

NMRによるタンパク質立体構造解析の最近の進歩

久米田 博之¹, 稲垣 冬彦^{1,2}

(¹北海道大学大学院薬学研究院構造生物学研究室, ²北海道大学大学院生命科学院構造生物学研究室)

はじめに

溶液 NMR 法によるタンパク質立体構造解析は、大まかに 1) 試料調製, 2) 各種 NMR 測定, 3) 信号帰属, 4) 構造計算の流れで行われる。1) の試料調製では、大腸菌等の微生物を用いた安定同位体元素 ¹³C/¹⁵N によるタンパク質の二重標識化に加え、無細胞タンパク質合成系技術の利用による多彩な標識体作成が可能になった。2) の測定に関しては装置開発や測定技術開発が行われ、測定感度の向上および測定時間の大幅な短縮が可能となった。3) の信号帰属では主鎖および側鎖帰属の自動化ソフトウェアの開発とともに、信号帰属とスペクトル閲覧の対話性を向上させたソフトウェアが開発され構造解析の高速化を推進した。4) の構造計算では距離情報となる核オーバーハウザー効果 (NOE) 信号の自動帰属ソフトウェアが広く用いられるようになり、構造計算にかかる時間が飛躍的に短縮された。これらの最近の技術革新によって、数年前までは 1 個につき 1 年から数年程度かかっていた NMR 法によるタンパク質立体構造解析に要する時間は、約 1 カ月程度に短縮され、解析可能な測定濃度・分子量限界も改善した。

本稿では、溶液 NMR 立体構造解析における最近の技術革新について実用的な観点より述べるとともに筆者の構造解析手順について紹介し、実際に解いたいくつかの構造について紹介する。

1. 試料調製

溶液 NMR 構造解析を目的としたタンパク質試料調製で

は、従来の微生物を利用した安定同位体 (¹³C および ¹⁵N) 標識化以外に無細胞タンパク質合成系が用いられるようになった。無細胞タンパク質合成系では発現を担う遺伝子およびアミノ酸を試験管内で混合し、大腸菌やその他の生物種から抽出した転写・翻訳に必須な成分(リボソームなど)を添加することでタンパク質を合成する。無細胞タンパク質合成系では、基質としたアミノ酸がそのまま目的タンパク質中に取り込まれるため、任意のアミノ酸を選択的に標識化することができる。また、菌体培養や形質転換体作成が不要であり、さらに鋳型遺伝子に PCR 産物を用いることも可能であるため、遺伝子から目的タンパク質を取得するまでの時間を大幅に短縮できる。これは網羅的構造解析におけるボトルネックとなっているコンストラクトスクリーニングにおいても非常に有用である。近年、立体特異的に標識化されたアミノ酸、SAIL (stereo-array isotope labelling) アミノ酸合成技術が開発された¹⁾。この SAIL アミノ酸を用いた無細胞タンパク質合成系により測定上の様々な利益がもたらされ、NMR 法による高分子量タンパク質の高精度構造解析が期待される。

2. 測定技術の進歩

溶液 NMR 法によるタンパク質立体構造解析における問題点として測定濃度と測定時間がある。数年前まで NMR 法による立体構造解析に最低限必要なタンパク質濃度は 1 mM 程度であり、三次元 NMR の測定には 3 日から 1 週間程度かかるため、解析に必要である全スペクトルの測定には 3 カ月程度を要した。そのため、測定試料は溶解度が高くかつ長期間安定な物に限られていた。測定感度を上昇させる装置開発の成果である低温プローブが広く普及した²⁾。検出コイルを極低温 (25K) まで冷却することにより、コイルの熱雑音を抑え従来と比較してエチルベンゼンの ¹H 信号強度比は 3 から 4 倍になった。実際のタンパク質測定においては溶媒中に含まれる塩など伝導性物質の影響から測定感度の向上は 2 倍程度のものである。この技術によりもたらされた測定感度の上昇は測定に必要な最低タンパク質濃度を飛躍的に改善し、天然比で存在する ¹³C や ¹⁵N を利

