

第9章

統計力学モデルによる Φ 値計算と フォールディング・メカニズム

<要旨>

Fersht らが提案した、実験による“ Φ 値解析”という手法は、どのようにして、フォールディング過程の遷移状態での構造を推定することができるか示している。

更に、統計力学的モデルである「A-W_NILSモデル」を用いて、“実際のタンパク質”の各アミノ酸残基の Φ 値を理論的に計算する方法を示し、具体的に、2つのタンパク質、プロテインA（1SS1）とキモトリプシンインヒビター2（3CI2）の各アミノ酸残基の Φ 値を求め、実験による Φ 値と比較・検討している。

1SS1 のフォールディング・メカニズムは、実験による Φ 値の分布から“核形成-凝縮モデル”であると推論しているが、「A-W_NILSモデル」による Φ 値の分布からは“フレームワークモデル”が妥当であると推論している。

また、キモトリプシンインヒビター2（3CI2）の実験による Φ 値の分布は、“核形成-凝縮モデル”の妥当性を示している。このタンパク質に対する「A-W_NILSモデル」の限界を述べている。



実験によるプロテイン A の遷移状態の Φ 値による核形成部位の推定

我々は、フォールディング・メカニズムを明瞭にする問題、つまり、『タンパク質は、ランダムコイル状態から如何にして天然構造へフォールドするのだろうか?』という問題に取り組んでいる。そのための第一歩として、タンパク質のフォールディング過程の“遷移状態の構造”を明らかにしたいが、遷移状態はきわめて寿命が短く、しかも不安定のため、通常の構造解析手法は使えない。

多くの研究グループが興味のある挙動を示すタンパク質である、“プロテイン A (1SS1)” (B domain of protein A) のコンピュータ・シミュレーションを行い、そのフォールディング過程の解明に取り組んでいる。

“プロテイン A” (「黄色ブドウ状球菌プロテイン A の B ドメイン」) は、アミノ酸残基数は 62 個で、3 個の α -ヘリックス

($\alpha 1$, $\alpha 2$, $\alpha 3$) をもつ小さなタンパク質である。右の表に、“プロテイン A” のアミノ酸配列を示している。

プロテイン A (1SS1) のアミノ酸配列

```

1 Gly - Ser - Thr - Ala - Asp - Asn - Lys - Phe - Asn - Lys - Glu - Gln - Gln -
Asn - Ala - Phe - Trp - Glu - Ile - Leu - His - Leu - Pro - Asn - Leu - Asn -
Glu - Glu - Gln - Arg - Asn - Gly - Phe - Ile - Gln - Ser - Leu - Lys - Asp -
40 Asp - Pro - Ser - Gln - Ser - Ala - Asn - Leu - Leu - Ala - Glu - Ala - Lys -
Lys - Leu - Asn - Asp - Ala - Gln - Ala - Pro - Lys - Ala
  
```

このタンパク質の特徴は、ある条件下では、フォールディング速度が 10μ 秒程度と非常に速いことである。これは、現在のコンピュータ・シミュレーションのタイムスケールに近い。タンパク質フォールディングのコンピュータ・シミュレーションは、フォールディングの道筋に関する情報を提供できる強力な方法である。実際、Wolynes は、“プロテイン A” に対する、全原子レベルの分子動力学 (Molecular dynamics; MD) によるコンピュータ・シミュレーションを行い、次のような結果を得た：『“プロテイン A” は、まず $\alpha 3$ が形成され、その後、 $\alpha 1$, $\alpha 2$, あるいはターンが形成される。』

<文献> P.G. Wolynes, *Proc. Natl Acad. Sci. USA* **101**, 6837-6838 (2004).

この結果が妥当であるのか、実験による解明が切望されていた。

タンパク質のフォールディング過程の“遷移状態の構造”を、実験的に明らかにしようとする試みとして、Fersht らによって導入された“ Φ 値解析”がある。彼らは、遺伝子工学的手法を用いて、もとの“野生型タンパク質”のフォールディング反応と、任意の部位のアミノ酸残基を他のアミノ酸残基で置換した“変異体タンパク質”のフォールディング反応とにおける、“反応速度の相違”と“自由エネルギーの変化”を観測することによって、フォールディング過程の遷移状態の構造を推定しようとする解析方法を提案し、“実際のタンパク質”に対して実行している (Appendix I: 「実験における Φ 値解析とは?」参照)。

彼らは、具体的に、“プロテイン A” のフォールディングの遷移状態におけるアミノ酸残基の Φ 値を、組み換え DNA 技術によるアミノ酸置換の変異体を利用して、系統的に、そして網羅的に測定する実験を行った。そして、コンピュータ・シミュレーションにおける遷移状態での構造と、実験による Φ 値解析によって得られた遷移状態の構造とを比較した。

<文献> Sato, S., Religa, T.L., Daggett, V. & Fersht, A.R. Testing protein-folding simulations by experiment: B domain of protein A. *Proc. Natl Acad. Sci. USA* **101**, 6952 - 6956 (2004).

<文献> 佐藤 聡; 「アミノ酸変異導入による蛋白質折れたたみ遷移状態の解析」, 蛋白質・核酸・酵素 Vol.49, No.14 (2004).

