「計算生命科学の基礎 II」 2015年11月18日

タンパク質の量子化学計算

神戸大学大学院システム情報学研究科
計算科学専攻　計算生物学講座
田中　成典

KOBE UNIVERSITY

Contents

- 量子化学の基礎のおさらい
- タンパク質の量子化学計算概観
- フラグメント分子軌道(FMO)法の概要
- さまざまなFMOソフトウェア
- インシリコ創薬への応用
- その他の応用事例(光励起系など)

Reference
S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, K. Fukuzawa, "Electron-Correlated Fragment-Molecular-Orbital Calculations for Biomolecular and Nano Systems", Phys. Chem. Chem. Phys. 16 (2014) pp. 10310-10344. DOI: 10.1039/c4cp00316k

Hartree-Fock(HF)法(1)

- 分子軌道法の基礎となる「平均場近似」

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_i,\cdots\mathbf{r}_j,\cdots,\mathbf{r}_N)=-\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_j,\cdots\mathbf{r}_i,\cdots,\mathbf{r}_N)$$

反対称性

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\phi_1(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_1)\\ \vdots & \ddots & \vdots\\ \phi_1(\mathbf{r}_N) & \cdots & \phi_N(\mathbf{r}_N)\end{vmatrix}$$

Slater行列式
φ:1電子軌道

2電子の場合:

$$\Psi(r_1,r_2)=\frac{1}{\sqrt{2}}\begin{vmatrix}\varphi_1(r_1) & \varphi_1(r_2)\\ \varphi_2(r_1) & \varphi_2(r_2)\end{vmatrix}=\frac{1}{\sqrt{2}}\{\varphi_1(r_1)\varphi_2(r_2)-\varphi_1(r_2)\varphi_2(r_1)\}$$

ハミルトニアン:

$$\hat{H}=-\sum_{i=1}^{N_{elec}}\frac{1}{2}\nabla_i^2-\sum_{i=1}^{N_{elec}}\sum_{I=1}^{N_{nuc}}\frac{Z_I}{|r_i-R_I|}+\sum_{i>j}^{N_{elec}}\frac{1}{|r_i-r_j|}+\sum_{I>J}^{N_{nuc}}\frac{Z_IZ_J}{|R_I-R_J|}$$

Hartree-Fock(HF)法(2)

エネルギー:

$$E = \frac{\int d\tau \Psi^* \hat{H} \Psi}{\int d\tau \Psi^* \Psi} = 2\sum_{i=1}^{n} h_i + \sum_{i,j=1}^{n} \left(2J_{ij} - K_{ij} \right)$$

$$h_i = \int d^3\mathbf{r}\, \phi_i^*(\mathbf{r}) \left\{ -\frac{1}{2}\nabla^2 + V_{\mathrm{ne}}(\mathbf{r}) \right\} \phi_i(\mathbf{r})$$

$$J_{ij} = \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \phi_i^*(\mathbf{r}_1)\phi_i(\mathbf{r}_1) \frac{1}{r_{12}} \phi_j^*(\mathbf{r}_2)\phi_j(\mathbf{r}_2)$$

$$K_{ij} = \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1) \frac{1}{r_{12}} \phi_j^*(\mathbf{r}_2)\phi_i(\mathbf{r}_2)$$

ここで、1電子演算子:

$$\hat{h} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i \sum_I \frac{Z_I}{|r_i - R_I|}$$

など

Lagrange未定乗数法:

$$L\left(\{\varphi_i\}\right) = E\left(\{\varphi_i\}\right) - \sum_{i,j} \varepsilon_{j,i} \left(\int \varphi_i^* \varphi_j - \delta_{i,j} \right)$$

Hartree-Fock（HF）法（3）

拘束条件付き極小化：$\delta L/\delta\varphi=0$より

$$\hat{F}\phi_i = \varepsilon_i\phi_i$$ ハートリー・フォック方程式

$$\hat{F} = \hat{h} + \sum_j (2\hat{J}_j - \hat{K}_j)$$ フォック演算子

$$\hat{J}_j(\mathbf{r}_1)\phi_i(\mathbf{r}_1) = \int d^3\mathbf{r}_2 \phi_j^*(\mathbf{r}_2)\phi_j(\mathbf{r}_2)\frac{1}{r_{12}}\phi_i(\mathbf{r}_1)$$ クーロン演算子

（2電子演算子）

$$\hat{K}_j(\mathbf{r}_1)\phi_i(\mathbf{r}_1) = \int d^3\mathbf{r}_2 \phi_j^*(\mathbf{r}_2)\phi_i(\mathbf{r}_2)\frac{1}{r_{12}}\phi_j(\mathbf{r}_1)$$ 交換演算子

Roothaan法による行列計算：

$$\phi_i = \sum_p^{N_{\text{basis}}} \chi_p C_{pi} = \boldsymbol{\chi}\mathbf{C}_i$$

X_p：基底関数

$$\hat{F}\phi_i = \varepsilon_i\phi_i \quad \Rightarrow \quad \mathbf{F}\mathbf{C}_i = \varepsilon_i\mathbf{S}\mathbf{C}_i$$

$$|\mathbf{F} - \varepsilon_i\mathbf{S}| = 0$$

行列の固有値問題に帰着

密度汎関数(DFT)法(1)

Hohenberg-Kohnの定理

第1定理:
ある系の基底状態の電子密度が決まると、それを基底状態にもつ外部ポテンシャルがもし存在すれば、それはただ1通りに定まる。
その外部ポテンシャルから導かれるハミルトニアンのシュレディンガー方程式を解けば、その外部ポテンシャルのもとで許される電子系の波動関数がわかり、全ての物理量を求めることができる。

第2定理:
外部ポテンシャルをパラメータにもつ電子密度のエネルギー汎関数が存在して、この汎関数は与えられた外部ポテンシャルのもとでの基底状態の電子密度で最小値(基底状態のエネルギー)をもつ。
電子密度関数を変化させて最小のエネルギーを与える電子密度を探索すれば、基底状態の電子密度を求めることができる。

「電子密度」を中心に考えることで「自由度の縮約」を目指す

密度汎関数(DFT)法(2)

Kohn-Sham方程式

$$E = \sum_i \int d^3\mathbf{r}\,\varphi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\varphi_i(\mathbf{r}) + \int d^3\mathbf{r}\,\rho(\mathbf{r})V(\mathbf{r}) + J[\rho] + E_{xc}$$

クーロン　交換相関

(Hohenberg-Kohnのエネルギー汎関数)

拘束条件付き変分原理 → $\left[-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\varphi_i = \varepsilon_i\varphi_i$　HF方程式と類似

$$V_{\text{eff}}(\mathbf{r}) = V(\mathbf{r}) + \int d^3\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{xc}(\mathbf{r})$$

$$\rho = \sum_i^{\text{occ}} |\phi_i|^2$$　電子密度

$V_{xc}(r) = \delta E_{xc}/\delta\rho(r)$ ：交換相関ポテンシャル

電子相関効果の適切な考慮

- HF法をベースにした場合
  - Moeller-Plesset (MP)摂動法、Configuration Interaction (CI)法、CASSCF法、Coupled Cluster (CC) 法などによる系統的な改善
  - 相応の基底関数の改善
- DFT法をベースにした場合
  - 交換相関汎関数をいかに選ぶか？（経験的）
  - 局所密度近似（LDA）、一般化勾配近似（GGA）、メタGGA、混成GGA、半経験的汎関数など
  - 長距離補正、自己相互作用補正、分散力補正

生体高分子の量子化学計算

- 通常、$O(N^3)$以上の計算コスト
- 化学反応に重要な「活性領域」のみに対してコストのかかる量子化学計算を行い、それ以外は古典力学的に計算する → QM/MM法（2013年ノーベル化学賞）
- タンパク質全体をコストパフォーマンスの良いDFT法で計算する → ProteinDFなど
- 並列計算を活用した分割統治型のアプローチ → Divide-and-Conquer、FMO法など

フラグメント分子軌道 (FMO) 法とは?

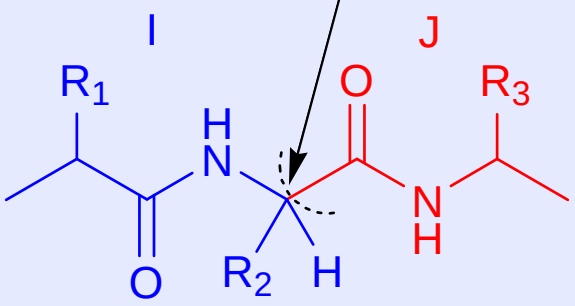
- 北浦和夫氏らが**1999**年に提唱
  - K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, “Pair interaction molecular orbital method: an approximate computational method for molecular interactions”, *Chem. Phys. Lett.* **312**, 319-324 (1999).
  - K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, “Fragment molecular orbital method: an approximate computational method for large molecules”, *Chem. Phys. Lett.* **313**, 701-706 (1999).
- 巨大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、トリマー等の分子軌道（**MO**）計算から分子全体の電子状態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 or Pair Interaction Energy; PIE)**解析が可能

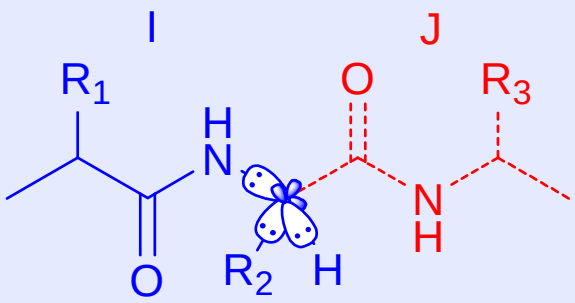
どうやって分子をフラグメントに分割するか？

- **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（デフォルト）
- **sp^2**炭素や**sp^3**ケイ素で分割することも可能（拡張が進められている）

Bond Detached Atom (BDA)

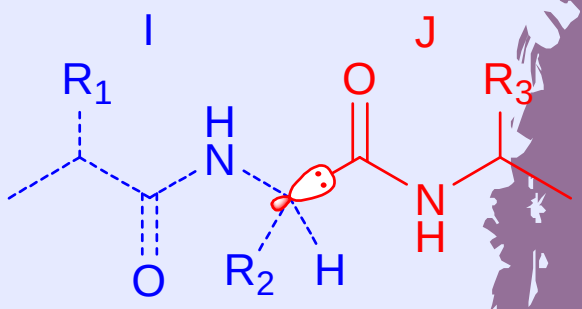
人為的な**H**キャップ等は不要





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

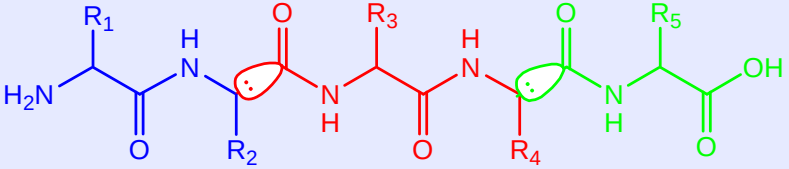


フラグメントJの変分空間

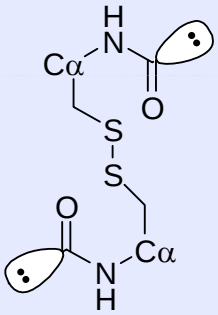
フラグメントJの形式電荷:　-1
BDAの炭素原子の核荷電:　1

タンパク質・DNAのフラグメント分割

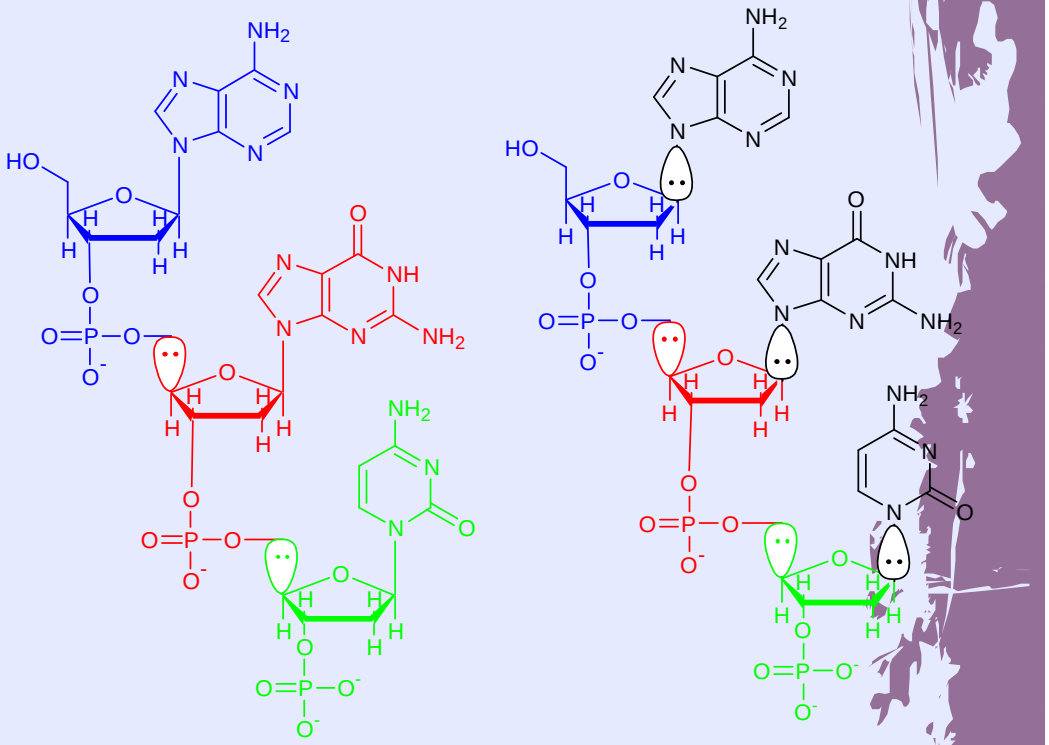
- タンパク質はCαの位置でフラグメントに分割



- 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



ジスルフィド結合



FMO計算の流れ

分子をフラグメントに分割し、電子を割り当てる

↓

モノマーの初期電子密度を計算する

↓

与えられた電子密度を用いて、周囲のモノマーによる環境静電ポテンシャル下での、モノマーのエネルギーと電子密度を計算する

↓

与えられた電子密度と得られた電子密度の差が閾値よりも小さい

No → 与えられた電子密度を用いて、周囲のモノマーによる環境静電ポテンシャル下での、モノマーのエネルギーと電子密度を計算する

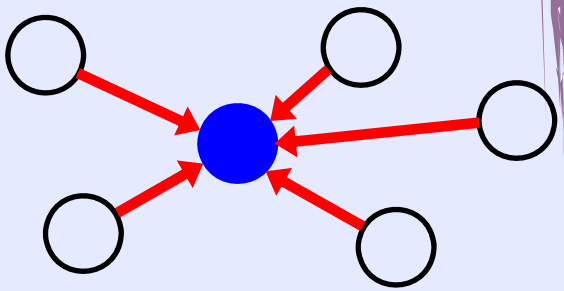
Yes

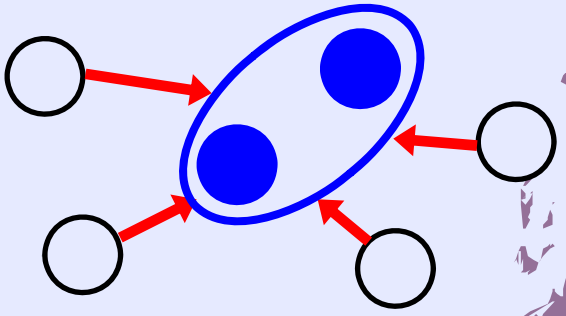
↓

収束したモノマーの電子密度(Self-Consistent Charge; SCC)を用いて、周囲のモノマーによる環境静電ポテンシャル下での、ダイマーのエネルギーと電子密度を計算する

↓

分子の全エネルギーおよび全電子密度を計算





FMO2法による分子の全エネルギー および全電子密度

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

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

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

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

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)　　2電子積分のMulliken近似

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

esp-ptc近似 (Mulliken atomic charge)　　点電荷近似

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

- ダイマー計算の高速化（**dimer-es近似**）

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

モノマー間の静電相互作用

➡ 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

L_{aoc}、L_{ptc}、$L_{dimer\text{-}es}$には適当なデフォルト値があてがわれているが、問題によっては確認が必要

フラグメント間相互作用エネルギー
Inter-Fragment Interaction Energy (IFIE)
Pair Interaction Energy (PIE)

- **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 $\Delta\tilde{E}_{IJ}$ 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。

$$E = \sum_{I>J} \Delta\tilde{E}_{IJ} + \sum_{I} E'_I$$

IFIE解析の基本式

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

IFIE

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

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

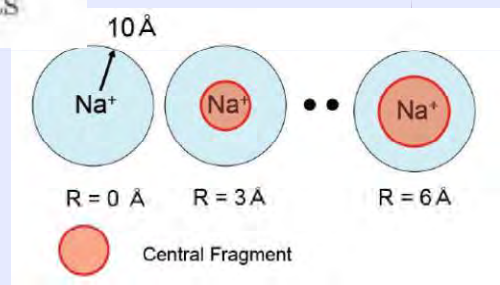
エネルギー分割 ⇒ **PIEDA**

FMO3: Application to hydrated sodium

Total energy errors of FMO calculations with the electrostatic potential and electrostatic interaction approximations (esp-aoc and $L_{\text{dimer-es}} = 2.0$) as functions of the central fragment radius in units of kcal/mol for 6-31G and 6-31G* basis sets

Fragment Radius (Å)	0	3	4	5	6
No. of Water molecules[a]	0	6	8	19	31
(FMO2)					
6-31G	−107.2	−96.9	−95.4	−76.8	−65.4
	−100.9[b]	−93.7[b]			
	−111.3[c]	−97.5[c]			
	−104.5[d]	−94.2[d]			
6-31G*	−106.8	−97.2	−95.3	−77.7	−67.1
	−92.6[b]	−83.8[b]			
	−111.3[c]	−97.9[c]			
	−96.6[d]	−84.4[d]			
(FMO3)					
6-31G	5.6	5.5	4.8	4.5	3.9
	4.2[b]	5.0[b]			
	−5.8[c]	−2.5[c]			
	−3.7[d]	0.2[d]			
6-31G*	1.7	1.8	2.0	2.3	1.9
	3.5[b]	3.6[b]			
	−5.5[c]	−1.5[c]			
	−4.6[d]	−1.2[d]			

[a] Number of water molecules included in the same fragment as Na^+.
[b] esp-non, $L_{\text{dimer-es}} = 2.0$. (See text.)
[c] esp-aoc, $L_{\text{dimer-es}} = 4.0$. (See text.)
[d] esp-non, $L_{\text{dimer-es}} = 4.0$. (See text.)



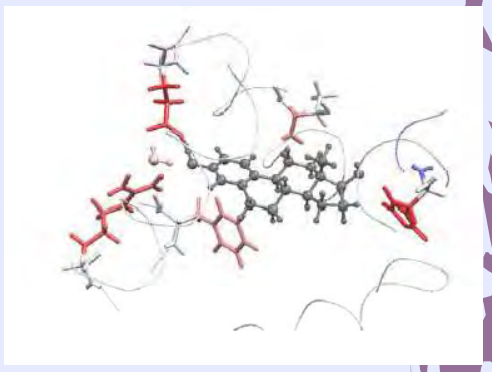
(T. Fujita *et al*., Chem. Phys. Lett. 478 (2009) 295)

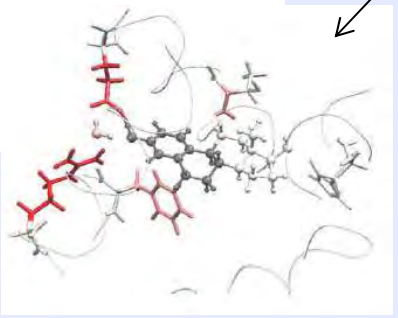
水クラスター:3体相互作用が重要

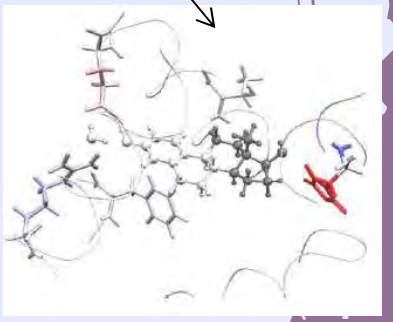
FMO4 Approximation

$$
\begin{aligned}
E^{\mathrm{FMO4}} &= \sum_{I>J>K>L} E_{IJKL} - (N-4)\sum_{I>J>K} E_{IJK} + \frac{(N-3)(N-4)}{2}\sum_{I>J} E_{IJ} \\
&\quad - \frac{(N-2)(N-3)(N-4)}{6}\sum_{I} E_I \\
&= \sum_{I>J>K>L} \{\Delta E_{IJKL} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{IL} - \Delta E_{JK} - \Delta E_{JL} - \Delta E_{KL} \\
&\quad - [\Delta E_{IJK} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{JK}] - [\Delta E_{IJL} - \Delta E_{IJ} - \Delta E_{IL} - \Delta E_{JL}] \\
&\quad - [\Delta E_{IKL} - \Delta E_{IK} - \Delta E_{IL} - \Delta E_{KL}] - [\Delta E_{JKL} - \Delta E_{JK} - \Delta E_{JL} - \Delta E_{KL}]\} \\
&\quad + \sum_{I>J>K} [\Delta E_{IJK} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{JK}] \\
&\quad + \sum_{I>J} \Delta E_{IJ} + \sum_{I} E_I
\end{aligned}
$$

$$
\Delta E_{IJKL} \equiv E_{IJKL} - E_I - E_J - E_K - E_L
$$

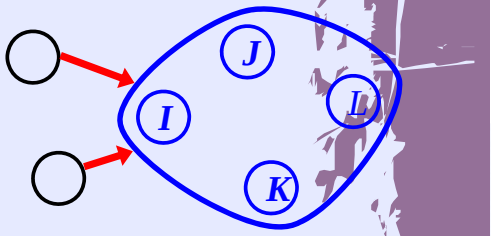






4体フラグメントまでを考慮に入れてFMO計算を行う。

→ エネルギー精度の向上
フラグメント分割の微細化

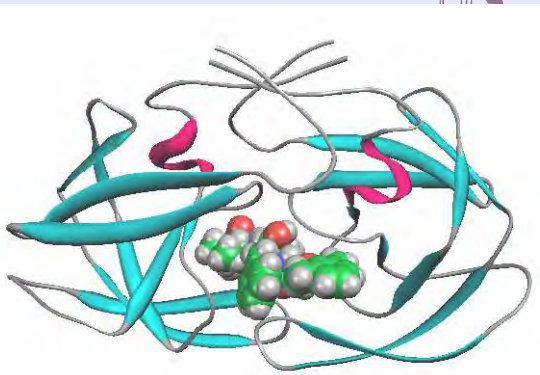


(T. Nakano et al., Chem. Phys. Lett. 523 (2012) 128.)

FMO4-MP2/ES2: HIV-1プロテアーゼ

	主鎖/側鎖分割		主鎖分割 （慣用）	
Total energy (au)	Main/side frg. HF	(total 363) MP2	Main frg. HF	(total 203) MP2
FMO2	-77554.2779	-77709.0268	-77589.3751	-77744.7417
Diff.(4)[a]	35.4754	36.1544	0.3059	0.3737
FMO3	-77590.0557	-77745.5971	-77589.6024	-77745.0359
Diff.(4)[a]	-0.3024	-0.4159	0.0786	0.0796
FMO4	-77589.7533	-77745.1812	-77589.6810	-77745.1155
Diff.(M)[b]	-0.0723	-0.0657		
Timing[c] (sec)				
Monomer	356.0	359.2	252.0	253.1
Dimer	120.8	141.2	131.7	175.3
Trimer	765.3	1040.1	648.2	1110.0
Tetramer	2352.3	3374.4	1358.2	2663.2
Total	3594.4	4914.9	2390.1	4201.6
Diff.(M)[b]	1204.3	713.3		

[a] Relative to the corresponding FMO4 energy.
[b] Breakdown timing analysis of the FMO4 jobs with 1024 processors of ES2.



- 198 残基
- 1水和水
- リガンド4分割
- 3225原子
- 17423軌道

- FMO2では主鎖/側鎖分割では大きなエラーが出てしまう （HF段階）
- FMO4の計算ではテトラマーの処理が支配的で超並列実行が望ましい

FMO法のソフトウェア

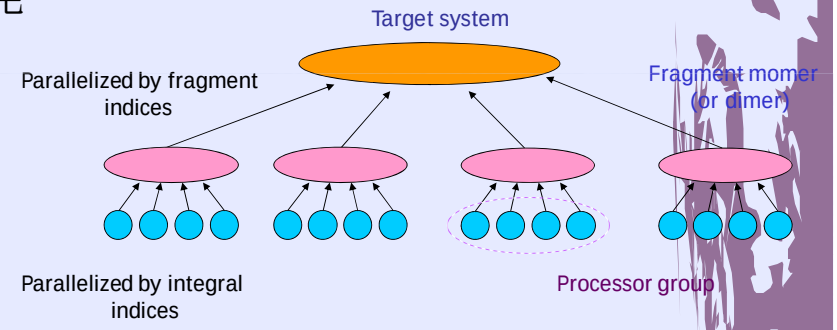
- 産業技術総合研究所（Fedorovら）⇒ **GAMESS/FMO**
- ナノ・グランドチャレンジ ⇒ **MP-FMO**
- 長崎大学（石川ら）⇒ **PAICS**
- 九州大学（稲富ら）⇒ **OpenFMO**
- **JST-CREST**プロジェクト ⇒ **ABINIT-MPX**
  http://insilico.h.kobe-u.ac.jp/tanaka/
- **CISS**プロジェクト（東大生産技術研究所）⇒ **ABINIT-MP**
  CISS公開版ダウンロード（無償）
  http://www.ciss.iis.u-tokyo.ac.jp/dl/index.php
  （**FMO-HF, FMO-MP2, FMO-MP3, FMO-MP2.5, CMM, MP2**エネルギー勾配, **BSSE**補正, **IFIE**解析, **IFIE Map, FILM**解析 **etc.**）

FMO法プログラム ABINIT-MP(X)

- RISSおよびJST-CREST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

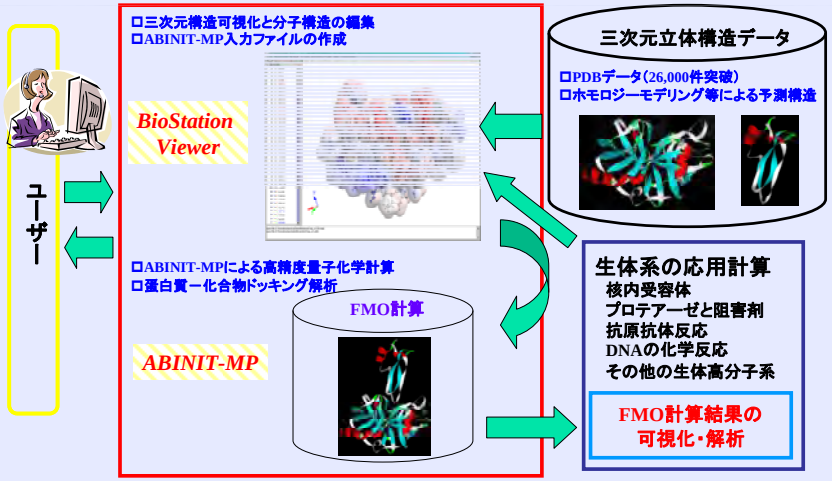
    Y. Mochizuki, M. Matsumura, T. Yokura, Y. Hirahara, T. Imamura, *J. Nucl. Sci. Tech.* **39**, 195-199 (2002).
- MP2, MP3計算のベクトル化
  - ⇒ 最深部はLevel 3 BLAS (DGEMM)



みずほ版MIZUHO/BioStation（有償）
独自機能(PIEDA等)および技術サポートつき

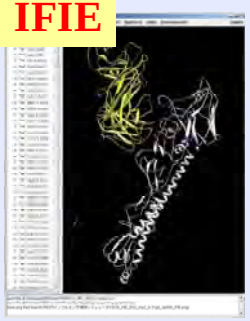
可視化・解析ソフトウェアBioStation Viewer

➢ ABINIT−MPによるFMO計算のためのGUI環境を整備

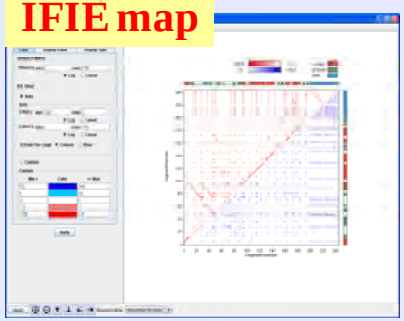


➢ フラグメント間相互作用エネルギー(IFIE)、軌道相互作用解析（FILM）等の結果を可視化表示

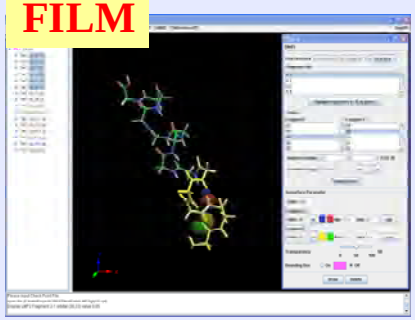
IFIE



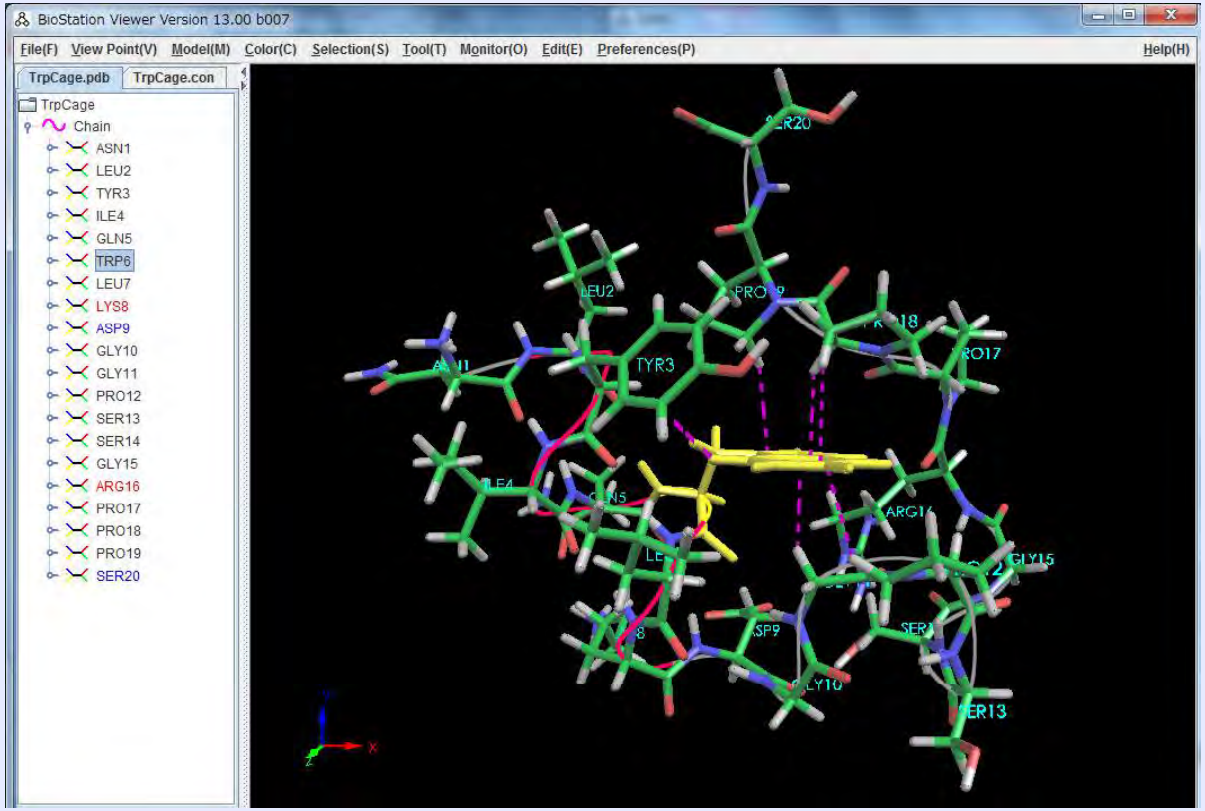
IFIE map



FILM



BioStation ViewerとCHPIプログラムの連携



Trp6-Tyr3
Gly11, Pro12, Pro18(2本), Pro19-Typ6

- Y. Umezawa, M. Nishio, Bioorganic Med. Chem. 6, 493-504 (1998).
- M. Nishio, M., Hirota, Y. Umezawa, “The CH/π Interaction Evidence, Nature, and Consequences” (Wiley-VCH, New York, 1998).
- 西尾元宏「新版　有機化学のための分子間力入門」(講談社、2008)

GAMESS/FMO

- アイオワ州立大学を中心に開発されている GAMESSにFMO機能を実装

  http://staff.aist.go.jp/d.g.fedorov/fmo/main.html
  http://fmo.alcf.anl.gov/ (FMO Portal)

- (階層的)並列化:GDDI (Generalized Distributed Data Interface)
- エネルギー解析:PIEDA (Pair Interaction Energy Decomposition Analysis)
- その他さまざまな機能を実装
- モデリングと可視化:Facio

PAICS

- Parallelized Ab Initio Calculation System based on FMO
- http://www.paics.net/index.html
- C言語、MPI並列、入力作成PaicsView
- FILM（局在化MP2）、エネルギー勾配、構造最適化、RI（Resolution of Identity）などの機能
- 2011年よりソースコードを無償で公開

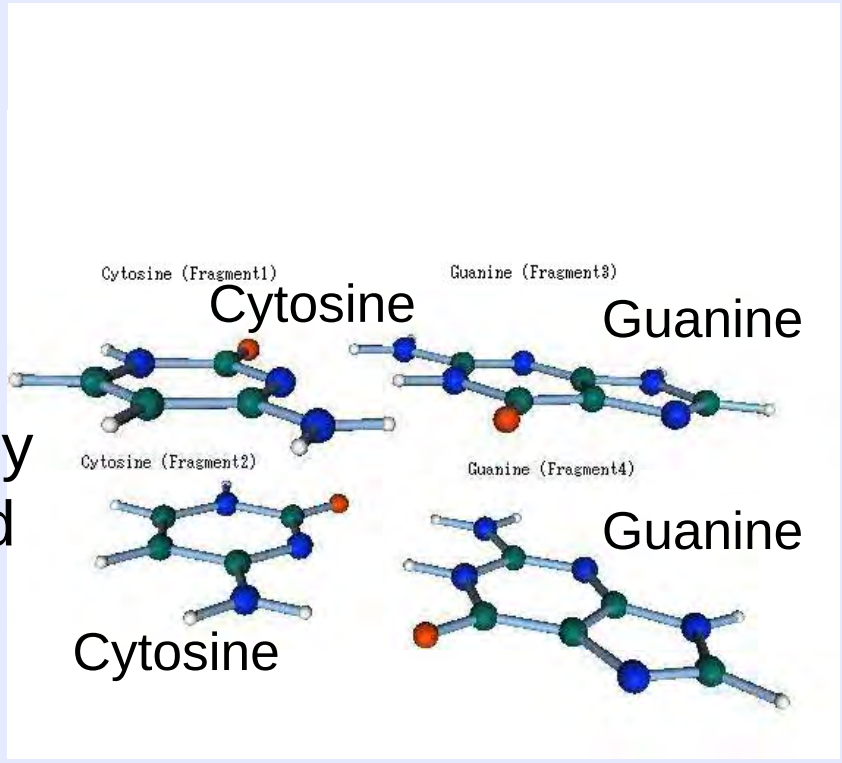
OpenFMO

- 九州大学、筑波大学、理研
- http://www.openfmo.org
- 「京」での実行を想定した並列FMOプログラム
- OpenMP+MPIによるハイブリッド並列化
- GPGPUの利用も進められている
- Hartree-Fock(HF)法に特化
- 負荷分散が超並列化の鍵 ⇒ 動的負荷分散

なぜ、電子相関効果が重要か？

DNA Base Pair Stacking

塩基スタッキング：分散力の重要性



HF or DFT cannot describe the weak attraction due to van der Waals (dispersion) interaction appropriately.

