実験と理論のハイブリッドから機能に関わる タンパク質のダイナミクスを観る

京都大学大学院理学研究科化学専攻 北尾彰朗

A novel method has been developed in order to determine both structure and dynamics of proteins using a molecular simulation and experimental data. This is based on the concept of the Jumping-Among-Minima model. In this method, inter-substate fluctuations are determined by molecular simulation and inter-substate motions are determined so as to reproduce experimental data. Slow dynamics of proteins and its relation to protein functions are discussed.

protein dynamics / function / collective motion / jumping-among-minima / NMR

1. なぜ今ダイナミクスなのか？

現在、生物の遺伝子を解読するプロジェクトが急速に進んでおり、それに伴ってタンパク質の立体構造もつぎつぎと決定されつつある。こうしたプロジェクトは遅かれ早かれ数年のうちにはかなりの成果をあげるであろうと期待されるから、その成果をもとにして今後どのように生命現象の理解が進んでいくのかを予想してみよう。リンネ的な分類学の進展、すなわち生物圏を形成している個々の種を系統的に分類することが近代生物学のひとつのスタート地点であったと考えると、遺伝子が解読されタンパク質の立体構造が決められて形態的・機能的に分類されていくことも新しい生物物理学のスタートとなるのではないか。

生物圏の構成要素（種）が明らかになることで生態学的な観点が生まれてきたように、生体系を構成する遺伝子やタンパク質などの要素がすべてリストアップされれば、それらがシステムとしてどのように機能するかという、よりマクロな観点が重要となってくるのではないか。一方、生体系を構成する各要素がどのようなメカニズムで機能を発現するかという、よりミクロな観点も必要となってくる。タンパク質が固有の立体構造をもつことは機能を発現する上で重要ではあるが、立体構造がわかっただけでは機能がなぜ発現されるのかを十分理解することはできない。生体系の構成要素がわかったときに、その各要素の働く原理をブラ

ックボックスとしておくのではなく、分子機械としての動作メカニズム、すなわちダイナミクスを解明していくことが、生命現象を理解するためのひとつの鍵となる。この総説では、今後重要性を増していくであろうタンパク質ダイナミクス研究の現状の一端をご紹介します。

2. タンパク質の個性とタンパク質に共通する物性

タンパク質を研究対象として観るときには、さまざまな興味に基づいた見方がある。タンパク質は多くの場合、生物学的な観点からその機能が注目されることが多い。すなわち、タンパク質それぞれの個性が問題にされることが多いのである。しかし、個々のタンパク質の性質を問題とするならば、タンパク質一般に共通する性質をも理解する必要がある。一般的な性質を理解していればこそ、何が個性であるかがわかるのだ。

タンパク質の個性と共通する物性に関して、ダイナミクスの観点から一つの例を挙げてみよう。図1は、タンパク質がどのような振動数の振動モードをどのくらいもっているかを表したものである。タンパク質の大きさや構造が異なっても、状態密度のスペクトルのかたちはピークの高さに至るまではほぼ一致している。つまり、ここで示した状態密度の概形は、タンパク質の種類・大きさや構造に依存しない一般的な性質を表している。しかし、スペクトルの微細構造に注目すれ

A Novel Method to Investigate Functionally-relevant Protein Dynamics

Akio KITAO

Department of Chemistry, Graduate School of Science, Kyoto University

図1 基準振動解析から得られたタンパク質の振動数分布
(a) メリチン (アンフォールド状態), (b) メリチン, (c) クランピン, (d) BPTI, (e) ヒトリゾチーム. (a) ~ (d) の縦軸はヒトリゾチームの自由度との比でスケールしてある.

