



Protonation spécifique des méthyles : un outil pour l'étude structurale des assemblages moléculaires par Résonance Magnétique Nucléaires

Rémy Sounier

► To cite this version:

Rémy Sounier. Protonation spécifique des méthyles : un outil pour l'étude structurale des assemblages moléculaires par Résonance Magnétique Nucléaires. Biologie structurale [q-bio.BM]. Université de Grenoble, 2008. Français. NNT : . tel-00785223

HAL Id: tel-00785223

<https://theses.hal.science/tel-00785223>

Submitted on 5 Feb 2013

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.

UNIVERSITE JOSEPH FOURIER – GRENOBLE 1

ÉCOLE DOCTORALE DE PHYSIQUE

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER

Discipline : Physique pour les Sciences du Vivant

présentée par

Rémy SOUNIER

**Protonation spécifique des méthyles : un outil pour
l'étude structurale des assemblages moléculaires
par Résonance Magnétique Nucléaire.**

Thèse soutenue publiquement le 12 Février 2008 devant le jury :

Frédéric DARDEL

Président

Eric GUITTET

Rapporteur

Christian ROUMESTAND

Rapporteur

Jean-Pierre SIMORRE

Directeur de thèse

Jérôme BOISBOUVIER

Co-directeur de thèse

*"Rien ne vaut la recherche lorsque l'on veut trouver
quelque chose."*

J.R.R. Tolkien [Bilbo le Hobbit]

A ma Mamie,

Remerciements

Voici venu le temps des remerciements, prémisse de la fin d'un long épisode. Ces 4 années passées à l'IBS et plus particulièrement au LRMN avec leurs hauts et leurs bas se terminent malgré tout avec beaucoup de regrets. Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à cette thèse.

Je remercie cordialement les membres du jury pour avoir accepté de porter leurs jugements sur ce travail de thèse. Tout d'abord, un grand merci à mes deux rapporteurs Eric Guittet et Christian Roumestand pour avoir bien voulu évaluer ce manuscrit. Je tiens à remercier également Frédéric Dardel pour son expertise et sa participation à mon jury.

Merci à mes directeurs de thèse. Tout d'abord à Jean-Pierre Simorre, "mon vrai faux directeur", merci de m'avoir accueilli au laboratoire il y a 4 ans, merci aussi pour toutes les discussions scientifiques (ou non) et pour avoir été le premier à utiliser la méthode des "méthyles-méthyles à longue portée" au laboratoire.

A Jérôme Boisbouvier, "mon chef de tous les jours", merci de m'avoir proposé des sujets aussi exaltants, d'être là durant les moments difficiles, et de m'avoir supporté dans ton bureau pendant plus de 3 ans. Je lui suis également redevable d'avoir pu trouver en lui un exemple d'honnêteté et de rigueur intellectuelles. Je le remercie enfin pour m'avoir laissé une grande liberté, et de m'avoir redirigé lorsque je m'écartais du bon chemin.

Je tiens à remercier aussi Isabel Ayala pour m'avoir formé à devenir un petit despote des bactéries, pour sa patience (bientôt légendaire) dans l'encadrement de futurs biochimistes, pour sa bonne humeur et pour son sourire qui illumine le laboratoire. Merci également à Laurence Blanchard, pour m'avoir appris le marquage spécifique des méthyles et pour m'avoir corrigé mon anglais lors des répétitions des présentations.

Merci aussi à Paul Schanda, Ewen Lescop et Bernhard Brutscher pour la table ronde qui m'a permis d'appréhender de nouveaux aspects de la spectroscopie. Pour leur disponibilité et les discussions, je tiens à remercier tous les gens du LRMN, en particulier Catherine Bougault (toujours présente pour répondre à des questions), Adrien Favier (pour l'aide apportée sur les postes de travail et surtout sur les spectros), Pierre Gans, Dominique Marion, Beate Bersch et Martin Blackledge.

Je tiens à remercier "l'équipe du PSB", Lionel Imbert (pour l'aide apporté sur la préparation d'échantillons de Spc 24-Spc 25), Cécile Giustini, Fito (pour son expertise sur l' α -synucléine), Nathalie Use (pour la production en chaîne de l'ubiquitine marquée spécifiquement) et Isabel Ayala (pour les préparations de la Malate Synthase G et de TET2 à la fin de ma thèse).

Un grand merci à notre groupe de thésards et de post-docs, Julien (la Yaute attitude!!!), Guillaume, Paul, Hélène, Denis, Monika, Anna, Pau, Klaartje, Matthew, Nicolas, Malene, Fito, Florence, Cédric, Ewen, Loic, Gabrielle, Michael et à l'ensemble des stagiaires. Un grand merci pour la bonne ambiance au labo et toutes les soirées que l'on a passées ensemble autour d'une bonne bière chez les uns et les autres ou au O'Callaghan.

Je voudrais remercier aussi mes copains, camarades et amis avec qui j'ai passé d'agréables moments universitaires à Grenoble. Je pense à Gérard pour notre collocation ; à Hélène pour toutes les cartes postales ; à l'ensemble de notre promo de Physique-Chimie, avec Nelly, Kelly, Will, Camille ; à Jérôme, dit Chapi (ou Chapo, je ne sais plus), pour les heures de TP et de ciné passé ensemble (Bonne chance pour ta fin de thèse).

Je n'oublie pas également l'équipe de water-polo, qui, durant ces années passées à Grenoble, a égaillé mes nombreuses soirées et a réussi à maintenir une certaine sérénité dans mon travail. Un grand merci à Aude, notre mamie préférée, pour les heures passées à relire les fautes de ce manuscrit ; au premier Pélot (Pierre) pour les soirées restaurants et à Barbu pour nous avoir rejoints dans nos soirées à 3 pélots ; à Etienne pour sa bibliothèque de mangas et de BD et à Anne-Lyse pour tous ses encouragements (Bonne chance à vous deux pour la suite); à Emily pour les soirées "aux tord-boyaux" et pour m'avoir accueilli chez elle à York ; à Sarah pour les petits goûters ; à Audrey pour toutes les discussions ; à Titi ; à Hélène pour nos débats ; à Claire pour m'avoir accepté durant sa collocation ; à Nico pour être présent quoi qu'il arrive ; à Cécile, enfin, pour m'avoir aidé et soutenu depuis presque 2 ans et surtout durant la rédaction de cette thèse.

Pour finir, je remercie également mes parents et mon frère pour leurs encouragements et pour avoir été toujours présents dans les bons moments comme dans les moins bons. Merci à vous pour m'avoir permis d'arriver jusqu'ici.

Résumé

PROTONATION SPECIFIQUE DES METHYLES : UN OUTIL POUR L'ETUDE STRUCTURALE DES ASSEMBLAGES MOLECULAIRES PAR RESONANCE MAGNETIQUE NUCLEAIRE

Les méthodes RMN standard appliquées à l'analyse structurale des protéines sont basées sur la mesure d'informations géométriques locales entre protons. Le manque de contraintes à longue portée est un facteur limitant pour l'étude des protéines modulaires, des complexes protéiques et des assemblages supramoléculaires. La deutération des échantillons est nécessaire pour l'étude de ces systèmes de grande taille. Néanmoins, le remplacement des protons par du deutérium limite le nombre de contraintes structurales détectables, en particulier des contraintes de distances de types NOEs. Pour résoudre ce problème, une stratégie basée sur la protonation de quelques sites discrets a été mise en place. La protonation spécifique des méthyles dans des protéines constitue un choix optimal pour l'obtention de contraintes à longue portée. Les protocoles de marquage spécifique des méthyles des Isoleucines, des Valines et des Leucines ont été implémentés et optimisés, et une nouvelle méthode de protonation spécifique des Alanines a été développée. Dans des systèmes de taille modérée, l'utilisation de ce type de marquage combiné aux développements d'expériences RMN adaptées permet la détection de contraintes (NOEs et RDCs) entre paire de méthyles séparés par plus de 12 Å. Une méthode robuste a été développée pour extraire des distances avec une très haute précision. Cette approche s'applique également à des protéines de grande taille, ainsi des NOEs entre méthyles distants par plus de 7 Å sont détecté dans un assemblage supramoléculaire de 468 kDa.

Mot Clés : Résonance Magnétique Nucléaire, Marquage Isotopique, Protonation Spécifique des Méthyles, Protéines, Effet Overhauser Nucléaire, Couplage Dipolaire Résiduel.

Laboratoire de these : Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075, CEA; CNRS; UJF. Laboratoire de Résonance Magnétique Nucléaire. 41, rue Jules Horowitz, 38027 Grenoble cedex.

Abstract

SPECIFIC PROTONATION OF METHYLS : A TOOL FOR STRUCTURAL STUDY OF MOLECULAR ASSEMBLIES BY NUCLEAR MAGNETIC RESONANCE.

Standard NMR methods applied to the structural analysis of proteins are based on the measurement of local geometric information between protons. The lack of long range restraints can be limiting factor for the study of the modular proteins, protein complexes and supramolecular assemblies. Perdeuteration of samples is necessary for the study of these large systems. However, the exchange of protons with deuterium limit the number of detectable structural restraints, particularly NOEs distance restraints. To resolve this problem, a strategy based on the protonation of some discrete sites was developed. The specific protonation of methyls in proteins constitutes an optimum choice for detection of long range restraints. Protocols for specific labelling of methyls of Isoleucines, Valines and Leucines were implemented and optimized, and a new method for specific protonation of Alanines was introduced. In moderate size systems, the use of this specific labelling strategy combined with the development of NMR experiments allows the detection of restraints (NOEs and RDCs) between methyls separated by more than 12 Å. A robust method was developed to extract distances with very high precision. This approach can also be applied to large proteins. NOEs have been observed between methyls separated by more than 7 Å in a 468 kDa supramolecular assembly.

Key words : Nuclear Magnetic Resonance, Isotopic Labelling, Specific Protonation of Methyls, Proteins, Nuclear Overhauser Effect, Residual Dipolar Coupling.

Table des matières

CHAPITRE 1. INTRODUCTION	17
1.1 LA PREPARATION D'ECHANTILLONS DE PROTEINE EN SOLUTION	18
1.1.1 <i>Production d'échantillons de protéine en solution</i>	18
1.1.2 <i>Le marquage isotopique</i>	19
1.1.2.1 <i>Le marquage isotopique uniforme</i>	20
1.1.2.2 <i>Le marquage isotopique spécifique</i>	21
1.2 ATTRIBUTION DES RESONANCES	21
1.3 MESURES DE CONTRAINTES CONFORMATIONNELLES	22
1.3.1 <i>Les contraintes spatiales</i>	23
1.3.1.1 <i>L'effet Overhauser nucléaire (NOE)</i>	23
1.3.1.2 <i>Les liaisons hydrogènes</i>	25
1.3.2 <i>Les contraintes angulaires locales</i>	25
1.3.2.1 <i>Le couplage scalaire 3J</i>	25
1.3.3 <i>Les informations orientationnelles globales</i>	27
1.3.3.1 <i>Les relaxations longitudinale et transversale hétéronucléaires</i>	27
1.3.3.2 <i>Les couplages résiduels dipolaires (RDCs)</i>	29
1.3.4 <i>Les contraintes obtenues dans les systèmes paramagnétiques</i>	31
1.3.4.1 <i>Les vitesses de relaxation R_1 et R_2</i>	32
1.3.4.2 <i>Les déplacements chimiques dipolaires (PCS)</i>	33
1.3.4.3 <i>La corrélation croisée Curie-Dipolaire</i>	34
1.4 DETECTIONS DE NOUVELLES INFORMATIONS A LONGUE PORTEE.	36
CHAPITRE 2. LA PROTONATION SPECIFIQUE DES METHYLES DANS LES PROTEINES	39
2.1 INTRODUCTION	40
2.2 LES METHODES DE MARQUAGES SELECTIFS	42
2.2.1 <i>Le marquage par inhibition</i>	42
2.2.2 <i>Les souches auxotrophes</i>	45
2.2.3 <i>La deutération sélective par précurseur d'acides aminés</i>	46
2.2.4 <i>Le marquage in-vitro</i>	47
2.3 CHOIX DU SITE DE PROTONATION SPECIFIQUE	49
2.4 PROTONATION SPECIFIQUE DU METHYLE δ_1 DES ISOLEUCINES	51
2.4.1 <i>Choix du méthyle δ_1 des Isoleucines</i>	51
2.4.2 <i>Principe de marquage spécifique du méthyle δ_1 des Isoleucines</i>	52
2.4.3 <i>Implémentation du protocole</i>	53
2.4.3.1 <i>Le solvant</i>	53
2.4.3.2 <i>Choix de la source de carbone</i>	53
2.4.3.3 <i>Préparation de l'α-cétobutyrate protoné spécifiquement</i>	55

2.4.3.4 Mise en place du protocole	56
2.4.3.5 Points critiques du protocole.....	58
2.4.4 Caractérisation de l'efficacité du marquage spécifique.....	59
2.4.5 Différents types de marquages envisageables	60
2.4.6 Applications.....	62
2.4.6.1 Préparation de protéines U-[² H], U-[¹² C / ¹³ C], [¹² CH ₃]-Ile- δ_1	62
2.4.6.2 Préparation de protéines U-[² H], U-[¹² C], [¹³ CH ₃]-Ile- δ_1	63
2.5 PROTONATION SPECIFIQUE DES METHYLES DES VALINES ET DES LEUCINES.....	65
2.5.1 Principe de marquage spécifique du méthyle des Valines et des Leucines	65
2.5.2 Préparation de l' α -cétoisovalérate protoné spécifiquement	66
2.5.3 Différents types de marquages envisageables	67
2.5.4 Application : MSG	68
2.6 PROTONATION SPECIFIQUE DU METHYLE DE L'ALANINE	69
2.6.1 Introduction	69
2.6.2 Principe du protocole de surexpression	70
2.6.3 Optimisation de l'incorporation du C β de l'Alanine	72
2.6.4 Limitation des fuites métaboliques.....	73
2.6.4.1 Estimation des fuites métaboliques.....	73
2.6.4.2 Dilution des fuites à l'aide du pyruvate.....	74
2.6.4.3 Limitations du marquage résiduel des autres méthyles.....	75
2.6.4.4 Limitation des fuites métaboliques au niveau du cycle de Krebs et de la glycolyse.....	76
2.6.4.5 Caractérisation du marquage	78
2.6.5 Protonation spécifique du méthyle de l'Alanine	79
2.6.5.1 Passage au solvant deutéré	79
2.6.5.2 Synthèse de l'Alanine spécifiquement protonée.....	80
2.6.5.3 Résultats.....	81
CHAPITRE 3. MESURE PRECISE DES DISTANCES ENTRE METHYLES ELOIGNES	
DANS LES PROTEINES SPECIFIQUEMENT PROTONEES.....	83
3.1 DETECTION DE L'EFFET OVERHAUSER NUCLEAIRE (NOE) A LONGUE PORTEE.....	84
3.1.1 Introduction	84
3.1.2 Principe de l'effet Overhauser nucléaire	85
3.1.2.1 Découverte de l'effet Overhauser.....	85
3.1.2.2 Description de l'effet NOE dans un système de 2 protons isolés	85
3.1.2.3 Description de l'effet NOE dans un système de n protons	88
3.1.2.4 Application aux protéines spécifiquement protonées sur les méthyles	90
3.1.3 Principe de l'expérience NOESY	91
3.1.3.1 Le temps de récupération (d ₁)	92
3.1.3.2 Le temps de transfert (τ_m)	93
3.1.4 Optimisation des paramètres d'acquisition pour la mesure de NOEs interméthyles dans les protéines spécifiquement protonées	94

3.1.4.1 Détermination de la relaxation longitudinale T_1	94
3.1.4.2 Le temps de récupération optimal (d_1)	95
3.1.4.3 Le temps de transfert optimal (τ_m^0)	96
3.1.5 Détection de faibles NOEs dans les protéines spécifiquement protonées	97
3.1.6 Applications à la détection de NOEs entre méthyles éloignés dans les protéines de tailles modérées.	100
3.2 QUANTIFICATION DES DISTANCES INTERMETHYLES	102
3.2.1 Matrice de relaxation totale	102
3.2.2 Applications	104
3.2.2.1 Résultats	105
3.2.2.2 Importance de la faible densité de protons dans l'extraction des distances interméthyles	106
3.3 CONCLUSIONS ET PERSPECTIVES	108
3.4 MESURE PRECISE DE DISTANCE ENTRE METHYLES SEPARES DANS DES PROTEINES SPECIFIQUEMENT PROTONEES	109
CHAPITRE 4. LA DETECTION DE NOES ENTRE LES METHYLES DANS LES ASSEMBLAGES SUPRAMOLECULAIRES	113
4.1 INTRODUCTION	114
4.2 LES ASPECTS EXPERIMENTAUX	115
4.2.1 Expérience NOESY éditée ^{13}C	115
4.2.2 Quantification des NOEs	116
4.3 DETECTION DE NOES A LONGUE PORTEE DANS LA MALATE SYNTHASE G (MSG), UNE PROTEINE GLOBULAIRE DE 82 KDA.	117
4.4 DETECTION DE NOES INTERMETHYLES DANS LES ASSEMBLAGES MOLECULAIRES DE GRANDES TAILLES : ETUDE THEORIQUE DES LIMITES ET DES POSSIBILITES.	119
4.4.1 La relaxation longitudinale T_1	119
4.4.2 Etude du transfert d'aimantation NOE en fonction de la taille de la molécule	122
4.4.3 La relaxation transversale limite la distance effective des NOEs à longue portée	124
4.5 APPLICATION A DES ASSEMBLAGES MOLECULAIRES DE GRANDE TAILLE	126
4.5.1 Utilisation de MSG comme un système modèle	126
4.5.2 Application à l'étude de TET2, un assemblage moléculaire de 468 kDa	128
4.6 LES NOES INTERMETHYLES POUR LES ETUDES STRUCTURALES	131
4.6.1 Application à l'étude d'une protéine globulaire de 21 kDa	131
4.6.2 Perspectives : applications aux protéines modulaires ou assemblages moléculaires de grandes tailles.	133
4.6.2.1 Les domaines de MSG	133
4.6.2.2 Vers l'affinement des modèles structuraux de protéines de grandes tailles : marquages spécifiques des méthyles et nouvelles contraintes interméthyles	134
4.6.2.3 Les combinaisons de marquages spécifiques : source de nouvelles contraintes interméthyles	139

4.7 CONCLUSIONS	140
CHAPITRE 5. DETECTION DE CONTRAINTES TRANSLATIONNELLES A LONGUE PORTEE PAR RMN EN MILIEU CRISTAL LIQUIDE	141
5.1 LES COUPLAGES DIPOLAIRES RESIDUELS (RDCs).....	142
5.1.1 Théorie	142
5.1.2 Les RDCs fournissent des informations orientationnelles	145
5.1.3 Les RDCs ^1H - ^1H : une source d'information translationnelles	146
5.2 LA DETECTION RAPIDE DES RDCs H_N - H_N DANS LES PROTEINES PROTONEES.....	148
5.2.1 L'expérience Semi-Selective-HMQC (SS-HMQC)	148
5.2.1.1 Le Principe du transfert de cohérence	148
5.2.1.2 Quantification des RDCs ^1H - ^1H	149
5.2.1.3 La séquence d'impulsion SS-HMQC2	150
5.2.2 Diminution du temps de relaxation longitudinale	151
5.2.3 Compensation de la disparité de couplage	153
5.2.4 Expérience optimisé pour la mesure de couplage dipolaire résiduel entre les protons amides.....	154
5.3 LA DETECTION DES RDCs A LONGUE PORTEE EST LIMITEE PAR LA TRONCATION DIPOLAIRE	164
5.3.1.1 Suppression de couplages forts par des impulsions à bandes sélectives	164
5.3.1.2 Limitation de la troncation dipolaire dans les acides nucléiques	166
5.4 DETECTION DE RDCs INTERMETHYLES A LONGUE PORTEE DANS LES PROTEINES SPECIFIQUEMENT PROTONEES	168
5.4.1 Propriétés de relaxation des méthyles en milieu isotrope.....	168
5.4.2 Evolution des cohérences transversales simple-quantum ^1H des méthyles en milieu cristal liquide.....	170
5.4.3 Evolution de l'aimantation transversale des méthyles en milieu cristal liquide.....	170
5.4.3.1 La contribution du deutérium à la relaxation des méthyles en milieu aligné.....	170
5.4.3.2 Contribution des autres protons	172
5.4.3.3 Evolution des couplages dipolaires résiduels ^1H - ^1H intraméthyles	174
5.4.4 Détection de RDCs ^1H - ^1H entre méthyles éloignés.....	176
5.4.4.1 La séquence d'impulsion : "3D Selective CT-COSY"	176
5.4.4.2 Application : ubiquitine U- ^{13}C , U- ^{13}C , $^{12}\text{CH}_3$ -Ile- δ_1	178
Corrélations des RDCs ^1H - ^1H avec la structures	180
5.4.4.3 Application : γ S-crystalline U- ^{13}C , $^{12}\text{CH}_3$ -Ile- δ_1	182
CHAPITRE 6. CONCLUSIONS.....	185
ANNEXE 1.....	191
ANNEXE 2.....	195
ANNEXE 3.....	203
REFERENCES BIBLIOGRAPHIQUES	207

Chapitre

1

Introduction

Jusqu'en 1985, la seule méthode permettant de déterminer la structure tridimensionnelle d'une macromolécule était la cristallographie par diffraction des rayons X. La résolution de la première structure de protéine en solution, par Résonance magnétique nucléaire (RMN) (Williamson *et al*, 1985), a ouvert de nouvelles perspectives dans le domaine de la biologie structurale. La RMN ne permet pas d'accéder directement à une image tridimensionnelle d'une macromolécule mais d'obtenir des informations structurales et dynamiques à l'échelle atomique. Pour établir la structure tridimensionnelle par RMN d'une macromolécule biologique en vue de comprendre et d'expliquer les fonctions de cette dernière, plusieurs étapes sont nécessaires :

- (i) Préparation d'un échantillon en solution.
- (ii) Attribution des fréquences de résonances des noyaux détectables par RMN.
- (iii) Mesure de contraintes conformationnelles.
- (iiii) Calculs de structures.

Grâce aux progrès réalisés dans les différents domaines entourant la détermination de structure tridimensionnelle par RMN en solution, les systèmes biologiques étudiés sont de plus en plus variés : peptides, acides nucléiques, protéines globulaires, modulaires, membranaires et complexes protéiques.

1.1 La préparation d'échantillons de protéine en solution

1.1.1 Production d'échantillons de protéine en solution

La préparation d'un échantillon de protéine est une étape cruciale pour l'étude structurale par RMN. Dû à la faible sensibilité de la spectroscopie RMN, l'échantillon doit avoir une concentration en protéines de l'ordre du millimolaire, soit, pour une protéine de 10 kDa, 5 milligrammes pour un volume effectif de 500 μ L. Afin d'optimiser la qualité des spectres RMN, l'échantillon doit être pur et stable pendant plusieurs semaines. Pour cela, plusieurs étapes de purification peuvent être

requis et les conditions physico-chimiques de l'échantillon sont adaptées en ajustant le pH de la solution, la nature du tampon, la force ionique, la température et tous paramètres susceptibles d'augmenter le seuil de solubilité de la protéine. L'ensemble de ces étapes délicates peut être long mais est indispensable.

Les études structurales par RMN des protéines requièrent en plus des critères décrits ci-dessus, un marquage isotopique. Pour préparer de tels échantillons, une des solutions consiste à surexprimer la protéine dans des cultures bactériennes, en milieu contenant des produits marqués isotopiquement. La culture des bactéries en milieu minimum permet de contrôler les différentes sources d'isotopes nécessaires à la production des protéines.

1.1.2 Le marquage isotopique

La spectroscopie RMN est basée sur l'étude d'interaction entre spins nucléaires. Le signal RMN d'un spin est caractérisé par sa fréquence de résonance, exprimée en Hz ou en ppm. La sensibilité d'une espèce atomique pour la RMN dépend de sa concentration et de son rapport gyromagnétique (γ) qui, en première approximation, reflète sa propension à interagir avec le champ magnétique. Le rapport gyromagnétique définit une sensibilité relative pour un nombre de noyaux donnés. Quant à la concentration de noyaux, elle dépend de l'abondance naturelle de l'isotope stable considéré. La sensibilité absolue pour la RMN est le produit de l'abondance naturelle par la sensibilité relative. Le proton est habituellement pris comme référence.

Les molécules du vivant, en particulier les protéines, sont constituées d'atomes d'hydrogène, de carbone, d'oxygène, d'azote et de soufre. Historiquement, les premières structures tridimensionnelles de protéines ont été établies par la RMN homonucléaire du proton, isotope ^1H de l'atome d'hydrogène présent en abondance naturelle à 99,98 %. Pour les autres atomes, du fait leur faible sensibilité, l'enregistrement en abondance naturelle d'expérience hétéronucléaire requerrait une concentration en protéines de l'ordre de 10 mM, ce qui est généralement incompatible avec leurs solubilités. De plus, ceci nécessiterait d'accumuler les

expériences pendant plusieurs semaines. Un moyen pour augmenter leurs sensibilités consiste à enrichir la protéine en isotopes détectables pour la RMN.

1.1.2.1 Le marquage isotopique uniforme

Le marquage uniforme d'une protéine en RMN est le remplacement de tous les noyaux d'une espèce atomique par l'un de ses isotopes requis pour l'étude RMN : ^{12}C par ^{13}C , ^{14}N par ^{15}N , et le proton par le deutérium. Dans cette thèse, pour indiquer que le marquage est uniforme, j'utiliserai la nomenclature suivante : U-[X], où X correspond à l'isotope de l'atome. Ainsi, pour indiquer qu'un échantillon est marqué uniformément ^{13}C , par exemple, la notation U- ^{13}C sera utilisée.

Le marquage U- ^{15}N permet l'enregistrement d'une expérience de corrélation ^1H - ^{15}N , l'empreinte de la protéine et ainsi de pouvoir déterminer si la protéine est en partie repliée ou non. Toutefois, ce marquage n'est, en général, pas suffisant en lui-même pour établir la structure tridimensionnelle des protéines. Pour l'étude de protéines de masses moléculaires supérieures à 10 kDa, les marquages combinés U- ^{13}C , U- ^{15}N sont recommandés. Ce double marquage uniforme, utilisé avec des séquences d'impulsions RMN triple-résonances, permet de distribuer les signaux RMN dans plusieurs dimensions (2, 3, 4, etc.), afin de faciliter l'étape d'attribution des résonances.

Le marquage U- ^2H (dernier marquage uniforme possible pour les protéines) est nécessaire lorsque la taille de la protéine étudiée est élevée (> 20 kDa). En effet, l'étude structurale de macromolécules de taille importante se heurte à deux difficultés majeures : l'élargissement des signaux RMN induit par les interactions dipolaires (d'autant plus grand que le temps de corrélation croît) et l'augmentation du nombre de signaux dans les spectres RMN. Le remplacement des protons par des ^2H permet d'augmenter significativement la sensibilité et de diminuer le nombre de résonances de protons dans les spectres RMN. Les effets de relaxation engendrés par les couplages dipolaires ^1H - ^1H à travers l'espace sont diminués, ainsi que ceux dus aux interactions dipolaires ^1H - ^{13}C , qui interviennent sur la relaxation du ^{13}C . Les interactions principales responsables de la relaxation des cohérences ayant été

supprimées, on peut utiliser les expériences triple-résonances standard pour l'attribution de protéines de taille moléculaire plus élevée.

1.1.2.2 Le marquage isotopique spécifique

D'autres techniques de marquages ont été proposées pour réduire la complexité des spectres RMN : les marquages sélectifs d'acides aminés. Quelques exemples de marquages sélectifs en RMN sont indiqués :

(i) le marquage sélectif du ^{15}N d'un type particulier d'acide aminé, qui est présenté dans le chapitre 2.

(ii) le marquage uniforme ^{15}N et sélectif des ^{13}C de la chaîne latérale d'un type d'acide aminé.

(iii) le marquage sélectif proton de certaine position dans un échantillon deutéré, comme la protonation sélective des méthyles décrite précisément dans le chapitre 2.

Un grand nombre de marquages sélectifs peuvent être envisagés (l'incorporation de segments de protéines marquées isotopiquement, par exemple). Le principal avantage de ce type de marquage est de simplifier le système pour détecter des informations non observables autrement. Toutefois, le coût d'un échantillon marqué spécifiquement est souvent le facteur limitant. Ainsi, proposer des protocoles de production d'échantillons spécifiquement marqués à des coûts relativement peu onéreux reste un défi à relever pour bénéficier des avantages de ces nouveaux échantillons.

1.2 Attribution des résonances

Avant de pouvoir mesurer des contraintes permettant la détermination de structures de biomolécules à une résolution atomique par RMN, l'attribution des résonances des spins détectables (^1H , ^{13}C et ^{15}N) est indispensable. Les stratégies d'attribution des atomes constituant une protéine impliquent l'utilisation de

techniques multidimensionnelles qui permettent de distribuer les signaux dans les différentes dimensions des spectres RMN.

Les stratégies d'attribution dépendent du marquage isotopique utilisé. Ainsi, pour une protéine non-enrichie isotopiquement, des expériences homonucléaires proton sont utilisées (2D COSY, 2D TOCSY) : elles permettent de corréler les protons couplés ensemble et ainsi d'identifier chaque système de spin correspondant à un acide aminé. Ensuite, des expériences de type 2D NOESY sont enregistrées pour déterminer l'enchaînement des résidus (Wüthrich, 1986).

Dans le cas de protéines marquées isotopiquement, les stratégies d'attribution sont divisées en deux étapes : l'attribution de la chaîne principale et, par suite, l'attribution des chaînes latérales. L'attribution du squelette d'une protéine U-[^{15}N] et U-[^{13}C] est réalisée à partir d'une série d'expériences (^{15}N -HSQC, HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, ...) basées sur des transferts de cohérence à travers la liaison peptidique par des couplages scalaires 1J et 2J . Cette partie de l'étape d'attribution est, aujourd'hui, plus aisée grâce à l'utilisation de logiciels d'attribution semi-automatique. Les fréquences de résonances ^1H et ^{13}C des chaînes latérales sont attribuées grâce à d'autres d'expériences (CBCA(CO)NH, H(C)CC(CO)NH, HCCH-TOCSY, ...). L'attribution de l'ensemble des résonances des noyaux de la molécule étudiée (^1H , ^{13}C , ^{15}N) peut être une étape plus longue (voire laborieuse).

1.3 Mesures de contraintes conformationnelles

Lorsque l'on a un échantillon exploitable pour la RMN et que l'étape d'attribution est réalisée, d'autres expériences RMN sont enregistrées afin de mesurer des informations structurales. Une grande variété de grandeurs physiques mesurables par RMN contient des informations structurales. Parmi celles-ci, le déplacement chimique, l'interaction du couplage scalaire et la relaxation dipolaire sont sensibles aux conformations moléculaires et à la distance entre les spins. La quantification de ces paramètres RMN permet l'analyse structurale. Les interactions dipolaires et de couplages scalaires sont à la base de tous les calculs de structure des protéines utilisant des données RMN.

1.3.1 Les contraintes spatiales

1.3.1.1 L'effet Overhauser nucléaire (NOE)

Le paramètre RMN observable le plus largement utilisé pour la détermination de structure de protéines est l'effet Overhauser nucléaire. Par interaction de type dipôle-dipôle, deux noyaux peuvent se transmettre de l'énergie à travers l'espace; ce phénomène est noté NOE. Cet effet de relaxation croisée est mesuré par RMN en observant la différence d'aimantation longitudinale (expérience NOESY) ou transversale (expérience ROESY) qu'il induit. Pour les macromolécules biologiques, les intensités des pics de corrélations NOEs, I_{ij} , sont proportionnelles au terme de relaxation croisée, σ_{ij} , entre les aimantations de chaque couple de protons i et j et au temps de transfert, τ_m , au cours duquel l'interaction dipolaire est active.

$$I_{ij} \propto \tau_m \sigma_{ij} \quad (1.1)$$

$$\sigma_{ij} = \frac{1}{10} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_H^2 \hbar \right)^2 [6J(2\omega_H) - J(0)] \quad (1.2)$$

où γ_H est le rapport gyromagnétique du proton, r_{ij} la distance entre les deux noyaux i et j , ω_H est la fréquence de Larmor associée aux protons, J est la fonction de densité spectrale.

Cet effet est proportionnel à l'inverse de la distance à la puissance six entre deux protons, r_{ij} . Ainsi, par exemple, si une distance entre 2 protons est connue (r_{ref}) à partir d'une géométrie covalente, alors les autres distances interprotons non connues (r_i) sont déterminées (en ne considérant aucune mobilité interne) par la relation suivante :

$$r_i = r_{ref} \left(\frac{S_{ref}}{S_i} \right)^{1/6} \quad (1.3)$$

où S_{ref} et S_i sont les intégrales des intensités des pics de corrélation. En pratique, du fait de la forte densité de protons, les intensités des pics de corrélations NOEs ne sont pas directement proportionnelles au terme de relaxation croisée entre des spins

interagissant entre eux. En effet, l'aimantation est aussi transférée entre spins à travers le réseau de protons via la diffusion de spins.

Pour la détermination de structure standard de protéines, les NOEs sont interprétés de manière semi-quantitative. Les pics de corrélations NOEs sont regroupés sur la base de leur volume en plusieurs catégories, comme montré dans le Tableau 1.1 (Wüthrich, 1986).

Intensité du nOe	Distances
Très faible	1,8 – 6,0 Å
Faible	1,8 – 5,0 Å
Moyen	1,8 – 3,3 Å
Fort	1,8 – 2,7 Å

Tableau 1.1 : Classification des NOEs suivant leur intensité et contraintes de distances lors d'un calcul de structure.

Dans les protéines protonées, la diffusion de spins affecte l'intensité des NOEs, qui sont alors utilisés de manière semi-quantitative. Les pics de corrélation NOEs sont regroupés sur la base de leur volume en plusieurs catégories, comme montré dans le Tableau 1.1 (Wüthrich, 1986). Une analyse des NOEs entre protons dont la distance est fixe (entre les protons géminaux du méthylène, ou les protons vicinaux des cycles aromatiques ou les protons dans des régions de structures secondaires régulières) est effectuée afin de disposer d'une référence interne pour affecter chaque NOE à une catégorie donnée. Pour obtenir un bon modèle de structure par RMN, un grand nombre de NOEs est nécessaire. On discerne les NOEs intrarésidus et séquentiels qui peuvent apporter des informations sur la structure locale. Les NOEs entre acides aminés proches dans la séquence (séparés par 3 résidus au maximum) complètent les NOEs séquentiels pour obtenir des informations sur les structures secondaires. A ceci s'ajoutent les NOEs à longue portée entre acides aminés distants dans la séquence. Cette dernière classe de NOEs est la plus importante et permet de déterminer le repliement global des protéines globulaires.

1.3.1.2 Les liaisons hydrogènes

Une autre source d'information structurale à courte portée est l'observation de liaisons hydrogènes. Elle permet souvent de relier des éléments de structures secondaires. En effet, l'échange lent d'un hydrogène dans les protéines est le plus souvent la preuve que cet atome est stabilisé par une liaison hydrogène. L'atome d'oxygène accepteur est fréquemment identifié par une analyse précise des connectivités NOEs entre les protons voisins appartenant aux éléments réguliers de structures secondaires. Une preuve expérimentale plus précise est obtenue à partir de l'observation de couplages scalaires à travers les liaisons hydrogènes ($^3J_{NC'}$). Dans des protéines U-[^{15}N], U-[^{13}C], U-[^2H], ce faible couplage (< 1 Hz) permet le transfert d'aimantation entre l'azote, lié au H_N , donneur de la liaison hydrogène, et le carbone du groupement carbonyle accepteur (Cordier *et al*, 1999 ; Wang *et al*, 1999). Le couplage scalaire confirme le recouvrement entre les orbitales électroniques des atomes impliqués et définit, de manière non ambiguë, les paires d'atomes formant une liaison hydrogène. Les liaisons hydrogènes sont utilisées comme contraintes de distance dans les calculs de structure, typiquement en contenant la distance accepteur-hydrogène $r(\text{O}\cdots\text{H})$ de 1,7-2,1 Å et la distance accepteur-donneur $r(\text{O}\cdots\text{N})$ de 2,7-3,0 Å. Lorsqu'une liaison hydrogène est détectée entre atomes éloignés dans la séquence, son impact dans la détermination de structure est considérable. Les liaisons hydrogènes dans les structures secondaires augmentent la régularité de ces éléments de structure dans les modèles calculés.

1.3.2 Les contraintes angulaires locales

1.3.2.1 Le couplage scalaire 3J

Le couplage scalaire 3J , entre atomes séparés par trois liaisons covalentes, est utilisé pour la détermination de structures de macromolécules biologiques afin d'affiner la conformation locale. Cette source d'information structurale a été décrite pour la première fois par Karplus (Karplus, 1959). La grandeur du couplage scalaire 3J est fonction de l'angle de torsion (dièdre), θ , formé par l'orientation relative des trois liaisons covalentes de ce couplage (Figure 1.1).

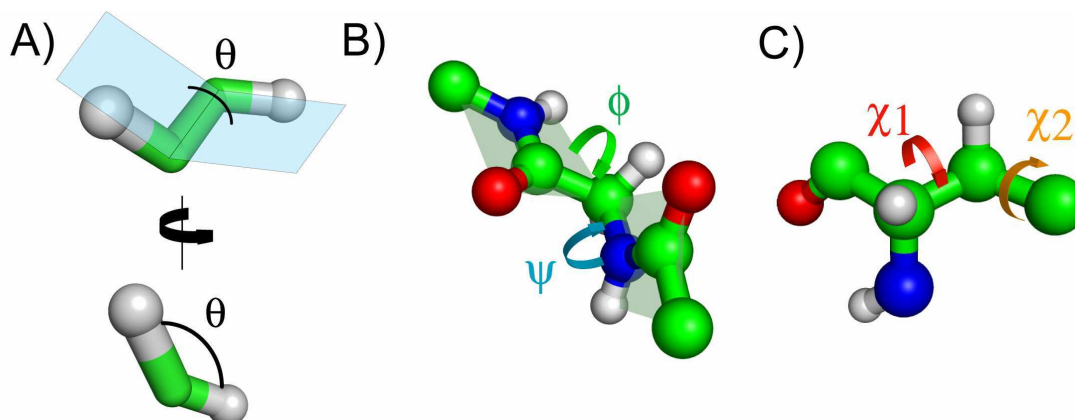


Figure 1.1 : A) Angle dièdre θ entre 3 liaisons covalentes. B) Les angles dièdres ϕ et ψ dans la chaîne principale des acides aminés d'une protéine. C) les angles dièdres χ_1 et χ_2 dans une chaîne latérale d'un acide aminé.

La relation reliant le couplage scalaire 3J à l'angle θ est montrée ci-dessous:

$$^3J = A \cos^2 \theta + B \cos \theta + C \quad (1.4)$$

où A, B et C sont des constantes dépendant des noyaux impliqués dans les liaisons covalentes.

Les angles dièdres ϕ et χ_1 sont dérivés, respectivement, à partir des constantes de couplages $^3J_{H^N H^\alpha}$ et $^3J_{H^\alpha H^\beta}$, permettant d'obtenir une meilleure définition de l'orientation des liaisons peptidiques et des chaînes latérales. La précision des paramètres de l'équation de Karplus est nécessaire afin de corrélérer les valeurs mesurées des constantes de couplages scalaires 3J aux valeurs des angles dièdres. Des approches théoriques ont été réalisées pour obtenir un rapport entre les constantes de couplages scalaires 3J et les angles dièdres correspondants. Toutefois, ces relations sont complexes et souvent difficilement applicables. Au lieu de cela, les constantes A, B et C de l'Equation 1.4 sont obtenues par des méthodes semi-empiriques.

Les valeurs les plus pertinentes, pour les constantes, sont dérivées par la corrélation entre les valeurs 3J observées pour des angles dièdres connus. Ces derniers sont mesurés dans les structures de protéines déterminées par cristallographie de diffraction des rayons X ou par RMN, (Figure 1.2).

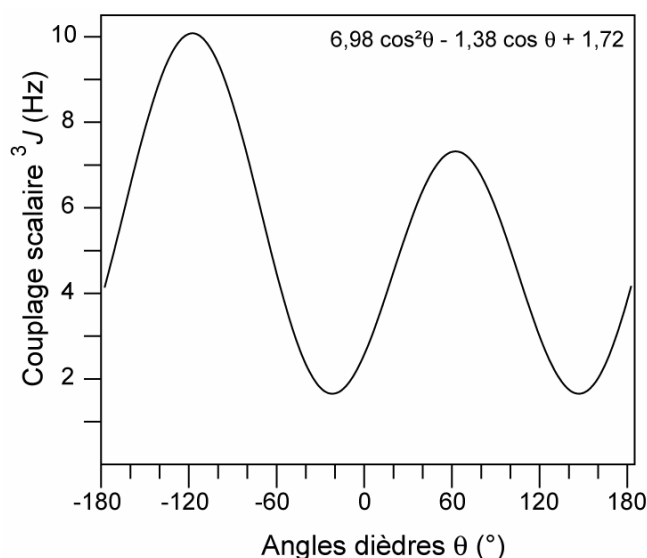


Figure 1.2 : Un exemple d'une Courbe de Karplus.

L'équation de Karplus est indiquée sur le graphe. Les valeurs A, B et C ont été paramétrées pour un couplage $^3J_{H^N H^\alpha}$ par Wang *et coll.* (Wang *et al*, 1996).

Cependant, la courbe de Karplus fait correspondre plusieurs valeurs d'angles dièdres pour une valeur donnée de couplage 3J . Pour des valeurs de couplages $^3J_{H^N H^\alpha}$ inférieures à 7 Hz, l'équation de Karplus admet 4 solutions distinctes, alors que pour des valeurs $^3J_{H^N H^\alpha}$ supérieures à 7 Hz, 2 valeurs de l'angle dièdre sont possibles. Il est donc nécessaire de compléter l'information par la mesure d'autres constantes de couplage reliées au même angle de torsion, et ce, afin de lever cette dégénérescence.

1.3.3 Les informations orientationnelles globales

1.3.3.1 Les relaxations longitudinale et transversale hétéronucléaires

La stratégie de détermination de structure de biomolécules par RMN repose sur la détermination de contraintes à courte portée, ($< 6 \text{ \AA}$), complétée par des informations angulaires contenues dans les constantes de couplages scalaires. Cette méthode a montré son efficacité sur des macromolécules de forme globulaire. En revanche, pour des protéines modulaires ou pour l'étude de complexes de protéines, l'orientation relative des différents modules dépend d'un petit nombre de contraintes définissant les zones charnières de ces systèmes. L'utilisation de contraintes à courte portée

permet d'obtenir la structure de ces molécules à plusieurs domaines avec une haute précision au niveau local. Cependant, elle ne permet pas d'avoir l'orientation relative de ces domaines entre eux. Le développement de techniques pour l'obtention de contraintes à longue portée par RMN est nécessaire. Une première technique permettant d'obtenir des informations orientationnelles à longue portée repose sur la dépendance du rapport des temps de relaxation hétéronucléaire (^{15}N) longitudinaux et transversaux (T_1/T_2), avec l'orientation du vecteur de la liaison N-H dans le tenseur d'anisotropie de diffusion rotationnelle de la molécule. Pour une protéine, avec un tenseur de diffusion de symétrie axiale ($D_{zz} \neq D_{xx} = D_{yy}$, où D_{zz} , D_{xx} et D_{yy} sont les éléments diagonaux du tenseur de diffusion), en absence de mouvements internes de large amplitude et sans élargissement des raies des signaux RMN par échange conformationnel, le rapport T_1/T_2 des noyaux ^{15}N dépend seulement de trois termes : $D_{\perp} = [D_{xx} + D_{yy}]/2$, $D_{\parallel} = D_{zz}$ (les valeurs propres du tenseur d'anisotropie de diffusion du système moléculaire) et θ , l'angle entre le vecteur de la liaison N-H et l'axe unique du tenseur de diffusion (Figure 1.3) (Brüschweiler *et al*, 1995 ; Tjandra *et al*, 1997c ; Cordier *et al*, 1998). $D_{\perp}$ et $D_{\parallel}$ sont déterminées à partir de l'ensemble des temps de relaxation ^{15}N , T_1 et T_2 .

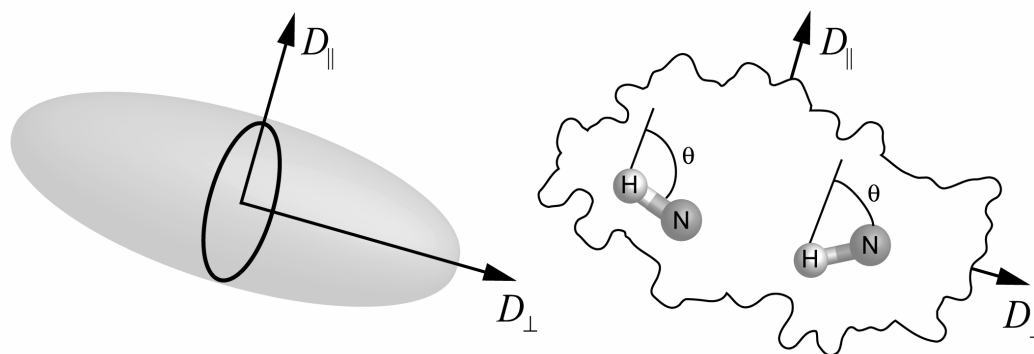


Figure 1.3 : Tenseur de diffusion et exemple d'orientation de vecteur N-H par rapport au tenseur de la molécule.

Ainsi dans l'hypothèse d'un tenseur de CSA axial symétrique, colinéaire à l'axe d'un vecteur NH, un seul rapport T_1/T_2 fourni une mesure directe de l'angle θ entre le vecteur de la liaison N-H et l'axe unique du tenseur de diffusion. Cette méthode est applicable si la molécule diffuse de manière anisotrope ($D_{\perp}/D_{\parallel} > 1$). Le rapport minimum $D_{\perp}/D_{\parallel}$ pour laquelle l'amélioration du rapport T_1/T_2 hétéronucléaire sera

utile dépend entièrement de la précision et des incertitudes des rapports T_1/T_2 mesurés. En pratique, dans les protéines globulaires, le rapport $D_{\perp}/D_{\parallel}$ est rarement supérieur à 1,5 ; l'utilisation de ce paramètre en tant que contrainte structurale précise nécessite des mesures très fines de ses paramètres de relaxation et une analyse rigoureuse. Cette approche a été complétée par d'autres paramètres orientationnels à longue portée, plus faciles à exploiter, avec notamment les couplages dipolaires résiduels présentés au paragraphe suivant.

1.3.3.2 Les couplages résiduels dipolaires (RDCs)

Une autre méthode permettant d'obtenir des contraintes à longue distance est basée sur l'analyse des constantes de couplage dipolaire résiduel. Dans les conditions classiques de la RMN liquide, la réorientation moléculaire est isotrope et le couplage dipolaire est moyenné à zéro. Cependant, pour des molécules possédant une orientation préférentielle par rapport au champ magnétique, l'interaction dipolaire n'est alors plus moyennée à zéro et peut donc être mesurée. Des systèmes moléculaires particuliers comme les protéines paramagnétiques, la myoglobine par exemple (Tolman *et al*, 1995), mais aussi les acides nucléiques ou les complexes moléculaires impliquant un oligonucléotide, peuvent s'aligner de façon suffisante de par leur forte anisotropie de susceptibilité magnétique intrinsèque. Toutefois, l'alignement obtenu reste très faible, en générale, ($< 0,1$ ‰). Une macromolécule ne possédant pas de telles propriétés peut être alignée dans l'échantillon lorsqu'elle est placée dans un milieu cristal liquide (Tjandra *et al*, 1997a). Beaucoup d'efforts ont été déployés pour développer des milieux orientants (voir la revue : Prestegard *et al*, 2001). Plusieurs de ces milieux sont utilisés avec des macromolécules biologiques : des bactériophages filamenteux, des membranes cellulaires de type rhodopsine (ou membranes pourpres), des cristallites de cellulose, du pentoxyde de vanadium, des mélanges organiques de type bromure de cétypyridinium/hexanol (ou n-alkyl-poly(éthylène-glycol)/hexanol), ou des solutions de DMPC/DHPC. Plus récemment, des nanotubes d'ADN ont été développés (Douglas *et al*, 2007) pour permettre la mesure de RDCs dans les protéines membranaires. Le choix d'un milieu repose sur les caractéristiques thermodynamiques de stabilité de la protéine (point isoélectrique,

température de dépliement/repliement), ainsi que sur la zone de pH ou de température pour laquelle la phase cristal liquide est stable.

En enregistrant deux fois une expérience pour un échantillon anisotrope (aligné) et un échantillon isotrope, le couplage dipolaire résiduel, D_{PQ} , est obtenu par différence entre la mesure en milieu orientant ($J_{PQ} + D_{PQ}$) et la mesure de couplage en milieu isotrope (J_{PQ}). Pour des interactions entre spins séparés par une liaison ^1H - ^{15}N , ^1H - ^{13}C ou ^{13}C - ^{13}C , les couplages scalaires sont assez larges et leurs signes sont connus. Le problème de la mesure des RDCs se réduit donc à la mesure de constantes de couplage. Lorsque 1 % des molécules étudiées sont dans une orientation magnétique préférentielle, la valeur maximale du couplage dipolaire résiduel mesuré pour une paire ^1H - ^{15}N est de l'ordre de 10 Hz. La valeur correspondante des couplages dans les autres paires de spins nucléaires est reportée dans le Tableau 1.2.

RDC mesuré	Distance internucléaire effective	D	J
$\text{H}_\alpha\text{-N}$	1.02 Å	10 Hz	93 Hz
$\text{C}_\alpha\text{-H}_\alpha$	2.05 Å	3.06 Hz	0-6 Hz
$\text{C}_\alpha\text{-C}_\alpha$ et $\text{C}_\alpha\text{-C}_\beta$	1.53 Å	1.85 Hz	55/35 Hz
$\text{C}_\alpha\text{-N}$	1.33 Å	1.13 Hz	15 Hz
C-H	1.12 Å	18.74 Hz	140 Hz
$\text{H}_\alpha\text{-H}_\beta$	2.5 - 4.5 Å	1-5 Hz	2-10 Hz

Tableau 1.2 : Valeurs de RDCs pour différentes paires de spins dans les protéines pour des milieux orientant standard.

Pour une paire de protons qui n'est pas couplée par un couplage scalaire, seul l'amplitude de D_{PQ} est obtenue ; d'autres expériences sont nécessaires pour obtenir le signe du couplage. Ainsi la mesure de ce type de couplages dipolaires résiduels peut également être extraite entre atomes dont les distances ne sont pas fixées. Ce qui permet d'obtenir aussi des informations de distances. J'introduirai plus précisément les RDCs dans le chapitre 5 de ce manuscrit ainsi que leur utilisation afin d'obtenir des informations à longue portée.

1.3.4 Les contraintes obtenues dans les systèmes paramagnétiques

Jusqu'à présent, toutes les informations structurales obtenues par RMN liquide décrites ci-dessus s'observent dans les macromolécules diamagnétiques, mais aussi paramagnétiques. Dans les protéines paramagnétiques, l'interaction électron-noyaux contient plusieurs sources d'informations structurales. Ces informations structurales provenant de cette interaction électron-noyaux ne peuvent pas être obtenues dans des protéines diamagnétiques. Le paramagnétisme dans une biomolécule provient d'un ou de plusieurs centres paramagnétiques formés soit par des ions métalliques, soit par des radicaux.

Dans un échantillon paramagnétique, où le temps de relaxation électronique T_e est beaucoup plus petit que le temps de corrélation de la molécule ($T_e \ll \tau_C$), il existe quatre types de contraintes structurales permettant d'avoir des informations à longue portée. Elles sont résumées dans la Figure 1.4 :

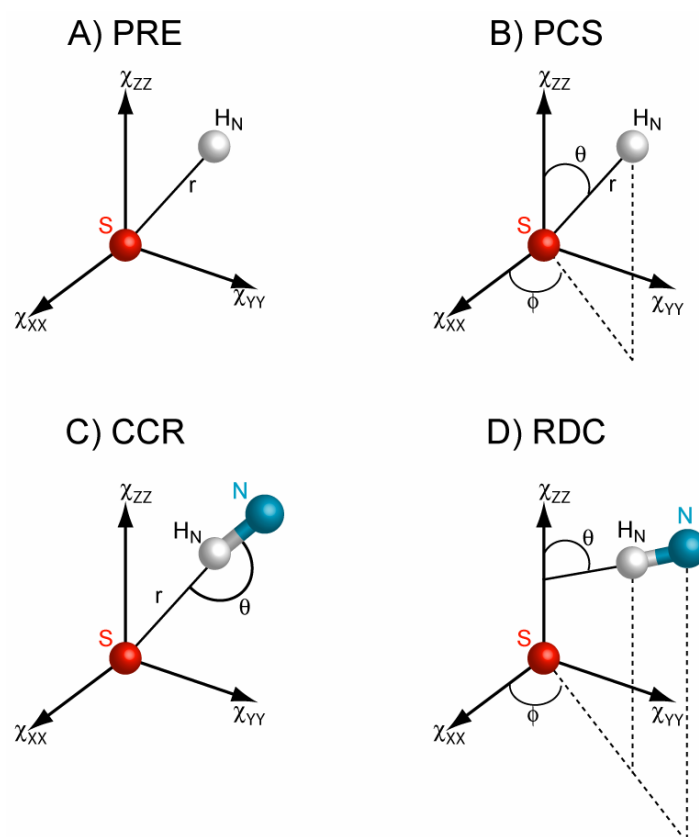


Figure 1.4 : Représentation schématique des dépendances géométriques des quatre effets paramagnétiques :

Les vitesses de relaxations longitudinale ou transversale (PRE) en A, les déplacements chimiques dipolaires (PCS) en B, la relaxation croisée Curie-Dipolaire (CCR) en C et les couplages dipolaires résiduels (RDCs) en D.

Les déplacements chimiques dipolaires, comme la relaxation croisée Curie-dipolaire, permettent d'obtenir des informations angulaires et de distances entre atomes, alors que les couplages dipolaires résiduels fournissent une information orientationnelle, si l'on s'intéresse à des atomes liés de manière covalente. Enfin les vitesses de relaxation (R_1^{DD} , R_2^{DD} ou R_2^{CS}) donnent des distances entre le centre paramagnétique et les atomes. Les couplages dipolaires résiduels ont été introduits dans § 1.2.3.2.

1.3.4.1 Les vitesses de relaxation R_1 et R_2

Dans le cas où, le temps de relaxation électronique est grand devant le temps de corrélation rotationnel ($T_e \gg \tau_c$), les vitesses de relaxation longitudinale et transversale (dû au moment paramagnétique moyen) s'écrivent sous la forme suivante (Guéron, 1975) :

$$R_1^{DD} = 2 \frac{K}{r_{IS}^6} \left[\frac{3\tau_c}{1 + \omega_I^2 \tau_c^2} \right] \quad (1.5)$$

$$R_2 = \frac{K}{r_{IS}^6} \left[4\tau_c + \frac{3\tau_c}{1 + \omega_I^2 \tau_c^2} \right] \quad (1.6)$$

$$K = \frac{1}{5} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\omega_I^2 g_e^4 \mu_B^2 S^2 (S+1)^2}{(3k_B T)^2} \quad (1.7)$$

où r_{IS} est la distance entre le métal S et le noyau I (Figure 1.4.A), S est le nombre quantique du spin électronique, τ_c est le temps de corrélation totale, ω_I est la fréquence de précession du spin nucléaire, μ_0 est la constante de perméabilité magnétique du vide, g_e est le facteur de Landé de l'électron, μ_B est le moment magnétique du proton.

Ces équations montrent que les vitesses de relaxation sont proportionnelles à l'inverse de la distance électron-noyau à la puissance six. Elles peuvent être utilisées comme contraintes de distances à la condition que la contribution diamagnétique soit

soustraite. Cette approche a été utilisée pour caractériser plusieurs complexes (Iwahara *et al*, 2003 ; 2006 ; Tang *et al*, 2006 ; 2007).

1.3.4.2 Les déplacements chimiques dipolaires (PCS)

Les déplacements chimiques dipolaires ("pseudo contact shift" PCS) résultent de la susceptibilité magnétique du métal χ anisotrope. Les déplacements chimiques dipolaires dépendent de la distance et de l'orientation du vecteur électron-atome (Figure 1.4.B) dans le tenseur de susceptibilité magnétique (χ) suivant l'équation :

$$\delta^{PC} = \frac{1}{12 \pi r_i^3} \left\{ \Delta\chi_{ax} (3 \cos^2 \theta - 1) + \frac{3}{2} \Delta\chi_{rh} \sin^2 \theta \cos 2\phi \right\} \quad (1.8)$$

où r_i est la distance entre un noyau i et le métal, $\Delta\chi_{ax}$ et $\Delta\chi_{rh}$ sont les composantes axiales et rhombiques de l'anisotropie du tenseur de susceptibilité magnétique, θ est l'angle entre le vecteur métal-noyau et l'axe Z du tenseur χ , et ϕ est l'angle entre l'axe X du tenseur χ et la projection du vecteur métal-noyau sur le plan XY (Figure 1.4.B).

Les premières tentatives pour exploiter l'Equation 1.11 date de plus de 30 ans. Bien que les déplacements chimiques dipolaires puissent donner un grand nombre de contraintes, ils n'ont jamais fournis une structure de protéine en solution par eux-mêmes. Des approches ont été développées pour inclure à la fois les PCS avec les contraintes de NOEs conventionnelles et des angles dièdres dans le calcul de structures en solution (Gochin *et al*, 1995 ; Banci *et al*, 1996 ; Hus *et al*, 2000). Dans cette voie, les PCS peuvent être utilisés comme un jeu de contraintes supplémentaires dans les protéines paramagnétiques.

Afin de bien utiliser les PCS, une analyse du déplacement chimique est nécessaire. Dans une molécule paramagnétique, le déplacement chimique peut s'écrire :

$$\delta^{obs} = \delta^{dia} + \delta^{CS} + \delta^{PCS} \quad (1.9)$$

où δ^{dia} est la contribution diamagnétique et les deux contributions, δ^{CS} et δ^{PCS} , proviennent de l'interaction hyperfine.

La séparation de ces deux contributions, nécessaires pour l'exploitation ultérieure dans l'étude structurale, n'est, en principe, pas évidente. Néanmoins, comme la contribution du "Contact Shift" décroît qualitativement à mesure que le nombre de liaisons σ liant l'ion métallique avec les noyaux augmente, cette contribution peut être considérée comme négligeable pour tous les résidus, exceptés ceux liant directement l'ion métallique.

Pour obtenir le δ^{PCS} , il est nécessaire de mesurer les valeurs δ^{dia} . Il faut donc pouvoir préparer la molécule dans des isoformes diamagnétique et paramagnétique. En pratique, la meilleure stratégie est de remplacer l'ion paramagnétique avec un métal diamagnétique de même charge. Par exemple, les ions paramagnétiques de lanthanide (III) sont substitués aux ions Ca^{2+} de protéines capables de lier le calcium. Dans ce cas, les PCS peuvent être utilisés comme contraintes dans les calculs de structures (Bentrop *et al*, 1997 ; Bertini *et al*, 2001). La forme associée au Ca^{2+} fournie une référence convenable.

La pertinence des contraintes de déplacement chimiques dipolaires est évidemment relative aux contraintes métal-noyaux, en analogie avec le cas des vitesses de relaxation dipolaire. Les déplacements chimiques dipolaires ont une dépendance avec la distance métal-noyau en r^{-3} , plutôt qu'en r^{-6} comme les vitesses de relaxation. Par conséquent, les contraintes obtenues à partir des déplacements chimiques dipolaires se propagent plus loin dans l'espace que celles obtenues avec les vitesses de relaxation. Un autre avantage est que les contraintes des PCS sont significativement plus nombreuses que les contraintes basées sur les vitesses de relaxations longitudinale et transversale.

1.3.4.3 La corrélation croisée Curie-Dipolaire

Les phénomènes de corrélation croisée contribuent aux déplacements et à la relaxation des propriétés des signaux RMN. Bien qu'ils soient connus depuis des années, c'est seulement récemment que l'on a insisté sur le fait que ces effets aient

un grand potentiel pour fournir des informations structurales et dynamiques sur les biomolécules en solution. Ils permettent d'entreprendre des études de macromolécules biologiques de grande taille. L'interaction entre les spins nucléaires et le moment magnétique statique de l'électron (induit par le champ magnétique) donne lieu à une relaxation non négligeable appelé "relaxation de Curie" (Guéron, 1975) ou "relaxation de susceptibilité" (Vega *et al*, 1976). Elle est décrite dans l'Equation 1.10. En toute rigueur, la susceptibilité magnétique doit être considérée comme une grandeur tensorielle et non comme un scalaire. Cependant, tant que l'anisotropie de cette susceptibilité reste faible par rapport à sa valeur moyenne, ses effets sur la relaxation de Curie reste négligeable (Vega *et al*, 1976). L'interaction entre le spin de Curie et le spin nucléaire est modulée par les mouvements rotationnels de la molécule, comme l'est l'interaction dipôle–dipôle entre spins nucléaires (DD). En conséquence, les effets de corrélation croisée entre l'interaction de Curie (*Curie*) et l'interaction dipôle–dipôle internucléaire (*DD*) peuvent donner lieu à des phénomènes de relaxation. Pour des molécules U-[¹⁵N], et dans le cas simple d'une molécule rigide avec une rotation isotrope ainsi qu'un tenseur de susceptibilité magnétique χ , la corrélation croisée entre l'interaction spin de Curie (S) - spins nucléaires (H) et l'interaction dipolaire entre deux spins (H-N) des groupes amide NH fournit des informations sur l'orientation relative entre les vecteurs S-H et N-H, ainsi que sur la distance S-H. Ainsi, sur la base de l'Equation 1.10 :

$$\Gamma_{H,H-N}^{Curie,DD} = \frac{2}{5} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\hbar \gamma_H^2 \gamma_N}{\langle r_{H-N}^3 \rangle} \chi B_0 \tau_c \frac{(3 \cos^2 \theta - 1)}{r_{H-S}^3} \quad (1.10)$$

$$\chi = \frac{\gamma_M^2 \hbar^2 S(S+1)}{3k_B T} \quad (1.11)$$

où θ est l'angle défini par le triangle N-H-S (Figure 1.4.C), r_{H-N} et r_{H-S} représentent respectivement la distance entre le proton et l'azote ainsi que la distance entre le proton amide et le spin électronique (lequel est localisé dans l'ion métallique), τ_c est le temps de corrélation de la molécule, μ_0 est la constante de perméabilité magnétique du vide, B_0 est la constante du champ magnétique. La susceptibilité magnétique χ est considérée comme un tenseur isotrope et est définie par l'Equation

1.11, où S est le nombre quantique du spin de Curie. γ_H , γ_N et γ_S sont les rapports gyromagnétiques du proton, de l'azote et du spin électronique S , $\hbar$ est la constante de Planck divisée par 2π , k_b est la constante de Boltzmann et T la température de l'échantillon.

$\Gamma_{H,H-N}$ est la contribution de corrélation croisée de la relaxation R_2 . Cette contribution à la relaxation R_2 est responsable de différentes largeurs de raies des composantes individuelles α et β , du doublet du proton résultant du couplage dipolaire NH selon :

$$R_2^{\alpha,\beta} = R_2^{all} \pm \Gamma_{H,H-N} \quad (1.12)$$

où R_2^{all} est la somme de toutes les autres contributions de R_2 . La constante de vitesse de relaxation $\Gamma_{H,H-N} = \Gamma_{H,H-N}^{CSA,DD} + \Gamma_{H,H-N}^{Curie,DD}$ est mesurée à partir du rapport des intensités des composantes du doublet du H_N . $\Gamma_{H,H-N}^{Curie,DD}$ ne peut pas être séparé de la corrélation croisée entre l'anisotropie de déplacement chimique du proton (CSA) et le couplage dipolaire HN (DD), laquelle est indépendante du paramagnétisme, $\Gamma_{H,H-N}^{CSA,DD}$. Pour différencier les deux phénomènes, $\Gamma_{H,H-N}^{CSA,DD}$ est mesurée sur une référence diamagnétique de la protéine étudiée de la même manière que pour les déplacements chimiques dipolaires. La faisabilité de cette approche a été montrée pour l'obtention de nouvelles informations structurales à longue portée (Boisbouvier *et al*, 1999).

1.4 Détections de nouvelles informations à longue portée.

Nous avons vu précédemment que, pour des protéines non globulaires, il demeure important d'obtenir des informations à longue portée (§1.2.3.1). Les méthodes utilisées, de manière routinière, font surtout appel à des informations orientationnelles, rapport T_1/T_2 hétéronucléaire, et à des couplages dipolaires résiduels (RDCs) entre 2 atomes liés de manière covalente (liaisons N-H, C-H). Toutefois, l'obtention de nouvelles informations translationnelles est très importante pour la détermination de structures de complexes protéiques, de protéines modulaires

et d'acides nucléiques. Parmi l'ensemble des contraintes déterminables par RMN pour les protéines diamagnétiques, deux approches ont été envisagées et font partie du travail de cette thèse : les NOEs et les RDCs ^1H - ^1H .

La première approche a été montrée comme particulièrement utile pour obtenir des informations locales ($< 6 \text{ \AA}$). La principale limitation pour obtenir des informations à longue portée à partir des NOEs est due à la diffusion de spins. Une manière de limiter ces chemins de relaxation dipolaire est la deutération des sites non échangeables. Il a été montré que des NOEs entre des protons amides séparés par des distances atteignant $8 \text{ \AA}$ sont détectables (Koharudin *et al*, 2003). Toutefois, ces NOEs restent des contraintes semi-quantitatives et sont utilisés par l'approche standard (vue précédemment § 1.3.1.1), lors du calcul de structure. Diminuer la densité de protons pour limiter encore plus la diffusion de spins est une solution pour obtenir des informations plus précises et peut permettre de détecter des NOEs plus faibles.

Généralement, l'information obtenue par les RDCs est extraite à partir des couplages dipolaires résiduels entre deux spins séparés par une liaison. Il est possible d'observer de manière précise des RDCs entre paires d'atomes non liés de manière covalente. En raison du rapport gyromagnétique élevé du proton, les couplages dipolaires ^1H - ^1H sont les plus grands. La mesure de couplage ^1H - ^1H , au-delà de grandes distances, demeure difficile, même si les couplages dipolaires résiduels ne sont pas affectés par un effet équivalent à la diffusion de spins. Un couplage dipolaire entre une paire de protons éloignés est affecté par les couplages avec leurs plus proches voisins et n'est pas détectable. Pour résoudre ce problème, il est possible d'utiliser des échantillons perdeutérés (comme pour la mesure de NOEs à longue portée) et des impulsions sélectives (pour découpler les interactions avec la plupart des protons les plus proches). En éliminant ainsi les plus proches voisins, des couplages directs entre ^1H - ^1H , séparés jusqu'à $7 \text{ \AA}$, ont été mesurés précisément dans des protéines perdeutérées (Wu *et al*, 2002 ; Meier *et al*, 2003). Egalement, ceux-ci ont été mesurés dans les acides nucléiques (couplages entre protons séparés par $12 \text{ \AA}$) (Boisbouvier *et al*, 2003).

Une approche possible pour détecter des distances à plus longue portée (avec la même sensibilité vis à vis des sondes RMN et les champs magnétiques actuels) est l'utilisation d'échantillons perdeutérés et spécifiquement protonés sur quelques sites choisis, afin de diminuer la densité de protons. Le coût d'un échantillon perdeutéré n'est pas négligeable. Ainsi, produire des échantillons perdeutérés et spécifiquement protonés sur certains sites à moindre coût est indispensable en vue d'une utilisation dans les approches standard de détermination de structure.

Le travail de cette thèse a consisté à développer des protocoles de marquages isotopiques spécifiques sur les méthyles dans les protéines, (Chapitre 2), et ce en vue d'utiliser les propriétés avantageuses de ces échantillons dans l'observation de nouvelles informations à longue portée, dans des systèmes de taille modérés à partir de NOE (Chapitre 3) et de RDCs ^1H - ^1H (Chapitre 5). Les résultats obtenus par l'approche développée sur la mesure de NOEs à longue portée ont permis d'envisager son utilisation dans des systèmes supramoléculaires et elle sera développée dans le chapitre 4.

Chapitre

2

La protonation spécifique des méthyles dans les protéines

2.1 Introduction

L'étape initiale pour résoudre la structure tridimensionnelle d'une protéine par RMN est l'attribution séquentielle des fréquences de résonances du squelette peptidique. Cette étape est d'autant plus difficile que la taille de la molécule augmente. En effet, la gamme de déplacement chimique des différents atomes d'une protéine est d'environ 10 ppm pour le ^1H , 30 ppm pour le ^{15}N , 30 ppm pour le C_α et 10 ppm pour le C_O . Comme le nombre de résonances dans une gamme de déplacements chimiques est proportionnel à la taille de la protéine, le degré de superposition et l'ambiguïté d'attribution le sont également. Le marquage isotopique uniforme ^{15}N et ^{13}C des protéines permet de simplifier énormément cette étape d'attribution en distribuant les signaux RMN dans plusieurs dimensions. L'attribution séquentielle est alors réalisée à partir d'expériences multidimensionnelles triple-résonance utilisant les couplages scalaires pour transférer l'aimantation à travers les liaisons du squelette peptidique (Marion *et al*, 1989a ; Ikura *et al*, 1990 ; Kay *et al*, 1990). Parmi toutes les stratégies de marquages isotopiques des protéines, le marquage uniforme en deutérium (ou perdeutération), c'est-à-dire le remplacement de l'isotope ^1H des hydrogènes non échangeables avec le solvant par l'isotope ^2H des atomes d'hydrogène, est une stratégie permettant d'augmenter simultanément la sensibilité et de diminuer le nombre de résonances dans les spectres RMN (Crespi *et al*, 1968 ; Markley *et al*, 1968 ; Kalbitzer *et al*, 1985 ; Griffey *et al*, 1985 ; Le Master *et al*, 1988 ; Torchia *et al*, 1988). Du fait d'un rapport gyromagnétique significativement plus faible pour le ^2H , en comparaison avec celui du ^1H ($\gamma[^2\text{H}]/\gamma[^1\text{H}] = 0,15$), le remplacement des protons par des deutérons diminue les contributions des interactions dipolaires ^1H - ^1H et des couplages scalaires ^1H - ^1H sur la largeur de raie des signaux des protons amides (H_N). Associé à des développements méthodologiques, tels que les expériences de type TROSY (Pervushin *et al*, 1997 ; Salzmann *et al*, 1999), le marquage au ^2H a permis de repousser les limites en taille et la complexité des systèmes de protéines étudiables par RMN. Ainsi, cette méthode a permis d'attribuer la chaîne principale d'une protéine monomérique de masse moléculaire de 82 kDa, la Malate Synthase G (Tugarinov *et al*, 2002 ; 2003a ; 2003b ; 2005a).

