

Title	フラグメント分子軌道法による量子生命情報基盤の構築：タンパク質基本フォールドと生体分子動的挙動の解析
Author(s)	福澤, 薫
Citation	サイバーメディアHPCジャーナル. 2024, 14, p. 109-112
Version Type	VoR
URL	<a href="https://doi.org/10.18910/96537">https://doi.org/10.18910/96537</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

# フラグメント分子軌道法による量子生命情報基盤の構築

## ～タンパク質基本フォールドと生体分子動的挙動の解析～

福澤 薫

大阪大学 大学院薬学研究科

### 1. はじめに

日本発の理論手法であるフラグメント分子軌道(FMO)法は、タンパク質や核酸などの生体高分子の全電子計算を高速に実施できる世界最先端の手法である。さらに生体内の分子認識の要である分子内・分子間相互作用を量子化学的に解析できるため、構造生物学者や創薬化学者の化学的直観にも沿う理論解析手法として国際的にも高く評価されている [1, 2]。我々は、FMO法を実用的なインシリコ創薬技術として発展させ、今までにない高精度な論理的創薬を実現させることを目指して、2015年度から現在に至るまで、産学官連携の FMO 創薬コンソーシアムを母体にして、HPCI を活用した研究開発を行っている。また、その成果を用いて FMO 創薬のための情報基盤を構築するために、世界初のタンパク質の量子化学計算データベースである「FMO データベース (FMO DB)」を開発し、2019 年から一般公開している [3, 4]。FMO DB では、構造生物学者や創薬化学者が利用しやすいように、web インターフェイスを整え、現在までに約 37,000 構造の FMO 計算結果を公開している。

FMO DB が世界最大のタンパク質量子化学計算データを保持している一方で、世界中の研究者が利活用している生体高分子の構造データベースである Protein Data Bank (PDB) では、本申請時点で約 22 万エントリーの構造が公開されている。FMO DB で公開されている FMO データのうち、独立した PDB 番号(PDBID)を持つものは約 7,700 構造であり、実験的に構造解析されている多様なタンパク質の計算を網羅しているとは言い難い。そこで本研究では、PDB の網羅的 FMO 計算向け

た第 1 歩として、代表となるタンパク質フォールド (立体構造のパターン) を持つ薬 6,000 の PDB データに対する FMO 計算を実施し、構造生物学の新たなインシリコ基盤となるデータベースを構築する。

また、FMO 法によるタンパク質の量子化学計算の現状は、エネルギー計算は実用レベルに達しているものの、エネルギー微分を用いた構造最適化 [5, 6] や *ab initio* MD (FMO-MD) [7] は膨大な計算資源を要するために殆ど実施例がない。しかしながら、タンパク質構造の精密化やドラッグデザイン、酵素反応等の解析のためにはこれらの実用化が必須であるため、本研究では SQUID 上での実行を試みた。

### 2. フラグメント分子軌道法

FMO 法では、生体高分子を残基単位などの「フラグメント」に分割し、フラグメントごとの量子化学計算を実行する。その際には、各フラグメントは他の全てのフラグメントからの環境静電ポテンシャルの存在下にあるとする。これを、全てのフラグメント (モノマー) だけでなく、そのペア (ダイマー) についても計算する。式の詳細は省略するが、全エネルギーは、以下の式で表すことができる。

$$E_{total} = \sum E'_i + \sum \Delta \tilde{E}_{IJ} \quad (1)$$

$E'_i$  は環境静電ポテンシャルの寄与を除いたモノマーのエネルギー、 $\Delta \tilde{E}_{IJ}$  はフラグメント間の相互作用エネルギー (Inter-Fragment Interaction Energy; IFIE) である。IFIE は網羅的に算出され、

それだけでも解析に用いることができるが、さらに、PIEDA (Pair Interaction Energy Decomposition Analysis)法によって、4つのエネルギー成分 (静電項 (ES)、交換反発項(EX)、電荷移動項 (CT+mix)、分散項 (DI)) に分割することができる [8]。

$$\Delta\tilde{E}_{IJ} = \Delta\tilde{E}_{IJ}^{ES} + \Delta\tilde{E}_{IJ}^{EX} + \Delta\tilde{E}_{IJ}^{CT+mix} + \Delta\tilde{E}_{IJ}^{DI} \quad (2)$$

