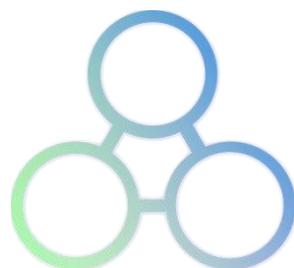


フラグメント分子軌道(FMO)計算プログラム ABINIT-MPの現状と今後

望月祐志（立教大学） fullmoon@rikkyo.ac.jp



ABINIT-MP



謝辞：講演の機会をいただいた(一社)スーパーコンピューティング・
ジャパンの運営の皆様、中島研吾先生(東京大)に感謝いたします。

ABINIT-MPの開発体制

◇ABINIT-MP Ver. 2系の主たるコラボレータ(敬称略)

中野達也(国立衛研)、坂倉耕太(FOCUS)、加藤季広(NEC)、
土居英男(立教大)、奥脇弘次(JSOL/立教大)、
片桐孝洋(名古屋大学)、大島聡史(九州大学)

◇研究開発支援

CISS/HPCIの4プロジェクト; 東大生研&文科省 / FY2002-2015
FS2020(ポスト「京」)プロジェクト; 東大&文科省 / FY2014-2019
CRESTプロジェクト(“田中FMO”); JST / FY2004-2009
科研費(“櫛-特定領域”); 文科省 / FY2008-2009
科研費(基盤B:代表); 文科省 / FY2016-2018
SFR; 立教大 / FY2006-2007, 2010-2014, 2019-2020, 2022-
NEC様; SX-AT高速化&機能強化の共同開発 / FY2020-
その他企業様からの立教大宛のご寄付
JHPCN課題; jh210036-NAH, jh220010, jh230001

ABINIT-MPの開発関係の論文リスト

(赤字: HPC分野の先生方 / 紫字: HPC分野の技術者方)

■ 英文

- (1) "Electron-correlated fragment-molecular-orbital calculations for biomolecular and nano systems", S. Tanaka*, Y. Mochizuki*, Y. Komeiji, Y. Okiyama, K. Fukuzawa, *Phys. Chem. Chem. Phys.*, **16** (2014) 10310-10344.
- (2) "The ABINIT-MP Program", Y. Mochizuki*, T. Nakano, K. Sakakura, Y. Okiyama, H. Watanabe, K. Kato, Y. Akinaga, S. Sato, J. Yamamoto, K. Yamashita, T. Murase, T. Ishikawa, Y. Komeiji, Y. Kato, N. Watanabe, T. Tsukamoto, H. Mori, K. Okuwaki, S. Tanaka, A. Kato, C. Watanabe, K. Fukuzawa (pp. 53-67) in *Recent Advances of the Fragment Molecular Orbital Method - Enhanced Performance and Applicability*, ed. Y. Mochizuki, S. Tanaka, K. Fukuzawa (January 2021, Springer).

■ 邦文

- (1) "FMOプログラムABINIT-MPの開発状況と機械学習との連携", 望月祐志*, 坂倉耕太, 秋永宜伸, 加藤幸一郎, 渡邊啓正, 沖山佳生, 中野達也, 古明地勇人, 奥沢明, 福澤薫, 田中成典, *J. Comp. Chem. Jpn.*, **16** (2017) 119-122.
- (2) "FMOプログラムABINIT-MPのOakForest-PACS上での多層並列化と性能評価", 渡邊啓正*, 佐藤伸哉, 坂倉耕太, 齊藤天菜, 望月祐志, *J. Comp. Chem. Jpn.* **17** (2018) 147-149.
- (3) "ABINIT-MP Openシリーズの最新の開発状況について", 望月祐志*, 秋永宜伸, 坂倉耕太, 渡邊啓正, 加藤幸一郎, 渡辺尚貴, 奥脇弘次, 中野達也, 福澤薫, *J. Comp. Chem. Jpn.*, **18** (2019) 129-131.
- (4) "FMOプログラムABINIT-MPの整備状況2020", 望月祐志*, 坂倉耕太, 渡邊啓正, 奥脇弘次, 加藤幸一郎, 渡辺尚貴, 沖山佳生, 福澤薫, 中野達也, *J. Comp. Chem. Jpn.*, **19** (2020) 142-145.
- (5) "FMOプログラムABINIT-MPの整備状況2021", 望月祐志*, 中野達也, 佐藤伸哉, 坂倉耕太, 渡邊啓正, 奥脇弘次, 大島聡史, 片桐孝洋, *J. Comp. Chem. Jpn.*, **20** (2021) 132-136.
- (6) "FMOプログラムABINIT-MPの整備状況2022", 望月祐志*, 中野達也, 坂倉耕太, 渡邊啓正, 佐藤伸哉, 奥脇弘次, 秋澤和輝, 土居英男, 大島聡史, 片桐孝洋, *J. Comp. Chem. Jpn.*, **21** (2022) 106-110.
- (7) "FMOプログラムABINIT-MPの整備状況2023", 望月祐志*, 中野達也, 坂倉耕太, 奥脇弘次, 土居英男, 加藤季広, 滝沢寛之, 成瀬彰, 大島聡史, 星野哲也, 片桐孝洋, *J. Comp. Chem. Jpn.*, accepted (10.2477/jccj.2024-0001).

■ 動画解説

- (1) 理研 配信講義 計算科学技術特論B(2022) 2022/6/9 & 2022/6/16 - "ABINIT-MPプログラムによるフラグメント分子軌道(FMO)計算1,2", 担当: 望月祐志, <<https://www.r-ccs.riken.jp/outreach/schools/20220407-1/>>.

近年のFMO計算の発展をまとめた本

2021年1月刊行

springer.com/gp/book/9789811592348

Springer

Search

Home Subjects Services Springer Shop About us

» Chemistry » Theoretical and Computational Chemistry

© 2021

**Recent Advances of the
Fragment Molecular Orbital
Method**

Enhanced Performance and Applicability

Editors: **Mochizuki, Yuji, Tanaka, Shigenori, Fukuzawa, Kaori** (Eds.)

Compiles a number of realistic and illustrative applications with color figures

» see more benefits

Buy this book

▼ eBook **96,29 €**
price for Spain (gross)

Buy eBook

- ISBN 978-981-15-9235-5
- Digitally watermarked, DRM-free
- Included format: PDF, EPUB
- ebooks can be used on all reading devices
- Immediate eBook download after purchase

▶ Hardcover **124,79 €**

VISA MasterCard AMERICAN EXPRESS PayPal INVOICE

» FAQ » Policy

FMO法とABINIT-MPプログラム

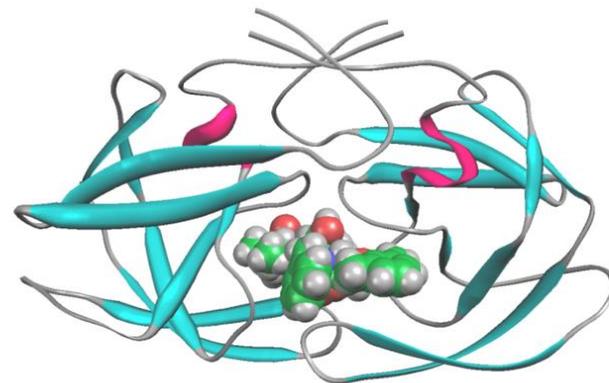
フラグメント分子軌道(FMO)法

◇巨大分子系

生体高分子や凝集系では一般的

⇒ タンパク質、DNA (水和状態)

数千～数万原子、数千～数十万軌道



【HIVプロテアーゼとロピナビル】

◇分割&統合系のアプローチの一つ

北浦らが22年程前に2体展開で提案

⇒ 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)

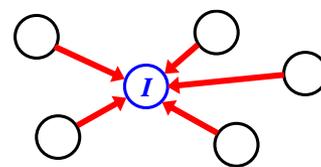
⇒ 定量性・信頼性を高める電子相関の導入も直截

フラグメント間の相互作用エネルギー (IFIE)

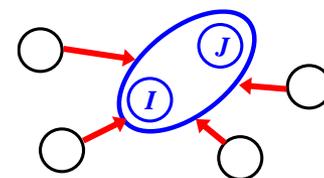
⇒ 計算対象の解析ツール

⇒ 生物物理や創薬に向く

⇒ 材料系にも適用可能



モノマー (アミノ酸単位など)



ダイマー

フラグメント間相互作用エネルギー(IFIE)

- FMO-HF法の全エネルギー E と全電子密度 ρ の計算式

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

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

N_f : フラグメント(モノマー)の個数

E_f : モノマーのエネルギー, E_{IJ} : ダイマーのエネルギー

ρ_f : モノマーの電子密度, ρ_{IJ} : ダイマーの電子密度

- FMO-HF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** $\Delta\tilde{E}_{IJ}$ 、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。
- 電子相関エネルギーの補正は加成的に行えばよい。

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

IFIE解析 ⇒ 対象系内部の詳細情報

(Inter-Fragment Interaction Energy)

FMO計算のためのプログラム

◇GAMESS-US [米国Gordonグループ]; Fedorov、Gordon、北浦ら
GAUSSIANに抗し得る有力なフリーソフト、世界規模 (Fortran)
⇒ 様々な機能をFMO化、多彩な計算、GDDI並列

◇ABINIT-MP; 望月、中野ら

実用機能は十分、東大系PJ・CREST-PJなどで開発 (Fortran)

⇒ IOレス、MPI、OpenMP/MPI混成並列、スパコンと好相性

◇PAICS; 石川

FMO-MP2(RI)に特化 (C)

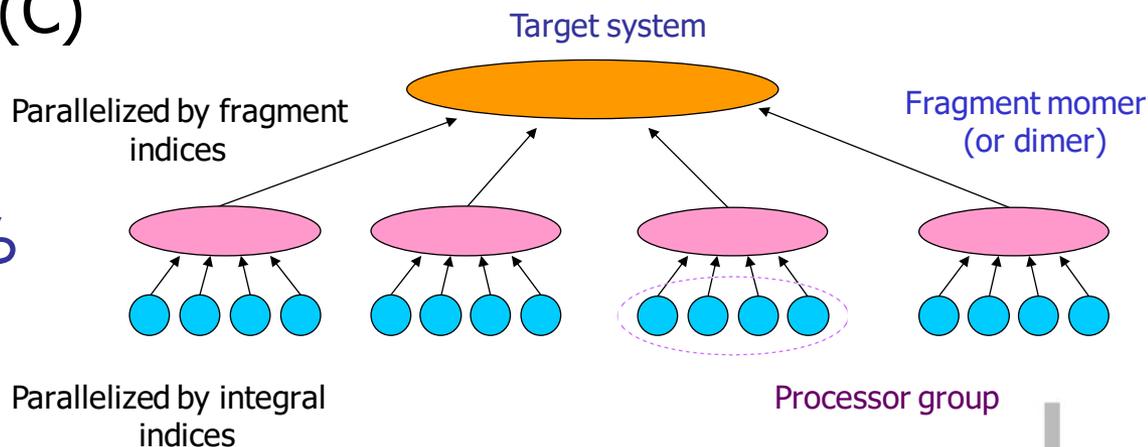
⇒ MPI並列

◇OpenFMO; 稲富、鬼頭ら

FMO-HF (C)

⇒ 超並列指向

2階層の並列処理で計算資源を有効利用



ABINIT-MPの主な機能（オープンシリーズとして整備中）

・エネルギー

- FMO4: HF, MP2
- FMO2: HF~CCSD(T), LRD
- FMO2: CIS/CIS(D), GF2

・エネルギー微分

- FMO4: HF, MP2
- FMO2: MP2構造最適化, MD

・その他機能

- SCIFIE, PB, sp2-BDA, $\alpha(\omega)$
- 電子密度生成, CAFI, FILM

・並列化環境(PC~スパコン)

- MPI, OpenMP/MPI混成
- 最深部はBLAS処理

http://www.cenav.org/abinit-mp-open_ver-2-rev-8/

計算工学ナビ

ものづくりにHPCを活用するためのツールとケーススタディー

ログイン | 検索

レポート ニュース 解析事例データベース ライブラリ Q&A コミュニティ

ABINIT-MP Openシリーズ (Ver.2 Rev.8)

※2021年9月版(Ver.2 Rev.4)に関するページはこちらです

※2020年6月版(Ver.1 Rev.22)に対応するBSV (BioStation Viewer) のWindows用インストールはこちらからダウンロードできます

はじめに

Open Ver. 1（ポスト「京」のPJで整備）

- ・Rev. 5 (2016年12月)
- ・Rev. 10 (2018年2月)
- ・Rev. 15 (2019年3月)
- ・Rev. 22 (2020年6月); 当面は併存

Open Ver. 2（「富岳」の時代に移行）

- ・Rev. 4 (2021年9月)
- ・Rev. 8 (2023年8月)

(注記: Ver. 2系では、BioStation Viewer へのデータファイルの書出しを廃止した)

FMOの主たる計算

◇FMO-HF (Hartree-Fock/SCF)

環境静電ポテンシャル(ESP)印加条件で占有分子軌道を最適化

- ⇒ モノマ一段階では**SCC(自己無撞着電荷)条件**まで反復
- ⇒ ダイマ一段階は反復なしで処理
- ⇒ **基底関数の2電子積分の生成**とFock行列の構築
- ⇒ ESP計算の近似処理、遠隔モノマ対の簡易計算アリ

◇FMO-MP2 (2nd-order Møller-Plesset)

電子相関補正の基本、相互作用エネルギーの定量性を向上

- ⇒ **2電子積分の添字変換** (基底関数→分子軌道)
- ⇒ DGEMMで処理可能

◇FMO-MP3 (3rd-order Møller-Plesset)

MP2より精度が改善、演算量とメモリ要求は大きく増加

- ⇒ **テンソル量の縮約**、DGEMM処理を多用

基本のHF計算

$$\mathbf{F}^x \mathbf{C}^x = \mathbf{S}^x \mathbf{C}^x \boldsymbol{\varepsilon}^x \quad \mathbf{F}^x = \mathbf{H}^x + \mathbf{G}^x \quad \Leftrightarrow \text{HFの一般化固有値問題 (要反復計算)}$$

$$H_{\mu\nu}^x = H_{\mu\nu}^{\text{core } x} + V_{\mu\nu}^x + \sum_k B_k \langle \mu | \theta_k \rangle \langle \theta_k | \nu \rangle \quad V_{\mu\nu}^x = \sum_{K \neq x} (u_{\mu\nu}^K + v_{\mu\nu}^K) \quad \Leftrightarrow \text{1電子部分の修飾}$$

$$u_{\mu\nu}^K = \sum_{A \in K} \langle \mu | (-Z_A / |\mathbf{r} - \mathbf{A}|) | \nu \rangle \quad v_{\mu\nu}^K = \sum_{\lambda\sigma \in K} P_{\lambda\sigma}^K (\mu\nu | \lambda\sigma) \quad \Leftrightarrow \text{環境静電ポテンシャル (ESP)}$$

$$P_{\mu\nu} = 2 \sum_{i=1}^{\text{occ}} C_{\mu i}^* C_{\nu i} \quad G_{\mu\nu}^x = \sum_{\lambda\sigma \in x} P_{\lambda\sigma}^x \left[(\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\sigma | \lambda\nu) \right] \quad \Leftrightarrow \text{2電子部分 } O(N^4) \text{ (並列処理)}$$

