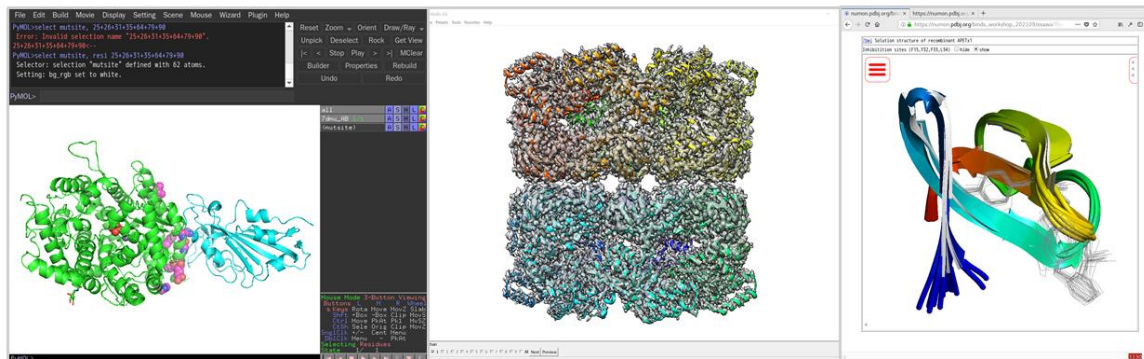




創薬等先端技術支援基盤プラットフォーム  
Basis for Supporting Innovative Drug Discovery and Life Science Research

BINDS

# PDBデータの使い方・ 分子ビューワの使い方の基礎



川端 猛

蛋白質研究奨励会 研究員・大阪大 生命機能 招へい准教授

kawabata@protein.osaka-u.ac.jp

2021年9月30日(木) 13:15-14:15

大阪大学 蛋白質研究所 本館2F大講義室 Zoomウェビナー開催

# 本日の話題

## PDBデータのなりたち

どんな情報が収納されているのか？

どういうファイルフォーマットで書かれているのか？

## 3つの分子ビューアの基本的使用方

PyMOL

UCSF Chimera

Molmil

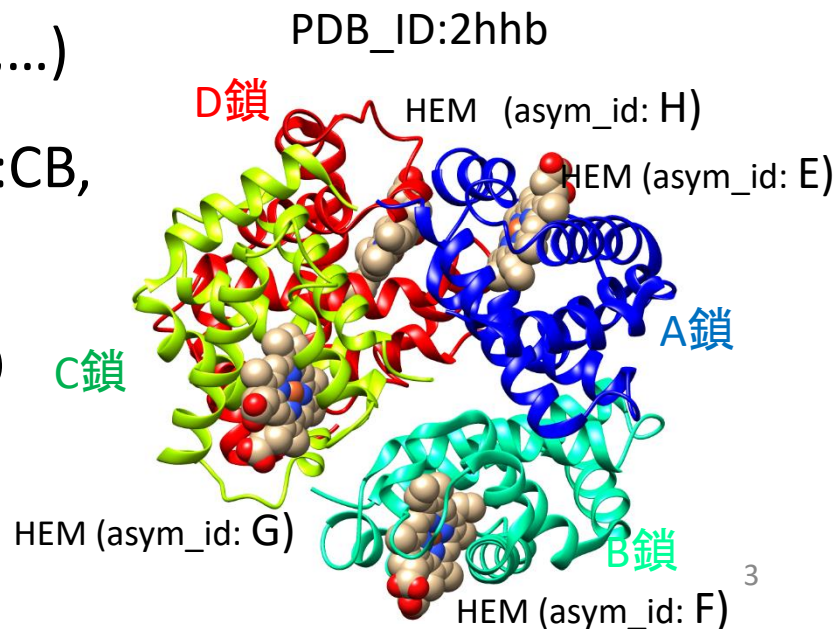
# Protein Data Bank (PDB)

- ・生体高分子の立体構造のデータベース
- ・現在のデータ数 (2021/09/22): 1,82,418件

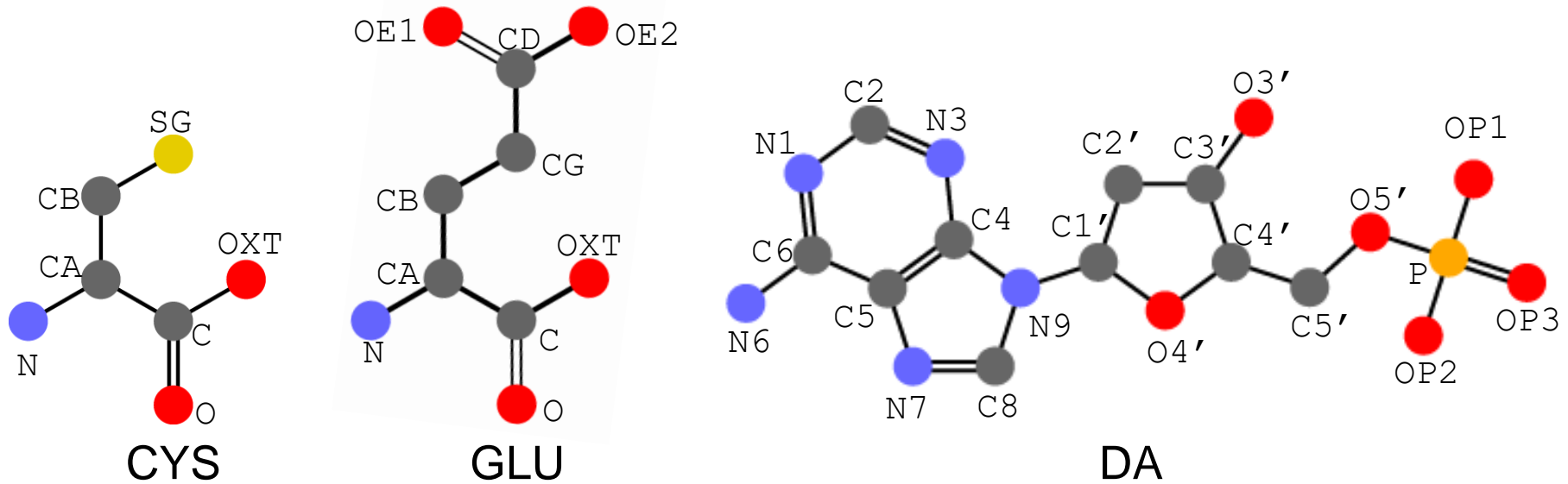
(taken from PDBj 2021/09/22)

X線結晶解析	NMR	電子顕微鏡	その他	合計
159849(87.6 %)	13487(7.4%)	8548 (4.7%)	534 (0.3%)	182418

- ・各エントリには**4文字の英数字のPDB\_ID**を持つ(2hhb, 1a5kなど)
- ・各高分子には、**鎖識別子**(A,B,C,D...)が付く
- ・各残基には**残基番号**がつく(1, 2, 3, 4,...)
- ・各原子には**固有の原子名**(C $\alpha$ :CA, C $\beta$ :CB, C $\gamma$ :CGなど)が付加
- ・低分子化合物には**化合物ID(残基名)**(ALA, HEM, ATP,...)が付く
- ・各分子ごとに**asym\_id**が付く



# 残基名と原子名



PDBで使われている原子名の例。重原子のみ示した。CYSはシステイン、GLUはグルタミン酸、DAは、DNAの構成要素となる2'-deoxyadenosine-5'-monophosphate.

※mmCIFでは、残基名はlabel\_comp\_idという名前、原子名はlabel\_atom\_idと呼ばれる

※残基名は、これまでは3文字以下。しかし、今後、拡張される予定

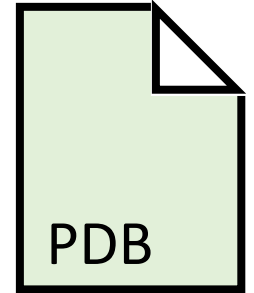
※Chemical Components Dictionaryというデータベースで管理

※PDBjのWEBサイトで、残基名で検索すると、これらの原子名の命名法を見ることができます

# 立体構造データのフォーマット

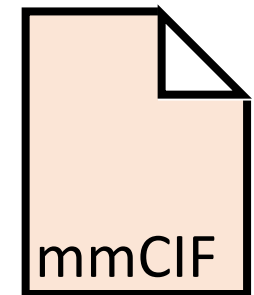
## • 旧PDBフォーマット

- ・PDB創立のころ(1971年)から使われていたフォーマット
- ・1行1原子。1行80文字の固定長のフォーマット。
- ・鎖識別子は1文字なので、62鎖を超えると扱えない。
- ・原子番号は5文字なので、99999個を超える原子は扱えない。
- ・座標以外のデータの記述は場当たりの



## • mmCIFフォーマット

- ・STARフォーマットの種類。可変長。
- ・キー・バリュー形式と表形式を併用
- ・値の長さは可変なので、鎖や原子数の上限はない。
- ・座標以外の様々なデータも、柔軟に記述できる
- ・2014年から、PDBの正規なフォーマットとなる
- ・2019年7月から、X線のデータ登録にmmCIFが必須となる
- ・主要な分子表示、モデリング、構造決定のソフトウェアは対応済み

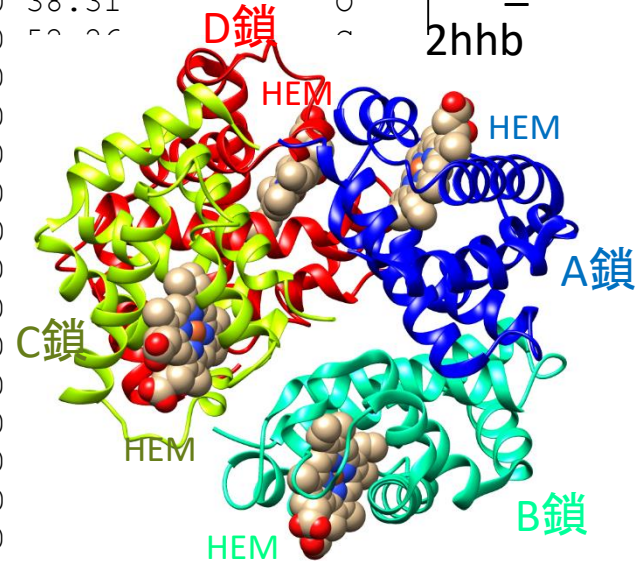


# 旧PDBフォーマットの立体構造データの例

PDB

PDB\_ID:  
2hhb

HEADER	OXYGEN TRANSPORT				07-MAR-84		2HHB				
TITLE	THE CRYSTAL STRUCTURE OF HUMAN DEOXYHAEMOGLOBIN AT 1.74 ANGSTROMS										
ATOM	1	N	VAL	A	1	6.130	16.559	4.905	1.00	41.29	N
ATOM	2	CA	VAL	A	1	6.870	17.784	4.702	1.00	41.33	C
ATOM	3	C	VAL	A	1	8.377	17.548	4.913	1.00	31.64	C
ATOM	4	O	VAL	A	1	8.820	16.980	5.922	1.00	38.31	O
ATOM	5	CB	VAL	A	1	6.345	18.763	5.731	1.00	50.00	~
ATOM	6	CG1	VAL	A	1	6.745	20.188	5.356	1.00		
ATOM	7	CG2	VAL	A	1	4.826	18.612	5.847	1.00		
ATOM	8	N	LEU	A	2	9.146	18.005	3.962	1.00		
ATOM	9	CA	LEU	A	2	10.599	17.914	4.153	1.00		
ATOM	10	C	LEU	A	2	11.062	19.085	5.062	1.00		
ATOM	11	O	LEU	A	2	10.829	20.254	4.744	1.00		
ATOM	12	CB	LEU	A	2	11.269	18.078	2.776	1.00		
ATOM	13	CG	LEU	A	2	10.986	16.983	1.769	1.00		
ATOM	14	CD1	LEU	A	2	11.735	17.168	0.427	1.00		
ATOM	15	CD2	LEU	A	2	11.276	15.630	2.404	1.00		
ATOM	16	N	SER	A	3	11.674	18.766	6.158	1.00		
ATOM	17	CA	SER	A	3	12.286	19.774	7.034	1.00		
ATOM	18	C	SER	A	3	13.529	20.322	6.334	1.00		
ATOM	19	O	SER	A	3	13.995	19.754	5.344	1.00	27.40	O
ATOM	20	CB	SER	A	3	12.719	19.060	8.326	1.00	23.83	C
ATOM	21	OG	SER	A	3	13.844	18.245	8.107	1.00	29.07	O
:											
HETATM	4389	CHA	HEM	A	142	8.634	7.898	-18.334	1.00	17.24	C
HETATM	4390	CHB	HEM	A	142	10.386	9.965	-14.276	1.00	23.17	C
HETATM	4391	CHC	HEM	A	142	8.337	6.403	-11.611	1.00	11.04	C
:											
:	原子番号	原子名	残基名	残基番号	鎖識別子	X座標	Y座標	Z座標	占有率	温度因子	元素名



80文字の固定長

パンチカード時代の名残り

# mmCIFフォーマットの立体構造データの例

mmCIF

```
data_2HHB
#
#_entry.id      2HHB
#
#_struct.title      'THE CRYSTAL STRUCTURE OF HUMAN DEOXYHAEMOGLOBIN AT 1.74 ANGSTROMS
#_struct.pdbx_descriptor  'HEMOGLOBIN (DEOXY)'
```

キー・バリュー形式

```
loop_
_atom_site.group_PDB
_atom_site.id
_atom_site.type_symbol
_atom_site.label_atom_id
_atom_site.label_alt_id
_atom_site.label_comp_id
_atom_site.label_asym_id
_atom_site.label_entity_id
_atom_site.label_seq_id
_atom_site.pdbx_PDB_ins_code
_atom_site.Cartn_x
_atom_site.Cartn_y
_atom_site.Cartn_z
_atom_site.occupancy
_atom_site.B_iso_or_equiv
_atom_site.pdbx_formal_charge
_atom_site.auth_seq_id
_atom_site.auth_comp_id
_atom_site.auth_asym_id
_atom_site.auth_atom_id
_atom_site.pdbx_PDB_model_num
ATOM 1 N N . VAL A 1 1 ? 6.130 16.559 4.905 1.00 41.29 ? 1 VAL A N 1
ATOM 2 C CA . VAL A 1 1 ? 6.870 17.784 4.702 1.00 41.33 ? 1 VAL A CA 1
ATOM 3 C C . VAL A 1 1 ? 8.377 17.548 4.913 1.00 31.64 ? 1 VAL A C 1
ATOM 4 O O . VAL A 1 1 ? 8.820 16.980 5.922 1.00 38.31 ? 1 VAL A O 1
ATOM 5 C CB . VAL A 1 1 ? 6.345 18.763 5.731 1.00 52.26 ? 1 VAL A CB 1
ATOM 6 C CG1 . VAL A 1 1 ? 6.745 20.188 5.356 1.00 52.75 ? 1 VAL A CG1 1
ATOM 7 C CG2 . VAL A 1 1 ? 4.826 18.612 5.847 1.00 58.75 ? 1 VAL A CG2 1
ATOM 8 N N . LEU A 1 2 ? 9.146 18.005 3.962 1.00 27.63 ? 2 LEU A N 1
ATOM 9 C CA . LEU A 1 2 ? 10.599 17.914 4.153 1.00 33.62 ? 2 LEU A CA 1
:
:
:
#
```

項目は[カテゴリ名].[アイテム名]