Progress in NMR structure analysis of protein
Hiroyuki Kumeta¹, Fuyuhiko Inagaki^{1,2} (Laboratory of Structural Biology, Graduate School of Pharmaceutical Sciences, Hokkaido University, ²Laboratory of Structural Biology, Graduate School of Life Science, Hokkaido University, Frontier Research Center for Post-genomic Science and Technology, Kita 21 Nishi 11, Kita-ku, Sapporo 001-0021, Japan)

テクニカルノート

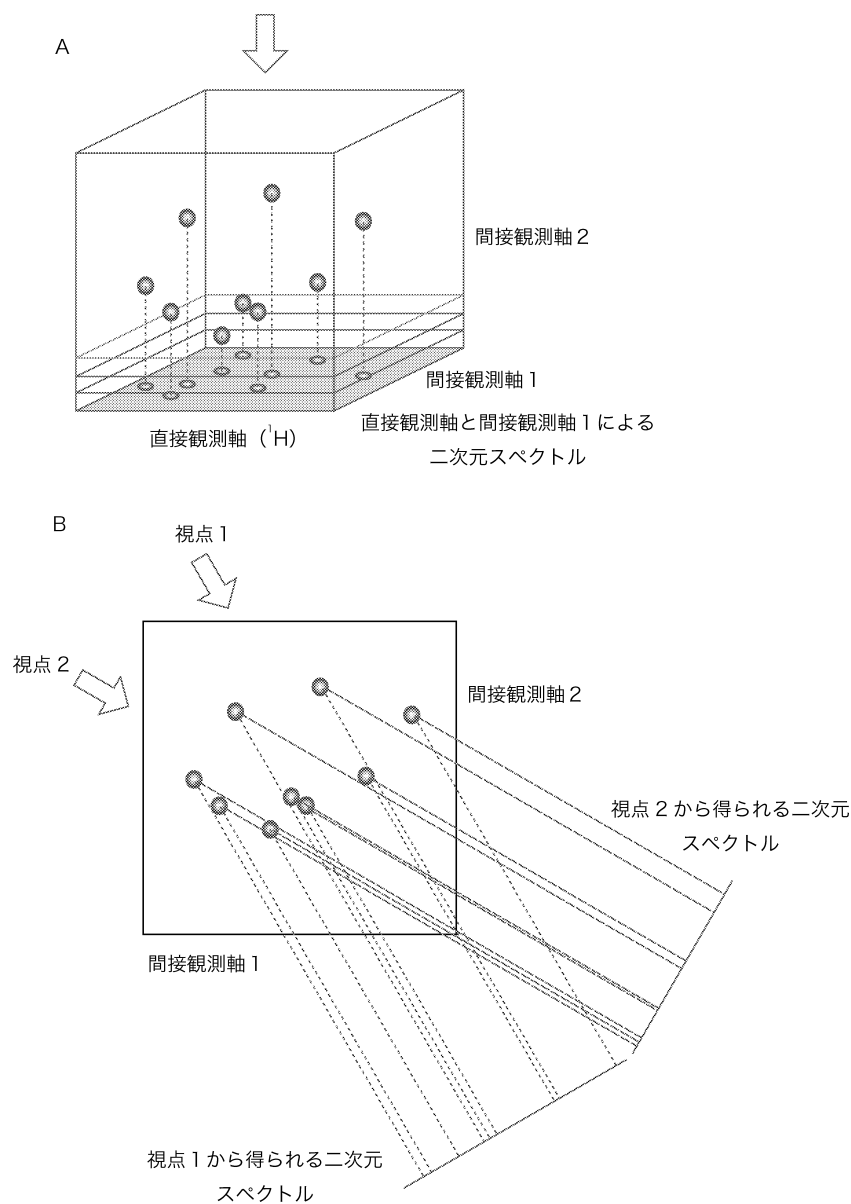


図 1 Projection Reconstruction の概念図

- A. 通常の測定手法による三次元立方体スペクトル
一定方向からの視点でスライスした複数の二次元平面スペクトルを得る。三次元立方体スペクトル上で観測される各信号は直接観測軸と間接観測軸 1 による二次元平面スペクトル上に投影される。
- B. 三次元 Projection Reconstruction 法に利用される投影二次元の概念図
A の三次元立方体スペクトルをある角度から (視点 1 および 2) 俯瞰した投影二次元を得る。