高速化のための幾つかの工夫

$$V_{\mu\nu}^L \cong \sum_{\lambda \in L} (\mathbf{P}^L \mathbf{S}^L)_{\lambda\lambda} (\mu\nu, \lambda\lambda) \quad \text{for } R_{\min}(X, L) \geq L_{\text{aoc}} \quad \Leftrightarrow \text{ESP-AOC近似 (実用精度高し)}$$

$$V_{\mu\nu}^L \cong \sum_{A \in L} \langle \mu | (Q_A / |\mathbf{r} - \mathbf{A}|) | \nu \rangle \quad \text{for } R_{\min}(X, L) \geq L_{\text{ptc}} \quad Q_A = \sum_{\lambda \in A} (\mathbf{P}^L \mathbf{S}^L)_{\lambda\lambda} \quad \Leftrightarrow \text{ESP-PTC近似 (速い)}$$

$$E'_{IJ} \cong E'_I + E'_J + \text{Tr}(\mathbf{P}^I \mathbf{u}^J) + \text{Tr}(\mathbf{P}^J \mathbf{u}^I) + \sum_{\mu\nu \in I} \sum_{\lambda\sigma \in J} \mathbf{P}_{\mu\nu}^I \mathbf{P}_{\lambda\sigma}^J (\mu\nu | \lambda\sigma) \quad \Leftrightarrow \text{Dimer-ES近似 (HF計算無)}$$

種々の工夫により、FMO2の計算コストのシステムサイズ依存性は2乗より低い

2電子積分の生成について

- ・ 小原のVertical Recurrence Relation (VRR)がベース
- ・ ジェネレータでF90コード群を自動生成（{s, p, d}関数では $3^4=81$ 種）

TABLE I. Recurrence expressions* for the electron repulsion integrals over *s* and *p* Cartesian Gaussian functions.

$$\begin{aligned}
 (ss, ss)^{(0)} &= (\zeta + \eta)^{-1/2} K(\zeta_a, \zeta_b, \mathbf{A}, \mathbf{B}) K(\zeta_c, \zeta_d, \mathbf{C}, \mathbf{D}) F_0(T) \\
 (p_i s, ss)^{(0)} &= (P_i - A_i)(ss, ss)^{(0)} + (W_i - P_i)(ss, ss)^{(1)} \\
 (p_i s, p_k s)^{(0)} &= (Q_k - C_k)(p_i s, ss)^{(0)} + (W_k - Q_k)(p_i s, ss)^{(1)} \\
 &\quad + \frac{\delta_{ik}}{2(\zeta + \eta)} (ss, ss)^{(1)} \\
 (p_i p_j, ss)^{(0)} &= (P_j - B_j)(p_i s, ss)^{(0)} + (W_j - P_j)(p_i s, ss)^{(1)} \\
 &\quad + \frac{\delta_{ij}}{2\zeta} \left\{ (ss, ss)^{(0)} - \frac{\rho}{\zeta} (ss, ss)^{(1)} \right\} \\
 (p_i p_j, p_k s)^{(0)} &= (Q_k - C_k)(p_i p_j, ss)^{(0)} + (W_k - Q_k)(p_i p_j, ss)^{(1)} \\
 &\quad + \frac{1}{2(\zeta + \eta)} \left\{ \delta_{ik}(sp_j, ss)^{(1)} + \delta_{jk}(p_i s, ss)^{(1)} \right\} \\
 (p_i p_j, p_k p_l)^{(0)} &= (Q_l - D_l)(p_i p_j, p_k s)^{(0)} + (W_l - Q_l)(p_i p_j, p_k s)^{(1)} \\
 &\quad + \frac{1}{2(\zeta + \eta)} \left\{ \delta_{il}(sp_j, p_k s)^{(1)} + \delta_{jl}(p_i s, p_k s)^{(1)} \right\} \\
 &\quad + \frac{\delta_{kl}}{2\eta} \left\{ (p_i p_j, ss)^{(0)} - \frac{\rho}{\eta} (p_i p_j, ss)^{(1)} \right\} \\
 (i, j, k, l = x, y, z)
 \end{aligned}$$

* For the definition of the variables, see the text.

SCHEME I.

(First step)

```

DO ICS = 1, n_CS  loops for the contracted shells
DO JCS = 1, ICS
DO IPS = 1, m_ICS  loops for the primitive shells
DO JPS = 1, m_JCS
The calculation of the parameters P, ζ, and K(ζ, ζ', R, R')
for each pair of primitive shells
CONTINUE

```

(Second Step)

```

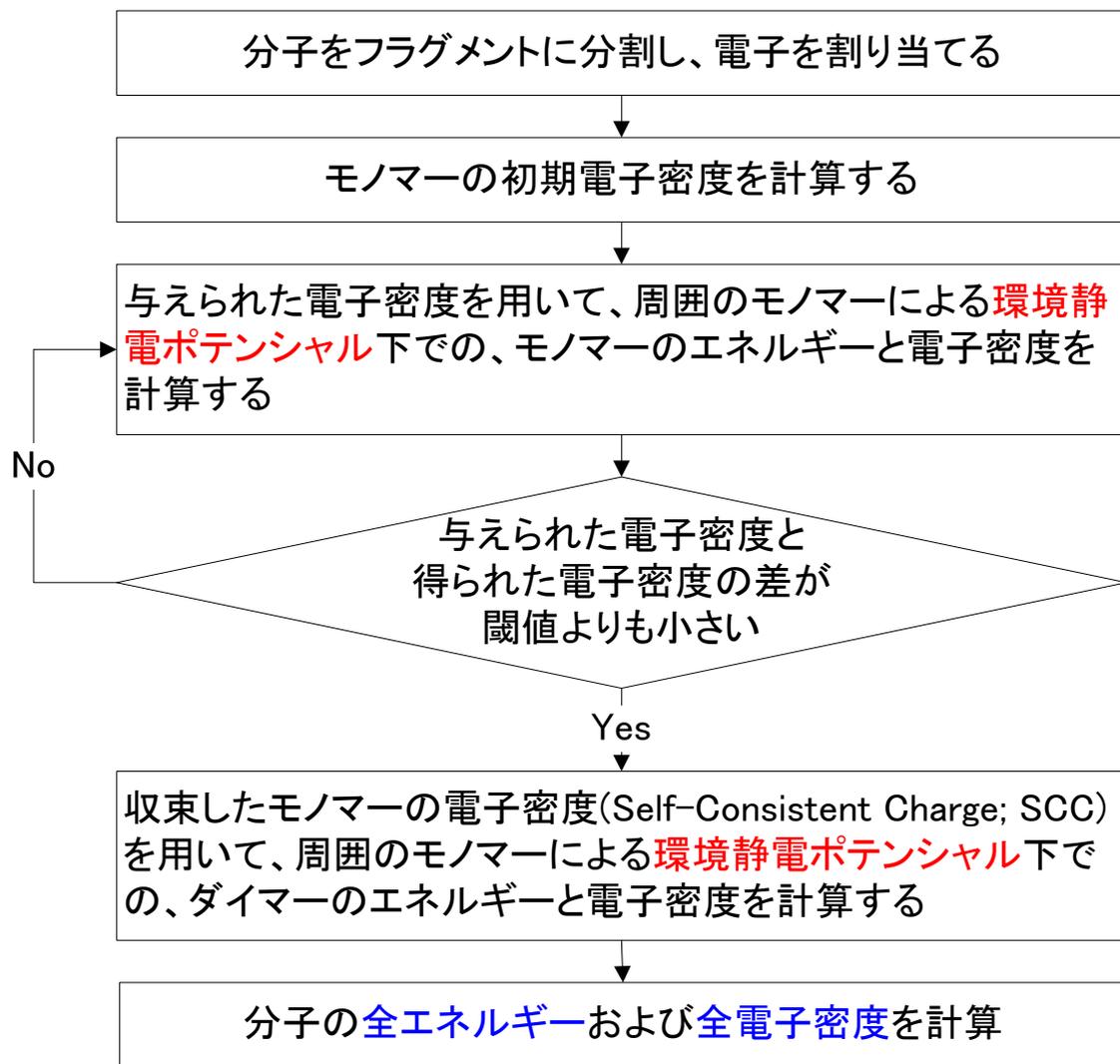
DO IPPS = 1, N_PPS  a loop for the first pair of primitive shells
DO JPSS = 1, IPPS  a loop for the second pair of primitive shells
The evaluation of ERI's
CONTINUE

```

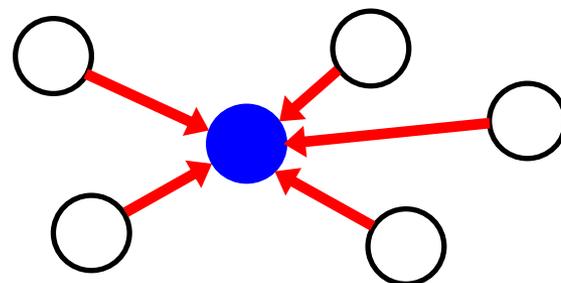
- ・ Gauss型関数の角運動量の昇降を利用
- ・ 並列化は短縮シェルの対のループで

$$G_{ijk;\alpha}(r) = Nx^i y^k z^l \exp(-\alpha r^2)$$

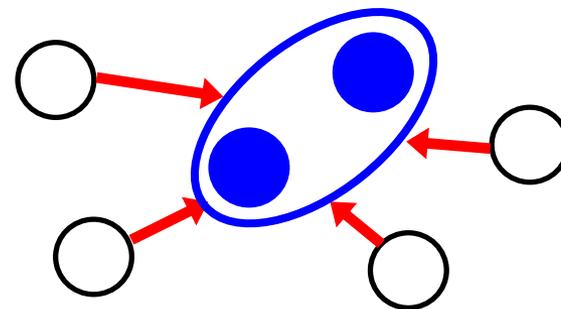
FMO-HF計算の流れ



モノマーの段階で自己無撞着電荷を課するのが特徴



モノマー

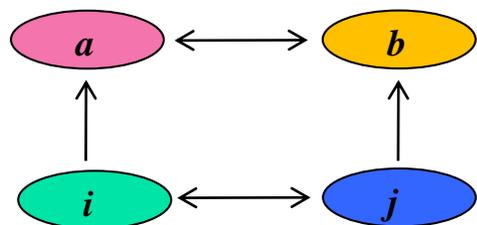


ダイマー

摂動論による電子相関の導入

揺動ポテンシャルが摂動

2次摂動(MP2)の相関: 占有軌道*i*と*j*の電子対が“衝突”して非占有軌道*a*と*b*に“散乱”



(電子対の相互作用)



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

(分子は電子間の反発積分の積、分母は軌道エネルギーの差)

3次摂動(MP3)の相関: 電子対間の相互作用も3通りで追加的に考慮して補正



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

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

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

- MP3の方がMP2よりも縮約が複雑
- 空間軌道の式に落として実装
- DGEMM処理を上手く利用する

MP2計算の基本アルゴリズム2種：形式的にN⁵

Loop over *i*-batch [parallelizable when needed]

Loop over σ [to be parallelized for worker processes]

Loop over λ

Generate $(\mu\nu, \lambda|\sigma)$ list [canonical relation for $\mu\nu$]

Do 1/4 transformation of $\mu \rightarrow i$ [screening & DAXPY]

Do 2/4 transformation of $\nu \rightarrow a$ [DDOT]

Do 3/4 transformation of $\lambda \rightarrow j$ [screening & DAXPY]

End of loop over λ

Do 4/4 transformation for $\sigma \rightarrow b$ [screening & DAXPY]

End of loop over σ ["all-reduce" must be done for $(ia|jb)$]

Calculate partial MP2 energy

End of loop over *i*-batch

Loop over *ij*-batch ! size depending on available memory

Loop over σ ! to be parallelized

Loop over λ

Preparing $(\mu\nu|\lambda\sigma)$! for canonical $\mu\nu$ -pair

Forming $(i\nu|\lambda\sigma)$! DGEMM, fixed $\lambda\sigma$, running over μ

Forming $(ia|\lambda\sigma)$! DGEMM, fixed $\lambda\sigma$, running over ν

Forming $(ia|j\sigma)$! DGEMM, fixed σ , direct-product for fixed λ

End of loop over λ

Forming $(ia|jb)$! DGEMM, direct-product for fixed σ

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

Calculate partial MP2 energy with respect to *ij*-batch

End of loop over *ij*-batch

MP2関連エネルギー補正

$$E_{\text{MP2}} = \sum_{ijab} \frac{(ia|jb)[2(ia|jb) - (ib|ja)]}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

添字の変換：4N⁵コスト

$$(ia|jb) = \sum_{\sigma} C_{\sigma b} \left(\sum_{\lambda} C_{\lambda j} \left(\sum_{\nu} C_{\nu a} \left(\sum_{\mu} C_{\mu i} (\mu\nu|\lambda\sigma) \right) \right) \right)$$

【第一版】

- ・ DAXPYとDDOTを使い、閾値判断を優先
- ・ 実効的演算数を下げる方針、15年程前のチップでは有効

【第二版】

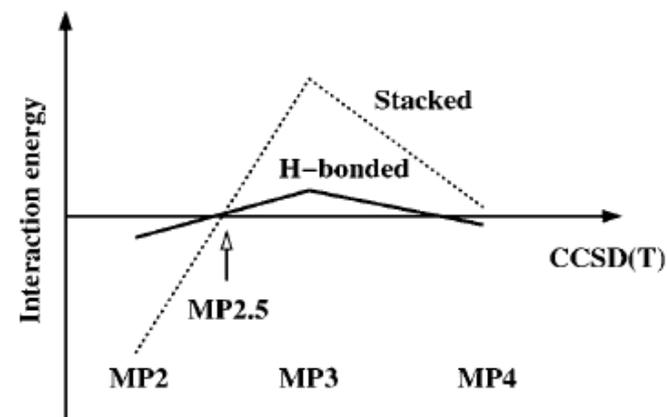
- ・ DGEMMの高性能に期待

【現在：混成型がデフォルト】

- ・ スカラ型のCPUでもステップ2とステップ4はDGEMMで

3次摂動論 (MP3) による相関エネルギー補正

- MP2に無い電子対間の相互作用を考慮
- OpenMP/MPI混成並列を前提に実装
- メモリ要求はMP2より増加、 $O(N^6)$ コスト
- 縮約/変換処理にはDGEMMを多用
- MP2.5スケールングを推奨



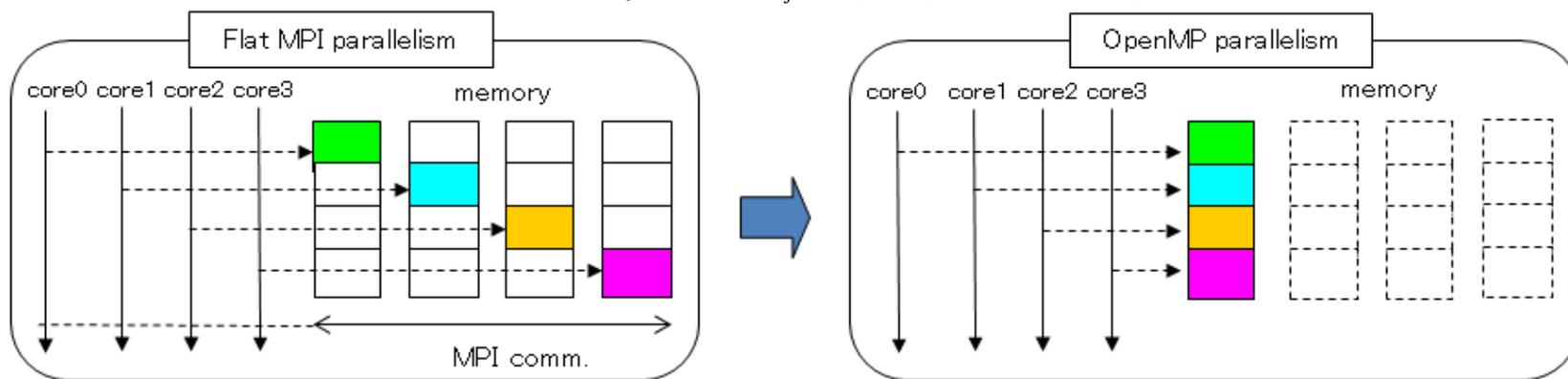
$$E^{MP3(2p-4h)} = \sum_{ijklab} \frac{(ia | jb)(ik | jl)[2(ka | lb) - (kb | la)]}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)(\epsilon_k + \epsilon_l - \epsilon_a - \epsilon_b)}$$

$$E^{MP3(4p-2h)} = \sum_{ijabcd} \frac{(ia | jb)(ac | bd)[2(ic | jd) - (id | jc)]}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)(\epsilon_i + \epsilon_j - \epsilon_c - \epsilon_d)}$$

$$E^{MP3(3p-3h)-1} = \sum_{ijkabc} \frac{[2(ia | jb) - (ij | ab)][2(kc | ia) - (ka | ic)][2(kc | jb) - (kb | jc)]}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)(\epsilon_i + \epsilon_k - \epsilon_a - \epsilon_c)}$$

$$E^{MP3(3p-3h)-2} = -3 \sum_{ijkabc} \frac{(ij | ab)(ka | ic)(kb | jc)}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)(\epsilon_i + \epsilon_k - \epsilon_a - \epsilon_c)}$$

