

## II部 タンパク質のコンピュータ・シミュレーション

# 7章 全原子模型のタンパク質のフォールディング

おかもと ゆうこう

岡本 祐幸

岡崎国立共同研究機構分子科学研究所  
総合研究大学院大学

### 1. はじめに

タンパク質の自然の立体構造(自由エネルギーが最小の状態)が、どのような物理的原理や機構によって構築されるかを理解しようとする問題は、タンパク質の折り畳み問題(protein folding problem)とよばれる。この問題はタンパク質が細胞内で合成される事実を考えると、環境因子が複雑すぎて、一見、理論家の挑戦を絶望的に退けているように思われる。しかし、1960年代初頭のいわゆる Anfinsen の実験<sup>1)</sup>は、試験管内でタンパク質の変性状態から自然の立体構造を再構築するものであり、細胞内のもろもろの複雑な環境因子を忘れて、タンパク質1分子とそのまわりの溶媒のみを考慮すればよいことを示唆していた。よって、この実験は、コンピュータ・シミュレーションによってタンパク質の折り畳み問題が解決できるという希望を多くの理論家に与えたが、折り畳み問題は依然として未解決の難問である。

タンパク質の折り畳み問題を理解するには、タンパク質の自然の立体構造ばかりでなく、変性状態や折り畳みの中間状態など、配位空間全体の情報が必要となる。よって、分子動力学法(MD)やモンテカルロ法(MC)などによるコンピュータ・シミュレーションで折り畳み問題を議論するとき、まず、エネルギー関数の精度を議論する必要がある。ここでは、原子レベルの詳細を取り入れたタンパク質の系を考える。このとき、われわれは大きな困難に突き当たる。それは、系の自由度が多いことに起因する。とくに、1個のタンパク質分子につき、考慮しなければならない溶媒分子の個数が膨大であるので、溶媒の寄与を精度よく見積もるのに多くの計算時間を要することになる。また、たとえ溶媒を忘れて、タンパク質分子1個に注目したとしても、エネルギー極小状態に対応する準安定構造が無数に存在する。よって、従来の手法によっては、シミュレーションが、それらエネルギー極小状態に留まってしまって、最小状態に対応する自然の構造に到達するのが至難の業となる。この困難を克服するための筆者らの戦略は、シミュレーションがエネルギー極小状態に留まらない強力な手法を導入することであ

#### Anfinsen の実験

水溶液中で折り畳まれているタンパク質に変性剤を加えることによって、立体構造を完全に壊して、ランダムコイル状態の一本鎖にしてしまったあと、変性剤を取り除いて、元の天然条件に戻してやると、自然の立体構造が回復することを示した実験のこと。1960年代初頭にリボヌクレアーゼAというタンパク質において、Anfinsen らによって初めて示されたが、多くのタンパク質において成り立つことが示されており、「Anfinsen のドグマ」という言葉でタンパク質の折り畳み問題の指導原理となっている。

る。本章では、原子レベルの詳細を取り入れた分子シミュレーションによるタンパク質の折り畳み問題について議論する。

## 2. タンパク質の折り畳みシミュレーション

タンパク質の折り畳み問題を研究するには、自然の立体構造の近傍だけではなく、変性状態や折り畳みの中間状態などを含む、配位空間の全体像を調べる必要がある。よって、分子シミュレーションでは、2章で紹介した、拡張アンサンブル法 (generalized-ensemble algorithm) がとくに適した手法とすることができる。拡張アンサンブル法では、ただ1回のシミュレーションの結果から、エネルギー最小状態ばかりでなく、物理量の任意の温度におけるカノニカルアンサンブル平均を得ることができるので、高温における変性状態から低温における自然の状態までの情報を一度に求めることができるわけである。

本章では、タンパク質の立体構造がどのような原理や機構によって構築されるのかを理解するためには、どのような物理量を計算すればよいのかを示すことに主眼を置く。よって簡単のため、計算対象はおもに気相(真空)中の小ペプチドを例にとるが、あとは計算機時間をかけることによって、水中の普通のサイズ(アミノ酸数300程度)のタンパク質においても同様の解析をすればよいことになる。ここで扱ったペプチド系は5個のアミノ酸からなる Met-enkephalin である。アミノ酸配列は Tyr-Gly-Gly-Phe-Met で与えられる。

モンテカルロシミュレーションではポテンシャルエネルギーとして、ECEPP/2<sup>2)</sup>のものを使用した。文献3で開発された KONF 90というプログラムに拡張アンサンブル法が組み込まれて実行された。また、分子動力学シミュレーションでは、ポテンシャルエネルギー(力場)として、AMBER<sup>4)</sup>のものを使用した。文献5で開発されたプログラム(PRESTO<sup>6)</sup>に基づいている)に拡張アンサンブル法が組み込まれて実行された。

たとえば、ECEPP/2の場合のポテンシャルエネルギー  $E_{\text{tot}}$  (単位は kcal/mol) は、以下のように、静電相互作用項  $E_{\text{C}}$ 、12-6レナード・ジョーンズ (LJ) 項  $E_{\text{LJ}}$ 、水素結合項  $E_{\text{HB}}$  の分子内のすべての原子対についての和に、すべての結合まわりの回転角についての和である、ねじれエネルギー項  $E_{\text{tor}}$  を足したもので与えられる。

$$\begin{aligned}
 E_{\text{tot}} &= E_{\text{C}} + E_{\text{LJ}} + E_{\text{HB}} + E_{\text{tor}} \\
 E_{\text{C}} &= \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}} \\
 E_{\text{LJ}} &= \sum_{(i,j)} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \\
 E_{\text{HB}} &= \sum_{(i,j)} \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\
 E_{\text{tor}} &= \sum_i U_i \{1 \pm \cos(n_i \chi^i)\}
 \end{aligned} \tag{1}$$

ここで、 $r_{ij}$  は  $i$  番目の原子と  $j$  番目の原子の間の距離(単位は Å)である。また、 $\epsilon$

は誘電率,  $\chi^i$  はボンド  $i$  における二面角である. 各原子ではその重心に相互作用の中心があるとし, 部分電荷  $q_i$  (単位は電子電荷) がそこに集中しているとする. さらに,  $E_c$  中の 332 という因子はエネルギーを kcal/mol の単位で表すための係数である. 残りのエネルギー関数としては, 溶媒の寄与があるが, ここでは簡単のため, とくに断らない限り, 気相(真空)中の計算結果を紹介する.

