



# 分子動力学シミュレーションソフトウェア 「GENESIS」ハンズオン講習会 part1 GENESISの概要

小林千草

理化学研究所・計算科学研究センター・  
運用技術部門ソフトウェア開発技術ユニット

2022/05/11



## 今日の内容

- 13:00-13:10 連絡事項
- 13:10-13:30 GENESISの概要
- 13:30-14:20 GENESISでのMDハンズオン (前半)
- 14:20-14:35 休憩
- 14:35-15:15 GENESISでのMDハンズオン (後半)
- 15:15-16:00 フリーディスカッション

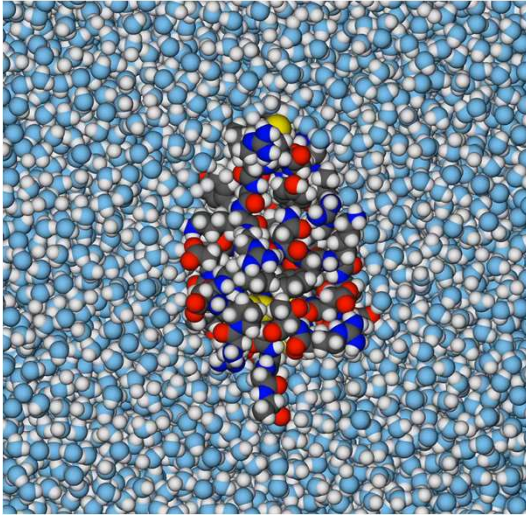


## 講義の内容

- 分子動力学法ソフトウェアGENESIS
- GENESISの高速化、性能
- GENESISでできること
- 今日のハンズオンの簡単な説明



# 分子動力学法(MD)シミュレーション



タンパク質(水溶液中)

ニュートン方程式による計算

1. **原子に働く力**を計算
2. 数値積分(時間発展)の繰り返しにより運動の解析が可能

**1msの計算には $5 \times 10^{11}$ 回の繰り返しが必要**  
ボトルネックは力の計算

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m}$$

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i \text{ 力}$$

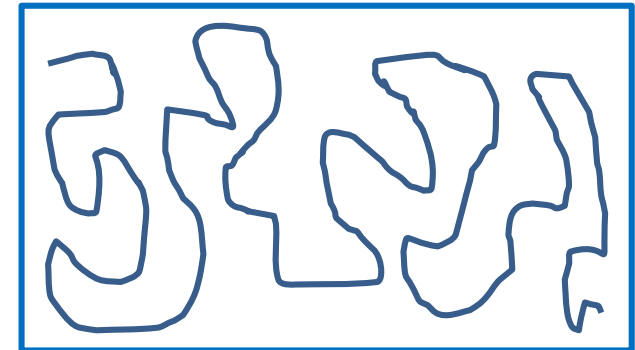
運動方程式



$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \frac{\mathbf{p}_i}{m} \Delta t$$

$$\mathbf{p}_i(t + \Delta t) = \mathbf{p}_i(t) + \mathbf{F}_i \Delta t$$

数値積分



長時間の原子・分子運動



# Generalized Ensemble Simulation Systems (GENESIS)

Version 1.0: Jung, Mori, *et al.* *WIREs Comput. Mol. Sci.* **5**:310-323, (2015)

Version 1.1.x: Kobayashi, Jung, *et al.* *J. Comput. Chem.* **38**, 2193-2206 (2017)

- 理化学研究所にて2010年から開発
- スーパーコンピュータの性能を最大限に引き出すことを目的に独自の高速化、超並列手法
- HPCIシステムの多くのスパコンにプリインストール  
([https://www.hpci-office.jp/pages/hardware\\_software?tab=software](https://www.hpci-office.jp/pages/hardware_software?tab=software))
- フリーソフト(LPGLv3)として公開中 → 最新は公開版1.7.1と開発版2.0β

GENESISウェブページ

<https://www.r-ccs.riken.jp/labs/cbrt/>

“**GENESIS riken**”で  
検索してください！



## ソフトウェアGENESIS

- 2つのMDエンジンと30を超えるトラジェクトリ解析用ツールからなるアプリケーションスイート
- 異なるアルゴリズムを持つ2つのMDエンジン
  - ✓ 超並列・高速化コードを持つspdyn (富岳利用はこちら)
  - ✓ 粗視化モデルやQM/MM計算も可能で、ユーザが比較的容易に機能を導入可能なatdyn
- 言語はFortran90、一部がC言語
- spdyn/atdyn共にMPIとOpenMPによるハイブリッド並列