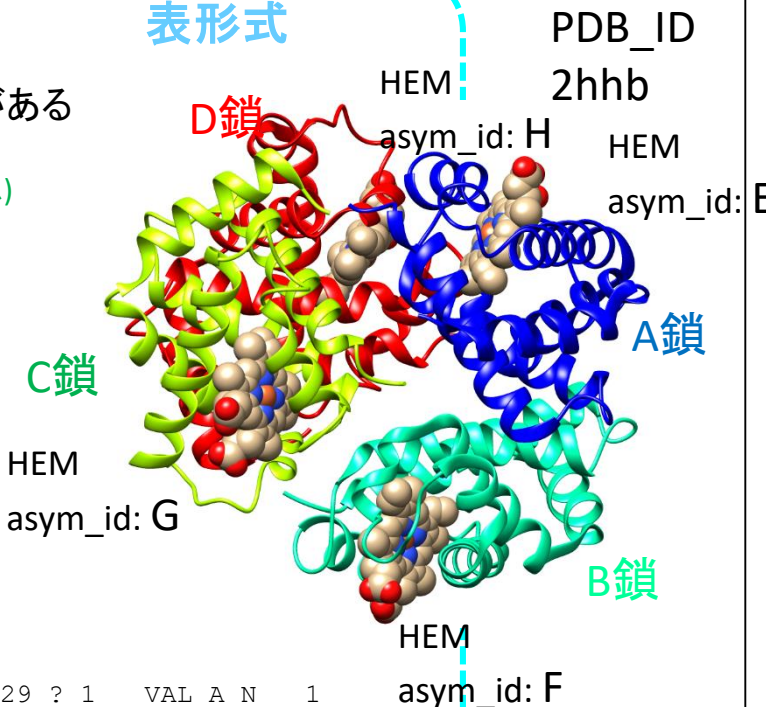
※2hhbの場合58個のカテゴリがある

データベース管理者が入力(label\_のアイテム)

原子座標

登録者が入力する(auth\_のアイテム)

表形式



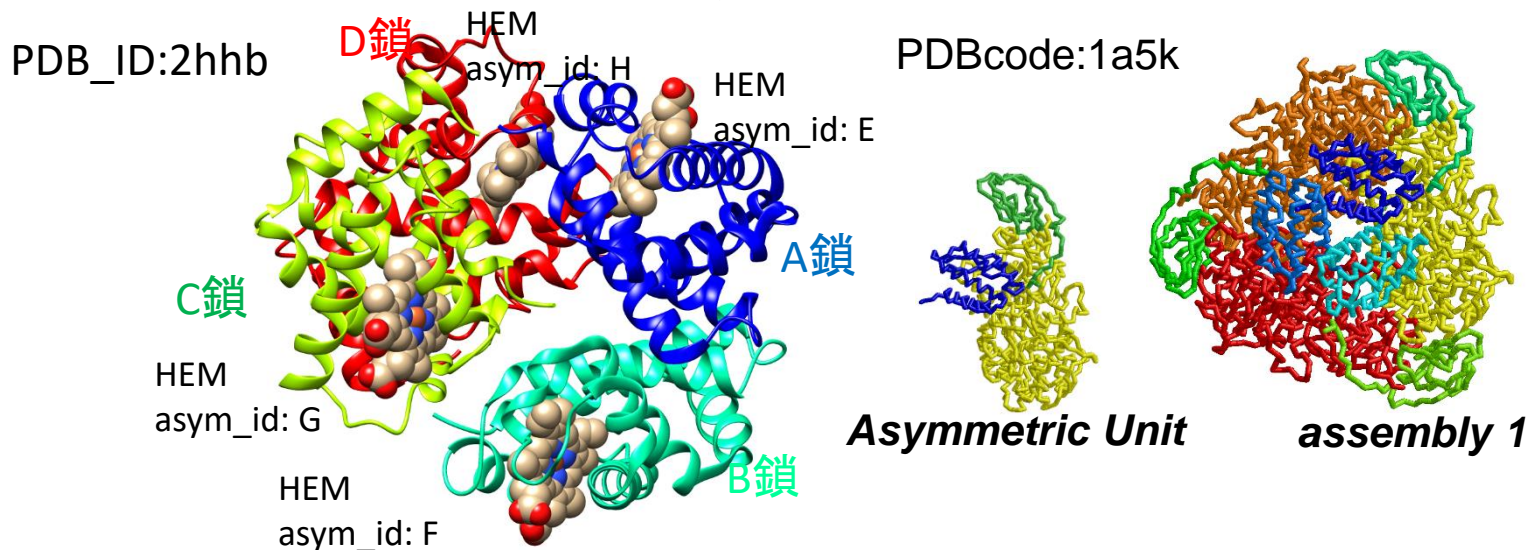
原子番号 元素名 原子名 残基名 残基番号 鎖識別子 x座標 y座標 z座標 占有率 温度因子 残基番号 残基名 原子名 asym\_id(鎖識別子)



# mmCIFファイルの特長

- STARフォーマット。キーバリュー形式と表形式の併用で、人間にも比較的読みやすい。
- [カテゴリ名].[アイテム名]で整理された項目名
- カテゴリatom\_site以外に、50-80ほどのカテゴリ(表)
- 鎖識別子、原子番号などの文字数は無制限
- 残基番号、鎖識別子については、読みやすさと統一性を両立するため、著者入力とデータベース管理者入力の2種が書かれる。
- Biological unitなど、非対称単位以外の複合体(assembly)の生成法が正確に記載されている。

Relionの出カファイルもSTARフォーマットです！





# mmCIFを使わなければならない場合

## 1. 巨大構造を扱う場合は、mmCIFを読み込むべき

巨大な構造の場合、mmCIF形式のみで公開されています。

鎖識別子が2文字以上の場合、巨大でなくてもmmCIFだけの構造はあります

## 2. 立体構造を扱うプログラムの開発者は、入力をmmCIFに対応するよう改良すべき。(出力は任意)

計算結果の出力は、旧PDBフォーマットでもかまいません。

## 3. PDBデータベース全体に対して、統計的・包括的な解析を行う場合は、mmCIFをベースにすべき

mmCIFのデータベース管理者による残基番号、鎖識別子は、統一的な基準で決まっているので、PDBデータベース全体の統計をとる場合に向いています。

# PDB形式変換不可エントリー : mmCIFだけで公開されているエントリー

182418  
件を2021-09-22に公開中

English 日本語 简体中文 繁體中文 한국어

Search pdbj.org

2021年9月22日時点で  
2259エントリー

### PDB形式変換不可の PDB エントリー

ID	Title	Author	Registration Date	Release Date	Final Update Date	Method	Resolution	Reference
1BOS	SHIGA-LIKE TOXIN COMPLEXED WITH ITS RECEPTOR	Ling, H, Boodhoo, A, Hazes, B, Cummings, M.D, Armstrong, G.D, Brunton, J.L, Read, R.J.	1998-01-13	1999-02-02	2020-07-29	X-RAY DIFFRACTION	2.8 Å	Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. <i>Biochemistry</i> , 37, 1998
1VVJ	Crystal Structure of Frameshift Suppressor tRNA SufA6 bound to Codon CCC-G on the Ribosome	Maehigashi, T, Dunkle, J.A, Dunham, C.M.	2013-05-24	2014-08-06	2014-12-10	X-RAY DIFFRACTION	3.4400011574 Å	Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. <i>Proc.Natl.Acad.Sci.USA</i> , 111, 2014
1VY4	Crystal structure of the Thermus thermophilus 70S ribosome in the pre-attack state of peptide bond formation containing acylated tRNA-substrates in the A and P sites.	Polikanov, Y.S, Steitz, T.A, Innis, C.A.	2014-05-13	2014-08-20	2019-07-03	X-RAY DIFFRACTION	2.6 Å	A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. <i>Nat.Struct.Mol.Biol.</i> , 21, 2014

**PDB形式変換不可エントリー**

クイックリンク  
ヘルプ  
PDB形式変換不可エントリー  
グループ登録エントリー  
化合物一覧

検索結果  
全ヒット件数: 2259  
表示件数: 25  
表示順: PDBID昇順 (0→9, a→z)

表示順  
PDBID昇順 (0→9, a→z)  
PDBID降順 (z→a, 9→0)  
登録日の古い順  
登録日の新しい順  
公開日の古い順  
公開日の新しい順  
分解能の高い順

Lists  
更新された PDB エントリー  
新しい PDB エントリー  
廃止された全ての PDB エントリー  
廃止された最新の PDB エントリー  
PDB形式変換不可の PDB エントリー  
PDBグループ登録エントリー  
X線構造因子データを含む全てのPDBエントリー  
X線構造因子データを含む新しいPDBエントリー  
X線構造因子データが更新されたPDBエントリー  
X線構造因子データを含む全ての廃止されたPDB エントリー  
X線構造因子データを含む最新の廃止された

鎖識別子(auth\_asym\_id)を2文字以上使うと、原子数が小さくてもmmCIFだけの公開となり、「PDB形式変換不可エントリー」とみなされる。

# なんとなく構造を眺めるだけなら...

- PDBのウェブページをクリックして、3D構造を眺めるだけなら、ファイルフォーマットを意識する必要はありません
- PyMOL、UCSF Chimeraなど最近のビューワは、直接PDBサーバから構造を読みに行く機能("fetch")があるので、ファイルフォーマットを意識する必要はありません
- ただし、残基名、残基番号、鎖識別子などの事情はある程度理解しておいたほうがいいかもしれません

# PDB IDの探し方

## (1) 文献検索 構造を決定した論文には必ずPDB IDが書いてある

GoogleやPubMedなどで、関心があるタンパク質・ファミリー名といっしょに、3D structure, X-ray, NMR, Cryo-EMなどのキーワードとともに文献を検索

## (2) PDBjのキーワード検索



PDBjのトップ画面の上のフォームにタンパク質名、ファミリー名などを入力

詳細検索(PDBjのRDBを直接検索する)も可能

## (3) PDBjのアミノ酸配列検索

関心のあるタンパク質のアミノ酸配列から配列相同性検索で類似した蛋白質の立体構造を検索

配列が類似した蛋白質は構造も類似しているので、ホモロジーモデリングの鋳型としても使える

検索サービス

- ヘルプ
- PDB検索 (PDBj Mine)
- PDB詳細検索
- 化合物検索 (Chemie)
- BMRB検索
- Sequence-Navigator
- DASH
- EM Navigator
- Omokage検索
- SeSAW
- wwPDB/RDF
- NBDC RDFポータル
- 未公開エントリーのステータス

Sequence Navigator

Mode: PDB entry Custom sequence

Sequence: MDMFQKVERIGEGTYGVVYKAKNRETGQLVALKKIRLDLEMEGVFSTAIREISLLKELKH  
PNIVRLLDVVHNERKLYLVFEFLSQDLKKYMDSTPGSELPLHLIKSYLFLQLQGVSPCHS  
HRVHRLDKPQNLLINELGAIKLADFLARAFVGLRTYTHEVVTLWYRAPEILLGSKFY  
TTAVDIWSIGCIFAEMVTRKALFPGDSEIDQLFRIFRMLGTSPSEDTPGVTQLPDYKGSF  
PRWTRKGLLEEIVPNLEPEGRDLLMQLLOYDPSQRITAKTALAHPYFSSPEPSPAARQVVL  
QRFH

Sequence type: Protein Nucleic acid

Clustering: No clustering

Find homologues

# HOMCOSを利用した配列からのPDB IDの検索

(3) 類似蛋白質の立体構造が複合体の種類別に表示される

https://homcos.pdbj.org

HOMCOS: 相同複合体の検索・モデリングサーバ

(Go to English page) ヘルプページ

HOMCOS(HQMology modeling of CQmplex Structure) は、PDBに収録されている複合体の立体構造データを利用して、分子の類似性・相同性から、構造未知の分子ペアの構造を予測するためのサーバです。アミノ酸配列や化学構造から、PDB内から複合体の立体構造が解けている他の分子を探索したり、PDB内の構造を鏡型にして複合体の立体構造を予測するサービスがあります。タンパク質の配列種別性検索にはBLASTや、化合物の構造種別性検索にはKCOMBIを使用しています。

サービス

(1) 「タンパク質に対する結合分子の検索」を選ぶ

結合分子の検索	タンパク質に対する結合分子の検索	アミノ酸配列	1本のアミノ酸配列をクエリとして、それと類似したタンパク質と結合している分子を検索します
	化合物に対する結合タンパク質の検索	化学構造	1つの化合物構造をクエリとして、それと類似した化合物と結合しているタンパク質を検索します
複合体立体構造のモデリング	ホモ多量体のモデル	アミノ酸配列	1本のアミノ酸配列をクエリとして、そのホモ多量体の立体構造をホモジーマデリングによって予測します
	ヘテロ多量体のモデル	2本のアミノ酸配列	2本のアミノ酸配列をクエリとして、そのヘテロ多量体の立体構造をホモジーマデリングによって予測します
	化合物-タンパク質複合体のモデル	アミノ酸配列と化学構造	1本のアミノ酸配列と1つの化学構造をクエリとして、その複合体の立体構造をホモジーマデリングによって予測します

[Download HOMCOS data]: Release of PDB:20210908, Release of UniProt:2021\_03 Release of KEGG:2021-09-06 3D of KEGG COMPOUND and DRUG

新型コロナウイルスSARS-CoV-2に関連するタンパク質と化合物のページを作りました!(PDB UPDATE:20210908)

https://homcos.pdbj.org/cgi-bin/prot\_sch\_inp.cgi?&LANG=ja

HOMCOS: タンパク質に対する結合分子の検索

(Go to English page) ヘルプページ

入力したタンパク質のある分子と類似した分子の検索

(2) UniProtIDかアミノ酸配列を入力

アミノ酸配列... (氨基酸配列を入力してください。1文字表記のアップロードのどちらの方法で入力してください。)

タンパク質配列のID (UniProt:ID/AC, INSDC/RefSeq:protein_id)	NCAP_SARS2 (CDK3_HUMAN, Q00526, AAV40830.1, NP_001249.1)
アミノ酸配列 (一文字表記)	<pre> LPYGANKDGIWVATEGALNTPKDHIGTRNPANNAALVQLPQGITLPKGFYAEGRGGG QASSRSSRSRNSRNSTPGSSRGTSPARMAGNGGDAALLLDLLRNOLESKMSGKGGQ IQGGTIVTKKSAEASKKPRKRTATKAYNVOAFGRGPEQTGQNFQDLEIRGDTYKH WPOIADFPASAFYFGRSRIQMEVTPSGTILTYGAIKLDKDPKDFKDDVILLNKHIDAY KTFPTPEPKKDKKKKADETQALPRQKRGQITVLLPAADLDFSKLQGSMSADSTQA                     </pre>

https://homcos.pdbj.org/cgi-bin/prot\_sch\_conbars\_demo.cgi?QUINIPROT\_AC...

