore et de ses collègues (Clore et al. [1990a]), on peut conclure que les mouvements internes rapides ont une origine physique et une amplitude similaires à ceux ayant lieu dans la plupart des résidus "normaux". Les mouvements intermédiaires ont pour conséquence de réduire le paramètre d'ordre S^2 d'un facteur S_s^2 .

Analyse des résidus en échange chimique.

Tensor2 a jugé nécessaire l'introduction d'un terme d'échange chimique dans la fonction de densité spectrale de 8 résidus : C21, T33, G48, M60, C107, T114, D115 et M119. Parmi ceux-ci, seul le résidu T33 présente de l'échange chimique d'après Modelfree. Il est vrai que l'élargissement de raie de ce résidu est particulièrement flagrante. Modelfree n'a pas introduit

de terme d'échange pour le résidu 119, il a conservé le modèle 1 mais la valeur du χ^2 obtenue est plusieurs fois supérieure aux χ^2 des autres ajustements. Pour les autres résidus, Modelfree a préféré ne pas fournir de modèles d'interprétation des paramètres de relaxation en terme de mouvements. Nous pouvons donc raisonnablement en déduire que seul le vecteur de liaison N-H de la thréonine 33 puisse adopter deux conformations avec des déplacements chimiques différents. Toute interprétation sur l'état conformationnel des autres résidus sera à considérer avec précaution.

Tensor2 vs FAST-modelfree

L'analyse des vitesses de relaxation et de l'effet nOe hétéronucléaire d'ASP2 par deux logiciels a montré qu'en première approche, les résultats issus sont compatibles. A l'exception de quelques résidus, le choix des modèles et le résultats des ajustements sont identiques aux incertitudes près. Les résidus pour lesquels des résultats significativement différents ont été obtenus sont apparemment des résidus dont la dynamique interne du vecteur NH est complexe (utilisation du modèle Lipari-Szabo étendu ou présence d'échange chimique). J'attire tout de même l'attention du lecteur sur le fait que le jeu de paramètres de relaxation d'un seul résidu n'a pu être interprété par Tensor2 tandis que FAST-modelfree en a exclu 7. Les critères de ce dernier sont donc plus sévères. Le futur utilisateur de Tensor2 prendra donc des précautions lors de l'analyse des résultats.

4.4 Influence de la température sur la dynamique.

La série d'expériences discutée précédemment a été menée à 283K dans les mêmes conditions. Les figures 4.1, 4.2 et 4.3 rassemblent les valeurs de R_1 , R_2 et de l'effet nOe pour les 89 résidus non superposés. La moyenne des R_1 est d'environ de 1,25Hz avec une incertitude moyenne de 0,2Hz (soit 16%). La moyenne des R_2 est de 12Hz pour une incertitude moyenne de 1Hz (soit 13%). Comme attendu, la relaxation longitudinale est plus lente et la relaxation transversale plus rapide à basse température. L'incertitude plus importante dans les expériences à 283K traduit la variation du rapport signal sur bruit entre les deux températures.

Le profil de variation des constantes de relaxation et de l'effet nOe hétéronucléaire est conservé entre les deux températures, indiquant que les mouvements présents à haute température ne sont pas réduits à basse température. C'est en particulier vrai pour les résidus de la région flexible de la protéine ainsi que pour la boucle entre les hélices H5 et H6.

En conséquence, le profil de variation des S^2 le long de la séquence est très proche entre les deux températures (figure 4.6). On note tout de même qu'à basse température, les mouvements internes des résidus 41 et 43 ont une amplitude plus faible mais toujours significativement supérieure à l'amplitude des mouvements internes des résidus impliqués dans une structure secondaire stable. Le résidu T33 présente toujours une valeur de R_2 anormalement élevée (25,2Hz contre 5,7Hz à 308K) : l'échange chimique qui affecte ce résidu n'est donc pas éliminé en baissant la température. La durée de vie réduite τ_{ex} d'un échange chimique gouverné par la loi d'Arrhénius diminue avec la température. En négligeant un changement de population relatives des sites, la baisse de la température conduit à une augmentation du terme d'échange dans la relaxation R_2 (équation 5.11). La diminution du terme R_{ex} entre 283K et 308K calculé par Tensor2 est donc cohérente avec la différence de température.

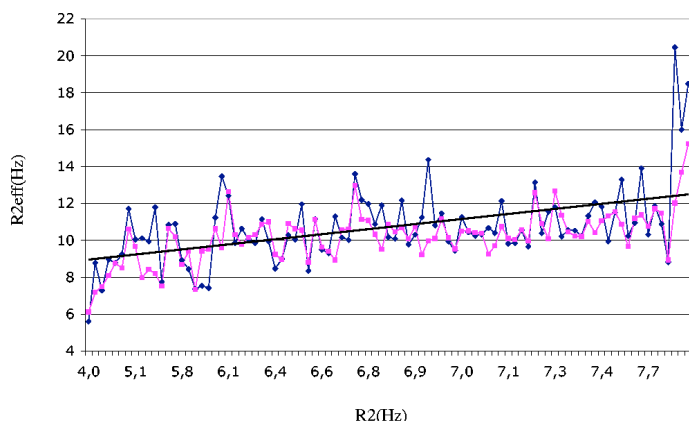


Fig. 4.7: Corrélation entre $R_{2,eff}$ et R_2 pour $\nu_{cpmg} = 80Hz$ (en bleu) et $\nu_{cpmg} = 800Hz$ (en violet). La droite de régression à 80Hz est représentée en noir

A 283K, le temps de corrélation d'ASP2 est de $\tau_c = 9,3 \pm 0,1ns$. L'augmentation de la température a pour effet de diminuer la viscosité de l'eau et d'accélérer la réorientation globale d'une protéine en solution. Il est donc satisfaisant de constater une baisse du temps de corrélation lorsque la température passe de 283K à 308K. Le rapport des τ_c expérimentaux est donc de 0,51. D'après la loi de Stoke-Einstein, la variation de τ_c avec la température est liée à la variation de la viscosité de l'eau :

$$\frac{\tau_c(308K)}{\tau_c(283K)} = \frac{\eta_{308K}^*}{\eta_{283K}^*} \frac{283}{308} = 0,71 \quad (4.1)$$

La variation du temps de corrélation globale d'ASP2 est donc relativement cohérente avec la baisse de température. La structure d'ASP2 n'est pas parfaitement sphérique, ceci explique l'écart entre 0,51 et 0,71.

4.5 Mise en évidence de l'échange chimique à 308K

J'ai appliqué la séquence d'impulsions publiée par Loria (voir paragraphe 5.2.1.4 de la partie Matériel et Méthodes) sur l'échantillon qui a servi à l'étude de la relaxation à 283K et 308K. En appliquant la procédure décrite au paragraphe 5.2.1.4, j'ai déterminé les rapports d'intensité entre les expériences avec CPMG (à 80Hz et 756Hz) et sans CPMG (expériences doublées), à 308K.

La valeur moyenne des $R_{2,eff}$ sur l'ensemble de la séquence est sensiblement la même à 80Hz ($10,89 \pm 2,6Hz$) et à 756Hz ($10,3 \pm 1.6Hz$). Il était attendu que la valeur de $R_{2,eff}$ soit supérieure à R_2 (6,7Hz) où $R_{2,eff}$ est une moyenne entre la relaxation transversale de l'état antiphase $I_z S_y$ et de l'état en phase S_x et que les états antiphase relaxent en général plus vite que les états en phase du fait de la corrélation croisée dipolaire-dipolaire proton-proton : la vitesse de relaxation moyenne de l'état antiphase vaut ici environ 14,3Hz.

4.5.1 Comparaison de $R_{2,eff}$ et R_2

Dans l'expérience de mesure de R_2 (paragraphe 4.1), le délai τ_{cp} était fixé à 1ms, soit une fréquence $\nu_{CPMG} = 500Hz$, fréquence intermédiaire entre 80 et 756Hz. La contribution de l'échange chimique de fréquence inférieure à 500Hz est donc éliminée de la mesure de R_2 . En

négligeant la variation du terme d'échange avec la fréquence CPMG et en faisant l'hypothèse que la vitesse de relaxation de l'état antiphase est proportionnelle à celle de l'état en phase (i.e. R_2 hors échange), on s'attendrait à ce que $R_{2,eff}$ ait la même évolution le long de la séquence que R_2 . La figure 4.7 représente les données de $R_{2,eff}$ aux deux fréquences CPMG en fonction de R_2 , excepté de résidu 2 dont le comportement semble être particulier (voir paragraphe suivant). Bien que les variations par rapport à la courbe de tendance soient importantes, on note tout de même une certaine corrélation entre les deux types de vitesse de relaxation. La pente des droites de régression ($R_{2,eff} = aR_2 + b$) est plus basse à 80Hz (1,21) qu'à 800Hz (1,43). Les données à 800Hz semblent être plus fortement corrélées aux R_2 ($r_{800}^2 = 0,37 > r_{80}^2 = 0,14$) comme on s'y attendait du fait que la mesure de R_2 élimine l'échange sur une gamme de fréquence ($k_{ex} < 500Hz$) proche de l'expérience CPMG à 756Hz ($k_{ex} < 756Hz$). Si $R_{2,anti}$ est proportionnelle à R_2 ($R_{2,anti} = \alpha R_2$), la pente des droites de régression serait théoriquement égale à $(1 + \alpha)/2$ (formule 5.24 de la partie Matériel et Méthodes, avec $R_2 = R_{2,in}$) et la constante b serait nulle. Par conséquent, d'après nos données, le coefficient α est de 1,42 et 1,86, soit supérieure à 1.

$$\begin{aligned} R_{2,in} = R_2(S_+) &= \frac{2}{3}(3d + c).J(0) + \frac{1}{2}(3d + c).J(\omega_N) + \frac{d}{2}.J(\omega_H + \omega_N) \\ &\quad + 3d.J(\omega_H - \omega_N) + 3d.J(\omega_H) \\ R_{2,anti} = R_{2z}(2I_z S_+) &= \frac{2}{3}(3d + c).J(0) + \frac{1}{2}(3d + c).J(\omega_N) + \frac{d}{2}.J(\omega_H + \omega_N) \\ &\quad + 3d.J(\omega_H - \omega_N) + \rho_{HH} \end{aligned} \quad (4.2)$$

Les relations 4.2 reliant la vitesse de relaxation des cohérences $I_z S_+$ et S_+ avec la valeur de la densité spectrale aux fréquences de transition du systèmes (hors échange chimique) indiquent que la différence de relaxation entre les états en phase et en antiphase dérive de la présence d'un terme de relaxation croisée DD entre le proton amide et les protons spatialement proches et du terme $3dJ(\omega_H)$ dans $R_2(S_+)$, soit $R_{2,anti} = R_{2,in} - 3dJ(\omega_H) + \rho_{HH}$. Nous pouvons déjà observer que $R_{2,anti}$ est significativement supérieur à $R_{2,in}$, par conséquent le terme de relaxation croisée proton-proton est plus grand que $3dJ(\omega_H)$. Ceci confirme les difficultés d'interprétation de la relaxation des états mixtes en termes de mouvement. Intuitivement, le terme de relaxation croisée devrait être peu corrélé à la relaxation transversale du spin azote 15 car la variation de la distance entre les protons devrait suffisamment influencer sur la relaxation dipôle-dipôle proton-proton (en $1/r^6$) pour compenser la corrélation entre l'orientation des vecteurs H_N-H et l'orientation du vecteur H_N-N . Le fait qu'une courbe de tendance s'établisse semblent montrer qu'il existe une certaine corrélation entre ρ_{HH} et $R_{2,in}$.

Bien entendu, il faut souligner que la mesure de $R_{2,eff}$ est sans doute entachée d'une incertitude plus importante car elle est obtenue pour deux points uniquement, contrairement au R_2 qui lui est calculé sur un ajustement de huit points. Une mesure plus précise de $R_{2,eff}$ à la même fréquence CPMG que la mesure de R_2 est indispensable pour pouvoir confirmer ou infirmer la corrélation entre ρ_{HH} et $R_{2,in}$.

4.5.2 La région D30 à A38 présente de l'échange chimique

En utilisant la relation 5.26 (partie Matériel et Méthodes) et l'incertitude associée pour comparer les spectres à 80Hz et à 756Hz, j'ai construit le profil de la figure 4.9. Formellement, le résultat est le même que d'effectuer la différence entre les valeurs de $R_{2,eff}$ aux deux fréquences.

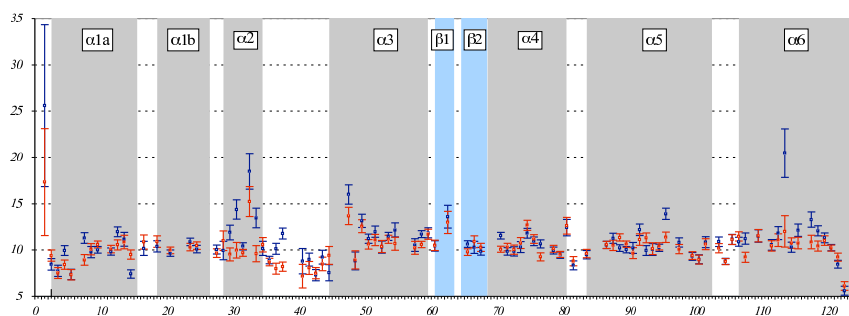


Fig. 4.8: Valeurs de $R_{2,eff}$ à 80Hz (en bleu) et à 756Hz (en rouge)

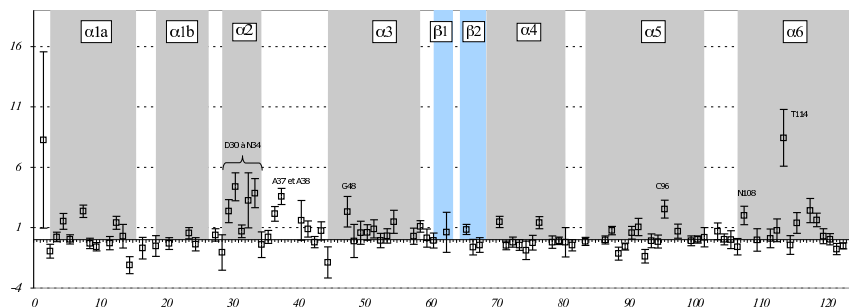


Fig. 4.9: Valeurs de $R_{2,eff}(80Hz) - R_{2,eff}(756Hz)$

Cependant, le calcul de l'erreur donne des intervalles plus grands car non seulement on tient compte de l'incertitude sur les spectres avec et sans CPMG mais on doit également sommer les incertitudes sur les $R_{2,eff}$ lors de la soustraction des $R_{2,eff}$. La barre d'erreur montrée sur la figure 4.9 n'est liée qu'à l'écart-type du bruit sur les spectres aux deux fréquences (et choisi identique pour les deux).

Il apparaît clairement alors que la région comprise entre les acides aminés D30 à A38 correspondant à l'hélice α_2 sont impliqués dans de l'échange chimique de l'ordre de la milliseconde.

Le résidu D2 présente à nouveau un comportement surprenant. Les valeurs de $R_{2,eff}$ mesurées (17,3 et 25,6Hz) sont considérablement supérieures au R_2 (4Hz). La grande barre d'erreur provient d'un signal initialement faible. On peut considérer qu'on a sous-estimé la barre d'erreur pour ce résidu. La différence de $R_{2,eff}$ entre les deux fréquences ne peut alors être significatif d'un échange chimique quelconque. Le résidu T114 quant à lui se comporte typiquement comme un résidu avec un vecteur NH en échange : la valeur de $R_{2,eff}(80Hz)$ est significativement supérieure à $R_{2,eff}(800Hz)$ avec la valeur $R_{2,eff}(800Hz)$ de l'ordre de grandeur du $R_{2,eff}(800Hz)$ des résidus voisins dans la séquence. Dans ce cas, la séquence d'impulsions CPMG a détruit la composante de relaxation d'échange. Le vecteur de liaison N-H de l'acide aminé T114 est donc susceptible d'être dans deux ou plusieurs sites chimiquement non équivalents et la durée de vie réduite τ_{ex} de cet équilibre est de l'ordre de la milliseconde.

4.5.3 CPMG et Modelfree

Les résultats du CPMG sont à mettre en lumière de l'ajustement des paramètres de relaxation. Si la première méthode donne une preuve directe de la présence d'échange dans la gamme de la milliseconde, la seconde méthode n'introduit l'échange que par l'intermédiaire d'un paramètre dans l'ajustement des données. Le R_{ex} déduit de la mesure de R_2 contient les contributions de tous les échanges chimiques de durée de vie inférieure à la milliseconde. Il doit donc être supérieur à l'échange que l'on a mesuré par CPMG. Ce n'est visiblement pas

le cas pour les résidus D30 à A38 pour lesquels le modèle de Lipari-Szabo n'a pas permis de détecter de l'échange chimique, à l'exception notable de T33. Une seule mesure de R_1 , R_2 et η (i.e. à un seul champ magnétique \vec{B}_0) ne permet d'accéder qu'à trois paramètres de mouvements simultanément. On ne peut donc pas reconnaître par les cinq modèles disponibles les résidus présentant à la fois une dynamique interne complexe avec deux gammes de temps de corrélation (model-free étendu) et plusieurs environnements magnétiques différents. Les mouvements internes intermédiaires ont tendance à baisser la constante de relaxation transversale et l'échange chimique l'augmente. Si leurs effets se compensent, ce type de résidu aura apparemment un comportement dynamique similaire à des résidus rigides sans échange. L'effet nOe hétéronucléaire des résidus D32 à A38 est inférieur à 0,7 ce qui témoigne de l'effet non négligeable de la dynamique interne sur la relaxation. Il est donc très probable que ces résidus présentent à la fois une dynamique interne riche et de l'échange chimique. Cependant, l'absence de R_{ex} , au sens model-free, pour ces résidus peut aussi être expliquées si la vitesse de l'échange est dans la gamme très serrée autour de 1000Hz : dans l'expérience CPMG, les déplacements chimiques des différents sites ne sont pas entièrement refocalisés tandis qu'ils le sont dans la mesure du R_2 .

Pour essayer d'apporter une réponse à cette question, j'ai soustrait à la mesure de R_2 , la composante d'échange détectée par la séquence CPMG. Ceci m'a donné un nouveau jeu de données de relaxation que j'ai soumis à Tensor2. La diminution du R_2 des azotes de la région 30 à 38 n'a pas conduit à une variation significative des paramètres de mouvement (S^2 , τ_e).

Le cas de la thréonine 114 est plus simple. L'analyse de la dynamique a nécessité l'utilisation du terme d'échange pour répondre à un R_2 anormalement élevé. L'expérience CPMG a confirmé l'existence de cet échange et a précisé qu'il apparaissait sur la gamme de la milliseconde. La composante d'échange est au minimum de 8Hz d'après l'expérience CPMG que nous avons utilisée alors que Model-free l'estime à 2,5Hz. Il est donc probable qu'une partie de l'échange que nous observons dans la CPMG a lieu sur une gamme de fréquence inférieure à 500Hz, ce qui nous la rend invisible lors de la mesure de R_2 . La figure 4.10 propose un profil de R_{ex} en fonction de la fréquence CPMG appliquée. A fréquence nulle, tous les phénomènes d'échange contribuent à la relaxation. Les échanges sur la gamme inférieure à 500Hz sont majoritaires et sont détruits lors de la mesure du R_2 . Le terme d'échange que l'on mesure par R_2 (2,5Hz) contient le résiduel des échanges à haute fréquence. Dans la CPMG, on n'observe qu'une variation du terme d'échange, soit 8Hz entre 76 et 756Hz.

4.6 Structure et dynamique

L'étude de la dynamique d'ASP2 à pH 6 enrichit notre compréhension de la structure obtenue au chapitre précédent. La dynamique d'ASP2 présente des caractéristiques déjà observées sur la plupart des protéines structurées. Nous verrons que la région E29 à S45 est à traiter séparément car sa dynamique particulière explique les difficultés rencontrées lors de la reconstruction.

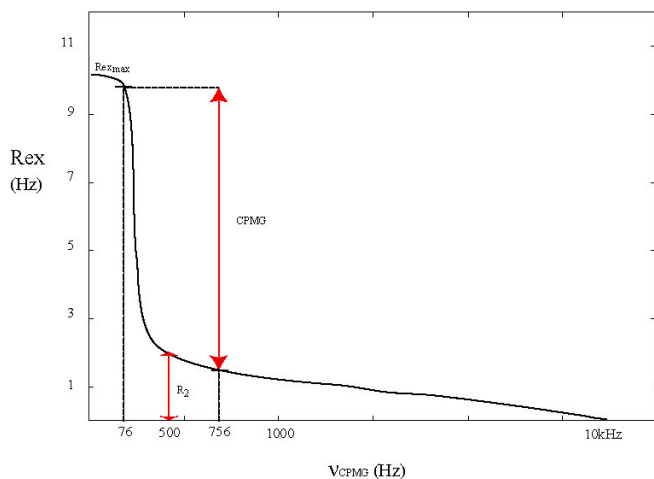


Fig. 4.10: Variation probable du terme d'échange chimique du résidu T114 en fonction de la fréquence CPMG et mesures expérimentales par CPMG ou R_2

4.6.1 Analyse du temps de corrélation τ_c

Etat oligomérique La valeur de τ_c calculée à partir de l'équation de Stoke-Einstein pour un monomère de poids moléculaire 13,7kDa est de 6,7ns à 308K et de 11,7ns pour un dimère de 27,4kDa. Dans ce modèle simplifié, le monomère a un temps de corrélation de 8,6ns et le dimère, un temps de corrélation de 15,1ns à 283K. La valeur de τ_c mesurée sur ASP2 à 308K ($\tau_c = 4,87ns$) indique clairement son état d'oligomérisation : ASP2 en complexe avec le TSP est un monomère en solution. L'équation de Stoke-Einstein est valable pour une sphère rigide en rotation isotrope dans de l'eau pure. Dans le cas d'une protéine, des différences peuvent apparaître : anisotropie du tenseur de réorientation de la structure, équilibre rapide monomère-dimère, variation de la viscosité de la solution par une forte concentration en protéine. Dans le cas d'ASP2, l'anisotropie du tenseur d'inertie et donc du tenseur de diffusion rotationnelle est certainement à l'origine de l'écart observé, sans toutefois mener à des différences notables en terme de relaxation. A 283K, le temps de corrélation d'ASP2 est de 9,3ns à comparer avec la valeur de 11,7ns, calculée à partir de la loi de Stoke-Einstein. L'état oligomérique est donc conservé à basse température.

Comparaison avec les temps de corrélation d'autres protéines Le temps de corrélation nous indique qu'ASP2 est monomérique. Nous pouvons donc mettre en lumière la valeur de τ_c (4,87ns) avec celles mesurées sur des protéines de taille comparable dans les mêmes conditions de température. Ces exemples (tableau 4.2) sont extraits de la mise au point bibliographique de la thèse de Carine Van Heijenoort (Van Heijenoort [1994]) à l'exception de la dernière. Il apparaît qu'ASP2 se réoriente particulièrement rapidement sur elle-même par rapport aux autres protéines citées.

4.6.2 La dynamique des structures secondaires

Les hélices α ainsi que le β -hairpin sont rigides, comme le montre la valeur élevée du paramètre d'ordre S^2 aux deux températures. La modélisation de la structure a convergé vers des structures secondaires effectivement bien définies. Si le résidu D2 est très mobile,

Protéine	taille	pH	T	$\tau_c(ns.rad^{-1})$
Staphyloccocal Nucléase	148aa	6,4	308K	9,1±0,4
Interleukine-4	129aa	5,6	308K	9,1±0,4
Domaine ribonucléase H de la transcriptase inverse de HIV-1	137aa	5,4	308K	8,35
Interleukine-1 β	153aa	5,4	309K	8,30±0,05
Thiorédoxine réduite et oxydée	108aa	?	308K	6,5
Domaine C-terminal SH2 de la phospholipase C γ 1 complexé à un phospholipide	13,7kDa	6,4	308K	6,5±0,3
Protéine "jaune" photoactive (Dux et al. [1998])	121aa	?	311K	6,4
ASP2	123aa	6	308K 283K	4,87 9,3

Tab. 4.2: Valeur de τ_c mesurée sur quelques protéines de taille comparable à ASP2 et à 308K, dans le cadre du formalisme de Lipari-Szabo.

la première hélice n'atteint sa rigidité maximale qu'à partir du résidu A8. Le squelette de la première partie de l'hélice est en effet moins bien défini que la seconde partie. L'hélice C-terminale quant à elle, est très bien structurée jusqu'à l'acide aminé F122, jugé sur la structure et par la dynamique, malgré la baisse sensible du S^2 dès la leucine 121. On peut également noter que le vecteur amide du résidu E59 en fin d'hélice α_3 a une amplitude de mouvement plus large qu'au sein de l'hélice, ce qui confirme la fin de cette hélice.

4.6.3 La dynamique des boucles

La structure d'ASP2 contient trois boucles (H26-S28, H79-D82 et K102-D105) aux comportements différents. La sérine 28 de la première boucle ainsi que les résidus G102 à D105 ont une amplitude de mouvement interne importante sur la gamme pico-nanoseconde. La boucle reliant les hélices α_4 à α_5 est à peine plus mobile que les hélices. Ceci s'explique par sa conformation en coude. La présence d'une liaison hydrogène entre les résidus H79 et N82 ainsi que sans doute, entre les chaînes latérales de cette asparagine et de l'aspartate D85, explique la rigidité de cette région. En revanche, les vecteurs NH de la boucle reliant les hélices α_5 et α_6 (A100 à D105) ont une amplitude relativement élevée qui se traduit par un S^2 inférieur à 0,8 pour les résidus G102 à D105 et par une contribution importante des mouvements rapides à la relaxation (τ_e de l'ordre de 40ps). La superposition des 20 structures montre en effet que cette région est légèrement moins bien définie que les hélices. L'écart quadratique moyen par résidu est également très légèrement supérieur entre les résidus A100 à D105 (moyenne de 0,7Å) par rapport à celui observé sur l'hélice α_5 (moyenne de 0,35Å) et l'hélice α_6 (moyenne de 0,55Å).

4.6.4 Ponts disulfure

Le cas des trois ponts disulfure est intéressant à discuter : ils sont susceptibles d'avoir une dynamique complexe. Le paramètre d'ordre des six cystéines est élevé, leurs vecteurs amides ont donc des mouvements relativement restreints. Cependant, plusieurs éléments suggèrent qu'elles sont en échange chimique : les vitesses de relaxation transversale R_2 des cystéines

49 et 107 sont sensiblement supérieures à la moyenne, un terme R_{ex} est introduit pour les cystéines 21 et 107 (par Tensor2 uniquement), l'expérience CPMG indique que la cystéine 96 ainsi que les résidus voisins des cystéines 107 et 49 ont un terme d'échange (faible) dans la gamme de la milliseconde. Le pont disulfure 49-107 semble être le plus affecté.

S'il faut prendre des précautions dans l'interprétation car le faisceau d'indices n'est pas clairement convergent, il est probable que l'isomérisation des ponts disulfure soit à l'origine de ces observations. Ce phénomène avait déjà été remarqué sur BPTI par plusieurs auteurs (Otting et al. [1993], Millet et al. [2000], Palmer III et al. [2001]).

4.6.5 La région D30 à S45

La reconstruction d'ASP2 à pH 6 a mis en évidence une grande flexibilité de la structure entre les résidus E29 et S45, traduisant à la fois une structuration secondaire fragile (écart de déplacement chimique faible par rapport au *random coil*) et le peu de nOe à moyenne distance et l'absence de nOe à longue distance. L'analyse fine des mouvements des vecteurs amides de ces acides aminés apporte des explications et montre que leur dynamique n'est pas homogène sur cette région. On peut distinguer les zones D30 à A38 (région I) et A37 à S45 (région II). La région II est clairement dynamique de part un effet nOe hétéronucléaire et un S^2 faibles. Les vecteurs amides sont affectés de mouvements importants sur une gamme large : de la picoseconde à la nanoseconde. La région I est aussi rigide que les structures en hélice α ($S^2 > 0,75$), ce qui peut confirmer la présence d'une hélice dans cette région, détectée par les déplacements chimiques. Cependant, la présence d'échange chimique sur la gamme de la milliseconde observé par l'expérience CPMG appuie l'idée de structure secondaire fragile. Bien que l'échange chimique peut être *a priori* lié à un changement de conformation du squelette ou à un équilibre entre plusieurs orientations des chaînes latérales affectant le déplacement chimique de l'amide, la faiblesse des nOe caractéristiques des hélices du tour d'hélice entre D30 et N34 favorise l'hypothèse de la déstructuration partielle de l'hélice α_2 . S'il est clair que les résidus D30 à A38 sont en échange chimique, cela n'exclut pas la possibilité d'une dynamique interne importante car l'échange chimique peut compenser en relaxation R_2 la contribution des mouvements rapides. La faiblesse de l'effet nOe hétéronucléaire dans la région I appuie cette idée.

Par ailleurs, nous avons noté que l'environnement chimique de l'azote du résidu T114 varie rapidement. Cet acide aminé appartient à l'hélice terminale et est orientée vers la partie ouverte de la structure, susceptible donc d'être en contact vers la région II. Aucun nOe n'a été attribué entre ces résidus. La présence d'échange chimique sur ce résidu est un argument fort en faveur d'un équilibre en solution entre une structure fermée et une structure plus ouverte. Dans une moindre mesure, le résidu A8, lui aussi orienté vers cette région, est susceptible de présenter de l'échange conformationnel, mais la barre d'erreur sur ce résidu ne permet pas de conclure. On peut noter également la faiblesse du signal issu de la chaîne latérale de l'isoleucine 18 dans l'ensemble des expériences. Orientée vers la région I, on peut à nouveau supposer que l'équilibre structure fermée-structure ouverte affecte le déplacement chimique de la chaîne latérale, menant à la coalescence des raies, à la manière d'un échange intermédiaire sur la gamme des déplacements chimiques.

La flexibilité de cette région n'est pas liée à la haute température à laquelle la reconstruction a été effectuée. En effet, à basse température (283K), physiologiquement plus significative, les

résidus D30 à S45 ne montrent pas plus de corrélations nOe qu'à haute température. Les mouvements internes sont toujours d'une amplitude importante à 283K, quoique légèrement plus faible qu'à 308K.

4.7 Conclusion

En conclusion, l'étude de la dynamique d'ASP2 a confirmé le comportement original du segment 30-45 de la protéine. Les difficultés rencontrées lors de la reconstruction s'expliquent par la conjonction de deux phénomènes : de l'échange chimique pour la première partie du segment (30-38) et des mouvements internes de grande amplitude dans la gamme pico-nanoseconde pour la seconde partie du segment. Si le jeu de structures obtenues montre que la dispersion de ce segment est très importante, il est raisonnable de penser qu'en réalité, l'espace conformationnel balayé par ce segment est beaucoup plus restreint. En effet, les titrations en pH et en température du complexe ASP2-TSP ont indiqué que l'échange entre les protons amides et l'eau n'est pas très rapide, à l'exception des résidus 41 à 45. Les protons amides de 30 à 40 sont donc protégés du solvant. Ceci suggère que cette partie du segment conserve une structuration secondaire. De plus, la présence de quelques nOes reliant ces résidus avec l'hélice α_{1b} ainsi que le comportement de résidus orientés vers ce segment, qui présentent soit de l'échange chimique (azote de A8, T114), soit des signaux très faibles (I18) confortent l'idée que la structure moyenne "réelle" est plus fermée que celle calculée. Enfin, comment expliquer une telle ouverture de la structure alors que de très nombreux acides aminés sont hydrophobes (I31, I35, A37, A38) ?

Ce n'est pas la première fois que le chercheur en RMN est confronté à ce genre de situation. Ainsi, D'amelio et al. [2002] ont étudié la dynamique d'échange de la forme apo de la cytochrome b_{652} . Cette dernière diffère la protéine holo, contenant l'hème, par une mauvaise définition de l'hélice C-terminale. Les données RMN étaient contradictoires sur le segment C-terminal de la forme apo : échange H/D lent, présence de liaisons hydrogènes, paramètre d'ordre S^2 élevé, CSI favorable à une hélice, qui indiquent toutes une bonne structuration secondaire du segment incriminé mais la difficulté à déterminer des contraintes nOe ont mené à cette mauvaise définition. La question fut close par des expériences de type CPMG et par une simulation de dynamique stochastique qui ont détecté sur le segment C-terminal d'une part, de l'échange sur la gamme de la milliseconde et d'autre part, des mouvements importants dans la gamme de la nanoseconde. Le désordre conformationnel sur différentes gammes de temps expliquent donc les observations. Les similitudes entre ce travail et le nôtre sont très fortes.

