



ナノバイオ界面での相互作用解析のための計算手法の開発 ～インプラントやバイオセンサーの設計に有用なツール～

「HPCI戦略プログラム」分野4 次世代ものづくり;

FMOグループ

立教大学 望月祐志*

日本大学松戸歯学部 福澤薫

みずほ情報総研 加藤幸一郎

FMO法とプログラム

◇分割-統合系のアプローチの一つ



北浦らが16年前に2体展開で提案

- ⇒ 全量子論的な扱い (分極や電荷移動の記述)
- ⇒ フラグメントとその対で系のエネルギーを評価 (FMO2)
- ⇒ 環境静電ポテンシャル (ESP)、直接結合切断 (BDA)
- ⇒ 階層的な並列処理 (フラグメントリスト&内部処理)

フラグメント間の相互作用エネルギー (IFIE)

- ⇒ 計算対象の解析ツール

◇プログラム

機能拡張・翻案

- ⇒ Fedorov・北浦・Gordon-G: 米GAMESS、関野ら: NWCHEM

フルスクラッチ・国産

- ⇒ 望月ら: ABINIT-MP (東大生産研を拠点に継続的開発)
- ⇒ 石川: PAICS、稲富ら: OpenFMO

$$E_{(FMO2)} = \sum_{I>J} E_{IJ} - (N_f - 2) \sum_I E_I = \sum_{I>J} \Delta E_{IJ} + \sum_I E'_I$$

ABINIT-MPの主な機能 (バイナリ公開中のVer.7)

・エネルギー

→ **MP2**: FMO2, FMO3, **FMO4**

→ 近似化2電子積分 (CD)

・エネルギー微分

→ FMO2-MP2: 部分構造最適化

→ 分子動力学 (MD)

・解析機能

→ **IFIE/SCIFIE**, 電子密度生成

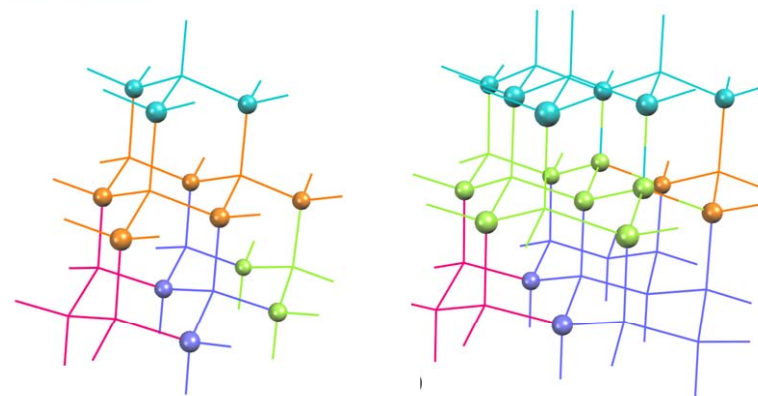
→ 軌道相互作用 (CAFI, FILM)

→ **専用GUI**: BioStation Viewer

・計算機環境

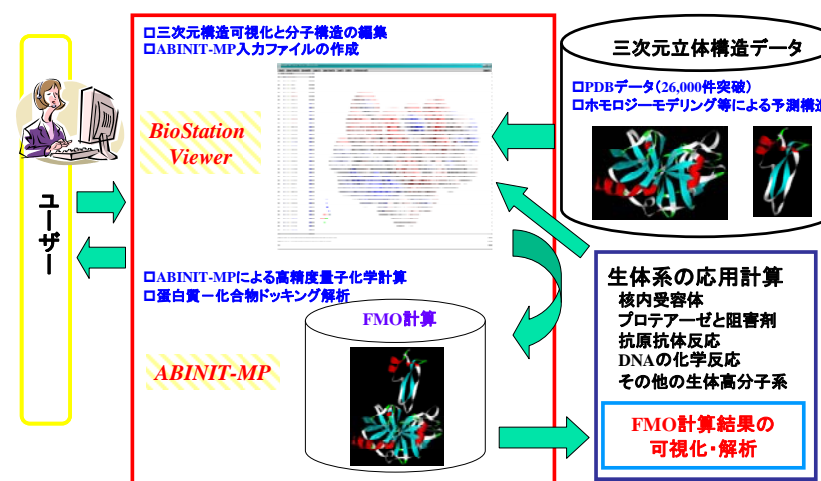
→ MPI, OpenMP/MPI混成

→ 小型サーバ~**「京」**等のスパコン



【FMO4法では固体・結晶も計算可能】

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



【BioStation Viewer】

FMO法の応用計算の方向性

◇創薬系

- ・ 標的タンパク質の**アミノ酸**と**薬品分子**の相互作用の解析
→ 普及期に突入、今後は構造最適化がポイント
- ・ 古典分子動力学シミュレーションに対する相補的な手段

◇ナノバイオ工学系

- ・ 高品質のインプラント、新規デバイスの開発（ものづくり）
- ・ 特定の固体表面を認識する**アミノ酸並び**がカギ
- ・ 実験では遺伝子工学的なアプローチでスクリーニング
- ・ 理論計算による解析は未だに発展途上（**フロンティア領域**）