ば、ピークの高さや位置に微妙な違いがある。これはタンパク質の種類や構造の違いによってあらわれてくる個性である。

この状態密度に対応する物理量を何らかの方法で測定したとする。スペクトルの概形がわかる程度の解像度であれば、タンパク質に一般的な性質を観ていることになる。もし、ピークの微細構造を議論できる解像度があれば、さらにタンパク質の構造変化や個性を議論できるのである。また、各モードが具体的にどの部位のどのようなゆらぎであるかを観測できれば、それはタンパク質の個性をあらわす。例えば、中性子非弾性散乱からは状態密度と対応する動的構造因子を測定することができる。動的構造因子の解像度は、以前はスペクトルの大まかな形を議論できる程度であったが、最近ではピークの微細構造を議論できるまでに解像度が上がってきた¹⁾。このようにタンパク質のダイナミクスを個性まで知るためには、高解像度で測定をおこなう必要がある。

タンパク質のダイナミクスは 10^{-15} 秒から1秒程度までの広い時間領域で起こる。分子シミュレーションは、原子レベルの空間解像度をもったモデルを用いると、通常 10^{-15} 秒から 10^{-9} 秒程度までの時間領域を観測できるという特徴がある。最近では 10^{-6} 秒までのシミュレーションがおこなえるようになってきたが²⁾、これには膨大な計算量が必要となる。分子シミュレーションのもう一つの特徴は、1分子だけを観測すること

ができることである。通常の実験手段での観測量はアンサンブル平均であるが、分子シミュレーションではアンサンブル平均ではわからない個々のイベントを直接観ることができる。分子シミュレーションを用いるとタンパク質の個性と一般的な性質の両方を研究することができるのである。

3. タンパク質のダイナミクスをどう観るか?

— 集団座標の効用 —

タンパク質のダイナミクスは、空間的にも時間的にも広い領域で起こる。しかも、タンパク質は数千から数万の内部自由度 (原子数を N として $(3N-6)$) をもつ。タンパク質のダイナミクスはこれらの自由度で張られる多次元空間の中で発展する。分子シミュレーションをおこなうと、自由度がどのように変化していくかをシミュレーションのステップにそって記録することになる。この膨大な情報をどのように認識すればいいのか、タンパク質のダイナミクスをどのように観るかという問題には工夫を要する。つまり、多数の自由度の中から、注目すべき重要な自由度を抽出する必要があるのだ。そのためにはゆらぎの周期性を利用して特定の周波数領域のモードのみを抽出する方法³⁾と、ゆらぎの大きさに注目して集団座標を用いる方法がある⁴⁾。前者の方法は、システムのゆらぎが周期的な場合には有効であるが、タンパク質の場合、遅いゆらぎは確率過程であるので適用できるとは限らない。後者の方法は、ゆらぎの時間依存性によらずに用いることができるので、タンパク質のダイナミクスを観察するのに非常に有効である。集団座標を用いたタンパク質ダイナミクスの解析法に関しては総説^{5), 6)}を参照していただきたい。

集団座標を決定するには主成分分析を用いる。これは統計処理の一般的手法であるが、ゆらぎに対して用いると物理的な意味を生じる。主成分分析で決定される集団座標 (主成分座標) は、ゆらぎの大きさを基準に決定される。第1番目の集団座標軸 (第1主成分軸) はゆらぎが最大となるように定義される。すなわちタンパク質が最もゆらぐ方向を表す座標軸である (後述のタンパク質CD2の第1主成分のゆらぎを図2に示した)。第2主成分軸は、ゆらぎが2番目に大きくなる方向として定義される。集団座標は、原子座標の線形結合で定義され、互いに正規直交系を成す。集団座標の導入はゆらぎを観る座標の取り方を変えているだけで、なんら情報は失われない。

集団座標の導入がなぜ有効なのか? それは、タンパク質のダイナミクスの特徴による。タンパク質のゆ

図2 タンパク質CD2の構造とそのゆらぎを特徴的に表す集団座標（第1主成分座標）
集団座標における各原子のゆらぎの向きと大きさをベクトル（各原子からのびる線）の方向と長さで示している。

らぎは集団座標で表してみると異方性が高く、その振幅には最大 10^5 程度の大きさの違いがある。だから、ゆらぎの大きさを基準に集団座標を決定する主成分分析が有効なのである。重要な少数の主成分軸（主成分の数で1%未満）で張られる部分空間内のゆらぎは、全体のゆらぎに対して80%以上の寄与をもつ。そうした少数の主成分であられてくる、振幅の大きな座標軸方向でのゆらぎがもし機能的に重要であれば、それら少数の自由度だけに注目すればよいことになる。これを示すにはさらにダイナミックドメイン^{7,8,9}という集団座標より単純化した概念を導入することが有用である。

