



GENESIS

Generalized-ensemble simulation system

GENESISの概要

岩橋一小林 千草

理化学研究所 計算科学研究センター

2021/09/28 GENESIS講習会 -Cygnusを用いたハンズオン-



GENESIS

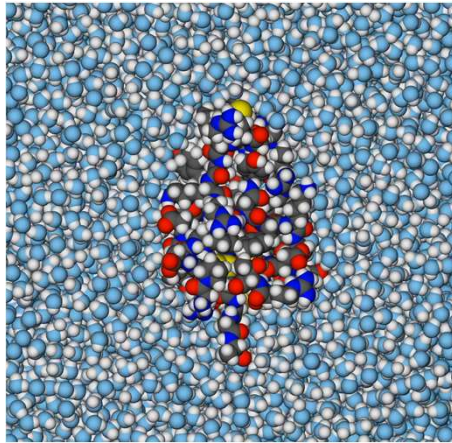
Generalized-ensemble simulation system

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 5×10^{11} 回の繰り返しが必要

ボトルネックは力の計算

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

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

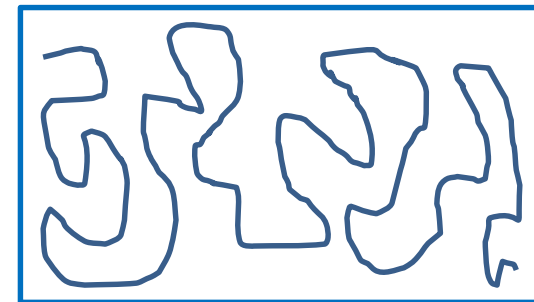
運動方程式



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

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

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)
- フリーソフト(LPGLv3)として公開中

Please visit our website

<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 (GPU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

