配布用参考資料

※当日の内容とは若干異なる場合がございます。

FMODBの開発および相互作用データ解析

高谷 大輔

理化学研究所 生命機能科学研究センター

http://eniac.scitec.kobe-u.ac.jp/fmodd/index.html.

K computer (project ID: hp190119)

BINDS from AMED, JP19am0101113

本日の内容

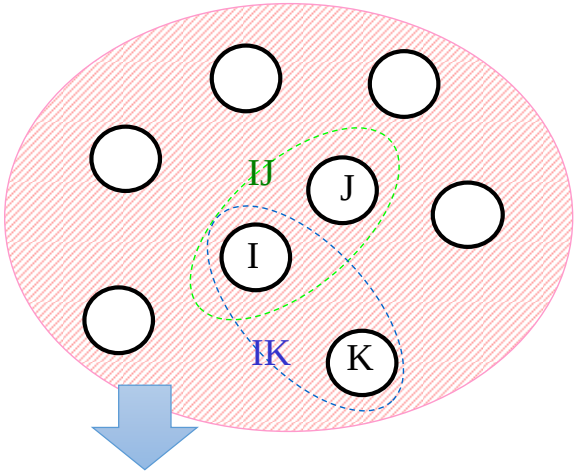
FMODB
2018-2020

- FMO法について
  - 相互作用解析例について
- FMODBの開発
  - Auto-FMO protocolの開発
  - FMODBの紹介
- IFIE/PIEDAデータを使った相互作用解析

フラグメント分子軌道(FMO)法

FMODB 2018-2020

- 1999年に北浦和夫教授（京大）らが提案した日本初のfull QM手法。
- **数千残基のタンパク質の全体構造を迅速に量子化学計算**することが可能な手法。



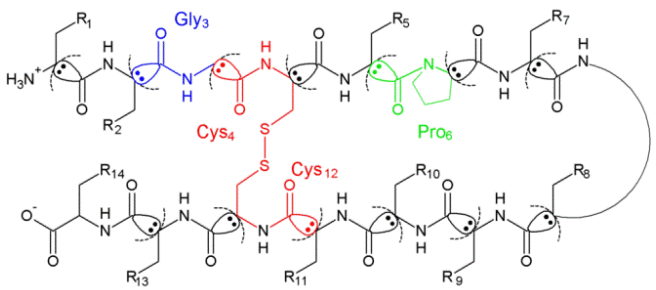
巨大分子をフラグメントに分割（アミノ酸残基、リガンド単位）

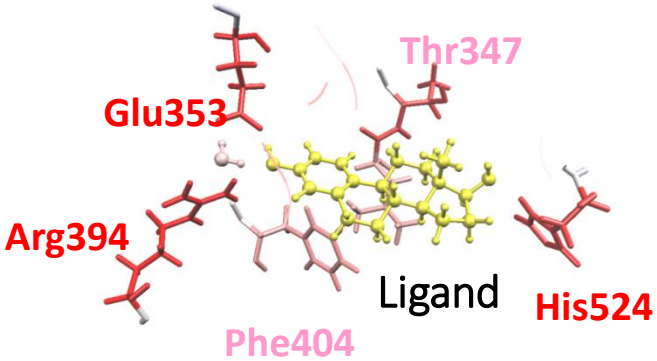
全エネルギー:

$$E_{\text{total}} = \sum_{I} E'_{I} + \sum_{I>J} \Delta\tilde{E}_{IJ} + \sum_{I>J>K} \Delta\tilde{E}_{IJK} + \sum_{I>J>K>L} \Delta\tilde{E}_{IJKL} + \cdots$$

FMO2ではモノマーとダイマーのエネルギーから全エネルギーを算出

フラグメント間相互作用エネルギー(IFIE, PIE):