用した二次元または三次元 NMR 測定も現実的な測定時間で観測可能になった。

測定時間短縮のための新たな測定技術として projection reconstruction (PR) 法が確立されつつある³⁾。通常の三次元 NMR 測定では二つの間接観測軸の展開時間をそれぞれ独立に変化させ測定する。これは、周波数軸における三次

元立方体スペクトルをある一定の方向からの視点でスライスした複数の二次元平面スペクトルを測定していることになる。PR 法では二つの間接観測核の展開時間を三角関数に従って連動させることにより、三次元立方体スペクトルをある角度から俯瞰した投影二次元スペクトルを得る (図 1)。視点の角度を変更した複数枚の投影二次元スペクトル

から三次元立方体スペクトルを再構築する。この技術によって観測信号数が少ない三次元測定の場合（例えばHNCOやHN(CO)CAなど）は、最低3枚の二次元測定を行うだけですべての観測信号の帰属を行うことができる。理論上、10倍程度の測定時間の短縮が期待され、観測信号数が多い場合を含めても実用的には3~5倍程度の時間短縮が可能になった。PR法は複数の間接観測核軸の展開時間を任意の法則に従って変動させるG-matrix Fourier transform (GFT) 法⁹⁾の一種である。GFT法を利用した他の測定手法によって、四次元以上の高次元測定も複数の低次元測定の組み合わせにより補完することも可能となった。また、1種類の多次元測定による主鎖連鎖帰属も可能になった。他に測定時間の短縮方法として、複数の展開時間の間接観測核測定を一度に行うsingle scan法⁵⁾も報告されている。

3. 帰属操作

NMR信号の帰属は、主鎖帰属と側鎖帰属に大別される。主鎖帰属では¹H-¹⁵N HSQCによって観測される主鎖アミド由来の信号がどの残基に由来するのかを決定する。主鎖帰属が行われると、測定対象タンパク質とその標的分子との残基レベルでの相互作用解析や、分子内部の運動性を議論するための緩和時間解析を行うことができる。側鎖帰属では主鎖以外のNMR観測可能なほぼ全ての信号について帰属を行い、より詳細な相互作用解析や立体構造計算などに用いる。主鎖連鎖帰属用測定(¹H-¹⁵N HSQC, HNCO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HN(CA)CO, HN(CA)HA, HN(CO)HAやHBHA(CO)NHなど)から得られる信号情報を元にした自動帰属ソフトウェアが多数報告されている。さらに、脂肪族側鎖を含めた自動帰属⁹⁾や、すでにX線法などによって立体構造が明らかになっている場合にその構造情報を踏まえて主鎖の自動帰属も行えるようになった^{7,8)}。また従来の主鎖連鎖帰属用測定以外に、パルス系列によるアミノ酸選択的な信号観測法⁹⁾や前述の試料調製の項で述べた選択的にアミノ酸を標識化したタンパク質を用い、残基特異的に信号を観測することも行われるようになった。現在では100残基程度の低分子量タンパク質で良好なスペクトルを与える場合には、前述の自動帰属ソフトウェアを用いてほぼ完全な帰属が可能になったが、より大きな分子量を持つ場合や解析が困難である場合には、解析者の手動による帰属補正が必要になる。この帰属補正を支援するソフトウェアがいくつか開発されている。IBIS¹⁰⁾, Smartnotebook¹¹⁾, AutoLink/CARA¹²⁾は主鎖連鎖帰属と各種スペクトルとの相関を支援する。当研究室で開発中であるOlivia (<http://fermi.pharm.hokudai.ac.jp/olivia/>にて公開)は、主鎖のみならず側鎖も含めた全ての帰属操作と各種スペクトルとの相関を支援している。

