

多自由度複雑系で有効な拡張アンサンブル法の開発

提案代表者	分子科学研究所	助教授	岡本 祐幸
提案者及び共同研究者	分子科学研究所	助手	杉田 有治（現東京大学講師）
	分子科学研究所	研究技官	長島 剛宏（現国立遺伝学研究所研究員）
	慶應義塾大学	助手	光武亜代理
	奈良女子大学	助教授	中沢 隆
	奈良女子大学	大学院生	弓長 浩子
	フロリダ州立大学	教授	Bernd A. BERG



本課題研究は、平成13年10月から平成15年3月まで実施された。以下に研究内容の概要及び主な成果について詳しく述べることにする。

1. 研究概要

生体高分子系などの多自由度の複雑系では、系にエネルギー極小状態が無数に存在するため、従来のカノニカルアンサンブル上の一定温度の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うという困難があった。拡張アンサンブル法 (generalized-ensemble algorithm) は従来一定温度のシミュレーションがボルツマン因子に基づいて状態を発生させるのに対し、非ボルツマン因子による人工のアンサンブルに基づいて、ポテンシャルエネルギー空間上の一次元酔歩を実現する手法である。よって、ポテンシャルエネルギーがいろいろな値を取っていくので、シミュレーションがどこかエネルギー極小状態に陥っても、いずれそこから脱出できる訳で、従来の手法よりも、はるかに幅広い配位空間を探索することができる。拡張アンサンブル法の利点は、唯一回のシミュレーションの結果から、最小エネルギー状態ばかりでなく、任意の温度における物理量の平均値が求められることである。マルチカノニカル法 (multicanonical algorithm: MUCA) や

焼き戻し法 (simulated tempering: ST) などが良く知られている拡張アンサンブル法であるが、これらの手法では、最初に (非ボルツマン的な) 重み因子を短い試行シミュレーションで決定する必要がある。この操作は系が複雑になると、なかなか難しくなり、大変な熟練が必要になる。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。我々は以上の見解のもとで、より適用が簡便な拡張アンサンブル法を求めてきた。そして、レプリカ交換法 (replica-exchange method: REM) という拡張アンサンブル法に到達した。この手法では、M個のオリジナルの系のコピー (レプリカ) とそれらを取るM個の温度を用意し、平行して、従来一定温度のカノニカルシミュレーションを走らせながら、途中で、レプリカ対 (温度対) を交換していくことによって、一つのレプリカを見れば、温度が酔歩し、エネルギーの酔歩も得られることになる。この方法では、従来一定温度のシミュレーションを平行して走らせるだけなので、重み因子の決定の必要がなく、非常に使い易い方法である。この手法はモンテカルロ法において開発されたが、タンパク質系などでは分子動力学法が広く使われているので、我々はまず、レプリカ交換法の分子動力学法版 (REMD) を開発し、¹⁾ 幅広い支持を得た。

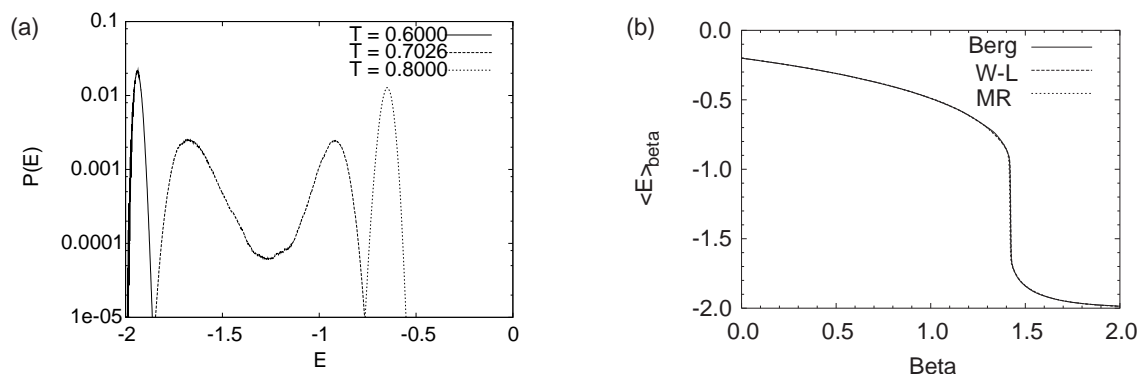


図1 二次元10状態Potts模型における3つの温度 ($T = 0.6000, 0.7026, 0.8000$) でのエネルギーの確率分布(a)と温度の逆数の関数としての平均エネルギー(b)。

