

# 福岡工業大学 学術機関リポジトリ

## スケール変換したエントロピーサンプリング法によるペプチド分子の構造予測

メタデータ	言語: Japanese 出版者: 公開日: 2021-02-08 キーワード (Ja): キーワード (En): Scale Transformation, Entropic Sampling, Global Optimization, Met-enkephalin 作成者: 中村, 秀昭 メールアドレス: 所属:
URL	<a href="http://hdl.handle.net/11478/00001616">http://hdl.handle.net/11478/00001616</a>

# スケール変換したエントロピーサンプリング法による ペプチド分子の構造予測

中 村 秀 昭 (機能材料工学科)

## Prediction of peptide conformation using a scale-transformed entropic sampling algorithm

Hideaki NAKAMURA (Department of Functional Materials Engineering)

### Abstract

A scale-transformed entropy sampling algorithm for searching a conformational energy space of polypeptides is presented. This algorithm is based on the assumption that high energy barriers can be overcome by a uniform sampling of the logarithmically transformed energy space. This algorithm is tested with Met-enkephalin, a peptide with a known global minimum structure. Using a new Monte Carlo sampling method, energy barriers of 3000 kcal/mol are frequently overcome and, therefore, many minima are searched more efficiently than with the usual entropy sampling simulations. This method can be utilized as the optimization algorithm. The global minimum conformations obtained in optimizations agree well with those obtained by the simulated annealing (SA) method and by the systematic stepsize variation (SSV) method. The entropy, the average energy and the specific heat are estimated with one Monte Carlo run.

Key words: *Scale Transformation, Entropic Sampling, Global Optimization, Met-enkephalin*

### 1. はじめに

蛋白質やペプチド分子のエネルギー最小構造を求めるとき、膨大な数のエネルギーミニマムとその間を遮っている高いエネルギー障壁が問題の解決を困難にしてきた。最近、高いエネルギー障壁に打ち勝つための有効なモンテカルロ法、すなわち、マルチカノニカル法<sup>1), 2), 3)</sup>とエントロピーサンプリング法<sup>4), 5), 6)</sup>が開発された。両者とも、すべてのエネルギー状態の確率が一定になるように遷移確率に重みを付けてサンプリングするので、どのように高いエネルギー障壁も原理的

には越えられる。しかしながら、計算時間の制約もあって大きなエネルギーの障壁の存在する系でエネルギー最小構造を求めることはまだそれほど簡単ではない。最近、我々は大きなエネルギーの障壁に打ち勝つ有効なモンテカルロ法を提案した<sup>7)</sup>。この方法では、エネルギースケールを対数的に縮小したエネルギー空間が導入され、系の統計性が失われないように状態がエントロピー的にサンプリングされている。この方法の効率性は、ペプチド結合のまわりの回転を固定したエンケファリン分子に対してテストされた。この論文では、エンケファリン分子のより正確な生物活性構造を求めるために、ペプチド結合のまわりの回転を許した計算を行った。この計算では、回転自由度が増えるだけでなく、よりエネルギー障壁の高いエネルギー一面を探索

平成13年10月31日受付

しなければならないので、新しいモンテカルロアルゴリズムの有効性を試すのに都合がよい。

## 2. 計算方法

### エントロピーサンプリング法

カノニカル分布では、エネルギー  $E$  をとる確率  $P_C(E)$  は

$$P_C(E) \propto N(E)e^{-\beta E} \propto e^{(S(E)/k - \beta E)} \quad (1)$$

で与えられる。ここで、温度  $T$  として  $\beta = 1/kT$  である。 $N(E)$  はエネルギー  $E$  をとるコンフォメーションの数で、 $S(E) = k \ln N(E)$  とおいた。 $S(E)$  はエントロピーを与える。

