

平成 28 年度 博士学位論文

分子動力学法を用いたラミニン由来ペプチドの構造解析

東京薬科大学 生命科学研究科

生命物理科学研究室

山田 寛尚

<b>1. 序論</b> .....	<b>5</b>
1.1. 分子動力学法の背景.....	5
1.1.1. 分子動力学法の役割.....	5
1.1.2. 分子動力学法による構造予測.....	6
1.1.3. 拡張アンサンブル法.....	7
1.2. ラミニンの実験的背景.....	8
1.2.1. 基底膜.....	8
1.2.2. ラミニン.....	9
1.2.3. ラミニン由来ペプチドの応用.....	13
1.2.4. ペプチドの環状化の例.....	14
1.2.5. 研究の目的.....	15
<b>2. 手法</b> .....	<b>16</b>
2.1. 分子動力学法.....	16
2.1.1. 分子シミュレーションの中での位置付け.....	16
2.1.2. 分子動力学法の概要.....	18
2.2. 周期境界.....	22
2.2.1. 周期境界条件.....	22
2.3. 粒子間に働く相互作用.....	24
2.3.1. レナード・ジョーンズポテンシャルによる相互作用.....	24
2.3.2. 静電（クーロン）相互作用.....	25
2.3.3. エワルド法.....	26
2.3.4. パーティクルメッシュエワルド法（PME法）.....	28
2.4. 差分近似.....	30
2.4.1. ベルレ法.....	30
2.4.2. 蛙飛び法.....	31
2.5. 温度制御および圧力制御法.....	32
2.5.1. 速度スケージング法.....	32
2.5.2. 能勢の方法.....	33
2.5.3. 能勢・フーバーの方法.....	35

2.5.4.	パリネロ・ラーマンの方法.....	36
2.6.	粒子の束縛.....	38
2.6.1.	SHAKE 法.....	38
2.6.2.	SHAKE 条件下の運動方程式.....	39
2.6.3.	Linear constraint solver (LINCS).....	40
2.7.	分子動力学シミュレーションにおいてよく用いられる解析.....	42
2.7.1.	根平均二乗変位 (Root Mean Square Deviation: RMSD).....	42
2.7.2.	慣性半径 (Radius of gyration: $R_g$ ).....	42
2.7.3.	根平均二乗揺らぎ (Root Mean Square Fluctuation: RMSF).....	43
2.8.	自由エネルギー.....	44
2.8.1.	自由エネルギーとタンパク質.....	44
2.8.2.	統計力学における自由エネルギー.....	45
2.8.3.	独立な系におけるハミルトンニアン.....	47
2.8.4.	$T$ - $P$ 分布 ( $NPT$ アンサンブル) における分配関数.....	48
2.8.5.	ギブスのパラドックス.....	49
2.8.6.	分子シミュレーションにおける自由エネルギーの算出.....	51
2.9.	シミュレーションによる構造探索.....	53
2.9.1.	シミュレーテッドアニーリング.....	54
2.9.2.	レプリカ交換法.....	55
2.9.3.	温度間隔の決定.....	57
2.10.	シミュレーションによる自由エネルギー計算.....	58
2.10.1.	アンブレラサンプリング法.....	58
2.10.2.	自由エネルギー計算.....	59
2.10.3.	平均力ポテンシャル (Potential of means force: PMF).....	60
2.10.4.	シミュレーションにおける PMF の計算例.....	60
2.11.	シミュレーションの条件と手順.....	63
2.11.1.	EF1 および EF2 ペプチドの初期構造.....	63
2.11.2.	レプリカ交換法シミュレーションの条件と手順.....	64
2.11.3.	シミュレーションの条件と手順.....	65
3.	結果.....	66

<b>3.1.</b>	<b>レプリカ交換法シミュレーション</b> .....	<b>66</b>
<b>3.1.1.</b>	レプリカの温度空間の遷移の様子.....	66
<b>3.1.2.</b>	精度の確認.....	67
<b>3.1.3.</b>	自由エネルギー地形.....	68
<b>3.1.4.</b>	着目した各レプリカへの遷移の様子.....	72
<b>3.1.5.</b>	自由エネルギー地形.....	73
<b>3.2.</b>	<b>通常の分子動力学シミュレーション</b> .....	<b>76</b>
<b>3.2.1.</b>	通常の分子動力学シミュレーションにより得られた軌跡の解析.....	76
<b>3.2.2.</b>	自由エネルギー地形による評価.....	78
<b>3.2.3.</b>	根平均二乗揺らぎ (RMSF) による評価.....	80
<b>3.2.4.</b>	溶媒露出面積による評価.....	81
<b>3.2.5.</b>	非共有結合性相互作用.....	82
<b>3.2.5.1.</b>	シミュレーション中における水素結合の評価.....	82
<b>3.2.5.2.</b>	水素結合の基準.....	82
<b>3.2.5.3.</b>	水素結合の寿命.....	83
<b>3.2.5.4.</b>	EF1 および EF2 の水素結合数および距離の計算.....	83
<b>3.2.5.5.</b>	水素結合の自己相関関数.....	87
<b>3.2.5.6.</b>	DSSP による二次構造予測.....	89
<b>3.2.5.7.</b>	極性・非極性アミノ酸残基.....	90
<b>4.</b>	<b>まとめと今後の課題</b> .....	<b>92</b>
<b>4.1.</b>	まとめ.....	92
<b>4.2.</b>	今後の課題.....	96
	<b>謝辞</b> .....	<b>97</b>
	<b>参考文献</b> .....	<b>98</b>

## 用語・略語一覧

ミクロカノニカルアンサンブル： $N$  (原子数)  $V$  (体積)  $E$  (エネルギー) 一定

カノニカルアンサンブル： $N$  (原子数)  $V$  (体積)  $T$  (温度) 一定

$T$ - $P$  分布 ( $NPT$  アンサンブル)： $N$  (原子数)  $P$  (圧力)  $T$  (温度) 一定

REMD：replica exchange molecular dynamics, レプリカ交換法

RMSD：root mean square deviation, 根平均二乗変位

$R_g$ ：radius of gyration, 慣性半径

RMSF：root mean square fluctuation, 根平均二乗揺らぎ

PME：particle mesh Ewald

LINCS：linear constraint solver

PMF：potential of mean force, 平均力ポテンシャル

GROMACS：Groningen machine for chemical simulations

DSSP：define secondary structure of protein

SASA：solvent accesible surface area, 溶媒露出面積

H-bonds：hydrogen bonds, 水素結合

# 1. 序論

本章では、研究に用いるシミュレーション手法（分子動力学法）の背景、実験的な背景についておよび研究目的について説明する。

## 1.1. 分子動力学法の背景

本節では生物学、特に生体高分子における分子動力学法の役割、分子動力学法を用いた構造予測について説明する。また、構造探索において有用な方法である拡張アンサンブル法について説明する。

### 1.1.1. 分子動力学法の役割

生体内にはオルガネラ、細胞、組織、器官が存在している。それらを構成している分子一つ一つはそれぞれの構造を持っており、止まっているわけではなく少しずつ動いている。また、激しく動く分子も存在しており、例としては膜を構成するリン脂質が挙げられる。初めは、生体膜は静的と思われていたが、シンガーとニコルソンが提唱した流動モザイクモデルにより流動性のある膜であることが明らかとなった[1]。生体分子の構造を決定する方法として、X線構造解析や核磁気共鳴（NMR）法がある。これらの方法は、静的な分子構造を見る上では有用であり、NMR法においては小さな揺らぎを見ることができる。しかし、NMR法では小さな揺らぎを解析することができても、時系列に沿ったダイナミクスを見ることはできない。そのようなダイナミクスを解析することを可能にするのが分子動力学法である。

分子動力学法は分子一つ一つの動きを時系列に沿って再現するシミュレーション方法であり、構造の小さな揺らぎから大きな構造変化を調べることができる。最初の分子動力学法を用いた研究が報告されたのは1957年のアルダーらによる剛体球系における固相-液相転移の研究である[2]。1964年にラーマンによって、質点系における液体アルゴンの運動が再現された[3]。2.3.1節で説明するが、希ガスであるアルゴンの原子間に働く力はレナード・ジョーンズポテンシャルでよく記述されており、ラーマンはこのポテンシャルを用いてシミュレーションを行った。ラーマンが行ったシミュレーションではアルゴン原子の数が864個の系であったが、今日、コンピュータの発展により、より多くの原子が含まれる系（ $10^4 \sim 10^6$ 個）を計算できるようになってきている。また、目に見える時間スケールではないが、より長い時間スケール（ミリ秒オーダー）における分子の動きに関してシミュレーションを行うことができるようになってきている。例えば、より大きなスケールでのシミュレーションとして、理化

学研究所はスーパーコンピューター「京」を用いて、細胞スケールでの様々な分子が混み合った環境におけるシミュレーションを行っている[4]。このように、分子動力学法は動的な分子の動きを調べるための有用であり、広く利用されている方法である。

### 1.1.2. 分子動力学法による構造予測

分子動力学法は、少なくとも一次構造が分かっている分子の高次の構造を決定する方法として有用である。分子の構造決定は上述した NMR や X 線構造解析を用いて実験的に調べることができ、分子の結晶化がうまくできないなど、構造を調べることが困難なケースがある。このような場合、シミュレーションを用いた構造予測が有用となってくる。しかし、従来の分子動力学法を用いたシミュレーションでは、巨大分子は、局所安定構造に陥りやすいため非常に長い時間スケールで計算する必要がある。それを解決する方法として、金属の加工の焼きなましに由来するシミュレーテッドアニーリング[5]やモンテカルロ法に基づく拡張アンサンブル法(マルチカノニカル法[6]やシミュレーテッドテンパリング[7, 8], レプリカ交換法[9])が挙げられる。これらの方法を用いることで、最安定構造の探索や大域的に構造を探索することができる。

シミュレーテッドアニーリングは、温度を上げて下げることで、局所安定に陥った構造から脱出させ、最安定構造を探索する方法である。この方法は、最安定構造を探索するには有用な方法であるが、大域的に構造を探索するには計算時間が掛かりすぎるため不向きである。

そこで、最安定構造を含め大域的かつ効率的に構造を探索する方法が、拡張アンサンブル法である。現在、数ある拡張アンサンブル法の中で最も一般的な手法がレプリカ交換法である。この方法は、シミュレーテッドアニーリングと比べ、効率的な構造探索を可能とするが、多くの計算資源を必要とする。拡張アンサンブル法は、従来の分子動力学法を用いたシミュレーションと比べ、計算量が多く容易に行えなかった。最近では、並列計算を得意とするグラフィックカード (GPU) を用いた GPGPU (General purpose computing on graphics processing units) 計算機が登場し、拡張アンサンブル法を用いたシミュレーションが高速に行えるようになってきている。

これらの方法は、従来の分子動力学法を用いたシミュレーションと異なり、シミュレーションの途中で温度を変えてしまうため、時系列を追うことはできない。しかし、特に拡張アンサンブル法を用いれば短いシミュレーション時間で最安定構造を含め幅広く構造を探索することができるため、有用な方法である。

### 1.1.3. 拡張アンサンブル法

モンテカルロ法に基づく拡張アンサンブル法は、分子構造の探索において有用な方法である。拡張アンサンブル法としてマルチカノニカル法やシミュレーテッドテンパリング、レプリカ交換法などが挙げられるが、その中で代表的なものがレプリカ交換法である。

マルチカノニカル法は初めはモンテカルロ法において実装された。モンテカルロ法を用いたシミュレーションはある確率を用いて原子の位置を更新していく方法である（メトロポリス法[10]）。通常はボルツマン分布で重み付けをしていくため、設定する温度の範囲内で実現可能な領域の構造を探索していく。マルチカノニカル法ではその確率分布を幅広くフラットにとるよう人工的に作成し、構造探索を行う。ただし、この人工的な確率分布の作成には時間が掛かることに注意する必要がある。

シミュレーテッドテンパリングは、シミュレーションの途中で温度を変えることにより、幅広いエネルギー空間を遷移させ、局所安定に陥ることを防ぐ方法である。手順としては、ある温度間隔に区切られた複数の温度を用意する。この温度空間は、低い温度から局所安定から脱出できるほど十分高い温度に設定する。次に、ある温度 $T$ の条件下で、あるステップ数のシミュレーションを行う。その後、隣り合う温度（現在の温度に最も近い前後の温度）に変更する。この変更はメトロポリス法を用いて行う。変更効率を無視すれば隣り合う温度以外で交換しても良いが、通常は、変更する温度差が大きいほど交換確率が下がるため隣り合う温度で行う。この操作を繰り返すことにより、幅広いエネルギー空間をランダムに動かし、構造探索を行う。この方法の欠点として、任意の温度空間の構造を得るため重みを求めるのが困難であることが挙げられる。

レプリカ交換法はシミュレーテッドテンパリングと同様、シミュレーションの途中で温度を変えて行う。レプリカ交換法の手順としては、異なる温度を持つ空間（系）を複数用意し、それぞれ独立にシミュレーションを行い、途中で異なった温度の系同士の交換を行う。レプリカ交換法は短いシミュレーション時間で幅広い構造を探索することができる。レプリカ交換法は各温度において同時にシミュレーションを行うため、任意の温度空間の構造を得ることは容易である。しかし、計算量が膨大であるところが欠点である。本研究では、温度を交換するレプリカ交換法（パラレルテンパリング）を用いたが、現在では圧力を交換パラメーターとするものやポテンシャルエネルギーを交換のパラメーターとするハミルトニアンレプリカ交換法がある。

これらの方法は構造探索において大変有用であり、本研究では着目するラミニン由来ペプチドの構造探索を行うためにレプリカ交換法を用いた。

## 1.2. ラミニンの実験的背景

本節では研究対象となるタンパク質であるラミニンタンパク質を含む、基底膜の機能について説明する。また、ラミニンの機能およびラミニン由来ペプチドの実験的な背景について述べる。ラミニン由来ペプチドを用いた応用例を紹介する。本研究の目的を述べる。

### 1.2.1. 基底膜

基底膜は真皮と上皮の間、筋肉、神経の周囲に存在する薄い膜状の細胞外マトリックスであり、膜ではあるが生体膜のようにリン脂質は含まない。基底膜は物理的に組織の安定性を保つとともに、生物活性において重要な役割を果たしており[11-13]、組織において構造的な安定性を与えるだけでなく、細胞接着や浸潤、増殖、分化などの多くの生物活性に重要な役割を果たしている。そのため、構成成分の機能や作用メカニズムの解明が注目されているとともに、医薬分野への応用が期待されている。基底膜の構成成分として、代表的なものとしてラミニン、ニドジェン（エンタクチン）、インテグリン、IV型コラーゲン、パールカン（ヘパラン硫酸プロテオグリカン）などの巨大分子が知られている[12]。これらの巨大分子がお互いに結合した超分子ネットワークを形成しており（図 1.2.1）、細胞接着や分化、がんの浸潤転移といった様々な生物学的な活性、細胞の足場という物理的な機能に深く関与していると考えられている。基底膜が存在する部位によって、構成成分も異なっている。

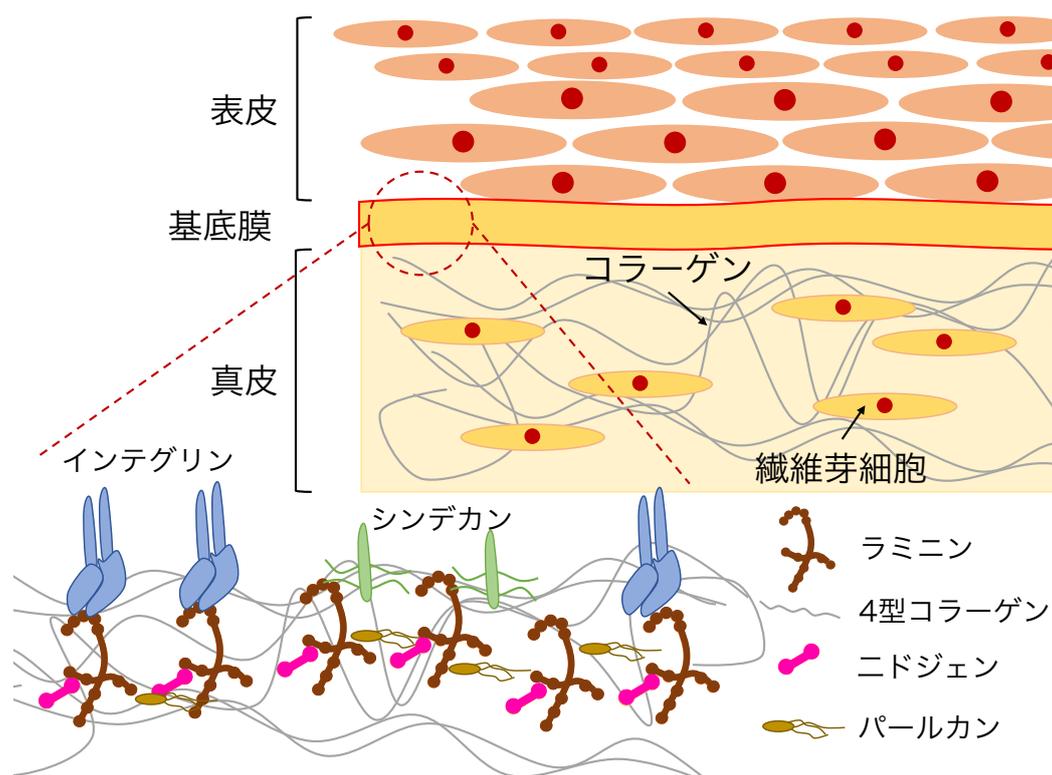


図 1.2.1 基底膜のモデル図 ([13]を参考に山田が作図). ラミニンは赤色で示されている. エンタクチンは黄色, 4型コラーゲンは青色, パールカンは緑色の数珠状のもので示されている. 濃い緑と薄い緑の二つの桿体状のものがインテグリンである.

## 1.2.2. ラミニン

ラミニンは基底膜を構成するタンパク質で, 基底膜中に量的には最も多く含まれており, 細胞接着, 細胞伸展, 神経突起伸長, 細胞増殖, 血管新生, がんの浸潤・転移, 創傷治癒など様々な生物活性を持つ[11-14]. ラミニンは $\alpha$ 鎖,  $\beta$ 鎖,  $\gamma$ 鎖の3つのサブユニットからなり, サブユニットが3本鎖コイルドコイル構造部分で会合し, 全体が十字架構造を持つ分子量約50万~90万の糖タンパク質である[11-15]. ラミニン $\alpha$ 鎖は $\alpha 1$ 鎖,  $\alpha 2$ 鎖,  $\alpha 3A$ 鎖,  $\alpha 3B$ 鎖,  $\alpha 4$ 鎖,  $\alpha 5$ 鎖の6種が知られており, それぞれが組織特異的に発現している(図1.2.2).  $\alpha 1$ 鎖は初期胚,  $\alpha 2$ 鎖は神経や筋肉組織,  $\alpha 3$ 鎖は皮膚組織,  $\alpha 4$ 鎖は血管内皮や脂肪細胞, そして $\alpha 5$ 鎖は上皮細胞に多く発現している. 現在, それぞれのサブユニットの組み合わせにより, 19種類のラミニン・アイソフォーム(ラミニン-111~ラミニン-523)が知られている[16-23]. ラミニンは様々な活性を持つことがわかっているが, これらの活性は受容体を介したシグナル伝達に依存している. 最初に発見されたラミニン-111(LN-111)は $\alpha 1$ ,  $\beta 1$ ,  $\gamma 1$ の3つのサブユニットで構成され, 細胞接着, 器官形成, 神経網再生, 血管新生, 創傷治癒やが

んの増殖転移などに深く関与している[12]. また, ラミニン-111 の受容体としてインテグリンやシンデカンなどの複数の分子が同定され[24], ラミニン-111 より数多くの活性部位が同定されている[25, 26].

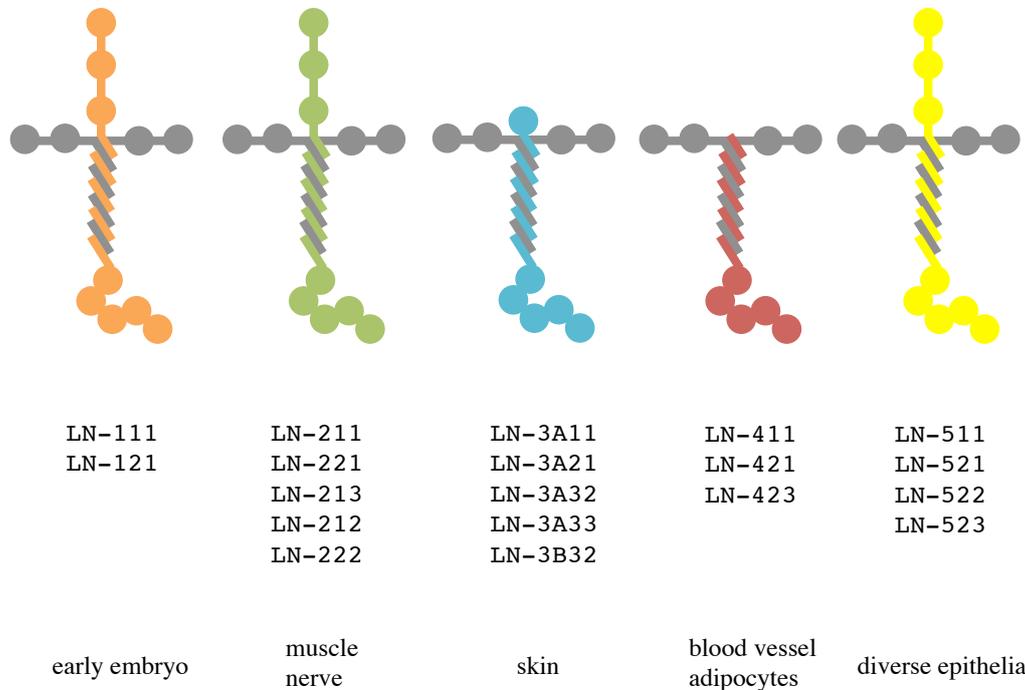


図 1.2.2 ラミニンアイソフォームの種類と局在箇所[16-23].  $\alpha 1$  鎖は橙色,  $\alpha 2$  鎖は緑色,  $\alpha 3$  鎖は青色,  $\alpha 4$  鎖は赤色,  $\alpha 5$  鎖は黄色で示されている. 各  $\beta$ ,  $\gamma$  鎖は灰色で示されている.

近年, ラミニンは人工多能性幹細胞 (iPS 細胞) の培養における接着基質として注目されている分子でもある. ラミニンアイソフォームは 19 種類見つかっているが, LN-511 と LN-512 のアイソフォームが iPS 細胞の培養において利用されている[27, 28]. 現在, 胚性幹細胞 (ES 細胞) や iPS 細胞の培養において, 標準的に利用されているのがマトリゲルやビトロネクチンであるが, LN-511 や LN-512 の接着基質としての接着の強さはそれらよりも強いことがわかっている. しかし, 接着基質として有効な反面, コストが他の接着基質と比べ高いことが問題点である.

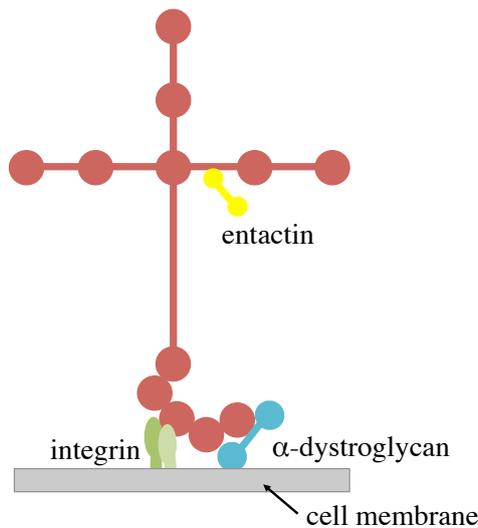


図 1.2.3 ラミニンと受容体の結合部位. 各受容体はラミニンにおける, おおよその結合部位が示されている. エンタクチンは  $\gamma$  鎖, インテグリンと  $\alpha$  ジストログリカンは  $\alpha$  鎖の G ドメインに結合部位がある.

ラミニン  $\alpha$  鎖の C 末端領域には G ドメインと呼ばれる 5 つの LG モジュール (LG1-LG5) が存在する. LG モジュールはそれぞれ, 約 200 アミノ酸残基からなり,  $\alpha$  鎖の生物活性を担っているとされており, インテグリンや  $\alpha$  ジストログリカンなどの細胞膜上に存在する受容体タンパク質と結合することが知られている (図 1.2.3). また, 細胞接着活性などを有する, 様々な活性配列が G ドメインから同定されている [29-34].  $\alpha 1$ ,  $\alpha 4$ , そして,  $\alpha 5$  鎖の G ドメインについて解析が行なわれ, それぞれの LG4 モジュールがヘパリン結合活性と細胞接着活性に重要な働きをしていることが明らかになっている [33-35]. さらに LG モジュールは, 14 本の  $\beta$  ストランド (A-N) からなっており [36], LG モジュールのループ領域においていくつかの活性ペプチドが同定されている [37]. それらは相同性を持ったペプチド (EF1, EF2, EF3, EF4, EF5) であり鎖特異的に存在する. また, それぞれが異なる受容体 ( $\alpha 2\beta 1$  インテグリン, シンデカン-2) と相互作用することが明らかとなっている [37]. EF1 ペプチド (DYATLQLQEGRLHFMDLG) はマウスラミニン  $\alpha 1$  鎖 LG4 モジュールの E-F ストランド間のループ領域の配列を切り出しており,  $\alpha 2\beta 1$  インテグリンと相互作用し, 細胞接着と細胞伸展を促進することが明らかとなっている [37].

EF1 に対する, 細胞接着活性を有するために絶対必要な最小限の配列の探索により, その配列 (LQLQEGRLHFMD, EF1m) が決定された. EF1m ペプチドの活性を測定したところ, EF1 に比べてその活性は低下した. そこで, 図 1.2.4 に示すように, この短縮ペプチド (EF1m) を環状化した. これは cyc-EF1Xm と呼ばれている [37]. cyc-EF1Xm の活性は EF1 配列の活性と同程度までに回復したと報告されている [37]. EF1 ペプチ

ドは図に示す様に、折りたたまれたラミニンタンパク質の構造内ではヘアピン様構造を形成している。実験的に EF1 単体の高次構造は明らかとなっていないが、同様の配列をもつポリペプチドのタンパク質内の二次構造としてヘアピン様構造を持つことから (図 1.2.5), ヘアピン様構造が活性発現に重要であることが示唆された。[37]の論文では短縮ペプチドを環状化の際、メチオニンが酸化されやすいため、ノルロイシンに置換したペプチド (LQLQEGRLHFXFD, X: norleucine, EF1Xm) を用いて、環状化 (CLQLQEGRLHFXFDC, X: norleucine, cyc-EF1Xm) している。ノルロイシンに置換した際、細胞接着活性に違いは出ていない[37]。また、環状化はシステイン-システインによるジスルフィド結合が用いられている。

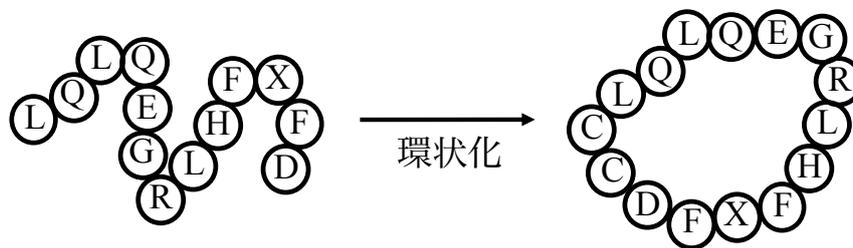


図 1.2.4 ペプチドの環状化。

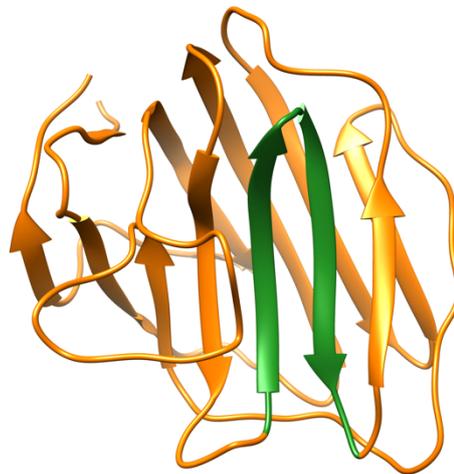


図 1.2.5 ラミニン  $\alpha 1$  鎖 LG4 モジュールの X 線結晶解析により得られた構造 (PDB ID: 2JD4 [38])。緑色で表されている領域に EF1 配列が存在する。

環状化による活性の向上はラミニン由来ペプチドの他にも報告されている。例えば、アリらはインターロイキン 2 (IL-2) の生産を阻害する活性を持つ T 細胞抗原受容体 (TCR) の  $\alpha$  鎖の膜貫通領域由来のペプチド (GLRILLKLV, CP) について研究を行っており、CP をアミド結合で環状化した C1 ペプチドが IL-2 生産の阻害活性を増加

させたと報告している[39]. また, アリらは, このペプチドを環状化することで, ジミリストイルホスファチジルコリン (DMPC) に対して, より強く結合することを明らかにした. C1 ペプチド単体の構造は解析していないが, DMPC に対する結合は脂質二重膜中に C1 ペプチドを含んだ膜と含んでいない膜を,  $^2\text{H}$  NMR スペクトルと  $^{31}\text{P}$  NMR スペクトルによる化学シフトを調べ, C1 ペプチドが含まれる膜における化学シフトは二つの測定 ( $^2\text{H}$  NMR,  $^{31}\text{P}$  NMR スペクトル) で共に増加していることから明らかとなった[39]. アリらの実験からも, 構造 (形) と生物活性 (機能) の関連性が示唆されている.

一方, LM-211 や LM-221, LM-213, LM-212 などの構成要素であるラミニン  $\alpha 2$  鎖は骨格筋, 末梢神経, 脳, 毛細血管, 顎下腺などに発現しており, 筋ジストロフィーの発症や末梢神経障害に関連している[40-42].  $\alpha 2$  鎖はインテグリンや  $\alpha$ -ジストログリカン ( $\alpha$ -dystroglycan,  $\alpha\text{DG}$ ), ヘパラン硫酸などと結合し[43-46], LG4-5 モジュールに  $\alpha\text{DG}$  の結合部位があることが示唆されている[43]. 鈴木らは  $\alpha 2$  鎖 LG4-5 モジュールより,  $\alpha\text{DG}$  と結合するペプチドを同定している[46]. また, 片桐らの報告より,  $\alpha 2$  鎖に存在する EF1 の相同配列 (相同率は約 52.6%) である EF2 ペプチド (DFATVQLRNGFPYFSYDLG) は, 細胞接着活性を示さないことが明らかとなっている[47]. EF2 もまた EF1 と同様に, 折りたたまれたラミニンタンパク質の構造内ではヘアピン様構造を形成している.

### 1.2.3. ラミニン由来ペプチドの応用

生物活性を有するラミニン由来ペプチドをキトサン膜に固定したペプチド-キトサン膜 (図 1.2.6) が基底膜の機能を模倣することが示唆されており, ラミニン由来ペプチドは生体材料や医薬品への応用が期待されている. 基底膜そのものを使うのではなく, 人工的な膜を使う利点として, 選択的に基底膜の機能を持った膜を作製可能な点が挙げられる. 基底膜そのものは様々な分子が絡み合った超分子ネットワークを形成しており, 細胞接着や分化といったプラスの面だけでなく, がんの浸潤転移といったマイナスの面にも関与している. そのため, 基底膜そのものを使うことはできない場合もある. 細胞接着や細胞伸展を有するペプチドをキトサン膜に固定したペプチド-キトサン膜が個々のペプチドで測定した活性よりも高い活性を示したと報告されている[48]. また, 神経突起伸長を有する 4 つのペプチドを固定したキトサン膜の場合においても, 個々のペプチドで測定した活性よりも向上していることが確認されている[48].

