

# Le marquage isotopique spécifique

## Un outil pour repousser les frontières de la RMN biomoléculaire

Pierre Gans et Jérôme Boisbouvier

### Résumé

La résonance magnétique nucléaire est une méthode de choix pour étudier, à la résolution atomique, les caractéristiques structurales et dynamiques des biomacromolécules en solution. Cependant les techniques RMN standard se caractérisent par une faible sensibilité et sont limitées à des objets biologiques de taille modeste. Le développement de nouvelles techniques de marquage isotopique spécifique a permis ces dernières années de repousser les limites de la RMN biomoléculaire. Des protocoles permettant d'incorporer des groupes méthyles  $^{13}\text{CH}_3$  de façon sélective ont été mis au point pour les acides aminés isoleucine, alanine, méthionine, leucine et valine au sein des protéines perdeutérées. Ce marquage augmente considérablement la qualité des données RMN. Ainsi, le gain en sensibilité apporté constitue une base simple et efficace pour la caractérisation structurale et dynamique des interactions moléculaires au sein des édifices supramoléculaires complexes pouvant atteindre 1 MDa. Par ailleurs, ce gain de sensibilité permet la détection d'interactions très faibles dans les protéines telles que des nOe entre protons séparés par plus de 10 Å ou des couplages scalaires de quelques dixièmes, voire centièmes de hertz.

### Mots-clés

**Résonance magnétique nucléaire, RMN, structure de biomolécules, marquage isotopique.**

### Abstract

**Specific isotopic labelling: a tool to extend the limits of biomolecular NMR**

Nuclear magnetic resonance spectroscopy is a method of choice to study the structural and dynamic properties of biomacromolecules in solution at atomic resolution. However, standard NMR techniques are characterized by low sensitivity and are limited to biological objects of modest size. The development of new specific isotope labeling techniques has allowed the limits of biomolecular NMR to be extended in recent years. Protocols have been developed to selectively incorporate  $^{13}\text{CH}_3$  labeled methyl groups in isoleucine, alanine, methionine, leucine and valine amino acids within otherwise perdeuterated proteins. These labeling schemes significantly improve the quality of NMR data and this gain in sensitivity and resolution allows the characterization of the structure and dynamics of molecular interactions within supramolecular complexes of up to 1 MDa. Moreover, the improvement in sensitivity allows detection of very weak interactions in proteins such as nOe between protons separated by more than 10 Å or scalar couplings of a few tenths or hundredths of a hertz.

### Keywords

**Nuclear magnetic resonance, NMR, biomolecular structure, isotopic labelling.**

## Le marquage isotopique des protéines

Depuis les années 1980, date de la résolution des premières structures tridimensionnelles de protéines par résonance magnétique nucléaire, la structure tridimensionnelle des protéines a largement bénéficié de l'amélioration technique des spectromètres : augmentation du champ magnétique statique et introduction de sondes de détection cryogéniques. Toutefois, l'avancée qui a principalement permis le passage d'une structure d'une taille de quelques kDa à celle de plus de 50 kDa a été l'introduction successive de différentes techniques de marquage isotopique dans les protéines. En effet, la base de la détermination structurale est la collecte d'informations de distances proton-proton (ou nOe, « nuclear overhauser effect ») servant de contraintes expérimentales pour des algorithmes de modélisation moléculaire. Toutefois, le prérequis de cette méthode est la détermination de la fréquence de résonance pour la quasi-totalité des protons de la protéine. Dans le cas de protéines de petite taille, cela est possible pour des échantillons obtenus sans marquage particulier. Cela a permis d'obtenir les structures de protéines jusqu'à une dizaine de kDa dans des cas favorables [1-2].

Cette stratégie ne suffit toutefois plus pour des objets de plus grande taille à cause de la superposition des résonances et il faut augmenter la résolution en dispersant les signaux

selon une nouvelle fréquence, ce qui est fait grâce à l'introduction d'hétéroatomes comme l'azote-15 et le carbone-13. Cette technique a permis dans un premier temps un bond dans la taille des protéines accessibles à la RMN en permettant de résoudre la structure de protéines jusqu'à 20 kDa [3]. L'augmentation de la taille des protéines étudiées induit par contre un phénomène très défavorable pour la RMN. Du fait de l'accroissement du temps de corrélation de la protéine, les échanges d'aimantation par effet dipolaire d'un proton donné vers ses voisins deviennent très efficaces. L'augmentation de la taille des biomacromolécules s'accompagne ainsi d'un élargissement des raies des signaux RMN et d'un phénomène de diffusion de spins qui limite l'extraction de contraintes de distances précises. Le remplacement partiel uniforme des protons de la protéine par des deutérons avec lesquels ces échanges d'aimantation ne peuvent pas se produire permet de pallier ces phénomènes de fuite d'aimantation et de recouvrer la sensibilité des expériences. Il a été ainsi possible d'obtenir la structure tridimensionnelle de complexes protéiques de plus de 40 kDa comme celui entre l'enzyme I et HPr [4].

Toutefois, si l'on veut étudier des objets de plus grande taille (de l'ordre d'une centaine de kDa ou plus) comme des complexes du protéasome, cette méthode n'est plus assez efficace. En effet, le surcroît d'efficacité des interactions

dipolaires entre protons – qui conduit à une relaxation encore plus rapide du signal – ne peut plus être compensé par une nouvelle dilution de ces protons par des deutérons, car on atteindrait la limite où il n'y aura plus assez de protons par site pour donner un signal observable. Cet état de fait a conduit à proposer ces dernières années une nouvelle stratégie de marquage pour ce type d'objet qui peut se résumer ainsi : pour une densité donnée de protons dans une protéine, il vaut mieux que quelques sites soient entièrement protonés dans un contexte entièrement deutéré (perdeutération) plutôt que tous les sites soient faiblement et uniformément protonés. L'inconvénient de ce marquage est que l'on ne peut plus observer qu'un nombre restreint de sites, de plus difficilement identifiables. Par ailleurs, on ne peut alors mesurer que les distances entre des sites occupés, ce qui réduit grandement le nombre de contraintes utiles pour la détermination de structures pouvant être obtenues pour un type d'échantillon. Par contre, les avantages de cette méthode sont que les intensités des signaux observés sont maximales puisque les sites sont occupés à 100 % par les protons, et que les spectres sont grandement simplifiés car un seul type de site est observable et la relaxation des signaux est lente puisque les fuites par échange d'aimantation sont faibles, les sites spécifiquement protonés étant bien isolés les uns des autres. De fait, l'introduction de ce type de marquage a permis d'aller vers l'obtention de structures de protéines de plus grande taille comme celle de la malate synthase G (MSG) de 82 kDa [5] ou même de complexes de 200 kDa [6].