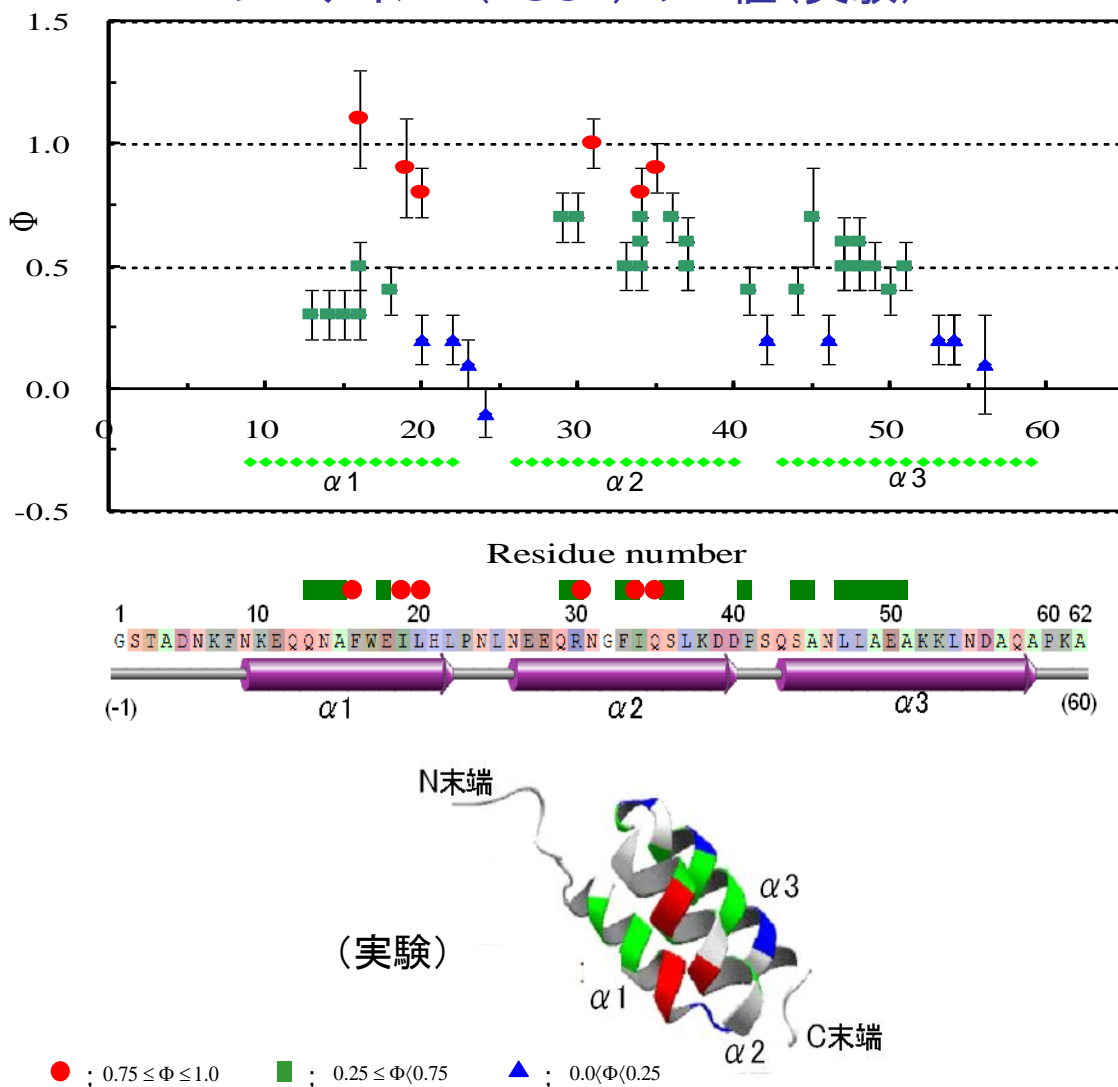
次のページの図は、Fersht らによる“1SS1” (プロテイン A) の実験による Φ 値の結果を、アミノ酸配列上と天然構造に重ね合わせて、3 色で分類して描いている。

この結果によって推定される遷移状態での構造形成度の考察は次の通りである：

- (1) “ $\alpha 1$ -ヘリックスの領域” の C 末端側の残基での Φ 値は高いが、他の残基での Φ 値は 0.3~0.5 と低い。このことは、遷移状態においては $\alpha 1$ -ヘリックス領域自体の構造形成度は低く、“構造形成の途上である” とみられる。
- (2) “ $\alpha 2$ -ヘリックス領域” での Φ 値は全般的に高い値である。このことは、 $\alpha 2$ -ヘリックスの領域は、“遷移状態においてほぼ天然構造を形成している” とみなせる。

- (3) “ $\alpha 3$ -ヘリックスの領域”のN末端側の Φ 値は0.4~0.7という比較的大きい値を示すものの、C末端側に向かって徐々に減少している。このことは、遷移状態においては $\alpha 3$ -ヘリックスのN末端側ではある程度天然構造を形成しているとみられるが、この領域自体の“構造形成度は低い”とみられる。
- (4) “ターン領域 (T1)” ($\alpha 1$ -ヘリックスと $\alpha 2$ -ヘリックスを結ぶ領域)の Φ 値は低く、遷移状態においては、“ほぼ変性状態”とみなせる。
- (5) “ターン領域 (T2)” ($\alpha 2$ -ヘリックスと $\alpha 3$ -ヘリックスを結ぶ領域)の Φ 値は中間的な値であり、遷移状態においては、“ある程度天然構造を形成している”とみなせる。

プロテインA(1SS1)の Φ 値(実験)



(無色は実験値が無い残基を示す)

◎実験による、『プロテインAの遷移状態での核形成部位の推定』(まとめ)

- “プロテイン A (1SS1)”の遷移状態では、“ $\alpha 2$ -ヘリックス領域がほぼ天然構造を形成”していて、その $\alpha 2$ -ヘリックスが、 $\alpha 1$ -ヘリックスと $\alpha 3$ -ヘリックスに位置する多数の疎水性アミノ酸による相互作用によって安定化されている。このことは、シミュレーションによる結果である、「まず $\alpha 3$ が形成され、その後、 $\alpha 1$ 、 $\alpha 2$ 、あるいはターンが形成される」との予測はみられなかったことになる。
- 具体的には、 $\alpha 2$ -ヘリックス領域の疎水性アミノ酸残基である、“Ile-34 と Leu-37” (ともに高い Φ 値を示している) と相互作用している $\alpha 1$ -ヘリックス領域の疎水性アミノ酸残基 “Ile-19 と

Leu-20”（ともに高いΦ値を示している）が天然構造形成過程の引き金となる重要な核であるとみなせる。

佐藤らは、実験で求められた“プロテイン A (ISS1)”（プロテイン A）のΦ値の分布から、フォールディングのメカニズムは、“核形成-凝縮モデル”に近いと、次のように推論している：

『遷移状態で、α2-ヘリックス領域のアミノ酸残基のΦ値は高く、2次構造を形成しているものの、α1-ヘリックスとα3-ヘリックスのアミノ酸残基のΦ値は中間的な値を示していて、これらの領域の2次構造形成度は低い。“フレームワーク モデル”が該当するためには、α1-ヘリックスとα3-ヘリックスも高度に2次構造を形成しているはずである。プロテイン A (ISS1) では、むしろ、“核形成-凝縮モデル”に近い、つまり、α2-ヘリックス領域を含む拡大したフォールディング核が、多数の相互作用で安定化され、その核を中心に残りの2次構造と3次構造が協同的に形成されているとみなせる。』

＜コメント＞ “フレームワーク モデル”と“核形成-凝縮モデル”については、第10章の「タンパク質のフォールディング・メカニズムのモデル」を参照のこと



“実際のタンパク質分子”のフォールディング過程に、“A-W_NILS モデル”を適用するにはどうすればよいだろうか？

我々が提案している、タンパク質フォールディングの“統計力学的モデル”である「A-W_NILS モデル」を、“実際のタンパク質”に適応するにはどうすればよいのだろうか？

3次元格子タンパク質と比較して、“実際のタンパク質分子”は、あまりにも複雑な系である。たとえアミノ酸残基を単純な球（ユニット）とみなしても、そのポリペプチド鎖の取り得るコンフォメーションの数は莫大で天文学的である。この複雑な系の分配関数を簡単なモデルによって近似的に求めるには、どうすればよいだろうか？ 「A-W_NILS モデル」を“実際のタンパク質分子”に適用するために、新たに次のような“定義”を採用しよう：

(1) 2個のアミノ酸残基（ユニット）の間の“接触”（contact）の定義

2個のアミノ酸残基（ユニット）を構成している原子の、少なくとも1つの原子ペアの距離が、あらかじめ定めた距離 D_c （接触判断距離, cutoff distance）以内ならば、これらの2個のアミノ酸残基（ユニット）は“接触”（contact）しているとみなす。 D_c の具体的な値は、次の4通り；
 $D_c = 4.0, 5.0, 5.5, 6.0 \text{ (Å)}$