従来のドメインは、1つの立体構造から定義される。すなわち、静的な構造から定義されたドメインである。ダイナミックドメインは、タンパク質が大きな構造変化を起こした際にも構造がほとんど変化しないかたいユニットとして定義される。例えば、リガンドと結合した構造と単体の構造の2つがX線結晶解析によって決定されている場合にも定義できるし、分子動力学計算から主成分分析をおこなった場合にも定義することができる。リガンドとの結合によって生じる結晶構造変化と分子シミュレーションからそれぞれ独立にダイナミックドメインを決定しそれらの対応を調べることで、シミュレーションで観ているダイナミクスが機能

的に重要かどうかをチェックできるのである。

4. “JAM モデル” によるダイナミクスの描像

前節の方法から明らかになったタンパク質に一般的な性質はどのようなものだろうか。それはJAM (Jumping-Among-Minima)¹⁰ という概念に凝縮されている。タンパク質の天然状態にはエネルギー的に準安定な状態が多数ある。タンパク質はほとんどの時間あるエネルギー極小点の近傍にエネルギー的にトラップされている。ただし、ここで言うエネルギー極小点の近傍とは、周りのエネルギー障壁を越えずにタンパク質がゆらぐことのできる範囲であって、それぞれがひとつの準安定な状態に対応する。しかし、まれな現象として大きな熱ゆらぎが起こり、周りのエネルギー障壁を越えて他のエネルギー極小点の近傍へと構造がジャンプする。これは多くの物質にも当てはまる一般性の高い概念である。それでは実際のタンパク質ではどのような特徴があるのか。まず重要な第1番目の点は、極小点近傍のエネルギー曲面の形状はどのエネルギー極小点近傍でもほとんど変わらないことである。これは図1での議論と関係する。さらに極小点近傍のエネルギー曲面の形状はパラボラ形であって非調和性はほとんどない。したがって、極小点近傍のゆらぎは調和的なゆらぎである（周期は 10^{-14} から 10^{-12} 秒程度）。第2の点は、非調和的なエネルギー極小点間のジャンプ（JAM）は少数の自由度で張られる部分空間内で起こることである。これらのジャンプは確率的に発生するまれな現象である。

表1にJAMの描像によるタンパク質のゆらぎの特徴を示した。ほとんどの自由度は、ジャンプと関係しない、調和的な（すなわちパラボラ形の）エネルギー面をもっている。ヒトリゾチームの場合、調和モードの数は全体の自由度の95%を占めるが、この分布は図1ですでに示したように、タンパク質の種類や構造に依存しない。非調和な自由度（多重階層モードと単階層モード）はごく少数であり、エネルギー極小間のジャンプに関わるモード（JAMモード）である。ゆらぎが遅くなるほど、また、ゆらぎが大きくなるほど、それにかかわる自由度は少なくなる。単階層モードのエネルギー面の形状は大まかにはパラボラ様であるが微細な極小点が多数あるので調和モードのような振動的な性質は失われ、ダイナミクスは確率過程となる。多重階層モードはドメイン運動などの非局所的な大きなゆらぎで、リガンドとの結合に伴う構造変化など、機能と密接に関係するモードである（ただし化学反応の過程では局所的なゆらぎが重要な場合もある）。

表1 タンパク質のゆらぎ

	多重階層モード	単階層モード	調和モード
自由エネルギー面
どのようなゆらぎか?	ドメイン運動などの大きなゆらぎ・機能と関係する	局所的なゆらぎ	局所的なゆらぎ
原子パッキングの変化	大	なし	なし
二面角の変化	主鎖および側鎖に大きな変化	主に表面の側鎖に変化	微小
ゆらぎの時間領域	10^{-13} 秒以上 (いろいろな時間スケールの運動が共存)	$10^{-13} \sim 10^{-11}$ 秒	$10^{-14} \sim 10^{-12}$ 秒
ヒトリゾチームの場合 ¹⁰⁾			
全体に占めるモード数の割合	0.5%	4.5%	95%
ゆらぎへの寄与率	82%	15%	3%
ゆらぎの振幅	0.08 Å以上	0.08~0.015 Å	0.015 Å以下

