

# PDBjとwwPDBの活動について

大阪大学蛋白質研究所

日本蛋白質構造データバンク: PDBj

中村 春木



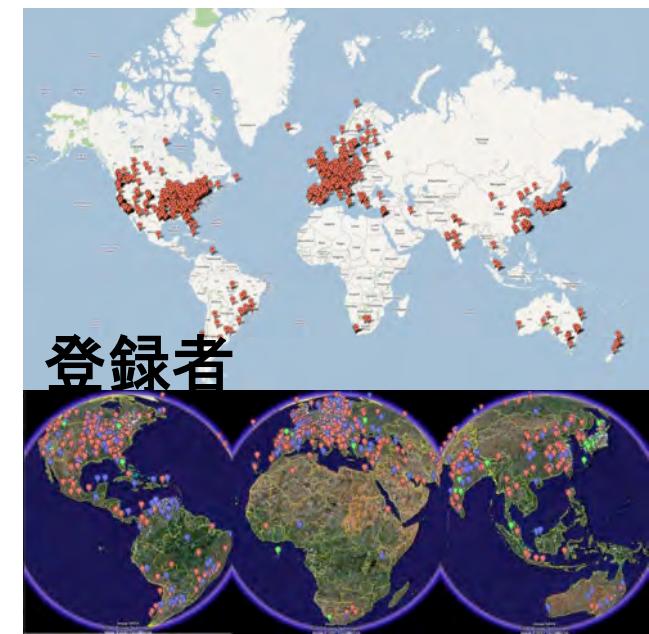
*<http://pdbj.org/>*



*<http://wwpdb.org/>*

# 蛋白質の形のデータバンク:PDB

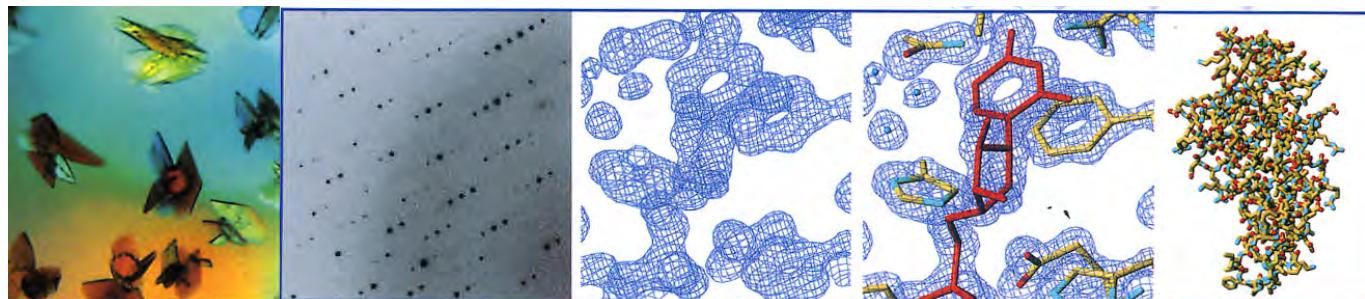
- ・ 生体高分子の構造(形)に関する全ての情報を集めた国際的データベース:データバンク方式
- ・ 情報は、研究者・教育者・学生・企業を問わず無償で利用される
- ・ 運営は各国(米国、欧州、日本)の研究費用でまかなわれている



利用者  
(2014年には5億件のダウンロード)

# 蛋白質の形のデータバンク:PDB

- X線結晶回折実験

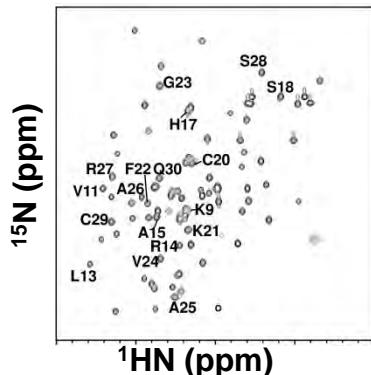


結晶作成 → X線回折像 → 電子密度マップ → 原子モデル作成



SPring-8 放射光施設

- 核磁気共鳴(NMR)法



超高磁場 NMR装置

- 超低温電子顕微鏡

# 蛋白質の形のデータバンク:PDB

## PDBの歴史

1960年代:蛋白質結晶学が始まる

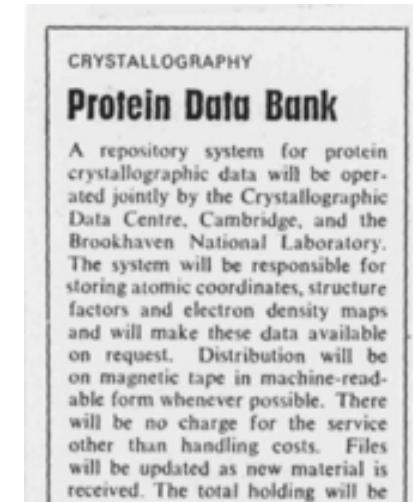
ミオグロビン、ヘモグロビン、  
リゾチームなど



マックス ペルーツ &  
ジョン ケンドリュー(1962  
年ノーベル化学賞)

1970年代:PDB 開始 7件(1971年10月)

米国ブルックヘブン国立研究所  
大阪大学蛋白質研究所においても  
シトクロムCの構造を決定→登録  
磁気テープを配布



# 蛋白質の形のデータバンク:PDB

1980年代:データ登録の急増(国際結晶学会:  
データ登録が論文投稿の必須条件)

1990年代:RCSB-PDB(米国)が運営を開始

2000年代:wwPDBが設立され国際化  
構造ゲノム科学による構造急増

2010年代:新規の手法が開発  
100,000件以上のデータ公開  
ビッグ・データの時代へ

# 蛋白質の形のデータバンク:PDB

WORLDWIDE PROTEIN DATA BANK

The worldwide Protein Data Bank

www.wwPDB.org • info@wwPDB.org

RCSB PDB PROTEIN DATA BANK

Research Collaboratory for Structural Bioinformatics  
[www.pdb.org](http://www.pdb.org)

BioMagResBank  
[www.bmrb.wisc.edu](http://www.bmrb.wisc.edu)

PDBe PROTEIN DATA BANK EUROPE  
[pdbe.org](http://pdbe.org)

Protein Data Bank in Europe

PDBj Protein Data Bank Japan  
[www.pdbj.org](http://www.pdbj.org)

WORLDWIDE PROTEIN DATA BANK

VALIDATION - DEPOSITION - DATA DICTIONARIES - DOCUMENTATION - TASK FORCES - STATISTICS - ABOUT

Validate Structure  
View validation reports

Deposit Structure  
All Deposition Resources

Download Archive

Since 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies.

The Worldwide PDB (wwPDB) organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global community.

Learn more about PDB HISTORY and FUTURE.

wwPDB Members

Each site offers tools for searching, visualizing, and analyzing PDB data:

- PDBj
- Protein Data Bank Japan

PDBj

Supports browsing in multiple languages such as Japanese, Chinese, and Korean; SeSAV identifies functionally or evolutionarily conserved motifs by locating and annotating sequence and structural similarities, tools for bioinformaticians, and more.

RCSB PDB

Research Collaboratory for Structural Bioinformatics Protein Data Bank

Simple and advanced searching for macromolecules and ligands, tabular reports, specialized visualization tools, sequence-structure comparisons, RCSB PDB Mobile, Molecule of the Month and other educational resources at PDB-101, and more.

BMRB

Biological Magnetic Resonance Bank

Collects NMR data from any experiment and captures assigned chemical shifts, coupling constants, and peak lists for a variety of macromolecules; contains derived annotations such as hydrogen exchange rates, pKa values, and relaxation parameters.