(2) 野生型タンパク質のアミノ酸残基間の“接触エネルギー”（contact energy）の値の定義

野生型タンパク質において、 m 番目から j 番目のアミノ酸残基（ユニット）からなる“局所構造内のエネルギーの値”は、次のようになる：

$$E(m, j) = \sum_{m \leq k, l \leq j} U(\xi_k, \xi_l) \Gamma_{k,l}$$

ここで、 $\Gamma_{k,l}$ は、次式で示すように、2つのユニットが“接触”している場合に、アミノ酸残基間の相互作用が発生する；

$\Gamma_{k,l} = 1$: ユニット k, l が“接触”しているとき

$\Gamma_{k,l} = 0$: ユニット k, l が“接触”していないとき

また、3次元格子タンパク質に倣って、ユニット k, l が $|k-l| \leq 2$ の場合； $\Gamma_{k,l} = 0$

$U(\xi_k, \xi_l)$ は、アミノ酸残基 ξ_k と ξ_l との相互作用エネルギーの値である。

以下の“実際のタンパク質”を扱う問題では、アミノ酸残基のタイプ ξ_k と ξ_l によらず、

一定の値を ε と仮定する： $U(\xi_k, \xi_l) = \varepsilon$. 具体的には、 $\varepsilon = -0.10$ とする（homogeneous contact-energy approximation）。

<コメント> 3次元格子タンパク質の場合は、アミノ酸残基のタイプ ξ_k と ξ_l で決まるエネルギー値 (Miyazawa & Jernigan [1985] による値) を採用している (第5章, Appendix G を参照のこと) .

(3) 変異体タンパク質のアミノ酸残基間の“接触エネルギー”(contact energy)の値の定義

野生型のタンパク質から、変異体タンパク質を作成する. これは、野生型タンパク質のアミノ酸配列の、ある一カ所のアミノ酸残基 (ξ_k) を他のアミノ酸残基 (ω_k) に置換したもので、このとき、変異体タンパク質のアミノ酸残基間の“接触エネルギー”(contact energy)の値は、次のように変化する: $U(\xi_k, \xi_l) \rightarrow U(\omega_k, \xi_l)$

$U(\omega_k, \xi_l)$ の値は、次の二つの“接触エネルギー”と取ると仮定する:

$$U(\omega_k, \xi_l) = \varepsilon + 0.01, \quad U(\omega_k, \xi_l) = \varepsilon - 0.01$$

変異体タンパク質では、これらの二つの“接触エネルギー”対して Φ 値を求める. そして、それらの Φ 値の平均値を、変異体タンパク質のアミノ酸残基(置換したアミノ酸残基)に対する Φ 値とする.

(4) “実際のタンパク質”のコンフォメーションに対するエントロピーの定義

現在のところ、実際のタンパク質のコンフォメーションに対するエントロピー項、 $f(m, j)$ の正確な関数形は確立していない. 我々は、3次元格子タンパク質に対して採用した次のような関数形を、そのまま採用しよう(最善の関数形であるとは言い難いが・・・):

$$f(m, j) = 1.4084 \times (4.750)^{j-m-2}$$

(1)~(4)の定義を採用すると、調節パラメータ (adjustable parameter) を含まない統計力学モデルである「A-W_NILS モデル」を、“実際のタンパク質”のフォールディング過程にも適用できる. その際の手順に従って、分配関数を求め、その分配関数より、フォールディング・レート k_f と、アンフォールディング・レート k_u を計算し、最終的に、 Φ 値を求めることができる.

Coffee Break

一般に、タンパク質の天然構造のコンフォメーション・エネルギーは最小値をとるであろう.

“実際のタンパク質”の天然構造において、2個のアミノ酸残基間の“接触”の情報から、“接触”しているアミノ酸残基ペアを求め、アミノ酸残基のタイプで決まるエネルギー値 (Miyazawa & Jernigan [1985] による値) を採用して、天然構造のコンフォメーション・エネルギー値を求めた. その結果は、天然構造で“接触”しているアミノ酸残基ペアの数は最大にもかかわらず、天然構造のコンフォメーション・エネルギー値が最小であるという条件を満たさなかった.

新たな試みとして、アミノ酸残基のタイプによらず、“接触”しているアミノ酸残基ペアの接触エネルギー値を、一定の値 ε とすることを試みた. そして、予備的に、 $-0.05 \sim -0.40$ の範囲で幾つかの ε の値を決定し、 Φ 値計算を実行した.

その結果は ε の値には、ほぼ依存しないことが明らかになった.
最終的に、 $\varepsilon = -0.10$ として、 Φ 値を計算することにした.

各アミノ酸残基のΦ値は、タンパク質フォールディングの転移温度において、次のような二つの量が計算できれば理論的に求められる：

- ・ 野生型タンパク質の自由エネルギーのプロファイルから求まるフォールディング・レート；

$$k_f^{wild}(T_m)$$

- ・ 変異体タンパク質の自由エネルギーのプロファイルから求まるアンフォールディング・レート；

$$k_u^{mut}(T_m)$$

図は、フォールディング・レート (k_f) とアンフォールディング・レート (k_u) の温度 (の逆数) 依存曲線 (シェブロン・プロット, Chevron plot) のイラストである。

転移温度 (T_m) においては、図に示しているように、次式が成立する；

$$\ln k_f^{wild}(T_m) = \ln k_u^{wild}(T_m)$$

故に、

$$\begin{aligned} \Delta\Delta F_{N-D} &= -k_B T_m ((\ln k_f^{mut}(T_m) - \ln k_f^{wild}(T_m)) - (\ln k_u^{mut}(T_m) - \ln k_u^{wild}(T_m))) \\ &= -k_B T_m (\ln k_f^{mut}(T_m) - \ln k_u^{mut}(T_m)) \end{aligned}$$

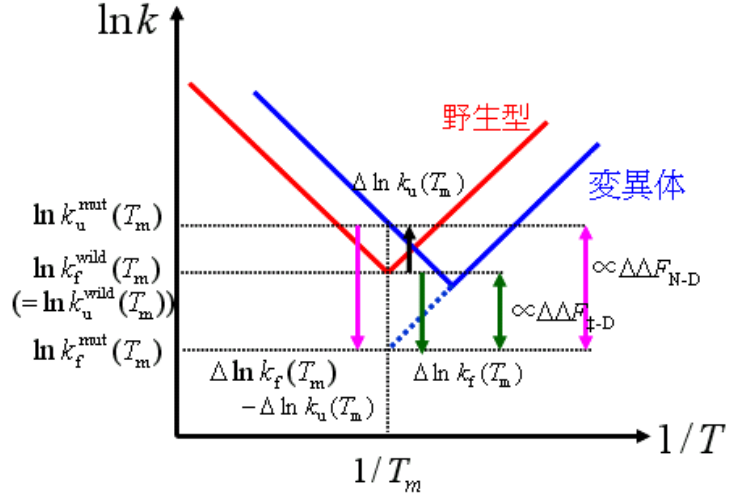
さらに、

$$\Delta\Delta F_{\ddagger-D} = -k_B T_m \Delta \ln k_f(T_m)$$

結局、転移温度 (T_m) におけるΦ値は次のように表される；

$$\begin{aligned} \Phi &= \frac{\Delta\Delta F_{\ddagger-D}}{\Delta\Delta F_{N-D}} \\ &= \frac{\Delta \ln k_f(T_m)}{\ln k_f^{mut}(T_m) - \ln k_u^{mut}(T_m)} \\ &\quad \left(\text{ただし, } \Delta \ln k_f(T_m) = \ln k_f^{mut}(T_m) - \ln k_f^{wild}(T_m) \right) \end{aligned}$$

