

地球シミュレータ産業利用シンポジウム2008
2008年9月5日

フラグメント分子軌道プログラムABINIT-MP による超大規模バイオシミュレーション

中野 達也

国立医薬品食品衛生研究所／東京大学生産技術研究所



RSS 21
Revolutionary Simulation Software



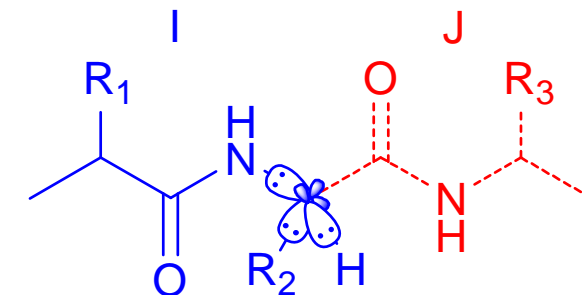
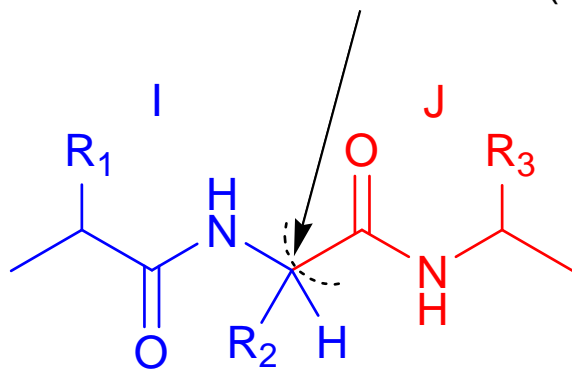
フラグメント分子軌道 (FMO) 法とは?

- 京都大学大学院薬学研究科 北浦和夫 教授により提唱
 - K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, “Pair interaction molecular orbital method: an approximate computational method for molecular interactions”, *Chem. Phys. Lett.* **312**, 319-324 (1999).
 - K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, “Fragment molecular orbital method: an approximate computational method for large molecules”, *Chem. Phys. Lett.* **313**, 701-706 (1999).
- 分子をフラグメントに分割し、フラグメントのモノマー、ダイマー、トリマーの分子軌道計算から分子全体を計算する近似計算法
- 全体を一度に扱う必要がない
- フラグメント単位で並列処理が可能
- フラグメント間相互作用エネルギー (Inter-Fragment Interaction Energy, IFIE)解析が可能

どうやって分子をフラグメントに分割するか？

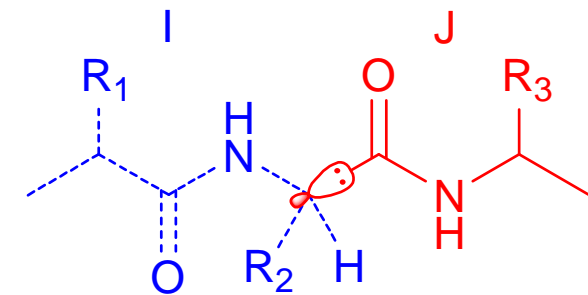
- C-H間の原子間距離を1.09Åに固定したメタン分子の局在化軌道(sp^3 混成軌道に相当)を使った射影演算子を用いて、フラグメントの分子軌道をフラグメント内に局在化する。
- sp^3 炭素でフラグメントに分割
- sp^2 炭素や sp^3 ケイ素で分割することも可能

Bond Detached Atom (BDA)



フラグメントIの変分空間

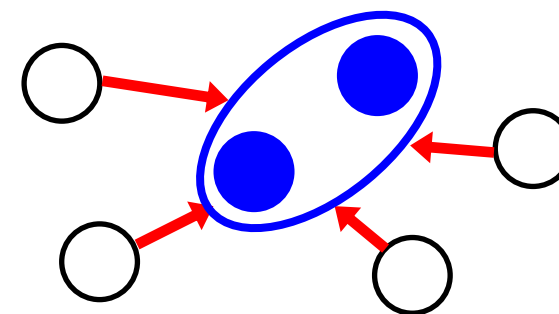
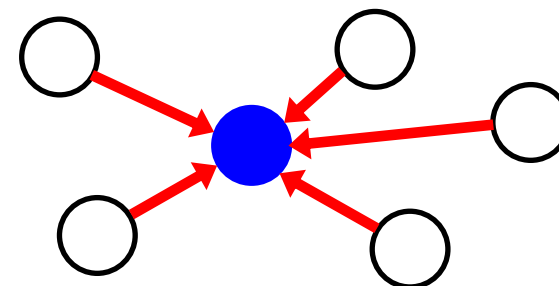
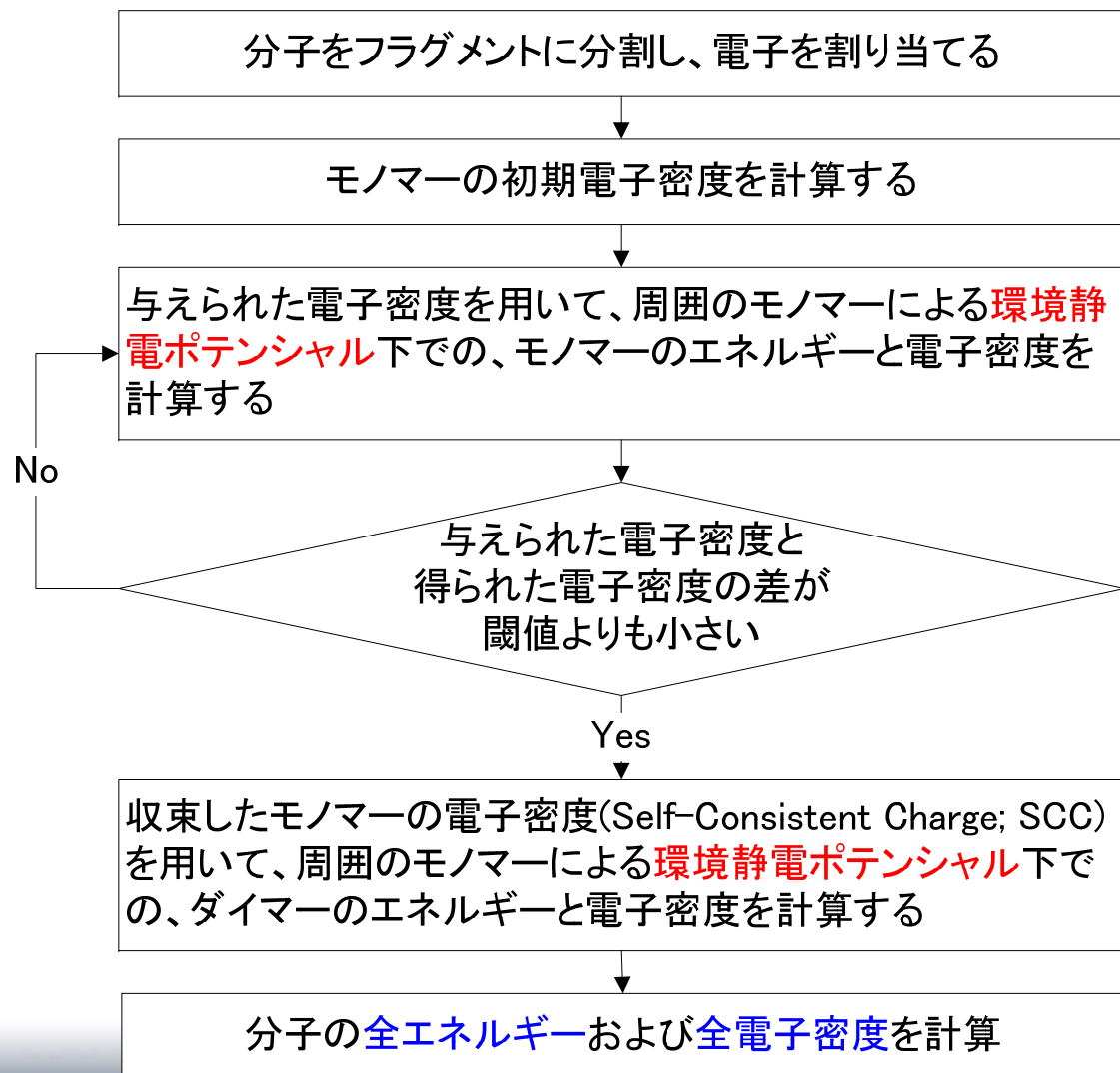
フラグメントIの形式電荷: +1
BDAの炭素原子の核荷電: 5



