

PDBj (PDB Japan) と wwPDB の 活動について

中村 春木

Haruki Nakamura

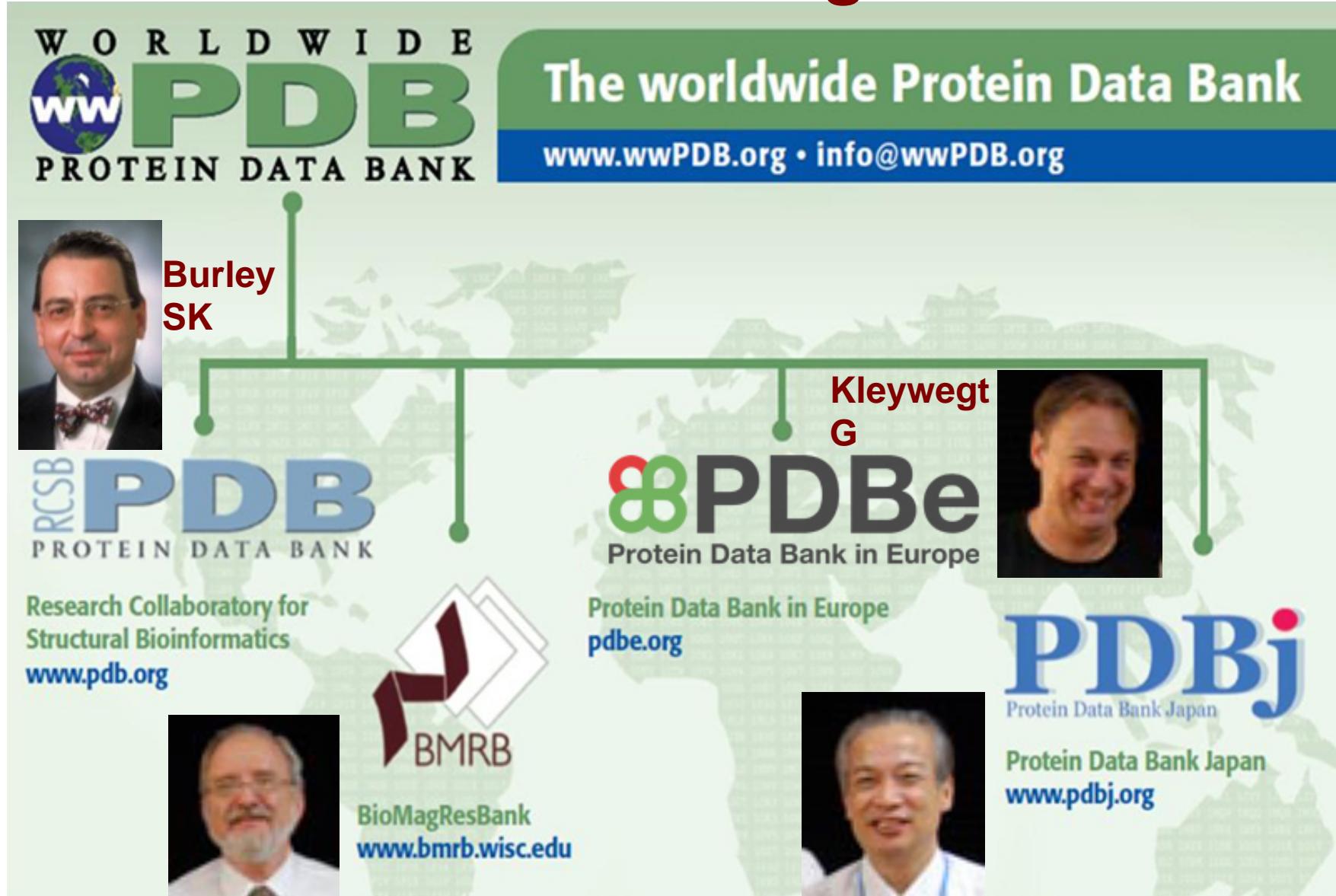


大阪大学蛋白質研究所

Institute for Protein Research

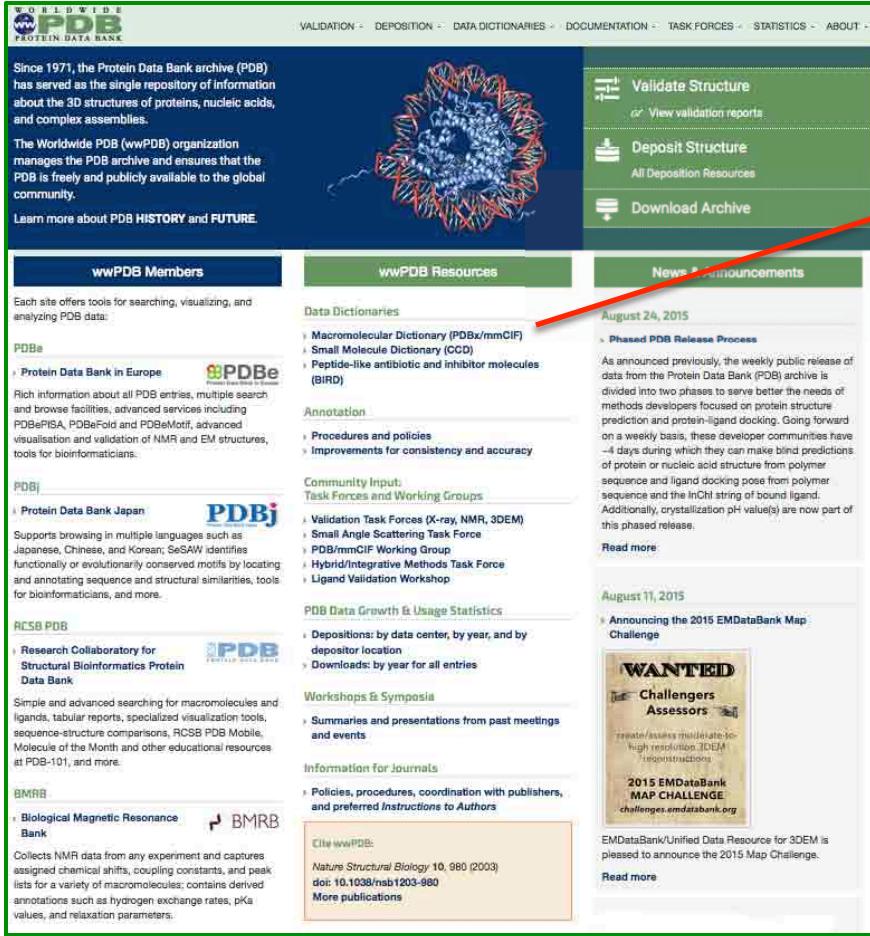
<http://pdbj.org/>

<http://wpdb.org/>



wwPDB partners and their heads

The wwPDB web page



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

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:

- PDBe**
 - Protein Data Bank in Europe
 - Rich information about all PDB entries, multiple search and browse facilities; advanced services including PDBePISA, PDBeFold and PDBeMotif, advanced visualization and validation of NMR and EM structures, tools for bioinformaticians.
- PDBj**
 - Protein Data Bank Japan
 - Supports browsing in multiple languages such as Japanese, Chinese, and Korean; SeSAW identifies functionally or evolutionarily conserved motifs by locating and annotating sequence and structural similarities, tools for bioinformaticians, and more.
- RCSD 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.

wwPDB Resources

Data Dictionaries

- Macromolecular Dictionary (PDBx/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

wwPDB Data Growth & Usage Statistics

- Depositions: by data center, by year, and by depositor location
- Downloads: by year for all entries

Workshops & Symposia

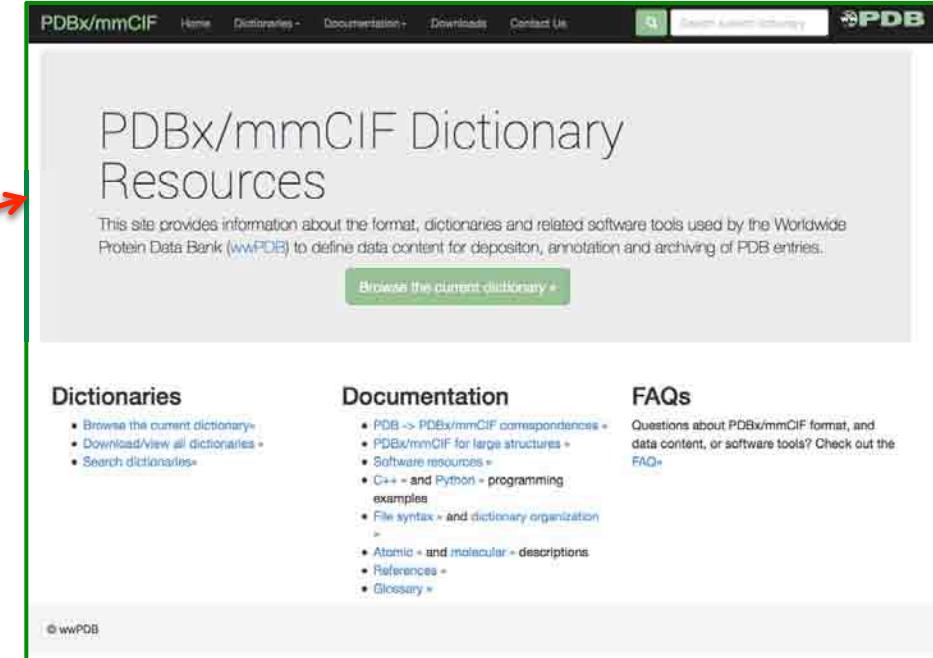
- Summaries and presentations from past meetings and events

Information for Journals

- Policies, procedures, coordination with publishers, and preferred Instructions to Authors

Cite wwPDB:

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



PDBx/mmCIF Home Dictionaries Documentation Downloads Contact Us Search Advanced Search

PDBx/mmCIF Dictionary Resources

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.

[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 [FAQs](#).

