

2016年度 ポスト「京」重点課題1

# GENESIS講習会(講義)

理化学研究所 生命システム研究センター  
分子機能シミュレーション研究チーム

李秀栄、尾嶋拓、新津藍、杉田有治

(資料提供: 小林千草)

2016年9月8日 FOCUS(神戸)

# 今日の内容

## ■ 13:30 – 13:50 講義

- 生体分子の分子動力学計算
- GENESISの特徴

## ■ 休憩(10分)

## ■ 14:00 – 16:00 実習

- GENESISのコンパイル
- 実行例(1) 分子動力学法計算と解析
- 実行例(2) レプリカ交換分子動力学計算と解析

# 今日の内容(改定)

## ■ 13:30 – 13:45 講義

- 生体分子の分子動力学計算
- GENESISの特徴

## ■ 13:45 – 14:45 実習(1)

- GENESISのコンパイル
- 分子動力学法計算と解析

## ■ 休憩(15分)

## ■ 15:00 – 16:00 実習(2)

- レプリカ交換分子動力学計算と解析

# 分子動力学法(Molecular Dynamics; MD)

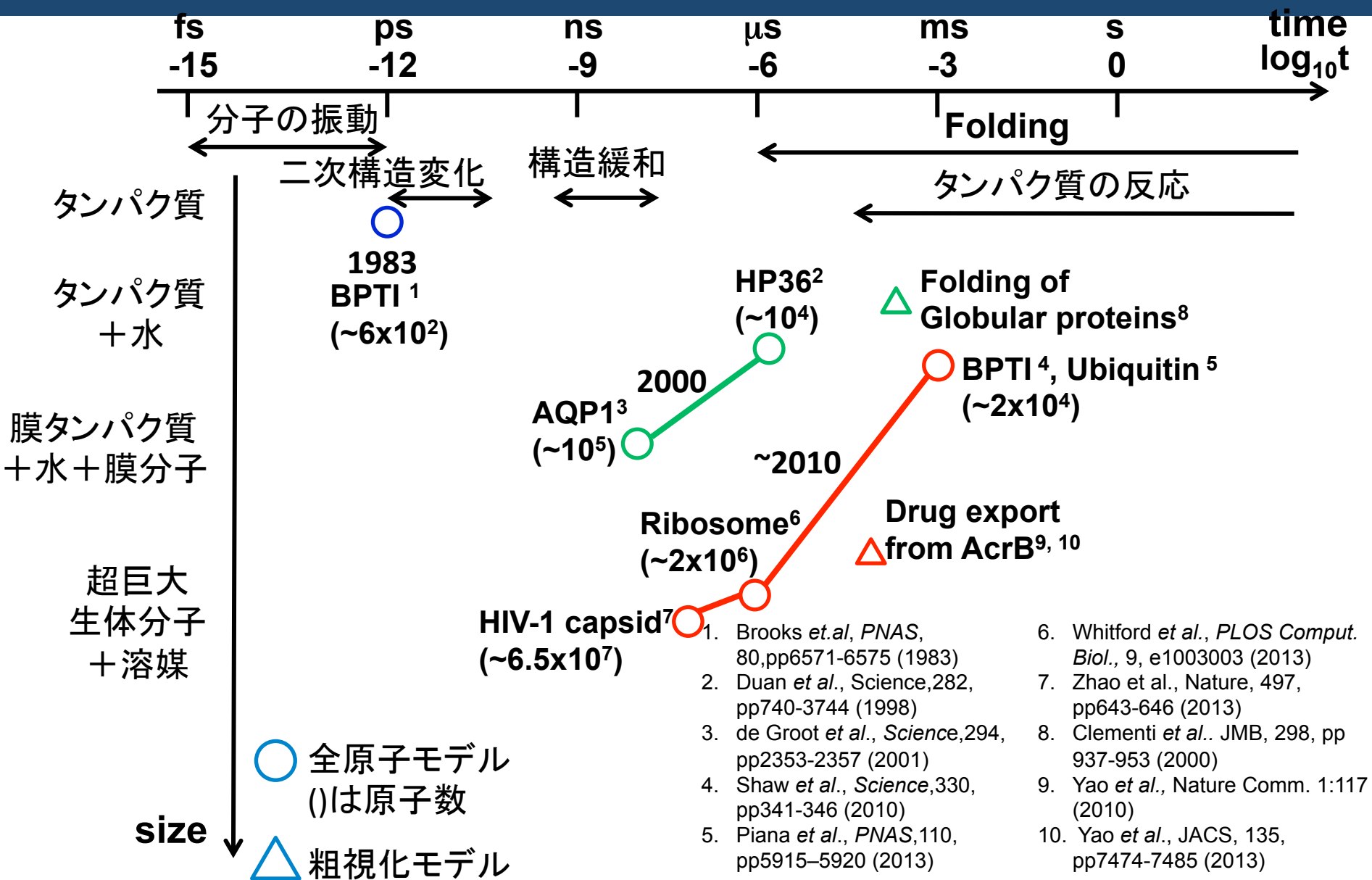
- 粒子間の相互作用力を計算し、ニュートンの運動方程式を解く

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

相互作用エネルギー  
粒子の座標

- 粒子の運動から、構造の安定性、構造ゆらぎ、更にアンサンブルを計算し、自由エネルギー面(結合エネルギーなど)を解析できる
- しかし、MD計算はとても時間がかかる...