フラグメントJの変分空間

フラグメントJの形式電荷: -1
BDAの炭素原子の核荷電: 1

FMO法計算の流れ



FMO2法による分子の全エネルギー および全電子密度

$$E_I = E_I^{\text{elec}} + \sum_{\substack{A>B \\ A,B \in I}} \frac{Z_A Z_B}{|\mathbf{B} - \mathbf{A}|}$$

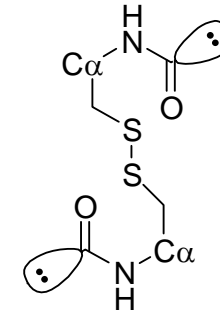
$$E_{IJ} = E_{IJ}^{\text{elec}} + \sum_{\substack{A>B \\ A,B \in IJ}} \frac{Z_A Z_B}{|\mathbf{B} - \mathbf{A}|}$$

$$E \cong \sum_{I>J} E_{IJ} - (N_f - 2) \sum_I E_I$$

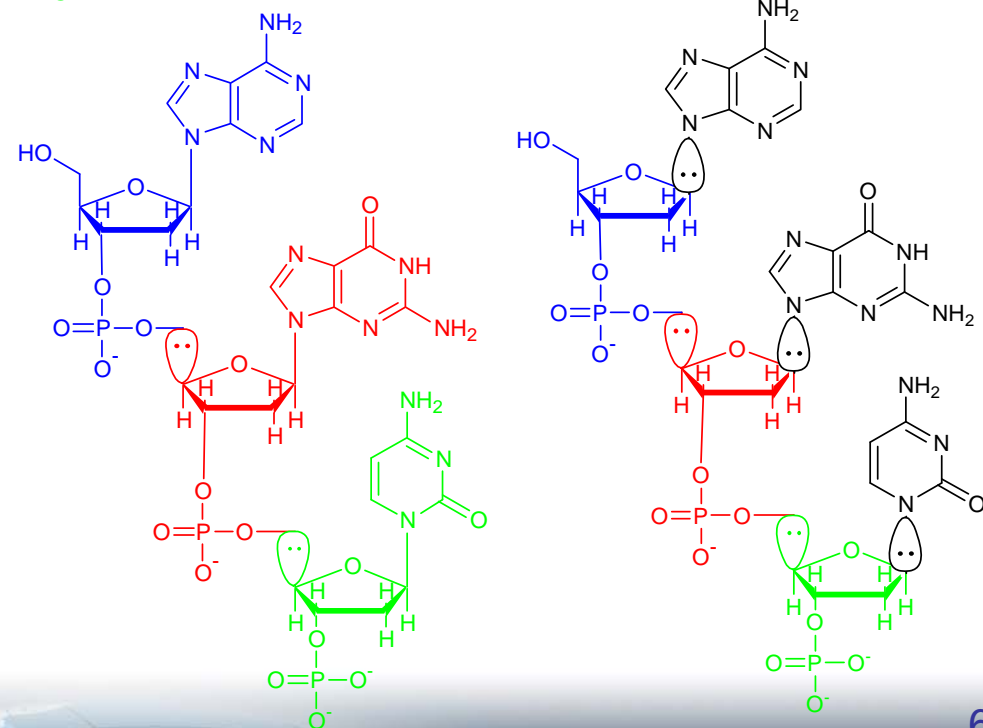
$$\rho(\mathbf{r}) \cong \sum_{I>J} \rho_{IJ}(\mathbf{r}) - (N_f - 2) \sum_I \rho_I(\mathbf{r})$$

タンパク質・DNAのフラグメント分割

- タンパク質はCaの位置でフラグメントに分割



- ペプチドを2残基単位でフラグメントに分割した場合の、全エネルギーの誤差は、数kcal/mol以下
- タンパク質・DNA(RNA)については自動分割可能



高速化のための近似

- 環境静電ポテンシャル計算の高速化
esp-aoc近似 (Mulliken AO population)

$$v_{pq}^L \cong \sum_{r \in L} (\mathbf{P}^L \mathbf{S}^L)_{rr} (pq, rr) \quad \text{for } R_{\min}(X, L) \geq L_{\text{aoc}}$$

- esp-ptc近似 (Mulliken atomic charge)

$$v_{pq}^L \cong \sum_{A \in L} \langle p | (Q_A / |\mathbf{r} - \mathbf{A}|) | q \rangle \quad \text{for } R_{\min}(X, L) \geq L_{\text{ptc}}$$

$$Q_A = \sum_{r \in A} (\mathbf{P}^L \mathbf{S}^L)_{rr}$$

- ダイマー計算の高速化 (dimer-es近似)

$$E'_{IJ} \cong E'_I + E'_J + \text{Tr}(\mathbf{P}^I \mathbf{u}^J) + \text{Tr}(\mathbf{P}^J \mathbf{u}^I) + \sum_{pq \in I} \sum_{rs \in J} \mathbf{P}_{pq}^I \mathbf{P}_{rs}^J (pq, rs)$$

➡ 演算量が $O(N^3)$ から $O(N^2)$ に減少
これらの近似を導入したエネルギー勾配も計算可能

Inter-Fragment Interaction Energy (IFIE)

- FMO法の全エネルギー E を書き換えると、全エネルギーをフラグメント間相互作用エネルギー ΔE_{IJ} と、モノマーのエネルギー E_I から周囲のフラグメントとの静電相互作用エネルギーを除いたエネルギー E'_I の和で表せる。

$$E = \sum_{I>J} \Delta \tilde{E}_{IJ} + \sum_I E'_I \quad \text{IFIE解析の基本式}$$

$$\Delta \tilde{E}_{IJ} = (E'_{IJ} - E'_I - E'_J) + \text{Tr}(\Delta \mathbf{P}^{IJ} \mathbf{V}^{IJ})$$

$$E'_I = E_I - \text{Tr}(\mathbf{P}^I \mathbf{V}^I)$$

$$E'_{IJ} = E_{IJ} - \text{Tr}(\mathbf{P}^{IJ} \mathbf{V}^{IJ})$$

FMO-MP2法による分散力の取り込み

- 生体系分子においては、静電相互作用や水素結合の他に分散力に基づいたファンデルワールス相互作用やCH/ π 相互作用が重要
- Hartree-Fock法では、分散力を記述できない。
- 一般的な密度汎関数法(B3LYP等)は、分散力の記述が不得手
- 生体系分子の計算に分散力を取り込むには、Size-consistencyと計算コストの面から2次のMøller-Plesset摂動論(MP2法)が有用
- 立教大 望月祐志准教授による高速なMP2計算エンジンの開発
⇒ **生体高分子のMP2計算が実用に!**

Y. Mochizuki, T. Nakano, S. Koikegami, S. Tanimori, Y. Abe, U. Nagashima, K. Kitaura, *Theor. Chem. Acc.* **112**, 442-452 (2004).

