

## タンパク質の構造空間：ファネルとエネルギーランドスケープ

神戸大学理学部 高田彰二

### イントロダクション：

タンパク質を物理学の対象として捉えるときに、主に2つの観点が考えられる。一つは、35億年の生物進化の所産としてのタンパク質という観点であり、ここでは例えば、現存するタンパク質のアミノ酸配列をいかに特徴づけるかに焦点がある。もう一つは、タンパク質を、与えられたアミノ酸配列をもつ高分子と捉え、その高分子の立体構造形成や物性を議論する観点である。しかし実際には、両者を切り離して考えることは不可能である。生物は、そのタンパク質がおのおの特異な立体構造を形作り、そこで柔らかく動けることができるように進化してきたのだから、進化的観点には物性的観点が不可欠である。また、(現存する)タンパク質の物性は、ランダム配列をもつポリペプチドとは著しく異なるので、その配列の特殊性を抜きにしてはタンパク質のタンパク質らしい物性を議論することは出来ない。配列空間(進化的観点)と構造空間(物性的観点)との相互関係を徹底的に議論しよう、というのがこの基研セミナーの主旨であった。その中で私の役目は、後者すなわち、タンパク質高分子のもつ物性的観点から、構造空間におけるタンパク質のおりたたみ(立体構造形成)研究についての現状を概観することである。

大きな自由度を持つ柔らかいシステムは、今の物理学研究の進む方向を示す旗印とも言える。その中の研究対象として、タンパク質分子は格好の好条件を兼ね備えている。一つには、圧倒的な量のデータベースである。ゲノムプロジェクトの成果により、アミノ酸配列情報は爆発的に増加しているし、また最近の技術革新により多くのタンパク質立体構造が決定され、構造データ数の方も1万件を超えた。また、分子生物学やタンパク質工学のパワーは目覚ましいものがある。まったく均一なタンパク質を比較的大量に作ることも、また配列の任意の位置を任意のアミノ酸に置換することも、実験室のなかでタンパク質を進化させることも、などなど思いのままである。また、大きな自由度系ではあるが、現在のコンピュータである程度扱えるサイズである点も重要である。問題によっては、比較的経験性の低いコンピュータシミュレーションで定量的に正しい答えが得られる。このような事情もあって、90年代にタンパク質折りたたみ研究に多くの物理学者が参入した。そうして、統計物理的解析から、折りたたみのエネルギーランドスケープ理論[1]とか、ファネル理論とか呼ばれるものが出来あがった。これら物理学者による見方は、new viewと呼ばれ、それまでのold viewと対比的に見られることも多いが、両者は矛盾するものではなく、違う観点(view)を強調していると考えた方が適切である。次節以降ではこのnew viewを概説する。

エネルギーランドスケープ理論のファネル的側面とガラス的側面：

ファネル的側面：

ファネルのコンセプトは、約 20 年前に郷によって提唱された consistency principle に基づいている。それは、「タンパク質は、進化の所産として、ほとんど全ての相互作用がネイティブ構造で安定になるように出来ている」という洞察である。ネイティブ構造でタンパク質鎖の有効エネルギー  $E$  が最小値をとり、コンフォメーションがそこから離れるほど  $E$  が大きくなる。これをマンガ的に描けば、図 1 のようになる。図の横軸は、タンパク質のコンフォメーションをあらわす  $r$  で、本来超多次元であるがここでは簡単に 1 次元で描いてある。縦軸は  $E$  である。変性状態は上方の  $E$  の大きなあたりに対応し、コンフォメーションとしては広い。一方ネイティブ状態は底、 $E$  が最小値のところである。全体としてエネルギー面はファネル（漏斗）状の形をとり、フォールディングは、図の上方から底への遷移と捉えることができる。さて、ある  $E$  のところでのファネルの横幅は、コンフォメーション空間の広さを表し、その対数は鎖エントロピー  $S$  である。

