

2020年1月31日
理化学研究所
東京工業大学
科学技術振興機構

タンパク質の構造や動きを解析する新技術を開発

—情報・数理科学の応用による NMR 法の革新—

理化学研究所（理研）生命機能科学研究センター細胞構造生物学研究チームの葛西卓磨研究員（科学技術振興機構（JST）さきがけ研究者）、木川隆則チームリーダー（東京工業大学情報理工学院特定教授）、東京工業大学情報理工学院情報工学系の小野峻佑准教授（JST さきがけ研究者）らの共同研究グループ^{*}は、核磁気共鳴（NMR）法^[1]に情報・数理科学の手法を応用することで、従来は解析が困難だった重なり合う NMR 信号^[1]を分離し、タンパク質の構造や動きなどに関する情報を得る新たな方法「SiPex（Stable-isotope-assisted Parameter extraction）法」を開発しました。

本研究成果は、タンパク質の機能に関する基礎研究に貢献し、タンパク質と医薬品候補分子の結合状態の解析に基づく創薬研究を加速させると期待できます。

NMR 法は、強い磁場中に置かれた原子核から発せられる信号（NMR 信号）を観測し、分子の構造を解析する手法です。タンパク質の解析では、NMR 法で観測可能な安定同位体^[2]で標識した試料を用いることが標準的です。

今回共同研究グループは、先行研究で開発した「符号化標識法^[3]」と数理科学の応用により、重なり合う複数の NMR 信号からでも、アミノ酸の情報とタンパク質の性質の情報を取得することに成功しました。

本研究は、科学雑誌『*Journal of Biomolecular NMR*』のオンライン版（1月30日付）に掲載されました。

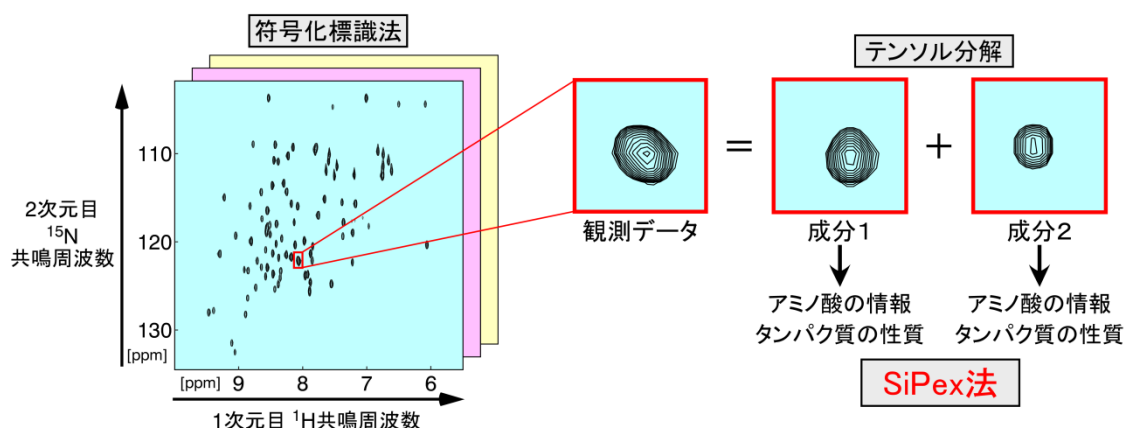


図 情報科学（符号化標識法）と数理科学（テンソル分解）を応用した SiPex 法

※共同研究グループ

理化学研究所 生命機能科学研究センター 細胞構造生物学研究チーム

研究員 葛西 卓磨 (かさい たくま)

(科学技術振興機構 (JST) さきがけ研究者)

チームリーダー 木川 隆則 (きがわ たかのり)

(東京工業大学情報理工学院特定教授)

東京工業大学 情報理工学院 情報工学系

准教授 小野 峻佑 (おの しゅんすけ)

(科学技術振興機構 (JST) さきがけ研究者)

東北大学 東北メディカル・メガバンク機構

機構長 山本 雅之 (やまもと まさゆき)

