

岡本祐幸, 「拡張アンサンブル法」, 超多自由度系の最適化, 「計算科学講座」第9巻第2章, pp. 119-241 (金田行雄, 笹井理生 (監修), 古橋武, 笹井理生 (編), 共立出版, 2013).

第1章

拡張アンサンブルの方法

前章では, 進化的計算法による最適化法が議論された. ここでも, 進化的計算法と同じように, 自然界の「知恵」を使うのであるが, 具体的には, 自然界の物理法則, 特に統計力学の法則に基づく最適化法について議論する. そこでは, 統計力学的集団 (統計力学的アンサンブル) や温度という概念が利用される.

生体系などの多自由度複雑系では, エネルギー極小状態が無数に存在するため, 温度一定のカノニカルアンサンブル上の従来のモンテカルロ法 (Monte Carlo: MC) や分子動力学法 (molecular dynamics: MD) による計算機シミュレーションは, それらのエネルギー極小状態に留まってしまうという難点がある. すなわち, シミュレーションの結果が初期状態に強く依存してしまうので, 計算結果の信頼性が極端に低くなってしまうのである.

本章では, 拡張アンサンブル法 (generalized-ensemble algorithm) と総称される手法を解説する. このシミュレーション法は, 非ボルツマン因子による人工の統計力学的集団に基づいており, ポテンシャルエネルギー空間や他のパラメータ空間上の酔歩 (random walk) を実現することで, エネルギー極小状態に留まるのを避ける. そして, 従来の手法と比べて格段に幅広い状態空間を探索することを可能とするのである. 拡張アンサンブル法では, 唯一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, 様々な物理量の任意の温度や系を指定する他のパラメータ値における期待値を得ることができるという特徴がある. 本章では, 特に広く使われている三つの拡張アンサンブル法 (すなわち, マルチカノニカル法, 焼き戻し法, レプリカ交換法) のモンテカルロ法および分子動力学法に基づくシミュレーション手法とその更なる拡張版について解説する.

2 第 1 章 拡張アンサンブルの方法

1.1 統計力学の基礎

統計力学で扱う系は物質（原子）の集団である．この集団がどのような環境（外的条件）のもとにあるかによって，系の振る舞いが規定されている．個々の統計力学的集団（統計力学的アンサンブル）とは，それらの環境を指定した物質の集団を表している．ここでは，モンテカルロ法や分子動力学法の理解に必要な統計力学の基礎を復習しておこう．

1.1.1 統計力学的集団

ここでは，よく知られている統計力学的集団（statistical mechanical ensemble）の例をあげよう．まず，ミクロカノニカルアンサンブル（microcanonical ensemble）は，孤立系を表す統計力学的集団である．すなわち，系が外界と熱など一切のやり取りをしない孤立系である．よって，系の粒子数，体積，全エネルギー（運動エネルギーとポテンシャルエネルギーの和）が時間とともに保存される．例えば，熱を一切通さない理想的な魔法瓶の中のお湯全体の系がミクロカノニカルアンサンブルである．図 1.1 にミクロカノニカルアンサンブルの概念図を示す．

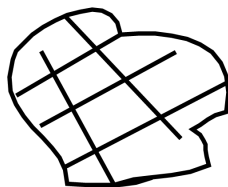


図 1.1 ミクロカノニカルアンサンブル．孤立系である．

次に，カノニカルアンサンブル（canonical ensemble）は，温度一定の熱浴中の小さい系を表す統計力学的集団である．ここで，「小さい」とは，熱浴と比較して小さいという意味である．熱浴を構成する物質質量が今問題とする系の物質質量よりはるかに大きいので，熱浴から系への影響は大きいとその逆はほとんど無視できるということである．特に，系と熱浴は物質のやり取りをしない（よって，系の粒子数は保存される）が，熱のやり取りはするので，系の温度が

1.1 統計力学の基礎 3

熱浴と同じ一定温度に保たれる。また、系の体積も一定に保たれる。カノニカルアンサンブルでは、粒子数 (N)・体積 (V)・温度 (T) が一定なので、 NVT アンサンブルとも呼ばれる。ミクロカノニカルアンサンブルとの関係は、この系 (カノニカルアンサンブル) と熱浴を足したものがミクロカノニカルアンサンブルになっている。図 1.2 にカノニカルアンサンブルの概念図を示す。

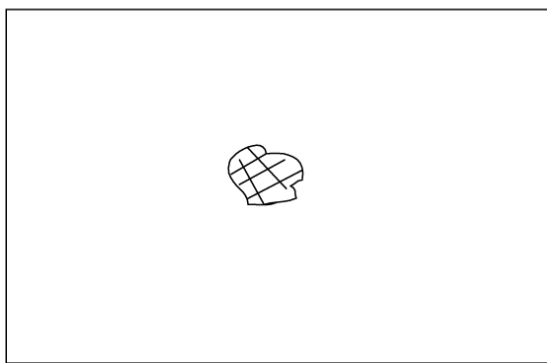


図 1.2 カノニカルアンサンブル。温度 T の熱浴の中に系がある。

自然の統計力学的集団には、他にもいろいろある。例えば、系が外界と熱ばかりでなく、粒子をやり取りするものが、グランドカノニカルアンサンブル (grandcanonical ensemble) である。また、系の圧力が熱浴と同じ圧力になるように、体積が変化するものが、定圧定温アンサンブル (isobaric-isothermal ensemble) である。これは、系の境界の一面がピストンにつながっているというようにイメージできる。系の圧力が下がれば、ピストンが系の体積が減るように動き、逆に系の圧力が熱浴より上がれば、ピストンが系の体積を増やすように動くわけである。すなわち、定圧定温アンサンブルでは、粒子数 (N)・圧力 (P)・温度 (T) が保存され、 NPT アンサンブルとも呼ばれる。

これらのいろいろな統計力学的集団を再現するシミュレーション手法は、普通、カノニカルアンサンブルを再現するモンテカルロ法や分子動力学法の手法をそれぞれのアンサンブルに一般化することによって得られる。よって、以下

4 第1章 拡張アンサンブルの方法

では、カノニカルアンサンブルの統計力学とこのアンサンブルを実現する計算機シミュレーション法について述べる。

1.1.2 カノニカルアンサンブルの統計力学

粒子数 (N)・体積 (V)・温度 (T) が一定な系であるカノニカルアンサンブルの系として、質量が m_k ($k = 1, \dots, N$) の N 個の粒子の系を考え、それらの座標と運動量がそれぞれ $q \equiv \{q_1, \dots, q_N\}$ と $p \equiv \{p_1, \dots, p_N\}$ で表されるとする (ここからは、体積は V で表す)。系のハミルトニアン $H(q, p)$ は運動エネルギー $K(p)$ とポテンシャルエネルギー $E(q)$ の和で次のように与えられる。

$$H(q, p) = K(p) + E(q) . \quad (1.1)$$

ここで、運動エネルギーは次式のとおりである。

$$K(p) = \sum_{k=1}^N \frac{p_k^2}{2m_k} . \quad (1.2)$$

温度一定のカノニカルアンサンブルでは、各状態 $x \equiv \{q, p\}$ は以下のボルツマンの重み因子 (Boltzmann weight factor) の確率で実現される (単にボルツマン因子と呼ばれることが多い)。

$$\Omega_B(x) = e^{-\beta H(q, p)} . \quad (1.3)$$

ここで、 β はボルツマン定数 k_B と絶対温度 T の積の逆数であり、以下で与えられる。

$$\beta = \frac{1}{k_B T} . \quad (1.4)$$

このとき、この系の分配関数 (partition function) は以下のように、ボルツマ

ン因子の $6N$ 次元積分で与えられる（前に掛る定数係数を 1 とおいた）.

$$\begin{aligned}
 Z(N; \mathcal{V}; T) &= \int dx \Omega_B(x) \\
 &= \int dq \int dp e^{-\beta H(q,p)} \\
 &= \prod_{k=1}^N \int d^3 \mathbf{q}_k \prod_{k=1}^N \int d^3 \mathbf{p}_k e^{-\beta H} \\
 &= \int_{-\infty}^{\infty} dq_{1x} \int_{-\infty}^{\infty} dq_{1y} \int_{-\infty}^{\infty} dq_{1z} \cdots \\
 &\quad \cdots \int_{-\infty}^{\infty} dp_{Nx} \int_{-\infty}^{\infty} dp_{Ny} \int_{-\infty}^{\infty} dp_{Nz} e^{-\beta H} .
 \end{aligned} \tag{1.5}$$

ここで、カノニカルアンサンブルでは、粒子数 (N)・体積 ($\mathcal{V}$)・温度 (T) が一定なので、それらの値ごとに分配関数が決まることをはっきりさせるために、 $Z(N; \mathcal{V}; T)$ と書いた。しかし、以下では、簡単のために、粒子数 N と体積 $\mathcal{V}$ への依存性は省略して、温度依存性のみ露わに書いて、 $Z(T)$ と書くことにする。系の分配関数 $Z(T)$ が求められれば、温度 T のカノニカルアンサンブルにおける任意の統計力学的物理量が $Z(T)$ から求められるので、この系は「厳密に解けた」と言われる。例えば、ヘルムホルツ自由エネルギー (Helmholtz free energy) $F(T)$ は

$$Z(T) = e^{-\beta F(T)} \tag{1.6}$$

で定義されるので、以下で与えられる。

$$F(T) = -k_B T \log Z(T) . \tag{1.7}$$

ここで、 $\log$ は自然対数関数である（以下、同様）。また、系の内部エネルギー (internal energy) は以下で与えられる。

$$\begin{aligned}
 U(T) &= \langle H \rangle_T \\
 &= \frac{\int dq dp H(q,p) e^{-\beta H(q,p)}}{Z} \\
 &= - \frac{\partial \log Z}{\partial \beta} \\
 &= - \frac{\frac{\partial Z}{\partial \beta}}{Z} .
 \end{aligned} \tag{1.8}$$

6 第1章 拡張アンサンブルの方法

ここで, $\langle A \rangle_T$ は物理量 A の温度 T における平均値 (期待値)

$$\langle A \rangle_T = \frac{\int dq dp A(q, p) e^{-\beta H(q, p)}}{Z}, \quad (1.9)$$

を表す (式 (1.8) では, $A = H$). 更に, 熱力学的関係式

$$F(T) = U(T) - TS(T), \quad (1.10)$$

から, 系のエントロピー (entropy) は以下で与えられる.

$$S(T) = [U(T) - F(T)]/T. \quad (1.11)$$

また, 式 (1.4) と式 (1.8) より, 定積比熱は以下で与えられる.

$$\begin{aligned} C_V(T) &= \frac{dU}{dT} \\ &= -\frac{d\beta}{dT} \frac{\partial}{\partial \beta} \left(\frac{\frac{\partial Z}{\partial \beta}}{Z} \right) \\ &= \frac{1}{k_B T^2} \left[\frac{\frac{\partial^2 Z}{\partial \beta^2}}{Z} - \frac{\left(\frac{\partial Z}{\partial \beta} \right) \left(\frac{\partial Z}{\partial \beta} \right)}{Z^2} \right] \\ &= \frac{1}{k_B T^2} \left\{ \frac{\int dq \int dp (-H)^2 e^{-\beta H}}{Z} - \left[\frac{\int dq \int dp (-H) e^{-\beta H}}{Z} \right]^2 \right\} \\ &= \frac{1}{k_B T^2} [\langle H^2 \rangle_T - \langle H \rangle_T^2]. \end{aligned} \quad (1.12)$$

式 (1.1) では座標 q と運動量 p が分離されているので, 式 (1.5) の分配関数は以下のように「座標分配関数」と「運動量分配関数」の積で表すことができる.

$$Z(T) = Z_q(T) Z_p(T). \quad (1.13)$$

ここで, Z_q と Z_p は以下で定義される.

$$Z_q(T) = \int dq e^{-\beta E(q)}, \quad (1.14)$$

$$Z_p(T) = \int dp e^{-\beta K(p)}. \quad (1.15)$$

1.1 統計力学の基礎 7

特に、物理量 A が q のみの関数、または、 p のみの関数のとき、式 (1.9) は、共通部分が分母と分子で打ち消し合って、それぞれ以下のように書くことができる。

$$\langle A(q) \rangle_T = \frac{\int dq A(q) e^{-\beta E(q)}}{Z_q}, \quad (1.16)$$

$$\langle A(p) \rangle_T = \frac{\int dp A(p) e^{-\beta K(p)}}{Z_p}. \quad (1.17)$$

また、運動量分配関数については、積分を実行できる。すなわち、以下のガウス積分の公式

$$\begin{cases} I_1(\alpha) = \int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}}, \\ I_2(\alpha) = \int_{-\infty}^{\infty} dx x^2 e^{-\alpha x^2} = -\frac{dI_1}{d\alpha} = \frac{1}{2\alpha} \left(\frac{\pi}{\alpha}\right)^{\frac{1}{2}} = \frac{1}{2\alpha} I_1(\alpha), \end{cases} \quad (1.18)$$

の第一式が使える。よって、 Z_p は式 (1.2) と式 (1.15) より以下のように書ける。

$$\begin{aligned} Z_p(T) &= \prod_{k=1}^N \int d^3 \mathbf{p}_k e^{-\beta \frac{\mathbf{p}_k^2}{2m_k}} \\ &= \prod_{k=1}^N \left[I_1 \left(\frac{\beta}{2m_k} \right) \right]^3 \\ &= \left[(2\pi k_B T)^N \left(\prod_{k=1}^N m_k \right) \right]^{\frac{3}{2}}. \end{aligned} \quad (1.19)$$

温度 T における系の平均運動エネルギーは次で与えられる。

$$\begin{aligned} \langle K(p) \rangle_T &= \left\langle \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m_k} \right\rangle_T \\ &= \sum_{k=1}^N \frac{\langle \mathbf{p}_k^2 \rangle_T}{2m_k} \\ &= \sum_{k=1}^N \frac{\langle p_{kx}^2 \rangle_T + \langle p_{ky}^2 \rangle_T + \langle p_{kz}^2 \rangle_T}{2m_k} \\ &= \frac{3}{2} N k_B T. \end{aligned} \quad (1.20)$$

8 第1章 拡張アンサンブルの方法

ここで、式 (1.17), 式 (1.18) および式 (1.19) から、以下が成り立つことを使った。

$$\langle p_{k_x}^2 \rangle_T = \langle p_{k_y}^2 \rangle_T = \langle p_{k_z}^2 \rangle_T = k_B T m_k . \quad (1.21)$$

式 (1.20) は、どんなポテンシャルエネルギー $E(q)$ のときでも成り立つ式であることに注目されたい。

以上の結果から、内部エネルギーは以下のように書ける。

$$\begin{aligned} U(T) &= \langle K(p) \rangle_T + \langle E(q) \rangle_T \\ &= \frac{3}{2} N k_B T + \frac{\int dq E(q) e^{-\beta E(q)}}{Z_q} . \end{aligned} \quad (1.22)$$

この第二項において、式 (1.12) の導出を $H \rightarrow E$, $Z \rightarrow Z_q$ の置き換えをして繰り返すと、定積比熱は以下で与えられる。

$$C_V(T) = \frac{3}{2} N k_B + \frac{1}{k_B T^2} \left[\langle E^2 \rangle_T - \langle E \rangle_T^2 \right] . \quad (1.23)$$

ここで、厳密に解ける系の例として N 個の質量が m の単原子分子からなる理想気体 (ideal gas) を考えよう。この系では各粒子は全く相互作用をしないので、

$$E(q) = 0 \quad (1.24)$$

が成り立つ。すると、系の分配関数は、式 (1.13), (1.14), (1.19) から以下で与えられる。

$$\begin{aligned} Z(T) &= \frac{1}{N! h^{3N}} \left(\int dq 1 \right) \left(\int dp e^{-\beta K(p)} \right) \\ &= \frac{\mathcal{V}^N}{N! h^{3N}} (2\pi m k_B T)^{\frac{3N}{2}} \\ &= \frac{1}{N!} \left[\mathcal{V} \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3}{2}} \right]^N . \end{aligned} \quad (1.25)$$

ここで、統計力学で習ったように、位相空間の単位を指定するプランク定数 h の寄与と同一粒子の区別ができないことによる係数 $N!$ の寄与を係数として掛

けた．この分配関数から，以下のように，いろいろな物理量が計算できる．まず，ヘルムホルツ自由エネルギーは式 (1.7) から次で与えられる．

$$F(T) = -Nk_B T \left[1 + \log \frac{\mathcal{V}}{N} + \frac{3}{2} \log \left(\frac{2\pi m k_B T}{h^2} \right) \right] . \quad (1.26)$$

ここで， N は大きい数であるので，スターリングの公式

$$\log N! = N \log N - N + O(\log N) \quad (1.27)$$

を使って， N に比べて $\log N$ 以下の項を無視した．また，内部エネルギーは式 (1.22) と式 (1.24) から次で与えられる．

$$U(T) = \frac{3}{2} N k_B T . \quad (1.28)$$

また，エントロピーは式 (1.11) から次で与えられる．

$$S(T) = N k_B \left[\frac{5}{2} + \log \frac{\mathcal{V}}{N} + \frac{3}{2} \log \left(\frac{2\pi m k_B T}{h^2} \right) \right] . \quad (1.29)$$

更に，定積比熱は内部エネルギー $U(T)$ の温度 T についての微分なので，式 (1.28) より，次で与えられる．

$$C_V(T) = \frac{3}{2} N k_B . \quad (1.30)$$

ここで，式 (1.29) を以下のように書き換える．

$$S(T) = k_B \log \left[e^{\frac{5}{2}} \frac{\mathcal{V}}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3}{2}} \right]^N . \quad (1.31)$$

式 (1.28) より， $k_B T$ を内部エネルギー U で置き換えて，この式は更に次のように書き換えられる．

$$S(U) = k_B \log \left[e^{\frac{5}{2}} \left(\frac{4\pi m}{3h^2} \right)^{\frac{3}{2}} \frac{\mathcal{V}}{N} \left(\frac{U}{N} \right)^{\frac{3}{2}} \right]^N . \quad (1.32)$$

一方，エントロピーは以下のようにも書ける．

$$S(U) = k_B \log n(U) . \quad (1.33)$$

10 第1章 拡張アンサンブルの方法

ここで、 $n(U)$ は状態密度 (density of states) である。すなわち、 $n(U)\Delta U$ は U と $U + \Delta U$ の間に存在する状態の数である。よって、式 (1.32) と式 (1.33) から、状態密度は以下のように書ける。

$$\begin{aligned} n(U) &= e^{\frac{5}{2}N} \left(\frac{4\pi m}{3h^2} \right)^{\frac{3}{2}N} \left(\frac{\mathcal{V}}{N} \right)^N \left(\frac{U}{N} \right)^{\frac{3}{2}N} \\ &\propto \mathcal{V}^N U^{\frac{3}{2}N}. \end{aligned} \quad (1.34)$$

1.2 統計力学に基づいた計算機シミュレーション

計算機シミュレーションは自然現象を計算機上で再現する手法であるが、本章では、特に、統計力学に基づいた計算機シミュレーション法を紹介する。前節で取り上げた個々の統計力学的集団を計算機上で再現するわけであるが、モンテカルロ法と分子動力学法が主な手法となる。

1.2.1 カノニカルアンサンブルを再現するシミュレーション手法

粒子数 (N)・体積 ($\mathcal{V}$)・温度 (T) が一定な系であるカノニカルアンサンブルにおいて、理想気体のように厳密解が存在する場合は少なく、多くの場合は、数値シミュレーションによらざるを得ない。カノニカルアンサンブルを再現するシミュレーション手法としては、乱数 (random number) を用いて確率的に次から次へと状態を発生させていくモンテカルロ法 (MC) とニュートンの運動方程式を差分として表したものを解いていく分子動力学法 (MD) の二つの手法が広く使われている。モンテカルロ法は本シリーズの第2巻で、また、分子動力学法は第6巻でそれぞれ紹介されているので、ここでは、その詳細は繰り返さず、概略を述べることにする。

まず、モンテカルロ法では、普通、運動量 p は考慮せず、座標 q のみを変化させていく。すなわち、物理量 A も q のみの関数を扱い、平均値は式 (1.16) で計算する。式 (1.16) は座標 q に関する積分をポテンシャルエネルギー $E(q)$ の

積分に置き換えて、以下のように書き直すことができる．

$$\langle A \rangle_T = \frac{\int dq A(q) e^{-\beta E(q)}}{Z_q}, \quad (1.35)$$

$$= \frac{\int dE A(E) P_B(E)}{\int dE P_B(E)}. \quad (1.36)$$

ここで、ポテンシャルエネルギー E のカノニカル分布 $P_B(E)$ は系の状態密度 (density of states) $n(E)$ とボルツマン因子 $W_B(E)$ の積で以下のように定義される．

$$P_B(E) = n(E) W_B(E), \quad (1.37)$$

$$W_B(E) = e^{-\beta E}. \quad (1.38)$$

状態密度はエネルギーと共に急速に増加する関数（例えば、式 (1.34) 参照）であり、ボルツマン因子は指数関数的に減少する関数（式 (1.38) 参照）であるから、カノニカル分布は一般にベル型をしている．図 1.3 にポテンシャルエネルギー E のカノニカル分布 $P_B(E)$ のスケッチを示す．

図のベル型のピークのあたりのエネルギー値がその温度における系の平均ポテンシャルエネルギー値であり、ベルの幅が、この温度におけるポテンシャルエネルギー値の熱的揺らぎ幅を示している．

モンテカルロ法では、以下のように、次から次へと状態 $x \equiv \{q\}$ を発生させていく．

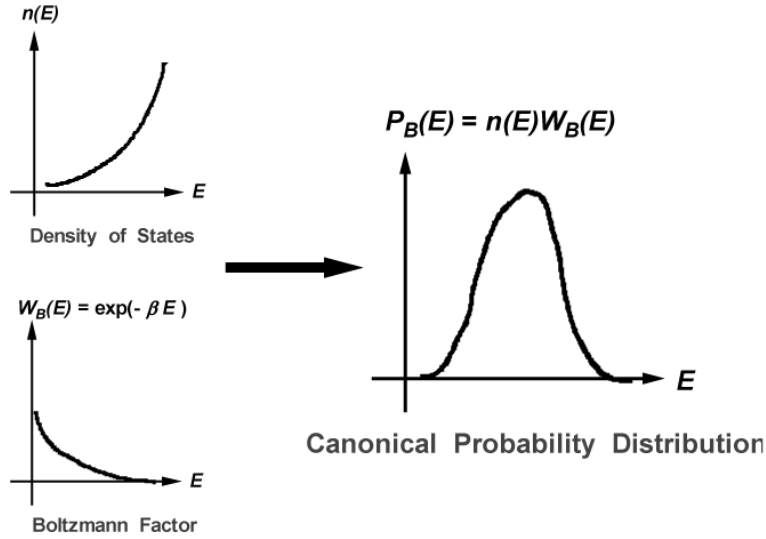
$$x^{(1)} \rightarrow x^{(2)} \rightarrow x^{(3)} \rightarrow x^{(4)} \rightarrow x^{(5)} \dots\dots\dots (1.39)$$

今、モンテカルロシミュレーションが ν 番目の MC ステップまで進み、状態 x_j にあったとする．そして、 $(\nu+1)$ 番目の状態 x_k に遷移するとする．この遷移確率 w は本来過去の全ての状態が何であったかに依存するはずである．よって、

$$w = w\left(x^{(\nu)} = x_j \rightarrow x^{(\nu+1)} = x_k; x^{(1)}, x^{(2)}, \dots, x^{(\nu-1)}\right), \quad (1.40)$$

と書ける．しかし、これでは、不確実性があり過ぎて、遷移確率を決めること

12 第1章 拡張アンサンブルの方法

図 1.3 ポテンシャルエネルギー E のカノニカル分布 $P_B(E)$.

はできない．よって，過去の状態に依存しないと仮定するとともに， ν にも依存しないと仮定する．すなわち，遷移確率を以下のように書くことにする．

$$w\left(x^{(\nu)} = x_j \rightarrow x^{(\nu+1)} = x_k; x^{(1)}, x^{(2)}, \dots, x^{(\nu-1)}\right) = w(x_j \rightarrow x_k) . \quad (1.41)$$

このような過程をマルコフ過程 (Markov process) という．そして，式 (1.39) の状態の流れをマルコフ鎖 (Markov chain) という．

このとき，確率の定義から，以下が成り立つ．

$$P^{(\nu+1)}(x_k) = \sum_j P^{(\nu)}(x_j) w(x_j \rightarrow x_k) . \quad (1.42)$$

ここで， $P^{(\nu)}(x)$ は状態 x の ν 番目のステップにおける確率分布である．そして，このマルコフ過程が ν が無限大の極限で，ある平衡分布 $P_{eq}(x)$ に収束していくことを要求する．つまり，以下が成り立つとする．

$$P_{eq}(x) = \lim_{\nu \rightarrow \infty} P^{(\nu)}(x) . \quad (1.43)$$

分布 $P^{(\nu)}(x)$ が平衡分布 $P_{eq}(x)$ に収束する（熱平衡に達する）とき、式 (1.42) は以下のように書くことができる。

$$P_{eq}(x_k) = \sum_j P_{eq}(x_j) w(x_j \rightarrow x_k) . \quad (1.44)$$

この条件式で大分可能性が狭まり、遷移確率 $w(x_j \rightarrow x_k)$ が決まりそうだが、実はもう一つだけ条件が必要である。例えば、以下の詳細つり合いの条件 (condition of detailed balance) を課することによって問題は解ける。

$$P_{eq}(x_j) w(x_j \rightarrow x_k) = P_{eq}(x_k) w(x_k \rightarrow x_j) . \quad (1.45)$$

ここで、詳細つり合いの条件式 (1.45) が熱平衡の式 (1.44) を満たすことを示そう。すなわち、式 (1.45) を式 (1.44) の右辺に代入すると、以下のように、式 (1.44) の等号が成り立つことを示すことができる。

$$\begin{aligned} \text{(式 (1.44) の右辺)} &= \sum_j P_{eq}(x_k) w(x_k \rightarrow x_j) \\ &= P_{eq}(x_k) \sum_j w(x_k \rightarrow x_j) \\ &= P_{eq}(x_k) \\ &= \text{(式 (1.44) の左辺)} \end{aligned} \quad (1.46)$$

ここで、確率の規格化条件式

$$\sum_j w(x_k \rightarrow x_j) = 1, \quad (1.47)$$

を使った。よって、詳細つり合いの条件の式 (1.45) はこのマルコフ過程が平衡状態に到達することを保障する十分条件である。

この仮定のもとで、メトロポリス等は以下の遷移確率が詳細つり合い条件を満たすことを示した [1].

$$w(x_j \rightarrow x_k) = \begin{cases} 1, & \text{if } \frac{P_{eq}(x_k)}{P_{eq}(x_j)} \geq 1, \\ \frac{P_{eq}(x_k)}{P_{eq}(x_j)}, & \text{if } \frac{P_{eq}(x_k)}{P_{eq}(x_j)} < 1. \end{cases} \quad (1.48)$$

14 第1章 拡張アンサンブルの方法

それでは、この解が詳細つり合い条件の式 (1.45) を満たしていることを示そう。まず、 $\frac{P_{eq}(x_k)}{P_{eq}(x_j)} \geq 1$ の場合を考える。すると、式 (1.48) の第一の等号が成り立ち、式 (1.45) の等号が以下のように成り立つことが分かる。

$$\begin{aligned} \text{(式 (1.45) の左辺)} &= P_{eq}(x_j) \times 1 = P_{eq}(x_j) \\ \text{(式 (1.45) の右辺)} &= P_{eq}(x_k) \times \frac{P_{eq}(x_j)}{P_{eq}(x_k)} = P_{eq}(x_j) = \text{(式 (1.45) の左辺)} . \end{aligned} \quad (1.49)$$

ここで、二つ目の式では、 $w(x_k \rightarrow x_j)$ は $w(x_j \rightarrow x_k)$ で x_j と x_k を入れ替えたものなので、式 (1.48) の第二の等号が成り立ち、 $w(x_k \rightarrow x_j) = P_{eq}(x_j)/P_{eq}(x_k)$ となることを使った。次に、 $\frac{P_{eq}(x_k)}{P_{eq}(x_j)} < 1$ の場合も同様に、式 (1.45) の等号が成り立つことを以下のように示すことができる。

$$\begin{aligned} \text{(式 (1.45) の左辺)} &= P_{eq}(x_j) \times \frac{P_{eq}(x_k)}{P_{eq}(x_j)} = P_{eq}(x_k) \\ \text{(式 (1.45) の右辺)} &= P_{eq}(x_k) \times 1 = P_{eq}(x_k) = \text{(式 (1.45) の左辺)} . \end{aligned} \quad (1.50)$$

特に、カノニカルアンサンブルの場合は、平衡分布は以下で与えられる。

$$P_{eq}(x) = \frac{W_B(E(x))}{Z} . \quad (1.51)$$

このとき、分配関数 Z は打ち消し合って、式 (1.48) の遷移確率は以下で与えられる。

$$w(x_j \rightarrow x_k) = \begin{cases} 1 , & \text{if } \frac{W_B(E(x_k))}{W_B(E(x_j))} \geq 1 , \\ \frac{W_B(E(x_k))}{W_B(E(x_j))} , & \text{if } \frac{W_B(E(x_k))}{W_B(E(x_j))} < 1 . \end{cases} \quad (1.52)$$

これは式 (1.38) より、更に以下のようにも書ける。

$$w(x_j \rightarrow x_k) = \begin{cases} 1 , & \text{if } \Delta E \leq 0 , \\ \exp(-\beta \Delta E) , & \text{if } \Delta E > 0 . \end{cases} \quad (1.53)$$

ここで、 ΔE は次で定義される。

$$\Delta E \equiv E(x_k) - E(x_j) . \quad (1.54)$$

1.2 統計力学に基づいた計算機シミュレーション 15

実際の計算は以下のように行われる．今，モンテカルロシミュレーションが ν 番目の MC ステップまで進み，状態 x_j にあったとする．そして， $(\nu+1)$ 番目の状態の候補 $x_{cand}^{(\nu+1)}$ として，状態 x_k を用意したとする．

$$x^{(1)} \rightarrow x^{(2)} \rightarrow \cdots \rightarrow x^{(\nu)} = x_j \rightarrow x_{cand}^{(\nu+1)} = x_k . \quad (1.55)$$

まず， $\Delta E = E(x_k) - E(x_j)$ を計算する．もし， $\Delta E \leq 0$ なら，この候補 $x_{cand}^{(\nu+1)}$ を採択する．逆に， $\Delta E > 0$ なら，0 から 1 の間の値をとる一様乱数 r を発生させ， $r \leq \exp(-\beta\Delta E)$ なら，候補 $x_{cand}^{(\nu+1)}$ を採択するという手順を踏むことになる．ここで，候補 $x_{cand}^{(\nu+1)}$ を採択する (accept)，棄却する (reject) とは， $(\nu+1)$ 番目の状態 $x^{(\nu+1)}$ を，それぞれ，以下のように置くことである．

$$x^{(\nu+1)} = \begin{cases} x_{cand}^{(\nu+1)} = x_k , & (\text{採択の場合}) , \\ x^{(\nu)} = x_j , & (\text{棄却の場合}) . \end{cases} \quad (1.56)$$

すなわち，ポテンシャルエネルギーが下がれば ($\Delta E \leq 0$)，確率 1 で候補を受け入れ，ポテンシャルエネルギーが上がってしまえば ($\Delta E > 0$)，1 より小さい確率 $\exp(-\beta\Delta E)$ で候補を受け入れるわけである．このフローチャート（この一連の操作を「メトロポリス判定」と呼ぶ）を図 1.4 に示した．

メトロポリスモンテカルロ法では，この操作を繰り返すことによって，カノニカルアンサンブルを計算機上に発生させる．

式 (1.52) と式 (1.53) は，両方の場合をまとめて以下のように書くことができる．

$$w(x_j \rightarrow x_k) = \min \left(1, \frac{W_B(E(x_k))}{W_B(E(x_j))} \right) = \min (1, \exp(-\beta\Delta E)) . \quad (1.57)$$

ここで， $\min(A, B)$ は二つの数字 A と B のうち，小さい方を表す関数であり，以下で定義される．

$$\min(A, B) = \begin{cases} A , & \text{if } A \leq B , \\ B , & \text{if } B \leq A . \end{cases} \quad (1.58)$$

以後，このような書き方を採用することにする．

16 第1章 拡張アンサンブルの方法

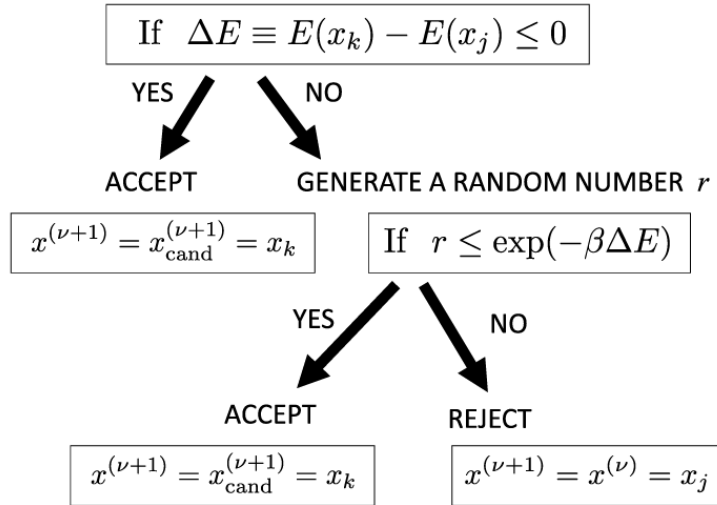


図 1.4 メトロポリスモンテカルロ法のフローチャート：メトロポリス判定.

ここで、式 (1.55) のマルコフ鎖において、 $x^{(\nu)} = x_j$ から候補 $x_{\text{cand}}^{(\nu+1)} = x_k$ を発生させるときに注意すべきことを述べる。まずは、この状態発生操作は任意の二つの状態を有限回数の操作でつなぐようでないといけない（エルゴード性とよばれる）。また、候補を発生させる操作においては、 $x_j \rightarrow x_k$ の生成確率とその逆 $x_k \rightarrow x_j$ の生成確率は普通同じにとる（同じでない場合は、上のメトロポリス判定の式 (1.48) に、この二つの状態発生確率の比の係数が掛ることになる）。

以上のように、メトロポリスのモンテカルロ法では、値域が 0 から 1 の間の一様乱数を頻繁に使う。計算機でこのような乱数を発生させる方法は幾つもあるが、ここでは最も簡単な方法の一つを紹介する。

M_n を n ビットの計算機が扱える最大整数とする。1 ビットは符号を表すのに使うので、32 ビットの計算機なら、 $M_{32} = 2^{31} - 1$ であり、64 ビットの計算機なら、 $M_{64} = 2^{63} - 1$ である。今、次から次へと整数の掛け算をするに

```
(a)
      DOUBLE PRECISION R,RAND
      INTEGER IX
C*** SEED FOR RANDOM NUMBER
      IX=30001
C***
      -----
      CALL RANDOM(IX,RAND)
      R = RAND
      -----
      STOP
      END
C***
      SUBROUTINE RANDOM(IX,RAND)
      DOUBLE PRECISION RAND
      INTEGER M,K,IX
C*** M=2**31-1, K=5**11
      DATA M,K /2147483647, 48828125/
      IX=IAND(IX*K,M)
      RAND=DBLE(IX)/DBLE(M)
      RETURN
      END
```

```
(b)
double random(int *seed); //random number
int IX = 30001; //seed for random number
-----
r = random(&IX)
-----

double random(int *seed){
    int M = 2147483647; //2^31-1
    int K = 48828125; //5^11
    double rand; //random number
    *seed = ((*seed)*K)&M;
    rand = (double)*seed/M;
    return rand;
}
```

図 1.5 乱数生成プログラムの例。プログラム言語は、(a) FORTRAN, (b) C.

18 第1章 拡張アンサンブルの方法

よって、新しい整数を作っていく．

$$I_{\ell+1} = I_{\ell} \times K \quad (1.59)$$

ここで、 K はある大きな正の整数である．また、 I_1 はこの整数列の初期値であり、乱数の種 (seed) と呼ばれる (普通、正の奇数が選ばれる)．ステップ数 ℓ が大きくなるにつれて、理論的には最大整数 M_n を越えてしまうが、これを計算機は無理やり絶対値が M_n 以下の整数として表してしまうわけである (符号は変わり得る)．よって、これらの整数は非常に乱雑になり、以下の数が正の乱数 $r_{\ell+1}$ ($0 < r_{\ell+1} < 1$) となる．

$$r_{\ell+1} = \frac{|I_{\ell+1}|}{M_n} \quad (1.60)$$

ここで、FORTRAN 言語では、乱数のための絶対値は以下のようにビット演算子「IAND」(C 言語での「&」に対応する) を使う．(組み込み関数「ABS」を使わないのは、生成される乱数の一様性を増すためである．) まず、 n ビットの計算機で整数 I を以下のように二進法で表す．

$$I = \{b_n, b_{n-1}, b_{n-2}, \dots, b_3, b_2, b_1\} \quad (1.61)$$

ここで、各桁の数 b_i ($i = 1, 2, \dots, n$) は 0 または 1 の値を取る．このとき、最大整数 M_n は以下で表せる．

$$M_n = \{0, 1, 1, \dots, 1, 1, 1\} \quad (1.62)$$

ここで、最初の 0 は符号を指定するビットで正の符号を表している (負の符号のときは 1)．ビットごとの AND 演算は両方のビット数が 1 のときのみ 1 となるので、任意の b_i について以下が成り立つ．

$$\begin{aligned} 0 &= AND(b_i, 0) \\ b_i &= AND(b_i, 1) \end{aligned} \quad (1.63)$$

よって、式 (1.61) と式 (1.62) の各成分ビット数を比較して、整数対についての IAND 演算子を使うと以下が成り立つ．

$$IAND(I, M) = \{0, b_{n-1}, b_{n-2}, \dots, b_3, b_2, b_1\} \quad (1.64)$$

すなわち, 最初の符号のビットだけ 0 (つまり, 正) になり, 残りのビットは変更なしである. よって, 式 (1.61) の整数 I の符号が如何にかかわらず, $IAND(I, M_n)$ は整数 I の符号を正に変換するわけで, 絶対値記号の役目を果たす. すなわち, 式 (1.60) の代わりに, 結局以下の数が $(\ell + 1)$ 番目の乱数 $r_{\ell+1}$ ($0 < r_{\ell+1} < 1$) として使われる.

$$r_{\ell+1} = \frac{IAND(I_{\ell+1}, M_n)}{M_n} \quad (1.65)$$

ここで, $I_{\ell+1}$ は I_ℓ から式 (1.59) で得られる. つまり, 乱数は式 (1.59) と式 (1.65) を交互に繰り返すことだけで次から次への発生されていくわけである.

