インシリコ解析ユニット

補助事業代表者（所属・職名）
本間　光貴（理化学研究所・チームディレクター）

AIとFMO法を融合したインシリコスクリーニングと分子間相互作用解析

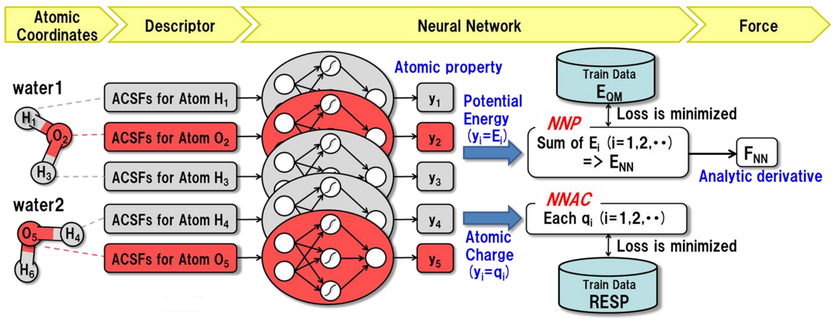
補助事業分担者（所属・職名）
福澤　薫（大阪大学大学院薬学研究科・教授）、加藤　幸一郎（九州大学大学院工学研究院・教授）

高度化における研究背景・目的

●FMO計算データと機械学習・AIを活用した相互作用予測の高度化と応用

- インシリコスクリーニングにおいては高精度かつ高速な分子間相互作用の予測が望ましい
- フラグメント分子軌道（FMO）法はタンパク質・リガンド複合体のような大規模複雑系の量子化学計算を可能としており、加えてタンパク質を構成するアミノ酸とリガンドの相互作用エネルギーも評価できる
- タンパク質・リガンド複合体を構成する原子座標に基づいて、上記のような相互作用エネルギーをFMO法と同精度で高速に予測する機械学習・AIモデルを構築する
- そのようなAIモデルを推し進めることにより、機械学習力場によるタンパク質の分子動力学計算を実現する

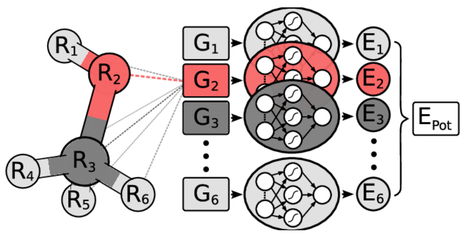
●FMO計算データの機械学習応用



高度化の成果

FMO計算から得られるリガンド・アミノ酸残基相互作用の機械学習モデルによる予測

●機械学習を用いた相互作用予測



- FMO計算から相互作用エネルギーである IFIE (inter-fragment interaction energy)や、そのエネルギー分割法である PIEDA (pair interaction energy decomposition analysis)により静電項(ES)、交換反発項(EX)、電荷移動項(CT)、分散項(DI)が得られる
- 機械学習の方法を用いて原子座標から上記を予測する

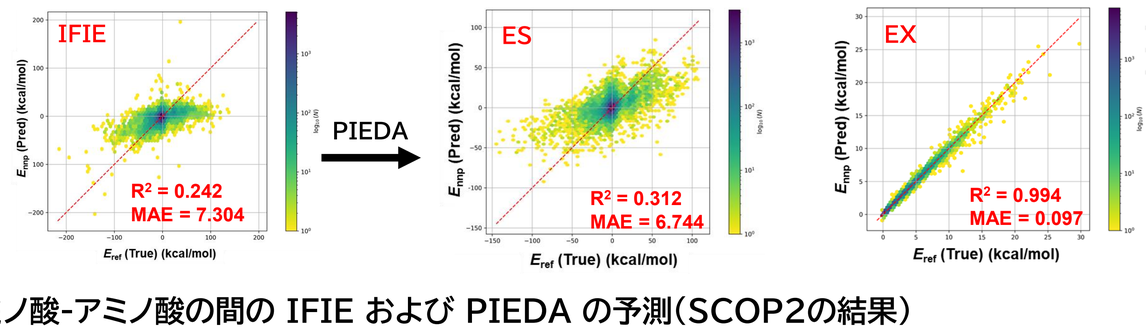
●FMO計算を実行する対象としたデータベースとその内容

- DUD-E(ドッキング→FMO計算でデータ生成)