  - タンパク質のダイナミクスをミリ秒計算するには、ニュートンの運動方程式を $10^{12}$ 回計算する必要がある
  - 例えば1stepに1msかかるなら、約31.7年必要！

# タンパク質のシミュレーション



# 重点課題1: 創薬計算の高精度化・高速化

- 基質結合とタンパク質構造変化の両方を取り込む
- スパコンを活用した高並列スクリーニング

創薬計算に主に用いられているMDソフトと力場

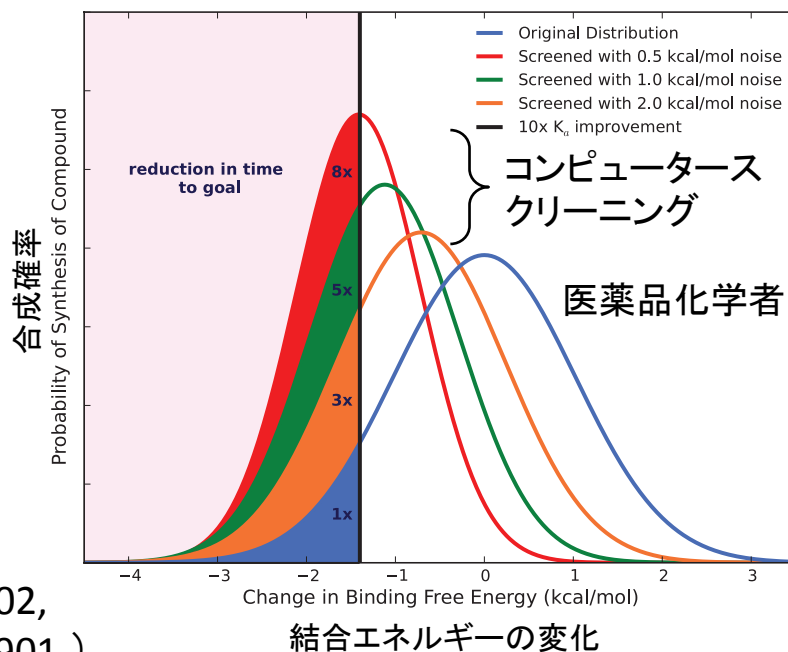
TABLE 1

## Main software and force-fields used in MD.

Developing group	MD suite	Force field
Groningen (NL)	GROMACS [1]	GROMOS [16]
Harvard (USA)	CHARMM [3]	CHARMM [15]
San Francisco (USA)	AMBER [4]	AMBER [13]
Zurich (CH)	GROMOS [22]	GROMOS [16]
Purdue/Yale University (USA)	–	OPLS [14]
Urbana-Champaign (USA)	NAMD [2]	–
Barcelona (ES)	ACEMD [23]	–
New York (USA)	DESMOND	–

(Mortier, J. et al. *Drug Discovery Today*, **2015**, 20:686-702,  
Mobley, D. & Klimovich, P. V. *J. Chem. Phys.* **2012**, 137:230901.)

