

# 計算生命科学の基礎II



## 分子動力学計算によるタンパク質の機能解析

---

京都大学大学院医学研究科  
臨床システム腫瘍学  
中津井雅彦



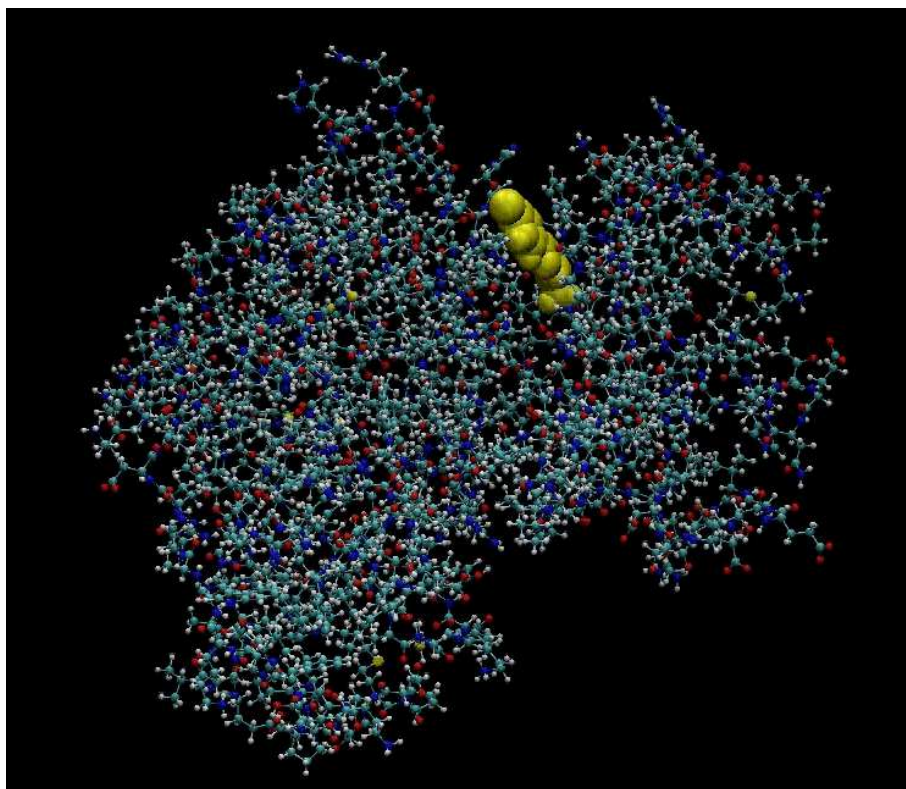


- (古典)分子動力学計算とは
  - (古典)分子動力学計算の概要
  - 運動方程式
  - 数値解法
  - 力場
  - 境界条件
- 統計アンサンブル
  - 運動方程式の一般化
  - NVTアンサンブル(カノニカルアンサンブル)
  - NPTアンサンブル(T-Pアンサンブル)
- 実例
  - タンパク質・リガンド間の結合自由エネルギー予測 (MP-CAFEE法)
- 拡張アンサンブル
  - 必要性
  - マルチカノニカル法
  - レプリカ交換法
  - 実例

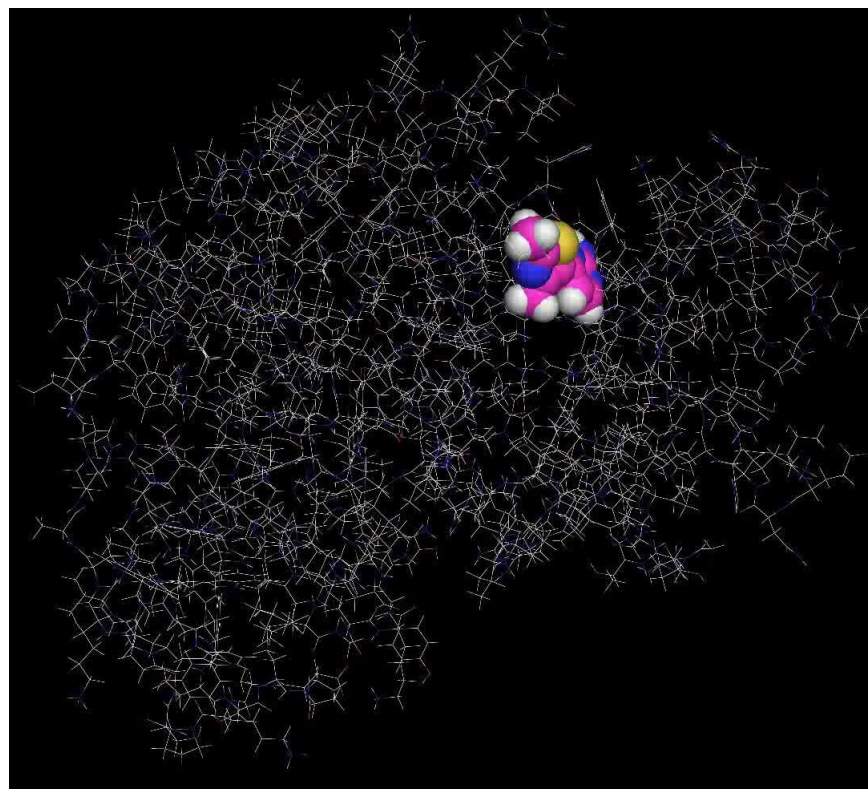




## 分子動力学計算



## 従来型の結合シミュレーション(Docking)



タンパク質は、動的な系である



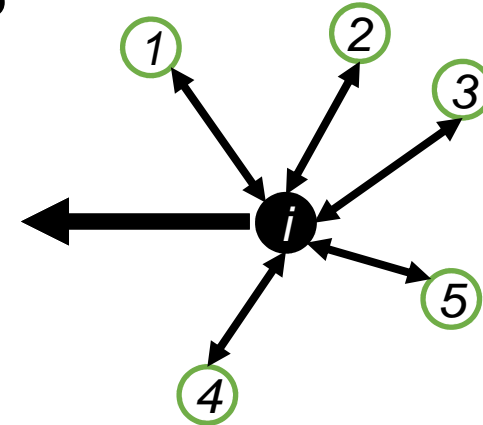
# (古典)分子動力学計算の概要

- 原子を「質量を持つ粒子」とみなし、古典力学の運動方程式を解くことで位置を求める

例) ニュートンの第二法則

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} = m_i a_i$$

$F_i$  質点*i*が受ける力  
 $m_i$  質点*i*の質量  
 $r_i$  質点*i*の位置  
 $a_i$  質点*i*の加速度



古典力学に基づいているため、  
量子的な効果は計算しない

加速度・力・質量の時刻に対する微分方程式  
初期座標・初期速度・力・質量がわかれば、  
任意の時刻の座標を計算できる

ただし、**解析的に解けない**ので、**数值的に解く**必要がある

粒子を動かすための力は、  
経験的に決められた**ポテンシャルの空間微分**

“The science of simulating the motions of a system of particles” (Karplus & Petsko)

# (古典)分子動力学計算の概要



- 決定論的な手法
  - 将来の時間における系の状態が、現在の状態から予測できる
- 分子動力学計算のサイクル
  1. 一定の短い時間刻みの間に、各々の原子にかかる力を定数として予測する
  2. 各々の原子の現在の座標および速度と、1. で計算した各々の粒子へかかる力から、一定の短い時間刻み後の原子の座標および速度を計算する

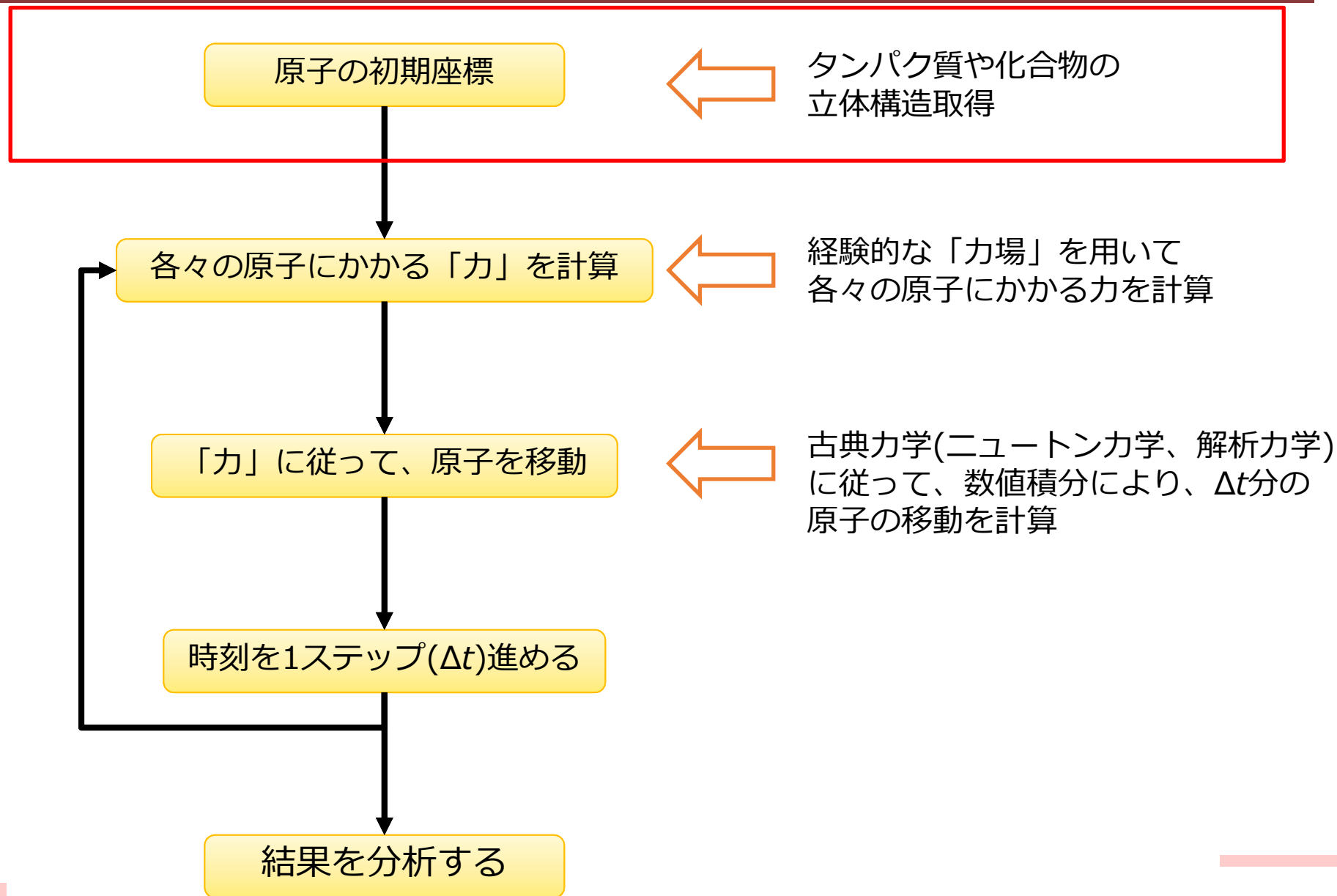
短い時間刻みごとに、原子の座標のスナップショットを得る



トラジェクトリ



# (古典)分子動力学計算の流れ



# (古典)分子動力学計算に必要な情報



---

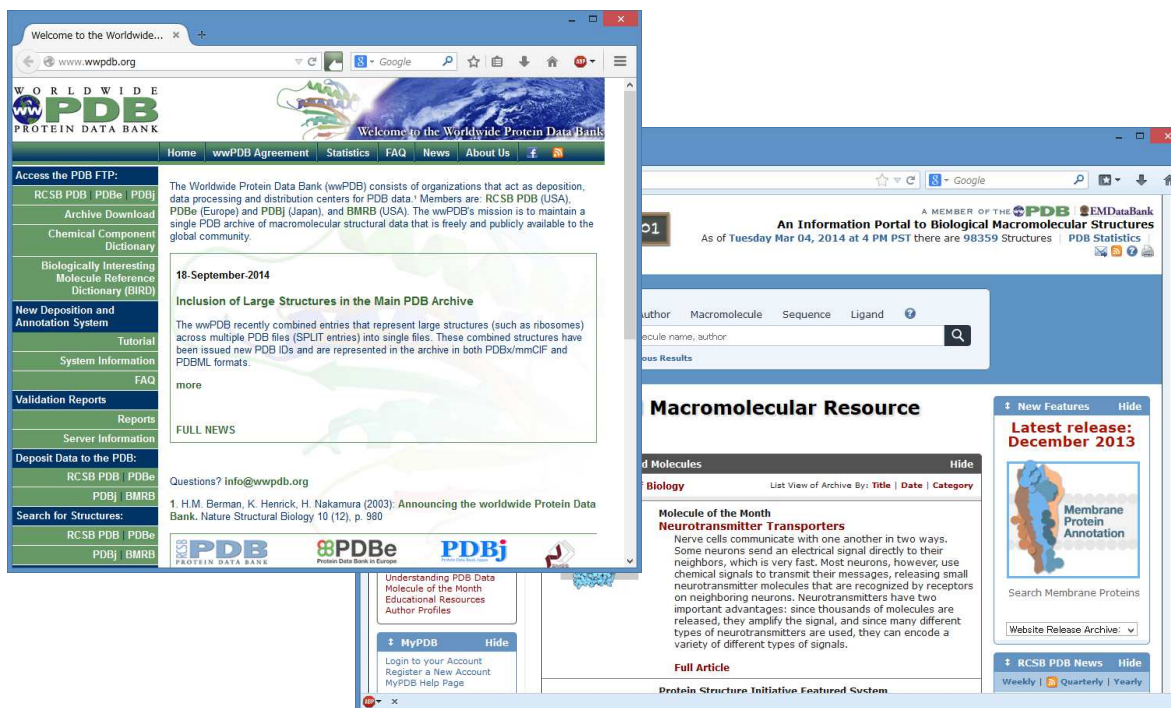
- 原子の三次元座標
  - X線結晶解析, NMR等
  - 無料の公的データベース
    - PDB
  
- 原子の電荷情報 (原子にかかる力を計算する際に使用)
  - タンパク質(アミノ酸や主要な金属原子等)
  - それ以外の化合物
    - 量子計算により求める



# タンパク質の立体構造を取得する



国際タンパク質データバンク (wwPDB) <http://www.wwpdb.org/>



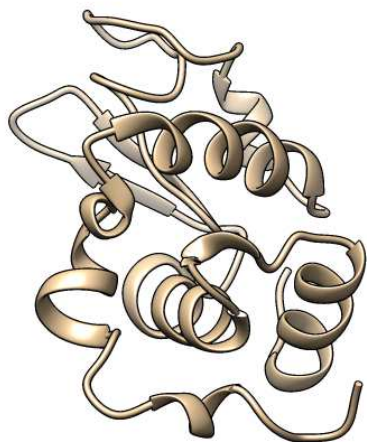
X線決勝解析、NMR等の手法によって原子レベルの分解能で解析された  
タンパク質・核酸・糖などの生体高分子の立体構造