Contact Molecules for Homologous Proteins

Summary Bars[0.0 %]

seq\_id(%) [0] [30] [40] [50] [60] [70] [80] [90] [95] [100]

PID	QueryLength	Homologous Sequence in PDB	UniProt Query	TITLE
2670477	419	142	P0DTC9(NCAP_SARS2)	RecName: Full=Nucleoprotein ; Short=N;AltName: Full=Nucleo... Short=Protein N ;

QUERYSEQ: MCEWDPQWNPRIFFGSPDSSTGSDNNGERSGARNRRRQPLPINTACHTALTNQHEQLKFRGRRDPIVNTMSFSDQIYIQRATRIIRGGSDNKLSPRMFVYLGTFEAGLPDQANWIDLIWIKETGALNTPKDHIGTRNPANNAALVQLPQGITLPKGFYAEGRGGGQASSRSSRSRNSRNSTPGSSRGTSPARMAGNGGDAALLLDLLRNOLESKMSGKGGQIQGGTIVTKKSAEASKKPRKRTATKAYNVOAFGRGPEQTGQNFQDLEIRGDTYKHWPOIADFPASAFYFGRSRIQMEVTPSGTILTYGAIKLDKDPKDFKDDVILLNKHIDAYKTFPTPEPKKDKKKKADETQALPRQKRGQITVLLPAADLDFSKLQGSMSADSTQA

[BLAST file for PDB] (plain) (bar) (multiple alignment) [BLAST for UniProt: (plain) (bar) (multiple alignment) (PSSM file)]

MONOMER 単量体

pdb_id	a <sup>1</sup>	identity[%] <sup>2</sup>	description
6y13	A	99.3	NCAP_SARS2 Nucleoprotein
6yun	B	100.0	NCAP_SARS2 Nucleoprotein

HETERO ヘテロ複合体

contact mol	homologue
a <sup>3</sup> description	a <sup>4</sup> identity[%] <sup>5</sup> No
7zcs[1] B monoclonal antibody chain H[216 aa]	A 100.0 /100.0 7 /7
7zcs[1] C monoclonal antibody chain L[213 aa]	A 100.0 /100.0 15 /15

NUCLEOTIDE 核酸との複合体

contact mol	homologue
a <sup>3</sup> description	a <sup>4</sup> identity[%] <sup>5</sup> No
7zct[1] B ssRNA	A 100.0 /100.0 28 /28
7zcs[1] B RNA (5'-R(P*(C*P*AP*CP*UP*GP*AP*C)-3')	A 100.0 /99.3 5 /5

COMPOUND 化合物との複合体

contact mol	homologue
a <sup>3</sup> description	a <sup>4</sup> identity[%] <sup>5</sup> No
DJU N,N-dimethyl-1-(5-phenylmethoxy-1H-indol-3-yl)meth...	B 83.3 /61.2 6 /6

# PyMOL

- ・Warren Lyford Delano (1972-2009) によって2000年ごろから開発された分子描画ソフト。2003年からDelano Scientific社によって配布・販売。
- ・2009年にDelano氏が死去。その後は、Schrödinger社で維持・配布
- ・PyMOLは、商品(commercial product)だが、そのソースコードは無料で公開。つまり、ソースコードをダウンロードして、自分でコンパイルして使うのは無償。
- ・Webページから試用版(ライセンス無し版)の実行バイナリのダウンロードは可能(Windows, macOS, Linux) ただし、試用版では、画面に“No License File – For Evaluation Only”のメッセージが常に画面に表示される。
- ・正式版ライセンスは有料。Academic 1人/1年で\$99。他にもいろいろなプラン。
- ・教育用無料ライセンスでは“Education-use-only”のメッセージが常に画面に表示される。



# PyMOL2.5の配布法・ライセンスの種類

PyMOLは、オープンソースの分子ビューアソフトで、現在はSchrödinger社によって開発・管理が行われています。配布には、オープンソース版とIncentive PyMOLに分かれています。Incentive PyMOLにはいくつかのライセンスの種類があります。

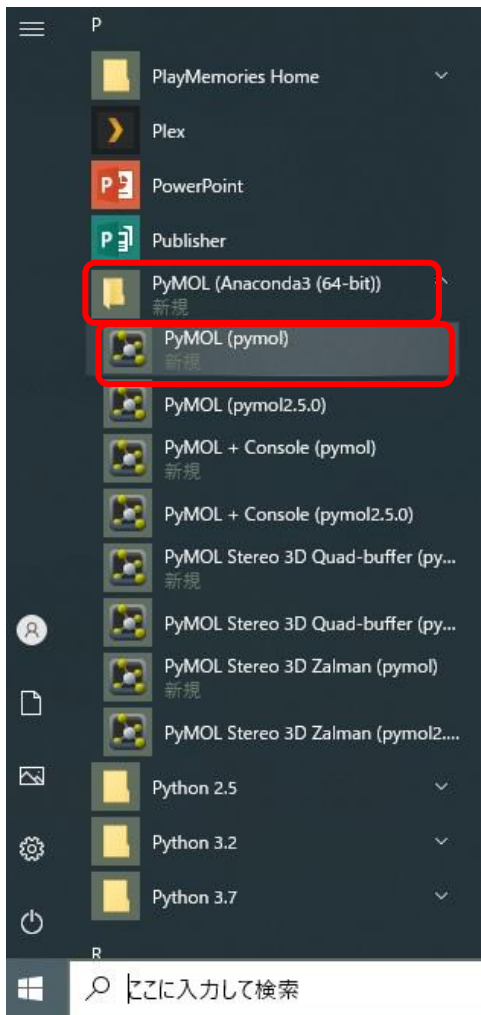
配布法	ライセンス	費用	説明
オープンソース	----	無料	機能制限あり。ソースのみ無料配布のため、自分でコンパイルするか、コンパイル済みバイナリを公開している第3者のサイトからダウンロードする必要。
incentive	ライセンスなし	無料	機能制限あり。30日間の限定使用。“No License File”の表示
	Commercial	要問合せ	
	Government	\$199/1人, 年	
	Academic/Non-profit	\$99/1人, 年	
	Student/Teacher (Education-use-only)	無料	機能制限あり。“Educational Use Only”の表示。

今回の講習会用だけの使用なら、「ライセンスなし」か「Student/Teacher」ライセンスを使ってください

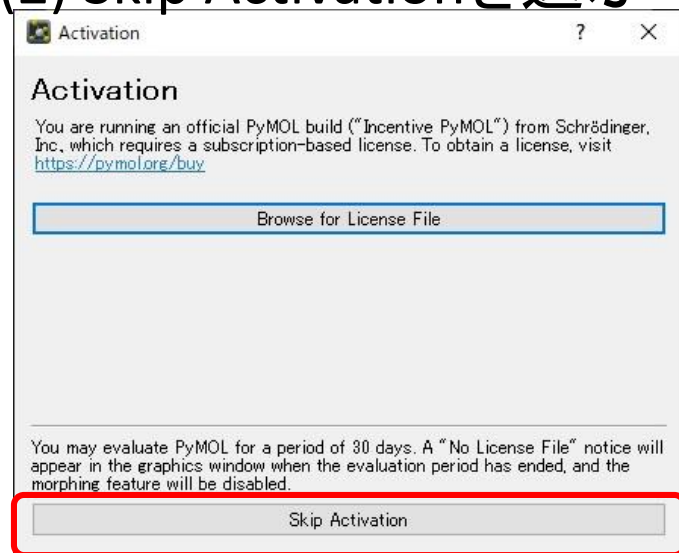


# PyMOLの起動

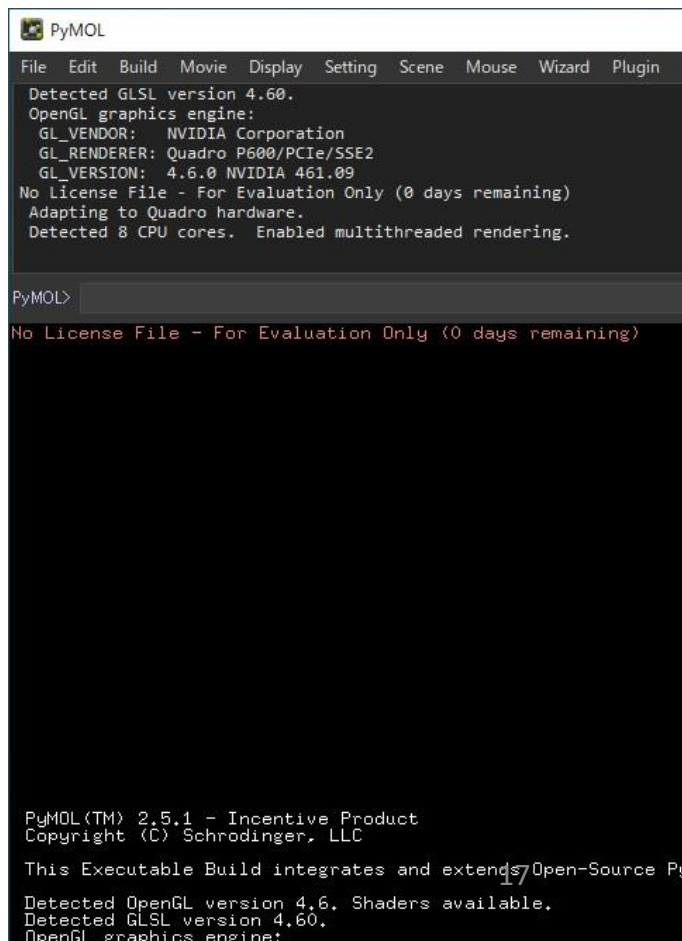
(1) スタートメニューからPyMOL(pymol)を選ぶ



(2) Skip Activationを選ぶ



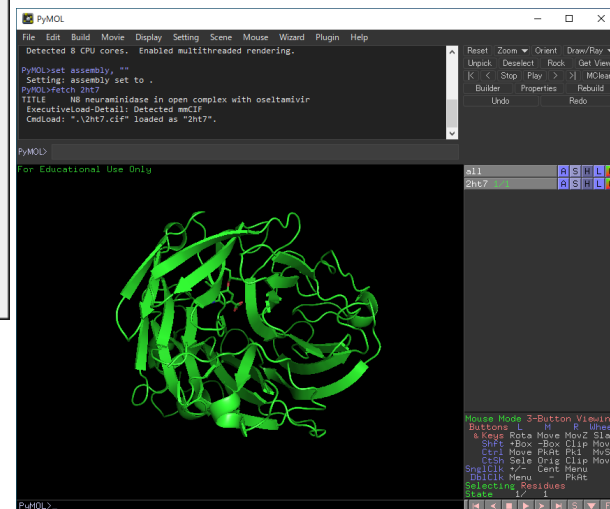
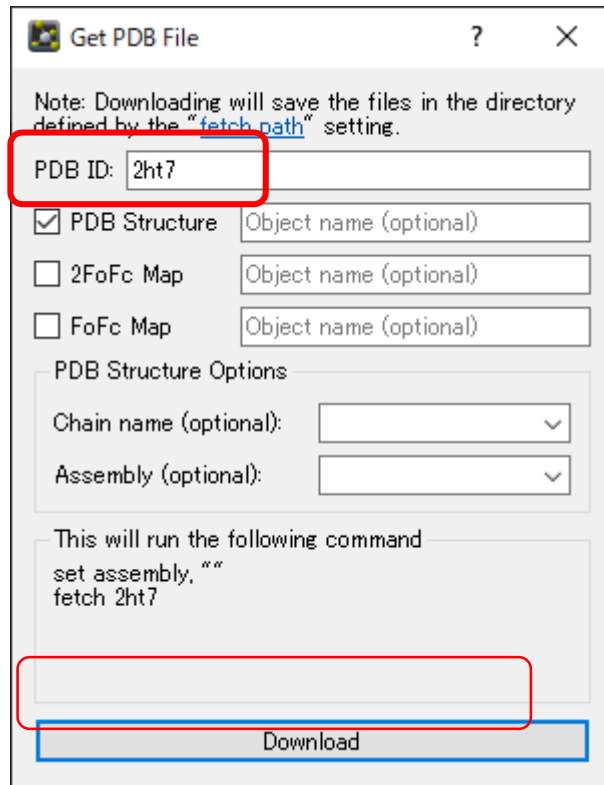
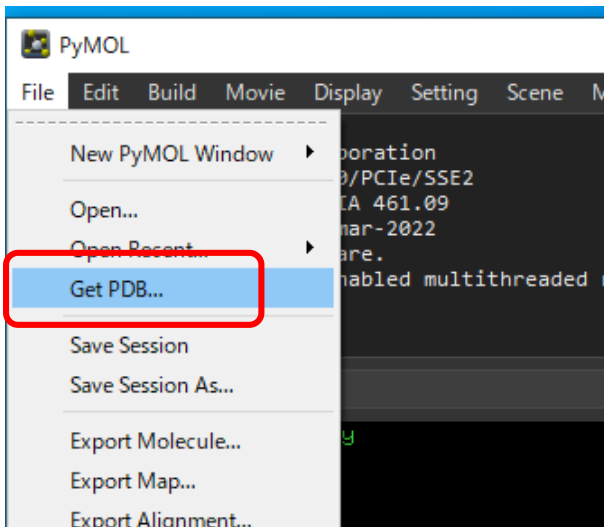
(3) 「ライセンスなし」のPyMOLが起動



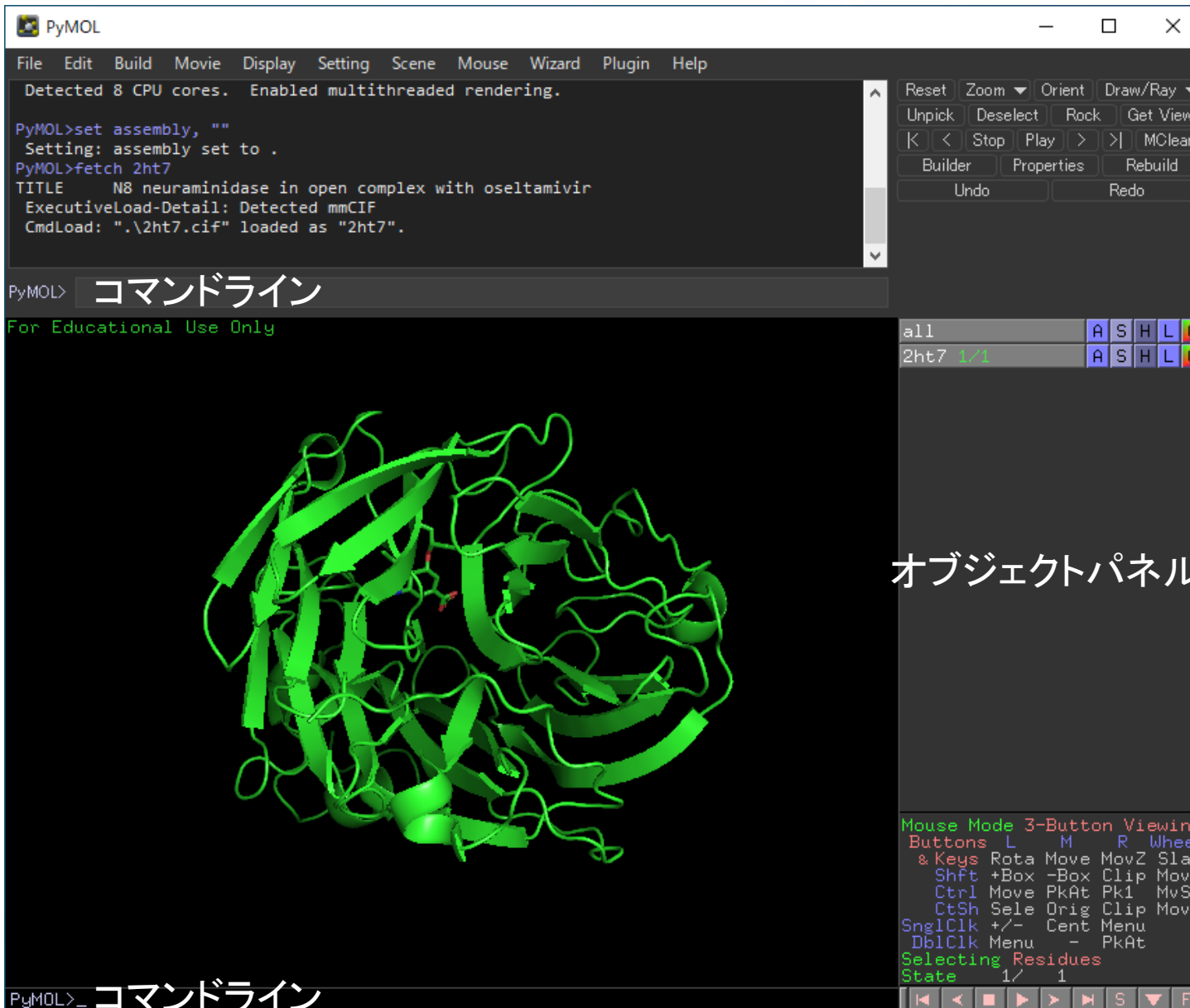
※ただし、機能限定版。使用日数の制限もあり。また、画面に常に  
“No License File”  
と表示される。