コンピュータスクリーニングの導入で  
合成効率はどう変わる？





# GENESIS

Generalized-ensemble simulation system



理化学研究所 計算科学研究機構  
粒子系生物物理研究チーム  
Computational Biophysics Research Team  
RIKEN Advanced Institute for Computational Science

- 生体系において、効率的で精度の良い自由エネルギー計算手法の開発が目的
- 高並列計算 - 京コンピュータなどでの超並列計算が可能  
(ただし、普通のPCクラスタでも動きます)
- 巨大な生体系も可能
- 全原子モデルのみでなく、粗視化モデル等異なる分子モデルへも応用できるアルゴリズムを採用
- レプリカ交換法(拡張アンサンブル法の一つ)による自由エネルギー計算も可能



# GENESIS

Generalized-ensemble simulation system



## ■ 開発チーム

計算科学研究機構(AICS), 粒子系生物物理研究チーム

プロジェクトリーダー: 杉田有治

主開発者: 鄭載運、森貴治、小林千草、松永康佑、安藤格士(理科大)、神谷基司、依田隆夫(長浜バイオ大)、Michael Feig(ミシガン州立大)

公開バージョン : Ver. 1.1.0

## ■ ホームページ

<http://www.aics.riken.jp/labs/cbrt/>

ソースコードとマニュアルは上記ページ内のリンクよりダウンロード可

## ■ 参考文献

J. Jung, T. Mori, C. Kobayashi, Y. Matsunaga, T. Yoda, M. Feig, and Y. Sugita, WIREs Comput. Mol. Sci., 5, 310-323 (2015).



# 生体分子MD計算のボトルネック

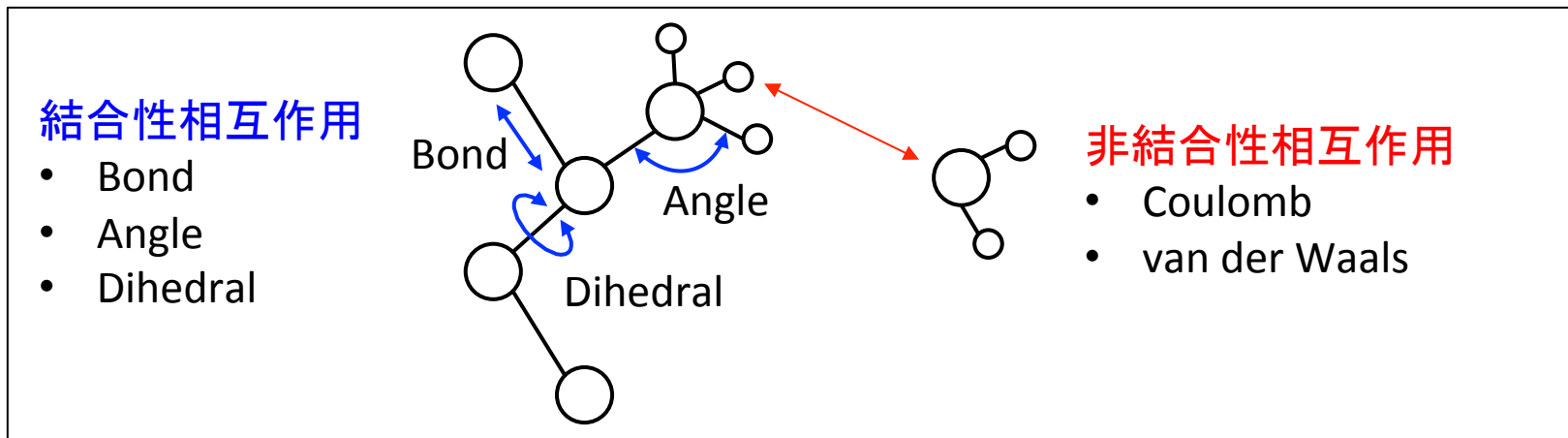
- 分子動力学法は膨大な回数のニュートン方程式を解く
- 最も時間がかかる部分は粒子間の相互作用計算の部分である