© wwPDB

wwpdb.org

PDB Depositions

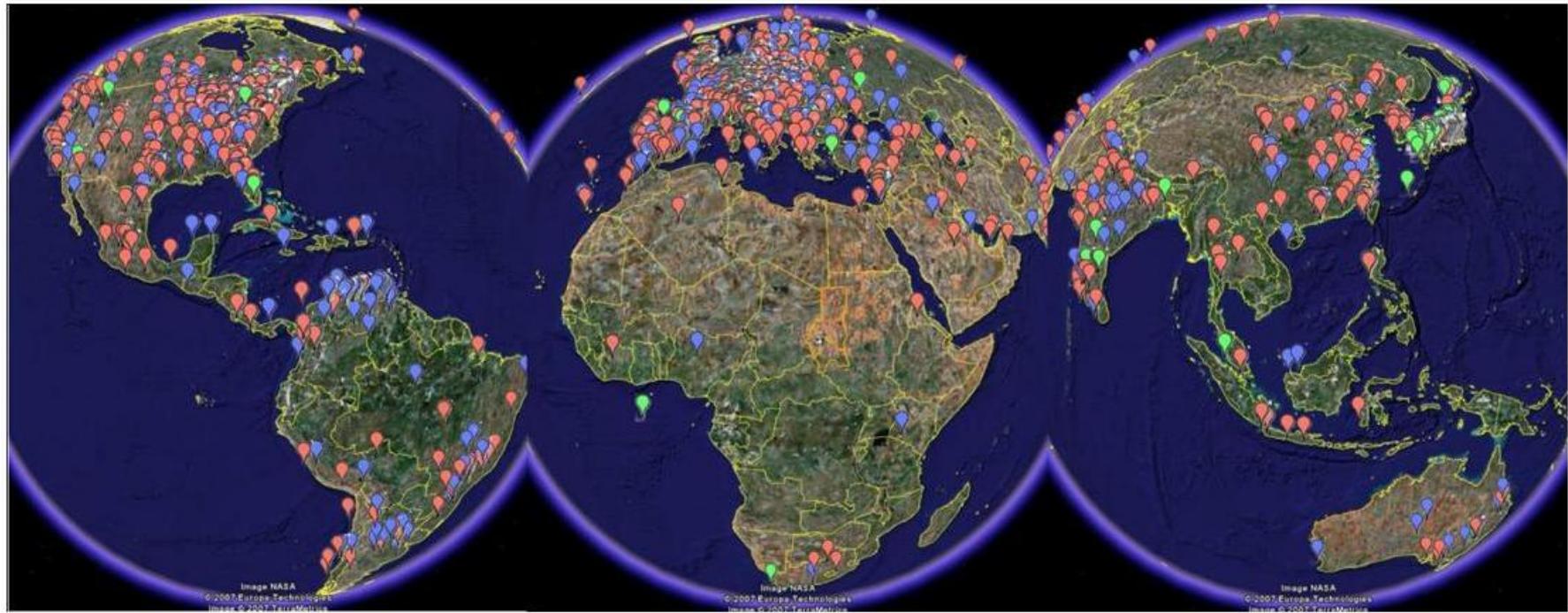
- As of 2015 region-based processing of D&A-deposited entries:
 - RCSB PDB: Americas & Oceania
 - PDBe: Europe & Africa
 - PDBj: Asia



Year	Total Depositions	Processed By		
		RCSB PDB	PDBj	PDBe
2000	2983	2297	158	528
2001	3287	2408	383	496
2002	3565	2401	657	507
2003	4830	3135	1026	669
2004	5508	3082	1614	812
2005	6678	3563	2110	1005
2006	7282	4252	1945	1085
2007	8130	4703	2299	1128
2008	7073	4106	1994	973
2009	8300	5069	2173	1058
2010	8878	5464	2041	1373
2011	9250	5938	1816	1496
2012	9972	6408	1888	1676
2013	10566	6652	2128	1786
2014	10364	6040	1779	2545
2015	7575	3492	1336	2748
TOTAL	114241	69010	25347	19885

114K 69K 25K 20K □

FTP + Rsync Entry Downloads



■ RCSB PDB

2014: 347 million
2013: 312 million
2012: 298 million
2011: 282 million
2010: 159 million

■ PDBe

2014: 100 million
2013: 81 million
2012: 46 million
2011: 59 million
2010: 34 million

■ PDBj

2014: 57 million
2013: 40 million
2012: 21 million
2011: 38 million
2010: 16 million

wwPDB Task Forces

Method-specific (Validation) Task Forces have been convened to collect recommendations and develop consensus on additional validation that should be performed, and to identify software applications to perform validation tasks, and to discuss archival needs and opportunities for non-traditional techniques.



Task Force	Meeting/Workshop	Chair(s)/Membership	Outcome
X-ray Validation Task Force	2008 (2015)	Randy Read (Univ of Cambridge) 17 members	(2011) <i>Structure</i> 19: 1395-1412
NMR Validation Task Force	2009, 2011, 2013 (x2), 2015	Gaetano Montelione (Rutgers) Michael Nilges (Institut Pasteur) 10 members	(2013) <i>Structure</i> , 21: 1563-1570□
3DEM Validation Task Force	2010	Richard Henderson (MRC-LMB) Andrej Sali (UCSF) 21 members	(2012) <i>Structure</i> 20: 205-214
Small-Angle Scattering Task Force	2012, 2014	Jill Trewella (Univ Sydney) 6 members	(2013) <i>Structure</i> 21: 875-881
Hybrid Methods Task Force	2014	Andrej Sali (UCSF), Torsten Schwede (Univ Basel), Jill Trewella (Univ Sydney) 27 members	(2015) <i>Structure</i> 23: 1156-1167

wwPDB/CCDC/D3R Ligand Validation Workshop

Meeting Objectives: To bring together co-crystal structure determination experts from Academe and Industry with X-ray Crystallography and Computational Chemistry Software Developers to discuss, develop, and recommend:

- Best practices for PDB archive deposition/validation of co-crystal structures
- Editorial/Refereeing/Publication standards for co-crystal structures
- Improvements in ligand representation across the PDB Archive

Rutgers July 30-31, 2015

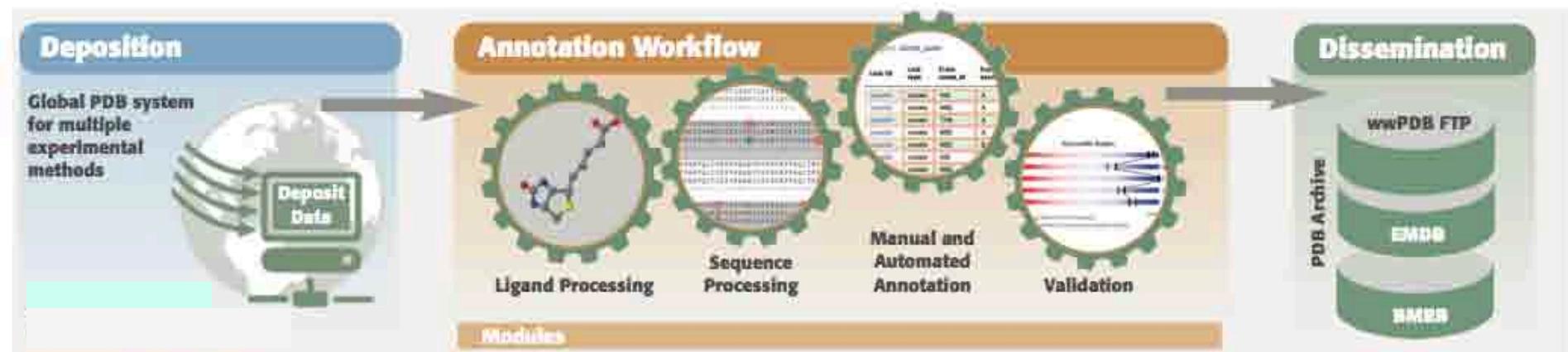




wwPDB Advisory Committee 2015

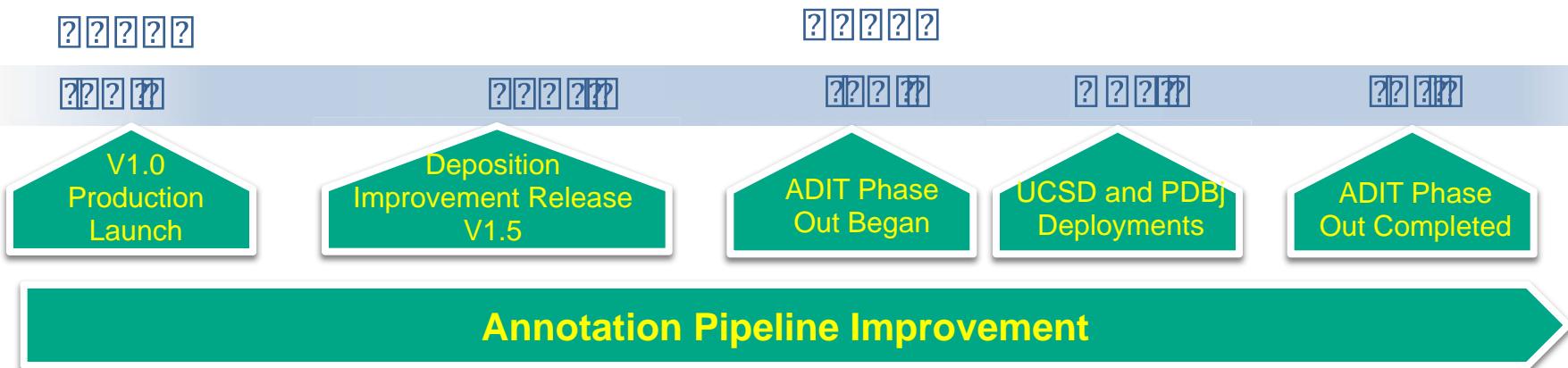
**on Oct 2, 2015
at IPR, Osaka Univ.**

新システムによる データ処理の効率化と品質化



- Enables workload balancing and increased productivity
- Better quality assurance of ligand chemistry and polymer sequences
- PDBx/mmCIF is the master file format
- Validation suites based on recommendations from expert task forces; X-ray validation pipeline is available as a stand-alone server
- System will support all accepted experimental methods

X-ray D&A Milestones



- **January 2014: X-ray production release (V1.0)**
- **Fall 2014: X-ray improvement release (V1.5)**
- **January 2015: RCSB PDB and PDBj ADIT systems phase out began**
- **April 2015: System deployments at UCSD and PDBj**
- **July 2015: RCSB PDB and PDBj ADIT systems phase out concluded**
- **2015: Annotation pipeline improvement throughout year**

Release of sequence and bound ligand information prior to release of structure

- **Phase I: Every Saturday by 3:00 UTC**, for every new entry, the wwPDB website provides: **sequence(s)** (amino acid or nucleotide) for each distinct polymer and, where appropriate, the **InChI string(s)** for each distinct ligand and the **crystallization pH value(s)**
- **Phase II: Every Wednesday by 00:00 UTC**, all new and modified data entries will be updated at each of the wwPDB FTP sites.