PIEDA を用いることで、分子間相互作用の物理化学的な解釈が可能となる。例えば水素結合では、主要成分の静電項に加えて電荷移動項が関与する。CH/ $\pi$ や $\pi$ - $\pi$ 相互作用では、分散項が主な成分となる。また交換反発項からは、立体障害を検出できる。

本研究でのフラグメント分割は、タンパク質はアミノ酸単位、RNA は主鎖/塩基単位で行い、低分子化合物や水分子は単一フラグメントとして扱った。なお、FMO 計算プログラムは ABINIT-MP を用いた。

### 3. タンパク質のフォールド代表構造の網羅的解析

PDB に登録されているタンパク質の立体構造は、X 線結晶構造解析やクライオ電子顕微鏡、NMR などの実験手法を用いて長い年月をかけて解析されてきた。近年の AlphaFold2 に代表されるタンパク質立体構造予測手法の精度の飛躍的向上により、実験構造未知の静的な立体構造予測も可能となった。今後は構造情報に対して構造生物学上有用な情報の付加が求められており、FMO DB への期待は大きい。そこで本研究項目では、フォールドデータベースである SCOP 2.0 に登録されているファミリーから代表となる 5,936 構造を選定し、FMO 計算結果を実施した。

PDB に登録されている実験構造は、欠損ループや欠損原子などを含んでいるものが多い。また、X 線結晶構造解析やクライオ電子顕微鏡では水素原子の位置が通常は特定できない。そのため、FMO 計算を実施するための自動モデリング機能 (Automatic complementation) を開発し、タンパク

質構造の前処理を行った (図 1)。

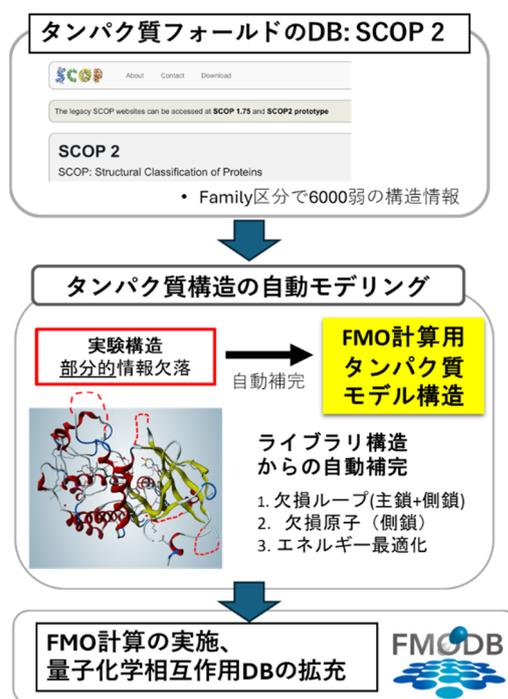


図 1 タンパク質基本フォールドの網羅的計算の流れ

SCOP2.0 からの 5,936 構造のうち、重複を除いた 5,352 構造を処理した結果、5,332 構造の自動処理に成功し、MP2/6-31G\*レベルの FMO 計算を実施した。用いたタンパク質には、例えば図 2 に示すような様々なフォールドが含まれる。

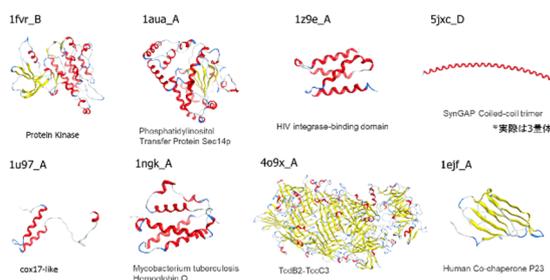


図 2 SCOP2.0 に登録されている代表構造の例

得られた FMO 計算結果から、各フォールドのタンパク質における残基間相互作用エネルギーを取り出すことができる。例えば N 残基からなるタンパク質には、 $N(N-1)/2$  個の残基ペアが内在しており、それらの全てに対して IFIE および PIEDA の相互作用エネルギーが得られる。全計

