

平成23年6月8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

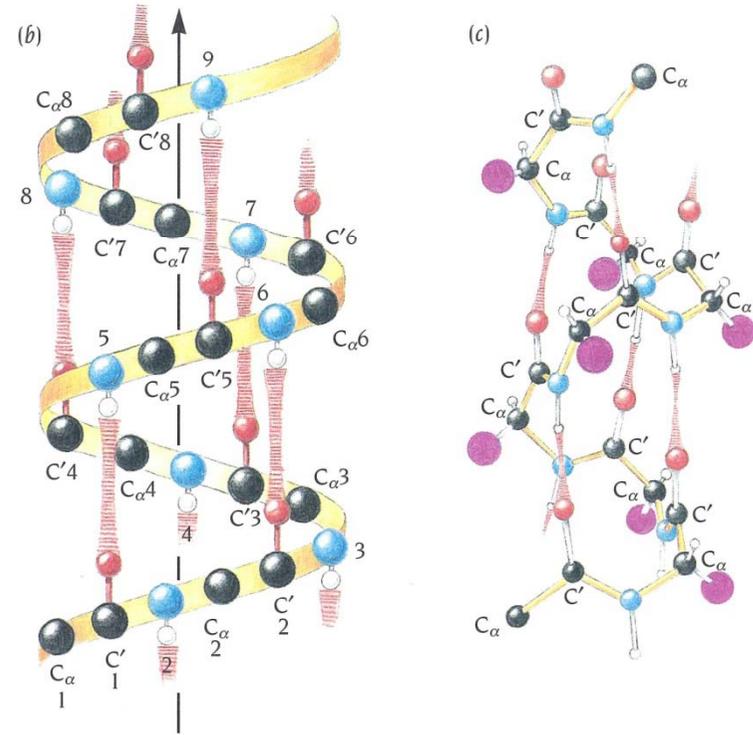
- ペプチドのエネルギー最小化
- 水溶液環境のモデル化
- ペプチドの分子動力学シミュレーション
 - 課題1
- シミュレーションの高速化
- タンパク質の分子動力学シミュレーション
 - 課題2
- 応用事例の紹介

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する
 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

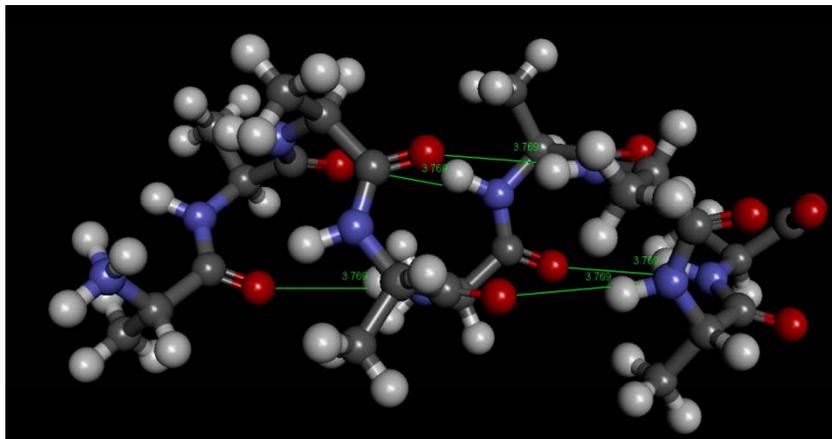


α helixでは*i*番目のカルボニル酸素と*i*+4番目のアミド窒素が水素結合を形成する

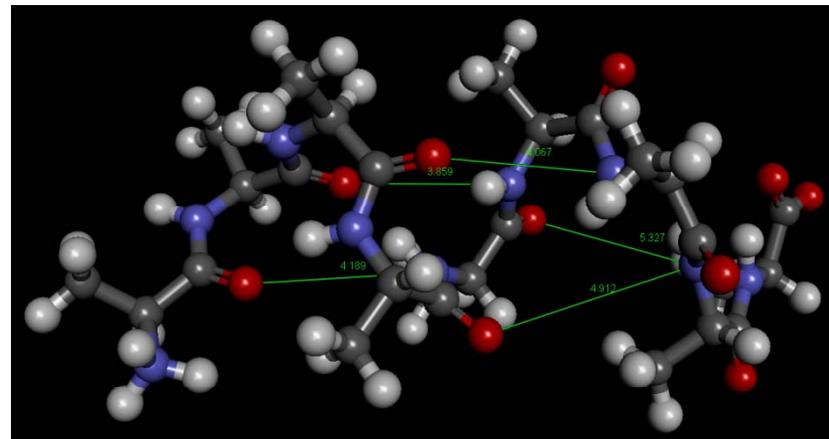
ペプチドのエネルギー最小化

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

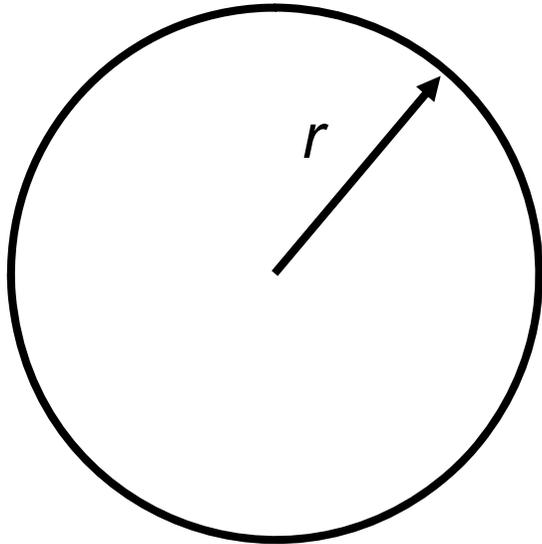
水溶液環境のモデル化(1)

- Alanineペプチドは水溶液中で α helix構造を安定にとりうると考えられている
- 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



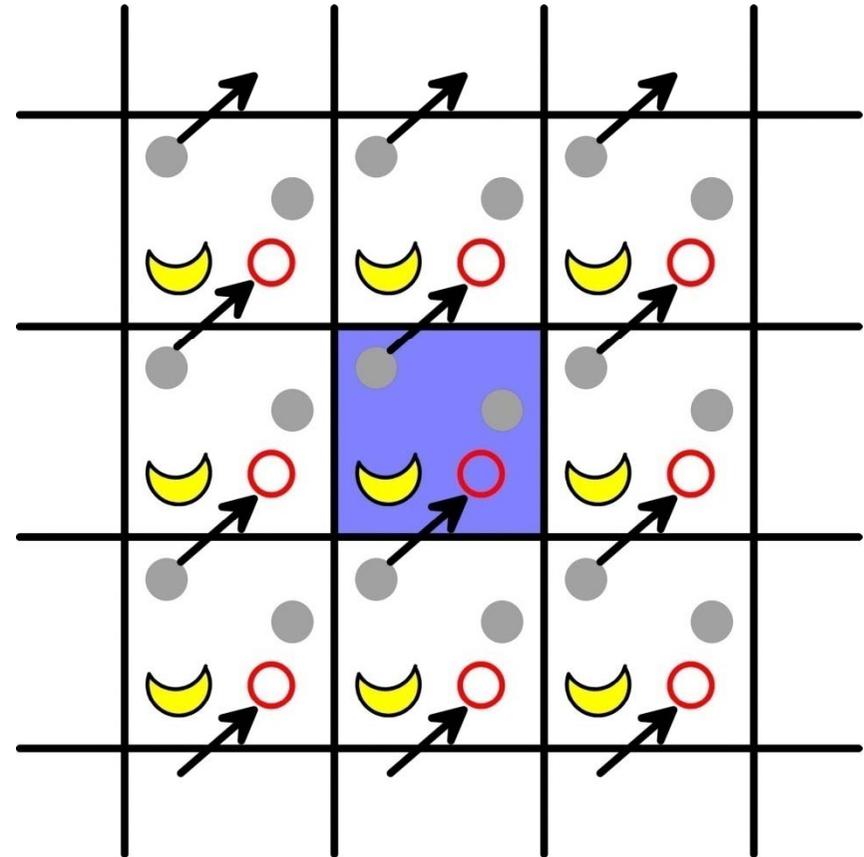
$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

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

- 分子が系の中心 (x_0, y_0, z_0) から半径 r_{Cap} の外側に出て行こうとすると系の中心に向けて束縛力が働く
- 水分子を生体高分子の周りに球状に配置することを”Cap”と呼ぶことがある

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

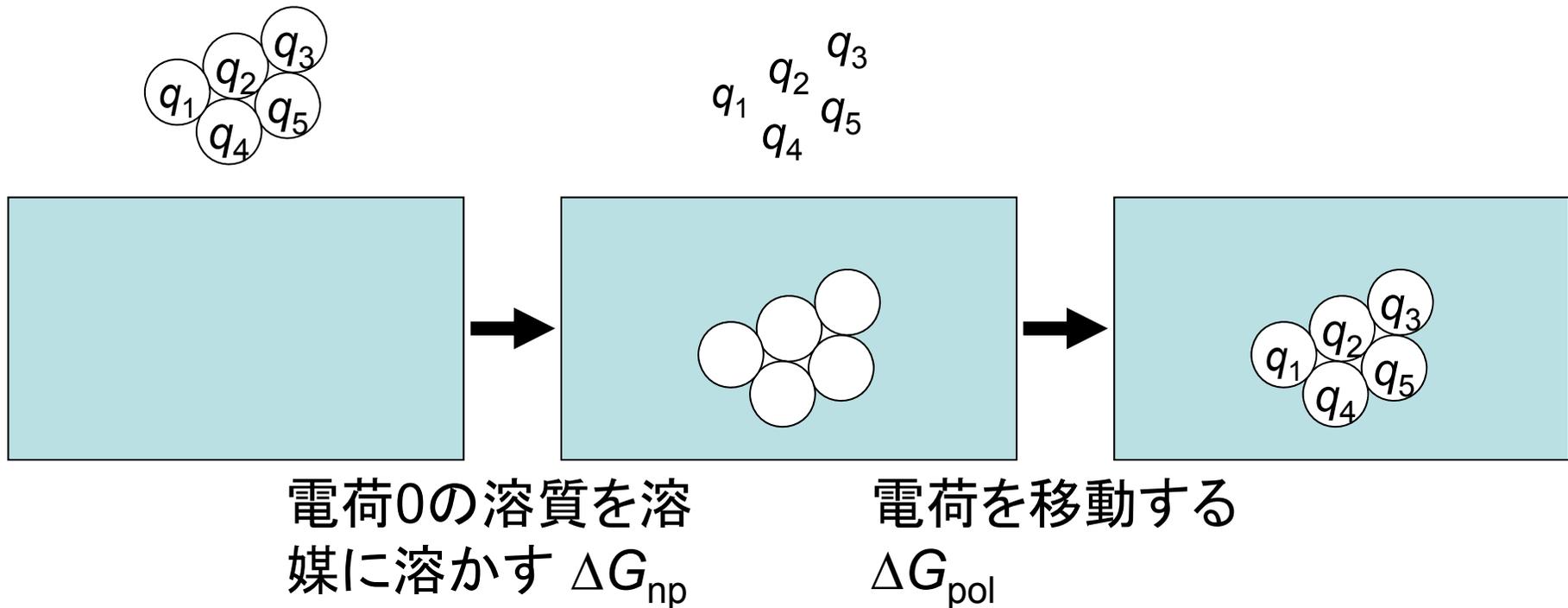


水溶液環境のモデル化(3)

- 水分子を球状に配置
 - 系の外側は真空であるので、水が蒸発していかないように束縛する必要がある
 - 系の表面に位置する水分子はバルクの水とは異なる環境に置かれる
- 周期境界条件
 - どの分子も同じ環境
 - 系の分子が周期性を感じないように、系を十分に大きくする必要がある

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

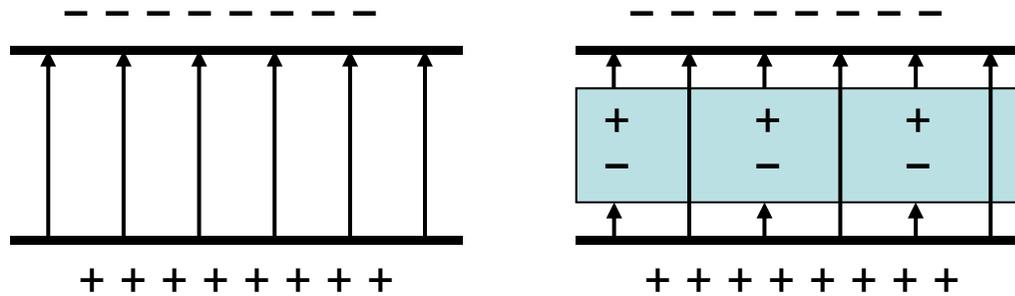
$$\Delta G_{np} = \sigma A + b \quad A: \text{SASA}, \sigma, b: \text{経験的パラメータ}$$

