

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 9 日現在

機関番号：14401

研究種目：基盤研究(C)

研究期間：2010～2013

課題番号：22570116

研究課題名(和文)データベースを有効活用した学習型NMRシグナル自動解析システムの開発

研究課題名(英文)Automated NMR signal assignment with self-learning system efficiently using database

研究代表者

小林 直宏 (Kobayashi, Naohiro)

大阪大学・たんぱく質研究所・研究員

研究者番号：80272160

交付決定額(研究期間全体)：(直接経費) 3,600,000円、(間接経費) 1,080,000円

研究成果の概要(和文)：タンパク質の核磁気共鳴(NMR)による解析は生体高分子の構造に関する情報に加え、リガンド相互作用、動的状態など創薬に関する重要な情報を与える。本研究課題ではNMRシグナルの解析を全自動化するプログラム群を開発する。当初の目標としていた自動的なNMR主鎖シグナルの連鎖帰属を実用的な精度で実行できる性能を実現した。本研究課題の研究開発過程で、NMR構造と化学シフト値との整合性の定量的な評価について、およびNMR構造の重ね合わせ自動化について2報の論文を国際誌に発表できた。開発されたプログラムは統合解析プラットフォームに搭載し、すでに大阪大学蛋白質研究所のホームページよりデモバージョンを公開した。

研究成果の概要(英文)：NMR analysis provides not only structural information of macromolecules but also the important information about ligand interaction and dynamics which are related to drug design. In this project, a system including programs which automate NMR signal assignments. The performance that are expected from beginning of the project has been achieved a practical automated sequence specific assignments of backbone NMR signals. During this project, a quantitative assessment of consistency between NMR structure and assigned chemical shifts, and a method for automated identification of ordered region in NMR structure ensemble were published in the international journals. The programs developed in this project have been implemented in an integrate platform for NMR analysis and its demo version can be available from the website on the institute for protein research in University Osaka.

研究分野：生物学

科研費の分科・細目：生物科学・構造生物化学

キーワード：NMR データベース シグナル帰属 自動化 機械学習

1. 研究開始当初の背景

(1) タンパク質の核磁気共鳴 (NMR) による解析は生体高分子の3次元構造に関する情報に加え、リガンド相互作用、動的ゆらぎ状態など創薬に関する重要な情報を与える。NMR シグナルの帰属には三重共鳴法による3次元スペクトル群に対してアミノ酸配列特異的連鎖帰属(以下連鎖帰属)を行う方法がほぼ標準化しており、100残基程度のタンパク質で1~2日、200残基程度のタンパク質で5~10日あれば解析が終了する。しかしながらこの方法はスペクトルの感度、分離の状態に強く依存しているため場合によっては数ヶ月かかることもある。一方、既知の立体構造情報を用いることでシグナル解析を容易にすることも可能であるが、既知情報を取り入れての解析は極めて慎重に行う必要があり、多くの実験研究者から敬遠されてきた。

(2) 近年のPDBデータベースの充実化によりターゲットとする生体高分子のホモログに対する立体構造情報を得る可能性は年々増加する傾向にあり、それらを用いたNMRデータの解析は実用的なレベルに達している。更に解析結果が既知情報と一致しない場合、知見としての高い新規性が期待できる。これは既知情報に頼り過ぎないバランスの良い解析システムを構築できた場合に有効であるが、現在一般的に利用できるツール群では達成できていない。

2. 研究の目的

本研究では研究代表者が既にプロトタイプを完成し、その有効性が実証されているNMRシグナル解析用のサブルーチン群を用いることでNMRシグナル解析を全自動化するプログラム群を開発する。更にそれらのプログラム群を統合管理し、教師データを学習させることで解析パラメータを自動的に最適化する学習型自動解析システムを開発する。このシステムは既知データを利用する上での恣意性を極力排除させ、不注意によるエラーの検出ばかりではなく新規性の高い発見の実験的裏付けに対して威力を発揮させることが期待できる。