Y. Mochizuki, S. Koikegami, T. Nakano, S. Amari, K. Kitaura, *Chem. Phys. Lett.* **396**, 473-479 (2004).

Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, *Chem. Phys. Lett.* **457**, 396-403 (2008).

- IFIEのMP2法による補正 $\Delta E_{IJ}^{\text{MP2}}$ (分散力に相当)

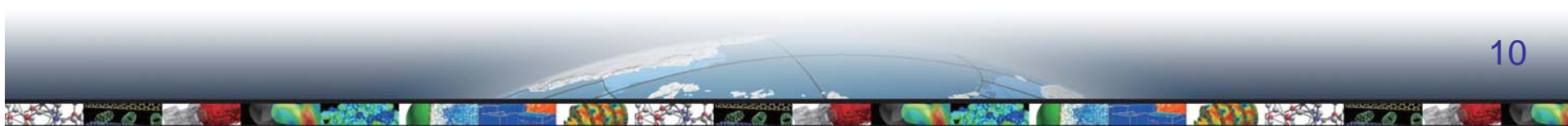
$$\Delta E_{IJ}^{\text{MP2}} = E_{IJ}^{\text{MP2}} - E_I^{\text{MP2}} - E_J^{\text{MP2}}$$

FMO法の実装系

- 産業技術総合研究所 Fedorovら ⇒ GAMESS/FMO
- 豊橋技術科学大学 関野ら ⇒ NWChem/FMO
- RSS21/JST-CRESTプロジェクト ⇒ ABINIT-MP

ダウンロード(無償)

<http://www.ciss.iis.u-tokyo.ac.jp/rss21/>
(FMO-HF, FMO-MP2, IFIE解析, VISCANA)



FMO法プログラム ABINIT-MP

- RSS21およびJST-CREST(神戸大 田中成典教授)プロジェクトにより開発

- 開発言語はFortran

- MPIを用いた、フラグメント内・フラグメント間の二階層並列化

- 積分計算のベクトル化

⇒ 小原法の漸化式をループ展開

S. Obara and A. Saika, *J. Chem. Phys.* **84**, 3963-3974 (1986).

⇒ 初期積分は4項Taylor展開+漸近展開

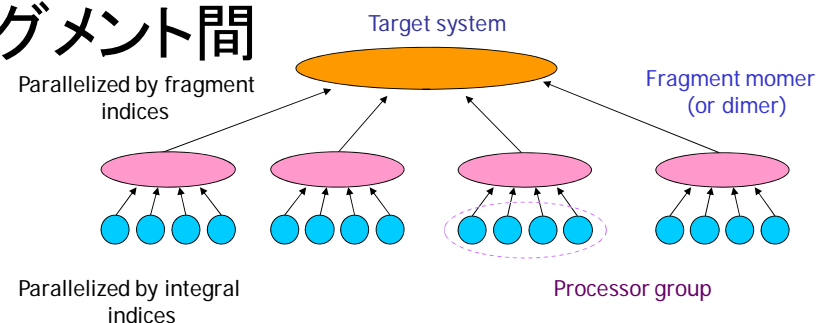
- Fock行列計算のベクトル化

⇒ 止まり木アルゴリズムを採用

Y. Mochizuki, M. Matsumura, T. Yokura, Y. Hirahara, T. Imamura, *J. Nucl. Sci. Tech.* **39**, 195-199 (2002).

- MP2計算のベクトル化

⇒ 最深部はLevel 3 BLAS (DGEMM)



DGEMMによるMP2エンジンのベクトル化

Loop over ij -batch ! size depending on available memory

Loop over s ! to be parallelized

Loop over r

Preparing (pq,rs) ! for canonical pq -pair

① Forming (iq,rs) ! DGEMM, fixed rs , running over p

② Forming (ia,rs) ! DGEMM, fixed rs , running over q

③ Forming (ia,js) ! DGEMM, fixed s , direct-product for fixed r

End of loop over r

④ Forming (ia,jb) ! DGEMM, direct-product for fixed s

End of loop over s ! all-reduce operation as barrier

Calculate partial MP2 energy with respect to ij -batch

End of loop over ij -batch



地球シミュレータ
640ノード
(5,120 VPU)
8GFLOPS/VPU

ステップ②～③は実測5.7GFLOPS(地球シミュレータ1VPU
(8GFLOPS)で測定)

Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, "Large scale FMO-MP2 calculations on a massively parallel-vector computer", *Chem. Phys. Lett.* **457**, 396-403 (2008).

