

平成24年6月4日
分子モデリングと分子シミュレーション

分子動力学法の応用

東京大学大学院農学生命科学研究科
アグリバイオインフォマティクス
教育研究プログラム
寺田 透

本日の講義内容

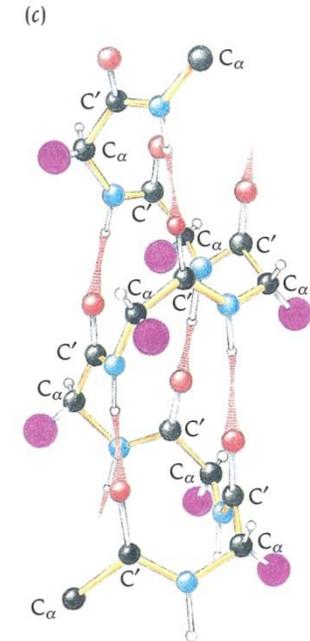
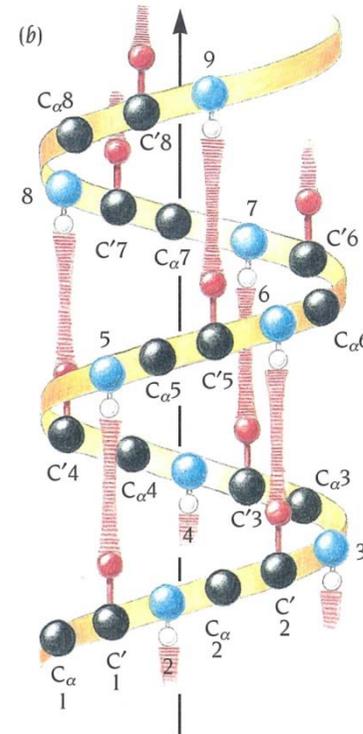
- ペプチドのエネルギー最小化
- 水溶液環境のモデル化
- ペプチドの分子動力学シミュレーション
 - 課題1
- タンパク質の分子動力学シミュレーション
 - 課題2
- シミュレーションの高速化
- シミュレーションの手順

ペプチドの生成

1. Discovery Studio 3.0 Clientを起動し、新しいMolecule Windowを開く
2. 「Macromolecules」ボタンを左クリック
3. 「Build and Edit Protein」を展開
4. Build Actionを「Create/Grow Chain」とする
5. Conformationを「Right-hand Alpha Helix」とする
6. Choose Amino Acidで「Ala」を9回クリックしてalanine 9-merペプチドを作成せよ

水素結合距離の測定

- 主鎖のアミド基の窒素原子とカルボニル基の酸素原子間で形成されている水素結合ペアについて以下に従って距離を測定しておく
 1. 原子間距離を測りたい原子のペアをShiftキーを押しながら左クリックで選択
 2. Measureボタン  を左クリックすると原子のペアが緑色の線で結ばれ、距離がÅ単位で表示される

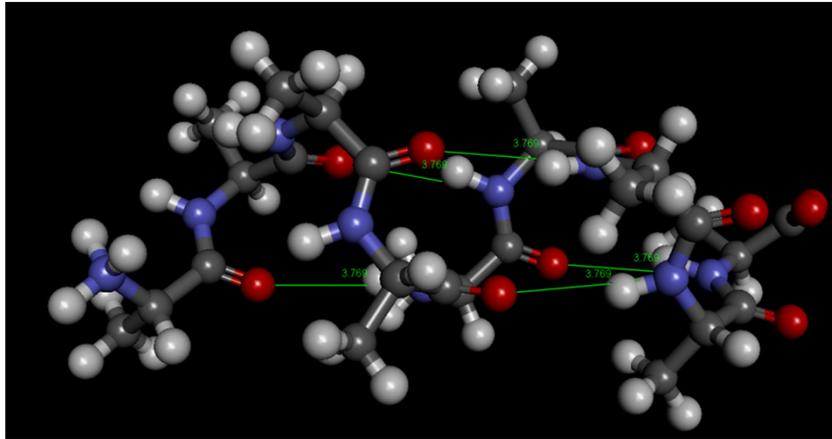


α helixでは*i*番目のカルボニル酸素と*i*+4番目のアミド窒素が水素結合を形成する

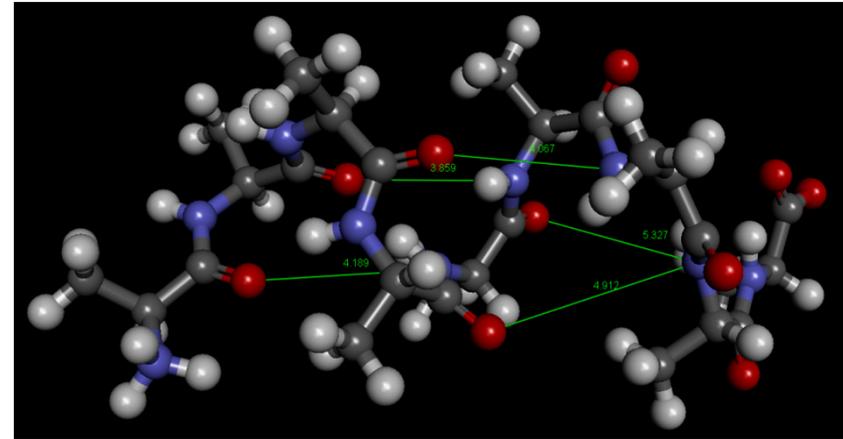
ペプチドのエネルギー最小化

1. デスクトップに「ala9.dsv」として保存
2. 「Simulation」ボタンを左クリック
3. 「Change ForceField」を展開し、Forcefieldに「charmm27」を指定
4. 「Apply Forcefield」を左クリック
5. 「Run Simulations」を展開し、Toolsにある「Minimization」を左クリックし「Run」
6. Jobが完了したら、Data Tableでエネルギーと水素結合距離を確認

エネルギー最小化の結果



初期構造のエネルギー
268.307 kcal mol⁻¹



最小化後のエネルギー
141.347 kcal mol⁻¹

エネルギー最小化後に水素結合が壊れていることに注意

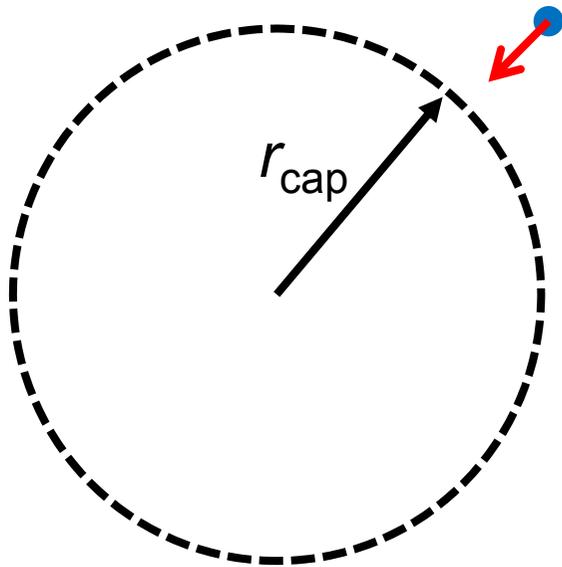
水溶液環境のモデル化(1)

- 今回のエネルギー最小化計算は真空中で行われており、水分子による溶媒効果は考慮されていない
- 生体分子のシミュレーションにおいては、水溶液環境を適切なモデルを用いて再現する必要がある

水溶液環境のモデル化(2)

- 現在以下の方法がよく用いられている
- 水分子を陽に配置
 - 球状に配置
 - 直方体状に配置→周期境界条件
- 溶媒和自由エネルギーを近似的に求める
 - 非極性項→溶媒接触表面積に比例
 - 極性項→連続誘電体モデル
 - Poisson-Boltzmann方程式
 - Generalized Bornモデル

球状の配置



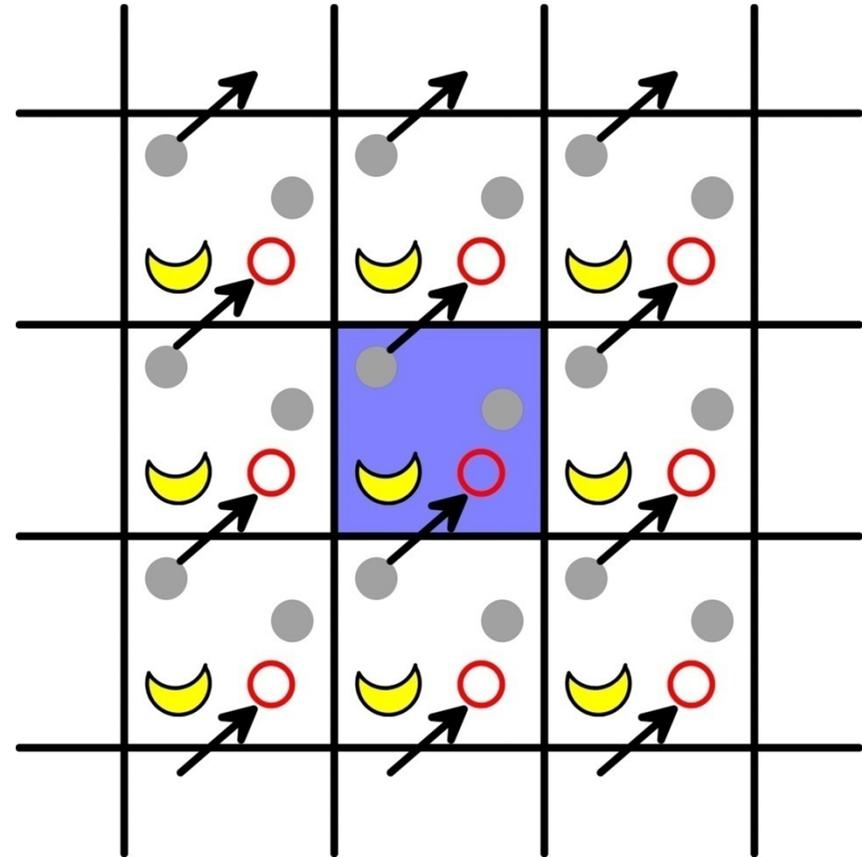
$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

$$E_{\text{cap}} = \begin{cases} 0 & r < r_{\text{cap}} \\ k(r - r_{\text{cap}})^2 & r \geq r_{\text{cap}} \end{cases}$$

- 水分子の“蒸発”を防ぐため、分子が半径 r_{cap} の球の外側に出て行こうとすると、系の中心に向けて束縛力をかける
- 系の表面に位置する水分子は中心付近の水とは異なる環境に置かれる

周期境界条件

- 中央のセルと同じものが無限に繰り返す
- セルから出て行った分子は、そのセルの反対側から入る
- どの分子も同じ環境
- 系が隣接セルからの影響を感じないように、系のサイズを十分に大きくする必要がある



圧力の計算

$$F = E - TS, \quad dE = -PdV + TdS$$

$$dF = dE - TdS - SdT = -PdV - SdT$$

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = kT \left(\frac{\partial \ln Z}{\partial V}\right)_T = \frac{kT}{Z} \left(\frac{\partial Z}{\partial V}\right)_T$$