## GENESISの特徴



GENESIS  
Generalized-ensemble simulation system

独自に開発された高速化・並列手法

レプリカ交換MD法(REMD)などの  
サンプリング計算

全原子・粗視化モデルによるマルチ  
レゾリューション手法



# 内容

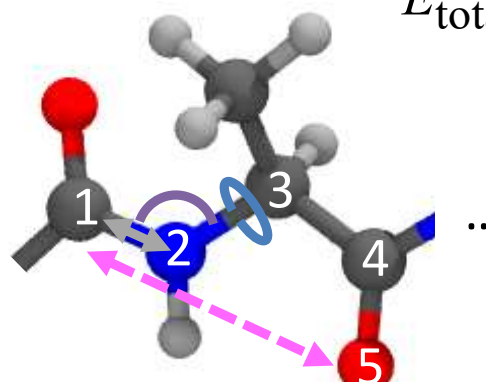
- 分子動力学法ソフトウェアGENESIS
- GENESISの高速化、性能
- GENESISでできること
- 今日のハンズオンの簡単な説明





# 生体分子MDの力の計算

力の計算は結合性、非結合性相互作用により計算


$$E_{\text{total}} = \sum_{\text{bonds}} k_b (b - b_0)^2 \quad \text{Bond (ex. 1-2, 2-3, ...)}$$
$$+ \sum_{\text{angles}} k_a (\theta - \theta_0)^2 \quad \text{Angle (ex. 1-2-3, 2-3-4, ...)}$$
$$+ \sum_{\text{dihedrals}} V_n [1 + \cos(n\omega - \gamma)] \quad \text{Dihedral (ex. 1-2-3-4, ...)}$$

結合性  
 $O(N)$

水素: 薄灰  
炭素: 濃灰  
酸素: 赤  
窒素: 青

$$+ \sum_{i,j \notin \text{bonding}} \left\{ \epsilon_{ij} \left[ \left( \frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{0ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \right\}$$

van der Waals      Coulomb  
(ex. 1-5, ...) :      (ex. 1-5, ...) :

非結合性  
 $O(N^2)$

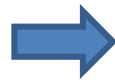
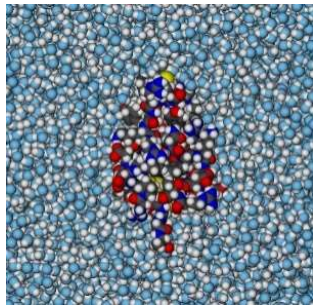
演算時間のボトルネックは非結合性相互作用計算



## 非結合性相互作用

- ボトルネックとなる非結合相互作用をカットオフ長より近い部分と遠い部分に分け、遠い部分をフーリエ変換により逆空間で計算
- 直接のペア計算はNのオーダーとなる

全ペア  
 $O(N^2)$



カットオフ長より近い部分  
 $O(CN)$

+ 遠い部分

$$\sum_{|r_{ij}| < R \& i, j \notin \text{bonding}} \left\{ \varepsilon_{ij} \left[ \left( \frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{0ij}}{r_{ij}} \right)^6 \right] f(r_{ij}) + \frac{q_i q_j \text{erfc}(\alpha r_{ij})}{r_{ij}} \right\} + \sum_{\mathbf{k} \neq 0} \frac{\exp(-\mathbf{k}^2 / 4\alpha^2)}{\mathbf{k}^2} \text{FFT}(Q(\mathbf{k}))$$

