

# 次世代分子理論の開発と展開

---

第十回PCクラスタシンポジウム  
2010年12月9日  
秋葉原コンベンションホール

計算科学研究機構  
中嶋 隆人

# 理論分子科学

原子・分子の世界は量子論によって支配されている。SchrödingerやDiracの波動方程式を解けば原子や分子の動きを記述することができる。

如何に近似して波動方程式を解くかが理論分子科学における理論開発の目的。

深く+大きく+幅広く

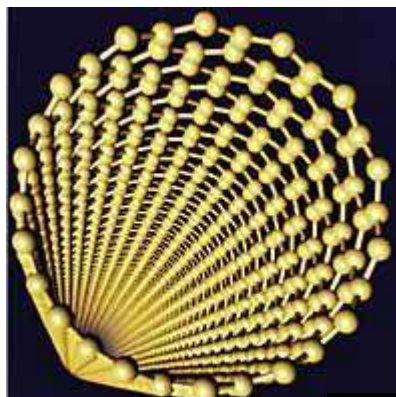
## 現状

- 計算手法の発展とコンピュータの進歩により、10年前とは比較にならないほど大規模な分子の性質を高い信頼度をもって計算できるようになっている。
- 対象とすることのできる現象や系は大きく拡がり、比較的大きな分子系に対してでも実験結果をうまく説明したり解釈できるようになっている。数原子分子系に対しては十分な予測性と定量性も備えている。
- それでもなお、現実系の分子に対しては十分な予測性を備えるには至っていない。

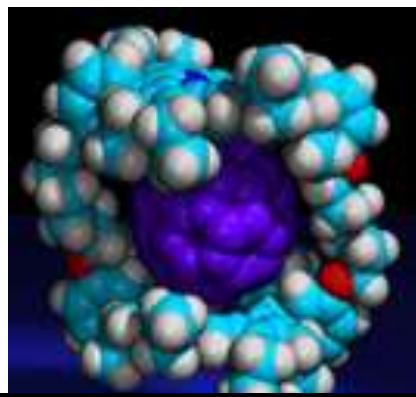
予測性を備えた理論分子科学による理論先導の科学

# 次世代分子理論

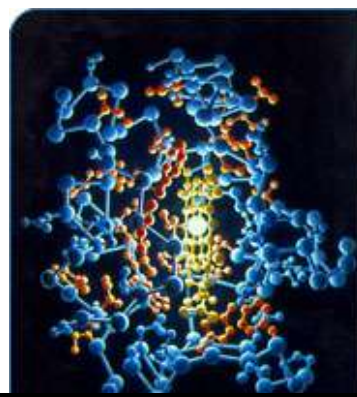
近年、新しい機能・特性の発現を目指して、科学者の取り扱う物質は多種多様な元素を含む複雑で大規模な分子系へと拡がりを見せている



ナノマテリアル



現在・次世代の科学や産業の基盤となる物質

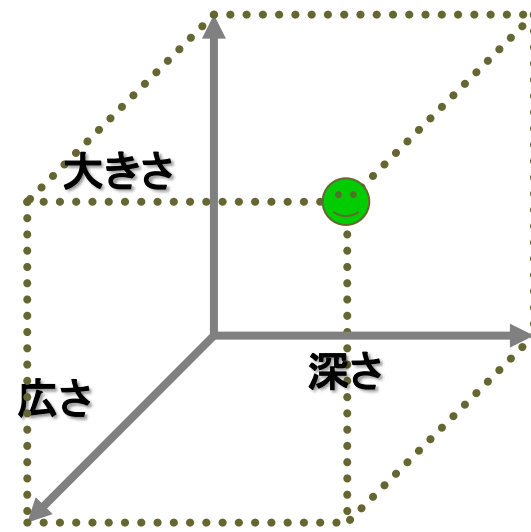


バイオマテリアル

科学で取り扱う物質の多様化・複雑化・大規模化に伴い、理論分子科学も科学的に興味のある多彩な系に対し積極的に貢献すべき

深さ + 大きさ + 幅広さ

幅広い環境における複雑で大規模な物質を高精度に取り扱える分子理論の開発



# 相對論的分子理論



# REL4Dプロジェクト

- 相対論的分子理論を新たに開発し展開することで、幅広い種類の元素を含む分子系を同じ精度でかつ高精度に取り扱うことのできる理論分子科学を構築する

## 次世代分子理論の展開

PERIODIC CHART OF THE ELEMENTS

すべての元素を  
同じ精度で!!

IA	IIA	IIIB	IVB	VB	VIB	VII B	VIII	IB	II B	IIIA	IVA	VA	VIA	VIIA	INERT GASES					
1 H 1.00794														2 He 4.0026						
3 Li 6.941	4 Be 9.0122													5 B 10.811	6 C 12.0112	7 N 14.0067	8 O 15.9994	9 F 18.9984	10 Ne 20.183	
11 Na 22.9898	12 Mg 24.304													13 Al 26.9815	14 Si 28.086	15 P 30.9738	16 S 32.06	17 Cl 35.453	18 Ar 39.948	
19 K 39.102	20 Ca 40.08	21 Sc 44.956	22 Ti 47.88	23 V 50.942	24 Cr 51.996	25 Mn 54.938	26 Fe 55.847	27 Co 58.9332	28 Ni 58.71	29 Cu 63.54	30 Zn 65.37	31 Ga 69.72	32 Ge 72.59	33 As 74.9216	34 Se 78.96	35 Br 79.904	36 Kr 83.60			
37 Rb 85.47	38 Sr 87.62	39 Y 88.905	40 Zr 91.22	41 Nb 92.906	42 Mo 95.94	43 Tc (99)	44 Ru 101.07	45 Rh 102.905	46 Pd 106.4	47 Ag 107.87	48 Cd 112.40	49 In 114.82	50 Sn 118.69	51 Sb 121.75	52 Te 127.6	53 I 126.904	54 Xe 131.30			
55 Cs 132.905	56 Ba 137.34	57 La 138.91	72 Hf 178.49	73 Ta 180.948	74 W 183.85	75 Re 186.2	76 Os 190.2	77 Ir 192.22	78 Pt 195.08	79 Au 196.967	80 Hg 200.59	81 Tl 204.37	82 Pb 207.19	83 Bi 208.980	84 Po (210)	85 At (210)	86 Rn (222)			
87 Fr (223)	88 Ra (226)	89 Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (264)	108 Hs (265)	109 Mt (266)	110 ? (271)	111 ? (272)	112 ? (273)									

Numbers in parenthesis are mass numbers of most stable or most common isotope.