# PDB ID:2ht7をGetPDBでfetchする

(1) [File]→[GetPDB...]を選ぶ (2) PDB ID:に2ht7と入力。[Download]をクリック



# PyMOLのウィンドウ構成



Action  
Show  
Hide  
Label  
Color

オブジェクトパネル

# マウスの使い方

表示法	<i>PyMOL</i>	<i>UCSF Chimera</i>	<i>Molmil</i>
分子の回転	左ボタンで画面をドラッグ	左ボタンで画面をドラッグ	左ボタンで画面をドラッグ
分子の並進	中ボタン(ホイール)で画面をドラッグ	中ボタン(ホイール)で画面をドラッグ	(1)中ボタン(ホイール)で画面をドラッグ (2)[Shift]キー+左ボタン
ズームイン・アウト	右ボタンで画面をドラッグ	(1)右ボタンで画面をドラッグ (2)ホイールを回す	(1)右ボタンで画面をドラッグ (2)ホイールを回す
断面の表示 (slab)	ホイールを回す	[Tools]→[Viewing Controls]→[Side View]	[View]→[Configure slab]
残基名・原子名の確認	左ボタンで原子をクリックすると、アップコントロール・ウィンドウに表示	画面上で原子の上にもウスポインタをしばらくかざしておくと、原子名のラベルが表示	左ボタンで原子をクリックすると、画面下に表示



# PyMOL 選択コマンド

書式	例	意味
[実行], chain [鎖]	color red, chain A	A鎖を赤に
[実行], resn [残基名]	color red, resn CYS	システインを赤に
[実行], atom [原子名]	color red, atom CB	C $\beta$ 原子を赤に
[実行], resi [番号]	color red, resi 104	104番目の残基を赤に
[実行], resi [番号]+[番号]	color red, resi 104+212	104番目と212番目の残基を赤に
[実行], resi [番号]-[番号]	color red, resi 104-212	104~212番目の残基を赤に
[実行], chain [鎖] & resn [残基名] & atom [原子名]	color red, chain A & resn CYS & atom CB	A鎖のシステインのC $\beta$ 原子を赤に
[実行], [鎖]/[残基名]/[原子名]	color red, A/CYS/CB	A鎖のシステインのC $\beta$ 原子を赤に
[実行], chain [鎖]& resi=[番号]	color red, chain A & resi 104	A鎖の104番目を赤に
[実行], [鎖]/[番号]/	color red, A/104/	A鎖の104番目を赤に
[実行], [条件] around [距離]	color red, resn ATP around 5	ATPから5Å以下の原子を赤に
[実行], byres([条件] around [距離])	color red, byres (resn ATP around 5)	ATPから5Å以下の残基を赤に
[実行], [条件] & [条件]	color red, polymer & resi 104	高分子の104番目を選択
[実行], [条件]   [条件]	color red, resn SER   resn THR	セリンかスレオニンを赤に
[実行], [条件] [値]+[値]	color red, resn SER+THR	セリンかスレオニンを赤に
[実行], polymer	color red, polymer	高分子を赤に
[実行], organic	color red, organic	有機低分子化合物を赤に
[実行], solvent	color red, solvent	水分子を赤に
select [選択原子名],[条件]	select actsite, resi 104	104番目をactsiteと命名

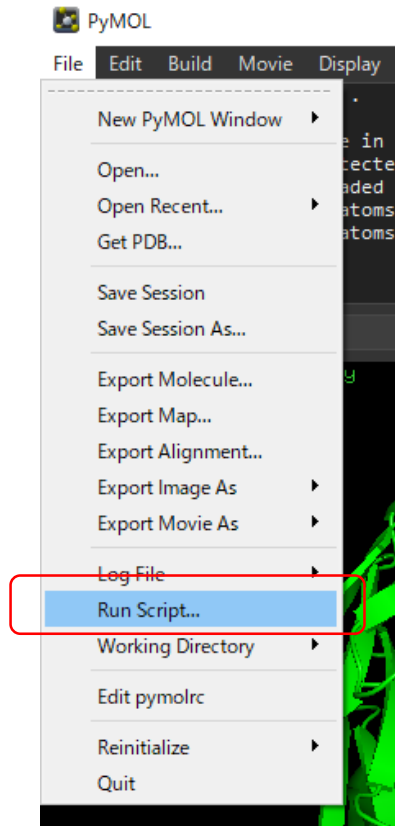
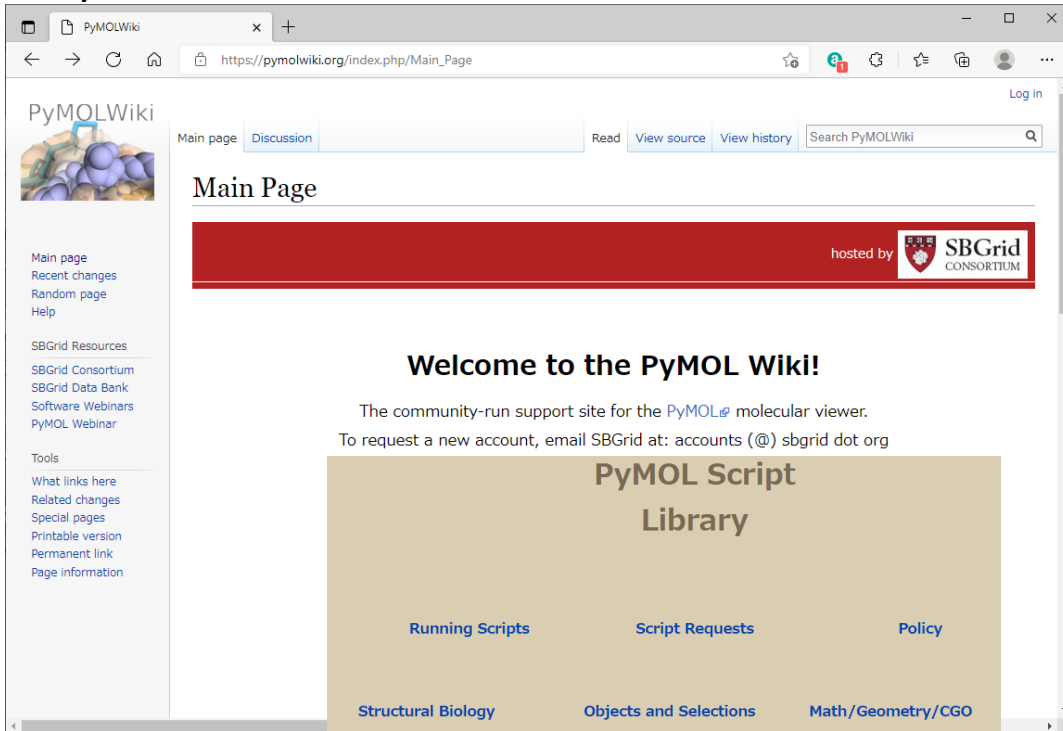
# PyMOL 実行コマンド

書式	例	意味
<code>show</code> [表示法]	<code>show sphere</code>	球モデルの非表示
<code>hide</code> [表示法]	<code>hide sphere</code>	球モデルの表示
※ [表示法]は、球: <code>sphere</code> 、線: <code>lines</code> 、スティック: <code>sticks</code> 、リボン: <code>ribbon</code> 、カートゥーン: <code>cartoon</code> 、表面: <code>surface</code> を使うことができる。		
<code>color</code> [色]	<code>color blue</code>	青色にする
<code>util.cbag</code>		元素ごとに色分けする
<code>bg_color</code> [色]	<code>bg_color white</code>	背景を白に
<code>rotate</code> [xyz] [回転角(°)]	<code>rotate y 180</code>	Y軸のまわりに180° 回転
<code>reset</code>		分子を元の向きに戻す

# PyMOL Wiki

[https://pymolwiki.org/index.php/Main\\_Page](https://pymolwiki.org/index.php/Main_Page)

PyMOLのマニュアル、チュートリアルに加え、様々なPyMOL Scriptのライブラリがある



Pythonスクリプトをダウンロードして  
[File]→[Run Script...]で選択して実行すると、  
新しいコマンドが使えるようになる

# UCSF Chimera

UCSF CHIMERA

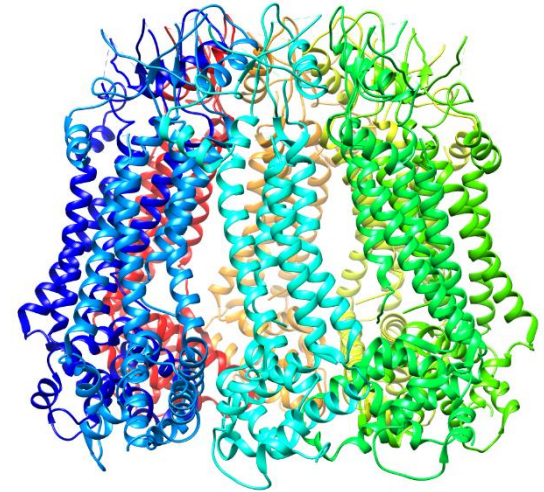
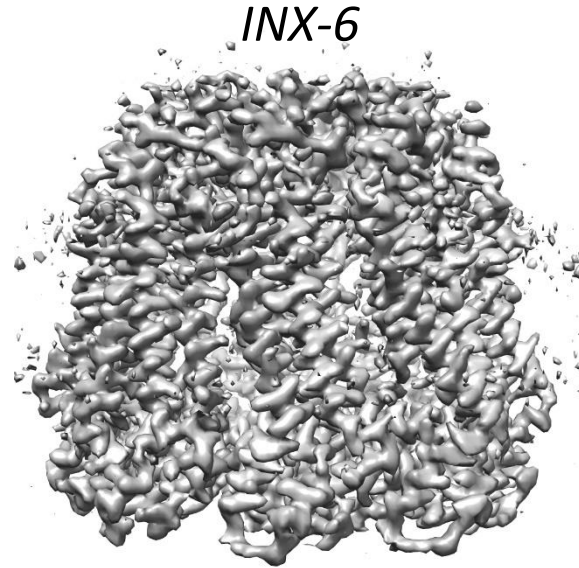
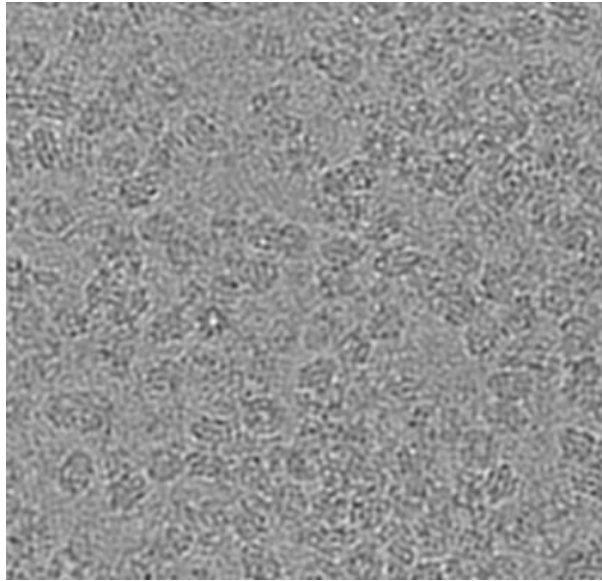
an Extensible Molecular Modeling System

- ・UCSF (カリフォルニア州立大サンフランシスコ校)の Thomas Ferrin の研究室で開発
- ・NIH(アメリカ国立衛生研究所)の研究費を用いて開発
- ・非商用(Non-commercial)な使用のみ無償。商用目的は有償で別途ライセンス契約を結ぶ必要
- ・Windows版、Mac版、Linux版の実行バイナリを配布
- ・電子顕微鏡の3Dマップの表示・操作の機能が充実
- ・次期プログラムChimera Xの開発が進行中



2014/02/19 Ferrin研でTom Goddardと

# Data processing for EM Single Particle Analysis



3D reconstruction

Atomic modeling

2D images

3D image

Atomic model

EMPIAR-10291

300 images

3710 x 3838 x 1; 32 bit real

15.9 GB

**EMPIAR**

**EMPIAR-PDBj**

300 movies

3710 x 3838 x 30; 32 bit real

(not registered in EMPIAR.

personal communication)

478 GB

180 x 180 x 180 voxels

emd\_9973.map

(3.6Å)

22.24 MB

**EMDB**

20840 atoms

PDB\_ID: 6kfh

1.7 MB

**wwPDB**

# 電顕で決定された原子モデル3Dマップを重ねて観察

PDB\_ID: 2c7e EMDB\_ID:1047 (9.7 Å; 128<sup>3</sup> voxels)

contourLevel : 0.0839

UCSF Chimeraを起動



2c7e - 概要 - 日本蛋白質構造データバンク

https://pdbj.org/mine/summary/2c7e

Protein Data Bank Japan

2C7E

REVISED ATOMIC STRUCTURE FITTING INTO A GROEL(D398A)-ATP7 CRYO-EM MAP (EMD 1047)

ダウンロード

- Sequence (fasta)
- PDBx/mmCIF
- PDBML (ヘッダのみ (no-atom))
- PDB形式 (全ての情報)
- 検証レポート (PDF)

構造

非対称単位を表示

他のデータベース情報

- wwPDB DOI Landing Page
- Yorodumi
- CATH
- VAST
- PISA
- eF-site
- 2c7e-A,B,C,D,E,F,G,H,I,J,K,L,M,N
- 2c7e-A
- 2c7e-B
- 2c7e-C
- 2c7e-D
- 2c7e-E
- 2c7e-F
- 2c7e-G
- 2c7e-H
- 2c7e-I
- 2c7e-J

2C7E の概要

エントリー-DOI [10.2210/pdb2c7e/pdb](https://doi.org/10.2210/pdb2c7e/pdb)

関連するPDBエントリー [1PF9](#) [1SS8](#) [1SVT](#) [1SX3](#) [1SX4](#) [1XCK](#)

EMDBエントリー [1047](#)

分子名称 60 KDA CHAPERONIN, POTASSIUM ION, MAGNESIUM ION, ... (5 entities in total)

機能のキーワード cell cycle, atp-binding, chaperone, chaperonin, d398a, hp60 class, cell division, nucleotide-binding, phosphorylation

由来する生物種 ESCHERICHIA COLI

タンパク質・核酸の鎖数 14

化学式組合計 [804631\\_83](#)

構造登録者 [Ranson, N.A., Farr, G.W., Roseman, A.M., Gowen, B., Fenton, W.A., Horwich, A.L., Saibil, H.R.](#) (登録日: 2005-11-22, 公開日: 2006-02-16, 最終更新日: 2019-10-23)

主引用文献 [Farr, G.W., Fenton, W.A., Gowen, B., Horwich, A.L., Ranson, N.A., Roseman, A.M., Saibil, H.R. ATP-Bound States of Groel Captured by Cryo-Electron Microscopy Cell \(Cambridge, Mass.\), 107:869-, 2001 PubMed: 11779463 DOI: 10.1016/S0092-8674\(01\)00617-1 主引用文献が同じPDBエントリー](#)

実験手法 ELECTRON MICROSCOPY (9.7 Å)

UCSF Chimera

File Select Actions Presets Tools Favorites Help

Scenes

Recent Data Sources

Save scene...

Named Selections

Name current selection...