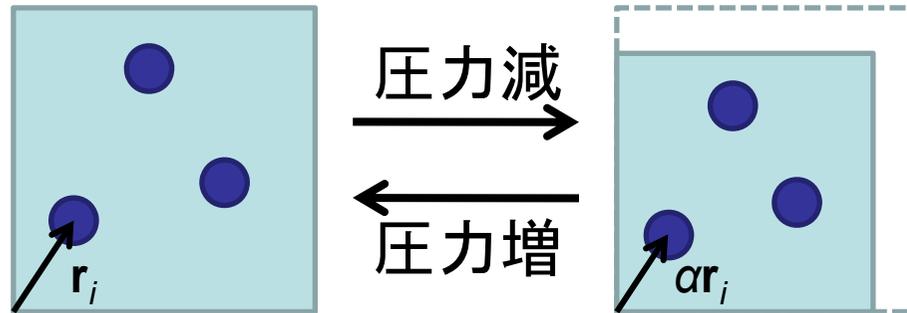
$$= \frac{NkT}{V} + \frac{1}{3V} \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i \quad \text{ビリアル定理}$$

$$= \frac{NkT}{V} + \frac{1}{3V} \sum_{i=1}^N \sum_{j=i+1}^N \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \quad \text{周期境界条件ではこちらを使う}$$

↓
相互作用のない系(理想気体)では、 $PV = NkT = nRT$

圧力の制御

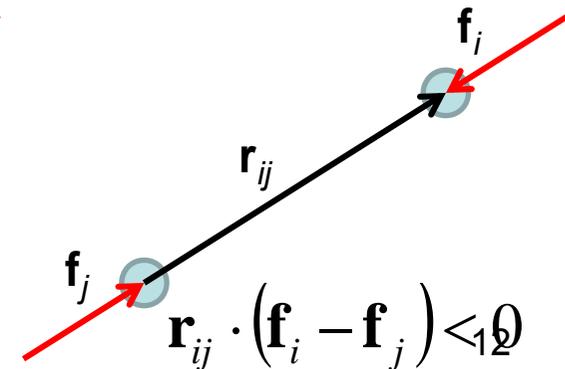
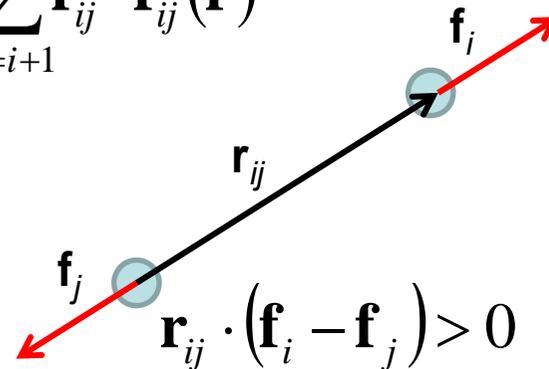
- 周期境界条件における、セルの大きさを変化させることで圧力を制御する



分子の重心位置も同様にスケールされる
分子内の原子の相対位置は変化しない

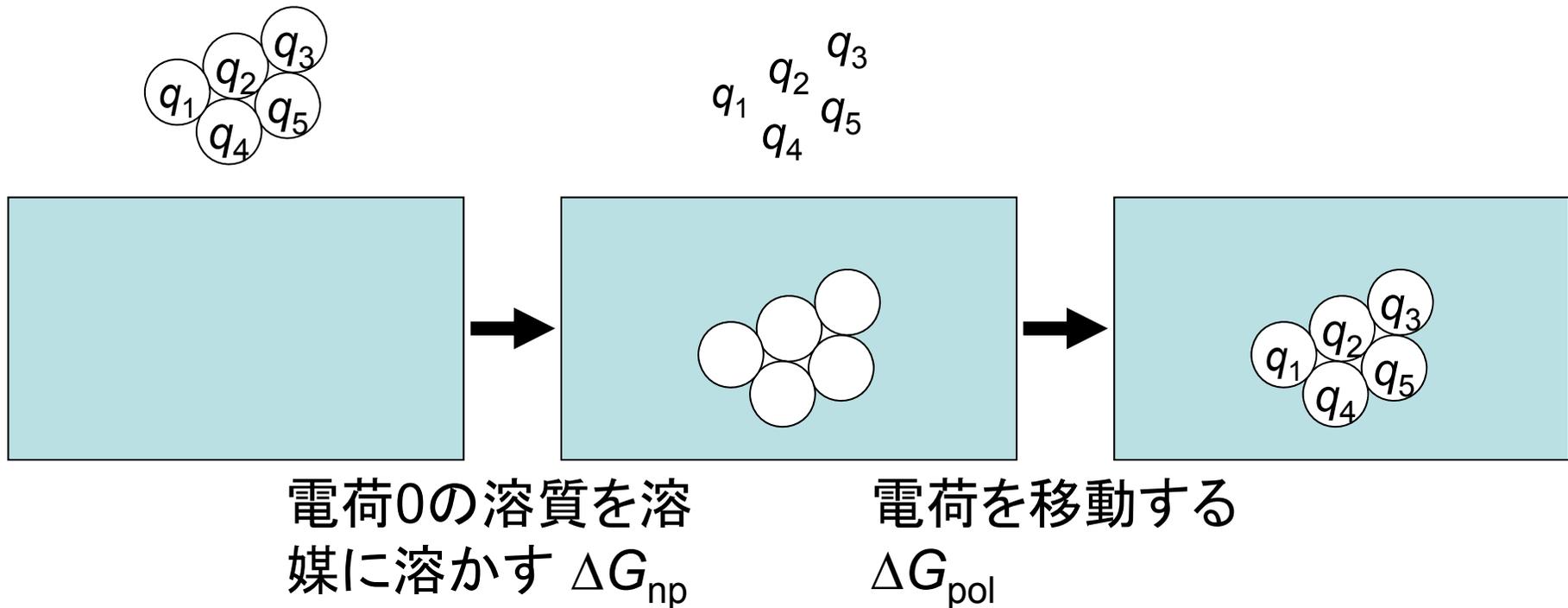
- 瞬間的にが $P < 0$ となることがある

$$P = \frac{NkT}{V} + \frac{1}{3V} \sum_{i=1}^N \sum_{j=i+1}^N \mathbf{r}_{ij} \cdot \mathbf{f}_{ij}(\mathbf{r})$$



溶媒和自由エネルギーの近似(1)

- 以下のような熱力学過程を考える



- 溶媒和自由エネルギー $\Delta G_{solv} = \Delta G_{np} + \Delta G_{pol}$

溶媒和自由エネルギーの近似(2)

- 非極性項 (ΔG_{np}) は、炭化水素の溶媒和自由エネルギーの実験データから、溶媒接触面積 (solvent-accessible surface area, SASA) に比例すると近似できる

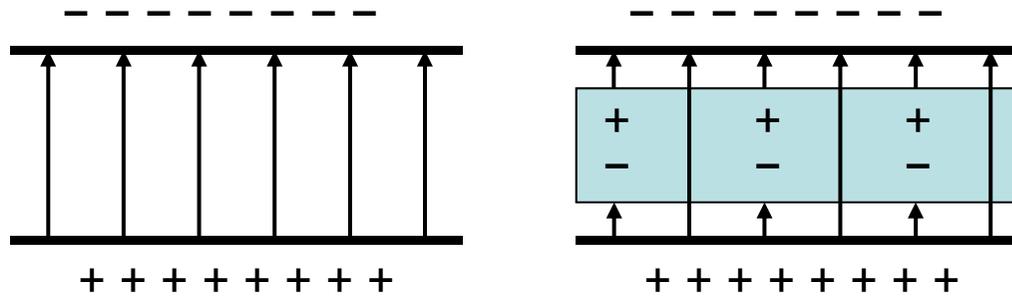
$$\Delta G_{np} = \sigma A + b \quad A: \text{SASA}, \sigma, b: \text{経験的パラメータ}$$

- 極性項は、溶媒を連続誘電体とみなして、電磁気学の理論を用いて求める

$$\Delta G_{pol} = \frac{1}{2} \int [\phi(\mathbf{r}) - \phi_{vac}(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r}$$

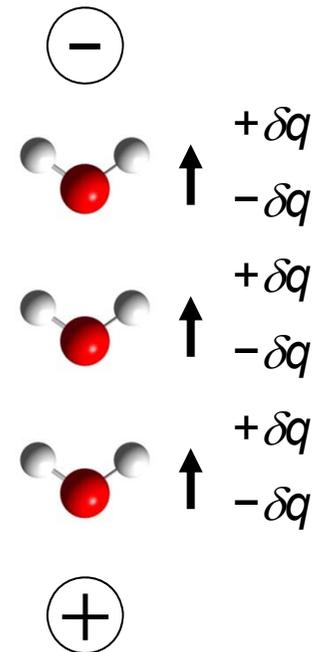
静電ポテンシャル 溶質の電荷分布

誘電体



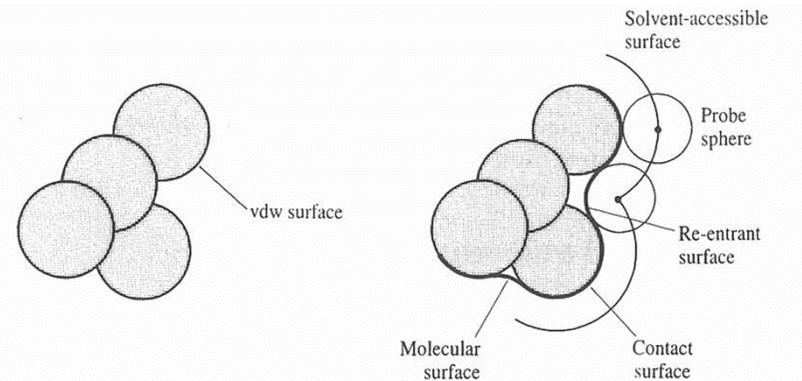
コンデンサーに比誘電率 ϵ の誘電体を挿入すると、誘電体の表面に電荷が現れ、極板間の電場を打ち消す
→静電ポテンシャルは $1/\epsilon$ となる

水溶液中では水分子が配向して誘電体として働き、静電相互作用を弱める



連続誘電体モデル

- 分子表面にプローブ球(水の場合半径1.4 Å)を転がした時、球の中心が作る軌跡→溶媒接触表面(solvent-accessible surface, SAS)
- SASからプローブ球の半径分内側の点がつくる表面→分子表面(molecular surface, MS)
- MSの内側を低誘電率($\epsilon = 1 \sim 4$)、外側を溶媒の誘電率(水の場合 $\epsilon = 80$)の誘電体とみなす



Poisson-Boltzmann方程式

- 連続誘電体モデルにおいて、静電ポテンシャルを与える

- 塩がない場合→Poisson方程式

$$\nabla \cdot [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] = -4\pi \rho(\mathbf{r})$$

静電ポテンシャル 溶質の電荷分布

- 塩が存在する場合→塩の電荷分布は Boltzmann分布に従う

$$\nabla \cdot [\varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r})] = -4\pi [\rho(\mathbf{r}) + \rho_{\text{ion}}(\mathbf{r})]$$

塩の電荷分布

Generalized Bornモデル

- Poisson-Boltzmann方程式の問題点
 - 力の計算ができない
 - 計算コストが高い
- Generalized Bornモデルの特徴
 - イオンの溶媒和自由エネルギーの式を拡張

$$\Delta G_{\text{pol}} = -\frac{q^2}{2a} \left(1 - \frac{1}{\varepsilon}\right) \quad \Rightarrow \quad \Delta G_{\text{pol}} = -\frac{1}{2} \left(1 - \frac{1}{\varepsilon}\right) \sum_{i,j=1}^N \frac{q_i q_j}{f_{\text{GB}}}$$

– 計算コストが低い

$$f_{\text{GB}} = \sqrt{r_{ij}^2 + a_i a_j} \exp\left(-r_{ij}^2 / 4a_i a_j\right)$$

– 力を解析的に求めることが可能

非極性項のモデル

- 横軸に溶媒接触表面積、縦軸にモル溶解度の対数をプロット*
- モル溶解度 s と自由エネルギー

$$\mu_w^\circ - \mu_o^\circ = -RT \ln s$$

μ_w° : 水溶液中での標準化学ポテンシャル

μ_o° : 炭化水素の標準化学ポテンシャル

- 現在では比例定数 σ に $5 \text{ cal mol}^{-1} \text{ \AA}^{-2}$ が使われる**
- この項に極性項を合わせて GB/SA (PB/SA) モデルと呼ばれる

*Hermann *J. Phys. Chem.* **76**, 2754 (1972).

