

myPresto/cosgene 方式設計書
(構造探索エンジン機能追加 編)

第4版

目次 -

1. はじめに	2
1.1 構造探索エンジンの概要	2
1.2 構造探索エンジンの構成	3
2. RESPA 法	4
3. CAP 拘束ポテンシャルの複数指定	5
4. Generalized Simulated Tempering 法	6
5. 自由エネルギー計算 (Filling potential 法 + WHAM 法)	9
6. 自動リスタートファイル出力	12
7. 解析ツール	14
7.1 Canonical 分布への Reweighting	14
7.2 主成分分析 / クラスタ分析	15
参考文献	17

1. はじめに

本設計書では、分子動力学シミュレーションソフト(cosgene)の「構造探索エンジン機能追加」の方式について説明する。

1.1 構造探索エンジンの概要

Multicanonical methodなどの統計物理学的に厳密な高効率の構造探索を行うアルゴリズムに基づく構造探索エンジンを実現する。構造探索エンジンは主に、エネルギー最小化処理、MD計算処理、および、トラジェクトリー解析処理から成る。

(1) エネルギー最小化処理

対象とする分子の初期構造(トポロジー情報と原子座標)からエネルギーが最小になるように構造を変化させていき、エネルギー最小点での分子構造を探索する。

エネルギー最小化処理は、結合距離、結合角など、初期構造が持っている不自然な構造の歪みを取り除き、より安定な構造へと緩和させる。求められた分子構造は、MD計算の初期座標となる。これにより、MD計算の初期における時間積分の発散を避けることができる。

(2) MD計算処理

初期状態での分子の構造(トポロジー情報と原子座標:エネルギー最小化処理の結果を使う)と速度を入力として、分子のconformationalエネルギーを計算する。

多数の原子から構成される仮想的な系(原子の集合体のバルクと同じ性質を持った系)を考え、この系で与えられた原子間の相互作用を用いて各原子に働く力を求め、運動方程式の時間に関する近似式により全ての原子の運動を追跡する。そして、時間に対する原子の位置と速度などの情報からトラジェクトリー(ポテンシャルエネルギー、座標、速度、自由エネルギーなどの物理量)を結果として得る。