PDBj, RCSB PDB, EBI PDBe





# タンパク質の立体構造を取得する PDBフォーマット



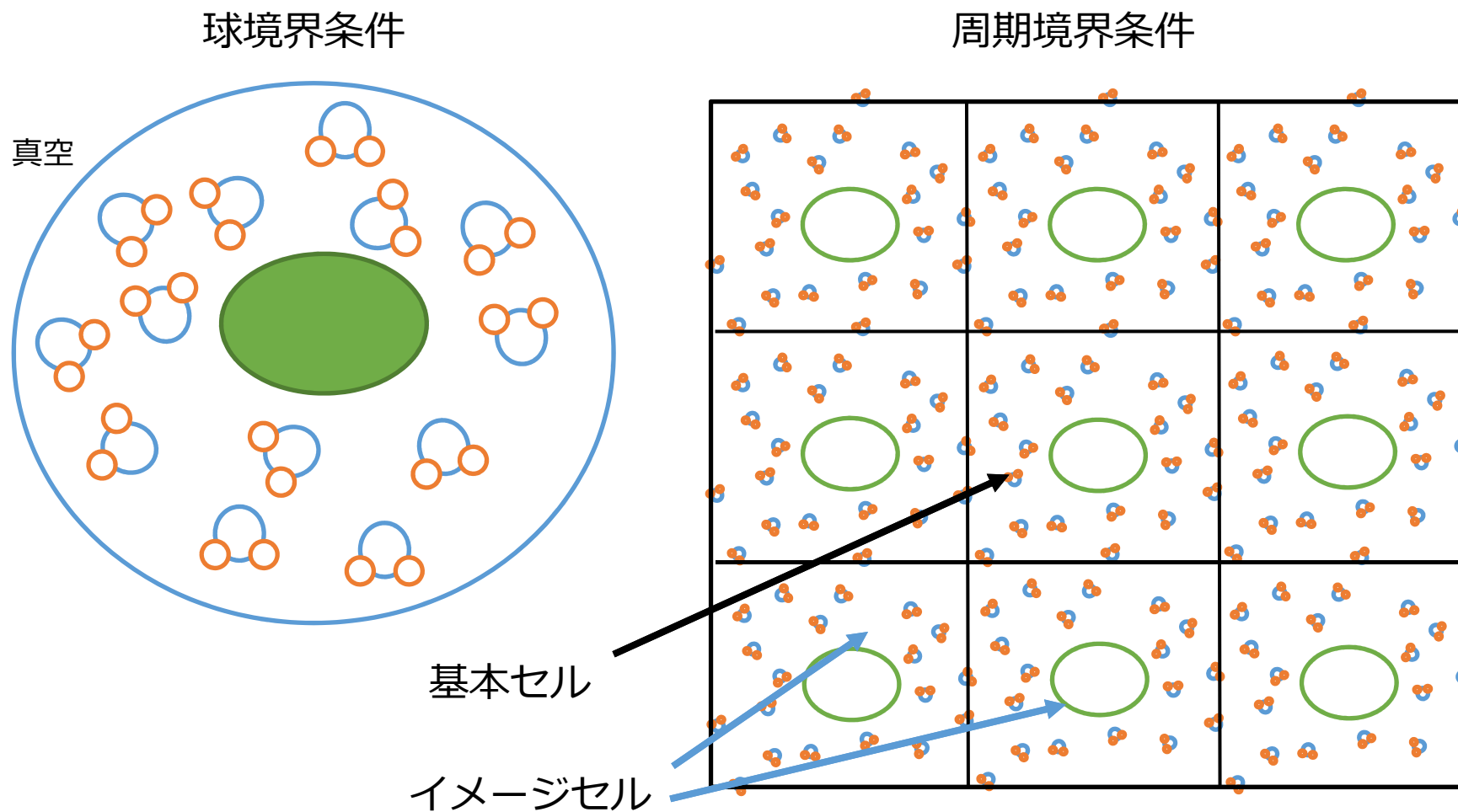
Lysozyme  
(1AKI.pdb)

ID	Atom	Res	Chain	Alt	X (Å)	Y (Å)	Z (Å)	Occupancy	B-factor	Element		
347	ATOM	1	N	LYS	A	1	35.365	22.342	-11.980	1.00	22.28	N
348	ATOM	2	CA	LYS	A	1	35.892	21.073	-11.427	1.00	21.12	C
349	ATOM	3	C	LYS	A	1	34.741	20.264	-10.844	1.00	16.85	C
350	ATOM	4	O	LYS	A	1	33.945	20.813	-10.081	1.00	18.94	O
351	ATOM	5	CB	LYS	A	1	36.872	21.435	-10.306	1.00	20.78	C
352	ATOM	6	CG	LYS	A	1	37.453	20.248	-9.565	1.00	18.47	C
353	ATOM	7	CD	LYS	A	1	38.688	20.649	-8.775	1.00	20.32	C
354	ATOM	8	CE	LYS	A	1	39.057	19.508	-7.837	1.00	24.76	C
355	ATOM	9	NZ	LYS	A	1	40.423	19.771	-7.299	1.00	28.27	N
356	ATOM	10	N	VAL	A	2	34.739	18.961	-11.042	1.00	19.96	N
357	ATOM	11	CA	VAL	A	2	33.903	17.998	-10.333	1.00	18.10	C
358	ATOM	12	C	VAL	A	2	34.800	17.312	-9.294	1.00	19.39	C
359	ATOM	13	O	VAL	A	2	35.759	16.605	-9.665	1.00	22.14	O
360	ATOM	14	CB	VAL	A	2	33.140	17.034	-11.232	1.00	16.81	C
361	ATOM	15	CG1	VAL	A	2	32.251	16.084	-10.434	1.00	21.92	C
362	ATOM	16	CG2	VAL	A	2	32.294	17.714	-12.290	1.00	19.46	C
363	ATOM	17	N	PHE	A	3	34.491	17.546	-8.038	1.00	19.89	N
364	ATOM	18	CA	PHE	A	3	35.185	16.903	-6.918	1.00	17.43	C
365	ATOM	19	C	PHE	A	3	34.742	15.441	-6.771	1.00	15.70	C
366	ATOM	20	O	PHE	A	3	33.525	15.162	-6.862	1.00	18.52	O
367	ATOM	21	CB	PHE	A	3	34.967	17.632	-5.594	1.00	17.94	C
368	ATOM	22	CG	PHE	A	3	35.944	18.737	-5.375	1.00	16.78	C
369	ATOM	23	CD1	PHE	A	3	35.666	20.050	-5.798	1.00	15.97	C
370	ATOM	24	CD2	PHE	A	3	37.000	18.557	-4.473	1.00	19.95	C
371	ATOM	25	CE1	PHE	A	3	36.577	21.076	-5.568	1.00	17.32	C
372	ATOM	26	CE2	PHE	A	3	37.869	19.589	-4.157	1.00	17.65	C
373	ATOM	27	CZ	PHE	A	3	37.636	20.873	-4.666	1.00	17.91	C
374	ATOM	28	N	GLY	A	4	35.724	14.639	-6.331	1.00	16.79	N
375	ATOM	29	CA	GLY	A	4	35.366	13.280	-5.870	1.00	16.34	C
376	ATOM	30	C	GLY	A	4	34.924	13.420	-4.415	1.00	11.91	C

原子ごとに、**三次元座標**が記述されている

X, Y, Z座標 (Å単位)

# 原子を配置する



球境界条件では、球境界付近での水の挙動が不自然になることがある



# 初期速度



$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} = m_i \mathbf{a}_i$$

数値積分には、初期座標・初期速度の両方が必要

初期速度の与え方

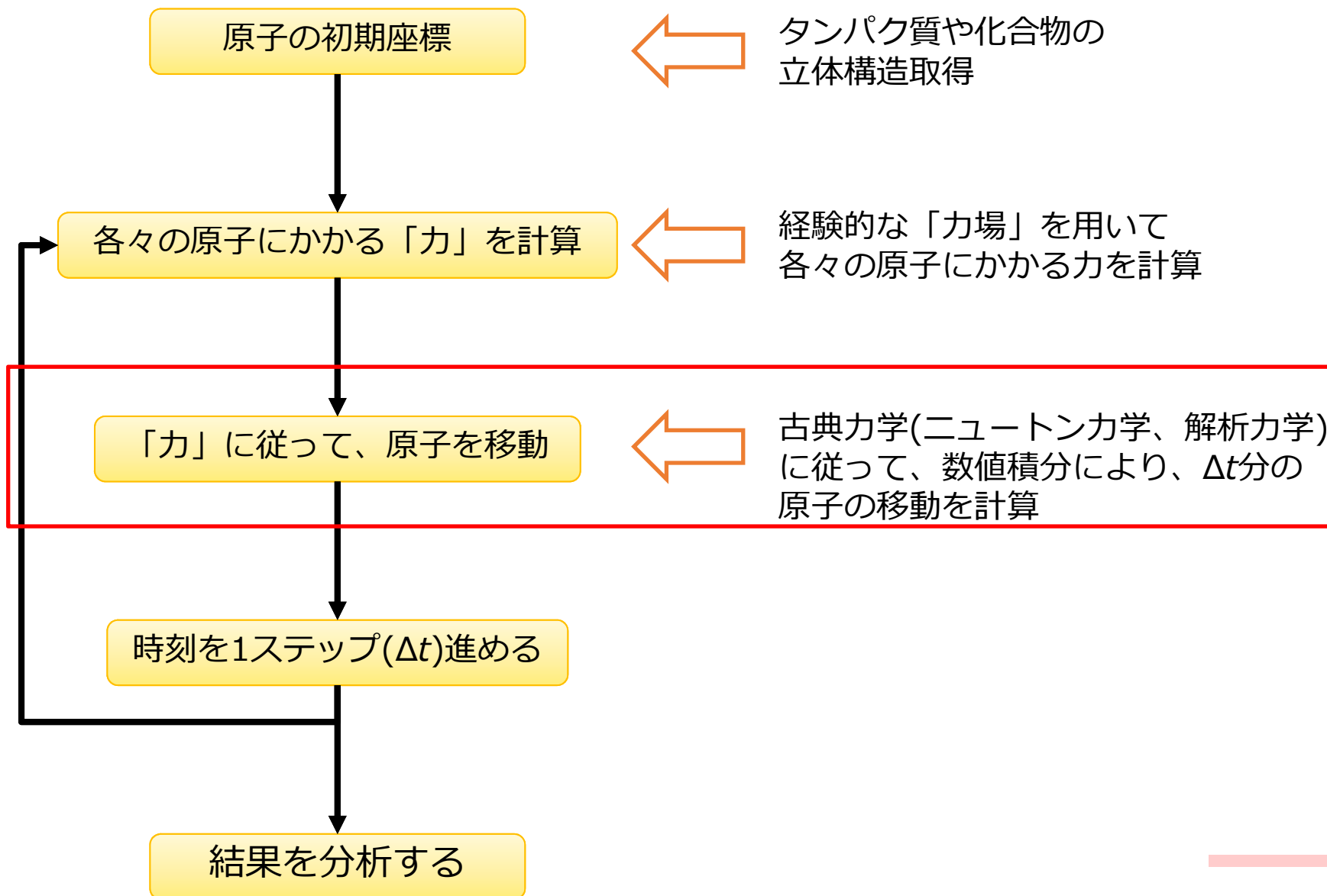
- 初期速度を与えずにMDを行う
- ボルツマン分布に従うように初期速度を発生させる

$$f(v_x, v_y, v_z) = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \exp\left(-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2kT}\right)$$

$$f_1(v_x) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} \exp\left(-\frac{mv_x^2}{2kT}\right) \quad (\text{正規分布})$$



# (古典)分子動力学計算の流れ





## ニュートンの法則

- 第一法則 (慣性の法則)
  - 外力を受けない質点は、等速直線運動を行う
  - (数学的には、第二法則に含まれる)
- 第二法則 (運動方程式)
  - 質量 $m$ の質点の座標 $\mathbf{r}$ は、質点に働く力 $\mathbf{F}$ を用いて、以下のよう  
に表すことができる

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} = m_i \mathbf{a}_i$$

$\mathbf{F}_i$  質点*i*が受ける力  
 $m_i$  質点*i*の質量  
 $\mathbf{r}_i$  質点*i*の位置  
 $\mathbf{a}_i$  質点*i*の加速度

- 第三法則 (作用反作用の法則)

$$\mathbf{F}_{ij} = -\mathbf{F}_{ji}$$





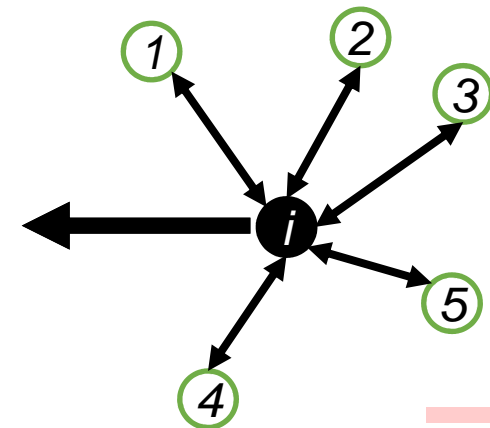
## その他の重要な法則

- 力の重ね合わせの原理
  - 質点に複数の力  $F^a, F^b, F^c, \dots$  が働くときに、質点はその和  $F$  が働いた場合と同様にふるまう

$$F = F^a + F^b + F^c + \dots$$

以上の基本原理を使って、質量を持つ粒子を動かしていく

“The science of simulating the motions of a system of particles” (Karplus & Petsko)



# 運動方程式の数値解法



- 運動方程式は解析的に解けないため、数値積分を行う。
- 数値積分により、次のステップの位置を計算する

$$\mathbf{r}_i(t_0) \rightarrow \mathbf{r}_i(t_0 + \Delta t) \rightarrow \mathbf{r}_i(t_0 + 2\Delta t) \rightarrow \cdots \mathbf{r}_i(t_0 + n\Delta t).$$

$$\mathbf{F} = m\ddot{\mathbf{r}}(t) = m\mathbf{a}$$

テイラー展開

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t\mathbf{v}(t) + \frac{\Delta t^2}{2!}\ddot{\mathbf{r}}(t) + O(\Delta t^3)$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t\mathbf{v}(t) + \frac{\Delta t^2}{2!}\frac{\mathbf{F}(t)}{m} + O(\Delta t^3)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \Delta t\dot{\mathbf{v}}(t) + O(\Delta t^2)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \Delta t\frac{\mathbf{F}(t)}{m} + O(\Delta t^2)$$

- 初期座標・初期速度があれば、数值的に解ける



# 運動方程式の数値解法 (Verletの方法)



Verletの方法(  $t+\Delta t, t-\Delta t$  のテイラー展開から導出 )

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}(t) + \frac{\Delta t^2}{2!} \ddot{\mathbf{r}}(t) + \frac{\Delta t^3}{3!} \dddot{\mathbf{r}}(t) + O(\Delta t^4)$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \Delta t \mathbf{v}(t) + \frac{\Delta t^2}{2!} \ddot{\mathbf{r}}(t) - \frac{\Delta t^3}{3!} \dddot{\mathbf{r}}(t) + O(\Delta t^4)$$

$$\mathbf{r}(t + \Delta t) + \mathbf{r}(t - \Delta t) = 2\mathbf{r}(t) + \Delta t^2 \ddot{\mathbf{r}} + O(\Delta t^4)$$

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \Delta t^2 \ddot{\mathbf{r}}(t) + O(\Delta t^4)$$