4. 主鎖の帰属補正手法

前述の主鎖連鎖帰属用測定はすべて¹H-¹⁵N HSQCを基本としている。そのため、¹H-¹⁵N HSQC上で重なっている信号に関して信号情報の取り違えが起こりうる。また、高分子量タンパク質では分子運動性の減少からHNCACBやCBCA(CO)NHの信号強度、特にCβの信号強度が極端に減少することがある。そのため筆者は、連鎖帰属に分子量によらず測定感度が良好な¹H-¹³C HSQCやCCH-TOCSYを用いて信号情報の補正を行っている。¹H-¹⁵N HSQC上で重なりあう信号についてHN(CO)CA, HNCA, HN(CA)HA, HBHA(CO)NHの候補となるCαおよびHαの信号情報の組み合わせの交差信号を¹H-¹³C HSQCにおいてそれぞれ確認する。Cβの情報を補完するためにはCα-Hα交差信号におけるCCH-TOCSYを用いる。連鎖帰属の最終確認には¹⁵N edited NOESYを用いる。互いに隣接する主鎖アミドにおける¹⁵N edited NOESYは、間にある側鎖水素や互いのアミド水素が共通して観測されるため似た信号出現になる。

5. 構造計算の自動化/高速化

NMR法によるタンパク質立体構造計算では、計算機上で疑似高温環境下における伸びたタンパク質構造から、NOE信号によって得られる距離束縛情報や他の束縛情報を満たすよう環境温度を下げながら構造を収束させる焼きなまし法が用いられる。NOE信号の帰属操作は構造計算に与える原子間距離束縛情報に直接関係しており、ここでの帰属間違いは計算結果である構造に大きな影響を与える。NOE信号を観測するNOESY測定には通常、起点となる水素と直接結合する¹³Cや¹⁵Nで展開するNOESY HSQCが用いられる。NOESY HSQCによって観測される信号数は数千個に達し、手動による帰属操作は解析者の多大な作業時間と労力を必要とした。このNOE信号の自動帰属を行うソフトウェアとして、ARIA¹³⁾およびCYANA¹⁴⁾がある。両ソフトウェアともに、構造計算と構造に基づくNOE信号の帰属を繰り返し最終構造を得る。また、化学シフト値が近いために重複したNOE信号に対して、複数の帰属を試みるアルゴリズムを持つ。これらのソフトウェアによって、人為性をなるべく排したNOE信号帰属および構造計算が行えるようになった。各ソフトウェアの特徴としてARIAは、NOE信号以外の有力な構造計算束縛情報の一つである残余双極子相互作用(RDC)情報を容易に計算に組み込むことができる。一方、CYANAはNOE

テクニカルノート

信号帰属効率を上昇させる二つのアルゴリズム (network anchoring, constraint combination) を有しているため初期構造を求めやすい。最近公開された CYANA 2.0 以降では、化学シフトが未帰属である原子の NOE 信号情報からの自動帰属や光学異性化原子の自動帰属なども可能になった。また、ARIA 2.2 においても NOE 信号帰属効率を上昇させる 2 種類のアルゴリズムの実装が今年度の ICMRBS にて発表されたが、執筆時ではまだ公開されていない。また ARIA, CYANA とともに Linux および構造計算の分散化に対応している。そのため、高価な UNIX 計算機や他の高性能計算機に頼らずとも安価な複数台の Linux 計算機による高速化が可能である。現在では 1 台数万円程度の計算機 10 台で構成される構造計算クラスターによって、両ソフトウェアともに 1 回の標準的な構造計算時間は 1 時間程度となっている。

正しい NOE 信号情報および正しい帰属情報を揃えることができれば、どちらの自動構造計算ソフトウェアを用いても正しい構造計算結果を与える。しかし実際の NOESY スペクトル中には測定ノイズが多く含まれており、そのノイズを NOE 信号情報として与えたがために間違った構造へ収束することが多い。特に NOE 信号帰属効率をあげるアルゴリズムはノイズの影響を受けやすい。そこでノイズ除去を含めた NOE 信号情報の見直しが必要になる。見直しには自動構造計算ソフトウェアによる NOE 信号帰属情報、相補性信号の有無、信号の強度および線形などを総合的に判断しながら修正を行う。前述の Olivia は構造計算結果の NOE 信号情報および構造情報を取り込むことができ、スペクトルを確認しながらの NOE 信号情報の見直し作業を支援する。