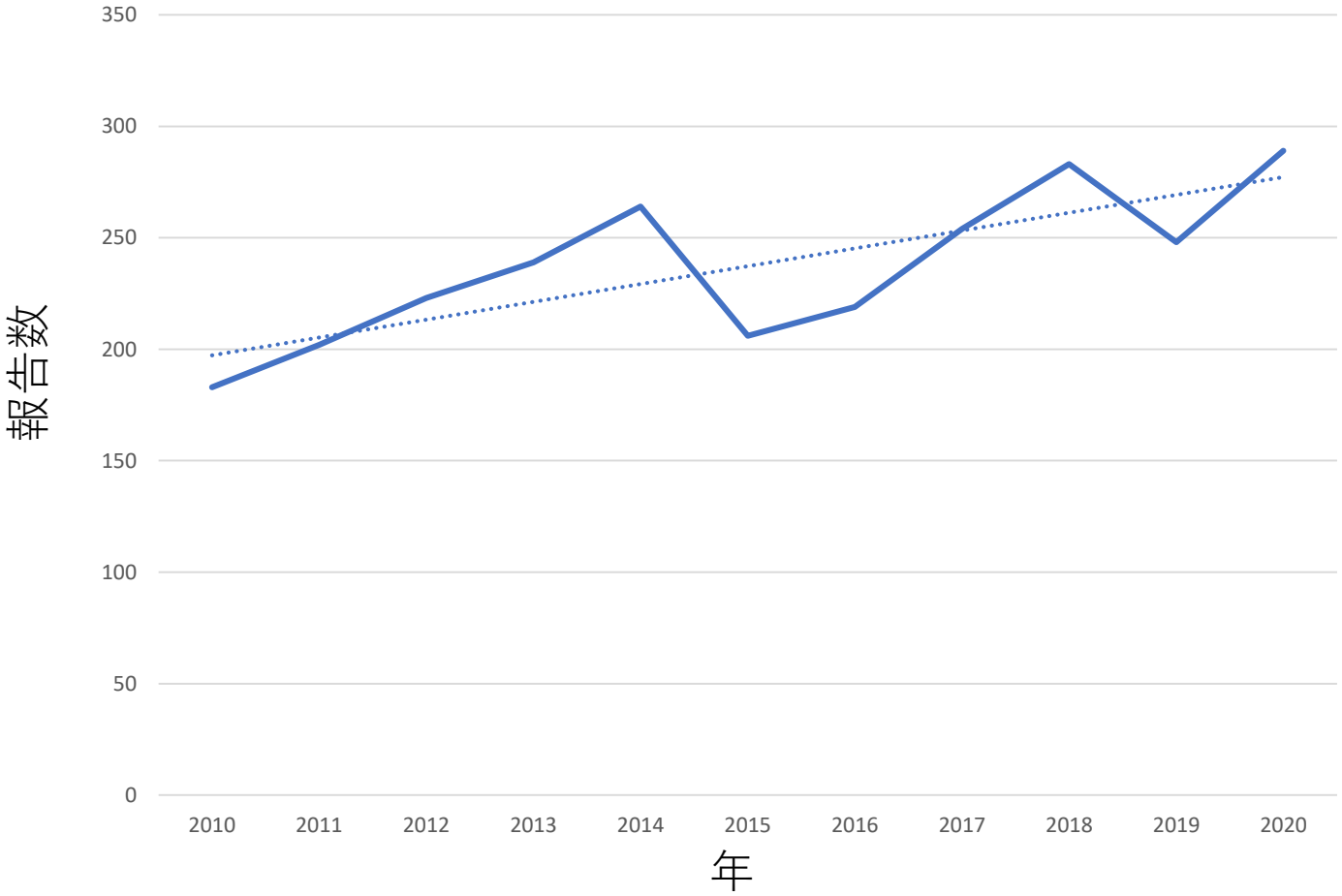
モノマー間の相互作用エネルギー

電子的な挙動まで考慮した相互作用解析が定量的に可能に！

フラグメント分子軌道(FMO)法

FMODB
2018-2020

年ごとのFMOに関する報告数の推移



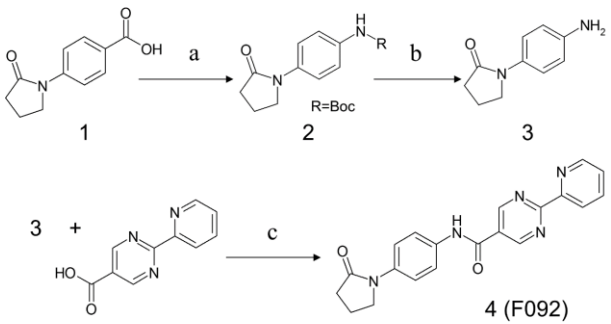
- Google Scholarにて「”fragment molecular orbital “」を年毎に検索
- 2020/10/26時点
- 今年はすでに300件近い！

FMO計算に基づく相互作用解析例(1.1)

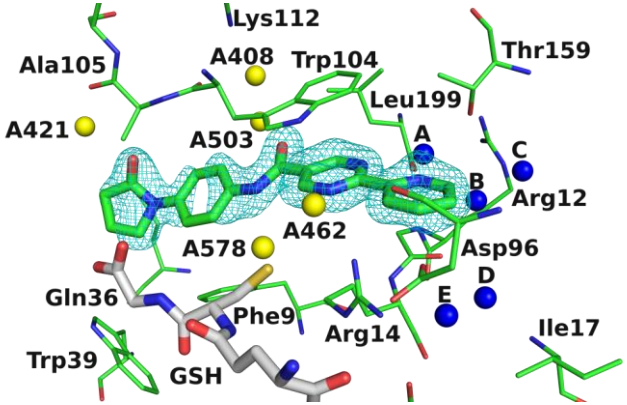
FMODB 2018-2020

H-PGDSの阻害剤設計研究概要

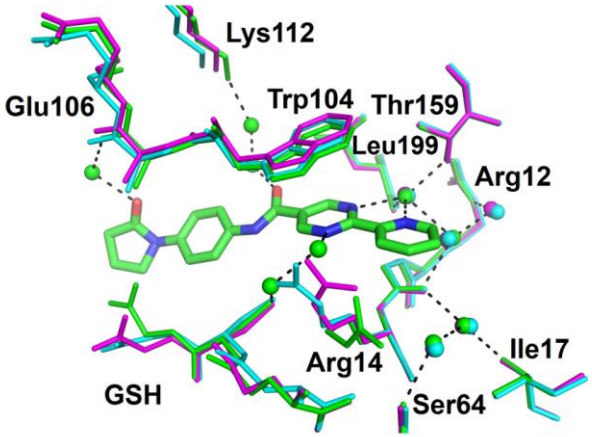
H-PGDS: プロスタグランジンD2合成を触媒する酵素の1つ
アレルギーおよび炎症反応などの**ターゲットタンパク質**