	Train	Validation	Test
HIVP	205262 (469)	25644 (66)	25764 (57)
CDK2	42770 (176)	5033 (21)	5815 (26)
AKT1	33106 (57)	3544 (8)	5919 (12)
AmpC	19343 (66)	2266 (10)	2821 (8)
EGFR	18218 (52)	2179 (8)	2426 (8)
HIVRT	9571 (54)	1085 (7)	1427 (9)

- SCOP2(FMO計算でデータ生成)
  - ✓ 基本フォールド　約6,000構造
  - ✓ アミノ酸-アミノ酸間のIFIE, PIEDAを予測
  - ✓ train: validation: test = 250,000: 31,200: 31,200

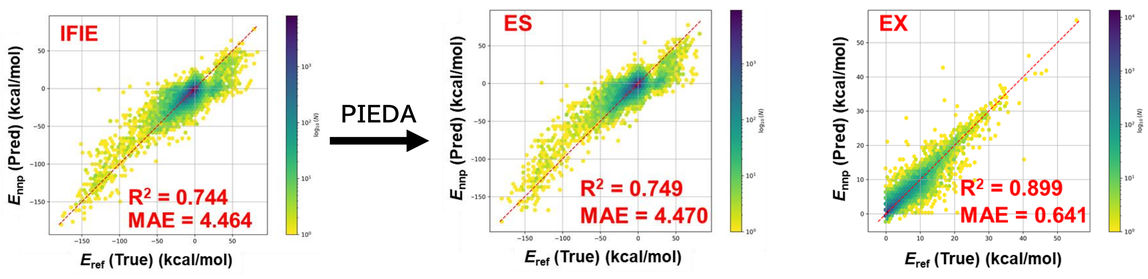
●リガンド-アミノ酸の間の IFIE および PIEDA の予測(DUD-Eの結果: HIV-Pを例に)



- IFIE の予測精度は ES 予測精度が支配的
- フラグメントに異なる電荷がある場合であっても、電荷を分けることなく学習・予測
- ES 以外の予測精度はリガンド電荷によらない

タンパク質を固定した場合、AI によるリガンド分子生成の評価関数に利用可能と期待できる

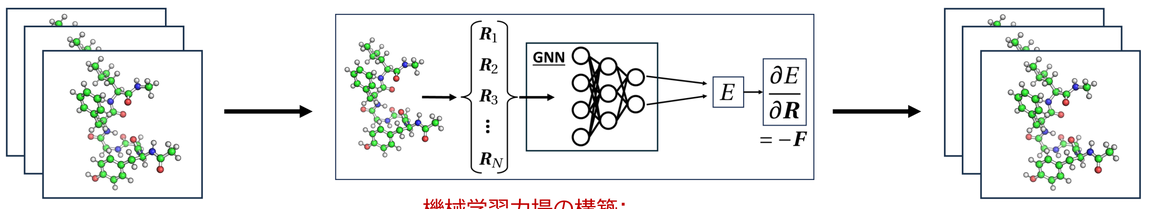
●アミノ酸-アミノ酸の間の IFIE および PIEDA の予測(SCOP2の結果)