As for FMO-MP2 description for vdW interaction, see, e.g.,
K. Fukuzawa *et al.*, J. Comput. Chem. 27 (2006) 948.

Electron Correlations

Electron correlations play important roles for the descriptions of weak molecular interactions associated with hydrogen bonding and van der Waals (dispersion) forces.

- Watson-Crick pair
- Ligand binding
- DNA base stacking

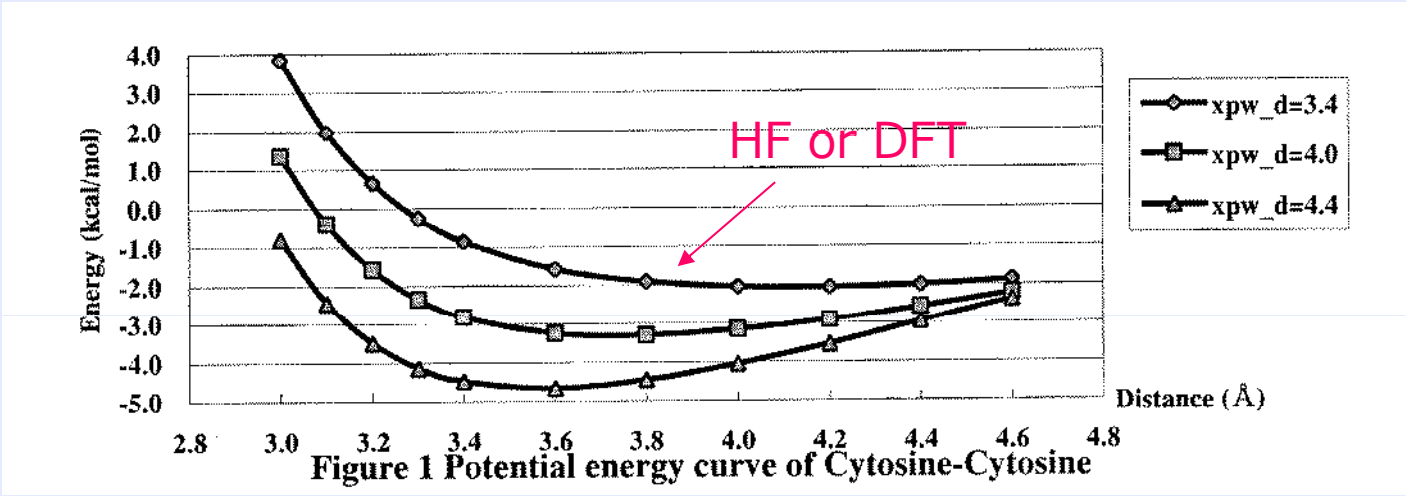


Figure 1 Potential energy curve of Cytosine-Cytosine

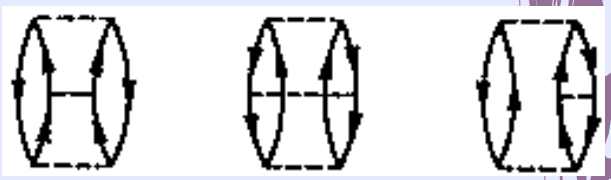
→ Implementation of MP2 and MP3 into FMO

FMO Calculations with MP2 and MP3

$$E^{MP2(2p-2h)} = \frac{1}{4}\sum_{ijab}\frac{\langle ij \| ab\rangle\langle ab \| ij\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

$$E_{MP2} = E^{MP2(2p-2h)}$$

$$E^{MP3(4p-2h)} = \frac{1}{8}\sum_{ijabcd}\frac{\langle ij \| ab\rangle\langle ab \| cd\rangle\langle cd \| ij\rangle}{(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)(\varepsilon_i + \varepsilon_j - \varepsilon_c - \varepsilon_d)}$$



$$E^{MP3(2p-4h)} = \frac{1}{8}\sum_{ijklab}\frac{\langle ij \| ab\rangle\langle ab \| kl\rangle\langle kl \| ij\rangle}{(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)(\varepsilon_k + \varepsilon_l - \varepsilon_a - \varepsilon_b)}$$

Cost for MP3 is higher by
- ⇒ 8-10 times @ PC cluster
- ⇒ less than 2 times @ ES2 (efficient vectorization)

$$E^{MP3(3p-3h)} = \sum_{ijkabc}\frac{\langle ij \| ab\rangle\langle kb \| cj\rangle\langle ac \| ik\rangle}{(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)(\varepsilon_i + \varepsilon_k - \varepsilon_a - \varepsilon_c)}$$

- integral direct
- DGEMM

$$E_{MP3} = E^{MP2(2p-2h)} + E^{MP3(2p-4h)} + E^{MP3(4p-2h)} + E^{MP3(3p-2h)}$$

(Y. Mochizuki *et al.,* Chem. Phys. Lett. 493 (2010) 346.)

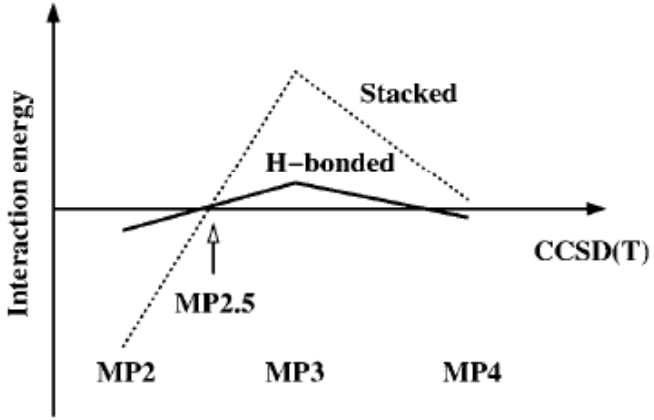
MP2.5 Approximation

Scaled MP3 Non-Covalent Interaction Energies Agree Closely with Accurate CCSD(T) Benchmark Data

Michal Pitoňák,[a] Pavel Neogrády,[b] Jiří Černý,[a] Stefan Grimme,*[c] and Pavel Hobza*[a]

Chem. Phys. Chem. **10** (2009) 282

Scaled MP3 interaction energies calculated as a sum of MP2/CBS (complete basis set limit) interaction energies and scaled third-order energy contributions obtained in small or medium size basis sets agree very closely with the estimated CCSD(T)/CBS interaction energies for the 22 H-bonded, dispersion-controlled and mixed non-covalent complexes from the S22 data set. Performance of this so-called MP2.5 (third-order scaling factor of 0.5) method has also been tested for 33 nucleic acid base pairs and two stacked conformers of porphine dimer. In all the test cases, performance of the MP2.5 method was shown to be superior to the scaled spin-component MP2 based methods, e.g. SCS–MP2, SCSN–MP2 and SCS(MI)–MP2. In particular, a very balanced treatment of hydrogen-bonded compared to stacked complexes is achieved with MP2.5. The main advantage of the approach is that it employs only a single empirical parameter and is thus biased by two rigorously defined, asymptotically correct ab-initio methods, MP2 and MP3. The method is proposed as an accurate but computationally feasible alternative to CCSD(T) for the computation of the properties of various kinds of non-covalently bound systems.



$$E(MP2.5) = E(MP2) + 0.5\, E^{corr}(MP3)$$

$$E(MP2.5/large\ basis) \doteq E(MP2/large\ basis) + \Delta E(MP2.5\text{-}MP2/small\ basis)$$

Comparable accuracy to CCSD(T)

MP2計算の高速化：RI & CDAM

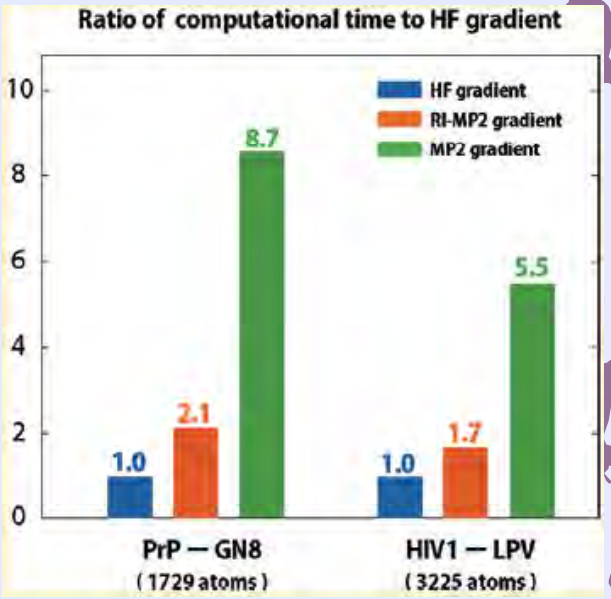
$$E^{\mathrm{MP2}} = \sum_{ij}\sum_{ab} \frac{2(ia|jb)^{\mathrm{RI}} - (ib|ja)^{\mathrm{RI}}}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}(ia|jb)^{\mathrm{RI}}$$

$$(ia|jb)^{\mathrm{RI}} = \sum_{P} B_{ia}^{P} B_{jb}^{P} \qquad B_{ia}^{P} = \sum_{Q} (ia|Q)V^{-1/2}{}_{QP}$$

$$V_{PQ} = (P|Q)$$

P, Q: Auxiliary basis functions

(T. Ishikawa *et al*., J. Phys. Chem. Lett. 3 (2012) 375.)



Acceleration of fragment molecular orbital calculations with Cholesky decomposition approach

Yoshio Okiyama [a,*], Tatsuya Nakano [a,b,c], Katsumi Yamashita [d], Yuji Mochizuki [a,c,e], Naoki Taguchi [e], Shigenori Tanaka [c,f]

Chemical Physics Letters 490 (2010) 84–89

$$(pq,rs) \approx \sum_{K=1}^{M} L_{K,pq} L_{K,rs},$$

$$(pq,rs) \approx \sum_{IJ} (pq|I) G_{IJ}^{-1} (J|rs),$$

$$G_{IJ}^{-1} \approx \sum_{K}^{M} Z_{IK} Z_{JK}$$

$$L_{K,pq} = \sum_{I}^{M} Z_{IK} (I|pq),$$

BSSE(基底関数重ね合わせ誤差)の補正

Counterpoise-corrected interaction energy analysis based on the fragment molecular orbital scheme

Yoshio Okiyama [a,*], Kaori Fukuzawa [a,b], Haruka Yamada [c], Yuji Mochizuki [a,d], Tatsuya Nakano [a,e], Shigenori Tanaka [f]

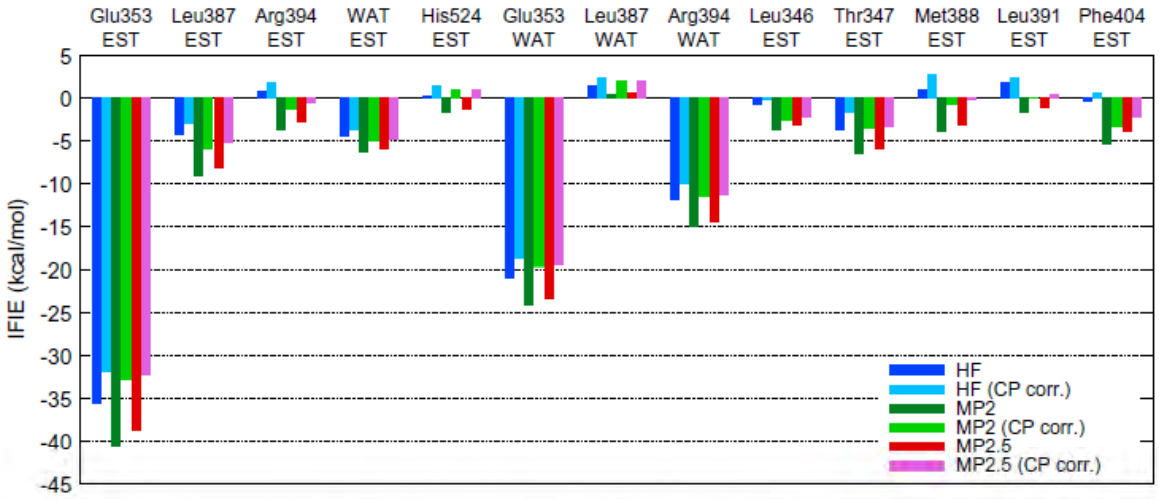
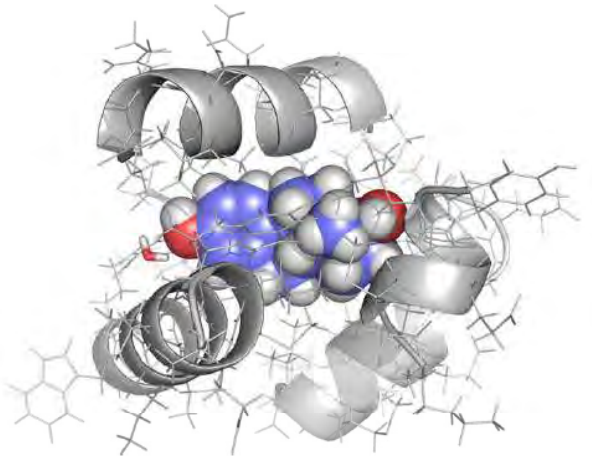


Figure 6. IFIEs and the CP-corrected ones of hydrogen-bonding networks in ER and EST complex.



Estrogen receptor

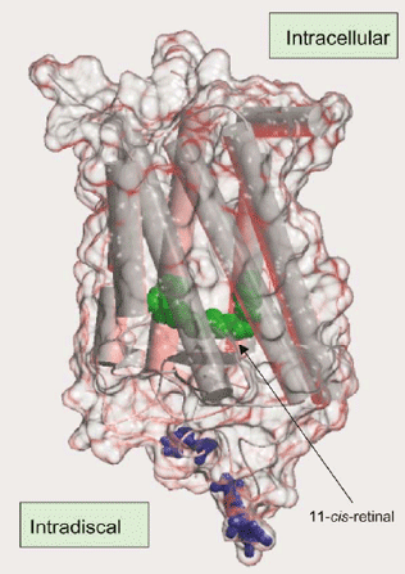
Chemical Physics Letters 509 (2011) 67–71

励起状態の計算（光吸収や発光）

$$\Psi^{CIS} = \sum_{ia} T_i^a \Psi_i^a$$

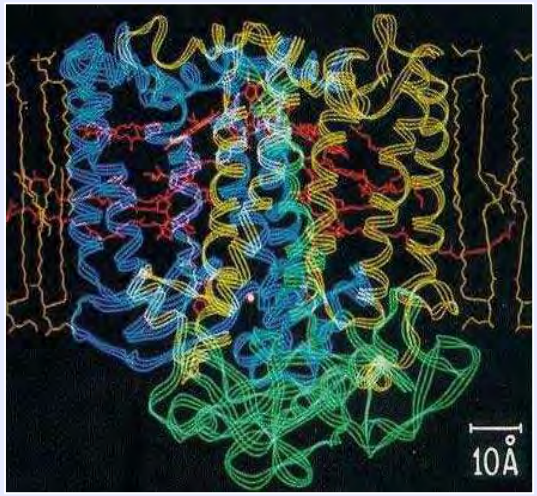
$$\Delta E^{CIS} = E^{HF} - E^{CIS}$$

・1電子励起のみを考慮して励起状態を表現
・HOMO-LUMO遷移などでは定性的にOK
・定量性には電子相関の考慮が必要



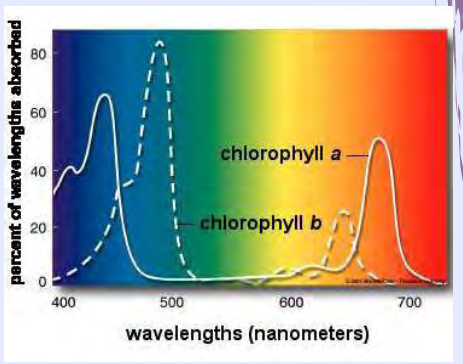
http://mcdb.colorado.edu/courses/3280/lectures/class14-2.html

（視物質レチナール蛋白質）



Allen *et al*., Proc. Natl. Acad. Sci. USA 84 (1987) 5730.

（光合成系）



http://ntri.tamuk.edu/bio/photo/

生命現象とも関係が深い

分子の励起状態の理論的記述（CIS近似）

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{n!}} \left| \chi_1 \chi_2 \cdots \chi_i \chi_j \cdots \chi_n \right|$$

：Hartree-Fock単一行列式波動関数
→ 基底状態を近似

$$\Psi_i^a = \frac{1}{\sqrt{n!}} \left| \chi_1 \chi_2 \cdots \chi_a \chi_j \cdots \chi_n \right|$$

$$E_{ia} = E_{\mathrm{HF}} + \varepsilon_a - \varepsilon_i - (ia||ia)$$

：1電子励起行列式
とそのエネルギー
（$i \rightarrow a$：MOの励起）

$$\chi_i = \sum_{\mu}^{N} c_{\mu i} \phi_\mu$$

$$(pq||rs) = \sum_{\mu\nu\lambda\sigma} c_{\mu p} c_{\nu q} c_{\lambda r} c_{\sigma s} (\mu\nu||\lambda\sigma)$$

$$(\mu\nu||\lambda\sigma) = \iint \phi_\mu(\boldsymbol{r_1}) \phi_\nu(\boldsymbol{r_2}) \frac{1}{r_{12}} \left[\phi_\lambda(\boldsymbol{r_1}) \phi_\sigma(\boldsymbol{r_2}) - \phi_\lambda(\boldsymbol{r_2}) \phi_\sigma(\boldsymbol{r_1}) \right] d\boldsymbol{r_1} d\boldsymbol{r_2}$$

$$\Psi_{\mathrm{CIS}} = \sum_{ia} a_{ia} \Psi_i^a$$

→ 励起状態のエネルギー E_{CIS}
（Configuration Interaction）

$$A_{ia,jb} = \left\langle \Psi_i^a \left| H \right| \Psi_j^b \right\rangle = \left[E_{\mathrm{HF}} + \varepsilon_a - \varepsilon_i \right] \delta_{ij} \delta_{ab} - (ja||ib)$$

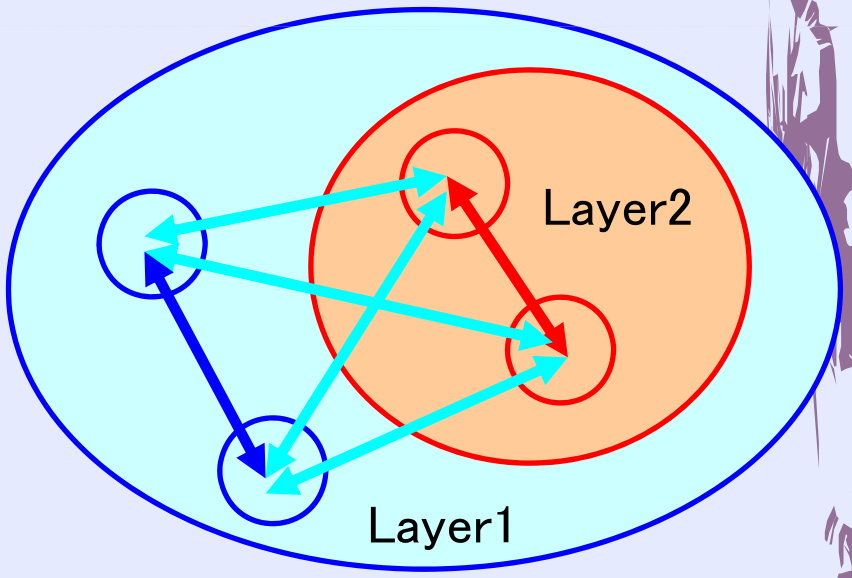
を対角化

Multi-Layer FMO (MLFMO) Method

$$E = \sum_{L_i}\left[\left(\sum_{I\in L_i} E'^{L_i}_I + \sum_{\substack{I>J \\ I,J\in L_i}} \Delta\tilde{E}^{L_i}_{IJ}\right) + \sum_{L_j>L_i}\sum_{I\in L_i}\sum_{J\in L_j}\Delta\tilde{E}^{L_i}_{IJ}\right]$$

Layer1

Layer2



E_I : Monomer energy, E_{IJ} : Dimer energy

$$E_I = E'_I + Tr\left(\mathbf{D}^I\mathbf{V}^I\right), \quad E_{IJ} = E'_{IJ} + Tr\left(\mathbf{D}^{IJ}\mathbf{V}^{IJ}\right)$$

$$\Delta E'_{IJ} = E'_{IJ} - E'_I - E'_J, \quad \Delta\mathbf{D}^{IJ} = \mathbf{D}^{IJ} - \mathbf{D}^I \oplus \mathbf{D}^J$$

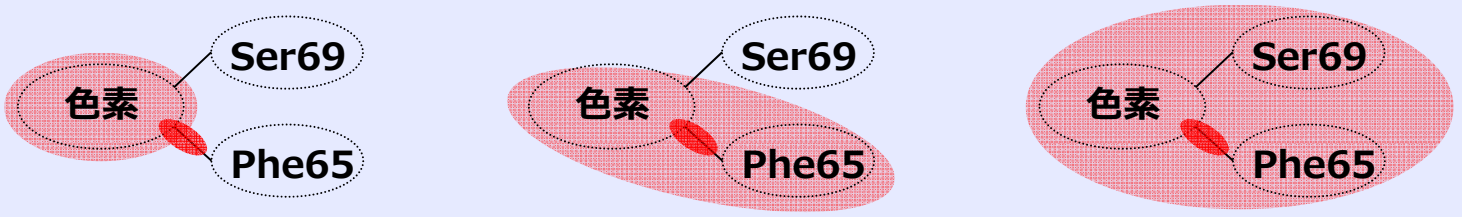
$$\Delta\tilde{E}_{IJ} = \Delta E'_{IJ} + Tr\left(\Delta\mathbf{D}^{IJ}\mathbf{V}^{IJ}\right)$$

- Layer 2: Important region to which accurate (correlated) methods are applied.
- Layer 1 is analyzed by FMO-HF or MM (XUFF) method, while layer 2 is analyzed under the environmental electrostatic potentials by layer 1.

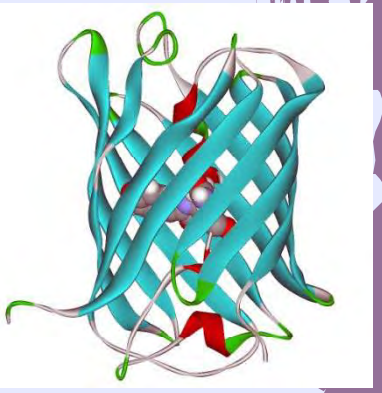
  ⇒ MLFMO-MP2, MLFMO-CIS, MLFMO-CIS(D) methods available.

D.G. Fedorov and K. Kitaura, *J. Chem. Phys.* **122**, 054108 (2005).

赤色蛍光タンパク質(RFP:DsRed)の計算値



eV	CIS	CIS(D)	CIS(Ds)	PR-CIS(D)	PR-CIS(Ds)
Excitation energy					
pigment (X)	3.35	2.49	2.54	2.36	2.41
F65+X	3.27	2.30	2.37	2.18	2.24
F65+X+S69	3.26	2.28	2.34	2.16	2.22
Experiment					2.22
Emission energy					
pigment (X)	3.25	2.41	2.45	2.28	2.32
F65+X	3.20	2.21	2.27	2.09	2.15
F65+X+S69	3.18	2.21	2.26	2.09	2.14
Experiment					2.13



(Y. Mochizuki et al., Chem. Phys. Lett. 433 (2007) 360; N. Taguchi et al., J. Phys. Chem. B)

“最良の扱い”では吸収・発光エネルギーをほぼ完全に再現!!

Bioluminescence Spectra of Firefly

Fig. 2. Structure of luciferase (PDBID: 2D1S). Model 1 with the amino acid residues within the distance of 4.5 Å from oxyluciferin and AMP is included in red circle. Layer 2 is included in pink circle, while the other region corresponds to Layer 1.

Table 2
Calculated values of emission energy and corresponding wavelength of WT and three mutant forms (model 1) obtained by MUFMO-CIS(D)/6-31G calculations in comparison with experimental values.

		WT	I288V	S286N	I288A
Emission energy (eV)	Calc.	2.49	2.49	2.39	2.37
	Expt.	2.21	2.02, 2.21	2.05	2.02
Wavelength (nm)	Calc.	498	498	518	523
	Expt.	560	560, 613	605	613
Color	Expt.	Yellow–green	Orange	Red	Red

Table 5
Results of emission energy (eV) for WT and I288A with model 2 obtained by MUFMO-CIS(D)/6-31G and MUFMO-PR-CIS(Ds)/6-31G calculations in comparison with the experimental values. (A) Assigning only oxyluciferin to the Layer 2; (C) Assigning oxyluciferin, ILE288 (or ALA288), PHE249 and ALA350 to the Layer 2.

Layer 2	A		C		
Method	CIS(D)	PR-CIS(Ds)	CIS(D)	PR-CIS(Ds)	Expt.
WT	2.30	2.27	2.26	2.23	2.21
I288A	2.25	2.22	2.15	2.12	2.02

(A. Tagami et al., Chem. Phys. Lett. 472 (2009) 118.)

励起状態への溶媒効果

Polarizable Continuum Model with the Fragment Molecular Orbital-Based Time-Dependent Density Functional Theory

MAHITO CHIBA,[1]* DMITRI G. FEDOROV,[1] KAZUO KITAURA[1,2]
[1]*Research Institute for Computational Sciences, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*
[2]*Graduate School of Pharmaceutical Sciences, Kyoto University, Sakyo-Ku, Kyoto 606-8501, Japan*

Received 5 January 2008; Revised 11 March 2008; Accepted 17 March 2008
DOI 10.1002/jcc.21000
Published online 16 May 2008 in Wiley InterScience (www.interscience.wiley.com).

Abstract: The polarizable continuum model (PCM) for describing the solvent effect was combined with the fragment molecular orbital-based time-dependent density functional theory (TDDFT). Several levels of the many-body expansion were implemented, and the importance of the many-body contributions to the singlet-excited states was discussed. To calibrate the accuracy, we performed a number of the model calculations using our method and the regular TDDFT in solution, applying them to phenol and polypeptides at the long-range corrected BLYP/6-31G* level. It was found that for systems up to 192 atoms the largest error in the excitation energy was 0.006 eV (vs. the regular TDDFT/PCM of the full system). The solvent shifts and the conformer effects were discussed, and the scaling was found to be nearly linear. Finally, we applied our method to the lowest singlet excitation of the photoactive yellow protein (PYP) in aqueous solution and determined the excitation energy to be in reasonable agreement with experiment. The excitation energy analysis provided the contributions of individual residues, and the main factors as well as their solvent shifts were determined.

 J Comput Chem 29: 2667–2676, 2008

TDDFT + PCM

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$$

(GAMESS/FMO)

Large-Scale MP2 Calculations on the Blue Gene Architecture Using the Fragment Molecular Orbital Method

Graham D. Fletcher,[†] Dmitri G. Fedorov,[‡] Spencer R. Pruitt,[§] Theresa L. Windus,[§] and Mark S. Gordon[*,§]

[†]Argonne Leadership Computing Facility, Argonne, Illinois 60439, United States
[‡]National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
[§]Iowa State University and Ames Laboratory, Ames, Iowa 50011, United States

J. Chem. Theory Comput. 2012, 8, 75–79

ABSTRACT: Benchmark timings are presented for the fragment molecular orbital method on a Blue Gene/P computer. Algorithmic modifications that lead to enhanced performance on the Blue Gene/P architecture include strategies for the storage of fragment density matrices by process subgroups in the global address space. The computation of the atomic forces for a system with more than 3000 atoms and 44 000 basis functions, using second order perturbation theory and an augmented and polarized double-ζ basis set, takes ~7 min on 131 072 cores.

Table 1. The Performance of FMO2-MP2 Force Calculations on the Blue Gene/P[a]

		racks:	1	2	4	8	16	32
		cores:	4096	8192	16 384	32 768	65 536	131 072
waters	atoms	basis functions	wall time (min)					
128	384	2432	0.5	0.4				
256	768	4864	1.1	0.7	0.5			
512	1536	9728	3.6	2.0	1.2	0.8	0.7	
1024	3072	19 456	10.7	6.3	3.4	2.1	1.5	1.2
2048	6144	38 912	34.6	18.5	11.1	6.1	3.7	2.6

[a] The atomic basis set is 6-31G(d).

Table 2. The Performance of FMO2-MP2 Force Calculations on the Blue Gene/P[a]

		racks:	1	2	4	8	16	32
		cores:	4096	8192	16 384	32 768	65 536	131 072
waters	atoms	basis functions	wall time (min)					
128	384	5504	8.6	4.8	2.7	1.8		
256	768	11 008	19.8	10.5	5.8	3.4	2.2	
512	1536	22 016		28.9	15.4	8.6	4.9	3.2
1024	3072	44 032			41.1	22.0	12.2	7.1

[a] The atomic basis set is aug-cc-pVDZ.

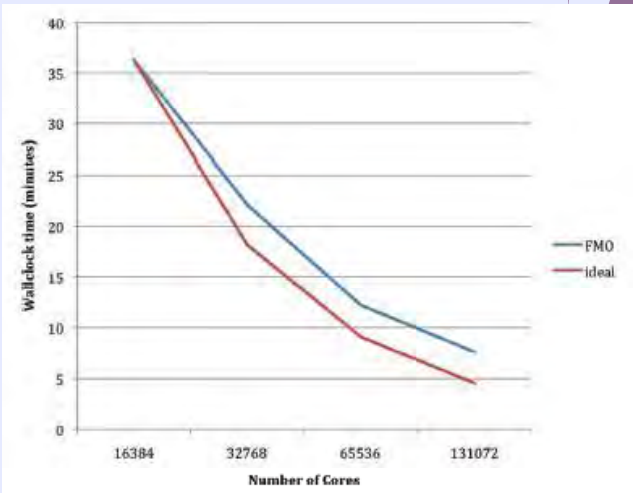


Figure 1. FMO2-MP2/6-31G(d) forces calculation of 12 288 atoms on Blue Gene/P.

