From SK model to Protein Folding

東大教養 時田恵一郎 (tokita@complex.c.u-tokyo.ac.jp)

タンパク質の立体構造は遺伝子の1次元配列にコードされていて、生合成されるアミノ酸の配列を知ることで立体構造ひいてはそのタンパク質の機能が予測可能であると一般に信じられている。最近の「タンパク質の折れたたみの統計力学[2, 1, 3, 4, 5]」は、種類の限られたアミノ酸を単純な素子群とみなし、それらが大域的に結合した系の協同現象として「折れたたみ」のメカニズムを知ろうというものである。しかしこの立場の研究は、立体構造の究極的予測法を提供するということとはスタンスを異にするように感じられるし、「複雑系」という観点から言えば、むしろ我々は、生物が進化の過程で獲得したエレガントな情報処理方式の本質を知りたいという欲求を持っている。タンパク質は、折れたたんでいないランダムコイルの状態から必ず生理的活性のある構造へと折れたたむというわけではなくて、生理的活性という機能の無い様々な異なる立体構造を取りうる。実験[6, 7]からも示唆されるこのタンパク質の自由エネルギーの多谷構造(ragged landscape)は、スピングラスの平均場理論で築き上げられてきた様々な知見[8]との対応が議論されることが多い[9]。しかし、現在までの「タンパク質≈スピングラス」という見方から(少しだけ?)脱却して、「ランドスケープ」という観点から折れたたみのメカニズムの本質を探りたいというのが筆者の目的の一つ[10]である。

さて、タンパク質のエネルギー構造は複雑な谷をなしているという直観があるけれども、1個ないしは非常に少数の折れたたみ状態だけが生理的活性を持ち得るという状況は、膨大な数の(準)安定状態を持つスピングラス描像とは相容れない。そこで、知られているエネルギーランドスケープを与えるということを目的としたおもちゃモデル(Translational Invariant Sherrington Kirkpatrick model : TISK model)を提案したい。TISKモデルはDNA塩基配列に観測された(準)周期性[11]をデフォルメして、それが残基間の相互作用にも並進対称性として反映しているという大胆な仮定を基礎にして以下のようなハミルトニアンを定義する。

$$H = \frac{1}{N}\sum_{i,j} J_{|i-j|} S_i S_j \tag{1}$$

ただし、$J_{|i-j|}$は1次元格子上に並ぶi,j番目のアミノ酸間の距離のみに依存するGaussランダム変数で、S_iは折れたたみの角度をsymbolicに表現するIsing変数である。ただし、$i,j = 1,2,\cdots,N$(system size)。Ikegami and Kaneko[12]は彼らの提案する「Genetic Fusion(GF)」において上記のようなエネルギー(コスト)関数曲面上ではGFの最適化プロセスが有効であることを示した。そこで示唆されたのは、ハミルトニアン(1)は少数の深い谷を持つ基底状態と多数の準安定状態を与えるエネルギーランドスケープを与えるということである。これはSKモデルとは異なるエネルギーランドスケープの描像を与えていて、実験的にも予想されているゴルフホール状のエネルギー構造により近い。

さらに上記モデルのランダムボンドは以下のようにWalsh関数の相関行列で表現できる。

$$J_{|i-j|} = \frac{1}{N^2}\sum_{n=0}^{N-1} a(n) \sum_{k=0}^{N-1} \phi^{(n)}_{(i+k)\%N}\,\phi^{(n)}_{(j+k)\%N} \tag{2}$$

ここで、$\phi_i^{(n)}$はn番目Walsh基底のi成分、$a(n)$は各Walsh基底の重みで、$x\%y \equiv x \bmod y$である。Walsh関数のいくつかの性質からJの集合が実際に与えられると$a(n)$をすべて知ることができる。この時Walsh関数及びそれらの並進対称な関数群全てを学習パターンと考えた場合のHopfield Modelと等価になり、どのWalsh基底がどの程度の重みでエネルギーに寄与して深い谷を持つかを直接知ることが出来る。さらに$N\to\infty$の極限ではJ_{ij}の分布から$a(n)$の分布を知ることができる。計算の詳細は省略するが、$a(n)$の分布はやはりガウス分布になることがわかる。(2)式は並進対称なWalsh関数群全体の重ね合わせであるので、直観的には、

$n = 2^m (m = 0,1,2\cdots)$ 番目の Walsh 関数に対応する状態が低いエネルギーを与えること、および基底状態はそれらの重ね合わせで表されることが予想される。