<コメント>： 詳細は、第7章：『フォールディングのキネティクスにおけるレートと緩和時間の関係は？』、第8章：『タンパク質フォールディングの遷移状態の構造の推定するΦ値解析とは？』、及び、Appendix H：「フォールディング・レートの計算方法」 参照。

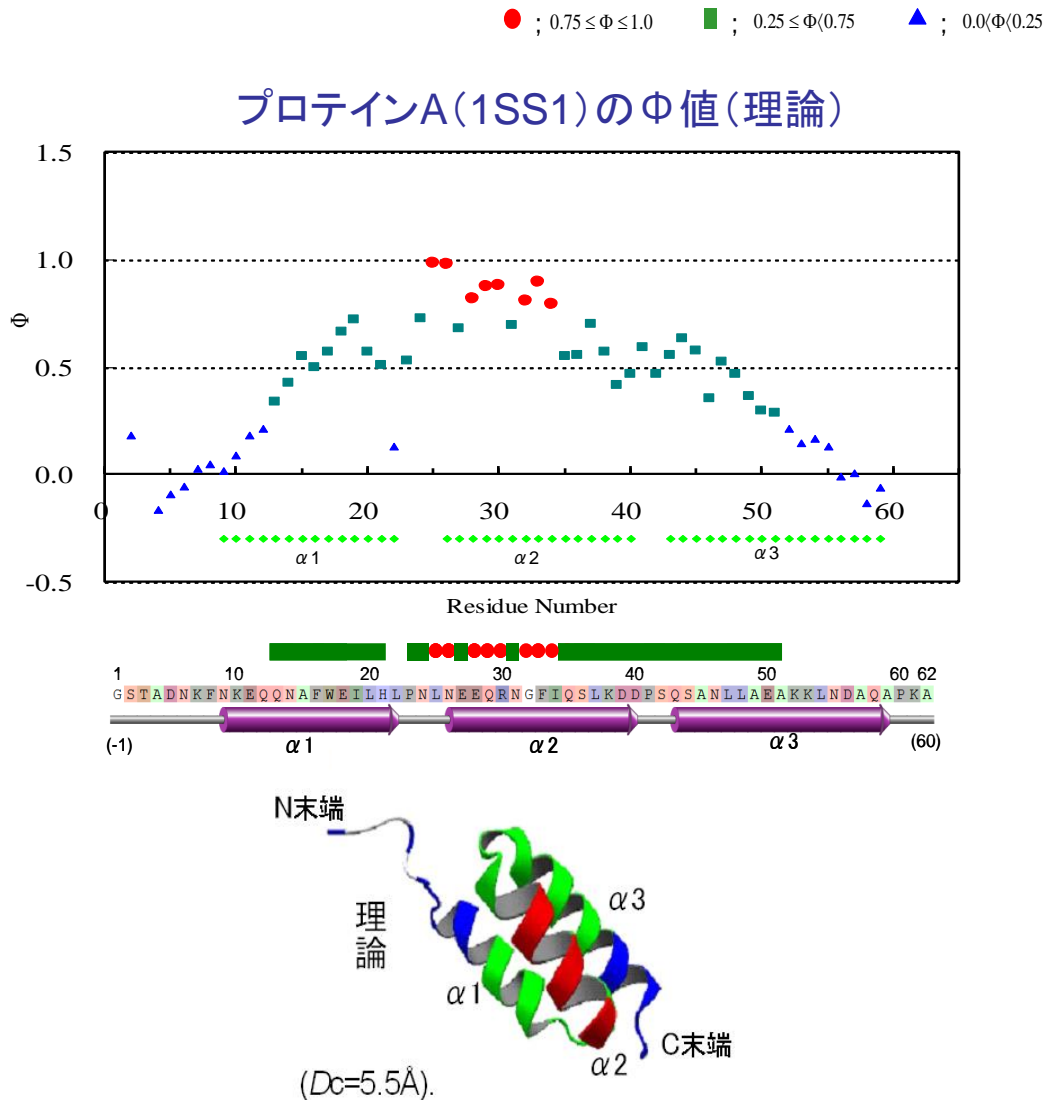




理論的に求めたプロテイン A の遷移状態の Φ 値による核形成部位の推定

統計力学的モデルである「A-W_NILS モデル」を用いて、具体的に、“実際のタンパク質”である“プロテイン A (1SS1)”に対して、そのアミノ酸残基の Φ 値を理論的に計算した結果を示そう。そして、この理論的に求めた Φ 値と、実験的に観測された Φ 値とを比較しよう。両方の Φ 値の相関を調べることで、プロテイン A のフォールディング・メカニズムとして、「フレームワーク モデル」が良いか、あるいは、「核形成-凝縮モデル」が適当かを検討しよう。

下図は、プロテイン A の理論による結果を、アミノ酸配列上に 3 色でアミノ酸残基を分類して描いている。更に、二つのアミノ酸残基間の接触判断距離 $D_c = 5.5(\text{\AA})$ の場合に、理論による Φ 値を天然構造に重ね合わせて分類した色で描いている。



理論的に求めたプロテイン A (1SS1) に対するアミノ酸残基毎の Φ 値から推定される“遷移状態での構造形成度”を考察しよう：

- (1) “ $\alpha 1$ -ヘリックスの領域”での Φ 値は“中間的な値”であり、N 末端側に向かって徐々に低い値になっている。このことは、遷移状態においては $\alpha 1$ -ヘリックスは、“構造形成の途上である”とみられる。
- (2) “ $\alpha 2$ -ヘリックス領域”で、“N 末端側の Φ 値は高い値”である。このことは、 $\alpha 2$ -ヘリックスの領域は、遷移状態においてほぼ“天然構造を形成している”とみなせる。
- (3) “ $\alpha 3$ -ヘリックスの領域”の Φ 値は、 $\alpha 1$ -ヘリックスの領域での Φ 値と同様に“中間的な値”で

あり、C 末端側に向かって徐々に低い値になっている。このことは、遷移状態においては $\alpha 3$ -ヘリックスは、“構造形成の途上にある”とみなせる。

- (4) “ターン領域 (T1)” ($\alpha 1$ -ヘリックスと $\alpha 2$ -ヘリックスを結ぶ領域) の Φ 値は高く、遷移状態においては、“ほぼ構造形成している”とみなせる。
- (5) “ターン領域 (T2)” ($\alpha 2$ -ヘリックスと $\alpha 3$ -ヘリックスを結ぶ領域) の Φ 値は“中間的な値”であり、遷移状態においては、“ある程度天然構造を形成している”とみなせる。

◎理論による『プロテイン A の遷移状態での核形成部位の推定』(まとめ)