(東北大学大学院医学系研究科教授)

教授 小柴 生造 (こしば せいぞう)

京都大学大学院 情報学研究科

教授 田中 利幸 (たなか としゆき)

統計数理研究所

教授 池田 思朗 (いけだ しろう)

※研究支援

本研究の一部は、科学技術振興機構 (JST) 戦略的創造研究推進事業個人型研究 (さきがけ) 「計測技術と高度情報処理の融合によるインテリジェント計測・解析手法の開発と応用 (研究総括: 雨宮慶幸、副研究総括: 北川源四郎)」の研究課題「試料への情報の符号化を活用する NMR 計測・解析法 (研究者: 葛西卓磨)」および「統合的凸最適化による In Hand な成分分離型信号情報再構成 (研究者: 小野峻佑)」、日本学術振興会 (JSPS) 科学研究費補助金若手研究(B) 「立体構造とアミノ酸判別に基づく NMR シグナル帰属法の開発 (研究代表者: 葛西卓磨)」および新学術領域研究「スパースモデリングによる NMR 計測・解析の高速高精度化 (研究代表者: 木川隆則)」、JST 戦略的創造研究推進事業 CREST 「ライフサイエンスの革新を目指した構造生命科学と先端的基盤技術」の研究課題「NMR と計算科学の融合による in situ 構造生物学の確立と真核細胞内蛋白質の動態研究への応用」による支援を受けて行われました。

1. 背景

アミノ酸が連なってできたタンパク質は、その構造と動きによって分子機能が決まります。そのため、タンパク質の構造や動きの解析は、その機能を解明する上で重要です。核磁気共鳴 (NMR) 法は、強い磁場中に置かれた原子核に電磁波を照射すると、原子の置かれた状況によって特徴的な信号 (NMR 信号) が観測される物理現象 (核磁気共鳴) を応用した、分子構造などの解析法です。生体環境に近い溶液中でタンパク質分子に含まれる原子を直接観測できるという特長から、タンパク質の構造・機能・動き・作用機序の解明に欠かせません。

NMR 法でタンパク質の解析を行うには、まず、複数の原子から発せられるそれぞれの NMR 信号が、どの原子に由来するかを決める「信号帰属」を行います。タンパク質の解析では、炭素-13 (^{13}C) や窒素-15 (^{15}N) の安定同位体でアミノ

酸を標識し、信号帰属を効率良く行うことが標準的な方法となっています。

測定されたNMR信号は二次元平面上のスペクトル^[4]で表すことができますが、複数の信号が同じ位置に重なると、信号帰属が困難になります。NMR信号の重なりは、異なる位置の原子に由来する信号が分離しにくい天然変性タンパク質^[5]、信号の数が多量な高分子量タンパク質、一つ一つの信号の広がりが大きくなってしまいう生細胞内のタンパク質などで著しく生じ、タンパク質の一部の構造や動きについての情報しか得られず、解析は不完全なものになってしまいます。

信号が重なった場合にも信号帰属を補助する情報を提供し、重なった信号を分離しつつ、タンパク質の性質を解析する方法があれば、これまで解析が困難であったタンパク質の部位をもNMR法によって解析することができます。また、解析は可能であっても信号帰属に時間がかかっていたタンパク質を、迅速に解析できるようになることも期待されます。

2. 研究手法と成果

葛西卓磨研究員らは、2015年に「符号化標識法」を開発しました^{注)}。NMR解析における安定同位体標識の標準的な方法では、タンパク質中の安定同位体の有無の情報だけを利用するのに対し、符号化標識法では、アミノ酸の情報をタンパク質に「符号化」し、アミノ酸ごとに異なる安定同位体標識パターン(符号語)を割り当てたタンパク質試料を用います。一例としては、¹⁵Nでの三つの標識率(100%、75%、50%)と、¹³Cでの三つの標識率(100%、50%、0%)の組み合わせでアミノ酸を標識したタンパク質(標識体)を3パターン用意し、それぞれをNMRで計測します。これにより、観測されたNMRスペクトルから、アミノ酸の情報を「復号」して取り出せるようにしておきます(図1)。