- 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

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

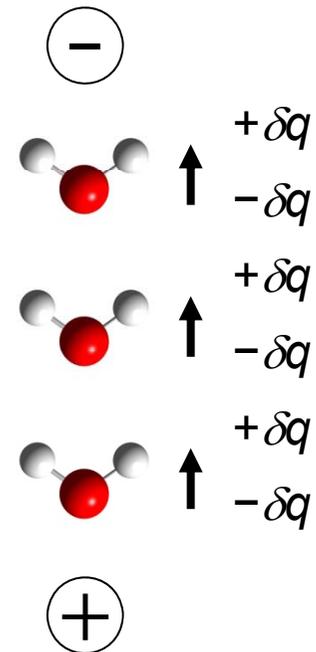
静電ポテンシャル 溶質の電荷分布

誘電体



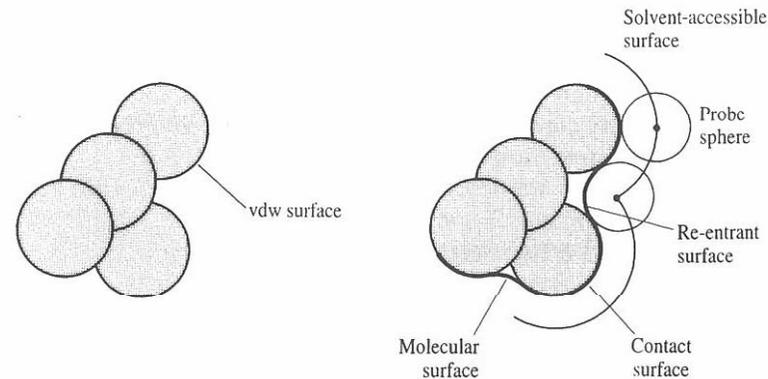
コンデンサーに比誘電率 ϵ の誘電体を挿入すると、誘電体の表面に電荷が現れ、極板間の電場を打ち消す
→静電ポテンシャルは $1/\epsilon$ となる

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

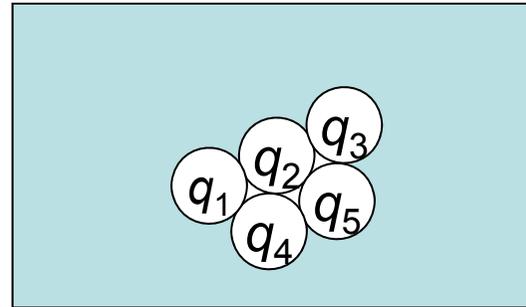
Generalized Bornモデル(1)

- Poisson-Boltzmann方程式は数値的にしか解けないので、ポテンシャルエネルギーの勾配(すなわち力)が必要な分子動力学シミュレーションでは使いにくい
- 空間を細かいグリッドに切って計算するため、計算コストが非常に大きい
- Generalized BornモデルはPoisson-Boltzmann方程式の近似解を低い計算コストで与えることができるため最近良く用いられている
- 力を解析的に求めることが可能

Generalized Bornモデル(2)

- イオンの溶媒和自由エネルギー (Bornの式)

$$\Delta G_{\text{pol}} = -\frac{q^2}{2a} \left(1 - \frac{1}{\epsilon} \right)$$



q : イオンの電荷、 a : イオン半径、 ϵ : 溶媒の比誘電率

- 一般の分子に拡張

$$\Delta G_{\text{pol}} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon} \right) \sum_{i,j=1}^N \frac{q_i q_j}{f_{\text{GB}}} \quad \begin{array}{l} R_i: \text{原子} i \text{の有効Born半径} \\ \text{溶質の形に依存} \end{array}$$

$$f_{\text{GB}} = \sqrt{r_{ij}^2 + R_i R_j} \exp\left(-r_{ij}^2 / 4R_i R_j\right) \quad r_{ij}: \text{原子} i, j \text{間距離}$$

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).

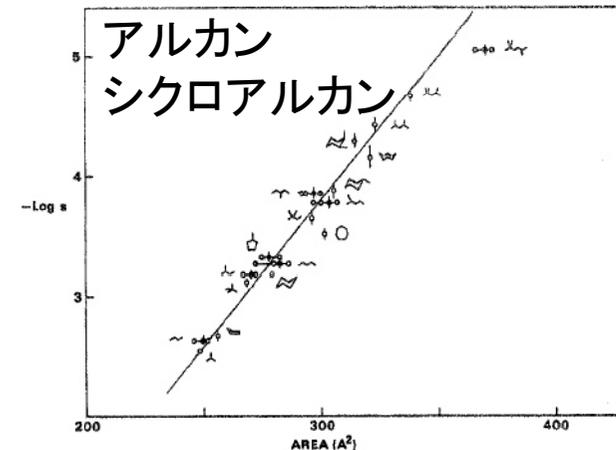


Figure 2. Relationship of cavity surface area to solubility for alkanes and cycloalkanes.

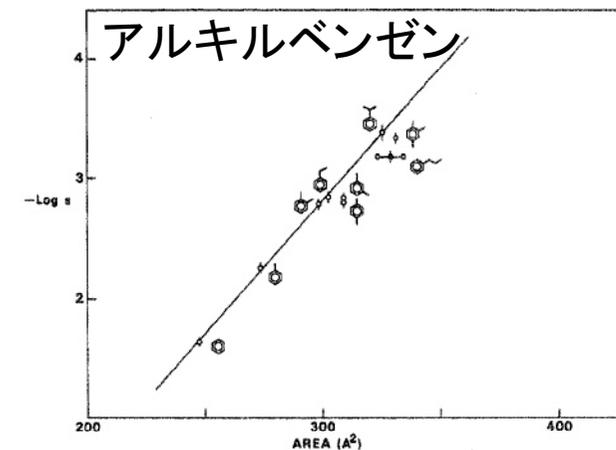


Figure 3. Relationship of cavity surface area to solubility for alkylbenzenes.

Discovery Studioでの操作(1)

- 水分子を配置する場合
 1. Molecule Windowに計算対象の分子を表示する
 2. 「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
 3. 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」
 - 最急降下法によるエネルギー最小化 (Minimization)、共役勾配法によるエネルギー最小化 (Minimization2)、加熱 (Heating)、平衡化 (Equilibration)、プロダクションラン (Production) の順に実行される
 - 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

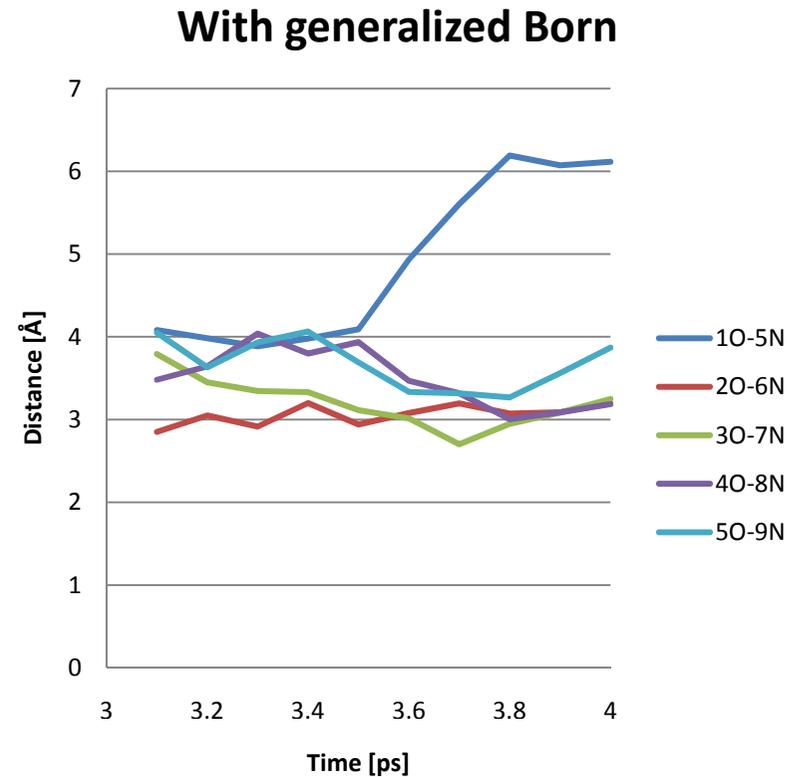
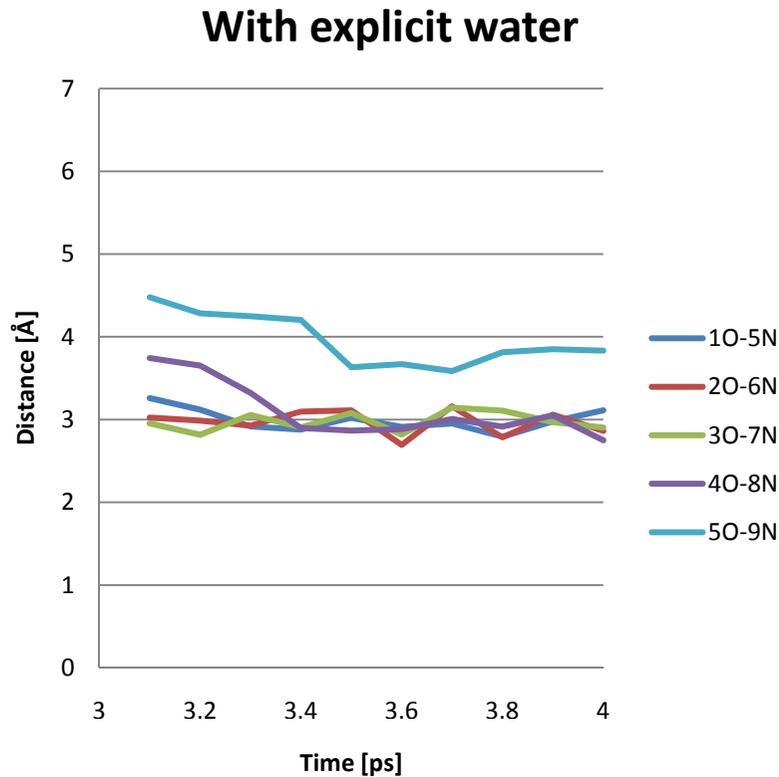
ペプチドのMDシミュレーション(2)

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

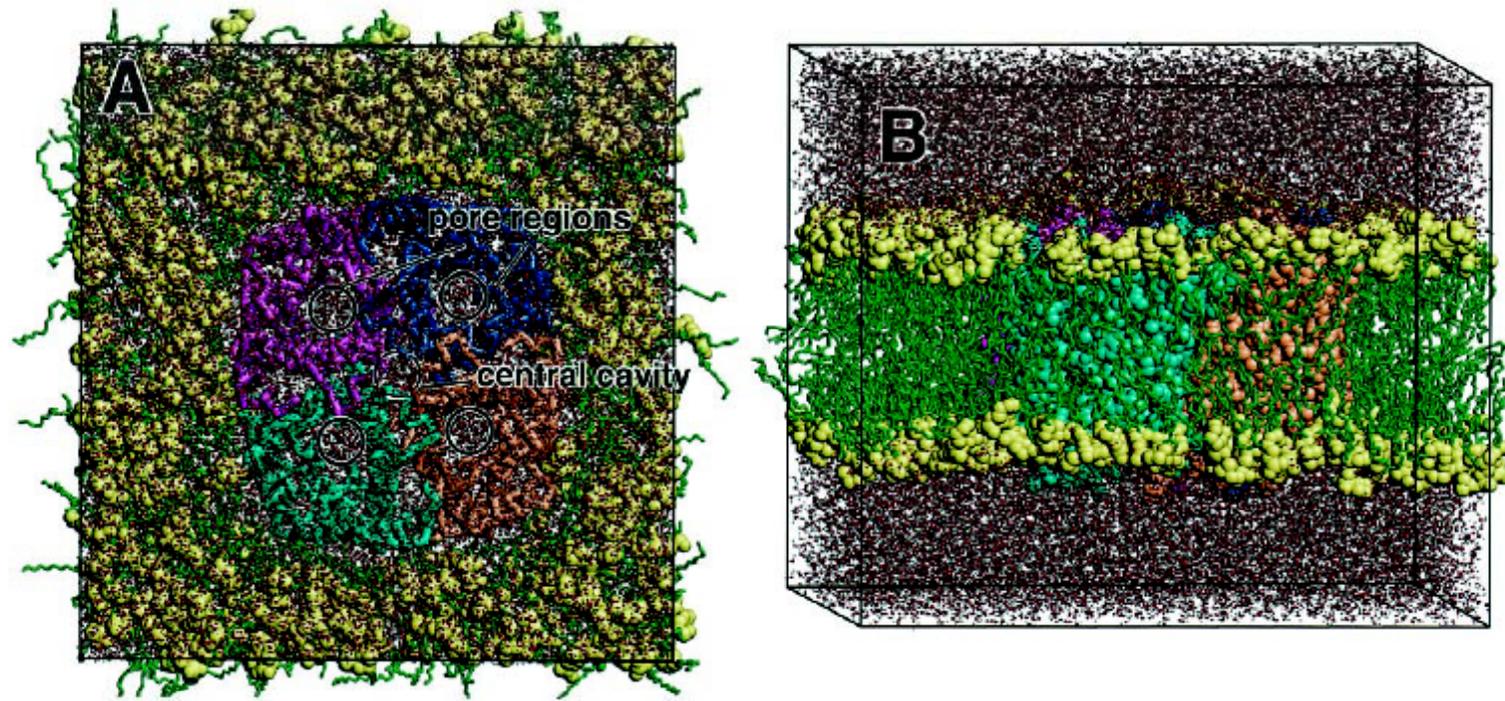
水素結合距離の時間変化



- 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

参考：膜環境のモデル化(1)