- ・プロテイン A (ISS1) の遷移状態では、“ $\alpha 2$ -ヘリックス領域”がほぼ天然構造を形成して、この領域が“核形成部位”となり、両側の $\alpha 1$ -ヘリックスと $\alpha 3$ -ヘリックスに位置する多数の疎水性側鎖による相互作用によって安定化されて、最終的に天然構造へフォールドするとみなせる。
- ・具体的には、 $\alpha 2$ -ヘリックス領域のアミノ酸残基である、“Arg-28, Asn-29, Gly-30, Ile-32, Gln-33, Ser-34”と、ターン領域 (T1) のアミノ酸残基である、“Glu-25, Glu-26” (これらのアミノ酸残基は高い Φ 値を示している) が、天然構造形成過程の引き金となる重要な核部位であるとみなせる。



プロテイン A (ISS1) に対する、 Φ 値解析の実験と理論の比較

プロテイン A (ISS1) Φ 値に関して、実験の結果と理論的に求めた結果とを比較・検証しよう。そして、そのフォールディング・メカニズムを考察しよう。

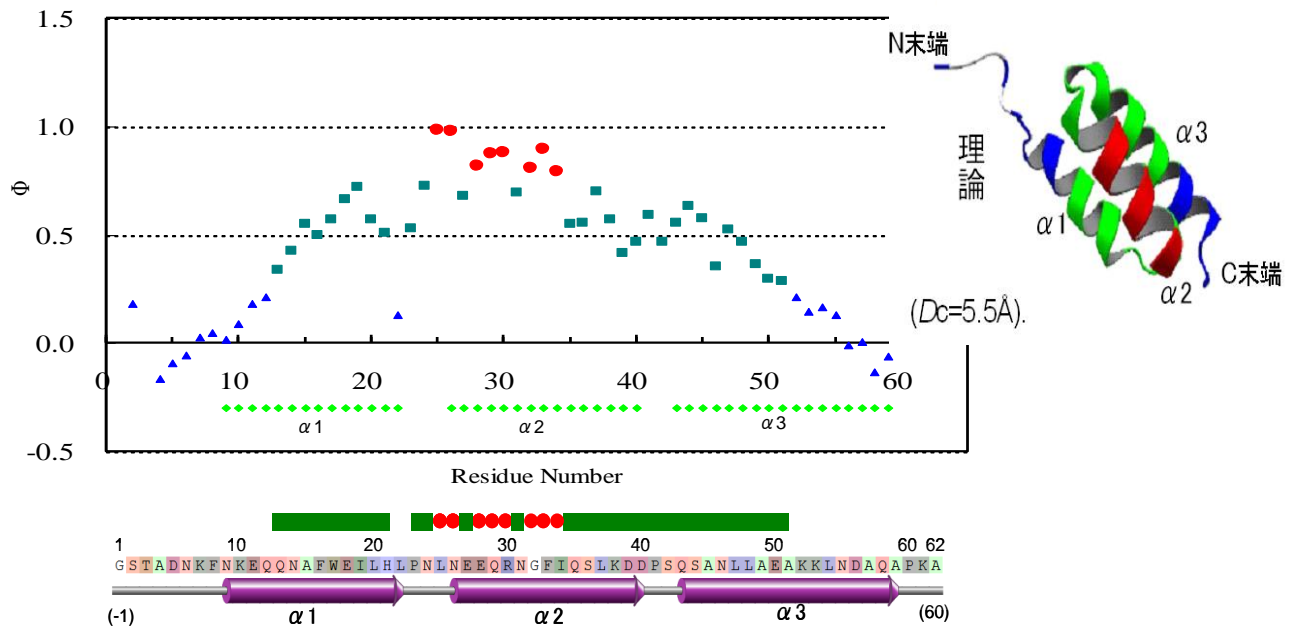
次ページの図では、実験的に Φ 値解析によって調べられたプロテイン A のフォールディング遷移状態構造と、それを「A-W_NILS モデル」を用いて理論的に求めた結果の両方を並べて示している。ただし、理論計算では、二つのアミノ酸残基間の接触判断距離 D_c (Cutoff distance) を、 $D_c = 5.5 \text{ \AA}$ としている。

この図を一瞥すると、“理論と実験との結果は、細かい差異は認められるが、ほぼ良好な一致を示している。実際、プロテイン A に対して、理論的に計算した Φ 値と、実験的に観測された Φ 値との相関係数 (CC; Correlation Coefficient) は、 $CC = 0.71$ であった。この CC の高い値は、プロテイン A は、我々が採用している統計力学モデルである「A-W_NILS モデル」で仮定したスキームに沿ってフォールドしているとみなせることを示唆している。

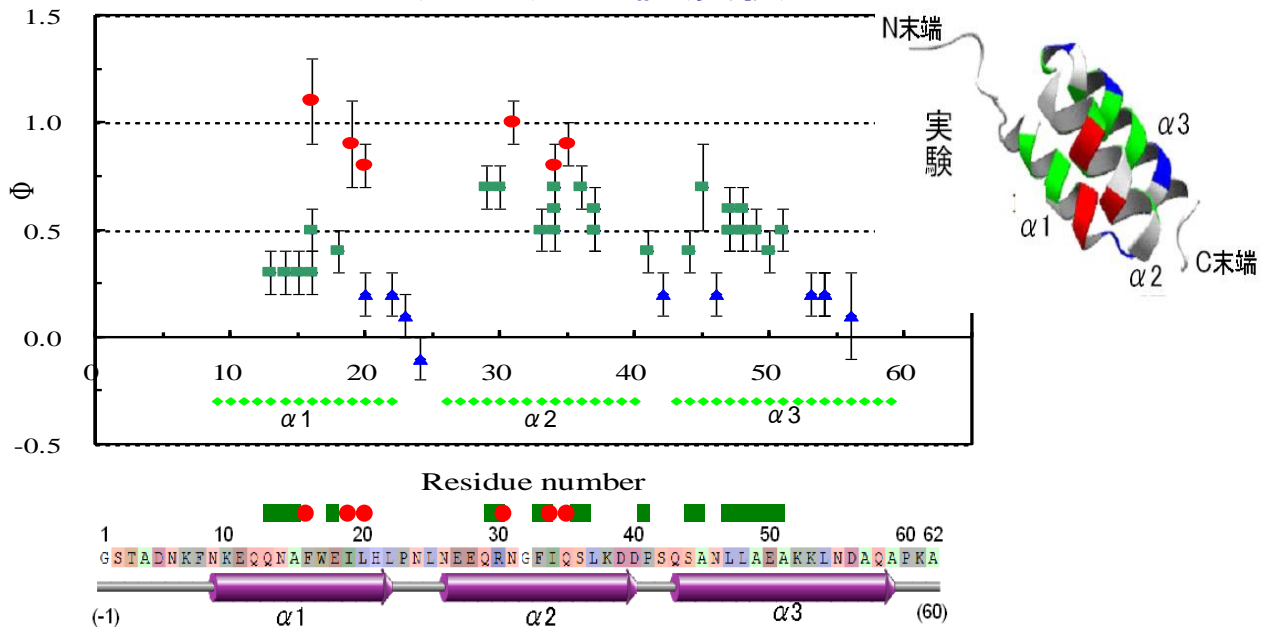
理論計算では、タンパク質立体構造データベース (PDB; Protein Data Bank) から得られた天然構造での天然接触ペアの情報を用いて、天然接触ペアの相互作用エネルギーの値は、アミノ酸残基の種類にかかわらず、近似として、ある一定値を割り当てる。つまり、アミノ酸残基ごとの個性をもっていることを省略している。また、置換したアミノ酸残基と他のアミノ酸残基間の相互作用エネルギーの変化量は、正または負の小さい値を採用している。このことは、理論計算における結果への天然構造のトポロジーの寄与が多きいこと、つまり、その重要性を示している。しかも、「A-W_NILS モデル」という非常に単純化された統計力学モデルにもかかわらず、プロテイン A に対する実験で観測された Φ 値がかなり再現されていることは、統計力学モデルである「A-W_NILS モデル」の仮定が妥当であることを示唆していると見なせるであろう。