Chapitre 5

Étude des deux formes libres d'ASP2

Comme nous l'avons observé dans le chapitre 2 au paragraphe 2.5.1, ASP2 adopte deux états en équilibre lent vis-à-vis de l'échelle des déplacements chimiques au-dessus de pH 6,6. J'ai entrepris quelques expériences afin de comprendre les différences et l'origine de ces deux formes.

5.1 Les couplages dipolaires et ASP2 libre

5.1.1 Mesure des couplages dipolaires résiduels.

Le choix du milieu orientant a été basé sur les caractéristiques d'ASP2. Le point isoélectrique théorique d'ASP2 étant inférieur à 5, la charge globale d'ASP2 à pH 6 est positive. Les interactions coulombiennes risquent donc d'être trop fortement attractives avec les bactériophages pour qu'ASP2 subsiste libre en solution. Jérôme Coutant (au laboratoire) ayant déjà l'expérience nécessaire sur les milieux bicellaires DHPC/DMPC, j'ai donc essayé ces milieux. Le protocole de préparation des bicelles est celui publié par Raffard (Raffard et al. [2000]) et mis en place au laboratoire par Jérôme Coutant (Coutant [2003]). Il est présenté en annexe D.

Dans un premier temps, j'ai essayé le milieu DHPC/DMPC classique ($q=3$, concentration totale 3%(w/v)). L'écartement différent des raies aux deux températures (313K et 293K) montrait que ASP2 s'orientait. Cependant, la valeur maximale de l'écart était faible. J'ai donc essayé le milieu bicellaire contenant les liaisons éther ($q=3$, concentration totale 3%(w/v)), plus stable vis-à-vis de l'hydrolyse à pH 6,4. L'orientation était identique. Bien que faibles (inférieurs à 10Hz pour la plupart), les couplages dipolaires résiduels ont pu être mesurés dans les deux milieux orientant. L'ordre induit dans les deux tubes est le même comme le traduisent les valeurs de RDCs. Nous ne traiterons donc par la suite que les données issues du tube contenant les bicelles-éther.

Les spectres à 293K et 313K (figure 5.1) indiquent qu'entre ces deux températures, les phospholipides subissent une transition de phase que l'on peut attribuer au passage d'un état où les phospholipides sont relativement mobiles en solution à un état où ces phospholipides sont en phase bicellaire. L'orientation des bicelles a été contrôlée par l'apparition d'un doublet sur le signal du deutérium, doublet issu du couplage quadrupolaire (environ 20Hz). La figure 5.2 montre les spectres ^{15}N -IPAP-HSQC enregistré à 313K et 293K sur l'échantillon d'ASP2

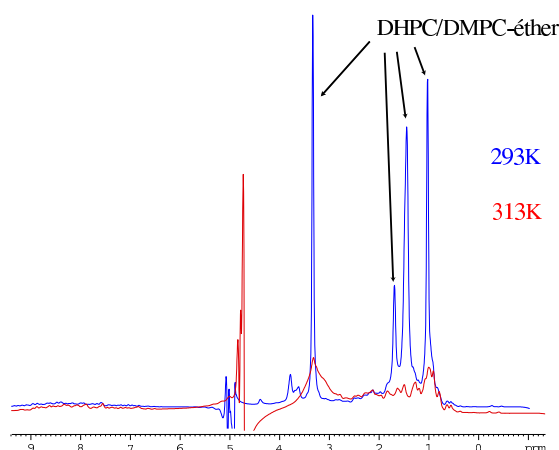


Fig. 5.1: Spectres 1d-Watergate à 293K (bleu) et à 313K (rouge). A 293K, les signaux des phospholipides sont fins et intenses, signe d'une dissolution importante. A 313K, les phospholipides sont engagés dans des objets de grande taille (bicelle), leurs signaux deviennent larges et peu intenses (vers 3,5ppm).

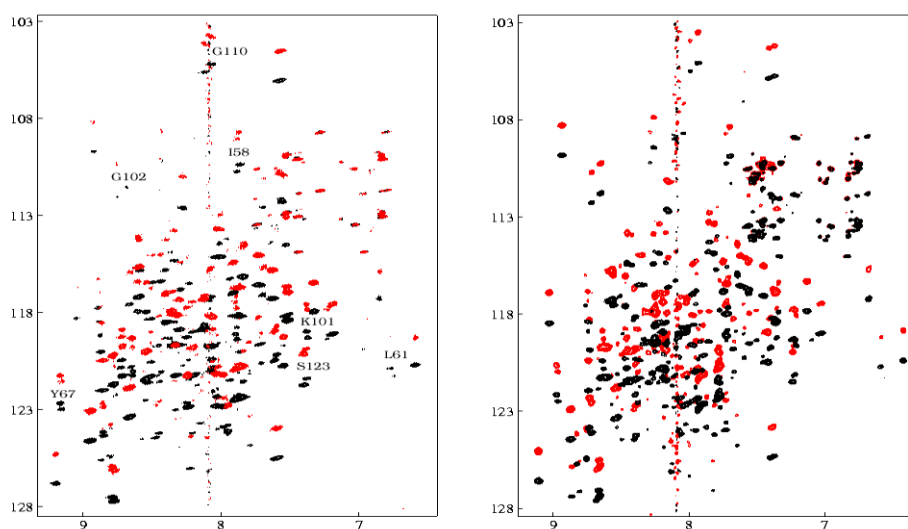


Fig. 5.2: Spectres ^{15}N -IPAP-HSQC d'ASP2 dissoute dans les bicelles éther à 3%(w/v). Les deux composantes de chaque doublet sont en rouge et en noir. A gauche, les spectres enregistrés à 313K (orienté). A droite, les spectres enregistrés à 293K. Quelques résonances dédoublées sont attribuées.

orienté dans la phase bicellaire éther. Tout d'abord, on constate que les raies sont plus larges à haute température qu'à basse température. Ceci est donc une indication d'une interaction avec les bicelles. Cette interaction se traduit par une orientation préférentielle pour ASP2 vis-à-vis des bicelles.

A pH 6,4, ASP2 est en équilibre entre deux formes comme le montre le dédoublement de certaines raies, visibles sur le spectre aux deux températures. Ces deux formes sont attribuées à la forme B (majoritaire) et A (minoritaire).

5.1.2 Analyse des RDCs

Les spectres IPAP ont été traités sous XWINNMR et analysés avec le logiciel NmrView. La figure 5.5 rassemble les valeurs de couplages dipolaires résiduels mesurés à partir de la différence d'écartement entre les raies aux deux températures en tenant compte du signe

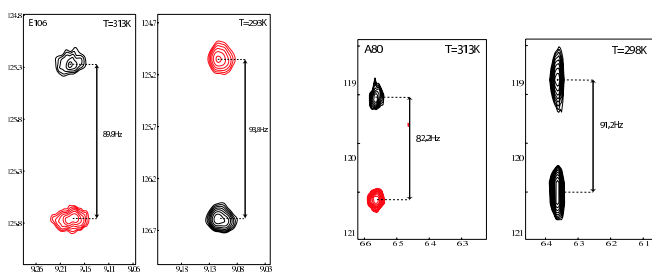


Fig. 5.3: Mesure des couplages dipolaires résiduels pour les acides aminés E106 (à gauche) et A80 (ci-dessus). Le couplage dipolaire résiduel du résidu E106 est la différence entre l'éclatement des raies dans l'IPAP à 313K et à 293K dans le milieu orienté par DMPC/DHPC-éther, soit 99Hz. La mesure sur le résidu A80 est effectuée entre 313K et 298K dans le milieu DMPC/DHPC-ester.



Fig. 5.4: Distribution des couplages dipolaires résiduels mesurés sur la structure d'ASP2 : en rouge, les résidus pour lesquels une mesure de RDC a été effectuée.

négatif du couplage scalaire $^1H-^{15}N$. Au total, les couplages dipolaires résiduels de 82 acides aminés ont pu être mesurés. La plupart des résidus de la partie bien définie de la structure ont pu donner des couplages observables du fait de la bonne séparation des raies. Très peu de couplages dipolaires résiduels ont été mesurés dans le segment 29-47 du fait de superpositions ou de l'absence de signaux, due à la largeur de raie importante (figure 5.4) induit par l'échange chimique. Pour les résidus dédoublés (attribués sur la figure 5.5), les deux valeurs de RDCs sont également données.

Les RDCs sur les hélices α sont en principe proches du fait de l'orientation des vecteurs $N-H$ sur un cône de demi-angle 15° . C'est ce l'on voit sur les hélices α_4 et α_6 et dans une moindre mesure, sur les hélices α_1 , α_3 et α_5 . Cependant, la figure 4.10 (partie Matériel et Méthodes) indique que pour certaines orientations de l'hélice par rapport au tenseur d'orientation, la variation des RDCs au sein de l'hélice peut être d'une grande amplitude (quand $\sin^2 \theta_h \cos 2\phi_h$ est maximal, voir équation 4.10). D'autre part, on constate que les RDCs des résidus impliqués dans les boucles interhélices sont très différents des RDCs dans les hélices qu'ils rejoignent

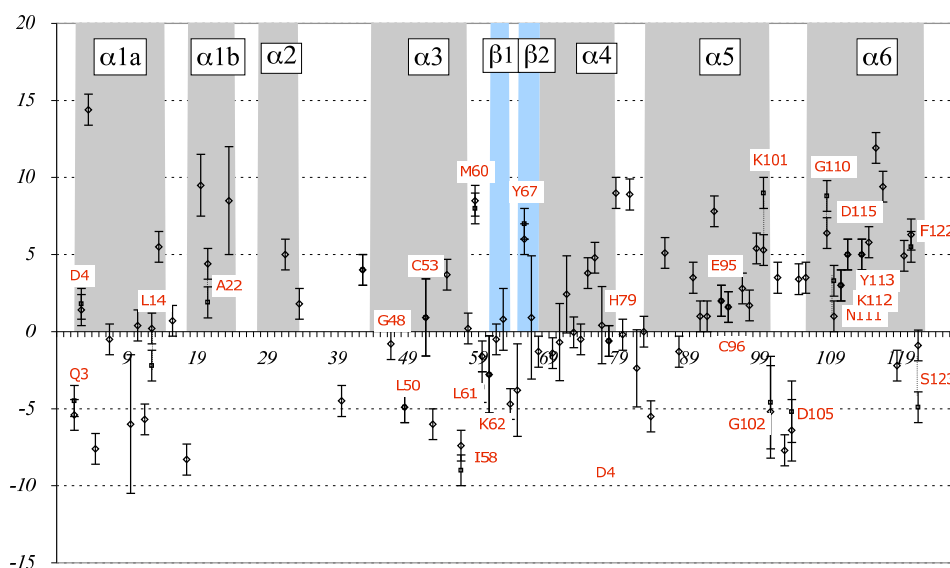


Fig. 5.5: Valeurs des couplages dipolaires en fonction de la séquence. Les structures secondaires sont marquées en couleur. Les résidus présentant deux taches de corrélations sont attribués. Lorsque deux couplages dipolaires résiduels fortement différents ont été mesurés sur un résidu, ces points sont reliés en pointillés (L14, A22, K101, G110 et S123).

(boucles α_{4-5} et α_{5-6}). Ceci traduit un changement radical dans l'orientation de ces vecteurs.

On peut noter aussi que les couplages dipolaires résiduels mesurés sur les deux formes A (minoritaire) et B (majoritaire) sont identiques (à l'incertitude près) pour une grande partie des résidus pour lesquels les déplacements chimiques de la liaison amide sont suffisamment différents : Q3, D4, M60, Y67, G102, D105 et F122. Par contre, les RDCs mesurés sur les résidus L14, A22, I58, K101, G110, N111 et S123 sont significativement différents dans les deux formes. Nous y reviendrons au paragraphe 5.5.

La distribution des RDCs est représentée sur la figure 5.6. Les valeurs extrêmes sont atteintes pour V6 (14,4Hz) et I58 (-9Hz) et la valeur la plus fréquente est 1,5Hz. En suivant l'analyse de Clore, on calcule : $D_a = 7,2\text{Hz}$ et $R = 0,49$. Cependant, cette méthode n'est pas complètement fiable pour ASP2, parce que l'espace des orientations des vecteurs $N-H$ n'est pas entièrement échantillonné du fait de la présence de nombreuses hélices. Comme la structure du complexe ASP2-TSP est connue, il est possible de déterminer les paramètres du tenseur d'orientation par Module (voir paragraphe suivant).

5.2 Rafinement de la structure d'ASP2 par les RDCs.

5.2.1 Objectif de la modélisation

La structure d'ASP2 n'est pas considérablement modifiée lorsque le TSP est présent dans la poche hydrophobe comme le montrent les spectres HSQC dans les deux situations. Il est donc raisonnable de considérer que la plupart des corrélations nOe observées sur le complexe seront conservées dans les formes A et B, avec cependant quelques variations possibles d'intensité.

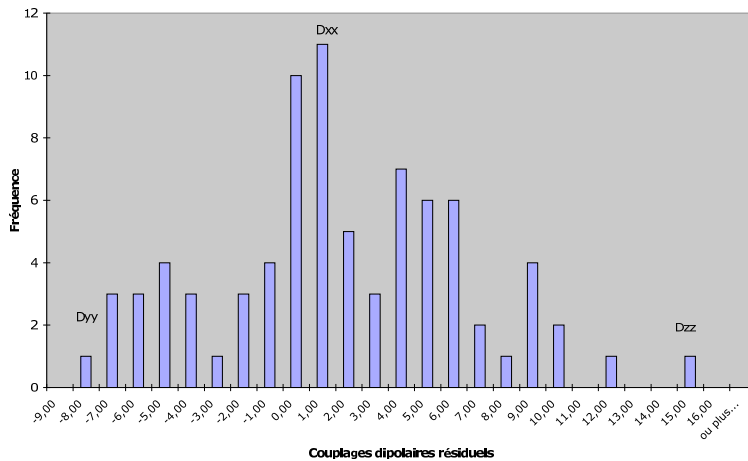


Fig. 5.6: Distribution des couplages dipolaires mesurés sur ASP2

Pour modéliser la forme libre d'ASP2, j'ai donc utilisé les mêmes contraintes que pour la reconstruction du complexe ASP2/TSP. L'introduction des couplages dipolaires résiduels dans le calcul devait en principe mettre en évidence les différences structurales.

5.2.2 Méthode

Les valeurs de RDCs sont sensibles à la fois à la dynamique du vecteur dipolaire sur la gamme pico-milliseconde et à l'existence d'échange conformationnel. Or les méthodes d'utilisation des RDCs comme contrainte angulaire impliquent que le paramètre d'ordre S_{dip} doit être homogène le long de la séquence. D'autre part, si le vecteur dipolaire dispose de deux ou plusieurs orientations différentes en échange rapide (vis-à-vis des couplages dipolaires résiduels), on ne peut plus corrélérer simplement la valeur mesurée à l'orientation du vecteur dipolaire dans ses différentes orientations. Par conséquent, j'ai éliminé les résidus dont la dynamique interne et le terme R_{ex} , susceptibles de traduire un échange conformationnel, étaient trop importants. En particulier, cette propriété interdit l'utilisation des couplages dipolaires résiduels dans la reconstruction pour le segment 30-45. D'autre part, je n'ai pas utilisé les RDCs dont la valeur était différente dans les deux formes A et B.

Le logiciel Aria permet d'incorporer directement les couplages dipolaires résiduels dans le calcul de structures. J'ai utilisé le mode SANI (optimisation des RDCs calculés par rapport aux RDCs expérimentaux). Un repère d'axes représentant l'orientation du tenseur est créé et les contraintes RDCs sont définies par les angles (ϕ , ψ) de chacun des vecteurs NH dans ce repère. L'orientation du repère sera optimisée durant la minimisation. Il s'agit ensuite de définir la magnitude D_a et la rhomboïté R du tenseur. Pour cela, j'ai tout d'abord obtenu par Module un premier jeu de paramètres à partir de la structure du complexe (7,2Hz et 0,5). J'ai ensuite recalculé les paramètres du tenseur sur la nouvelle structure moyenne (6,5Hz et 0,58) et j'ai relancé un calcul de 200 structures.

5.2.3 Résultats

Le tenseur d'orientation versus le tenseur d'inertie Les bicelles non chargées induisent une orientation des protéines par contact stérique. Ainsi, en théorie, le tenseur d'orientation d'une molécule sphérique dans un tel milieu bicellaire reste isotrope. L'utilisation de ces bicelles

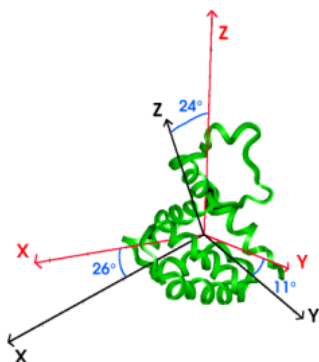


Fig. 5.7: Axes principaux des tenseurs d'inertie et d'orientation pour une structure.

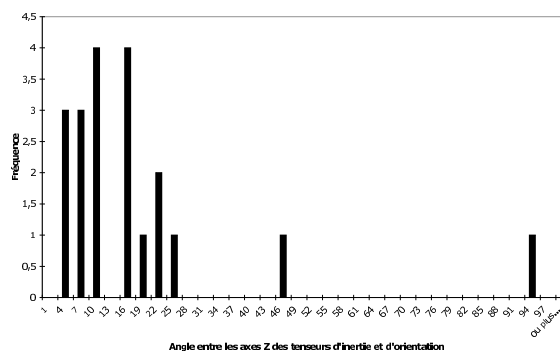


Fig. 5.8: Distribution des angles entre les axes x et Z des tenseurs d'inertie et d'orientation dans chacune des 20 structures.

est donc conseillée pour des protéines non sphérique, et le tenseur d'orientation est en principe colinéaire avec le tenseur d'inertie de la protéine. Nous pouvons donc comparer le tenseur d'orientation déduit des valeurs de couplage dipolaire résiduel avec le tenseur d'inertie d'ASP2 calculé au paragraphe 3.2.2.5. Nous pouvons dès à présent noter la parfaite adéquation entre la rhomboïté des tenseurs d'orientation et d'inertie (0,6).

Les axes du tenseur d'orientation ont été déduits des fichiers pdb (repère OO,X,Y,Z) après minimisation sous contraintes RDC. En fin de calcul, la structure est orientée dans le repère du tenseur d'inertie. Cependant, l'axe x du tenseur d'inertie est défini par rapport à la plus grande valeur principale du tenseur. Cet axe joue donc le rôle d'axe z (direction correspondant à la plus grande valeur de moment d'inertie). La figure 5.7 illustre l'orientation des tenseurs d'inertie et d'orientation pour une des structures. J'ai ensuite calculé l'angle formé par l'axe Z du tenseur d'orientation avec l'axe x du fichier pdb.

L'angle formé par les axes Z et x des deux tenseurs dans les 20 structures de plus basse énergie est relativement dispersé (figure 5.8). A l'exception de deux structures pour lesquelles les tenseurs d'orientation et d'inertie ont des orientations très distinctes, l'angle entre les axes Z et x est compris entre 2° et 24° . L'interaction avec les bicelles étant gouvernée par les contacts stériques, cela suggère que le tenseur d'inertie d'ASP2 calculé à partir du monomère est très proche du tenseur d'orientation. Cette observation est cohérente avec une forme monomérique d'ASP2 en solution, car les tenseurs d'inertie du monomère et de l'homodimère devraient être très différents.

Comme nous l'avons vu au paragraphe 3.2.2.5, la conformation du segment 30-45 influe sur l'orientation du tenseur d'inertie. Dans le jeu de structures dont nous disposons, ce segment est très mal défini, ce qui implique une distribution assez large des orientations du tenseur d'inertie. Si, en réalité, les différentes conformations adoptées par ce segment (observés par l'expérience CPMG) ont une distribution plus restreintes (par exemple, collées au coeur de la protéine), nous devrions retrouver une plus forte corrélation entre le tenseur d'orientation et le tenseur d'inertie des structures "fermées" qu'avec celui des structures "ouvertes". Il n'en est rien. Les segments 29-45 des structures correspondant aux angles les plus faibles ont une conformation aussi hétérogène que l'ensemble des structures.

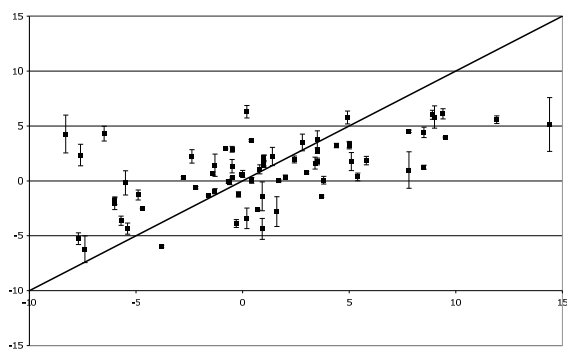


Fig. 5.9: Valeurs des RDCs calculées pour chaque résidu en fonction des valeurs expérimentales, sur la structure d'ASP2 de plus basse énergie avant introduction des contraintes RDCs.

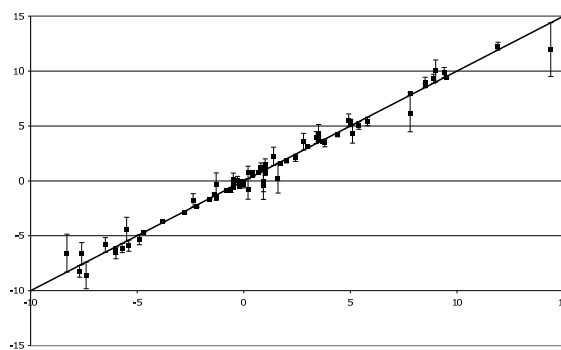


Fig. 5.10: Valeurs calculées des RDCs pour chaque résidu en fonction des valeurs expérimentales, sur la structure d'ASP2 de plus basse énergie après introduction des contraintes RDCs.

La minimisation sous contraintes La recherche par Module du tenseur d'orientation sur la structure de plus basse énergie à partir des RDCs sélectionnés (figure 5.9) montre que globalement, les couplages mesurés sont assez peu compatibles avec la structure initiale. Après introduction des contraintes issues des RDCs, la recherche du tenseur mène à des valeurs recalculées bien plus proches des valeurs expérimentales (figure 5.10). Ce résultat indique que la minimisation du potentiel harmonique dans Aria décrivant le couplage dipolaire résiduel a bien fonctionné.

La superposition des 20 structures de plus basse énergie (figure 5.11) suggère qu'à première vue, les contraintes de distances obtenues sur le complexe s'accordent relativement bien avec les RDCs mesurés sur la protéine libre. Cependant, l'écart quadratique moyen mesuré sur le squelette hors résidus 24 à 48 est supérieur ($0,98\text{\AA}$ par rapport à $0,71\text{\AA}$). Cette augmentation de l'écart type est liée à l'augmentation des violations dans le nouveau jeu de structures (123 ± 12 par rapport à 99 ± 11 violations par structure). Cependant, l'augmentation du nombre de violations ne peut être associée à une incompatibilité locale avec les contraintes de distance. En effet, comme pour les structures initiales, il n'est mesuré que 11 violations supérieures à $0,5\text{\AA}$, dont aucune n'est présente de façon récurrente dans les 20 structures. D'autre part, l'analyse des angles (ϕ , ψ) de l'ensemble des 20 structures par Procheck n'indique ni amélioration ni détérioration significatives de la géométrie du squelette.

La superposition des structures moyennes avant et après introduction des RDCs mènent à un écart quadratique moyen de $0,74\text{\AA}$ sur le squelette (excepté le segment E24 à G48). Cet écart se réduit à $0,54\text{\AA}$ sur les hélices α_3 à α_6 . La superposition des deux structures moyennes (figure 5.12) montre clairement que les plus grandes différences structurales sont localisées sur l'hélice N-terminale. En effet, l'angle formé par les hélices α_{1a} et α_{1b} diminue significativement de 135° à 120° . Ce changement d'orientation n'est à mon sens pas significatif car l'écart-quadratique par résidu est déjà important dans la structure du complexe, résultat, sans doute, de la mauvaise définition du segment suivant dans la séquence.

Compte tenu de la bonne convergence de la structure sur la partie structurée, nous pouvons conclure que les structures d'ASP2 libre en solution et en complexe avec le TSP sont identiques, du moins en ce qui concerne le squelette.

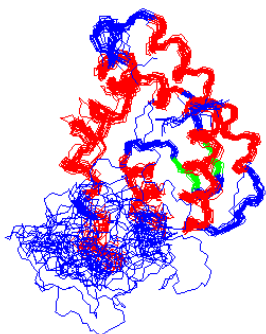


Fig. 5.11: Ensemble des 20 structures de plus basse énergie superposées entre les segments 1-23 et 49-123.

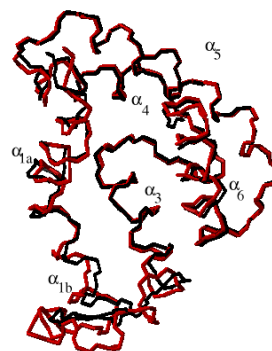


Fig. 5.12: Vue de la superposition de la structure moyenne d'ASP2 avant (rouge) et après (noir) introduction des couplages dipolaires résiduels.

Hélice	Résultats bruts			Résultats après prise en compte des dégénérescences			Angles observés dans les structures	
	$\theta(^{\circ})$	$\phi(^{\circ})$	ph($^{\circ}$)	$\theta(^{\circ})$	$\phi(^{\circ})$	ph($^{\circ}$)	$\theta(^{\circ})$	$\phi(^{\circ})$
α_{1a}	83,9	201,3	171,6	83,9	21,3	171,6	46,3	70,0
α_3	63,8	36,1	13,6	116,2	143,9	193,6	123,6	67,5
α_4	71,5	186,6	174,2	108,5	173,4	-5,8	97,1	168,4
α_5	59,4	176,2	105,5	59,4	-3,8	105,5	52,1	-23,8
α_6	126,9	3,3	19,6	126,9	183,3	19,6	145	140

Tab. 5.1: Résultat des ajustements de l'onde dipolaire sur les hélices d'ASP2

5.3 Les ondes dipolaires

Afin de valider une hypothèse que nous avons formulée au laboratoire et qui a été publiée dans le même temps par Mesleh, j'ai voulu étudier la périodicité des couplages dipolaires sur les hélices d'ASP2.

La magnitude D_a et la rhomboïcité R du tenseur d'orientation d'ASP2 dans les bicelles éther ont été déterminées. Il reste donc trois inconnues dans la formule 4.10 (partie Matériel et Méthodes) : les angles θ et ϕ déterminant l'orientation de l'hélice par rapport au tenseur d'orientation, et la phase ph au sein de l'hélice. J'ai alors déterminé ces paramètres à partir de l'ajustement de l'onde dipolaire sur les hélices α_{1a} et α_3 à α_6 . Etant donné le peu de résidus formant l'hélice α_{1b} , il n'est pas possible d'obtenir un ajustement raisonnable. Les meilleurs ajustements sont illustrés sur la figure 5.13. La colonne de gauche du tableau 5.1 recense le résultat des ajustements pour chacune des hélices. Les angles θ et ϕ peuvent alors être comparés avec les angles mesurés sur la structure. Pour cela, j'ai utilisé InsightII. La direction des hélices a été définie par un vecteur normé \vec{u} reliant deux carbones aux extrémités des hélices dont la direction est visuellement parallèle à l'axe de l'hélice. J'ai ensuite calculé les coordonnées polaires de \vec{u} par les relations $\cos(\theta) = \vec{u} \cdot \vec{z}$, $\cos(\phi) = \vec{u} \cdot \vec{x} / \sin(\theta)$ et $\sin(\phi) = \vec{u} \cdot \vec{y} / \sin(\theta)$. Ceci permet de déterminer la valeur et le signe des angles. La colonne de droite du tableau rassemble

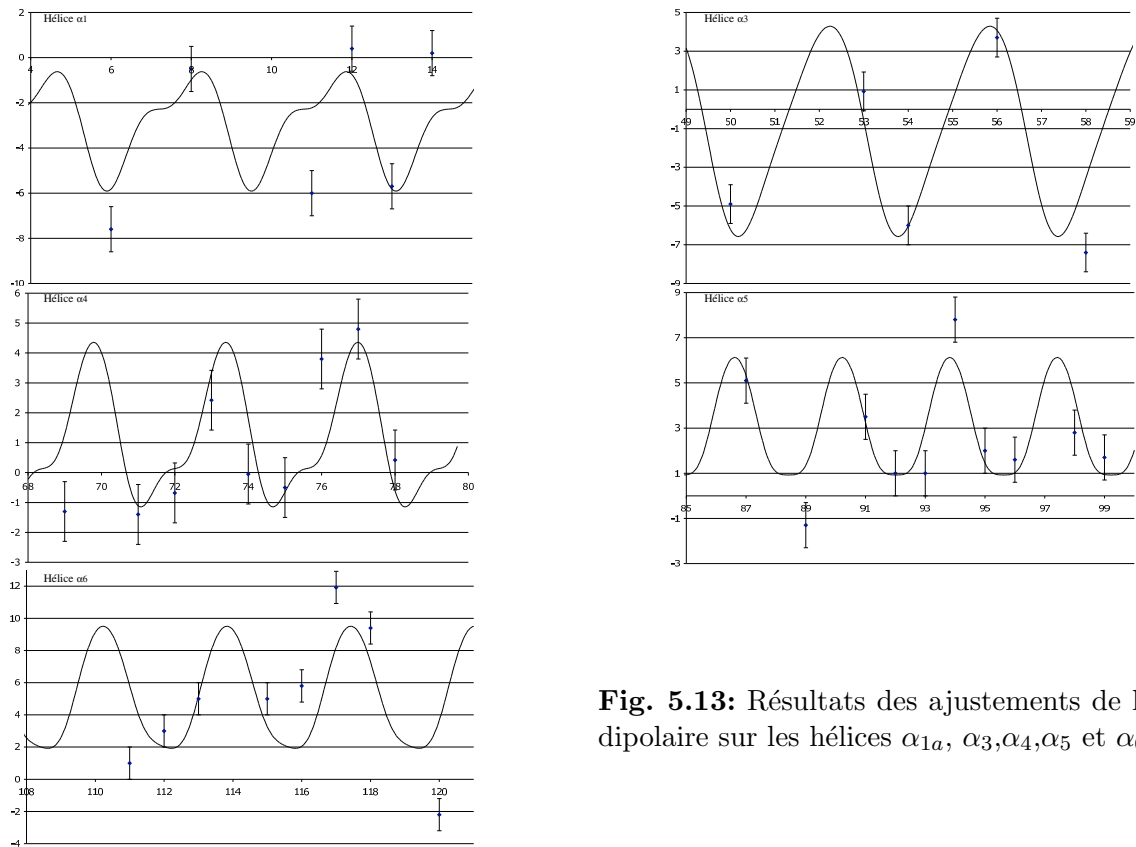


Fig. 5.13: Résultats des ajustements de l'onde dipolaire sur les hélices α_{1a} , α_3 , α_4 , α_5 et α_6 .

les angles des cinq hélices analysées. Bien entendu, il faut prendre en compte l'incertitude sur la direction de l'hélice, sur un jeu de structures. On peut estimer raisonnablement que suivant les méthodes employées pour définir ces angles, l'angle de ces hélices peut avoir une dispersion de quelques degrés.

Il est clair sur les courbes des meilleurs ajustements qu'aucune des séries de couplages dipolaires résiduels sur une hélice ne reflète parfaitement l'onde dipolaire. Les ondes dipolaires sur les hélices α_3 , α_4 et α_5 sont celles qui s'ajustent le mieux aux données, bien que pour un ou deux résidus centraux de l'hélice (E76, et V89-N94), les RDCs sont assez éloignées de leurs valeurs déterminées par l'ajustement. En manipulant les angles issus de l'ajustement et les dégénérescences, j'ai déterminé ceux qui se rapprochaient le plus des valeurs observées dans la structure. A l'exception de l'angle ϕ de l'hélice α_3 , les angles produits déduits par l'onde dipolaire sont identiques dans une marge de moins de 20° , ce qui est tout à fait satisfaisant. L'hélice α_4 est celle qui donne des corrélations les plus probantes : environ 10° d'écart. En revanche, les ajustements sur les hélices α_{1a} et α_6 donnent des corrélations plutôt médiocres. La mauvaise qualité de l'ajustement se traduit également par une différence importante entre l'orientation prédite pour ces hélices et celle observée dans la structure.