## Intérêt de la protonation spécifique des groupements méthyles dans des protéines perdeutérées

Cette stratégie de marquage spécifique pourrait s'appliquer à tous les sites où se trouvent des protons non échangeables avec le solvant, dès lors qu'il est possible d'obtenir des acides aminés ou des précurseurs protonés spécifiquement en une position de la molécule. Elle a été particulièrement développée ces dernières années pour les groupements méthyles des acides aminés des protéines car ils bénéficient en effet de trois propriétés très favorables pour la RMN :

- L'aimantation transférée entre deux groupes méthyles est multipliée par neuf, du fait de la multiplicité des méthyles, par rapport à l'intensité du signal qui serait transféré entre deux protons séparés par la même distance. D'une part, l'aimantation du méthyle est trois fois ce qu'elle serait pour un seul proton ; d'autre part, l'aimantation d'un premier méthyle peut être transférée vers chacun des trois protons du second méthyle.
- Du fait de la rotation rapide du groupement méthyle suivant la liaison carbone-carbone, les vitesses de relaxation transversale et longitudinale sont beaucoup plus favorables que celles d'un proton du squelette ou d'une chaîne latérale. Cela conduit à un rapport signal sur bruit favorable permettant l'acquisition de spectres RMN 2D en quelques secondes [7].
- Enfin, dans une protéine globulaire, alors que les groupes polaires restent plutôt en surface, les acides aminés hydrophobes (dont les leucines, valines et isoleucines) se rassemblent pour former un ou plusieurs cœurs hydrophobes au centre de la structure tertiaire. Les groupes méthyles seront donc concentrés dans des zones intérieures et la mesure de leurs proximités permettra l'obtention de nombreuses contraintes de distance indispensables pour le calcul du repliement global de la protéine.

## Obtention des méthyles spécifiquement protonés pour un type d'acide aminé dans une protéine perdeutérée

Il est donc nécessaire que la protéine soit entièrement deutérée à l'exception d'un type de méthyle qui restera protoné. Par voie de synthèse chimique, les six acides aminés (Ala, Val, Leu, Ile, Thr et Met) peuvent être obtenus marqués  $^{13}\text{CH}_3$  spécifiquement sur les méthyles, et dans le cas des leucines et valines, sur un seul des méthyles de façon stéréospécifique, mais le coût de tels acides aminés limite l'application en RMN biomoléculaire. Il serait possible d'obtenir une protéine spécifiquement marquée sur un méthyle d'un acide aminé particulier en la surproduisant en milieu perdeutéré dans une souche d'*E. coli* auxotrophe pour cet acide aminé particulier et en fournissant l'acide aminé spécifiquement marqué sur le groupe méthyle [8]. Toutefois, la croissance de ces souches auxotrophes<sup>(1)</sup> est généralement trop faible pour surexprimer en milieu deutéré la protéine désirée en quantité suffisante. De plus, il est nécessaire de fournir tous les autres acides aminés sous forme deutérée, ce qui augmente considérablement les coûts.

Une autre voie pourrait être la synthèse des protéines *in vitro* [9] (ou méthode « cell-free ») en apportant les acides aminés correctement enrichis. Grâce à cette méthode, il a été possible de synthétiser des protéines où les carbones des chaînes latérales ne portent qu'un seul proton – méthode SAIL [10] –, ce qui permet de grandement simplifier et d'améliorer les spectres. En apportant les acides aminés deutérés et spécifiquement marqués  $^{13}\text{CH}_3$  sur un méthyle donné, on pourrait donc obtenir le type d'échantillon souhaité. Ce type d'approche reste toutefois extrêmement onéreux et difficile à mettre en œuvre.

De ce fait, les différents groupes (dont le nôtre) qui ont développé de tels marquages spécifiques se sont plutôt tournés vers une approche où la protéine est produite *in vivo* par *E. coli* dans un milieu de culture minimum perdeutéré et où sont fournis les acides aminés ou leurs précurseurs spécifiquement protonés. Cette méthode est d'un coût beaucoup plus raisonnable, mais nécessite la maîtrise des voies métaboliques par lesquelles *E. coli* assimile l'acide aminé ou le synthétise à partir d'un précurseur. En effet, il faut autant que possible fournir un intermédiaire proche dans la voie de biosynthèse de l'acide aminé pour pouvoir diluer au maximum, sinon rétroinhiber, le flux provenant du métabolisme, et d'autre part que les réactions enzymatiques dans cette voie soient irréversibles, de façon à ce que le marquage ne se retrouve pas dans les autres acides aminés. Il est possible de pallier ce type de problème comme nous le verrons dans le cas du marquage de l'alanine, mais cela peut présenter un surcoût important et se montrer d'une optimisation délicate [11-12]. Certains de ces précurseurs sont disponibles commercialement, mais en règle générale, préalablement à la culture dans *E. coli*, il faut synthétiser le composé avec le marquage isotopique désiré. Nous allons maintenant détailler quelques exemples de modes de marquage développés récemment.

### Marquage spécifique des isoleucines en position $\delta 1$

Dans le cas de l'isoleucine, il est simple d'obtenir une protonation spécifique sur le méthyle terminal ( $\delta 1$ ) de la chaîne latérale [13]. L'isoleucine est produite par une voie de biosynthèse qui met en jeu le 2-oxobutanoate produit par déamination de la thréonine (voir figure 1). Celui-ci, après une

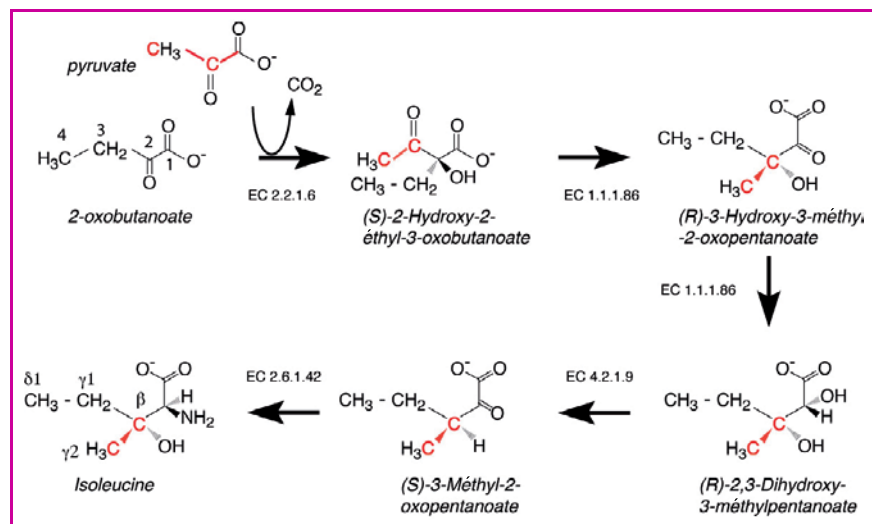


Figure 1 - Voie de biosynthèse de l'isoleucine.

