

AICS公開ソフト講習会 第4回 「GENESIS」講義

理研、計算科学研究機構(AICS)

粒子系生物物理研究チーム

2014/06/17

今日の内容

- 13:30 – 14:00 講義
 - 生体分子の分子動力学法
 - GENESISの特徴
- 14:00 – 14:10 休憩
- 14:10 – 16:00 実習
 - GENESISのコンパイル
 - 実行例(1) 分子動力学法計算
 - 実行例(2) 分割I/Oツールを用いた解析

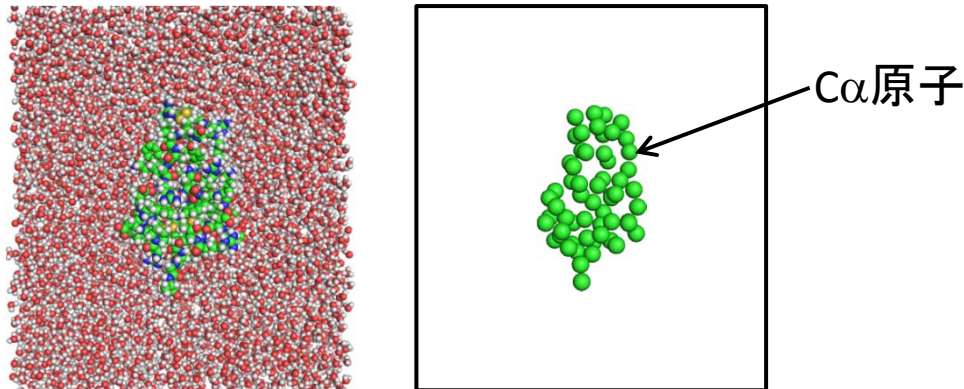
分子動力学法(Molecular Dynamics; MD)

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