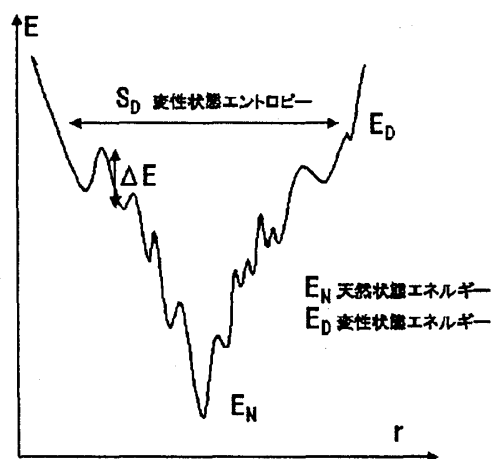


図 1 蛋白質のエネルギーランドスケープの模式図

このように、ファネル理論は、「フォールディング転移は、 $E$  の大きな減少（安定化）と、 $S$  の大きな減少（不安定化）によって特徴づけられる。」ということを強調する。変性温度  $T_f$  では、

$$\Delta F = \Delta E - T_f \Delta S = 0$$

となり、エネルギーの安定化とエントロピーの不安定化がちょうどキャンセルする（ $\Delta$  は天然状態と変性状態との差という意味）。生理条件はそれよりやや低温であるから、大きなエネルギーの安定化がエントロピーによる不安定化と大部分キャンセルしながら、トータルでわずかにネイティブ構造の方が安定である。ここで、 $E$ （または  $S$ ）が

反応座標の役割をしていることが分かる。

ネイティブ構造は、ある程度のゆらぎを含めた意味で、一つの構造であるから  $S(N) \sim 0$  と考えて良い。ファネルを上昇すると、ネイティブ以外では、 $S > 0$  となるから、本質的に多数のコンフォメーションの集団ということになる。重要な帰結として、フォールディング経路は、ある特定のタンパク質構造を経て起こるといようなものでは決してありえない。フォールディング経路は、統計的な分布の変移経路としてのみ意味を持つ。逆に、統計的な経路を考える限りそれは、ファネル的な見方と矛盾するものではない。

図1に描かれた斜面のでこぼこには後で触れることにして、でこぼこを取り去ってもう少し定量的な議論を行う。思考実験的に、タンパク質の鎖エントロピー  $S(E)$  を計算し(図2左)、それを横軸にエントロピー、縦軸にエネルギーをとってプロットすると漏斗の片側だけきりとしたものが出来る。それを、図1に似せるために図形を左右対称になるように変形したのが、図2中である。同じ  $S(E)$  を使って、自由エネルギー  $F(E) = E - TS$  をプロットしたのが、図2右である。これはよく見慣れた (old view 的な) プロットであり、反応座標 ( $E$ ) を横軸にとって、変性状態とネイティブ状態の両極小点が、自由エネルギー障壁によって隔てられている。

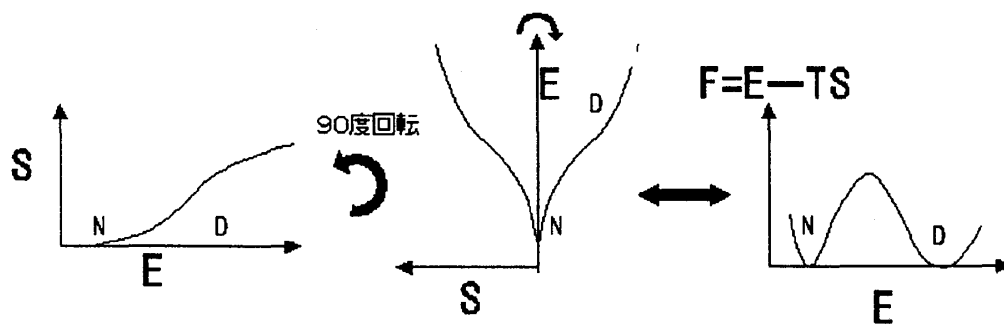


図2 ファネル (真中) と自由エネルギー曲線 (右)