Atomic weights corrected to conform to the 1963 values of the Commission on Atomic Weights.

The group designations used here are the former Chemical Abstract Service numbers.

\* Lanthanide Series

58 Ce 140.12	59 Pr 140.907	60 Nd 144.24	61 Pm (147)	62 Sm 150.35	63 Eu 151.96	64 Gd 157.25	65 Tb 158.928	66 Dy 162.50	67 Ho 164.931	68 Er 167.26	69 Tm 168.934	70 Yb 173.04	71 Lu 174.967
--------------------	---------------------	--------------------	-------------------	--------------------	--------------------	--------------------	---------------------	--------------------	---------------------	--------------------	---------------------	--------------------	---------------------

† Actinide Series

90 Th 232.038	91 Pa (231)	92 U 238.03	93 Np (237)	94 Pu (242)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (249)	99 Es (254)	100 Fm (253)	101 Md (256)	102 No (256)	103 Lr (257)
---------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	-------------------	--------------------	--------------------	--------------------	--------------------



REL4D

- 様々な化学的現象に適用して、化学における相対論効果の重要性をひろく認識する
- 大規模な重原子分子系を高精度計算することのできるプログラムを完成させる

# Dirac方程式

外場中を運動するひとつのFermi粒子を記述することのできる相対論的方程式

$$H_D \psi \equiv (c \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m c^2 + V_{\text{ext}}) \psi = E \psi$$

$H_D$  : 1粒子Dirac Hamilton演算子

$\boldsymbol{\alpha}, \beta$  : Dirac行列

$$\alpha_t \equiv \begin{pmatrix} 0_2 & \sigma_t \\ \sigma_t & 0_2 \end{pmatrix}, t = (x, y, z), \quad \beta \equiv \begin{pmatrix} \mathbf{I}_2 & 0_2 \\ 0_2 & -\mathbf{I}_2 \end{pmatrix}$$

$\sigma_t$  : 2×2 Pauli spin行列

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dirac方程式は4成分の波動関数 (4成分スピノル) を持つ

$$\psi = \begin{pmatrix} \varphi^1 \\ \varphi^2 \\ \varphi^3 \\ \varphi^4 \end{pmatrix} = \begin{pmatrix} \varphi^L \\ \varphi^S \end{pmatrix}, \varphi^L = \begin{pmatrix} \varphi^1 \\ \varphi^2 \end{pmatrix}, \varphi^S = \begin{pmatrix} \varphi^3 \\ \varphi^4 \end{pmatrix}$$

$\varphi^L$  : large成分 (電子状態)  
 $\varphi^S$  : small成分 (陽電子状態)

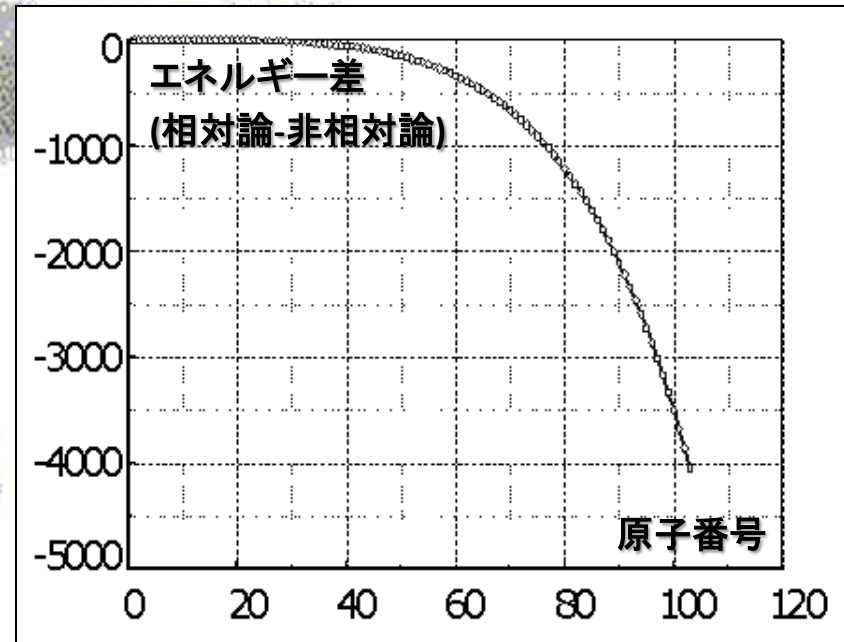
$$\begin{pmatrix} V_{\text{ext}} + c^2 & c \boldsymbol{\sigma} \cdot \mathbf{p} \\ c \boldsymbol{\sigma} \cdot \mathbf{p} & V_{\text{ext}} - c^2 \end{pmatrix} \begin{pmatrix} \varphi^L \\ \varphi^S \end{pmatrix} = E \begin{pmatrix} \varphi^L \\ \varphi^S \end{pmatrix}$$

# 原子・分子における相対論効果

## 重原子効果

原子核近傍で高速運動している電子に起因する相対論効果. 重原子効果は原子番号の増加につれて急激に重要になる.

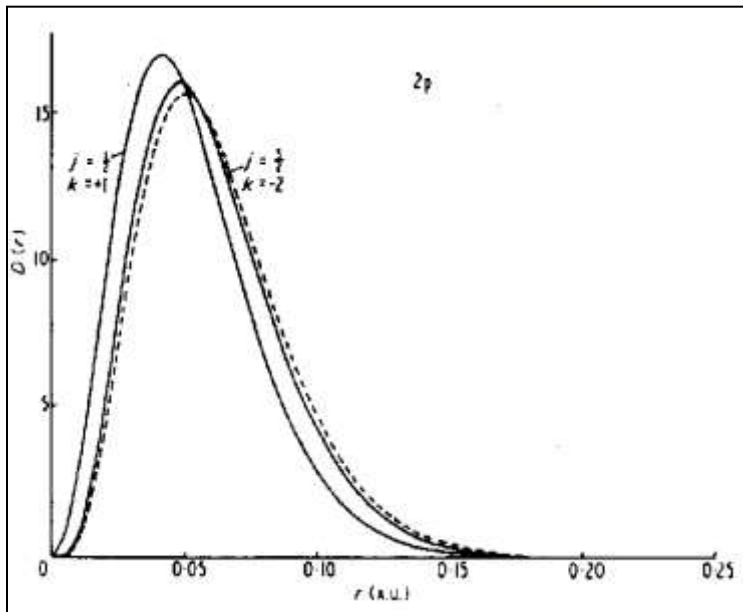
**s軌道の相対論的収縮と軌道の安定化**



## スピン-軌道効果

電子のLSカップリングに由来する相対論効果. 項間交差を通して様々な化学反応やプロパティにおいて重要になってくる相対論効果.