$$m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} = \mathbf{F}_i(t) = - \frac{dU(\mathbf{r}^N)}{d\mathbf{r}_i} \quad \begin{array}{l} U(\mathbf{r}^N) : \text{相互作用エネルギー} \\ \mathbf{r}^N : \text{粒子の座標} \end{array}$$

粒子の大きさ(粒度)で全原子モデルと粗視化モデルに分けられる

GENESISは
どちらにも対応



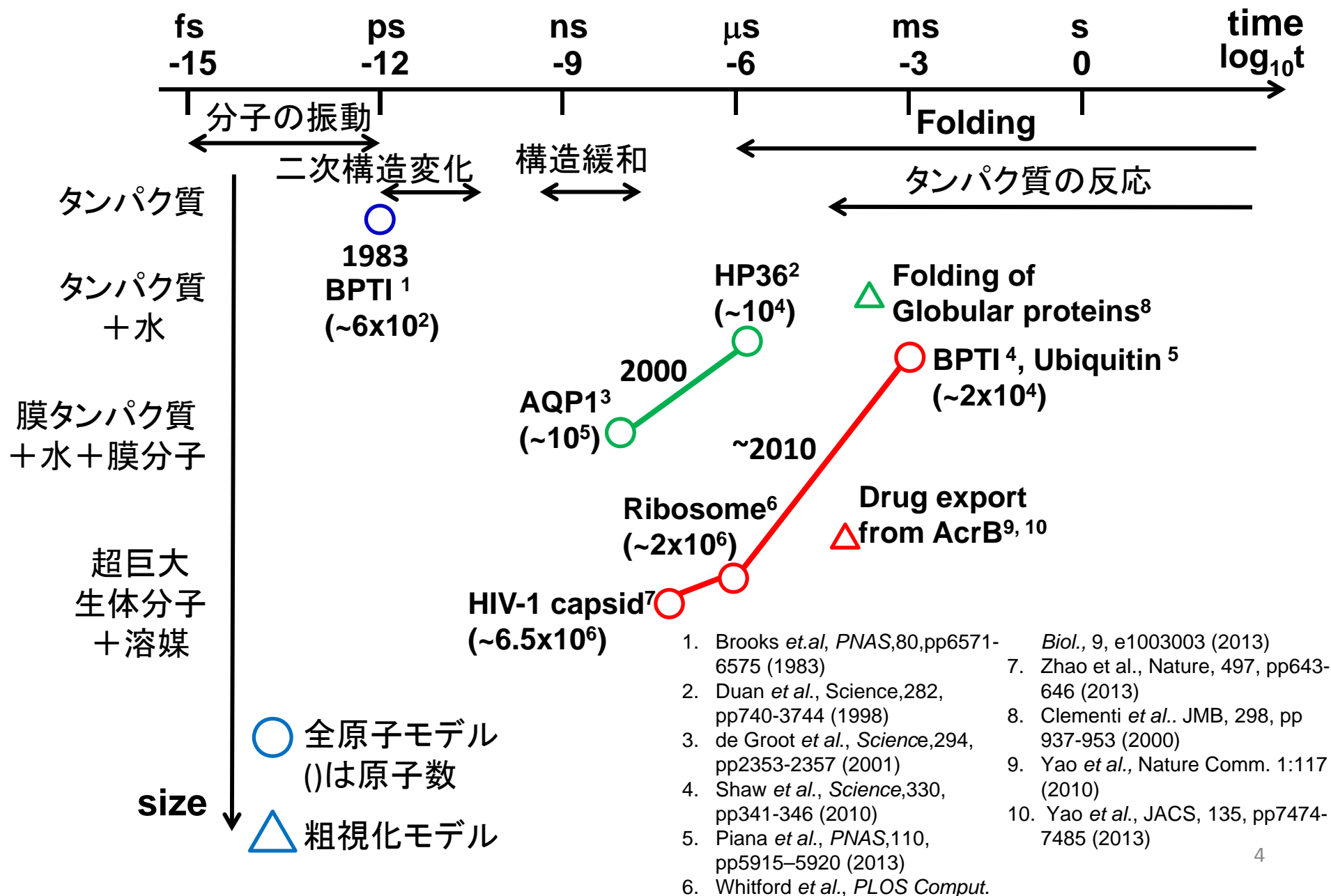
粒子の運動から、構造の安定性、構造ゆらぎ、更にアンサンブルを計算し、自由エネルギー面(結合エネルギーなど)を解析できる