ここで注意しなければならないのは, 以上の導出から分かるように, この乱数発生操作には周期性があり, 最大で $M_n (= 2^{n-1} - 1)$ 回乱数を発生させると元の初期乱数に戻ることである. 図 1.5 に 32 ビット計算機用のこの乱数生成プログラムの例を FORTRAN および C のプログラム言語の場合において示した.

実は, 図 1.5 のプログラムは 64 ビット計算機のようにビット数が 32 より大きい計算機でもそのまま利用できる. これは, 例えば, 64 ビット計算機の場合, 整数は二進法で以下のように表される (式 (1.61) 参照).

$$I = \{b_{64}, b_{63}, b_{62}, \dots, b_3, b_2, b_1\} \quad (1.66)$$

このとき, 式 (1.62) の 32 ビット計算機の場合の最大整数 M_{32} は以下で表わされる.

$$M_{32} = \{0, 0, \dots, 0, 1, \dots, 1, 1\} \quad (1.67)$$

すなわち, 最初の 33 個のビットには全て 0 がはいっている. すると, 式 (1.63) より, 任意の整数 I について, 以下が得られる.

$$IAND(I, M_{32}) = \{0, 0, \dots, 0, b_{31}, b_{30}, \dots, b_2, b_1\} \quad (1.68)$$

これは, 正の整数であり, 絶対値が M_{32} より小さいので, 式 (1.65) で $n = 32$ としてそのまま 64 ビットの計算機の乱数として使うことができる (ただし, 乱数の周期が, 最大で $2^{63} - 1$ ではなく, $2^{31} - 1$ 以下になってしまう).

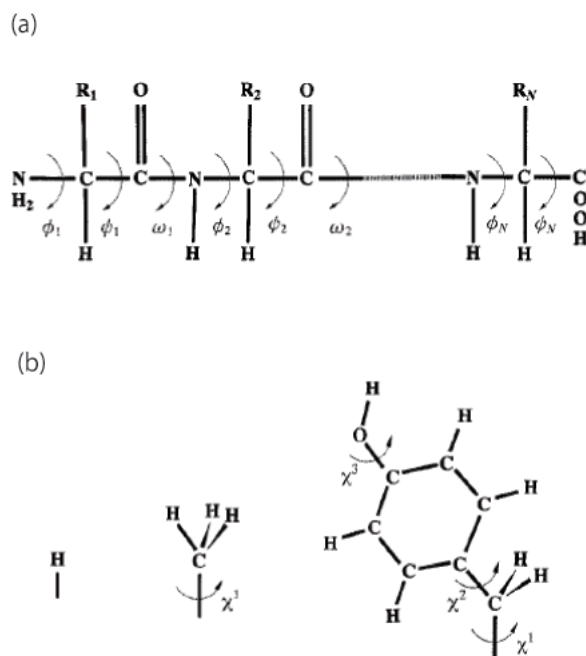


図 1.6 (a) 蛋白質の模式図。 N 個のアミノ酸が一列に重合している。 R の部分は側鎖と呼ばれ、残りが主鎖と呼ばれる。 ϕ , ψ , ω は主鎖の立体構造を決める二面角である。(b) 蛋白質の側鎖の例。左はグリシン (Gly), 中はアラニン (Ala), 右はチロシン (Tyr)。 χ は側鎖の立体構造を決める二面角である。

1.2 統計力学に基づいた計算機シミュレーション 21

これらの乱数が一様分布になっているかどうかの最も簡単な確認事項としては、以下のようなテストが考えられる。理論的には、一様乱数の確率分布は以下で与えられる。

$$p(r) = 1. \quad (1.69)$$

よって、乱数の期待値と分散は以下で与えられる。

$$\begin{aligned} \langle r \rangle &= \int_0^1 r p(r) dr = \int_0^1 r dr = \frac{1}{2} \\ \langle r^2 \rangle - \langle r \rangle^2 &= \int_0^1 r^2 p(r) dr - \left(\int_0^1 r p(r) dr \right)^2 = \frac{1}{3} - \left(\frac{1}{2} \right)^2 = \frac{1}{12}. \end{aligned} \quad (1.70)$$

図 1.5 のプログラムを使って、 $\mathcal{N}$ 回乱数を発生させて、 r_ℓ ($\ell = 1, \dots, \mathcal{N}$) を得たとする。この乱数の平均と分散は以下で得られる。

$$\begin{aligned} \bar{r} &= \frac{1}{\mathcal{N}} \sum_{\ell=1}^{\mathcal{N}} r_\ell \\ \overline{r^2} - \bar{r}^2 &= \frac{1}{\mathcal{N}} \sum_{\ell=1}^{\mathcal{N}} r_\ell^2 - \left(\frac{1}{\mathcal{N}} \sum_{\ell=1}^{\mathcal{N}} r_\ell \right)^2. \end{aligned} \quad (1.71)$$

この二つの量が $\mathcal{N}$ が大きくなるにつれて、式 (1.70) の理論値に収束していくことを確認するのが一つのテストである。また、得られた $\mathcal{N}$ 個の乱数 r_ℓ ($\ell = 1, \dots, \mathcal{N}$) の分布を例えば 10 個の領域 $(0.0, 0.1)$, $(0.1, 0.2)$, $(0.2, 0.3)$, $\dots$, $(0.9, 1.0)$ に分けて、それぞれの領域にはいる乱数の数が、 $\mathcal{N}$ が大きくなるにつれて、理論値 ($\mathcal{N}/10$) に収束していくことを確認するのがもう一つのテストである。

ここで紹介した乱数作成プログラムは最も簡単なものの一つであり、最高で M_n 回乱数を発生させると元の初期乱数に戻るという周期性がある。それで、イジング模型のような対称性が高い系などにおける MC シミュレーションでは注意を要する場合がある（より厳密な一様性と長い周期をもつ乱数作成プログラムについては、例えば、文献 [2] を参照されたい）。しかし、分子シミュレーションの粒子系のような対称性の低い系などでは、十分に使えるものである。

22 第1章 拡張アンサンブルの方法

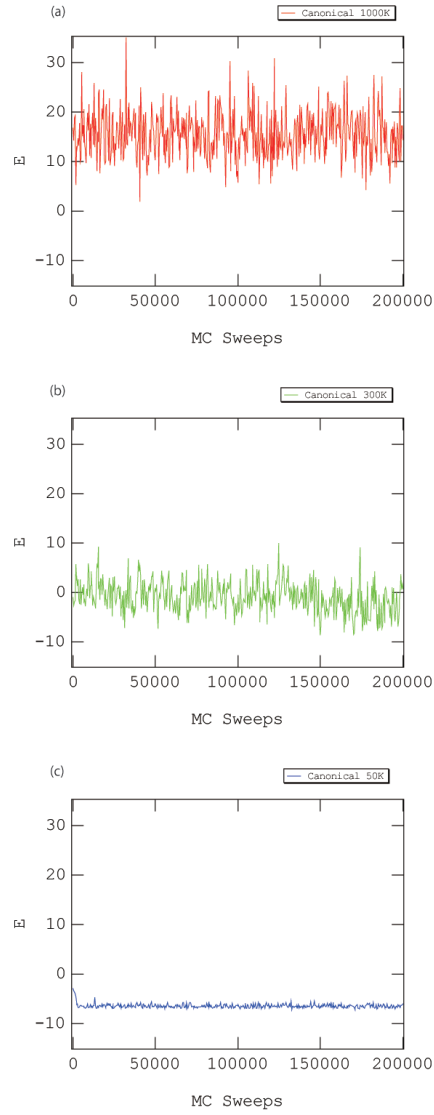


図 1.7 温度一定のカノニカルアンサンブルを再現するモンテカルロシミュレーションにおける真空中の Met-enkephalin のポテンシャルエネルギー E (kcal/mol) の「時間発展」. (a) $T = 1000$ K, (b) $T = 300$ K, (c) $T = 50$ K. ここで, kcal/mol はエネルギーの単位 (1 モル当たり何キロカロリーか) であり, $23 \text{ kcal/mol} \approx 1 \text{ eV}$ が成り立つ. また, 1 MC Sweep とは, 全ての座標自由度 (ここでは, 二面角と呼ばれる回転角のこと) を 1 回ずつ更新することに対応する. そして, 更新の度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を課す操作に対応する.

1.2 統計力学に基づいた計算機シミュレーション 23

ここで、式 (1.55) のマルコフ鎖において、 $x^{(\nu)} = x_j$ から候補 $x_{cand}^{(\nu+1)} = x_k$ を発生させる操作を、具体的な例で実際の計算ではどのようにするかを示そう。最初の例は、磁性体のイジング模型 (Ising model) である。この系の状態は $x = \{S_1, S_2, \dots, S_N\}$ で与えられる。ここで、 $S_k (= \pm 1)$ は k 番目のスピンを表し、 N は全スピンの数である。そして、ゼロ外磁場の場合のポテンシャルエネルギーは以下で与えられる。

$$E = -J \sum_{\langle i, j \rangle} S_i S_j . \quad (1.72)$$

ここで、 J は結合定数 ($J > 0$ のとき強磁性、 $J < 0$ のとき反強磁性)、 $\sum_{\langle i, j \rangle}$ は最隣接のスピン対についての和を表す。 $x^{(\nu)} = \{S_1^{(\nu)}, S_2^{(\nu)}, \dots, S_N^{(\nu)}\}$ を与えられたとき、次の状態の候補 $x_{cand}^{(\nu+1)}$ は $x^{(\nu)}$ のスピンのうちから、一つ $S_i^{(\nu)}$ を (ランダムに) 選び、その符号を入れ替えることで得られる (スピンは ± 1 の値のみを取ることに注意されたい)。

$$S_{cand\ i}^{(\nu+1)} = -S_i^{(\nu)} . \quad (1.73)$$

その他のスピンはそのままにする。

なぜ、一つのスピンだけを変化させるのかというと、多くのスピンを同時に変化させると、メトロポリス判定の受諾率が低くなってしまうからである。つまり、エネルギーが高い状態の数の方が低い状態の数より、一般に圧倒的に多いので (例えば、式 (1.34) 参照)、 $\Delta E \gg k_B T$ となって、式 (1.53) より、候補が受諾される確率が非常に低くなってしまふのである。

二つ目の例は、分子シミュレーションにおける、 N 個の粒子系である。この系の状態は $x = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ で表される。 $x^{(\nu)}$ を与えられたとき、次の状態の候補 $x_{cand}^{(\nu+1)}$ は $x^{(\nu)}$ の座標のうちから、一つ $\mathbf{q}_i^{(\nu)}$ を (ランダムに) 選び、それを以下のように変化させることで得られる。

$$\mathbf{q}_{cand\ i}^{(\nu+1)} = \mathbf{q}_i^{(\nu)} + \eta \boldsymbol{\rho} \quad (1.74)$$

ここで、 $\eta (> 0)$ はある定数であり、 $\boldsymbol{\rho}$ は値域が 0 から 1 の一様乱数を三つ (r_x , r_y , r_z) 発生させることによって、以下で定義される。

$$\boldsymbol{\rho} = (2r_x - 1, 2r_y - 1, 2r_z - 1) . \quad (1.75)$$

24 第1章 拡張アンサンブルの方法

次に、蛋白質などの生体高分子系を考えよう．すなわち、原子間が化学結合でつながっている場合を考える．この場合、その化学結合ボンドを両端の原子を結びつけるバネで表すことができる（このバネの自然長がそのボンドの自然の長さになるようにする）．更に、二つの隣り合うボンド間のなす角度もある自然の角度の周りを振動するバネで結びつけることを考える．つまり、高分子系をそれを構成する N 個の原子とそれらをつなぐ質量ゼロのバネで表すわけである．すると、 N 個の粒子系のポテンシャルエネルギーにこれらのバネに対応する調和振動子項を加えれば、後は、上と同じように N 個の原子系として扱い、モンテカルロシミュレーションの座標の更新は、式 (1.74) に従えば良いことになる．

もし、他のポテンシャルエネルギー項に比べて化学結合が強くて、ボンドがバネではなく、長さが固定されるとともに、二本の隣り合うボンド間の角度も固定されているとみなすことができる場合は、問題が更に簡単になり、系の状態は N 個の原子の座標 $x = \{q_1, \dots, q_N\}$ の代わりに、 N_D 個のボンドの周りの回転角で次のように表すことができる ($N_D \ll N$)． $x = \{\theta_1, \dots, \theta_{N_D}\}$ ．この回転角は、対象とするボンドの前後のボンドを加えて、合計三本のボンドを構成する四個の原子のうち、最初の三個で定義される平面と終わりの三個で定義される平面との成す角度のことなので、二面角 (dihedral angle) とも呼ばれる．図 1.6 に蛋白質の概略図と二面角の定義を示す．主鎖の二面角は ϕ , ψ , ω の三種類あるが、 ω に対応するボンドはペプチド結合のボンドであり、大体 180° に固定されているので、モンテカルロシミュレーションでは、しばしば、主鎖の ϕ , ψ および側鎖の χ を系の状態を指定する変数として扱われる．この場合、状態 $x^{(\nu)}$ を与えられたとき、次の状態の候補 $x_{cand}^{(\nu+1)}$ は $x^{(\nu)}$ の二面角のうちから、一つ $\theta_i^{(\nu)}$ を (ランダムに) 選び、それを以下のように変化させることで得られる．

$$\theta_{cand\ i}^{(\nu+1)} = \theta_i^{(\nu)} + 180\eta(2r - 1) . \quad (1.76)$$

ここで、 $\eta (> 0)$ はある定数であり、 r は値域が 0 から 1 の一様乱数である．また、 θ_i は、 ϕ , ψ , χ のいずれかを表す．

以上、 $x^{(\nu)} = x_j$ から候補 $x_{cand}^{(\nu+1)} = x_k$ を発生させる操作を、三つの例にお

いて示した (式 (1.73), (1.74), (1.76) 参照). これらの操作の後, 式 (1.57) のメトロポリス判定により, 候補が受諾 (または棄却) されて, モンテカルロシミュレーションが続いていくわけである.

ところで, 式 (1.74) と式 (1.76) では, 調整可能な定数 η が存在する. この値は, メトロポリス判定の受諾率 ACP をある値に固定するために調節される. ここで, 受諾率は以下で定義される.

$$ACP = \frac{N_A}{N_{TOT}}. \quad (1.77)$$

ここで, N_A と N_{TOT} はそれぞれ, メトロポリス判定が受諾された回数とメトロポリス判定が課された総回数である. 受諾率はできるだけ高い方がよいように思われるかも知れないが, 受諾率が高いということは, 状態があまり変わっていないということである (すなわち, 状態がほとんど変わらないということとは, $|\Delta E| \ll 1$ ということであり, 受諾率は高くても状態空間のほんの一部しかサンプルできないことを意味する). 筆者の経験では, $ACP \approx 0.2$ (つまり, 約 20%) が良い値である. 文献 [3] では, ある条件の下では, 熱平衡への収束が最も速い最適の受諾率 ACP_{opt} は以下の数値 (または, それ以下の数値) であることが数学的に示されている.

$$ACP_{opt} = 0.234. \quad (1.78)$$

式 (1.74) と式 (1.76) の中の定数 η を調節することによって, 受諾率 ACP がちょうど約 0.2 になるようにするわけである. η を小さくすると, 状態変化が小さいので, ΔE が小さくなり, 受諾率が上がる. 逆に, η を大きくすると, 状態変化が大きくなり, ΔE も大きくなる可能性が高いので受諾率が下がる (式 (1.53) 参照). 普通, 短い試行シミュレーションで受諾率が約 0.2 になるような η を求め, その後の本シミュレーションでは, その値を固定する (式 (1.45) の詳細つり合いの条件を破らないためである).

ちなみに以下で紹介する徐冷法では, それぞれの温度で最適な η の値が違うので, シミュレーション中に η の値を変えていく必要がある. よって, 徐冷法では詳細つり合いの条件が破れる (η で議論しなくても, そもそも $P_{eq}(x)$ は温度に依存しているので, シミュレーションの途中で温度を変えると, $P_{eq}(x)$ も

26 第1章 拡張アンサンブルの方法

変わっていき、詳細つり合いの条件が破れてしまう)。よって、徐冷モンテカルロシミュレーションから得られる、いろいろな温度における物理量の期待値などは信頼性を伴わないことに注意されたい。

次に、分子動力学法をごく簡単に説明する。ここでは、運動量も考慮に入れる必要があり、以下のニュートンの運動方程式に対応する差分方程式を解くことを考える。

$$\dot{\mathbf{q}}_k = \frac{\mathbf{p}_k}{m_k}, \quad (1.79)$$

$$\dot{\mathbf{p}}_k = -\frac{\partial E}{\partial \mathbf{q}_k} = \mathbf{F}_k, \quad (1.80)$$

ここで、 $\mathbf{F}_k$ は k 番目 ($k = 1, \dots, N$) の粒子に働く力である。

しかし、解析力学で学んだように (系の作用積分が時間の平行移動について不変の場合)、この方程式では、全エネルギー ($K(p) + E(q)$) が保存されるので、ミクロカノニカルアンサンブルになってしまう。

温度一定のカノニカルアンサンブルを実現する分子動力学法は、例えば、能勢の方法 [4] に従って、以下の微分方程式を差分方程式にして解けば良いことが知られている。

$$\dot{\mathbf{q}}_k = \frac{\mathbf{p}_k}{m_k}, \quad (1.81)$$

$$\dot{\mathbf{p}}_k = -\frac{\partial E}{\partial \mathbf{q}_k} - \frac{\dot{s}}{s} \mathbf{p}_k = \mathbf{F}_k - \frac{\dot{s}}{s} \mathbf{p}_k, \quad (1.82)$$

$$\dot{s} = s \frac{P_s}{Q}, \quad (1.83)$$

$$\dot{P}_s = \sum_{k=1}^N \frac{\mathbf{p}_k^2}{m_k} - 3Nk_B T = 3Nk_B (T(t) - T). \quad (1.84)$$

ここで、 s は能勢のスケールパラメータ、 P_s はその共役運動量、 Q はその「質量」を表す。また、「瞬間温度」 $T(t)$ は以下で定義される。

$$T(t) = \frac{1}{3Nk_B} \sum_{k=1}^N \frac{\mathbf{p}_k(t)^2}{m_k}. \quad (1.85)$$

上のモンテカルロ法も分子動力学法も温度一定のカノニカルアンサンブルを実現するので、各状態はボルツマン因子に比例して発生させられる。よって、

1.2 統計力学に基づいた計算機シミュレーション 27

式 (1.35) の物理量 A の温度 T における平均値は、以下の算術平均で求めることができる。

$$\langle A \rangle_T = \frac{1}{n_s} \sum_{\nu=1}^{n_s} A(x^{(\nu)}) . \quad (1.86)$$

ここで、 $x^{(\nu)}$ はモンテカルロまたは分子動力学シミュレーションで発生された ν 番目の状態であり、 n_s はその総サンプル数である。普通、これらのデータは一定の MC (または、MD) ステップ毎に定間隔に抽出されることが多い。MC でも MD でも、隣り合う状態間の変化は小さいので、毎回状態を保存する意味がないからである (勿論、計算機のハードディスクと計算時間に余裕があれば、毎回保存しても良いわけである)。

ここで注意しておかなければならないのは、モンテカルロ法でも分子動力学法でもシミュレーションのごく初期の段階は熱平衡に達していないということである。例えば、(各粒子の座標が) 完全にランダムな初期状態を用意したとすると、これは温度が非常に高い状態の一つである (よって、「ホットスタート (hot start)」と呼ぶ)。よって、この場合は、シミュレーションの最初のうちは、ポテンシャルエネルギーが非常に高く、図 1.3 で言えば、正しい温度 T のカノニカル分布の右側に大きく外れたところからシミュレーションがスタートしたことを意味する。同様に、(各粒子の座標が) 結晶構造であるような初期状態を用意したとすると、これは温度が非常に低い状態の一つである (よって、「コールドスタート (cold start)」と呼ぶ)。そして、この場合は、シミュレーションの最初のうちは、ポテンシャルエネルギーが非常に低く、図 1.3 で言えば、正しい温度 T のカノニカル分布の左側に大きく外れたところからシミュレーションがスタートしたことを意味する。いずれの場合も、シミュレーションが暫く続いて熱平衡に達すると、正しいカノニカル分布に所属する状態が発生するようになるので、ポテンシャルエネルギーの揺らぎが図 1.3 のベル型の分布内にとどまるようになる。式 (1.86) の和は、このように正しいカノニカルアンサンブルのベル型の中にエネルギーがおさまるようになってから (熱平衡に達してから) のデータだけを考慮することに注意しなければならない。また、計算結果が初期状態に依存しないことを確かめるために、ホットスタートとコールド

28 第1章 拡張アンサンブルの方法

スタートの両方のシミュレーションを行って、式 (1.86) の期待値がエラーの範囲内で一致することを確認することが多い。

さて、ここで、実際のカノニカルアンサンブルを再現するシミュレーション手法の適用例を示してみよう。系は、真空中のアミノ酸数 5 個の小ペプチド Met-enkephalin であり、メトロポリスモンテカルロ法を用いた。特に、ボンド長とボンド角が固定されているとして、 N_D 個の二面角を自由度として、状態を $x = \{\theta_1, \dots, \theta_{N_D}\}$ で指定した。そして、モンテカルロ法の状態の発生操作は式 (1.76) によった。図 1.7 に三つの温度におけるポテンシャルエネルギー E の「時間発展」の様子を示す。

図からまず分かることは、全ての温度において、ポテンシャルエネルギーはシミュレーションの初期において既に平均値の周りを一定の幅を持って熱的に揺らいでいることである。一番熱平衡への到達が遅い最低温度 (50 K) においても、約 5000 MC Sweep の後には熱平衡に達しているように見える。1000 K, 300 K, 50 K における平均ポテンシャルエネルギーとその周りの熱揺らぎの幅は、それぞれ、15 と 15, 0 と 10, -7 と 2 ぐらいである (図 1.3 のカノニカル分布を参照されたい)。しかし、低温 (50 K) では熱的揺らぎが非常に小さいので、実は、この温度のシミュレーションはエネルギー極小状態に留まってしまって、間違った答えを出してしまっている。後に示すように、50 K の平均ポテンシャルエネルギー値は -7 ではなく、約 -11 である。

このように、温度一定のカノニカルアンサンブルを再現するモンテカルロ法や分子動力学法は広く使われているが、多自由度複雑系ではエネルギー極小状態が無数に存在するために、精度の高いシミュレーションが非常に難しくなる。すなわち、熱的揺らぎが小さい低温ではシミュレーションがそれらエネルギー極小状態に落ち込んでしまうとそこから脱出するのが非常に難しくなるために、間違った答えを出してしまう (正しい熱平衡に達するまで膨大な計算時間を必要とする)。この困難を克服する有効な手法の一つが次の節で紹介する徐冷法である。

1.2.2 徐冷法

徐冷法 (simulated annealing: SA) [5] は物質の結晶構造を作る過程に基づいている。すなわち、融解温度より十分高温でシミュレーションを始め、シミュレーションを通じて徐々に温度を下げていき、最終的には最小エネルギー状態 (結晶) に到達しようとする方法である。もし、温度の下げ方が熱平衡を保つ程緩やかならば、最終温度が 0 K の時、最小エネルギー状態のみが得られる。しかし、急冷すれば、エネルギー極小状態に留まってしまうわけである。徐冷法の原理を図 1.8 に示した。

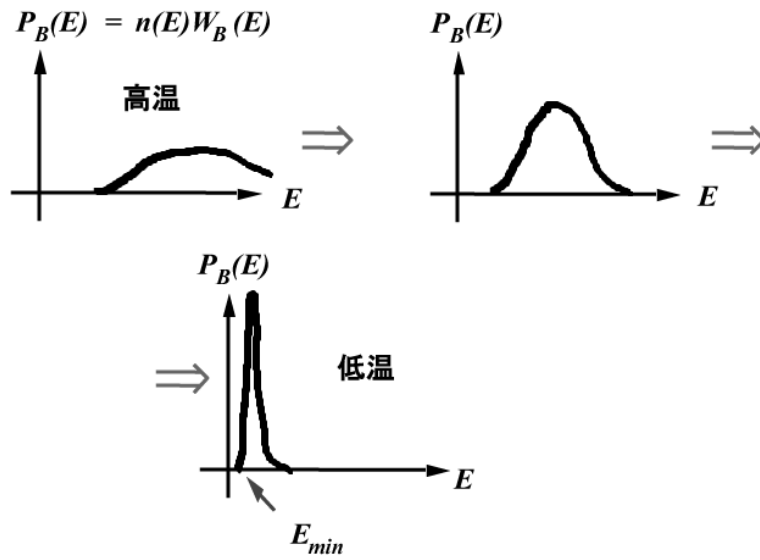


図 1.8 徐冷法.

温度の下げ方を以下のように対数関数的にすると、最小エネルギー状態に到達するのが数学的に保証されている [6].

$$T_m = T_1 \frac{\log 2}{\log(1+m)} . \quad (1.87)$$

ここで、 T_1 は初期温度、 T_m は m 番目のステップの温度である。しかし、実際の計算では、有限の計算時間の下で、このように (熱平衡を保つ程) ゆっくりとは温度を下げていられないので、以下の式のように、指数関数的に下げる場

30 第1章 拡張アンサンブルの方法

合が多い [5, 7]. すなわち, m 番目のステップでの温度は次で与えられる.

$$T_m = T_1 \gamma^{m-1}. \quad (1.88)$$

ここで, T_1 は初期温度, 定数 γ は以下で与えられる.

$$\gamma = \left(\frac{T_M}{T_1} \right)^{\frac{1}{M-1}}. \quad (1.89)$$

ここで, T_M は最終温度, M は一つの試行の全シミュレーションステップ数である.

徐冷法は, モンテカルロ法でも分子動力学法でも同じように実行できるが, ここでは, 上の Met-enkephalin の例と同じ系で徐冷モンテカルロシミュレーションを実行した結果を示す. 図 1.9 にポテンシャルエネルギー E がシミュレーションが進むにつれて, どのように変化していくかを示した.

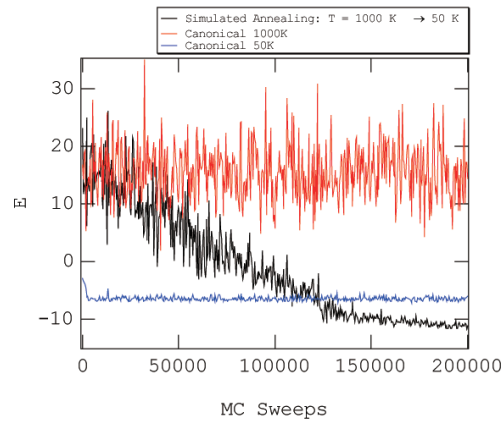


図 1.9 徐冷モンテカルロシミュレーションにおける真空中の Met-enkephalin のポテンシャルエネルギー E (kcal/mol) の「時間発展」. 温度を $T = 1000$ K から $T = 50$ K まで下げた. 図 1.7 の温度が $T = 1000$ K (赤色) と $T = 50$ K (青色) におけるカノニカルシミュレーションの結果も重ねて示した.

まず, 徐冷シミュレーションの初期では, 初期温度 $T_1 = 1000$ K における平均値 ($E \approx 15$) の周りをその温度における熱的揺らぎ幅 (約 15) で揺らいでい

るが、温度が下がるにつれて、ポテンシャルエネルギー値が下がるとともに熱的揺らぎ幅も小さくなっていく。また、図 1.7 のカノニカルシミュレーションでは、 $T = 50$ K でエネルギー極小状態 ($E \approx -7$) に留まってしまったが、徐冷シミュレーションでは、正しく温度 $T = 50$ K の平均値 ($E \approx -11$) まで到達していることが分かる。

徐冷法は従来のカノニカルアンサンブルを再現するシミュレーション手法をそのまま使いながら、シミュレーションの途中で温度を下げていくだけなので、非常に使い易い。よって、様々な複雑系の最適化問題で広く使われている。しかし、温度を変えていくということは、熱平衡を常に破っているわけであり、物理量の平均値は精度良く計算することができない。また、熱平衡が保てないということは、低温でエネルギー極小状態に留まってしまう有限の確率があるということを意味しているので、徐冷法は初期条件を変えて、何回も試行を繰り返し、結果を比較検証する必要がある。

徐冷法ほど使い易くはないが、徐冷法のこの問題点を解決する、更に有効な最適化法が次節で紹介する拡張アンサンブル法である。

1.3 拡張アンサンブル法

生体系などの多自由度複雑系では、エネルギー極小状態が無数に存在するため、温度一定のカノニカルアンサンブル上の従来のモンテカルロ法や分子動力学法による計算機シミュレーションでは、それらのエネルギー極小状態に留まってしまい、誤った答えを出してしまうという難点がある。この困難を克服するのが、本節で紹介する拡張アンサンブル法 (generalized-ensemble algorithm) と総称される手法である (これまでに書かれた日本語の解説としては、例えば、文献 [8] がある)。この方法は、非ボルツマン因子による人工の統計力学的集団に基づいており、ポテンシャルエネルギー空間や他のパラメータ空間上の酔歩 (random walk) を実現することで、エネルギー極小状態に留まるのを避ける。そして、従来の手法と比べて格段に幅広い構造空間を探索することを可能とするのである。ここでは、広く使われている三つの拡張アンサンブル法 (すなわち、マルチカノニカル法、焼き戻し法、レプリカ交換法) を紹介する。

32 第1章 拡張アンサンブルの方法

1.3.1 マルチカノニカル法

拡張アンサンブル法の最初の例として、マルチカノニカル法 (multicanonical algorithm: MUCA) [9] について述べる (エントロピーサンプル法 [10] と呼ばれる). マルチカノニカル法では、各状態は次の非ボルツマン因子 $W_{MUCA}(E)$ に比例する確率で発生し、ポテンシャルエネルギーの一様分布が得られる.

$$P_{MUCA}(E) = n(E) W_{MUCA}(E) \equiv \text{constant} . \quad (1.90)$$

すなわち、この人工の統計力学的アンサンブルではポテンシャルエネルギー空間上の一次元自由酔歩が実現される. よって、シミュレーションがどのようなエネルギー極小状態に落ち込んでしまっても、いずれそれから脱出してくるわけで、従来の温度一定の手法よりはるかに短いステップ数で幅広い配位空間をサンプルすることができる.

図 1.10 にポテンシャルエネルギー E のマルチカノニカル分布 $P_{MUCA}(E)$ のスケッチを示す.

式 (1.90) から分かるように、マルチカノニカル重み因子は状態密度に逆比例するので、以下のように書くことができる.

$$W_{MUCA}(E) \equiv e^{-\beta_0 E_{MUCA}(E; T_0)} = \frac{1}{n(E)} . \quad (1.91)$$

ここで、任意の参照温度 $T_0 (= 1/k_B \beta_0)$ を導入し、「マルチカノニカルポテンシャルエネルギー」 $E_{MUCA}(E; T_0)$ を次式で定義した.

$$E_{MUCA}(E; T_0) = k_B T_0 \log n(E) = T_0 S(E) . \quad (1.92)$$

ここで、

$$S(E) = k_B \log n(E) \quad (1.93)$$

は (ミクロカノニカルアンサンブルにおける) 「エントロピー」である.

すると、マルチカノニカル MC 法はいつものメトロポリス判定により実行される [9]. すなわち、ポテンシャルエネルギー E をもつ状態 x から E' をもつ状態 x' への遷移確率は、式 (1.57) において $W_B(E)$ を式 (1.91) の $W_{MUCA}(E)$

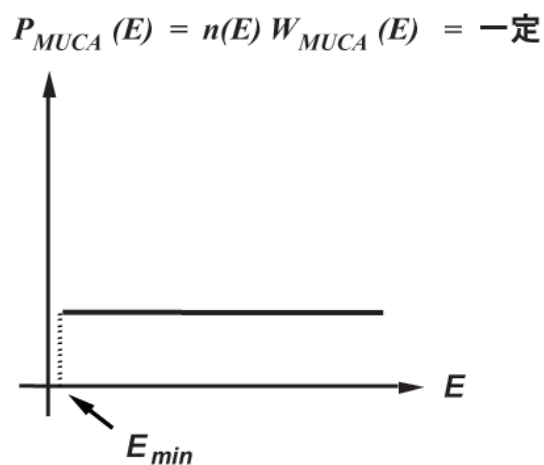


図 1.10 ポテンシャルエネルギーのマルチカノニカル分布. これは厳密なマルチカノニカル重み因子 $W_{MUCA}(E)$ が得られた場合の理想的なスケッチであり, 実際のシミュレーションではこのように完全な一様分布は得られることはなく, ある程度でこぼこがある. マルチカノニカル重み因子 $W_{MUCA}(E)$ は普通短いシミュレーションを繰り返し行うことによって逐次的に決められるが, このイテレーションをいつ止めるかは, ポテンシャルエネルギー分布が大体一様 (普通, 同じオーダーなら良しとする) になり, ポテンシャルエネルギー空間の十分な酔歩が得られるようになったときに止めることになる.

34 第1章 拡張アンサンブルの方法

で置き換えることによって、次で与えられる.

$$w(x \rightarrow x') = \min \left(1, \frac{W_{MUC A}(E')}{W_{MUC A}(E)} \right) \quad (1.94)$$

$$= \min \left(1, \frac{n(E)}{n(E')} \right) \quad (1.95)$$

$$= \min (1, \exp(-\beta_0 \Delta E_{MUC A})) . \quad (1.96)$$

ここで、 $\Delta E_{MUC A}$ は次で定義される.

$$\Delta E_{MUC A} \equiv E_{MUC A}(E'; T_0) - E_{MUC A}(E; T_0) . \quad (1.97)$$

実際の計算では、状態 x から次の状態の候補 x' を発生させ、 $\Delta E_{MUC A}$ を計算する. もし、 $\Delta E_{MUC A} \leq 0$ なら、この候補 x' を採択する. 逆に、 $\Delta E_{MUC A} > 0$ なら、0 から 1 の間の値をとる一様乱数 r を発生させ、 $r \leq \exp(-\beta_0 \Delta E_{MUC A})$ なら、候補 x' を採択するという、いつものメトロポリス法の手順を踏めばよいことになる (図 1.4 参照).

マルチカノニカルアンサンブル上の MD 法も、温度一定 ($T = T_0$) の条件に基づきながら、式 (1.81)~式 (1.84) を解くことで実現されるが、その中で、式 (1.82) は以下の「修正されたニュートン方程式」に置き換わる (式 (1.91) 参照) [11].

$$\dot{\mathbf{p}}_k = -\frac{\partial E_{MUC A}(E; T_0)}{\partial \mathbf{q}_k} - \frac{\dot{s}}{s} \mathbf{p}_k = \frac{\partial E_{MUC A}(E; T_0)}{\partial E} \mathbf{F}_k - \frac{\dot{s}}{s} \mathbf{p}_k . \quad (1.98)$$

ここで、 $\mathbf{F}_k$ は k 番目 ($k = 1, \dots, N$) の原子に働く力である.

まず、厳密なマルチカノニカル重み因子 $W_{MUC A}(E)$ (つまり、状態密度 $n(E)$) が分かっているとする (このとき、図 1.10 に対応して、完全な一様分布になっている). このとき、物理量 $A(E)$ の任意の温度 $T (= 1/k_B \beta)$ における期待値は、シミュレーションを実行しなくても、以下の式から得られる (式 (1.36), (1.37), (1.38) 参照).

$$\langle A \rangle_T = \frac{\sum_E A(E) P_B(E; T)}{\sum_E P_B(E; T)} = \frac{\sum_E A(E) n(E) e^{-\beta E}}{\sum_E n(E) e^{-\beta E}} . \quad (1.99)$$

ここで、状態密度 $n(E)$ は以下で与えられる (式 (1.91) 参照).

$$n(E) = \frac{1}{W_{MUCA}(E)}. \quad (1.100)$$

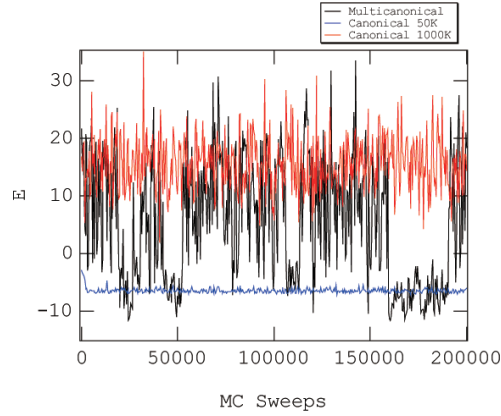


図 1.11 マルチカノニカルモンテカルロシミュレーションにおける真空中の Met-enkephalin のポテンシャルエネルギー E (kcal/mol) の「時間発展」. 図 1.7 の温度が $T = 1000$ K (赤色) と $T = 50$ K (青色) におけるカノニカルシミュレーションの結果も重ねて示した.