算結果を統計的に解析すると、中性残基間と荷電性残基間の相互作用のエネルギー分布は図 3 のようになっていた。中性残基間の相互作用 (図 3A) では、最近接原子間の距離が、1.6~2.0 Å のところでのエネルギー的に安定な相互作用 (-15~-20 kcal/mol) をしているペアが多く、これらは主に水素結合によるものであると考えられる。一方で、荷電性残基間の相互作用ペアのエネルギーは、同程度の距離で-120 kcal/mol 程度であり、これらの残基間では塩橋を形成しているため大きく安定化する (図 3B)。今後、残基の種類ごとに分類した同様の解析を進め、タンパク質に内在する残基間相互作用の分布を明らかにする予定である。また、これらのデータは FMO DB にも順次登録し、公開する予定である。

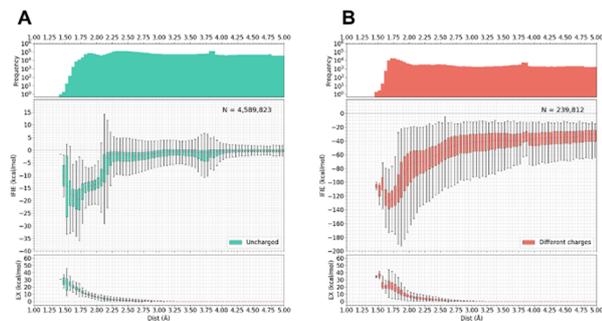


図 3 残基間の距離 (Å) と相互作用エネルギー (kcal/mol) の関係。(A) 中性残基間、(B) 荷電性残基間 (酸性-塩基性アミノ酸残基) の組み合わせ。

#### 4. 量子化学計算による生体高分子の動的挙動の解明

ABINIT-MP プログラムによる FMO 計算では、結晶構造や古典 MD からのサンプリング構造に対するエネルギーの一点計算は既に実用レベルに達しており、アカデミアのみならず製薬企業等の研究現場でも用いられている。しかしエネルギー微分を用いた構造最適化や *ab initio* MD (FMO-MD) は膨大な計算資源を要するために未だ実用化されていない。実験的な分解能の低い PDB データの精密化や精密なドラッグデザイン、さらには酵素反応や DNA/RNA などの新規創薬モダリティの解析において、構造最適化の実用化は必須

である。特に構造決定の難しい RNA や環状ペプチドに対しては、力場が十分に整備されていないことから、その構造解析や創薬応用への期待が高まっている。

構造最適化では、タンパク質-リガンド系の結合サイトの部分構造最適化を Frozen Domain (FMO/FD) 法 [5] を用いて HF/6-31G\* レベルで実施した。核内受容体、キナーゼ、アンジオテンシン II 受容体などの典型例に対して、リガンドと周辺数残基の構造最適化が実施できることを確認できた。さらに、FMO 最適化構造を用いたリガンド結合性の評価では、分子力場や QM/MM 法による最適化構造を用いるよりも阻害活性値をよく再現できることが示された[9]。