Les positions et la stéréochimie dans chaque intermédiaire des carbones issus du pyruvate sont indiquées en rouge. Chaque intermédiaire dans la voie de biosynthèse a été nommé suivant KEGG (*Kyoto Encyclopedia of Genes and Genomes*). Les enzymes catalysant les différentes réactions sont désignées par leur numéro EC (*Enzyme Commission number*). EC 2.2.1.6 : acétolactate synthase ; EC 1.1.1.86 : acétolactate isoméroréductase ; EC 4.2.1.9 : dihydroxyacide déshydratase ; EC 2.6.1.42 : glutamate amino transférase, spécifique des acides aminés à chaînes latérales (branched-chain-amino-acid transaminase). Des informations complémentaires sur la voie de biosynthèse de l'isoleucine peuvent être trouvées sur [www.genome.jp/kegg](http://www.genome.jp/kegg).

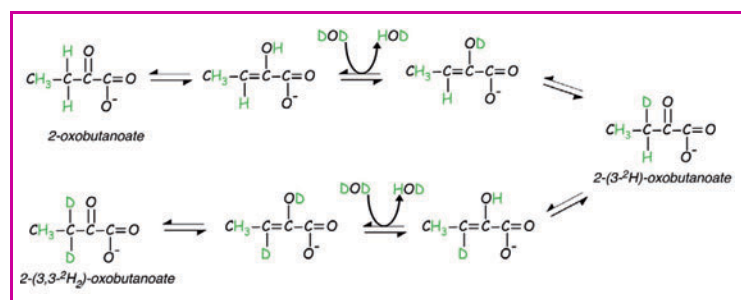


Figure 2 - Réaction d'échange proton-deuteron en position 3 du 2-oxobutanoate.

La réaction se fait dans D<sub>2</sub>O, pD : 10,5 pendant 72 heures à 40 °C [13].

condensation irréversible avec une molécule de pyruvate, est transformé en (S)-2-acéto-2-hydroxybutanoate, puis en (S)-3-méthyl-2-oxopentanoate, et enfin en isoleucine par transamination, les carbones 3 et 4 du 2-oxobutanoate se retrouvant respectivement en position γ1 et δ1 de l'isoleucine.

À partir du 4-<sup>13</sup>C 2-oxobutanoate disponible commercialement, on obtient aisément le 4-<sup>13</sup>C-3-<sup>2</sup>H<sub>2</sub>-2-oxobutanoate par échange dans l'eau lourde à pH basique (figure 2). Ce composé est ajouté dans le milieu de culture deutéré, où il constitue l'unique source de protons, une heure avant induction de l'expression de la protéine par *E. coli* ; cela permet d'obtenir le marquage souhaité de l'isoleucine dans la protéine d'intérêt avec des taux de deutération de la protéine et d'incorporation du groupement <sup>13</sup>C<sup>1</sup>H<sub>3</sub> sur les sites δ1 des isoleucines supérieurs à 98 % [13].

### Marquage stéréospécifique de la leucine et de la valine

La protonation des deux méthyles de la valine et de la leucine peut être obtenue à partir du 2-oxo-3-méthylbutanoate (ou α-cétoisovalérate) protoné spécifiquement sur les deux méthyles de la même manière que pour l'isoleucine. Comme pour le 2-oxobutanoate, ce composé spécifiquement protoné

s'obtient aisément directement dans l'eau lourde par échange proton-deuteron en milieu basique [14]. Mais ce marquage ne présente que peu d'intérêt pratique pour l'application de la spectroscopie RMN à l'étude des protéines de grande taille. En effet, la proximité des deux méthyles dans le même résidu conduit à des transferts d'aimantation entre ces deux groupes, ce qui induit une diminution beaucoup plus rapide du signal et interdit pratiquement les transferts d'aimantation vers des méthyles éloignés. Pour résoudre ce problème, un 2-oxo-3-méthylbutanoate, le 2-oxo-3-(<sup>2</sup>H<sub>3</sub>)méthyl-3-(<sup>2</sup>H)-4-(<sup>13</sup>C)-butanoate où statistiquement un méthyle sur deux est deutéré, a été développé [15] et est disponible commercialement.

Toutefois, l'utilisation de ce composé, s'il permet effectivement d'obtenir des signaux exploitables dans des protéines de grande taille, présente un certain nombre d'inconvénients pour les applications RMN. Certes, il est possible d'accéder à des contraintes de distances, mais celles-ci étant obtenues entre deux sites occupés par moitié, le signal est le

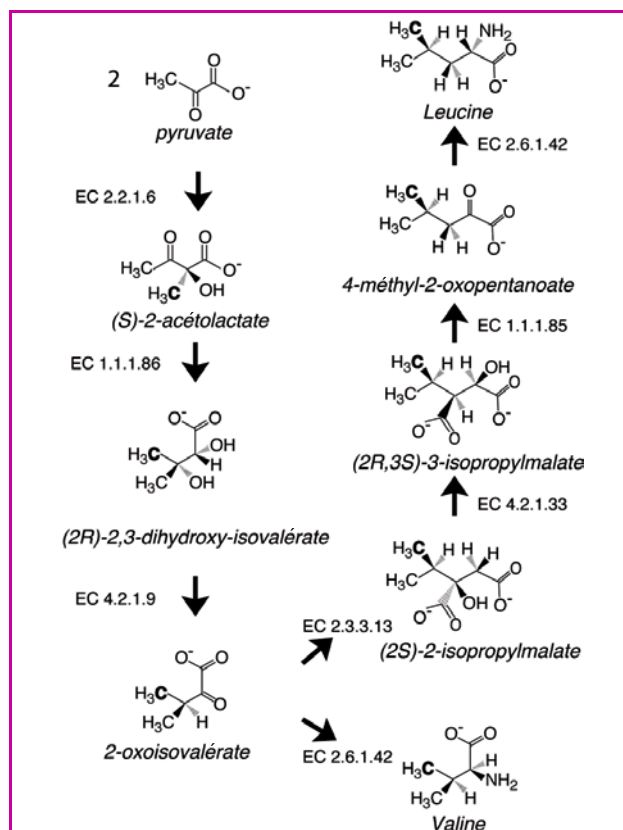


Figure 3 - Voie de biosynthèse de la valine et de la leucine.