実空間,  $O(CN)$

計算・メモリアクセス・1対1通信

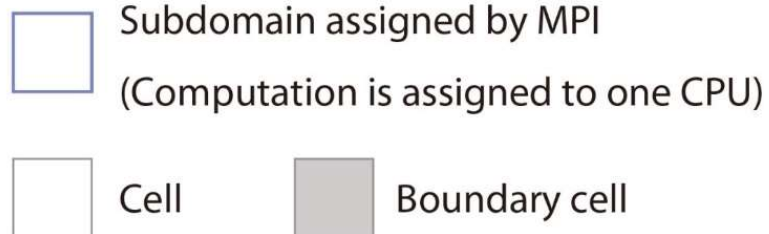
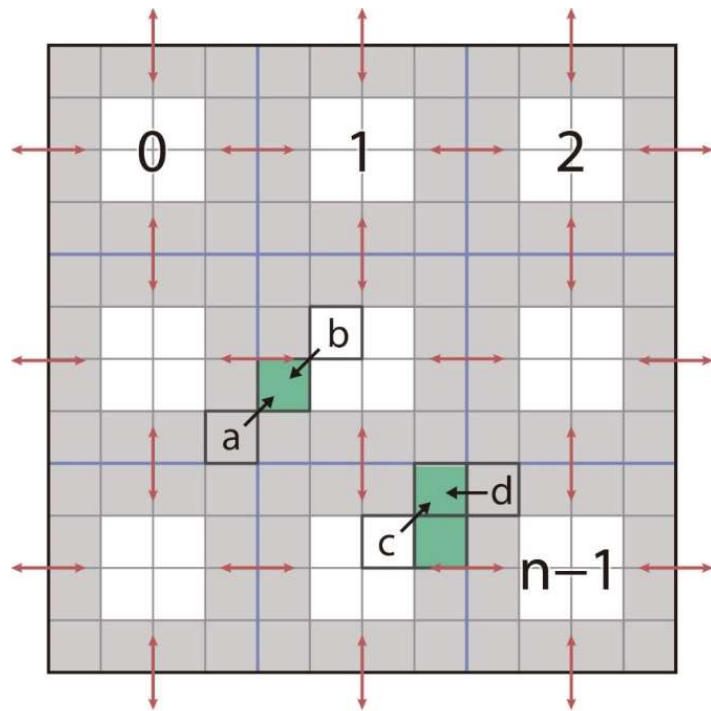
逆空間,  $O(M \log M)$

全体通信



## ドメイン分割: Midpoint cell method

Jung et al. *J. Comput. Chem.* **35**, 1064 (2014)



1. シミュレーション空間をMPIのプロセス数に応じてサブドメインに分ける
2. サブドメインを更に小さなセルに分ける
3. MPIコミュニケーションは隣接プロセスと境界のセルの部分について行う

### 利点

- 高い並列効率
- プロセス数が増えると演算、メモリ双方が減少



GENESIS

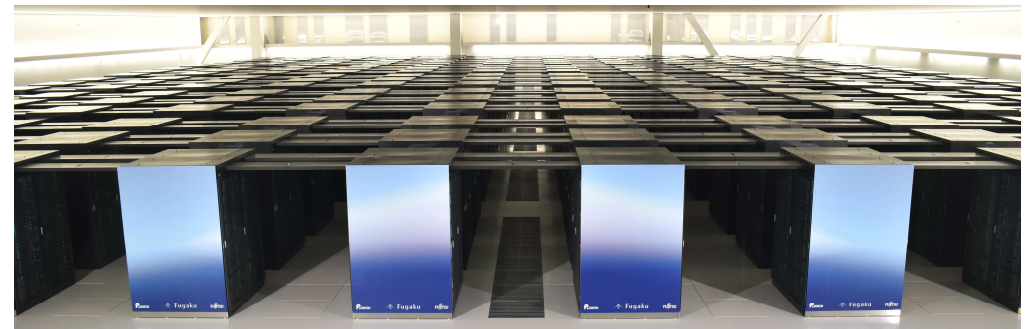
Generalized-ensemble simulation system

# スーパーコンピュータ「富岳」



1<sup>st</sup> in TOP500, HPCG, HPL-AI and Graph500 (Jun. & Nov. in 2020 & 2021)

- 158,976 nodes (1 CPU/node)
- CPU: A64FX
  - Armv3.2-A SVE 512 bit SIMD
  - 48 cores (12cores/CMG)
  - Normal 2.0 GHz (boost 2.2GHz)
- メモリ: 32 GiB/node
- ストレージ: 3階層(LLIO, FEFS, cloud)
- I/O: PCIe Gen3 x 16 lane
- Interconnect: Tofu Interconnect D (28 Gbps x 2 lane x 10 port)





# 「富岳」に向けた高速化

## 短距離相互作用

Jung et al. *J. Comput. Chem.* **42**, 231-240 (2021)  
<https://doi.org/10.1002/jcc.26450>