**p, d, f軌道の軌道分裂**



# 相対論的分子理論

Dirac方程式は多大な計算量が必要, 従来は小さな分子に適用するのも困難

## 4成分相対論的分子理論

相対論的方程式を効率的に解くため, 2成分型スピノル基底を導入

## 相対論的分子計算の高速計算法

世界最速の積分計算を実現

相対論積分の計算時間

### Dirac-Hartree-Fock法

高速な4成分相対論分子理論

### Dirac-Kohn-Sham密度汎関数法

世界最初のGauss型4成分密度汎関数理論

	LLLL+LLSS+SSSS	LLLL+LLSS	LLLL
Present	1.37	0.77	0.21
MOLFDIR	76.35	21.16	1.63

60倍高速

## 相対論的近似分子理論

重原子分子に対する高精度・大規模計算が可能

### RESC法

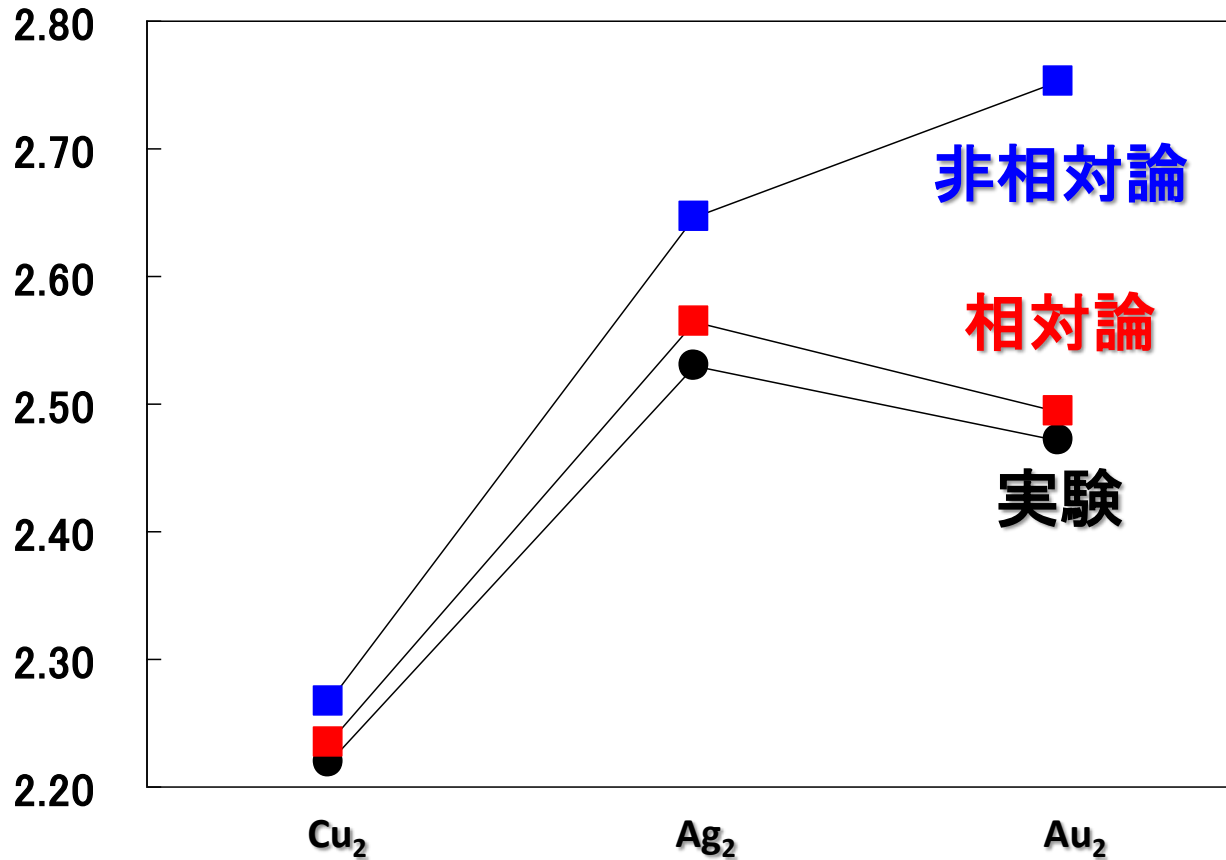
本質を衝くモデル, シンプルなアイデアにもかかわらず高精度

### 高次Douglas-Kroll (DK) 法

2成分相対論的分子理論の決定版



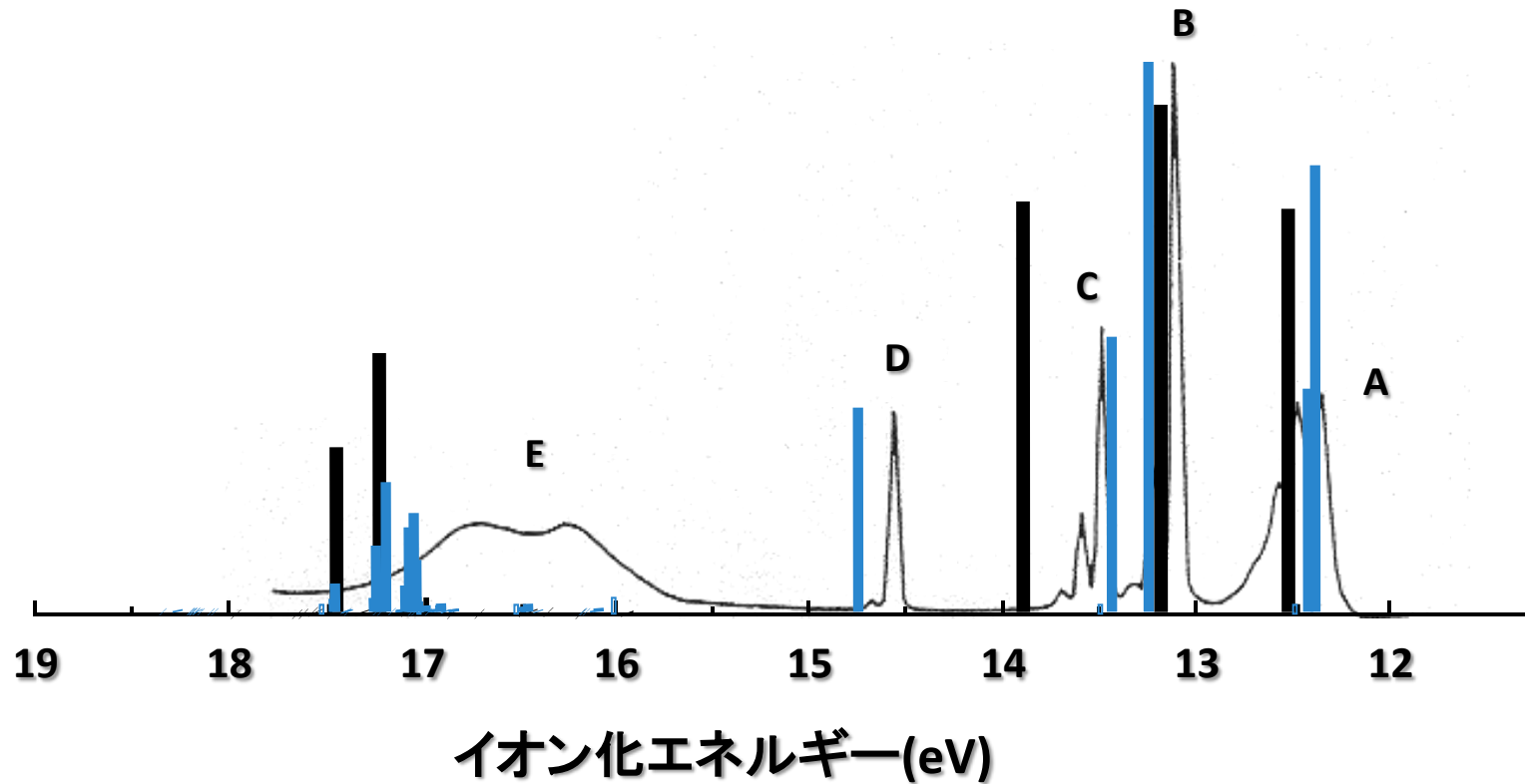
# 分子に対する相対論効果



11族遷移金属二量体の平衡核間距離 (Å)

- 相対論効果を考慮してはじめて実験値の傾向を再現
- 重原子を含む分子系では相対論効果が重要

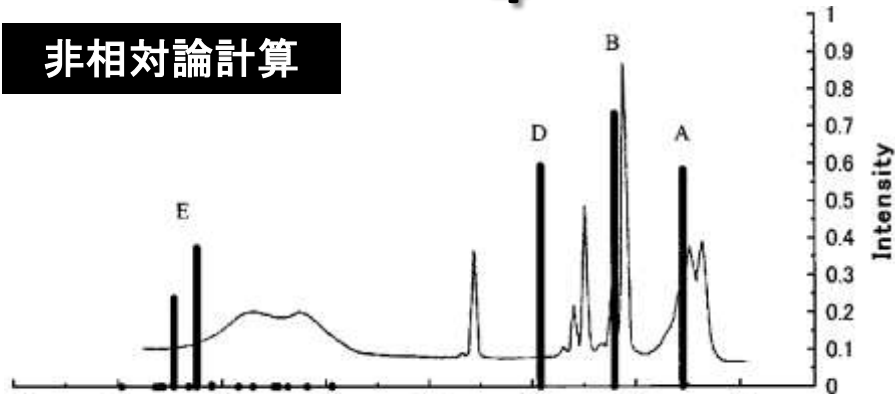
# OsO<sub>4</sub>の光電子スペクトル



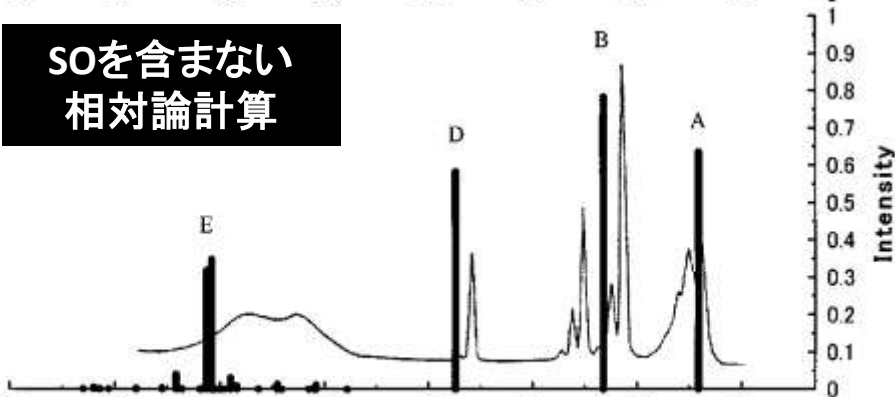
- 非相対論 --- ピークCとDを同定することができない
- 相対論 --- 実験と良く一致する