Sitkoff *et al. J. Phys. Chem.* **98, 1978 (1994).

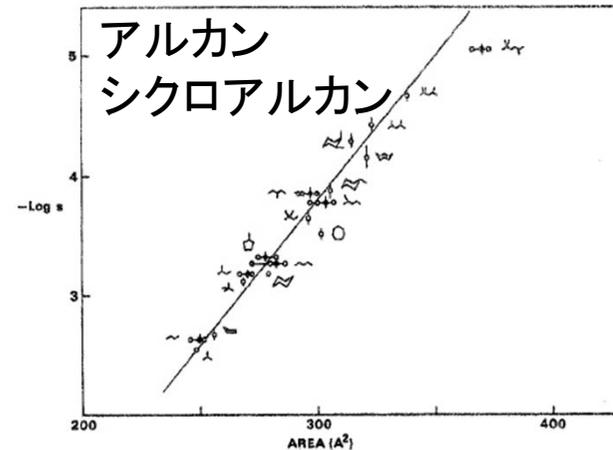


Figure 2. Relationship of cavity surface area to solubility for alkanes and cycloalkanes.

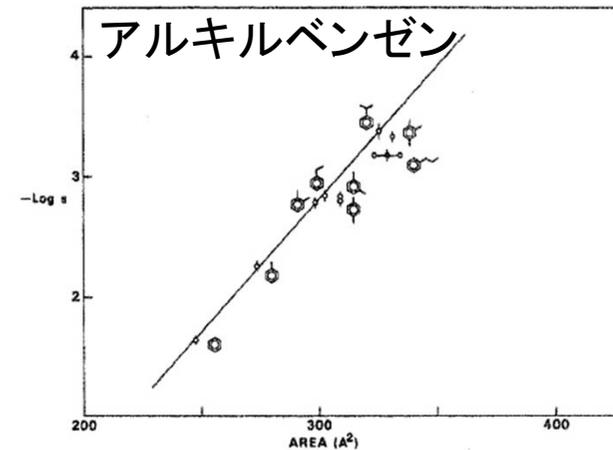


Figure 3. Relationship of cavity surface area to solubility for alkylbenzenes.

Discovery Studioにおける設定(1)

- 水分子を配置する場合
 1. 「Simulation」ボタンを左クリックし、「Run Simulations」を展開し、Advancedにある「Solvation」を左クリック
 2. Solvation Modelに「Explicit Periodic Boundary」もしくは「Explicit Spherical Boundary with Harmonic Restraint」を選び「Run」
- Generalized Bornモデルを使用する場合
 1. エネルギー最小化計算等で、Implicit Solvent Modelに「Generalized Born with a simple SWitching (GBSW)」を指定する

ペプチドのMDシミュレーション(1)

1. 保存しておいた「ala9.dsv」を開く
2. ペプチドの周りに球状に水分子を配置する
3. 新しいMolecule Windowが開くので、もう一度Force fieldをcharmm27に設定する
4. 「Run Simulations」のDynamicsにある「Standard Dynamics Cascade」を左クリックする
5. Electrostaticsを「Spherical Cutoff」に設定し「Run」
 - 2回のエネルギー最小化に続いて、加熱(Heating)、平衡化(Equilibration)、プロダクション(Production)のMDシミュレーションが順に実行される
 - エラーになる場合はForce fieldをクリアしてcharmm22に変更した後、もう一度クリアしてcharmm27に再設定すること

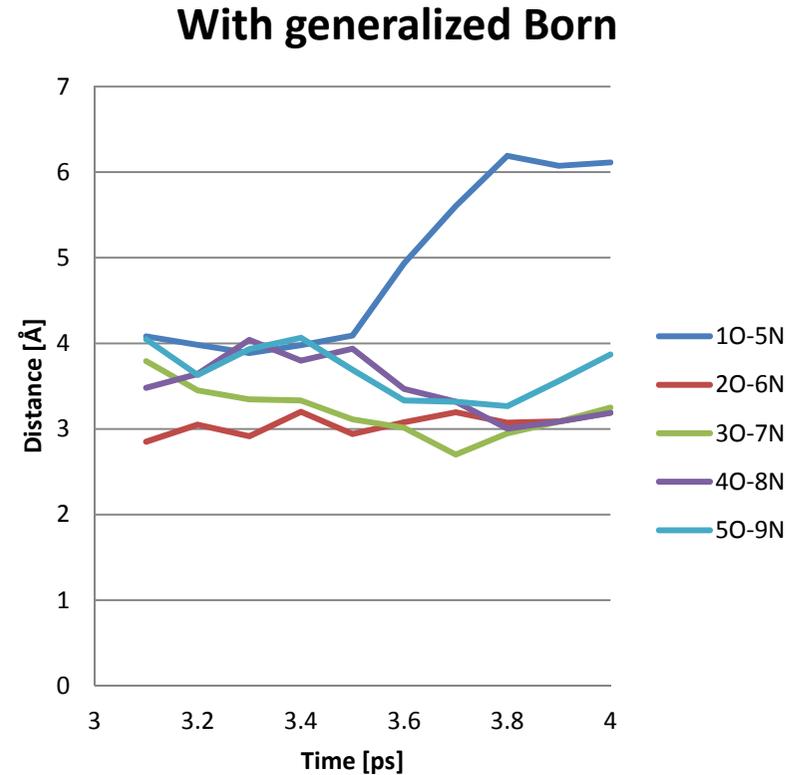
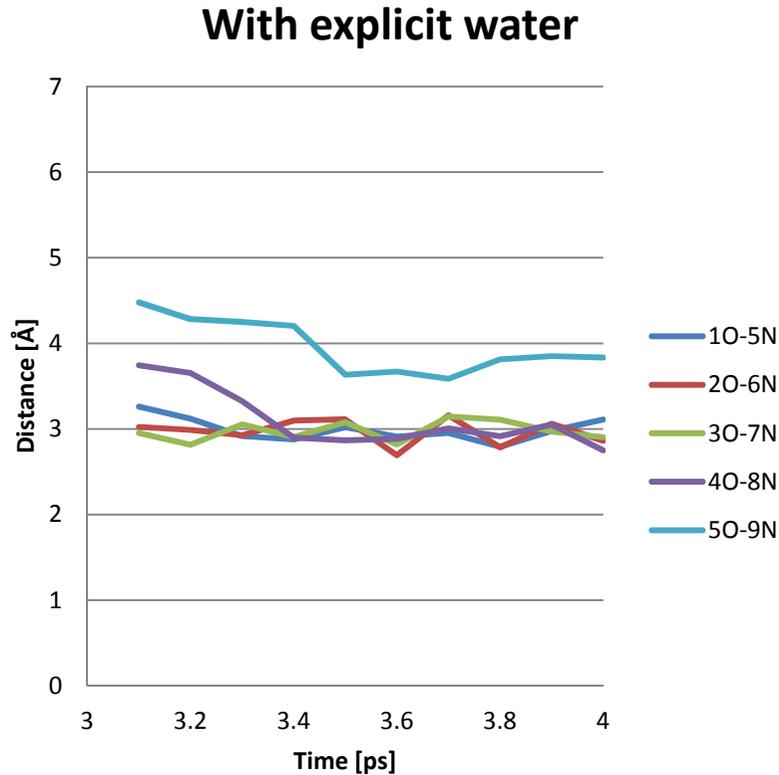
ペプチドのMDシミュレーション(2)

6. 計算が終わったら(計算には15分程度かかる)、新しいMolecule Windowに結果が表示される
7. メニューの「View」→「Toolbars」→「Animation」を選択→Animation Toolbarが表示される
8. 「Start Animation」ボタン  でAnimationを再生
9. 「Analyze Trajectory」を展開し、Analyzeにある「Analyze Trajectory」を左クリックし、「Run」
10. Data TableのConformationタブにプロダクションランにおける水素結合距離が表示される

課題1

- Alanine 9-merの水溶液中のMDシミュレーションについて、プロダクションランにおける水素結合長の変化をExcelでプロットせよ
 - TimeとDistanceのカラムをCtrlキーを押しながら選択してコピーし、Excelに貼り付け
 - どの系列がどの原子間の距離か明示すること
- Generalized Bornモデル(GBSW)を用いて同様の計算を行い、水素結合長の変化をプロットせよ
- これらのプロットから何が言えるか考察せよ