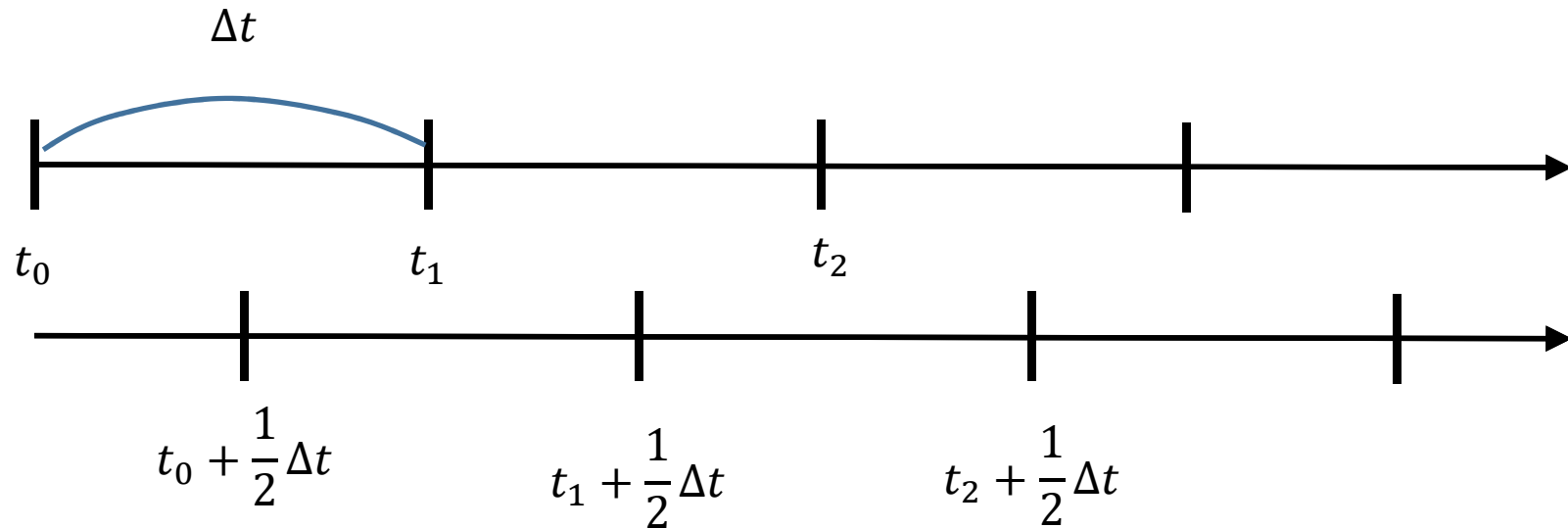
$$\mathbf{F} = m\ddot{\mathbf{r}}(t) = m\mathbf{a}$$

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{\Delta t^2 \mathbf{F}(t)}{m} + O(\Delta t^4)$$





# 運動方程式の数値解放 (蛙飛び法)

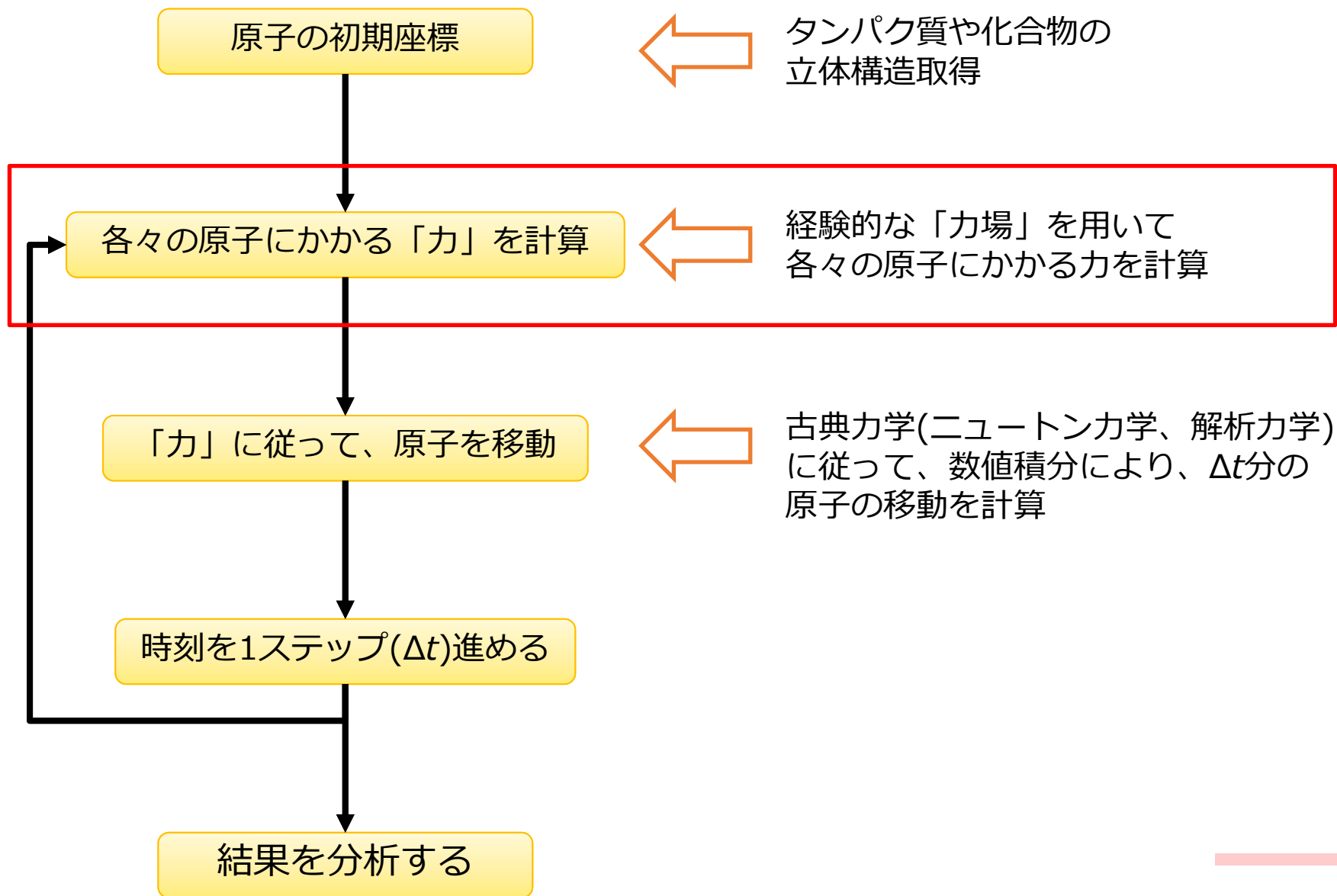


$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{1}{2}\Delta t\right)\Delta t$$

$$\mathbf{v}\left(t + \frac{1}{2}\Delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\Delta t\right) + \left[\frac{\mathbf{F}(t)}{m}\right]\Delta t$$



# (古典)分子動力学計算の流れ

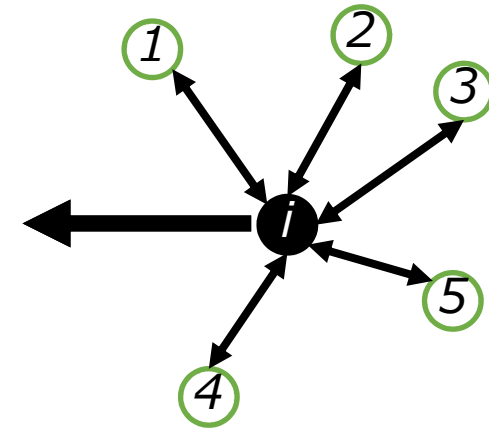


# 各々の原子にかかる「力」を計算する



## タンパク質に働く力

- 静電相互作用
- ファンデルワールス相互作用
- 水素結合
- 疎水性相互作用  
(水の存在が重要)



## ポテンシャルエネルギー

経験的な「力場」関数により、原子の座標を元に計算する  
ポテンシャルエネルギーを微分することで、原子に働く力が得られる

$$F_i = -\text{grad}_i E_i$$
$$\text{grad}_i = \mathbf{e}_x \frac{\partial}{\partial x_i} + \mathbf{e}_y \frac{\partial}{\partial y_i} + \mathbf{e}_z \frac{\partial}{\partial z_i}$$



# 分子力場計算法



Molecular Mechanics Method (MM法)

分子構造に関する経験的な概念を数学的に記述

様々な力場が提案されている

OPLS, Amber, CHARMM 等

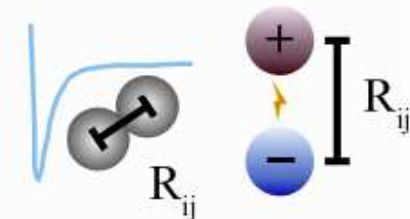
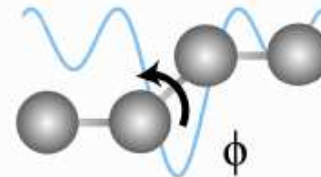
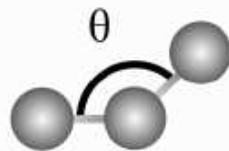
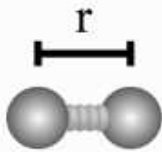
**計算量が非常に大きい**

クーロン力,  
ファンデルワールスカ  
**Non-bonded**

力場の例

共有結合  
**Bonded**

$$E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$



Molecular dynamics simulations and drug discovery, Jacob D Durrant and J Andrew McCammon, BMC Biology 2011, 9:71

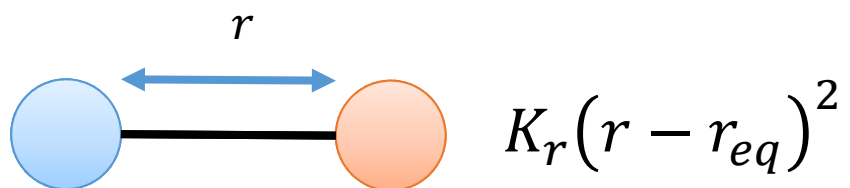
原子の位置によって、エネルギーが決まる



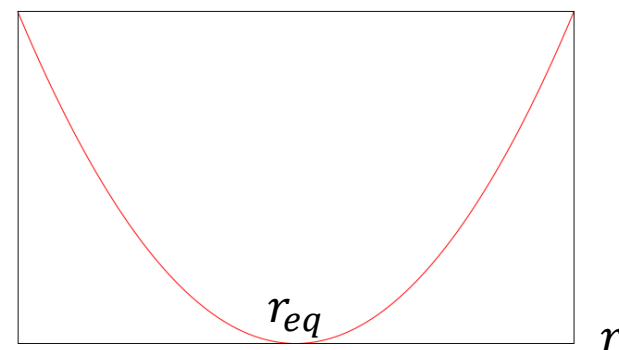
# 共有結合に関するポテンシャル



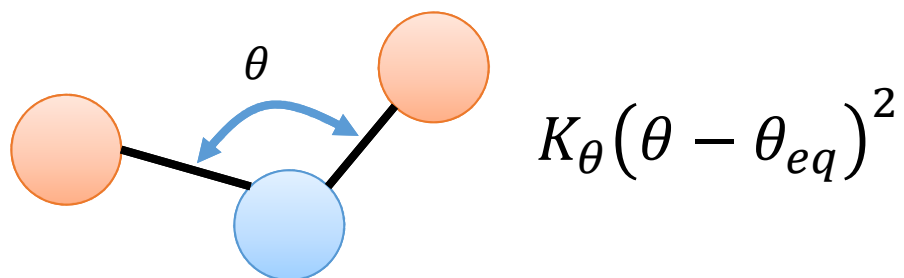
結合長 (bond length)



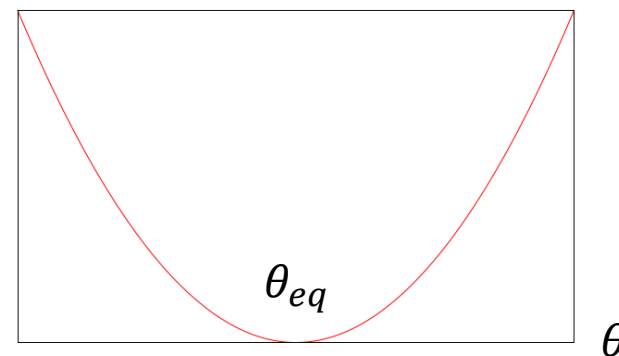
$E_{bond}$



結合角 (bond angle)



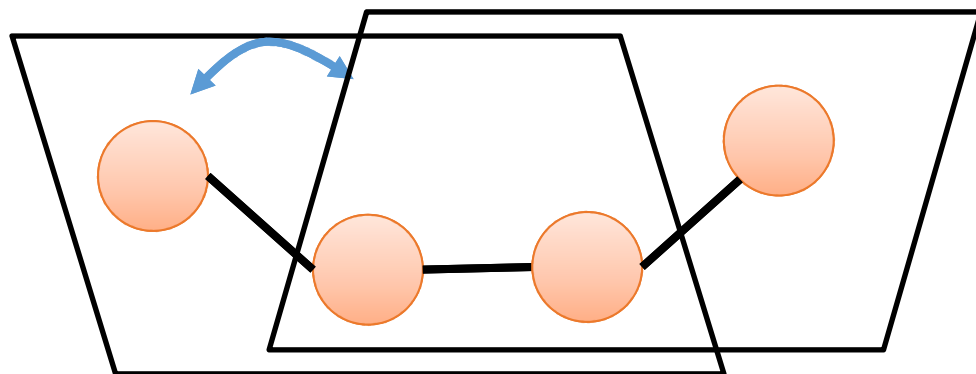
$E_{angle}$



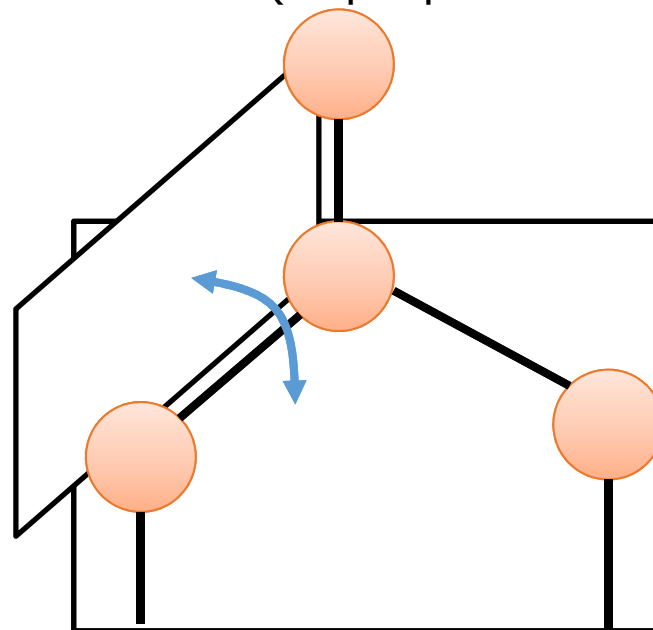
# 共有結合に関するポテンシャル



二面角 (torsion)

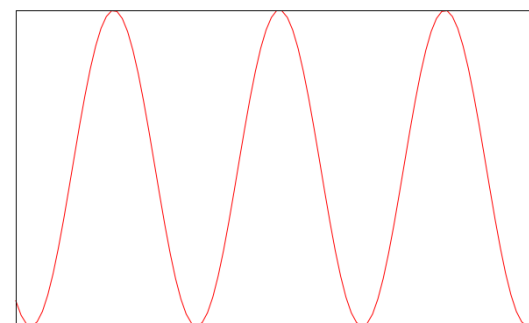


二面角 (improper torsion)



$$E_{\text{torsion}} = \frac{1}{2} V_{\phi} [1 + \cos(n\phi - \gamma)]$$

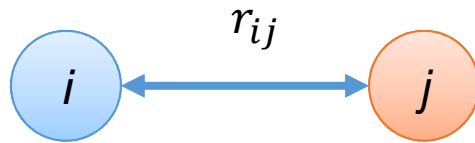
$$E_{\text{improper torsion}} = \frac{1}{2} V_{\phi} [1 + \cos(n\phi - \gamma)]$$