6. CYANA による立体構造決定までの手法

筆者は主に CYANA を用いて構造計算を行っている。構造計算に与える NOE 信号情報は、1) あらかじめ出現が予測される NOE 信号 (残基内や隣接残基間) を拾い上げておき、2) HSQC 上に観測される交差信号を中心とした短冊状の信号拾い上げを行う (図 2A)。このときの信号拾い上げは、1) ですでに拾い上げた NOE 信号を除外している。3) 得られた信号情報から構造計算を行い、初期構造を求める。4) 正しく折り畳まれた初期構造および NOE 信号の自動帰属情報から NOESY スペクトル上での信号情報の見直しを行い、再度の構造計算を行う。3) の見直し作業では 2) で新たに拾い上げた信号について検証を行う (図 2B)。NOESY HSQC から得られるそれぞれ数千個の信号の大部分は残基内および隣接残基間であり、1) により見直すべき信号数は劇的に減少する。3) の短冊状

の信号拾い上げによって NOE 信号が確実に観測されない領域におけるノイズを大幅に軽減できる。また、1) で拾い上げた NOE 信号の除去は NOE 信号の間違った自動帰属を防ぐこともでき、安定した初期構造を求める上でも有用であった。4) の見直し作業はまず CYANA によって構造計算に利用されなかった NOE 信号について検討を行い、最終的には全ての信号について検討を行う。

7. 立体構造解析例

筆者らがここ数年の間に NMR 法によって決定した立体構造について紹介する。

図 3-A はイネ由来葉緑体局在型 NifU 様タンパク質である NifU1A の domain II (L154-S226) の立体構造である。葉緑体局在型 NifU 様タンパク質には互いに配列相同性のある二つのドメインがある。N 末端側ドメインには NifU 様タンパク質の活性部位である鉄硫黄クラスター結合モチーフが存在するが、C 末端側ドメインにはモチーフが存在せず機能未知であった。NifU1A domain II 表面は、他の NifU 様タンパク質と異なり全体に正の静電ポテンシャルを持つ。葉緑体局在型 NifU 様タンパク質の鉄硫黄クラスター移送先であるフェレドキシンが負の静電ポテンシャルをタンパク質表面を持つことから、domain II はフェレドキシンとの相互作用に寄与することが考えられた。

図 3-B はイネ由来チオレドキシン (M1-L132) の立体構造である。¹³C/¹⁵N 標識タンパク質を取得してから測定時間を含めて 3 週間で決定した。この立体構造は既知の大腸菌由来およびヒト由来チオレドキシンと同じ骨格構造を持ち、さらに N 末端に伸びた構造を持つ。この N 末端領域は、植物内細胞間輸送に関わる構造体である原形質連絡の透過性に寄与することが報告¹⁵⁾されており、今後の研究展開が期待される。

図 3-C はイネ由来ユビキチン結合酵素の一つ、UBC2 (M1-D152) の立体構造である。イネ完全長 cDNA データベースより 20 種類を超えるユビキチン結合酵素アイソフォームが確認されており、これらアイソフォームの系統解析により植物固有種の存在が示唆されている。イネ由来ユビキチン結合酵素アイソフォーム立体構造の網羅解析により、植物特有ユビキチン系の原子レベルでの機構解明への寄与が期待される。

図 3-D はマウス由来 NIPP1 の FHA ドメイン (M1-T132) の立体構造である。Ser/Thr 残基特異的プロテインホスファターゼ PP1 の核内制御サブユニットである NIPP1 は、PP1 との結合を介しての遺伝子発現の転写後調節や pre-mRNA のスプライシング制御に関与している。NIPP1 の N 末端に位置する FHA ドメインは転写調節因子やスプライ