● ; $0.75 \leq \Phi \leq 1.0$ ■ ; $0.25 \leq \Phi < 0.75$ ▲ ; $0.0 < \Phi < 0.25$

プロテインA(1SS1)の Φ 値(理論)



プロテインA(1SS1)の Φ 値(実験)



佐藤らは、プロテイン A (1SS1) のフォールディングのメカニズムが、“核形成-凝縮モデル”に近いと推論した際に、『フレームワーク モデルが該当するためには、 $\alpha 1$ -ヘリックスと $\alpha 3$ -ヘリックスも高度に2次構造を形成しているはずである』と述べているが、我々の「A-W_NILS モデル」を用いた結果が実験結果をほぼ再現していることから、プロテイン A のフォールディングのメカニズムは、「核形成-凝縮モデル」よりも、「フレームワーク モデル」(「A-W_NILS モデル」)の方が妥当であるとみなしてよいであろう。

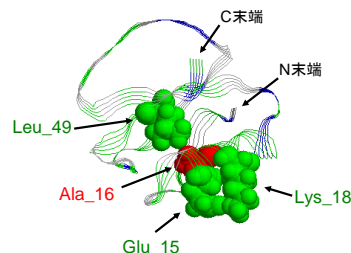
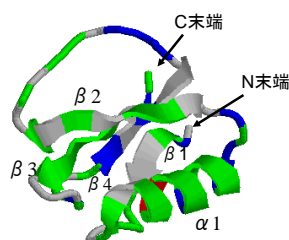
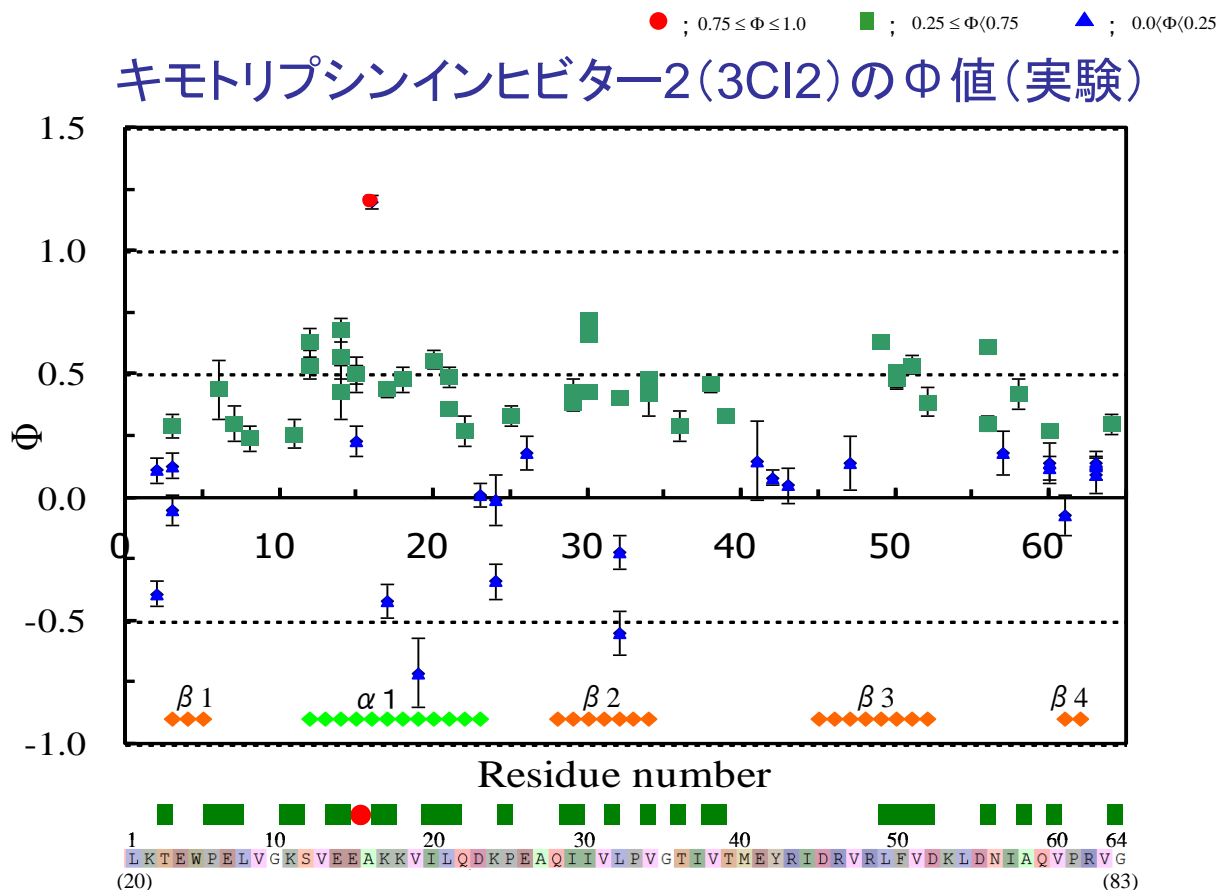
キモトロプシンインヒビター2 (3CI2) の実験による Φ 値とそのフォールディング・メカニズム

Fersht らは, “キモトリプシンインヒビター2 (3CI2)” (chymotrypsin inhibitor) のフォールディング過程のキネテクスの実験を行った. キモトロプシンインヒビター2 では, 中間体は観測されず, 変性状態 (D 状態) から天然状態 (N 状態) への二状態転移であることを観測した. その後, キモトロプシンインヒビター2 と同程度の大きさをもつタンパク質が集中的に実験され, 多くが, 同様にフォールディング過程では中間体を作らない, 二状態転移を示すことが明らかになった.

中間体を形成しないタンパク質ではそのフォールディング過程の途中で何が起きているのかを調べるために, キモトロプシンインヒビター2 のアミノ酸変異体を 100 個以上作り Φ 値解析を実行してアミノ酸残基の Φ 値を求め, 遷移状態の構造が推定された.

下図は, Fersht らによるキモトロプシンインヒビター2 の実験による Φ 値の結果 (実験条件; “4M GdmCl concentration”) を, アミノ酸配列上と天然構造に重ね合わせて, 3 色で分類して描いている.