# OsO<sub>4</sub>の光電子スペクトル

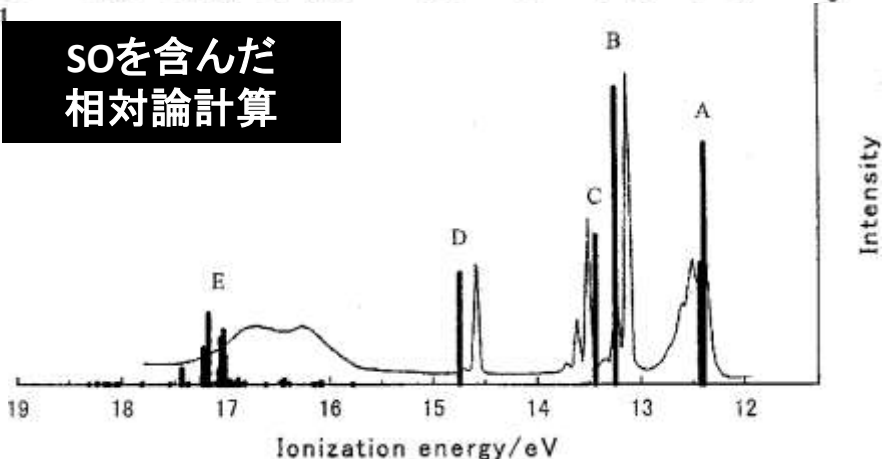
非相対論計算



SOを含まない  
相対論計算



SOを含んだ  
相対論計算



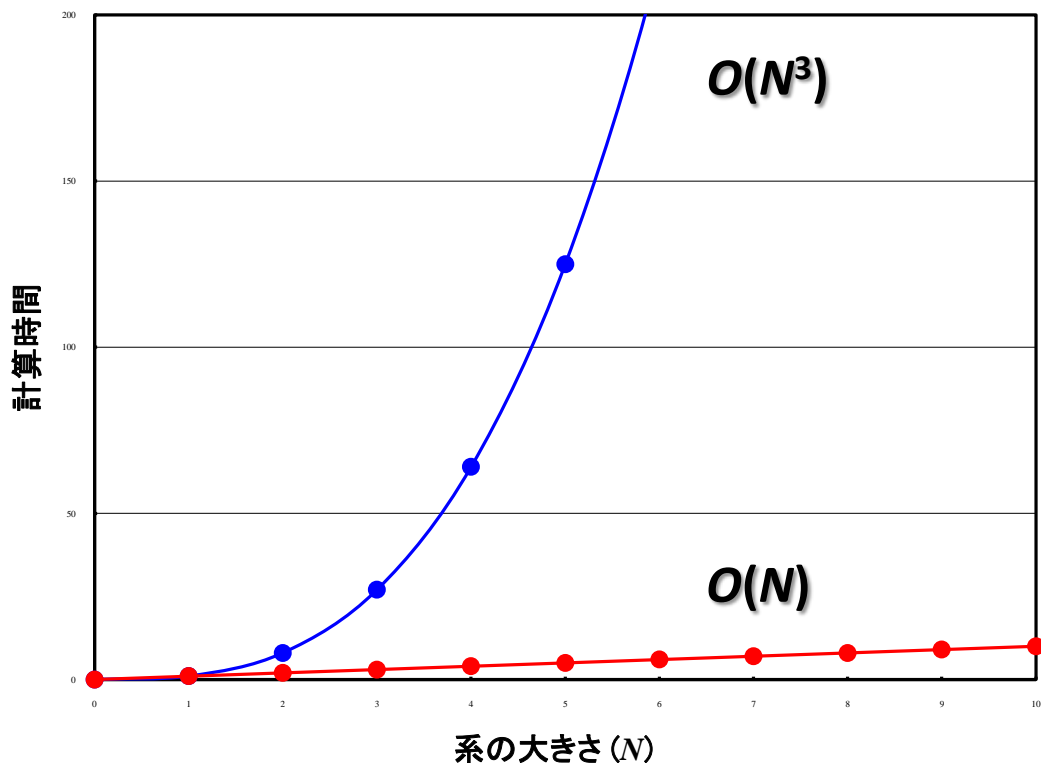
- ピークDは**相対論的収縮**による軌道の安定化により高エネルギー側にシフト (DはOsの価電子s軌道からのイオン化)
- ピークBとCは**スピン-軌道相互作用**による**軌道分裂**により得られる

重原子分子においては価電子の寄与する化学的性質に対してでさえ相対論効果は重要

# 大規模分子理論



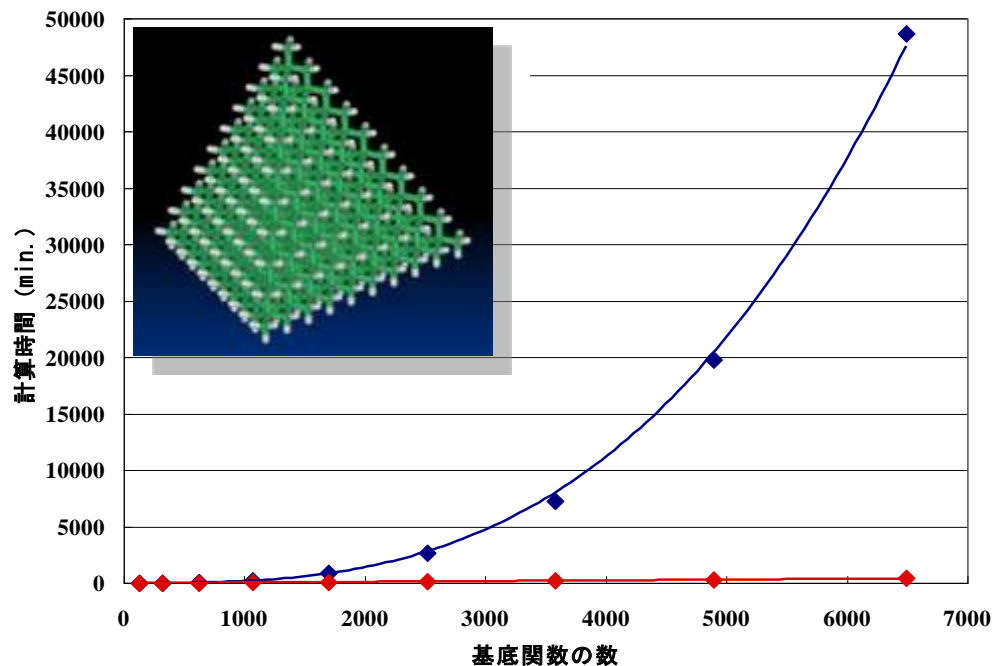
# 大規模分子理論



- 従来の分子理論は、系の大きさ(N)に対して計算時間が $O(N^3)$ 以上でスケールするので、大規模分子になるほど計算コストが必要
- 大規模分子理論の目的は、なんらかの近似を導入することで系の大きさに対してスケールラブルな計算 ( $O(N)$ , linear-scaling) を達成して、高速かつ大規模な分子計算を実現することにある

# 独自の大規模分子理論

- Pseudospectral法
- Gaussian Auxiliary Plane-Wave法
- Gaussian Auxiliary Finite-Element法
- Gaussian & Finite-Element Coulomb法
- Dual-Level密度汎関数法
- PS-GAP法
- RILL-MP2法
- SPOT法
- Gauss+平面波ONIOM



# Gaussian & FE Coulomb法

JCP, 126, 144106 (2007) & JCP, 128, 054105 (2008).