(3) トラジェクトリー解析処理

MD計算処理によって出力されるトラジェクトリーを解析する。

生体高分子の安定性に関連する原子の変動、自由エネルギーの変化、生体高分子の周りの水分子の動きなどを明らかにすることができる。

1.2 構造探索エンジンの構成

構造探索エンジンは主に、エネルギー最小化処理、MD 計算処理、および、トラジェクトリー解析処理を中心としたユーティリティ群で構成する。

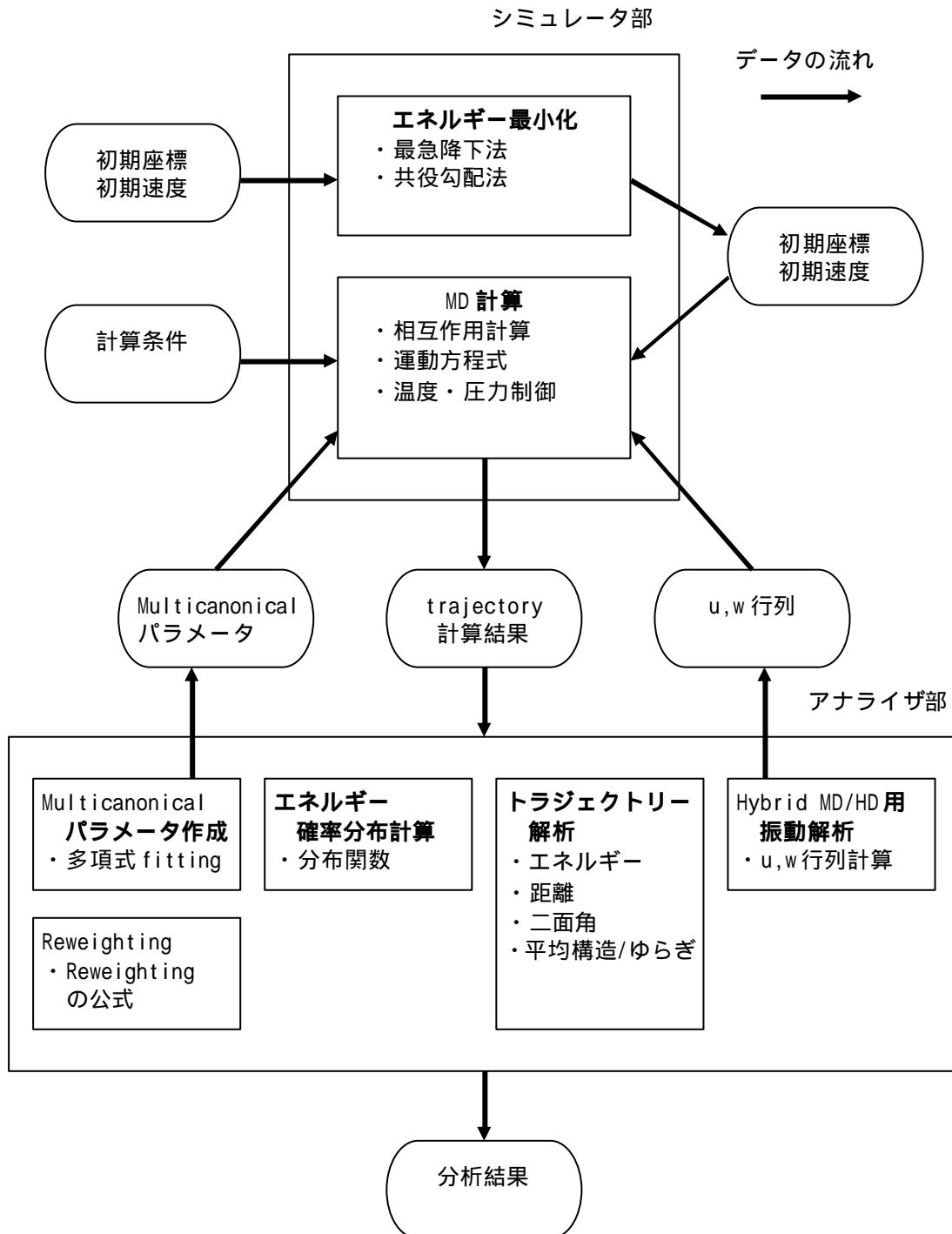


図 1 構造探索エンジンの構成

2. RESPA 法

共有結合の結合伸縮の周期は、特に水素原子が関連しているとき短くなる。単一のタイムステップ t に基づく MD 計算では、最も周期の短い運動にタイムステップを合わせる必要がある ($0.1 \sim 0.5 \text{ fs}$)。しかし、クーロン力や van der Waals 力のなどの力はそう速く変化するわけではなく、 $1 \sim 2 \text{ fs}$ 程度のタイムステップでも十分である。

力の種類に合わせてタイムステップを変える方法として Multiple time step method がある。RESPA (Reference System Propagator Algorithm) は、異なるタイムステップ t をいくつか用いても、時間に対して可逆的な軌跡が得られる数値解法である。この場合の擬似プログラムを以下に示す。

```

given [r(t),v(t)][m,dta,dtb,dtc]
call force
do loopa = 1, nstep
  v = v + dta/2 * fa/m
  do loopb = 1, L
    v = v + dtb/2 * fb/m
    do loopc = 1, K
      v = v + dtc/2 * fc/m
      r = r + dtc * v
      call force(fc)
      v = v + dtc/2 * fc/m
    enddo
    call force(fb)
    v = v + dtb/2 * fb/m
  enddo
  call force(fa)
  v = v + dta/2 * fa/m
enddo

```

変化の遅い力 f_a は毎ステップに一回、中間の力 f_b はより多く、変化の速い力 f_c はさらに多く、というように、それぞれに違うタイムステップ (dta, dtb, dtc) を与える。この方法により、単一のタイムステップを与える場合に比べて、計算精度をあまり落とさずに計算時間を短縮することができる。

NOTE

RESPA は、正確には、Tuckerman らの方法によって導出される任意のアルゴリズムのことを指す。その中で、上述のアルゴリズムは、drRESPA (double reversible REference System Propagator Algorithm) と呼ばれる。

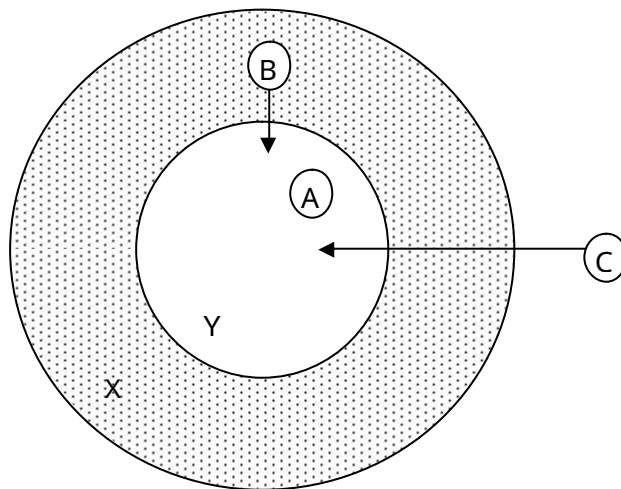
3. CAP 拘束ポテンシャルの複数指定

ある原子をある範囲にとどめておくために、CAP 拘束を使用することができる。CAP 拘束は、任意の位置に球体を定義し、その球体の外にある原子に対して球体の内側に戻るように軽いバネの力 (CAP 拘束力) がかかるようにする。

ここでは、原子が力を受ける CAP を予め指定しておき、1つの系の中に複数の CAP を指定できるようにする。ただし、対象の原子に指定できる CAP は1つだけとする。