ここで示したJAMモデルによるダイナミクスの描像は、タンパク質に共通する物性を示したものである。タンパク質の個性は、例えば機能に関する多重階層モードが具体的にどのようなゆらぎであるかにあらわれてくる (図2)。

5. 実験とシミュレーションのハイブリッド

さて、タンパク質に共通する物性を知る事は、タンパク質の個性、例えば構造や機能を研究するのにどのように役立つのだろうか? タンパク質のダイナミクスは複雑で、ひとつの実験手段ですべてを明らかにすることはできない。しかし、複数の手段から得られるデータを総合的に解釈すれば、より広い時空でのタンパク質のふるまいが明らかになる。タンパク質の比較的スローなダイナミクスを観測するためには、いろいろなデータをどのように接続すればいいのか? そのためにJAMの描像が役立つ。そのヒントは以下の点である。(1) タンパク質は多数の調和的な自由度とごく少数の非調和的な自由度をもつ。(2) エネルギー極小点近傍のゆらぎは比較的速いゆらぎである。(3) エネルギー極小点近傍のゆらぎは調和的でどの極小点にいるかにほとんど依存しない。(4) 非調和なゆらぎは確率的な運動である。(5) 分子シミュレーションを用いると、比較的速いゆらぎを高解像度で観測できる。(6) 分子シミュレーションより遅い時間領域を観測できる実験手段は存在するが、複雑なダイナミクスを反映しているので解釈がむずかしい。

これらの情報から、分子シミュレーションからはエ

ネルギー極小点近傍の速い動きを決定し、非調和なゆらぎを実験から決定するというアイデアが生まれてくる。非調和なゆらぎを記述するには少数の自由度でよいのだから、実験から決定すべきパラメーターの数は少数ですむ。これがJAMモデルに基づくハイブリッド法の特徴である。

それでは、実験と理論をどのように接続すればいいのか? 実験からある物理量の時間相関関数 $C(t)$ が求まったとしよう。分子シミュレーションからはエネルギー極小点近傍での時間相関関数 $C^{min}(t)$ を決めておく。ヒント (3) の知見と、外部運動からの寄与 $C^{ext}(t)$ ・調和的なゆらぎ・非調和なゆらぎがそれぞれ独立であれば、

$$C(t) = C^{ext}(t)C^{JAM}(t)C^{min}(t) \quad (1)$$

となる。外部運動が単純なモデルで与えられれば、 $C^{JAM}(t)$ を実験と理論のハイブリッドから決めることができるのである。 $C^{JAM}(t)$ は機能的に重要であると期待される比較的ゆっくりとしたタンパク質のダイナミクスの情報を与えてくれる。このアイデアはいろいろな実験データとのハイブリッドを可能とする。中性子非弾性散乱の場合は、インコヒーレントな動的構造因子 $S_{inc}(Q, \omega)$ が観測される (Q は波数, ω は振動数)。これは外部運動からの寄与 $S_{inc}^{ext}(Q, \omega)$ も考慮すると、

$$S_{inc}(Q, \omega) = S_{inc}^{ext}(Q, \omega) \otimes S_{inc}^{JAM}(Q, \omega) \otimes S_{inc}^{min}(Q, \omega) \quad (2)$$

となり、 $S_{inc}^{ext}(Q, \omega)$ を決めることが可能となる (⊗は振

実験と理論のハイブリッドから機能に関わるタンパク質のダイナミクスを観る

図3 CD2/CD58 複合体形成の概念図

動数に関する畳み込み積分)⁶⁾。

実験から時間依存性ではなく物理量 A のアンサンブル平均 $\langle A \rangle$ が得られる場合には、JAMモデルの描像から、アンサンブル平均を

$$\langle A \rangle = \sum_k f_k \langle A \rangle_k \quad (3)$$

のように、各エネルギー極小点 k の近傍での平均 $\langle A \rangle_k$ をさらにその統計重率で平均したもので置き換えられる¹¹⁾。ヒント(2)を考慮すれば、平均 $\langle A \rangle_k$ は分子シミュレーションから十分計算することができる。あとは重率 f_k を実験データ $\langle A \rangle$ を再現するように決定すればよい。決めるべきパラメーターの数はエネルギー極小点の数であるので、実験データがそれより十分にあればこの方法が使える。この方法はさまざまな実験データの解析に使えるが、X線結晶解析やNMRなどが有望である。