分配関数は、

$$Z(\beta) = \int e^{(S(E)/k - \beta E)} dE \quad (2)$$

と書くことができる。通常の方法では、ペプチド分子がコンフォメーション  $x$  からコンフォメーション  $x'$  へ遷移するとき、確率、

$$P_T(x \rightarrow x') = \min[1, e^{-(\beta E(x') - \beta E(x))}] \quad (3)$$

で受け入れてやると状態のカノニカル分布が得られる。

いま、温度無限大 ( $\beta = 0$ ) の場合を考える。(2)より、

$$Z(\beta = 0) = \int e^{S(E)/k} dE \quad (4)$$

となり、温度無限大の分配関数はエントロピー  $S(E)$  だけで決定される。エントロピーサンプリングモンテカルロシミュレーション (ESMC) 法は  $S(E)$  を求めるための計算方法である。すなわち、(1)式でボルツマン因子  $e^{-\beta E}$  を  $e^{-J(E)}$  で置き換えた任意の確率分布

$$P_{ES}(E) \propto N(E)e^{-J(E)} = e^{(S(E)/k - J(E))} \quad (5)$$

を導入し、 $P_{ES}(E) = \text{const.}$  になるように  $J(E)$  を決めてやることにより  $S(E) = J(E)$  を求めることができる。 $P_{ES}(E) = \text{const.}$  を実現するためには、(3)式との類推から

$$P_T(x \rightarrow x') = \min[1, e^{-\{J(E(x')) - J(E(x))\}}] \quad (6)$$

の遷移確率のもとで MC を行い、MC の過程でヒストグラム  $H(E)$  を作成し、 $J(E)$  を

$$\begin{aligned} J_{\text{new}}(E) &= J_{\text{old}}(E) + \ln H(E) & \text{if } H(E) > 0 \\ &= J_{\text{old}}(E) & \text{if } H(E) = 0 \end{aligned} \quad (7)$$

のように更新していけばよい。ESMC 法を用いると、どのエネルギー状態も出現する確率が同じになるので、エネルギーの高い状態も容易に乗り越え、エネルギーの低い状態を効率的にサンプリングすることができる。

### 最適化アルゴリズム

一般に、ポリペプチド分子が長くなるとエネルギー極小の数が急速に増大するだけでなく、それらの間に存在するエネルギー障壁の高さも著しく増大する。従ってこのような系の MC シミュレーションでは、ほとんどの時間が高いエネルギー状態のサンプリングに費やされ、エネルギー障壁を乗り越えてエネルギーミニマムを見つけることが困難である。このような問題に対して、ESMC 法では通常のエネルギー空間を一樣にサンプリングして、等間隔な幅  $\Delta E$  でヒストグラム  $H(E)$  を作成する。しかしながら、この方法でも 1000 kcal/mol を超えるような大きなエネルギー障壁がある系ではエネルギーの低い状態を効率的にサンプリングしているとは言えない。そこで、エネルギー空間  $E$  (以後  $E$  空間と呼ぶ) を

$$E'(x) = \frac{1}{\alpha} \ln(N(x) - E_0) \quad (8)$$

によって対数的にスケール変換したエネルギー空間  $E'$  (以後  $E'$  空間と呼ぶ) を導入する。ここで、 $\alpha$  と  $E_0$  はそれぞれ適当な定数で  $E_0 < E$  を満たす。 $E'$  空間では、エネルギーの高さのみが対数的に縮小されるだけでエネルギー極小のコンフォメーションは保持される。

いま  $E'$  空間で、(5)式と同様なスケール変換された確率分布

$$P_M(E') \propto e^{(S'(E')/k - J'(E'))} \quad (9)$$

を導入し、この分布が一定になるように  $J'(E')$  を(6)、(7)と同じ手続きの ESMC 法を適用することを考える。ただし、 $S'(E')$  は  $S(E)$  を  $E'$  空間に射影したものである。このとき、遷移確率

$$P_T(x \rightarrow x') = \min[1, e^{-\{J'(E'(x')) - J'(E'(x))\}}] \quad (10)$$

のもとで MC を行い、MC の過程でヒストグラム  $H'(E')$  を作成し、 $J'(E')$  を

$$\begin{aligned} J'_{\text{new}}(E') &= J'_{\text{old}}(E') + \ln H'(E') & \text{if } H'(E') > 0 \\ &= J'_{\text{old}}(E') & \text{if } H'(E') = 0 \end{aligned} \quad (11)$$