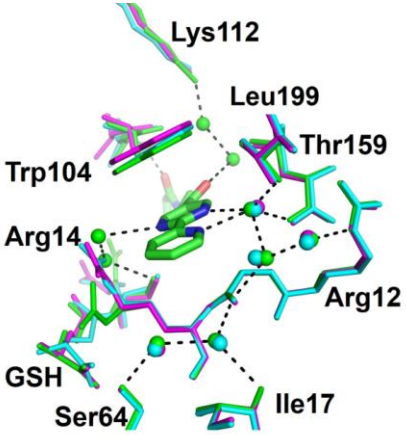


F092: Kd(nM) = **0.14** ± 0.010



本研究で得られたH-PGDS-F092複合体構造;PDBID 5YWX; 解像度1.74 Å; 各球表示は結晶水（A-EはほとんどのH-PGDS結晶で保存している結晶水(=**保存水**)）





他のH-PGDSの結晶構造（PDBID:5YWX (緑)、1IYH (マゼンタ)、3KXO（シアン））を重ねて表示。水素結合を破線で表示。**保存水を含めた結晶水がリガンド分子も含めて多くの水素結合ネットワークを形成していた。**

Takaya, D., Inaka, K., Omura, A. *et al. Bioorg Med Chem.* **26**, 4726-4734 (2018)

FMO計算に基づく相互作用解析例(1.2)

FMODB 2018-2020

F092-結晶構造のIFIE/PIEDA解析

量子化学計算による定量的相互作用評価

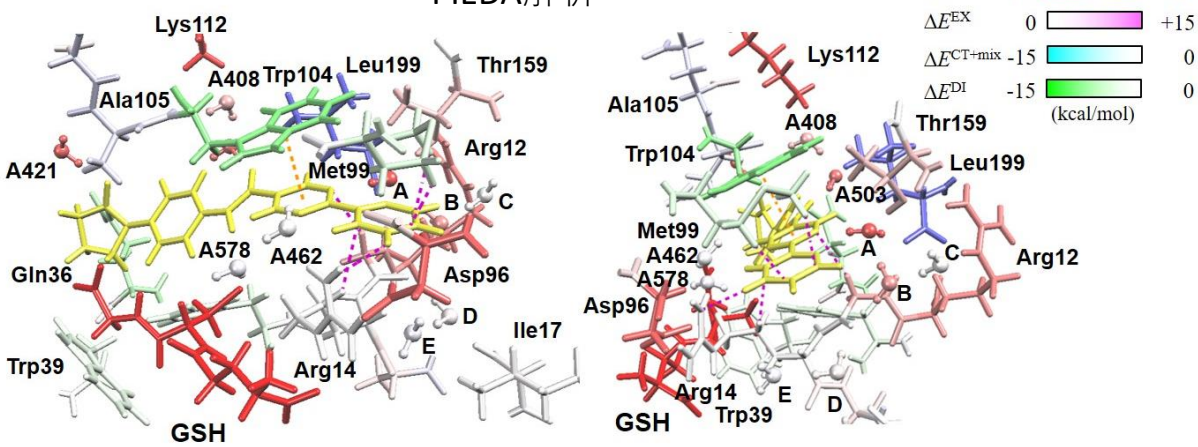
water-amino acid(AA) interaction		IFIE (kcal/mol)
water	AA	AA-water
$Water_A$	Leu199	-21.60
$Water_B$	Arg14	-8.70
$Water_C$	Arg12	-18.13
$Water_D$	Arg14	-8.27
$Water_E$	Ser64	-8.62

アポ構造における**保存水-アミノ酸相互作用ペアー**（最も強いIFIE値を表示;　PDBID: 1IYHについて計算した）

water-amino acid(AA) interaction		IFIE (kcal/mol)			
water	AA	AA-water	water-F092	AA-F092	AA-water-F092**
$Water_A$	Leu199	-16.72	-11.54	10.09	-18.16
$Water_B$	Arg14	-9.68	-6.75	2.07	-14.35
$Water_C$	Arg12	-16.3	-0.09	-7.91	-24.29
$Water_D$	Arg14	-8.66	-0.65	2.07	-7.24
$Water_E$	Ser64	-8.06	0.76	0.3	-7.01
HOHA408	Lys112	-14.37	-4.67	-12.5	-31.54
HOHA421	Glu106	-11.15	-10.57	2.6	-19.13
HOHA462	Arg14	-12	-4.42	2.07	-14.35
HOHA503	Lys112	-4.94	-9.11	-12.5	-26.55
HOHA578	Arg14	-7.04	0.68	2.07	-4.29

** **左3つのカラムの合計値**

PIEDA解析



- 保存水を介することで受容体とエネルギー的に強く結合
- F092とH-PGDSが安定に結合することがエネルギー的に示唆された。

Takaya, D., Inaka, K., Omura, A. *et al.* *Bioorg Med Chem.* **26**, 4726-4734 (2018)