Dans un article très récent (Mesleh and Opella [2003]) traitant des effets de déformations des hélices (courbure de l'hélice et coude induit par une proline) sur l'onde dipolaire, Mesleh et ses collaborateurs montrent que la déviation d'une hélice par rapport à la géométrie idéale se traduit par une déformation de l'onde dipolaire, soit en phase (ph), soit en magnitude (θ_h, ϕ_h), soit les deux effets simultanément. Un bon ajustement de l'onde dipolaire est donc une façon de

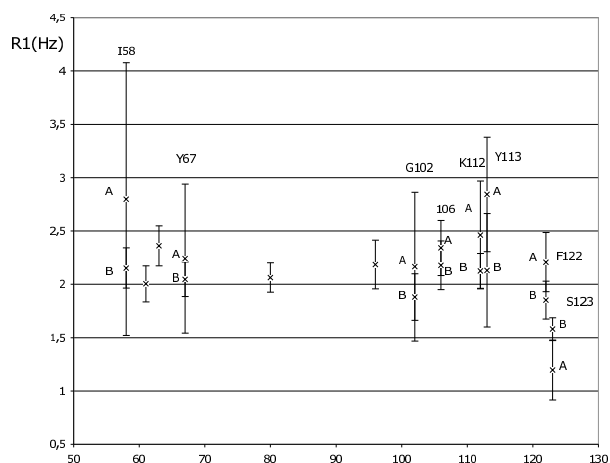


Fig. 5.14: Valeurs de R_1 mesurées sur les formes A et B.

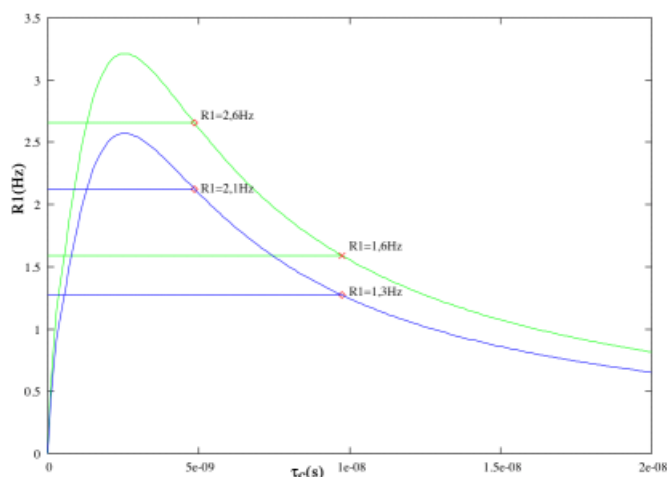


Fig. 5.15: Evolution du R_1 en fonction du temps de corrélation global τ_c pour un S^2 de 1 (en vert) et un S^2 de 0,8 (en bleu) à un champ B_0 de 600MHz.

détecter des déformations, même très faibles, par rapport à la structure idéale. Nous pouvons alors mettre nos résultats en regard de cette étude.

Les cartes NOESY-HSQC éditées ^{15}N et ^{13}C ainsi que l'expérience d'échange $^1H/^2H$ montrent que les nOe caractéristiques des hélices et les liaisons hydrogènes sont constamment conservés sur les hélices α_4 et α_5 . Je n'ai observé de corrélations $H_{\alpha}^{i-4} - H_N^i$ que dans ces hélices. La forte dispersion dans la carte ^{15}N -HSQC des taches correspondant à ces résidus est également en faveur d'une très bonne structuration de ces hélices. En revanche, les signaux issus des hélices α_{1a} , α_3 (partie N-terminale de l'hélice) et α_6 dans les différents spectres (faibles), indiquent que ces hélices ne sont pas aussi parfaites que les deux premières. Il semble donc qu'un mauvais ajustement de l'onde dipolaire trahit une déformation de ces hélices.

En conclusion de cette étude, nous pouvons dire que l'onde dipolaire est une bonne méthode pour détecter les déformations dans l'hélice. Cependant, l'utilisation de cette analyse des couplages dipolaires nécessite une bonne qualité des couplages mesurés et surtout un nombre important de mesures sur une hélice. En effet, moins nous disposons de couplages, plus l'ajustement sera mauvais et plus l'erreur sur la détermination de l'orientation de l'hélice sera grande.

5.4 La dynamique d'ASP2 libre en solution : monomère ou dimère ?

L'équilibre entre les deux formes A et B peut tout à fait correspondre à un équilibre lent (vis-à-vis des déplacements chimique) entre un monomère et un dimère d'ASP2. La différence de masse et donc de temps de corrélation τ_c entre le monomère et le dimère est susceptible de pouvoir être détectée par la mesure de la relaxation. Afin d'approfondir l'hypothèse de deux états oligomériques, j'ai mené une expérience de relaxation R_1 . Bien que n'étant pas la vitesse de relaxation la plus sensible vis-à-vis du temps de corrélation, du moins dans la gamme de τ_c

d'une protéine, elle présente l'avantage de ne pas être contaminée par la relaxation d'échange, contrairement au R_2 .

La mesure du R_1 a été effectuée sur un tube contenant 0,3mM d'ASP2 marquée ^{15}N à pH 5,8, en absence de TSP et de sel. La carte ^{15}N -HSQC montre que la forme B est majoritaire et que la forme A est très minoritaire. J'ai utilisé le même protocole de mesure de R_1 qu'au paragraphe 4.1 (expériences et traitement). La valeur de R_1 a été mesurée sur les résidus présents sous deux formes et isolés sur les spectres (I58, Y67, G102, E106, K112, Y113, F122 et S123, figure 5.14) ainsi que sur les résidus L61, G63, A80 dont les déplacements chimiques sont uniques dans les deux formes et sur la résonance de la forme B de C96.

L'intensité des résonances sous la forme A est très faible en comparaison avec celle de la forme B. Ceci mène à une incertitude plus grande pour les R_1 de la forme A. Les intervalles de confiance de R_1 sur les deux formes se recouvrent pour tous les résidus dans les deux formes. On remarque cependant que les R_1 de la forme A sont systématiquement supérieurs aux R_1 mesurés sur la forme B. Il est donc possible que nous ayons surestimé l'incertitude. Nous pouvons alors considérer que la relaxation longitudinale des azotes du squelette de la forme A est plus lente que dans la forme B.

La figure 5.15 rappelle l'évolution du R_1 en fonction du temps de corrélation pour un paramètre d'ordre S^2 de 1 et de 0,8, en se plaçant dans le modèle simplifié de Lipari-Szabo. Les valeurs de R_1 au temps de corrélation du complexe ASP2/TSP et au double du temps de corrélation correspondant à un dimère, sont marquées pour les deux valeurs de paramètre d'ordre.

Le R_1 mesuré sur la forme B est supérieure à 1,9Hz et est en moyenne supérieure à 2Hz, à l'exception des deux résidus en position C-terminale de R_1 plus faible. Ceci avait déjà été observé dans l'étude du complexe et interprété par la présence de mouvements internes de forte amplitude. La valeur moyenne des R_1 de la forme B est sensiblement identique à celle observée sur le complexe (2Hz). La valeur attendue pour un dimère étant de 1,3Hz, soit très éloignée de 2Hz, il est raisonnable de conclure que la forme B correspond à un monomère d'ASP2.

De même, du fait d'une valeur plus élevée de R_1 , la forme A ne peut en aucun cas être une forme dimérique. Cette différence de vitesse de relaxation longitudinale, si elle est vraiment significative, peut indiquer, soit un temps de corrélation légèrement plus faible, soit à une restriction plus importante des mouvements dans cette forme que dans la forme B.

Nous pouvons conclure de cette étude que les formes A et B correspondent à des états monomériques d'ASP2. Cependant, nos données indiquent qu'il est possible qu'une différence très légère de comportement hydrodynamique ou de dynamique interne existe entre elles.

5.5 Discussion sur les formes A et B

Les deux formes sont monomériques La titration en pH menée au premier chapitre nous a montré qu'à haut pH (8-8,5), ASP2 adoptait deux états en échange lent entre eux et qu'en baissant le pH à 6,1, la forme A devenait majoritaire. L'environnement chimique des protons amides est très proche dans les deux formes à l'exception des résidus autour du pont disulfure C_{96} - C_{116} , ce qui indique que les formes A et B ont une structure tridimensionnelle très proche l'une de l'autre et celle-ci est compatible globalement avec la structure

du complexe ASP2/TSP, comme nous l'a montré la minimisation sous contraintes issues des couplages dipolaires résiduels. Par ailleurs, la question de l'équilibre monomère-dimère peut être exclue (colinéarité des tenseurs d'orientation et d'inertie et valeurs de R_1). Les différences qui s'établissent entre les deux ne peuvent donc résider qu'au sein d'un monomère.

Il existe une faible différence de conformation du squelette La mesure des couplages dipolaires résiduels a mis en évidence des différences significatives pour les résidus L14, A22, I58, K101, G110, N111 et S123. Si une valeur identique de RDC pour un vecteur ne nous donne pas d'information directement exploitable, le fait que les RDCs soient différents pour une partie des résidus et identiques pour les autres implique plusieurs possibilités : soit le tenseur d'orientation est différent pour les deux formes, soit l'orientation du vecteur NH est différente dans les deux formes, soit la dynamique de ce vecteur est différente. La première hypothèse pourrait expliquer le profil de RDCs obtenu, tout en conservant une conformation du squelette identique entre les deux formes. Du fait de la dégénérescence de l'expression du couplage dipolaire résiduel vis-à-vis de l'orientation des vecteurs, il est possible que des valeurs différentes de RDCs ne soient mesurées que sur un nombre faible de résidus. Cependant, les deux tenseurs d'orientation devraient alors être très proches et la différence d'orientation des vecteurs suffisamment grandes. Dans la cadre de la seconde hypothèse, le tenseur d'orientation est le même pour les deux formes, et la différence de couplage provient d'une orientation différente des vecteurs dans ce tenseur. Comment expliquer l'existence de deux tenseurs d'orientation ? L'hypothèse de l'équilibre monomère-dimère ne tient plus, et un changement de volume global de la structure est improbable. Je déduis donc que les vecteurs amides des résidus incriminés ont une orientation différente dans les deux formes. Il est possible également que d'autres liaisons de la séquence aient une orientation différente dans les deux formes, bien que donnant des couplages dipolaires résiduels identiques.

Les vecteurs amides ayant une orientation différente dans les deux formes ne sont pas concentrés sur la figure. Ceci ne permet pas de détecter sur la structure, une zone particulière qui subirait un changement de conformation du squelette mesurable. Je pense donc que la conformation du squelette dans les deux formes est globalement modifiée.

L'isomérisation du pont disulfure C96-C116 : le siège initial de la déformation.

La différence de déplacement chimique entre les deux formes n'est pas la même pour tous les résidus. Les résidus les plus affectés en terme de déplacement chimique sont localisés autour du pont disulfure C96-C116. L'environnement magnétique est donc très différent dans cette région entre les deux formes. On peut attribuer cela au phénomène d'isomérisation du pont disulfure. Le changement de conformation de la liaison S-S pourrait expliquer, par un effet de levier, les déformations significatives observées sur tout le segment reliant ces deux cystéines, et en particulier, la boucle α_5 - α_6 (résidus 101 à 106), ainsi que sur la partie N-terminale de l'hélice α_3 et sur le β -hairpin, dont plusieurs taches sont présentes en double.

L'isomérisation de ce pont disulfure permet ainsi d'expliquer les faibles différences de conformations observées sur les deux formes.

L'histidine H79 : un candidat pour l'interconversion A \leftrightarrow B induite par le pH L'expérience de titration en pH a clairement montré que le déplacement chimique du cycle aro-

matique de la tyrosine 117 et de la phénylalanine 122 étaient différents dans les deux formes. L'environnement chimique de ces deux cycles est donc différent. Cela peut s'expliquer soit par une orientation différente d'une chaîne latérale proche de ces noyaux (voire eux-mêmes), soit par les états de protonation d'un groupe acido-basique.

Le fait que les populations des formes A et B soient inversées par une baisse du pH, et que ce phénomène soit réversible en fonction du pH, suggère que l'état de protonation d'un groupe proche de ces aromatiques est relié à l'existence de ces deux formes. Disposant de la structure, nous pouvons proposer un candidat. La gamme de pH dans laquelle nous nous situons suggère que le pKa du groupe en question se situe autour de 7. Le cycle aromatique de l'histidine possède un pKa de cette valeur, et il existe deux histidines dans ASP2 : H26 et H79. La chaîne latérale de H26 est orientée vers le solvant et ne peut donc avoir aucun effet sur la structure globale. En revanche, le groupe aromatique de H79 est située au niveau du point de convergence du faisceau d'hélices. Par ailleurs, le signal du proton H_ϵ subit un déplacement caractéristique d'une titration de pKa dans la gamme de pH étudiée : nous suivons donc la protonation d'un des deux azotes du cycle imidazole. Nous pouvons alors proposer que la protonation de l'histidine 79 stabilise la forme A au détriment de la forme B.

L'hypothèse d'un équilibre entre les formes, dépendant de l'état de protonation de l'histidine 79 est à mettre en regard avec la vitesse de l'équilibre entre ces deux formes. L'interconversion entre les deux formes est lente ce qui induit une barrière énergétique importante à franchir pour le passage d'une forme à l'autre. On peut donc proposer un modèle dans lequel la protonation du groupe imidazole permet de former un pont salin entre l'azote protoné de H79 et un groupe accepteur de la protéine (D85 ?). La formation de ce pont salin déclenche ensuite une série de réarrangements, dont peut-être, l'isomérisation du pont disulfure C96-C116, globalement coûteux en énergie, et donc lents par rapport à l'échelle des déplacements chimiques.

Conclusion Nous avons en partie caractérisé les différences entre les deux formes A et B, nous avons identifié un phénomène responsable potentiellement de la différence de conformations et un candidat expliquant l'inversion de populations A et B induite par le pH. En résumé, ASP2 peut adopter deux conformations A et B. Les populations de ces deux états seraient liées à la formation d'un pont salin, dépendant de l'état de protonation de l'histidine H79. Ce modèle permet d'expliquer nos données mais il reste encore à l'état de proposition. D'autres données, comme la détermination du pKa de l'histidine H79 et de la constante d'équilibre de la réaction d'interconversion, seraient indispensables pour conclure définitivement sur ce point.

Chapitre 6

Étude des interactions d'ASP2 avec le TSP et l'IBMP

Nous atteignons maintenant le troisième et dernier objectif que nous nous étions fixé au début de mon travail de thèse. Lors des étapes d'attribution et de reconstruction, il s'est avéré qu'ASP2 possédait une propriété qui a modifié mon approche du sujet. Le TSP est utilisé pour la calibration des spectres de RMN et n'interagit pas en général avec les objets biologiques. Le fait qu'ASP2 lie le TSP et que nous ayons essentiellement travaillé sur le complexe a donc forcé des interconnexions entre l'étude structurale et dynamique et la partie interaction. Le lecteur trouvera dans la section 6.1, les données propres à l'interaction ASP2/TSP dont quelques résultats principaux ont déjà été cités précédemment. Le TSP n'a que peu d'intérêt biologique et j'ai donc choisi d'étudier les interactions avec une molécule plus significative : la 2-isobutyl-3-méthoxypyrazine (IBMP). C'est l'objet de la section 6.3 est l'étude de la compétition de l'IBMP contre le complexe ASP2/TSP. Enfin, la section 6.4. La dernière section expérimentale (6.5) traite d'une expérience dont les résultats sont décevants mais j'ai choisi de la rapporter car elle apporte tout de même certaines informations. Nous discuterons dans la dernière section des conclusions partielles que nous pouvons tirer des propriétés d'interaction d'ASP2.

6.1 Une expérience “TSP contre ASP2- ^{15}N ”

6.1.1 Objectif et protocole expérimental

Cette expérience a pour objectif d'observer l'évolution d'ASP2 libre en solution avec l'ajout de TSP pur.

Une solution de $500\mu l$ contenant $0,7mM$ d'ASP2- ^{15}N de la préparation P02, $100mM$ de tampon phosphate a été préparée. Le tampon a fixé le pH à 6,1. Le TSP en poudre a été ajouté de façon à obtenir un excès de 3 :1 puis de 28 :1 par rapport à la protéine. Les excès ont été déterminés à partir du pic du TSP et du proton H_ϵ de l'histidine H79. Chaque point a été contrôlé par une ^{15}N -HSQC et une Watergate 1d.

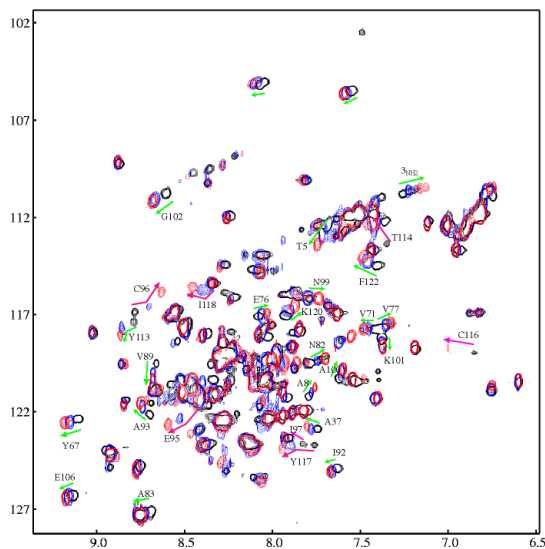


Fig. 6.1: Titration en TSP contrôlée par ^{15}N -HSQC. En noir, sans TSP. En bleu, le TSP est en excès de 3 :1 et en rouge, en excès de 28 :1. Le déplacement de quelques taches est suivi par une flèche. Les résonances caractéristiques des formes A et B sont suivies par une flèche violette tandis que les autres sont en vert.

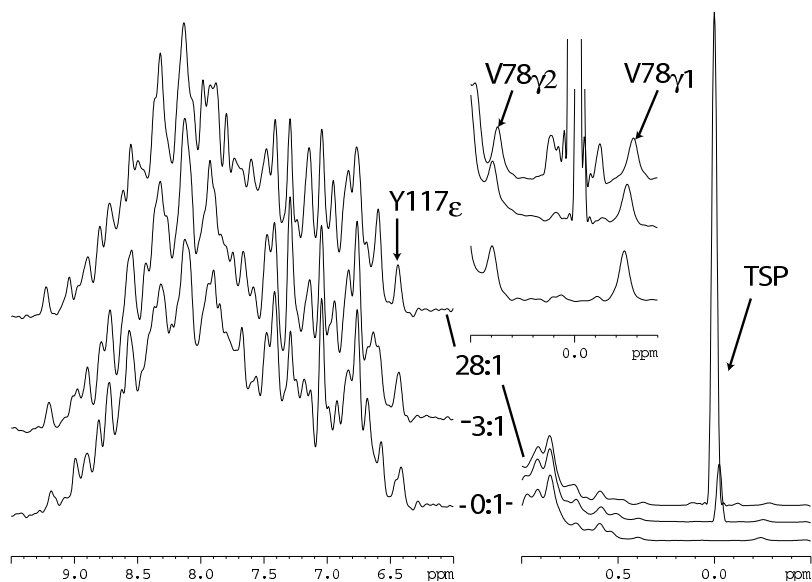


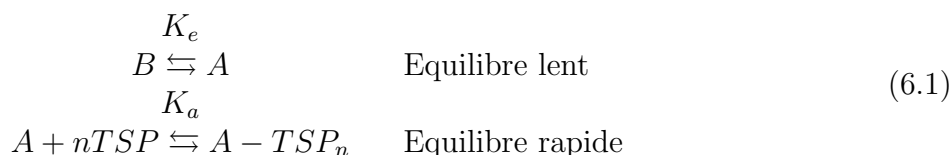
Fig. 6.2: Titration en TSP contrôlée par des spectres 1d Watergate calibrés sur l'eau. De bas en haut, sans TSP, en excès de 3 :1 et en excès de 28 :1.

6.1.2 Analyse des spectres

Le premier spectre d'ASP2 seule en solution à pH 6,1 suggère que la protéine est présente sous la forme B majoritaire et la forme A minoritaire (d'après les intensités des deux formes de E95, C96, C116, Y117). Contrairement à la titration en pH de la forme libre, la forme B est donc majoritaire à ce pH. Il est possible que la présence de sel modifie le pH de transition entre les deux formes.

Par ajout successif de TSP, il est évident que cette petite molécule interagit avec ASP2 : un grand nombre de raies sont déplacées, à la fois sur le spectre HSQC et sur le spectre 1d Watergate (figures 6.1 et 6.2). De plus, le signal du TSP qui apparaît dans les spectres 1d-Watergate est un déplacement chimique différent entre les spectres à 3 :1 et à 28 :1. Contrairement à ce qui était attendu dans le cas d'une titration simple du type $P+L \rightleftharpoons PL$, où seules les résonances proches du site de fixation sont affectées, nous observons ici un déplacement de très grand nombre de résonances, associées à des résidus dispersés sur la structure (par exemple A37, E106, Y67, Q3_{NH2}, V78_γ). La présence de TSP dans le coeur hydrophobe est sans ambiguïté comme le montrent les nOes entre les protons aromatiques des tyrosines Y113, Y117, Y10 (spectre non montré) et les méthyles du TSP. Cependant, les protons ϵ et/ou ζ de la phénylalanine 122 donnent également une corrélation avec le TSP alors que celle-ci n'est pas dans le coeur hydrophobe.

Le mécanisme de fixation du TSP n'est donc pas simple. Il est remarquable que les amides des acides aminés caractéristiques E95, C96, I97, Y117 et I118 se déplacent progressivement vers des déplacements chimiques proches de la forme A. Cette expérience suggère donc que la fixation de TSP induit une inversion de populations entre les deux formes A et B. L'explication la plus probable est que seule la forme A fixe le TSP. Bien que nous n'ayons que deux points de titration, le déplacement progressif du pic du TSP traduit que la vitesse d'échange entre le TSP libre et le TSP lié est rapide. Les réactions que nous observons ici sont donc les suivantes :



Ces équilibres contiennent un rapport stoechiométrique n que nos données ne permettent de déterminer précisément. Une observation attentive de la HSQC montre que certaines résonances sont stabilisées en fréquence dès l'excès 3 :1 (G91, I92, K101, G102, E106) alors que d'autres (Q3_{NH2}, A8, Y67, V89, A93, E95, C96, I97, I117, K120, F122) se sont déplacées entre 3 :1 et 28 :1. Compte tenu de cette observation et des nOes observés entre le TSP et ASP2, non compatibles avec un seul site de fixation, je conclus que le TSP se fixe à au moins deux endroits de la protéine, dont à l'intérieur de la poche hydrophobe.

6.2 IBMP : un bon modèle de molécule odorante

L'IBMP (2-isobutyl-3-méthoxy-pyrazine) est une petite molécule qui pour l'homme a l'odeur de poivron. Son seuil de perception chez l'homme est l'un des plus faibles mesurés. Elle fait partie des molécules olfactives rentrant dans la composition du poivron, du café, des petits

pois, de la résine, du vin et des produits issus de la pomme de terre²³. On la trouve également dans le tabac et elle est utilisée en parfumerie pour ses notes poivrées, métalliques, irritantes et vertes.

Il a été montré que l'IBMP pouvait se lier avec des constantes de dissociation de l'ordre du micromolaire à des protéines de la famille, au sens large, des OBPs de vertébrés (OBP de porc, de boeuf, lipocaline de la glande salivaire du verrat, Monte et al. [1993], Marchese et al. [1998], Pelosi [1994], Burova et al. [1999]). Le travail de l'équipe de l'INRA sur les études de fixation de petites molécules sur ASP2 montrait que ASP2 liait l'IBMP également. Dans le cadre de l'étude par RMN des interactions d'ASP2 avec des odeurs, cette molécule présentait sur les autres molécules fixées par ASP2 l'avantage de posséder des protons aromatiques qui devaient *a priori* résonner dans une région isolée du spectre. Nous avons donc mené une série d'expériences afin de mieux comprendre l'interaction entre ASP2 et l'IBMP.

6.2.1 Caractéristiques de l'IBMP

La structure de l'IBMP contient un cycle pyrazine substitué par des groupements méthoxy et isobutyle (figure 6.3). Son poids moléculaire est de 166,22g/mol et sa densité par rapport à l'eau est de 0,886. En fonction de l'orientation des groupes isobutyl et méthoxy, le volume de la molécule peut différer. La molécule peut être grossièrement schématisée par un pavé de côtés 7Å*5Å*3Å.

6.2.2 Attribution de l'IBMP dans l'eau

L'attribution de raies de l'IBMP libre est importante pour l'interprétation des cartes d'interaction avec la protéine. J'ai donc accumulé un spectre 1d-Watergate (figure 6.3) de la solution à partir de laquelle j'ai mené une des titrations (expérience IBMP contre ASP2-¹⁵N, voir paragraphe suivant), diluée dans 450µl de D₂O afin de ne pas être gêné par le signal de l'eau. La concentration en IBMP est alors de environ 4mM et de 0,05% en méthanol. A partir des déplacements chimiques attendus pour le type de protons et des couplages scalaires observés, ce spectre est suffisant pour attribuer les résonances de l'IBMP (tableau 6.1). Les deux protons aromatiques ont un déplacement chimique légèrement différent provoqué par l'absence de symétrie de la molécule. Les deux méthyles résonnent à la même fréquence, ce qui n'est pas étonnant, compte tenu des symétries des trois rotamères et de l'absence de centre asymétrique. Les seuls nOes visibles à cette température impliquent les méthyles H3 et H4 vers le proton H5 et les protons H6 et H6'. Ces nOes sont de signe inverse de la diagonale. On peut donc conclure que l'IBMP en solution dans l'eau n'adopte pas un arrangement créant un objet de grande taille (empilement, ...).

²³<http://www.thegoodscentscompany.com/data/rw1013601.html>

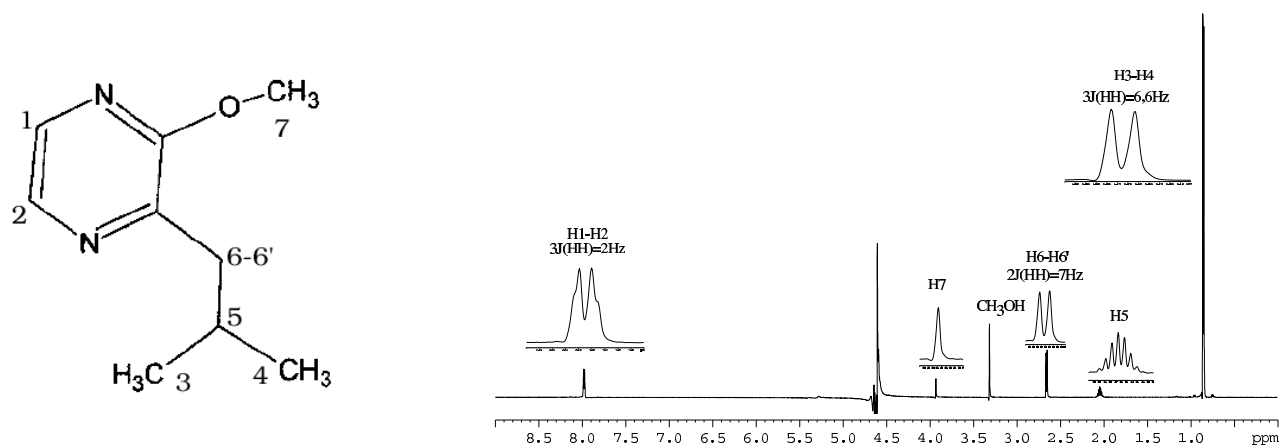


Fig. 6.3: A gauche, la structure de l'IBMP. A droite, le spectre 1D Watergate de l'IBMP (4mM) dans un mélange $\text{H}_2\text{O}/\text{D}_2\text{O}$ (90/10), à 0,05% de méthanol à 308K. Les pics de l'IBMP sont agrandis et les valeurs de constantes de couplage sont reportés.

Hydrogène	δ
H1-H2	8,00-7,98
H3-H4	0,87
H5	2,06
H6-H6'	2,67
H7	3,94

Couplage scalaire	Valeur
${}^3J_{H1-H2}$	2Hz
${}^2J_{H6-H6'}$	7Hz
${}^3J_{H3-H5}$	6,6Hz

Tab. 6.1: Attribution de l'IBMP et valeurs des constantes de couplage scalaire.

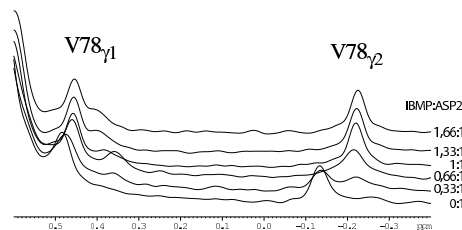
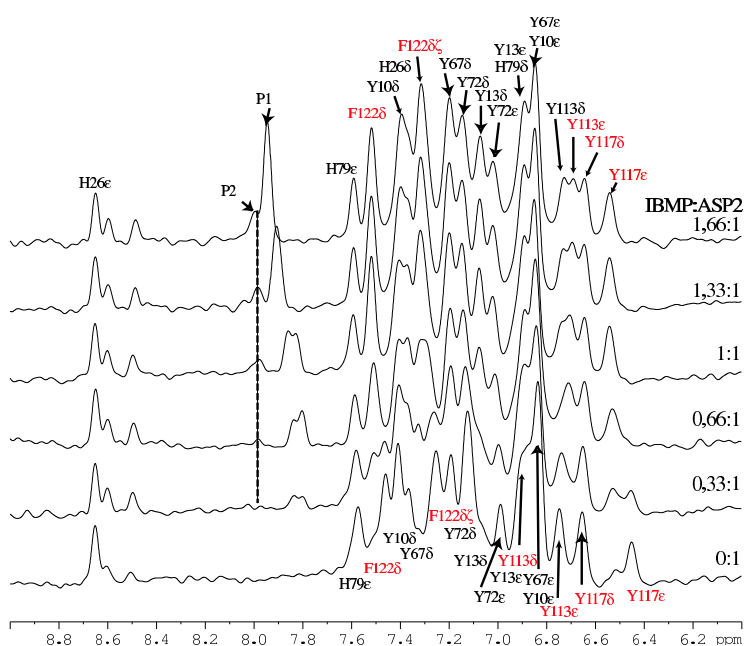


Fig. 6.4: (ci-dessus) Titration des méthyles de la valine 78 (Watergate filtrée)

Fig. 6.5: (à gauche) Titration des aromatiques (Watergate filtrée)

6.3 Une expérience “IBMP contre ASP2- ^{15}N ”

6.3.1 Protocole expérimental

Cette expérience a été menée sur un tube d'ASP2 marqué ^{15}N (0,31mM) contenant de l'azide (60mM) et 50mM de tampon phosphate. Une solution de 700 μ l d'eau légère contenant l'IBMP a été préparée : 39mM en IBMP, 0,5% (v/v) de méthanol. Le méthanol aide à dissoudre l'IBMP dans l'eau. A ces concentrations, chaque ajout de 2 μ l de la solution d'IBMP augmente le rapport de concentration IBMP :ASP2 de 0,33. La dissolution totale de l'IBMP ajouté dans la solution d'ASP2 a été confirmée par les rapports entre les signaux des aromatiques de l'IBMP et de la protéine. Au bout de trois ajouts, l'état 1 :1 est atteint. J'ai ajouté l'IBMP pour arriver à l'excès 1,66 :1.

La titration a été suivie par une série de ^{15}N -HSQC, de Watergate filtrée ^{15}N et de NOESY filtrée ^{15}N .

6.3.2 Analyse des spectres

Le premier spectre Watergate ainsi que la 2D-NOESY filtrée ^{15}N traduit l'existence de deux formes (A et B) pour les chaînes latérales aromatiques des résidus Y113, Y117 et F122 (figure 6.5). Les ajouts successifs d'IBMP se sont traduits par l'apparition de trois pics : un pic double vers 7,8 ppm (P1) et un pic simple vers 7,95ppm (P2), qui se déplacent vers les déplacements chimiques des aromatiques de l'IBMP seul en solution. Ces pics ne peuvent être attribués qu'à l'IBMP. Cependant, leurs intensités relatives suggèrent qu'ils ne peuvent appartenir à la même molécule. L'IBMP est donc au moins sous deux formes liées.

Le spectre des aromatiques de la protéine est considérablement affecté par l'augmentation en concentration en IBMP. L'attribution des aromatiques est rendue possible par les NOESY filtrées. Au cours de la titration, on observe une inversion entre les deux formes (figure 6.7).

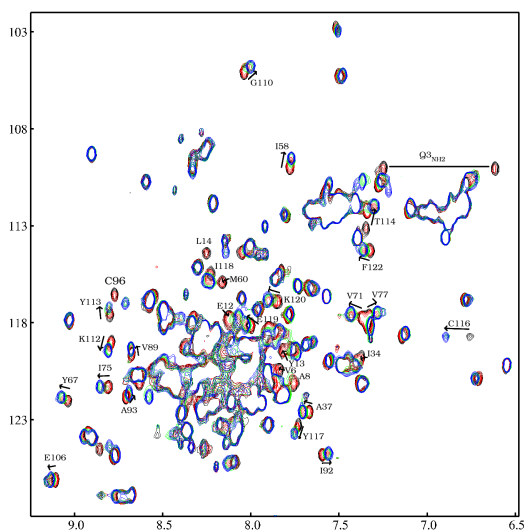


Fig. 6.6: Titration d'ASP2 par IBMP à 308K suivie par HSQC (0 :1 en noir, 0,33 :1 en rouge, 0,66 :1 en vert, 1 :0 en bleu).