PDBe

Protein Data Bank in Europe

Rich information about all PDB entries, multiple search and browse facilities, advanced services including PDBeISA, PDBeFold and PDBeMotif, advanced visualisation and validation of NMR and EM structures, tools for bioinformaticians.

wwPDB Resources

Data Dictionaries

- Macromolecular Dictionary (PDB/mmCIF)
- Small Molecule Dictionary (CCD)
- Peptide-like antibiotic and inhibitor molecules (BIRD)

Annotation

- Procedures and policies
- Improvements for consistency and accuracy

Community Input: Task Forces and Working Groups

- Validation Task Forces (X-ray, NMR, 3DEM)
- Small Angle Scattering Task Force
- PDB/mmCIF Working Group
- Hybrid/Integrative Methods Task Force
- Ligand Validation Workshop

Read more

April 11, 2015

Phased PDB Release Process

As announced previously, the weekly public release of data from the Protein Data Bank (PDB) archive is now divided into two phases to serve better the needs of methods developers focused on protein structure prediction and protein-ligand docking. Going forward on a weekly basis, these developer communities will have ~4 days during which they can make blind predictions of protein or nucleic acid structure from polymer sequence and ligand docking pose from polymer sequence and the InChI string of bound ligand.

January 27, 2015

Time-stamped Copies of the PDB Archive

A snapshot of the PDB archive (<ftp://ftp.wwpdb.org>) as of January 2, 2015 has been added to <http://snapshots.wwpdb.org/>. Snapshots have been archived annually since January 2005 to provide readily identifiable data sets for research on the PDB archive.

Read more

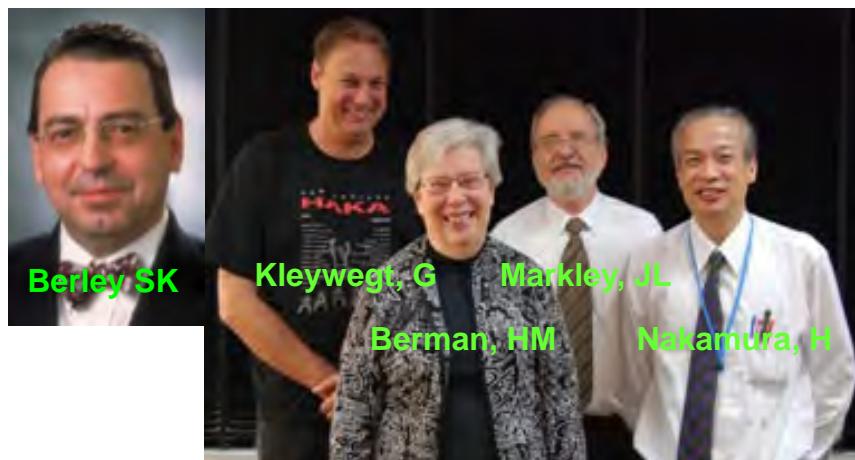
December 10, 2014

Integration of Large Structures with the Main PDB Archive

With this week's update, large structures (containing >62 chains and/or 99999 ATOM lines) represented as single files have been fully integrated into the main PDB FTP archive in both PDB/mmCIF and PDBML formats, as announced in September. Previously, large structures were represented in multiple "SPLIT" entries, which have now been removed (obsoleted). Users searching for ID codes of "SPLIT" entries at wwPDB member websites will be automatically redirected to the combined entry.

wwPDB

Nature Structural Biology 10, 980 (2003)  
doi: 10.1038/nsb1203-980  
More publications



<http://wwpdb.org/>

# 大阪大学蛋白質研究所とPDBj

Protein Data Bank  
Japan  
日本蛋白質構造データバンク  
*<http://pdbj.org/>*



PDBj トップページ(日英中韓)

2001年度から、(独) 科学技術振興機構(JST-NBDC:  
バイオサイエンスデータベースセンター)  
の支援を受けて活動中



PDBj スタッフ(2015年4月)

# ビッグデータ: データ主導の科学

## 科学における新しいパラダイム

実験、理論



アイザック ニュートン  
(1642-1727)

第三の科学: 計算科学



第四の科学: データ主導の科学

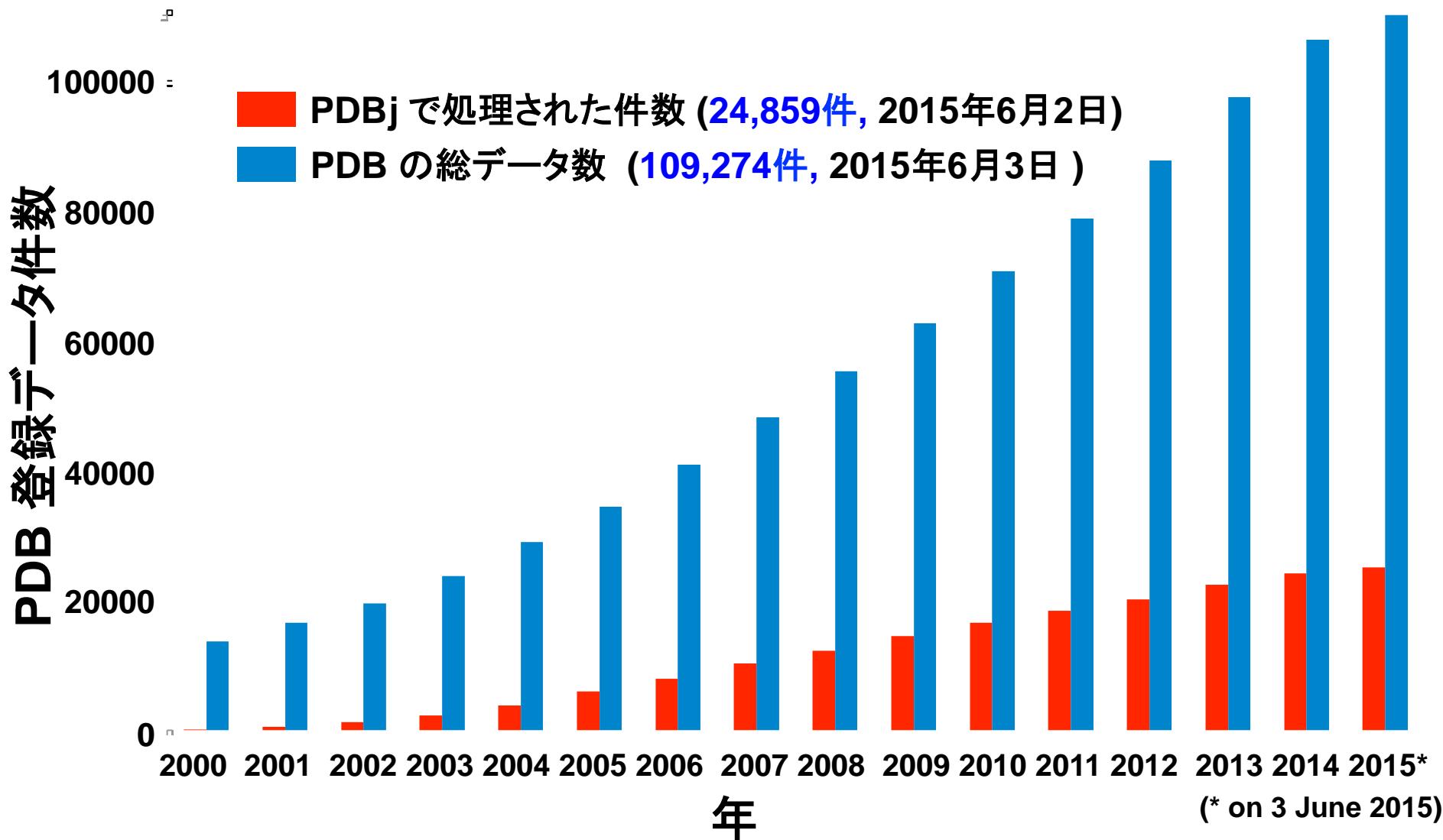
Wisdom  
Knowledge  
Information  
Data

DIKW ピラミッド  
(J. Rowley: The wisdom hierarchy, 2007)