6. CD2に見るダイナミクスと機能

それでは実際にこの方法がどのように使えるかをみてみよう。NMRの緩和データからは、ピコ(10^{-12})秒からミリ(10^{-3})秒程度までのダイナミクスの情報が得られる。その情報と分子シミュレーションを組み合わせ、タンパク質のダイナミクスを詳細にみてみよう。

CD2はT細胞やナチュラルキラー細胞の表面に存在し、CD58と結合する。この過程はMHCによって提示されたペプチドがT細胞レセプター(TCR)に認識されるのに先立っており、CD2とCD58の結合が阻害されると免疫反応そのものが阻害されることが知られている。我々はすでに発表されているNMRの緩和時間のデータ¹²⁾と分子動力学を用いてCD2のダイナミクスと機能のかかわりを研究した¹¹⁾。

これによってCD2がどのような構造変化をするかが明らかになった。CD2はおもに2つの準安定な状態をもつ。それぞれの状態の統計重率 f_k は0.42と0.24であるから、CD2の構造は42%および24%の確率でそれぞれの状態をとることがわかった。これ以外にもより重率の低い状態が存在する。このように現実のタンパク質も、シミュレーションで見られるように少数の準安定な状態間をジャンプしているのである。CD2がCD58と結合するには結合面がより露出された形に構造変化する必要がある。CD2が単独で存在する時に起こっている準安定な状態間のジャンプは実はこの構造変化と深く関係している。CD2は単独では結合面が比較的閉じた形で存在しやすい。しかし、ナノ(10^{-9})秒程度の時間スケールで確率的なジャンプがおり、しばしば結合面がより露出された形に構造変化してい

る。CD2はCD58の存在しない状態でも、CD58との結合モードをもっているのである。CD58との結合はこれよりゆっくりとした動きでミリ秒程度の時間スケールでおこると見られている。つまり、CD2とCD58の結合は、頻繁に出現する結合面の露出した構造をとらえて、よりゆっくりとした時間スケールでおこると考えられる(図3)。

7. 今後の可能性

ここではNMRの緩和データと分子動力学との組み合わせを示したが、NOE強度を用いることもできる。また、NMR以外にもさまざまな実験と組み合わせることが可能である。今後、タンパク質のダイナミクスは多面的に研究されていくだろう。さまざまなタンパク質のダイナミクスが調べられ、機能のメカニズムが解明されることが期待される。

JAMモデルは、現The University of East AngliaのSteven Hayward博士・京都大学の郷 信広教授との共同研究であり、NMRと分子シミュレーションのハイブリッドは、執筆者が1998年4月から99年3月までに文部省在外研究員として滞在したHarvard Medical SchoolのGerhard Wagner教授との共同研究である。

文 献

- 1) Lamy, A. V., Smith, J. C., Yunoki, J., Parker, S. F. and Kataoka, M. (1997) *J. Am. Chem. Soc.* **119**, 9268-9273.
- 2) Duan, Y. and Kollman, P. A. (1998) *Science* **282**, 740-744.
- 3) Sessions, R. B., Dauber-Osguthorpe, P. and Osguthorpe, D. J. (1988) *J. Mol. Biol.* **209**, 617-633.
- 4) Kitao, A., Hirata, F. and Go, N. (1991) *Chem. Phys.* **158**, 447-472.
- 5) Hayward, S. and Go, N. (1995) *Annu. Rev. Phys. Chem.* **46**, 223-250.
- 6) Kitao, A. and Go, N. (1999) *Curr. Opin. Struc. Biol.* **9**, 164-169.
- 7) Wriggers, W. and Schulten, K. (1997) *Proteins* **29**, 1-14.
- 8) Hayward, S., Kitao, A. and Berendsen, H. J. C. (1997) *Proteins* **27**, 425-437.
- 9) Hayward, S. and Berendsen, H. J. C. (1998) *Proteins* **30**, 144-154.
- 10) Kitao, A., Hayward, S. and Go, N. (1998) *Proteins* **33**, 496-517.
- 11) Kitao, A. and Wagner, G. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 2064-2068.
- 12) Wyss, D. F., Dayie, K. T. and Wagner, G. (1997) *Protein Science* **6**, 534-542.