3. 研究の方法

本研究課題ではNMRシグナルの解析を全自動化する以下のプログラム群を開発する：1. ホモロジーモデリングを統合的に実行する model_robot、2. 推定構造からの化学シフトを予測させるプログラムの出力プログラム shift_robot、3. NMR データからのシグナルを自動検出させる spec_robot、4. アニールによる連鎖帰属を実現する anneal_robot、5. 磁化移動グラフを自動構

築する graph_robot、6. 1~5のプログラム群が入出力するパラメータを自動学習により最適化させる par_robot を開発し、GPU演算ユニットなどによる並列演算によりパラメータ群を最適化しつつプログラム群を統合的にコントロールするNMRシグナルの自動帰属プログラムとして完成させる。

4. 研究成果

平成22年度ではグラフ理論を適用し、帰属確率行列を評価関数として用いる anneal_robot が完成し、Modeller による構造予測と、TALOS による主鎖二面角予測法と組み合わせることで、難度の高い実験データについてもアミノ酸配列特異的連鎖帰属を極めて高い精度で実現できることを示し、その研究成果は国際学会において発表された[学会発表]。平成23年度においては完成したプロトタイプのシステムを立命館大学、京都大学、広島大学など数拠点の外部研究者による試用を実施した。興味深いことに、2~3個の構造状態を化学交換するサンプルデータにおいてもコンホマーごとにその存在確率を明示できる計算結果を与えることに成功した。この研究成果は国際誌で発表された[雑誌論文]。また、平成23年度ごろから、Intel プロセッサのプロセッサーが22nm 台に入るなど高密度化が進みマルチコア化、低価格化が進んだ。この理由でシステムを拡張性、管理運用性の高い仮想OSへ移行し、CPU コア数を積極的に増やす構成に開発計画を最適化した。平成24年度現在では合計50コアを超える構成となり、研究課題開始直後の性能から20~40倍程度のパフォーマンスを発揮する、GPU演算ユニットによる演算性能に匹敵する計算機環境を実現した。研究代表者が独自に開発したKUJIRAおよび原野陽子氏(大阪大学)が開発したBESSを用い、タンパク3000プロジェクトにおいて解析された600件以上のNMRデータのBMRBへの登録を理化学研究所との共同研究により実現した。これらの登録されたタンパク質立体構造座標についてSPARTA+を用いた化学シフトの逆計算を行い、帰属された化学シフトのデータのうち、 $^{13}\text{C}\alpha$ 、 $^{13}\text{C}\beta$ シグナルとの整合性を確認したところ、解析された立体構造と帰属された化学シフトとの整合性は他の同時期に登録されたデータ群に比べ極めてよいことが示された[図1]。この研究成果により、化学シフトとNMR構造あるいはモデル構造との整合性を定量的に評価可能になったためNMRシグナル自動帰属

におけるパラメータ最適化などへの応用が可能になったといえる。この研究成果に関

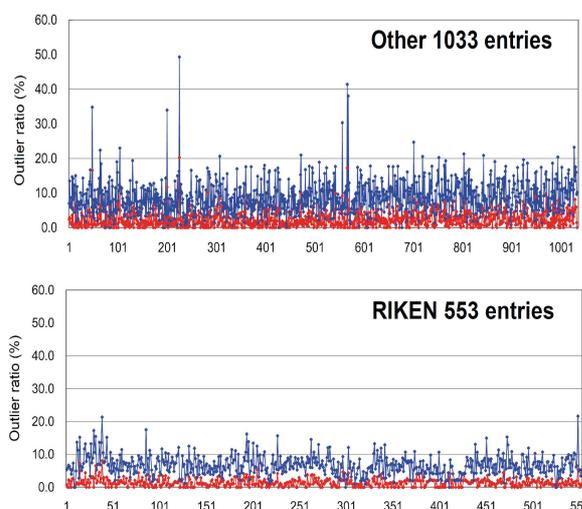


図 1 化学シフトの整合性評価の結果比較

しては雑誌論文 に掲載された。また平成 25 年度においては NMR 構造の収斂領域を完全自動的に検出し、重ね合わせを自動的に実行する fit_robot を開発し、その詳細は雑誌論文 に掲載された。開発されたプログラム群はシグナル自動帰属プラグインとしてパッケージ化され、研究代表者がすでに開発している統合解析プラットフォームである MagRO-NMRView に搭載させあるいはスタンドアロンなツールとして、大阪大学蛋白質研究所のホームページよりデモバージョンを公開している。