さて, Met-enkephalin の (ECEPP/2) の最小ポテンシャルエネルギー値は  $E_{GS} = -12.2$  kcal/mol であることが知られている<sup>7)</sup>. この構造を構造 A とよぶことにし, 図 1 に示した. この図は球棒図と空間充填図を重ね合わせたものであり, 後

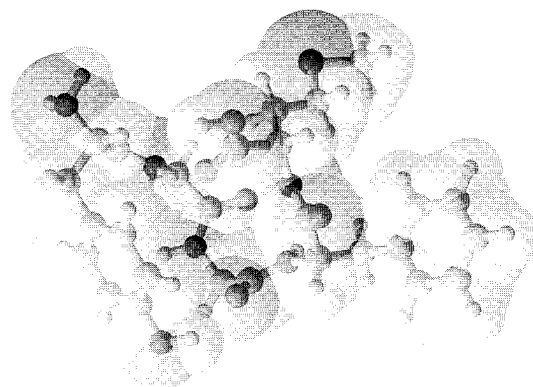


図 1 Met-enkephalin の最小エネルギー構造(構造 A)

者の図はのちに議論されるこのペプチドの体積についての実感をもてるように加えられた.

この系において, 完全にランダムな初期構造から始めて, Tsallis 統計に基づく拡張アンサンブル法<sup>8)</sup>による MC シミュレーションを 1 回だけ実行した (MC sweep 数<sup>\*1</sup>は 1,000,000 回であった).

拡張アンサンブル法では構造空間の探索が大きく加速されることを調べるため, まず, ポテンシャルエネルギーの MC sweep 数の関数としての「時間発展」を比較する. 図 2 に温度  $T = 50$  K における従来のカノニカル MC シミュレーション(破線)と拡張アンサンブルシミュレーション(実線)の結果を示した. カノニカルシミュレーションは, エネルギー値が  $E = -7$  kcal/mol の近傍に留まっているが, この温度における平均値は  $\langle E \rangle_T = -11.1$  kcal/mol であることが知られているの

\* 1 ここで, 1 Monte Carlo sweep はすべての回転角(二面角)を 1 回ずつ更新することに対応する. そして, 更新のたびごとにメトロポリス判定を課す. よって, 回転角が  $N$  個あるとき, 1 MC sweep は  $N$  回のメトロポリス判定に対応する.

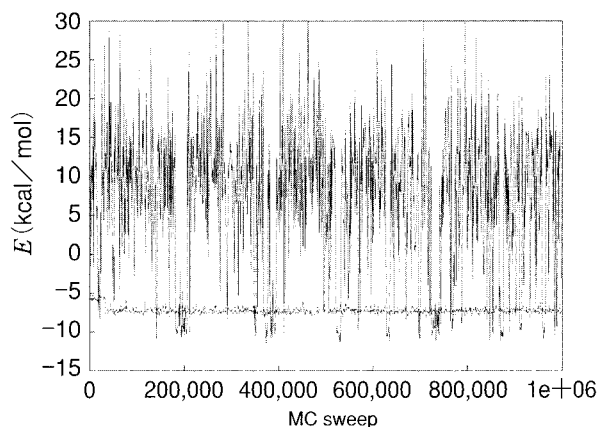


図 2 全ポテンシャルエネルギー  $E_{tot}$  (kcal/mol) の温度  $T = 50$  K における従来のカノニカルシミュレーション(破線)と拡張アンサンブルシミュレーション(実線)における時間発展

で<sup>7)</sup>, エネルギー極小状態に留まってしまっていることがわかる. 一方, 拡張アンサンブルシミュレーションは非常に幅広いエネルギー領域をカバーしており, エネルギー空間の酔歩を実現して, エネルギー極小状態に留まるのを避けているのがよくわかる. 実際, このシミュレーションでは, 1,000,000 MC sweep の間に最小エネルギー領域( $E \leq -11$  kcal/mol)に何度も到達している.

2章で詳しく述べたように, 拡張アンサンブル法では, 再重法を使って, ただ1回のシミュレーションの結果から, いろいろな熱力学量を温度の関数として求めることができる. その例として, Met-enkephalin の系の平均ポテンシャルエネルギーと比熱を温度の関数として表したものを図3に示す. ここで, 温度  $T$  における比熱  $C(T)$  は次式で定義される.

$$C \equiv \frac{1}{Nk_B} \frac{d\langle E_{\text{tot}} \rangle_T}{dT} = \beta^2 \frac{\langle E_{\text{tot}}^2 \rangle_T - \langle E_{\text{tot}} \rangle_T^2}{N} \quad (2)$$

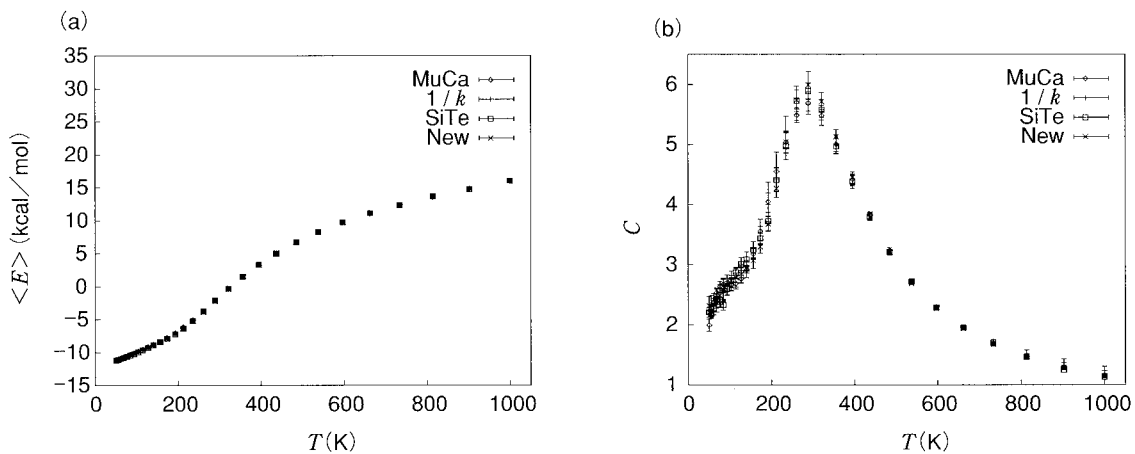


図3 Met-enkephalin の全構造エネルギー  $E_{\text{tot}}$  の, (a) 平均値(kcal/mol), (b) 比熱  $C$  を温度  $T$  (K) の関数として表したもの  
これらは四つの異なる拡張アンサンブル法の結果を重ね合わせたものである.