La position et la stéréochimie dans chaque intermédiaire du carbone issu du méthyle situé en 2 dans l'acétolactate (ou 2-hydroxy-2-méthyl-3-oxobutanoate) sont indiquées en gras. Chaque intermédiaire dans la voie de biosynthèse a été nommé suivant KEGG (*Kyoto Encyclopedia of Genes and Genomes*). Les enzymes catalysant les différentes réactions sont désignées par leur numéro EC (*Enzyme Commission number*). EC 2.2.1.6 : acétolactate synthase ; EC 1.1.1.86 : céto-acide réductoisomérase ; EC 4.2.1.9 : dihydroxy-acide déshydratase ; EC 2.6.1.42 : aminotransférase d'acide aminé à chaîne ramifiée ; EC 2.3.3.13 : 2-isopropylmalate synthase ; EC 4.2.1.33 : 3-isopropylmalate déshydratase ; EC 1.1.1.85 : 3-isopropylmalate déshydrogénase. Des informations complémentaires sur la voie de biosynthèse des leucines et valines peuvent être trouvées sur [www.genome.jp/kegg](http://www.genome.jp/kegg).

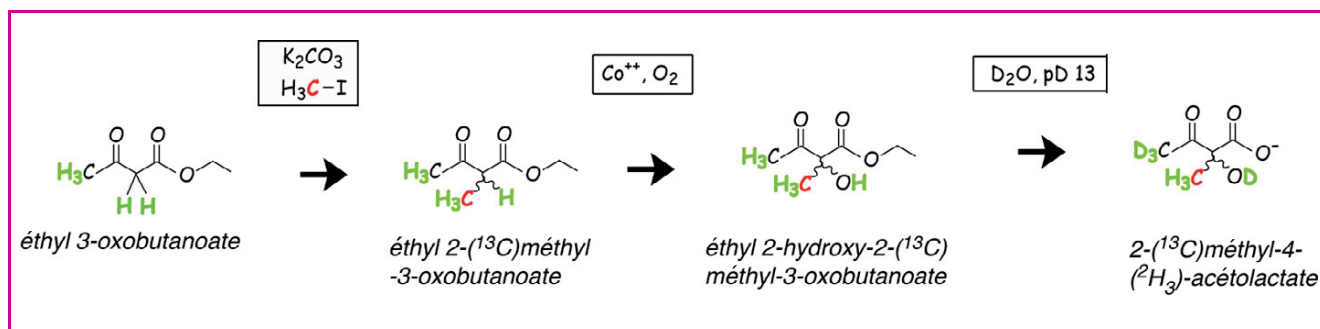


Figure 4 - Protocole de synthèse chimique du 2-( $^{13}\text{C}$ )méthyl-4-( $^2\text{H}_3$ )-acétolactate.

Le noyau  $^{13}\text{C}$  est indiqué en rouge. La réaction de méthylation se fait à 40 °C dans l'éthanol sous argon pendant 90 h. L'oxydation est effectuée dans un mélange acétonitrile-isopropanol pendant 28 h à 60 °C. La deutération du 4-méthyle se fait en moins de 30 min dans NaOD pD : 13 en même temps que la déprotection du groupe carboxylate.

quart de ce qu'il serait si les deux sites étaient entièrement occupés. D'autre part, des signaux sont obtenus pour chacun des deux méthyles, ce qui conduit à de nombreuses superpositions dans les spectres pour les protéines de grande taille. Nous avons donc cherché à développer un mode simple de marquage stéréospécifique des leucines et des valines. L'examen des voies de biosynthèse montre que la stéréochimie du radical isopropyl est établie dès l'étape de conversion de l'acétolactate en (R)-3,4-dihydroxy-3-méthylbutanoate (figure 3, p. 7).

Cette conversion se fait de façon stéréosélective en transférant le groupe méthyle en 2 en position proR du groupement isopropyle du 2-cétoisovalérate (voir figure 3) [16-17]. Il est donc possible d'obtenir la stéréosélectivité désirée de

la leucine et de la valine à partir d'un acétolactate différemment marqué sur les deux méthyles. Par contre, l'acétolactate étant synthétisé par une condensation de deux pyruvates dans l'étape de l'acétolactate synthase, les deux méthyles de ce composé ne peuvent pas être marqués *in vivo* de façon différente. Nous nous sommes donc tournés vers la synthèse chimique de cet intermédiaire et l'acétolactate a été synthétisé suivant les méthodes rapportées dans les références [18-20] par une suite de réactions décrite dans la figure 4.

La première étape part du 3-oxobutanoate (ou acétoacétate) dont la fonction carboxylique est protégée par un méthyle ou un éthyle. Celui-ci est traité en milieu basique par de l'iodure de méthyle marqué  $^{13}\text{C}$  ( $^{13}\text{CH}_3\text{-I}$ ), ce qui permet d'obtenir le 2-( $^{13}\text{C}$ )méthyl 3-oxobutanoate avec un bon rendement (90 %). Ce composé est ensuite oxydé à l'air en présence de catalyseur métallique pour donner le 2-hydroxy-2-( $^{13}\text{C}$ )méthyl-3-oxobutanoate. La fonction carboxylique est déprotégée en même temps qu'est deutérée la position 4 de l'acétolactate par un échange proton/deuteron dans l'eau lourde en milieu basique. L'ajout du composé obtenu dans un milieu de culture deutéré de *E. coli* permet d'obtenir des protéines uniquement protonées sur la position proS des valines et des leucines comme le montre la comparaison des spectres de la malate synthase G [21] (figure 5). Il est à noter que la synthèse de l'acétolactate marqué n'est pas stéréosélective et que seul le stéréoisomère 2-(S) est incorporé *in vivo*, le stéréoisomère 2-(R) n'étant pas métabolisé. Cela permet une simplification importante du spectre, une seule moitié des méthyles des leucines et valines étant encore observable. D'autre part, les sites étant maintenant pleinement protonés, les intensités des signaux permettant de quantifier les distances méthyles-méthyles sont multipliées par un facteur 4, ce qui permet de mesurer des contraintes jusqu'à des distances proches de 12 Å dans des protéines de taille allant jusqu'à 20 kDa [22] comme l'illustre la figure 6.

