

SS研:HPCM2001

産業技術総合研究所における計算科学:
ナノからマクロまで

産業技術総合研究所
先端情報計算センター
計算科学研究部門

寺倉清之
k-terakura@aist.go.jp

2001. 11. 2

計算科学の進歩

- 計算機の進歩
 - ベクトル計算機
 - 並列計算機
 - 演算速度: 1年半で2倍
 - 10年で100
 - 自分の研究生生活35年間の進歩は 10^7 倍
 - 広域分散処理: network computing
- アルゴリズムの進歩
 - 計算機の進歩と同様に重要

計算科学

- 理論、実験と並ぶ、研究の第3の柱
- 理論、実験と同様に、真理の探究、工学における技術の進歩への貢献を目指す共通基盤性のつよい研究手段
- ただし、計算科学はまだ若い分野であり、計算科学自身の技術向上が重要

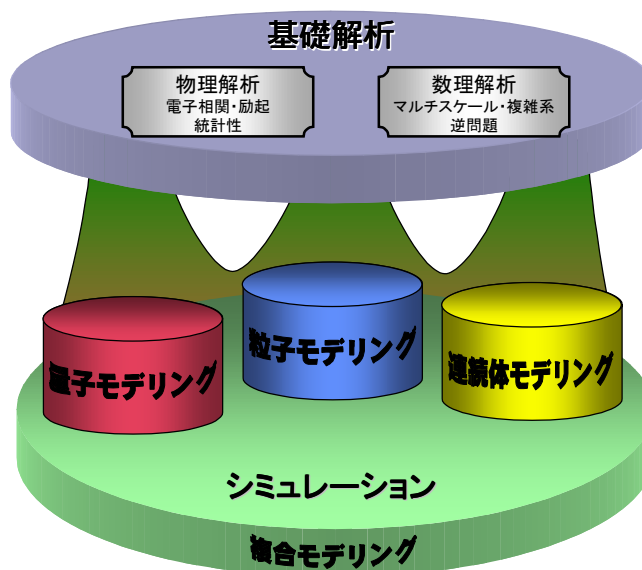
産総研における計算科学

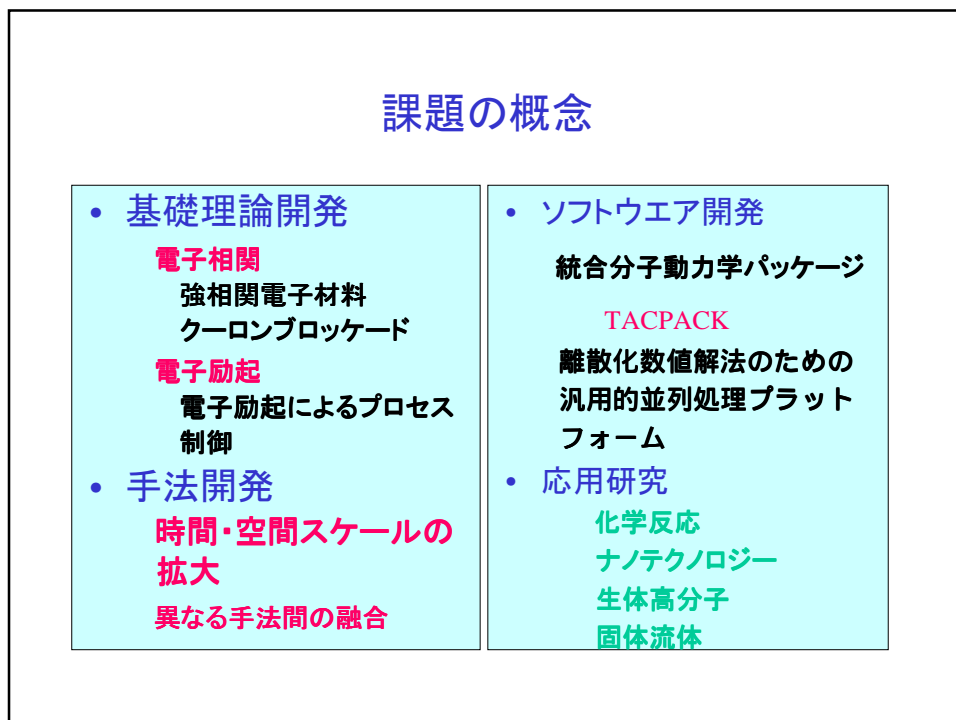
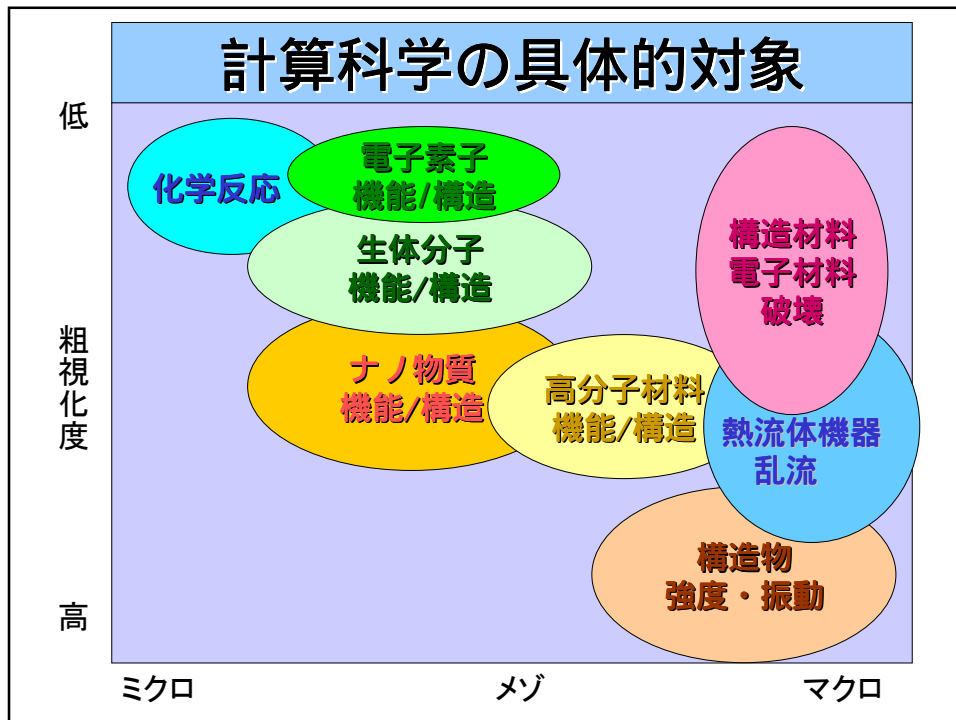
- 集中部分(専任研究員:26名)
計算科学研究部門(RICS)
理論、手法の開発・発展
応用(他ユニット、外部との協力を重視)
- 分散部分(兼任研究員:17名)
他の研究ユニットを本務
本務先研究についての適用研究
手法開発について、RICSと協力
一部はRICSに兼任
先端情報計算センターの研究グループ
JRCAT理論グループ(H14.3.31で終了)
より一般的にはCOSMOSで情報交換:再開予定