(Started in April 2015. Crystallization pH information was added in August 2015)

Data correspondences between the PDB and CSD archives

- A new data resource is available containing correspondences between the biopolymer components and ligand molecules found in the PDB Chemical Component Dictionary (CCD) that exactly match small-molecule X-ray structures in the Cambridge Structural Database (CSD) archive
- The new **PDB Chemical Component Model** data file complements information in the PDB by providing the CSD information for matching molecular entries: **accession code** correspondences, **Cartesian coordinates** and **R-value**, **data-collection temperature** and a **disorder flag**, **SMILES** and **InChI** descriptors, and a Digital Object Identifier (**DOI**) for the citation associated with the CSD entry.

New Proposals

1. Introduction of ORCID

The Open Researcher and Contributor ID (ORCID; orcid.org) will be used to uniquely identify researchers and link them to their research work product. With definitive deployment of D&A V2.1, the wwPDB proposes to capture an ORCID identifier for each contact author at the time of deposition

A placeholder definition for the ORCID is already available in the PDBx/mmCIF V5 dictionary,

_pdbx_contact_author.identifier_ORCID

2. Federation of Data Resources for integrative/hybrid 3D structural models

- **Following-up on the Outcome of the Hybrid/Integrative Methods Task Force Workshop**

The wwPDB aims to provide a Deposition and Annotation System for collecting 3D structure model coordinates plus primary data and associated metadata (sufficient to validate the 3D structural model) from the current methods (X-ray, Neutron, NMR, and EM) plus additional methods specific to Federated Data Resources

- **Data Storage Formatting**

PDBx/mmCIF is the principal archival format for storage of 3D structure model coordinates, associated metadata, and selected primary data supporting model validation in the PDB Archive.

Federated Data Resources are free to store their method specific primary data and associated metadata in whatever way they choose to do so.

3. Versioning of PDB entries

Policy Proposal

- Every accession code includes the version number after the PDB code
- The version number is updated, only when the group of the original author(s) with the same PI revises the atomic coordinates, in addition to change ligand codes or any other metadata
 - When new experimental data are used to change the atomic coordinates, a new PDB code is adopted instead of updating the version number
 - When another group contributes an extension or reinterpretation without any additional experimental data, such a structure can be deposited with a new PDB code, only after publication from a peer-reviewed journal
 - Incorporates the archive name (PDB) explicitly within the accession code to simplify recognition in scientific literature

[Examples of proposed versioning]

Template: PDB-< pdb_code >-< version_number >

Example: PDB-1ABC-1 (original structure)

PDB-1ABC-2 (updated structure)

Regex: PDB-[0-9][A-Z0-9]{3}-[0-9]{1,n} (n~3)

PDBj @ 大阪大学蛋白質研究所

Protein Data Bank
Japan

日本蛋白質構造データバンク

<http://pdbj.org/>

The screenshot shows the PDBj homepage with a sidebar containing links for Home, Top Page, Help, FAQ, Contact, Data Sources, Help, PDBx/mmCIF, Download, New Products, and Help. The main content area includes sections for '112387' (entry ID), 'PDBj Protein Data Bank Japan', '初めての利用者のためのガイド' (Guide for new users), '必要なサービスを探す' (Find services), and search boxes for PDB, BMRB, EMDB, NMR, X-ray, Compounds, and others. Logos for DBCLS, NBDC, and BMRB are also present.

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)

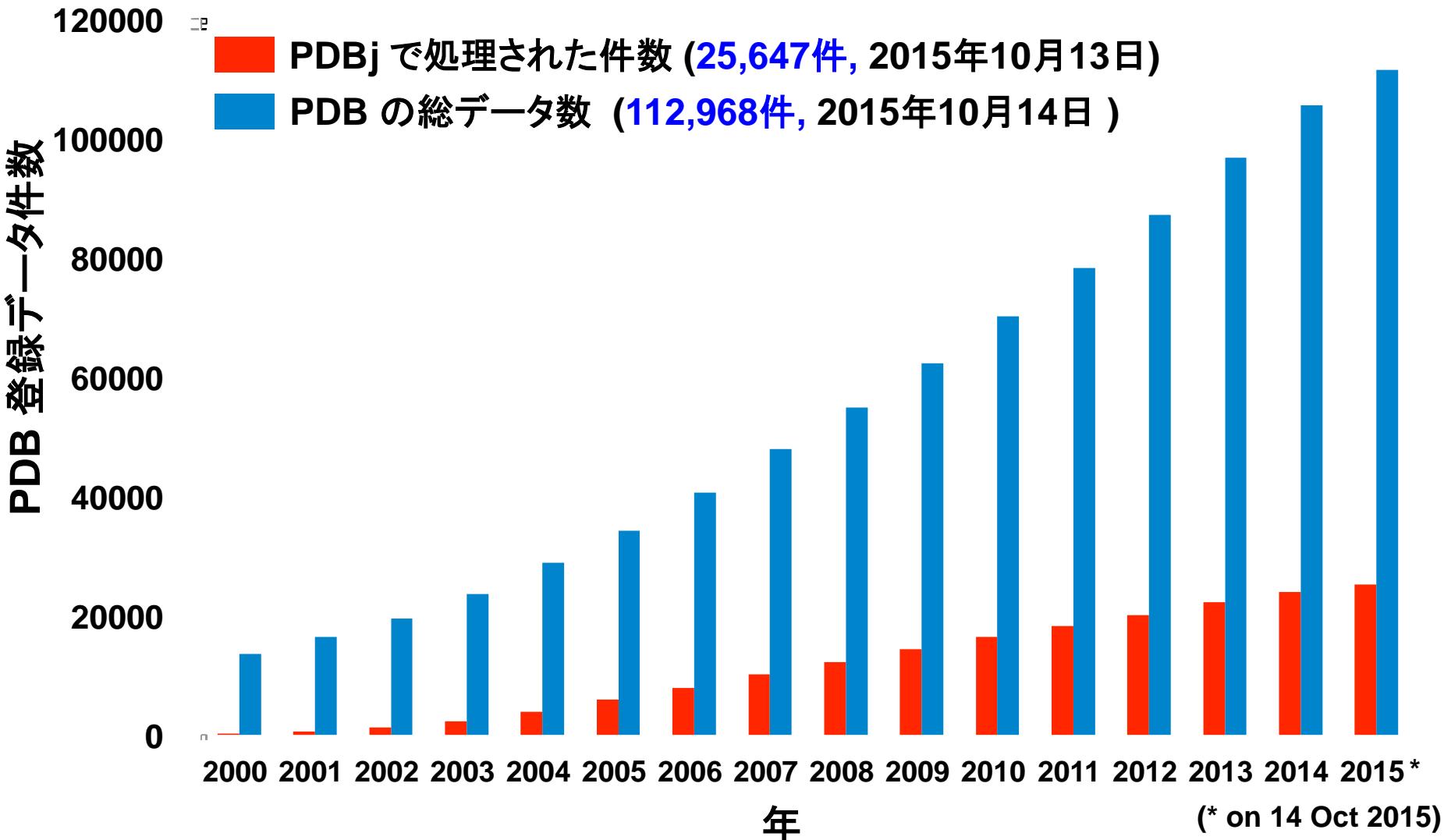


PDBj-BMRBスタッフ

PDBjの活動

- 1) “Data-in” の活動: wwPDBの一員として
 - ・品質管理がなされた登録作業の実施
 - ・新たな統一的登録システムの開発
 - ・新たな共通フォーマットの開発(XML, RDF)
- 2) “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
 - ・データベース統合化環境の開発・提供

PDBj と wwPDB の Data-in 活動



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Data-out活動

1GOF
NOVEL THIOETHER BOND REVEALED BY A 1.7 ANGSTROMS CRYSTAL STRUCTURE OF GALACTOSE OXIDASE

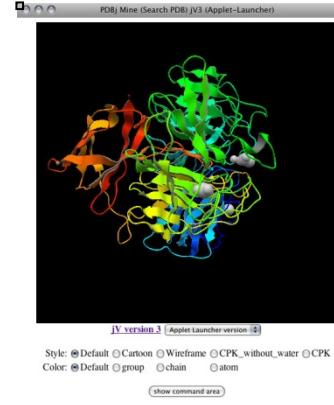
Metric **Percentile Ranks** **Value**

- Clashscore: Worse (0.3%)
- Ramachandran outliers: Worse (2.9%)
- Sidechain outliers: Worse (100.0%)
- RSRZ outliers: Better (5%)

Legend: █ Percentile relative to all X-ray structures █ Percentile relative to X-ray structures of similar resolution

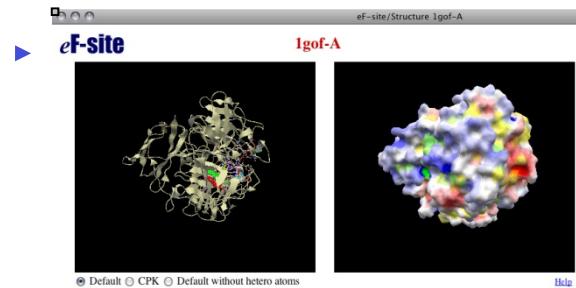
Data viewer at PDBj

Graphic viewer: jV
<http://pdbj.org/jV/>



Amino acid sequence (FASTA)

```
>1GOF:GALACTOSE_OXIDASE
ASAPIGSAISRNHWAVTCDSAQSGNECKNAIDGNKIDTPWHFTFYGANGDPPCPHPTYTDIMK
TTONVNGLNSMLPQDGDNQNGWIGIREVEYLSSDCGTNWGSFVASGSFADFDTTKYSNFETRP
ARYVRVLVAAITEEANGQFWTSIAEINVFQASASYTAPQFGLGRNGPTIDLPIPVAAAIEPTS
GRVLWMSWSYRNDAFGGSPFGGITLTTSWDFSTGIVSUDRTVTVKHOMPCFGISMODNGQIV
VSVTSDAKETKTFVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV
VVSFSVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQ
GSCDVKSAGKRSNRSVGAAPDAMGNMVAHYDAVKGKILTFGCGSPDYQOSDATTNAMIIITLG
EGFTCSNTVPSANGLYPARTHTPSVLPDGSTFTTGQRRBGPFPEDSTVPTVPEIYVPEQ
DTFYKQNPNSIVRVYHSISLLLPDGVRVNCGGGLCDDCTNNHFDAQIFTPNYLNNSGNL
ATRKPKITRTSOTSVKVGRIT1STDSSISAKSLIRYGTATHTVNTDQRRIPLTLNNNGGN
SYSFQVQPSDSDGVALPGYWMLFVMHNSAGVPSVASTIRVTO
```