# PDBjの活動とデータバンクの使い方

- “**Data-in**” の活動: wwPDBの一員として品質管理をしつつ登録作業を実施  
新たな共通フォーマットの作成(XML, RDF)  
にも寄与
- “**Data-out**” の活動: 共通データのダウンロード  
サイト(毎週アプデート)の運営  
種々のサービスや二次データベースの提供

# PDBjの活動とデータバンクの使い方



世界中で決定された構造の約1／4の登録処理をPDBjで実施

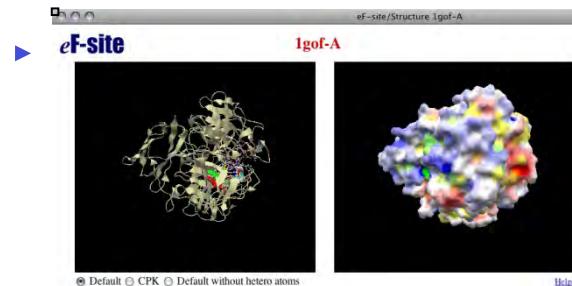
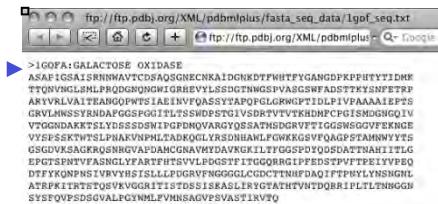
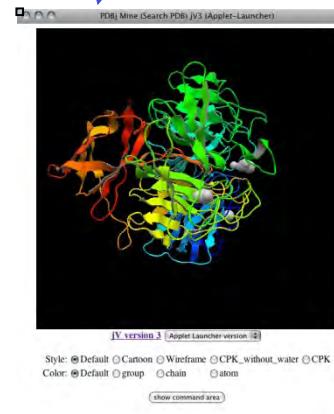
<http://pdbj.org/>

# PDBjの活動とデータバンクの使い方

## Amino acid sequence (FASTA)

# Data viewer at PDBj

## Graphic viewer: jV

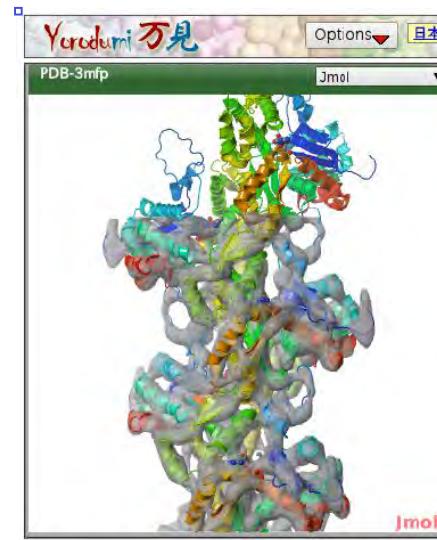
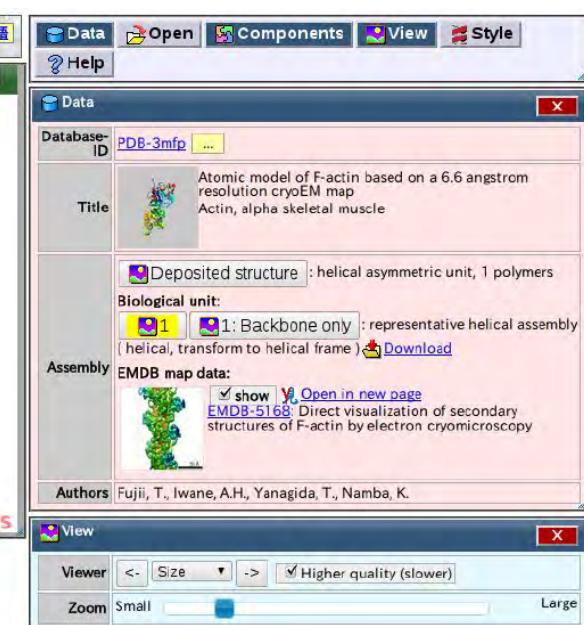


Functional site	focus & de
1) A:495	on of
2) A:272	on of
3) A:495	on of
4) A:496	on of
5) A:581	on of
6) A:228	on of
7) A:290	on of
8) A:272	on of
9) A:492	on of
10) A:75-87	on of
11) A:194	on of
12) A:227-228	on of
13) A:272	on of

## Molecular surface DB: *eF-site*

<http://ef-site.hgc.jp/eF-site/>

# PDBjの活動とデータバンクの使い方

The screenshot shows the 'Data' viewer interface for the PDB entry 'PDB-3mfp'. The top menu includes 'Data', 'Open', 'Components', 'View', and 'Style'. The main panel displays the following details:

- Title:** Atomic model of F-actin based on a 6.6 angstrom resolution cryoEM map
- Biological unit:** 1: Backbone only : representative helical assembly (helical, transform to helical frame)
- Assembly:** EMDB map data: EMDB-5168: Direct visualization of secondary structures of F-actin by electron cryomicroscopy
- Authors:** Fujii, T., Iwane, A.H., Yanagida, T., Namba, K.

Below this, there's a 'View' panel with options for 'Viewer', 'Size', 'Zoom' (set to 'Small'), and a checked checkbox for 'Higher quality (slower)'.

*EM Navigator*: 電子顕微鏡  
画像のデータベースEM-DBの  
ビューア

*Yorodumi*: PDBの原子構造と  
電子顕微鏡画像情報の統合化

# EM Navigator

3次元電子顕微鏡データナビゲーター [ [English](#) ] / [日本語](#)

[TOP](#) [キャラリーリスト](#) [分布図](#) [統計情報](#) [ビューア](#) [解説](#)

[PDBj](#) > [EM Navigator](#)

データを見る 詳しく

- さがす : (キーワード / EMDB ID / PDB ID)  EMDBかPDBのID、あるいはキーワードを入力

実行

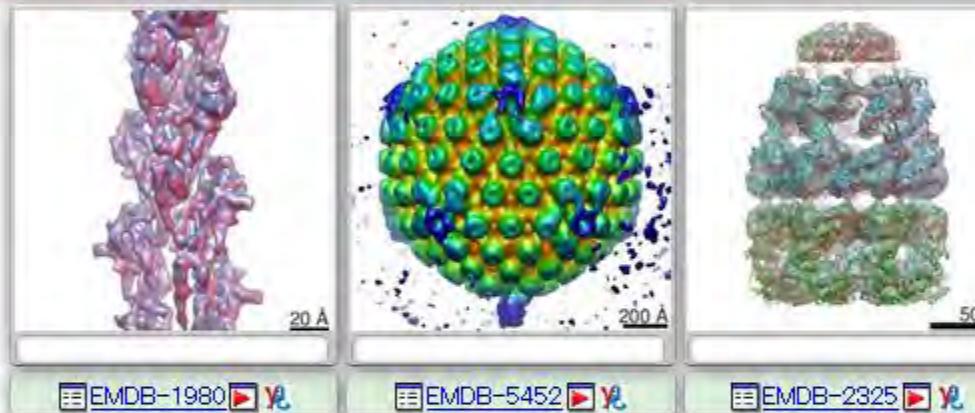
- 眺める : [ギャラリー](#) [リスト](#) [表](#)

ムービースロット

再生

方向

シャッフル



## 情報

### EM Navigatorとは？

- 生体分子や生体組織の3次元電子顕微鏡データを、気軽にわかりやすく眺めるためのウェブサイトです。
- EMDBとPDBのデータを利用しています（[統計情報](#)）
- 分子・構造生物学の専門家にも、初心者や専門外のかたにも利用していただけるサイトを目指しています。
- PDBが運営しています。