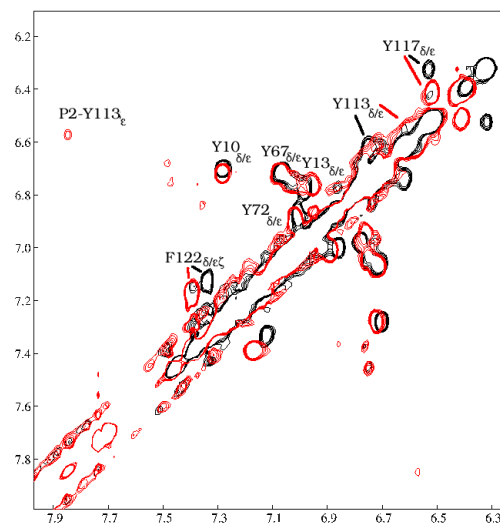


Fig. 6.7: Superposition de la région aromatique-aromatique de la NOESY filtrée sans IBMP (noir) et dans un rapport 1 :1 (rouge).

Ce phénomène pourrait a priori être interprété comme étant la titration d'un équilibre lent vis-à-vis des déplacements chimiques. Cependant, le fait que les deux formes existent avant tout ajout d'IBMP exclut cette hypothèse. Les aromatiques dont les déplacements chimiques sont les plus différents entre les deux formes, et dont les populations s'inversent, sont F122, Y113 et Y117. Les chaînes latérales de Y10, Y13, Y67, Y72 ainsi que V78 (figure 6.4) sont également faiblement déplacées lors de la formation du complexe. Si Y10, Y113 et Y117 font partie du coeur hydrophobe, les autres aromatiques n'en font pas partie.

La titration suivie par HSQC (figure 6.6) indique que la forme B est majoritaire avant ajout d'IBMP et que les amides d'ASP2 sont perturbés par la présence d'IBMP. Le spectre en fin de titration montre qu'une seule forme est présente et celle-ci a été attribuée grâce à un spectre 3D-NOESY-HSQC. A l'exception de quelques résidus (N34, A37) dont les résonances

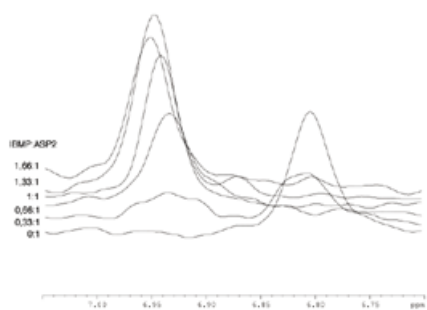


Fig. 6.8: Titration du proton amide du résidu C116. Les spectres sont extraits de HSQC au déplacement chimique de l'azote (118,6ppm).

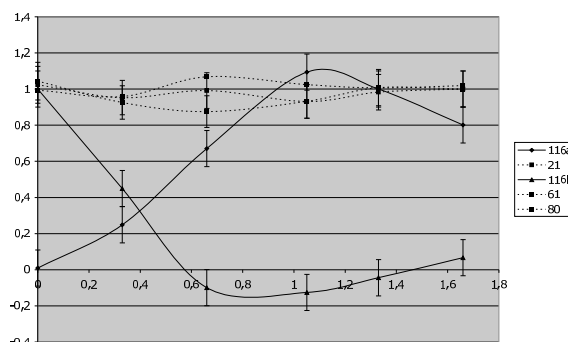
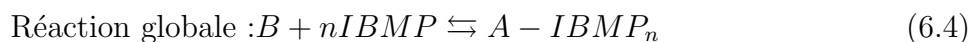


Fig. 6.9: Variation de l'intensité des pics de la cystéine C116 en fonction du rapport IBMP :ASP2. Sont aussi représentées les intensités des pics de C21, L61 et A80.

n'étaient pas observables lors de la titration en pH, les résidus montrant une évolution avec l'IBMP sont également impliqués dans l'équilibre entre la forme A et B. A l'exception de quelques résidus (V6, A8, L14, C96), les déplacements chimiques en fin de titration sont compatibles avec la forme A. On peut remarquer que les résidus qui, en fin de titration, ont un déplacement chimique significativement différent par rapport à la forme A, sont proches du coeur hydrophobe, il n'est donc pas étonnant que leur environnement chimique dans la forme A libre et dans la forme liée soit fortement différents. Il est donc séduisant d'envisager que l'IBMP, comme le TSP, induit la sélection de la forme A (liée à l'IBMP) au détriment de la forme B. Le mécanisme de fixation de l'IBMP suivrait alors le même schéma que pour le TSP.



Comme on n'observe plus qu'un seul jeu de résonances en fin de titration, celui-ci correspond nécessairement à la forme A liée à l'IBMP. A aucun moment, il n'est observé la forme A libre.

La cystéine 116 est utile car elle permet de contrôler simplement dans le spectre HSQC le déplacement chimique d'ASP2. Si pour la plupart des résidus, la différence de déplacement chimique en début et fin de titration est trop faible pour discriminer entre une vitesse rapide ou lente pour la réaction globale, le déplacement chimique de C116 entre les deux formes B et A liée est suffisamment différent pour observer clairement la baisse de l'intensité de la forme B (libre) et l'augmentation en intensité de la forme A liée (figure 6.8). L'équilibre $B + nIBMP \xrightleftharpoons{} A - IBMP_n$ observée est donc lent.

La vitesse d'échange de la réaction 6.2 est lente vis-à-vis des déplacements chimiques. Ce mécanisme est à mettre en lumière avec l'évolution des spectres. En effet, lorsque l'IBMP devient en excès, les spectres HSQC et Watergate 1d restent inchangés à l'exception des pics P1 et P2. Par exemple, l'intensité du pic de C116 reste constante lorsque le rapport 1 :1 est atteint (figure 6.9). Cette observation devrait impliquer que la vitesse d'échange de l'équilibre 6.3 est lente et qu'à 1 :1, la protéine est saturée en ligand. Ces propriétés devraient donc se retrouver sur l'IBMP. Aucun des pics P1 et P2 ne satisfait à ces observations car en échange rapide. Cependant, les spectres NOESY contiennent des pics de corrélation entre ces signaux et ceux de la poche hydrophobe d'ASP2. P1 et P2 sont donc liés à une complexation d'ASP2 avec l'IBMP. Nous devrions alors observer un déplacement progressif des résonances de la forme A libre vers la forme A liée. Ce qui n'est manifestement pas le cas.

Nous devons donc pouvoir expliquer pourquoi l'évolution de l'IBMP traduit un échange rapide avec la protéine alors que les résonances de la protéine (amides et aromatiques) traduisent un échange apparemment lent. Dans le cadre d'une réaction simple, en échange rapide, du type $P + L \xrightleftharpoons{} P - L$ et titrée par ajout progressif de ligand L, la concentration totale en protéine P est constante. La population de la protéine libre est donc de 1 au départ, puis diminue jusqu'à devenir nulle à saturation. La population de la protéine liée subit l'évolution inverse. Nous ne sommes pas tout à fait dans la même situation car, ici, la concentration totale en protéine

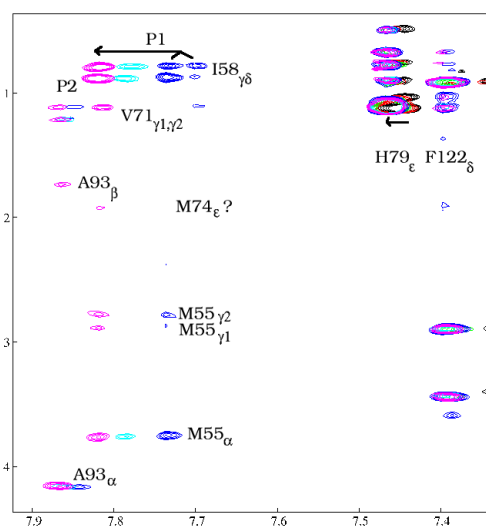


Fig. 6.10: Superposition de la région aromatique-aliphatique des 2D-NOESY filtrées ¹⁵N sans IBMP (noir) et en présence d'IBMP (0,33 :1 en rouge, 0,66 :1 en vert, 1 :1 en bleu, 1,33 :1 en cyan, 1,66 :1 en magenta)

capable de lier le ligand (i.e. A) n'est pas constante. L'équilibre entre A et B est très favorable à la forme B. Par conséquent, la forme A libre est toujours très minoritaire par rapport à la quantité totale de protéine en solution. La population de A est donc toujours très faible. Il est donc probable que dès le premier ajout de ligand, la forme A liée soit très majoritaire par rapport à la forme A libre. Il est donc cohérent de n'observer que les déplacements chimiques de la forme A liée.

6.3.3 Un modèle d'interaction à deux sites

Compte tenu de l'existence des deux signaux de l'IBMP (P1 et P2), je propose l'existence de deux sites de fixation de l'IBMP sur ASP2 en échange lent entre eux. La présence de deux sites est cohérente avec les signaux P1 et P2. En effet, la série de spectres Watergate suggère que le rapport d'intensité entre P1 et P2 est constant (environ 4). L'équilibre de complexation peut alors s'écrire :



S_1 et S_2 sont les deux sites de fixation proposés et K_{a1} et K_{a2} les affinités de l'IBMP pour les deux sites. Du fait du rapport de population entre les deux sites, les affinités ne sont pas identiques. Soit S_1 le site de plus forte affinité, correspondant aux déplacements chimiques du pic P1 le plus intense et S_2 le site de plus faible affinité (P2). On peut estimer le rapport d'intensité des deux pics à environ 4. Par conséquent, si la stoechiométrie des deux sites est de 1 :1, le site S_1 est quatre fois plus affin que le site S_2 . Compte tenu des volumes de la molécule d'IBMP et de la poche hydrophobe, il est difficilement envisageable que deux molécules d'IBMP soient présentes en même temps dans la poche.

Les spectres NOESY filtrés apportent des éléments très intéressants pour la détermination de ces sites sur la structure (figure 6.10). Il faut signaler que les autres protons de l'IBMP ont des déplacements chimiques qui les situent dans la région spectrale des protons aliphatiques. Ceux-ci ne peuvent donc être exploités simplement. Analysons les corrélations impliquant les

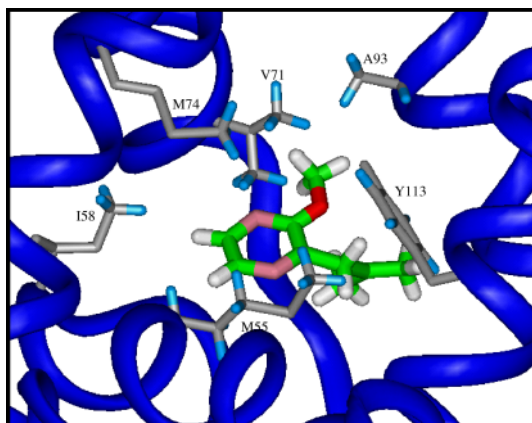


Fig. 6.11: Vue du site S_1 de fixation de l'IBMP sur ASP2 à partir de l'ouverture. Seuls les hydrogènes donnant des corrélations nOes avec l'IBMP sont affichés (en bleu).

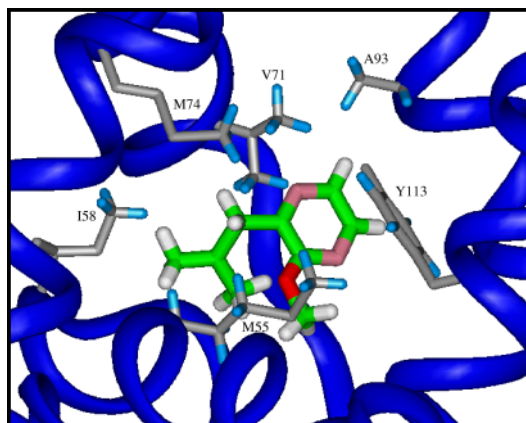


Fig. 6.12: Vue du site S_2 de fixation de l'IBMP sur ASP2 à partir de l'ouverture. Seuls les hydrogènes donnant des corrélations nOes avec l'IBMP sont affichés (en bleu).

aromatiques de l'IBMP. Une observation rapide du spectre montre que les aromatiques placés dans les deux sites ont des corrélations différentes avec la protéine.

L'attribution des nOes que donne le pic P_2 est simple du fait de la présence d'une corrélation nOe vers les protons H_c de la tyrosine Y113. En regardant les déplacements des protons proches de Y113 dans la structure, j'ai attribué les quatre nOes. Il s'agit des protons H_α et H_β de l'alanine A93, les méthyles de la valine V71 ainsi que donc, les protons H_c de la tyrosine Y113. Cela permet d'orienter la molécule au sein de la poche hydrophobe (figure 6.11), à une rotation près autour du cycle aromatique de l'IBMP.

Les sept corrélations reliant l'IBMP du site 1 avec la protéine sont plus délicates à attribuer, car aucune n'est reliée à des protons aromatiques de la protéine. Plusieurs attributions sont possibles pour chaque résonance, mais une seule solution est cohérente par rapport à la structure, pour l'ensemble des sept corrélations : les protons aromatiques de l'IBMP placé dans le site 1 sont en proximité spatiale avec les protons $H_{\alpha,\gamma_{1,2}}$ de la méthionine M55, le méthyle H_{δ_1} de l'isoleucine I58, le groupe S-méthyle de la méthionine M74 et les méthyles de la valine V71. Je propose un modèle structural sur la figure 6.12. Ce modèle respecte les distances observées.

Comme je ne peux observer les corrélations des groupes méthoxy et isobutyle vers la protéine, l'orientation de ces groupes reste ambiguë, en particulier, la molécule peut subir une rotation par rapport au cycle pyrazine tout en satisfaisant les nOes.

Il est donc remarquable que la poche hydrophobe d'ASP2 puisse recevoir l'IBMP dans au moins deux orientations possibles. Ces deux orientations sont tête-bêche. Ceci explique alors parfaitement pourquoi la vitesse d'échange entre les deux sites est lente et donc, pourquoi nous observons séparément les deux résonances P1 et P2. En effet, durant l'édition du déplacement chimique, une molécule placée dans le site 1 n'a pas le temps de sortir de la poche, et d'y entrer à nouveau mais dans l'orientation du site 2.

Nous sommes encore en face d'une contradiction apparente : l'existence de deux sites en échange lent devrait se traduire sur la protéine par l'existence de deux formes liées de A en échange également lent. Or je n'observe qu'une seule forme. J'attribue ceci à la différence

de population entre les protéines dont le site 1 est occupé et les protéines dont le site 2 est occupé. Le rapport des populations étant d'environ 4, il est suffisant pour n'observer que la forme d'ASP2 dont l'IBMP est placée au site fort.

6.3.4 Conclusion de l'expérience

A l'instar de la titration en TSP, cette expérience a montré que la fixation d'IBMP sur ASP2 induit une stabilisation de la forme A au détriment de la forme B. Il existe deux sites d'interaction et les nOes intermoléculaires prouvent que l'IBMP s'oriente de deux façons dans la poche hydrophobe. Les régimes d'échange sont particulièrement complexes compte tenu de l'existence de plusieurs équilibres. Cependant, les contradictions ont pu être levées par le biais de mécanismes et de populations respectives. En revanche, devant la complexité des mécanismes en jeu, nous ne pouvons pas calculer d'ordre de grandeur de la constante de dissociation des deux sites de fixation. La vitesse d'échange de la fixation d'une molécule sur les deux sites étant rapide, nous pouvons tout de même considérer que la constante de dissociation est de l'ordre du millimolaire.

6.4 Une expérience : “IBMP contre le complexe ASP2-¹⁵N/TSP”

6.4.1 Objectif et protocole expérimental

Les études de microcalorimétrie ont montré que l'IBMP était très affin envers ASP2 (K_d de l'ordre du micromolaire). Afin de comparer l'affinité de l'IBMP et du TSP, j'ai mené une expérience de compétition. L'échantillon initial contient le complexe ASP2-TSP et je le titre par une solution d'IBMP. Nous nous attendions à une évolution différente de cette série de spectres par rapport à celle des deux expériences de titration directe de la protéine libre. En effet, la forme A de la protéine étant déjà majoritaire (liée au TSP), nous ne devons plus observer d'interconversion en échange lent, du type B vers A.

Le tube utilisé est celui qui a servi à la détermination des nOes des groupes aromatiques du complexe ASP2-TSP (tube 4). Je rappelle qu'il est dans D₂O et que l'excès en TSP est de 32 :1. L'ajout de IBMP s'est fait à partir d'une solution de IBMP pur et d'éthanol (10 :1 v/v). J'ai ajouté progressivement environ 1 μ l de cette solution dans le tube. La concentration totale d'IBMP à un instant donné est difficile à estimer car j'ai laissé la solution diffuser dans le tube (sans agitation). Je l'ai estimée à partir des rapports d'intensité entre les pics de l'IBMP et de l'histidine 79. La titration a été suivie par une série de Watergate (non filtrée) et ponctuellement par des NOESYs.

6.4.2 Analyse des spectres

La principale modification des spectres Watergate (figure 6.13) est l'apparition de deux pics à environ 7,8ppm. Ces deux pics, absents avant ajout d'IBMP, non seulement augmentent en intensité (en dépassant largement l'intensité des pics de la protéine) mais ils se déplacent au fur et à mesure que l'IBMP est ajouté. Il s'agit donc de signaux de l'IBMP. D'autre part,

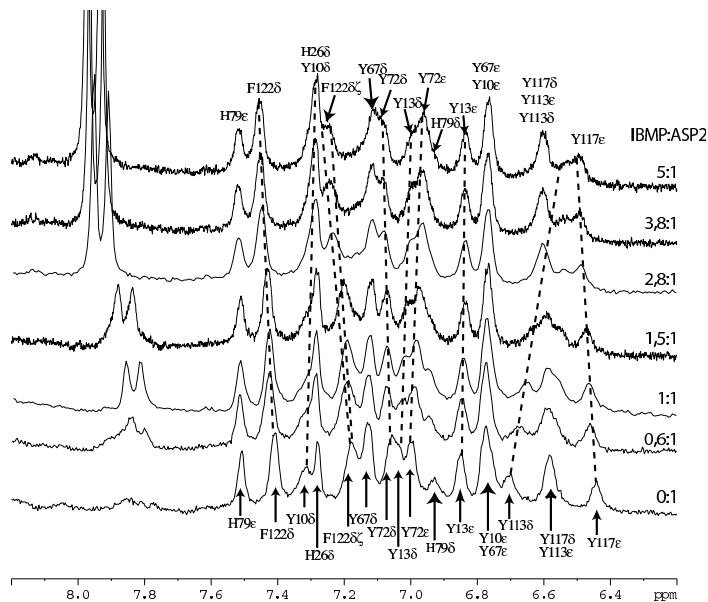


Fig. 6.13: Titration du complexe ASP2-TSP par l'IBMP suivie par des Watergate 1D.

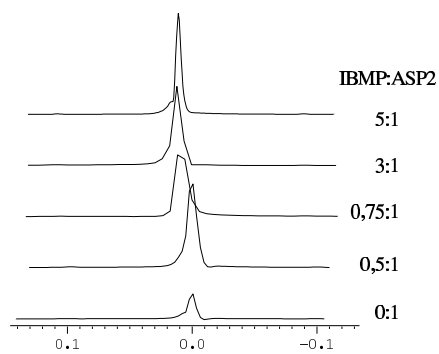


Fig. 6.14: Effet de la présence d'IBMP sur le déplacement chimique du TSP.

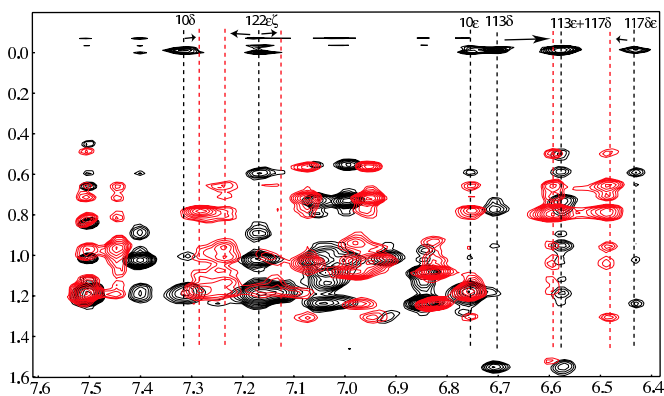


Fig. 6.15: Superposition des spectres NOESYs sans IBMP (en noir) et avec IBMP (5 :1, en rouge). Les nOes entre le TSP et ASP2 disparaissent.

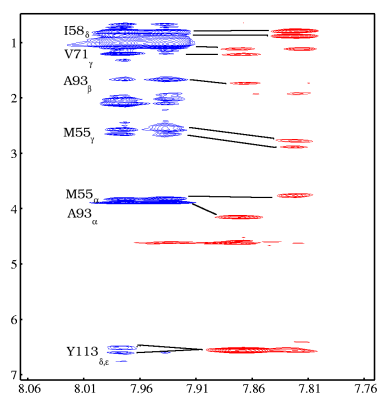


Fig. 6.16: Superposition des spectres NOESY menés sur le complexe ASP2-IBMP (en rouge, excès 1,66 :1 en IBMP) et lors de l'expérience de compétition (en bleu, excès de 5 :1 en IBMP).

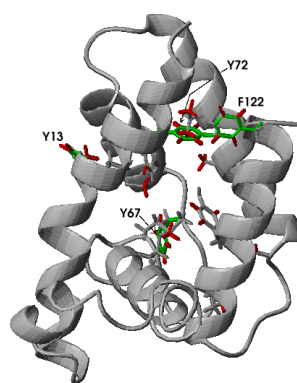


Fig. 6.17: Représentation sur la structure d'ASP2 des protons (en rouge) dont le déplacement chimique est affecté par la titration en IBMP.

une partie des résonances des aromatiques de la protéine se déplace avec l'ajout d'IBMP. Le proton H_c de la tyrosine 113 est celui dont le déplacement chimique est le plus affecté (0,20ppm). Les résonances des protons aromatiques de Y10, Y13, Y67, Y72, Y117 et F122 ont également un déplacement significatif. De plus, le signal du TSP est également déplacé. Par conséquent, le complexe ASP2-TSP et l'IBMP interagissent. Analysons maintenant plus en détail, les caractéristiques de l'interaction.

L'IBMP chasse le TSP de la poche hydrophobe Comme nous l'avions pensé, le TSP est moins affiné pour ASP2 que l'IBMP. En effet, la comparaison des spectres NOESY du complexe ASP2-TSP et en fin de titration (figure 6.15) est concluante : les nOes qui existaient entre le TSP et les aromatiques de la protéine disparaissent. D'autre part, le déplacement de la raie du TSP sur les spectres 6.14 est une preuve supplémentaire au fait que le TSP soit chassé de la poche hydrophobe. En revanche, le spectre NOESY du complexe ASP2-TSP contient des taches de corrélations TSP-protons aliphatiques qui peuvent être attribués à des nOes intermoléculaires mais la qualité du spectre dans cette région pose des réserves à toutes conclusions. Ces taches restent présentes en fin de titration avec l'IBMP.

Le TSP augmente la vitesse d'échange de l'IBMP entre les deux sites Contrairement à l'expérience précédente, l'IBMP ne produit ici que deux pics, de même intensité. La titration de ces pics est très semblable à celle du pic majoritaire de l'expérience précédente "IBMP contre ASP2", ceci suggère que le pic minoritaire est absent. Cependant, l'écart en déplacement chimique des deux pics est supérieur ici (26Hz) que pour le pic P1 de l'expérience de titration d'ASP2 libre par l'IBMP seul (21Hz). Par conséquent, le signal que nous observons à 7,8ppm ne peut pas être simplement celui d'un équilibre rapide entre une molécule d'IBMP libre en solution et la même dans le site fort d'ASP2. La présence de TSP modifie donc l'interaction entre l'IBMP et ASP2. Les expériences NOESY (figure 6.16) montrent de plus, que les nOes impliquant le pic à 7,8ppm de la présente expérience différent de ceux observés sur le

Résonance	δ (ppm)	Résonance	δ (ppm)	Résonance	δ (ppm)	Résonance	δ (ppm)
Y113 δ	0,20	Y113 α	0,10	M60 α	0,05	Y67 ϵ	0,03
M55 ϵ	0,16	Y13 α	0,07	A100 β	0,05	Y10 α	0,03
F122 ϵ	0,13	Y113 δ	0,07	I109 γ 12	0,05	Y10 δ	0,03
F122 α	0,13	L121 δ	0,06	V89 γ	0,04	Y67 δ	0,02
Y13 β	0,13	V6 γ	0,06	Y72 ϵ	0,04	Y72 δ	0,02
F122 δ	0,12	Y13 δ	0,05	A93 β	0,04	Y13 ϵ	0,02
						M74 α	0,02

Tab. 6.2: Tableau des variations de déplacement chimique après ajout d'IBMP (excès 5 :1).

pic majoritaire de l'expérience précédente : il y en a davantage et ils ne peuvent être attribués à une seule molécule dans la cavité. Il est remarquable que le jeu de nOes donné par le pic unique de notre expérience rassemble à la fois les nOes correspondant au site minoritaire et ceux du site majoritaire. Tout se passe donc comme si les deux sites de fixation S_1 et S_2 étaient toujours présents et que le TSP augmentait pour, une molécule d'IBMP, la vitesse de passage d'un site à l'autre. Ceci expliquerait aussi pourquoi nous n'observons qu'un seul signal pour l'IBMP et que la forme des raies est modifiée. Cependant, on ne peut pas exclure l'hypothèse où les deux sites aient fortuitement le même déplacement chimique, variation induite par la présence de TSP.

Les nOes observés suggèrent que certaines résonances de la protéine ont un déplacement chimique très différent dans les deux types d'expériences (compétition versus titration directe). C'est particulièrement vrai pour le proton H_α de l'alanine 93 (de 4,25 à 3,9ppm). En réalité, l'excès en IBMP n'est pas le même dans les deux spectres. Lors de l'expérience "IBMP contre ASP2", nous sommes allés jusqu'à un excès de 1,66 :1 en IBMP tandis que dans l'expérience de compétition, l'excès final est d'environ 5 :1. L'attribution de l'alanine est non ambiguë car contrôlée par le suivi de la tache de corrélation entre les protons aromatiques de la tyrosine 117 et les protons de l'alanine. D'autre part, il n'est pas étonnant que les protons de cette alanine voient leurs déplacements chimiques être très déplacés car la proximité du groupe pyrazine de l'IBMP induit des courants de cycle importants.

Carte des sites affectés par la présence d'IBMP Les cartes NOESY disponibles sur cette titration permettent de contrôler l'effet de la présence d'IBMP dans le tube sur les protons d'ASP2 proches des groupes aromatiques de la protéine. En particulier, il est intéressant de comparer l'évolution des déplacements chimiques de protons éloignés du site de fixation et ceux situés dans la poche hydrophobe, qui donnent des nOes avec les tyrosines 10, 113 et 117. Les protons qui permettent de sonder le comportement des protons éloignés du coeur hydrophobe sont contrôlés par les tyrosines 13, 67, 72, l'histidine 79 et la phénylalanine 122. Le tableau 6.17 rassemble les variations les plus importantes que j'ai pu détecter en proximité des aromatiques. Les noyaux correspondants sont représentés sur la structure d'ASP2 (figure 6.17).

Comme on pouvait s'y attendre, la majorité des plus grandes variations sont observées sur les protons de la poche hydrophobe. En revanche, je ne m'attendais pas à ce que les protons des

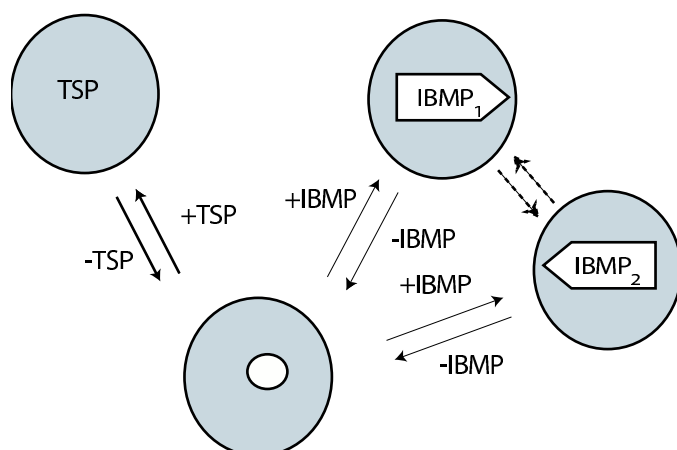
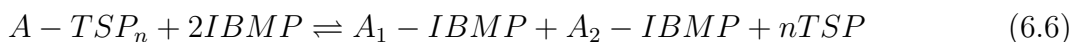


Fig. 6.18: Mécanisme de compétition entre la fixation du TSP et de l’IBMP sur ASP2. La réaction en trait pointillé est indirecte.

cycles de Y13, Y67, Y72 et F122 soient affectés, car ils sont éloignés du coeur de la protéine. Or c’est manifestement le cas. On peut donc supposer que l’introduction d’IBMP dans la poche hydrophobe induit une déformation de la poche qui se transmet sur l’ensemble de la structure.

6.4.3 Discussion

L’expérience de compétition entre le TSP et l’IBMP pour les sites de fixation de la protéine nous renseigne un peu plus sur les mécanismes d’interaction d’ASP2. Nous avons conclu de cette expérience que les deux sites de fixation S_1 et S_2 de l’IBMP sont conservés. La réaction globale que nous avons titrée peut être alors écrite de la façon suivante, la suivante :



Il faut préciser que l’état $A - TSP$ n’a pu être déterminé en termes de sites de fixation pour le TSP. Cette équilibre est donc plus une vision qualitative qu’une réelle réaction. De plus, des réactions intermédiaires interviennent qui correspondent à la libération du TSP et à l’entrée de l’IBMP dans la poche hydrophobe. Le mécanisme probable est représenté sur la figure 6.18. La première étape de libération du TSP est simplifiée. Ce mécanisme fait intervenir la forme libre de la protéine. Compte tenu de l’excès en TSP, celle-ci est toujours très minoritaire.

Nous avons observé un changement de vitesse dans l’équilibre de fixation de l’IBMP par la forme A libre (intermédiaire) ainsi qu’un effet notable du remplacement du TSP par l’IBMP sur des noyaux éloignés de la poche hydrophobe. Il est donc probable que la protéine ne fixe pas ces deux molécules de la même façon et qu’un réarrangement global de la protéine est nécessaire pour permettre à la protéine de s’adapter au ligand. Compte tenu des tailles relatives de ces deux molécules, ceci n’est pas totalement surprenant. On peut proposer que les chaînes latérales de la poche hydrophobe s’écartent légèrement pour augmenter le volume de la poche afin de pouvoir recevoir l’IBMP. Ceci a un effet global sur la structure.

Il faut également expliquer le changement de vitesse de passage de la molécule d’IBMP d’un site à l’autre. La différence principale entre la situation où on ajoute l’IBMP sur la protéine libre et celle où le TSP est déjà présent dans la poche, est que la forme A est déjà majoritaire dès le début, sous sa forme liée au TSP. De plus, l’équilibre entre A et B est ici négligeable car l’excès en TSP fait que toute la protéine est sous la forme A liée. On peut donc supposer que la sélection au préalable de la forme A augmente la vitesse globale de sortie

d'une molécule d'IBMP d'un site et d'entrée dans l'autre site. Cependant, on ne peut pas exclure l'hypothèse dans laquelle le TSP, en dehors de la poche, induit des modifications dans la poche, qui conduiraient à la possibilité pour l'IBMP de changer de site, tout en restant dans la poche.

Ce mécanisme à nouveau est très complexe et il faudrait le confronter à d'autres titrations, telle l'expérience de titration inversée.

6.5 Une expérience "IBMP contre ASP2 ^{15}N - ^{13}C "

6.5.1 Objectif et protocole expérimental

L'expérience "IBMP contre ASP2" a montré ses limites dans la détection des corrélations spatiales entre les protons autres qu'aromatiques de l'IBMP et ceux de la protéine. Ces nOes sont importants car ils permettraient de préciser l'orientation de l'IBMP dans les deux sites. J'ai donc été encouragé à étudier les complexes ASP2-IBMP en utilisant une protéine doublement marquée. Ainsi, je pouvais utiliser des expériences filtrées isotopiquement qui devaient me permettre d'observer ces corrélations.