Generalized-ensemble simulation system

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



GENESIS

Generalized-ensemble simulation system

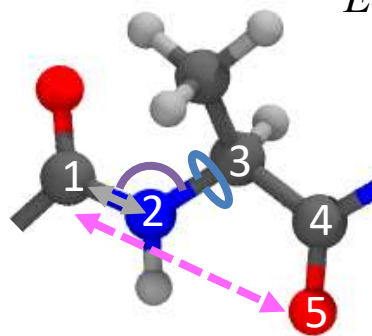
内容

- 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



水素: 薄灰
炭素: 濃灰
酸素: 赤
窒素: 青

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

結合性
 $O(N)$

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

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

非結合性
 $O(N^2)$

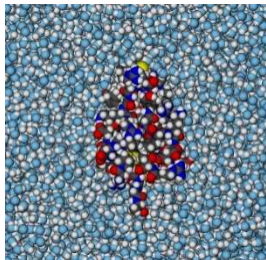
演算時間のボトルネックは非結合性相互作用計算



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

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

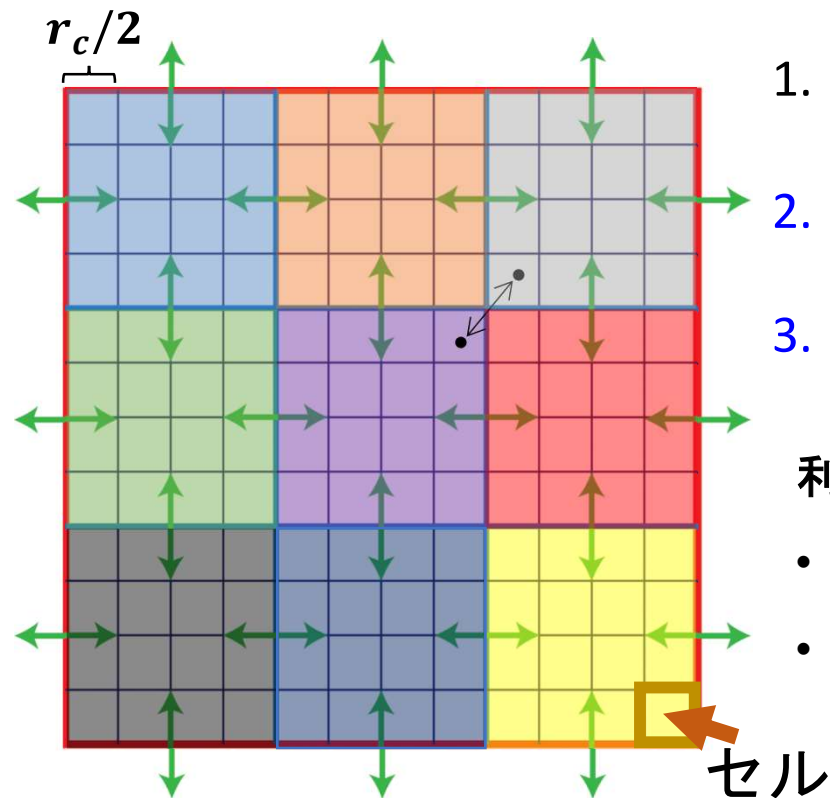
実空間, $O(CN)$
計算・メモリアクセス・1対1通信

逆空間, $O(N \log N)$
全体通信



ドメイン分割: Midpoint cell method (1)

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

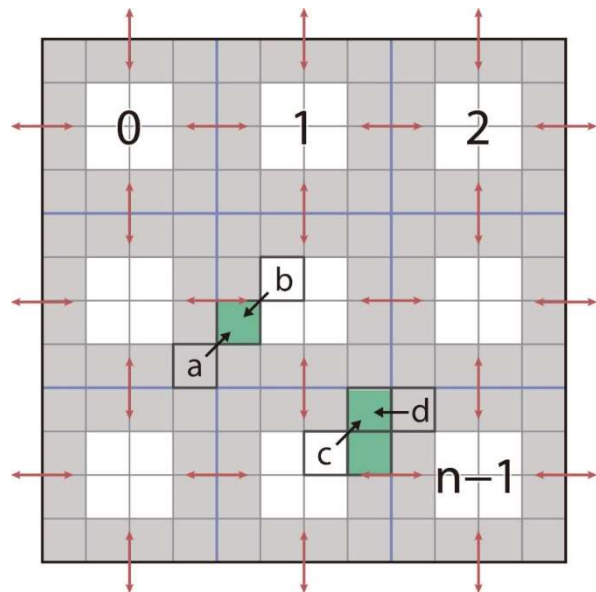
利点


- 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



ドメイン分割: Midpoint cell method (2)

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



 Subdomain assigned by MPI
(Computation is assigned to one CPU)

 Cell  Boundary cell

相互作用計算

1. 同じセル: セルを持つプロセスが計算
2. 異なるセル: セル間の中間セルを持つプロセスが計算
3. 中間セルが一意に決まらない時には、粒子数などから自動的に決定
4. それぞれのサブドメインの境界部分 (灰色) のセルのデータの通信のみで計算可能



GPUでの高速化(1)