# 非結合項のエネルギー (クーロン力)



静電エネルギー



$$E_{\text{electrostatic}} = \frac{q_i q_j}{4\pi\epsilon r_{ij}}$$

$q_i$  : 原子*i*の電荷

$q_j$  : 原子*j*の電荷

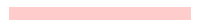
符号が異なる場合: 斥力

符号が等しい場合: 引力

力の大きさは、距離に反比例する

計算系全体の静電エネルギーは、系を構成するすべての原子ペアに対する静電エネルギーの和で表される

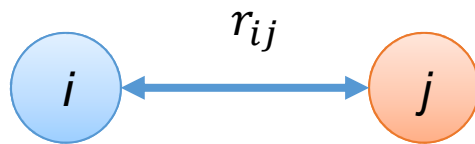
$$E_{\text{electrostatic}} = \sum_{i < j} \frac{q_i q_j}{4\pi\epsilon r_{ij}}$$



# 非結合項のエネルギー (ファンデルワールスカ)

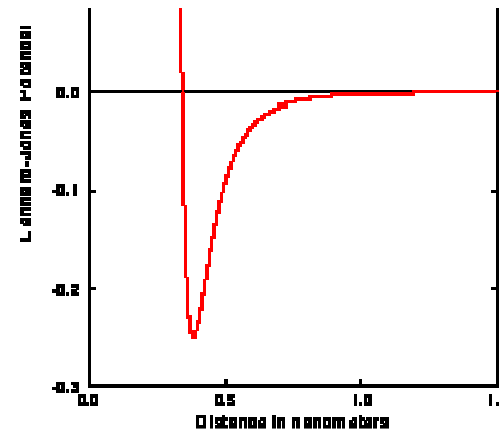


ファンデルワールス・エネルギー



$$E_{\text{van der Waals}} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

Lennard-Jones Energy



計算系全体のファンデルワールス・エネルギーは、系を構成するすべての原子ペアに対するファンデルワールス・エネルギーの和で表される

$$E_{\text{van der Waals}} = \sum_{i < j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$$





# 粗視化モデルと全原子モデル



---

- 粗視化モデル(ユナイテッド原子モデル)
  - 複数の原子を一つの粒子として扱う
  - 重元素と、それに結合する水素原子など
  - CHARMM19, GROMOS96等
  
- 全原子モデル
  - 水素原子も含め、全ての原子を別個の粒子として扱う
  - OPLS, Amber, CHARMM22, CHARMM27等



# 分子力場の種類



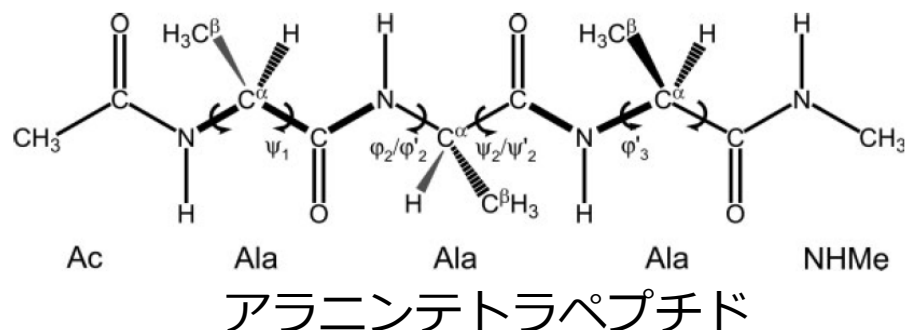
- OPLS
  - 結合と結合角のパラメータはAmber ff94と同じ
  - 二面角・非結合のパラメータを独自に決定
- Amber
  - ff94, ff96, ff99, ff03, ff99SB, ff99SB-ILDN等
  - ff94, ff99: ヘリックス構造を取りやすい
  - ff96 シート構造を取りやすい
- CHARMM
  - CHARMM19: 粗視化原子モデル, タンパク質
  - CHARMM22: 全原子モデル, タンパク質&水
  - CHARMM27: 全原子モデル, DNA, RNA, 脂質



# Amber分子力場の比較



実質的な自由度が二面角 $\varphi$ ,  $\phi$ のみの最小単位であるアラニンジペプチドや、グリニンジペプチドを用いて $\varphi$ ,  $\phi$ のパラメータを決定する

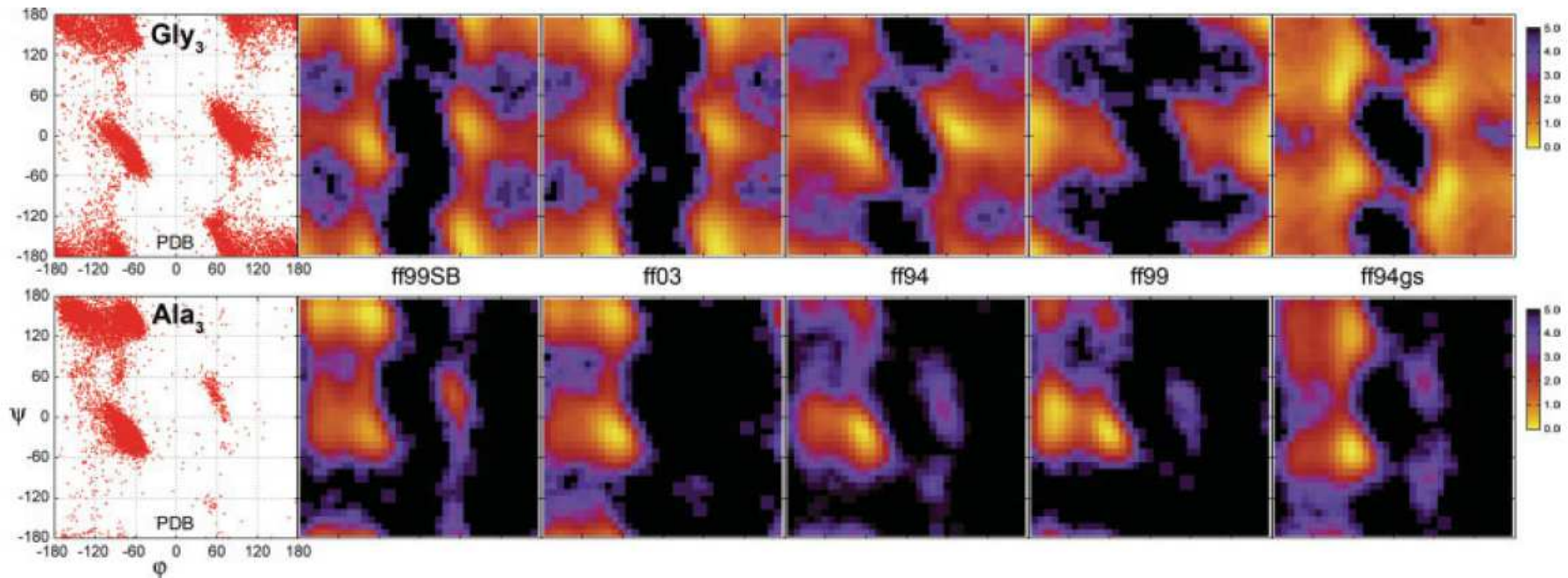
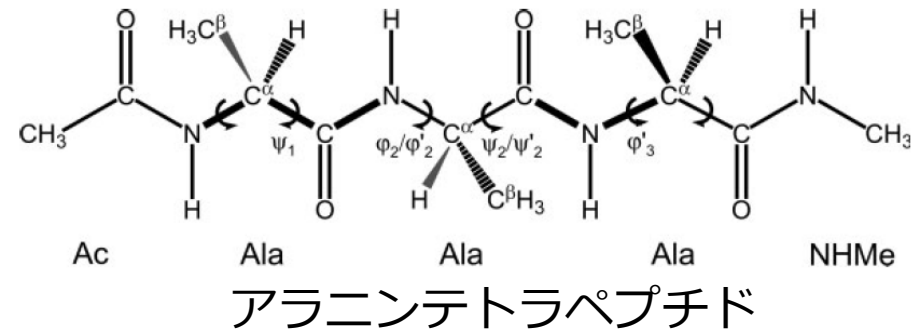


- ff94                    グリニンジペプチドにて主鎖二面角 $\varphi$ ,  $\phi$ の値を推定後、アラニンジペプチドにて二面角 $\varphi'$ ,  $\phi'$ の値を推定
- ff96                     $\varphi$ ,  $\phi$  ( $\varphi = \phi$ )の値を、アラニンテトラペプチドの伸張・ヘリックス状態のエネルギー差を再現するように調整
- ff99                    アラニンジペプチドの高精度量子計算により $\varphi$ ,  $\phi$ の値を推定
- ff99SB                グリニンジペプチドにて $\varphi$ ,  $\phi$ を推定後、アラニンジペプチドにより $\varphi'$ ,  $\phi'$ の値を推定
- ff03                    溶媒中での環境を直接QMで計算し、電荷と二面角を求める
- ff99SB-ILDN        ff99SBで $\varphi'$ ,  $\phi'$ が再現できないイソロイシン、ロイシン、アスパラギン酸、アスパラギンを用いて、 $\varphi'$ ,  $\phi'$ を最適化



# 分子力場の比較

## Hornakらの比較



Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters, Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling, PROTEINS: Structure, Function, and Bioinformatics, 65:712-725 (2006)

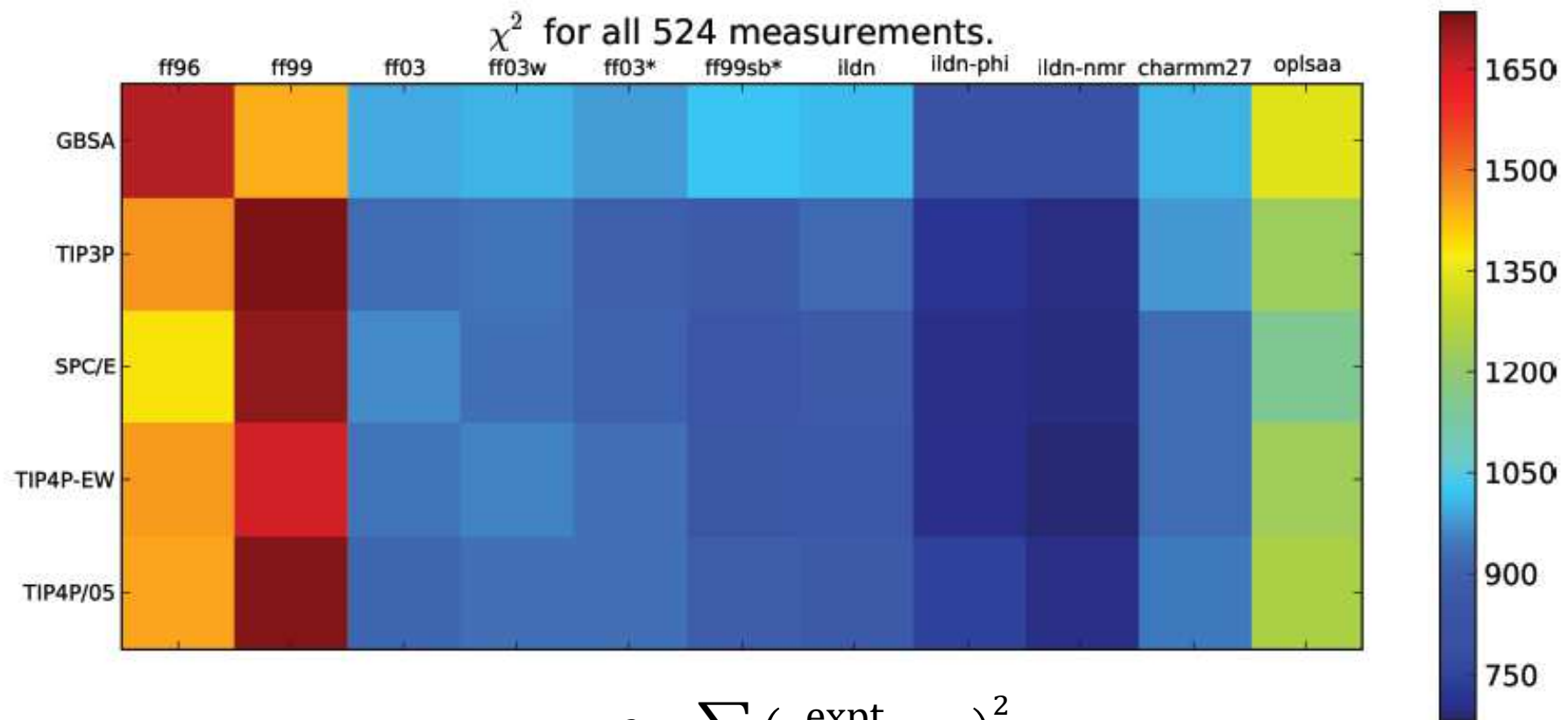
# 分子力場の比較



Are Protein Force Fields Getting Better?

A Systematic Benchmark on 524 Diverse NMR Measurements,

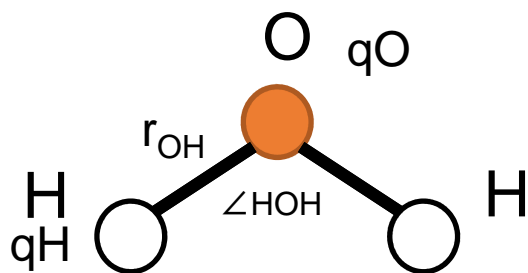
Kyle A. Beauchamp, Yu-Shan Lin, Rhiju Das, and Vijay S. Pande, Journal of Chemical Theory and Computation, 8. 1409-1414, 2012



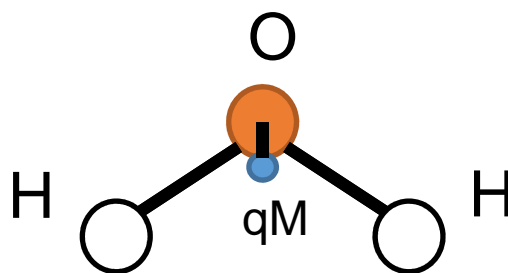
$$\chi^2 = \sum_i (x_i^{\text{expt}} - x_i)^2$$



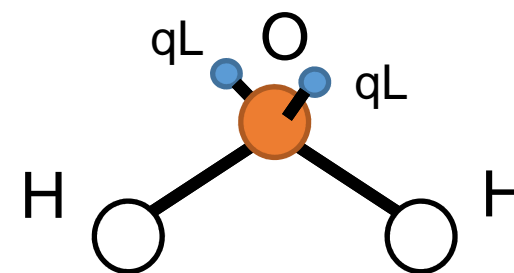
# 水分子の力場



TIP3P

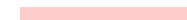


TIP4P

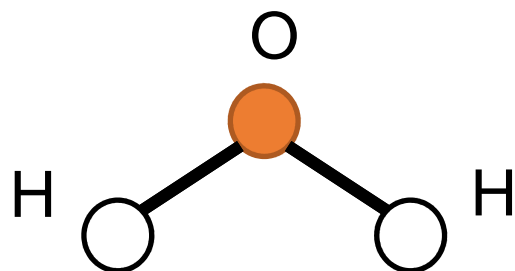


TIP5P

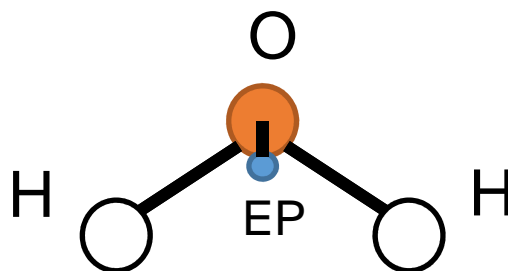
	点数	結合長 (Å)	結合角 ( $\angle\text{HOH}^\circ$ )	qH	qO	qM	qL
SPC	3点モデル	1	109.47	0.41	-0.82		
TIP3P				0.417	-0.834		
TIP4P	4点モデル	0.9572	104.52	0.52	0	-1.04	
TIP5P	5点モデル			0.241	0		-0.241