和の添字は6つ



MP3計算のループ (*ij*バッチと*k*バッチの2重構造)

- ・メモリが少ないとバッチ数が増える
- ・残念ながら「京」では使えなかった

Loop over *ij*-batch ! size depending on available memory

Loop over σ ! parallelized

Form $\{(ia|jb), (ij|ab), (ik|jl)\}$! 1st-4th quarter transformations

End of loop over σ

Calculate partial MP2 energy ! parallelized for *ij*

Prepare half-backtransformed 1st-order amplitude ! parallelized for *ij*

Perform EEO processing for “2h-4p” contribution ! parallelized for $\mu\nu$

Calculate partial “2h-4p” energy ! parallelized for *ij*

Loop over *k*-batch ! size depending on available memory

Loop over σ ! parallelized

Form $(kc|ld)$! 1st-4th quarter transformations

End of loop over σ

Calculate partial “4h-2p” and “3h-3p” energies ! parallelized for *ij*

End of loop over *k*-batch

End of loop over *ij*-batch

Sum up “2h-4p”, “4h-2p” and “3h-3p” energies for final MP3 energy

$$E^{MP3(4p-2h)} = \sum_{ij} \left(\sum_{ab} \frac{(ia|jb)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \left(\sum_{cd} (ac|bd) \left(\frac{2(ic|jd) - (id|jc)}{\varepsilon_i + \varepsilon_j - \varepsilon_c - \varepsilon_d} \right) \right) \right)$$

Fock行列的处理で
(*ac|bd*)構築を回避

「富岳」の新型コロナ特別PJ(2020年度)

試行的運用の「富岳」を使った特別PJ



新型コロナウイルス対策に関する貢献 - Society5.0的 社会要求に対する迅速な対応 -



医学的側面からの研究

「富岳」による 新型コロナウイルスの治療薬候補同定



分子動力学計算により、約2000種の既存医薬品の中から、新型コロナウイルスの標的タンパク質に高い親和性を示す治療薬候補を探索・同定する。

(課題代表者; 理化学研究所/京都大学 奥野 恭史)

「富岳」を用いた新型コロナウイルス 表面のタンパク質動的構造予測



クライオ電子顕微鏡によって解かれたウイルス表面タンパク質の立体構造を初期モデルとして、その立体構造の動きを「富岳」を用いた分子動力学計算で予測する。

(課題代表者; 理化学研究所 杉田 有治)

新型コロナウイルス関連タンパク質に対する フラグメント分子軌道計算

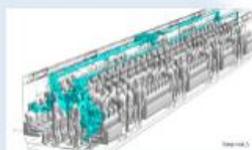


新型コロナウイルス関連タンパク質に対するフラグメント分子軌道計算を系統的に実施し、詳細な相互作用解析を行う。

(課題代表者; 立教大学 望月 祐志)

室内環境におけるウイルス飛沫感染の 予測とその対策

通勤列車内、オフィス、教室、病室といった室内環境において、新型コロナウイルスの特性を考慮した飛沫の飛散シミュレーションを行い、感染リスク評価を行った上で、感染リスク低減対策の提案を行う。

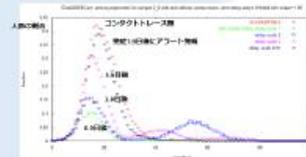


(課題代表者; 理化学研究所/神戸大学 坪倉 誠)

社会的側面からの研究

パンデミック現象および対策の シミュレーション解析

今後生じうる社会経済活動への影響を評価し、収束シナリオとその実現方法を探る。あわせてウイルスの変異などにより感染・発病の経過が変化した場合に起こりうる事象への対応を立案する。

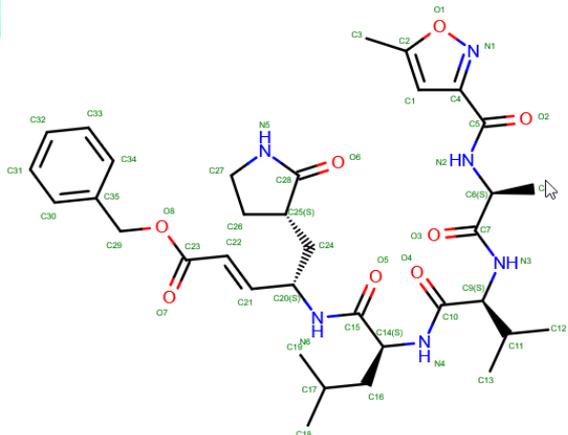


(課題代表者; 理化学研究所 伊藤 伸泰)

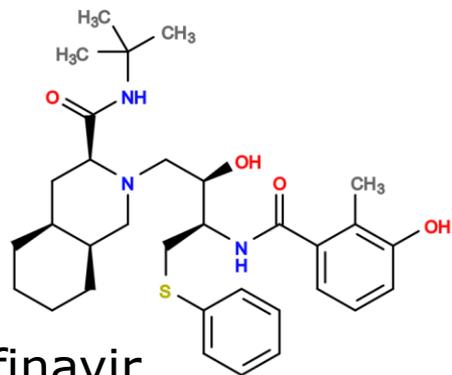
理化学研究所 計算科学研究センターは、スーパーコンピュータを用いた新型コロナウイルス対策研究のため本年3月に米国にて設立されたコンソーシアム (COVID-19 High Performance Computing Consortium) に参加

メインプロテアーゼ (Mpro) と阻害剤

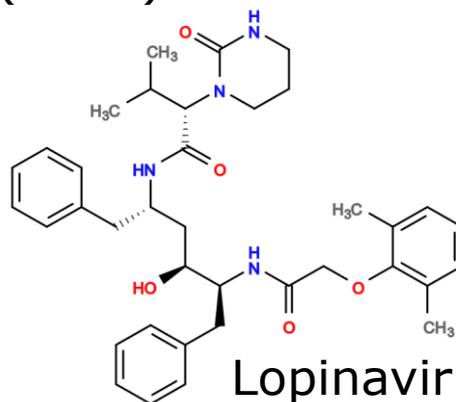
メインプロテアーゼは増殖に関わる



N3 SARS-Cov-2 Mpro (6LU7)

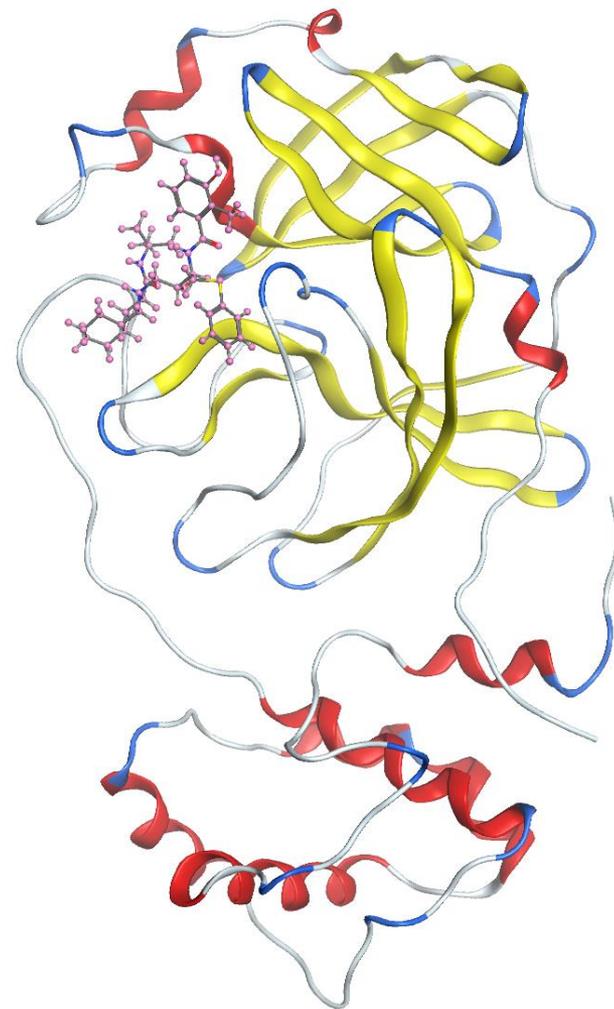


Nelfinavir
HIV-1 Protease (3EL5)



Lopinavir
HIV-1 Protease (6DJ1)

	Nelfinavir	Lopinavir
EC ₅₀ (μ M)	1.13	5.73

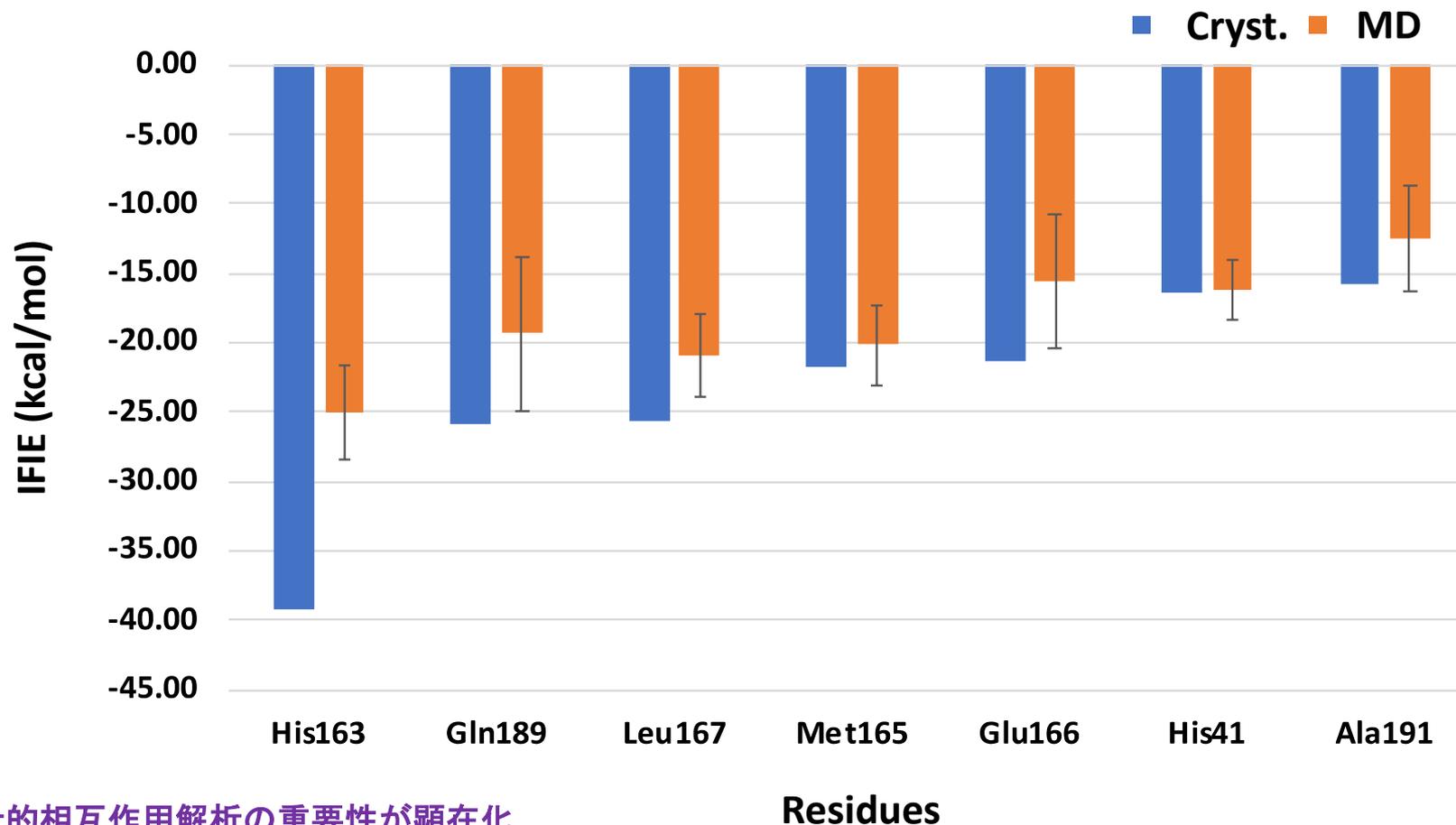


Nelfinavir-Mpro系の全体図

Mpro + N3阻害剤の統計的相互作用解析

395フラグメント(タンパク質:306)