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**10¹²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



Generalized Ensemble Simulation Systems (GENESIS)

1. 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
2. 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)
3. 巨大な生体系も可能
4. 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
5. 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 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

Project Leader : 杉田 有治



Main developers: Jaewoon Jung

森 貴治(和光, TMS)

岩橋 一 小林 千草

松永 康佑



GENESIS web site

<http://www.riken.jp/TMS2012/cbp/en/research/software/genesis/index.html>

生体分子の相互作用計算

1. 分子動力学法は膨大な回数のニュートン方程式を解く
2. 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
3. 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
4. force fieldはおおまかに2つの部分に分けられる -

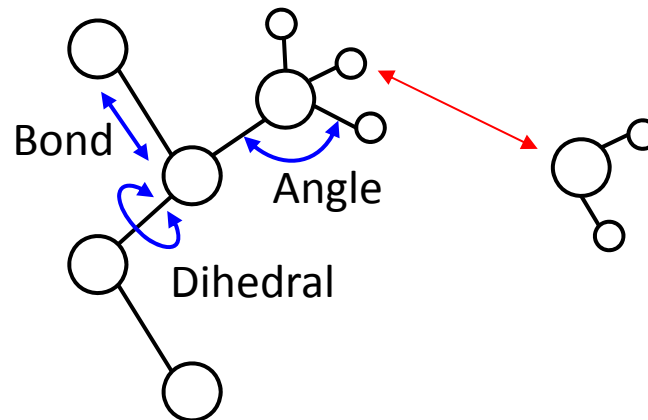
結合性相互作用: 原子間の共有結合によるもの

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である

結合性相互作用

- Bond
- Angle
- Dihedral



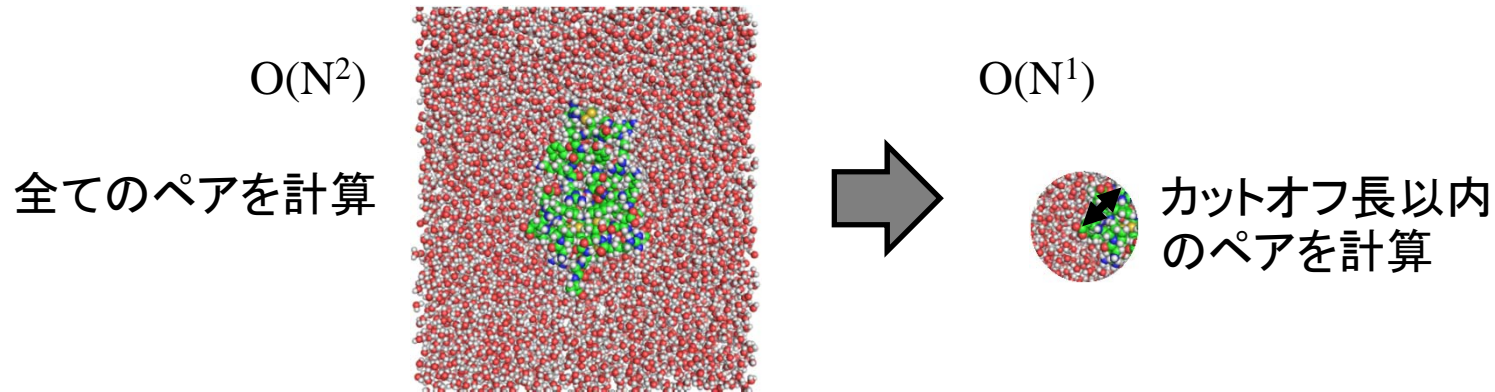
非結合性相互作用

- Coulomb
- van der Waals

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. 非結合相互作用はカットオフ長を導入することで $O(N^2)$ から $O(N^1)$ に変更できる



2. カットオフ長より長い部分のクーロン相互作用はフーリエ変換(FFT)により逆空間で計算する

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

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
(Jung *et al.*, *J. Comput. Chem.*, **34**:2412–2420, (2013)) (近距離相互作用計算の高速化)
- Midpoint cell methods
(Jung *et al.*, *J. Comput. Chem.*, **35**:1064–1072, (2014)) (近・長距離相互作用計算の高並列化)
- Parallelization of FFT
(Jung *et al.*, (in submitted)) (長距離相互作用計算の高並列化)

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

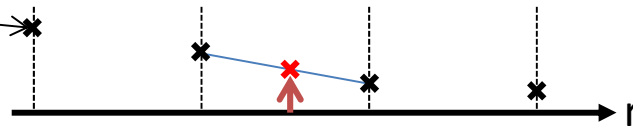
Jung et al., *J. Comput. Chem.*, **34**:2412–2420, (2013)

Lookup table法とは

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。

MDステップ中では、実際の距離に近いtableの値から内挿して求める

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

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

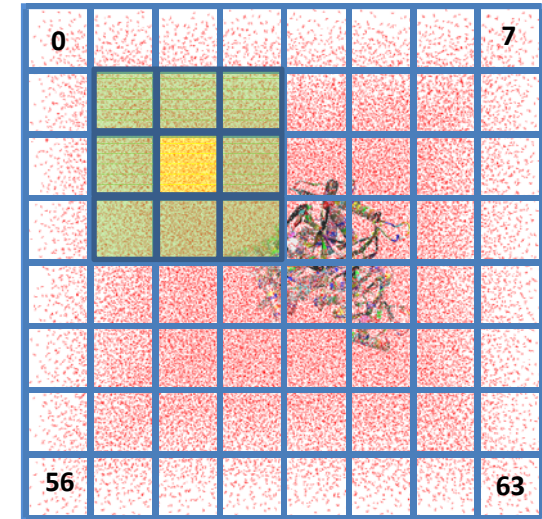
従来の方法では r^2 の線形関数、3次関数で内挿していた物を、GENESISでは $1/r^2$ の線形関数で内挿し、高速で精度の良い計算を可能とする

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つの原子の中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