しかし、一般的には系の状態密度は厳密に分かっていないので、マルチカノニカル重み因子 $W_{MUCA}(E)$ (よって、マルチカノニカルポテンシャルエネルギー $E_{MUCA}(E; T_0)$) を前もって決定する必要がある. マルチカノニカル重み因子は普通短いシミュレーションのイテレーションによって逐次的に決められる. 詳細については、例えば文献 [12, 13] を参照されたい. また、Wang-Landau 法 [14] と呼ばれる手法が、マルチカノニカル重み因子決定法として特に有効であることをコメントしておく. ある十分な精度の重み因子が求まれば、長いマルチカノニカル (MC または MD) シミュレーションを一回だけ実行する. このシミュレーション中のポテンシャルエネルギーを追跡していくことによって、最小エネルギー構造が分かる. また、ポテンシャルエネルギー分布 $P_{MUCA}(E)$ のヒストグラム $N_{MUCA}(E)$ をとることによって、物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値が式 (1.99) から得られる. ここで、この式中の状態

36 第1章 拡張アンサンブルの方法

密度 $n(E)$ の最適解は次の単ヒストグラム再重法 (single-histogram reweighting techniques) [15] から与えられる (式 (1.90) 参照).

$$n(E) = \frac{N_{MUCA}(E)}{W_{MUCA}(E)}. \quad (1.101)$$

更に, (ポテンシャルエネルギーの関数で表せないものを含む) 一般の物理量 A の任意の温度 $T (= 1/k_B\beta)$ における期待値は, シミュレーションの途中の状態 (すなわち, 全粒子の座標) をある間隔で保存することによって, 以下の式で計算できる.

$$\langle A \rangle_T = \frac{\sum_{k=1}^{n_s} A(x_k) W_{MUCA}^{-1}(E(x_k)) \exp[-\beta E(x_k)]}{\sum_{k=1}^{n_s} W_{MUCA}^{-1}(E(x_k)) \exp[-\beta E(x_k)]}. \quad (1.102)$$

ここで, x_k は k 番目の MC (または, MD) のステップにおける全粒子の座標であり, n_s はその総サンプル数である. また, A がポテンシャルエネルギー E の関数として表せるときは, 式 (1.102) は式 (1.101) より, 式 (1.99) に帰着することに注意されたい.

さて, 実際の計算では, 式 (1.102) などのように, 指数関数の和を含む場合は, 多くの場合, 数値の発散 (オーバーフロー) などの問題が起こる. この問題は以下の式を繰り返し使うことによって解決される [16]. すなわち, $C = A + B$ ($A > 0, B > 0$) のとき, 以下が成り立つ.

$$\begin{aligned} \log C &= \log \left[\max(A, B) \left(1 + \frac{\min(A, B)}{\max(A, B)} \right) \right], \\ &= \max(\log A, \log B) + \log \{1 + \exp[\min(\log A, \log B) - \max(\log A, \log B)]\}. \end{aligned} \quad (1.103)$$

ここで, $\max(A, B)$ は二つの数字 A と B のうち, 大きい方を表す関数であり, 以下で定義される.

$$\max(A, B) = \begin{cases} A, & \text{if } A \geq B, \\ B, & \text{if } B \geq A. \end{cases} \quad (1.104)$$

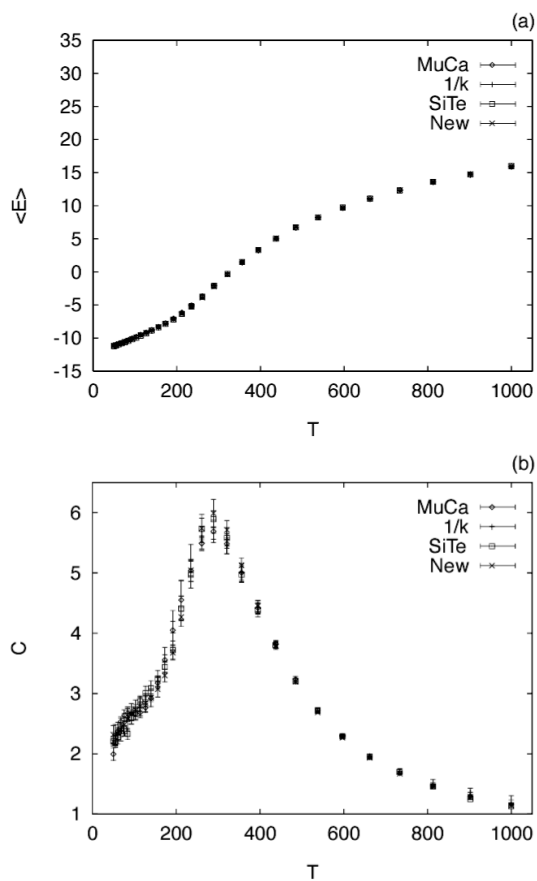


図 1.12 真空中の Met-enkephalin の系の (a) 平均ポテンシャルエネルギー $\langle E \rangle_T$ (kcal/mol) と (b) 比熱 C . マルチカノニカル法, 次節の焼き戻し法など四つの拡張アンサンブル法に基づくモンテカルロシミュレーションの結果を重ね描きした.

38 第1章 拡張アンサンブルの方法

実際のマルチカノニカルシミュレーションの例として、上の図 1.7 や図 1.9 と同じ真空中の Met-enkephalin の系においてマルチカノニカル MC シミュレーションを実行した結果を示す。図 1.11 にポテンシャルエネルギー E がシミュレーションが進むにつれて、どのように変化していくかを示した。マルチカノニカルシミュレーションは、エネルギー極小状態に留まることなく、上は $T = 1000$ K における平均値 ($E \approx 15$) を越えるとともに、下は最小エネルギー状態 ($E \approx -12$) に何度も到達し、ポテンシャルエネルギー空間の酔歩を実現していることが分かる。

この1回のマルチカノニカルシミュレーションの結果を元に式 (1.99) と式 (1.101) から、平均ポテンシャルエネルギー $\langle E \rangle_T$ および比熱 C を計算することができる。ここで、後者は (アミノ酸数 5 とボルツマン因子 k_B で割って) 次式で定義した (式 (1.23) 参照)。

$$C(T) \equiv \frac{1}{5k_B} \frac{d(\langle E \rangle_T)}{dT} = \frac{\beta^2}{5} [\langle E^2 \rangle_T - \langle E \rangle_T^2] . \quad (1.105)$$

これらの結果を図 1.12 に示した。温度 $T = 50$ K における平均ポテンシャルエネルギー値 ($E \approx -11$) が正しく得られていることに注目されたい。

また、最小ポテンシャルエネルギー E_0 は絶対温度ゼロにおけるポテンシャルエネルギーなので、非常に低温では、ポテンシャルエネルギーは、その近傍にあり、以下が成り立つ。

$$E(x) \approx E_0 + \sum_{i=1}^{N_D} \frac{1}{2} k_i (\theta_i - \theta_i^0)^2 \quad (1.106)$$

ここで、 θ_i は系の二面角であり、 N_D は二面角の総数である。また、 θ_i^0 ($i = 1, \dots, N_D$) は、以下のように、最小ポテンシャルエネルギーにおける二面角である。

$$E_0 = E(\theta_1^0, \dots, \theta_{N_D}^0) . \quad (1.107)$$

このような調和振動子項では式 (1.20) の運動エネルギー項のときと同じように、以下が成り立つ。

$$\langle E \rangle_T \approx E_0 + \frac{N_D}{2} k_B T . \quad (1.108)$$

これを式 (1.105) に代入することによって以下が得られる.

$$\lim_{T \rightarrow 0} C(T) = \frac{N_D}{10} . \quad (1.109)$$

この Met-enkephalin のシミュレーションでは, $N_D = 19$ であるので, 図 1.12(b) では, $C(T)$ が正しいゼロ温度極限 1.9 に近づいていることが確認できる.

1.3.2 焼き戻し法

次に焼き戻し法 (simulated tempering: ST) [17] を説明しよう. この方法では, 温度自体が動変数となり, シミュレーションの間に粒子の座標 (と運動量) および温度が変更されていく. 各状態は次の重み因子 $W_{ST}(E; T)$ に比例した確率で発生する.

$$W_{ST}(E; T) = e^{-\beta E + a(T)} . \quad (1.110)$$

ここで, 関数 $a(T)$ は以下のように, 温度の一様分布が得られるように導入された.

$$P_{ST}(T) = \int dE n(E) W_{ST}(E; T) = \int dE n(E) e^{-\beta E + a(T)} = \text{constant} . \quad (1.111)$$

すなわち, この手法では「温度空間」上の自由酔歩が実現される. そして, それがポテンシャルエネルギー空間上の酔歩を誘起する. これによって, マルチカノニカル法と同様に, シミュレーションがエネルギー極小状態に留まるのを避けることができるわけである.

実際の数値計算においては, 温度を M 個の値 T_m ($m = 1, \dots, M$) に離散化する. ここで, 一般性を失うことなく, $T_1 < T_2 < \dots < T_M$ と仮定できる. 最低温度 T_1 は最低エネルギー領域を探索できる程十分低くとる. また, 最高温度 T_M はエネルギー極小状態に留まらないよう, 十分高くとる必要がある. このとき, $a_m = a(T_m)$ ($m = 1, \dots, M$) として, 式 (1.110) の重み因子は,

$$W_{ST}(E; T_m) = e^{-\beta_m E + a_m} , \quad (1.112)$$

40 第1章 拡張アンサンブルの方法

と書けるので、式 (1.111) と式 (1.112) から次のことが言える.

$$e^{-a_m} \propto \int dE n(E) e^{-\beta_m E} = Z_q(T_m). \quad (1.113)$$

すなわち、焼き戻し法のパラメータ a_m は温度 T_m における「無次元化されたヘルムホルツ自由エネルギー」(すなわち、 β_m とヘルムホルツ自由エネルギーの積) である (式 (1.6) 参照. ここでいうヘルムホルツ自由エネルギーとは、運動量の寄与を無視したものである).

上のパラメータ a_m ($m = 1, \dots, M$) は、マルチカノニカル法のとおりと同じように、普通短いシミュレーションのイテレーションによって決められる. その手順の詳細については、例えば、文献 [18, 19, 13] を参照されたい. 重み因子が求まれば、焼き戻しシミュレーションは、次の二つのステップを交互に繰り返すことで実行される.

1. 一定温度 T_m のカノニカルアンサンブル上の従来の MC または MD シミュレーションをあるステップ数実行する.
2. 状態 (座標と運動量) を固定したまま、温度 T_m を隣接温度 $T_{m\pm 1}$ に更新する. この更新の遷移確率は、次のようにメトロポリス判定で与えられる.

$$w(T_m \rightarrow T_{m\pm 1}) = \min \left(1, \frac{W_{ST}(E; T_{m\pm 1})}{W_{ST}(E; T_m)} \right) = \min (1, \exp(-\Delta)) . \quad (1.114)$$

ここで、 Δ は次で与えられる.

$$\Delta = (\beta_{m\pm 1} - \beta_m) E - (a_{m\pm 1} - a_m) . \quad (1.115)$$

ステップ2では、温度更新の受け入れ確率が十分高いことを保証するために、隣接する温度値へのみ更新することに注意されたい (サンプリングの効率が悪くなることを除けば、原理的には離れた温度値へ更新しても良いわけである).

ステップ1で分子動力学法を採用する場合、運動量 p も考慮する必要がある. その場合、ステップ2で温度 $T_{m0\pm 1}$ が受諾されるときに、全粒子の運動量を以下のように変更させる

[20, 21, 22].

$$p_k' = \sqrt{\frac{T_{m_0 \pm 1}}{T_{m_0}}} p_k. \quad (1.116)$$

すると、運動エネルギー項が式 (1.114) で打ち消し合って、ステップ2のメトロポリス判定でモンテカルロ法でも分子動力学法でも同じ式 (1.115) の Δ を使うことになる。同様の運動量スケール法が他の熱浴法でも導かれている [23].

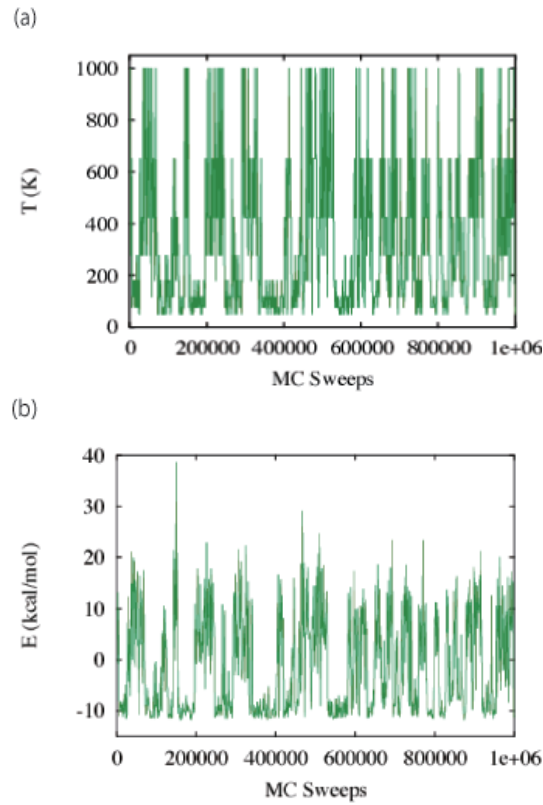


図 1.13 真空中の Met-enkephalin の焼き戻しモンテカルロシミュレーションにおける (a) 温度 T (K) と (b) ポテンシャルエネルギー E (kcal/mol) の「時間発展」.

長い焼き戻しシミュレーションの後、物理量 A の温度 T_m ($m = 1, \dots, M$) に

42 第1章 拡張アンサンブルの方法

におけるカノニカルアンサンブル平均は、以下の算術平均で与えられる (式 (1.86) 参照).

$$\langle A \rangle_{T_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)) , \quad (1.117)$$

ここで, $x_m(k)$ ($k = 1, \dots, n_m$) は温度 T_m における k 番目の MC (または, MD) のステップにおける状態 (全粒子の座標) であり, n_m はその総サンプル数である.

また, 任意の中間の温度 T ($= 1/k_B\beta$) における A のカノニカルアンサンブル平均は, 式 (1.99) で与えられるが, その中の状態密度 $n(E)$ の最適解は以下の多ヒストグラム再重法 (multiple-histogram reweighting techniques) [24, 25] で与えられる (化学の分野では, Weighted Histogram Analysis Method: WHAM とも呼ばれている [25]). すなわち, 焼き戻しシミュレーションにおいて, 温度 T_m ($= 1/k_B\beta_m$) ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を, それぞれ $N_m(E)$ と n_m とする. このとき, 状態密度 $n(E)$ の最適解は次の二つの WHAM 方程式を自己無撞着に逐次的に解くことによって得られる.

$$n(E) = \frac{\sum_{m=1}^M g_m^{-1} N_m(E)}{\sum_{m=1}^M g_m^{-1} n_m e^{f_m - \beta_m E}} , \quad (1.118)$$

$$e^{-f_m} = \sum_E n(E) e^{-\beta_m E} , \quad (m = 1, \dots, M) . \quad (1.119)$$

ここで, $g_m = 1 + 2\tau_m$ であり, τ_m は温度 T_m における積分された自己相関時間^{†1}である. しかし, 多くの場合, この量は幅広い範囲において温度によらず定数とおいても多ヒストグラム再重法の結果に影響がないことが示されてお

^{†1}ある温度 T において物理量 A をいろいろな時刻で観測するとき, その規格化された自己相関関数は $\phi_A(t) = C_A(t)/C_A(0)$ で与えられる. ここで, $C_A(t) = \langle A(0)A(t) \rangle_T - \langle A(0) \rangle_T^2$ である. 「積分された自己相関時間」は $\tau_A = \int_0^\infty \phi_A(t) dt$ で定義される. ここでは, 物理量としてポテンシャルエネルギー E を考える. よって, 式 (1.118) 中の $g_m (= 1 + 2\tau_m)$ は温度 T_m に依存する量である.

り [25], 一般に式 (1.118) では g_m^{-1} が落ちた形で計算される (つまり, $g_m = 1$ とおく). 以下でも, この変数は省略することにする. また, f_m は無次元化されたヘルムホルツ自由エネルギーである. 上の方程式は例えば次のように逐次的に解かれる. 先ず, 全ての f_m ($m = 1, \dots, M$) を 0 にセットする. 次に式 (1.118) を使って, $n(E)$ を求める. 得られた $n(E)$ を式 (1.119) に代入して, 新しい f_m の値を得る. この操作を $n(E)$ と f_m の値が収束するまで繰り返す.

更に, (ポテンシャルエネルギーの関数で表せないものを含む) 一般の物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値は, シミュレーションの途中の状態 (全粒子の座標) をある間隔で保存することによって, 以下のように計算することができる. まず, 式 (1.118) と式 (1.119) を自己無撞着に逐次的に解くことによって, f_m ($m = 1, \dots, M$) を得る. すると, 以下の式で任意の温度における期待値が求められる [26].

$$\langle A \rangle_T = \frac{\sum_{m=1}^M \sum_{k=1}^{n_m} A(x_m(k)) \frac{1}{\sum_{\ell=1}^M n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}{\sum_{m=1}^M \sum_{k=1}^{n_m} \frac{1}{\sum_{\ell=1}^M n_\ell \exp[f_\ell - \beta_\ell E(x_m(k))]} \exp[-\beta E(x_m(k))]}, \quad (1.120)$$

ここで, $x_m(k)$ ($k = 1, \dots, n_m$) は温度 T_m における k 番目の MC (または, MD) のステップにおける全粒子の座標であり, n_m はその総数である.

実際の焼き戻しシミュレーションの例として, 上と同じ真空中の Met-enkephalin の系において焼き戻し MC シミュレーションを実行した結果を示す. 図 1.13 に温度 T とポテンシャルエネルギー E がシミュレーションが進むにつれて, どのように変化して行くかを示した. 図から分かるように, 温度の酔歩が得られているが, それに従って, ポテンシャルエネルギーもマルチカノニカル法のとときと同じように酔歩が実現されている. これによって, シミュレーションがエネルギー極小状態に留まらずに, 幅広い状態空間がサンプルされていることが分かる. また, 温度とポテンシャルエネルギーの変化に強い正の相関があるこ

44 第1章 拡張アンサンブルの方法

とに注目されたい.

最後に焼き戻し法における M 個の温度をどのように分布させるかについてコメントする. 上では, 最低温度 T_1 は最低エネルギー領域を探索できる程十分低くとり, 最高温度 T_M はエネルギー極小状態に留まらないよう, 十分高くとる必要があると述べたが, それではそれ以外の温度はどうするのかということである. 基準としては, 焼き戻しシミュレーションのステップ2において, どの隣り合う温度対についても, 温度の更新が受け入れられる確率が一様に約0.2 (式 (1.78) 参照) になるように温度を分布させたいわけである. 経験的には, 系に強い相転移がない場合, 徐冷法と同じように, 以下のように指数関数的に温度を分布させるとうまくいくことが多い (式 (1.88), (1.89) 参照) [19].

$$T_m = T_1 \gamma^{m-1}, \quad (m = 1, \dots, M) \quad (1.121)$$

ここで, 定数 γ は, 最低温度 T_1 と最高温度 T_M によって, 以下で与えられる.

$$\gamma = \left(\frac{T_M}{T_1} \right)^{\frac{1}{M-1}}. \quad (1.122)$$

1.3.3 レプリカ交換法

レプリカ交換法 (replica-exchange method: REM) [27] は焼き戻し法 (simulated tempering) の改良版として開発された (よって, parallel tempering [28] とも呼ばれる).

レプリカ交換法における系は, M 個の違う温度 T_m ($m = 1, \dots, M$) をもつ, 相互作用しない M 個の独立なコピー (レプリカ) から成る. それぞれのレプリカには常に一つの温度が対応するとする. よって, レプリカと温度には1対1対応が成り立ち, レプリカを表すラベルを i ($= 1, \dots, M$), 温度を表すラベルを m ($= 1, \dots, M$) とするとき, i と m は置換関係にある.

$$\begin{cases} i = i(m) \equiv f(m), \\ m = m(i) \equiv f^{-1}(i). \end{cases} \quad (1.123)$$

ここで, $f(m)$ は m の置換であり, $f^{-1}(i)$ はその逆置換である.

1.3 拡張アンサンブル法 45

この系における「状態」を $X = \{x_1^{[i(1)]}, \dots, x_M^{[i(M)]}\} = \{x_{m(1)}^{[1]}, \dots, x_{m(M)}^{[M]}\}$ で表すことにしよう．すなわち，系の状態 X は各レプリカの状態

$$x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m \quad (1.124)$$

により指定される．ここで，上付きのラベル i はレプリカを表し，下付きのラベル m は温度を表す．そして，温度 T_m にあるレプリカ i の状態 $x_m^{[i]}$ は N 個の原子の座標 $q^{[i]}$ と運動量 $p^{[i]}$ により決まる．

レプリカ同士は相互作用しないので，状態 X の重み因子はボルツマン因子の積で以下のように与えられる．

$$\begin{aligned} W_{REM}(X) &= \prod_{i=1}^M \exp \left\{ -\beta_{m(i)} H(q^{[i]}, p^{[i]}) \right\} = \prod_{m=1}^M \exp \left\{ -\beta_m H(q^{[i(m)]}, p^{[i(m)]}) \right\} \\ &= \exp \left\{ -\sum_{i=1}^M \beta_{m(i)} H(q^{[i]}, p^{[i]}) \right\} = \exp \left\{ -\sum_{m=1}^M \beta_m H(q^{[i(m)]}, p^{[i(m)]}) \right\}, \end{aligned} \quad (1.125)$$

ここで， $i(m)$ と $m(i)$ は式 (1.123) の置換関数である．

温度がそれぞれ T_m と T_n に対応するレプリカ i と j を交換することを考える．

$$X = \{\dots, x_m^{[i]}, \dots, x_n^{[j]}, \dots\} \longrightarrow X' = \{\dots, x_m^{[j]'}, \dots, x_n^{[i]'}, \dots\}. \quad (1.126)$$

ここで， i, j, m, n は式 (1.123) の置換関数による関係を持つ．よって，レプリカの交換は以下のような新しい置換関数 f' を導入することになる．

$$\begin{cases} i = f(m) \longrightarrow j = f'(m), \\ j = f(n) \longrightarrow i = f'(n). \end{cases} \quad (1.127)$$

このレプリカの交換は次のように書くことができる．

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m \longrightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]'})_m, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n \longrightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]'})_n. \end{cases} \quad (1.128)$$

ここで， $p^{[i]'}$ と $p^{[j]'}$ の定義は後ほど与えることにする．ちなみに，この操作はレプリカがそれぞれ i と j に対応する温度 T_m と T_n を交換することと同等で

46 第1章 拡張アンサンブルの方法

ある．よって，式 (1.128) は次のように書いても良い．

$$\begin{cases} x_m^{[i]} \equiv (q^{[i]}, p^{[i]})_m \longrightarrow x_n^{[i]'} \equiv (q^{[i]}, p^{[i]'})_n, \\ x_n^{[j]} \equiv (q^{[j]}, p^{[j]})_n \longrightarrow x_m^{[j]'} \equiv (q^{[j]}, p^{[j]'})_m. \end{cases} \quad (1.129)$$

レプリカ交換法が開発されたときは，各レプリカの状態の更新はモンテカルロ法に基づいていた．よって，座標 q （そして，ポテンシャルエネルギー $E(q)$ ）のみを考慮すれば済んだ．一方，各状態の更新が分子動力学法に基づく場合，運動量 p も考慮する必要が出てくる．運動量と運動エネルギーの関係を表す式 (1.20) から，式 (1.128)（と式 (1.129)）にある運動量は次のように指定すれば良いことが分かる [20]．

$$\begin{cases} p^{[i]'} \equiv \sqrt{\frac{T_n}{T_m}} p^{[i]}, \\ p^{[j]'} \equiv \sqrt{\frac{T_m}{T_n}} p^{[j]}. \end{cases} \quad (1.130)$$

すなわち，式 (1.20) が満たされるように，各原子の速度を二つの温度の比の平方根を掛けることによって，一様にリスケールすれば良いことになる．同様の運動量スケールリング法がいろいろな熱浴法で導かれている [23]．

このレプリカ対の交換の遷移確率 $w(X \rightarrow X')$ はいつものメトロポリス判定で得られ，以下で与えられる．

$$w(X \rightarrow X') \equiv w(x_m^{[i]} | x_n^{[j]}) = \min \left(1, \frac{W_{REM}(X')}{W_{REM}(X)} \right) = \min(1, \exp(-\Delta)), \quad (1.131)$$

ここで，2つ目の表現 $w(x_m^{[i]} | x_n^{[j]})$ では，交換されるレプリカ対（と温度対）を陽に書いた．式 (1.1)，(1.2)，(1.125)，(1.130) より以下が得られる．

$$\begin{aligned} \frac{W_{REM}(X')}{W_{REM}(X)} &= \exp \left\{ -\beta_m [K(p^{[j]'}) + E(q^{[j]})] - \beta_n [K(p^{[i]'}) + E(q^{[i]})] \right. \\ &\quad \left. + \beta_m [K(p^{[i]}) + E(q^{[i]})] + \beta_n [K(p^{[j]}) + E(q^{[j]})] \right\} \\ &= \exp \left\{ -\beta_m \frac{T_m}{T_n} K(p^{[j]}) - \beta_n \frac{T_n}{T_m} K(p^{[i]}) + \beta_m K(p^{[i]}) + \beta_n K(p^{[j]}) \right. \\ &\quad \left. - \beta_m [E(q^{[j]}) - E(q^{[i]})] - \beta_n [E(q^{[i]}) - E(q^{[j]})] \right\} \\ &= \exp(-\Delta). \end{aligned}$$

$$(1.132)$$

よって、 Δ は次で与えられる.

$$\Delta = \beta_m \left(E(q^{[j]}) - E(q^{[i]}) \right) - \beta_n \left(E(q^{[j]}) - E(q^{[i]}) \right), \quad (1.133)$$

$$= (\beta_m - \beta_n) \left(E(q^{[j]}) - E(q^{[i]}) \right). \quad (1.134)$$

すなわち、運動エネルギー項は全て打ち消しあって、ポテンシャルエネルギーだけの式になっている [20]. ラベル i, j, m, n は交換前の置換関数 (式 (1.123)) で次のように関係している.

$$\begin{cases} i = f(m), \\ j = f(n). \end{cases} \quad (1.135)$$

レプリカ交換シミュレーションは、次の二つのステップを交互に繰り返すことで実行される ($T_1 < T_2 < \dots < T_M$ とする).

1. 各レプリカ i ($i = 1, \dots, M$) において、一定温度 T_m ($m = 1, \dots, M$) の独立なカノニカルアンサンブル上の従来の MC または MD シミュレーションを同時にあるステップ数実行する.
2. 隣接した温度 (T_m と T_{m+1}) に対応するレプリカ対 (i と j) を式 (1.131) (と式 (1.134)) の遷移確率 $w(x_m^{[i]} | x_{m+1}^{[j]})$ に従って交換する.

ステップ2では隣り合う温度値に対応するレプリカ対を交換することに注意されたい. これは、レプリカ交換の受け入れ確率が温度差に従って指数関数的に減少するので、できるだけレプリカ交換の受諾率が極端に小さくならないようにしたいからである (式 (1.131), (1.134) 参照). また、ステップ2において、レプリカ交換が採択されるたびに、式 (1.123) の置換関数が更新されることになる.

レプリカ交換シミュレーションは並列計算機に特に適している. 各レプリカを各計算機 (CPU) に対応させることによって、計算機間の情報通信量を最小にできるわけである (座標と運動量の代わりに温度を交換するだけなので). これは、シミュレーションの間に、MC または MD ステップ t の関数として置換

48 第1章 拡張アンサンブルの方法

関数 $m(i; t) = f^{-1}(i; t)$ を記録していくことになる．あるステップ数の並列の MC または MD シミュレーション（ステップ1）の後，隣り合う温度に対応する $M/2$ 組のレプリカ対を同時に交換する（ステップ2）．そして，2通りの組み合わせを交代させる．ここで，2通りの組み合わせとは， $(T_1, T_2), (T_3, T_4), \dots$ と $(T_2, T_3), (T_4, T_5), \dots$ のことである．

焼き戻し法の場合と同じように，長いレプリカ交換シミュレーションの後，物理量 A の温度 T_m ($m = 1, \dots, M$) におけるカノニカルアンサンブル平均は，式 (1.117) の算術平均で与えられる．

また，任意の中間の温度 $T (= 1/k_B\beta)$ における A のカノニカルアンサンブル平均は，式 (1.99) で与えられるが，その式中の状態密度 $n(E)$ の最適解は，焼き戻しシミュレーションのときと同じように，多ヒストグラム再重法で与えられる．すなわち，レプリカ交換シミュレーションにおいて，温度 $T_m = 1/k_B\beta_m$ ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を，それぞれ $N_m(E)$ と n_m とする．このとき，状態密度 $n(E)$ の最適解は式 (1.118) と式 (1.119) の二つの WHAM 方程式を自己無撞着に逐次的に解くことによって得られる．ここで，これらの式中のヒストグラム $N_m(E)$ は，レプリカごとではなく，温度 $T_m (= 1/k_B\beta_m)$ ($m = 1, \dots, M$) ごとに得られたものを使うことに注意されたい．

更に，（ポテンシャルエネルギーの関数で表せないものを含む）一般の物理量 A の任意の温度 $T (= 1/k_B\beta)$ における期待値は，式 (1.120) から計算することができる．

ここで，レプリカ交換法の大きな利点として強調しておきたいのは，マルチカノニカル法や焼き戻し法では最初に重み因子を短い試行シミュレーションで決めなければならないのに対し，レプリカ交換法では重み因子はボルツマン因子の積なので最初から分かっていることである．しかし，前者では1個のレプリカでシミュレーションを行うのに対し，後者では必要とするレプリカ数が，系の自由度の数 N が増加するに従って，その平方根 $\sqrt{N}$ に比例して増えてしまう（よって，計算時間が増えてしまう）という困難が存在する [27]．

さて，実際のレプリカ交換シミュレーションの例として，上と同じ Met-

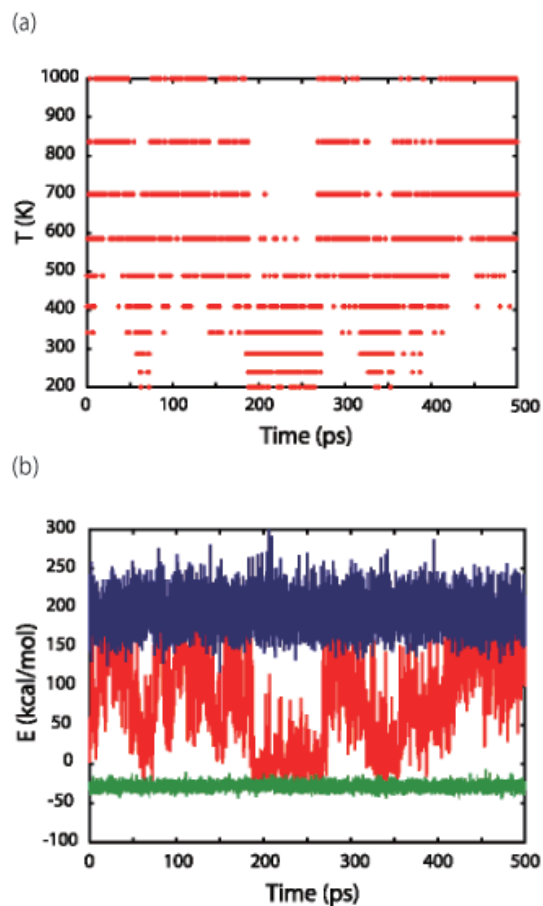


図 1.14 真空中の Met-enkephalin のレプリカ交換分子動力学シミュレーションの一つのレプリカにおける (a) 温度 T (K) と (b) ポテンシャルエネルギー E (kcal/mol) の「時間発展」. (b) では、温度が $T_1 = 200$ K (緑色) と $T_{10} = 1000$ K (青色) における従来のカノニカルシミュレーションの結果も重ねて示した.

50 第1章 拡張アンサンブルの方法

enkephalin の系においてレプリカ交換分子動力学 (replica-exchange molecular dynamics: REMD) [20] シミュレーションを実行した結果を示す. レプリカ数は 10 であり, 温度は最低温度が $T_1 = 200$ K, 最高温度が $T_{10} = 1000$ K であった. 図 1.14 に一つのレプリカにおける温度 T とポテンシャルエネルギー E がシミュレーションが進むにつれて, どのように変化していくかを示した. 真空中の同じペプチドの系ではあるが, 今回は MC 法ではなく, レプリカ交換分子動力学法の結果であり, ポテンシャルエネルギー関数は上のモンテカルロ法の場合と違ったものを使っているため, ポテンシャルエネルギー値は以前のもの (図 1.7, 図 1.9, 図 1.11, 図 1.12, 図 1.13 参照) とは直接には比較できない. しかし, 図から分かるように, 温度の十分な酔歩が得られており, それに従って, ポテンシャルエネルギーもマルチカノニカル法や焼き戻し法のときと同じように酔歩が実現されている. これによって, シミュレーションがエネルギー極小状態に留まらずに, 幅広い構造空間がサンプルされていることが分かる. また, 焼き戻し法のときと同じように, 温度とポテンシャルエネルギーの変化に強い正の相関があることが分かる.

最後にレプリカ交換法における M 個の温度をどのように分布させるかについてコメントする. 基準としては, レプリカ交換シミュレーションのステップ 2 において, どの隣り合う温度対についても, レプリカ交換が受け入れられる確率が一様に約 0.2 (式 (1.78) 参照) になるように温度を分布させたいわけである. 経験的には, 系に強い相転移がない場合, (徐冷法や) 焼き戻し法と同じように, 指数関数的に温度を分布させるとうまくいくことが多い (式 (1.121), (1.122) 参照) [20]. 系に強い相転移があって, 焼き戻し法やレプリカ交換法において, この指数関数的な温度分布では, 温度更新や温度交換の遷移確率が一様でなく, (相転移点近傍で) 大きく下がるような場合は, 遷移確率の低い領域に温度分布を密にするように温度の再分布をすれば良く, 多くの場合, 何回かの試行による手による温度再分布によって, 満足のいく一様な遷移確率が得られる (より厳密で系統的な最適な温度分布の求め方としては, 例えば, 文献 [29] を参照されたい).

1.4 マルチカノニカル法, 焼き戻し法, レプリカ交換法の合体

前節では, 良く知られている三つの拡張アンサンブル法 (マルチカノニカル法, 焼き戻し法, レプリカ交換法) を紹介した. マルチカノニカル法や焼き戻し法などに比べて, レプリカ交換法の利点は重み因子が最初から分かっていることである (すなわち, ボルツマン因子の積である). 一方, マルチカノニカル法や焼き戻し法では, 重み因子の決定のために, しばしば大変な労力を必要とする. しかし, レプリカ交換法では, 系の自由度が大きくなるにつれて, 非常に多くのレプリカ (よって多くの計算時間) を必要とするという困難が存在する. 本節では, マルチカノニカル法や焼き戻し法の難点 (重み因子の決定が困難) とレプリカ交換法の難点 (多くの計算時間が必要) の両方を克服する拡張アンサンブル法を紹介する.

1.4.1 レプリカ交換マルチカノニカル法とレプリカ交換焼き戻し法

最初の手法はレプリカ交換マルチカノニカル法 (replica-exchange multicanonical algorithm: REMUCA) [30, 26, 31] である. この手法では, まず短い (レプリカ数 M の) レプリカ交換シミュレーションを実行し, その結果からマルチカノニカル重み因子を決め, その後で長いマルチカノニカルシミュレーションを一回だけ実行する. 式 (1.91) から分かるように, マルチカノニカル重み因子は状態密度 $n(E)$ と等価である. そして, $n(E)$ は最初のレプリカ交換シミュレーションの結果に多ヒストグラム再重法を使うことによって次のように求めることができる. このレプリカ交換シミュレーションにおいて, 温度 $T_m (= 1/k_B\beta_m)$ ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を, それぞれ $N_m(E)$ と n_m とする. このとき, 状態密度 $n(E)$ の最適解は式 (1.118) と式 (1.119) の二つの WHAM 方程式を自己無撞着に逐次的に解くことによって得られる.