<http://bmrdep.protein.osaka-u.ac.jp/en/nmrtoolbox/>

統合的な自動帰属システム全体に関する詳細についても学会発表 ~ のとおり国内外の学会への発表を行い、国際誌への投稿も予定している。解析システムを搭載したデモバージョンを用いて多くの NMR 解析が行われそれらの研究成果は雑誌論文、~、~ においてすべて査読ありの国際誌にて掲載された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 11 件)

Naohiro Kobayashi, A robust method for quantitative identification of ordered cores in an ensemble of biomolecular structures by non-linear multi-dimensional scaling using inter-atomic distance variance matrix., J. Biomol. NMR, 査読

有、58巻、2014、61-67

DOI: 10.1007/s10858-013-9805-z

Takashi Nagata T, Kengo Tsuda, Naohiro Kobayashi, Peter Guntert, Shigeyuki Yokoyama, and Yutaka Muto, 1H, 13C, and 15N resonance assignments of the dsRBDs of mouse RNA helicase A., Biomolecular NMR Assignment, 査読有、7巻、2013、69-72

DOI:10.1007/s12104-012-9380-3

小林直宏、ビッグデータ時代の生体高分子 NMR データベース、BioMagResBank (BMRB) その有効利用とわが国における現状、日本核磁気共鳴学会機関誌、査読有、58巻、2013、121-124

FaHu He, Makoto Inoue, Takanori Kigawa, Mari Takahashi, Kanako Kuwasako, Kengo Tsuda, Naohiro Kobayashi, Takako Terada, Mikako Shirouzu, Peter, Guntert, Shigeyuki Yokoyama, and Yutaka Muto, Solution structure of the splicing factor motif of the human Prp18 protein., Proteins, 査読有、80巻、2012、968-974
DOI:10.1002/prot.24003

Takako Ohyama, Takashi Nagata, Kengo Tsuda, Naohiro Kobayashi, Takako Imai, Hideyuki Okano, Toshio Yamazaki, and Masato Katahira, Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions, Nucleic Acids Research, 査読有、40巻、2012、3218-3221
DOI:10.1093/nar/gkr1139

Takashi Nagata, Kengo Tsuda, Naohiro Kobayashi, Mikako Shirouzu, Takanori, Kigawa, Peter Guntert, Shigeyuki Yokoyama, and Yutaka Muto, Solution structures of the double-stranded RNA-binding domains from rna helicase A., Proteins, 査読有、80巻、2012、1699-1706
DOI:10.1002/prot.24059

Ayako Furukawa, Hideyasu Okamura, Ryo Morishita, Satoko Matsunaga, Naohiro Kobayashi, Takahisa Ikegami, Tsutomu Kodaki, Akifumi Takaori-Kondo, Ryo Akihida, Takashi Nagata, Masato Katahira, NMR study of xenotropic murine leukemia virus-related virus protease in a complex with amprenavir., Biochem Biophys Res Commun., 査読有、24巻、2012、284-289
DOI:10.1016/j.bbrc.2012.07.083

Naohiro Kobayashi, Yoko Harano, Naoya Tochio, Eiichi Nakatani, Takanori Kigawa, Shigeyuki Yokoyama, Steven Mading, Eldon L. Ulrich, John L. Markley, Hideo Akutsu and Toshimichi Fujiwara, An automated system designed for large scale NMR data deposition and

annotation: Application to over 600 assigned chemical shift data entries to the BioMagResBank from the RIKEN Structural Genomics/Proteomics Initiative internal database., *J. Biomol. NMR*, 査読有、53巻、2012、311-320
DOI: 10.1007/s10858-012-9641-6.