生体分子の相互作用計算は「力場(force field)」と呼ばれる経験的な関数で記述される

force fieldはおおまかに2つの部分に分けられる

- 結合性相互作用: 原子間の共有結合によるもの
- 非結合性相互作用: 長距離相互作用(電荷、van der Waals力)によるもの

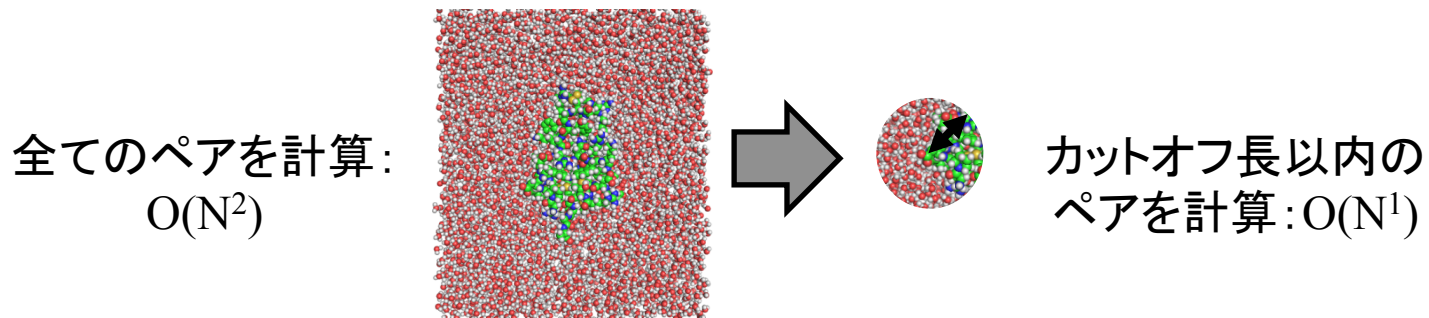
粒子数(N) に対して $O(N^2)$ で時間がかかるため、最も時間のかかる部分



→ 非結合相互作用の高速化が必要

# 非結合相互作用の高速化

1. カットオフ長を導入することで $O(N^2)$ から $O(N^1)$ に時間短縮



2. カットオフ長より長い部分のクーロンはフーリエ変換により逆空間で計算

$$U_{elec} = \underbrace{\sum_{|i-j| < R} \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erfc}(\alpha_{ij})}{r_{ij}}}_{\text{実空間(近距離)}} + \underbrace{\frac{2\pi}{V} \sum_{\mathbf{G} \neq 0} \frac{\exp(-|\mathbf{G}|^2/4\alpha^2)}{|\mathbf{G}|^2} \sum_{ij} \frac{q_i q_j}{4\pi\epsilon_0} \cos(\mathbf{G} \cdot \mathbf{r}_{ij})}_{\text{逆空間(長距離)}} - \underbrace{\sum_i \frac{q_i q_i}{4\pi\epsilon_0} \frac{\alpha}{\sqrt{\pi}}}_{\text{Self energy(定数)}}$$

3. 1,2に加えて、非結合相互作用の高速化、高並列化は近距離項、長距離項の双方で更に行われる

# GENESISの高速化・高並列化

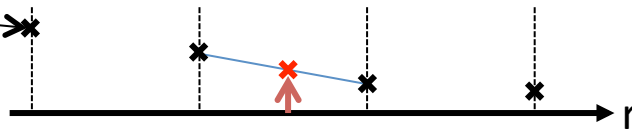
- 最も時間のかかる非結合相互作用計算を高速化するため  
GENESISでは下のアルゴリズムを新規に開発、組み込む
  - **Inverse lookup table approach** (近距離相互作用計算の高速化)  
(Jung *et al.*, *J. Comput. Chem.*, **34**:2412–2420 (2013))
  - **Midpoint cell methods** (近・長距離相互作用計算の高並列化)  
(Jung *et al.*, *J. Comput. Chem.*, **35**:1064–1072 (2014))
  - **Parallelization of FFT** (長距離相互作用計算の高並列化)  
(Jung *et al.*, *Comp. Phys. Comm.*, **200**: 57-65 (2016))
- 上の3つに加えて、トラジェクトリなどの書き出し・読み込みを  
高速化するため並列I/Oスキームを独自に持つ