Ab Initio FMO-MD

◇*Ab initio* molecular dynamics

Forces by FMO @ ABINIT-MP ⇔ Trajectory @ PEACH

⇒ FMO2-HF （MP2 and FMO3 in progress）

Comparison with Car-Parrinello MD？

◇Application to hydrated formaldehyde

Blue shift in excitation energy

⇒ Benchmark system

Droplet model with 128 waters

FMO-HF/6-31G

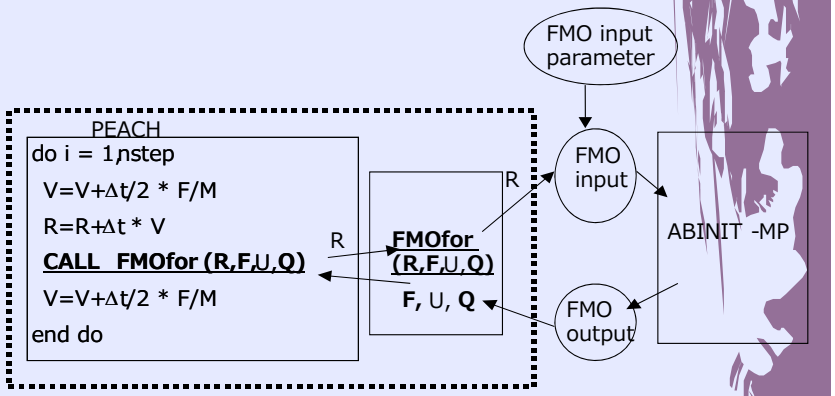
⇒ 1 step/1.5 min. @ 20cores

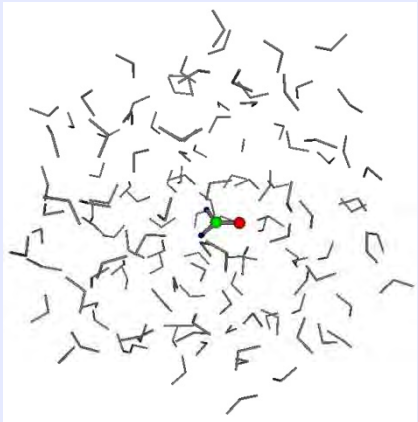
⇒ several thousands step MD

Hundreds of sampling

⇒ MLFMO-CIS(D)/6-31G*

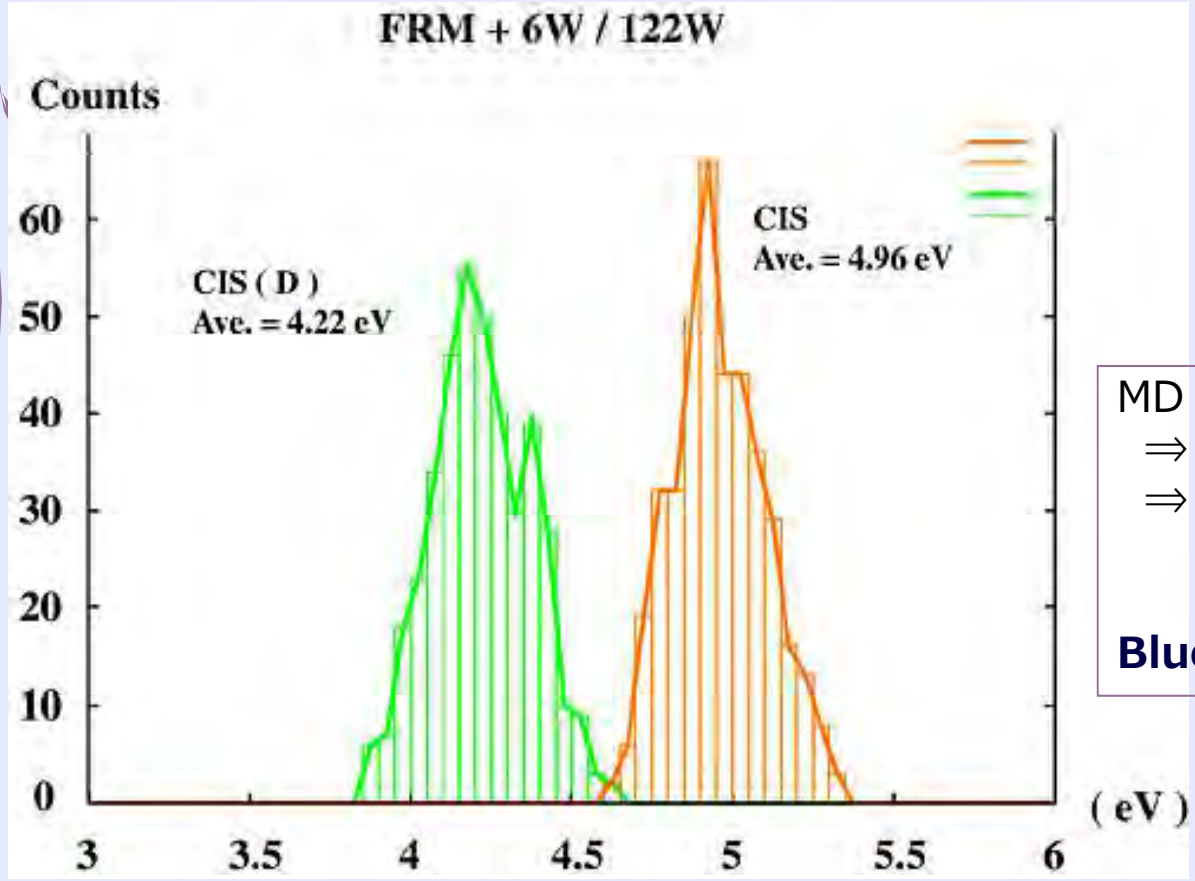
Good agreement with experiment





Ab Initio FMO-MD

400 samples



MD for gas-phase FRM
⇒ 4.56eV @ CIS
⇒ 4.08eV @ CIS(D)
(Expt.: 4.07eV)

Blue shift: +0.14eV @ CIS(D)

(Y. Mochizuki *et al*., Chem. Phys. Lett. 437 (2007) 66.)

Periodic Boundary Condition for FMO

$$E^{cell} = \sum_{I_0} E'_{I_0} + \frac{1}{2}\sum_{\mathbf{n}}\sum_{I_0 J_\mathbf{n}}{}' \Delta\widetilde{E}_{I_0 J_\mathbf{n}} + \sum_{\mathbf{n},\mathbf{n}'}\sum_{I_0 J_\mathbf{n} K_{\mathbf{n}'}}{}'' \Delta\widetilde{E}_{I_0 J_\mathbf{n} K_{\mathbf{n}'}}$$

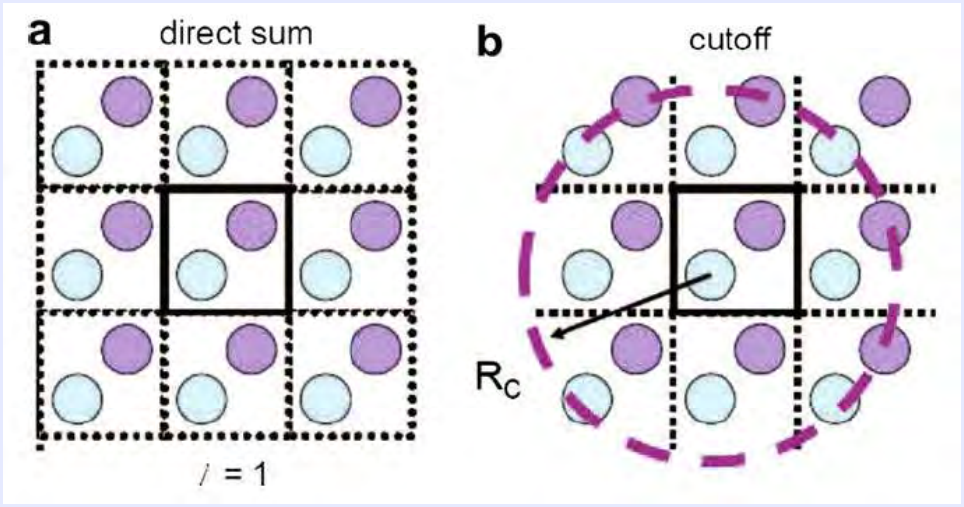
$$\Delta\widetilde{E}_{I_0 J_\mathbf{n}} = E'_{I_0 J_\mathbf{n}} - E'_{I_0} - E'_{J_\mathbf{n}} + \mathrm{Tr}\left(\Delta\mathbf{D}^{I_0 J_\mathbf{n}}\mathbf{V}^{I_0 J_\mathbf{n}}\right)$$

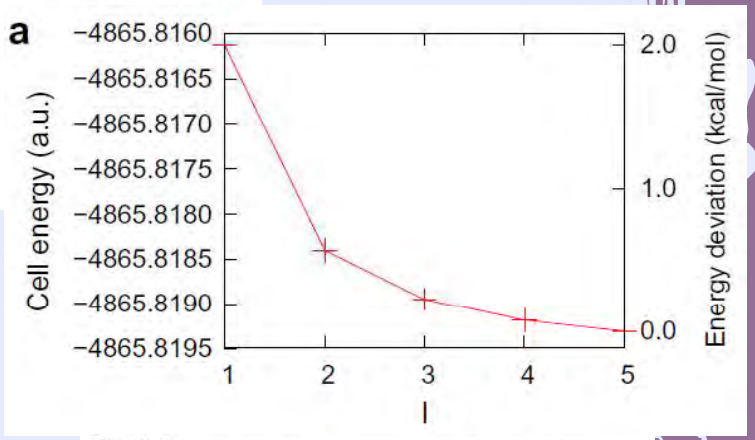
$$\Delta\widetilde{E}_{I_0 J_\mathbf{n} K_{\mathbf{n}'}} = E'_{I_0 J_\mathbf{n} K_{\mathbf{n}'}} - E'_{I_0} - E'_{J_\mathbf{n}} - E'_{K_{\mathbf{n}'}} + \mathrm{Tr}\left(\Delta\mathbf{D}^{I_0 J_\mathbf{n} K_{\mathbf{n}'}}\mathbf{V}^{I_0 J_\mathbf{n} K_{\mathbf{n}'}}\right) - \Delta\widetilde{E}_{I_0 J_\mathbf{n}} - \Delta\widetilde{E}_{I_0 K_{\mathbf{n}'}} - \Delta\widetilde{E}_{J_\mathbf{n} K_{\mathbf{n}'}}$$

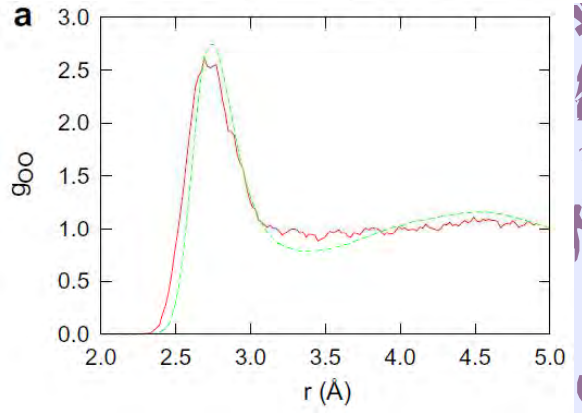
$$\mathbf{n} = (n_x, n_y, n_z)$$

$$n_x, n_y, n_z = 0, \pm 1, \pm 2, \ldots$$

水溶液中のFMO-MD
計算はピコ秒レベル





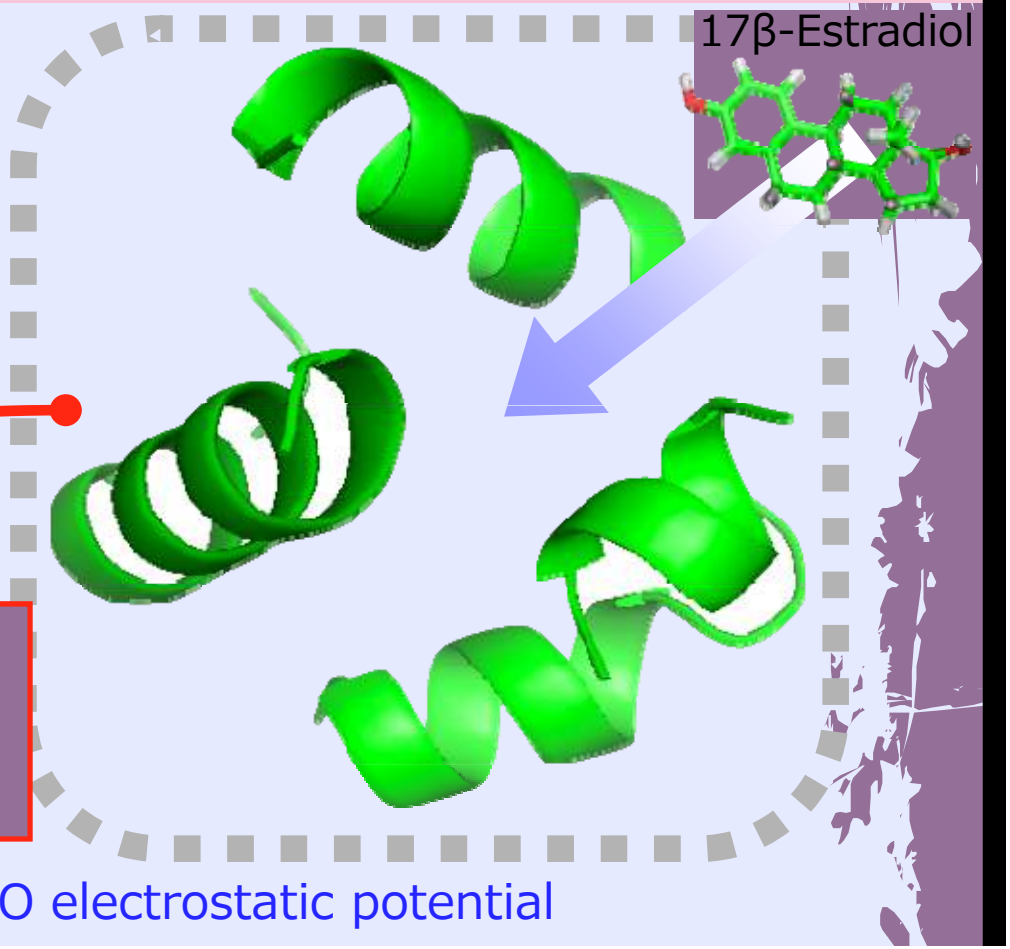


(T. Fujita *et al.,* Chem. Phys. Lett. 506 (2011) 112.)

Relevance of FMO-ESP charges

Reproduction of electrostatic field

FMO-ESP charges
in estrogen
receptor

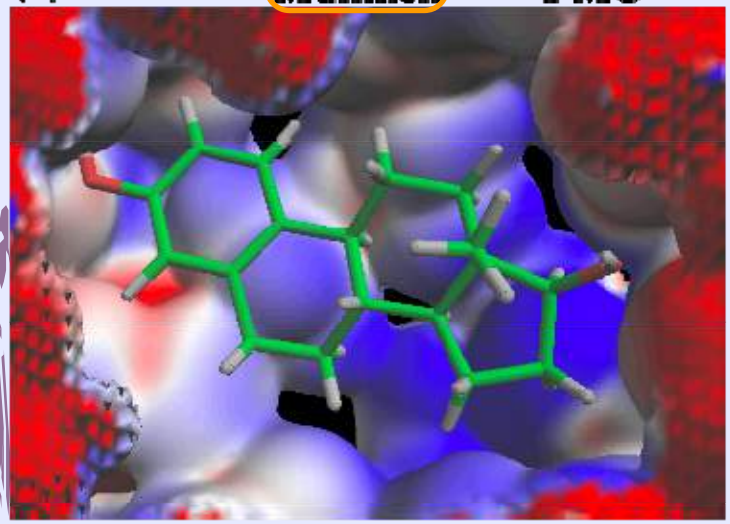


Estrogen receptor
50 amino acid residues
model

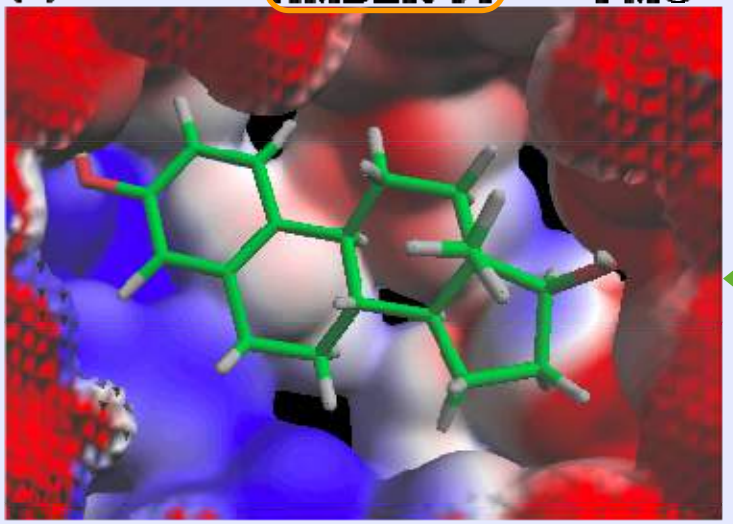
FMO: Excellent reproduction of MO electrostatic potential

FMO-ESP charge: Difference in ligand binding pocket of ER-EST

(a) $\Delta V = V_{\text{Mulliken}} - V_{\text{FMO}}$



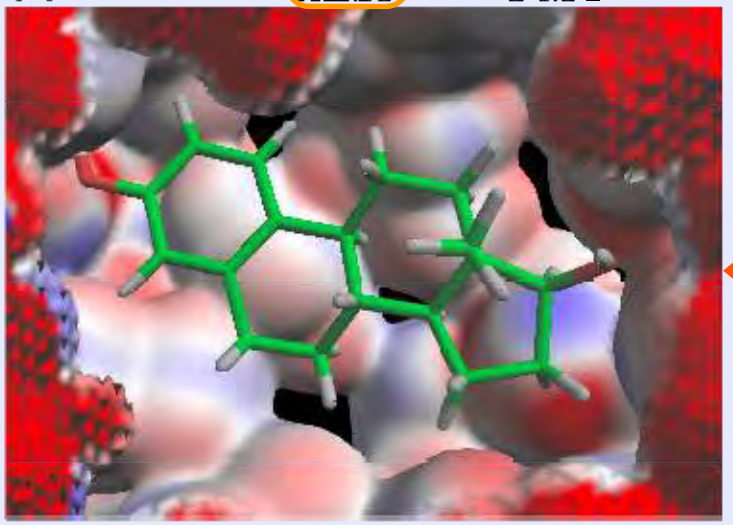
(b) $\Delta V = V_{\text{AMBER 94}} - V_{\text{FMO}}$



Deep:
Large diff.

(c) $\Delta V = V_{\text{Merz-Kollman}} - V_{\text{FMO}}$

(d) $\Delta V = V_{\text{RESP}} - V_{\text{FMO}}$

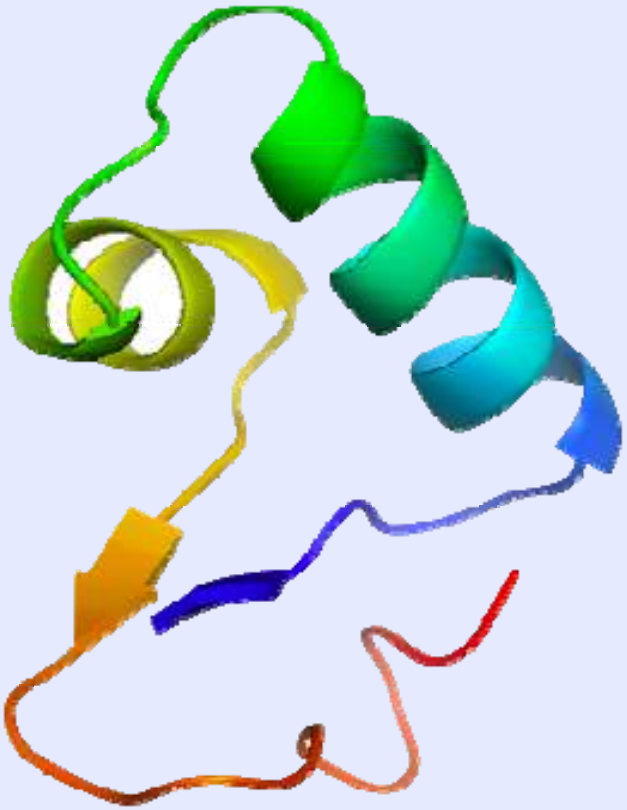


Light:
Small diff.

(a.u.)

(Y. Okiyama *et al*., Chem. Phys. Lett. 449 (2007) 329.)

MK vs. AMBER Charges in Crambin



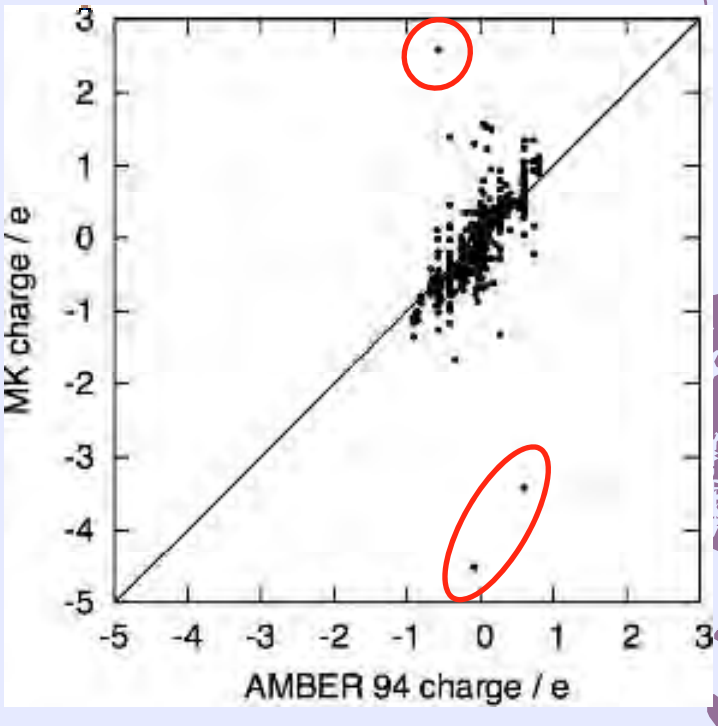
crambin (PDB ID: 1CCM)

46 residues (642 atoms)

Accuracy of ES potential

Charge	Deviation from QM
Merz-Kollman	1.9%
AMBER 94	27.5%

MK vs. AMBER charges



Restrained Fitting

W. Cornell et al., J. Am. Chem. Soc. 115 (1993) 9620.

Add a cost function due to the deviation from reference charges

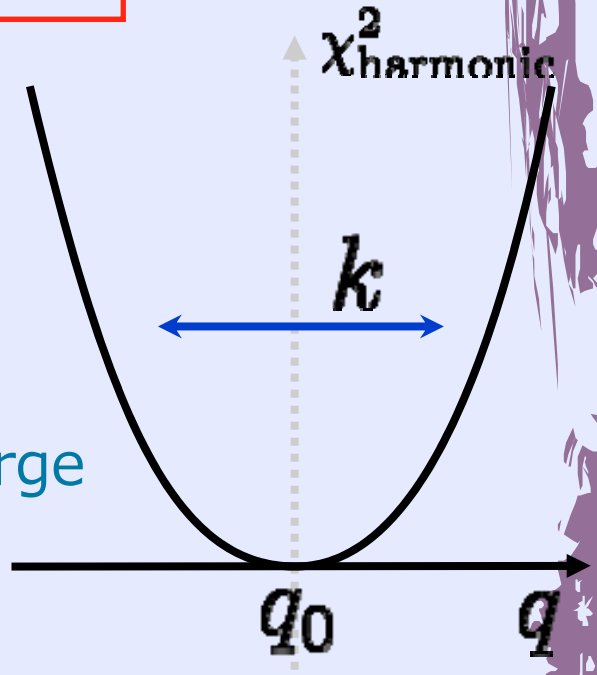
$$\chi^2_{ESP} = \chi^2_{ESP} + \chi^2_{harmonic}$$

Minimization

$$\chi^2_{harmonic} = k \sum_j (q_j - q_{0j})^2$$

Force constant (weight)

Reference charge

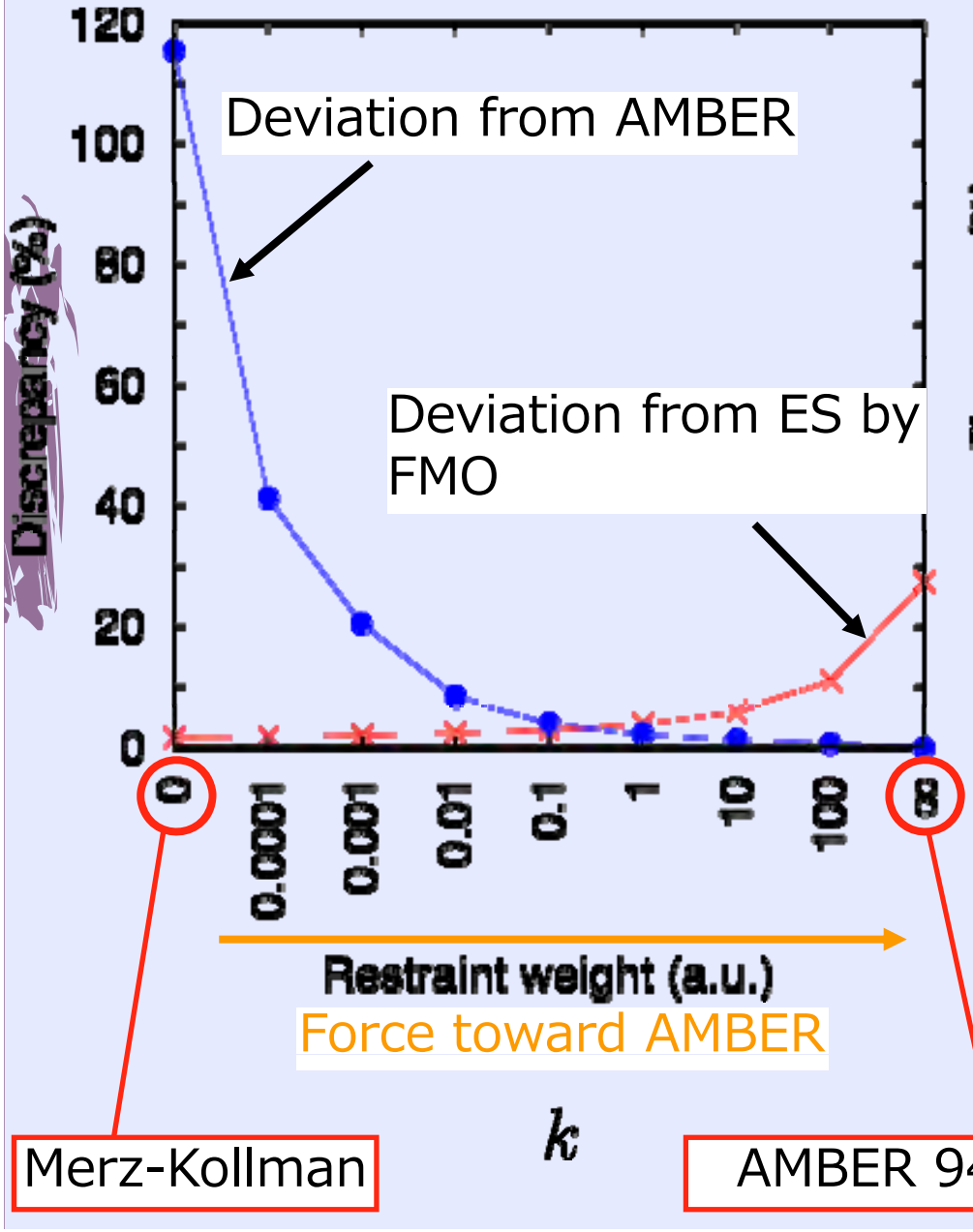


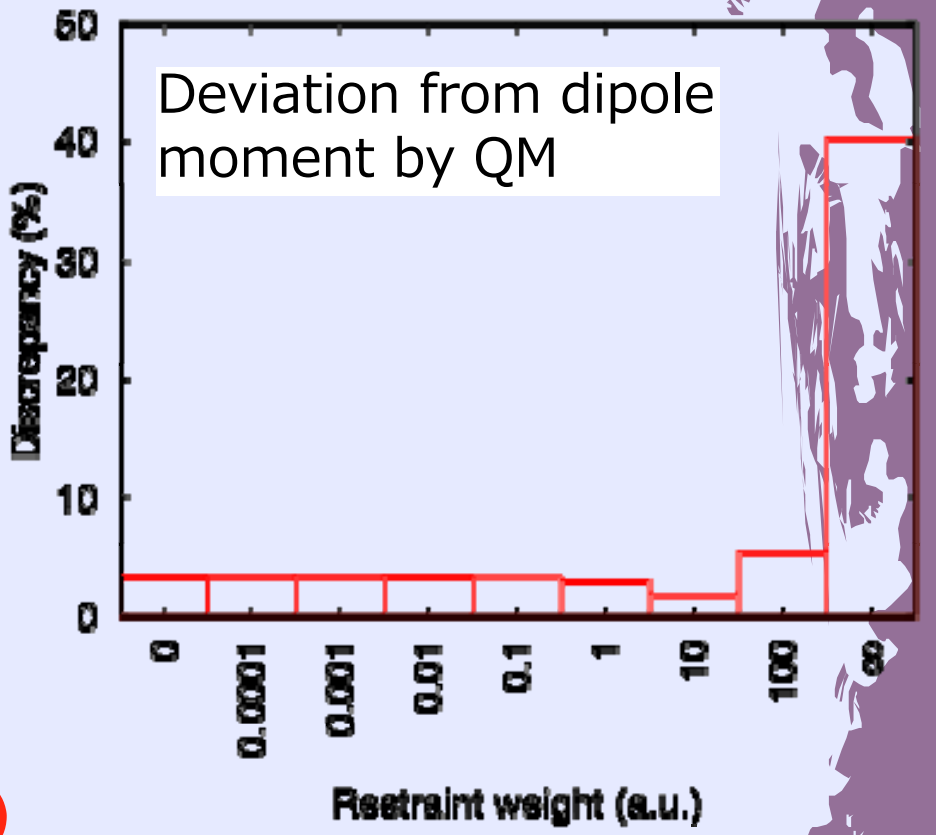
Reference charge: AMBER 94

・AMBER: well established force field

・Use of ESP charge (RESP)

Reproducibility of ES & Similarity of Charge



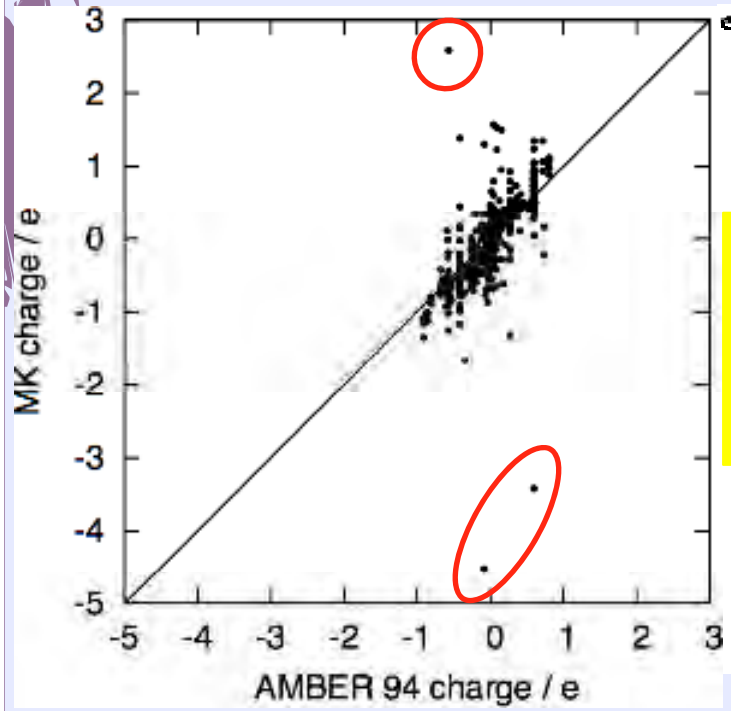


Approaching AMBER charges without loss of accuracy in ES

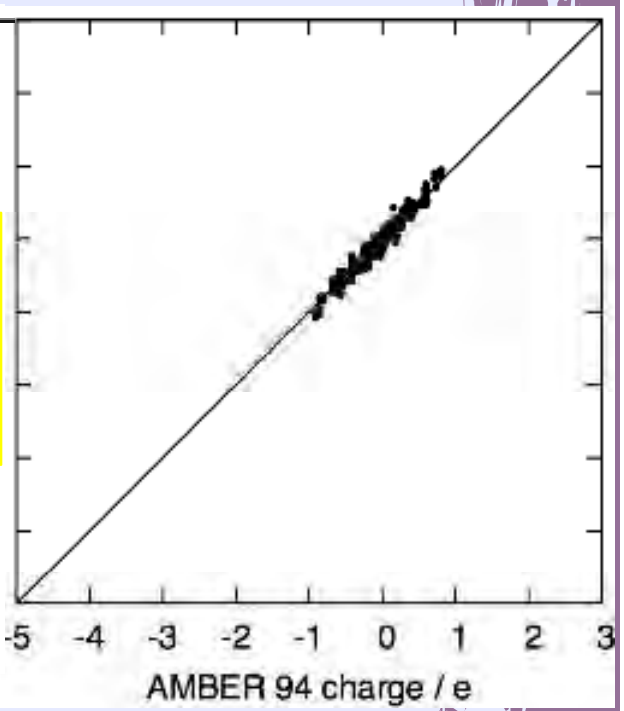
ESP vs. AMBER charges

Force toward AMBER charge

k=0 k=0.0001 k=0.001



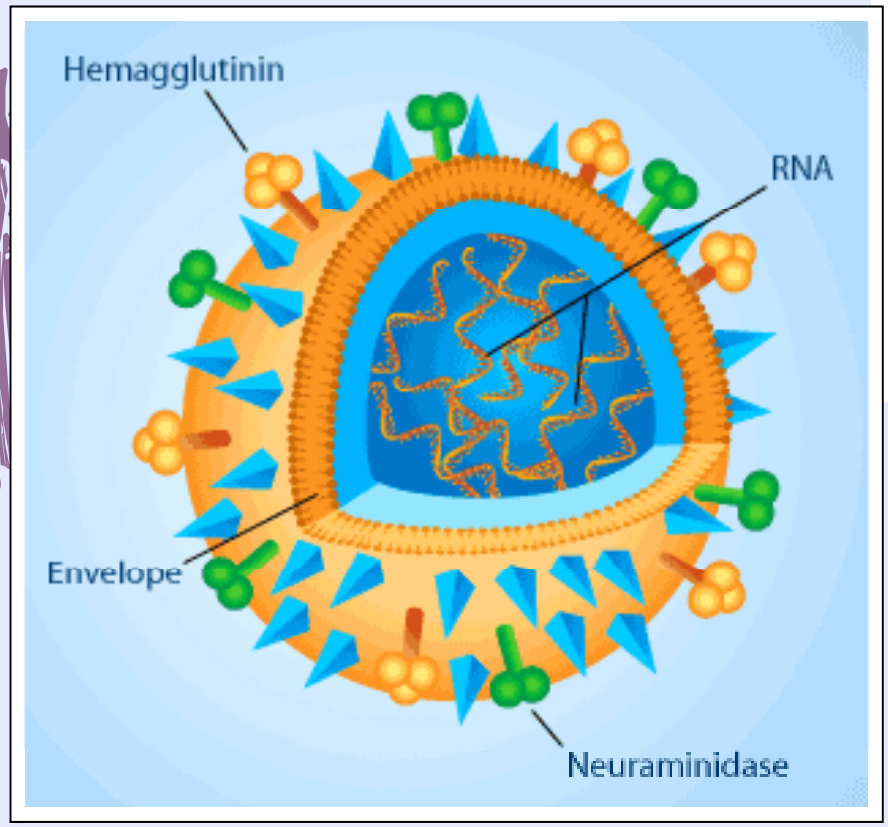
Good reproduction of ES without significant deviation from AMBER charges



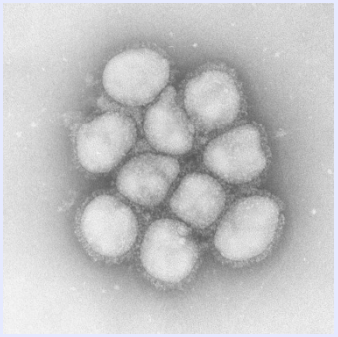
(Y. Okiyama *et al.,* Chem. Phys. Lett. 467 (2009) 417.)