[詳しくはこちら](#)

### お知らせ

- 2013-09-11: 公開データ



他の最新データ: [EMDB付随情報](#), [EMDB更新](#), [PDB更新](#)

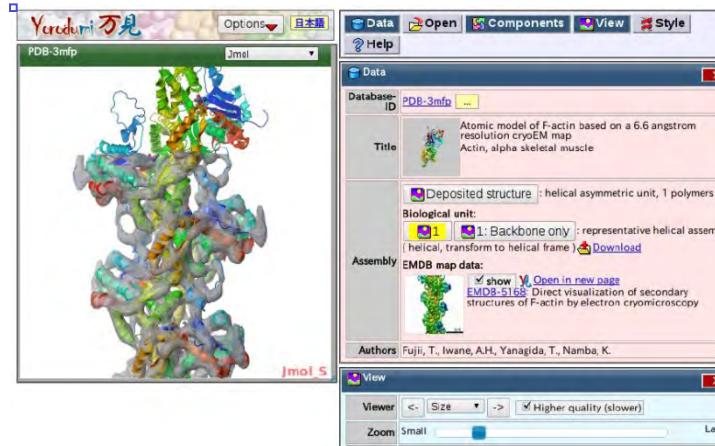
- 2013-09-04: 公開データ



# PDBj services



**EM Navi:** EMDB browser



**Yorodumi:** PDB & EMDB

## GIRAF query upload

Interface type

- Nonpolymer binding
- DNA and RNA binding
- All types of ligands (nonpolymer, DNA, RNA, peptides, and others).
- PPI (protein-protein interfaces).
- All (ligands + PPI).

Input PDB ID:

or upload a PDB file:

Chain IDs (optional):  (comma-separated multiple IDs [e.g., "A,B"] or "all" are allowed.)

Limit target PDB entries (optional):  (comma-separated multiple IDs [e.g., "101m,1a00"] or "all" are allowed.)

Number of displayed results (optional):  [100]

Your email address (optional):

DB version: 2013-09-21: 713107 interfaces

**GIRAF:** Similar binding site



**ProMode:** Normal Mode



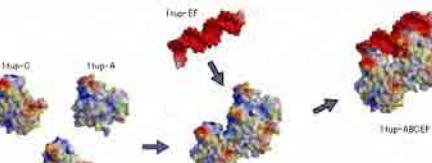
[About eF-site](#) | [References](#) | [Links](#) | [Acknowledgements](#) | [Feedback](#)

175033 Entries. Last Update: 20-Aug-2005

Keyword Search   
 PDB code only  and  or  
 Search  Reset

- Category Search
- Antibody
  - Protein
  - Active Site
  - Membrane
  - Binding Site

Examples of molecular surface



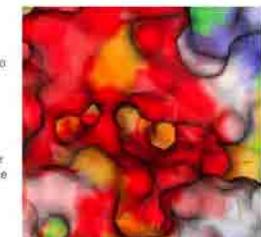
**eF-site:** Mol. Surface DB

[TOP](#) [Help](#) [FAQ](#) [References](#) [Links](#)



### ABOUT eF-seek:

Molecular function of proteins are determined by their three dimensional structures, thus the similarity of protein structure can give some clues to infer their functions. In many cases, the molecular functions are begun with the molecular interaction with small molecules (ligands). eF-seek is a web server to search for the similar ligand binding sites for the uploaded coordinate file with PDB format. The representative binding sites in eF-site database are search by our own algorithm based on the clique search algorithm.



### Submission STEP-1:

Specify a PDB format file:

E-mail address:

Keyword: #

Title: (optional)

**eF-seek:** Similar surface

# Integration with EMDB: Search for similar SHAPE

Query: human RNA polymerase II with RNA (EMDB: 2190)

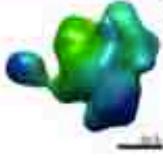
Similar shapes from 195,658 images

 Omokage search - Shape similarity search of macromolecules - (English) 日本語

**Search query**

**Subject structure**

Database: EMDB / ID: 2190  
human RNA polymerase II in complex with AluRA RNA  
[Quick](#), [Yorodumi](#), [EM Navigator](#)



**Search result**

Showing 1 - 100 of 2000 structures found from all (195658 structures)

Pages: [1](#) [2](#) [3](#) [4](#) [10](#) [20](#) [Previous](#) [Next](#)

Display: [images only](#) [as list](#)

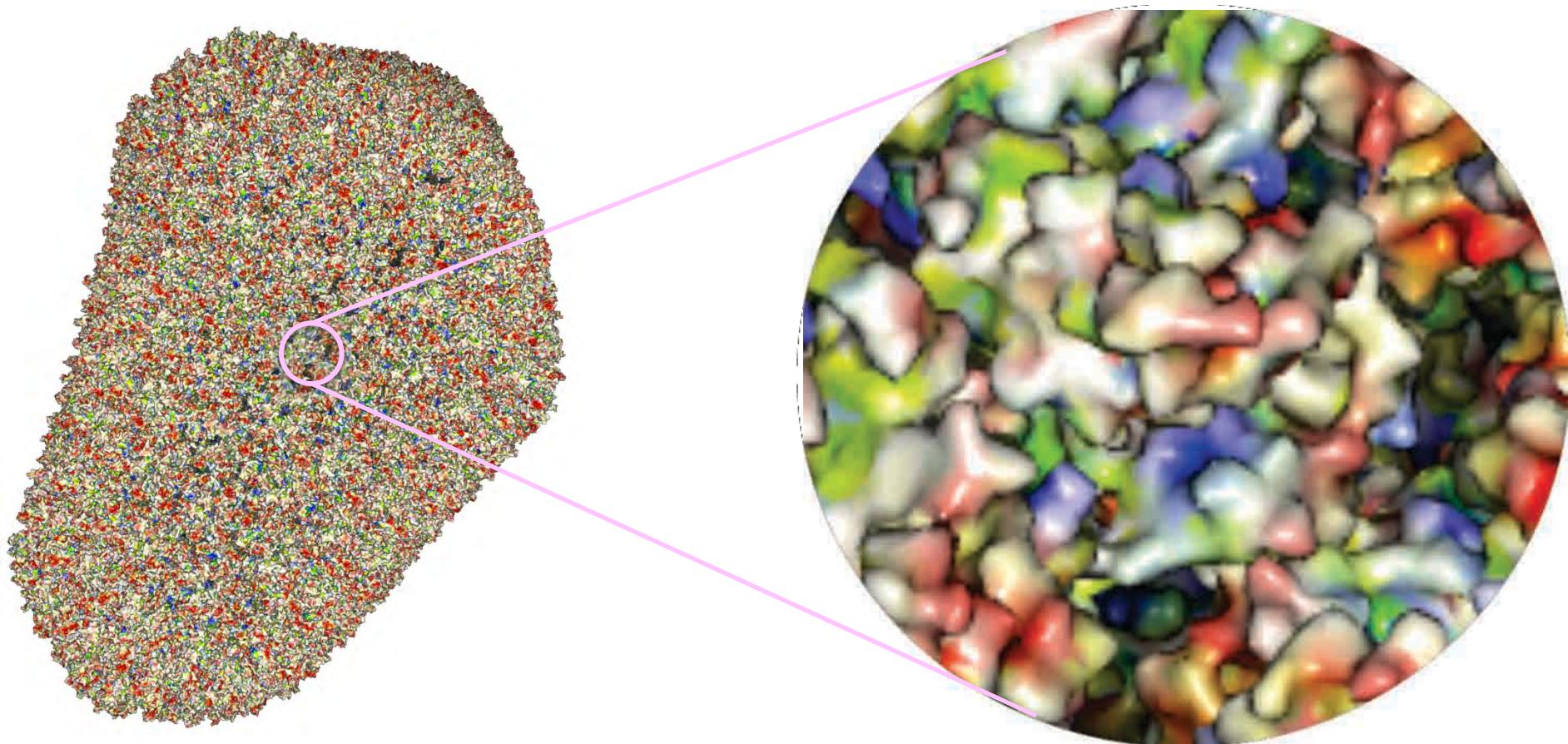
# Molmil: Molecular Viewer based on WebGL