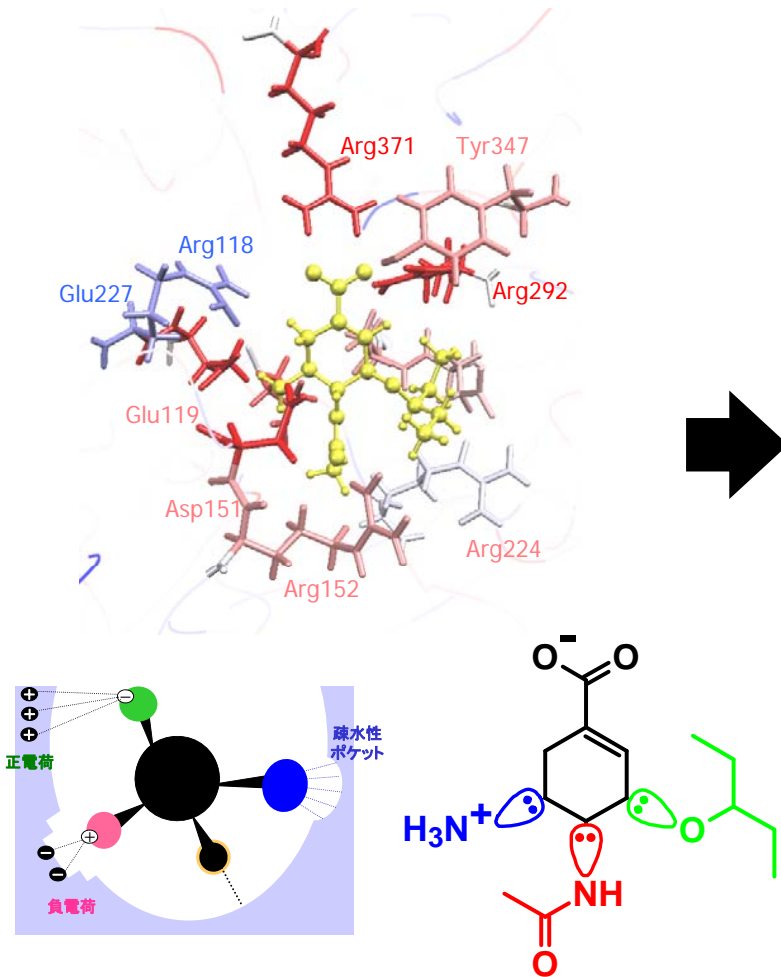
◇化学工学系

- ・ エンジニアリングポリマーや膜、粗視化シミュレーションと連携
→ ポスト「京」/FS2020プロジェクトに関連（**サブ課題⑥E**）
- ・ 分子性固体などの機能性材料の扱い

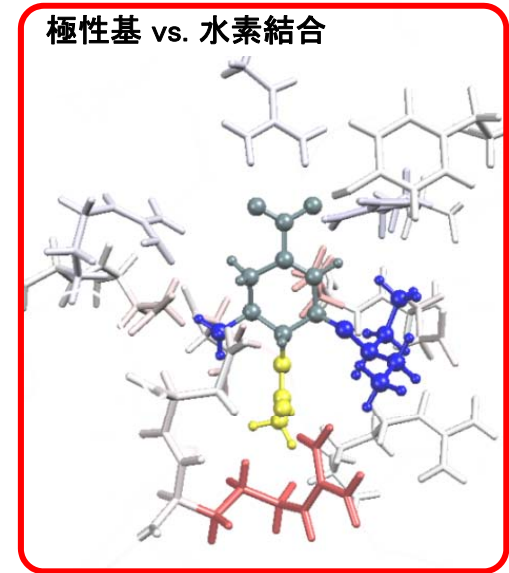
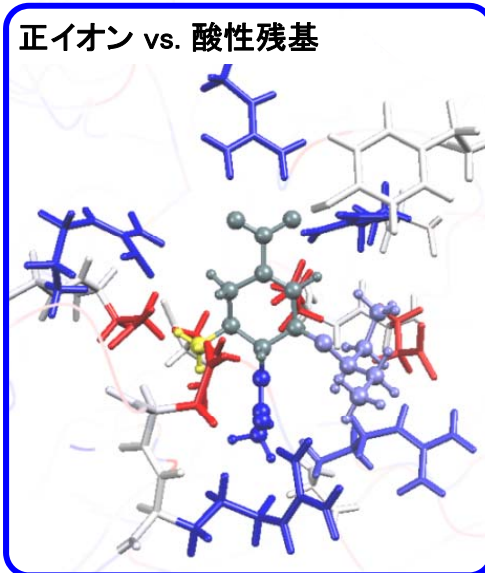
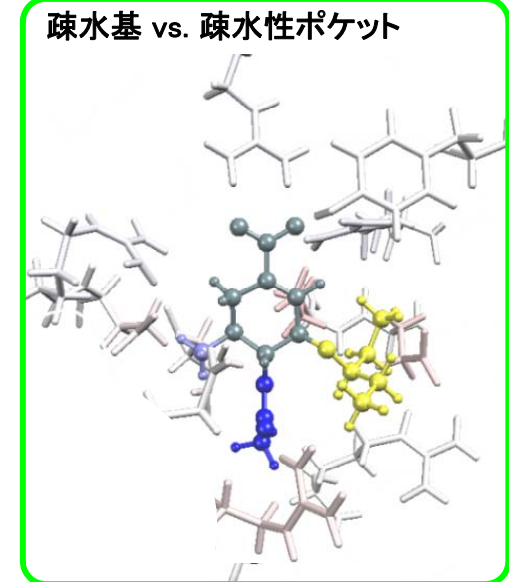
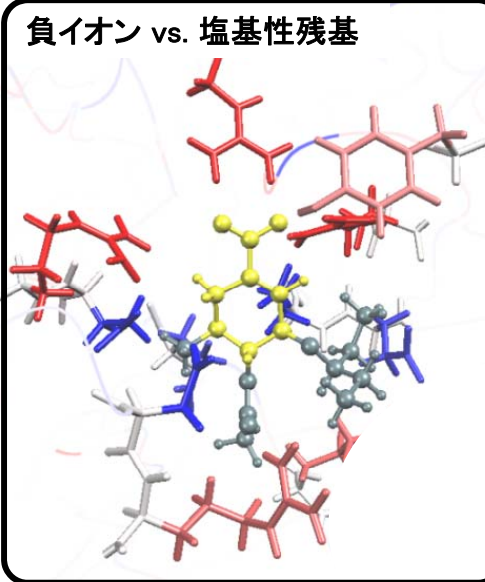
FMO4によるIFIE解析例 (インフルエンザNA-タミフル)