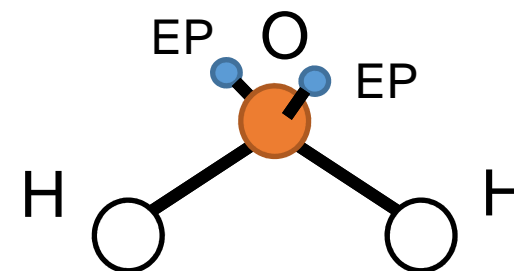
# 水分子の力場



TIP3P



TIP4P



TIP5P

Table 2. Computed properties for liquid water at 25°C and 1 atm

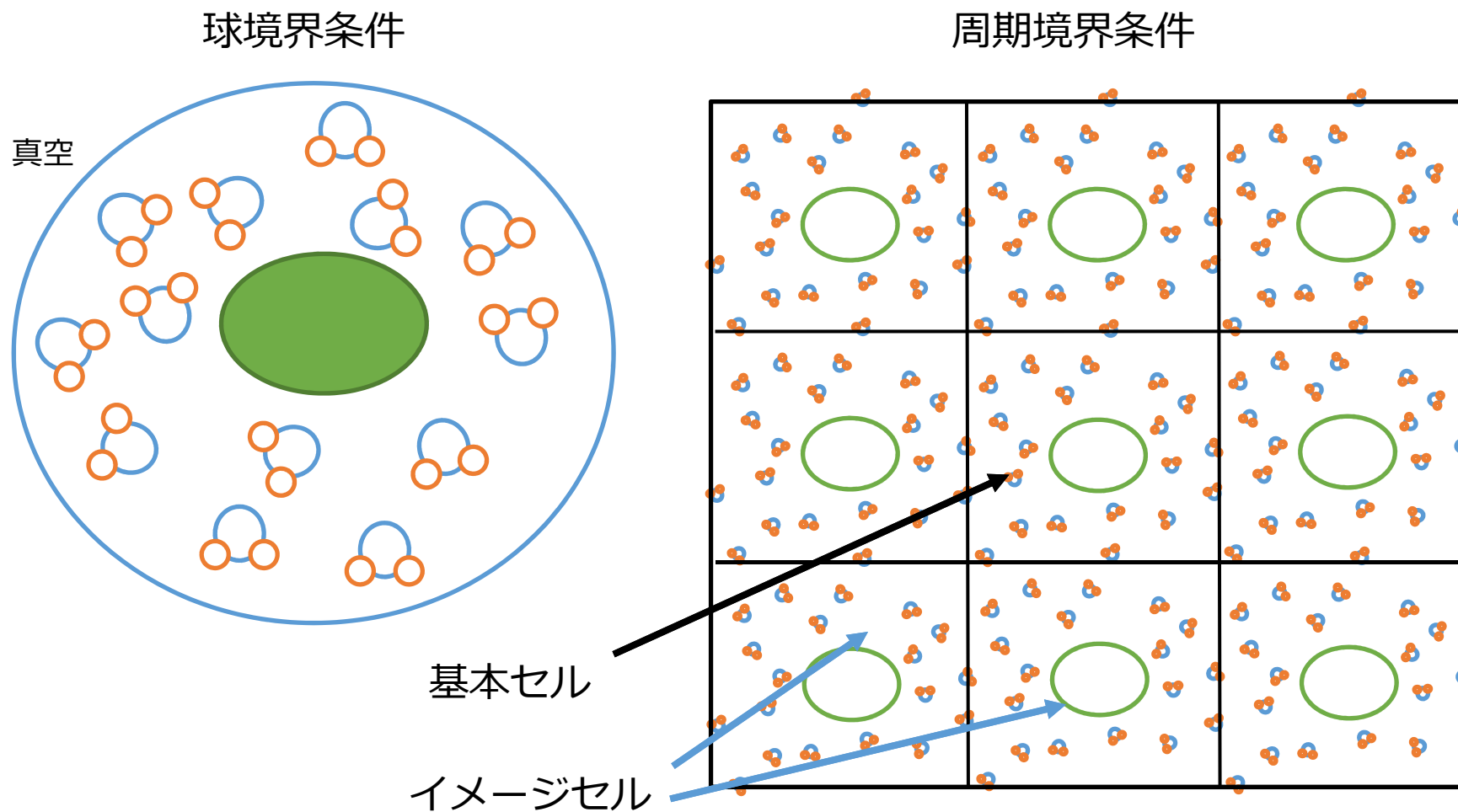
Model	$\rho$ , g/cm <sup>3</sup>	$\Delta H_{\text{vap}}$ , kcal/mol	$C_p$ , cal/mol-deg	$10^6 \kappa$ , atm <sup>-1</sup>	$10^5 \alpha$ , deg <sup>-1</sup>	$\epsilon$	$10^5 D$ , cm <sup>2</sup> /s
SPC	0.985	10.74	20	$60 \pm 4$	$106 \pm 8$	$60 \pm 10$	3.9
TIP3P	1.002	10.41	20	$64 \pm 5$	$92 \pm 8$	$88 \pm 6$	5.1
TIP4P	1.001	10.65	20	$60 \pm 5$	$44 \pm 8$	$60 \pm 10$	3.3
TIP5P	0.999	10.46	29	$41 \pm 2$	$63 \pm 6$	$82 \pm 2$	2.6
Exptl.	0.997	10.51	18.0	45.8	25.7	78.3	2.30

See refs. 70–72. Values are  $\pm$  SD. Exptl., experimental.

Jorgensen, W. L. *et al.*: *Proc. Natl. Acad. Sci. USA*, **102**, 6665 (2005)



# 遠距離相互作用の計算: 境界条件



球境界条件では、球境界付近での水の挙動が不自然になる





# 長距離相互作用における クーロン力・ファンデルワールス力の減衰



クーロン相互作用・ファンデルワールス相互作用の計算量  $\propto N^2$   
結合に関する計算量  $\propto N$

カットオフ法: もっとも単純な近似法

クーロン力

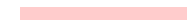
ファンデルワールス力

$$E_{\text{electrostatic}} = \frac{q_i q_j}{4\pi\epsilon r_{ij}}$$

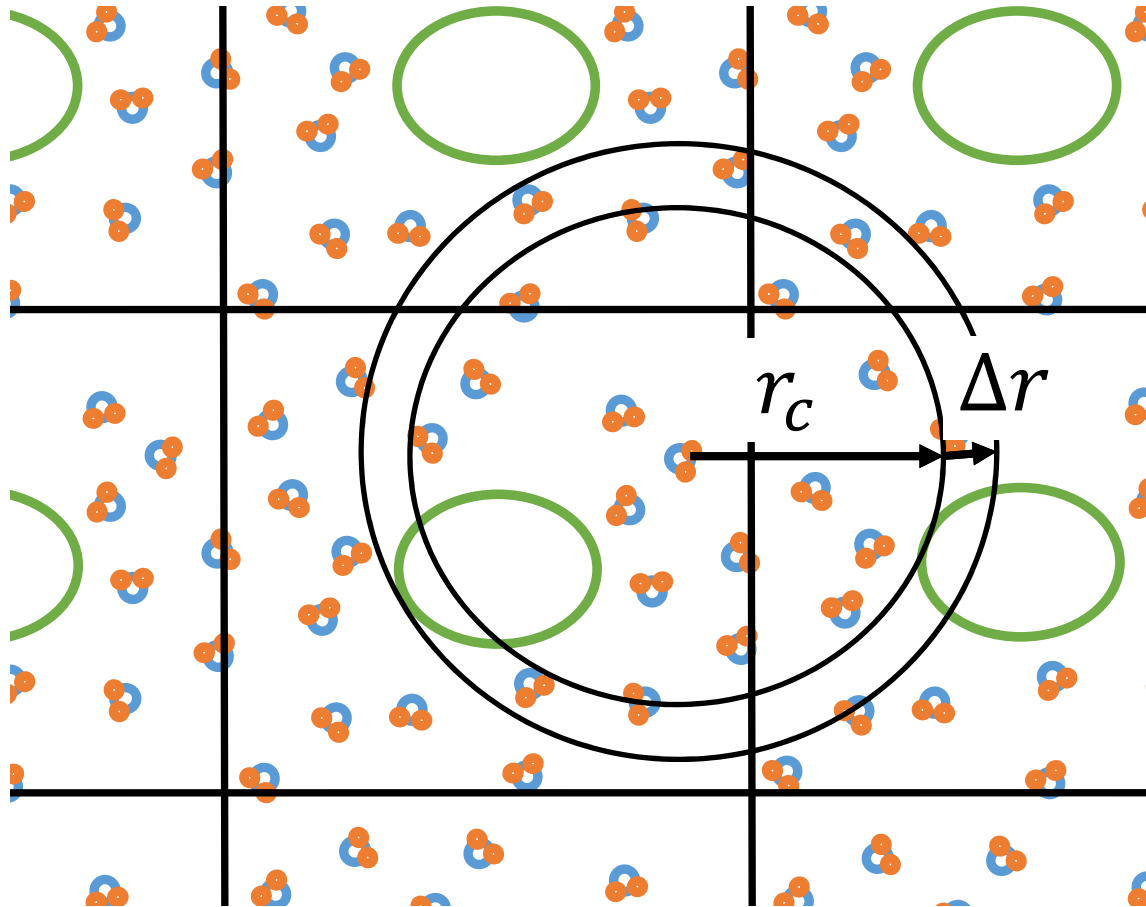
$$E_{\text{van der Waals}} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

クーロン力は、 $r_{ij}^{-1}$ 、ファンデルワールス力は $r_{ij}^{-6}$ に比例して減衰する

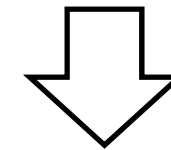
クーロン力は減衰が遅いため、  
長距離相互作用を考慮する必要がある



# ベルレの帳簿法 (距離によるカットオフ)



カットオフ距離 $r_c + \Delta r$ は、  
ボックスサイズの1/2以下



ボックスサイズは、  
カットオフ距離 $r_c + \Delta r$ の  
倍以上を指定する

1. 相互作用を計算しようとしている分子からの距離が $r_c + \Delta r$ よりも短い距離にある全ての分子をリストアップ
2. 一定のステップ数が経過するまでは、1. でリストアップした分子のみをエネルギー計算の対象とする



# 長距離相互作用における クーロン力・ファンデルワールス力の減衰



クーロン相互作用・ファンデルワールス相互作用の計算量  $\propto N^2$

カットオフ法: もっとも単純な近似法

クーロン力

ファンデルワールス力

$$E_{\text{electrostatic}} = \frac{q_i q_j}{4\pi\epsilon r_{ij}}$$

$$E_{\text{van der Waals}} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

クーロン力は、 $r_{ij}^{-1}$ 、ファンデルワールス力は $r_{ij}^{-6}$ に比例して減衰する

クーロン力は減衰が遅いため、  
長距離相互作用を考慮する必要がある



# 静電相互作用の計算(Ewald法)



系に周期性があることを利用して、エネルギーや力をフーリエ級数展開する  
無限遠からの影響を考慮する

周期的境界条件下での静電相互作用

$$E_{\text{electrostatic}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{2} \sum_n' \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon |\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|} \quad \text{収束が遅い}$$

Ewald法では、上記の式を以下のように表す

$$E_{\text{electrostatic}}(\mathbf{r}_N) = E_{\text{real}} + E_{\text{wave}} + E_{\text{self}}$$

$$E_{\text{real}} = \frac{1}{2} \sum_n' \sum_i \sum_j \frac{q_i q_j \operatorname{erfc}(\alpha |\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|)}{4\pi\epsilon |\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}|} \quad (\text{補助関数により}) \text{収束が速い}$$

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

$$E_{\text{wave}} = \frac{2\pi}{V} \sum_{\mathbf{m} \neq 0} \frac{\exp\left(-\frac{|\mathbf{m}|^2}{4\alpha^2}\right)}{|\mathbf{m}|^2} \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon} \cos[\mathbf{m} \cdot (\mathbf{r}_i - \mathbf{r}_j)] \quad \text{収束が速い}$$

$$E_{\text{self}} = \sum_i \frac{q_i^2}{4\pi\epsilon} \frac{\alpha}{\sqrt{\pi}} \quad \text{定数}$$



# 系の電荷を中和する



---

Ewald法やParticle Mesh Ewald法(PME法)は、系の電荷の和が0であることが前提  
系の電荷が0でない場合は、電荷を中和する(0にする)必要がある

- **イオンを発生させる**
  - 電荷を中和させるだけのイオン( $\text{Na}^+$ ,  $\text{Cl}^-$ など)を、溶媒に加える  
(溶媒分子を置き換える)
- 溶質の全原子に電荷を分散させる
- アミノ酸の電荷状態を調節する



# 分子動力学シミュレーションの実際の流れ



1. 構造情報の取得(PDB等)
2. 構造情報の確認、欠失領域の補完  
電荷、水素原子
3. シミュレーションボックスの定義  
境界条件の設定(カットオフ距離を考慮)
4. 水分子の追加
5. 電荷の中和  
イオンの追加等
6. エネルギー最小化 (最急勾配法等)
7. 系の平衡化(NVT -> NPT)
8. プロダクトラン  
実際の分子動力学計算
9. トrajエクトリの解析

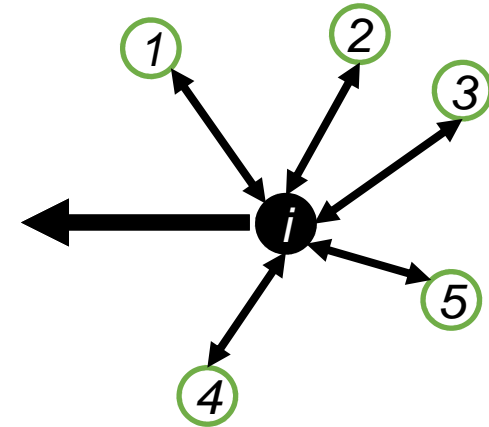


# 運動方程式の一般化



ニュートンの第二法則

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} = m_i \mathbf{a}_i$$



ラグランジアン  $L$

$$L = K - U$$

ポテンシャル  
運動エネルギー

ラグランジュの運動方程式

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0$$

一般化座標

$$\mathbf{q} = (q_1, q_2, \dots, q_n)$$



# ハミルトニアン



ハミルトニアン( $H$ ) : 系の全エネルギー

$$H(\mathbf{q}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}})$$

$\mathbf{p}$ : 一般化座標  
 $\mathbf{q}$ : 一般化運動量

ハミルトニアンの正準方程式

$$\dot{q} = \frac{\partial H}{\partial p}, \dot{p} = -\frac{\partial H}{\partial q}$$

全エネルギーの保存

$$\frac{dH}{dt} = \sum_{i=1}^n \left\{ \frac{\partial H}{\partial q_i} \frac{dq_i}{dt} - \frac{\partial H}{\partial p_i} \frac{dp_i}{dt} \right\} = \sum_{i=1}^n \left\{ \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} \right\} = 0$$

全エネルギーの時間変化が0なので、エネルギー保存則が成立する





# 代表的なアンサンブル



## アンサンブル

個々の座標や運動量は異なるが、熱力学的には同じ状態にある座標・運動量の集団  
分子動力学計算では、運動方程式を解いて得られるタンパク質構造のスナップショットの集合(トラジェクトリ)