CAP 拘束力の計算対象は「チェーン」、「残基」、「原子」の単位で指定できるようにする。また、CAP の球体の内外判定をそれぞれ次のように定義する。

チェーン : チェーンの重心
 残基 : 残基の重心
 原子 : 原子の座標



2つのCAPが重なって、
 以下のように力を受ける
 CAPを指定した場合
 CAP-X : A
 CAP-Y : B, C

指定したCAPだけから
 力を受ける。
 AはYの力を受けない
 CはXの力を受けない

CAPの外になる場合は、指定単位内 (チェーン、残基、原子) の原子に力がかかる。

NOTE

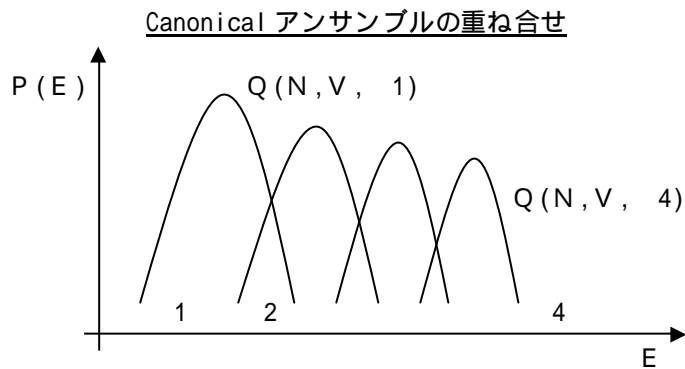
同じ残基または分子内の原子に、それぞれ異なる CAP を指定した場合は、正しいシミュレーション結果を得られない (構造が壊れてしまう) 可能性がある。

4. Generalized Simulated Tempering 法

一般的なエネルギー分布 $P(E)$ は、様々な温度 $\beta_i (=1/k_B T_i)$ での Canonical 分布を重み F_i で重ね合わせたもので表現できる。これまでは、 β で決まるサブアンサンブル Q が重み $F(\beta)$ で重ね合わされたアンサンブル Ξ を考えてきた。

$$\Xi(N, V) = \sum_i Q(N, V, \beta_i) F_i$$

$$\Xi(N, V) = \int d\mathbf{b} Q(N, V, \mathbf{b}) F(\mathbf{b}) \quad (\text{温度 } \beta \text{ を連続に拡張})$$



Generalized Simulated Tempering では、パラメータ \mathbf{I} を導入し、 β で決まる一般のサブアンサンブルの重ね合わせで表す。

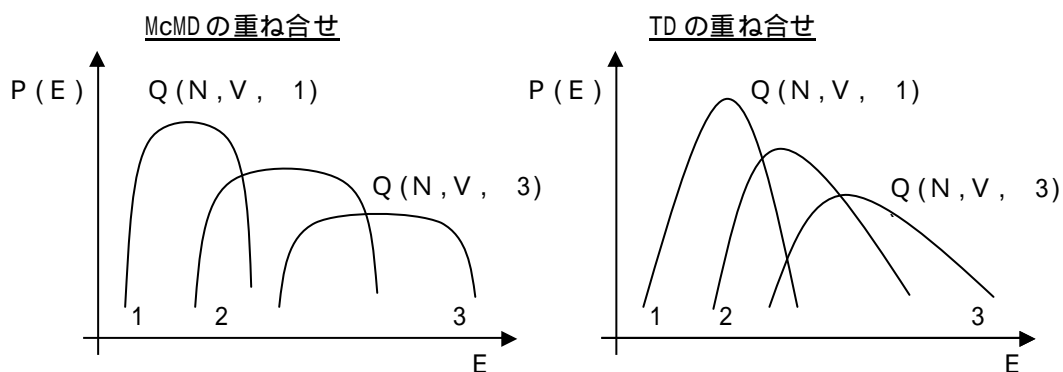
$$\Xi(N, V) = \int d\mathbf{I} Q(N, V, \mathbf{I}) F(\mathbf{I}) \quad \text{--(1)}$$

$$Q(N, V, \beta_i) = \int dE \Omega(E) W_{\mathbf{I}}(E) : \text{温度 } \beta_i \text{ での Canonical 統計}$$

$$W_{\mathbf{I}}(E) = e^{-\beta_0 a_{\mathbf{I}}(E)} : \text{エネルギー } E \text{ での自由度当りの実現確率 (重み)}$$

$$\Omega(E) : \text{エネルギー } E \text{ の状態となる場合の数 (density of state)}$$

$$a_{\mathbf{I}}(E) : \text{effective potential}$$



ある から出発して平衡状態を保ったまま を変化させ、かつ、長時間平均で $F()$ に比例するようにアンサンブルを実現するためには、一定時間間隔で A を計算し、 $0 \sim 1$ の範囲で発生させた一様乱数と比較して確率的に $(x,)$ $(x', ')$ への遷移を行えばよい。

$$A(xI; x'I') = \min \left[1, e^{-b_0(a_{I'}(E') - a_I(E))} \frac{F(I')}{F(I)} \right] \quad --(2)$$

最終的には、 $P() = \text{一定}$ 、となる幅広いサンプリングを実現しなければならないが、そのためには $F()$ を決定する必要がある ($F()$ は一定ではない)。

$$P(I) = \int dE P(E, I) = F(I) Lw[\Omega(E)]$$

同一の の空間座標の遷移は、以下の式で与えられるが、これは、ポテンシャルスケールリングした MD の運動方程式を解くことにより実現される。

$$A(x \rightarrow x'; I) = \min \left[1, e^{-b_0\{a_I(E') - a_I(E)\}} \right]$$

$$\dot{p}_i = -\frac{\partial a_I(E)}{\partial x_i} = \frac{\partial a_I(E)}{\partial E} f_i = v_I(E) f_i$$