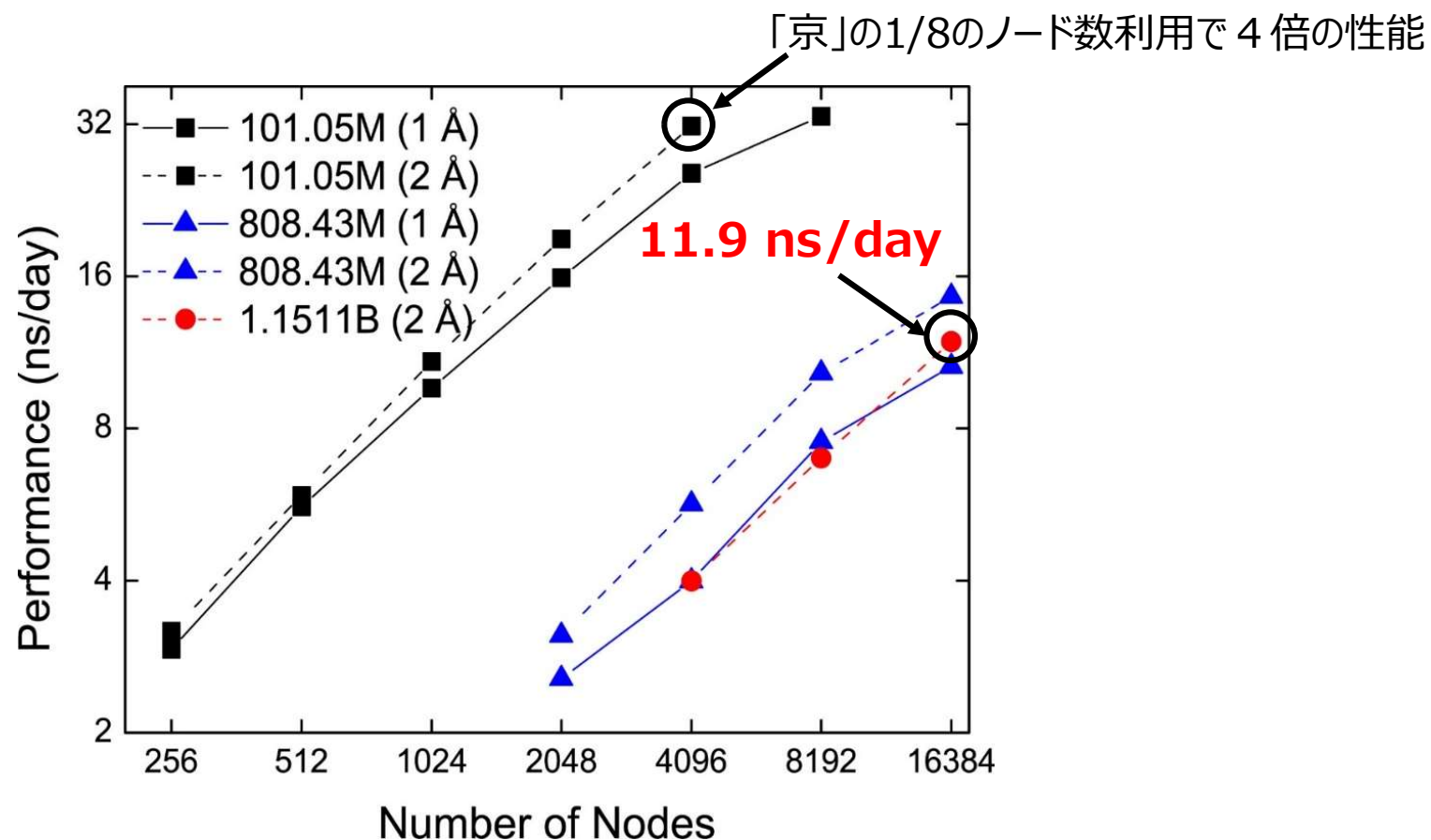
1. 力の計算のループ長を伸ばす
2. 力の計算と対生成での座標等の配列構造を変える
3. L1 cacheに対してprefetchを掛ける
4. 構造体内の配列に“CONTIGUOUS”属性を付ける

## 長距離相互作用

1. FFT計算やグリッド計算の手法を複数用意し、計算サイズや並列度により実行時に選択

$\beta$ バージョン (GENESIS2.0 $\beta$ )として公開中(現在正式公開に向け準備中)

# 「富岳」でのGENESIS 2.0 $\beta$ の性能



Jung *et al.* *J. Comput. Chem.* **42**, 231-241 (2021).