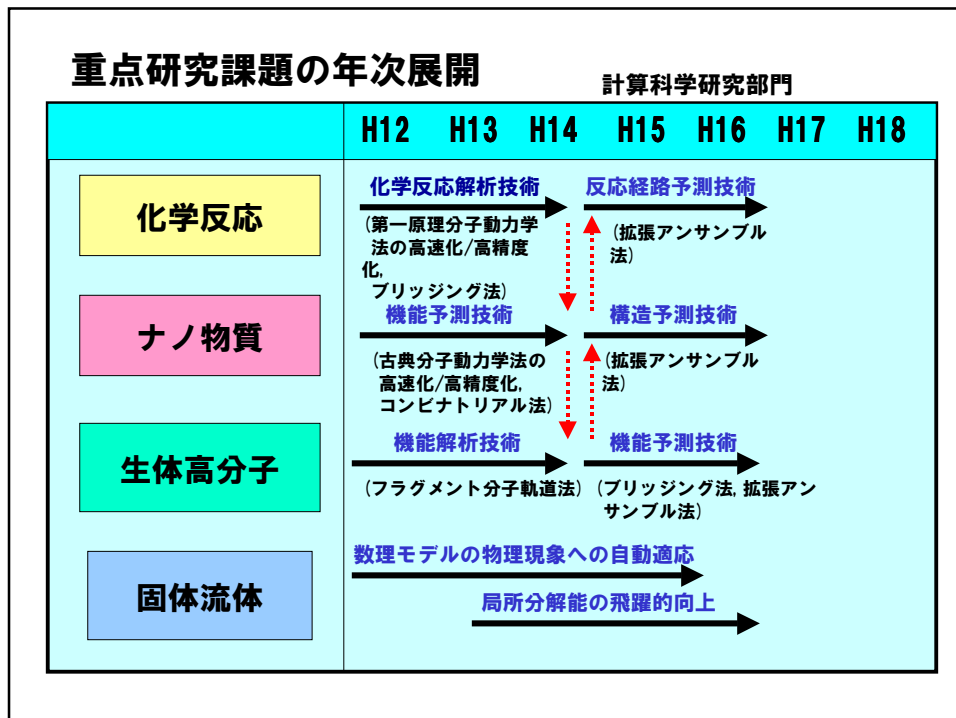
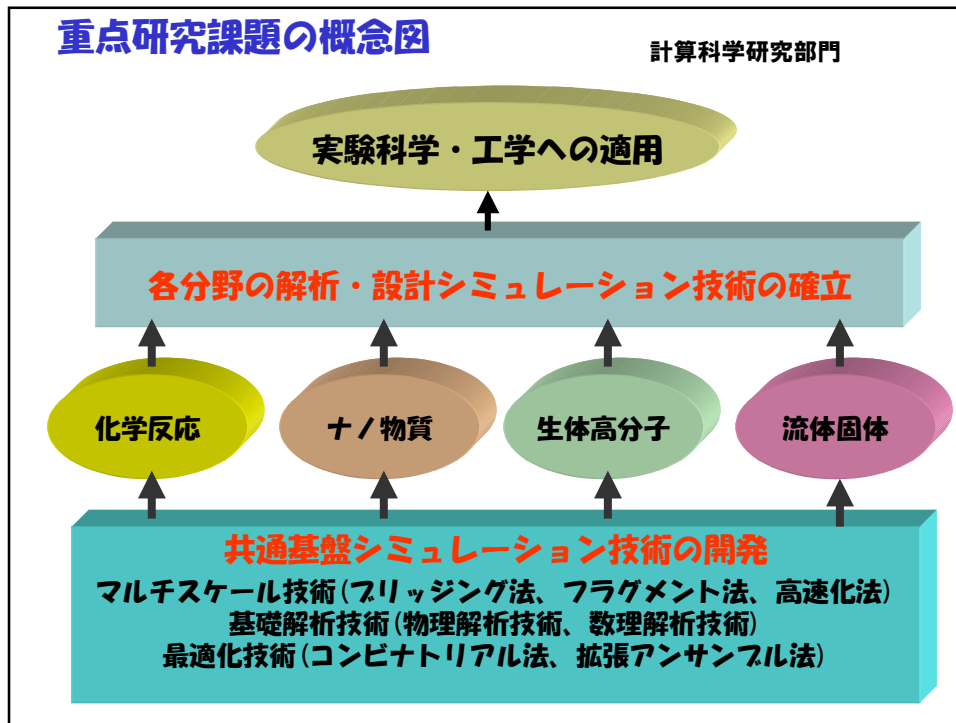
応用例

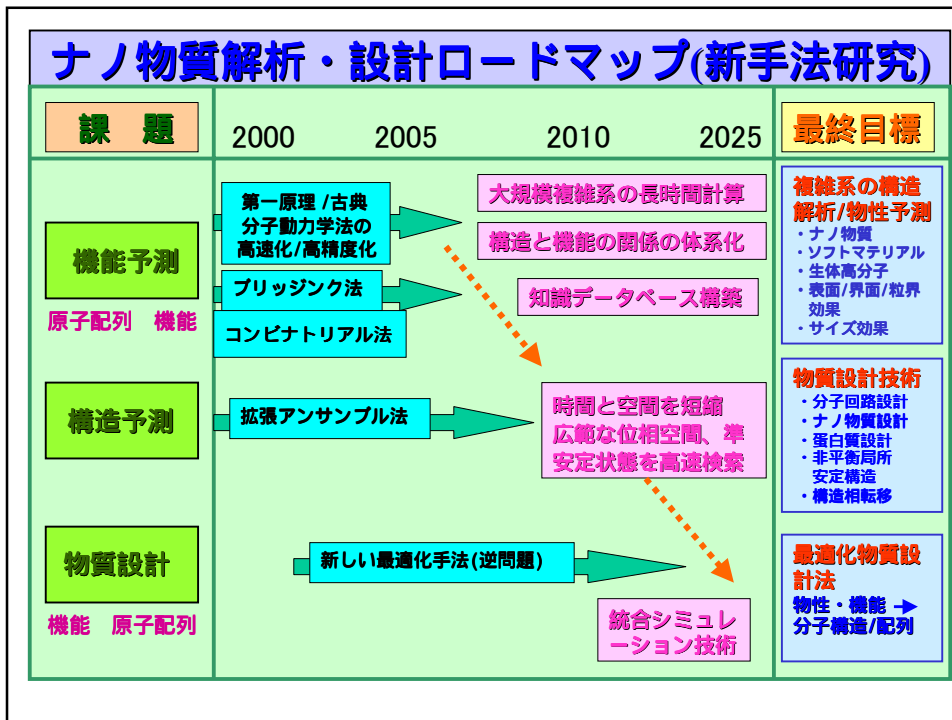
- **RICSとして設定した課題**
 - 化学反応
 - ナノテクノロジー
 - 生体高分子
 - 固体流体連成
- **産総研内プロジェクト**
 - TACCライブラリー開発**
 - ・統合分子動力学パッケージ
TACPACK
 - ・離散化数値解法のための汎用的並列処理プラットフォーム
- **TACCの情報基盤研究開発室**
 - MOエンジン(分子軌道計算専用機作成)
 - 高速分子軌道計算手法開発
- **JRCATでの成果**
 - 半導体表面・界面
 - 原子間顕微鏡の原子像形成機構
 - 触媒反応シミュレーション
 - 遷移金属酸化物

組織図





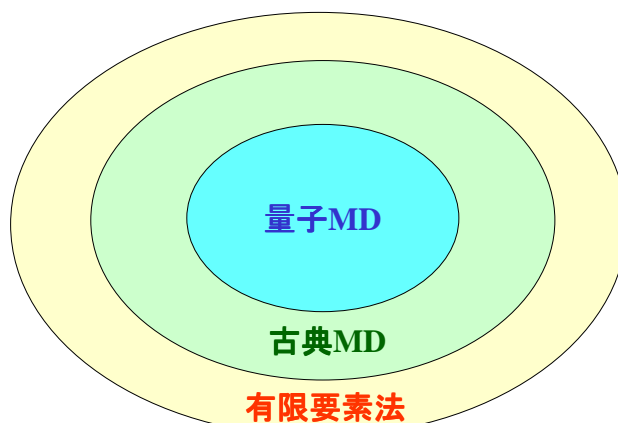




複雑系への対応

- ナノテクノロジー
 原子間距離: 0.3 nm → 3原子/nm
 ナノテクノロジーのスケール: 1~100 nm → 3~300原子
 1次元系: 3~300原子
 2次元系: 10~10万原子
 3次元系: 30~3000万原子
- バイオ
 少なくとも、数千から数万原子

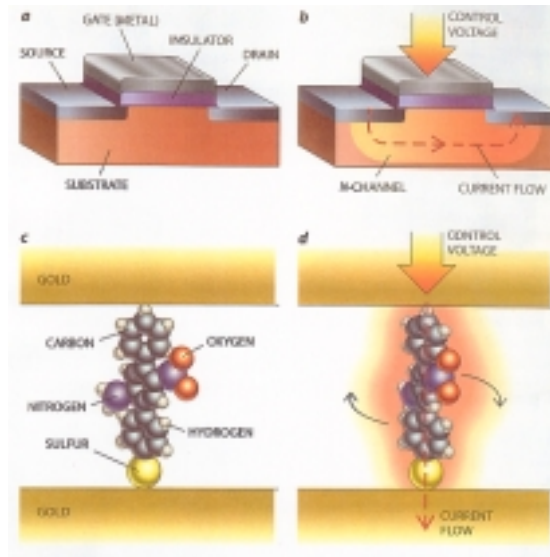
量子計算を基盤とした
マイクロからマクロまでのシームレスな接続



量子力学的計算手法(1)

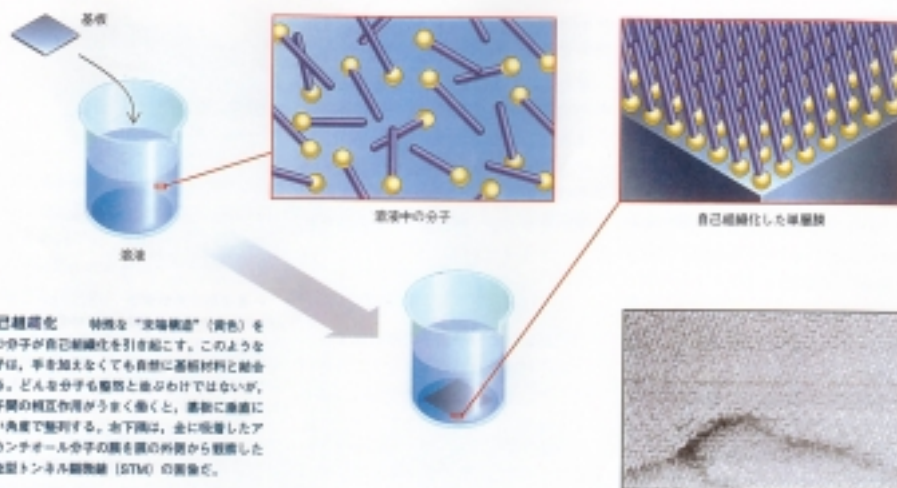
- **高信頼性標準的手法**
 - 平面波基底の第一原理分子動力学法 中心人物
 - Ultra-soft pseudopotential: 開発済み 森川
 - STATE: Simulation Tool for Atom TEchnology
 - Projector Augmented-Wave (PAW): 開発中 石橋
- **大規模系**
 - 有限要素基底の第一原理分子動力学法: 開発済み 土田
 - 超並列計算、オーダーN法
 - 非直交局在軌道基底のリカーゾン法: 開発中 尾崎
 - オーダーN法(金属系にも適用可)
 - タイトバインディング分子動力学法: 開発済み 織田
 - オーダーN法
 - Fragment Molecular Orbital 法: 開発済み 北浦
 - オーダーN法、生体高分子への適用
 - 量子・古典ハイブリッド法: 開発中 JRCAT