古典MDで1000構造をサンプルしてFMO-MP2/6-31G*(PR)レベルで計算

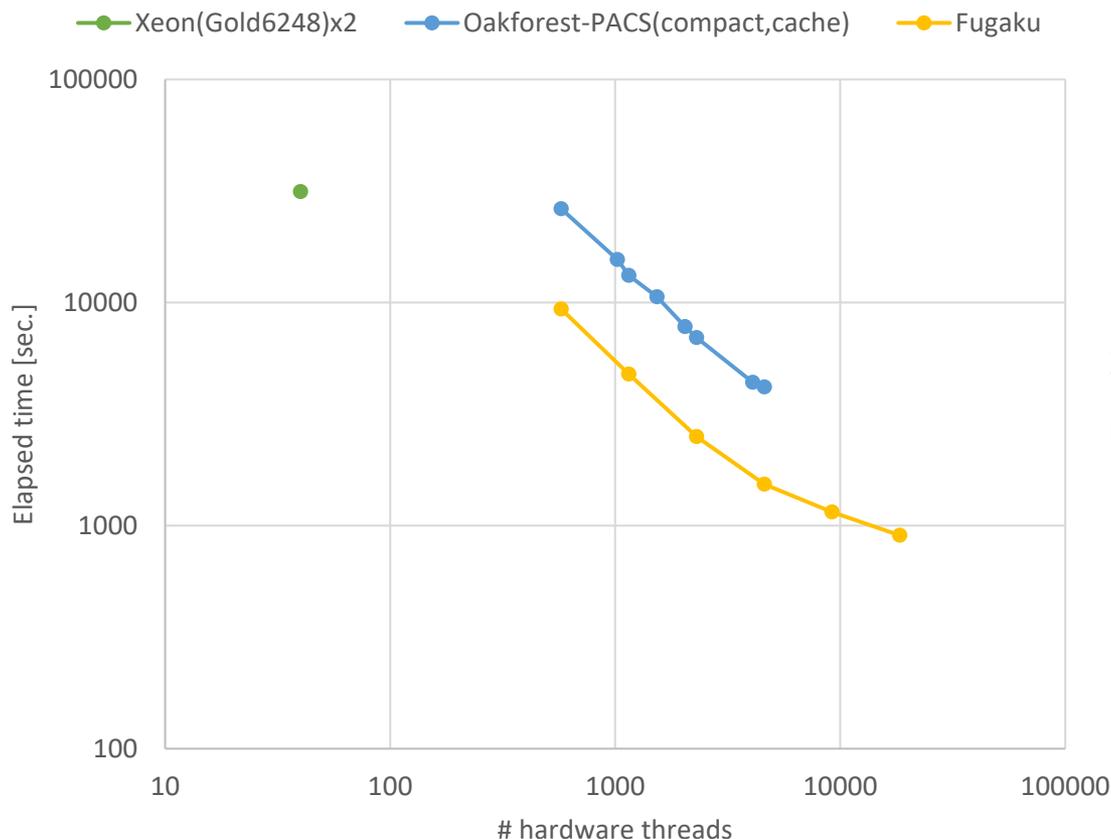


統計的相互作用解析の重要性が顕在化

1構造あたり0.6時間 @ 「富岳」192ノード / 一斉投入で5時間で処理

FMO-MP2/6-31G*ジョブのスケーリング

6LU7 - FMO2-MP2/6-31G* - Elapsed time



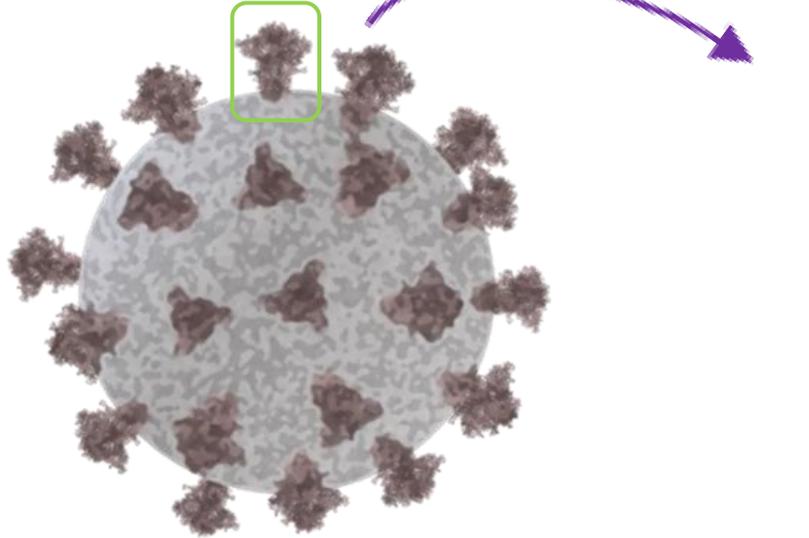
Xeon; 40 cores
OFP; up to 4608 cores
Fugaku; up to 18432 cores

September 2020

- PDB ID: 6LU7 = SARS-CoV-2 Mpro + N3 ligand の系
- MP2の積分変換は全てDGEMMで実行
- Dimer-ESのCMM近似は (>5のリージョンで使用)
- 「富岳」はOakforest-PACSよりも2.8倍ほど速い

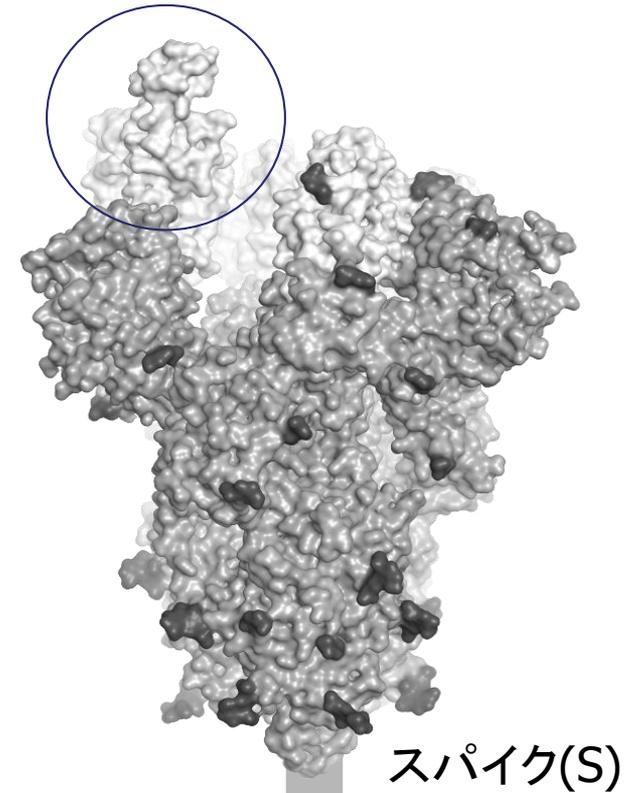
スパイクタンパク質

SARS-CoV-2



拡大

RBD→ヒト受容体に結合



スパイク(S)

脂質二重膜(E)

【感染の仕組み】

スパイクタンパク質のReceptor Binding Domain (RBD)ヒトのAngiotensin Converting Enzyme 2 (ACE2)以上の2つのタンパク質同士が相互作用することで感染

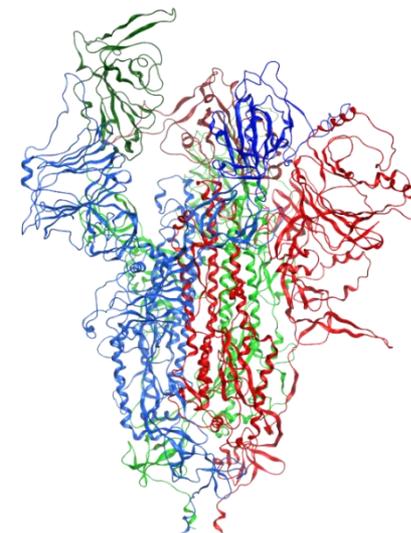
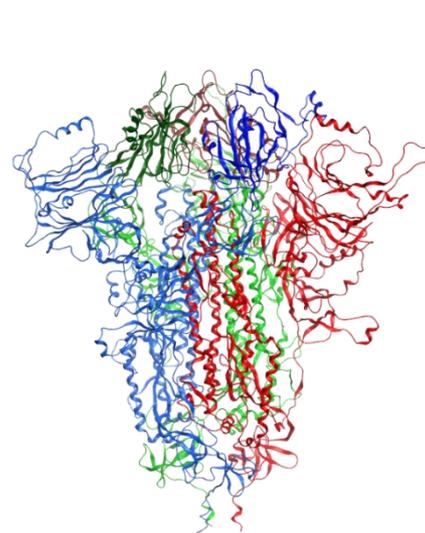
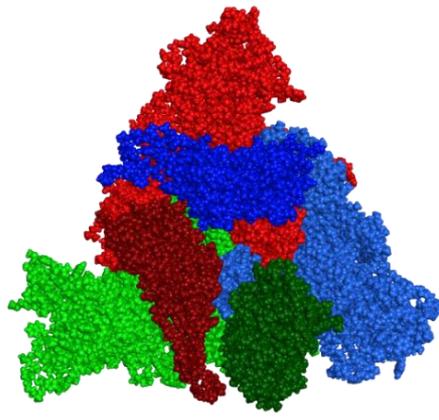
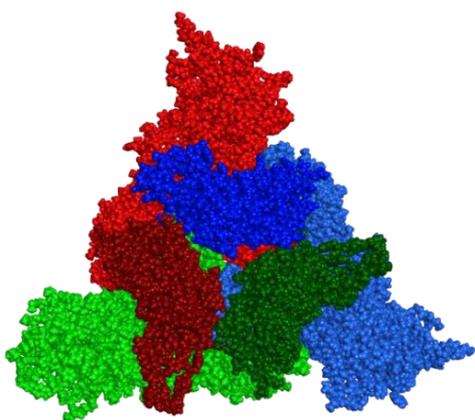
RBDに変異が起きている「変異ウイルス」は、ACE2との結合形態が野生型のSARS-CoV-2と異なっている可能性がある

クローズ構造とオープン構造

注: 糖鎖は削除

S-タンパク質をRBD側から見た図

S-タンパク質を横から見た図



closed構造 (6VXX)

open構造 (6VYB)

closed構造 (6VXX)

open構造 (6VYB)

■ Chain A ■ Chain B ■ Chain C 色が濃い領域は各ChainにおけるRBD (Thr333-Pro527) 領域

S-タンパク質は1.1千残基のタンパク質鎖の3量体で構成

Closed構造はすべてのRBD領域が閉じている構造open構造はB (RBD) のみが開いた構造

クローズ型(6VXX)のタイミング

3.3千残基(フラグメント)

FMO-MP2/cc-pVDZ

```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =          3892.7 seconds
Elapsed time: Monomer MP2      =           27.7 seconds
Elapsed time: Monomer (Total)  =          3965.3 seconds
Elapsed time: Dimer ES        =           282.4 seconds
Elapsed time: Dimer SCF       =           838.0 seconds
Elapsed time: Dimer MP2       =           473.1 seconds
Elapsed time: Dimer (Total)   =          2001.4 seconds
Elapsed time: FMO (Total)     =          5966.6 seconds
```

```
*** Write check point file (CPF) open1.0 rev10 ***
```

```
WriteGeom = 6vxx_nonag-min0430_50k-mp2-ccpvdz.cpf
Write coordinate data: done
```

```
## Time profile
```

```
Number of cores (total)   =    3072
Number of cores (fragment) =         1

THREADS (FRAGMENT)       =         48
```

```
Total time =          6135.0 seconds          1.7時間
```

FMO-MP3/cc-pVDZ

```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =          3885.7 seconds
Elapsed time: Monomer MP3      =           280.6 seconds
Elapsed time: Monomer (Total)  =          4211.3 seconds
Elapsed time: Dimer ES        =           296.1 seconds
Elapsed time: Dimer SCF       =           763.6 seconds
Elapsed time: Dimer MP3       =          5701.6 seconds
Elapsed time: Dimer (Total)   =          7907.9 seconds
Elapsed time: FMO (Total)     =         12119.1 seconds
```

```
*** Write check point file (CPF) open1.0 rev10 ***
```

```
WriteGeom = 6vxx_nonag-min0430_50k-mp3-ccpvdz.cpf
Write coordinate data: done
```

```
## Time profile
```

```
Number of cores (total)   =    3072
Number of cores (fragment) =         1

THREADS (FRAGMENT)       =         48
```

```
Total time =         12291.0 seconds          3.4時間
```

- ・ 総計では8ラックで3072プロセス×48スレッド、147,456コアを使用した超並列計算
- ・ 1ノード/フラグメント条件で実行、エネルギー成分分析(PIEDA)はオン、CPF書き込み
- ・ MP2ジョブではモノマーSCF(HF)段階がジョブ時間の半分以上を占める
- ・ FMO-MP2に比してFMO-MP3の相対コストは2倍で済む → スパコン利用のメリット
- ・ Dimer-ESは連続多重極展開で近似 → 未使用では計算は不可

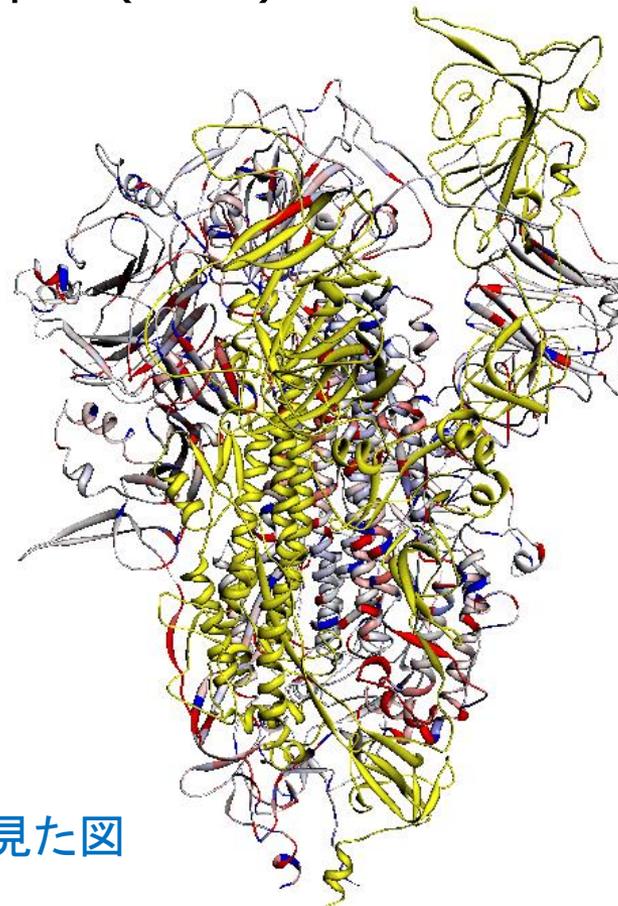
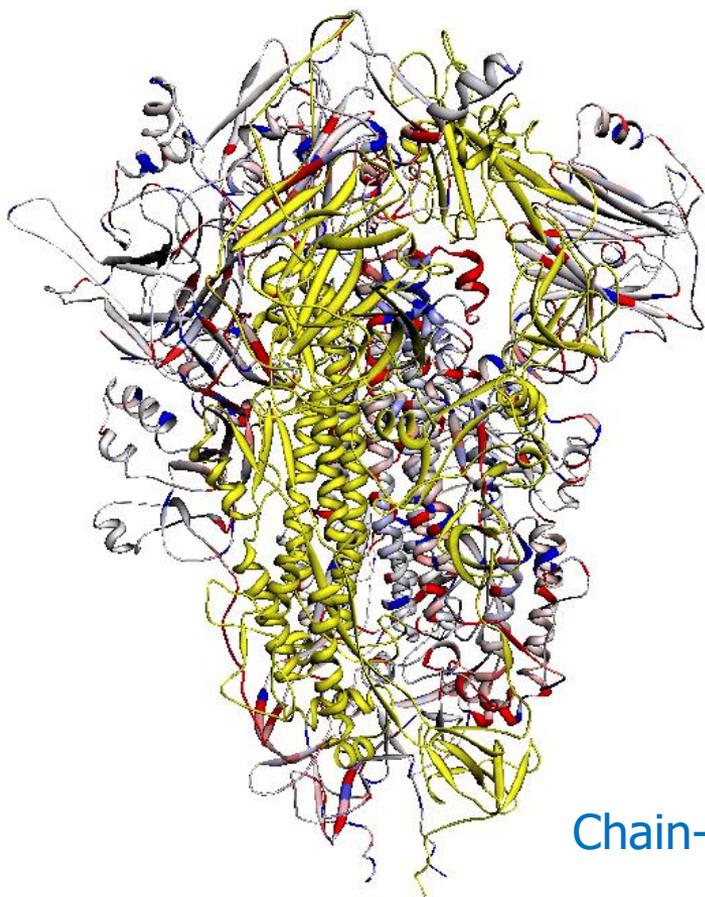
May 2020

スパイクタンパク質の相互作用エネルギーの可視化

MP2.5/cc-pVDZレベル

Closed (6VXX)

Open (6VYB)



Chain-Bから見た図

可視化はBioStation Viewerを64GBメモリのPCで利用して実行

Ver. 1 Rev. 22の「限界」から液滴モデルでは扱えなかった...