FMO4法: 官能基毎の局所的相互作用を描写

FMO2法: リガンド全体が均一化された描像



リガンドを官能基で分割可能



ABINIT-MPによる「京」での1.2万ノード計算例

インフルエンザノイラミニダーゼ (386残基)

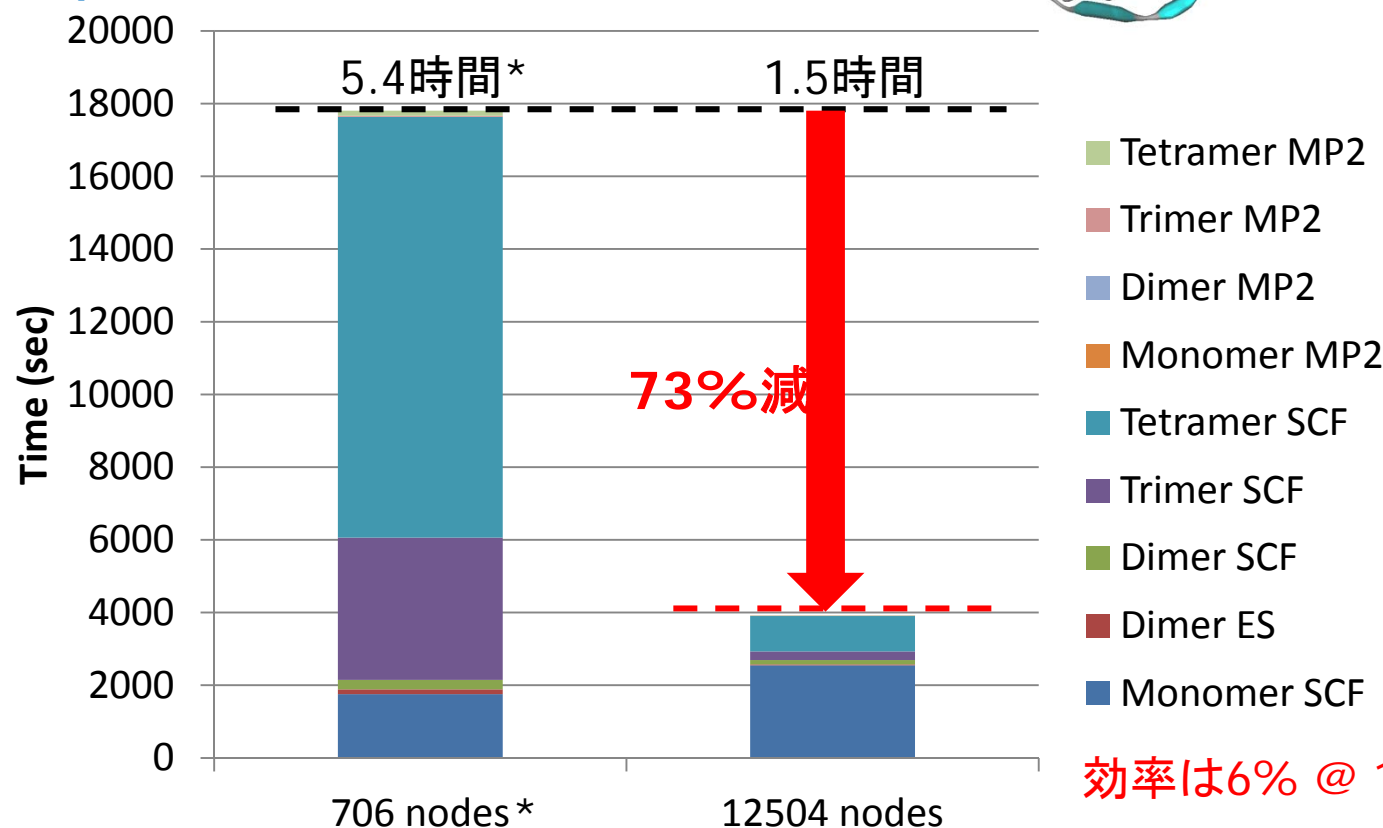
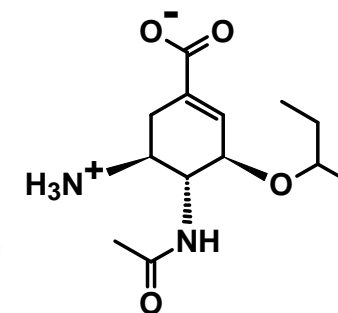
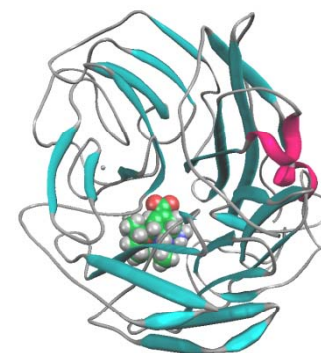
FMO4-MP2/6-31G(1C-CD)