分子コンピュータ



M. A. Reed and J. M. Tour, Scientific American, June 2000

自己組織化を用いて分子素子を作る



日経サンエンス 2000年10月号

STATEの応用例

**第一原理分子動力学法による金表面上の
自己組織化膜の研究**

森川 良忠 (産業技術総合研究所)

林 智広 (ハイデルベルク大)

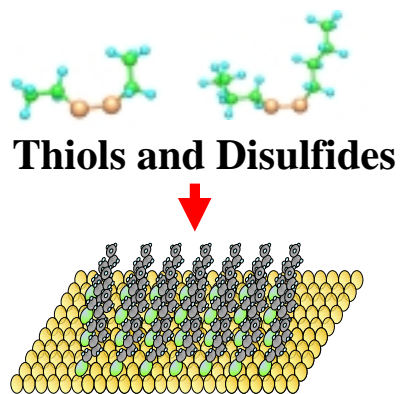
野副 尚一 (産業技術総合研究所)

STATEコードの効率

- **SR8000**
16ノードから32ノードを用いた種々のジョブの
1月平均: 約30GFLOPS
- **VPP500**
SR80001ノードとVPP500の4PEが互角

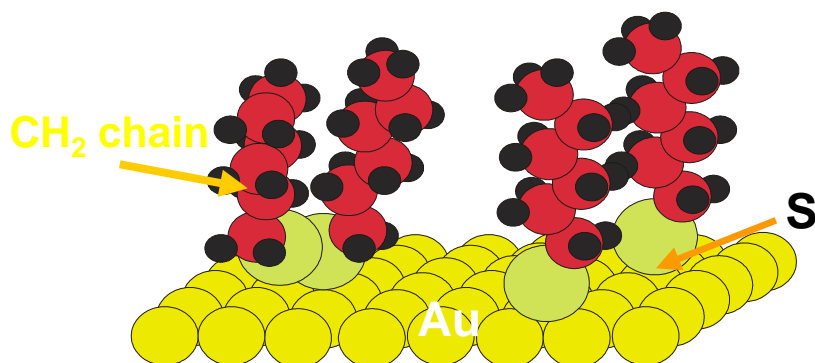
Self-Assembled Monolayers

- Corrosion inhibition
- Lubrication
- Coating
- Catalysis
- Molecular devices
- Nonlinear optical devices



Self-Assembled Monolayers

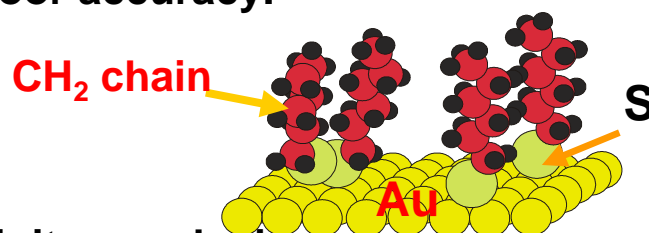
Structure Problem



Adsorption as Disulfide or Thiolate?
The S-S bond cleaved or not?
Adsorption Site?

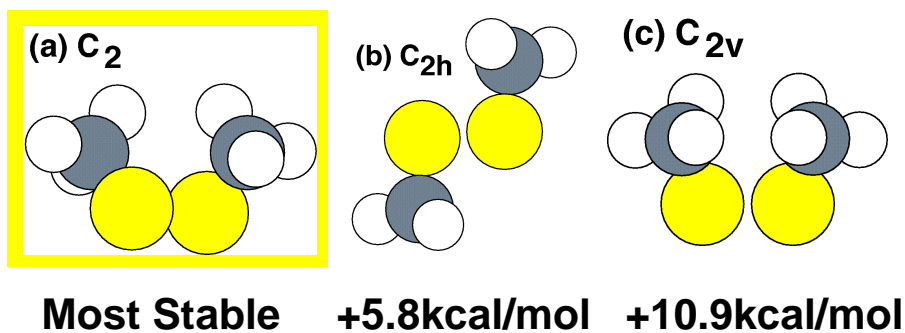
Previous Classical MD Simulations

- Interactions among CH₂ chains
→ Accurate force fields.
- Au-S interaction
→ Poor accuracy.

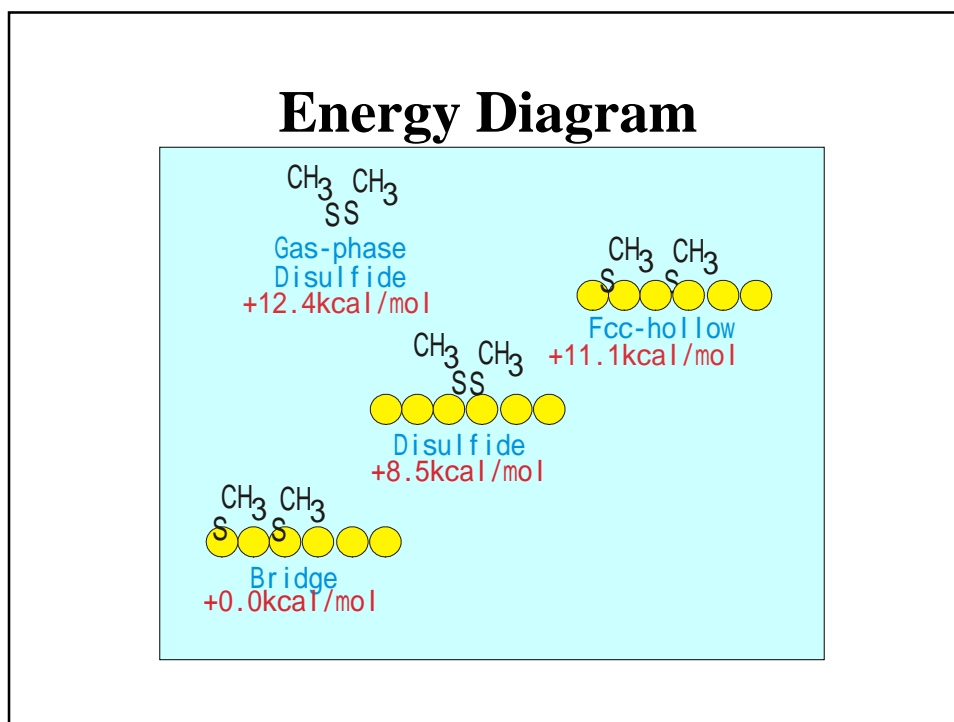
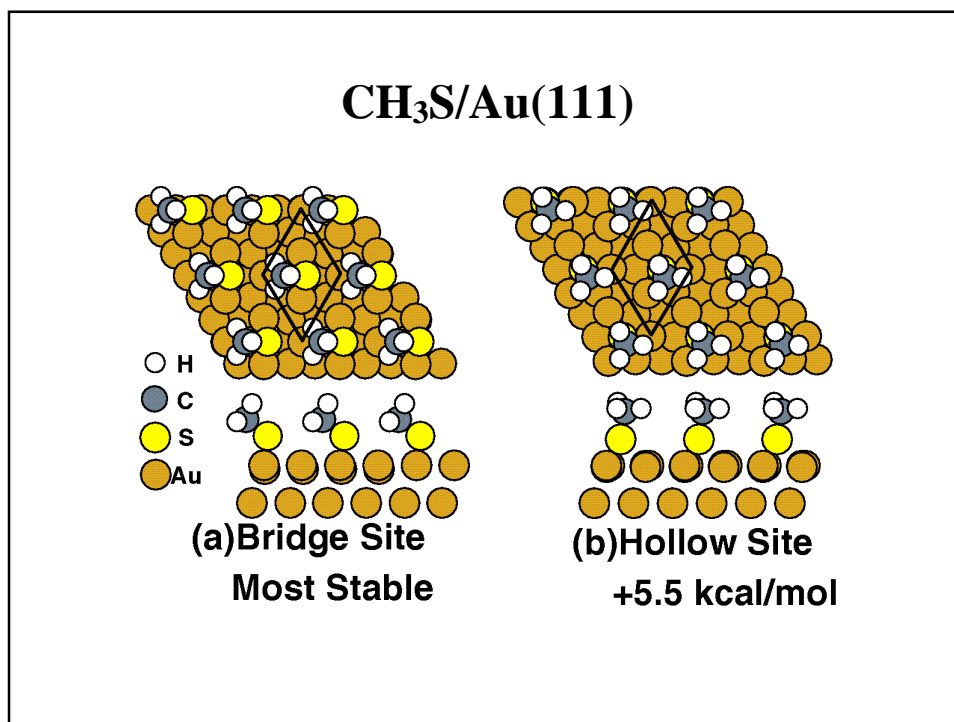


→ No definite conclusion.