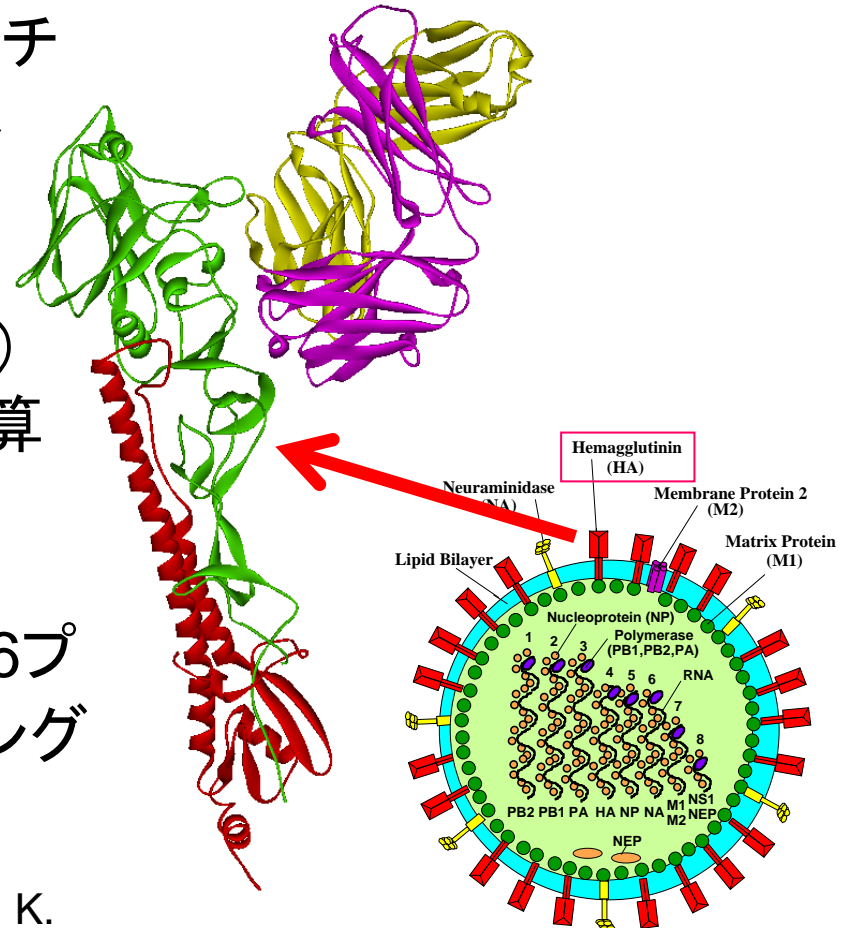
ABINIT-MPのベクトル化率および並列化率

- (Trp)₁₂₇His a helix, FMO-MP2/6-31G
 - 128ノード(1,024プロセッサ) 1.65 TFLOPS (20.2%)
 - 256ノード(2,048プロセッサ) 3.22 TFLOPS (19.7%)
- エストロゲン受容体タンパク質リガンド結合部位(241残基)のFMO-MP2/6-31G計算が128ノードで20.9分
 - ⇒ **低分子化合物のab initio MO計算と同程度の計算時間で, タンパク質のab initio FMO計算が可能に!**

VPUs	N _p	Time (s)	Acceleration Factor	VPU (GFLOPS)	Total (TFLOPS)	Peak Ratio (%)
(Trp) ₁₂₇ His	16					
	1024	1309.2	1.00	1.61	1.65	20.2
	2048	672.3	1.95	1.57	3.22	19.7
EST-ER	32					
	1024	1254.6	1.00	1.09	1.11	13.6
	2048	793.5	1.58	0.86	1.76	10.7
	3072	677.4	1.85	0.67	2.06	8.4
	4096	570.4	2.20	0.60	2.45	7.5

インフルエンザウイルス抗原-抗体 タンパク質複合体のFMO-MP2計算 (1/2)

- インフルエンザウイルス ヘマグルチニン(HA)タンパク質(抗原)と抗体 Fabフラグメントの複合体
 - ⇒ 抗原: HA1(緑)、HA2(赤)
 - 抗体: H鎖(ピンク)、L鎖(黄)
- 世界最大のFMO-MP2/6-31G計算
 - ⇒ **921残基、14,086原子、78,390基底**
- 地球シミュレータ 512ノード(4,096プロセッサ)でのベンチマークタイミング
 - ⇒ **53.4分!**



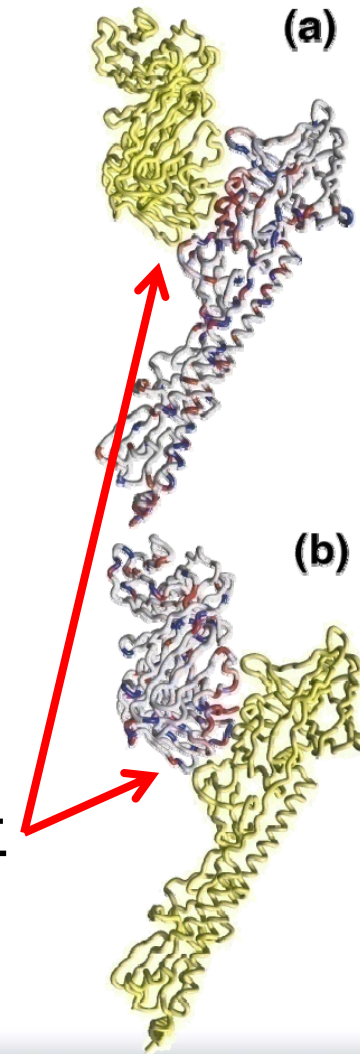
インフルエンザ
ウイルス模式図

Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, *Chem. Phys. Lett.* **457**, 396-403 (2008).

インフルエンザウイルス抗原-抗体 タンパク質複合体のFMO-MP2計算 (2/2)

Timing data of influenza HA antigen-antibody @ FMO-MP2/6-31G

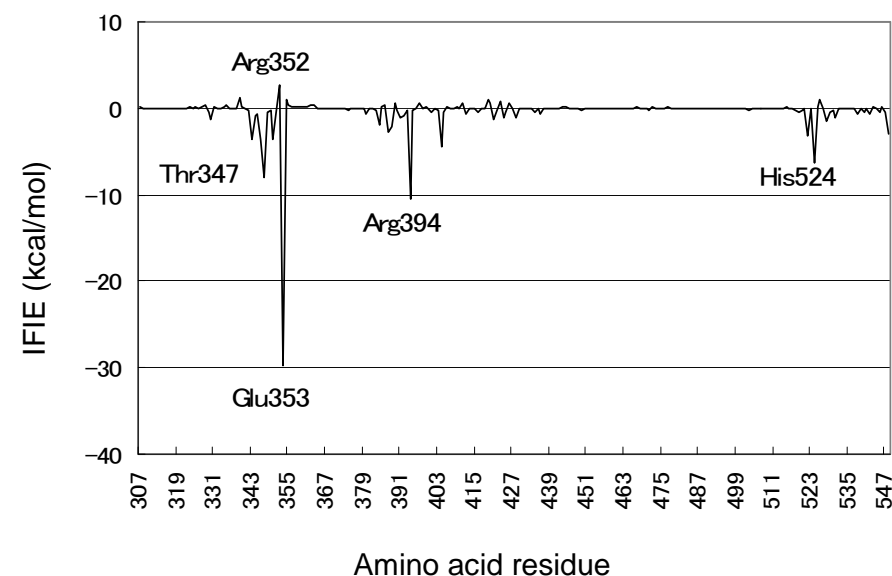
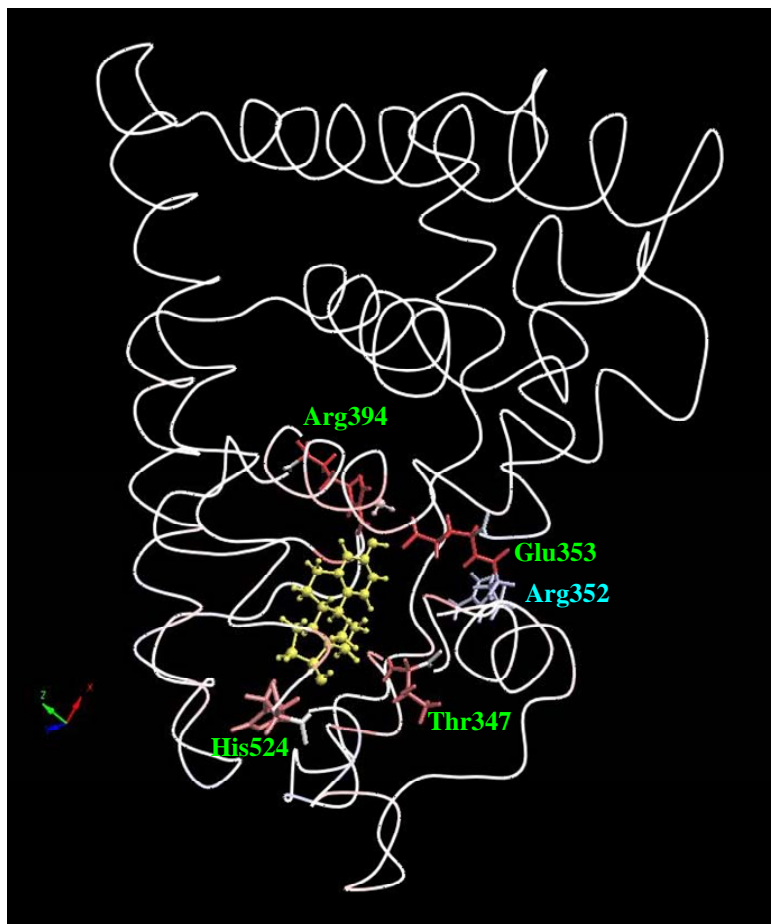
VPU@ ES	Time (s)	Accel.	TFLOPS	Effic. (%)	Rel. cost
FMO-MP2					
1024	10084.6		1.19	14.5	2.42
2048	5486.2	1.84	2.19	13.3	2.57
3072	3927.1	2.57	3.06	12.4	2.56
4096	3204.4	3.15	3.75	11.5	2.66
FMO-HF					
1024	4164.9		1.36	16.6	
2048	2131.7	1.95	2.67	16.3	
3072	1533.5	2.72	3.72	15.2	
4096	1205.9	3.45	4.75	14.5	