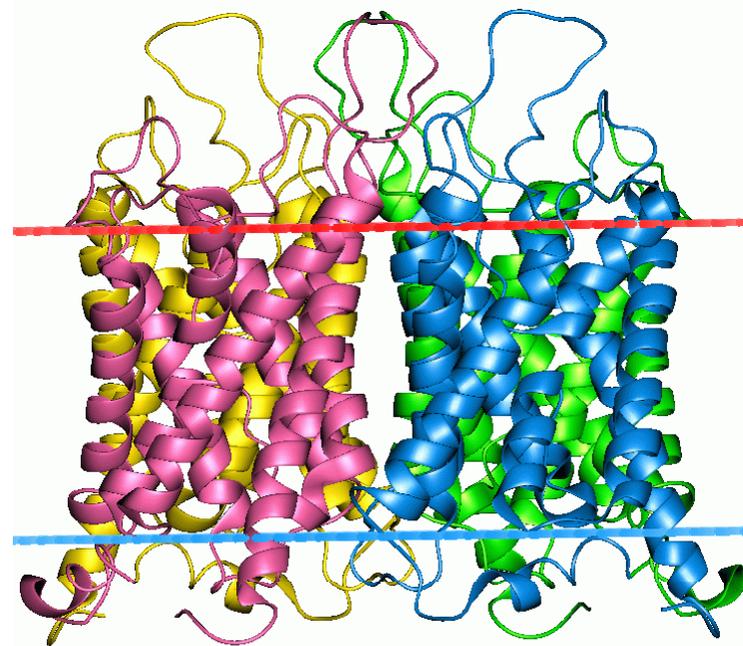
- タンパク質を脂質2重膜に埋め込む
- 膜の両側に水分子を配置



de Groot & Grubmuller, *Science* **294**, 2353 (2001)

参考：膜環境のモデル化(2)

- Orientations of Proteins in Membranes (OPM) database
 - URL: <http://opm.phar.umich.edu/>
 - 膜タンパク質が膜にどのような向き・位置で埋め込まれているか予測しデータベース化している



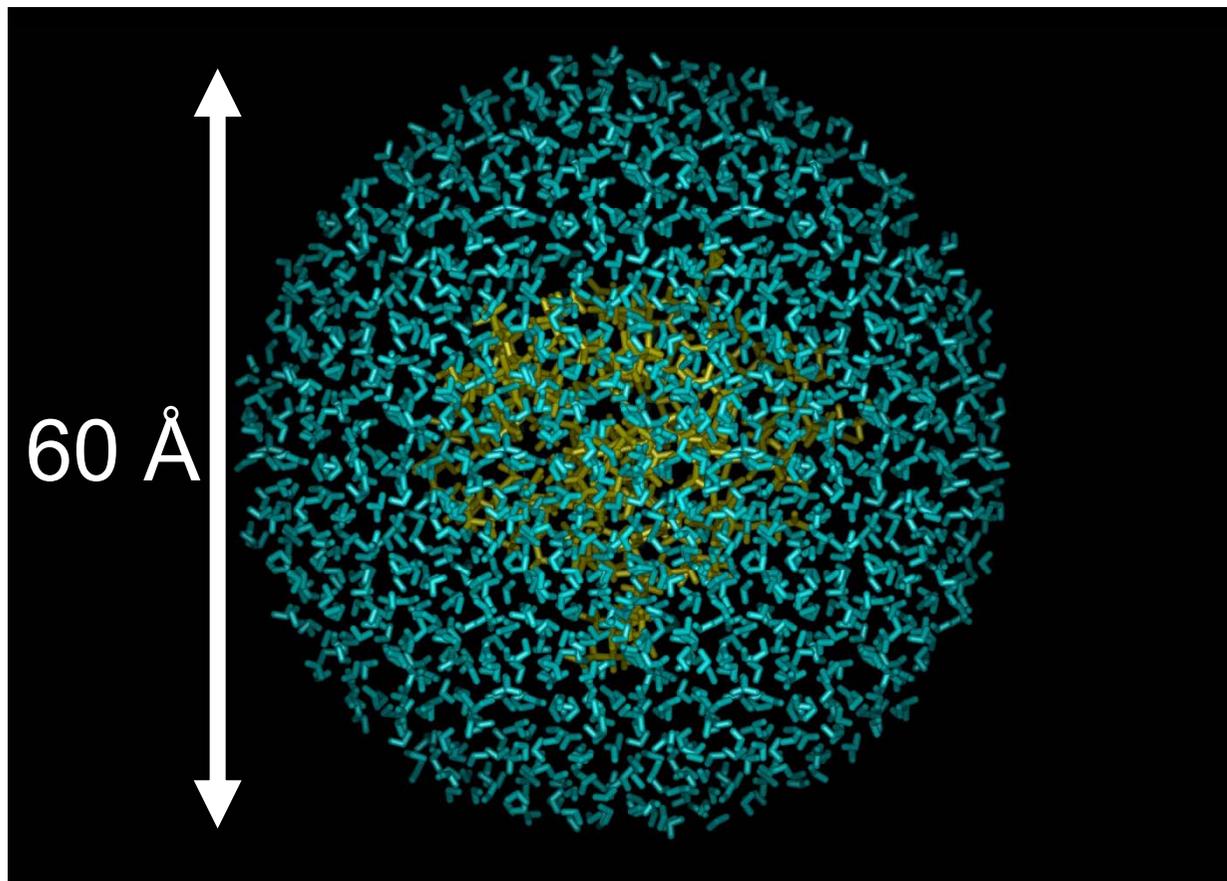
計算時間の例

- 対象: 球状に配置した水分子 (TIP3Pモデル)
- Amber 8のSanderモジュール使用
- 計算にはIntel Xeon Processor 2.8 GHzを備えたPCを使用
- 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	計算時間	比率	非共有結合相互作用の計算時間	割合
3087	912	1.0	911	0.9983
6066	3955	4.3	3951	0.9992
10608	16918*	18.5	16911	0.9996

*4.7時間 1 nsのシミュレーションに196日かかる

計算対象の例



タンパク質: 1123原子、水分子: 3363分子
合計: 11212原子

分子運動の時間スケール(1)

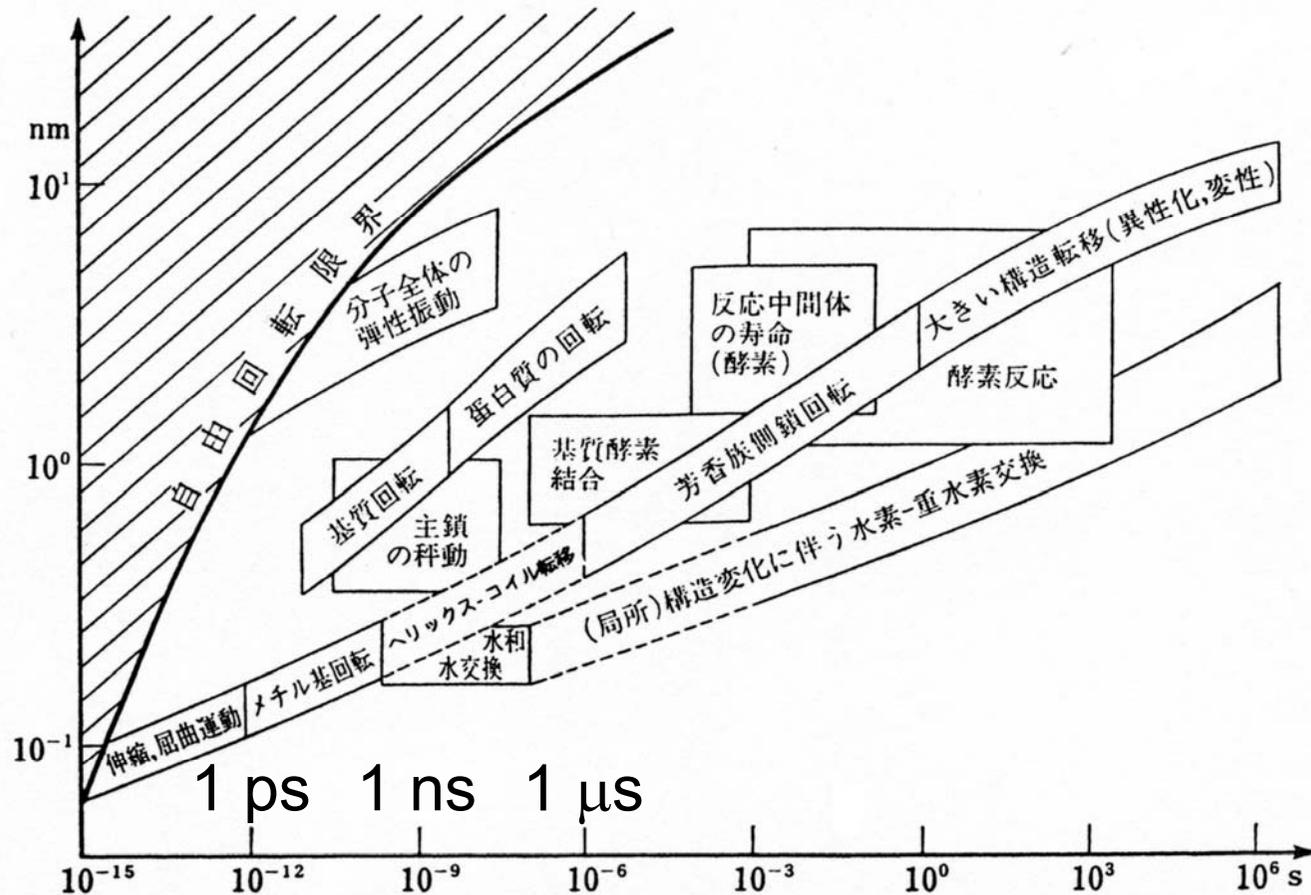


図 4.17 タンパク質分子の構造変化や化学反応事象の時間-空間相関プロット

時間軸と空間軸のケタ数の違いに注目。

永山國昭 「生命と物質 生物物理学入門」より引用

分子運動の時間スケール(2)

- 現状の計算機的能力では、何らかの工夫をしなければ、生物学的に意味のある時間スケール(~10 ns以上)のシミュレーションはできない
- 分子動力学法になんらかの近似を導入して、効率化を図る必要がある

分子シミュレーションの効率化

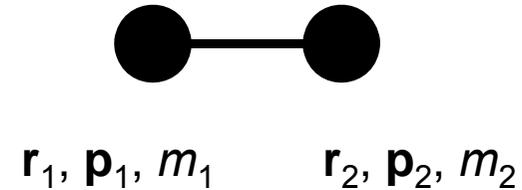
- 時間刻み Δt を長くする
 - SHAKE/RATTLE法
 - 多重時間積分法
- 非共有結合相互作用の計算の近似
 - カットオフ法
 - 多重極子展開法
 - Ewald法
- 本講義では赤枠の3つの方法について解説

時間刻みの選択

- 一般的に、最も早い運動の周期の10分の1から20分の1程度の時間刻みがいられる
- 最も速い運動は、X-H伸縮運動(X=C、N、O、S)で、 3000 cm^{-1} 程度(周期は約10 fs)
- 時間刻み Δt には0.5~1.0 fsを用いる
- X-H伸縮運動の次に速い運動はX-X伸縮運動で 1500 cm^{-1} 程度
- 従って、X-H伸縮運動を除くことができれば時間刻みは2倍にできる
- X-H結合長を固定→SHAKE/RATTLE法

SHAKE/RATTLE法(1)

簡単のために2原子分子を考える
束縛力をかけて2原子間の距離を
 d に固定する



1. 2原子間に働く束縛力を未定係数 λ を用いて表し、velocity Verlet法に従って、時刻 $t+\Delta t$ における座標を計算する

$$\begin{cases} \mathbf{v}'_1 = \mathbf{v}_1(t) + [\mathbf{f}_1(\mathbf{r}(t)) + \lambda \mathbf{r}_{12}(t)] \frac{\Delta t}{2m_1} \\ \mathbf{r}_1(t + \Delta t) = \mathbf{r}_1(t) + \mathbf{v}'_1 \Delta t \end{cases} \quad \begin{cases} \mathbf{v}'_2 = \mathbf{v}_2(t) + [\mathbf{f}_2(\mathbf{r}(t)) - \lambda \mathbf{r}_{12}(t)] \frac{\Delta t}{2m_2} \\ \mathbf{r}_2(t + \Delta t) = \mathbf{r}_2(t) + \mathbf{v}'_2 \Delta t \end{cases}$$

2. 以下の2次方程式を解き、2原子間の距離が d になるように未定係数 λ を求め、1の式に代入する→SHAKE

$$d^2 = |\mathbf{r}_1(t + \Delta t) - \mathbf{r}_2(t + \Delta t)|^2$$

SHAKE/RATTLE法(2)

3. 新しい座標を用いて力 $\mathbf{f}(\mathbf{r}(t+\Delta t))$ を計算する
4. 2原子間に働く束縛力を未定係数 λ を用いて表し、時刻 $t+\Delta t$ における運動量を計算する

$$\begin{cases} \mathbf{v}_1(t+\Delta t) = \mathbf{v}_1 + [\mathbf{f}_1(\mathbf{r}(t+\Delta t)) + \lambda \mathbf{r}_{12}(t+\Delta t)] \frac{\Delta t}{2m_1} \\ \mathbf{v}_2(t+\Delta t) = \mathbf{v}_2 + [\mathbf{f}_2(\mathbf{r}(t+\Delta t)) - \lambda \mathbf{r}_{12}(t+\Delta t)] \frac{\Delta t}{2m_2} \end{cases}$$