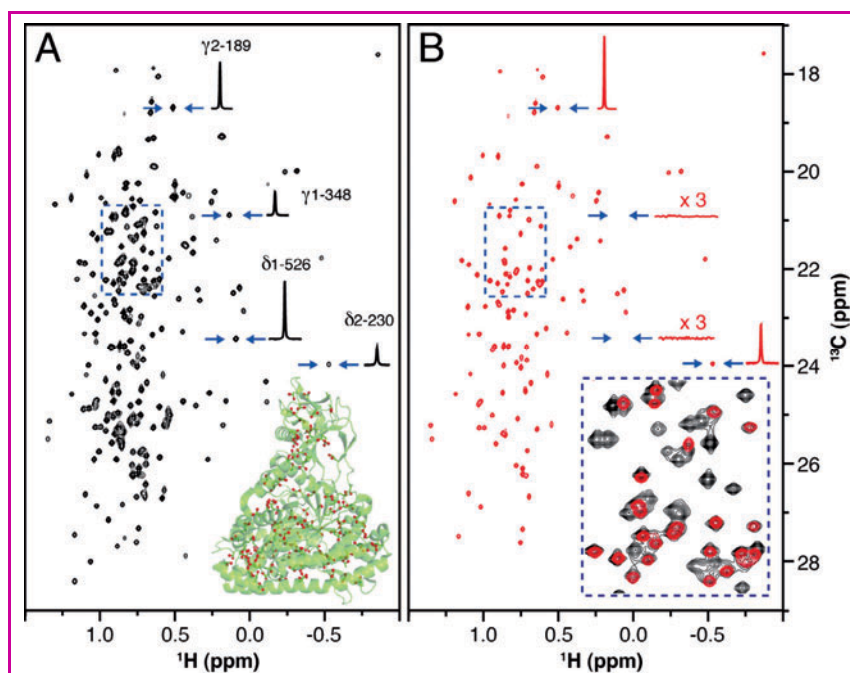


Figure 5 - Comparaison des spectres méthyl-TROSY (« transverse relaxation optimized spectroscopy ») enregistrés sur des échantillons de malate synthase G (82 kDa) spécifiquement marquée sur les leucines et les valines, et préparée **A** : avec du 2-oxo-3-( $^2\text{H}_3$ )méthyl-3-( $^2\text{H}$ )-4-( $^{13}\text{C}$ )-isovalérate donnant de l'U-[ $^2\text{H}$ ], U-[ $^{13}\text{C}$ ], Leu/Val-[ $^{12}\text{C}_2\text{H}_3$ ,  $^{13}\text{C}_1\text{H}_3$ ] MSG ; **B** : avec du 2-( $^{13}\text{C}$ )méthyl-4-( $^2\text{H}_3$ )-acétolactate donnant de l'U-[ $^2\text{H}$ ], U-[ $^{13}\text{C}$ ], Leu/Val-[ $^{13}\text{C}_1\text{H}_3$ ]proS MSG.

Les corrélations des méthyles  $\delta 1$  des leucines et  $\gamma 1$  des valines ne sont plus détectables dans le spectre B alors que l'intensité de celles des méthyles  $\delta 2$  des leucines et  $\gamma 2$  des valines est multipliée par un facteur proche de 2. L'insert de la figure A montre la position dans la structure 3D de MSG des groupes méthyles des leucines et valines (sphères rouges).

Reproduit avec permission de [21], © Wiley-VCH Verlag GmbH & Co. KGaA.

### Marquage spécifique de l'alanine

L'alanine est statistiquement l'acide aminé le plus abondant dans les protéines avec la leucine, mais contrairement à cette dernière où le méthyle

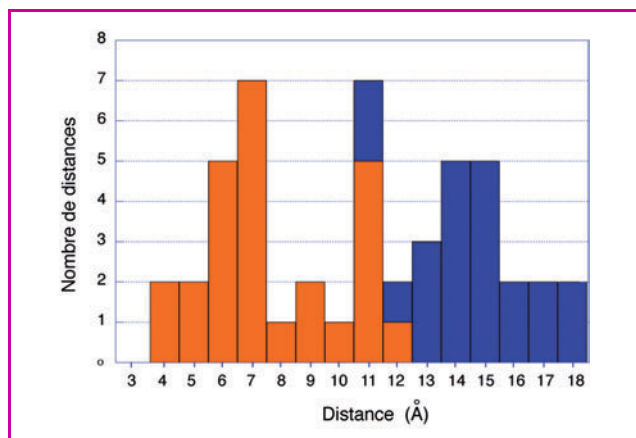


Figure 6 - Distribution des distances interméthyles en fonction des distances proton-proton calculées à partir des structures de l'ubiquitine (1D3Z) et de la  $\gamma$ S-crystalline (2A5M). Les NOEs observés expérimentalement sont indiqués par des barres orange.

est en bout de chaîne latérale, le marquage de l'alanine permet d'introduire une sonde proche du squelette protéique. Par contre, à la différence de l'isoleucine, la valine et la leucine où l'oxoacide donnant l'acide aminé par transamination se trouve en fin de voie métabolique, l'oxoacide se trouve être ici le pyruvate qui occupe une position centrale dans le métabolisme carboné (voir figure 7). De ce fait, le protocole de marquage de l'alanine s'est révélé complexe à mettre au point.

Il n'était donc pas possible d'utiliser du pyruvate protoné comme précurseur sans introduire de multiples sites de protonation dans la protéine. Il a fallu dans ce cas fournir lors de la culture de l'alanine convenablement marquée – c'est-à-dire où le groupement méthyle était à la fois protoné et marqué  $^{13}\text{C}$  alors que le reste de la molécule était deutéré et  $^{12}\text{C}$ . Partant de la 3- $^{13}\text{C}$  alanine commerciale, en utilisant la tryptophane synthase – une protéine à pyridoxal phosphate qui a la capacité d'échanger les protons alpha sans changer la configuration absolue du carbone alpha de l'alanine [23] –, nous avons obtenu en milieu deutéré le composé de base, la 3- $^{13}\text{C}$ , 2- $^2\text{H}$  alanine qui, une fois introduit dans le milieu de culture permet d'obtenir le marquage désiré dans la protéine (figure 8).

Par contre, la réaction d'amination du pyruvate étant réversible, comme un certain nombre de réactions enzymatiques, l'ajout de 3- $^{13}\text{C}$ , 2- $^2\text{H}$  alanine conduit à la formation de 3- $^{13}\text{C}$  pyruvate protoné par déamination de l'alanine, et donc à l'apparition de protonations indésirables en de nombreux sites. L'ajout de pyruvate entièrement deutéré pour diluer le pyruvate produit par déamination permet de pallier ce phénomène mais au prix d'une diminution importante du marquage final de l'alanine. Nous avons donc apporté pour chaque voie de biosynthèse une quantité d'intermédiaires deutérés ( $\alpha$ -cétoglutarate,  $\alpha$ -cétosovalérate, isoleucine, oxaloacétate et glycérone) suffisante pour fortement diluer le flux d'intermédiaires ou rétroinhiber la voie de biosynthèse. Ce protocole permet d'obtenir des taux de marquage de la position C $\beta$  de l'alanine supérieurs à 95 % sans que les autres sites soient marqués à plus de 3 % [12], et d'observer ainsi tous les signaux des alanines dans des protéines de plus 80 kDa comme l'illustre la figure 9.