次に、 ' の遷移を求める。式(2)に従ってランダムに が変化するとき、この の変化はランジュバン方程式と理論的に同じである。

$$\frac{\partial I}{\partial t} = \Pi(I) + z(t)$$

(t) は時間 t でランダムに働く力、 ($)$ は場所 に平均的に働く力である。

$= 0$ のとき、 は完全にランダムに変化し、 $P() = \text{一定}$ 、 が実現される。

$$\Pi(I) = -U(I) + \mathbf{m}(I) ; U(I) = b_0 \left\langle \frac{\partial \mathbf{a}}{\partial I} \right\rangle_I ; \mathbf{m}(I) = \frac{\partial \ln F(I)}{\partial I}$$

ここで、 $F_u(I) = e^{\{dI' u'(x')\}}$ のとき、 $P() = \text{一定}$ 、のサンプリングが行われることになるが、 U はわからないので、反復解法で求めなければならない。

$$\mathbf{m}^{i+1}(I) = \mathbf{m}^i(I) - \Pi^i(I) ; \Pi^i(I) = \frac{\partial \ln P^i(I)}{\partial I}$$

ここで、 のランダムウォークの条件、

$$\Pi(I) = -U(I) + \mathbf{m}(I) = 0 ; U(I) = \mathbf{m}(I)$$

から収束の速い式にする。

$$\mathbf{m}^{i+1}(I) = U^i(I) - \Pi^i(I)$$

$U^i()$ は、 に関するヒストグラムから $b \left(\frac{\partial \mathbf{a}}{\partial I} \right)$ を平均したものである。

Generalized Simulated Tempering 法のスキーム：

- (1) の範囲を定義し、 をタイムステップ N_{it} ごとに变化させながら MD 計算を行う。
ある温度 (、 F) での MD 計算によって様々な構造をサンプリングする。
ある構造のときに、 の遷移を行う。

$$A(I \rightarrow I'; x) = \begin{cases} I_{\min} & \text{for } I' < I_{\min} \\ \min \left[1, e^{-b_0(a_{I'}(E') - a_I(E))} \frac{F(I')}{F(I)} \right] & \text{for } I_{\min} \leq I' \leq I_{\max} \\ I_{\max} & \text{for } I' > I_{\max} \end{cases}$$

- (2) に関するヒストグラムから、 と U を計算する。

$$\Pi^i(I) = \frac{\partial \ln P^i(I)}{\partial I}$$

$$U_A^i(I) = \frac{\partial a}{\partial I} \text{ の平均}$$

- (3) μ () および重み F をタイムステップ N_{it} ごとに更新することにより、確率分布関数 P () を平らにしていく。

$$\mathbf{m}^{i+1}(I) = \begin{cases} U_A^i(I) - \Pi^i(I) \\ U_A^i(I) \end{cases} \quad \text{for } otherwise$$

$$F^i(I) = e^{\int a_I' U'(I')} \text{、 } \mathbf{m}^i(I) = \frac{\partial \ln F^i(I)}{\partial I} \text{、 から } F() \text{ が求められる。}$$

- (4) 最終的に、得られた温度のヒストグラムから、任意の温度のデータを集計することによって Canonical Sampling を再構成する。

P () がある閾値の範囲にあるときに更新する。

(P () が小さすぎるとノイズになるのでサンプリングの少ないところでは更新しない)

5. 自由エネルギー計算 (Filling potential 法 + WHAM 法)

Filling potential 法と WHAM 法を組合わせて、蛋白質 - リガンドの複合体から解離した状態までの自由エネルギーを計算する。

以下の図の部品群 (網掛け部) を開発し、全体を連動させる。

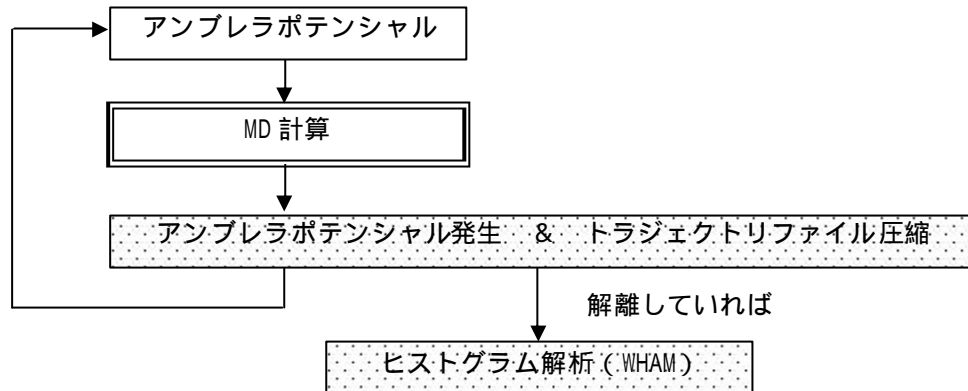


図 . 自由エネルギー計算の手順

(1) アンブレラポテンシャルの発生

過去に経験した座標にトラジェクトリーが戻ってこないように、過去にとった構造周辺に反発ポテンシャルを発生させる。以下に関数系を示す。

$$V_{fillpot}^k = c(k) \exp \left[\sum_i^{landmark} \{ a_x(k,i)(x_i - x_i^k)^2 + a_y(k,i)(y_i - y_i^k)^2 + a_z(k,i)(z_i - z_i^k)^2 \} \right]$$