5. 運動量に関する以下の束縛条件を満たすように未定係数 λ を決定し、4に代入する→RATTLE

$$\begin{aligned} \frac{d}{dt} |\mathbf{r}_1(t+\Delta t) - \mathbf{r}_2(t+\Delta t)|^2 &= 2(\mathbf{r}_1(t+\Delta t) - \mathbf{r}_2(t+\Delta t)) \cdot (\dot{\mathbf{r}}_1(t+\Delta t) - \dot{\mathbf{r}}_2(t+\Delta t)) \\ &= 2(\mathbf{r}_1(t+\Delta t) - \mathbf{r}_2(t+\Delta t)) \cdot (\mathbf{v}_1(t+\Delta t) - \mathbf{v}_2(t+\Delta t)) = 0 \end{aligned}$$

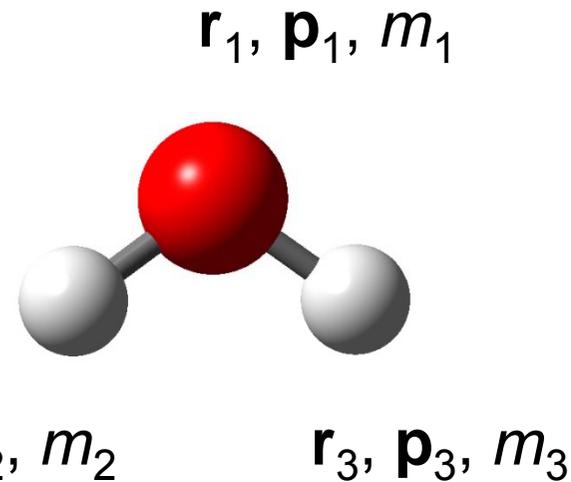
SHAKE/RATTLE法(3)

水分子の例(TIP3P)

$$d_{12} = |\mathbf{r}_1 - \mathbf{r}_2| = 0.9572 \text{ \AA}$$

$$d_{23} = |\mathbf{r}_2 - \mathbf{r}_3| = 1.5136 \text{ \AA}$$

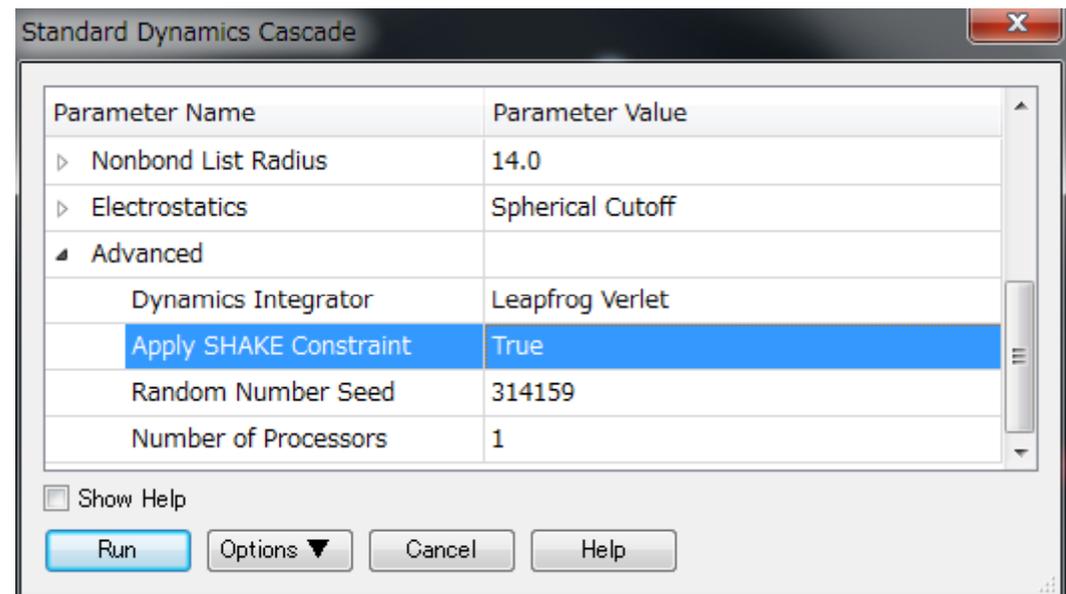
$$d_{31} = |\mathbf{r}_3 - \mathbf{r}_1| = 0.9572 \text{ \AA}$$



- このような場合は、連立2次方程式が現れるので解析的に解けない→誤差が一定値(たとえば 10^{-6} Å)以下になるまで繰り返し計算により求める
- TIP3Pモデルは剛体を前提に力場パラメータが決められているので必ずSHAKE/RATTLEを使う

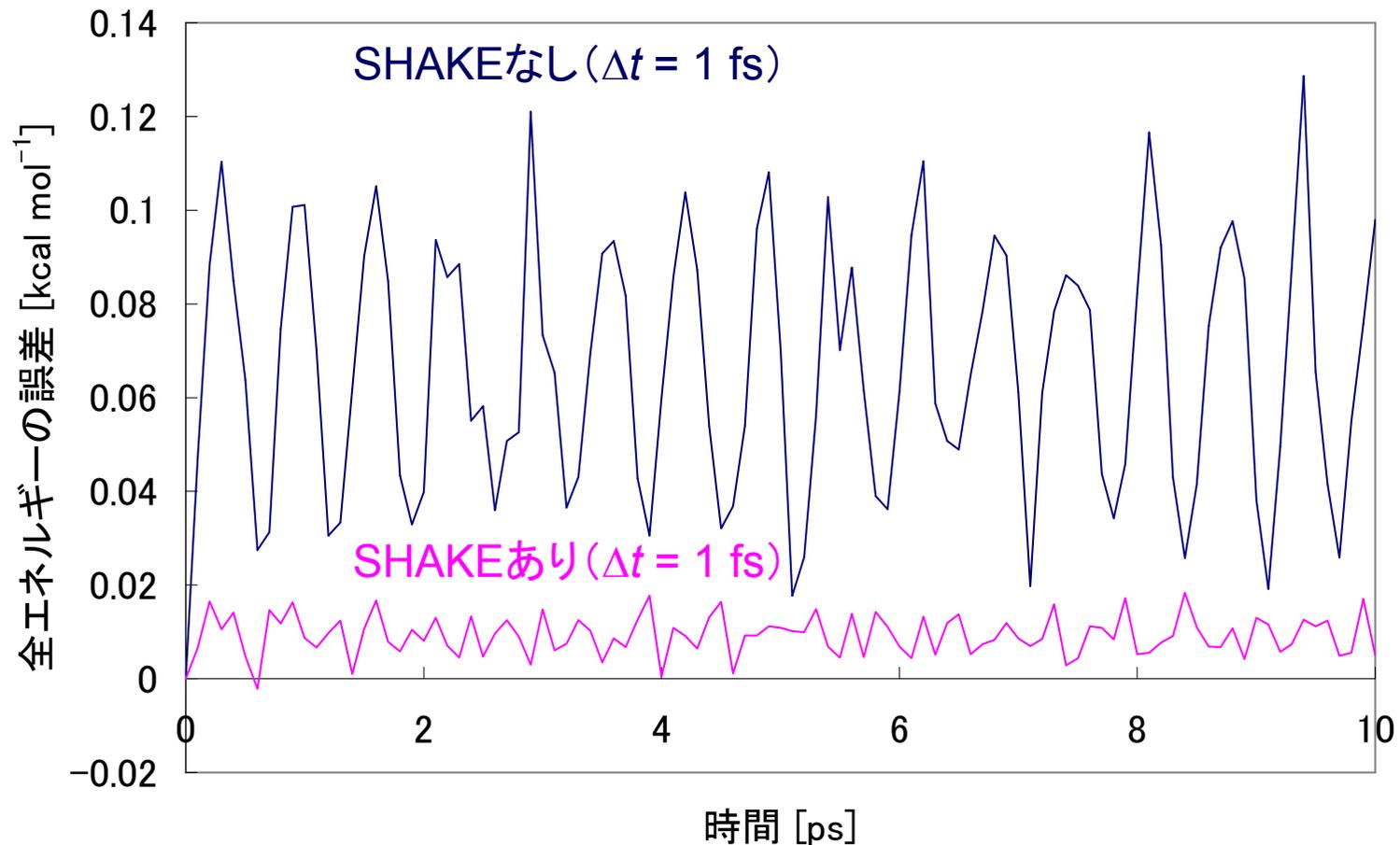
Discovery Studioでの操作(2)