J. Jung et al. J. Chem. Theory Comput. 12, 4947 (2016)

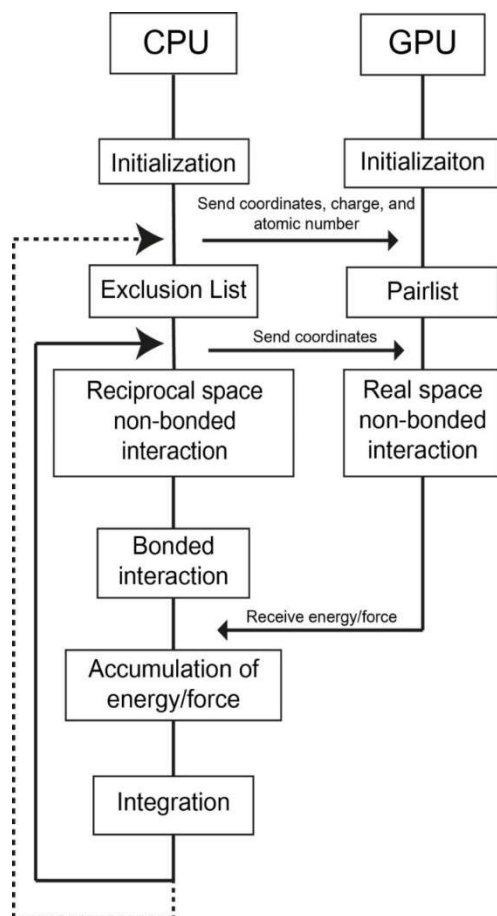
計算を2つの部分に分け、CPUとGPUで分担

演算が主な部分：GPU

- 実空間部分 (ペアリスト)

それ以外：CPU

- 逆空間部分(FFT)
- 結合性相互作用
- I/O



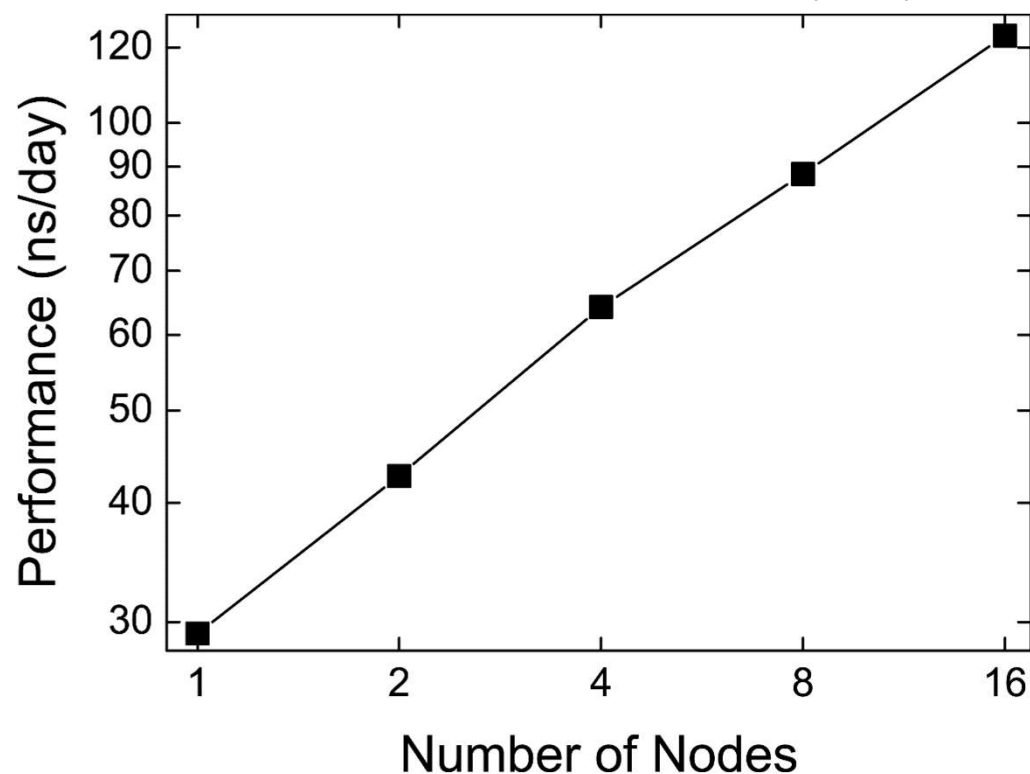


Cygnusでの性能



Dr. J. Jung

10万原子でのベンチマーク (NVT)



4 GPU 利用
高い並列性
一方で単体の計算ではGPUの性能
を生かし切れていない部分もある



GENESIS

Generalized-ensemble simulation system

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	—	○



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)

The screenshot shows the GENESIS website interface. The 'Tutorials & Samples' menu item is highlighted with a red box and a red arrow. A callout box on the right provides a detailed list of tutorial levels and topics:

- 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 Preparation of the force field parameters
 - 2.2 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-tripeptide 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 own programming
 - 4.2 Analysis of the statistics of the trajectory data