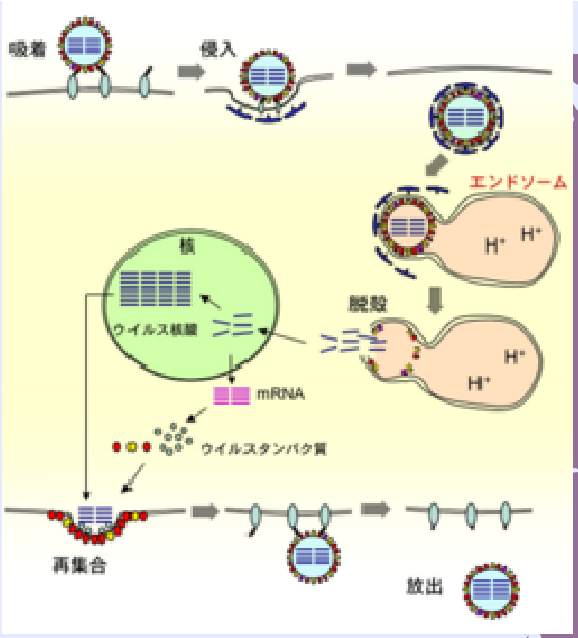
インフルエンザウイルスの構造

インフルエンザウイルス



- ウイルスは膜に覆われた球状の構造
- 膜表面からは，三量体タンパク質ヘマグルチニン(HA)や四量体タンパク質ノイラミニダーゼ(NA)がスパイク状に突き出している
- HAはウイルスの細胞侵入時に宿主細胞表面のシアル酸を認識し，細胞膜に結合する
- HAは抗原として中和抗体に認識される（抗原抗体反応）

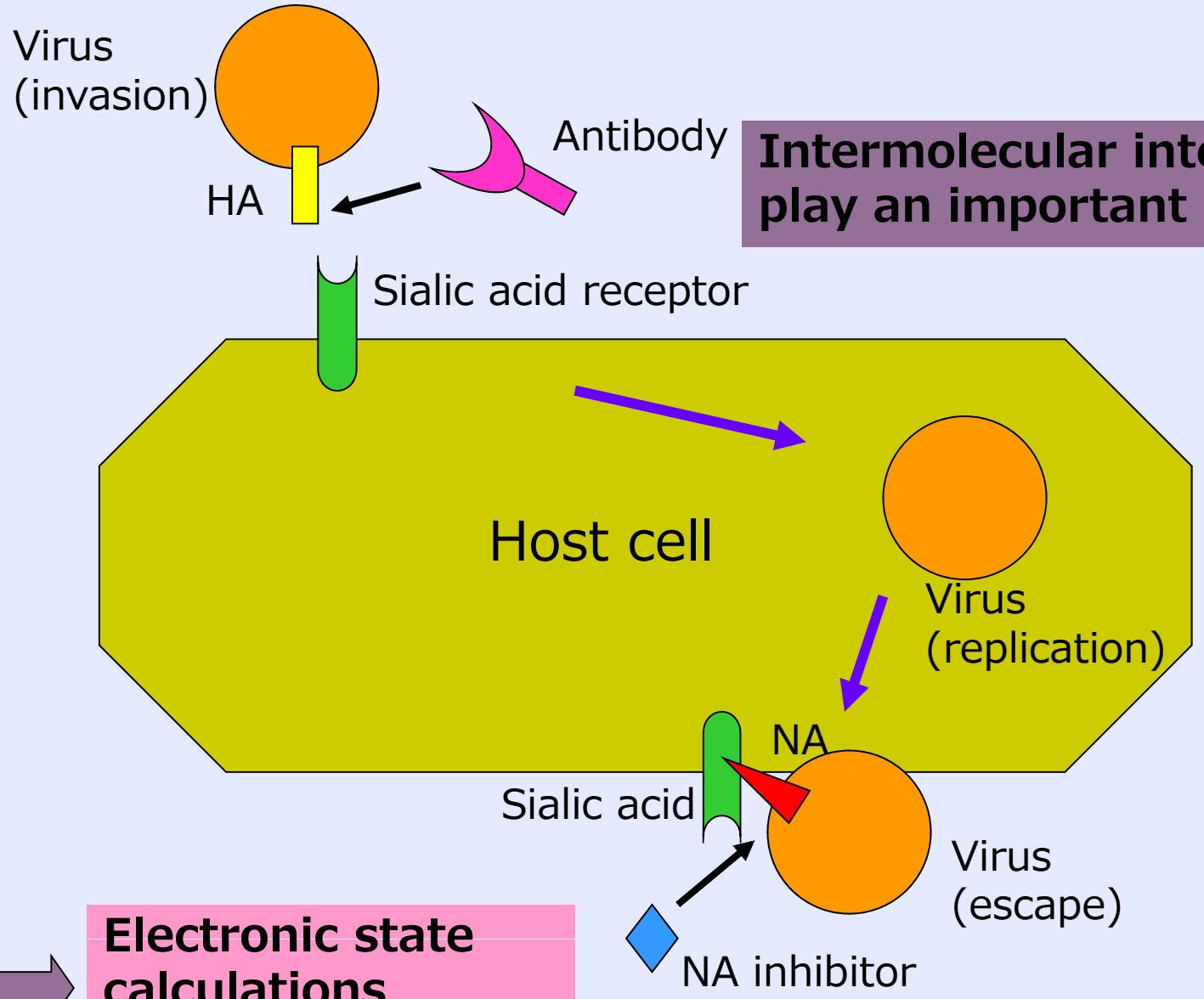




HA：ウイルスが吸着する過程で重要

NA：増殖したウイルスが脱出する過程で重要

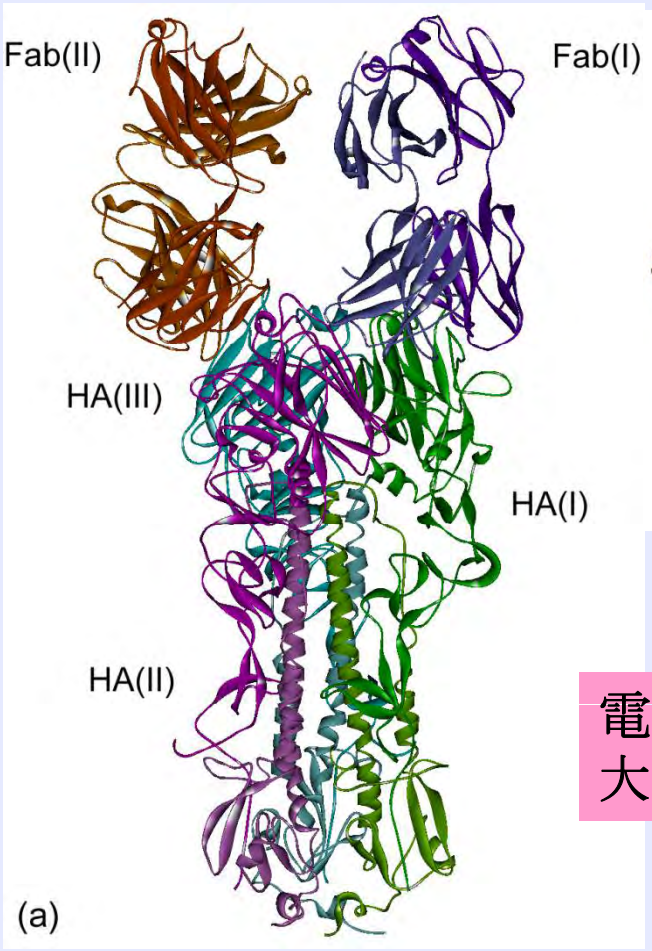
Host cell invasion and replication by influenza virus



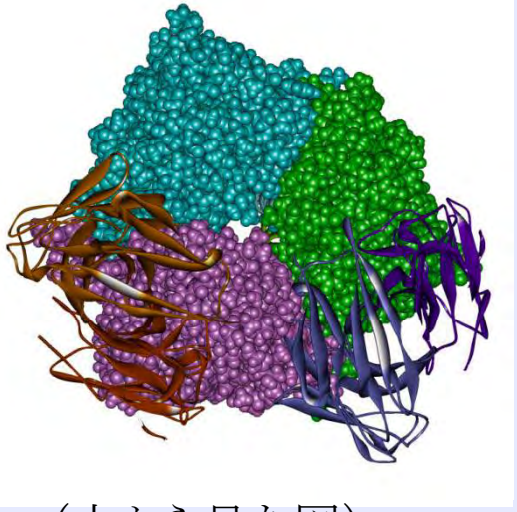
Intermolecular interactions play an important role.

Electronic state calculations

地球シミュレータ(ES2)を用いた大規模FMO-MP3計算

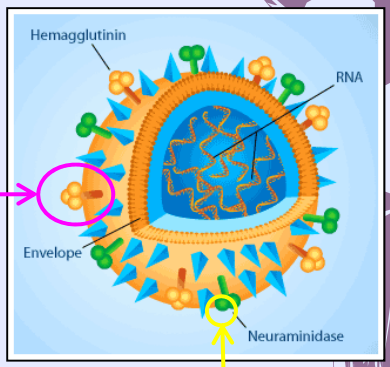


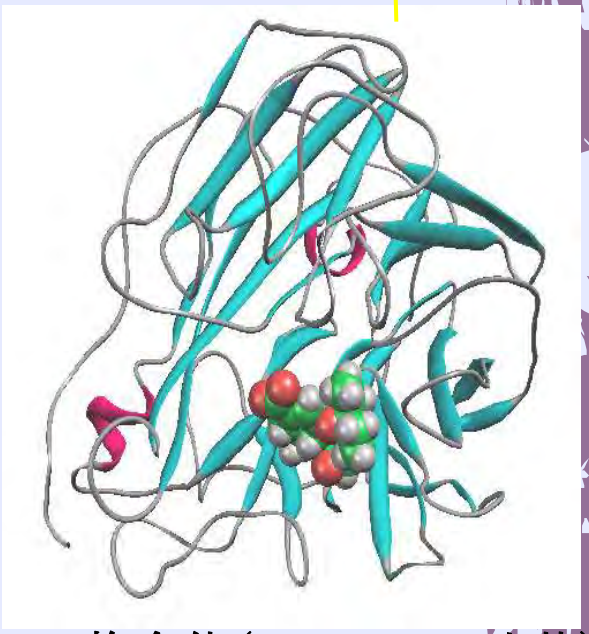
HA3量体－Fab抗体(1KEN;2351残基)



(上から見た図)

電子相関を取り入れた
大規模量子化学計算



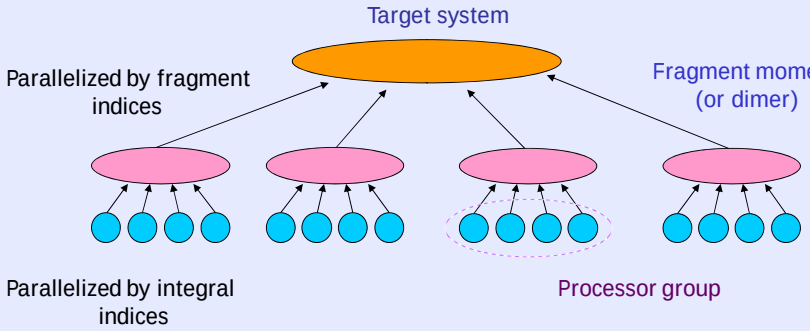


NA－タミフル複合体(2HU4;386残基)

HA３量体：128ノード(計1024プロセッサ)を用いてFMO-MP3/6-31Gが5.8時間
NA単量体：64ノード(計512プロセッサ)を用いてFMO-MP3/6-31Gが1.0時間

Hybrid Parallelization with MPI-OpenMP

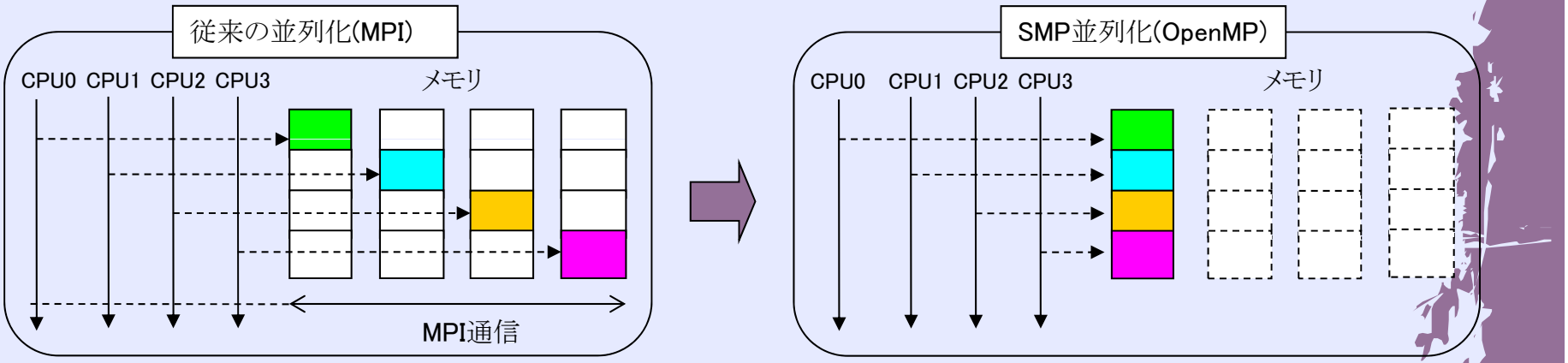
- Inter-fragment: MPI; Intra-fragment: OpenMP
  ⇒ Hybrid parallelization



←MPI

←OpenMP

- Memory saving by sharing MP3 arrays among threads
  ⇒ OpenMP is promising for acceleration on multi-core chips



ES2 was used as a massively parallel-vector computational platform.

地球シミュレータ(ES2)での計算時間

系	計算レベル	ノード数	計算時間(h)		倍率	TFLOPS
HA1	FMO-MP2	64	1.7			0.97
	FMO-MP3	64	2.7	(x1.6)		2.27
	FMO-MP2*	64	4.4			1.19
	FMO-MP3*	64	8.7	(x2.0)		3.02
	FMO-MP2	128	0.8		2.1	2.06
	FMO-MP3	128	1.3	(x1.6)	2.1	4.67
HA3	FMO-MP2	64	9.4			0.83
	FMO-MP3	64	11.9	(x1.3)		1.66
	FMO-MP2	128	4.3		2.2	1.83
	FMO-MP3	128	5.8	(x1.3)	2.1	3.44
NA	FMO-MP3	64	1.0			3.04
	FMO-MP3*	64	4.4			3.09

・64ノード=512VPUs、Cys-Cysは1フラグメント化、NAはタミフル込み
・HA1（14086原子、 921残基、 911フラグメント、 78390関数（121314: 6-31G*））
・HA3（36160原子、2351残基、2325フラグメント、201276関数）
・NA （ 5792原子、 386残基、 378フラグメント、 32549関数（50447: 6-31G*））
・＊は基底関数として6-31G*を使ったことを表す。

HA三量体と抗体との結合系のFMO-MP3計算

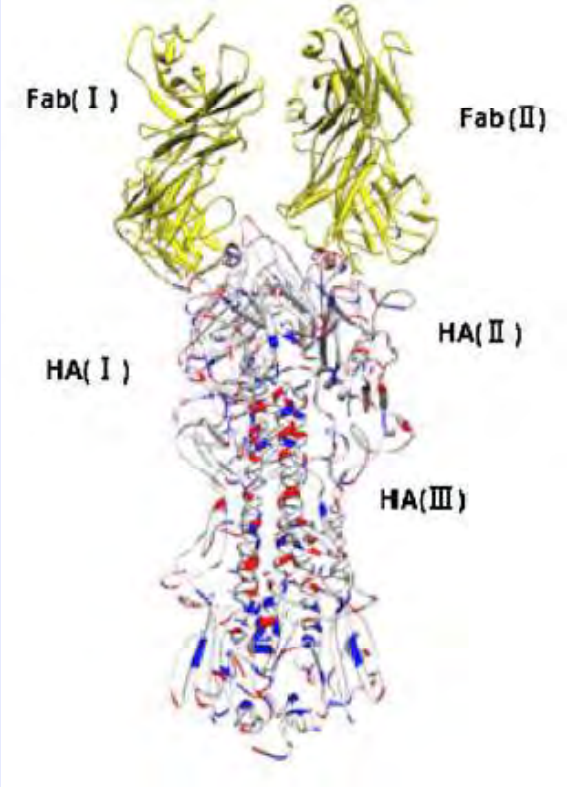
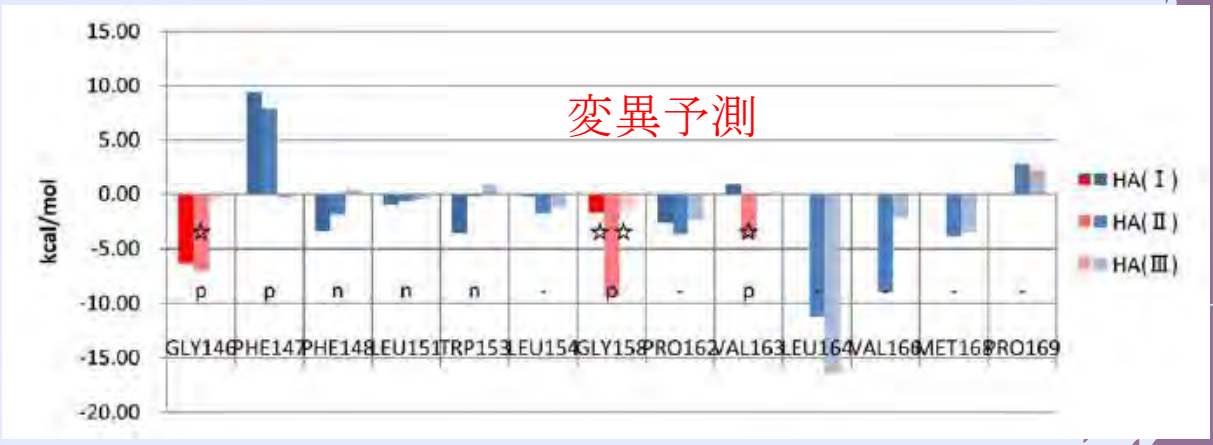


Table 1
IFIE results (in units of kcal/mol) for inter–domain interactions in the complex of HA-trimer (I, II and III) and Fab-dimer (I and II). The calculated values by the HF, MP2, MP3 and MP2.5 methods with the basis set of 6-31G are shown.

Inter–domain	HF	MP2	MP3	MP2.5
Fab dimer–HA trimer	38.1	−163.5	−127.0	−145.3
Fab(I)–HA(I)	−288.8	−367.0	−352.8	−359.9
Fab(I)–HA(II)	177.5	155.5	144.5	150.0
Fab(I)–HA(III)	134.3	134.3	134.3	134.3
Fab(II)–HA(I)	137.0	137.0	137.0	137.0
Fab(II)–HA(II)	−292.7	−380.4	−363.7	−372.0
Fab(II)–HA(III)	170.8	157.0	159.5	158.2
HA(I)–HA(II)	−1022.4	−1280.4	−1237.1	−1258.7
HA(II)–HA(III)	−981.7	−1245.7	−1200.6	−1223.1
HA(I)–HA(III)	−1189.0	−1469.7	−1421.3	−1445.5
Fab(I)–Fab(II)	210.8	197.7	199.5	198.6
Fab dimer–HA(I)	−151.8	−230.0	−215.8	−222.9
Fab dimer–HA(II)	−115.3	−224.9	−205.0	−214.9
Fab dimer–HA(III)	305.1	291.3	293.8	292.6

IFIE value (kcal/mol).

インフルエンザHAタンパク質と抗体の複合体
2351残基、36160原子
FMO calculation

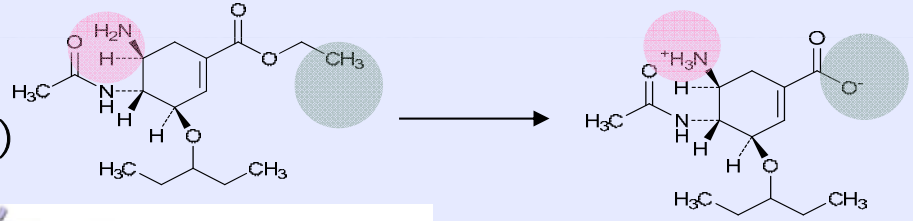


(A. Yoshioka *et al.,* J. Mol. Graph. Model. 30 (2011) 110.)

NAタンパク質と阻害剤(タミフル)の相互作用

ノイラミニダーゼ阻害薬(抗インフルエンザ薬)

タミフル
(**oseltamivir**)



活性体構造

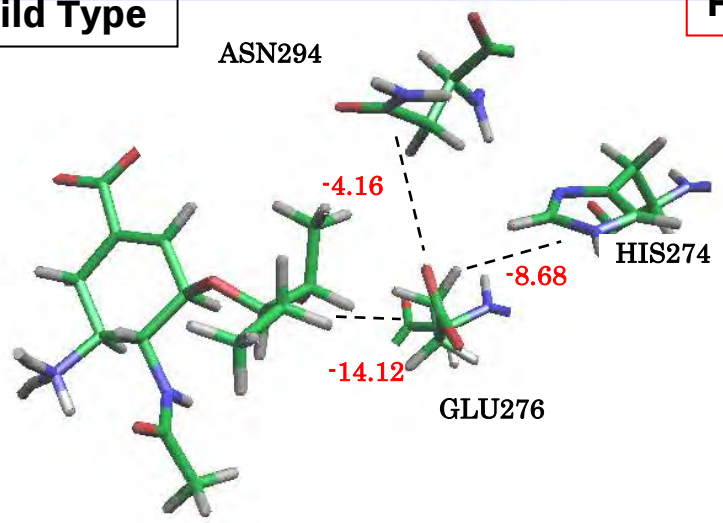
黄色の分子:タミフル(黄色)と
NAの各アミノ酸残基との**IFIE**

[FMO-MP3/6-31G]

- His274Tyr, Asn294Serは、タミフル耐性を持つことが知られている
- 赤色のアミノ酸残基が突然変異を起こすと薬剤耐性を持つ可能性が高い
  ⇒ウイルス変異の予測
- 赤色・青色のアミノ酸残基との相互作用の強さを調節することで、より効果の高い阻害薬の設計ができる
  ⇒論理的創薬を加速

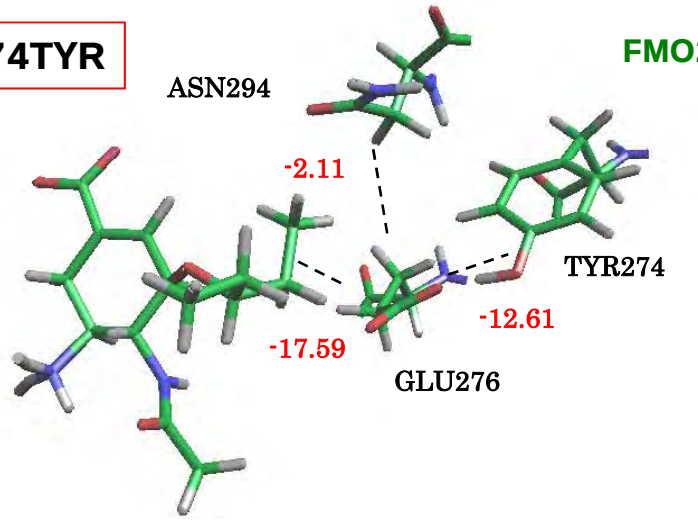
野生株およびタミフル耐性変異株の相互作用

Wild Type

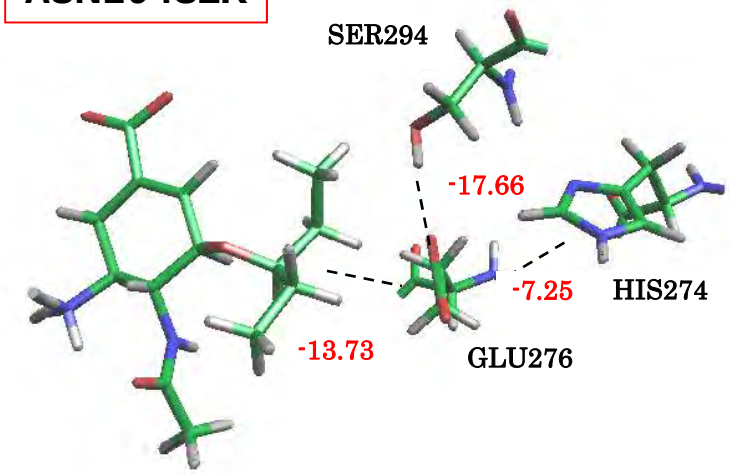


HIS274TYR

FMO2-MP2/6-31G



ASN294SER



赤の数値･･･IFIE(kcal/mol)

$$IFIE-SUM=\sum_{I}\Delta E_{LI}$$

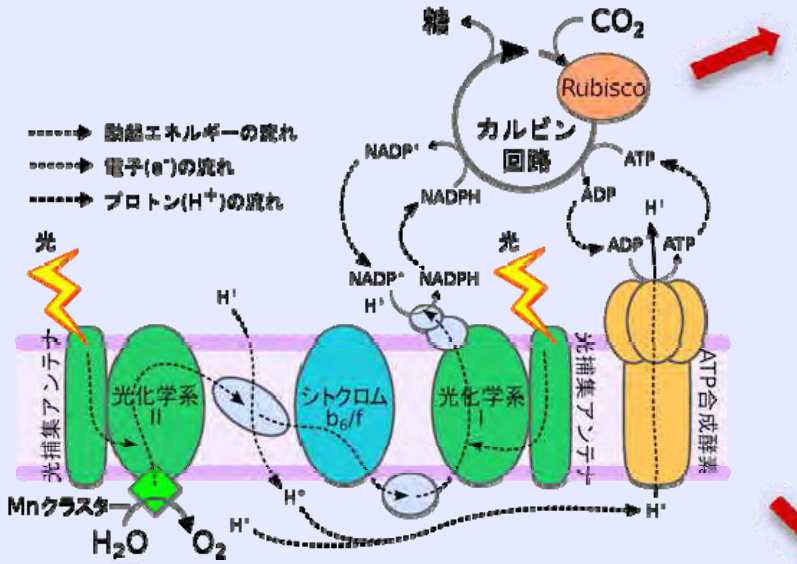
	Wild Type	His274Tyr	Asn294Ser
IFIE-SUM (kcal/mol)	-321.31	-318.80	-321.12
実験値(阻害定数)	1	265	81

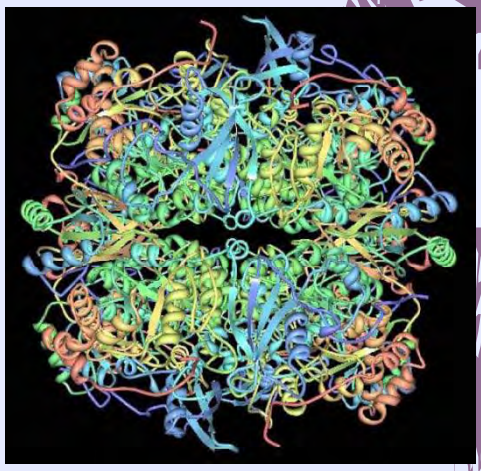
IFIE-SUMと実験値の比較

変異体の計算において結合実験の順番を再現できる

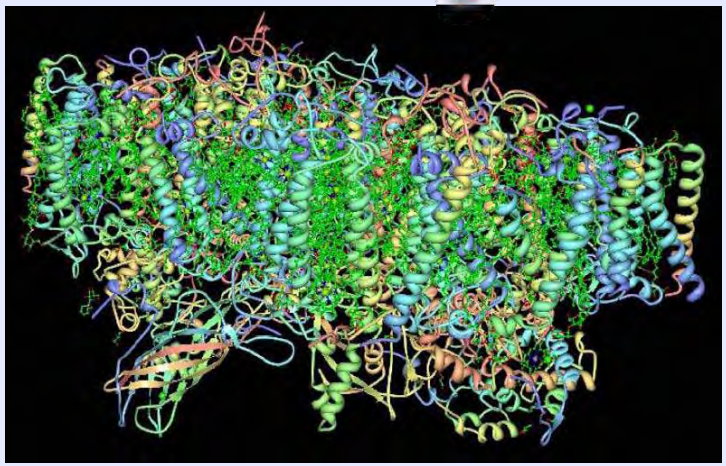
Systems Approach to Photosynthesis

Oxygen-evolving photosynthetic system





Rubisco



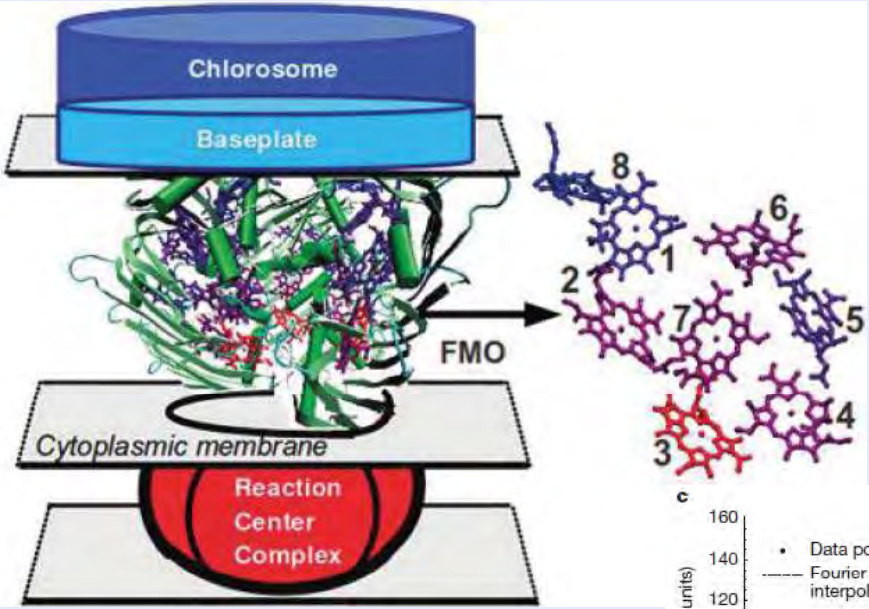
PS II

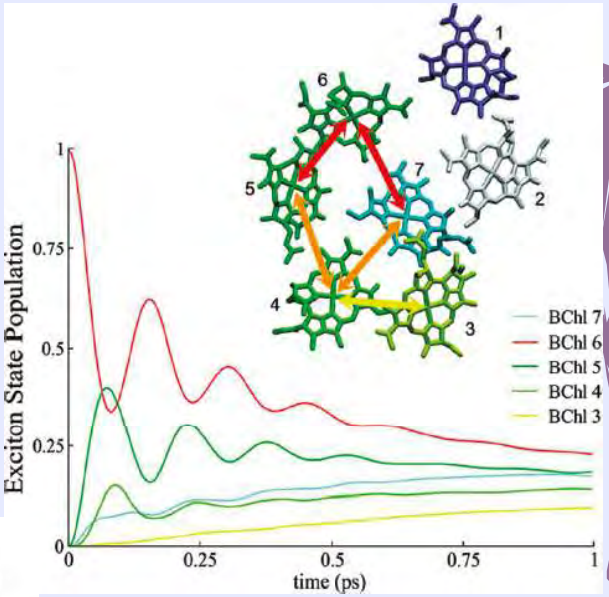
ATP synthase

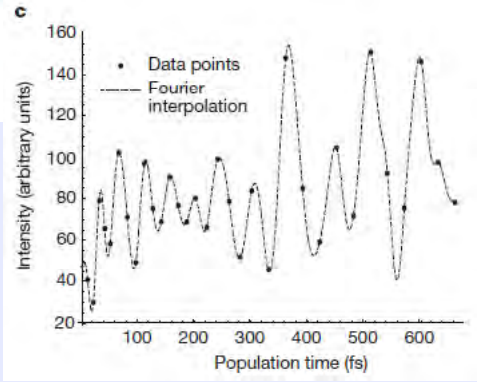
励起エネルギー移動、
電子移動が重要

Excitation Energy Transfer in FMO Protein

Fenna-Matthews-Olson (FMO) light-harvesting protein connects the outer antenna system (chlorosome/baseplate) with the reaction center in green sulfur bacteria.







Quantum beat in two-dimensional Fourier transform electronic spectroscopy

(Engel *et al.,* 2007)

The FMO complex has recently become a paradigmatic model system in terms of the long-lived electronic quantum coherence associated with highly efficient excitation energy transfer that has been experimentally observed in photosynthetic systems.

Description of excitation energy transfer (I)

$$H = H_0 + V_0,$$ Homodimer embedded in biomolecular environment

$$H_0 = H_D|D\rangle\langle D| + H_A|A\rangle\langle A|$$

$$V_0 = J\left(|A\rangle\langle D| + |D\rangle\langle A|\right)$$

H_D and H_A contain phonon modes.

$$i\hbar\frac{\mathrm{d}\rho_I(t)}{\mathrm{d}t} = [V_I(t), \rho_I(t)] \equiv L_I(t)\rho_I(t)$$ Liouville-von Neumann equation

$$\begin{aligned} & i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{P}\rho_I(t)\right] \\ &= \mathcal{P}L_I(t)\mathcal{P}\rho_I(t) \\ &\quad + \mathcal{P}L_I(t)\exp_+\left[-\frac{i}{\hbar}\int_0^t \mathrm{d}t_1 \mathcal{Q}L_I(t_1)\right]\mathcal{Q}\rho_I(0) \\ &\quad - \frac{i}{\hbar}\int_0^t \mathrm{d}t_1 \mathcal{P}L_I(t)\exp_+\left[-\frac{i}{\hbar}\int_{t_1}^t \mathrm{d}t_2 \mathcal{Q}L_I(t_2)\right]\mathcal{Q}L_I(t_1)\mathcal{P}\rho_I(t_1) \end{aligned}$$

P, Q=1-P : Projection operators

Description of excitation energy transfer (II)

$$\frac{\mathrm{d}}{\mathrm{d}t}P_D(t) = \int_0^t \mathrm{d}t_1 \left[-M\left(t,t_1\right) P_D\left(t_1\right) + M\left(t,t_1\right) P_A\left(t_1\right)\right]$$

$$\frac{\mathrm{d}}{\mathrm{d}t}P_A(t) = \int_0^t \mathrm{d}t_1 \left[M\left(t,t_1\right) P_D\left(t_1\right) - M\left(t,t_1\right) P_A\left(t_1\right)\right]$$

$$\begin{aligned} M(t,t_1) &= \frac{1}{\hbar^2}\langle V_{DA}(t)V_{AD}(t_1) + V_{DA}(t_1)V_{AD}(t)\rangle \\ &= \frac{1}{\hbar^2}\langle V_{AD}(t)V_{DA}(t_1) + V_{AD}(t_1)V_{DA}(t)\rangle \end{aligned}$$

Renormalized (Reduced) Equation

Memory function can be evaluated through *ab initio* calculation.