注) Kasai T, Koshiba S, Yokoyama J, Kigawa T (2015), Stable isotope labeling strategy based on coding theory. *J Biomol NMR*, 63:213-221

アミノ酸	標識体 1		標識体 2		標識体 3	
	¹³ C	¹⁵ N	¹³ C	¹⁵ N	¹³ C	¹⁵ N
G (グリシン)	● 100	● 100	● 100	● 100	● 100	● 100
F (フェニルアラニン)	● 100	● 100	● 100	● 100	◐ 50	◑ 75
Q (グルタミン)	● 100	● 100	● 100	● 100	○ 0	◑ 50
Y (チロシン)	● 100	● 100	◐ 50	◑ 75	● 100	● 100
M (メチオニン)	● 100	● 100	◐ 50	◑ 75	◐ 50	◑ 75
V (バリン)	● 100	● 100	◐ 50	◑ 75	○ 0	◑ 50
N (アスパラギン)	● 100	● 100	○ 0	◑ 50	● 100	● 100
D (アスパラギン酸)	● 100	● 100	○ 0	◑ 50	◐ 50	◑ 75
S (セリン)	● 100	● 100	○ 0	◑ 50	○ 0	◑ 50
H (ヒスチジン)	◐ 50	◑ 75	● 100	● 100	● 100	● 100
E (グルタミン酸)	◐ 50	◑ 75	● 100	● 100	○ 0	◑ 50
P (プロリン)	◐ 50	◑ 75	◐ 50	◑ 75	● 100	● 100
I (イソロイシン)	◐ 50	◑ 75	○ 0	◑ 50	● 100	● 100
K (リジン)	○ 0	◑ 50	● 100	● 100	● 100	● 100
R (アルギニン)	○ 0	◑ 50	● 100	● 100	◐ 50	◑ 75
A (アラニン)	○ 0	◑ 50	● 100	● 100	○ 0	◑ 50
L (ロイシン)	○ 0	◑ 50	◐ 50	◑ 75	● 100	● 100
T (トレオニン)	○ 0	◑ 50	○ 0	◑ 50	● 100	● 100

● 100% ¹⁵N ◑ 75% ¹⁵N ◑ 50% ¹⁵N
 ● 100% ¹³C ◐ 50% ¹³C ○ 0% ¹³C

図 1 符号化標識法における安定同位体標識パターン（符号語表）の例

アミノ酸ごとに異なる安定同位体標識パターンを対応させる「符号語表」に従い、3 パターンのタンパク質試料（標識体）を用意する。ユビキチンタンパク質に含まれる 18 種類のアミノ酸を判別可能な符号語表の一例を示す。アミノ酸の情報を「復号」するには、標識体ごとに NMR 計測を行い、同位体の比率に伴う NMR 信号の強度の違いを利用する。例えば標識体 1 では、¹⁵N の標識率が 100% であるアミノ酸は 9 種類まで絞られる。さらに、標識体 2 と標識体 3 の情報を合わせると、全ての標識体において ¹⁵N の標識率が 100% の強度として観測される NMR 信号は、G（グリシン）に由来すると判断できる。

本研究で開発した「SiPex 法」は、符号化標識法を応用したもので、この標識パターンを単にアミノ酸の情報を取り出すだけではなく、重なった信号を分離する手掛かりとしても用います。まず、符号化標識したタンパク質試料を用いて、アミノ酸の情報を得るための測定と、¹⁵N 緩和速度^[6]などタンパク質の性質を得るための測定を行い、これらを組み合わせで解析します。測定ごとに得られる NMR スペクトルは 2 次元ですが、それらを集めて、全体では 4 次元のテンソル^[7]であると見なし、テンソル分解^[7]という数理解析手法で信号ごとに分離します。分離された各成分の 3 次元目のシグナル強度にはアミノ酸の情報が含まれており、符号語表と照らし合わせて解析することで、アミノ酸の情報を復号することができます。また、各成分の 4 次元目の強度は、¹⁵N 緩和速度の情報を持っており、タンパク質の動きの情報を取り出すことができます。このように、タンパク質の NMR 解析において障害となっていた信号の重なりの問題を解決しつつ、信号帰属の助けとなるアミノ酸情報と、タンパク質の動きなど性質に関する情報を同時に取り出すことができるのが、SiPex 法の特長です（図 2）。