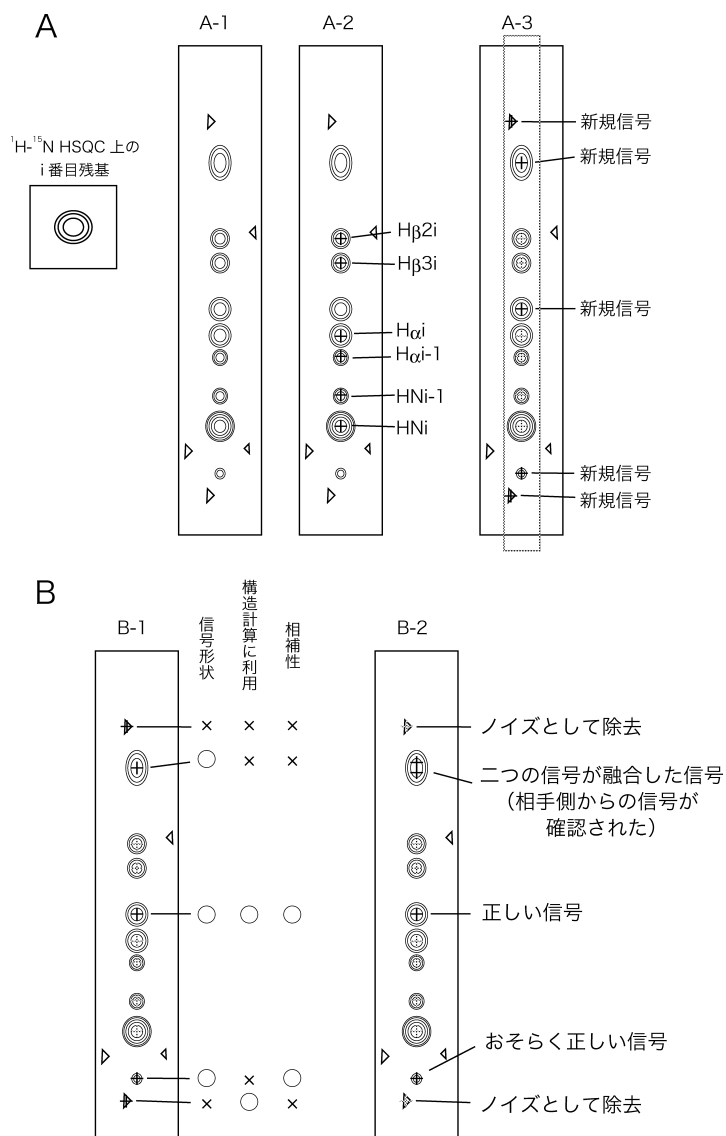


図2 NOE 信号情報の作成(A)および見直し作業(B)

- A-1. i 番目の主鎖アミドプロトンから ^{15}N edited NOESY 上で観測される信号
- A-2. 配列および化学シフト情報から出現が予測される信号の拾い上げ
 i 番目の主鎖アミドプロトン近傍にある i 番目残基の $\text{H}\alpha$, $\text{H}\beta$ および $i-1$ 番目の $\text{H}\alpha$, HN について拾い上げている。
- A-3. ^1H - ^{15}N HSQC 上の信号を中心とした短冊状の拾い上げ
 A-2 で拾い上げた信号を除外している。
- B-1. 信号の見直しのための参考情報
 A-3 で新規に拾い上げた信号について検証している。それぞれの信号について信号形状、構造計算時の帰属情報、相補性について検証する。
- B-2. 次回構造計算に用いる信号情報
 B-1 の検証から、信号別に可否判断を行う。検証後の信号情報による構造計算を行い、検証を繰り返す。