Ansatz form for memory function:

$$M(t,t_1) = \frac{2J_0^2}{\hbar^2}\exp\left(-\frac{t-t_1}{\tau_0}\right)\frac{1+\delta\cos[\omega_0(t-t_1)]}{1+\delta}$$

GME can be analytically solved through solution to quartic equation, while essential features can be grasped in the limit of δ >> 1.

➡ Cubic equation:

$$s^3 + \frac{2}{\tau_0}s^2 + \left(\frac{1}{\tau_0^2} + \omega_0^2 + \frac{4J_0^2}{\hbar^2}\right)s + \frac{4J_0^2}{\hbar^2}\cdot\frac{1}{\tau_0} = 0,$$

Description of excitation energy transfer (III)

$$M(t,t_1) = \frac{2J_0^2}{\hbar^2}\exp\left(-\frac{t-t_1}{\tau_0}\right)\cos\left[\omega_0\left(t-t_1\right)\right]$$

$$F(t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \mathrm{d}s\; e^{st}\left[s + \frac{4J_0^2}{\hbar^2}\cdot\frac{s+1/\tau_0}{\left(s+1/\tau_0\right)^2+\omega_0^2}\right]^{-1}$$

$$= \frac{(\alpha+1/\tau_0)^2+\omega_0^2}{(\alpha-\beta)(\alpha-\gamma)}e^{\alpha t} + \frac{(\beta+1/\tau_0)^2+\omega_0^2}{(\beta-\alpha)(\beta-\gamma)}e^{\beta t} + \frac{(\gamma+1/\tau_0)^2+\omega_0^2}{(\gamma-\alpha)(\gamma-\beta)}e^{\gamma t}$$

Here, for notational simplicity, we introduce parameters, $a = (1/\tau_0)^2$, $b = \omega_0^2$, and $c = 4J_0^2/\hbar^2$. The solutions to the cubic Eq. (20) are then given by

$$\alpha = (-q+\sqrt{r})^{1/3} + (-q-\sqrt{r})^{1/3} - 2\sqrt{a}/3, \quad (22)$$

$$\beta = \omega_2(-q+\sqrt{r})^{1/3} + \omega_3(-q-\sqrt{r})^{1/3} - 2\sqrt{a}/3, \quad (23)$$

$$\gamma = \omega_3(-q+\sqrt{r})^{1/3} + \omega_2(-q-\sqrt{r})^{1/3} - 2\sqrt{a}/3, \quad (24)$$

where we employ $\omega_2 = \left(-1+\sqrt{3}i\right)/2$, $\omega_3 = \left(-1-\sqrt{3}i\right)/2$, and

$$r = q^2 + p^3, \quad (25)$$

$$q = -\sqrt{a}(2a+18b-9c)/54, \quad (26)$$

$$p = (-a+3b+3c)/9. \quad (27)$$

The discriminant for the cubic Eq. (20) is given by

$$\Delta = -4a^2b - a(8b^2 - 20bc - c^2) - 4(b+c)^3, \quad (28)$$

$$P_A(t) = \frac{1}{2}\left[1 - F(t)\right]$$

$$P_D(t) = \frac{1}{2}\left[1 + F(t)\right]$$

When $\Delta > 0$, α, β, γ are real. Otherwise, β and γ become complex conjugates.

Description of excitation energy transfer (IV)

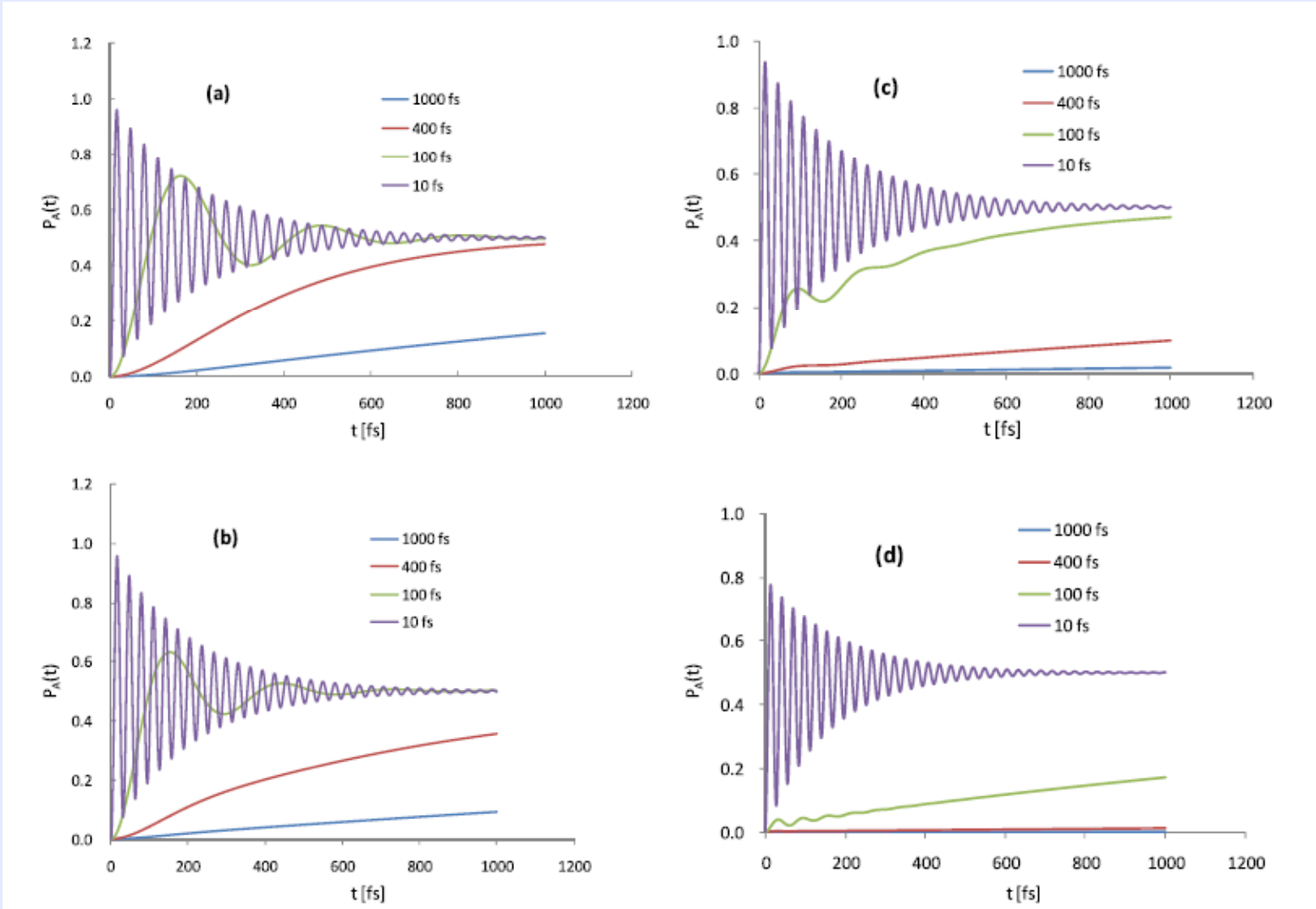


Figure 1. Temporal evolutions of the population density of exciton at the acceptor site, $P_A(t)$, for $\tau_0 = 100$ fs and $\hbar/J_0 = 1000$ (blue), 400 (red), 100 (green), 10 (purple) fs. (a) $\omega_0^{-1} = 1000$ fs. (b) $\omega_0^{-1} = 100$ fs. (c) $\omega_0^{-1} = 30$ fs. (d) $\omega_0^{-1} = 10$ fs.

(ST, Chem. Phys. Lett. 508 (2011) 139.)

Nuclear Quantum Effect on Electron Transfer

- ET in DNA

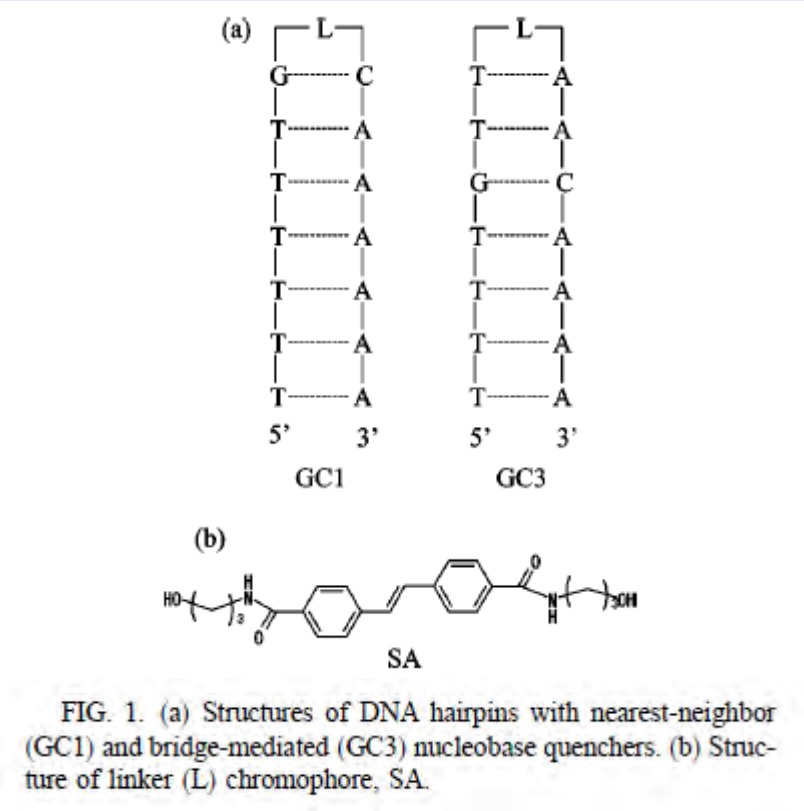
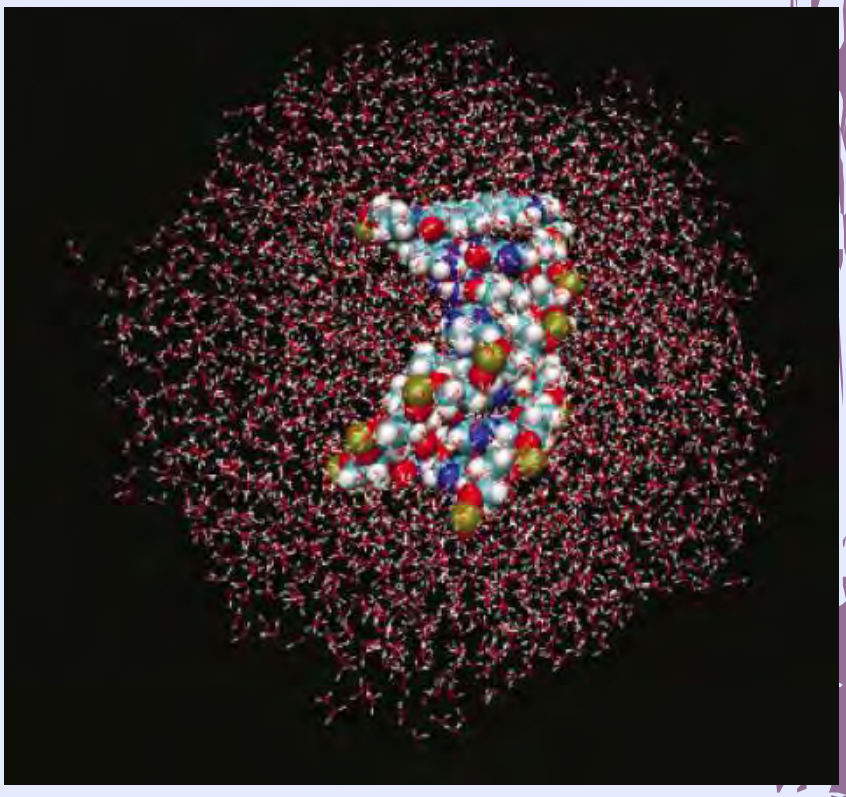


FIG. 1. (a) Structures of DNA hairpins with nearest-neighbor (GC1) and bridge-mediated (GC3) nucleobase quenchers. (b) Structure of linker (L) chromophore, SA.



(F.D. Lewis *et al.,* J. Am. Chem. Soc. 122 (2000) 12346.)

Hole transfer in DNA

$$k = \frac{V^2}{\hbar^2}\mathrm{Re}\left(\int_{-\infty}^{\infty} dt \exp[F(t)] \right)$$

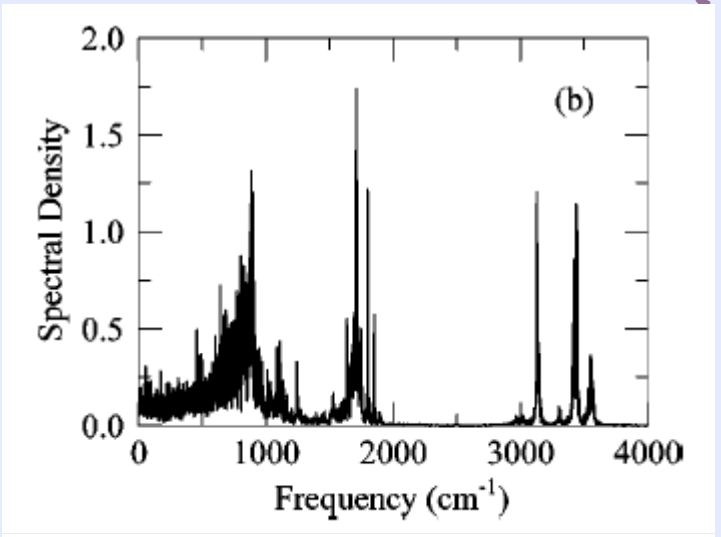
$$F(t) = \frac{i}{\hbar}\Delta G t + \frac{2}{\pi\hbar}\int_0^{\infty} d\omega \frac{J(\omega)}{\omega^2} \times \left[(\cos\omega t - 1)\coth\left(\frac{\hbar\omega}{2k_B T}\right) + i\sin\omega t \right],$$

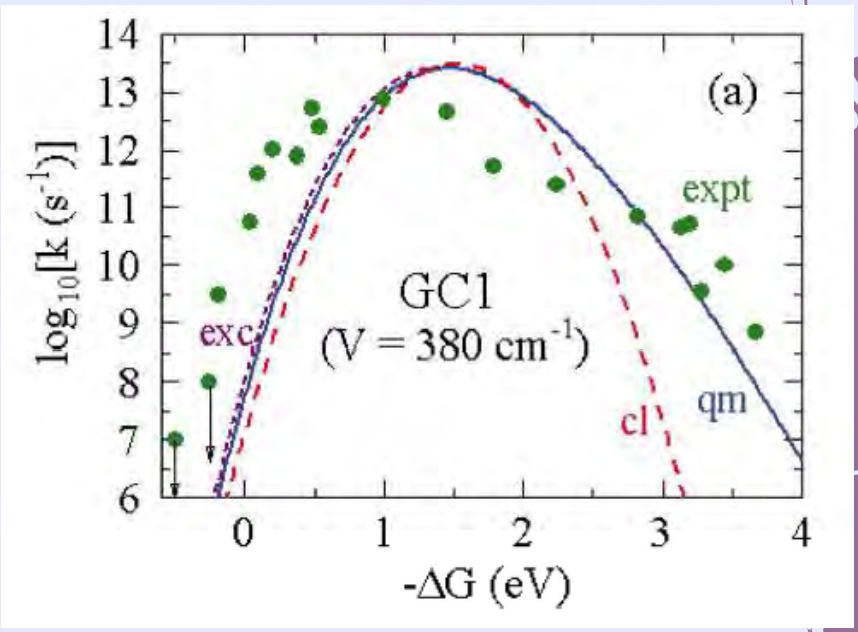
$$\langle \delta\Delta V(0)\,\delta\Delta V(t)\rangle = \frac{4k_B T}{\pi}\int_0^{\infty} d\omega \frac{J(\omega)}{\omega}\cos\omega t.$$

spectral density

$$\Delta V = H_i - H_f,$$

$$\lambda = \frac{2}{\pi}\int_0^{\infty} d\omega \frac{J(\omega)}{\omega}.$$





Microscopic information is incorporated through some parameters (V, λ, ΔG) and functions (J(ω)).

(ST & Y. Sengoku, Phys. Rev. E 68 (2003) 031905.)

Electron Transfer Rate in Biomolecular Systems

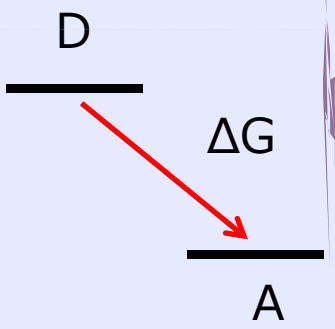
$$k_{\mathrm{DA}} = \frac{2\pi}{\hbar}\Big\langle \sum_{\nu} |\langle \Psi_{\mathrm{iu}}(\vec{r},\vec{R})|\hat{T}^{\mathrm{DA}}|\Psi_{\mathrm{fv}}(\vec{r},\vec{R})\rangle_{\vec{r},\vec{R}}|^2 \delta(E_{\mathrm{iu}} - E_{\mathrm{fv}})\Big\rangle_{\mathrm{T}},$$

$$k_{\mathrm{DA}} = \frac{1}{\hbar^2}\int_{-\infty}^{\infty} dt\ \langle T_{\mathrm{DA}}^{\mathrm{qc}}(t) T_{\mathrm{AD}}^{\mathrm{qc}}(0)\rangle_{\mathrm{T}} \langle I(t)\rangle_{\mathrm{T}}.$$

$$\langle I(t)\rangle_{\mathrm{T}} = \exp\left[-\frac{i}{\hbar}(\Delta G + \lambda)\, t - \frac{1}{\hbar^2}\int_0^t d\tau \int_0^{\tau} d\tau' C(\tau')\right],$$

$$C(t) = \frac{2\hbar}{\pi}\int_0^{\infty} d\omega\ J(\omega)\left[\coth\left(\frac{\hbar\omega}{2k_{\mathrm{B}}T}\right)\cos\omega t - i\sin\omega t\right]$$

Spectral density



Electronic tunneling

$$k_{\mathrm{DA}} = \frac{2\pi}{\hbar^2}\int_{-\infty}^{\infty} d\varepsilon \frac{2}{1+\exp(-\varepsilon/k_{\mathrm{B}}T)} \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dt\ \langle T_{\mathrm{DA}}(t) T_{\mathrm{DA}}(0)\rangle_{\mathrm{T}} \exp\left(\frac{i\varepsilon t}{\hbar}\right)$$
$$\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau \exp\left[-Q_2(\tau) - iQ_1(\tau)\right] \exp\left[-i\frac{(\Delta G + \varepsilon)}{\hbar}\tau\right],$$

Bath coupling

ET rate can be estimated in terms of temporal correlation functions obtained through MD and FMO.

(ST and E.B. Starikov, Phys. Rev. E 81 (2010) 027101.)

$$Q_1(t) = \frac{2}{\pi\hbar}\int_0^\infty \frac{d\omega}{\omega^2} J(\omega)\sin\omega t$$

$$Q_2(t) = \frac{2}{\pi\hbar}\int_0^\infty \frac{d\omega}{\omega^2} J(\omega)\coth\left(\frac{\hbar\omega}{2k_BT}\right)(1-\cos\omega t).$$

$$\lambda = \frac{2}{\pi}\int_0^\infty \frac{d\omega}{\omega} J(\omega),$$

$$A(t) = [\langle T_{DA}(t)T_{DA}(0)\rangle_T - \langle T_{DA}\rangle_T^2]/(\langle T_{DA}^2\rangle_T - \langle T_{DA}\rangle_T^2),$$

$$A(t) = \exp[-\gamma|t|] \qquad \gamma = 1/\tau_c.$$

(inelastic electron tunneling and nuclear quantum effects)

Test calculations using the dielectric function of water

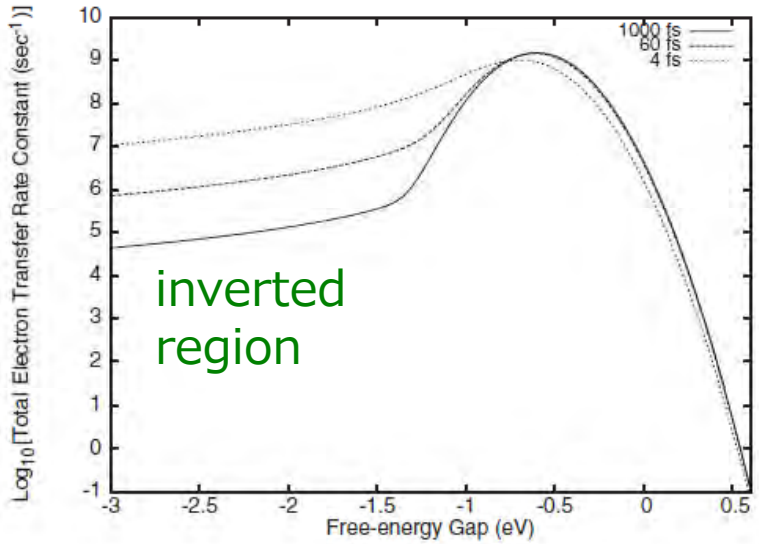


FIG. 1. ET rate constant $k_{DA}(\Delta G)$ calculated by Eq. (18) in the case of λ=0.6 eV for τ_c=4 fs (dotted line), 60 fs (dashed line), and 1000 fs (solid line).

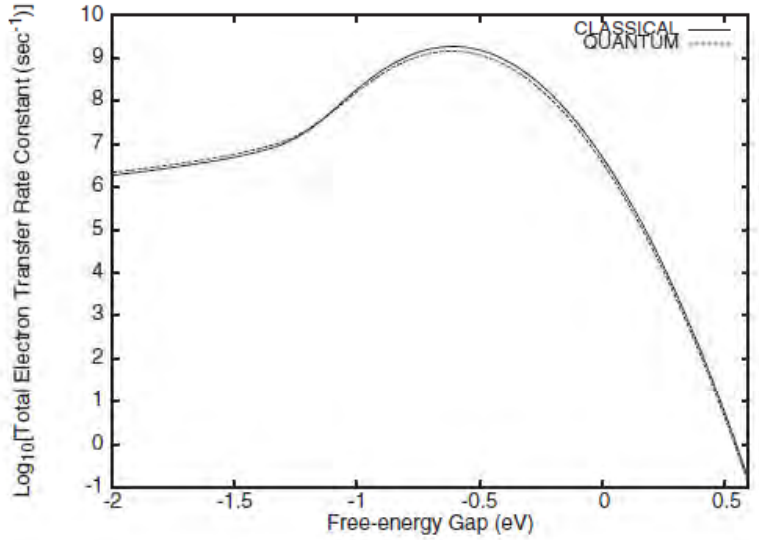


FIG. 2. Comparison of ET rate constants calculated by Eq. (18) (dashed line) and Eq. (23) (solid line) in the case of τ_c=60 fs and λ=0.6 eV.

FMO計算の創薬への応用に向けて

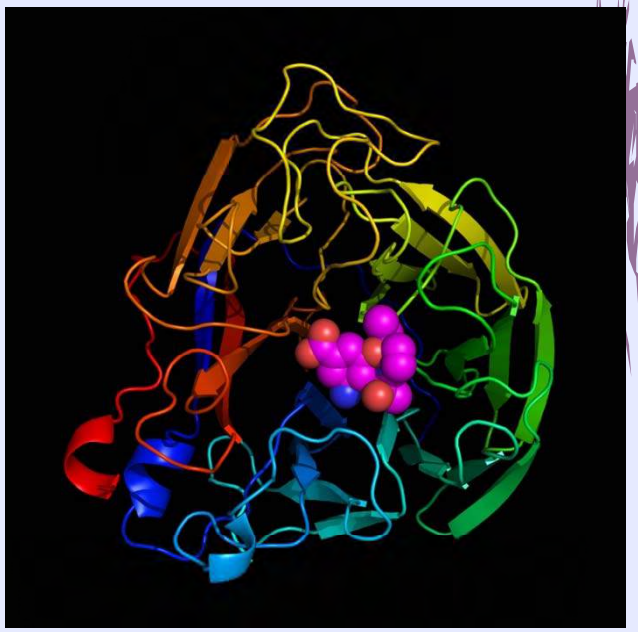
- リガンドータンパク質相互作用の「分解」と解析の高精度化
  - PIEDA、FMO4、BSSE補正
- 溶媒効果の取り入れと自由エネルギー計算
  - Poisson-Boltzmann、エントロピー効果
- 効率的なドッキング計算への展開
  - 力場の改良（電荷、二面角部分、プロトン化状態）
  - 構造最適化（QM vs. MM）
- 新規な技術展開
  - 実験と協働した構造精密化（電子密度解析）

Structure-Based Drug Design

Drug design based on protein-ligand 3D structure

Recent technological developments

- Advancement of computer power
- Elucidation of protein 3D structures
  - Advances in X-ray crystallography & NMR
  - Developments of associated infrastructures (SPring-8, K-Computer, etc.)



Efficient computer-aided (*in silico*) protein-ligand docking simulation

Benefits of computer-aided SBDD

- Calculation of protein-ligand interaction (free) energies
- Detailed analysis of intermolecular interactions due to three-dimensional binding information

Accurate description of molecular force field is a prerequisite for SBDD.

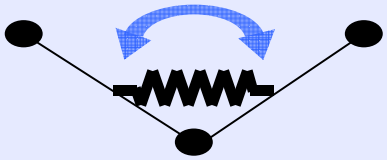
例えば、リガンドの結合自由エネルギーΔGを 1 kcal/mol 以下の精度で計算する必要。

Classical Force Field

Molecular Mechanics (MM)

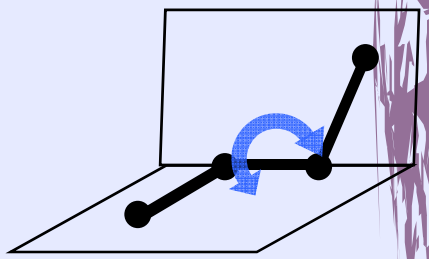
Bond stretching

$$\sum_{\text{bonds}} K_r \left(r - r_{eq}\right)^2$$



Angle bending

$$\sum_{\text{angles}} K_\theta \left(\theta - \theta_{eq}\right)^2$$



Torsion

$$\sum_{\text{dihedrals}} \frac{V_n}{2}\left(1+\cos\left[n\phi-\gamma\right]\right)$$

Electrostatic interaction

$$\sum_{\text{atoms}} \left(\frac{q_i q_j}{\varepsilon R_{ij}} \right)$$

van der Waals

$$\sum_{\text{atoms}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right)$$

Parameters should depend on global conformations and surrounding environment.

Problem: Significance of quantitative accuracy

Classical molecular mechanics/ dynamics based on conventional force fields often give inaccurate results for ligand docking.

従来の力場ではTorsion、ES、vdWなどの記述があまり良くないことが知られている。
また、「力場を変えるとΔGがどのくらい変化するか」、「プロトン化状態によってΔGがどのくらい変化するか」等はあまり詳しく調べられていない。

Protein-Specific Force Field Derived from the Fragment Molecular Orbital Method can Improve Protein–Ligand Binding Interactions

Le Chang,[a,b] Takeshi Ishikawa,[c,d] Kazuo Kuwata,[b,c] and Shoji Takada*[a,b]

Accurate computational estimate of the protein–ligand binding affinity is of central importance in rational drug design. To improve accuracy of the molecular mechanics (MM) force field (FF) for protein–ligand simulations, we use a protein-specific FF derived by the fragment molecular orbital (FMO) method and by the restrained electrostatic potential (RESP) method. Applying this FMO-RESP method to two proteins, dodecin, and lysozyme, we found that protein-specific partial charges tend to differ more significantly from the standard AMBER charges for isolated charged atoms. We did not see the dependence of partial charges on the secondary structure. Computing the binding affinities of dodecin with five ligands by MM PBSA protocol with the FMO-RESP charge set as well as with the standard AMBER charges, we found that the former gives better correlation with experimental affinities than the latter. While, for lysozyme with five ligands, both charge sets gave similar and relatively accurate estimates of binding affinities. © 2013 Wiley Periodicals, Inc.

Journal of Computational Chemistry **2013**, *34*, 1251–1257

DOI: 10.1002/jcc.23250

Table 1. The free energy values (in kcal/mol) from experiments and computations with FMO-RESP charges and AMBER charges for dodecin.

Ligand	Experimental value	FMO-RESP charges	AMBER charges
FAD	−8.7	−13.0 (−10.3)	0.4
FMN	−6.6	−8.0 (−6.6)	−11.3
Riboflavin	−10.2	−12.2 (−15.2)	−19.9
Lumiflavin	−10.6	−13.9 (−13.5)	−14.9
Lumichrome	−10.9	−10.0 (−6.6)	−12.3

The data in brackets are values with ligand partial charges from first stage RESP fitting.

Table 2. The free energy values (in kcal/mol) from experiments and computations with FMO-RESP charges and AMBER charges for lysozyme.

Ligand	Experimental value	FMO-RESP charges	AMBER charges
NAG	−4.5	−22.6 (−19.4)	−6.3
NAG2	−7.2	−21.4 (−29.4)	−30.9
NAG3	−9.3	−41.1 (−35.9)	−37.6
NAG4	−9.6	−61.0 (−56.3)	−73.9
NAG-MUB	−4.1	−17.5 (−19.6)	−15.2

The data in brackets are values with ligand partial charges from first stage RESP fitting.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics 307 (2004) 269–283
Chemical Physics
www.elsevier.com/locate/chemphys

Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations

Takao Yoda [a,1], Yuji Sugita [b], Yuko Okamoto [a,c,*]

[a] Department of Theoretical Studies, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan
[b] Institute of Molecular and Cellular Biosciences, University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan
[c] Department of Functional Molecular Science, The Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan

Received 1 March 2004; accepted 9 August 2004
Available online 25 September 2004

Abstract

Secondary-structure forming tendencies are examined for six well-known protein force fields: AMBER94, AMBER96, AMBER99, CHARMM22, OPLS-AA/L, and GROMOS96. We performed generalized-ensemble molecular dynamics simulations of two peptides. One of these peptides is C-peptide of ribonuclease A, and the other is the C-terminal fragment from the B1 domain of streptococcal protein G. The former is known to form α-helix structure and the latter β-hairpin structure by experiments. The simulation results revealed significant differences of the secondary-structure forming tendencies among the force fields. Of the six force fields, the results of AMBER99 and CHARMM22 were in accord with experiments for C-peptide. For G-peptide, on the other hand, the results of OPLS-AA/L and GROMOS96 were most consistent with experiments. Therefore, further improvements on the force fields are necessary for studying the protein folding problem from the first principles, in which a single force field can be used for all cases.

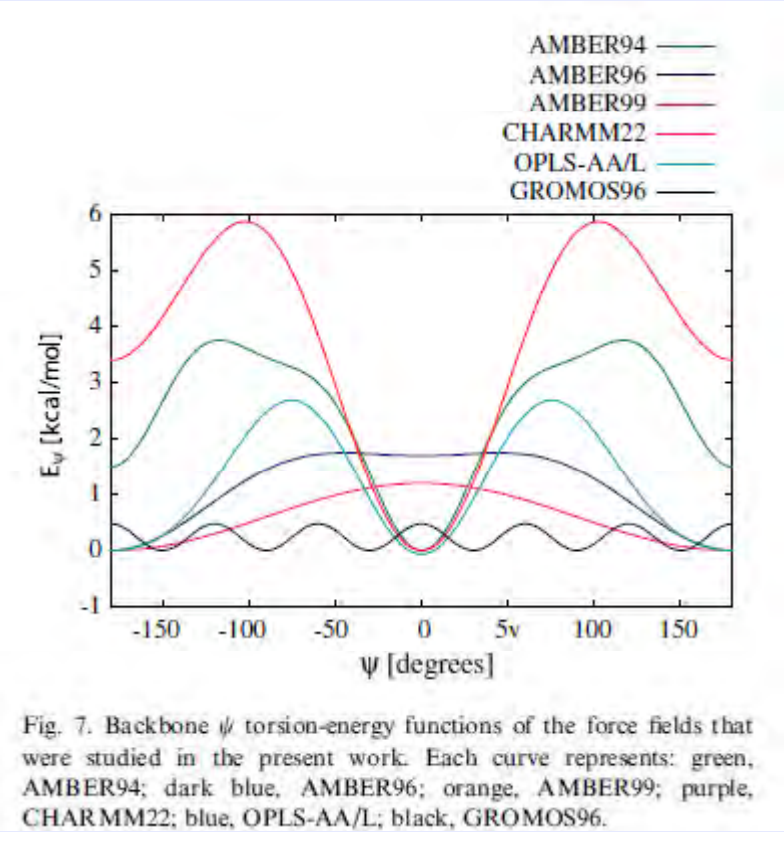


Fig. 7. Backbone ψ torsion-energy functions of the force fields that were studied in the present work. Each curve represents: green, AMBER94; dark blue, AMBER96; orange, AMBER99; purple, CHARMM22; blue, OPLS-AA/L; black, GROMOS96.

Backbone torsional energies in various force fields are diverse substantially.

Only torsional parts are different between AMBER94 and AMBER96.

Torsion terms are a main cause for the difference in preferred secondary structures among various force fields.

(T. Yoda, Y. Sugita & Y. Okamoto, Chem. Phys. 307 (2004) 269.)

(a) C-peptide for ribonuclease A (13 residues)
(b) B1 domain of streptococcal protein G
(G-peptide;16 residues)

C-peptide prefers α-helix structure.
G-peptide prefers β-hairpin structure.

→ **Force field dependence is significant.**

Fragment molecular orbital (FMO) method can provide a nice compromise between cost and accuracy for the description of molecular force field.

生体高分子の効率的な力場計算手法としてFMO法を位置付けることもできる。

Fragment Molecular Orbital (FMO) method

Problems of existing approaches to biomolecules:

- Classical force field is **empirical** for inter-molecular interactions.
- Conventional quantum mechanical calculations is too **expensive**.