水素結合距離の時間変化



- 末端を除いて水素結合が概ね維持されている
- Generalized Bornでは距離の変動が大きい

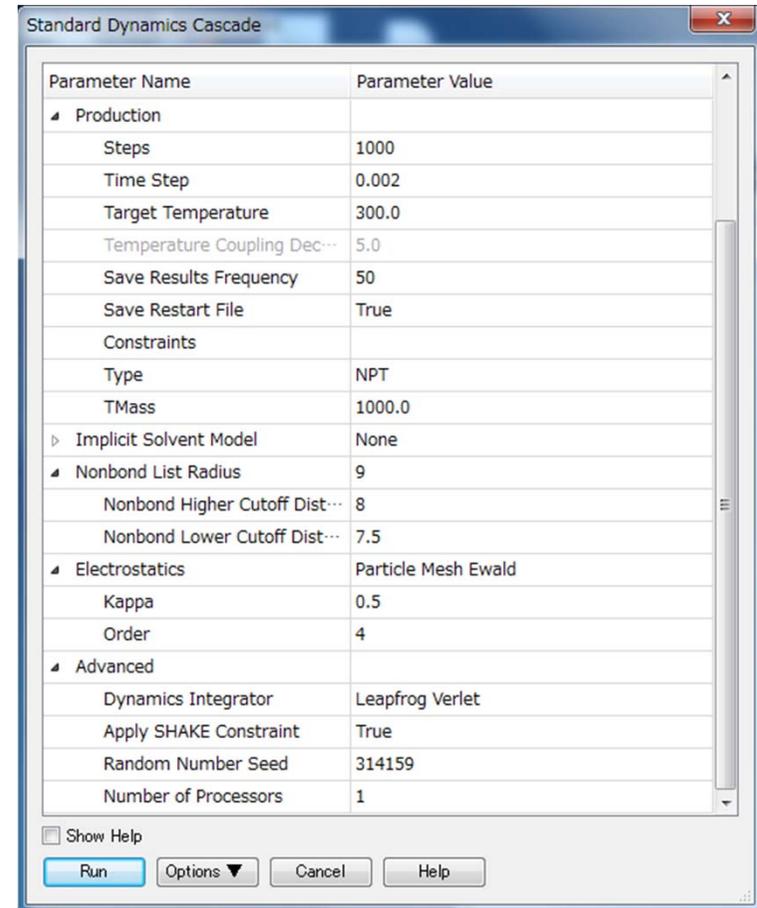
タンパク質のMDシミュレーション(1)

1. PDB ID 1CRNの構造を開く
2. Display StyleをLineに変更する
3. Force fieldにcharmm27を指定(水素が付加される)
4. Implicit Solvent ModelにGBSWを指定してエネルギー最小化を行う
5. タンパク質の周りに周期境界条件の下で、直方体状に水分子を配置する
6. もう一度Force fieldをcharmm27に設定

タンパク質のMDシミュレーション(2)

7. 「Standard Dynamics Cascade」で右の通り指定し「Run」

- 18分ほどかかる
- ステップ数: 1000
- 時間刻み: 2 fs
- 定温・定圧
- Particle Mesh Ewald
- SHAKE



課題2

- 講義のページから、プロダクションを10 ps分続けたトラジェクトリを含む1CRN_10ps.dsvをダウンロードして開く
- 1CRNの結晶構造を観察し、水素結合を2つ以上見つけよ
- このトラジェクトリから、この水素結合の距離を計算し、時間を横軸としてプロットせよ(どの残基のどの原子間か明示すること)
- このプロットから何が言えるか考察せよ

計算時間(1)

- 対象: 球状に配置した水分子 (TIP3Pモデル)
- Amber 11のSanderモジュール使用
- 計算にはIntel Xeon Processor 8コア使用
- 時間刻み Δt は0.5 fs
- 1 psの計算にかかる時間(単位は秒)を計測

原子数	T_{total} [s]	比率	T_{nb} [s]	$T_{\text{nb}}/T_{\text{total}}$
3087	35	1.0	35	0.983
6066	137	3.9	136	0.995
10608	420*	12.0	419	0.998

*1 nsあたり4.9日かかる

分子シミュレーションの効率化

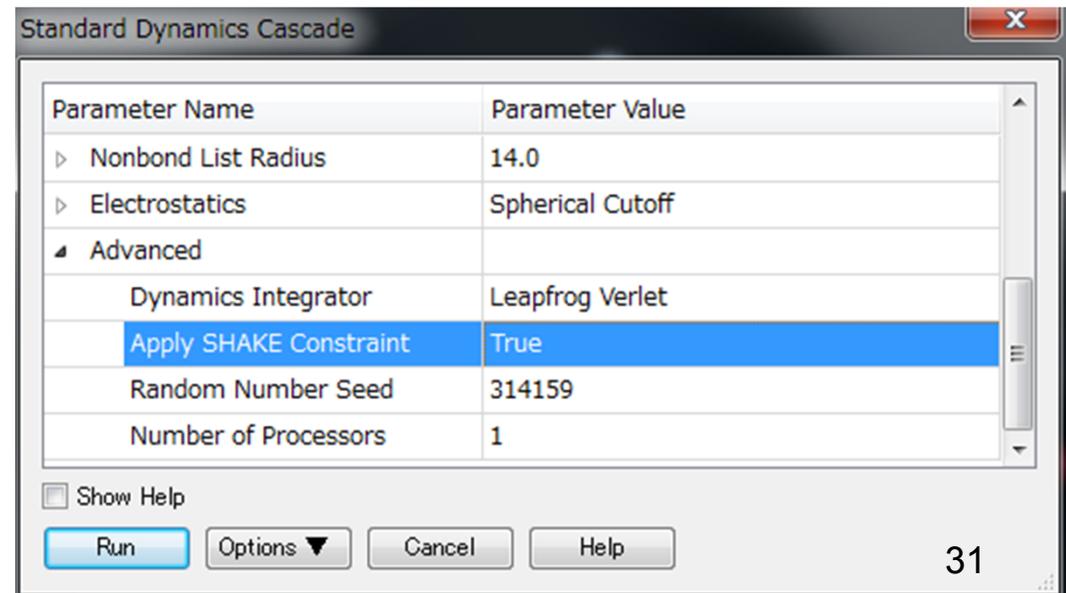
- 時間刻み Δt を長くする
 - SHAKE法
 - 多重時間積分法
- 非共有結合相互作用の計算の近似
 - カットオフ法
 - 多重極子展開法
 - Particle mesh Ewald法
- 本講義では赤枠の3つの方法について解説

SHAKE法

- 時間刻みは、最も速い運動の周期の10分の1から20分の1
- 最も速い運動は、X-H伸縮運動
→周期は約10 fs → $\Delta t = 0.5 \sim 1$ fs
- 次に速い運動は、X-X伸縮運動
→周期は約20 fs
- SHAKE法によりX-H結合長を固定
→長い時間刻み ($\Delta t = 2$ fs) の使用が可能

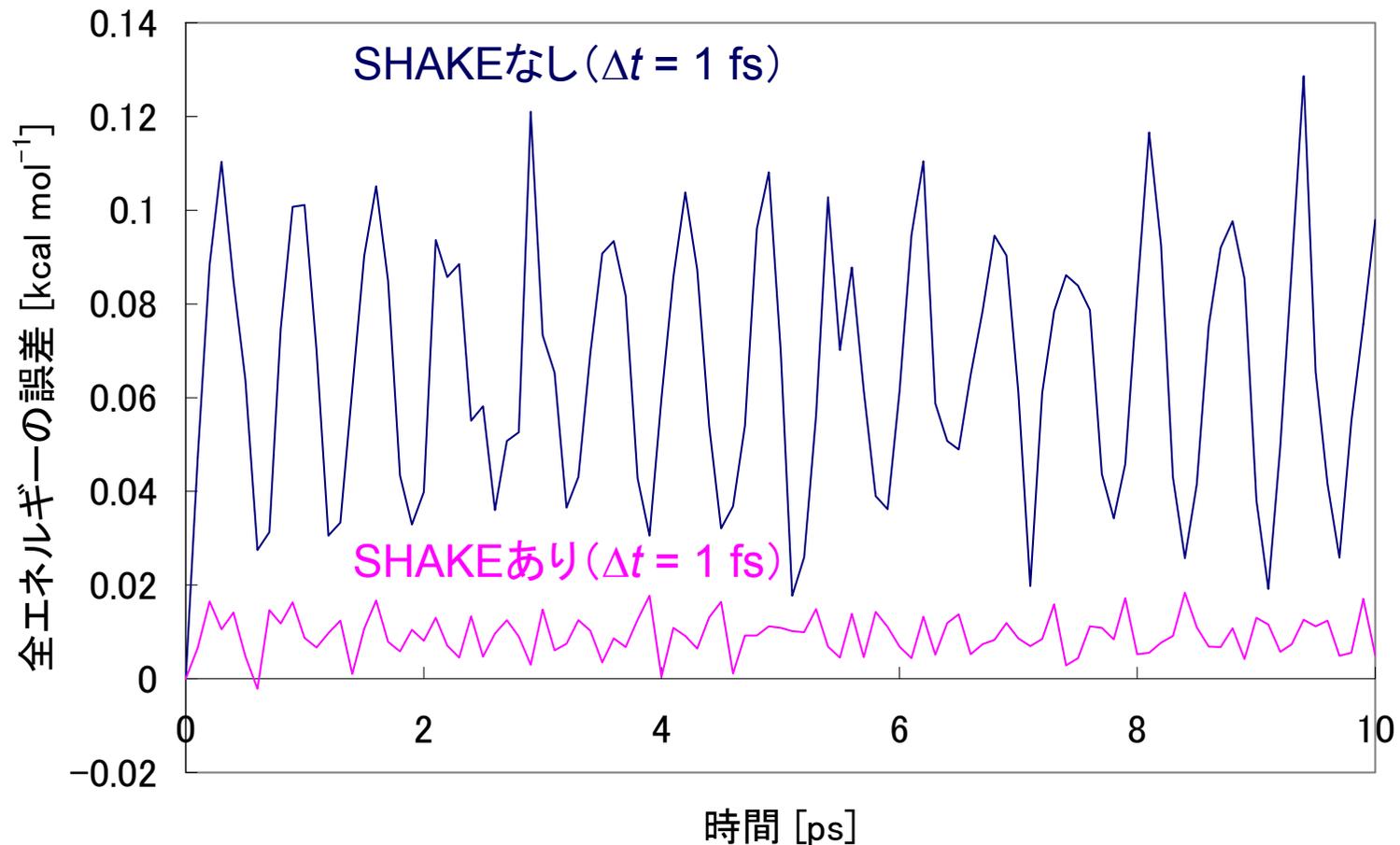
Discovery Studioにおける設定(2)

- 標準ではSHAKEは使用されない
- SHAKEを利用する場合は、Standard Dynamics Cascade、Dynamics (Production)などのパラメータのうち、Advancedの左の▶をクリックして展開し、右のように設定する