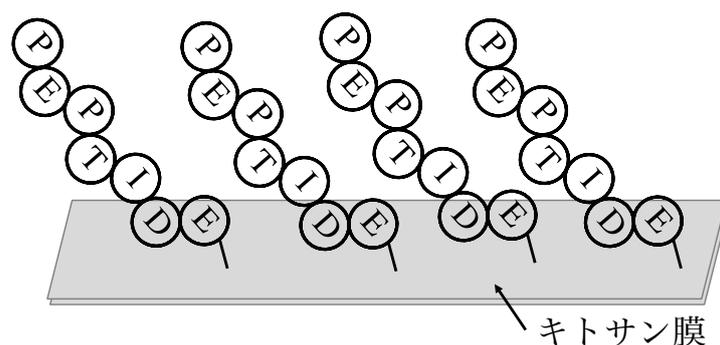


図 1.2.6 ペプチド-キトサン膜のモデル

#### 1.2.4. ペプチドの環状化の例

ハーヤウカからは、LEDGF/p75 タンパク質由来の HIV-1 インテグラーゼ阻害ペプチドについて、そのペプチドの構造を環状化した場合、活性にどれだけ影響するかの研究を行った[49]。彼らが行った環状化では末端間を架橋しており、架橋するループ構造の長さの違いによる活性への影響を調べている。ハーヤウカからは環状化することにより、ペプチドの安定性と活性が向上することを明らかにした。環状化ペプチドの構造は NMR を用いて測定しており、主にループ構造の安定性を根平均二乗変位 (RMSD) の値を用いて評価している。RMSD は構造変化の指標としてよく用いられ、値が大きいほど構造変化していることを表す。HIV-1 のコアを構成するキャプシドタンパク質の複製阻害実験により、環状化したペプチドは 8 日後も阻害活性を維持していた。2 日毎に投与したケースでは直鎖ペプチドも 8 日後においても阻害活性を維持していたが、環状化したペプチドの方がより、タンパク質の複製を阻害することが報告されている。

ハーヤウカからはまた、HIV-1 インテグラーゼ阻害ペプチドの環状化について、さらに研究を行った[50]。その研究で彼らはペプチドの環状化において、ペプチドの主鎖を対象に環状化した場合と、側鎖を対象に環状化した場合を比較した。その結果、環状化の際、側鎖を対象にした場合は、主鎖を対象にした場合に比べ、阻害活性は低下した。この環状ペプチドにおいても、NMR を用いて構造を決定しており、ループ構造の安定性について RMSD 値を用いて評価していた。また、ハーヤウカからは、環状化の作業において、主鎖を対象にする場合には、側鎖を対象とした環状化に比べコストが掛かると述べており、主鎖、側鎖ともにメリット、デメリットがあることを報告した。

### 1.2.5. 研究の目的

鈴木らの実験結果から、EF1 ペプチドは  $\alpha 2\beta 1$  インテグリンを介した細胞接着活性を持つことが明らかとなり、その活性はヘアピン様構造と関連することが示唆された。一般に相同性が高い配列ならば、同様の活性を持つ可能性が高いが、EF1 の相同配列である EF2 ペプチドは細胞接着活性を持たないことが示された。このような実験背景の下、本研究では、EF1 および EF2 ペプチドが水溶液中でどのような 2 次構造を形成しているか、EF1 および EF2 の構造の違いを調べるため、分子動力学シミュレーションによる構造解析を行う。また、EF1 と相互作用する受容体である  $\alpha 2\beta 1$  インテグリンとの結合様式がどのようなものかを予測する。

## 2. 手法

本章では分子シミュレーション手法のひとつである分子動力学法，その手法を用いる上で必要な物理的な相互作用とコンピュータによる相互作用の計算方法について述べる．分子動力学法を用いたシミュレーションにおいて，アンサンブルという単語が出てくるが，これはある微視的状态が多数集まったものを言う．分子動力学法における重要なアンサンブルとして，ミクロカノニカル（エネルギー一定： $NVE$ ）アンサンブル，カノニカル（等温定積： $NVT$ ）アンサンブル，圧力（等温定圧： $NPT$ ）アンサンブルが挙げられる．ミクロカノニカルアンサンブルはもっとも単純な系で，シミュレーションは系のエネルギーを一定に保ちつつ行われる．しかし，多くの場合，温度一定，圧力一定の系が求められる．これは温度や圧力が一定の下で，実験が行われているからである．カノニカルアンサンブルや  $NPT$  アンサンブルを実現するために，温度制御や圧力制御をする必要がある．そこで温度や圧力を制御する手法を紹介する．また本研究において，ペプチドの解析に自由エネルギーを用いた．ここで述べるのは熱力学的な自由エネルギーではなく，統計力学における自由エネルギーである．そこでここでは，統計力学における自由エネルギーの扱いとシミュレーションにおける自由エネルギーの計算方法について説明する[51-57]．最後に，本研究のシミュレーション条件および手順を述べる．

### 2.1. 分子動力学法

本節では分子の動きを時系列に沿って追跡する手法である分子動力学法の原理について説明する．

#### 2.1.1. 分子シミュレーションの中での位置付け

分子シミュレーションとは，コンピュータを用いて物質科学における現象を探る方法である．例えば，分子のエネルギー計算や最安定構造の探索など，物質を分子レベルで解析することができる．また近年，計算機の進歩は目覚ましく，タンパク質や生体膜などの生体高分子の構造変化に関してシミュレーションを行うことが比較的容易になってきている．メリットとして，コンピュータグラフィックスにより立体構造を見ることができること（図 2.1.1），実験系において困難なことを実行可能であること（任意に状態・条件を決められる），そしてコストの低下に繋がることなどが挙げられる．デメリットとして，構成する系によって計算時間が異なり，サイズが大きくなるほど時間が掛かる．また，シミュレーション時間スケールについて，多くはナノ

–マイクロ秒（現在、現実的に計算可能な時間スケール）であり、実験系の時間スケールと一致しないなどが挙げられる。これはおもに、溶媒を含めた全原子系におけるシミュレーションであり、原子の自由度が増えれば増えるほど時間が掛かるためである。このような問題は、今後計算機の進歩により解決の可能性がある。例えば、2009年に分子動力学専用計算機 Anton を用いた、これまで計算ができていなかったミリ秒スケールの長時間シミュレーションが報告されている[58]。近年におけるスーパーコンピュータの性能向上は超並列により図られており、例えば、京コンピュータは64万個以上のプロセッサコアから構成されている。このような超並列計算機の性能を十分に利用するためにはある程度の大きな規模の計算を行う必要がある。しかし、分子動力学計算は1ステップあたりの演算数が比較的少なく、通常のスーパーコンピュータでは有効に利用することができない。この問題を解決したのが Anton である。日本では分子動力学専用計算機として、MDGRAPE が開発・推進されている。これにより、タンパク質の長時間の揺らぎやフォールディングの解析などが実現可能なものとなってきている。しかし、このような大型計算機はコストが非常に掛かることや、多くの人々が利用するため、計算待ちをする必要があり、自由に計算することができないというデメリットもある。また、上述した Anton は、米国以外での利用について制限が設けられている。

分子シミュレーションを実現する代表的な方法には、分子動力学法と確率的な手法であるモンテカルロ法が挙げられる。分子動力学法は系を構成する原子や分子1つ1つに対し、運動方程式を数値的に解くことにより、位置、速度の時間変化やエネルギーなどを追跡することができる方法である。モンテカルロ法は系を構成する原子や分子の位置や速度をランダムに変化させ、統計平均から熱平衡状態における値を求める方法である。大きな違いとして、分子動力学法の場合は時間変化を直接知ることができる点が挙げられる。また、水溶液中の構造をサンプリングする場合、分子動力学法が用いられることが多い。これは、分子動力学法において、タンパク質周りの水との相互作用は直接計算することができるからである。モンテカルロ法の場合においても、溶媒を直接導入することは可能であるが、採択率（系が異なる状態に移れる確率）が非常に悪いため、溶媒効果として間接的に導入されることが多い。

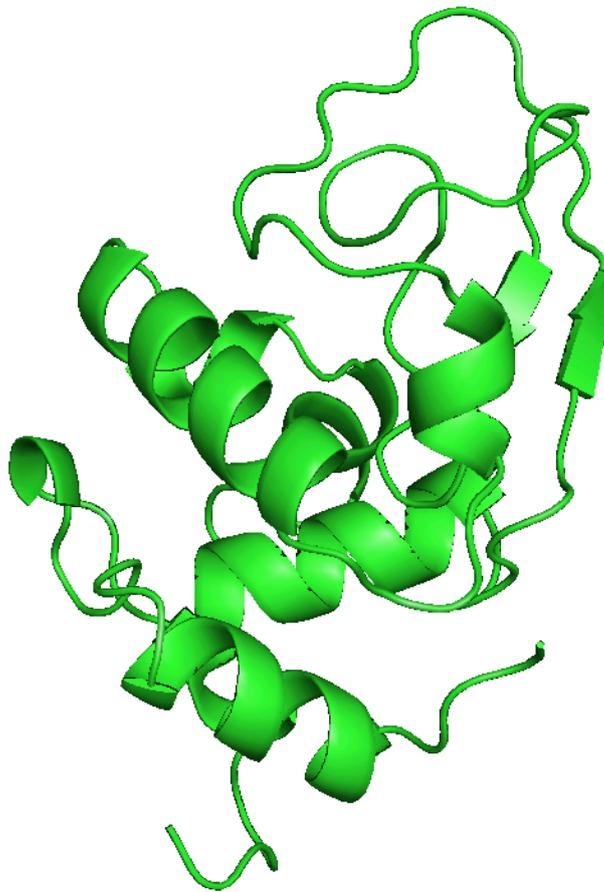


図 2.1.1 生体高分子の立体構造の可視化例：リゾチームの構造 (PDB ID: 1AKI [59])

### 2.1.2. 分子動力学法の概要

分子動力学法は $N$ 個の粒子からなる系に対し、ニュートンの運動方程式  $\mathbf{F} = m\mathbf{a}$  を解くことにより、粒子の動き（軌跡）を追跡することができる。図 2.1.2 にその概要を示す。

$N$ 個の粒子からなる物質系の運動方程式は

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2}, \quad i = 1, 2, \dots, N. \quad (2.1)$$

と表すことができる。 $m_i$  は  $i$  番目の原子の質量、 $\mathbf{r}_i$  は  $i$  番目の原子位置、 $t$  は時刻を示し、そして  $\mathbf{F}_i$  は  $i$  番目の原子に働く力である。また、 $\mathbf{F}_i$  はポテンシャル関数  $V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  についての導関数に負の符号を付けたもので表せ、

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad (2.2)$$

となる。通常は式 (2.1) を初期値問題として解く。式 (2.1) を時間について積分することにより、任意時間での各粒子の速度、位置を予測することができる。当然、取り扱おうとする系全体の自由度の数だけ運動方程式を解かねばならない。しかし、一般に三体以上の系を解析的に解くのは不可能であるため、運動方程式を時間に沿って、数値的に解いていく必要がある。方法として差分近似法（ベルレ法）、予測子-修正子法、蛙飛び法（Leap-Frog 法）などが挙げられる。また、計算精度が高い、ルンゲ・クッタ法という方法もあるが、計算時間が他と比べ何倍もかかるため、通常、粒子数が多い系に対する分子動力学法では用いない。

運動方程式は、時間に関する 2 階微分方程式で与えられるため、それを解くには初期条件（時刻  $t = 0$  における位置  $\{\mathbf{r}_{i0}\}$  と速度  $\{\mathbf{v}_{i0}\}$ ）を与える必要がある。初期速度  $\{\mathbf{v}_{i0}\}$  を与えるために、温度  $T$  と  $N$  個の原子からなる系の全運動エネルギーを利用する。

$N$  個の原子からなる系の平均の運動エネルギー  $E_{\text{kin}}$  は

$$E_{\text{kin}} = \frac{1}{2} \left\langle \sum_{i=1}^N m_i \mathbf{v}_i^2 \right\rangle \quad (2.3)$$

と書ける。原子の速度は  $x, y, z$  成分に分けられ、 $\langle \rangle$  は長時間平均またはアンサンブル平均である。 $\mathbf{v}_i^2 = v_x^2 + v_y^2 + v_z^2$  であるので、 $x$  成分における速度の二乗  $\langle v_x^2 \rangle$  は

$$\langle v_x^2 \rangle = \frac{1}{N} \sum_{i=1}^N v_{xi}^2 \quad (2.4)$$

となる。 $y, z$  成分についても同様である。そして、温度  $T$  と運動エネルギーの関係式（エネルギー等分配則）

$$\frac{3}{2} N k_B T = \frac{1}{2} \left\langle \sum_{i=1}^N m_i \mathbf{v}_i^2 \right\rangle \quad (2.5)$$

を満たすように平衡状態の速度分布関数であるマクスウェル分布を適用した。ここで、 $k_B$  はボルツマン定数である。温度  $T$  での速度の  $x$  成分  $v_x$  に対する速度分布関数であるマクスウェル=ボルツマン分布は

$$f(v_x) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{mv_x^2}{2k_B T}\right) \quad (2.6)$$

である。 $y, z$  成分についても同様の形となる。また、各成分の分布は互いに独立である。各速度成分を含めた式は、次のようになる。

$$f(v_x, v_y, v_z) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T}\right) \quad (2.7)$$

式 (2.7) に従うように、乱数を使用して各原子に速度が割り振られる。平均の速度は速度分布から求めることができ、速度の  $x$  成分  $v_x$  の平均は以下の式を積分すると得られる。

$$\langle v_x \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v_x f(v_x, v_y, v_z) dv_x dv_y dv_z \quad (2.8)$$

この積分の値はゼロとなる。  $y, z$  成分の速度の平均も同様であり、

$$\langle v_x \rangle = \langle v_y \rangle = \langle v_z \rangle = 0 \quad (2.9)$$

となる。これは、マクスウェル=ボルツマン分布における速度分布が等方的であることを表している。

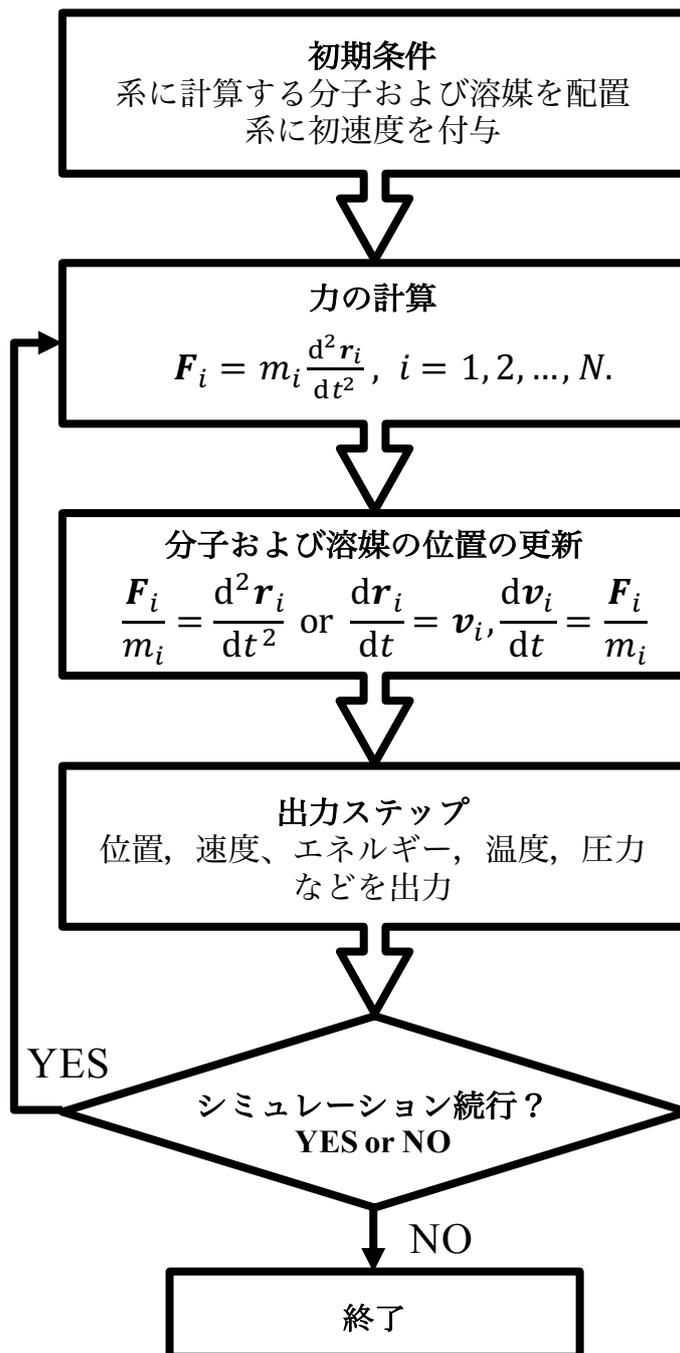


図 2.1.2 分子動力学シミュレーションの流れ

## 2.2. 周期境界

分子動力学法は分子の動きを一つ一つ追跡することで、時系列に沿った分子のダイナミクスをみることができする方法である。実際の観測において、目に見えるスケールにおける分子の数はアボガドロ数個 ( $6.02 \times 10^{23}$  個) と非常に大きな個数である。この数を現在あるコンピュータを用いて直接計算することは不可能であり、大きさとしては  $10^4 \sim 10^6$  個程度の系が現実的である。最新の大型計算機 (スーパーコンピュータ) を用いれば、 $10^7$  個程度ぐらいまでは計算することが可能である。

また、数が少ないと別の問題も発生してくる。それは、分子の壁との相互作用である。壁近傍とそうでないところでは、力の加わり方が変わり、分子のふるまいが異なってくる。この影響は、分子の数が少なくなればなるほど大きくなる。計算する系の大きさ ( $10^4 \sim 10^6$  個) でのこの影響は、目に見えるスケール ( $10^{23}$  個) と比較すると、無視できなくなるほど大きい。そこで、これらの問題を解決するために、周期境界条件と呼ばれる手法が広く用いられている。

### 2.2.1. 周期境界条件

$N$  個の粒子が一辺  $L$  の長さを持った箱で運動していると仮定する。各粒子の座標を  $x_i, y_i, z_i$  とする。もし箱の内壁が弾性壁ならば、粒子は壁にぶつかったとき弾性的に跳ね返される。この場合、粒子の運動範囲は、

$$0 \leq x_i, y_i, z_i \leq L \quad (2.10)$$

である。この条件において、粒子は他の粒子以外に、壁からも力を受けることになる。このように、粒子の運動を箱内に制限する条件を、固定弾性壁の境界条件と呼ぶ。

そこで、周期境界条件では、粒子は壁からの力を受けることなく、そのまま箱を出て、対称となる位置から同じ速度で入ってくる。この運動を2次元平面内で見ると、図 2.2.1 のようになる。中央を着目する系とし、これを基本セルと呼ぶ。基本セル以外、周りのセルをイメージセルと呼ぶ。イメージセルは基本セル内の運動と全く同じ運動をするため、飛び出た粒子は反対側の境界から入ってくることになる。このようにして擬似的に無限に大きな系を扱うことができる。

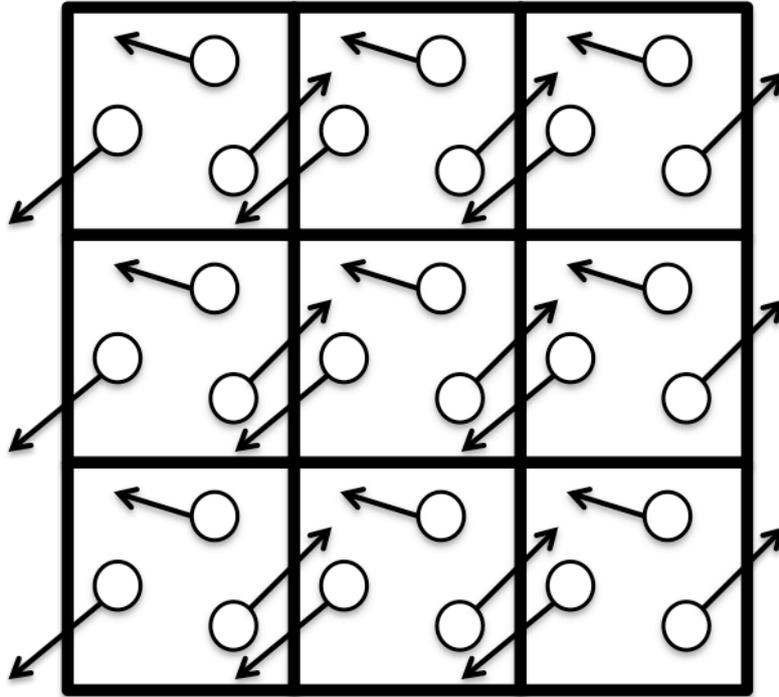


図 2.2.1 二次元における周期境界条件

## 2.3. 粒子間に働く相互作用

一般に、結合していない原子間に働く相互作用として、引力および原子半径程度より近づかないように近距離で働く斥力を含む分子間力と長距離まで働く静電（クーロン）相互作用を主として考慮し、それらポテンシャル関数の和として表現される。

$$V(r_{ij}) = V_{LJ}(r_{ij}) + V_C(r_{ij}) \quad (2.11)$$

$V(r_{ij})$  が結合していない原子間に働く相互作用の和、 $V_{LJ}(r_{ij})$  と  $V_C(r_{ij})$  はそれぞれ、レナード・ジョーンズポテンシャルと静電ポテンシャルを表す。

### 2.3.1. レナード・ジョーンズポテンシャルによる相互作用

原子間に近距離において働く斥力と、中距離に働く引力を含む分子間力は、レナード・ジョーンズポテンシャル（図 2.3.1）として表現される

$$V_{LJ}(r_{ij}) = 4\varepsilon_{ij} \left\{ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right\}. \quad (2.12)$$

12 乗の項は斥力項、6 乗の項は引力項を表す。 $\sigma_{ij}$  と  $\varepsilon_{ij}$  は粒子  $i$  と  $j$  との間の LJ パラメーターである。 $\sigma_{ij}$  は距離の次元を持っており、 $r_{ij} \rightarrow +\infty$  のときと  $r_{ij} = \sigma_{ij}$  のときにポテンシャルがゼロになることがわかる。 $\varepsilon_{ij}$  はエネルギーの次元を持ち、ポテンシャルエネルギーの深さを表している。原子間距離が  $r_{ij} = 2^{1/6}\sigma_{ij}$  のとき  $V = -\varepsilon_{ij}$  になる。これは底におけるポテンシャルを表し、 $r_{ij} = 2^{1/6}\sigma_{ij}$  のときが安定した原子間距離となる。つまり、2 個の粒子が近づく（ $r_{ij}$  が小さくなる）と、引力項より斥力項が大きくなり、粒子が反発する。2 個の粒子が離れる（ $r_{ij}$  が大きくなる）と斥力項より引力項が大きくなり、粒子同士が引き合う。さらに離れると、相互作用は 0 になる。

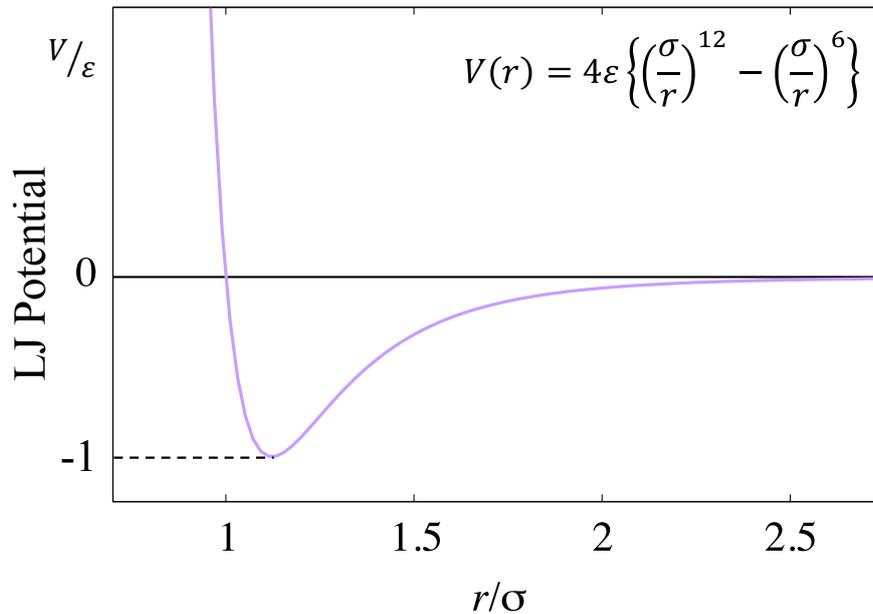


図 2.3.1 レナード・ジョーンズポテンシャル.

### 2.3.2. 静電（クーロン）相互作用

静電ポテンシャルは以下の式で表される

$$V_C(r_{ij}) = \frac{1}{4\pi\epsilon} \frac{q_i q_j}{r_{ij}}. \quad (2.13)$$

$q_i$  と  $q_j$  はそれぞれ粒子  $i$  と  $j$  の電荷,  $\epsilon$  は空間の誘電率を表している. クーロン相互作用はレナード・ジョーンズポテンシャルのような短距離力とは異なり, 距離に対して, 減衰が極めてゆっくりである. その分, 計算すべき距離が長くなる. そのため, 分子シミュレーションにおいて, このクーロン相互作用の計算が最も時間を要する. もし基本セルの半分の長さでカットしたとすると, 大きな誤差を招いてしまう. それを避けるために遠方のイメージセルの相互作用も評価する必要がある. しかし, これを計算するのは現実的に不可能であるため, 効率よく見積もる方法が必要である. その方法にエwald法 (2.3.3 節), パーティクルメッシュエwald法 (Particle Mesh Ewald: PME 法) がある (2.3.4 節). 本研究においては, PME 法を用いて静電相互作用を計算した.

### 2.3.3. エワルド法

エワルド法は、粒子間のクーロン相互作用を精度よく計算するためにエワルドが開発した方法であり、イオン結晶のポテンシャルエネルギーを求めるために開発された。距離  $r$  に反比例する静電ポテンシャルは、距離に対するポテンシャルの減衰（収束）が非常に遅く、レナード・ジョーンズポテンシャルのような短距離力の計算と比べ、膨大であることは前節で述べた。エワルド法の考え方の要点は、収束の遅い静電ポテンシャルを誤差関数と補誤差関数を用いて、収束が速くなる形に分解することにある。

イオン結晶は周期的に並んでおり、電荷の総和が 0 になる格子が周期的に並んでいる。シミュレーションにおいては、それぞれの格子が周期境界条件における基本セルとイメージセルに相当することとなる。体積  $V = L^3$  のユニットセルにおける、電荷の数を  $N$  個とし、イオン  $i$  の電荷を電気素量  $e$  を用いて  $Z_i e$  とおくと、それら電荷の総和はゼロになるので、

$$\sum_{i=1}^N Z_i e = 0 \quad (2.14)$$

となる。クーロンポテンシャル  $\phi_c$  は、長距離相互作用であるため、系内のイオン同士およびイメージセルのイオンとの和として表される

$$\phi_c = \frac{e^2}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\nu} \frac{Z_i Z_j}{4\pi\epsilon_0 |L\nu + \mathbf{r}_j - \mathbf{r}_i|}. \quad (2.15)$$

ベクトル  $\nu = (\nu_1, \nu_2, \nu_3)$ ,  $\nu_i = 0, \pm 1, \pm 2, \dots$  ( $i = 1, 2, 3$ ) は基本セルとすべてのイメージセルの位置を表し、 $\sum_{\nu}$  はすべてのセルの和をとることを示す。ただし、 $\nu = (0, 0, 0)$  のとき（基本セルにおいて）、 $j = i$  の項は除く。 $\epsilon_0$  は真空の誘電率である。

次に、系が熱膨張（系の体積が変化）する場合について、考えてみる。基本セルの一辺を  $L$  として、イオンの位置  $\mathbf{r}$  を、

$$\mathbf{r} = L\mathbf{r}^* = (Lx^*, Ly^*, Lz^*), \quad 0 \leq x^*, y^*, z^* \leq 1 \quad (2.16)$$

のように、無次元ベクトル  $\mathbf{r}^*$  に変換する。これにより、スケールされた系の体積は 1 となる。クーロンポテンシャル  $\phi_c$  は

$$\phi_c = \frac{e^2}{4\pi\epsilon_0 L} \phi_c^* \quad (2.17)$$

$$\phi_c^* = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\nu} \frac{Z_i Z_j}{|\nu + \mathbf{r}_j^* - \mathbf{r}_i^*|} \quad (2.18)$$

と表すことができる。式を見ると、クーロンポテンシャル  $\phi_c^*$  は体積  $L$  によらないことがわかる。もし、体積が変わる場合は、式 (2.17) の  $1/L$  という係数のみが影響を受けることとなる。 $\phi_c^*$  の式は 0 への収束が遅くなるため、エワルドの方法を用いて導くと、次のような公式としてまとめることができる。

$$\phi_c = \frac{e^2}{4\pi\epsilon_0 L} \Phi_c^* = \frac{e^2}{4\pi\epsilon_0 L} (\phi_c^{*(1)} + \phi_c^{*(2)} + \phi_c^{*(3)}) \quad (2.19)$$

$$\phi_c^{*(1)} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\nu^*} \frac{Z_i Z_j}{|\nu + \mathbf{r}_j^* - \mathbf{r}_i^*|} \operatorname{erfc}[\chi |\nu + \mathbf{r}_j^* - \mathbf{r}_i^*|] \quad (2.20)$$

$$\begin{aligned} \phi_c^{*(2)} &= \frac{1}{\pi} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{h}^*} \frac{Z_i Z_j}{|\mathbf{h}^*|^2} \exp\left[-\frac{\pi^2 |\mathbf{h}^*|^2}{\chi^2}\right] \\ &\quad \times \cos[2\pi \mathbf{h}^* \cdot (\mathbf{r}_j^* - \mathbf{r}_i^*)] \end{aligned} \quad (2.21)$$

$$\begin{aligned} &= \frac{1}{\pi} \sum_{\mathbf{h}^*} \frac{1}{|\mathbf{h}^*|^2} \exp\left[-\frac{\pi^2 |\mathbf{h}^*|^2}{\chi^2}\right] \\ &\quad \times \left\{ \left[ \sum_{i=1}^N Z_i \cos 2\pi \mathbf{h}^* \cdot \mathbf{r}_i^* \right]^2 + \left[ \sum_{i=1}^N Z_i \sin 2\pi \mathbf{h}^* \cdot \mathbf{r}_i^* \right]^2 \right\} \end{aligned} \quad (2.22)$$

$$\phi_c^{*(3)} = -\frac{\chi}{\sqrt{\pi}} \sum_{i=1}^N Z_i^2 \quad (2.23)$$

$\phi_c^{*(1)}$  は直接項、 $\phi_c^{*(2)}$  は逆格子空間の項、 $\phi_c^{*(3)}$  は定数項である。 $\phi_c^{*(2)}$  の計算は基本セルも含まれており、重複を防ぐために  $\phi_c^{*(3)}$  の項が導入されている。式の  $\operatorname{erfc}$  は補誤差関数 (complementary error function) である。このとき、パラメーター  $\chi$  は  $0 < \chi < \infty$  の範囲になっている。このパラメーターは電荷の広がりを表している。補誤差関数  $\operatorname{erfc}$  は誤差関数  $\operatorname{erf}$  と次のような関係となっている。

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt \quad (2.24)$$

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt \quad (2.25)$$

補誤差関数は  $x$  が 1 に近づくほど減少していき、2 でほぼゼロになる関数である。  $\mathbf{h}$  は逆格子ベクトルであり、  $\nu$  と同様に

$$\mathbf{h} = (h_1, h_2, h_3), \quad h_i = 0, \pm 1, \pm 2, \dots \quad (2.26)$$

と定義される。  $\mathbf{h} = (0, 0, 0)$  の項は、総電荷がゼロである条件式により除かれる。

誤差関数と補誤差関数による、ポテンシャルの分解の概要を図 2.2.2 に示す。図 2.2.2 から、距離に反比例するポテンシャルの収束は遅いことがわかる。補誤差関数において、距離が 1 に近づくにつれ急激に減少していることがわかる。これは  $\phi_c^{*(1)}$  に対応している。誤差関数は  $x$  の増加するにつれ、ゆっくりと減衰していく関数であり、  $x \rightarrow 0$  の極限において、  $2/\sqrt{\pi}$  となる。これが  $\phi_c^{*(2)}$  に対応することから、  $\mathbf{h}$  に依存してポテンシャルが寄与し、  $\mathbf{h}$  の増加とともに減衰していくことがわかる。

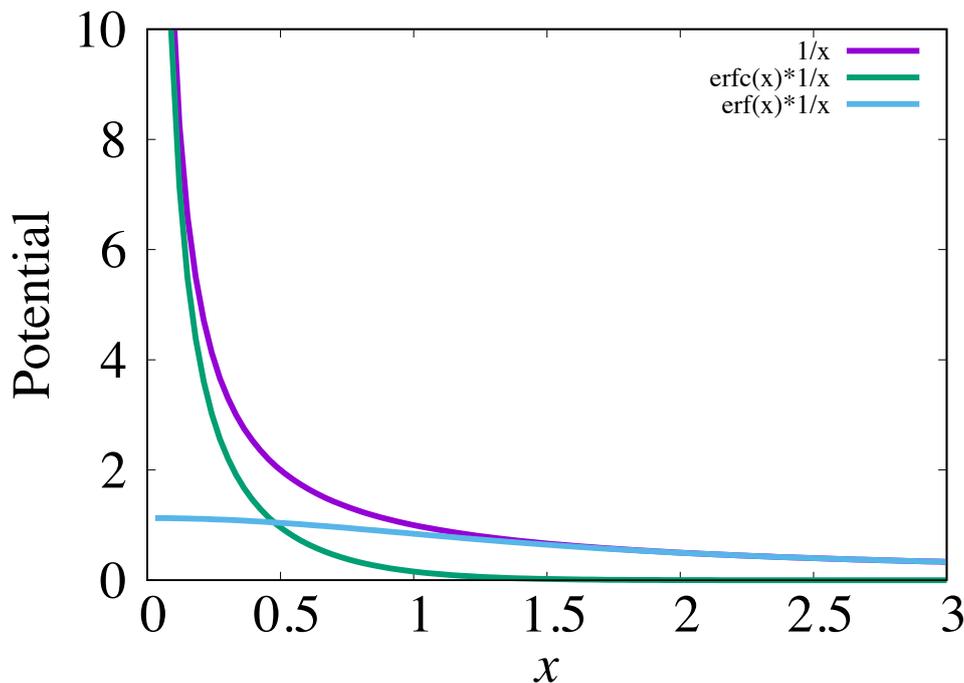


図 2.2.2 距離に反比例するポテンシャルの誤差関数と補誤差関数による分解。紫色のポテンシャル曲線は距離に反比例するポテンシャルの減衰の様子を示す。緑と青色は、誤差関数と補誤差関数を用いて分解されたポテンシャルの減衰の様子である。

### 2.3.4. パーティクルメッシュエワルド法 (PME 法)

通常のエワルド法は無限遠までの静電相互作用を、現実的な計算時間で計算可能である。しかし、系の粒子数  $N$  に対し  $N^{2/3}$  に比例するため、大規模計算には向かない。

パーティクルメッシュエワルド法 (PME 法) は逆格子空間の計算に高速フーリエ変換を用いる相互作用計算の高速化技術であり, 本研究ではこの方法を用いて計算を行った.

PME 法を用いることにより,  $N \log N$  に比例する計算量に押さえることが可能となる. PME 法では, 式 (2.21) を, 電荷分布を格子点上に分ける形にし, 高速フーリエ変換を行うことで, 計算の高速化を実現している.

よって,  $\phi_c^{*(2)}$  の式は次のようになる.

$$\phi_c^{*(2)} = \frac{1}{\pi} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{h}^*} \frac{Z_i Z_j}{|\mathbf{h}^*|^2} \exp \left[ -\frac{\pi^2 |\mathbf{h}^*|^2}{\chi^2} \right] \times \cos [2\pi \mathbf{h}^* \cdot (\mathbf{r}_j^* - \mathbf{r}_i^*)] \quad (2.27)$$

$$= \frac{1}{\pi} \sum_{\mathbf{h}^*} \frac{1}{|\mathbf{h}^*|^2} \exp \left[ -\frac{\pi^2 |\mathbf{h}^*|^2}{\chi^2} \right] \times \sum_{i=1}^N Z_i \exp(i\mathbf{h}^* \cdot \mathbf{r}_i^*) \sum_{j=1}^N Z_j \exp(-i\mathbf{h}^* \cdot \mathbf{r}_j^*) \quad (2.28)$$

$$= \frac{1}{\pi} \sum_{\mathbf{h}^*} \frac{1}{|\mathbf{h}^*|^2} \exp \left[ -\frac{\pi^2 |\mathbf{h}^*|^2}{\chi^2} \right] S(\mathbf{h}^*) S(-\mathbf{h}^*) \quad (2.29)$$

式 (2.29) の  $S(\mathbf{h}^*) = \sum_{i=1}^N Z_i \exp(i\mathbf{h}^* \cdot \mathbf{r}_i^*)$  は構造因子と呼ばれ, 電荷密度のフーリエ変換に対応する.