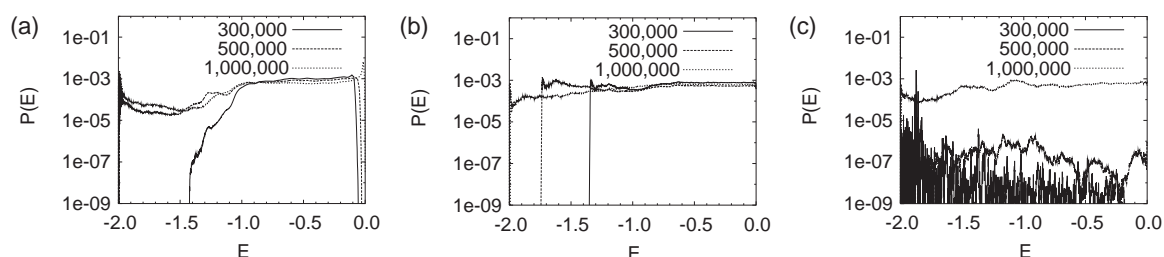


図2 二次元10状態Potts模型におけるマルチカノニカル重み因子の決定の様子。(a) MUCAREM、(b)Bergの方法、(c)Wang-Landauの方法。

しかし、レプリカ交換法にも一つだけ問題点があった。それは、系が複雑になると必要となるレプリカの数が多くなって、一個の系(レプリカ)だけを使うマルチカノニカル法や焼き戻し法に比べて、計算時間が大幅に増えてしまうことである。この問題も最近、次のように解決された。すなわち、まず、短いレプリカ交換シミュレーションによって、マルチカノニカル法や焼き戻し法の重み因子を決定して、その後、長い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を実行するのである。前者はレプリカ交換マルチカノニカル法(replica-exchange multicanonical algorithm: REMUCA)²⁾ 後者はレプリカ交換焼き戻し法(replica-exchange simulated tempering: REST)³⁾と名付けられた。更には、カバーするエネルギー空間の領域を制限したマルチカノニカルアンサンブルをレプリカの個数だけ用意し、途中でレプリカ交換を実行する新手法であ

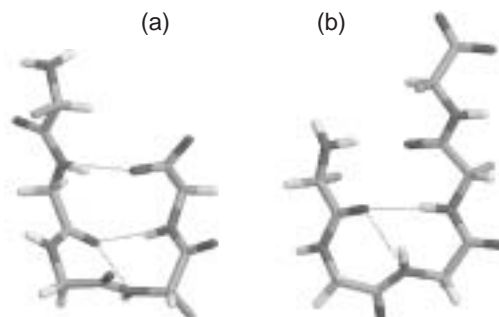
る、マルチカノニカルレプリカ交換法(multicanonical replica-exchange method: MUCAREM)も開発した。²⁾ 本課題研究では、これらの新手法の有効性をスピン系やタンパク質系のような多自由度複雑系で確認するとともに、より有効な拡張アンサンブル法の新たな開発を目指した。

2. 主な研究成果

2.1 Potts模型へのMUCAREMの適用

スピン系であるPotts模型を対象として、我々によって開発されたマルチカノニカルレプリカ交換法(MUCAREM)²⁾ Bergにより開発された手法、更には、Wang-Landauの手法の有効性の詳しい比較を行った。^{4),5)} 二次元 q 状態Potts模型は、 q が4以下の場合、2次相転移、4より大きい場合、1次相転移を引き起こすことが解析的に示されている。ここでは、格子サイズが 34×34 の二次元10状態Potts模型の

図3 構造1(a)と構造2(b)。



拡張アンサンブルモンテカルロシミュレーションを実行した。よって、この系は1次相転移を示す。1次相転移を持つ系では、従来のレプリカ交換法は使えないので、マルチカノニカル法を適用することにする。マルチカノニカル法では、非ボルツマン因子(マルチカノニカル重み因子)によって状態が発生し、エネルギーの一樣分布が得られる。問題は、短い試行シミュレーションのイテレーションによって、まず、マルチカノニカル重み因子を決定する必要があることである。この重み因子の決定法としては、BERGの方法、Wang-Landauの方法、及びMUCAREMが特に有望である。よって、これらの3手法の有効性を比較した。そして、十分精度の良いマルチカノニカル重み因子が求められたら、長いプロダクションランを一回だけ実行する。このプロダクションランの結果に、ヒストグラム再重法によって、任意の温度におけるカノニカル分布を再現でき、物理量の平均値が得られる訳である。図1(a)に、ヒストグラム再重法で得られた、3つの温度(相転移温度より低温、相転移温度、相転移温度より高温)における、エネルギーの確率分布を示す。相転移温度では、分布が2つに分かれて1次相転移の傾向が現れている。図1(b)には、温度の逆数の関数としての平均エネルギーを示す。相転移温度($T = 0.7026$, or $\beta = 1.423$)で潜熱によるギャップが現れており、強い1次相転移の傾向が顕著に出ている。