# 内容

- 分子動力学法ソフトウェアGENESIS
- GENESISの高速化、性能
- **GENESISでできること**
- 今日のハンズオンの簡単な説明



# GENESISでできること (MD編)

GENESISのマニュアルより(chapter 3)

Table 3.1: Available functions in **atdyn** and **spdyn**

Function	atdyn	spdyn
Energy minimization	○ (SD and LBFGS)	○ (SD)
All-atom molecular dynamics	○	○
Coarse-grained molecular dynamics	○	○ (All-atom Go-model)
Implicit solvent model	○	—
Replica-exchange method	○	○
Gaussian accelerated MD	○	○
Reaction path search	○ (MEP and MFEP)	○ (MFEP)
QM/MM calculation	○	—
Vibrational analysis	○	—
Cryo-EM flexible fitting	○	○
Precision	double	double/mixed
GPGPU calculation	—	○ (All-atom MD)
Parallel I/O	—	○

Version 1.7.1

Version 1.7.1 & 2.0b





## GENESISでできること (解析編)

トラジェクトリ解析用ツール(CPUマシンで利用可)の一部

解析ツールのほとんどはデスクトップ、ラップトップでも利用可能

- Distance/angle/dihedral angles計算 : trj\_analysis
- RMSD計算: rmsd\_analysis
- PCA計算: avecrd\_analysis, flccrd\_analysis, eigmat\_analysis, prjcrd\_analysis  
(avecrd\_analysis, flccrd\_analysisでRMSF計算も)
- 自由エネルギー解析: mbar\_analysis, wham\_analysis, pmf\_analysis ...
- 拡散計算: msd\_analysis, diffusion\_analysis
- 膜分子の計算: lipidthink\_analysis, tilt\_analysis

...



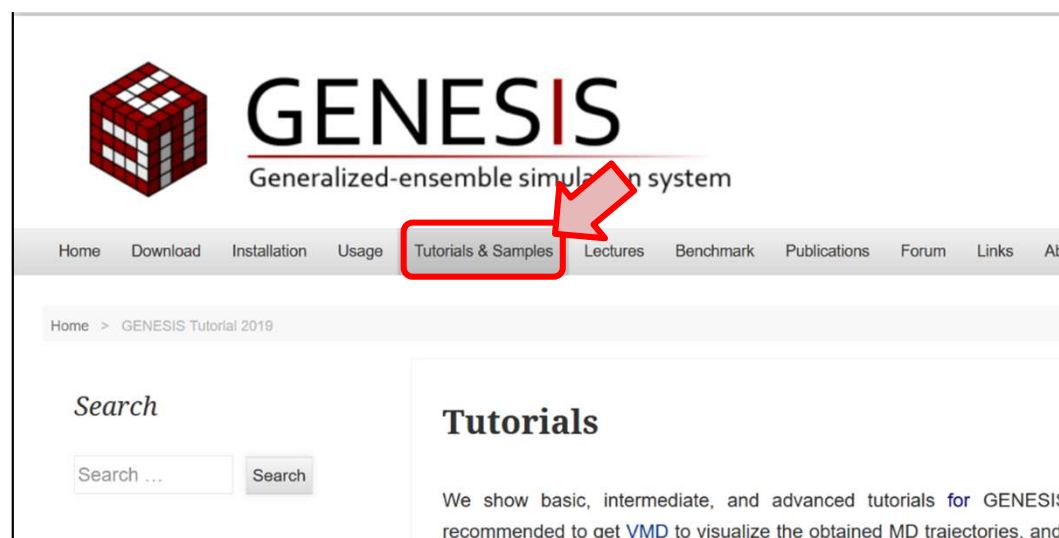
# GENESIS 2.0 $\beta$ について

- 高速化されたspdyn (マシンやMPI数により最適な方法を自動的に選択)
- より長い時間幅の計算、新しいmultiple time step(MTS)手法が可能  
(Jung & Sugita *J. Chem. Phys.* 153 234115 (2020), [Jung et al. J. Chem. Theory Comput.](#) 17, pp5312-5321 (2021))
- 可能な計算 **SPDYN & all atom FF** (CHARMM/AMBER FF)
  - MD, Minimization
  - Simulations with replicas: REMD, REUS, gREST, GaMD, String method
- **2.0 $\beta$ のSPDYNで出来ない機能**
  - Leapfrog integrator (削除)
  - Langevin thermostat (削除)
  - FEP ([1.7.1を利用ください](#))
- 2022年初夏に2.0として正式公開予定



# GENESISの使い方

チュートリアル、サンプルインプットを用意(バージョンに合わせて更新)