フラグメント数 706

(主鎖/側鎖分割、リガンド分割)

OpenMP版 THREADS=8・Kfast+部分的O2指定

(*OpenMP版 THREADS=8・Kfast指定)



効率は6% @ 100,032コア

ナノバイオ工学系問題の扱いの難しさ

■ 遺伝子工学の組み合わせ爆発 ■

200 残基のタンパク質 【推算: 林重彦氏(京大)】

- ・ 2 変異: ~7,000,000 通り
- ・ 3 変異: ~10,000,000,000 通り
- ・ DNA 合成のコスト: \$200 / 変異体?!

従来の ランダムな変異 +
スクリーニングでは限界...

■ 計算・シミュレーションの困難 ■

古典MD

- ・ 相対コストは小さい、電子的な議論は不可

量子論計算

- ・ 固体用の密度汎関数法は扱える原子数が少ない
- ・ 量子/古典の混成でも量子領域が大きく取れない
- ・ 結果からアミノ酸単位の詳細な情報が得にくい

<日本科学未来館のビデオ>

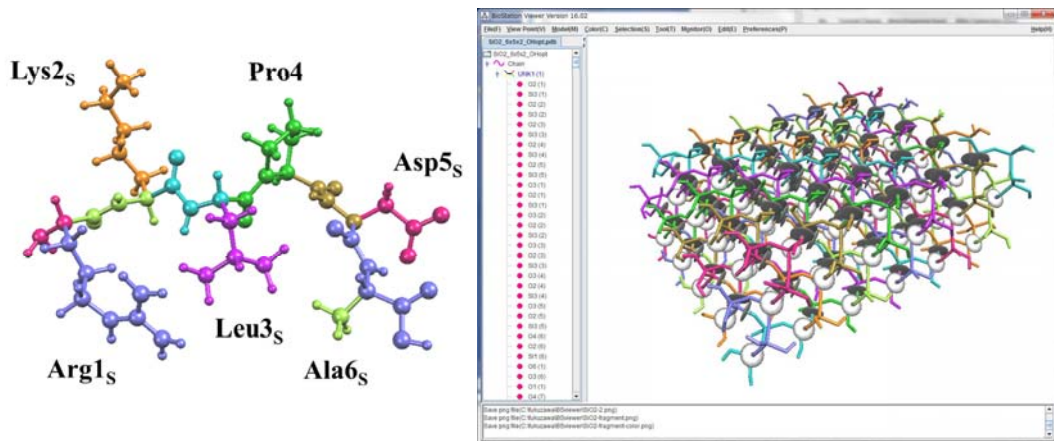


<https://www.youtube.com/watch?v=Q4gTV4r0zRs>

**FMO4法を武器
に開拓したい!**

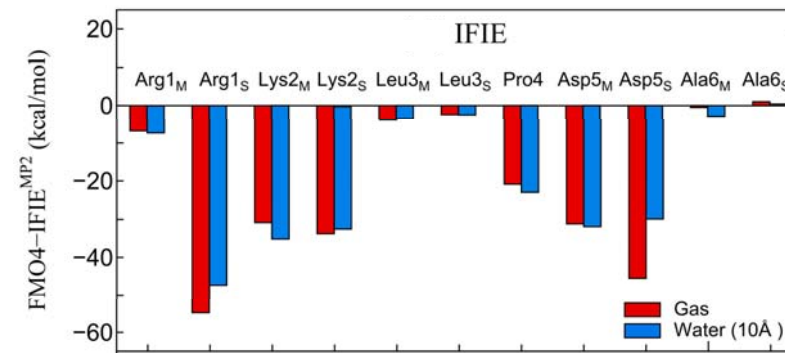
(ABINIT-MP/
BS-Viewer
システム)