Functional site	focus & details
1) A:345	<input checked="" type="radio"/> on <input type="radio"/> off
2) A:372	<input checked="" type="radio"/> on <input type="radio"/> off
3) A:495	<input checked="" type="radio"/> on <input type="radio"/> off
4) A:496	<input checked="" type="radio"/> on <input type="radio"/> off
5) A:581	<input checked="" type="radio"/> on <input type="radio"/> off
6) A:228	<input checked="" type="radio"/> on <input type="radio"/> off
7) A:390	<input checked="" type="radio"/> on <input type="radio"/> off
8) A:273	<input checked="" type="radio"/> on <input type="radio"/> off
9) A:395	<input checked="" type="radio"/> on <input type="radio"/> off
10) A:75-87	<input checked="" type="radio"/> on <input type="radio"/> off
11) A:194	<input checked="" type="radio"/> on <input type="radio"/> off
12) A:227-228	<input checked="" type="radio"/> on <input type="radio"/> off
13) A:272	<input checked="" type="radio"/> on <input type="radio"/> off

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

EM Navigator

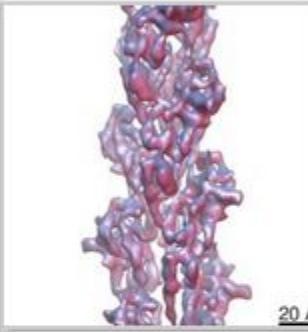
3次元電子顕微鏡データナビゲーター [[English](#)] / [日本語](#)

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

データを見る 詳しく

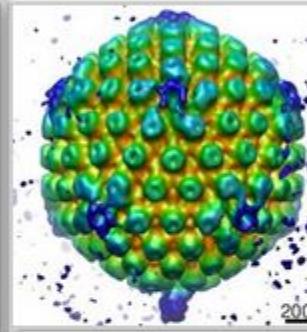
- さがす : (キーワード / EMDB ID / PDB ID) EMDBがPDBのID、あるいはキーワードを入力
- 眺める : [ギャラリー](#) [リスト](#) [表](#)

ムービースロット 方向



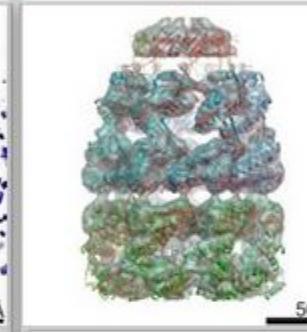
20 Å

[EMDB-1980](#)



200 Å

[EMDB-5452](#)



50 Å

[EMDB-2325](#)

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](#)

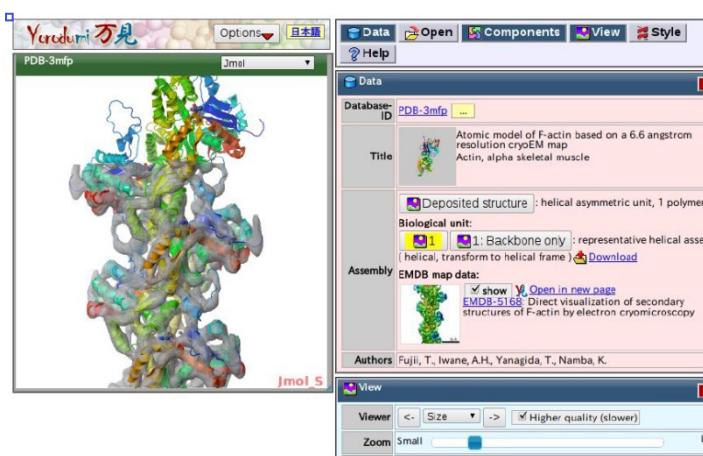
- 2013-09-04:** 公開データ



PDBjのサービス・ツール



EM Navi: 電子顕微鏡画像EMDB



Yorodumi: PDBとEMDBの統合化 **GIRAF:**類似リガンド結合部位

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:

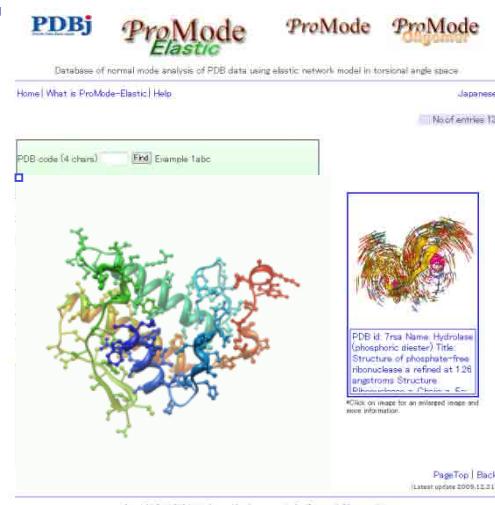
Chain IDs (optional): (comma-separated multiple IDs [e.g., "A,B"] or "all" are allowed.)

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



ProMode:基準振動解析



eF-site:分子表面DB



Submission STEP-1:

Specify a PDB format file:

E-mail address:

Keyword #1:

Title: (optional)

eF-seek:類似分子表面検索

Integration with EMDB: Search for similar SHAPE

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

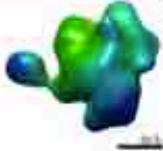
Similar shapes from 195,658 images

 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](#)

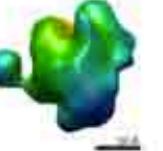
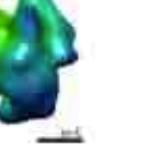


Search result

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](#)

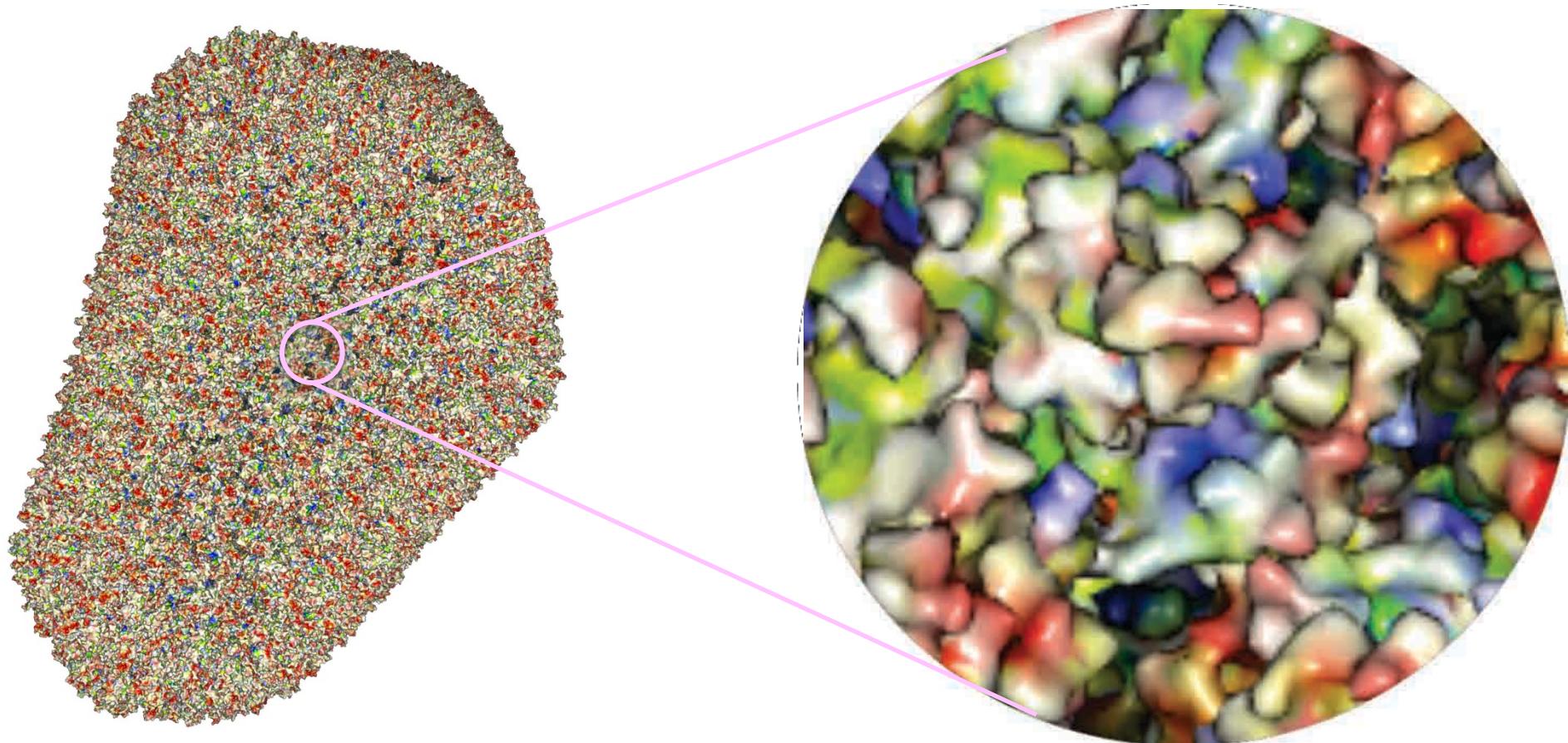
Molmil: Molecular Viewer based on WebGL



- 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)