# Inverse lookup table法

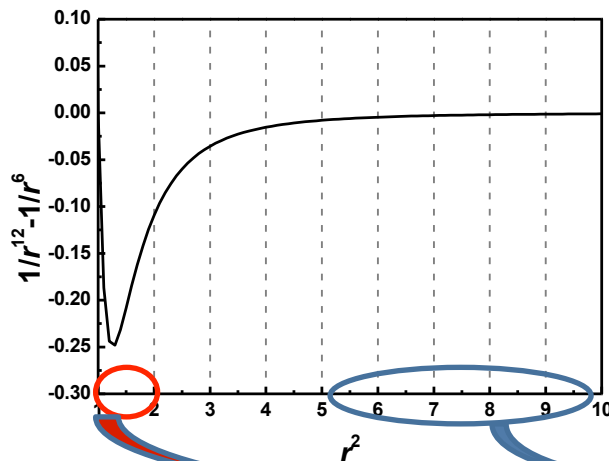
Jung et al., *J. Comput. Chem.*, **34**:2412–2420, (2013)

- 近距離相互作用は距離の関数であるため、カットオフ長までの距離の代表点での相互作用の値を計算、メモリ(table)に記憶しておく。

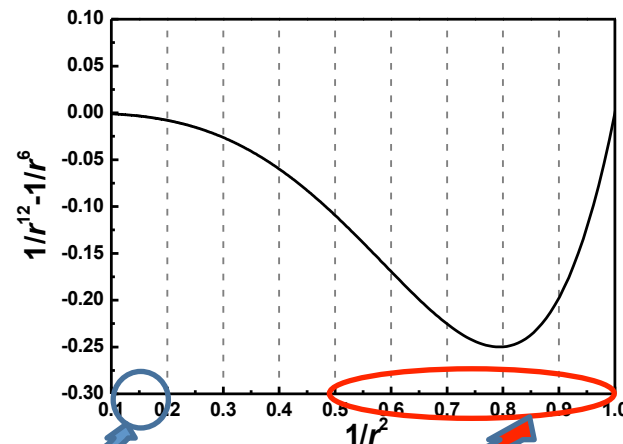
メモリに記録されて  
いる相互作用値



従来の方法



GENESISの方法



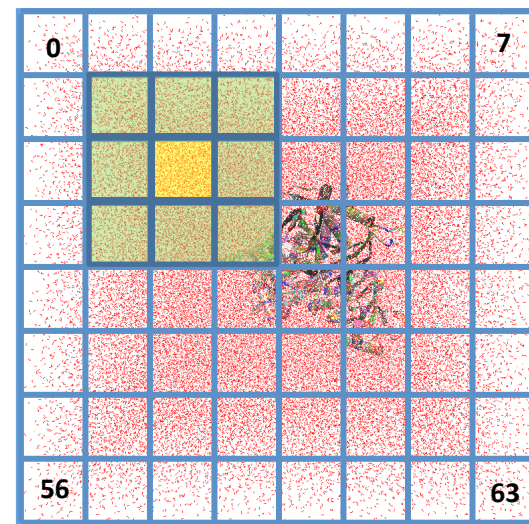
- 従来の方法では $r^2$ の線形関数、3次関数で内挿していた物を、GENESISでは $1/r^2$ の線形関数で内挿し、高速で精度の良い計算を可能とする

# Midpoint cell 法

Jung et al., *J. Comput. Chem.*, **35**:1064–1072, (2014)