FMO法について

FMODB
2018-2020

まとめ

- FMO法を用いた解析例について紹介
  - タンパク質ーリガンド間相互作用解析事例
  - 年々と報告数は増加傾向にある。

- こういったFMO計算結果があらかじめ提供されているならば有益ではないだろうか？

- FMO計算の相互作用エネルギーの傾向はどうか。

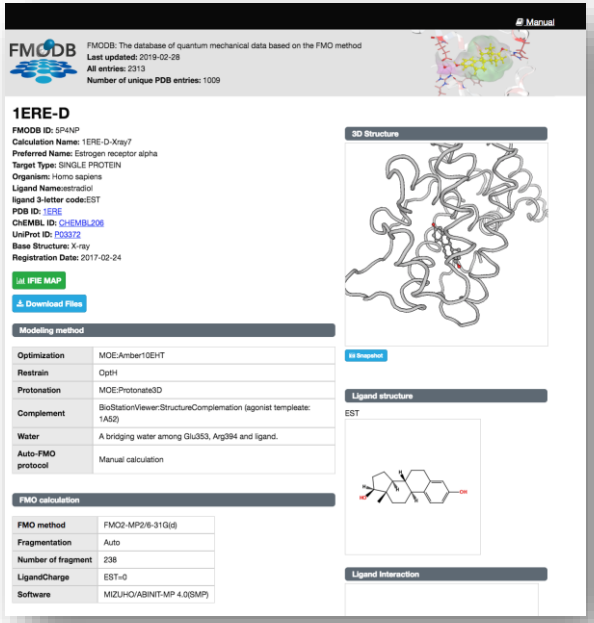
本日の内容

FMODB
2018-2020

- FMO法について
  - 相互作用解析例について
- FMODBの開発
  - Auto-FMO protocolの開発
  - FMODBの紹介
- IFIE/PIEDAデータを使った相互作用解析