Browse... Fetch... Edit

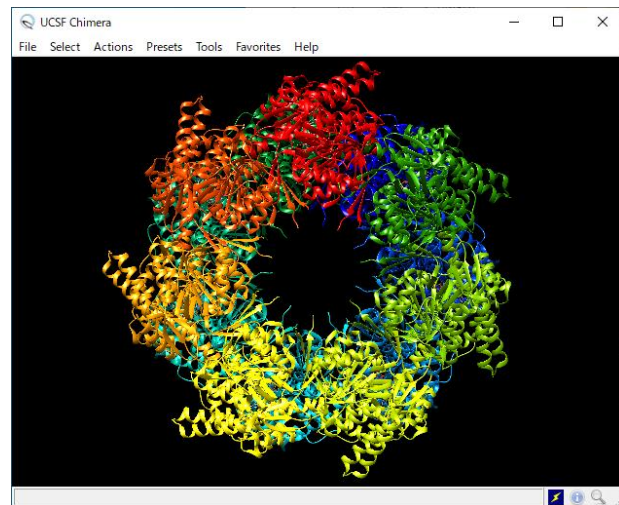
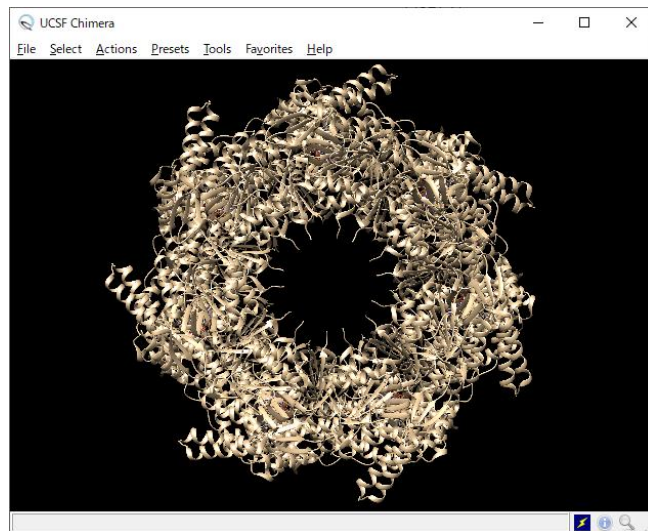
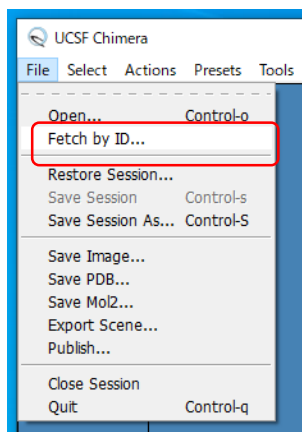


# 原子モデル(PDB ID:2c7e)をfetchして表示

(1)[File]→[Fetch by ID]

(3) 2c7eの原子モデルが表示

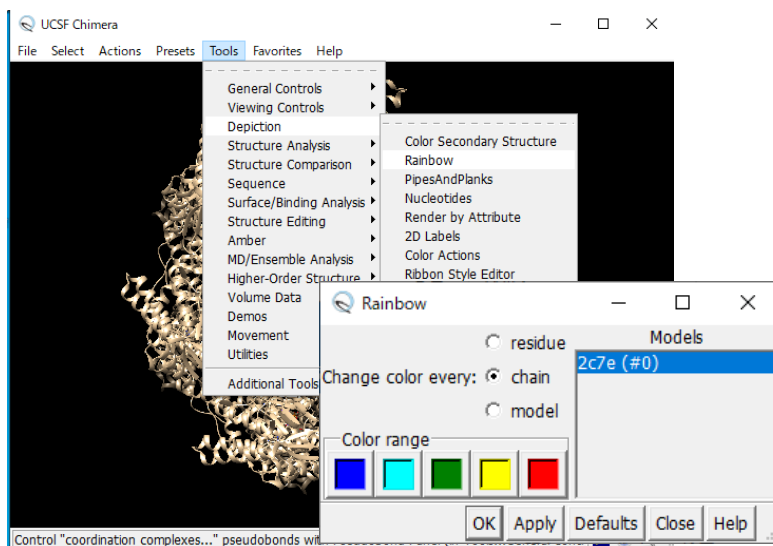
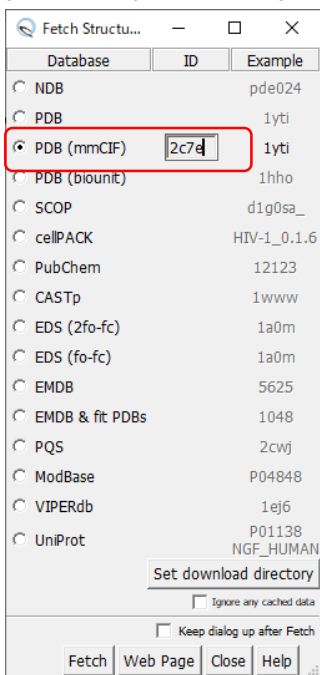
(6) 鎖ごとに色分け



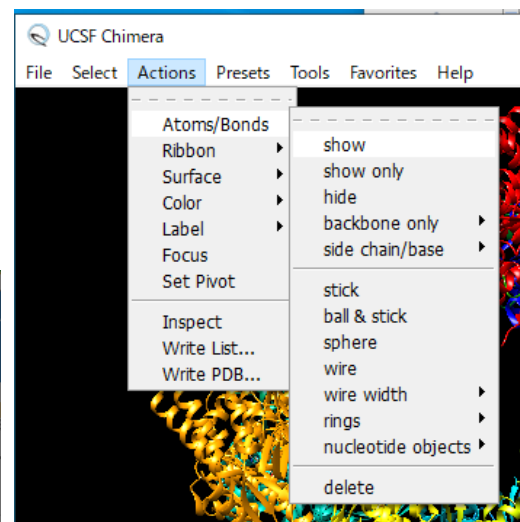
(2)[PDB (mmCIF)]に2c7e

(4)[Tools] → [Depiction] → [Rainbow]

(7)[Actions] → [Atoms/Bonds] → [Show]



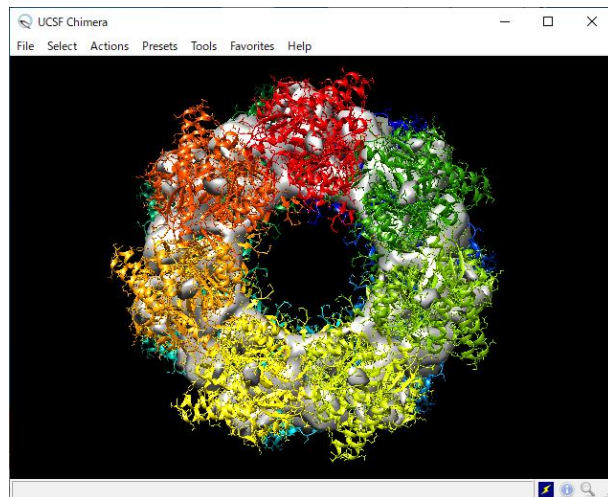
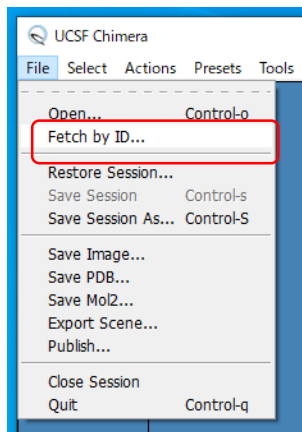
(5)[chain]を選んで [OK]



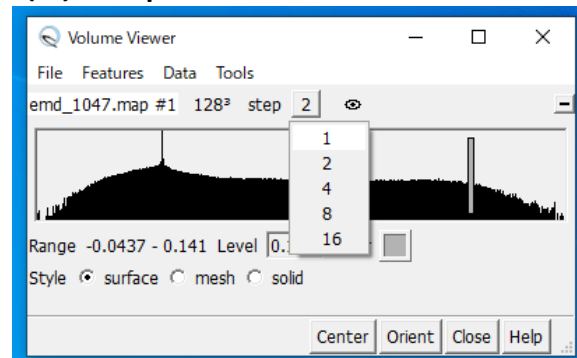
→原子がワイアで表示される

# 3Dマップ(EMDB:1047)をfetchして表示

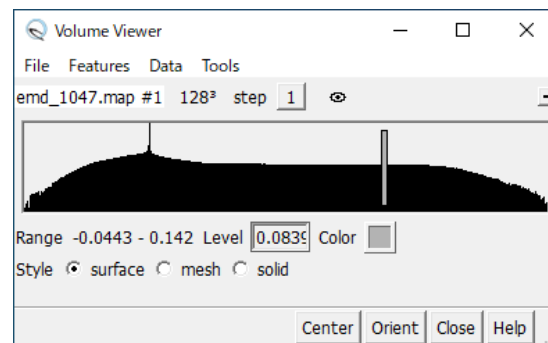
(1)[File]→[Fetch by ID]      (3) 1047の3Dマップが表示



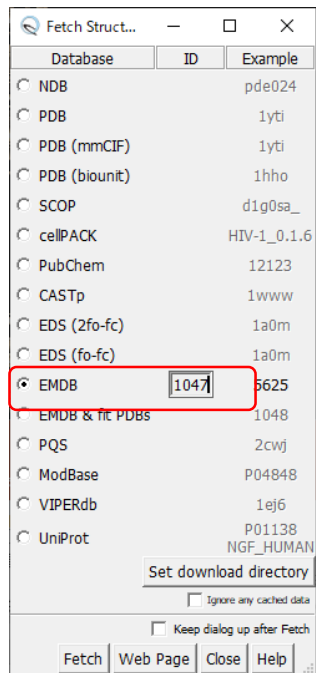
(5) stepを1にする



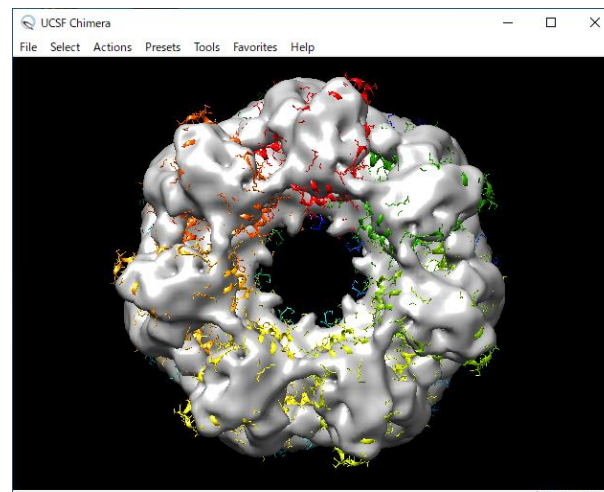
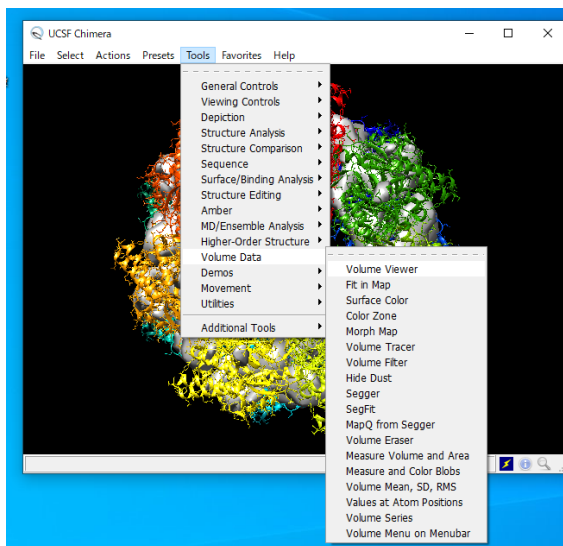
(6) Levelを0.0839にする



(2)[EMDB]に1047

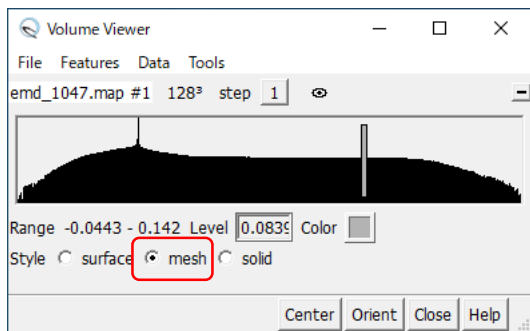


(4)[Tools] → [Volume Data] → [Volume Viewer]

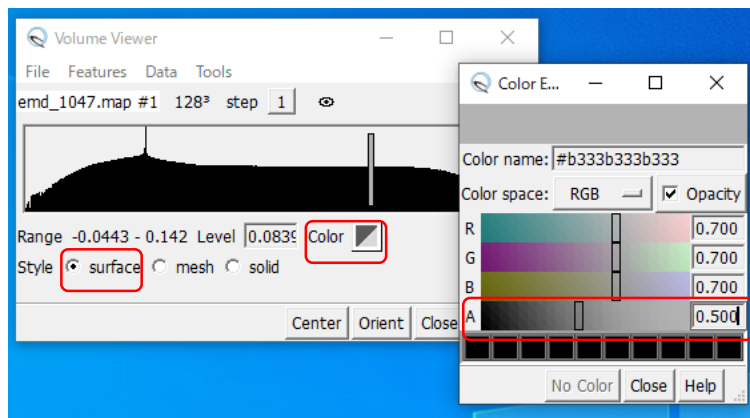


# 3Dマップを透明にする方法

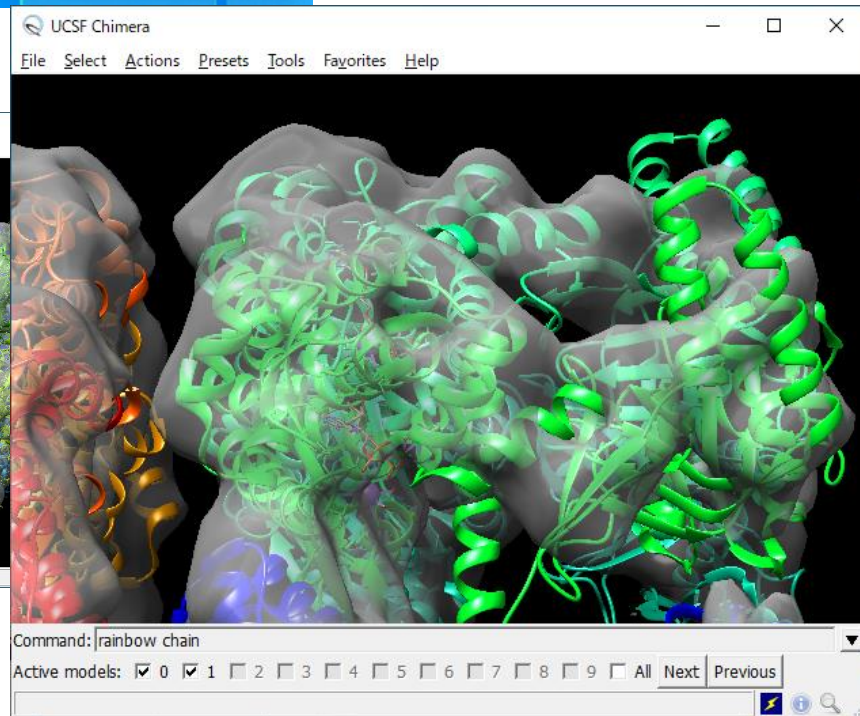
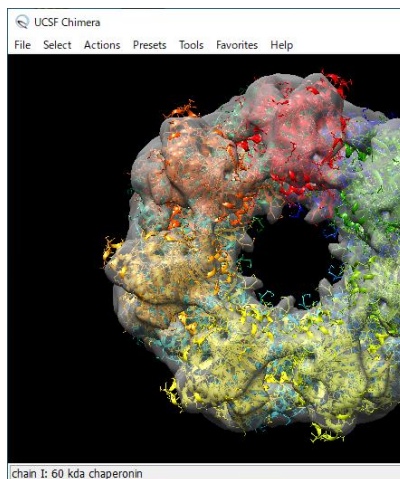
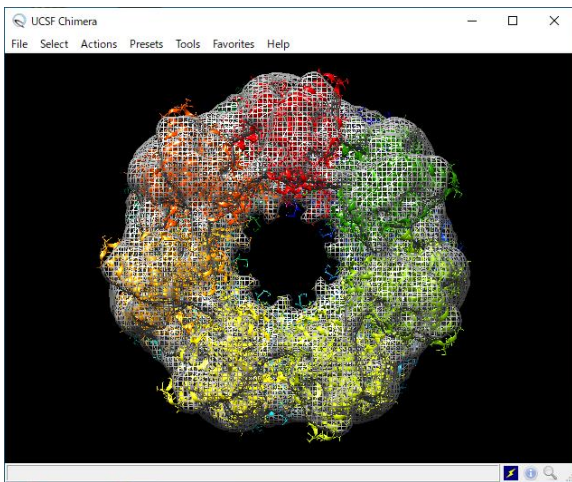
(1) Volume ViewerのStyleを  
"mesh"に