Level 2: Standard MD tutorials

5. Preparation of the input files for various systems

<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



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 Preparation of the force field parameters
 - 2.2 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-tripeptide 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 own programming
 - 4.2 Analysis of the statistics of the trajectory data

vmd¹によるセットアップ
(CHARMM力場:水溶性タンパク質)

CHARMM-GUI²によるセットアップ
(CHARMM力場:水溶性タンパク質,DNA-protein, 膜分子)

AmberTools³によるセットアップ
(AMBER力場:水溶性タンパク質,DNA-protein)

Level 2: Standard MD tutorials

5. 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
6. MD simulations of various biomolecules with the all-atom model
 - 6.1 POPC lipid bilayers with the CHARMM force field

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

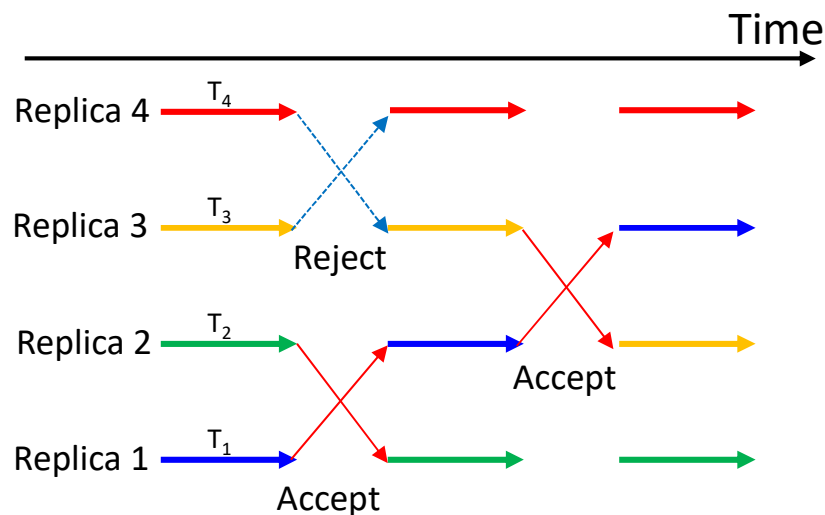
1. <https://www.ks.uiuc.edu/Research/vmd/>
2. <https://www.charmm-gui.org/>
3. <http://ambermd.org/AmberTools.php>



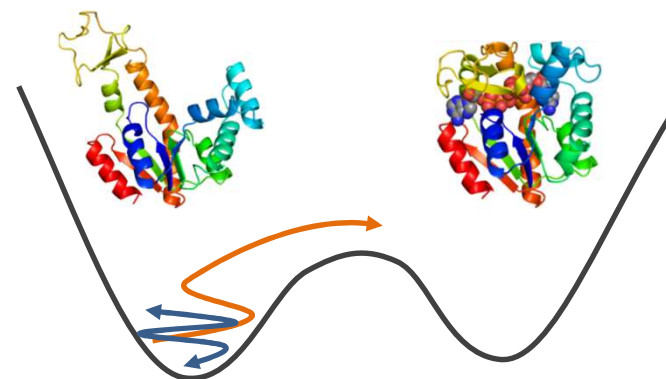
GENESIS計算の例 #1

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).



T_{1-4} : 物理パラメータ(温度等)