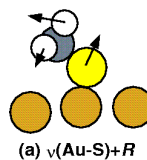
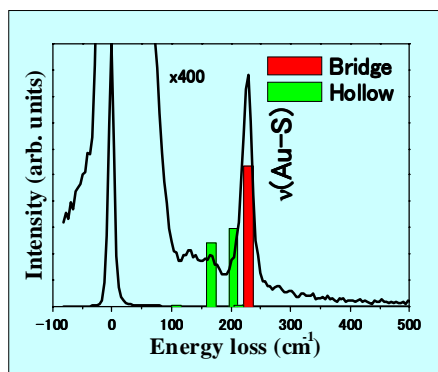
Dimethyl Disulfide



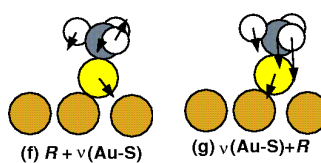
In Good Agreement with Experiment.



HREELS: CH₃S/Au(111)

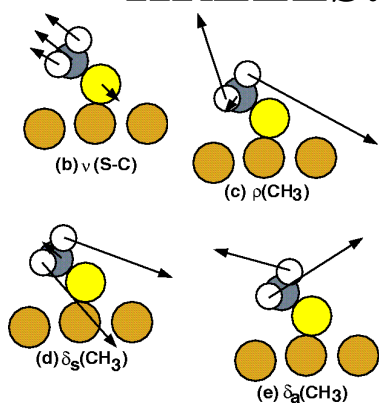


Bridge: Agree Well!

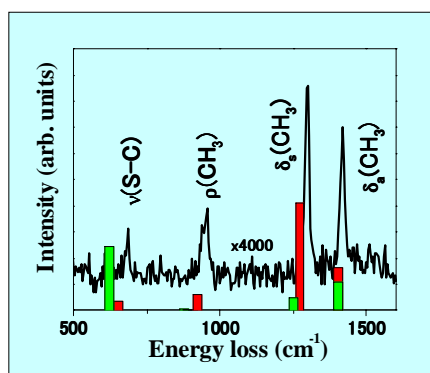


Hollow: Disagree!

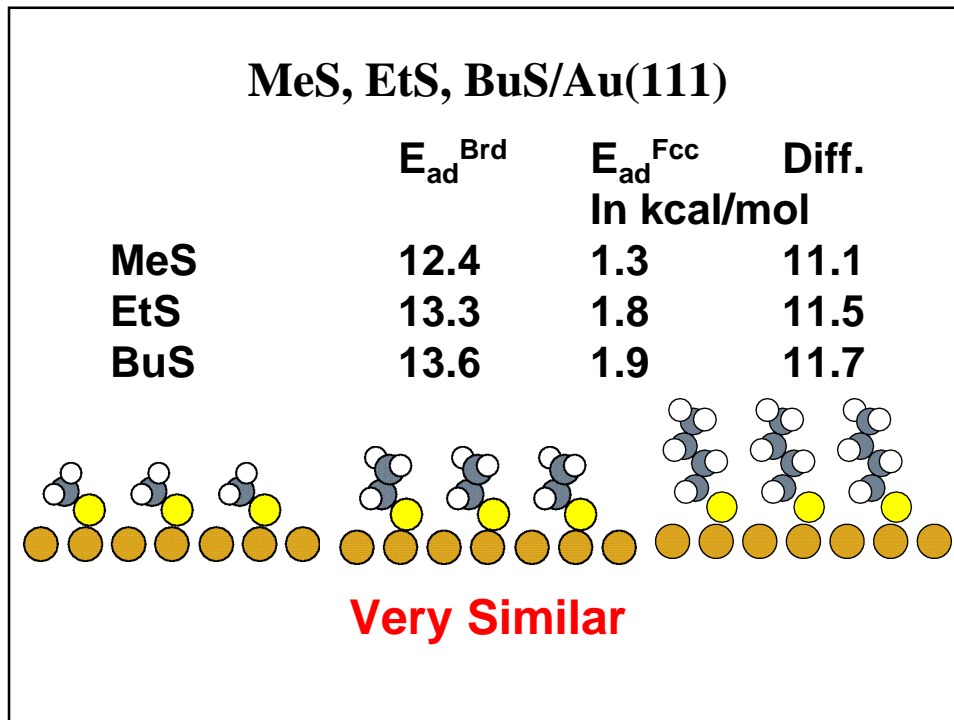
HREELS: CH₃S/Au(111)



Bridge: Agree Well



Hollow: Disagree



有限要素基底の第一原理分子動力学法(土田)

- 平面波基底と同様に、系統的に精度が制御できる。
- 孤立系の計算が容易。
- 変分原理を満たす。
- 並列計算に適している。

Figure 3.4: The shapes of $S_0(x)$ and $S_1(x)$ are shown. $S_1(x)$ is magnified, and the basis functions on neighboring grid points are omitted for convenience.

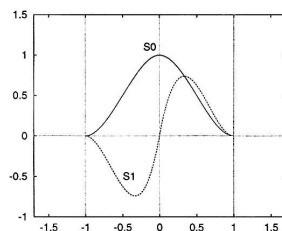
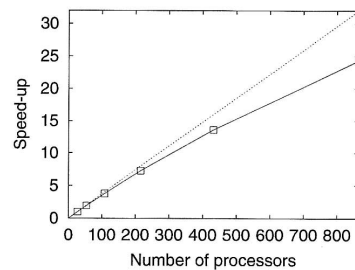


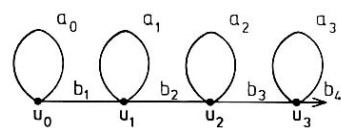
Fig.2 有限要素法を用いたプログラムの自動効果を示す。27プロセッサ上での性能を1としている。



リカーゾン法による第一原理分子動力学法(尾崎)

- リカーゾン法を非直交の局在軌道関数基底に拡張
- オーダーN法
金属にも適用可能

$$H\mathbf{u}_n = a_n\mathbf{u}_n + b_{n+1}\mathbf{u}_{n+1} + b_n\mathbf{u}_{n-1}$$

$$H = \begin{pmatrix} a_0 & b_1 & 0 & 0 & 0 & \cdot & \cdot \\ b_1 & a_1 & b_2 & 0 & 0 & \cdot & \cdot \\ 0 & b_2 & a_2 & b_3 & 0 & \cdot & \cdot \\ 0 & 0 & b_3 & a_3 & b_4 & \cdot & \cdot \\ \cdot & \cdot & & & & & \cdot \end{pmatrix}$$


オーダー(N) SCF密度汎関数計算の概要

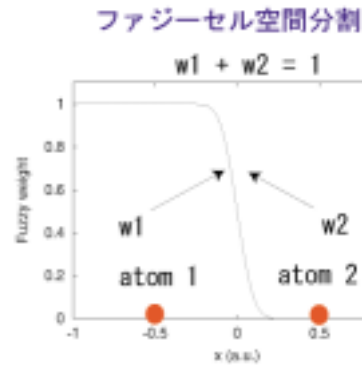
- 一電子波動関数の擬原子数値基底展開
- KB分離されたノルム保存擬ポテンシャル
- ファジーセル空間分割とGauss-Legendre-Lebedevグリッドによる数値積分
- ファジーセル分割と有限差分法によるPoisson方程式のオーダー(N)解法
- ブロックLanczos変換とGreen関数法による固有値問題のオーダー(N)解法
- ネーバーリスト法によるオーダー(N)の使用メモリ

ファジーセル空間分割とGauss-Legendre-Lebedevグリッドによる数値積分

行列要素の積分をファジーセル分割、Gauss-Legendreグリッド及び、Lebedevグリッドを用いて離散化し計算

例えば、3中心積分は以下のように離散化される。

$$\int \chi_{ia}(\mathbf{r} - \mathbf{r}_i) V(\mathbf{r} - \mathbf{r}_k) \chi_{j\beta}(\mathbf{r} - \mathbf{r}_j) d\mathbf{r} \\ = \sum_{s=i,j,k} \sum_{p,q} w_s^F(\mathbf{r}_{p,q}) w_p^{GL} w_q^{Leb} \chi_{ia}(\mathbf{r}_{p,q} - \mathbf{r}_i) V(\mathbf{r}_{p,q} - \mathbf{r}_k) \chi_{j\beta}(\mathbf{r}_{p,q} - \mathbf{r}_j)$$