図1はマルチカノニカルモンテカルロシミュレーションの結果であるが、上に述べたように、マルチカノニカル重み因子は3つの方法で求めた。(エネルギーの一樣分布を与える)精度の良いマルチカノ

ニカル重み因子へ収束していく様子を図2に示した。それぞれ、途中の振る舞いは大きく違うが、3手法とも100万回のモンテカルロステップを経ると全エネルギー領域で一樣な分布が得られていることが分かる。よって、Potts模型においては、3手法は大体同等の有効性を持つと言える。

2.2 マルチオーバーラップ法の開発

マルチカノニカル法はエネルギーの一樣分布を与える手法であるが、今回、ある状態とのオーバーラップ関数について一樣分布を与える手法を新たに開発して、マルチオーバーラップ法と名付けた。⁶⁾ この手法を適用する系の例として、真空中のMet-enkephalinを考えた。このペプチドのアミノ酸数は5個であり、配列は、Tyr-Gly-Gly-Phe-Metである。この系は以前、マルチカノニカル法で詳しく調べられた。⁷⁾そして、最小エネルギー構造はGly-2とMet-5が水素結合している β ヘアピン構造であり、次にエネルギーが低い極小状態はTyr-1とPhe-4が水素結合している β ヘアピン構造であることが分かっている。⁷⁾これらをそれぞれ構造1と構造2と呼ぶこととして、図3にその主鎖構造を示す。

ここで、問題とする構造とある参照構造とのオーバーラップ関数を対応する全ての二面角の差の二乗和で定義される「距離」 d の関数として定義する。ここで、2つの構造が完全に一致すると、 $d = 0$ である。すると、マルチオーバーラップ法では、 d の一樣分布が得られる。例えば、構造1とのオーバーラップ(または、距離 d_1)について新手法を適用す



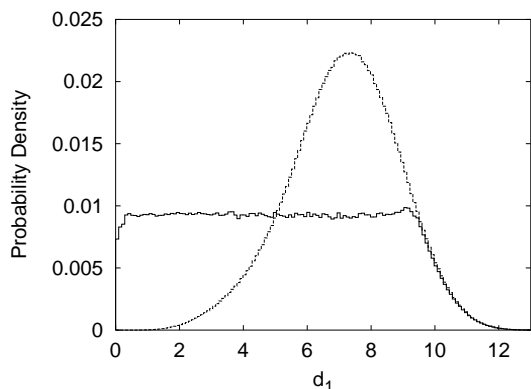


図4 構造1を参照構造とする場合の二面角の距離 d_1 の温度 $T = 400$ Kにおける確率分布。カノニカル分布（ベル型）とマルチオーバーラップ分布（平坦）を示す。

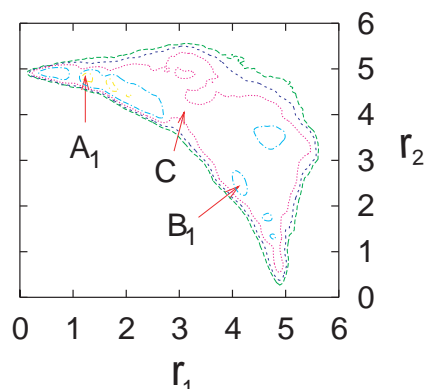
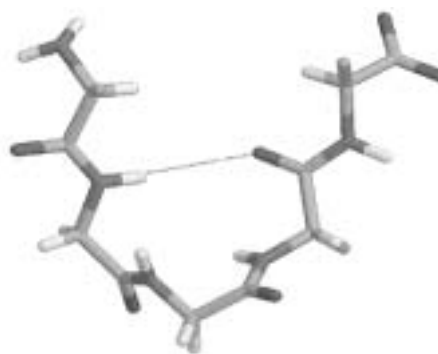


図5 折り畳み温度 ($T = 250$ K)における真空中のMet-enkephalinの自由エネルギー地形。 r_1 、 r_2 はそれぞれ構造1、構造2からの根二乗平均距離である。A₁、B₁はそれぞれ、この温度における構造1、構造2に対応する自由エネルギー極小状態を表す。また、鞍点Cは遷移状態を表す。