- Based on JavaScript/WebGL.
- Supports PDB, mmCIF, PDBML formats.
- Links to PDB, chem\_comp (Compound), and ProMode Elastic.
- Outputs screenshots.
- Available for iOS8 (iPAD, iPhone etc)

# Molmil Viewer for eF-site

even for large structures (coming soon)



HIV-1 Capsid 3J3Q, 1356 chains, >2M atoms

# PDBjの活動とデータバンクの使い方

## eProtS: 日本語と英語による蛋白質構造の解説

**PDBj** 一覧 最近の更新 

[eProtS](#)

**Topics**

- [タンパク質入門](#)
- [タンパク質の構造入門](#)
- [タンパク質リスト](#)
- [オリジナルeProtSの分類](#)
- [生物学的機能による分類](#)
- [生化学的機能による分類](#)
- [サブマップ](#)
- [お問い合わせ](#)
- [RSS](#)
- [用語解説](#)
- [Molecule of the Month](#)
- [リンク集](#)
- [エントリーの追加・編集](#)

[English version eProtS](#)

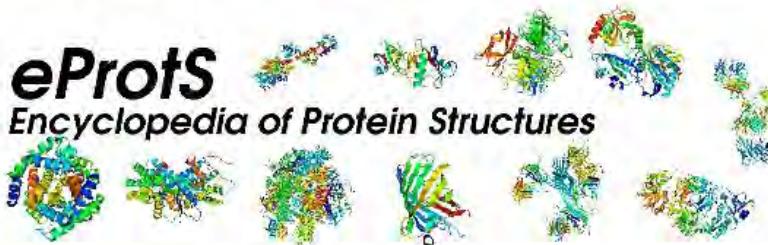
---

**Search**

**Recent Changes**

PDB:3p0c  
PDB:2yad  
PDB:2ic8  
PDB:3aek  
[生物学的機能による分類](#)  
[生物学的機能による分類](#)  
[オリジナルeProtSの分類](#)  
PDB:1gq  
PDB:3rs\_jirg  
PDB:2rc  
PDB:3afm  
PDB:3el8  
PDB:3ae5  
PDB:3ae9  
PDB:2efr  
PDB:3kdr  
[エントリーの追加・編集](#)  
PDB:2HYD  
PDB:2EY4  
PDB:2R4P

**eProtS Encyclopedia of Protein Structures**



English version eProtS

**eProtS: タンパク質構造百科事典**

eProtS(タンパク質構造百科事典)とは、生物学的に重要なタンパク質を選び、その立体構造を表示するとともに、タンパク質の構造と機能についてわかりやすく解説したものです。タンパク質構造の専門家でない方にも利用しやすいように、日本語版と英語版を用意しています。Jmolという表示ソフトを使えるシステムでは、マウスによるタンパク質分子の回転・移動・拡大縮小操作をリアルタイムで行うこともできます。(IE 5.0 or NN 6.0 以上推奨)

**目次**

- [eProtS: タンパク質構造百科事典](#)
  - [目次](#)
  - [一般的の読者の方へ](#)
  - [構造生物学の専門家の方へ](#)
- [タンパク質入門](#)
- [タンパク質の構造入門](#)
- [タンパク質リスト](#)
- [オリジナルeProtSの分類](#)
- [生物学的機能による分類](#)
- [生化学的機能による分類](#)
- [用語解説](#)
- [エントリーの追加・編集](#)

**一般的の読者の方へ**

# タンパク質名

ミオグロビン

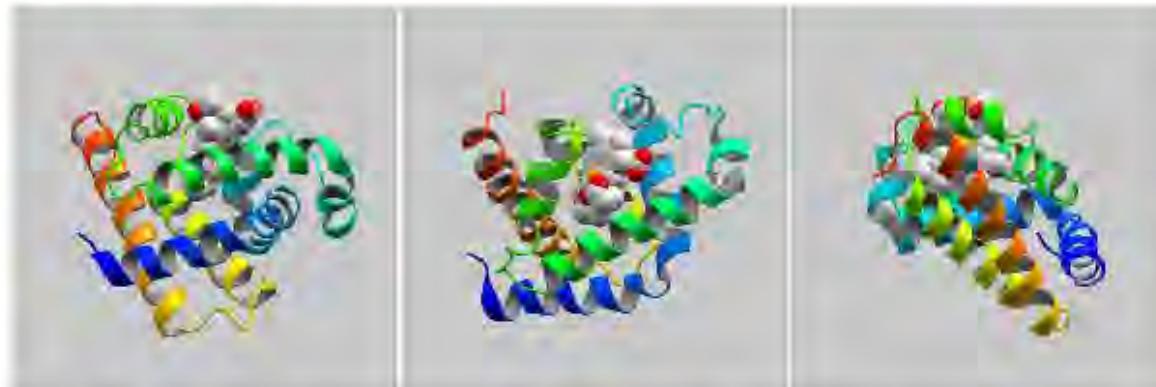
## 生物種

マッコウクジラ

## 生物学的役割

酸素がなくても生育できる生物もいれば、酸素がない中で起こる生命活動もある。嫌気性細菌や嫌気的代謝がその例である。しかしながら、大多数の生物にとって酸素は生命活動には必要不可欠な要素である。哺乳動物ではヘモグロビンが酸素を肺から末梢器官にまで運んでいる。しかし、時によっては酸素を蓄積しておくことができなくてはならないこともある。たとえば、鯨のような海生哺乳類は一度水面に酸素補給に来るとその後は酸素補給なしで水中にもぐっていられる。このような行動のために自然はミオグロビンという蛋白質を作り出したのである。ミオグロビンは主に筋肉組織に分布している。ミオグロビンはその構造が解かれた2番目の蛋白質であり、そのときのミオグロビンはマッコウクジラの筋肉組織からとったものであった。ミオグロビンは単量体の単一ポリペプチド鎖でできた蛋白質で、ヘモグロビンβサブユニットに非常に似ている。ミオグロビンはヘモグロビンと同じように酸素を結合する。中央に酸素を結合する鉄原子を持つプロトボルフィリン(ヘム)グループが分子にくっついている。ヘモグロビン内のヘムのように、ミオグロビンのヘムも一酸化炭素、二酸化炭素、シアノ化物のような酸素以外の分子を結合することができる。一酸化炭素やシアノ化物への親和性は酸素への親和性よりも大きいので、これらによる酸素結合の阻害により死に至ることもある。

## 立体構造の特徴



ミオグロビンは円盤のような形をしている。その厚さは約20 Å (1 Åは100億分の1メートル)で直径は約35 Åである。ヘムグループは2つのヘリックスでできている表面ポケットの中に円盤の平面に対して約90°の角度で入り込んでいる。ヘモグロビンで酸素結合による効果として見られる現象がミオグロビンでは見られないものもある。ミオグロビンは単量体であるので、酸素結合の際にサブユニット間での共同化がみられない。ヘモグロビンは四量体であり、おのののサブユニットがヘムグループを持っているのでこういった共同化がみられる。ヘモグロビンでもミオグロビンでもいわゆる末端ヒスチジンと呼ばれるヒスチジン残基がヘムグループの近くにある。しかしながら、象のミオグロビンではこのヒスチジンがグルタミンに置き換わっており、隣のロイシンがフェニルアラニンに置き換わっている([PDBjMire:1EMY](#))。そのように置き換わっているにも関わらず酸素結合特性は通常のミオグロビンと非常に似ている。