Poisson方程式のオーダー(N)解法

(Beckeの方法(J. Chem. Phys., 892993(1988))を採用した)

ファジーセル分割によるPoisson方程式の分解

$$\rho(\mathbf{r}) = \sum_I \rho_I(\mathbf{r}) \quad \Delta V_I(\mathbf{r}) = -4\pi\rho_I(\mathbf{r}) \\ \rho_I(\mathbf{r}) = w_I^F(\mathbf{r})\rho(\mathbf{r}) \quad \longrightarrow \quad V(\mathbf{r}) = \sum_I V_I(\mathbf{r})$$

球面調和関数での展開

$$\rho_I(\mathbf{r}) = \sum_{lm} \rho_{lm}(r_I) Y_{lm}(\theta_I, \phi_I) \quad V_I(\mathbf{r}) = \sum_{lm} r^{-1} U_{lm}(r) Y_{lm}(\theta_I, \phi_I)$$

その結果、次式の動径1次元微分方程式を得る

境界条件

$$\frac{d^2 U_{lm}}{dr^2} - l(l+1) \frac{U_{lm}}{r^2} = -4\pi r \rho_{lm} \quad \lim_{r \rightarrow \infty} V(r) = \frac{Q}{r} \\ U_{lm}(r=0) = 0$$

この微分方程式を有限差分法で離散化

$$-(1+h)U_{lm,j-2} + (16+8h)U_{lm,j-1} - (30+12h^2(l+1))U_{lm,j} \\ + (16-8h)U_{lm,j+1} + (-1+h)U_{lm,j+2} = -48h^2\pi r_j^3 \rho_{lm,j}$$

固有値問題のオーダー(N) 解法

ボンドオーダーとグリーン関数の関係

$$\Theta_{\alpha,j\beta} = -\frac{2}{\pi} \text{Im} \int G_{\alpha,j\beta}(E+i0^+) f\left(\frac{E-\mu}{k_B T}\right)$$

非直交基底におけるグリーン関数を求めるためにハイブリッド表示を導入する。

ハイブリッド表示

$$H'_{\alpha,j\beta} = \langle i\alpha | \hat{H} | j\beta \rangle$$

双対基底

$$|i\alpha\rangle = \sum_{j\beta} |j\beta\rangle S_{j\beta,i\alpha}^{-1}$$

ハイブリッド表示されたハミルトニアンを二分岐ブロックLanczos変換でブロック三重対角化する

中心方程式

$$\hat{H}|U_n\rangle = |U_n\rangle A_n + |U_{n-1}\rangle B_n + |U_{n+1}\rangle C_{n+1},$$

$$\langle \hat{U}_n | \hat{H} = \langle \hat{U}_n | A_n + \langle \hat{U}_{n-1} | B_n + \langle \hat{U}_{n+1} | C_{n+1}$$

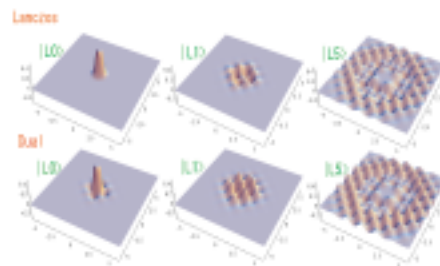
出発状態

$$|U_0\rangle = (|i1\rangle, |i2\rangle, \dots, |iM\rangle),$$

$$|\hat{U}_0\rangle = (|\hat{i}1\rangle, |\hat{i}2\rangle, \dots, |\hat{i}M\rangle)$$

固有値問題のオーダー(N) 解法(続き)

Lanczos基底と双対Lanczos基底は双直交性を保ちながら拡散する。右図は正方格子上的Lanczos基底の発展の様子を示したものである。



ブロック三重対角化されたハイブリッドハミルトニアンから対角・非対角グリーン関数は次式で計算される。

$$G_{0n}^L(Z) = (G_{0n-1}^L(Z)(Z1 - A_{n-1}) - G_{0n-2}^L(Z)C_{n-1} - \delta_{1n}I)(B_n)^{-1}$$

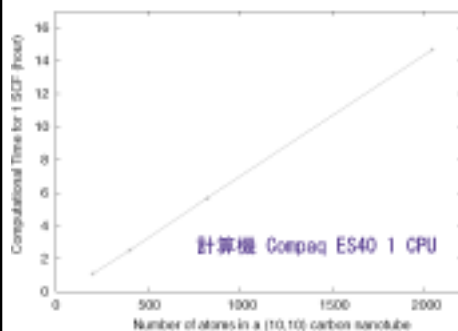
$$G_{0n}^L(Z) = [Z1 - A_0 - C_1[Z1 - A_1 - C_2[\dots]^{-1}B_2]^{-1}B_1]^{-1}$$

最後に逆Lanczos変換によりボンドオーダーを得る。

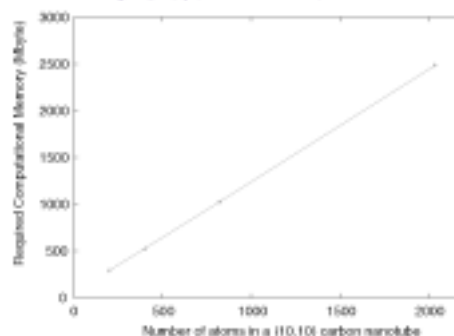
$$\Theta_{ij} = \sum_{\alpha,\beta} \Theta_{0i}^L \hat{U}_{\alpha\beta} S_{\beta j}^{-1}$$

計算時間とメモリサイズ

原子数と計算時間



原子数とメモリサイズ

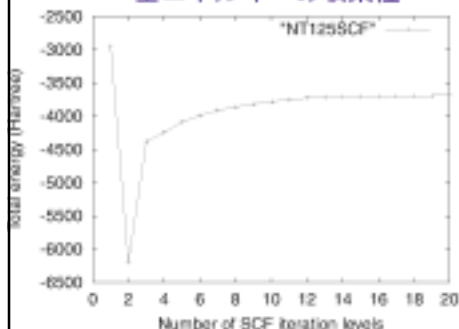


計算条件

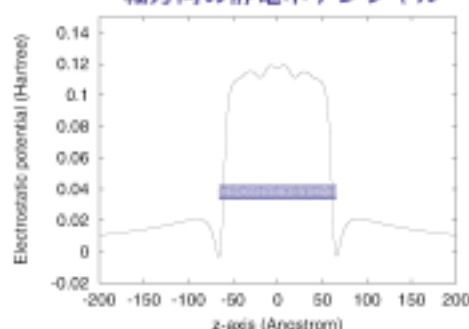
(10, 10)カーボンナノチューブの長さを変えて原子数を調節
 基底=ミニマル基底, 交換相関ポテンシャル=LDA (Ceperly-Alder)
 数値積分グリッド (Gauss-Legendre=32, Lebedev=172)
 Lanczos変換の際のクラスターサイズ=3, シェルクラスター, Lanczos変換のリカージョンレベル=10

オープンエッジを持ったカーボン ナノチューブのSCF計算

全エネルギーの収束性



軸方向の静電ポテンシャル



計算条件

長軸125Åの(10, 10)カーボン(2040原子を含む)
 基底=ミニマル基底, 交換相関項=LDA (Ceperly-Alder)
 数値グリッド (Gauss-Legendre=32, Lebedev=172)
 Lanczos変換の際のクラスターサイズ=3, リカージョンレベル=10

計算機 Compaq ES40, 1CPU
 使用メモリ 2.5GB
 20 SCF計算に要した時間
 133時間

研究目的

密度汎関数法を数千原子の大規模系に適用するために様々なオーダー(N)法が開発されている。しかしこれらのオーダー(N)法の多くは重なり積分の逆行列計算を要求する。通常逆行列計算法は重なり積分の疎構造を考慮してもオーダー(N^2)の計算量を必要とする。それゆえ計算量が原子数Nに比例した効率的なオーダー(N)逆行列計算法の開発を行う。

4つのオーダー(N)逆行列計算法

1. リカージョン法(今回の提案)

ブロックランチョス変換とResolvent法

2. 分割法

系のクラスター分割とLU分解

3. テイラー展開法

$0(=S-I)$ 行列によるテイラー展開

4. Hotelling法

漸化式による繰り返し計算

リカージョン法

Resolventの定義

$$R(Z) = (S - ZI)^{-1}. \longrightarrow S^{-1} = \text{Re}R(0).$$

ブロックLanczos変換による重なり積分Sの
ブロック三重対角化

$$\hat{S}|U_n\rangle = |U_n\rangle A_n + |U_{n-1}\rangle B_n + |U_{n+1}\rangle B_{n+1}$$