のように更新していけばよい。この方法では、エネ

ギー障壁が対数的に低められ、その結果、従来のESMC法より効率よくエネルギーの低い状態をサンプリングすることができる。しかしながら、この方法は高いエネルギー領域において正しいエントロピー $S(E)$ を与えないので、熱力学的量を計算することに適さない。むしろ、最小エネルギーを探索する最適化アルゴリズムとして利用できることに注目すべきである。この最適化アルゴリズムの詳細については論文(7)に記述されている。ここでは、この最適化アルゴリズムを用いてエンケファリンのエネルギー最小構造と $E_0$ を計算する。

### スケール変換されたエントロピーサンプリング(STES)法

熱力学性を保証するためには、ボルツマン因子 $e^{-\beta E}$ が寄与しない無限大の温度( $\beta=0$ )で(4)式と同じ分配関数になるように変換すればよい。すなわち、新しいエネルギー変数を用いて、

$$Z(\beta=0) = \int e^{S'(E')/\alpha} dE' \quad (12)$$

と書ける。正しい $S'(E')$ を決定するためには、以下の様な確率分布

$$P_{sr}(E') \propto e^{(S'(E')/\alpha - J'(E'))} \quad (13)$$

を導入して、 $E'$ 空間でESMC法を実行してやればよい。(5)式との類推から遷移確率を

$$P_T(x \rightarrow x') = \min[1, e^{-\{J'(x') - \alpha E'(x') - J'(x) + \alpha E(x)\}}] \quad (14)$$

として与え、 $E'$ 空間でのMCを行い、得られたヒストグラム $H'(E')$ より、次の関係

$$J'_{\text{new}}(E') = J'_{\text{old}}(E') + \ln H'(E') \quad \text{if } \ln H'(E') > 0 \quad (15) \\ = J'_{\text{old}}(E') \quad \text{if } \ln H'(E') = 0$$

によって $J'(E')$ を更新していくと、(13)式の $P_{sr}(E')$ はしだいに平坦な分布になっていく。最終的には、 $S'(E') = J'(E') + \alpha E'$ によって正しいエントロピーを求めることができる。この方法をスケール変換されたエントロピーサンプリング(STES)法と呼ぶ。STES法の特徴は、スケール変換によってエネルギーの低い状態がエネルギーの高い状態より相対的に多くサンプリングされることである。一方で、エネルギーの高い障壁を乗り越える機会は十分与えられており、局所的なミニマムに留まる危険性を少なくしている。

### 計算の詳細

エンケファリンはアミノ酸配列 H-Tyr-Gly-Gly-Phe-Met-OH を持っている。主鎖のN末端側は $\text{NH}_2$ 基で、C末端側は $-\text{COOH}$ 基で終わる。このペプチド分子のコンフォメーションエネルギーをECEPP/2<sup>(8), (9)</sup>を用いて計算する。コンフォメーションを与える独立変数(2面角)の数は24個である。

エンケファリンのエネルギー最小構造と最適な $E_0$ を求めるために、最適化アルゴリズムを用いたモンテカルロ(MC)計算が $E_0=1000 \text{ kcal/mol}$  および $J'(E')=0$ から出発して行われた。ただし、(8)式の変換では、 $\alpha=1$ が仮定されている。MC計算が1000MCステップ終了するごとに0.1kcal/molのエネルギー幅でヒストグラムが作られ、 $J'(E')$ が(11)式によってアップデートされた。このアップデートはさらに1000回繰り返された。