## ■ ドメイン分割法 (ほとんどのMDプログラムで採用)

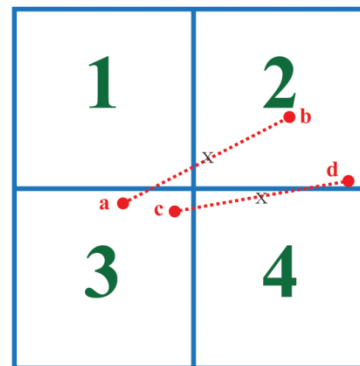
- 系全体をカットオフ長より長い一辺をもつセルで区切る
- エネルギー計算は隣接のboxのみを考えればよい (黄色のセル内の原子の相互作用は黄と緑のセル内の原子のみをカウントすればよい)
- 通信回数が減少されるため、並列度が良くなる
- ドメイン分割法で異なるセル間の相互作用計算をどのCPUコアに割り振るかが並列度に重要な問題になる



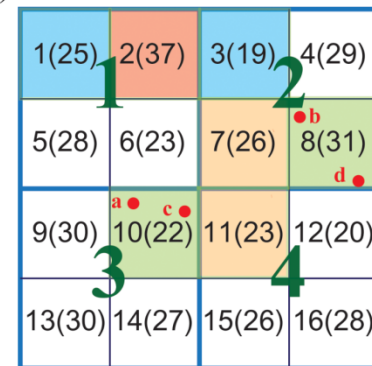
## ■ Midpoint cell法

- 従来のmidpoint法では、2つの原子の間地点のセルを受け持つコアが計算する
- それぞれの原子が存在する二つのセル (8, 10とする) の中間のセル (7, 11のどちらか) を受け持つコアが計算を担当する
- 通信回数が減少されるため、並列度が良くなる

(a) Midpoint法<sup>1</sup>



(b) Midpoint cell法



<sup>1</sup> KJ. Bowers et al, JCP 124, 184109 (2006)