A la suite de la tentative d'observation du couplage scalaire à travers les liaisons hydrogènes, la protéine marquée ^{13}C - ^{15}N montrait des signes de coupure : des signaux fins et intenses étaient apparus sur le spectre ^{15}N -HSQC. Loïc Briand effectua un séquençage N-terminal qui confirma que deux séquences N-ter étaient présentes en solution, toutes deux correspondant à des segments d'ASP2 : la séquence N-ter du peptide natif et à partir de N40. La solution contenant environ 10mg de protéine en complexe avec le TSP fut donc repurifiée par colonne HPLC en phase inverse. La quantité finale de protéine doublement marquée obtenue après dialyse fut décevante : 0,9mg. Cependant, en dissolvant la poudre dans 400 μ l d'eau (concentration en protéine de 0,3mM, environ 10% de D₂O, pH entre 5,5 et 6) et en utilisant le spectromètre 600MHz disposant de la cryosonde, le rapport signal/bruit des expériences était suffisant pour mener une titration (spectres 1d et 2d, filtrés ou non).

L'IBMP a été ajouté à partir de la même solution que dans l'expérience "IBMP contre ASP2". Cependant, elle a été diluée trois fois pour tenir compte des rapports de concentration en protéine entre les deux échantillons. Pour suivre la titration, j'ai utilisé un spectre 1d-Watergate filtré ^{15}N , des spectres NOESY filtrés ^{15}N ou ^{13}C en dimension F2 et l'expérience filtrée ^{13}C - ^{15}N dans les deux dimensions pour observer les nOes entre molécules non marquées.

6.5.2 Analyse des spectres.

Le spectre ^{15}N -HSQC de la protéine fraîchement dissoute ne fait pas apparaître les doublements ni les signaux issus des acides aminés coupés. Le spectre 1d-Watergate filtré ^{15}N indique la présence de TSP dans un rapport environ de 1 :6 par rapport à la protéine, ainsi que la présence d'un pic très intense à 1,9ppm que j'ai attribué au méthyle d'acétonitrile (CH_3CN) résiduel.

Les ajouts successifs d'IBMP ont fait apparaître dans le spectre Watergate filtré un pic à environ 8ppm, correspondant aux protons aromatiques de l'IBMP. Et, chose étonnante, ce pic ne s'est pas déplacé lorsque la concentration en IBMP a augmenté. D'autre part, les protons

aromatiques de la protéine n'ont pas subi de changement d'environnement chimique. Tout se passe donc comme si la protéine ne fixe pas l'IBMP.

La présence de compétiteurs potentiels (TSP et acétonitrile) m'a poussé à essayer de déterminer des contacts entre ceux-ci et la protéine. Pour cela, j'ai effectué un spectre NOESY filtré ^{15}N - ^{13}C dans la dimension d'acquisition et découplée ^{15}N - ^{13}C dans les deux dimensions. Je n'ai observé aucun contact intermoléculaire.

6.5.3 Discussion

En premier lieu, la purification n'a pas permis d'éliminer complètement le TSP. D'autre part, et pour la première fois, la lyophilisation n'a pas éliminé l'acétonitrile. Cet échantillon contient donc des contaminants qui rendaient difficile l'expérience de titration avec l'IBMP. En effet, l'IBMP ne s'est pas fixé sur ASP2, même à forte concentration. Je doute fortement que le TSP puisse expliquer ce phénomène car sa concentration est très faible en comparaison avec celle de l'expérience de compétition. En revanche, l'homogénéité d'ASP2 dans la carte HSQC suggère qu'elle n'est que sous une seule forme. Il est possible que l'histoire de la protéine influe sur des propriétés de fixation. Si la protéine n'est capable de fixer un ligand, qu'une fois au cours de son "existence", ceci expliquerait pourquoi la protéine est présente en si grande concentration.

On peut aussi imaginer l'hypothèse que l'acétonitrile interagit avec la protéine. Compte tenu de sa forte concentration (excès d'environ 10 :1 vis-à-vis de la protéine), même si la constante de dissociation est très faible vis-à-vis de celle de l'IBMP, la loi d'action de masse complique les réactions dans le tube.

6.6 Discussion finale sur les interactions

6.6.1 Comment interpréter l'inversion de population entre A et B ?

Les expériences de titration de la protéine libre par le TSP et l'IBMP ont mis en évidence un phénomène très important : la présence de ligand induit un changement non négligeable dans la structure d'ASP2. La forme qui est majoritaire en absence de ligand devient minoritaire. Il semble donc que le ligand mime le changement de populations entre les formes A et B dû au pH. Cette inversion de population en fonction du pH a été interprétée dans le chapitre 4 par la création d'un pont salin impliquant l'histidine H79. S'en était suivi, un changement de conformation du squelette d'ASP2 pour aboutir à la forme A. Le changement de conformation a été attribué, entre autres, à une isomérisation du pont disulfure C96-C116. Il semble donc que la présence de ligand induise le même type de changement de conformation mais par un mécanisme totalement différent. L'expérience de compétition a montré que la poche hydrophobe de la forme A liée devait subir une modification de son volume pour s'accommoder des volumes du ligand. La littérature fait état de plusieurs possibilités de comportement du volume d'une poche hydrophobe de protéines impliquées dans le transport de petites molécules hydrophobes : pour certains systèmes, la poche est inexistante en absence de ligand (Charvolin et al. [1999], Han et al. [2001], Lerche et al. [1997]), tandis que pour d'autres systèmes, la poche préexiste (Curry et al. [1998], Shin et al. [1995]). Si nous ne pouvons conclure de nos données

que la poche est absente en absence de ligand, il semble que celle-ci subit des modifications importantes. Il est possible alors de penser que la poche hydrophobe dans la forme B est de volume très faible, voire nulle tandis que la forme A contient une poche de volume nécessaire à l'entrée d'une petite molécule. La poche et par voie de conséquence, l'ensemble de la structure subit ensuite de légères déformations en fonction du ligand présent.

6.6.2 RMN versus microcalorimétrie

Les résultats des études d'interaction sont susceptibles de modifier profondément les conclusions des expériences de microcalorimétrie menées sur les différents ligands. En effet, le dégagement d'énergie a été interprété dans le cadre d'un système simple où le ligand se fixe sur plusieurs sites sur la protéine sans considérer le coût énergétique d'un changement de conformation de la protéine. Or, les informations collectées en RMN suggèrent que ce dernier ne peut être négligé. En effet, les deux formes de la protéine sont en échange lent vis-à-vis des déplacements chimiques. La barrière d'énergie de passage de la forme B à A est donc importante. Si cela ne nous dit pas que la différence d'énergie entre les formes est importante, il faut tout de même considérer cette hypothèse. L'équipe de l'INRA a commencé à analyser les courbes de titration en fonction de cette information. Il est donc possible que les valeurs de constantes de dissociation soient erronées.

En revanche, nos données confirment l'existence de deux sites forts, de fixation d'affinité comparable. Ces deux sites avaient été interprétés à la lumière du dimère d'ASP2. Nous pouvons réfuter cette conclusion car les différentes expériences de relaxation ont montré que ASP2 était sous une forme monomérique. Nous corrigeons donc les conclusions issues de la microcalorimétrie : il existe bien deux sites de fixation, mais ils sont situés au sein d'un même monomère. La vitesse d'échange entrée-sortie de l'IBMP dans la poche nous laisse supposer que la constante de dissociation du complexe est bien plus faible que celle déterminée par microcalorimétrie. Les résultats de la nouvelle analyse des données de microcalorimétrie nous confirmera peut-être cela.

L'ajustement des courbes de microcalorimétrie a nécessité l'introduction d'un troisième site d'affinité plus faible (mM). Les données de RMN ne permettent pas de conclure définitivement sur son existence. A mon avis, l'introduction du paramètre d'échange B-A dans les ajustements sera suffisant pour éliminer ce site.

6.6.3 Les sites de fixation de l'IBMP

Nous avons montré l'existence de deux sites de fixation de l'IBMP au sein de la poche hydrophobe, dont les affinités sont de même ordre de grandeur (un rapport de 4). Ces deux sites diffèrent par l'orientation de la molécule dans la poche. Les deux orientations sont tête-bêche. Il est donc étonnant que les chaînes latérales formant la poche puissent créer des liaisons stables avec des groupes aussi différents que des protons aromatiques de pyrazine, les protons de méthoxy et d'isobutyle. Le caractère commun à ces groupes sont leur hydrophobicité. Il semble donc que les interactions d'ASP2 avec l'IBMP et sans doute ses autres ligands sont gouvernées en premier lieu par l'hydrophobicité. Ceci confirme les expériences de fixation menées par l'équipe de l'INRA. Les trois molécules autres que l'IBMP testées (1,8-cinéol, isoamylacétate et 2-heptanone) ont une hydrophobicité importante. Les quatre molécules étudiées possèdent

également un groupe polaire (carbonyle ou méthoxy). La poche dégagée dans la structure d'ASP2 ne montre que très peu de liaisons polaires. Il s'agit des hydroxyles des tyrosines Y10, Y113, et Y117 et de la thréonine 114. Ces liaisons sont donc susceptibles de stabiliser le complexe.

6.6.4 Un mécanisme d'entrée du ligand

La structure du complexe ASP2-TSP a montré que la poche hydrophobe disposait d'une ouverture. Celle-ci est présente, car dans la majorité de nos structures, le segment 30-45 est éloigné de la poche. Cependant, les expériences de titration en TSP et en IBMP ont montré que les résonances de respectivement T33, et de N34 et A37 étaient affectées par le ligand. Ce segment a donc un rôle à joué dans l'interaction. Bien que nos données soient très limitées, il est possible d'imaginer que l'ouverture de la poche soit gouvernée par l'orientation du segment. Dans une structure fermée, la poche est fermée tandis que l'ouverture apparaît lorsque ce segment a une conformation plus ouverte. Il est alors envisageable que le ligand entre dans la poche par cette ouverture.

Chapitre 7

Conclusion générale et perspectives

7.1 Rappel des observations et des modèles proposés issus de l'étude d'ASP2 par RMN

⇔ L'apoprotéine

- ASP2 libre en solution adopte deux formes en échange lent.
- Les deux structures en équilibre sont très similaires : les deux formes sont monomériques et leurs squelettes présentent une conformation légèrement différente. L'état d'isomérisation du pont disulfure $C_{96} - C_{116}$ est peut-être la principale différence structurale entre les deux formes.
- Les populations relatives des deux formes sont dépendantes du pH. Le pH de transition est compris entre 6 et 7. La formation à bas pH d'un pont salin impliquant l'histidine 79 est sans doute à l'origine de cette transition. La dépendance en pH n'a été observée que sur un seul échantillon. Dans les autres, la forme majoritaire à haut pH reste majoritaire à pH 6.

⇔ La structure d'ASP2

- la structure du complexe ASP2/TSP a été résolue : elle est constituée de 6 hélices formant un coeur essentiellement hydrophobe dans lequel est fixé le TSP.
- le segment D30 à S45 présente une dynamique originale qui explique la mauvaise définition de la structure dans cette région : le segment D30 à A38 subit le phénomène d'échange conformationnel dans la gamme de la milliseconde tandis que le squelette de A37 à S45 subit une dynamique interne très importante dans la gamme piconoseconde.
- Le complexe ASP2/TSP est monomérique.
- ASP2 ne subit pas de modification importante de la structure lors de la fixation de ligand.

⇔ Les interactions ASP2-ligand

- Le TSP et l'IBMP se lient à ASP2 et sélectionnent la forme minoritaire à pH basique.
- La stoechiométrie de fixation de ligand est supérieure à 1 :1 et l'équilibre de fixation du ligand est rapide.
- L'IBMP se fixe de deux façons possibles, orientées tête-bêche dans la poche hydrophobe.

➤ La cavité s'adapte au volume du ligand.

Discussion sur l'état oligomérique d'ASP2 A partir d'expériences de gel filtration, Loïc Briand et ses collaborateurs concluent qu'ASP2 libre en solution est un dimère en solution à pH 7 et un mélange monomère-dimère en échange lent (par rapport à la durée de l'élution) à pH 3 ou en présence d'un agent chaotropique à une concentration de l'ordre de $40\mu\text{M}$. Dans nos conditions (environ $0,7\text{mM}$), l'équilibre monomère-dimère devrait donc être déplacé vers la forme dimérique.

Toute l'étude par RMN du complexe est consistante avec un état monomérique (largeur de raie, temps de corrélation, tenseur de diffusion orientationnelle faiblement anisotrope, reconstruction de la structure). Les données dont je dispose sur ASP2 sans ligand (voir chapitre 5) suggèrent que la protéine libre est également monomérique.

Les différentes études menées jusqu'à présent sur l'état oligomérique sur les OBPs sont ponctuelles et parfois contradictoires. Ainsi, la PBP de *Bombyx mori* (paragraphe 2.1.4.1) est dimérique à pH physiologique en gel filtration et dans la structure cristallographique tandis que, bien que la publication de Lee ne discute pas de ce point, elle est monomérique dans la structure RMN. Nous sommes donc confrontés à un problème similaire dans les deux systèmes.

Discussion sur les constantes d'affinité Je rappelle ici également que l'étude par RMN n'a pas permis la détermination précise des constantes d'affinité des ligands TSP et IBMP. En revanche, elle a mis en évidence la nécessité de revoir l'interprétation des courbes de microcalorimétrie car l'existence de l'équilibre conformationnel de la protéine libre et des enthalpies associées est à prendre en compte. Les constantes de dissociation déduites (de l'ordre du micromolaire pour les quatre ligands étudiés) sont donc susceptibles d'être modifiées.

7.2 Des données structurales des OBPs vers leur(s) fonction(s)

Il est temps à présent de mettre en lumière nos résultats avec les systèmes d'OBPs déjà connues. La comparaison des trois OBPs très proches structuralement d'ASP2 permet d'extraire des informations sur leurs rôles.

7.2.1 Comparaison des structures d'ASP2 et d'autres OBPs

Ces derniers mois, deux nouvelles structures de protéines d'invertébrés apparentées aux OBPs ont été publiées : une PBP de la blatte (Lartigue et al. [2003]) et l'OBP LUSH de la drosophile (Kruse et al. [2003]). L'arrangement en hélices α des quatre structures confirme la profonde différence qui existe au niveau structural entre les OBPs de vertébrés (boeuf, porc) et les OBPs d'invertébrés.

Le serveur DALI²⁴ (Holm and Sander [1993]) est dédié à la comparaison de domaines tridimensionnels de protéines. Basé sur l'analyse des matrices de distances entre les carbones α des structures et sur une base de données de repliements, l'algorithme permet d'avoir accès aux

²⁴disponible à l'adresse : <http://www.ebi.ac.uk/dali/>

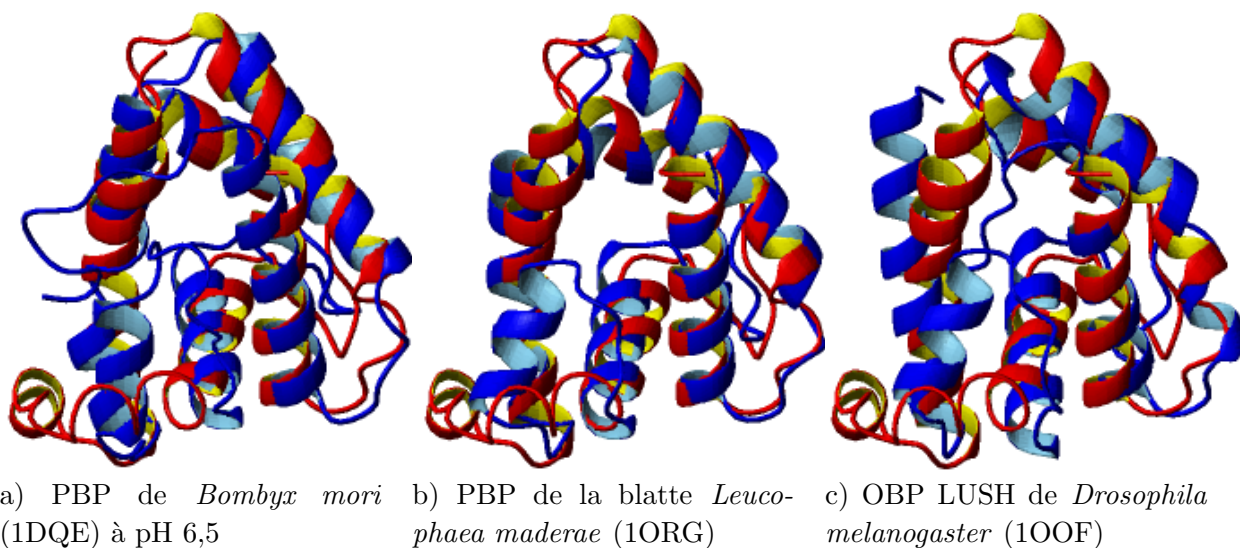


Fig. 7.1: Superposition de la structure moyenne d'ASP2 en complexe avec le TSP (en rouge) avec les structures connues d'OBPs. Les références PDB des structures sont indiquées.

ASP2	2-6 ; 8-27 ; 40-43 ; 44-66 ; 67-79 ; 82-122	6-27 ; 43-62 ; 63-68 ; 69-77 ; 78-82 ; 84-101 ; 102-121	8-27 ; 44-64 ; 65-79 ; 82-122
PBP de <i>Bombyx mori</i>	1-5 ; 6-25 ; 26-29 ; 45-67 ; 69-81 ; 83-123		
PBP de la blatte		1-22 ; 37-56 ; 58-63 ; 65-73 ; 75-79 ; 80-97 ; 99-118	
OBP LUSH			4-23 ; 41-61 ; 63-77 ; 78-118
rmsd ¹	2,6Å	2,1Å	3Å

Tab. 7.1: Alignement deux à deux utilisé lors de la superposition des structures.

¹ : Ecart quadratique moyen lors de la superposition du squelette sur les résidus structurellement équivalents.

protéines de structure proche de celle étudiée. L'application de DALI sur la structure moyenne d'ASP2 a donné les meilleurs scores aux trois OBPs précédemment citées. L'alignement des résidus "structuralement équivalents" entre deux structures proposés par DALI est rassemblé dans le tableau 7.1. J'ai enlevé les alignements impliquant le segment D30-S45 d'ASP2, compte-tenu de sa flexibilité dans la structure. Les superpositions des trois structures d'OBPs et d'ASP2, à partir de l'alignement, sont représentées sur la figure 7.1.

L'écart quadratique moyen entre les structures, compris entre 2 et 3Å, s'explique essentiellement par l'orientation de l'hélice α_1 , du β -*hairpin* et de la boucle α_4 - α_5 . C'est particulièrement visible sur la figure 7.1 où l'hélice α_1 de LUSH s'écarte notablement de la position occupée par cette hélice dans ASP2. La position du β -*hairpin* par rapport à la boucle α_5 - α_6 dans les quatre structures est dispersée : la boucle est bien superposée dans les quatre structures, en revanche les coude des β -*hairpin* de LUSH et d'ASP2 sont spatialement nettement plus proches de cette boucle que dans la structure de la PBP de *Bombyx mori*. La boucle α_4 - α_5 des quatre structures ont des positions relativement dispersées de part et d'autre du plan formé par les hélices α_4 et α_5 . Dans ASP2, cette boucle est située dans ce plan et l'alignement de séquences suggère qu'elle contient une insertion d'un résidu.

La dispersion de positionnement des structures secondaires au travers des quatre structures permet d'expliquer en partie les propriétés de fixation de ces OBPs car la taille de la cavité de fixation dépend bien entendu de l'orientation des hélices. Cependant, l'extrémité C-terminale de la PBP de *Bombyx mori*, qui forme, à bas pH, l'hélice chassant le ligand, n'est présente que dans l'OBP de la drosophile. Le mécanisme de relargage du ligand doit donc être différent en fonction de la longueur de la séquence protéique de l'OBP.

7.2.2 Les propriétés d'interaction des complexes OBPs/ligands

7.2.2.1 Description des interactions OBPs/ligands

PBP de *Bombyx mori* Comme nous l'avons déjà précisé, la structure de la PBP de *Bombyx mori* a été obtenue par RMN à pH physiologique et à pH 4,5 sous la forme apo et par cristallographie sous sa forme liée au bombykol. Le bombykol ((E,Z)-10,12-hexadecadiénol) est en contact avec les chaînes latérales de la protéine : une liaison hydrogène entre l'hydroxyle et la sérine 56 et des interactions d'empilement entre les doubles liaisons conjuguées et les phénylalanines 12 et 118. A l'exception de S9, S56, E98 et de T111, les chaînes latérales formant la poche sont hydrophobes (Met, Phe, Trp, Ile, Val, Leu et Ala).

PBP de la blatte *Leucophaea maderae* L'étude par cristallographie de la PBP de la blatte a permis de déterminer sa structure sous la forme apo mais en présence de glycérol (agent cryoprotectant) dans la poche, et sous la forme liée avec le 3-hydroxy-butan-2-one (H3B2, composé phéromonal hydrophile) et l'ANS (acide 1-anilinonaphtalène-8-sulfonique, rapporteur de fluorescence utilisé pour l'étude de fixation des ligands naturels des OBPs en général). La poche est formée par des résidus hydrophobes mais également par des chaînes latérales polaires (Y5, Y75 et T111). La cavité est ouverte et l'ouverture est formée par 6 résidus (R33, N34, P35, K85, V114 et R115), dont trois sont chargés positivement. Le ligand phéromonal est en interaction avec des chaînes polaires (Y5 et Y75), chargées (K89) et hydrophobes (F110). La carte de densité montre que les deux énantiomères de H3B2 sont présents dans la cavité. Les

auteurs ont également noté une augmentation de 50% de la cavité lorsque l'ANS y est présente par rapport à la forme apo.

OBP de *Drosophila melanogaster* La structure de l'OBP LUSH à pH 4,6 a été déterminée par cristallographie sous sa forme apo, et en complexe avec l'éthanol, le propan-2-ol et le butan-2-ol. Le complexe OBP-butanol a été étudié également à pH 6,5. Il ressort qu'à ces deux pH, les structures sont très similaires alors que la séquence de LUSH contient le même type de séquence C-terminale que la PBP de *Bombyx mori*. Le mécanisme déduit de l'étude structurale de la PBP de *Bombyx mori* n'est donc pas généralisable aux OBPs contenant l'extension C-terminale. De plus, l'OBP conserve sa faculté de fixation à bas pH. Les auteurs notent une différence importante sur les orientations relatives des hélices α_1 et de l'extension C-terminale entre LUSH et la PBP du papillon *Bombyx mori*, ce qui affecte la formation de la cavité. Les cartes de densité électronique contiennent plusieurs molécules d'alcool mais un seul des sites est significatif. Le groupe hydroxyle des alcools forme des liaisons hydrogènes avec les résidus S52 et T57 de l'hélice α_3 . L'alcool n'occupe qu'une partie de la cavité. Les spectres ^{15}N -HSQC des complexes LUSH-éthanol et LUSH-butanol ainsi que de l'échantillon dont l'alcool a été retiré, montrent que la liaison d'alcool stabilise la structure de la protéine (présence d'échange conformationnel sur la plupart des résidus détectée par des résonances larges et peu intenses) et que le degré de stabilisation dépend de la longueur de la chaîne. Par ailleurs, le facteur cristallographique B du complexe avec le butanol montre que la densité électronique est moins bien définie à pH 6,5 qu'à pH 4,5, signe d'une plus grande homogénéité conformationnelle à bas pH.

7.2.2.2 Discussion sur les propriétés de OBPs

Caractéristiques des ligands Contrairement à la PBP de *Bombyx mori* et à l'instar de la PBP de la blatte, la poche formée par les hélices d'ASP2, bien qu'essentiellement hydrophobe, présente des liaisons polaires (Y10, Y113, Y117 et T114) concentrées dans une moitié de la cavité. Dans le cas de la PBP de la blatte, cela se justifie par la forte polarité du composé phéromonal (groupes hydroxyle et cétone). Pour ASP2, les quatre ligands qui ont montré par microcalorimétrie une affinité de l'ordre du micromolaire sont faiblement polaires (liaisons éther et carbonyle). En revanche, la β -ionone, qui est également polaire par la liaison carbonyle ne se fixe pas sur ASP2. La présence des chaînes polaires dans la cavité d'ASP2 suggère donc qu'ASP2 peut lier *a priori* des molécules odorantes peu hydrophobes. Mais l'absence de fixation de la β -ionone indique que la polarité du ligand n'est pas le seul critère. On peut noter que le site de plus forte affinité observé sur l'IBMP oriente le groupe méthoxy de l'IBMP vers la partie plus polaire de la cavité tandis que le site de plus faible affinité l'oriente vers une région très peu polaire. La différence d'affinité entre les deux sites peut donc être expliquée par les interactions de type polaire.

La fixation du TSP est une propriété étonnante d'ASP2. En effet, cette molécule est très soluble et chargée. Cependant, la chaîne aliphatique de cette molécule est suffisante pour instaurer des interactions hydrophobes entre ses méthyles et les aromatiques du coeur de la protéine. En revanche, l'affinité d'ASP2 pour le TSP est sans doute assez faible par rapport à l'IBMP.

Si les interactions de type hydrophobe sont un prérequis à la fixation d'un ligand sur ASP2 et la PBP de la blatte, celui-ci peut donc être aussi légèrement polaire, voire chargé. On peut rappeler ici que certaines phéromones sont des acides gras non saturés et chargés, comme le composé de la phéromone sexuelle de l'abeille (acide 9-céto-2(E)-décénoïque). Le cas de LUSH montre qu'effectivement, des OBPs sont capables de lier des molécules très solubles. Le TSP est biologiquement peu significatif mais il montre que les OBPs peuvent lier des molécules aux propriétés physico-chimiques très différentes.

La cavité de liaison et spectre de liaison La présence de l'ouverture dans la cavité de la PBP de la blatte est intéressante. En effet, la structure d'ASP2 met aussi en évidence une telle ouverture dans la même région de la cavité. Cependant, on ne peut conclure définitivement sur son existence du fait de la mauvaise définition de la région D30-S45. L'ouverture dans la PBP de la blatte fait intervenir trois résidus (R33, N34 et P35) qui correspondent aux résidus A37, A38, K39 de la séquence d'ASP2. Dans une structure d'ASP2 plus "fermée", ces trois résidus sont susceptibles de fermer, au moins partiellement, la cavité. La comparaison des deux structures suggèrent donc que la cavité d'ASP2 peut présenter une ouverture plus restreinte que celle observée sur l'ensemble des structures calculées. Il est intéressant de noter que K39 est chargé. Il est donc possible que l'ouverture ainsi que le résidu chargé autour de l'ouverture aient un rôle à jouer dans le mécanisme de fixation-relarguage.

D'autre part, nous avons proposé à partir des données de compétition IBMP contre TSP que la taille de la cavité d'ASP2 s'adaptait au ligand. Compte tenu de l'observation du même phénomène sur la PBP de la blatte entre la forme apo et complexée avec l'ANS, on peut imaginer que ce processus est généralisé et que la fluidité des chaînes latérales du coeur de la protéine est un critère important dans les propriétés de liaison des OBPs. En suivant cette idée, le critère stérique n'est plus un paramètre crucial du spectre de fixation d'une OBP donnée car la poche de l'OBP peut s'accommoder plus facilement aux ligands de différente taille et les interactions de type hydrophobe, polaire ou électrostatique ont alors un poids très important.

On peut avancer une hypothèse sur la très grande variabilité de séquences des OBPs alors que leurs fonctions sont très proches. En effet, les propriétés de fixation d'une OBP ne dépendent pas uniquement de la présence de telle ou telle chaîne latérale dans la poche. Elles sont peut-être aussi reliées à la taille de la cavité et à la façon dont celle-ci peut se déformer. Cette flexibilité ne peut être réduite à quelques résidus localisés autour de la poche mais elle doit impliquer des variations importantes sur toute la structure. Les contraintes minimales sur la séquence d'une OBP sont alors de créer les hélices et de générer le repliement caractéristique par le biais entre autres, des trois ponts disulfure. La variabilité de séquence permet alors de créer un arrangement d'hélices dont les propriétés "élastiques" seraient en rapport avec le spectre de fixation de l'OBP.

Dynamique structurale et fonction L'étude par RMN de la PBP de *Bombyx mori*, de l'OBP de la drosophile et d'ASP2 a montré que les OBPs ont un comportement conformationnel intéressant. Dans le cas de la PBP du papillon et du complexe ASP2/TSP, le segment correspondant à A37-S45 subit des mouvements de large amplitude dans la gamme picoseconde. Dans le cas de LUSH et d'ASP2, des phénomènes d'échange conformationnel dans la gamme supérieure à la milliseconde affectent les résonances. Ceux-ci sont très im-

portants sur la forme apo de LUSH où la très grande majorité des résonances de la carte ^{15}N -HSQC sont élargies. En présence de ligand, le désordre conformationnel de LUSH est fortement diminué. Dans le cas d'ASP2, un tel équilibre entre plusieurs conformations de la forme libre a également été observé bien que la vitesse d'échange entre les formes libres soit plus faible. Dans le cas de la PBP de la blatte, si le facteur B est assez élevé pour l'apoprotéine et le complexe avec l'ANS, il est fortement abaissé en présence de H3B2.

Ces résultats indiquent que la liaison de ces trois OBPs à un ligand stabilisent une conformation particulière parmi les conformations stables de la protéine libre. Les données sur ASP2 suggèrent que les différentes conformations stables de la protéine libre sont très proches les unes des autres, du moins au niveau du squelette. Comme dans le cas de LUSH, l'ensemble de la structure subit de légères modulations. Il est donc séduisant d'envisager que la conformation sélectionnée dans ces deux systèmes joue un rôle dans la fonction de l'OBP : la stabilisation d'une conformation particulière est peut-être à l'origine de l'augmentation d'affinité de l'OBP avec une autre protéine. Celle-ci pourrait être soit directement le récepteur olfactif, soit une protéine membranaire intermédiaire du type SNMP. Par ce processus, l'activation terminale du récepteur serait reliée à l'état libre-lié de l'OBP et donc à la présence de ligand dans la lymphe sensillaire.

L'existence d'une telle interaction protéine-protéine nécessite la formation d'une zone de contact. Ni les données sur ASP2, ni les différences structurales sur la PBP de la blatte ou de LUSH entre les formes apo et liées ne permettent d'isoler sur la structure une région du squelette dont la conformation est significativement modifiée lors de la complexation avec un ligand.

7.3 Conclusion finale

Les travaux que j'ai menés sur une OBP d'abeille ont apporté de nouveaux éléments à la discussion sur la fonction des protéines de liaison aux odeurs. Certains résultats observés sur ASP2 sont communs à d'autres OBPs (structure, désordre conformationnel de la forme libre). Ils autorisent donc à généraliser ces propriétés aux OBPs. En revanche, nous avons décrit précisément le comportement dynamique du complexe et en particulier, celui du squelette compris entre les hélices α_1 et α_3 . Nous ne pouvons à l'heure actuelle dire si cette région a une réelle importance dans la fonction d'ASP2. Compte tenu de la possibilité de liaison des OBPs à des molécules polaires et faiblement ou très solubles (alcools, TSP), et d'un changement global de conformation lors de la complexation, il devient de plus en plus clair que le rôle des OBPs ne se limite pas à permettre la solubilisation des molécules odorantes dans la lymphe sensillaire ou le mucus.

Le modèle dans lequel le complexe OBP-odeur active directement le récepteur olfactif est séduisant. En revanche, le mécanisme par lequel cette activation se déclenche reste encore très incertain. Il est probable qu'une fois le complexe arrimé au récepteur, l'odeur vient au contact du récepteur. Si le mécanisme de relarguage proposé pour la PBP de *Bombyx mori* semble réaliste, il ne peut s'appliquer à LUSH (où la formation de l'hélice C-terminale n'a pas été observée) et encore moins à ASP2 et la PBP de la blatte, dont la séquence est plus courte. Ces trois protéines doivent donc relarguer le ligand d'une manière plus classique. Il semble donc que les OBPs adoptent plusieurs méthodes de relarguage.

7.4 Perspectives

Il reste encore un grand nombre de points d'interrogation sur ASP2. Ceux-ci concernent la forme libre de la protéine et les propriétés d'interaction de la protéine.

La difficulté à reproduire la transition en pH de la protéine est à analyser, et les origines de la mauvaise reproductibilité sont à rechercher. Lorsque ce mécanisme sera maîtrisé, on peut alors envisager d'étudier la dynamique du squelette sur différentes gammes de temps. Bien que d'après mon expérience, il me semble que la dynamique rapide (inférieure à la milliseconde) ne soit pas considérablement modifiée lors de la fixation, il serait intéressant de le confirmer. Les résultats pourraient apporter des renseignements précieux sur le mécanisme de fixation de ligand.