状態密度 $n(E)$ が得られれば, マルチカノニカル重み因子は式 (1.91) から直接得られる (式 (1.92) 参照). 実際には, そのようにして決められた状態密度 $n(E)$ とマルチカノニカルポテンシャルエネルギー $E_{MUC}(E; T_0)$ は以下の範

52 第1章 拡張アンサンブルの方法

囲内のみ信頼できるといえる.

$$E_1 \leq E \leq E_M, \quad (1.136)$$

ここで, E_1 と E_M は以下で定義される.

$$\begin{cases} E_1 = \langle E \rangle_{T_1}, \\ E_M = \langle E \rangle_{T_M}, \end{cases} \quad (1.137)$$

また, T_1 と T_M は, この最初のレプリカ交換シミュレーションにおける, それぞれ最低温度と最高温度である. この範囲外では, マルチカノニカルポテンシャルエネルギーを以下のように線形に外挿する [30].

$$\mathcal{E}_{MUCA}(E) \equiv \begin{cases} \left. \frac{\partial E_{MUCA}(E; T_0)}{\partial E} \right|_{E=E_1} (E - E_1) + E_{MUCA}(E_1; T_0), & \text{for } E < E_1, \\ E_{MUCA}(E; T_0), & \text{for } E_1 \leq E \leq E_M, \\ \left. \frac{\partial E_{MUCA}(E; T_0)}{\partial E} \right|_{E=E_M} (E - E_M) + E_{MUCA}(E_M; T_0), & \text{for } E > E_M. \end{cases} \quad (1.138)$$

モンテカルロ法においては, REMUCA の重み因子は式(1.138) の $\mathcal{E}_{MUCA}$ を式 (1.91) の E_{MUCA} に代入することによって以下のように得られる [30, 26].

$$W_{MUCA}(E) = \exp[-\beta_0 \mathcal{E}_{MUCA}(E)] = \begin{cases} \exp(-\beta_1 E), & \text{for } E < E_1, \\ \frac{1}{n(E)}, & \text{for } E_1 \leq E \leq E_M, \\ \exp(-\beta_M E), & \text{for } E > E_M. \end{cases} \quad (1.139)$$

最後に実行する長いマルチカノニカル MC または MD シミュレーションは, それぞれ, 式 (1.94) のメトロポリス判定または式 (1.98) の「修正されたニュートン方程式」によって実行される. 後者においては, 式 (1.138) の $\mathcal{E}_{MUCA}(E)$ を式 (1.98) の $E_{MUCA}(E; T_0)$ に代入することになる. これらのシミュレーションの結果, 式 (1.136) の範囲内で一様なポテンシャルエネルギー分布が得られることが期待される. また, $E < E_1$ では温度 T_1 の, $E > E_M$ では温度 T_M

1.4 マルチカノニカル法, 焼き戻し法, レプリカ交換法の合体 53

のカノニカル分布が得られることになる。そして、このシミュレーション結果は、式 (1.101) (と式 (1.99)) の単ヒストグラム再重法で解析される。

更に、(ポテンシャルエネルギーの関数で表せないものを含む) 一般の物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値は、式 (1.102) によって計算できる。

二つ目の手法はレプリカ交換焼き戻し法 (replica-exchange simulated tempering: REST) [32] である。この手法では、REMUCA と同じように、まず短い (レプリカ数 M の) レプリカ交換シミュレーションを実行し、その結果から焼き戻し法の重み因子を決め、その後で長い焼き戻しシミュレーションを一回だけ実行する。式 (1.113) から分かるように、焼き戻し法の重み因子のパラメータ a_m は無次元化されたヘルムホルツ自由エネルギー f_m と等価である。そして、 f_m は最初の REM シミュレーションの結果に多ヒストグラム再重法を適用することによって次のように求めることができる。このレプリカ交換シミュレーションにおいて、温度 T_m ($= 1/k_B\beta_m$) ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を、それぞれ $N_m(E)$ と n_m とする。このとき、無次元化されたヘルムホルツ自由エネルギー f_m の最適解は式 (1.118) と式 (1.119) の二つの WHAM 方程式を自己無撞着に逐次的に解くことによって得られる。

無次元化されたヘルムホルツ自由エネルギー f_m が得られれば、焼き戻し重み因子は式 (1.112) から直接得られる。すなわち、 $a_m = f_m$ とおく (式 (1.113) と式 (1.119) 参照)。

最後に実行する長い焼き戻しシミュレーションの結果を解析するために、温度 T_m ($= 1/k_B\beta_m$) ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を、それぞれ $N_m(E)$ と n_m とする。式 (1.118) と式 (1.119) の多ヒストグラム再重法を再び適用するにより、状態密度 $n(E)$ の最適解が得られる。これを式 (1.99) に代入することによって、物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値を精度良く求めることができる。

更に、(ポテンシャルエネルギーの関数で表せないものを含む) 一般の物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値は、式 (1.120) によって計算できる。

54 第1章 拡張アンサンブルの方法

これらの手法は、原理が簡単でほとんど自明といえることができる。しかし、実際の数値計算における改善は極めて大きいものであり、有効な手法といえることができる。

1.4.2 マルチカノニカルレプリカ交換法と焼き戻しレプリカ交換法

更なる改良版として、マルチカノニカルレプリカ交換法 (multicanonical replica-exchange method: MUCAREM) [30, 26, 31] を紹介しよう。この手法では、(REMUCA などによって) マルチカノニカル重み因子が与えられた後、少ないレプリカによるマルチカノニカルアンサンブル上のレプリカ交換シミュレーションを実行する。

MUCAREM では最後の長いシミュレーションは、各レプリカが異なったエネルギー領域をカバーするマルチカノニカルシミュレーションを実行するレプリカ交換シミュレーションである。

このレプリカ交換シミュレーションのレプリカ数を M とする。ここでは、それぞれのレプリカには (温度ではなく) 異なったエネルギー領域をカバーする一つのマルチカノニカル重み因子が対応する。マルチカノニカルシミュレーションは従来のカノニカルシミュレーションより幅広いエネルギー領域をカバーするので、MUCAREM に必要なレプリカ数は普通の REM よりはるかに少なく済む (つまり、 $M \ll M$ が成り立つ)。ここで、 M は普通のレプリカ交換シミュレーションで MUCAREM と同じエネルギー領域をカバーするのに必要なレプリカ数のことである。この拡張アンサンブルの重み因子は以下で与えられる (式 (1.125) 参照)。

$$W_{MUCAREM}(X) = \prod_{i=1}^M W_{MUCA}^{\{m(i)\}}(E(x_{m(i)}^{[i]})) = \prod_{m=1}^M W_{MUCA}^{\{m\}}(E(x_m^{[i(m)]})), \quad (1.140)$$

ここで、マルチカノニカル重み因子 (よって、状態密度) は違う領域について別々に準備する (式 (1.91) 参照)。

$$W_{MUCA}^{\{m\}}(E(x_m^{[i]})) = \exp[-\beta_m \mathcal{E}_{MUCA}^{\{m\}}(E(x_m^{[i]}))] \equiv \frac{1}{n^{\{m\}}(E(x_m^{[i]}))}.$$

(1.141)

ここで, $\mathcal{M}$ 個の任意の参照温度 $T_m (= 1/k_B\beta_m)$ ($m = 1, \dots, \mathcal{M}$) を用意したが, 式(1.141)の2つ目の等式から分かるように, 最終結果はこれらの温度の値には依らないはずである(これらの参照温度は分子動力学法に基づく場合のみ必要である).

それぞれのマルチカノニカル重み因子 $W_{MUCA}^{\{m\}}(E)$ または状態密度 $n^{\{m\}}(E)$ は以下で定義される. それぞれの $m (= 1, \dots, \mathcal{M})$ の値について, 温度対 $(T_L^{\{m\}}, T_H^{\{m\}})$ を用意する. ここで, $T_L^{\{m\}} < T_H^{\{m\}}$ を仮定し, 温度分布が隣り合う領域において十分な重なりを持つようにする. 今, 一般性を失うことなく, $T_L^{\{1\}} < \dots < T_L^{\{\mathcal{M}\}}$ と $T_H^{\{1\}} < \dots < T_H^{\{\mathcal{M}\}}$ を仮定することができる. そして, 以下の量を定義する.

$$\begin{cases} E_L^{\{m\}} = < E >_{T_L^{\{m\}}} , \\ E_H^{\{m\}} = < E >_{T_H^{\{m\}}} , \end{cases} \quad (m = 1, \dots, \mathcal{M}). \quad (1.142)$$

今, 対象としている全エネルギー領域 ($E_L^{\{1\}} < E < E_H^{\{\mathcal{M}\}}$) において, マルチカノニカル重み因子 $W_{MUCA}(E)$ (または, 式(1.92)のマルチカノニカルポテンシャルエネルギー $E_{MUCA}(E; T_0)$) が得られているとする. すると, それぞれの $m (= 1, \dots, \mathcal{M})$ について, 以下のマルチカノニカルポテンシャルエネルギーが定義できる [30].

$$\mathcal{E}_{MUCA}^{\{m\}}(E) = \begin{cases} \frac{\partial E_{MUCA}(E_L^{\{m\}}; T_m)}{\partial E} (E - E_L^{\{m\}}) + E_{MUCA}(E_L^{\{m\}}; T_m) , & \text{for } E < E_L^{\{m\}}, \\ E_{MUCA}(E; T_m) , & \text{for } E_L^{\{m\}} \leq E \leq E_H^{\{m\}}, \\ \frac{\partial E_{MUCA}(E_H^{\{m\}}; T_m)}{\partial E} (E - E_H^{\{m\}}) + E_{MUCA}(E_H^{\{m\}}; T_m) , & \text{for } E > E_H^{\{m\}}. \end{cases} \quad (1.143)$$

モンテカルロ法においては, MUCAREM の重み因子は式(1.143)の $\mathcal{E}_{MUCA}^{\{m\}}$

56 第1章 拡張アンサンブルの方法

を式 (1.141) に代入することによって以下のように得られる [30, 26].

$$W_{MUCA}^{\{m\}}(E) = \begin{cases} \exp(-\beta_L^{\{m\}} E), & \text{for } E < E_L^{\{m\}}, \\ \frac{1}{n(E)}, & \text{for } E_L^{\{m\}} \leq E \leq E_H^{\{m\}}, \\ \exp(-\beta_H^{\{m\}} E), & \text{for } E > E_H^{\{m\}}. \end{cases} \quad (1.144)$$

ここで, $\beta_L^{\{m\}} = 1/k_B T_L^{\{m\}}$, $\beta_H^{\{m\}} = 1/k_B T_H^{\{m\}}$ である.

MUCAREM シミュレーションは次の二つのステップを交互に繰り返すことによって実行される.

1. 各レプリカ i ($i = 1, \dots, M$) において, それぞれ, 一定のマルチカノニカル重み因子 $W_{MUCA}^{\{m\}}(E)$ ($m = 1, \dots, M$) に基づく MC または MD シミュレーションを同時にあるステップ数実行する.
2. (m 番目と $(m+1)$ 番目の) 隣接したマルチカノニカルアンサンブルに対応するレプリカ対 (i と j) を交換する. つまり, $X = \{\dots, x_m^{[i]}, \dots, x_{m+1}^{[j]}, \dots\} \rightarrow X' = \{\dots, x_m^{[j]}, \dots, x_{m+1}^{[i]}, \dots\}$ を考える. このレプリカ交換の遷移確率は以下のメトロポリス判定で与えられる.

$$w(X \rightarrow X') = \min(1, \exp(-\Delta)). \quad (1.145)$$

ここで, 式 (1.133) から以下が成り立つ [30].

$$\Delta = \beta_m \left\{ \mathcal{E}_{MUCA}^{\{m\}}(E(q^{[j]})) - \mathcal{E}_{MUCA}^{\{m\}}(E(q^{[i]})) \right\} - \beta_{m+1} \left\{ \mathcal{E}_{MUCA}^{\{m+1\}}(E(q^{[j]})) - \mathcal{E}_{MUCA}^{\{m+1\}}(E(q^{[i]})) \right\}. \quad (1.146)$$

また, $E(q^{[i]})$ と $E(q^{[j]})$ はそれぞれ i 番目のレプリカと j 番目のレプリカのポテンシャルエネルギーである.

$m \neq n$ の場合, $\mathcal{E}_{MUCA}^{\{m\}}(E)$ と $\mathcal{E}_{MUCA}^{\{n\}}(E)$ は一般に異なる関数なので, 式 (1.146) において, マルチカノニカルポテンシャルエネルギー $\mathcal{E}_{MUCA}^{\{m\}}(E(q^{[j]}))$ と $\mathcal{E}_{MUCA}^{\{m+1\}}(E(q^{[i]}))$ を新たに計算しなければならないことに注意されたい.

この方法においては, m 番目のマルチカノニカルアンサンブルは, 実際には, $E < E_L^{\{m\}}$ の領域では温度 $T = T_L^{\{m\}}$ のカノニカルシミュレーション,

1.4 マルチカノニカル法, 焼き戻し法, レプリカ交換法の合体 57

$E_L^{\{m\}} \leq E \leq E_H^{\{m\}}$ の領域ではマルチカノニカルシミュレーション, そして, $E > E_H^{\{m\}}$ の領域では温度 $T = T_H^{\{m\}}$ のカノニカルシミュレーションを実行することになる. 一方, レプリカ交換によって, 全エネルギー領域 ($E_L^{\{1\}} \leq E \leq E_H^{\{M\}}$) がサンプルされることになる.

任意の中間の温度 T におけるカノニカル分布を得るためには, 多ヒストグラム再重法がここでも利用される. マルチカノニカル重み因子 $W_{MUCA}^{\{m\}}(E)$ ($m = 1, \dots, M$) で得られたポテンシャルエネルギー分布のヒストグラムと全サンプル数を, それぞれ $N_m(E)$ と n_m とする. 物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値は式 (1.99) で与えられるが, その式中の状態密度 $n(E)$ の最適解は以下の二つの WHAM 方程式を自己無撞着に逐次的に解くことによって得られる [30].

$$n(E) = \frac{\sum_{m=1}^M N_m(E)}{\sum_{m=1}^M n_m \exp(f_m) W_{MUCA}^{\{m\}}(E)} = \frac{\sum_{m=1}^M N_m(E)}{\sum_{m=1}^M n_m \exp\left(f_m - \beta_m \mathcal{E}_{MUCA}^{\{m\}}(E)\right)}, \quad (1.147)$$

また, それぞれの m ($= 1, \dots, M$) について,

$$\exp(-f_m) = \sum_E n(E) W_{MUCA}^{\{m\}}(E) = \sum_E n(E) \exp\left(-\beta_m \mathcal{E}_{MUCA}^{\{m\}}(E)\right). \quad (1.148)$$

ここでは, 式 (1.118) と式 (1.119) 中のボルツマン因子 $\exp(-\beta_m E)$ の代わりに $W_{MUCA}^{\{m\}}(E)$ が使われていることに注意されたい.

更に, (ポテンシャルエネルギーの関数で表せないものを含む) 一般の物理量 A の任意の温度 T ($= 1/k_B\beta$) における期待値は, シミュレーションの途中の状態 (全粒子の座標) をある間隔で保存することによって, 以下のように計算することができる. まず, 式 (1.147) と式 (1.148) を自己無撞着に逐次的に解くことによって, f_m ($m = 1, \dots, M$) を得る. すると, 以下の式で任意の温度

58 第1章 拡張アンサンブルの方法

における期待値が求められる [26].

$$\langle A \rangle_T = \frac{\sum_{m=1}^{\mathcal{M}} \sum_{k=1}^{n_m} A(x_m(k)) \frac{1}{\sum_{\ell=1}^{\mathcal{M}} n_{\ell} \exp(f_{\ell}) W_{MUC A}^{\{\ell\}}(E(x_m(k)))} \exp[-\beta E(x_m(k))]}{\sum_{m=1}^{\mathcal{M}} \sum_{k=1}^{n_m} \frac{1}{\sum_{\ell=1}^{\mathcal{M}} n_{\ell} \exp(f_{\ell}) W_{MUC A}^{\{\ell\}}(E(x_m(k)))} \exp[-\beta E(x_m(k))]} , \quad (1.149)$$

ここで, $x_m(k)$ ($k = 1, \dots, n_m$) はそれぞれのマルチカノニカル重み因子 $W_{MUC A}^{\{m\}}(E)$ ($m = 1, \dots, \mathcal{M}$) に基づくマルチカノニカルシミュレーションによって別々に得られた k 番目の MC (または, MD) のステップにおける状態 (全粒子の座標) であり, n_m はその総数である.

同様に, MUCAREM においてマルチカノニカル法を焼き戻し法で置き換えるのが, 焼き戻しレプリカ交換法 (simulated tempering replica-exchange method: STREM) [33] である. STREM では最後の長いシミュレーションは, 各レプリカが異なった温度領域をカバーする焼き戻しシミュレーションを実行するレプリカ交換シミュレーションである.

マルチカノニカル重み因子や焼き戻し重み因子が求まれば, 従来のマルチカノニカルシミュレーションや焼き戻しシミュレーションで十分のようにも思えるが, この新手法で更にまたレプリカ交換を導入する理由は, 普通のモンテカルロや分子動力学シミュレーションは配位の局所的な変更 (local update) しかないのに対し, レプリカ交換の操作は大局的な変更 (global update) を可能にするからである. これによって, 配位空間の探索効率を大幅に上げることが期待できるわけである.

最後に, MUCAREM と STREM は次節で紹介する, 多次元レプリカ交換法 [34] の特別な例であるということをコメントしておく.

1.5 多次元拡張アンサンブル法の一般定式化

前々節では、広く使われている三つの拡張アンサンブル法（マルチカノニカル法、焼き戻し法、レプリカ交換法）を紹介した。これらの手法は、ポテンシャルエネルギーまたはそれと共役関係にある温度の酔歩を実現して、シミュレーションがエネルギー極小状態に留まるのを避ける手法であるが、本節では、これらの三手法の多次元（多変数）への拡張版を紹介する [21]。これらの手法では、温度やポテンシャルエネルギーばかりでなく、任意の系のパラメータやその共役量の酔歩を実現して、構造空間の探索効率の更なる向上を図る。

1.5.1 一般化されたポテンシャルエネルギー

L 個のパラメータ $\lambda = (\lambda^{(1)}, \dots, \lambda^{(L)})$ に依存する、状態 x における一般化されたポテンシャルエネルギー $E_{\lambda}(x)$ を考える。一般には、 $E_{\lambda}(x)$ は λ の任意の関数であっても良いが、ここでは、簡単のために、以下の特別な関数形を考える。

$$\begin{aligned} E_{\lambda}(x) &= E_0(x) + V_{\lambda}(x), \\ V_{\lambda}(x) &= \sum_{\ell=1}^L \lambda^{(\ell)} V_{\ell}(x). \end{aligned} \quad (1.150)$$

ここでは、 $(L+1)$ 個のエネルギー項がある。すなわち、 $E_0(x)$ と $V_{\ell}(x)$ ($\ell = 1, \dots, L$) である。また、 $\lambda^{(\ell)}$ は $V_{\ell}(x)$ に対する結合定数である。

運動量の自由度を積分してしまったあと、温度の値が T 、パラメータの値が λ のときの系の分配関数は以下で与えられる。

$$\begin{aligned} Z(T, \lambda) &= \int dx \exp(-\beta E_{\lambda}(x)) \\ &= \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \dots, V_L) \exp(-\beta E_{\lambda}), \end{aligned} \quad (1.151)$$

ここで、 $n(E_0, V_1, \dots, V_L)$ は多次元の状態密度であり、以下で定義される。

$$n(E_0, V_1, \dots, V_L) = \int dx \delta(E_0(x) - E_0) \delta(V_1(x) - V_1) \cdots \delta(V_L(x) - V_L). \quad (1.152)$$

60 第1章 拡張アンサンブルの方法

ここで、状態を指定する変数 x が離散的な値を取るときは、この式の積分が和で表される。

式 (1.150) のような表示は計算機シミュレーションではよく使われる。例えば、スピン系のシミュレーションでは、 $L = 1$ として、状態 x は $x = \{S_1, S_2, \dots\}$ で指定することができる。ここで、 S_k は k 番目のスピンを表す。そして、 $E_0(x)$ と $V_1(x)$ はそれぞれ、ゼロ場項と外場 $\lambda^{(1)}$ の影響下の磁化項と考えることができる。特に、イジング模型の場合は、 $E_0 = -J \sum_{\langle i,j \rangle} S_i S_j$, $V_1 = -\sum_i S_i$, $\lambda^{(1)} = h$ で与えられる (h は外部磁場である)。また、分子シミュレーションにおける傘サンプリング法 (umbrella sampling: US) [35] では、 $E_0(q)$ は元々のポテンシャルエネルギー、 $V_\ell(q)$ ($\ell = 1, \dots, L$) は結合パラメータ $\lambda^{(\ell)}$ ($\lambda = (\lambda^{(1)}, \dots, \lambda^{(L)})$) を持つ傘ポテンシャルである。ここで、状態は $x = \{q\} = \{q_1, \dots, q_N\}$ で指定されるが、 q_k は k 番目の粒子の座標ベクトルである (N は全粒子数)。更には、定圧定温アンサンブルにおける分子シミュレーションでは、 $E_0(x)$ と $V_1(x)$ (ここでは、 $L = 1$) はそれぞれ、ポテンシャルエネルギー E と体積 V に対応し、後者には圧力 P が掛っている。つまり、系の状態は $x = \{q, V\}$ で指定され、 $E_0 = E$, $V_1 = V$, $\lambda^{(1)} = P$ と書くことができる。すなわち、 $E_\lambda (= E + PV)$ は運動エネルギーからの寄与を無視した「エンタルピー」である。また、グラントカノニカルアンサンブルのシミュレーションでは、状態は $x = \{q, N\}$ で指定される (N は粒子数)。そして、 $E_0(x)$ と $V_1(x)$ (ここでも、 $L = 1$) はそれぞれ、ポテンシャルエネルギー E と全粒子数 N に対応し、後者には化学ポテンシャル μ が掛っている。つまり、 $E_0 = E$, $V_1 = N$, $\lambda^{(1)} = -\mu$ と書くことができる。

上で述べたようなよく知られた統計力学的アンサンブルから離れて、任意の物理量またはその関数を追加ポテンシャルエネルギー V_ℓ として導入することができる。例えば、 V_ℓ はスピングラスの系の参照状態とのオーバーラップとしてもよいし、分子系の末端間距離や慣性半径としてもよい。そのような場合は、結合定数 $\lambda^{(\ell)}$ の取り得る値の範囲を、新しいエネルギー項 $\lambda^{(\ell)} V_\ell$ の値が元々のポテンシャルエネルギー E_0 と大体同じオーダーの値になるように、微調整する必要がある。そうすることによって、 E_0 空間ばかりでなく V_ℓ 空間での酔歩を実現したいのである。以下に紹介する多次元の REM, ST, MUCA はこ

のような多次元の酔歩を可能にする.

1.5.2 多次元レプリカ交換法

まず, レプリカ交換法の多次元 (多変数) への拡張版の一般定式化である, 多次元レプリカ交換法 (multidimensional replica-exchange method: MREM) [34] を紹介しよう (温度以外のパラメータを交換する場合は, 特に, ハミルトニアンレプリカ交換法とも呼ばれる [36]). この手法の開発に至った重要なポイントは, M 個のレプリカが相互作用しない限り, 各レプリカのハミルトニアン $H(q, p)$ は同一である必要がないということである. 例えば, ハミルトニアンはあるパラメータに依存し, それぞれのレプリカが違うパラメータ値をとっていても良いことになる. MREM の系は相互作用しない M 個の元々の系のコピー (レプリカ) からなり, それぞれのレプリカが違うパラメータ値 Λ_m ($m = 1, \dots, M$) における「カノニカルアンサンブル」を構成する. ここで, $M = M_0 \times M_1 \times \dots \times M_L$ であり,

$$\Lambda_m \equiv (T_{m_0}, \lambda_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)}), \quad (1.153)$$

と書くことができる. また, それぞれのパラメータのラベルは以下のように書ける.

$$\begin{cases} m_0 = 1, \dots, M_0, \\ m_\ell = 1, \dots, M_\ell, \quad (\ell = 1, \dots, L). \end{cases} \quad (1.154)$$

従来のレプリカ交換法では, レプリカ i と温度 T_m が 1 対 1 対応しているのに対し, ここでは, レプリカ i と「パラメータセット」 Λ_m が 1 対 1 対応している. 従来のレプリカ交換法のパラメータ空間が 1 次元 ($\Lambda_m = T_m$) であるので, この手法はその多次元 (つまり, $(L+1)$ 次元) への拡張版ということができる.

さて, レプリカ同士は相互作用しないので, 状態 X の重み因子は各レプリカに対応するボルツマン因子の積で以下のように与えられる (式 (1.125) 参照).

$$W_{MREM} \equiv \prod_{m_0=1}^{M_0} \prod_{m_1=1}^{M_1} \dots \prod_{m_L=1}^{M_L} \exp(-\beta_{m_0} E_{\lambda_m}). \quad (1.155)$$

62 第1章 拡張アンサンブルの方法

よって、従来のレプリカ交換の採択判定式を導いたのと同じことをすることによって、パラメータ値 $\mathbf{\Lambda}_m$ をもつレプリカ i とパラメータ値 $\mathbf{\Lambda}_n$ をもつレプリカ j を交換する遷移確率は以下で与えられる。

$$w\left(x_{\mathbf{\Lambda}_m}^{[i]} \mid x_{\mathbf{\Lambda}_n}^{[j]}\right) = \min(1, \exp(-\Delta)) . \quad (1.156)$$

ここでは Δ は以下のようなになる (式 (1.133) 参照)。

$$\Delta = \beta_{m_0} \left(E_{\mathbf{\Lambda}_m} \left(q^{[j]} \right) - E_{\mathbf{\Lambda}_m} \left(q^{[i]} \right) \right) - \beta_{n_0} \left(E_{\mathbf{\Lambda}_n} \left(q^{[j]} \right) - E_{\mathbf{\Lambda}_n} \left(q^{[i]} \right) \right) . \quad (1.157)$$

一般に、 $E_{\mathbf{\Lambda}_m}$ と $E_{\mathbf{\Lambda}_n}$ は違う関数なので、新たに、交換後の $E_{\mathbf{\Lambda}_m} \left(q^{[j]} \right)$ と $E_{\mathbf{\Lambda}_n} \left(q^{[i]} \right)$ を評価する必要があることに注意されたい。

一般性を失うことなく、 $(L+1)$ 個のパラメータの大きさが次のようになっていると仮定できる。 $T_1 < T_2 < \dots < T_{M_0}$, $\lambda_1^{(\ell)} < \lambda_2^{(\ell)} < \dots < \lambda_{M_\ell}^{(\ell)}$ ($\ell = 1, \dots, L$). MREM シミュレーションは、次の二つのステップを交互に繰り返すことで実行される。

1. 各レプリカ i ($i = 1, \dots, M$) において、一定のパラメータ値 $\mathbf{\Lambda}_m$ ($m = 1, \dots, M$) をもつ従来の「カノニカル」MC または MD シミュレーションを同時にあるステップ数実行する。
2. 隣接したパラメータ値 ($\mathbf{\Lambda}_m$ と $\mathbf{\Lambda}_{m+1}$) に対応するレプリカ対 (i と j) を以下の遷移確率で交換する。

$$w\left(x_{\mathbf{\Lambda}_m}^{[i]} \mid x_{\mathbf{\Lambda}_{m+1}}^{[j]}\right) = \min(1, \exp(-\Delta)) , \quad (1.158)$$

ここで、 T 交換の場合は

$$\Delta = (\beta_{m_0} - \beta_{m_0+1}) \left(E_{\mathbf{\Lambda}_m} \left(q^{[j]} \right) - E_{\mathbf{\Lambda}_m} \left(q^{[i]} \right) \right) , \quad (1.159)$$

であり、 $\lambda^{(\ell)}$ 交換 (ℓ は $\ell = 1, \dots, L$ の中から一つの値を選んでいる) の場合は

$$\Delta = \beta_{m_0} \left[\left(E_{\mathbf{\Lambda}_{m_\ell}} \left(q^{[j]} \right) - E_{\mathbf{\Lambda}_{m_\ell}} \left(q^{[i]} \right) \right) - \left(E_{\mathbf{\Lambda}_{m_\ell+1}} \left(q^{[j]} \right) - E_{\mathbf{\Lambda}_{m_\ell+1}} \left(q^{[i]} \right) \right) \right] ,$$

$$(1.160)$$

で与えられる．また， $q^{[i]}$ と $q^{[j]}$ はそれぞれ，レプリカ交換前のレプリカ i とレプリカ j の粒子の座標である．

ここでは，レプリカ交換率を上げるために，それぞれ一つだけのパラメータの隣接値だけを交換して，それ以外のパラメータ値はレプリカ対間で同じにとっている．すなわち， T 交換の場合は，

$$\begin{cases} \mathbf{\Lambda}_m & \equiv (T_{m_0}, \mathbf{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)}), \\ \mathbf{\Lambda}_{m+1} & \equiv (T_{m_0+1}, \mathbf{\lambda}_m) \equiv (T_{m_0+1}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)}), \end{cases} \quad (1.161)$$

と温度の値だけが違っており， $\lambda^{(\ell)}$ 交換の場合は，

$$\begin{cases} \mathbf{\Lambda}_m & \equiv (T_{m_0}, \mathbf{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_\ell}^{(\ell)}, \dots, \lambda_{m_L}^{(L)}), \\ \mathbf{\Lambda}_{m+1} & \equiv (T_{m_0}, \mathbf{\lambda}_{m+1}) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_\ell+1}^{(\ell)}, \dots, \lambda_{m_L}^{(L)}), \end{cases} \quad (1.162)$$

と $\lambda^{(\ell)}$ の値だけが違っている．

特に，ポテンシャルエネルギー $E_{\mathbf{\lambda}_{m_\ell}}$ と $E_{\mathbf{\lambda}_{m_\ell+1}}$ が似た関数である場合， $\lambda^{(\ell)}$ 交換は，式 (1.162) とともに，「隣合う」ポテンシャルエネルギーについて，以下の形でも行うことができる．

$$\begin{cases} \mathbf{\Lambda}_m & \equiv (T_{m_0}, \mathbf{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_\ell}^{(\ell)}, \lambda_{m_\ell+1}^{(\ell+1)}, \dots, \lambda_{m_L}^{(L)}), \\ \mathbf{\Lambda}_{m+1} & \equiv (T_{m_0}, \mathbf{\lambda}_{m+1}) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_\ell+1}^{(\ell+1)}, \lambda_{m_\ell}^{(\ell)}, \dots, \lambda_{m_L}^{(L)}), \end{cases} \quad (1.163)$$

以下で紹介するレプリカ交換傘サンプリング法では，この $\lambda^{(\ell)}$ 交換を行う．

1.5.3 多次元焼き戻し法

次に，多次元焼き戻し法 (multidimensional simulated tempering: MST) の一般定式化 [21] を紹介する．この手法では，温度 T とパラメータ $\mathbf{\lambda}$ の酔歩が実現される．全パラメータセット $\mathbf{\Lambda} = (T, \mathbf{\lambda}) \equiv (T, \lambda^{(1)}, \dots, \lambda^{(L)})$ が動変数となって，シミュレーション中に状態とパラメータセットが以下の重み因子

64 第1章 拡張アンサンブルの方法

に従って更新されていく.

$$W_{MST}(\mathbf{\Lambda}) \equiv \exp(-\beta E_{\mathbf{\Lambda}} + f(\mathbf{\Lambda})) , \quad (1.164)$$

ここで, 関数 $f(\mathbf{\Lambda}) = f(T, \boldsymbol{\lambda})$ は $\mathbf{\Lambda}$ の確率分布が一様になるように選ばれる.

$$\begin{aligned} P_{MST}(\mathbf{\Lambda}) &= \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \cdots, V_L) \exp(-\beta E_{\mathbf{\Lambda}} + f(\mathbf{\Lambda})) \\ &\equiv \text{constant} . \end{aligned} \quad (1.165)$$

すなわち, 焼き戻し法のパラメータ $f(\mathbf{\Lambda})$ は「無次元化されたヘルムホルツ自由エネルギー」である.

$$\exp(-f(\mathbf{\Lambda})) \propto \int dE_0 dV_1 \cdots dV_L n(E_0, V_1, \cdots, V_L) \exp(-\beta E_{\mathbf{\Lambda}}) . \quad (1.166)$$

実際の数値計算においては, パラメータセット $\mathbf{\Lambda}$ を $M(= M_0 \times M_1 \times \cdots \times M_L)$ 個の値 $\mathbf{\Lambda}_m \equiv (T_{m_0}, \boldsymbol{\lambda}_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \cdots, \lambda_{m_L}^{(L)})$ に離散化する ($m_0 = 1, \cdots, M_0, m_\ell = 1, \cdots, M_\ell$ ($\ell = 1, \cdots, L$)). そして, パラメータの大小関係は以下のようになっているとす. $T_1 < T_2 < \cdots < T_{M_0}$, $\lambda_1^{(\ell)} < \lambda_2^{(\ell)} < \cdots < \lambda_{M_\ell}^{(\ell)}$ ($\ell = 1, \cdots, L$). すると, 自由エネルギー $f(\mathbf{\Lambda}_m)$ は $f_{m_0, m_1, \cdots, m_L} = f(T_{m_0}, \lambda_{m_1}^{(1)}, \cdots, \lambda_{m_L}^{(L)})$ で与えられる.

初期状態とパラメータの初期値がセットされたら, 多次元焼き戻しシミュレーションは以下の二つのステップを交互に繰り返すことによって実行される.

1. 固定されたパラメータ値 $\mathbf{\Lambda}_m = (T_{m_0}, \boldsymbol{\lambda}_m) = (T_{m_0}, \lambda_{m_1}^{(1)}, \cdots, \lambda_{m_L}^{(L)})$ を持つ重み因子 $\exp(-\beta_{m_0} E_{\mathbf{\Lambda}_m})$ に基づく「カノニカル」MC または MD シミュレーションを数ステップ実行する (固定された $\mathbf{\Lambda}_m$ では, 式 (1.164) の $f(\mathbf{\Lambda}_m)$ は定数であり, 寄与しない).
2. 状態と他のパラメータ値を固定したまま, パラメータセット $\mathbf{\Lambda}_m$ を新しい値 $\mathbf{\Lambda}_{m \pm 1}$ に更新する. このパラメータ更新の遷移確率は以下で与えられる.

$$w(x_{\mathbf{\Lambda}_m} \rightarrow x_{\mathbf{\Lambda}_{m \pm 1}}) = \min \left(1, \frac{W_{MST}(\mathbf{\Lambda}_{m \pm 1})}{W_{MST}(\mathbf{\Lambda}_m)} \right) = \min(1, \exp(-\Delta)) .$$

(1.167)

ここで、パラメータセットの更新には二つの可能性がある．すなわち、 T 更新と $\lambda^{(\ell)}$ 更新である． T 更新では、 $\Lambda_{m\pm 1} = (T_{m_0\pm 1}, \lambda_m)$ であり、以下が得られる．

$$\Delta = (\beta_{m_0\pm 1} - \beta_{m_0}) E_{\lambda_m} - (f_{m_0\pm 1, m_1, \dots, m_L} - f_{m_0, m_1, \dots, m_L}) . \quad (1.168)$$

また、 $\lambda^{(\ell)}$ 更新では ($\ell = 1, \dots, L$ の中の一つの ℓ を選んでいる)、 $\Lambda_{m\pm 1} = (T_{m_0}, \lambda_{m_\ell\pm 1})$ であり、以下が得られる．

$$\Delta = \beta_{m_0} (E_{\lambda_{m_\ell\pm 1}} - E_{\lambda_{m_\ell}}) - (f_{m_0, \dots, m_\ell\pm 1, \dots} - f_{m_0, \dots, m_\ell, \dots}) , \quad (1.169)$$

ここで、 $\lambda_{m_\ell\pm 1} = (\dots, \lambda_{m_{\ell-1}}^{(\ell-1)}, \lambda_{m_\ell\pm 1}^{(\ell)}, \lambda_{m_{\ell+1}}^{(\ell+1)}, \dots)$ 、 $\lambda_{m_\ell} = (\dots, \lambda_{m_{\ell-1}}^{(\ell-1)}, \lambda_{m_\ell}^{(\ell)}, \lambda_{m_{\ell+1}}^{(\ell+1)}, \dots)$ である．特に、ポテンシャルエネルギー $E_{\lambda_{m_\ell}}$ と $E_{\lambda_{m_\ell\pm 1}}$ が似た関数である場合、 $\lambda^{(\ell)}$ 更新は、上の更新とともに、「隣合う」ポテンシャルエネルギーへの更新をしても良い (式 (1.163) 参照)．

1.5.4 多次元マルチカノニカル法

更に、多次元マルチカノニカル法 (multidimensional multicanonical algorithm: MMUCA) の一般定式化 [21] を紹介する．この手法では、 $E_0(x)$ と $V_\ell(x)$ ($\ell = 1, \dots, L$) の $(L+1)$ 次元空間の酔歩が実現される．

多次元マルチカノニカルアンサンブルでは、各状態が重み因子 $W_{MMUCA}(E_0, V_1, \dots, V_L)$ に基づき、 E_0 と $V_1, \dots, V_L$ の一様エネルギー分布が得られる．

$$\begin{aligned} P_{MMUCA}(E_0, V_1, \dots, V_L) &\propto n(E_0, V_1, \dots, V_L) W_{MMUCA}(E_0, V_1, \dots, V_L) \\ &\equiv constant , \end{aligned} \quad (1.170)$$

ここで、 $n(E_0, V_1, \dots, V_L)$ は多次元状態密度である．この式から以下が得ら

66 第1章 拡張アンサンブルの方法

れる.

$$W_{MMUCA}(E_0, V_1, \dots, V_L) \equiv \exp(-\beta_a E_{MMUCA}(E_0, V_1, \dots, V_L)) \quad (1.171)$$

$$\propto \frac{1}{n(E_0, V_1, \dots, V_L)},$$

つまり, 任意の参照温度 $T_a = 1/k_B \beta_a$ を導入し, 重み因子をボルツマン因子様の形に書いた. ここで, 「マルチカノニカルポテンシャルエネルギー」は以下で定義される.

$$E_{MMUCA}(E_0, V_1, \dots, V_L; T_a) \equiv k_B T_a \log n(E_0, V_1, \dots, V_L). \quad (1.172)$$

多次元マルチカノニカル MC シミュレーションは以下のメトロポリスの遷移確率で実行される. すなわち, エネルギー値 $E_{\boldsymbol{\lambda}} = E_0 + \sum_{\ell=1}^L \lambda^{(\ell)} V_{\ell}$ をもつ状態 x からエネルギー値 $E_{\boldsymbol{\lambda}'} = E_0' + \sum_{\ell=1}^L \lambda^{(\ell)} V_{\ell}'$ をもつ状態 x' への遷移確率は以下で与えられる.

$$w(x \rightarrow x') = \min \left(1, \frac{W_{MMUCA}(E_0', V_1', \dots, V_L')}{W_{MMUCA}(E_0, V_1, \dots, V_L)} \right) \quad (1.173)$$

$$= \min \left(1, \frac{n(E_0, V_1, \dots, V_L)}{n(E_0', V_1', \dots, V_L')} \right).$$

多次元マルチカノニカルアンサンブル上の MD 法も式.(1.171) から自然な拡張として得られ, 従来の温度 $T = T_a$ における一定温度の MD シミュレーションが式.(1.82) において全ポテンシャルエネルギー $E_{\boldsymbol{\lambda}}$ をマルチカノニカルポテンシャルエネルギー E_{MMUCA} で置き換えることによって実行される.

$$\dot{\mathbf{p}}_k = - \frac{\partial E_{MMUCA}(E_0, V_1, \dots, V_L; T_a)}{\partial \mathbf{q}_k} - \frac{\dot{s}}{s} \mathbf{p}_k. \quad (1.174)$$

ここで, MMUCA シミュレーションの E_0 と V_{ℓ} の酔歩は, MREM や MST シミュレーションの β と $\beta \lambda^{(\ell)}$ の酔歩に以下のようにそれぞれ対応している.

$$\begin{cases} E_0 \longleftrightarrow \beta, \\ V_{\ell} \longleftrightarrow \beta \lambda^{(\ell)}, \quad (\ell = 1, \dots, L). \end{cases} \quad (1.175)$$

これらは共役関係にあるといえる.