STES法によるモンテカルロシミュレーションは、2段階に分けて実行された。第1段階では、50,000MCステップの短いMC計算を行うことによって $J'(E')$ のアップデートがさらに20回程度繰り返された。この計算で十分に低いエネルギー状態が得られ、かつ(13)式の分布が一樣になることを確認する。第2段階では、1,000,000MCステップの長いMC計算が第1段階で得られた $J'(E')$ を用いて実行された。この計算は、再び(13)式の分布が一樣になるまで繰り返され、エントロピー $S(E)$ が最終的に求められた。

## 3. 結果と考察

### 最適化

最適化アルゴリズムを用いて、エンケファリンの最小エネルギー構造と最適な $E_0$ の値を求める計算が2回実行された。図1は、1回目の最適化計算におけるエネルギー変化の軌跡を示したものである。-10kcal/mol前後の低エネルギー状態と数1000kcal/molの高エネルギー領域の間を何度も往復しながらエネルギーの低い状態を探索している様子がわかる。この計算では、約450,000MCステップでエネルギー最小構造(表1のOP1)を見出すことができた。2回目の最適化計算で得られたエネルギー最小構造は表1でOP2として与えられている。両構造の違いは、メチオニン側鎖の先端部分の向きが異なるだけである。これら2つの構造は、別の方法によって得られたエネルギー最小構造(表中でLSとSSVとして記載されている構造)

TABLE 1.  
Lowest-Energy Conformations Obtained by Optimization Method.

	L.S. <sup>a</sup>	SSV <sup>b</sup>	OP 1 <sup>c</sup>	OP 2 <sup>c</sup>
Tyr				
φ	-86.0	-83.5	-86.8	-82.5
Ψ	156.0	155.8	157.0	155.8
ω	-177.0	-177.1	-178.8	-175.8
χ <sup>1</sup>	-173.0	-173.2	-172.8	-172.6
χ <sup>2</sup>	79.0	79.3	79.8	76.7
χ <sup>3</sup>	-166.0	-166.3	-158.2	-156.1
Gly				
φ	-154.0	-154.3	-156.9	-152.4
Ψ	83.0	85.8	87.9	77.9
ω	169.0	168.5	170.1	169.3
Gly				
φ	84.0	83.0	80.8	83.1
Ψ	-74.0	-75.0	-78.0	-73.5
ω	-170.0	-170.0	-169.5	-170.6
Phe				
φ	-137.0	-136.9	-134.0	-133.8
Ψ	19.0	19.1	19.0	18.4
ω	-174.0	-174.1	-174.4	-174.4
χ <sup>1</sup>	59.0	58.9	58.3	57.8
χ <sup>2</sup>	95.0	-85.5	93.5	-87.2
Met				
φ	-164.0	-163.5	-162.7	-162.2
Ψ	160.0	160.9	162.9	162.2
ω	180.0	-179.8	-179.9	180.0
χ <sup>1</sup>	53.0	53.0	51.3	52.5
χ <sup>2</sup>	175.0	175.3	174.7	175.1
χ <sup>3</sup>	180.0	-179.9	-179.1	179.5
χ <sup>4</sup>	-59.0	61.4	61.4	-60.2
Energy (kcal/mol)	-11.7	-11.7	-16.0	-16.0

<sup>a</sup>Taken from ref. 10.

<sup>b</sup>Taken from ref. 11.

<sup>c</sup>Global minimum conformations obtained by the optimization.