図6 構造1と構造2の遷移状態（構造3）。

ると、 d_1 の一樣分布が得られる。図4に $T = 400$ Kにおける d_1 のカノニカル分布（ベル型）とマルチオーバーラップ分布（平坦）を示す。 d_1 の一樣分布が得られたということは、そのシミュレーションは構造1 ($d_1 = 0$)とランダムコイルの間を行ったり来たりしていることを意味する。同様に、構造2との距離 d_2 についてのマルチオーバーラップシミュレーションも実行することができる。我々は更に、構造1と構造2の間を行ったり来たりするようなマルチオーバーラップシミュレーションにも成功した。⁶⁾すると、このペプチドの折り畳みに関する構造1と構造2の間の遷移状態を詳しく調べることができることになる。このシミュレーションの結果から、我々はこのペプチド系の折り畳み温度 ($T = 250$ K)における自由エネルギー地形（平均力ポテンシャル）を構造1からの根二乗平均距離 r_1 と構造2からの根二乗平均距離 r_2 との関数として求めた（図5）、図5における自由エネルギーの鞍点Cが遷移状態を表す。遷移状態を詳しく調べると、図6の



ような立体構造（構造3）を持つことが判明した。この構造はGly-2とPhe-4が水素結合している。自由エネルギー地形ばかりでなく、内部エネルギー地形及びエントロピー地形もプロットすることができる訳であり、それらを図7に示した。内部エネルギーは構造1 ($d_1 = 0$)及び構造2 ($d_2 = 0$)のところで、それぞれ最小及び極小になっているが、有限温度のエントロピー効果のために、自由エネルギーの最小状態及び極小状態は、それぞれ、図5のA₁、B₁に位置がずれている訳である。また、内部エネルギーとエントロピーはそれぞれでばこが多いが、それらが相殺して、自由エネルギーは滑らかとなっていることが分かる。

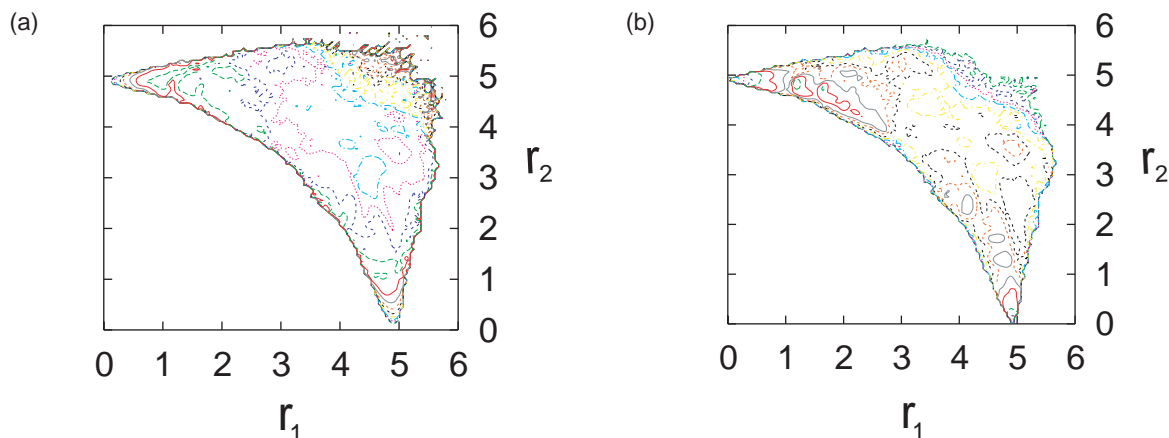


図7 折り畳み温度 ($T = 250$ K) における真空中の Met-enkephalin の内部エネルギー地形(a)とエントロピー地形(b)。

3. おわりに

本課題研究に関する予算により、マルチカノニカル法の元祖である、フロリダ州立大学の Berg 教授を招へいすることができ、研究が大いに進んだ。ここに感謝の意を表したい。

参考文献

- 1) Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- 2) Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **329**, 261 (2000).
- 3) A. Mitsutake and Y. Okamoto, *Chem. Phys. Lett.* **332**, 131 (2000).
- 4) T. Nagasima, Y. Sugita, A. Mitsutake, and Y. Okamoto, *Comput. Phys. Commun.* **146**, 69 (2002).
- 5) Y. Okamoto, in Proceedings of the Los Alamos Workshop, *The Monte Carlo Method in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm*, J. E. Gubernatis, Ed., American Institute of Physics; Melville, pp.248–260 (2003).
- 6) B. A. Berg, H. Noguchi and Y. Okamoto, *Phys. Rev. E* **68**, 036126 (2003).

- 7) A. Mitsutake, U. H. E. Hansmann and Y. Okamoto, *J. Mol. Graphics Modell.* **16**, 226 (1998).