1.5.5 多次元焼き戻し法および多次元マルチカノニカル法の重み因子の決定法

前の三つの節で紹介された三つの多次元拡張アンサンブル法の一般定式化のうち、多次元レプリカ交換法のみ、ほとんど準備なしで実行できる。それは、多次元レプリカ交換法の重み因子はいつものボルツマン因子の積で与えられるので、既に分かっているからである。一方、多次元焼き戻し法と多次元マルチカノニカル法では、重み因子は前もって分かっているわけではなく、短い試行シミュレーションで決める必要がある。ここでは、上の節で紹介した、REST および REMUCA における重み因子決定法を多次元へ一般化する [21]。

まず、 $M(= M_0 \times M_1 \times \cdots \times M_L)$ 個の違うパラメータセットの値 $\mathbf{\Lambda}_m$ ($m = 1, \dots, M$) に対応する M 個のレプリカを持つ、短い多次元レプリカ交換シミュレーションを実行する。 $N_{m_0, m_1, \dots, m_L}(E_0, V_1, \dots, V_L)$ と $n_{m_0, m_1, \dots, m_L}$ を、それぞれ、パラメータセット値 $\mathbf{\Lambda}_m = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ におけるこの試行シミュレーションで得られたポテンシャルエネルギーの $(L+1)$ 次元ヒストグラムと総サンプル数とする。一般化された多ヒストグラム再重法の WHAM 方程式の二式は以下で与えられる [21]。

$$n(E_0, V_1, \dots, V_L) = \frac{\sum_{m_0, m_1, \dots, m_L} N_{m_0, m_1, \dots, m_L}(E_0, V_1, \dots, V_L)}{\sum_{m_0, m_1, \dots, m_L} n_{m_0, m_1, \dots, m_L} \exp(f_{m_0, m_1, \dots, m_L} - \beta_{m_0} E_{\mathbf{\Lambda}_m})}, \quad (1.176)$$

$$\exp(-f_{m_0, m_1, \dots, m_L}) = \sum_{E_0, V_1, \dots, V_L} n(E_0, V_1, \dots, V_L) \exp(-\beta_{m_0} E_{\mathbf{\Lambda}_m}). \quad (1.177)$$

ここで、状態密度 $n(E_0, V_1, \dots, V_L)$ はパラメータセット $\mathbf{\Lambda}_m = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ に依存しないことに注意されたい。無次元化された自由エネルギー $f_{m_0, m_1, \dots, m_L}$ (すなわち、多次元焼き戻し法の重み因子のパラメータ) と状態密度 $n(E_0, V_1, \dots, V_L)$ (すなわち、多次元マルチカノニカル法の重み因子の逆数) は式(1.176) と式(1.177) を自己無撞着に逐次的に解くことによって得られる。

1.5.6 多次元拡張アンサンブル法における物理量の期待値の計算

ここで、多次元拡張アンサンブル法の結果から任意の温度 T と任意のパラメータ値 λ における期待値を求める方法を示す。これらは、上で述べた「一次元」の拡張アンサンブル法における方法の多次元への一般化である。

温度を含めたパラメータセットの値 Λ_m ($m = 1, \dots, M$) は $\Lambda_m \equiv (T_{m_0}, \lambda_m) \equiv (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ [$m_0 = 1, \dots, M_0, m_\ell = 1, \dots, M_\ell$ ($\ell = 1, \dots, L$), $M (= M_0 \times M_1 \times \dots \times M_L)$] のように書ける。長い MREM または MST シミュレーションのあと、物理量 A のパラメータセット値 Λ_m ($m = 1, \dots, M$) における「カノニカルアンサンブル」の期待値は以下のいつもの算術平均で与えられる。

$$\langle A \rangle_{T_{m_0}, \lambda_m} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)) , \quad (1.178)$$

ここで、 $x_m(k)$ ($k = 1, \dots, n_m$) は、パラメータセットの値 Λ_m で得られた状態であり、 n_m はその総サンプル数である ($m = 1, \dots, M$)。任意の中間のパラメータセットの値 T ($= 1/k_B\beta$) と λ における A の期待値は以下で与えられる。

$$\langle A \rangle_{T, \lambda} = \frac{\sum_{E_0, V_1, \dots, V_L} A(E_0, V_1, \dots, V_L) n(E_0, V_1, \dots, V_L) \exp(-\beta E_\lambda)}{\sum_{E_0, V_1, \dots, V_L} n(E_0, V_1, \dots, V_L) \exp(-\beta E_\lambda)} , \quad (1.179)$$

ここで、状態密度 $n(E_0, V_1, \dots, V_L)$ は多ヒストグラム再重法から得られる。すなわち、まず、長い MREM または MST シミュレーションの結果から、式 (1.176) 中のポテンシャルエネルギーのヒストグラム $N_{m_0, m_1, \dots, m_L}(E_0, V_1, \dots, V_L)$ とサンプル総数 $n_{m_0, m_1, \dots, m_L}$ を求める。状態密度 $n(E_0, V_1, \dots, V_L)$ と無次元化された自由エネルギー $f_{m_0, m_1, \dots, m_L}$ は、式 (1.176) と式 (1.177) を自己無撞着に逐次的に解くことによって得ることができる。そして、得られた状態密度 $n(E_0, V_1, \dots, V_L)$ を式 (1.179) に代入することによって、任意の T と任意の λ における物理量 A の期待値が得られるわけである。

更に、(E_0 または V_ℓ ($\ell = 1, \dots, L$) の関数として表せない場合を含む) 一般の物理量 A の期待値はシミュレーションで得られる途中の状態から求め

1.5 多次元拡張アンサンブル法の一般定式化 69

ることができる [21]. すなわち, まず, それぞれの $(m_0 = 1, \dots, M_0, m_1 = 1, \dots, M_1, \dots, m_L = 1, \dots, M_L)$ について, 式 (1.176) と式 (1.177) を自己無撞着に解いて $f_{m_0, m_1, \dots, m_L}$ を求める. すると, 任意の T と任意の λ における物理量 A の期待値は以下で与えられる.

$$\langle A \rangle_{T, \lambda} = \frac{\sum_{m_0=1}^{M_0} \cdots \sum_{m_L=1}^{M_L} \sum_{x_m} A(x_m) \frac{\exp(-\beta E_{\lambda}(x_m))}{\sum_{n_0=1}^{M_0} \cdots \sum_{n_L=1}^{M_L} n_{n_0, \dots, n_L} \exp(f_{n_0, \dots, n_L} - \beta_{n_0} E_{\lambda_{n_0}}(x_m))}}{\sum_{m_0=1}^{M_0} \cdots \sum_{m_L=1}^{M_L} \sum_{x_m} \frac{\exp(-\beta E_{\lambda}(x_m))}{\sum_{n_0=1}^{M_0} \cdots \sum_{n_L=1}^{M_L} n_{n_0, \dots, n_L} \exp(f_{n_0, \dots, n_L} - \beta_{n_0} E_{\lambda_{n_0}}(x_m))}}, \quad (1.180)$$

ここで, x_m はパラメータセットの値が $\Lambda_m = (T_{m_0}, \lambda_m) = (T_{m_0}, \lambda_{m_1}^{(1)}, \dots, \lambda_{m_L}^{(L)})$ のときに得られた状態である. また, x_m はそれぞれの Λ_m について別々に保存されることに注意されたい.

多次元マルチカノニカル法については以下になる. 重み因子 $W_{MMUCA}(E_0, \dots, V_L)$ に基づく長い MMUCA シミュレーションの結果から, 任意の T と任意の λ における物理量 A の期待値は式 (1.179) で与えられるが, 式中の状態密度は, 単ヒストグラム再重法によって以下のように求められる. $N_{MMUCA}(E_0, V_1, \dots, V_L)$ をポテンシャルエネルギーの分布 $P_{MMUCA}(E_0, V_1, \dots, V_L)$ の MMUCA シミュレーションから得られたヒストグラムとすると, 状態密度 $n(E_0, V_1, \dots, V_L)$ の最適解は以下で与えられる.

$$n(E_0, V_1, \dots, V_L) = \frac{N_{MMUCA}(E_0, V_1, \dots, V_L)}{W_{MMUCA}(E_0, \dots, V_L)}. \quad (1.181)$$

更に, $(E_0$ または V_ℓ ($\ell = 1, \dots, L$) の関数として表せない場合を含む) 一般の物理量 A の期待値はシミュレーションで得られる途中の状態から以下の式

70 第1章 拡張アンサンブルの方法

で求められる.

$$\langle A \rangle_{T,\lambda} = \frac{\sum_{k=1}^{n_s} A(x_k) W_{MMUCA}^{-1}(E_0(x_k), \dots, V_L(x_k)) \exp(-\beta E_{\lambda}(x_k))}{\sum_{k=1}^{n_s} W_{MMUCA}^{-1}(E_0(x_k), \dots, V_L(x_k)) \exp(-\beta E_{\lambda}(x_k))} . \quad (1.182)$$

ここで, x_k は k 番目の MC (または MD) ステップにおける状態であり, n_s は保存された状態の総数である.

1.6 多次元拡張アンサンブル法の具体例

前節では, 広く使われている三つの拡張アンサンブル法 (マルチカノニカル法, 焼き戻し法, レプリカ交換法) の多次元 (多変数) への拡張版の一般的な定式化を与えた. ここでは, この一般定式化の具体的な例を幾つか紹介する.

1.6.1 レプリカ交換自由エネルギー摂動法とレプリカ交換傘サンブル法

多次元レプリカ交換法は, 例えば, 自由エネルギー計算に使うことができる. まず, 自由エネルギー摂動法的扱いに基づく, レプリカ交換自由エネルギー摂動法 (replica-exchange free energy perturbation: REFEP) [34] を述べる. ここでは, ポテンシャルエネルギーを以下で与える.

$$E_{\lambda}(q) = E_I(q) + \lambda(E_F(q) - E_I(q)) . \quad (1.183)$$

ここで, E_I と E_F はそれぞれ「野生型」分子と「突然変異型」分子のポテンシャルエネルギーである. これが式 (1.150) と同じ形をしていることに注意されたい ($L=1$, $V_1 = E_F - E_I$, $\lambda^{(1)} = \lambda$ に対応している).

多次元レプリカ交換シミュレーションは M 個の異なるパラメータ $\Lambda_m = (T_m, \lambda_m)$ を持つ M 個のレプリカについて実行される. ここで, $E_{\lambda=0}(q) = E_I(q)$, $E_{\lambda=1}(q) = E_F(q)$ であるから, 十分なレプリカ交換が実現されるためには, 0 と 1 の間に十分な数の λ_m を用意する必要がある. この多次元レ

1.6 多次元拡張アンサンブル法の具体例 71

プリカ交換シミュレーションの結果, エネルギー分布として, M 個のヒストグラム $N_m(E_I, E_F - E_I)$ (または, 同等な量として, $N_m(E_I, E_F)$) が得られる. 温度 T における, 突然変異に関するヘルムホルツ自由エネルギーの差 $\Delta F \equiv F_{\lambda=1} - F_{\lambda=0}$ は以下の式から求めることができる.

$$\exp(-\beta\Delta F) = \frac{Z_{T,\lambda=1}}{Z_{T,\lambda=0}} = \frac{\sum_{E_I, E_F} P_{T,\lambda=1}(E_I, E_F)}{\sum_{E_I, E_F} P_{T,\lambda=0}(E_I, E_F)}. \quad (1.184)$$

ここで, 確率分布 $P_{T,\lambda}(E_I, E_F) = n(E_I, E_F) e^{-\beta E_\lambda}$ は式 (1.176) と式 (1.177) の多ヒストグラム再重法から得られる.

もう一つ, 今度は傘サンプル法に基づく自由エネルギーの計算法を紹介しよう. よって, レプリカ交換傘サンプル法 (replica-exchange umbrella sampling: REUS) [34] と呼ぶことができる. ポテンシャルエネルギーは式 (1.150) で与えられ, 以下のように書ける.

$$E_{\boldsymbol{\lambda}}(q) = E_0(q) + V_{\boldsymbol{\lambda}}(q) = E_0(q) + \sum_{\ell=1}^L \lambda^{(\ell)} V_{\ell}(q). \quad (1.185)$$

ここで, $E_0(q)$ は元々のポテンシャルエネルギーであり, $V_{\ell}(q)$ ($\ell = 1, \dots, L$) は結合パラメータ $\lambda^{(\ell)}$ ($\boldsymbol{\lambda} = (\lambda^{(1)}, \dots, \lambda^{(L)})$) を持つ傘ポテンシャルである. 反応座標 (reaction coordinate) ξ を導入することによって, 傘ポテンシャルは普通次のような調和振動子のポテンシャルとして表すことが多い.

$$V_{\ell}(q) = k_{\ell} (\xi(q) - d_{\ell})^2, \quad (\ell = 1, \dots, L). \quad (1.186)$$

ここで, d_{ℓ} は拘束ポテンシャルの中心であり, k_{ℓ} は拘束ポテンシャルの強度 (バネ定数) を表す. そして, M 個の異なるパラメータセット $\boldsymbol{\Lambda}_m = (T_m, \boldsymbol{\lambda}_m)$ に対応する M 個のレプリカを用意して, 多次元レプリカ交換シミュレーションを実行する.

レプリカ交換傘サンプル法の実現例として, 以下の2例についてもう少し詳しく述べることにする. まず, 1 個の温度 T と L 個の傘ポテンシャルを用意する. そして, それぞれのレプリカには 1 個の傘ポテンシャルを対応させる (よっ

72 第1章 拡張アンサンブルの方法

て、レプリカの個数は $M = L$ 。このとき、式 (1.185) において $\lambda = \lambda_m$ とすると以下が得られる。

$$\lambda_m^{(\ell)} = \delta_{\ell,m} . \quad (1.187)$$

ここで、 $\delta_{k,l}$ はクロネッカーのデルタ関数である。すると、レプリカ i と j の (式 (1.185) の) ポテンシャルエネルギーはレプリカ交換の前では次のように書くことができる。

$$\begin{cases} E_{\lambda_m}(q^{[i]}) = E_0(q^{[i]}) + V_m(q^{[i]}) , \\ E_{\lambda_m}(q^{[j]}) = E_0(q^{[j]}) + V_{m+1}(q^{[j]}) . \end{cases} \quad (1.188)$$

また、レプリカ交換の操作は以下のようにポテンシャルエネルギーを置き換えることである (式 (1.163) 参照)。

$$\begin{cases} E_{\lambda_m}(q^{[j]}) = E_0(q^{[j]}) + V_m(q^{[j]}) , \\ E_{\lambda_m}(q^{[i]}) = E_0(q^{[i]}) + V_{m+1}(q^{[i]}) . \end{cases} \quad (1.189)$$

よって、このレプリカ交換傘サンプルシミュレーションは、(温度の交換はせず) 「隣接する」 傘ポテンシャル V_m と V_{m+1} に対応するレプリカ対を交換する操作を繰り返すことで実現される。このレプリカ交換の操作の採択条件は式 (1.158) で与えられるが、ここでは式 (1.160) の Δ は以下ようになる ($\beta = 1/k_B T$ は固定されている)。

$$\Delta = \beta \left[\left(V_m(q^{[j]}) - V_m(q^{[i]}) \right) - \left(V_{m+1}(q^{[j]}) - V_{m+1}(q^{[i]}) \right) \right] . \quad (1.190)$$

ここで、レプリカ i と j はそれぞれ交換の前に傘ポテンシャル V_m と V_{m+1} に対応していたことになる。

次に、 N_T 個の温度と L 個の傘ポテンシャルを用意し、レプリカの個数を $M = N_T \times L$ とする。すなわち、パラメータセットが一辺の長さがそれぞれ N_T と L の二次元格子状に値をもつとする。よって、レプリカを特徴づけるパラメータセットに以下のようなラベルを導入することができる。

$$\begin{aligned} \Lambda_m = (T_m, \lambda_m) &\longrightarrow \Lambda_{I,J} = (T_I, \lambda_J) . \\ (m = 1, \dots, M) &\quad (I = 1, \dots, N_T, J = 1, \dots, L) \end{aligned} \quad (1.191)$$

1.6 多次元拡張アンサンブル法の具体例 73

このとき、ポテンシャルエネルギーは式 (1.188) の中で $m \rightarrow J$ と置き換えた形で与えられる。すると、レプリカ交換傘サンブルシミュレーションは以下の二つのレプリカ交換の操作を交互に繰り返すことによって実現される。

1. 隣接する温度 T_I と T_{I+1} に対応するレプリカ対を交換する（すなわち、それぞれパラメータセット $\mathbf{\Lambda}_{I,J}$ と $\mathbf{\Lambda}_{I+1,J}$ に対応するレプリカ i と j を交換する）。(この操作を「 T 交換」と呼ぶことにする。)
2. 「隣接する」傘ポテンシャル V_J と V_{J+1} に対応するレプリカ対を交換する（すなわち、それぞれパラメータセット $\mathbf{\Lambda}_{I,J}$ と $\mathbf{\Lambda}_{I,J+1}$ に対応するレプリカ i と j を交換する）。(この操作を「 λ 交換」と呼ぶことにする。)

これらのレプリカ交換の操作の採択条件は式 (1.131) で与えられるが、ここでは式 (1.146) の Δ は以下ようになる。 T 交換については、

$$\Delta = (\beta_I - \beta_{I+1}) \left[\left(E_0(q^{[j]}) + V_J(q^{[j]}) \right) - \left(E_0(q^{[i]}) + V_J(q^{[i]}) \right) \right], \quad (1.192)$$

であり、 λ 交換については、

$$\Delta = \beta_I \left[\left(V_J(q^{[j]}) - V_J(q^{[i]}) \right) - \left(V_{J+1}(q^{[j]}) - V_{J+1}(q^{[i]}) \right) \right], \quad (1.193)$$

である。これらの操作を繰り返すことによって、温度空間および反応座標空間上の酔歩が得られて、エネルギー極小状態に留まるのを避けられるわけである。

さて、式 (1.186) の傘ポテンシャル $V_\ell(q)$ は全て反応座標 ξ のみの関数であるので、エネルギー分布のヒストグラムは $N_{I,J}(E_0, V_1, \dots, V_L)$ の代わりに $N_{I,J}(E_0, \xi)$ をとればよいことになる。ここで、 $N_{I,J}(E_0, \xi)$ は温度が T_I でポテンシャルエネルギーが $E_0 + V_J$ のときに得られた E_0 と ξ の分布のヒストグラムである ($I = 1, \dots, N_T$, $J = 1, \dots, L$)。すると、式 (1.176) と式 (1.177) の WHAM 方程式は以下のように書くことができる。

$$n(E_0, \xi) = \frac{\sum_{I=1}^{N_T} \sum_{J=1}^L N_{I,J}(E_0, \xi)}{\sum_{I=1}^{N_T} \sum_{J=1}^L n_{I,J} e^{f_{I,J} - \beta_I(E_0 + V_J)}}, \quad (1.194)$$

74 第1章 拡張アンサンブルの方法

$$e^{-f_{I,J}} = \sum_{E_0, \xi} n(E_0, \xi) e^{-\beta_I(E_0 + V_J)} . \quad (1.195)$$

よって、物理量 A の傘ポテンシャルによるバイアスを受けていない、元々の系の任意の温度 T ($= 1/k_B\beta$) におけるカノニカルアンサンブル平均は次式で与えられる.

$$\langle A \rangle_T = \frac{\sum_{E_0, \xi} A(E_0, \xi) P_T(E_0, \xi)}{\sum_{E_0, \xi} P_T(E_0, \xi)} . \quad (1.196)$$

ここで、確率分布 $P_T(E_0, \xi) = n(E_0, \xi) e^{-\beta E_0}$ は式 (1.194) と式 (1.195) の多ヒストグラム再重法から得られる.

傘ポテンシャルによるバイアスを受けていない、元々の系の任意の温度 T における平均力ポテンシャル (potential of mean force: PMF) ^{†2}は以下の式で与えられる.

$$\mathcal{W}_T(\xi) = -k_B T \log \left[\sum_{E_0} P_T(E_0, \xi) \right] . \quad (1.197)$$

特に、温度が一つの値 T_0 ($= 1/k_B\beta_0$) に固定されている場合を考えよう ($N_T = 1$). このとき、式 (1.194) と式 (1.195) は以下のように書ける.

$$n(E_0, \xi) = \frac{\sum_{J=1}^L N_J(E_0, \xi)}{\sum_{J=1}^L n_J e^{f_J - \beta_0(E_0 + V_J)}} , \quad (1.198)$$

$$e^{-f_J} = \sum_{E_0, \xi} n(E_0, \xi) e^{-\beta_0(E_0 + V_J)} . \quad (1.199)$$

ここで、 $N_J(E_0, \xi)$ は温度が T_0 でポテンシャルエネルギーが $E_0 + V_J$ のときに得られた E_0 と ξ の分布のヒストグラムである ($J = 1, \dots, L$).

^{†2}平均力ポテンシャルとは、ここでは反応座標 ξ の関数としてのヘルムホルツ自由エネルギーのことである.

1.6 多次元拡張アンサンブル法の具体例 75

また、傘ポテンシャルは反応座標 ξ の関数なので、以下のように書ける（式 (1.186) 参照）.

$$V_J(q) = V_J(\xi(q)) , \quad (J = 1, \dots, L) . \quad (1.200)$$

式 (1.198) の両辺に $\exp(-\beta_0 E_0)$ を掛けて、 E_0 について和を取ると以下が得られる.

$$\sum_{E_0} n(E_0, \xi) e^{-\beta_0 E_0} = \frac{\sum_{E_0} \sum_{J=1}^L N_J(E_0, \xi) e^{-\beta_0 E_0}}{\sum_{J=1}^L n_J e^{f_J - \beta_0 (E_0 + V_J)}} = \frac{\sum_{J=1}^L \left[\sum_{E_0} N_J(E_0, \xi) \right]}{\sum_{J=1}^L n_J e^{f_J - \beta_0 V_J(\xi)}} . \quad (1.201)$$

また、式 (1.195) は以下のように書き換えることができる.

$$e^{-f_J} = \sum_{\xi} \left[\sum_{E_0} n(E_0, \xi) e^{-\beta_0 E_0} \right] e^{-\beta_0 V_J(\xi)} . \quad (1.202)$$

よって、これらの WHAM 方程式は結局以下のように書き換えることができる [37, 38].

$$\rho_0(\xi) = \frac{\sum_{J=1}^L N_J(\xi)}{\sum_{J=1}^L n_J e^{f_J - \beta_0 V_J(\xi)}} , \quad (1.203)$$

$$e^{-f_J} = \sum_{\xi} \rho_0(\xi) e^{-\beta_0 V_J(\xi)} . \quad (1.204)$$

ここで、 $\rho_0(\xi)$ は傘ポテンシャルによるバイアスを受けていない、元々の系の温度 T_0 における反応座標 ξ の確率分布であり、以下で定義される.

$$\rho_0(\xi) \equiv \sum_{E_0} n(E_0, \xi) e^{-\beta_0 E_0} . \quad (1.205)$$

76 第1章 拡張アンサンブルの方法

また, $N_J(\xi)$ は (温度 T_0 で) ポテンシャルエネルギーが $E_0(q) + V_J(\xi)$ のときに得られた ξ の分布のヒストグラムで, 以下の関係が成り立つ.

$$N_J(\xi) = \sum_{E_0} N_J(E_0, \xi) . \quad (1.206)$$

更に, n_J は (温度 T_0 で) ポテンシャルエネルギーが $E_0(q) + V_J(\xi)$ のときに得られたサンプル (状態) の総数である.

よって, 傘ポテンシャルによるバイアスを受けていない, 元々の系の温度 T_0 における平均力ポテンシャルは以下の式で与えられる.

$$\mathcal{W}_{T_0}(\xi) = -k_B T_0 \log \rho_0(\xi) . \quad (1.207)$$

すなわち, 温度が固定されている場合は, 二次元のヒストグラム $N_J(E_0, \xi)$ の代わりに一次元のヒストグラム $N_J(\xi)$ を考慮するだけで良いことになる.

1.6.2 ファンデルワールスレプリカ交換法

多次元レプリカ交換法のもう一つの例として, ファンデルワールスレプリカ交換法 (van der Waals Replica-Exchange Method: vWREM) [39] を紹介する.

水分子をあらわに取り入れた溶質分子からなる系を考える. このとき, 系の全ポテンシャルエネルギーは以下のように書ける.

$$E_\lambda(q) = E_p(q_p) + E_{ps}(q_p, q_s) + E_s(q_s) , \quad (1.208)$$

ここで, E_p は溶質の原子だけについてのポテンシャルエネルギー項, E_{ps} は溶質の原子と溶媒の原子の間の相互作用を表すポテンシャルエネルギー項, E_s は溶媒の原子だけについてのポテンシャルエネルギー項である. また, 全原子の座標は $q = \{q_p, q_s\}$ であり, その中身は, q_p が溶質原子の座標, q_s が溶媒原子の座標である. これらは, それぞれ次のように書ける. $q_p \equiv \{q_1, \dots, q_{N_p}\}$, $q_s \equiv \{q_{N_p+1}, \dots, q_N\}$ (N は全原子数, N_p は溶質分子の全原子数である).

われわれは, 溶媒よりも溶質分子の構造空間の効率的なサンプリングに興味がある. 溶質分子構造における立体障害は, 溶質分子の各原子のファンデルワールス半径によって規定されている. すなわち, ファンデルワールス半径が大きい

1.6 多次元拡張アンサンブル法の具体例 77

いと、溶質分子はあちこちぶつかるので、より多くの立体障害を受ける。一方、ファンデルワールス半径が小さいと、溶質分子は自由に動き回れる。よって、溶質原子のファンデルワールス半径をスケールするパラメータ λ を以下のように導入する。

$$\sigma_{k\ell} \longrightarrow \lambda \sigma_{k\ell} . \quad (1.209)$$

そして、式 (1.208) の E_p 内のうち、溶質原子間のレナード・ジョーンズ相互作用エネルギー項を以下のように書く。

$$V_\lambda(q_p) = \sum_{k=1}^{N_p-1} \sum_{\ell=k+1}^{N_p} 4\epsilon_{k\ell} \left[\left(\frac{\lambda \sigma_{k\ell}}{r_{k\ell}} \right)^{12} - \left(\frac{\lambda \sigma_{k\ell}}{r_{k\ell}} \right)^6 \right] , \quad (1.210)$$

ここで、 $r_{k\ell}$ は溶質分子の原子対 k と ℓ の間の距離である。また、 $\epsilon_{k\ell}$ と $\sigma_{k\ell}$ は対応するレナード・ジョーンズパラメータである。元々のポテンシャルエネルギーは $\lambda = 1$ のときに再現される。また、 $\lambda < 1$ のとき、立体障害が減少する。式 (1.208) の E_λ の中で、この V_λ 項だけが λ に依存していることに注意されたい。更に、式 (1.210) の λ 依存性は、式 (1.150) のように比例定数の形になっていないことに注目されたい。

そして、 M 個の λ の値 λ_m ($m = 1, \dots, M$) を用意する ($\lambda_1 < \dots < \lambda_M$ とする)。また、温度については、 $T_0 = 1/k_B\beta_0$ に固定する場合をまず考える。

ファンデルワールスレプリカ交換法は以下の二つのステップを交互に繰り返すことによって実行される。

1. 違う λ_m の値をもつそれぞれのレプリカにおいて式 (1.3) のボルツマン因子に基づくカノニカル MC または MD シミュレーションを数ステップ同時にまた独立に実行する。
2. 隣合うパラメータ値 λ_m と λ_{m+1} に対応するレプリカ i とレプリカ j を交換する。このレプリカ交換の遷移確率は式 (1.158) で与えられるが、その中の Δ (式 (1.160) 参照) は以下で与えられる。

$$\Delta = \beta_0 \left[\left(V_{\lambda_m} \left(q_p^{[j]} \right) - V_{\lambda_m} \left(q_p^{[i]} \right) \right) - \left(V_{\lambda_{m+1}} \left(q_p^{[j]} \right) - V_{\lambda_{m+1}} \left(q_p^{[i]} \right) \right) \right] . \quad (1.211)$$

78 第1章 拡張アンサンブルの方法

ここで、 V_λ は式 (1.210) の中の溶質原子間のみのレナード・ジョーンズポテンシャルエネルギー項である。

E_λ の λ 依存性は V_λ のみに存在するので、残りの項は式 (1.160) で打ち消し合ったことに注意されたい。

式 (1.211) は溶質分子の原子の座標 q_p のみを含み、溶媒分子の座標 q_s に依存しない。普通、 $N_p \ll N$ が成り立つので、系の自由度の数が増えるに従って、必要とされるレプリカ数が増加するという、従来のレプリカ交換法の困難は大きく緩められるわけである。

式 (1.159) と式 (1.160) に基づいて、温度 T とパラメータ λ についての二次元レプリカ交換シミュレーションを実行することによって、更に構造サンプリングの効率を上げることができることをコメントしておく。

このファンデルワールスレプリカ交換シミュレーションの結果は、多ヒストグラム再重法によって解析することができる。まず、一定温度 T_0 において、 M 個のパラメータ値 λ_m ($m = 1, \dots, M$) に対応する、 M 個のレプリカの vWREM シミュレーションを実行したとする。

適宜に選んだ反応座標 (例えば、 ξ_1, ξ_2 とする) について、任意のパラメータ値 λ と任意の温度 T における、「カノニカル分布」 $P_{T,\lambda}(\xi_1, \xi_2)$ についての WHAM 方程式は以下で与えられる。

$$P_{T,\lambda}(\xi_1, \xi_2) = \sum_{E_{\lambda_1}, \dots, E_{\lambda_M}} \frac{\sum_{m=1}^M N_m(E_{\lambda_1}, \dots, E_{\lambda_M}; \xi_1, \xi_2) e^{-\beta E_\lambda}}{\sum_{m=1}^M n_m e^{f_{T_0, \lambda_m} - \beta_0 E_{\lambda_m}}} \quad (1.212)$$

$$e^{-f_{T_0, \lambda_m}} = \sum_{\xi_1, \xi_2} P_{T_0, \lambda_m}(\xi_1, \xi_2) . \quad (1.213)$$

ここで、 $N_m(E_{\lambda_1}, \dots, E_{\lambda_M}; \xi_1, \xi_2)$ はパラメータ値 λ_m と反応座標値 (ξ_1, ξ_2) をもつ M 次元エネルギー分布のヒストグラムである。また、 n_m はパラメータ値 λ_m における全サンプル数である。ここで、確率分布は規格化されていないことに注意されたい。式 (1.212) と式 (1.213) が自己無撞着に逐次的に解かれ

1.6 多次元拡張アンサンブル法の実例 79

る。また、これらの方程式は任意の反応座標のセット $(\xi_1, \xi_2, \dots)$ に簡単に一般化できることをコメントしておく。

このように得られた式 (1.212) の確率分布から、任意のパラメータ値 λ および任意の温度 T における物理量 A の期待値は以下で得られる。

$$\langle A \rangle_{T,\lambda} = \frac{\sum_{\xi_1, \xi_2} A(\xi_1, \xi_2) P_{T,\lambda}(\xi_1, \xi_2)}{\sum_{\xi_1, \xi_2} P_{T,\lambda}(\xi_1, \xi_2)} . \quad (1.214)$$

また、任意のパラメータ値 λ と任意の温度 T における、反応座標 ξ_1 と ξ_2 についての平均力ポテンシャルは以下で得られる。

$$F_{T,\lambda}(\xi_1, \xi_2) = -k_B T \log P_{T,\lambda}(\xi_1, \xi_2) . \quad (1.215)$$

式 (1.212), (1.214), (1.215) で、「任意の温度」と書いたが、計算結果が信頼できる温度領域は T_0 の近傍に限られる。幅広い温度値 T について精度が良い結果を得るためには、式 (1.159) の T 交換の操作も必要になる。

1.6.3 マルチオーバーラップ法とマルチカノニカル・マルチオーバーラップ法

マルチカノニカル法はポテンシャルエネルギー空間の酔歩を実現するが、原理的には任意の別の変数の酔歩を実現することも可能である。そのような例の一つがマルチオーバーラップ法 (multioverlap algorithm: MUOV) [40, 41] である。

ここでは、蛋白質のような高分子系を考える。そして、二面角空間の「オーバーラップ」を以下で定義する [42]。

$$O = 1 - d , \quad (1.216)$$

ここで、 d は二面角距離であり、以下で定義される (n 個の二面角を対象としている)。

$$d = \frac{1}{n\pi} \sum_{i=1}^n d_a(\theta_i, \theta_i^0) . \quad (1.217)$$

80 第1章 拡張アンサンブルの方法

ここで、 θ_i は i 番目の二面角であり、 θ_i^0 は参照構造の対応する二面角である。二面角対の「距離」は以下で与えられる。

$$d_a(\theta_i, \theta_i^0) = \min(|\theta_i - \theta_i^0|, 2\pi - |\theta_i - \theta_i^0|) . \quad (1.218)$$

ここで、角度はラジアン単位で与えている。また、式 (1.217) の二面角距離 d の値域は $0 \leq d \leq 1$ である。もし、 $d = 0$ なら、問題としている立体構造と参照構造の全ての二面角が一致していることになる。よって、二面角距離は対象構造がどれくらい参照構造に似ているかを表す指標である。式 (1.216) から分かるように、二面角距離 d はオーバーラップ O と同等である。よって、以下ではオーバーラップの代わりに二面角距離で議論を進める。

一定温度 T_0 におけるマルチオーバーラップアンサンブルでは、確率分布は以下の非ボルツマン因子（「マルチオーバーラップ重み因子」と呼ぶことにする）で与えられる。

$$W_{MUOV}(d, E; T_0) = e^{-\beta_0 E_{MUOV}} , \quad (1.219)$$

ここで、「マルチオーバーラップポテンシャルエネルギー」 E_{MUOV} は以下で定義される。

$$E_{MUOV}(d, E; T_0) = E - k_B T_0 f(d; T_0) . \quad (1.220)$$

また、 $f(d; T_0)$ は二面角距離 d における無次元化された自由エネルギーである。

多次元二面角空間への一般化も容易であり、そのマルチオーバーラップ重み因子は以下で与えられる。

$$\begin{aligned} W_{MUOV}(d_1, \dots, d_L, E; T_0) &= e^{-\beta_0 E_{MUOV}} \\ &\equiv e^{-\beta_0 E + f(d_1, \dots, d_L; T_0)} , \end{aligned} \quad (1.221)$$

ここで、 L は参照構造の総数であり、 d_i は i ($= 1, \dots, L$) 番目の参照構造との二面角距離である。関数 $f(d_1, \dots, d_L; T_0)$ は二面角距離 $d_1, \dots, d_L$ における無次元化された自由エネルギーであり、二面角の確率分布 $P_{MUOV}(d_1, \dots, d_L; T_0)$

が一様になるように以下で定義される.

$$\begin{aligned}
 P_{MUOV}(d_1, \dots, d_L; T_0) &= \int dE P_{MUOV}(d_1, \dots, d_L, E; T_0) \\
 &= \int dE n(d_1, \dots, d_L, E) W_{MUOV}(d_1, \dots, d_L, E; T_0) \\
 &= \int dE n(d_1, \dots, d_L, E) e^{-\beta_0 E + f(d_1, \dots, d_L; T_0)} \\
 &\equiv \text{constant} ,
 \end{aligned} \tag{1.222}$$

ここで, $P_{MUOV}(d_1, \dots, d_L, E; T_0)$ はポテンシャルエネルギー E と二面角距離 $(d_1, \dots, d_L)$ の確率分布であり, $n(d_1, \dots, d_L, E)$ はその状態密度である.

マルチオーバーラップ MC シミュレーションは以下のメトロポリス判定に基づいて実行される. ポテンシャルエネルギーが E で二面角距離が $(d_1, \dots, d_L)$ の状態 x からポテンシャルエネルギーが E' で二面角距離が $(d_1', \dots, d_L')$ の状態 x' への遷移確率は, 式(1.57)において, ボルツマン重み因子 $W_B(E)$ をマルチオーバーラップ重み因子 $W_{MUOV}(d_1, \dots, d_L; E)$ で置き換えることによって得られる [40].

$$w(x \rightarrow x') = \min \left(1, \frac{W_{MUOV}(d_1', \dots, d_L'; E')}{W_{MUOV}(d_1, \dots, d_L; E)} \right) . \tag{1.223}$$

マルチオーバーラップアンサンブルにおける MD 法も同様に式(1.221)から導かれる. すなわち, いつもの温度一定 ($T = T_0$) の MD シミュレーションが式 (1.82) において E を E_{MUOV} で置き換えて実行される [41].

$$\begin{aligned}
 \dot{\mathbf{p}}_k &= - \frac{\partial E_{MUOV}}{\partial \mathbf{q}_k}(d_1, \dots, d_L, E; T_0) - \frac{\dot{s}}{s} \mathbf{p}_k \\
 &= \mathbf{F}_k + k_B T_0 \frac{\partial f}{\partial \mathbf{q}_k}(d_1, \dots, d_L; T_0) - \frac{\dot{s}}{s} \mathbf{p}_k .
 \end{aligned} \tag{1.224}$$

マルチオーバーラップ重み因子 (または, 無次元化された自由エネルギー) は前もって知られていないので, 短いシミュレーションのイテレーションによって決定する必要がある. 適当な重み因子が温度 T_0 において求まったあと, この温度において長いマルチオーバーラップシミュレーションを実行する. このシミュレーションの結果は単ヒストグラム再重法で解析できる. すなわち, 任

82 第1章 拡張アンサンブルの方法

意の温度 T における物理量 A の期待値は以下で与えられる.

$$\begin{aligned}\langle A \rangle_T &= \frac{\sum_{d_1, \dots, d_L, E} A(d_1, \dots, d_L, E) N_{MUOV}(d_1, \dots, d_L, E) W_{MUOV}(d_1, \dots, d_L, E; T_0)^{-1} e^{-\beta E}}{\sum_{d_1, \dots, d_L, E} N_{MUOV}(d_1, \dots, d_L, E) W_{MUOV}(d_1, \dots, d_L, E; T_0)^{-1} e^{-\beta E}} \\ &= \frac{\sum_{d_1, \dots, d_L, E} A(d_1, \dots, d_L, E) N_{MUOV}(d_1, \dots, d_L, E) e^{-(\beta - \beta_0)E - f(d_1, \dots, d_L; T_0)}}{\sum_{d_1, \dots, d_L, E} N_{MUOV}(d_1, \dots, d_L, E) e^{-(\beta - \beta_0)E - f(d_1, \dots, d_L; T_0)}}, \quad (1.225)\end{aligned}$$

ここで, $N_{MUOV}(d_1, \dots, d_L, E)$ はポテンシャルエネルギーと二面角距離の確率分布 $P_{MUOV}(d_1, \dots, d_L, E; T_0)$ のヒストグラムである.