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

通常分子動力学計算はNEVアンサンブル  
ニュートンの運動方程式, ハミルトニアン  
の正準方程式はエネルギー保存則が成り立つため

- カノニカルアンサンブル (NVTアンサンブル)
  - 粒子数( $N$ ), 体積( $V$ ), 温度( $T$ )一定
- T-Pアンサンブル (NPTアンサンブル)
  - 粒子数( $N$ ), 圧力( $P$ ), 温度( $T$ )一定



# 温度一定のMD (NVTアンサンブル)



## カノニカルアンサンブル

運動エネルギーと温度の関係

$$\sum_{i=1}^N \frac{m_i}{2} \mathbf{v}_i^2 = \frac{3}{2} (N - 1) k_B T$$

## 速度スケーリング法

温度を制御するためには、速度を制御すればいい

$$\mathbf{v}_i = s \mathbf{v}'_i$$

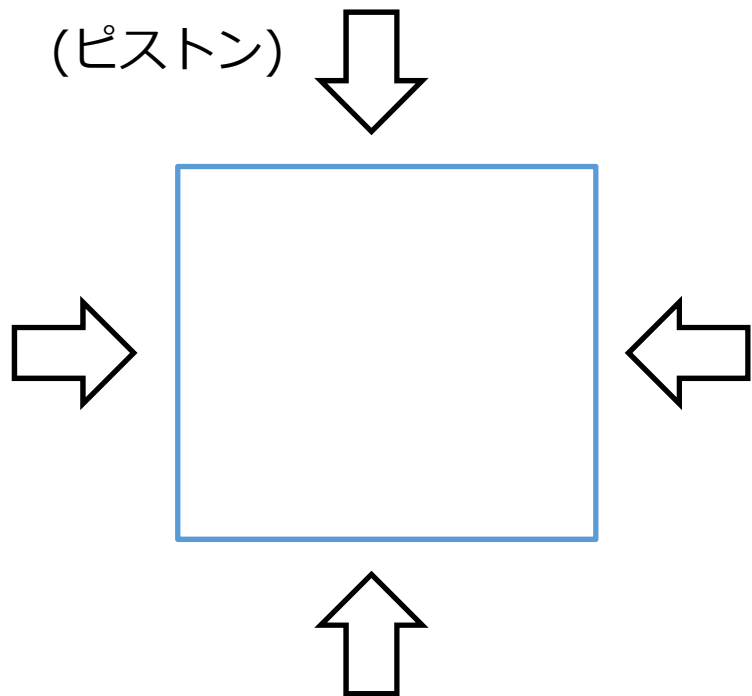
$$s = \left\{ \frac{(3N - 1) k_B T_0}{\sum_i m_i \mathbf{v}_i^2} \right\}$$

# 温度・圧力一定のMD (NPTアンサンブル)



能勢・アンダーソンの方法  
(温度制御) (圧力制御)

圧力を制御する方法  
(ピストン)



外界からの圧力 $P_0$ を考える

ポテンシャルエネルギー

$$V_V = P_0 V$$

全ての座標を、 $V^{\frac{1}{3}}$ でスケーリング

座標

$$\mathbf{r}_i = V^{\frac{1}{3}} \mathbf{q}_i$$

速度

$$\dot{\mathbf{r}}_i = V^{\frac{1}{3}} \dot{\mathbf{q}}_i$$

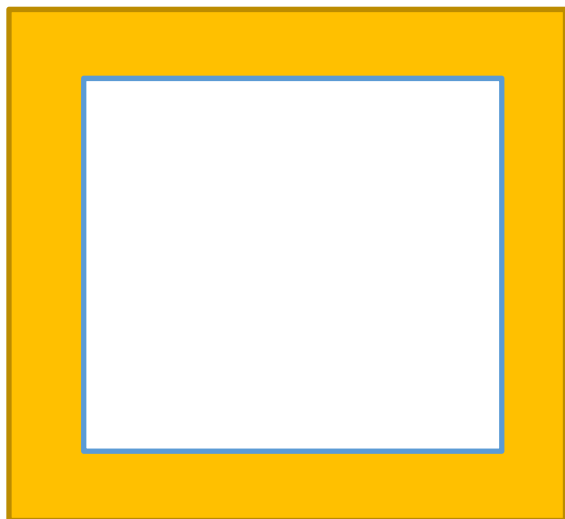


# 温度・圧力一定のMD (NPTアンサンブル)

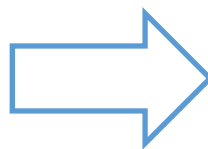


## 温度を制御する方法

熱浴



時間の進み方が異なる( $t \rightarrow t'$ )と考える



$$dt = \frac{dt'}{s}$$

運動量の関係

$$p_i = \frac{p'_i}{sV^{1/3}}$$



# 温度・圧力一定のMD (NPTアンサンブル)



仮想系のハミルトニアン  $H'$

$$H' = \sum_{(i)} \frac{\mathbf{p}'_i{}^2}{2m_i V^{2/3} s^2} + U(V^{1/3} \mathbf{q}^N) + \frac{P_s^2}{2Q} + 3Nk_B T_0 \ln s + P_{V^2} + P_O V$$

現実系の運動方程式

$$\ddot{\mathbf{q}}_i = -\frac{1}{m_i V^{2/3}} \frac{\partial U}{\partial \mathbf{q}_i} - \left( \frac{2\dot{V}}{3V} + \frac{\dot{s}}{s} \right) \dot{\mathbf{q}}_i \quad \ddot{s} = \frac{2s}{Q} \frac{3Nk_B}{2} (T - T_0) + \frac{\dot{s}}{s}$$

$$\ddot{V} = \frac{s^2}{W} (P - P_0) + \frac{\dot{s}}{s} \dot{V}$$

$V, pv, s, p_s$  に関する項を消去

: ミクロカノニカルアンサンブル (NEV)



$V, pv$  に関する項を消去

: カノニカルアンサンブル (NVT)

**バイオグリッドHPCIプロジェクト「新薬開発を加速する「京」インシリコ創薬基盤の構築」  
KBDD (K supercomputer-based drug discovery project by Biogrid pharma consortium)**

MP-CAFEE法\*による、  
タンパク質-リガンド間結合自由エネルギーの計算

\*H. Fujitani, et. al., “Massively parallel computation of absolute binding free energy with well-equilibrated states”,  
Physical Review E, 79, 021914, (2009)



# 分子動力学シミュレーションの課題(1)

## シミュレーション時間



1ヶ月程度で到達可能なシミュレーション時間  
(タンパク質、水系)

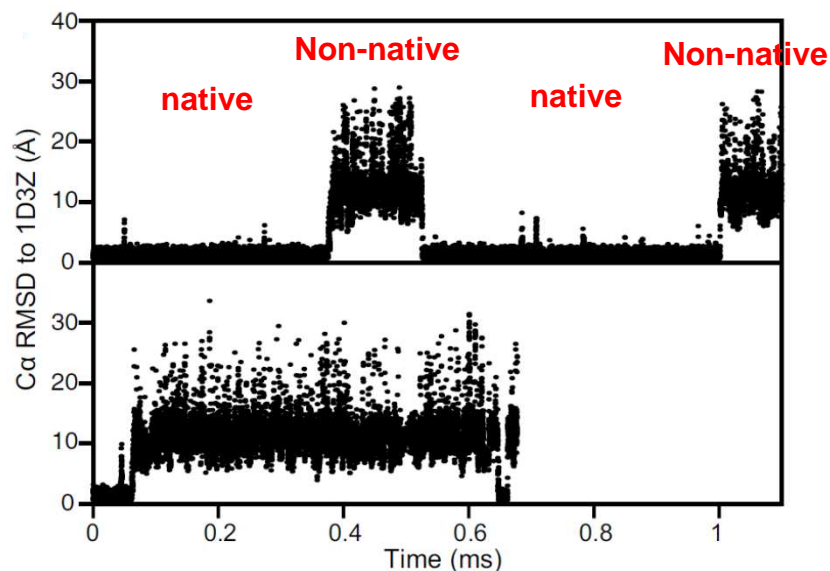
- ~100ns: 通常の計算機サーバ
- ~10 $\mu$ s: GPU搭載型計算機(20日  $\div$  10<sup>6</sup>秒)
- ~10 $\mu$ s: 京(並列計算の合計値)
- ~1ms: Anton (データ通信がボトルネック)



生命化学反応の時定数  
(ミリ秒~秒)

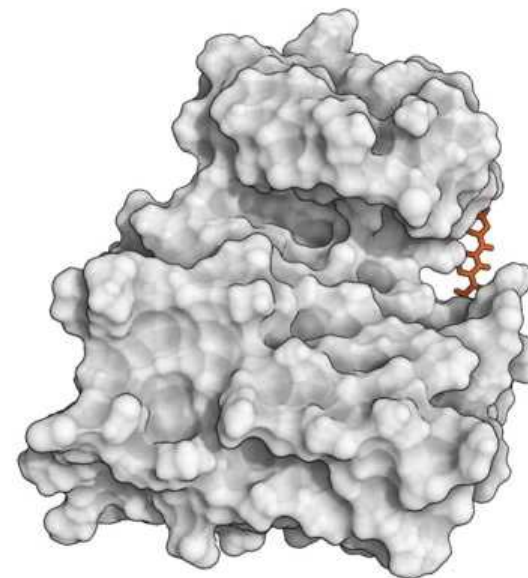
### Antonを用いた計算例

#### ユビキチンのfolding, unfolding



Piana et al., PNAS 2013, 110, 5915–5920

#### リガンドのタンパク質結合過程のシミュレーション



Shan et al., J. Am. Chem. Soc. 2011, 133, 9181–9183

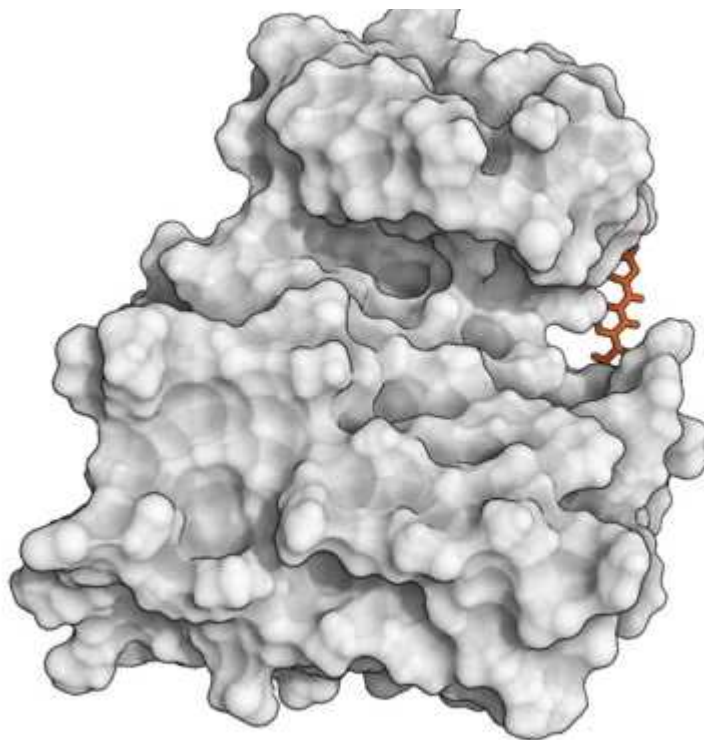
## How Does a Drug Molecule Find Its Target Binding Site?

Yibing Shan,<sup>†</sup> Eric T. Kim,<sup>†</sup> Michael P. Eastwood,<sup>†</sup> Ron O. Dror,<sup>†</sup> Markus A. Seeliger,<sup>§</sup> and David E. Shaw<sup>\*,†,‡</sup>

<sup>†</sup>D. E. Shaw Research, New York, New York 10036, United States

<sup>‡</sup>Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032, United States

<sup>§</sup>Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York 11794, United States



Src kinase(チロシンキナーゼ)と  
dasatinibの結合



# 分子動力学シミュレーションの課題(2)

## 分子数



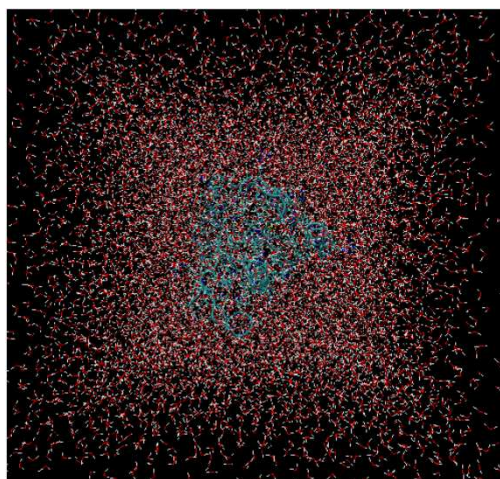
タンパク質のシミュレーションの場合

システム構成:

タンパク質: 1個 (水:  $10^4$ - $10^5$ 個)

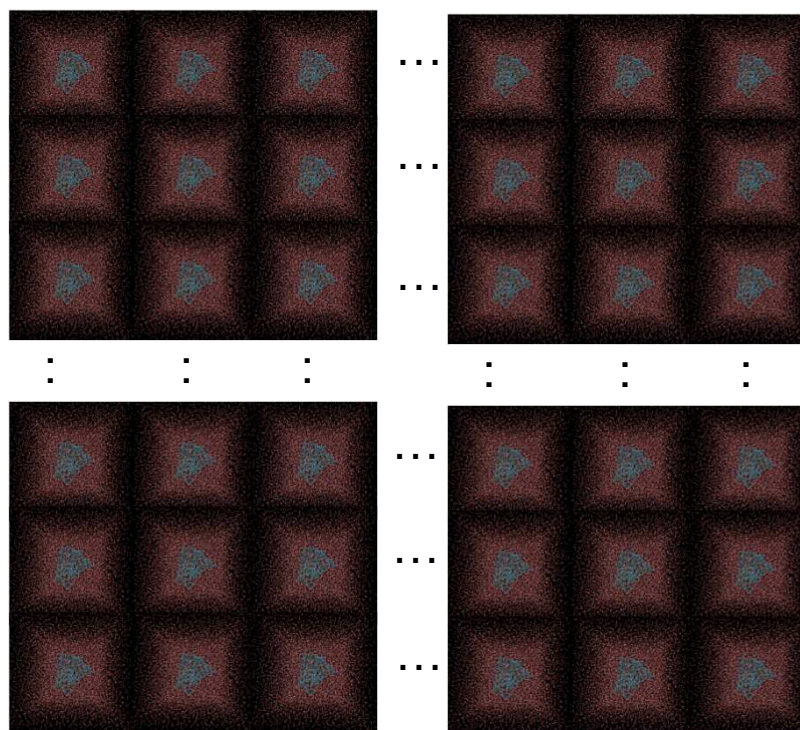
<<

アボガドロ数 ( $10^{23}$ 個)



~10nm

1分子シミュレーションから算出された物理量で熱力学量を推定しなければならない



京: 88000ノードを使用して、1mMのタンパク質溶液1mLの挙動を1秒間シミュレートしようとする、...