L'incorporation de ^2H dans les protéines a été employée depuis de nombreuses années pour simplifier les spectres et pour améliorer les propriétés de relaxation des protons restants. La deutération peut également améliorer la qualité des expériences NOESY. En particulier, la substitution de protons non-échangeables par des deutérons a comme conséquence des gains impressionnants de sensibilité dans la détection de corrélation H_N - H_N dans des spectres NOESY (Le Master *et al*, 1988 ; Torchia *et al*, 1988). Démontrée sur des protéines perdeutérées (Venters *et al*, 1995 ; Mal *et al*, 1998), cette approche peut aussi être exploitée sur des protéines ou des peptides en solution dans de l'eau lourde, D_2O , si certains sites aliphatiques sont spécifiquement protonés (Tsang *et al*, 1990 ; Reisman *et al*, 1991).

Plus récemment, en combinant la stratégie de perdeutération et l'utilisation d'aimants modernes opérant à des champs magnétiques élevés, des NOEs entre H_N séparés par des distances allant même jusqu'à 8 Å ont pu être détectés dans un échantillon totalement perdeutééré (Koharudin *et al*, 2003). Le remplacement de tous les protons non-échangeables par du deutérium dans les protéines réduit la densité de protons d'environ 85 %. En diminuant la densité de protons, ce type de marquage a permis de mesurer précisément des couplages dipolaires résiduels entre $^1\text{H}_\text{N}$ - $^1\text{H}_\text{N}$ séparés jusqu'à 7 Å (Wu *et al*, 2002 ; Meier *et al*, 2003).

D'autres sites que les protons H_N peuvent être utilisés pour obtenir des contraintes spatiales indispensables à la détermination de structures de protéines. Dans ce chapitre, les différentes méthodes de marquages isotopiques de protéines, uniformément marquées au deutérium (^2H) et spécifiquement protonées sur certains sites, seront présentées. Pour l'étude par RMN des structures de protéines, les méthyles sont de bons candidats pour les sites de protonation spécifique. Dans la suite de ce chapitre, je vais détailler les différents protocoles que j'ai mis en place au laboratoire pour la préparation d'échantillons deutérés et spécifiquement protonés sur les méthyles. Il s'agira, en particulier, de ceux mis en place sur le méthyle δ_1 des Isoleucines, des méthyles, γ des Valines et δ des Leucines, ainsi que du méthyle β des Alanines. Les applications de ces différents types de marquages spécifiques, permettant la détection de NOEs ou de RDCs à longue portée ainsi que l'étude des assemblages biomoléculaires de grande taille, seront exposées aux chapitres 3, 4 et 5.

2.2 Les méthodes de marquages sélectifs

2.2.1 Le marquage par inhibition

Nous allons nous intéresser ici aux stratégies de protonations spécifiques de certaines positions dans un échantillon de protéine deutérée sur tous les autres sites. La méthode la plus simple est de produire des cultures bactériennes dans un milieu ne contenant aucune source de protons, hormis l'acide aminé que l'on souhaite marquer. Dans la majorité des cas, la voie métabolique de biosynthèse de l'acide aminé en question est régulée, au niveau d'une ou de plusieurs enzymes de cette voie, par la présence de l'acide aminé. C'est un rétrocontrôle ou une régulation par inhibition (Lu *et al*, 1976). En incluant dans le milieu de culture un acide aminé à une concentration suffisamment élevée (Tableau 2.1), il est donc possible d'inhiber sa biosynthèse. La bactérie incorpore alors l'acide aminé exogène. Ainsi, l'incorporation de l'acide aminé biosynthétisé par la bactérie est réduite au minimum. La Figure 2.1 représente un exemple de marquages spécifiques ^{15}N par inhibition de différents acides aminés (Lee *et al*, 1995).

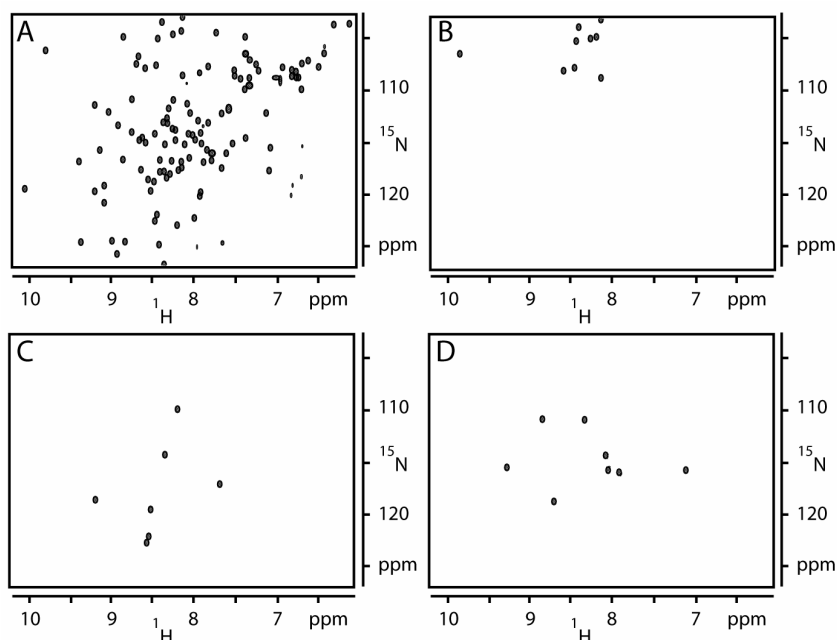


Figure 2.1 : Spectre de corrélation HSQC ^1H - ^{15}N de la protéine E2.

Le spectre de la protéine E2 U- ^{15}N qui sert de référence est représenté en A. Les autres spectres représentent la protéine E2 spécifiquement marquée ^{15}N sur les Glycines en B, les Leucines en C, et les Phénylalanines en D (Lee *et al*, 1995). Pour chacun des marquages spécifiques, l'acide aminé ^{15}N est ajouté à sa concentration d'inhibition dans le milieu, saturé par les 19 autres acides aminés non marqués à leur concentration d'inhibition.

Acide Aminé		Inhibition (mg/L) ^a	Acide Aminé		Inhibition (mg/L) ^a
Alanine	Ala	400 (800) ^b	Leucine	Leu	100 (200) ^b
Arginine	Arg	400	Lysine	Lys	100
Asparagine	Asn	100	Méthionine	Met	250 (500) ^b
Acide Aspartique	Asp	250	Phénylalanine	Phe	50
Cystéine	Cys	50	Proline	Pro	100
Glutamate	Glu	400 (800) ^b	Sérine	Ser	1600
Glutamine	Gln	400	Thréonine	Thr	100
Glycine	Gly	400	Tryptophane	Trp	50
Histidine	His	100	Tyrosine	Tyr	100
Isoleucine	Ile	100 (200) ^b	Valine	Val	100

Tableau 2.1 : Concentration en acides aminés pour les marquages sélectifs par inhibition dans le milieu. ^a d'après (Griffey *et al*, 1985) ; ^b d'après (Ramesh *et al*, 1994).

Dans le cas de marquages spécifiques ¹⁵N, la présence des transaminases peut engendrer des fuites métaboliques depuis le squelette de l'acide aminé protoné vers les dix-neuf autres acides aminés (Figure 2.2.A).

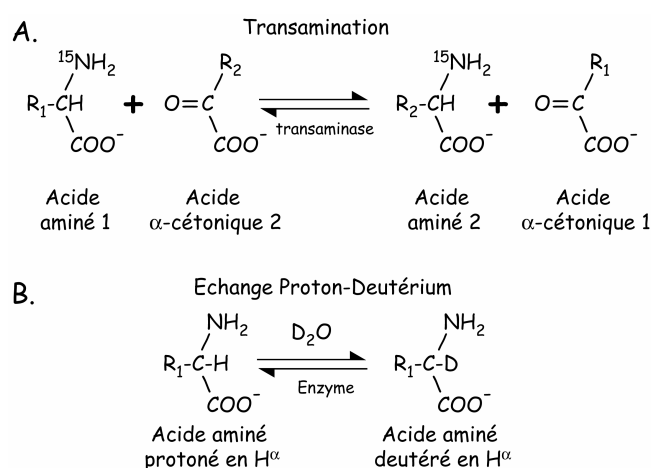


Figure 2.2 : En A) Schéma de transamination et en B) Schéma de l'échange entre le proton en α de l'acide aminé et le deutérium du solvant.

Au cours de ce mécanisme, il y a formation d'un carbanion avec départ du H $_{\alpha}$. Si le but recherché est la protonation spécifique d'un proton en alpha d'un acide aminé donné, les transaminases catalyseront son échange avec les molécules de D₂O du

milieu de culture. Il en résultera une diminution du taux de marquage spécifique de l'acide aminé recherché (Figure 2.2.B). Pour la majorité des protons des chaînes latérales protonées, les transaminases ne jouent pas de rôle et les fuites métaboliques sont faibles.

Ce marquage par inhibition simple présente l'inconvénient de ne pas être applicable à tous les acides aminés. Par exemple, seuls les acides aminés : H, G, C, W, F, Y, P, R, M, K, I, V, L peuvent être utilisés pour marquer spécifiquement leurs chaînes latérales. En effet, certains acides aminés sont recyclés dans la synthèse d'autres acides aminés (comme le Glutamate, par exemple). Ainsi, les protons de l'acide aminé considéré donnent un marquage secondaire sur d'autres acides aminés. Afin de limiter les fuites métaboliques, l'anabolisme des acides aminés doit être inhibé par la complémentation du milieu avec tous les dix-neuf autres acides aminés deutérés à leur concentration d'inhibition, Tableau 2.1.

Une stratégie alternative consiste à utiliser des acides aminés en fin de voies métaboliques, Figure 2.3. Ainsi, ces résidus peuvent être marqués avec une fuite limitée des protons vers d'autres acides aminés (Smith *et al*, 1996).

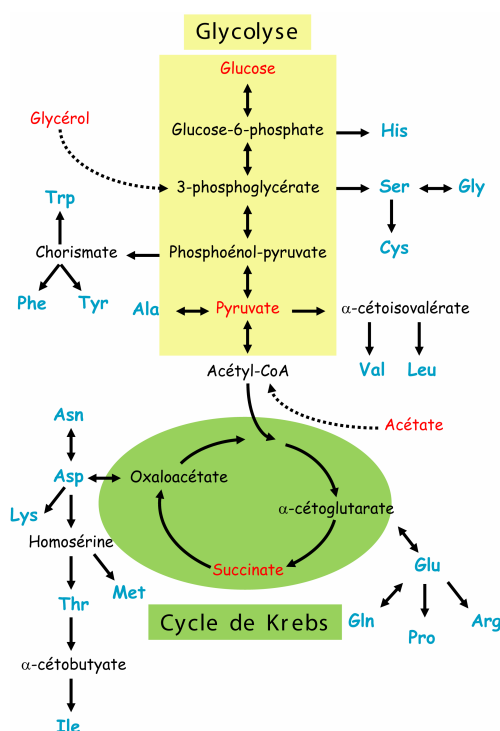


Figure 2.3 : Schéma simplifié de la biosynthèse des acides aminés dans *E. coli*.

Les acides aminés sont notés en bleu et les sources de carbone, utilisables comme uniques sources de carbone lors d'une expression en milieu minimum, en rouge.

2.2.2 Les souches auxotrophes

Alternativement au marquage par inhibition, l'utilisation de souches bactériennes auxotrophes permet de marquer les acides aminés donnant un marquage secondaire. Le phénotype d'auxotrophie pour un acide aminé est dû à une (des) mutation(s) dans un ou plusieurs gènes de son anabolisme. Ainsi, l'utilisation de souches bactériennes auxotrophes limite la dilution isotopique et permet d'optimiser l'incorporation de l'acide aminé (Tableau 2.2). Cependant, les autres voies de biosynthèse sont toujours présentes et ne limitent pas le recyclage de l'acide aminé considéré vers les autres acides aminés (Tableau 2.2). Comme pour le marquage par inhibition, le milieu doit être supplémenté avec les autres acides aminés deutérés afin d'inactiver leur anabolisme. Donc, le coût d'une culture par ces deux méthodes d'un échantillon deutéré et spécifiquement marqué sur un type d'acide aminé n'est pas négligeable. Aujourd'hui, de nombreuses souches auxotrophes sont disponibles pour le marquage sélectif. Toutefois, la construction de génotypes complexes et la faible croissance bactérienne qui lui est liée sont une limite de cette méthode.

Acide Aminé			Auxotrophe (mg/L) ^a		
Alanine	Ala	600	Leucine	Leu	60
Arginine	Arg	380 / E	Lysine	Lys	160
Asparagine	Asn	200 / D	Méthionine	Met	150
Acide Aspartique	Asp	100	Phénylalanine	Phe	50-75
Cystéine	Cys	<i>n.d.</i>	Proline	Pro	<i>n.d.</i>
Glutamate	Glu	550	Sérine	Ser	500 / G
Glutamine	Gln	380 / E	Thréonine	Thr	150
Glycine	Gly	250-375	Tryptophane	Trp	35
Histidine	His	<i>n.d.</i>	Tyrosine	Tyr	35
Isoleucine	Ile	50	Valine	Val	60-100

Tableau 2.2 : Concentration en acides aminés pour les marquages sélectifs par inhibition ou par souches auxotrophes dans le milieu.

^a d'après (Muchmore *et al*, 1989). La lettre en capitale, à droite de la concentration d'un acide aminé, correspond à l'acide aminé et à sa concentration ajoutée ; *n.d.* non déterminé.

2.2.3 La deutération sélective par précurseur d'acides aminés

Cette approche est une variante du marquage par inhibition. La surexpression de la protéine par *E. coli* se fait avec l'unique source d'hydrogène ^1H , un (des) précurseur(s) spécifiquement protoné(s) pour un acide aminé donné, qui est (ou sont) supplémenté(s) au milieu. Toutes les autres sources d'hydrogène doivent être deutérées (^2H). Ainsi, la culture est effectuée en eau lourde, D_2O , et la source de carbone est deutérée (glycérol, glucose, acétate ou pyruvate Figure 2.3). On peut, par cette méthode, marquer spécifiquement certains sites de quelques acides aminés. Le choix des intermédiaires doit être basé sur plusieurs critères :

(i) L'existence d'une voie biosynthétique d'incorporation, sans réaction d'échange entre les deutérons du solvant de la culture et les protons du précurseur.

(ii) La disponibilité et le coût du précurseur métabolique spécifiquement protoné ou la difficulté de le produire sous forme spécifiquement protoné.

(iii) Le précurseur métabolique doit être en fin de voie métabolique irréversible. Il ne sera ainsi pas nécessaire de saturer la culture avec les dix-neuf autres acides aminés deutérés (en opposition au marquage par inhibition ou par souches auxotrophes).

(iiii) Le dernier point est l'efficacité de l'assimilation du précurseur dans *E. coli*. La pénétration du précurseur métabolique dans la cellule est primordiale pour être utilisée dans les voies de biosynthèse des acides aminés.

D'après ces critères, cette méthode ne peut être appliquée qu'à un nombre restreint d'acides aminés. Comme on peut le voir sur la Figure 2.3, les acides aminés aromatiques Phe, Tyr, et Trp ont comme précurseur commun le chorismate (Rajesh *et al*, 2003), et les acides aminés contenant des groupements méthyles Val, Leu, et Ile(δ_1) ont, eux, plusieurs précurseurs (Gardner *et al*, 1997a ; Goto *et al*, 1999). L'un des avantages à choisir ces acides aminés est qu'ils font partis du cœur hydrophobe des protéines solubles et peuvent, ainsi, fournir des informations structurales importantes sur leur repliement (Gardner *et al*, 1997b). La limite de cette méthode est

de ne pouvoir produire des protéines spécifiquement protonées que pour un nombre restreint d'acides aminés.

2.2.4 Le marquage *in-vitro*

La dernière méthode présentée ici est très différente des approches décrites précédemment. En effet, jusqu'à présent, les approches présentées étaient basées sur des productions *in vivo* (cellules bactériennes). Dans l'approche *in vitro*, la méthode "Cell-free", mise au point dans les années 70-80 (Spirin *et al*, 1988) et largement développée au Japon, utilise des extraits cellulaires de différents organismes contenant toute la machinerie nécessaire à la transcription de l'ADN en ARN et à la traduction de l'ARN messenger en protéine (Kigawa *et al*, 1991 ; 1995 ; Vinarov *et al*, 2004 ; 2005 ; 2006 ; Toriza *et al*, 2004 ; Yabuki *et al*, 1998 ; Klammt *et al*, 2006 ; Sawasaki, *et al*, 2002 ; Busso *et al*, 2003 ; Torizawa *et al*, 2004 ; Matsuda *et al*, 2007). L'un des avantages des systèmes de marquages *in-vitro* est de pouvoir synthétiser des protéines qui ne s'expriment pas par des méthodes *in-vivo*, du fait de leur toxicité pour les cellules hôtes, de leur dégradation par des phénomènes de protéolyse, ou encore de leur agrégation sous forme de corps d'inclusion.

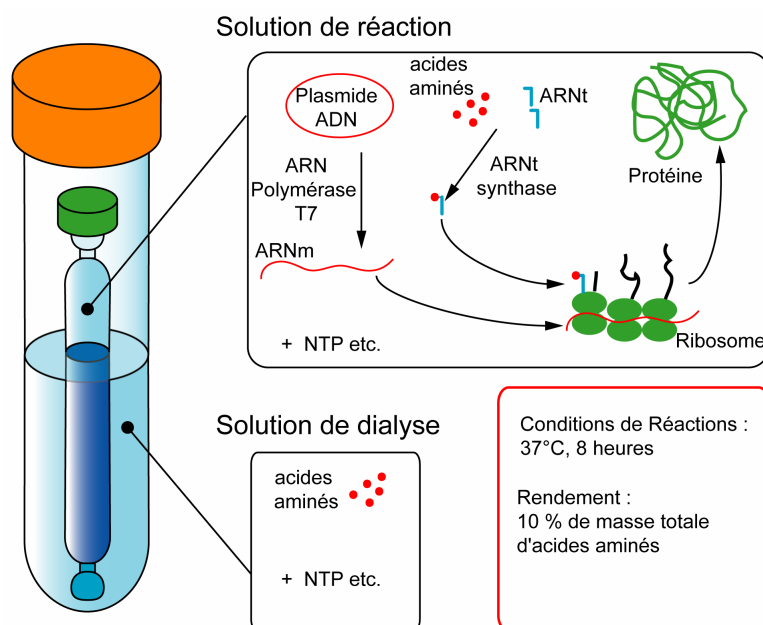


Figure 2.4 : Schéma de l'approche de synthèse *in-vitro* de protéine.

Dans les années 90, il a été montré que des acides aminés marqués isotopiquement peuvent être incorporés avec une grande efficacité pour des applications RMN. Le

système idéal de marquages *in-vitro* pour la spectroscopie RMN est un système robuste, capable de produire des quantités de l'ordre du milligramme d'échantillons de protéines marquées, reproductibles et à un coût raisonnable. Dans les systèmes *in-vitro* optimisés, la masse de protéines purifiées correspond à environ 10 % de la masse totale d'acides aminés marqués introduits. En comparaison, l'obtention de 30 mg/L de protéines marquées pures, lors d'une surexpression dans *E. coli*, à partir de trois grammes de matériel isotopiquement enrichi (glucose- ^{13}C , $^{15}\text{NH}_4\text{Cl}$, milieux riches marqués, ...) constitue un bon résultat. Le rendement du marquage *in-vitro* est donc bien plus élevé. De plus, cette approche s'accompagne d'une fuite minimale du marquage. En effet, les enzymes (transaminases,...) ne sont présentes que sous forme de traces dans la solution *in-vitro* et des inhibiteurs peuvent être utilisés pour les inactiver (Morita *et al*, 2004). Les systèmes de marquages *in-vitro* sont maintenant considérés comme des voies alternatives aux méthodes conventionnelles d'expression *in-vivo*. Actuellement, les systèmes commerciaux de kit de marquage *in-vitro* restent relativement onéreux. Une approche plus économique est la préparation de ses propres extraits de *E. coli* pour la synthèse de protéine (Torizawa *et al*, 2004).

Comme chaque acide aminé est ajouté dans le milieu, les marquages *in-vitro* présentent une grande flexibilité et un nombre important de marquages possibles. Récemment, dans l'objectif de déterminer la structure de protéines globulaires de manière automatisée, il a été proposé d'utiliser des acides aminés stéréospécifiquement marqués (SAIL = "Stereo-array isotope labelling"). En effet, un seul des atomes d'hydrogène des méthylènes ou des méthyles est présent sous forme ^1H et les autres sont remplacés par ^2H (Figure 2.5, (Kainosho *et al*, 2006). Ainsi La technique de SAIL permet d'améliorer grandement la résolution des spectres RMN, en rendant possible l'analyse automatique des spectres NOESY et le calcul de structure (Kainosho *et al*, 2006). Cependant, même pour des rendements de l'ordre de 10 % de protéines pures, le coût de ces acides aminés stéréospécifiquement marqués reste une limitation. Il faut compter de l'ordre de 3000 € d'acides aminés marqués pour préparer un échantillon RMN souhaitable afin d'appliquer l'approche SAIL.

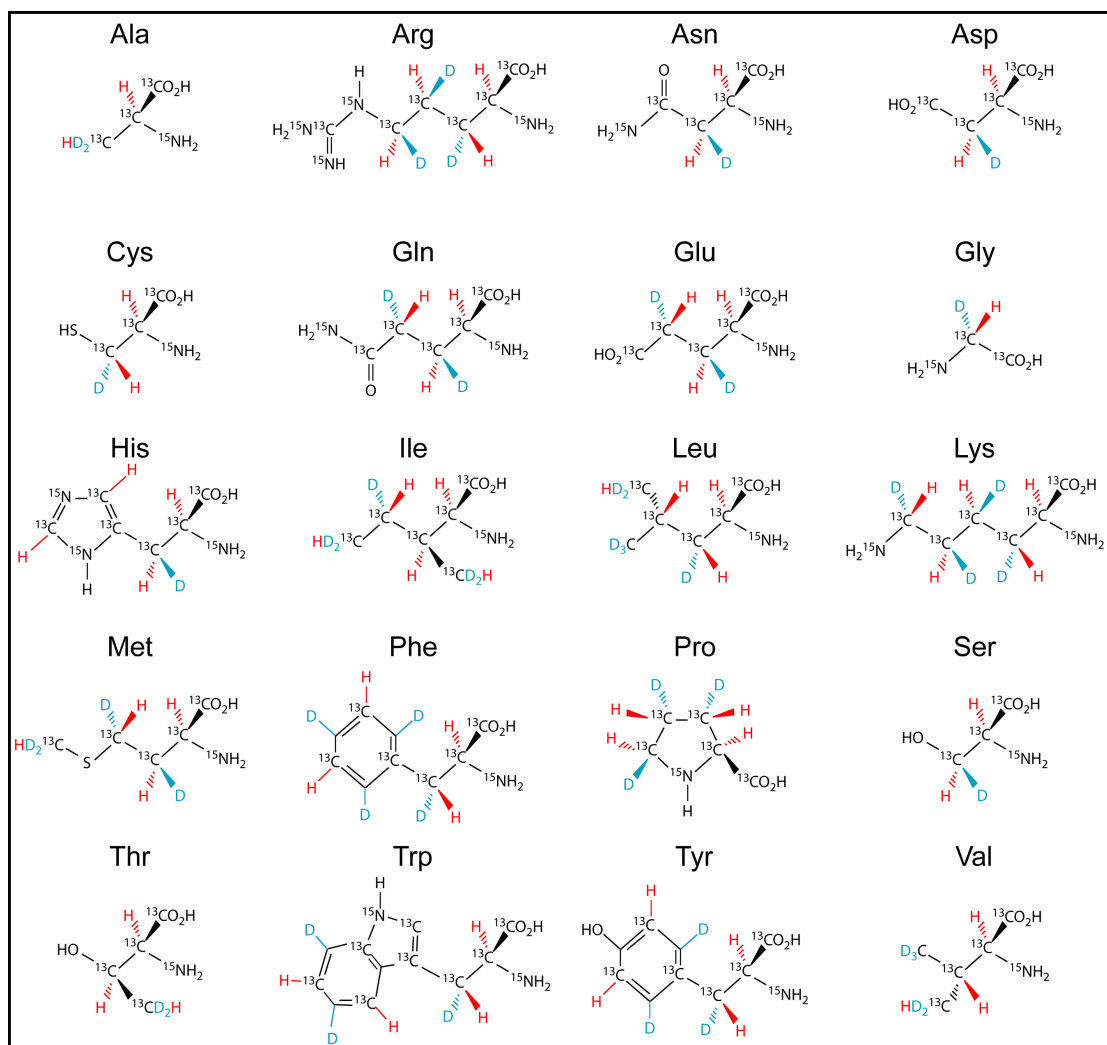


Figure 2.5 : Schéma des acides aminés stéréo-spécifiquement marqués pour la technique de SAIL.

Les acides aminés sont marqués uniformément ^{13}C et ^{15}N . Les protons sont notés en rouge et les deutériums en bleu.

2.3 Choix du site de protonation spécifique

Afin de détecter des informations structurales à longue portée, il est important de rechercher le site le plus intéressant pour introduire des protons de manière sélective dans une protéine perdeutérée. Les méthyles des chaînes latérales bénéficient de trois propriétés majeures pour la RMN :

(i) Dans la limite des grandes distances ($d_{HH}^{\text{intra-méthyle}} \ll d_{HH}^{\text{inter-méthyle}}$), l'aimantation transférée entre deux méthyles est multipliée par neuf, du fait de la multiplicité du méthyle, par rapport à l'intensité du signal qui serait transféré entre deux protons, séparés par la même distance. Un premier facteur trois provient de

l'aimantation de Boltzmann, disponible pour un méthyle, en comparaison avec un proton unique. Un autre facteur trois vient du fait que l'aimantation d'un méthyle peut être transférée vers chacun des trois protons du second méthyle.

(ii) Dans les protéines spécifiquement protonées, les protons des méthyles sont caractérisés par des vitesses de relaxations transversale (Tugarinov *et al*, 2003c ; 2004a ; 2007) et longitudinale favorables, donnant lieu à une bonne résolution et à un bon rapport signal sur bruit dans les spectres RMN.

(iii) Une protonation spécifique s'accompagnera d'une densité de protons faibles. Les interactions dipolaires entre protons voisins seront ainsi considérablement réduites, permettant de limiter la contribution de la diffusion de spins dans la détection d'effets NOE (Koharudin *et al*, 2003) et de limiter les phénomènes de troncations dipolaires lors de l'observation de couplages dipolaires résiduels (RDCs) (Wu *et al*, 2002 ; Meier *et al*, 2003). Ce faible taux de protonation est un prérequis indispensable pour l'extraction par RMN de contraintes structurales précises entre protons éloignés.

Tous les méthyles des différents acides aminés (Ala, Val, Leu, Ile, Thr et Met) peuvent potentiellement être marqués spécifiquement. Le marquage *in-vitro* est une solution pour protoner spécifiquement ces méthyles avec toutes les combinaisons possibles des acides aminés précédents, pourvu qu'ils soient disponibles sous formes spécifiquement protonées. Cependant, cette approche reste onéreuse. Une stratégie *in-vivo* efficace, avec un coût plus raisonnable, a été proposée pour trois acides aminés : ceci concerne les méthyles (γ_1 / γ_2) de la Valine, (δ_1 / δ_2) de la Leucine et (δ_1) de l'Isoleucine (Gardner *et al*, 1997a ; Goto *et al*, 1999). J'ai implémenté et optimisé cette stratégie au laboratoire ; les protocoles de production et les résultats sont présentés dans ce chapitre. Sur les trois acides aminés restants, l'Alanine est un très bon candidat pour une protonation spécifique. En effet, il est l'acide aminé le plus représenté dans les protéines (13 % contre moins de 10 % pour les dix-neuf autres acides aminés). Un protocole a été proposé par Isaacson *et coll.* (Isaacson *et al*, 2007). Cependant, il ne permet pas d'avoir un taux de protonation spécifique en position β des Alanines aussi élevé que les marquages spécifiques des Isoleucines, des Valines et des Leucines ; ce taux est inférieur à 50 % et n'est pas compatible avec

la mesure de paramètres structuraux. Je présenterai, à la fin de ce chapitre, un autre protocole permettant le marquage spécifique des Alanines.

2.4 Protonation spécifique du méthyle δ_1 des Isoleucines

2.4.1 Choix du méthyle δ_1 des Isoleucines

L'analyse des voies métaboliques des acides aminés montre que les méthyles de l'Ala, Val, Leu et Ile- γ_2 sont des dérivés du pyruvate. Par conséquent, la surexpression d'une protéine dans *E. coli* ayant le pyruvate comme source de carbone et de proton conduit à la protonation spécifique des groupements méthyles de l'Ala, Val, Leu et Ile- γ_2 (Rosen *et al*, 1996). Ce type de marquage s'accompagne d'une incorporation des protons sur les positions C_α et C_β de l'ordre de 5 %. De plus, les niveaux de protonation s'échelonnent de 80 % pour les méthyles des Leucines et à 40 % pour les Alanines, conduisant à la présence d'un mélange d'isotopomères sur les méthyles (CH_3 , CH_2D , CHD_2 , CD_3). Ces différents isotopomères sont dus à l'échange des protons du pyruvate avec les deutérons du solvant utilisé pendant la culture (temps caractéristique d'échange *in-vivo* de l'ordre de deux jours dans *E. coli* ; Rosen *et al*, 1996). Compte tenu des déplacements chimiques distincts pour chacun de ces isotopomères, la sensibilité et la résolution des spectres RMN se trouvent affectées.

Les protocoles de marquages utilisant des précurseurs d'acides aminés, pour protoner les méthyles des Valines et des Leucines à partir de l' α -cétoisovalérate (Goto *et al*, 1999), ou le méthyle δ_1 des Isoleucines à l'aide de l' α -cétobutyrate (Gardner *et al*, 1997a), sont beaucoup plus spécifiques et ne sont pas affectés par un échange H/D avec les deutérons du solvant. Cependant, que ce soit pour la Valine ou pour la Leucine, les deux méthyles sont très proches et, localement, la densité de protons sera élevée. Ceci peut être contourné par l'utilisation de précurseurs marqués spécifiquement sur l'un des deux méthyles (un des méthyles est protoné, l'autre est deutéré). Cependant, ce type de marquage n'est pas stéréosélectif et cette dilution de la densité de protons s'accompagne d'une chute du transfert d'aimantation par un facteur 4 (la moitié de l'aimantation d'un premier méthyle peut être transférée sur un

second méthyle protoné dans 50 % des cas, soit un transfert réduit à 25 %). En prenant en compte les protocoles de marquages sélectifs disponibles et les caractéristiques des acides aminés, la protonation sélective des Isoleucines en position δ_1 apparaît comme la meilleure stratégie pour détecter des informations translationnelles à longue portée entre méthyles.

2.4.2 Principe de marquage spécifique du méthyle δ_1 des Isoleucines

L'implémentation du marquage spécifique sur les méthyles δ_1 des Isoleucines peut se faire par le protocole mis au point par Gardner *et coll.* (Gardner *et al.*, 1997a), utilisant comme précurseur à l'Isoleucine l' α -cétobutyrate. Les carbones 4 et 3 de l' α -cétobutyrate, en passant par une chaîne de réactions enzymatiques, seront incorporés respectivement en position δ_1 et γ_1 de l'Isoleucine (Figure 2.6). Toutes ces réactions sont réalisées *in-vivo* dans *E. coli*. Pour ce faire, il suffit d'ajouter, une heure avant l'induction de la surexpression de la protéine d'intérêt, une quantité suffisante (~ 75 mg/L) du précurseur métabolique marqué isotopiquement.

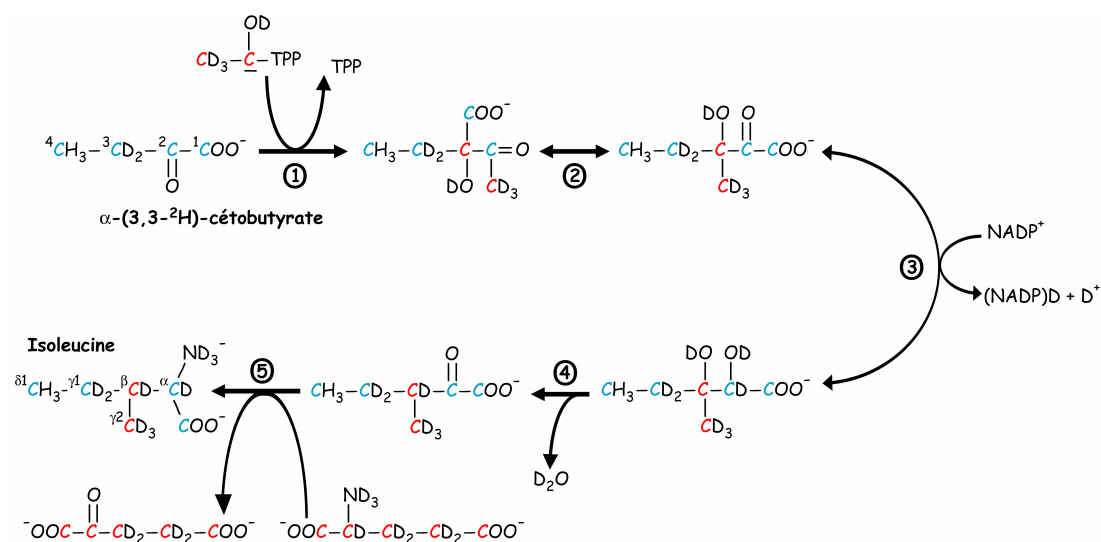


Figure 2.6: Réaction de conversion du (3,3- 2 H)-2-cétobutyrate en Isoleucine spécifiquement protonée dans *E. coli*.

Les carbones provenant du précurseur sont indiqués en bleu et ceux provenant des autres sources de carbones en rouge. Les mécanismes réactionnels sont disponibles sur le site "<http://www.genome.ad.jp/kegg/pathway.html>". Les numéros correspondent aux enzymes permettant la réaction : 1) *Acetolactate pyruvate-lyase* ; 2) *Isomerase* ; 3) *Ketol-acid reductoisomerase* ; 4) *Dihydroxy-acid-dehydrase* ; 5) *Transaminase B*. TPP = Thiamine diphosphate.

2.4.3 Implémentation du protocole

2.4.3.1 Le solvant

La production d'un échantillon perdeutééré et spécifiquement protoné sur certains sites nécessite des conditions particulières dues au milieu minimum utilisé lors de la culture. Ceci reste indépendant du choix du site spécifiquement protoné. L'eau étant la première source d'hydrogène de la bactérie et donc pour les protéines, il faut utiliser de l'eau lourde lors de la culture. Le choix du pourcentage d'enrichissement du D₂O est important. Si l'on veut deutérer à 99 % les protéines, le D₂O utilisé doit être supérieur à 99 % de deutération. Un taux d'enrichissement du D₂O entre 99.7 et 99.9 % est un bon compromis en ce qui concerne tant du point de vue du coût (~200 €/L) et du pourcentage de deutération.

Il est important d'éliminer les différentes sources d'entrée d'eau résiduelle dans le milieu de culture. Une première source est l'humidité de l'air : il est idéal de travailler avec un air sec dans les incubateurs. Les différents composés du milieu de culture peuvent également être des sources d'hydrogène possibles, même si elles restent faibles. Sur tous les produits contenant des atomes d'hydrogènes échangeables avec le solvant, ou partiellement hydratés, une attention particulière est portée afin d'éliminer toutes sources isotopiques ¹H. Pour ceux-ci, les composés sont soit achetés anhydres, soit lyophilisés plusieurs fois après avoir été solubilisés dans du D₂O.

2.4.3.2 Choix de la source de carbone

Lors de culture bactérienne, plusieurs sources de carbones peuvent être utilisées : le pyruvate, l'acétate, le glycérol et le glucose (Figure 2.3). Durant les années 90, l'acétate a été utilisé pour les marquages ¹³C uniformes ou partiels (Venters *et al*, 1991), mais aussi pour la deutération totale (Venters *et al*, 1995), en raison du coût élevé du glucose ¹³C à l'époque (deutééré ou non). Le pyruvate peut également être une source de carbone deutéré facile à obtenir, puisque ses protons s'échangent à pH basique avec ceux du solvant. Le glycérol ¹²C deutéré est aussi une source de carbone bon marché pour la perdeutération des protéines. Ce dernier, utilisé en bioréacteur, permet d'obtenir une biomasse beaucoup plus importante. Avec la

généralisation des méthodes de marquages isotopiques pour la RMN, le prix du D-glucose perdeuté, ou uniformément marqué ^{13}C et perdeuté, a beaucoup diminué ces dernières années (~ 100 €/g). Le glucose est désormais la source de carbone ^{13}C perdeutérée la plus abordable pour la surexpression des protéines dans *E. coli* et est donc privilégié pour les études RMN.

Afin d'obtenir un échantillon avec un niveau de deutération le plus haut possible à partir de la même source de D_2O , le taux de deutération du glucose doit être le plus élevé possible. L'utilisation de D-glucose- d_7 deutéré à plus de 98 % dans un milieu de culture, où l'on aura pris soin d'éliminer toutes les traces potentielles des protons ^1H de l'eau (§ 2.4.3.1), permet d'obtenir un niveau de deutération de l'ordre de 98-99 % dans la protéine surexprimée. Il est possible de simuler une deutération virtuellement proche de 100 % pour les application RMN en combinant la perdeutération avec un marquage uniforme ^{13}C . En associant la production d'échantillons $\text{U-}[^2\text{H}]$, $\text{U-}[^{13}\text{C}]$, $[^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$ et des séquences d'impulsion filtrant sélectivement les signaux ^{13}C , il est possible d'éliminer le signal des protons résiduels liés à un ^{13}C . La combinaison de ces marquages isotopiques avec ces outils spectroscopiques doit permettre d'obtenir un spectre RMN où le signal des protons est réduit à moins de 0.2 %.

Toutefois, cette stratégie suppose que les marquages ^2H et ^{13}C soient indépendants. Or, pour une source de glucose $\text{U-}[^{13}\text{C}]$, $\text{U-}[^2\text{H}]$ donnée, la présence d' ^1H et d'un ^{12}C peut être corrélée. Pour le marquage spécifique $\text{U-}[^2\text{H}]$, $\text{U-}[^{13}\text{C}]$, $[^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$ d'une protéine, des contrôles qualités ont été effectués sur quatre lots de glucose $\text{U-}[^{13}\text{C}]$, $\text{U-}[^2\text{H}]$ de différentes compagnies. Les taux de marquages isotopiques spécifiques ont été déterminés par RMN 1D ^1H et ^{13}C et sont reportés dans le tableau 2.2. Si l'on s'intéresse à un marquage $[^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$, la présence d'isotopomères ^{12}C , ^1H , sur les autres sites de la protéine, ne pourra pas être éliminée par les filtres RMN. Le Tableau 2.3 indique qu'il est particulièrement important de choisir une source de glucose comportant une proportion d'isotopomère ^{12}C , ^1H , la plus faible possible afin d'éliminer plus de 99.9 % du signal des protons résiduels.

Glucose	$^{13}\text{C}-^1\text{H}$	$^{12}\text{C}-^1\text{H}$	$^{13}\text{C}-^2\text{H}$	$^{12}\text{C}-^2\text{H}$
Lot 1 ^a	2,18 %	0,82 %	95,82 %	1,18 %
Lot 2 ^a	2,63 %	0,37 %	95,37 %	1,63 %
Lot 3 ^b	1,00 %	< 0,01 %	98,00 %	1,00 %
Lot 4 ^c	1,30 %	< 0,01 %	98,10 %	0,60 %

Tableau 2.3 : Taux de marquage isotopique corrélé entre ^2H et ^{13}C du D-Glucose (^{13}C , d_7). ^a Spectra Stable Isotope ; ^b Cambridge Isotope Laboratories ; ^c Isotec.

2.4.3.3 Préparation de l' α -cétobutyrate protoné spécifiquement

Les atomes d'hydrogène du groupement méthylène en position 3 du 2- α -cétobutyrate s'échangent progressivement avec le deutérium du solvant lors de la formation de la forme tautomère énolate du 2- α -cétobutyrate en milieu basique (Figure 2.7). La réaction doit se faire de manière diluée, à température modérée et pH faiblement basique. Ces conditions permettent d'éviter une addition de l'ion énolate sur le carbonyle d'une autre molécule. De manière pratique, l'échange est réalisé en plaçant 75 mg de 2- α -cétobutyrate dans 150 mL de 99,9 % $^2\text{H}_2\text{O}$. Avant la réaction, le pH de la solution est ajusté à 10,5 par l'addition de deutéroxyde de sodium à 40 % (NaOD). La réaction s'effectue, ensuite, en incubant le 2- α -cétobutyrate à 40°C pendant 72 heures.

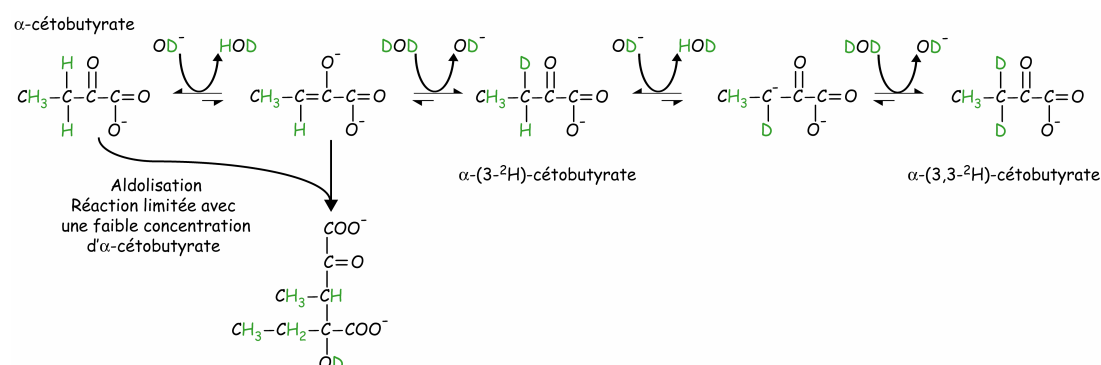


Figure 2.7 : Réaction d'échange H/D en position 3 du 2- α -cétobutyrate.

La réaction se fait à pH basique (pH = 10,5) pendant 72 heures à 40°C.

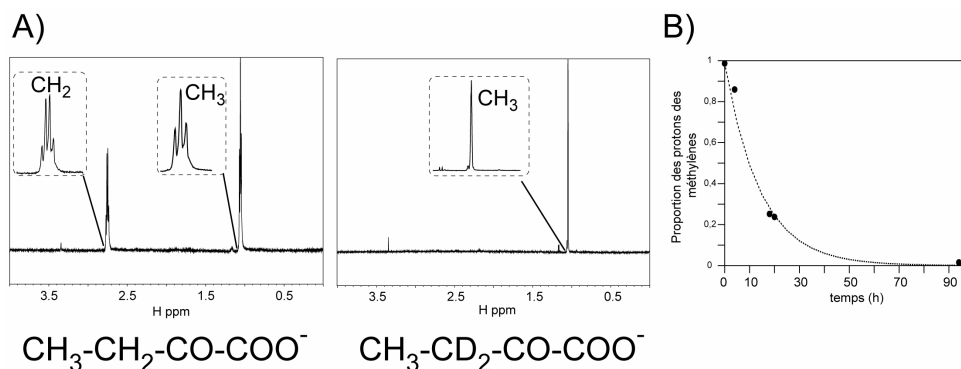


Figure 2.8 : Spectres RMN 1D du proton du 2- α -cétobutyrate en A et courbe de suivi de la déprotonation du méthylène au cours de la réaction en B.

Les spectres ont été enregistrés sur un spectromètre 600 MHz à 25°C dans du D₂O avant réaction (à gauche) et après réaction (à droite).

L'échange total H/D a été suivi par spectroscopie RMN 1D du proton (Figure 2.7). L'arrêt de la réaction se fait en abaissant le pH de la solution à 7,2 par une solution d'acide chlorhydrique deutéré (DCl) ; cette solution est alors conservée à 4°C. Il est important de noter que le D₂O utilisé pour cet échange peut être recyclé dans le milieu de culture des bactéries, n'entraînant ainsi pas de surcoût pour la deutération spécifique du précurseur.

2.4.3.4 Mise en place du protocole

Le protocole décrit ci-dessous a été implémenté à partir du protocole proposé initialement par Gardner *et coll.* (Gardner *et al.*, 1997a). Il a été optimisé pour obtenir le marquage isotopique spécifique le plus élevé possible. L'ubiquitine a été utilisée comme protéine modèle pour mettre en place ce marquage spécifique. La production de l'échantillon se fait en plusieurs étapes et dure, au total, trois jours (sans l'étape de purification) :

(i) Des cellules de *E. coli* BL21(DE3) sont transformées avec le plasmide d'intérêt (pET41c/ubiquitine humaine dans cet exemple).

(ii) Les bactéries se développent dans 3 précultures successives afin de les acclimater progressivement au milieu minimum M9 (précultures I et II), puis à la culture en eau lourde (préculture III).

Jour -1 :	Echange de D/H du 2-céto-3,3-d ₂ -butyrate (§ 2.4.3.3). Préparation des produits des milieux de culture M9 à lyophiliser. Lyophilisation sur la nuit.
Jour 0 :	Transformation d' <i>E.coli</i> BL21 (DE3) avec le plasmide d'intérêt.
Jour 1 :	
Préculture I : durée 4 heures	1 clone $\Rightarrow$ 25 mL LB / antibiotiques (antibio.). Placer à 37°C, 220 rpm pendant 4 heures. $DO_{600nm} = 0,7 - 0,8$. Centrifugation 10 min, 3220 g.
Préculture II : durée 2 heures	Resuspendre le culot dans 50 mL M9/H ₂ O (Annexe 1) (erlen 250 mL). $DO_{600nm} = 0,3 - 0,4$. Placer à 37°C, 220 rpm pendant 2 heures. $DO_{600nm} = 1$. Centrifugation 10 min, 3220 g.
Préculture III : durée 12 heures	Resuspendre le culot dans 100 mL M9/D ₂ O ₉₉ % (Annexe 1) (erlen 500 mL). $DO_{600nm} = 0,2 - 0,3$. Placer à 37°C, 220 rpm pendant la nuit. $DO_{600nm} = 1,5$. Centrifugation 10 min, 3220 g.
Jour 2 :	
Culture 1 : durée 2 - 3 heures	Resuspendre le culot dans 350 mL M9/D ₂ O _{99,85} % (Annexe 1) (erlen 3 L). $DO_{600nm} = 0,2-0,3$. Placer à 37°C, 220 rpm pendant 2 heures. $DO_{600nm} = 0,7-0,8$.
Culture 2 : durée 1 heure	Ajout de 150 mL de milieu avec le précurseur. $DO_{600nm} = 0,5-0,6$. Placer à 37°C, 220 rpm pendant 1 heure. $DO_{600nm} = 0,8$. Induction à l'IPTG (1 mM). Placer à 37°C, 220 rpm pendant 4-6 heures.
Durée 7 - 8 heures	$DO_{600nm} = 1,5$. Arrêt des cultures, récolte des cellules. Centrifugation 15 min, 5500 g à 4°C. Congélation du culot de cellules ou purification.

Figure 2.9 : Protocole de production de l'ubiquitine deutérée spécifiquement protonée sur les méthyles δ_1 des Isoleucines pour 500 mL de culture.

Les milieux minimums utilisés dans ce protocole (M9/H₂O, M9/D₂O₉₉ % et M9/D₂O_{99,2} %) sont décrits dans l'Annexe 1. IPTG= Isopropyl β -D-thiogalactopyranoside.

(iii) Dans l'étape finale de ce protocole, les bactéries se développent à 37°C dans 350 mL de milieu minimum M9 (99.2 % D₂O), contenant 1 g.L⁻¹ ¹⁵ND₄Cl comme unique source d'azote ¹⁵N, et 2 g.L⁻¹ D-glucose-d₇ comme unique source de carbone.

(iiii) Quand la DO_{600nm} atteint 0,8, on ajoute 150 mL d'une solution contenant 75 mg de 2-céto-3,3-d₂-butyrate (préparée selon le protocole décrit au § 2.4.3.3), 1 g.L⁻¹ ¹⁵ND₄Cl, 2 g.L⁻¹ D-glucose-d₇ ainsi que tous les compléments du milieu M9 (Annexe 1). Après 1 heure, l'expression de la protéine est induite par l'addition d'IPTG pour une concentration finale de 1 mM. En fin de culture, la DO_{600nm} est typiquement de 1,5. Les cellules sont collectées par centrifugation et le culot de cellules peut alors être congelé à -20°C pour être conservé avant la purification. L'ubiquitine est ensuite purifiée en deux étapes : une chromatographie utilisant une colonne échangeuse d'anions (Q-Sepharose, Amersham Pharmacia Biotech) et une colonne de gel filtration (Superdex-75, Amersham Pharmacia Biotech). Le rendement final de la production après purification est de l'ordre de 60 mg.L⁻¹ de culture, du même ordre de grandeur qu'une culture en non deutérée (Tableau 2.4).

2.4.3.5 Points critiques du protocole

Le protocole permettant une protonation spécifique des protons des méthyles δ₁ des Isoleucines dans une protéine perdeutérée a été optimisé. En effet, le protocole implémenté au départ n'était pas optimal sur certains points. Ces détails ne concernent pas l'efficacité du marquage, mais plutôt des parties techniques et expérimentales.

Tout d'abord, par souci de reproductibilité des cultures en milieu deutéré, l'utilisation de vaisselle jetable est indispensable. L'usage d'une verrerie, utilisée pour d'autres cultures, peut affecter le développement des bactéries et/ou contaminer le milieu de culture (présence de contaminants sous forme de traces : antibiotiques, ...).

Lors des précultures, la masse cellulaire ne doit pas être trop élevée. En effet, une biomasse trop importante peut entraîner un début de lyse des cellules, ce qui doit être évité pour l'expression, car nous cherchons à rester dans la phase de croissance exponentielle des bactéries. Cette remarque d'ordre générale reste particulièrement importante dans les cultures en milieu deutéré.

Une des étapes critique est de limiter la phase de surexpression, après induction à moins de huit heures, à 37°C (Tugarinov *et al*, 2006), ceci pour prévenir la mise en place des voies de synthèse parallèle et donc d'entraîner une fuite du marquage ou une dilution isotopique.

2.4.4 Caractérisation de l'efficacité du marquage spécifique

Le taux d'incorporation du deutérium dans l'ubiquitine a été quantifié par RMN 1D du proton. Seuls sept signaux intenses, entre 0,3 et 0,9 ppm, correspondants aux méthyles δ_1 des Isoleucines, sont visibles sur le spectre RMN proton de l'échantillon d'ubiquitine U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 . L'analyse quantitative des intensités des signaux résiduels pour les autres protons aliphatiques, en comparaison avec le spectre de l'ubiquitine totalement protonée, révèle que le taux d'incorporation du deutérium est de 98,5 % (Figure 2.10).

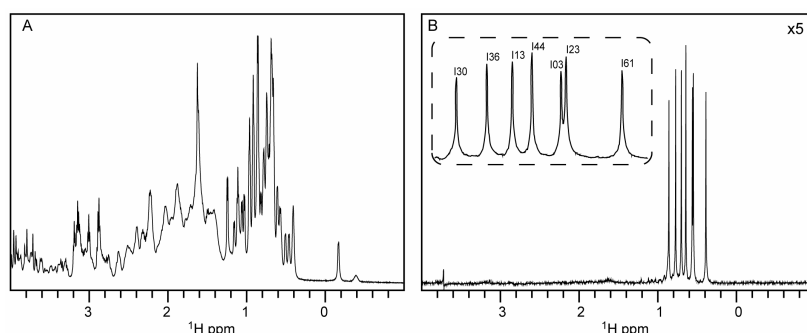


Figure 2.10 : Spectre RMN ^1H de l'ubiquitine protonée en A et U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 en B.

En comparaison avec (A), l'échelle verticale en (B) a été multipliée par un facteur 5. L'attribution des méthyles δ_1 des Isoleucines est notée sur le zoom du spectre B.

Aucun signal pour les carbones δ_1 des Isoleucines n'est observé sur le spectre 1D RMN en détection directe du carbone-13 de l'échantillon d'ubiquitine U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 (Figure 2.11.B). La comparaison entre les spectres ^{13}C , enregistrés en détection directe sur l'ubiquitine U-[^{13}C] (Figure 2.10.A) et U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 (Figure 2.11.B), montre que l'incorporation des méthyles à partir du précurseur, l' α -cétobutyrate, est supérieure à 98 %. Seuls, les signaux des carbones δ_1 des Isoleucines sont observés dans le spectre ^{13}C , enregistré sur l'échantillon d'ubiquitine U-[^2H], U-[^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 (Figure 2.11.C), préparé en utilisant (3,3- ^2H)-2-(4- ^{13}C)-cétobutyrate comme précurseur. Ceci indique que le

méthyle $^{13}\text{CH}_3$ de l' α -cétobutyrate n'est pas incorporé dans d'autres sites que celui de l'Isoleucine δ_1 .

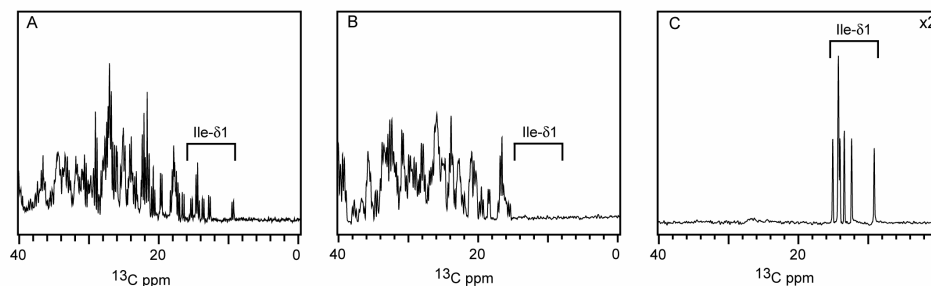


Figure 2.11 : Spectre RMN ^{13}C avec découplage ^1H de l'ubiquitine.

En A) l'ubiquitine U- ^{13}C , en B) de l'ubiquitine U- ^{2}H , U- ^{13}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 et en C) de l'ubiquitine U- ^{2}H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$ -Ile- δ_1 . En comparaison avec (A) et (B), l'échelle verticale en (C) a été multipliée par un facteur 2.

2.4.5 Différents types de marquages envisageables

Plusieurs marquages spécifiques des Isoleucines sont envisageables en fonction de l' α -cétobutyrate spécifiquement marqué et de la source de carbone utilisée. Chacun de ces types de marquages spécifiques peut être utilisé pour une application particulière. Comme le montre la Figure 2.12, les possibilités de marquages spécifiques sont multiples, dépendent de la source principale de carbone (^{13}C ou ^{12}C / ^2H) et surtout du marquage du précurseur. On peut envisager tous ces marquages avec les différents isotopomères pour le méthyle du précurseur (CH_3 / CH_2D / CHD_2). Ces précurseurs sont disponibles commercialement et sont utiles pour l'étude des propriétés de relaxation des méthyles (Tugarinov *et al*, 2006). Sur l'ensemble des marquages spécifiques des Isoleucines possibles, j'ai réalisé trois types d'échantillons différents (cadre rouge sur la Figure 2.11). Les avantages de ces marquages sont détaillés ci-dessous.

(i) Dans un échantillon U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 (Figure 2.12.A), seuls les atomes d'hydrogène du méthyle δ_1 des Isoleucines sont protonés ; il n'y a pas de marquage spécifique des carbones. Le faible coût de production de cet échantillon est l'intérêt principal de ce type de marquage.

(ii) Dans un échantillon U- ^{2}H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$ -Ile- δ_1 (Figure 2.12.B), comme le précédent, seul le site δ_1 des Isoleucines est protoné. Cet échantillon est également marqué ^{13}C sur le carbone δ_1 des Isoleucines. Ce type de marquage offre

la possibilité de réaliser des expériences NOESY, éditées et filtrées ^{13}C , pour distribuer les résonances dans plusieurs dimensions (3D ou 4D).

(iii) Enfin, le type de marquage spécifique $\text{U}-[^2\text{H}]$, $\text{U}-[^{13}\text{C}]$, $[^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$ (Figure 2.12.D), où tous les carbones sont marqués ^{13}C sauf le carbone δ_1 des Isoleucines et où seuls les atomes d'hydrogène des méthyles des Isoleucines sont protonés. On remarquera que les carbones γ_1 , α et le carbonyle des Isoleucines sont aussi présents sous forme ^{12}C , car ils proviennent du site du précurseur non enrichi en carbone-13 (Figure 2.6). S'il n'y a pas de superposition dans la dimension ^1H , le coût est un premier avantage pour ce type de marquage. De plus, la relaxation du méthyle n'est pas affectée par l'interaction dipolaire $^{13}\text{C}-^1\text{H}$. Ce type de marquage combine ainsi les deux avantages des marquages précédents : il permet toujours l'utilisation de filtre $^{12}\text{C}/^{13}\text{C}$ et la conservation d'une relaxation optimale pour le méthyle.

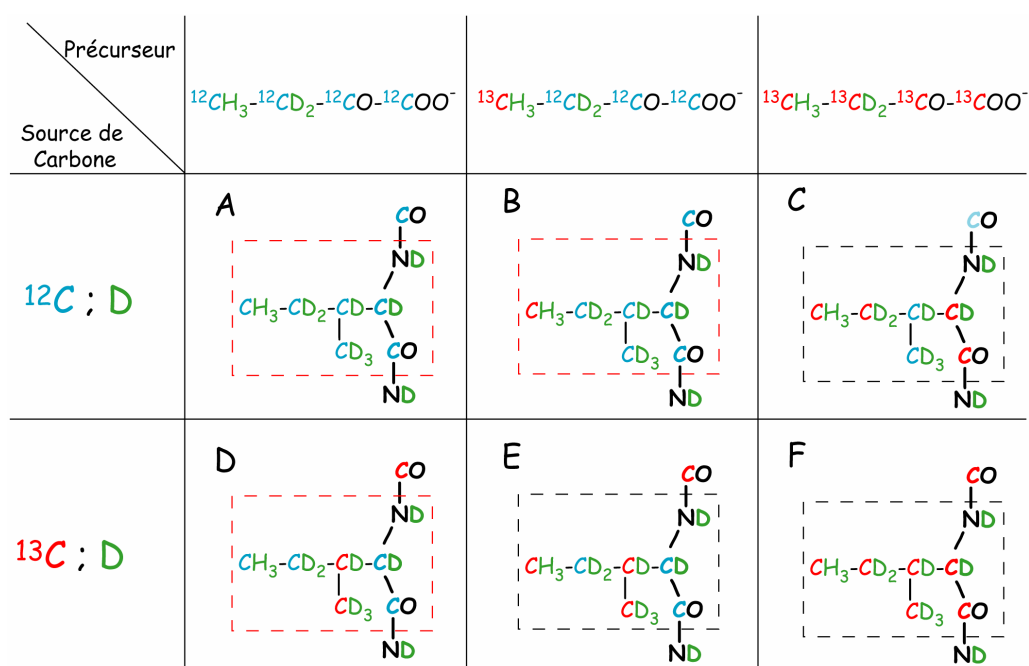


Figure 2.12 : Schéma des marquages isotopiques pour l'Isoleucine.

Ils dépendent de l' α -cétobutyrate spécifiquement marqué isotopiquement choisi et de la source principale de carbone (^{12}C ou ^{13}C) deutérée. Les ^{12}C (^{13}C) sont indiqués en bleu (rouge) et les Isotopes de l'hydrogène en vert. Les marquages utilisés sont encadrés en rouge.

Les spectres 1D proton des différents échantillons d'ubiquitine marqués spécifiquement sont présentés dans la Figure 2.13.

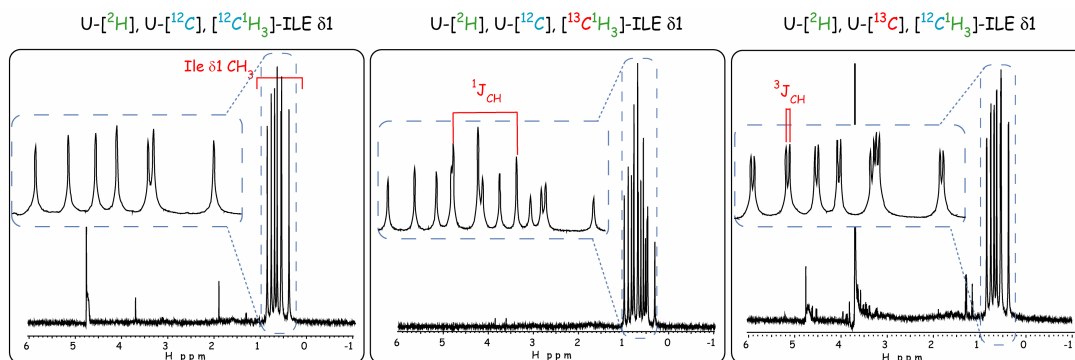


Figure 2.13 : Spectres RMN 1D du proton sans découplage ^{13}C de l'ubiquitine. Chaque type de marquage est écrit au-dessus de son spectre 1D proton. Un zoom indique la structure fine des signaux due aux couplages scalaires $^nJ_{\text{CH}}$ observés.

2.4.6 Applications

2.4.6.1 Préparation de protéines U- ^{2}H , U- ^{12}C / ^{13}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1

Ce protocole a permis de préparer deux sortes d'échantillons d'ubiquitine : un U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 ainsi qu'un U- ^{2}H , U- ^{13}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 . En collaboration avec le Professeur Zhengrong Wu (Université de l'Ohio, Columbus, USA), un échantillon de la γS -crystalline U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 a été préparé suivant le protocole optimisé à l'IBS. La γS -crystalline est une protéine modulaire de 19 kDa, contenant neuf Isoleucines. Le marquage isotopique spécifique $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 est envisageable pour les protéines contenant peu d'Isoleucines et pour lesquelles il n'y a pas de recouvrement entre les résonances ^1H des méthyles δ_1 des Isoleucines.

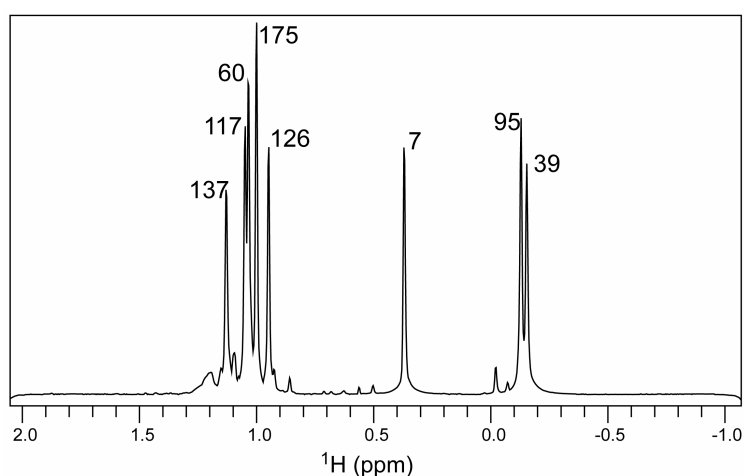


Figure 2.14 : Spectres RMN 1D du proton de la γS -crystalline U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 . L'attribution des méthyles δ_1 des Isoleucines est notée sur le spectre.

2.4.6.2 Préparation de protéines U-[²H], U-[¹²C], [¹³CH₃]-Ile- δ_1

Pour la production de protéines de plus grandes tailles, le marquage isotopique spécifique [¹³CH₃]-Ile- δ_1 est indispensable afin d'avoir une résolution suffisante lors de l'analyse des spectres. Grâce à ce marquage isotopique spécifique, l'obtention de spectres en 3D ou 4D est possible pour analyser les corrélations interméthyles. Trois protéines de tailles différentes ont été surexprimées selon le protocole optimisé avec le (3,3-²H)-2-(4-¹³C)-cétobutyrate comme précurseur :

(i) yajG est une protéine globulaire, étudiée au laboratoire ; sa masse moléculaire est de 21 kDa et elle possède quatorze Isoleucines (Boudet, 2007).

(ii) la Malate synthase G (MSG) est une protéine de 82 kDa contenant quarante-quatre Isoleucines. Sa structure a été résolue par cristallographie aux rayons X (Howard *et al*, 2000). L'attribution et la structure RMN basse résolution sont aussi connues (Tugarinov *et al*, 2002, 2003a, 2003b, 2005a). La MSG est une bonne protéine modèle pour l'étude de protéines de grandes tailles (> 50 kDa).

(iii) TET2 est un assemblage macromoléculaire de grande taille (>468 kDa) issu d'un organisme hyperthermophile. La structure cristallographique de TET2 a été résolue en 2005 (Borissenko *et al*, 2005). TET2 est un dodécamère symétrique dont l'unité principale a une masse moléculaire de 39 kDa. Chaque sous-unité contient trente-quatre Isoleucines.

Les rendements de surexpression de ces protéines sont présentés dans le Tableau 2.4, avec l'ubiquitine indiquée comme référence.

Rendement	Ubiquitine	yajG	MSG	TET2
MW (τ_c)	9 kDa (5 ns)	21 kDa (13 ns)	82 kDa (55 ns)	468 kDa (310 ns)
M9/H ₂ O	60 mg.L ⁻¹	140 mg.L ⁻¹	80 mg.L ⁻¹	-
M9/D ₂ O _{99,2%}	60 mg.L ⁻¹	100 mg.L ⁻¹	76 mg.L ⁻¹	~20 mg.L ⁻¹

Tableau 2.4 : Rendement de surexpression des différentes protéines pour 1 L de culture en milieu minimum M9/H₂O et en M9/D₂O_{99,2%} avant purification.

Le temps de corrélation de chacune des protéines dans 100 % de D₂O et à une température de 37°C est aussi indiqué.

Les spectres de corrélation HMQC ^1H - ^{13}C correspondant à chacune des protéines citées ci-dessus sont présentés sur la Figure 2.15.

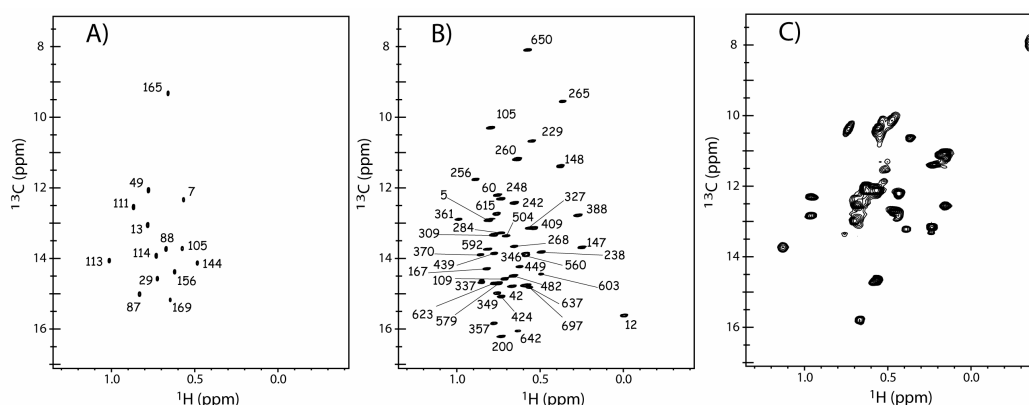


Figure 2.15 : Spectres 2D de corrélation HMQC ^1H - ^{13}C des méthyles δ_1 des Isoleucines de yajG, MSG et TET2.

Les spectres ont été enregistrés sur un spectromètre 800 Mhz, équipé d'une sonde cryogénique, sur des échantillons $\text{U}-[^2\text{H}]$, $\text{U}-[^{12}\text{C}]$, $[^{13}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$: A) yajG à 37°C, B) MSG à 37°C et C) TET2 à 50°C. Les attributions des quatorze Isoleucines de yajG et des quarante-quatre Isoleucines de MSG sont notées sur leur spectre respectif.

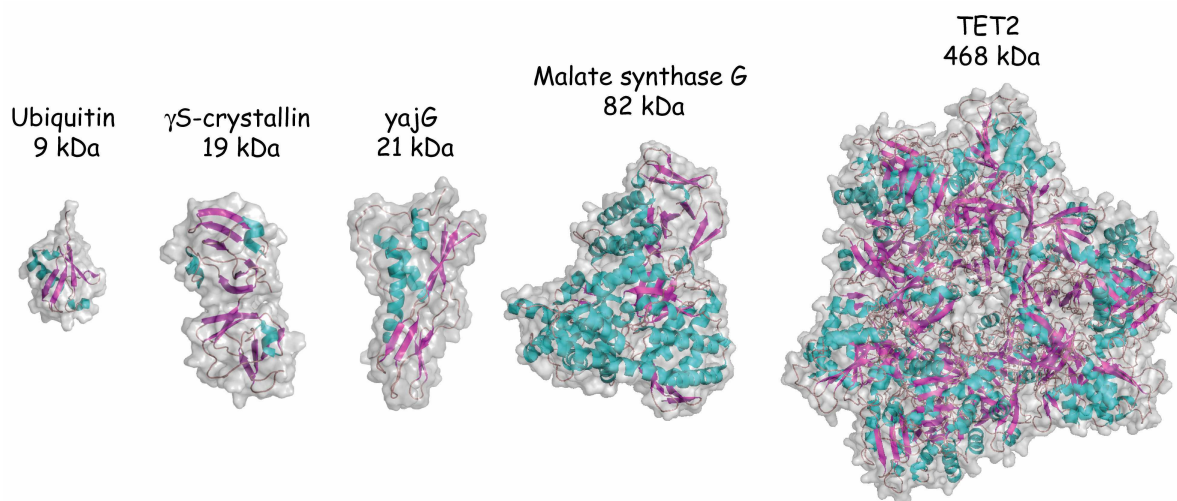


Figure 2.16 : Structure et taille relative des protéines marquées sélectivement sur les méthyles δ_1 des Isoleucines.