SCF中で計算コストの高いCoulomb積分を高速に計算する方法

$$J_{pq} = \sum_{rs} D_{rs} \int \chi_p(r_1) \chi_q(r_1) \frac{1}{r_{12}} \chi_r(r_2) \chi_s(r_2) dr = \int \chi_p(r) V(r) \chi_q(r) dr$$

- Gauss型基底をCoulomb積分と波動関数を表現するために用いる
- Coulombポテンシャルを表現するためにGauss型関数と有限要素関数の混合基底を用いる

$$V(r) \simeq \sum_i c_i^{\text{Gauss}} \cdot f_i^{\text{Gauss}}(r) + \sum_i c_i^{\text{FE}} \cdot f_i^{\text{FE}}(r)$$

$$J_{pq} \simeq \sum_i c_i \cdot \underbrace{\int f_i(r) \cdot \chi_p(r) \chi_q(r) dr}_{\text{overlap integral}}$$

- 重なり積分は漸近的に $O(N)$ の計算が可能であるので、Coulomb積分も $O(N)$ で評価することができる

# GFC法

## 展開係数の決定

- Coulombポテンシャル中の展開係数を決定するために、代数的な形でPoisson方程式を解く

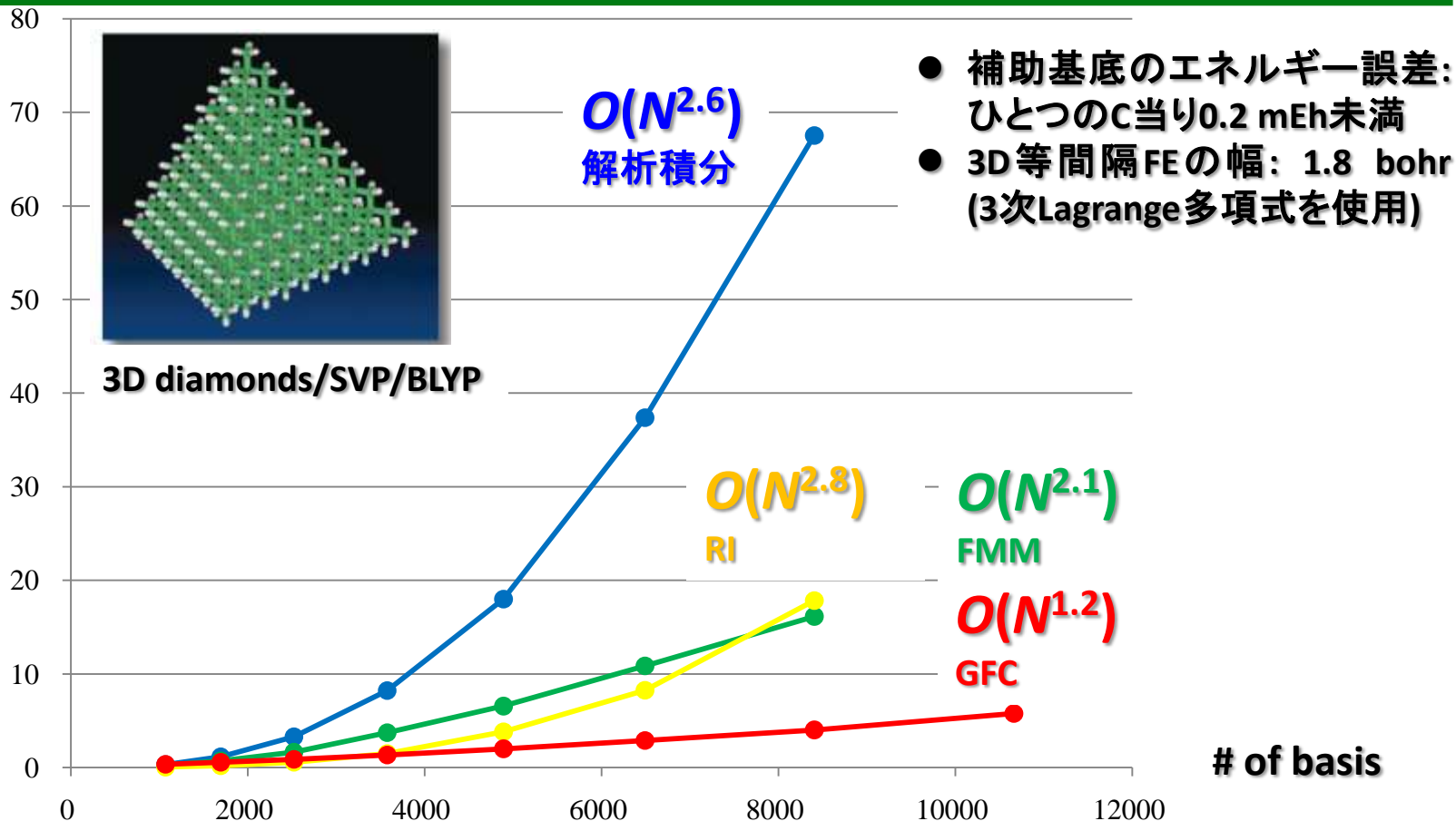
$$-\nabla^2 V(r) = 4\pi\rho(r) \xrightarrow{\text{auxiliary functions}} \begin{bmatrix} \ddots & & 0 \\ & \mathbf{A} & \\ 0 & & \ddots \end{bmatrix} \times \begin{bmatrix} \vdots \\ \mathbf{c} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathbf{b} \\ \vdots \end{bmatrix},$$
$$\left\{ \begin{array}{l} \mathbf{A}_{ij} = \frac{\int \nabla f_i(r) \cdot \nabla f_j(r) dr}{\text{kinetic integral}} \\ \mathbf{b}_i = \frac{\int \rho(r) \cdot f_i(r) dr}{\text{overlap integral}} \\ V(r) \simeq \sum_i \mathbf{c}_i \cdot f_i(r) \end{array} \right.$$

- 有限要素と原子様Gauss型関数の局在化の性質から行列Aは疎であるので、このPoisson方程式は $O(N)$ で解くことができる
- 線型方程式を解く際に共役勾配法を用いる
- 補助Gauss関数の軌道指数は原子のCoulombエネルギーに対して解析解との誤差を小さくするように、あらかじめ決定しておく



# GFC法: スケーリング

CPU time/  
hour



- 系が大きくなるのにつれ, GFCは他の方法よりも効率が良い
- GFCは $O(N^{1.2})$ で線形スケーリング計算を実現できる

GFCは分子の形状によらず線形スケーリングを達成

# GFCエネルギー微分

JCP, 132, 244107 (2010).