シリカとRKLPGAペプチドの解析例 (2013年4月 発表)



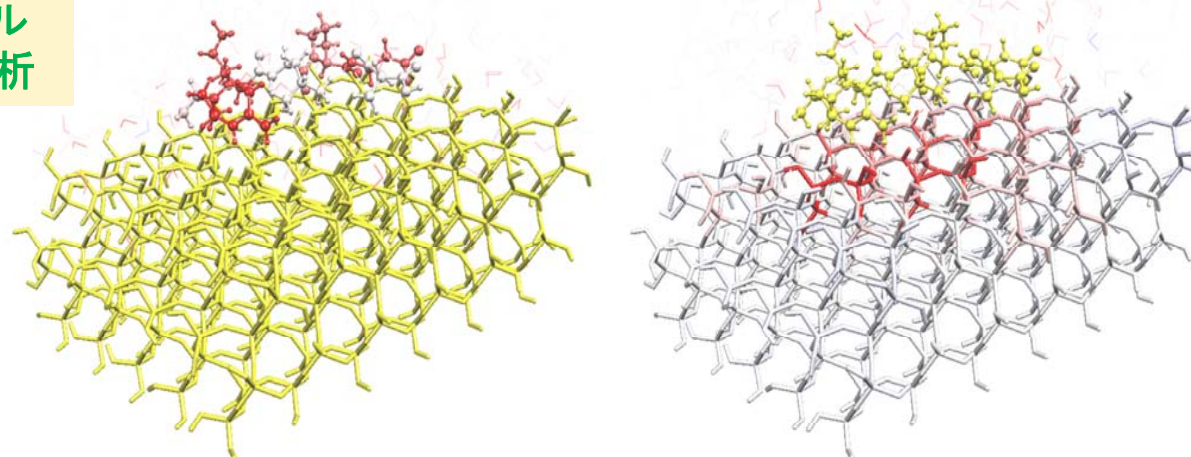
【佐野&芝らのRKLPGAペプチドとSi257個を含むシリカクラスター(分割済)】

構造サンプルは1つのみ、
FMO4-MP2(CD)レベル
で相互作用の詳細を解析



【各アミノ酸のシリカとの相互作用エネルギー】

FMO4法と分割法の工夫に
よって初めて計算が可能に
(FOCUSスパコンを利用)



【相互作用エネルギーの可視化 — 左側;ペプチド側 / 右側;シリカ側 (赤は安定化、青は不安定化)】

HAとESQESペプチドの解析例（今回の発表）

$[\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2]$ の4x4x2超格子

◇計算対象

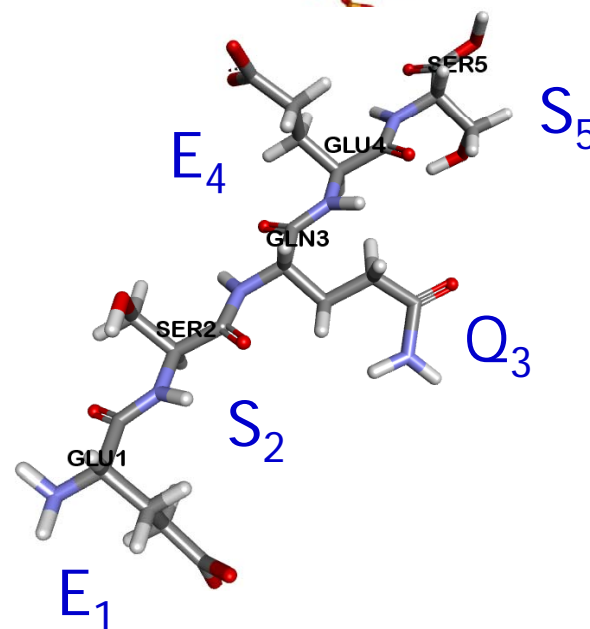
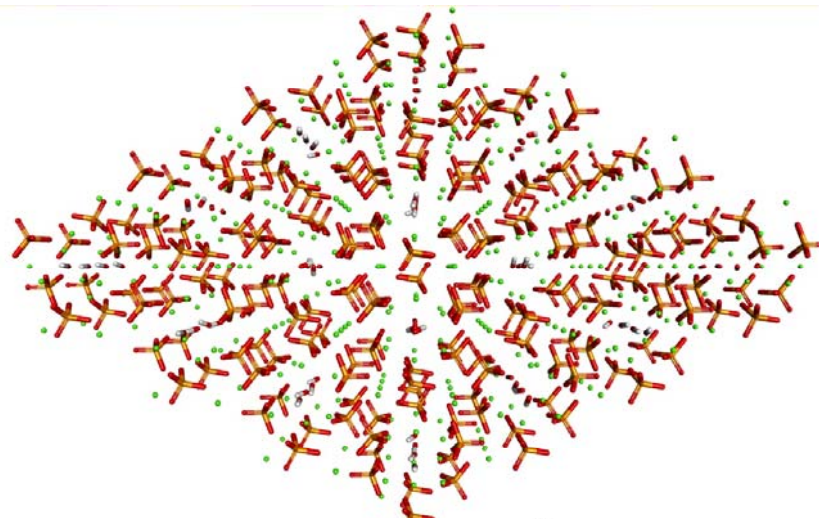
- ・ 高生体親和性の**ヒドロキシアパタイト** (HA)
- ・ 計1408原子の巨大なHA結晶モデル
→ **過去最大**（これまでは数十個程）
- ・ HAに特異吸着する**ESQES**人造ペプチド
→ 各アミノ酸残基の役割に興味