Le protocole de production de protéines perdeutérées et spécifiquement protonées sur les méthyles δ_1 des Isoleucines a été optimisé. A partir de ces échantillons, de nouvelles informations à longue portée peuvent être détectées et les applications seront présentées dans les chapitres suivants.

2.5 Protonation spécifique des méthyles des Valines et des Leucines

2.5.1 Principe de marquage spécifique du méthyle des Valines et des Leucines

L'implémentation du marquage spécifique sur les méthyles des Valines et des Leucines peut se faire par le protocole mis au point par Goto, (Goto *et al*, 1999), utilisant comme précurseur à ces deux acides aminés l' α -cétoisovalérate. Les deux méthyles en position 4 et le carbone en position 3 de l' α -cétoisovalérate seront incorporés enzymatiquement en fin de chaîne latérale de la Leucine et de la Valine. Toutes ces réactions sont réalisées *in-vivo* dans *E. coli*. La transformation de l' α -cétoisovalérate en Valine est une réaction en une seule étape qui peut être catalysée par trois transaminases différentes (Figure 2.17).

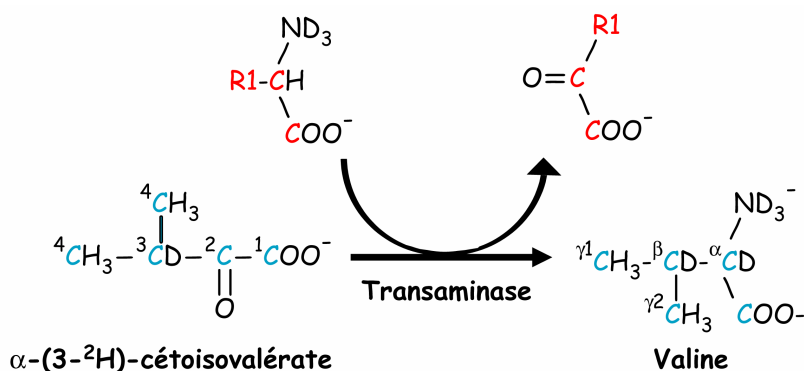


Figure 2.17 : Réaction de conversion du (3- ^2H) α -cétoisovalérate en Valine spécifiquement protonée dans *E. coli*.

Les carbones provenant du précurseur sont indiqués en bleu et ceux provenant des autres sources de carbones en rouge. Les trois transaminases peuvent être : *Transaminase C*, *Transaminase B* et *Leucine dehydrogenase*. Les mécanismes réactionnels sont disponibles sur le site "<http://www.genome.ad.jp/kegg/pathway.html>".

La synthèse de la Leucine se fait en quatre étapes et nécessite l'apport de deux carbones supplémentaires provenant de l'acétyl-CoA (Figure 2.18). Pour réaliser le marquage spécifique des méthyles des Valines et des Leucines, il suffit d'ajouter, une heure avant l'induction, une quantité suffisante (125 mg/L) du précurseur métabolique marqué isotopiquement. Le protocole utilisé est le même que pour le marquage spécifique des Isoleucines, seul le précurseur change (Figure 2.9).

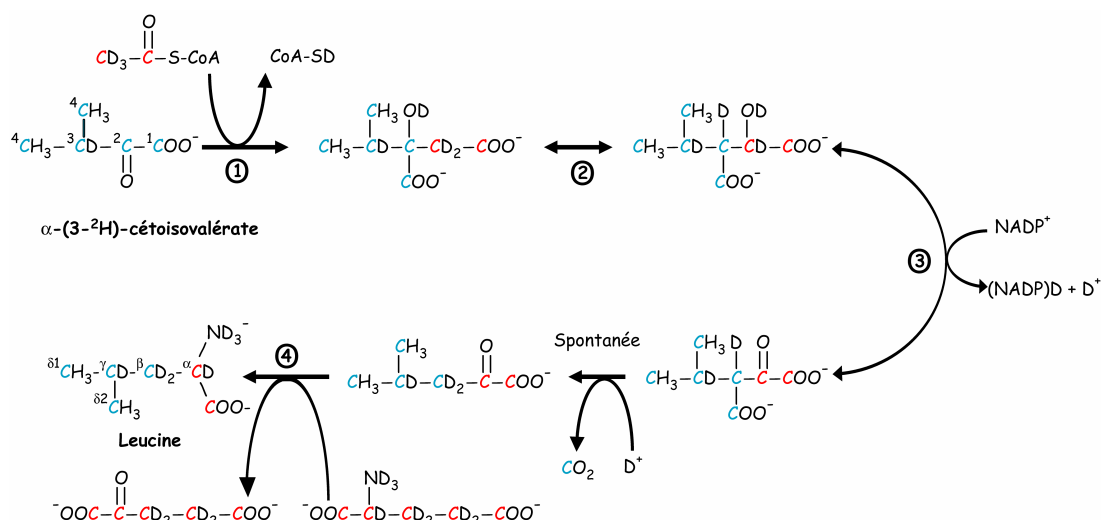


Figure 2.18 : Réaction de conversion du (3- ^2H) α -cétoisovalérate en Leucine spécifiquement protonée dans *E. coli*.

Les carbones provenant du précurseur sont indiqués en bleu et ceux provenant des autres sources de carbones en rouge. Les numéros correspondent aux enzymes permettant les réactions pour la Leucine : 1 *Isopropylmalate synthase* ; 2 *Isopropylmalate isomerase* ; 3 *Isopropylmalate dehydrogenasesynthase* ; 4 *Leucine Transaminase*. CoA-S = Coenzyme A. Les mécanismes réactionnels sont disponibles sur le site "<http://www.genome.ad.jp/kegg/pathway.html>".

2.5.2 Préparation de l' α -cétoisovalérate protoné spécifiquement

L'atome d'hydrogène en position 3 du 2- α -cétoisovalérate s'échange progressivement avec le deutérium du solvant en milieu basique (Figure 2.19). De manière pratique, l'échange est réalisé en plaçant 100 mg de 2-cétobutyrate dans 60 mL de 99,9 % $^2\text{H}_2\text{O}$. Avant la réaction, le pH de la solution est ajusté à 12,5 par l'addition de deutéroxyde de sodium à 40 % (NaOD) et en incubant à 45°C pendant trois heures. L'arrêt de la réaction se fait en abaissant le pH de la solution à 7,2 par une solution d'acide chlorhydrique deutérée (DCl). La solution est ensuite conservée à 4°C. Comme pour le précurseur de l'Isoleucine, l'eau lourde utilisée pour cet échange peut être recyclée dans le milieu de culture des bactéries.

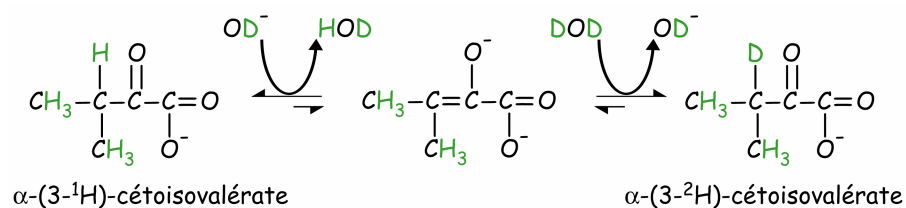


Figure 2.19 : Réaction d'échange H/D en position 3 du α -cétoisovalérate.

La réaction se fait à pH basique (pH = 12,5), pendant trois heures, à 45°C.

2.5.3 Différents types de marquages envisageables

De la même manière que pour l' α -cétobutyrate, plusieurs types de marquages spécifiques sont envisageables pour la Valine et la Leucine (Figure 2.20), en fonction de l' α -cétisovalérate spécifiquement marqué et de la source de carbone utilisée (Tugarinov *et al*, 2006).

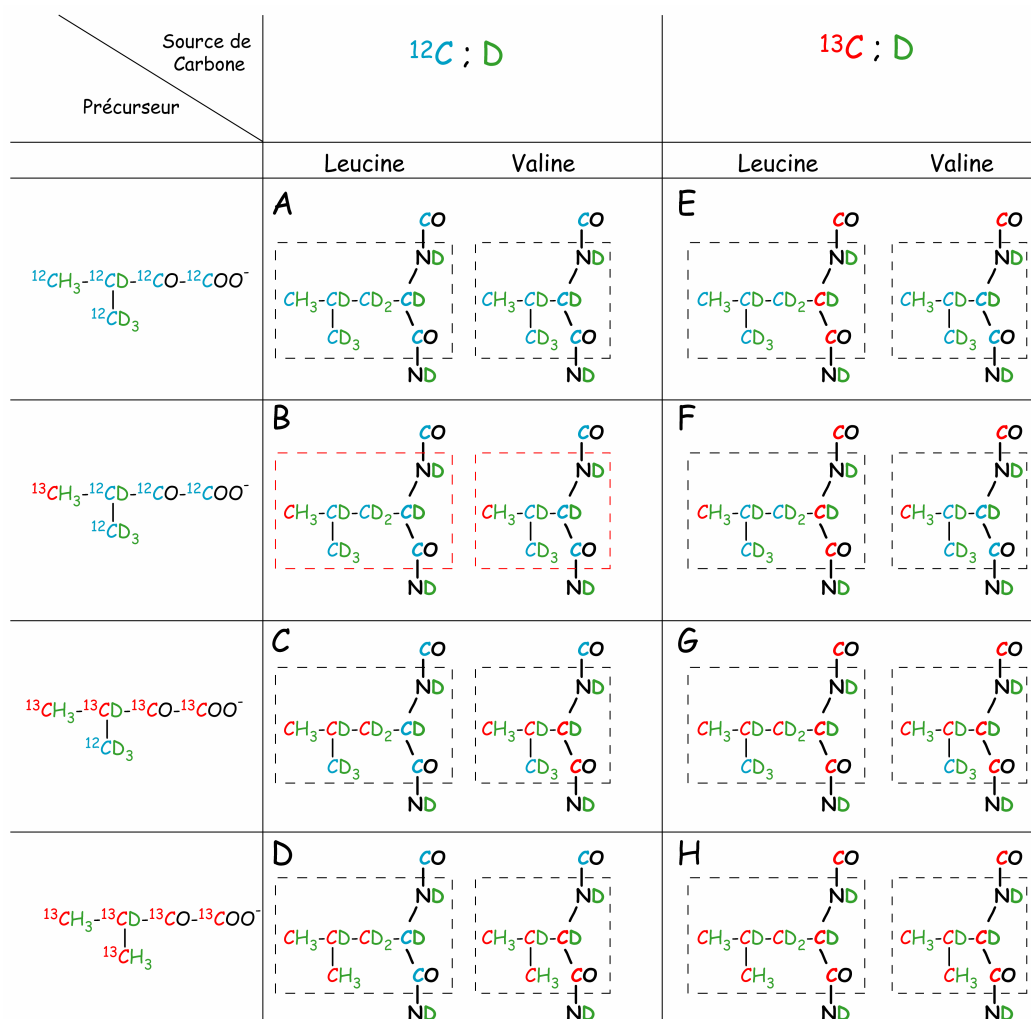


Figure 2.20 : Schéma des marquages isotopiques pour la Leucine et la Valine.

Ils dépendent de l' α -cétisovalérate spécifiquement marqué isotopiquement, choisi, et de la source principale de carbone (^{12}C ou ^{13}C) deutérée. Les ^{12}C (^{13}C) sont indiqués en bleu (rouge) et les Isotopes de l'hydrogène en vert. Le marquage utilisé est encadré en rouge.

Cependant, un marquage spécifique où les deux méthyles de ces acides aminés sont protonés n'est pas intéressant pour l'obtention de contraintes à longue portée, c'est-à-dire les schémas D, H sur la Figure 2.20. En effet, localement, la densité de proton sera très élevée et des interactions dipolaires entre les méthyles du même résidu limiteront le transfert vers des méthyles plus éloignés (diffusion de spins ou

truncation dipolaire). Il est donc intéressant de deutérer l'un des deux méthyles de ces acides aminés (Schémas A, B, C, E, F et G sur la Figure 2.20). Le précurseur nécessaire pour ce marquage est disponible commercialement. Ce type de composé, incorporé dans les Leucines et les Valines, permet d'éliminer les interactions fortes entre les méthyles du même résidu. Puisque les méthyles des Valines et des Leucines partagent la même plage de fréquences protons, pour la plupart des applications, il sera nécessaire de marquer également en ^{13}C ces méthyles, afin d'éviter les superpositions dans les spectres RMN. Comme pour l'Isoleucine, les marquages avec les différents isotopomères sont possibles et la plupart de ces précurseurs sont disponibles commercialement (Tugarinov *et al*, 2006).

2.5.4 Application : MSG

Le marquage spécifique des Valines et des Leucines a été réalisé sur la Malate Synthase G. L' α -cétoisovalérate utilisé est le 2-céto-3-méthyle- d_3 -3- d_1 -4- ^{13}C -butyrate, correspondant au deuxième précurseur dans la Figure 2.20. Un spectre de corrélation HMQC ^1H - ^{13}C , enregistré à partir d'un échantillon à 1 mM de MSG U- $[\text{}^2\text{H}]$, U- $[\text{}^{12}\text{C}]$, $[\text{}^{13}\text{CH}_3, \text{}^{12}\text{CD}_3]$ -Leu,Val, est présenté sur la Figure 2.21.

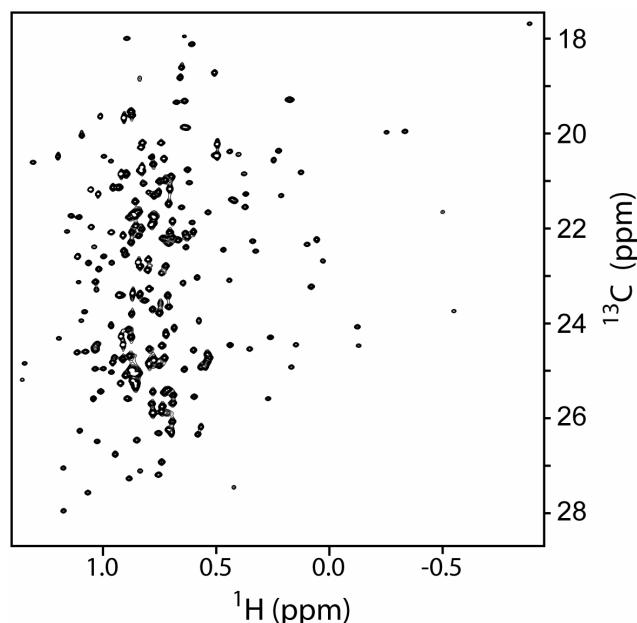


Figure 2.21 : Spectres 2D de corrélation HMQC ^1H - ^{13}C de MSG correspondants aux méthyles δ des Leucines et γ des Valines.

Le spectre a été enregistré dans D_2O sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, sur un échantillon à 1 mM de MSG U- $[\text{}^{15}\text{N}]$, U- $[\text{}^2\text{H}]$, $[\text{}^{13}\text{CH}_3, \text{}^{12}\text{CD}_3]$ -Leu,Val à 37°C .

2.6 Protonation spécifique du méthyle de l'Alanine

2.6.1 Introduction

Comme nous venons de le voir, les protocoles permettant de marquer spécifiquement les méthyles des résidus Valine, Leucine et Isoleucine sont maintenant connus. En revanche, aucun protocole de marquage spécifique de l'Alanine dans une protéine n'a été mis au point, à ce jour. Ce type de marquage présenterait pourtant de nombreux avantages. En premier lieu, l'Alanine est statistiquement l'acide aminé le plus abondant dans les protéines. Le marquage de l'Alanine permettrait d'apporter des informations structurales qui ne soient pas affectées par la dynamique propre aux chaînes latérales longues. De plus, n'impliquant que des expériences de type HNCACB facilement analysables, l'attribution de ces méthyles est à priori plus aisée que celle des chaînes longues.

Nous nous sommes donc attachés à la mise au point d'un protocole de marquage spécifique de cet acide aminé. Nous avons écarté l'utilisation de précurseurs marqués, voie que nous avons précédemment utilisée dans les protocoles de marquage des Isoleucines, des Valines et des Leucines. En effet, le précurseur de l'Alanine est le pyruvate, qui est un intermédiaire entre la glycolyse et le cycle de Krebs ; il est donc un composant central dans le métabolisme carboné. Son utilisation conduirait aux marquages des méthyles des Valines, des Leucines et des méthyles γ_2 des Isoleucines (Rosen *et al*, 1996), d'une part, comme au marquage aspécifique du squelette de la plupart des acides aminés, d'autre part.

L'utilisation d'Alanine marquée est donc la seule possibilité pour une production *in-vivo*. Cependant, chez *E. coli*, il existe plusieurs réactions réversibles de conversion entre le pyruvate et l'Alanine catalysés par trois transaminases qui peuvent conduire à des fuites des protons de l'Alanine (Figure 2.22). Ces réactions sont catalysées par des transaminases indispensables pour le métabolisme cellulaire et, de ce fait, l'existence de souches strictement auxotrophes pour l'Alanine est impossible.

Pour réduire au maximum l'incorporation de squelettes protonés issus de la conversion Alanine-pyruvate, nous avons utilisé différents précurseurs spécifiques, soit pour inhiber la voie de synthèse et donc incorporer préférentiellement le précurseur, soit pour diluer fortement les espèces protonées et donc leurs incorporations.

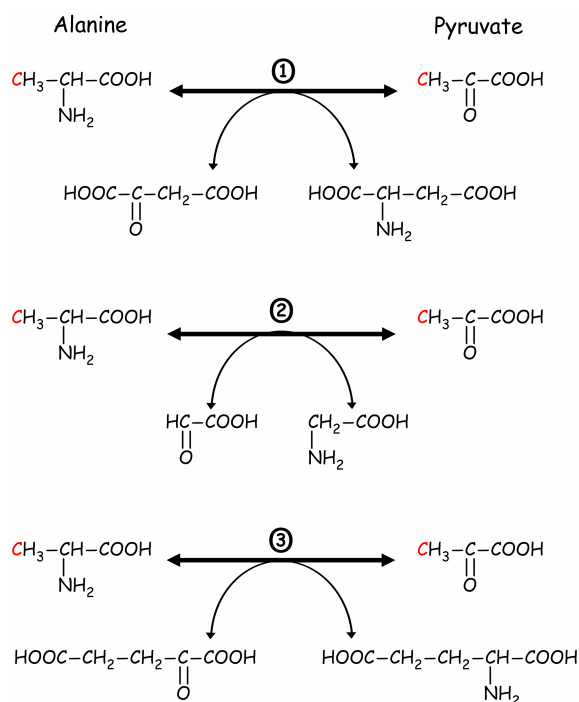


Figure 2.22 : Réaction de conversion de l'Alanine en pyruvate par les trois transaminases de *E. coli*.

En rouge, le carbone β de l'Alanine allant vers le carbone du pyruvate. 1) *Alanine-oxo acid aminotransferase*, 2) *L-Alanine-glycine transaminase*, 3) *Alanine transaminase*. Les réactions sont disponibles sur le site "<http://www.genome.ad.jp/kegg/pathway.html>".

2.6.2 Principe du protocole de surexpression

La mise au point du protocole de marquage spécifique des méthyles des Alanines dans les protéines nécessite plusieurs points importants :

(i) Pour des raisons de coût, il n'est pas envisageable de faire un grand nombre de tests en milieu deutéré. Donc, pour palier à ce problème, nous utiliserons de l'Alanine marquée spécifiquement ^{13}C en position β et réaliserons les cultures en milieu protoné. Nous analyserons l'incorporation de ce ^{13}C dans l'Alanine, comme dans les fuites éventuelles vers d'autres acides aminés. Cela implique que les protons liés au ^{13}C restent toujours liés aux ^{13}C lors de la culture, même si le carbone est

recyclé dans différents acides aminés. Cette hypothèse pourra être vérifiée, une fois le protocole optimisé, en comparant les différents pourcentages des isotopomères du méthyle de l'Alanine lors d'une culture en D₂O.

(ii) Les précurseurs utilisés lors de la mise au point du protocole ne sont pas deutérés mais devront l'être dans le protocole final. Il faut donc utiliser des précurseurs métaboliques, disponibles commercialement, sous forme deutérés et /ou facilement préparables sous forme deutérés.

Jour 0 :	
Transformation d' <i>E. coli</i> BL21 (DE3) par le plasmide pET41C.	
Jour 1 :	
Préculture 1 : durée : 8 heures	1 clone ⇒ 5 mL LB / antibiotique. Placer à 37°C, 220 rpm pendant 8 heures. Centrifugation 10 min, 3220 g.
Préculture 2 : durée : 12 heures	Resuspendre le culot dans 20 mL M9/H ₂ O (Annexe 1) (erlen 100 mL). Placer à 37°C, 220 rpm pendant la nuit.
Jour 2 :	
Culture 1 : durée : 3 heures	Centrifugation 10 min, 3220 g. Resuspendre le culot dans 70 mL M9/H ₂ O (Annexe 1) (erlen 500 mL). Placer à 37°C, 220 rpm pendant 2 heures. <i>DO</i> ₆₀₀ = 0,8-1,0.
Culture 2 : durée : 1 heure	Ajout de 30 mL de milieu M9/H ₂ O avec de l'Alanine ¹³ C _β et les précurseurs métaboliques. Attendre 1 heure avant induction à l'IPTG (1 mM).
Culture 2 : durée : 3 heures	Placer à 37°C, 220 rpm pendant 3 heures. Arrêt des cultures et récolte des cellules. Centrifugation 15 min, 5500 g à 4°C. Resuspendre dans le tampon.

Figure 2.23 : Protocole de production d'un échantillon U-[¹²C], [¹³C_β]-Ala de l'ubiquitine pour 100 mL de culture.

(iii) Il faut utiliser une protéine modèle avec des étapes de purification simples. Nous avons choisi l'ubiquitine avec une étiquette histidine en position N-terminale qui possède deux Alanines et se produit avec un rendement supérieur à 50 mg/L de culture. Il est ainsi facile de pouvoir analyser les fuites métaboliques par spectroscopie RMN. L'expression a été réalisée en cultivant des cellules de *E. coli* BL21(DE3)/pET41c contenant le gène codant pour l'ubiquitine humaine, dans un milieu minimum M9. Le principe du protocole est récapitulé sur la Figure 2.23. Lors

de ces cultures, les cellules de *E. coli* sont cultivées dans un milieu minimum M9, utilisant comme principale source de carbone le glycérol. En effet, pour la perdeutération sans marquage uniforme ^{13}C , le glycerol- d_5 est meilleur marché que le glucose- d_7 (§ 2.4.3.2).

Dans l'étape de surexpression de la culture, le milieu est supplémenté avec l'Alanine $^{13}\text{C}_\beta$ et tous les précurseurs métaboliques nécessaires une heure avant l'induction à l'IPTG (culture 2).

2.6.3 Optimisation de l'incorporation du C_β de l'Alanine

Notre but est de produire un échantillon perdeutééré spécifiquement protoné sur le méthyle C_β de l'Alanine. Nous avons donc, dans une première étape, déterminé la quantité nécessaire et minimale d'Alanine à ajouter pour obtenir un taux de marquage de la position β proche de 100% dans les protéines surexprimées. L'estimation de cette quantité permettra également de limiter le coût de production d'un tel échantillon.

Sur la Figure 2.24 est indiqué le taux de marquage en ^{13}C de la position β en fonction de la quantité d'Alanine $^{13}\text{C}_\beta$ ajoutée à la culture. Le taux d'incorporation du ^{13}C en C_β des Alanines dans l'ubiquitine a été mesuré grâce à des expériences de corrélation 2D ^1H -TOCSY. La proportion de ^{13}C incorporé en C_β est déterminée en comparant l'intensité du signal central des protons β liés à un $^{12}\text{C}_\beta$ avec celles des satellites correspondant au couplage entre les $^{13}\text{C}_\beta$ et les H_β des Alanines.

Comme cela peut être observé sur la Figure 2.24, un taux de marquage ^{13}C de 95% est atteint pour une concentration d'Alanine supérieure à 700 mg/L dans le milieu de culture minimum M9. Nous avons vérifié qu'en cas d'utilisation du glucose comme source de carbone, un taux de marquage équivalent (97%) était obtenu pour 2 g/L de glucose supplémenté par 800 mg/L d'Alanine- $^{13}\text{C}_\beta$. Pour la suite, une concentration de 800mg/L d'Alanine- $^{13}\text{C}_\beta$ a été utilisée dans les cultures afin d'obtenir un taux d'incorporation supérieur à 95 % dans les protéines surexprimées.

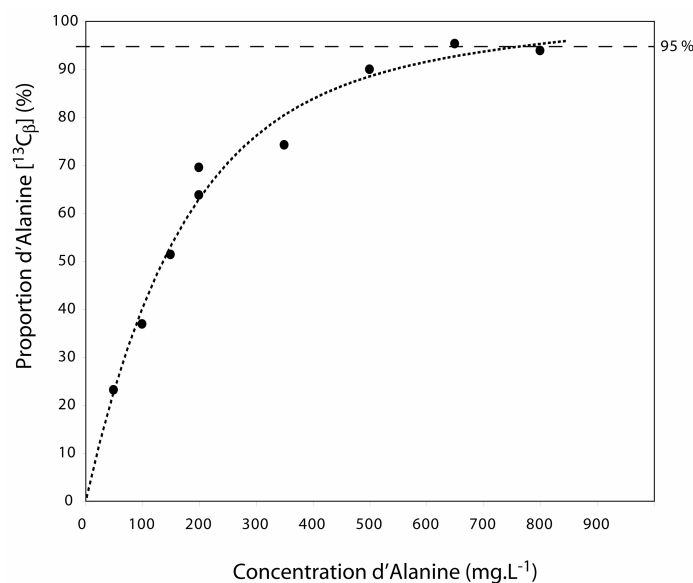


Figure 2.24 : Proportion de l'Alanine- $^{13}\text{C}_\beta$ dans l'ubiquitine en fonction de la concentration d'Alanine ajoutée dans le milieu de culture.

2.6.4 Limitation des fuites métaboliques

2.6.4.1 Estimation des fuites métaboliques

L'échantillon, obtenu à partir d'un milieu minimum contenant une concentration de 2 g/L de glucose et 800 mg/L d'Alanine marquée ^{13}C sur son C_β (noté U- ^{13}H], U- ^{12}C], [$^{13}\text{CH}_3$]-Ala- β), a été utilisé afin de déterminer la nature et l'importance des fuites métaboliques du ^{13}C vers les autres acides aminés. La quantification a été réalisée en comparant le spectre HSQC obtenu pour l'ubiquitine U- ^{12}C], [$^{13}\text{CH}_3$]-Ala- β , (Figure 2.25 A), avec le spectre obtenu pour l'ubiquitine U- ^{13}C] (non présenté). Après une correction prenant en compte la concentration relative des échantillons, les pourcentages des différentes fuites métaboliques ont pu être estimés pour chaque type de groupement.

Le carbone-13, issu de l'Alanine, se retrouve essentiellement dans les méthyles des acides aminés dont les taux de marquage sont les suivants : Isoleucine, $\text{C}\delta_1 \sim 10\%$ et $\text{C}\gamma_2 \sim 25\%$; Valine, $\text{C}\gamma \sim 25\%$; Leucine, $\text{C}\delta \sim 25\%$ et Thréonine, $\text{C}\gamma \sim 10\%$. Ces taux importants d'incorporation s'expliquent par le fait que le pyruvate est directement incorporé dans la voie de synthèse de la Valine, de la Leucine (Figure 2.3, 2.17, 2.18) et du méthyle en γ_2 de l'Isoleucine (Fig. 2.6). Le moindre taux de

marquage de la thréonine et du méthyle en δ_1 de l'Isoleucine qui en dérive provient probablement du fait que la voie de synthèse de cet acide aminé n'utilise pas directement le pyruvate- ^{13}C , mais un intermédiaire via le cycle de Krebs. On peut noter que les différents groupements des méthyles, méthynes ou méthylènes des résidus issus du cycle de Krebs (K, D, N, M, T, I, E, P, R, Q) incorporent le ^{13}C avec un taux voisin (entre 10 et 15 %). Enfin, le pyruvate peut également être réincorporé dans les voies métaboliques de la glycolyse ; on observe alors pour les méthynes et méthylènes des résidus issus de cette voie (V, L, F, Y, W, H, S, G, C) et un marquage ^{13}C plus faible (5 à 10 %).

2.6.4.2 Dilution des fuites à l'aide du pyruvate

Pour diluer au maximum la fuite métabolique de carbone-13 de l'Alanine vers les autres acides aminés, nous avons rajouté, lors de la culture, du pyruvate et de l'acétate non marqués (afin de diluer davantage l'acétate marqué issu du pyruvate vers les acides aminés provenant du cycle de Krebs). Une culture est réalisée en ajoutant du pyruvate (3 g.L $^{-1}$) et de l'acétate (3 g.L $^{-1}$) sous cette forme ^{12}C . Cette culture a permis de réduire considérablement l'incorporation de ^{13}C dans l'ensemble des acides aminés. Le taux de marquage est inférieur à 5 % (Figure 2.25). Cependant, nous observons une forte diminution du marquage de la position β de l'Alanine (~ 60 %). Cette dilution est certainement due aux réactions de transamination du pyruvate en Alanine (Figure 2.22).

Cette dilution isotopique étant trop limitante pour les applications structurales, nous avons décidé de nous tourner vers l'ajout de mélanges de précurseurs, de façon à diluer les fuites de carbone-13 dans les différentes voies métaboliques. Il aurait été possible d'augmenter la concentration en Alanine marquée, de façon à compenser ce phénomène, mais cela aurait entraîné obligatoirement une augmentation de la concentration en pyruvate marqué via les réactions de transamination et donc une augmentation du marquage des autres acides aminés.

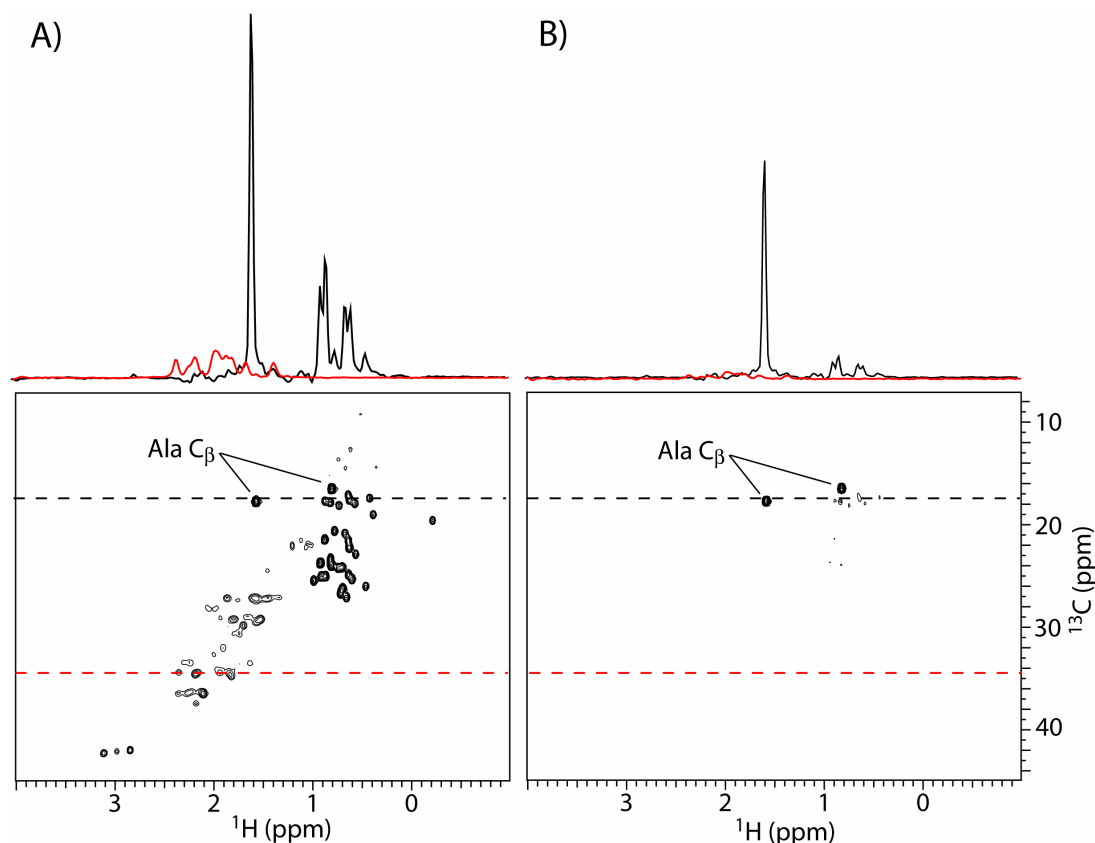


Figure 2.25 : Spectres 2D de corrélation HSQC ^1H - ^{13}C de l'ubiquitine.

Les spectres ont été enregistrés à 25°C dans D_2O sur un spectromètre 600 MHz, équipé d'une sonde cryogénique. Le premier contour correspond à un signal égal à 5 % de l'intensité des pics de corrélation des méthyles des Alanines. A) culture supplémentée par 800 mg.L $^{-1}$ d'Alanine et B) Culture supplémentée par 800 mg.L $^{-1}$ d'Alanine, 3 g.L $^{-1}$ de pyruvate et 3 g.L $^{-1}$ d'acétate. Les sections transversales F2 à deux fréquences ^{13}C sont représentées pour chacun des spectres. Le pic de l'Alanine (section transversale noire) correspond à 100 % de marquage ^{13}C en A et à 60 % en B.

2.6.4.3 Limitations du marquage résiduel des autres méthyles

Nous avons donc supplémentés le milieu avec de l' α -cétobutyrate et l' α -cétisovalérate. Cela a permis de diminuer l'incorporation de carbone-13, en particulier vers les méthyles des Valines, Leucines et vers le C δ_1 de l'Isoleucine. En revanche, le marquage de la Thréonine reste inchangé car la voie métabolique allant de la Thréonine à l' α -cétobutyrate est irréversible. De même, l'incorporation de ^{13}C dans le C γ_2 de l'Isoleucine reste inchangée car ce méthyle provient directement du méthyle du pyruvate (Figure 2.6).

Pour limiter cette incorporation, nous avons supplémenté le milieu avec de l'Isoleucine. Compte tenu du coût de cet acide aminé deutéré (~ 2000 €/gr), nous avons cherché à en optimiser la concentration dans les cultures.

Le glycérol est utilisé comme source principale de carbone et le milieu est supplémenté en Alanine- $^{13}\text{C}_\beta$ à 800 mg.L^{-1} ainsi qu'en α -cétoisovalérate à 200 mg.L^{-1} (pour limiter le marquage ^{13}C des méthyles des Leucines et Valines). Le taux d'incorporation de carbone ^{13}C sur l'Isoleucine, en fonction de la concentration en Isoleucine ajoutée, est présenté sur la Figure 2.26. On voit qu'une concentration de 60 mg.L^{-1} d'Isoleucine est suffisante pour réduire l'incorporation de carbone-13 en position delta et gamma à moins de 3 %. Toutefois, dans ces conditions, le taux de fuite vers les méthyles des Thréonines (et Méthionines) reste évidemment inchangé car la voie de synthèse de l'Isoleucine est irréversible.

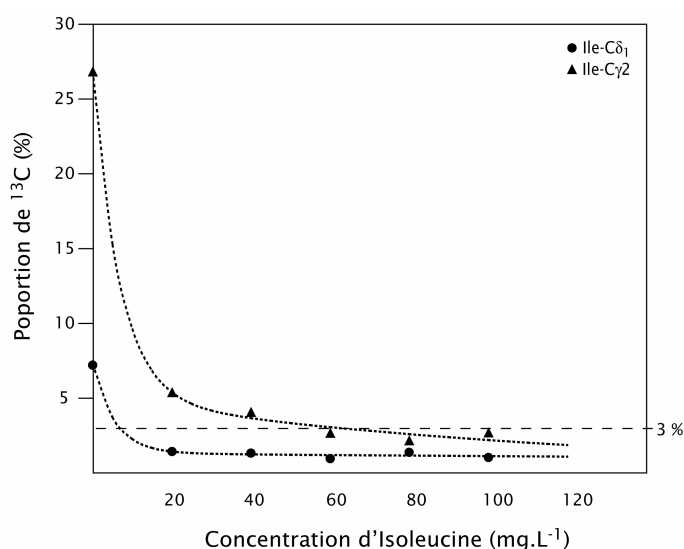


Figure 2.26 : Proportion des fuites métaboliques de ^{13}C vers les méthyles γ_2 et δ_1 de l'Isoleucine dans l'ubiquitine en fonction de la concentration d'Isoleucine.

2.6.4.4 Limitation des fuites métaboliques au niveau du cycle de Krebs et de la glycolyse

Afin de limiter les fuites métaboliques vers les acides aminées issus du cycle de Krebs (Asn, Lys, Asp, Thr, Met, Ile, Glu, Pro, Gln, Arg), nous avons testé l'effet de l'ajout d'intermédiaires de cette voie métabolique. Pour ce faire, l' α -cétoglutarate et l'oxaloacétate, deux de ces intermédiaires (Figure 2.3), ont été ajoutés avec l'Alanine et les autres précurseurs.

Sur la Figure 2.26 est représentée le taux d'incorporation du ^{13}C sur les CH_2 et CH des différents acides aminés, en fonction de la quantité de ces deux précurseurs ajoutés au milieu de culture. Les conditions de concentration des différents précurseurs sont 800 mg.L^{-1} d'Alanine- $[\text{}^{13}\text{C}_\beta]$, 200 mg.L^{-1} d' α -cétoisovalérate, 60 mg.L^{-1} d'Isoleucine et du glycérol (5 g.L^{-1}) comme source principale de carbone. Comme prévu, on observe une baisse du niveau d'incorporation du ^{13}C pour les CH et CH_2 des résidus synthétisés à partir des intermédiaires du cycle de Krebs. Une valeur plateau, inférieure à 5 %, est obtenue pour des concentrations en α -cétooglutarate et d'oxaloacétate supérieures à 500 mg/L . Le méthyle de la Thréonine, provenant d'un intermédiaire du cycle de Krebs, voit son niveau de marquage ^{13}C diminuer de façon similaire aux CH et CH_2 .

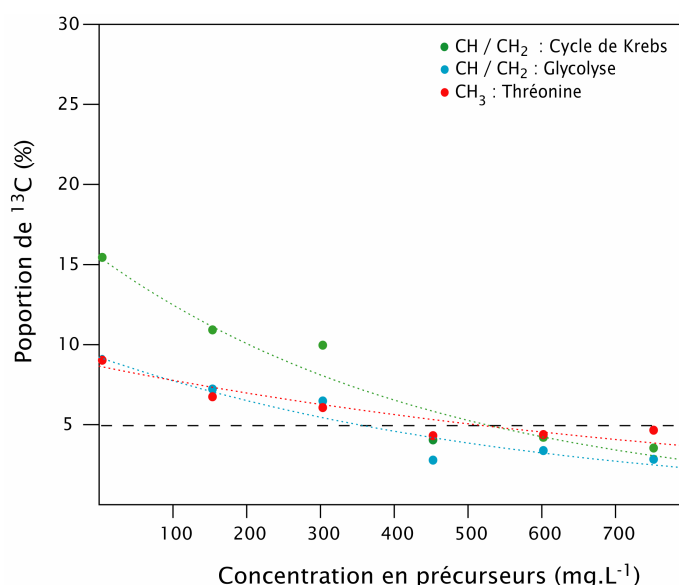


Figure 2.27 : Taux de marquage en carbone-13 des groupes CH_2 , CH et le méthyle de la Thréonine dans l'ubiquitine, en fonction de la quantité d'oxaloacétate et d' α -cétooglutarate, ajoutés dans le milieu de culture.

A chaque point, la concentration correspond à la concentration de chacun des précurseurs. Les CH / CH_2 du Cycle de Krebs correspondent aux résidus D, K, N, E, P, Q et R et les CH / CH_2 de la Glycolyse aux résidus L, Y, G, S.

De manière surprenante, le niveau d'incorporation du ^{13}C dans les résidus, produits à partir d'intermédiaires de la glycolyse, est aussi considérablement réduit par l'ajout d' α -cétooglutarate et d'oxaloacétate. Ceci est lié au fait que ces voies métaboliques ne sont pas indépendantes. La vue générale présentée sur la Figure 2.3, est très simpliste ; il existe de nombreuses voies métaboliques interconnectant la

glycolyse et le cycle de Krebs. Le glutamate, par exemple, est impliqué dans la synthèse de nombreux acides aminés rattachés à des intermédiaires de la glycolyse.

2.6.4.5 Caractérisation du marquage

En conclusion, l'ajout des précurseurs tels que l' α -cétobutyrate et l' α -cétoisovalérate est nécessaire pour limiter le recyclage du ^{13}C du carbone β de l'Alanine vers les méthyles des autres résidus. Cependant, l' α -cétobutyrate n'est pas suffisant, car il n'empêche pas la fuite métabolique vers le méthyle γ_2 des Isoleucines provenant du pyruvate. L'ajout d'Isoleucine est indispensable pour limiter ce recyclage. De même, l'utilisation d' α -cétoglutarate et d'oxaloacétate permet de limiter l'incorporation de carbone-13 dans les résidus issus du métabolisme de la glycolyse ou du cycle de Krebs. L'utilisation du glycérol, comme source principale de carbone, paraît aussi intéressante dans le cas de la culture en D_2O . Le glycérol ^{12}C , ^2H est en effet deux fois moins onéreux que le glucose ^{12}C , ^2H . Enfin, l'utilisation du pyruvate est à éviter, si l'on désire obtenir un échantillon avec un marquage maximal des Alanines. En revanche, il permet d'obtenir, pour un moindre coût, un échantillon avec un très faible taux d'incorporations parasites, dans les cas où la dilution isotopique sur la position β n'est pas rédhibitoire pour les expériences envisagées.

En résumé, un protocole de marquage spécifique ^{13}C en position C_β de l'Alanine a été mis au point. Un marquage de la position C_β de l'Alanine à plus de 95 % est possible avec des fuites métaboliques inférieures à 3 % (voir Figure 2.28) en supplémentant le milieu de culture M9 par :

- | | |
|--|---|
| - 800 mg.L^{-1} d'Alanine- $^{13}\text{C}_\beta$]. | - 500 mg.L^{-1} d' α -cétoglutarate |
| - 200 mg.L^{-1} d' α -cétoisovalérate. | - 500 mg.L^{-1} d'oxaloacétate |
| - 60 mg.L^{-1} d'Isoleucine. | - 5 g.L^{-1} de glycérol |

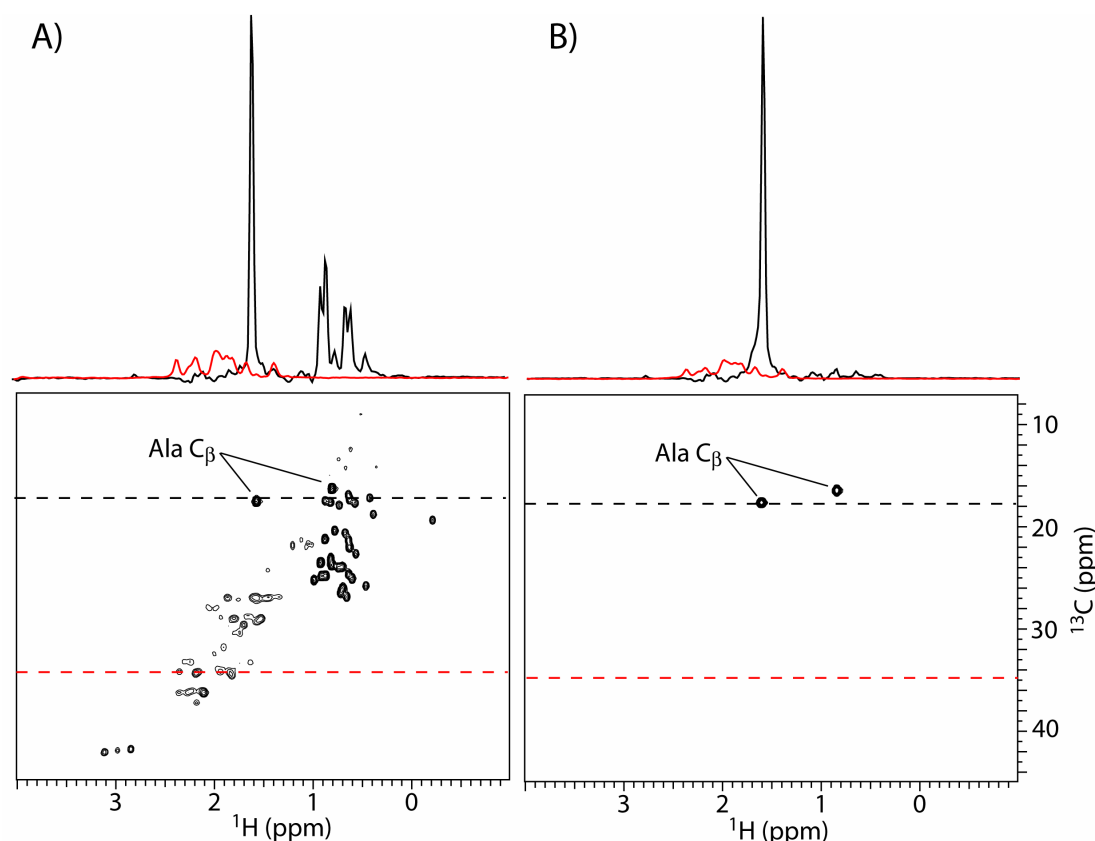


Figure 2.28 : Spectres 2D de corrélation HSQC ^1H - ^{13}C de l'ubiquitine.

Les spectres ont été enregistrés à 25°C dans D_2O sur un spectromètre 600 MHz, équipé d'une sonde cryogénique. Le premier contour correspond à un signal égal à 5 % de l'intensité des pics de corrélation des méthyles des Alanines. A) Culture supplémentée par 800 mg.L^{-1} d'Alanine (équivalent à la Figure 2.25.A) et B) Culture supplémentée par 800 mg.L^{-1} d'Alanine, 60 mg.L^{-1} d'Isoleucine, 200 mg.L^{-1} d' α -cétoisovalérate, 5 g.L^{-1} de glycérol, 500 mg.L^{-1} d' α -cétoglutarate, et 500 mg.L^{-1} d'oxaloacétate. Les sections transversales F2 à deux fréquences ^{13}C sont représentées pour chacun des spectres. Le pic de l'Alanine (section transversale noire) correspond à 100 % de marquage ^{13}C en A et en B.

2.6.5 Protonation spécifique du méthyle de l'Alanine

2.6.5.1 Passage au solvant deutéré

La préparation d'un échantillon perdeutééré et spécifiquement protoné sur le méthyle de l'Alanine nécessite la deutérioration des précurseurs métaboliques du protocole ci-dessus. De plus, l'hypothèse que nous avons faite au début de cette étude, concernant le fait que les protons liés au ^{13}C ne soient pas échangés pendant la culture, reste encore à vérifier. Pour cela, la production d'un échantillon d'ubiquitine en D_2O , avec de l'Alanine protonée, ou en H_2O , avec de l'Alanine deutérée, doit être

entreprise. Des méthodes peu onéreuses doivent être utilisées pour l'obtention de l'ensemble des précurseurs métaboliques deutérés.

2.6.5.2 Synthèse de l'Alanine spécifiquement protonée

L'Alanine spécifiquement deutérée en α et marquée ^{13}C en β n'est pas disponible commercialement. Toutefois, plusieurs voies de synthèse ont été décrites dans la littérature, soit par échange enzymatique du proton α (Níbeilliú *et al*, 2004), soit par échange chimique (Tenenbaum *et al*, 1974). Nous avons envisagé de réaliser la synthèse par la première voie. Le protocole de synthèse de ce composé se fait à partir de la L-Alanine [$^{13}\text{CH}_3$]- β , disponible commercialement (150 €/g), et en présence de la *tryptophan synthase*. Cette enzyme est capable d'échanger le H_α de l'Alanine par un deutérium du solvant en conservant la configuration absolue du carbone. L'échange est fait directement à partir de l'extrait cellulaire brut d'une culture d'*E. coli* contenant la *tryptophan synthase* surexprimée (vecteur d'expression commercial : ATCC 37845). L'échange se réalise en 48 heures, à 25°C, en présence de 6 mL d'extrait cellulaire brut contenant la *tryptophan synthase* (équivalent à 1 L de culture dans *E. coli*) pour 1 g d'Alanine. L'échange est suivi directement par RMN et arrêté par un choc thermique à 75 °C. L'ensemble du mélange réactionnel contenant l'Alanine spécifiquement protonée sur son méthyle est ajouté au milieu de culture une heure avant l'induction.

Les précurseurs et les acides aminés limitant les fuites métaboliques

L'Isoleucine et l' α -cétoisovalérate sont tous deux disponibles commercialement sous forme perdeutérée (respectivement 1500 €/g et 500 €/g). L'oxaloacétate est facilement deutérée par échange avec le solvant comme cela a été décrit pour l' α -cétobutyrate ou l' α -cétoisovalérate. Pour l' α -cétooglutarate, cet échange n'est pas aussi trivial. Toutefois, au regard de la Figure 2.3, ces deux précurseurs peuvent être remplacés par le succinate, qui est disponible sous forme deutéré (~ 30 €/g). Les cultures sont donc réalisées avec le glycérol deutéré (2,5 g/L) et le succinate (2,5 g/L) comme sources de carbone, afin de diluer toutes les fuites métaboliques de l'Alanine vers les acides aminés synthétisés, à partir des précurseurs de la glycolyse et du cycle de Krebs.

2.6.5.3 Résultats

Le protocole de protonation des méthyles de l'Alanine a été utilisé pour produire un échantillon d'ubiquitine U-[^2H], U-[^{12}C], U-[$^{13}\text{C}^1\text{H}_3$]-Ala- β . Un spectre de corrélation ^1H - ^{13}C est présenté sur la Figure 2.29.B. Les fuites métaboliques sont toutes inférieures à 0,5 % ; ceci s'explique par l'utilisation du double marquage isotopique ^2H et ^{13}C . Un seul pic de corrélation, bien résolu pour l'isotopomère CH_3 du méthyle de l'Alanine, est observé. Les deux autres isotopomères (CHD_2 et CD_2H) ne sont pas détectés. Ce résultat valide notre hypothèse initiale (§ 2.6.2) : les protons du méthyle de l'Alanine restent liés au carbone lors de l'incorporation de l'Alanine.

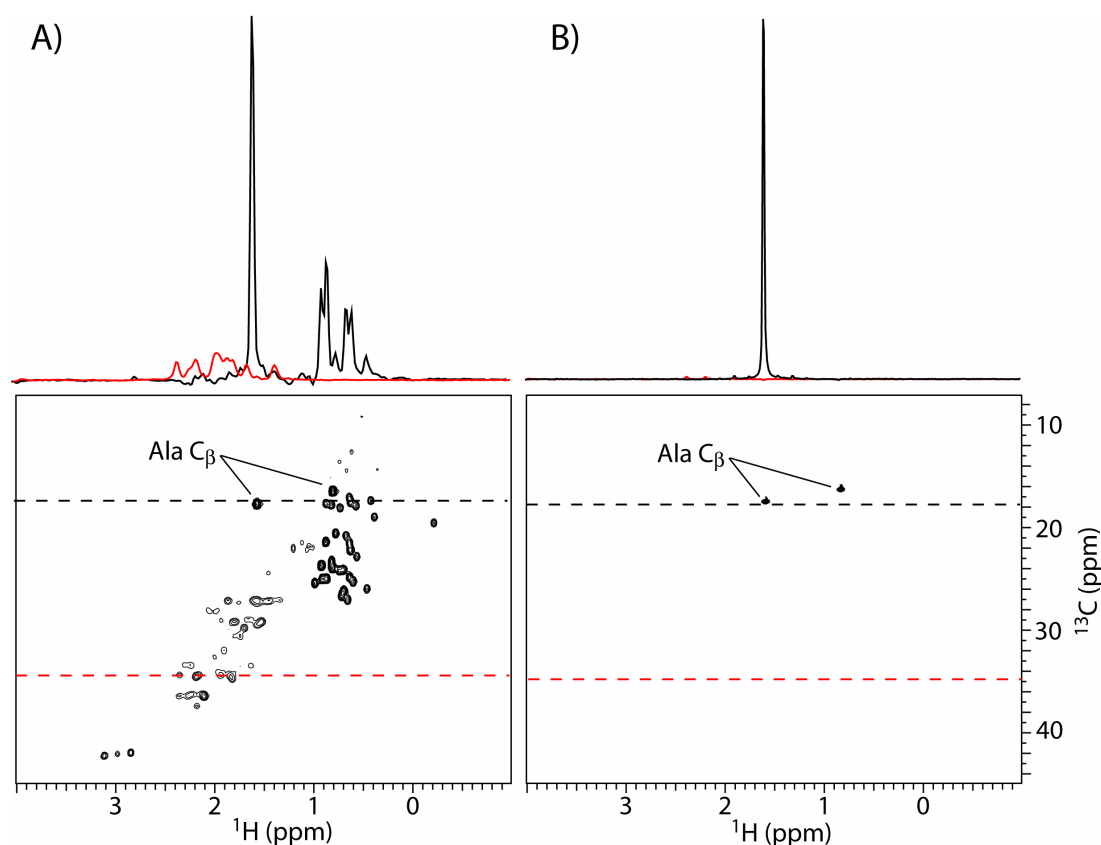


Figure 2.29 : Spectres 2D de corrélation HSQC ^1H - ^{13}C de l'ubiquitine.

Les spectres ont été enregistrés à 25°C, dans D_2O , sur un spectromètre 600 MHz, équipé d'une sonde cryogénique. Le premier contour correspond à un signal égal à 5 % de l'intensité des pics de corrélation des méthyles des Alanines. A) Culture supplémentée avec 800 mg.L^{-1} d'Alanine (équivalent à la Figure 2.25.A) et B) ubiquitine U-[^2H], U-[^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ala- β . Les sections transversales F2 à deux fréquences ^{13}C sont représentées pour chacun des spectres. Le pic de l'Alanine (section transversale noire) correspond à 100 % de marquage ^{13}C en A et en B.

Il faut noter que tout le travail présenté ici, sur ce protocole de marquage spécifique, a été réalisé de manière indépendante aux travaux de Isaacson *et coll.* (Isaacson *et al*, 2007) publiés 15 jours avant la fin de la rédaction de ce manuscrit. Notre protocole est meilleur sur deux points. Tout d'abord, l'incorporation des méthyles est bien supérieure à celle réalisée par leur protocole utilisant un milieu riche. Enfin, le coût de production est réduit en utilisant une culture en milieu minimum par rapport à l'utilisation d'un milieu riche deutéré.

Chapitre

3

*Mesure précise des distances
entre méthyles éloignés dans
les protéines spécifiquement
protonées*

3.1 Détection de l'effet Overhauser nucléaire (NOE) à longue portée

3.1.1 Introduction

Il a été démontré que l'utilisation de protéines perdeutérées permettait la détection de NOEs entre protons H_N distants jusqu'à 7, voire 8 Å (Koharudin *et al*, 2003). Peut-on détecter des NOEs au-delà de 8 Å ? Dans ce chapitre, je démontrerai que des NOEs, d'un à deux ordres de magnitude plus faible, peuvent être détectés. En effet, en utilisant des protéines perdeutérées et protonées spécifiquement sur certains méthyles, on bénéficie de trois effets majeurs (§ 2.3) :

(i) L'aimantation transférée entre deux méthyles éloignés est neuf fois plus intense que pour une paire de protons séparés par une distance équivalente, grâce à la multiplicité des protons du méthyle.

(ii) Dans les protéines totalement deutérées, les temps caractéristiques de relaxation transversale (T_2) et longitudinale (T_1) des méthyles sont favorables, donnant lieu à une bonne résolution et à un bon rapport signal sur bruit dans les spectres RMN.

(iii) Enfin, la diffusion de spins est réduite, du fait de la faible densité de protons. Ce prérequis est indispensable pour la précision des mesures.

Pour démontrer la faisabilité de cette approche, j'ai choisi d'utiliser deux protéines : l'ubiquitine, un système modèle de 9 kDa, pour lequel une structure haute résolution en solution est disponible (PDB 1D3Z, Cornilescu *et al*, 1998), et la γ S-crystalline, une protéine de 19 kDa comprenant deux domaines, pour laquelle une structure a été déterminée de manière indépendante en solution (Grishaev *et al*, 2005 ; Wu *et al*, 2005). Ces deux protéines ont été préparées selon le protocole décrit au § 2.4.3, avec comme unique source de protons lors de l'expression, l' α -cétobutyrate, précurseur pour le méthyle δ_1 des Isoleucines. Dans cette partie, je rappellerai le principe de l'effet Overhauser nucléaire (NOE), les bases de l'expérience NOESY ainsi que l'optimisation des différents paramètres pour la détection de NOEs de faible

magnitude dans un système spécifiquement protoné. Ensuite, je démontrerai que, dans ces systèmes, la détection de NOEs entre méthyles séparés par plus de 10 Å est possible. Cependant, compte tenu de la faible intensité de ces NOEs, j'insisterai sur l'attention particulière dont il faut faire preuve pour analyser les spectres NOESY. Enfin, j'exposerai la méthode mise au point pour extraire des contraintes de distances précises de ces NOEs, afin de les utiliser par la suite dans un calcul de structure.

3.1.2 Principe de l'effet Overhauser nucléaire

3.1.2.1 Découverte de l'effet Overhauser

Le fait de saturer les spins électroniques dans un système paramagnétique, tel que les métaux, peut conduire à une augmentation de la polarisation des spins nucléaires par un facteur 1000. Ceci vient, en partie, du fait que le rapport gyromagnétique de l'électron est trois ordres de magnitude plus élevé que le rapport gyromagnétique d'un proton. Ce phénomène fut prédit par Albert Overhauser, au début des années 50, et démontré expérimentalement par C. P. Slitcher et T. R. Carver, en 1953 (Carver *et al*, 1953). La physique sous jacente à l'effet Overhauser est une interaction dipolaire entre deux moments magnétiques.

3.1.2.2 Description de l'effet NOE dans un système de 2 protons isolés

Si l'on considère une molécule contenant deux protons isolés 1 et 2, sans couplage scalaire, le spectre RMN ^1H consiste en deux singulets aux fréquences Ω_1 et Ω_2 ppm. Supposons maintenant que, pendant la période précédant l'excitation des spins, les spins 2 soient saturés par l'application d'un champ radiofréquence à la fréquence de résonance Ω_2 . Le premier effet de la saturation est l'équilibre des populations α et β ainsi que la disparition du signal du proton 2. Le second effet possible est la variation d'intensité du signal du proton 1, si les deux spins ont une interaction dipolaire non négligeable (Figure 3.1). C'est ce phénomène qui est connu sous le nom d'*effet Overhauser nucléaire*, NOE. La mesure du NOE entre deux spins peut donc être réalisée par deux expériences. Une première expérience de référence, où l'irradiation sélective n'est pas appliquée ; une seconde expérience, où l'aimantation d'un des deux

spins est irradiée jusqu'à saturation et suivie d'une impulsion de lecture. Enfin, le spectre NOE est obtenu en faisant la différence de ces deux spectres (Figure 3.1.C).

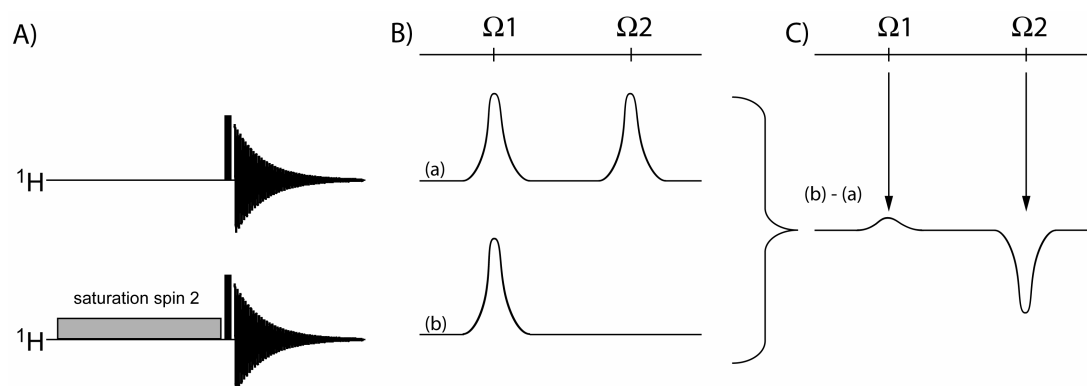


Figure 3.1 : Représentation de la mesure de NOE entre deux spins.

A) Les deux séquences d'impulsion pour enregistrer un spectre de différence de NOE permanent. En B) figurent les spectres obtenus par chacune des expériences et en C) est présentée la différence des deux spectres.

Pour comprendre l'origine du NOE, il est nécessaire de regarder les chemins possibles de relaxation spin-réseau pour la paire de spins nucléaires 1 et 2. La Figure 3.2 présente les quatre états d'énergie. Chaque niveau d'énergie a été écrit avec l'état de spin de chacun des spins couplés, où $\alpha\alpha$ (n_1) et $\beta\beta$ (n_4) sont à une énergie respective de $-\omega_0$ et $+\omega_0$, et $\alpha\beta$ (n_2) ainsi que $\beta\alpha$ (n_3) sont à une énergie nulle (seulement dans le cas de deux noyaux identiques).

A l'équilibre thermique, les populations relatives de ces quatre états sont $1+2\Delta$ pour $\alpha\alpha$, 1 pour $\alpha\beta$ et $\beta\alpha$, et $1-2\Delta$ pour $\beta\beta$ avec $\Delta = \frac{1}{2} \hbar \omega_0 / k_B T$.

Sur les six chemins possibles de relaxation des deux spins couplés, quatre correspondent à l'inversion d'un seul spin et aux processus de relaxation spin-réseau. Leurs vitesses sont notées $W_1^{(X,Y)}$; l'indice inférieur indique le changement du nombre de spin (± 1 dans ce cas) et les indices supérieurs, dans la parenthèse, indiquent quel spin est inversé (X) ainsi que l'état de spin du spin passif (Y). Les deux chemins restants correspondent aux processus de relaxation croisée pour lesquels les deux spins relaxent ensemble. Ces processus comportent une transition double-quantum (où le changement total du nombre de spins est de ± 2 , avec la probabilité de transition W_2) et une transition zéro-quantum (pour laquelle le nombre

de spins dans l'état α et β reste globalement identique, avec la probabilité de transition W_0). Ces deux processus sont des chemins supplémentaires permettant aux populations des états de spins de revenir à l'équilibre suivant certaines perturbations.

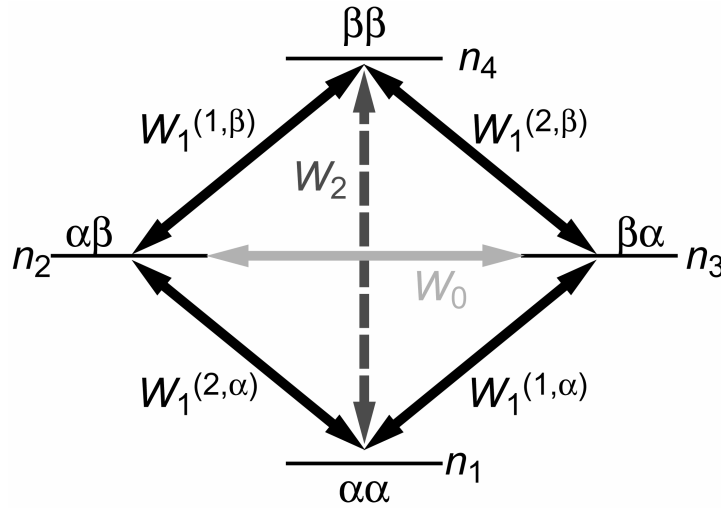


Figure 3.2 : Les transitions et les niveaux d'énergie pour un système à deux spins $\frac{1}{2}$ identiques.

Intéressons nous aux signaux RMN des deux spins : $M_{1,z}$ et $M_{2,z}$ sont proportionnelles à la différence des populations entre les états de spin.

$$M_{1,z} \propto (n_1 - n_3) + (n_2 - n_4) \quad (3.1)$$

$$M_{2,z} \propto (n_1 - n_2) + (n_3 - n_4)$$

La dérivée de $M_{1,z}$ en fonction du temps s'écrit :

$$\frac{d M_{1,z}}{dt} \propto \left(\frac{dn_1}{dt} - \frac{dn_3}{dt} \right) + \left(\frac{dn_2}{dt} - \frac{dn_4}{dt} \right) \quad (3.2)$$

Et, de la même manière pour le spin 2, on obtient la dérivée de $M_{2,z}$:

$$\frac{d M_{2,z}}{dt} \propto \left(\frac{dn_1}{dt} - \frac{dn_2}{dt} \right) + \left(\frac{dn_3}{dt} - \frac{dn_4}{dt} \right) \quad (3.3)$$

La variation de la population de l'état de spin $\alpha\alpha$ (n_1) est donnée par les différentes transitions présentées sur la Figure 3.2 et peut être écrite sous la forme :

$$\begin{aligned}
\frac{dn_1}{dt} \propto & \underbrace{-W_1^{(2,\alpha)}(n_1 - n_1^0) - W_1^{(1,\alpha)}(n_1 - n_1^0) - W_2(n_1 - n_1^0)}_{\text{perte à partir du niveau 1}} \\
& + \underbrace{W_1^{(2,\alpha)}(n_2 - n_2^0)}_{\text{gain à partir du niveau 2}} + \underbrace{W_1^{(1,\alpha)}(n_3 - n_3^0)}_{\text{gain à partir du niveau 3}} + \underbrace{W_2(n_4 - n_4^0)}_{\text{gain à partir du niveau 4}}
\end{aligned} \tag{3.4}$$

où n_i est la population du $i^{\text{ème}}$ niveau d'énergie à l'instant t , et n_i^0 est la population à l'état d'équilibre de ce même niveau d'énergie. Les trois premiers termes sont tous négatifs et correspondent à la diminution de la population de spins sur le niveau d'énergie n_1 . Chacun de ces termes dépend de la population du niveau d'énergie n_1 . Les termes positifs représentent respectivement les processus dans lesquels la population de l'état de spin n_1 est augmentée, aux dépens des autres niveaux d'énergie. Ces trois derniers termes dépendent de la population de chacun de ces niveaux d'énergie (n_2 , n_3 et n_4).

En décrivant, par ces équations, l'évolution de la population pour chacun des quatre niveaux d'énergie et en considérant le cas simple où les probabilités de transition simple quanta pour un spin sont égales ($W_1^{(1,\alpha)} = W_1^{(1,\beta)} = W_1^{(1)}$ et $W_1^{(2,\alpha)} = W_1^{(2,\beta)} = W_1^{(2)}$), on peut réécrire l'Equation 3.2 pour le spin 1 et l'Equation 3.3 pour le spin 2 sous la forme suivante (Solomon, 1955) :

$$\begin{aligned}
\frac{dM_{1,z}}{dt} &= -\rho^{(1)}(M_{1,z} - M_{1,z}^0) - \sigma^{21}(M_{2,z} - M_{2,z}^0) \\
\frac{dM_{2,z}}{dt} &= -\sigma^{21}(M_{1,z} - M_{1,z}^0) - \rho^{(2)}(M_{2,z} - M_{2,z}^0)
\end{aligned} \tag{3.5}$$

où

$$\begin{aligned}
\rho^{(1)} &= 2W_1^{(1)} + W_2 + W_0 \\
\rho^{(2)} &= 2W_1^{(2)} + W_2 + W_0 \\
\sigma^{12} &= \sigma^{21} = W_2 - W_0
\end{aligned} \tag{3.6}$$

3.1.2.3 Description de l'effet NOE dans un système de n protons

Si l'on considère maintenant la relaxation croisée, σ , pour un système de n spins en absence d'interaction scalaire spin-spin, la relaxation croisée de la composante de

l'aimantation longitudinale ΔM_z peut s'écrire, d'après les équations de Solomon (Equations 3.5) :

$$\frac{d\Delta M_z(t)}{dt} = -\sigma \Delta M_z(t) \quad (3.7)$$

où $\Delta M_z(t)$ représente la déviation de l'aimantation de sa valeur à l'équilibre thermique et σ est la matrice de relaxation. σ est une matrice à $n \times n$ éléments qui peut s'écrire :

$$\sigma = \begin{bmatrix} \rho_1 & \sigma_{12} & \sigma_{13} & \cdot \\ \sigma_{21} & \rho_2 & \cdot & \cdot \\ \sigma_{31} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \rho_n \end{bmatrix} \quad (3.8)$$

où l'élément σ_{ij} représente la vitesse de relaxation croisée entre le spin i et le spin j . L'élément ρ_i représente la contribution de la relaxation directe à partir de toutes les autres sources pour le spin i .

Ces termes peuvent être écrits en fonctions des probabilités de transition à zéro-, simple- ou double-quantum, comme précédemment :

$$\rho_i = \sum_{\substack{j=1 \\ j \neq i}}^n (W_0^{ij} + 2W_1^{ij} + W_2^{ij}) + R_{1i} \quad (3.9)$$

$$\sigma_{ij} = (W_2^{ij} - W_0^{ij})$$

où R_{1i} représente la relaxation longitudinale due aux autres mécanismes de relaxation (échanges, paramagnétiques, CSA, ...).