初期ベクトル

$$|U_0\rangle = (|i1\rangle, |i2\rangle, \dots, |iM_i\rangle)$$

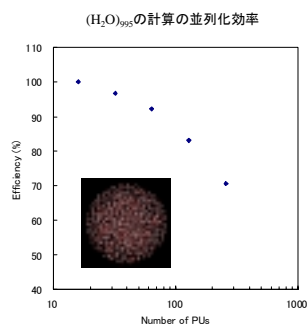
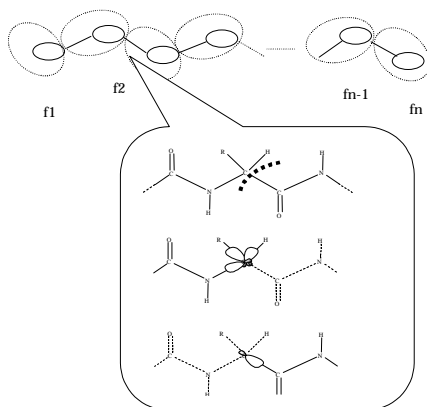
Fragment molecular orbital (FMO) method

- 北浦を中心に開発
巨大分子をfragmentsに分解し、個々のfragment,
fragment pair について量子化学的計算 → 巨大分子の
全エネルギー、原子にかかる力を計算
これまでの最大規模の計算例:
estrogen receptor (240 amino acid residues;
about 4000 atoms) の全エネルギーの1点計算
HF/STO-3G
7.8 h on SR8000/256PU
4.2 h on SR8000/512PU

フラグメント分子軌道法 (Fragment MO method) 巨大分子・分子集合体の電子状態計算法

概要：分子を n 個のフラグメントに分割し、フラグメントとフラグメントペアの計算を行うだけで分子のエネルギーと性質を計算する

特徴：フラグメントやフラグメントペアの計算は独立に行える
高い効率で並列計算ができる



Estrogen Receptor (約 4 0 0 0 原子のタンパク質)のFM0計算結果

Table Energy of Estrogen receptor alpha Complex (PDB:3ERD, 3ERT)						
PDB	Complex	Receptor	Ligand	ΔE	Elapsed time	Number of pu
	/ Hartree	/ Hartree	/ Hartree	/ kcal/mol	/ second	(Complex)
3ERD	-98709.911684	-97877.297156	-832.567099	-29.8	16926	256
3ERT	-99068.853263	-97878.025334	-1190.766958	-38.3	17759	256

ΔE : binding energy, Elapsed time: Hitachi SR8000



3ERD



3ERT

SR8000/32node で約5時間

FMO法による分子内非結合相互作用の解析

FMO法の全エネルギーは一体 (E^{def}) と二体 (ΔE^{int}) エネルギーの和になる → ΔE^{int} は fragment間の相互作用エネルギー

Alanine 10量体 α -helix
の残基間相互作用

$$E = \sum_i E_i + \sum_{i,j} \{ (E_{ij} - E_i - E_j) + Tr(\Delta \mathbf{D}^{ij} \mathbf{V}^{ij}) \}$$

E^{def}

ΔE^{int}

$$E_i = E_x - Tr(\mathbf{D}^i \mathbf{V}^i)$$

$$\Delta \mathbf{D}_{\mu\nu}^{ij} = \mathbf{D}_{\mu\nu}^{ij} - \mathbf{D}_{\mu\nu}^i \oplus \mathbf{D}_{\mu\nu}^j$$

$$\mathbf{D}_{\mu\nu}^i \oplus \mathbf{D}_{\mu\nu}^j = \mathbf{D}_{\mu\nu}^i + \mathbf{D}_{\mu\nu}^j \quad \text{for } \mu, \nu \in I, J,$$

$$= \mathbf{D}_{\mu\nu}^i \quad \text{for } \mu, \nu \in I \text{ and } \mu, \nu \notin J,$$

$$= \mathbf{D}_{\mu\nu}^j \quad \text{for } \mu, \nu \in J \text{ and } \mu, \nu \notin I,$$

$$= 0 \quad \text{for all other cases.}$$

FMO 法によるリゾチームとインヒビター間の相互作用エネルギーの計算

全エネルギーおよび相互作用エネルギー (HF/STO-3G)

	Total energy (a.u.)	ΔE (kcal/mol)	Expt. ΔH (kcal/mol)
Complex (neutral)	-53708.33703	-7.5	-10.4 (PH=2.0)
Complex (charged)	-53710.88591	-11.1	-14.3 (PH=7.6)
Lysozyme(neutral)	-51444.15004		
Lysozyme(charged)	-51446.69326		
(GlcNAc) ₃	-2264.17499		

(GlcNAc)₃ のサイト(A,B,C)と各アミノ酸残基との相互作用

Tri-[N-acetyl-D-glucosamine] (GlcNAc)₃

Lysozyme-(GlcNAc)₃ complex
(PDB:1HEW)

Complex の計算: PCクラスター(18 cpu: 550MHz)で、約半日

古典分子動力学法の高速度化手法の開発

大規模系を精度良く長時間にわたり高速に計算する。

長距離力計算法

高速多重極子展開法 (開発済み)

三次元粒子メッシュエバルト法 (開発済み)

高速二次元エバルト法 (一部開発済み、継続中)

長時間高速積分法

多重時間数値積分法 (開発済み)

シンプレティック積分法 (開発済み)

多重時間シンプレティック積分法 (今後開発予定)

拡張アンサンブルの開発

複雑物質系/Rare Eventの位相空間を短時間でもれなく探索する。

マルチカノニカル法

(NTP系を開発済み、今後、高分子のバルク系に適用)

レプリカ交換法

(NTP系を開発済み、今後、複雑物質、化学反応系に適用)

自己誘導力法 (NTP系を開発済み、今後、生体高分子に適用)

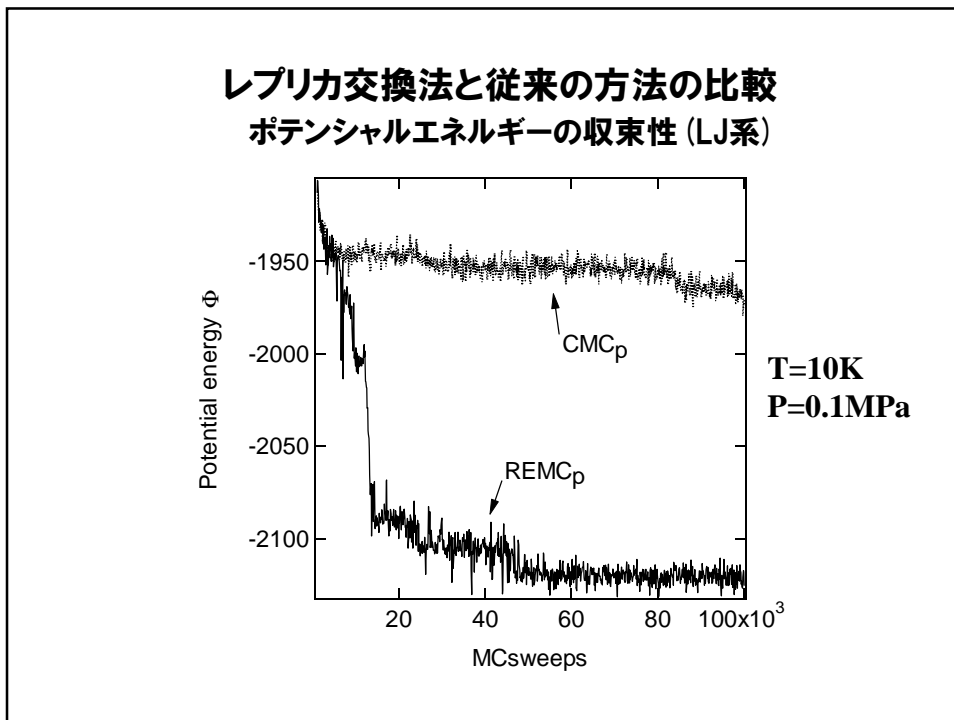
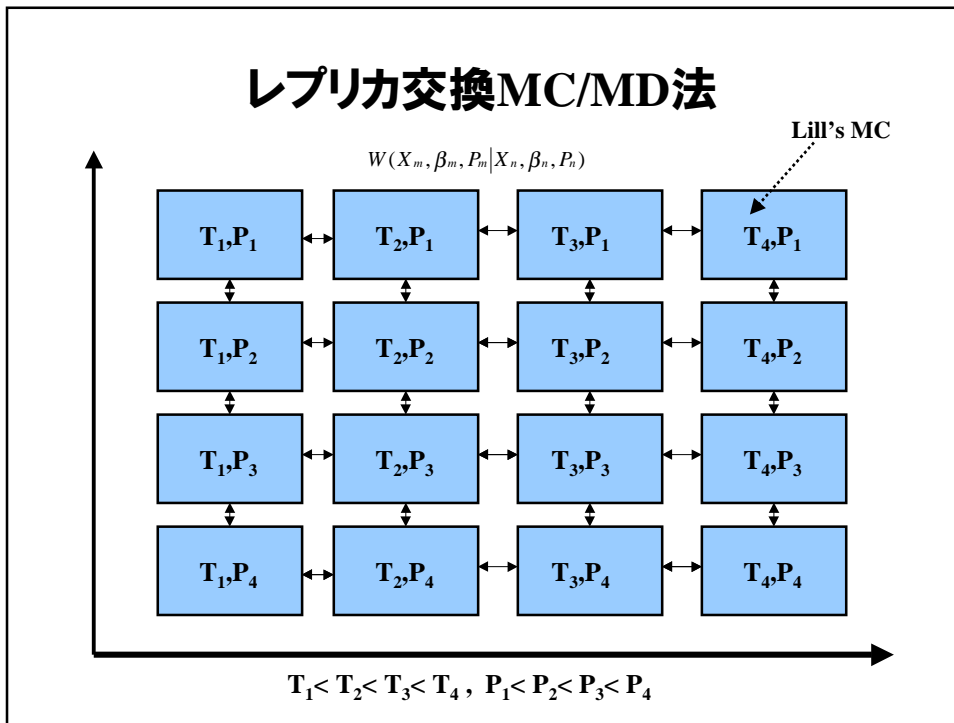
配置バイアスモンテカルロ法