- **HF計算の約2.7倍の計算時間でMP2計算が可能**
- MP2-IFIEによって抗原-抗体反応の様相を詳細に解析予定 (JST-CREST@神戸大)
 - ⇒ 新型鳥インフルエンザウイルスの変異予測に貢献も!

K. Fukuzawa et al., *J. Phys. Chem. B* 110, 16102–16110 (2006).

Inter-fragment interaction energy (IFIE)



Attractive interaction: Glu353, His524,
Arg394, (water molecule), etc.

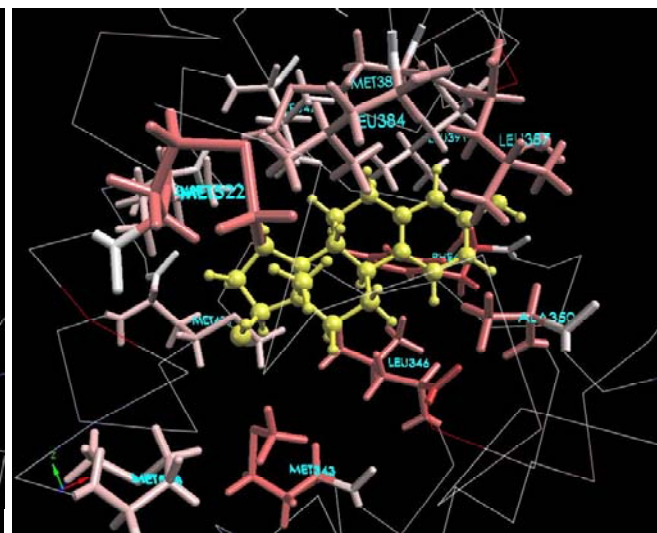
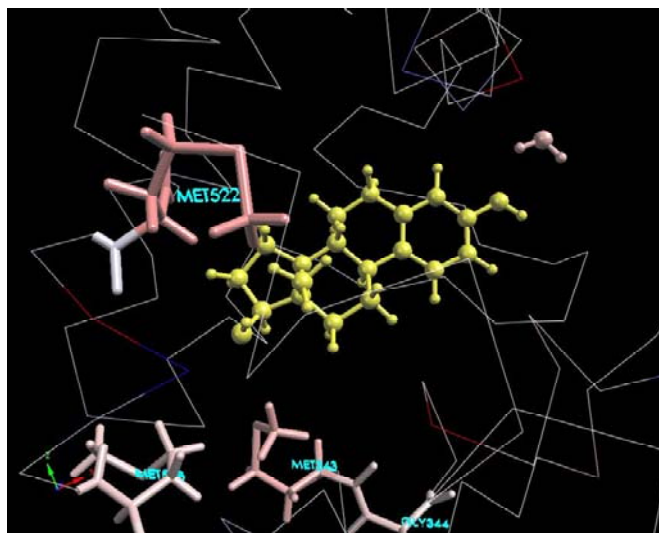
Repulsive interaction: Arg352, etc.

[MODEL1, FMO-MP2/6-31G]

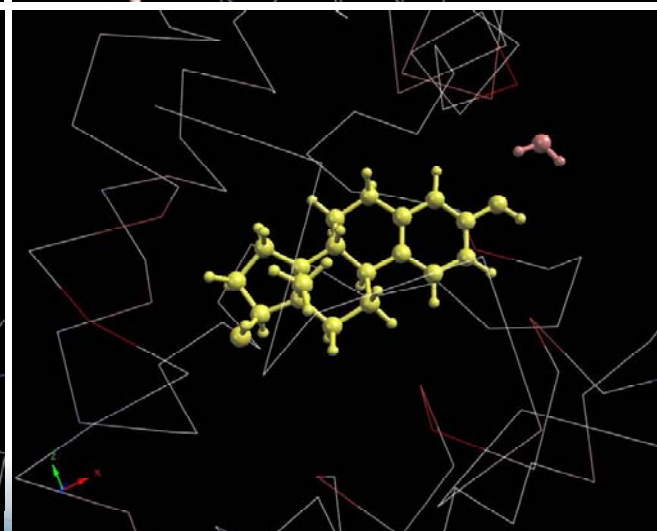
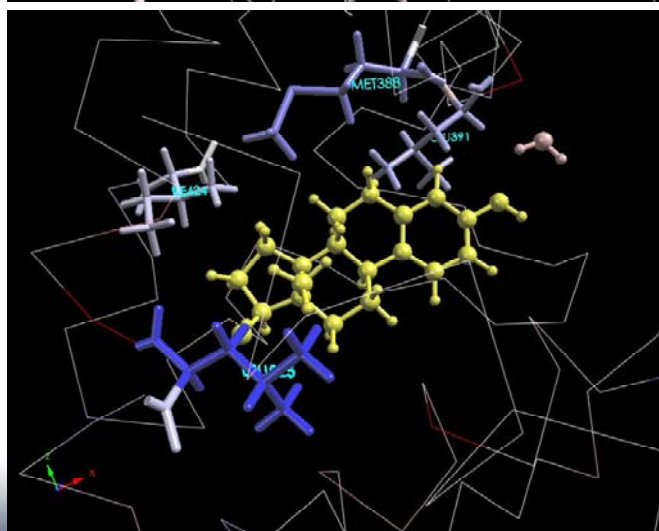
FMO-HF/6-31G