## 2.4. 差分近似

数値計算を行う上で、コンピュータ上での計算は欠かせない。コンピュータでは微分を直接計算することはできない。計算を行うためには、微分の式を差分化の形にする必要がある。差分化の簡単な例は、高校で習う微分の定義で極限をとる前の形である。微分方程式を差分商の形にしたものを差分方程式という。ここでは、分子動力学法においてよく使われる、ベルレ法および蛙飛び法を紹介する。本研究において、レプリカ交換法においては速度ベルレ法、通常の分子動力学法において蛙飛び法を用いた。蛙飛び法と速度ベルレ法は計算が簡単であり、計算誤差が小さく安定して微分方程式の解を計算できる方法である。レプリカ交換法においては温度制御の計算をより精確なものとするために、より精確に計算できる速度ベルレ法を選んだ。

### 2.4.1. ベルレ法

時刻  $t$  から  $\Delta t$  後 ( $t + \Delta t$ ) と  $\Delta t$  前 ( $t - \Delta t$ ) の粒子  $i$  の位置  $\mathbf{r}_i(t - \Delta t)$  を、時刻 ( $t = t$ ) のまわりで  $\Delta t$  に関するテイラー展開をする。

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \frac{d\mathbf{r}_i(t)}{dt} + \frac{\Delta t^2}{2!} \frac{d^2\mathbf{r}_i(t)}{dt^2} + \frac{\Delta t^3}{3!} \frac{d^3\mathbf{r}_i(t)}{dt^3} + O(\Delta t^4) \quad (2.30)$$

$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \frac{d\mathbf{r}_i(t)}{dt} + \frac{\Delta t^2}{2!} \frac{d^2\mathbf{r}_i(t)}{dt^2} - \frac{\Delta t^3}{3!} \frac{d^3\mathbf{r}_i(t)}{dt^3} + O(\Delta t^4) \quad (2.31)$$

両辺を足し合わせると、

$$\mathbf{r}_i(t + \Delta t) + \mathbf{r}_i(t - \Delta t) = 2\mathbf{r}_i(t) + \Delta t^2 \frac{d^2\mathbf{r}_i(t)}{dt^2} + O(\Delta t^4) \quad (2.32)$$

となる。得られた式に、式 (2.1) を代入すると、

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i(t) + O(\Delta t^4) \quad (2.33)$$

と書ける。これは粒子  $i$  の時刻  $t$  での位置と力、および時刻  $t - \Delta t$  での位置を用いて、時刻  $t + \Delta t$  での位置が決定することを表す。 $O(\Delta t^4)$  は  $\Delta t$  の4次以上の項を示している。また、粒子の速度は、式 (2.30) から式 (2.31) を差し引き、

$$\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t) = 2\Delta t \frac{d\mathbf{r}_i(t)}{dt} + O(\Delta t^3) \quad (2.34)$$

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t) = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t} + O(\Delta t^2) \quad (2.35)$$

と表すことができる. 式 (2.35) において,  $O(\Delta t^2)$  の誤差が含まれており,  $\Delta t^3$  を含む  $d^3\mathbf{r}_i(t)/dt^3$  の項は  $\Delta t$  の割り算により,  $O(\Delta t^2)$  の誤差となっている.

実際のプログラムにおいて, 式 (2.33), 式 (2.35) はそのまま用いない. 桁落ちを防ぐために, 速度ベルレ法と呼ばれる形式を用いる. 式 (2.33), 式 (2.34) を変形すると,

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\Delta t \mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)}{m_i} + O(\Delta t^3) \quad (2.36)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i(t) + O(\Delta t^4) \quad (2.37)$$

となる.

## 2.4.2. 蛙飛び法

ニュートンの運動方程式を数値積分する方法として, 蛙飛び法 (leap frog integrator) と呼ばれる方法がある. これは, ベルレ法と同様に運動方程式を差分することにより, 位置と速度を求める方法である. 位置と速度の時間軸は,  $\Delta t/2$  だけずれており, 時刻  $t$  における力  $\mathbf{F}(t)$  を決定するために, 蛙飛び法は時刻  $t$  における位置  $\mathbf{r}$  と時刻  $t - \Delta t/2$  における速度  $\mathbf{v}$  を用いる.  $t + \Delta t/2$  における速度と位置は

$$\mathbf{v}\left(t + \frac{1}{2}\Delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m} \mathbf{F}(t) + O(\Delta t^3) \quad (2.38)$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}\left(t + \frac{1}{2}\Delta t\right) + O(\Delta t^3) \quad (2.39)$$

と表せる. このアルゴリズムは蛙のように跳ぶことから蛙飛び法と呼ばれている.

## 2.5. 温度制御および圧力制御法

本節では温度制御および圧力制御法について説明する。ある実験系を考えた場合、多くは決められた温度、圧力または体積などの条件下で行なわれている。シミュレーションを用いてその実験系を再現する際、温度や圧力の制御が必要となってくる。ここでは、カノニカル ( $NVT$ ) アンサンブルや  $NPT$  アンサンブルにおけるシミュレーションを行うために、温度制御法として速度スケーリング法[60]、能勢の方法[61]、能勢・フーバーの方法[62]、圧力制御法としてパリネロ・ラーマンの方法[63]を紹介する。本研究において、レプリカ交換法では、温度制御をより精確に制御できる能勢・フーバーの方法、圧力制御をパリネロ・ラーマンの方法を用いた。通常の分子動力学法では温度制御を速度スケーリング法、圧力制御をパリネロ・ラーマンの方法を用いた。速度スケーリングは最も簡便な温度制御法であるが、本研究において計算に用いる GROMACS では速度スケーリングを改良した形で導入されており、温度制御の精度は十分である。

### 2.5.1. 速度スケーリング法

速度スケーリング法は温度制御法の中で最も単純な方法である。速度スケーリング法における温度の制御は、式 (2.5) の等分配則に基づいて行われる。蛙飛び法を例に用いると、 $t + \Delta t/2$  での運動エネルギー  $K$  は

$$K = \frac{1}{2} \left\langle \sum_{i=1}^N m_i \mathbf{v}_i^2 \left( t + \frac{1}{2} \Delta t \right) \right\rangle = \frac{3}{2} N k_B T \quad (2.40)$$

となる。 $K$  は  $t + \Delta t/2$  後の温度  $T$  に対応し、この温度が指定した温度  $T_t$  になるように速度を修正 (スケール) する必要がある。そこでスケール因子  $s$  を用いて、

$$\mathbf{v}_{ti} \left( t + \frac{1}{2} \Delta t \right) = s \mathbf{v}_i \left( t + \frac{1}{2} \Delta t \right) \quad (2.41)$$

速度をスケールする。運動エネルギーも等しいことから、

$$\frac{1}{2} \sum_{i=1}^N m_i \mathbf{v}_{ti}^2 \left( t + \frac{1}{2} \Delta t \right) = \frac{s^2}{2} \sum_{i=1}^N m_i \mathbf{v}_i^2 \left( t + \frac{1}{2} \Delta t \right) \quad (2.42)$$

となり、等分配則より、

$$\begin{aligned}\frac{3}{2}Nk_{\text{B}}T_t &= s^2\frac{3}{2}Nk_{\text{B}}T \\ s &= \sqrt{\frac{T_t}{T}}\end{aligned}\tag{2.43}$$

スケール因子  $s$  を得ることができる。この  $s$  を用いて式 (2.41) をスケールする。GROMACS では、この速度スケーリングをブッシらが修正した形で導入しており、本質的には Berendsen の方法を拡張した形となっている。まず、スケール因子を

$$s = \sqrt{\frac{K_t}{K}}\tag{2.44}$$

という形でスケール因子を導入している。 $K_t$  は指定したい運動エネルギーにおける平衡状態のカノニカル分布により得られる。

$$\bar{P}(K_t)dK_t \propto K_t^{(N_f/2-1)}\exp(-\beta K_t)dK_t\tag{2.45}$$

$$dK = (\bar{K} - K)\frac{dt}{\tau} + 2\sqrt{\frac{K\bar{K}}{N_f}}\frac{dW}{\sqrt{\tau}}\tag{2.46}$$

$\tau$  は任意のパラメーターであり、時間の次元をもつ。 $N_f$  は自由度の数であり、 $dW$  はウィーナー過程である。ウィーナー過程はブラウン運動を確率で表現した確率過程のことを言う。また、理想的なランダムウォークがブラウン運動である。

## 2.5.2. 能勢の方法

速度スケーリング法による温度制御は、スケール因子を用いて指定した温度になるよう制御していた。能勢の方法では、実際の実験系のように系を熱浴に入れエネルギーのやりとりをすることで、温度制御を行う。そこで、分子動力学法においては、熱浴とのエネルギーのやりとりがある仮想系を考え、仮想系におけるハミルトニアン  $\mathcal{H}'(\mathbf{p}', \mathbf{r}', p_s, s)$  を導入する。仮想系におけるハミルトニアンにより拡張された系の粒子の座標と正準共役な運動量にはプライムで表し、 $s$  を導入した座標の自由度とすると、それぞれ、次のような関係式が成り立つと仮定する。

$$\mathbf{r}_i = \mathbf{r}'_i, \quad \mathbf{p}_i = \frac{\mathbf{p}'_i}{s},\tag{2.47}$$

ここで、導入した座標  $s$  のポテンシャルエネルギー  $V$  は  $s$  の関数として表すと、

$$V(s) = gk_B T \log s \quad (2.48)$$

となる。温度  $T$  は指定する温度であり、パラメーターとして与えられる、そして、 $g$  は系の自由度である。よって、仮想系におけるハミルトニアン  $\mathcal{H}'$  は

$$\mathcal{H}'(\mathbf{p}', \mathbf{r}', p_s, s) = \sum_i \frac{\mathbf{p}'_i{}^2}{2ms^2} + U(\mathbf{r}') + \frac{p_s^2}{2Q} + gk_B T \log s \quad (2.49)$$

と表すことができる。  $p_s$  は導入した座標  $s$  の正準共役な運動量であり、 $Q$  はその質量である。  $g$  は自由度であり、自由度が  $3N$  の時、カノニカル分布におけるアンサンブル平均と一致する。式 (2.49) において、系または熱浴から与えられるポテンシャルエネルギーは、次式としてみる事ができる。

$$W(s) = \sum_i \frac{\mathbf{p}'_i{}^2}{2ms^2} + 3Nk_B T \log s \quad (2.50)$$

自由度  $g$  は  $3N$  とした。ここで、現実系と仮想系との間に次の関係があると仮定する。

$$t = \int_0^t \frac{dt'}{s}, \quad dt = \frac{dt'}{s} \quad (2.51)$$

これは、分子動力学法のような小さな系で熱浴とのやりとりをすると、時間依存量に大きな影響を与えてしまうので、その調整のために導入している。式 (2.47), 式 (2.51) の関係式より速度は

$$\frac{d\mathbf{r}_i}{dt} = \frac{d\mathbf{r}_i}{dt'} \frac{dt'}{dt} = s \frac{d\mathbf{r}'_i}{dt'} \quad (2.52)$$

と表せる。仮想系の正準方程式は、式 (2.49) より得られる。

$$\frac{d\mathbf{r}'_i}{dt'} = \frac{\partial \mathcal{H}'}{\partial \mathbf{p}'_i} = \frac{\mathbf{p}'_i}{ms^2} \quad (2.53)$$

$$\frac{d\mathbf{p}'_i}{dt'} = -\frac{\partial \mathcal{H}'}{\partial \mathbf{r}'_i} = -\frac{\partial U}{\partial \mathbf{r}'_i} \quad (2.54)$$

$$\frac{ds}{dt'} = \frac{\partial \mathcal{H}'}{\partial p_s} = \frac{p_s}{Q} \quad (2.55)$$

$$\frac{dp_s}{dt'} = -\frac{\partial \mathcal{H}'}{\partial s} = -\frac{\partial W(s)}{\partial s} = \sum_i \frac{\mathbf{p}'_i{}^2}{ms^3} - \frac{3Nk_B T}{s} \quad (2.56)$$

この得られた方程式を、現実系における方程式に書き直すと ( $s\mathbf{p}_i = \mathbf{p}'_i, sdt = dt'$  の点に注意する) ,

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m} \quad (2.57)$$

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial U}{\partial \mathbf{r}_i} - \mathbf{p}_i \frac{p_s}{Q} \quad (2.58)$$

$$\frac{ds}{dt} = s \frac{p_s}{Q} \quad (2.59)$$

$$\frac{dp_s}{dt} = \sum_i \frac{\mathbf{p}_i^2}{m} - 3Nk_B T \quad (2.60)$$

この運動方程式が、能勢の方法における運動方程式である。

### 2.5.3. 能勢・フーバーの方法

能勢・フーバーの方法は、能勢の運動方程式を少し変形したものである。フーバーは、現実系における方程式に書き直す際、

$$\zeta \equiv \frac{1}{s} \frac{ds}{dt} = \frac{ds}{dt'} = \frac{p_s}{Q} \quad (2.61)$$

とし、

$$\frac{d\zeta}{dt} = \frac{1}{s} \frac{d}{dt} \left( \frac{ds}{dt} \right) = \frac{1}{s} \frac{d}{dt} \left( s \frac{p_s}{Q} \right) = \frac{1}{Q} \frac{dp_s}{dt} \quad (2.62)$$

$$\frac{d\zeta}{dt} = \frac{1}{Q} \left( \sum_i \frac{\mathbf{p}_i^2}{m} - 3Nk_B T \right) \quad (2.63)$$

とした。よって、運動方程式は次のようになる。

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m} \quad (2.64)$$

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial U}{\partial \mathbf{r}_i} - \zeta \mathbf{p}_i \quad (2.65)$$

$$\frac{d\zeta}{dt} = \frac{1}{Q} \left( \sum_i \frac{\mathbf{p}_i^2}{m} - 3Nk_B T \right) \quad (2.66)$$

これが能勢・フーバーの方法による運動方程式となる。

#### 2.5.4. パリネロ・ラーマンの方法

パリネロ・ラーマンの方法は圧力を制御するアルゴリズムの一つである。平行六面体セルの各辺をベクトル  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  とし、行列を用いて表記すると、

$$L = (\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}) = \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} \quad (2.67)$$

となり、粒子の座標  $\mathbf{r}_i$  を  $L$  とセル内部における座標  $\mathbf{q}_i$  を用いて、

$$\mathbf{r}_i = L\mathbf{q}_i \quad 0 \leq q_{ia}, q_{ib}, q_{ic} \leq 1 \quad (2.68)$$

と表す。速度は、

$$\frac{d\mathbf{r}_i}{dt} = L \frac{d\mathbf{q}_i}{dt} \quad (2.69)$$

となる。ここで計量テンソル  $\mathbf{G}$  を導入する。

$$\mathbf{G} = L^T L = \begin{pmatrix} \mathbf{a}^T \\ \mathbf{b}^T \\ \mathbf{c}^T \end{pmatrix} (\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}) \quad (2.70)$$

ここで、 $T$  は転置である。

$$\mathbf{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}, \mathbf{a}^T = (a_x \quad a_y \quad a_z) \quad (2.71)$$

$$= \begin{pmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{a} \cdot \mathbf{b} & \mathbf{a} \cdot \mathbf{c} \\ \mathbf{b} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{b} & \mathbf{b} \cdot \mathbf{c} \\ \mathbf{c} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{b} & \mathbf{c} \cdot \mathbf{c} \end{pmatrix} \quad (2.72)$$

計量テンソルを用いて、速度の二乗を表すと、

$$\left(\frac{d\mathbf{r}_i}{dt}\right)^2 = \frac{d\mathbf{r}_i^T}{dt} \frac{d\mathbf{r}_i}{dt} = \left(L \frac{d\mathbf{q}_i}{dt}\right)^T L \frac{d\mathbf{q}_i}{dt} = \frac{d\mathbf{q}_i^T}{dt} L^T L \frac{d\mathbf{q}_i}{dt} = \frac{d\mathbf{q}_i^T}{dt} \mathbf{G} \frac{d\mathbf{q}_i}{dt} \quad (2.73)$$

となる。これらを用いて、拡張したラグランジアン  $\mathcal{L}$  を

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N m_i \frac{d\mathbf{r}_i^T}{dt} \mathbf{G} \frac{d\mathbf{q}_i}{dt} - U(\mathbf{q}_i, L) + \frac{1}{2} M \text{Tr} \left( \frac{dL^T}{dt} \frac{dL}{dt} \right) - P_0 V \quad (2.74)$$

と導入する。これを解くことにより、以下の運動方程式が得られる。

$$\frac{d^2 L}{dt^2} = \frac{1}{M} \left[ \frac{1}{V} \left\{ \sum_{i=1}^N m_i \left( L \frac{d\mathbf{q}_i}{dt} \right) \left( L \frac{d\mathbf{q}_i}{dt} \right)^T + \sum_{i=1}^N \mathbf{F}_i \mathbf{r}_i^T \right\} - P_0 I \right] \sigma \quad (2.75)$$

ここで、 $I$  は 3 行 3 列の単位行列であり、 $\sigma$  は  $L$  の逆行列を求める際の余因子行列  $\tilde{L}$  における行列である。

$$\begin{aligned} L^{-1} &= \frac{1}{|L|} (\tilde{L}) \\ L^{-1} &= \frac{1}{V} (\tilde{L}) \\ &= \frac{1}{V} (\mathbf{b} \times \mathbf{c} \quad \mathbf{c} \times \mathbf{a} \quad \mathbf{a} \times \mathbf{b})^T = \frac{1}{V} \sigma^T \end{aligned} \quad (2.76)$$

$$\sigma = (\mathbf{b} \times \mathbf{c} \quad \mathbf{c} \times \mathbf{a} \quad \mathbf{a} \times \mathbf{b})$$

平行六面体においては、行列式はその体積になるので、 $V$  とした。

## 2.6. 粒子の束縛

分子動力学法では、粒子それぞれに運動方程式を与え、粒子の運動の時間発展を追う。タンパク質内における原子間距離や結合角は決まっており、それに沿うように束縛条件付きで計算を行う必要がある。タンパク質のシミュレーションにおいて、水素原子は高速に振動しているため、前述した差分近似における時間刻み  $\Delta t$  を細くとる必要がある。しかし、時間刻みを細くするほど、時間発展を追う上で計算回数が増え、長時間の計算を必要とする。また、計算回数が増えるため丸め誤差が積み重なり、かえって最終の計算結果の誤差が大きくなる可能性がある。これを解決するために、拘束を用いて水素原子を含む結合の速い振動による粒子位置の変化について、平均的な位置で代表させることをする。これにより高速振動を考えなくて良くなるため、 $\Delta t$  を水素原子の振動周期より長い時間に設定できるようになる。代表的な拘束法として、SHAKE 法や LINCS (Linear constraint solver) 法が挙げられる。

### 2.6.1. SHAKE 法

SHAKE はよく用いられる拘束法の 1 つであり、原子間距離を拘束する方法である。SHAKE の概要を図 2.6.1 に示しているが、SHAKE は原子 AB 間、AD 間の共有結合長の距離や破線で示した BD 間や CD 間の距離を拘束する。拘束の対象は水を含めた系全ての分子となっている。

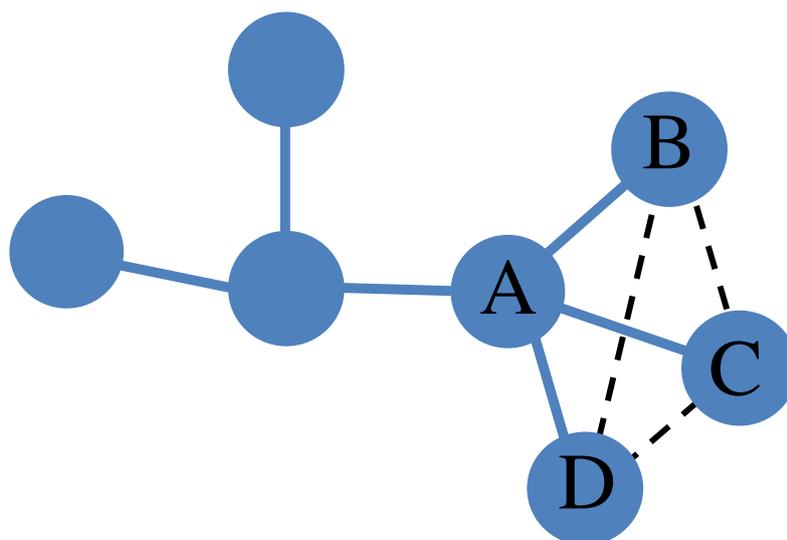


図 2.6.1 SHAKE に用いる原子の例。破線は拘束する原子間ペアを示す。

## 2.6.2. SHAKE 条件下の運動方程式

SHAKE による距離を拘束するために、運動方程式に拘束力を加える必要がある。図 2.6.2 に示すような、原子 AB 間の距離を  $d$  として、それが時刻  $t + \Delta t$  後もその距離を維持するための運動方程式を導く。まず、拘束条件は、

$$\phi = (\mathbf{r}_A - \mathbf{r}_B)^2 - d^2 = 0 \quad (2.77)$$

となる。 $\mathbf{r}_A, \mathbf{r}_B$  は原子 A と B の位置に対応する。この拘束条件を  $\mathbf{r}_A, \mathbf{r}_B$  それぞれで偏微分を行うと

$$\frac{\partial}{\partial \mathbf{r}_A} \phi = 2(\mathbf{r}_A - \mathbf{r}_B) \quad (2.78)$$

$$\frac{\partial}{\partial \mathbf{r}_B} \phi = -2(\mathbf{r}_A - \mathbf{r}_B) \quad (2.79)$$

となり、拘束力は  $(\mathbf{r}_A - \mathbf{r}_B)$  に比例する関数として表される。よって、原子 A, B 間に働く拘束力は、比例係数  $\lambda_A, \lambda_B$  を用いて、 $2\lambda_A(\mathbf{r}_A - \mathbf{r}_B), -2\lambda_B(\mathbf{r}_A - \mathbf{r}_B)$  となる。この  $\lambda_A, \lambda_B$  をうまく決めることで、原子間距離を保つことができる。そして、拘束力を運動方程式 (2.1) に加えると

$$\mathbf{F}_A + 2\lambda_A(\mathbf{r}_A - \mathbf{r}_B) = m_A \frac{d^2 \mathbf{r}_A}{dt^2} \quad (2.80)$$

$$\mathbf{F}_B - 2\lambda_B(\mathbf{r}_A - \mathbf{r}_B) = m_B \frac{d^2 \mathbf{r}_B}{dt^2} \quad (2.81)$$

となる。あとは、前述した差分近似法を用いて、時間発展を追っていき、ラグランジュ未定数法を用いて、 $\lambda_A, \lambda_B$  を決定すれば良い。

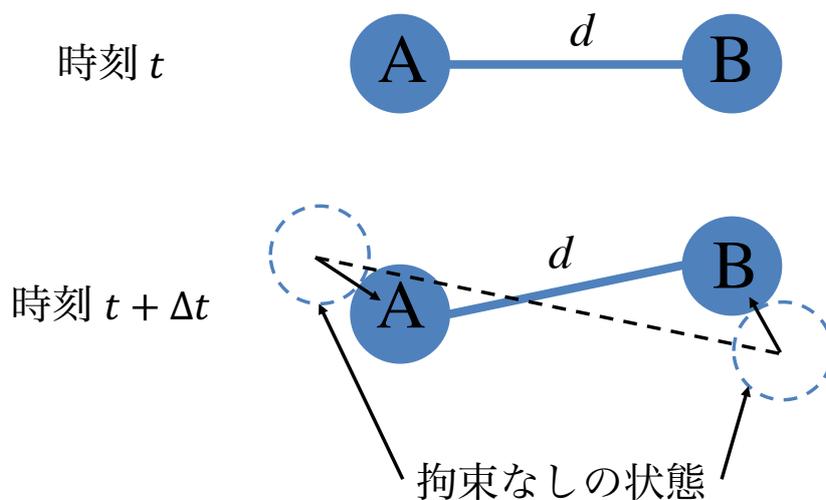


図 2.6.2 SHAKE による位置拘束の概要

### 2.6.3. Linear constraint solver (LINCS)

LINCS は拘束法の 1 つであり、この方法は非拘束条件下で 1 ステップ計算を行い、その後正しい距離に修正するという方法である。GROMACS ではこの拘束法がデフォルト設定になっている。LINCS は SHAKE アルゴリズムに比べ、速く安定しているという利点があり、GROMACS が速いと言われるのは、LINCS アルゴリズムを用いている点が挙げられる。LINCS は結合数が少ない分子において適切に働くアルゴリズムである。LINCS はヤコビ行列の逆数を近似する形で処理が行われており、ヤコビ行列の固有値が 1 より小さくなるものにも適用されるアルゴリズムである。

図 2.6.3 に LINCS の概要を示した。射影した後の結合長は  $l = d \cos \theta$  となるが、これは実際の結合長ではないので、次のステップ (回転) でその修正を行う。修正 (回転) を行うとき、結合の射影は  $p = (2d^2 - l^2)^{1/2}$  となる。

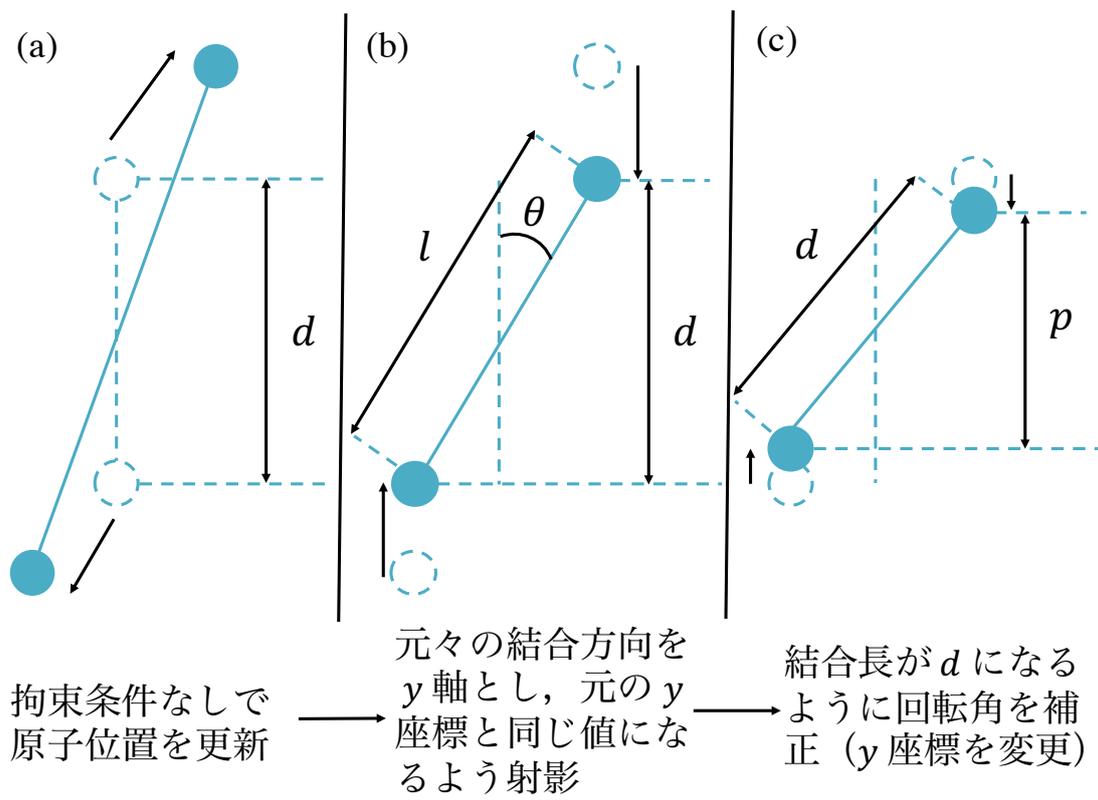


図 2.6.3 LINCS による位置拘束の概要

## 2.7. 分子動力学シミュレーションにおいてよく用いられる解析

本節では、シミュレーションの解析によく用いられる手法や統計量を紹介する。

### 2.7.1. 根平均二乗変位 (Root Mean Square Deviation: RMSD)

シミュレーション時間内において構造全体がどの程度変化したかを知ることができ量として RMSD がある。RMSD は対応する 2 点の位置それぞれの差を二乗し、その相加平均の平方根として表されている。シミュレーション時刻  $t_{\text{ref}}$  での参照構造とシミュレーション時刻  $t$  での構造の間での構造全体のずれを数値化したものが RMSD である。ここでいう参照構造とは、変化の基準となる構造で、結晶構造の座標 (原子位置) もしくは、緩和を行った後の座標を用いることが多い。

$$RMSD(t, t_{\text{ref}}) = \left[ \frac{1}{M} \sum_{i=1}^N m_i \{ \mathbf{r}_i(t) - \mathbf{r}_i(t_{\text{ref}}) \}^2 \right]^{\frac{1}{2}} \quad (2.82)$$

$$M = \sum_{i=1}^N m_i$$

$M$  は全質量、 $m_i$  は  $i$  番目の原子の質量、 $\mathbf{r}_i$  は原子位置を表す。 $t$  と  $t_{\text{ref}}$  の関数として RMSD が決定される。本研究では、基準となる構造との RMSD が最小となるように、X 線構造解析により決定された結晶構造全体を回転させてフィッティングしている。

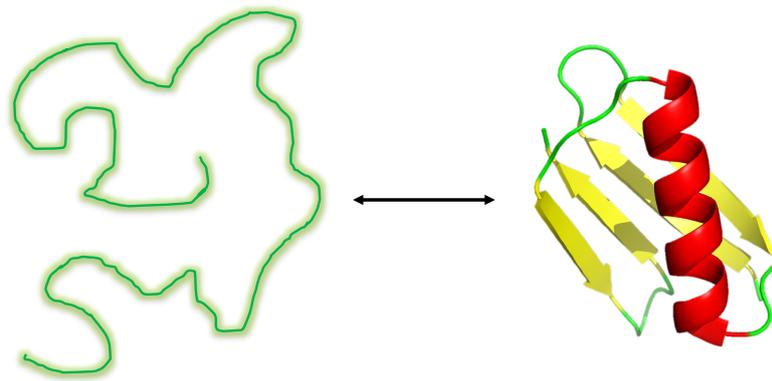
### 2.7.2. 慣性半径 (Radius of gyration: $R_g$ )

慣性半径  $R_g$  は、構造の広がり具合を表す。例えば、タンパク質はフォールディング状態では、折り畳まれコンパクトな構造を持つが、アンフォールディング状態 (ほどけた状態) になると、図 2.7.1 に示すような長い紐状に広がった形になる。 $R_g$  はこのような状態変化の指標となる量である。

$$R_g = \left[ \frac{1}{M} \sum_{i=1}^N m_i \{ \mathbf{r}_i(t) - \mathbf{r}_G(t) \}^2 \right]^{\frac{1}{2}} \quad (2.83)$$

$$M = \sum_{i=1}^N m_i$$

$\mathbf{r}_G(t)$  はある時刻  $t$  における構造の重心位置を表す。



アンフォールディング状態

フォールディング状態

図 2.7.1 タンパク質のフォールディング-アンフォールディング

### 2.7.3. 根平均二乗揺らぎ (Root Mean Square Fluctuation: RMSF)

根平均二乗揺らぎ (RMSF) は構造の動的な指標の一つとして役に立つ。例えば、計算する対象をアミノ酸の  $C_\alpha$  (不斉炭素原子) とすれば、タンパク質のどの残基が揺らぎやすいのかを調べることができる。

$$\text{RMSF}_i = \left[ \frac{1}{T - T_0} \int_{T_0}^T \{ \mathbf{r}_i(t) - \langle \mathbf{r}_i \rangle_t \}^2 dt \right]^{\frac{1}{2}} \quad i = 1, 2, \dots, N \quad (2.84)$$

$\langle \mathbf{r}_i \rangle_{T_0, T}$  は  $i$  番目の原子の時刻  $t = T_0 \sim t = T$  の時間の平均位置である。この式を用いて対象としたい各原子の数だけそれぞれで計算を行う。

## 2.8. 自由エネルギー

自由エネルギーは、ある状態を表す量として知られる。自由エネルギーは乱雑さ、状態数であるエントロピーと結び付けられている。エントロピーを説明するための例えとして、ブラックコーヒーにミルクを加えることが挙げられる。ブラックコーヒーにミルクを加えると、ミルクが次第に広がり混ざり合う。ミルクを入れた瞬間ある所にミルクが集中（状態数が少ない）しているが、時間が経つにつれてミルクが広がり混ざり合う。この広がり、混ざり合うことがエントロピーの増大と言える。つまり、この状態の変化が自由エネルギーの変化に対応する。本節では、この自由エネルギーとタンパク質の関係について述べる。

### 2.8.1. 自由エネルギーとタンパク質

自然界において、タンパク質は様々な構造を取っており、それぞれの構造の状態を自由エネルギーで表すことができる。（ヘルムホルツ）自由エネルギーは内部エネルギーからエントロピーと温度の積を引いたものである。タンパク質における内部エネルギーは、分子構造の揺らぎによる運動エネルギーとボンド間の結合エネルギー（主にバネポテンシャル）、角度ポテンシャル、二面角のエネルギーなどの和である。エントロピーは乱雑さ（もしくは状態数）を表す量であり、自然界において増大する方向に進む。よって、熱力学第二法則に従えば、自由エネルギーは減少する方向に進む。自然界におけるタンパク質の構造は平衡状態にあると考えられており[64]、それは自由エネルギーが最小となる構造である。このことから、自由エネルギーの変化は構造の変化と結び付いていると言える。タンパク質の構造は熱的に揺らぎ、その自由エネルギーも同様にゆらぐが、その中で自由エネルギーの最も低い構造を最安定構造と言う。この最安定構造におけるタンパク質の自由エネルギーと、その周辺を揺らいでいる構造の自由エネルギーとの差が大きい場合、構造は最安定構造に向かって変化しやすくなる。よって自由エネルギー差は、ある温度におけるタンパク質の構造変化のしやすさの指標となる。

以上のことから、自由エネルギーは構造安定性の基準としてみることができ、常温において天然の構造が最も安定なのは、自由エネルギーが最も低いからであると言い換えることができる。例えば、図 2.8.1 に示すように、タンパク質がアンフォールディング状態にあるとき、タンパク質の構造はより安定であるような、自由エネルギーの低い方へとシフトしていくと考えられる。