Les probabilités de transitions W_0 , W_1 et W_2 ont été établies théoriquement par Bloch, Wangsness et Redfield (Wangsness *et al*, 1953 ; Bloch, 1956 ; Redfield, 1957 ; 1965). Elles peuvent être exprimées à l'aide des fonctions de densité spectrale

contenant l'information sur la corrélation entre les interactions responsables de la relaxation des spins.

$$\begin{aligned}
 W_0^{ij} &= \frac{1}{10} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_i \gamma_j \hbar \right)^2 J(\omega_i - \omega_j) \\
 W_2^{ij} &= \frac{6}{10} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_i \gamma_j \hbar \right)^2 J(\omega_i + \omega_j) \\
 W_1^{ij} &= \frac{3}{20} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_i \gamma_j \hbar \right)^2 J(\omega_i)
 \end{aligned} \tag{3.10}$$

Ceci permet de réécrire les termes de la matrice de relaxation totale :

$$\begin{aligned}
 \rho_i &= \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{10} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_i \gamma_j \hbar \right)^2 [3J(\omega_i) + J(\omega_i - \omega_j) + 6J(\omega_i + \omega_j)] + R_{1i} \\
 \sigma_{ij} &= \frac{1}{10} \frac{1}{r_{ij}^6} \left(\frac{\mu_0}{4\pi} \gamma_i \gamma_j \hbar \right)^2 [6J(\omega_i + \omega_j) - J(\omega_i - \omega_j)]
 \end{aligned} \tag{3.11}$$

où γ_i et γ_j sont les rapports gyromagnétiques des noyaux i et j , r_{ij} la distance entre les deux noyaux, ω_i et ω_j sont les fréquences de Larmor associées à ces noyaux, μ_0 est la constante de perméabilité magnétique du vide et $\hbar$ est la constante de Planck divisée par 2π . Enfin $J(\omega)$ est la fonction de densité spectrale et contient l'information de dynamique.

3.1.2.4 Application aux protéines spécifiquement protonées sur les méthyles

Dans notre cas, j'ai considéré la protéine avec un mouvement isotrope ; la densité spectrale peut donc s'écrire :

$$J(\omega) = S^2 \frac{\tau_c}{1 + (\omega \tau_c)^2} \tag{3.12}$$

Le paramètre d'ordre S^2 tient compte des mouvements entre les deux méthyles. En sachant que dans une protéine structurée, la distance interméthyle entre un méthyle A et B est grande, comparée à la distance entre les protons intraméthyle (1.8 Å), le

paramètre d'ordre S^2 tend vers 1 (Annexe 2). Ainsi, en négligeant toutes les sources de dynamique interne, exceptée la rotation rapide des méthyles (Tropp, 1980), et en tenant compte de la multiplicité de ces derniers, la distance moyenne interméthyle $\langle r_{AB}^3 \rangle$ peut être extraite à partir de :

$$\sigma_{AB} \cong \frac{3\tau_C}{10} \left(\frac{\mu_0 \gamma_H^2 \hbar}{4\pi \langle r_{AB}^3 \rangle} \right)^2 \left(\frac{6}{1 + (2\omega\tau_C)^2} - 1 \right) \quad (3.13)$$

où γ_H est le rapport gyromagnétique du proton, τ_c est le temps de corrélation totale et ω est la fréquence du spectromètre.

3.1.3 Principe de l'expérience NOESY

La séquence d'impulsion NOESY à deux dimensions est présentée sur la Figure 3.3 (Jeener *et al*, 1979). Elle permet l'obtention d'un spectre à deux dimensions NOESY, où la présence d'un pic de corrélation NOESY à $\{\Omega_i, \Omega_j\}$ nous informe de l'existence d'un transfert d'aimantation du spin i vers le spin j , due au terme de la relaxation croisée σ_{ij} .

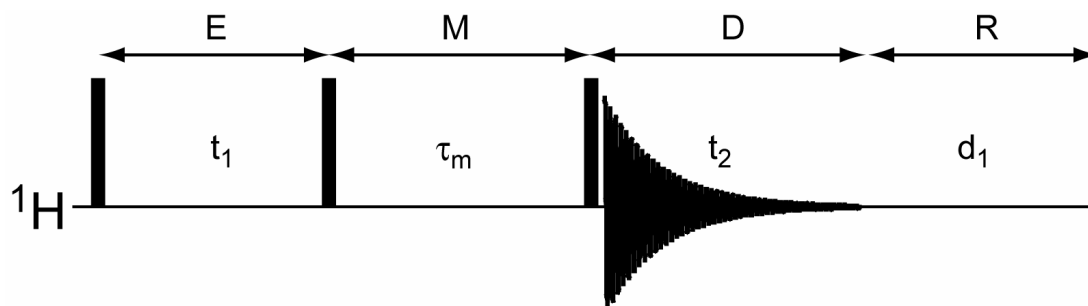


Figure 3.3 : Séquence d'impulsion de base de l'expérience NOESY.

Les quatre périodes sont notées avec une lettre : (E) Evolution, (M) Mélange, (D) Détection et (R) Récupération. Chacune de ces périodes durent respectivement t_1 , τ_m , t_2 et d_1 . Les impulsions rectangulaires sont appliquées avec des angles d'impulsion de 90° et le long de l'axe X.

La séquence d'impulsion est divisée en quatre parties. La première période (E), c'est-à-dire 90° - t_1 - 90° , est la période d'évolution et d'édition de la dimension indirecte. La première impulsion 90° crée une aimantation transversale. Les composantes de l'aimantation vont alors précesser librement pour permettre l'édition des fréquences de ces composantes durant le délai d'évolution t_1 . La deuxième

impulsion 90° restitue une partie de l'aimantation sur l'axe z. L'aimantation ainsi placée sur z est modulée en $\cos(\Omega_i t_1)$ et diffère pour chaque spin. Cette deuxième impulsion 90° marque le début de la période de mélange (M) de durée τ_m (Equation 3.7). Pendant cette période, les composantes longitudinales de l'aimantation de chaque spin placé hors d'équilibre peuvent évoluer sous l'influence de relaxations croisées durant le temps τ_m . A la fin du temps de mélange, une troisième impulsion 90° est appliquée pour créer une aimantation transversale qui sera détectée pendant la durée t_2 . Enfin, la dernière période est le délai de récupération (R). C'est un délai d'attente (d_1) qui permet, dans des conditions stationnaires, de restaurer l'aimantation M_z après la détection du signal par la relaxation longitudinale.

3.1.3.1 Le temps de récupération (d_1)

Pendant la période d'acquisition (t_2) et le délai de récupération (d_1), l'aimantation longitudinale M_z est (partiellement) restituée par les processus de relaxation le long de l'axe z. Dans les conditions stationnaires, l'aimantation $M_z(t)$ peut être exprimée par :

$$M_z(T_{rec}) = M_z(0)(1 - e^{(-T_{rec}/T_1)}) \quad (3.14)$$

où T_{rec} correspond à la somme des périodes d'acquisition et de récupération ($t_2 + d_1$) et $M_z(0)$ est l'aimantation de Boltzmann à l'équilibre.

Plus T_{rec} sera élevé, plus le signal détecté dans chaque acquisition sera élevé. Mais, inversement, plus le temps de récupération sera élevé, plus le nombre d'accumulations sera faible à temps expérimental constant. Si on accumule n acquisitions, le bruit augmente avec $\sqrt{n}$, alors que le signal, lui, augmente en n . Le rapport signal sur bruit d'une expérience RMN peut s'écrire, si l'on considère qu'une accumulation dure ($T_{rec} + T_{seq}$) :

$$\left(\frac{S}{N}\right) \propto (1 - e^{(-T_{rec}/T_1)}) \frac{1}{\sqrt{T_{rec} + T_{seq}}} \quad (3.15)$$

avec T_{seq} la somme des périodes d'évolution et de mélange.

En pratique, pour la plupart des expériences RMN standard (COSY, TOCSY, HSQC, ...), T_{seq} est négligeable devant T_{rec} , et on peut estimer le temps de récupération par la relation suivante (Ernst *et al*, 1987) :

$$T_{\text{rec}}((S/N)^{\text{max}}) \approx 1,256 \times T_1 \quad (3.16)$$

Dans des expériences NOESY, T_{seq} n'est plus négligeable devant T_{rec} ; le délai de récupération optimale (d_1) peut donc être directement déterminé à partir de la mesure de la relaxation longitudinale T_1 et en utilisant l'Equation 3.15.

3.1.3.2 Le temps de transfert (τ_m)

Le second paramètre important à optimiser est le temps de mélange pour que le transfert NOE, entre les méthyles Ile- δ_1 protonés, soit optimal. Dans le cas d'un échantillon marqué spécifiquement sur les méthyles δ_1 des Isoleucines, la densité de protons est considérablement réduite. Pour des protéines de tailles modérées avec un temps de corrélation faible, en considérant les méthyles comme isolés ($d_{HH}^{\text{intra-méthyle}} \ll d_{HH}^{\text{inter-méthyle}}$), l'intensité du signal transféré d'un méthyle i vers un méthyle j peut être approximée par l'Equation 3.17 :

$$M_{ij} \approx M_i^0 \sigma_{ij} \tau_m e^{-(R_{1i} + \sum_j \sigma_{ij}) \tau_m} \quad (3.17)$$

Où les termes, σ_{ij} et R_{1i} , sont définis dans l'équation 3.9 et M_i^0 représente l'aimantation initiale du spin i .

Le temps de mélange optimal durant lequel les deux méthyles échangent de l'aimantation est :

$$\tau_m^0 = \frac{1}{(R_{1i} + \sum_k \sigma_{ik})} < T_1 \quad (3.18)$$

pour lequel l'aimantation maximale transférée est $M_i^0 \frac{\sigma_{ij}}{(R_{1i} + \sum_j \sigma_{ij})} e^{-1}$.

3.1.4 Optimisation des paramètres d'acquisition pour la mesure de NOEs interméthyles dans les protéines spécifiquement protonées

3.1.4.1 Détermination de la relaxation longitudinale T_1

Comme nous l'avons vu ci-dessus, le délai de récupération (d_1) et le temps de mélange (τ_m) de l'expérience NOESY doivent être optimisés pour tenir compte de la vitesse de la relaxation longitudinale.

Le temps de relaxation longitudinale ^1H est mesuré par une expérience d'inversion-récupération avec un délai de récupération $\gg T_1$ (en pratique ~ 10 s). L'aimantation en fonction du temps prend la forme suivante :

$$M_z(t) = M_z^0 [1 - 2e^{(-t/T_1)}] \quad (3.19)$$

La valeur de la relaxation longitudinale T_1 peut être déterminée par un ajustement de l'Equation 3.14 ou en déterminant la valeur pour laquelle $M_z(t) = 0$. Lorsque le signal est nul, on détermine le temps de relaxation longitudinale T_1 , avec la relation suivante :

$$T_1 = \tau_{nul} / \ln 2 \quad (3.20)$$

où τ_{nul} est le délai pour lequel le signal est égal à zéro. Cette simple relation permet d'estimer rapidement le temps de relaxation longitudinale T_1 moyen pour un ensemble de spins, dans une protéine, par une série d'expériences à une dimension. L'ajustement individuel de chaque signaux permet de déterminer précisément la valeur du temps de relaxation T_1 pour chaque proton. Les valeurs de relaxation spin-réseau proton, mesurées sur chacun des méthyles δ_1 des Isoleucines dans l'échantillon d'ubiquitine $\text{U-}[^2\text{H}]$, $\text{U}[^{13}\text{C}]$, $[^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$, sont présentées sur le Tableau 3.1.

$[\text{CH}_3]\text{-Ile-}\delta_1$	T_1
3	2,2 s
13	1,6 s
23	0,9 s
30	1,9 s
36	1,7 s
44	1,7 s
61	1,1 s
Moyen	$1,6 \pm 0,4$ s

Tableau 3.1 : Relaxation longitudinale pour les méthyles δ_1 des Isoleucines de l'ubiquitine U- $[\text{}^2\text{H}]$, U $[\text{}^{13}\text{C}]$, $[\text{}^{12}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$ à 800MHz.

Les mesures ont été réalisées à 25°C. L'incertitude sur les valeurs de T_1 est de $\pm 0,1$ s.

3.1.4.2 Le temps de récupération optimal (d_1)

On utilise l'Equation 3.15 pour déterminer les temps de récupération optimaux, en fonction de T_{seq} , pour différents temps de relaxation longitudinale (Figure 3.4). Le temps de récupération optimal croît à la fois avec le temps de relaxation longitudinale et le temps de la séquence.

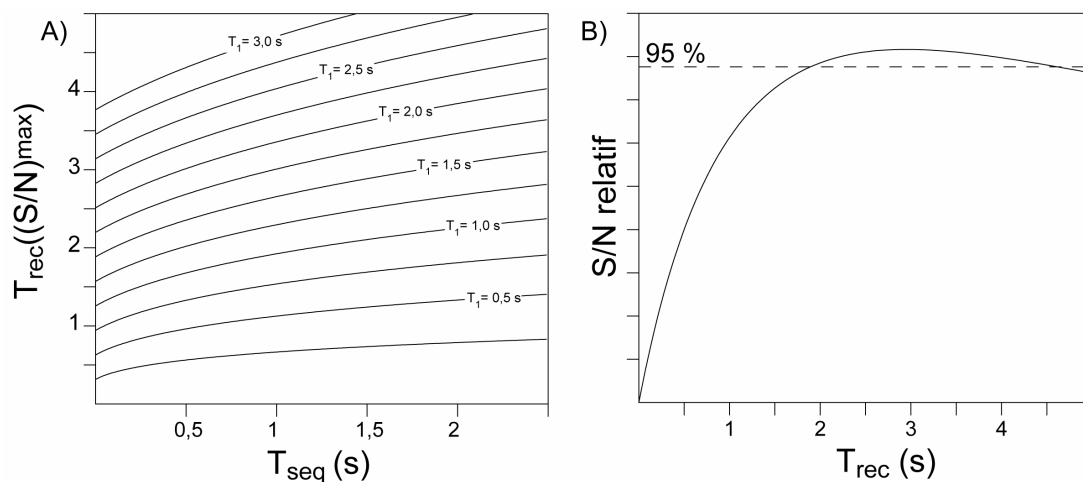


Figure 3.4 : A) Temps de récupération optimaux en fonction du temps d'expérience, pour différents temps de relaxation longitudinale. B) Intensité du signal relatif en fonction du temps de récupération.

Les valeurs de $T_{\text{rec}}((S/N)^{\text{max}})$ ont été déterminées à partir de l'Equation 3.17 pour différents temps de relaxation longitudinale indiqués sur les courbes. L'intensité du signal sur bruit a été déterminée à partir de l'Equation 3.15 et en prenant un temps de relaxation longitudinale de 1,6 s, équivalent à l'ubiquitine et $T_{\text{seq}} = 1,3$ s.

Le délai de récupération optimal pour l'ubiquitine spécifiquement protonée est de l'ordre de 3 s pour une expérience NOESY de durée $T_{\text{seq}} = 1,3$ s. On peut réduire ce délai à 1,9 s en ayant toujours 95 % du signal (Figure 3.4.B).

3.1.4.3 Le temps de transfert optimal (τ_m^0)

En pratique, une série d'expériences NOESY a été enregistrée, pour chacune des protéines, afin de déterminer le temps de mélange optimal. La Figure 3.5 représente la variation d'intensité de différents pics de corrélation NOEs interméthyles pour l'ubiquitine. Lorsque le bruit n'est pas pris en compte, le transfert NOE est maximal pour des temps de mélanges de $1,5 \pm 0,5$ s, dans le cas de l'ubiquitine. Cette valeur moyenne est proche des valeurs de T_1 mesurées pour les méthyles de cette protéine ($T_1^{\text{moy}} = 1,6$ s $\pm$ 0,4 s ; Tableau 3.1). Cependant, un temps de mélange élevé augmente considérablement la durée de chaque accumulation, T_{sec} . Si nous tenons compte de la durée de l'expérience (Figure 3.5.B), les temps de mélanges optimaux passent à $1,1 \pm 0,4$ s. Les NOEs entre méthyles éloignés étant bien plus faibles que la vitesse de relaxation longitudinale (ρ), le transfert est maximal pour des temps de mélange proche de T_1 .

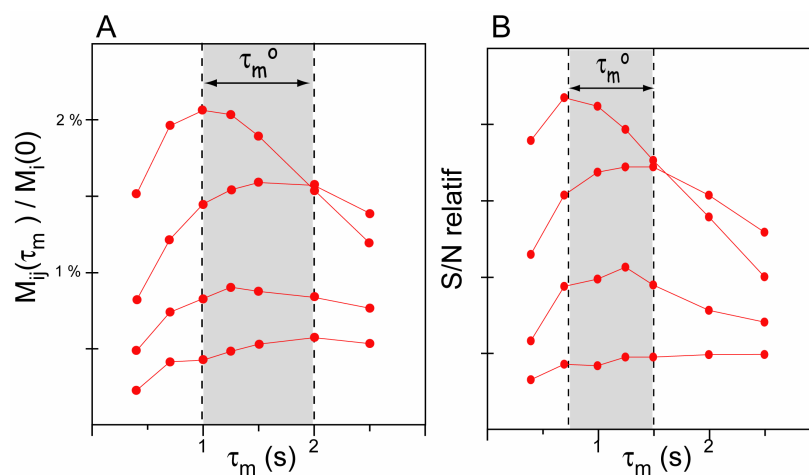


Figure 3.5 : Les courbes typiques des "Build-up" de NOEs de l'ubiquitine.

Les données ont été extraites à partir d'expériences NOESY enregistrées avec différents τ_m , sur l'ubiquitine U- ^{2}H], U- ^{13}C], $^{12}\text{C}^{1}\text{H}_3$]-Ile- δ_1 , à 25°C, dans 100% D_2O , sur un spectromètre fonctionnant à une fréquence ^1H de 800 MHz, équipé d'une sonde cryogénique. Le maximum de l'intensité relative des NOEs est pour un temps de mélange entre 1 et 2 s en A. Le rapport signal sur bruit relatif, pour tenir compte du temps expérimental, est représenté en B ; le temps de mélange optimal est de $1,1 \pm 0,4$ s.

3.1.5 Détection de faibles NOEs dans les protéines spécifiquement protonées

Pour l'observation de faibles NOEs entre des groupes méthyles spécifiquement protonés, dans des échantillons deutérés sur les autres sites, une attention particulière doit être prise sur les protons résiduels. Même à un niveau de deutération aussi haut que 98%, un méthyle, avec un niveau de protonation résiduelle de 2% et séparé par 5 Å d'un méthyle Ile- δ_1 totalement protoné, donne lieu à un pic de corrélation NOE plus intense qu'un pic de corrélation NOE entre deux méthyles Ile- δ_1 , totalement protonés et séparés par plus de 10 Å. Une analyse statistique de plusieurs protéines globulaires indique que chaque méthyle Ile- δ_1 est entouré, en moyenne, par sept autres méthyles éloignés par moins de 5 Å. Une attention particulière doit donc être prise pour l'analyse de NOEs entre des méthyles très éloignés, car des méthyles proches, faiblement protonés, peuvent donner lieu à des NOEs indésirables, qualifiés d'artéfacts par la suite.

Pour démontrer l'importance de tels artéfacts, j'ai enregistré une expérience NOESY (Figure 3.3), avec un temps de mélange optimal, sur un échantillon d'ubiquitine U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 , d'une concentration de 1 mM. Le spectre obtenu est représenté sur la Figure 3.6. Plusieurs pics de corrélation observés (carrés rouges sur la Figure 3.6) ont une seule fréquence proton en commun avec les fréquences ^1H des 7 méthyles δ_1 des Isoleucines. Ils correspondent aux artéfacts décrits ci-dessus.

Connaissant l'attribution et la structure pour l'ubiquitine (Weber *et al*, 1987 ; Cornilescu *et al*, 1998), j'ai déterminé que certains de ces artéfacts pouvaient apparaître aux fréquences potentielles des NOEs, entre les méthyles δ_1 des Isoleucines (carrés bleus sur la Figure 3.6). Il est donc impératif de pouvoir éliminer ces artéfacts afin d'identifier les NOEs de manière non ambiguë, c'est-à-dire les NOEs entre deux méthyles Ile- δ_1 éloignés.

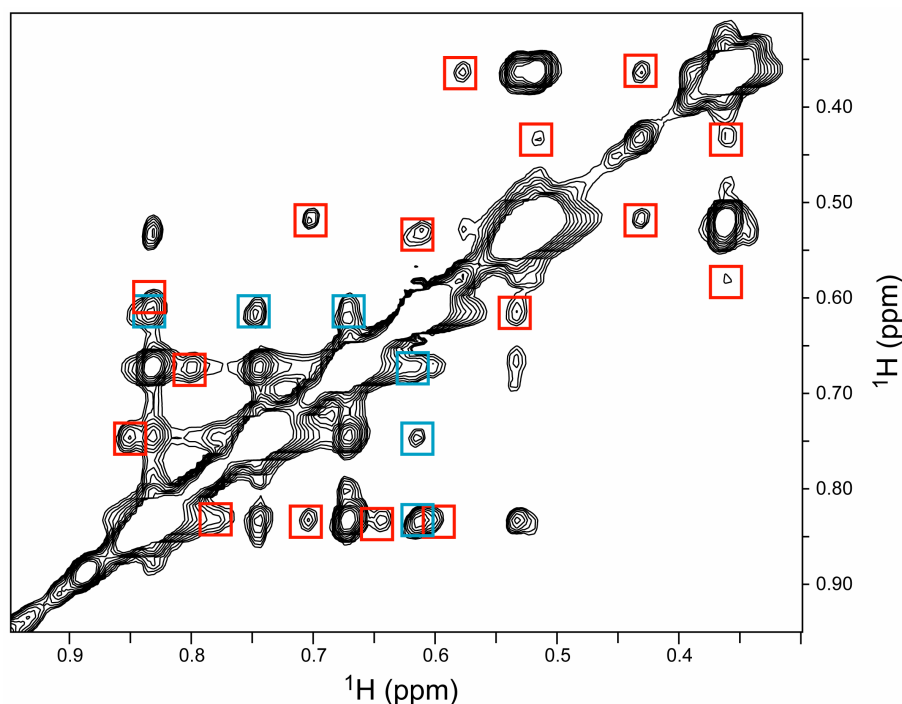


Figure 3.6 : Spectre NOESY de l'ubiquitine U-[^2H], U-[^{12}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 .

Il a été enregistré durant 14 heures avec $\tau_m = 1,25$ s et $d_1 = 1,5$ s, sur un échantillon à 1 mM, sur un spectromètre 800 MHz, équipé d'une sonde cryogénique. Les artefacts, ayant une de leurs fréquences identiques à une fréquence des méthyles δ_1 des Isoleucines, sont indiqués par un cadre rouge. Les carrés bleus correspondent à des artefacts ayant leurs deux fréquences identiques aux fréquences des méthyles δ_1 des Isoleucines.

Pour supprimer ces artefacts, j'ai développé une approche générale basée sur l'utilisation d'expériences NOESY filtrées carbones, en combinaison avec l'utilisation de marquages spécifiques à la fois sur le carbone et les hydrogènes (§ 2.4.4). Une première stratégie consiste à utiliser un marquage U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 , puis à ajouter un filtre ^{12}C dans l'expérience NOESY (Figure 3.7), pour éliminer le signal des protons résiduels liés à un carbone-13, c'est-à-dire tous les CH, CH₂ et CH₃ autres que les méthyles δ_1 des Isoleucines. Alternativement, des expériences NOESY filtrées ^{13}C ou éditées en 3D voire 4D (Tugarinov *et al*, 2005b), en combinaison avec un échantillon U-[^2H], U-[^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 , peuvent aussi être utilisées pour filtrer ces NOEs indésirables.

Le premier type de marquage est plus souhaitable pour observer des NOEs à longue distance, car la relaxation proton du méthyle n'est pas affectée par la forte interaction dipolaire ^{13}C - ^1H . La seconde stratégie, plus coûteuse, doit être utilisée pour les plus grosses protéines, en cas de superposition des signaux des Ile- δ_1 dans la dimension proton. L'efficacité de l'expérience NOESY filtrée ^{12}C , pour éliminer ces

artéfacts, est démontrée pour l'ubiquitine. L'utilisation de ce type de filtre permet de réduire considérablement ces artéfacts en dessous du seuil de détection (Figure 3.8).

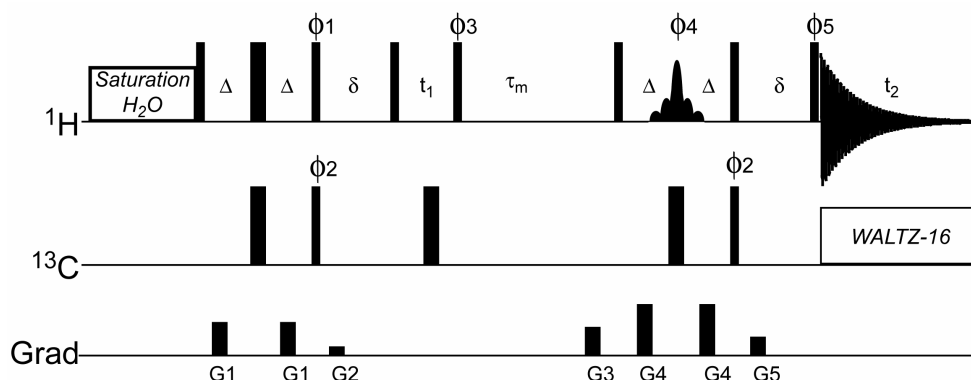


Figure 3.7 : Schéma d'impulsion de l'expérience NOESY avec filtre ^{13}C .

Les impulsions rectangulaires fines et larges sont appliquées avec des angles d'impulsion respectivement de 90° et de 180° . Sauf indication contraire, les impulsions sont appliquées le long de l'axe X. La fréquence de la porteuse ^1H est placée au centre de la région méthyle des Ile- δ_1 et la fréquence ^{13}C est placée au centre de la région C_β des Isoleucines à 27 ppm. L'impulsion sélective ^1H est de type 180° REBURP (Geen *et al*, 1991) (la durée est de 3 ms à 800 MHz, pour une largeur de bande d'inversion de 1600 Hz, centrée sur la région des méthyles Ile- δ_1 à 0,6 ppm). La durée et la puissance des gradients (en unités ms, G/cm) sont : G1 = (0.2 ; 26), G2 = (1 ; 6), G3 = (1 ; 22), G4 = (0.2 ; 58), G5 = (1 ; 14). Le cycle de phase est : $\phi_1 = 4(x), 4(-x)$; $\phi_2 = x, -x, x, x, -x, x, -x, -x, x, -x, x, x, -x, x, -x, -x$; $\phi_3 = 8(x)$; $8(-x)$; $\phi_4 = 2(x), 2(y)$; $\phi_5 = x, -x$; rec = $x, -x, -x, x, -x, x, x, -x, -x, x, x, -x, x, -x, -x, x$. La détection en quadrature en t_1 est obtenue en incrémentant ϕ_3 de la manière usuelle du States-TPPI (Marion *et al*, 1989b). Les délais de récupération des gradients ont été placés à $\delta = 1$ ms et le délai de transfert $\Delta = 2\text{ms}$.

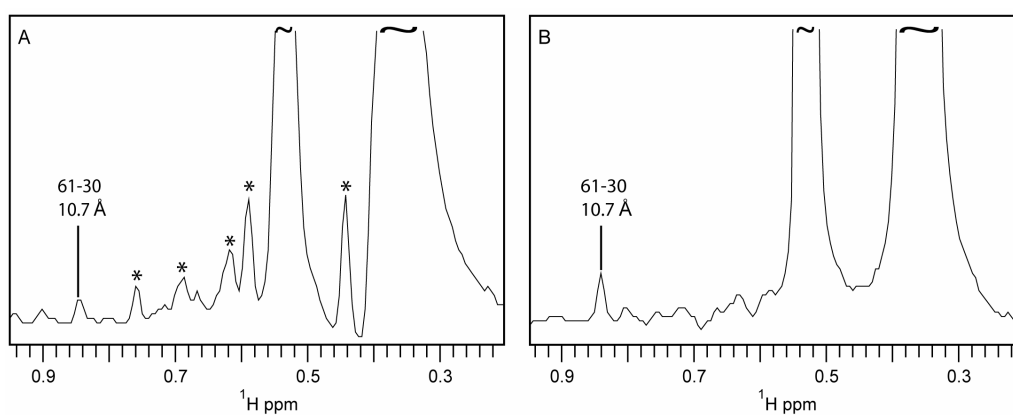


Figure 3.8 : Suppression des artéfacts par combinaison d'un marquage isotopique spécifique et de la séquence d'impulsion NOESY avec un filtre ^{13}C .

Les sections transversales F2 à la fréquence du méthyle δ_1 de Ile 61 de l'ubiquitine sont présentées. Le spectre NOESY a été enregistré sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, durant 14 heures, avec la séquence d'impulsions de la Figure 3.6, $\tau_m = 1.25$ s et $d_1 = 1.5$ s sur un échantillon à 1 mM d'ubiquitine U- ^{2}H , U- ^{12}C , $^{12}\text{C}^{13}\text{H}_3$ -Ile- δ_1 (spectre A) ainsi que sur un échantillon à 1mM d'ubiquitine U- ^{2}H , U- ^{13}C , $^{12}\text{C}^{13}\text{H}_3$ -Ile- δ_1 (spectre B). En (A) les pics représentant des artéfacts sont marqués d'un astérisque.

3.1.6 Applications à la détection de NOEs entre méthyles éloignés dans les protéines de tailles modérées.

Deux systèmes ont été étudiés : l'ubiquitine, comme protéine modèle, et la γ S-crystalline, une protéine de taille modérée. Des spectres 2D NOESY, avec un temps de mélange et de recyclage optimisés (§ 3.1.4), ont été enregistrés sur un spectromètre 800 MHz, équipé avec une sonde cryogénique, à 25°C, pendant 14 heures. Pour l'ubiquitine, un échantillon U- ^{2}H , U- ^{13}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 d'une concentration de 1 mM a été utilisé en combinaison de la séquence d'impulsion NOESY avec filtre ^{12}C , afin d'éliminer les artéfacts présents aux fréquences correspondants aux méthyles δ_1 des Isoleucines. L'analyse de la structure de la γ S-crystalline démontre que dans une sphère de 6 Å autour de chacun des méthyles Ile- δ_1 , il n'y a aucun méthyle présent avec une fréquence proton pouvant être confondue avec celles des méthyles δ_1 des Isoleucines, spécifiquement protonés. Par conséquent, l'analyse des NOEs à longue portée entre les méthyles Ile- δ_1 , a été réalisée en utilisant un échantillon à 1 mM de γ S-crystalline U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 .

Un total de dix paires de pics de corrélations interméthyles pour l'ubiquitine et onze pour la γ S-crystalline a été détecté. Tous les NOEs interméthyles observés pour la γ S-crystalline sont représentés sur la Figure 3.9.A. Les traces 1D montrant les quatre NOEs impliquant le méthyle δ_1 de l'Isoleucine 95 de la γ S-crystalline sont montrées sur la Figure 3.9.B. L'Isoleucine 169 est sujet à un échange conformationnel responsable d'un élargissement important des raies. Aucun NOE impliquant ce résidu ne peut être détecté.

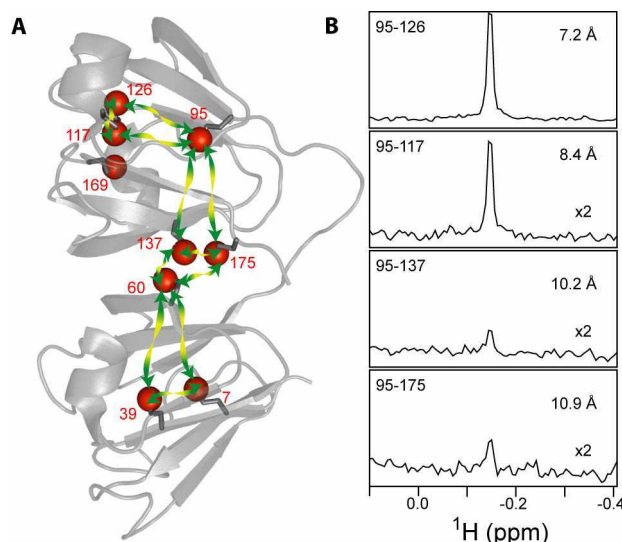


Figure 3.9 : Structure secondaire de la γ S-crystalline avec la répartition des méthyles δ_1 des Isoleucines et les NOEs interméthyles détectés pour le méthyle δ_1 de l'Ile-95.

En A) les NOEs observés sur la γ S-crystalline entre les méthyles δ_1 des Isoleucines (sphères rouges) sont indiqués par des flèches vertes. En B) les sections F1 aux fréquences des méthyles δ_1 des Isoleucines 126, 117, 137 et 175 sont représentées. Le spectre 2D-NOESY a été enregistré sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, pendant 14 heures, avec $\tau_m = 0.8$ s et $d_1 = 1.5$ s, sur un échantillon de 1 mM de γ S-crystalline U- ^{12}C , U- ^{13}C , $^{13}\text{C}^{13}\text{H}_3$ -Ile- δ_1 .

La comparaison de la distribution des distances entre méthyles δ_1 des Isoleucines, prédites à partir des structures de l'ubiquitine (Cornislec *et al*, 1998) et de la γ S-crystalline (Grishaev *et al*, 2005), avec les pics de corrélation observés dans les spectres NOESY, indique que tous les NOEs observés correspondent à des distances interméthyles comprises entre 4 Å et 12 Å, Figure 3.10. Toutes les paires de méthyles distantes par moins de 10,5 Å donnent lieu à des NOEs observables. Pour les distances interméthyles comprises entre 10,5 Å et 12,5 Å, plus de 60 % des NOEs sont encore détectés des deux côtés de la diagonale du spectre NOESY, avec un rapport moyen de signal sur bruit supérieur à 5. Il est intéressant de noter que les trois pics de corrélation non détectés entre 10,5 et 12,5 Å (en gris sur la Figure 3.10) correspondent aux NOEs impliquant le méthyle δ_1 de l'Isoleucine 44 de l'ubiquitine. Alors que toutes les chaînes latérales des Isoleucines se situent dans le cœur hydrophobe et sont caractérisées par une faible mobilité, cette chaîne latérale se situe à la surface de la protéine et est caractérisée par une dynamique importante (Lee *et al*, 1999). L'efficacité du transfert NOE est diminuée à cause de la mobilité de cette chaîne latérale, réduisant l'intensité des pics de corrélation attendue en dessous du seuil de détection.

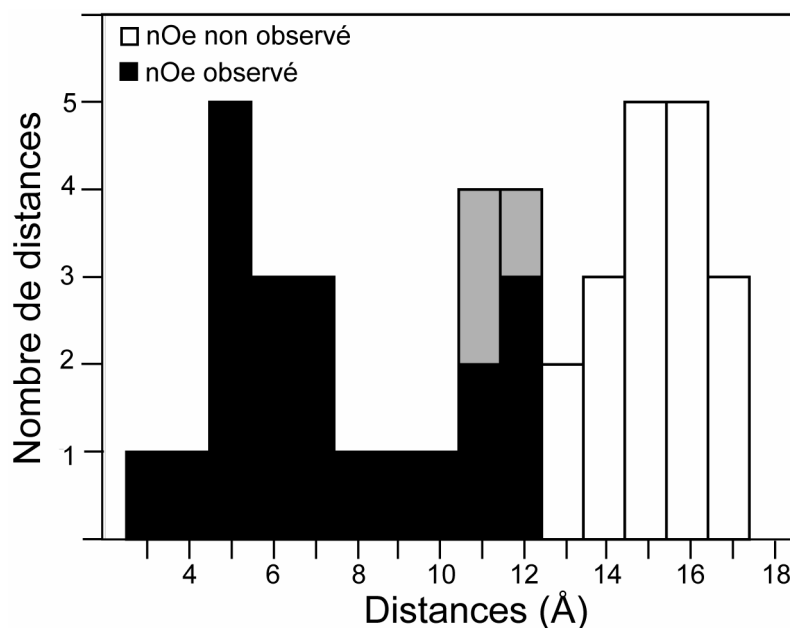


Figure 3.10 : Distribution des distances interméthyles entre méthyles Ile- δ_1 dans l'ubiquitine et la γ S-crystalline.

L'axe des abscisses correspond à la distance moyenne interprotons déterminée à partir des structures de l'ubiquitine (pdb : 1D3Z) et de la γ S-crystalline (pdb : 2A5M). Les NOEs impliquant Ile-169 ont été omis, du fait de l'échange conformationnel caractérisant ce résidu. Les NOEs détectés expérimentalement sont représentés par des barres noires. Les NOEs non détectés, impliquant le méthyle δ_1 mobile de Ile-44, sont représentés en gris.

3.2 Quantification des distances interméthyles

3.2.1 Matrice de relaxation totale

L'analyse quantitative des NOEs entre protons est utilisée, pour les études structurales dans les molécules biologiques, et rencontre trois difficultés majeures diminuant la précision de mesures des intensités des pics :

- (i) Le transfert de l'aimantation par des voies relayées (la diffusion de spins).
- (ii) La dynamique interne du système.
- (iii) Le recouvrement des pics de corrélation.

L'utilisation de la matrice de relaxation dans l'interprétation des valeurs NOEs s'avère être pratique pour tenir compte du problème de la diffusion de spins (Keepers *et al*, 1984 ; Lefèvre *et al*, 1987 ; Boelens *et al*, 1988 ; Borgias *et al*, 1990 ; Koehl *et al*, 1990). Des programmes itératifs permettant d'affiner les structures, incluant le calcul des paramètres de la matrice de relaxation, ont été utilisés pour

résoudre la structure d'oligomère d'ADN (Boelens *et al*, 1989) de même que la structure de protéine (Nilges *et al*, 1991). Dans le cas de petits peptides, où l'ensemble des résonances est bien résolu, c'est-à-dire les pics diagonaux ainsi que les pics de corrélation, la matrice de relaxation peut être calculée à partir d'un seul spectre NOESY (Mirau, 1988). Toutefois dans le cas de protéines, le recouvrement des pics diagonaux et de corrélation est une difficulté qui limite l'utilisation de la matrice de relaxation totale. Plus le réseau de spins est dense, plus la propagation d'erreurs est importante dans le calcul. Pourtant, si l'on utilise un marquage isotopique spécifique, on simplifie le réseau de spins à sept méthyles pour l'ubiquitine U-[²H], U-[¹³C], [¹²C¹H₃]-Ile- δ_1 , ce qui est une simplification considérable en comparaison avec les 700 protons présents dans l'ubiquitine protonée. Les pics diagonaux et de corrélation sur le spectre NOESY sont tous résolus. Si l'on considère une protéine globulaire, dont les chaînes latérales intervenant dans le cœur hydrophobe sont soumises à une faible dynamique, l'utilisation de la matrice de relaxation totale paraît alors être une méthode de choix pour extraire précisément des distances entre protons dans un tel système simplifié.

Dans une expérience NOESY, les volumes des NOEs entre les paires de protons peuvent être décrits par l'équation de Bloch :

$$\frac{d}{d\tau_m} V(\tau_m) = -\sigma \cdot V(\tau_m) \quad (3.21)$$

où $V(\tau_m)$ est la matrice des volumes des pics de corrélation ainsi que des pics diagonaux, extraits à partir des spectres NOESY enregistrés avec le temps de mélange τ_m et σ est la matrice de relaxation. L'intégration de l'équation de Bloch peut alors être effectuée analytiquement afin d'obtenir la matrice des volumes au temps de mélange τ_m :

$$V(\tau_m) = \exp(-\sigma \cdot \tau_m) V(0) \quad (3.22)$$

$V(\tau_m)$ et $V(0)$ sont mesurables par des expériences NOESY ; il est donc possible de déterminer σ ainsi que les distances entre les protons. L'équation 3.20 peut être réarrangée pour donner σ en fonction des termes de $V(\tau_m)$ et $V(0)$:

$$\sigma = -(\ln[V(t_m).V(0)^{-1}])/\tau_m \quad (3.23)$$

L'obtention de la matrice de relaxation σ à partir de $V(\tau_m)$ et de $V(0)$ revient à un simple problème de détermination de matrice. La matrice des distances interméthyles est alors extraite à partir des termes hors diagonaux de la matrice de relaxation σ , grâce à l'Equation 3.13.

3.2.2 Applications

Pour évaluer la précision de mon approche et afin d'analyser les effets éventuels de diffusion de spins, l'analyse de la matrice de relaxation totale a été utilisée pour extraire les distances interméthyles à partir des spectres NOESY. Etant donné la haute résolution des spectres (§ 2.4.5) et la simplicité du système de spins dans l'échantillon protoné spécifiquement sur les Ile- δ_1 , la matrice de relaxation σ peut être directement déterminée à partir des volumes des pics diagonaux et des pics de corrélation expérimentaux.

Les volumes des pics de $V(\tau_m)$ (diagonaux et de corrélation) sont mesurés sur le spectre NOESY, enregistré avec un temps de mélange optimal, pendant 14 heures. Les volumes des pics de $V(0)$ (diagonaux) sont extraits à partir d'une autre expérience NOESY, enregistrée avec un temps de mélange nul ; compte tenu du fort rapport signal sur bruit caractérisant les pics diagonaux, le temps expérimental de cette expérience NOESY est réduit au minimum (le nombre d'accumulation = 16, pour 1 heure d'acquisition). Un facteur correctif est appliqué sur les volumes de $V(0)$ afin de compenser cette différence.

Les systèmes de spins dans des échantillons de l'ubiquitine et de la γ S-crystalline protonées spécifiquement sur les méthyles Ile- δ_1 sont composés, respectivement, de sept et de neuf spins équivalents. J'ai écrit un programme nommé RELOAD ("*full-RElaxation-matrix analysis for LOn-g-rAnge methyl methyl Distance measurement*") permettant de calculer directement les distances interméthyles à partir des fichiers d'intensités des matrices NOESY (le programme est disponible sur <http://www.ibs.fr/ext/labos/LRMN/softs/welcome.htm> ; le protocole d'utilisation est détaillé en Annexe 3). Ce programme demande, comme paramètres d'entrée : deux

fichiers contenant la matrice des volumes des pics NOESY de $V(\tau_m)$ et de $V(0)$. De plus, afin d'extraire directement les distances interméthyles des termes hors diagonaux de la matrice de relaxation σ , le temps de corrélation total de la protéine, τ_C , doit être fourni comme paramètre d'entrée dans le calcul. Il peut être estimé par des expériences de mesures des paramètres de relaxation (T_1 , T_2) standard.

3.2.2.1 Résultats

Dans le cas simplifié d'une rotation rapide des groupes méthyles par rapport au mouvement globale de la protéine, la distance moyenne entre les méthyles i et j peut être estimée directement à partir des éléments σ_{ij} de la matrice de relaxation.

La comparaison des distances expérimentales avec des distances extraites à partir des structures de l'ubiquitine et de la γ S-crystalline montre une très bonne corrélation (Figure 3.11). Le coefficient de corrélation de Pearson R_p est de 0,96. Le plus grand écart se produit entre les résidus 3 et 13 de l'ubiquitine. Une étude indépendante a montré que les chaînes latérales de ces résidus peuplent différents rotamères (Chou *et al*, 2003) et ne peuvent donc pas être représentées par une structure unique.

Les distances supérieures à 5 Å peuvent être déterminées avec une précision moyenne de 9 %, sans l'utilisation d'une quelconque référence interne de distance. En dépit de la simplicité du modèle dynamique, les distances dérivées des NOEs demeurent précises. L'incertitude est réduite à 4 %, si l'on ne considère que les paires de méthyles appartenant à des chaînes latérales, ne peuplant qu'un seul conformère et pour laquelle une structure haute-résolution est disponible. En raison de la dépendance en $\langle r^{-3} \rangle^2$ du NOE (Equation 3.13), les distances extraites sont relativement peu sensibles à l'erreur sur le τ_C . Dans la limite d'une diffusion rotationnelle globale lente, une erreur de 30 %, sur le temps de corrélation total τ_C , aura comme conséquence une erreur de 5 % sur les distances expérimentalement déterminées. Ceci nous permet de justifier l'utilisation du modèle de mouvement global isotrope pour les protéines globulaires. Dans le cas de distances interméthyles courtes, l'approximation ($d_{HH}^{\text{intra-méthyle}} \ll d_{HH}^{\text{inter-méthyle}}$) n'est plus valide et le paramètre

d'ordre S^2 ne tend pas vers 1 ; on a donc une surestimation des distances expérimentales.

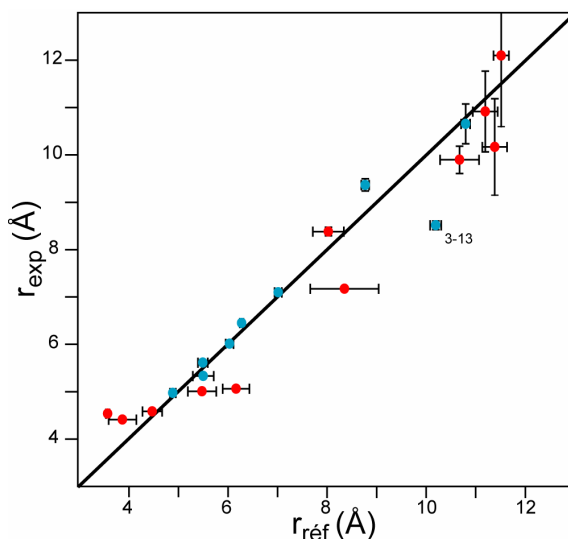


Figure 3.11 : Comparaison des distances expérimentales calculées à partir des volumes des pics NOEs avec les distances références à partir des structures de l'ubiquitine (cercles bleus, PDB : 1D3Z) et de la γ S-crystalline (cercles rouges, PDB : 2A5M).

La déviation standard des distances références ($r_{\text{réf}}$) à partir de l'ensemble de structures est représentée par une barre d'erreur horizontale. Pour calculer la déviation standard pour les distances expérimentales (r_{exp}), un ensemble de mille volumes a été généré en ajoutant aléatoirement du bruit, correspondant aux bruits expérimentaux de l'expérience NOESY (représentée par une barre d'erreur verticale).

3.2.2.2 Importance de la faible densité de protons dans l'extraction des distances interméthyles

L'analyse de la matrice de relaxation totale permet également d'évaluer la contribution de la diffusion de spins sur les pics de corrélation NOEs observés.

Pour cela, la matrice de relaxation, déterminée expérimentalement pour les méthyles Ile- δ_1 de la γ S-crystalline, a été utilisée pour simuler les volumes des pics de corrélations NOEs. Pour évaluer la contribution de la diffusion de spins des NOEs entre les méthyles i et j , les éléments σ_{ij} et σ_{ji} correspondants, dans la matrice de relaxation totale, ont été arbitrairement fixés à une valeur nulle avant le calcul des volumes des NOEs. Ceci a pour effet de modéliser uniquement la partie des NOEs observés, due à la diffusion de spins.

Bien que des temps de mélange relativement longs aient été employés (0,8 s pour un temps de corrélation globale de la molécule de 16,7 ns), la diffusion de spins

représente seulement 12 % des volumes des pics de corrélation, pour des méthyles séparés par plus de 10 Å dans la γ S-crystalline U- ^{2}H , U- ^{12}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 . Ceci s'explique par le fait que la densité de protons, dans cette protéine spécifiquement marquée, est réduite à 2 % (Figure 3.12).

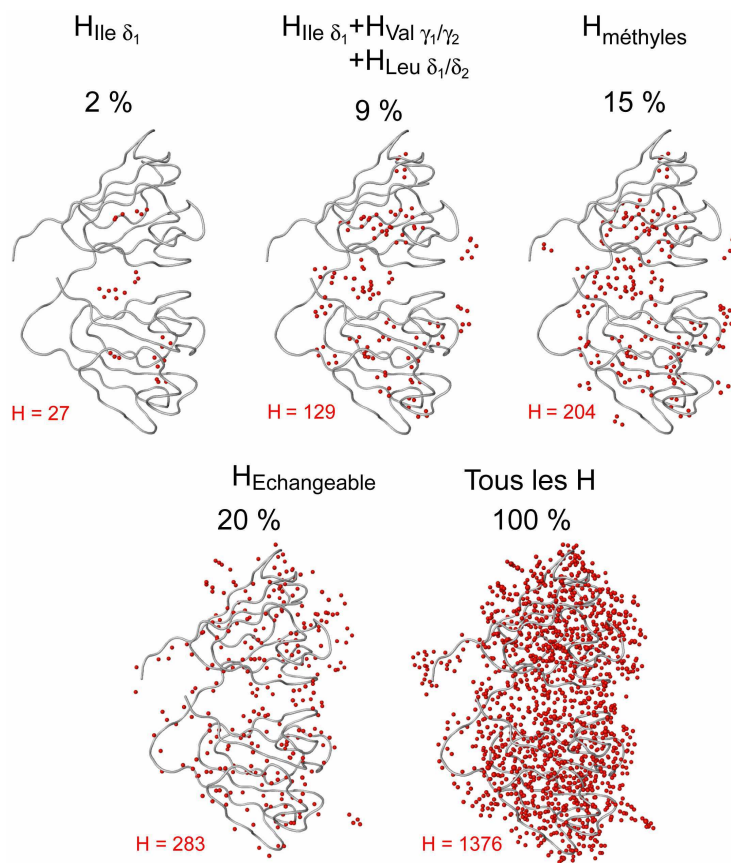


Figure 3.12 : Densité de proton dans la γ S-crystalline en fonction des sites spécifiquement protonés.

Les protons sont représentés par des sphères rouges pour un marquage : A) uniquement sur les méthyles δ_1 des Isoleucines ; B) les méthyles des Valines, des Leucines et les méthyles δ_1 des Isoleucines ; C) l'ensemble des méthyles ; D) tous les protons échangeables et E) tous les protons. Le nombre d'atomes d'hydrogène est noté en rouge et le pourcentage correspond à la densité de protons.

Afin d'établir l'importance de la faible densité de protons dans la précision de la quantification des distances interméthyles, j'ai considéré deux cas supplémentaires. Dans le premier, on considère qu'en plus des méthyles δ_1 des Isoleucines, les méthyles des acides aminés des Leucines et des Valines sont protonés, la densité de protons est alors de 9 %. Dans le second cas, tous les groupements méthyles sont considérés protonés dans la γ S-crystalline, la densité de protons est alors de 15%. La simulation de la matrice de relaxation totale indique que la contribution de la

diffusion de spins augmente ainsi respectivement à 76% et à 92% (Figure 3.13) pour des pics de corrélation NOEs observés entre des méthyles Ile- δ_1 éloignés par plus de 10 Å. Dans de tels systèmes, compte tenu de la propagation importante des erreurs lors de la détermination de la matrice σ , il devient alors difficile de déterminer précisément ces distances.

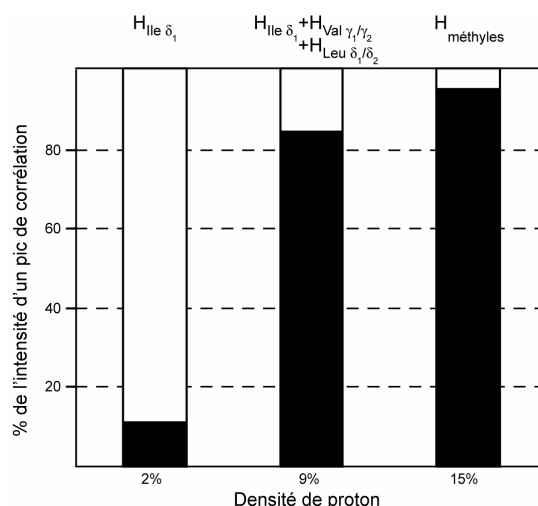


Figure 3.13 : Evaluation de la contribution de la diffusion de spins dans un pic de corrélations NOEs interméthyles.

La contribution de la diffusion de spins (en noire) des pics de corrélations NOEs observés entre les méthyles Ile- δ_1 séparés par plus de 10 Å dans la γ S-crystalline. Les densités de protons correspondent aux valeurs de la Figure 3.11.

En comparaison, dans une protéine perdeutérée où l'on a réintroduit des protons par échange avec le solvant, la densité de protons est de l'ordre de 20 %. Cette densité de protons est cinq à dix fois plus importante que pour la protonation spécifique des méthyles δ_1 des Isoleucines. Ceci explique le fait que les distances H_N-H_N ne peuvent être quantifiées précisément, même dans les protéines perdeutérées (Koharudin *et al*, 2003), et justifie notre approche de la protonation spécifique de quelques sites dans une protéine perdeutérée.

3.3 Conclusions et Perspectives

La méthode proposée ici pour la mesure des contraintes précises à longue portée entre des méthyles se fonde sur la réduction de l'effet de diffusion de spins par une protonation spécifique de quelques sites dans la protéine. J'ai montré que des NOEs entre méthyles distants jusqu'à 12 Å peuvent être détectés et donnent lieu à des

contraintes de distances précises. L'efficacité de la méthode a été démontrée sur une petite protéine modèle, l'ubiquitine, aussi bien que sur une protéine de taille modérée la γ S-crystalline, une protéine de masse moléculaire de 19 kDa. Cette approche permet d'augmenter la précision sur les modèles structuraux de protéines allongées, modulaires ou des complexes macromoléculaires étudiés par RMN.

Les bons résultats, obtenus avec cette stratégie de protonation spécifique, justifient le développement d'autres sondes que les méthyles δ_1 des Isoleucines et donc le marquage spécifique des autres acides aminés contenant des groupements méthyles. Par exemple, la protonation spécifique de seulement l'un des deux méthyles des Leucines et des Valines peut être réalisée, en utilisant α -cétoisovalérate comme unique source de protons, lors de la surexpression d'une protéine dans *E. coli* (§ 2.5). Aux vues des propriétés dynamiques favorables du méthyle de l'Alanine et de l'attribution, à priori plus aisée, la mise au point d'un protocole de protonation spécifique du méthyle de l'Alanine (§ 2.6) est la plus intéressante parmi les acides aminés restants. Toutefois, le marquage spécifique de la Thréonine et de la Méthionine peut être envisagé, par la suite, pour disposer de l'ensemble des sondes permettant, ainsi, d'adapter cette stratégie à un plus grand nombre d'études de protéines.

3.4 Mesure précise de distance entre méthyles séparés dans des protéines spécifiquement protonées

High-Accuracy Distance Measurement between Remote Methyls in
Specifically Protonated ProteinsRemy Sounier,[§] Laurence Blanchard,[§] Zhengrong Wu,[‡] and Jérôme Boisbouvier^{*§}IBS, Institut de Biologie Structurale Jean-Pierre Ebel, 41 rue Jules Horowitz, F-38027 Grenoble, CEA, CNRS,
Université Joseph Fourier, and Biochemistry Department, The Ohio State University, Columbus, Ohio 43210

Received October 10, 2006; E-mail: jerome.boisbouvier@ibs.fr

NMR structure determination of biological macromolecules in solution has relied on the use of semiquantitative NOE distance restraints and dihedral angle restraints. The introduction of orientational information derived from residual dipolar couplings has allowed a significant improvement in the precision of structural models of globular proteins.¹ Nevertheless the lack of long-range distances remains a major limiting factor for the accuracy of structure determination of modular and elongated biomolecules as well as for biomolecular complexes. The use of perdeuterated samples and improvement in spectrometer hardware have allowed in the last years the measurement of distance restraints between protons separated by up to 8 Å.² Detection of ¹H–¹H NOEs between more remote protons has been hampered by the small magnitude of such effects due to the r^{-6} dependence. Here we demonstrate that observation of ¹H–¹H NOEs over distances of up to 12 Å is feasible using a highly deuterated protein in which protons have been selectively incorporated into isoleucine δ_1 -methyl sites. By reducing the proton density to 2–3% compared to fully protonated samples, the measurement of accurate long-range distance restraints becomes possible. The detection of NOEs between such remote protons benefits from three major effects. First, because of the proton multiplicity, the choice of methyls for selective protonation enhances signal transferred between two sites by a factor of 9, in the limit of long intermethyl distances. Second, in highly deuterated samples, methyl protons are characterized by favorable transverse and longitudinal relaxation rates providing high resolution and signal-to-noise ratio of the NMR spectra.³ And last, by reducing the proton density, spin diffusion is considerably reduced, a prerequisite for accurate distance measurements. Long mixing times can then be used to transfer substantial amounts of magnetization through very weak NOEs.

The efficiency of the method for observing the long distance NOE is demonstrated on two proteins for which high resolution solution structures are available: ubiquitin and γ S-crystallin, a 19 kDa protein comprising two domains.⁴ Highly deuterated samples with selective incorporation of protonated methyls on Ile- δ_1 methyl sites of ubiquitin and γ S-crystallin, have been prepared using 2-keto-3,3- d_2 -butyrate as the unique proton source during protein overexpression in *E. coli* as previously proposed by Gardner and Kay.⁵ A protocol has been implemented to incorporate selectively protonated methyls on Ile- δ_1 sites with an efficiency larger than 98% while the average deuteration level at all other sites in the overexpressed proteins remains larger than 98% (Supporting Information S1). As all the ¹H resonances of Ile- δ_1 methyls (7 for ubiquitin and 9 for γ S-crystallin) are well resolved, all the NOESY spectra have been recorded using ¹H edited 2D-experiments (Supporting Information S2).

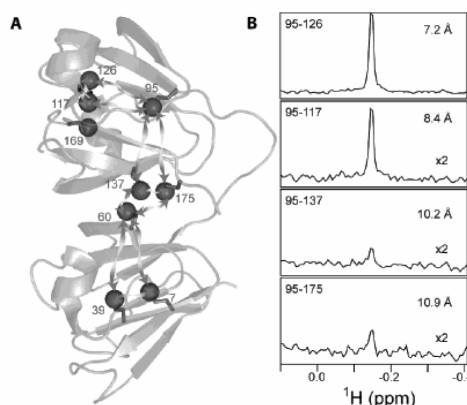


Figure 1. (A) Ile- δ_1 intermethyl NOEs observed in γ S-crystallin are indicated by green arrows. Ile-169 does not give rise to detectable intermethyl NOEs owing to extensive line broadening. (B) Intermethyl NOEs detected for Ile-95- δ_1 . F_1 cross-sections at δ_1 methyl frequency of Ile 126, 117, 175, and 137 are displayed. The 800 MHz NOESY spectrum has been acquired in 14 h on a 1 mM sample of U-[²H]₃[CH₃]-Ile- δ_1 γ S-crystallin.

The optimal NOE mixing time for transfer between protonated Ile- δ_1 methyls was determined from the build-up of cross-peak intensities in a series of short 2D NOESY spectra. Maximum cross-peak intensities were obtained for mixing times of 1.25 and 0.8 s for ubiquitin and γ S-crystallin, respectively (data not shown). 2D NOESY spectra with optimized mixing times have been acquired at 25 °C in 14 h for each protein on a 800 MHz spectrometer equipped with a cryogenic probe. A total of ten pairs of intermethyl cross-peaks for ubiquitin and eleven for γ S-crystallin has been detected. All the intermethyl NOEs observed for γ S-crystallin are displayed in Figure 1A. 1D traces showing the four observed NOEs involving methyl δ_1 of Ile-95 are displayed in Figure 1B. Comparison of the distance distribution of Ile- δ_1 methyls predicted from the ubiquitin and γ S-crystallin structures⁴ with cross-peaks detected in the NOESY spectra, revealed that all observed NOEs correspond to intermethyl distances ranging from 4 to 12 Å (Figure 2). All methyl pairs distant by less than 10.5 Å give rise to observable NOEs. For intermethyl distances ranging from 10.5 to 12.5 Å, more than 60% of the expected NOEs are still detected on both sides of the NOESY diagonal with an average signal-to-noise ratio of 5:1.

To evaluate the accuracy of the method, a full-relaxation-matrix analysis was employed to extract distance restraints from NOESY spectra.⁶ Given the high resolution of the spectra and the simplicity of the spin system in Ile- δ_1 methyl specific protonated samples, the relaxation matrix σ can be directly determined from diagonal- and cross-peak volumes in NOESY spectra (Supporting Information S3). For the simplified case of fast rotating methyl groups in an isotropically tumbling macromolecule,⁷ the average distance be-

[§] Institut de Biologie Structurale[‡] The Ohio State University.

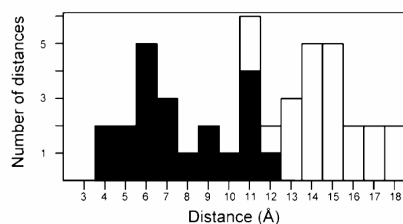


Figure 2. Distribution of Ile- δ_1 intermethyl distances versus average interproton distance calculated from structures of ubiquitin and γ S-crystallin (except for Ile-169).⁴ The experimentally detected NOEs are displayed by black bars.

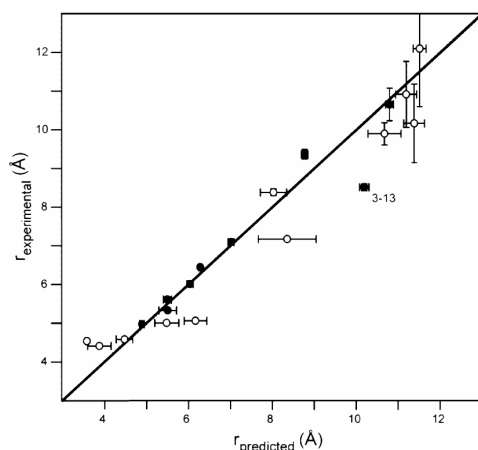


Figure 3. Comparison of the experimental distances calculated from NOE peak volumes using full-relaxation-matrix analysis with distances predicted from ubiquitin (filled circles) and γ S-crystallin (open circles) structures.⁴ The standard deviation of the distances predicted from the structure ensemble are represented by a horizontal error bar, with $r = (\langle r^{-3} \rangle)^{-1/3}$. A thousand sets of noise-corrupted NOESY peak volumes have been generated (using experimental noise) and used to extract the standard deviation on $r_{\text{experimental}}$ (displayed as vertical error bars). The largest discrepancy occurs between residues 3 and 13 of ubiquitin, which have been shown to populate different side-chain rotamers.¹¹

tween methyls i and j can be estimated directly from the relaxation matrix elements σ_{ij} . Comparison of experimental distances with distances extracted from the ubiquitin and γ S-crystallin structures shows a very good agreement (Figure 3), with a Pearson's correlation coefficient $R_P = 0.96$. The distances exceeding 5 Å can be determined with an average accuracy of 9% without the use of any internal distance reference. Despite the simplicity of the dynamic model, the NOE-derived distances remain accurate. Because of the $\langle r^{-3} \rangle^2$ dependence of the NOE, the extracted distances are relatively insensitive to error on τ_C . In the slow tumbling limit, an error of 30% in τ_C will result in an error of 5% in the experimentally determined distances.

Full-relaxation-matrix analysis also allows us to evaluate the contribution of spin diffusion to the NOE cross-peaks observed between Ile- δ_1 methyls. Although relatively long mixing times were used (0.8 s for a 17 ns tumbling molecule), spin diffusion only accounts for 12% of the cross-peak volumes, for methyls separated by more than 10 Å in U-[^2H], [CH_3]-Ile- δ_1 - γ S-crystallin (proton density 2%). Protonation of the methyls of Val and Leu residues (proton density 9%) or of all methyl sites in γ S-crystallin (proton

density 15%) would increase the contribution of spin diffusion to respectively 76% and 92% of the NOEs observed between methyls distant by more than 10 Å (Supporting Information S3), making accurate distance determination difficult. The method proposed here for measurement of accurate long-distance restraints between methyls relies on the reduction of spin diffusion effects by the protonation of only few specific sites in a protein. This specific protonation strategy can be generalized to other methyl-containing residues. For example, specific protonation of only one of the two methyls of Leu and Val residues can be achieved using 2-keto-3-methyl- d_3 -3- d_1 -butyrate as a precursor during *E. coli* overexpression.⁸ The selective protonation of other methyl-containing residues can be obtained alternatively using a cell-free approach.⁹ The efficiency of the method was demonstrated on the small model protein ubiquitin as well as on the 19 kDa protein γ S-crystallin. The proposed approach can be transposed to larger proteins using ^{13}C , ^1H -methyl specific labeling. Simulation indicates that in a 45 ns tumbling methyl-specific protonated protein, NOE cross-peaks will still be detectable between methyls separated by up to 10 Å using the recently proposed 4D methyl-TROSY-NOESY experiment.¹⁰ In such large proteins, spin diffusion will contribute for less than 10% of the NOE cross-peak volumes, provided that the total proton density remains lower than 2–3%. The possibility to measure accurate translational restraints between distant methyls offers new opportunities for the study of large proteins and protein complexes by solution NMR.

Acknowledgment. This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, and Human Frontier Science Program Organization. The authors thank Drs. Simorre, Schanda, Devany, Bersch, Bax, and Brutscher for helpful discussions, C. Jaroniec for acquisition of the initial NOESY data set on γ S-crystallin, and I. Ayala and K. Treche for help in ubiquitin preparation.

Supporting Information Available: Protocol for specific protonation of Ile- δ_1 methyls, experimental details for detection of long-range methyl-methyl NOEs, and equations used to calculate experimental distances from NOESY spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111–1114.
- (2) (a) Mueller, G. A.; Choy, W. Y.; Daiwen, Y.; Forman-Kay, J. D.; Venters, R. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *300*, 197–212. (b) Wu, Z.; Bax, A. *J. Am. Chem. Soc.* **2002**, *124*, 9672–9673. (c) Koharudin, L. M. I.; Bonvin, A. M. J. J.; Kaptein, R.; Boelens, R. *J. Magn. Reson.* **2003**, *163*, 228–235. (d) Meier, S.; Häussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, *125*, 44–45.
- (3) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420–10428.
- (4) (a) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836–6837. (b) Grishaev, A.; Wu, J.; Trewella, J.; Bax, A. *J. Am. Chem. Soc.* **2005**, *127*, 16621–16628.
- (5) Gardner, K. H.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 7599–7600.
- (6) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95–107.
- (7) Tropp, J. *J. Chem. Phys.* **1980**, *72*, 6035–6043.
- (8) Tugarinov, V.; Kay, L. E. *J. Biomol. NMR* **2004**, *28*, 165–172.
- (9) (a) Kigawa, T.; Muto, Y.; Yokoyama, S. *J. Biomol. NMR* **1995**, *6*, 129–134. (b) Torizawa, T.; Shimizu, M.; Taoka, M.; Miyano, H.; Kainosho, M. *J. Biomol. NMR* **2004**, *30*, 311–325. (c) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Ono, A. M.; Güntert, P. *Nature* **2006**, *440*, 52–57.
- (10) Tugarinov, V.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2005**, *127*, 8214–8225. (b) Frueh, D.; Vosburg, D. A.; Walsh, C. T.; Wagner, G. *J. Biomol. NMR* **2006**, *34*, 3–172.
- (11) Chou, J. J.; Case, D. A.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 8959–8966.

JA067260M

Chapitre

4

La Détection de NOEs entre les méthyles dans les assemblages supramoléculaires

4.1 Introduction

Depuis la résolution de la première structure de protéine par RMN en 1985 (Williamson *et al*, 1985), de nombreux progrès ont été faits pour transposer ces techniques RMN standard à des systèmes dont la taille est de plus en plus importante. La détermination de structures par RMN en solution a généralement été limitée aux protéines de faibles masses moléculaires (< 30kDa), en raison de l'augmentation des vitesses de relaxation transversale des signaux RMN et de l'élargissement des raies qui en dérivent par rapport à la taille des macromolécules. En combinant l'utilisation d'aimants modernes opérant à des champs magnétiques élevés, les innovations en marquages isotopiques, telles que la perdeutération (Sattler *et al*, 1996) et la protonation spécifique de certains sites (Gardner *et al*, 1998), et les avancées spectroscopiques, le développement d'expériences adaptées pour les macromolécules, telles que TROSY (Pervushin *et al*, 1997) ou CRINEPT (Riek *et al*, 1999), ont permis à la RMN en solution de s'adresser à des complexes protéiques de plus en plus grands. Depuis les années 2000, l'étude de plusieurs complexes de plus de 50 kDa a été réalisée par différents groupes RMN (Fiaux *et al*, 2002 ; Rüdiger *et al*, 2002 ; Christodoulou *et al*, 2004 ; Sprangers *et al*, 2005 ; 2007). A ce jour, la plus grosse protéine monomérique, dont l'attribution et la structure ont été établies par RMN, est la Malate Synthase G de masse moléculaire de 82 kDa (Tugarinov *et al*, 2002 ; 2003a ; 2003b ; 2005a). Toutefois, sans avoir obtenu la structure tridimensionnelle par RMN en solution, des études d'interactions ont été réalisées sur des systèmes encore plus grands, tels que le complexe GroEL-GroES 900 kDa (Fiaux *et al*, 2002 ; 2004) ou le complexe de plus de 1 MDa entre la particule du cœur du protéasome 20S et son activateur, la protéine 11S (Sprangers *et al*, 2007). Dans des assemblages de 360, 470 et 670 kDa, des NOEs entre protons proches dans l'espace ont pu être observés (Horst *et al*, 2006 ; Sprangers *et al*, 2007). Compte tenu des propriétés favorables de relaxation des méthyles dans les protéines spécifiquement protonées (§ 2.3) et des résultats obtenus pour des protéines de tailles modérées (§ 3.1.6), l'utilisation de ce type de marquage apparaît comme une stratégie optimale pour détecter des NOEs à longue portée dans les systèmes de tailles moléculaires élevées. Dans ce chapitre, je présenterai la transposition de la stratégie exposée au chapitre précédent pour l'étude des complexes macromoléculaires de grandes tailles.

Pour cette étude, j'ai choisi d'utiliser le marquage sélectif des méthyles δ_1 de l'Isoleucine (§ 2.4.6.2) sur plusieurs protéines dont la taille est comprise entre 21 et 468 kDa. J'étudierai les possibilités et les limites de la détection de NOEs interméthyles dans de tels assemblages moléculaires.

4.2 Les aspects expérimentaux

4.2.1 Expérience NOESY éditée ^{13}C

Tout d'abord, l'étude de gros complexes protéiques oblige l'utilisation d'un marquage au ^{13}C sur le carbone du méthyle. L'intérêt que présente ce type de marquage spécifique est l'amélioration de la résolution des spectres RMN en distribuant les signaux RMN dans plusieurs dimensions. Si l'on s'intéresse à un échantillon perdeutééré et spécifiquement protoné sur les méthyles δ_1 des Isoleucines, seul le carbone δ_1 est marqué isotopiquement ^{13}C . A partir de ces types d'échantillons, on peut enregistrer des expériences NOESY multidimensionnelles, c'est-à-dire à trois ou quatre dimensions.