Our approach: The FMO method

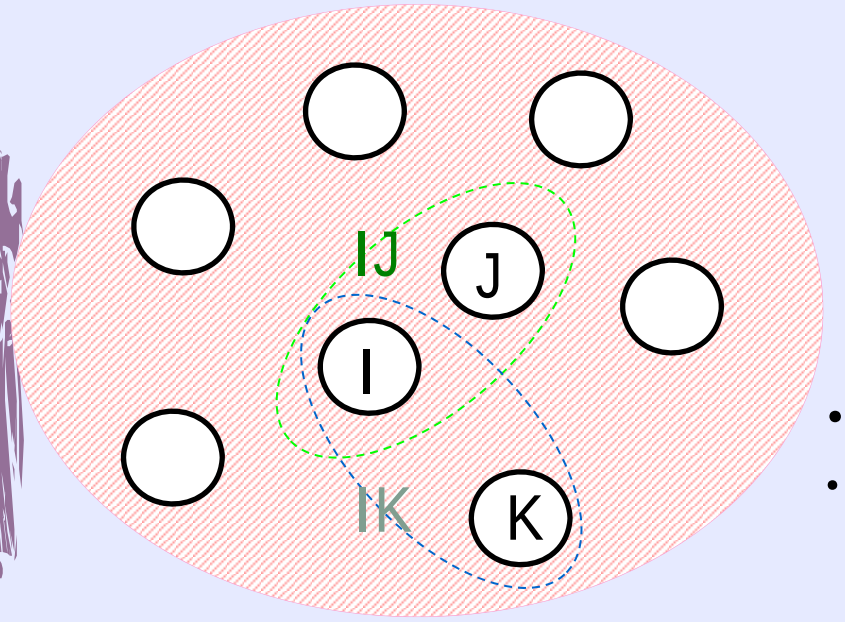
- FMO method was proposed by Kitaura *et al.* (1999).
- Molecules are divided into **fragments.**

- Whole molecular energies are calculated by the sum of fragment and fragment-pair energies.
  → Drastic speed up has been attained !
- Error of FMO total energy is within 0.4 kcal/mol for crambin (46 residues).

Advantage of FMO method: Molecular interaction analysis

- Molecular interaction analysis : receptor-ligand binding
- Inter-fragment interaction energy (IFIE) analysis

FMO Method and Its Energy Analysis (IFIE)



Divide a molecule into fragments

- N pieces of fragments
- [N(N-1)/2] pieces of fragment pairs

trimer, tetramer → FMO3, FMO4

Total Energy: Calculated from energies of fragments and fragment pairs

$$E = \sum_{I>J} E_{IJ} - (N-2)\sum_{I} E_I \qquad \text{(FMO2)}$$

Inter-Fragment Interaction Energy (IFIE):

$$\Delta E_{IJ} = (E'_{IJ} - E'_I - E'_J) + \mathrm{Tr}(\Delta P_{IJ} V_{IJ})$$

E_X : energies of a fragment and a fragment pair

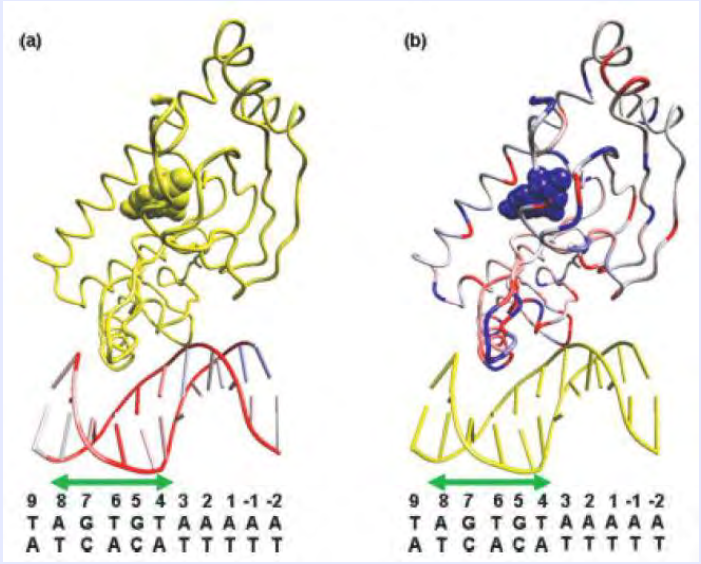
V_X : Electrostatic potential from surrounding fragments

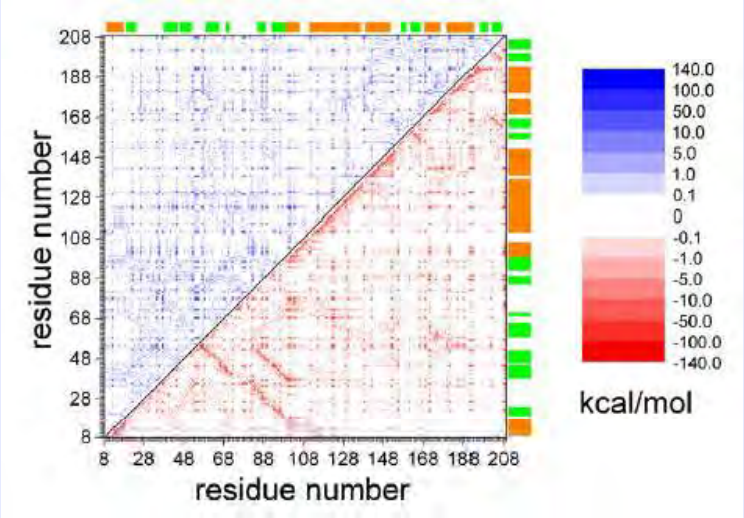
$$E'_X = E_X - V_X$$

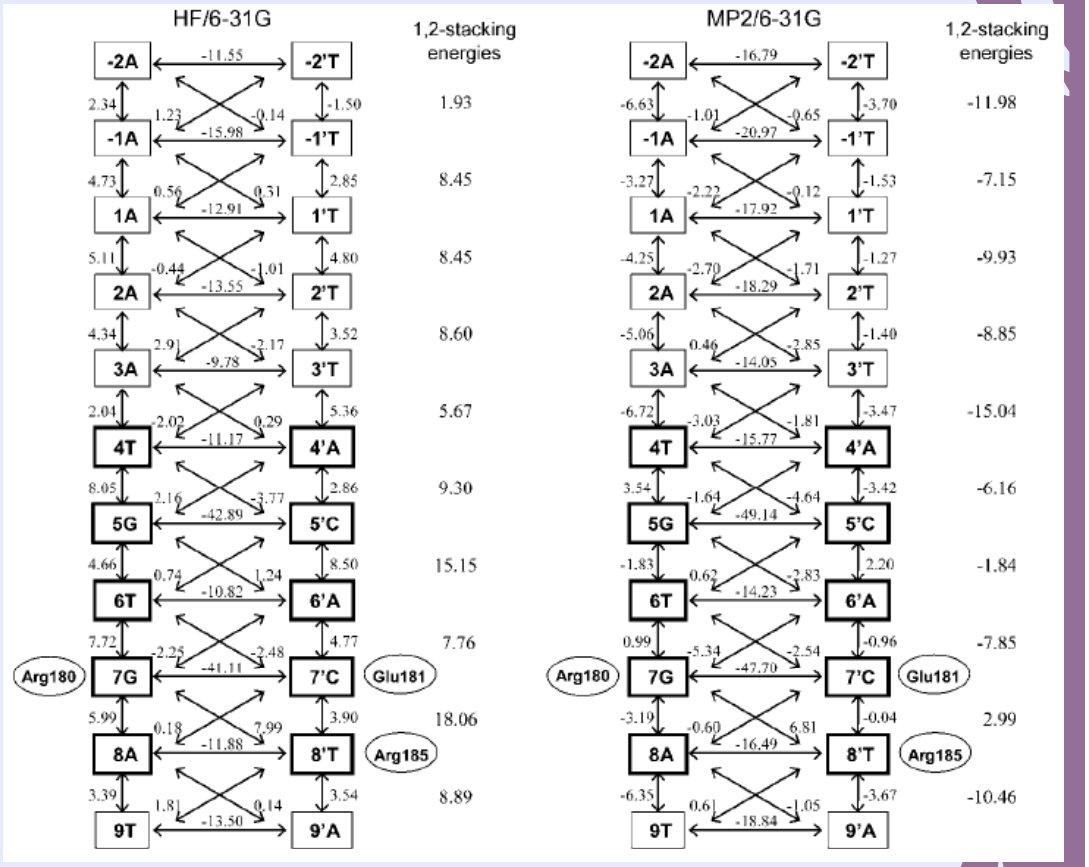
Incorporation of electron correlation effects is essential in the accurate description of biomolecular interactions.

少なくともMP2かDFT-Dレベル、可能ならばCCSD(T)レベルの計算。

DNA-CRP-cAMP system

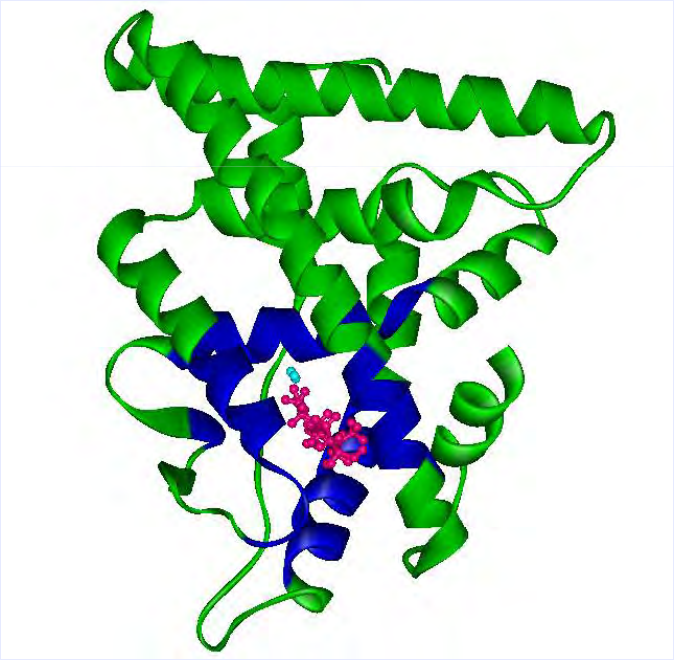




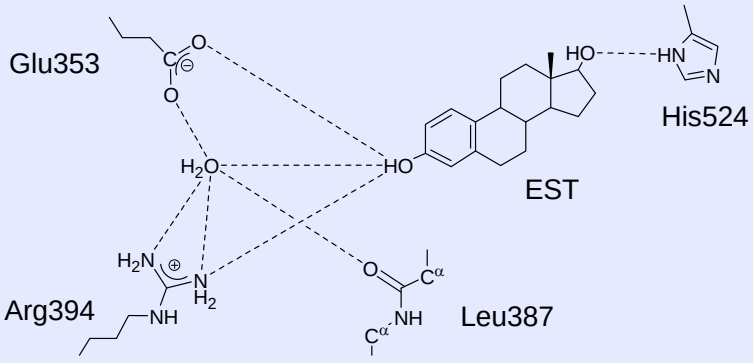


K. Fukuzawa *et al*., J. Comput. Chem. 27 (2006) 948.
I. Kurisaki *et al*., Biophys. Chem. 130 (2007) 1.

Binding Energy Calculations Between ER and Ligands

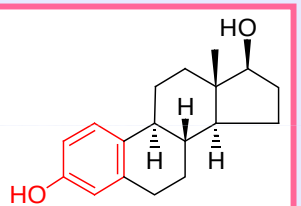


MODEL1（green） (241residues, whole LBD)
MODEL2（blue） (50 residues)

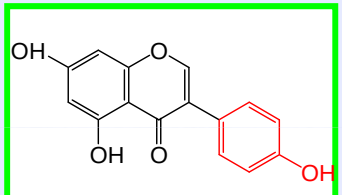


Hydrogen bond network among ER and ligand

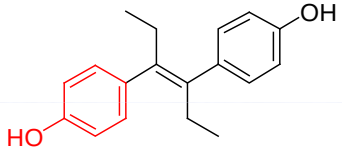
Estrogen, Phytoestrogens, Synthetic estrogens and Industrial chemicals



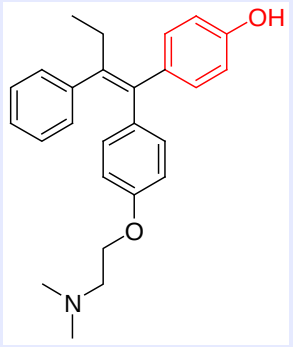
17β-Estradiol (EST)



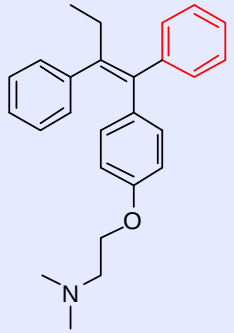
Genistein (GEN)



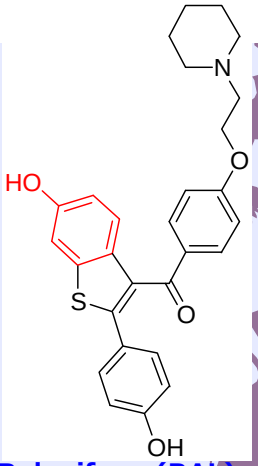
Diethylstilbestrol (DES)



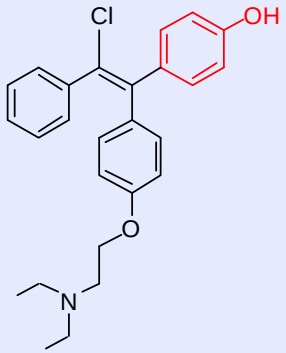
4-hydroxytamoxifen (OHT)



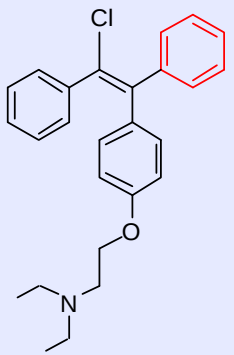
Tamoxifen (TAM)



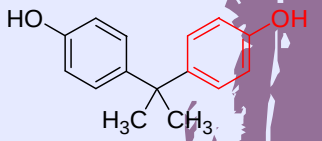
Raloxifene (RAL)



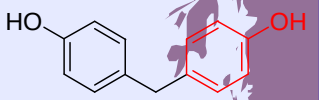
4-hydroxyclomifene (OHC)



Clomifene (CLO)



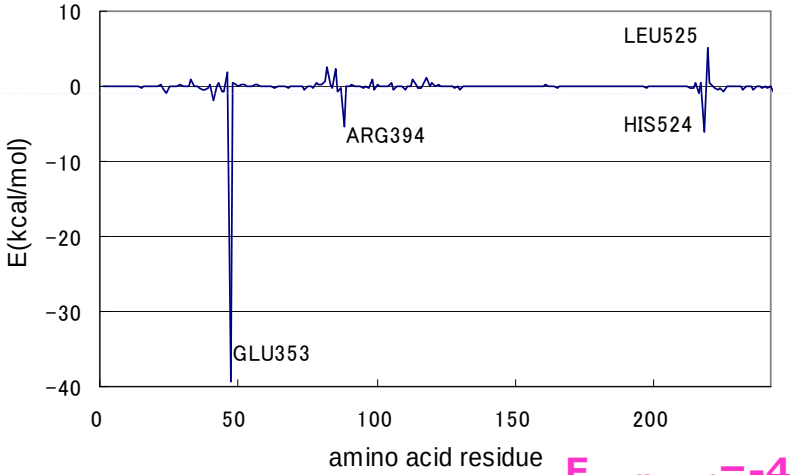
Bisphenol A (BISA)



Bisphenol F (BISF)

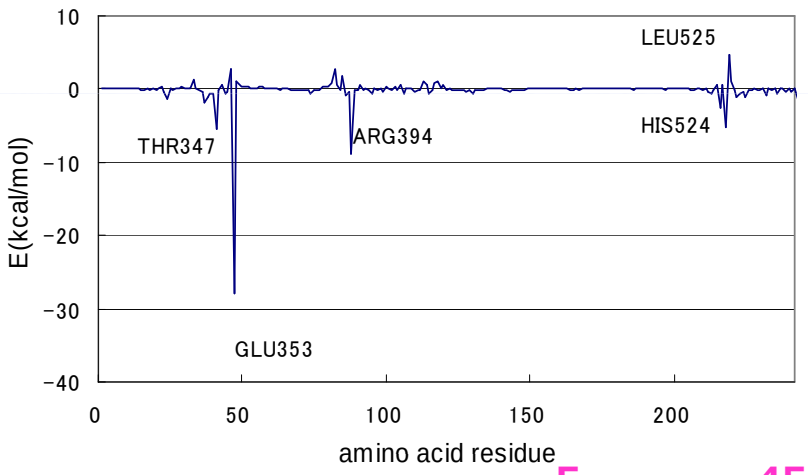
IFIE analysis with various methods and basis sets

FMO-HF/STO-3G



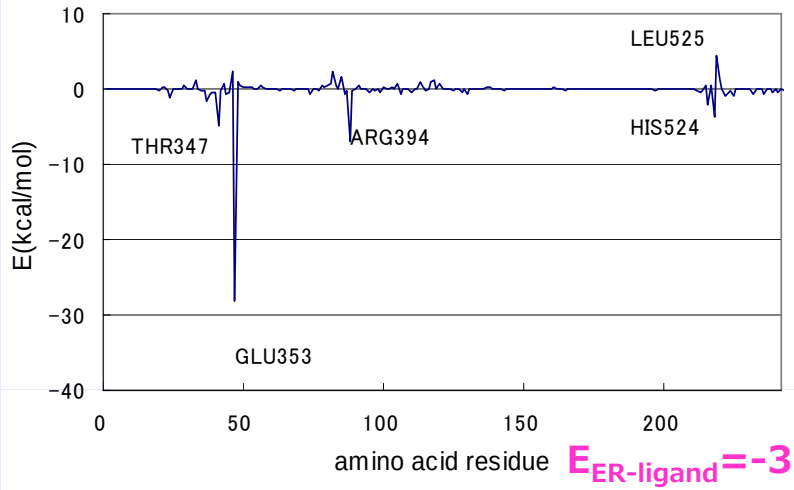
$E_{ER\text{-}ligand}$=-42.456

FMO-HF/6-31G



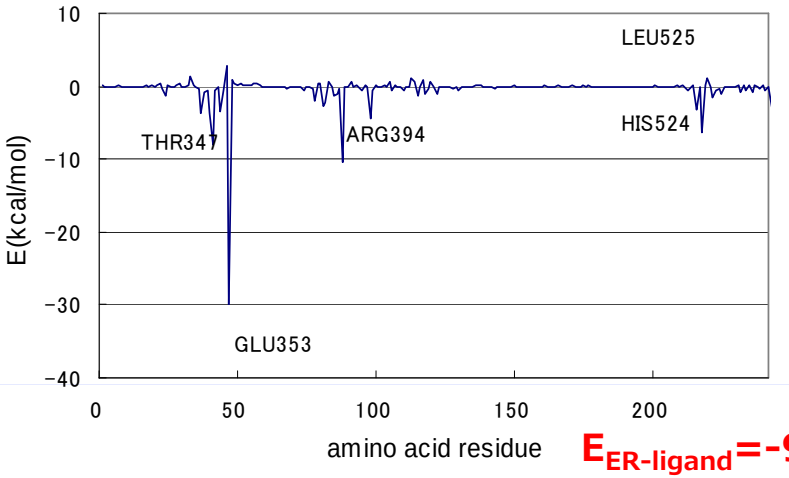
$E_{ER\text{-}ligand}$=-45.476

FMO-HF/6-31G(d,p)



$E_{ER\text{-}ligand}$=-39.340

FMO-MP2/6-31G



$E_{ER\text{-}ligand}$=-93.172

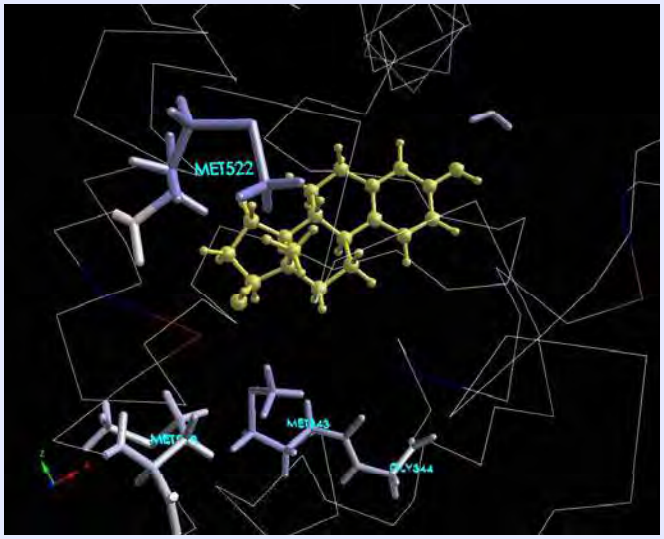
[MODEL1]

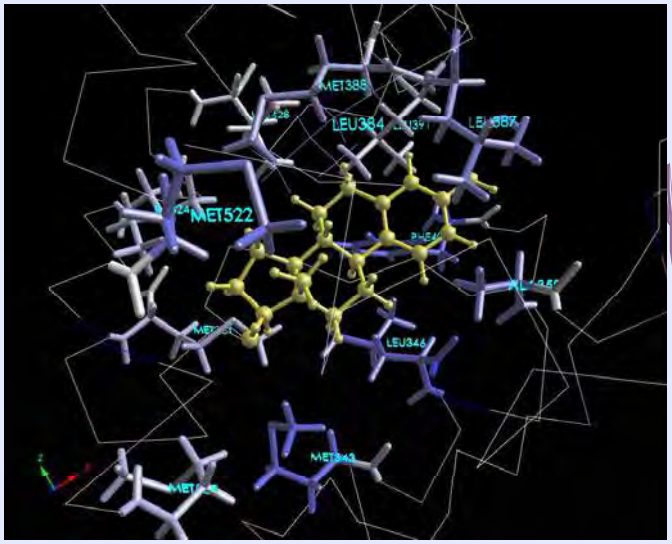
Interaction with surrounding hydrophobic residues

FMO-HF/6-31G

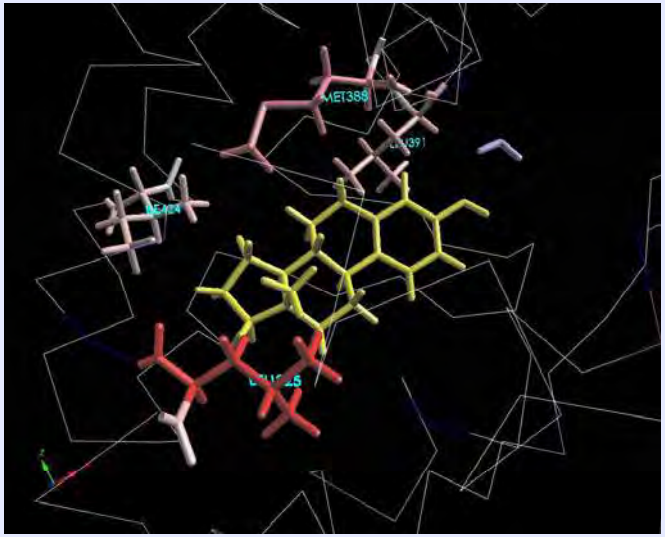
FMO-MP2/6-31G

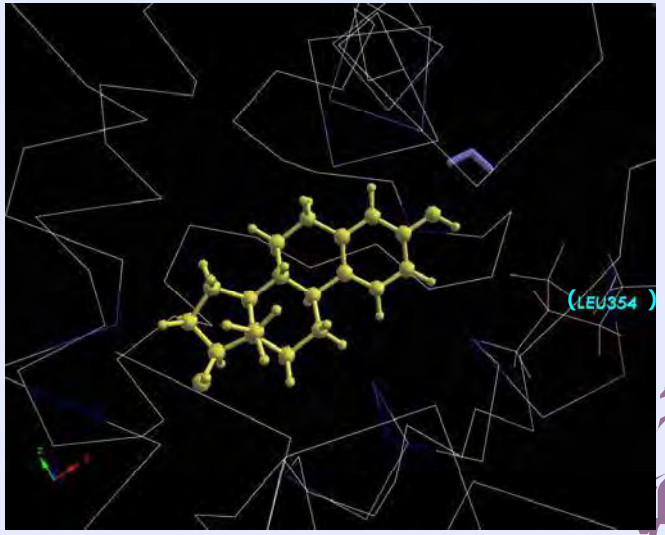
Attractive
E<-1kcal/mol





Repulsive
E>+1kcal/mol





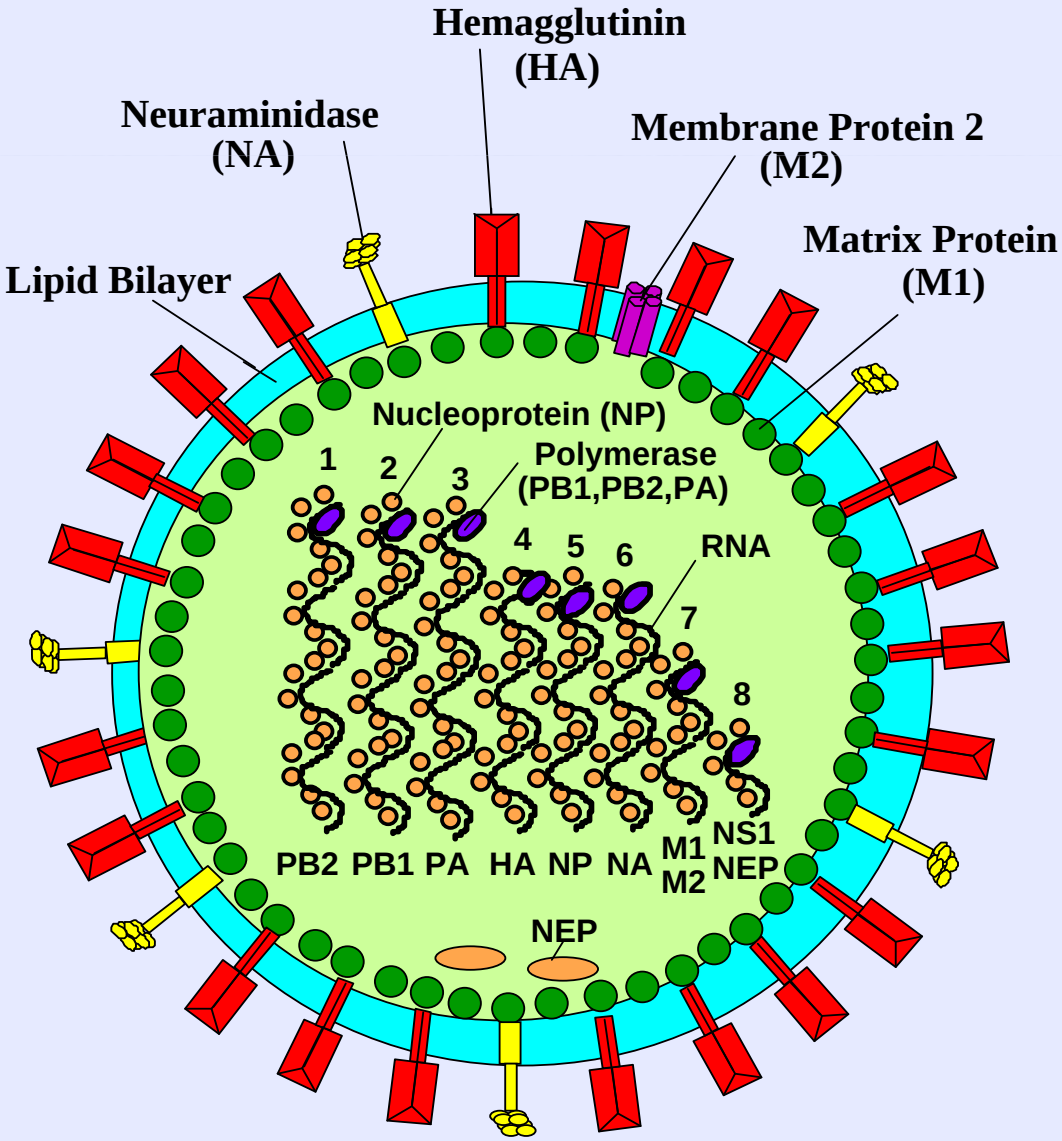
Electron correlation effect is important in ER-ligand binding.

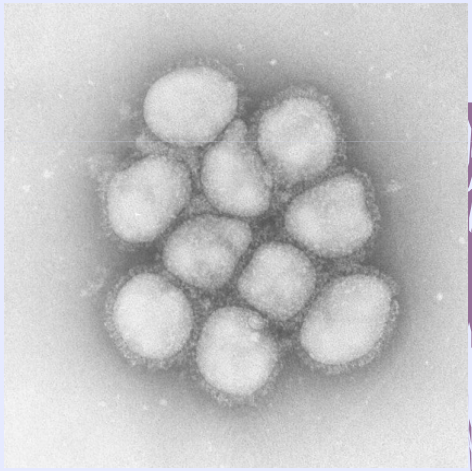
(Fukuzawa *et al*. 2006)

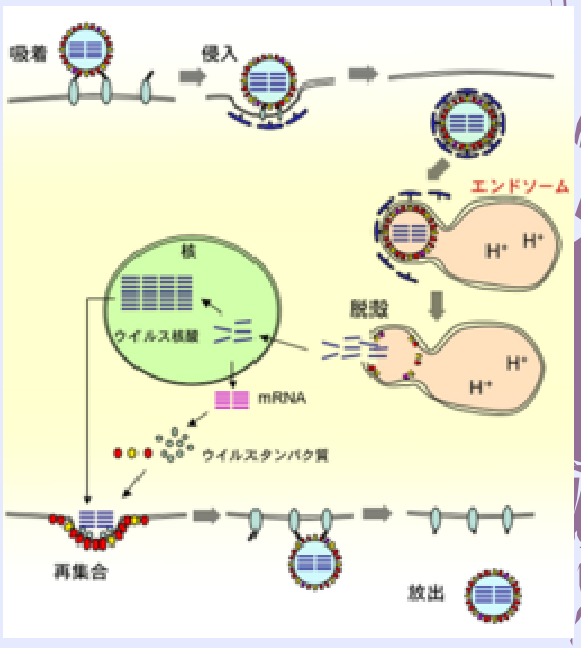
FMO can be applied to fairly large systems composed of biomolecules.

この10年ほどで様々な生体分子系に対して適用。

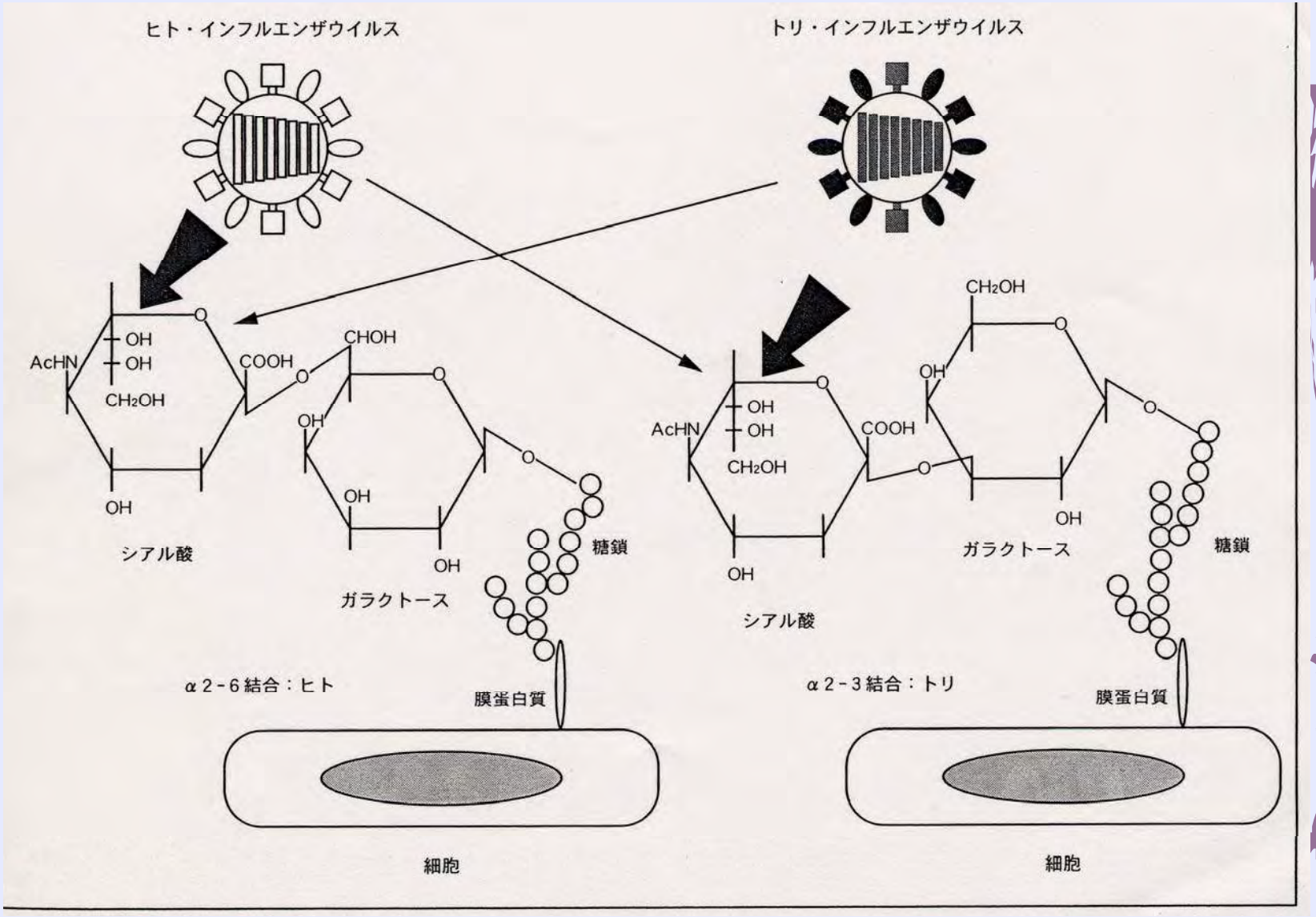
Structure of Influenza Virus







インフルエンザウイルスHAとシアル酸の結合



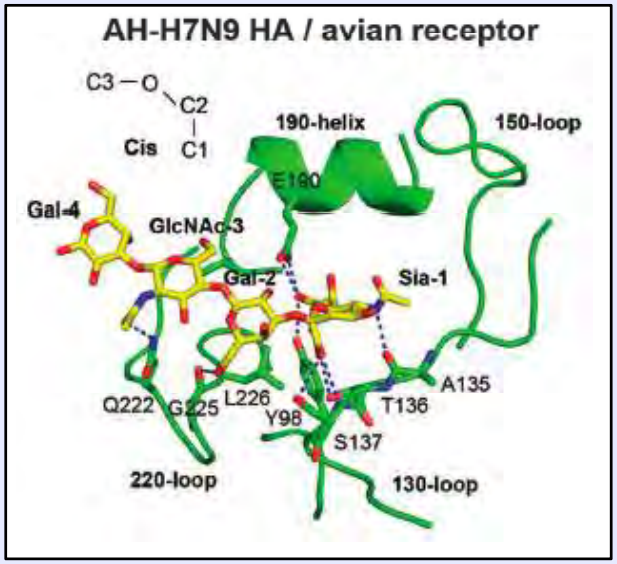
Avian to Human Infectious Transmission

FMO-MP2/6-31G calculations for the complexes of
Avian HA: H7N3 (A/Turkey/Italy/214845/2002)
Human HA: H7N9 (A/Anhui/1/2013)
Avian receptor: 3'SLN (α2-3)
Human receptor: 6'SLN (α2-6)

(S. Anzaki *et al*., J. Mol. Graph. Model. 53 (2014) pp. 48-58.)