FMO-MP2/6-31G

Attractive residues
IFIEE < -1kcal/mol



Repulsive residues
IFIEE > +1kcal/mol



Y. Mochizuki et al., *Chem. Phys Lett.* **410**, 247-253 (2005).

◇軌道レベルの相互作用解析

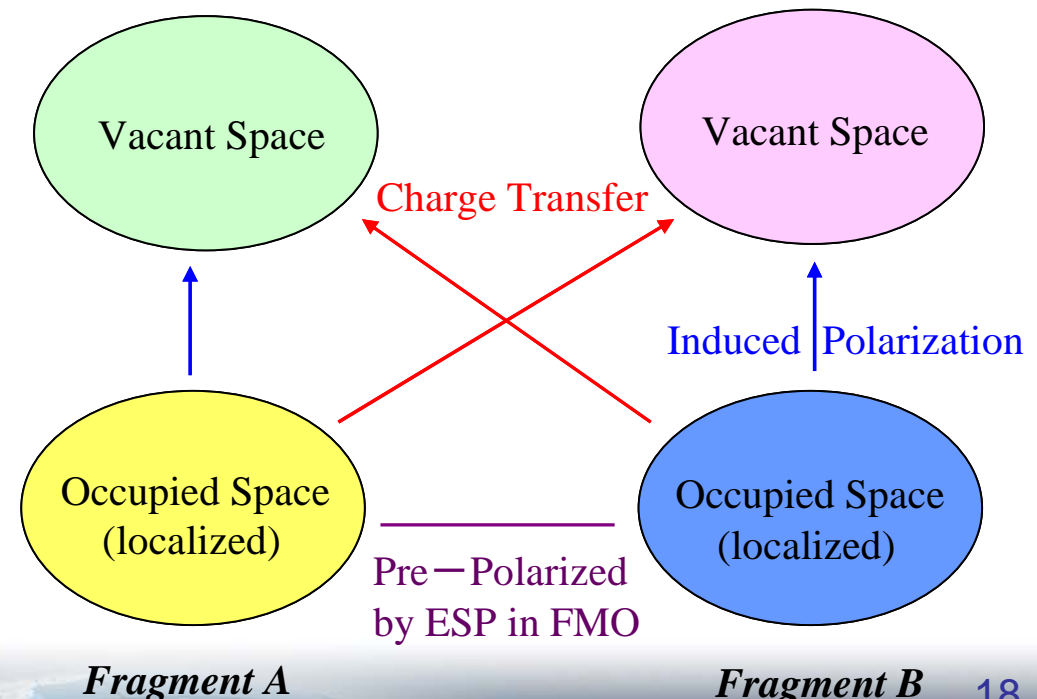
水素結合 \leftrightarrow O上の孤立電子対 \rightarrow O-Hの反結合性軌道

\Rightarrow 軌道レベルでの化学的な描像

モノマーの“集積”に対するCERF緩和エネルギー計算

\Rightarrow 解析と可視化

- #1 モノマー計算は通常のFMO
- #2 リガンドと重要な残基を選択
- #3 局在化、重み付Lowdin直交化
- #4 CERFによる多体緩和の導入
(N**4、多電子論的な近似)
- #5 軌道毎の緩和エネルギー評価
- #6 供与/受容の相互作用軌道
(可視化が有効)

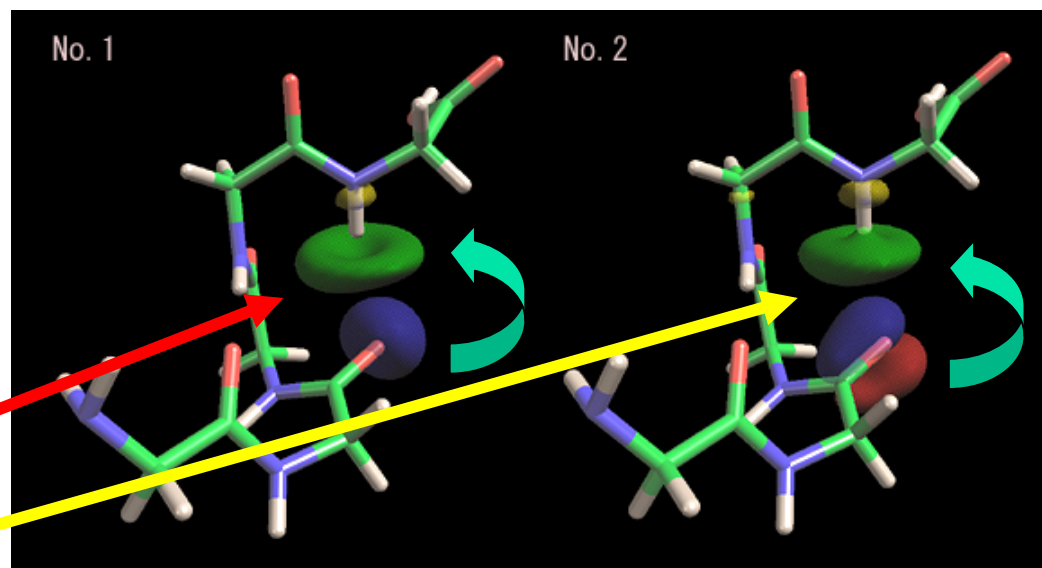


CAFI: Glycine pentamer

分子内水素結合の解析例

- 分子構造: HF/6-31G**最適化構造
- FMO計算: FMO-HF/6-31G** (一残基分割)

水素結合(CT)
による安定化は
それぞれ
-1.9kcal/mol,
-1.4kcal/mol



電子の移動
方向及び電
荷の移動量
まで解析で
きる。

Table CAFI energies (in hartree) for Gly3 – Gly5 hydrogen bonding in glycine pentamer

No.	Energy (CT)	Energy (POL)	Donor orbital	Acceptor orbital	Occupation number
1	-0.003055	0.000023	σ lone-pair	σ^*_{NH}	0.00245
2	-0.002260	0.000057	π lone-pair	σ^*_{NH}	0.00332

FILM: Fragment Interaction based on Local MP2

T. Ishikawa, Y. Mochizuki, S. Amari, T. Nakano, H. Tokiwa, S. Tanaka, K. Tanaka,
"Fragment interaction analysis based on local MP", *Theor. Chem. Acc.* **118**, 937-945 (2007).

IFIE with (L)MP2 method

$$\Delta\tilde{E}_{IJ}^{\text{corr}} = \Delta\tilde{E}_{IJ}^{\text{HF}} + \Delta E_{IJ}^{(\text{L})\text{MP2}}$$

Canonical MP2

$$\Delta E_{IJ}^{\text{MP2}} = E_{IJ}^{\text{MP2}} - E_I^{\text{MP2}} - E_J^{\text{MP2}}$$

Local MP2 (LMP2)