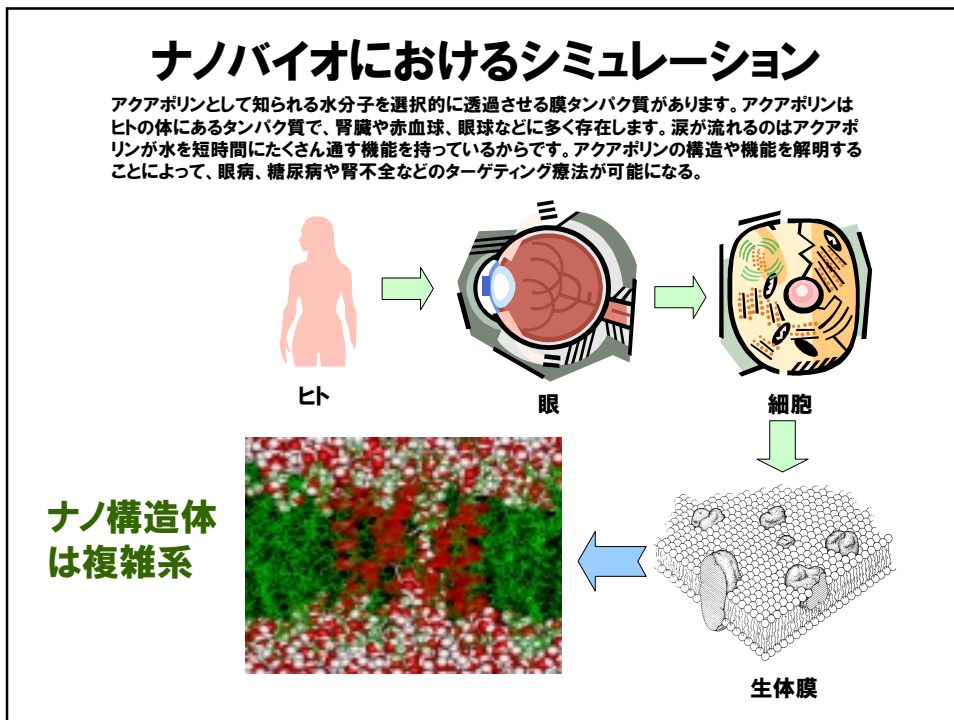
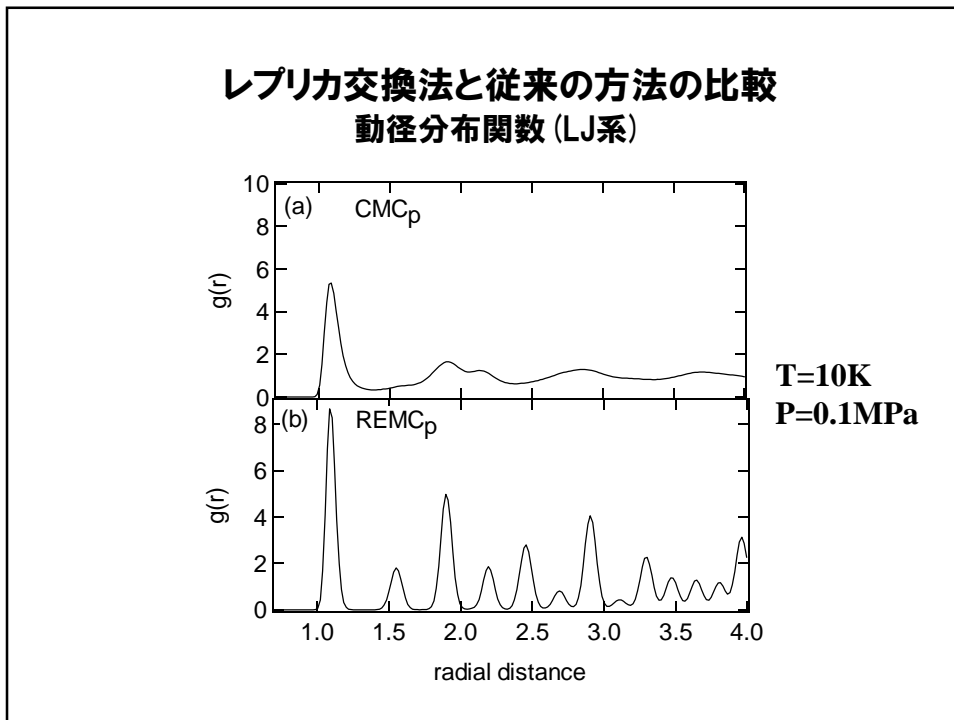
(開発予定、分子の溶解度予測、膜蛋白質の水透過に適用)

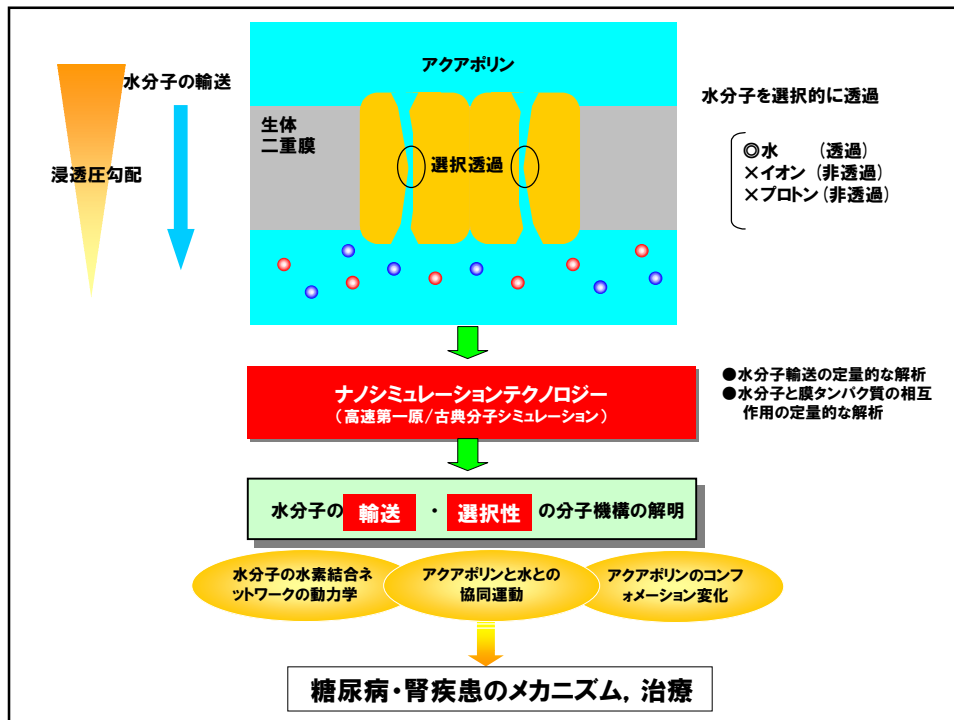
遷移経路サンプリング法 (今後開発予定)

一般化除冷法/サリス統計法 (今後開発予定)

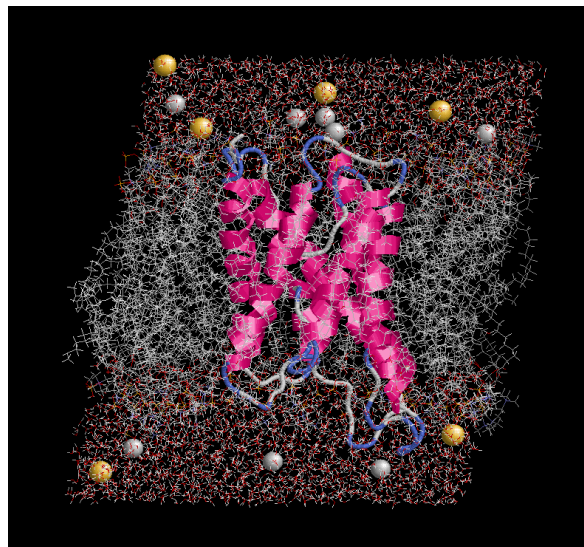
上記の方法を組み合わせた新しい方法 (今後開発予定)