マルチオーバーラップ法はマルチカノニカル法と合わせて二次元の拡張アンサンブル法に一般化できる. よって, この方法はマルチカノニカル・マルチオーバーラップ法 (multicanonical-multioverlap algorithm: MUCA-MUOV) と呼ばれる [43]. この方法では, ポテンシャルエネルギーと二面角距離の一樣分布が得られるように, 非ボルツマン因子 $W_{MCMO}(d_1, \dots, d_L, E)$ (マルチカノニカル・マルチオーバーラップ重み因子と呼ばれる) が以下のように定義される.

$$P_{MCMO}(d_1, \dots, d_L, E) = n(d_1, \dots, d_L, E) W_{MCMO}(d_1, \dots, d_L, E) \equiv \text{constant}. \quad (1.226)$$

よって, この方法では, 二面角距離空間ばかりでなくポテンシャルエネルギー空間の酔歩が実現される.

1.6.4 定圧定温アンサンブルにおける拡張アンサンブル法

前節の多次元拡張アンサンブル法の一般定式化の例として, 圧力と温度が一定の定圧定温アンサンブルにおけるシミュレーション法の拡張アンサンブル法への一般化について議論する.

有限体積 $\mathcal{V}$ の箱にはいった N 個の粒子系を考える. この系の状態 x は粒子の座標 $q \equiv \{q_1, q_2, \dots, q_N\}$, 運動量 $p \equiv \{p_1, p_2, \dots, p_N\}$, 箱の体積 $\mathcal{V}$ で指定される (つまり, $x = (q, p, \mathcal{V})$). ポテンシャルエネルギー $E(q, \mathcal{V})$ は q と $\mathcal{V}$ の関数である.

1.6 多次元拡張アンサンブル法の具体例 83

定圧定温アンサンブルでは、温度 T と圧力 $\mathcal{P}$ におけるポテンシャルエネルギー E と体積 $\mathcal{V}$ の確率分布 $P_{NPT}(E, \mathcal{V}; T, \mathcal{P})$ は以下で与えられる [4, 44, 45].

$$P_{NPT}(E, \mathcal{V}; T, \mathcal{P}) = n(E, \mathcal{V}) W_{NPT}(E, \mathcal{V}; T, \mathcal{P}) = n(E, \mathcal{V}) e^{-\beta \mathcal{H}}. \quad (1.227)$$

ここで、状態密度 $n(E, \mathcal{V})$ はポテンシャルエネルギー E と体積 $\mathcal{V}$ の関数である。また、 $\mathcal{H}$ は（運動エネルギー項以外の寄与による）「エンタルピー」であり、以下で定義される。

$$\mathcal{H} = E + \mathcal{P}\mathcal{V}. \quad (1.228)$$

これは、一般定式化の式 (1.150) において、 $L = 1$, $E_0 = E$, $V_1 = \mathcal{V}$, $\lambda^{(1)} = \mathcal{P}$ の場合に対応していることに注意されたい。

今、粒子が一辺 $\mathcal{V}^{-1/3}$ の立方体の箱にはいついて周期境界条件を課すとする。定圧定温 MC シミュレーション [44] では、立方体の一辺の長さを単位とするスケールされた座標 $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ と体積 $\mathcal{V}$ をメトロポリス判定に従って更新していく。ここで、

$$\sigma_k = \mathcal{V}^{-1/3} \mathbf{q}_k \quad (k = 1, 2, \dots, N) \quad (1.229)$$

である ($\mathbf{q}_k$ は実座標)。この表現は、圧力を制御するピストンは一方向に移動するのではなく、立方体の三方向から等方的に圧縮または膨張する仮想的なピストンを考えていることを意味する。

座標 σ と体積 $\mathcal{V}$ をもつ状態 x から座標 σ' と体積 $\mathcal{V}'$ をもつ状態 x' へのメトロポリスの遷移確率は以下で与えられる。

$$w(x \rightarrow x') = \min \left(1, \exp \left[-\beta \left(\mathcal{H}' - \mathcal{H} - N k_B T \log \frac{\mathcal{V}'}{\mathcal{V}} \right) \right] \right). \quad (1.230)$$

ここで、最後の対数の項は以下の関係式から出てきたことに注意されたい。

$$e^{-\beta \mathcal{H}} d\mathbf{q} = e^{-\beta \mathcal{H}} \mathcal{V}^N d\sigma = \exp[-\beta(\mathcal{H} - N k_B T \log \mathcal{V})] d\sigma. \quad (1.231)$$

定圧定温アンサンブルにおける MD シミュレーションとして、ここでは能勢・アンダーソンの方法を紹介する [4, 45]。式 (1.81)～式 (1.84) の運動方程式は以

84 第1章 拡張アンサンブルの方法

下のように一般化される.

$$\dot{\mathbf{q}}_k = \frac{\mathbf{p}_k}{m_k} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \mathbf{q}_k, \quad (1.232)$$

$$\dot{\mathbf{p}}_k = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_k} - \left(\frac{\dot{s}}{s} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \right) \mathbf{p}_k \quad (1.233)$$

$$= \mathbf{F}_k - \left(\frac{\dot{s}}{s} + \frac{\dot{\mathcal{V}}}{3\mathcal{V}} \right) \mathbf{p}_k, \quad (1.234)$$

$$\dot{s} = s \frac{P_s}{Q}, \quad (1.235)$$

$$\dot{P}_s = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - 3Nk_B T = 3Nk_B (T(t) - T), \quad (1.236)$$

$$\dot{\mathcal{V}} = s \frac{P_{\mathcal{V}}}{M_{\mathcal{V}}}, \quad (1.237)$$

$$\dot{P}_{\mathcal{V}} = s \left[\frac{1}{3\mathcal{V}} \left(\sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - \sum_{i=1}^N \mathbf{q}_i \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} \right) - \frac{\partial \mathcal{H}}{\partial \mathcal{V}} \right] \quad (1.238)$$

$$= s(\mathcal{P}(t) - \mathcal{P}), \quad (1.239)$$

ここで, $M_{\mathcal{V}}$ は体積に対応する「質量」, $P_{\mathcal{V}}$ は体積の共役運動量, 「瞬間圧力」 $\mathcal{P}(t)$ は以下で定義される.

$$\mathcal{P}(t) = \frac{1}{3\mathcal{V}} \left(\sum_{i=1}^N \frac{\mathbf{p}_i(t)^2}{m_i} + \sum_{i=1}^N \mathbf{q}_i(t) \cdot \mathbf{F}_i(t) \right) - \frac{\partial E}{\partial \mathcal{V}}(t). \quad (1.240)$$

定圧定温アンサンブルにおけるレプリカ交換法 (isobaric-isothermal replica-exchange method: PTREM) [46] では, 互いに相互作用しない M ($M = M_0 \times M_1$) 個のレプリカを用意する. ここで, M_0 と M_1 は, それぞれ, シミュレーションで使われる温度値と圧力値の数である. 各変数のラベルとしては以下を使う. レプリカは, i ($= 1, 2, \dots, M$), 温度は m_0 ($= 1, 2, \dots, M_0$), 圧力は m_1 ($= 1, 2, \dots, M_1$) で表すことにする.

一般性を失うことなく, パラメータの大きさが次のようになっていると仮定できる. $T_1 < T_2 < \dots < T_{M_0}$, $P_1 < P_2 < \dots < P_{M_1}$. 定圧定温レプリカ交換シミュレーションは, 次の二つのステップを交互に繰り返すことで実行される.

1.6 多次元拡張アンサンブル法の実例 85

1. 各レプリカ i ($i = 1, \dots, M$) において, 一定温度 T_{m_0} と一定圧力 $\mathcal{P}_{m_1}$ における従来の定圧定温 MC または MD シミュレーションを同時にあるステップ数実行する.
2. 隣接したパラメータ値 (m と n) に対応するレプリカ対 (i と j) を以下の遷移確率で交換する.

$$w_{PTREM} = \min[1, \exp(-\Delta_{PTREM})], \quad (1.241)$$

ここで, Δ_{PTREM} は以下で与えられる.

$$\begin{aligned} \Delta_{PTREM} = & (\beta_{m_0} - \beta_{n_0}) [E(\sigma^{[j]}, \mathcal{V}^{[j]}) - E(\sigma^{[i]}, \mathcal{V}^{[i]})] \\ & + (\beta_{m_0} \mathcal{P}_{m_1} - \beta_{n_0} \mathcal{P}_{n_1}) (\mathcal{V}^{[j]} - \mathcal{V}^{[i]}) . \end{aligned} \quad (1.242)$$

ここで, 式 (1.231) の $\log \mathcal{V}$ の項は打ち消し合ったことに注意されたい. そして, T 交換の場合は

$$\begin{aligned} n_0 &= m_0 + 1 \\ n_1 &= m_1 \end{aligned} \quad (1.243)$$

と置き, $\mathcal{P}$ 交換の場合は

$$\begin{aligned} n_0 &= m_0 \\ n_1 &= m_1 + 1 \end{aligned} \quad (1.244)$$

と置く.

定圧定温アンサンブルにおける焼き戻し法 (isobaric-isothermal simulated tempering: PTST) [22] においては, 関数 $g(T, \mathcal{P})$ を導入し, 以下のように, 温度 T と圧力 $\mathcal{P}$ の一様分布が得られるような重み因子 $W_{PTST}(E, \mathcal{V}; T, \mathcal{P}) \equiv \exp[-\beta(E + \mathcal{P}\mathcal{V}) + g(T, \mathcal{P})]$ を使う.

$$P_{PTST}(T, \mathcal{P}) = \int_0^\infty d\mathcal{V} \int_{\mathcal{V}} dq W_{PTST}[E(q, \mathcal{V}), \mathcal{V}; T, \mathcal{P}] = \text{constant} . \quad (1.245)$$

式 (1.245) から, 以下が得られる.

$$e^{-g(T, \mathcal{P})} \propto \int_0^\infty d\mathcal{V} \int_{\mathcal{V}} dq \exp[-\beta(E(q, \mathcal{V}) + \mathcal{P}\mathcal{V})], \quad (1.246)$$

86 第1章 拡張アンサンブルの方法

よって, $g(T, \mathcal{P})$ は (運動エネルギー項の寄与を除いた) 無次元化されたギブス自由エネルギー (Gibbs free energy) である.

定圧定温アンサンブルにおける焼き戻しシミュレーションでは, レプリカ交換法の時と同じように, 温度と圧力を以下のように離散化する. $(T_{m_0}, \mathcal{P}_{m_1})$ ($m_0 = 1, \dots, M_0, m_1 = 1, \dots, M_1$). また, $M = M_0 \times M_1$ である. そして, 大小関係は $T_1 < T_2 < \dots < T_{M_0}, \mathcal{P}_1 < \mathcal{P}_2 < \dots < \mathcal{P}_{M_1}$ とする.

初期状態とパラメータの初期値が用意されたら, 定圧定温アンサンブル焼き戻しシミュレーションは以下の二つのステップを交互に繰り返すことによって実行される.

1. 固定されたパラメータ値 $(T_{m_0}, \mathcal{P}_{m_1})$ を持つ重み因子 $\exp(-\beta_{m_0}(E + \mathcal{P}_{m_1}\mathcal{V}))$ に基づく従来の定圧定温 MC または MD シミュレーションを数ステップ実行する.
2. 状態を固定したまま, パラメータセット $(T_{m_0}, \mathcal{P}_{m_1})$ を新しい値 $(T_{n_0}, \mathcal{P}_{n_1})$ に更新する. このパラメータ更新の遷移確率は以下で与えられる.

$$w_{PTST} = \min[1, \exp(-\Delta_{PTST})], \quad (1.247)$$

ここで, Δ_{PTST} は以下で与えられる.

$$\Delta_{PTST} = (\beta_{n_0} - \beta_{m_0})E(\sigma, \mathcal{V}) + (\beta_{n_0}\mathcal{P}_{n_1} - \beta_{m_0}\mathcal{P}_{m_1})\mathcal{V} - (g_{n_0, n_1} - g_{m_0, m_1}). \quad (1.248)$$

ここで, 式 (1.231) の $\log \mathcal{V}$ の項は打ち消し合ったことに注意されたい. そして, T 更新の場合は

$$\begin{aligned} n_0 &= m_0 \pm 1 \\ n_1 &= m_1 \end{aligned} \quad (1.249)$$

と置き, $\mathcal{P}$ 更新の場合は

$$\begin{aligned} n_0 &= m_0 \\ n_1 &= m_1 \pm 1 \end{aligned} \quad (1.250)$$

と置く.

1.6 多次元拡張アンサンブル法の具体例 87

ここで、定圧定温レプリカ交換法（または、定圧定温焼き戻し法）のステップ1において、MD シミュレーションを実行する場合は、ステップ2で温度が交換（更新）された場合、粒子の運動量をスケールしなければならないことに注意されたい（式 (1.116), 式 (1.130) 参照）[20, 21, 22].

定圧定温アンサンブルにおける長いレプリカ交換や焼き戻しシミュレーションのあと、物理量 A の $(T_{m_0}, \mathcal{P}_{m_1})$ ($m_0 = 1, \dots, M_0$, $m_1 = 1, \dots, M_1$) における期待値は以下のいつもの算術平均で得られる.

$$\langle A \rangle_{T_{m_0}, \mathcal{P}_{m_1}} = \frac{1}{n_m} \sum_{k=1}^{n_m} A(x_m(k)) , \quad (1.251)$$

ここで、 $x_m(k)$ ($k = 1, \dots, n_m$) はパラメータ値 $(T_{m_0}, \mathcal{P}_{m_1})$ において得られた状態であり、 n_m はその総数である.

また、任意の中間の温度 $T (= 1/k_B \beta)$ と圧力 $\mathcal{P}$ における A の期待値は以下で与えられる.

$$\langle A \rangle_{T, \mathcal{P}} = \frac{\sum_{E, \mathcal{V}} A(E, \mathcal{V}) n(E, \mathcal{V}) \exp(-\beta(E + \mathcal{P}\mathcal{V}))}{\sum_{E, \mathcal{V}} n(E, \mathcal{V}) \exp(-\beta(E + \mathcal{P}\mathcal{V}))} , \quad (1.252)$$

ここで、状態密度 $n(E, \mathcal{V})$ は多ヒストグラム再重法で得られる. すなわち、 $(T_{m_0}, \mathcal{P}_{m_1})$ におけるレプリカ交換または焼き戻しシミュレーションで得られたポテンシャルエネルギーと体積の分布のヒストグラムとサンプル総数を、それぞれ $N_{m_0, m_1}(E, \mathcal{V})$ と n_{m_0, m_1} とするとき、状態密度 $n(E, \mathcal{V})$ と無次元化されたギブス自由エネルギー g_{m_0, m_1} は以下の WHAM 方程式を自己無撞着に逐次的に解くことによって得られる（式 (1.176) と式 (1.177) 参照）.

$$n(E, \mathcal{V}) = \frac{\sum_{m_0, m_1} N_{m_0, m_1}(E, \mathcal{V})}{\sum_{m_0, m_1} n_{m_0, m_1} \exp(g_{m_0, m_1} - \beta_{m_0}(E + \mathcal{P}_{m_1}\mathcal{V}))} , \quad (1.253)$$

$$\exp(-g_{m_0, m_1}) = \sum_{E, \mathcal{V}} n(E, \mathcal{V}) \exp(-\beta_{m_0}(E + \mathcal{P}_{m_1}\mathcal{V})) . \quad (1.254)$$

88 第1章 拡張アンサンブルの方法

最後に、定圧定温アンサンブルにおけるマルチカノニカル法に対応するマルチバーリック・マルチサーマル法 (multibaric-multithermal algorithm: MUBATH) [47, 48] を紹介する. この拡張アンサンブル法では、ポテンシャルエネルギー空間と体積空間の酔歩が実現される.

MUBATH アンサンブルでは、ポテンシャルエネルギー E と体積 $\mathcal{V}$ の一様分布が得られるように、各状態は MUBATH 重み因子 $W_{MBT}(E, \mathcal{V}) \equiv \exp\{-\beta_a \mathcal{H}_{MBT}(E, \mathcal{V})\}$ に従って発生させられる [47]. ここで、任意の参照温度 T_a ($= 1/k_B \beta_a$) を導入した. また、 $\mathcal{H}_{MBT}$ は「マルチバーリック・マルチサーマルエンタルピー」と呼ばれる. つまり、以下を要求する.

$$\begin{aligned} P_{MBT}(E, \mathcal{V}) &= n(E, \mathcal{V}) W_{MBT}(E, \mathcal{V}) \\ &= n(E, \mathcal{V}) \exp\{-\beta_a \mathcal{H}_{MBT}(E, \mathcal{V})\} \\ &\equiv \text{constant} . \end{aligned} \quad (1.255)$$

MUBATH MC シミュレーションは、式 (1.230) において $\mathcal{H}$ を $\mathcal{H}_{MBT}$ で置き換えることによって実行される. すなわち、遷移確率は以下で与えられる [47].

$$w(x \rightarrow x') = \min \left(1, \exp \left[-\beta_a \left(\mathcal{H}'_{MBT} - \mathcal{H}_{MBT} - N k_B T_a \log \frac{\mathcal{V}'}{\mathcal{V}} \right) \right] \right) , \quad (1.256)$$

また、MUBATH MD シミュレーションは、任意の参照温度 T_a において、いつもの定圧定温アンサンブルの運動方程式 (式 (1.232)~式 (1.239)) の中で、式 (1.233) と式 (1.238) のエンタルピー $\mathcal{H}$ をマルチバーリック・マルチサーマルエンタルピー $\mathcal{H}_{MBT}$ で置き換えることによって実行される [48].

長い MUBATH シミュレーションの結果は単ヒストグラム再重法で解析される. すなわち、任意の温度 T と圧力 P における物理量 A の期待値は式 (1.252) に以下の状態密度を代入することによって得られる.

$$n(E, \mathcal{V}) = \frac{N_{MBT}(E, \mathcal{V})}{W_{MBT}(E, \mathcal{V})} , \quad (1.257)$$

ここで、 $N_{MBT}(E, \mathcal{V})$ は長い MUBATH シミュレーションで得られたポテンシャルエネルギーと体積の確率分布 $P_{MBT}(E, \mathcal{V})$ のヒストグラムである.

1.7 拡張アンサンブルシミュレーションの実例

前節で述べた幾つかの拡張アンサンブル法に基づく、生体分子系のシミュレーション結果の例を紹介しよう。

まず、従来のカノニカルアンサンブル上のシミュレーションに比べて、拡張アンサンブル上のシミュレーションがどれぐらい広く配位空間をサンプルすることができるかを見てみよう。

最初の例は5個のアミノ酸からなる Met-enkephalin の拡張アンサンブルシミュレーションである。アミノ酸配列は Tyr-Gly-Gly-Phe-Met で与えられる。このペプチドは生物学的機能をもつ最小のペプチドの一つであるため、多くのシミュレーション法のベンチマークの系とされてきた。我々は真空中の Met-enkephalin のレプリカ交換 MD (REMD) シミュレーションを実行した [20]。計算プログラムとしては、PRESTO[49] バージョン 2 に基づいて、文献 [50] で開発されたプログラムに REMD を組み込んで使われた。ポテンシャルエネルギー (力場 (force field) と呼ばれる) は文献 [51] の AMBER のものを使った。MD シミュレーションの単位時間ステップは 0.5 フェムト秒とした (1 フェムト秒とは 10^{-15} 秒のことである)。8 個のレプリカを用意し、完全に伸びた構造を初期構造として、それぞれのレプリカにおいて、1 ナノ秒 (よって、 2×10^6 回の時間ステップ) の間、時間発展させた (1 ナノ秒とは 10^{-9} 秒のことである)。8 個の温度は徐冷法のときの徐冷スケジュール [5, 7] と同じように、指数関数的に分布させ (式 (1.121), (1.122) 参照), 以下の値 T_m ($m = 1, \dots, 8$) をとった。200, 239, 286, 342, 409, 489, 585, 700 K。そして、レプリカ交換は 10 フェムト秒毎に試行し、レプリカ交換の寸前の状態を後の結果解析のために保存した。

図 1.15 に一つのレプリカにおける温度とポテンシャルエネルギーの時間発展を示す。それぞれ、温度空間とポテンシャルエネルギー空間の酔歩が得られていることが分かる。図 1.15(a) と図 1.15(b) を比較すると、これらの酔歩には強い正の相関があることに注目されたい。

レプリカ交換法が効率良く実行されるためには、それぞれの温度対についてのレプリカ交換の受諾率が十分に一樣で大きいことである (例えば, $> 10\%$)。

90 第 1 章 拡張アンサンブルの方法

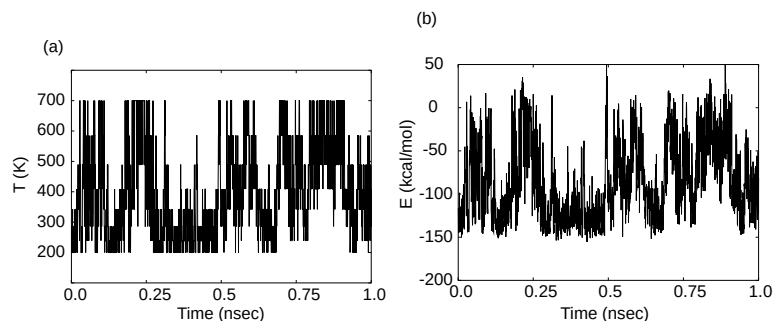


図 1.15 真空中の Met-enkephalin のレプリカ交換 MD シミュレーションの一つのレプリカにおける (a) 温度 T (K) と (b) ポテンシャルエネルギー E (kcal/mol) の「時間発展」.

本シミュレーションでは、全ての温度対について、レプリカ交換の受諾率が約 15 %で一様であった。

図 1.16 にはこのレプリカ交換 MD シミュレーションで得られた八つの温度におけるポテンシャルエネルギーのカノニカル分布を示す。全ての隣り合う温度対の分布が十分な重なりをもっており（これは、レプリカ交換シミュレーションが効率良く実行されるための確認事項である）、レプリカ交換が頻繁に行われたことが分かる。

図 1.17 では、このペプチドの 2 番目のアミノ酸の主鎖の二面角対 (ϕ, ψ) の分布について、従来の一定温度のカノニカル MD シミュレーションとレプリカ交換 MD シミュレーションの結果を低温 ($T = 200$ K) で比較した。従来のシミュレーションがエネルギー極小状態に留まってしまって、二面角の分布が一つの大きなピークを持つだけであるが、レプリカ交換シミュレーションで得られた分布は上のピーク以外にも幾つか他のピークが得られていて、配位空間が広くサンプルされていることが分かる。

物理量のいろいろな温度における期待値については、式 (1.118) と (1.119) の多ヒストグラム再重法が使われた。図 1.18 に平均ポテンシャルエネルギーを温度の関数として示した。比較のために、8 個の温度において、それぞれ、従来のカノニカルアンサンブル上の 1 ナノ秒の MD シミュレーションの結果も同じ図

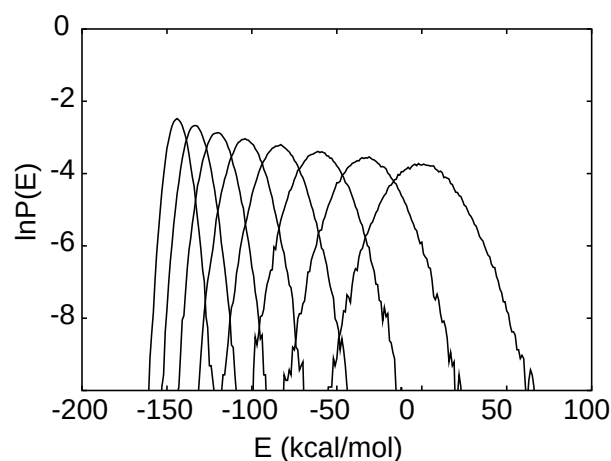


図 1.16 真空中の Met-enkephalin のレプリカ交換 MD シミュレーションから得られたポテンシャルエネルギーのカノニカル分布. 対応する温度は左から 200, 239, 286, 342, 409, 489, 585, 700 K である.

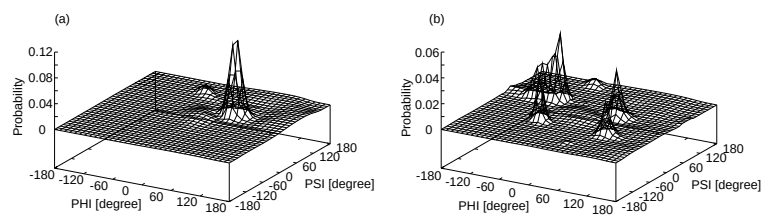


図 1.17 温度 $T = 200$ K における真空中の Met-enkephalin の 2 番目のアミノ酸の主鎖の二面角対 (ϕ, ψ) の分布. (a) 従来の温度一定のカノニカル MD シミュレーションから得られた結果. (b) レプリカ交換 MD シミュレーションから得られた結果.

92 第1章 拡張アンサンブルの方法

に示した．高温領域では両者のシミュレーションの結果が一致しているが，低温の領域では，従来のカノニカルシミュレーションの結果が少し数値が大きくなって，シミュレーションがエネルギー極小状態に留まってしまっていることが分かる．

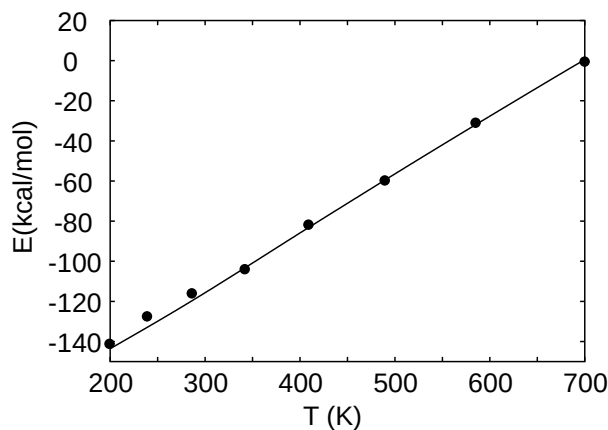


図 1.18 温度の関数としての真空中の Met-enkephalin の平均ポテンシャルエネルギー．実線がレプリカ交換 MD シミュレーションの結果に多ヒストグラム再重法を適用して得られたものであり，点は従来のカノニカル MD シミュレーションの結果である．

二つ目の例は，リボヌクレアーゼ T_1 という蛋白質の一部である 17 残基のペプチドフラグメントのマルチカノニカルシミュレーションである．このペプチドフラグメントは，実験によって α ヘリックス構造と呼ばれる螺旋構造を取ることが知られている [52]．ペプチドのポテンシャルエネルギーとしては ECEPP/2[53] を使った．また，溶媒効果は（水素以外の）重い原子の溶媒接触表面積に比例する項の和で与えた [54]．そして，この表面積の計算プログラムとしては NSOL[55] を使った．モンテカルロプログラムとしては，KONF90 [7] を修正して，マルチカノニカル法を組み入れた．

この系において，低温 ($T = 200$ K) のカノニカル MC シミュレーションとマルチカノニカル MC シミュレーションを実行した [13, 31]．図 1.19 にこれらのシミュレーションにおけるポテンシャルエネルギーの「経時変化」を示す．従来のカノニカルシミュレーションは，なかなか平衡に達しないのに対し，マル

1.7 拡張アンサンブルシミュレーションの実例 93

チカノニカルシミュレーションでは、ポテンシャルエネルギー空間上の幅広い酔歩を実現していることが分かる。

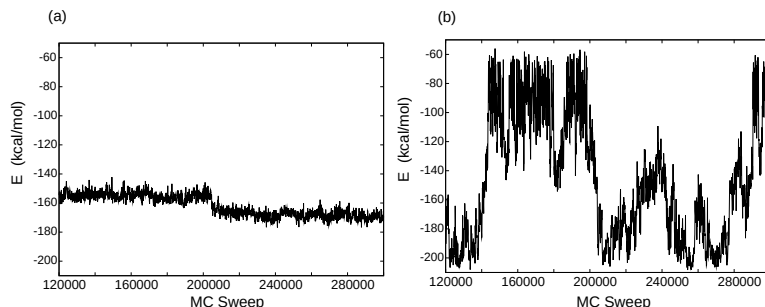


図 1.19 リボスクレアーゼ T_1 のペプチドフラグメントのポテンシャルエネルギーの「経時変化」(120000 MC sweeps から 300000 MC sweeps まで). (a) 温度 $T = 200$ K における従来のカノニカル MC シミュレーション. (b) マルチカノニカル MC シミュレーション.

120000 MC sweeps から 300000 MC sweeps に至る間からそれぞれ四つの構造を抽出して、図 1.20 と図 1.21 に示した. 図 1.20 のカノニカルシミュレーションがエネルギー極小状態に留まってしまっているのに対し、マルチカノニカルシミュレーションでは、図 1.21(a), (b), (c), (d) と時間発展するにつれて、 α ヘリックス構造とランダムコイル状態の間を行き来して、幅広い構造空間をサンプルしていることが分かる。

三つ目の例は、リボスクレアーゼ A という蛋白質の N 末端の 13 残基からなる C ペプチドの拡張アンサンブルシミュレーションである [56]. C ペプチドのアミノ酸配列は次のとおりである. Ala-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Leu-Arg-Ala-His-Ala [57]. 力場は AMBER parm99 [58] を用いた. MD シミュレーションの単位時間ステップは 0.5 フェムト秒とした. まず、完全に伸びた構造から高温 ($T = 1000$ K) で真空中のシミュレーションを実行して一つの (α ヘリックスや β シートなどの二次構造を持たない) ランダムコイル構造を得た. この立体構造を持つ C ペプチド 1 個を半径 22 Å の水球 (水分子は 1387 個) に入れて、初期構造とした (全原子数は、4365 個). 水分子としては TIP3P 模型 [59] を使用した. 図 1.22 にこの初期構造を示した.

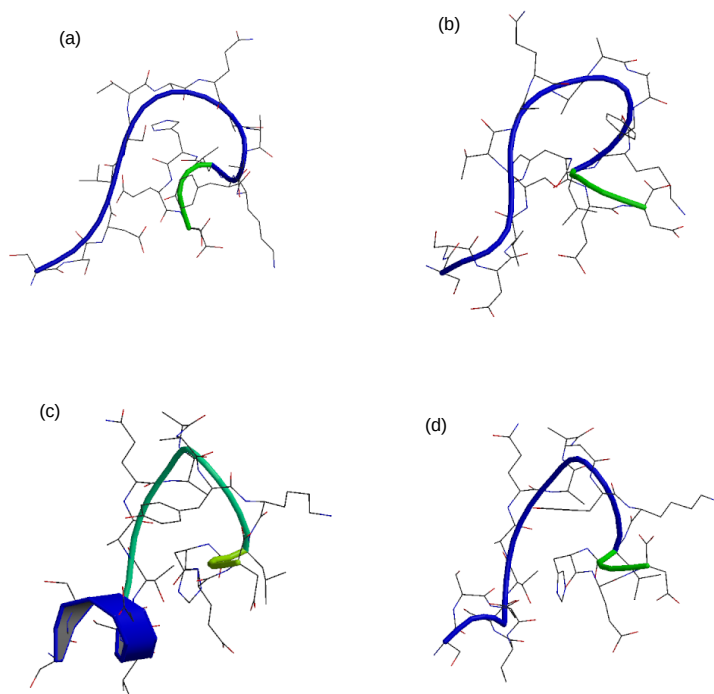


図 1.20 図 1.19(a) のリボヌクレアーゼ T_1 のペプチドフラグメントのカノニカル MC シミュレーションのスナップショット. 対応する MC Sweep 数は (a) 138000, (b) 190000, (c) 243000, (d) 295000.

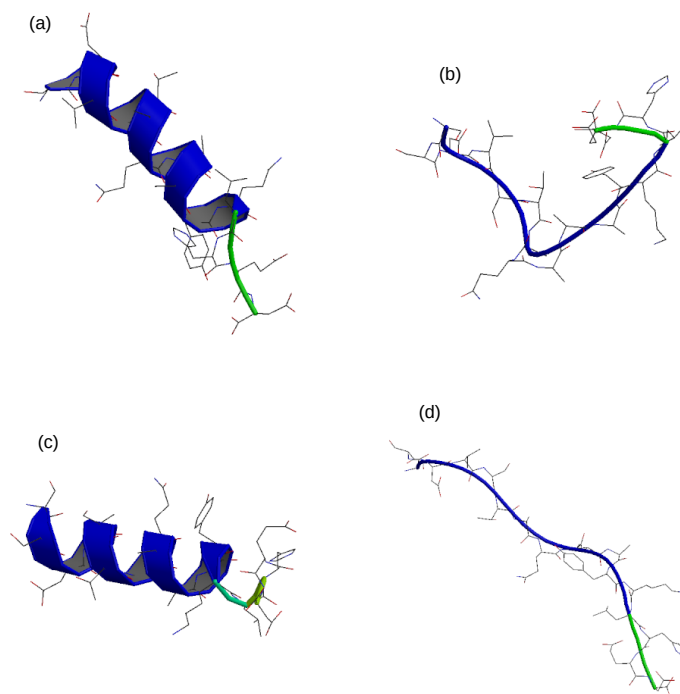


図 1.21 図 1.19(b) のリボヌクレアーゼ T_1 のペプチドフラグメントのマルチカノニカル MC シミュレーションのスナップショット. 対応する MC Sweep 数は図 1.20 と同じである.

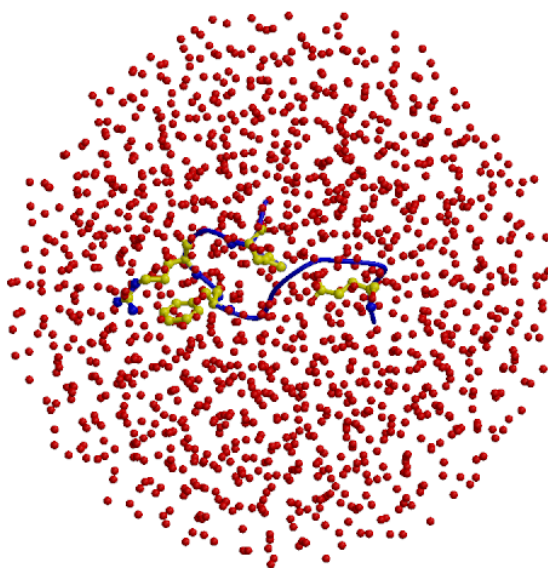


図 1.22 リボヌクレアーゼ A の C ペプチドのレプリカ交換 MD シミュレーションの初期構造. ペプチドの主鎖は青い線で表し, 側鎖は Glu-2, Phe-8, Arg-10, His-12 のものだけを黄色の線で表した. また, 赤い小さな球は 1387 個の水分子の酸素原子を表す.

1.7 拡張アンサンブルシミュレーションの実例 97

表 1.1 リボヌクレアーゼ A の C ペプチドの拡張アンサンブルシミュレーションのまとめ.

シミュレーション	レプリカ数 M	温度 T_m (K) ($m = 1, \dots, M$)	レプリカ毎の MD ステップ数
REMD1	32	250, 258, 267, 276, 286, 295, 305, 315, 326, 337, 348, 360, 372, 385, 398, 411, 425, 440, 455, 470, 486, 502, 519, 537, 555, 574, 593, 613, 634, 655, 677, 700	2.0×10^5
MUCAREM1	4	360, 440, 555, 700	2.0×10^6
REMUCA1	1	700	3.0×10^7

また, 表 1.1 にこの系で実行された三つの拡張アンサンブルシミュレーションの条件をまとめた.

最初のシミュレーションはレプリカ数が 32 個のレプリカ交換 MD シミュレーションである (表 1.1 の REMD1). 各レプリカ毎に 100 ピコ秒 (よって, 2.0×10^5 MD ステップ) の MD シミュレーションを実行した (1 ピコ秒とは 10^{-12} 秒のことである). レプリカ交換は, 200 MD ステップ毎に試行された. その結果に式 (1.118) と式 (1.119) の多ヒストグラム再重法を適用して, マルチカノニカル重み因子 (つまり, 状態密度の逆数) の 1 回目の推測値を決めた. このマルチカノニカル重み因子を四つの違うエネルギー領域に分けて定義した (式 (1.143) 参照). 四つの (隣同士に重なりを持つ) エネルギー領域は, 以下のとおりであった. $-13791.5 \sim -11900.5$ kcal/mol, $-12962.5 \sim -10796.5$ kcal/mol, $-11900.5 \sim -9524.5$ kcal/mol, $-10796.5 \sim -8293.5$ kcal/mol.