高速化と大規模系への対応 (Ver. 1系からVer. 2系へ)

「計算工学ナビ」での記載

https://www.cenav.org/abinit-mp-open_ver-1-rev-22/

https://www.cenav.org/abinit-mp-open_ver-2-rev-4/

https://www.cenav.org/abinit-mp-open_ver-2-rev-8/

2020年度の「富岳」の利用で認識したポイント

■ Ver. 1 Rev. 22からの改良の必要性

・高速化と大規模化

- ⇒ 多数のサンプル構造の扱い（Ver. 1 Rev. 22に比して数倍を目標）
- ⇒ 大型の液滴モデルの扱い（水を含めて数万フラグメントを視野）

・HPCI拠点のスパコン

- ⇒ 「富岳」を頂点とする富士通A64FX系
- ⇒ Intel Xeon、NEC SX-Aurora TSUBASA、AMD EPYCなども

・BioStation Viewerの可視化解析の「放棄」

- ⇒ 数千フラグメント系は「フツアのPC」では利用不可
- ⇒ 可視化用の配列を全削除（フラグメント数の自乗性で複数で存在）
- ⇒ 多数構造による統計的な相互作用解析が常態化
- ⇒ Pythonスクリプトやライブラリでの準自動処理（属人性の排除）

■ 新しい開発体制

・HPCの計算機科学の専門家とのコラボレーション（名大の片桐先生G）

- ⇒ 性能評価やチューニングの強化

・応用計算ではMDの専門家との共同作業（慶應大の泰岡先生G）

- ⇒ 液滴モデルの調製の迅速化

A64FX向けの高速度化

■2電子積分の生成

- ・ {SSSS, PSSS, SPSS, SSPP, SSSP, PPSS, PSPS, PSSP, SPPS, SPSP, SSPP, DSSS, SDSS, SSDS, SSSD}の積分タイプを(手動で)SIMD化
- ・ OCL指示詞の追加
- ・ コンパイラオプション: -O3 -Knosimd -Kocl
- ・ **改良の効果は20~30%程度** (Ver. 2 Rev. 4に反映)
- ・ レジスタスピル対策でループ分割 (Ver. 2 Rev. 8)
- ・ モノマー段階での積分のインコアバッファリングも実装 (Ver. 2 Rev. 8)

■Fock行列の構築

- ・ 添字同値性を $(1/2)n$ ($n=1,2,3$)で反映
- ・ **改良の効果は30%程度** (Ver. 2 Rev. 8)

■その他の改良

- ・ **通信量の削減**
- ・ 「無駄なプリント出力」の抑制オプション (logファイルが1千万行になることも...)
- ・ モノマーSCC反復の外挿法の改良

```
do p=ixi1,ixi2
  do q=ixj1,ixj2
    do r=ixk1,ixk2
      do s=ixl1,ixl2
        ix=ix+1
        val = sint(ix)
        if((abs(val) <= tv)) cycle
        fock(q,p)=fock(q,p)+dc(s,r)*val*2. d0! クーロン項
        fock(s,r)=fock(s,r)+dc(q,p)*val*2. d0
        fock(r,p)=fock(r,p)-dc(s,q)*val*0.5d0! 交換項
        fock(s,p)=fock(s,p)-dc(r,q)*val*0.5d0
        fock(r,q)=fock(r,q)-dc(s,p)*val*0.5d0
        fock(s,q)=fock(s,q)-dc(r,p)*val*0.5d0
      end do
    end do
  end do
end do
```

Ver. 2 Rev. 4での速度向上の一例

スパイクタンパク質

6VXX / FMO-MP2/6-31G* / Benchmark 8 racks @ Fugaku

Ver. 1 Rev. 22

```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =          2028.7 seconds
Elapsed time: Monomer MP2      =           15.0 seconds
Elapsed time: Monomer (Total)  =          2068.6 seconds
Elapsed time: Dimer ES        =           353.9 seconds
Elapsed time: Dimer SCF       =           362.4 seconds
Elapsed time: Dimer MP2       =           302.6 seconds
Elapsed time: Dimer (Total)   =          1603.4 seconds
Elapsed time: FMO (Total)     =          3672.1 seconds
```

Time profile

```
Number of cores (total)   =    3072
Number of cores (fragment) =         1

THREADS (FRAGMENT)       =         48
```

Total time = 3759.3 seconds

Ver. 2 Rev. 4

```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =          1801.6 seconds
Elapsed time: Monomer MP2      =           14.2 seconds
Elapsed time: Monomer (Total)  =          1839.1 seconds
Elapsed time: Dimer ES        =           314.2 seconds
Elapsed time: Dimer SCF       =           335.7 seconds
Elapsed time: Dimer MP2       =           294.6 seconds
Elapsed time: Dimer (Total)   =          1188.5 seconds
Elapsed time: FMO (Total)     =          3027.7 seconds
```

Time profile

```
Number of cores (total)   =    3072
Number of cores (fragment) =         1

THREADS (FRAGMENT)       =         48
```

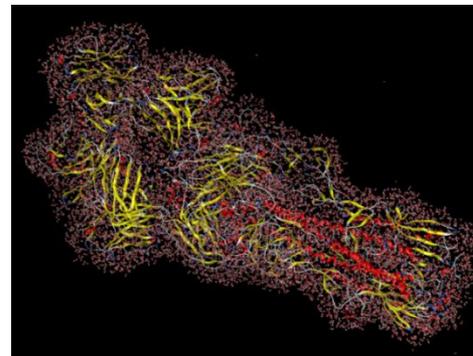
Total time = 3090.8 seconds

- Ver. 2 Rev. 4はA64FX向け積分のSIMD化、「不要配列」の整理などを反映済み
- 対Ver. 1 Rev.22で1.2倍の加速 (cc-pVDZ; 8769.9秒→6356.6秒で1.4倍)

Ver. 2 Rev. 4での超大規模系への対応例

従前の2倍の系が計算可能に

- ・インフルHA+Fab抗体×2 (PDB id: 1KEN) の水和モデル
- ・フラグメント総数は11307、水と対イオンを含む
- ・「不老」の1ラック、FMO-MP2/cc-pVDZは9.2時間で完走
- ・「富岳」の8ラック、FMO-MP3/cc-pVDZは6.7時間で完走
- ・FMO-MP2ではモノマーSCCが半分弱のコスト ⇒ 要対応
- ・水クラスターでは2万フラグメントのMP2ジョブも完走確認



```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =      14546.6 seconds
Elapsed time: Monomer MP2      =           32.5 seconds
Elapsed time: Monomer (Total)  =      14741.5 seconds
Elapsed time: Dimer ES         =       4021.8 seconds
Elapsed time: Dimer SCF       =       7215.9 seconds
Elapsed time: Dimer MP2       =       2492.4 seconds
Elapsed time: Dimer (Total)   =      18240.6 seconds
Elapsed time: FMO (Total)     =      32982.1 seconds
```

```
## Time profile
```

```
Number of cores (total)   =    384
Number of cores (fragment) =     1
```

```
THREADS (FRAGMENT)      =    48
```

```
Total time =          33120.9 seconds
```

```
=====
## TIME PROFILE
=====
```

```
Elapsed time: Monomer SCF      =       7114.0 seconds
Elapsed time: Monomer MP3      =       343.1 seconds
Elapsed time: Monomer (Total)  =      7532.4 seconds
Elapsed time: Dimer ES         =       534.8 seconds
Elapsed time: Dimer SCF       =       891.4 seconds
Elapsed time: Dimer MP3       =      4265.7 seconds
Elapsed time: Dimer (Total)   =     16306.3 seconds
Elapsed time: FMO (Total)     =     23838.7 seconds
```

```
## Time profile
```

```
Number of cores (total)   =    3072
Number of cores (fragment) =     1
```

```
THREADS (FRAGMENT)      =    48
```

```
Total time =          24203.2 seconds
```

Ver. 2 Rev. 8での高速化の例

HIV-protease + Lopinavir = 203 frag.

Timing of FMO-MP2/cc-pVDZ job (24 threads / 204 process).

「不老」Type I上

Timing in second	V1 R22	V2 R4	V2 R8
Elapsed time: Monomer SCF	864.9	617.8	387.4
Elapsed time: Monomer MP2	21.6	20.1	18.3
Elapsed time: Monomer (Total)	890.9	642.2	410.0
Elapsed time: Dimer ES	198.0	204.5	178.6
Elapsed time: Dimer SCF	421.8	344.0	260.8
Elapsed time: Dimer MP2	354.1	344.2	344.4
Elapsed time: Dimer (Total)	1205.6	1091.7	909.8
Elapsed time: FMO (Total)	2096.5	1733.9	1319.8
Total time	2101.4	1737.2	1323.2
Relative Acceleration	1.0	1.2	1.6

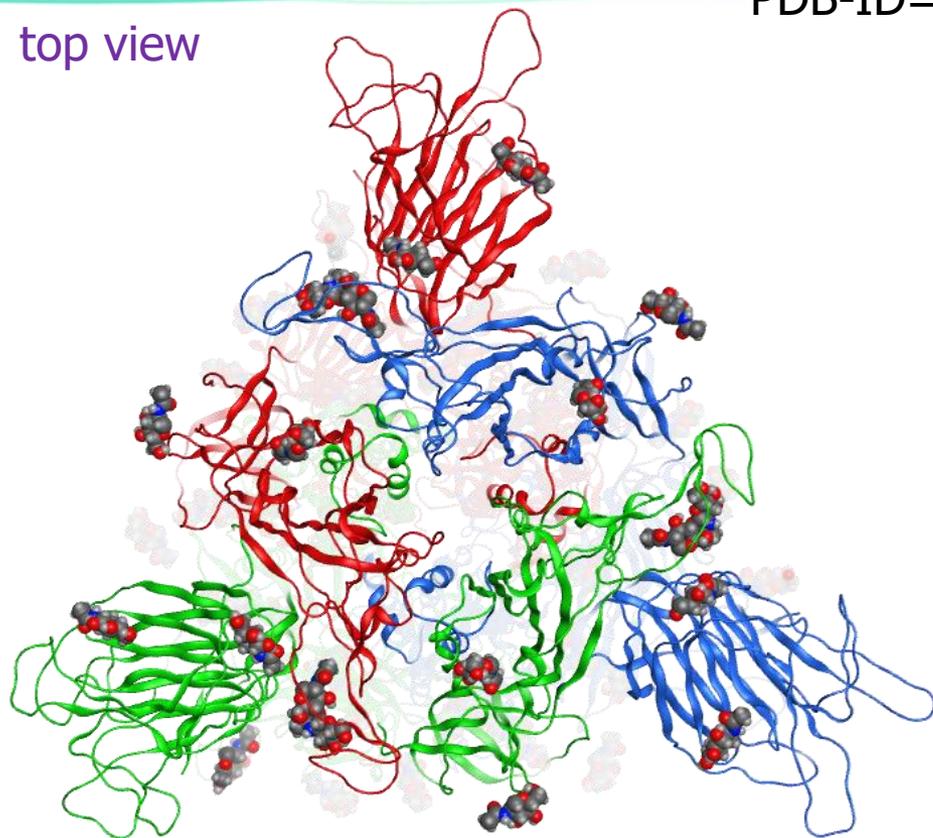
- ・ 積分の高速化が奏功（ダイマーSCFでも有効）
- ・ モノマーSCC段階での積分バッファリングも効果的
- ・ ダイマーES近似も若干の高速化

Ver. 2 Rev. 8は「富岳」では2024年1月から利用可能(Spack)

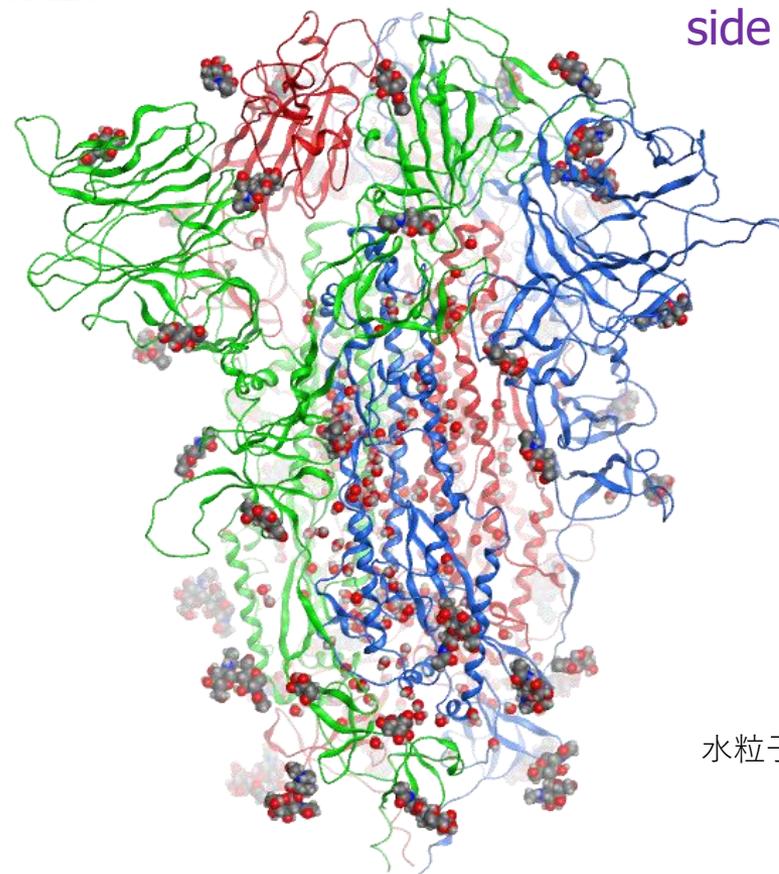
スパイクタンパク質の水とモデルの扱い

PDB-ID=6XLU

top view



side view

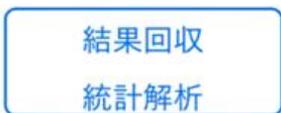
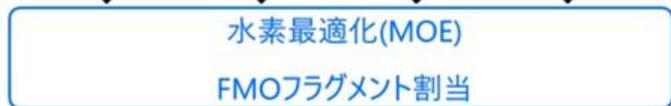
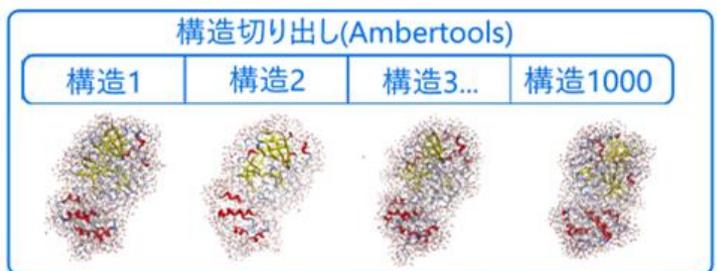


水粒子非表示

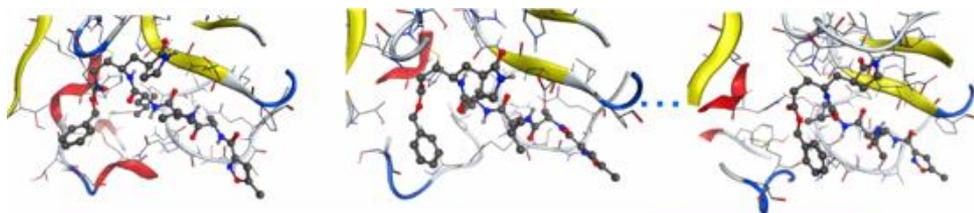
- ・液滴モデルの総数1.8万フラグメントを水のクラスタリングスクリプトで1万弱に低減
- ・糖鎖部分も考慮、Glnとの結合もスクリプトで切断
- ・FMO-MP2/6-31G*ジョブ、8ラック(3072ノード)で約2時間で完走
- ・現時点で世界最大規模のFMO計算、100サンプル構造は済み

MD軌跡からの多構造サンプルの一括処理

「富岳」利用 {hp210026, hp220025, hp230017}; “新規感染症のための計算科学的解析環境の整備”



- * MD由来の多サンプルの液滴構造の扱い
- * 新型コロナウイルスのRBD-ACE2系
- * インフルエンザウイルスのHA(3)-Fab系
- * 処理は準自動的、解析もスクリプト処理