(2) Volume ViewerのStyleを"surface"にして  
Colorの□をクリック、"A"の値を0.5にする



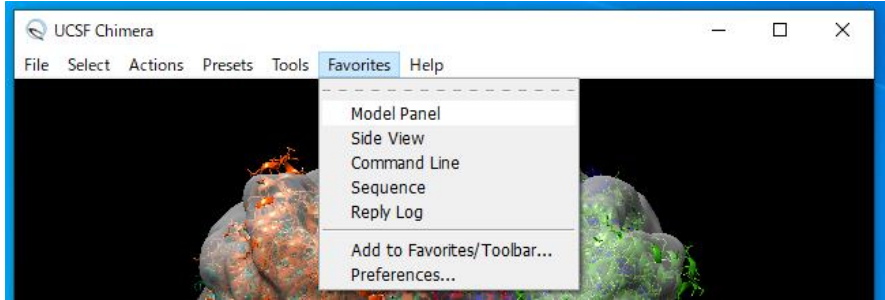
分解能9.7 Åなので  
ドメイン程度まで判別できる



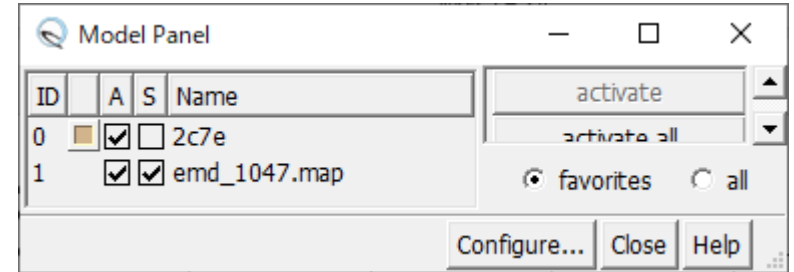


# 原子モデル・3Dマップを隠す方法

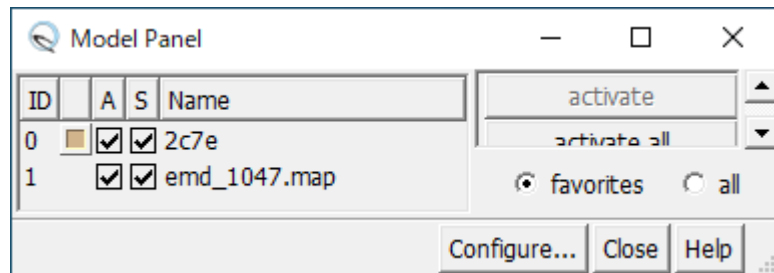
(1)[Favorites]→[Model Panel]



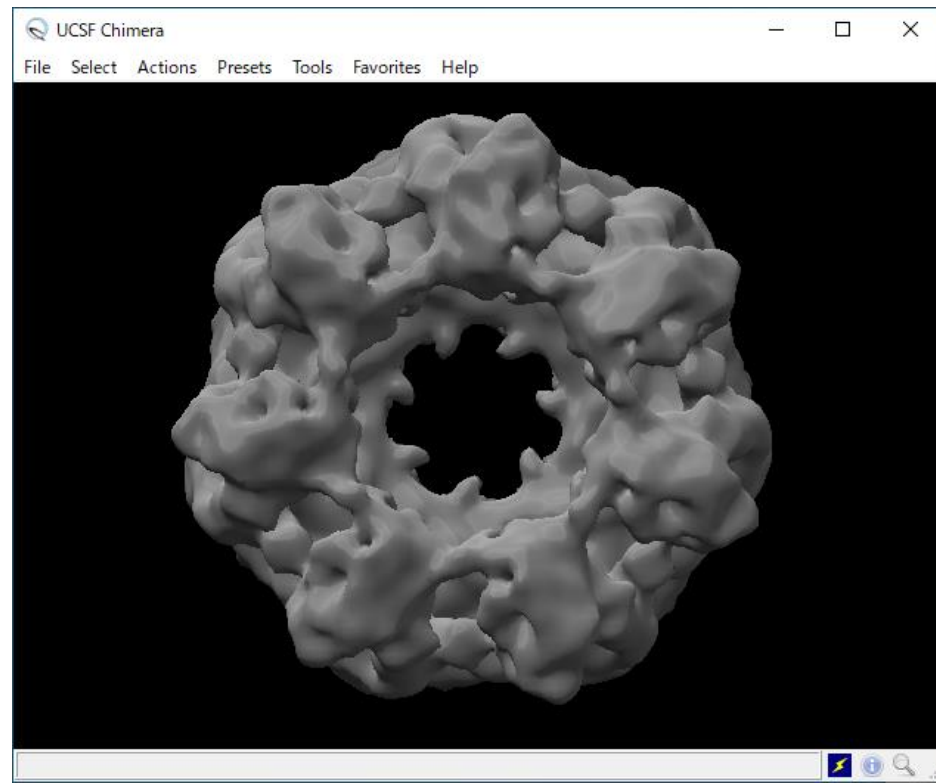
(3)2c7eのをはずす



(2)[Model Panel]が表示される



(4)マップ(emd\_1047.map)だけが表示される



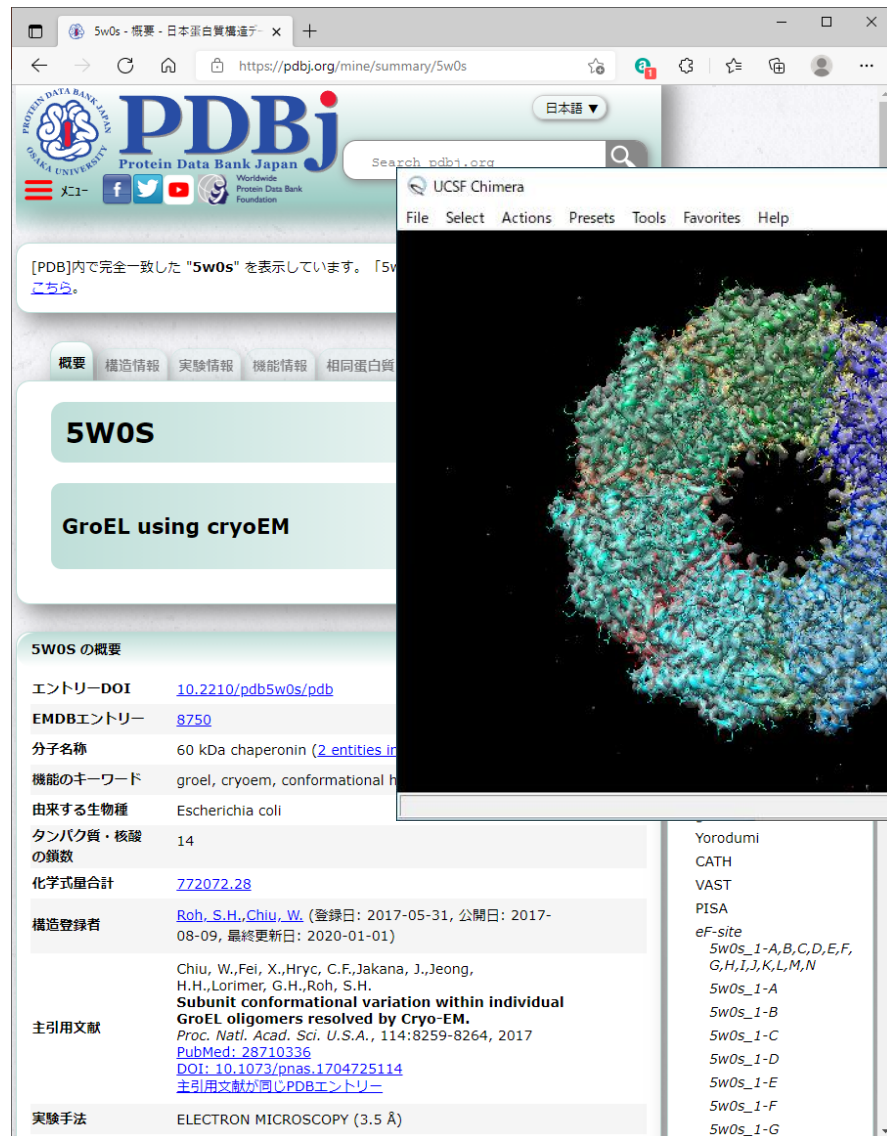
※AはActive, SはShowを表す

# PDB ID:5w0s, EMDB:8750も見てみましょう

PDB\_ID: 5w0s EMDB\_ID: 8750 (3.5 Å; 240<sup>3</sup> voxels)

contourLevel : 0.04

分解能3.5 Åなので  
二次構造まで判別できる



5W0S

GroEL using cryoEM

5W0S の概要

エントリー-DOI [10.2210/pdb5w0s/pdb](https://doi.org/10.2210/pdb5w0s/pdb)

EMDBエントリー [8750](https://emdb.ebi.ac.uk/emdb/8750)

分子名称 60 kDa chaperonin (2 entities in assembly)

機能のキーワード groel, cryoem, conformational h...

由来する生物種 Escherichia coli

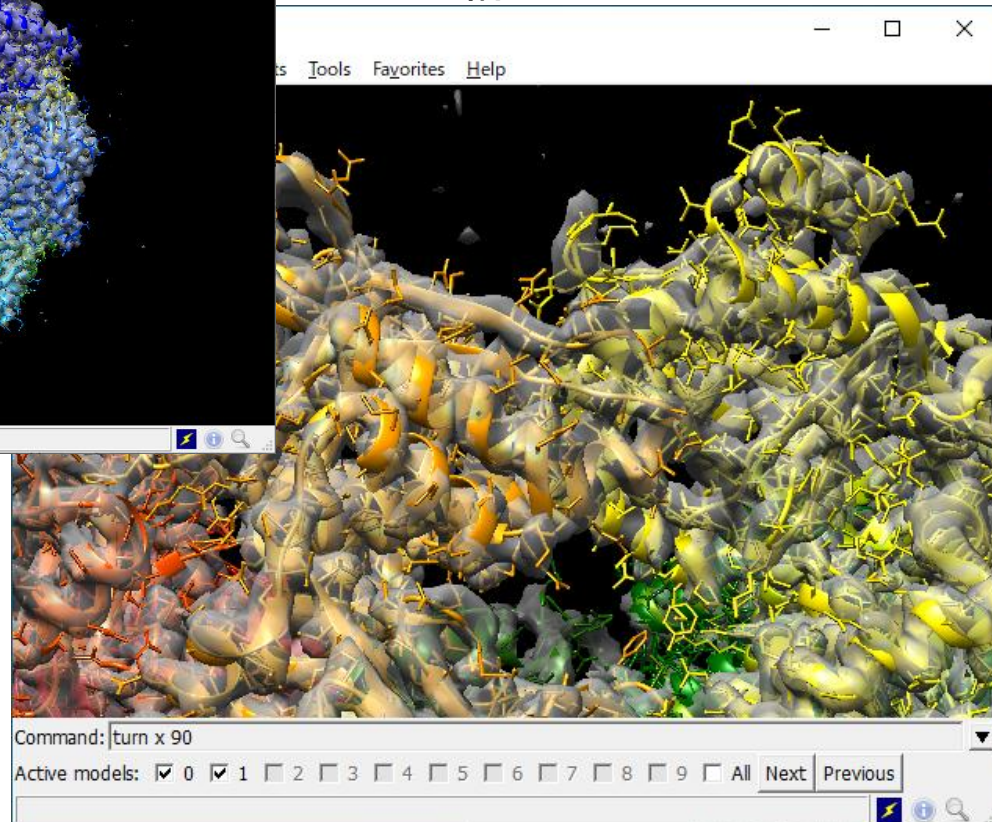
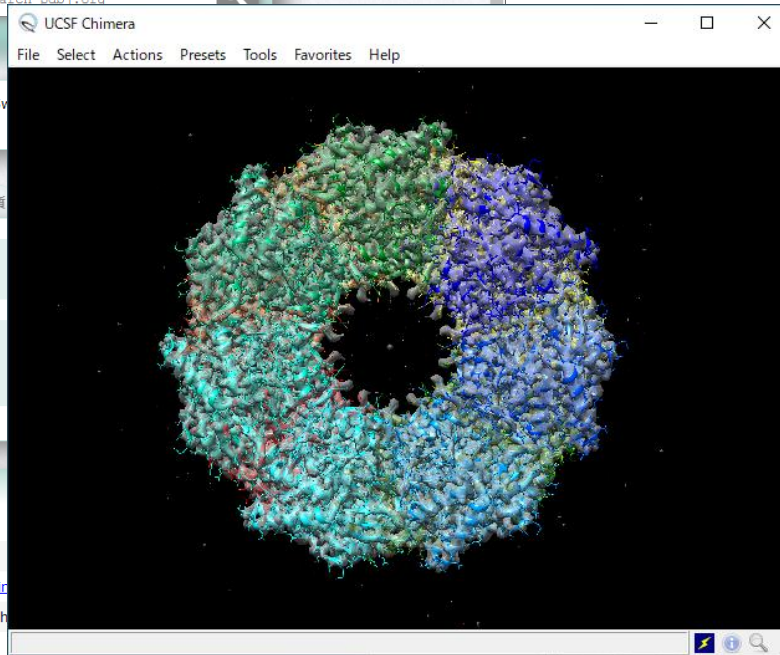
タンパク質・核酸の鎖数 14

化学式量合計 [772072.28](https://pubchem.ncbi.nlm.nih.gov/compound/77207228)

構造登録者 [Roh, S.H., Chiu, W.](#) (登録日: 2017-05-31, 公開日: 2017-08-09, 最終更新日: 2020-01-01)

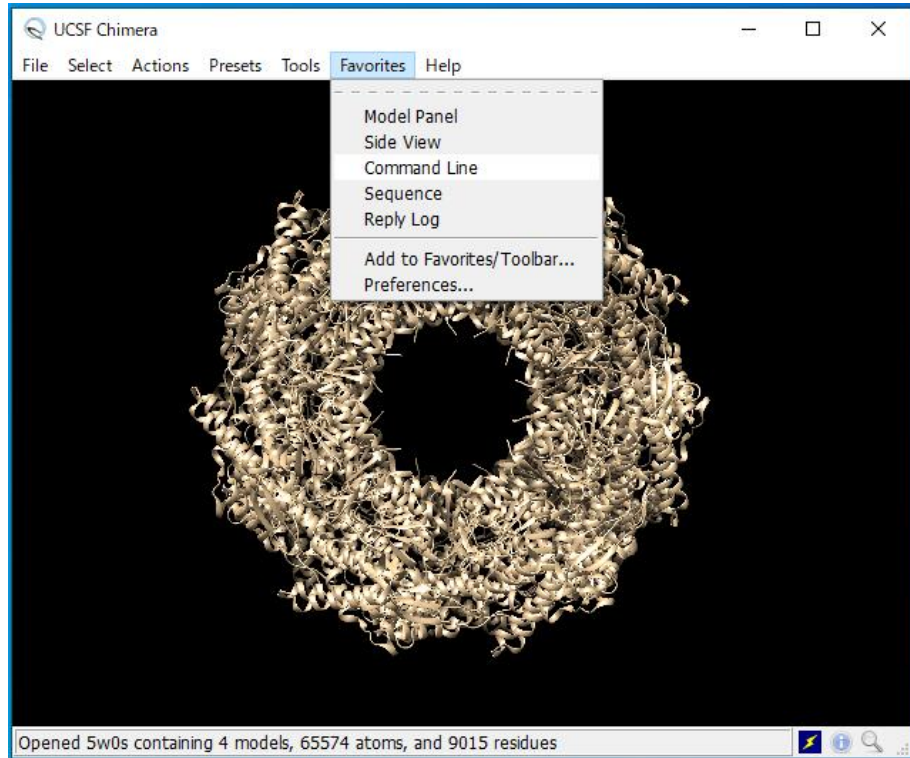
主引用文献 Chiu, W., Fel, X., Hryc, C.F., Jakana, J., Jeong, H.H., Lorimer, G.H., Roh, S.H. Subunit conformational variation within individual GroEL oligomers resolved by Cryo-EM. Proc. Natl. Acad. Sci. U.S.A., 114:8259-8264, 2017. PubMed: 28710336. DOI: 10.1073/pnas.1704725114. 主引用文献が同じPDBエントリー