ここで,  $k_B$  はボルツマン定数,  $\beta = 1/k_B T$ ,  $N (= 5)$  はこのペプチドのアミノ酸数である. 図3は四つの拡張アンサンブル法の結果<sup>8,9)</sup>を重ね書きしたものであるが, よい一致が得られている.

図4には構造エネルギーの各成分の平均値を温度  $T$  (K) の関数として表したものを示す[式(1)参照]. この図からわかることは, 各項の中では, LJ項が温度とともに最も大きく変化することであり, 気相中の Met-enkephalin においては, このエネルギー項が折り畳みに一番効いていることである. これは, ペプチドが低温でコンパクトになるとき, ファンデルワールス接触が大幅に増えることに起因する.

タンパク質の折り畳みでは, 最低三つの段階があると考えられる. 一つの段階はランダムコイル状態からコンパクトなグロビュール状態に移行する段階である. この相転移は高分子の分野で広く知られる, コイル・グロビュール転移温度  $T_\theta$

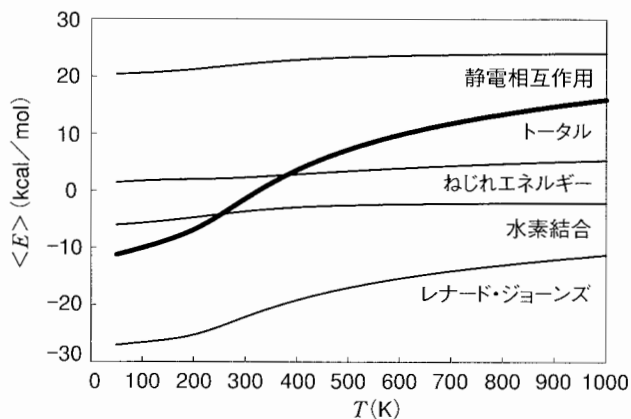


図4 Met-enkephalinの構造エネルギーの各成分の  
平均値(kcal/mol)

で特徴づけられる。二つ目の段階は、コンパクトな構造の中から自然の立体構造(最小自由エネルギー構造)に折り畳まれる段階である。コンパクトな構造がいろいろと探索され、最終的に自然の立体構造へと転移する。この転移は「折り畳み温度」 $T_f$ で特徴づけられる。そして、三つ目の段階は、低温のため熱的ゆらぎが小さく、系がエネルギー極小状態に長く留まってしまいう段階である。この転移はガラス転移温度  $T_g$  で特徴づけられる。

一般に、 $T_f, T_g \leq T_0$  が成り立つ。 $T_f$  と  $T_g$  については、 $T_f > T_g$  ではタンパク質は自然の立体構造へ折り畳むことが保証されるが、反対に  $T_f \leq T_g$  では唯一の立体構造に折り畳むことはなく、スピングラス理論でいうところの「フラストレーション」の多い「ガラス状態」になってしまう。よって、二つの温度の比  $T_g/T_f$  が系のフラストレーションを表す量であり、自然のタンパク質では、

$$\frac{T_g}{T_f} < 1 \quad (3)$$

になっていると主張された<sup>10)</sup>。これは、「最小フラストレーションの原理(principle of minimal frustration)」とよばれる<sup>11)</sup>。なお、この原理は現在では、文献12の「整合性原理(consistency principle)」の再発見と位置づけられている。

最小フラストレーションの原理が成り立っている系の自由エネルギー地形は、自然の立体構造へ向かうような傾きをもった、漏斗(funnel)状をしている<sup>11)</sup>。また、以下の関係が近似的に導かれている<sup>13)</sup>。

$$\frac{T_g}{T_f} \propto \frac{\Delta E}{\delta E} \quad (4)$$

ここで、 $\Delta E$  は変性状態における特徴的なエネルギーゆらぎ、 $\delta E$  は最小エネルギー状態とほかの準安定な状態とのエネルギー差である。よって、 $T_g/T_f$  を最小にする最小フラストレーション原理は  $\Delta E/\delta E$  を最小にすることに対応するといってもよい。すなわち、最小エネルギー構造へ移行しようとする「力」がエネルギー地形の「でこぼこ」に比べて大きいとき ( $\delta E > \Delta E$ )、タンパク質は折り畳むことができるわけである。

タンパク質の折り畳みやすさを表す量として、次の量を考察する場合もある<sup>14)</sup>.

$$\sigma = \frac{T_\theta - T_f}{T_\theta} \quad (5)$$

この $\sigma$ の値が小さければ小さいほど、タンパク質は折り畳みやすい<sup>14)</sup>. すなわち、この条件は $T_f$ が $T_\theta$ と値が近いことを意味し、タンパク質がコンパクトなグロビュールになってから、それほど温度を下げなくても自然の立体構造に折り畳んでくれることを意味するからである. とくに、 $T_f \approx T_\theta$ の場合、タンパク質の折り畳みは二状態転移であるという.

これらの温度の計算の例として、筆者らは気相中の Met-enkephalin において、文献 8 の Tsallis 統計に基づく新しい拡張アンサンブル法による MC シミュレーションを実行した<sup>15)</sup>.

とくに、体積とオーバーラップが計算された. ここで、体積は、各原子をそのファンデルワールス半径の球としたときの溶媒排除体積(単位は $\text{\AA}^3$ )で表したが、NSOL<sup>16)</sup>というプログラムを使って求めた. オーバーラップは、スピングラス理論でよく使われるが、ここでは、問題とする立体構造が基準となる立体構造(R)にどれぐらい似ているかを表す量として、以下で定義した.

$$O_R = 1 - \frac{1}{90m_F} \sum_{i=1}^{m_F} |\alpha_i - \alpha_i^{(R)}| \quad (6)$$

ここで、 $\alpha_i$ と $\alpha_i^{(R)}$  ( $i = 1, \dots, m_F$ )は、それぞれ問題とする立体構造と基準となる立体構造の二面角である. 差 $\alpha_i - \alpha_i^{(R)}$ はつねに $(-180^\circ, 180^\circ)$ の間に入るように射影した. この定義によれば、次式が成り立つ.

$$0 \leq \langle O \rangle_T \leq 1 \quad (7)$$

温度 $T$ における平均オーバーラップ $\langle O \rangle_T$ は、温度があがるにつれて、 $T \rightarrow \infty$ の極限值0にゆっくり近づく. しかし、たとえば、温度 $T = 1000 \text{ K}$ においても、最小エネルギー構造(構造A)とのオーバーラップ $O_A$ の平均値は依然 $\langle O_A \rangle \approx 0.3$ とかなり大きい. これは、 $\langle O \rangle_T = 0$ が二面角の完全にランダムな分布に対応するが、このような条件は主鎖および側鎖の立体障害のために、エネルギー的に実現が困難なためである. また、温度が $T \rightarrow 0 \text{ K}$ の極限で、極限值 $O_A \rightarrow 1$ が得られる. すなわち、 $T = 0 \text{ K}$ では最小エネルギー構造のみが許される.