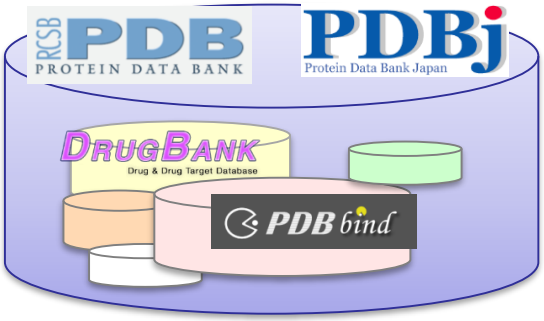
世界初のタンパク質の量子化学計算値DB

PDBなどの構造データのFMO計算結果を収載するデータベース



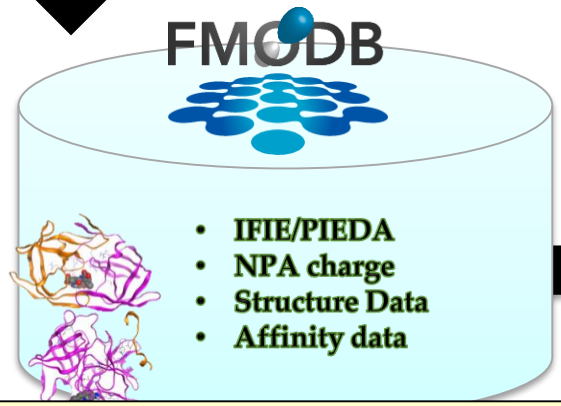
Webページで公開中

https://drugdesign.riken.jp/FMODB



$ rsync -avz --delete ftp.pdbj.org::ftp_data/structures/divided/pdb/ $topdir/pdb

目標
全データをFMO計算！



主な用途

- タンパク質の精密な相互作用エネルギーを介した生命現象の解明
- **FMO**力場開発の基礎データ
- 医薬品設計への利用（新しい置換基の提案など）

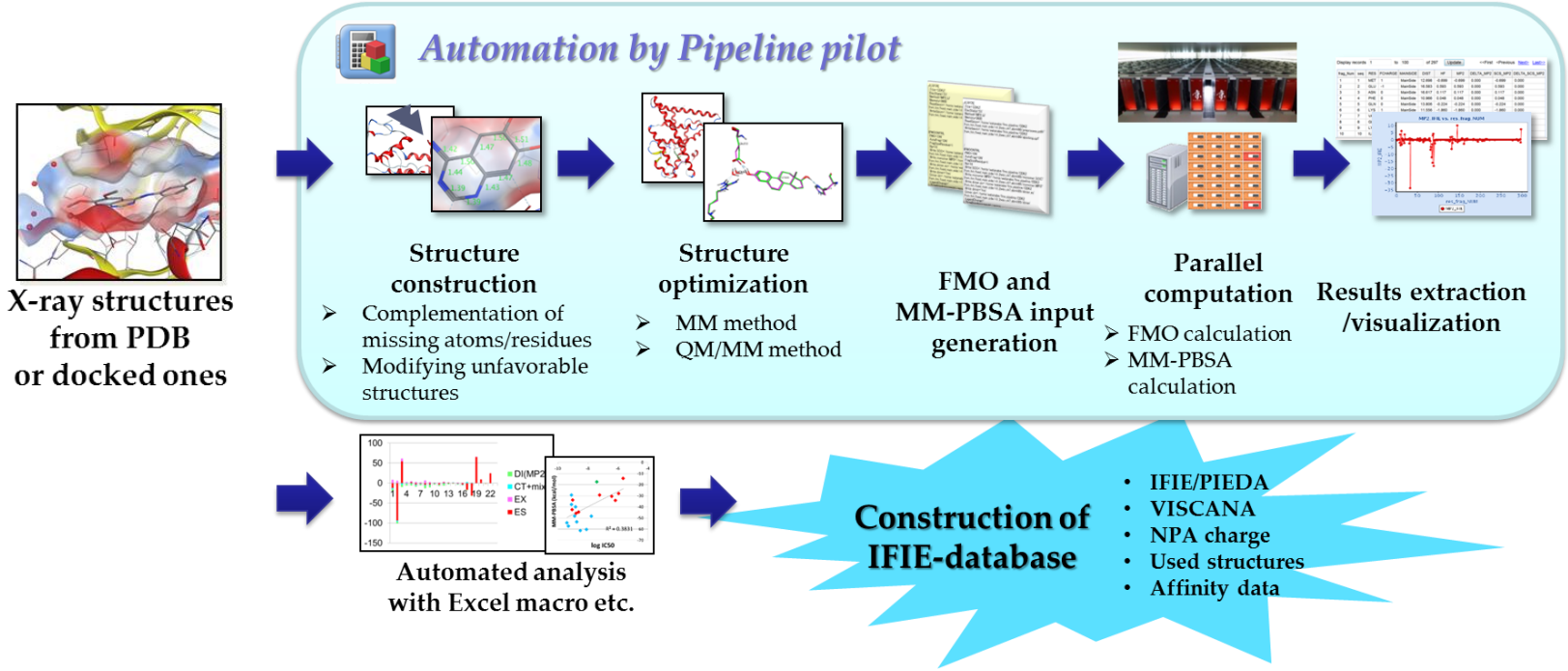
Last updated: 2020-10-19
All entries: 12732
Number of unique PDB entries: 1204

FMO計算自動化プロトコルの開発

手作業には限界がある。

モデリング、入力ファイル作成、計算実行、解析 etc.