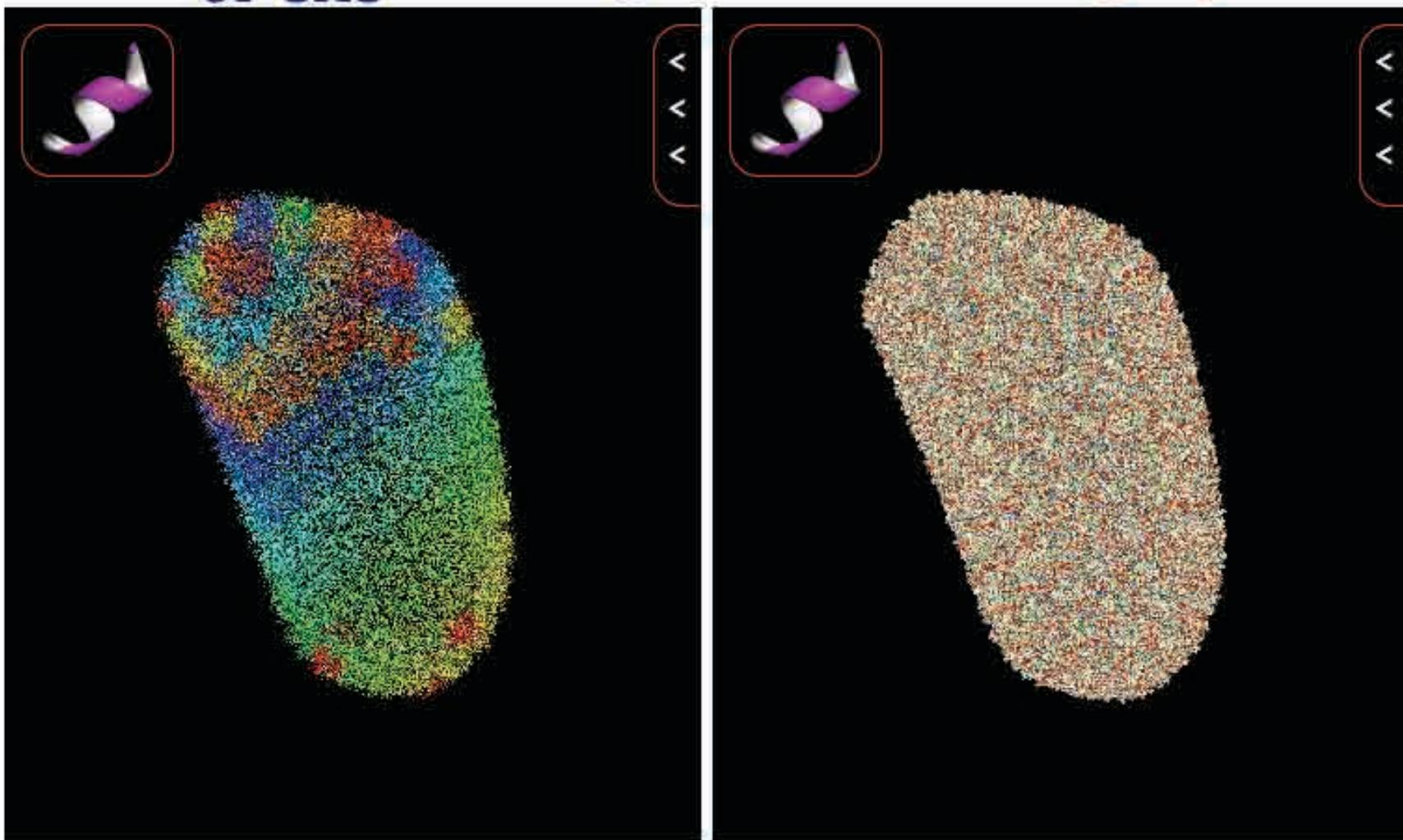
Molmil Viewer for eF-site

even for large structures



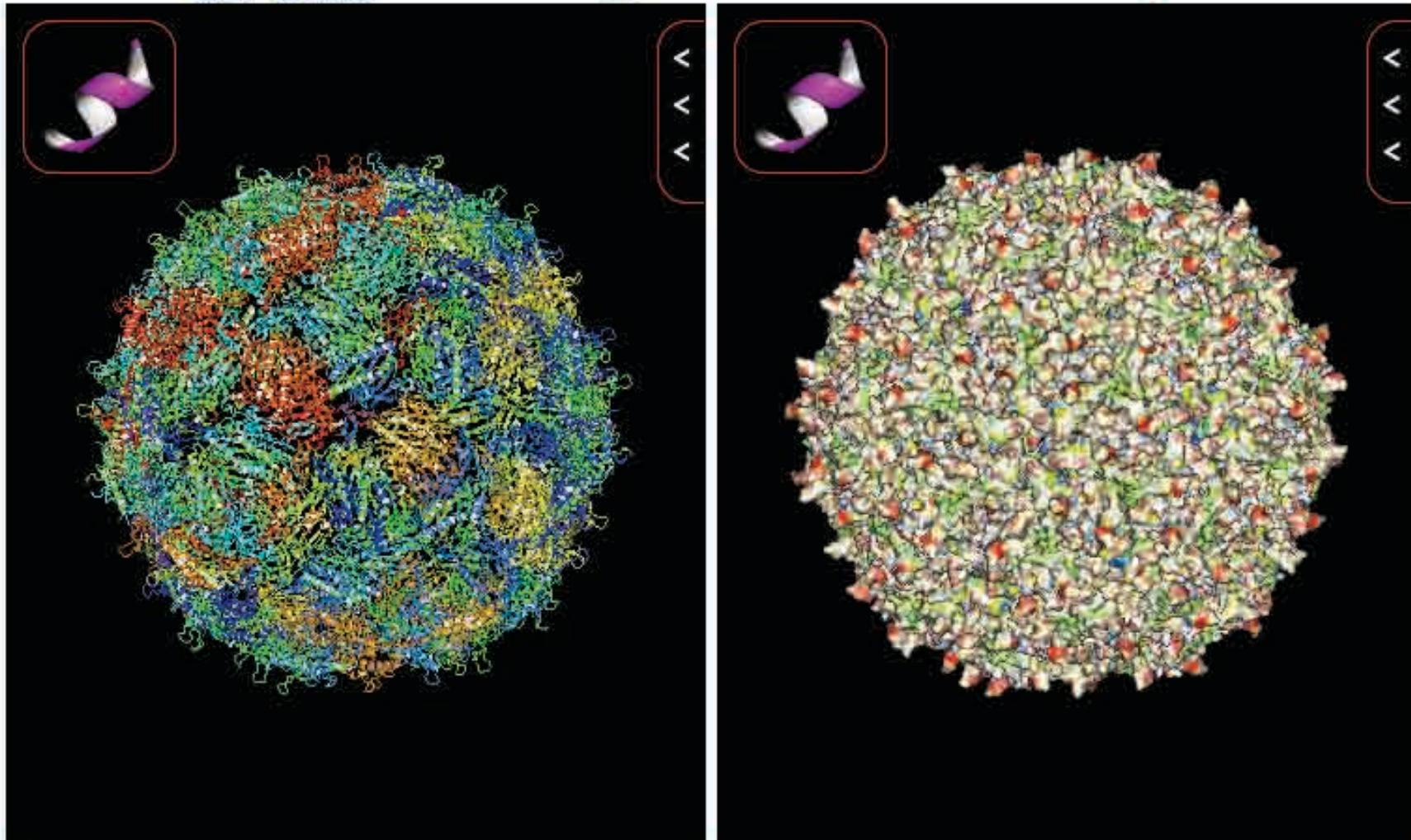
HIV-1 Capsid 3J3Q, 1356 chains, >2M atoms

eF-site for large structures 3j3q

 Cartoon CPK Cartoon without hetero atoms[Help](#)

ATOMIC-LEVEL STRUCTURE OF THE ENTIRE HIV-1 CAPSID

eF-site for large structures 4bp7



Cartoon

CPK

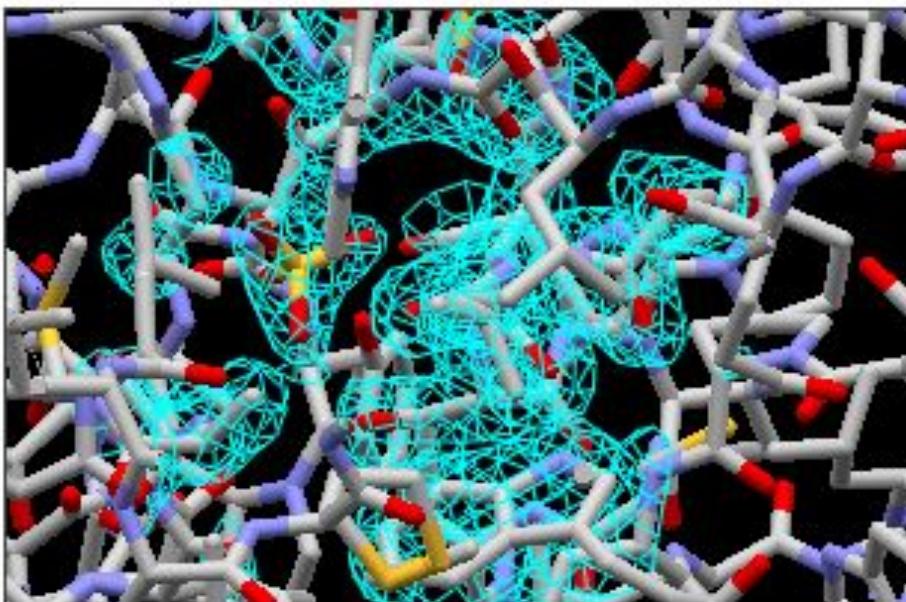
Cartoon without hetero atoms

[Help](#)

ASYMMETRIC STRUCTURE OF A VIRUS-RECEPTOR COMPLEX

ED map (電子密度マップ)

EDmap: 2xvy



jV>

2149 Atoms Selected.

Style:

Wireframe

Color:

Atom



電子密度マップのパラメータ

電子密度マップについて

マップのタイプ:

緩密級の網 密度平面

マップの座標:

マップの中心に一番近い原子

原子ID: , ,

座標 (x: y: z:)

マップ半径: 10 Å

緩密級レベル: 1 σ

色: R: 0.0 G: 1.0 B: 1.0

密度平面の透過レベル: 0.5

電子密度マップ作成

リセット

Electron Density Map Download/Delete

file format

filename

structure factor

r2xvysf.ent.gz

Download

refinement file

2xvy_ref.tar.gz

Download

2013108172932_2xvy.c.xml.gz

Delete

Download

PDBの新たなフォーマット: PDBx/mmCIF

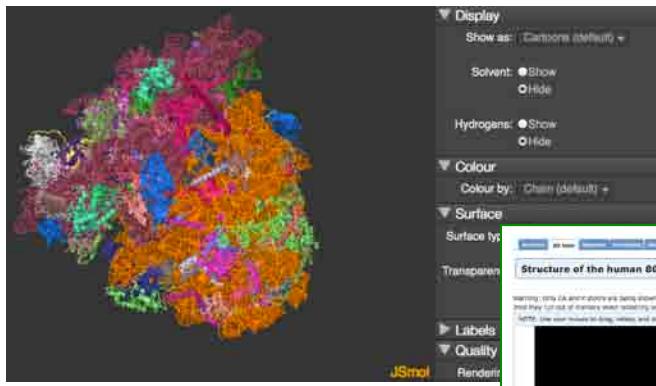
- 問題点: PDBフォーマットは **40 年前**に開発されたものであり、現代科学の要求には適合できなくなっている
- 記述における現実的な限界
 - 最大 **62** の高分子鎖
 - 最大 **99,999** の原子数
 - 低分子リガンドにおける結合次数や光学活性を特定できない
 - NMR, EM(電子顕微鏡), それらの融合法の記載が困難
 - メタデータの記載が厄介で柔軟性がない