- 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

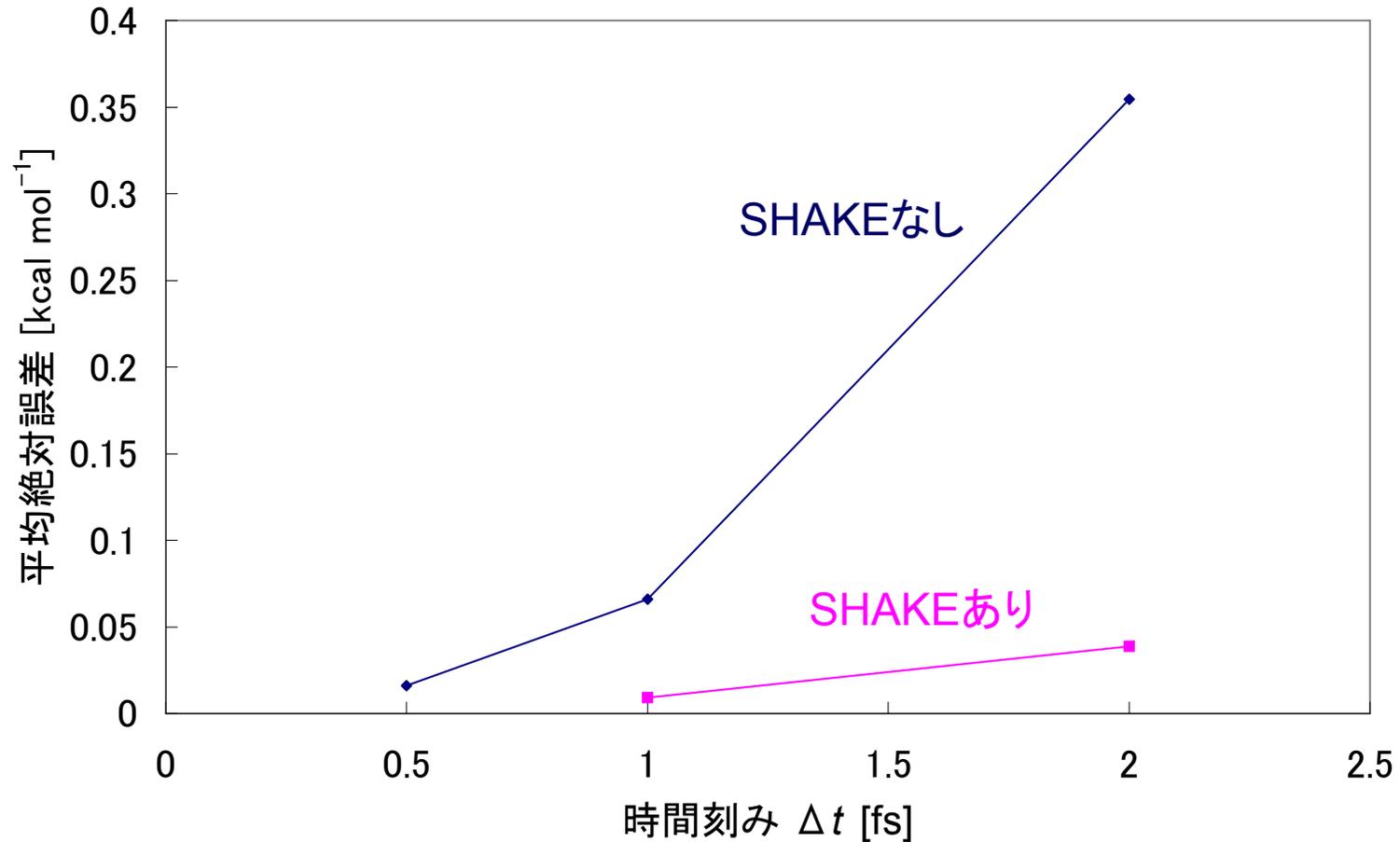


SHAKEの適用例

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

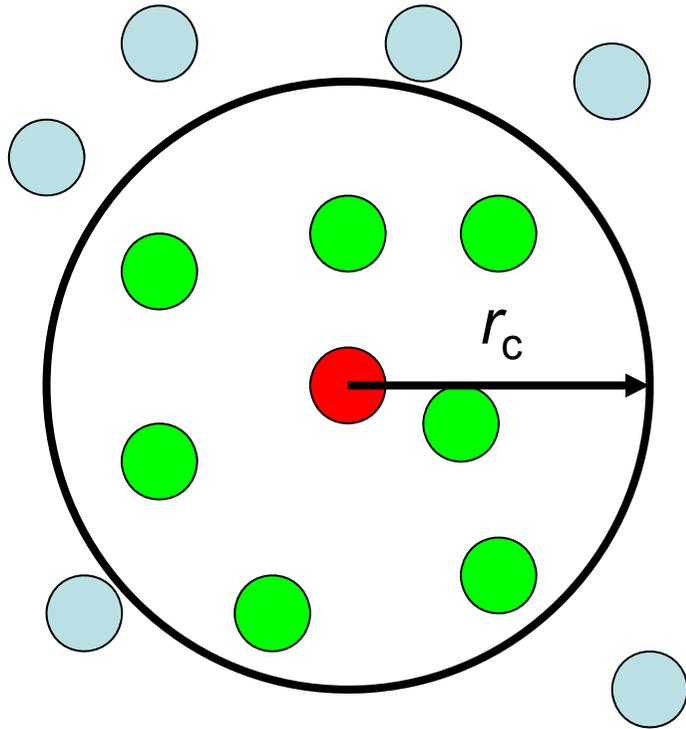


SHAKEを用いると時間刻み2 fsでもSHAKEなしの0.5 fsに匹敵する精度が得られる

非共有結合相互作用の扱い

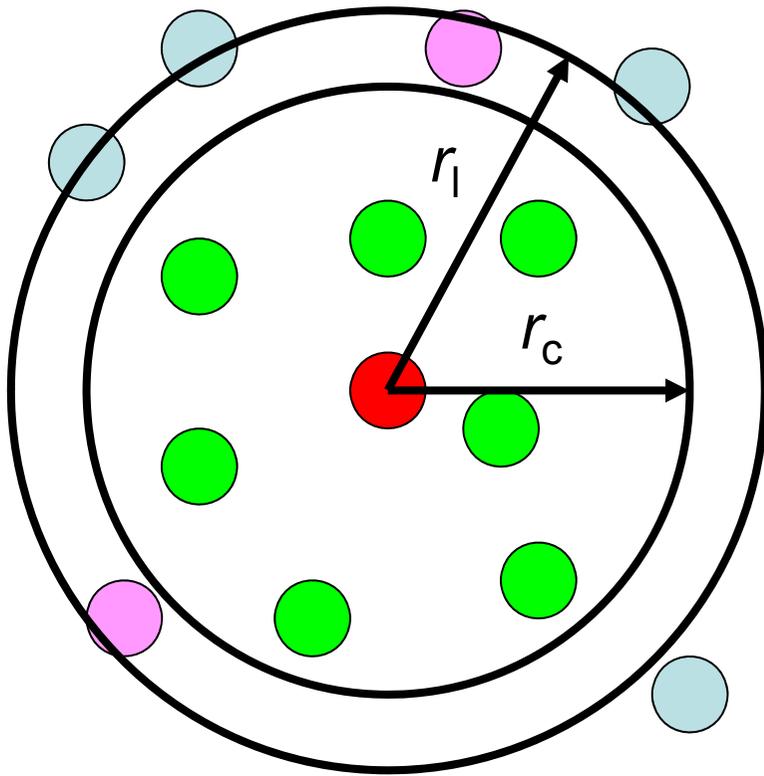
- 非共有結合相互作用は、原子のペアについて計算する必要がある
→ N 原子系では $N(N-1)/2$ のペア
- 非共有結合相互作用は距離が離れるほど弱くなる (van der Waals 引力は r^{-6} に比例、静電相互作用は r^{-1} に比例)
- 離れている原子同士は相互作用しないとみなす → カットオフ法

カットオフ法



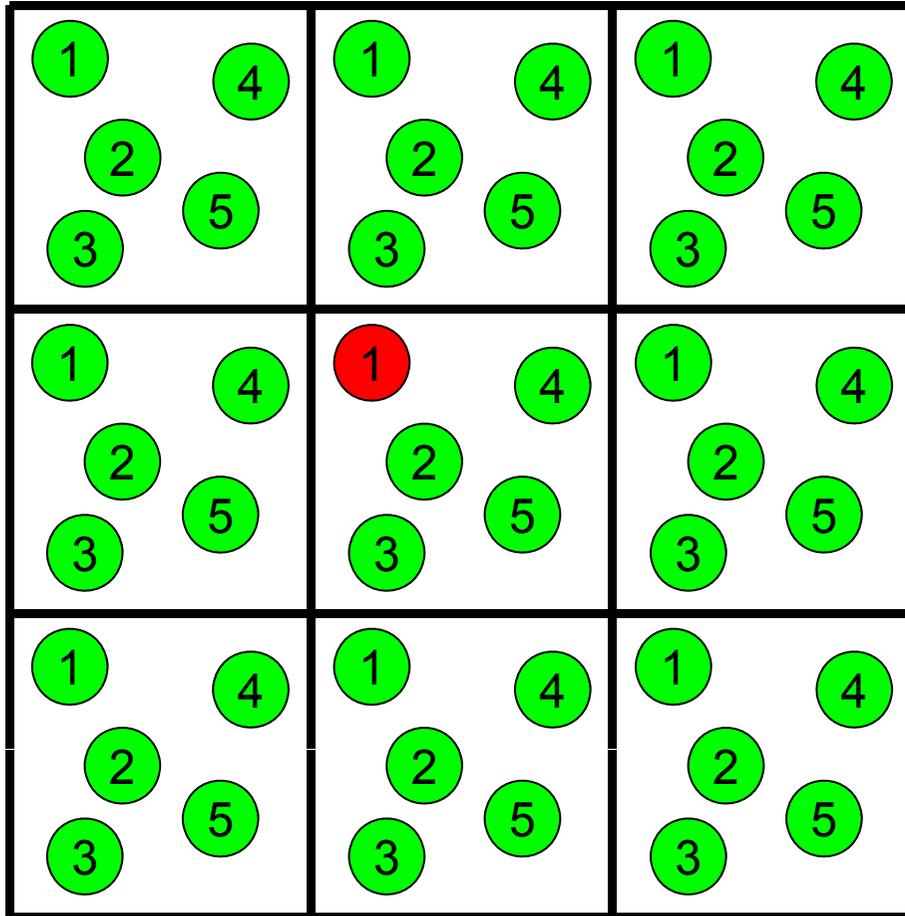
- 原子から半径 r_c の範囲内にある原子との非共有結合相互作用の計算を行う
- この範囲にある原子の平均個数を M とすると、非共有結合相互作用の計算量は $N(N-1)/2$ から NM に減少する

ペアリストの作成



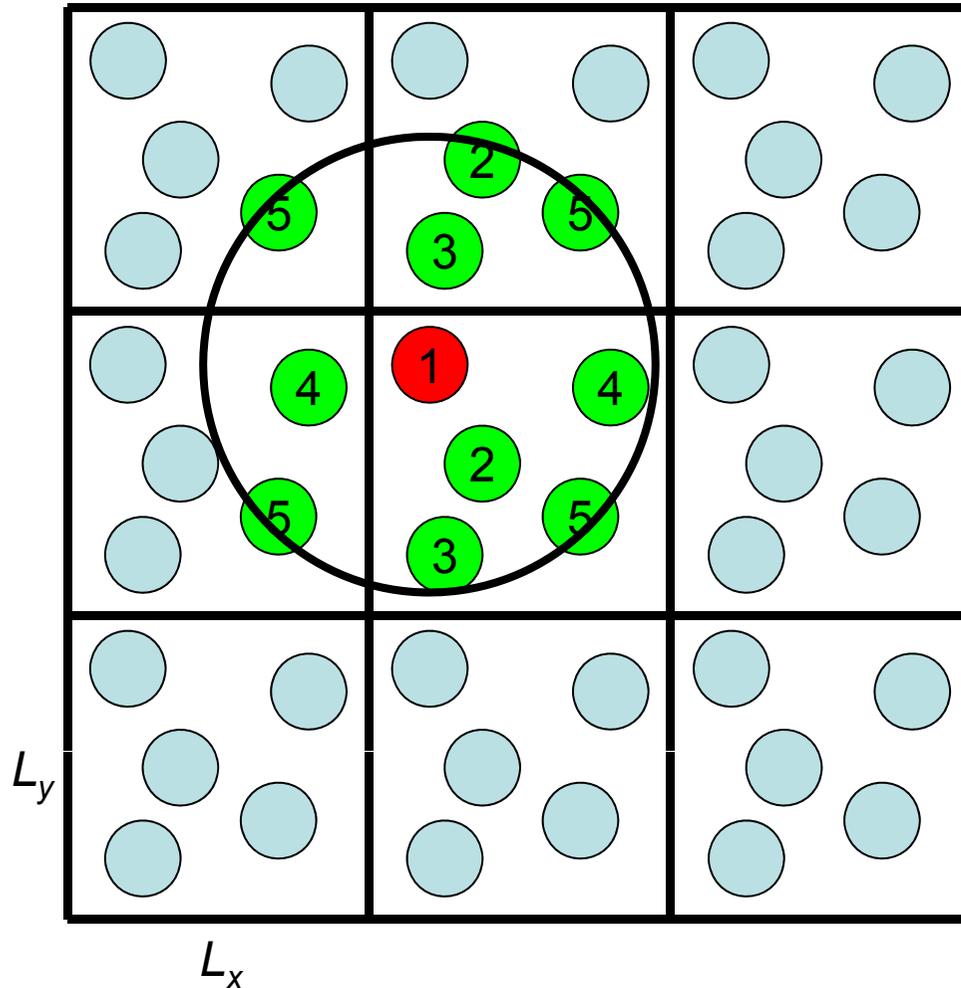
- カットオフ半径 r_c 以内にある原子ペアのリストを作成する必要がある
- この計算量は $N(N-1)/2$
- カットオフ半径 r_c より外側の半径 r_l の範囲でリストを作っておき、原子の最大移動度が $r_l - r_c$ を超えた時にリストを更新するようにすると計算量を削減できる

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
- カットオフを用いる必要がある

Minimum image convention

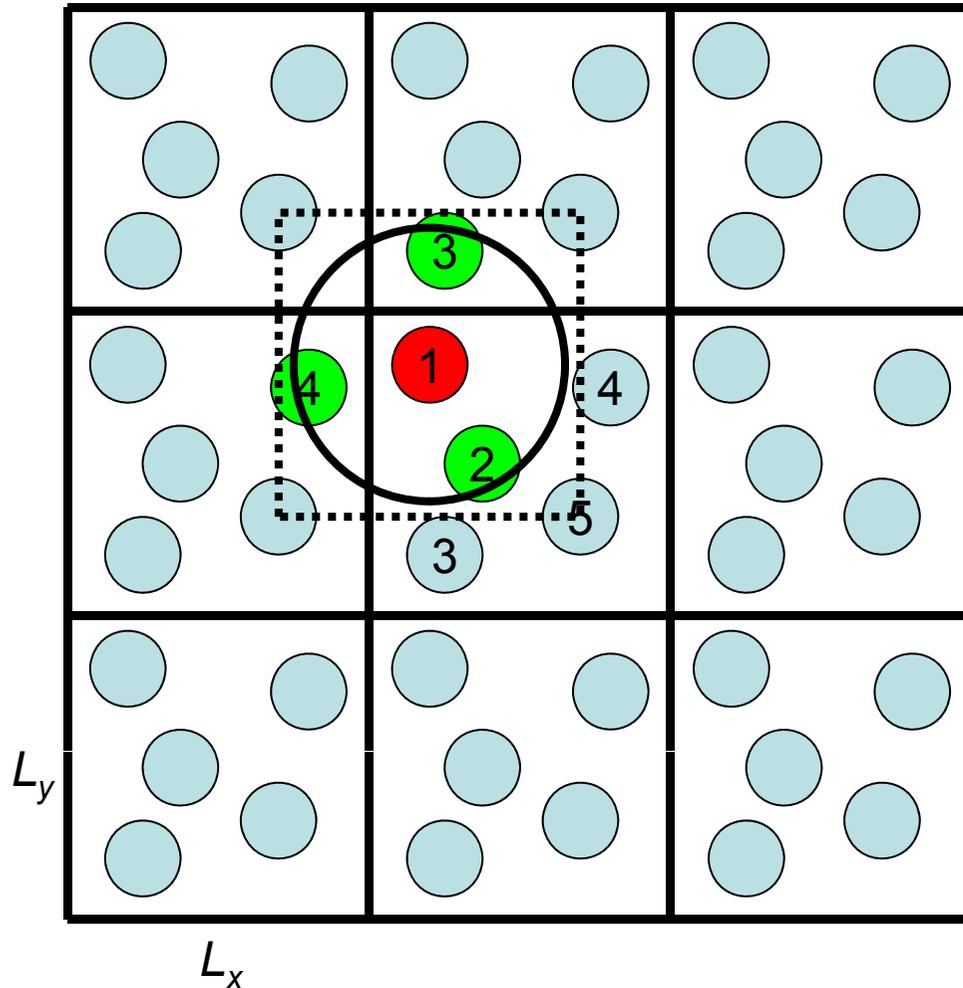


- カットオフ半径によっては、基本セルの周辺のイメージセルも考慮する必要がある
(左の例では $26N^2 + N(N-1)/2$ ペアの計算が必要)
- カットオフ半径 r_c を最も短い基本セルの1辺の長さの2分の1以下にすれば考慮すべきペア数は $N(N-1)/2$ でよい
→minimum image convention
- イメージセルの原子座標は基本セルの座標をセルの長さだけずらして計算する