<https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/>

- PDBからの入力ファイルの作成
- システムのセットアップ、シミュレーション、解析と一連の計算の流れを確認できます

- : Suitable for laptop or small desktop machine (atdyn, less than 4 CPU cores)
- : Suitable for workstation or GPU machine (spdyn, ~16 CPU cores)
- : Suitable for cluster machine or super-computer (spdyn, more than 64 CPU cores)

## Level 1: Basic tutorials レベル別 (Level1-3)

1. Getting started
  - 1.1 Installation of GENESIS for Tutorials
  - 1.2 Let's take a quick look at the source code of GENESIS
2. Preparation of the input files for GENESIS
  - 2.1 3D structure of biological molecules
  - 2.2 Preparation of the force field parameters
  - 2.3 Building the initial structure of the target system
3. MD simulations of peptides and proteins with the all-atom CHARMM force field
  - 3.1 Ala-dipeptide in the gas-phase ← 必要な計算環境
  - 3.2 (Ala)<sub>3</sub> in water
  - 3.3 Protein G in NaCl solution
4. Analysis of the MD trajectories
  - 4.1 Analysis of DCD file by the user's Fortran programming
  - 4.2 Analysis of the statistics of the trajectory data



# MDのセットアップ

## チュートリアル

<https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/>

### Level 1: Basic tutorials

#### 1. Getting started

- 1.1 Installation of GENESIS for Tutorials
- 1.2 Let's take a quick look at the source code of GENESIS

#### 2. Preparation of the input files for GENESIS

- 2.1 3D structure of biological molecules
- 2.2 Preparation of the force field parameters
- 2.3 Building the initial structure of the target system

#### 3. MD simulations of peptides and proteins with the all-atom CHARMM force field

- 3.1 Ala-dipeptide in the gas-phase
- 3.2 (Ala)<sub>3</sub> in water
- 3.3 Protein G in NaCl solution

#### 4. Analysis of the MD trajectories

- 4.1 Analysis of DCD file by the user's Fortran programming
- 4.2 Analysis of the statistics of the trajectory data

### Level 2: Standard MD tutorials

#### • Preparation of the input files for various systems

- 5.1 Creating input files for the CHARMM force field
- 5.2 Creating input files for the AMBER force field

#### • MD simulations of various biomolecules with the all-atom model

- 6.1 POPC lipid bilayers with the CHARMM force field

vmd<sup>1</sup>によるセットアップ  
(CHARMM力場:水溶性タンパク質)

vmd<sup>1</sup>によるセットアップ  
(CHARMM力場:水溶性タンパク質,DNA-protein)

AmberTools<sup>2</sup>によるセットアップ  
(AMBER力場:水溶性タンパク質,DNA-protein)

Charmm-gui<sup>3</sup>によるセットアップ  
(CHARMM力場:膜分子)

以下のソフトのライセンス・利用条件は各自で  
ご確認ください

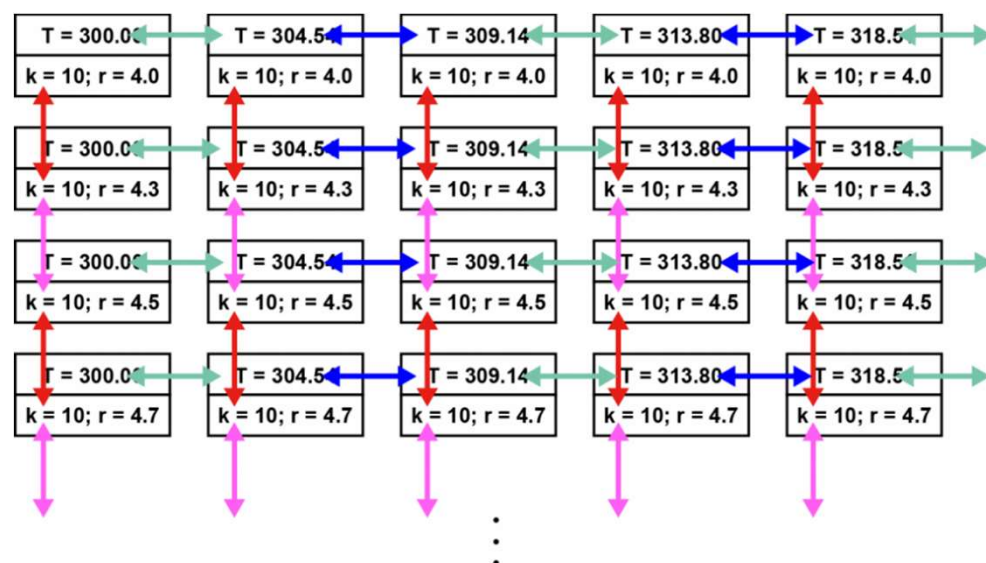
1. <https://www.ks.uiuc.edu/Research/vmd/>
2. <http://ambermd.org/AmberTools.php>
3. <https://www.charmm-gui.org/>