Crystal structures are available from:
Shi et al. Science 342 (2013) 243.

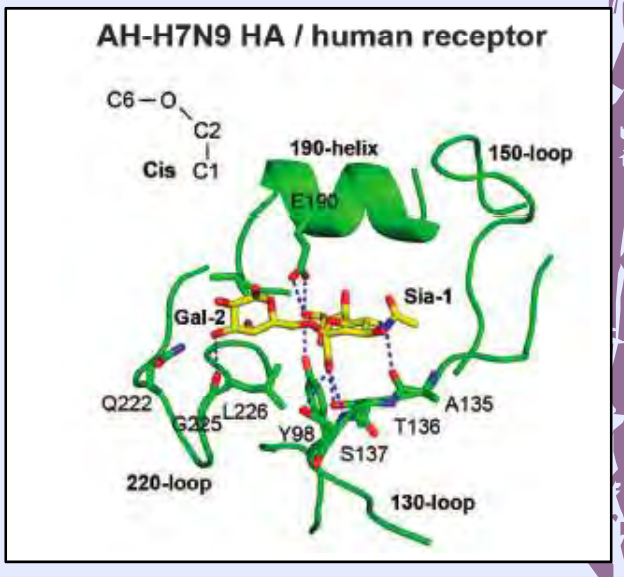
3'SLN



Binding of avian receptor and H7N9-HA

Binding of human receptor and H7N9-HA

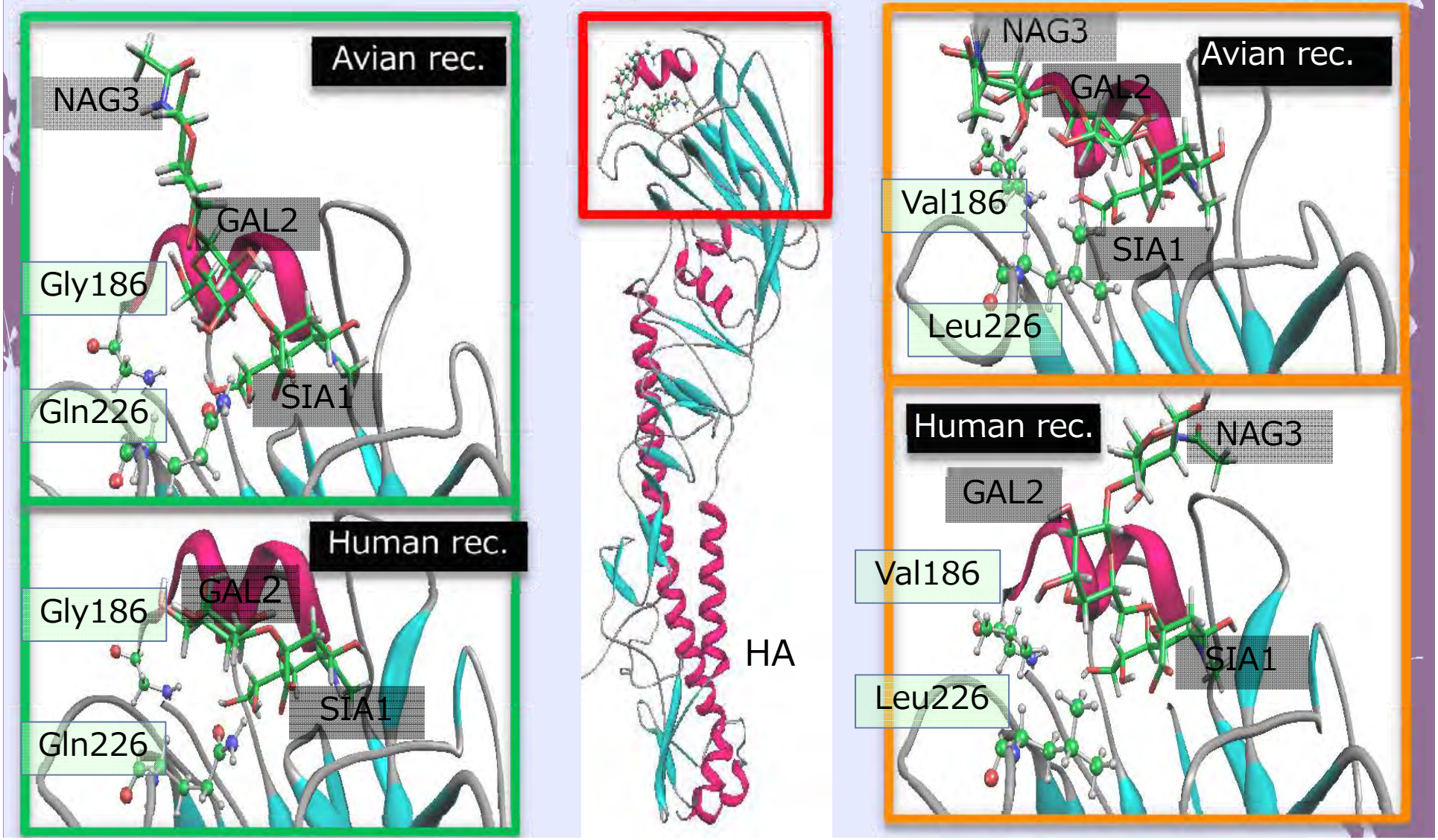
6'SLN



Avian (3'SLN) and human (6'SLN) receptors with HAs

Mutations of G186V and Q226L (hydrophilic to hydrophobic) in H7N9 play important roles for binding.

H7N3 (Avian type) **Binding** **H7N9 (Human type)**



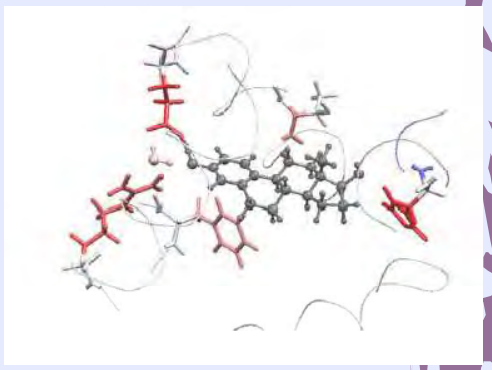
FMO provides novel methods to decompose ligand-protein interactions.

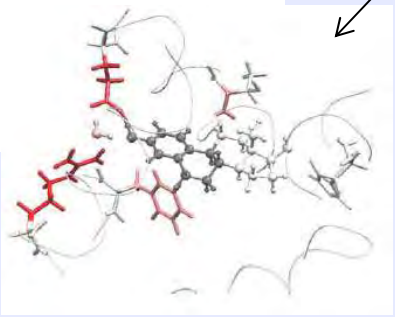
単に結合エネルギーの評価だけでなく、様々な「分解」解析が可能。

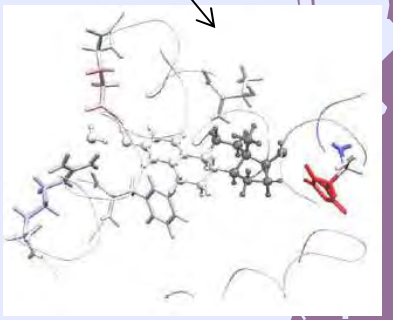
FMO4 Approximation

$$
\begin{aligned}
E^{\mathrm{FMO4}} &= \sum_{I>J>K>L} E_{IJKL} - (N-4)\sum_{I>J>K} E_{IJK} + \frac{(N-3)(N-4)}{2}\sum_{I>J} E_{IJ} \\
&\quad - \frac{(N-2)(N-3)(N-4)}{6}\sum_{I} E_I \\
&= \sum_{I>J>K>L} \{\Delta E_{IJKL} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{IL} - \Delta E_{JK} - \Delta E_{JL} - \Delta E_{KL} \\
&\quad - [\Delta E_{IJK} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{JK}] - [\Delta E_{IJL} - \Delta E_{IJ} - \Delta E_{IL} - \Delta E_{JL}] \\
&\quad - [\Delta E_{IKL} - \Delta E_{IK} - \Delta E_{IL} - \Delta E_{KL}] - [\Delta E_{JKL} - \Delta E_{JK} - \Delta E_{JL} - \Delta E_{KL}]\} \\
&\quad + \sum_{I>J>K} [\Delta E_{IJK} - \Delta E_{IJ} - \Delta E_{IK} - \Delta E_{JK}] \\
&\quad + \sum_{I>J} \Delta E_{IJ} + \sum_{I} E_I
\end{aligned}
$$

$$
\Delta E_{IJKL} \equiv E_{IJKL} - E_I - E_J - E_K - E_L
$$

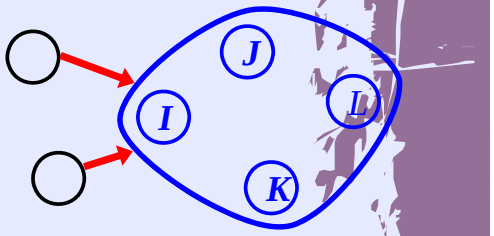






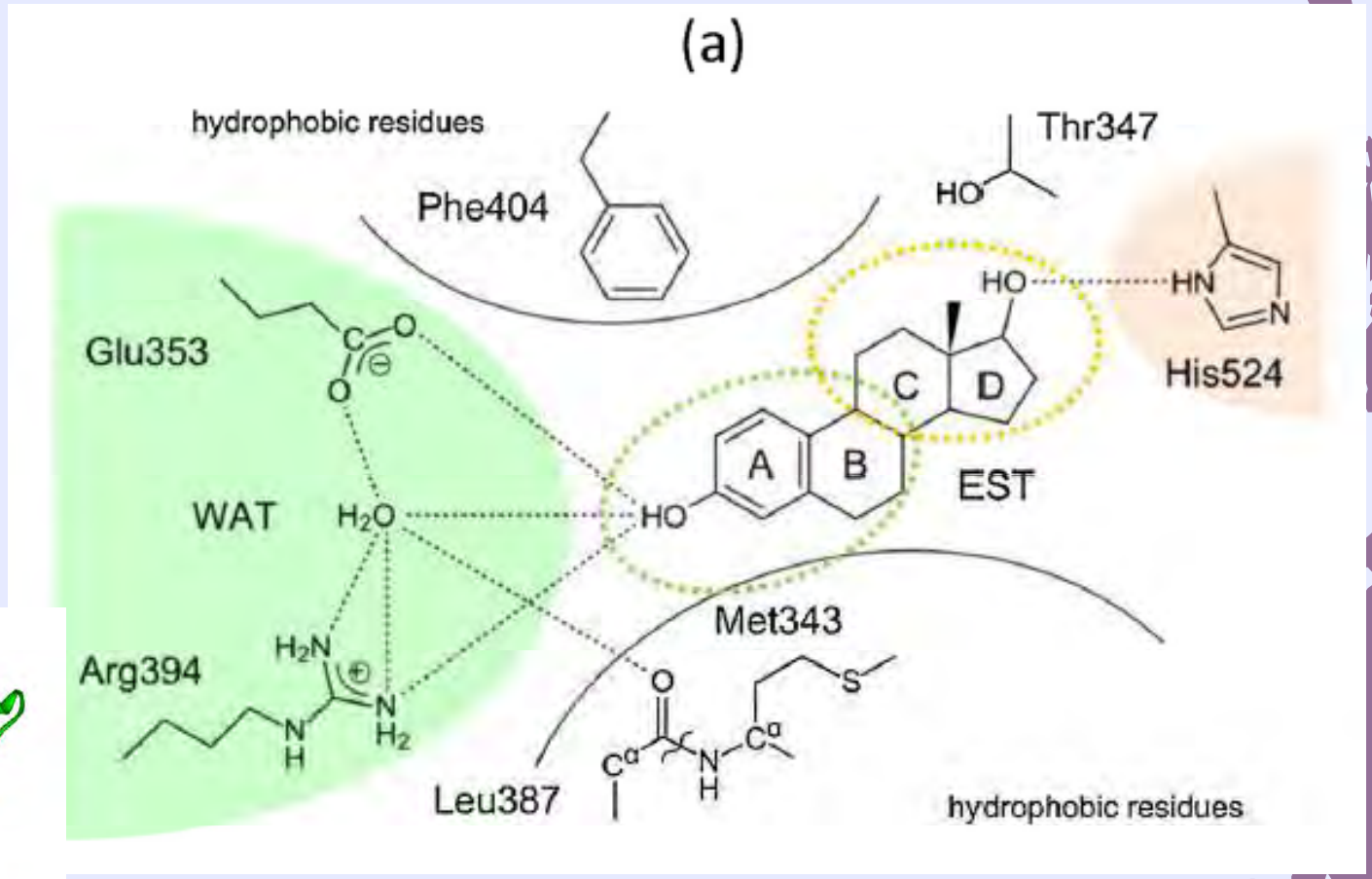
FMO calculations up to four-body terms

Improvement of energy accuracy
More refined fragment divisions



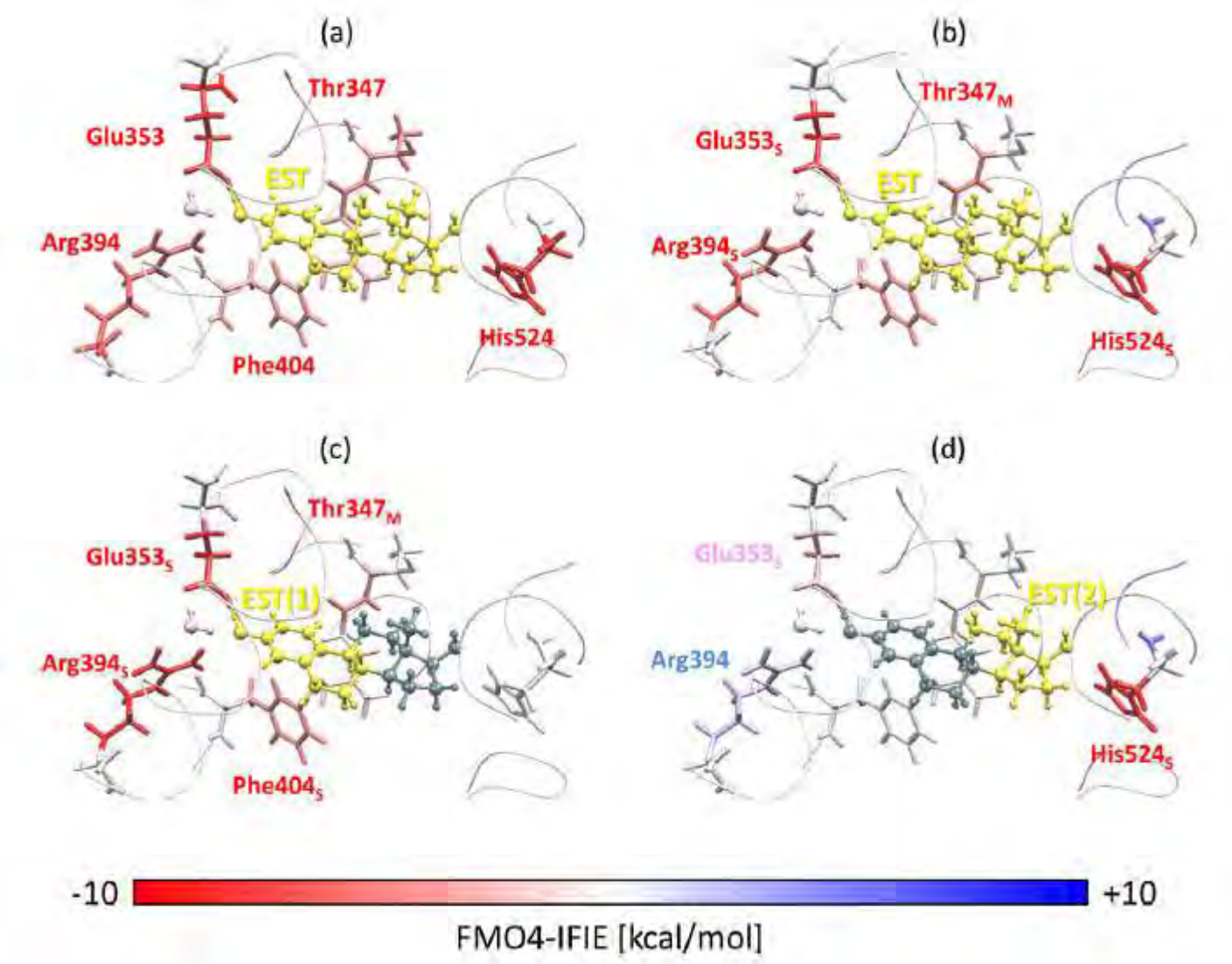
(T. Nakano *et al*., Chem. Phys. Lett. 523 (2012) 128.)

Various Interactions in ER-EST Complex



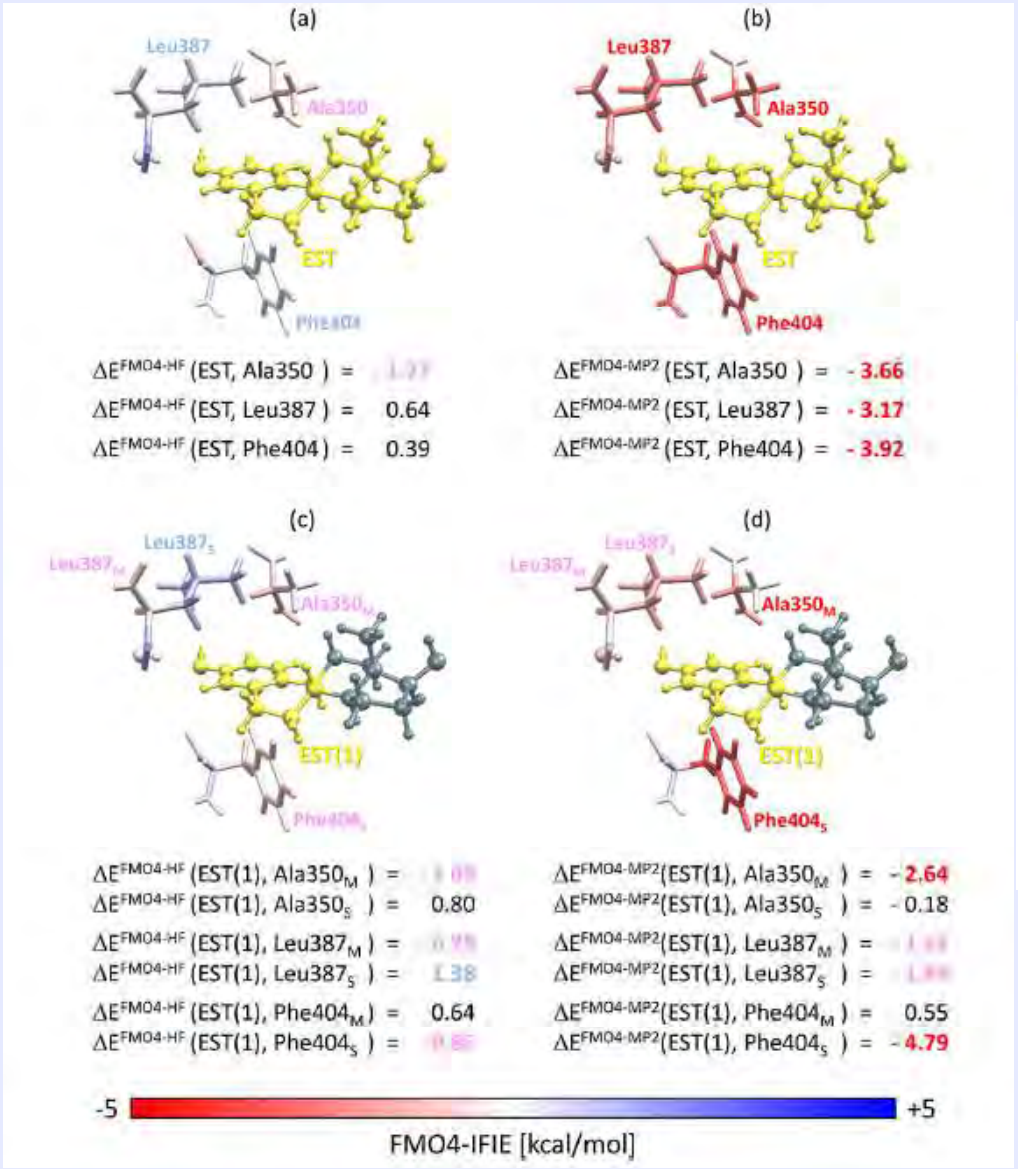
Estrogen Receptor (ER) – Estradiol (EST) Complex

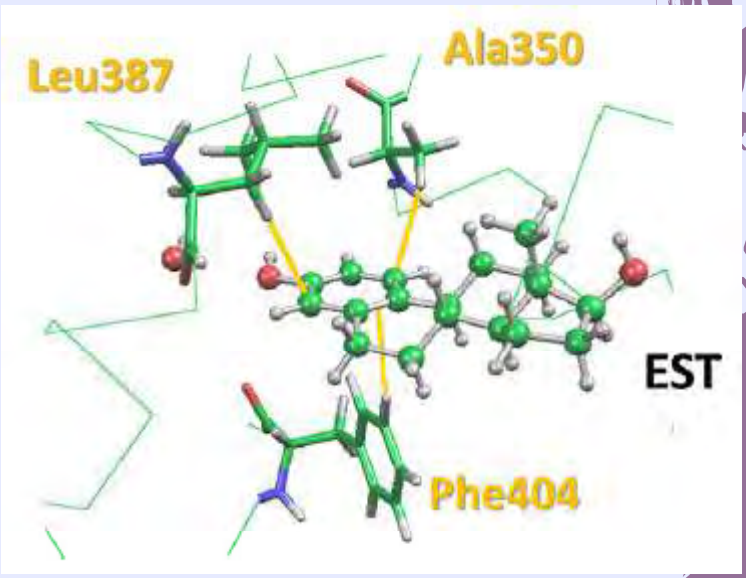
Higher Resolution Interaction Analysis



CH-π Interactions by Electron Correlations

HF MP2





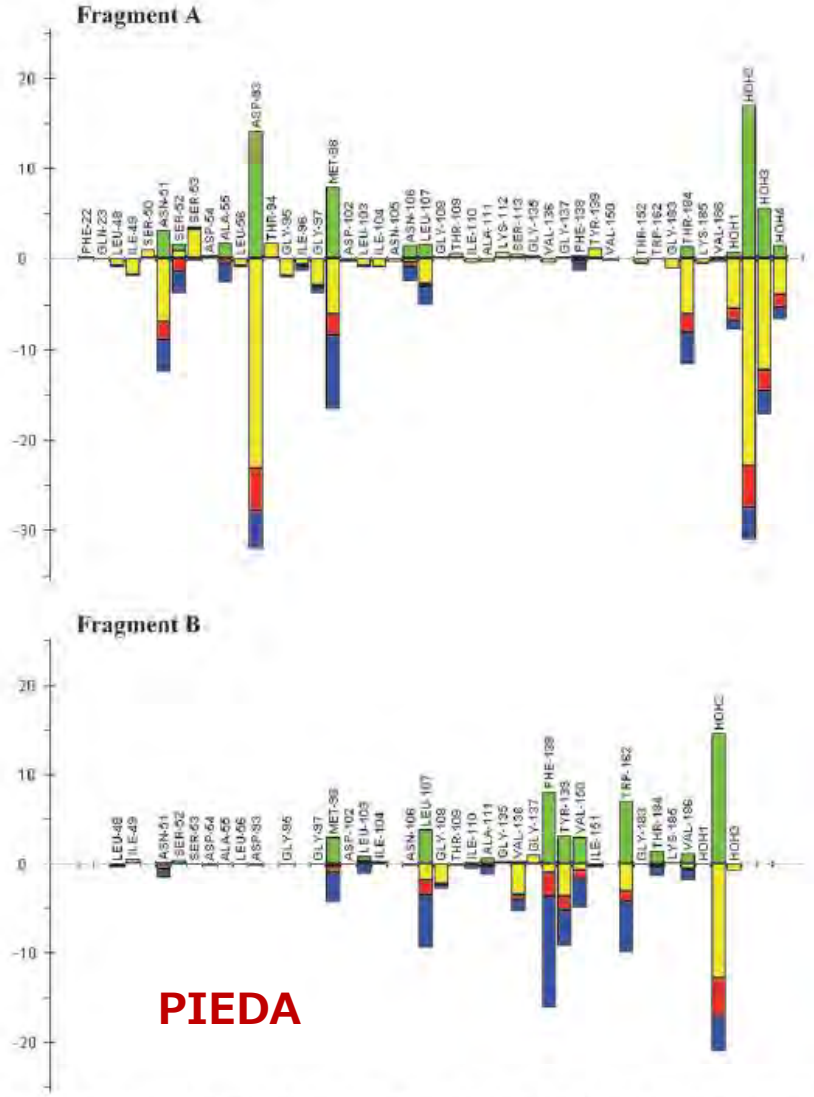
(C. Watanabe *et al*., J. Mol. Graph. Model. 41 (2013) 31.)

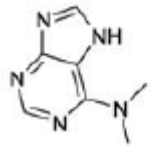
Compound Design by Fragment-Linking

Mol. Inf. 2011, 30, 298 – 306

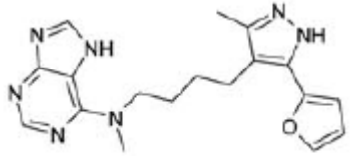
Osamu Ichihara,[a] John Barker,[a] Richard J. Law,[a] and Mark Whittaker*[a]

(Evotec) Hsp90

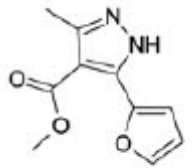




$IC_{50} = 1.5$ mM



$IC_{50} = 1.5$ μM



$IC_{50} = 1$ mM

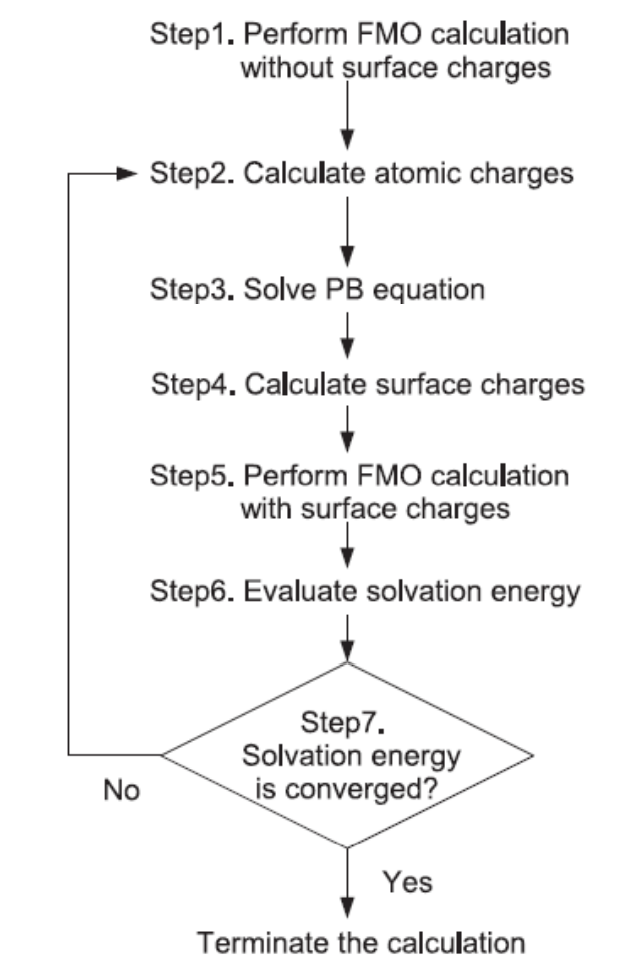
FBDD

Figure 1. FMO PIEDA diagrams for Fragment A and B in complex with Hsp90 (Table 1, Entry 5). Interaction energies are shown in kcal/mol, green (exchange repulsion), yellow (electrostatic), red (charge transfer and mixed terms), blue (dispersion). The calculations were performed using the fragment-protein complex structures, PDB ID 3HYZ and 3HYY, respectively. The diagrams were prepared using Facio software,[34] and the results are tabulated in the Supporting Information.

Inclusions of solvent effects and entropic contributions are important for quantitatively dependable predictions.

溶媒効果やエントロピー効果の記述は一点量子化学計算が主体のFMO法の弱みであるが、最近は様々な対処法が開発されている。

FMO-PB (Poisson-Boltzmann)



$$\nabla[\epsilon(\mathbf{r})\nabla\Psi(\mathbf{r})] = \rho(\mathbf{r}) - \sum_i c_i z_i q \lambda(\mathbf{r}) \exp\left[\frac{-z_i q \Psi(\mathbf{r})}{k_B T}\right],$$

$$\Delta G^{total} = \Delta G^{es} + \Delta G^{np}.$$

$$\Delta G^{es} = \langle \psi^s | H^0 | \psi^s \rangle - \langle \psi^g | H^0 | \psi^g \rangle + \frac{1}{2}[\langle \psi^s | H' | \psi^s \rangle + H''],$$

$$\Delta G^{np} = 1.09 + 0.005A,$$

Electrostatic contribution and non-polar contribution to solvation free energies (in kcal/mol) of polyalanines and Alpha-1, where FMO-PB, MO-PB and classical PB mean the electrostatic contributions by each method. FMO and MO calculations were performed with HF/6-31G* method.

	FMO-PB	MO-PB	Classical PB	Non-polar
ALA5 α-helix	186.7	188.4	187.4	2.9
ALA5 β-strand	162.2	163.2	160.2	3.4
ALA10 α-helix	268.5	269.1	238.5	4.4
ALA10 β-stand	212.2	213.7	195.1	6.2
Alpha-1	419.6	419.7	374.6	8.9

(H. Watanabe *et al.*, Chem. Phys. Lett. 500 (2010) 116.)

Comparison among Various Docking Scores

MP2/6-31G* by GAMESS/FMO

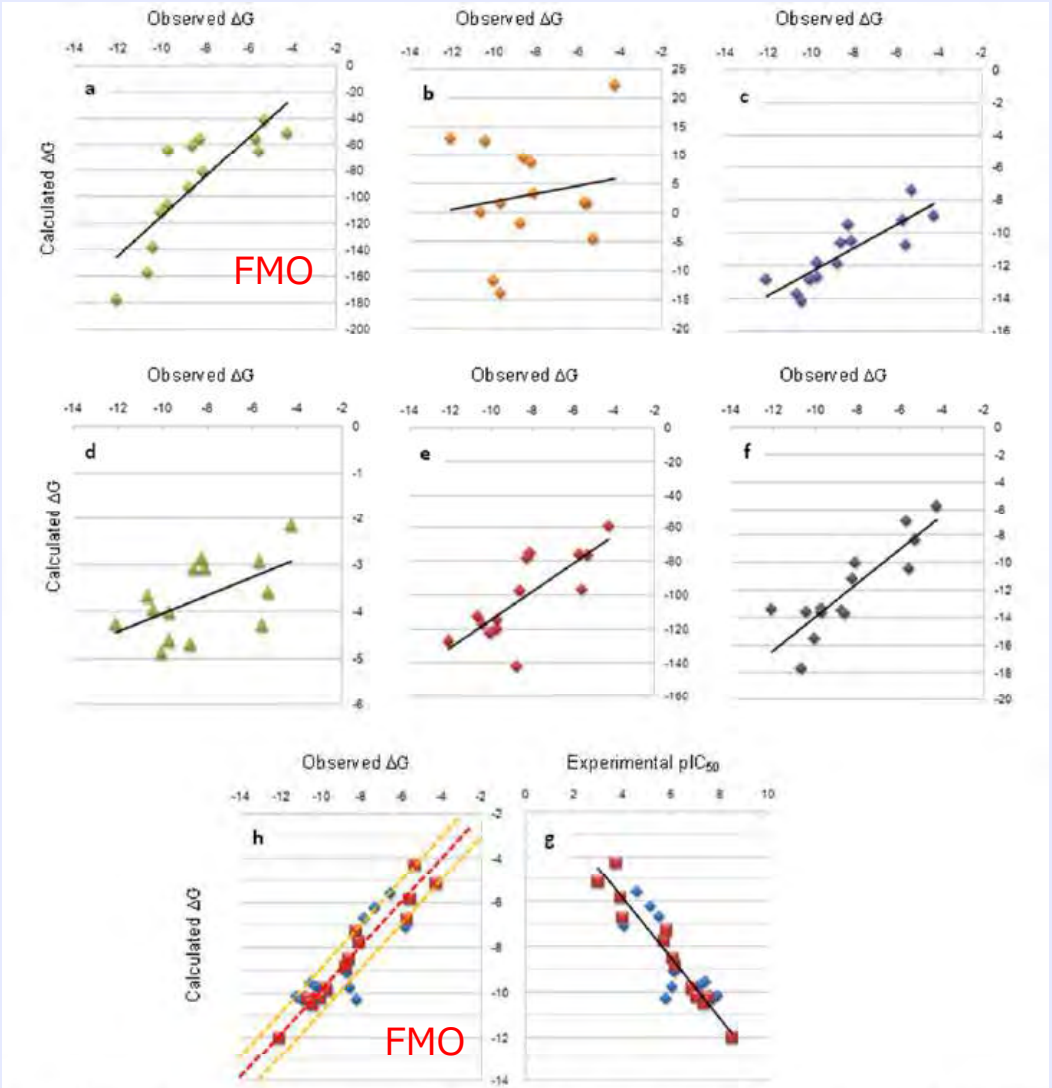


Figure 3 Calculated versus observed free energy of binding for 14 CDK2 inhibitors assessed using seven different methods. Methods include a) FMO (green diamonds), $r^2 = 0.68$; b) GBVI (orange diamonds), $r^2 = 0.03$; c) London dG (purple diamonds), $r^2 = 0.73$; d) Affinity dG (green triangles), $r^2 = 0.31$; e) Alpha HB (red diamonds), $r^2 = 0.61$; f) ASE (black diamonds) $r^2 = 0.75$; and g) QM-based scoring function (red squares) together with the 14 compound test set (blue diamonds) for the QM-scoring function, and h) the calculated versus the experimental pIC_{50} values for the QM-scoring function, $r^2 = 0.94$. For graphs a-f, and g, the line of best fit is shown in black. Graph h shows the line of best fit as a dotted red line and the two dotted yellow lines correspond to 1 log unit boundaries.