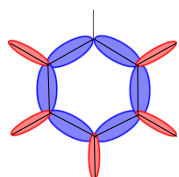
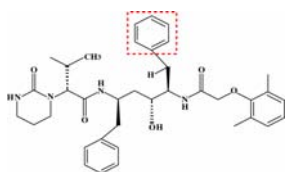
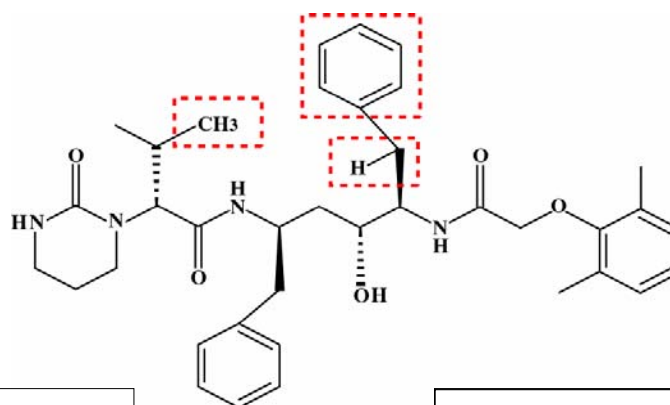
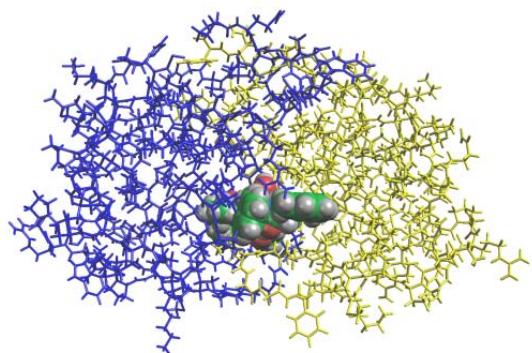
$$\Delta E_{IJ}^{\text{LMP2}} = \sum_{i \in I} \sum_{j \in J} \varepsilon_{ij}$$

フラグメント間の分散力をペア相関エネルギーの和で計算
⇒置換基単位のファンデルワールス相互作用や
CH/ π 相互作用の見積もりが可能

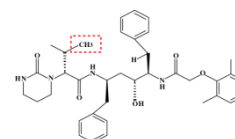
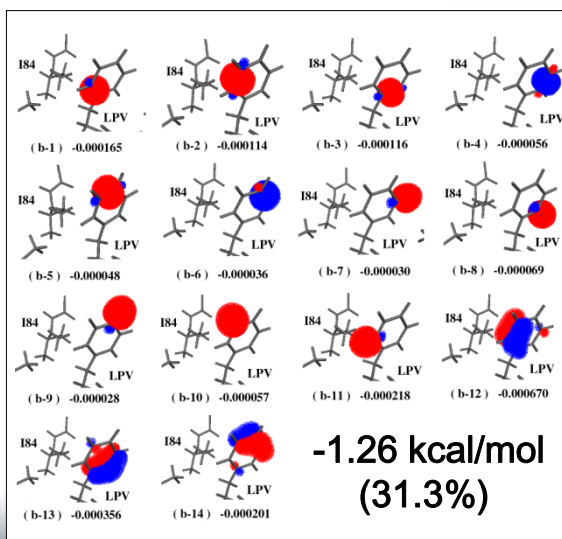
HIV-1 Protease – Lopinavir (FILM)

FILM analysis of Isoleucine 84 of HIV-1 Protease – Lopinavir

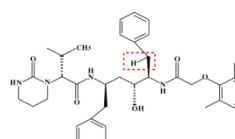
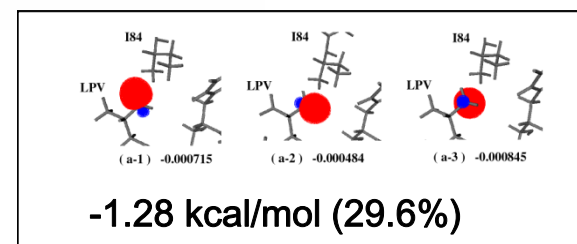
T. Ishikawa, Y. Mochizuki, S. Amari, T. Nakano, S. Tanaka, K. Tanaka, "An application of fragment interaction analysis based on local MP2", *Chem. Phys. Lett.* **463**, 189-194 (2008).



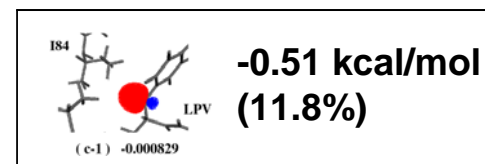
C-C σ bonds: 6
C-H σ bonds: 5
C-C π bonds: 3