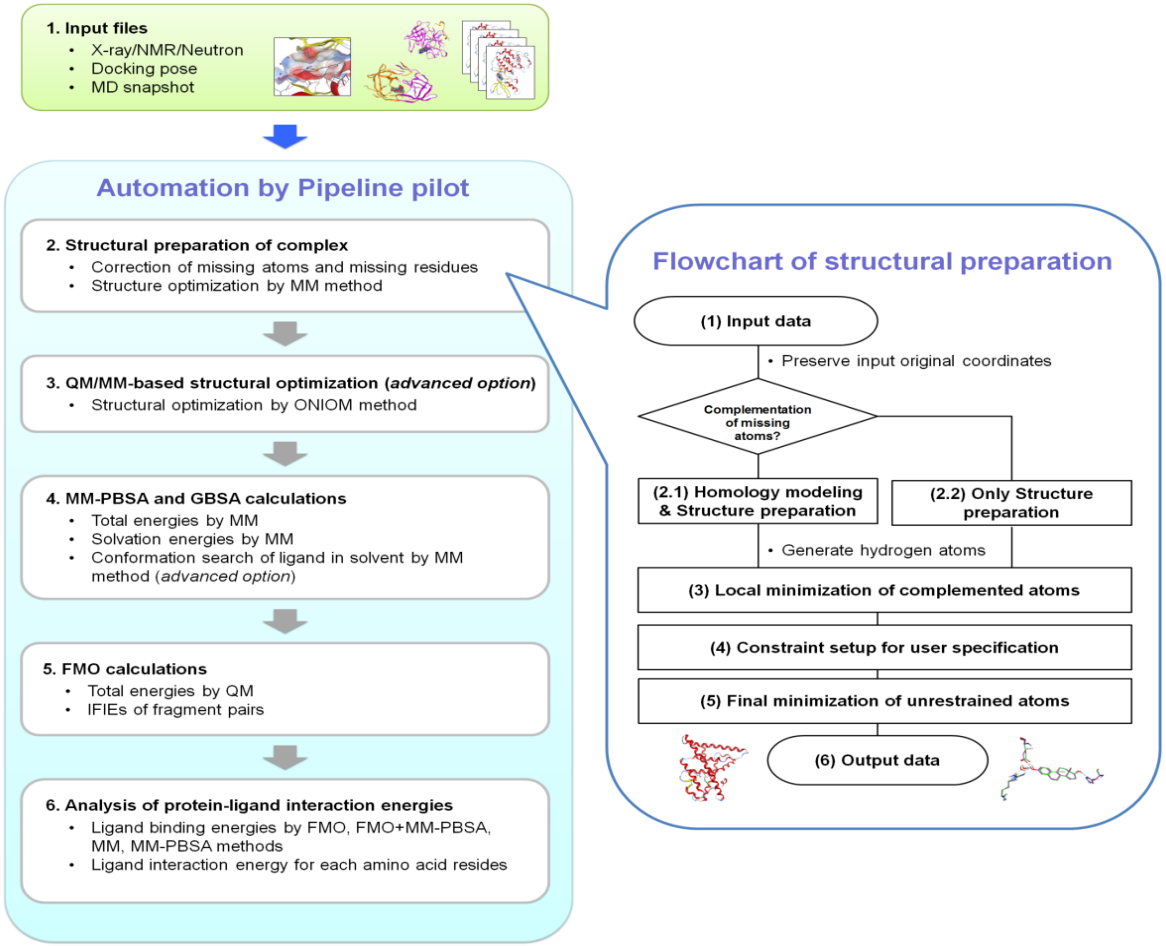
Auto-FMO protocolの開発



- 自動処理でFMO計算実行数を飛躍的に増やすことが可能
- FMODBの構造増加

Auto-FMOプロトコルの機能

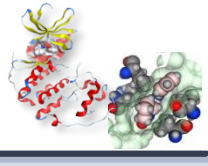
自動で前処理から計算実行までをおこなう。



Watanabe C *et al. CBI Journal* **19**, 5-18, *2019*

- ユーザーが用意するのはタンパク質構造
- 入力構造の前処理
  - 欠損残基の補完
  - キャップ（ACE, NME)導入
  - リガンド分子のMM
  - etc
- FMO計算の実施
  - ABINIT-MP
  - MP2/6-31G*
  - IFIE/PIEDA解析
  - etc
- その他（DBには含まれない）
  - QM/MM Gaussian
  - MM-PBSA

Auto-FMOプロトコル精度検証



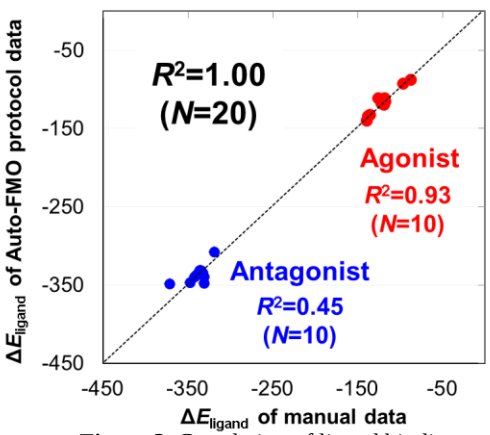
Success ratio and computational time of FMO calculations

Target	# of residue /fragment	PDB with IC_{50}	Constructed structures by MM-based minimization	Completed FMO calculations	Average of total time for each structure
Estrogen receptor α	236/238	22	38	38	***9 hours
Estrogen receptor β	234/235	29	55	*54	***6 hours
CHK1 kinase	261/262	42	44	44	***9 hours
Aurora kinase A	260/261	37	55	**53	***12 hours
p38-MAP kinase	346/347	85	222	**202	***14 hours
CDK2	298/299	11	*94	*94	***8 hours
Total	-	226	508	**485 (95.5%)	-

* Including 83 docking pose. ** Several failed data. *** K-computer using 960cores.

手動データ vs. FMOプロトコルデータ

(A)ERα



(B)p38α kinase

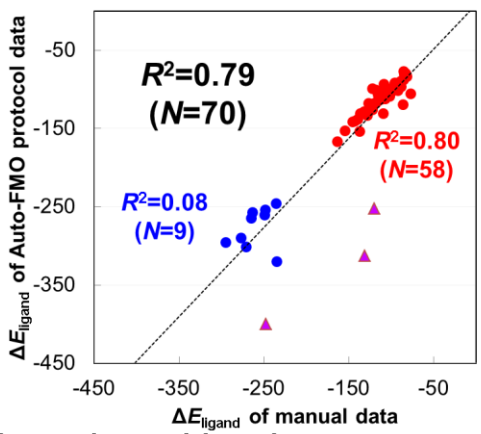


Figure 3. Correlation of ligand binding energies between the manual data and the Auto-FMO protocol data for ERα (A) and p38α (B).

- FMODDのデータを使用
- プロトコルで作成した約500複合体（6ターゲット）のうち**95%**の構造が問題なくFMO計算が完走！
- X線結晶構造、ドッキング構造、MD構造もプロトコルの初期構造として利用可能。
- 手動作成構造データ（FMODD提供）ともIFIE値は遜色ない（R^2=0.9）。

FMODBへのデータ登録の流れ

主にFMODDコンソーシアムで実施

①FMO計算実行

各自研究室のクラスター等

②計算条件を記述したexcelファイルを作成

F	G	H	I	J	K	L
LastModifie	pdb_file	out_file	cpf_file	ajf_file	job_script	PDB_ID
	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	4DEA
	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	4DED
	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	4DHF
	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	/data/hp1 ξ	4DHF

③指定された
フォルダにデータを置く

※今後は相談して多くの研究者の
計算データを設置可能にする予定

RIKEN 研究室内部

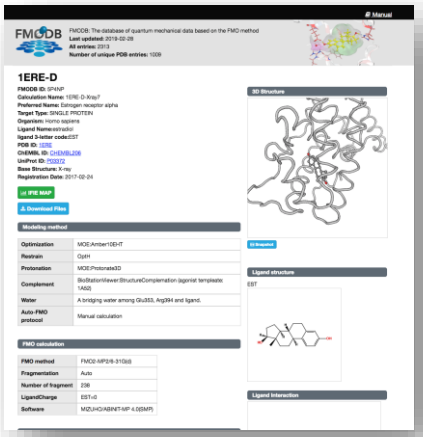
1. Errorチェック

正常終了した計算
データのみを保存

2. MySQLデータ化
計算生データのコピー

drugdesign
(公開用サーバー,http)