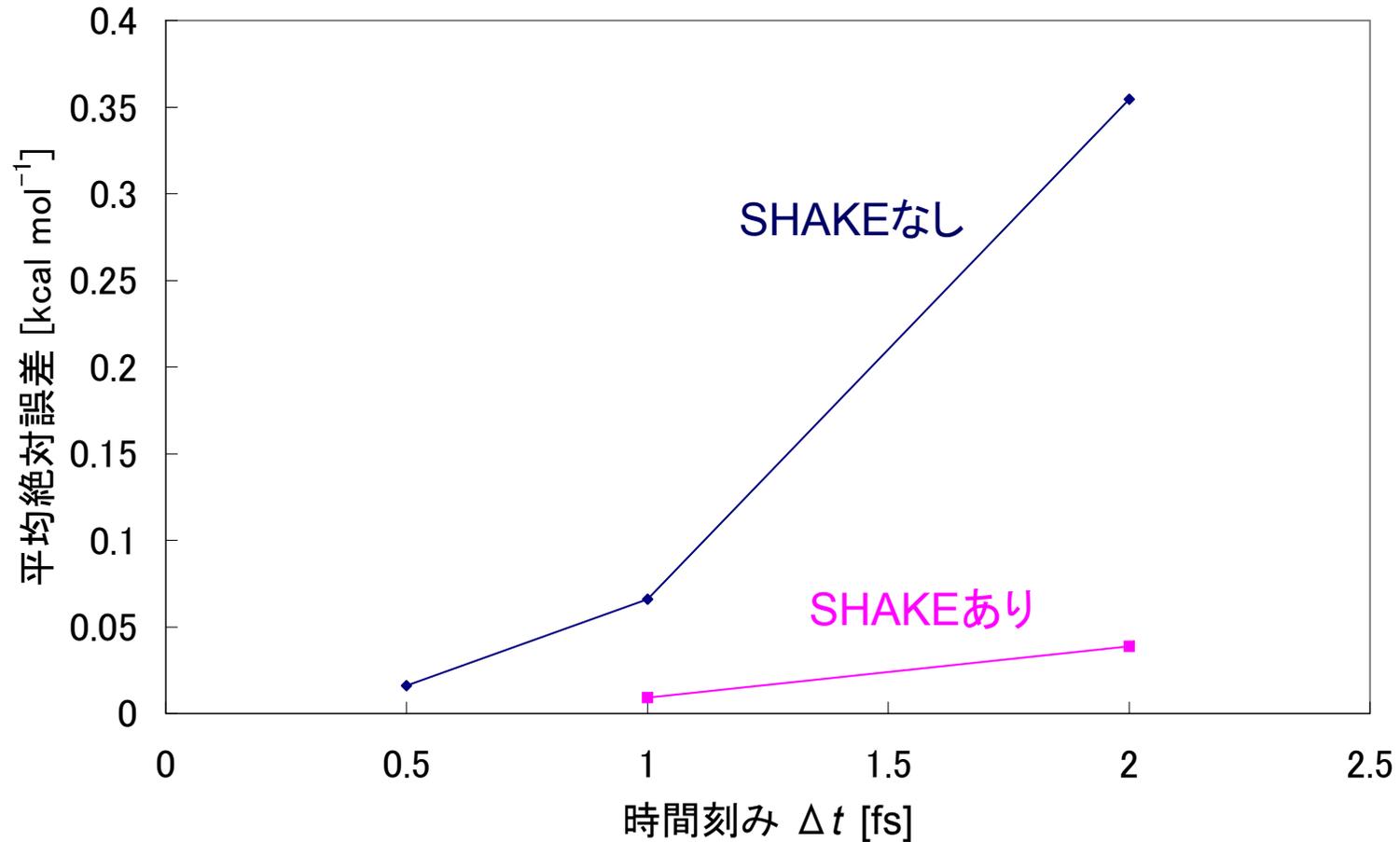


SHAKEの適用例

Methanolの分子動力学シミュレーション(温度制御なし)における全エネルギーの誤差(初期値との差)の推移



SHAKEの適用例

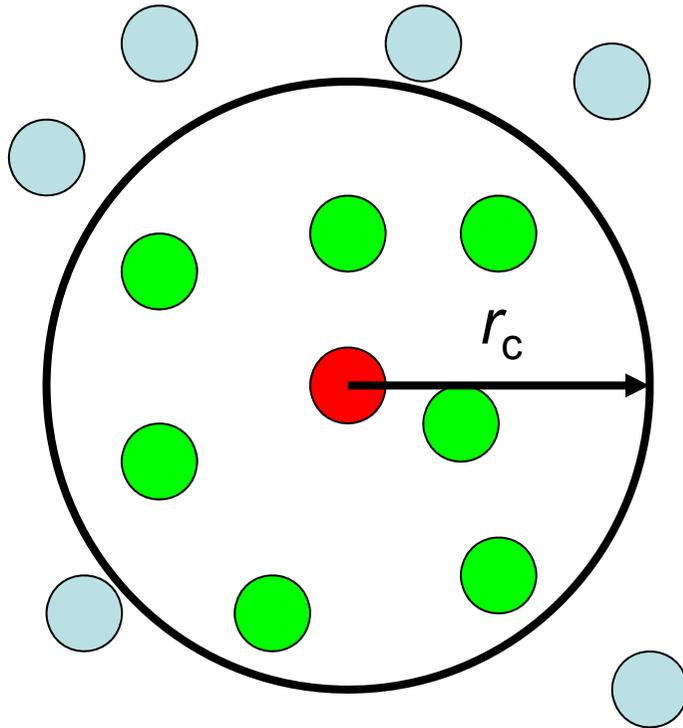


SHAKEを用いると時間刻み2 fsでもSHAKEなしの0.5 fsに匹敵する精度が得られる

非共有結合相互作用の扱い

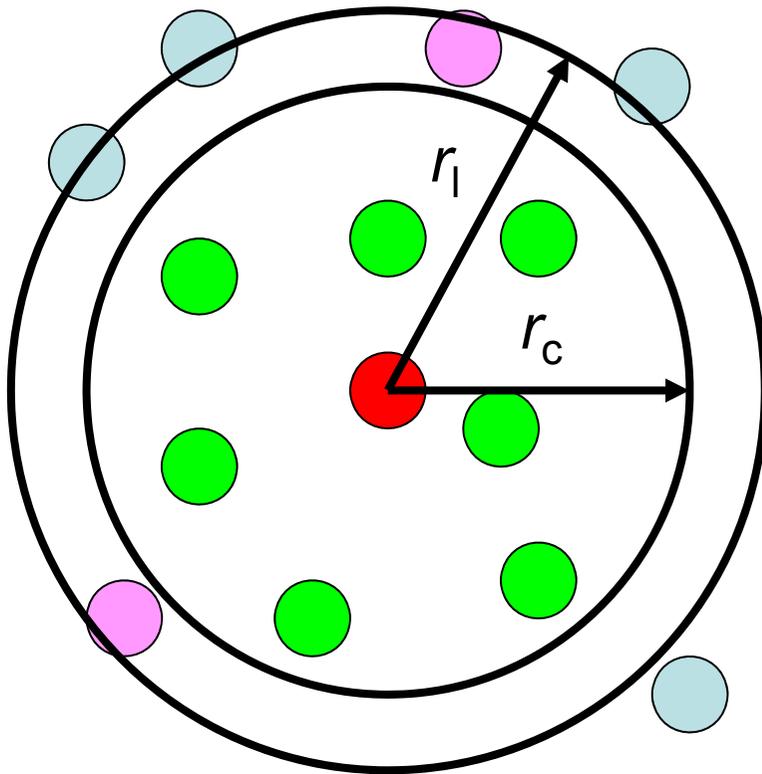
- 非共有結合相互作用は、原子のペアについて計算する必要がある
→ N 原子系では $N(N-1)/2$ のペア
- 非共有結合相互作用は距離が離れるほど弱くなる (van der Waals 引力は r^{-6} に比例、静電相互作用は r^{-1} に比例)
- 離れている原子同士は相互作用しないとみなす → カットオフ法

カットオフ法



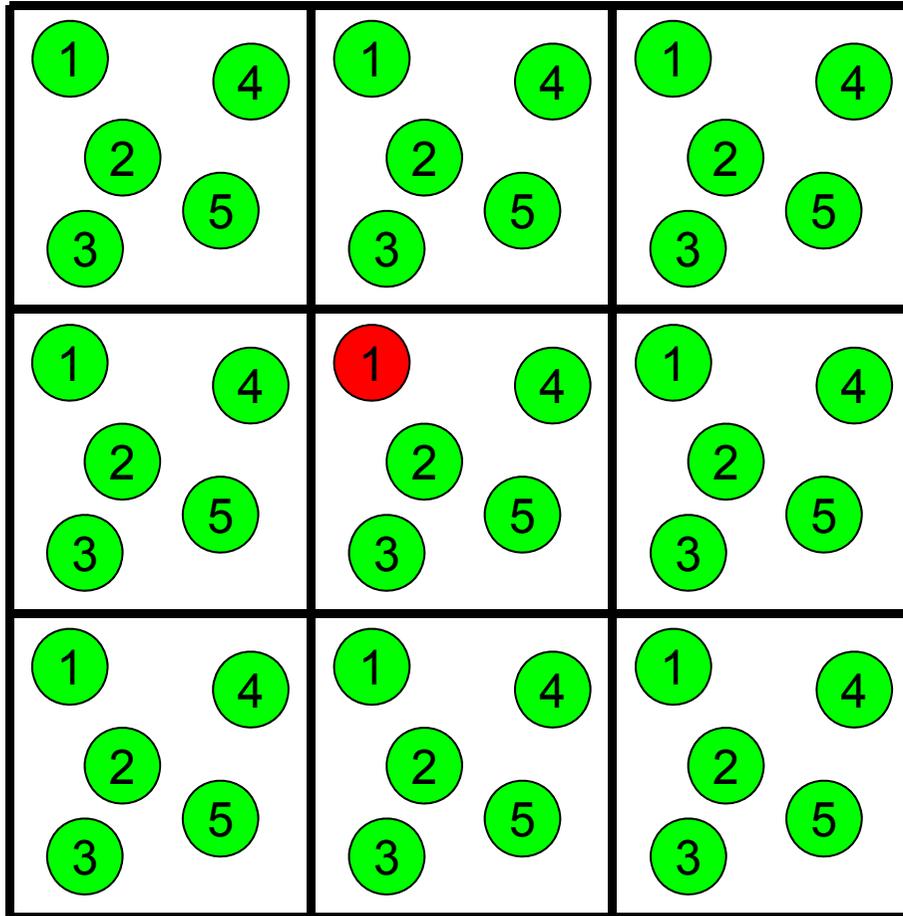
- 原子から半径 r_c の範囲内にある原子との非共有結合相互作用の計算を行う
- この範囲にある原子の平均個数を M とすると、非共有結合相互作用の計算量は $N(N-1)/2$ から NM に減少する