最初の転移は伸びたランダムコイル構造からコンパクトな構造への転移であるから、転移温度 $T_\theta$ において平均体積が急速に変化すると期待される. よって、 $T_\theta$ は次式の微分の最大値をとる温度として求めることができる.

$$\frac{d\langle V \rangle_T}{dT} \equiv \beta^2 (\langle VE_{\text{tot}} \rangle_T - \langle V \rangle_T \langle E_{\text{tot}} \rangle_T) \quad (8)$$

平均体積とその微分を温度の関数として図5に示した. これより、次式のコイル・グロビュール転移温度を得た<sup>15)</sup>.

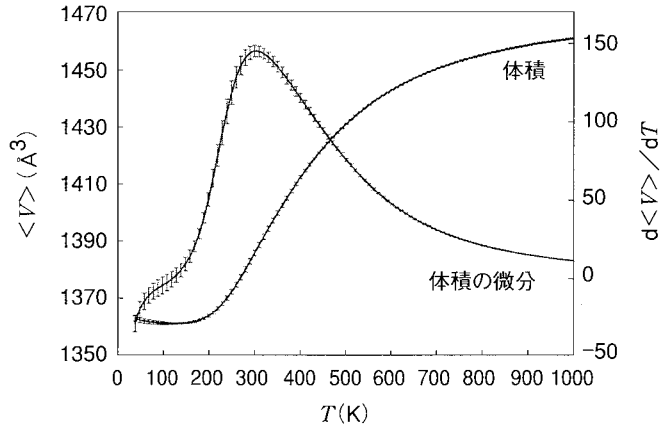


図5 Met-enkephalin の平均体積  $\langle V \rangle_T$  ( $\text{\AA}^3$ ) とその温度  $T$  (K) による微分

$$T_\theta = 310 \pm 20 \text{ K} \tag{9}$$

実は、図3 (b)からもわかるように、この温度は比熱の最大値を取る温度に対応している。この量からは、 $T_\theta = 280 \pm 20 \text{ K}$  が得られ、上の値と誤差の範囲内で一致している。

二つ目の転移はもう少し低い温度  $T_i$  で起こる。この転移は最小エネルギー構造との平均オーバーラップ  $\langle O_A \rangle_T$  の急速な変化によって特徴づけられる[式(6)を参照]。よって、次式の微分の絶対値の最大値をとる温度として求めることができる。

$$\frac{d\langle O_A \rangle_T}{dT} \equiv \beta^2 (\langle O_A E_{\text{tot}} \rangle_T - \langle O_A \rangle_T \langle E_{\text{tot}} \rangle_T) \tag{10}$$

平均オーバーラップ  $\langle O_A \rangle_T$  とその微分を温度の関数として、図6に示した。これより、次式の折り畳み温度を得た<sup>15)</sup>。

$$T_i = 230 \pm 30 \text{ K} \tag{11}$$

ガラス転移温度  $T_g$  の計算は上の二つの転移温度ほど容易ではないが、筆者らは以下を示した<sup>17)</sup>。

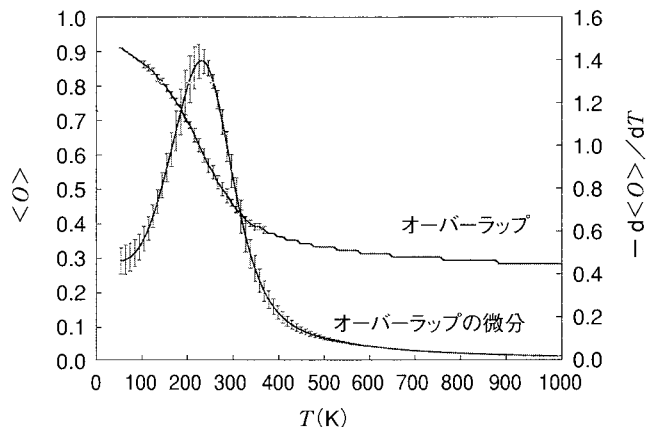


図6 Met-enkephalin の最小エネルギー構造との平均オーバーラップ  $\langle O_A \rangle_T$  とその温度  $T$  (K) による微分

$$T_g \leq 150 \text{ K} \quad (12)$$

よって、Met-enkephalin は式(3)の最小フラストレーションの条件を満たす。また、式(5)の量は

$$\sigma \approx 0.2 \quad (13)$$

となり、Met-enkephalin が折り畳みやすいことを示唆している。

筆者らはさらに、うへの拡張アンサンブルシミュレーションの結果を使って、直接 Met-enkephalin の自由エネルギー地形を求めた<sup>17)</sup>。このペプチドの立体構造のいろいろな温度における分布は詳しく調べられている<sup>7)</sup>。そして、低温ではとくに二つの立体構造グループが存在することが知られている。図7にこの二つの立体構造を示した。構造Aは最小エネルギー構造であり、Gly-2とMet-5の間の

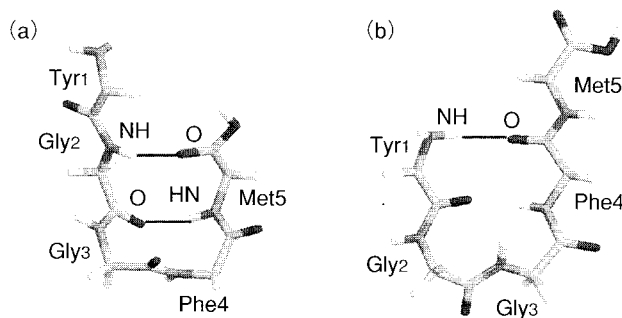


図7 Met-enkephalin の気相中における、(a) 最小エネルギー構造(構造A:  $E = -12.2 \text{ kcal/mol}$ ), (b) 二番目にエネルギーが低い構造(構造B:  $E = -11.0 \text{ kcal/mol}$ )の主鎖の構造