この重み因子を用いて, 次に, マルチカノニカルレプリカ交換 MD シミュレー

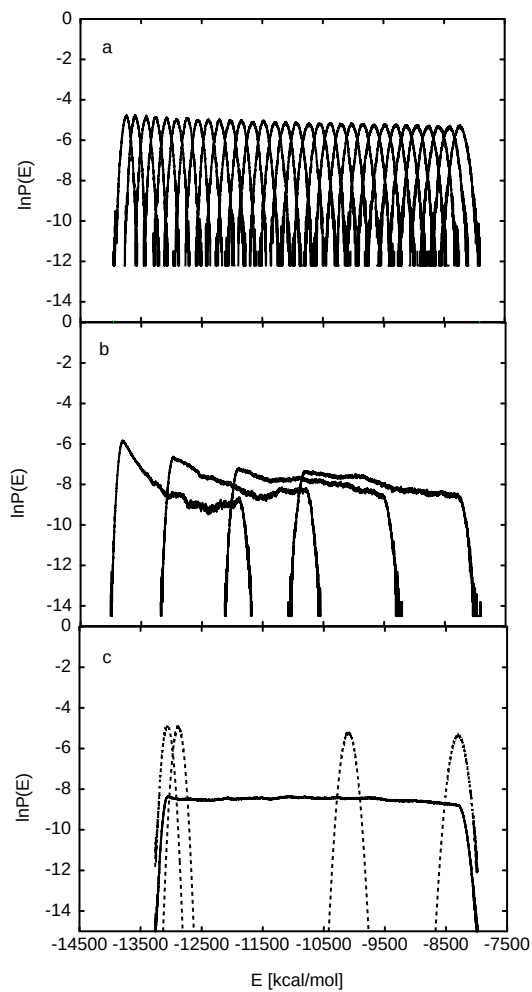


図 1.23 三つの拡張アンサンブルシミュレーションの結果から得られた水中の C ペプチドのポテンシャルエネルギーの確率分布. 拡張アンサンブル法は, (a) REMD1, (b) MUCAREM1, and (c) REMUCA1 であり, 詳細は表 1.1 に示した. (c) の中のダッシュ線は再重法で得られた温度 290, 300, 500, 700 K (左から右) におけるカノニカル分布である.

1.7 拡張アンサンブルシミュレーションの実例 99

ション (表 1.1 の MUCAREM1) を実行した. 各レプリカ毎に 1 ナノ秒 (よって, 2.0×10^6 MD ステップ) の MD シミュレーションを実行した. レプリカ交換は, 1000 MD ステップ毎に試行された. そして, その結果に多ヒストグラム再重法を再び適用して, より精度の高いマルチカノニカル重み因子を決めた.

そして, この重み因子を用いて, 最後に, マルチカノニカル MD シミュレーション (表 1.1 の REMUCA1) を 15 ナノ秒間 (よって, 3.0×10^7 MD ステップ) 実行し, 結果を解析した.

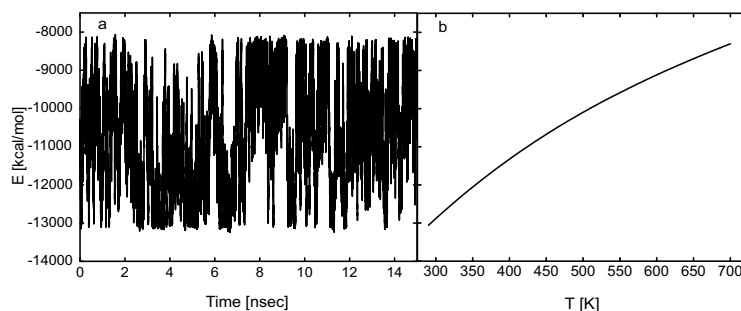


図 1.24 REMUCA1 シミュレーションにおける水中の C ペプチドの (a) ポテンシャルエネルギーの「経時変化」と (b) 温度の関数としての平均ポテンシャルエネルギー. 後者は REMUCA1 の結果に単ヒストグラム再重法を適用して得られた.

三つの拡張サンサンブルシミュレーション REMD1, MUCAREM1, REMUCA1 で得られたポテンシャルエネルギーの確率分布を図 1.23 に示した. 図 1.23(a) では, 全ての隣り合うカノニカル分布に十分な重なりがあり, REMD1 において, 十分なレプリカ交換が実施されたことが示唆される. 図 1.23(b) では, 隣り合うマルチカノニカル分布に十分な重なりがあり, MUCAREM1 も正しく実行されたことが示唆される. そして, 図 1.23(c) では, -13000 kcal/mol と -8000 kcal/mol の間において一様な分布が得られており, このエネルギー領域において自由酔歩が実現されていることが示唆される.

図 1.24(a) に REMUCA1 におけるポテンシャルエネルギーの「経時変化」を示す. 実際, 5000 kcal/mol にも及ぶ領域をカバーする酔歩が得られている. 図 1.24(b) には, REMUCA1 の結果に単ヒストグラム再重法を適用して得られた温度の関数としての平均ポテンシャルエネルギーを示す. 温度とともに単調増

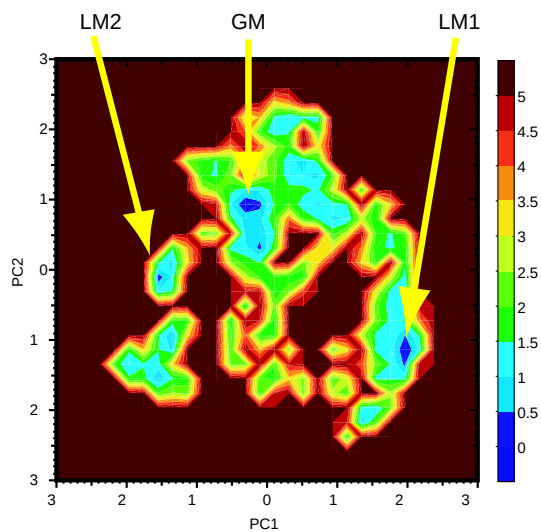


図 1.25 温度 300 K における水中の C ペプチドの最初の二つの主成分についての平均力ポテンシャル (kcal/mol). 自由エネルギーは REMUCA1 の結果に単ヒストグラム再重法を適用して得られた. GM, LM1, LM2 は三つの自由エネルギー極小状態を示す.

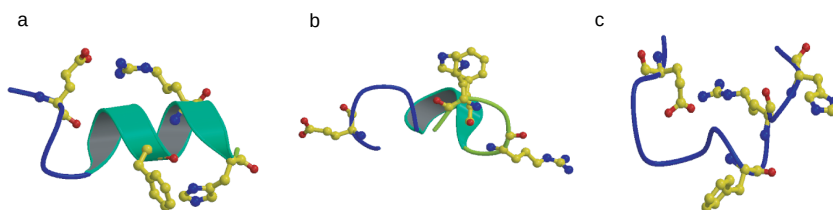


図 1.26 C ペプチドの自由エネルギー最小状態 ((a) GM) と二つの局所的自由エネルギー極小状態 ((b) LM1 and (c) LM2) の代表的立体構造. 主鎖の構造とともに、四つのアミノ酸 Glu⁻-2, Phe-8, Arg⁺-10, His⁺-12 の側鎖のみが棒球図で示されている.

加していることが分かる。

図 1.25 にこの系の温度 300 K における、主成分解析 (principal component analysis: PCA)[60] の結果得られた二つの主成分についての平均力ポテンシャル (反応座標の関数としての自由エネルギー) を示す (主成分解析法については、本シリーズの第 6 巻第 2 章に解説があるので参照されたい)。この自由エネルギー地形には三つの自由エネルギー極小状態が存在する。これらの自由エネルギー極小状態の代表的立体構造は図 1.26 に示した。最小自由エネルギー状態 (GM) の構造は、二周の α ヘリックス構造と Glu⁻-2 と Arg⁺-10 の側鎖間の塩橋 (salt bridge) を形成している。また、Phe-8 と His⁺-12 の側鎖のコンタクトも示唆される。これらは、X 線回折実験や NMR 実験などの結果と良く一致している。局所的自由エネルギー極小状態の一つ (LM1) では、Phe-8 と His⁺-12 の側鎖のコンタクトは存在するが、Glu⁻-2 と Arg⁺-10 の側鎖間の塩橋は存在しない。一方、局所的自由エネルギー極小状態のもう一つ (LM2) では、逆に、Glu⁻-2 と Arg⁺-10 の側鎖間の塩橋は存在するが、Phe-8 と His⁺-12 の側鎖のコンタクトは存在しない。よって、最小自由エネルギー状態 GM は実験結果と最も一致しているわけである。

四つ目の例は、水中の小蛋白質の折り畳みの MUCAREM シミュレーションである [61]。系は水分子をあらわに取り入れた、トリヴィリンヘッドピースサブドメイン (chicken villin headpiece subdomain) である (以下では、ヴィリンと呼ぶことにする)。アミノ酸数は 36 個である。力場としては、蛋白質では CMAP [62] を取り入れた CHARMM22 [63] を、また、水では TIP3P[59, 63] を用いた。計算プログラムとしては、PRESTO[49] バージョン 2 に基づいて、文献 [50, 20, 30] で開発されたプログラムを更に CHARMM 力場でも計算できるように改良して使われた。MD シミュレーションの単位時間ステップは 1.0 フェムト秒であった。

まず、水分子なしで、距離に依存した誘電率による粗い溶媒効果を取り入れた 90 個のレプリカに基づくレプリカ交換分子動力学 (REMD) シミュレーションを実行した。温度は 250 K~2000 K の範囲をカバーした。初期構造は完全に伸びきった構造であった。このシミュレーションの高温で α ヘリックスなどの二次構造を一切持たないほどけた構造を (ランダムに) 選び、それを 3513 個

102 第1章 拡張アンサンブルの方法

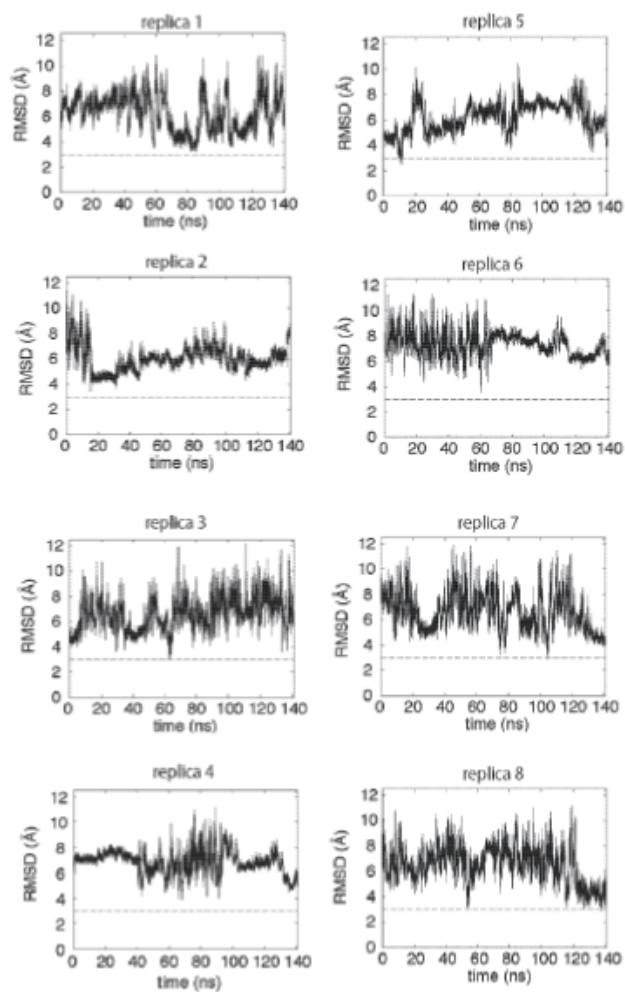


図 1.27 MUCAREM1 シミュレーションにおける水中のヴィリンの主鎖の自然構造からの RMSD の「時間発展」.

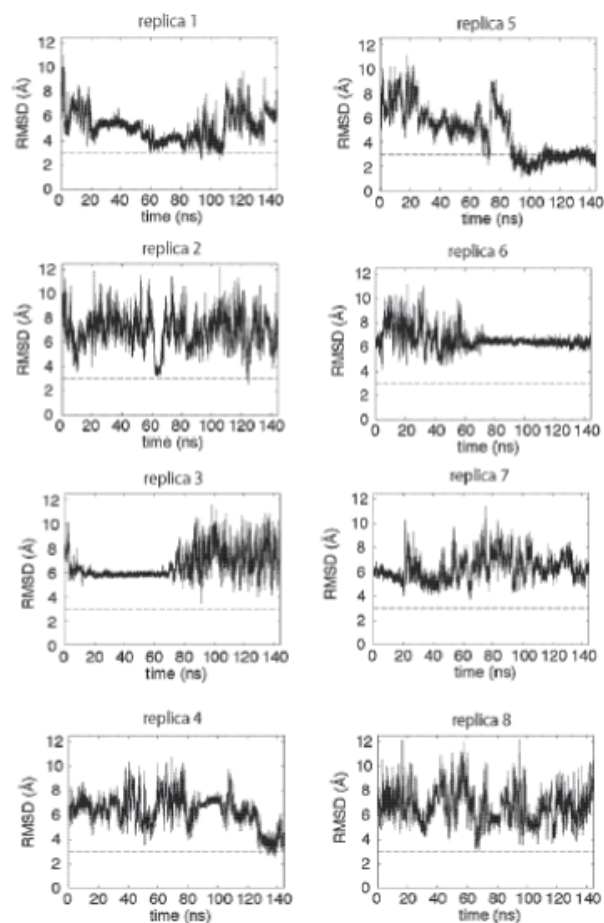


図 1.28 MUCAREM2 シミュレーションにおける水中のヴィリンの主鎖の自然構造からの RMSD の「時間発展」.

104 第1章 拡張アンサンブルの方法

の水分子からなる半径 30 Å の水球の中に入れた．そして，128 個のレプリカからなる REMD シミュレーションを実行した．ここでは，温度は 250 K～700 K の範囲をカバーした．

この水中の REMD シミュレーションの結果に多ヒストグラム再重法を適用してマルチカノニカル重み因子を求め，MUCAREM シミュレーションに移った．その結果にまた多ヒストグラム再重法を適用することを続けて，マルチカノニカル重み因子の精度を上げていった．

そして，最終のシミュレーションとしては，約 1 マイクロ秒の 8 個のレプリカからなる MUCAREM シミュレーションが初期条件を変えて二回実行された．それらを，MUCAREM1 と MUCAREM2 と呼ぶことにする．前者は 269 K～699 K の温度領域をカバーした 1.127 マイクロ秒（レプリカ毎には 140.9 ナノ秒）のシミュレーションであり，後者は 289 K～699 K の温度領域をカバーした 1.157 マイクロ秒（レプリカ毎には 144.6 ナノ秒）のシミュレーションであった．最初の水中の REMD シミュレーションでは 128 個ものレプリカを要したのに対し，最後の水中の MUCAREM シミュレーションでは 8 個のレプリカのみで十分であったことに注目されたい．

図 1.27 と図 1.28 に，それぞれ，MUCAREM1 と MUCAREM2 における，主鎖の自然の天然構造^{†3}からの根平均二乗距離（root-mean square distance: RMSD）の時間発展を示す．ここで，問題とする立体構造 q の参照構造 q^0 からの根平均二乗距離は以下で定義される．

$$RMSD = \min_{\text{回転}} \sqrt{\frac{1}{N} \sum_{k=1}^N (\mathbf{q}'_k - \mathbf{q}^0_k)^2}. \quad (1.258)$$

ここで， q' は q を平行移動によって，参照構造 q^0 と幾何中心（重心の計算において全ての原子の質量を 1 とおいたもの）が一致するように座標変換したものである．ここで，最小化の操作は， q' を幾何中心の周りに回転させることによる最小化のことである．この定義から分かるように，もし， $RMSD = 0$ なら，問題としている立体構造と参照構造が平行移動と回転の後に，完全に重なり合

^{†3}天然の蛋白質は生理的な条件下において、自由エネルギー最小の状態にあるときに秩序のある立体構造をとることが多いが、そうした立体構造を天然構造と呼ぶ。

1.7 拡張アンサンブルシミュレーションの実例 105

うことを意味しており、全く同じ立体構造をしていることを意味する。よって、根平均二乗距離は上で定義されたオーバーラップと同じように、対象構造がどれぐらい参照構造に似ているかを表す指標である。

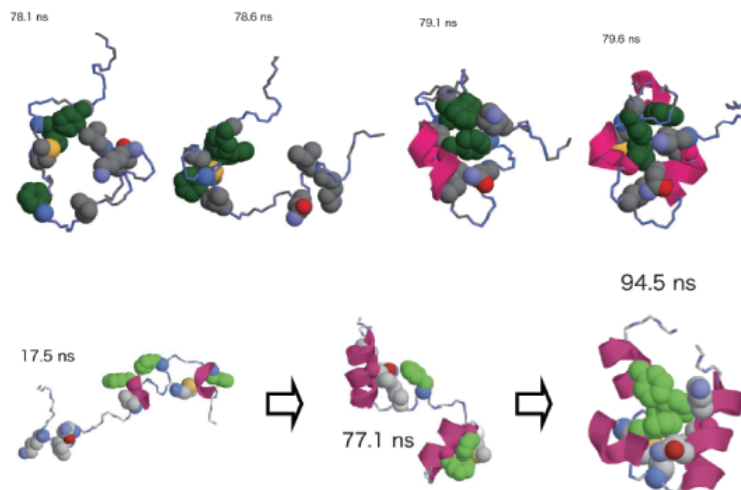


図 1.29 水中のヴィリンの MUCAREM シミュレーション中に自然構造に似た構造へ折り畳んだスナップショット. MUCAREM1 (上) と MUCAREM2 (下).

天然構造からの主鎖の RMSD が、 $6.5 \text{ \AA}$ 以上のほどけた状態から $3.0 \text{ \AA}$ 以下になったとき、天然構造に折り畳んだと考えることにする。この基準によれば、7 個のレプリカにおいて、合計 11 回の折り畳みを観察したことになる (MUCAREM1 ではレプリカ 5, 7, 8, MUCAREM2 ではレプリカ 1, 2, 4, 5 においてである)。図 1.29 に二つの MUCAREM シミュレーションによって観察された自然構造に似た構造への折り畳み過程のスナップショットを示す。

図 1.30 では、天然構造と MUCAREM シミュレーションで得られた RMSD が小さい構造を比較する。これらの構造の類似性は明らかである。

五つ目の例は、レプリカ交換傘サンプル法 (REUS) のシミュレーションである [34]。系はアラニンというアミノ酸が 3 個並んだペプチド alanine-trimer ((Ala)³ と書くことにする) であった。力場は文献 [51] の AMBER のものを用

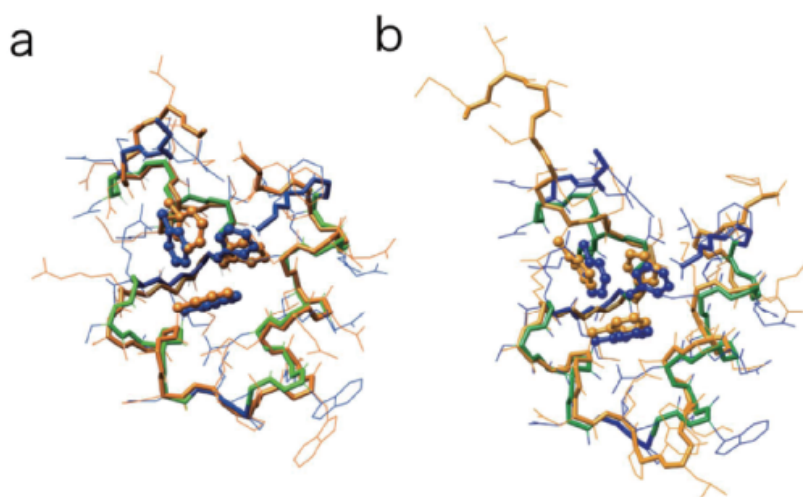


図 1.30 ヴィリンの MUCAREM1 と MUCAREM2 シミュレーションで得られた小 RMSD 構造 (オレンジ色). X 線回折実験で決められた構造 (PDB ID: 1YRF) も重ね描きしている (青色と緑色). 疎水コアを形成する三つのフェニルアラニンの側鎖 (Phe7, Phe11, Phe18) も棒球図で示した. (a) MUCAREM2 のレプリカ 5. (末端の二つの残基を除いた 34 残基の) 主鎖の RMSD は 1.1 Å であった. (b) MUCAREM1 のレプリカ 8. (9 番目~32 番目の残基の) 主鎖の RMSD は 1.0 Å であった.

1.7 拡張アンサンブルシミュレーションの実例 107

い、距離に依存した誘電率で溶媒効果を取り入れた (AMBER 力場については、本シリーズの第 6 巻第 2 章に解説があるので参照されたい)。MD の単位時間ステップは 0.5 フェムト秒とし、 4×10^6 の時間ステップ (よって、合計 2.0 ナノ秒) の計算を行った。データは 20 MD ステップ毎に保存したので、サンプル数の合計は 2×10^5 個であった。

表 1.2 にシミュレーションのパラメータのまとめを示す。REM1 では、レプリカ交換は 20 MD ステップ毎に試行された。REUS シミュレーションでは、レプリカ交換は 400 MD ステップ毎に試された。末端間距離 (O1 と H5 の間の距離) が反応座標 ξ とされ、式 (1.186) の傘ポテンシャルが REUS と US で使われた。

それでは、まず、レプリカ交換シミュレーションがどのように実行されたかを見てみよう。図 1.31 に REUS2 シミュレーションの間の温度と末端間距離 ξ の時間発展を示す。両方とも酔歩が実現されていることが分かる。(Ala)³ は末端

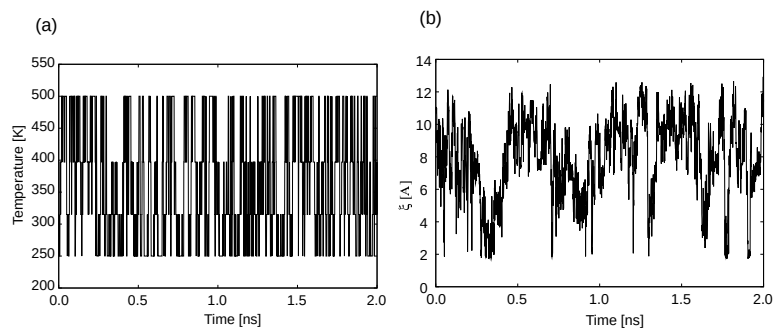


図 1.31 (Ala)³ の REUS2 シミュレーションの間の一つのレプリカの (a) 温度 T と (b) 反応座標 ξ の時間発展。

間距離 ξ の値によって、 α ヘリックス構造 ($\xi < 3 \text{ Å}$)、ターン構造 ($3 \text{ Å} < \xi < 7 \text{ Å}$)、伸びた構造 ($\xi > 7 \text{ Å}$) をとっていることが知られている [37]。図 1.31(b) はヘリックス・コイル転移がシミュレーションの間に頻繁に起こっていることを示している。

図 1.32 には温度 $T = 300 \text{ K}$ における、反応座標 ξ についての平均力ポテンシャル (PMF) を示す。この PMF の計算は、シミュレーションの結果から式

表 1.2 US, REM, REUS シミュレーションのパラメータ

シミュレーション ^a	M^b	N_T^b	T [K]	L^b	d_ℓ (k_ℓ) ^c
REM1	16	16	200, 229, 262, 299, 342, 391, 448, 512, 586, 670, 766, 876, 1002, 1147, 1311, 1500	0	
REUS1, US1	14	1	300	14	0.0 (0.0) ^d , 1.8 (1.2), 2.8 (1.2), 3.8 (1.2), 4.8 (1.2), 5.8 (1.2), 6.8 (1.2), 7.8 (1.2), 8.8 (1.2), 9.8 (1.2), 10.8 (1.2), 11.8 (1.2), 12.8 (1.2), 13.8 (1.2)
REUS2, US2	16	4	250, 315, 397, 500	4	0.0 (0.0), 7.8 (0.3), 10.8 (0.3), 13.8 (0.3)

^a REM, REUS, US は従来のレプリカ交換法, レプリカ交換傘サンプル法, 従来の傘サンプル法である.

^b M , N_T , L は全レプリカ数, 温度の数, 傘ポテンシャルの数である (式 (1.150), (1.155) 参照).

^c d_ℓ [$\text{\AA}$] と k_ℓ [kcal/mol· $\text{\AA}^2$] ($\ell = 1, \dots, L$) はそれぞれ, 傘ポテンシャルの中間点とバネ定数である (式 (1.186) 参照).

^d パラメータの値 0.0 (0.0) は傘ポテンシャルがゼロであることを意味する (つまり, $V_\ell = 0$).

(1.194) と式 (1.195) の WHAM 方程式を解き, その結果を式 (1.197) に代入することによって得られた.

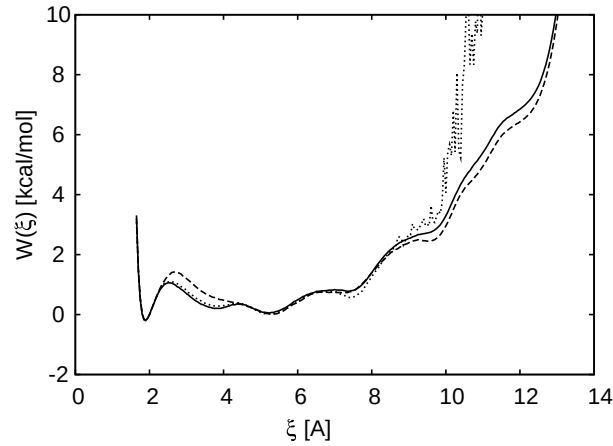


図 1.32 $(\text{Ala})^3$ の温度 $T = 300 \text{ K}$ におけるバイアスなしのオリジナルの系の反応座標 ξ についての平均力ポテンシャル. 点線, 実線, ダッシュ線は, それぞれ, REM1, REUS1, US1 の結果に対応する.

図 1.32 から分かることは, REM1 と REUS1 の結果は ξ が小さい領域 ($\xi < 7 \text{ \AA}$) でほぼ一致していることである. しかし, ξ が大きくなるにつれて ($\xi > 7 \text{ \AA}$), 少しずつ値がずれ始め, $\xi > 9 \text{ \AA}$ の領域では完全に不一致になる. この不一致の原因は, REM1 の最高温度 ($T_{16} = 1500 \text{ K}$) における反応座標の平均値は $\langle \xi \rangle_{T_{16}} \approx 8.0 \text{ \AA}$ なので, それより大きい領域は温度交換だけによる REM1 では到達できないためである. REM1 と REUS1 はとても違う条件の下での計算であるので, それらの結果が一致する領域 ($\xi < 9 \text{ \AA}$) では, 計算結果の精度が高いと思われる.

一方, US1 の結果は α ヘリックス構造とターン構造の遷移状態の領域 ($2 \text{ \AA} < \xi < 4 \text{ \AA}$) において, REM1 および REUS1 の結果よりも大きめに出ており, US1 がエネルギー極小状態に留まってしまっていることを示唆している. 温度 $T = 300 \text{ K}$ では, 最小自由エネルギー状態は $\xi \approx 2 \text{ \AA}$ に対応しており, α ヘリックス構造をとっていることが分かる.

110 第1章 拡張アンサンブルの方法

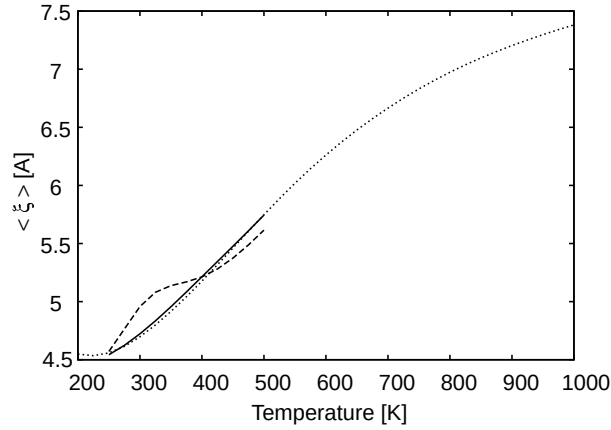


図 1.33 (Ala)³ のバイアスなしのオリジナルの系の温度の関数としての反応座標 ξ の平均値. 点線, 実線, ダッシュ線は, それぞれ, REM1, REUS2, US2 の結果に対応する.

図 1.33 にオリジナルの系の反応座標 ξ の平均値を温度の関数として示す. 平均値は式 (1.196) により計算された. REM1 と REUS2 の結果は一致しているが, US2 の結果は不正確であることが分かる.

六つ目の例は, ファンデルワールスレプリカ交換分子動力学法 (van der Waals replica-exchange molecular dynamics method: vWREMD) のシミュレーションである [39]. 系としては水中のアラニンジペプチド (alanine dipeptide) を採用した. 水分子の数は 67 個で, 力場はペプチドについては AMBER parm96[64], 水分子については TIP3P[59] を用いた. MD シミュレーションの温度制御はシンプレクティックである能勢・ポアンカレ熱浴を使った [65, 66] (シンプレクティック法については, 本シリーズの第 6 巻第 2 章に解説があるので参照されたい). 系は一辺の長さが 13.4 Å の立方体の箱の中に入れられ, 周期境界条件が課された. 静電相互作用については, エwald法を適用した (エwald法については, 本シリーズの第 6 巻第 2 章に解説があるので参照されたい). また, MD の単位時間ステップは 0.5 フェムト秒とした.

vWREMD シミュレーションでは, レプリカ数は 4 個で十分であった ($M = 4$). すなわち, 以下の四つのパラメータ値 λ_m ($m = 1, \dots, 4$) を用いた. $\lambda_1 = 0.85$,

1.7 拡張アンサンブルシミュレーションの実例 111

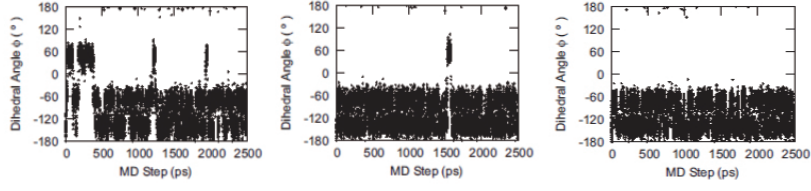


図 1.34 水中のアラニンジペプチドのシミュレーション中の主鎖の二面角 ϕ の時間発展. 左は vWREM シミュレーション, 中は REMD シミュレーション, 右は従来のカノニカル MD シミュレーションの結果である.

$\lambda_2 = 0.9$, $\lambda_3 = 0.95$, $\lambda_4 = 1.0$. オリジナルのポテンシャルエネルギーは $\lambda_4 = 1.0$ の場合に対応する. vWREM シミュレーションでは温度は全てのレプリカにおいて $T_0 = 300$ K に固定された. vWREM 構造サンプリングの効率を調べるために, 4 個のレプリカ (温度は 300, 315, 335, 360 K) の REMD シミュレーションも実行した. 更に, 温度が 300 K における従来のカノニカル MD シミュレーションも 4 回違う初期速度で実行された. 初期構造は全てのシミュレーションで同じであり, 完全に伸びた構造を採用した. MD シミュレーションの総時間ステップは vWREM と REMD においては, レプリカ当たり 2.5 ナノ秒であり, 1 回のカノニカルシミュレーションも同じ長さであった (つまり, 全てのシミュレーションの長さは同じ 10 ナノ秒であった). データは 50 フェムト秒毎に保存され, vWREM と REMD においては, レプリカ交換は 250 フェムト秒毎に試行された.

図 1.34 に主鎖の二面角 ϕ の時間発展を三つのシミュレーションについて示す. 図 1.34 から分かることは, ϕ 空間のサンプリングは vWREM が一番効率が良い, 次に REMD, 最後に従来のカノニカル MD となっていることである.

七つ目の例は, マルチオーバーラップ法 (MUOV) のシミュレーションである [41]. 系は再び真空中の Met-enkephalin である. 力場は CHARMM param 22[63] を採用し, CHARMM プログラム [67] を修正して MUOV を組み込んだ.

まず, 温度を 1000 K から 300 K まで下げる徐冷 MD シミュレーションで得られた二つのエネルギー極小状態を参照構造とした. 図 1.35 にこの二つの参照構造を示す. 式 (1.221) において $L = 2$ とおいて, 無次元化された自由エネルギー

112 第 1 章 拡張アンサンブルの方法

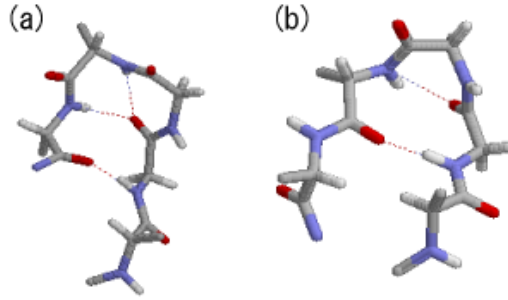


図 1.35 真空中の Met-enkephalin の (a) 参照構造 1 と (b) 参照構造 2. 主鎖のみを表した. 点線は水素結合である.

ギーを $f(d_1, d_2; T_0)$ で表した. MUOV MD シミュレーションは $T_0 = 300$ K で実行された. ここでも, MD の単位時間ステップは 0.5 フェムト秒とした.

図 1.36 にそれぞれの参照構造についての主鎖の二面角距離の時間発展を示す. $d_1 = 0$ のとき, 問題とする立体構造の二面角は参照構造 1 のそれと完全に一致し, 二つの主鎖構造が同じであることを意味する. ここで, 参照構造 2 の参照構造 1 に関する二面角距離は $d_1 = 0.122$ である. よって, 同様に, $d_2 = 0$ (参照構造 2 と同じ構造) のとき $d_1 = 0.122$ が成り立つ. そして, d_1 (d_2) がゼロに近いとき, 構造は参照構造 1 (参照構造 2) に似ている. 図 1.36 から分かることは, 従来のカノニカル MD シミュレーションが参照構造 2 の近傍に留まっているのに対し, MUOV MD シミュレーションは二面角距離の酔歩を実現して, 参照構造 1 と参照構造 2 の間を行き来していることである.

ここで, 温度 T における自由エネルギー (平均力ポテンシャル) $F(d_1, d_2; T)$ を以下で定義する.

$$F(d_1, d_2; T) = -k_B T \log P_B(d_1, d_2; T). \quad (1.259)$$

ここで, $P_B(d_1, d_2; T)$ は, 温度 T における, 再重法で得られた d_1 と d_2 の確率

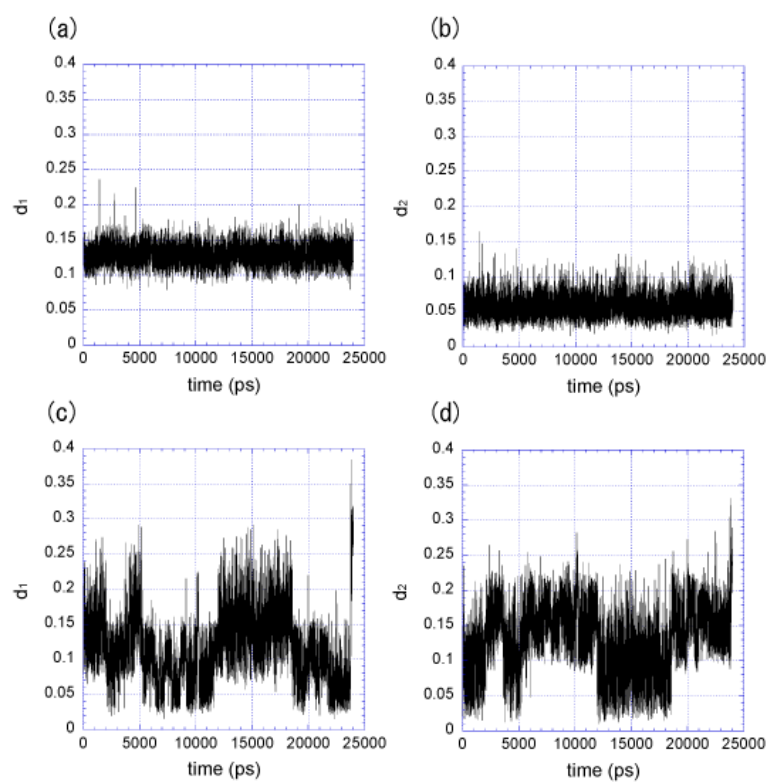


図 1.36 真空中の Met-enkephalin の主鎖の二面角距離 d_1 と d_2 の時間発展. (a) と (b) は温度 300 K における従来のカノニカル MD シミュレーションの結果, (c) と (d) は温度 300 K における MUOV MD シミュレーションの結果である.

114 第1章 拡張アンサンブルの方法

分布であり以下で与えられる (式. (1.225) 参照).

$$P_B(d_1, d_2; T) = \frac{\sum_E N_{muov}(d_1, d_2, E) e^{-(\beta - \beta_0)E - f(d_1, d_2; T_0)}}{\sum_{d_1, d_2, E} N_{muov}(d_1, d_2, E) e^{-(\beta - \beta_0)E - f(d_1, d_2; T_0)}} \cdot (1.260)$$

図 1.37 にこの自由エネルギー地形を示す. 図 1.37(a) では, 参照構造 2 の近傍の自由エネルギー極小状態だけが観察されているのに対し, 図 1.37(b) では, 両方の参照構造近傍の自由エネルギー極小状態が得られている. 極小状態 B と遷移状態 C の間の自由エネルギー差は約 5 kcal/mol, 極小状態 A と遷移状態 C の間の自由エネルギー差は約 2 kcal/mol であった. 温度 300 K における熱エネルギーは $k_B T \approx 0.6$ kcal/mol なので, これらのエネルギー障壁はかなり高いといえる. よって, 従来のカノニカル MD シミュレーションは最小自由エネルギー状態 B の近傍に留まってしまったわけである.