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/>

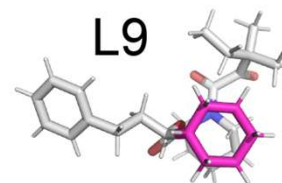
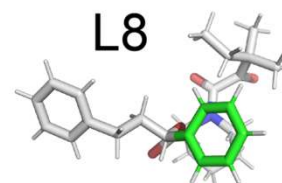
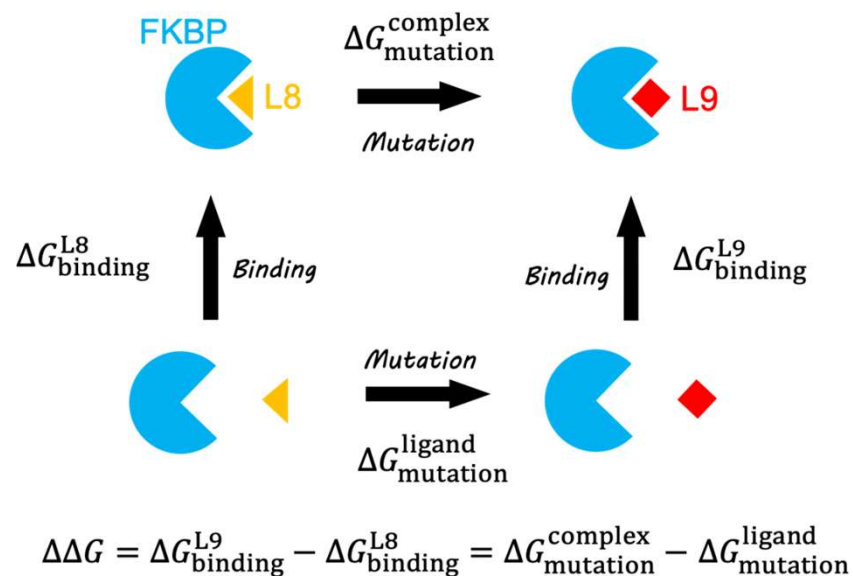


GENESIS計算の例 #2

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).

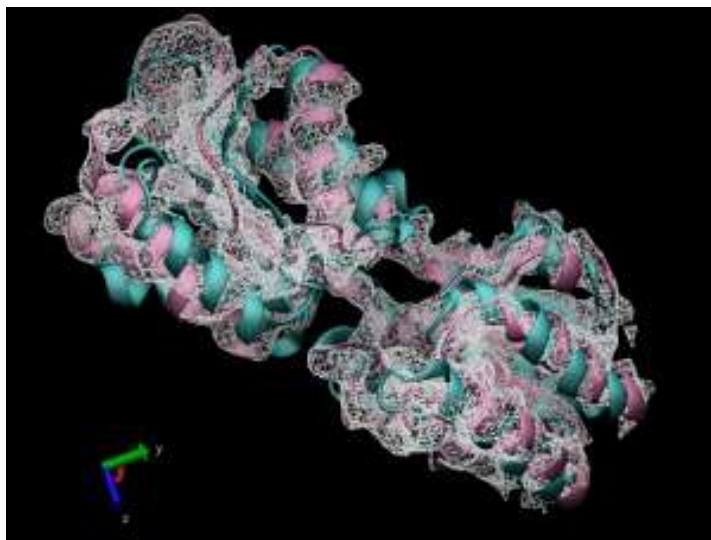




GENESIS計算の例 #3

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.



GENESIS

Generalized-ensemble simulation system

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
- 今日の手ズオンの簡単な説明



今日のハンズオン

1. GPUでの高速化

- GENESIS, spdynが持つ高速計算方法を使い、GPU計算を高速化する方法を学びます

2. レプリカ交換MD(REMD)計算

- GENESISの特徴であるREMD計算をGPU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GENESISのGPU計算を高速化する方法についてハンズオンを行います



謝辞

- GENESIS開発

開発プロジェクトリーダー: 杉田 有治 博士 (理化学研究所)

開発者: (順不同、特に1.5.x、1.6.x)

Jaewoon Jung 博士 (理化学研究所)

八木 清 博士 (理化学研究所)

森 貴治 博士 (理化学研究所)

尾嶋 拓 博士 (理化学研究所)

伊東 真吾 博士 (理化学研究所)

神谷 基司 博士 (理化学研究所、現・分子科学研究所)

松永 康佑 博士 (理化学研究所、現・埼玉大学)

松岳 大輔 博士 (理化学研究所、現・RIST)

優 乙石 博士 (理化学研究所、現・前橋工科大学)