Toutefois, l'édition d'une dimension supplémentaire dans une séquence d'impulsion entraîne une perte d'un facteur $\sqrt{2}$, si on prend le soin d'optimiser le temps d'édition maximale dans la dimension indirecte à une valeur proche du temps de relaxation transversale T_2 . Le signal, enregistré dans une expérience NOESY 4D par rapport à une expérience NOESY 2D, est donc diminué d'un facteur deux puisqu'il y a deux dimensions supplémentaires. Cependant, cette perte est compensée par une durée d'enregistrement nécessairement plus longue, requise pour l'acquisition d'une expérience 4D. En effet, 12 heures, environ, sont nécessaires pour enregistrer les expériences NOESY 2D, préalablement présentées (§ 3.1.6), alors qu'il faut 48 à 72 heures pour acquérir une expérience 4D HMQC-NOESY-HMQC avec une résolution suffisante.

L'expérience 4D HMQC-NOESY-HMQC, développée par Tugarinov *et coll.* (Tugarinov *et al*, 2005b), a été implémentée au laboratoire afin de réaliser les expériences NOESY sur de grands systèmes, pour détecter des NOEs entre paires de méthyles (Figure 4.1).

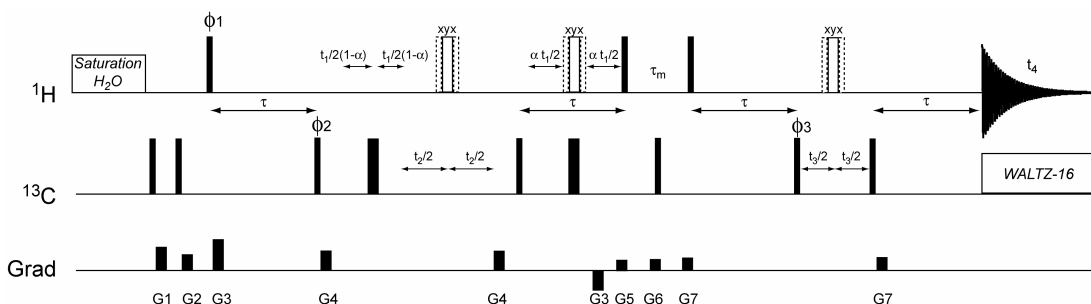


Figure 4.1 : Schéma d'impulsion de l'expérience HMQC-NOESY-HMQC.

Les impulsions rectangulaires fines et larges sont respectivement appliquées avec des angles d'impulsion de 90° et de 180° . À moins que cela ne soit indiqué, les impulsions sont appliquées le long de l'axe X. La fréquence ^1H est placée au centre de la région méthyle des Ile- δ_1 et la fréquence ^{13}C est placée au centre de la région C_{δ_1} des Isoleucines. La durée et la puissance des gradients (en unités ms, G/cm) sont : G1 = (0.5 ; 26), G2 = (0.5 ; 18.2), G3 = (0.2 ; 48.2), G4 = (1 ; 22), G5 = (0.5 ; 16), G6 = (0.5 ; 14), G7 = (0.5 ; 6). Le cycle de phase est : $\phi_1 = 4(x)$; $\phi_2 = 2(x, -x)$; $\phi_3 = 2(x) ; 2(-x)$; rec = x, -x, -x, x. La détection en quadrature en t_1 , t_2 et t_3 est achevée en incrémentant respectivement ϕ_1 , ϕ_2 et ϕ_3 de la manière usuelle du States-TPPI. Les délais de recouvrement des gradients ont été placés à $\Delta = 200 \mu\text{s}$, le délai de transfert $\tau = 4\text{ms}$ et $\alpha = \tau / t_1^{\text{max}}$.

4.2.2 Quantification des NOEs

L'enregistrement d'une expérience 4D NOESY ne permet pas d'utiliser l'approche proposée précédemment (§ 3.2.2) pour déterminer les volumes des pics diagonaux $V(0)$. En effet, l'enregistrement de l'expérience de référence pour mesurer les volumes des pics diagonaux représente une perte de temps non négligeable du point de vue du spectromètre.

J'ai développé une solution alternative pour obtenir le rapport entre les volumes des pics diagonaux, $V(\tau_m)/V(0)$. Elle consiste à enregistrer deux expériences 2D NOESY, basées sur la séquence d'impulsion 4D HMQC-NOESY-HMQC, où seule la fréquence du carbone du premier bloc HMQC (t_2) est éditée dans une dimension indirecte. La première expérience est enregistrée avec un temps de mélange égal à zéro et la seconde avec τ_m égal à τ_m^0 , utilisé pour l'expérience NOESY 4D. Ce rapport, obtenu en quelques heures grâce à ces deux expériences 2D, nous permet de recalculer la matrice $V(0)$ correspondant à une expérience 4D, sans temps de mélange, à partir de la matrice des volumes des pic diagonaux $V(\tau_m)$, extraite de l'expérience 4D enregistrée avec un temps de mélange τ_m . Ces deux matrices peuvent ensuite être utilisées pour déterminer la matrice des distances interméthyles à partir du programme RELOAD (§ 3.2.2).

4.3 Détection de NOEs à longue portée dans la malate synthase G (MSG), une protéine globulaire de 82 kDa.

La Malate Synthase G est une protéine globulaire qui comporte 44 Isoleucines sur 782 acides aminés. Un échantillon perdeutééré et spécifiquement protoné sur les méthyles δ_1 des Isoleucines a été produit comme décrit au § 2.4.6.2.

Le temps de relaxation longitudinale globale $^1\text{H Ile-}\delta_1$ de la MSG $\text{U-}[^2\text{H}]$, $\text{U-}[^{12}\text{C}]$, $[^{13}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$ a été déterminé à partir d'une expérience d'inversion-récupération. Le délai de récupération optimale d_1 est déterminé de la même manière que pour des protéines de tailles modérées (Equation 3.17). La valeur de la relaxation longitudinale, déterminée à 800 MHz pour cet échantillon à 37°C, est de 1,3 s ; le délai de recyclage optimal a été calculé de l'ordre de 1,6 s.

Des spectres 2D NOESY avec différents temps de mélange ont été enregistrés à 37°C sur un spectromètre 800 MHz, équipé avec une sonde cryogénique, afin de déterminer le temps de mélange optimal ($\tau_m^0 = 1,3$ s). On voit que ce dernier est proche du temps de relaxation longitudinale, ce qui a également pu être observé pour des protéines de tailles modérées (§ 3.1.4.3).

A partir de l'optimisation des différents paramètres de l'expérience NOESY, des spectres 4D HMQC-NOESY-HMQC ont été enregistrés à 37°C pendant 62 heures sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec un échantillon à 0,9 mM de MSG $\text{U-}[^2\text{H}]$, $\text{U-}[^{12}\text{C}]$, $[^{13}\text{C}^1\text{H}_3]\text{-Ile-}\delta_1$. A partir de cette expérience et en utilisant la procédure décrite au § 4.2.2, la quantification des distances interméthyles montre que des NOEs sont toujours détectables entre des protons séparés de plus de 10 Å. Une illustration de pic de corrélation, détecté entre les méthyles δ_1 des résidus 337 et 409, dans les plans de l'expérience 4D NOESY, est présentée sur la Figure 4.2 ; la distance entre cette paire de méthyles a été déterminée à 11 ± 1 Å, à l'aide du programme RELOAD (Annexe 3).

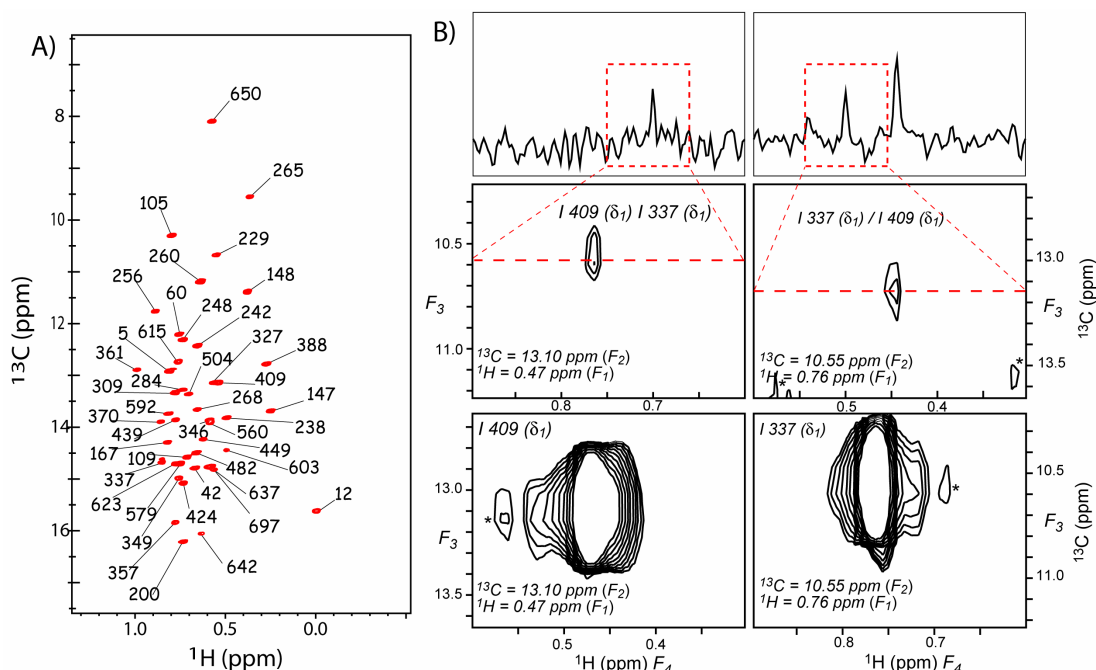


Figure 4.2 : A) Spectres 2D de corrélation HMQC ^1H - ^{13}C de la MSG. B) Plans de l'expérience 4D NOESY des pics diagonaux de I337 et I409 ainsi que les pics de corrélation correspondants.

Les deux expériences ont été enregistrées à 37°C sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec un échantillon concentré à 0,9mM de U-[^2H], U-[^{12}C], [^{13}C - $^1\text{H}_3$]-Ile- δ_1 dans 99,9 % de D_2O , 25 mM NaPO_4 à pH = 7,4 (corrigé), 20 mM MgCl_2 et 5 mM DTT ainsi que 0,05% NaN_3 . En A), l'attribution des pics de corrélation est notée sur le spectre HMQC. En B), les sélections, à droite, correspondent aux plans $F_1(^1\text{H})$ - $F_2(^{13}\text{C})$ de I337 et, à gauche, aux plans $F_1(^1\text{H})$ - $F_2(^{13}\text{C})$ de I409. Les valeurs des paramètres d'acquisition et les délais d'incrémentation de l'expérience 4D NOESY sont : $t_{1\text{max}} = 29.4 \text{ ms}$, $t_{2\text{max}} = 22 \text{ ms}$, $t_{3\text{max}} = 15.8 \text{ ms}$ et 80 ms de temps d'acquisition, qui a été utilisé dans la dimension directe ^1H . Les largeurs spectrales sont respectivement : $F_1(^1\text{H}) = 0.43 \text{ ppm}$, $F_2(^{13}\text{C}) = 4 \text{ ppm}$, $F_3(^{13}\text{C}) = 4 \text{ ppm}$ et $F_4(^1\text{H}) = 1 \text{ ppm}$. Les pics annotés d'une étoile correspondent soit à un épaulement du pic diagonal, soit à des pics se trouvant dans d'autres plans. La distance au sein de la paire de méthyles a été déterminée à $11 \pm 1 \text{ \AA}$.

La structure de la MSG a été déterminée par cristallographie de diffraction des rayons X (PDB : 1D8C ; Howard *et al*, 2000). Nous pouvons donc comparer nos données expérimentales avec cette structure. Ainsi, tous les pics de corrélation, observés dans les spectres de l'expérience 4D HMQC-NOESY-HMQC, correspondent à des distances interméthyles comprises entre 3 Å et 10,5 Å (Figure 4.3). Tous les NOEs sont observés entre paires de méthyles distantes de moins de 9 Å. Les distances interméthyles comprises entre 8,5 Å et 10,5 Å, de l'ordre de 50 % des NOEs, ont été détectés dans les plans 4D NOESY, avec un rapport moyen de signal sur bruit de 3/1.

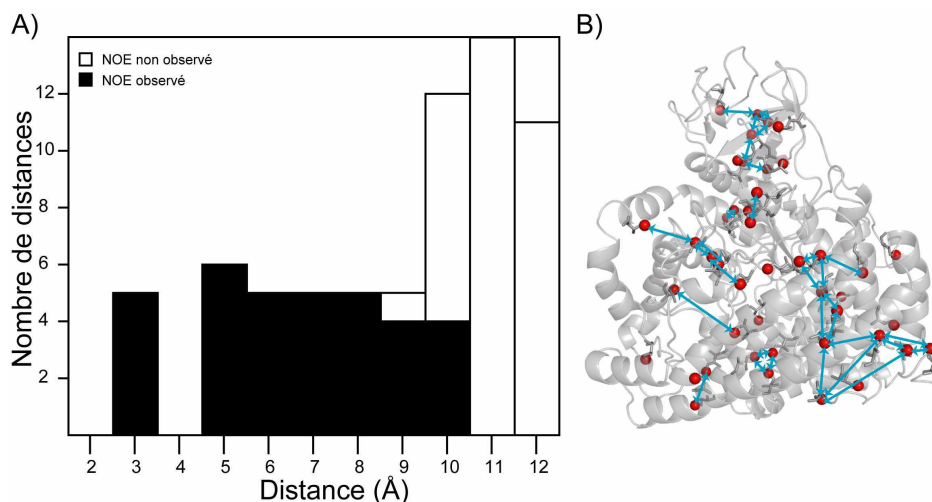


Figure 4.3 : Distribution des distances interméthyles des Ile- δ_1 dans MSG.

A) L'axe des abscisses correspond à la distance moyenne interprotons déterminée à partir des structures de MSG (1D8C). Les NOEs détectés expérimentalement sont représentés par des barres noires. B) Structure secondaire de MSG avec la répartition des Isoleucines (sphères rouges) et les NOEs interméthyles détectés (flèches bleues).

4.4 Détection de NOEs interméthyles dans les assemblages moléculaires de grandes tailles : étude théorique des limites et des possibilités.

4.4.1 La relaxation longitudinale T_1

Comme nous l'avons présenté dans le chapitre précédent (§ 3.1.4), l'optimisation des expériences NOESY consiste à prendre en compte le temps de relaxation longitudinale. Les interactions principales, pour la relaxation longitudinale des spins ^1H dans des groupements méthyles $^{13}\text{CH}_3$, sont de nature dipolaire entre les spins intraméthyles ^1H - ^1H et ^1H - ^{13}C (Werbelow *et al*, 1973 ; 1977 ; Kay *et al*, 1992). La contribution de l'anisotropie du déplacement chimique des protons des méthyles peut être négligée compte tenu de la faible amplitude d'anisotropie de déplacement chimique (CSA) de ces protons ($\Delta\sigma \approx 1$ ppm ; Tugarinov *et al*, 2004b). La relaxation longitudinale ^1H du méthyle provient alors d'interactions avec les différents atomes dans la protéine ; on peut l'écrire sous la forme :

$$R_1(^1\text{H}) = R_1^{(\text{intra-méthyle})} + R_1^{HH(\text{inter-méthyle})} + R_1^{HD} \quad (4.1)$$

où le premier terme correspondant aux contributions intraméthyles, qui peut lui-même se décomposer en deux termes liés aux interactions dipolaires entre les protons du méthyle, ainsi qu'entre les protons et le carbone-13:

$$R_1^{(\text{intra-méthyle})} = R_{1,i}^{HH} + R_{1,i}^{HC} \quad (4.2)$$

Les deux autres termes de l'Equation 4.1 correspondent respectivement aux interactions entre les protons des méthyles et les autres méthyles marqués spécifiquement ainsi que les deutériums de la molécule. La description de ces quatre termes, $R_{1,i}^{HH}$, $R_{1,i}^{HC}$, $R_1^{HH(\text{inter-méthyle})}$ et R_1^{HD} , résultant des Equations 4.1 et 4.2, est exposée dans l'Annexe 2.

Dans les protéines de tailles modérées, si l'on considère que les distances interméthyles sont grandes devant les distances intraméthyles ($d_{HH}^{\text{intra-méthyle}} \ll d_{HH}^{\text{inter-méthyle}}$) et que le taux de protonation résiduelle sur les sites non échangeables est négligeable, les termes prépondérants, sur la vitesse de relaxation longitudinale, sont les vitesses de relaxation intraméthyles ($R_{1,i}^{HH}$ et $R_{1,i}^{HC}$) ainsi que le terme R_1^{HD} . Ces trois contributions dépendent des fonctions de densité spectrale aux fréquences ω_H , ω_C et ω_D (voir les Equations A2.3, A2.4 et A2.6 dans l'Annexe 2).

Dans les grands assemblages moléculaires (> 50 kDa), la relaxation induite par ces trois interactions ($R_{1,i}^{HH}$, $R_{1,i}^{HC}$ et R_1^{HD}) devient d'autant plus faible que le temps de corrélation de la molécule augmente. En revanche, dans le même temps, le terme $R_1^{HH(\text{inter-méthyle})}$ augmente de manière linéaire avec le τ_C (Equation A2.5) : il devient donc prépondérant dans les assemblages moléculaires de grandes tailles. La valeur de ce terme est certes proportionnelle au temps de corrélation de la molécule, mais elle est également très dépendante de la densité de protons localisés autour d'un méthyle. Puisque ces deux paramètres, le temps de corrélation et la densité de protons, sont dépendants de l'échantillon utilisé, il est très difficile de déterminer, de manière théorique, l'effet de chacune de ces interactions avec les protons des méthyles.

Néanmoins, j'ai déterminé expérimentalement le temps de relaxation longitudinale ^1H des méthyles sur différentes protéines spécifiquement marquées, afin d'évaluer

l'effet du temps de corrélation sur la relaxation spin-réseau. Ces valeurs sont résumées sur le Tableau 5.1. De manière générale, on observe une très faible dispersion des valeurs moyennes des temps de relaxation longitudinale ($T_1^{\text{moyen}} = 1,33 \pm 0,16$ s) pour les différents échantillons marqués spécifiquement sur les méthyles δ_1 des Isoleucines, alors que le τ_c varie de 5,5 à 590 ns.

Protéine	Marquages Isotopiques	M	T °C	τ_c	T_1
ubiquitine	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	9 kDa	25°C	5,5 ns	1,4 s
	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	9 kDa	5°C	9 ns	1,2 s
yajG	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	21 kDa	37°C	14 ns	1,3 s
MSG	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	82 kDa	37°C	55ns	1,3 s
	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	82 kDa	5°C	118 ns	1,1 s
	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3/^{12}\text{C}^2\text{H}_3$]-Leu, Val	82 kDa	37°C	55 ns	0,65 s
TET2	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	468 kDa	50°C	240 ns	1,6 s
	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	468 kDa	30°C	370 ns	1,4 s
	U- ^{2}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1	468 kDa	10°C	590 ns	1,2 s

Tableau 4.1 : Les temps de relaxation longitudinale moyens pour différentes protéines à 800 MHz.

M désigne la masse moléculaire, T est la température à laquelle l'expérience a été réalisée, τ_c est le temps de corrélation estimé pour l'échantillon à la température T et T_1 est le temps de relaxation longitudinale déterminée expérimentalement.

De plus, pour un échantillon donné, on observe une diminution de l'ordre de 15 % du temps de relaxation longitudinale T_1 lorsque le τ_c double. Pour un faible taux de protonation correspondant à un marquage spécifique des méthyles δ_1 des Isoleucines, on peut considérer que la baisse de la contribution des interactions dipolaires ^1H - ^{13}C , ^1H - ^1H intraméthyles et ^1H - ^2H compense partiellement l'effet des interactions ^1H - ^1H interméthyles. Cette dépendance, observée pour T_1 en fonction du τ_c , peut être exprimée de manière empirique sous la forme suivante :

$$\frac{1}{R_{1i}} = T_1 = T_1^{\text{ref}} \times 0,85^{\frac{\ln(\tau_c / \tau_c^{\text{ref}})}{\ln(2)}} \quad (4.3)$$

Toutefois, pour un même temps de corrélation, on observe une forte dépendance du temps de relaxation longitudinale avec la densité de protons dans la protéine. Ceci est illustré à partir de deux échantillons de MSG, l'un marqué spécifiquement sur les méthyles δ_1 des Isoleucines (§ 2.4.6.2) et l'autre sur l'un des deux méthyles γ des Valines et δ des Leucines (§ 2.5.4). Si l'on compare le marquage isotopique spécifique des Isoleucines (44 méthyles protonés), contre le double marquage spécifique des Leucines et des Valines (116 méthyles protonés), le temps de relaxation longitudinale est diminué par un facteur deux alors que la densité de protons est multipliée par un facteur 2,6.

4.4.2 Etude du transfert d'aimantation NOE en fonction de la taille de la molécule

Comme nous l'avons vu au § 3.1.4.3, le transfert d'aimantation dans une expérience NOESY dépend en partie du temps de mélange de l'expérience NOESY. Dans un système caractérisé par une faible densité de protons, l'aimantation transférée entre un méthyle i et j peut être approximée par l'Equation 3.19. Le temps de transfert optimal est donné par :

$$\tau_m^0 = \frac{1}{(R_{ii} + \sum_j \sigma_{ij})} \quad (4.4)$$

De plus, comme nous l'avons montré au § 4.4.1, la vitesse de relaxation longitudinale d'un méthyle i , R_{ii} , dépend principalement de la densité de protons entourant ce méthyle, peu du temps de corrélation de la molécule. Pour les assemblages moléculaires de grandes tailles (> 50 kDa), les termes σ_{ij} sont, eux, directement proportionnels au τ_c . Dans des systèmes faiblement protonés et de taille modérée, on a : $R_{ii} \gg \sum \sigma_{ij}$. Le temps de mélange optimal ne varie alors pas beaucoup et on peut approximer : $\tau_m^0 \approx (R_{ii})^{-1}$ (§ 3.1.4.3). En revanche, pour des assemblages moléculaires de haut poids moléculaire, on a : $R_{ii} \ll \sum \sigma_{ij}$; ainsi, le temps de mélange optimal sera directement proportionnel au temps de corrélation : $\tau_m^0 \approx (\sum \sigma_{ij})^{-1} \approx f(\tau_c)$.

En tenant compte de la dépendance en τ_c du terme R_{1i} (Equation 4.3), mais aussi des termes σ_{ij} , il est possible de simuler le transfert d'aimantation pour différents temps de mélange en utilisant l'Equation 3.19. Pour un système de spins isolés dans MSG, en considérant un marquage isotopique spécifique sur les méthyles δ_1 des Isoleucines et la structure (PDB : 1D8C), la densité autour d'un méthyle, c'est-à-dire la distance moyenne effective entre les protons d'un méthyle et les autres protons des méthyles, est de 5,5 Å (Tugarinov *et al*, 2003b). La simulation de l'aimantation transférée entre deux méthyles séparés par 5.5 Å nous indique que le transfert d'aimantation, $a_{ij}(\tau_m)$, est plus efficace lorsque le temps de corrélation des molécules augmente. Le temps de mélange optimal, τ_m^0 , est cependant de plus en plus court (Equation 4.4). Ceci s'interprète par le fait que le terme σ_{ij} croît beaucoup plus vite que $R_{1,i}$ lorsque que τ_c augmente.

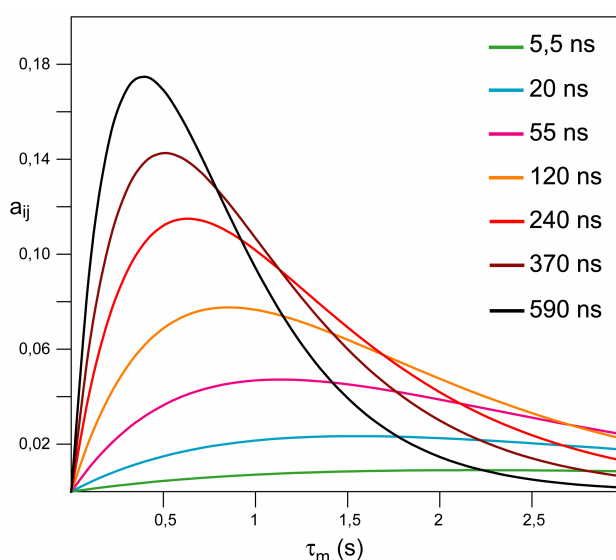


Figure 4.4 : Simulation de l'évolution du transfert d'aimantation $a_{ij}(\tau_m)$ entre deux méthyles séparés par 5.5 Å, en fonction du temps de mélange (τ_m), pour différents temps de corrélation.

L'évolution du transfert a été simulée en utilisant les Equations 3.19 et 3.11, en tenant compte de la dépendance de T_1 avec le τ_c de l'Equation 4.3, à partir des valeurs de référence de MSG à 37°C ($T_1 = 1.3$ s et $\tau_c = 55$ ns). La distance moyenne interméthyle effective utilisée a été déterminée à 5,5. Les temps de corrélation utilisés sont estimés pour des échantillons dans 100 % de D_2O et sont respectivement de : 5,5 ns pour l'ubiquitine à 25°C, 20 ns pour une protéine de taille modérée à 25°C, 55 ns pour MSG à 37°C, 120 ns pour MSG à 5°C, 240 ns pour TET2 à 50°C, 370 ns pour TET2 à 30°C et 590 ns pour TET2 à 10°C.

Le temps de transfert optimal est estimé à 1,1 s pour une protéine caractérisée par un temps de corrélation de 55 ns, telle que la MSG (courbe rose sur la Figure 4.4).

Ce temps de mélange est très proche du temps de mélange expérimental estimé pour MSG ($\tau_m^0 = 1,3$ s ; § 4.3). Lorsqu'on multiplie le τ_C par un facteur dix (courbe noire sur la Figure 4.4), on observe une augmentation de la fraction d'aimantation transférée pour un NOE d'approximativement un facteur quatre, alors que le temps de transfert optimal est réduit à 0,4 s. Néanmoins, la traduction de ce gain de transfert d'aimantation, en terme de distance maximale observable entre deux méthyles, requiert de considérer la dépendance du temps de relaxation transversale des méthyles avec le τ_C de la molécule.

4.4.3 La relaxation transversale limite la distance effective des NOEs à longue portée

Comme nous venons de le voir pour les grosses molécules (>40 kDa) spécifiquement protonées sur les méthyles, l'aimantation transférable par l'effet Overhauser entre deux méthyles augmente avec le temps de corrélation de la molécule. L'aimantation transférée se traduit, en terme de volume, par un pic de corrélation NOE observé dans la matrice 4D. Pour déterminer la distance maximale observable par cette approche, il nous faut déterminer le seuil limite de sensibilité correspondant à l'intensité minimale que doit avoir un pic de corrélation pour être détectable. Il nous faut donc tenir compte de la variation de la relaxation transversale des signaux RMN en fonction du τ_C .

La distance maximale pour laquelle on peut détecter un transfert NOE est fonction du temps de corrélation du système. En effet, même si la quantité d'aimantation transférée est plus importante lorsque le temps de corrélation augmente (Figure 4.4), la largeur de raie des signaux va également augmenter. En prenant soin d'optimiser le temps d'édition maximal dans chaque dimension indirecte, comme décrit précédemment (§ 4.2.1), en utilisant une valeur proche du temps de relaxation transversale $\langle R_2 \rangle^{-1}$, le rapport signal sur bruit, dans l'expérience 4D HMQC-NOESY-HMQC, va diminuer avec $R_2^{-3/2}$. Dans la dimension de détection directe, la perte du rapport signal sur bruit est proportionnelle à la vitesse de relaxation transversale. En remarquant que la vitesse de relaxation transversale, R_2 , est proportionnelle à τ_C (Annexe 2) et en utilisant les Equations 3.11, 4.3 et 4.4, on peut

déterminer la distance maximale qui correspond à la limite de détection de NOEs, en fonction du τ_C :

$$r_{\max} = r_{\max}^{\text{ref}} \left(\frac{\tau_C^{\text{ref}}}{\tau_C} \right)^{1/4} \left(\frac{(R_{1i} + \sum \sigma_{ij})^{\text{ref}}}{(R_{1i} + \sum \sigma_{ij})} \right)^{1/6} \quad (4.5)$$

où $r_{\max}$ correspond à la distance maximale observable pour un τ_C donné. Pour la simulation, les valeurs obtenues à partir de MSG à 37°C ($T_1 = 1,3$ s, $\tau_C = 55$ ns et $r_{\max}^{\text{ref}} = 10$ Å) m'ont servi de référence. Toutefois, l'Equation 4.5 n'est valide que dans le cadre des molécules de grandes tailles ayant la fonction de densité spectrale $J(0)$ prédominante devant les fonctions de densité spectrale aux fréquences ω_C et ω_H .

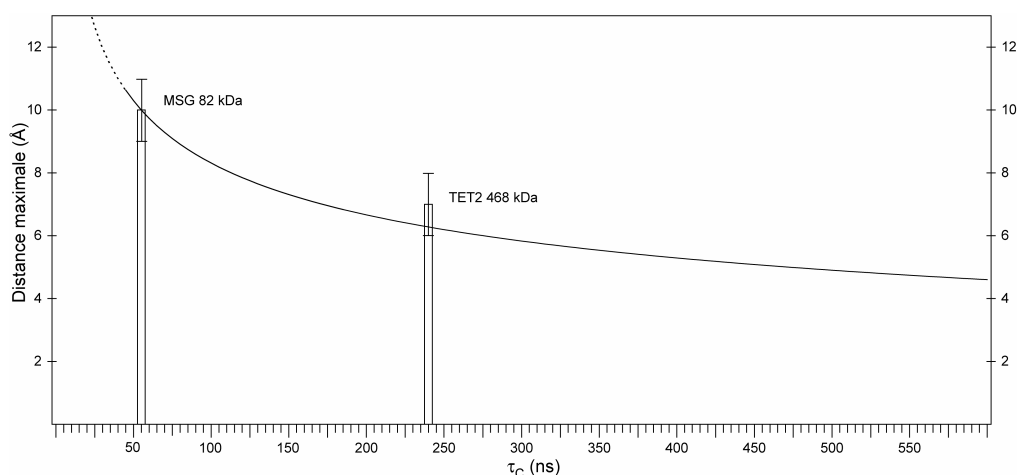


Figure 4.5 : Distance maximale observable au sein d'une paire de méthyles en fonction du temps de corrélation de la protéine.

La courbe noire correspond à l'Equation 4.5 avec les valeurs de référence suivantes : $T_1 = 1,3$ s, $\tau_C = 55$ ns et $r_{\max}^{\text{ref}} = 10$ Å. Les barres correspondent à la distance limite expérimentale pour laquelle des paires de méthyles donnent lieu à un NOE détectable. L'incertitude sur la détermination des distances est notée sur les barres.

Cette simulation nous indique, qu'avec un marquage spécifique des méthyles, de nombreux NOEs interméthyles peuvent être détectés dans des assemblages moléculaires de grande taille. Par extrapolation, si l'on considère un assemblage supramoléculaire thermophile de plus de 1 MDa, correspondant à un temps de corrélation de l'ordre de 500 ns à une température de 50°C, il paraît possible de détecter des NOEs entre de paires de méthyles distants de 5 Å.

4.5 Application à des assemblages moléculaires de grande taille

4.5.1 Utilisation de MSG comme un système modèle

En abaissant la température de l'échantillon de MSG, la viscosité du milieu augmente, entraînant une augmentation du temps de corrélation de la protéine. Le τ_c , pour une molécule sphérique, est donné par la loi de Stoke-Einstein :

$$\tau_c = \frac{4\pi V \eta}{3k_B T} \quad (4.6)$$

où η est la viscosité de la solution et dépend de la température T , V le volume de la molécule assimilée à une sphère, k_B est la constante de Boltzmann et, enfin, T est la température en Kelvin. Le temps de corrélation pour la MSG à 5°C est ainsi estimé à 118 ns, ce qui correspond au temps de corrélation d'une protéine de 200 kDa à 37°C (Tugarinov *et al*, 2003c).

La comparaison des spectres de corrélation HMQC ^1H - ^{13}C , enregistrés sur le même échantillon de MSG que précédemment, est présentée sur la Figure 4.6. On observe un élargissement des raies dans la dimension de détection (^1H). Cependant, cet élargissement est plus important pour certains pics de corrélation, ce qui conduit à leur disparition. On observe également un dédoublement de certains pics de corrélation (Cadre rouge sur la Figure 4.6). Ces deux phénomènes sont la signature de phénomènes d'échanges conformationnels, qui affectent notablement la qualité des spectres RMN de MSG à basse température. L'attribution des fréquences de résonances des pics de corrélation établie à 37°C a été reportée sur les signaux observables à 5°C, en acquérant successivement plusieurs expériences HMQC, à des températures de plus en plus faibles. Sur les 44 Isoleucines attribuées sur le spectre à 37°C, seulement 35 pics de corrélation sont attribués à 5°C, soit de l'ordre de 80%.

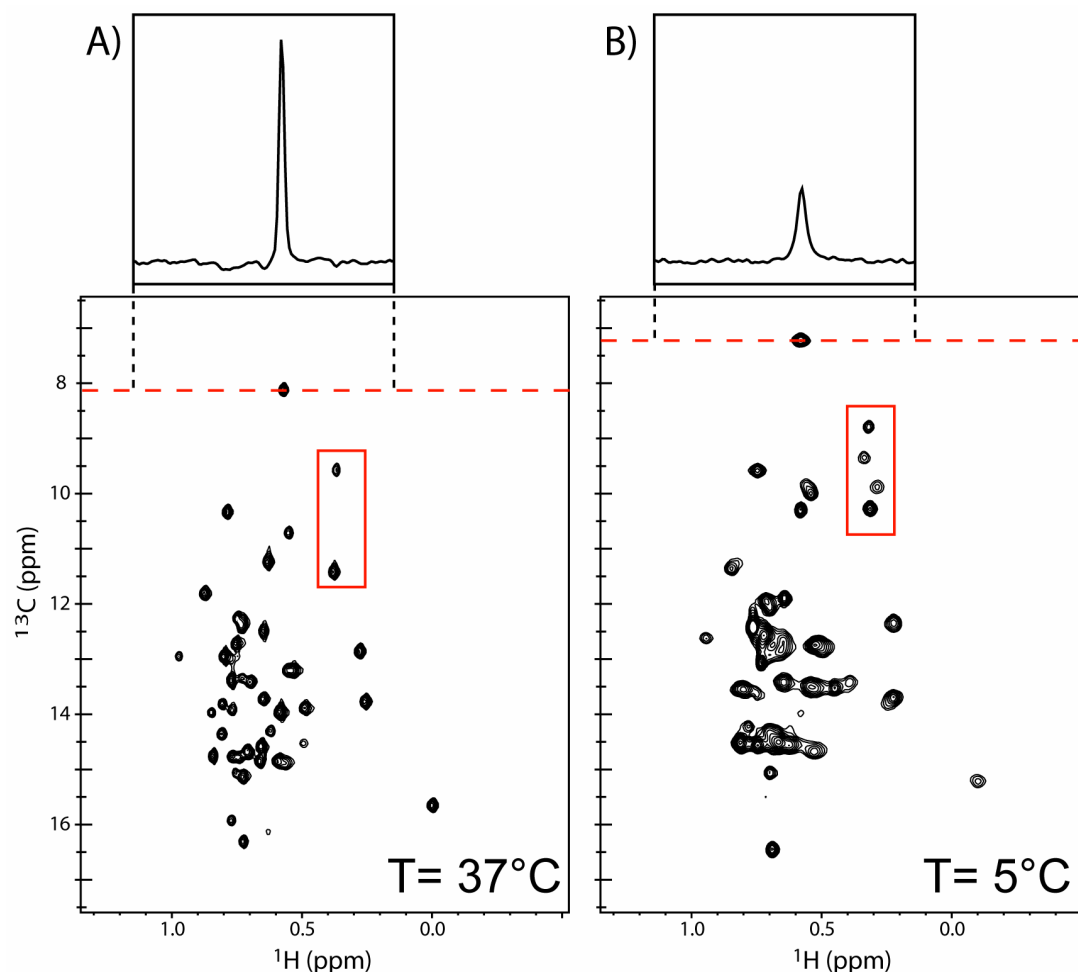


Figure 4.6 : Spectres 2D de corrélation HMQC ^1H - ^{13}C de MSG à 37°C et 5°C.

Les spectres ont été enregistrés sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, dans D_2O , sur un échantillon à 0.9mM de MSG U- ^{12}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 , à deux températures : 37°C en A) et 5°C en B). Le dédoublement des pics de corrélation entre les deux spectres est présenté dans le cadre rouge.

Une expérience d'inversion-récupération a permis de déterminer la valeur du temps de relaxation longitudinale globale à 1,1 s. Le temps de mélange optimal a été estimé à partir d'une série d'expériences 2D NOESY à $\tau_m^0 = 400$ ms. Cette valeur est deux fois plus petite que la valeur simulée (courbe orange sur la Figure 4.4). Ceci peut s'expliquer par une rotation des méthyles plus lente due à la diminution de la température. Une expérience 4D HMQC-NOESY-HMQC a été enregistrée à 5°C, pendant 68 heures, sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec un échantillon à 0,9 mM de MSG U- ^{12}H , U- ^{12}C , [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 .

J'ai comparé la distribution des distances entre méthyles δ_1 des Isoleucines, prédites à partir de la structure de la MSG, et les pics de corrélation observés dans les

spectres 4D HMQC-NOESY-HMQC à 5°C, après avoir éliminé toutes les distances impliquant des méthyles qui ne sont plus observables, suite aux phénomènes d'échanges conformationnels (en gris sur la Figure 4.8).

Cette analyse nous révèle que des NOEs entre les méthyles séparés par 7.5 Å sont toujours observables dans des macromolécules caractérisées par un τ_c de plus de 100 ns. Cette valeur est légèrement plus faible que la valeur prédite de l'ordre de 8 Å pour un tel τ_c (Figure 4.5). De plus, seuls 50 % des NOEs correspondants aux paires de méthyles distants par moins 7 Å sont observés. Cet écart peut être expliqué par les propriétés spécifiques de cette protéine à 5 °C. Pour pallier à ce problème, nous avons choisi d'étudier le transfert NOE à plus haute température dans un assemblage moléculaire de près de 500 kDa.

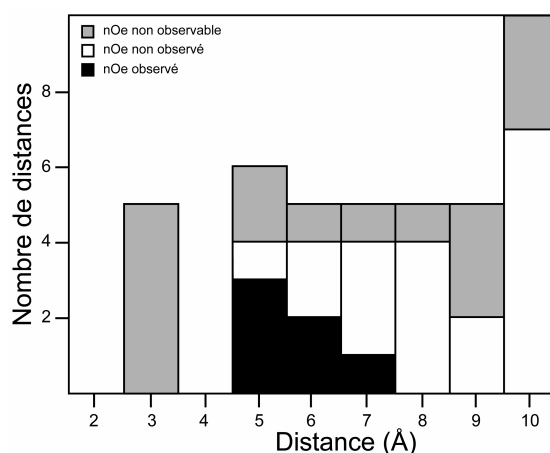


Figure 4.7 : Distribution des distances interméthyles des Ile- δ_1 dans MSG à 5°C.

Les paires de méthyles donnant lieu à un NOE observable sont indiquées en noir, les paires de méthyles pour lesquelles l'un des deux partenaires n'est plus observable, suite aux phénomènes d'échanges conformationnels, en gris et les paires de méthyles, pour lesquelles aucun NOE n'est observé, en blanc.

4.5.2 Application à l'étude de TET2, un assemblage moléculaire de 468 kDa.

TET2, pour "*Tetrahedral aminopeptidase 2*", est un homo dodécamère dont l'unité principale comporte 358 acides aminés (39 kDa), contenant 34 Isoleucines. C'est une aminopeptidase hyperthermophile, découverte au laboratoire de Biophysique Moléculaire de l'IBS, ayant son activité maximum au-delà de 95°C. La structure cristallographique de TET2 a été résolue en 2005 par Borissenko *et coll.* (Borissenko *et al*, 2005). TET2 a été surexprimé avec le protocole de marquage

spécifique sur les méthyles δ_1 des Isoleucines de la même manière que MSG (§ 2.4.6.2). Un spectre 2D de corrélation HMQC ^1H - ^{13}C , enregistré à 50 °C, est présenté sur cet échantillon (Figure 4.8). 30 pics de corrélation sont observés dans le spectre sur 34 attendus, néanmoins plusieurs recouvrements sont observés. En somme, environ 90 % des méthyles δ_1 des Isoleucines donnent lieu à un pic de corrélation observable.

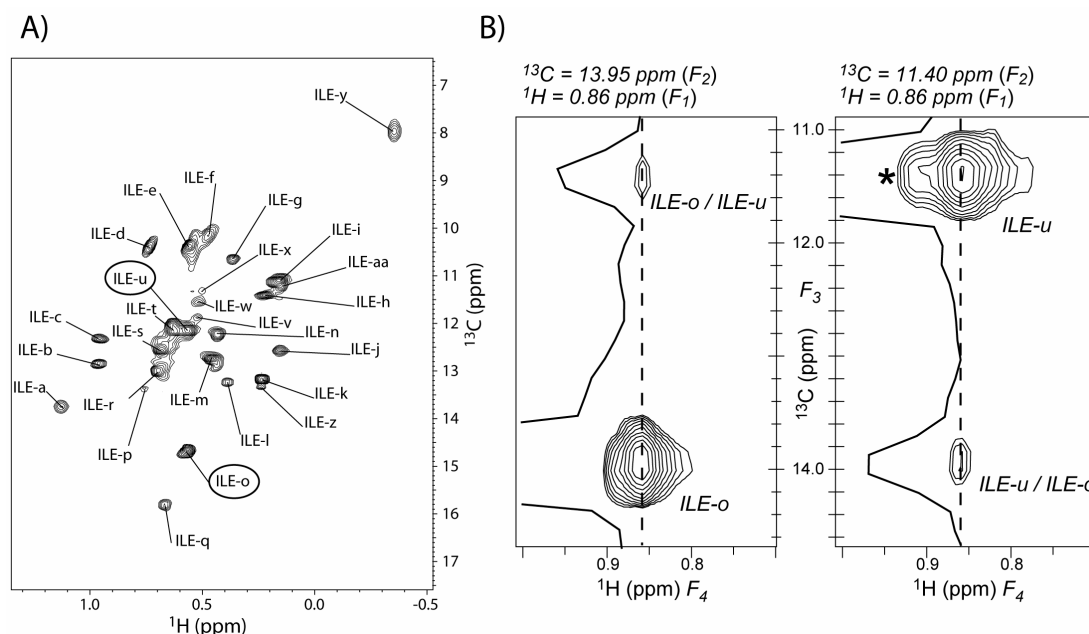


Figure 4.8 : A) Spectres 2D de corrélation HMQC ^1H - ^{13}C de TET2. B) Plans de l'expérience 4D NOESY des pics diagonaux de Ile-o et Ile-u ainsi que les pics de corrélation correspondants.

Les deux expériences ont été enregistrées sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, à 50°C, avec un échantillon à 0,6 mM de U-[^2H], U-[^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 , dans 99,9 % de D_2O . En A), la numérotation des pics de corrélation est notée sur le spectre à deux dimensions. En B), les sélections de droite correspondent aux plans $F_1(^1\text{H})$ - $F_2(^{13}\text{C})$ de Ile-o et celles de gauche aux plans $F_1(^1\text{H})$ - $F_2(^{13}\text{C})$ de Ile-u. Les valeurs des paramètres d'acquisition et les délais d'incrémentation de l'expérience 4D NOESY sont : $t_{1\text{max}} = 18.9 \text{ ms}$, $t_{2\text{max}} = 18.5 \text{ ms}$, $t_{3\text{max}} = 13 \text{ ms}$ et 80 ms de temps d'acquisition a été utilisé dans la dimension directe ^1H . Les largeurs spectrales sont respectivement : $F_1(^1\text{H}) = 0,92 \text{ ppm}$, $F_2(^{13}\text{C}) = 5,6 \text{ ppm}$, $F_3(^{13}\text{C}) = 5,6 \text{ ppm}$ et $F_4(^1\text{H}) = 2,0 \text{ ppm}$. Les pics annotés d'une étoile correspondent à des pics se trouvant dans d'autres plans. La distance au sein de la paire de méthyles a été déterminée à $7 \pm 1 \text{ \AA}$.

Comme pour les autres échantillons, j'ai déterminé le temps de relaxation longitudinale globale de TET2 par une expérience d'inversion-récupération. La valeur est déterminée à 1,6 s. Une série d'expériences NOESY a permis d'estimer le temps de mélange optimal à $\tau_m^0 = 500 \text{ ms}$ pour cet assemblage dont le τ_c est estimé à 240 ns à 50°C. Cette valeur est en très bon accord avec la simulation (courbe rouge

sur la Figure 4.4). Un spectre 4D HMQC-NOESY-HMQC a été enregistré à 50°C, pendant 63 heures, sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec un échantillon à 0,6 mM de TET2 U- ^2H], U- ^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 .

L'analyse quantitative des NOEs observés dans les spectres de l'expérience 4D HMQC-NOESY-HMQC a permis la quantification des distances entre onze paires de méthyles, distants de 5 à 7 Å, et une douzième paire de pics de corrélation, correspondant à une distance de plus de 9 Å (cette distance peut être surestimée par la dynamique d'un des méthyles). Comme l'attribution de cette protéine n'est pas connue, on ne peut pas comparer directement ces distances avec la structure cristallographique de TET2. Néanmoins, si on fait l'hypothèse que les NOEs observés correspondent aux paires de méthyles les plus proches dans la protéine, l'analyse de la distribution des distances entre les méthyles δ_1 dans TET2 indique que les 12 premières paires sont séparées par des distances comprises entre 3.3 et 7.0 Å (Figure 4.9). Ainsi, il est permis d'établir que des NOEs atteignant 7 Å peuvent être observés dans de tels assemblages, en très bon accord avec la prédiction effectuée (Figure 4.5).

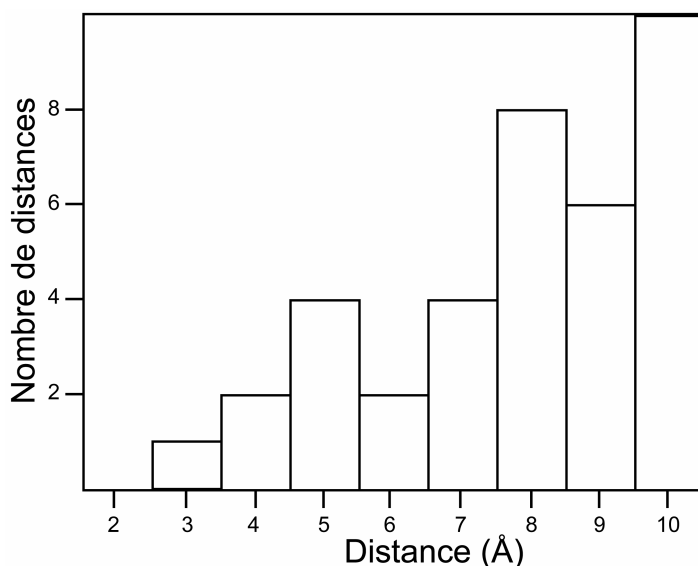


Figure 4.9 : Distribution des distances interméthyles des Ile- δ_1 dans la protéine TET2. L'axe des abscisses correspond à la distance moyenne interprotons déterminée à partir de la structure de TET2 (PDB : 1Y0Y).

4.6 Les NOEs interméthyles pour les études structurales

4.6.1 Application à l'étude d'une protéine globulaire de 21 kDa

La stratégie mise au point et présentée dans les paragraphes précédents a été appliquée pour mesurer les distances interméthyles sur yajG, une protéine globulaire de 21 kDa, dont la structure était en cours d'étude au laboratoire. Un échantillon U-[²H], U-[¹²C], [¹³C¹H₃]-Ile- δ_1 de yajG a été produit (§ 2.4.6.2) et utilisé pour détecter des NOEs interméthyles. Onze corrélations NOEs entre paires de méthyles ont été observées. Les distances expérimentales interméthyles, extraites à l'aide du programme RELOAD, sont présentées dans le Tableau 4.2.

Résidu 1	Résidu 2	Distance expérimentale	Borne sup.	Borne inf
7	105	6,33 $\pm$ 0.04	6,86	6,20
7	144	7,56 $\pm$ 0.10	8,15	7,37
13	156	5,57 $\pm$ 0.02	6,06	5,48
87	156	4,68 $\pm$ 0.12	5,12	3,68
88	113	4,54 $\pm$ 0.01	4,98	3,54
105	144	6,83 $\pm$ 0.04	7,38	6,68
111	113	6,47 $\pm$ 0.21	7,00	6,34
111	156	7,25 $\pm$ 0.16	7,82	7,08
113	165	5,19 $\pm$ 0.26	5,66	5,12
114	165	6,02 $\pm$ 0.47	6,69	5,75
165	169	6,04 $\pm$ 0.49	6,73	5,75

Tableau 4.2 : Table des distances interméthyles à longue portée utilisée dans le calcul de structure de yajG.

La valeur de l'erreur expérimentale déterminée en fonction du rapport signal sur bruit dans les spectres est indiquée dans la colonne des distances expérimentales. Les bornes inférieures et supérieures ont été déterminées comme indiqué dans le texte. Une correction de + 0,2 Å est appliquée aux distances expérimentales, car, dans le calcul, les distances impliquant les protons des méthyles sont moyennées en $1/\langle r^6 \rangle$ au lieu de $1/\langle r^3 \rangle^2$ de manière théorique (Tropp, 1980).

Les distances dérivées de ces NOEs ont été incorporées dans un protocole de calcul utilisant le programme CNS. Les bornes inférieures et supérieures des contraintes de distance ont été obtenues à partir de l'erreur expérimentale. Cette erreur représente l'incertitude liée au rapport signal sur bruit des spectres NOESY. Dans le cas des distances expérimentales extraites avec précision, une erreur minimale de 5 % a été considérée, afin de tenir compte des 5 % d'inexactitude des distances expérimentales observées sur l'ubiquitine (§ 3.2.2.1). Pour les distances inférieures à 5 Å, la borne inférieure a été fixée à 1 Å de la distance expérimentale pour tenir compte du biais systématique observé, concernant les faibles distances (§ 3.2.2.1).

Afin d'évaluer l'influence de ces contraintes sur la qualité des structures, deux calculs ont été lancés en parallèle dans les mêmes conditions. Le premier calcul sert de référence ; dans le second, les onze NOEs interméthyles ont été ajoutés. L'incorporation de ces onze contraintes supplémentaires, soit moins de 0,5 % du nombre de contraintes NOEs utilisées (Tableau 4.3), permet d'augmenter de manière significative la précision du calcul, sans inclure de violation supplémentaire. Une amélioration du RMSD est observée aussi bien sur le squelette que sur l'ensemble des atomes lourds (Tableau 4.3).

	Calcul 1	Calcul 2
Nombre de NOEs non ambigus	2723	2723
CH ₃ -CH ₃ Ile δ_1 NOEs	0	11
RMSD chaîne principale	1,96 $\pm$ 0,27 Å	1,76 $\pm$ 0,25 Å
RMSD atome lourd	2,70 $\pm$ 0,25 Å	2,52 $\pm$ 0,19 Å

Tableau 4.3 : Les paramètres utilisés et les résultats des calculs de structure.

Les nombres de NOEs non ambigus utilisés dans le Calcul 1, sans les contraintes à longue portée entre paires de méthyles δ_1 des Isoleucines, et dans le Calcul 2, avec ces contraintes, sont indiqués. Les RMSD moyens, entre toutes les paires constituant l'ensemble des 5 structures sélectionnées par rapport aux 20 structures calculées, de la chaîne principale et des atomes lourds des résidus 28 à 173, sont indiqués.

Ceci est d'autant plus étonnant, car dans une protéine globulaire telle que yajG, un très grand nombre de NOEs standard peut être extrait et suffit à définir le repliement

d'une telle protéine. Cependant, ces contraintes interméthyles à longue portée auront un effet beaucoup plus important dans le cas de protéines modulaires, allongées, ou de complexes de grandes tailles, pour lesquels très peu de NOEs définissent la position relative des différents domaines, extrémités ou partenaires de l'édifice.

4.6.2 Perspectives : applications aux protéines modulaires ou assemblages moléculaires de grandes tailles.

L'approche utilisant le marquage spécifique des méthyles des Isoleucines et la détection de NOEs à longue portée peut donc servir à affiner les structures de protéines. En particulier, ces contraintes sont très importantes dans la détermination de structures de protéines modulaires ou de complexes protéiques de grandes tailles. Dans ces systèmes, les approches RMN classiques permettent l'obtention de contraintes locales et, ainsi, la détermination précise de la structure des différents domaines globulaires avec une résolution relativement bonne. En revanche, le manque d'informations à longue portée ne permet pas d'avoir la position relative des différents domaines.

4.6.2.1 Les domaines de MSG

L'exemple de la MSG est intéressant pour illustrer ce problème. Sur la structure proposée par Tugarinov *et coll.* (Tugarinov *et al*, 2005a) ; les RMSD de chacun des quatre domaines de la MSG sont compris entre 1 et 3 Å :

- "α-helical clasp" (3-88) : 1.40 ± 0.14 Å.
- "Core" (116-132 ; 266-295 ; 334-550) : 3.37 ± 0.41 Å.
- "α/β domain" (135-262 ; 296-333) : 1.45 ± 0.11 Å.
- "C-terminal domain" (589-722) : 1.98 ± 0.32 Å.

Néanmoins, l'ensemble de la protéine ne peut être superposé qu'avec un RMSD de 4.06 ± 0.45 Å. Ceci est directement lié au manque de contraintes translationnelles entre chacun des domaines. Ce manque d'informations à longue portée limite la résolution des modèles structuraux proposés pour des systèmes biologiques de

grande taille. La mesure d'un plus grand nombre de contraintes entre les différents domaines est envisageable, en utilisant un marquage spécifique des méthyles de différents résidus. L'approche présentée jusqu'alors, concernant la mesure de distances interméthyles, utilise le marquage spécifique sur les méthyles δ_1 des Isoleucines. Dans le cas de MSG, seules 7 distances interméthyles mesurées sur 34 concernent des distances entre les différents domaines (§ 4.3). Tout seul, ce type de marquage ne sera donc pas suffisant pour extraire le nombre nécessaire de contraintes interdomaines. Cependant, la méthode développée pour détecter des NOEs à longue portée entre méthyles peut être généralisée aux autres résidus comportant des méthyles.

4.6.2.2 Vers l'affinement des modèles structuraux de protéines de grandes tailles : marquages spécifiques des méthyles et nouvelles contraintes interméthyles

Comme nous l'avons vu dans le chapitre 2, différents protocoles existent aujourd'hui pour marquer spécifiquement les méthyles des Isoleucines, des Valines, des Leucines et des Alanines. Le marquage spécifique d'un résidu peut ainsi fournir de nouvelles contraintes à longue portée permettant d'améliorer la précision des modèles de structure des protéines de grandes tailles.

Afin d'étudier les types de marquage les plus adaptés pour l'affinement de la structure de MSG, j'ai considéré le marquage spécifique des méthyles des différents résidus : le méthyle β de l'Alanine, les méthyles δ_1 et γ_2 de l'Isoleucine, le méthyle γ_2 de la Thréonine, le méthyle ϵ de la Méthionine et les méthyles δ de la Leucine ainsi que des γ de la Valine. Pour déterminer le nombre de NOEs interdomaines qui pourraient être détectés pour chaque type de marquage, j'ai utilisé les résultats obtenus avec le marquage des méthyles δ_1 des Isoleucines comme référence et j'ai pris en compte la différence de la densité de protons, d'un marquage à l'autre, pour déterminer la distance maximale interméthyle pouvant donner lieu à des NOEs observables. L'ensemble de ces données simulées est représenté dans le Tableau 4.4.

Marquage	Ile δ_1	Ile γ_2	Val / Leu	Met	Ala	Thr
Densité de protons	2 %	2 %	6 %	1 %	4 %	2 %
NOEs interdomaines	7	3	137*	2	19	6
NOEs Total	32	24	427*	8	48	14
Distance maximale	10 Å	10 Å	7 Å	11 Å	10 Å	10 Å

Tableau 4.4 : Nombres de NOEs interméthyles observables en fonction du marquage spécifique utilisé dans la MSG.

La densité de protons est déterminée en considérant la perdeutération de l'échantillon sauf pour le méthyle du résidu marqué spécifiquement, dans le cas des Valines et des Leucines, seulement l'un des deux méthyles est protoné *. Les distances ont été calculées à partir de la structure de MSG résolue par RMN (PDB : 1Y8B). Les résultats obtenus avec le marquage spécifique des Isoleucines ont été utilisés comme référence. Le nombre de NOEs a ainsi été déterminé, en considérant comme limite la distance où tous les NOEs sont observés. Cette distance a été modulée pour chaque combinaison afin de tenir compte de la densité de protons. * nombre de NOEs total, mais 4 NOEs sont détectables, en général, par paires de méthyles, nombre à diviser par 4 pour obtenir le nombre de NOEs indépendants.

Le marquage spécifique des méthyles des Isoleucines

Le premier marquage spécifique effectué est le marquage du méthyle δ_1 de l'Isoleucine. Comme nous l'avons vu au § 4.3, le marquage spécifique du méthyle δ_1 de l'Isoleucine a permis d'observer expérimentalement 34 NOEs, dont 7 interdomaines. Ceci concorde parfaitement avec les 32 NOEs observables, prédits théoriquement et auxquels viennent s'ajouter 2 NOEs supplémentaires à plus grandes distances.

Le marquage des méthyles γ_2 des Isoleucines apporte une information redondante avec l'information qui peut être obtenue à l'aide d'un marquage des méthyles δ_1 . De plus, il n'existe pas de protocole de production de manière spécifique des méthyles γ_2 de l'Isoleucine. Comme nous l'avons vu au § 2.4.2, le méthyle γ_2 provient directement de la conversion du pyruvate et est donc difficile à incorporer spécifiquement. Par conséquent, de manière générale, l'utilisation du méthyle γ_2 de l'Isoleucine ne semble pas pertinente dans l'obtention de nouvelles contraintes interdomaines.

Le marquage spécifique d'un des deux méthyles des Leucines et des Valines

Comme présenté dans le § 2.5.4, la protonation spécifique d'un seul des deux méthyles des Leucines et des Valines peut être effectuée avec un protocole proche de celui utilisé pour le marquage des méthyles δ_1 des Isoleucines. Afin de limiter la diffusion de spins entre les deux méthyles du même résidu, il ne faut utiliser, comme précurseur, l' α -cétoisovalérate protoné spécifiquement que sur l'un de ces deux méthyles. A ce jour, il est disponible commercialement, sous une forme non stéréosélective. L'utilisation d'un tel précurseur conduit donc à une protonation à 50% de chaque méthyle. En prenant en compte la plus forte densité de protons correspondante aux 232 méthyles marqués à 50 % des Leucines et des Valines, le seuil de détection des NOEs chute à 7 Å car l'intensité des pics de corrélations n'est que de 25 % de la valeur obtenue pour un marquage uniforme. Néanmoins, du fait du grand nombre de méthyles répartis dans la protéine, 137 NOEs interdomaines sont à priori détectables. L'analyse préliminaire d'une expérience 4D NOESY, enregistrée avec un échantillon U- ^2H], U- ^{12}C], [$^{13}\text{C}^1\text{H}_3$ / $^{12}\text{C}^2\text{H}_3$]-Leu, Val de MSG, révèle que la plupart des NOEs interdomaines sont observés. Chaque paire de Leucine/Valine pouvant donner jusqu'à 4 NOEs, on peut donc considérer, en première approximation, que seul un quart de ces distances sont indépendantes. Toutefois, le fait d'avoir plusieurs distances interméthyles entre deux acides aminés permettra de positionner, relativement, ces deux résidus avec une plus grande précision. Ce type de marquage paraît donc potentiellement très intéressant pour positionner les domaines relativement les uns par rapport aux autres. De plus, un développement intéressant, pour limiter la perte du signal du pic de corrélation présent à 25 %, serait un marquage stéréosélectif des méthyles à partir de l' α -cétoisovalérate préparé chimiquement. Dans l'hypothèse d'obtenir un échantillon avec les méthyles des Valines et des Leucines stéréosélectivement protonés, des NOEs interméthyles distants de 9 Å pourraient être détectés.

Le marquage spécifique des méthyles des Méthionines

En ce qui concerne la Méthionine, son utilisation ne semble pas être un choix très pertinent dans le cas de MSG (les Méthionines ne représentent que 3 % des acides aminés dans MSG). Toutefois, ce marquage pourrait être intéressant dans le cas de

protéines plus riches en Méthionines. La limite ne concerne pas le marquage spécifique du méthyle de la Méthionine. En effet, le marquage par inhibition est envisageable, pour la Méthionine, en utilisant un acide aminé totalement deutéré, sauf sur les protons du méthyle. Il existe aussi la possibilité de préparer un échantillon selon le protocole proposé par Fischer *et coll.*, utilisant un précurseur des Méthionines, l' α -céto-méthylthio-butyrates (Fischer *et al.*, 2007).

Le marquage spécifique des méthyles des Alanines

L'utilisation de l'Alanine est vraiment très intéressante. En effet, c'est l'acide aminé le plus abondant, à l'inverse de la Méthionine, et un grand nombre de NOEs est donc observable entre les méthyles des Alanines. Ces NOEs apportent également une information très précise grâce à la faible mobilité de la chaîne latérale (§ 2.6.1). Un premier protocole a été proposé par Isaacson *et coll.* (Isaacson *et al.*, 2007) mais il ne permet pas de marquer à plus de 50 % les méthyles des Alanines. Nous avons un autre protocole qui permet un marquage spécifique des Alanines avec un taux d'incorporation à plus de 95 % (§ 2.6). Prochainement, un échantillon de MSG deutéré et spécifiquement protoné sur les Alanines, en utilisant le protocole exposé au § 2.6.5, est en cours de préparation. A priori, 19 nouvelles contraintes interdomaines entre paires de méthyles séparés par moins de 9 Å seront détectables dans cet échantillon.

Le marquage spécifique des méthyles des Thréonines

Enfin, la Thréonine est aussi avantageuse que l'Isoleucine, du point de vue de son abondance et de la présence d'un unique méthyle protoné dans l'acide aminé. Ce type de marquage devrait permettre de détecter des NOEs entre méthyles distants par 10 Å. Aucun protocole de marquage spécifique n'a encore été proposé dans la littérature. Un protocole utilisant un marquage par inhibition (§ 2.2), en utilisant la Thréonine spécifiquement protonée, est aussi en développement au laboratoire. Il faut noter qu'il est nécessaire de saturer le milieu de culture avec de l' α -cétobutyrate deutéré pour éviter le marquage en δ_1 de l'Isoleucine (Figure 2.3).

Les contraintes interdomaines dans MSG

La répartition des méthyles et des distances interdomaines détectables pour un marquage spécifique Ile- δ_1 , Ala- β , Val & Leu, ou Thr- γ est représentée sur la Figure 4.10. Ces quatre marquages indépendants devraient permettre d'extraire entre 70 à 80 NOEs interdomaines. En ce qui concerne la réalisation expérimentale de ces marquages, j'ai effectué le marquage sur les méthyles des Isoleucines, des Leucines et des Valines, mais le marquage du méthyle des Thréonines est en cours de développement à l'IBS, quant au marquage spécifique des Alanines, l'échantillon est en cours de production. Ces échantillons seront employés à l'extraction des distances interméthyles qui seront utilisées pour déterminer précisément la structure de cette protéine modulaire de plus de 80 kDa.

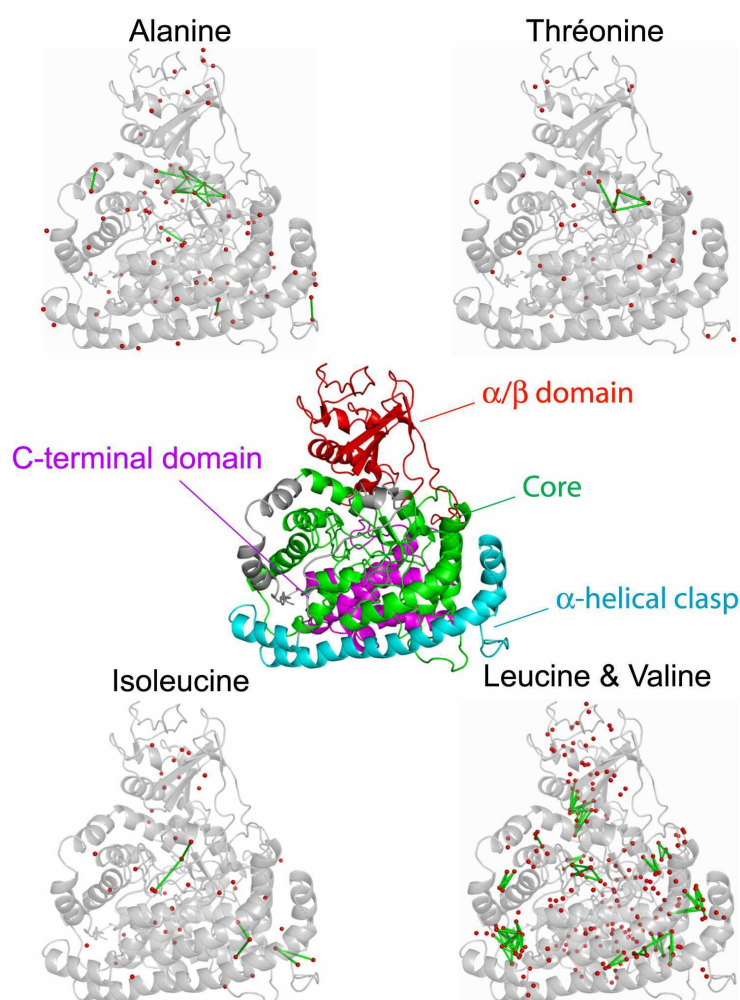


Figure 4.10 : Structures secondaires de la MSG avec la répartition des méthyles (sphères rouges) et des NOEs interdomaines (barres vertes).

Seuls les NOEs interméthyles entre les différents domaines de la MSG sont présentés en fonction du type de marquage spécifique utilisé : U- ^2H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$ -Ala- β pour l'Alanine, U- ^2H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$ -Thr- γ_2 pour la Thréonine, U- ^2H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$ -Ile- δ_1 pour l'Isoleucine, U- ^2H , U- ^{12}C , $^{13}\text{C}^1\text{H}_3$, $^{12}\text{C}^2\text{H}_3$ -Leu,Val pour la Leucine et la Valine.

4.6.2.3 Les combinaisons de marquages spécifiques : source de nouvelles contraintes interméthyles

Suite à la description des marquages indépendants, nous pouvons envisager les combinaisons des approches. Sur le Tableau 4.5, j'ai représenté toutes les combinaisons de deux marquages spécifiques envisageables *in-vivo*, pour la MSG, à condition que les précurseurs spécifiquement marqués ^1H , ^2H et ^{13}C soient disponibles. La première combinaison envisageable consiste au marquage (Ile + Leu/Val) (Tugarinov *et al*, 2002). Cette combinaison est la plus aisée à réaliser. Même si Tugarinov *et coll.* ont produit un tel échantillon, il n'ont détecté que 386 NOEs interméthyles sur 606 estimés (Tugarinov *et al*, 2005a). 18 distances interdomaines indépendantes supplémentaires pourraient être observés. La combinaison (Alanine ; Leucine/Valine) désormais réalisable permettrait d'obtenir 20 contraintes interdomaines supplémentaires. Seules trois autres combinaisons permettent d'obtenir plus de dix informations interdomaines : (Ala + Met), (Ala + Thr) et (Ile + Thr). Au total, ces cinq combinaisons pourraient permettre d'obtenir approximativement 70 nouvelles contraintes interdomaines afin d'améliorer la précision de la structure de cette protéine modulaire.

	Val / Leu	Met	Ala	Thr
Ile δ_1	36 / 147*	3 / 13	8 / 30	11 / 32
Val / Leu		7 / 23*	40 / 163*	16 / 49*
Met			10 / 28	2 / 10
Ala				10 / 28

Tableau 4.5 : Nombre de nouveaux NOEs interméthyles observables en fonction de la combinaison de marquage spécifique utilisé dans la MSG.

Les distances ont été calculées à partir de la structure de MSG résolue par RMN (PDB : 1Y8B). La première valeur, dans chaque colonne, représente le nombre de nouvelles distances interdomaines entre les méthyles de type A/B (les NOEs A/A et B/B sont déjà reportés dans le Tableau 4.4), et la deuxième valeur, le nombre total de nouvelles distances détectées. Les résultats obtenus avec le marquage spécifique des Isoleucines ont été utilisés comme référence. Le nombre de NOEs a ainsi été déterminé, en considérant comme limite la distance où tous les NOEs sont observés. Cette distance a été modulée pour chaque combinaison afin de tenir compte de la densité de protons. * nombre de NOEs total, mais 2 NOEs sont détectables, en général, par paire de méthyles ; nombre à diviser par 2 pour obtenir le nombre de NOEs indépendants.

4.7 Conclusions

A partir des résultats obtenus sur les protéines de tailles modérées spécifiquement protonées sur les méthyles, les simulations nous ont indiquées que, pour une protéine avec un temps de corrélation de 50 ns, des pics de corrélation NOEs sont observables entre des méthyles séparés par des distances allant jusqu'à 10 Å. En outre, ces simulations nous ont montré que si la densité de protons demeure faible (2-3%), la diffusion de spins contribue pour moins de 10% des volumes de ces pics de corrélation NOEs (§3.2.2.2). J'ai pu démontrer, expérimentalement, que cette approche était possible pour détecter des NOEs entre paires de méthyles δ_1 des Isoleucines sur une protéine monomérique de masse moléculaire de 82 kDa, la Malate Synthase G. Cette stratégie paraît particulièrement prometteuse pour déterminer la structure précise de protéines modulaires, de forme allongée ou de complexes protéiques de plus de 50 kDa.