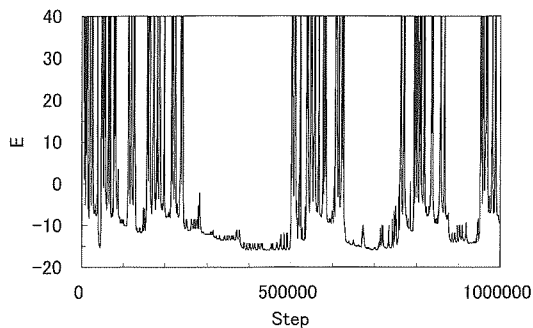


図1 最適化アルゴリズムによるエンケファリンのエネルギー変化の軌跡

と非常によく一致していることがわかる。L.S.構造とSSV構造のエネルギー最小値(-11.7kcal/mol)が我々の結果(-16.0kcal/mol)と異なるのは単に計算に用いたECEPPのバージョンが異なるためである。このように、対数的にスケール変換されたエネルギー空間で最適化を実行すると、非常に効率的にエネルギー最小構造を見出すことができる。我々の計算で得られたエネルギー最小値より、以下のSTES法によるMCシミュレーションでは $E_0 = -17.0$ が使用された。

#### STES法によるモンテカルロシミュレーション

STES法によるモンテカルロシミュレーションは、

計算方法に記述されたような手順で2段階に分けて実行された。図2は、第2段階の長いMCシミュレーションにおけるエネルギー変化の軌跡を示したものである。MCシミュレーションは主鎖の伸びた構造から出発し、高いエネルギー障壁を何度か越えた後に、350,000MCステップ付近で表1のOP2とほぼ同じエネルギー最小構造 (-15.9kcal/mol) に到達している。その後、再度高いエネルギー領域を経て、別の低エネルギー領域を探索している。エネルギー最小値から10 kcal/mol 以内にある低エネルギー領域をサンプリングした比率は、約60%であった。逆に、1000kcal/mol以上の高いエネルギー領域をサンプリングした比率は、約10%であった。このように、サンプリングの大半は低エネルギー領域でなされているが、非常に高いエネルギー領域のサンプリングもかなりの比率でなされている。

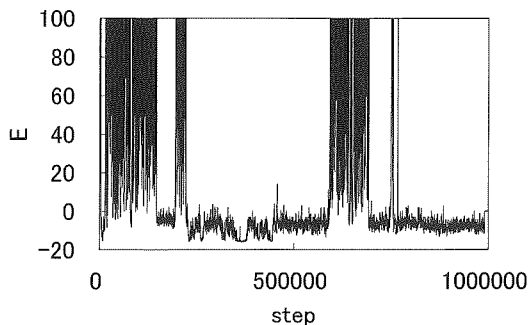


図2 STES 法モンテカルロシミュレーションにおけるエネルギー変化の軌跡

図3は、このMCシミュレーションで得られたエネルギーの確率分布を示したものである。多少起伏が見られるものの分布はスケール変換されたエネルギー空間全体にわたって一様になっていることがわかる。

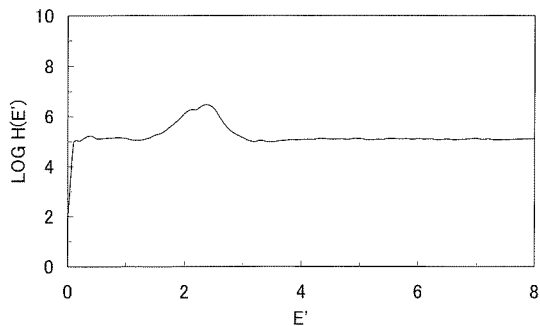


図3 スケール変換したエネルギー空間での確率分布

熱力学的な量

STES 法で得られた  $S(E)$  より、種々の熱学的な量が計算できる。エネルギー  $E$ 、温度  $T$  における分子の自由エネルギーは

$$F(E,T) = E - TS(E) \tag{16}$$

として計算され、分子の平均エネルギー  $\langle E \rangle$  とその2乗平均  $\langle E^2 \rangle$  は、

$$\langle E \rangle = \frac{\int E \exp[-\beta F(E,T)]}{\int \exp[-\beta F(E,T)]} \tag{17a}$$

$$\langle E^2 \rangle = \frac{\int E^2 \exp[-\beta F(E,T)]}{\int \exp[-\beta F(E,T)]} \tag{17b}$$