テクニカルノート

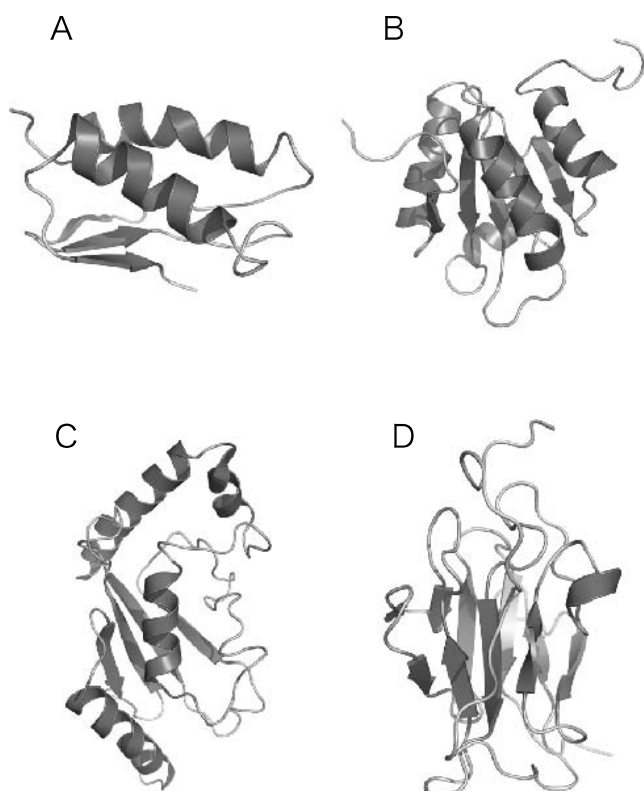


図3 NMR法により筆者らが決定した立体構造

- A. イネ由来 NifU1A domain II の立体構造 (PDB ID: 1TH5)
 B. イネ由来チオレドキシンの立体構造 (PDB ID: 1WMJ)
 C. イネ由来ユビキチン結合酵素, UBC2 の立体構造
 D. マウス由来 NIPP1 の FHA ドメインの立体構造

シング調節因子との結合が確認されており, NIPP1 FHA ドメインの立体構造解析や分子レベルでの認識機構解明により, 転写やスプライシング調節における新たな知見が期待される。

謝辞

本稿で紹介した立体構造は, 農業生物資源研究所の加藤悦子博士ならびに北海道大学遺伝子病制御研究所の菊池九二三教授, 田沼延公助手との共同研究により行われたものであり, 感謝の意を表します。

- 1) Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Ono, A.M., & Güntert, P. (2006) *Nature*, 440, 52–57.
- 2) Kovacs, H., Moskau, D., & Spraul, M. (2005) *Prog. Nucl. Magn. Reson. Spectrosc.*, 46, 131–155.
- 3) Kupce, E. & Freeman, J. (2003) *J. Biomol. NMR*, 27, 383–387.
- 4) Kim, S. & Szyperski, T. (2003) *J. Am. Chem. Soc.*, 125, 1385–1393.
- 5) Frydman, L., Scherf, T., & Lupulescu, A. (2002) *Proc. Natl. Acad. Sci. USA*, 99, 15858–15862.
- 6) Eghbalian, H.R., Bahrami, A., Wang, L., Assadi, A., & Markley, K.L. (2005) *J. Biomol. NMR*, 32, 219–233.
- 7) Pristovsek, P., Rüterjans, H., & Jerala, R. (2002) *J. Comput. Chem.*, 23, 335–340.
- 8) Jung, Y. & Zweckstetter, M. (2004) *J. Biomol. NMR*, 30, 11–23.
- 9) Schubert, M., Labudde, D., Leitner, D., & Oschkinat, H. (2005) *J. Biomol. NMR*, 31, 115–127.
- 10) Hyberts, S.G. & Wagner, G. (2003) *J. Biomol. NMR*, 26, 335–344.
- 11) Slupsky, C.M., Boyko, R.F., Booth, V.K., & Sykes, B.D. (2003) *J. Biomol. NMR*, 27, 313–321.
- 12) Masse, J.E. & Keller, R. (2005) *J. Magn. Reson.*, 174, 133–151.
- 13) Linge, L.P., Habeck, M., Rieping, W., & Nilges, M. (2003) *Bioinformatics*, 19, 315–316.
- 14) Güntert, P. (2003) *Progr. NMR Spectrosc.*, 43, 105–125.
- 15) Ishiwatari, Y., Fujiwara, T., McFarland, K.C., Nemoto, K., Hayashi, H., Chino, M., & Lucas, W.J., (1998) *Planta*, 205, 12–22.