八つ目の例は, マルチバーリック・マルチサーマル法 (MUBATH) のシミュレーションである [48]. ここでは, レナード・ジョーンズポテンシャルで相互作用する N 個の同一粒子系を扱った. 系のポテンシャルエネルギーは以下で与えられる.

$$E = 4\epsilon \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (1.261)$$

ここで, ϵ はエネルギーの単位 (エネルギーの深さ) を表し, σ は距離の単位 (大雑把に言って, ファンデルワールス半径の 2 倍と考えると良い) を表す. (また, r_{ij} は粒子対 i と j の間の距離である.) 以下では, これらの単位でスケールされた単位系で表し, そのことを示すために物理量には星印 (*) を付けることにする. たとえば, 距離は $r^* = r/\sigma$, 温度は $T^* = k_B T/\epsilon$, 圧力は $P^* = P\sigma^3/\epsilon$ などと表される.

粒子数は 500 ($N = 500$) であり, それらを立方体の箱に入れ, 周期境界条件を課した. MUBATH 重み因子決定のために, 従来の定圧定温シミュレーションを $T_0^* = 2.0$, $P_0^* = 3.0$ で始めた. そして, 重み因子が求められた後に, 1 回 10^6 MD ステップの MUBATH シミュレーションを実行した.

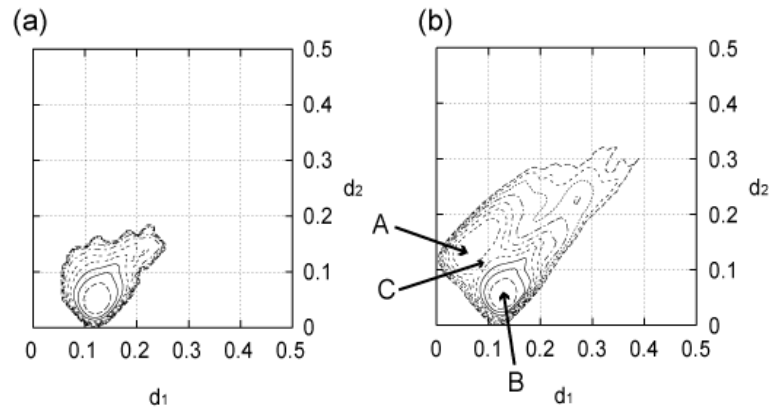


図 1.37 真空中の Met-enkephalin の温度 300 K における二面角距離 d_1 , d_2 についての平均力ポテンシャル. (a) は従来のカノニカル MD シミュレーション, (b) は MUOV MD シミュレーションの結果である. 等高線は 1 kcal/mol 毎に記された. ラベル A と B はそれぞれ参照構造 1 と参照構造 2 の近傍の自由エネルギー極小状態を示す. また, ラベル C はこの二つの自由エネルギー極小状態を結ぶ遷移状態 (鞍点) を示す.

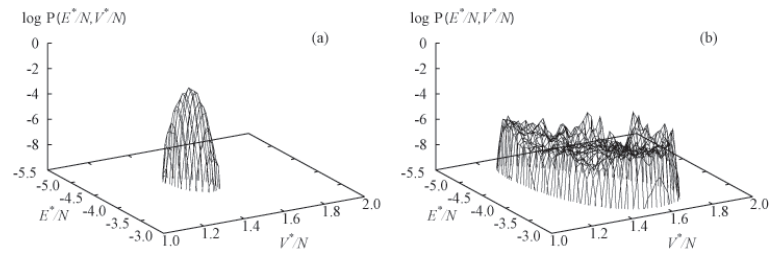


図 1.38 シミュレーションで得られたレナード・ジョーンズ粒子系のポテンシャルエネルギーと体積の確率分布. (a) は $(T_0^*, P_0^*) = (2.0, 3.0)$ における従来の定圧定温 MD シミュレーションの結果である. (b) は MUBATH MD シミュレーションの結果である.

116 第1章 拡張アンサンブルの方法

図 1.38(a) に最初に行われた従来の定圧定温シミュレーションで得られたエネルギーと体積の確率分布 $P_{NPT}(E^*/N, \mathcal{V}^*/N)$ を示す. この分布はベル型をしている. また, 図 1.38(b) に最後に実行された MUBATH シミュレーションで得られたエネルギーと体積の確率分布 $P_{MBT}(E^*/N, \mathcal{V}^*/N)$ を示す. 定圧定温シミュレーションに比べて広いエネルギーと体積の分布が得られていることが分かる.

二つの条件 $(T_0^*, P_0^*) = (1.6, 3.0)$, $(2.4, 3.0)$ における従来の定圧定温 MD シミュレーションにおけるポテンシャルエネルギー E^*/N の時間発展を図 1.39(a) に示した. 高温 ($T_0^* = 2.4$) では, ポテンシャルエネルギーは $E^*/N = -4.0 \sim -3.5$ の狭い範囲を揺らいでいる. また, 低温 ($T_0^* = 1.6$) では, ポテンシャルエネルギーは $E^*/N = -5.1 \sim -4.7$ の狭い範囲を揺らいでいる. 一方, MUBATH MD シミュレーションにおけるポテンシャルエネルギー E^*/N の時間発展を図 1.39(b) に示した. 幅広いエネルギー空間の酔歩が得られていることが分かる.

同様に, 二つの条件 $(T_0^*, P_0^*) = (2.0, 2.2)$, $(2.0, 3.8)$ における従来の定圧定温 MD シミュレーションにおける体積 $\mathcal{V}^*/N$ の時間発展を図 1.40(a) に示した. 高压 ($P_0^* = 3.8$) では, 体積は $\mathcal{V}^*/N = 1.3 \sim 1.4$ の狭い範囲を揺らぎ, 高压 ($P_0^* = 2.2$) では, 体積は $\mathcal{V}^*/N = 1.5 \sim 1.6$ の狭い範囲を揺らいでいるのに対し, 図 1.40(b) に示したように, MUBATH MD シミュレーションでは, 幅広い体積空間の酔歩が得られていることが分かる.

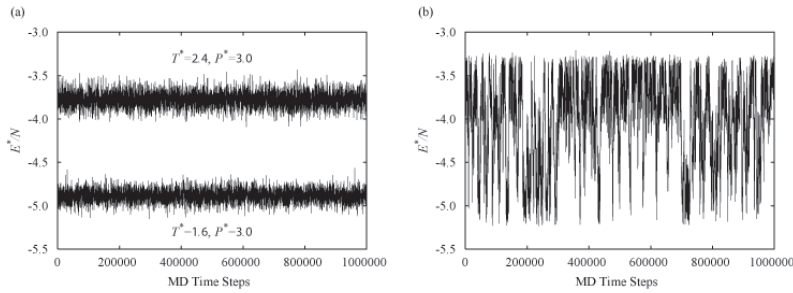


図 1.39 シミュレーションで得られたレナード・ジョーンズ粒子系のポテンシャルエネルギー E^*/N の時間発展. (a) は $(T_0^*, P_0^*) = (2.4, 3.0)$, $(T_0^*, P_0^*) = (1.6, 3.0)$ における従来の定圧定温 MD シミュレーションの結果である. (b) は MUBATH MD シミュレーションの結果である.

1.7 拡張アンサンブルシミュレーションの実例 117

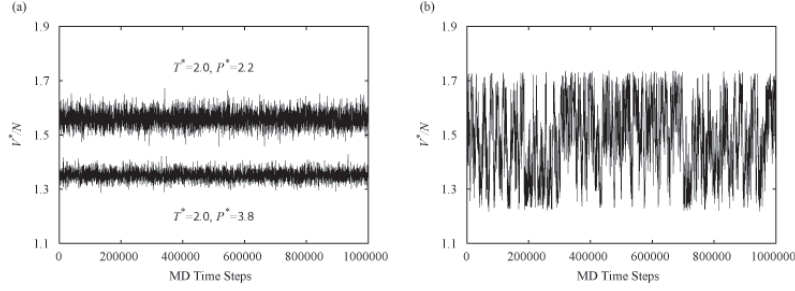


図 1.40 シミュレーションで得られたレナード・ジョーンズ粒子系の体積 V^*/N の時間発展.
 (a) は $(T_0^*, P_0^*) = (2.0, 2.2)$, $(T_0^*, P_0^*) = (2.0, 3.8)$ における従来の定圧定温 MD シミュレーションの結果である. (b) は MUBATH MD シミュレーションの結果である.

九つ目の例は、また、多次元レプリカ交換法と多次元焼き戻し法の MC シミュレーションである [21]. 系は再び 17 残基のリボヌクレアーゼ T_1 のペプチドフラグメントである [52, 31]. ここでは、二次元レプリカ交換シミュレーションと二次元焼き戻しシミュレーションを実行した. ポテンシャルエネルギーは以下のものを使った.

$$E_\lambda = E_C + \lambda E_{\text{SOL}}. \quad (1.262)$$

これは、一般の定式化の式 (1.150) において $L = 1$, $E_0 = E_C$, $V_1 = E_{\text{SOL}}$, $\lambda^{(1)} = \lambda$ とおいたことになる. ここで、 E_C は溶質であるペプチドのポテンシャルエネルギー、 E_{SOL} は溶媒和自由エネルギーである. 溶質のポテンシャルエネルギーとしては ECEPP/2[53] を使った. 溶媒項 E_{SOL} は (水素以外の) 重い原子の溶媒接触表面積に比例する項の和で与えられる [54]. この表面積の計算プログラムとしては NSOL[55] を使った.

モンテカルロプログラムとしては、KONF90 [7] を修正して、この拡張アンサンブル法を取り入れた. シミュレーションはランダムな初期構造から始められた. 温度は 8 個用意し ($M_0 = 8$), それらは、 $T_1 = 300$ K と $T_{M_0} = 700$ K の間に指数関数的に分布させられた. 値は以下のとおりである. 300.00, 338.60, 382.17, 431.36, 486.85, 549.49, 620.20, 700.00 K. また、パラメータ λ は 4 個 ($M_1 = 4$) 等間隔に以下のように分布させた. $\lambda_1 = 0$, $\lambda_2 = 1/3$, $\lambda_3 = 2/3$, λ_4

118 第1章 拡張アンサンブルの方法

= 1. これらが二次元レプリカ交換法と二次元焼き戻し法で使われた．ここで， $\lambda = 0$ （よって， $E_\lambda = E_C$ ）は真空中のシミュレーションであり， $\lambda = 1$ （よって， $E_\lambda = E_C + E_{SOL}$ ）は水中のシミュレーションに対応する．すなわち，この拡張アンサンブルシミュレーションでは，真空中のシミュレーションと水中のシミュレーションを同時に実行するわけである．

まず，二次元レプリカ交換 MC シミュレーションの結果を述べる．値が上のように分布する 8 個の温度と 4 個の λ パラメータに対応する 32 個のレプリカにおいて，各レプリカについて 10^6 MC sweeps の二次元レプリカ交換シミュレーションを実行した．各 20 MC sweeps 毎に T 交換または λ 交換が試行された．

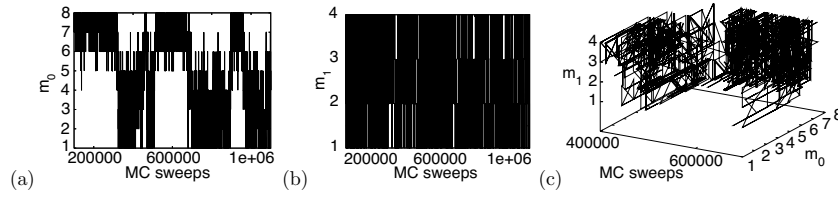


図 1.41 リボヌクレアーゼ T_1 のペプチドフラグメントの二次元レプリカ交換 MC シミュレーションにおける一つのレプリカについての温度 T_{m_0} とパラメータ λ_{m_1} のラベル (m_0 と m_1) の「時間発展」．(a) は m_0 ，(b) は m_1 の結果である．(c) は両者を二次元の「時間発展」として表したものである．

図 1.41 にこのシミュレーション中の一つのレプリカにおける，温度 T_{m_0} とパラメータ λ_{m_1} のラベル (m_0 と m_1) の時間発展を示す．このレプリカでは，温度とパラメータの二次元酔歩が実現されている．他のレプリカの結果も同様であった．

図 1.42 に図 1.41 と同じレプリカにおける，温度 T ，全ポテンシャルエネルギー E_{TOT} ，ペプチドの構造エネルギー E_C ，溶媒和自由エネルギー E_{SOL} ，末端間距離 D の同様の時間発展を示す．図 1.42(a) と図 1.42(b) から分かることは，エネルギーが低いときは，末端間距離が α ヘリックス構造に対応する約 8 Å の値をとり，エネルギーが高いときは，末端間距離が 7 Å～14 Å の間を揺らいでいることである．よって，このペプチドは低温では α ヘリックス構造，高温ではランダムコイル状態にあることが分かる．

1.7 拡張アンサンブルシミュレーションの実例 119

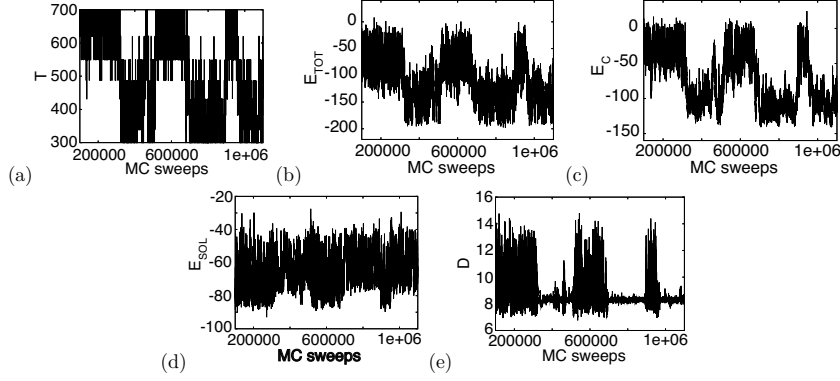


図 1.42 リボヌクレアーゼ T_1 のペプチドフラグメントの二次元レプリカ交換 MC シミュレーションにおける一つのレプリカ (図 1.41 のものと同じ) についての (a) 温度 T , (b) 全ポテンシャルエネルギー E_{TOT} , (c) ペプチドの構造エネルギー E_C , (d) 溶媒和自由エネルギー E_{SOL} , (e) 末端間距離 D の「時間発展」、パラメータ λ_{m_1} のラベル (m_0 と m_1) の「時間発展」. 温度の単位は K, エネルギーの単位は kcal/mol, 末端間距離の単位は Å である.

この二次元レプリカ交換 MC シミュレーションの結果に多ヒストグラム再重法を適用して、二次元焼き戻し法の重み因子を決定することができる. すなわち, 式 (1.176) と式 (1.177) の一般化された WHAM 方程式を解くことによって, 32 個の二次元焼き戻し法のパラメータ f_{m_0, m_1} ($m_0 = 1, \dots, 8; m_1 = 1, \dots, 4$) を求めた.

二次元焼き戻し重み因子 $W_{ST} = \exp(-\beta_{m_0}(E_C + \lambda_{m_1}E_{SOL}) + f_{m_0, m_1})$ が求められたので, それを使って, 10^6 MC sweeps の二次元焼き戻し MC シミュレーションを実行した. 各 20 MC sweeps 毎に T_{m_0} から $T_{m_0 \pm 1}$ への温度更新または λ_{m_1} から $\lambda_{m_1 \pm 1}$ へのパラメータ更新が試行された.

この二次元焼き戻しシミュレーションの結果から得られた, 全ポテンシャルエネルギー E_{TOT} の平均, 構造エネルギー E_C の平均, $\lambda \times E_{SOL}$ の平均, 末端間距離 D の平均を図 1.43 に示す. これらの結果は, 二次元レプリカ交換シミュレーションの結果と良い一致を示している.

ここでは, 二次元焼き戻しシミュレーションと二次元レプリカ交換シミュレーションの結果に良い一致を見た. しかし, 二次元レプリカ交換法では 32 個のレ

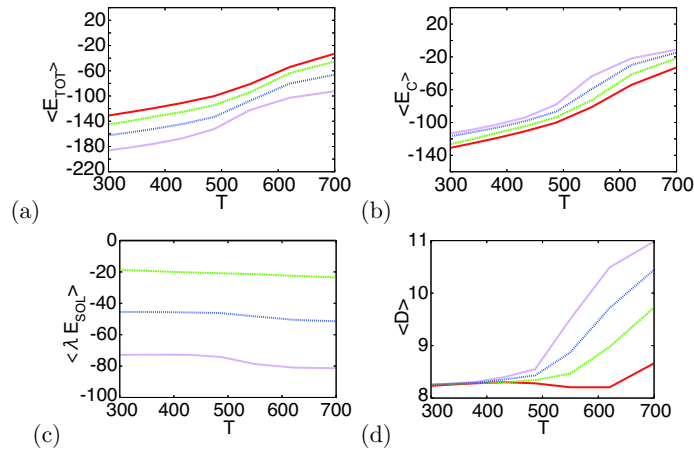


図 1.43 リボヌクレアーゼ T_1 のペプチドフラグメントの二次元焼き戻し MC シミュレーションの結果から得られた (a) 全ポテンシャルエネルギー E_{TOT} の平均, (b) 構造エネルギー E_C の平均, (c) $\lambda \times E_{SOL}$ の平均, (d) 末端間距離 D の平均. λ の値によって別々に計算した. それぞれの曲線の色は赤, 緑, 青, 紫がそれぞれ, λ_1 , λ_2 , λ_3 , λ_4 に対応している. それらは, (a) と (c) では, 上から下に向かって, また, (b) と (d) では, 下から上に向かって, λ_1 , λ_2 , λ_3 , λ_4 に対応している.

1.7 拡張アンサンブルシミュレーションの実例 121

プリカを必要としたのに対し、二次元焼き戻し法では一つのレプリカのみの必要であったことに注目されたい。

十番目の例は、定圧定温アンサンブルにおける圧力に関する焼き戻し法 (pressure simulated tempering: PST) [22] の MD シミュレーションである [68]. このシミュレーションは一次元の圧力空間の酔歩を実現する. 系は水分子をあらわに取り入れた小蛋白質ユビキチン (ubiquitin) である. この系は, NMR 実験によって高圧において変性することが知られている [69]. ユビキチンは 76 個のアミノ酸残基から構成されている. ここでは, 1 個のユビキチンと 6232 個の水分子は立方体の箱に入れて, 周期境界条件を課した. 温度は 300 K に固定した. 1 バール (bar) から 10000 バールの間の圧力の値を 100 個用意した. 温度と圧力の制御はフーバー・ランジバン法 [70] を用いた. また, 静電相互作用については, パーティクル・メッシュ・エwald法 [71] を採用した. MD の単位時間ステップは 2.0 フェムト秒とした. 力場としては, 蛋白質では CMAP [62] を考慮した CHARMM22[63] を, また, 水分子では TIP3P[59, 63] を採用した. また, 計算プログラムとしては, NAMD バージョン 2.7b3[72] を用い, それを修正して PST を取り入れた.

焼き戻し重み因子の決定のために, まず, 従来の定圧定温 MD シミュレーションを一つの温度 $T = 300$ K と上の 100 個の圧力値において, 別々に独立に, それぞれ, 4 ナノ秒実行した. 得られたポテンシャルエネルギーと体積の分布のヒストグラム $N_{m_0, m_1}(E, V)$ を式 (1.253) と式 (1.254) の WHAM 方程式に代入して, 焼き戻しパラメータ f_{m_0, m_1} ($m_0 = 1, m_1 = 1, \dots, 100$) を求めた. これを使って 500 ナノ秒の PST シミュレーションを初期条件 (乱数の種) を変えて 10 回実行した (よって, 全データは 5 マイクロ秒の MD 計算に対応する).

図 1.44 にこの PST シミュレーションの圧力とポテンシャルエネルギーの時間発展を示す. 図 1.44 から分かるように, 圧力が 1 バールから 1 万バールの間を酔歩している. また, これに従って, ポテンシャルエネルギーの酔歩も得られており, この二つの酔歩には逆相関が存在する.

この PST シミュレーションの結果から, アミノ酸対に対応する C^α 原子対の間の距離 d の揺らぎ $\sqrt{\langle d^2 \rangle - \langle d \rangle^2}$ を計算した. 図 1.45 に結果を示す.

この結果はアミノ酸番号 7~10 と 20~40 の間のアミノ酸残基間の距離の揺

122 第 1 章 拡張アンサンブルの方法

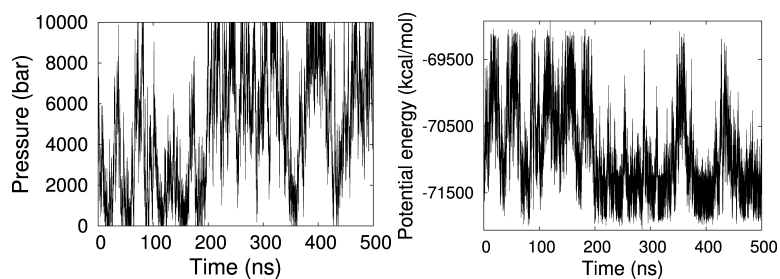


図 1.44 水中のユビキチンの圧力についての焼き戻し MD シミュレーションにおける圧力 (左) とポテンシャルエネルギー (右) の時間発展.

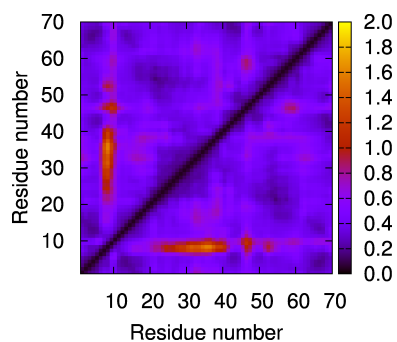


図 1.45 水中のユビキチンの圧力についての焼き戻し MD シミュレーションの結果から得られた C^α 原子対間の距離の揺らぎ.

らぎが大きいことを示唆しているが、それは高圧 NMR 実験の結果と一致している [69].

図 1.46 のスナップショットでは、この揺らいでいる距離は β ヘアピンのターンの部分と α ヘリックスの末端の間の距離に対応している。低圧ではこの距離は短いのに対し、高圧では長くなっており、間にできた空隙に水分子が入り込んで高圧変性を引き起こしていることが示唆される。

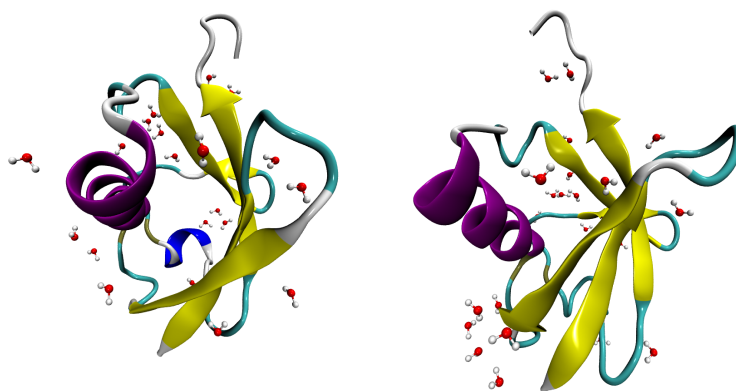


図 1.46 水中のユビキチンの圧力についての焼き戻し MD シミュレーション中のスナップショット。低圧（左）と高圧（右）の場合の 2 つを示した。紫が α ヘリックス構造、黄色が β シート構造を表している。蛋白質近傍の重要と思われる水分子も描いた。

1.8 おわりに

本章では、多自由度複雑系の計算機シミュレーションにおける有効な手法として、拡張アンサンブル法について解説した。計算例としては、主に古典力学に基づいた生体分子系のものを扱ったが、この手法の適用可能範囲は極めて広く、例えば、磁性体のスピングラスの系や量子系のシミュレーションでも有効である。拡張アンサンブル法は強力であるが、系が複雑になればなるほど、その適用が難しくなってしまう。良い手法というものは、強力であるばかりでなく、万人が容易に使えるものでなければならない。その意味において、本章で

124 第 1 章 拡張アンサンブルの方法

紹介した, レプリカ交換法とマルチカノニカル法または焼き戻し法を合体させた拡張アンサンブル法や温度ばかりでなく多次元 (多変数) に一般化した拡張アンサンブル法が, 特に有望であると考ええる.

謝辞

本章で紹介したいろいろな手法とシミュレーション結果は, 以下の人々との共同研究によって得られた: 川合光, 菊地武司, 中沢隆, 福来正孝, Ulrich Hansmann, 光武亜代理, Bernd Berg, 杉田有治, 北尾彰朗, 三上益弘, 西川武志, 野口博司, 依田隆夫. 奥村久士, 伊藤暁, 森義治. これらの各氏にこの場を借りて感謝したい. 更に, 榮慶丈 (蛋白質系の力場開発に関する共同研究およびその他計算機のハード・ソフトについての諸々の助言), 伊東真吾 (乱数発生プログラムの C 言語への翻訳), Walter Nadler (メトロポリス法の受諾率についての議論) の 3 氏にも感謝する.

参考文献

- [1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087–1092 (1953).
- [2] a) W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge University Press, New York, 1992).
b) G. Marsaglia, A. Zaman, and W.W. Tsang, *Stat. Prob. Lett.* **9**, 35–39 (1990).
- [3] a) G.O. Roberts, A. Gelman, and W.R. Gilks, *Ann. Appl. Probab.* **7**, 110–120 (1997).
b) M. Bédard, *Stochastic Process. Appl.* **118**, 2198–2222 (2008).
- [4] a) S. Nosé, *Mol. Phys.* **52**, 255–268 (1984).
b) S. Nosé, *J. Chem. Phys.* **81**, 511–519 (1984).
- [5] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, *Science* **220**, 671–680 (1983).
- [6] S. Geman and D. Geman, *IEEE Trans. Patt. Anal. Mach. Intel.* **6**, 721–741 (1984).
- [7] a) H. Kawai, Y. Okamoto, M. Fukugita, T. Nakazawa, and T. Kikuchi, *Chem. Lett.* **1991**, 213–216 (1991).
b) Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai, *Protein Eng.* **4**, 639–647 (1991).
- [8] a) 岡本祐幸, 日本物理学会誌 **51** 卷 4 月号, 279–287 (1996).
b) 杉田有治, 光武亜代理, 岡本祐幸, 日本物理学会誌 **56** 卷 8 月号, 591–599

126 参考文献

- (2001).
- c) 岡本祐幸, 化学フロンティア No. 8: 生体系のコンピュータ・シミュレーション: タンパク質の立体構造をどこまで予測できるか, (岡崎進・岡本祐幸編, 化学同人, 2002) pp. 25–42.
- d) 分子シミュレーション研究会会誌アンサンブル **12** 巻 4 月号, 特集「拡張アンサンブル法」(2010) pp. 6–35.
- [9] a) B.A. Berg and T. Neuhaus, *Phys. Lett.* **B267**, 249–253 (1991).
b) B.A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9–12 (1992).
- [10] J. Lee, *Phys. Rev. Lett.* **71**, 211–214 (1993); *ibid.* **71**, 2353 (1993).
- [11] a) U.H.E. Hansmann, Y. Okamoto, and F. Eisenmenger, *Chem. Phys. Lett.* **259**, 321–330 (1996).
b) N. Nakajima, H. Nakamura, and A. Kidera, *J. Phys. Chem. B* **101**, 817–824 (1997).
- [12] a) B.A. Berg, *Nucl. Phys. B (Proc. Suppl.)* **63A-C**, 982–984 (1998).
b) W. Janke, in *Computer Simulations of Surfaces and Interfaces*, NATO Science Series, II. Mathematics, Physics and Chemistry, vol. 114, Proceedings of the Advanced Study Institute, B. Dünweg, D.P. Landau, and A. Milchev (eds.) (Kluwer Academic, Dordrecht, 2003) pp. 137–157.
- [13] A. Mitsutake, Y. Sugita, and Y. Okamoto, *Biopolymers*, **60**, 96–123 (2001).
- [14] a) F. Wang and D.P. Landau, *Phys. Rev. Lett.* **86**, 2050–2053 (2001).
b) F. Wang and D.P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
- [15] A.M. Ferrenberg and R.H. Swendsen, *Phys. Rev. Lett.* **61**, 2635–2638 (1988); *ibid.* **63**, 1658 (1989).
- [16] a) B.A. Berg, *Comp. Phys. Commun.* **153**, 397–406 (2003).
b) B.A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore, 2004).
- [17] a) A.P. Lyubartsev, A.A. Martinovski, S.V. Shevkunov, and P.N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776–1783 (1992).
b) E. Marinari and G. Parisi, *Europhys. Lett.* **19**, 451–458 (1992).
- [18] A. Irbäck and F. Potthast, *J. Chem. Phys.* **103**, 10298–10305 (1995).
- [19] U.H.E. Hansmann and Y. Okamoto, *J. Comput. Chem.* **18**, 920–933 (1997).
- [20] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141–151 (1999).
- [21] a) A. Mitsutake and Y. Okamoto, *Phys. Rev. E* **79**, 047701 (2009).
b) A. Mitsutake and Y. Okamoto, *J. Chem. Phys.* **130**, 214105 (2009).
c) A. Mitsutake, *J. Chem. Phys.* **131**, 094105 (2009).

- [22] Y. Mori and Y. Okamoto, *J. Phys. Soc. Jpn.* **79**, 074003 (2010).
- [23] Y. Mori and Y. Okamoto, *J. Phys. Soc. Jpn.* **79**, 074001 (2010).
- [24] A.M. Ferrenberg and R.H. Swendsen, *Phys. Rev. Lett.* **63**, 1195–1198 (1989).
- [25] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, and J.M. Rosenberg, *J. Comput. Chem.* **13**, 1011–1021 (1992).
- [26] A. Mitsutake, Y. Sugita, and Y. Okamoto, *J. Chem. Phys.* **118**, 6664–6675 (2003).
- [27] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604–1608 (1996).
- [28] E. Marinari, G. Parisi, J.J. Ruiz-Lorenzo, in *Spin Glasses and Random Fields*, A.P. Young (ed.) (World Scientific, Singapore, 1998) pp. 59–98.
- [29] a) H.G. Katzgraber, S. Trebst, D.A. Huse, and M. Troyer, *J. Stat. Mech.* **6**, P03018 (2006).
b) W. Nadler and U.H.E. Hansmann, *Phys. Rev. E* **75**, 026109 (2007).
- [30] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **329**, 261–270 (2000).
- [31] A. Mitsutake, Y. Sugita, and Y. Okamoto, *J. Chem. Phys.* **118**, 6676–6688 (2003).
- [32] A. Mitsutake and Y. Okamoto, *Chem. Phys. Lett.* **332**, 131–138 (2000).
- [33] A. Mitsutake and Y. Okamoto, *J. Chem. Phys.* **121**, 2491–2504 (2004).
- [34] Y. Sugita, A. Kitao, and Y. Okamoto, *J. Chem. Phys.* **113**, 6042–6051 (2000).
- [35] G.M. Torrie and J.P. Valleau, *J. Comput. Phys.* **23**, 187–199 (1977).
- [36] F. Fukunishi, O. Watanabe, and S. Takada, *J. Chem. Phys.* **116**, 9058–9067 (2002).
- [37] E.M. Boczko and C.L. Brooks, III, *J. Phys. Chem.* **97**, 4509–4513 (1993).
- [38] B. Roux, *Comp. Phys. Commun.* **91**, 275–282 (1995).
- [39] S.G. Itoh, H. Okumura, and Y. Okamoto, *J. Chem. Phys.* **132**, 134105 (2010).
- [40] B.A. Berg, H. Noguchi, and Y. Okamoto, *Phys. Rev. E* **68**, 036126 (2003).
- [41] a) S.G. Itoh and Y. Okamoto, *Chem. Phys. Lett.* **400**, 308–313 (2004).
b) S.G. Itoh and Y. Okamoto, *J. Chem. Phys.* **124**, 104103 (2006).
- [42] U.H.E. Hansmann, M. Masuya, and Y. Okamoto, *Proc. Natl. Acad. Sci. USA* **94**, 10652–10656 (1997).
- [43] S.G. Itoh and Y. Okamoto, *Phys. Rev. E* **76**, 026705 (2007).
- [44] I.R. McDonald, *Mol. Phys.* **23**, 41–58 (1972).
- [45] F.G. Andersen, *J. Chem. Phys.* **72**, 2384–2393 (1980).
- [46] a) T. Nishikawa, H. Ohtsuka, Y. Sugita, M. Mikami, and Y. Okamoto,

128 参考文献

- Prog. Theor. Phys. (Suppl.)* **138**, 270–271 (2000).
- b) T. Okabe, M. Kawata, Y. Okamoto, and M. Mikami, *Chem. Phys. Lett.* **335**, 435–439 (2001).
- c) Y. Sugita and Y. Okamoto, in *Lecture Notes in Computational Science and Engineering*, T. Schlick and H. H. Gan (eds.) (Springer-Verlag, Berlin, 2002) pp. 304–332; arXiv: cond-mat/0102296.
- [47] a) H. Okumura and Y. Okamoto, *Chem. Phys. Lett.* **383**, 391–396 (2004).
b) H. Okumura and Y. Okamoto, *Phys. Rev. E* **70**, 026702 (2004).
- [48] a) H. Okumura and Y. Okamoto, *Chem. Phys. Lett.* **391**, 248–253 (2004).
b) H. Okumura and Y. Okamoto, *J. Comput. Chem.* **27**, 379–395 (2006).
- [49] K. Morikami, T. Nakai, A. Kidera, M. Saito, and H. Nakamura, *Comput. Chem.* **16**, 243–248 (1992).
- [50] a) Y. Sugita and A. Kitao, *Proteins* **30**, 388–400 (1998).
b) A. Kitao, S. Hayward, and N. Gō, *Proteins* **33**, 496–517 (1998).
- [51] S.J. Weiner, P.A. Kollman, D.T. Nguyen, and D.A. Case, *J. Comput. Chem.* **7**, 230–252 (1986).
- [52] J.K. Myers, C.N. Pace, J.M. Scholtz, *Proc. Natl. Acad. Sci. USA*, **94**, 2833–2837 (1997).
- [53] a) F.A. Momany, R.F. McGuire, A.W. Burgess, and H.A. Scheraga, *J. Phys. Chem.* **79**, 2361–2381 (1975).
b) G. Némethy, M.S. Pottle, H.A. Scheraga, *J. Phys. Chem.* **87**, 1883–1887 (1983).
c) M.J. Sippl, G. Némethy, H.A. Scheraga, *J. Phys. Chem.* **88**, 6231–6233 (1984).
- [54] T. Ooi, M. Oobatake, G. Némethy, and H.A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3086–3090 (1987).
- [55] M. Masuya, unpublished; see <http://biocomputing.cc/nsol/>.
- [56] Y. Sugita and Y. Okamoto, *Biophys. J.* **88**, 3180–3190 (2005).
- [57] a) K.R. Shoemaker, P.S. Kim, E.J. York, J.M. Stewart, and R.L. Baldwin, *Nature* **326**, 563–567 (1987).
b) K.R. Shoemaker, R. Fairman, D.A. Schultz, A.D. Robertson, E.J. York, J.M. Stewart, and R.L. Baldwin, *Biopolymers* **29**, 1–11 (1990).
- [58] J. Wang, P. Cieplak, and P.A. Kollman, *J. Comput. Chem.* **21**, 1049–1074 (2000).
- [59] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, *J. Chem. Phys.* **79**, 926–935 (1983).
- [60] a) A. Kitao, F. Hirata, N. Gō, *Chem. Phys.*, **158**, 447–472 (1991).

- b) R. Abagyan and P. Argos, *J. Mol. Biol.* **225**, 519–532 (1992).
- c) A.E. Garcia, *Phys. Rev. Lett.* **68**, 2696–2699 (1992).
- d) A. Amadei, A.B.M. Linssen, and H.J.C. Berendsen, *Proteins* **17**, 412–425 (1993).
- [61] T. Yoda, Y. Sugita, and Y. Okamoto, *Biophys. J.* **99**, 1637–1644 (2010).
- [62] a) A.D. MacKerell, Jr., M. Feig, and C.L. Brooks, III, *J. Am. Chem. Soc.* **126**, 698–699 (2004).
 b) A.D. MacKerell, Jr., M. Feig, and C.L. Brooks, III, *J. Comput. Chem.* **25**, 1400–1415 (2004).
- [63] A.D. MacKerell, Jr., D. Bashford, M. Bellott, R.L. Dunbrack, Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- [64] P. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille, in *Computer Simulation of Biomolecular Systems Vol. 3*, W.F. van Gunsteren, P.K. Weiner, A.J. Wilkinson, Eds., (KLUWER/ESCOM, Dordrecht, 1997) pp. 83–96. .
- [65] a) S.D. Bond, B.J. Leimkuhler, and B.B. Laird, *J. Comput. Phys.* **151**, 114–134 (1999).
 b) S. Nosé, *J. Phys. Soc. Jpn.* **70**, 75–77 (2001).
- [66] H. Okumura, S.G. Itoh, and Y. Okamoto, *J. Chem. Phys.* **126**, 084103 (2007).
- [67] B.R. Brooks R.R. Brucoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187–217 (1983).
- [68] 森義治, 名古屋大学大学院理学研究科博士論文 (2011); Y. Mori and Y. Okamoto, in preparation.
- [69] a) R. Kitahara and K. Akasaka, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3167–3172 (2003).
 b) R. Kitahara, S. Yokoyama, and K. Akasaka, *J. Mol. Biol.* **347**, 277–285 (2005).
- [70] D. Quigley and M.I.J. Probert, *J. Chem. Phys.* **120**, 11432–11441 (2004).
- [71] a) T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089–10092 (1993).
 b) U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G.

130 参考文献

Pedersen, *J. Chem. Phys.* **103**, 8577–8593 (1995).

- [72] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten, *J. Comput. Chem.* **26**, 1781–1802 (2005).