- GFC法を使うと、エネルギー以外にエネルギー微分も効率的に計算できる。

## 核座標 $X$ に関するCoulombエネルギーの一次微分

$$\frac{\partial E_J}{\partial X} = \frac{\partial}{\partial X} \left\{ \int dr_1 \int dr_2 \rho(r_1) \frac{1}{r_{12}} \rho(r_2) \right\} = 2 \int dr_1 \left( \frac{\partial \rho(r_1)}{\partial X} \right) V(r_1)$$

- GFC法ではCoulombポテンシャル $V(r_1)$ を補助基底で展開する。エネルギーの場合と同様に、エネルギー微分にも補助基底展開を採用する。

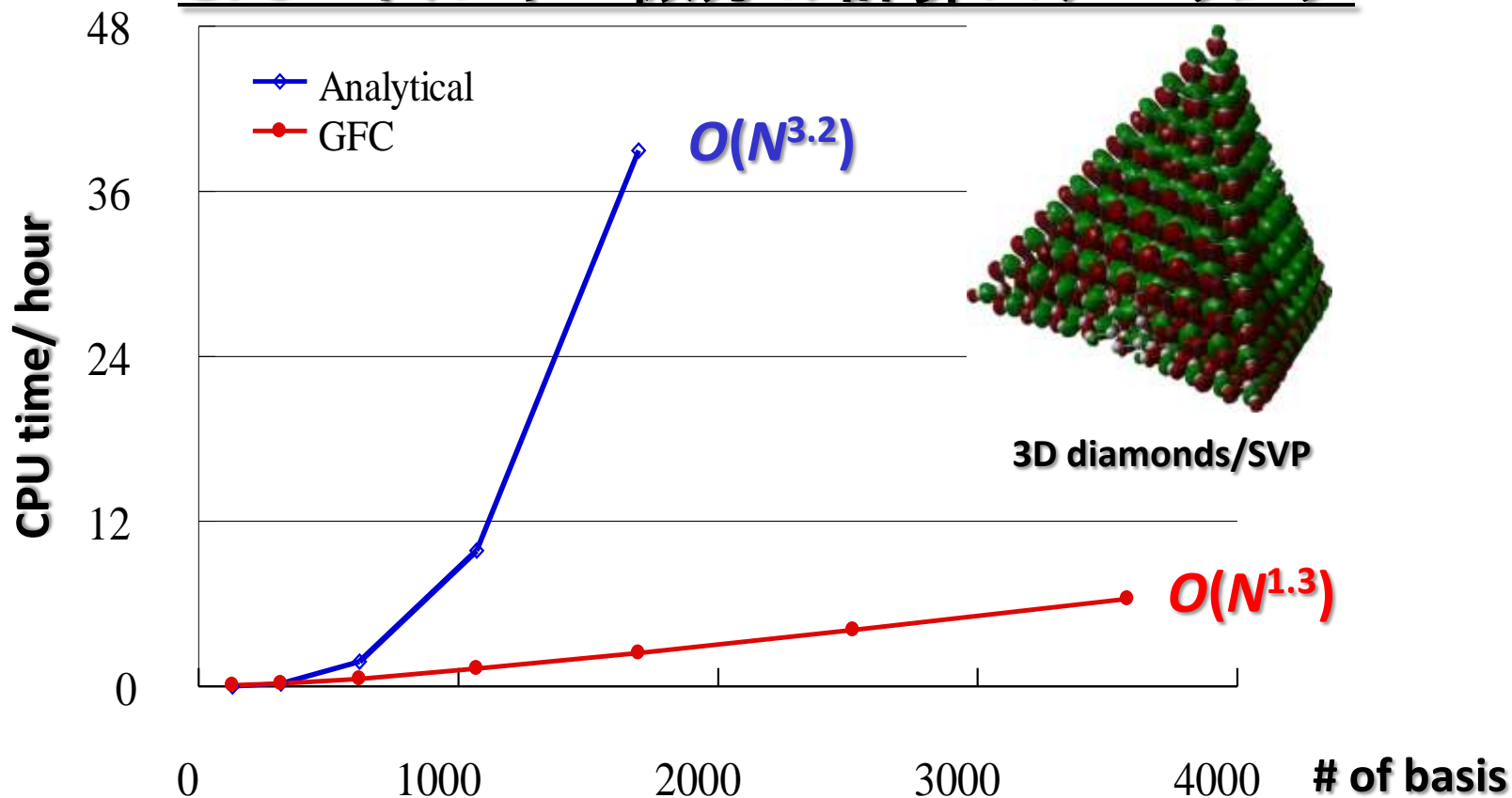
$$\begin{aligned} \frac{\partial E_J}{\partial X} &= 2 \sum_i c_i^G \int dr \left( \frac{\partial \rho(r)}{\partial X} \right) f_i^G(r) + 2 \sum_i c_i^{\text{FE}} \int dr \left( \frac{\partial \rho(r)}{\partial X} \right) f_i^{\text{FE}}(r) \\ &= 2 \sum_{pqi} D_{pq} c_i \int dr \left( \frac{\partial \chi_p}{\partial X} \right) \chi_q f_i(r) \end{aligned}$$

- 3中心重なり積分の形で与えられるので、エネルギー微分についても線形スケールリングを達成できる。

大規模分子の化学反応の追跡やダイナミクス計算が実現可能

# GFCエネルギー微分

## GFCエネルギー微分の計算スケーリング



- 系が大きくなるのにしたがって、GFCエネルギー微分は従来の解析的方法よりも効率的

エネルギー微分計算に対しても線形スケーリングを達成

# 本プロジェクトの研究目標

本プロジェクトの目標はナノマテリアルや生体分子のような大規模で複雑な分子系の機能を**ミクロの立場から理論先導で解明**するため、**次世代の理論分子科学の基盤を構築し諸問題に適用**すること

## 5年後の達成目標

- 現状の分子理論が抱える問題点を克服し、ブレークスルーを達成することで、大規模分子系を定量的に取り扱える「次世代分子理論」を構築する
- この「次世代分子理論」をもとに理論分子科学の革新を狙う。次世代スパコンの能力を最大限に発揮することのできる理論分子科学の計算手法とその計算プログラム「**NTChem**」を開発する

## 長期的目標

- 複雑で大規模な実在分子系を化学的精度をもって取り扱うことのできる分子理論の開発。アルゴリズム・ハードウェアレベルにまで特化した次々世代スパコンのアプリソフトとしての新しい分子計算プログラムの開発
- 開発した「次々世代分子理論」を化学や物理の様々な現象に適用することで、一般的な指導原理を実験に先行して確立する

# NTChem

日本発の理論+プログラム+スパコンで理論先導の科学, 理論予測から実験への流れを牽引!