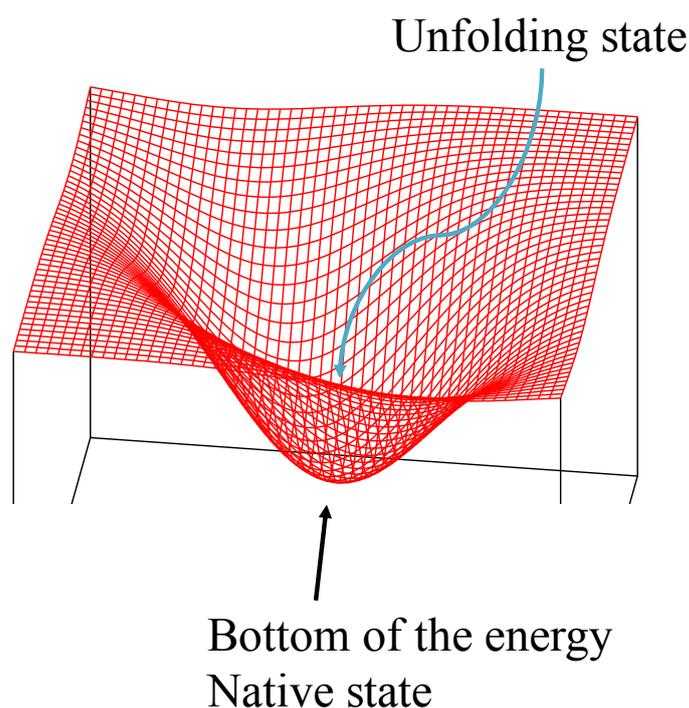


図 2.8.1 タンパク質の自由エネルギー地形の概要

## 2.8.2. 統計力学における自由エネルギー

自由エネルギーは熱力学における状態量であり、ヘルムホルツの自由エネルギーとギブスの自由エネルギーがある。ヘルムホルツの自由エネルギーは等温等積過程における自由エネルギーであり、ギブスの自由エネルギーは等温定圧過程における自由エネルギーである。

ヘルムホルツの自由エネルギーは、

$$F = U - TS \quad (2.85)$$

と表すことができ、ここで  $U$  は内部エネルギー、 $T$  は絶対温度、 $S$  はエントロピーを示す。自然界において、エントロピーは増大する方向へ進む（不可逆）ため、熱力学第二法則にある通り、自由エネルギーは減少する方向に進む。

ギブスの自由エネルギーは、

$$G = H - TS \quad (2.86)$$

と表され、 $H$  はエンタルピーを示す。エンタルピーは内部エネルギー  $U$  に圧力  $p$  と体積  $V$  の積 (体積変化) を加えたものである。ヘルムホルツの自由エネルギーと同様に、自由エネルギーは減少する方向に進むことがわかる。統計力学においては、分配関数  $Z$  を用いて関係付けることができる。

ミクロカノニカルにおける分配関数  $Z$  は次のような式で表すことができる。

$$Z(T) = \int d\Gamma \exp(-\mathcal{H}(\Gamma)/k_B T) \quad (2.87)$$

$\mathcal{H}$  はハミルトニアンである。

ここで、ある状態  $\Gamma$  がエネルギー  $E$  をとる確率密度は、

$$P(E) = C\delta(E - \mathcal{H}(\Gamma)) \quad (2.88)$$

となる。 $\delta(E - \mathcal{H}(\Gamma))$  はデルタ関数である。この確率密度を全区間  $(-\infty, \infty)$  で積分すると次式が成り立つ。

$$1 = \int_{-\infty}^{\infty} dE \delta(E - \mathcal{H}(\Gamma)) \quad (2.89)$$

これを用いると、

$$Z(T) = \int_{-\infty}^{\infty} dE \int d\Gamma \delta(E - \mathcal{H}(\Gamma)) \exp(-\mathcal{H}(\Gamma)/k_B T) \quad (2.90)$$

となる。さらにエネルギー  $E$  をもつ状態の状態密度  $W(E)$  の式を用いて、

$$W(E) = \int d\Gamma \delta(E - \mathcal{H}(\Gamma)) \quad (2.91)$$

$$Z(T) = \int_{-\infty}^{\infty} dE W(E) \exp(-E/k_B T) \quad (2.92)$$

と変形する。またここで、ボルツマンの関係

$$S = k_B \log W \quad (2.93)$$

を用いると、

$$Z(T) = \int_{-\infty}^{\infty} dE \exp\left(-\frac{1}{k_B T}(E - TS(E))\right) \quad (2.94)$$

となる。  $k_B$  はボルツマン定数である。自由エネルギーが変化しない条件（関数のピーク）を用いて、

$$\left. \frac{\partial(TS(E) - E)}{\partial E} \right|_{E=E^*} = 0 \quad (2.95)$$

$$\left. \frac{\partial S(E)}{\partial E} \right|_{E=E^*} = \frac{1}{T} \quad (2.96)$$

積分の値は  $\Delta E^*$  を用いて、評価することができ

$$Z(T) = \exp\left(-\frac{1}{k_B T} (E^* - TS(E^*))\right) \Delta E^* \quad (2.97)$$

となる。両辺の自然対数をとることで、

$$-k_B T \log Z(T) = E^* - TS(E^*) - k_B T \log(\Delta E^*) = F \quad (2.98)$$

となり、自由エネルギーの式と一致していることが確認できる（左辺はヘルムホルツ自由エネルギー  $F$  である）。右辺最後の項は、最初の二項に比べ、非常に小さい値のため無視することができる。

### 2.8.3. 独立な系におけるにおけるハミルトンニアン

$N$  個の粒子からなる系における、ハミルトンニアンは、

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} u(|\mathbf{r}_i - \mathbf{r}_j|) + \sum_{i=1}^N u_e(\mathbf{r}_i) \quad (2.99)$$

となる。式の第 2 項は粒子  $i, j$  間に働く相互作用であり、第 3 項は壁から働く力や、重力などの、各粒子に働く力を表す。一般に、粒子間には相互作用があるが、系の密度  $N/V$  を小さくするほど、その相互作用は小さくなる。理想気体はこの極限を考えたものである。理想気体においては、粒子間の相互作用は無視できるので、ハミルトンニアンは次のようになる。

$$\mathcal{H} = \sum_{i=1}^N \mathcal{H}(\mathbf{r}_i, \mathbf{p}_i) \quad (2.100)$$

$$\mathcal{H}(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + u_e(\mathbf{r}) \quad (2.101)$$

よって、理想気体の各粒子は独立な系とみなせる。

以上のことから、 $N$ 個の粒子からなる分配関数は、 $N$ 個の独立な系を合わせたものとみなすことができる。したがって、分配関数は

$$Z_N = z^N \quad (2.102)$$

となる。 $z$ は粒子1つの分配関数である。

$$z = \int d\mathbf{r} \int d\mathbf{p} \exp(-\mathcal{H}(\mathbf{r}, \mathbf{p})/k_B T) \quad (2.103)$$

ここで、とくに外力が働いていない場合、粒子のポテンシャルエネルギーは系に閉じ込めるためのポテンシャルのみである。よって、ハミルトンニアンは運動エネルギーの項のみとなるため、式に粒子の位置が含まれなくなるため、位置に対する積分は体積  $V$  となり、分配関数は、

$$z = V \int d\mathbf{p} \exp\left(-\frac{\mathbf{p}^2}{2mk_B T}\right) \quad (2.104)$$

となる。運動量に対する積分は、ガウス積分の公式を用いることで、次のようになる。

$$\begin{aligned} \int d\mathbf{p} \exp\left(-\frac{\mathbf{p}^2}{2mk_B T}\right) &= \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \exp\left(-\frac{p_x^2 + p_y^2 + p_z^2}{2mk_B T}\right) \\ &= (2\pi mk_B T)^{\frac{3}{2}} \end{aligned} \quad (2.105)$$

よって、系全体の分配関数は、

$$Z_N = V^N (2\pi mk_B T)^{\frac{3N}{2}} \quad (2.106)$$

となる。

#### 2.8.4. $T$ - $P$ 分布 ( $NPT$ アンサンブル) における分配関数

$T$ - $P$  分布は粒子数  $N$ 、温度  $T$ 、圧力  $P$  一定に保たれた系である。本シミュレーション ( $T$ - $P$  分布) における分配関数  $\Upsilon$  は、

$$\Upsilon(T, P, N) = \int_0^{\infty} \exp\left(\frac{-PV}{k_B T}\right) Z_N(T, V, N) dV \quad (2.107)$$

となる。  $Z_N(T, V, N)$  はカノニカル分布における分配関数である。  $T$ - $P$  分布における自由エネルギーは分配関数  $\Upsilon(T, P, N)$  を用いて次のような式で与えられる。

$$G(T, P, N) = -k_B T \log \Upsilon(T, P, N) \quad (2.108)$$

式の形はカノニカル分布における分配関数を用いた自由エネルギーの式 (2.98) の左辺と同じである。  $T$ - $P$  分布における自由エネルギーはギブスの自由エネルギーと呼ばれる。

### 2.8.5. ギブスのパラドックス

古典統計力学においては、粒子を互いに区別できるものとして、理論が組み立てられているが、状態数や分配関数から熱力学的な量を計算するとき、同種の粒子を互いに区別できないものとして、状態を考える必要がある。図 2.8.2 に示すように、古典力学的には (a) の状態  $(\mathbf{r}, \mathbf{p})$  と (b) の状態  $(\mathbf{r}', \mathbf{p}')$  は異なっている。しかし、熱力学的にはどちらも同じであり、その位置に粒子があるという状態のみが意味をもつ。これは、ギブスのパラドックスといい、古典力学的には異なった状態であるが、熱力学的には同じ状態であるとして見なくてはならないということである。よって、これを元に状態の数を考えると、粒子が 2 個の場合は、今までの状態の数を半分にする事となる。

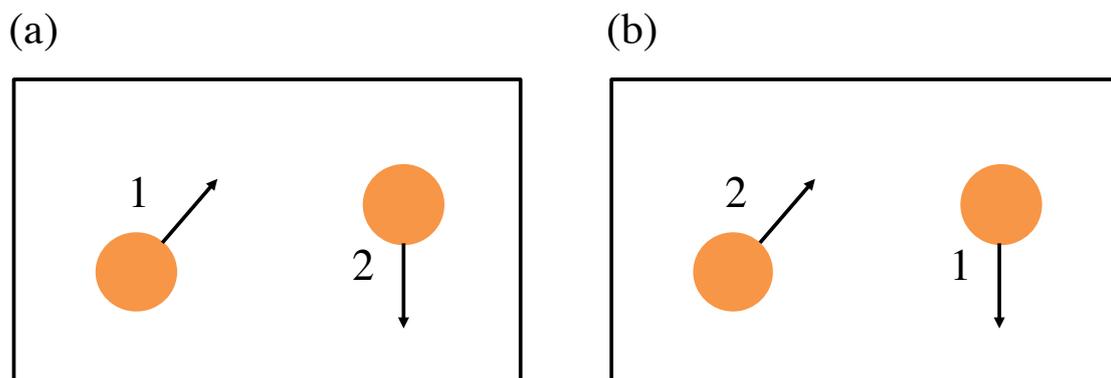


図 2.8.2 ギブスのパラドックス

同種粒子が  $N$  個ある系を考えたとき、粒子  $N$  個の並べ替えの問題は  $N!$  であるから、これまでの数え方に対して、熱力学的に同等な状態数は  $N!$  回重複することとなる。すなわち、状態数を  $N!$  で割れば良い。

古典統計力学は以上のような曖昧さがあるが、量子統計力学においては生じない。それは、量子力学においては、同種粒子の区別ができないものとして、理論が組み立

てられているからである。そこで、量子統計力学に基づいた分配関数は、次のようになる。

$$Z(T) = \frac{1}{(2\pi\hbar)^f N!} \int d\Gamma \exp(-\mathcal{H}(\Gamma)/k_B T) \quad (2.109)$$

ここで、 $f$  は系の自由度であり、 $(2\pi\hbar)^f$  は位置と運動量に関する量子力学的な不確定性原理より由来する。 $\hbar$  はディラック定数であり、プランク定数  $h$  を  $2\pi$  で割ったものである。プランク定数  $h$  は光子一つが持つエネルギーと振動数の関係式における比例定数である。ただし、エントロピーや自由エネルギーはこれまで通り、式 (2.93)、式 (2.98) のまま与えられる。

等温等積における分配関数は、体積  $V$  の箱に  $N$  個の粒子が入っているとして、次のように表せる。

$$Z_N(V, T) = \frac{z(V, T)^N}{N!} \quad (2.110)$$

$z(V, T)$  は粒子 1 つの分配関数である。 $N!$  は各粒子を独立に積分しているのので、数えすぎを防ぐために導入している。

$$z(V, T) = \frac{1}{(2\pi\hbar)^3} \int d\mathbf{r} \int d\mathbf{p} \exp\left(-\frac{\mathbf{p}^2}{2mk_B T}\right) \quad (2.111)$$

ここで、式 (2.105) を用いて、

$$z(V, T) = \frac{V(2\pi mk_B T)^{\frac{3}{2}}}{(2\pi\hbar)^3} = \frac{V}{\lambda_T^3} \quad (2.112)$$

となる。 $\lambda_T^3$  は熱的ドブロイ波という長さを持った定数であり、

$$\lambda_T^3 = \frac{2\pi\hbar}{\sqrt{2\pi mk_B T}} \quad (2.113)$$

となる。式 (2.112) を式 (2.110) に代入すると、その分配関数は、

$$Z_N(V, T) = \frac{V^N}{\lambda_T^{3N} N!} \quad (2.114)$$

となる。

## 2.8.6. 分子シミュレーションにおける自由エネルギーの算出

通常、自由エネルギーを用いて生体高分子の複数の状態を評価する場合、自由エネルギー差を用いる。例えば、天然構造のギブスの自由エネルギーを  $G_N$ 、変性構造（構造が変化）のギブスの自由エネルギーを  $G_D$  としたときの自由エネルギー差  $\Delta G$  が構造の安定性を決定する。

$$\Delta G = G_D - G_N \quad (2.115)$$

変性の自由エネルギー  $G_D$  が天然構造の自由エネルギー  $G_N$  を下回ったとき、構造は変化（変性）する。

本研究の解析において、自由エネルギーは生体高分子の構造の出現頻度分布関数を用いて計算することができる。すなわち、 $T, p$  一定の条件下における分子動力学シミュレーションで系の状態をうまくサンプリングできているならば、生体高分子の構造の出現頻度の比は  $T$ - $P$  分布における分配関数  $\Upsilon(T, P, N)$  とみなすことができる

$$\Upsilon(T, P, N) \approx \frac{P(x, y)}{P_{\max}}. \quad (2.116)$$

ここで  $P(x, y)$  は構造の出現頻度を反応座標  $(x, y)$  に射影した分布関数、 $P_{\max}$  は  $P(x, y)$  の最大値を示す。すると自由エネルギー差  $\Delta G$  は、式 (2.108) より、

$$\Delta G(x, y) = -k_B T \log \frac{P(x, y)}{P_{\max}} \quad (2.117)$$

と表すことができる[65]。頻度分布の例を図 2.8.3 に示す。実際に計算する際は、反応座標をメッシュに切り分け、メッシュ内に射影された構造の数の状態数を数える。本研究では、メッシュサイズを  $64 \times 64$  として計算を行った。この式によりギブスの自由エネルギーを評価できるのは、シミュレーションしている系が等温定圧だからである。

自由エネルギー地形は、頻度分布関数の反応座標  $(x, y)$  の選び方により、見られる地形が異なってくる。例として、RMSD(根平均二乗変位)、慣性半径 ( $R_g$ )、末端間距離  $d_{\text{end}}$ 、二面角、主成分分析により得られる第一主成分、第二主成分などがある。これらは状況に応じて、選び方を決める必要がある。

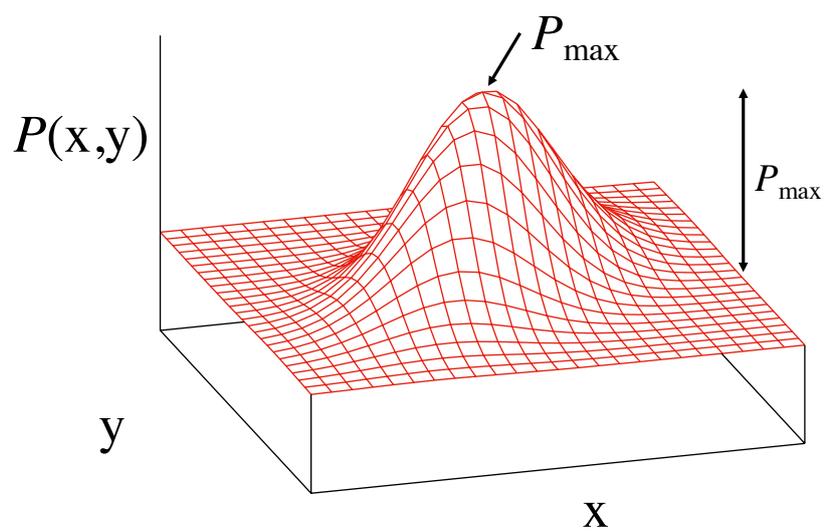


図 2.8.3 頻度分布関数の例.  $x, y$  は反応座標を示す.

## 2.9. シミュレーションによる構造探索

タンパク質は様々な構造を取ると考えられるが、図 2.9.1a に示すように、タンパク質の立体構造には局所安定(ローカルミニマム)構造が数多くある。そのため、図 2.9.1a に示す通り、通常の分子動力学シミュレーションを用いると、ローカルミニマムにトラップされやすい。理論上は全ての構造を通常の分子動力学シミュレーションでみることは可能だが、一度、ローカルミニマムにトラップされてしまうと、そこから抜け出すには、非常に長いシミュレーションを行う必要がある。よって、通常の分子動力学シミュレーションを用いて、タンパク質の構造を網羅するのは、現実的ではない。これを解決する方法の一つに、レプリカ交換分子動力学 (REMD) 法[9]がある。大域的最低値 (グローバルミニマム) を求める方法としては、シミュレーテッドアニーリング (焼き鈍し法) [5]も有効な手段であるが、グローバルミニマムを含め、広域における構造を得るためには、レプリカ交換法が大変有用である。レプリカ交換法は、温度の異なる系 (レプリカ) を複数用意し、あるステップ数ごとに各レプリカ間の温度をメトロポリス法に従い交換することで、様々な温度空間 (エネルギー空間) をランダムウォークさせる方法である。レプリカ交換法を用いることで、図 2.9.1b に示すように、構造がローカルミニマムにトラップされることなく様々な構造を得ることができ。しかし、レプリカ交換法はシミュレーション中に系の温度を交換する (温度が変わる) ため、通常の分子動力学法と異なり、時間発展を見ることができない。よって、対象とする分子のダイナミクスを見るためには通常の分子動力学法を用いる必要がある。

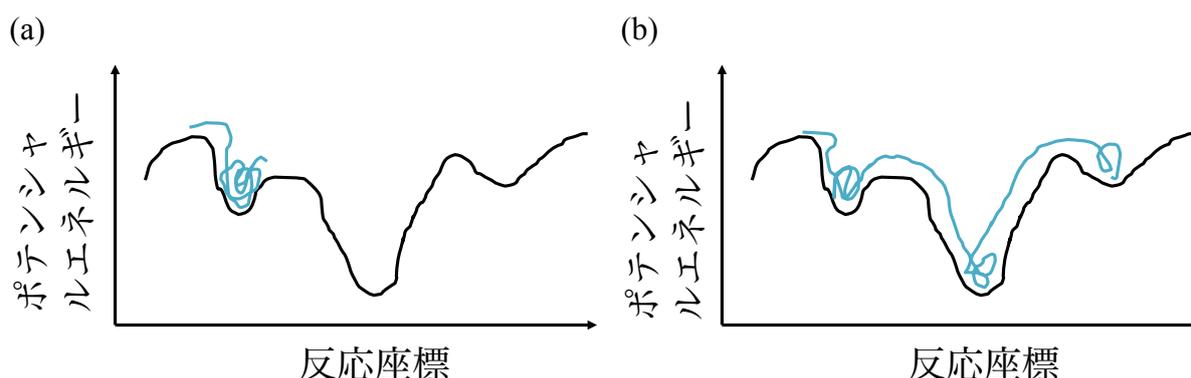


図 2.9.1 ポテンシャルエネルギー地形上におけるタンパク質の軌跡の概念図。(a) は通常の分子動力学シミュレーション、(b) は REMD シミュレーションの軌跡を示す。

## 2.9.1. シミュレーテッドアニーリング

シミュレーテッドアニーリング[5, 66], 別名焼きなまし法は局所探索を行うに当たって, 通常は図 2.9.1a のようにローカルミニマムに陥ってしまうが, 高温ではいろいろな構造をとる性質を用いることでローカルミニマムに陥らないようにした手法である. つまり, 探索結果に偏りが無いようにするための手法である. シミュレーテッドアニーリングの名前は, 金属加工で用いられる焼きなましに由来している. 熱を加えることにより, 原子が初期位置から離れ, 高エネルギー状態で構造空間をランダムに動く. これは, 図 2.9.1a に示されるようなローカルミニマムに陥っている状態に熱を加えることで, あるエネルギー障壁を越えさせることに値する. その後, ゆっくり冷却することにより, 原子の再配列が行われる. これによって, エネルギーが初期構造より低い状態を得る可能性がある. これに基づき, グローバルミニマムの探索を可能とする.

シミュレーテッドアニーリングはヒーティング段階, 高温サンプリング, アニーリング (冷却段階) の 3 段階に分けて行う (図 2.9.2) .

### 1. ヒーティング

温度を上げることにより, 折り畳まれた構造のエネルギー障壁を越え, 様々な構造をサンプリングすることが可能となる, これにより, 初期構造の依存性を払拭することができる. パラメーターとしての温度を 1000K 程度まで上昇させれば,  $\beta$ -シート構造のような硬い構造でも, エネルギー障壁を越えるのに十分な状態であると考えられる.

### 2. 高温サンプリング

温度を上昇させ, 高温状態のままシミュレーションを行い, 構造をサンプリングする. この段階より, 上述したように様々な構造を得ることができる.

### 3. アニーリング

サンプリングにより得られた初期構造をそれぞれ, 冷却 (アニーリング) する. これにより, 最終構造が得られる. 冷却は, その都度状況に応じて, 決める必要がある.

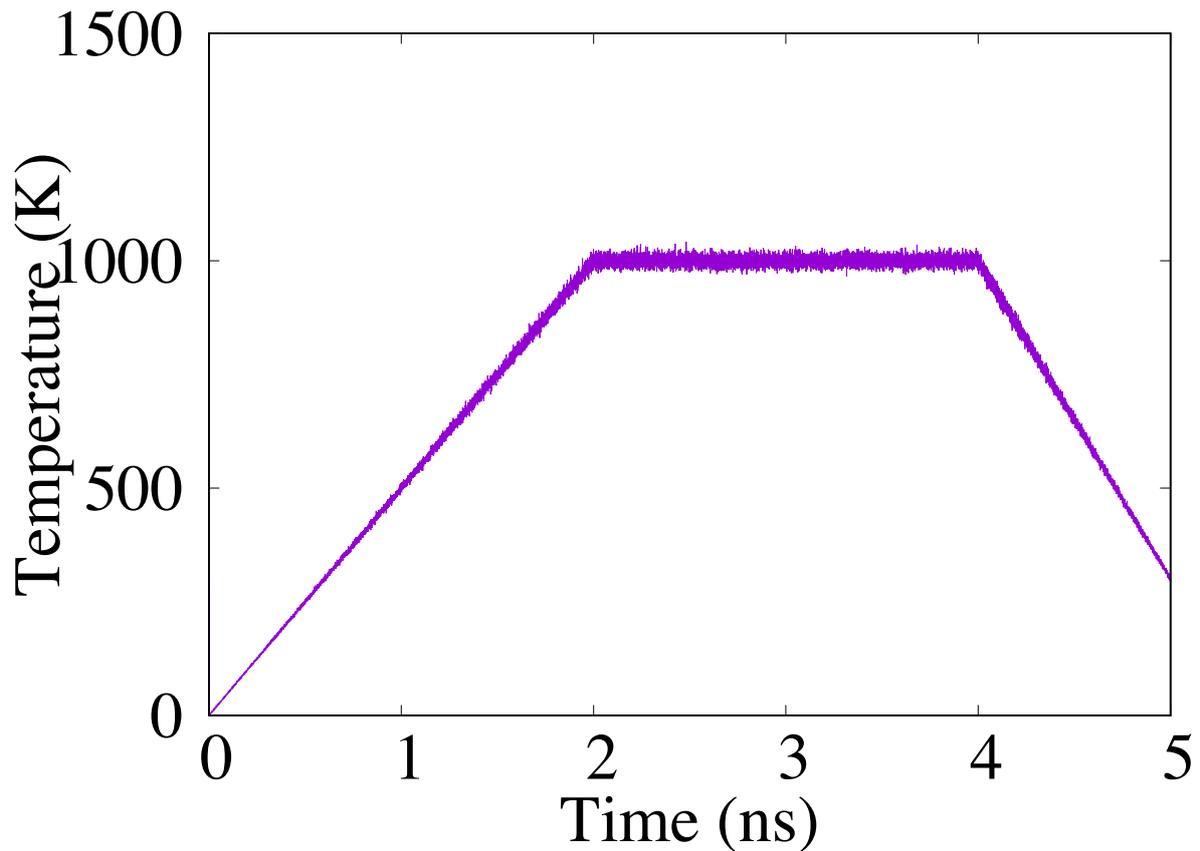


図 2.9.2 シミュレーテッドアニーリングにおける温度変化の様子。

## 2.9.2. レプリカ交換法

異なる温度  $T_1, T_2, \dots, T_i, \dots, T_j, \dots, T_M$  (本研究においては 300 K–450.5 K) を持つ系のレプリカを複数  $x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_M$  (本研究においては 48 個) 用意し, それぞれ独立した分子動力学シミュレーションを行う. そして, あるステップごとに各レプリカ間の温度交換を試み, 交換が成立した場合, それぞれのレプリカは交換後の温度で独立して分子動力学シミュレーションを行う. 多くの場合, レプリカの交換過程において, 十分条件となる詳細釣り合いの条件が用いられる. 詳細釣り合いとは, ある状態  $x_i$  から状態  $x_j$  に遷移する確率とその逆のある状態  $x_j$  から状態  $x_i$  に遷移する確率が等しいことを示す. ここで言う状態は系全体を指す. このようにして, 様々な温度空間 (エネルギー空間) のランダムウォークを可能とする方法である. 図 2.9.3 にその概念図を示す. この図は各レプリカの他の温度への遷移の様子を表している. 紫色のレプリカを例にすると, 最初の交換で, 隣り合う空色のレプリカと交換 (つまり, 紫色のレプリカは温度が上昇し, 空色のレプリカは温度が下降する) しており, 続いて, 緑色のレプリカと温度を交換し, 次はそのままの温度を維持している様子がわか

る。ここで示す交換は温度のみであるが、エネルギーや圧力を交換してシミュレーションを行う場合もある。

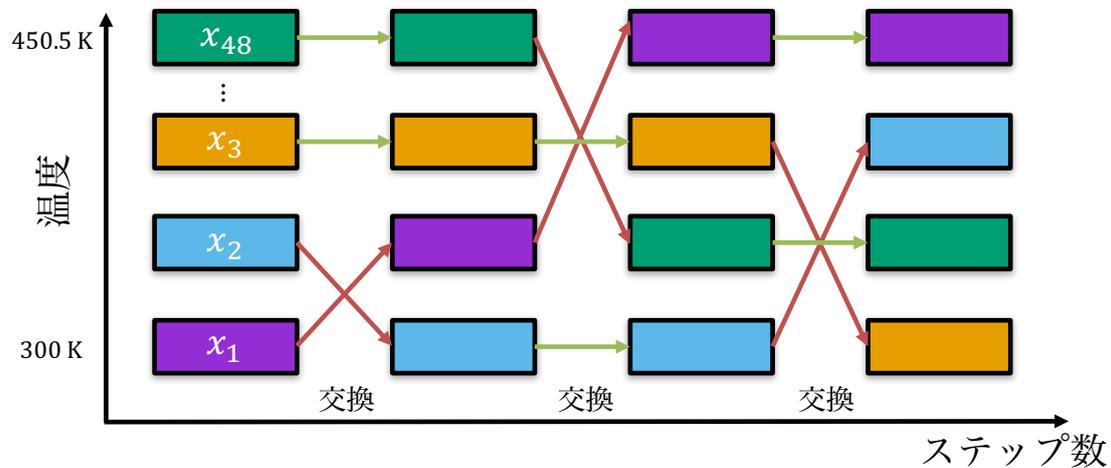


図 2.9.3 レプリカ交換分子動力学法の概念図

交換はメトロポリス法[10]を基に行っており、各レプリカの交換確率  $\mathcal{W}$  (ある状態  $x_i$  から状態  $x_j$  に遷移する確率) は次のようになる。

$$\mathcal{W}(x_i \rightarrow x_j) = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ \exp(-\Delta) & \text{for } \Delta > 0 \end{cases} \quad (2.118)$$

$$\Delta = \left( \frac{1}{k_B T_i} - \frac{1}{k_B T_j} \right) (E_i - E_j)$$

ここで、 $k_B$  はボルツマン定数、 $T_i$ 、 $T_j$  は状態  $x_i$ 、 $x_j$  の温度、 $E_i$ 、 $E_j$  は状態  $x_i$ 、 $x_j$  のポテンシャルエネルギーである。

上述したエネルギーを交換する場合の交換確率  $\mathcal{W}$  は、

$$\mathcal{W}(x_i \rightarrow x_j) = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ \exp(-\Delta) & \text{for } \Delta > 0 \end{cases} \quad (2.119)$$

$$\Delta = \frac{1}{k_B T} \{ (E_i(x_j) - E_j(x_i)) + (E_i(x_i) - E_j(x_j)) \}$$

となる。

### 2.9.3. 温度間隔の決定

各レプリカ間の温度間隔を決定するためには、各レプリカ間の状態（本研究においては系のポテンシャルエネルギー）を見ることが有効である。これは、交換確率にレプリカ間のポテンシャルエネルギー差が効いてくるためである。ここでは、本研究で行った計算を例に、各レプリカのポテンシャルエネルギーの分布を図 2.9.4 に示す。このように、ポテンシャルエネルギーの分布がオーバーラップしていることが重要であり、それぞれの分布が十分に重なるよう、温度を設定する必要がある。

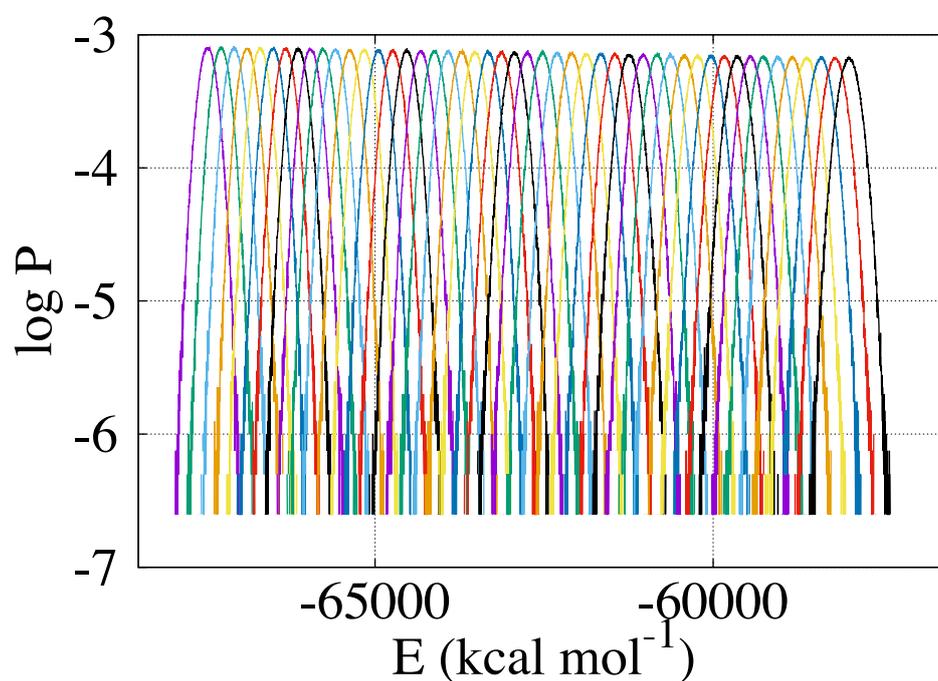


図 2.9.4 各レプリカのポテンシャルエネルギーの分布

## 2.10. シミュレーションによる自由エネルギー計算

本節ではシミュレーションを用いて、自由エネルギーを計算する方法の一つであるアンブレラサンプリング法[67]について説明する。

### 2.10.1. アンブレラサンプリング法

アンブレラサンプリング法は反応座標 (Reaction coordinate)  $\zeta$  を用いて、構造のサンプリングを行う手法である。また、この方法によりある反応座標  $\zeta$  に沿った PMF (Potential of Mean Force: 平均力ポテンシャル) を計算することが可能である。平均力ポテンシャルについては、2.10.3 項で説明する。アンブレラサンプリング法は、ポテンシャルエネルギー  $E$  に人工的なポテンシャルエネルギー (アンブレラポテンシャル)  $E(\zeta)$  を導入することで、もとのポテンシャルエネルギーに  $X(\zeta)$  を加えた、 $E_{\text{umb}}^* = E(\zeta) + X(\zeta)$  を用いることで、サンプリング効率を上げることができる。図 2.10.1 のように  $\zeta$  に沿って幅広くサンプリングできることから、傘 (アンブレラ) という名称が付けられた。

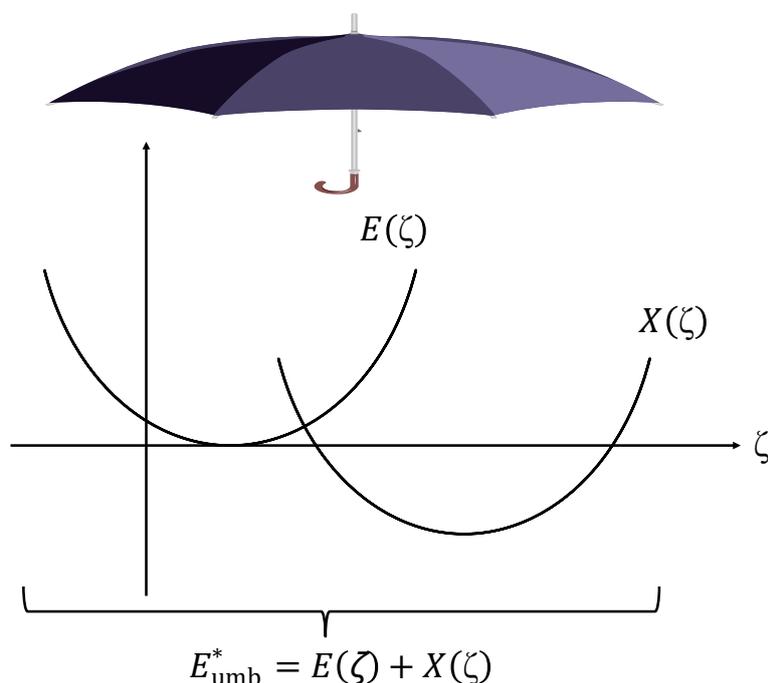


図 2.10.1 アンブレラサンプリングの概念図。

## 2.10.2. 自由エネルギー計算

分配関数を用いて自由エネルギーを計算できることは2.8節で紹介した。ここではその計算法を説明する。また、カノニカルアンサンブルにおける自由エネルギーについて説明しているが、 $NPT$  アンサンブルにおけるギブスの自由エネルギーについても同様の形となる。これは式 (2.98) と式 (2.108) で示されるように自由エネルギーの式は一致しており、状態がうまくサンプリングされていれば、カノニカルアンサンブルにおける自由エネルギーも  $NPT$  アンサンブルにおける自由エネルギーも同様の形で計算することができるからである。