⇒  $1\text{s} / (10\text{ns/day}) \times 10^{17}\text{個} / 10^5\text{個} = 10^{20}\text{年かかる!!}$

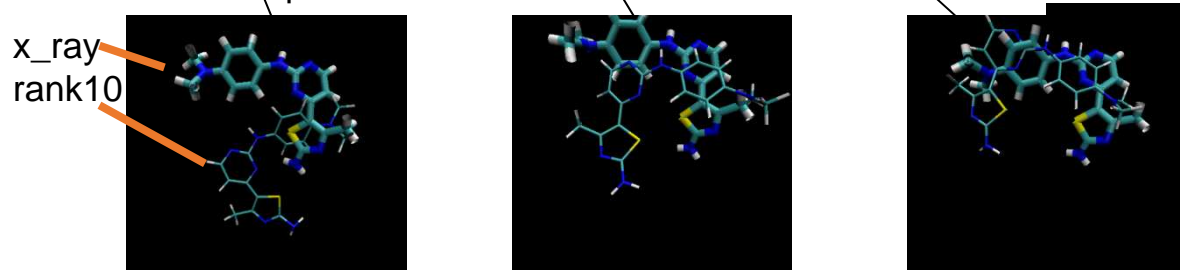
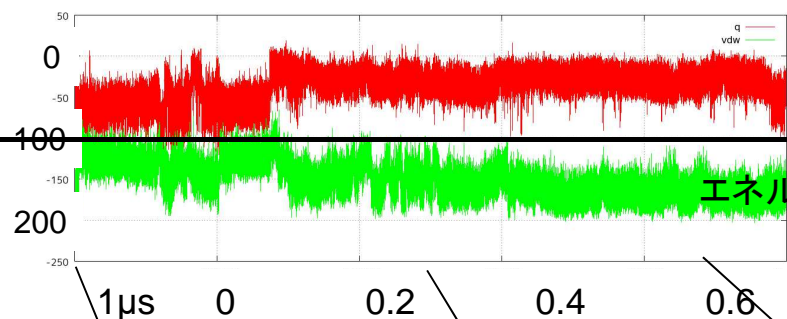
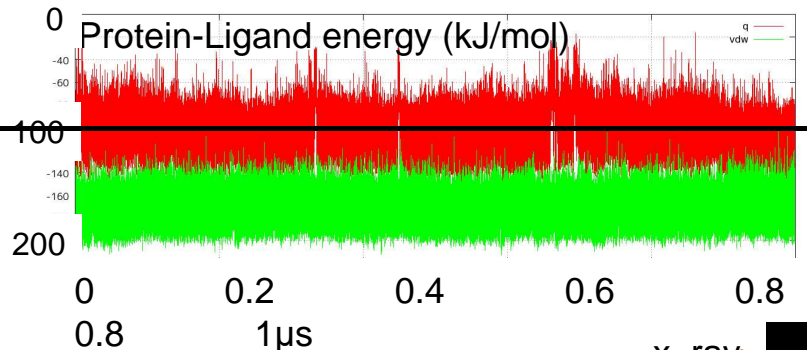
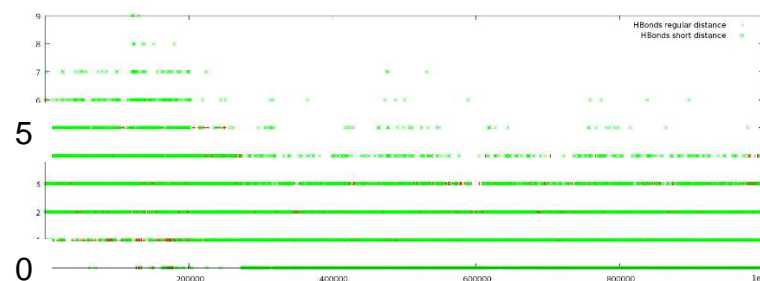
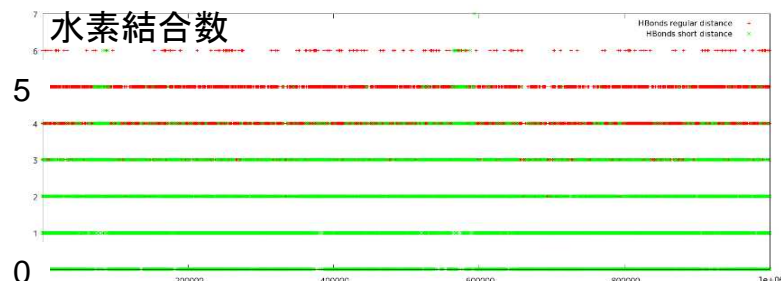
# 準安定構造からの長時間MDで最安定構造にたどり着けるか？



長時間MDによる化合物結合ポーズの追跡 (CDK2\_CS18)

結晶構造 (equil\_2)

docking rank10 (equil\_3)



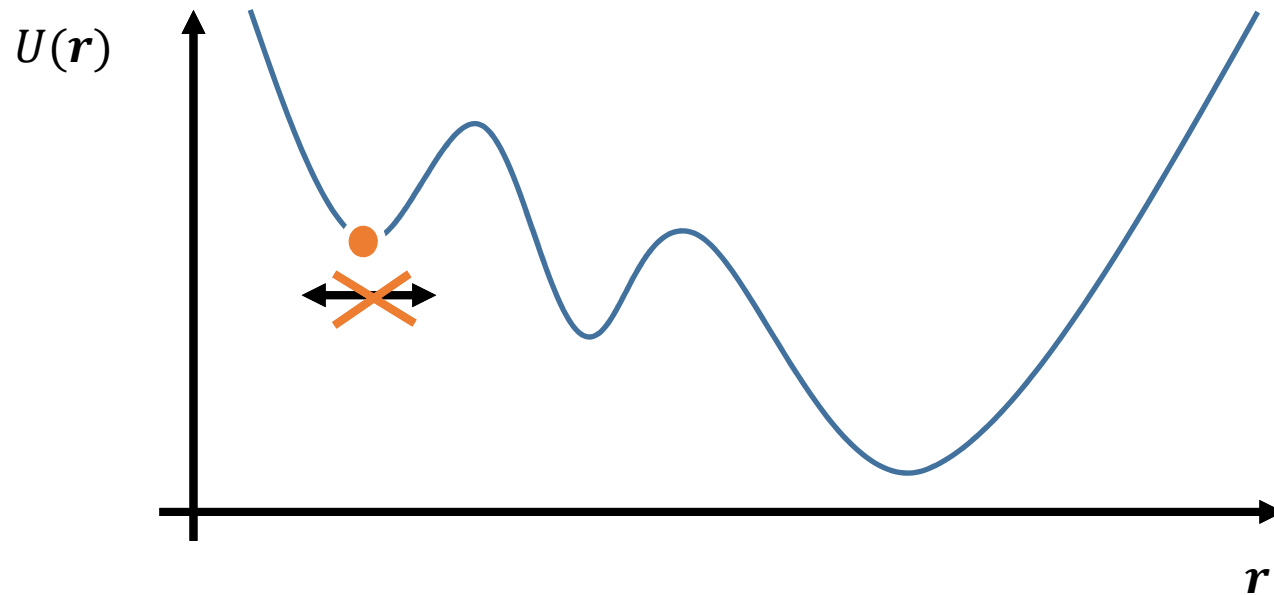
結合ポーズは1 $\mu$ s後も同じにならない

⇒正しい結合ポーズ(最安定構造)を予測するには温度レプリカ等の拡張サンプリングが必要

# トラップからの脱却



通常のMDシミュレーションでは、タンパク質の構造がエネルギー極小の領域にトラップされることがある



エネルギーが上がれば、トラップから抜け出すことができる  
原子の速度を上げる、温度を上げる

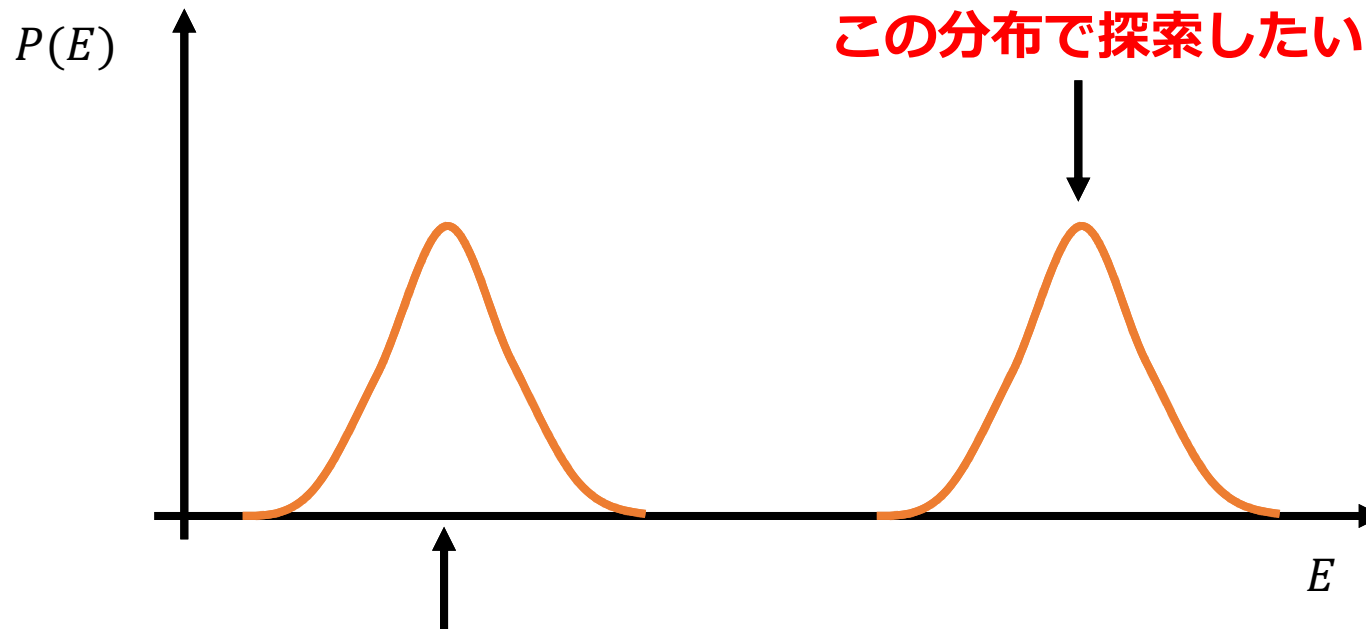


# カノニカルアンサンブルでのエネルギー分布



温度一定のMD

高エネルギーでの探索  
トラップから抜けられるが、  
求めるエネルギー(温度)での構造は  
得られない



**この分布で探索したい**

このエネルギー分布内で座標が移動  
なかなか高エネルギーにはならない

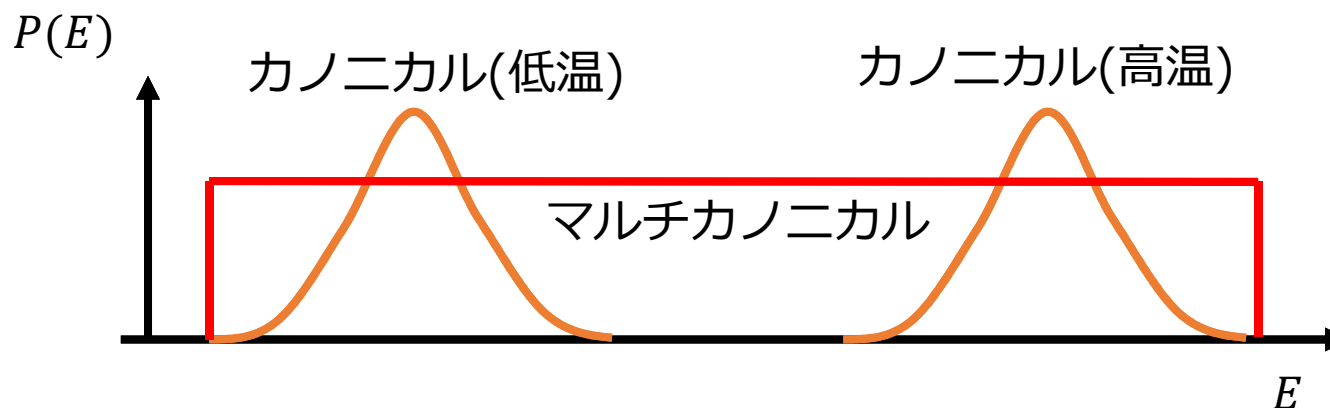
**この分布での構造が欲しい**



# 拡張サンプリング

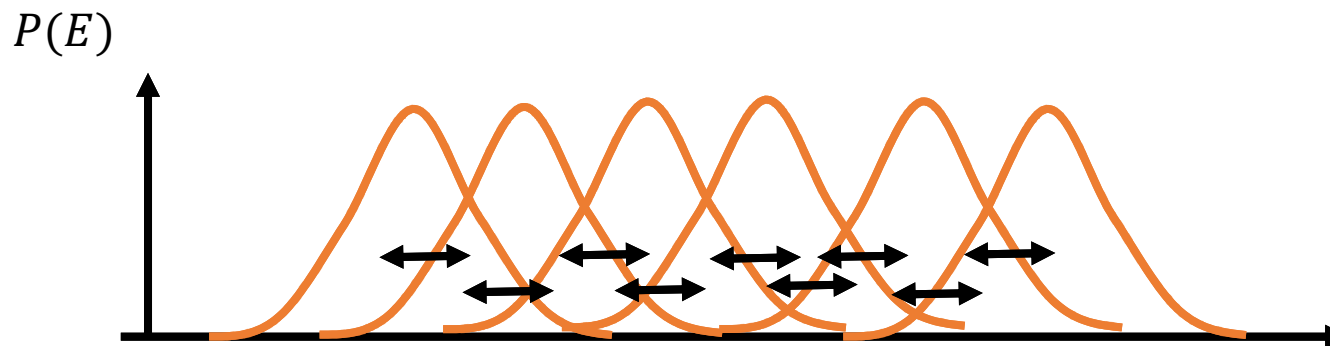


## マルチカノニカル法



一様のエネルギー分布空間において、構造探索を行う

## レプリカ交換法



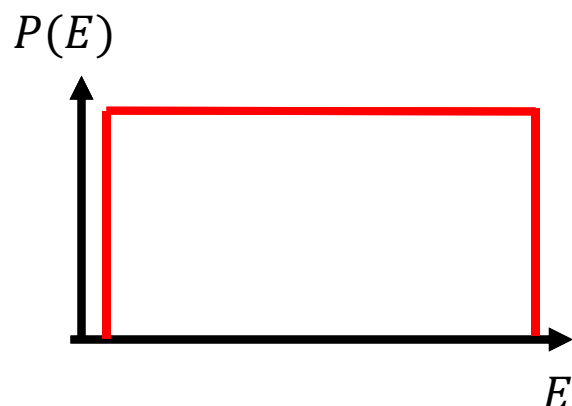
複数のエネルギー分布でMDを実行し、それらの間で温度または構造の交換を行いつつ、構造探索を行う