過去の記憶 (タブー長) が長くなるとメモリーや計算に支障をきたす為、代表的な特徴 (ランドマーク) のみを記憶する。k はその構造の番号、 x_i^k , y_i^k , z_i^k は、構造 k でのランドマークの原子の座標である。 $a_x(k,i)$, $a_y(k,i)$, $a_z(k,i)$ は、 x_i^k , y_i^k , z_i^k の揺らぎ rmsd の逆数に相当する値、 $c(k)$ (>0) はポテンシャルの高さである。ガウス型関数にしたのは、極小点でのポテンシャルに $V_{fillpot}$ を加算することによりポテンシャルをかなり平滑にできる (1次微分と2次微分が0) ことが原理的に可能だからである。

この関数系は、全てのランドマーク原子が過去の座標のランドマーク原子と重なるときのみ、大きな値を持ち、1つでも原子が重ならなければ急速にポテンシャル障壁は減少する性質をもっている。リガンドの母核が蛋白に吸着した場合、母核の位置が動かなくても側鎖が回転すれば反発ポテンシャルは消失し、部分構造の最適化が生じうるように作られている。

$$V_{harmonic}^k = c(k) \sum_i^{landmark} \exp\{a_x(k,i)(x_i - x_i^k)^2 + a_y(k,i)(y_i - y_i^k)^2 + a_z(k,i)(z_i - z_i^k)^2\}$$

これは、1つでもランドマーク原子が過去の座標の原子と重なれば、その1原子分の反発ポテンシャルが加算される。

単純な $V_{fillpot}$ では、極小を埋めたとき、もともと吸着に活性化エネルギーが必要な場合や、一度に埋める potential の高さが高すぎる場合、最後の potential を埋めたときに結果的にポテンシャル面が盛り上がる場合がある。このとき、系は遠い場所まで転がり落ち、トラジェクトリーの分布に重なりがなくなることでヒストグラム法が破綻する。そこで、反応座標が求まった後で、構造 m の V として構造 1 ~ m までの V に構造 m を中心に求心的な調和ポテンシャルを加算することでトラジェクトリーの重なりを失わないようにする。

(2) トラジェクトリの圧縮

Filling potential 法による MD を数回行い、その座標トラジェクトリを保存しておく。このデータ量は大きくなるので、必要なデータだけを取り出して圧縮しておく。

(3) WHAM 法

WHAM 法は、 R 個の trajectory のデータ数を考慮し、推定誤差が最小になるように各 trajectory に重みをつけてまとめ、全体のヒストグラムを以下の式

$$P_{\{I\},b}(\{V\},\mathbf{z}) = \frac{\sum_{k=1}^R N_k(\{V\},\mathbf{z}) \exp(-\mathbf{b} \sum_{j=0}^L I_j V_j)}{\sum_{m=1}^R n_m \exp(f_m - \mathbf{b} \sum_{j=0}^L I_{j,m} V_j)}$$

$$\exp(-f_j) = \sum_{\{V\},\mathbf{z}} P_{\{I\},b_j}(\{V\},\mathbf{z})$$

を同時に解くことで求める方法である。摂動ポテンシャルは L 種類で m 番目のシミュレーションでは、 $\sum_{j=0}^L I_{j,m} V_j$ のポテンシャルを用いているとする。 m 番目の計算の Boltzman 因子は f_m 、サンプル数は n_m 、 $\text{bin}\{V\}$ 、 \mathbf{z} に含まれるサンプル数 (ヒストグラム) N_m とする。 f_m は、 m 番目の状態の自由エネルギー (kbT 単位) である。

ここで分布関数 P は各摂動ポテンシャルエネルギーの関数であり、座標 \mathbf{z} 以外にエネルギーの座標も bin にとらねばならない。分布関数 P が余りに高次の空間で定義されると、