自由エネルギー障壁が現れたことが重要である。その原因を考えてみると、図2a)でファネルがわずかに曲線を描いていることに行着く。すなわち、変性状態からエネルギーとエントロピーが減少する際、エントロピーの方がやや先に減少していることによる。その差が最大になったところが、自由エネルギーで見ると遷移状態をつくる。従って折りたたみ転移は、有限系の (一次) 相転移と捉えられる。フォールディング反応の律速は、エネルギー的な安定化を駆動力にしながらエントロピーを減らす過程である。

以上の議論から、ファネルを量的に特徴付けるものとして二つ重要なものがあることに気づく。一つはグローバルなファネルの斜度であり、もう一つはカーブしている斜面の曲がり具合である。前者は、ファネルの高さが $\Delta E$ で横幅が $\Delta S$ であることから、折りたたみの相転移温度 $T_F$ に等しいことがわかる。生物進化の所産として急勾配のファネル構造を作り上げたタンパク質はある程度高い相転移温度 $T_F$ （高い熱安定性と置き換えてもよい）を持つことが出来ることを示している。一方、斜面の曲がり具合は自由エネルギー障壁の高さに対応し、これがアーレニウスの意味での折りたたみ速度を決定する因子である。

折りたたみの速度定数を決めている要因に関して、最近非常に重要な発見があった[2]。Bakerらは、比較的小さなタンパク質の中で折りたたみ転移が二状態的に起こるものについて、実験による折りたたみ速度データをリストアップし、その特徴を調べた。その結果、タンパク質の天然状態の構造トポロジーを反映した非常に簡単な指数、コンタクトオーダー、と折りたたみ速度とにかなり強い（負の）相関があることを見出した。彼らが定義したコンタクトオーダーとは、天然構造において側鎖の接触しているアミノ酸対について、「対の高分子鎖に沿った距離」の（全対にわたる）平均が、タンパク質の全アミノ酸数に比べて何パーセントか、で定義される。例えば、 $\alpha$ ヘリックスだけで出来ているタンパク質では、（高分子鎖にそって）局所的な接触が多いのでコンタクトオーダーは小さくなる。既に指摘したように、コンタクトオーダーは天然構造のトポロジーの複雑さを表す一つの指標であり、Bakerらの発見はその値が大きいほど（複雑なトポロジーであるほど）折りたたみ速度が遅くなることを示している。この発見と前後する時期に、折りたたみの反応経路が、詳細なアミノ酸にそれほどよらず、折りたたみ最終構造（天然構造）のトポロジーでおおむね決まっているらしいというデータが実験、理論両方面から集まってきた[3]。これらを総合して、折りたたみ機構のトポロジー支配仮説とでも呼べるようなコンセプトが出来あがった。

これを、上に述べたファネルのモデルで解釈してみよう。たとえば $\beta$ シート（ $\alpha$ ヘリックス）主体のタンパク質では、折りたたみ初期に高分子鎖にそって遠い（近い）アミノ酸対の接触が必要であるとする。その対の接触によるエントロピーの減少は高分子鎖にそった距離の対数程度であるのに対し、接触によって得られるエネルギー安定化は高分子鎖にそった距離によらず一定である。これは、 $\beta$ シート主体のタンパク質のファネル斜面が、 $\alpha$ ヘリックスのものより強く曲がることを示唆している。その結果、 $\beta$ タンパク質の折りたたみは遅くなる。つまり、タンパク質のトポロジーがファネル斜面の曲がり具合を支配し、それによって折りたたみ反応機構が決まってくる。

ガラス的側面：

ここで、先送りしていたファネルの斜面に描かれたでこぼこの話に戻る。これは、エネルギー面に多谷構造があることを表している。実際タンパク質分子のコンピュータシミュレーションをしてみると、エネルギー面が如何にでこぼこしているか、を嫌というほど思い知らされる。

直感的に明らかなように、エネルギー面のでこぼこさは、その系のガラスらしさと関係している。実際簡単な理論モデルによるとガラス転移温度  $T_G$  は

$$T_G = \Delta E / \sqrt{2S_0 k_B}$$

のように、エネルギー面のでこぼこさの尺度  $\Delta E$  (図 1 参照、この  $\Delta$  は天然状態と変性状態の差と関係ない) に比例している (ここで  $S_0$  は全状態数の対数)。