001100.log 001200.log ... 101000.log

TIME(ps)	TYR161(161)	MET162(162)	HIS163(163)	HIS164(164)
1100	0.178	-0.252	-7.358	-0.350
1200	0.190	-0.269	-5.538	-1.361
1300	0.092	-0.240	-7.403	-0.914
1400	0.164	-0.267	-5.661	-1.498
1500	0.176	-0.183	-5.502	-0.603
1600	0.176	-0.208	-7.010	-0.442
		⋮		
100900	0.046	-0.197	-5.061	-0.713
101000	0.167	-0.277	-5.315	-1.042
101100	0.184	-0.144	-6.252	-0.306

リリース状況のまとめ(2023年秋時点)

◇Ver. 1系の最終リリース版はRev. 22 (2020年6月)

f関数のサポート (直交10成分のみ)

MFMOによる指定領域のみでのMP2&MP3計算

加速されたPIEDA (「富岳」の新型コロナPJの中で改良)

◇「富岳」時代でVer. 1からVer. 2へ移行 - Open Ver. 2 Rev. 4 (2021年9月)

単一構造サンプルでの可視化的解析の「限界」

統計的な相互作用解析の重要性 (MDベースで百~千サンプル構造を対象)

大規模系の扱いの必要性 (タンパク質の液滴モデルは数万フラグメントに)

可視化ツールによる解析を「諦め」(ダンプ用データがメモリを圧迫)

HPC分野の専門家とコラボしての高速化 (特に「富岳」などのA64FX向け)

基本のMP2計算ではVer. 1 Rev. 22比で**1.2~1.4倍**の加速

◇最新リリース版 - Open Ver. 2 Rev. 8 (2023年8月)

PIEDA機能の強化 (分散力寄与の区別、静電相互作用のRESP評価)

MFMOでの励起エネルギーとイオン化エネルギーの算定

MP2計算はVer. 1 Rev. 22比で**1.5~2倍**の加速

2万フラグメントの液滴モデルの扱い (「富岳」では多サンプル処理が可能)

HPCI拠点でのライブラリ整備(2024年1月時点)

赤:A64FX, 緑: SX-Aurora TSUBASA, 紫: Xeon

■登録サイトと版

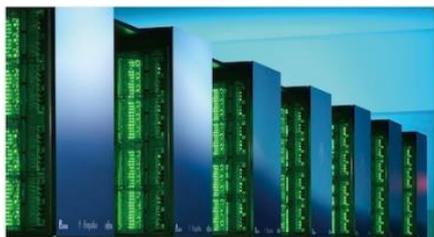
- ・北大「**Grand Chariot**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4
- ・東北大「**AOBA-A/S**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4 (Vectorized version)
- ・東大「**Wisteria / Odyssey & Aquarius**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4
- ・分子研「**RCCS**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4
- ・名大「不老 / Type I & II」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4 & Ver. 2 Rev. 8
- ・京大「**Camphor**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 8
- ・阪大「**SQUID**」 : Ver. 1 Rev. 22 (Vectorized version)
- ・R-CCS「**富岳**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4 & Ver. 2 Rev. 8
- ・計算科学振興財団「**FOCUSスパコン**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4
- ・九大「**ITO Subsystem-A**」 : Ver. 1 Rev. 22 & Ver. 2 Rev. 4 (to be replaced)

- ・東工大「**TSUBAME4.0**」 : 2024 1Q (scheduled)

- ・ Ver. 1 Rev. 22はBioStation Viewerとの関係で当面は併存
- ・ Ver. 2 Rev. 8も2024年度内に随時整備の予定
- ・ HPCI拠点のスパコン機種更新に応じた対応も適宜行う方針

「富岳」のOnDemand(Beta)に登録済

https://www.hpci-office.jp/for_users/appli_software#openondemand



Welcome to the supercomputer Fugaku



OnDemand provides an integrated, single access point for all of your HPC resources.

Pinned Apps A featured subset of all available apps

Batch Jobs

 ABINIT-MP	 FDS	 FrontISTR	 GROMACS
 LAMMPS	 MODYLAS	 OpenFOAM	 OpenMX
			

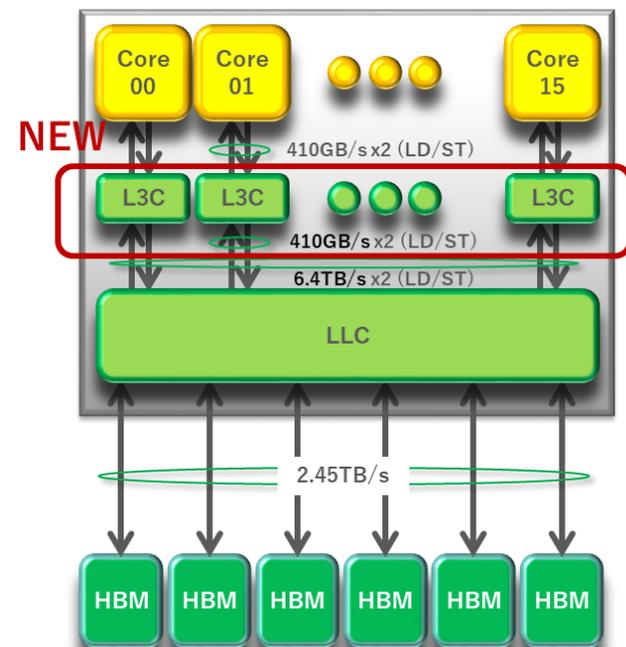
SX-Aurora TSUBASAへの対応 (NECとのコラボレーション)

SX-Aurora TSUBASA (SX-AT)

東北大 AOBA-A

東北大 AOBA-S

	VE20B		VE30A
CORE/CPU	8	x2.0	16
CORE Performance	~307GF(DP) ~614GF(SP)		~307GF(DP) ~614GF(SP)
CPU Performance	~2.45TF(DP) ~4.91TF(SP)	x2.0	~4.91TF(DP) ~9.83TF(SP)
Level3 Cache Capacity	-	NEW	2MB/Core SW Controllable
Last Level Cache capacity	16MB Shared	x4	64MB Shared
MEMORY bandwidth	1.53TB/s	x1.6	2.45TB/s
MEMORY capacity	48GB	x2	96GB
PCIe	Gen3	x2	Gen4



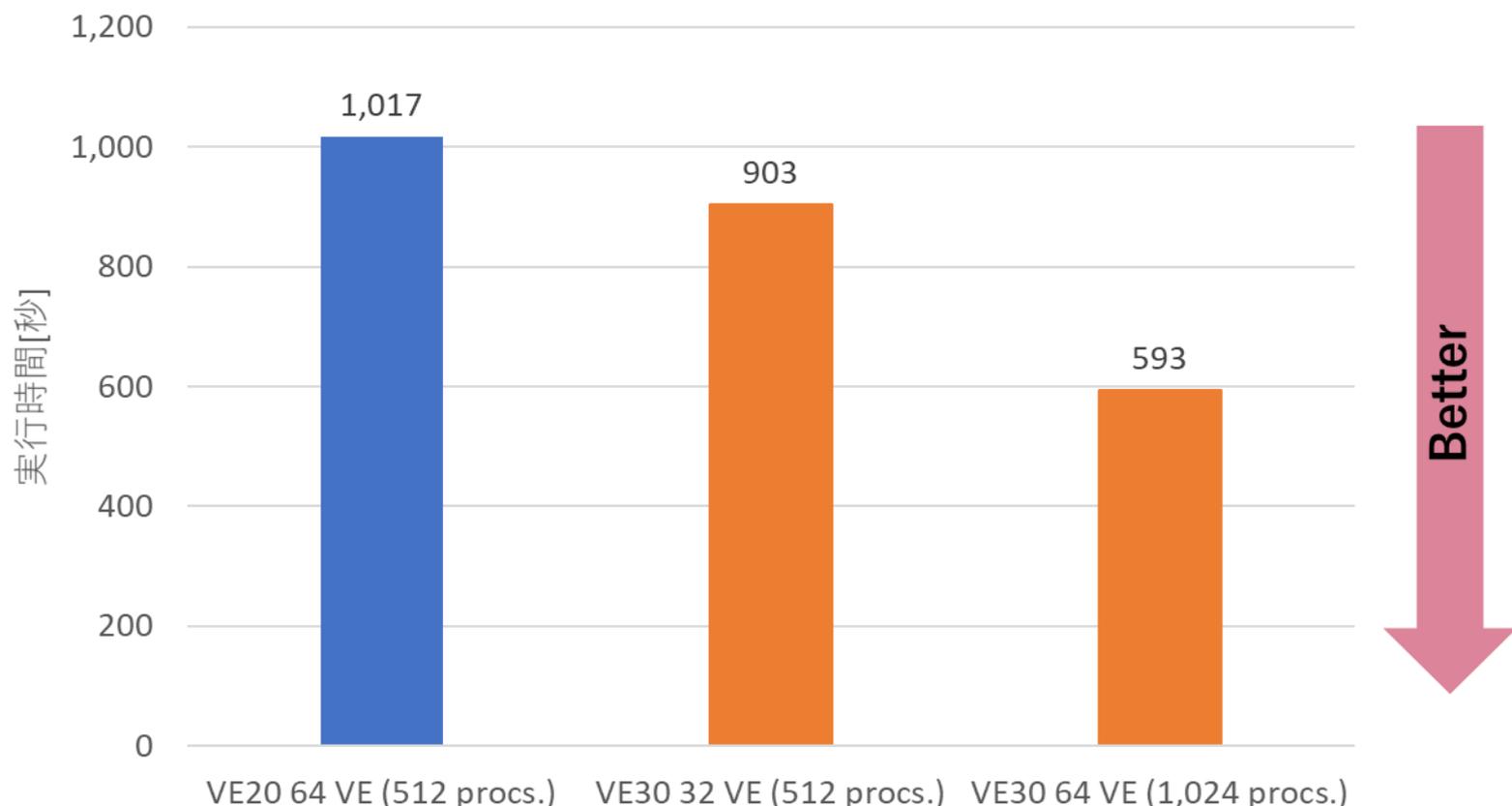
- ・ VE30でコア数は8から16に倍増
- ・ メモリも48GBから96GBに倍化
- ・ メモリバンド幅は1.6倍
- ・ L3キャッシュに新規増設も改善

Orchestrating a brighter world. NEC

SX-ATでのMP2の評価

東北大 AOBA-S

6LU7のFMO-MP2/6-31G*ジョブ / Ver. 2 Rev. 4のベクトル化対応版



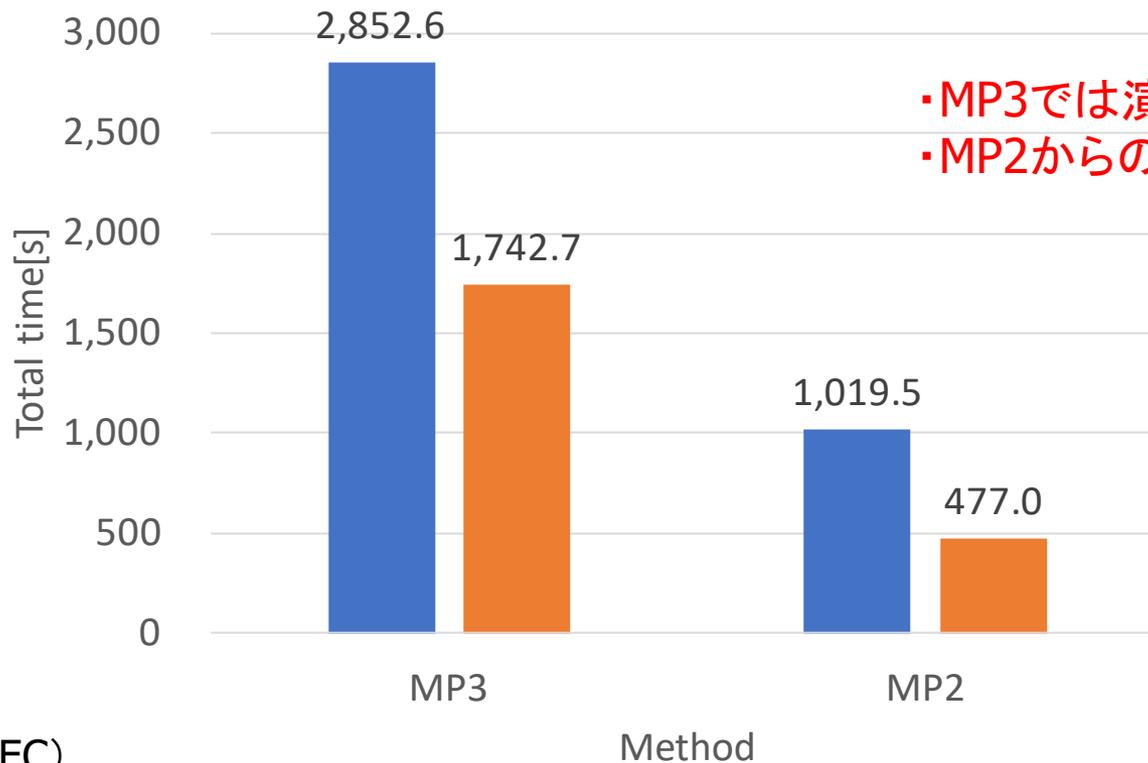
NP=4条件

計測は加藤氏 (NEC)

- ・ 同一VE数ではVE30はVE20の1.7倍高速
- ・ 同一プロセス数では1.1倍の改善
- ・ L3キャッシュの効果
- ・ Ver. 2 Rev. 4のベクトル化版を導入予定

SX-ATでのMP2とMP3の比較

HW	Method	Memory	Total time[s]	GFLOPS	演算量[TFLOP]	メモリ[GB/VE]
VE20B 64 VE	MP3	40000	2,852.6	12,307.4	35,108.0	46.8
	MP2	40000	1,019.5	6,266.1	6,388.3	46.3
VE30A 64 VE	MP3	40000	1,742.7	20,169.6	35,149.6	92.3
	MP2	40000	477.0	13,478.2	6,429.1	91.4



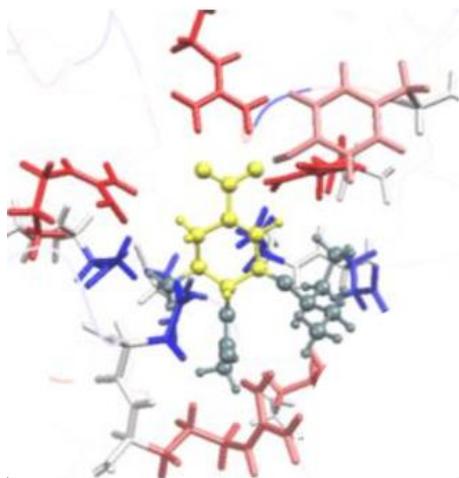
- ・MP3では演算量が約6倍
- ・MP2からのコスト増は約3倍

計測は加藤氏(NEC)

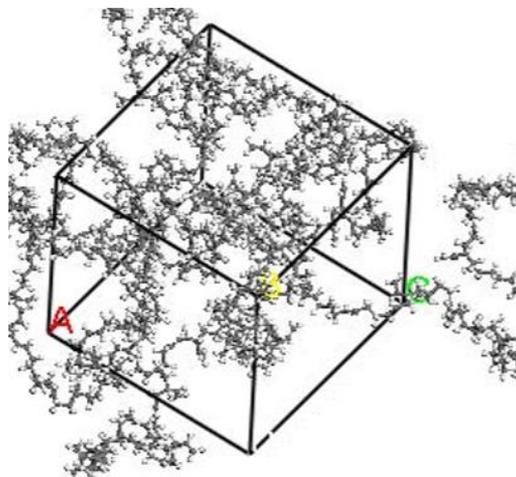
粗視化シミュレーションとの接続 (JSOLとのコラボレーション)