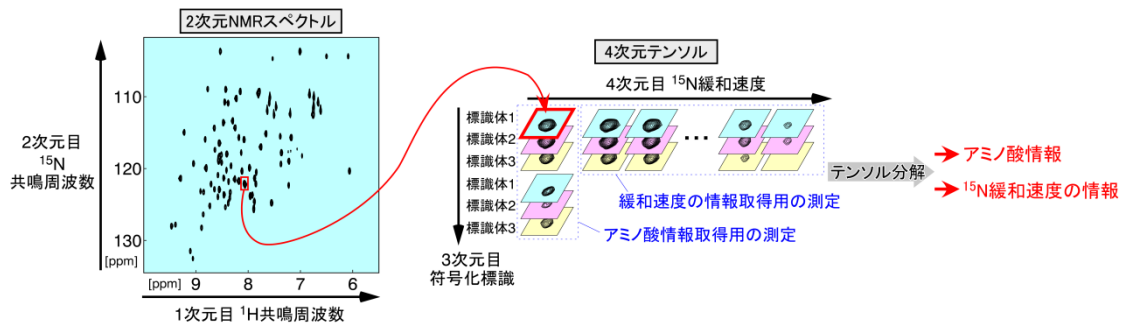


図2 符号化標識法で得られたスペクトルを用いた、信号の分離と情報の取出し

左図は、特定のアミノ酸に由来する NMR 信号の分布 (スペクトル) を、横軸を ^1H 共鳴周波数、縦軸を ^{15}N 共鳴周波数として表したもの。赤四角で囲んだ点は、二つの NMR 信号が同じ位置で重なったものだが、テンソル分解により二つのアミノ酸情報に分離することができる。さらに、テンソル分解を利用して ^{15}N 緩和速度を測定することで、タンパク質の動きの情報が得られる。

共同研究グループは、SiPex 法の性能を評価するため、ユビキチンと呼ばれるタンパク質をモデルにした実証実験を行いました。従来の NMR 法によるタンパク質の測定では、15~16 番目の LE (ロイシン-グルタミン酸) というアミノ酸の並びと、25~26 番目の NV (アスパラギン-バリン) というアミノ酸の並びに対応するシグナルが重なってしまうことが分かっています。SiPex 法による解析でも、重なったシグナルが確かに二つの成分で構成されていることが分かり、それぞれの成分から正しいアミノ酸情報を取り出し、かつ、 ^{15}N 緩和速度の情報を取り出すことができました (図 3)。このことは、従来の技術では解析が困難であったタンパク質に対して、SiPex 法が有効であることを示しています。