C-H σ bonds: 3

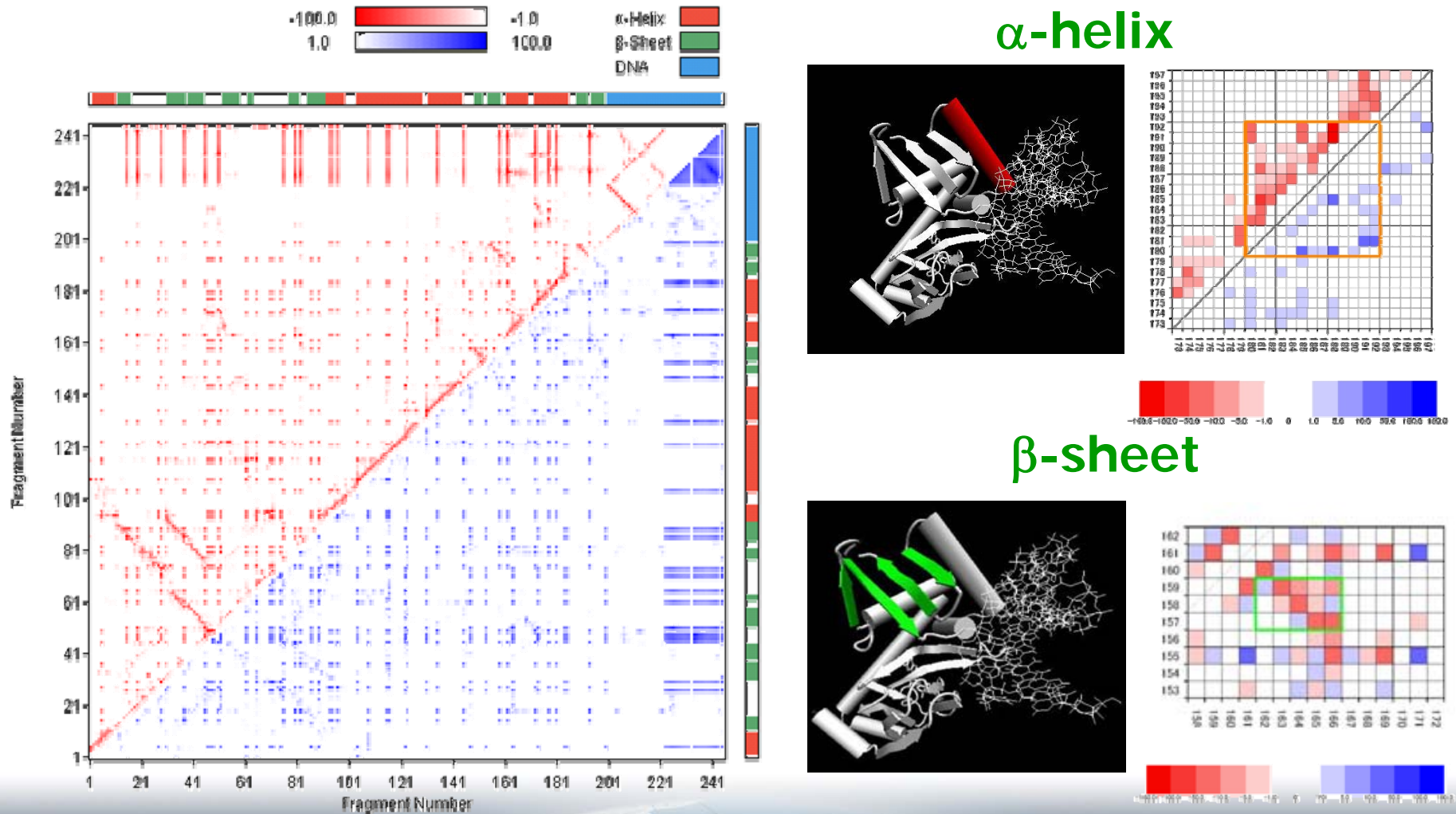


C-H σ bond: 1



IFIE-MAP of CRP-DNA complex

I. Kurisaki et al., Biophys. Chem. 130, 1-9 (2007).

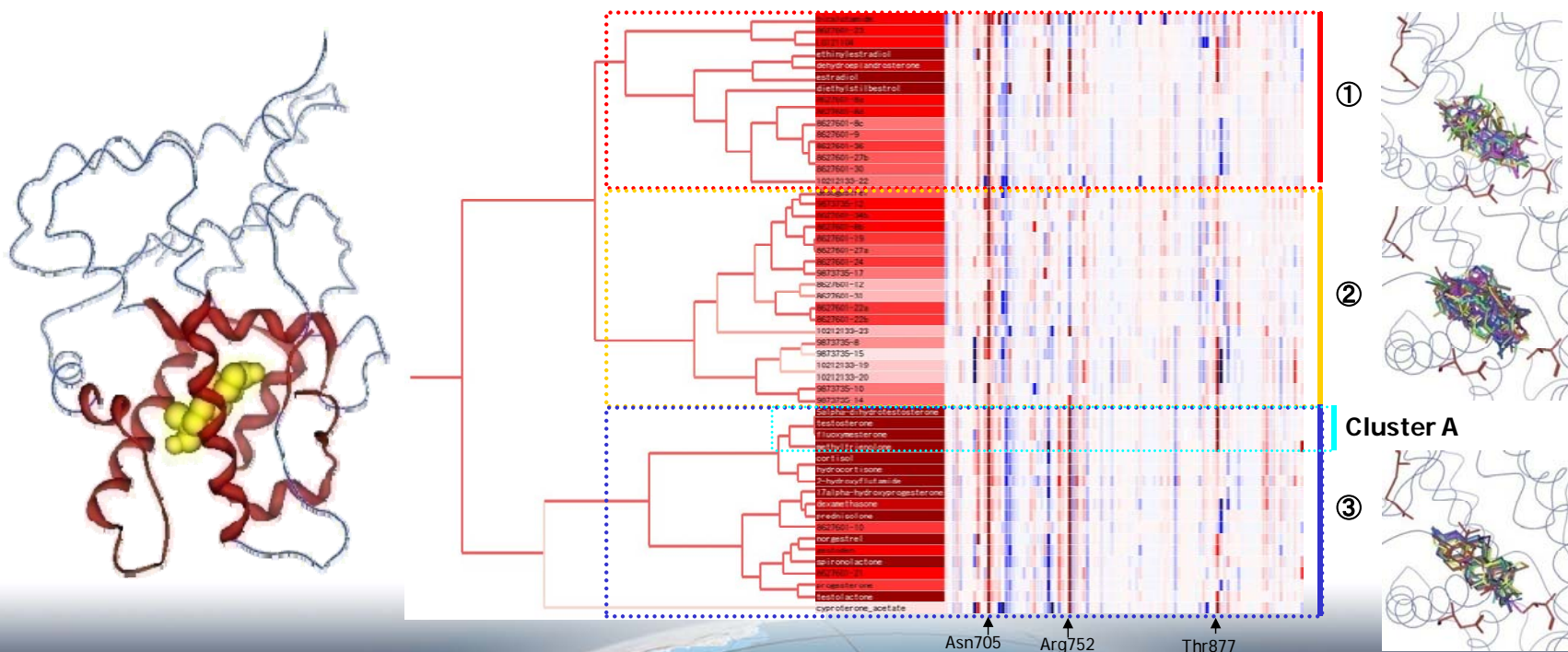


地球シミュレータのバーチャルリガンドスクリーニング(VLS)への応用

- VISCANA: タンパク質の各アミノ酸残基と化合物間の相互作用エネルギー(IFIE)のパターンに基づいた階層的クラスタ解析
S. Amari et al., *J. Chem. Inf. Model.* **46**, 221-230 (2006).

⇒ 地球シミュレータの高速性を活用すれば、高速かつ高精度なVLSが可能に!

FMO-MP2/6-31G法を用いたアンドロゲン受容体(AR)の VISCANAによる解析例

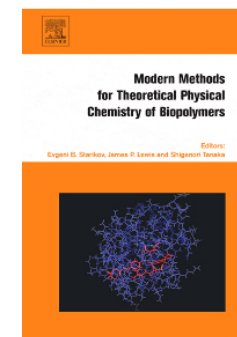


FMO法のレビュー #1

- 中野達也, 望月祐志, 甘利真司, 小林将人, 福澤薫, 田中成典, "フラグメント分子軌道法に基づいた生体巨大分子の電子状態計算の現状と今後の展望", *J. Comput. Chem. Jpn.* **6**, 173-184 (2007).
- 福澤薫, 中野達也, 加藤昭史, 望月祐志, 田中成典, "フラグメント分子軌道法による生体高分子の応用計算", *J. Comput. Chem. Jpn.* **6**, 185-198 (2007).
<http://www.sccj.net/publications/JCCJ/v6n3/indexj.html>
- 北浦和夫, "フラグメント分子軌道法の概要", CBI学会 計算化学研究会ワークショップ資料 (2008/5/22)
http://www.cbi.or.jp/cbi/jigyoku/FMO_gaiyo.pdf
- 佐藤文俊・中野達也・望月祐志 編著, プログラムで実践する生体分子量子化学計算, 森北出版(9月刊行)

FMO法のレビュー #2

- D. G. Fedorov, K. Kitaura, “Theoretical development of the fragment molecular orbital (FMO) method”, in Modern methods for theoretical physical chemistry of biopolymers, E. B. Starikov, J. P. Lewis, S. Tanaka, Eds., Elsevier (2006) pp. 3-38.
- T. Nakano, Y. Mochizuki, K. Fukuzawa, S. Amari, S. Tanaka, “Developments and applications of ABINIT-MP software based on the fragment molecular orbital method”, in Modern methods for theoretical physical chemistry of biopolymers, E. B. Starikov, J. P. Lewis, S. Tanaka, Eds., Elsevier (2006) pp. 39-52.
- D. G. Fedorov, K. Kitaura, “Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method”, *J. Phys. Chem. A*, **111**, 6904-6914 (2007) .



謝辞

本研究は、文部科学省次世代IT基盤構築のための研究開発プログラム「革新的シミュレーションソフトウェアの研究開発(RSS21)」プロジェクトおよびJST-CREST「フラグメント分子軌道法による生体分子計算システムの開発」プロジェクトの支援を受けています。また、海洋研究開発機構の支援により、地球シミュレータを利用しています。