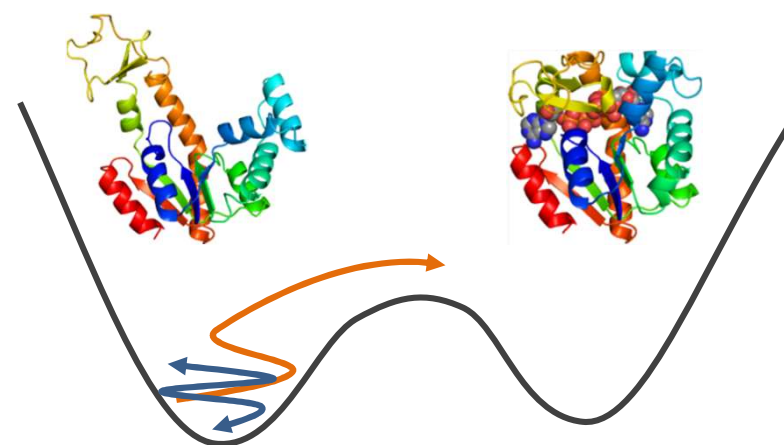
# GENESIS計算の例 #1

## レプリカ交換MD法によるサンプリング

Sugita & Okamoto, *Chem. Phys. Lett.* **314**, 141–151 (1999),  
Sugita *et al.*, *J. Chem. Phys.* **113**(15), 6042–6051 (2000).



2D REMD <https://www.r-ccs.riken.jp/labs/cbrt/samples/sample-of-2d-remd/>



異なる物理パラメータを持つ複数のレプリカ(=シミュレーションシステムのコピー)を同時に計算することでより効率の良いサンプリング計算を達成

Replica-exchange molecular dynamics simulation (REMD) : <https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/tutorial-10-1/>

Replica-exchange umbrella sampling (REUS) : <https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/tutorial-11-1/>

Replica-exchange solute tempering (gREST) : <https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/genesis-tutorial-12-1/>

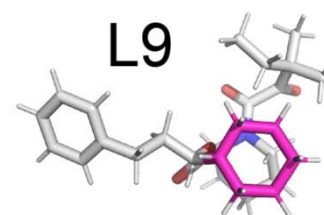
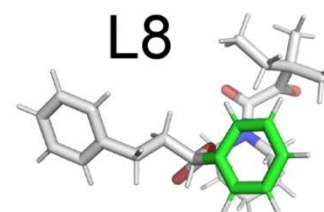
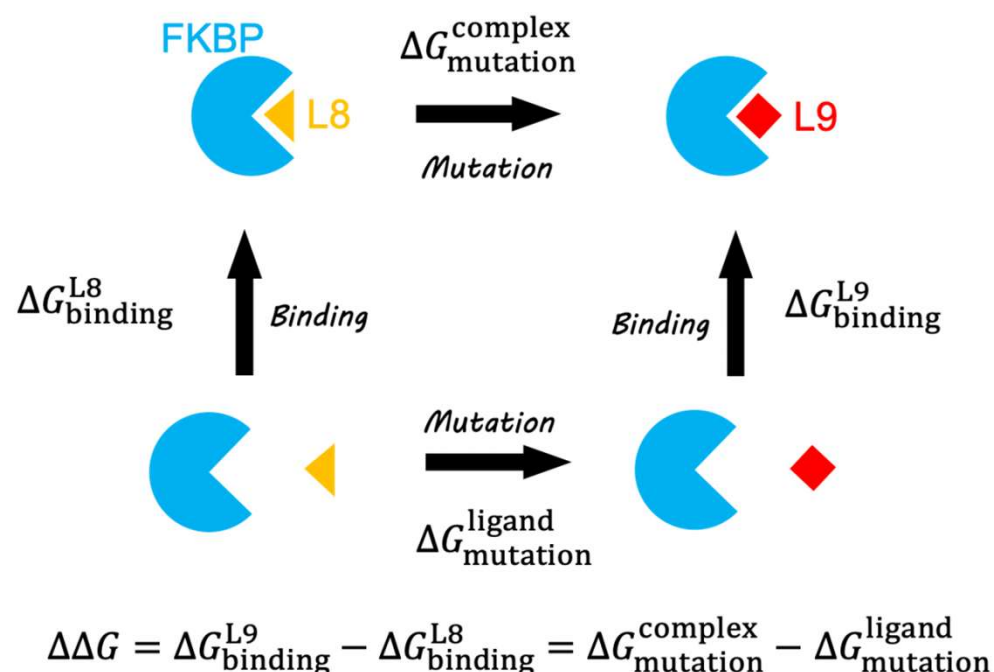