$$y'_3 = y_3 + L_y$$

$$x'_4 = x_4 - L_x$$

Minimum image convention



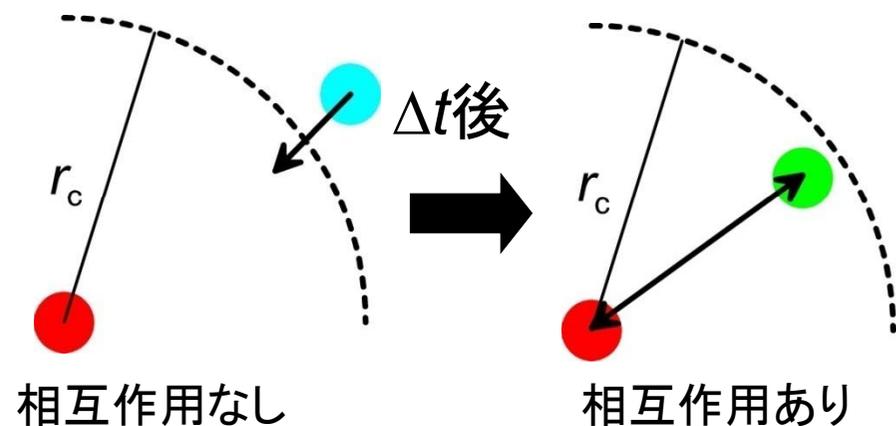
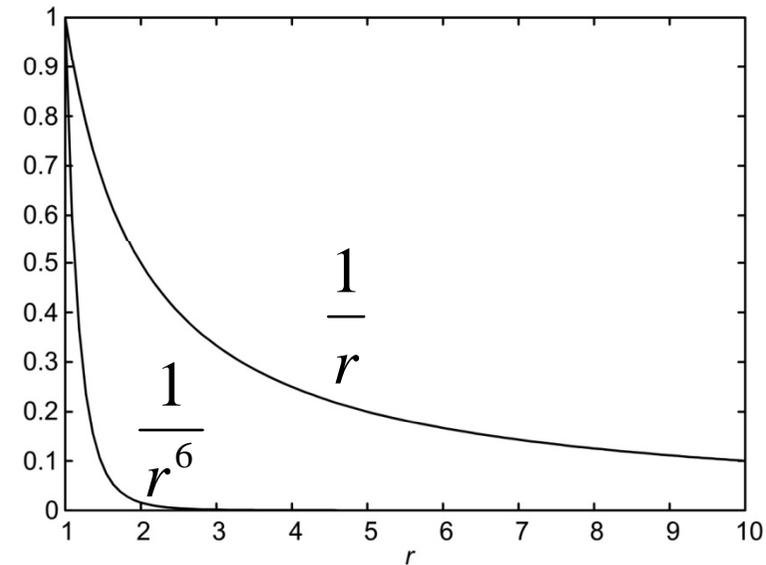
- カットオフ半径によっては、基本セルの周辺のイメージセルも考慮する必要がある
(左の例では $26N^2 + N(N-1)/2$ ペアの計算が必要)
- カットオフ半径 r_c を最も短い基本セルの1辺の長さの2分の1以下にすれば考慮すべきペア数は $N(N-1)/2$ でよい
→minimum image convention
- イメージセルの原子座標は基本セルの座標をセルの長さだけずらして計算する

$$y'_3 = y_3 + L_y$$

$$x'_4 = x_4 - L_x$$

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
→ 5 Åの距離にある原子に比べて、10 Åの距離にある原子は相互作用が $1/2^6 = 1/64$
→ 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
→ 5 Åの距離にある原子に比べて、10 Åの距離にある原子は相互作用が $1/2$
→ 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

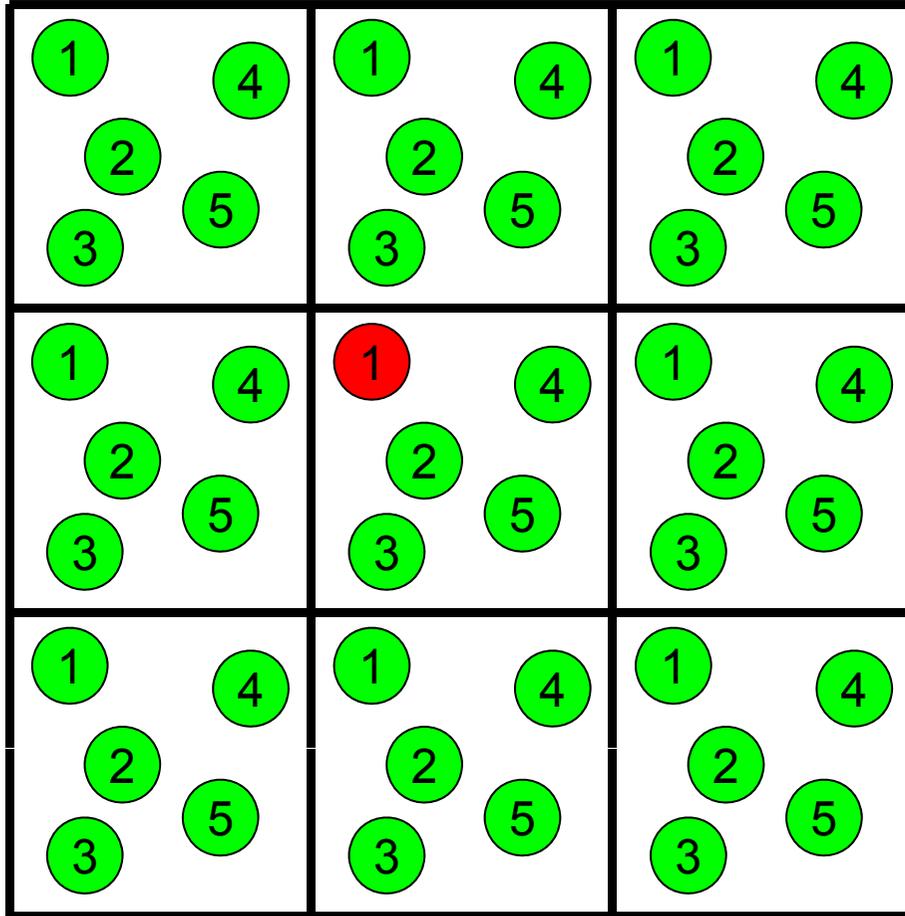


カットオフしない計算法

- 遠距離まで相互作用が及ぶ静電相互作用は、カットオフせずに計算する必要がある
- 遠距離の相互作用を多重極子で近似する
 - Cell multipole法
 - Fast multipole法

→ 主に球状に水分子を配置する場合
- 周期境界条件において、イメージセルからの寄与を近似的に取り込む
 - Ewald法
 - Particle Mesh Ewald法

周期境界条件の場合(2)



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したい

原子*i*の位置 \mathbf{r}_i における静電ポテンシャル:

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

$\mathbf{n} = 0$ 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)

- 静電ポテンシャルもそれぞれの電荷分布由来の和で書ける

$$\varphi(\mathbf{r}) = \varphi_1(\mathbf{r}) + \varphi_2(\mathbf{r})$$

φ_1 : 残りの電荷分布由来
 φ_2 : ガウス分布に従う電荷分布由来

- 静電相互作用エネルギーは以下の通り

$$E = E_{\text{dir}} + E_{\text{rec}} + E_{\text{self}}$$

$$E_{\text{dir}} = \frac{1}{2} \sum_i q_i \varphi_1(\mathbf{r}_i), \quad E_{\text{rec}} = \frac{1}{2} \sum_i q_i \varphi_2(\mathbf{r}_i)$$

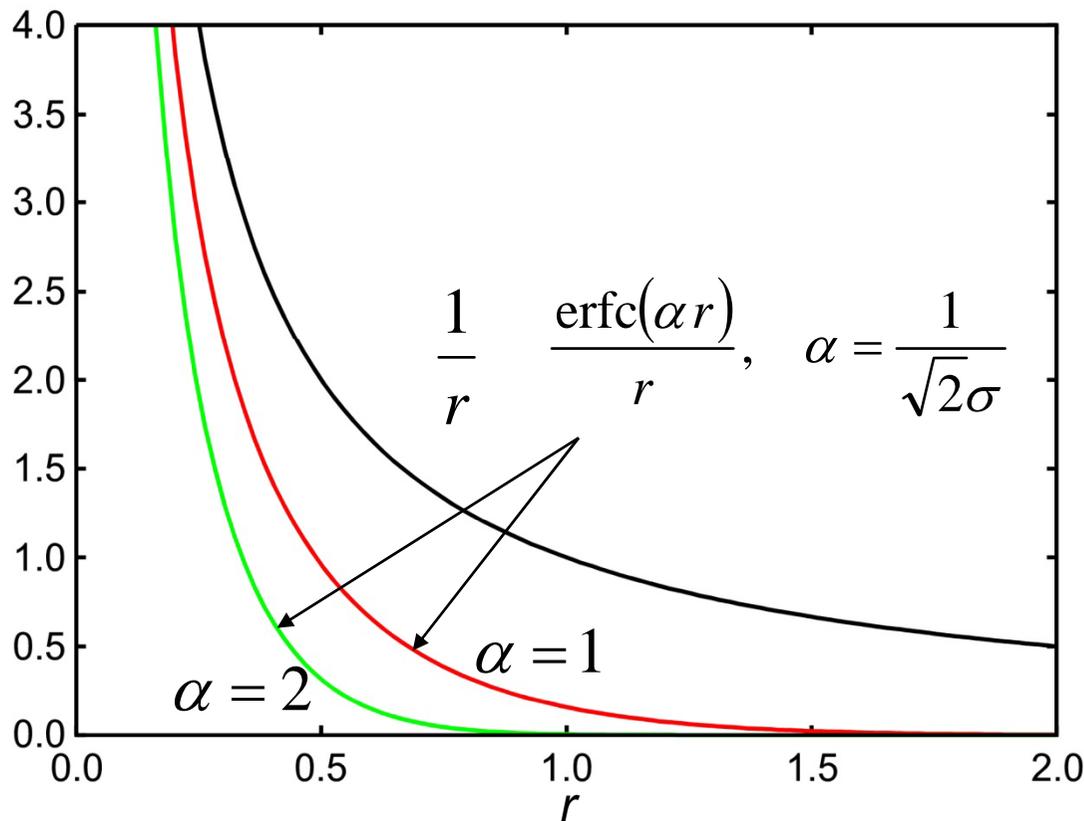
$$E_{\text{self}} = -\frac{1}{\sqrt{2\pi\sigma^2}} \sum_i q_i^2$$

同じ原子によるガウス分布に従う電荷分布同士の相互作用を差し引く

Particle Mesh Ewald法 (3)

$i = j$ のとき $n = 0$ を除く

$$E_{\text{dir}} = \sum_{j \geq i} \sum_{\mathbf{n}} \frac{q_i q_j \operatorname{erfc}\left(\frac{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|}{\sqrt{2}\sigma}\right)}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|} \approx \sum_{\substack{j > i \\ \text{MIC}}} \frac{q_i q_j \operatorname{erfc}\left(\frac{|\mathbf{r}_i - \mathbf{r}_j|}{\sqrt{2}\sigma}\right)}{|\mathbf{r}_i - \mathbf{r}_j|}$$



erfc(r)/ r は $1/r$ と比べて速く減衰する
 →パラメータ σ を調整し、 $L/2$ 以内で十分減衰するようにする
 →minimum image conventionで相互作用を計算すればよい

Particle Mesh Ewald法 (4)

- Poisson方程式 ($\varepsilon = 1$)

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

- 周期境界条件下では、電荷分布が十分に滑らかであれば、離散フーリエ変換を用いて解くことができる

$$\tilde{\varphi}(\mathbf{k}) = \frac{4\pi}{|\mathbf{k}|^2} \tilde{\rho}(\mathbf{k}), \quad \mathbf{k} = \frac{2\pi\mathbf{h}}{L}$$

$\mathbf{k}=0$ で発散を防ぐため、 $\tilde{\rho}(0)$ 、すなわち全電荷は0にする

Particle Mesh Ewald法 (5)

- エネルギー E_{rec} は以下で計算できる

$$E_{\text{rec}} = \frac{1}{2} \sum_i q_i \varphi_2(\mathbf{r}_i) = \frac{h^3}{2} \sum_m \rho'(\mathbf{r}_m) \varphi'(\mathbf{r}_m)$$

h はグリッドの幅で、 \mathbf{r}_m はグリッドの点である
また、 ρ' は以下を満たし、 φ' は ρ' による静電ポテンシャルである

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

Particle Mesh Ewald法(6)

- 以上をまとめると以下のようなになる
 - ガウス分布に従う電荷分布による静電ポテンシャルは、離散Fourier変換を用いて計算する
 - 全電荷は0でなければならない
 - 残りの電荷分布による静電相互作用はパラメータ σ を調整し、この項が小さいカットオフ半径で十分減衰するようにし、minimum image conventionに従って計算する
- 離散Fourier変換に高速フーリエ変換を用いることで、原子数 N の系で $N\log N$ に比例する計算量に削減できる

参考: Fourier変換

周期 L の関数のFourier変換

$$f(x) = \sum_{n=-\infty}^{\infty} \tilde{f}(n) \exp\left(\frac{2\pi i n x}{L}\right)$$

Fourier変換

$$\tilde{f}(n) = \frac{1}{L} \int_0^L f(x) \exp\left(-\frac{2\pi i n x}{L}\right) dx$$