Nous avons déterminé la présence de deux sites de fixation de l'IBMP dans la cavité d'ASP2. Cependant, l'orientation précise de ces deux molécules n'a pas été obtenue. L'utilisation d'un échantillon marqué ^{15}N - ^{13}C d'ASP2 contenant l'IBMP analysé par d'expériences filtrées isotopiquement permettraient de reconnaître les contacts des groupes méthoxy et isobutyle avec les chaînes latérales de la protéine.

La détermination des constantes de dissociation des complexes a été impossible par RMN du fait de l'équilibre des deux formes en solution. De plus, le mécanisme de fixation de l'IBMP reste encore à l'état de proposition. Une étude plus poussée sur la cinétique et la thermodynamique des nombreuses réactions mériterait d'être menée plus avant.

Les études d'interaction d'ASP2 par microcalorimétrie et par RMN ont montré que la protéine pouvait lier des molécules aux propriétés très différentes mais toujours suffisamment hydrophobes, au moins localement (comme le TSP). Il serait intéressant de préciser par différentes méthodes les classes de molécules qu'elle lie et quels sont les contacts mis en jeu. En particulier, je propose de rechercher un ligand soluble mais ayant une signification biologique plus importante que le TSP. La découverte d'une telle molécule amènerait à abandonner le modèle simple de transporteur.

Bibliographie

- A. Annala, H. Aitio, E. Thulin, and T. Drakenberg. Recognition of proteins folds via dipolar couplings. *J. Biomol. NMR*, 1999 :223–230, 1999.
- P.W. Atkins. *Chimie Physique*. Number Vol 2. Vuibert - Technique & Documentation - Lavoisier, 1982.
- L. Ban, A. Scaloni, C. D'Ambrosio, L.Zhang, Y. Yan, and P. Pelosi. Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cell. Mol. Life Sci.*, 60 :390–400, 2003.
- L. Banci, I.C. Felli, and R. Kümmeler. Direct detection of hydrogen bonds in monomeric superoxyde dismutase. *Biochemistry*, 41 :2913–2920, 2002.
- G. Barbato, M. Ikura, L.E. Kay, R.W. Pastor, and A. Bax. Backbone dynamics of Calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy : the central helix is flexible. *Biochemistry*, 31 :5269–5278, 1992.
- A. Bax, G. Kontaxis, and N. Tjandra. Dipolar couplings in macromolecular structure determination. *Methods in Enzymology*, 339 :127–174, 2001.
- A. Bax and N. Tjandra. High-resolution heteronuclear ubiquitin in an aqueous liquid crystalline medium. *J. Biomol. NMR*, 10 :289–292, 1997.
- L. Briand, E. Lescop, V. Bézirard, N. Birlirakis, J. C. Huet, C. Henry, E. Guittet, and J. C. Pernollet. Isotopic double-labeling of two honeybee odorant-binding proteins secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.*, 23 :167–74, 2001a.
- L. Briand, E. Lescop, V. Bézirard, N. Birlirakis, J. C. Huet, C. Henry, E. Guittet, and J. C. Pernollet. Isotopic double-labeling of two honeybee odorant-binding proteins secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.*, 23(1) :167–74, 2001b.
- L. Briand, V. Perez, J. C. Huet, E. Danty, C. Masson, and J. C. Pernollet. Optimization of the production of a honeybee odorant-binding protein by *Pichia pastoris*. *Protein Expr. Purif.*, 15 :362–9, 1999.
- T. Burova, Y. Choiset, C.K. Jankowski, and T. Haertlé. Conformational stability and binding properties of porcine odorant binding protein. *Biochemistry*, 38 :15043–15051, 1999.
- H.Y. Carr and E. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.*, 94 :630–638, 1954.

- J. Cavanagh, W. Fairbrother, A. Palmer, and N. Skelton. *Protein NMR spectroscopy - Principles and Practice*. Edition Academic Press Inc., 1996.
- D. Charvolin, J.P. Douliez, D. Marion, C. Cohen-Addad, and E. Pebay-Peroula. The crystal structure of a wheat nonspecific lipid transfer protein (ns-LPT1) complexed with two molecules of phospholipid at 2.1Å resolution. *Eur. J. Biochem.*, 264 :562–568, 1999.
- G.M. Clore, P.C. Driscoll, P.T. Wingfield, and A.M. Gronenborn. Analysis of the backbone dynamics of interleukine-1 β using two-dimensional inverse detected heteronuclear ^{15}N - ^1H NMR spectroscopy. *Biochemistry*, 29(32) :7387–7401, 1990a.
- G.M. Clore, A.M. Gronenborn, and A. Bax. A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.*, 133 :216–221, 1998a.
- G.M. Clore, A.M. Gronenborn, and N. Tjandra. Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. reson.*, 131 :159–162, 1998b.
- G.M. Clore, A. Szabo, A. Bax, L.E. Kay, P.C. Driscoll, and A.M. Gronenborn. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.*, 112 :4989–91, 1990b.
- R. Cole and J.P. Loria. FAST-Modelfree : a program for rapid automated analysis of solution NMR spin-relaxation data. *J. Biomol. NMR*, 26 :203–213, 2003.
- F. Cordier and S. Grzesiek. Direct evidence of hydrogens bonds in proteins by interresidue $^3\text{hJNC}'$ scalar couplings. *J. Am. Chem. Soc.*, 121 :1601–1602, 1999.
- G. Cornilescu, F. Delaglio, and A. Bax. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR*, 13(3) :289–302, 1999a.
- G. Cornilescu, J.-S. Hu, and A. Bax. Identification of the hydrogen bonding network in a protein by scalar coupling. *J. Am. Chem. Soc.*, 121 :2949–2950, 1999b.
- G. Cornilescu, B.E. Ramirez, M.K. Frank, G.M. Clore, A.M. Gronenborn, and A. Bax. Correlation between $^3\text{hJ}_{\text{NC}'}$ and hydrogen bond length in proteins. *J. Am. Chem. Soc.*, 121 :6275–6269, 1999c.
- J. Coutant. *Etudes par RMN de la protéine orf78 et des complexes actine/thymosine et actine/domaine 1 de ciboulot. Utilisation des milieux faiblement orientants pour la détermination de structure*. PhD thesis, Université Paris VI, 2003.
- S. Curry, H. Mandelkow, P. Brick, and N. Franks. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Biol.*, 5 :827–835, 1998.
- N. D'amelio, A.M. Bonvin, M. Czisch, P. Baker, and R. Kaptein. The C terminus of Apocytochrome b562 undergoes fast motions and slow exchange among ordered conformations resembling the folded state. *Biochemistry*, 41 :5505–5514, 2002.
- E. de Alba and N. Tjandra. NMR dipolar couplings for the structure determination of biopolymers in solution. *Prog. in Nucl. Magn. Reson. Spectr.*, 40 :175–197, 2002.

- F. Delaglio, G. Kontaxis, and A. Bax. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.*, 122 :2142–2143, 2000.
- J.K. Denny, J. Wang, T.A. Cross, and J.R. Quine. PISEMA powder patterns and PISA wheels. *J. Magn. Reson.*, 152 :217–226, 2001.
- C. Deverell, R.E. Morgan, and J.H. Strange. Studies of chemical exchange by nuclear magnetic relaxation in the rotating frame. *Mol. Phys.*, 18 :553–559, 1970.
- K. Dijkstra, G.J.A. Kroon, N.A.J. Van Nuland, and R.M. Scheek. The COCAH experiment to correlate intraresidue carbonyl, C_α , and H_α resonances in proteins. *J. Magn. Reson.*, 107A : 102–105, 1994.
- A.J. Dingley and S. Grzesiek. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings. *J. Am. Chem. Soc.*, 120 :8293–8297, 1998.
- P. Dosset, J.-C. Hus, M. Blackledge, and D. Marion. Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR.*, 16 :23–28, 2000.
- P. Dosset, J.-C. Hus, D. Marion, and M. Blackledge. A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J. Biomol. NMR*, 20 :223–231, 2001.
- P. Dux, G. Rubinstenn, GW Vuister, R. Boelens, F.A. Mulder, K. Hard, W.D. Hoff, A.R. Kroon, W. Crielaard, K.J. Hellingwerf, and R. Kaptein. Solution structure and backbone dynamics of the photoactive yellow protein. *Biochemistry*, 37 :12689–99, 1998.
- D. Eliezer, J. Dao, H.J. Dyson, and P.E. Wright. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.*, 5 :148–155, 1998.
- M.W.F. Fischer, J.A. Losonczi, J.L. Weaver, and J.H. Prestegard. Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. *Biochemistry*, 38 :9013–9022, 1999.
- S. Gaemers and A. Bax. Morphology of three lyotropic liquid crystalline biological NMR media studied by translational diffusion anisotropy. *J. Am. Chem. Soc.*, 123 :12343–12352, 2001.
- G. Goussard. *Etude par RMN de la dynamique en solution des domaines de liaison à l'ADN de deux protéines, AlcR et FruR*. PhD thesis, Université PARIS XI, 2000.
- S. Grzesiek and A. Bax. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched protein. *J. Mag. Reson.*, 99 :201–207, 1992a.
- S. Grzesiek and A. Bax. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.*, 96 :432–40, 1992b.
- S. Grzesiek and A. Bax. The importance of not saturating H₂O in protein NMR. application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.*, 115 :12593–12594, 1993.
- S. Grzesiek, F. Cordier, and A.J. Dingley. Scalar couplings across hydrogen bonds. *Methods in enzymology*, 338 :111–133, 2001.

- G.W. Han, J.Y. Lee, H.K. Song, C.S. Chang, K. Min, J. Moon, D.H. Shin, M.L. Kopka, M.R. Sawaya, H.S. Yuan, T.D. Kim, J. Choe, D. Lim, H.J. Moon, and S.W. Suh. Structural basis of non-specific lipid binding in maize lipid-transfer protein-complexes revealed by high-resolution X-ray crystallography. *J. Mol. Biol.*, 308 :263–278, 2001.
- F.J.M Van de Ven. *Multidimensional NMR in liquids*. p339/340.
- C. Van Heijenoort. *Etude de la dynamique en solution d'une protéine de 98 résidus, la capsicéine, à partir de l'analyse de la relaxation des spins azote-15 par résonance magnétique nucléaire*. PhD thesis, Université Paris XI, 1994.
- C.W. Hilbers and C. MacLean. Electric field effects in NMR. *Chem. Phys. Lett.*, 2 :445, 1968.
- L. Holm and C. Sander. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, 233 :123–138, 1993.
- J.-S. Hu, S. Grzesiek, and A. Bax. Two-dimensional NMR methods for determining χ_1 angles of aromatic residues in proteins from three-bond $J_{C'C\gamma}$ and $J_{NC\gamma}$ couplings. *J. Am. Chem. Soc.*, 119 :1803–1804, 1997.
- M. Karplus. Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.*, 30 :11–15, 1959.
- L.E. Kay, G.Y. Xu, and T. Yamazaki. Enhanced-sensitivity triple-resonance spectroscopy with minimal H₂O saturation. *J. Magn. Reson.*, 109A :129–133, 1994.
- D.M. Korzhnev, M. Billeter, A.S. Arseniev, and V.Y. Orekhov. NMR studies of Brownian tumbling and internal motions in proteins. *Progress in NMR spectroscopy*, 38 :197–266, 2001.
- J. Krieger, K. Raming, G. D. Prestwich, D. Frith, S. Stabel, and H. Breer. Expression of a pheromone-binding protein in insect cells using a baculovirus vector. *Eur. J. Biochem.*, 203(1-2) :161–166, 1992.
- S.W. Kruse, R. Zhao, D.P. Smith, and D.N. Jones. Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. *Nat. Struct. Biol.*, 10 :694–700, 2003.
- Y. Laroche, V. Storme, J. de Meutter, J. Messens, and M. Lauwereys. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast *Pichia pastoris*. *Bio/Technology*, 12 :1119–1124, 1996.
- A. Lartigue, A. Gruez, S. Spinelli, S. Rivière, R. Brossut, M. Tegoni, and C. Cambillau. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. *J. Biol. Chem.*, 278 :30213–30218, 2003.
- D. Lee, F. Damberger, G. Peng, R. Horst, P. Guntert, L. Nikonova, W. Leal, and K. Wüthrich. NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. *FEBS Letters*, 531 :314–318, 2002.
- L.K. Lee, M. Rance, W.J. Chazin, and A.G. Palmer. Rotational diffusion anisotropy of proteins from simultaneous analysis of ¹⁵N and ¹³C_α nuclear spin relaxation. *J. Biomol. NMR*, 9 :287–298, 1997.

- M.H. Lerche, B.B. Kragelund, L.M. Bech, and F.M. Poulsen. Barley lipid-transfer protein complexed with malpitoyl CoA : the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands. *Structure*, 5 :291–306, 1997.
- E. Lescop, L. Briand, J. C. Pernollet, C. Van Heijenoort, and E. Guittet. ^1H , ^{13}C and ^{15}N chemical shift assignment of the honeybee odorant-binding protein ASP2. *J. Biomol. NMR*, 21(2) :181–2, 2001.
- G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation. 2. Analysis of experimental results. *J. Am. Chem. Soc.*, 104 :4559–4570, 1982a.
- G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecule. 1 Theory and range of validity. *J. Am. Chem. Soc.*, 104 :4546–4559, 1982b.
- P.L. Loria, M. Rance, and A.G. Palmer III. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterising chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.*, 121 : 2331–2332, 1999.
- J.A. Losonczi, M. Andrec, M.W.F. Fischer, and J.H. Prestegard. Order matrix analysis of residual couplings using singular value decomposition. *J. Magn. Reson.*, 138 :334–342, 1999.
- A.M. Mandel, M. Akke, and A.G. Palmer. Dynamics of ribonuclease H : temperature dependence of motions on multiple time scales. *Biochemistry*, 35(50) :16009–23, 1996.
- S. Marchese, D. Pes, A. Scaloni, V. Carbone, and P. Pelosi. Lipocalins of boar salivary glands binding odours and pheromones. *Eur. J. Biochem.*, 252 :563–568, 1998.
- D. Marion, M. Ikura, R. Tschudin, and A. Bax. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange protein. *J. Magn. Reson.*, 85 :393–399, 1989.
- S. Meiboom and D. Gill. Modified Spin-Echo method for measuring relaxation times. *D. Rev. Sci. Instrum.*, 29 :688–691, 1958.
- J. Meiler, N. Blomberg, M. Nilges, and C. Griesinger. A new approach for applying residual dipolar couplings as restraints in structure elucidation. *J. Biomol. NMR*, 16 :245–252, 2000a.
- J. Meiler, W. Peti, and C. Griesinger. Dipocoup : a versatile program for 3D-structure homology comparison based on residual dipolar couplings and pseudocontact shifts. *J. Biomol. NMR*, 17 : 283–294, 2000b.
- M.F. Mesleh and S.J. Opella. Dipolar waves as NMR maps of helices in proteins. *J. Magn. Reson.*, 163 :288–299, 2003.
- M.F. Mesleh, G. Veglia, T.M. DeSilva, F.M. Marassi, and S.J. Opella. Dipolar Waves as NMR maps of proteins structure. *J. Am. Chem. Soc.*, 124 :4206–4207, 2002.
- O. Millet, J.P. Loria, C.D. Kroenke, M. Pons, and A.G. Palmer III. The static magnetic field dependance of chemical exchange linebroadening defines the NMR chemical shift time scale. *J. Am. Chem. Soc.*, 122 :2867–2877, 2000.

- S. Moltke and S. Grzesiek. Structural constraints from residual tensorial couplings in high resolution NMR without an explicit term for the alignment tensor. *J. Biomol. NMR*, 15 :77–82, 1999.
- M. Dal Monte, M. Centini, C. Anselmi, and P. Pelosi. Binding of selected odorants to bovine and porcine odorant-binding proteins. *Chem. Senses*, 18 :713–721, 1993.
- F.A. Mulder, N.R. Skrynnikov, B. Hon, F.W. Dahlquist, and L.E. Kay. Measurement of slow (μ s) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy : application to Asn and Gln residues in a cavity mutant of T4 Lysozyme. *J. Am. Chem. Soc.*, 123 : 967–975, 2001.
- L.K. Nicholson, L.E. Kay, D.M. Baldissari, J. Arango, P.E. Young, A. Bax, and D.A. Torchia. Dynamics of methyl groups in proteins as studied by proton-detected ^{13}C -NMR spectroscopy. application to the leucine residues of Staphylococcal Nuclease. *Biochemistry*, 31 :5253–5263, 1992.
- J.H. Noggle and R.E. Shirmer. *The Nuclear Overhauser Effect*. Academic Press, New York, 1971.
- M. Ottiger and A. Bax. Bicelle-based liquid crystals for NMR-measurement of dipolar couplings at acidic and basic pH values. *J. Biomol. NMR*, 13 :187–191, 1999.
- M. Ottiger, F. Delaglio, and A. Bax. Measurements of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, 131 :373–378, 1998.
- G. Otting, E. Liepinsh, and K. Wüthrich. Disulfide bond isomerization in BPTI and BPTI(G36S) : an NMR study of correlated mobility in proteins. *Biochemistry*, 32 :3571–3582, 1993.
- A.G. Palmer III. *Protein NMR spectroscopy - principles and Practice*, chapter Relaxation and dynamic processes. edition academic press inc., 1996.
- A.G. Palmer III, J. Cavanagh, P.E. Wright, and M. Rance. Sensivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.*, 93 :151–170, 1991a.
- A.G. Palmer III, C.D. Kroenke, and J.P. Loria. NMR methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in enzymology*, 239 (part B) :204–238, 2001.
- A.G. Palmer III, M. Rance, and P.E. Wright. Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance ^{13}C heteronuclear NMR spectroscopy. *J. Am. Chem. Soc.*, 113(12) :4371–4380, 1991b.
- A. Pardi, M. Billeter, and K. Wüthrich. Calibration of the angular dependence of the amide proton- C_{α} proton coupling constants, $^3J_{HN-H_{\alpha}}$, in a globular protein. use of $^3J_{HN-H_{\alpha}}$ for identification of helical secondary structure. *J Mol Biol*, 180(3) :741–51, Dec 15 1984.
- A. Pardi, G. Wagner, and K. Wüthrich. Protein conformation and proton nuclear-magnetic-resonance chemical shifts. *Eur. J. Biochem.*, 137 :445–454, 1983.
- W.R. Pearson and D.J. Lipman. Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci.*, 85 :2444–2448, 1988.
- P. Pelosi. Odorant-binding proteins. *Critic. Reviews Biochem. Molec. Biol.*, 29(3) :199–228, 1994.

- J.W. Peng and G. Wagner. Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry*, 31 :8571–8586, 1992.
- K. Pervushin. Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Quarterly reviews of biophysics*, 33(2) :161–197, 2000.
- W. Peti, J. Meiler, R. Brüschweiler, and C. Griesinger. Model-free analysis of protein backbone motion from residual dipolar couplings. *J. Am. Chem. Soc.*, 124 :5822–5833, 2002.
- D. Picone, O. Crescenzi, S. Angeli, S. Marchese, A. Brandazza, L. Ferrara, P. Pelosi, and A. Scalonì. Bacterial expression and conformational analysis of a chemosensory protein from *Schistocerca gregaria*. *Eur. J. Biochem.*, 268 :4794–4801, 2001.
- M. Piotto, V. Saudek, and V. Sklenar. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, 2(6) :661–5, 1992.
- E. Plettner, J. Lazar, E.G. Prestwich, and G.D. Prestwich. Discrimination of pheromone enantiomers by two pheromone binding proteins from the gypsy moth *Lymantria dispar*. *Biochemistry*, 39 :8953–8962, 2000.
- J.H. Prestegard, H.M. Al-Hashimi, and J.R. Tolman. NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Quarterly Reviews of Biophysics*, 33 :371–424, 2000.
- G. Raffard, S. Steinbruckner, A. Arnold, J.H. Davis, and E.J. Dufourc. Temperature-composition diagram of dimyristoylphosphatidylcholine-dicaproylphosphatidylcholine "bicelles" self-orienting in the magnetic field. a solid state ^2H and ^{31}P study. *Langmuir*, 16 :7655–7662, 2000.
- B.E. Ramirez and A. Bax. Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium. *J. Am. Chem. Soc.*, 120 :9106–9107, 1998.
- C.A. Rohl and D. Baker. De novo determination of protein backbone structure from residual dipolar couplings using Rosetta. *J. Am. Chem. Soc.*, 124 :2723–2729, 2002.
- S. Rothmund, Y.-C. Liou, P.L. Davies, E. Krause, and F.D. Sönnichsen. A new class of hexahelical insect proteins revealed as putative carriers of small hydrophobic ligands. *Structure*, 7 :1323–1332, 1999.
- C.R. Sanders and J.P. Schwonek. Characterization of magnetically orientable bilayers in mixtures of dihexanoylphosphatidylcholine and dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry*, 31 :8898–8905, 1992.
- M. Sattler, J. Schleucher, and C. Griesinger. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradient. *Progress in NMR Spectroscopy*, 34 :93–158, 1999.
- D.H. Shin, J.Y. Lee, K.Y. Hwang, K.K. Kim, and S.W. Suh. High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure*, 3 :189–199, 1995.
- N.R. Skrynnikov, N.K. Goto, D.W. Yang, W.Y. Choy, J.R. Tolman, G.A. Mueller, and L.E. Kay. Orienting domains in proteins using dipolar couplings measured by liquid-state NMR : differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin. *J. Mol. Biol.*, 295 :1265–1273, 2000.

- N. Tjandra and A. Bax. Direct measurement of distances and angles in biomolecular by NMR in a dilute liquid crystalline medium. *Science*, 278 :1111–1114, 1997.
- N. Tjandra, J.G. Omichinski A.M. Gronenborn, G.M. Clore, and A. Bax. Use of dipolar ^1H - ^{15}N and ^1H - ^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.*, 4 :732–738, 1997.
- N. Tjandra, J. Marquardt, and G.M. Clore. Direct refinement against proton-proton dipolar couplings in NMR structure determination of macromolecules. *J. Magn. Reson.*, 2000 :393–396, 2000.
- J.R. Tolman, J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard. Nuclear magnetic dipolar interactions in field-oriented proteins : information for structure determinations in solution. *Proc. Natl. Acad. Sci USA*, 92 :9279–9283, 1995.
- D. A. Torchia, S. W. Sparks, and A. Bax. Staphylococcal nuclease : sequential assignments and solution structure. *Biochemistry*, 28(13) :5509–24, Jun 27 1989.
- G.W. Vuister and A. Bax. Quantitative J correlation : a new approach for measuring homonuclear three-bond $J(H_N H_\alpha)$ coupling constants in ^{15}N enriched proteins. *J. Am. Chem. Soc.*, 115 :7772–7777, 1993.
- G.W. Vuister and A. Bax. Measurement of four-bond $H_N - H_\alpha$ J-couplings in staphylococcal nuclease. *J. Biomol. NMR*, 4(2) :193–200, 1994.
- H. Wang, M. Eberstadt, E. T. Olejniczak, R.P. Meadows, and S.W. Fesik. A liquid crystalline medium for measuring residual dipolar couplings over a wide range of temperatures. *J. Biomol. NMR*, 12 : 443–446, 1998.
- J. Wang, J. Denny, C. Tian, S. Kim, Y. Mo, F. Kovacs, Z. Song, K. Nishimura, Z. Gan, R. Fu, J.R. Quine, and T.A. Cross. Imaging membrane protein helical wheels. *J. Magn. Reson.*, 2000 : 162–167, 2000.
- Y.X. Wang, J. Jacob, F. Cordier, P. Wingfield, S.J. Stahl, S. Lee-Huang, D. Torchia, S. Grzesiek, and A. Bax. Measurement of $^3hJ(NC')$ connectivities across hydrogen bonds in a 30kDa protein. *J. Biomol. NMR*, 14 :181–184, 1999.
- D. Wishart and B. Sykes. The ^{13}C Chemical Shift Index : a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol NMR*, 4 :171–180, 1994.
- D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol*, 222(2) :311–33, Nov 20 1991.
- D. S. Wishart, B. D. Sykes, and F. M. Richards. The chemical shift index : a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, 31 (6) :1647–51, Feb 18 1992.
- M.J. Wood and E.A. Komives. Production of large quantities of isotopically labeled protein in *Pichia pastoris* by fermentation. *J. Biomol. NMR*, 13 :149–159, 1999.
- C.H. Wu, A. Ramamoorthy, and S.J. Opella. High resolution heteronuclear dipolar solid-state NMR spectroscopy. *J. Magn. Reson.*, A109 :270–274, 1994.

K. Wüthrich. *NMR of Proteins and Nucleic Acids*. John Wiley, 1986.

M. Zweckstetter and A. Bax. Prediction of sterically induced alignment in a dilute liquid crystalline phase : aid to protein determination by NMR. *J. Am. Chem. Soc.*, 122 :3791–3792, 2000.

Annexe A

Paramètres d'acquisition et de traitement des expériences 3D.

Séquence	HNCA	HN(CO)CA	CBCA(CO)NH	CBCANH	HNCO	HN(CA)CO	¹³ C-HCCH-TOCSY	¹³ C-NOESY-HSQC	¹⁵ N-DIPSI-HSQC	¹⁵ N-NOESY-HSQC
temps de mélange							33ms	120ms	70ms	150ms
ns	32	16	16	32	16	32	16	16	24	24
dimension ¹³ C										
TD	50	50	50	100	50	50	60	80		
SW(ppm)	35	35	77,65	77,65	17,5	12	70	70		
filtre	S	Q	S	S			S	S		
prédiction linéaire	LPPc(20)	LPPc(20)	LPPc(20)	non			non			
zero-filling	256	256	256	256	128	128				
dimension ¹⁵ N										
TD	50	50	50	50	50	42			32	32
SW(ppm)	30	30	30	30	30	30			30	30
filtre	S	Q	S	S						
prédiction linéaire	LPPc(20)	LPPc(20)	LPPc(20)	LPPc(20)						
zero-filling	64	64	64	128	128	128				
dimension ¹ H (indirecte-directe)										
TD	512	512	512	512	512	512	150-512	200-512	146-512	334-512
SW(ppm)	3,58	3,58	3,58	3,58	3,58	3,58	5,5-6,62	6,6-12	11-3,58	12-3,58
filtre	Q	Q	Q	Q			Q			
prédiction linéaire	non	non	non	non						
zero-filling	512	512	512	512	512	512				

Annexe B

Tableau des déplacements chimiques d'ASP2

L'attribution d'ASP2 a été menée à 308K et à pH 6. Les résonances superposées sont marquées par une étoile (*). Les résonances non attribuées sont marquées par x (chaînes latérales de K39, K51, Q36 et Q87). L'absence d'attribution est due à la superposition des résonances.

Résidu	amide	α	β	γ	δ	autre	C'
I1		3.748 59.759	1.991 39.063	1.231 1.604 0.955 26.659 17.000	0.820 13.000		171.700
D2	8.797 126.796	4.861 53.051	2.928 2.683 41.043			C'_γ 180.4	175.500
Q3	8.806 125.042	3.776 59.540	1.880* 28.200	1.93* 33.441		H_ϵ 6.739 7.060 $N_{\epsilon 2}$ 110.428 C'_δ 179.3	176.700
D4	8.513 117.139	4.413 57.126	2.727* 39.986				178.760
T5	7.698 113.328	4.308 64.027	4.288 69.530	1.250 21.677			176.500
V6	7.942 120.369	3.804 66.277	2.326 31.499	1.148* 22.209*			177.400
V7	8.488 120.707	3.768 68.055	2.209 31.221	1.165 1.036 24.106 21.670			177.600
A8	7.703 120.220	4.166 55.207	1.568 17.854				180.700
K9	7.796 118.410	4.214 59.104	2.017 1.756 32.640	1.551 1.710 25.438	1.704* 28.749	H_ϵ 2.971 C_ϵ 41.500	180.300
Y10	8.407 120.397	4.217 62.867	3.280 2.948 38.416		H_δ 7.300 C_δ 130.7	H_ϵ 6.750 C_ϵ 118.2	178.500
M11	8.568 116.695	4.339 57.356	1.986 2.320 31.010	2.786* 32.652		H_ϵ 2.044 C_ϵ 16.176	178.300
E12	8.158 118.482	4.023 59.400	2.187* 29.867	2.210 2.416 36.336			178.800
Y13	7.786 118.990	4.022 61.292	3.010 3.296 38.837		H_δ 7.040 C_δ 130.9	H_ϵ 6.840 C_ϵ 115.4	177.500
L14	8.299 115.532	4.561 55.097	2.067 1.605 43.799	2.286 27.222	1.127 0.956 22.479 26.400		178.620
M15	8.306 124.135	4.554 59.150	2.214 2.594 28.252	2.600 2.846 32.320		H_ϵ 2.021 C_ϵ 16.197	173.770
P16		4.361 65.234	1.705 2.328 30.965	1.900* 28.162	3.544 3.630 51.000		176.130
D17	8.017 113.823	4.854 56.036	2.538* 44.690				175.950
I18	8.421 120.469	3.474 65.690	1.800 37.900	0.916 1.706 0.823 30.124 17.633	0.918 14.523		177.660
M19	7.827 115.253	4.430 59.600	2.120* 29.200	2.560 2.760 32.000		H_ϵ 2.020	175.950
P20		4.481 65.740	1.758 2.352 30.837	2.118* 28.228	3.386 3.539 50.613		180.300
C21	6.770 116.652	4.420 59.726	3.264 2.767 41.326				175.300
A22	8.860 123.711	3.840 55.605	1.455 16.800				179.300
D23	7.953 117.366	4.400 57.100	2.740* 41.100				179.400
E24	8.012 120.191	3.920 59.100	2.100 2.250 29.400	2.380* 36.200			177.960
L25	7.639 116.484	4.239 54.433	1.641* 42.220	1.761 27.498	0.849 0.759 22.796 25.437		175.800
H26	7.751 115.650	4.306 56.338	3.288 3.473 26.350		$H_{\delta 2}$ 7.310	$H_{\epsilon 1}$ 8.600	174.000
I27	8.471 118.987	4.292 59.693	1.553 39.986	1.092 1.510 0.976 27.404 17.300	0.821 14.410		175.400

Résidu	amide	α	β	γ	δ	autre	C'
S28	8.358 120.822	4.270 58.579	4.050 3.899 63.789				175.600
E29	8.795 123.708	3.890 59.700	2.030* 29.000	2.330 2.220 35.600			177.150
D30	8.312 118.140	4.291 55.879	2.614* 39.780				177.220
I31	7.480 119.999	3.949 62.471	1.920 37.800	1.195 1.507 0.960 28.000 17.500	0.868 12.368		177.140
A32	8.031 120.492	3.828 54.521	1.275 19.271				176.600
T33	7.523 104.892	4.251 61.890	4.429 69.244	1.260 21.600			174.160
N34	7.524 118.321	4.860 53.000	2.738 2.934 38.893			C'_γ 174.8	175.960
I35	8.334 122.158	3.960 63.100	1.920 37.800	1.195 1.507 0.960 28.000 27.995	0.868 12.385		176.570
Q36	8.059 119.815	3.940 56.700	2.105* 29.600	2.370* 34.0		H_ϵ 6.80 7.51 $N_{\epsilon 2}$ 111.6 C'_δ 180	176.570
A37	7.814 122.347	4.220 52.838	1.390 18.400				178.430
A38	8.065 121.214	4.200 54.802	1.413 18.400				178.700
K39	7.998 118.529	4.230 57.000	1.760* 32.700	1.440 * 27.400	1.640 X	H_ϵ 2.970 C_ϵ 41.000	176.700
N40	7.956 117.098	4.741 53.132	2.766 2.910 38.800			H_δ 6.880 7.520 N_δ 112.397 C'_γ 177.03	175.700
G41	8.106 108.284	4.069 3.890 45.501					174.300
A42	7.899 123.183	4.301 52.352	1.365 19.428				177.210
D43	8.325 118.921	4.616 54.228	2.791 2.694 41.100				176.800
M44	8.497 123.161	4.491 56.168	2.112* 32.071	2.577 2.769 32.071			178.760
S45	8.583 118.813	4.250 61.261	3.998* 62.571				176.330
Q46	8.411 120.966	4.333 57.670	2.190* 28.734	2.392 * 33.971		H_ϵ 6.810 7.510 N_ϵ 111.629 C'_δ 180	177.890
L47	7.928 119.053	4.195 56.888	1.564 1.953 41.368	1.720 28.900	0.934 0.881 25.257 25.200		178.700
G48	8.300 110.251	3.690 4.166 48.075					176.420
C49	8.330 118.987	4.706 56.078	3.064 3.228 36.686				177.900
L50	8.030 125.664	3.984 58.744	1.530 2.273 40.613	1.599 27.590	0.900 0.772 24.200 26.100		177.700
K51	8.252 118.813	3.770 60.100	1.820 1.820 29.500	1.425 * X	1.550 * X		178.620
A52	8.383 119.923	4.110 54.802	1.413 18.400				178.500
C53	8.069 118.781	4.175 60.543	3.279 3.577 38.199				177.100
V54	8.137 123.260	3.369 67.009	2.283 31.548	0.958 * 22.368 *			177.300
M55	8.258 117.549	3.754 59.700	1.932 2.209 34.248	2.835 * 32.600		H_ϵ 1.898 C_ϵ 17.775	179.000
K56	8.553 120.320	4.551 59.388	2.010* 32.200	1.534 1.786 25.563	1.737* 29.357	H_ϵ 2.800 2.940 C_ϵ 41.800	181.000
R57	8.210 122.117	4.072 60.784	1.990 2.040 30.622	1.590 2.110 29.916	3.434 3.608 44.056		179.100
I58	7.802 109.836	4.558 61.518	2.347 37.648	1.229 1.613 0.960 26.013 16.511	0.811 13.019		174.900
E59	7.501 111.580	3.763 58.279	2.105 2.434 26.045	2.400* 32.100			174.500
M60	8.208 116.146	4.821 56.631	1.680 2.252 34.058	2.388* 32.147		H_ϵ 1.215 C_ϵ 17.096	173.200
L61	6.677 120.257	5.222 52.900	1.419 0.943 46.277	1.410 27.100	0.836 0.700 26.139 23.359		175.000
K62	8.734 126.899	4.520 55.582	1.817* 33.861	1.173 1.274 24.435	1.438* 29.010	H_ϵ 2.688 C_ϵ 41.689	176.500
G63	9.045 117.665	3.657 4.187 46.893					173.900
T64	8.598 116.658	4.249 60.842	4.591 68.404	1.113 21.688			173.000
E65	7.979 122.209	4.407 55.345	2.103 1.906 32.086	2.185 2.250 36.201			174.100
L66	7.825 121.613	5.254 53.000	1.572 0.930 45.443	1.447 27.117	0.752 0.569 26.726 25.730		176.400
Y67	9.147 122.111	4.550 56.937	2.805 3.154 40.245		H_δ 7.130 C_δ 130.9	H_ϵ 6.770 C_ϵ 115.8	175.000
V68	8.076 117.985	3.899 62.683	1.746 33.347	0.722 0.572 21.878 19.423			177.000
E69	8.919	3.835	2.186 1.998	2.315 2.505			