粒子の運動エネルギーの他に系のポテンシャル  $U$  も考慮した時、分配関数は、

$$Z_N(V, T) = \frac{1}{\lambda_T^{3N} N!} \int d\mathbf{r}^N \exp(-U/k_B T) \quad (2.120)$$

となる。ここで

$$V^N = \int d\mathbf{r}^N = \int d\mathbf{r}^N \exp(U/k_B T) \exp(-U/k_B T) \quad (2.121)$$

この関係式を用いて、分配関数を書き直すと、

$$Z_N = \frac{\int d\mathbf{r}^N \exp(U/k_B T) \exp(-U/k_B T)}{\lambda_T^{3N} N! \int d\mathbf{r}^N \exp(U/k_B T) \exp(-U/k_B T)} \times \int d\mathbf{r}^N \exp(-U/k_B T) \quad (2.122)$$

$$= \frac{\int d\mathbf{r}^N \exp(-U/k_B T)}{\lambda_T^{3N} N! \int d\mathbf{r}^N \exp(U/k_B T) \exp(-U/k_B T)} \times \int d\mathbf{r}^N \exp(U/k_B T) \exp(-U/k_B T) \quad (2.123)$$

$$= \frac{V^N}{\lambda_T^{3N} N! \langle \exp(U/k_B T) \rangle_U} \quad (2.124)$$

となる。 $\langle \exp(U/k_B T) \rangle_U$  はポテンシャル  $U$  空間の状態数を平均するという意味をもつ。この得られた分配関数を自由エネルギーの式 ( $-k_B T \log Z(V, T)$ ) に代入すると、

$$F = -k_B T \log Z(V, T) = -k_B T \log \frac{V^N}{\lambda_T^{3N} N! \langle \exp(U/k_B T) \rangle_U} \quad (2.125)$$

$$= k_B T \log \langle \exp(U/k_B T) \rangle_U - k_B T \log(V^N) + k_B T \log(\lambda_T^{3N} N!) \quad (2.126)$$

となる。この式は、ポテンシャルエネルギー  $U/k_B T$  に依存して急速に増加する式であり、このままの式の形を用いるには精度が悪く、シミュレーションにそのまま用いるには現実的でないため、少しばかり工夫する必要がある。

### 2.10.3. 平均力ポテンシャル (Potential of means force: PMF)

前節で自由エネルギーの計算について紹介した。ここでは平均力ポテンシャル[68, 69]による自由エネルギー計算を説明する。ある状態を表すパラメーター  $\zeta$  があつたとき、 $\zeta$  の mean force は  $\mathbf{r}^N$  でポテンシャル  $U$  を積分することで、次のような式で得られる。

$$\text{mean force} = \frac{\int d\mathbf{r}^N \delta(\zeta' - \zeta) \exp(-U/k_B T)}{\int d\mathbf{r}^N \exp(-U/k_B T)} = \frac{\partial}{\partial \zeta} [k_B T \log \rho(\zeta)] \quad (2.127)$$

$\rho(\zeta)$  は状態  $\zeta$  の確率分布であり、

$$\rho(\zeta) = \frac{\int d\mathbf{r}^N \delta(\zeta' - \zeta) \exp(-U/k_B T)}{\int d\mathbf{r}^N \exp(-U/k_B T)} = \langle \delta(\zeta' - \zeta) \rangle_U \quad (2.128)$$

となる。ここで、

$$W(\zeta) = -k_B T \log \rho(\zeta) \quad (2.129)$$

が導入でき、

$$\frac{\partial}{\partial \zeta} W(\zeta) = \frac{\partial}{\partial \zeta} [k_B T \log \rho(\zeta)] \quad (2.130)$$

mean force の式と一致することがわかる。

### 2.10.4. シミュレーションにおける PMF の計算例

上述したように、反応座標  $\zeta$  を用いて、自由エネルギーを計算できることを示した。ここでは、シミュレーションにおける、計算例を示す。シミュレーションにおいてよく用いられる反応座標  $\zeta$  は、図 2.10.2 に示すように、対象とする二つの分子の内、片方を反応座標に沿って引っ張りそれぞれの重心間距離を反応座標とするものである。このとき、反応座標の方向を固定するため、もう片方は位置拘束をする。

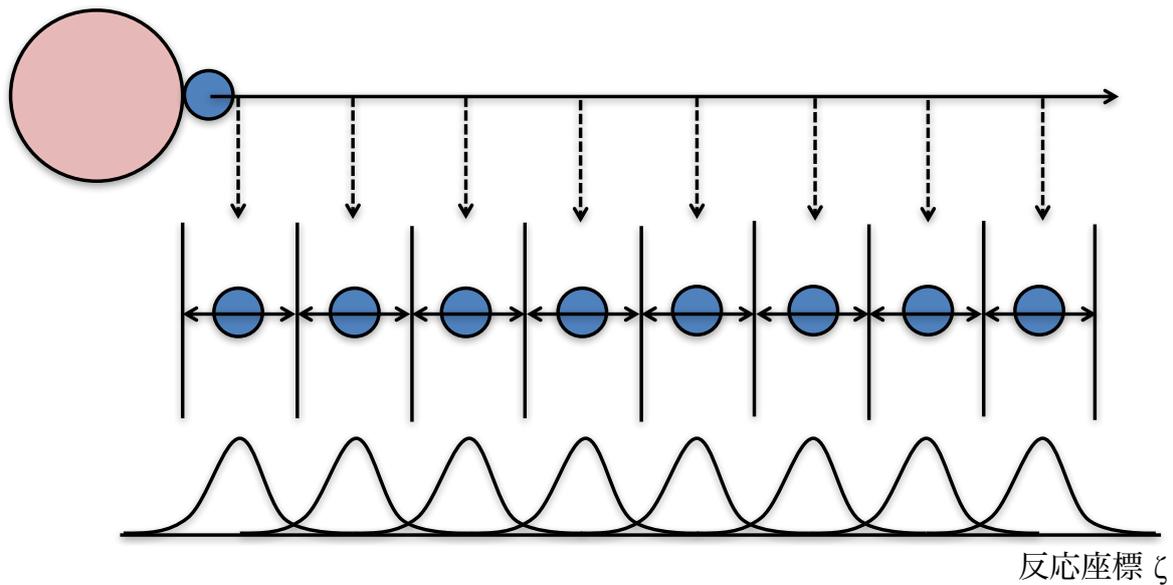


図 2.10.2 アンブレラサンプリングの概要.

図 2.10.3 に実際にシミュレーションを行った例として、凝集したプリオンタンパク質の一つの  $\beta$  シート構造が外れるのにどれくらいのエネルギーが必要かを示す. ここで示す反応座標  $\zeta$  は緑で示した  $\beta$  シート領域の重心と、赤色で示した  $\beta$  シート領域の重心間距離としている. この結果は凝集したプリオンタンパク質を崩すには非常に強い力が必要であることを示している.

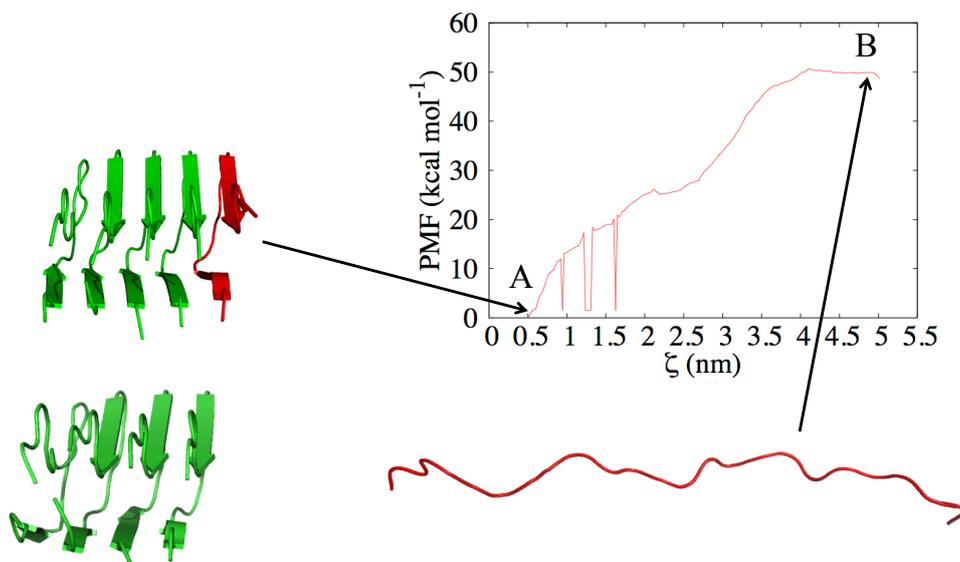


図 2.10.3 プリオンタンパク質の破壊過程における PMF の見積りの例.

図 2.10.3 の結果は粗く見積もっているが，このように自由エネルギーを見積もることができる．アンブレラサンプリング法は受容体とタンパク質の結合部位の予測において，有用である．

## 2.11. シミュレーションの条件と手順

本節では、レプリカ交換法と通常の分子動力学シミュレーションの条件と手順について述べる。本研究は分子動力学ソフトウェア、GROMACS[70]を用いて行った。GROMACS はフローニンゲン大学で開発された分子動力学シミュレーションソフトウェアパッケージである。

### 2.11.1. EF1 および EF2 ペプチドの初期構造

シミュレーションにおける EF1 および EF2 の初期構造は、図 2.11.1 に示した X 線結晶構造回折より得られた、マウスラミニニン  $\alpha 1, 2$  鎖 LG4 モジュールより切り出した構造を用いた[38, 71]。PDB ID はそれぞれ、2JD4 と 1DYK である。(a) の青色で示した領域が EF1 の配列部位であり、(b) の黄色で示した領域は EF2 の配列部位である。EF1 と EF2 は共に折り畳まれたタンパク質内では、 $\beta$ -シート構造を形成していることがわかる。また、N 末端および C 末端は不安定であるため、キャップ修飾を行った。N 末端側にはアセチル基 (Ace) , C 末端側には N メチル基 (Nme) を修飾した。EF1 および、EF2 の配列は表 1 に示した。この初期構造のことを、以降、結晶構造と呼ぶ。

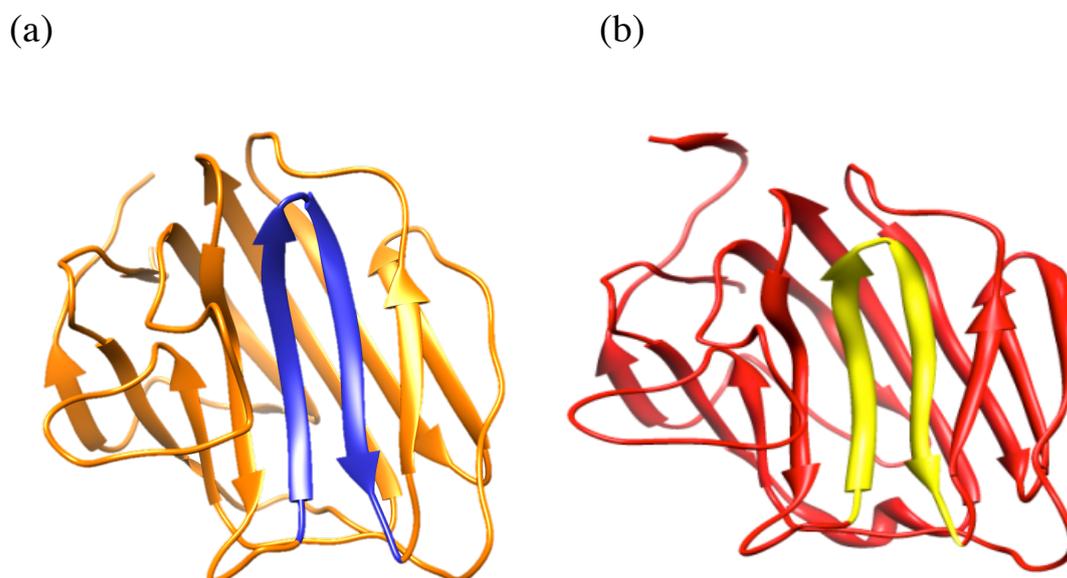


図 2.11.1 マウスラミニニン  $\alpha 1, 2$  鎖 LG4 モジュール。(a) はマウスラミニニン  $\alpha 1$  鎖 LG4 モジュール、(b) はマウスラミニニン  $\alpha 2$  鎖 LG4 モジュールの結晶構造である。(a) の青は EF1 ペプチドの配列領域、(b) の黄は EF2 ペプチドの配列領域を示す。

表1 EF1 および EF2 ペプチドのアミノ酸配列

Peptides	Amino acid sequences
EF1	Ace-DYATLQLQEGRLHFMFDLG-Nme
EF2	Ace-DFATVQLRNGFPYFSYDLG-Nme

## 2.11.2. レプリカ交換法シミュレーションの条件と手順

本研究において、レプリカは48個用意した。温度は300 K から 450.5 K の範囲に設定した。それぞれのレプリカの温度間隔は2.5 K から 4 K おきに設定した。各レプリカの温度は表2に示した。ここで示す温度間隔は、レプリカ間の交換確率が約0.3にするために、短いレプリカ交換シミュレーションを繰り返し行い、経験的に得られた間隔である。力場はAmber99sb-ildnを用いた[72]。水はTIP3Pモデル[73]を使用し、水分子の数は、EF1においては7,953個、EF2においては7,011個を加えた。系の境界条件は周期境界条件を使用した。水を含めた系を用意する際、系を中性に保つため、対イオンとして、EF1においてはナトリウムイオンを2つ、EF2においては、ナトリウムイオンを1つ加えた。

温度制御と圧力制御はそれぞれ能勢・フーバー熱浴、パリネロ・ラーマン法を用いた。ただし、緩和過程においてのみ、温度制御は速度スケーリング法を用いて行った。温度、体積一定の条件下 (*NVT*: 粒子, 体積, 温度一定) のシミュレーションにおいては、体積が一定になるようにシミュレーションを行う系の大きさを固定した。

はじめに、系を用意したのち、配置した水およびペプチドが無理のない位置になるよう、エネルギー最小化を行い、300 K, 1 bar の条件下 (*NPT*: 粒子, 圧力, 温度一定) で1 ns 緩和のシミュレーションを行った。次に300 K, *NVT* の条件下で1 ns 緩和のシミュレーションを行った。その後、緩和後に48個のレプリカを作成し、各レプリカのそれぞれの温度 (300 K から 450.5 K) で *NVT* の条件下で緩和を行った。各レプリカの緩和が終わった後、各レプリカを *NVT* の条件下で独立に60 ns のシミュレーションを行う。この時、各レプリカ間の交換頻度は2 ps に設定した。解析には、初期構造の依存があるため、5-60 ns の軌跡を用いた。

表 2 各レプリカの温度

Replica 1	300 K	Replica 18	346 K	Replica 35	402 K
2	302.5 K	19	349 K	36	405.5 K
3	305 K	20	352 K	37	409 K
4	307.5 K	21	355 K	38	412.5 K
5	310 K	22	358 K	39	416 K
6	312.5 K	23	361 K	40	419.5 K
7	315 K	24	364 K	41	423 K
8	317.5 K	25	367 K	42	426.5 K
9	320 K	26	370.5 K	43	430.5 K
10	322.5 K	27	374 K	44	434.5 K
11	325 K	28	377.5 K	45	438.5 K
12	328 K	29	381 K	46	442.5 K
13	331 K	30	384.5 K	47	446.5 K
14	334 K	31	388 K	48	450.5 K
15	337 K	32	391.5 K		
16	340 K	33	395 K		
17	343 K	34	398.5 K		

### 2.11.3. シミュレーションの条件と手順

EF1 および EF2 の初期構造はレプリカ交換法と同様の方法で準備した。配列は、表 1 に示してある。水は TIP3P モデルを使用し、水分子の数は、EF1 においては 7,118 個、EF2 においては 7,100 個を加えた。系の境界条件は周期境界条件を使用した。系を中性に保つため、対イオンとして、EF1 においてはナトリウムイオンを 2 つ、EF2 においては、ナトリウムイオンを 1 つ加えた。系の緩和において、温度制御および圧力制御はそれぞれ速度スケージング法、パリネロ・ラーマン法を用いた。緩和後の分子動力学シミュレーションにおいて、圧力制御は緩和と同様にパリネロ・ラーマン法を用いた。温度制御は能勢・フーバー熱浴を用いた。

系を用意したのち、配置した水およびペプチドが接近しすぎない位置になるよう、エネルギー最小化を行った。その後、系が設定温度 (300 K) になるよう等温定積下で緩和を 1 ns 行い、系が設定する圧力 (1 bar) になるよう常温常圧下で緩和を 50 ns 行った。緩和終了後、等温定圧下で 2  $\mu$ s のシミュレーションを行った。シミュレーションは、EF1 および EF2、それぞれ 2 回ずつ行った。それぞれ 2 回のシミュレーションの違いは、水の初期配置および初期速度である。

## 3. 結果

本章では、レプリカ交換法および通常の方法により得られた計算結果について述べる。EF1 および EF2 の構造を幅広くサンプリングするために、(3.1) 節ではレプリカ交換法を用いたシミュレーションを行なった。レプリカ交換法はシミュレーション中に系全体を異なる温度間で交換しており、構造が連続的ではないため、ダイナミクスを調べることができない。そこで、(3.2) 節では等温定圧下におけるシミュレーションを行い、水溶液中におけるダイナミクスの解析を行なった。EF1 は受容体である  $\alpha 2\beta 1$  インテグリンとどのように結合するか分かっていない。

### 3.1. レプリカ交換法シミュレーション

ここでは、レプリカ交換法を用いたシミュレーションの結果を述べる。

#### 3.1.1. レプリカの温度空間の遷移の様子

レプリカ交換法によるシミュレーションは EF1 と EF2、それぞれ 60 ns 行った。図 3.1.1 にレプリカの異なった温度空間への遷移の様子の例を示す。温度のインデックスが小さいほど低い温度であり、インデックスが大きくなるほど高い温度を示す。ここでは、例として、レプリカ 1 (300 K)、レプリカ 17 (343 K)、レプリカ 31 (388 K)、レプリカ 47 (446.5 K) を示した。レプリカ交換法は、各レプリカが最も低い温度空間と最も温度が高い温度空間を行き来することが理想的であるが、ここで示したレプリカの例を見てみると、それを満たしていることがわかる。例えば、図 3.1.1 の (a) における黄線を見てみると、はじめは高い温度にあり、10 ns 付近で低い温度まで遷移した後、再び高い温度に遷移している。40 ns 付近で低い温度まで遷移していることがわかる。紫線を見てみると、25 ns 付近で一番高い温度に遷移し、その後低い温度に戻り、38 ns 付近でまた高い温度に遷移しており、低い温度から高い温度を行き来している様子が見られる。また、レプリカ交換法における交換率は 0.2-0.3 が理想的とされるが、本研究においても、交換率は約 0.2-0.3 を満たしていた。

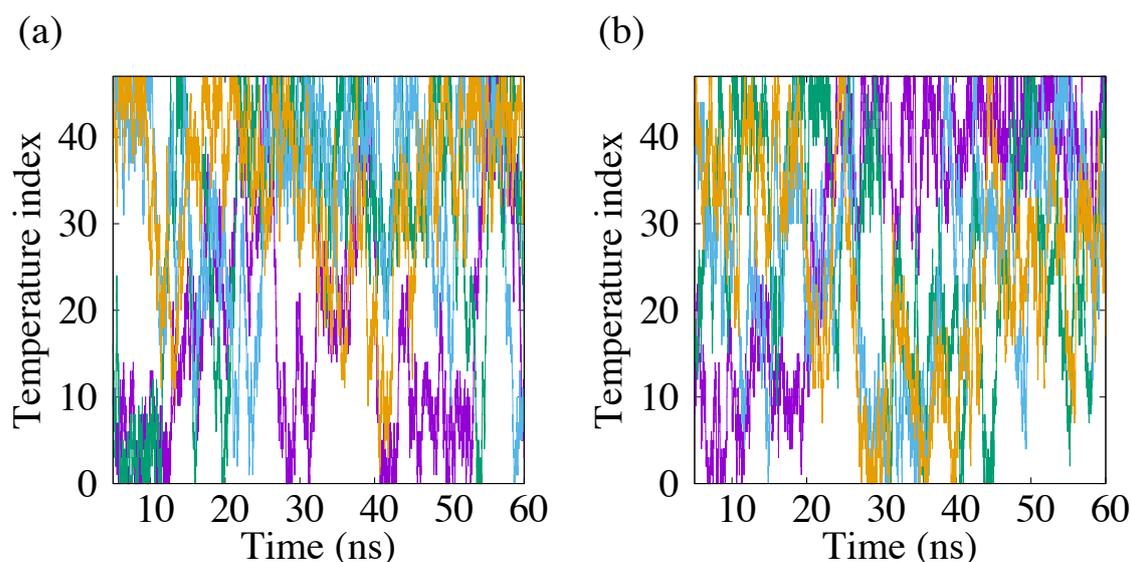


図 3.1.1 各レプリカの異なる温度への遷移の様子. (a) が EF1, (b) が EF2 の温度の遷移の様子を示す. それぞれの線において, 紫色が 300 K, 青色が 343 K, 黄色が 388 K, 緑色が 446.5 K のレプリカに相当する.

### 3.1.2. 精度の確認

はじめに, レプリカ交換法によるシミュレーションの精度を確認する[75, 76]. これは, 温度交換が行われている場合でも, 初期構造の依存が残り続ける可能性があるためである. それを確認するため,  $\beta$ -sheet 構造の形成率の自己相関関数および平均値を計算した. 自己相関関数は次式に基づき計算した.

$$C(t) = \frac{1}{N-j} \sum_{i=0}^{N-1-j} f(i\Delta t) f((i+j)\Delta t), \quad (3.1)$$

ここで,  $t = j\Delta t$ ,  $f$  は物理量である (ここでは  $\beta$ -sheet 構造の形成率).  $N$  はシミュレーション中のフレーム数である.

確認は通常分子動力学シミュレーションにより得られた結果と比較する. 自己相関関数と平均値は図 3.1.2 および表 3 に示す. レプリカ交換法により得られた自己相関関数は, 通常分子動力学シミュレーションにより得られた自己相関関数と比べ, 非常に早く減衰している. また,  $\beta$ -sheet 構造の形成率の平均値を見てみると, レプリカ交換法と通常の方法により得られた値に大きな差異は見られないが, 値のばらつき (標準偏差) を見てみると, レプリカ交換法のばらつきが通常の方法と比べて大きな値となっている. これらのことから, レプリカ交換法における初期構造依存は非

常に早い段階でなくなっていると考えられ、レプリカ交換法の精度は十分であると思われる。

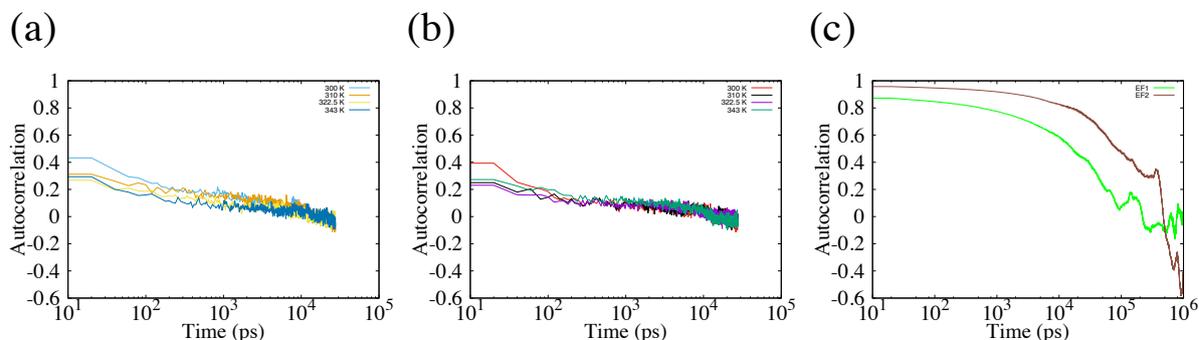


図 3.1.2  $\beta$ -sheet 構造の形成率の自己相関関数関数. (a) と (b) が EF1 と EF2 のレプリカより得られた自己相関関数を示し, (c) は通常の分子動力学シミュレーションにより得られた自己相関関数を示す.

表 3  $\beta$ -sheet の形成率の平均値と標準偏差

		REMD	通常の分子動力学シミュレーション
EF1	300 K	0.444 (0.165)	0.438 (0.068)
	310 K	0.439 (0.165)	
	322.5 K	0.412 (0.169)	
	343 K	0.382 (0.193)	
EF2	300 K	0.415 (0.212)	0.471 (0.148)
	310 K	0.390 (0.218)	
	322.5 K	0.342 (0.217)	
	343 K	0.310 (0.213)	

括弧内の値は標準偏差.

### 3.1.3. 自由エネルギー地形

タンパク質の構造は上述したように、自由エネルギーを使って表すことができるため、レプリカ交換法により得られた構造から、自由エネルギー地形を描いた。自由エネルギー計算には式 (2.117) を用いた。はじめに、EF1 および EF2 の各レプリカより得られた自由エネルギー地形の温度変化による違いをみた。反応座標は根平均二乗変位 (RMSD) と慣性半径  $R_g$  を用いた。RMSD の計算は式 (2.82) を用いて計算した。RMSD および  $R_g$  は主鎖 (窒素, 不斉炭素, 炭素, 酸素) について計算した。それらを組み合わせた頻度分布より得られた自由エネルギー地形を図 3.1.3 (EF1) および

図 3.1.4 (EF2) に示す。自由エネルギー地形は 300 K-450.5 K の間をある程度等間隔になるよう選んで描いた。EF1 の自由エネルギー地形において、温度変化による自由エネルギー地形の変化は 322.5 K まで大きくはなかった。また、頻度が高い領域 ( $\Delta G < 20$  kJ/mol) の広がりがあり、構造は幅広くサンプリングされているが、構造の頻度がより高い領域 ( $\Delta G < 2.5$  kJ/mol) の変化はあまり見られなかった。このことから、EF1 は温度上昇によって変性しにくく、熱安定性があることが示唆される。一方、EF2 の自由エネルギー地形を見てみると、343 K まで、地形に大きな変化はなかったが、温度変化による自由エネルギー地形上の構造の頻度が高い領域の広がりも顕著であった。高温になるにつれて、エネルギー障壁が崩れ、ランダム構造になることがわかる。また、図 3.1.5 に異なる温度における自由エネルギー地形の面積を示す。しきい値を  $\Delta G < 2.5$  kJ/mol,  $\Delta G < 3.75$  kJ/mol,  $\Delta G < 5$  kJ/mol,  $\Delta G < 20$  kJ/mol とし、面積を計算したところ、全体的に EF2 の自由エネルギー地形の面積は EF1 における自由エネルギー地形の面積と比べると、広いことがわかる。これらのことから、EF2 は温度上昇により変性（熱安定性がない）しやすいことが示唆され、EF1 に比べ構造が不安定であることが予想される。

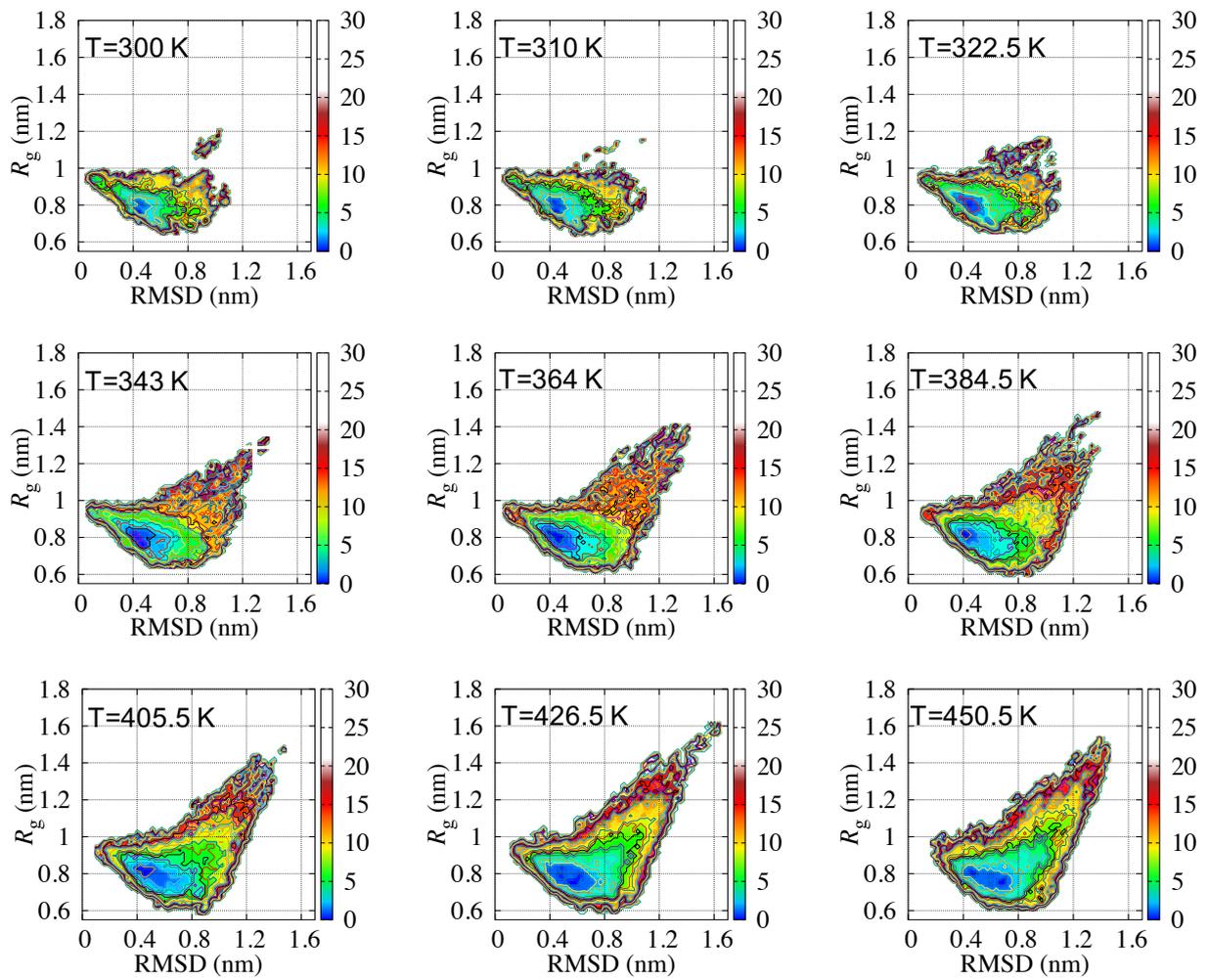


図 3.1.3 EF1 の異なる温度における自由エネルギー地形. カラーバーは自由エネルギー差を示し, 単位は kJ/mol である.