で計算される。1残基あたりの比熱は、

$$C = \frac{\beta^2}{n} (\langle E^2 \rangle - \langle E \rangle^2) \tag{18}$$

より求められる。ただし、 $n$  はポリペプチド分子の残基数である。

図4は、STES 法によるMCシミュレーションから計算されたエントロピー  $S(E)$  である。 $S(E)$  曲線は、単調に増加していることがわかる。図5は、(17)式より計算された平均エネルギー  $\langle E(T) \rangle$  を示す。 $T=370K$  付近の温度で曲線は変曲点を持つ。図6は、比熱の温度変化を示したもので、この温度付近で比熱のピークがあることがわかる。この結果から、この温度付近でなんらかの構造変化が起きていると考えられる。

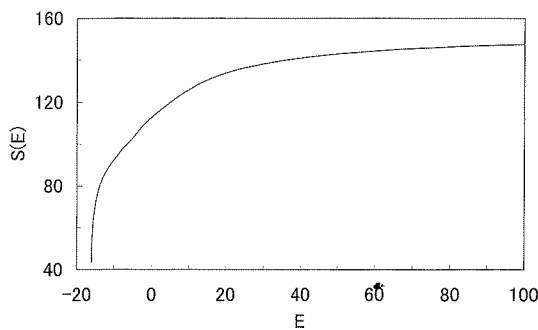


図4 エネルギーに対するエントロピー曲線

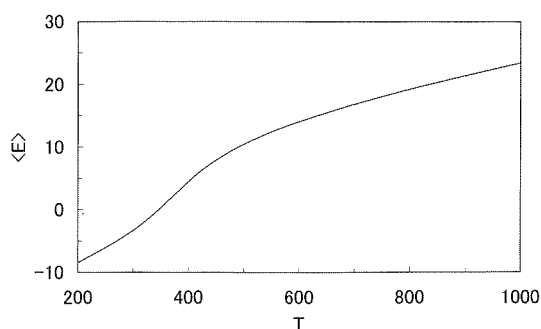


図5 温度に対する平均エネルギー

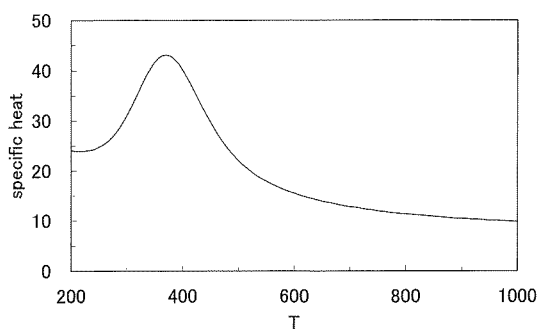


図6 温度に対する比熱の変化

10. J. Lee, H. A. Scheraga and S. Rackovsky, *J. Comput. Chem.*, 18,1222(1997)

11. C. T. Klein, B. Mayer, G. Kohler and P. Wolschann, *J. Comput. Chem.* 19, 1470(1998)

#### 4. 参考文献

1. B. A. Berg and T. Neuhaus, *Phys. Lett. B* 267, 24 (1991)
2. B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* 68, 9 (1992)
3. U. H. E. Hansmann and Y. Okamoto, *J. Comput. Chem.* 14, 1333(1993)
4. J. Lee, *Phys. Rev. Lett.* 71, 211(1993)
5. M. Hao and H. A. Scheraga, *J. Phys. Chem.* 98, 4940 (1994)
6. M. Hao and H. A. Scheraga, *J. Phys. Chem.* 98, 9882 (1994)
7. H. Nakamura, *J. Comput. Chem.* In press.
8. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.* 79, 2361(1995)
9. G. Nemethy, M. S. Pottle, and H. A. Scheraga, *J. Phys. Chem.* 87, 1883(1983)