<文献> Itzhaki, L.S., Otzen, D.E. & Fersht, A.R. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**, 260-288 (1995).



実験による“3CI2”の Φ 値は、ほとんどの部位で、おおむね、0.20～0.5の間に集中している。より大きな Φ 値は、 α 1-ヘリックスのアミノ酸残基にみられる。特に、『16番目のアミノ酸残基，“Ala_16”が大きな Φ 値をもっている』ことがわかる。彼らは、『“Ala_16”が，“フォールディングの核形成部位”である』と提案している。まとめると、 Φ 値が0.5以上のアミノ酸残基は、 α 1-ヘリックス中のN末端側の数残基と、このヘリックスにドッキングしている β シート中の数残基のみで、他のアミノ酸残基はおおむね0.2～0.5に収まっている。

Φ 値をその大きさを分類して、図のように天然構造上に表示した。特にアミノ酸残基“Ala_16”に近接していて Φ 値が0.5以上の残基である，“Glu_15”，“Lys_18”，“Leu_49”を天然状態の構造に重ねてその構成原子をファンデルワース半径で表示している。

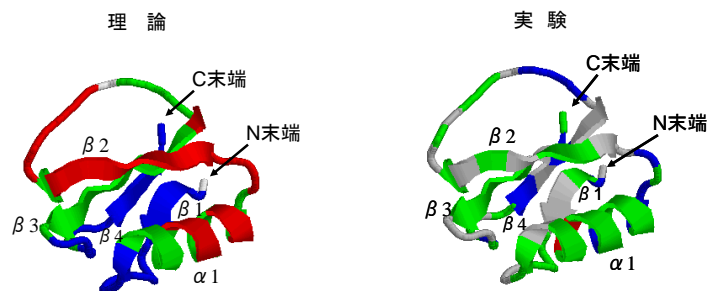
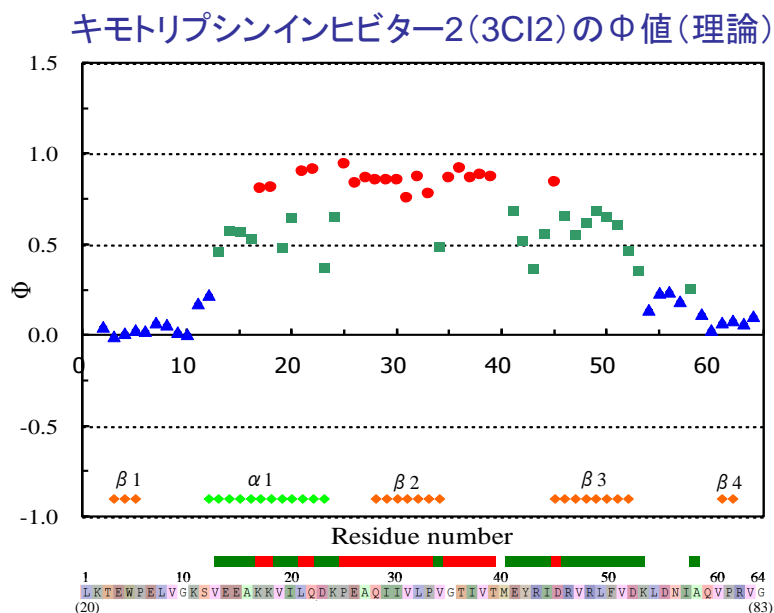
彼らは、このキモトリプシンインヒビター2 (3CI2) の遷移状態では、幾つかのアミノ酸残基が互いに近づいて、しかも、アミノ酸の一次配列上で離れたアミノ酸残基も含んで“フォールディング核”を形成していると推定した。そして、そのフォールディング過程を、次のように描いている：

『キモトロプシンインヒビター2 のフォールディングの初期では変性状態 (D 状態) にあり、何らかの揺らぎ運動を行っている。ある瞬間、特定の数個のアミノ酸残基が衝突して合体し“フォールディング核”が形成され、さらに、その周りのアミノ酸残基が“凝集”を起こして成長し、最終的に天然構造 (N 構造) を形成する。“フォールディング核”は、特定のアミノ酸残基“Ala16”から形成されるが、多くのアミノ酸残基の部位で独立にフォールディングが進むことはない。つまり、遷移状態では、各部位で独立に安定な2次構造が形成されることはない。』

彼らは、キモトロプシンインヒビター2 のフォールディング過程の考察より、小さいタンパク質のフォールディングの一般的なメカニズムとして、次のような、“核形成-凝縮モデル” (nucleation-condensation model) を提案した：『小さいタンパク質のフォールディングは、“核形成” (nucleation) と呼ぶ遷移状態に至るまでの過程と、“凝縮” (condensation) というそれ以降の過程から成る。』

キモトリプシンインヒビター2 (3CI2) の理論による Φ 値

我々は、キモトリプシンインヒビター2 の Φ 値を、統計力学モデル「A-W_NILS モデル」を用いて、理論的に計算した。その結果、キモトリプシンインヒビター2 の、理論的に計算された Φ 値と、実験的に観測された Φ 値との相関係数 (CC; Correlation Coefficient) の値は、 $CC = 0.45$ （ただし、理論：二つのアミノ酸残基間の接触判断距離 $D_c = 4.0 \text{ \AA}$ ，実験条件：4M GdmCl）である。中間的なケースで、タンパク質のアミノ酸残基のある領域 (α 1-ヘリックス) では実験を再現しているが、他の領域ではほとんど再現していない。結局、キモトロプシンインヒビター2 のフォールディング・メカニズムは、「A-W_NILS モデル」が基盤とする“フレームワーク モデル” (framework model) ではなく、むしろ、“核形成-凝縮モデル” (nucleation-condensation model) と推定できる。



次章で、我々は、27 個の“実際のタンパク質”に対して、統計力学モデル“A-W_NILS モデル”を適用して理論的に計算した Φ 値と、実験的に観測された Φ 値と比較した結果をまとめて表 (Table 1) に示そう。その際、二つのアミノ酸残基間の接触判断距離 (Cutoff distance, D_c) を次のような 4 通りの場合に対して Φ 値を理論的に計算している： $D_c = 4.0, 5.0, 5.5, 6.0 (\text{\AA})$ 。

27 個の“実際のタンパク質”に対して、理論的に計算された Φ 値と、実験的に観測された Φ 値との相関係数 (CC; Correlation Coefficient) の値を求めているが、この CC の値によって、“A-W_NILS モデル”で仮定したスキームに沿ってタンパク質がフォールドしているかどうかを推察しよう。