## GENESIS計算の例 #2

Free-energy perturbation によるタンパク質-リガンド結合エネルギー計算

<https://www.r-ccs.riken.jp/labs/cbrt/tutorials2019/genesis-tutorial-15-1/>

Kim et al. *J. Chem. Theory Comput.*, **16**, 7207-7218 (2020).

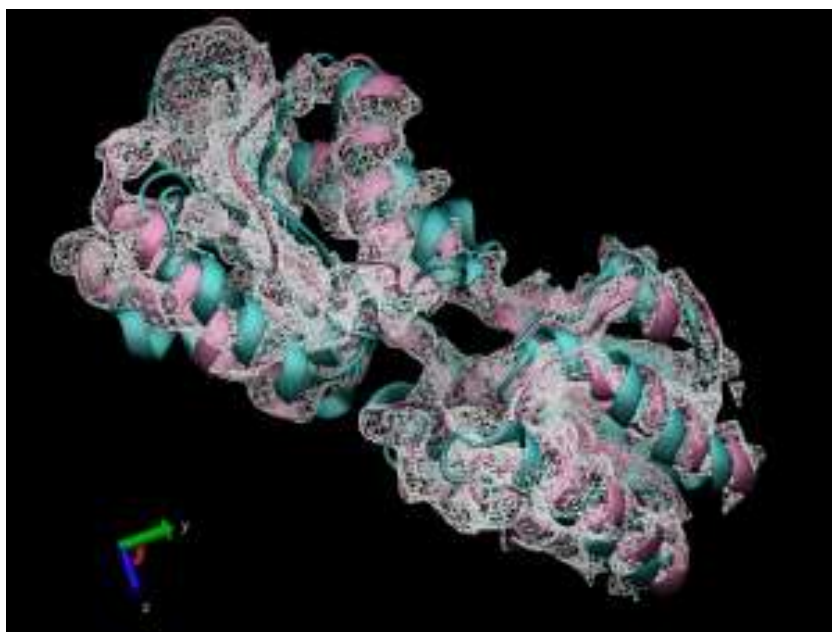




## GENESIS計算の例 #3

### Cryo-EM flexible fitting

- [GGBP using a simulated density map](#)
- [Multi-scale cryo-EM flexible fitting](#)
- [Flexible fitting refinement for \*de novo\* models](#)



GGBP using a simulated density mapより  
初期構造 (シアン)と fit後の構造 (ピンク) の比較

T. Mori *et al.*, *Structure*, **27**, 161-174.e3 (2019).  
O. Miyashita *et al.*, *J. Comput. Chem.*, **38**, 1447-1461 (2017).  
M. Kulik *et al.*, *Front. Mol. Biosci.*, **8**, 631854 (2021).  
T. Mori *et al.*, *J. Chem. Inf. Model.*, DOI: 10.1021/acs.jcim.1c00230.



## 内容

- 分子動力学法ソフトウェアGENESIS
- GENESISの高速化、性能
- GENESISでできること
- 今日のハンズオンの簡単な説明





## 今日のハンズオン

### 1. 富岳でのシミュレーション

- 富岳用に高速化されたGENESIS2.0 $\beta$ を使ってspdynでの計算を行います

### 2. レプリカ交換MD(REMD)計算

- GENESISの特徴であるREMD計算を富岳で行います



## ここまでのまとめ

- GENESISは
  - スーパーコンピュータの性能を最大限に引き出すことを目的に独自の高速化、超並列手法を持つフリーソフト
  - 2つのMDエンジンと30を超える解析ツールを持つアプリケーションスイート
  - 利用方法、サンプルは公式サイトに用意
- 今日は富岳でのGENESIS計算のハンズオンを行います