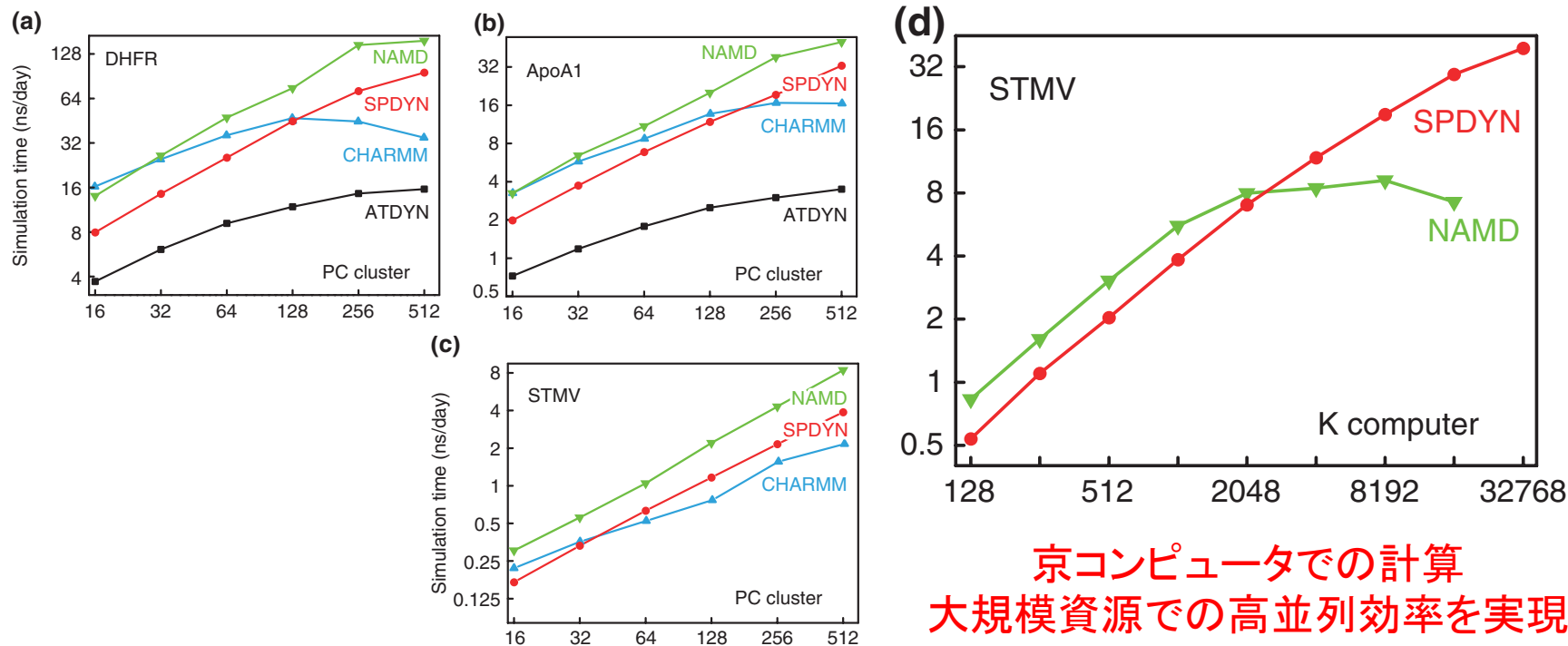
# ベンチマークデータ

## ■ ベンチマーク詳細

参照: <http://www.aics.riken.jp/labs/cbrt/benchmark/>

J. Jung, T. Mori, C. Kobayashi, Y. Matsunaga, T. Yoda, M. Feig, and Y. Sugita, WIREs Comput. Mol. Sci., 5, 310-323 (2015).

## ■ 例



# GENESISの内容物(1)

## ■ MDプログラム

### – **ATDYN** (ATomic decomposition DYNamics simulator):

- atomic decompositionを使用
- 粗視化モデル(C $\alpha$ -GO, all atom GO)が計算可能
- わかり易いコードでユーザによる開発可能

### – **SPDYN** (SPatial decomposition DYNamics simulator):

- domain decompositionを使用
- 超並列・高速なルーチン (Midpoint cell/3次元分割FFT/並列 I/O)

 今回の実習ではこちらを使用

特徴	ATDYN	SPDYN
システムの分割法	原子分割	ドメイン分割
New lookup table	○	○
レプリカ交換法	○	○
粗視化モデル	○	×
3次元分割FFT	×	○
並列I/O	×	○
r-RESPA (1.1.0)	×	○

# GENESISの内容物(2)

## ■ ATDYN/SPDYN共通

- 最適化
  - ✧ Steepest Decent法
- Integrator
  - ✧ Leapfrog
  - ✧ Velocity Verlet
- アンサンブル
  - ✧ NVE
  - ✧ NVT
    - Langevin
    - Bussi
    - Berendsen
  - ✧ NPT
    - Langevin Piston
    - Bussi
    - (Isotropy of Simulation box:

Isotropic, Semi-iso, An-iso,  
XY-fixed)

- 拘束計算(constraint)
  - ✧ SHAKE (Leapfrog)
  - ✧ RATTLE (Velocity Verlet)
  - ✧ SETTLE
- FFT (PME)
  - ✧ FFTE
- Restraint functions
  - ✧ Position
  - ✧ Bond
  - ✧ Angle
  - ✧ Dihedral angle
  - ✧ RMSD
- Target/Steered MD
- String method

(赤字は1.1.0の新機能)



# GENESISの内容物(3)

## ■ その他解析ツール

- トラジェクトリ変換(trjcnv)
  - ✧ crd\_convert
  - ✧ pcrd\_convert(並列I/O用計算用)
  - ✧ remd\_convert(REMD計算用)
- トラジェクトリ解析(trjana)
  - ✧ mbar\_analysis
  - ✧ pathcv\_analysis
  - ✧ pmf\_analysis
  - ✧ gval\_analysis
  - ✧ rmsd\_analysis(RMSD、RMSF)
  - ✧ trj\_analysis(距離、角度など)
  - ✧ wham\_analysis
- リスタートファイル変換(rstcnv)
  - ✧ rst\_convert
  - ✧ rst\_upgrade
- PCA解析(pcaana)
  - ✧ avecrd\_analysis
  - ✧ eigmat\_analysis
  - ✧ flccrd\_analysis
  - ✧ prjcrd\_analysis
- 反応経路探索関連(rpath\_generator)
- 解析共通(libana)