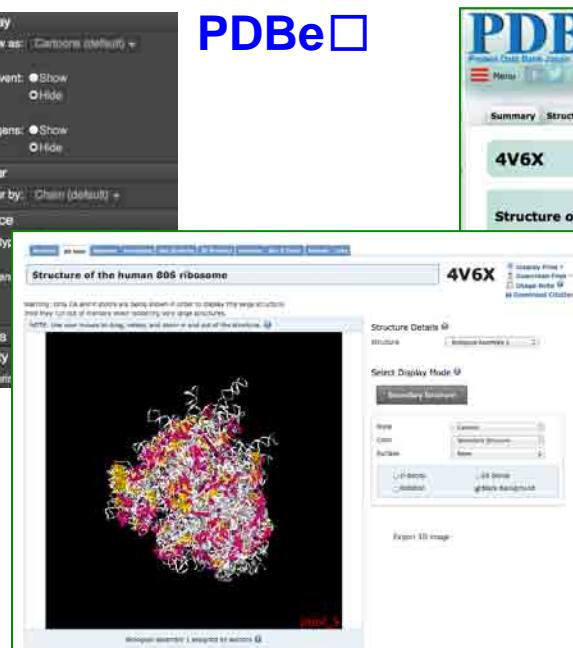
- **backward compatibility** をできるだけ保持
- これまでの **PDB** フォーマット dataを作れるWeb site
- 2013年中にスタートし、2016年から徐々にfade-out

Large Structures in PDB

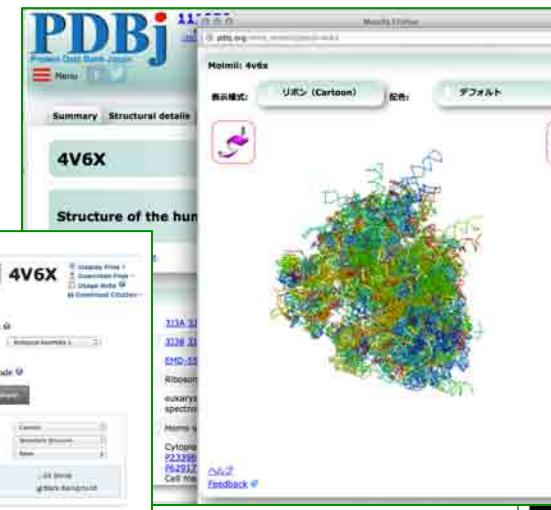
- As of December 2014, large structures are no longer split into multiple PDB entries but are **released as single PDBx and PDBML files**
 - Existing split entries were consolidated into single entries
 - A bundle of best-effort PDB files (PDB bundle) is available for every large structure through the wwPDB FTP sites
- Each wwPDB member provides **tools to view large structures**