図 1 は $N = 64$ の有限系に対して行ったモンテカルロ・シミュレーションの結果である。図 1(a) はランダムな初期状態 1000 個からの温度 0 の非同期ダイナミクスで得られた安定状態のベーシンの大きさである (横軸エネルギー、縦軸頻度)。これから (少なくとも見つかった) 最低エネルギーを与える状態は比較的大きな basin をもっていることがわかる。一方図 1(b) は縮退の数を示す (横軸エネルギー、縦軸縮退数)。これによると最低エネルギーを与える状態は縮退していないが、比較的エネルギーの高い (準) 安定状態は多くの状態が縮退していることを示している。これは既に述べた通り、多数の準安定状態とごく少数の基底状態という実験からも与えられる描像をサポートするものである。図 1(c) は見つかった最低エネルギーを与える状態の Walsh パワースペクトル (横軸 Walsh 基底の番号、縦軸パワー) で、4 個の Walsh 関数の重ね合わせになっていることがわかる。これも Walsh 関数を用いた解析からの予想を裏付けるものである。なお、同じ J_{ij} のサンプルに対する GF での基底状態探索はモンテカルロでみつかったものと同じ状態を見つけだしている。

最後に本研究は神戸大池上高志氏、東大教養金子邦彦氏との協同研究であり、特に GF での基底状態探索は神戸大において池上氏のシミュレーションプログラムを使用して行いました。両氏に感謝致します。

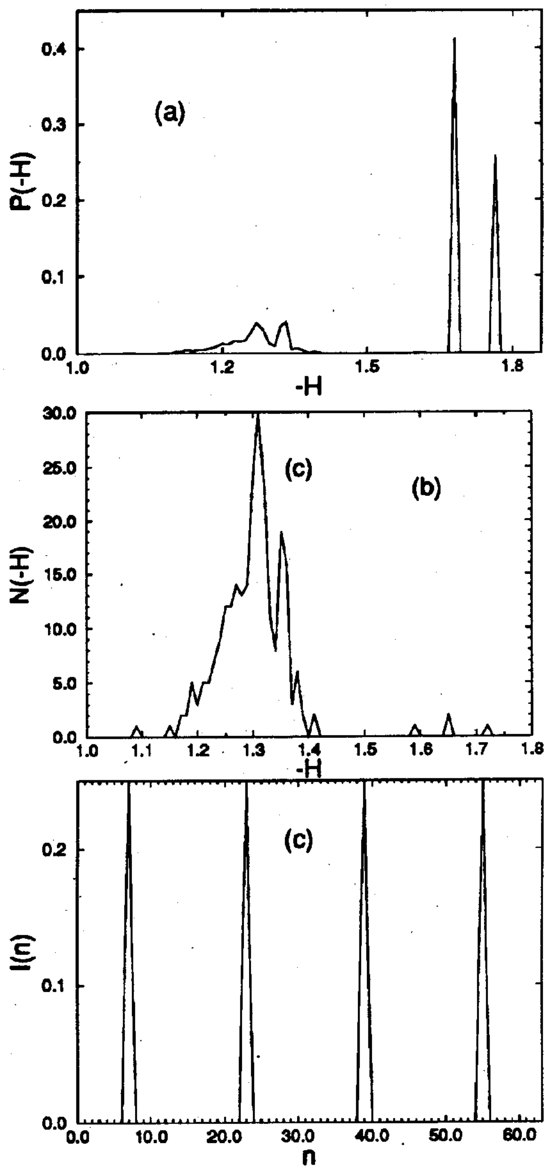

図 1:

参考文献

[1] Garel and Orland, 1988, *Europhys. Lett.*, **6**, 597

[2] Shaknovich and Gutin, 1990, *Nature*, **346**, 773

[3] Sasai and Wolynes, 1990, *PRL*, **65**, 2740

[4] Iori, Marinari and Parisi, 1991, *J. Phys. A*, **24**, 5349

[5] Stillinger, Head-Gordon and Hirshifeld, 1993, *Phys. Rev. E*, **48**, 1469

[6] Ansari, et al, 1985, *Proc. Natl. .Acad. Sci. USA*, **82**, 5000

[7] Iben, et al, 1989, *PRL*,**62**, 1916

[8] Mézard, Parisi and Virasoro, 1987 *Spin Glass Theory and Beyond* (World Scientific)

[9] ed. by D. L. Stein, 1992, *Spin Glasses and Biology* (World Scientific)

[10] 一見単純に思われるモデル化が、より広いクラスの概念を生み出すことがある。例えば、SK モデルから始まった純状態分布の超計量性に関する議論もその一つであって、そのような「創発的概念」の出現を期待するというのが密かな目的でもある。

[11] Li and Kaneko, 1992, *Europhys. Lett.*, **17**, 655

[12] Ikegami and Kaneko, 1990, *PRL*, **65**, 3352