水素結合で安定化されている。また、構造Bは二番目にエネルギーが低い構造であり、Tyr-1とPhe-4の水素結合で安定化されている。

自由エネルギー  $G$  をいくつかの適当な「反応座標 (reaction coordinate)」(オーダーパラメータ)の関数として表すことにする。一つの可能性は  $G(V, O_A)$  である。すなわち、体積  $V$  と最小エネルギー構造(構造A)とのオーバーラップ  $O_A$  の関数として表すことである。もう一つの可能性は  $G(O_A, O_B)$  であり、構造Aと構造Bとのオーバーラップ  $O_A$  および  $O_B$  の関数として表すことである。これらは以下の式で与えられる。

$$G(V, O_A) = -k_B T \ln P(V, O_A) \quad (14)$$

$$G(O_A, O_B) = -k_B T \ln P(O_A, O_B) \quad (15)$$

ここで、 $P(V, O_A)$  と  $P(O_A, O_B)$  は、このペプチドの立体構造がそれぞれ  $(V, O_A)$  と  $(O_A, O_B)$  の値をとる確率であり、拡張アンサンブルシミュレーションで得られる立体構造の分布から、ある特定の温度  $T$  における値を再重法によって求めることができる。とくに、温度  $T = 1000 \text{ K}$ ,  $T = T_\theta = 300 \text{ K}$ ,  $T = T_i = 230 \text{ K}$ ,  $T =$



150 Kにおける結果を図8と図9に示した<sup>17)</sup>.

等高線は $k_B T$ の単位ごとに書き込んだ(よって、温度によって数値が違うが、折り畳み機構を探るにはよい単位である). 図8(a)は $G(V, O_A)$ の高温( $T = 1000$  K)における結果である. 自由エネルギーの最小値は、大きな体積( $\approx 1470$  Å<sup>3</sup>)で小さなオーバーラップ( $O_A \approx 0.3$ )のところにある. そして、体積が小さく、オーバーラップが大きい領域は $k_B T$ で何十倍も自由エネルギーが大きくなっている. よって、この温度では、伸びた(体積の大きい)ランダムコイル構造が安定である. 図8(b)からわかるように、この様子はコイル・グロビュール転移温度 $T_\theta$ では大きく変化する. この温度では、 $V, O_A$ 空間の非常に幅広い領域が到達可能である. 等高線は体積が小さいところも大きいところもカバーしており、ほとんどすべての $O_A$ の値において、自由エネルギーが $2 k_B T$ の値の差しかない. これは、この温度では伸びた構造とコンパクトな構造の間には熱力学的障壁が小さく、ペプチドが伸びたり縮んだりしていることを示している. さらに、図8(c)に示されているように、温度を折り畳み温度 $T_f = 230$  Kまで下げると、自由エネルギー地形は漏斗状になる<sup>11)</sup>. この温度では最小エネルギー構造へ折り畳もうとする傾向が強くなり、ほかのエネルギー極小状態にトラップされる傾向が少なくなる. 最小エネルギー構造( $O_A = 1$ )への傾斜が明らかであるが、同じぐらいの体積のほかの構造(たとえば、 $O_A \approx 0.5$ )との自由エネルギー障壁は $1 k_B T$ のオーダーしかない. この温度より低くなると、 $T = 150$  Kに対応する図8(d)にあるように、最小エネルギー状態がほかの低エネルギー状態とはっきりと区別されてくる. 自由エネルギー障壁は $k_B T$ の何倍にもなってくる. 実際、最小エネルギー状態はほかのエネルギー極小状態より、自由エネルギーで $3 k_B T$ も低く、その障壁は $2 k_B T$ ぐらいである.

同様の振舞いが図9の $G(O_A, O_B)$ にも示されている. 図9(a)は高温( $T = 1000$  K)における結果であるが、自由エネルギーが小さいところは、両方のオーバーラップ値も小さく、この温度ではランダムコイル構造が安定であることを示唆している. コイル・グロビュール転移温度 $T_\theta = 300$  Kでは、図9(b)にあるように、 $O_A$ と $O_B$ の幅広い空間が、自由エネルギーで $2 k_B T$ 以内にあり、いろいろな立体構造が許されることがわかる. 図9(c)にあるように、折り畳み温度 $T_f = 230$  Kでは、最小エネルギー構造への傾斜をもった漏斗状の自由エネルギー地形が得られる. しかし、ほかのいろいろな構造も互いに $1 k_B T$ 以下の自由エネルギー障壁で隔てられているとともに、最小エネルギー構造からも自由エネルギーで大きく隔たっているわけではない. たとえば、構造Bに対応する領域( $O_B \approx 1$ )は最小エネルギー構造より $3 k_B T$ だけ高い. よって、この温度では漏斗状の自由エネルギー地形は滑らかであり、いろいろな経路を経て最小エネルギー構造へ折り畳むことができる. 実際に $T = 230$  Kにおける従来のカノニカルMCシミュレーションを実行したところ、伸びた構造から直接最小エネルギー構造へ折り畳む場合や、構造Bの領域に立ち寄ったあと、最小エネルギー構造に折り畳む場合などが確認できた<sup>17)</sup>. 図9(d)では低温( $T = 150$  K)での結果が示されているが、ガラス的振舞いが顕著になっている. すなわち、最小エネルギー構造はふたたび漏斗