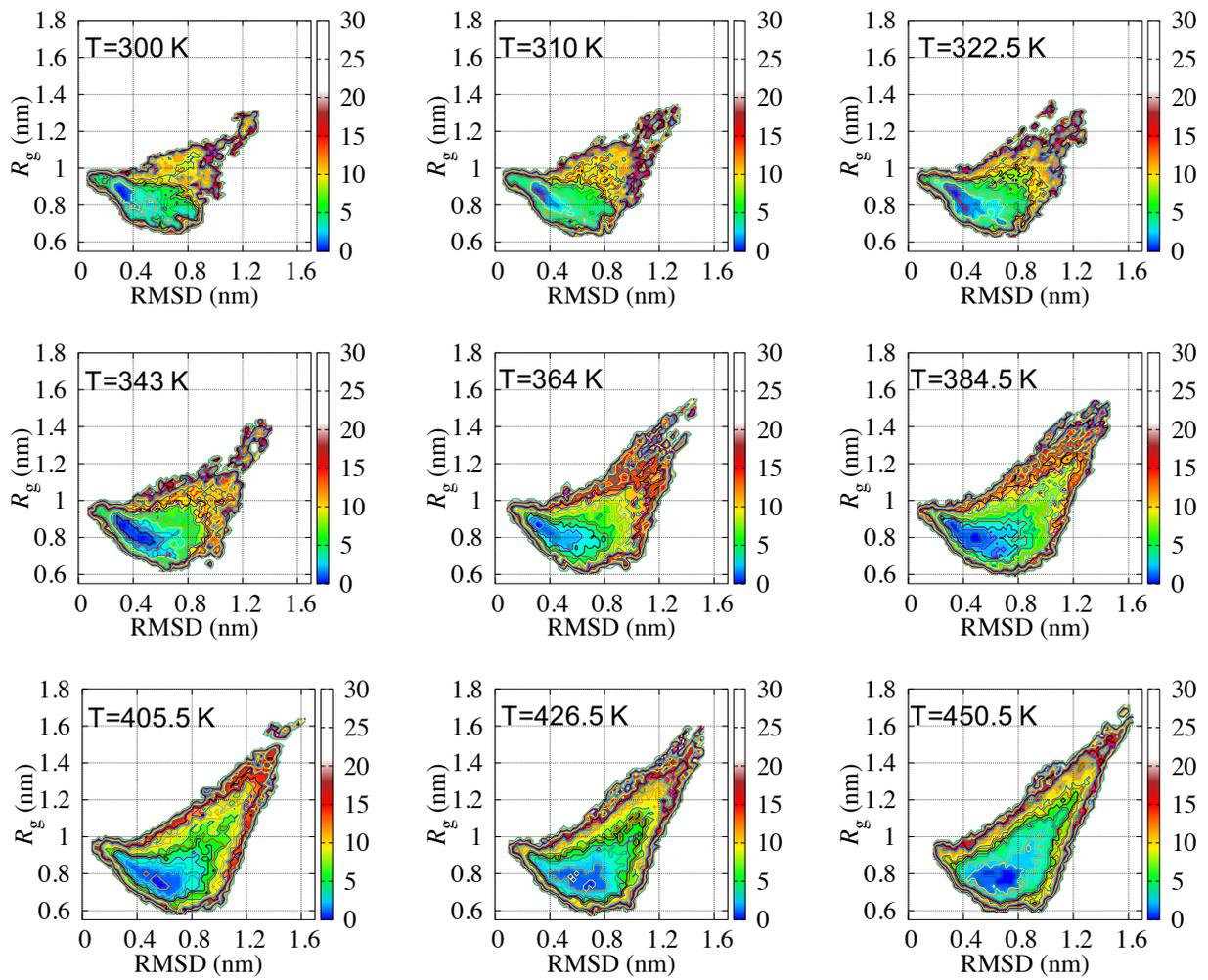


図 3.1.4 EF2 の異なった温度における自由エネルギー地形

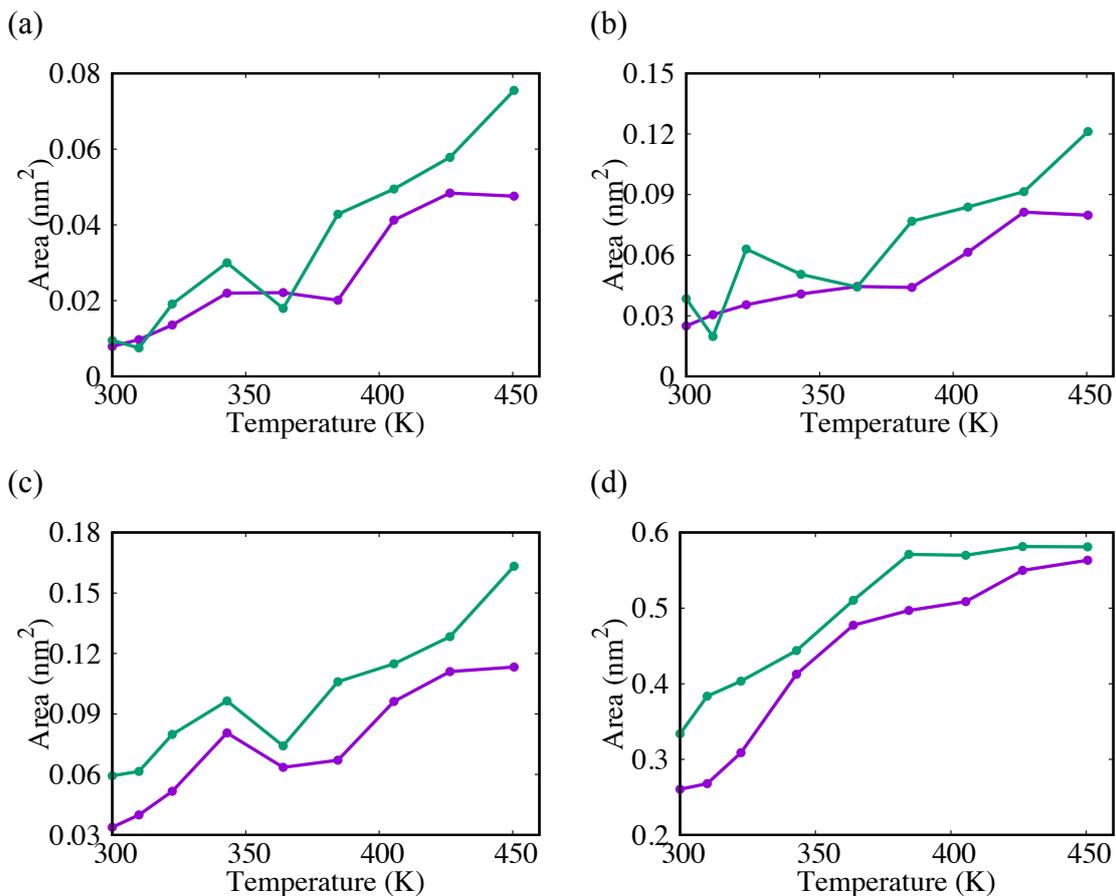


図 3.1.5 異なる温度における自由エネルギー地形の面積. (a) が  $\Delta G < 2.5$  kJ/mol における面積を示し, (b) が  $\Delta G < 3.75$  kJ/mol, (c) が  $\Delta G < 5$  kJ/mol, (d) が  $\Delta G < 20$  kJ/mol の面積である. 紫線が EF1, 緑線が EF2 を示す.

### 3.1.4. 着目した各レプリカへの遷移の様子

前節では, 温度による自由エネルギー地形の変化を見た. ここでは, 常温より温度が数十 K 程度まで高いレプリカ間で比較をすることで, EF1 と EF2 の違いを見る. そこで, 300 K, 310 K (室温), 322.5 K, 343 K に着目し, 解析を行った. 図 3.1.6 は解析に用いたレプリカへ遷移してきた元々のレプリカの様子を示している. 赤色は 300 K, ピンク色は 310 K, 黒色は 322.5 K, 青色は 343 K のレプリカに遷移してきた元々のレプリカのインデックスを示す. この図から解析に用いたレプリカが色々な温度空間 (エネルギー空間) からサンプリングできていることがわかる.

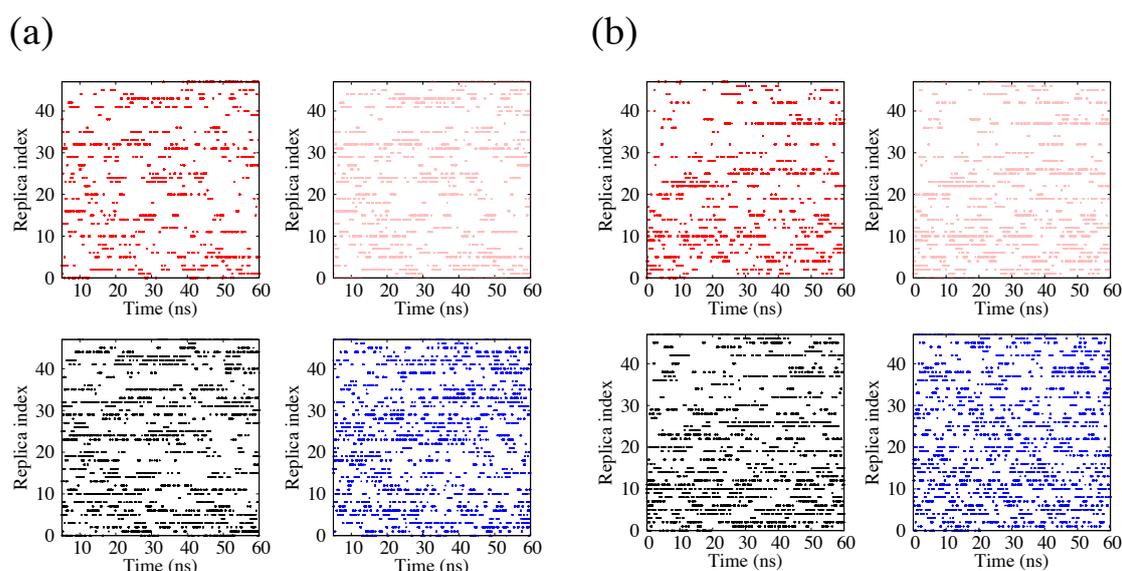


図 3.1.6 解析に用いたレプリカへ遷移してきた元々のレプリカインデックス。(a) が EF1, (b) が EF2 の解析に用いたレプリカへ遷移してきた様子を示す。ここで、赤色は 300 K, ピンク色は 310 K, 黒色は 322.5 K, 青色は 343 K である。

### 3.1.5. 自由エネルギー地形

EF1 および EF2 を比較するために, 300 K, 310 K, 322.5 K, 343 K の自由エネルギー地形のみを選び, 図 3.1.7 (EF1) および図 3.1.8 (EF2) に示した。ここで, 自由エネルギー地形のグローバルミニマム近傍を自由エネルギー差  $\Delta G$  が約 1.5 kJ/mol 以下と定義した。図中に青色で示してある構造は, グローバルミニマム近傍で得られた構造を示してある。水色で示してある構造は, 約 1.5-10.0 kJ/mol の範囲にある構造である。グローバルミニマム近傍で得られた構造は, ヘアピン様構造をしており, 二次構造として  $\beta$ -シート構造を持つものが多かった。高い温度 (343 K) において, EF2 の C 末端側に  $\alpha$ -ヘリックスを持つ構造も得られた (図 3.1.8d)。EF1 は温度変化によるグローバルミニマム近傍の構造に大きな違いは見られなかった。

グローバルミニマム周辺を自由エネルギー差  $\Delta G$  が約 2.5 kJ/mol 以下と定義し, その定義の元で, グローバルミニマム周辺における, RMSD および  $R_g$  のその範囲を表 4 に示した。表より, EF1 はグローバルミニマム周辺における温度による RMSD および  $R_g$  の範囲の変化は小さいことがわかった。EF2 の場合は, 310 K と 322.5 K の間で, グローバルミニマム周辺における, RMSD および  $R_g$  の温度による範囲の変化が著しいことがわかった。これらのことから, EF2 は温度変化により構造が変化しやすく, 広がった構造の頻度が高くなっていることがわかる。

次に自由エネルギー地形上のエネルギーが高い(頻度が低い)領域に焦点を当てた。RMSD,  $R_g$ ともに大きい値をとる領域において、構造はアンフォールディング状態となっている。EF2 においては、頻度は下がるが、図 3.1.8 に示した各温度の自由エネルギー地形でアンフォールディング状態の構造が得られた。これは、EF2 だけでなく EF1 でも得られるが、EF1 の場合、温度が高い条件でようやく得ることができた。アンフォールディング状態の構造は図 3.1.7 と図 3.1.8 にそれぞれ、赤色の構造で示されている。これは RMSD,  $R_g$ ともに大きい値をとる領域であるが、RMSD の値は大きい、 $R_g$  の値は小さいという領域も自由エネルギー地形上で見られた。その領域は EF1 においては、温度上昇に伴い、その領域の頻度は低くなっていった。一方、EF2 では温度上昇に伴い、その領域の頻度は高くなっていった。これらのことから、EF2 は小さい温度変化で異なった構造へ変化しやすいことが示唆される。

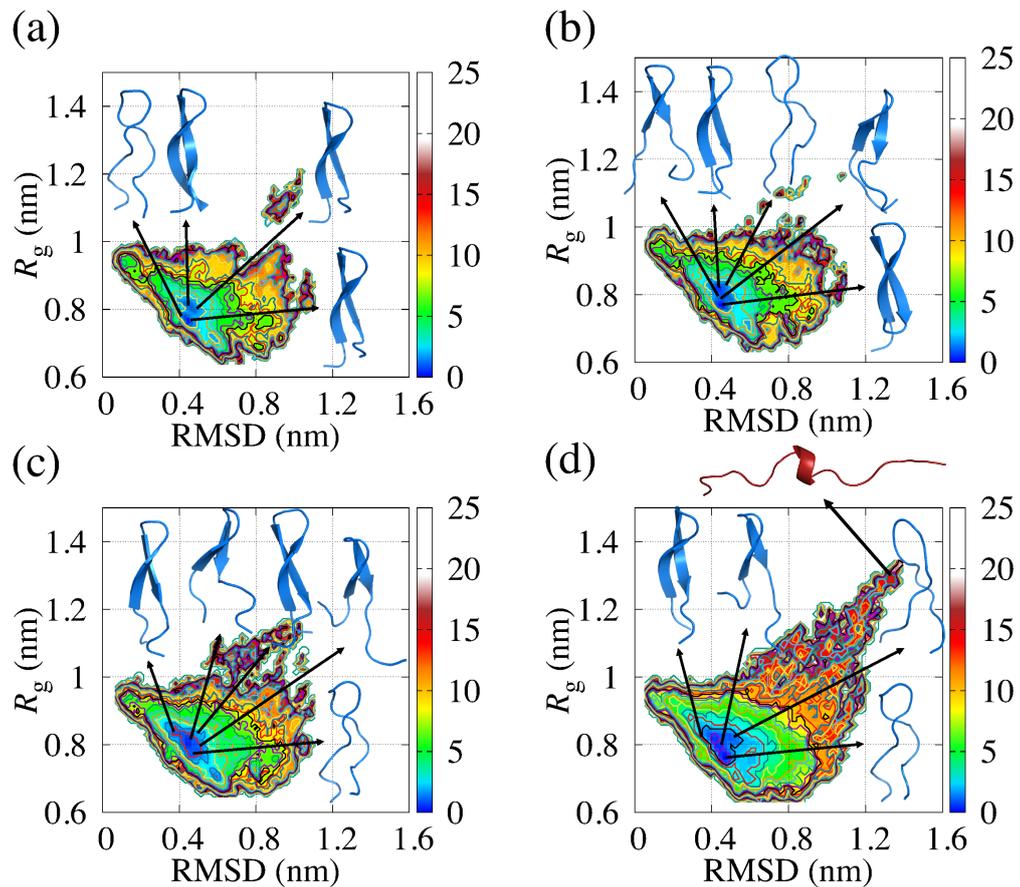


図 3.1.7 頻度分布から得た EF1 の自由エネルギー地形。それぞれの自由エネルギー地形は 300 K (a) , 310 K (b) , 322.5 K (c) , 343 K (d) のものを示す。

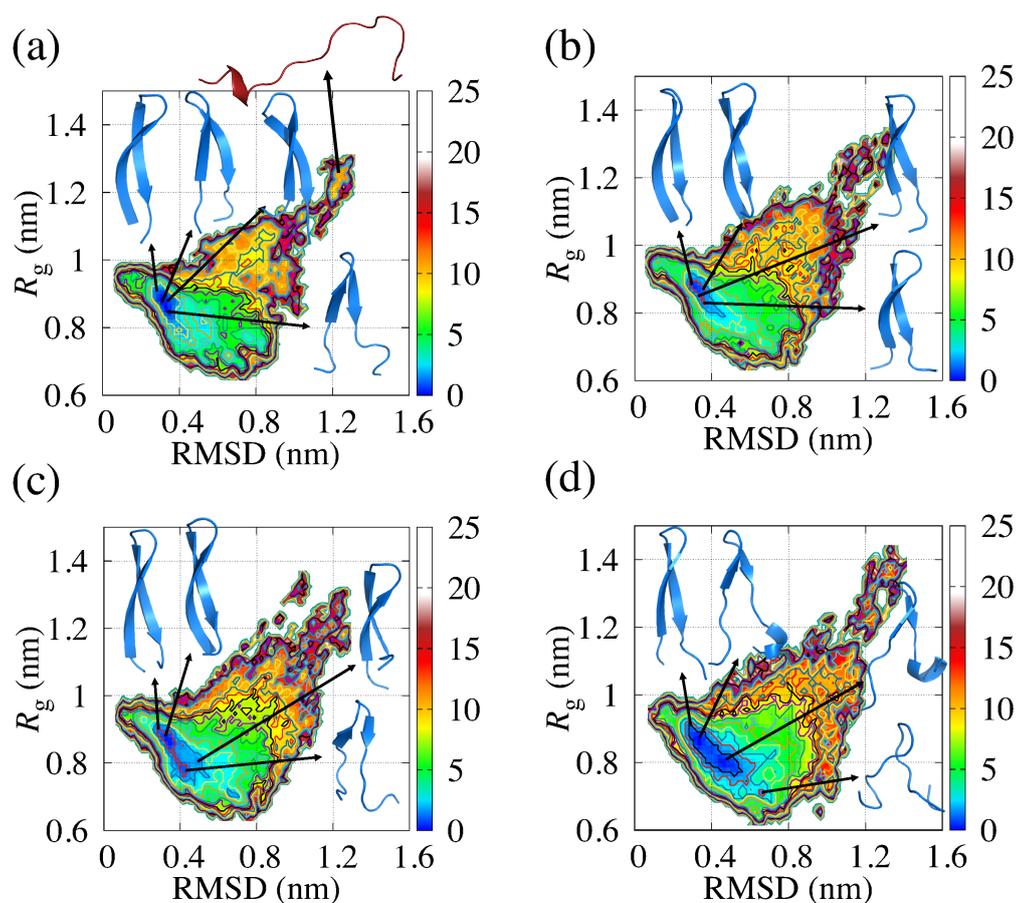


図 3.1.8 頻度分布から得た EF2 の自由エネルギー地形. それぞれの自由エネルギー地形は 300 K (a), 310 K (b), 322.5 K (c), 343 K (d) のものを示す.

表 4 グローバルミニマム周辺の RMSD および  $R_g$  の範囲

	Temperature (K)	RMSD (nm)	$R_g$ (nm)
EF1	300	0.404-0.539	0.702-0.840
	310	0.368-0.602	0.690-0.850
	322.5	0.331-0.613	0.696-0.863
	343	0.287-0.607	0.688-0.880
EF2	300	0.300-0.659	0.719-0.931
	310	0.283-0.468	0.791-0.916
	322.5	0.277-0.707	0.700-0.911
	343	0.261-0.667	0.706-0.917

## 3.2. 通常の分子動力学シミュレーション

前節では拡張アンサンブルのひとつである、レプリカ交換法に基づいたシミュレーションを行った。しかし、レプリカ交換法だとそのままでは時間に依存した量を求めることができない。例えば、レプリカ交換法により得られた構造で、ある原子間に水素結合が形成されることが分かった場合、その水素結合が時系列に沿って連続で存在するかはわからない。ここでは、EF1 および EF2 の水溶液中におけるダイナミクスを調べるため、等温定圧下における通常の分子動力学シミュレーションを行い、その解析結果を示す。

### 3.2.1. 通常の分子動力学シミュレーションにより得られた軌

#### 跡の解析

等温定圧下におけるシミュレーションを 2  $\mu$ s 行った。通常の分子動力学シミュレーション中において、EF1 および EF2 が結晶構造より、どれだけ構造が変化したかを調べるため、主鎖（窒素、不斉炭素、炭素、酸素）の RMSD を計算した。また自由エネルギー地形を計算するため、レプリカ交換法で解析したときと同様に慣性半径  $R_g$  を計算した。それぞれの平均値を表 5 に示した。それぞれの平均値に大きな違いは見られなかった。

通常の分子動力学シミュレーションで得られた軌跡がレプリカ交換法より得られた自由エネルギー地形上をどの程度網羅しているかを調べるために、通常の分子動力学シミュレーションで得られた軌跡を RMSD と  $R_g$  を反応座標として自由エネルギー地形を描いた図上にプロットした。その結果を図 3.2.1 に示す。図より、通常の分子動力学シミュレーションの軌跡は、自由エネルギーが比較的低い領域を推移しているように見える。図 3.2.1a より、EF1 の軌跡は、図 3.1.7 に示されているグローバルミニマムより、少し離れた反応座標上にトラップされてしまっている様子が見られる。図 3.2.1 に示すそれぞれの矢印は、レプリカ交換法により得られた自由エネルギー地形上ではエネルギーが比較的低い領域（構造の頻度が高い）であるにも関わらず、通常の分子動力学シミュレーションの時間スケール（2  $\mu$ s）では到達しなかった領域を示している。

表 5 RMSD および  $R_g$  の平均値と標準偏差.

	RMSD		$R_g$	
EF1	0.531 (0.085)	0.455 (0.133)	0.793 (0.049)	0.826 (0.058)
EF2	0.454 (0.129)	0.445 (0.109)	0.829 (0.069)	0.845 (0.062)

括弧内の値は標準偏差. それぞれの値の単位は nm である.

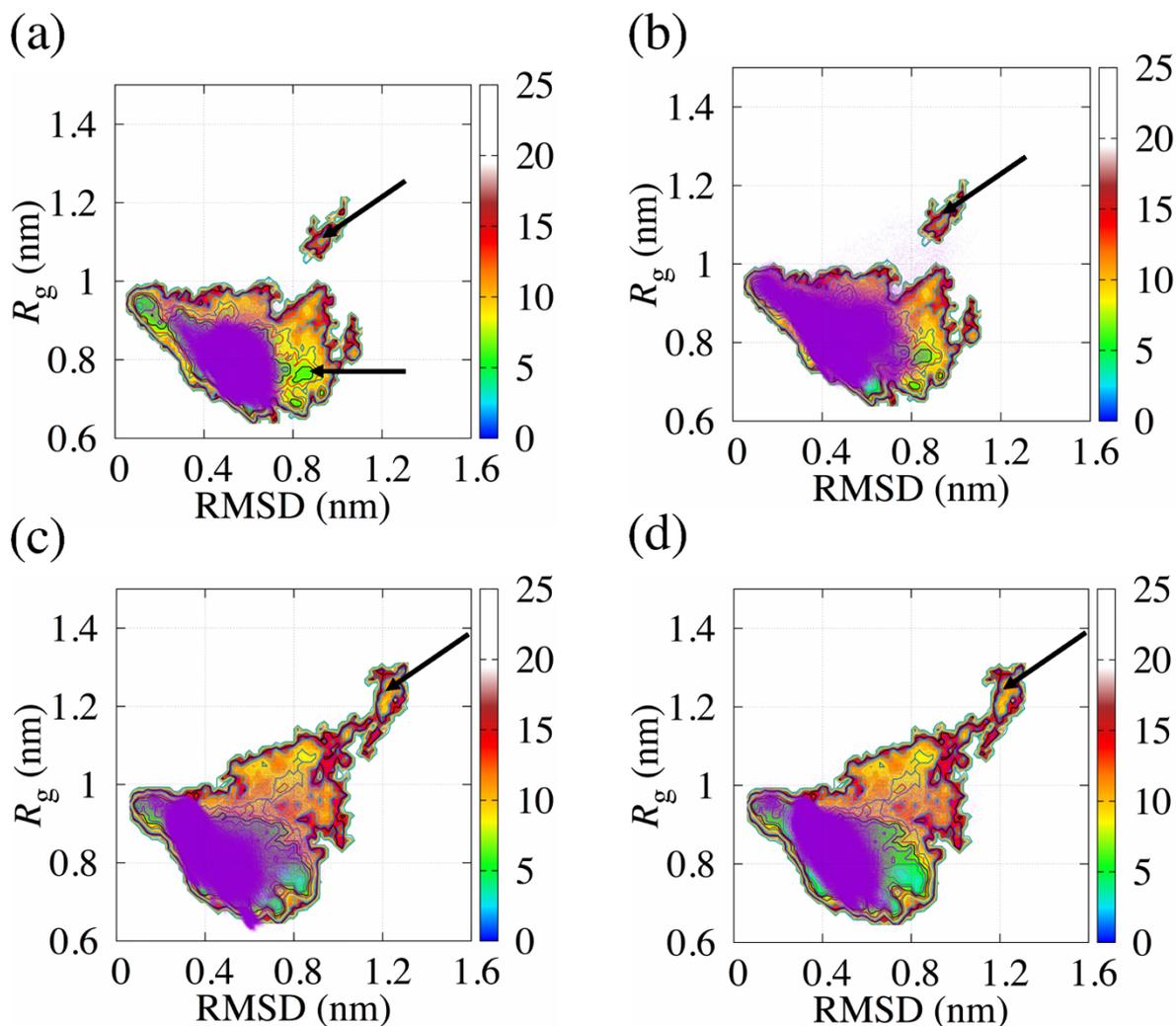


図 3.2.1 レプリカ交換法により得られた自由エネルギー地形上における, 通常の分子動力学シミュレーションの軌跡. (a) と (b) は EF1 の軌跡, (c) と (d) は EF2 の軌跡を示す. 紫色の点が通常の分子動力学シミュレーションの軌跡である. 矢印は通常の分子動力学シミュレーションの時間スケールでは得られなかった領域を示す.

### 3.2.2. 自由エネルギー地形による評価

通常の分子動力学シミュレーションにおける EF1 および EF2 のローカルミニマム構造を得るために、前節の反応座標を用いて、通常の分子動力学シミュレーションの軌跡での頻度分布に基づく自由エネルギー地形を計算した。各シミュレーションにおける自由エネルギー地形を計算した結果を図 3.2.2 に示す。EF1 は 2 つのシミュレーションで、自由エネルギー地形が異なっている。第一のシミュレーションでは、大きなローカルミニマムを 1 つ、第二のシミュレーションでは、ローカルミニマムを 2 つ持っていた (図 3.2.2a, b)。EF1 のローカルミニマム A, B の自由エネルギー差  $\Delta G$  は約 1.5 kJ/mol 以下であった。EF2 において、第一のシミュレーションでは、3 つのローカルミニマムが見られ (図 3.2.2c)、第二のシミュレーションでは、1 つのローカルミニマムが得られた (図 3.2.2d)。図 3.2.2c において、ローカルミニマム A, B の  $\Delta G$  は約 1.5 kJ/mol 以下であった。また、ローカルミニマム C の  $\Delta G$  は約 2.0 kJ/mol 以下であった。EF1 におけるローカルミニマムは比較的近い領域に集まっているが、EF2 のローカルミニマムは異なる領域、広くに存在していた。これは、ローカルミニマムにおいて、EF2 の構造が変化しやすいことを示唆する。図 3.2.2 に示している構造はローカルミニマム付近で得られた構造である。得られた構造のそれぞれが、ヘアピン様構造を形成しており、二次構造として  $\beta$ -シート構造を持っていた。

EF1 の 2 つシミュレーションを比較すると、図 3.2.2a の結果は反応座標の範囲が狭く、構造がトラップされていることを示している。これは、EF1 がローカルミニマム A の構造を持つ領域にトラップされたためと考えられる。また、EF1 は結晶構造に近い構造の頻度も高かった (図 3.2.2b)。一方、EF2 では結晶構造に近い構造の頻度は低かった (図 3.2.2c, d)。

EF1 および EF2 とともに、レプリカ交換法により得られた自由エネルギー地形におけるグローバルミニマム周辺を揺らいでいることが示唆される。

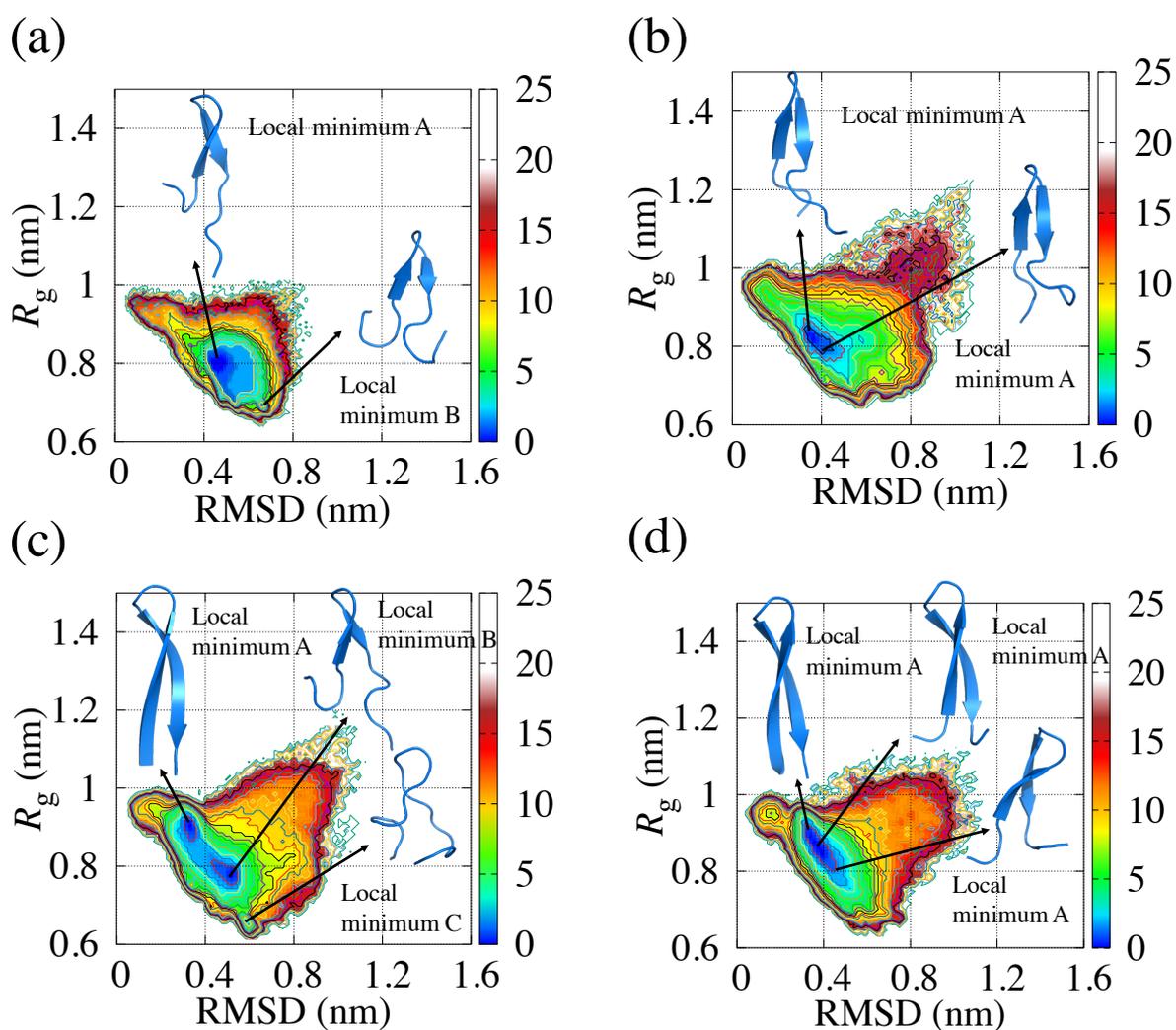


図 3.2.2 通常の分子動力学シミュレーションより得られた EF1 および EF2 の自由エネルギー地形. (a) と (b) は EF1, (c) と (d) は EF2 の自由エネルギー地形を示す. 示されている構造は, 各地形のローカルミニマム付近より得られた構造である.