従来のプログラムと比べて, 数倍から数百倍の高速化と大規模計算が目標  
扱える分子系が周期表全体へと格段に広がるのも特徴

## HF & DFT

- 閉殻系, 開殻系 (1+2成分)
- 各種汎関数:  
LDA, GGA, 混成GGA, LC-GGA
- 各種数値グリッド
- エネルギー微分 (1+2成分)

## 高速SCF計算

- 積分計算高速化
- 各種収束法:  
DIIS, 2次収束, 直接最小化
- 各種SCF技術:  
非整数占有数, レベルシフト
- 各種初期分子軌道

## Order-N化法

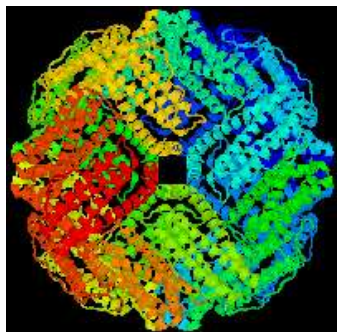
- クーロン積分Order-N化(GFC)
- Resolution of Identity(RI)法
- Dual-level DFT
- 局在化MP2法
- 領域分割法 (ONIOM, SPOT)

## その他

- 波動関数法 (MP2, CIS)
- 擬ポテンシャル (ECP, MCP)
- 各種相対論的方法:  
DKn, RESC, RA, スピン-軌道効果
- NMR, EPR, 磁化率計算 (1+2成分)

# 本プロジェクトの優位性

- 従来の分子理論では少なくとも系の大きさの3乗に比例して計算時間が必要になるので、大規模な分子系の理論計算は不可能である。
- われわれの大規模分子理論は、系のサイズの増加に対し線形スケーリングを示す理論である。十分な精度も持っている。得られている結果から外挿すると、インスリン6量体が20個集合したような**6000残基のタンパク質 (約十萬原子)**の**DFT計算を十萬coreの並列化により2時間**で計算することが可能。この大きさの分子計算は、これまでに第一原理計算されたもっとも大きな分子の20倍にあたり、その計算が32coreで3週間程度かかっていることを考えると大規模分子シミュレーションに大きな進歩をもたらすことになる。
- 機能性分子の自己集合過程の**10nsにおよぶ第一原理分子ダイナミクス計算を64万coreの並列化により2週間で実現**することが可能。数百～数千原子系の第一原理ダイナミクス計算という未踏の理論計算が実現できる。



6000残基タンパク質(約10万原子)の  
第一原理計算: 2時間/10万core  
従来, 300残基計算が3週間/32core

機能性分子の自己集合過程の第一  
原理MD計算(10ns):2週間/64万core



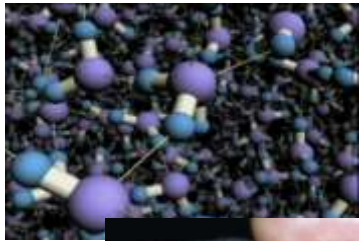
# 本プロジェクトの独自性

## 研究領域を横断した融合研究を積極的に実施

- 計算科学研究機構では大学での研究とは異なり、学科を超えた研究分野にとらわれない領域横断の研究が実施できる
- 横断・融合研究に誘発される新しい分子理論とアルゴリズムのブレークスルーを達成
- 理論分子科学の枠を超えた革新的シミュレーション (理論+アルゴリズム+計算機) を基盤にした新規研究領域の創成
- 横断・融合研究により初めて見えてくる新しい科学的知見の発見

# 分子科学シミュレーション

日本発の分子理論+ソフトウェア+スパコンで分子科学の革新を狙う！



機能性ナノ分子

## 大規模分子理論

- ・線形スケールング
- ・領域分割法
- ・局在化アプローチ

## 高精度分子理論

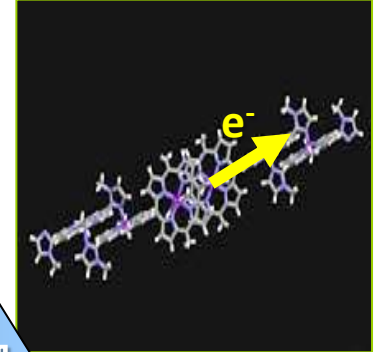
- ・電子相関・QMC
- ・励起状態・線形応答
- ・相対論効果

## ソフトウェア

- ・NTChem
- ・SMITH2
- ・R4QMC

## 分子設計

## 反応制御



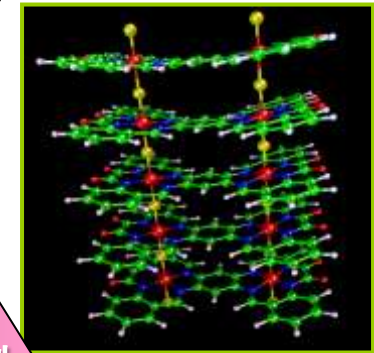
人口光合成

## ダイナミクス理論

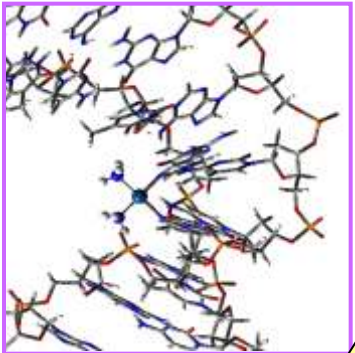
- ・QM/MM・メタCP
- ・励起状態ダイナミクス
- ・非断熱, SO遷移

## 物性計算理論

- ・電気伝導・電気性質
- ・磁気性質・固体
- ・溶媒効果・振動状態



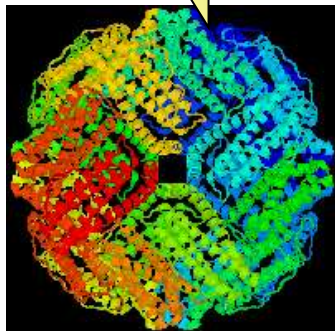
分子エレクトロニクス



たんぱく質の動的機能発現

## 機能発現

## 物性予測



## 期待される成果

6000残基タンパク質(約10万原子)の  
第一原理計算: 2時間/10万core  
従来, 300残基計算が3週間/32core

機能性分子の自己集合過程の第一  
原理MD計算(10ns):2週間/64万core