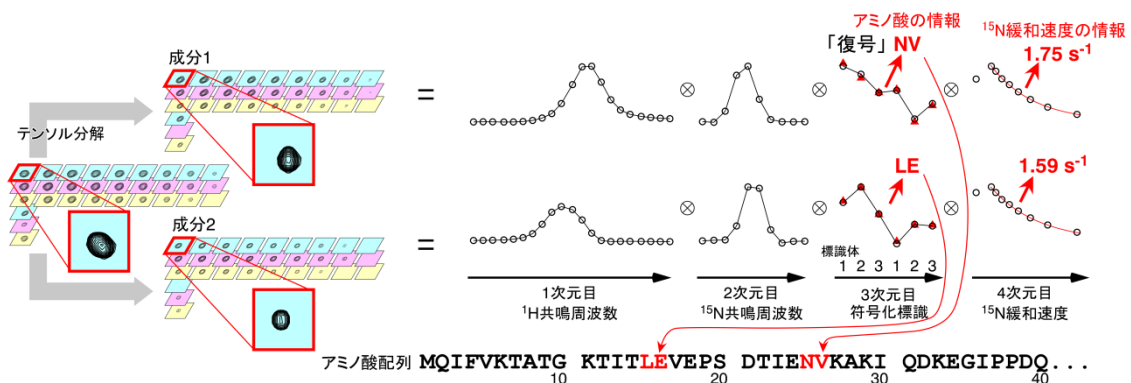


図3 SiPex によるユビキチンタンパク質の重なった信号の分離と情報の取出し

84 個のアミノ酸から成るユビキチンの変異タンパク質を用いた SiPex 法の実証実験。重なり合った信号が、テンソル分解により二つの成分に分けられた。各成分に含まれるアミノ酸情報を取り出すと、成分 1 に対応するのは 25~26 番目の NV (アスパラギン-バリン) であり、成分 2 に対応するのは 15~16 番目の LE (ロイシン-グルタミン酸) であることが分かった。また、 ^{15}N 緩和速度の情報を取り出すと、どちらも同じくらい動き (揺らぎ) のある領域であることが読み取れた。

3. 今後の期待

SiPex 法は、NMR を用いて信号の分離を行いながらタンパク質の情報を取得する新しい手法であり、タンパク質の構造、動き、他の分子との結合などの解析に用いることができます。符号化・復号という情報科学の手法と、テンソル分解という数理科学の手法を組み合わせることで、タンパク質の NMR 解析という生命科学の手法を革新したもので、異分野の研究者の共同研究によって実現されました。

SiPex 法を用いることで、従来の NMR 法では信号が重なることでタンパク質の性質に関する情報が得られなかったタンパク質の解析が可能になるだけでなく、SiPex 法で得られるアミノ酸の情報は信号帰属にも有用なため、信号帰属が困難であったタンパク質の解析も可能になります。特に、天然変性タンパク質の解析や、生きている細胞内のタンパク質を直接観測し、解析する in-cell NMR 法^[8]などにおいて、解析可能なタンパク質の種類を増やすことにつながり、タンパク質の機能や作用機構の解明という基礎研究に役立ちます。また、信号帰属に役立つアミノ酸の情報と、タンパク質の性質に関する情報がリンクした形で得られるため、従来の NMR 法で解析可能であったタンパク質についても、より迅速・効率的に解析することができます。生命科学における基礎研究の加速はもちろんのこと、医薬品候補分子との結合を迅速に評価し改善につなげる必要がある創薬研究など応用研究にも役立つと期待できます。

4. 論文情報

<タイトル>

Amino-acid selective isotope labeling enables simultaneous overlapping signal decomposition and information extraction from NMR spectra

<著者名>

Takuma Kasai, Shunsuke Ono, Seizo Koshiba, Masayuki Yamamoto, Toshiyuki Tanaka, Shiro Ikeda, and Takanori Kigawa

<雑誌>

Journal of Biomolecular NMR

<DOI>

10.1007/s10858-019-00295-9

5. 補足説明

[1] 核磁気共鳴 (NMR) 法、NMR 信号

強い磁場中に置かれた原子核に電磁波を照射すると、核スピンの共鳴現象により、原子核の性質や周囲の環境に応じた周波数 (共鳴周波数) の電磁波の吸収や放出が起こるが、その電磁波を NMR 信号として捉えることで、物質の分子構造の解析や物性の解析を行う手法。分子の相互作用などの情報も得られるため、生命科学、医薬、化学、

食品、材料物性といった幅広い分野で利用されている。NMR は Nuclear Magnetic Resonance の略。

[2] 安定同位体

原子番号が同じで質量の異なる同位体のうち、放射性崩壊を起こさず安定に存在するもの。タンパク質の主要な構成元素は水素、炭素、窒素であるが、それぞれ自然界では水素-1 (^1H)、炭素-12 (^{12}C)、窒素-14 (^{14}N)がほとんどを占めている。このうち、 ^1H は NMR 法で観測可能だが、 ^{12}C 、 ^{14}N は観測不能か困難である。そのため、物理化学的性質がほとんど変わらない安定同位体であり NMR 観測が可能な ^{13}C 、 ^{15}N に置き換えたタンパク質試料を用いることが標準的な方法となっている。安定同位体に置き換えることで NMR 観測が可能になることから、興味のある原子を観測するための標識としても利用できる。