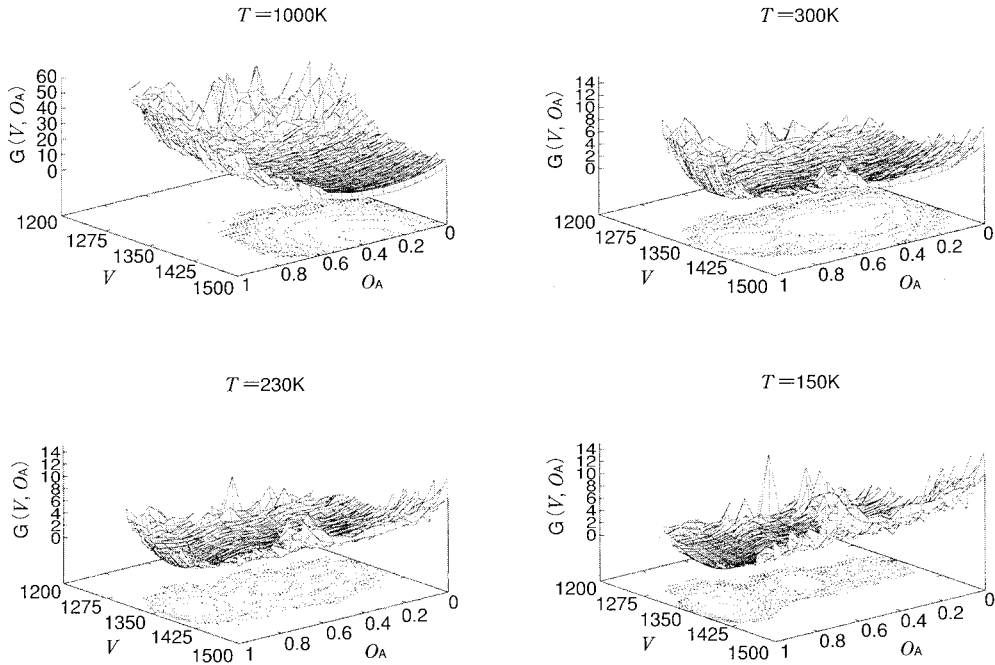


図8 体積  $V$  ( $\text{\AA}^3$ ) および構造 A とのオーバーラップ  $O_A$  の関数としての自由エネルギー  $G(V, O_A)$  (kcal/mol)  
 対応する温度は, (a)  $T = 1000$  K, (b)  $T = 300$  K, (c)  $T = 230$  K,  
 (d)  $T = 150$  K.

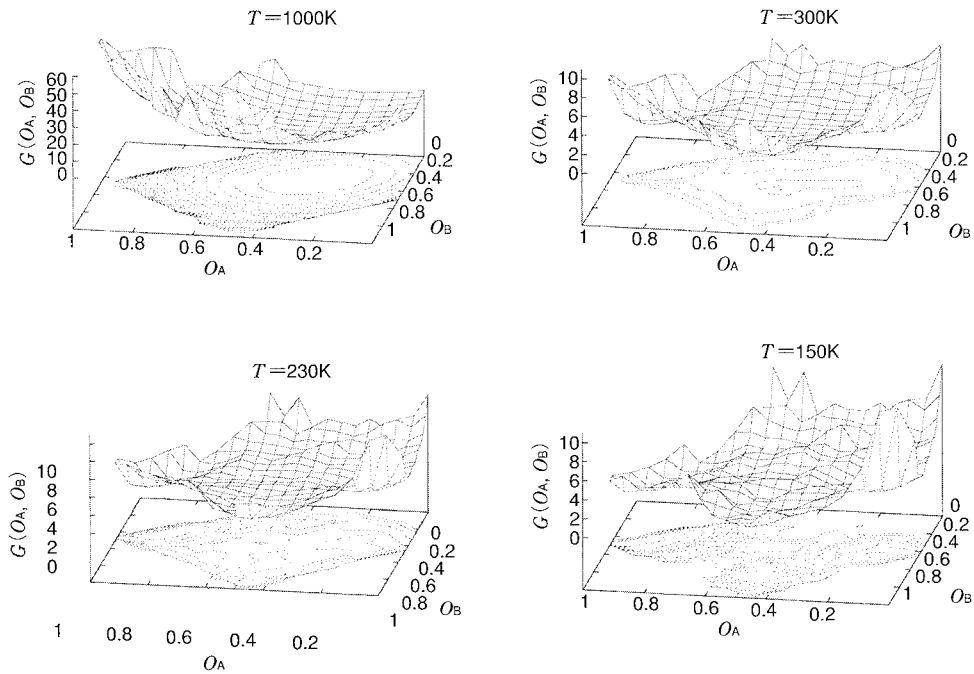


図9 構造 A とのオーバーラップ  $O_A$  および構造 B とのオーバーラップ  $O_B$  の関数としての自由エネルギー  $G(O_A, O_B)$  (kcal/mol)  
 対応する温度は, (a)  $T = 1000$  K, (b)  $T = 300$  K, (c)  $T = 230$  K,  
 (d)  $T = 150$  K.

状の地形をしているが、すでに滑らかではなく、でこぼこ状態になっている。エネルギー極小状態を隔てる障壁は  $k_B T$  の何倍にもなっており、この温度における折り畳みを非常に困難にしていることがわかる。

ここまでは、気相(真空)中のシミュレーションの結果を述べてきた。しかし、タンパク質は普通水溶液中に存在し、実験もそのような環境で行われる。よって、厳密な溶媒効果をシミュレーションに取り入れる必要がある。ここでは、二つの拡張アンサンブルシミュレーションの結果を紹介する。最初の例は RISM 理論<sup>18)</sup>によって溶媒効果を取り入れたマルチカノニカル MC シミュレーションである<sup>19)</sup>。図10に温度の関数としての、このペプチドの平均末端間距離を気相中と水中につ

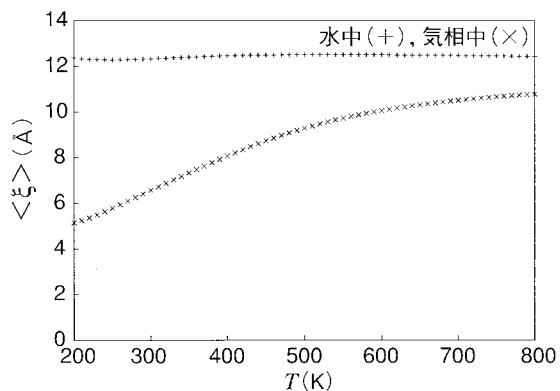


図 10 マルチカノニカル MC シミュレーションで得られた温度の関数としての気相中(x)および水中(+)  
Met-enkephalin の平均末端間距離

いて示す。気相中では低温でコンパクトな構造が温度があがるにつれて、ペプチドがほどけて伸びてゆくのに対し、水中ではすべての温度で伸びている(約12 Åの長さ)ことがわかる。ちなみに、NMR 実験の結果も主鎖が完全に伸びた構造を示唆している<sup>20)</sup>。

同じペプチドが水分子をあらわに取り入れたレプリカ交換 MD シミュレーションでも調べられた<sup>21,22)</sup>。温度は200 K から500 K の間に36個用意された。36個の