Webインターフェイス



FMO計算
研究者等

一般ユーザー

登録用Excelファイルの内容（一部）

※赤は必須項目

カラム名	説明
Lig_Res	リガンドの残基
BaseStructure	ベース構造（X-ray/MD/docking）
Water	水処理（YES/NO）
Protonation	プロトン化の方法（MOE:Protonate3D）
Optimization	最適化の方法（MOE:Amber10EHT）
Restrain	拘束
Complement	構造補完（Homology modeling）
Fragmentation	分割方法（Auto）
Ligand_Fragment_number	リガンドのフラグメント番号
ajf_file	*.ajfファイルのファイル名（パス無）
pdb_file	PDBファイルのファイル名（パス無）
out_file	ABINIT-MPのログファイル（*.out）のファイル名（パス無）
cpf_file	CPFファイルのファイル名（パス無）
job_script	ジョブ投入のスクリプトのファイル名（パス無）

カラム名	説明
Affinity	活性値
LigandCharge	リガンド電荷
Mutation	変異有無
PDB_Chain	計算に使用したPDBの鎖
PROTEIN_ACCESSION	ターゲット
Method	計算方法（FMO2-MP2/6-31G*-PIEDA）
UserID	ユーザID
Total_Core_Thred	使用ノード
total_time	全体のFMO計算実行時間
start_time	FMO計算の実行開始時間
stop_time	FMO計算の実行終了時間
RecordData	FMODBへの登録日
LastModifiedDate	最終更新日（修正した場合など）

→　Auto-FMO プロトコルを経由すればほぼ自動で記載される。

FMODBへのデータの登録

今後の展開予定

- 現状ではFMODD参加者の計算データを公開している。
  - コンソ内部で相談できるので計算結果はある程度信頼できる。

- FMODDコンソーシアム参加者以外からの受付
  - 現状ではFMODD参加者以外は登録は難しい。
  - 必要項目を精査してWebから登録？
  - ライセンスについての同意など
  - 必要項目/仕様の策定など

  → FMODDコンソーシアムに入会するのが楽

- Webサーバーの性能
  - 現状 一台なので一度に大量の通信は限界がある。

- 何かご意見がありましたらZoomのチャット等からお知らせください。

FMO計算結果の自動エラー検出

- 多くのFMO計算（ABINIT-MPのみ）を実施していると計算が進まない原因は概ねパターン化してくる。

プログラム: abinitmp_status.py (理研で開発中)

```
[takaya@designbase16 linepop]$ python3 abinitmp_status.py -h
usage: abinitmp_status.py [-h]
                          {checkoutfiles,checklog,get_ligand_ifie,yourifie,apoifie}
                          ...

positional arguments:
  {checkoutfiles,checklog,get_ligand_ifie,yourifie,apoifie}
                        commands
    checkoutfiles       check files. (ajf,pdb,cpf,out)
    checklog            check abinitmp.log
    get_ligand_ifie     get ligand ifie!!
    yourifie            your interest ifie will be obtained
    apoifie             Apo IFIE

optional arguments:
  -h, --help            show this help message and exit
[takaya@designbase16 linepop]$
```

ログの失敗を示す部分を表示

```
Failed: sample1.out, msg='Invalid
number of fragment: 261'


Failed: sample2.out, msg='ABINIT-MP
is NOT DONE.; MONOMOER SCC ierr =
120'

  …
```