各スケールの分子シミュレーション

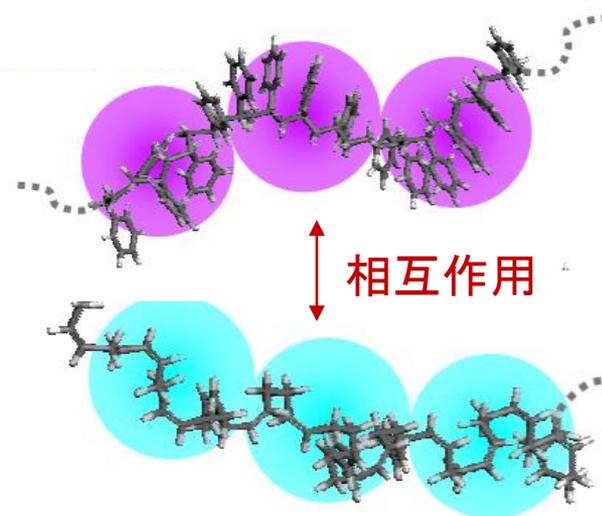
◇量子化学 (FMO)



◇分子動力学



◇粗視化シミュレーション



[電子状態の計算]

- ・タンパク質とリガンドの相互作用
- ・結晶中の分子間相互作用

Å単位

[原子単位の運動]

- ・結晶のパッキング
- ・弾性率(応力-ひずみ)
- ・ガラス転移点の挙動

数nm単位

[原子集団単位の運動]

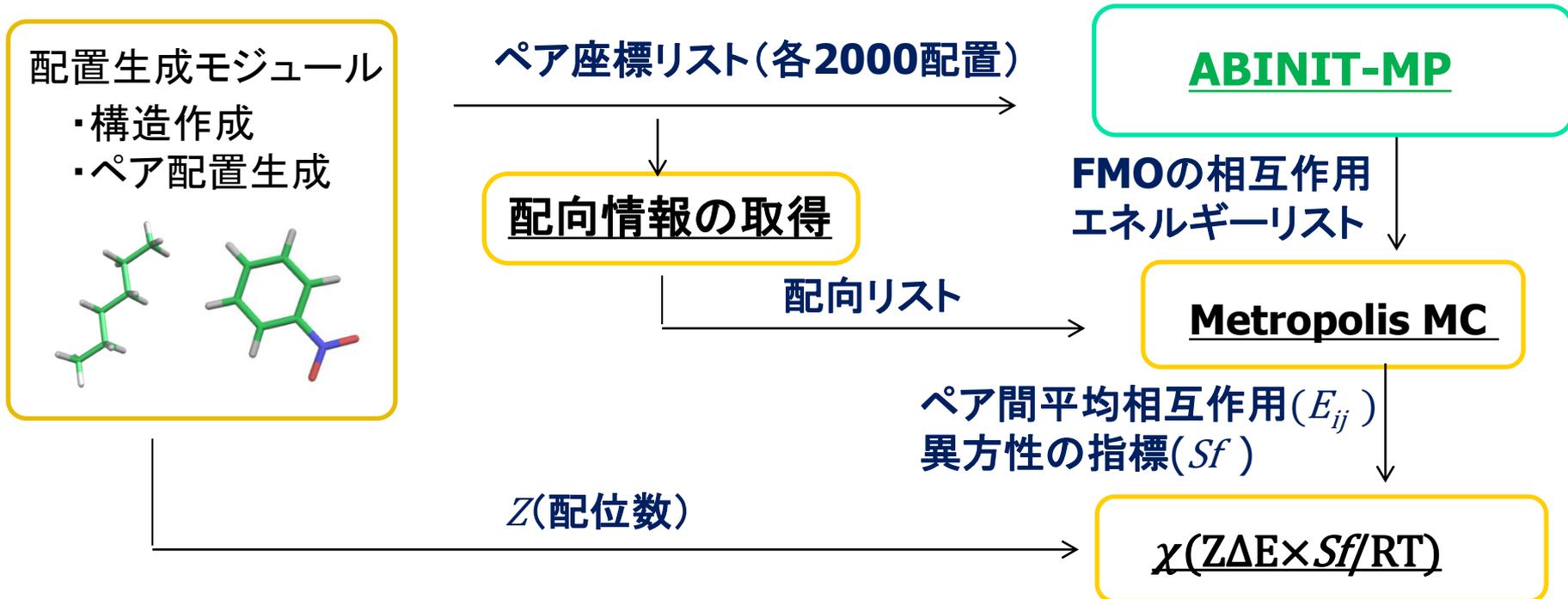
- ・長時間の運動
- ・リポソームの融合
- ・ポリマーの相分離

数10~100nm単位

- ・粗視化粒子間の有効相互作用パラメータをFMOを使って算出
- ・そのパラメータで散逸粒子動力学シミュレーションを行う

FMO計算に基づくパラメータ算定のワークフロー

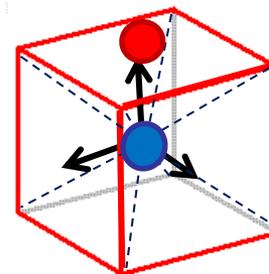
◆配置生成からパラメータ算出処理までの自動化



◆異方性の指標(Sf)の導入:

- ・特定の配向で強い相互作用を示す系: 平均相互作用を過大評価
- ・MC法採用配置の配向の偏りから相互作用をスケールリング

⇒単純な二成分系の相転移臨界温度が実験と良好な一致

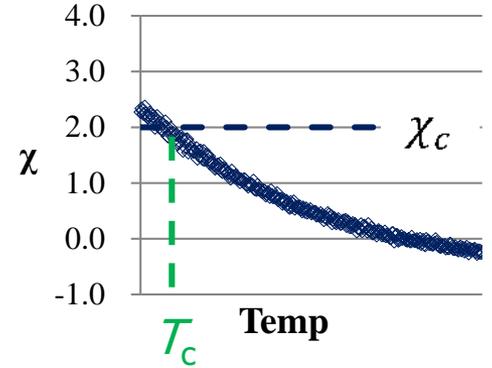


ワークフローを汎用システム化してリリース (名称はFCEWS)

FMOによる算定値の検証

◇温度上昇： χ が減少（一般的に）→相溶しやすく
 臨界点（相が分離しなくなる閾値）の温度 T_c ：実験と比較

臨界点 χ (χ_c) $\chi_c = \frac{(\frac{1}{\sqrt{n_a}} + \frac{1}{\sqrt{n_b}})^2}{2}$ $n_a, n_b =$ 各成分の重合度
 重合度： $\frac{\text{ポリマー分子量 (Mw)}}{\text{セグメント分子量}}$



Hexane - Nitrobenzene

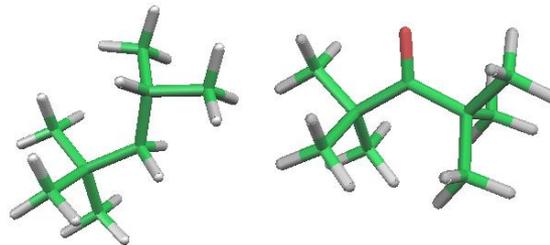
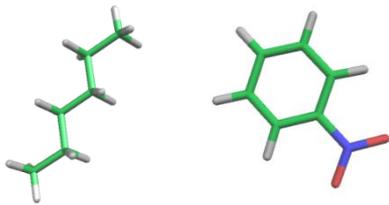
Diisobutyl ketone - Polyisobutylene

Polyisoprene - Polystyrene

$T_c(K)$		
χ_c	exptl.	FMO
2.0	293	286

$T_c(K)$				
Mw	χ_c	exptl.	FMO	
22700	0.57	292	328	
285000	0.52	319	346	
6000000	0.50	329	354	

Mw		$T_c(K)$		
pip	ps	χ_c	exptl.	FMO
1000	1000	0.34	243	255
2000	2700	※0.15	329	420
2700	2100	0.15	408	420
2700	2700	0.12	448	489

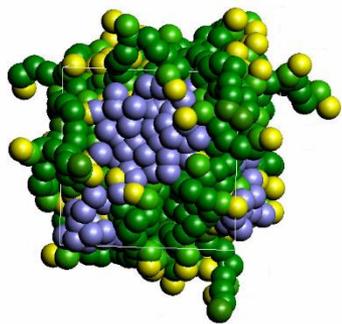


実験値を10%程の誤差で再現

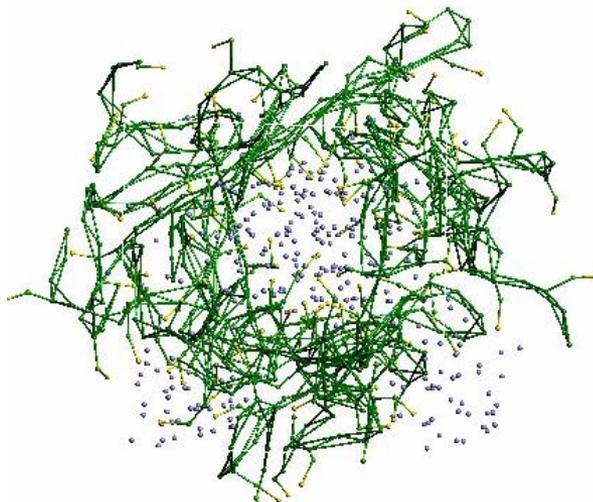
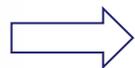
※理論値が実験値と非対応

リバースマッピング

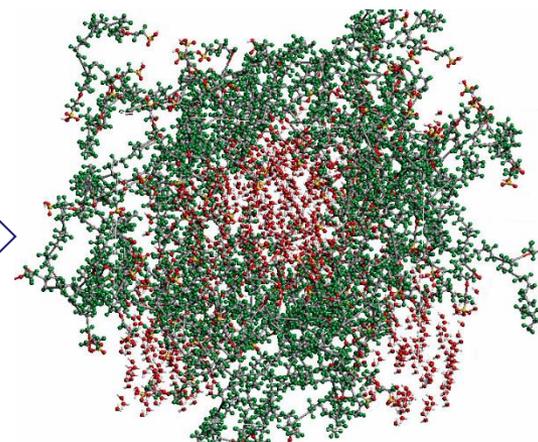
◇DPDの粗視化構造を再原子化し、FMO計算を行うスキームを確立（DSRMS）



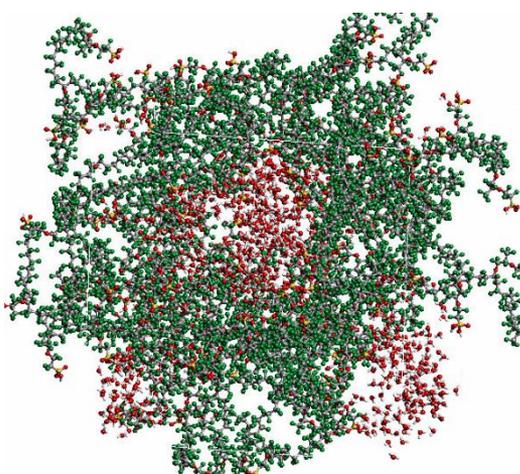
DPD結果
(約1000粒子)



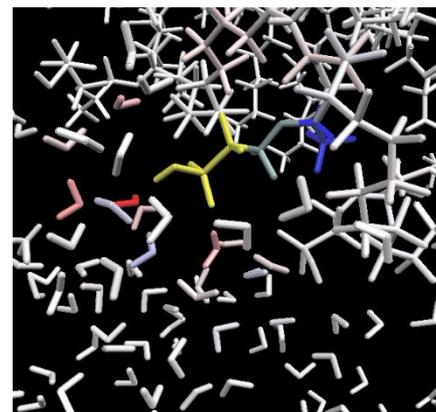
アフィン変換(体積拡大)



原子化
(約10000原子)



MDによる緩和(NPT)

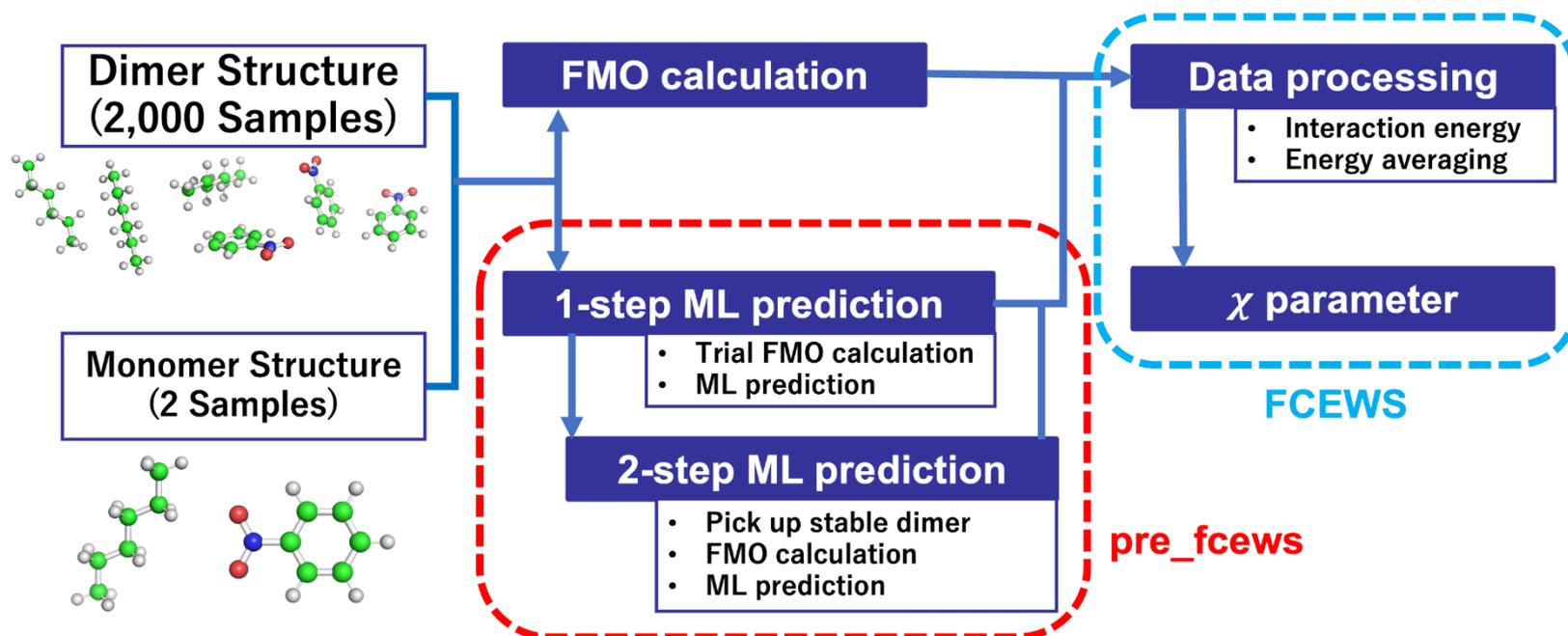


FMO(IFIE解析)

機械学習による χ パラメータ算定の効率化

pre_fcwvs

- ・FCEWSを用いた χ パラメータ算定において、FMO計算の実行回数を減らすべく作成されたワークフローシステム
- ・2,000構造のうちの一部のFMO計算の結果を教師データとし、その他の構造の相互作用エネルギーの予測を行う
⇒ 得られたエネルギー予測値を利用して χ の算定を行う