ヒストグラムのデータが疎になり、まともな計算ができない。 をタブー探索のランダムマークをもとにした値とし、摂動ポテンシャル V が のみの関数であれば、 の次元は、 $V_{fillpot}$ を構成した座標の数だけあることになり、上の式は以下ようになる。各シミュレーションの温度・サンプル数を揃えるなどすると更に簡単になる。

$$P_{\{I\},b}(\mathbf{z}) = \frac{\sum_{k=1}^R N_k(\mathbf{z}) \exp(-\mathbf{b} \sum_{j=0}^L \mathbf{I}_j V_j)}{\sum_{m=1}^R n_m \exp(f_m - \mathbf{b}_m \sum_{j=0}^L \mathbf{I}_{j,m} V_j)}$$

$$\exp(-f_j) = \sum_{\{V\},\mathbf{z}} P_{\{I\},b_j}(\mathbf{z})$$

filling potential 法の計算に 20 構造使ったとすると、 は 20 次元のベクトルとなり、これでも多い。そこで精度を上げるため、ヒストグラムを構成する前のデータを用いる。

$$\exp(-f_i) = \sum_{k=1}^R \sum_{t=1}^{nk} \frac{\exp(-\mathbf{b}_i \sum_{j=0}^L \mathbf{I}_{jt} V_{jt}^{(k)})}{\sum_{m=1}^R n_m \exp(f_m - \mathbf{b}_m \sum_{j=0}^L \mathbf{I}_{jm} V_{jt}^{(k)})}$$

プログラムは、データ量を考慮して以下の方法を選択できるようにする。

- (a) ヒストグラムにデータを溜め込まないように、 1 回の計算あたりトラジェクトリの読み込みを R^2 回繰り返しながら処理する。
(メモリ使用量：小、処理時間：長)
- (b) トラジェクトリの読み込みを 1 回にし、中間データをメモリ中に収めて再利用する。
(メモリ使用量：大、処理時間：短)

NOTE

N kcal/mol 以下の binding energy のものに興味がないなら、 $V_{fillpot}$ をあらかじめ N kcal/mol 程度に設定し、直ちに解離するものを省くことも可能である。

$$V_{fillpot} = \sum_{k=1}^{no_of_pot} c(k) \exp\left(\sum_i^{landmark} a_x(k,i)(x_i - x_i^k)^2 + a_y(k,i)(y_i - y_i^k)^2 + a_z(k,i)(z_i - z_i^k)^2 \right)$$

$$+ 2c(last) \sum_i^{landmark} (a_x(last,i)(x_i - x_i^{last})^2 + a_y(last,i)(y_i - y_i^{last})^2 + a_z(last,i)(z_i - z_i^{last})^2)$$

6. 自動リスタートファイル出力

一般に、コンフォメーション・サンプリングなどは、長時間にわたる MD 計算を行わなければならない。このとき、ハード不良、停電、ユーザによる誤停止などの外的要因により、まれに MD 計算が中断されることがある。

このような外的要因での MD 計算の中断の可能性に対して、MD 計算の中間結果のリスタートファイルを定期的に出力するように機能を拡張することで対応する。ユーザは中断直前のリスタートファイルを使用して、リスタートファイル出力時点からの MD 計算を再開することができる。これにより、システムの可用性を向上させる。

リスタートに必要な情報

MD 計算を正確にリスタートさせるためには、MD 計算の各ステップで定義、更新されていく物理量をリスタートファイル経由で引き継ぐ必要がある。

各種計算条件の適用に従い、リスタートファイルに以下の情報を設定する。

表 . 計算条件と必要なリスタート情報

計算条件	リスタート情報
LeapFrogVerlet	タイトル
	原子数
	自由原子数
	MD ループ回数
	シミュレート時間
	総エネルギー
	ポテンシャルエネルギー
	リスタート時の自由原子速度
	リスタート時の原子座標
VelocityVerlet	原子への force
Hybrid	仮想領域の粒子数
	仮想領域粒子の座標
	仮想領域粒子の速度
Nose-Hoover の温度制御	熱浴の座標
	熱浴の運動量
	熱浴の係数
剛体モデル	剛体の速度
	剛体の座標
	剛体の角度
	剛体のモーメント
	剛体への力
	剛体のトルク
	剛体構成原子の相対位置

Tsallis	拡張変数	
	積分 check 値	
NPT	直方体セルのサイズ	
	斜方セルのセル行列	
	セル行列の逆行列	
	セル行列の行列式	
	ピストン	
	熱浴 1	
	熱浴 2	
Force-Biased マルチカノニカル	ベリアル	
	エネルギー	
	当該計算区間のループ回数	
	総ループ回数	
	当該計算区間のヒストグラム	
	総計算区間のヒストグラム	
	力のスケールファクター	
	ヒストグラムの下限	
	ヒストグラムの上限	
	直前のヒストグラムの下限	
	直前のヒストグラムの上限	
	Simulated-Tempering マルチカノニカル	当該スケージング用温度
		エネルギー
当該計算区間のループ回数		
総ループ回数		
当該計算区間の温度ヒストグラム		
総計算区間の温度ヒストグラム		
力のスケールファクター		
当該計算区間のエネルギーヒストグラム		
総計算区間のエネルギーヒストグラム		
ヒストグラムの下限		
ヒストグラムの上限		
直前のヒストグラムの下限		
直前のヒストグラムの上限		
Effective-Temperature マルチカノニカル	有効温度	
	エネルギー	
	当該計算区間のループ回数	
	総ループ回数	
	計算区間増分値	
	力のスケールファクター	
	当該計算区間のエネルギーヒストグラム	
	総計算区間のエネルギーヒストグラム	
	ヒストグラムの下限	
	ヒストグラムの上限	
	直前のヒストグラムの下限	
	直前のヒストグラムの上限	