◇構造サンプル

- ・ 古典MD計算の3系列で**計30個**を調製
→ 統計的な評価が可能
- ・ 水分子は露わに配置

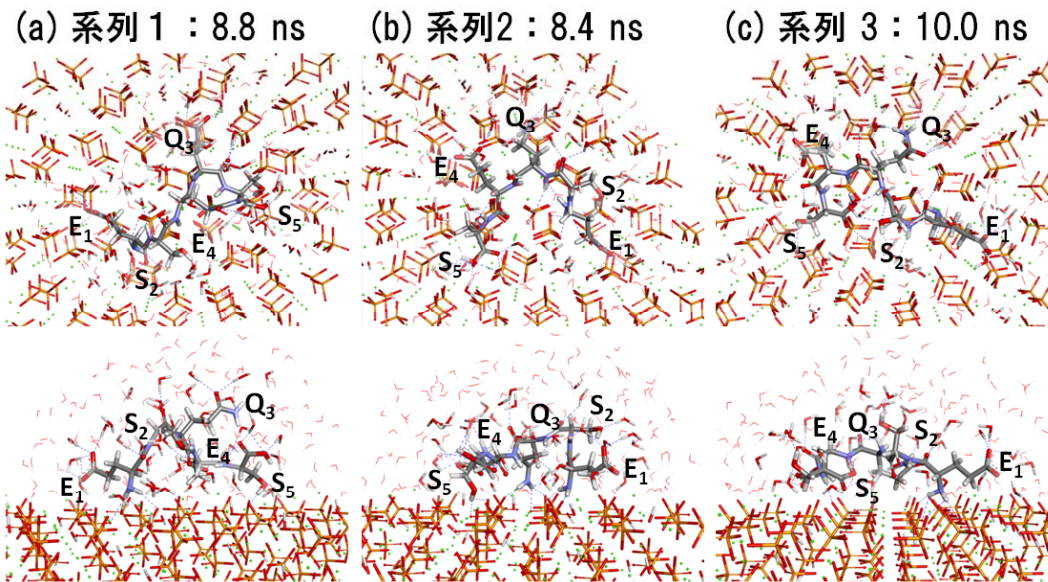
◇FMO4-MP2計算

- ・ 高イオン性を考慮した条件セッティング
→ **シリカの場合よりも困難**
- ・ 96コアの小型サーバで敢えて実施
→ 1構造あたりで約3時間
- ・ 解析の**ワークフロー**を確立