χ パラメータ算定に「富岳」の計算力を活用

■ 圧倒的な計算力

- ・ セグメント対あたりで2千ショットのFMO計算も容易に
- ・ 多成分系の算定を小規模サーバで行うと数週間～月の時間に
- ・ capacity computing文脈で大量ジョブをFCEWSで自動投入
- ・ 中規模系なら(待ち行列が短ければ)数日以内に完了、複数対象も平行処理

■ 産学連携のhp210261-「富岳」機動的(一般)課題(2021年11月から1年間)

- ・ FCEWS開発元の立教大が主導/サポート、JSOLが幹事
- ・ 東レ、日本触媒、東洋紡、ブリヂストンが参加
- ・ 化学工学系企業でのFCEWSの知名度向上を期待
- ・ 月に一回程度Zoomで会合 (オンラインのメリットを実感)
- ・ 前半で慣れ/助走段階は終了、後半で各社が個別テーマに取り組み
- ・ FMOで求めたパラメータを使った DPDシミュレーションも推進

■ 産学連携のhp230016-「富岳」産業課題(2023年度)

- ・ コミュニティの維持、慶應義塾大学の泰岡先生-荒井先生らが参画
- ・ 各社のテーマを推進、立教大はタンパク質用のパラメータ算定を実施
- ・ 後継のhp240013課題が採択済

GPU対応

GPU対応の必要性と対応

■背景

- ・GPUスパコンの遍在化
 - ⇒ 「Wisteria」 Aquarius&Mercury、「不老」 Type II、Oakforest PACS-II...
 - ⇒ 使われている分野はCG、AI、量子シミュレーション、MD...
- ・ポスト「富岳」
 - ⇒ 「アクセラレータ」の搭載は(ほぼ)確定
- ・MO系の計算もGPUに対応
 - ⇒ TeraChemは高名、GAMESS-USも対応 (GPU化FMO専用コードも在り)

■対応策

- ・ABINIT-MPも2電子積分生成などのGPU化が必要
 - ⇒ FMO-HF計算をまず改良
 - ⇒ MP2計算はDGEMM化で相対的には対応は容易と想定(?!)
- ・ABINIT-MP本体は巨大に過ぎる
 - ⇒ HF計算のみを切り出したミニアプリを作成
 - ⇒ NVIDIAの成瀬彰氏のチューニングで6-8倍の加速を確認
- ・ABINIT-MPのGPU化を検討中
 - ⇒ 積分ルーチン群の再構成(81種→21種へ低減)から着手

ABINIT-MPのミニアプリ

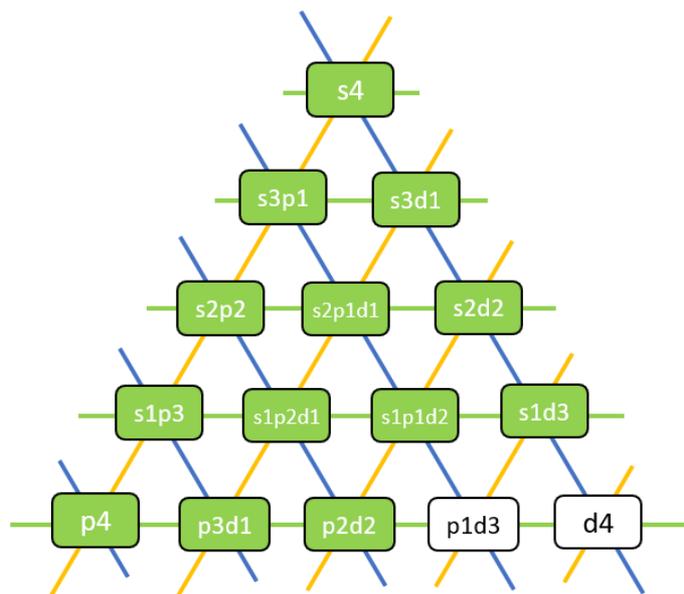
2022年度下期の東大情基セの支援PJに参加

- ・ FMOなしでHF計算のみを実行
- ・ 2電子積分の性能評価とチューニング用
- ・ ネームリストで積分タイプを使い分け可能
 - ObaraはABINIT-MPと同じVRR法で軌道の組み合わせ毎のサブルーチン
 - Obara-generalは汎用ループでの計算、最適化は特にされていない
 - HGP-generalはHRR法での汎用ループでの計算、最適化は未
 - Obara-vectorはNEC SX向けにループ構造をベクトル計算用に変更してある
- ・ Obara-vectorは、GPUでの処理に向く?!
- ・ 演算数的にはHRRにも可能性あり
- ・ コンパクトなルーチンで性能が出れば最良...
- ・ 4タイプでの評価をお願いできれば...
 - 筋のよさそうなものをチューニング
 - 基底関数は6-31G*が基本

```
&CNTRL
  ReadGeom='adamantane.pdb'
  eri_type=0 ! Obara
! eri_type=1 ! Obara-general
! eri_type=2 ! HGP-general
! eri_type=3 ! Obara-vector
! memory=4000
/
&SCF
!maxscfcyc=10
/
&BASIS
  BasisSet='6-31G*'
! BasisSet='6-31G'
! BasisSet='cc-pVDZ*'
/
```

ミニアプリのチューニング (NVIDIA成瀬彰氏)

- ◇ 先ず、ABINIT-MPの通常の積分生成と同じ、**小原・雑賀のVRR式**のルーチン群を対象に
 - 元は{x, y, z, w}={s, p, d}関数の組み合わせからsub_xyzwは81種類 (自動生成)
- ◇ 多数の sub_xyzw をバッチ化して、GPU の多並列性を活用する
 - 以下の理由から単純に全部をバッチ化すればいいわけではない
 - ・ sub_ssss: 計算手順が簡単、一時配列数が少なく、そのサイズは小さい (レジスタ使用量: 少): データ並列性が高い (IJKL_ex が大きめ)
 - ・ sub_dddd: 計算手順が複雑、多数の一時配列が必要で、そのサイズは大きい (レジスタ使用量: 多): データ並列性が低い (IJKL_ex が小さい)
 - 計算特性の近いものをまとめる (**構造の改変**)、**OpenACC指示文**の挿入
 - **Fock行列への加算**も、バッチ化積分ルーチンの内部で実施 (メモリアクセス量削減)



- ・ 緑色が GPU 化の対象
- ・ "s3d1"等は4通りを含む
- ・ "s1d3"、"p3d1"、"p2d2"等はレジスタ要求量が多く、レジスタスピル発生
- ・ "p1d3"、"d4" は CPU 実行

Phe-TrpとTrpでのミニアプリのテスト(HF)

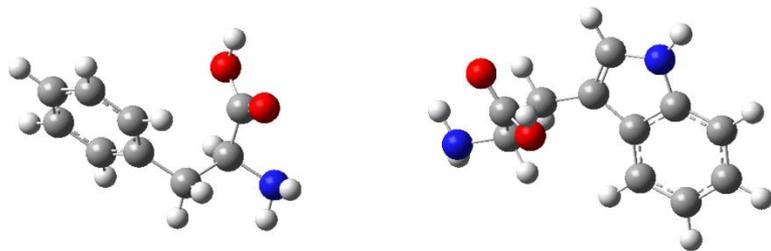
東大Wisteria-Aquariusを使用

- 下記の条件で再度実行(flat MPI)

export NNODES=1

export NGPUS=1

export NPROCS=8



- Trp-Pheのテスト

6-31G (#Orb = 289, #SCF = 17) ; 27.9 s / 4.3 s → Acc = 6.5

6-31G* (#Orb = 451, #SCF = 18) ; 69.4 s / 10.8 s → Acc = 6.4

cc-pVDZ (#Orb = 520, #SCF = 19) ; 179.0 s / 21.8 s → Acc = 8.2

6-311G** (#Orb = 651, #SCF = 21) ; 257.4 s / 37.7 s → Acc = 6.8

- Trpのテスト

6-31G (#Orb = 159, #SCF = 17) ; 8.4 s / 1.8 s → Acc = 4.7

6-31G* (#Orb = 249, #SCF = 18) ; 21.9 s / 3.6 s → Acc = 6.1

cc-pVDZ (#Orb = 285, #SCF = 17) ; 51.4 s / 6.2 s → Acc = 8.3

6-311G** (#Orb = 357, #SCF = 21) ; 76.4 s / 10.5 s → Acc = 7.3

- GPUは「加速装置」としての価値アリ

量子計算との関係 (シミュレータ利用)

cuQuantumシミュレータの利用例

<https://iopscience.iop.org/article/10.1088/2516-1075/acf909>

Electronic Structure



PAPER

OPEN ACCESS

RECEIVED
14 April 2023

REVISED
8 August 2023

ACCEPTED FOR PUBLICATION

Bayesian phase difference estimation algorithm for direct calculation of fine structure splitting: accelerated simulation of relativistic and quantum many-body effects

Kenji Sugisaki^{1,2,3,*}, V S Prasanna³, Satoshi Ohshima⁴, Takahiro Katagiri⁵, Yuji Mochizuki^{6,7}, B K Sahoo⁸ and B P Das^{3,9,*}

Abstract

Despite rapid progress in the development of quantum algorithms in quantum computing as well as numerical simulation methods in classical computing for atomic and molecular applications, no systematic and comprehensive electronic structure study of atomic systems that covers almost all of the elements in the periodic table using a single quantum algorithm has been reported. In this work, we address this gap by implementing the recently-proposed quantum algorithm, the Bayesian phase difference estimation (BPDE) approach, to determine fine structure splittings of a wide range of boron-like atomic systems. Since accurate estimate of fine structure splittings strongly depend on the relativistic as well as quantum many-body effects, our study can test the potential of the BPDE approach to produce results close to the experimental values. Our numerical simulations reveal that the BPDE algorithm, in the Dirac–Coulomb–Breit framework, can predict fine structure splittings of ground states of the considered systems quite precisely. We performed our simulations of relativistic and electron correlation effects on Graphics Processing Unit by utilizing NVIDIA's cuQuantum, and observe a ×42.7 speedup as compared to the Central Processing Unit-only simulations in an 18-qubit active space.

- ・ 慶應大の杉崎研司先生が主導の研究例
- ・ cuQuantumを「不老」Type IIで使用
- ・ 片桐先生と大島先生も共著
- ・ Bの等電子系の $2p_{(1/2)}-2p_{(3/2)}$ エネルギー

- ・ 相対論的な計算(Lr⁹⁸⁺まで)
- ・ 計算値の精度は十分
- ・ GPUによる加速は42.7倍

Quantum Physics

[Submitted on 28 Feb 2024]

Unitary coupled-cluster singles and doubles (UCCSD) calculations in conjunction with fragment molecular orbital (FMO) scheme

Kenji Sugisaki, Tatsuya Nakano, Yuji Mochizuki

The fragment molecular orbital (FMO) method is one of the popular methods to efficiently treat macromolecular systems by dividing the system of interest into small fragments based on embedding under the electrostatic potential (ESP). Such a fragmentation method has the potential advantage of making the circuit flat in quantum chemical calculations on quantum computers. In this study, we used a GPU-accelerated quantum simulator (cuQuantum) to perform the electron correlation part of the FMO calculation as a unitary coupled-cluster singles and doubles (UCCSD) with the variational quantum eigensolver (VQE), using hydrogen-bonded systems $(\text{FH})_3$ and $(\text{FH})_2\text{-H}_2\text{O}$ as testbeds. From the numerical simulations, we found that the Trotter decomposition of the UCCSD ansatz breaks the orbital-invariance condition, and it can yield different total energies for symmetrically equivalent molecules. We also observed that the Trotterized UCCSD does not satisfy the size-consistency, which is an essential requirement in the FMO scheme, when the molecular orbitals delocalized to dimers are used. The GPU acceleration was substantial for the simulations with larger numbers of qubits, and it was about a factor of 5.8--7.5 for 18 qubit systems.

Comments: 15 pages, 2 figures, 5 tables

Subjects: **Quantum Physics (quant-ph)**; Chemical Physics (physics.chem-ph)

Cite as: [arXiv:2402.17993](https://arxiv.org/abs/2402.17993) [quant-ph]

(or [arXiv:2402.17993v1](https://arxiv.org/abs/2402.17993v1) [quant-ph] for this version)

Submission history

From: Kenji Sugisaki [[view email](#)]

[v1] Wed, 28 Feb 2024 02:16:14 UTC (1,087 KB)

- 相関計算部分をUCCSD計算で置き換え
- cuQuantumシミュレータでのGPU加速
- 実は通常のFMO-CCSD(T)より数万倍遅い...
- size-consistencyやorbital-invarianceにも注意
- 基礎研究の位置づけ(jh230001課題)
- Ver. 2を準備中

FMO計算の試行#2

TABLE II. Correlation energies of $(\text{FH})_3^a$ in units of Hartree.

Unit	MP2	CCSD	CCSD(T)	UCCSD:CB	UCCSD:PW	CAS-CI
Monomer						
“1”	-0.017933	-0.026945	-0.026945	-0.026884	-0.026914	-0.026945
“2”	-0.017526	-0.026216	-0.026216	-0.026164	-0.026192	-0.026216
“3”	-0.017933	-0.026929	-0.026929	-0.026899	-0.026875	-0.026929
Dimer						
“21”	-0.035493	-0.051856	-0.051933	-0.049880	-0.050452	-0.051963
“31”	-0.035980	-0.052778	-0.052850	-0.051554	-0.051527	-0.052879
“32”	-0.035446	-0.053122	-0.053123	-0.051952	-0.053067	-0.053124
Sum.	-0.053527	-0.077666	-0.077816	-0.073439	-0.075065	-0.077876

^a The HF energies (in units of Hartree) of monomer “1”, “2”, and “3” are -103.815720, -103.995064, -103.563842, respectively. In contrast, the HF energies of dimer “21”, “31”, “32” are -228.173251, -227.542281, -218.792929, respectively. The sum of Eq. (1) is -363.133835 Hartree. Eq. (1) is also used for the sum of correlation energies.

TABLE IV. Timings of UCCSD job (in second) of $(\text{FH})_3$ with/without GPU.^a

Unit	UCCSD:CB			UCCSD:PW		
	with GPU	without GPU	Acceleration	with GPU	without GPU	Acceleration
Monomer						
“1”	14.1	30.1	2.13	25.4	52.3	2.06
“2”	16.1	29.1	1.80	27.9	68.9	2.47
“3”	14.8	31.5	2.13	22.2	48.4	2.18
Dimer						
“21”	11901.3	70531.8	5.93	31276.4	216982.7	6.94
“31”	12871.9	74618.2	5.80	30315.4	210808.9	6.95
“32”	11209.1	83915.6	7.49	27740.4	203298.3	7.33

^a All the calculations were carried out on ‘Flow’ Type-II subsystem.

おわりに

まとめ/今後

■まとめ/現状

- ・Ver. 2 Rev. 8までリリース
 - ⇒ 従前のVer. 1 Rev. 22に比してFMO-MP2計算で1.5倍以上の高速化
 - ⇒ 2万フラグメントのタンパク質の液滴モデルの扱いが可能
- ・FMO計算の広がり
 - ⇒ 粗視化シミュレーションのパラメータ算定、量子シミュレータの試行的利用

■今後(ネクスト「富岳」も意識)

- ・実効数で4万フラグメントの扱いを目標
 - ⇒ フラグメント情報の分散保持
 - ⇒ 入力データ作成の支援と結果の集計と解析のためのスクリプト類の整備
- ・GPU対応
 - ⇒ FMO-MP2まで (2025年度上期を目標)
- ・開殻系の扱い
 - ⇒ ROHF/RMP2の実装 (スピン汚染の回避)
- ・機械学習の導入
 - ⇒ サロゲート化 (計算無し/低コスト計算から高コスト計算の結果を予測)

ABINIT-MPによるFMO計算のロードマップ