Fourier逆変換

$$f \otimes g(x) = \frac{1}{L} \int_0^L f(x-y)g(y)dy$$

畳み込み

$$\frac{1}{L} \int_0^L f \otimes g(x) \exp\left(-\frac{2\pi i n x}{L}\right) dx = \tilde{f}(n)\tilde{g}(n)$$

畳み込みの
Fourier変換

参考：式の証明

- 電子密度 ρ の作る静電ポテンシャル ϕ は以下の
ように書くことができる

$$\phi(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' = \rho \otimes \gamma, \quad \gamma(\mathbf{r}) = \frac{1}{|\mathbf{r}|}$$

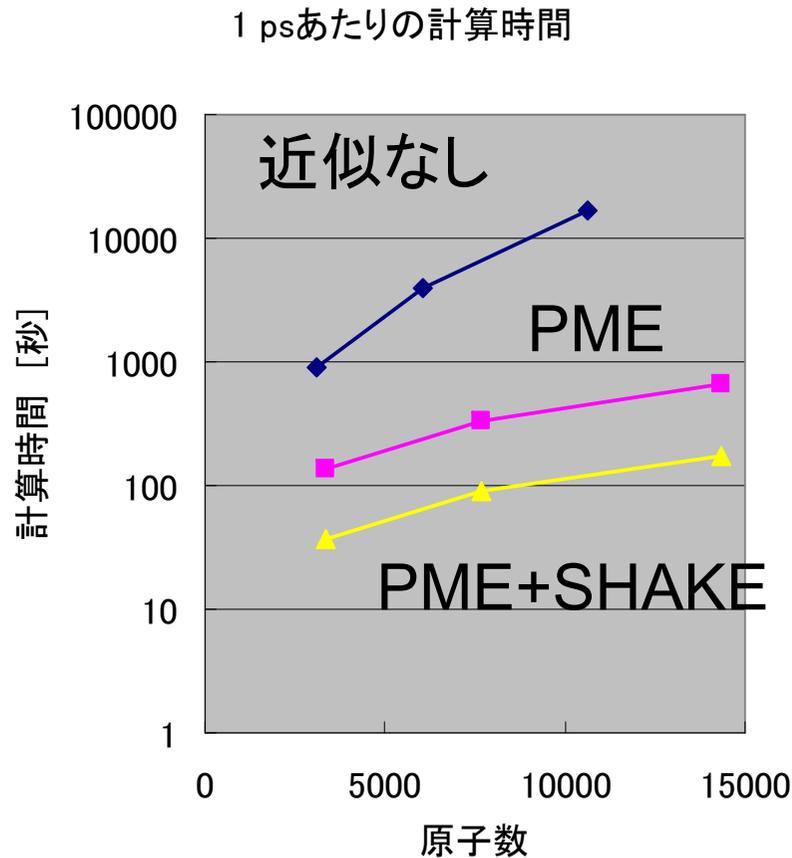
- これを用いると E_{rec} は以下のように書ける

$$\begin{aligned} E_{\text{rec}} &= \frac{1}{2} \rho \cdot \phi_2 = \frac{1}{2} \rho \cdot (\rho \otimes G^\sigma \otimes \gamma) \\ &= \frac{1}{2} \rho \cdot (\rho \otimes G^{\sigma/\sqrt{2}} \otimes G^{\sigma/\sqrt{2}} \otimes \gamma) \\ &= \frac{1}{2} (\rho \otimes G^{\sigma/\sqrt{2}}) \cdot (\rho \otimes G^{\sigma/\sqrt{2}} \otimes \gamma) \end{aligned}$$

分散 σ^2 のガウス分布に従う電荷分布

$A \cdot (B \otimes C) = (A \otimes B) \cdot C$

実際の計算時間



- 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3.7倍程度高速化した

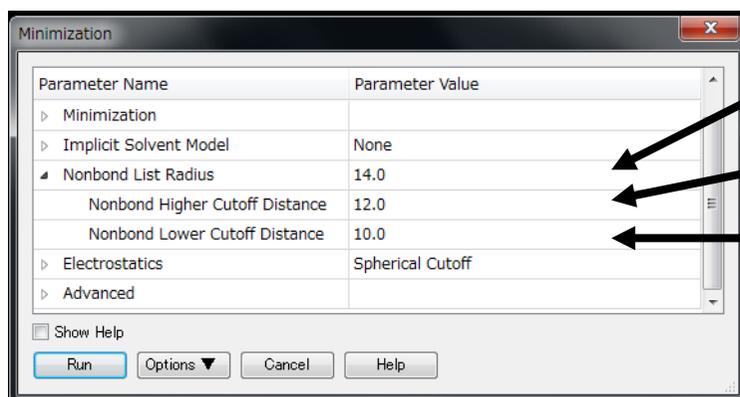
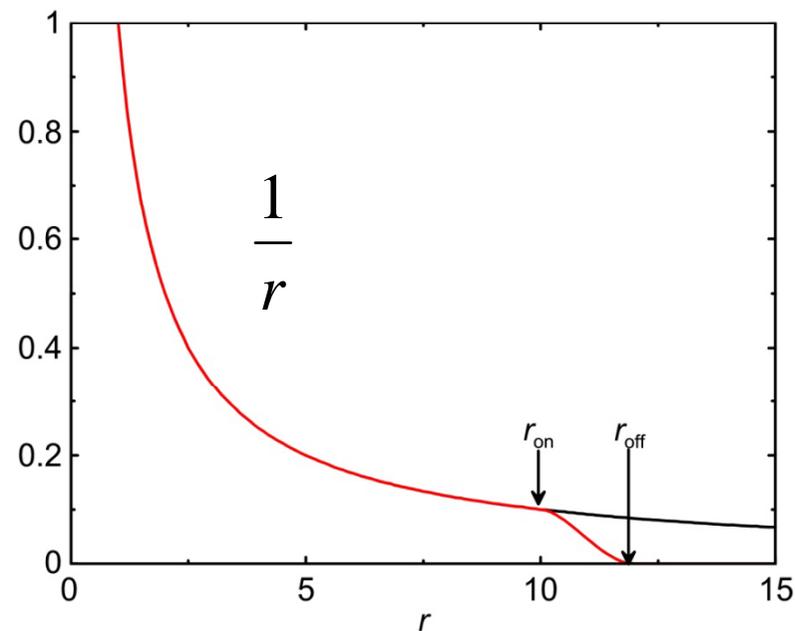
計算の並列化



- 分子動力学シミュレーションで用いるソフトウェアは、並列計算を前提としてプログラムされている
- PCを高速ネットワークで接続したPCクラスタを用いると、安価なシステムで比較的高速なシミュレーションが可能

Discovery Studioでの操作(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



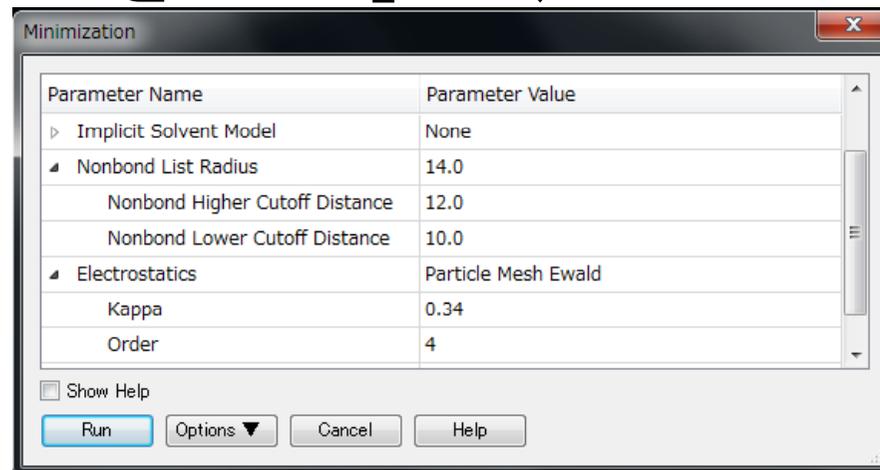
r_l (ペアリスト作成用)

r_{off}

r_{on}

Discovery Studioでの操作(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



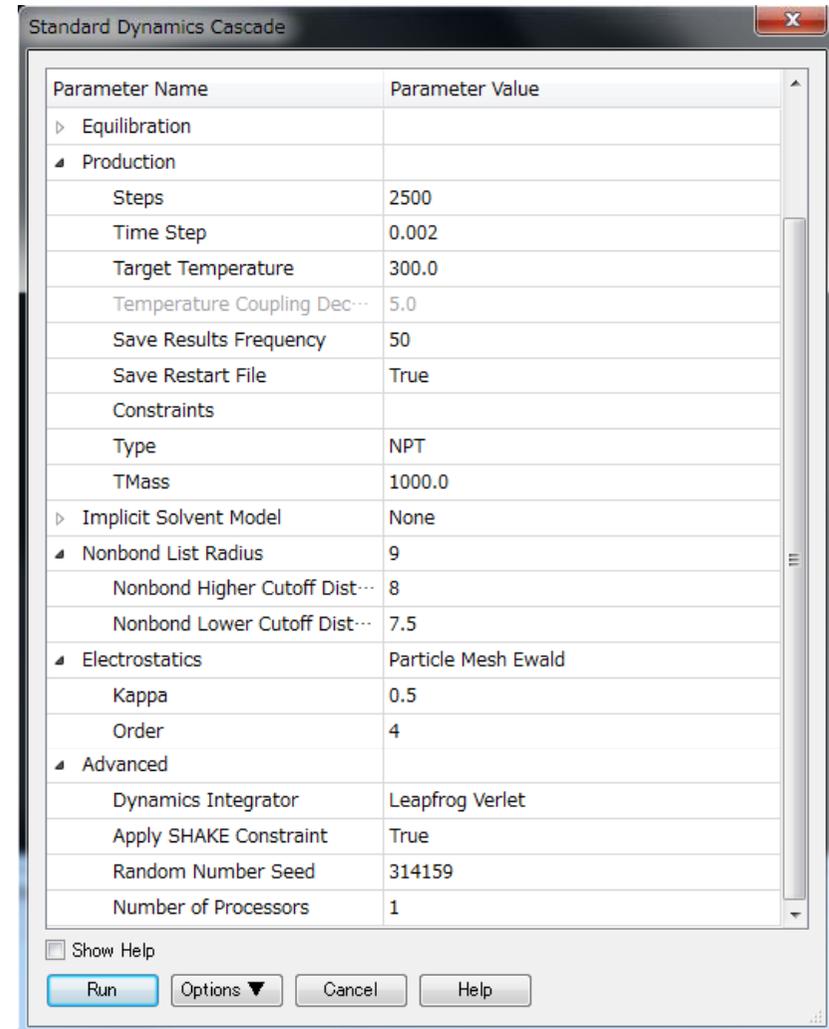
タンパク質のMDシミュレーション(1)

1. PDB ID 1CRNの構造を開く
2. Force fieldにcharmm27を指定(水素が付加される)
3. Implicit Solvent ModelにGBSWを指定してエネルギー最小化を行う
4. 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
5. 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)

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

- 30分ほどかかる
- ステップ数: 2500
- 時間刻み: 2 fs
- 定温・定圧
- Particle Mesh Ewald
- SHAKE



課題2

- 1CRNの結晶構造を観察し、水素結合を2つ以上見つけよ
- 1CRNのMDシミュレーション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

応用事例の紹介

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 44, pp. 30230–30239, October 30, 2009
© 2009 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

A Non-canonical DNA Structure Enables Homologous Recombination in Various Genetic Systems^{*[S]}

Received for publication, July 13, 2009, and in revised form, August 28, 2009. Published, JBC Papers in Press, September 3, 2009, DOI 10.1074/jbc.M109.043810

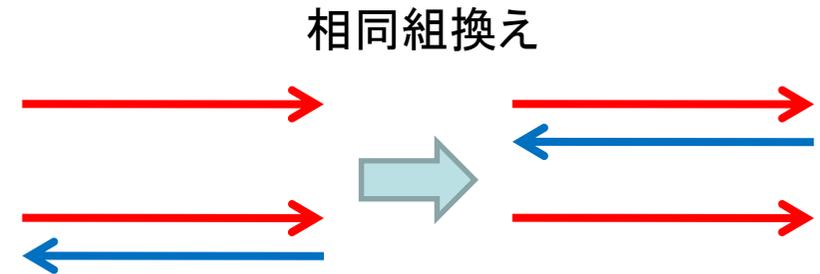
Tokiha Masuda^{‡S}, Yutaka Ito[¶], Tohru Terada^{||**}, Takehiko Shibata^{‡S}, and Tsutomu Mikawa^{‡S¶¶1}

From the [‡]Graduate School of Nanobioscience, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, the ^SRIKEN Advanced Science Institute and the ^{**}RIKEN Computational Science Research Program, 2-1 Hirosawa, Wako, Saitama 351-0198, the [¶]Department of Chemistry, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, the ^{||}Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, and the ^{¶¶}RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

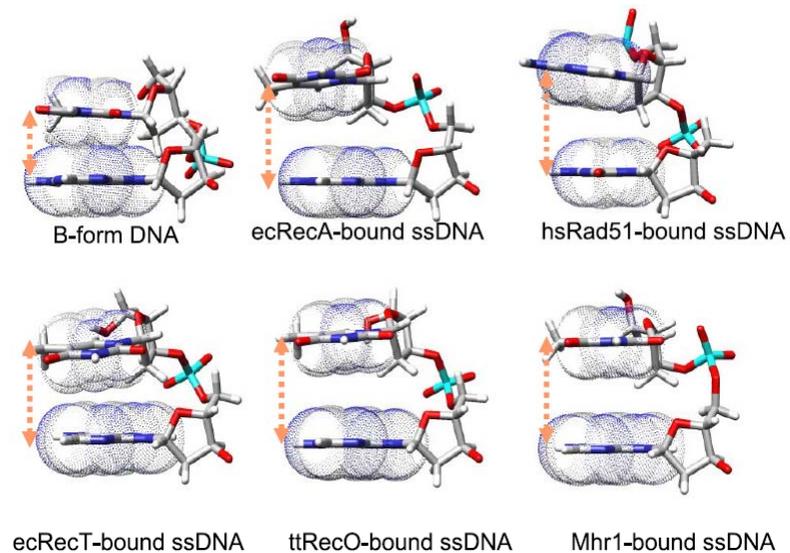
Masuda *et al.* *J. Biol. Chem.* **284**, 30230 (2009).