Fahu He, Kengo Tsuda, Mari Takahashi, Keiko Kuwasako, Takako Terada, Mikako Shirouzu, Satoru Watanabe, Takanori Kigawa, Naohiro Kobayashi, Peter Guntert, Shigeyuki Yokoyama, Yutaka Muto, Structural insight into the interaction of ADP-ribose with the PARP WWE domains., *FEBS Lett.*, 査読有、586巻、2012、3858-3864
DOI:10.1016/j.febslet.2012.09.009.

Shino Mizuno, Hirokazu Amida, Naohiro Kobayashi, Shin-Ichi Aizawa, Shin-Ichi Tate, The NMR Structure of FliK, the Trigger for the Switch of Substrate Specificity in the Flagellar Type III Secretion Apparatus, *J. Mol. Biol.* 409巻、2011、558-573
DOI: 10.1016/j.jmb.2011.04.008

Kengo Tsuda, Tatsuhiro Someya, Kanako Kuwasako, Mari Takahashi, FaHu He, Satoru Unzai, Makoto Inoue, Takushi Harada, Satoru Watanabe, Takako Terada, Naohiro Kobayashi, Mikako Shirouzu, Takanori Kigawa, Akiko Tanaka, Sumio Sugano, Peter Güntert, Shigeyuki Yokoyama, Yutaka Muto, Structural basis for the dual RNA-recognition modes of human Tra2-beta RRM, *Nuc. Acids, Res.* 39、2011、1538-1553
DOI:10.1093/nar/gkq854

〔学会発表〕(計 8件)

天野 亮,野村祐介,永田崇,小林直宏, 森瑶子, 福永淳一, 田中陽一郎, 片平正人, 中村義一, 神津知子, 坂本泰一、転写因子AML1に対して複数の結合部位を持つ高親和性RNAアプタマーの解析、第36回日本分子生物学会年会、2013年12月06日～2013年12月06日、神戸市

天野亮,野村祐介,永田崇,小林直宏, 森瑶子, 福永淳一, 田中陽一郎, 片平正人, 中村義一, 神津知子, 坂本泰一、Multi-site binding of high affinity aptamer against AML1 Runt domain、第40回国際核酸化学シンポジウム、2013年12月03日～2013年12月03日、横浜市

永田 崇, 白川貴恵, 小林直宏, 始平堂弘和, 片平正人, 堀澤健一, 土居信英, 柳川弘志、第52回NMR討論会、2013年11月12日～2013年11月12日、金沢市

Naohiro Kobayashi, International Conference on Structural Genomics 2013、2013年07月29日～2013年07月29日、札幌市

小林直宏、仮想化技術を用いたNMRデータの統合的解析環境、第50回日本生物物理学会年会、2012年09月23日～2012年09月23日、名古屋市

小林直宏、化学シフトデータベースを利用した高精度なNMR構造評価システム、第51回NMR討論会、2012年11月09日～2012年11月09日、名古屋市

Naohiro Kobayashi, Takeshi Iwata, Yoko Harano, Chojiro Kojima, Haruki Nakamura, Toshimichi Fujiwara, New server systems based on virtual computing for the database BMRB and MagRO-Tools for NMR analysis、第50回記念NMR討論会国際シンポジウム、2011年11月15-18日、横浜

Naohiro Kobayashi, Ayako Furukawa, Soichiro Kitazawa, Takashi Nagata, Ryo Kitahara, Masato Katahira and Toshimichi Fujiwara, MagRO-Assign: a tool for automated sequence specific assignment of chemical shifts for NMR signals of protein backbone, EUROMAR2011, 2011年8月21-25、Frankfurt (Germany)

〔その他〕

ホームページ等

<http://bmrdep.protein.osaka-u.ac.jp/en/nmrtoolbox/>

6. 研究組織

(1)研究代表者

小林 直宏 (KOBAYASHI, Naohiro)

研究者番号：80272160

(2)研究分担者 なし

()

研究者番号：

(3)連携研究者 なし

()

研究者番号：