正常終了すると各種計算結果サマリーを表示
→DB登録へ

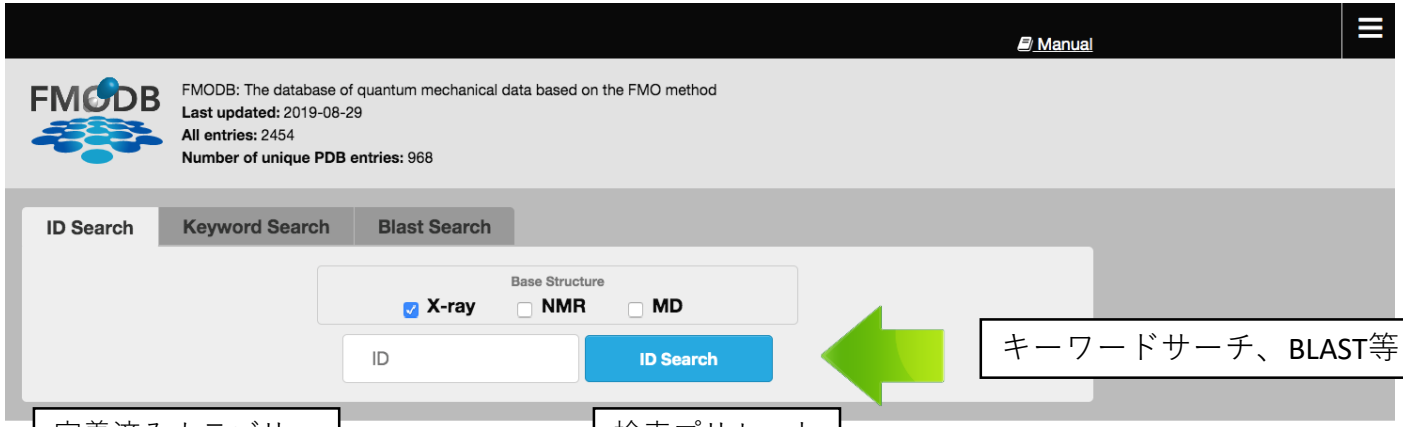
{'FMO2-HF:Nuclear_repulsion': '3055498.6796883377', 'obj_logfile_rel': 'tests/Ablog/ERa_1ERE-A_abinitmp.log', 'abinitmp_done': 'ok', 'np_fragment': 4, 'Grimme_SCS_energy::Total_energy': '-98400.7616948316', 'basis_set': '6-31G(d)', 'total_time_sec': 65583.1, 'pass_ReadGeom': 'ok', 'input_type': 'File', 'FMO2-HF:Total_energy': '-98132.2665254665', 'Hill_SCS_energy::Total_energy': '-98264.0884646828', 'abinitmp_calc_finish_date': '2015/02/18 13:10:11', 'abinitmp_calc_start_date': '2015/02/17 18:57:08', 'abinitmp_ver': 'MIZUHO/ABINIT-MP 3.0', 'Jung_(Head-Gordon)_SCS_energy::E(FMO2-MP2)': '-263.8229674149', 'FMO2-MP2:Total_energy': '-98410.1060987319', 'FMO2-HF:Electronic_energy': '-3153630.9462138042', 'n_total_core': 16, 'LigandCharge': '', 'AtomCheck': 'ok', 'Fragmentation': 'Auto:Res1:Main-Chain', 'Hill_SCS_energy::Electronic_energy': '-3153762.7681530206', 'obj_logfile_abs': '/home/takaya/Work/FMODB_ALL/linepop/tests/Ablog/ERa_1ERE-A_abinitmp.log', 'Grimme_SCS_energy::E(FMO2-MP2)': '-268.4951693651', 'number_of_fragment': '237', 'dimer_SCF_calculation': 'converged.', 'qm_method': 'MP2', 'Grimme_SCS_energy::Electronic_energy': '-3153899.4413831695', 'ori_logfile': 'tests/Ablog/ERa_1ERE-A_abinitmp.log', 'Hill_SCS_energy::E(FMO2-MP2)': '-131.8219392163', 'Jung_(Head-Gordon)_SCS_energy::Total_energy': '-98396.0894928814', 'FMO_method': 'FMO2-MP2/6-31G(d)', 'FMO2-MP2:E(FMO2-MP2)': '-277.8395732655', 'n_total_proc': 16, 'FMO2-MP2:Electronic_energy': '-3153908.7857870697', 'n_basis_functions': '32630', 'Jung_(Head-Gordon)_SCS_energy::Electronic_energy': '-3153894.7691812189', 'monomer_SCC_calculation': 'converged.'}

- エラー検出の自動化による効率化
- 他にも色々機能あり。
- より親切なエラー表示、解決法の提案を目指す。
  - エネルギーベースの警告など

FMODB Web user interface

FMODB 2018-2020

https://drugdesign.riken.jp/FMODB/

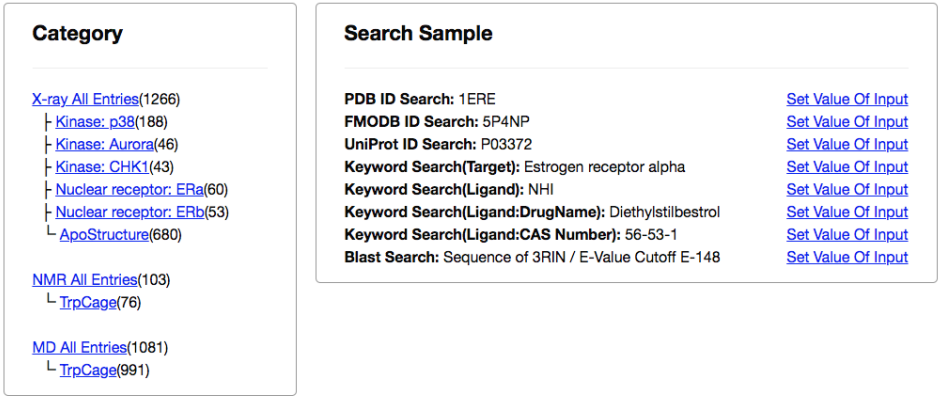


マニュアル等

キーワードサーチ、BLAST等

定義済みカテゴリー

検索プリセット



後でFMODBの使い方のデモを行います。

All of calculated data by FMO Drug Design (FMODD) Consortium are licensed under CC BY-SA 4.0.
Citing Us.

FMODB calculation data: CC BY-SA 4.0
https://creativecommons.org/licenses/by-sa/4.0/deed.ja

まとめ

FMODB
2018-2020

FMODBの開発

- Auto-FMO Protocol
- FMODBの内容
- Webインターフェイス

FMODBデータを使った相互作用解析

→ 引き続きFMODBのデモを行います。