# PDBjの活動とデータバンクの使い方

## MoM: RCSB-PDBの記事を日本語で同時配信

### 今月の分子

このサイトはRCSBの David S. Goodsell博士による「[Molecule of the Month](#)」を日本語に訳したものです。社会で話題となっている内容に関わる分子を蛋白質構造データバンク（PDB）から選び、機能と構造に関して解説しています。転載・引用については[利用規約](#)をご覧下さい。

2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2004 2003 2002 2001 2000

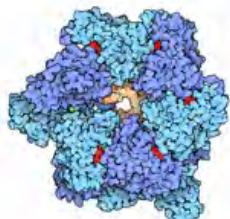
**2014**



- 180: TALエフェクター (TAL Effectors)
- 179: メチル補酵素M還元酵素 (Methyl-coenzyme M Reductase)
- 178: エボラウイルスたんぱく質 (Ebola Virus Proteins)
- 177: アポトソーム (Apoptosomes)
- 176: ダイニン (Dynein)
- 175: 微小管 (Microtubules)
- 174: GFP様たんぱく質 (GFP-like Proteins)
- 173: アクアポリン (Aquaporin)
- 172: RecAとRad51 (RecA and Rad51)
- 171: 神経伝達物質輸送体 (Neurotransmitter Transporters)
- 170: 広域中和抗体 (Broadly Neutralizing Antibodies)
- 169: HIV外被糖たんぱく質 (HIV Envelope Glycoprotein)

[ページのトップへ](#)

**2013**



- 168: DNAヘリカーゼ (DNA Helicase)
- 167: SNAREたんぱく質 (SNARE Proteins)
- 166: プロテアソーム (Proteasome)
- 165: 人工設計されたたんぱく質製の籠 (Designed Protein Cages)
- 164: セロトニン受容体 (Serotonin Receptor)
- 163: HIVカプシド (HIV Capsid)
- 162: ダームシジン (Dermcidin)
- 161: リシン (Ricin)
- 160: アクチノマイシン (Actinomycin)
- 159: エリスロクロルオリン (Erythrocyruorin)
- 158: プロトン開閉型尿素チャネル (Proton-Gated Urea Channel)
- 157: 転移伝令RNA (Transfer-Messenger RNA, tmRNA)

[ページのトップへ](#)

# 今月の分子(Molecule of the Month) 2014 / 10: No. 178□

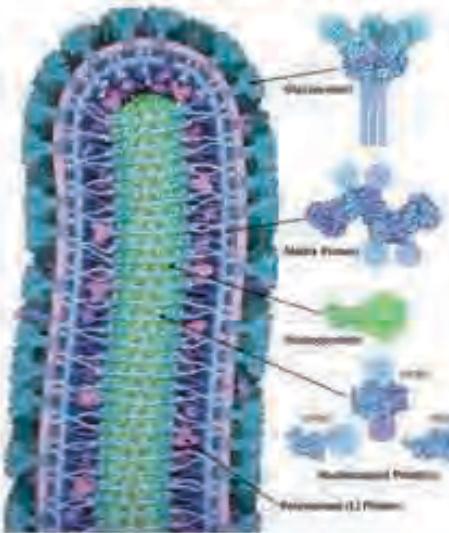
## 178: エボラウィルスたんぱく質 (Ebola Virus Proteins)

このページはRCSB の David S. Goodsell博士による「Molecule of the Month」2014年10月の記事を日本語に訳したもので、転載・引用については利用規約をご覧下さい。

「今月の分子」一覧に戻る / この記事のRCSBオリジナルサイト(英語)を見る □

エボラウイルス (ebola virus)のゲノムには7種類のたんぱく質を作るための指示情報が含まれていて、これらのたんぱく質とゲノムRNAが集まって最も致死的なウイルスの一つになる。エボラウイルスは感染先の細胞から盗んできた膜と、その表面から突き出したウイルス自身の複たんぱく質とで囲まれている。基質たんぱく質の層が膜を内側から支え、中心部分にあるRNAゲノムの保持と遺伝を行う筒状のヌクレオカプシド (nucleocapsid)をつかんでいる。

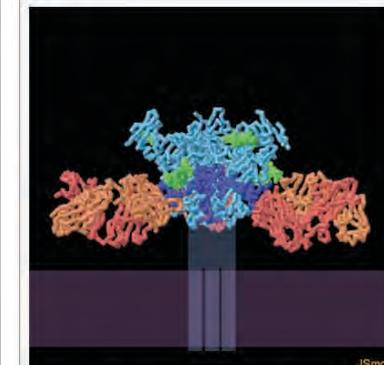
翻訳 工藤高裕 (PDBj)



### 細胞への侵入

エボラウイルスの複たんぱく質は細胞表面の受容体に結合し、ワイルスゲノムを細胞の中に入れる役割を担う。このたんぱく質を持つ特徴は、インフルエンザの赤血球凝集素 (hemagglutinin、ヘマグルチニン)や HIV外被膜たんぱく質 (HIV Envelope Glycoprotein)のような他のウイルスが持つ複合たんぱく質と多くの点で共通している。ウイルス表面から突き出し、免疫系から隠すために偽装でおおわれている。これは非常に活動的なたんぱく質でもあり、細胞表面に結合すると形が変わり、ウイルスと細胞を引き寄せたままの膜が融合できるようになる。ここに示す構造(PDBエントリー 3csy)には、複たんぱく質の受容体結合部分と膜融合機能部分が含まれる。通常は膜鎖がこの複たんぱく質をおおっているが、ここでは総合化できるようにするために、膜鎖の大半を含む小さなドメインは除去されている。

構造を見る



\*上の画像をクリックすると画像を対話的操作ができるモードに切り替えることができます。

エボラウイルスの表面に対する方法は、IgM、ワクチンとともになかなか見つかっていない。複たんぱく質はウイルスの表面にあって抗体 (antibody)が標的があるので、ワクチンの主な作用対象となる。ここに示す構造(PDBエントリー 3csy)は、エボラウイルスに感染し生き残ったヒトから得られた中和抗体 (IgG、抗体)を含んでいる。抗体は複たんぱく質の下側に結合するが、この部分は通常炭水化物ではありません。結合過程に不可欠な部分である。うるさいのは、ワクチン接種によって患者にこの種の抗体を作らさせることができるのがどう。図の下のボタンをクリックして対話的操作ができる準備に切り替えるとより詳しくみることができます。



□□□□2015年2月初版□□

# PDB flat ファイルの具体例1

```

HEADER      HYDROLASE          21-AUG-00   1FN8
TITLE       FUSARIUM OXYSPORUM TRYPSIN AT ATOMIC RESOLUTION
COMPND     MOL_ID: 1;
COMPND     2 MOLECULE: TRYPSIN;
COMPND     3 CHAIN: A;
COMPND     4 EC: 3.4.21.4;
COMPND     5 MOL_ID: 2;
COMPND     6 MOLECULE: GLY-ALA-ARG;
COMPND     7 CHAIN: B;
COMPND     8 ENGINEERED: YES
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: FUSARIUM OXYSPORUM;
SOURCE      3 ORGANISM_COMMON: FUNGUS;
SOURCE      4 MOL_ID: 2;
SOURCE      5 SYNTHETIC: YES
KEYWDS     BETA BARREL
EXPDAT     X-RAY DIFFRACTION
AUTHOR     W.R.RYPNIEWSKI,P.OESTERGAARD,M.NOERREGAARD-MADSEN,M.DAUTER,
AUTHOR     2 K.S.WILSON
REVDAT     1 07-FEB-01 1FN8      0
JRNL        AUTH  W.R.RYPNIEWSKI,P.OESTERGAARD,M.NOERREGAARD-MADSEN,
JRNL        AUTH 2 M.DAUTER,K.S.WILSON
JRNL        TITL  FUSARIUM OXYSPORUM TRYPSIN AT ATOMIC RESOLUTION AT
JRNL        TITL 2 100 AND 283 K: A STUDY OF LIGAND BINDING
JRNL        REF   ACTA CRYSTALLOGR., SECT.D      V. 57      8 2001
JRNL        REFN  ASTM ABCRE6  DK ISSN 0907-4449
  