7. 解析ツール

7.1 Canonical 分布への Reweighting

Force-Biased Multicanonical Sampling 法でのシミュレーション結果は、multiple-histogram method により統合されている。このシミュレーション結果から、以下の Reweighting の公式により、任意の温度 T での Canonical Sampling を再構成することができる。

$$P_0(E, T) = P_{mu}(E) e^{b_0 a_{mu}(E) - bE}$$

P_{mu} : シミュレーションの最後の繰り返しでのエネルギー確率分布関数 P 。

mu : シミュレーションの最後の繰り返しでの Multicanonical ポテンシャル。

$$b_0 : \frac{1}{k_B T_0}$$

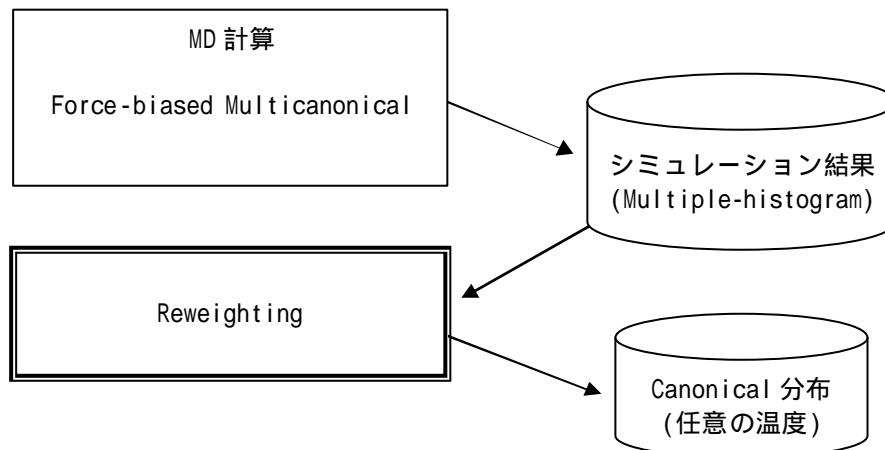


図 . Reweighting 処理の流れ

7.2 主成分分析 / クラスタ分析

(1) 主成分分析 (PCA: Principal Component Analysis)

ある問題に対していくつかの要因が考えられるとき、単に変数を独立に扱うのではなく、それらの総合的な指標を統計的に設定し、データに含まれる変数間の関係や特徴を把握するための手法である。p 変量 (p 次元) の観測値を m 個 (m 次元) の主成分に縮約させるという意味で、次元を減少させる方法と言うこともでき、多変量データを要約する一つの有力な方法である

P 個の変数 $\{x_p | p=1, 2, \dots, P\}$ の持つ情報を、情報の損失を最小限に抑えながら $\{x_p\}$ の一次結合として与えられる互いに独立な M ($M \leq P$) 個の主成分 $\{z_m\}$ を用いて表現する。

z_m は第 m 主成分と呼ばれる。

$$z_m = \sum_{p=1}^P w_{pm} x_p \quad (m=1, 2, \dots, M)$$

結合係数 $\{w_{pm} | p=1, 2, \dots, P; m=1, 2, \dots, M\}$ は、以下の条件を満足するように決定する。

- ・ 第 1 主成分 z_1 の分散は $\{x_p | p=1, 2, \dots, P\}$ のあらゆる 1 次式の中で最大であり、
- ・ 第 m 主成分 $\{z_m | m=1, 2, \dots, M\}$ の分散は $\{z_{m'} | m'=1, 2, \dots, m-1\}$ の全てと無相関な 1 次式の持つ分散の中で最大である。ただし、