Résidu	amide 124.003	α 62.820	β 26.254	γ 37.474	δ	autre	C' 176.100
P70		4.391 65.700	1.501 2.600 32.578	1.810 1.920 28.300	3.690 3.610 50.000		179.530
V71	7.453 117.585	3.827 65.210	2.340 31.600	1.123 1.204 23.325 22.219			178.060
Y72	8.288 120.630	4.414 58.370	3.169 3.654 35.693		H_δ 7.060 C_δ 127.8	H_ϵ 6.980 C_ϵ 116.5	177.830
K73	7.806 117.295	4.044 59.800	1.969* 32.160	1.569* 24.758	1.732* 29.312	H_ϵ 3.050 2.966 C_ϵ 41.865	177.980
M74	7.138 118.488	3.485 58.558	1.255 1.889 30.106	1.781 2.494 33.058		H_ϵ 1.970 C_ϵ 16.164	177.800
I75	8.831 121.638	3.625 65.958	2.086 38.508	1.339 2.108 1.021 30.108 18.588	1.007 16.456		177.250
E76	7.976 116.496	3.756 59.161	2.112 2.187 29.125	2.199 2.545 37.204			177.600
V77	7.304 117.132	3.933 65.430	2.113 32.549	1.217 1.071 22.287 21.126			178.600
V78	8.659 120.371	3.306 66.525	1.350 31.977	0.399 -0.259 22.346 20.128			176.800
H79	8.294 114.738	4.569 54.810	3.383 3.017 27.166		$H_{\delta 2}$ 6.905 $C_{\delta 2}$ 126.0	$H_{\epsilon 1}$ 7.510 $C_{\epsilon 1}$ 133.400	172.800
A80	6.537 120.007	3.981 54.183	1.515 18.994				178.600
G81	8.915 109.077	4.335 3.707 44.748					174.000
N82	7.610 118.428	4.997 52.065	3.148 2.711 39.116			H_δ 6.92 7.411 N_δ 113.572 C'_γ 177.4	174.000
A83	8.744 127.130	3.991 55.415	1.487 18.994				180.000
D84	8.261 118.360	4.413 57.126	2.780 2.716 39.986				178.800
D85	7.633 122.186	4.338 56.811	2.849 2.316 39.689				177.500
I86	8.000 119.067	3.173 66.279	1.808 37.960	0.671 1.916 1.006 31.328 18.564	0.856 13.866		177.000
Q87	7.575 116.604	3.867 58.880	2.138* 28.453	2.436* 33.488		H_ϵ 6.80 X N_ϵ 111.6 C'_δ 180	178.900
L88	7.868 121.820	4.042 58.150	1.876 1.754 42.450	1.462 27.025	0.901 0.840 23.922 25.111		179.800
V89	8.675 120.251	3.428 66.160	1.726 30.721	0.488 0.950 22.602 22.887			178.300
K90	8.498 120.748	3.881 61.590	1.491 1.757 32.108	0.695 1.840 27.750	0.774 1.208 29.603	H_ϵ 2.170 2.240 C_ϵ 41.300	179.000
G91	7.540 105.371	3.961* 47.361					176.800
I92	7.602 124.772	3.541 65.427	1.805 38.094	0.817 1.724 0.430 28.334 17.621	0.515 13.466		177.300
A93	8.755 121.042	4.202 54.531	1.753 18.798				181.000
N94	7.989 114.467	4.611 56.062	3.165 2.941 39.491			H_δ 7.000 7.580 N_δ 112.397 C'_γ 175.744	178.200
E95	8.606 122.549	4.100 59.450	2.107 2.010 29.991	2.230 2.420 36.200			180.100
C96	8.548 115.217	4.647 56.200	2.420 2.638 37.333				177.200
I97	7.915 123.316	3.909 65.183	2.356 39.000	1.540 2.040 0.958 28.514 16.709	0.728 12.237		178.300
E98	7.467 117.632	4.050 58.580	2.200* 28.990	2.400* 35.600			178.700
N99	7.684 116.119	4.532 55.189	2.832* 38.836			H_δ 7.104 7.419 N_δ 112.318 C'_γ 175.86	175.500
A100	7.528 119.344	4.385 51.986	1.536 19.963				176.400
K101	7.326 118.560	4.158 58.188	1.924* 32.600	1.590* 24.718	1.760 29.300	H_ϵ 2.960 3.030 C_ϵ 41.800	177.300
G102	8.705 111.297	3.766 4.165 44.961					174.400
E103	7.508 119.917	4.411 55.345	2.181 2.085 30.780	2.171* 37.011			176.200
T104	8.233 112.011	4.449 61.808	4.376 69.571	1.191 21.600			174.000
D105	8.422 123.606	4.846 53.000	2.932 2.609 42.930			C'_γ 179.7	177.200
E106	9.183 126.145	3.870 59.000	2.020* 29.049	2.200 2.333 35.600			177.110
C107	8.363 117.959	4.219 60.225	3.781 2.817 44.423				175.860
N108	7.822 118.074	4.657 55.215	2.795 3.056 37.526			H_δ 7.003 7.676 N_δ 112.196 C'_γ 175.5	178.420
I109	8.201 122.139	3.426 67.045	1.973 37.707	0.916 1.715 0.822 30.124 17.633	0.732 13.564		177.400
G110	8.109 104.848	3.872 3.872 47.602					175.920

Résidu	amide	α	β	γ	δ	autre	C'
N111	7.519 119.853	4.525 55.838	2.838* 38.836			H_δ 7.836 7.127 N_δ 112.263 C'_γ 174.4	176.160
K112	8.795 119.321	4.070 59.400	1.760 1.920 33.000	1.500 1.620 25.797	1.740 29.300	H_ϵ 3.000 C_ϵ 41.800	180.180
Y113	8.851 118.083	3.950 63.288	2.818* 39.500		H_δ 6.720 C_δ 130.2	H_ϵ 6.580 C_ϵ 115.6	176.700
T114	7.436 111.944	3.658 66.938	4.314 68.140	1.311 22.361			175.620
D115	8.543 120.965	4.316 57.356	2.545 2.684 39.780				178.160
C116	6.995 118.462	4.141 59.242	3.255 3.062 36.328				173.950
Y117	7.955 123.681	3.513 60.757	3.034 2.326 39.986		H_δ 6.590 C_δ 130.2	H_ϵ 6.430 C_ϵ 114.4	177.350
I118	8.446 115.143	3.654 64.186	2.024 37.000	1.350 1.760 0.950 X 17.000	0.800 11.000		178.060
E119	7.858 119.698	4.020 59.400	2.080 2.140 29.000	2.330* 36.400			179.450
K120	7.798 116.148	3.997 58.733	1.582 1.754 33.385	1.620* 26.588	1.579* 28.822	H_ϵ 2.829 2.900 C_ϵ 42.200	178.770
L121	8.146 120.540	3.763 57.281	1.078 1.309 42.717	1.110 26.200	0.660 0.725 26.000 25.000		177.530
F122	7.428 113.609	4.890 56.403	3.486 2.937 38.575		H_δ 7.410 C_δ 129.2	$H_{\epsilon\zeta}$ 7.170 7.130 $C_{\epsilon\zeta}$ 128.050 128.050	175.400
S123	7.351 120.934	4.345 60.722	3.973* 64.826				178.760

Annexe C

Résultats de l'expérience HNHA et de la prédiction par Talos

Constantes de couplage ${}^3J_{H_N H_\alpha}$ - Valeurs des angles ϕ et ψ

Légende : **i.c.** Impossible de conclure (corrélation absente, superposition des amides,...), **super.** Superposition de 2 ou plusieurs résidus, **Gly** Résidu glycine, **Pro** Résidu proline

Résidu	J_{corrig} (Hz)	Commentaire	Structure secondaire	Angles	ϕ (Talos)	ψ (Talos)
I1			Structure étendue	i.c.		
D2		absent	Structure étendue	i.c.	-85±20	113±15
Q3	3,26		Hélice α	-60±25	-62±11	-32±14
D4	5,439	$J_{mesur}=7.6\text{Hz}$	Hélice α	-80±25	-63±4	-37±4
T5	6,43		Hélice α	-80±25	-68±3	-41±6
V6	3,68		Hélice α	-60±25	-66±8	-43±6
V7	>2,5		Hélice α	i.c.	-67±7	-42±8
A8		absent	Hélice α	i.c.	-61±5	-40±5
K9	4,71		Hélice α	-60±25	-68±2	-39±3
Y10	2,71		Hélice α	-60±25	-67±5	-38±9
M11	<4,71		Hélice α	-60±25	-65±5	-42±8
E12	2,9		Hélice α	-60±25	-64±7	-44±9
Y13	>3,5		Hélice α	-80±25	-70±15	-30±9
L14		super. 79	Hélice α	i.c.		
M15	<2,04		Hélice α	-60±25		
P16			?	i.c.		
D17	7,63		?	i.c.		
I18	>2,2		Hélice α	i.c.	-65±12	-40±13
M19	>3,37		Hélice α	i.c.	-62±6	-43±6
P20			Hélice α	i.c.	-78±20	-25±21
C21	5,35		Hélice α	-60±25	-68±4	-36±8
A22		super. 22, 29 et 69	Hélice α	i.c.	-69±8	-32±15
D23	3,3		Hélice α	-60±25	-63±9	-37±6
E24		super. 32	Hélice α	i.c.	-65±7	-32±11
L25	7,21		Hélice α	-80±25	-96±19	4±17
H26	8,43		Structure étendue	-120±55		
I27	9,6		Structure étendue	-120±55		
S28	3,8		Hélice α	-60±25		
E29		super. 22, 29 et 69	Hélice α	i.c.	-65±8	-34±11
D30		absent	Hélice α	i.c.	-71±15	-33±15
I31		absent	Hélice α	i.c.	-72±10	-32±10
A32		super. 24	Hélice α	i.c.	-78±13	-29±15

Résidu	$J_{corr}(Hz)$	Commentaire	Structure secondaire	Angles	ϕ (Talos)	ψ (Talos)
T33		absent	?	i.c.		
N34		absent	?	i.c.		
I35		absent	?	i.c.		
Q36		absent	Hélice α	i.c.	-88 \pm 17	-16 \pm 33
A37	4,4		Hélice α	-60 \pm 25	-74 \pm 20 ¹	-32 \pm 21 ¹
A38		absent	?	i.c.	-76 \pm 14 ¹	-26 \pm 14 ¹
K39		super. 47	Hélice α	i.c.	-75 \pm 14 ¹	-21 \pm 19 ¹
N40		absent	Hélice α	i.c.	-97 \pm 14	3 \pm 14
G41		<Jcorr>=7,73	?	Gly	83 \pm 12	9 \pm 12
A42	4,27		?	-60 \pm 25		
D43	7,06		Structure étendue	i.c.		
M44	>4,6		Structure étendue	i.c.		
S45		absent	Hélice α	i.c.	-69 \pm 9	-35 \pm 11
Q46		absent	Hélice α	i.c.	-67 \pm 10	-39 \pm 13
L47		super. 39	Hélice α	i.c.	-65 \pm 12	-39 \pm 17
G48			Hélice α	Gly		
C49		absent	Hélice α	i.c.		
L50	4,3		Hélice α	-60 \pm 25	-66 \pm 8	-36 \pm 15
K51	>2,48		Hélice α	i.c.	-64 \pm 5	-40 \pm 6
A52		super. 10	Hélice α	i.c.	-73 \pm 16	-28 \pm 18
C53	3,3		Hélice α	-60 \pm 25	-69 \pm 6	-35 \pm 12
V54	4,5		Hélice α	-60 \pm 25	-67 \pm 4	-39 \pm 10
M55	>1,46		Hélice α	i.c.	-60 \pm 4	-45 \pm 8
K56	>2,9		Hélice α	i.c.	-68 \pm 6	-36 \pm 6
R57	>2,66		Hélice α	i.c.	-66 \pm 8	-31 \pm 11
I58	9,9		?	i.c.	-93 \pm 6	0 \pm 17
E59	6,5		Structure étendue	i.c.		
M60	10,3		?	-120 \pm 55		
L61	11,655		Feuillet β	-115 \pm 60	-131 \pm 18	128 \pm 14
K62	9,8124		Feuillet β	-120 \pm 25	-111 \pm 13	120 \pm 21
G63			Structure étendue	Gly		
T64	10		?Structure étendue	-120 \pm 25		
E65	10,2		Feuillet β	-120 \pm 25	-112 \pm 26	142 \pm 25
L66	9,5		Feuillet β	-120 \pm 35	-118 \pm 25	142 \pm 16
Y67	11,4		Feuillet β	-120 \pm 55	-101 \pm 24	147 \pm 23
V68	6,44		Structure étendue	-120 \pm 35		
E69		super. 22, 29 et 69	Structure étendue	i.c.	-68 \pm 8	-41 \pm 7
P70			Structure étendue	Pro	-61 \pm 6	-34 \pm 5
V71	6,3		Hélice α	-70 \pm 25	-69 \pm 14	-40 \pm 11
Y72	3,3		Hélice α	-60 \pm 25	-65 \pm 5	-43 \pm 9
K73	3,4		Hélice α	-60 \pm 25	-65 \pm 7	-43 \pm 6
M74	4,6		Hélice α	-60 \pm 25	-66 \pm 6	-43 \pm 8
I75	3,6		Hélice α	-60 \pm 25	-69 \pm 6	-42 \pm 6
E76	4,4		Hélice α	-60 \pm 25	-62 \pm 6	-39 \pm 8
V77	7,7		Hélice α	-80 \pm 25	-68 \pm 6	-45 \pm 5
V78		super. 89	Hélice α	i.c.	-67 \pm 10	-31 \pm 11
H79		super. 14	Hélice α	i.c.		
A80	1,5		Structure étendue	i.c.		
G81			Structure étendue	Gly		
N82		absent	Structure étendue	i.c.	-98 \pm 15	133 \pm 23
A83		absent	Hélice α	i.c.	-61 \pm 8	-35 \pm 11
D84	5,85		Hélice α	i.c.	-64 \pm 7	-37 \pm 7
D85	5,7		Hélice α	-70 \pm 25	-67 \pm 5	-41 \pm 8

Résidu	J_{corrig} (Hz)	Commentaire	Structure secondaire	Angles	ϕ (Talos)	ψ (Talos)
I86	>3,35		Hélice α	i.c.	-67±8	-43±7
Q87	3,9		Hélice α	-60±25	-61±5	-40±5
L88	3,5		Hélice α	-60±25	-63±6	-43±6
V89		super. 78	Hélice α	i.c.	-67±9	-42±8
K90	3,1		Hélice α	-60±25	-62±6	-36±11
G91			Hélice α	Gly	-67±4	-39±6
I92	4,1		Hélice α	-60±25	-65±8	-41±6
A93	2,5		Hélice α	-60±25	-65±4	-36±4
N94	5		Hélice α	-60±25	-64±6	-41±7
E95	3,9		Hélice α	-60±25	-62±5	-39±3
C96	4,4		Hélice α	-60±25	-66±2	-37±8
I97	4,02		Hélice α	-60±25	-67±6	-29±10
E98	4,1		Hélice α	-60±25	-68±13	-33±10
N99		absent	Hélice α	i.c.		
A100	9,4		Structure étendue	i.c.		
K101	3,1		Structure étendue	i.c.		
G102			Structure étendue	Gly		
E103		super. 111	Structure étendue	-120±35	-103±21	139±21
T104	10,6		Structure étendue	-120±55		
D105	9,6		Structure étendue	-120±35	-79±16	133±30
E106	2,3		Hélice α	-60±25		
C107	7,7		Hélice α	-80±25		
N108	>4,82		Hélice α	i.c.	-68±3	-36±14
I109	>3,08	super. 57	Hélice α	i.c.	-66±6	-40±6
G110			Hélice α	Gly	-66±5	-36±6
N111		super. 103	Hélice α	i.c.	-69±7	-33±14
K112	4,1		Hélice α	-60±25	-67±3	-39±5
Y113	4,8		Hélice α	-60±25	-66±5	-44±7
T114		absent	Hélice α	i.c.	-69±11	-39±9
D115		absent	Hélice α	i.c.	-76±20	-28±22
C116	5,1		Hélice α	-60±25	-70±5	-42±10
Y117		absent	Hélice α	i.c.	-65±5	-38±10
I118	6		Hélice α	-70±25	-65±6	-37±11
E119	3,47		Hélice α	-60±25	-64±7	-39±10
K120	5		Hélice α	-60±25	-66±8	-38±5
L121	3,7		Hélice α	-60±25	-74±13	-31±8
F122	8,6		Structure étendue	-120±55		
S123	8,4		Structure étendue	-120±55		

Annexe D

Protocole de préparation des bicelles ester et éther

Le protocole de préparation des bicelles que j'ai utilisé a été proposé par Raffard.

Les masses des phospholipides à chaîne longue et à chaîne courte sont pesées précisément dans des eppendorfs afin de respecter le rapport entre le nombre de moles de ces deux phospholipides. Pour 1ml de solution de bicelles (5%, $q=3$), la masse de phospholipides à chaîne longue est de 41mg, la masse de phospholipides à chaîne courte est de 9mg (pour les bicelles à liaison ester ou éther).

Le tampon de solubilisation est tout d'abord utilisé pour dissoudre les phospholipides à chaîne courte puis à chaîne longue. Cette étape de solubilisation peut s'avérer longue et nécessite la répétition d'un cycle de vortex, congélation, chauffage et centrifugation décrit ci-dessous.

La solution de bicelles est agitée vigoureusement. Elle est ensuite placée à $-80^{\circ}C$ pendant 15mn, puis dans un bain chauffant à $50^{\circ}C$ pendant 10mn et centrifugée à 8000 tours par minute pendant 10mn à $18^{\circ}C$.

Lorsque les phospholipides ont été solubilisés, la solution est soumise à un autre cycle permettant son homogénéisation. La solution est placée à $4^{\circ}C$ pendant 15mn puis chauffée à $40^{\circ}C$ pendant 15mn. Ce cycle est répété trois fois et doit être effectué avant la préparation d'un échantillon de RMN.

Lorsque les bicelles sont formées à $37^{\circ}C$, l'échantillon doit être très visqueux et translucide. En dessous de $25^{\circ}C$, le milieu est isotrope, translucide et présente une viscosité similaire à celle de l'eau.

Mesure des couplages dipolaires résiduels dans le milieu DHPC/DMPC-éther sur les formes A et B d'ASP2

Résidu	D_{HN}	D_{HN}	Commentaire	Résidu	D_{HN}	D_{HN}	Commentaire
3	-4,5	-5,4	dédoublé	75		-0,5	
4	1,8	1,4	dédoublé	76		3,8	
5		14,4		77		4,8	
6		-7,6		78		0,4	
8		-0,5		79		-0,6	dédoublé
11		-6,0		80		9,0	
12		0,4		81		-0,2	
13		-5,7		82		8,9	
14	-2,2	0,2	dédoublé	83		-2,4	
15		5,5		85		-5,5	
17		0,7		87		5,1	
19		-8,3		89		-1,3	
21		9,5		91		3,5	
22	1,9	4,4	dédoublé	92		1,0	
25		8,5		93		1,0	
33		5,0		94		7,8	
35		1,8		95		2,0	dédoublé
41		-4,5		96		1,6	dédoublé
44		4,0	dédoublé	98		2,8	
48		-0,8		99		1,7	
50		-4,9	dédoublé	100		5,4	
53		0,9	dédoublé	101		5,3	dédoublé
54		-6		102		-5,2	dédoublé
56		3,7		103		3,5	
58	-9,0	-7,4	dédoublé	104		-7,7	
59		0,2		105	-5,2	-6,4	dédoublé
60	8,0	8,5	dédoublé	106		3,4	
61	-3,6	-1,6	dédoublé	107		3,5	
62		-2,8	dédoublé	110	8,8	6,4	dédoublé
63		-0,5		111	3,3	1,0	dédoublé
64		0,8		112		3,0	dédoublé
65		-4,7		113		5,0	dédoublé
66		-3,8		115		5,0	dédoublé
67	7,0	6,0	dédoublé	116		5,8	
68		0,9		117		11,9	
69		-1,3		118		9,4	
71		-1,4		120		-2,2	
72		-0,7		121		4,9	
73		2,4		122	5,5	6,3	dédoublé
74		-0,1		123	-4,9	-0,9	dédoublé

Annexe E

Programmes Octave d'analyse des couplages dipolaires et de l'Onde Dipolaire

Algorithme de recherche par grille sur θ et ϕ (solasp2.m)

```
load asp2h4moins; rdcfilecomplet=asp2h4moins
clear Ki2 theta phi itheta iphi minKi2 nmin
disp("Recherche sur theta et phi pour le fit des ondes dipolaires");
itheta=input('entrer le pas de recherche pour theta en degre : ');
iphi=input('entrer le pas de recherche pour phi en degre : ');
iph=input('entrer le pas de recherche pour ph en degre : ');
Da=input('valeur de Da : '); R=input('valeur de R : ');
disp("L'affichage de la courbe necessite deux nombres n1 et n2 correspondant au premier residu (n1=1) et au dernier residu (n2=10)");
n1=input('n1 : '); n2=input('n2 : '); Ki2=0*ones((180/itheta+1)*(180/iphi), 9); iDa=1; iR=0.1;
xb=[Da-iDa Da+iDa; R-iR R+iR; -0.1 180; -0.1 360; -0.1 360];
icb=[-10000000 10000000;-1000000 1000000;-1000000 10000000;-1000000 10000000]; n=1;
for theta=0 :itheta :180
    for phi=0 :iphi :(360-iphi)
        for ph=0 :iph :(360-iph)
            x0=[Da R theta phi ph]';
            [a,oh,y]=solnp([x0,xb]);
            a=a';
            Ki2(n, :)=a(3) a(4) a(5) cost(a) a(1) a(2) theta phi ph ];
            n++;
        end end end
        %%%%%%%%%%%%%%% Analyse de la grille %%%%%%%%%%%%%%%
        [minKi2 nmin]=min(Ki2( :,4)); best=Ki2(nmin, :);
        disp("Les angles donnant le meilleur fit (Ki2 minimal, variable best) sont [theta phi ph minKi2 Da
R theta_depart phi_depart ph_depart] : "), disp (Ki2(nmin, :)); phicorr=Ki2(nmin,2); Rcorr=Ki2(nmin,6);
        if Rcorr<0
            Rcorr=-Rcorr; phicorr=Ki2(nmin,2)+90;
        end
        if phicorr > 355
            while ( phicorr > 355 )
                phicorr=phicorr-360;
            end
        end
    end
end
```



```

end
if phicorr < 0
    while ( phicorr <=0 )
        phicorr=phicorr+360;
    end
end
bestcorr=best ; bestcorr(2)=phicorr ;bestcorr(6)=Rcorr ; disp("Valeurs corriges (bestcorr :)", disp(bestcorr) ;
Ki2(nmin,2)=phicorr ;Ki2(nmin,6)=Rcorr ;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% affichage du resultat %%%%%%%%%%%%%%
figure(1) ; clg ; hold off ;
errorbar(rdcfilecomplet( :,1),rdcfilecomplet( :,2),rdcfilecomplet( :,3),"~ ;erreur ;") ;
hold on ;
p=[Ki2(nmin,5) Ki2(nmin,6) Ki2(nmin,1) Ki2(nmin,2) Ki2(nmin,3) ] ;
disp('Trace de la courbe rdc_local avec les valeurs du fit a Da et R constant') ;
plot([n1 :0.1 :n2],rdc([n1 :0.1 :n2],p ),'-g;trace du fit a Da et R constant ; ') ;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% affichage 3D de la grille %%%%%%%%%%%%%%
figure(2) ; clg ;
x=[0 :itheta :180]' ; y=[0 :iphi :360-iphi]' ; lmax=length(Ki2) ; z=0.*ones(360/iphi,180/itheta+1) ;
for l=1 :lmax
    z((Ki2(l,7))/iphi+1, Ki2(l,6)/itheta+1)=Ki2(l,4) ;
end
mesh (x, y, z) ; replot ;

```

Définition de la fonction χ^2 (cost.m)

```

function [f]=cost(x,par)
%calcul de Ki2
load asp2h4moins ;
rdcfile=asp2h4moins ;
v=length(rdcfile)-length(x) ;
f(1)=sum(((rdc(rdcfile( :,1),x)-rdcfile( :,2)).^2)/v) ;
return

```

Exemple de fichier d'entrée (asp2h4moins)

```

Format : numero du résidu dans l'hélice RDC erreur
1 -1.30 2 %res.num=69
3 -1.40 2 %res.num=71
4 -0.68 2.5 %res.num=72
5 2.42 2.5 %res.num=73
6 -0.05 2 %res.num=74
7 -0.50 2 %res.num=75
8 3.80 2 %res.num=76
9 4.80 2 %res.num=77
10 0.42 2.5 %res.num=78

```

Annexe F

Données de relaxation et de dynamique à 308K et 283K

Les temps de relaxation R_1 et R_2 sont en Hz, les grandeurs R_1/R_2 et noe sont sans unité.

Rés.	R_1	R_2	à 308K		à 283K			
			R_1/R_2	noe	R_1	R_2	R_1/R_2	noe
2	1,97 ± 0,29	4,06 ± 0,67	0,49 ± 0,15	0,466 ± 0,015	1,33 ± 0,24	8,61 ± 0,82	0,154 ± 0,043	0,497 ± 0,098
3	1,97 ± 0,23	5,81 ± 1,05	0,34 ± 0,10	0,604 ± 0,009	1,23 ± 0,21	8,87 ± 1,26	0,139 ± 0,044	0,681 ± 0,093
4	1,96 ± 0,16	5,44 ± 0,39	0,36 ± 0,06	0,642 ± 0,009	1,31 ± 0,19	9,48 ± 0,83	0,138 ± 0,032	0,652 ± 0,069
5	1,97 ± 0,15	5,25 ± 0,69	0,38 ± 0,08	0,652 ± 0,013	1,26 ± 0,24	12,82 ± 1,13	0,098 ± 0,028	0,733 ± 0,159
6	2,03 ± 0,11	5,92 ± 0,52	0,34 ± 0,05	0,680 ± 0,012				
8	2,06 ± 0,11	6,70 ± 0,28	0,31 ± 0,03	0,822 ± 0,014	1,19 ± 0,22	12,54 ± 0,79	0,095 ± 0,024	0,704 ± 0,161
9	2,06 ± 0,14	6,89 ± 0,26	0,30 ± 0,03	0,778 ± 0,013	1,18 ± 0,15	10,96 ± 1,24	0,107 ± 0,025	0,790 ± 0,118
10	2,02 ± 0,11	6,77 ± 0,44	0,30 ± 0,04	0,772 ± 0,012				
11					1,23 ± 0,21	11,01 ± 1,15	0,112 ± 0,031	0,747 ± 0,142
12	2,07 ± 0,12	7,06 ± 0,43	0,29 ± 0,03	0,763 ± 0,015				
13	2,02 ± 0,13	6,54 ± 0,34	0,31 ± 0,04	0,701 ± 0,013	1,16 ± 0,16	12,03 ± 0,76	0,096 ± 0,019	0,887 ± 0,189
14	2,07 ± 0,12	6,38 ± 0,49	0,32 ± 0,04	0,797 ± 0,022				
15	2,01 ± 0,14	5,96 ± 0,73	0,34 ± 0,06	0,802 ± 0,020				
17	2,10 ± 0,20	6,86 ± 0,43	0,31 ± 0,05	0,771 ± 0,019	1,08 ± 0,19	11,33 ± 1,19	0,095 ± 0,027	0,651 ± 0,140
19	2,08 ± 0,14	7,26 ± 0,40	0,29 ± 0,04	0,777 ± 0,019	1,09 ± 0,18	11,42 ± 1,43	0,095 ± 0,028	0,830 ± 0,183
21	1,96 ± 0,16	7,22 ± 0,36	0,27 ± 0,04	0,908 ± 0,023	1,22 ± 0,27	10,83 ± 1,17	0,112 ± 0,037	0,742 ± 0,198
22					1,15 ± 0,14	11,37 ± 1,35	0,102 ± 0,024	0,879 ± 0,140
23					1,21 ± 0,18	10,41 ± 0,90	0,116 ± 0,027	0,674 ± 0,167
24	1,89 ± 0,14	6,84 ± 0,34	0,28 ± 0,03	0,684 ± 0,012	1,13 ± 0,20	11,46 ± 0,86	0,098 ± 0,025	0,769 ± 0,137
25	1,96 ± 0,15	6,86 ± 0,38	0,29 ± 0,04	0,819 ± 0,016				
27					1,26 ± 0,22	10,81 ± 1,11	0,117 ± 0,032	0,744 ± 0,184
28	1,85 ± 0,16	5,06 ± 0,51	0,37 ± 0,07	0,639 ± 0,012				
29	2,03 ± 0,14	6,40 ± 0,44	0,32 ± 0,04	0,727 ± 0,024				
30	2,00 ± 0,13	6,85 ± 0,44	0,29 ± 0,04	0,696 ± 0,016	1,28 ± 0,20	12,14 ± 0,93	0,105 ± 0,025	0,638 ± 0,163
31	1,85 ± 0,18	6,91 ± 0,31	0,27 ± 0,04	0,729 ± 0,029				
32	1,95 ± 0,17	7,04 ± 0,29	0,28 ± 0,04	0,690 ± 0,012				
33	2,02 ± 0,21	11,95 ± 0,77	0,17 ± 0,03	0,631 ± 0,031	1,04 ± 0,24	14,99 ± 1,42	0,030 ± 0,014	0,750 ± 0,480
34	1,85 ± 0,19	6,13 ± 0,30	0,30 ± 0,05	0,662 ± 0,038	1,27 ± 0,26	10,98 ± 2,40	0,116 ± 0,051	0,807 ± 0,364
35	1,94 ± 0,15	6,72 ± 0,45	0,29 ± 0,04	0,622 ± 0,021	1,37 ± 0,21	9,10 ± 0,68	0,150 ± 0,034	0,412 ± 0,306
36	1,98 ± 0,14	5,79 ± 0,43	0,34 ± 0,05	0,682 ± 0,018	1,26 ± 0,17	9,74 ± 0,91	0,129 ± 0,029	0,576 ± 0,195
37	2,05 ± 0,14	5,09 ± 0,53	0,40 ± 0,07	0,574 ± 0,017	1,30 ± 0,20	8,99 ± 1,07	0,145 ± 0,039	0,661 ± 0,172
38	1,89 ± 0,17	5,38 ± 0,50	0,35 ± 0,06	0,501 ± 0,018	1,47 ± 0,21	10,40 ± 1,51	0,141 ± 0,040	0,516 ± 0,201
41	1,92 ± 0,28	4,07 ± 0,78	0,47 ± 0,16	0,380 ± 0,018	1,36 ± 0,17	7,41 ± 0,64	0,184 ± 0,039	0,504 ± 0,154
42	1,98 ± 0,19	4,58 ± 0,46	0,43 ± 0,09	0,549 ± 0,011	1,32 ± 0,23	8,81 ± 0,87	0,150 ± 0,040	0,631 ± 0,109
43	1,75 ± 0,21	4,08 ± 0,45	0,43 ± 0,10	0,429 ± 0,010	1,34 ± 0,15	4,93 ± 0,62	0,271 ± 0,065	0,278 ± 0,061
44	1,88 ± 0,18	4,66 ± 0,52	0,40 ± 0,08	0,449 ± 0,015				
45	2,17 ± 0,27	5,96 ± 0,31	0,36 ± 0,06	0,615 ± 0,013	1,32 ± 0,21	11,36 ± 1,01	0,116 ± 0,029	0,804 ± 0,152
48	2,14 ± 0,10	10,00 ± 0,68	0,21 ± 0,03	0,720 ± 0,020				
49	1,99 ± 0,19	8,47 ± 1,06	0,23 ± 0,05	0,778 ± 0,033				
50	2,23 ± 0,18	7,24 ± 0,49	0,31 ± 0,05	0,877 ± 0,019	1,28 ± 0,21	14,45 ± 2,21	0,088 ± 0,028	0,780 ± 0,226
51	1,93 ± 0,19	5,97 ± 0,56	0,32 ± 0,06	0,641 ± 0,014				
52	2,09 ± 0,12	6,80 ± 0,31	0,31 ± 0,03	0,760 ± 0,014	1,29 ± 0,21	11,27 ± 0,77	0,114 ± 0,026	0,719 ± 0,133
53	2,06 ± 0,15	7,02 ± 0,36	0,29 ± 0,04	0,783 ± 0,015	1,17 ± 0,18	10,91 ± 0,96	0,107 ± 0,026	0,909 ± 0,223
54	2,13 ± 0,11	6,93 ± 0,43	0,31 ± 0,03	0,782 ± 0,015				
55	2,06 ± 0,16	6,88 ± 0,47	0,30 ± 0,04	0,894 ± 0,024	1,26 ± 0,22	11,64 ± 0,87	0,108 ± 0,027	0,760 ± 0,268
57					1,16 ± 0,23	10,54 ± 0,93	0,110 ± 0,031	0,773 ± 0,105
58	2,06 ± 0,13	7,36 ± 0,42	0,28 ± 0,03	0,795 ± 0,016	1,18 ± 0,24	10,32 ± 0,99	0,115 ± 0,034	0,840 ± 0,149
59	1,89 ± 0,15	4,86 ± 1,30	0,39 ± 0,14	0,544 ± 0,009				
60	2,04 ± 0,12	7,95 ± 0,31	0,26 ± 0,02	0,767 ± 0,018	1,14 ± 0,21	13,18 ± 0,78	0,086 ± 0,021	0,760 ± 0,225