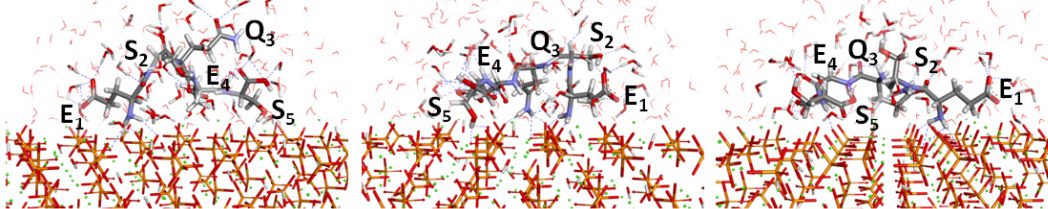


代表的サンプル構造と相互作用エネルギー（平均）

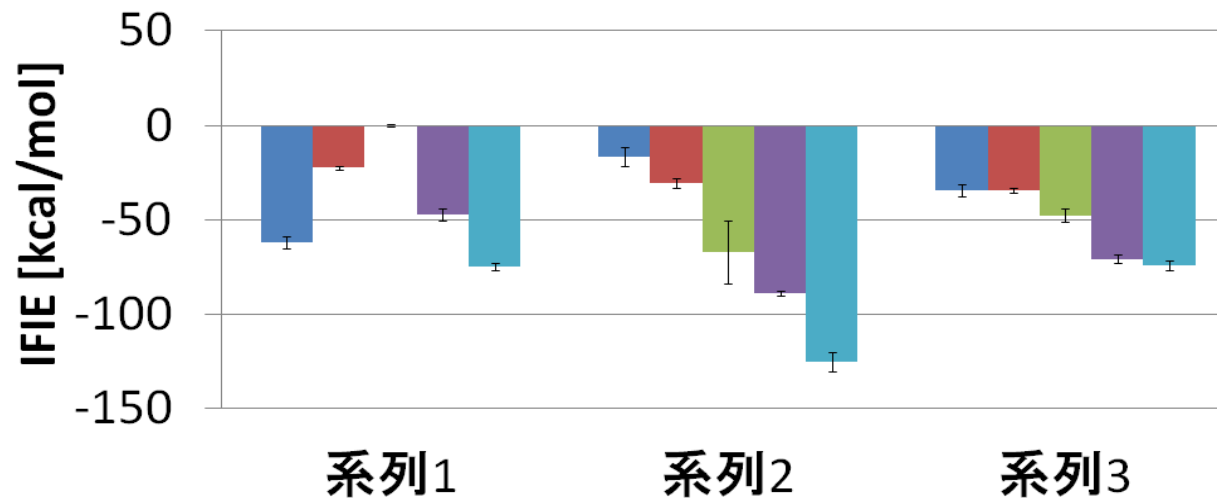
上視図



側面図



■ E1 ■ S2 ■ Q3 ■ E4 ■ S5

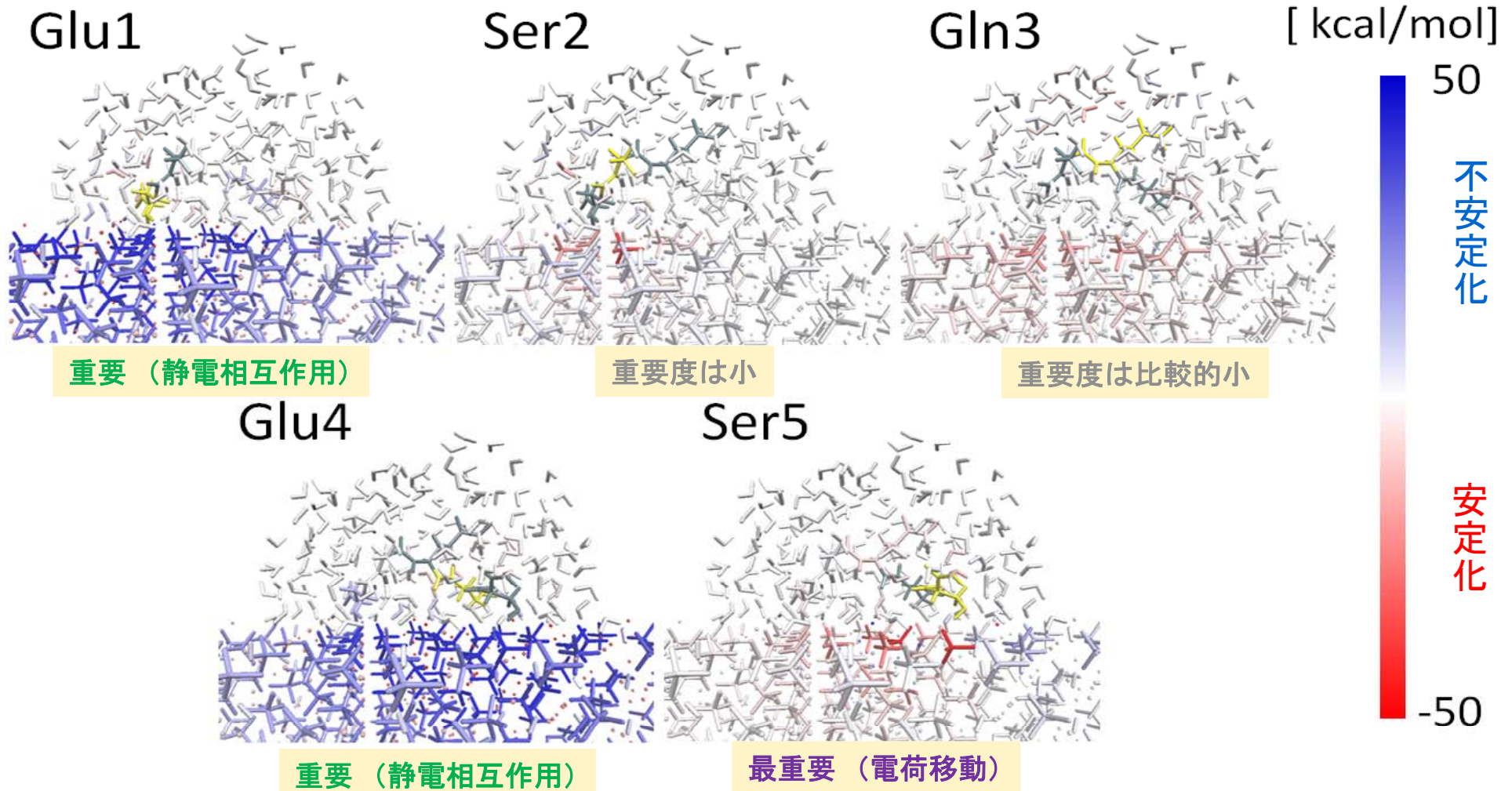


3系統でペプチドの構造にかなりの差が見られる (S₅は常に表面近接)

3系統でアミノ酸の役割に有意な差が見られる (S₅は常に最重要)

HA表面とアミノ酸毎の相互作用の可視化

遮蔽効果を含むSCIFIEの結果



まとめ、ABINIT-MPの関連状況、今後の展開

■ ナノバイオ技術関係のものづくり分野

- ・ 前回の“シリカ+RKLPDA”、今回の“ヒドロキシアパタイト+ESQES”の成功
 - 当該分野での合理的な設計ツールとしての有用性の実証
 - 「京」等の超並列資源の利活用によりコラーゲンなども扱い可能
- ・ 次のステップ: 数百サンプルの処理、“**計算機内変異**”、機械学習による知識導出