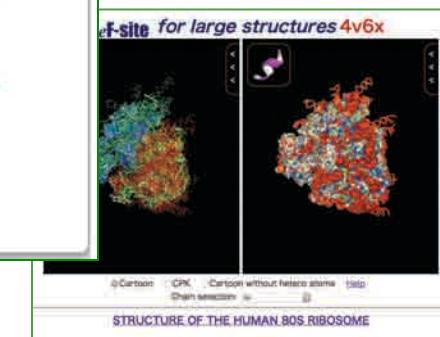
PDBe□



4v6x: structure of human 80S ribosome, 89 chains, 4 MDa□



PDBj□



RCSB-PDB□

```

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 is a navigation bar with links for "Home", "Dictionaries", "Documentation", "Downloads", and "Contact Us". On the right side of the navigation bar is the "wwPDB" logo. Below the navigation bar, there is a search bar with the placeholder text "Search current dictionary" and a magnifying glass icon. The main content area features a large title "PDBx/mmCIF Dictionary Resources" in a serif font. Below the title, a paragraph of text explains the purpose of the site: "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 the main content area is a green button with the text "Browse the current dictionary »".

PDBx/mmCIF Dictionary Resources

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.

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.

ファイルが選択されていません。

2. Confirm the contents of operations

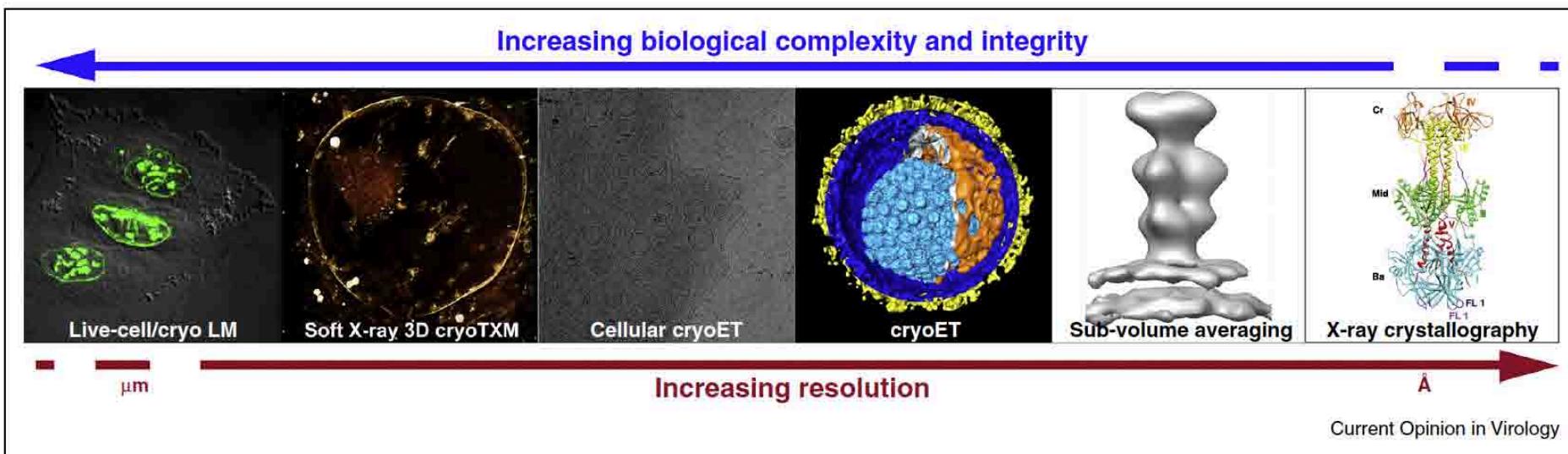
3. Execute conversion & Download the converted file

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.

Towards Structural Life Science

A new trend to reveal structures and functions of cellular machineries by the hybrid approach



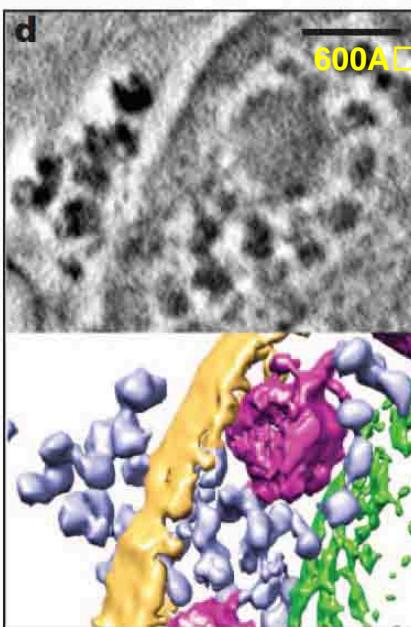
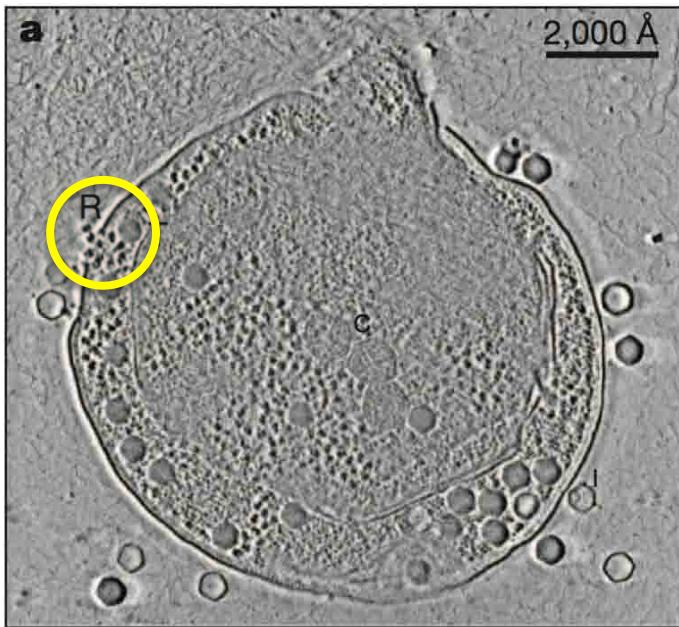
The spectrum of techniques applied to study the herpesvirus ‘life’ cycle. An integrated approach combining high-resolution structure determination methods and correlative light, soft X-ray cryo and electron cryo microscopy allows looking at dynamic processes at different resolution and complexity and ultimately leads to a better perception of those processes. LM, light microscopy; cryoTXM, transmission X-ray cryo microscopy; cryoET, electron cryo tomography.

Zeev-Ben-Mordehai et al. (2014)
***Current Opinion in Virology*, 5, 42-49**

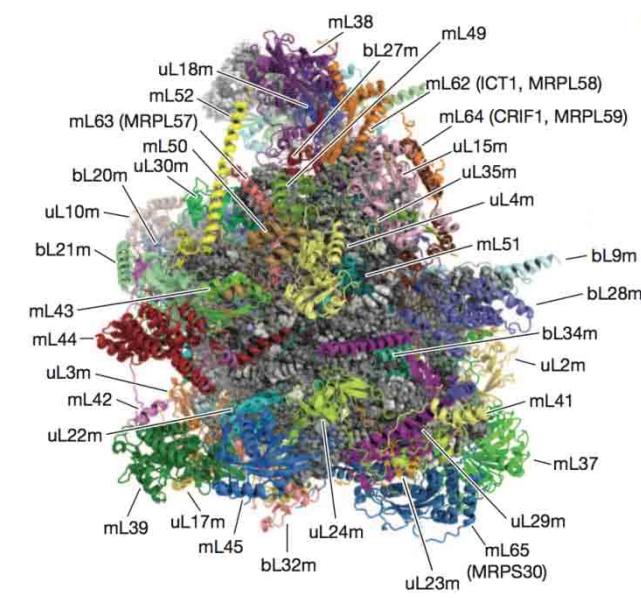
Towards Structural Life Science □

A new trend to reveal structures and functions of cellular machineries by the hybrid approach

by high-resolution cryoET □



by hybrid approach □

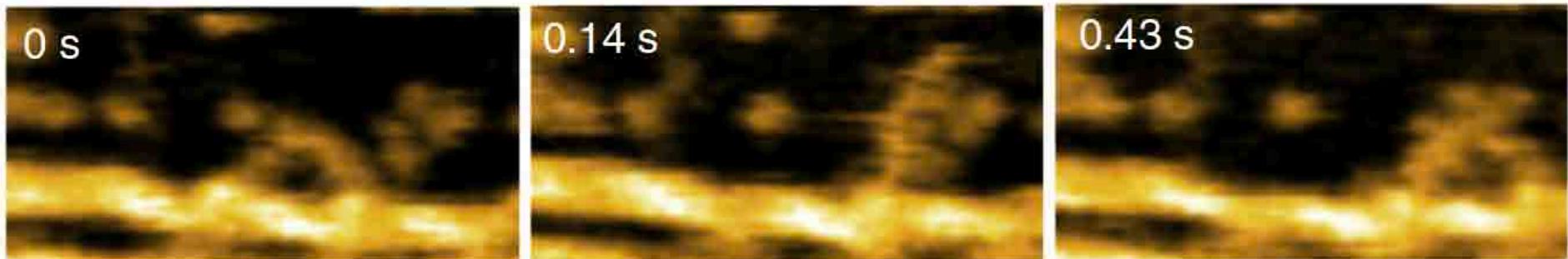


Dai et al. (2013) *Nature* 502, 707- 710 □

Greber et al. (2014) *Nature* 515, 283-286 □

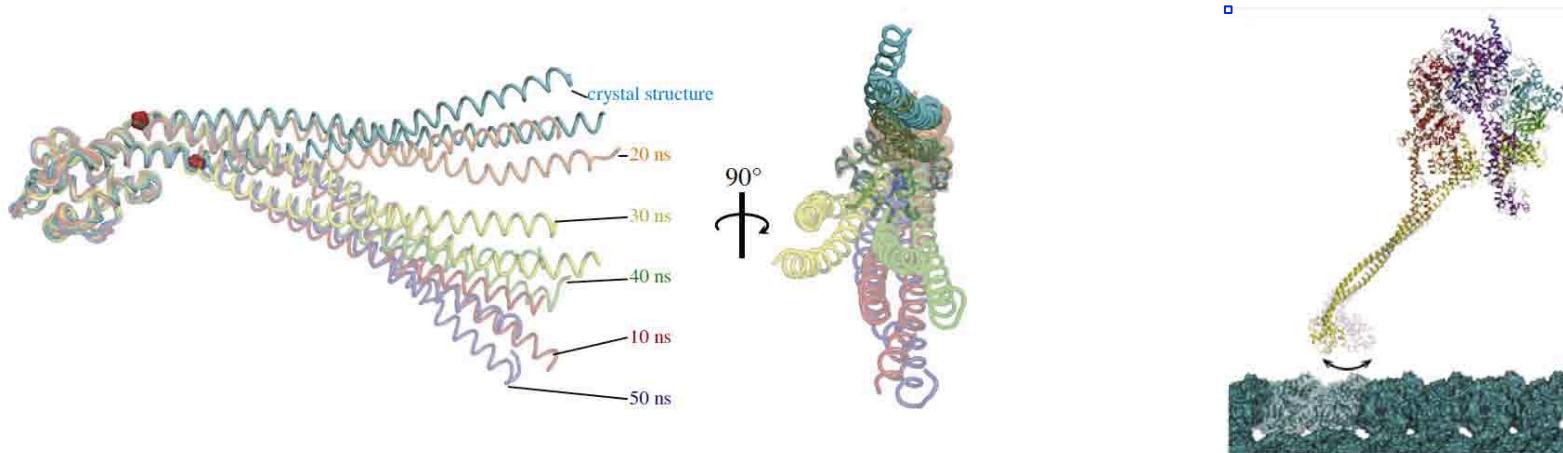
Dynamics of the machinery

“Walking” motion of Myosin V head along Actin by **high-speed AFM**



Endo T (2014) Curr. Opin Str Biol 28, 63-68

Swing Motion of the Stalk of Dynein by **X-ray Crystallography & MD**



Nishikawa et al. (2014) J. Mol. Biol. 426, 3232-3245

Task force of wwPDB for the hybrid approach was held in EBI on 6-7 October, where Iwasaki and Nakamura attended.

Data bank struggles as protein imaging ups its game

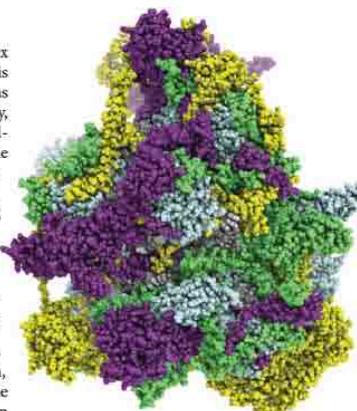
Hybrid methods to solve structures of molecular machines create a storage headache.

BY EWEN CALLAWAY

Structural biology, the mapping of complex biological molecules such as proteins, is in the grip of a revolution. The field has long been dominated by X-ray crystallography, a technique made iconic by its role in decoding the DNA double helix in the 1950s. But the need to tackle more complex structures and to watch ‘molecular machines’ function in real time is fuelling a shift towards hybrid imaging methods that can create moving models.

That is posing a challenge for the world’s official repository for protein structures: the Protein Data Bank (PDB), which relies almost exclusively on crystallography data and lacks the standards and software infrastructure to archive structures described by hybrid methods. This month, leaders of the four organizations around the world that host the data bank held a workshop in Hinxton, UK, to hatch a plan to ensure that hybrid models and their insights into fundamental biology and disease do not get lost.

Historically, structural biology has focused on generating three-dimensional (3D) descriptions of individual proteins. In many cases, this is a task perfect for crystallography, in which