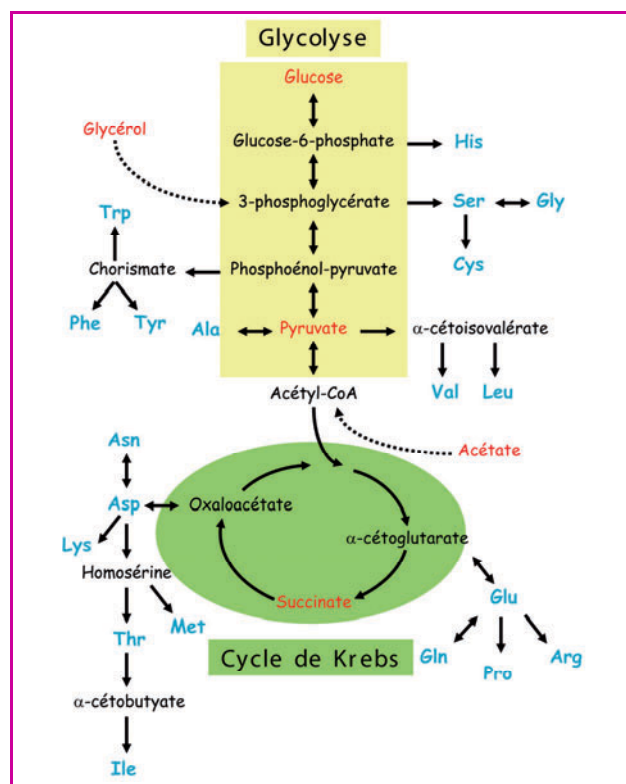


Figure 7 - Schéma simplifié des voies métaboliques conduisant aux vingt différents acides aminés.

En rouge : les composés carbonés utilisés comme source de carbone dans les cultures de bactéries ; en bleu : les acides aminés.

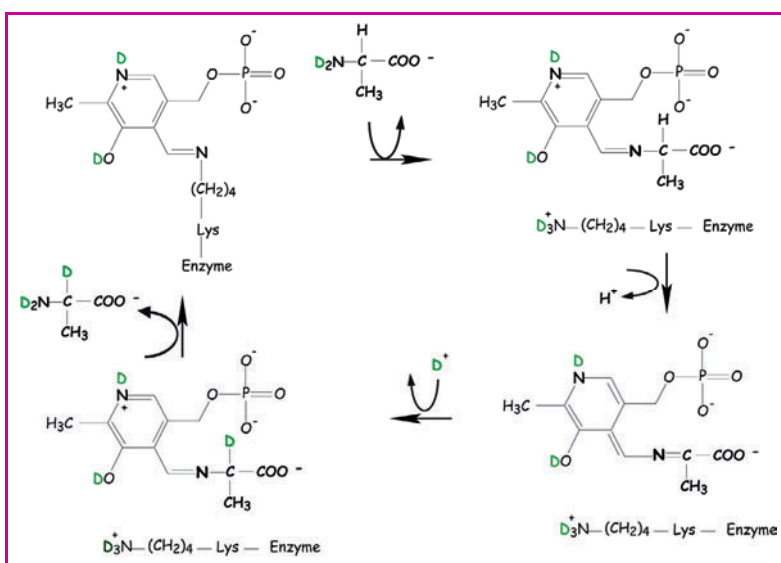


Figure 8 - Mécanisme de deutération de la 3- $^{13}\text{C}$ -alanine par la tryptophane synthase dans l'eau lourde.

Les atomes de l'alanine sont représentés en gras, ceux de deutérium en vert.

## Perspectives

Le nombre de types de méthyles différents dans une protéine est très limité et il est clair que les protocoles de marquage que nous venons de présenter constituent l'essentiel des protocoles utilisables *in vivo* pour des applications de résonance magnétique nucléaire des protéines. Pour la méthionine, un protocole de marquage a été également

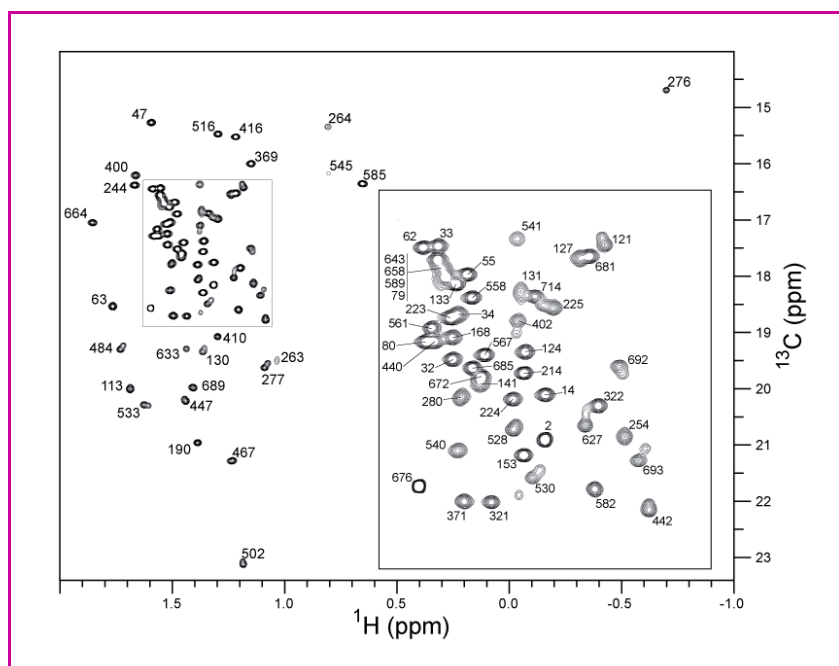


Figure 9 - Spectre 2D  $^1\text{H}$ - $^{13}\text{C}$  HMQC (« heteronuclear multiple quantum coherence ») de la malate synthase G (MSG, 82 kDa) perdeutérée et protonée spécifiquement sur les méthyles des alanines (U- $^2\text{H}$ ), U- $^{15}\text{N}$ ], U- $^{12}\text{C}$ ], U- $^{13}\text{C}_1\text{H}_3$ -Ala- $\beta$  MSG).