- 上記と同様に、IFIE の予測精度は ES 予測精度が支配的
- フラグメントに異なる電荷がある場合であっても、電荷を分けることなく学習・予測

FMO計算に基づいた機械学習力場の構築: 量子化学精度の力の予測と分子動力学計算

●FMO計算結果とグラフニューラルネットワークを用いたFMO機械学習力場の構築



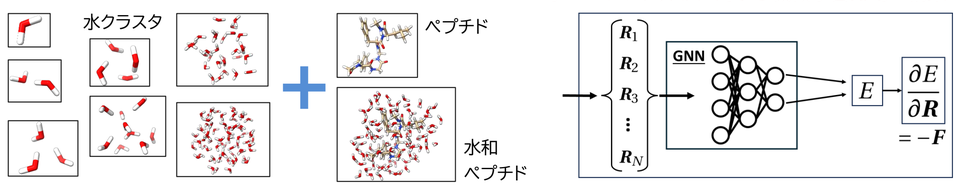
古典力場によるMD計算: 構造サンプリング

機械学習力場の構築: FMO計算のエネルギー・力の学習

機械学習力場によるMD計算

●水分子とペプチド分子構造を学習データとしたFMO力場の構築

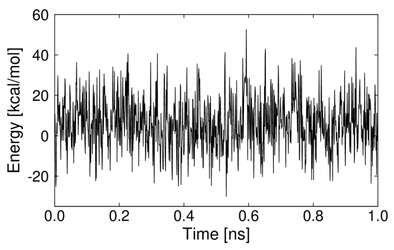
水クラスタと水和ペプチドを含む学習で水中のMD計算を可能に



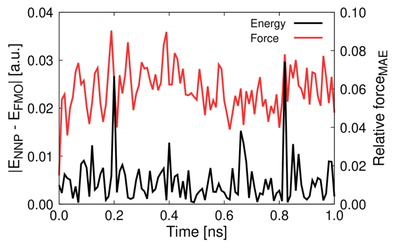
●ペプチド分子(enkephalin)のFMO機械学習力場による分子動力学計算

FMO計算に匹敵する精度でナノ秒（ns）程度の安定的なMD計算が可能に

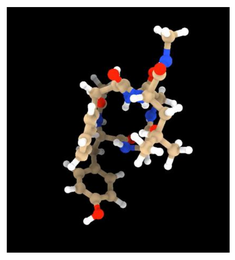
MD計算中のFMO力場によるエネルギー



FMOとの比較(E: 0.01 au, F: 5%程度の誤差)



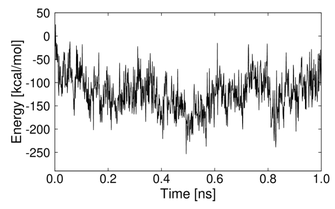
86 atoms, 500 K, 1 ns



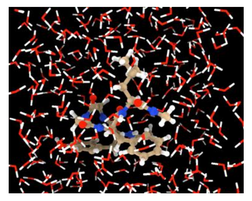
●水中の生体分子系に対するFMO力場を利用した分子動力学シミュレーション

水中のLeu-enkephalin：周期境界条件でのMD計算も可能

MD計算中のFMO力場による エネルギー

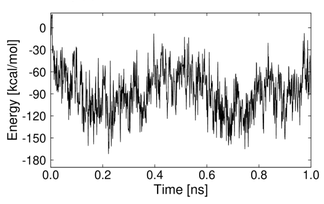


1,070 atoms, 310 K, 1 ns

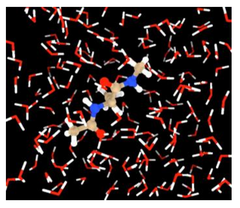


水中のアラニンジペプチド：非学習アミノ酸を含むMD計算

MD計算中のFMO力場による エネルギー



577 atoms, 310 K, 1 ns



今後の発展

- FMOデータを用いた機械学習モデル(FMOサロゲートモデル)の開発
- 相互作用(IFIE/PIEDA)予測モデル
- FMO-AI力場の開発
- MD-FMOのサンプリング高度化
- FMOサロゲートモデルを用いたAI創薬技術開発
- 分子生成AI → 複合体構造生成 → FMOサロゲートモデルのサイクル
- FMO相互作用の表現学習

など

本発表に関連し、開示すべき利益相反（COI）はありません。