Midpoint法¹

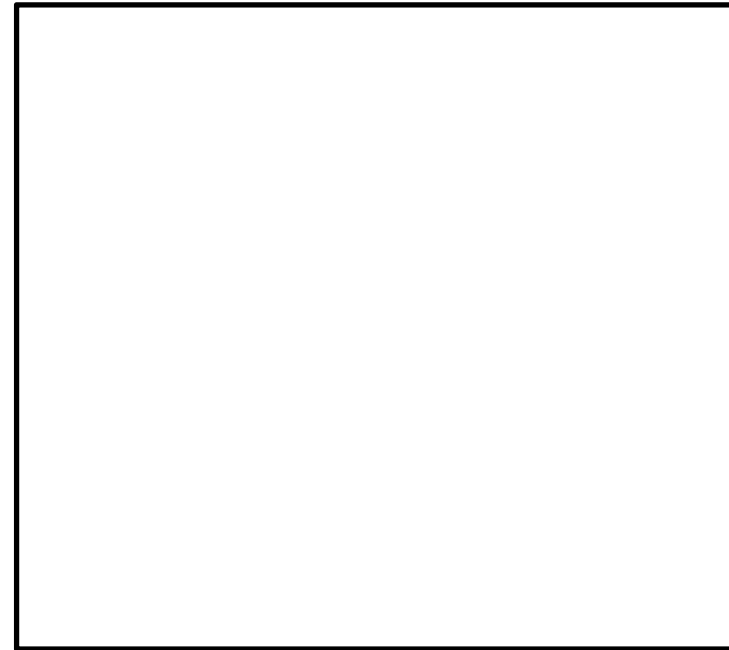
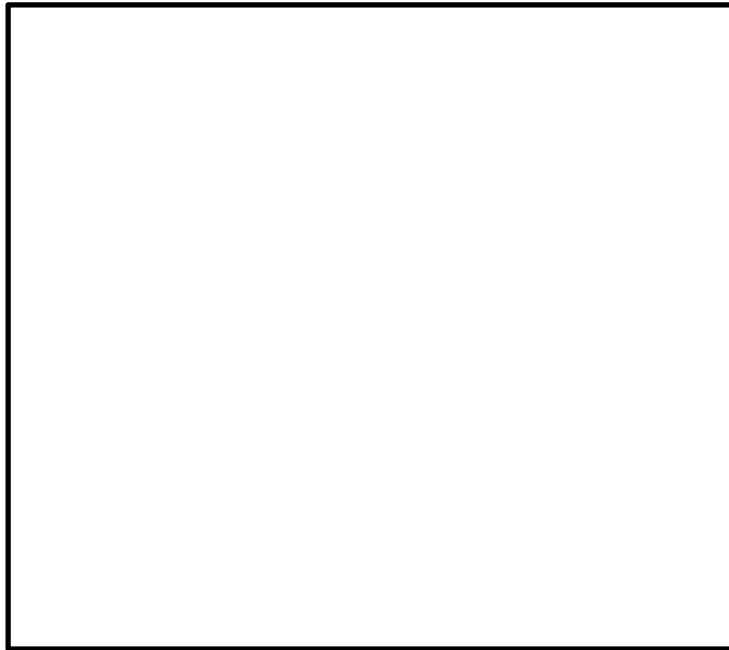
Midpoint cell法

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

System :

(time step = 2fs)

京コンピュータでの計算



(unpublished)

GENESISの内容物(1)

MDプログラム

- ATDYN (ATomic decomposition DYNamics simulator):
 - atomic decompositionを使用
 - レプリカ交換法が計算可能
 - 粗視化モデル(C α -GO, all atom GO)が計算可能
 - わかりやすいコードでユーザによる開発可能
(SPDYNでの開発の前にATDYNでテストをすることも可能)
 - 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	×	○

GENESISの内容物(2)

ATDYN/SPDYN共通に計算可能な物

- 最適化
 - Steepest Decent法
- Integrator
 - Leapfrog
 - Velocity Verlet
- アンサンブル
 - NVE
 - NVT
 - Langevin
 - Berendsen
 - NPT
 - Langevin Piston
 - (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
 - FFTW
- Restraint functions
 - Position
 - Bond
 - Angle
 - Dihedral angle

GENESISの内容物(3)

その他の主なツール

- trj_analysis
 - トラジェクトリを解析するツール
 - 距離・角度・二面角などが計算可能
- crd_convert
 - トラジェクトリを変換するツール
 - トラジェクトリから一部分のみを抜き出して、新しいトラジェクトリを作成 (例えば水分子を抜くとか、C α 原子のみにするなど)
 - 同時にfittingやRMSDなども計算可能
- prst_convert
 - GENESISのrestartファイルから並列I/O用のrestartファイルに交換
- pcrd_convert
 - 並列I/O用計算で出された複数のトラジェクトリファイルを変換するツール
 - crd_convertと同じ機能が使える

ダウンロードはGENESIS web siteから

<http://www.riken.jp/TMS2012/cbp/en/research/software/genesis/index.html>