```

## PDB flat ファイルの具体例2

```

CRYST1  58.390   86.700   46.270   90.00   90.00   90.00 P 21 21 2      4
ORIGX1   0.017126  0.000000  0.000000          0.000000
ORIGX2   0.000000  0.011534  0.000000          0.000000
ORIGX3   0.000000  0.000000  0.021612          0.000000
SCALE1   0.017126  0.000000  0.000000          0.000000
SCALE2   0.000000  0.011534  0.000000          0.000000
SCALE3   0.000000  0.000000  0.021612          0.000000
ATOM     1  N  PRO A   1    29.061  39.981   4.981  1.00 28.69
ATOM     2  CA PRO A   1    29.970  38.922   4.561  1.00 29.08
ATOM     3  C  PRO A   1    29.325  38.106   3.429  1.00 29.19
ATOM     4  O  PRO A   1    28.097  38.168   3.298  1.00 29.87
ATOM     5  CB PRO A   1    30.106  38.013   5.789  1.00 29.07
ATOM     6  CG PRO A   1    28.749  38.112   6.413  1.00 28.59
ATOM     7  CD PRO A   1    28.387  39.600   6.246  1.00 29.21
ATOM     8  N  GLN A   2    30.153  37.412   2.681  1.00 28.13
ATOM     9  CA GLN A   2    29.636  36.572   1.593  1.00 27.95
ATOM    10  C  GLN A   2    29.861  35.139   2.082  1.00 27.28
ATOM    11  O  GLN A   2    31.038  34.773   2.266  1.00 27.61
ATOM    12  CB GLN A   2    30.373  36.787   0.305  1.00 28.43
ATOM    13  CG GLN A   2    30.346  35.501  -0.539  1.00 29.40
ATOM    14  CD GLN A   2    30.921  35.844  -1.899  1.00 29.51
ATOM    15  OE1 GLN A   2    31.894  35.283  -2.340  1.00 30.56
ATOM    16  NE2 GLN A   2    30.288  36.839  -2.518  1.00 30.01
  
```

X座標, Y座標, Z座標 (Å=10<sup>-8</sup> cm 単位)

# Data Formats of PDB data

- **PDB (conventional and flat)**
- **PDB Exchange (mmCIF)**
  - Mechanism for extension based on new demands
- **PDBML**
  - Derived from mmCIF
  - All entries converted to XML
  - Automatic translation from mmCIF data files and dictionaries
  - 3-styles of translation released

*(Westbrook, Ito, Nakamura, Henrick, Berman (2005)  
Bioinformatics, 21, 988-992)*

# New standard PDB format: PDBx/mmCIF

- Current PDB format is almost **40 years old** and does not support today's science.
- PDB Record format limitations
  - **Max. 62 chains**
  - **Max. 99,999 atoms**
  - No bond orders or chirality specified for ligands
  - No support for NMR, EM, hybrid methods, ...
  - Meta-data specification cumbersome and inflexible



- **Preserve backward compatibility where possible**
- **PDBML (XML) and RDF format files are available.**
- **Start in 2014 and the current PDB format will be phased out in 2016.**

```

ATOM      1  N   GLN A  39       24.690 -27.754  24.275  1.00 60.76      N
ATOM      2  CA  GLN A  39       23.581 -26.768  24.416  1.00 60.98      C
ATOM      3  C   GLN A  39       23.990 -25.379  23.905  1.00 59.98      C
ATOM      4  O   GLN A  39       25.070 -25.209  23.330  1.00 60.25      O
ATOM      5  CB  GLN A  39       23.136 -26.685  25.878  1.00 60.69      C
ATOM      6  N   VAL A  40       23.115 -24.395  24.122  1.00 59.58      N
ATOM      7  CA  VAL A  40       23.342 -23.010  23.690  1.00 57.26      C
ATOM      8  C   VAL A  40       24.000 -22.152  24.778  1.00 56.00      C
ATOM      9  O   VAL A  40       23.992 -20.920  24.692  1.00 55.53      O
ATOM     10  CB  VAL A  40       22.015 -22.337  23.275  1.00 57.32      C

```

PDB

```

loop_
_atom_site.group_PDB
_atom_site.id
_atom_site.auth_atom_id
_atom_site.type_symbol
_atom_site.auth_comp_id
_atom_site.auth_asym_id
_atom_site.auth_seq_id
_atom_site.Cartn_x
_atom_site.Cartn_y
_atom_site.Cartn_z
_atom_site.pdbx_PDB_model_num
_atom_site.occupancy
_atom_site.pdbx_auth_alt_id
_atom_site.B_iso_or_equiv

```

PDBx/mmCIF

	1	N	N	GLN	A	39	24.690	-27.754	24.275	1	1.000	.	60.760
ATOM	2	CA	C	GLN	A	39	23.581	-26.768	24.416	1	1.000	.	60.980
ATOM	3	C	C	GLN	A	39	23.990	-25.379	23.905	1	1.000	.	59.980
ATOM	4	O	O	GLN	A	39	25.070	-25.209	23.330	1	1.000	.	60.250
ATOM	5	CB	C	GLN	A	39	23.136	-26.685	25.878	1	1.000	.	60.690
ATOM	6	N	N	VAL	A	40	23.115	-24.395	24.122	1	1.000	.	59.580
ATOM	7	CA	C	VAL	A	40	23.342	-23.010	23.690	1	1.000	.	57.260
ATOM	8	C	C	VAL	A	40	24.000	-22.152	24.778	1	1.000	.	56.000
ATOM	9	O	O	VAL	A	40	23.992	-20.920	24.692	1	1.000	.	55.530
ATOM	10	CB	C	VAL	A	40	22.015	-22.337	23.275	1	1.000	.	57.320
ATOM	11	N	N	ALA	A	41	24.560	-22.804	25.797	1	1.000	.	54.570

# wwPDB Service site for a new format

<http://mmcif.wwpdb.org/> or <http://mmcif.pdbj.org/>



The screenshot shows the homepage of the PDBx/mmCIF Dictionary Resources. At the top, there's a navigation bar with links for "Home", "Dictionaries", "Documentation", "Downloads", and "Contact Us". To the right of the navigation is the "wwPDB" logo. Below the navigation is a search bar with a magnifying glass icon and the placeholder text "Search current dictionary". The main content area has a light gray background and features a large, stylized title "PDBx/mmCIF Dictionary Resources" in blue and purple. Below the title is a descriptive paragraph: "This site provides information about the format, dictionaries and related software tools used by the Worldwide Protein Data Bank (wwPDB) to define data content for deposition, annotation and archiving of PDB entries." At the bottom of this section is a green button with the text "Browse the current dictionary »".