### 3.2.3. 根平均二乗揺らぎ (RMSF) による評価

EF1 および EF2 の動的なふるまいを調べるため、RMSF を計算した。対象は各アミノ酸残基の不斉炭素子  $C_\alpha$  とし、式 (2.84) を用いて計算を行った。その結果を図 3.2.3 に示した。図 3.2.3a は EF1 と EF2 の各シミュレーションにおける、 $C_\alpha$  についての RMSF の値を示している。赤と緑線は EF1 の結果を示し、黒と青線は EF2 の結果を示している。横軸は各アミノ酸残基に対応している。図 3.2.3b と図 3.2.3c は N 末側と C 末側のストランドに残基を分けて  $C_\alpha$  についての RMSF の計算結果を示している。線および横軸の対応は図 3.2.3a と同様である。これらの図から、EF1 は EF2 と比較して揺らぎが小さいことがわかる。ストランドごとに分けて計算した結果より、EF2 の N 末端側のストランドの揺らぎが大きいことがわかった。

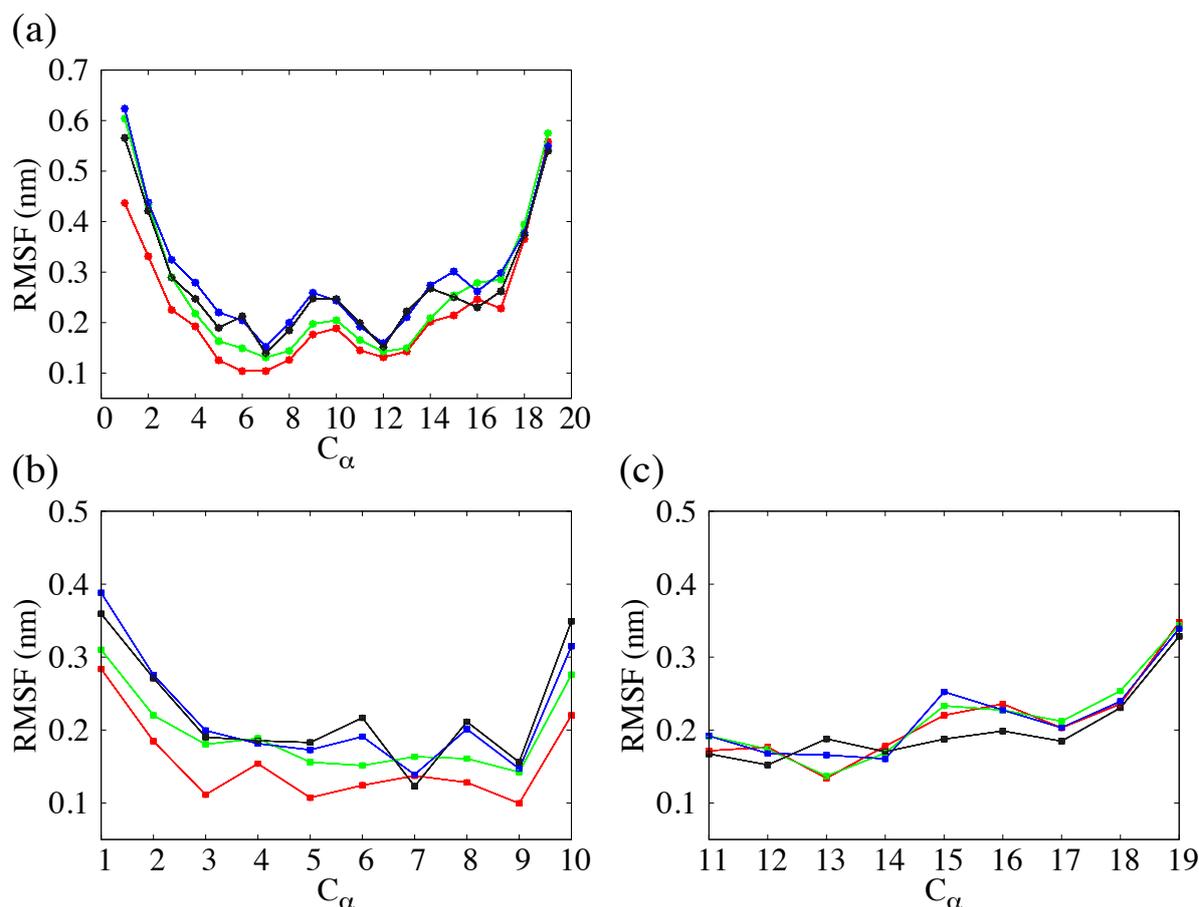


図 3.2.3 EF1 および EF2 の RMSF 値。(a) は EF1 と EF2 の各シミュレーションにおける RMSF の値を示している。(b) と (c) は EF1 と EF2 の 1-10 番目の残基 (N 末端側ストランド) と 11-19 番目の残基 (C 末側ストランド) に分けた RMSF の値を示している。赤と緑線が EF1, 青と黒線が EF2 の結果を示す。

### 3.2.4. 溶媒露出面積による評価

溶媒露出面積 (Solvent Accessible Surface Area: SASA) [77]はタンパク質を構成する原子に対して、溶媒が接触可能な曲面の面積である。接触可能かどうかはファンデルワールス半径に基づいている。図 3.2.4 に SASA のイメージを示した。溶媒分子を球のプローブとして、タンパク質を構成する原子の周りを動かしていき、その曲面の面積を計算する。点線はプローブの中心を通り、接触可能な表面を示している。溶媒が水の場合、プローブのサイズは半径 1.4 Å とする。

タンパク質において、天然構造から変性構造に遷移した時、SASA は増大する。これは、タンパク質は折り畳まれコンパクトな構造をしており、内側の原子と溶媒は接触できず、変性することにより構造が変化し、内側にあった原子が露出し、接触可能な原子が増えるためである。このことから、SASA の計算は構造の相違の指標として用いることができる。表 6 にシミュレーションを行う前であるエネルギー最小化後の構造における SASA の値とシミュレーション後の平均値を示す。EF1 はエネルギー最小化後の値と平均値の値に大きな差は見られなかったが、EF2 は平均値がエネルギー最小化後の値に対して増加していることがわかった。

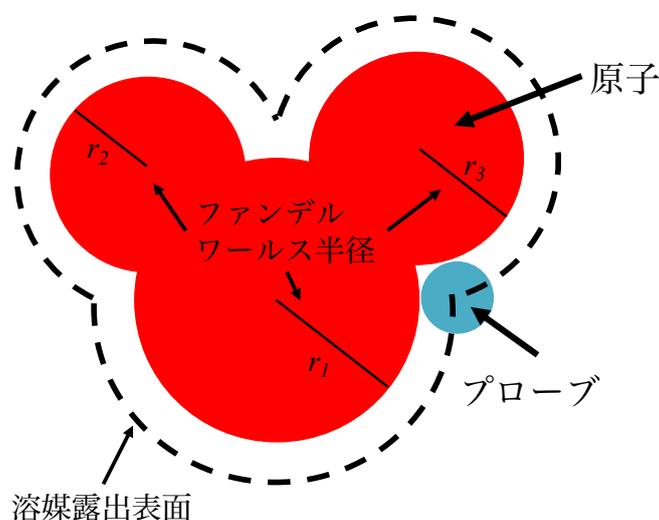


図 3.2.4 溶媒露出面積 (Solvent Accessible Surface Area: SASA)

表 6 EF1 および EF2 の溶媒露出面積

SASA (nm <sup>2</sup> )	After energy minimization	Average values	
EF1	7.992	7.896 (0.517)	7.913 (0.586)
EF2	7.695	8.307 (0.606)	8.418 (0.554)

### 3.2.5. 非共有結合性相互作用

一般に、非共有結合性相互作用には水素結合、疎水性相互作用、イオン結合が挙げられる。これらの相互作用はタンパク質やペプチドの構造形成や維持において重要な役割をしている。そこで、非共有結合性相互作用に焦点をあて、EF1 および EF2 の解析を行った。

#### 3.2.5.1. シミュレーション中における水素結合の評価

水素結合はタンパク質の安定化や二次構造の形成において重要な相互作用の一つである。ここでは、シミュレーションにより得られた軌跡を用いて、EF1 および EF2 のシミュレーション中における水素結合の評価を行った。

#### 3.2.5.2. 水素結合の基準

水素結合の存在は図 3.2.5 に示すような  $\theta, r$  を用いて、次の基準で計算した。

$$\begin{aligned} r &\leq r_{\text{HB}} = 0.35 \text{ nm} \\ \theta &\leq \theta_{\text{HB}} = 30^\circ \end{aligned} \quad (3.2)$$

$r_{\text{HB}} = 0.35 \text{ nm}$  の値は、SPC water モデルの動径分布関数の最初の極小値 ( $r = 0$  から増加させた時の) に一致する。

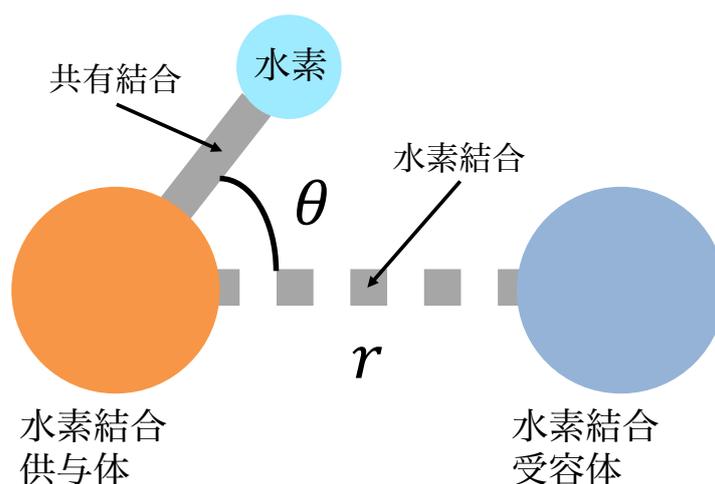


図 3.2.5 水素結合の幾何学的な基準。水素結合供与体と水素結合受容体間の結合が水素結合であり、水素結合供与体と水素間の結合は共有結合である。

### 3.2.5.3. 水素結合の寿命

水素結合の寿命は水素結合が存在するかどうかの関数 (0 か 1) における自己相関関数から計算することができる。

$$C(\tau) = \langle s_i(t)s_i(t+\tau) \rangle \quad (3.3)$$

上述した式における,  $s_i(t) = \{0, 1\}$  はある時刻  $t$  における  $i$  番目の水素結合が存在するかどうかの関数である.  $\tau$  はシミュレーション時の構造の保存間隔によって変わるので, もっと細かい時間間隔で相関を見たい場合は構造の保存間隔を短くすればいい. そして, 水素結合の寿命  $\tau_{\text{HB}}$  は  $C(\tau)$  を積分することで水素結合のおおよその寿命が得られる.

$$\tau_{\text{HB}} = \int_0^{\infty} C(\tau) d\tau \quad (3.4)$$

となる.

### 3.2.5.4. EF1 および EF2 の水素結合数および距離の計算

表 7, 図 3.2.6 に示した水素結合のペアについて, 水素結合を持つかどうかの判定を行った. これらのペアは EF1 と EF2 において  $\beta$ -シート構造を特徴づけるものであり, ラミニン LG4 モジュール内における結晶構造においても形成されている. 水素結合の判定は, 図 3.2.5 および式 (3.2) に示した条件で行った.

表 7 EF1 および EF2 の  $\beta$ -シート構造を特徴づける水素結合の位置.

Item	Pairs of H-bonds (EF1)			Pairs of H-bonds (EF2)		
HB1	TYR <sub>2</sub> -NH	•••	OC-ASP <sub>17</sub>	PHE <sub>2</sub> -NH	•••	OC-ASP <sub>17</sub>
HB2	TYR <sub>2</sub> -CO	•••	HN-ASP <sub>17</sub>	PHE <sub>2</sub> -CO	•••	HN-ASP <sub>17</sub>
HB3	THR <sub>4</sub> -NH	•••	OC-MET <sub>15</sub>	THR <sub>4</sub> -NH	•••	OC-SER <sub>15</sub>
HB4	THR <sub>4</sub> -CO	•••	HN-MET <sub>15</sub>	THR <sub>4</sub> -CO	•••	HN-SER <sub>15</sub>
HB5	GLN <sub>6</sub> -NH	•••	OC-HIS <sub>13</sub>	GLN <sub>6</sub> -NH	•••	OC-TYR <sub>13</sub>
HB6	GLN <sub>6</sub> -CO	•••	HN-HIS <sub>13</sub>	GLN <sub>6</sub> -CO	•••	HN-TYR <sub>13</sub>
HB7	GLN <sub>8</sub> -NH	•••	OC-ARG <sub>11</sub>	ARG <sub>8</sub> -NH	•••	OC-PHE <sub>11</sub>
HB8	GLN <sub>8</sub> -CO	•••	HN-ARG <sub>11</sub>	ARG <sub>8</sub> -NH	•••	OC-PHE <sub>11</sub>

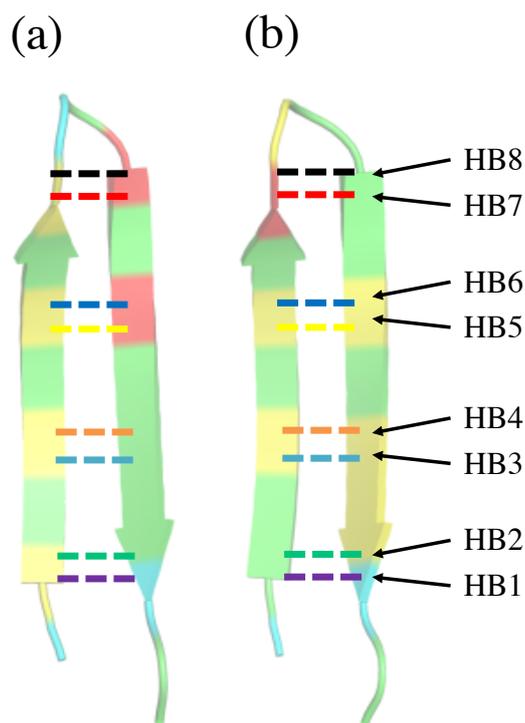


図 3.2.6 EF1 および EF2 の水素結合の位置

各時刻における水素結合数を計算した結果を図 3.2.7 に示した。また、平均の水素結合数を計算した結果を表 8 に示す。水素結合数の時系列は 5 ns 毎にプロットされているが、平均値の計算は 2 ps 毎に保存した構造を用いて行なっている。図および表より、EF1 の水素結合数は、EF2 よりもシミュレーション中、多く維持していることがわかる。EF1 は 1 回目、2 回目のシミュレーション共に、水素結合数が約 4-5 個の間を常に維持していることがわかる。EF2 の水素結合数は約 3-4 個維持していることがわかる。図 3.2.7d で見られるように、シミュレーションの早い段階（100 ns 未満）で水素結合数が減っていることがわかる。図 3.2.7c では、水素結合が減った後、元の水素結合数まで増えているが、500 ns の時点で水素結合数が約 3 個程度まで減っていることがわかる。

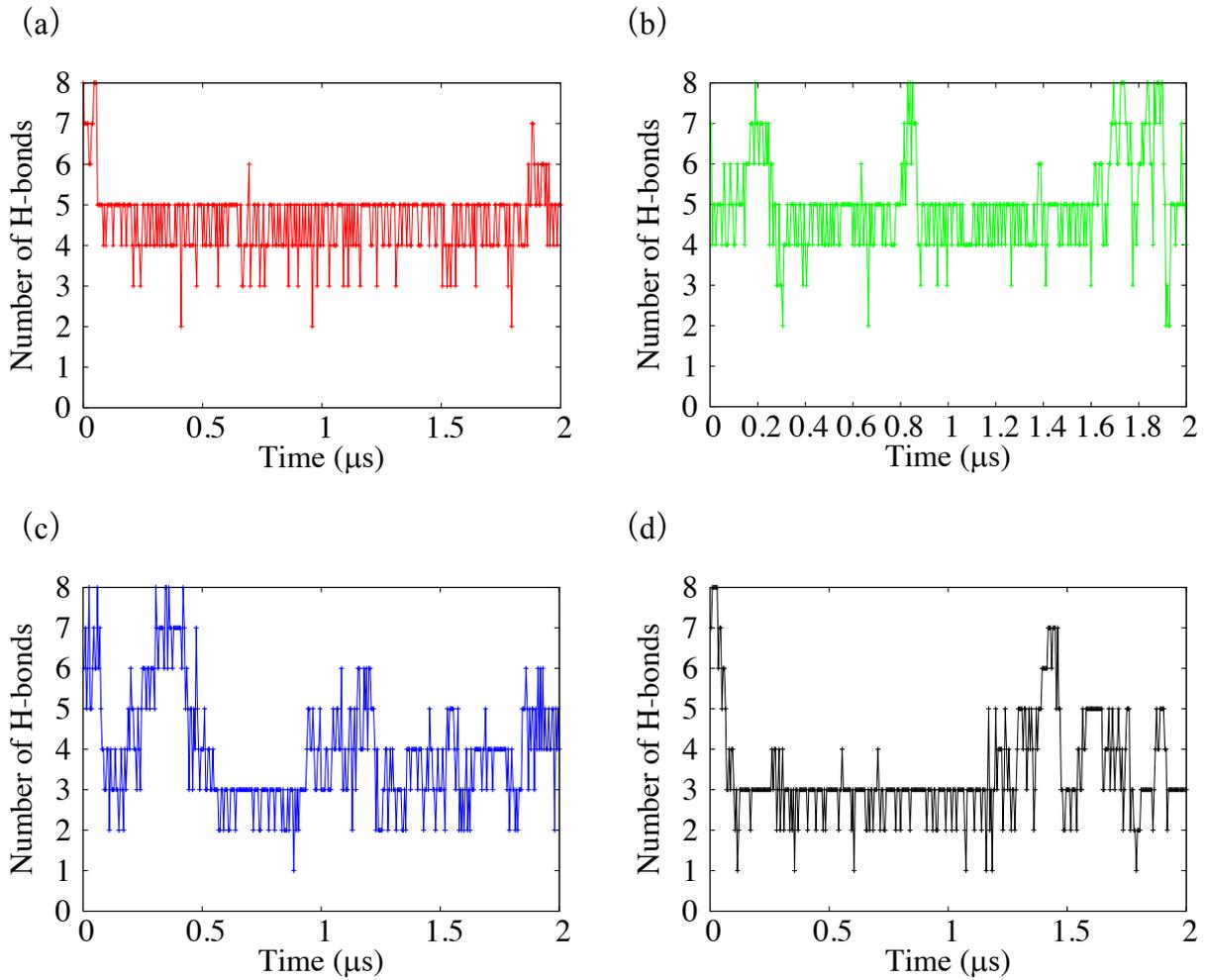


図 3.2.7 時間に伴う水素結合数の変化. (a) と (b) は EF1 の 1 回目と 2 回目のシミュレーションにおける水素結合数の時間変化を示す. (c) と (d) は EF2 の水素結合数の時間変化を示す. それぞれ, 5 ns 毎にプロットしている.

表 8 平均した水素結合数

Values and standard deviation		
	1 回目	2 回目
EF1	4.61 (0.80)	5.04 (1.15)
EF2	4.00 (1.45)	3.39 (1.18)

括弧内の値は標準偏差

次に, 表 7 に示した各ペアの原子間距離を計算した. その結果を, 図 3.2.8 に示した. (a) と (b) は EF1 の 1 回目と 2 回目のシミュレーションにおける各水素結合ペアの原子間距離を示す. (c) と (d) は EF2 における各水素結合ペアの原子間距離を示す. また, 表 9 に各水素結合ペアの原子間距離, それぞれの平均値と標準偏差を示す.

す。EF1 においては、8 個中 5 個のペア一定の距離を維持していた。EF2 は 8 個中 3 個のペアが一定の距離を維持していた。平均値の結果から、約 0.250 nm よりも近い距離で維持しているペアの数は、EF1 において、5 個のペアが維持しており、EF2 においては 3 個のペアであった。ここで、常に水素結合を維持しているペアは  $\beta$ -ターン付近のペアであり、N 末端側および C 末端側のペアは水素結合を形成していないことがわかる。しかし、EF1 は EF2 に比べ、N 末端側および C 末端側に近い水素結合のペア (HB4 と HB5) を多く維持している。

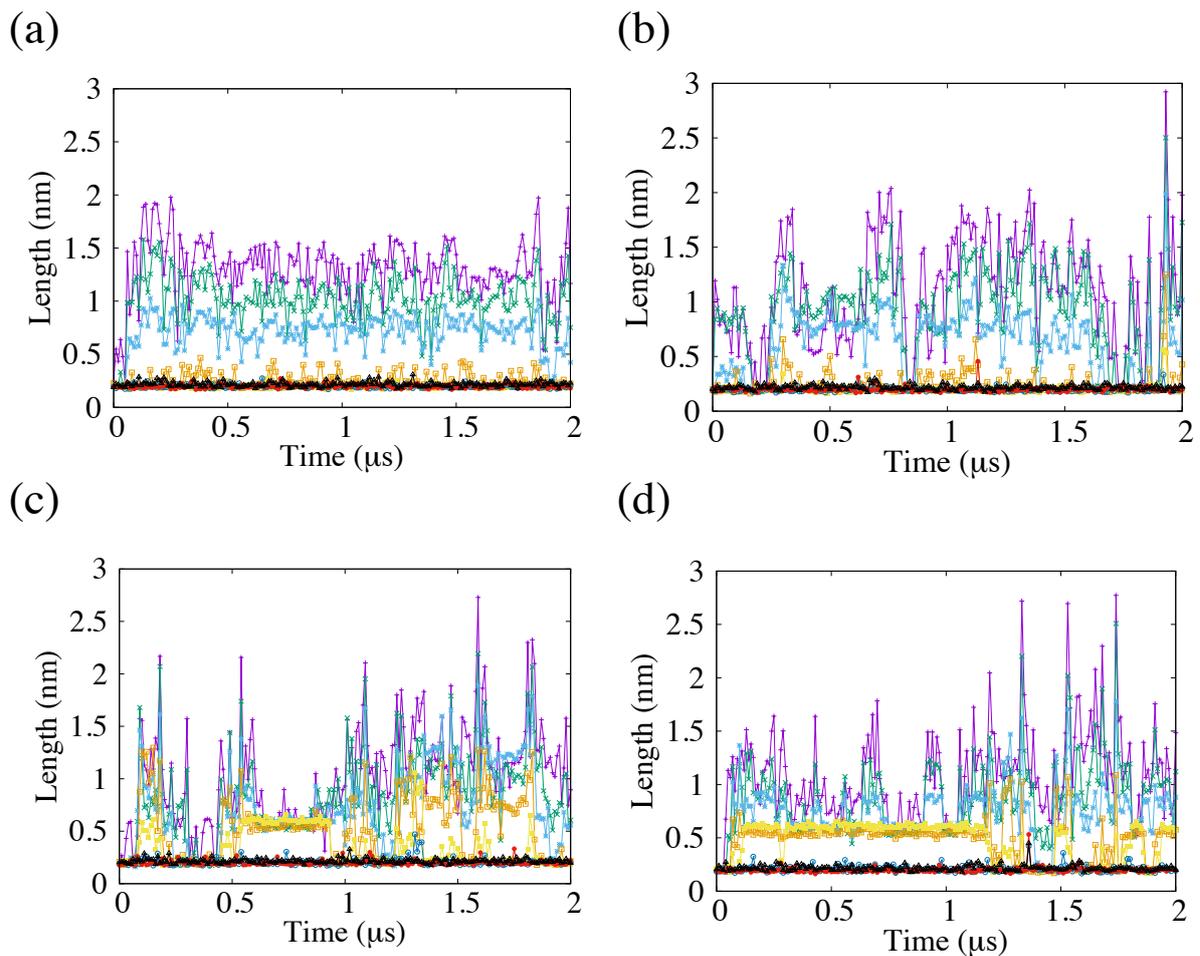


図 3.2.8 各水素結合ペアの原子間距離。(a) と (b) は EF1 の 1 回目と 2 回目のシミュレーションにおける表 7 に示されている各水素結合ペアの原子間距離の時間変化を表す。(c) と (d) は EF2 の各水素結合ペアの原子間距離の時間変化を示す。表 7 に示してある各水素結合のペアを HB1 から HB8 の順番に、紫色、緑色、水色、橙色、黄色、青色、赤色、黒色で表している。

表9 各水素結合ペアの原子間距離の平均値

	EF1_R1	EF1_R2	EF2_R1	EF2_R2
HB1	1.299 (0.302)	1.072 (0.505)	1.038 (0.459)	1.074 (0.397)
HB2	1.022 (0.252)	0.908 (0.399)	0.852 (0.403)	0.857 (0.337)
HB3	0.715 (0.171)	0.623 (0.298)	0.749 (0.396)	0.727 (0.249)
HB4	0.250 (0.070)	0.250 (0.121)	0.533 (0.305)	0.490 (0.188)
HB5	0.193 (0.140)	0.200 (0.053)	0.338 (0.184)	0.462 (0.186)
HB6	0.198 (0.016)	0.202 (0.029)	0.208 (0.038)	0.206 (0.025)
HB7	0.200 (0.023)	0.199 (0.024)	0.200 (0.023)	0.201 (0.030)
HB8	0.216 (0.023)	0.219 (0.025)	0.200 (0.023)	0.215 (0.025)

括弧内の値は標準偏差である。

### 3.2.5.5. 水素結合の自己相関関数

次に、水素結合の自己相関関数を計算した。水素結合の自己相関関数は式 (3.3) から計算した。この計算は、シミュレーション中に指定したペアの水素結合が常に維持しているかどうかを確認するためである。本研究においては、 $\tau$  は 0.2 ps で計算した。計算結果は図 3.2.9 に示した。また、2 回のシミュレーションにおいて、それぞれを平均した結果を図 3.2.10 に示す。これらの結果から、EF1 は HB5-8 のペアが常に水素結合を維持しており、EF2 は HB6-8 のペアが常に水素結合を維持していることがわかる。EF2 は常に水素結合を維持しているペア (HB6-8) かそうでないペア (HB1-5) に分かれていることがわかる。この結果は、水素結合の判定および水素結合のペアの原子間距離を直接計算した結果と一致している。EF1 の自己相関関数の図によると、HB4 の水素結合における自己相関関数がゆっくりと減衰している。これは、HB4 の水素結合のペアは付いたり離れたりを繰り返していることを示す。

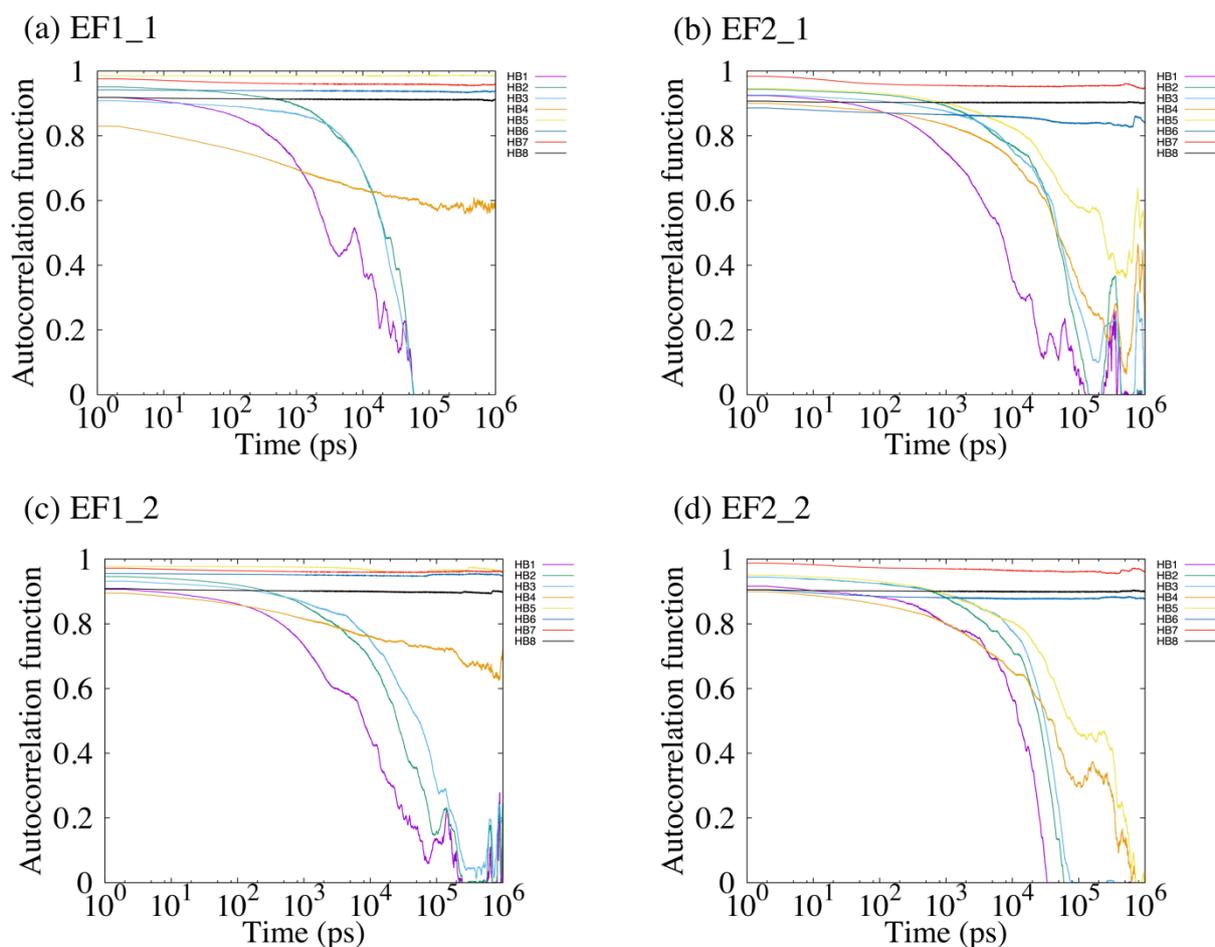


図 3.2.9 水素結合の自己相関関数

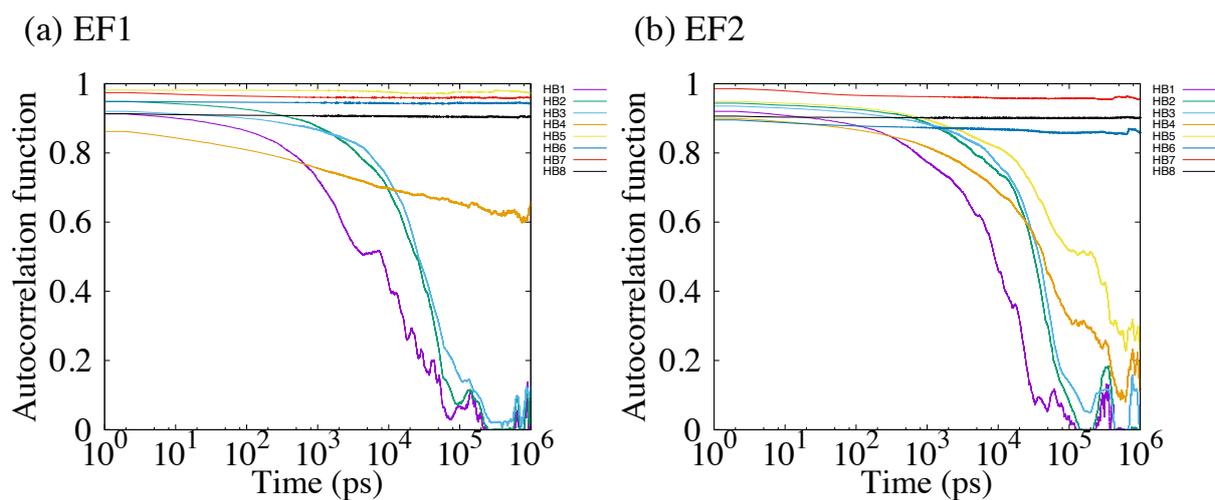


図 3.2.10 2回のシミュレーションにおける平均した水素結合の自己相関関数

### 3.2.5.6. DSSP による二次構造予測

二次構造を予測する DSSP (Define Secondary Structure of Proteins) プログラム[78]を用いて, シミュレーション中の二次構造変化を計算した. DSSP は水素結合のパターンに基づき評価しており, 水素結合の有無は静電相互作用のエネルギーを計算することにより判定が行われている. エネルギーの計算は以下の式に基づく.

$$E(\text{kcal/mol}) = q_1 q_2 \left( \frac{1}{r(\text{ON})} + \frac{1}{r(\text{CH})} - \frac{1}{r(\text{OH})} - \frac{1}{r(\text{CN})} \right) \times f \quad (3.5)$$

ここで,  $q_1 = 0.42e$ ,  $q_2 = 0.20e$ ,  $f = 332$  である. 距離  $r$  の単位は Å である. このエネルギーが  $-0.5 \text{ kcal/mol}$  であるならば, 水素結合が形成していると判定される. 水素結合が伴わない Bend の判定は, 主鎖の  $C_\alpha$  の位置に基づいて行われている.

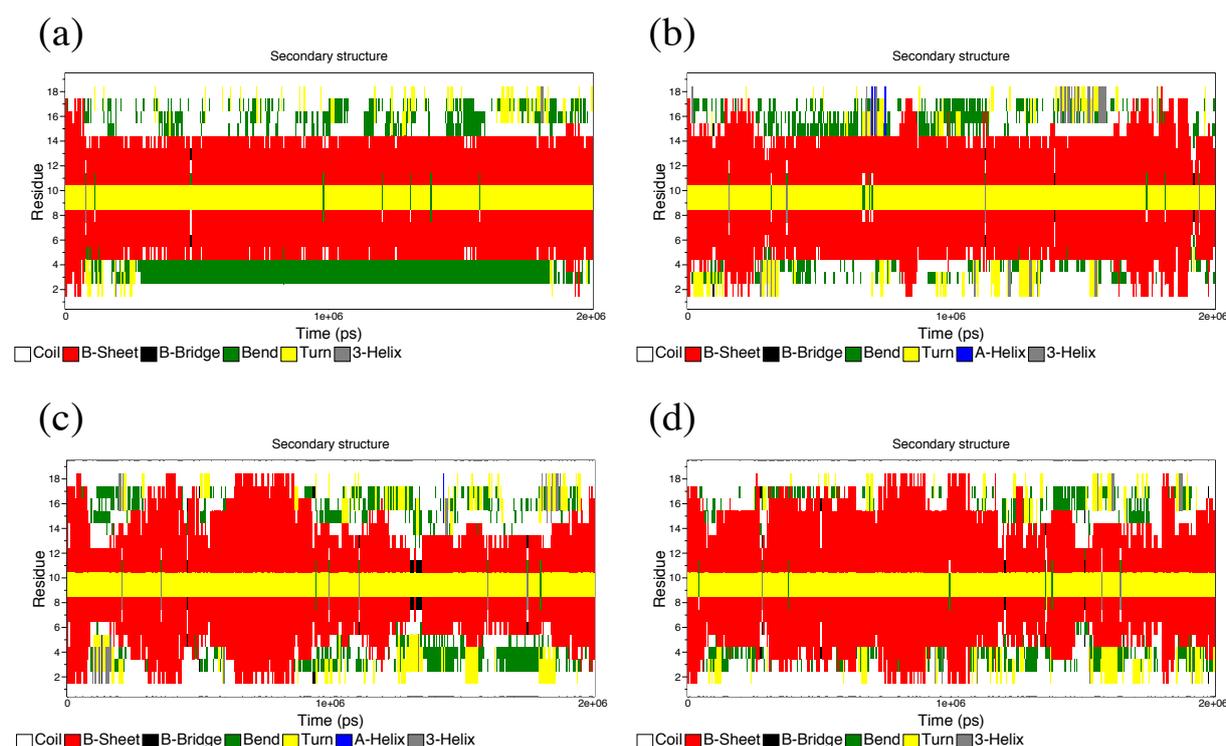


図 3.2.11 DSSP による二次構造予測. (a) と (b) は 1 回目と 2 回目のシミュレーションにおける EF1 の二次構造を予測した結果を示している. (c) と (d) は EF2 の二次構造を予測した結果を示す. Coil は二次構造なし, B-Sheet は  $\beta$ -シート構造, B-Bridge は 2 本のストランドからなる  $\beta$ -シート, Bend は水素結合を伴わないターン, Turn はターン構造, A-Helix は  $\alpha$ -ヘリックス, 3-Helix は 3-ターンヘリックスを表している.

DSSP による時刻毎に二次構造を予測した結果を図 3.2.11 に示した。(a) と (b) は 2 回のシミュレーションそれぞれにおける, EF1 の二次構造を予測した結果を示す。(c) と (d) は EF2 の二次構造を予測した結果を示す。図 3.2.11a, b より, EF1 は二次構造の変化はあまり見られなかった。これにより, 末端以外は常に  $\beta$ -シート構造を維持していることが示唆される。一方, EF2 は二次構造の変化が EF1 に比べ大きかった。DSSP の結果としては  $\beta$ -シート構造を持つことがわかるが, 常に維持しているわけではなかった。

### 3.2.5.7. 極性・非極性アミノ酸残基

EF1 および EF2 のアミノ酸残基毎の性質 (極性・非極性) に着目し, 解析した。EF1 および EF2 のアミノ酸残基それぞれを 4 つの型 (極性電荷アミノ酸, 極性負電荷アミノ酸, 極性無電荷アミノ酸, 非極性アミノ酸 (疎水性アミノ酸)) に分類し, 図 3.2.12 に示した。疎水性相互作用は水溶液中において疎水アミノ酸間で引き合う作用である。EF1 は 2 つのストランド間に疎水アミノ酸が向き合うペアが 3 つあった (図 3.2.12a)。EF2 においては, 疎水性アミノ酸が向き合う 2 つのペアが見られた (図 3.2.12b)。図 3.2.12a より, EF1 の 3 つ疎水性ペアは水素結合のペアと交互に並んでいることがわかる。一方, EF2 は  $\beta$ -ターン側では, 疎水性ペアと水素結合のペアが交互に並んでいるが, 末端側には交互になっていない (図 3.2.12b)。また, EF1 の配列の中心 ( $\beta$ -ターン側) において, 極性電荷アミノ酸と極性負電荷アミノ酸があるため, それらのアミノ酸間でイオン結合を持っている可能性が示唆される。

データベースに登録されているタンパク質の二次構造に対して統計的手法を用いることにより, 二次構造を形成する傾向のあるアミノ酸を解析する研究が行われている [79]。  $\beta$ -シート構造を形成する傾向にあるアミノ酸として, Val, Ile, Tyr があり,  $\beta$ -シート構造を破壊する傾向にあるアミノ酸は Glu, Asp, Pro がある [79]。 EF1 では,  $\beta$ -シート構造を形成する傾向にあるアミノ酸は Tyr, 1 つのみであった。一方, 破壊する傾向にあるアミノ酸は Asp の 2 つであった。 EF2 では,  $\beta$ -シート構造を形成する傾向にあるアミノ酸は Tyr の 2 つと Val の 1 つであった。破壊する傾向にあるアミノ酸は Asp の 2 つと Pro の 1 つであった。 EF1, EF2 ともに  $\beta$ -シート構造の形成, 破壊する傾向にあるアミノ酸はあまり含まれておらず, 特に EF1 においては 3 残基のみであった。 EF1 と EF2 において,  $\beta$ -シート構造を形成, 破壊する傾向にあるアミノ酸の影響は少ないと見られ, 上述したように疎水性ペアや水素結合ペアといった対となる相互作用が重要であり, それらが末端付近に存在することが大事である。タンパク質のデータベースに基づいて得られた二次構造を形成, 破壊する傾向のあるアミノ酸を

用いてペプチドの二次構造を議論する場合、このようなケースがあるため注意する必要がある。

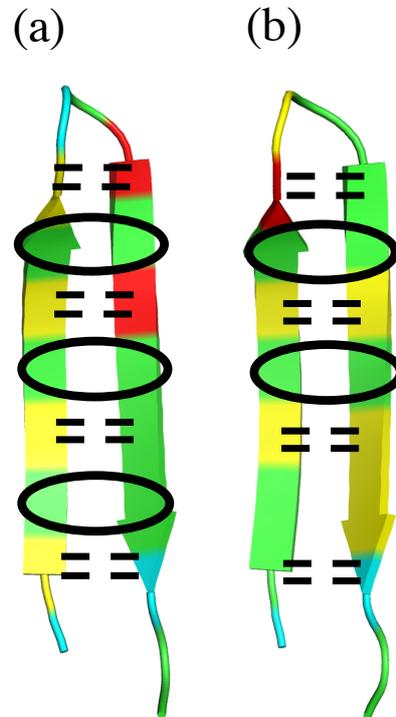


図 3.2.12 EF1 および EF2 のアミノ酸残基の分類. (a) が EF1, (b) が EF2 の分類を示す. 点線は水素結合の位置を示し, 黒丸は疎水性アミノ酸のペアを示す. 赤色は極性電荷アミノ酸, 青色は極性負電荷アミノ酸, 黄色は極性無電荷アミノ酸, 緑色は非極性アミノ酸 (疎水性アミノ酸) を示す.