RecAによるDNAの相同組換え

- RecAは大腸菌で、DNAの相同組換えを担うタンパク質
- 1本鎖DNAはRecAに結合すると、伸張した特異な立体構造をとる
- 本論文では、他の生物に存在する、別の相同組換えタンパク質に結合したDNAも同様な立体構造をとることを示した
- タンパク質同士に進化的な類縁関係はないことから、この構造をとることが、相同組換えに重要であると考えられる



組換えタンパク質に結合したDNAの構造

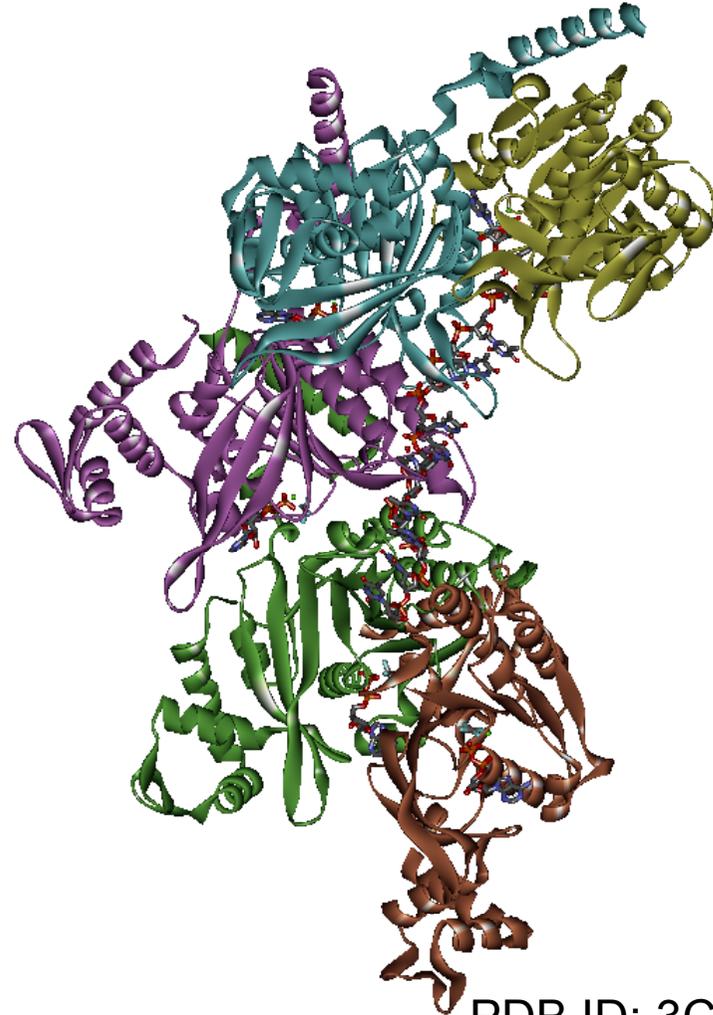


RecA・DNA複合体構造

- では、この構造をとるとなぜ相同組換えに有利なのか？



- RecA・DNA複合体の結晶構造を基に、分子動力学シミュレーションを用いてDNAの運動性を解析した



PDB ID: 3CMW

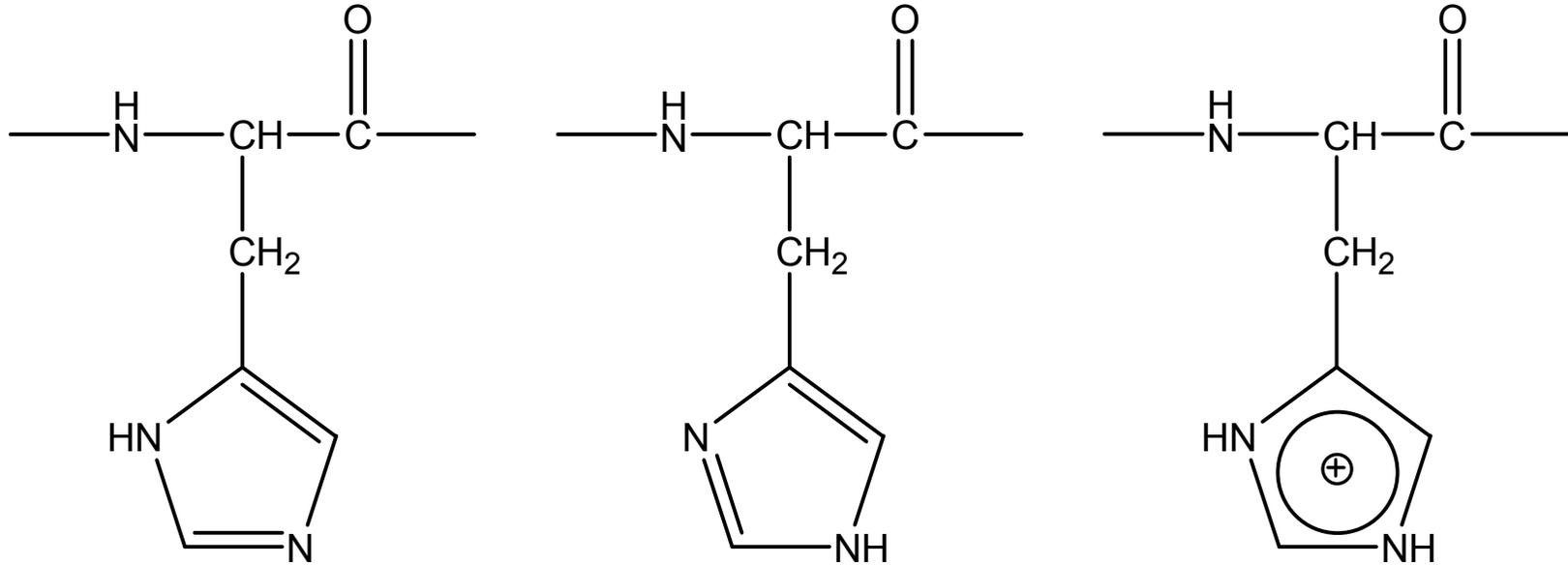
シミュレーションの手順

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
 - 座標の束縛
 - 水分子配置の最適化
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のサイト (<http://www.rcsb.org/pdb/>) からダウンロード
 - 非対称単位に2つの複合体が含まれているが、一方だけで機能していることは明らか→一方の複合体のみを選択
- 欠失残基への対応
 - 結晶構造に含まれる欠失残基はモデリングなどで補う必要がある
 - ここでは、欠失残基は人為的に付加されたリンカ配列であるので、欠失残基の前後の残基をacetyl基、N-methyl基でブロック
- 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

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
- 基本的には、His周りの水素結合ネットワークからプロトン化状態を決定する

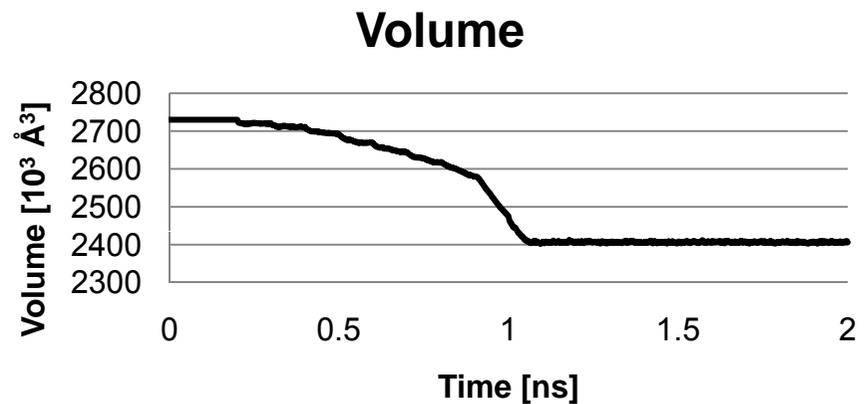
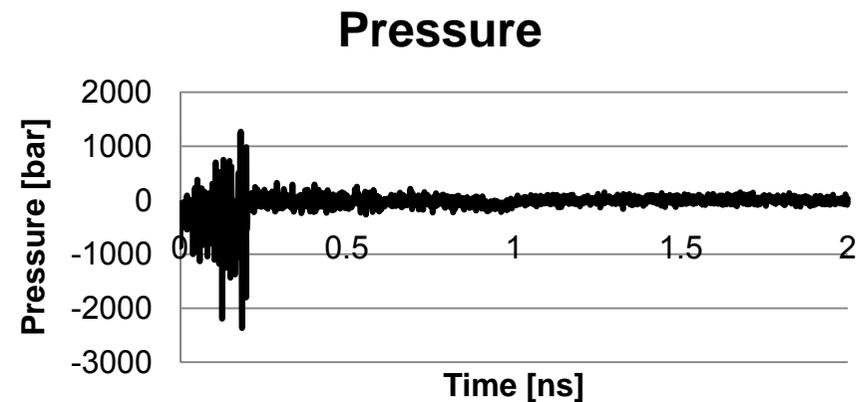
初期構造の作成(2)

- リガンドのモデリング
 - RecAにはATPが結合するが、ここでは反応中間体アナログADP·AlF₃が結合している
 - アナログを本来のATPに戻すモデリングを行う
- 力場パラメータの取得
 - リガンドの力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する
 - 電荷を中性にするためにカウンターイオンを配置する

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

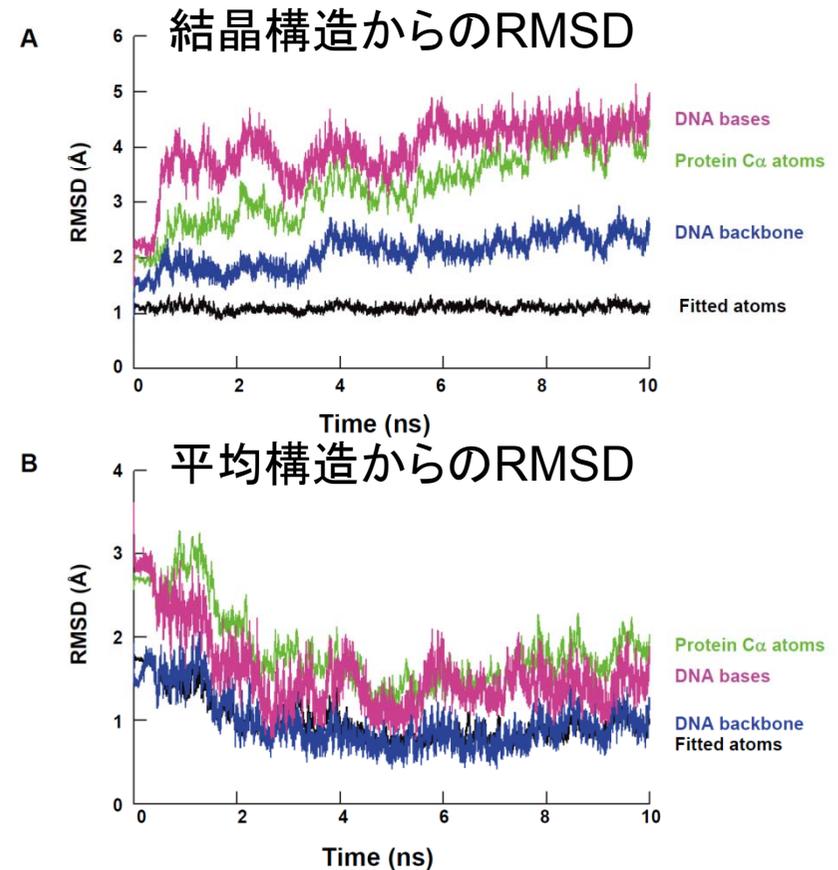
平衡化

- 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する



結果の解析

- DNAのbackboneに対してbaseのRMSDは大きく、baseの構造が大きく揺らいでいることが明らかとなった
- これは、RecAに結合したDNAが伸長した構造をとることにより、base間の相互作用が弱くなることによる
- この運動性が相同性探索に有利に働く可能性がある



課題の提出

- 課題1、課題2の結果のエクセルファイルを添付して、寺田宛tterada@iu.a.u-tokyo.ac.jpに送ること
- 課題1、課題2の考察はメールの本文に記載すること
- 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