# マルチカノニカル法



一様分布になるように、人工的なエネルギー分布 $P_{mu}(E)$ を定義



$$P_{mu}(E) = n(E)W_{mu}(E) = \text{定数}$$

$$W_{mu}(E) = n(E)^{-1}$$

$n(E)$ : 状態密度

系の状態密度に反比例するように、  
マルチカノニカル重み因子 $W_{mu}(E)$ を定義

任意の温度 $T_0$ におけるマルチカノニカルポテンシャルエネルギー

$$E_{mu}(E; T_0) = k_B T_0 \ln n(E) = T_0 S(E)$$

マルチカノニカル法の運動方程式

$$\dot{\mathbf{p}}_k = \frac{\partial E_{mu}(E; T_0)}{\partial E} \mathbf{f}_k$$

ニュートンの運動方程式

$$\dot{\mathbf{p}}_k = \mathbf{f}_k$$



※事前に、短いMDでマルチカノニカル重み因子 $W_{mu}(E)$ を決定しておく必要がある

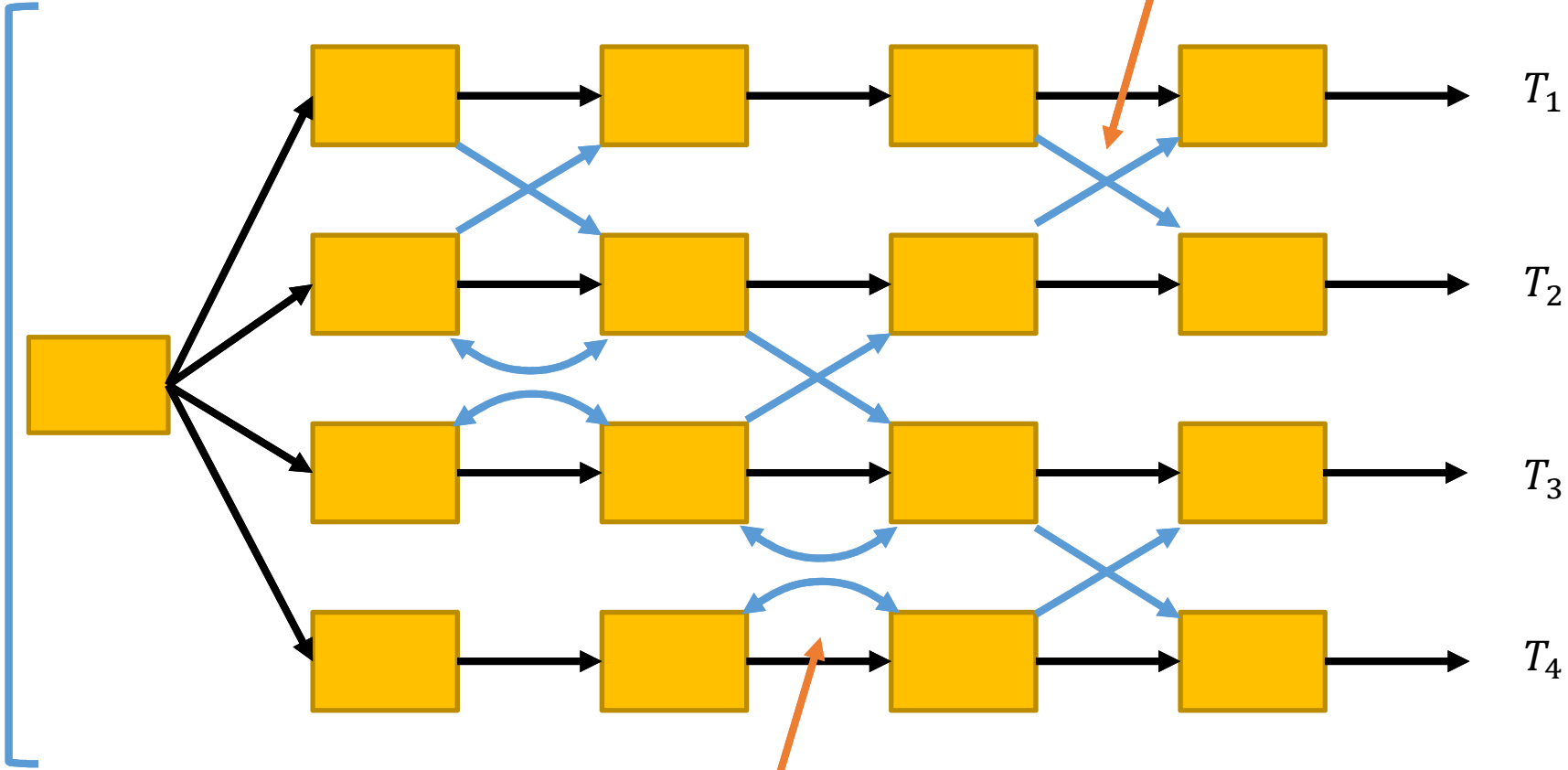


「生体系のコンピュータ・シミュレーション」, 岡崎進, 岡本祐幸 [編], 化学同人, 2002



# レプリカ交換法

複数の温度で、MD計算を並列して実行する  
温度を交換する



同じ系のレプリカ  
温度が異なるシミュレーション

交換不成立

シミュレーション時間



# レプリカ交換法の手順



1. 互いに相互作用しない系のコピー(レプリカ)を複数用意する
2. 各レプリカで、異なる温度において独立にカノニカルシミュレーションを実行する
3. 温度値が隣接した二つのレプリカを、詳細つり合いの条件に基づいて交換する

$$\Delta = (\beta_m - \beta_n) \{E(q^{[j]}) - E(q^{[i]})\}$$

$$\text{遷移確率 } w(X \rightarrow X') = \begin{cases} 1 & \Delta \leq 0 \\ \exp(-\Delta) & \Delta > 0 \end{cases}$$

4. 2. 3.を繰り返す。

事前に重み因子を決定する必要がない。

独立したMDシミュレーションを複数実行するため、並列コンピュータに適する。







## Hydrophobic Core Formation and Dehydration in Protein Folding Studied by Generalized-Ensemble Simulations

Takao Yoda,<sup>†\*</sup> Yuji Sugita,<sup>‡§</sup> and Yuko Okamoto<sup>¶</sup>

<sup>†</sup>Nagahama Institute of Bio-Science and Technology, Tamura, Nagahama, Shiga, Japan; <sup>‡</sup>CREST, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; <sup>§</sup>RIKEN Advanced Science Institute, Wako, Saitama, Japan; and <sup>¶</sup>Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan

- KBDDプロジェクト
  - CSAR CHK1\_89 re-docking #1からの、X線結晶構造の探索



# Hydrophobic Core Formation and Dehydration in Protein Folding Studied by Generalized-Ensemble Simulations

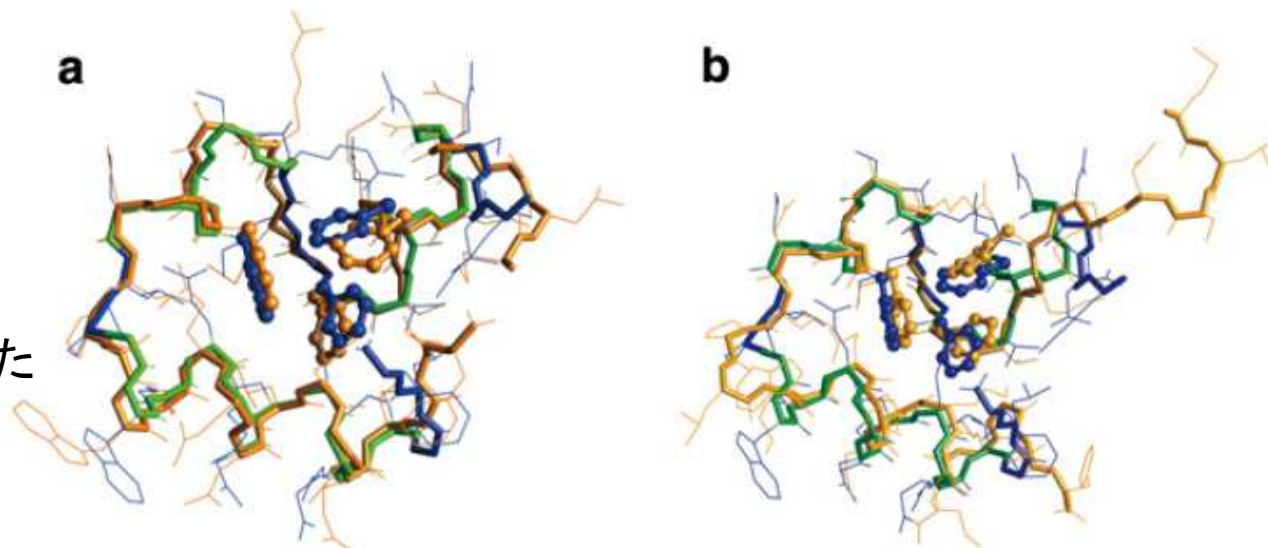
Takao Yoda,<sup>†\*</sup> Yuji Sugita,<sup>‡§</sup> and Yuko Okamoto<sup>¶</sup>

<sup>†</sup>Nagahama Institute of Bio-Science and Technology, Tamura, Nagahama, Shiga, Japan; <sup>‡</sup>CREST, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; <sup>§</sup>RIKEN Advanced Science Institute, Wako, Saitama, Japan; and <sup>¶</sup>Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan

HP36の完全に展開された構造から、  
MUCAREMシミュレーションによってフォールディングを行う

マルチカノニカル法と  
レプリカ交換法とを  
組み合わせたもの

全レプリカの合計2.28 $\mu$ sで  
フォールディングを再現できた



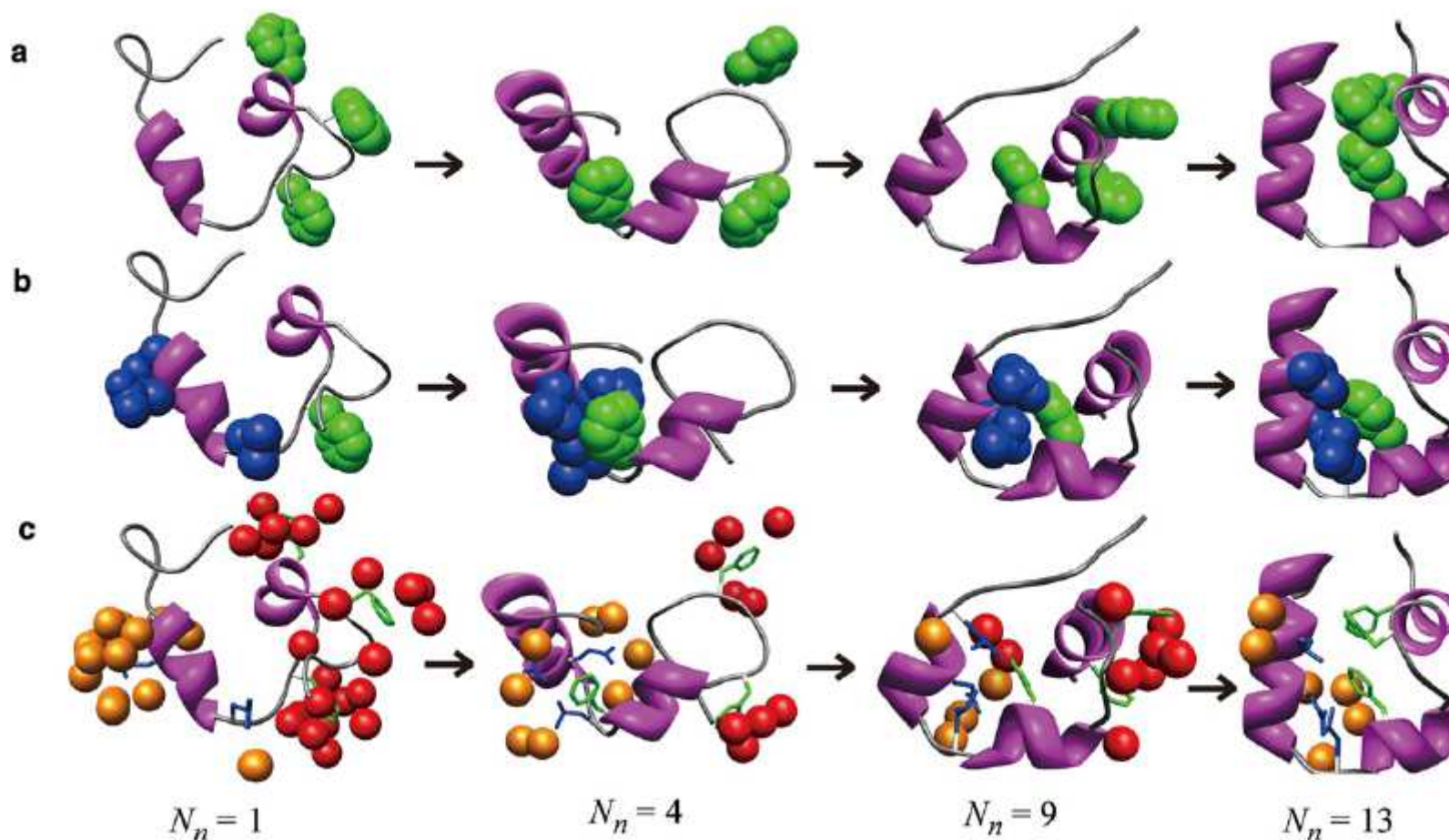
オレンジ: MUCAREMで得られた最小RMSDSの構造  
緑&青 : X線構造



## Hydrophobic Core Formation and Dehydration in Protein Folding Studied by Generalized-Ensemble Simulations

Takao Yoda,<sup>†\*</sup> Yuji Sugita,<sup>‡§</sup> and Yuko Okamoto<sup>¶</sup>

<sup>†</sup>Nagahama Institute of Bio-Science and Technology, Tamura, Nagahama, Shiga, Japan; <sup>‡</sup>CREST, Japan Science and Technology Agency, Kawaguchi, Saitama, Japan; <sup>§</sup>RIKEN Advanced Science Institute, Wako, Saitama, Japan; and <sup>¶</sup>Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan



# 謝辞



## KBDDコンソーシアムメンバー

NPO法人 バイオグリッドセンター

奥野恭史 教授, 荒木望嗣 様 (理化学研究所計算科学研究機構), J. B. Brown 先生 (京都大学医学研究科), 白石慧 様, 佐藤美和 様 (三井情報(株)), 広川貴次 研究チーム長 ((独)産業技術総合研究所), 金井千里 主任研究員((株)京都コンステラテクノロジーズ)

## 製薬企業 (23社):

アステラス製薬(株)、アスピオファーマ(株)、エーザイ(株)、小野薬品工業(株)、  
科研製薬(株)、(株)カネカ、キッセイ薬品工業(株)、杏林製薬(株)、協和発酵キリン(株)、参天製薬(株)、  
塩野義製薬(株)、千寿製薬(株)、大正製薬(株)、大日本住友製薬(株)、田辺三菱製薬(株)、  
帝人ファーマ(株)、東レ(株)、日産化学工業(株)、日本新薬(株)、日本たばこ産業(株)、マルホ(株)、  
MeijiSeikaファルマ(株)、持田製薬(株)

## IT企業 (2社):

(株)京都コンステラ・テクノロジーズ, 三井情報(株)

(独) 産業技術総合研究所

理研・計算科学研究機構

高度情報科学技術研究機構

<http://sciencechannel.jst.go.jp/M120001/detail/M120001027.html>

**JST サイエンスチャンネルにて放送中**

NHK Eテレ サイエンスZERO No. 519

これが世界一のシミュレーション！ 「スーパーコンピュータ 京」  
(2015年10月4日放送)



# 参考文献等



- 「タンパク質計算科学 基礎と創薬への応用」, 神谷成敏・肥後順一・福西快文・中村春木, 共立出版, 2009
- 「コンピュータ・シミュレーションの基礎[第2版]」, 岡崎進・吉井範行, 化学同人, 2011
- 「生体系のコンピュータ・シミュレーション」, 岡崎進・岡本祐幸 編, 化学同人, 2002
- Improved side-chain torsion potentials for the Amber ff99SB protein force field, Kresten Lindorff-Larsen, Stefano Piana, Kim Palmo, Paul Maragakis, John L Klepeis, Ron O Dror, and David E Shaw, *Proteins*, 78(8), 1950-1958, 2010
- 「生体分子の分子動力学シミュレーション(1) 方法」, 古明地勇人, 上林正巳, 長嶋雲兵, *J. Chem. Software*, Vol. 6, No. 1, pp. 1-36, 2000
- Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements, Kyle A. Beauchamp, Yu-Shan Lin, Rhiju Das, and Vijay S. Pande, *Journal of Chemical Theory and Computation*, 8. 1409-1414, 2012
- Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters, Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling, *PROTEINS: Structure, Function, and Bioinformatics*, 65:712-725 (2006)