De plus, j'ai démontré que des NOEs interméthyles étaient toujours détectables entre méthyles distants par plus de 7 Å dans des assemblages de près de 500 kDa. Par extrapolation, en considérant un temps de corrélation de 500 ns, c'est-à-dire un complexe protéique de 1 MDa à une température de 50°C, il serait possible de détecter des NOEs séparés par des distances allant jusqu'à 5 Å. En utilisant cette stratégie de marquage spécifique des méthyles, il est donc possible de détecter des NOEs à des distances équivalentes à ce qui se fait, par les méthodes RMN standard, dans des protéines protonées de moins de 30 kDa. Il a été démontré que l'utilisation des NOEs interméthyles était une information cruciale pour déterminer le repliement global de protéines globulaires (Gardner *et al*, 1997b ; Bersch *et al*, 2003). Ces travaux ouvrent donc la porte à la détermination, dans un futur proche, du repliement global de protéines par RMN au sein d'assemblages de plusieurs centaines de kDa. Néanmoins, la première limitation pour une telle étude par RMN réside dans l'étape d'attribution des fréquences de résonances des méthyles. Les techniques triples résonances standard développées pour l'attribution des petites protéines globulaires ne peuvent être transposées à de tels assemblages moléculaires. On peut imaginer que ce travail d'attribution soit résolu par l'utilisation combinée de mutagenèse dirigée et de marquage isotopique pour lever ce verrou.

*Chapitre**5*

*Détection de contraintes
translationnelles à longue portée
par RMN en milieu cristal
liquide*

5.1 Les couplages dipolaires résiduels (RDCs)

Ces dernières années la RMN en milieu anisotrope s'est imposée comme une technique de choix pour obtenir des informations structurales précises dans les biomolécules. Après avoir introduit le paramètre structural principal de la RMN en milieu cristal liquide, le couplage dipolaire résiduel (RDC), je montrerai que celui-ci apporte des informations orientationnelles précises, mais qu'il peut également donner lieu à des contraintes translationnelles. Je présenterai les difficultés liées à la détection de RDCs entre protons éloignés dans une biomolécule et les améliorations tant dans le domaine spectroscopique qu'au niveau des marquages isotopiques. Je démontrerai qu'il est possible de diminuer considérablement le temps expérimental nécessaire pour détecter des RDCs entre protons proches dans l'espace ($d_{HH} < 7 \text{ \AA}$) dans les protéines protonées. Dans un second temps, je montrerai que l'utilisation des marquages isotopiques des méthyles, présentés au chapitre 2, associé à des outils spectroscopiques adaptés permet de détecter précisément des RDCs entre méthyles distant au-delà de $10 \text{ \AA}$ dans les protéines spécifiquement protonées.

5.1.1 Théorie

L'interaction dipolaire est due au couplage dipolaire entre deux spins nucléaires $\vec{P}$ et $\vec{Q}$ placés dans un champ magnétique $\vec{B}_0$ (Figure 5.1). En utilisant un formalisme vectoriel par souci de simplification d'écriture, l'hamiltonien correspondant s'exprime sous la forme (Levitt, 2001):

$$\hat{H}_{PQ} = \frac{D_{PQ}^{\max}}{r_{PQ}^5} \left[3(\vec{P} \cdot \vec{r}_{PQ})(\vec{Q} \cdot \vec{r}_{PQ}) - r_{PQ}^2 \vec{P} \cdot \vec{Q} \right] \quad (5.1)$$

où $\vec{r}_{PQ}$ représente le vecteur internucléaire, et $\vec{P}$, $\vec{Q}$ les deux opérateurs de spin concernés. L'amplitude de cette interaction est donnée par la constante de couplage dipolaire :

$$D_{PQ}^{\max} = -\frac{\mu_0}{4\pi^2} \gamma_P \gamma_Q \hbar \quad (5.2)$$

où γ_P et γ_Q sont les rapports gyromagnétiques des deux spins, μ_0 est la constante de perméabilité magnétique du vide et $\hbar$ est la constante de Planck divisée par 2π .

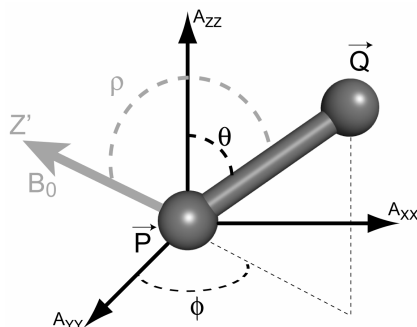


Figure 5.1 : Vecteur dipolaire internucléaire PQ.

Le vecteur du champ magnétique, B_0 , est indiqué en gris et dans le référentiel du tenseur d'alignement A (en noir). ρ est l'angle que fait le vecteur internucléaire avec le champ magnétique et θ et ϕ sont les coordonnées sphériques décrivant l'orientation du vecteur PQ dans le repère principal de A.

Considérons plusieurs paires de spins $\vec{P}_i$ et $\vec{Q}_j$ soumises à l'interaction dipolaire décrite ci-dessus, l'expression de cette interaction peut être simplifiée dans le cadre de l'approximation séculaire selon laquelle seule la composante Z du champ magnétique local du spin $\vec{P}_i$ (composante dans la direction du champ $\vec{B}_0$ appliqué ; Figure 5.1) sur le spin $\vec{Q}_j$ est significative (la condition mathématique équivalente à cette approximation est $\Delta\omega = \omega_{P_i} - \omega_{Q_j} \gg D_{P_i Q_j}^{\max}$, autrement dit l'interaction dipolaire peut être traitée comme une perturbation après l'interaction Zeeman) :

$$\hat{H}_{P_i Q_j} = \frac{D_{P_i Q_j}^{\max}}{2r_{P_i Q_j}^3} \langle 3 \cos^2 \rho_{ij} - 1 \rangle P_{iz} Q_{jz} = D_{ij} P_{iz} Q_{jz} \quad (5.3)$$

où D_{ij} est la constante de couplage dipolaire, ρ_{ij} est l'angle que fait le vecteur internucléaire avec le champ magnétique $\vec{B}_0$. P_{iz} et Q_{iz} sont les composantes selon Z des moments angulaires des spins P et Q. On notera dans cette expression, les "brackets" de part et d'autre de la transformation géométrique.

La molécule se réoriente par rapport au champ magnétique externe $\vec{B}_0$ selon un temps caractéristique de l'ordre de la nanoseconde, ns. A l'échelle de temps de la RMN ($> ms$) ; l'interaction dipolaire internucléaire détectée se trouve donc moyennée comme indiquée par les "brackets".

Dans les conditions standard de la RMN liquide, la réorientation moléculaire est isotrope et cette interaction est moyennée à zéro. Cependant, pour des molécules possédant une orientation préférentielle par rapport au champ magnétique, l'interaction dipolaire n'est alors plus moyennée à zéro et peut être mesurée. Le couplage dipolaire observé entre deux noyaux P et Q peut donc s'exprimer en fonction de la norme et de l'orientation du vecteur PQ dans le tenseur d'alignement préférentiel, A, de la molécule (Bothner-By, 1996 ; Prestegard, 1998 ; Prestegard *et al*, 2004 ; Bax *et al*, 2001 ; Blackledge, 2005) :

$$D_{PQ}(\theta, \phi) = S \frac{D_{PQ}^{\max}}{2r_{PQ}^3} \left[A_a (3 \cos^2 \theta - 1) + \frac{3}{2} A_r \sin^2 \theta \cos 2\phi \right] \quad (5.4)$$

où S est le paramètre d'ordre pour le mouvement interne du vecteur PQ, r_{PQ} est la distance entre P et Q, θ et ϕ sont les coordonnées sphériques décrivant l'orientation du vecteur PQ dans le repère principal de A (Figure 5.1). Les valeurs de composantes axiale (A_a) et rhombique (A_r) du tenseur d'alignement (A) dépendent à la fois de la forme de la protéine, de sa charge mais aussi du milieu orientant utilisé.

S'il existe un couplage scalaire J_{PQ} entre les deux noyaux P et Q, les interactions scalaires et dipolaires pour de faibles alignements conduisent simultanément à l'éclatement des résonances RMN. La Figure 5.2 illustre ce résultat dans le cas d'une paire de spins nucléaires ^1H - ^{15}N . Néanmoins, seul le couplage scalaire est indépendant de l'orientation de la molécule dans le champ magnétique. Ainsi, la mesure du couplage dipolaire, D_{PQ} , sera effectuée par différence entre une mesure en milieu anisotrope ($J_{PQ}+D_{PQ}$) et une mesure de couplage en milieu isotrope (J_{PQ}).

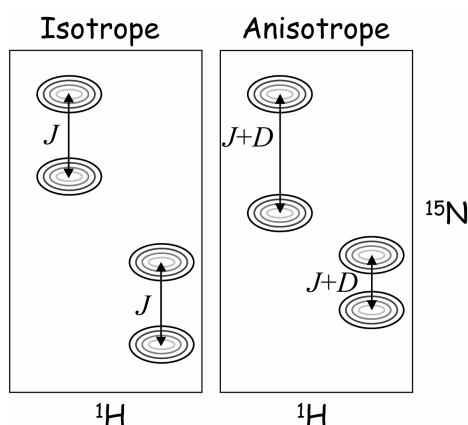


Figure 5.2 : Exemple de mesures des couplages dipolaires résiduels.

D'après l'Equation 5.4, les RDCs sont dépendants de cinq paramètres : r_{PQ} , θ , ϕ , A_a et A_r . Ils fournissent une information orientationnelle dégénérée ; il existe un ensemble de solutions pour une mesure de RDC (Figure 5.3). La mesure de plusieurs RDCs dans un même plan peptidique permet de réduire le nombre de solutions. Mais, compte tenu de la périodicité de la fonction angulaire de l'Equation 5.4, il existe ainsi huit solutions possibles pour un ensemble de RDCs mesurés dans un même plan peptidique et quatre pour un motif chiral (Hus *et al*, 2001). La mesure de RDCs, dans différents milieux d'alignement, permet de lever cette dégénérescence (Figure 5.3), en supposant que le changement de milieu d'alignement n'induit pas de modification sur la structure de la protéine.

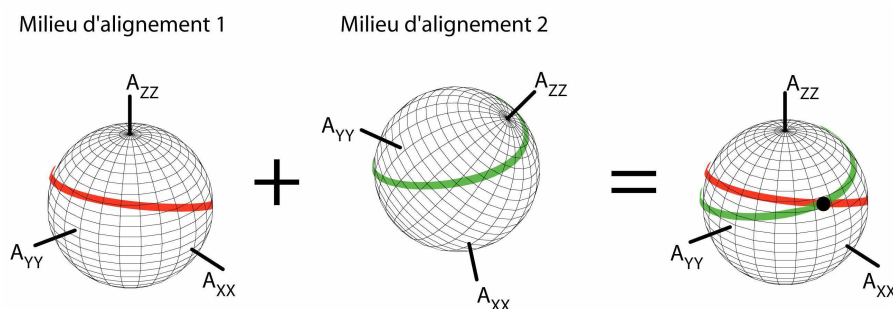


Figure 5.3 : Orientation d'un vecteur H-N en accord avec les valeurs de couplages dipolaires mesurés dans deux milieux d'alignement.

La ligne rouge (verte) correspond au couplage dipolaire mesuré dans le milieu d'alignement 1 (2).

5.1.2 Les RDCs fournissent des informations orientationnelles

Les couplages résiduels dipolaires (RDCs) peuvent être mesurés dans les protéines qui sont faiblement alignées par rapport au champ magnétique, à la suite de leur anisotropie de susceptibilité magnétique intrinsèque ou par l'utilisation de milieu anisotrope. Au fur et à mesure des développements concernant l'acquisition des RDCs, l'extraction de l'information structurale à partir de ces nouvelles données n'a cessé de s'accroître (de Alba *et al*, 2002). Les RDCs mesurés entre paires d'atomes dont la distance reste fixe, nous renseignent sur l'orientation du vecteur internucléaire. Ce type d'information orientationnelle est généralement mesuré entre deux atomes séparés par une liaison covalente. Le couplage dipolaire $^1D_{H-N}$ entre le proton amide et son azote a été le premier mesuré (Tolman *et al*, 1996) et de la même

manière, le couplage dipolaire $^1D_{H-C}$ a été détecté (Ottiger et al, 1998) sur des protéines alignées. Les couplages dipolaires entre deux hétéroatomes ont aussi été considérés ($^1D_{N-C'}$ ou $^1D_{C\alpha-C'}$ (Permi *et al*, 2000), $^1D_{C\alpha-C\beta}$, par exemple), même s'ils sont de 5 à 8 ordre de magnitude plus faible que le couplage dipolaire NH, dû aux différences de rapports gyromagnétiques. Toutefois, les couplages dipolaires ne sont pas limités à des paires d'atomes liés par une liaison covalente et des couplages entre atomes séparés par deux liaisons ou plus ont été mesuré : $^2D_{H-H}$ dans les groupements CH_2 (Miclet *et al*, 2004 ; 2005 ; Carlomagno *et al*, 2000) par exemple, ou différents couplages appartenant au même plan peptidique ($^2D_{C'-HN}$, $^2D_{C\alpha-HN}$, ...) (Brutscher, 2001 ; Prestegard *et al*, 2004), mais aussi dans le plan des bases nucléiques ($^3D_{H-H}$, $^2D_{H-C}$) (Boisbouvier *et al*, 2003).

Fournissant des informations orientationnelles par rapport à un repère lié à la molécule, les RDCs permettent de palier le manque de précision des modèles obtenus avec les paramètres semi-quantitatif standard que sont les NOEs. Elles sont utilisées pour affiner les modèles structuraux des protéines (Tjandra *et al*, 1997b ; Cornilescu *et al*, 1998 ; Clore *et al*, 1999), des acides nucléiques (Mollova *et al*, 2000 ; Bayer *et al*, 1999 ; Sibille *et al*, 2001), de systèmes de grandes tailles, telles que la Malate Synthase G (Tugarinov *et al*, 2005a ; Grishaev *et al*, 2008) et l'hémoglobine (Lukin *et al*, 2003). Elles permettent également de valider le repliement de structure (Annala *et al*, 1999). Dans des systèmes de taille modérée, pouvant être alignés dans plusieurs milieux liquide cristallin, la détermination simultanément du repliement et de la dynamique par les RDCs est possible (Bouvignies *et al*, 2006b ; 2007) mais aussi le repliement global en absence d'autres informations structurales (Hus *et al*, 2001 ; Delaglio *et al*, 2000).

5.1.3 Les RDCs 1H - 1H : une source d'information translationnelles

La RMN en milieu cristallin a également été utilisée pour détecter des couplages à travers l'espace, c'est-à-dire entre paires d'atomes dont la distance est inconnue. Plusieurs travaux ont permis de démontrer, la possibilité de détecter de tels RDCs entre des protons et des hétéroatomes que ce soit entre protons des sucres et le phosphore dans les acides nucléiques (Wu *et al*, 2001a ; Hennig *et al*, 2001), à

travers les liaisons hydrogènes (Wu *et al*, 2001b) ou entre protons et carbones à travers l'espace (Meier *et al*, 2003)

En raison du rapport gyromagnétique élevé du proton, les couplages ^1H - ^1H sont en soi les plus intenses. En effet, les RDCs ^1H - ^1H ont été considérés dès le début de l'utilisation des RDCs dans les protéines (Tjandra *et al*, 1997a). Différentes séquences d'impulsions standard permettent de détecter le transfert d'aimantation par l'intermédiaire de RDCs et de les quantifier. Ces séquences d'impulsions sont de type COSY (Tian *et al*, 1999 ; 2000 ; Wu *et al*, 2001c), de type E-COSY (Cai *et al*, 1999), TOCSY (Hansen *et al*, 1998) ou se combine avec un transfert de type NOE pour générer des pics de corrélation (Peti *et al*, 2000).

Malgré l'ensemble des séquences d'impulsion proposé, la mesure de RDCs ^1H - ^1H reste limitée à des protons séparés par des distances inférieures à 5 Å, car la plupart des transferts de cohérences se font vers les protons les plus proches. En effet, dans les protéines, un proton est entouré en moyenne de 25 autres protons dans un rayon de 5 Å. De ce fait, si l'on veut transférer de l'aimantation du proton P vers un proton Q éloigné, le transfert du couplage D_{PQ} va être atténué par l'ensemble des couplages dipolaires intenses avec les atomes proches de P.

Une des solutions pour résoudre ce problème est de limiter le transfert entre un groupe d'atomes. Ceci a été introduit en 2002 par Bax *et Coll.* en limitant le transfert entre protons amide par l'utilisation combinés (ou non) d'impulsions sélectives et de la perdeutération. L'élimination des interactions avec les plus proches voisins a permis, dans des protéines deutérées, la mesure précise de couplages directs entre $^1\text{H}_\text{N}$ - $^1\text{H}_\text{N}$ séparés jusqu'à 7 Å (Wu *et al*, 2002 ; Meier *et al*, 2003). Dans le prochain paragraphe, je présenterai le principe de l'expérience Semi-Selective-HMQC permettant la détection et la quantification de RDCs entre H_N . Ainsi que les améliorations qui ont été apporté à cette expérience pour la détection rapide de ces informations translationnelles.

5.2 La Détection rapide des RDCs H_N-H_N dans les protéines protonées

5.2.1 L'expérience Semi-Selective-HMQC (SS-HMQC)

5.2.1.1 Le Principe du transfert de cohérence

À la différence des expériences pour mesurer des couplages entre atomes séparés par une ou deux liaisons, les RDCs $^1H-^1H$ à longue portée sont mesurés par une expérience de corrélation J quantitative (Bax *et al*, 1994), où la partie de l'aimantation qui est transférée à partir du proton H^P vers le proton H^Q produit un pic de corrélation détecté aux fréquences (ω_P, ω_Q) et la partie non transférée, un pic diagonal à la fréquence de H^P dans le spectre 2D final.

La Figure 5.4 présente la séquence de base permettant le transfert d'aimantation entre deux protons amides (H_N) et la quantification des couplages dipolaire résiduels $^1H-^1H$ entre ces protons. Elle est composée de deux périodes d'évolution T_A et T_B séparées par une impulsion 90° , qui a la même fonction que l'impulsion 90° dans une expérience COSY classique. Tout d'abord pour comprendre le principe de la séquence, considérons un ensemble de protons dans un milieu anisotrope et intéressons nous au transfert de cohérence d'un proton H^P vers un proton H^Q .

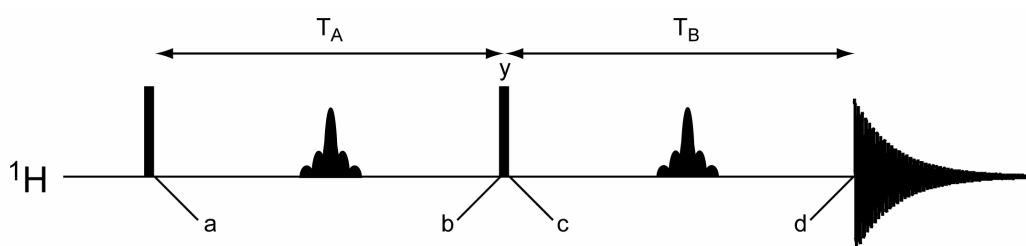


Figure 5.4 : La séquence d'impulsion pour la mesure de couplage dipolaire $^1H-^1H$ quantitative.

Les points a, b, c et d correspondent à des repères pour le texte. Les impulsions sélectives situées au centre des deux périodes de transfert, T_A et T_B , refocalisent l'évolution entre des couplages entre les H_N et les autres protons. Seuls les couplages entre deux protons amides simultanément inversés par cette impulsion donne lieu à une évolution.

Par souci de clarté, les termes liés à l'évolution du déplacement chimique sont omis. Après une première impulsion (a), l'aimantation initiale H_y^P évolue durant la période de transfert (T_A). L'impulsion sélective située au centre de T_A refocalise

l'évolution entre des couplages entre les H_N et les autres protons. Seuls les couplages entre deux protons H_N simultanément inversés par cette impulsion ne sont pas refocalisé. En (b), on a :

$$H_y^P \xrightarrow{T_A} \begin{cases} H_y^P \prod_j \cos(\pi D_{H^P H^j} T_A) e^{(-T_A \times R_2(H^P))} & [I] \\ 2H_x^P H_z^Q \sin(\pi D_{H^P H^Q} T_A) \prod_{j \neq Q} \cos(\pi D_{H^P H^j} T_A) e^{(-T_A \times R_2(H^P))} & [II] \end{cases}$$

où D_{HH} correspond au couplage dipolaire résiduel entre deux protons. L'impulsion 90° 1H suivante, (c), convertie le terme [II] en $-2H_z^P H_x^Q$ modulé par les fonctions dépendant de T_A , alors que le terme [I] n'est pas affecté. Durant la seconde période de transfert (T_B) une partie de ce terme est déphasée en (d):

$$[I] \xrightarrow{T_B} \begin{cases} H_y^P \prod_j \cos(\pi D_{H^P H^j} T_A) \prod_j \cos(\pi D_{H^P H^j} T_B) e^{(-(T_A+T_B) \times R_2(H^P))} & [III] \\ 2H_x^P H_z^Q \sin(\pi D_{H^P H^Q} T_B) \prod_j \cos(\pi D_{H^P H^j} T_A) \prod_{j \neq Q} \cos(\pi D_{H^P H^j} T_B) e^{(-(T_A+T_B) \times R_2(H^P))} & [IV] \end{cases}$$

et une partie du terme $-2H_z^P H_x^Q$ provenant de [II] est refocalisée en (d):

$$[II] \xrightarrow{T_B} \begin{cases} -2H_z^P H_x^Q \sin(\pi D_{H^P H^Q} T_A) \prod_{j \neq Q} \cos(\pi D_{H^P H^j} T_A) \\ \quad \times \prod_j \cos(\pi D_{H^Q H^j} T_B) e^{(-T_A \times R_2(H^P))} e^{(-T_B \times R_2(H^Q))} & [V] \\ H_y^Q \sin(\pi D_{H^P H^Q} T_A) \sin(\pi D_{H^P H^Q} T_B) \prod_{j \neq Q} \cos(\pi D_{H^P H^j} T_A) \\ \quad \times \prod_{j \neq P} \cos(\pi D_{H^Q H^j} T_B) e^{(-T_A \times R_2(H^P))} e^{(-T_B \times R_2(H^Q))} & [VI] \end{cases}$$

Les termes [III] et [VI] vont correspondre respectivement aux pics diagonal et de corrélation. Les deux autres termes [IV] et [V] donnent lieu à des termes en antiphasé qui peuvent être éliminés. De la même manière, il est possible d'obtenir l'évolution du signal pour le transfert de cohérence d'un proton Q vers un proton P en changeant les indices P et Q dans les termes [I] à [VI].

5.2.1.2 Quantification des RDCs 1H - 1H

Si pour deux protons P et Q, l'intensité du pic de corrélation correspondant au transfert de cohérence de P vers Q est I^{PQ} , le rapport d'intensité I^{PQ}/I^{PP} n'est pas

suffisant pour déterminer la valeur du RDC ^1H - ^1H entre P et Q. En effet, ce rapport ne tient pas compte de l'évolution de la relaxation transversale et des cohérences transversales concernant les spins P et Q pendant la durée de l'expérience. Pour être indépendant de ces termes de relaxation, l'extraction des valeurs des RDCs ^1H - ^1H entre méthyles est obtenue à partir des mesures d'intensité des 2 pics diagonaux (I^{PP} et I^{QQ}) et des 2 pics de corrélations correspondant (I^{PQ} et I^{QP}) en utilisant la relation suivante (Equation 5.5):

$$\tan^2(\pi D_{\text{H}^{\text{P}}\text{H}^{\text{Q}}} T_{\text{A}}) \tan^2(\pi D_{\text{H}^{\text{P}}\text{H}^{\text{Q}}} T_{\text{B}}) = \frac{I^{\text{PQ}} I^{\text{QP}}}{I^{\text{QQ}} I^{\text{PP}}} \quad (5.5)$$

Comme il a été présenté auparavant, en découplant des interactions entre protons de manière sélective et en combinant à la deutération des protons non-échangeables, des couplages directs entre $^1\text{H}_{\text{N}}$ - $^1\text{H}_{\text{N}}$ séparés jusqu'à 7 Å ont été mesurés précisément dans des protéines deutérées (Wu *et al*, 2002 ; Meier *et al*, 2003). En contraste avec les interactions NOEs, qui dans les protéines perdeutérées peuvent aussi être observées sur de telles distances (Venters *et al*, 1995 ; Mal *et al*, 1998, Koharudin *et al*, 2003), les couplages dipolaires ne sont pas affectés par la diffusion de spin et leur interprétation de manière structurale est donc directe (Tian *et al*, 2000).

5.2.1.3 La séquence d'impulsion SS-HMQC2

Dans les paragraphes précédents (§ 5.2.1.1 et § 5.2.1.2), nous avons négligé l'édition des fréquences pour un souci de clarté. Comme nous venons de le voir, la quantification des RDCs nécessite la mesure de l'intensité des pics diagonaux comme des pics de corrélation croisée. Une expérience à deux dimensions protons n'est donc pas exploitable dans les protéines, compte tenu du recouvrement des pics diagonaux dans les spectres. Alternativement, Wu *et coll.* ont proposé d'utiliser les temps T_{A} et T_{B} pour créer une cohérence multiquantum avec l'azote afin d'éditer la fréquence de l'azote pendant les temps de transfert T_{A} et T_{B} . Ainsi un pic diagonal sera caractérisé par un triplet de fréquence ($\omega_{\text{N}}^{\text{P}}$, $\omega_{\text{N}}^{\text{P}}$, $\omega_{\text{H}}^{\text{P}}$) et le pic de corrélation par ($\omega_{\text{N}}^{\text{P}}$, $\omega_{\text{N}}^{\text{Q}}$, $\omega_{\text{H}}^{\text{Q}}$). Cette expérience permet ainsi de quantifier l'intensité des pics diagonaux avec la résolution typique d'une expérience de corrélation 2D ^1H - ^{15}N et

d'extraire automatiquement des bandes pour chaque résidu avec l'ensemble des transferts interprotons détectés.

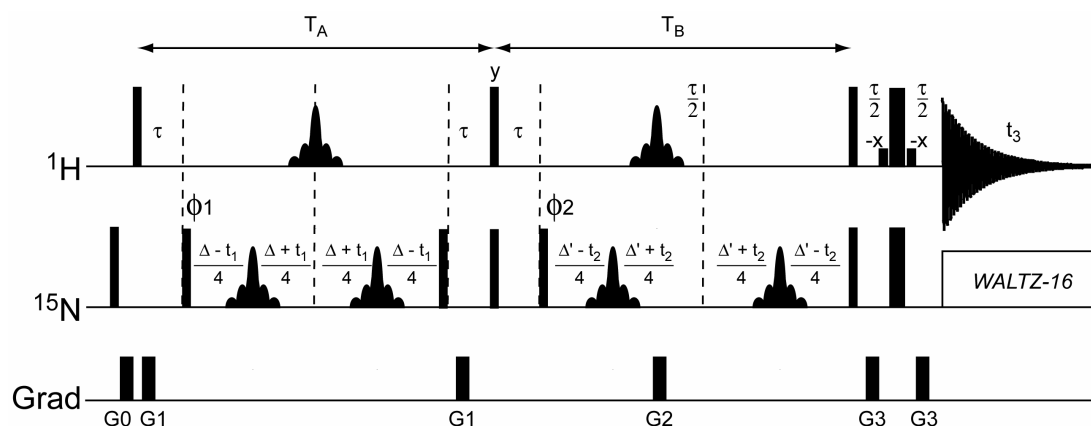


Figure 5.5 : Schéma de la séquence d'impulsion SS-HMQC2.

Les impulsions rectangulaires fines et larges sont appliquées avec des angles d'impulsion respectivement de 90° et de 180° . À moins que cela ne soit indiqué, les impulsions sont appliquées le long de l'axe X. Les quatre impulsions sélective ^{15}N sont de type 180° REBURP ainsi que l'impulsion sélective ^1H (Wu *et al.*, 2002). T_A et T_B sont les deux périodes de transfert de cohérences entre protons amides. $\Delta = T_A - 2\tau$ et $\Delta' = T_B - \tau$.

En collaboration avec Paul Schanda et Ewen Lescop, nous avons apporté des améliorations à la séquence d'impulsion SS-HMQC2 proposée par Wu et *coll.* (Wu *et al.*, 2002) (Figure 5.5). L'optimisation se fait en terme d'efficacité de relaxation longitudinale et de compensation du transfert de cohérence due à la disparité de couplage J. Cette nouvelle séquence permet la détection des RDCS H_N-H_N dans les protéines protonées de tailles modérées avec un gain considérable en sensibilité.

5.2.2 Diminution du temps de relaxation longitudinale

Dans les molécules qui ont un temps de corrélation de plusieurs nanosecondes, telles que les protéines, l'efficacité de la relaxation longitudinale ^1H est surtout déterminée par le réseau de spins ^1H dipolaire (^1H - ^1H NOEs). En effet, un spin excité à proximité d'un autre spin peut transférer une partie de son énergie à ce spin afin de revenir lui-même à son état d'équilibre. Ainsi, lors d'une expérience RMN, si seuls les protons amides (H_N) sont excités (les autres protons restant dans leur état d'équilibre), la vitesse de relaxation longitudinale effective des H_N est plus rapide, permettant alors l'utilisation de temps de récupération plus court. Ce phénomène est à la base de l'expérience SOFAST-HMQC (Schanda *et al*, 2005 ; Schanda, 2007). Ceci est obtenu en n'utilisant que des impulsions d'excitation sélectives sur les H_N et en

prenant soin que le nombre total d'impulsions d'inversion de l'aimantation soit toujours paire afin de ramener les spins que l'on ne désire pas exciter sur l'axe +Z à la fin de la séquence d'impulsion.

L'efficacité de la relaxation spin-réseau des protons amides labiles accessible par le solvant est aussi sous l'influence de l'échange chimique avec les hydrogènes de l'eau. La saturation partielle des protons de l'eau traduit, via ce mécanisme d'échange, à réduire la polarisation d'amides à l'état d'équilibre échangeant rapidement au début de chaque accumulation. La relaxation longitudinale est de nouveau améliorée quand l'eau est maintenue dans son état d'équilibre (Figure 5.6). Cette augmentation apparente de la vitesse de relaxation longitudinale permet de diminuer le temps de récupération entre chaque répétition de l'expérience et augmente ainsi la sensibilité de l'expérience pour une durée donnée (Figure 5.6). L'efficacité de la relaxation longitudinale a récemment été exploitée avec succès sur les protéines dans le contexte d'expériences d'attribution des résonances (Atreya *et al*, 2004 ; Diercks *et al*, 2005 ; Pervushin *et al*, 2002 ; Schanda *et al*, 2006) et l'acquisition de données ultrarapide à deux dimension (Schanda *et al*, 2005). Nous avons implémenté cette stratégie dans chacun des deux blocs HMQC de l'expérience présenté sur la Figure 5.5. Cette amélioration se traduit par un gain en sensibilité moyen de l'ordre de 2. Ces améliorations sont présentées dans l'article présenté au § 5.2.4.

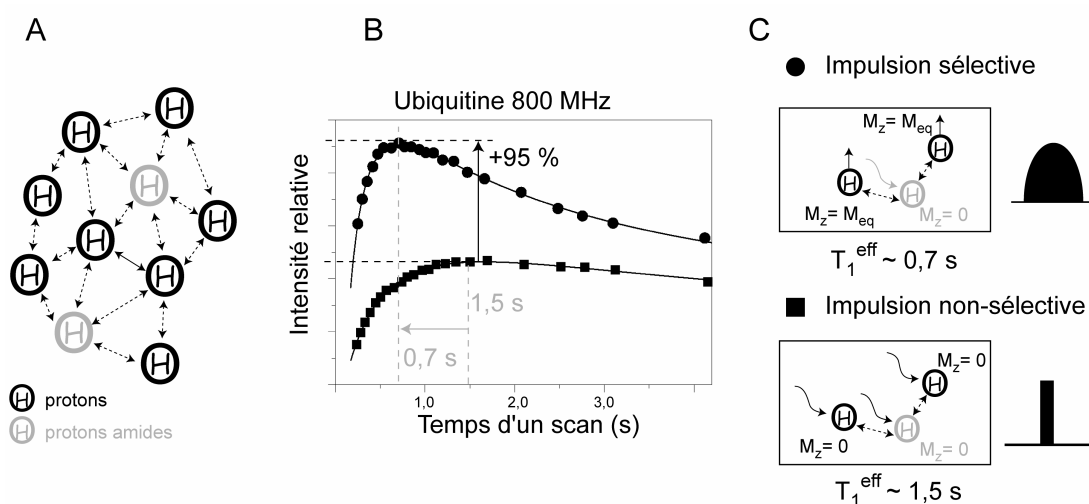


Figure 5.6 : A) Représentation schématique d'un réseau de spin proton. B) Intensité en fonction du temps d'une accumulation avec ou sans l'utilisation d'impulsion sélective dans la séquence d'impulsion SS-HMQC2. C) Schéma montrant l'efficacité de relaxation longitudinale en fonction d'utilisation ou non d'excitation sélective sur les protons amides.

5.2.3 Compensation de la disparité de couplage

Chacune des deux unités élémentaires de structure HMQC dans la séquence (Figure 5.5) exige un ajustement de deux délais de transfert $\tau = 1/(2J_{NH})$, où J_{NH} est la constante de couplage hétéronucléaire entre le 1H et le ^{15}N . En milieu isotrope, le couplage scalaire J peut être considéré de manière uniforme le long de la chaîne polypeptide, c'est-à-dire que d'un site amide à un autre, les variations sont faibles. Dans les molécules faiblement alignées, le couplage dipolaire résiduel s'ajoute au couplage J résultant dans la constante de couplage effectif $J_{NH}^{eff} = J_{NH} + D_{NH}$ qui couvrent une gamme allant typiquement $\pm 20-40\%$ par rapport à la valeur moyenne du couplage. Il n'est donc plus possible de régler les délais pour tous les sites amide simultanément, une perte de sensibilité s'ensuit qui dépend du degré de disparité de couplage J^{eff} .

$$f_{NH}^{Jmis} = 2\tau J_{NH}^{eff} \quad (5.6)$$

L'utilisation de schémas particulier "d'INEPT large bande" (Nielsen *et al*, 1989 ; Wimperis *et al*, 1986) permet de récupérer une partie de l'aimantation "perdu" en créant un deuxième chemin de cohérence qui s'ajoute aux signal RMN détecté (Figure 5.7). Ceci permet ainsi de compenser en partie la perte due à la disparité du couplage J . Compte tenu que l'expérience développé pour la détection des RDCs H_N-H_N utilise quatre transferts du type INEPT (ou équivalent), l'utilisation de ce transfert compensé dans les protéines partiellement alignées apporte un gain substantiel à l'expérience SS-HMQC2.

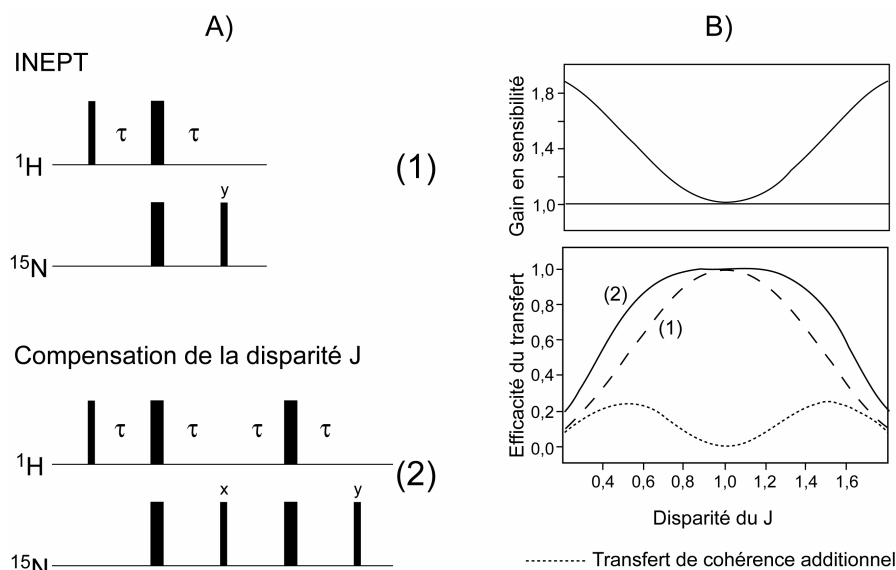


Figure 5.7 : A) Schéma d'impulsion INEPT et de compensation de disparité des couplages. B) En haut, le gain de signal pour un bloc HMQC en fonction de la compensation de disparité des couplages est représenté. En dessous, les simulations de l'efficacité de transfert ^1H - ^{15}N pour une expérience de type HMQC.

Les courbes continue (2) et discontinue (1) représentent respectivement la simulation avec et sans l'ajout de l'élément de la compensation de la disparité J. Le deuxième chemin de cohérence est représenté en pointillé.

5.2.4 Expérience optimisé pour la mesure de couplage dipolaire résiduel entre les protons amides

Ces améliorations ont été implémentées dans le schéma de la séquence d'impulsion (Wu *et al*, 2002) et ont permis d'améliorer la sensibilité d'un facteur 1,5 à 8 pour des pics individuels de corrélations en comparaison avec les séquences d'impulsions proposés précédemment (Wu *et al*, 2002). Cette approche permet ainsi d'accélérer l'expérience pour la mesure des RDCs ^1H - ^1H à longue portée par un facteur de l'ordre de 10. De plus ces informations à longue portée sont des contraintes importantes dans la détermination de structure des protéines protonées de petites tailles et de tailles modérées. Cette information est directement utilisable après l'attribution de la chaîne principale sans informations ni sur les chaînes latérales ni l'analyse de NOEs. Enfin, l'utilisation de ces RDCs est particulièrement intéressante dans le contexte de l'élucidation rapide du repliement des protéines (Bouvignies *et al*, 2006a).

J Biomol NMR (2007) 38:47–55
 DOI 10.1007/s10858-006-9138-2

ARTICLE

Sensitivity-optimized experiment for the measurement of residual dipolar couplings between amide protons

Paul Schanda · Ewen Lescop · Mirjam Falge ·
 Rémy Sounier · Jérôme Boissbouvier ·
 Bernhard Brutscher

Received: 26 October 2006 / Accepted: 20 December 2006 / Published online: 8 March 2007
 © Springer Science+Business Media B.V. 2007

Abstract High signal to noise is a necessity for the quantification of NMR spectral parameters to be translated into accurate and precise restraints on protein structure and dynamics. An important source of long-range structural information is obtained from ^1H – ^1H residual dipolar couplings (RDCs) measured for weakly aligned molecules. For sensitivity reasons, such measurements are generally performed on highly deuterated protein samples. Here we show that high sensitivity is also obtained for protonated protein samples if the pulse schemes are optimized in terms of longitudinal relaxation efficiency and J-mismatch compensated coherence transfer. The new sensitivity-optimized quantitative J-correlation experiment yields important signal gains reaching factors of 1.5 to 8 for individual correlation peaks when compared to previously proposed pulse schemes.

Keywords Protein · Structure · Residual dipolar couplings · Fast NMR · HMQC · J-mismatch compensation · Longitudinal relaxation enhancement

Introduction

Residual dipolar couplings (RDCs), measured in weakly aligned molecules in solution, provide an important source of structural information (Tjandra and Bax 1997; Prestegard et al. 2004; Blackledge 2005) that complement or replace more classical NMR data such as nuclear Overhauser effects (nOes) and scalar coupling constants (J). Whereas nOe-based distance restraints are affected by indirect spin interactions (spin diffusion), and J-coupling-derived torsion angles rely on empirical Karplus curves, RDCs directly reflect the interaction strength between two nuclear spins that can be translated into structural restraints of high accuracy using a well-known analytical formula. RDCs therefore provide access to more precise definition of local and long-range molecular structure. It has also been shown that RDCs, measured in several alignment media, are particularly attractive for RDC-based *de novo* protein fold determination (Kontaxis et al. 2005; Hus et al. 2001; Beraud et al. 2002), and the characterization of molecular dynamics at an atomic level occurring at time scales of pico- to milliseconds (Meiler et al. 2001; Bouvignies et al. 2005; Lakomek et al. 2005).

Nowadays RDCs are routinely measured between covalently attached nuclei, e.g. N–H, C–H, N–C, and C–C, mainly in the protein backbone, as part of the standard data set recorded for NMR structure elucidation. The distance and orientational information provided by long-range RDCs between more distant nuclei is still less exploited. It has been shown recently for proteins and nucleic acids that ^1H – ^1H RDCs can be accurately measured between protons separated by more than 7 Å, provided that the ^1H spin coupling

Paul Schanda and Ewen Lescop contributed equally to this work.

P. Schanda · E. Lescop · M. Falge · R. Sounier ·
 J. Boissbouvier · B. Brutscher (✉)
 Institut de Biologie Structurale – Jean-Pierre Ebel,
 UMR5075 CNRS-CEA-UJF, 41, rue Jules Horowitz, 38027
 Grenoble Cedex 1, France
 e-mail: bernhard.brutscher@ibs.fr

J. Boissbouvier
 e-mail: jerome.boissbouvier@ibs.fr

network in these molecules is simplified by means of either deuteration, band-selective homonuclear decoupling, or both (Boisbouvier et al. 2003; Meier et al. 2003; Wu and Bax 2002). Here, we will focus on the measurement of long-range RDCs between amide protons in small to medium sized proteins. The sensitivity of the experimental schemes proposed so far greatly benefits from perdeuteration followed by back-protonation of the labile (amide) hydrogen sites. Most NMR experiments, commonly used for the study of small to medium-sized proteins, do not require perdeuteration. In order to make amide ^1H – ^1H RDC measurements more attractive for fully protonated (or only partially deuterated) samples we have developed a new pulse sequence that provides greatly improved sensitivity over existing methods. Signal enhancements, ranging from a factor of 1.5 to 8, are observed for individual correlation peaks, as illustrated for ^{15}N labeled ubiquitin under weak alignment conditions.

Sensitivity-enhanced double ^{15}N -edited ^1H – ^1H correlation experiment

In contrast to experiments for the measurement of one- or two bond couplings, long-range ^1H – ^1H RDCs are most conveniently measured by a quantitative J-correlation experiment (Bax et al. 1994), where part of the magnetization is transferred from proton H^{A} to proton H^{B} yielding a cross peak at the resonance frequency of H^{B} and a diagonal peak at the frequency of H^{A} in the final spectrum. The spin coupling constant is then easily retrieved from the ratio of the measured intensities of cross- and diagonal peaks (see below). The new pulse sequence of Fig. 1a is a modified version of the Semi-Selective HMQC2 (SS-HMQC2) experiment (Wu and Bax 2002) that consists of two successive HMQC blocks sandwiching a COSY-type ^1H mixing pulse. In the following we will refer to our new experiment as BEST-Jcomp-HMQC2 (BEST: Band-selective-Excitation Short-Transient, Jcomp: J-mismatch compensated). Before describing in more detail the new features of BEST-Jcomp-HMQC2, we will briefly summarize the relevant coherence transfer pathways common to the double-HMQC experiment. Amide ^1H chemical shift, scalar and dipolar coupling evolutions with aliphatic protons are refocused by the band-selective 180° ^1H pulses applied in the middle of each HMQC building block. In contrast, H^{N} – H^{N} RDCs evolve during the transfer delay Δ_1 as $\text{H}^{\text{A}}_{\text{y}} \rightarrow \text{H}^{\text{A}}_{\text{y}} \cos(\pi D_{\text{HH}} \Delta_1) + 2\text{H}^{\text{A}}_{\text{x}} \text{H}^{\text{B}}_{\text{z}} \sin(\pi D_{\text{HH}} \Delta_1)$, where D_{HH} is the residual dipolar coupling constant between spins H^{A} and H^{B} and interactions with other amide

protons are neglected. The following 90° ^1H pulse then converts $2\text{H}^{\text{A}}_{\text{x}} \text{H}^{\text{B}}_{\text{z}}$ to $-2\text{H}^{\text{A}}_{\text{z}} \text{H}^{\text{B}}_{\text{x}}$ while leaving $\text{H}^{\text{A}}_{\text{y}}$ (which gives rise to the diagonal peak) unaffected. During the subsequent Δ_2 transfer period, the anti-phase coherence $-2\text{H}^{\text{A}}_{\text{z}} \text{H}^{\text{B}}_{\text{x}}$ partially refocuses to $\text{H}^{\text{B}}_{\text{y}} \sin(\pi D_{\text{HH}} \Delta_2)$ that is then detected as a cross peak. Remaining $2\text{H}^{\text{A}}_{\text{z}} \text{H}^{\text{B}}_{\text{x}}$ and $2\text{H}^{\text{A}}_{\text{x}} \text{H}^{\text{B}}_{\text{z}}$ coherences that are apparent as dispersive antiphase signals in the ^1H (ω_3) dimension of the spectrum can be suppressed by an additional z-filter element that will be explained in more detail below. Contrary to the experiment of Wu and Bax (2002), where ^1H antiphase suppression is an integral part of the pulse scheme, removing the z-filter element from the sequence of Fig. 1a does not alter the main coherence transfer pathways (see below). Therefore the experiment may also be performed without z-filter. During both ^1H – ^1H dephasing times Δ_1 and Δ_2 , the ^{15}N chemical shift of the covalently attached amide is labeled using HMQC-type ^1H – ^{15}N correlation schemes. The residual dipolar coupling $D = D_{\text{HH}}$ is then obtained from the 4 peak intensities measured at positions $(\omega_1, \omega_2, \omega_3)$: $I_{\text{cross}}^{\text{A}}(\omega_{\text{N}}^{\text{A}}, \omega_{\text{N}}^{\text{B}}, \omega_{\text{H}}^{\text{H}})$, $I_{\text{diag}}^{\text{A}}(\omega_{\text{N}}^{\text{A}}, \omega_{\text{N}}^{\text{A}}, \omega_{\text{H}}^{\text{H}})$, $I_{\text{cross}}^{\text{B}}(\omega_{\text{N}}^{\text{B}}, \omega_{\text{N}}^{\text{A}}, \omega_{\text{H}}^{\text{H}})$, $I_{\text{diag}}^{\text{B}}(\omega_{\text{N}}^{\text{B}}, \omega_{\text{N}}^{\text{B}}, \omega_{\text{H}}^{\text{H}})$ using the following relation:

$$\frac{I_{\text{cross}}^{\text{A}} I_{\text{cross}}^{\text{B}}}{I_{\text{diag}}^{\text{A}} I_{\text{diag}}^{\text{B}}} = \frac{\lambda^2 \sin^2(\pi D \Delta_1) \sin^2(\pi D \Delta_2)}{\{\lambda^2 \cos^2(\pi D \Delta_1) \cos^2(\pi D \Delta_2) + 2\lambda(1-\lambda) \cos(\pi D \Delta_1) \cos(\pi D \Delta_2) + (1-\lambda)^2\}} \quad (1)$$

with λ a correction factor taking into account the protonation level at the amide sites in the protein (Wu and Bax 2002), e.g. $\lambda = 0.9$ for a 10%/90% $\text{D}_2\text{O}/\text{H}_2\text{O}$ mixture.

Solving Eq. [1] for a measured intensity ratio yields the magnitude of the D_{HH} coupling constant, given by

$$|D_{\text{HH}}| = \frac{\gamma_{\text{H}}^2 \mu_0 \hbar}{16\pi^3 r_{\text{HH}}^3} \left| A_{\text{a}} (3 \cos^2 \theta - 1) + \frac{3}{2} A_{\text{r}} \sin^2 \theta \cos 2\varphi \right| \quad (2)$$

with A_{a} and A_{r} the axial and rhombic components of the alignment tensor, θ and φ the polar angles defining the H^{N} – H^{N} vector orientation with respect to the alignment frame, and r_{HH} the inter-proton distance.

The sensitivity of BEST-Jcomp-HMQC2 greatly benefits from two new features, longitudinal relaxation enhancement and J-mismatch compensation, that will be discussed in detail in the following.

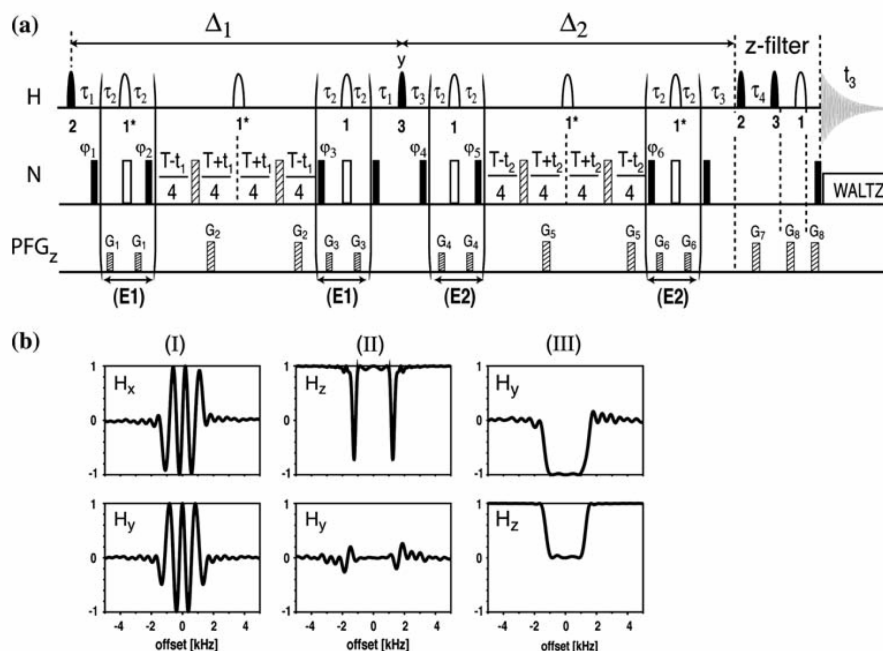


Fig. 1 (a) Pulse sequence of the 3D BEST-Jcomp-HMQC2 experiment for the measurement of amide ^1H - ^1H residual dipolar couplings. The inserts labeled (E1) and (E2) correspond to J-mismatch compensation elements as explained in the text. The optional z-filter at the end suppresses $2\text{H}_x\text{H}_z$ -type coherence prior to detection. All radio-frequency (rf) pulses are applied along the x -axis unless indicated. 90° and 180° rf pulses are represented by filled and open pulse symbols, respectively. The dashed ^{15}N pulses are broadband inversion pulses applied with a BIP-360-30-5 shape (Smith et al. 2001). All shaped ^1H pulses are centered at 8.2 ppm covering a band width of 4.0 ppm corresponding to the amide ^1H spectral region. The following shapes are used: (1) REBURP, (2) time-reversed EBURP2 (EBURP*), and (3) EBURP2 (Geen and Freeman 1991). For pairs of successive 180° rotations (1*) ISNOB5 pulses (Kupce et al. 1995) instead of REBURP are applied because of their cleaner band-selective inversion profile. The transfer delays are adjusted to $\tau_1 = 1/(2J_{\text{NH}}) - \delta_1/2$, $\tau_2 = 1/(4J_{\text{HN}}) - \delta_2/2$, $\tau_3 = 1/(2J_{\text{NH}})$, $\tau_4 = 1/(2J_{\text{NH}}) - \delta_1$, and $T \approx 30$ – 60 ms, with δ_1 and δ_2 the lengths of

the EBURP2 and REBURP pulses, respectively. The transfer delays Δ_1 and Δ_2 take into account ^1H - ^1H coupling evolution during the selective pulses: $\Delta_1 = 2\tau_1 + 4\tau_2 + T$ and $\Delta_2 = 2\tau_3 + 4\tau_2 + T$. Pulsed field gradients, G_1 – G_8 are applied along the z -axis (PFG_z) with a duration of $200 \mu\text{s}$ and field strengths ranging from 5–40 G/cm. Phase cycling: $\phi_1 = x$, $\phi_2 = y$, $\phi_3 = 2y$, $\phi_4 = x$, $\phi_5 = 4y$, $\phi_6 = 4y$, and the receiver $\phi_{\text{rec}} = x$. Quadrature detection in the t_1 (t_2) dimension is obtained by simultaneous time-proportional phase incrementation of ϕ_1 (ϕ_4) and decrementation of ϕ_2 (ϕ_5) according to TPPI-STATES. The pulse sequence code (Varian) is available from the authors upon request. (b) Bloch simulations illustrating the effect of EBURP2 and time-reversed EBURP2 pulses in the pulse sequence (a). Coupling evolution has not been taken into account for the simulations. The plotted offset profiles represent the ^1H spin state after different pulse combinations starting from pure z -magnetization (H_z): (I) EBURP_x^* , (II) $\text{EBURP}_x^* - \Delta - 180^\circ - \Delta - \text{EBURP}_x$, and (III) $\text{EBURP}_x^* - \Delta - 180^\circ - \Delta - \text{EBURP}_y$.

Longitudinal relaxation enhancement

In slowly tumbling diamagnetic molecules such as proteins the efficiency of ^1H spin lattice relaxation is mainly determined by the dipolar-coupled ^1H spin network (^1H - ^1H nOes). Therefore, if only a subset of protons is observed in a particular NMR experiment, spin-lattice relaxation can be enhanced by manipulating the remaining proton spins during the pulse sequence in a way that their spin state before detection is close to thermal equilibrium. The effective longitudinal spin-lattice relaxation of labile solvent-accessible amide ^1H is also influenced by chemical exchange with

water hydrogens. Partial saturation of the water proton spins translates via this exchange mechanism to a reduced steady state polarization of the fast exchanging amides at the beginning of each scan. Again, longitudinal relaxation is enhanced when the water is in a relaxed state between subsequent scans. Shorter longitudinal relaxation times allow for higher repetition rates of the pulse sequence, and thus provide increased signal to noise (S/N) in a given amount of experimental time (higher sensitivity). Longitudinal relaxation enhancement has recently been successfully exploited for proteins in the context of resonance assignment experiments (Pervushin et al. 2002; Atreya and

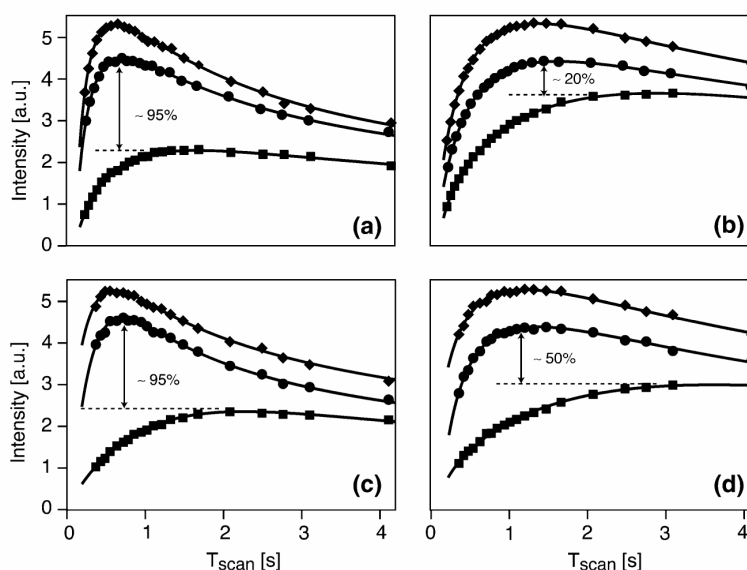
Szyperski 2004; Diercks et al. 2005; Schanda et al. 2006b), ultrafast two-dimensional data acquisition (Schanda and Brutscher 2005; Schanda et al. 2005; Schanda and Brutscher 2006), and fast characterization of structural compactness and heterogeneity of polypeptide chains (Schanda et al. 2006a).

In the pulse sequence of Fig. 1a, only shaped ^1H pulses are applied that selectively manipulate amide ^1H while leaving aliphatic ^1H mostly unaffected. The BEST concept, consisting of the use of amide ^1H -selective pulses only, has been shown to yield superior results in terms of leaving aliphatic ^1H polarization unperturbed compared to hard-pulse-based sequences using additional selective “flip-back” pulses (Schanda et al. 2006b). In the sequence of Fig. 1a, a single 180° ^1H pulse is applied with a REBURP shape, whereas for pairs of consecutive 180° ^1H pulses ISNOB5 pulse shapes are preferred because of their better off-resonance performance. A time-reversed EBURP2 pulse (EBURP2*), initially optimized for flip-back purposes, is used for ^1H excitation. As illustrated by the Bloch simulations in Fig. 1b, EBURP2* leaves chemical shift evolution (and spin-spin coupling) active during half of the pulse duration thereby creating a linear phase gradient. This phase gradient is refocused during the subsequent 90° pulse applied with an EBURP2 shape and a 90° phase shift with respect to the EBURP2* pulse. An inphase H_y coherence is created for the component that has not evolved under dipolar coupling. At the same time the EBURP2 creates inphase H_x coherence from H_z , and flips back orthogonal transverse coherence along the

z -axis (see Fig. 1b), thereby converting the antiphase coherence $2H_x^A H_z^B$ to $-2H_z^A H_x^B$. The combination EBURP2*– Δ – 180° – Δ –EBURP2 allows efficient use of the relatively long pulse durations (on the order of 2 ms), and avoids general rotation 90° pulses that are known to be more sensitive to B_1 -field inhomogeneities and pulse imperfections.

Figure 2 illustrates the effect of the BEST modifications on the sensitivity of the experiment for two protein samples at ^1H frequencies of 600 and 800 MHz: (a, c) 76-residue fully protonated ubiquitin and (b, d) a 167-residue fragment of the *Escherichia coli* sulfate reductase (SiR-FP18) deuterated at a level of $\sim 75\%$ (Sibille et al. 2005). 2D BEST-HMQC2 spectra were recorded without the J-mismatch compensation filters setting the effective evolution delays Δ_1 and Δ_2 to 26 ms. The average sensitivity (intensity per unit time) of BEST-HMQC2 is plotted as a function of the scan time (sum of pulse sequence duration and recovery delay), and compared to results obtained using the SS-HMQC2 sequence (Wu and Bax 2002) recorded under comparable conditions. As a result of the selective ^1H manipulation in BEST-HMQC2 the maximum in the sensitivity curves shifts towards shorter scan times, reflecting the significant decrease in effective longitudinal relaxation times. For protonated ubiquitin, scan times of ~ 600 ms are optimal, whereas for the highly deuterated sample of SiR-FP18 a scan time of ~ 1.5 s provides highest S/N ratios. These optimal repetition rates depend only little on the magnetic field strength. An average sensitivity gain of a factor of ~ 2 , and individual gains of up to a factor of 4 (data not

Fig. 2 Average S/N ratios per unit time (intensity) plotted as a function of the scan time (pulse sequence duration plus recycle delay) obtained with the BEST-Jcomp HMQC sequence without z-filter (diamonds), with additional z-filter (circles), and using the SS-HMQC2 sequence of Wu and Bax (squares) for fully protonated ubiquitin at (a) 600 MHz and (c) 800 MHz ^1H frequency, and for $\sim 75\%$ partially deuterated SiR-FP18 at (b) 600 MHz and (d) 800 MHz ^1H frequency. Intensities were obtained by adding the peak intensities of all detected amide (excluding NH_2) correlation peaks observed in the 2D spectra



shown) are observed for ubiquitin if the scan times of both experiments are optimized independently. Even for the deuterated sample of SiR-FP18, an average sensitivity gain of 20% at 600 MHz and 50% at 800 MHz is obtained. This observation can be explained by the fact that most methyl groups remain protonated at least at one of the three methyl hydrogen sites. ^1H – ^1H spin diffusion to methyl groups then still allows quite efficient energy dissipation at these sites of high mobility (methyl rotation). Additionally, and in contrast to the SS-HMQC2 experiment, water polarization is left unperturbed rather than dephased, which also contributes to faster effective relaxation of the exchangeable amide protons. Even higher sensitivity gains are observed in the fast pulsing regime when using short recycle delays for both SS-HMQC2 and BEST-HMQC2 experiments. This may be relevant for situations where short overall acquisition times become important, e.g. for sample stability reasons.

J-mismatch compensation

Each of the two HMQC building blocks in the sequence of Fig. 1a requires the adjustment of two transfer delays $\tau = 1/(2J_{\text{NH}})$, where J_{NH} is the heteronuclear coupling constant between amide ^1H and ^{15}N . While the scalar J couplings are quite uniform along the polypeptide chain with very little variation from one amide site to another, in weakly aligned molecules the RDC adds to the J coupling resulting in effective spin coupling constants $J_{\text{NH}}^{\text{eff}} = J_{\text{NH}} + D_{\text{NH}}$ that cover a range of typically ± 20 – 40% with respect to the mean value. It is therefore no longer possible to adjust the τ delays for all amide sites simultaneously, resulting in a sensitivity loss that depends on the degree of J-mismatch $f_{\text{NH}}^{\text{Jmis}} = 2\tau J_{\text{NH}}^{\text{eff}}$.

In a standard HMQC experiment the main coherence transfer pathway (P_I) is as follows:

$$H_y \xrightarrow{J_{\text{NH}}^{\text{eff}}} 2H_x N_y (t_1) \xrightarrow{J_{\text{NH}}^{\text{eff}}} H_y \sin^2(\pi f_{\text{NH}}^{\text{Jmis}}/2) \quad (3)$$

The efficiency of the heteronuclear transfer steps, given by $\sin^2(\pi f_{\text{NH}}^{\text{Jmis}}/2)$ neglecting spin relaxation, translates into a signal loss of $\sim 20\%$ for a J-mismatch of $f_{\text{NH}}^{\text{Jmis}} = 0.7$ or 1.3 (see Fig. 3). In order to increase the performance of the HMQC sequence for partially aligned protein samples, J-mismatch compensation elements (denoted (E1) and (E2) in the sequence of Fig. 1a) are inserted. The J-mismatch compensation technique used here is conceptually similar to previously proposed “broadband INEPT” schemes (Nielsen et al. 1989; Wimpey and Bodenhausen 1986).

J-mismatch compensation allows recovering part of the “lost” magnetization by creating a second coherence transfer pathway (P_{II}) that adds to the detected NMR signal:

$$H_y \xrightarrow{J_{\text{NH}}^{\text{eff}}} 2H_x N_x (t_1) \xrightarrow{J_{\text{NH}}^{\text{eff}}} H_y \cos^2(\pi f_{\text{NH}}^{\text{Jmis}}/2) \sin^2(\pi f_{\text{NH}}^{\text{Jmis}}/2) \quad (4)$$

The transfer efficiency in J-mismatch compensated HMQC is then given by the sum of the two pathways P_I and P_{II} . In the J-mismatch range $0.7 < f_{\text{NH}}^{\text{Jmis}} < 1.3$ the ^1H – ^{15}N transfer becomes almost independent of the effective coupling constant $J_{\text{NH}}^{\text{eff}}$ with transfer efficiencies of more than 0.95 over the whole range (Fig. 3). The amide ^1H – ^1H coupling evolution is not affected by these additional pulse sequence elements because no 90° ^1H pulse is required. The J-mismatch compensation elements can thus be inserted in the ^1H – ^1H coupling evolution periods of the BEST-HMQC blocks. Therefore, no additional delays are required and the improved heteronuclear transfer efficiency of the HMQC blocks directly translates into intensity gains for the cross- and diagonal peaks depending on the

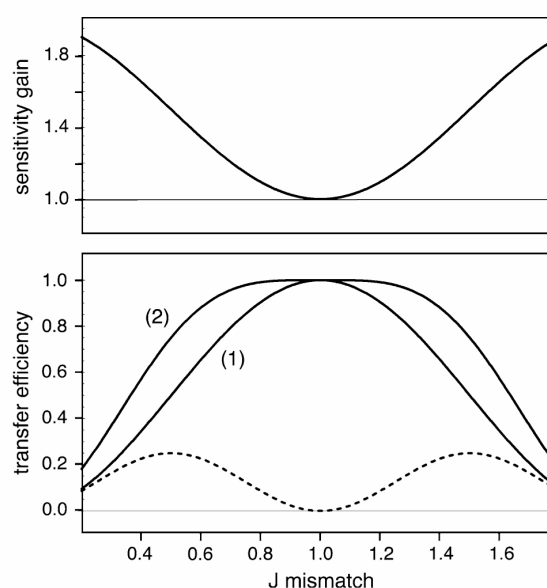


Fig. 3 Simulations of ^1H – ^{15}N transfer efficiencies (see Eqs. [3] and [4]) in HMQC-type experiments with (2) and without (1) the additional J-mismatch compensation elements of Fig. 1a. The transfer efficiency of the additional coherence transfer pathway (P_{II}) created by the J-mismatch compensation elements is plotted as a dashed line. In addition, the expected signal gain for a single HMQC block as a function of J-mismatch is shown on top

J-mismatch of the involved amide groups. The only price to pay for the J-mismatch compensation, besides a few additional pulses, is a reduction (by ~11 ms) in the available maximal t_1 and t_2 evolution times. This however does not present a strong limitation as long as transfer delays $\Delta_1, \Delta_2 > 30$ ms are chosen for the amide ^1H - ^1H coupling evolution.

Artefact suppression using an additional z-filter

To suppress undesired antiphase $2\text{H}_x\text{H}_x$ coherences before detection, while leaving inphase H_y coherence unperturbed, an optional z-filter element can be used. In short, the filter works as follows: the inphase H_y coherence is stored along the z -axis by a 90° EBURP2* pulse while antiphase $2\text{H}_x\text{H}_x$ coherence is converted to zero-quantum (ZQ) and double-quantum (DQ) coherences by the same pulse. The DQ coherence is destroyed by a pulsed field gradient. The ZQ coherence, that remains unaffected by the gradient pulse, evolves during a delay $1/2J_{\text{NH}}$ into $8\text{H}_x\text{H}_x\text{N}_z\text{N}_z$ -type coherence that is not detected in the presence of ^{15}N decoupling during data acquisition. The use of a high-performance ^{15}N composite decoupling sequence, such as WALTZ-16, is important to avoid partial detection of undesired coherence transfer pathways resulting in phase distortions in the ^1H dimension. The subsequent EBURP2 pulse then converts H_z to H_y for final detection. Finally, a WATERGATE sequence (Piotto et al. 1992) is added for water suppression purposes.

Experimental results

The performance of the BEST-Jcomp-HMQC2 experiment has been tested on a 1.25 mM sample of fully protonated ^{15}N -labeled ubiquitin. To create a weak alignment in the magnetic field, the ubiquitin sample was dissolved in a 15% w/v DMPC/DHPC mixture at a molar ratio of 3:1. The pH of the sample was adjusted to 6.6 in a 10 mM phosphate buffer. The resulting molecular alignment tensor, characterized on the basis of a set of D_{NH} coupling constants measured from IPAP-HSQC spectra (Andersson et al. 1998; Ottiger et al. 1998) and the ubiquitin solution structure (pdb entry 1D3Z), showed an axial component of $A_a = 17.8 \cdot 10^{-4}$ and a rhombic component of $A_r = 2.6 \cdot 10^{-4}$. Under these alignment conditions the J-mismatch for individual amide sites ranged from 0.6 to 1.3 assuming a scalar coupling constant $J_{\text{NH}} = 92$ Hz. All spectra were acquired at 800 MHz ^1H frequency

and 28°C sample temperature on a Varian INOVA spectrometer equipped with a cryogenic triple-resonance probe and shielded z-gradients. 3D spectra were recorded using different options of the pulse sequence displayed in Fig. 1a: with/without J-mismatch compensation, and with/without z-filter. In addition, a SS-HMQC2 spectrum (Wu and Bax 2002) was acquired for comparison using identical acquisition parameters except for the interscan delay, which was set to 450 ms in BEST-HMQC2 and to 1.4 s in SS-HMQC2 for optimal sensitivity in both experiments. Examples of 2D (ω_1, ω_2) strips extracted from these data sets are shown in Fig. 4a. On average, a sensitivity gain of approximately a factor of 2 is obtained from longitudinal relaxation enhancement when using BEST-Jcomp-HMQC2 instead of SS-HMQC2. This gain, however, is not uniform along the peptide chain and enhancement factors of 3 to 5 are observed for amide groups in the hydrophobic core of the protein characterized by a high local proton density and little internal dynamics. Interestingly, these are also the molecular regions where the most useful long-range structural information is obtained, as illustrated for the β -sheet structure of ubiquitin (Fig. 4b). J-mismatch compensation yields significant further signal enhancements for amide groups with large D_{NH} couplings. Sensitivity gains of up to a factor of 2 are observed when adding the J-mismatch compensation elements for cross and diagonal peaks were both amide groups involved in the correlation show large J-mismatch. The combination of both effects yields enhancement factors of the S/N ratio for individual correlation peaks in the range of ~1.5 to 8. The increased sensitivity allowed us to quantify additional peaks and thus obtain more structural restraints as compared to the SS-HMQC2 experiment. In addition to the 113 cross peaks that are observed in both experiments 41 additional cross peaks with intensities above a cutoff level of 3 times the noise level could be quantified. Exceptions are correlations between amide ^1H and the NH_2 groups of Asn and Gln side chains. These cross peaks have significantly reduced intensity in BEST-Jcomp-HMQC2 because of the additional J_{NH} coupling evolution of $2\text{N}_y\text{H}_x$ -type coherence during the J-compensation pulse sequence elements (see residue L15 in Fig. 4a). Figure 5a shows the correlation of D_{HH} values extracted from 3D spectra measured using the SS-HMQC2 and BEST-Jcomp-HMQC2 sequences, illustrating that the same level of accuracy is obtained from the two experiments, although with a significantly higher precision for the sensitivity-optimized BEST-Jcomp-HMQC2 version.