ペアリストの作成



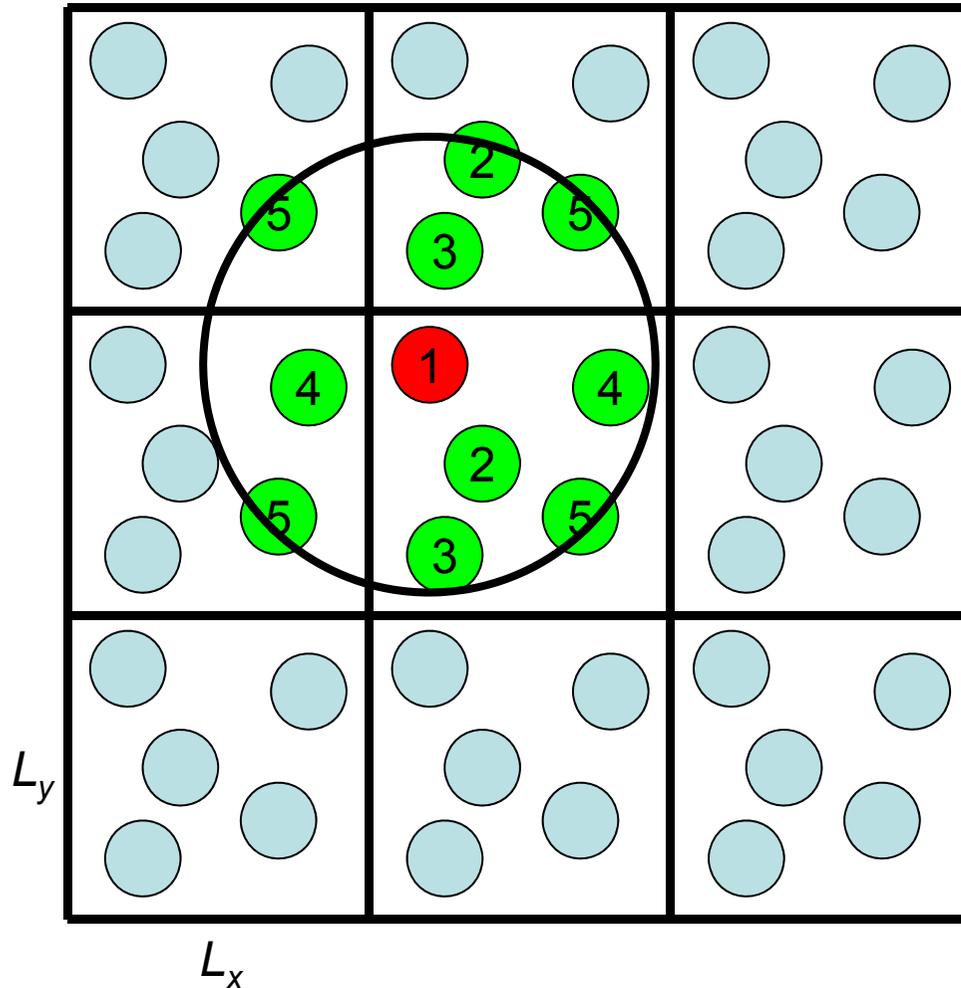
- カットオフ半径 r_c 以内にある原子ペアのリストを作成する必要がある
- この計算量は $N(N-1)/2$
- カットオフ半径 r_c より外側の半径 r_l の範囲でリストを作っておき、原子の最大移動度が $r_l - r_c$ を超えた時にリストを更新するようにすると計算量を削減できる

周期境界条件の場合(1)



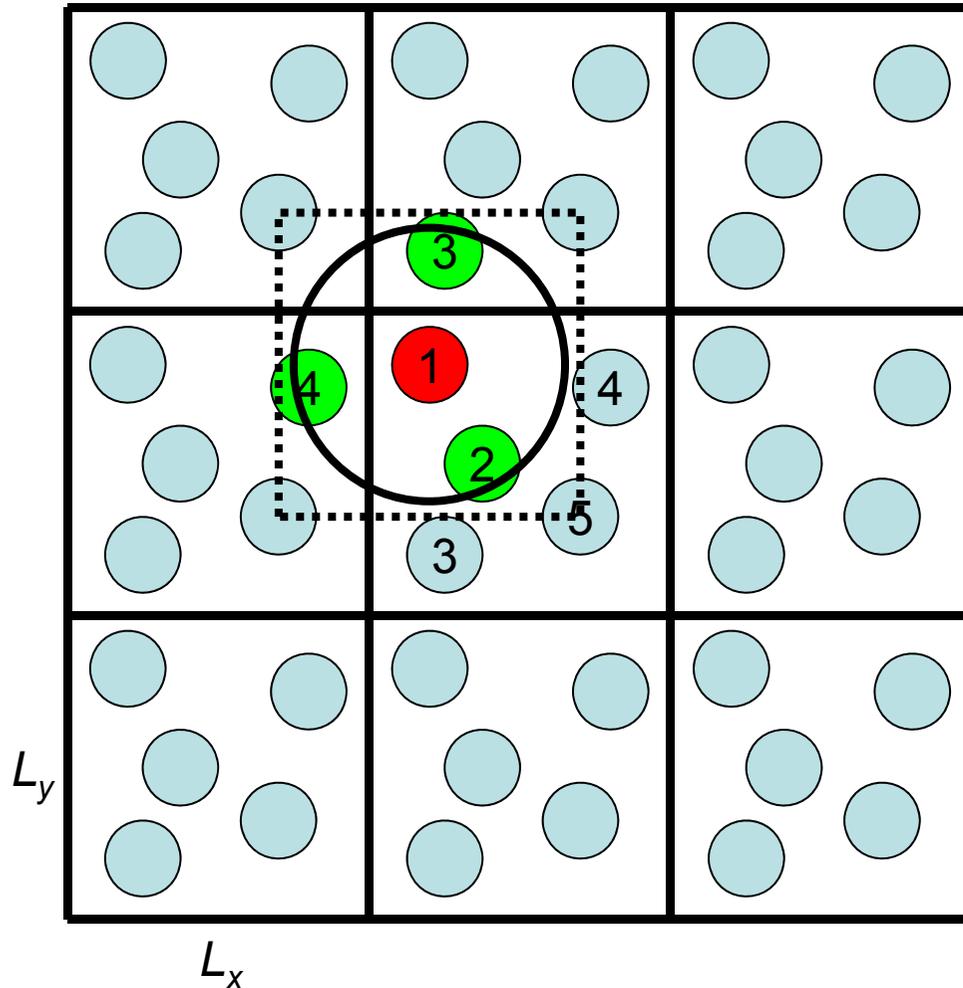
- 周期境界条件では基本セルのコピーが無限に続くので全ての原子ペアについて相互作用を**近似せずに**直接計算することは不可能

カットオフ法の適用



- カットオフ半径によっては、基本セルの周辺イメージセルも考慮する必要がある
(左の例では $26N^2 + N(N-1)/2$ ペアの計算が必要)

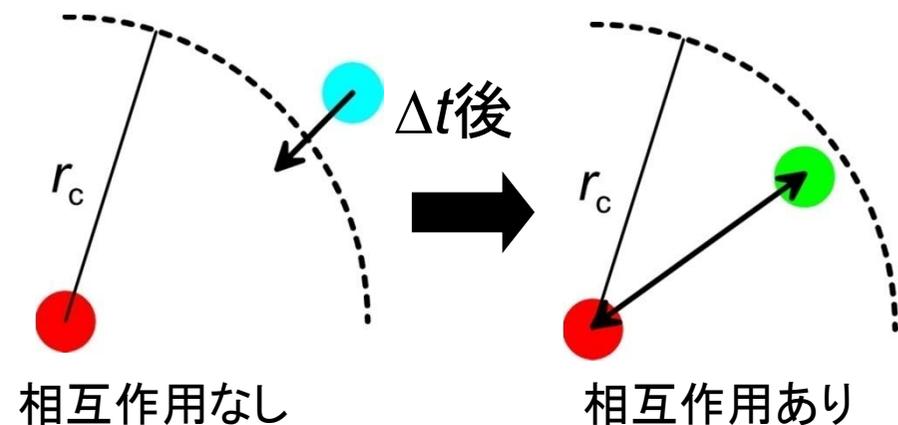
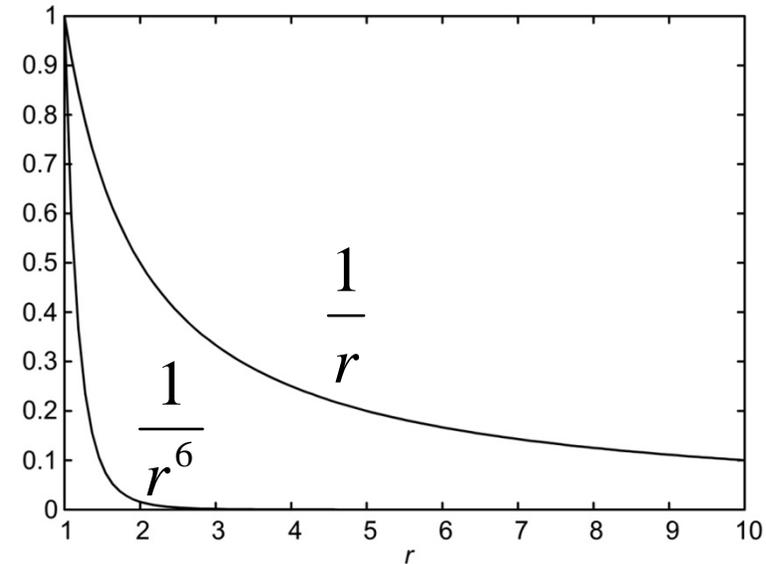
Minimum image convention



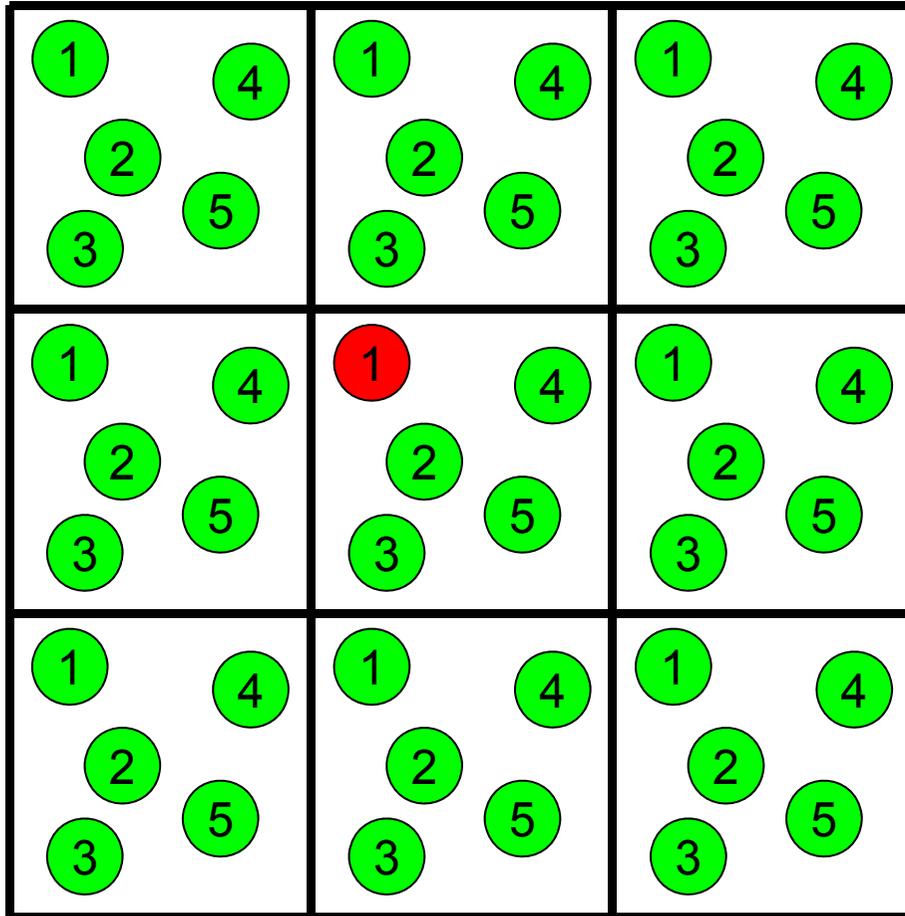
- カットオフ半径 r_c を最も短い基本セルの1辺の長さの2分の1以下にすれば考慮すべきペア数は $N(N-1)/2$ でよい
→ minimum image convention