CDK2 (cyclin-dependent kinase 2) inhibitors

MM score functions on MOE in comparison with experiments

Binding free energy

$$\Delta G_{bind} = \Delta G_{bind}^{gas} + \Delta G_{solv}^{complex} - \Delta G_{solv}^{receptor} - \Delta G_{solv}^{ligand}$$

$$\Delta G_{bind}^{gas} = \Delta H_{bind}^{gas} - T\Delta S_{bind}^{gas}$$

$$\Delta H_{bind}^{gas} = ES + EX + DI + CT + mix$$

$$T\Delta S_{bind}^{gas} = num(\ rot\_bonds\ )$$

$$\Delta G_{solv} = \Delta G_{psolv} + \Delta G_{npsolv}$$ PBSA

$$\Delta G_{npsolv} = \gamma SASA + b$$

(Mazanetz *et al.*, J. Cheminf. 3 (2011) 2.)

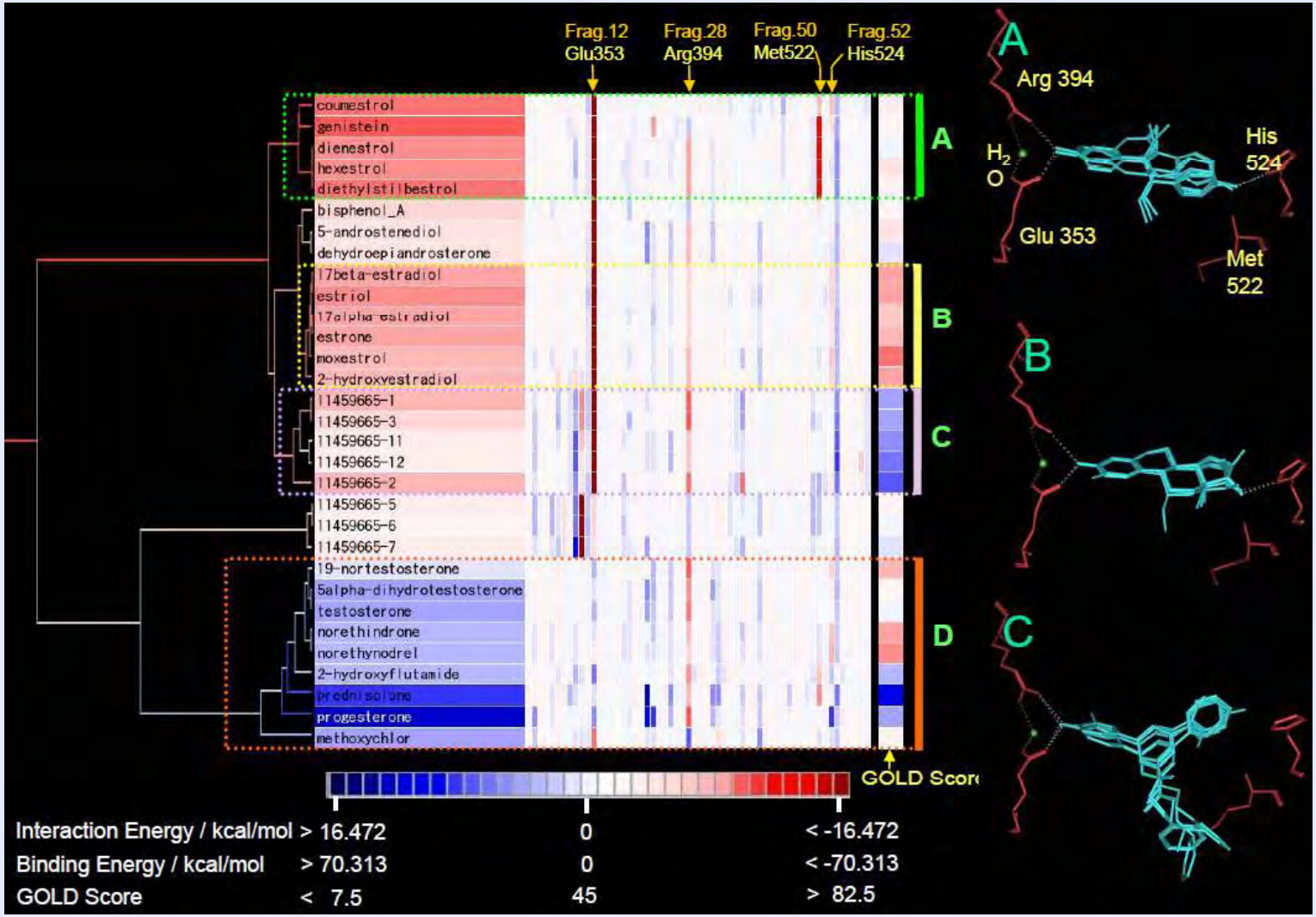
Detailed descriptions of ligand-receptor interactions based on FMO provide a new perspective for ligand screening.

FMO計算のアウトプットであるIFIE（Inter-Fragment Interaction Energy）は情報科学的な解析と組み合わせることで様々な新しい知見を与えてくれる。

VISCANA: Clustering due to IFIE patterns

$$d_{IJ} = \sum_{K=1}^{N} \left(\Delta\tilde{E}_{IK} - \Delta\tilde{E}_{JK} \right)^2$$

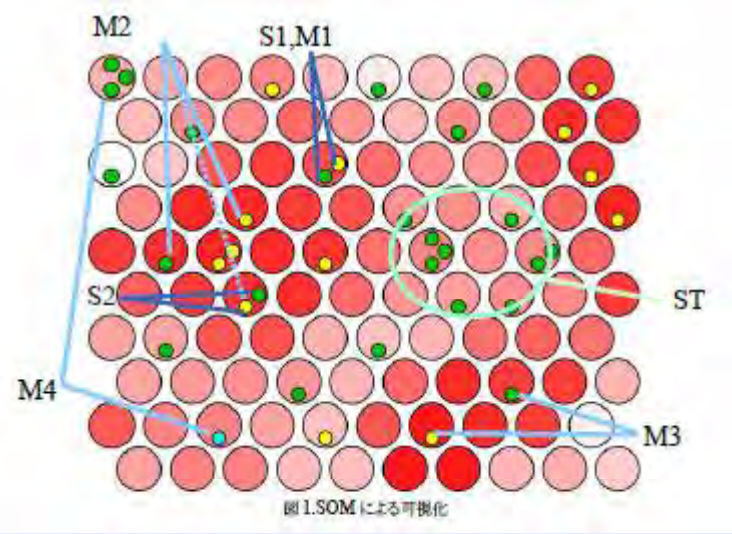
Distance matrix



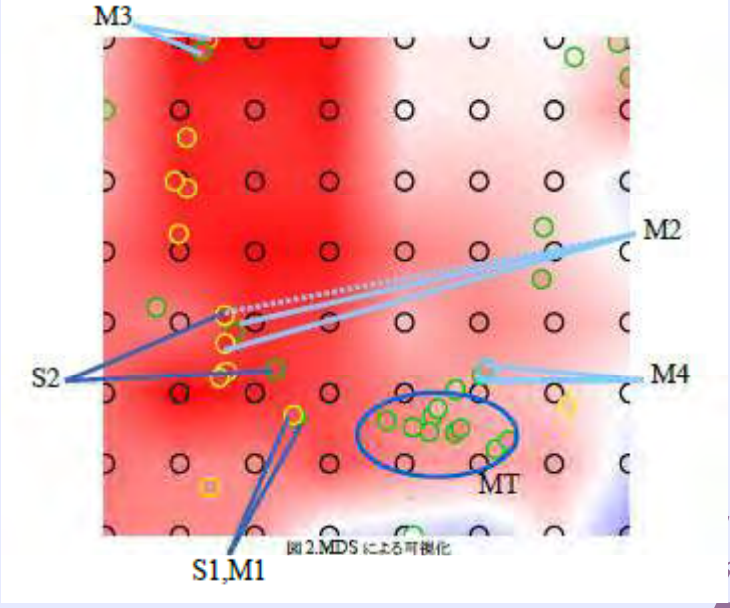
Extended clustering methods

- Multi-dimensional visualizations with SOM（Self-Organizing Map） and MDS（Multi-Dimensional Scaling）
- Inclusion of molecular information due to MDL MACCS keys in addition to FMO-IFIE

(Kurauchi *et al.* 2015)



SOM



MDS

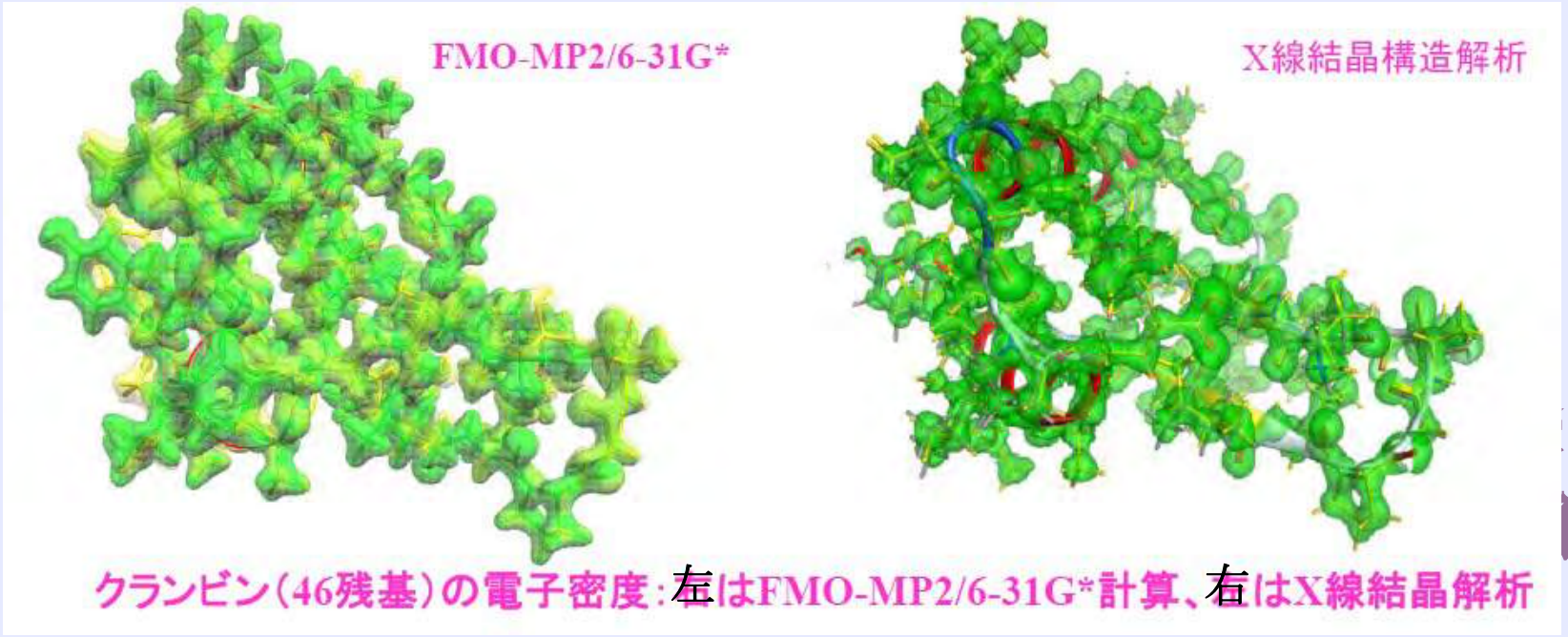
➡ Reduction of false positives and false negatives; Agonist vs. Antagonist

Accurate “PDB” structures are still essential as a starting point for reliable *in silico* screening.

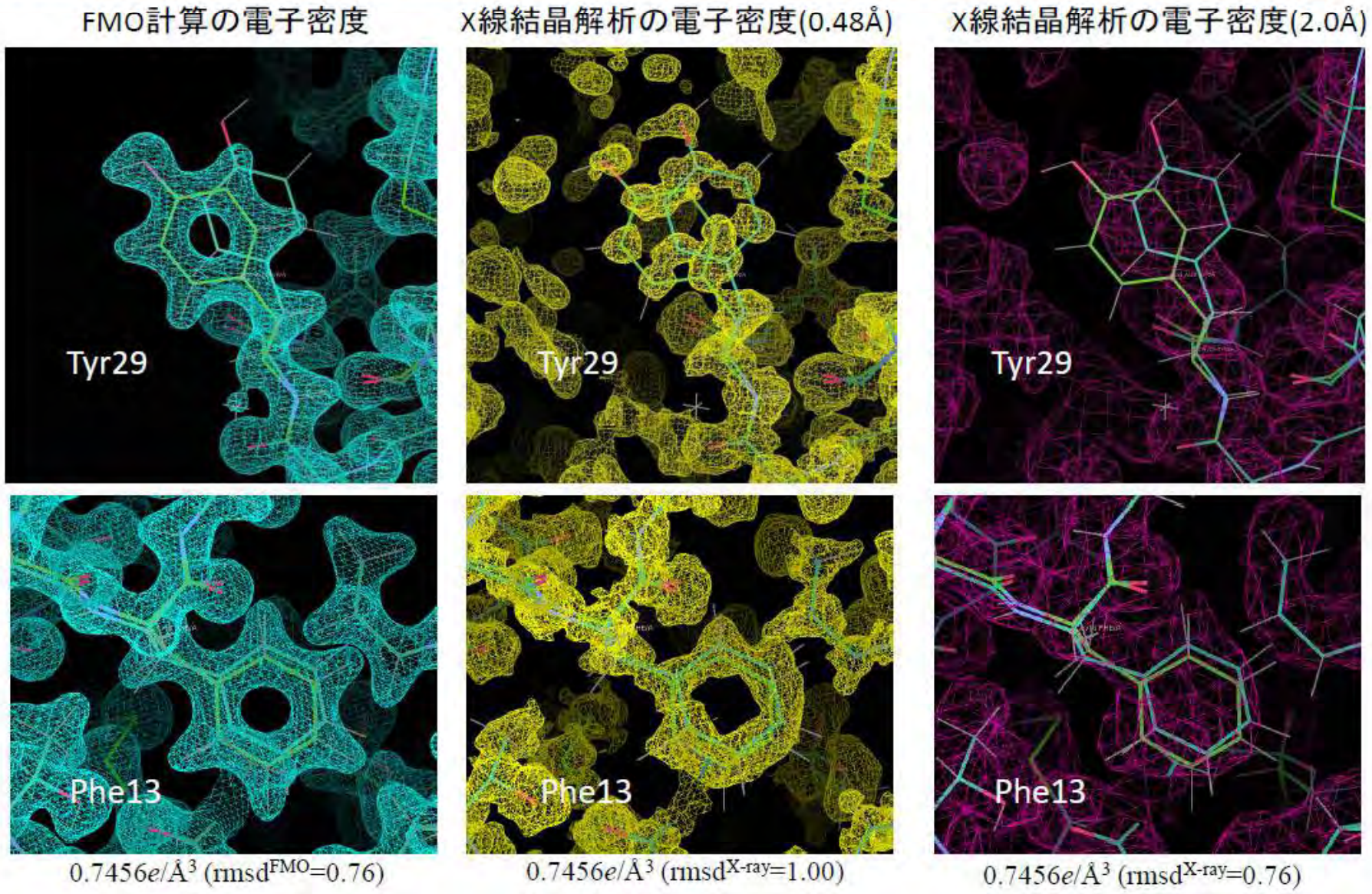
X線結晶構造の精密化そのものにFMO法を使えないか？

Refinement of Protein-Ligand Complex Structures

- Self-consistent determination of electron density and atomic coordinates in X-ray crystallography
  - Hydrogen atoms
  - Conformers



FMO計算の電子密度、解像度の異なるX線結晶解析の電子密度(0.48, 2.0Å)の比較



Ab initio quantum-mechanical dynamics would be realized in the future in the framework of FMO.

Ab initio MDの生体分子系への適用（酵素反応、光反応など）はFMO法の重要な将来課題。

Visualization of Molecular Information

- Visualization of “real-time” chemical reactions in biomolecular systems
- FMO-MD: *ab initio* dynamics
- FMO-LCMO (linear combination of MOs): efficient description of frontier orbitals (HOMO, LUMO, etc.)

FMO-LCMO

Chemical Physics Letters 476 (2009) 104–108

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Molecular orbital calculation of biomolecules with fragment molecular orbitals

Shinji Tsuneyuki [a,*], Tomoki Kobori [a], Kazuto Akagi [b], Keitaro Sodeyama [a], Kiyoyuki Terakura [c], Hidetoshi Fukuyama [d]

[a] Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[b] WPI-AIMR, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
[c] Research Center for Integrated Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
[d] Department of Applied Physics, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

ARTICLE INFO

Article history:
Received 14 March 2009
In final form 29 May 2009
Available online 6 June 2009

ABSTRACT

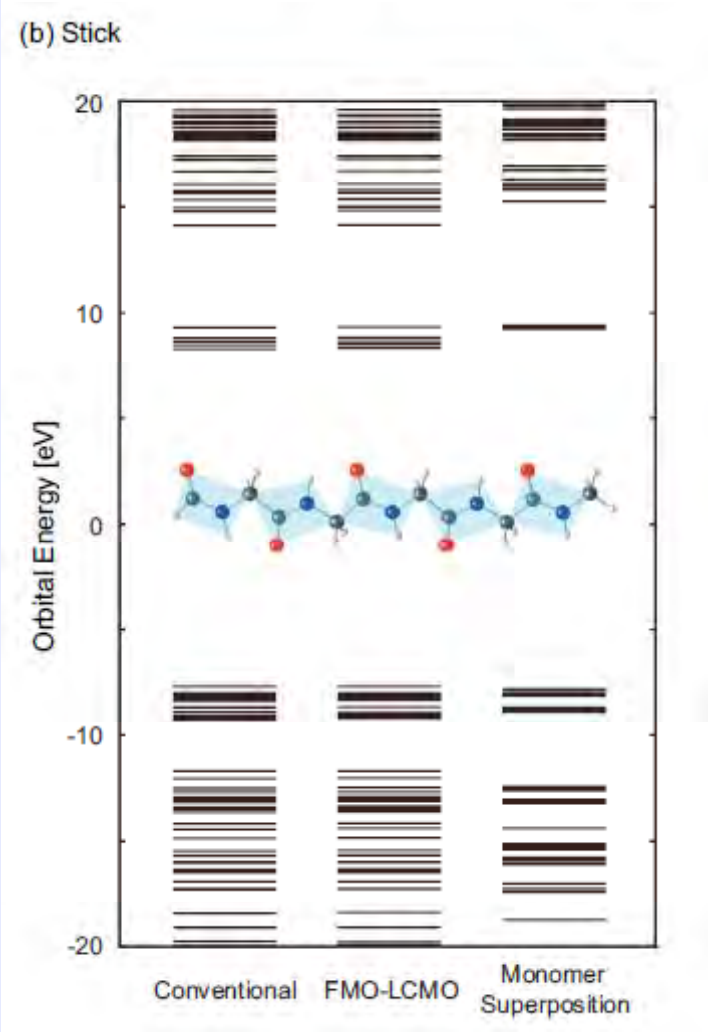
A new method of calculating molecular orbitals and orbital energy levels of biomolecules is proposed based on the fragment molecular orbital method (FMO), in which a huge biomolecule is first divided into 'fragments' for feasible calculation. In the present method, called FMO–LCMO, one-electron Hamiltonian of the whole molecule is derived from the output of FMO with little computational cost. The method gives fast and accurate description of the molecular orbitals of the whole system by a linear combination of molecular orbitals (LCMO) of the fragments and provides a good starting point for understanding the electronic structure of biomolecules.

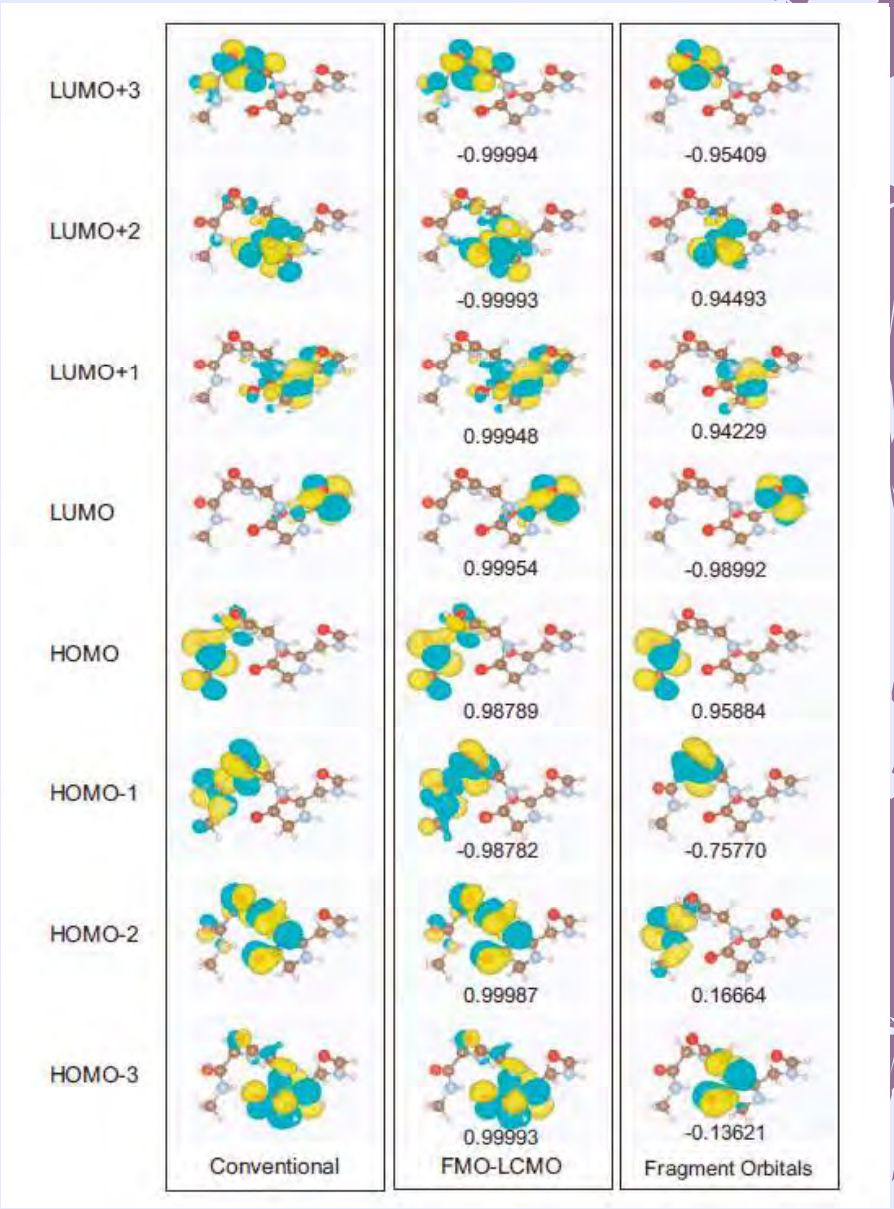
$$(H_I)_{pq} = \varepsilon_p \delta_{pq},$$
$$(H_{IJ})_{ij} = \varepsilon_i \delta_{ij},$$

$$(H_{IJ})_{Kp,K'q} = \sum_i \sum_j \langle \phi_p^K | \phi_i^{IJ} \rangle (H_{IJ})_{ij} \langle \phi_j^{IJ} | \phi_q^{K'} \rangle$$
$$= \sum_i \varepsilon_i \langle \phi_p^K | \phi_i^{IJ} \rangle \langle \phi_i^{IJ} | \phi_q^{K'} \rangle .$$

$$(H_{\text{total}})_{Ip,Jq} = (H_{IJ})_{Ip,Jq},$$

$$(H_{\text{total}})_{Ip,Iq} = \sum_{J \neq I} (H_{IJ})_{Ip,Iq} - (N-2)(H_I)_{pq},$$



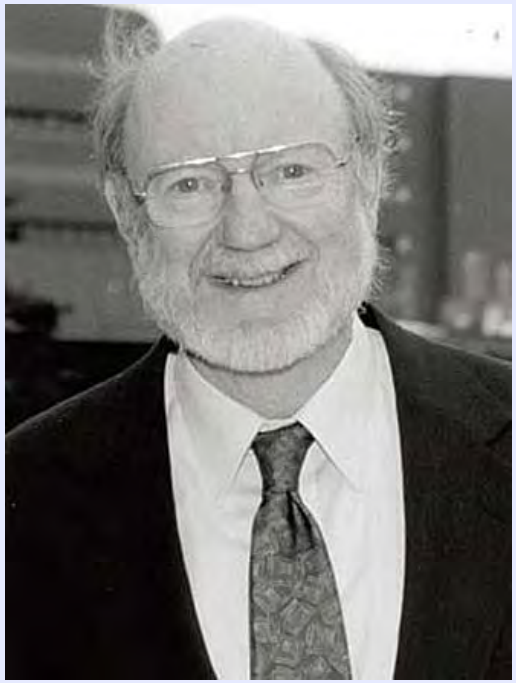


「アカデミア創薬」の意義
(特に、インシリコ創薬に関して)

- 製薬企業の現場では、スクリーニングなどのツール開発にまでは時間やコストを割けない。
- 一方、大学や国研は時間や手間暇のかかる基礎研究にその存在意義を見い出せる。
- 個々の利潤追求を超えて共有できる共通基盤技術(ソフトウェアなど)の開発やその基となる基礎的課題の解決・新規シーズの発見等が「アカデミア創薬」に期待される。

The Nobel Prize in Physiology or Medicine 2015

Satoshi Omura
エバーメクチン
（写真はノーベル財団）



William C. Campbell
イベルメクチン
（メクチザン；Merck）

for their discoveries concerning a novel therapy against infections caused by roundworm parasites
（回虫が原因となる感染症の有効成分の発見と治療薬の開発）

FMO創薬コンソーシアムとは

FMO Drug Design Consortium (FMODD)

FMO法に基づくインシリコ創薬手法を、実用的な技術として発展させるための集まり

代表：福澤薫（日大）
副代表：田中成典（神戸大）
本間光貴（理研）

第1期活動期間：

- 2014年11月1日～2017年3月31日

活動内容：

- 年に数回のミーティング
- FMO法利用ノウハウの共有
- 新規研究テーマの検討
- ユーザーインターフェイスの開発
- HPCI（High Performance Computing Infrastructure）の利用
- データベースの共有

参加について：

- 参加費：無料
- 参加資格：製薬企業、情報系企業、大学、研究所等、創薬研究に携わる研究者のどなたでも参加できます。

FMO創薬コンソーシアムと「京」コンピュータの利用

「HPCIを活用したFMO創薬プラットフォームの構築」(課題番号hp150160)

産学官連携で実用的なFMO創薬技術を開発する

FMO創薬コンソーシアム

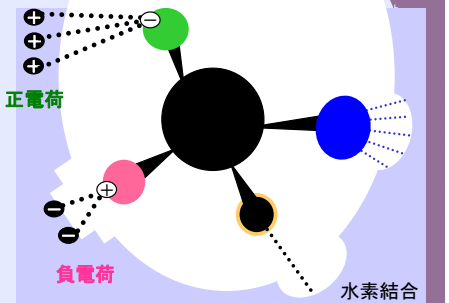
FMO consortium

大学・研究所　7機関
製薬企業　12社
IT企業　1社
(うちHPCI参加は7機関+10社)

問題設定、計算および解析の実施

「京」などのHPCI設備
超並列FMO計算

医薬品化合物とタンパク質の
相互作用デザイン



計算プログラム
の提供

成果の
フィードバック

評価指導

【ABINIT-MP開発機関】

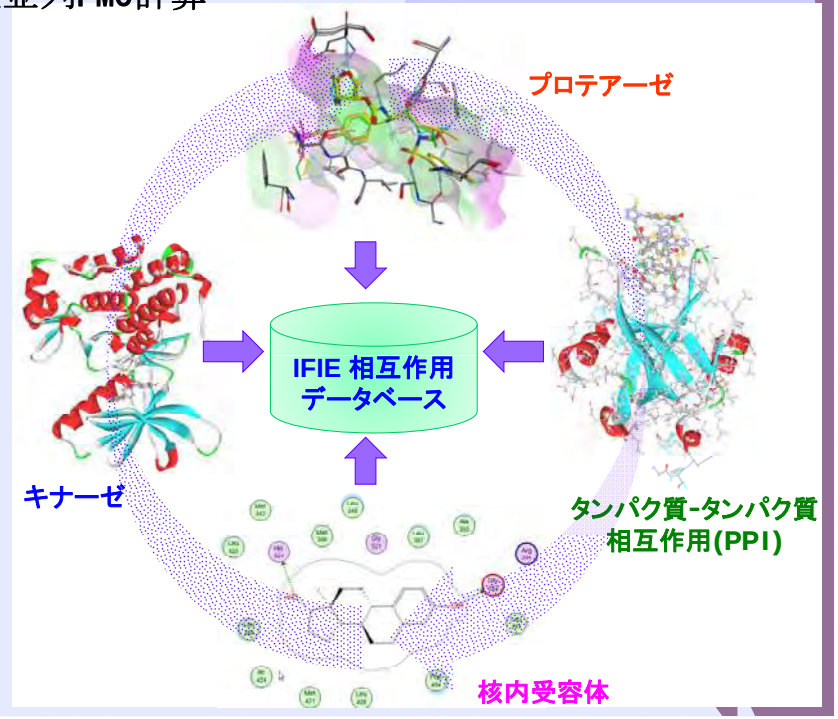
東京大学生産技術研究所
東京大学・立教大学

ABINIT-MPの開発・公開
HPCI利用に関するアドバイス

【アドバイス機関】

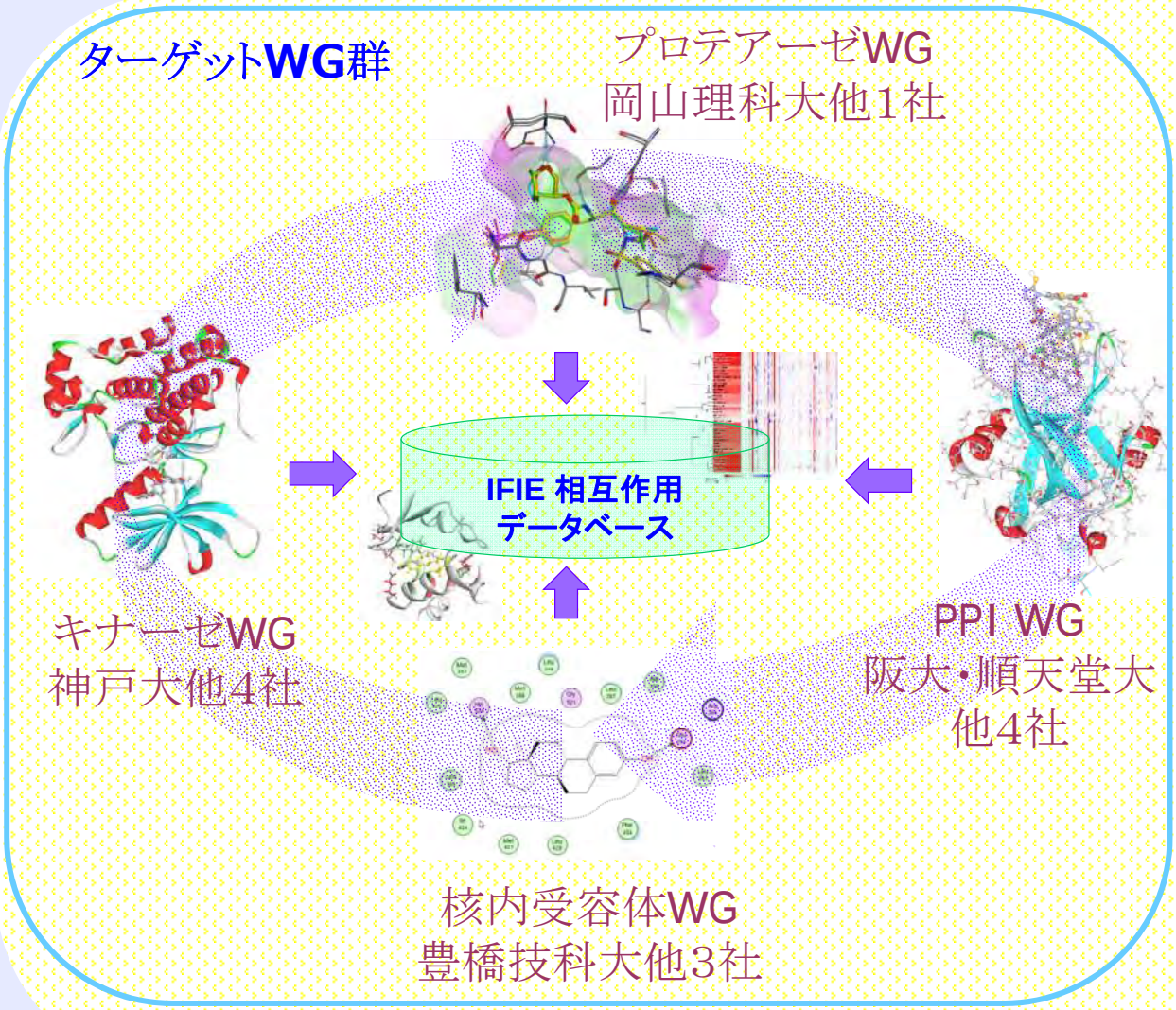
国立医薬品食品衛生研究所
大阪大学・熊本大学

化合物・安全性評価、化合物ライブラリー
量子化学計算データベース



- HPCI活用による大量データの生成と解析(1年間で複合体500個を計算予定)
- 幅広い創薬ターゲットに対するデータの収集・解析
  ⇒ 汎用的な技術を作り上げる
- FMOに基づく**新規スクリーニング手法**の開発
- **IFIEデータベース**の作成

FMODD: WGによる活動



先行事例WG
日大・神戸大・理研
計算・解析手法の試行錯誤

開発WG
理研・熊本大・日大他1社
研究基盤の構築

アクティブメンバー：
50名以上

- 幅広い創薬ターゲットへの適用
- 大量データの生成と解析

→ 汎用的な技術を作り上げる