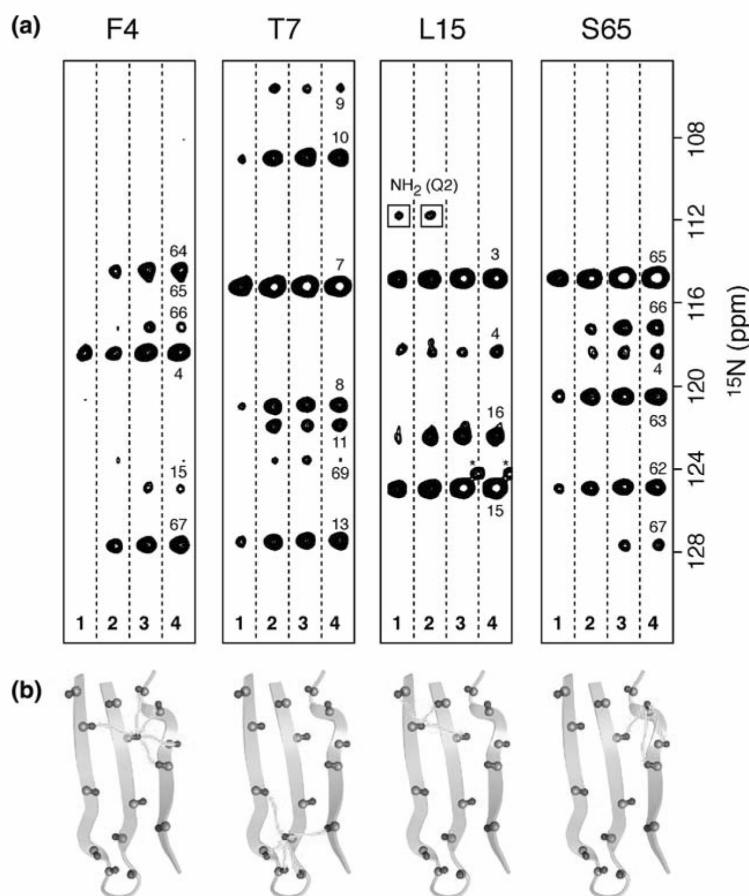


Fig. 4 2D ^{15}N (ω_1)- ^{15}N (ω_2) strips extracted at the amide ^1H (ω_3) frequency of several residues in the β -sheet of ubiquitin. 3D data sets were recorded at 800 MHz (28°C) on a sample of ^{15}N -labelled, fully protonated ubiquitin dissolved in a DMPC/DHPC mixture. The results obtained with the following pulse sequences are shown: (1) SS-HMQC2, (2) BEST-HMQC2 without J-mismatch compensation and z-filter, (3) BEST-Jcomp-HMQC2 with additional J-mismatch compensation, and (4) BEST-Jcomp-HMQC2 with J-mismatch compensation and z-filter. For all experiments the ^1H - ^1H transfer delays (corresponding to Δ_1 and Δ_2 in the sequence of Fig. 1a) were adjusted to 41 ms. $43(t_1) \times 43(t_2) \times 400(t_3)$ complex points were recorded

for spectral widths of 2400 Hz (ω_1), 2400 Hz (ω_2) and 10000 Hz (ω_3). Optimal recycle delays were used for all experiments: 450 ms for BEST-Jcomp-HMQC2, and 1.4 s for SS-HMQC2. The total acquisition time was 10 h for the BEST experiments, and 24 h for the SS-HMQC2 spectrum. The intensities in the SS-HMQC2 spectrum were scaled by a factor 0.65 to account for the unequal experimental times. A star indicates a residual peak from another ^1H plane (residue). (b) Schematic drawing of the β -sheet structure in ubiquitin. The arrows indicate amide pairs for which RDCs could be measured from the spectral regions displayed in (a), illustrating that important long-range structural information is obtained from these data

We have also experimentally evaluated the effect of the additional z-filter element on the overall sensitivity of the BEST-Jcomp-HMQC2 experiment and the accuracy of the measured amide ^1H - ^1H RDCs. Overall the intensity of most correlation peaks recorded with the additional z-filter is slightly reduced, by about 10–15% in the case of the weakly aligned ubiquitin sample used for this study (Fig. 4a), which is in agreement with the data shown in Fig. 2. There are two possible sources of signal losses: first, and probably

most importantly, ^1H longitudinal relaxation during the z-filter element and ^1H transverse relaxation during the WATERGATE sequence of ~2 ms duration; second, as can be seen from Fig. 2, the optimal scan time is slightly shifted towards longer delays as compared to the experiment without the z-filter element, indicating that the additional pulses decrease the steady-state polarization of the aliphatic protons, and thus the sensitivity of the experiment. The small intensity increase observed for some cross peaks is explained by

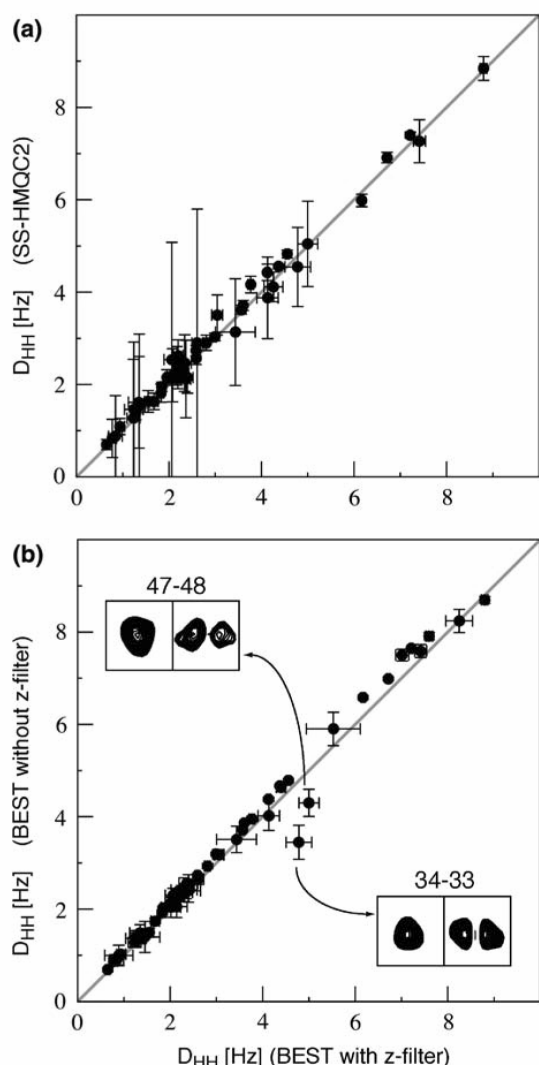


Fig. 5 (a) Correlation plot of amide ^1H - ^1H RDCs measured using the BEST-Jcomp-HMQC2 sequence of Fig. 1a (x-axis) and SS-HMQC2 (Wu and Bax 2002) (y-axis). The total acquisition time was 10 h for BEST-Jcomp-HMQC2 and 24 h for SS-HMQC2. (b) Correlation plot of amide ^1H - ^1H RDCs extracted from 3D BEST-Jcomp-HMQC2 spectra recorded with (x-axis) and without (y-axis) the additional z-filter. The experimental error bars were estimated on the basis of intensity uncertainties of 1.5 times the noise level. (b) The absolute values of the D_{HH} coupling constants were extracted from a single spectrum. Therefore all D_{HH} couplings were assumed to be in the 0 to $1/2\Delta = 12.2$ Hz range ($\Delta = \Delta_1 = \Delta_2$). The periodicity of the trigonometric functions in Eq. [1] does not allow the distinction between coupling constants D and $1/\Delta \cdot D$. In order to solve this ambiguity a second data set recorded using a different transfer delay Δ is required. The inserts in (b) show ^1H line shape distortions observed for some cross peaks (right spectrum) that are completely removed by addition of the z-filter (left spectrum)

the removal of artifacts with negative intensity in the $\omega_3(^1\text{H})$ dimension. The difference in the RDC values extracted from the two experiments recorded with and without z-filter is surprisingly small for most of the observed correlations (Fig. 5b), indicating that also without the z-filter meaningful structural restraints can be obtained. Only a few peaks show significant distortions along the ^1H dimension (see Inserts in Figs. 5b) that result in erroneous peak intensities and thus stronger deviations in the measured D_{HH} coupling constants. These peak distortions are absent in the spectra recorded with the additional z-filter illustrating that the antiphase $2\text{H}_x\text{H}_z$ coherence is efficiently suppressed. The z-filtered version of the experiment thus yields higher accuracy at the expense of a small loss in sensitivity. However, the outliers are easily identified from inspection of the peak shape and accurate coupling constants can also be obtained from the experiment without the z-filter element if these peaks are manually removed from the analysis.

Conclusions

We have presented BEST-Jcomp-HMQC2, a new pulse sequence that allows accurate measurement of amide ^1H - ^1H residual dipolar couplings in weakly aligned proteins. ^1H - ^1H RDC measurements provide important long-range restraints for NMR structure determination. This information is available immediately after backbone resonance assignment without the need for any additional tedious side chain resonance assignment or NOE analysis. ^1H - ^1H RDCs are therefore especially useful in the context of fast *de novo* fold elucidation using only a limited set of NMR-derived structural restraints. It has been shown recently (Bouvincies et al. 2006) that high structural precision is obtained by combining local orientational information from RDCs between covalently-bound nuclei in the protein backbone with long-range translational and orientational information from amide ^1H - ^1H RDCs. The high sensitivity provided by the BEST-Jcomp-HMQC2 experiment makes it attractive to measure amide ^1H - ^1H RDCs on a routine basis for partially aligned protein samples without the need for (per)deuteration.

Acknowledgments This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, the French Research Agency (ANR), Human Frontier Science Program Organization and the European Commission (EU-NMR). P.S. and R.S. acknowledge support from the French ministry of education, research, and technology.

We thank Beate Bersch, Isabel Ayala, and Jacques Covès (IBS Grenoble) for the preparation of the isotope-labeled protein samples.

References

- Andersson P, Weigelt J, Otting G (1998) Spin-state selection filters for the measurement of heteronuclear one-bond coupling constants. *J Biomol NMR* 12:435–441
- Atreya HS, Szyperski T (2004) G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. *Proc Natl Acad Sci USA* 101:9642–9647
- Bax A, Vuister GW, Grzesiek S, Delaglio F, Wang AC, Tschudin R, Zhu G (1994) Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol* 239, Pt C:79–105
- Beraud S, Bersch B, Brutscher B, Gans P, Barras F, Blackledge M (2002) Direct structure determination using residual dipolar couplings: reaction-site conformation of methionine sulfoxide reductase in solution. *J Am Chem Soc* 124:13709–13715
- Blackledge M (2005) Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog Nucl Magn Reson* 46:23–61
- Boisbouvier J, Delaglio F, Bax A (2003) Direct observation of dipolar couplings between distant protons in weakly aligned nucleic acids. *Proc Natl Acad Sci USA* 100:11333–11338
- Bouvignies G, Bernado P, Meier S, Cho K, Grzesiek S, Brüschweiler R, Blackledge M (2005) Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc Natl Acad Sci USA* 102:13885–13890
- Bouvignies G, Meier S, Grzesiek S, Blackledge M (2006) Ultrahigh-resolution backbone structure of perdeuterated protein GB1 using residual dipolar couplings from two alignment media. *Angew Chem Int Ed Engl* 45:8166–8169
- Diercks T, Daniels M, Kaptein R (2005) Extended flip-back schemes for sensitivity enhancement in multidimensional HSQC-type out-and-back experiments. *J Biomol NMR* 33:243–259
- Geen H, Freeman R (1991) Band-selective radiofrequency pulses. *J Magn Reson* 93:93–141
- Hus JC, Marion D, Blackledge M (2001) Determination of protein backbone structure using only residual dipolar couplings. *J Am Chem Soc* 123:1541–1542
- Kontaxis G, Delaglio F, Bax A (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol* 394:42–78
- Kupce E, Boyd J, Campbell ID (1995) Short selective pulses for biochemical applications. *J Magn Reson B* 106:300–303
- Lakomek NA, Fares C, Becker S, Carlomagno T, Meiler J, Griesinger C (2005) Side-chain orientation and hydrogen-bonding imprint supra- τ (c) motion on the protein backbone of ubiquitin. *Angew Chem Int Ed Engl* 44:7776–7778
- Meier S, Haussinger D, Jensen P, Rogowski M, Grzesiek S (2003) High-accuracy residual ^1H – ^{13}C and ^1H – ^1H dipolar couplings in perdeuterated proteins. *J Am Chem Soc* 125:44–45
- Meiler J, Prompers JJ, Peti W, Griesinger C, Brüschweiler R (2001) Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J Am Chem Soc* 123:6098–6107
- Nielsen NC, Bildsoe H, Jakobsen HJ, Sørensen OW (1989) Composite refocusing sequences and their application for sensitivity enhancement and multiplicity filtration in INEPT and 2D correlation spectroscopy. *J Magn Reson* 85:359–380
- Ottiger M, Delaglio F, Bax A (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* 131:373–378
- Pervushin K, Vogeli B, Eletsky A (2002) Longitudinal $(1)\text{H}$ relaxation optimization in TROSY NMR spectroscopy. *J Am Chem Soc* 124:12898–12902
- Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2:661–665
- Prestegard JH, Bougault CM, Kishore AI (2004) Residual dipolar couplings in structure determination of biomolecules. *Chem Rev* 104:3519–3540
- Schanda P, Brutscher B (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* 127:8014–8015
- Schanda P, Brutscher B (2006) Hadamard frequency-encoded SOFAST-HMQC for ultrafast two-dimensional protein NMR. *J Magn Reson* 178:334–339
- Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33:199–211
- Schanda P, Forge V, Brutscher B (2006a) HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn Reson Chem* 44:S177–S184
- Schanda P, van Melckebeke H, Brutscher B (2006b) Speeding up three-dimensional protein NMR experiments to a few minutes. *J Am Chem Soc* 128:9042–9043
- Sibille N, Blackledge M, Brutscher B, Coves J, Bersch B (2005) Solution structure of the sulfite reductase flavodoxin-like domain from *Escherichia coli*. *Biochemistry* 44:9086–9095
- Smith MA, Hu H, Shaka AJ (2001) Improved broadband inversion performance for NMR in liquids. *J Magn Reson* 151:269–283
- Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* 278:1111–1114
- Wimperis S, Bodenhausen G (1986) Heteronuclear coherence transfer over a range of coupling constants. A broadband-INEPT experiment. *J Magn Reson* 69:264–282
- Wu ZR, Bax A (2002) Measurement of long-range ^1H – ^1H dipolar couplings in weakly aligned proteins. *J Am Chem Soc* 124:9672–9673

5.3 La détection des RDCs à longue portée est limitée par la troncation dipolaire

Dans la limite de deux protons isolés P et Q, l'aimantation transférée dans le plan transversal, pendant une période de transfert T, entre ces deux protons est donné par :

$$I_{PQ} = \sin(\pi D_{PQ} T) e^{(-R_2 T)} \quad (5.7)$$

Pour deux protons éloignés, si la valeur de la constante de couplage est faible ($D_{PQ} T \ll 1$), on aura un transfert optimal pour $T = R_2^{-1}$ et $I_{PQ} = \pi D_{PQ} R_2^{-1} e^{(-1)}$.

Toutefois, dans un système à plusieurs protons, ce facteur est modulé par l'ensemble des couplages entre P et les protons proches. L'aimantation transférée peut s'exprimer alors par :

$$I_{PQ} = \sin(\pi D_{PQ} T) e^{(-R_2 T)} \prod_{J \neq Q} \cos(\pi D_{PJ} T) \quad (5.8)$$

Dans une protéine protonée, un proton est entouré en moyenne de 25 autres protons dans un rayon de 5 Å. En revanche, dans une protéine perdeutérée, le nombre de protons voisins d'un proton amide dans cette même sphère tombe à 4. Les RDCs étant inversement proportionnel à r^3 , le transfert d'aimantation entre deux protons séparés par des distances atteignant 7 Å sera limité par l'ensemble des couplages avec les protons les plus proches. Ce phénomène, nommé la troncation dipolaire, est la principale limitation à l'observation de RDCs à longue portée dans les protéines de faible taille. Dans la suite de ce paragraphe, je présenterai quelques exemples où l'on a pu repousser cette limite dans les biomolécules.

5.3.1.1 Suppression de couplages forts par des impulsions à bandes sélectives

Afin d'observer les couplages dipolaires les plus faibles, il est nécessaire de supprimer les interactions dipolaires entre les protons les plus proches. Une solution consiste à utiliser des impulsions très sélectives pour découpler des protons ayant des fréquences distinctes. En collaboration avec P. Schanda, nous avons implémenté des impulsions sélectives de découplage des protons amides dans la séquence SS-HMQC2 (Figure 5.5). Les impulsions à bandes sélectives n'agissent que sur des

protons amides se trouvant sur une bande de déplacements chimiques définie. Elles ont pour fonction de refocaliser les couplages H_N-H_N entre des protons qui ont une fréquence hors de cette bande de déplacement chimique et ceux à l'intérieure de cette bande. Un exemple d'implémentation de ce type d'impulsion sélective pour la détection de RDCs $^1H_N-^1H_N$ sur un échantillon d'ubiquitine U- $[^2H]$; U- $[^{15}N]$ est présenté dans la Figure 5.8. L'utilisation d'une impulsion à bande sélective, centrée sur 9,5 ppm avec une largeur d'inversion de $\pm 1,5$ ppm, permet de limiter une partie de la troncation dipolaire en éliminant les couplages avec les protons ayant une fréquence de résonance en dehors de cette bande. Le transfert peut avoir lieu seulement s'il se fait entre protons situés dans la bande de déplacement chimique qui est refocalisé par cette impulsion, ou entre paire de protons non affecté par l'impulsion sélective. Sur la Figure 5.8, dans le panneau de gauche une expérience SS-HMQC2 standard a été réalisée (Wu *et al*, 2002 ; Schanda *et al*, 2007). La présence d'un couplage intense entre les H_N de la Glycine 35 et de l'Isoleucine 36 (8,5 Hz) donne lieu à un pic de corrélation intense et limite le transfert vers d'autres protons. Lorsqu'on applique l'impulsion sélective à 9,5 ppm, le couplage entre le H_N de la Glycine 35 (8,53 ppm - fréquence 1H) et le H_N de l'Isoleucine 36 (6,19 ppm - fréquence 1H) est découplé. Ainsi, il est possible d'observer un transfert de la Glycine 35 vers le H_N du Glutamate 34 (8,74 ppm - fréquence 1H), correspondant à un couplage plus faible, $D_{HH} = 2,7$ Hz.

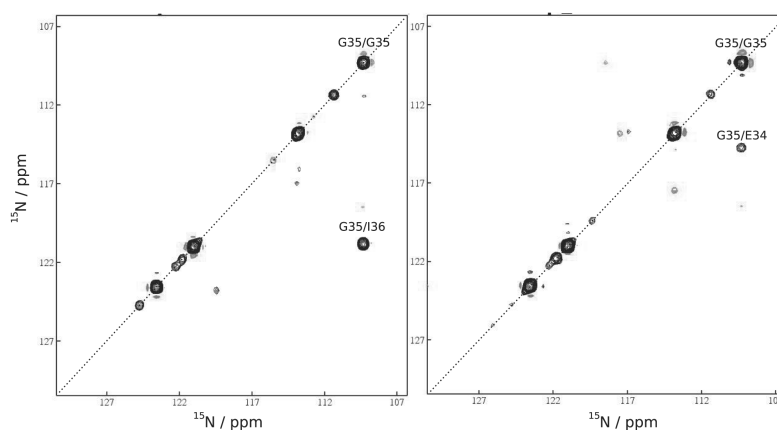


Figure 5.8: Plan 2D ^{15}N - ^{15}N à 8,53 ppm de 1H du spectre sans impulsion à bande sélective (à gauche) et avec impulsion à bande sélective (à droite).

Le pic de corrélation entre la Glycine 35 et l'Isoleucine 36 (6,19 ppm) est supprimé dans le spectre en B et un pic de corrélation entre la Glycine 35 et le Glutamate 34 (8,74 ppm) est détecté.

Dans cet exemple, la suppression de couplages intenses permet la détection de transfert entre des protons plus éloignés. Toutefois, le recouvrement des déplacements chimiques dans les protéines ne permet qu'une utilisation relativement limitée des impulsions à bandes sélectives. De plus, pour vraiment être spécifique d'une paire de résonance, il serait nécessaire d'appliquer des impulsions sélectives de l'ordre de quelques dixièmes de seconde. Ce qui n'est pas réalisable de manière pratique compte tenu des vitesses de relaxation transversale des protons dans les macromolécules biologiques.

5.3.1.2 Limitation de la troncation dipolaire dans les acides nucléiques

Comme présenté au § 5.3.1.1 pour un échantillon d'ubiquitine, de masse moléculaire de 9 kDa, on peut limiter le transfert grâce à l'utilisation d'impulsion sélective. Cependant, même pour un échantillon d'ubiquitine deutérée, la densité de protons reste trop importante ($\sim 20\%$; § 3.2.2.2), il reste 76 protons amides auxquels s'ajoute 55 NH_2 . Ainsi la troncation dipolaire est toujours limitante pour la mesure de faibles couplages entre protons séparés par des distances au-delà de 7 Å. En comparaison, un acide nucléique de 12 paires de bases, de masse moléculaire similaire, possède une densité de protons beaucoup plus faible. De plus, dans les acides nucléiques, les fréquences de résonance pour un type d'atomes sont bien séparées des autres types d'atomes, on peut ainsi limiter le transfert entre les H_1' et quelques H_5 , soit une trentaine de protons. Pour un ARN de structure standard, la distance moyenne séparant deux protons H_1' est supérieure à 5 Å. Cette faible densité de protons limite fortement l'effet de troncation dipolaire. De plus, ces mêmes protons possèdent des vitesses de relaxation très lente ($R_2 \approx 5 \text{ Hz}$), favorable à l'utilisation de temps de transfert long de l'ordre de 100 à 200 ms. Il a ainsi été démontré que, dans ce type d'ARN, des couplages sont détectables entre des protons séparés par des distances atteignant 12 Å (Boisbouvier *et al*, 2003). Un exemple de RDCs ^1H - ^1H détecté sur un ARN de 24 nucléotides est montré sur la Figure 5.9.

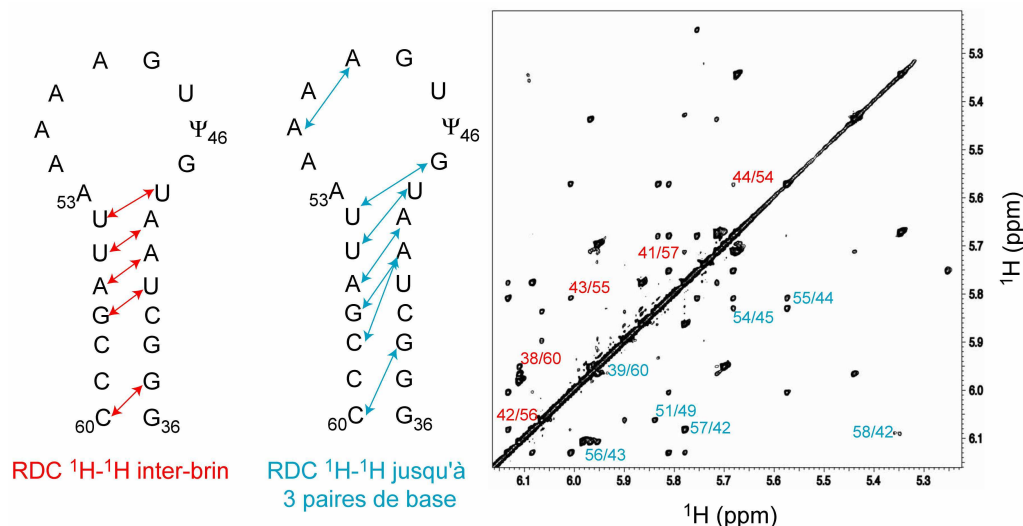


Figure 5.9: En A) Les couplages dipolaires résiduels ^1H - ^1H observés sont indiqués sur la séquence de l'ARN de 24 nucléotides. En B) Spectre de corrélation ^1H - ^1H de cet ARN dont sont extraits les RDCs.

Le spectre a été enregistré avec la séquence d'impulsion (Figure 5.3 ; Boisbouvier *et al*, 2003) sur un échantillon d'ARN à 1,2 mM. L'acquisition a été effectuée pendant 12 heures sur un spectromètre opérant à une fréquence de 600 MHz. Les couplages dipolaires résiduels ^1H - ^1H détectés entre les brins sont indiqués en rouge et ceux détectés entre 2 ou 3 paires de bases sont représentés en bleu.

Dans cet exemple, il est démontré que des transferts par RDC entre protons séparés par plus de 10 Å sont possibles, à la condition que la troncation dipolaire ne soit plus le facteur limitant. Pour réaliser cela, une solution, décrite au chapitre 2, consiste à diminuer la densité de protons en utilisant la perdeutération avec le marquage spécifique des protons des méthyles d'un type de résidu dans les protéines. Ce marquage spécifique peut permettre dans le cas d'une protonation spécifique des méthyles δ_1 des Isoleucines de réduire la densité de protons à 2-5 % (§ 2.4.4). Compte tenu de leurs propriétés favorables (§ 2.3), les méthyles apparaissent comme une sonde de choix pour détecter des RDCs à longue portée dans les protéines. Les possibilités et les difficultés spécifiques liées à la détection de RDCs entre méthyles seront discutées par la suite.

5.4 Détection de RDCs interméthyles à longue portée dans les protéines spécifiquement protonées

5.4.1 Propriétés de relaxation des méthyles en milieu isotrope

Dans les protéines perdeutérées et spécifiquement protonées sur les méthyles, la relaxation longitudinale ^1H des groupements méthyles est très favorable. Au Chapitre 4 (Tableau 4.1), j'ai démontré que le temps de relaxation longitudinale moyen, T_1 , des méthyles était de l'ordre de la seconde et ce même pour des systèmes de supramoléculaire ($> 100\text{kDa}$). En comparaison, pour des protons isolés (H_N ou H_α) au sein de protéines perdeutérées, les temps de relaxation longitudinale atteignent plusieurs secondes, voir la dizaine de secondes. Les temps de relaxation longitudinale rapide des méthyles permettent d'exécuter les expériences RMN avec un délai de recyclage court, optimal pour la sensibilité et la résolution des expériences RMN.

Ces dernières années, le groupe du Professeur Lewis Kay (Toronto, Canada) a démontré que les groupements méthyles sont également caractérisés par des propriétés de relaxation transversales très favorables. L'utilisation des groupements méthyles comme sondes structurales et dynamiques a été largement développé durant ces dernières années pour l'étude des protéines ou complexes de grandes tailles (Tugarinov *et al*, 2004a ; Sprangers *et al*, 2005 ; 2007). La relaxation des cohérences ^1H simple quanta d'un méthyle isolé dans une protéine perdeutérée dans un milieu isotrope est bi-exponentielles (Annexe 2). Les effets de corrélations croisée dipôle-dipôle ^1H - ^1H donne lieu à un premier phénomène d'interférence destructive de la relaxation pour la moitié du signal et vont pour l'autre moitié du signal contribuer à augmenté la vitesse de relaxation. Dans les macromolécules biologiques, la moitié du signal n'est plus utilisable dans les expériences RMN standard à cause de temps de relaxation transversale très rapide, $R_{2,H}^f$. Par contre l'autre partie du signal est caractérisée par des vitesses de relaxation transversale lentes. Il peut ainsi être montré que, dans la limite des macromolécules ($\omega\tau_c \gg 1$) et de la rotation rapide des méthyles, cette partie du signal n'est plus affectée par les interactions intraméthyles, mais uniquement par des interactions avec les spins externe aux protons du méthyles.

Dans un système perdeutééré et marqué spécifiquement sur quelques méthyles, la composante de la vitesse de relaxation transversale relaxant lentement s'exprime par (Annexe 2) :

$$R_{2,H}^s = \left(\frac{8}{15}\right) \left(\frac{\mu_0}{4\pi}\right)^2 \sum_i \frac{\gamma_H^2 \gamma_D^2 \hbar^2 \tau_c}{r_{HD,i}^6} + \left(\frac{1}{4}\right) \left(\frac{\mu_0}{4\pi}\right)^2 \sum_i \frac{\gamma_H^4 \hbar^2 \tau_c}{r_{HH,i}^6} \quad (5.9)$$

Cette équation n'est qu'une approximation, valable que dans le cas d'une macromolécule avec un temps de corrélation élevé. Sur la Figure 5.10, j'ai représenté l'évolution de l'aimantation du méthyle de δ_1 de l'Isoleucine 36 de l'ubiquitine après une impulsion d'excitation de 90° .

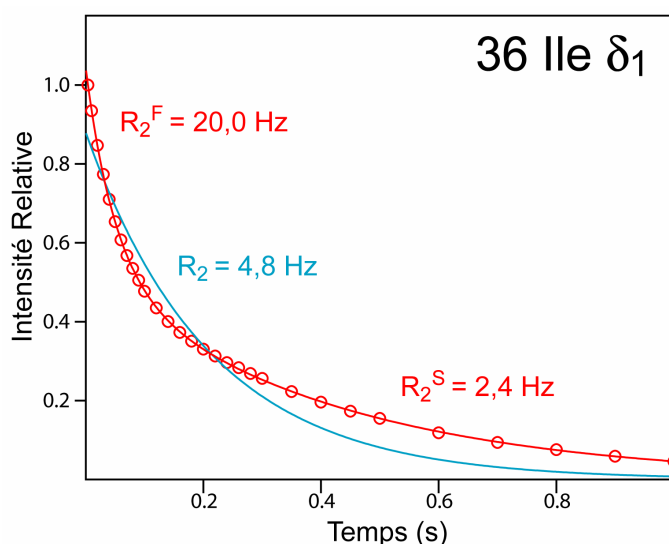


Figure 5.10: Exemple de relaxation transversale protons des méthyles δ_1 de l'Isoleucine 36 de l'ubiquitine U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 à une température de 25°C ($\tau_c = 5,5$ ns) et $B_0 = 18.8$ T.

Les ajustements des valeurs expérimentales sont représentés par la courbe en rouge pour une fonction bi-exponentielles et par la courbe bleu pour une fonction mono-exponentielle. Les vitesses de relaxation de la composante rapide $R_{2,H}^f$ et de la composante lente $R_{2,H}^s$ sont notées en rouge et en bleu la vitesse de relaxation transversale pour un ajustement avec un modèle mono-exponentiel.

Comme on peut le voir sur la Figure 5.10, même dans un système de petite taille la relaxation du signal d'un méthyle $^{12}\text{C}^1\text{H}_3$ reste bi-exponentielle avec l'une de ces composantes caractérisés par une vitesse de relaxation de l'ordre quelques Hz, $R_{2,H}^s$. La seconde composante $R_{2,H}^f$ (correspondant à la moitié du signal) a une vitesse de relaxation 10 fois plus rapide. Compte tenu des temps de transfert standard utilisés

pour la mesure de RDCs à longue portée supérieure à 100 ms (Boisbouvier *et al*, 2003), cette composante aura une contribution négligeable par rapport au signal détecté et ce même dans les protéines de petite taille.

Dans l'échantillon U-[^2H], U-[^{13}C], [$^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 , il n'y a pas de carbone-13 proche des protons du méthyle, les spins ^{13}C ont donc une contribution négligeable. En revanche. Dans un échantillon U-[^2H], U-[^{12}C], [$^{13}\text{C}^1\text{H}_3$]-Ile- δ_1 , où les méthyles sont enrichis en ^{13}C , il existe une relaxation dipolaire supplémentaire. Compte tenu de phénomènes de corrélation croisée et de la dynamique des groupements cette interaction reste assez faible pour la composante rapide, $R_{2,H}^f$. Cependant lorsque l'on s'intéresse à la composante lente, qui n'est pas affectée par les sources de relaxation ^1H - ^1H intraméthyles, cette faible contribution ^{13}C - ^1H devient prédominant et correspond à une augmentation de la vitesse de relaxation par un facteur 1,6 pour le méthyle δ_1 de l'Isoleucine 36 de l'ubiquitine. Par conséquent lorsque la résolution dans la dimension ^1H le permet, il sera souhaitable de travailler avec un échantillon sans marquage ^{13}C sur les méthyles afin de détecter les RDC ^1H - ^1H les plus faibles.

5.4.2 Evolution des cohérences transversales simple-quanta ^1H des méthyles en milieu cristal liquide

5.4.3 Evolution de l'aimantation transversale des méthyles en milieu cristal liquide

L'alignement de la protéine dans milieu cristal liquide induit des interactions dipolaires (^1H - ^1H , ^1H - ^2H , ^1H - ^{13}C) qui ne sont plus moyennées à zéro et qui peuvent affecter l'évolution des cohérences transversales simple quanta des protons des méthyles. Ces différentes interactions dipolaires sont détaillées ci-dessous.

5.4.3.1 La contribution du deutérium à la relaxation des méthyles en milieu aligné

Dans des échantillons U-[^2H], [$^{12}\text{CH}_3$]-Ile- δ_1 , les interactions dipolaires entre les protons des méthyles et les deutériums sont les plus nombreuses. Durant les périodes

de transfert d'une expérience Spin-Echo de durée totale τ , par exemple, l'aimantation d'un proton H_y évolue sous forme antiphase avec les ^2H (ou D) à cause du couplage dipolaire (D_{HD}).

$$H_y \xrightarrow{D_{HD}} \begin{cases} H_y \cos(\pi D_{HD} \tau) \\ 2H_x D_z \sin(\pi D_{HD} \tau) \end{cases}$$

La présence d'une impulsion 180° au milieu de la période de transfert devrait permettre de refocaliser l'évolution due à ces couplages. Toutefois, la relaxation du terme $2H_x D_z$ revêt la forme suivante:

$$R(2H_x D_z) \approx R_{2,H} + R_{1,D} \quad (5.10)$$

Due à une interaction quadripolaire particulièrement efficace, la vitesse de relaxation longitudinale du deutérium ($R_{1,D}$) est très rapide (quelques centaines de Hz) par rapport à la relaxation transversale du proton ($R_{2,H}$) (quelques Hz pour l'ubiquitine).

Dès qu'une partie de l'aimantation évolue en une cohérence $2H_x D_z$, cette partie du signal relaxera quasi instantanément, ainsi, ceci se traduira par une perte de signal. Il ne suffit donc pas de refocaliser l'aimantation par une impulsion 180° durant la période de transfert mais de découpler en continue tout le long de cette période, pour éviter à tout instant la formation des cohérences impliquant le Deutérium. Cet effet est illustré pour un échantillon d'ubiquitine U- $[^2\text{H}]$, U- $[^{13}\text{C}]$, $[^{12}\text{C}^1\text{H}_3]$ -Ile- δ_1 aligné dans les phages sur la Figure 5.11. On observe que pour un délai de transfert de 65 ms, la présence de couplages dipolaires entre les protons des méthyles et les deutériums peuvent conduire à une perte atteignant 65 % du signal. Compte tenu que les expériences permettant la détection de RDCs à longue portée nécessite deux périodes de transfert de plus de 100 ms, il est particulièrement important de supprimer ces interactions par un découplage continue.

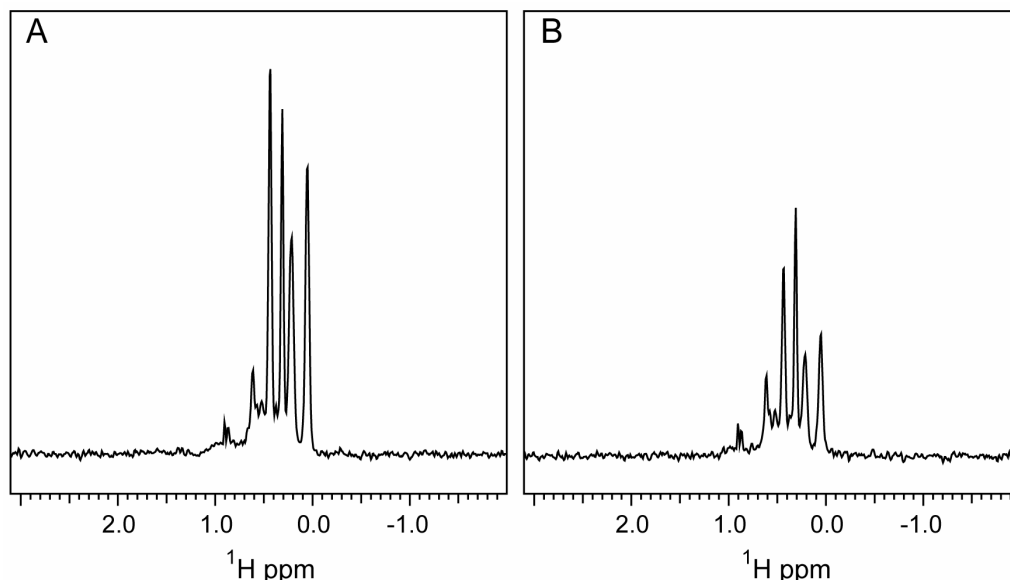


Figure 5.11: Spectres 1D des méthyles δ_1 de l'Ubiquitine U- ^2H], U- ^{13}C], $^{12}\text{C}^1\text{H}_3$]-Ile- δ_1 aligné après un écho d'une durée totale de 65 ms.

L'échantillon a été aligné dans une solution de 17 mg/ml de phages pf1, 40 mM NaCl à 25°C ($D_a = 21,5$ Hz ; $R = 0,541$). Le délai de transfert totale τ de l'expérience "Spin-Echo" est de 65 ms. Les expériences ont été enregistrées sur un spectromètre opérant à une fréquence proton de 600 MHz : en présence d'un découplage continu (type WALTZ-16 d'une puissance B_1 moyen de 1,3 kHz) en A ; et les interactions dipolaires ^1H - ^2H sont refocalisées par une seule impulsion 180° placée au milieu de la période de transfert en B.

5.4.3.2 Contribution des autres protons

En milieu aligné, l'ensemble des couplages dipolaires résiduels ^1H - ^1H vont contribuer à défocaliser l'aimantation pendant la période de transfert. On peut distinguer ces RDCs ^1H - ^1H en deux catégories :

(i) Les RDCs avec les protons résiduels présents de 1 à 2 % (§ 2.4.4). Dans l'ubiquitine perdeutérée, il y a plus de 600 sites qui peuvent être occupés par des protons de manière résiduelle. Autour de chaque proton, il y a en moyenne 70 sites de protons distants par moins de 7 Å. Ainsi, si l'on s'intéresse à des RDCs à longue portée entre méthyles éloignés par plus de 10 Å, statistiquement, cette protonation résiduelle donnera lieu à des RDCs plus intenses. Ceci se manifeste par une décroissance rapide du signal importante en milieu liquide cristallin (Figure 5.12).

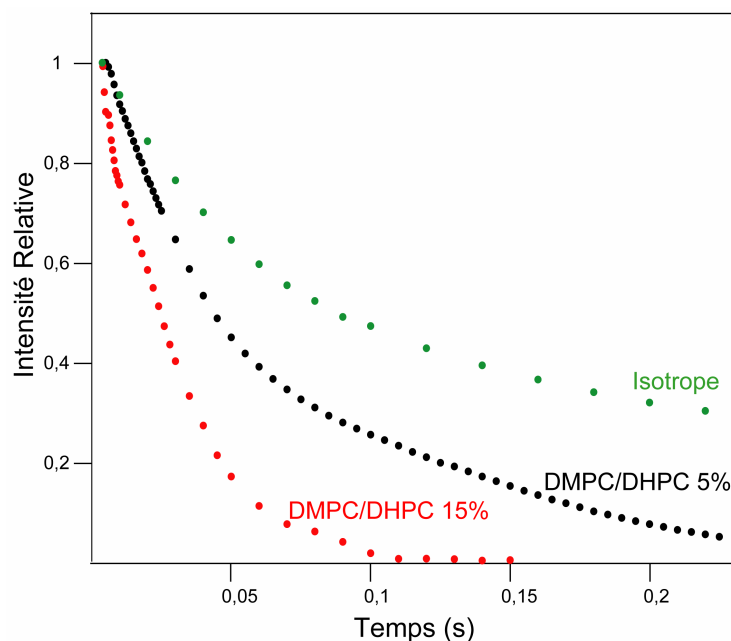


Figure 5.12: Exemple de relaxation transversale protons des méthyles δ_1 de l'Isoleucine 13 de l'ubiquitine enregistré sur un échantillon U- ^2H , U- ^{13}C , $^{12}\text{C}^1\text{H}_3$ -Ile- δ_1 .

L'intensité du signal du méthyle de l'Isoleucine 13 de l'ubiquitine enregistré à partir d'expérience Spin-Echo avec découplage du deutérium en continu avec différents temps de transfert. En vert en milieu isotrope, en rouge dans une solution de DMPC/DHPC (3:1) à 15 % ($D_a = 23,06$ Hz) et en noir dans une solution de DMPC/DHPC (3:1) à 5 % ($D_a = 7,93$ Hz).

Afin de limiter le phénomène de troncation dipolaire dus au proton résiduels, il est nécessaire d'utiliser des milieux orientant dilués. Comme on peut l'observer sur la Figure 5.12, 1/3 du signal est toujours détectable après 100 ms d'évolution transversales dans un échantillon d'ubiquitine faiblement aligner (DMPC/DHPC (3:1) à 5 %, $D_a = 7,93$ Hz ; $R=0,174$). Afin de limiter d'avantage la troncation dipolaire, on pourra pendant ce transfert appliquer des impulsions sélectives couvrant la plage de résonances ^1H des méthyles δ_1 des Isoleucines. Ainsi toutes les interactions dipolaires avec les protons résiduels ayant leur fréquence de résonance en dehors de cette zone seront découplées.

(i) Les RDCs entre méthyles spécifiquement protonés : les Isoleucines δ_1 dans le cas présent. Il s'agit des couplages que l'on veut utiliser pour transférer de l'aimantation à longue portée. Compte tenu du faible nombre de méthyles marqués ^1H et de la distance moyenne importante, ces couplages sont relativement faibles ($< 1\text{Hz}$ dans la majorité des cas). Ils contribuent donc très faiblement à la défocalisation du signal lors d'une expérience de Spin-Echo. Il est donc très important de limiter les autres effets pour pouvoir les détecter.

5.4.3.3 Evolution des couplages dipolaires résiduels ^1H - ^1H intraméthyles

La Figure 5.13 représente la courbe de décroissance de l'intensité des méthyles δ_1 des Isoleucines dans un échantillon d'ubiquitine en fonction de la période d'évolution en milieu aligné. Une fois un découplage ^2H appliqué en continu, ainsi que le découplage de la majorité des couplages dipolaires avec les protons résiduels à l'aide d'impulsions sélectives, on observe une oscillation atténuée de la décroissance de l'intensité des signaux des méthyles en fonction du temps de relaxation. Une explication probable de cette oscillation est qu'elles sont liées à l'évolution dus aux couplages dipolaires ^1H - ^1H intraméthyles. Il apparaît que l'effet d'oscillation est d'autant plus grand que l'alignement de l'ubiquitine est important (courbe rouge Figure 5.13).

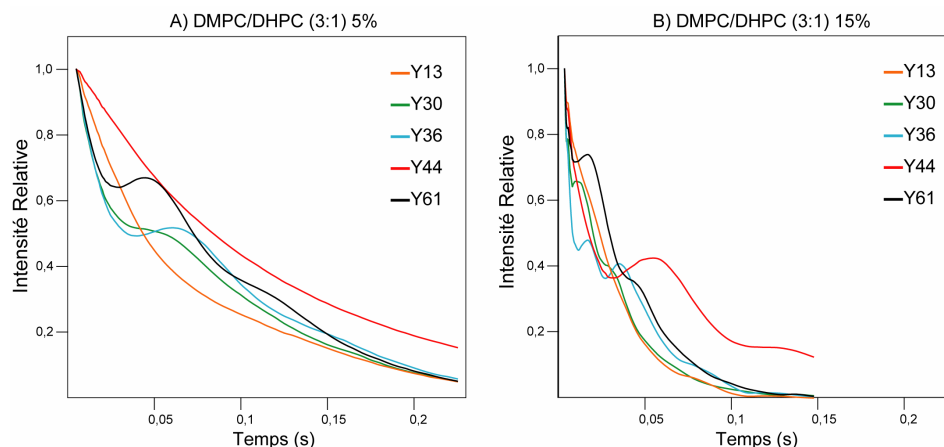


Figure 5.13: Les courbes de relaxation transversale ^1H des méthyles δ_1 de l'Ubiquitine U- $[\text{}^2\text{H}]$, U- $[\text{}^{13}\text{C}]$, $[\text{}^{12}\text{C}^1\text{H}_3]$ -Ile- δ_1 en milieu cristallin liquide.

Elles ont été enregistrées à partir d'une expérience Spin-Echo sur un spectromètre à 600 MHz, équipé d'une sonde cryogénique, avec une impulsion sélective de refocalisation appliquée sur les méthyles δ_1 en milieu de la période de relaxation et un découplage continue ^2H pendant toute la durée de l'expérience pour l'ubiquitine diluée dans un mélange de DMPC/DHPC (3:1) : à 5% ($D_a = 7,93$ Hz) en A et 15% ($D_a = 23,06$ Hz) en B. Compte tenu de la superposition des signaux des méthyles des Isoleucines 3 et 23, les données pour ces deux méthyles ne sont pas représentées.

Afin de vérifier cette hypothèse, nous avons déterminé les couplages dipolaires résiduels ^1H - ^1H intraméthyles. Compte tenu de la symétrie des méthyles et de leur rotation rapide, on peut directement recalculer ces couplages à partir des RDCs $^1D_{\text{CH}}$ intraméthyles (Sibille *et al*, 2002 ; Kaikkonen *et al*, 2001):

$$^2D_{\text{HH}} = 2,28 \times ^1D_{\text{CH}} \quad (5.11)$$

Résidus	$ ^2D_{HH} $	Résidus	$ ^2D_{HH} $
03	9,64 Hz	36	7,57 Hz
13	3,37 Hz	44	2,34 Hz
23	7,30 Hz	61	11,74 Hz
30	11,24 Hz		

Tableau 5.1 : RDCs 1H - 1H intraméthyles calculés à partir des RDCs 1H - ^{13}C intraméthyles mesurés sur un échantillon d'ubiquitine U- $[^2H]$, U- $[^{13}C]$, $[^{12}C^1H_3]$ -Ile- δ_1 dans DMPC/DHPC (3:1) à 5 % (Da= 7,93 ; R = 0,174).

On observe une très bonne corrélation entre les valeurs déterminées pour chacun des méthyles de l'ubiquitine U- $[^2H]$, U- $[^{13}C]$, $[^{12}C^1H_3]$ -Ile- δ_1 alignée dans une solution de DMPC/DHPC (3:1) à 5 %, et la période de l'oscillation observée sur la Figure 5.13.A.

A ce jour, les études théoriques disponibles permettant de décrire l'évolution et la relaxation des systèmes à trois spins concernent des systèmes de type AMX, ABX et A_3 . Elles ont été développées pour un milieu isotrope. Afin d'essayer d'expliquer ce phénomène, on peut se référer au diagramme d'énergie d'un système A_3 (Figure 5.14 ; Tugarinov *et al*, 2007).

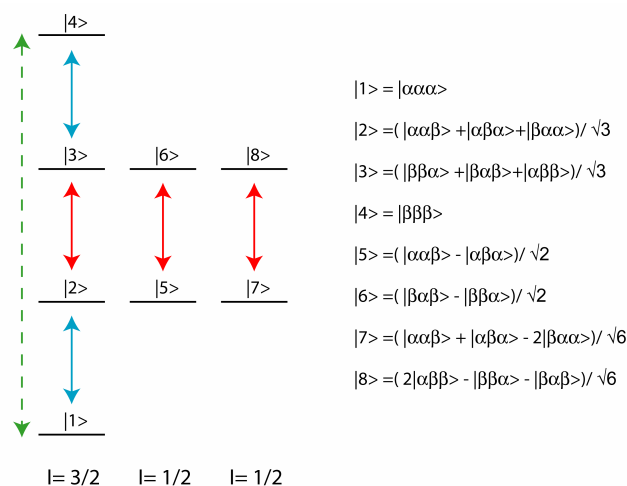


Figure 5.14: Diagramme des niveaux d'énergie d'un système A_3 d'un groupement méthyle.

Les transitions 1H simples quantum relaxants lentement (rapidement) sont représentées par des flèches rouges (bleues). Les états propre des niveaux d'énergie sont représentés par $|i, j, k\rangle$ ($i, j, k \in \{\alpha, \beta\}$). Le nombre quantique de spin, I , des trois

états multiples est indiqué sur le diagramme. La transition ^1H triple quanta est indiquée par une flèche verte en pointillée.

Les transitions dans un système à trois spins identiques, A_3 , se répartissent entre niveaux de spins $1/2$ (états singulets) et niveau de spin $3/2$ (état triplet). Alors que les états singulets ne sont pas sensibles aux couplages $^1D_{HH}$, l'état triplet va, lui, évoluer avec les couplages dipolaires résiduels $^1D_{HH}$. Cette interprétation paraît simple, mais il faut la prendre avec précaution, car, aujourd'hui, il n'existe aucune étude théorique permettant de décrire l'évolution et la relaxation d'un système à trois spins en milieu cristal liquide (Pr. Alfred G. Redfield, communication personnelle). Malgré cette incertitude, il est à noter que cette oscillation atténuée affecte aussi bien les pics de corrélation diagonaux que les pics de corrélation croisés. De ce fait, si l'on utilise une approche pour quantifier les RDCs similaire à ce qui a été présentée pour les RDCs H_N-H_N (§ 5.2.1.2), ces phénomènes n'auront pas d'influence sur la quantification des RDCs.

5.4.4 Détection de RDCs ^1H - ^1H entre méthyles éloignés

5.4.4.1 La séquence d'impulsion : "3D Selective CT-COSY"

Pour l'observation de RDCs ^1H - ^1H à longue portée, des attentions particulières doivent être prises en compte tant sur le type d'échantillon utilisé que sur la séquence d'impulsion :

(i) Comme la relaxation transversale est très importante dans ce type de mesure, il est judicieux de limiter tous les effets affectant cette relaxation. Comme montré sur la Figure 5.10, l'utilisation d'un marquage isotopique ^{12}C sur le carbone δ_1 des méthyles des Isoleucines est plus favorable que le marquage isotopique ^{13}C , supprimant ainsi l'interaction dipolaire ^1H - ^{13}C entre les protons et le carbone du méthyle.

(ii) De plus, comme pour la mesure de NOEs à longue portée, des artéfacts peuvent être détectés entre un méthyle δ_1 des Isoleucines marqué à 100% et des

protons résiduels à 2 %. L'utilisation d'un filtre lors de la séquence d'impulsion est nécessaire pour supprimer ces artéfacts.

(iii) Enfin, l'interaction dipolaire ^1H - ^2H doit être supprimée, pour ne pas perdre l'aimantation à cause de la relaxation longitudinale très rapide du deutérium. Pour ce faire un découplage ^2H durant toute la séquence doit être appliqué.

La séquence d'impulsion "3D Selective CT-COSY" (Boisbouvier *et al*, 2003) a été modifiée afin d'inclure ces trois modifications nécessaires pour la mesure de RDCs ^1H - ^1H entre méthyles dans des échantillons U- ^{12}H , U- ^{13}C , $^{12}\text{CH}_3$ -Ile- δ_1 (Figure 5.15). Les temps de transfert T_A et T_B sont optimisés pour la détection de couplages faibles, à un temps de transfert optimal égal à T_2 (§ 5.3). L'utilisation de l'impulsion sélective REBURP durant les périodes T_A et T_B , permet d'éliminer les couplages dipolaires avec la majorité des protons résiduels.

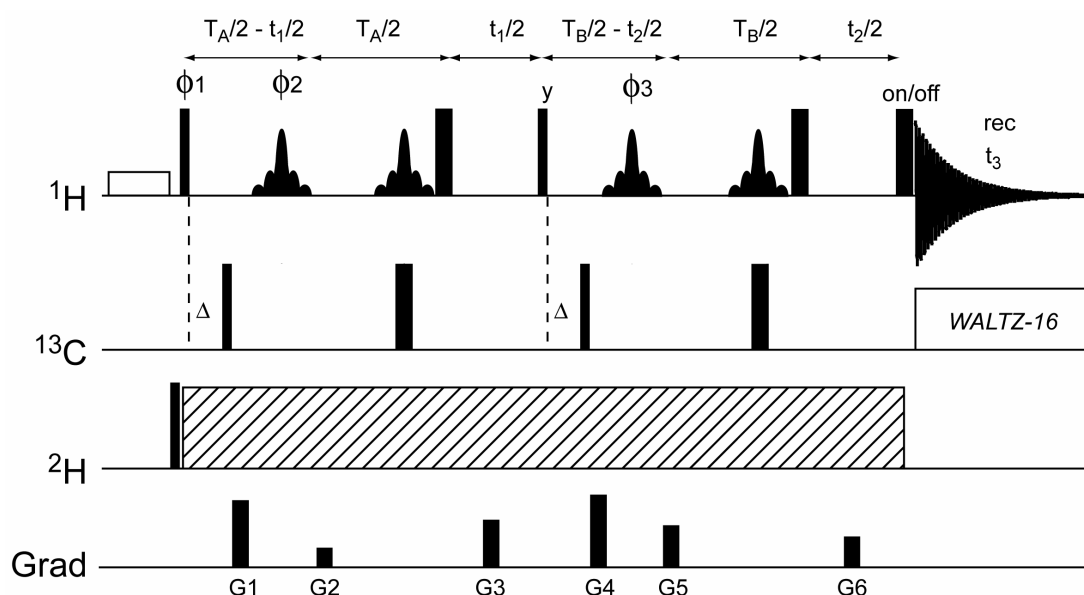


Figure 5.15: Schéma d'impulsion de l'expérience "3D Selective Constant-time COSY" avec filtre ^{13}C .

Les impulsions rectangulaires fines et larges sont appliquées avec des angles d'impulsion respectivement de 90° et de 180° . À moins que cela ne soit indiqué, les impulsions sont appliquées le long de l'axe X. La fréquence ^1H est placée au centre de la région méthyle des Ile- δ_1 et la fréquence ^{13}C est placée au centre de la région C_β des isoleucines. L'impulsion sélective ^1H est de type 180° REBURP (la durée est de 3 ms à 800 MHz, pour une largeur de bande d'inversion de 1600 Hz, centré sur la région des méthyles Ile- δ_1). La durée et la puissance des gradient (en unités ms, G/cm) sont : $G_1 = (0.2; 48)$, $G_2 = (0.2; 14)$, $G_3 = (0.2; 34)$, $G_4 = (0.2; 52)$, $G_5 = (0.2; 30)$, $G_6 = (0.2; 22)$. Le cycle de phase est: $\phi_1 = 4(x)$; $\phi_2 = 2(x)$, $2(y)$; $\phi_3 = x, y, x, y$; rec = $y, -y, -y, y$. La détection en quadrature en F_1 est achevée en incrémentant ϕ_1 de la manière usuel du States-TPPI et en F_2 le mode Echo/Anti-Echo est implémentée par l'insertion alternée entre les cycles d'une impulsion ^1H 180° juste avant l'acquisition. Les délais de recouvrement des gradients ont été placés à $\delta = 200 \mu\text{s}$, et le délai de transfert $\Delta = 4 \text{ ms}$.

L'expérience est enregistrée en mode 3D. La fréquence ^1H d'un premier méthyle est éditée pendant la première période de transfert. La fréquence ^1H du second méthyle est éditée pendant l'acquisition et la seconde période de transfert. Du fait de la présence d'un découplage homonucléaire et d'une édition à temps constant pendant une période de plus de 100 ms, la seconde période de transfert permet d'éditer la fréquence avec une très grande résolution. La dernière édition des fréquences ne sert qu'à la détection du signal. L'édition des fréquences pendant la seconde période de transfert (t_2) se fait sans perte de signal grâce à une édition en mode Echo/Anti-Echo réalisée grâce à une impulsion de 180° appliquée juste avant la détection pour la moitié des FID (Figure 5.15).

5.4.4.2 Application : ubiquitine U- ^{2}H], U- ^{13}C], $^{12}\text{CH}_3$]-Ile- δ_1

Des spectres tri-dimensionnels ^1H à partir de la séquence d'impulsion "3D Selective CT-COSY" sur des échantillons d'ubiquitine U- ^{2}H], U- ^{13}C], $^{12}\text{CH}_3$]-Ile- δ_1 dans différents milieux d'alignement ont été enregistrés. L'ensemble des milieux cristal liquide est en solution dans du D_2O . Trois milieux anisotropes ont été utilisés :

- (i) une solution de DMPC/DHPC (3:1) à 5% (Da = 7,93 ; R= 0,174).
- (ii) un mélange d'alcool n-alkyl-poly(éthylène-glycol)/hexanol (Da = 7,08 ; R= 0,280).
- (iii) une solution de bactériophages à 12 mg/mL avec 100mM NaCl (Da = 13,87 ; R = 0,610).

La Figure 5.16 représente des plans 2D extraits des spectres 3D mesurés dans les trois milieux d'alignement, où l'on observe les RDCs ^1H - ^1H entre le méthyle δ_1 de l'Isoleucine 30 et les autres méthyles δ_1 des Isoleucines de l'ubiquitine. Compte tenu de l'orientation différente des tenseurs d'alignement dans le repère moléculaire, les paires de méthyles de chacun de trois milieux sont caractérisées par une valeur de RDCs différentes. L'ensemble des RDCs ^1H - ^1H interméthyles détectés est montré sur la Figure 5.17.

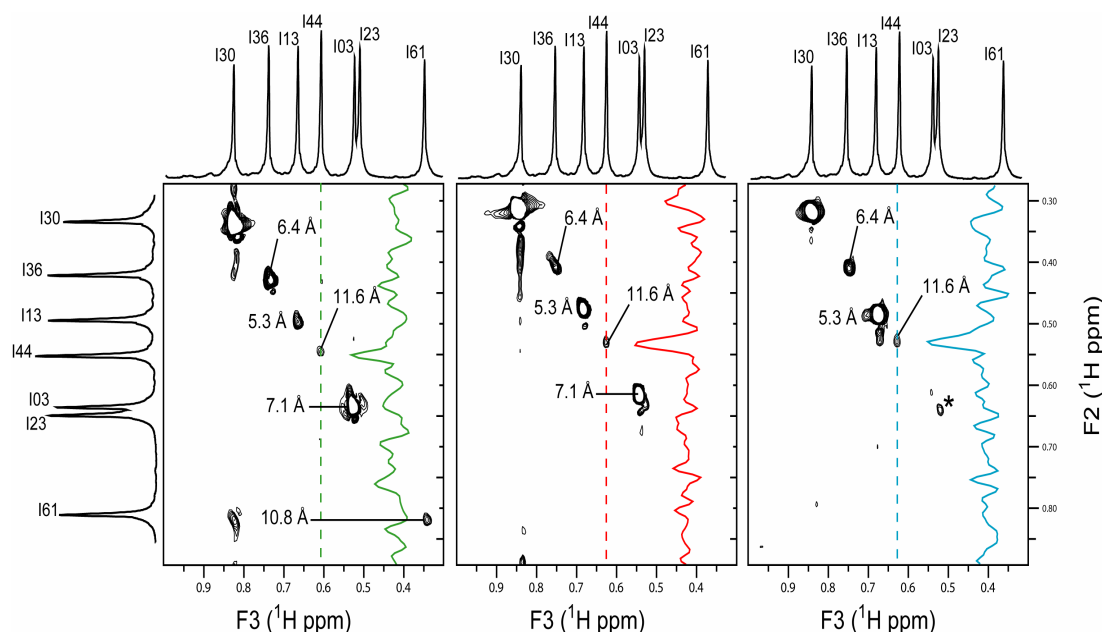


Figure 5.16: Plans 2D (F3,F2) extraits des spectres 3D enregistrés à partir de "3D Selective CT-COSY" avec un échantillon à 1 mM d'ubiquitine U-[^2H], U-[^{13}C], [$^{12}\text{CH}_3$]-Ile- δ_1 utilisé dans chacun des milieux d'alignement.

Les expériences ont été enregistrées pendant 12 heures sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec pour les temps de transfert : $T_A = T_B = 156$ ms. Les milieux liquides cristallins utilisés sont : à gauche une solution de DMPC/DHPC (3:1) à 5% ($D_a = 7,93$; $R = 0,174$) ; au centre un mélange de $\text{C}_{12}\text{E}_5/\text{n-hexanol}$ $D_a = 7,08$; $R = 0,280$) ; et à droite une solution de bactériophages (12 mg/mL avec 100mM NaCl, $D_a = 13,87$; $R = 0,610$). Les traces correspondant à la dimension F2 sont présentées sur chacun des plans et la distance interméthyle correspondant à chaque pic est notée sur les spectres. L'étoile correspond à des épaulements de pics se trouvant dans des plans voisins.

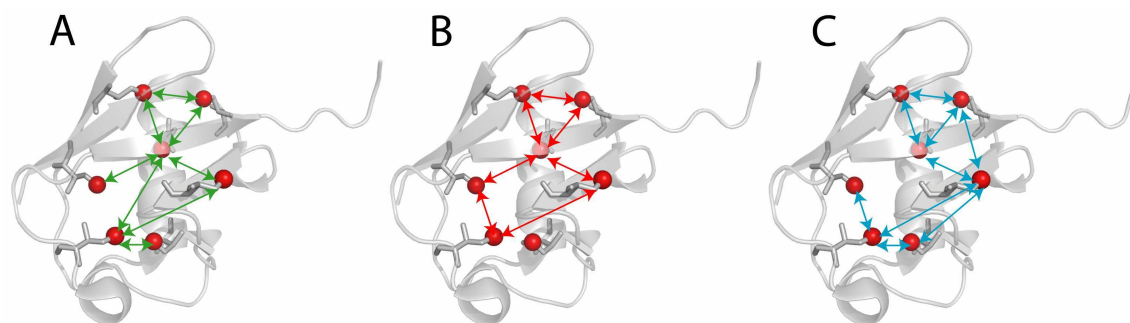


Figure 5.17: Les RDCs ^1H - ^1H entre méthyles δ_1 des Isoleucines de l'ubiquitine détectés dans les différents milieux d'alignement.

Ils sont indiqués par des doubles flèches. Les milieux liquides cristallins sont : en A une solution de DMPC/DHPC (3:1) à 5% ($D_a = 7,93$; $R = 0,174$) ; en B un mélange de $\text{C}_{12}\text{E}_5/\text{n-hexanol}$ $D_a = 7,08$; $R = 0,280$) ; en C une solution de bactériophages (12 mg/mL avec 100 mM NaCl, $D_a = 13,87$; $R = 0,610$).

Les couplages dipolaires entre méthyles non détectés sont dus à une mauvaise orientation du dipôle. Sur l'ensemble des mesures effectuées sur l'ubiquitine U- ^2H , U- ^{13}C , $^{12}\text{CH}_3$ -Ile- δ_1 , des couplages dipolaires entre paires de méthyles séparés jusqu'à 13 Å ont été mesurés. Le seuil de détection correspondant à la valeur des couplages dipolaires les plus faibles détectée est de $0,08 \pm 0,02$ Hz. Si l'on considère que l'orientation entre les deux méthyles est optimale, cette valeur seuil de couplage correspond à une distance maximale entre une paire de méthyle possible d'observer de 15 Å.

Corrélations des RDCs ^1H - ^1H avec la structures

Pour vérifier la précision des paramètres mesurés, les couplages entre méthyles δ_1 des Isoleucines (D_{HH} Expérimentale) de l'ubiquitine ont été calculés à partir d'un ensemble de structures. Pour déterminer ces couplages théoriques (D_{HH} Théorique), les paramètres de tenseur d'alignement de chacun des milieux d'alignement sont calculés à partir de la mesure des RDCs NH. Pour la mesure des RDCs $^1D_{\text{NH}}$, j'ai utilisé le même échantillon que pour la détection des RDCs interméthyles. Les échantillons utilisés sont placés dans une solution constituée à plus de 99% de D_2O . Les protons amides résiduels présent ont été utilisés pour mesurer les RDCs NH sur les trois échantillons. Pour un souci d'homogénéité avec la représentation habituelle, les signes des couplages expérimentaux ont été ajoutés en fonction du signe du couplage calculé (Wu *et al*, 2002). La comparaison des valeurs D_{HH} expérimentales avec les valeurs calculées est présentée sur la Figure 5.18 avec un coefficient de corrélation de Pearson $R_p = 0,90$.

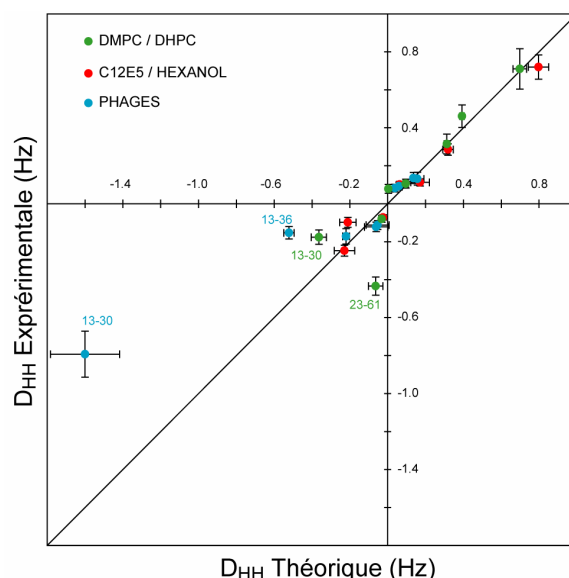


Figure 5.18: Comparaison des valeurs D_{HH} expérimentales et calculées dans l'ubiquitine U-[2H], U-[^{13}C], [$^{12}CH_3$]-Ile- δ_1 dans les différents milieux d'alignement.

Les différentes valeurs de D_{HH} obtenues à partir des trois milieux liquides cristallins sont représentées par le code couleur suivant : en vert la solution de DMPC/DHPC (3:1) à 5% ($Da = 7,93$; $R = 0,174$) ; en rouge le mélange de $C_{12}E_5$ /n-hexanol $Da = 7,08$; $R = 0,280$) et en bleu la solution de bactériophages (12 mg/mL avec 100mM NaCl, $Da = 13,87$; $R = 0,610$). La déviation standard des couplages prédits à partir de l'ensemble des structures (PDB: 1D3Z) est représentée par une barre d'erreur horizontale. La déviation standard pour les couplages expérimentaux $D_{HH}^{expérimentaux}$ est représentée par une barre d'erreur verticale et déterminée à partir de l'incertitude des mesures des intensités des pics.

On observe une bonne corrélation entre les valeurs expérimentales et valeurs calculées. Les valeurs de couplages corrélant le moins bien avec la structure concernent les paires de méthyles des Isoleucines suivantes : 13-30, 13-36 et 23-61. Elles correspondent aux distances les plus proches ($d_{HH} < 6 \text{ \AA}$), de plus l'orientation de ces vecteurs dans chacun des tenseurs d'alignement est proche de la valeur pour lesquels les couplages sont nuls. Cette double condition fait que les RDCs prédits pour ces vecteurs sont très sensibles à toutes les variations d'orientation. Pour s'en convaincre prenons l'exemple de la paire 23-61 (Figure 5.19). En supposant un modèle de rotation du méthyle par saut rapide entre les trois rotamères privilégiés, le couplage correspondant à cette paire peut se calculer comme la moyenne des neuf couplages entre paires de protons qui peuvent être formées entre ces deux méthyles. Les valeurs de chacun des couplages sont données à droite sur de la Figure 5.19. On remarque que les amplitudes de ces neuf couplages sont en moyenne 6 fois supérieures à la valeur moyenne. Toutes approximations sur la manière de moyenniser ces couplages, toutes variations dues à la dynamique de ces deux chaînes latérales

d'Isoleucines proches, ou toutes variations de l'orientation de ces deux méthyles aura une influence très importante sur la valeur moyenne. Cet exemple démontre la sensibilité de ces couplages aux variations angulaires.

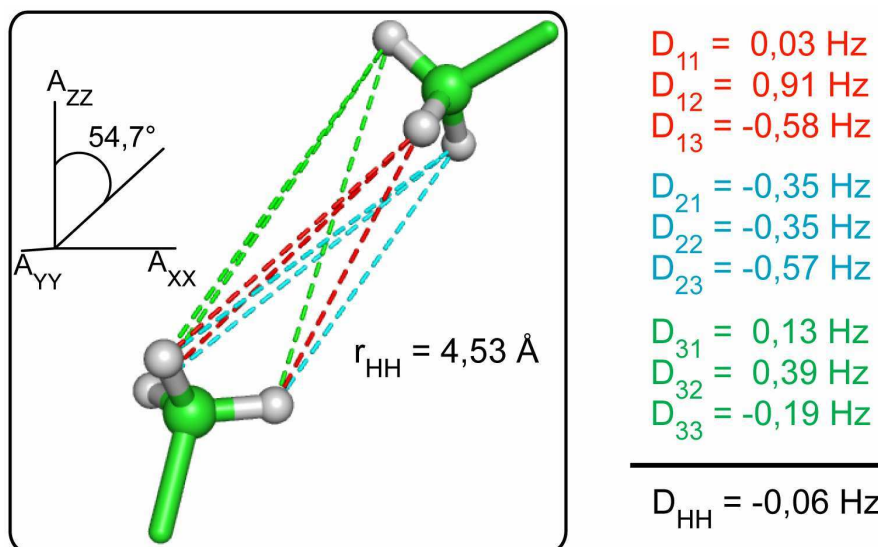


Figure 5.19: Schéma des 9 couplages de la paire de méthyle 23-61 de l'ubiquitine.

Les deux méthyles sont représentés relativement par rapport au tenseur d'alignement. La distance interméthyles est indiquée. Les valeurs des 9 couplages sont représenté à droite du schéma (le même code de couleur est utilisé pour le schéma et les valeurs de couplages) et la valeur moyenne du couplage est indiquée en noir. Les valeurs de couplages ont été calculé à partir de la structure de l'ubiquitine (PDB : 1D3Z) dans une solution de DMPC/DHPC (3:1) à 5% ($D_a = 7,93$; $R = 0,174$).

5.4.4.3 Application : γ S-crystalline U- $[^2\text{H}]$, $[^{12}\text{CH}_3]$ -Ile- δ_1

En collaboration avec le Professeur Zhengrong Wu (Université de l'Ohio, Columbus, USA), cette approche a été appliquée sur la γ S-crystalline U- $[^2\text{H}]$, $[^{12}\text{C}^1\text{H}_3]$ -Ile- δ_1 . C'est une protéine, de masse moléculaire de 19 kDa, avec un temps de corrélation de l'ordre de 17 ns en D_2O . La résolution est suffisante dans la région des méthyles δ_1 des Isoleucines pour la γ S-crystalline pour pouvoir utiliser un marquage isotopique ^{12}C sur le carbone. Le milieu d'alignement utilisé est un échantillon de phages dilué, placé dans un champ magnétique intense, qui a été figé dans son état orienté par l'amorce de la polymérisation de gel acrylamide. Par la suite, la solution de protéine avec la force ionique désirée est diffusée à l'intérieur de l'échantillon orienté. L'alignement de la γ S-crystalline est indépendant de la force du champ magnétique (Wu *et al*, 2005).