Rés.	R_1	R_2	R_1/R_2	noe	R_1	R_2	R_1/R_2	noe
61	1,97 ± 0,10	6,97 ± 0,34	0,28 ± 0,03	0,739 ± 0,020	1,14 ± 0,27	13,17 ± 1,17	0,086 ± 0,028	0,733 ± 0,202
62					1,14 ± 0,24	7,61 ± 0,94	0,149 ± 0,051	0,785 ± 0,095
63	2,09 ± 0,30	6,78 ± 0,34	0,31 ± 0,06	0,765 ± 0,011	1,20 ± 0,19	11,55 ± 1,03	0,104 ± 0,026	0,708 ± 0,094
64					1,20 ± 0,25	11,88 ± 1,47	0,101 ± 0,034	0,691 ± 0,120
65					1,17 ± 0,18	11,31 ± 1,09	0,103 ± 0,026	0,806 ± 0,092
66	1,84 ± 0,11	6,25 ± 0,33	0,29 ± 0,03	0,735 ± 0,011				
67	1,88 ± 0,16	6,43 ± 0,39	0,29 ± 0,04	0,774 ± 0,015	1,13 ± 0,23	10,24 ± 0,43	0,110 ± 0,027	0,743 ± 0,154
68	1,92 ± 0,14	6,17 ± 0,39	0,31 ± 0,04	0,667 ± 0,009				
69					1,18 ± 0,18	11,28 ± 1,15	0,104 ± 0,027	0,831 ± 0,094
71	2,02 ± 0,18	7,26 ± 0,30	0,28 ± 0,04	0,765 ± 0,011	1,19 ± 0,20	11,56 ± 0,85	0,103 ± 0,025	0,726 ± 0,143
72	2,00 ± 0,16	6,35 ± 0,36	0,32 ± 0,04	0,781 ± 0,010	1,24 ± 0,18	10,81 ± 0,95	0,115 ± 0,027	0,678 ± 0,086
73	2,07 ± 0,13	7,09 ± 0,30	0,29 ± 0,03	0,782 ± 0,011	1,17 ± 0,19	12,03 ± 0,83	0,097 ± 0,023	0,779 ± 0,110
74	2,07 ± 0,17	6,90 ± 0,35	0,30 ± 0,04	0,797 ± 0,012	1,14 ± 0,20	11,36 ± 1,27	0,101 ± 0,029	0,742 ± 0,128
75	2,09 ± 0,15	7,27 ± 0,34	0,29 ± 0,03	0,780 ± 0,012	1,20 ± 0,26	11,85 ± 0,94	0,102 ± 0,030	0,830 ± 0,159
76	2,17 ± 0,17	7,63 ± 0,32	0,28 ± 0,03	0,793 ± 0,014	1,16 ± 0,24	12,18 ± 1,04	0,095 ± 0,028	0,707 ± 0,161
77	1,94 ± 0,12	7,03 ± 0,30	0,28 ± 0,03	0,679 ± 0,012	1,10 ± 0,26	11,70 ± 1,11	0,094 ± 0,031	0,865 ± 0,139
78					1,22 ± 0,19	11,01 ± 1,27	0,111 ± 0,030	0,739 ± 0,112
79	2,07 ± 0,12	6,93 ± 0,43	0,30 ± 0,04	0,751 ± 0,013				
80	2,01 ± 0,21	6,94 ± 0,42	0,29 ± 0,05	0,804 ± 0,011	1,12 ± 0,20	11,54 ± 1,49	0,097 ± 0,030	0,770 ± 0,109
81	1,98 ± 0,22	6,14 ± 0,41	0,32 ± 0,06	0,759 ± 0,011	1,11 ± 0,22	10,23 ± 0,72	0,108 ± 0,029	0,779 ± 0,098
82	1,91 ± 0,18	6,59 ± 0,33	0,29 ± 0,04	0,770 ± 0,013	1,06 ± 0,24	11,10 ± 0,88	0,096 ± 0,029	0,738 ± 0,117
83					1,11 ± 0,23	12,19 ± 1,10	0,091 ± 0,027	0,759 ± 0,081
84	2,01 ± 0,14	6,61 ± 0,41	0,30 ± 0,04	0,794 ± 0,008	1,15 ± 0,15	10,53 ± 0,97	0,110 ± 0,025	0,818 ± 0,069
85	1,98 ± 0,13	6,48 ± 0,32	0,31 ± 0,04	0,782 ± 0,008	1,16 ± 0,24	11,21 ± 0,85	0,103 ± 0,029	0,782 ± 0,083
86					1,17 ± 0,18	11,56 ± 0,85	0,101 ± 0,023	0,835 ± 0,131
87	2,06 ± 0,14	7,19 ± 0,29	0,29 ± 0,03	0,797 ± 0,009	1,14 ± 0,18	11,51 ± 0,92	0,099 ± 0,023	0,845 ± 0,078
88	2,08 ± 0,11	6,96 ± 0,38	0,30 ± 0,03	0,771 ± 0,010	1,15 ± 0,19	11,30 ± 1,04	0,102 ± 0,026	0,802 ± 0,074
89	2,12 ± 0,16	7,27 ± 0,42	0,29 ± 0,04	0,789 ± 0,012	1,20 ± 0,21	12,50 ± 0,98	0,096 ± 0,024	0,897 ± 0,136
90	1,96 ± 0,16	6,47 ± 0,58	0,30 ± 0,05	0,730 ± 0,006				
91	2,08 ± 0,15	7,60 ± 0,38	0,27 ± 0,03	0,799 ± 0,014	1,18 ± 0,16	11,66 ± 1,16	0,101 ± 0,024	0,887 ± 0,133
92	2,06 ± 0,12	6,80 ± 0,37	0,30 ± 0,03	0,760 ± 0,015	1,12 ± 0,22	12,60 ± 0,86	0,089 ± 0,023	0,790 ± 0,156
93	2,06 ± 0,13	7,49 ± 0,37	0,28 ± 0,03	0,799 ± 0,011	1,20 ± 0,20	12,12 ± 0,95	0,099 ± 0,024	0,937 ± 0,145
94	2,00 ± 0,15	6,32 ± 0,75	0,32 ± 0,06	0,734 ± 0,013	1,15 ± 0,20	11,75 ± 1,25	0,098 ± 0,027	0,802 ± 0,118
95	2,04 ± 0,13	7,01 ± 0,40	0,29 ± 0,04	0,823 ± 0,012	1,20 ± 0,19	11,34 ± 1,03	0,106 ± 0,027	0,877 ± 0,122
96	2,03 ± 0,12	7,72 ± 0,37	0,26 ± 0,03	0,738 ± 0,013				
97					1,17 ± 0,22	11,24 ± 1,05	0,104 ± 0,029	0,895 ± 0,181
98	2,08 ± 0,13	6,92 ± 0,48	0,30 ± 0,04	0,802 ± 0,017	1,19 ± 0,23	11,78 ± 0,90	0,101 ± 0,027	0,715 ± 0,100
99	1,97 ± 0,11	6,50 ± 0,31	0,30 ± 0,03	0,708 ± 0,010	1,12 ± 0,17	11,98 ± 0,79	0,094 ± 0,020	0,724 ± 0,098
100	1,93 ± 0,16	6,62 ± 0,41	0,29 ± 0,04	0,706 ± 0,013	1,11 ± 0,21	11,55 ± 0,81	0,096 ± 0,025	0,729 ± 0,135
101	1,82 ± 0,14	6,41 ± 0,31	0,28 ± 0,04	0,690 ± 0,009	1,09 ± 0,14	10,69 ± 0,99	0,102 ± 0,023	0,656 ± 0,114
102	1,83 ± 0,15	5,50 ± 0,43	0,33 ± 0,05	0,622 ± 0,013	1,22 ± 0,23	9,87 ± 0,56	0,124 ± 0,030	0,641 ± 0,133
103	1,80 ± 0,20	6,24 ± 0,43	0,29 ± 0,05	0,720 ± 0,009	1,13 ± 0,20	10,78 ± 0,95	0,104 ± 0,028	0,692 ± 0,077
104	1,89 ± 0,26	5,65 ± 0,41	0,33 ± 0,07	0,705 ± 0,013	1,05 ± 0,19	10,08 ± 0,75	0,104 ± 0,026	0,724 ± 0,117
105	1,76 ± 0,18	4,66 ± 0,57	0,38 ± 0,08	0,529 ± 0,012				
106	2,04 ± 0,12	6,60 ± 0,39	0,31 ± 0,04	0,748 ± 0,010	1,15 ± 0,17	10,92 ± 0,86	0,106 ± 0,023	0,812 ± 0,122
107	2,07 ± 0,15	7,96 ± 0,35	0,26 ± 0,03	0,861 ± 0,013	1,21 ± 0,19	12,75 ± 1,09	0,095 ± 0,023	0,849 ± 0,201
108	1,98 ± 0,13	6,90 ± 0,31	0,29 ± 0,03	0,799 ± 0,017				
110	2,11 ± 0,18	7,50 ± 0,34	0,28 ± 0,04	0,759 ± 0,018	1,23 ± 0,23	11,55 ± 1,20	0,106 ± 0,031	0,861 ± 0,324
111	2,06 ± 0,15	7,72 ± 0,38	0,27 ± 0,03	0,808 ± 0,018				
112	2,11 ± 0,17	7,36 ± 0,39	0,29 ± 0,04	0,796 ± 0,016	1,18 ± 0,22	11,16 ± 0,99	0,105 ± 0,029	0,640 ± 0,259
113	2,13 ± 0,14	7,43 ± 0,44	0,29 ± 0,04	0,779 ± 0,017	1,21 ± 0,25	12,45 ± 1,23	0,097 ± 0,029	0,876 ± 0,206
114	2,23 ± 0,24	9,46 ± 0,55	0,24 ± 0,04	0,666 ± 0,024				
115	2,13 ± 0,11	7,74 ± 0,31	0,27 ± 0,03	0,891 ± 0,014	1,26 ± 0,20	10,56 ± 1,11	0,119 ± 0,032	0,739 ± 0,116
116	2,12 ± 0,22	7,05 ± 0,35	0,30 ± 0,05	0,732 ± 0,024	1,16 ± 0,23	11,27 ± 1,30	0,103 ± 0,032	1,088 ± 0,390
118	2,05 ± 0,13	7,54 ± 0,38	0,27 ± 0,03	0,813 ± 0,023				
119	2,09 ± 0,11	7,42 ± 0,29	0,28 ± 0,03	0,841 ± 0,014				
120	2,04 ± 0,13	7,41 ± 0,40	0,28 ± 0,03	0,756 ± 0,018				
121	1,92 ± 0,12	7,39 ± 0,42	0,26 ± 0,03	0,769 ± 0,014	1,22 ± 0,15	11,33 ± 1,00	0,107 ± 0,023	0,739 ± 0,190
122	1,98 ± 0,21	6,40 ± 0,27	0,31 ± 0,05	0,796 ± 0,018	1,10 ± 0,19	10,73 ± 1,29	0,102 ± 0,030	0,706 ± 0,133
123	1,38 ± 0,20	4,04 ± 0,63	0,34 ± 0,10	0,226 ± 0,010	1,17 ± 0,17	7,08 ± 0,70	0,165 ± 0,040	0,314 ± 0,091

Rés.	Mod.	S^2				Mod.	S^2			
		à 308K		à 283K			à 308K		à 283K	
2	2	0,289 ± 0,155	1992 ± 359	0 ± 0	1 ± 0	2	0,660 ± 0,058	42 ± 239	0 ± 0	1 ± 0
3	2	0,784 ± 0,071	71 ± 335	0 ± 0	1 ± 0	1	0,714 ± 0,078	0 ± 0	0 ± 0	1 ± 0
4	2	0,736 ± 0,038	41 ± 9	0 ± 0	1 ± 0	2	0,732 ± 0,055	30 ± 277	0 ± 0	1 ± 0
5	2	0,770 ± 0,052	46 ± 20	0 ± 0	1 ± 0	1	0,934 ± 0,066	0 ± 0	0 ± 0	1 ± 0
6	2	0,819 ± 0,033	49 ± 13	0 ± 0	1 ± 0					

Rés.	Mod.	S^2	τ_e	R_{ex}	S_f^2	Mod.	S^2	τ_e	R_{ex}	S_f^2
8	1	0,871 ± 0,029	0 ± 0	0 ± 0	1 ± 0	1	0,915 ± 0,053	0 ± 0	0 ± 0	1 ± 0
9	1	0,886 ± 0,029	0 ± 0	0 ± 0	1 ± 0	1	0,815 ± 0,070	0 ± 0	0 ± 0	1 ± 0
10	1	0,868 ± 0,035	0 ± 0	0 ± 0	1 ± 0					
11						1	0,825 ± 0,072	0 ± 0	0 ± 0	1 ± 0
12	1	0,899 ± 0,034	0 ± 0	0 ± 0	1 ± 0					
13	2	0,841 ± 0,032	46 ± 16	0 ± 0	1 ± 0	1	0,873 ± 0,052	0 ± 0	0 ± 0	1 ± 0
14	1	0,861 ± 0,042	0 ± 0	0 ± 0	1 ± 0					
15	1	0,836 ± 0,049	0 ± 0	0 ± 0	1 ± 0					
17	1	0,888 ± 0,048	0 ± 0	0 ± 0	1 ± 0	1	0,812 ± 0,072	0 ± 0	0 ± 0	1 ± 0
19	1	0,916 ± 0,038	0 ± 0	0 ± 0	1 ± 0	1	0,809 ± 0,078	0 ± 0	0 ± 0	1 ± 0
21	3	0,844 ± 0,000	0 ± 0	0,66 ± 0	1 ± 0	1	0,809 ± 0,080	0 ± 0	0 ± 0	1 ± 0
22						1	0,822 ± 0,074	0 ± 0	0 ± 0	1 ± 0
23						1	0,787 ± 0,057	0 ± 0	0 ± 0	1 ± 0
24	2	0,843 ± 0,032	56 ± 21	0 ± 0	1 ± 0	1	0,837 ± 0,056	0 ± 0	0 ± 0	1 ± 0
25	1	0,867 ± 0,031	0 ± 0	0 ± 0	1 ± 0					
27						1	0,818 ± 0,073	0 ± 0	0 ± 0	1 ± 0
28	2	0,706 ± 0,048	36 ± 10	0 ± 0	1 ± 0					
29	1	0,846 ± 0,047	0 ± 0	0 ± 0	1 ± 0					
30	2	0,857 ± 0,038	55 ± 31	0 ± 0	1 ± 0	1	0,897 ± 0,059	0 ± 0	0 ± 0	1 ± 0
31	1	0,868 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
32	2	0,881 ± 0,046	75 ± 524	0 ± 0	1 ± 0					
33	4	0,835 ± 0,086	85 ± 634	5,4 ± 0,9	1 ± 0	3	0,722 ± 0,157	0 ± 0	7,24 ± 0,72	1 ± 0
34	2	0,779 ± 0,034	45 ± 19	0 ± 0	1 ± 0	1	0,847 ± 0,111	0 ± 0	0 ± 0	1 ± 0
35	2	0,828 ± 0,048	86 ± 374	0 ± 0	1 ± 0	1	0,703 ± 0,049	0 ± 0	0 ± 0	1 ± 0
36	2	0,779 ± 0,041	37 ± 12	0 ± 0	1 ± 0	1	0,759 ± 0,065	0 ± 0	0 ± 0	1 ± 0
37	5	0,513 ± 0,162	1809 ± 589	0 ± 0	0,87 ± 0,05	1	0,726 ± 0,071	0 ± 0	0 ± 0	1 ± 0
38	2	0,713 ± 0,053	79 ± 174	0 ± 0	1 ± 0	1	0,865 ± 0,082	0 ± 0	0 ± 0	1 ± 0
41	2	0,347 ± 0,192	1553 ± 437	0 ± 0	1 ± 0	5	0,487 ± 0,124	1364 ± 1288	0 ± 0	0,78 ± 0,08
42	5	0,419 ± 0,167	1929 ± 587	0 ± 0	0,83 ± 0,06	1	0,691 ± 0,061	0 ± 0	0 ± 0	1 ± 0
43	5	0,393 ± 0,164	1373 ± 471	0 ± 0	0,78 ± 0,07	5	0,280 ± 0,082	1214 ± 143	0 ± 0	0,72 ± 0,06
44	2	0,657 ± 0,042	72 ± 17	0 ± 0	1 ± 0					
45	2	0,609 ± 0,110	2063 ± 574	0 ± 0	1 ± 0	1	0,857 ± 0,062	0 ± 0	0 ± 0	1 ± 0
48	4	0,903 ± 0,040	64 ± 441	2,9 ± 0,7	1 ± 0					
49	1	0,918 ± 0,065	0 ± 0	0 ± 0	1 ± 0					
50	1	0,942 ± 0,000	0 ± 0	0 ± 0	1 ± 0	1	0,969 ± 0,072	0 ± 0	0 ± 0	1 ± 0
51	2	0,778 ± 0,053	53 ± 27	0 ± 0	1 ± 0					
52	1	0,883 ± 0,031	0 ± 0	0 ± 0	1 ± 0	1	0,842 ± 0,050	0 ± 0	0 ± 0	1 ± 0
53	1	0,897 ± 0,038	0 ± 0	0 ± 0	1 ± 0	1	0,809 ± 0,060	0 ± 0	0 ± 0	1 ± 0
54	1	0,905 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
55	1	0,884 ± 0,000	0 ± 0	0 ± 0	1 ± 0	1	0,864 ± 0,057	0 ± 0	0 ± 0	1 ± 0
57						1	0,784 ± 0,064	0 ± 0	0 ± 0	1 ± 0
58	1	0,915 ± 0,041	0 ± 0	0 ± 0	1 ± 0	1	0,774 ± 0,068	0 ± 0	0 ± 0	1 ± 0
59	2	0,747 ± 0,051	78 ± 171	0 ± 0	1 ± 0					
60	3	0,875 ± 0,050	0 ± 0	1,1 ± 0,5	1 ± 0	1	0,949 ± 0,047	0 ± 0	0 ± 0	1 ± 0
61	2	0,863 ± 0,032	29 ± 17	0 ± 0	1 ± 0	1	0,942 ± 0,062	0 ± 0	0 ± 0	1 ± 0
62						1	0,596 ± 0,061	0 ± 0	0 ± 0	1 ± 0
63	1	0,874 ± 0,045	0 ± 0	0 ± 0	1 ± 0	1	0,850 ± 0,065	0 ± 0	0 ± 0	1 ± 0
64						1	0,868 ± 0,091	0 ± 0	0 ± 0	1 ± 0
65						1	0,830 ± 0,065	0 ± 0	0 ± 0	1 ± 0
66	2	0,792 ± 0,033	19 ± 6	0 ± 0	1 ± 0					
67	1	0,820 ± 0,042	0 ± 0	0 ± 0	1 ± 0	1	0,759 ± 0,032	0 ± 0	0 ± 0	1 ± 0
68	2	0,793 ± 0,034	46 ± 14	0 ± 0	1 ± 0					
69						1	0,830 ± 0,070	0 ± 0	0 ± 0	1 ± 0
71	1	0,920 ± 0,034	0 ± 0	0 ± 0	1 ± 0	1	0,850 ± 0,056	0 ± 0	0 ± 0	1 ± 0
72	1	0,830 ± 0,042	0 ± 0	0 ± 0	1 ± 0	1	0,815 ± 0,059	0 ± 0	0 ± 0	1 ± 0
73	1	0,905 ± 0,034	0 ± 0	0 ± 0	1 ± 0	1	0,876 ± 0,054	0 ± 0	0 ± 0	1 ± 0
74	1	0,887 ± 0,039	0 ± 0	0 ± 0	1 ± 0	1	0,827 ± 0,072	0 ± 0	0 ± 0	1 ± 0
75	1	0,924 ± 0,036	0 ± 0	0 ± 0	1 ± 0	1	0,871 ± 0,065	0 ± 0	0 ± 0	1 ± 0
76	1	0,969 ± 0,029	0 ± 0	0 ± 0	1 ± 0	1	0,885 ± 0,067	0 ± 0	0 ± 0	1 ± 0
77	2	0,866 ± 0,029	73 ± 61	0 ± 0	1 ± 0	1	0,849 ± 0,076	0 ± 0	0 ± 0	1 ± 0
78						1	0,827 ± 0,076	0 ± 0	0 ± 0	1 ± 0
79	2	0,884 ± 0,039	27 ± 373	0 ± 0	1 ± 0					
80	1	0,885 ± 0,045	0 ± 0	0 ± 0	1 ± 0	1	0,827 ± 0,081	0 ± 0	0 ± 0	1 ± 0
81	2	0,800 ± 0,044	11 ± 6	0 ± 0	1 ± 0	1	0,758 ± 0,049	0 ± 0	0 ± 0	1 ± 0
82	1	0,840 ± 0,041	0 ± 0	0 ± 0	1 ± 0	1	0,811 ± 0,065	0 ± 0	0 ± 0	1 ± 0
83						1	0,876 ± 0,066	0 ± 0	0 ± 0	1 ± 0
84	1	0,857 ± 0,037	0 ± 0	0 ± 0	1 ± 0	1	0,787 ± 0,059	0 ± 0	0 ± 0	1 ± 0
85	1	0,838 ± 0,031	0 ± 0	0 ± 0	1 ± 0	1	0,826 ± 0,057	0 ± 0	0 ± 0	1 ± 0
86						1	0,847 ± 0,059	0 ± 0	0 ± 0	1 ± 0

Rés.	Mod.	S^2	τ_e	R_{ex}	S_f^2	Mod.	S^2	τ_e	R_{ex}	S_f^2
87	1	0,913 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
88	1	0,895 ± 0,035	0 ± 0	0 ± 0	1 ± 0	1	0,838 ± 0,058	0 ± 0	0 ± 0	1 ± 0
89	1	0,927 ± 0,042	0 ± 0	0 ± 0	1 ± 0	1	0,827 ± 0,072	0 ± 0	0 ± 0	1 ± 0
90	2	0,829 ± 0,050	27 ± 224	0 ± 0	1 ± 0					
91	1	0,947 ± 0,035	0 ± 0	0 ± 0	1 ± 0	1	0,908 ± 0,061	0 ± 0	0 ± 0	1 ± 0
92	1	0,878 ± 0,037	0 ± 0	0 ± 0	1 ± 0	1	0,847 ± 0,065	0 ± 0	0 ± 0	1 ± 0
93	1	0,930 ± 0,035	0 ± 0	0 ± 0	1 ± 0	1	0,909 ± 0,061	0 ± 0	0 ± 0	1 ± 0
94	2	0,835 ± 0,054	26 ± 35	0 ± 0	1 ± 0	1	0,885 ± 0,055	0 ± 0	0 ± 0	1 ± 0
95	1	0,839 ± 0,065	0 ± 0	0 ± 0	1 ± 0					
96	2	0,931 ± 0,040	67 ± 720	0 ± 0	1 ± 0					
97						1	0,828 ± 0,065	0 ± 0	0 ± 0	1 ± 0
98	1	0,893 ± 0,036	0 ± 0	0 ± 0	1 ± 0	1	0,865 ± 0,059	0 ± 0	0 ± 0	1 ± 0
99	2	0,831 ± 0,029	38 ± 12	0 ± 0	1 ± 0	1	0,866 ± 0,052	0 ± 0	0 ± 0	1 ± 0
100	2	0,834 ± 0,041	41 ± 24	0 ± 0	1 ± 0	1	0,842 ± 0,053	0 ± 0	0 ± 0	1 ± 0
101	2	0,802 ± 0,036	39 ± 11	0 ± 0	1 ± 0	1	0,779 ± 0,055	0 ± 0	0 ± 0	1 ± 0
102	2	0,723 ± 0,040	44 ± 11	0 ± 0	1 ± 0	1	0,738 ± 0,041	0 ± 0	0 ± 0	1 ± 0
103	2	0,788 ± 0,047	25 ± 11	0 ± 0	1 ± 0	1	0,795 ± 0,065	0 ± 0	0 ± 0	1 ± 0
104	2	0,735 ± 0,042	23 ± 7	0 ± 0	1 ± 0	1	0,743 ± 0,055	0 ± 0	0 ± 0	1 ± 0
105	2	0,649 ± 0,045	50 ± 12	0 ± 0	1 ± 0					
106	2	0,856 ± 0,036	23 ± 15	0 ± 0	1 ± 0	1	0,807 ± 0,055	0 ± 0	0 ± 0	1 ± 0
107	3	0,889 ± 0,000	0 ± 0	1,04 ± 0	1 ± 0	1	0,914 ± 0,062	0 ± 0	0 ± 0	1 ± 0
108	1	0,875 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
110	1	0,950 ± 0,035	0 ± 0	0 ± 0	1 ± 0	1	0,854 ± 0,077	0 ± 0	0 ± 0	1 ± 0
111	1	0,953 ± 0,037	0 ± 0	0 ± 0	1 ± 0					
112	1	0,933 ± 0,040	0 ± 0	0 ± 0	1 ± 0	1	0,825 ± 0,067	0 ± 0	0 ± 0	1 ± 0
113	1	0,937 ± 0,037	0 ± 0	0 ± 0	1 ± 0	1	0,904 ± 0,065	0 ± 0	0 ± 0	1 ± 0
114	4	0,925 ± 0,204	244 ± 1380	2,2 ± 0,8	1 ± 0					
115	3	0,913 ± 0,000	0 ± 0	0,63 ± 0	1 ± 0	1	0,806 ± 0,074	0 ± 0	0 ± 0	1 ± 0
116	1	0,907 ± 0,035	0 ± 0	0 ± 0	1 ± 0	1	0,827 ± 0,082	0 ± 0	0 ± 0	1 ± 0
118	1	0,932 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
119	3	0,897 ± 0,000	0 ± 0	0,44 ± 0	1 ± 0					
120	1	0,920 ± 0,033	0 ± 0	0 ± 0	1 ± 0					
121	1	0,884 ± 0,038	0 ± 0	0 ± 0	1 ± 0	1	0,841 ± 0,061	0 ± 0	0 ± 0	1 ± 0
122	1	0,826 ± 0,034	0 ± 0	0 ± 0	1 ± 0	1	0,784 ± 0,078	0 ± 0	0 ± 0	1 ± 0
123	2	0,506 ± 0,055	67 ± 91	0 ± 0	1 ± 0	2	0,554 ± 0,044	44 ± 141	0 ± 0	1 ± 0

Résumé

ASP2 est une protéine de 123 acides aminés liant les odeurs (Odorant-Binding Protein). Elle est sans doute impliquée dans la première étape du processus de reconnaissance des odeurs chez l'abeille domestique au niveau de la lymph sensillaire.

Le présent travail de thèse a permis, par résonance magnétique nucléaire, d'élucider la structure ainsi que de sonder la dynamique du squelette d'ASP2 en complexe avec le TSP (triméthylesilyl propionate-d4). A l'exception du segment D30-S45, la structure d'ASP2, monomérique, est bien définie ($\text{rmsd} < 1\text{\AA}$). Quatre hélices $\alpha_{1\alpha}$ et α_{4-6} antiparallèles forment un faisceau convergent. L'extrémité ouverte du faisceau est fermée par l'hélice α_3 . L'arrangement de ces cinq hélices α crée une poche essentiellement hydrophobe et faiblement polaire. Le segment D30-S45 présente une dynamique originale expliquant sa mauvaise définition dans la structure : le segment D30 à A38 est en échange conformationnel sur la gamme de la milliseconde tandis que la région A37 à S45 contient des mouvements importants dans la gamme pico-nanoseconde.

L'apoprotéine est présente sous deux formes en équilibre lent et de conformations très similaires à celle du complexe. Les populations respectives dépendent du pH.

Lors de la fixation de ligand, une seule des conformations est sélectionnée. L'IBMP (2-isobutyle-3-méthoxy-pyrazine) se fixe au sein de la poche hydrophobe de deux orientations, tête-bêches, avec des affinités comparables. De plus, la poche hydrophobe d'ASP2 s'adapte à la taille du ligand.

L'implication fonctionnelle de ces résultats est discutée.

Mots-Clés : Résonance Magnétique Nucléaire, protéine transporteuse d'odeur, structure, dynamique, interaction, *Apis Mellifera*.

Abstract

ASP2 is a 123-amino-acid odorant-binding protein (OBP) from the honeybee. The transport of airborne, hydrophobic odorants and pheromones through the sensillum lymph is thought to be accomplished by OBPs.

The structure of the complex of ASP2 with TSP (triméthylesilylpropionate-d4) was solved by nuclear magnetic resonance. The structure is monomeric and well-defined ($\text{rmsd} < 1\text{\AA}$), apart from residues D30 to S45. The overall fold of ASP2 is roughly conical, with four antiparallel helices ($\alpha_{1\alpha}$ et α_{4-6}) converging to a point and enclosing a highly hydrophobic and weakly polar pocket, with the opposite side capped by a fifth helix (α_3). Residues D30 to S45 undergo an original dynamical behavior: segment D30 to A38 is conformationally disordered in millisecond-timescale exchange whereas backbone from A37 to S45 is highly mobile in the pico-nanosecond range. The global correlation time is consistent with a monomeric protein.

The ASP2 apoprotein is present under two stable forms whose conformations are very similar to the structure of the complex. Their relative populations are pH-dependent. Upon binding, one conformation is selected. The odorant IBMP (2-isobutyl-3-méthoxy-pyrazin) binds to ASP2 hydrophobic pocket with two orientations of similar affinity, in a head-to-foot fashion. Furthermore, the binding cavity adapts to the size of the ligand.

Functional considerations are discussed.

Key Words: Nuclear Magnetic Resonance, Odorant-Binding Protein, structure, dynamics, interaction, *Apis Mellifera*.

Laboratoire de RMN à haut-champ, ICSN-CNRS, Gif-sur-Yvette
Pôle : Ingénierie des Protéines
Université PARIS XI