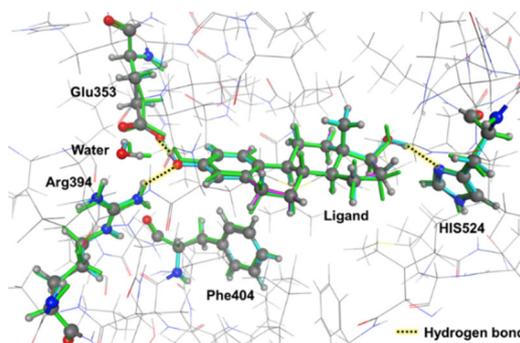


図 4 リガンド周辺の水素結合ネットワーク構造の最適化

次に、FMO-MD 計算の例として、バルジ RNA と低分子リガンドの複合体 (5 塩基対モデル) の計算を実施した。計算レベルは HF/6-31G\*、

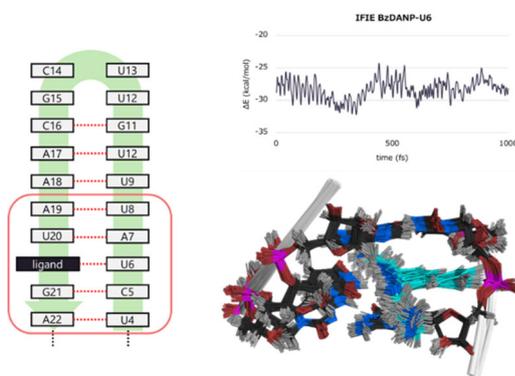


図 5 RNA-低分子複合体の FMO-MD 計算タイムステップを 0.5 fs とし、8 ノードを用いて約 26 日で 1.5 ps まで進行した。今後、古典 MD

との揺らぎ構造の違いや相互作用の変化などの評価・検討を実施する予定である。

## 5. おわりに

本研究の実施によって、世界に類を見ない、タンパク質基本フォールドを網羅した、量子化学計算データベースを構築することが可能となった。本研究で得られた膨大かつ網羅的なデータを基にした IFIE/PIEDA の分布を新標準として、例えば PDB 構造（実験構造）の妥当性を評価する手法の開発などの活用が期待される。量子化学的な観点から、現在結晶構造のモデルのクオリティ評価に用いられている結合長、結合角、二面角等の情報とは重複しない新たな視点を提供できる。また、SCOP 登録構造はフォールディング的に独立性が高いことから、機械学習によるモデル構築のためのデータセットとして期待でき、さらに創薬応用分野で課題となっているデータ不足問題を解決可能である。今後は、PDB 構造の網羅に向けてさらにデータ数を増やすとともに、応用例としては、FMO-AI 力場及び、相互作用予測 AI の 2 つの AI モデル構築等へも展開する予定である。加えて、本研究による網羅的なデータの拡充によって、FMO DB 公開データが広く 2 次利用されることになると期待できる。

また、本研究によって、構造精密化や酵素反応解析の基盤となる構造最適化の実利用がスタートし、生体高分子の *ab initio* MD 計算の実現可能性が示された。今後多くの生体分子の機能解明に貢献することが期待される。

## 謝辞

本研究は、大阪大学薬学研究科／薬学部 量子生命情報薬学分野の高谷大輔講師、田雨時助教、奥脇弘次研究員、大野修氏、宮岸澄真氏、田中蒼大氏、半田佑磨氏、理化学研究所の渡邊千鶴博士、九州大学の加藤幸一郎博士の協力の下で進められた。また、構造最適化計算は、中野達也博士（国立医薬品食品衛生研究所）および中山尚史博士（コンフレックス社）、FMO-MD 計算は古明地勇

人博士（産業技術総合研究所）と共同で行われた。

## 参考文献

- (1) K. Kitaura, et al., Chem. Phys. Lett., **313**, 701-706 (1999).
- (2) Y. Mochizuki, S. Tanaka, K. Fukuzawa, ed. “Recent Advances of the Fragment Molecular Orbital Method: Enhanced Performance and Applicability”, Springer, Singapore, (2021).
- (3) D. Takaya D, et al., J. Chem. Inf. Model. **61**, 777-794 (2021).
- (4) C. Watanabe, et al., Chem-Bio Info. J., **19**, 5-18 (2019).
- (5) D. G. Fedorov, et. al., Geometry Optimization of the Active Site of a Large System with the Fragment Molecular Orbital Method, J. Phys. Chem. Lett. **2**, 282-288 (2011).
- (6) T. Tsukamoto et al., Partial geometry optimization with FMO-MP2 gradient: Application to TrpCage, Chem. Phys. Lett., **535**, 157-162 (2012).
- (7) Y. Komeiji, et al., Fragment molecular orbital method: application to molecular dynamics simulation, ‘ab initio FMO-MD’ Chem. Phys. Lett., **372**, 342-347 (2003).
- (8) D. G. Fedorov, K. Kitaura, Pair interaction energy decomposition analysis. J. Comput. Chem. **28**, 222-237 (2006).
- (9) K. Okuwaki, et al., Geometry Optimization using the Frozen Domain and Partial Dimer Approach with the Fragment Molecular Orbital Method: Implementation, Benchmark, and Application for Ligand-Binding Site of Proteins, in preparation.