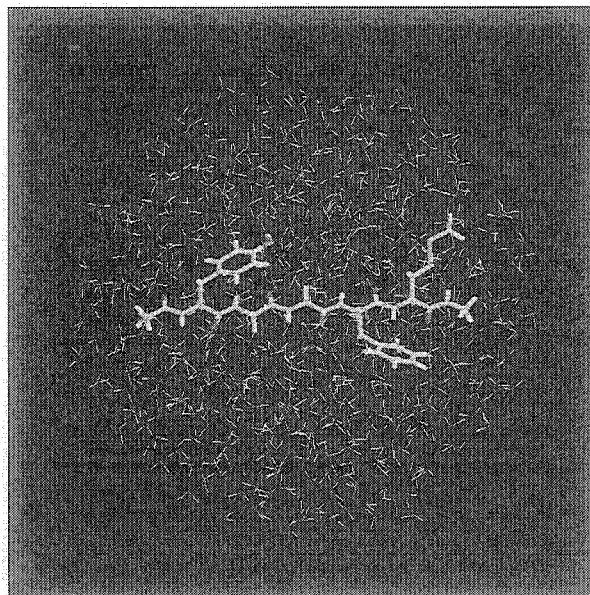


図 11 水中(526個の水分子)の Met-enkephalin のレプリカ交換 MD シミュレーションの初期配置(口絵④参照)

それぞれのレプリカでは、ペプチド1分子と526個の水分子が半径16 Åの球内に置かれた(図11参照).

このレプリカ交換 MD シミュレーションから得られた、温度の関数としての平均末端間距離を気相中の結果とともに図12に示す. 図10と同じように、このペプチドの気相中の低エネルギー構造はコンパクトでターン構造をしているが、水中では伸びていることが示された.

図13では、レプリカ交換 MD シミュレーションで得られた、主鎖の二面角対( $\phi, \psi$ )の温度  $T = 300$  K における分布を気相中と水中の場合で比較した<sup>21)</sup>. 気相中では分布が局在化しているが、水中では分布が広がっていることがわかる. すなわち、水は系の自由エネルギー地形を滑らかにしている.

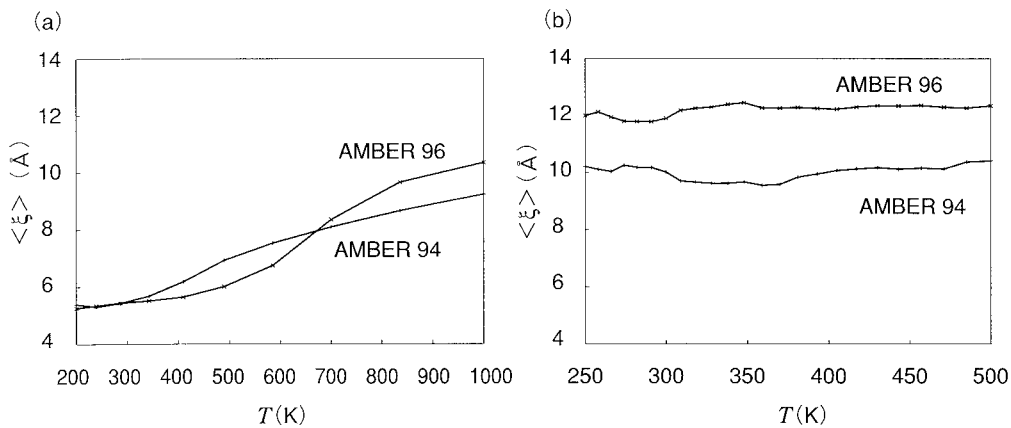


図 12 レプリカ交換 MD シミュレーションで得られた温度の関数としての気相中(a) および水中(b) の Met-enkephalin の平均末端間距離

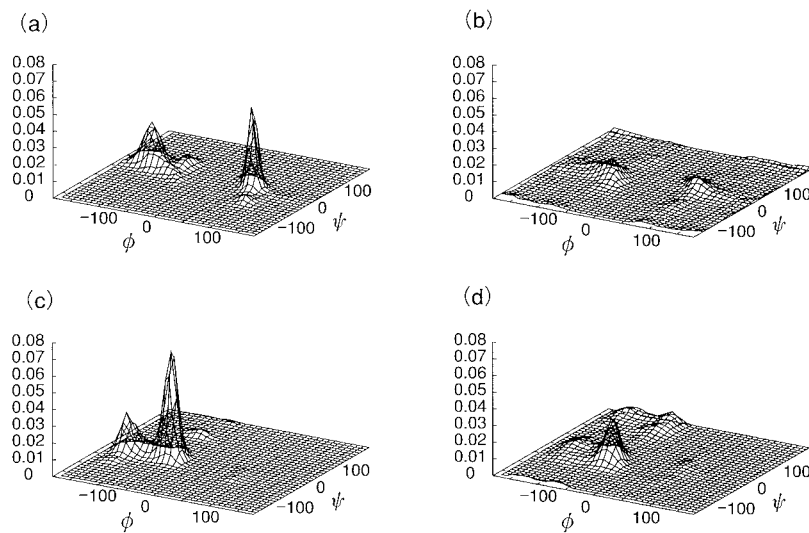


図 13 温度  $T = 300$  K における Met-enkephalin の主鎖の二面角対( $\phi, \psi$ )の分布 (a) 気相中の二番目のアミノ酸の場合, (b) 水中の二番目のアミノ酸の場合, (c) 気相中の四番目のアミノ酸の場合, (d) 水中の四番目のアミノ酸の場合.

### 3. おわりに

本章では、全原子模型のタンパク質の分子シミュレーションの結果を紹介し、タンパク質の折り畳み問題を考察した。このような問題では、とくに自由エネルギーの計算が重要である。タンパク質の系の自由エネルギーは高次元の「反応座標」の関数として表される。本章では、「反応座標」として、タンパク質の体積と二つの基準となる立体構造とのオーバーラップを使用した。ほかにもいろいろな反応座標が考えられる。たとえば、体積に代わるものとして、慣性半径(radius of gyration)や末端間距離(end-to-end distance)などがある。そして、自然の立体構造とのオーバーラップに代わるものとして、自然の立体構造の座標からの根平均二乗距離(root mean square distance)や自然の立体構造に存在する接触部分(native contact)の数などが考えられる。また、 $\alpha$ ヘリックスや $\beta$ シートなどの二次構造の情報も反応座標となりうる。さらには、主成分解析(principal component analysis)<sup>23)</sup>の固有座標も反応座標として広く使われている。こういうさまざまな反応座標の関数として自由エネルギー地形を多角的に考察することによって、タンパク質の折り畳み問題の全体像が見えてくるのである。