[3] 符号化標識法

葛西卓磨研究員らが 2015 年に発表したタンパク質の標識方法。少ない種類の標識体で 20 種類全てのアミノ酸を判別することを目指した。SiCode (Stable isotope encoding) 法ともいう。

[4] スペクトル

各周波数において、どれくらいの強度の信号が観測されたかを表すデータ。NMR 法はラジオ波の周波数領域の電磁波を扱う分光法の一つであり、NMR 信号を共鳴周波数ごとに分解したスペクトルで表すことができる。

[5] 天然変性タンパク質

溶液中での構造変化が大きく、タンパク質の一部、もしくは全体にわたって一定の構造をとらないタンパク質。従来の構造生物学では解析の対象になりにくかったが、近年では機能との関連が注目されている。

[6] ^{15}N 緩和速度

タンパク質分子に含まれるアミド ($-\text{NH}$) 基の窒素 NMR 信号の減衰の速さ (^{15}N 緩和速度) は、タンパク質の動きの大きさや速さを反映することが知られている。 ^{15}N 標識したタンパク質試料を用い、2 次元 NMR スペクトル上の各信号の減衰を測定する。タンパク質を構成するアミノ酸のほとんどにアミド基が含まれることから、タンパク質分子中のさまざまな箇所の動きを解析できる。

[7] テンソル、テンソル分解

テンソルは、本研究においては数値を 3 次元以上に並べたもの。1 個の数値 (0 次元) はスカラー、数値を 1 列 (1 次元) に並べたものはベクトル、数値を 2 次元に並べたものは行列と呼ばれる。テンソル分解 (テンソル因子分解ともいう) は、テンソルをより単純な形で表すこと。本研究においては、テンソルをベクトルの直積の和というより単純な形で表す方法 (polyadic decomposition) によってテンソル分解をしている。

[8] in-cell NMR 法

水溶液中のタンパク質を NMR 測定する通常の方法に対して、生きた細胞内に導入し

たんぱく質を測定する方法。細胞内はたんぱく質や核酸などの生体分子が高濃度で存在する「分子混雑」の状態、水溶液中とは異なった環境である。in-cell NMR は実際に機能している環境下のたんぱく質の解析が可能な方法として注目されている。

6. 発表者・機関窓口

<発表者> ※研究内容については発表者にお問い合わせください。
理化学研究所 生命機能科学研究センター 細胞構造生物学研究チーム
研究員 葛西 卓磨 (かさい たくま)
チームリーダー 木川 隆則 (きがわ たかのり)

東京工業大学 情報理工学院 情報工学系
准教授 小野 峻佑 (おの しゅんすけ)
TEL : 045-924-5089 E-mail : ono[at]c.titech.ac.jp

<JST事業に関すること>
科学技術振興機構 戦略研究推進部 グリーンイノベーショングループ
中村 幹 (なかむら つよし)
TEL : 03-3512-3525 FAX : 03-3222-2066
E-mail : presto[at]jst.go.jp

<機関窓口>
理化学研究所 生命機能科学研究センター センター長室 報道担当
山岸 敦 (やまぎし あつし)
TEL : 078-306-3095 FAX : 078-306-3090
E-mail : ayamagishi[at]riken.jp

理化学研究所 広報室 報道担当
TEL : 048-467-9272 FAX : 048-462-4715
E-mail : ex-press[at]riken.jp

東京工業大学 広報・社会連携本部 広報・地域連携部門
TEL : 03-5734-2975 FAX : 03-5734-3661
E-mail : media[at]jim.titech.ac.jp

科学技術振興機構 広報課
TEL : 03-5214-8404 FAX : 03-5214-8432
E-mail : jstkoho[at]jst.go.jp

※上記の[at]は@に置き換えてください。