実験手法 ELECTRON MICROSCOPY (3.5 Å)

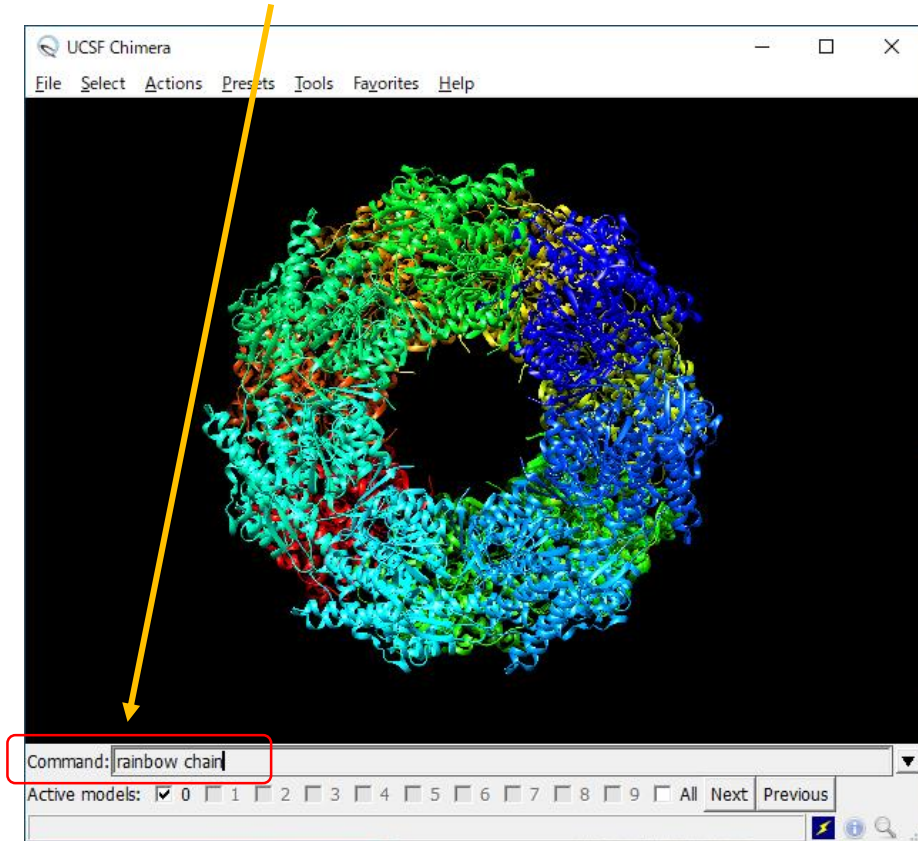


# UCSF Chimeraのコマンドライン

(1)[Favorites]→[Command Line]



(2) Command:にコマンド  
(例えば、**rainbow chain**)を入力





# UCSF Chimera 選択コマンド

書式	例	意味
[実行] :.[鎖]	color red :.A	A鎖を赤に
[実行] :[残基名]	color red :CYS	システインを赤に
[実行] @[原子名]	color red @CB	Cb原子を赤に
[実行] :[残基名]@[原子名]:[鎖]	color red :CYS.A@CB	A鎖のシステインのCb原子を赤に
[実行] :[番号]	color red :104	104番目の残基を赤に
[実行] :[番号],[番号]	color red :104,212	104番目と212番目の残基を赤に
[実行] :[番号]-[番号]	color red :104-212	104~212番目の残基を赤に
[実行] :[番号]-[番号].[鎖]	color red :104-212.A	A鎖の104~212番目を赤に
[実行] :[条件] za<[距離]	color red :ATP za<5	ATPから5Å以下の原子を赤に
[実行] :[条件] zr<[距離]	color red :ATP zr<5	ATPから5Å以下の残基を赤に
[実行] [条件] and [条件]	select protein and 104	タンパク質の104番目を選択
[実行] :[条件] or :[条件]	color red :SER    :THR	セリンかスレオニンを赤に
[実行] protein	color red protein	タンパク質を赤に
[実行] nucleic acid	color red nucleic acid	核酸を赤に
[実行] ligand	color red ligand	リガンド分子を赤に
[実行] solvent	color red solvent	水分子を赤に

# UCSF Chimera 実行コマンド

書式	例	意味
<code>display</code> と <code>~display</code>		原子の表示と非表示
<code>ribbon</code> と <code>~ribbon</code>		リボンモデルの表示と非表示
<code>surface</code> と <code>~surface</code>		分子表面の表示と非表示
<code>repr sphere</code>		原子を空間充填モデルで
<code>repr wire</code>		原子をワイアフレームモデルで
<code>repr stick</code>		原子をスティックモデルで
<code>repr bs</code>		原子をボール&スティックで
<code>color [色]</code>	<code>color blue</code>	青色にする
<code>color byelement</code>		元素ごとに色分けする
<code>rainbow chain</code>		鎖ごとに虹色に
<code>rainbow residue</code>		N末からC末へ虹色に
<code>set bg_color [色]</code>	<code>set bg_color white</code>	背景を白に
<code>turn [xyz] [回転角(°)]</code>	<code>turn y 180</code>	Y軸のまわりに180° 回転
<code>reset</code>		分子を元の向きに戻す

# PyMOL・UCSF Chimeraの操作法などが載っている本



見えてわかる 構造生命科学 —生命科学  
研究へのタンパク質構造の利用—  
中村春木 編 化学同人 2014年 税抜  
5000円

RasMol, UCSF Chimera, PyMOLの  
使い方を解説



Gert-Jan Bekker

# Molmil

- PDBjのGert-Jan Bekker氏が開発した分子ビューア
- ブラウザ内で分子を表示。WebGLを使用
- インストール作業が不要
- JavaScriptで開発
- ブラウザ内の標準的なビューアJSmolに比べて、表示がきれいで動作が速い
- 自分のWEBページに3次元の分子をかざることができる
- mmCIFファイルにきちんと対応

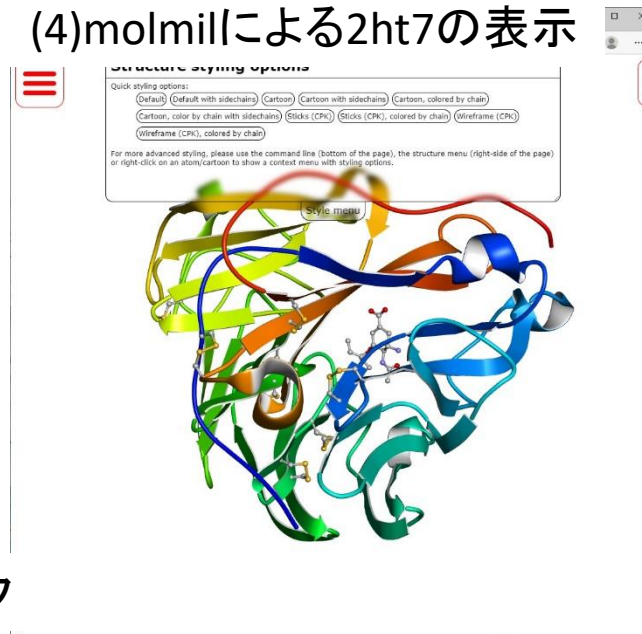
Gert-Jan Bekker, Haruki Nakamura & Akira R. Kinjo. Molmil: a molecular viewer for the PDB and beyond. Journal of Cheminformatics 8, Article number: 42 (2016)

# ノイラミニダーゼのPDBの構造の観察

(1) <https://pdj.org>のフォームに2ht7と入力



(4) molmillによる2ht7の表示

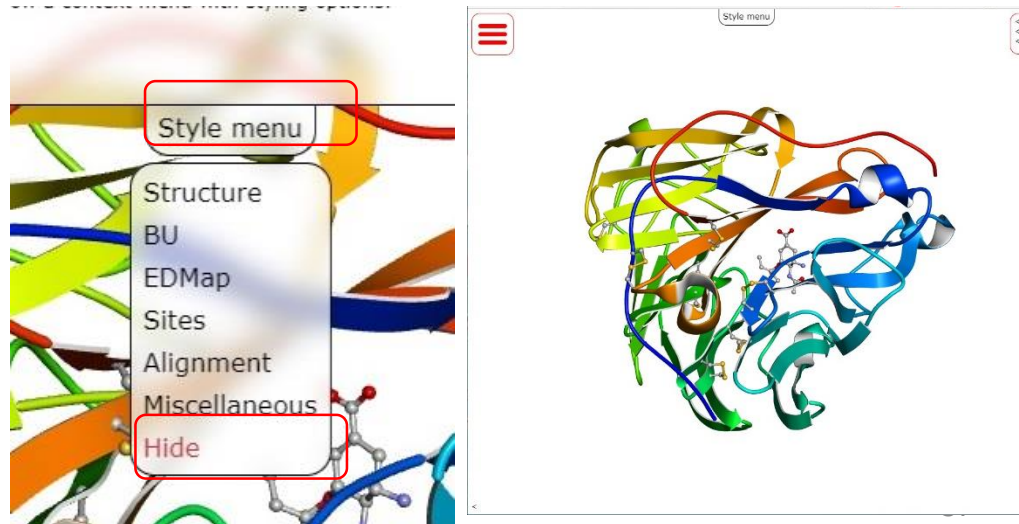


(2) PDB\_ID: 2ht7 のページが表示される



(3) [非対称単位を表示]をクリック

(5) [Style menu]->[Hide]でスタイルメニューを消せます



2ht7 chain A: UniProt: NRAM\_I63A3

Neuraminidase EC=3.2.1.18 GN Name=NA

Influenza A virus (strain A/Duck/Ukraine/1/1963 H3N8).

-|- FUNCTION: Catalyzes the removal of terminal sialic acid residues from viral and cellular glycoconjugates. Cleaves off the terminal sialic acids on the glycosylated HA during virus budding to facilitate virus release. -|- CATALYTIC ACTIVITY: Reaction=Hydrolysis of alpha-(2->3)-, alpha-(2->6)-, alpha-(2->8)- glycosidic linkages of terminal sialic acid residues in oligosaccharides, glycoproteins, glycolipids, colominic acid and synthetic substrates.; EC=3.2.1.18; -|- SUBUNIT: Homotetramer.

# メニューによる表示の変更

デフォルト

左クリック

左クリック

右クリック

Display:  
Ca trace

Color:  
structure

Structures: + 2HT7 (#1)

Structures: - 2HT7 (#1)

Chains: + A + B (A)

Chain A > Display

- Hidden
- Default
- Amino acid >
- Sidechain >
- Ca trace
- Tube
- Cartoon
- Rocket
- CG Surface
- Simple Surface
- Hydrogen bonds