タンパク質の折りたたみ可能性は、熱安定性と速度論と両方の要請を満たすことで決まってくる。天然状態が安定であるためには  $T < T_F$  であり、ガラス化して準安定状態にトラップされないためには  $T > T_G$  でなければならない。合わせると、 $T_F > T > T_G$  の温度範囲で折りたたみが可能であることになる。このような温度範囲が十分に広く取れること、すなわち  $T_F/T_G$  が 1 にくらべて十分に大きいことがタンパク質に課せられた要請であり、進化によって獲得されたアミノ酸配列はこの要請を満たしているはずである。これが、エネルギーランドスケープ理論の最重要ポイントである。これは、ファネルのマンガで言うと、ファネルの高さに比べてファネル斜面のでこぼこさが小さいことに対応する。

折りたたみ速度  $k_F$  は、Kramers の理論で

$$k_F \sim D \exp(-\Delta F / k_B T)$$

と書ける。ここで、 $\Delta F$  は反応座標 (オーダーパラメータ) に沿った自由エネルギー障壁の高さであり、また  $D$  は反応座標方向の拡散係数である。後者は、局所安定な構造からの脱出の速度であり、それは自由エネルギー面のでこぼこさに依存する。合わせると、折りたたみの速度は、反応座標方向 (縦方向) のエネルギー障壁と、拡散する際に超えなければならない (横方向とでもいえる) エネルギー障壁 (の平均) との大きいほうで決まってくる。前者が効いている場合、その振舞いは指数関数的、後者が効いている場合は、障壁の高さは非一様であり、べき的または stretched exponential になるはずである。通常の実験ではほとんどいつも折りたたみの起こり方は指数関数的であることが知られているが、最近 Gruebele らによって、縦方向のエネルギー障壁をゼロに (あるいは非常に小さく) することによって、非指数関数的振舞いが観測さ

れた[4]。この実験はエネルギーランドスケープのでこぼこさを実験で観測しようとする試みであり、今後の発展が興味深い。

#### タンパク質の立体構造予測：

最後に、タンパク質構造空間のエネルギーランドスケープ上における最適化問題、すなわち立体構造予測問題、の可能性と現状に触れてみたい。アミノ酸配列の情報だけからタンパク質の立体構造を予測することは、ほぼ50年にもわたるこの分野の夢である。ファネルのコンセプトは、構造予測問題がいわゆるNP完全な問題ではない、ことを主張している。また折りたたみ機構のトポロジー支配仮説の成功は、折りたたみ問題が比較的簡単な物理で決まっていることを暗示している。これらは、立体構造予測問題には朗報であり今、比較的簡単なモデルでも荒削りな予測が出来るのではないか、と多くの予測屋は考えている。

そういう機運の中、1994年以来2年ごとに、タンパク質の構造予測コンテストが行われている[5]。まだ構造の知られていないタンパク質のアミノ酸配列数十本がインターネット上で出題されて、予測屋がその構造を予測して投稿する。構造が実験的に決定されて後、各予測の成否を点数化して客観的な評価が下る。このコンテストへの参加者は回を追うごとに増加しており、非常に活況である。完全な解決にはまだ少しの年数を要するであろうが、予測の成績は全体的に上がってきているように見える。世紀を跨いだ大問題を解いてみようというあなた、参加してみませんか。

[1]J.D. Bryngelson, J.N. Onuchic, N.D. Socci, and P.G. Wolynes, *Proteins Struct. Funct. Genet.* 21 167 (1995).

[2]K.W. Plaxco, K.T. Simons, and D. Baker, *J. Mol. Biol.* 277, 985 (1998).

[3]J.J. Portman, S. Takada, and P.G. Wolynes, *Phys. Rev. Lett.* 81, 5237 (1998).

[4]J. Sabelko, J. Ervin, and M. Gruebele, *Proc. Natl. Acad. Sci. USA* 96 6031 (1999).

[5]Special issue on CASP3, *Proteins, Struct. Funct. Genet. Suppl* 3 (1999).