Une expérience 3D Selective Constant-time COSY (Figure 5.15) a été enregistrée pendant 12 heures sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec comme temps de transfert : $T_A = T_B = 76$ ms. L'échantillon utilisé n'étant pas marqué ^{13}C sur les autres carbones, les filtres n'ont aucun effet pour supprimer les pics de corrélation correspondant aux protons résiduels (§ 3.1.5). Des pics de corrélation correspondant à des RDCs ^1H - ^1H mesurés entre les méthyles δ_1 de la γS -crystalline séparés jusqu'à 9 Å peuvent être détectés (Figure 5.19).

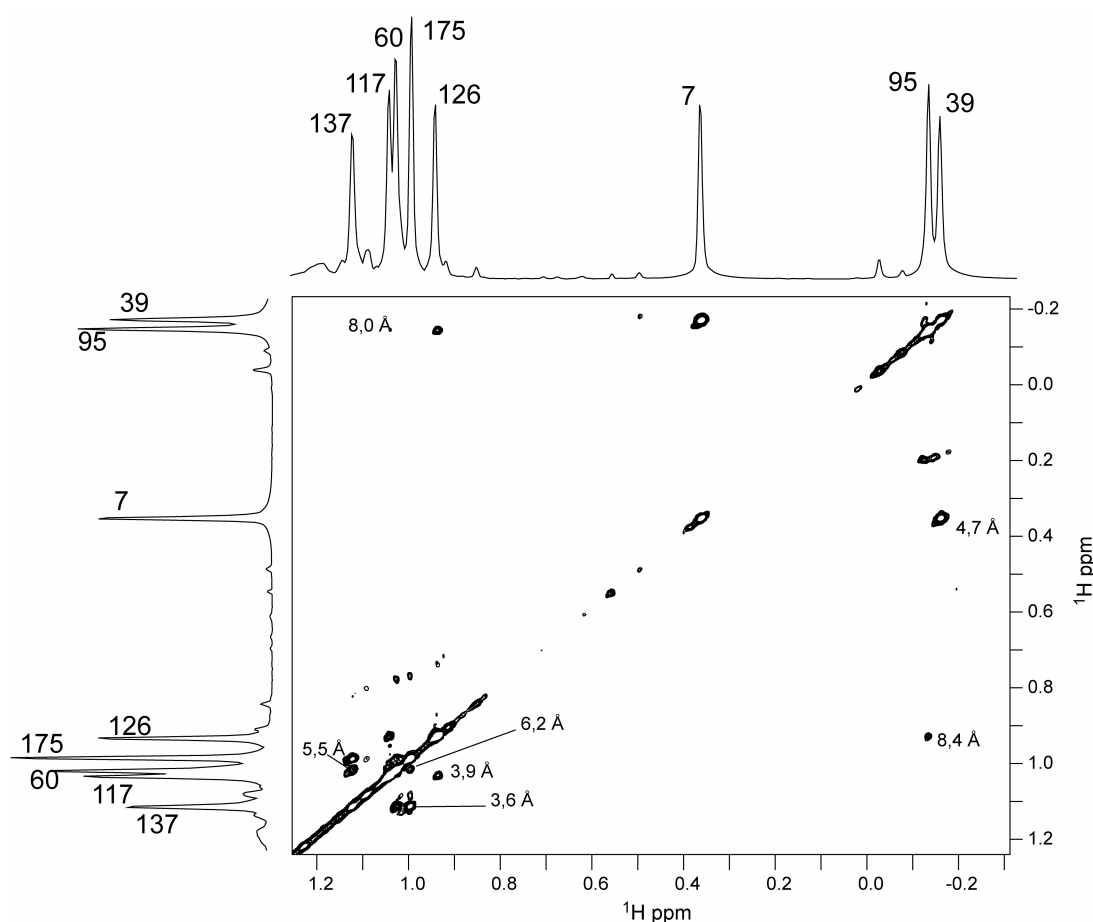


Figure 5.20: Spectre 2D de la région des méthyles δ_1 de la γS -crystalline enregistré à partir de la 3D Selective Constant-time COSY.

L'attribution des pics de corrélation est indiquée sur le spectre ainsi que la distance correspondante entre les paires de méthyles. L'expérience a été enregistrée à partir de la 3D Selective Constant-time COSY (Figure 5.15) à 25°C pendant 12 heures sur un spectromètre 800 MHz, équipé d'une sonde cryogénique, avec comme temps de transfert : $T_A = T_B = 76$ ms.

Cet exemple montre que cette approche n'est pas restreinte aux protéines de petite taille, et que des informations à longue portée peuvent toujours être observées dans les protéines de près de 20 kDa. Toutefois, ceci reste valable à condition d'avoir

suffisamment de résolution dans la région des méthyles δ_1 des Isoleucines pour utiliser un marquage U- ^2H , U- ^{13}C , [$^{12}\text{CH}_3$]-Ile- δ_1 , c'est-à-dire de conserver un carbone-12 sur le carbone δ_1 . Cette condition est importante pour garder les propriétés de relaxation les plus favorables. Pour des protéines de tailles plus importantes, le nombre de méthyles devient rapidement trop important pour pouvoir résoudre simultanément tous les pics diagonaux dans ces spectres 2D. Au contraire, des expériences de détection des NOEs à longue portée qui peuvent toujours être détectés dans des grosses protéines (chapitre 4). L'application des RDCs ^1H - ^1H interméthyles reste conseillée pour des protéines de taille modérée. La possibilité de marquer les différents types d'acides aminés contenant des méthyles permettra d'augmenter le nombre de contraintes longue portée interméthyles mesurables dans une protéine. En combinaison avec l'extraction de NOEs interméthyles cette approche pourra permettre de déterminer précisément la distance et l'orientation des vecteurs interméthyles dans les protéines.

Chapitre

6

Conclusions

Les méthodes standard sont, en générale, limitées à des informations locales (< 6 Å). Le manque de contraintes à longue portée est un des facteurs limitant pour les études structurales de protéines allongées, modulaires, de complexes protéiques et d'assemblages supramoléculaires. Pour étudier des systèmes de grande taille, il est nécessaire d'utiliser des échantillons deutérés. Toutefois, le remplacement des protons par des deutériums limite le nombre de données structurales, car les expériences NOESY ^1H - ^1H sont de manière générale la principale source de contraintes pour établir la structure de macromolécules par RMN.

C'est dans le but de répondre à ces deux problèmes, que s'inscrit le travail de cette thèse. La solution générale envisagée est de combiner un marquage isotopique spécifique d'un échantillon à des expériences RMN optimisées. Le choix des sondes pour détecter des contraintes à longue portée est primordial. En effet, obtenir un échantillon de protéine deutérée et spécifiquement protonée est onéreux. Il faut donc garder un nombre restreint de protons, dans des sites privilégiés, qui permettent d'extraire un maximum d'informations structurales importantes pour la détermination de structure. Ainsi, j'ai démontré que la protonation spécifique des méthyles bénéficie de plusieurs avantages et constitue un choix optimal pour l'obtention de contraintes structurales à longue portée. Pour ce faire, j'ai implémenté et amélioré des protocoles de marquage spécifique des méthyles des Isoleucines, des Valines et des Leucines. Pour un coût raisonnable, il m'a été possible de marquer spécifiquement les méthyles de ces résidus à un taux supérieur à 98% avec des fuites métaboliques limitées à moins de 2%.

L'utilisation de ce type de marquage a permis la détection de NOEs entre paire de méthyles séparés par des distances atteignant les 12 Å dans des systèmes de taille modérée. Ce type d'information ne peut pas être observé dans des échantillons de protéines standard. En effet, la détection de ces NOEs repose sur la diminution de la densité de proton à moins de 5% pour diminuer la diffusion de spin, source principale d'erreur sur la quantification des NOEs. L'utilisation d'échantillons faiblement protonés permet d'extraire les distances interméthyles avec précision. Dans le cas de protéines globulaires de taille modérée, ce type de contraintes, même peu nombreuses par rapport au nombre total de contraintes standard, permet d'améliorer les modèles de structures. L'utilisation de ce type d'échantillon en milieu

cristal liquide permet également d'extraire des informations à longue portée. J'ai démontré que des couplages dipolaires résiduels pouvaient être détectés entre des méthyles séparés plus de 13 Å. Bien que la mesure de ces RDCs reste limitée à des systèmes de taille modérée, elle offre des informations très sensibles sur l'orientation relative entre les méthyles.

La combinaison de la mesure de NOEs et de RDCs interméthyles permet de fournir des informations angulaires et translationnelles à longue portée non observable par les méthodes standard. Toutefois, le nombre limité de contraintes observables dans ce type d'échantillon spécifiquement marqué sur un seul type de méthyle est un désavantage de cette approche. C'est pourquoi, je me suis attaché à développer de nouveaux protocoles de marquage isotopiques. Un protocole efficace pour marquer spécifiquement les méthyles des Alanines a été développé, et d'autres concernant les Thréonines et les Méthionines sont en cours de mises au point. Il sera ainsi possible de marquer tous les méthyles des acides aminés dans une protéine de manière indépendante (ou combiner) afin d'extraire une multitude de contraintes translationnelles précises. Ceci permettra de déterminer les structures précises de protéines modulaires ou de complexes qui ne peuvent être résolues avec une très haute résolution avec les informations structurales standard.

Une analyse théorique et expérimentale du transfert NOE dans les systèmes de grandes tailles m'a permis de prédire que des contraintes structurales à longue portée étaient détectables dans des systèmes de plus de 100 kDa. Expérimentalement, j'ai pu vérifier ces résultats théoriques et détecter des NOEs entre méthyles séparés par plus de 7 Å dans des assemblages de près de 500 kDa. Ces résultats étaient encore inenvisageables en 2004 au début de cette thèse. Ces nouvelles méthodes de marquages isotopiques et les outils spectroscopiques associés, qui ont été développées au cours de cette thèse, ouvrent de nouveaux horizons à la RMN structurale pour l'étude de nano particules biologiques.

*Annexe**1**Les milieux M9*

Dans le chapitre 2, nous avons décrit différents protocoles de production de marquages spécifiques. Dans la suite de cette annexe, je vais détailler les milieux de cultures utilisés.

A1.1 Composition des milieux M9 pour la surexpression de protéines U-[²H]; U-[¹²C/¹³C] et spécifiquement protonée sur les méthyles des Isoleucines, des Valines, des Leucines et des Alanines.

A1.1.1 Milieu M9/H₂O

Quelques généralités à effectuer durant la culture. Une fois les solutions stocks autoclavées et les contenants stériles, il faut toujours travailler à la flamme.

Na ₂ HPO ₄ , 7H ₂ O	10 g	} Solutions Stock : H ₂ O 985 mL, pH = 7,2 et autoclaver Solution M9
KH ₂ PO ₄	3 g	
NaCl	0,5 g	
NH ₄ Cl	1 g	
Glucose (50 g/L)	8 mL	Solution stock Filtrée sur 0,22 µm
MgSO ₄ (1 M)	1 mL	Solution stock Autoclavée
CaCl ₂ (100 mM)	1 mL	Solution stock Autoclavée
MnCl ₂ (100 mM)	1 mL	Solution stock Autoclavée
ZnSO ₄ (50 mM)	1 mL	Solution stock Autoclavée
FeCl ₃ (50 mM)	1 mL	Solution stock Autoclavée
Thiamines (2 mg/mL)	1 mL	Solution stock Filtrée sur 0,22 µm
Antibiotiques (1 M) ^a	1 mL	Solution stock Filtrée sur 0,22 µm

^a Selon le système d'expression.

Lors des productions de l'ubiquitine, la quantité de glucose utilisée était 2 g.L⁻¹ 4 mL) ; par la suite, la production en milieu M9/H₂O s'est réalisée avec 4 g.L⁻¹ de glucose (8 mL).

A1.1.2 Milieu M9/D₂O_{99%}

Les quantités indiquées sont pour 1 L de M9/D₂O_{99%} ; ce milieu est utilisé pour l'adaptation en D₂O. Les différents produits lyophilisés sont solubilisé directement avec la solution M9/D₂O_{99%}. L'utilisation de 5 g.L⁻¹ de glycérol, comme source principale de carbone, peut être envisagée lors de la production d'un échantillon U-[²H]; U-[¹²C] à la place du glucose.

Na ₂ HPO ₄	10 g	} Solutions Stock :
KH ₂ PO ₄	3 g	
NaCl	0,5 g	
NH ₄ Cl	1 g	
		D ₂ O (99%) 1 L, pH = 7,2 et filtrer
		Solution M9/D ₂ O _{99%}
Glucose ^a	2 g	} Solutions stock Filtrées
MgSO ₄ (1 M)	1 mL	
CaCl ₂ (100 mM)	1 mL	
MnCl ₂ (100 mM)	1 mL	
ZnSO ₄ (50 mM)	1 mL	
FeCl ₃ (50 mM)	1 mL	
Thiamines (2 mg/mL)	1 mL	
Antibiotiques (1 M) ^b	1 mL	
		sur 0,22 µm et lyophilisées (3x)

^a Le glucose utilisé dépend de la culture : D-Glucose [²H, ¹³C] ou D-Glucose [²H, ¹²C].

^b Selon le système d'expression.

A1.1.3 Milieu M9/D₂O_{99,85%}

Les quantités indiquées sont pour 1 L de M9/D₂O_{99,85%} ; ce milieu est utilisé pour la culture en D₂O. Les différents produits lyophilisés sont solubilisés directement avec la solution M9/D₂O_{99,85%}. Toutefois, un volume prédéfini d'eau lourde, dépendant du précurseur, avant l'ajout des compléments, est gardé pour réaliser l'échange des protons du précurseur avec les deutériums du solvant (Figure 2.6 et 2.18). Une fois la réaction d'échange finie, les compléments du milieu M9 sont ajoutés.

Na ₂ HPO ₄ , anhydre	10 g	} <u>Solutions Stock :</u>
KH ₂ PO ₄ , anhydre	3 g	
NaCl, anhydre	0,5 g	
¹⁵ NH ₄ Cl ou NH ₄ Cl	1 g	
		D ₂ O (99,85 %) 1 L, pH = 7,2 et filtrer
		Solution M9/D ₂ O _{99,85%} (¹⁵ N)
Source de carbone ^a	X	} Solutions stock Filtrées
MgSO ₄ (1 M)	1 mL	
CaCl ₂ (100 mM)	1 mL	
MnCl ₂ (100 mM)	1 mL	
ZnSO ₄ (50 mM)	1 mL	
FeCl ₃ (50 mM)	1 mL	
Thiamines (2 mg/mL)	1 mL	
Antibiotiques (1 M) ^b	1 mL	
		sur 0,22 µm et lyophilisées (3x)

^a La source de carbone utilisé dépend du marquage : 2 g/L D-Glucose [²H, ¹³C] ou D-Glucose [²H, ¹²C] pour le marquage des Isoleucines, Valines et Leucines. 2,5 g/L de glycérol [²H, ¹³C] et 2,5 g/L de succinate [²H, ¹³C] pour le marquage des Alanines.

^b Selon le système d'expression.

*Annexe**2**Théorie de la relaxation pour
les méthyles*

A.2 Introduction

Lorsque l'on manipule, en RMN, un système de spins que l'on amène hors de son état d'équilibre par une (ou des) impulsion(s) radiofréquence(s), celui-ci évolue pour retourner vers son état d'équilibre thermodynamique. Ce retour à l'équilibre des spins nucléaires est un phénomène irréversible complexe, nommé sous le terme de relaxation. Lorsque l'on considère le système de spins couplés à son environnement, le plus souvent appelé "réseau", le phénomène de relaxation apparaît. Le système est composé des degrés de liberté de la molécule (mouvement de rotation, vibration et translation) et des autres systèmes présents dans l'échantillon. La présence du réseau modifie le champ local au niveau de chaque spin. Bien que faibles, ces champs locaux stochastiques possèdent des fréquences induisant des transitions entre les niveaux d'énergie de chaque spin, permettant ainsi le retour du système vers son état d'équilibre.

Les effets des corrélations, dans les systèmes composés de trois noyaux identiques de spin 1/2 (par exemple, le trio de proton d'un méthyle CH_3), ont attiré un fort taux d'intérêt dans des études passées. C'était le premier système de spins où des effets de corrélation croisée ont été étudiés par Hubbard, il y a presque un demi-siècle. Dans cette partie, je vais décrire les relaxations longitudinale et transversale d'un système AX_3 , équivalent à un méthyle $^{13}\text{CH}_3$ par exemple. Je vais présenter les équations des vitesses de relaxation de ce système ; elles sont particulièrement utilisées dans le chapitre 4. Ces vitesses de relaxation se simplifient pour un système A_3 , auquel on s'intéresse dans les chapitres 3 et 5, en ne tenant pas compte des contributions entre les protons et le carbone du méthyle.

A2.1.1 Le système AX_3 ($^{13}\text{CH}_3$)

La relaxation longitudinale

Comme nous l'avons vu dans le chapitre 4, les interactions principales responsables de la relaxation dans des groupements méthyles $^{13}\text{CH}_3$ sont de nature dipolaire entre les paires de spins ^1H - ^1H et ^1H - ^{13}C (Werbelow *et al*, 1973 ; 1977 ; Kay *et al*, 1992). Pour un méthyle dans une protéine perdeutérée, il faut tenir compte en plus des

interactions dipolaires ^1H - ^2H (que l'on notera dans la suite H-D) et ^1H - ^1H interméthyles.

Tous d'abord, avant de décrire les différentes interactions dipolaires, pour simplifier la notation, le coefficient d'interaction dipolaire ξ_{ab} (Equation A2.1) est utilisé dans les équations qui suivent.

$$\xi_{ab} = \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\hbar^2}{r_{ab}^6} \gamma_a^2 \gamma_b^2 \quad (\text{A2.1})$$

où γ_a et γ_b sont les rapports gyromagnétiques des noyaux a et b , r_{ab} la distance entre les deux noyaux a et b , μ_0 est la constante de perméabilité magnétique du vide et $\hbar$ est la constante de Planck divisée par 2π .

La vitesse de relaxation longitudinale $R_1(^1\text{H})$ des protons d'un méthyle (communément dénommé $1/T_1$), due aux interactions dipolaires, est donnée par:

$$R_1(^1\text{H}) = R_{1,i}^{HH} + R_{1,i}^{HC} + \sum R_1^{HH} + \sum R_1^{HD} \quad (\text{A2.2})$$

où l'indice i signifie que les interactions sont intraméthyles. Les autres interactions ont lieu entre les protons du méthyle et les autres noyaux, ^1H et ^2H , de la protéine. Nous considérons que les interactions avec les autres noyaux, ^{13}C et ^{15}N , sont négligeables comme ils ne font pas partie de la première sphère d'atomes voisins.

Détaillons maintenant ces différentes vitesses de relaxation longitudinale dues aux interactions dipolaires ; nous commençons par les contributions intraméthyles. Elles sont de deux types et sont données par les Equation A2.3 et A2.4.

$$R_{1,i}^{HH} = \frac{3}{2} \xi_{HH} [J_{HH}^{\text{intra}}(\omega_H) + 4J_{HH}^{\text{intra}}(2\omega_H)] \quad (\text{A2.3})$$

$$R_{1,i}^{CH} = \frac{\xi_{HC}}{4} [J_{HC}^{\text{intra}}(\omega_H - \omega_C) + 3J_{HC}^{\text{intra}}(\omega_H) + 6J_{HC}^{\text{intra}}(\omega_H + \omega_C)] \quad (\text{A2.4})$$

où $J_{ab}^{\text{intra}}(\omega)$ est la fonction de densité spectrale intraméthyle entre les atomes a et b . Les densités spectrales sont proportionnelles au temps de corrélation de la molécule, τ_C . On remarquera que la contribution décrite par l'Equation A2.4 disparaît dans le cas d'un système X_3 . On voit aussi, comme l'ensemble des vitesses de relaxation longitudinale dues aux interactions dipolaires, qu'elles sont dépendantes de la fonction de densité spectrale ($J_{ab}(\omega_H)$ et $J_{ab}(\omega_H \pm \omega_C)$).

Les deux autres contributions intervenant sur les protons des méthyles proviennent d'une part, des interactions avec les autres méthyles dans la molécule, $\sum R_1^{HH}$, et d'autre part, avec les deutériums, $\sum R_1^{HD}$.

$$R_1^{HH(\text{inter-méthyle})} = \frac{\xi_{HH}}{4} \left[J_{HH}^{\text{inter}}(0) + 3J_{HH}^{\text{inter}}(\omega_H) + 6J_{HH}^{\text{inter}}(2\omega_H) \right] \quad (\text{A2.5})$$

Le terme de densité spectrale $J_{HH}^{\text{inter}}(0)$ est le seul terme qui augmente avec le temps de corrélation de la molécule (τ_C).

Enfin, la dernière interaction entre les protons des méthyles et les deutériums est donnée par :

$$R_1^{HD} = \frac{\xi_{HD}}{4} \left[J_{HD}^{\text{inter}}(\omega_H - \omega_D) + 3J_{HD}^{\text{inter}}(\omega_H) + 6J_{HD}^{\text{inter}}(\omega_H + \omega_D) \right] \quad (\text{A2.6})$$

La relaxation transversale

Dans le cas d'un méthyle ayant une rotation rapide le long de son axe de symétrie dans des complexes protéiques de grande taille ($\tau_C \gg 10$ ns), il a été montré que la moitié de l'aimantation transversale des méthyles relaxe rapidement, suite à un phénomène de corrélation croisée (Werbelow *et al*, 1973 ; Kay *et al*, 1987). Ce phénomène est dû à une interférence entre les dipôles H_i-H_{ii} et H_i-H_{iii} dans le méthyle. L'autre moitié de l'aimantation ne relaxe que par les interactions $^1H-^1H$, $^1H-^2H$ interméthyle, ou $^1H-^{13}C$ intraméthyle.

La vitesse de relaxation transversale d'une cohérence simple quanta pour les protons du méthyles peut s'écrire comme la somme de deux composantes, la première relaxant rapidement $R_{2,H}^f$ et l'autre relaxant plus lentement $R_{2,H}^s$:

$$R_{2,H} = R_{2,H}^f + R_{2,H}^s \quad (\text{A2.7})$$

Si on fait l'hypothèse que $\omega_C \tau_C \gg 1$, le terme $J_{ab}(0)$ est prédominant devant tout les autres termes des fonctions de densité spectrale. Dans l'étude de systèmes de grandes tailles, cette hypothèse est vérifiée et l'on observe une dépendance avec la fonction de densité spectrale $J_{ab}(0)$ dans tous les termes des vitesses de relaxation transversale. La composante rapide peut s'écrire :

$$R_{2,H}^f = \left(\frac{9}{16}\right) \xi_{HH} [J_{HH}^{\text{intra}}(0)] + \left(\frac{1}{18}\right) \xi_{HC} [J_{HC}^{\text{intra}}(0)] + R_{2,\text{ext}} \quad (\text{A2.8})$$

Les deux premiers termes proviennent respectivement des interactions dipolaires intraméthyles $^1\text{H}-^1\text{H}$, et $^1\text{H}-^{13}\text{C}$, alors que le deuxième terme représente les contributions des spins externes H et D.

La seconde composante, celle qui relaxe plus lentement, s'écrit sous la forme:

$$R_{2,H}^s = \left(\frac{1}{18}\right) \xi_{HH} [J_{HC}^{\text{intra}}(0)] + R_{2,\text{ext}} \quad (\text{A2.9})$$

Toujours, dans le cas des systèmes de grandes tailles, la fonction de densité spectrale $J_{ab}(0)$ est prédominant devant les autres fonctions de densité spectrale $J_{ab}(\omega_C)$ et $J_{ab}(\omega_H)$. Ainsi, le terme tenant compte de la contribution des spins externes, $R_{2,\text{ext}}$, peut s'écrire de la manière suivante:

$$R_{2,\text{ext}} = \left(\frac{4}{3}\right) \sum_j \xi_{HD} [J_{HD}^{\text{inter}}(0)] + \left(\frac{5}{8}\right) \sum_j \xi_{HH} [J_{HH}^{\text{inter}}(0)] \quad (\text{A2.10})$$

A2.2 Fonctions de densité spectrale

Des expériences RMN de relaxation peuvent être employées pour sonder le mouvement inférieur de la nanoseconde des vecteurs (liaison) dans la structure moléculaire. Les paramètres RMN observables de relaxation, tels que les vitesses de relaxations longitudinale (R_1) et transversale (R_2) d'un spin nucléaire, sont directement liés à la densité spectrale, $J(\omega)$, décrivant le mouvement du vecteur d'interaction impliqué. La fonction de densité spectrale à la fréquence, ω , est directement liée, par transformée de Fourier, à la fonction de corrélation des interactions dipolaires dans le domaine des temps définissant ce mouvement dans le repère du laboratoire :

$$J(\omega) = \int_{-\infty}^{+\infty} C(\tau) e^{-i\omega\tau} d\tau \quad (\text{A2.11})$$

Si le mouvement global de réorientation de la molécule est isotrope et que l'on suppose que les mouvements internes de la molécule sont indépendants du mouvement global de réorientation, alors la fonction de corrélation peut s'exprimer comme le produit de deux fonctions de corrélation indépendantes. Cette hypothèse est satisfaite pour les macromolécules biologiques structurées dont le mouvement global est bien plus lent que les mouvements internes.

$$C(\tau) = C_g(\tau) \times C_i(\tau) \quad (\text{A2.12})$$

où $C_g(\tau)$ et $C_i(\tau)$ représentent respectivement le mouvement de diffusion rotationnelle de la molécule et le mouvement interne. Pour un vecteur diffusant librement, la fonction de corrélation du mouvement interne est caractérisée par un paramètre d'ordre (S^2) et un temps de corrélation caractéristique du mouvement interne τ_i (Lipari *et al*, 1982).

$$C_i(t) = S^2 + [P_2(\cos \theta) - S^2] e^{-t/\tau_i} \quad (\text{A2.13})$$

Où θ décrit l'angle du cône dans lequel le vecteur diffuse librement. $P_2(x)$ est le polynôme de Legendre d'ordre 2. Le paramètre d'ordre (S^2) représente la limite de la fonction de corrélation interne pour un temps infini, en l'absence de mouvement interne $S^2 = P_2(\cos\theta)$, et tend vers 0 pour un mouvement interne important.

Si l'on considère le modèle où le mouvement du groupement méthyle s'effectue par sauts aléatoires entre les trois sites équivalents et dans le cas d'un mouvement isotrope de la macromolécule, où est attaché le méthyle, alors la fonction de densité spectrale pour un groupement méthyle s'exprime :

$$J(\omega) = \frac{2}{5} \left[\frac{\tau_c S^2}{1 + (\omega\tau_c)^2} + \frac{\tau_{eff} [P_2(\cos\theta) - S^2]}{1 + (\omega\tau_{eff})^2} \right] \quad (\text{A2.14})$$

Avec $\tau_{eff}^{-1} = (\tau_c + \tau_i)^{-1}$. Pour un mouvement interne rapide, $\tau_i \ll \tau_c$, l'Equation A1.4 se simplifie en :

$$J(\omega) = \frac{2}{5} \frac{\tau_c S^2}{1 + (\omega\tau_c)^2} \quad (\text{A2.15})$$

Il faut remarquer qu'il existe un paramètre d'ordre par type de corrélation.

On distingue des corrélations intraméthyles (J_{HH}^{intra} et J_{HC}^{intra}) avec un S^2 faible, dû à la dynamique des méthyles et des corrélations interméthyles (J_{ab}^{inter}). Par défaut, on prend un $S^2 = 0,5$ (Tugarinov *et al*, 2003c ; 2005c) pour les corrélations intraméthyles. Compte tenu de la distance élevée entre les méthyles, le S^2 tend vers 1.

*Annexe**3**Protocole d'extraction des
distances interméthyles*

A3.1 Introduction

L'utilisation de la matrice de relaxation totale permet d'extraire les distances inter-méthyles en absence de référence interne aux spectres NOESY. L'obtention de la matrice de relaxation σ se fait à partir de $V(\tau_m)$ et de $V(0)$; cela revient à un simple problème de détermination de matrice.

$$\sigma = -(\ln[V(t_m).V(0)^{-1}])/\tau_m \quad (\text{A3.1})$$

Pour des raisons de calcul, l'équation A3.1 est plus facilement résolue en l'écrivant en termes de matrice des vecteurs propres, $[X]$, et en matrice diagonale des valeurs propres, $[D]$, de $V(\tau_m).V(0)^{-1}$:

$$\sigma = -([X](\ln[D])[X]^{-1})/\tau_m \quad (\text{A3.2})$$

La matrice des distances inter-méthyles est extraite à partir des termes hors diagonaux de la matrice de relaxation σ , grâce à l'Equation 3.13 (§3.1.2.4).

J'ai écrit un programme nommé RELOAD ("*full-RElaxation-matrix analysis for LQng-rAuge methyl methyl Distance measurement*") permettant de calculer directement les distances inter-méthyles à partir de l'Equation A3.2.

Le programme a été écrit en langage Python (Python 2.4 ou version supérieure) et est utilisé sur un système UNIX/Linux. Il nécessite la librairie numarray-1.5.1 (ou version supérieure, http://www.stsci.edu/resources/software_hardware/numarray).

A3.2 Les commandes

Le programme requiert des paramètres d'entrée et des options ; voici un exemple des commandes pour utiliser le programme :

```
RELOAD.py -inM1 Vmix.tab -inM0 V0.tab -outM out.txt -tc 16.7 -tmix
0.8 -w 800 -nb 8 -ile 7 39 60 95 117 126 137 175
```

A3.2.1 Les paramètres d'entrée requis

- -inM1 Le nom du fichier contenant les volumes des pics diagonaux et de corrélation, $V(\tau_m)$, de l'expérience NOESY, enregistrée à τ_m généré par le programme nmrDraw.

INTERMETHYLES

(voir en dessous pour le format du fichier)

ex: -inM1 Vmix.tab

- -inM0 Le nom du fichier contenant les volumes des pics diagonaux, $V(0)$, de l'expérience NOESY, enregistrée au temps de mélange égal à zéro, généré par le programme nmrDraw.
(voir en dessous pour le format du fichier)
ex: -inM0 V0.tab
- -outM Le nom du fichier de sortie.
ex: -outM out.txt
- -tc Le temps de corrélation globale de la protéine (τ_C) en ns.
ex: -tc 16.7
- -tmix Le temps de mélange (τ_m) de l'expérience NOESY en s.
ex: -tmix 0.8
- -nb Le nombre d'Isoleucines dans la protéine.
ex: -nb 8
- -ile L'attribution des Isoleucines dans la protéine.
ex: -ile 7 39 60 95 117 126 137
- -w La fréquence proton du spectromètre en MHz.
ex: -w 800

A3.2.2 Les options

- -int Utilise l'intensité des pics (par défaut, le programme utilise les volumes)
ex: -int
- -noise Estimation du bruit expérimental dans l'expérience inM1, requis pour la précision des distances dans le calcul Monte-Carlo.
ex: -noise 2000
- -it Le nombre d'itérations du calcul Monte-Carlo, si l'on utilise le bruit expérimental.
ex: -it 100

A3.2.3 Les formats des fichiers d'entrée

Comme nous venons de le voir, le programme RELOAD nécessite deux fichiers contenant les volumes des pics NOESY de $V(\tau_m)$ et de $V(0)$.

Il utilise l'attribution des Isoleucines donnée par le paramètre (-ile). Cette attribution doit être la même dans les fichiers d'entrée (-inM1 et -inM0).

Pour un pic diagonal *i* :

L'attribution : [nombre *i* du résidu] [nombre *i* du résidu].
ex: "7 7" (ligne 1 en dessous).

Pour un pic de corrélation NOE correspondant au transfert d'aimantation du méthyle *j* vers le méthyle *i* (*i* correspond au méthyle ayant sa fréquence proton détectée dans la dimension d'acquisition directe) :

L'attribution : [nombre *j* du résidu] [nombre *i* du résidu].
ex: "137 175" (ligne 9 en dessous).

Un exemple de ligne que l'on retrouve dans un fichier d'entrée :

```
VARS    INDEX X_AXIS Y_AXIS ... DHEIGHT VOL PCHI2 TYPE ASS ...
FORMAT %5d    %9.3f %9.3f ... %e      %e %5f %d %s ...

      1   552.410   276.563   ...   +7.719756e+04 +4.805753e+09 0.000000
1 007 007   ...
      9   345.660   152.109   ...   +7.410455e+04 +1.888468e+09 0.000000
1 137 175   ...
```

A3.2.4 Les formats des fichiers de sortie

```
Use the spectrometer proton frequency : 800.000000 MHz
Use this correlation time : 16.700000 ns
Use this mixing time : 0.800 s
the number of isoleucine : 8
Use this list of isoleucine delta 1 methyl : [7, 39, 60, 95, 117,
126, 137, 175]
Use an iteration procedure : 10
Noise using: 3500000
Distance   R (A)   dR(A)
  7-39      4.551   0.001
  7-60      9.836   0.205
 39-60     12.015   0.488
 60-137     4.970   0.001
 60-175     5.024   0.001
 95-117     8.293   0.080
 95-126     7.125   0.024
 95-137    10.140   0.291
 95-175    10.851   0.578
117-126     4.380   0.001
137-175     4.504   0.001
```

Références bibliographiques

A

- ♦Annala, A., Aitio, H., Thulin, E. and Drakenberg, T. **(1999)**. Recognition of protein folds via dipolar couplings. *J. Biomol. NMR.*, 14: 223-230.
- ♦Atreya, H. S. and Szyperski, T. **(2004)**. G-matrix fourier transform nmr spectroscopy for complete protein resonance assignment. *Proc. Natl. Acad. Sci. USA*, 101: 9642-9647.

B

- ♦Banci, L., Bertini, I., Bren, K. L., Cremonini, M. A., Gray, H. B., Luchinat, C. and Turano, P. **(1996)**. The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome c as an example. *J. B. I. C.*, 1: 117-126.
- ♦Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R. and Zhu, G. **(1994)**. Measurement of homonuclear and heteronuclear *J*-couplings from quantitative *J*-correlation. *Nuclear Magnetic Resonance, Pt C*, 239: 79-105.
- ♦Bax, A., Kontaxis, G. and Tjandra, N. **(2001)**. Dipolar couplings in macromolecular structure determination. *Methods Enzymol.*, 339: 127-74.
- ♦Bayer, P., Varani, L. and Varani, G. **(1999)**. Refinement of the structure of protein-RNA complexes by residual dipolar coupling analysis. *J. Biomol. NMR.*, 14: 149-155.
- ♦Bentrop, D., Bertini, I., Cremonini, M. A., Forsen, S., Luchinat, C. and Malmendal, A. **(1997)**. Solution structure of the paramagnetic complex of the N-terminal domain of calmodulin with two Ce³⁺ ions by ¹H NMR. *Biochemistry*, 36: 11605-11618.
- ♦Bersch, B., Rossy, E., Coves, J. and Brutscher, B. **(2003)**. Optimized set of two-dimensional experiments for fast sequential assignment, secondary structure determination, and backbone fold validation of ¹³C/¹⁵N-labelled proteins. *J. Biomol. NMR.*, 27: 57-67.
- ♦Bertini, I., Luchinat, C. and Piccioli, M. **(2001)**. Paramagnetic probes in metalloproteins. *Methods Enzymol.*, 339: 314-340.
- ♦Blackledge, M. **(2005)**. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog. Nucl. Magn. Reson. Spect.*, 46: 23-61.
- ♦Bloch, F. **(1956)**. Dynamical theory of nuclear induction II. *Phys. Rev.*, 102: 104-135.
- ♦Boelens, R., Koning, T. M. G. and Kaptein, R. **(1988)**. Determination of biomolecular structures from proton-proton NOEs using a relaxation matrix approach. *J. Mol. Struct.*, 173: 299-311.
- ♦Boelens, R., Koning, T. M. G., Van Der Marel, G. A., Van Boom, J. H. and Kaptein, R. **(1989)**. Iterative procedure for structure determination from proton-proton NOEs using a full relaxation matrix approach. Application to a DNA octamer. *J. Magn. Reson.*, 82: 290-308.
- ♦Boisbouvier, J., Gans, P., Blackledge, M., Brutscher, B. and Marion, D. **(1999)**. Long-range structural information in NMR studies of paramagnetic molecules from

electron spin-nuclear spin cross-correlated relaxation. *J. Am. Chem. Soc.*, 121: 7700-7701.

♦Boisbouvier, J., Delaglio, F. and Bax, A. (2003). Direct observation of dipolar couplings between distant protons in weakly aligned nucleic acids. *Proc. Natl. Acad. Sci. USA*, 100: 11333-8.

♦Borgias, B. A. and James, T. L. (1990). MARDIGRAS - A procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J. Magn. Reson.*, 87: 475-487.

♦Borissenko, L. and Groll, M. (2005). Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J. Mol. Biol.*, 346: 1207-1219.

♦Bothner-By, A. A. (1996). Magnetic field induced alignment of molecules. *Encyclopedia of Nuclear Magnetic Resonance*, 2932-2938.

♦Boudet, J. (2007) Etude RMN structurale et fonctionnelle de macromolécules biologiques impliquées dans des modulations de pathogénicité bactérienne et virale. *Thèse de doctorat de l'Université J. Fourier - Grenoble - France*

♦Bouvignies, G., Meier, S., Grzesiek, S. and Blackledge, M. (2006a). Ultrahigh-resolution backbone structure of perdeuterated protein GB1 using residual dipolar couplings from two alignment media. *Angew. Chem. Int. Ed.*, 45: 8166-8169.

♦Bouvignies, G., Markwick, P., Brüschweiler, R. and Blackledge, M. (2006b). Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. *J. Am. Chem. Soc.*, 128: 15100-151001.

♦Bouvignies, G. (2007) Etudes analytiques de la structure et de la dynamique des protéines en solution à partir des couplages dipolaires résiduels. *Thèse de doctorat de l'Université J. Fourier - Grenoble - France*

♦Brüschweiler, R., Liao, X. and Wright, P. E. (1995). Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science*, 268: 886-889.

♦Brutscher, B. (2001). Accurate measurement of small spin-spin couplings in partially aligned molecules using a novel J-mismatch compensated spin-state-selection filter. *J. Magn. Reson.*, 151: 332-338.

♦Busso, D., Kim, R. and KIM, S. H. (2003). Expression of soluble recombinant proteins in a cell-free system using a 96-well format. *J. Biochem. biophys. methods*, 55: 233-240.

C

♦Cai, M. L., Wang, H., Olejniczak, E. T., Meadows, R. P., Gunasekera, A. H., Xu, N. and Fesik, S. W. (1999). Accurate measurement of H_N - H_α residual dipolar couplings in proteins. *J. Magn. Reson.*, 139: 451-453.

♦Carlomagno, T., Peti, W. and Griesinger, C. (2000). A new method for the simultaneous measurement of magnitude and sign of $^1D_{CH}$ and $^1D_{HH}$ dipolar couplings in methylene groups. *J. Biomol. NMR*, 17: 99-109.

- ♦Carver, T. R. and Slitcher, C. P. **(1953)**. Polarization of nuclear spins in metal. *Phys. Rev.*, 92: 212.
- ♦Chou, J. J., Case, D. A. and Bax, A. **(2003)**. Insights into the mobility of methyl-bearing side chains in proteins from $^3J_{CC}$ and $^3J_{CN}$ couplings. *J. Am. Chem. Soc.*, 125: 8959-8966.
- ♦Christodoulou, J., Larsson, G., Fucini, P., Connell, S. R., Pertinhez, T. A., Hanson, C. L., Redfield, C., Nierhaus, K., Robinson, C. V., Schleucher, J. and Dobson, C. M. **(2004)**. Heteronuclear NMR investigations of dynamic regions of intact *Escherichia coli* Ribosomes. *Proc. Natl. Acad. Sci. USA*, 101: 10949-10954.
- ♦Clore, G. M., Starich, M. R., Bewley, C. A., Cai, M. L. and Kuszewski, J. **(1999)**. Impact of residual dipolar couplings on the accuracy of NMR structures determined from a minimal number of NOE restraints. *J. Am. Chem. Soc.*, 121: 6513-6514.
- ♦Cordier, F., Caffrey, M., Brutscher, B., Cusanovich, M. A., Marion, D. and Blackledge, M. **(1998)**. Solution structure, rotational diffusion anisotropy and local backbone dynamics rhodobacter capsulatus cytochrome c2. *J. Mol. Biol.*, 281: 341-361.
- ♦Cordier, F. and Grzesiek, S. **(1999)**. Direct observation of hydrogen bonds in proteins by interresidue $^3J_{NC'}$ scalar couplings. *J. Am. Chem. Soc.*, 121: 1601-1602.
- ♦Cornilescu, G., Marquardt, J. L., Ottiger, M. and Bax, A. **(1998)**. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.*, 120: 6836-6837.
- ♦Crespi, H. L., Rosenberg, R. M. and Katz, J. J. **(1968)**. Proton magnetic resonance studies of fully deuterated except for 1H -Leucine side chains. *Science*, 161: 795-796.

D

- ♦de Alba, E. and Tjandra, N. **(2002)**. NMR dipolar couplings for the structure determination of biopolymers in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 40: 175-197.
- ♦Delaglio, F., Kontaxis, G. and Bax, A. **(2000)**. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.*, 122: 2142-2143.
- ♦Diercks, T., Daniels, M. and Kaptein, R. **(2005)**. Extended flip-back schemes for sensitivity enhancement in multidimensional HSQC-type out-and-back experiments. *J. Biomol. NMR*, 33: 243-259.
- ♦Douglas, S. M., Chou, J. J. and Shih, W. M. **(2007)**. From the cover: DNA-nanotube-induced alignment of membrane proteins for NMR. *Proc. Natl. Acad. Sci. USA*, 104: 6644-6648.

E

- ♦Ernst, R. R., Bodenhausen, G. and Wokaun, A. **(1987)**. Principles of nuclear magnetic resonance in one and two dimensions. *Oxford University Press, Oxford*.

F

- ♦Fiaux, J., Bertelsen, E. B., Horwich, A. L. and Wüthrich, K. **(2002)**. NMR analysis of a 900 kDa GroEL GroES complex. *Nature*, 418: 207-211.
- ♦Fiaux, J., Bertelsen, E. B., Horwich, A. L. and Wüthrich, K. **(2004)**. Uniform and residue-specific ^{15}N -labeling of proteins on a highly deuterated background. *J. Biomol. NMR.*, 29: 289-297.
- ♦Fischer, M., Kloiber, K., Häusler, J., Ledolter, K., Konrat, R. and Schmid, W. **(2007)**. Synthesis of ^{13}C -methyl-group labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *Chembiochem*, 8: 610-612.

G

- ♦Gardner, K. H. and Kay, L. E. **(1997a)**. Production and incorporation of ^{15}N , ^{13}C , ^2H (^1H - δ_1 methyl) isoleucine into proteins for multidimensional NMR studies. *J. Am. Chem. Soc.*, 119: 7599-7600.
- ♦Gardner, K. H., Rosen, M. K. and Kay, L. E. **(1997b)**. Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry*, 36: 1389-1401.
- ♦Gardner, K. H. and Kay, L. E. **(1998)**. The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Ann. Rev. Biophys. Biomol. Struct.*, 27: 357-406.
- ♦Geen, H. and Freeman, R. **(1991)**. Band-selective radiofrequency pulses. *J. Magn. Reson.*, 93: 93-141.
- ♦Gochin, M. and Roder, H. **(1995)**. Protein structure refinement based on paramagnetic NMR shifts: Application to wild-type and mutants of cytochrome c. *Prot. Science*, 4: 296-305.
- ♦Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C. and Kay, L. E. **(1999)**. A robust and cost-effective method for the production of Val, Leu, Ile (δ_1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J. Biomol. NMR.*, 13: 369-374.
- ♦Griffey, R. H. and Redfield, A. G. **(1985)**. Nuclear magnetic resonance observation and dynamics of specific amide protons in T4 lysozyme. *Biochemistry*, 24: 817-822.
- ♦Grishaev, A., Wu, J., Trewella, J. and Bax, A. **(2005)**. Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. *J. Am. Chem. Soc.*, 127: 16621-16628.
- ♦Grishaev, A., Tugarinov, V., Kay, L. E., Trewella, J. and Bax, A. **(2008)**. Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. *J. Biomol. NMR*, 40, 95-106.
- ♦Guéron, M. **(1975)**. Nuclear relaxation in macromolecules by paramagnetic ions: A novel mechanism. *J. Magn. Reson.*, 19: 58-66.

H

- ♦Hansen, M. R., Rance, M. and Pardi, A. **(1998)**. Observation of long-range ^1H - ^1H distances in solution by dipolar coupling interactions. *J. Am. Chem. Soc.*, 120: 11210-11211.

- ♦Hennig, M., Carlomagno, T. and R., W. J. **(2001)**. Residual dipolar coupling TOCSY for direct through space correlations of base protons and phosphorus nuclei in RNA. *J. Am. Chem. Soc.*, 123: 3395-3396.
- ♦Horst, R., Wider, G., Fiaux, J., Bertelsen, E. B., Horwich, A. L. and Wüthrich, K. **(2006)**. Proton-proton Overhauser NMR spectroscopy with polypeptide chains in large structures. *Proc. Natl. Acad. Sci. USA*, 103: 15445-15450.
- ♦Howard, B. R., Endrizzi, J. A. and Remington, S. J. **(2000)**. Crystal structure of *E. coli* Malate Synthase G complexed with magnesium and glyoxylate at 2.0 Å resolution: Mechanistic implications. *Biochemistry*, 39: 3156-3168.
- ♦Hus, J. C., Marion, D. and Blackledge, M. **(2000)**. De novo determination of protein structure by NMR using orientational and long-range order restraints. *J. Mol. Biol.*, 298: 927-936.
- ♦Hus, J. C., Marion, D. and Blackledge, M. **(2001)**. Determination of protein backbone structure using only residual dipolar couplings. *J. Am. Chem. Soc.*, 123: 1541-1542.

I

- ♦Ikura, M., Kay, L. E. and Bax, A. **(1990)**. A novel approach for sequential assignment of proton, ^{13}C , and ^{15}N spectra of larger proteins: heteronuclear triple-resonance three dimensional NMR spectroscopy. *Biochemistry*, 29: 4659-4667.
- ♦Isaacson, R. L., Simpson, P. J., Liu, M., Cota, E., Zhang, X., Freemont, P. and Matthews, S. **(2007)**. A new labelling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of Alanine residues. *J. Am. Chem. Soc.*, 129: 15428-15429.
- ♦Iwahara, J., Anderson, E., Murphy, E. C. and Clore, M. **(2003)**. EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J. Am. Chem. Soc.*, 125: 6634-6635.
- ♦Iwahara, J., Zweckstetter, M. and Clore, G. M. **(2006)**. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proc. Natl. Acad. Sci. USA*, 103: 15062-15067.

J

- ♦Jeener, J., Meier, B. H., Bachmann, P. and Ernst, R. R. **(1979)**. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem Phys.*, 71: 4546-4553.

K

- ♦Kaikkonen, A. and Otting, G. **(2001)**. Residual dipolar ^1H - ^1H couplings of methyl groups in weakly aligned proteins. *J. Am. Chem. Soc.*, 123: 1770-1771.
- ♦Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Ono, A. M. and Güntert, P. **(2006)**. Optimal isotope labelling for NMR protein structure determinations. *Nature*, 440: 52-57.

- ♦Kalbitzer, H. R., Leberman, R. and Wittinghofer, A. **(1985)**. ^1H -NMR spectroscopy on elongation factor Tu from *E. coli*. *F.E.B.S. Lett.*, 180: 40-42.
- ♦Karplus, M. J. **(1959)**. Contact electron-spin couplings of nuclear magnetic moments. *J. Chem Phys.*, 30: 11-15.
- ♦Kay, L. E. and Prestegard, J. H. **(1987)**. Methyl-group dynamics from relaxation of double quantum filtered NMR signals - Application to deoxycholate. *J. Am. Chem. Soc.*, 109: 3829-3835.
- ♦Kay, L. E., Clore, M., Bax, A. and Gronenborn, A. **(1990)**. Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 β in solution. *Science*, 249: 411-414.
- ♦Kay, L. E. and Bull, T. E. **(1992)**. Heteronuclear transverse relaxation in AMX, AX₂ and AX₃ spin systems. *J. Magn. Reson.*, 99: 615-622.
- ♦Keepers, J. W. and James, T. L. **(1984)**. A theoretical study of distance determination from NMR. Two dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.*, 57: 404-426.
- ♦Kigawa, T. and Yokoyama, S. **(1991)**. A continuous cell-free protein synthesis system for coupled transcription-translation. *J. Biochem.*, 110: 166-168.
- ♦Kigawa, T., Muto, Y. and Yokoyama, S. **(1995)**. Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J. Biomol. NMR.*, 6: 129-134.
- ♦Klammt, C., Schwarz, D., Löhr, F., Schneider, B., Dötsch, V. and Bernhard, F. **(2006)**. Cell-free expression as an emerging technique for the large scale production of integral membrane protein. *F.E.B.S.*, 273: 4141-4153.
- ♦Koehl, P. and Lefèvre, J.-F. **(1990)**. The reconstitution of the relaxation matrix from an incomplete set of nuclear Overhauser effects. *J. Magn. Reson.*, 86: 565-583.
- ♦Koharudin, L. M. I., Bonvin, A. M. J. J., Kaptein, R. and Boelens, R. **(2003)**. Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination. *J. Magn. Reson.*, 163: 228-235.

L

- ♦Lee, K. M., Androphy, E. J. and Baleja, J. D. **(1995)**. A novel method for selective isotope labeling of bacterially expressed proteins. *J. Biomol. NMR.*, 5: 93-96.
- ♦Lee, A. L., Flynn, P. F. and Wand, A. J. **(1999)**. Comparaison of ^2H and ^{13}C relaxation techniques for the study of protein methyl group dynamics in solution. *J. Am. Chem. Soc.*, 121: 2891-2902.
- ♦Lefèvre, J.-F., Lane, A. N. and Jardetsky, O. **(1987)**. Solution structure of the Trp operator of *Escherichia coli* determined by NMR. *Biochemistry*, 26: 5076-5090.
- ♦LeMaster, D. M. and Richards, F. M. **(1988)**. NMR sequential assignment of *E. coli* thioredoxin utilizing random fractional deuteration. *Biochemistry*, 27: 142-150.
- ♦Levitt, M. H. **(2001)**. Spin dynamics: basics of Nuclear Magnetic Resonance.

- ♦Lu, P., Jarema, M., Mosser, K. and Daniel, W. E. J. **(1976)**. lac repressor: 3-Fluorotyrosine substitution for Nuclear Magnetic Resonance studies. *Proc. Natl. Acad. Sci. USA*, 73: 3471-3475.
- ♦Lukin, J. A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A. and Ho, C. **(2003)**. Quaternary structure of hemoglobin in solution. *Proc. Natl. Acad. Sci. USA*, 100: 517-520.

M

- ♦Mal, T. K., Matthews, S. J., Kovacs, H., Campbell, I. D. and Boyd, J. **(1998)**. Some NMR experiments and a structure determination employing a $\{^{15}\text{N}, ^2\text{H}\}$ enriched protein. *J. Biomol. NMR*, 12: 259-276.
- ♦Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. and Bax, A. **(1989a)**. three-dimensional heteronuclear NMR of ^{15}N -labelled proteins. *J. Am. Chem. Soc.*, 111: 1515-1517.
- ♦Marion, D., Ikura, M., Tschubin, R. and Bax, A. **(1989b)**. Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.*, 85: 393-399.
- ♦Markley, J. L., Putter, I. and Jardetsky, O. **(1968)**. High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. *Science*, 161: 1249-1251.
- ♦Matsuda, T., Koshiba, S., Tochio, N., Seki, E., Iwasaki, N., Yabuki, T., Inoue, M., Yokoyama, S. and Kigawa, T. **(2007)**. Improving cell-free protein synthesis for stable isotope labeling. *J. Biomol. NMR*, 37: 225-229.
- ♦Meier, S., Häussinger, D., Jensen, P., Rogowski, M. and Grzesiek, S. **(2003)**. High-accuracy residual $^1\text{H}_\text{N}$ - ^{13}C and $^1\text{H}_\text{N}$ - $^1\text{H}_\text{N}$ dipolar couplings in perdeuterated proteins. *J. Am. Chem. Soc.*, 125: 44-45.
- ♦Miclet, E., Williams, J., D.C., Clore, G. M., Bryce, D. L., Boisbouvier, J. and Bax, A. **(2004)**. Relaxation-optimized NMR spectroscopy of methylene groups in proteins and nucleic acids. *J. Am. Chem. Soc.*, 126: 10560-10570.
- ♦Miclet, E., Boisbouvier, J. and Bax, A. **(2005)**. Measurement of eight scalar and dipolar couplings for methine-methylene pairs in proteins and nucleic acids. *J. Biomol. NMR*, 31: 201-216.
- ♦Mirau, P. A. **(1988)**. Quantitative interpretation of a single NOESY spectrum. *J. Magn. Reson.*, 80: 439-444.
- ♦Mollova, E. T. and Pardi, A. **(2000)**. NMR solution structure determination of RNAs. *Current Opinion in Structural Biology*, 10: 298-302.
- ♦Morita, E. H., Shimizu, M., Ogasawara, T., Endo, Y., Tanaka, R. and Kohno, T. **(2004)**. A novel way of amino acid-specific assignment in ^1H - ^{15}N HSQC spectra with a wheat germ cell-free protein synthesis system. *J. Biomol. NMR*, 30: 37-45.
- ♦Muchmore, D. C., McIntosh, L. P., Russel, C. B., Anderson, D. E. and Dahlquist, F. W. **(1989)**. Expression and ^{15}N labelling of proteins for proton and ^{15}N Nuclear Magnetic Resonance. *Methods Enzymol*, 177: 44-73.

N

- ♦Nibeilliu, M. E. and Malthouse, P. G. **(2004)**. The stereospecificity and catalytic efficiency of the tryptophan synthase-catalysed exchange of the α -protons of amino acids. *Biochem. J.*, 381: 847-852.
- ♦Nielsen, N. C., Bildsoe, H., Jakobsen, H. J. and Sorensen, O. W. **(1989)**. Composite refocusing sequences and their application for sensitivity enhancement and multiplicity filtration in INEPT and 2D correlation spectroscopy. *J. Magn. Reson.*, 85: 359-380.
- ♦Nilges, M., Habazettl, J., Brünger, A. T. and Holak, T. A. **(1991)**. Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J. Mol. Biol.*, 219: 491-510.

O

- ♦Ottiger, M., Delaglio, F. and Bax, A. **(1998)**. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, 131: 373-8.

P

- ♦Permi, P. and Annala, A. **(2000)**. Transverse relaxation optimised spin-state selective NMR experiments for measurement of residual dipolar couplings. *J. Biomol. NMR.*, 16: 221-7.
- ♦Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. **(1997)**. Attenuated T2 relaxation by mutual cancellation of dipole- dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. USA*, 94: 12366-12371.
- ♦Pervushin, K., Vogeli, B. and Eletsky, A. **(2002)**. Longitudinal ^1H relaxation optimization in TROSY NMR spectroscopy. *J. Am. Chem. Soc.*, 124: 12898-12902.
- ♦Peti, W. and Griesinger, C. **(2000)**. Measurement of magnitude and sign of H,H-dipolar couplings in proteins. *J. Am. Chem. Soc.*, 122: 3975-3976.
- ♦Prestegard, J. H. **(1998)**. New techniques in structural NMR - Anisotropic interactions. *Nat. Struct. Biol.*, 5: 517-522.
- ♦Prestegard, J. H. and Kishore, A. I. **(2001)**. Partial alignment of biomolecules: an aid to NMR characterization. *Current Opinion in Chemical Biology*, 5: 584-590.
- ♦Prestegard, J. H., Bougault, C. M. and Kishore, A. I. **(2004)**. Residual dipolar couplings in structure determination of biomolecules. *Chem. rev.*, 104: 3519-3540.

R

- ♦Rajesh, S., Nietlispach, D., Nakayama, H., Takiob, K., Laue, E. D., Shibata, T. and Ito, Y. **(2003)**. A novel method for the biosynthesis of deuterated proteins with selective protonation at the aromatic rings of Phe, Tyr and Trp. *J. Biomol. NMR.*, 27: 81-86.

- ♦Ramesh, V., Frederick, R. O., Syed, S. E., Gibson, C. F., Yang, J. C. and Roberts, G. C. **(1994)**. The interactions of *Escherichia coli* trp repressor with tryptophan and with an operator oligonucleotide. NMR studies using selectively ^{15}N -labelled protein. *Eur. J. Biochem.*, 225: 601-608.
- ♦Redfield, A. G. **(1957)**. On the theory of relaxation processes. *Journal of Research and development*, 1: 19-31.
- ♦Redfield, A. G. **(1965)**. the theory of relaxation processes. *Ad. Magn. Reson.*, 1: 1-32.
- ♦Reisman, J., Jariel-Encontre, I., Hsu, V. L., Parello, J., Geiduschek, E. P. and Kearns, D. R. **(1991)**. Improving two dimensional ^1H NMR NOESY spectra of a large protein by selective deuteration. *J. Am. Chem. Soc.*, 113: 2787-2789.
- ♦Riek, R., Wider, G., Pervushin, K. and Wüthrich, K. **(1999)**. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proc. Natl. Acad. Sci. USA*, 96: 4918-4923.
- ♦Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T. and Kay, L. E. **(1996)**. Selective methyl group protonation of perdeuterated proteins. *J. Mol. Biol.*, 263: 627-636.
- ♦Rüdiger, S., Freund, S. M. V., Veprintsev, D. B. and Fresht, A. R. **(2002)**. CRINEPT-TROSY NMR reveals p53 core domain bound in an unfolded form to the chaperone Hsp90. *Proc. Natl. Acad. Sci. USA*, 99: 11085-11090.

S

- ♦Salzmann, M., Wider, G., Pervushin, K., Senn, H. and Wuthrich, K. **(1999)**. TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.*, 121: 844-848.
- ♦Sattler, M. and Fesik, S. W. **(1996)**. Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure*, 4: 1245-1249.
- ♦Sawasaki, T., Ogasawara, T., Morishita, R. and Endo, Y. **(2002)**. A cell-free protein synthesis systems for high-throughput proteomics. *Proc. Natl. Acad. Sci. USA*, 99: 14652-14657.
- ♦Schanda, P., Kupce, E. and Brutscher, B. **(2005)**. SOFAST-HMQC experiments for recording two dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR.*, 33: 199-211.
- ♦Schanda, P., Van Melckebeke, H. and Brutscher, B. **(2006)**. Speeding up three-dimensional protein NMR experiments to a few minutes. *J. Am. Chem. Soc.*, 128: 9042-9043.
- ♦Schanda, P. **(2007)** Development and application of fast NMR methods for the study of protein structure and dynamics. *de l'Université Joseph Fourier - Grenoble*
- ♦Schanda, P., Lescop, E., Falge, M., Sounier, R., Boisbouvier, J. and Brutscher, B. **(2007)**. Sensitivity-optimized experiment for the measurement of residual dipolar couplings between amide protons. *J. Biomol. NMR.*, 38: 47-55.
- ♦Sibille, N., Pardi, A., Simorre, J. P. and Blackledge, M. **(2001)**. Refinement of local and long-range structural order in theophylline-binding RNA using ^{13}C - ^1H residual

dipolar couplings and restrained molecular dynamics. *J. Am. Chem. Soc.*, 123: 12135-12146.

♦Sibille, N., Bersch, B., Coves, J., Blackledge, M. and Brutscher, B. (2002). Side chain orientation from methyl ^1H - ^1H residual dipolar couplings measured in highly deuterated proteins. *J. Am. Chem. Soc.*, 124: 14616-14625.

♦Smith, B. O., Ito, Y., Raine, A., Teichmann, S., Ben-Tovim, L., Nietlispach, D., Broadhurst, R. W., Terada, T., Kelly, M., Oschkinat, H., Shibata, T., Yokoyama, S. and Laue, E. D. (1996). An approach to global fold determination using limited NMR data from larger proteins selectively protonated at specific residue types. *J. Biomol. NMR.*, 8: 360-368.

♦Solomon, I. (1955). Relaxation processes in a system of two spins. *Phys. Rev.*, 99: 559.

♦Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y. and Alakhov, Y. B. (1988). A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*, 242: 1162-1164.

♦Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. and Kay, L. E. (2005). Quantitative NMR spectroscopy of supramolecular complexes: Dynamics side pores in ClpP are important for product release. *Proc. Natl. Acad. Sci. USA*, 102: 16678-16683.

♦Sprangers, R. and Kay, L. E. (2007). Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, 445: 618-622.

T

♦Tang, C., Iwahara, J. and Clore, G. M. (2006). Visualization of transient encounter complexes in protein-protein association. *Nature*, 444: 383-386.

♦Tang, C., Schwieters, C. D. and Clore, G. M. (2007). Open-to-closed transition in apo maltose binding protein observed by paramagnetic NMR. *Nature*, 449: 1078-1082.

♦Tenenbaum, S. W., Witherup, T. H. and Abbott, E. H. (1974). Selectivity in vitamin B-6 model reactions: Selective α or β deuteration of amino acids under transamination or racemization conditions. *Biochim. Biophys Acta*, 362: 308-315.

♦Tian, F., Bolon, P. J. and Prestegard, J. H. (1999). Intensity based measurement of homonuclear residual dipolar couplings from CT-COSY. *J. Am. Chem. Soc.*, in press.

♦Tian, F., Fowler, C. A., Zartler, E. R., Jenney, F. A., Adams, M. W. and Prestegard, J. H. (2000). Direct measurement of ^1H - ^1H dipolar couplings in proteins: A complement to traditional NOE measurements. *J. Biomol. NMR.*, 18: 23-31.

♦Tjandra, N. and Bax, A. (1997a). Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, 278: 1111-1114.

♦Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. and Bax, A. (1997b). 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: 732-8.

- ♦Tjandra, N., Garrett, D. S., Gronenborn, A. M., Bax, A. and Clore, G. M. **(1997c)**. Defining long range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy. *Nat. Struct. Biol.*, 4: 443-449.
- ♦Tolman, J. R., Flanagan, J. M., Kennedy, M. A. and Prestegard, J. H. **(1995)**. Nuclear magnetic dipole interactions in field-oriented proteins - information for structure determination in solution. *Proc. Natl. Acad. Sci. USA*, 92: 9279-9283.
- ♦Tolman, J. R. and Prestegard, J. H. **(1996)**. A quantitative *J*-correlation experiment for the accurate measurement of one-bond amide ^{15}N - ^1H couplings in proteins. *J. Magn. Reson.*, 112: 245-252.
- ♦Torchia, D. A., Sparks, S. W. and Bax, A. **(1988)**. NMR signal assignments of amide protons in the α -helical domains of staphylococcal nuclease. *Biochemistry*, 27: 5135-5141.
- ♦Torizawa, T., Shimizu, M., Taoka, M., Miyano, H. and Kainosho, M. **(2004)**. Efficient production of isotopically labeled proteins by cell-free synthesis: A practical protocol. *J. Biomol. NMR.*, 30: 311-325.
- ♦Tropp, J. **(1980)**. Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: The effect of fluctuating internuclear distances. *J. Chem Phys.*, 72: 6035-6043.
- ♦Tsang, P., Wright, P. E. and Rance, M. **(1990)**. Specific deuteration strategy for enhancing direct nuclear overhauser effects in high-molecular weight complexes. *J. Am. Chem. Soc.*, 112: 8183-8185.
- ♦Tugarinov, V., Muhandiram, D. R., Ayed, A. and Kay, L. E. **(2002)**. Four-dimensional NMR spectroscopy of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase G. *J. Am. Chem. Soc.*, 124: 10025-10035.
- ♦Tugarinov, V. and Kay, L. E. **(2003a)**. Side chain assignments of Ile δ_1 methyl groups in high molecular weight proteins: An application to a 46 ns tumbling molecule. *J. Am. Chem. Soc.*, 125: 5701-5706.
- ♦Tugarinov, V. and Kay, L. E. **(2003b)**. Ile, Leu, and Val methyl assignments of the 723-residue Malate Synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc.*, 125: 13868-13878.
- ♦Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. and Kay, L. E. **(2003c)**. Cross-correlated relaxation enhanced ^1H - ^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.*, 125: 10420-10428.
- ♦Tugarinov, V. and Kay, L. E. **(2004a)**. An isotope labeling strategy for methyl TROSY spectroscopy. *J. Biomol. NMR.*, 28: 165-172.
- ♦Tugarinov, V., Scheurer, C., Br schweiler, R. and Kay, L. E. **(2004b)**. Estimates of methyl ^{13}C and ^1H CSA values in proteins from cross-correlated spin relaxation. *J. Biomol. NMR.*, 30: 397.
- ♦Tugarinov, V., Choy, W. Y., Orekhov, V. Y. and Kay, L. E. **(2005a)**. Solution NMR-derived global folded of a monomeric 82-kDA enzyme. *Proc. Natl. Acad. Sci. USA*, 102: 622-627.

- ♦Tugarinov, V., Kay, L. E., Ibraghimov, I. and Orekhov, V. Y. **(2005b)**. High-resolution four-dimensional ^1H - ^{13}C NOE spectroscopy using methyl-TROSY, Sparse data acquisition, and multidimensional decomposition. *J. Am. Chem. Soc.*, 127: 2767-2775.
- ♦Tugarinov, V. and Kay, L. E. **(2005c)**. Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. *Chembiochem*, 6: 1567 – 1577.
- ♦Tugarinov, V., Kanelis, V. and Kay, L. E. **(2006)**. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Prot.*, 1: 749-754.
- ♦Tugarinov, V. and Kay, L. E. **(2007)**. Separating degenerate ^1H transitions in methyl group probes for single-quantum ^1H -CPMG relaxation dispersion NMR spectroscopy. *J. Am. Chem. Soc.*, 129: 9514-9521.

V

- ♦Vega, A. J. and Fiat, D. **(1976)**. Nuclear relaxation processes of paramagnetic complexes The slow-motion case. *Mol. Phys.*, 31: 347-355.
- ♦Venters, R. A., Calderonne, T. L., Spicer, L. D. and Fierke, C. A. **(1991)**. Uniform ^{13}C isotope labeling of proteins with sodium acetate for NMR studies: Application to human carbonic anhydrase II. *Biochemistry*, 30: 4491-4494.
- ♦Venters, R. A., Huang, C.-C., Farmer II, B. T., Trolard, R., Spicer, L. D. and Fierke, C. A. **(1995)**. High-level $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeling of proteins for NMR studies. *J. Biomol. NMR.*, 5: 339-344.
- ♦Vinarov, D. A., Lytle, B. L., Peterson, F. C., Tyler, E. M., Volkman, B. F. and Markley, J. L. **(2004)**. Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nature Met.*, 1: 149-153.
- ♦Vinarov, D. A. and Markley, J. L. **(2005)**. High-throughput automated platform for nuclear magnetic resonance based structural proteomics. *Expert Rev. Proteomics*, 2: 49-55.
- ♦Vinarov, D. A., Loushin Newman, C. L. and Markley, J. L. **(2006)**. Wheat germ cell-free platform for eukaryotic protein production. *F.E.B.S.*, 273: 4160-4169.

W

- ♦Wang, A. C. and Bax, A. **(1996)**. Determination of the backbone dihedral angles phi in human ubiquitin from reparametrized empirical Karplus equations. *J. Am. Chem. Soc.*, 118: 2483-2494.
- ♦Wang, Y. X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., Grzesiek, S. and Bax, A. **(1999)**. Measurement of $^3\text{h}J_{\text{NC}}$ connectivities across hydrogen bonds in a 30 kDa protein. *J. Biomol. NMR.*, 14: 181-184.
- ♦Wangsness, R. K. and Bloch, F. **(1953)**. The dynamical theory of nuclear relaxation. *Phys. Rev.*, 89: 728-739.

- ♦Weber, P. L., Brown, S. C. and Mueller, L. **(1987)**. Sequential ^1H NMR assignments and secondary structure identification of human ubiquitin. *Biochemistry*, 26: 7282-7290.
- ♦Werbelow, L. G. and Marshall, A. G. **(1973)**. Internal rotation and nonexponential methyl nuclear relaxation for macromolecules. *J. Magn. Reson.*, 11: 299-313.
- ♦Werbelow, L. G. and Grant, D. M. **(1977)**. Intramolecular dipolar relaxation in multispin systems. *Ad. Magn. Reson.*, 9: 189-299.
- ♦Williamson, M. P., Havel, T. F. and Wüthrich, K. **(1985)**. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^1H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.*, 182: 295-315.
- ♦Wimperis, S. and Bodenhausen, G. **(1986)**. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, 278: 1111-1114.
- ♦Wu, Z., Tjandra, N. and Bax, A. **(2001a)**. Measurement of $^1\text{H}_3$ - ^{31}P dipolar couplings in a DNA oligonucleotide by constant-time NOESY difference spectroscopy. *J. Biomol. NMR.*, 19: 367-70.
- ♦Wu, Z., Ono, A., Kainosho, M. and Bax, A. **(2001b)**. $\text{H}^{\cdots}\text{N}$ hydrogen bond lengths in double stranded DNA from internucleotide dipolar couplings. *J. Biomol. NMR.*, 19: 361-5.
- ♦Wu, Z. and Bax, A. **(2001c)**. Measurement of homonuclear proton couplings based on cross-peak nulling in CT-COSY. *J. Magn. Reson.*, 151: 242-52.
- ♦Wu, Z. and Bax, A. **(2002)**. Measurement of Long-Range ^1H - ^1H Dipolar Couplings in Weakly Aligned Proteins. *J. Am. Chem. Soc.*, 124: 9672-9673.
- ♦Wu, Z., Delaglio, F., Wyatt, K., Wistow, G. and Bax, A. **(2005)**. Solution structure of γS -crystallin by molecular fragment replacement NMR. *Prot. Science*, 14: 3101-3114.
- ♦Wüthrich, K. **(1986)**. NMR of Proteins and Nucleic Acids. *A Wiley-Interscience Publication, New York*.

γ

- ♦Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E. D., Cooper, J. A., Kainosho, M. and Yokoyama, S. **(1998)**. Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. *J. Biomol. NMR.*, 11: 295-306.