A subunit of a ribosome, a molecular machine.

crystallography. “These days, being a crystallographer is not good enough,” says Gerard Kleywegt, a structural biologist at the European Bioinformatics Institute in Hinxton, who heads the European annex of the PDB.

that alter the proteasome’s activity. This year, another team published a hybrid model of the key HIV proteins that sneak the virus into a cell, which may help in vaccine design (M. Pancera *et al. Nature* <http://doi.org/wtz>; 2014).

The hybrid approach has also tackled the ribosome, which produces proteins; the nuclear pore complex, which provides a gateway between the genome in the nucleus and the rest of the cell; and the molecular syringes made by bacteria that inject proteins into cells. Models of many more molecular machines are expected. “We’re going to enter a period of exponential growth in the generation of these hybrid structures,” says Stephen Burley, a structural biologist at Rutgers University in Piscataway, New Jersey, who heads one of the two US annexes of the PDB.

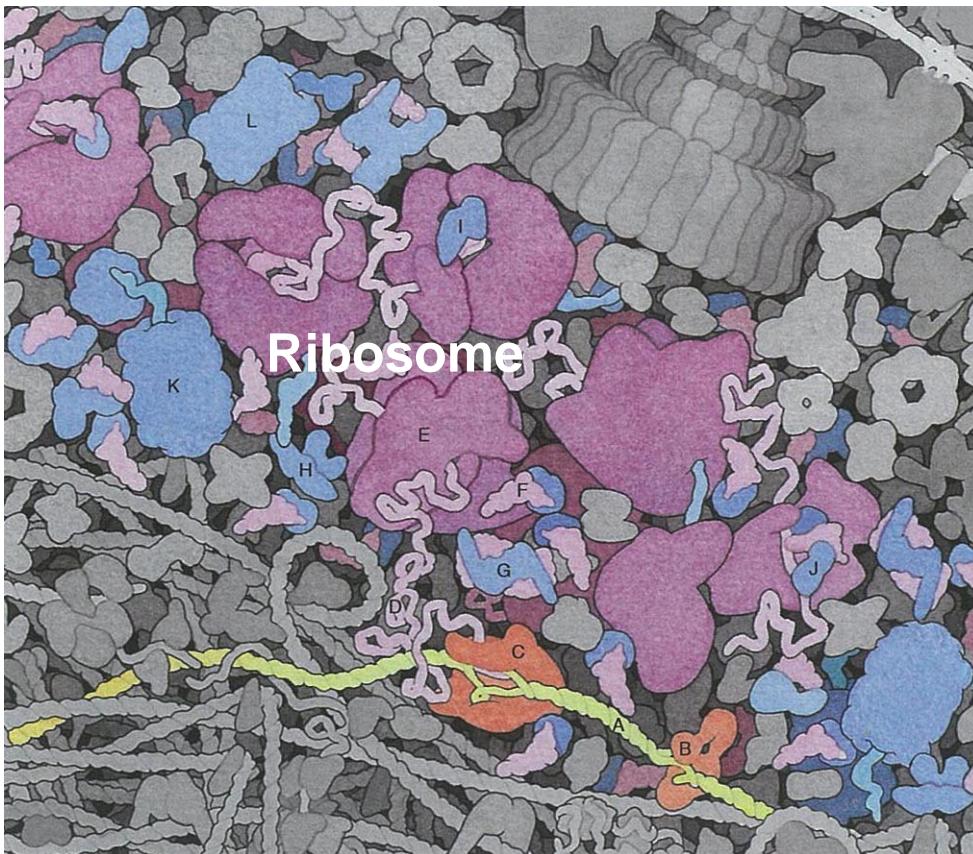
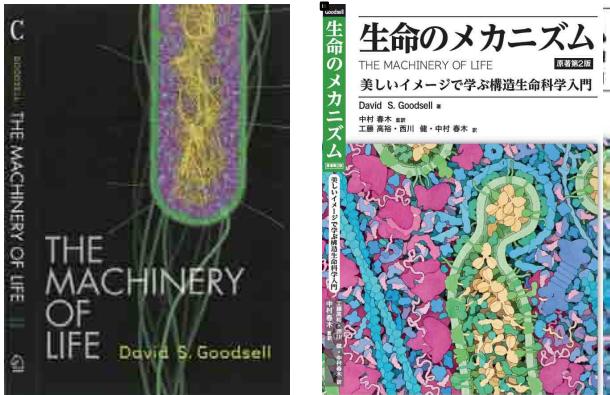
At the PDB workshop, on 6–7 October, Kleywegt, Burley and three dozen others hashed out the challenges that these triumphs are creating for the PDB. Crystallography yields a standardized set of data files in which a structure and its level of precision are self-evident; by contrast, the underlying data for the hybrid models exist in a mishmash of formats such as X-ray diffraction patterns or electron-micrograph pictures. And going from raw data to a



Task force of wwPDB at EBI on 6-7 October 2014

Nature (2014) 514, pp.416

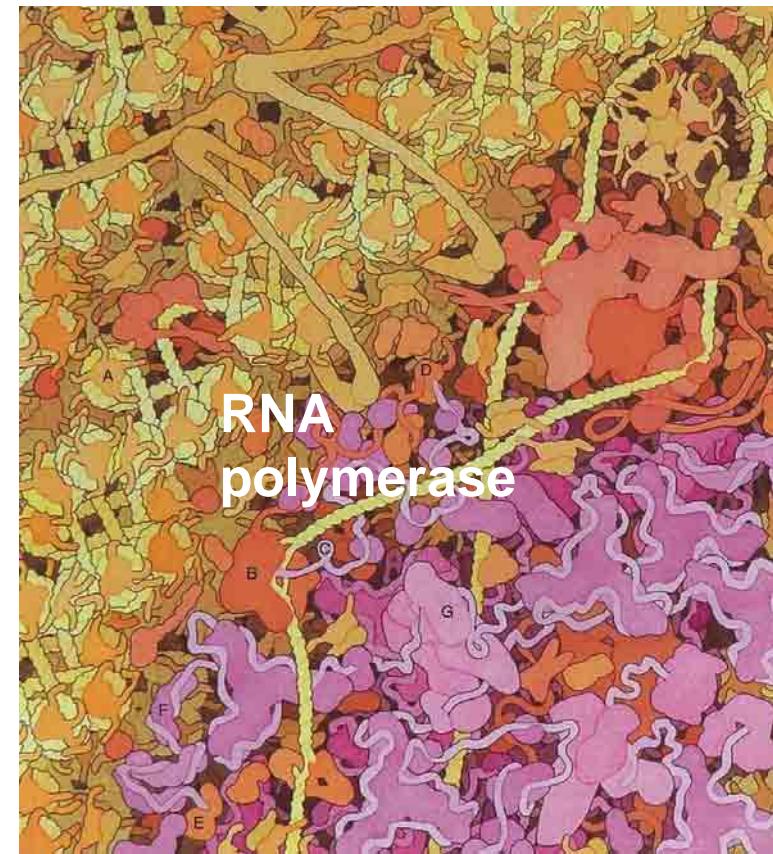
“Ultimately, we want to have the molecular structure of an entire cell. That’s still science fiction at the moment – but it’s some-where we can get to in 10, 20 years” (by Jan Ellenberg, EMBL)



Protein Synthesis in the Bacterial Cell

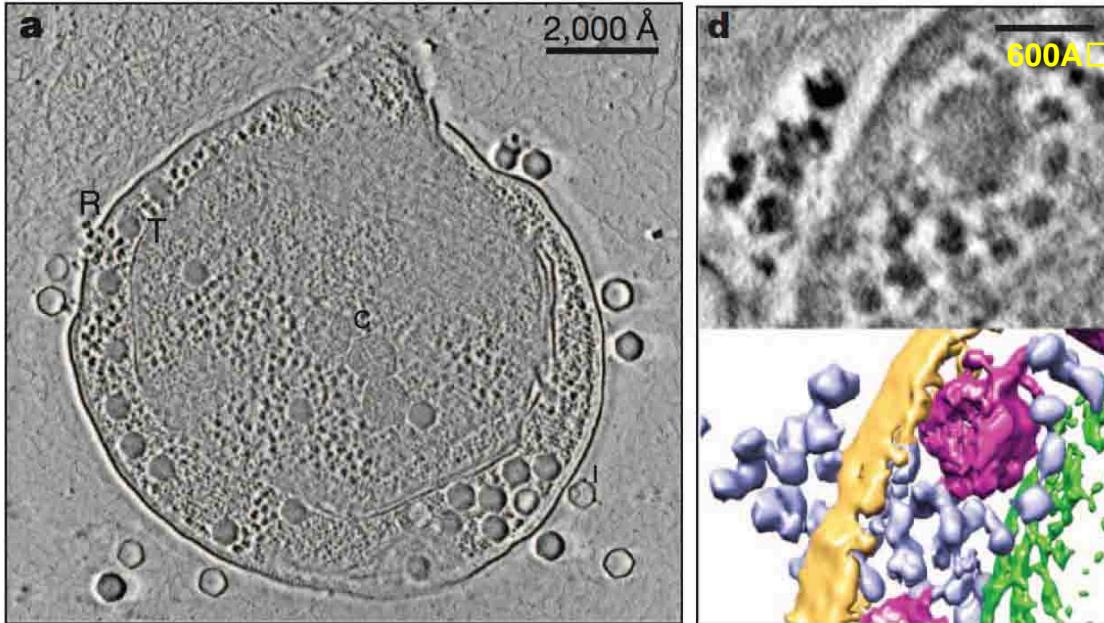
THE MACHINERY OF LIFE

(Second Edition, 2009)
by David S. Goodsell



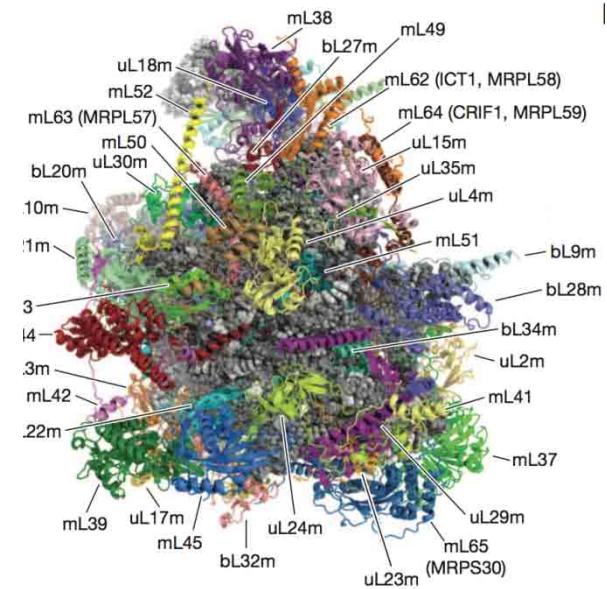
Nucleus

high-resolution cryoET □



Dai et al. (2013) *Nature* 502, 707- 710 □

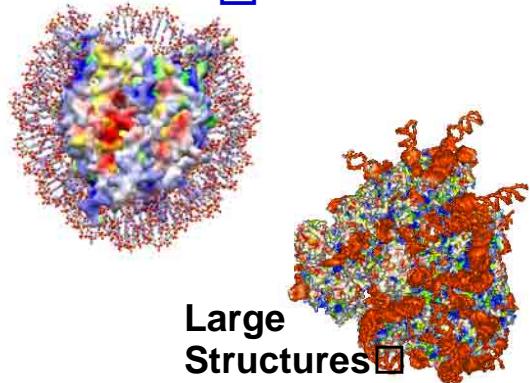
hybrid approach □



Greber et al. (2014) *Nature* 515, 283-286□

Databases and Portals□

PDB □



EMDB

PMP (protein model portal) □

SASBDB □

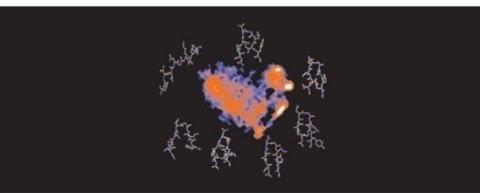
Themes in Integrative Structural Biology with Hybrid Methods

- Various methods and Integrations
- Modeling and Fitting
- Dynamics and Structural Changes
- Applications
- Data Validation
- Archives and Databases

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

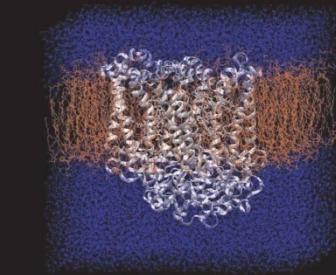
タンパク質計算科学



タンパク質計算科学

基礎と創薬への応用

神谷成敏・肥後順一・福西快文・中村春木 著



共立出版



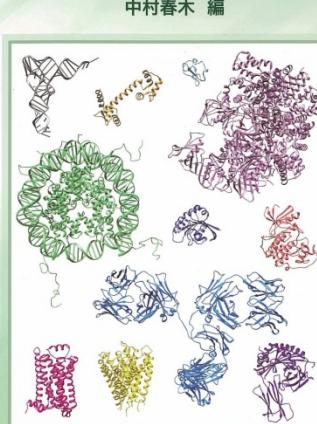
見てわかる 構造生命科学

▼ 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

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 ◀



化学同人

中村 編

工藤・西川・中村 訳

生命のメカニズム
Goodsell

生命のメカニズム

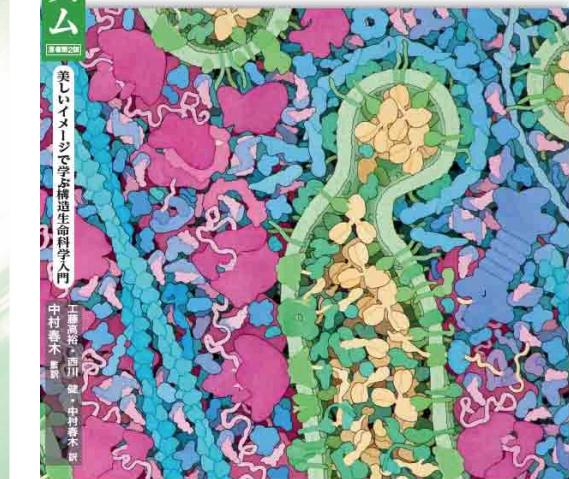
THE MACHINERY OF LIFE

原著第2版

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



(共立出版2009年8月初版)

(化学同人2014年4月初版)

(シナジー2015年2月初版)

PDBj Staffs

- 統括責任者

- [中村 春木](#) (大阪大学蛋白質研究所・教授)

H. Nakamura

- 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)

A. Nakagawa
R. Igarashi
Y. Kengaku
H. Cho
Y. Ikegawa
J. Sato

- 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 煉治 (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)

A. R. Kinjo
K. Iwasaki
H. Suzuki
R. Yamashita
T. Kudou
B. Gert-Jan

- BMRBデータベース管理運営グループ

- [藤原 敏道](#) (大阪大学蛋白質研究所・教授)
- 阿久津 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)

T. Fujiwara
H. Akutsu
C. Kojima
N. Kobayashi
T. Iwata
M. Yokochi

- 研究協力者

- [輪湖 博](#) (早稲田大学社会科学総合学術院・教授) for ProMode
- [猿渡 茂](#) (北里大学理学部・准教授) for ProMode
- 伊藤 暢聰 (東京医科歯科大学大学院・教授)
- [木下 賢吾](#) (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

H. Wako
S. Endo
N. Ito
K. Kinoshita
K. Katoh

- 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)

N. Haruki