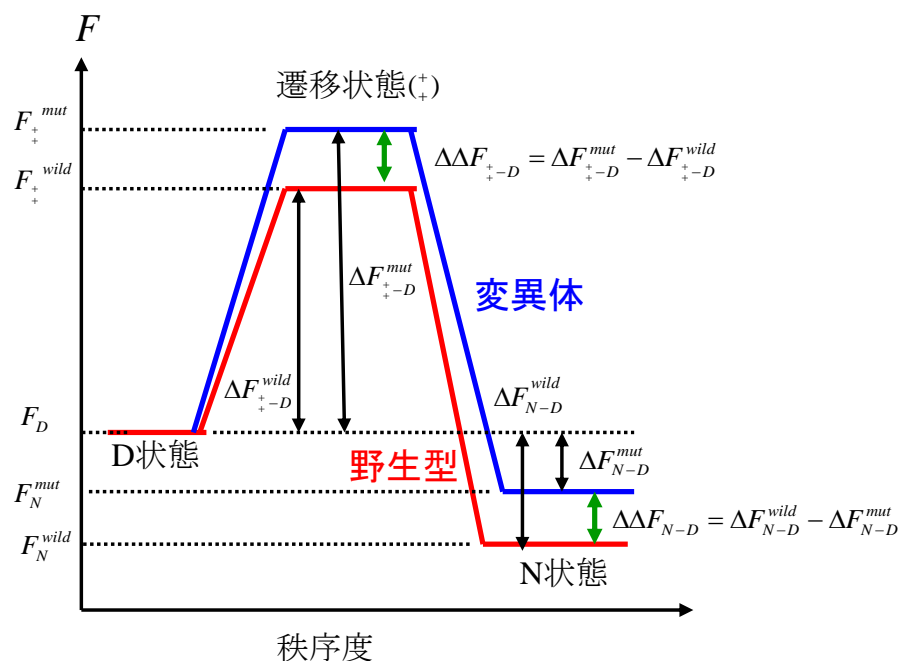
実験におけるΦ値解析とは？



実験における“Φ値解析法”とは？

Fersht らが提案した実験による“Φ値解析”という手法は、どのようにして、フォールディング過程の遷移状態での構造を推定することができるのだろうか？自由エネルギー曲線との関係を考察しよう。

下図は、野生型タンパク質の自由エネルギー曲線（赤色）と変異体タンパク質の自由エネルギー曲線（青色）を模式的に描いている。



実験における，“Φ値解析法”は、次の2つの熱力学的実測からなる：

- 熱力学的実測1；アミノ酸変異によって，“平衡論的な安定性に与える変化量”の実測
- 熱力学的実測2；アミノ酸変異によって，“遷移状態の自由エネルギーに与える変化量”の実測

それぞれの熱力学的実測の手順を示そう。

◎熱力学的実測1；

<手順1>

遷移状態の構造的特徴をアミノ酸残基レベルで観測するタンパク質を選ぶ（そのタンパク質を“野生型タンパク質”と呼ぶ）。

<手順2>

熱力学測定から、野生型タンパク質の変性状態（D 状態）と天然状態（N 状態）との自由エネルギー差（ $\Delta F_{N-D}^{wild} = F_D - F_N^{wild}$ ）を測定する。

<手順3>

“野生型タンパク質”のアミノ酸配列の特定の部位のアミノ酸残基を選び、その残基をグリシン（Gly）や、アラニン（Ala）などの体積の小さなアミノ酸残基に置き換える。

<コメント> 組み換え DNA 技術を用いてアミノ酸置換したタンパク質を“変異体タンパク質”と呼ぶ。置換するアミノ酸がグリシンやアラニンである理由については章末の <Coffee Break> を参照のこと。

<手順4>

この“変異体タンパク質”の、変性状態（D 状態）と天然状態（N 状態）との自由エネルギー差（ $\Delta F_{N-D}^{mut} = F_D - F_N^{mut}$ ）を測定する。

<手順5>

手順2と手順4の結果から、アミノ酸変異によって“平衡論的な安定性に与えた変化量”， $\Delta\Delta F_{N-D} = \Delta F_{N-D}^{wild} - \Delta F_{N-D}^{mut}$ を求める。

◎熱力学的実測2；

<手順1>

“野生型タンパク質”のフォールディングの反応速度（ k_f^{wild} ）を測定する。この反応速度 k_f^{wild} と、遷移状態の自由エネルギー（ F_+^{wild} ）と変性状態の自由エネルギー（ F_D ）の差（ ΔF_+^{wild} ）との関係は、次式で表される（Arrhenius's equation）。

$$k_f^{wild} = k_0 e^{-\Delta F_+^{wild} / k_B T}$$

（ここで、 k_0 ；任意の定数、 k_B ；Boltzmann 定数、 T ；絶対温度）

<手順2>

“変異体タンパク質”のフォールディングの反応速度（ k_f^{mut} ）を測定する。この反応速度 k_f^{mut} と、遷移状態の自由エネルギー（ F_+^{mut} ）と変性状態の自由エネルギー（ F_D ）の差（ ΔF_+^{mut} ）との関係は、次式で表される（Arrhenius's equation）。

$$k_f^{mut} = k_0 e^{-\Delta F_+^{mut} / k_B T}$$

（ここで、 k_0 ；任意の定数、 k_B ；Boltzmann 定数、 T ；絶対温度）

<手順3>

“野生型タンパク質”のフォールディングの反応速度（ k_f^{wild} ）と、“変異体タンパク質”のフォールディングの反応速度（ k_f^{mut} ）から、“遷移状態の自由エネルギーに与えた変化量”， $\Delta\Delta F_+^{mut}$ はを、次式のように求めることができる：

$$\Delta\Delta F_+^{mut} = \Delta F_+^{mut} - \Delta F_+^{wild} = k_B T \ln\left(\frac{k_f^{wild}}{k_f^{mut}}\right)$$

<手順4>

熱力学的実測1から求まる，“遷移状態の自由エネルギーに与えた変化量”， $\Delta\Delta F_+^{mut}$ と、熱力学的実測2から求まる，“アミノ酸変異によって平衡論的な安定性に与えた変化量”， $\Delta\Delta F_{N-D}$ との比

から、 Φ 値次式のように求めることができる：

$$\phi = \frac{\Delta\Delta F_{+D}}{\Delta\Delta F_{N-D}}$$

この Φ 値は、タンパク質フォールディングの遷移状態における構造形成度を 0 から 1 の数値で示すものである ($0 \leq \phi \leq 1$)。

Coffee Break

Fersht らは、“ Φ 値解析”において、どのようなアミノ酸置換を考慮したのだろうか？

彼らは、対象とするタンパク質において、『アミノ酸残基間の特定の相互作用を排除するようなアミノ酸置換』を行っている。

つまり、タンパク質の天然構造を変化させずに、特定のアミノ酸残基間相互作用にのみ影響を与えるようなアミノ酸変異を実行している。

このとき、新たなアミノ酸残基間相互作用が生じないように、大きなアミノ酸残基を小さなアミノ酸残基に置換することになる。D 状態においてはアミノ酸置換の影響は無視できるほど小さいとみなしている。

“理想的なアミノ酸置換”として、次のような具体例を考慮している：

イソロイシン (Ile) → バリン (Val)

アラニン (Ala) → グリシン (Gly)

トレオリン (Thr) → セリン (Ser)

これらの置換では、単一のメチル基が削除され、影響を受ける相互作用の数は少ない。

更に、タンパク質内部の疎水性相互作用に影響を与え、天然構造の形成度を反映する置換として、次のような具体例も採用している：

バリン (Val) → アラニン (Ala) ,

イソロイシン (Ile) → アラニン (Ala) ,

ロイシン (Leu) → アラニン (Ala)



“Autumn Morning”