■ ABINIT-MPバイナリ公開の現状

- ・ Ver. 6+を「京」のライブラリとして提供済
- ・ Ver. 7のIntel Xeon用をBioStation Viwerと共に東大生産研のサイトで提供

■ ソース公開の準備

- ・ **Ver. 7準拠版**: テストとドキュメントの整備を行い、2015年度中にリリース予定
- ・ 未公開の機能(高次相関計算等)を含む版: 2016年度以降に提供

■ コンソーシアム

- ・ **ソースの共有と機能強化のためのコミュニティ**、2015年度内に立ち上げ
- ・ 産官学の構成、商用利用については基準/規約を要検討
- ・ **ポスト「京」/課題⑥E**の研究開発と連動の予定
- ・ 既に確立している**創薬系は別コミュニティ**で活動
 - 「HPCIを活用したFMO創薬プラットフォームの構築(福澤薫 代表)」
- ・ バイナリは今後ともフリーで利用可能 (“計算工学ナビ”をポータルに)

ABINIT-MPによるFMO計算のロードマップ

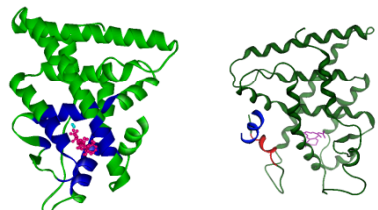
2004年

2008年

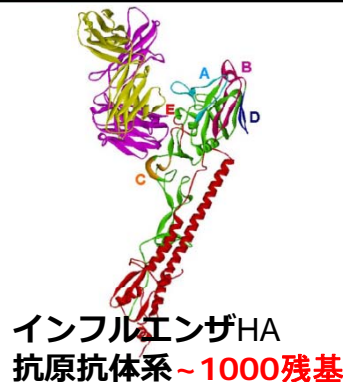
2010年

2013年

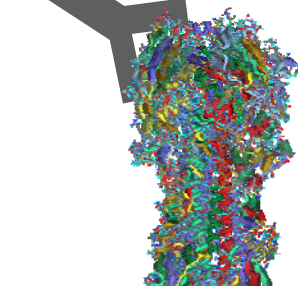
2015年



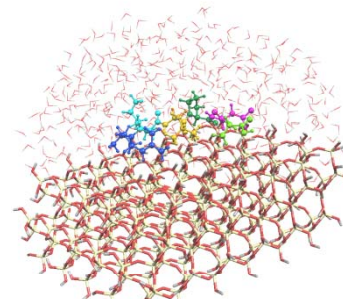
核内受容体(ER, RXR)
HIVプロテアーゼ ~300残基



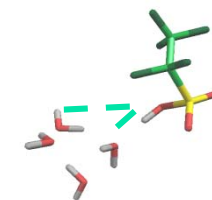
インフルエンザHA
抗原抗体系 ~1000残基



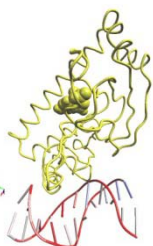
インフルエンザHA 3量体
抗原抗体系 ~2400残基



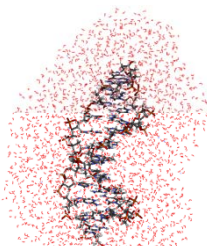
結晶-ペプチド複合系
~(SiO₂)₂₆₀-6残基-水



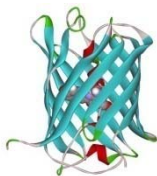
粗視化MD用
パラメータ
~1万サンプル



CRP
11塩基対+200残基



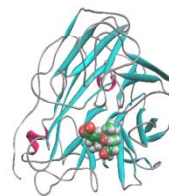
水和DNA
12塩基対+2500水



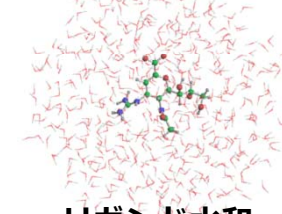
mFruits



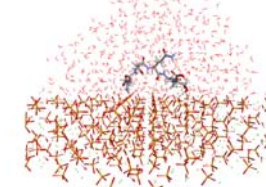
EGFR
チロシンキナーゼ



インフルエンザNA
タミフル ~400残基



リガンド水和
10 Å 水和水



結晶-ペプチド複合系
[Ca₅(PO₄)₃OH]₆₄-
5残基-水

計算レベル

MP2

CIS

CIS(D) CD近似

高次電子相関法
MPnスケールリング法

FMO4

複スケール
モデリング

構造最適化

PDB一点計算/
モデル埋め戻し

古典MD
サンプリング

ONIOM
(QM/MM)

FMO-MP2最適化(部分)

解析

IFIE

CAFI

FILM
IFIE map

励起エネルギー

BSSE

FMO4-IFIE

統計性

電荷、溶媒効果

Mulliken

ESP, RESP

自然密度

PB水和