Chain A > Color

- Default
- Structure
- Atom (CPK)
- Group
- ABEGO
- Custom

Chain A

- Display
- Color
- Label



# コマンドラインの表示

(1) メニューアイコン



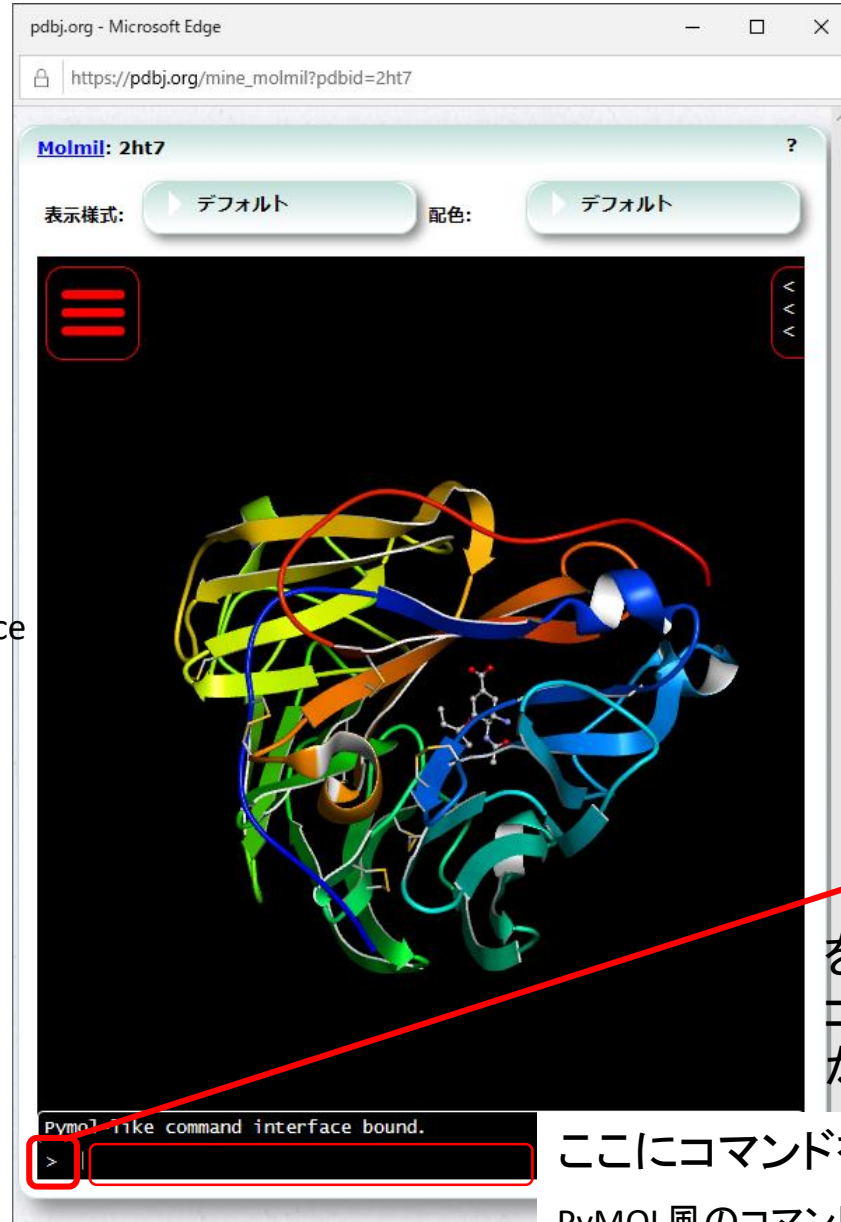
を左ボタンでドラッグし、  
[Toggle CLI]を選ぶ



※もう一度、  
[Toggle CLI]を  
選ぶと、  
コマンドライン  
が消える

CLI :  
Command Line Interface

(2) Pymol-like command line interfaceが表示される



をクリックすると  
コマンドライン  
が消える

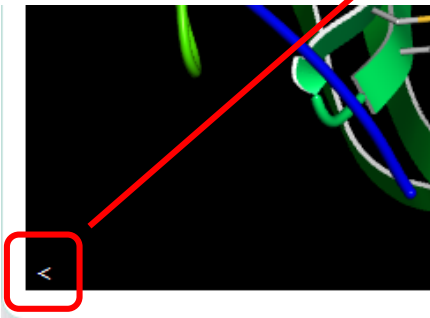
ここにコマンドを入力

PyMOL風のコマンドを入力できる

あるいは、画面下の



をクリック

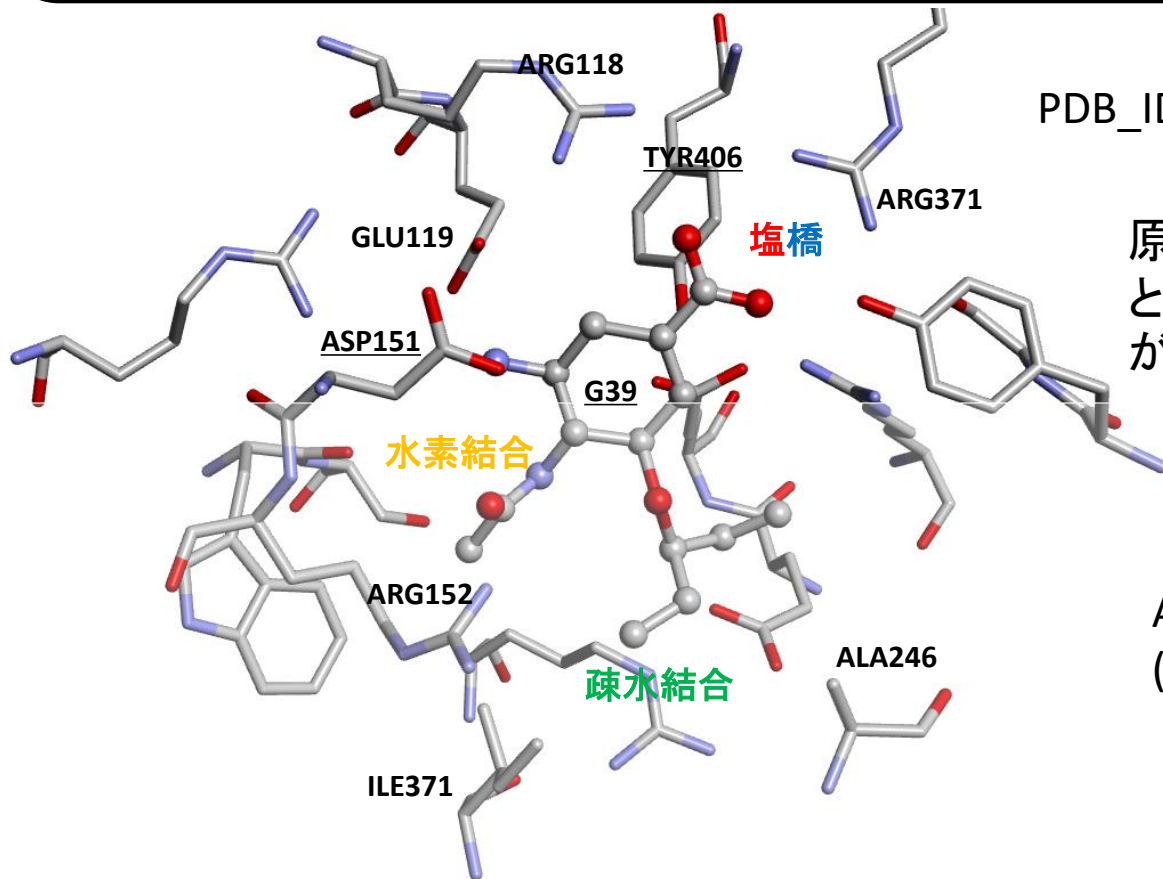


# コマンド入力:リガンド周辺原子の表示

スペースを必ず入れること!

コマンドラインインターフェースに以下のコマンドを入力する

```
hide cartoon, all #カートゥーン表示をオフ
hide ball_stick, all #スティック表示をオフ
show sticks, byres  (resn G39 around 5) #分子G39の周辺5Åの残基を選択
color cpk, all #原子の色をCPKに
show ball_stick, resn G39 #選択した原子を、太さ50のワイアフレームで表示
```



PDB\_ID:2HT7 のG39周辺の構造

原子をマウス左ボタンでクリックすると、その残基名・残基番号・原子名が、左下に表示される

ASP151, TYR406は活性部位  
(UniProt: NRAM\_I63A3)

ASP151: Proton Donor/Acceptor  
TYR406: Nucleophile

# MolmilコマンドとPyMOLコマンドの違い

デフォルトの蛋白質がカートゥーン表示のときに以下のコマンドを実行する場合

## 1. “,all”が必要

× `hide cartoon`

何も起きない...

○ `hide cartoon,all`

カートゥーンが消えてくれる！

## 2. “color”と“cartoon\_color”を区別

× `color red,all`

何も起きない...

○ `cartoon_color red,all`

カートゥーンの色が赤くなる！

# Molmil コマンド早見表 : 選択コマンド

選択コマンドの書式	例	意味
[実行], all	show sticks, all	全ての原子を棒表示
[実行], chain [鎖]	cartoon_color red, chain A	A鎖のカートウーンを赤に
[実行], resn [残基名]	show sticks, resn CYS	システインを棒表示
[実行], name [原子名]	show spheres, name CB	CB原子を球表示に
[実行], symbol [元素名]	show spheres, symbol S	硫黄元素を球表示に
[実行], entity [エンティティ番号]	show spheres, entity 3	エンティティ3を球表示に
[実行], resi [番号]	show spheres, resi 104	104番目を球表示に
[実行], resi [番号]-[番号]	show spheres, resi 104-212	104~212番目を球表示に
[実行],[条件] around [距離]	show sticks, resn ATP around 5	ATPから5Å以下の原子を棒表示に
[実行], byres ([条件] around [距離])	show sticks, byres (resn ATP around 5)	ATPから5Å以下の残基を棒表示に
[実行], [条件] and [条件]	show spheres, chain A and resi 104	A鎖の104番目を球表示に
[実行], [条件] or [条件]	show sticks, resn SER or resn THR	セリンかスレオニンを棒表示に
[実行], hetatm	show spheres, hetatm	HETATMを球表示に
[実行], hydro	show spheres, hydro	水分子を球表示に
[実行], snfg	show cartoon, snfg	SNFG表示可能な糖鎖をカートウーンに
select [選択原子名],[条件]	select actsite, resi 104+212	104と212番目をactsiteと命名
[実行], [選択原子名]	show spheres, actsite	選択したactsiteを球表示に

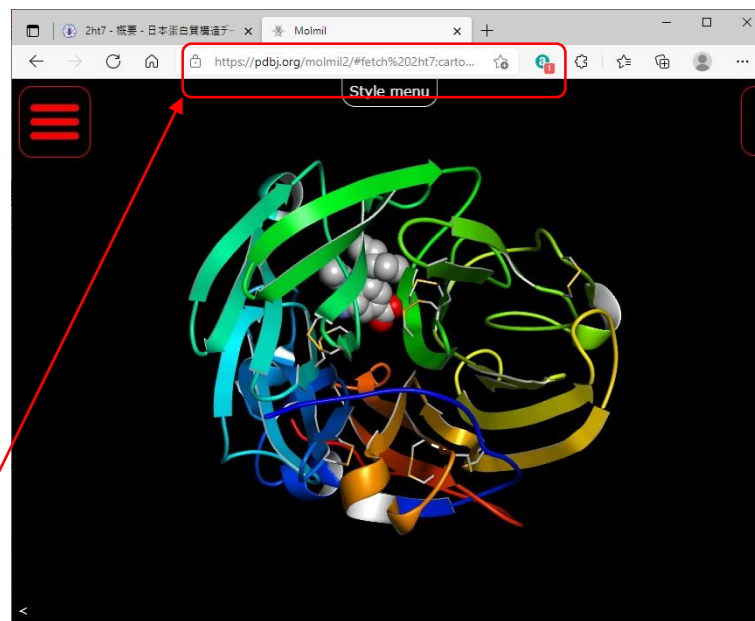
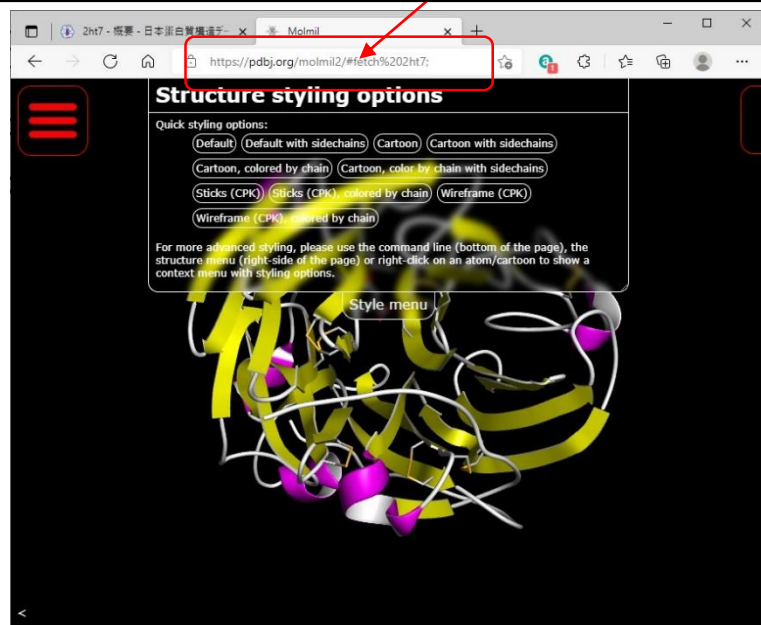
# Molmil コマンド早見表 : 実行コマンド

実行コマンドの書式	例	意味
show [表示法]	show spheres, all	球モデルの非表示
hide [表示法]	hide spheres, all	球モデルの表示
※ [表示法]は、球:spheres、線:lines、スティック:sticks、ボール&スティック:ball_stick、カートゥーン:cartoon、バックボーン:ca-trace、SNFGとスティック:snfg-icon		
color [色]	color blue, all	青色にする
cartoon_color [色]	cartoon_color red, all	カートゥーンの色を赤にする
color [r,g,b]	color [0,255,255],all	シアン色[0,255,255]にする
color cpk	color cpk, all	元素ごとに色分けする
※次の[色]も使用できる。N末からC末へ青から赤:group、二次構造による色分け:structure、温度因子による彩色:bfactor		
bg_color [色]	bg_color white	背景を白に
turn [xyz], [回転角(°)]	turn y, 180	Y軸のまわりに180°回転
reset	reset	分子を元の向きに戻す
set cif_use_auth, [onかoff]	set cif_use_auth,off	label_asym_idを鎖識別子に、label_seq_idを残基番号に

# WEBアドレス1行で構造が表示できる

WEBブラウザのアドレスに以下のコマンドを打つだけで構造を表示できます

```
https://pdbj.org/molmil2/#fetch 2ht7;
```



```
https://pdbj.org/molmil2/#fetch 2ht7;cartoon_color group,all;show spheres, resn G39;
```

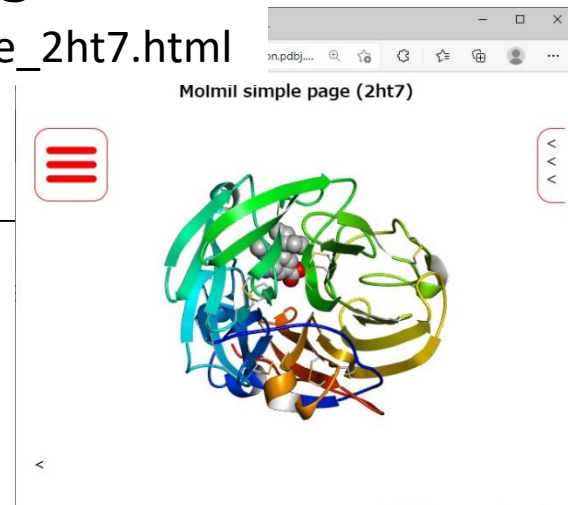
#の後に、Molmilのコマンドをセミコロン(;)区切りで入力することができます

電子メールやブログなどで、簡単に指定した表示の立体構造へのリンクを示すことができる



# Molmilを組み込んだシンプルなWEBページ

[https://numon.pdbj.org/binds\\_workshop\\_202109/molmil\\_simple\\_2ht7.html](https://numon.pdbj.org/binds_workshop_202109/molmil_simple_2ht7.html)



```
<html>
  <head>
    <script src="https://pdbj.org/molmil2/molmil.js"></script>
    <script>
      var molmil_settings = {src: "/molmil2/"}, canvas;

      function initView() {
        canvas = document.getElementById("molmilViewer");
        molmil.autoSetup(); PDBjサーバから2ht7を取得
        molmil.loadPDB("2ht7", null, null, canvas.molmilViewer);
        canvas.commandLine.run("cartoon_color group, all;show spheres, resn G39;");
        Molmilのコマンドをここに記入
      }
    </script>
  </head>

  <body onload="initViewer();">
    <center>
      <h3>Molmil simple page (2ht7)</h3>
      <span class="molmil_UI_container">
        <canvas id="molmilViewer" width="600" height="400"></canvas>
      </span>
    </center>
  </body>
</html>
```

# Molmilを組み込んだボタン付きWEBページ

[https://numon.pdbj.org/binds\\_workshop\\_202109/molmil\\_advanced\\_2ht7.html](https://numon.pdbj.org/binds_workshop_202109/molmil_advanced_2ht7.html)

```
<html>
<head>
<script src="https://pdbj.org/molmil2/molmil.js"></script>
<script>
  var molmil_settings = {src: "/molmil2/"}, canvas;

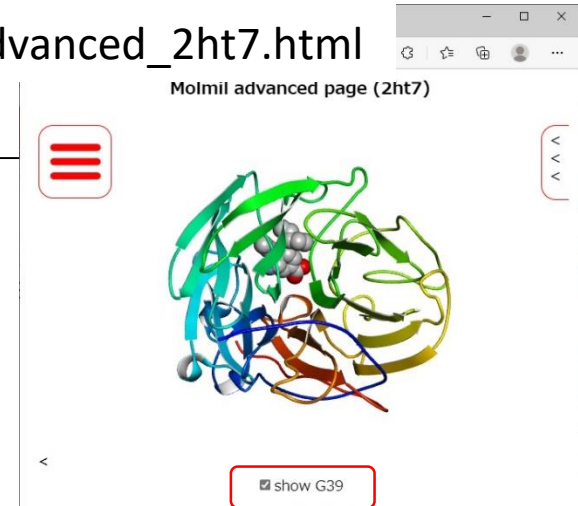
  function initView() {
    canvas = document.getElementById("molmilViewer");
    molmil.autoSetup();
    canvas.molmilViewer.loadStructure("2ht7_in_this_server.pdb", "pdb");
    canvas.commandLine.run("cartoon_color group, all;show spheres, resn G39;");
  }
  function show_G39(boolEnable) {
    if (boolEnable){ canvas.commandLine.run("show spheres, resn G39;"); }
    else{ canvas.commandLine.run("hide spheres, resn G39;"); }
  }
</script>
</head>

<body onload="initViewer();">
  <center>
    <h3>Molmil advanced page (2ht7)</h3>
    <span class="molmil_UI_container">
      <canvas id="molmilViewer" width="600" height="400"></canvas>
    </span>
    <BR> <input type="checkbox" onclick="show_G39(this.checked)" checked>show G39
  </center>
</body>
</html>
```

Webサーバに置いたPDBファイルを読み込む

G39を球表示したり、隠したりする関数

checkboxをクリックすると、関数show\_G39()を呼び出すように設定



# Molmilを活用したWEBページ

(1) 本講習会の大澤先生の発表のWEBページ

[https://numon.pdbj.org/binds\\_workshop\\_202109/osawa/](https://numon.pdbj.org/binds_workshop_202109/osawa/)



(2) 一般向けコロナウィルスの解説ページ

生命をささえるタンパク質の「かたち」- 新型コロナウイルスの立体構造 -

<https://numon.pdbj.org/covid19>

YouTubeに解説動画もあります

<https://www.youtube.com/user/PDBjmovie>