膜蛋白質の分子動力学シミュレーション 水チャンネル (アクアポリン)



システムサイズ

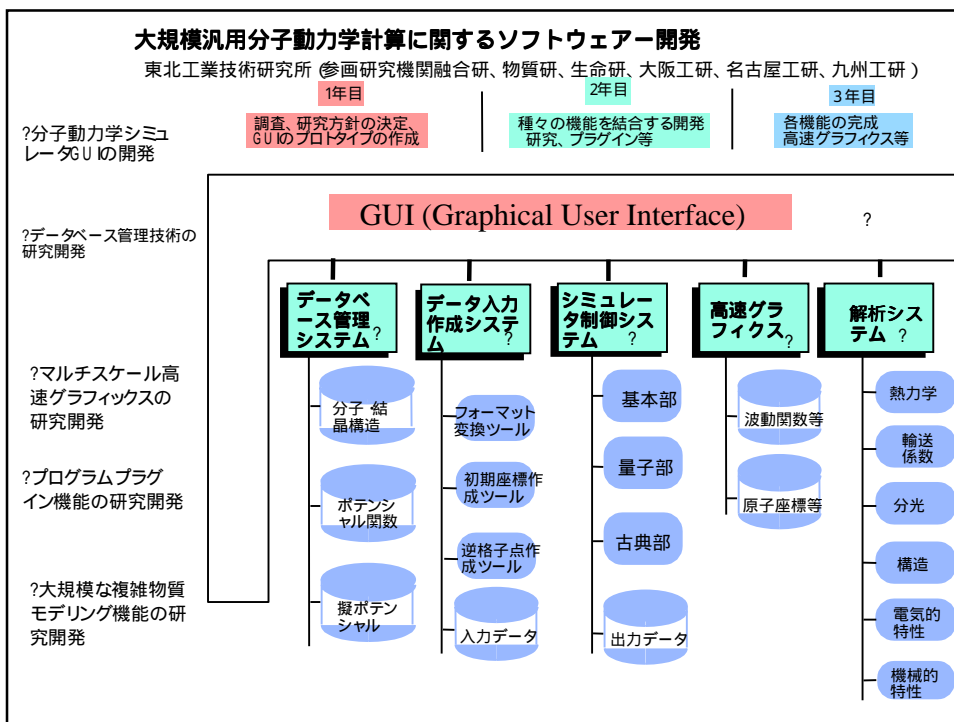
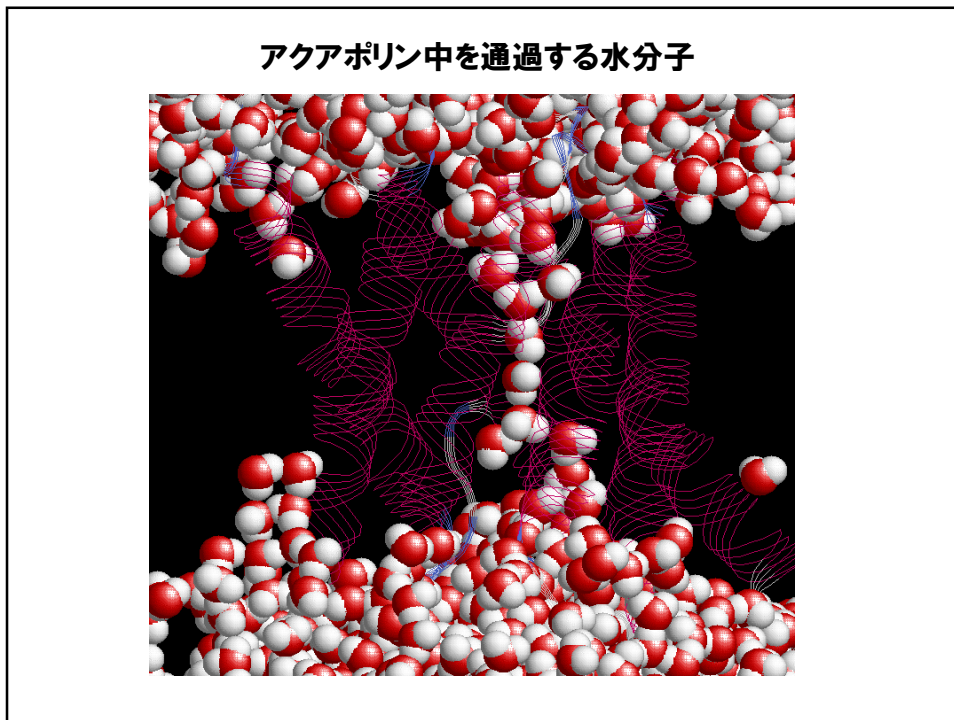
原子数 27,698
 タンパク 1分子
 脂質 94分子
 水 4029分子
 Na+ 6
 Cl- 8

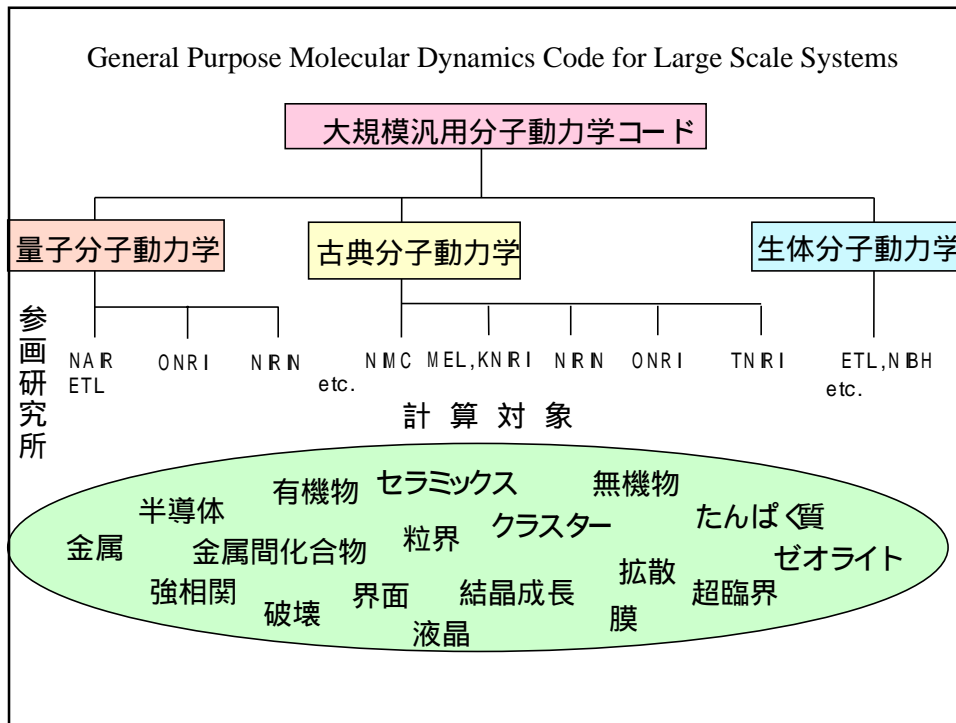
計算時間

3.45s / step

計算機

Compaq
 XP1000 Cluster
 15nodes





■ 大規模並列解析への敷居の高さの解消 ■ (富士総研との共同研究)

- ・逐次解析プログラムの並列化には**多大な作業時間**が必要
- ・PETSc, Aztec, GEOFEM, ADVENTURE等で研究開発された並列ソフトを自らの研究に取り込むには**並列計算のプロである必要**.

並列計算プラットフォームの開発

逐次解析プログラムが数時間で並列化可能

1. ユーザから供給すべき情報
 - ・有限要素モデルデータ
 - ・要素毎の係数マトリックス作成ルーチン
2. 行列計算ルーチン
 - GMRES, BiCGSTAB(前処理:ポイントヤコビ法, 加法的シュルツ法), バンド法
3. サブルーチンツール
 - 並列用行列インデクス作成, 各種境界条件作成
4. 並列解析のための領域分割
 - グラフ生成ルーチン+MeTiSを使用

並列化変更パターン

```

DIMENSION Aij( ),WORK( ),Bi( ),IDOMAIN( )
CALL MPI_INIT( ... ) ; MPI 初期化

set IDOMAIN(I) ; 要素の領域番号入力

CALL GM4_INDEX ; 並列用インデクス作成

DO I=1, 要素数 ; 剛性行列の作成
IF (IDOMAIN(I) .EQ. MYRANK) THEN ; 自分のみ処理
. . . . .
Aij( )= ユーザ供給部分 ; 要素剛性作成
CALL GM4_INDEX ; 全体剛性作成
END IF
END DO

CALL GM4_BOUNDARY ; 境界条件作成

CALL GM4_SOLVE(Aij,Bi,IDOMAIN,.WORK); ソルバ

CALL MPI_FINALIZE( ... ) ; MPI 終了
STOP
END
    
```