## Dictionaries

- [Browse the current dictionary»](#)
- [Download/view all dictionaries »](#)
- [Search dictionaries»](#)

## Documentation

- [PDB > PDBx/mmCIF correspondences »](#)
- [PDBx/mmCIF for large structures »](#)
- [Software resources »](#)
- [C++ » and Python » programming examples](#)
- [File syntax » and dictionary organization »](#)
- [Atomic » and molecular » descriptions](#)
- [References »](#)
- [Glossary »](#)

## FAQs

Questions about PDBx/mmCIF format, and data content, or software tools?  
Check out the [FAQ»](#)

# PDBx/mmCIF Software Support

- **Phenix and Refmac** – produce native PDBx files for deposition
- **MMDB** - macromolecular object library in CCP4
- **iotbx.cif/ucif** - CCTBx C++/Python IO library with dictionary validation
- **CCIF** – CCP4 C++ library with FORTRAN support and dictionary validation
- **CBFLib** - ANSI-C library for CIF & imgCIF files
- **mmLIB** - Python toolkit supporting CIF & mmCIF
- **BioPython** - Python toolkit for computational biology
- **PyCifRW** - Python CIF/mmCIF parsing tools
- **BioJava** - Java mmCIF IO package
- **STAR::Parser** – Perl mmCIF parser and molecular object library
- **RCSBTools** - C++/Python parsing and dictionary validation tools plus many other supporting format conversion and data management applications
- **Visualization** - **UCSF Chimera, Jmol, OpenRasMol, Coot, CCP4mg, jV, Molmil**

# wwPDB Service site for a new format

<http://mmcif.pdbj.org/converter/index.php?l=en>

## PDBx/mmCIF

Home

Dictionaries

Documentation

Downloads

Format Conversion

Contact Us

English 日本語

## PDB format - PDBx/mmCIF conversion service

You can convert a molecular structural data into another format. The type of uploaded file is determined automatically. When the type is mmCIF and PDB format, it is converted into PDB format and mmCIF, respectively. The gzip compressed files that end ".gz" of the name, are also available. When they are gzipped, the converted files are also gzipped.

### 1. Specify a source file to convert

Specify a source file to convert. The maximum size of the file is 1GB.

選択... ファイルが選択されていません。

### 2. Confirm the contents of operations

### 3. Execute conversion & Download the converted file

When you convert a large structure mmCIF file into the PDB format, It will be treated as following:

- When it includes more than 99999 atoms, all the atomids larger than 99999 are rewrited to 99999.
- When the chain id (auth\_asym\_id) has two letters, it will be described as it is by using unused 21th column and defined 22th column.

# 創薬への応用

## 蛋白質の形に基づく創薬

(SBDD: Structure Based Drug Design)

(SGDD: Structure Guided Drug Development)

**Viracept:** アグロン社, **Agenerase:** Vertex社 & USA-Kissei

AIDS

タミフル/ノイラミニダーゼ(2qwk)

標的分子:HIV-1 プロテアーゼ

**Gleevec:** ノバルティス ファーマ社

慢性骨髓性白血病

標的分子:BCA-Abl、Tyrキナーゼ

**Iressa:** アストラゼネカ社

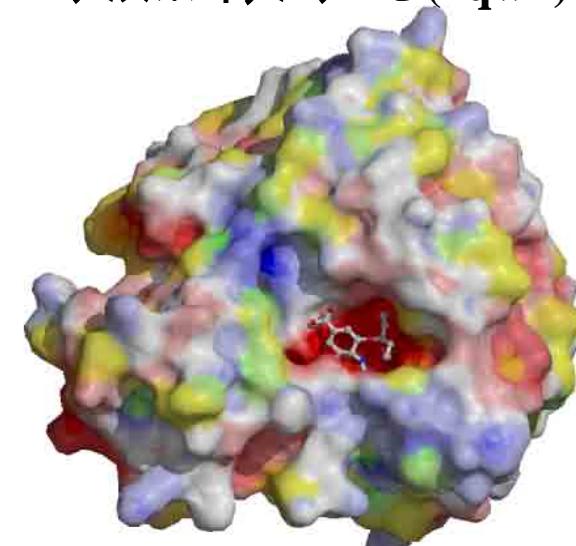
肺がん

標的分子:EGFR、Tyrキナーゼ

**Tamiflu:** スイスF.Hoffmann-La Roche社、米国Gilead Sciences社

インフルエンザ

標的分子:糖分解酵素



# 構造生命科学： 生体高分子の構造を基に進める生命科学

神谷・肥後・福西・中村 著

タンパク質計算科学



(共立出版2009年8月初版)

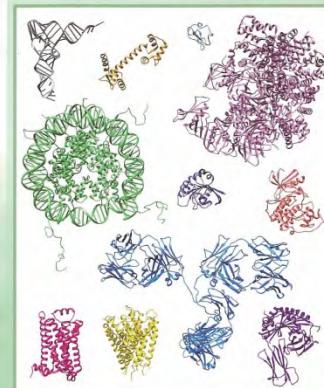
中村 編

見てわかる 構造生命科学

## 見てわかる 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



化学同人

(化学同人2014年4月初版)

工藤・西川・中村 訳

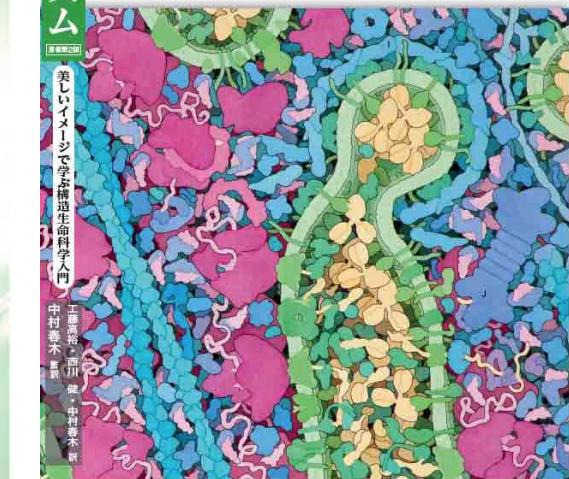
生命のメカニズム  
THE MACHINERY OF LIFE

原著第2版

## 生命のメカニズム

美しいイメージで学ぶ構造生命科学入門

David S. Goodsell ■  
中村春木 訳  
工藤高裕・西川健・中村春木 訳



(シナジー2015年2月初版)

# PDBjの スタッフ

## • 総括責任者

- [中村 春木](#) (大阪大学蛋白質研究所・教授)

## • PDBjデータベース構築グループ

- [中川 敦史](#) (大阪大学蛋白質研究所・教授)
- 五十嵐 令子 (大阪大学蛋白質研究所・特任研究員)
- 見学 有美子 (大阪大学蛋白質研究所・特任研究員)
- 張 羽澄 (大阪大学蛋白質研究所・特任研究員)
- 池川 恵代 (大阪大学蛋白質研究所・特任研究員)
- 佐藤 純子 (大阪大学蛋白質研究所・特任研究員)

## • PDBj国際的な運営高度化グループ

- [金城 玲](#) (大阪大学蛋白質研究所・准教授)
- 岩崎 繁治 (大阪大学蛋白質研究所・准教授)
- 鈴木 博文 (大阪大学蛋白質研究所・特任研究員)
- 山下 鈴子 (大阪大学蛋白質研究所・特任技術専門職員)
- 工藤 高裕 (大阪大学蛋白質研究所・特任研究員)
- Bekker, Gert-Jan (大阪大学蛋白質研究所・特任研究員)

## • BMRBデータベース管理運営グループ

- [鶴原 敏道](#) (大阪大学蛋白質研究所・教授)
- 陶久津 秀雄 (大阪大学蛋白質研究所・客員教授)
- 児嶋 長次郎 (大阪大学蛋白質研究所・准教授)
- 小林 直宏 (大阪大学蛋白質研究所・特任研究員)
- 岩田 武史 (大阪大学蛋白質研究所・特任研究員)
- 横地 政志 (大阪大学蛋白質研究所・特任研究員)

## • 研究協力者

- [鈴湖 博](#) (早稲田大学社会科学総合学術院・教授) for ProMode
- [猪瀬 広](#) (北里大学理学部・准教授) for ProMode
- 伊藤 暢嗣 (東京医科歯科大学大学院・教授)
- [木下 駿吾](#) (東北大学大学院情報科学研究科・教授) for eF-site
- [Standley, M.Daron](#) (京都大学ウイルス研究所・教授) for SeqNavi, StructNavi, SeSAW and ASH
- 加藤 和貴 (大阪大学免疫学フロンティア研究センター・准教授) for MAFFTash

## • 事務職員

- 晴氣 菜穂子 (大阪大学蛋白質研究所・特任事務職員)