最後に、タンパク質の折り畳み問題における水の役割について、考察してみよう。文献24で、筆者らは水の存在がLevinthalのパラドックスを解決するという推測を得た。ここでは、この推測をさらに発展させて、以下のように言い換える。まず、前節で見たように、タンパク質は分子内のLJ相互作用などの寄与で、真空中でもコンパクトに折り畳まれる。真空中では、普通のサイズ(アミノ酸が300程度)のタンパク質では、エネルギー地形は高いエネルギー障壁をもつため、ガラス状態に近い。一方、水は、このコンパクトになろうとするタンパク質本来の傾向に逆らって、むしろタンパク質を不安定にして、引き伸ばそうとする。小ペプチドでは、この水の効果が勝って、伸びたコイル状態になってしまう。しかし、アミノ酸数がある程度の数まで増えてくると、タンパク質本来のグロービュール状態になる傾向が水の伸ばそうとする効果に打ち勝って、タンパク質はコンパクトに折り畳まれる。このとき、水とタンパク質の間の相互作用が強く働いて、タンパク質の親水的な部分は外側に、疎水的な部分は内側に配置される。さらには、水の存在はタンパク質の自由エネルギー地形のでこぼこを減らし滑らかにする。よって、水の存在によって、タンパク質の系の自由エネルギー地形が唯一の立体構造(自然の立体構造)へと傾きをもった漏斗状になる。

以上のようなことを普通の大きさのタンパク質において、水分子を含めた分子シミュレーションを実行することによって定量的に示す必要がある。そのためには、自由エネルギーを温度の関数として求める必要があり、それを1回のシミュレーションの結果から精度よく求めることができる、拡張アンサンブル法の適用がとくに重要であると考えられる。

#### 参考文献

- (1) C. J. Epstein, R. F. Goldberger, C. B. Anfinsen, *Cold Spring Harbor Symp. Quant. Biol.*, **28**, 439

**Levinthalのパラドックス**  
 一個のタンパク質が取りうる立体構造の数を考慮するとき、アミノ酸1個当り、主鎖だけでも最低約10個の準安定な立体構造が考えられる。すると、 $N$ 個のアミノ酸からなるタンパク質全体では、最低10の $N$ 乗個の立体構造を考慮しなければならない。この個数は $N$ とともに急激に増大して、数十個のアミノ酸からなるタンパク質においてさえ、タンパク質分子が一つ一つ可能な立体構造を網羅するのに、宇宙の年齢よりも長い時間が要求されることになってしまう。それでは、なぜ、自然のタンパク質は決った立体構造にミリ秒から分のオーダーのタイムスケールで折り畳まれるのか、という疑問がLevinthalのパラドックスとよばれる。

- (1963).
- (2) a) F. A. Momany, R. F. McGuire, A. W. Burgess, H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).  
b) G. Némethy, M. S. Pottle, H. A. Scheraga, *J. Phys. Chem.*, **87**, 1883(1983).  
c) M. J. Sippl, G. Némethy, H. A. Scheraga, *J. Phys. Chem.*, **88**, 6231(1984).
- (3) a) H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, T. Kikuchi, *Chem. Lett.*, **1991**, 213.  
b) Y. Okamoto, M. Fukugita, T. Nakazawa, H. Kawai, *Protein Eng.*, **4**, 639(1991).
- (4) a) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179(1995).  
b) P. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, A. Pohorille, in "Computer Simulation of Biomolecular Systems," Vol. 3, ed. by W. F. van Gunsteren, P. K. Weiner, A. J. Wilkinson, KLUWER/ESCOM (1997), p. 83.
- (5) a) Y. Sugita, A. Kitao, *Proteins*, **30**, 388(1998).  
b) A. Kitao, S. Hayward, N. Gō, *Proteins*, **33**, 496(1998).
- (6) K. Morikami, T. Nakai, A. Kidera, M. Saito, H. Nakamura, *Comput. Chem.*, **16**, 243(1992).
- (7) A. Mitsutake, U. H. E. Hansmann, Y. Okamoto, *J. Mol. Graphics Mod.*, **16**, 226(1998).
- (8) U. H. E. Hansmann, Y. Okamoto, *Phys. Rev. E*, **56**, 2228(1997).
- (9) U. H. E. Hansmann, Y. Okamoto, *J. Comput. Chem.*, **18**, 920(1997).
- (10) J. D. Bryngelson, P. G. Wolynes, *Proc. Natl. Acad. Sci. USA*, **84**, 7524(1987).
- (11) J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, *PROTEINS: Struct. Funct. Genet.*, **21**, 167(1995).
- (12) N. Gō, *Ann. Rev. Biophys. Bioeng.*, **12**, 183(1983).
- (13) R. A. Goldstein, Z. A. Luthey-Schulten, P. G. Wolynes, *Proc. Natl. Acad. Sci. USA*, **89**, 4918 (1992).
- (14) C. J. Camacho, D. Thirumalai, *Proc. Natl. Acad. Sci. USA*, **90**, 6369(1993).
- (15) U. H. E. Hansmann, M. Masuya, Y. Okamoto, *Proc. Natl. Acad. Sci. USA*, **94**, 10652(1997).
- (16) M. Masuya, <http://mio.cc.kagoshima-u.ac.jp/masatom/>.
- (17) U. H. E. Hansmann, Y. Okamoto, J. N. Onuchic, *PROTEINS: Struct. Funct. Genet.*, **34**, 472 (1999).
- (18) a) D. Chandler, H. C. Andersen, *J. Chem. Phys.*, **57**, 1930(1972).  
b) F. Hirata, P. J. Rossky, *Chem. Phys. Lett.*, **83**, 329(1981).
- (19) A. Mitsutake, M. Kinoshita, Y. Okamoto, F. Hirata, *Chem. Phys. Lett.*, **329**, 295(2000).
- (20) W. H. Graham, S. E. Carter, II, P. R. Hickes, *Biopolymers*, **32**, 1755(1992).
- (21) Y. Sugita, Y. Okamoto, in preparation.
- (22) A. Mitsutake, Y. Sugita, Y. Okamoto, *Biopolymers(Pept. Sci.)*, **60**, 96(2001).
- (23) a) A. Kitao, F. Hirata, N. Gō, *Chem. Phys.*, **158**, 447(1991).  
b) S. Hayward, A. Kitao, F. Hirata, N. Gō, *J. Mol. Biol.*, **234**, 1207(1993).  
c) R. Abagyan, P. Argos, *J. Mol. Biol.*, **225**, 519(1992).  
d) A. E. Garcia, *Phys. Rev. Lett.*, **68**, 2696(1992).
- (24) M. Kinoshita, Y. Okamoto, F. Hirata, *J. Am. Chem. Soc.*, **120**, 1855(1998).