カットオフ法の問題点

- Van der Waals相互作用は遠距離では、 r^{-6} の項が支配的
→ van der Waals相互作用はカットオフ法で十分な精度で計算可能
- 静電相互作用は r^{-1} に依存
→ カットオフ法では精度良く評価することが困難
- 原子がカットオフ半径の範囲から出入りする際にエネルギーが変動するため、全エネルギーは保存しない



カットオフしない計算法



中央の基本セル内の原子同士だけでなく、基本セル内の原子と周囲のイメージセル内の原子との間の相互作用も計算する

原子*i*の位置 \mathbf{r}_i における静電ポテンシャル:

$$\varphi(\mathbf{r}_i) = \sum_n \sum_j \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|}$$

$\mathbf{n} = 0$ の時は*i* = *j*となるペアは計算しない

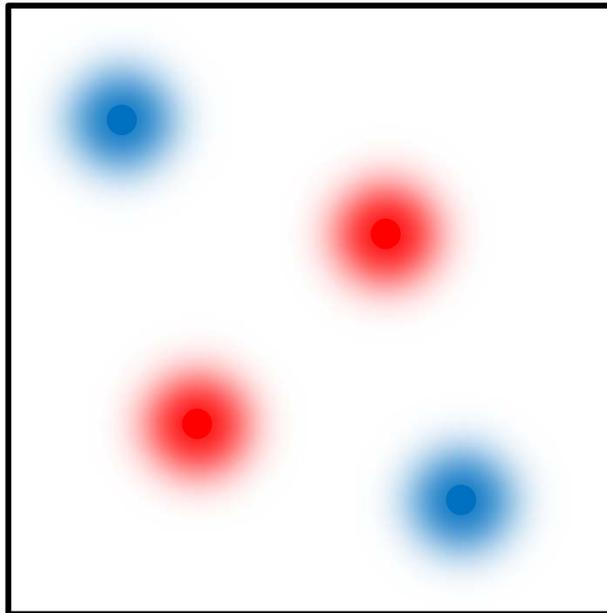
Particle Mesh Ewald法(1)

- 点電荷を以下の2つの電荷分布に分ける



$$\rho^\sigma(\mathbf{r}) = \sum_i q_i \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{3}{2}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_i|^2}{2\sigma^2} \right)$$

Particle Mesh Ewald法 (2)



- ガウス分布に従う電荷分布はなめらか
→ 高速Fourier変換を用いてPoisson方程式と解き、静電ポテンシャルを求める

$$\nabla^2 \varphi(\mathbf{r}) = -4\pi \rho^\sigma(\mathbf{r})$$

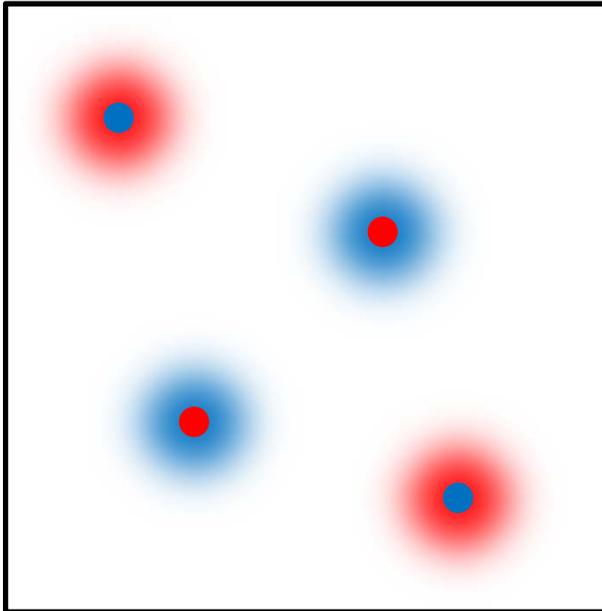


高速Fourier変換

$$\tilde{\varphi}(\mathbf{k}) = \frac{4\pi}{|\mathbf{k}|^2} \tilde{\rho}^\sigma(\mathbf{k})$$

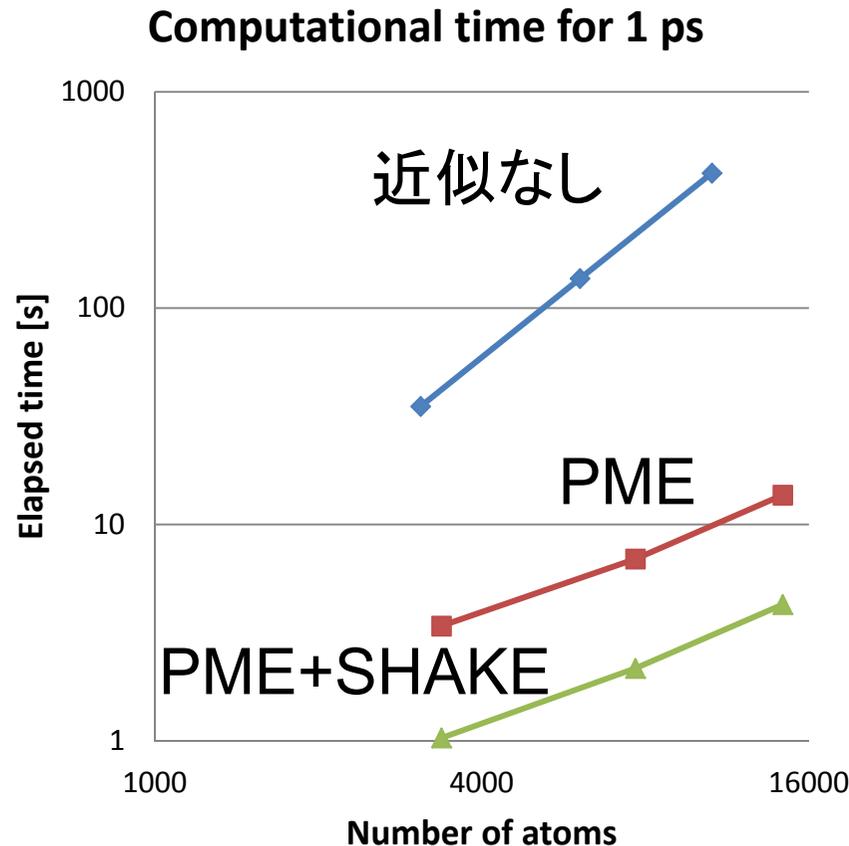
- 発散を防ぐため、全電荷は0にする必要がある

Particle Mesh Ewald法 (3)



- 残りの電荷では、点電荷のまわりに、これを打ち消す反対の符号の電荷が分布
 - 静電ポテンシャルは r^{-1} より速く0に減衰
 - カットオフ法でも精度よく計算できる

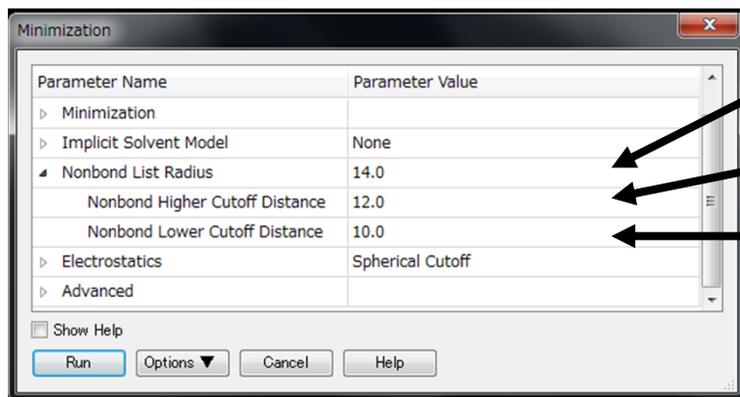
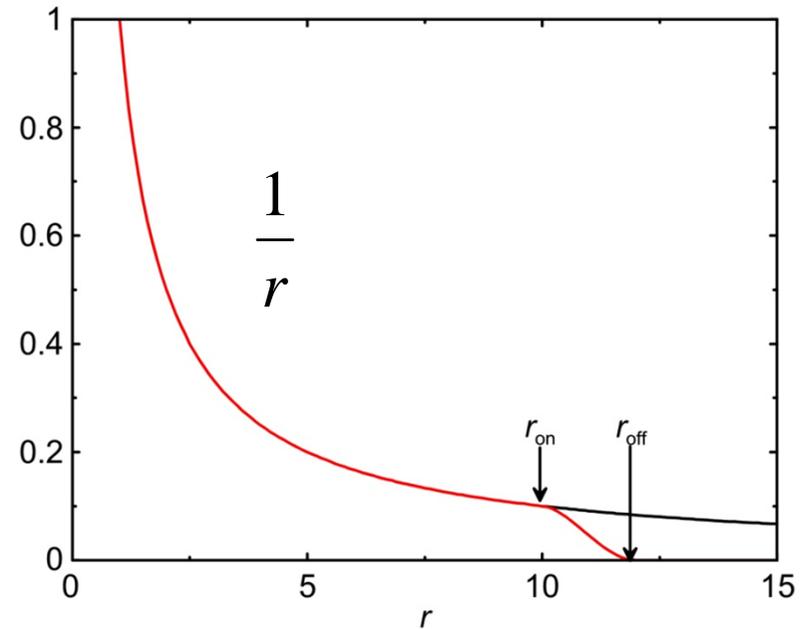
計算時間(2)



- 水分子の系で計算時間を計測
- 「近似なし」では原子数 N の2乗に比例
- PMEを使用することでほぼ $N \log N$ に比例
- SHAKEを併用することで時間刻みを4倍(2 fs)にでき、計算速度は3.2倍程度高速化した

Discovery Studioにおける設定(3)

- 標準では静電相互作用の計算にはカットオフ法が用いられる
- カットオフによるエネルギーの不安定性を緩和するため、スイッチング関数を用いられている



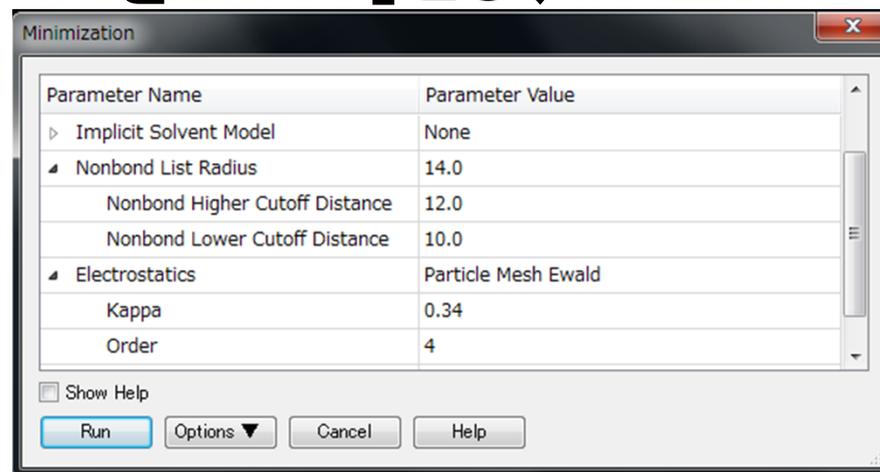
r_l (ペアリスト作成用)

r_{off}

r_{on}

Discovery Studioにおける設定(4)