L'encart correspond à la partie centrale du spectre. Les méthyles des alanines ont été identifiés à l'aide d'une expérience 3D HCC corrélant les fréquences  $\text{C}_\alpha$ ,  $\text{C}_\beta$  et  $\text{H}_\beta$  des résidus alanines.

décrit qui utilise comme précurseur de l'oxométhionine entièrement deutérée à l'exception du méthyle- $\epsilon$  [6, 24]. Il ne nous reste maintenant à mettre au point que des protocoles permettant d'introduire en position proR des leucines et valines,  $\gamma^2$  des isoleucines et  $\gamma$  des thréonines des groupes  $^{13}\text{CH}_3$ . Pour les positions proR et  $\gamma^2$ , cela ne constitue théoriquement qu'une variante du protocole présenté ci-dessus pour les positions proS des valines et leucines. Quant au marquage de la thréonine, pour des raisons de mécanismes réactionnels des enzymes impliquées dans les voies de biosynthèse, il paraît encore difficilement accessible en utilisant seulement des précurseurs marqués et une production *in vivo*. L'utilisation de la thréonine protonée uniquement sur le méthyle  $\gamma$ , obtenue par synthèse chimique, reste quant à elle possible, mais à un coût important étant donné l'existence de deux centres chiraux dans cette molécule et donc un rendement maximum de seulement 25 % en composé utile.

Les protocoles que nous venons de décrire ont déjà permis d'obtenir des résultats prometteurs. La dynamique du fonctionnement de complexes du protéasome d'Archaea a ainsi pu être étudiée [25-26]. Comme l'illustre la figure 10, il a aussi été possible grâce à une cartographie de déplacements chimiques, de déterminer les résidus leucine et valine affectés par la liaison d'un inhibiteur du site actif de la protéine TET2 – un oligomère de 480 kDa [21].

La maîtrise du marquage pour chaque type de méthyle devrait permettre d'accéder à des informations structurales, dynamiques et d'interactions au sein de machines biomoléculaires responsables des principales fonctions d'une cellule vivante. La combinaison de l'utilisation de précurseurs convenablement enrichis isotopiquement et aisément synthétisables avec celle des voies de biosynthèse pourrait également constituer une voie originale de synthèse de composés marqués difficilement accessibles par une autre méthode.

## Remerciements

Les auteurs tiennent à associer à ces résultats tous ceux qui ont contribué à ce projet : Olivier Hamelin, Rémy Sounier, Isabel Ayala, Michael Plevin, Carlos Amaro, Bruno Franzetti, et le personnel de la plateforme de marquage isotopique du PSB « Partnership for Structural Biology », Grenoble. Nous remercions les « Très Grandes Infrastructures de Recherche (TGIR)-RMN » pour l'accès à la plateforme RMN hauts champs de l'Institut de Biologie Structurale (IBS). Ce travail effectué à l'IBS est soutenu par le CNRS, le CEA et l'Université Joseph Fourier, par le Programme interdisciplinaire du CNRS 2007 « Interface physique, biologie et chimie : soutien à la prise de risque », ainsi que par l'Agence Nationale de la Recherche (ANR PIRibio2009-Nano-Life@Work).

## Note et références

- (1) L'*auxotrophie* est l'incapacité d'un organisme vivant de synthétiser un composé organique nécessaire à son développement. Par exemple, une *E. coli* portant une

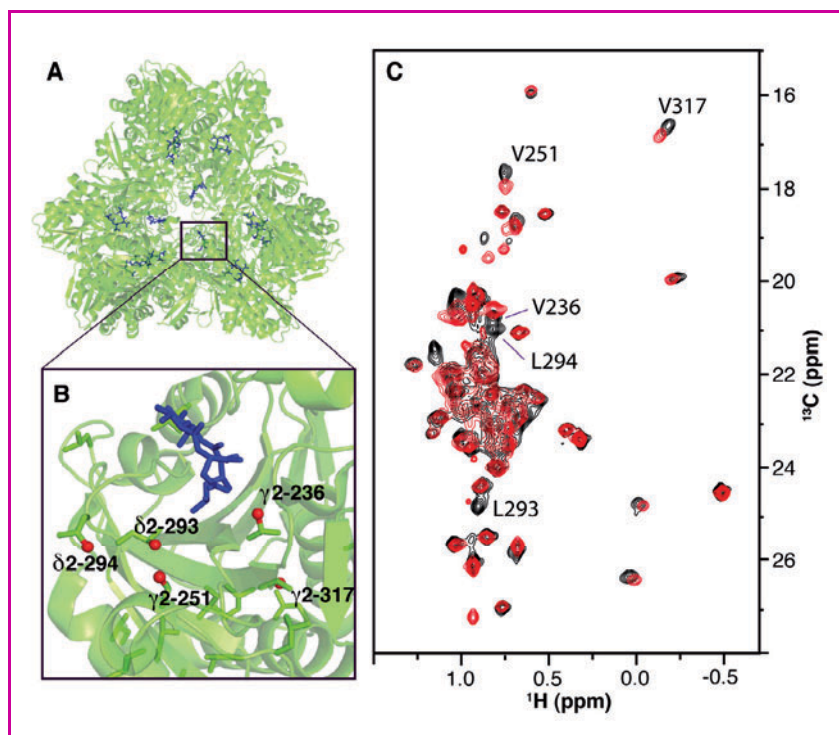


Figure 10 - Interaction entre TET2 (468 kDa), une peptidase impliquée dans la protéolyse chez les archaebactéries et un inhibiteur, l'amastatine.

**A** : structure quaternaire du dodécamère de TET2 en complexe avec l'amastatine (n° PDB : 1Y0Y). **B** : détail de la structure 3D montrant la position des méthyles proS des leucines et valines situés à moins de 6 Å de l'amastatine (sphères rouges). **C** : superposition des spectres 2D méthyl-TROSY de U- $^2\text{H}$ ], U- $^{12}\text{C}$ ], U- $^{13}\text{C}_1\text{H}_3$ ]proS-Leu/Val TET2 libre (en noir) ou en complexe avec l'amastatine (en rouge).

Reproduit avec permission de [21], © Wiley-VCH Verlag GmbH & Co. KGaA.

mutation bloquant la synthèse d'isoleucine est auxotrophe à l'isoleucine et ne pourra se développer dans un milieu que si de l'isoleucine y est apporté (d'après Wikipédia).