$$\sum_{p=1}^P w_{pm}^2 = 1 \quad (m=1, 2, \dots, M)$$

とする。

適当に標準化されたデータ行列が与えられているとき、その共有分散行列の固有ベクトルが主成分の結合係数になり、固有値が主成分の分散に等しくなる。

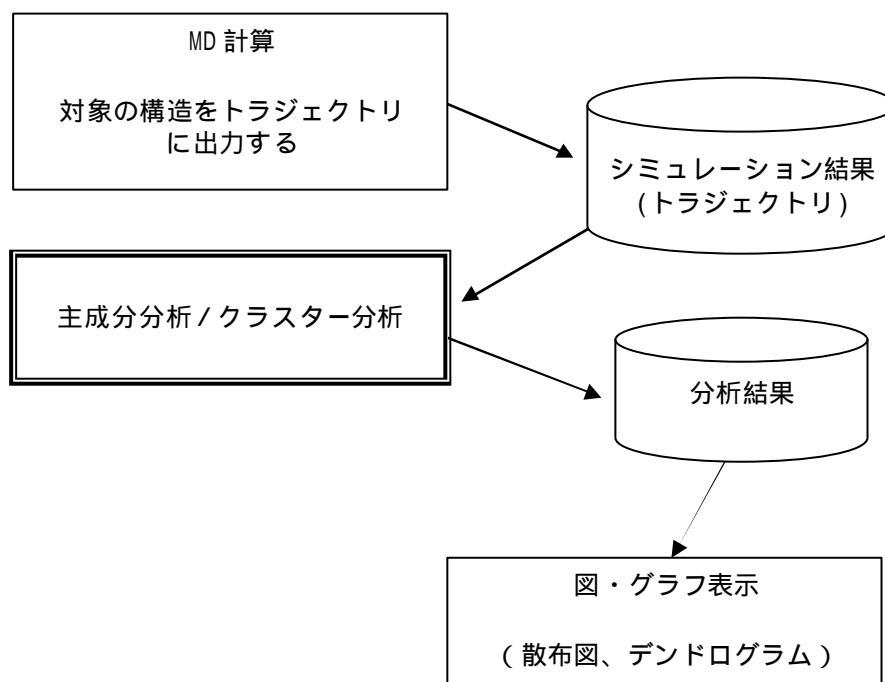
(2) クラスタ分析

異質なものが混じっている対象から、似ているものを凝集し、グループ分けを行うための方法である。ある類似性や距離の測度を利用して、逐次的に大きくなるグループ（クラスター）と一緒に結び付けていく。

- ・ユークリッド距離：多次元空間の幾何的な距離。

$$D(x, y) = \sqrt{\sum_{k=1}^r (x_k - y_k)^2}$$

- ・ワード法：形成される任意の2つのクラスターの平方和が最小になるように試みる。



図．主成分分析 / クラスタ分析処理の流れ

NOTE

ワード法は、小さいクラスターを作りやすい傾向がある。
クラスター間の距離を分散分析的アプローチで評価している。

参考文献

Multicanonical 法 :

- [1] N. Nakajima, H. Nakamura, A. Kidera, Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides, *J. Phys. Chem. B* (1997) 101, 817.
- [2] J. Higo, N. Nakajima, H. Shirai, A. Kidera, H. Nakamura, Two-component multicanonical Monte Carlo method for effective conformational sampling, *J. Comp. Chem.* (1997) 18, 2086.
- [3] N. Nakajima, A selectively enhanced multicanonical molecular dynamics method for conformational sampling of peptides in realistic water molecules, *Chem. Phys. Letters*, (1998) 288, 319.
- [4] H. Shirai, N. Nakajima, J. Higo, A. Kidera, H. Nakamura, Conformational sampling of CDR-H3 in antibodies by multicanonical molecular dynamics simulation, *J. Mol. Biol.* (1998) 278, 481.
- [5] S. T. Kim, H. Shirai, N. Nakajima, J. Higo, H. Nakamura, Enhanced conformational diversity search of CDR-H3 in antibodies: role of the first CDR-H3 residue, *PROTEINS: Structure, Function, and Genetics* (1999) 37, 683.
- [6] H. Nakamura, S. Ono, J. Higo, *Proc. Japan Acad.* A general ab initio approach for free energy landscapes of biological molecules around the transition states: Fusion of the classical molecular mechanics simulation and the quantum chemical calculation. (1999) *Proc. Japan Acad.* 75 Ser.B, 291.
- [7] N. Nakajima, J. Higo, A. Kidera, H. Nakamura, Free energy landscape of peptides by enhanced conformational sampling, *J. Mol. Biol.* (2000) 296, 197.
- [8] S. Ono, J. Higo, H. Nakamura, *J. Comp. Chem.* (2000) Peptide free energy profile is strongly dependent on the force field: Comparison of C96 and AMBER95. *J. Comp. Chem.* 9, 748.
- [9] J. G. Kim, Y. Fukunishi, H. Nakamura, Dynamical origin of uniform sampling in multicanonical ensemble, *Physical Review E* 67, 011105 (2003).

Multicanonical WHAM 法 :

- [10] S. Ono, N. Nakajima, J. Higo, H. Nakamura, The multicanonical weighted histogram analysis method for the free-energy landscape along structural transition paths, *Chem. Phys. Letters*, (1999) 312 , 247.

GB / ASA :

- [11] Ooi, T., Oobatake, M., Nemethy, G. & Scheraga, H. A. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc Natl Acad Sci U S A* 84, 3086-90. (1987).
- [12] Richmond, T. J. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J Mol Biol* 178, 63-89. (1984).
- [13] Kinjo, A. R., Kidera, A., Nakamura, H. & Nishikawa, K. Physicochemical evaluation of protein folds predicted by threading. *Eur Biophys J* 30, 1-10. (2001).
- [14] Still, W. C., Tempczyk, A., Hawley, R., C. & Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* 112, 6127 (1990).
- [15] Hawkins, D., Gregory., Cramer, J., Christopher. & Truhlar, G., Donald. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J Phys Chem* 100, 19824-39 (1996).
- [16] Hawkins, D., Gregory., Cramer, J., Christopher. & Truhlar, G., Donald. Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters* 246, 122-129 (1995).
- [17] Onufriev, A., Bashford, D. & Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J Phys Chem. B* 104, 3712-3720 (2000).
- [18] Bashford, D. & Case, D. A. Generalized born models of macromolecular solvation effects. *Annu Rev Phys Chem* 51, 129-52. (2000).
- [19] Schaefer, M. & Karplus, M. A Comprehensive Analytical Treatment of Continuum Electrostatics. *J Phys Chem* 100, 1578-1599 (1996).