- Electrostaticsに「Particle Mesh Ewald」を選ぶとこの方法を用いて静電相互作用の計算が行われる
- van der Waals相互作用はカットオフ法で計算される
- パラメータ σ はKappa ($\kappa = 1/\sqrt{2}\sigma$) で与えられ、 $3/r_{\text{off}}$ と $6/r_{\text{off}}$ の間に設定する
- 系が電氣的に中性でない場合は、SolvationでAdd Counterionを「True」とし、counterionを配置して中性にする

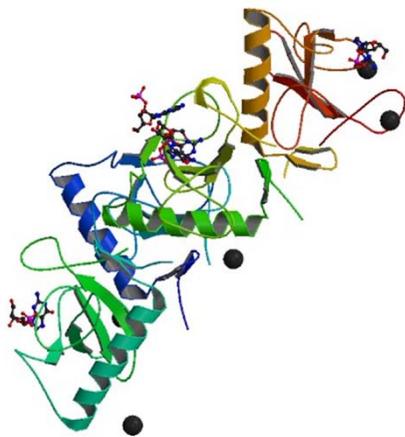


シミュレーションの手順

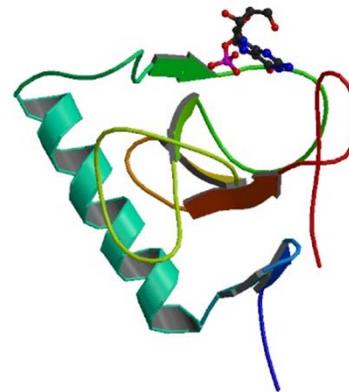
1. 初期構造の作成
 - 立体構造の取得
 - 欠失残基への対応
 - 水素原子付加
 - リガンドのモデリング
 - 力場パラメータの取得
 - 水分子の配置
2. 立体構造最適化
3. 初期速度の割り当て
4. 平衡化
5. プロダクションラン

初期構造の作成(1)

- 立体構造の取得
 - PDB (<http://www.rcsb.org/pdb/>) からダウンロード
 - 通常、生物学的に機能しうる単位である biological unit 構造に対してシミュレーションを行う
 - 例: Ribonuclease T1 (PDB ID: 1I0X)



非対称単位の構造



Biological Unitの構造

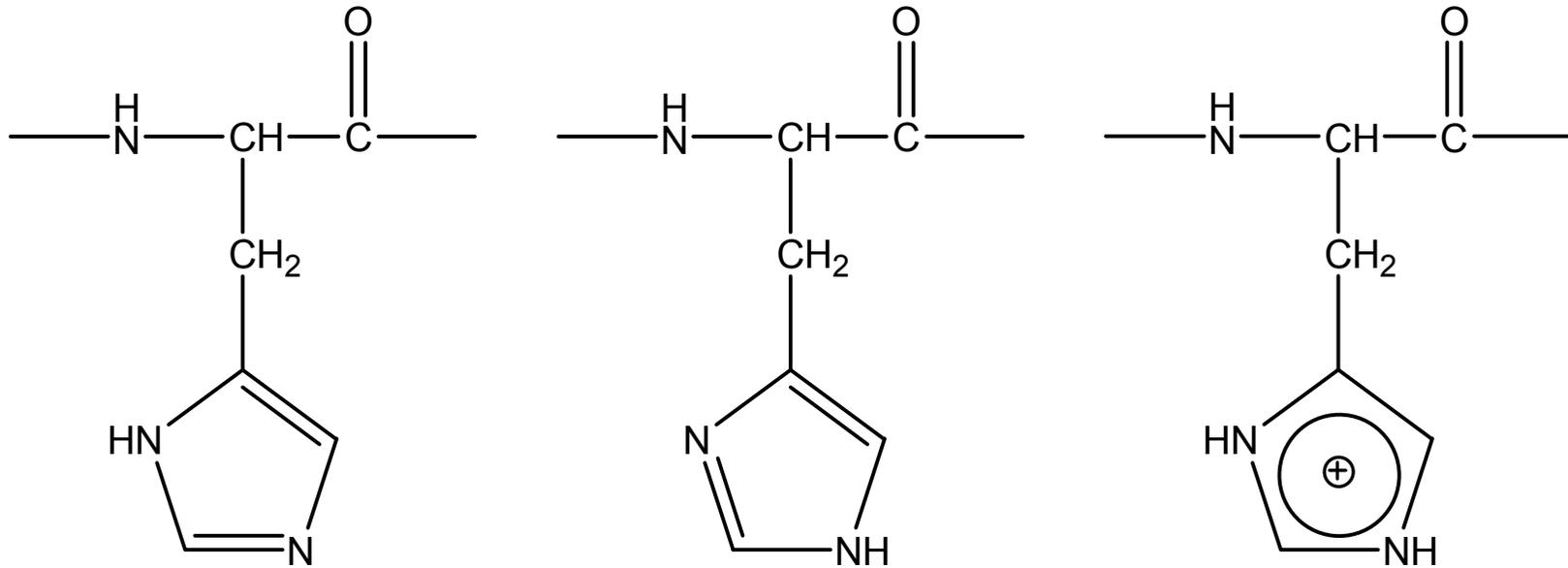
初期構造の作成(2)

- 欠失残基への対応
 - 欠失残基はモデリングなどで補う
 - N末端、C末端が欠失している場合は、欠失残基の前後の残基をacetyl基、N-methyl基でブロックしても良い
- 水素原子付加
 - 基本的に自動的に付加できる
 - SS結合の有無、Hisのプロトン化状態に注意

Discovery Studioでの操作(1)

1. 「File」→「Open URL」でIDに「1I0X」を指定して「Open」
2. 「Display Style」をLineに変更
3. Hierarchy WindowでB, C, D鎖を選択し削除
4. 「Macromolecules」ボタンをクリックし、Toolsタブに表示されるProtein Reportを展開する
5. 「Protein Report」をクリック
→Incomplete or Invalid Residues (Lys41、Asp49、Glu102; 紫色で表示される) に注意
6. ToolsタブのPrepare Proteinを展開し、Manual Preparationの「Clean Protein」をクリック→欠失原子が補われる

Hisのプロトン化状態



δ 位にプロトン化

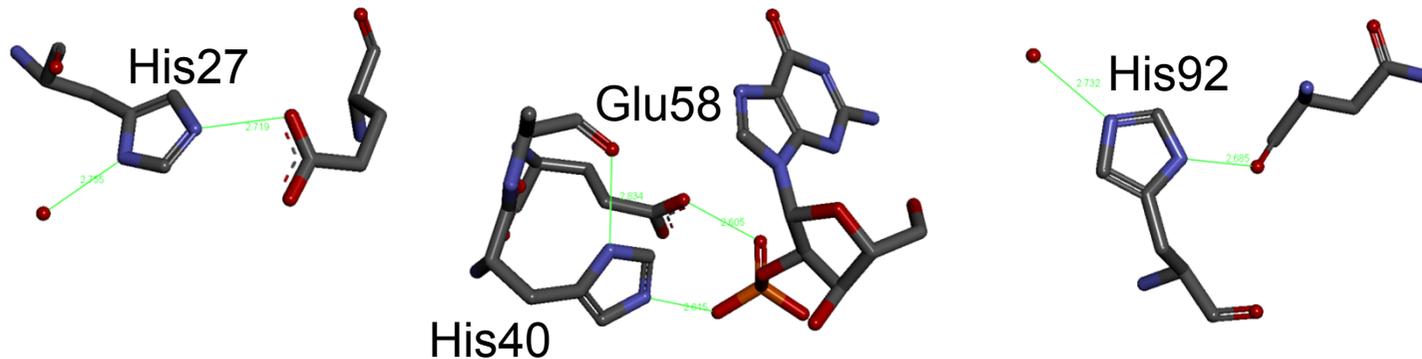
ϵ 位にプロトン化

δ, ϵ 位にプロトン化

- His側鎖の pK_a は中性付近であるため2つの窒素原子とも水素原子が結合した状態も十分にとりうる
- His周りの水素結合ネットワークからプロトン化状態がわかる

Discovery Studioでの操作(2)

7. His27、His40、Glu58、His92がどのような相互作用をしているか確認



8. Forcefieldを「CHARMm」とした後、Prepare ProteinのProtonate Proteinにある、「Calculate Protein Ionization and Residue pKa」をクリックし「Run」→ 上記残基のプロトン化状態を確認

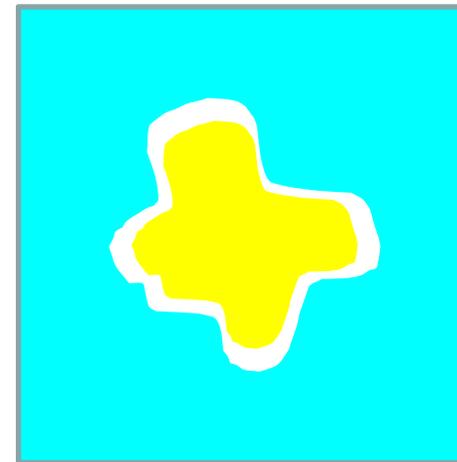
初期構造の作成(3)

- リガンドのカ場パラメータの取得
 - リガンドのカ場パラメータは分子動力学ソフトウェアに含まれていないので、自分で作成するか、Amber Parameter Database*等から取得する
- 水分子の配置
 - PMEを利用して高精度かつ高速にシミュレーションを行うため水分子を直方体状に配置
 - 電荷を中性にするためにカウンターイオンを配置

*<http://www.pharmacy.manchester.ac.uk/bryce/amber>

平衡化

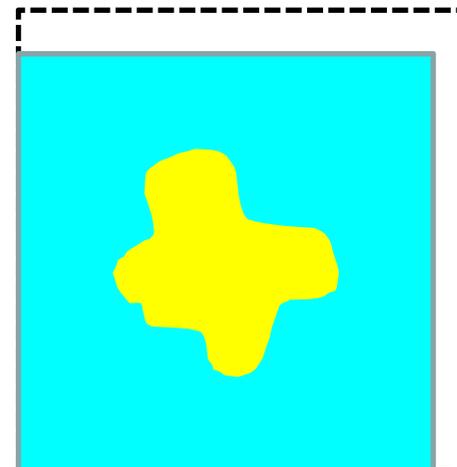
- 初期構造では、配置した水分子とタンパク質の間に隙間がある
- 定温定圧シミュレーションを行い、水分子の配置を最適化する
- その際、タンパク質の原子が初期位置からあまり動かないように束縛し、1 ns程度かけて徐々に束縛を緩める



タンパク質の周りに隙間がある



束縛付き定温定圧シミュレーション



体積が減少

課題の提出

- 課題1、課題2の結果と考察を1つのPowerPointファイルにまとめて、寺田宛 tterada@iu.a.u-tokyo.ac.jp に送ること
- その際件名は「分子モデリング課題」とし、本文に氏名と学生証番号を明記すること