- [1] Wüthrich K., *NMR of Proteins and Nucleic Acids*, Wiley, **1986**.
- [2] Boisbouvier J., Albrand J.P., Blackledge M., Jaquinod M., Schweitz H., Lazdunski M., Marion D., A structural homologue of colipase in black mamba venom revealed by NMR floating disulphide bridge analysis, *J. Mol. Biol.*, **1998**, 283(1), p. 205.
- [3] Ikura M., Kay L.E., Bax A., A novel approach for sequential assignment of  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin, *Biochemistry*, **1990**, 29(19), p. 4659.
- [4] Garrett D.S., Seok Y.J., Peterkofsky A., Gronenborn A.M., Clore G.M., Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr, *Nat. Struct. Biol.*, **1999**, 6(2), p. 166.
- [5] Tugarinov V., Kay L.E., Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G., *J. Mol. Biol.*, **2003**, 327(5), p. 1121.
- [6] Gelis I., Bonvin A.M., Keramisanou D., Koukaki M., Gouridis G., Karamanou S., Economou A., Kalodimos C.G., Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR, *Cell*, **2007**, 131(4), p. 756.
- [7] Amaro C., Schanda P., Durá M.A., Ayala I., Marion D., Franzetti B., Brutscher B., Boisbouvier J., Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies, *J. Am. Chem. Soc.*, **2009**, 131(10), p. 3448.
- [8] Muchmore D.C., McIntosh L.P., Russell C.B., Anderson D.E., Dahlquist F.W., Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, *Methods Enzymol.*, **1989**, 177, p. 44.
- [9] Yokoyama S., Protein expression systems for structural genomics and proteomics, *Curr. Opin. Chem. Biol.*, **2003**, 7(1), p. 39.
- [10] Kainosho M., Torizawa T., Iwashita Y., Terauchi T., Mei Ono A., Güntert P., Optimal isotope labelling for NMR protein structure determinations, *Nature*, **2006**, 440(7080), p. 52.
- [11] Isaacson R.L., Simpson P.J., Liu M., Cota E., Zhang X., Freemont P., Matthews S., A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues, *J. Am. Chem. Soc.*, **2007**, 129(50), p. 15428.
- [12] Ayala I., Sounier R., Usé N., Gans P., Boisbouvier J., An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein, *J. Biomol. NMR*, **2009**, 43(2), p. 111.
- [13] Gardner K.H., Rosen M.K., Kay L.E., Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR, *Biochemistry*, **1997**, 36(6), p. 1389.
- [14] Goto N.K., Gardner K.H., Mueller G.A., Willis R.C., Kay L.E., A robust and cost-effective method for the production of Val, Leu, Ile ( $\delta$  1) methyl-protonated  $^{15}\text{N}$ -,  $^{13}\text{C}$ -,  $^2\text{H}$ -labeled proteins, *J. Biomol. NMR*, **1999**, 13(4), p. 369.
- [15] Tugarinov V., Kay L.E., 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.*, **2003**, 125(45), p. 13868.
- [16] Sylvester S.R., Stevens C.M., Stereospecificity of the reductoisomerase-catalyzed step in the pathway of biosynthesis of valine and leucine, *Biochemistry*, **1979**, 18(21), p. 4529.
- [17] Crout D.H., Hedgecock C.J., Lipscomb E.L., Armstrong F.B., Stereochemistry of valine biosynthesis. Configuration of the product of rearrangement of  $\alpha$ -acetolactate, *Eur. J. Biochem.*, **1980**, 110(2), p. 439.
- [18] Lee M., Kim D.H., Syntheses and kinetic evaluation of racemic and optically active 2-benzyl-2-methyl-3,4-epoxybutanoic acids as irreversible inactivators for carboxypeptidase A, *Bioorg. Med. Chem.*, **2002**, 10(4), p. 913.
- [19] Baucherel X., Levoirier E., Uziel J., Jugé S., Monohydroxylation of cyclic and acyclic  $\beta$ -keto esters with molecular oxygen catalyzed by cobalt(II) chloride in neutral conditions, *Tetrahedron Letters*, **2000**, 41(9), p. 1385.
- [20] Crout D.H., Hedgecock C.J., The base-catalysed rearrangement of  $\alpha$ -acetolactate (2-hydroxy-2-methyl-3-oxobutanoate): a novel carboxylate ion migration in a tertiary ketol rearrangement, *J. Chem. Soc., Perkin Trans. 1*, **1979**, p. 1982.
- [21] Gans P., Hamelin O., Sounier R., Ayala I., Durá M.A., Amaro C.D., Noirclerc-Savoye M., Franzetti B., Plevin M.J., Boisbouvier J., Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins, *Angew. Chem. Int. Ed. Engl.*, **2010**, 49(11), p. 1958.
- [22] Sounier R., Blanchard L., Wu Z., Boisbouvier J., High-accuracy distance measurement between remote methyls in specifically protonated proteins, *J. Am. Chem. Soc.*, **2007**, 129(3), p. 472.
- [23] Nibeiliú M.E., Malthouse J.P., The stereospecificity and catalytic efficiency of the tryptophan synthase-catalysed exchange of the  $\alpha$ -protons of amino acids, *Biochem. J.*, **2004**, 381(Pt 3), p. 847.
- [24] Fischer M., Kloiber K., Hausler J., Ledolter K., Konrat R., Schmid W., Synthesis of a  $^{13}\text{C}$ -methyl-group-labeled methionine precursor as a

useful tool for simplifying protein structural analysis by NMR spectroscopy, *ChemBioChem*, **2007**, 8, p. 610.

- [25] Sprangers R., Kay L.E., Quantitative dynamics and binding studies of the 20S proteasome by NMR, *Nature*, **2007**, 445(7128), p. 618.
- [26] Religa T.L., Sprangers R., Kay L.E., Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR, *Science*, **2010**, 328(5974), p. 98.



J. Boisbouvier



P. Gans

**Jérôme Boisbouvier** est directeur de recherche au CNRS et **Pierre Gans**, ingénieur-chercheur CEA, au Laboratoire de Résonance Magnétique Nucléaire, Institut de Biologie Structurale « Jean-Pierre Ebel », Grenoble\*. Pour ses travaux, Jérôme Boisbouvier a reçu la **Médaille de bronze du CNRS en 2008**.

\* Laboratoire de Résonance Magnétique Nucléaire, Institut de Biologie Structurale « Jean-Pierre Ebel », UMR 5075 CEA/CNRS/Univ. Joseph Fourier, 41 rue Jules Horowitz, F-38027 Grenoble Cedex 1.

Courriels : jerome.boisbouvier@ibs.fr, pierre.gans@ibs.fr



## La RMN à la Portée de Tous

Fourier 300

### ● Haute Résolution, Ultra Compact, Coût Réduit Fourier 300

Le Fourier 300 est un nouveau Spectromètre de RMN Haute Résolution dédié, conçu pour l'analyse chimique de routine, le contrôle qualité, l'enseignement, à un coût abordable.

Découvrez le en ligne:  
[www.bruker.fr/fourier300](http://www.bruker.fr/fourier300)

think forward

RMN