## 4. まとめと今後の課題

この章では、これまで述べた結果のまとめ及び今後の課題を述べる。

### 4.1. まとめ

EF1 ペプチド (DYATLQLQEGRLHFMDLG) はマウスラミニン  $\alpha 1$  鎖 LG4 モジュール内の E-F スtrandに存在する配列で、 $\alpha 2\beta 1$  インテグリンと相互作用し、細胞接着および細胞伸展を促進する[37]。鈴木らは、EF1 の短縮したペプチド (最小活性配列) を同定し、そのペプチドの末端同士をジスルフィド結合による環状化することにより、細胞接着活性がもとの配列の程度まで回復したことを報告した[37]。また、EF1 の相同配列である、マウスラミニン  $\alpha 2$  鎖 LG4 モジュール内に存在する配列である EF2 ペプチド (DFATVQLRNGFPYFSYDLG) は細胞接着活性を持たないことが報告されている[47]。これらの実験結果から、細胞接着活性とペプチドの構造の関係性が示唆された。以前、シミュレーテッドアニーリング法を用いた分子動力学シミュレーションにより、EF1 および EF2 のグローバルミニマム構造の探索を行った[80, 81]。その結果より、EF1 はヘアピン様構造を持つことを示した。EF2 もまた、ヘアピン様の構造を持っていたが、EF1 に比べ、いびつな構造をしていた。これらの結果は、鈴木らの実験結果と一致している。本研究では、EF1 および EF2 がグローバルミニマムにおける構造を含め、取り得る構造を幅広く網羅する自由エネルギー地形を描くために、拡張アンサンブル法の1つである、レプリカ交換法による分子動力学シミュレーションを行った[81]。また、EF1 および EF2 の水溶液中の動的振る舞いを調べるため、等温定圧下における分子動力学シミュレーションを行った[81, 82]。ここでは、これをレプリカ交換法に対し、通常分子動力学シミュレーションと呼ぶ。

レプリカ交換法による分子動力学シミュレーションは、48 レプリカ (300 K–450.5 K) 用意し、各レプリカ 60 ns の計算を行った。解析は、初期構造の依存を防ぐため、5–60 ns の軌跡を用いて行った。まず、各レプリカの温度変化による自由エネルギー地形の変化を見た。自由エネルギー地形の反応座標は RMSD と慣性半径  $R_g$  を選んだ。その結果を、図 3.1.3 (EF1) および図 3.1.4 (EF2) に示した。EF1 は温度変化により、自由エネルギー地形域は広がっており、その変化は 322.5 K と 343 K の間が顕著であった。また、温度変化によるグローバルミニマム周辺の変化は小さかった。一方、EF2 は、300 K のレプリカにおいて、自由エネルギー地形域は EF1 と比べ広く、温度変化による自由エネルギー地形域は広がりやグローバルミニマム周辺の広がりも顕著であった。特に、322.5 K からのグローバルミニマム周辺の変化が顕著であり、それ以降の温度ではグローバルミニマム周辺はフラットになっていた。このとき、EF2 は変

性状態にあると考えられ、310 K と 322.5 K の間に 1 つの転移温度があると考えられる。これらのことから、EF1 の構造は EF2 に比べ、熱安定性をもつことが示唆され、EF2 の構造は温度変化に伴い、アンフォールドしやすいことが示唆される。

次に、常温を含めた、ある程度低い温度から高い温度、4 つの温度 (300K, 310 K, 322.5 K および 344 K) に着目し解析を行った。その結果は図 3.1.7 に示した。グローバルミニマム近傍より得られた EF1 の構造はヘアピン様構造を示し、多くは二次構造として  $\beta$ -シート構造を持っていた。この結果はシミュレーテッドアニーリング法による結果と一致していることがわかる。一方、EF2 もまた、グローバルミニマム近傍では、ヘアピン様構造を持ち、 $\beta$ -シート構造を持つものや、構造の一部 (C 末端側) に  $\alpha$ -ヘリックスを持つ構造も得られた。

EF1 および EF2 の動的振る舞いを解析するために、等温定圧下で 2  $\mu$ s のシミュレーション (通常の分子動力学シミュレーション) を行った。まず、レプリカ交換法で得られた自由エネルギー地形が通常の分子動力学シミュレーションの軌跡をカバーしているかどうか確認するため、REMD で描いた自由エネルギー地形上に軌跡を RMSD と慣性半径  $R_g$  の反応座標としてプロットした。基本的には、EF1, EF2 ともに自由エネルギー地形のグローバルミニマム周辺で揺らいでいることがわかった。EF1 の 1 つのシミュレーションではグローバルミニマムから少し離れたローカルミニマムに閉じ込められている様子が見られた。このローカルミニマムはグローバルミニマム周辺に含まれており、レプリカ交換法より得られた自由エネルギー地形と矛盾ないことがわかる。レプリカ交換法で得られた自由エネルギー地形からは、RMSD と  $R_g$  の値が大きい領域 (アンフォールディング状態) は得られたが、通常の分子動力学シミュレーションの時間スケールではその領域を観察することはできなかった。

レプリカ交換法により得られた構造を解析し、EF1 は EF2 と比較して、構造変化が小さいことを示した。また、通常の分子動力学シミュレーションにおいて EF1 および EF2 はグローバルミニマム周辺を揺らいでいることを示唆した。改めて、通常の分子動力学シミュレーションで得られた軌跡を用いて、自由エネルギー地形を描いた (図 3.2.2)。EF1 において、ローカルミニマムの構造はヘアピン様構造を示し、二次構造としては  $\beta$ -シート構造を形成していた。EF2 においても同様に、 $\beta$ -シート構造を伴ったヘアピン様構造を形成していた。これらのことから、本研究で行ったシミュレーションの時間スケールにおいて、EF1 および EF2 は  $\beta$ -シート構造を維持していることを示唆した。通常の分子動力学シミュレーションは 2 回行ったが、EF1 は 2 回のシミュレーションで自由エネルギー地形が異なっていた。これは 1 回目のシミュレーションにおいて、EF1 がローカルミニマム A の構造を持つ領域にトラップされてしまったためと考えられる。ローカルミニマムの数は 1 回目 2 つ、2 回目 1 つであった。EF2 においては、自由エネルギー地形に大きな違いはなかった。1 つはローカルミニマム

が3つ、2つはローカルミニマムが2つ見られた。自由エネルギー地形より、EF1は狭い範囲で構造が分布している一方、EF2は広い範囲で構造が分布していることが示された。初めに、EF1およびEF2の通常の分子動力学シミュレーションの軌跡をレプリカ交換法より得られた自由エネルギー地形上にプロットしたが、軌跡の分布の違いは小さく見えた。しかし、頻度分布を計算することによりEF1およびEF2の違いが明らかとなった。よって、違いを示すためにも頻度分布を計算することは重要なことである。

これまで、自由エネルギー地形の結果を示し、静的な構造の違いを見てきた。次にEF1およびEF2の動的な構造の違いを明らかにするため、 $C_{\alpha}$ について、RMSFを計算した。各残基において、EF2はEF1に比べて高い値を示しており、揺らぎやすいことが示され、特に、EF2のN末端側の残基（1から10番目の残基）が揺らいでいることがわかった。また、EF1およびEF2の溶媒露出面積（Solvent Accessible Surface Area: SASA）の計算も行った。SASAの計算は、エネルギー最小化後の構造と、シミュレーションより得られた構造より平均値を算出した。EF1はエネルギー最小化後のSASAの値と平均値の差は約 $0.1 \text{ nm}^2$ と大きな違いはなかった。一方、EF2は約 $0.6\text{--}0.7 \text{ nm}^2$ の違いがあった。EF1およびEF2でのこれらの違いは、構造の多様性に影響し、EF2のN末端側の揺らぎが構造の変動性に影響を与えていると言える。

レプリカ交換法より得られた結果から、EF1において、温度上昇に伴った自由エネルギー地形のローカルミニマム周辺における変化は小さかった。一方、EF2の変化は顕著であった。このことから、EF2はEF1に比べ、構造がアンフォールドしやすいと考えられる。また、EF1は水溶液中において、安定的に $\beta$ -シート構造を形成し、EF2は不安定であると考えられる。 $\beta$ -シート構造の維持に水素結合が重要であると示した。ハットフィールドらもまた、 $\beta$ -シート構造を持つCLN025（安定したシニョリン小型タンパク質）の構造維持において、分子動力学シミュレーションにより、水素結合の重要性を示唆している[84]。サントスらもまた、芳香族骨格をもつTri-ペプチドの構造は水素結合により安定化されていることを報告している[85]。また、いくつかの研究グループ[86-89]が水素結合数と慣性半径を反応座標とした自由エネルギー地形の描像を出している。そこで、通常の分子動力学シミュレーションにおいて、EF1およびEF2の $\beta$ -シート構造がどれだけ維持されているかどうか調べるため、主鎖のN-HとC=Oの間に形成される水素結合数と距離を計算した。EF1およびEF2は、主鎖間に8個の水素結合のペアを持っており、それぞれのペアは表7に示してある。水素結合数はEF1がEF2と比べ、シミュレーション中、多く維持していた。8個ある水素結合の内、HB4とHB5が構造維持において重要な水素結合であるとわかった。EF2において、SASAの平均値がエネルギー最小化後の値と比べ大きな値をとったのは、水素結合数が少なくなり、構造が崩れたためだと考えられる。水素結合とSASAの関係

は、シングらのアミロイドポリペプチドであるアミリン (islet amyloid polypeptide, IAPP) を用いた研究[90]により報告されている。EF1 の末端の揺らぎが EF2 と比べ小さく、以上の結果は EF1 の主鎖間にある水素結合が EF2 より維持されていることを示している。

水素結合の解析に加えて、EF1 と EF2 におけるアミノ酸間の非共有結合の相互作用についても考慮し、比較した。図 3.2.12 に示すように、EF1 は N 末端側と C 末端側のストランド間に 3 つの疎水性ペアを持ち、EF2 は 2 つ持っている。EF1 はストランドを結ぶループ付近に正電荷アミノ酸と負電荷アミノ酸があり、これらのアミノ酸間でイオン間相互作用を持つ可能性がある。よって、EF1 は水素結合を含め 8 対の非共有結合の相互作用を持つことになる。一方、EF2 は 6 対の非共有結合の相互作用を持つことになる。EF2 は末端に疎水性ペアを持たず、末端が揺らぎやすくなっていると考えられる。これらの要因が、EF1 と EF2 の揺らぎやすさの違いおよび、構造の多様性の違いに関連していると考えられる。また、HB4 と HB5 の水素結合の安定性は Ala<sup>3</sup> と Phe<sup>14</sup>間の疎水性相互作用により増加していると思われる。

鈴木らの実験より、EF1 が細胞接着活性を持つことが示され、その活性においてヘアピン様構造の重要性を示唆した。また、片桐らは EF1 が細胞接着活性を持つことに加え、EF2 が細胞接着活性を持たないことを示した[47]。結論として、N 末端と C 末端周辺の構造の揺らぎを抑えることが、 $\beta$ -シート構造 (ヘアピン様構造) を維持する上で重要であり、構造の揺らぎを抑えることが細胞接着活性において重要であると考えられる。また、構造安定性において、非共有結合性の相互作用が重要であり、主鎖間に水素結合を形成することに加え、末端に疎水性アミノ酸のペアを持つことの両方が構造を安定化する上で重要であると結論し、他のペプチドの構造安定性を評価するためには、これらを調査することを提案する。本研究が医薬品や生体材料といった分野への応用として、将来の研究に有益であることを望む。

## 4.2. 今後の課題

本研究において、レプリカ交換法を用いたシミュレーションを、各レプリカ 60 ns ずつ計算を行った。さらに、各レプリカのシミュレーション時間を 100 ns 程度まで増やすことにより、詳細な自由エネルギー地形の描像を得られることが予測できる。その地形から、EF1 および EF2 の構造について新たな知見が得られるかもしれない。また、細胞接着活性を持つペプチドは EF1 以外にも様々あり、様々な細胞接着ペプチドの構造を解析することで、それらに共通する性質がわかるかもしれない。さらに、EF1 は  $\alpha 1\beta 2$  インテグリンと相互作用することがわかっているので、EF1 と  $\alpha 1\beta 2$  インテグリンの結合をアンブレラサンプリング法により評価する。

## 謝辞

本研究の推進ならびに本論文の作成につきまして、御指導、御鞭撻を賜りました本学東京薬科大学生命物理科学研究室の教授高須昌子先生，講師森河良太先生，助教宮川毅先生に対しまして，心より感謝申し上げます。

また，共同研究でお世話になりました本学病態生化学教室の教授野水基義先生，同教室の准教授吉川大和先生，講師保住建太郎先生，助教片桐文彦先生に対しまして，心より感謝申し上げます。

また，生命物理科学研究室が研究棟3号館1階に所在している際，お世話になりました分子生物化学研究室内の助教尹永淑先生とは日ごろからよくして頂き，感謝申し上げます。

最後に，筆者の研究生活に関わりました多くの皆様に感謝申し上げます。

## 参考文献

1. Singer, J. C., Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science*, 1972, 175, 720–731.
2. Alder, B. J., Wainwright, T. E. Phase transition for a hard sphere system. *J. Chem. Phys.* 1957, 27, 1208–1209.
3. Rahman, A. Correlations in the motion of atoms in liquid. *Phys. Rev.* 1964, 136, 405–411.
4. Feig, M., Sugita, Y. Reaching new levels of realism in modeling biological macromolecules in cellular environments. *J. Mol. Graph. Model.* 2013, 45, 144–156.
5. Nilges, M., Clore, G. M., Gronenborm, A. M. Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. *FEBS Lett.* 1988, 239, 129–136.
6. Berg, B. A., Neuhaus, T. Multicanonical ensemble: A new approach to simulate first-order phase transitions. *Phys. Rev. Lett.* 1992, 68, 9–12.
7. Lyubartsev, A. P., Martsinovski, A. A., Shevkunov, S. V., Vorontsov-Velyaminov P. N. New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.* 1992, 96, 1776–1783.
8. Marinari, E., Parisi, G. Simulated tempering: A new Monte Carlo Scheme. *Europhys. Lett.* 1992, 19, 451–458.
9. Okabe T, Kawata M, Okamoto Y, Mikami M. Replica-exchange Monte Carlo method for the isobaric-isothermal ensemble. *Chem. Phys. Lett.* 2001, 335, 435–439.
10. Metropolis, N., Rosenbluth, W. A., Rosenbluth, N. M., Teller, H. A., Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 1953, 21, 1087–1092.
11. Sasaki, T., Fassler, R., Hohenester, E. J. Laminin: the crux of basement membrane assembly. *J. Cell Biol.* 2004, 164, 959–963.
12. Colognato, H., Yurchenco, P. D. Form and function: the laminin family of heterotrimers. *Dev. Dyn.* 2000, 218, 213–234.
13. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J. D. *Molecular biology of the cell* 5th ed. Garland Publishing Inc. New York. 2007, pp 1167.
14. Li, S., Edgar, D., Fassler, R., Wadsworth, W., Yurchenco, P. D. The role of laminin in embryonic cell polarization and tissue organization. *Dev. Cell.*, 2003, 4, 613–624.
15. Miner, J. H., Yurchenco, P. D. Laminin functions in tissue morphogenesis. *Annu. Rev. Cell Dev. Biol.*, 2004, 20, 255–284.

16. Burgeson, R. E., Chiquet, M., Deutzmann, R., Ekblom, P., Engel, J., Kleinman, H. K., Martin, G. R., Meneguzzi, G., Paulsson, M., Sanes, J., Timpl, R., Tryggvason, K., Yamada, Y., Yurchenco, P. D. A new nomenclature for the laminins. *Matrix Biol.* 1994, 14, 209–211.
17. Miner, J. H., Patton, B. L., Lentz, S. I., Gilbert, D. J., Snider, W. D., Jenkins, N. A., Copeland, N. G., Sanes, J. R. The laminin  $\alpha$  chains: expression, developmental transitions, and chromosomal locations of  $\alpha 1$ -5, identification of heterotrimeric laminins 8-11, and cloning of a novel  $\alpha 3$  isoform. *J. Cell Biol.* 1997, 137, 685–701.
18. Iivanainen, A., Morita, T., Tryggvason, K. Molecular cloning and tissue-specific expression of a novel murine laminin  $\gamma 3$  chain. *J. Biol. Chem.* 1999, 274, 14107–14111.
19. Libby, R. T., Champlaud, M. F., Claudepierre, T., Xu, Y., Gibbons, E. P., Koch, M., Burgeson, R. E., Hunter, D. D., Brunken, W. J. Laminin expression in adult and developing retinae: evidence of two novel CNS laminins. *J. Neurosci.* 2000, 20, 6517–6528.
20. Aumailley, M., Bruckner-Tuderman, L., Cater, W. G., Deutzmann, R., Edgar, D., Ekblom, P., Engels, J., Engvall, E., Hohenester, E., Jones, J. C. R., Kleinman, H. K., Marinkovich, M. P., Martin, G. R., Mayer, U., Meneguzzi, G., Miner, J. H., Miyazaki, K., Patarroyo, M., Paulsson, M., Quaranta, V., Sanes, J. R., Sasaki, T., Sekiguchi, K., Sorokin, L. M., Talts, J. F., Tryggvason, K., Uitto, J., Virtanen, I., von der Mark, K., Wewer, U., Yamada, Y., Yurchenco, P. D. A simplified laminin nomenclature. *Matrix Biol.* 2005, 24, 326–332.
21. Ramadhani, D., Tsukada, T., Fujiwara, K., Horiguchi, K., Kikuchi, M., Tashiro, T. Laminin isoforms and laminin-producing cells in rat anterior pituitary. *Acta Histochem. Cytochem.* 2012, 45, 309–315.
22. Hallmann, R., Horn, N., Selg, M., Wendler, O., Pausch, F., Sorokin, L. M. Expression and function of laminins in the embryonic and mature vasculature. *Physiol. Rev.* 2005, 85, 979–1000.
23. Durbeej, M. Laminins. *Cell Tissue Res.* 2010, 339, 259–268.
24. Powell, S. K., and Kleinman, H. K. Neuronal laminins and their cellular receptors. *Int. J. Biochem. Cell Biol.* 1997, 29, 401–414.
25. Yamada, K. M. Adhesive recognition sequences. *J. Biol. Chem.* 1991, 266, 12809–12812.
26. Yamada, Y., Kleinman, H. K. Functional domains of cell adhesion molecules. *Curr. Opin. Cell Biol.* 1992, 4, 819–823.
27. Miyazaki, T., Futaki, S., Hasegawa, K., Kawasaki, M., Sanzen, H., Hayashi, M., Kawase, E., Sekiguchi, K., Nakatsuji, N., Suemori, H. Recombinant human laminin isoforms can support the undifferentiated growth of human embryonic stem cells. *Biochem. Biophys. Res. Commun.* 2008, 375, 27–32.

28. Rodin, S. Antonsson, L., Niaudet, C., Simonson, O. E., Salmela, E., Hansson, E. M., Domogatskaya, A., Xiao, Z., Damdimopoulou, P., Sheikhi, M., Inzunza, J., Nilsson, A. S., Baker, D., Kuiper, R., Sun, Y., Blennow, E., Nordenskjöld, M., Grinnemo K. H., Kere, J., Betsholtz, C., Hovatta, O., Tryggvason, K. Clonal culturing of human embryonic stem cells on laminin-521/E-cadherin matrix in defined and xeno-free environment. *Nat. Commun.* 2014, 5, 3195.
29. Utani, A., Nomizu, M., Matsuura, H., Kato, K., Kobayashi, T., Takeda, U., Aota, S., Nielsen, P. K., Shinkai, H. A unique sequence of the laminin alpha 3 G domain binds to heparin and promotes cell adhesion through syndecan-2 and -4. *H. J. Biol. Chem.* 2001, 276, 28779–28788.
30. Kato, K., Utani, A., Suzuki, N., Mochizuki, M., Yamada, M., Nishi, N., Matsuura, H., Shinkai, H., Nomizu, M. Identification of neurite outgrowth promoting sites on the laminin alpha 3 chain G domain. *Biochemistry.* 2002, 41, 10747–10753.
31. Yamaguchi, H., Yamashita, H., Mori, H., Okazaki, I., Nomizu, M., Beck, K., Kitagawa, Y. High and low affinity heparin-binding sites in the G domain of the mouse laminin alpha 4 chain. *J. Biol. Chem.* 2000, 275, 29458–29465.
32. Makino, M., Okazaki, I., Kasai, S., Nishi, N., Bougaeva, M., Weeks, B. S., Otaka, A., Nielsen, P. K., Yamada, Y., Nomizu, M. Identification of cell binding sites in the laminin alpha5-chain G domain. *Exp. Cell Res.* 2002, 277, 95–106.
33. Nielsen, P. K., Gho, Y. S., Hoffman, M. P., Watanabe, H., Makino, M., Nomizu, M., Yamada, Y. Identification of a Major Heparin and Cell Binding Site in the LG4 Module of the Laminin  $\alpha$ 5 Chain. *J. Biol. Chem.*, 2000, 275, 14517–14523.
34. Okazaki, I., Suzuki, N., Nishi, N., Utani, A., Matsuura, H., Shinkai, H., Yamashita, H., Kitagawa, Y., Nomizu, M. Identification of Biologically Active Sequences in the Laminin  $\alpha$ 4 Chain G Domain. *J. Biol. Chem.* 2002, 277, 37070–37078.
35. Suzuki, N., Ichikawa, N., Kasai, S., Yamada, M., Nishi, N., Morioka, H., Yamashita, H., Kitagawa, Y., Utani, A., Hoffman, M. P., Nomizu, M. Syndecan Binding Sites in the Laminin  $\alpha$ 1 Chain G Domain. *Biochemistry.* 2003, 42, 12625–12633.
36. Hohenester, E., Tisi, D., Talts, J. F., Timpl, R. The crystal structure of a laminin G-like module reveals the molecular basis of alpha-dystroglycan binding to laminins, perlecan, and agrin. *Mol. Cell.* 1999, 4, 783–792.
37. Suzuki, N., Nakatsuka, H., Mochizuki, M., Nishi, N., Kadoya, Y., Utani, A., Oishi, S., Fujii, N., Kleinman K. H., Nomizu, M. Biological activities of homologous loop regions in the laminin alpha chain. *J. Biol Chem.* 2003, 278, 45697–45705.

38. Harrison, D., Hussain, S. A., Combs, A. C., Ervasti, J. M., Yurchenco, P. D., Hohenester, E. Crystal structure and cell surface anchorage sites of laminin alpha1LG4-5. *J. Biol. Chem.* 2007, 282, 11573-11581.
39. Ali, M., Amon, M., Bender, V., Bolte, A., Separovic, F., Benson, H., Manolios, N. Cyclization enhances function of linear anti-arthritic peptides. *Clin. Immunol.* 2014, 150, 121–133.
40. Helbling-Leclerc, A., Zhang, X., Topaloglu, H., Cruaud, C., Tesson, F., Weissenbach, J., Tomé, F. M., Schwartz, K., Fardeau, M., Tryggvason, K., Guicheney, P. Mutations in the laminin alpha 2-chain gene (LAMA2) cause merosin-deficient congenital muscular dystrophy. *Nat. Genet.* 1995, 11, 216–218.
41. Kuang, W., Xu, H., Vachon, P. H., Liu, L., Loechel, F., Wewer, U. M., Engvall, E. Merosin-deficient congenital muscular dystrophy. Partial genetic correction in two mouse models. *J. Clin. Invest.* 1988, 102, 844–852.
42. Miyagoe, Y., Hanaoka, K., Nonaka, I., Hayashima, Y., Arahata, K., Nabeshima, Y., Takeda, S. Laminin alpha2 chain-null mutant mice by targeted disruption of the Lama2 gene: a new model of merosin (laminin 2)-deficient congenital muscular dystrophy. *FEBS Lett.* 1997, 415, 33–39.
43. Wizemann, H., Garbe, J. H. O., Friedrich, M. V. K., Timpl, R., Sasaki, T., Hohenester, E. Distinct requirements for heparin and alpha-dystroglycan binding revealed by structure-based mutagenesis of the laminin alpha2 LG4-LG5 domain pair. *J. Mol. Biol.* 2003, 332, 635–642
44. Talts, J. F., Timpl, R. Mutation of a basic sequence in the laminin alpha2 LG3 module leads to a lack of proteolytic processing and has different effects on beta1 integrin-mediated cell adhesion and alpha-dystroglycan binding. *FEBS Lett.* 1999, 458, 319–323.
45. Cohn R. D., Mayer, U., Saher, G., Hermann, R., van der Flier, A., Sonnenberg, A., Sorokin, L., Voit, T. Secondary reduction of alpha7B integrin in laminin alpha2 deficient congenital muscular dystrophy supports an additional transmembrane link in skeletal muscle. *J. Neurol. Sci.* 1999, 163, 140–152.
46. Suzuki, N., Hozumi, K., Urushibata, S., Yoshimura, T., Kikkawa, Y., Gumerson, J. D., Michele, D. E., Hoffman, M. P., Yamada, Y., and Nomizu M. Identification of  $\alpha$ -dystroglycan binding sequences in the laminin  $\alpha$ 2 chain LG4-5 module. *Matrix Biol.* 2010, 29, 143–151.
47. Katagiri F, Hara T, Yamada Y, Hozumi K, Urushibata S, Kikkawa Y, Nomizu, M. The biological activities of the homologous loop regions in the laminin  $\alpha$  chain LG modules. *Biochemistry.* 2014, 53, 3699–3708.

48. Hozumi K, Sasaki A, Yamada Y, Otagiri D, Kobayashi K, Fujimori C, Katagiri F., Kikkawa, Y., Nomizu, M. Reconstitution of laminin-111 biological activity using multiple peptide coupled to chitosan scaffolds. *Biomaterials*. 2012, 33, 4241–4250.
49. Hayouka, Z., Hurevich, M., Levin, A., Benyamini, H., Iosub, A., Maes, M., Shalev, D. E., Loyter, A., Gilon, C., Friedler, A. Cyclic peptide inhibitors of HIV-1 integrase derived from the LEDGF/p75 protein. *Bioorg. Med. Chem.* 2010, 18, 8388–8395.
50. Hayouka, Z., Levin, A., Hurevich, M., Shalev, D. E., Loyter, A., Gilon, C., Friedler, A. A comparative study of backbone versus side chain peptide cyclization: application for HIV-1 integrase inhibitors. *Bioorg. Med. Chem.* 2012, 20, 3317–3322.
51. 上田顯 著 「コンピュータシミュレーション –マクロな系の中の原子運動–」 朝倉書店, 1990 年
52. 上田顯 著 「分子シミュレーション –古典系から量子系手法まで–」 裳華房, 2011 年
53. 岡崎進・吉井範行 著 「コンピュータ・シミュレーションの基礎 (第2版)」 化学同人, 2011 年
54. 神谷成敏・肥後順一・福西快文・中村春木 著 「タンパク質計算科学 (基礎と創薬への応用)」 共立出版, 2009 年
55. 土井正男 著 「統計力学」 朝倉書店, 2010 年
56. 宮下精二 著 「熱・統計力学」 培風館, 2007 年
57. 久保亮五 著 「大学演習 熱学・統計力学 修訂版」 裳華房, 2010 年
58. Shaw, E. D., Dror, O. R., Salmon, K. J., Grossman, P. J., Mackenzie, M. K., Bank, A. J., Young, C., Deneroff, M. M., Batson, B., Bowers, J. K., Chow, E., Eastwood, P. M., Ierardi, J. D., Klepeis, L. J., Kuskin, S. J., Larson, H. R., Lindorff-Larsen, K., Maragakis, P., Moraes, A. M., Piana, S., Shan, Y., Towles, B. Millisecond-scale molecular dynamics simulations on Anton. *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*. Association for Computing Machinery, 2009, DOI: 10.1145/1654059.1654126.
59. Wang, Z., Li, C., Ellenburg, M., Soistman, E., Ruble, J., Wright, B., Ho, J. X., Carter, D. C. Structure of human ferritin L chain. *Acta Crystallogr. D. Biol. Crystallogr.* 2006, 62, 800–806.
60. Bussi, G., Donadio, D., Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 2007, 126, 014101.
61. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* 1984, 81, 511–.
62. Hoover, G. W. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*. 1985, 31, 1695–1697.

63. Parrinello, M., Rahman, A. Crystal Structure and Pair Potentials: A Molecular- Dynamics Study. *Phys. Rev. Lett.* 1980, 45, 1196–1199.
64. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science.* 1973, 181, 223-230.
65. Altis, A., Otten, M., Nguyen, P. H., Hegger, R., Stock, G. Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis. *J. Chem. Phys.* 2008, 128, 245102–245111.
66. Umezawa, K., Ikebe, J., Nomizu, M., Nakamura, H., Higo, J. Conformational requirement on peptides to exert laminin's activities and search for protein segments with laminin's activities. *Biopolymers.* 2009, 92, 124–131.
67. Torrie, M. G., Valleau, P. J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* 1977, 23, 187–199.
68. Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* 1935, 3, 300.
69. Kirkwood, J. G. Statistical mechanics of liquid solutions. *Chem. Rev.* 1936, 19, 275–307.
70. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 2013, 29, 845–854.
71. Tisi, D., Talts, J. F., Timpl, R., Hohenester, E. Structure of the C-terminal laminin G-like domain pair of the laminin alpha2 chain harbouring binding sites for alpha-dystroglycan and heparin. *EMBO J.* 2000, 19, 1432–1440.
72. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dorr, R. O., Shaw, D. E. Improved side-chain torsion potentials for the AMBER ff99SB protein force field. *PROTEINS. Struct Funct Gen.* 2010, 78, 1950–1958.
73. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983, 79, 926–935.
74. Emsley, J., King, S. L., Bergelson, J. M., Liddington, R. C. Crystal structure of the I domain from integrin alpha2beta1. *J. Biol. Chem.* 1997, 272, 28512–28517.
75. Zhang, W., Wu, C., Duan, Y. Convergence of replica exchange molecular dynamics. *J. Chem. Phys.* 2005, 123, 154105–154113.
76. Chodera, J. D., Swope, W. C., Pitera, J. W., Seok, C., Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering. *J. Chem. Theory Comput.* 2007, 3, 26–41.

77. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C., Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* 1995, 16, 273–284.
78. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 1983, 22, 2577–2637.
79. Chou, P. Y., Fasman, G. D. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* 1983, 47, 45–148.
80. Yamada, H., Komatsu, Y., Miyakawa, T., Morikawa, R., Katagiri, F., Hozumi, K., Kikkawa, Y., Nomizu, M., Takasu, M. Conformation analysis of loop region peptides in the laminin  $\alpha$  chain LG4 modules by molecular dynamics simulations. *Peptide Science 2011, Jap. Peptide Soc.* 2012, 201–204.
81. Yamada, H., Miyakawa, T., Morikawa, R., Katagiri, F., Hozumi, K., Kikkawa, Y., Nomizu, M., Takasu, M. Molecular dynamics simulations of peptides derived from laminin  $\alpha$ 2 Chain. *Peptide Science 2012, Jap. Peptide Soc.* 2013, 339–342.
82. Yamada H., Mori S., Miyakawa T., Morikawa R., Katagiri, F., Hozumi, K., Kikkawa, Y., Nomizu, M., Takasu, M. Structural study of cell attachment peptide derived from laminin by molecular dynamics simulation. *PLoS One.* 2016, 11, e0149474.
83. Yamada, H., Fukuda, M., Miyakawa, T., Morikawa, R., Takasu, M. Conformation analysis of peptides derived from laminin alpha 1-2 chain using molecular dynamics simulation. *JPS Conf. Proc.*, 2014, 1, 016016.
84. Hatfield, M. P. D., Murphy, R. F., Lovas, S. Molecular dynamics analysis of the conformations of a  $\beta$ -hairpin miniprotein. *J. Phys. Chem. B.* 2010, 114, 3028–3037.
85. Csontos, J., Murphy, R. F., Lovas, S. The role of weakly polar and H-bonding interactions in the stabilization of the conformers of FGG, WGG, and YGG: an aqueous phase computational study. *Biopolymers.* 2008, 89, 1002–1011.
86. Pande, V. S., Roshkar, D. S. Molecular dynamics simulations of unfolding and refolding of a  $\beta$ -hairpin fragment of protein G. *Proc. Natl. Acad. Sci. USA.* 1999, 96, 9062–9067
87. Garcia, A. E., Sanbonmatsu, K. Y. Exploring the energy landscape of a  $\beta$ -hairpin in explicit solvent. *Proteins.* 2001, 42, 345–354.
88. Zhou, R., Berne, B. J., Germain, R. The free energy landscape for  $\beta$ -hairpin folding in explicit water. *Proc. Natl. Acad. Sci. USA.* 2001, 98, 14931–14936.
89. Zhou, R., Exploring the protein folding free energy landscape: coupling replica exchange method with P3ME/RESPA algorithm. *J. Mol. Graph. Model.* 2004, 22, 451–463.
90. Singh, G., Brovchenko, I., Oleinikova, A., Winter, R. Aggregation of fragments of the islet amyloid polypeptide as a phase transition: a cluster analysis. *NIC series.* 2007, 36, 275–278.