

PDBjとwwPDBの活動方針について

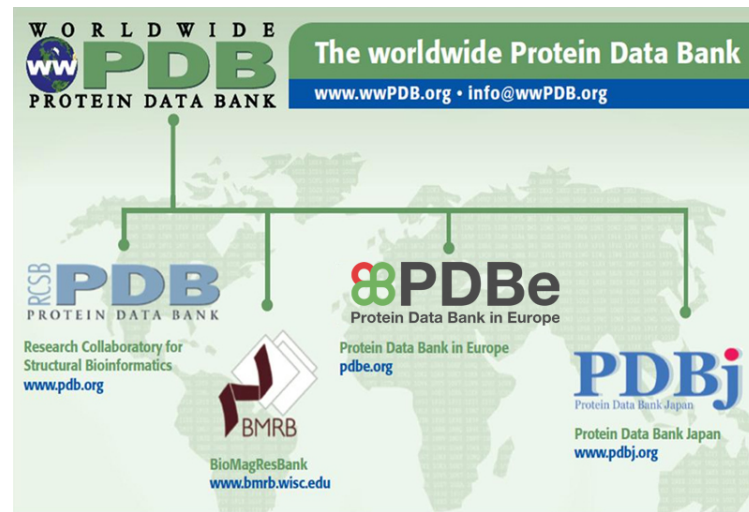
大阪大学蛋白質研究所
蛋白質データベース開発研究室(PDBj)
栗栖源嗣



Protein Data Bank Japan

<http://pdbj.org/>

Since 2001, PDBj has been managed at **Institute for Protein Research, Osaka University** as a member of the **wwPDB**, to curate and process the deposited data for an open and single archive.



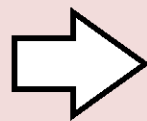
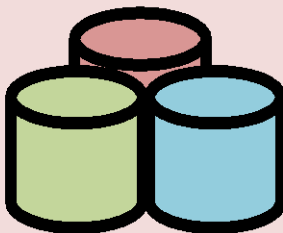
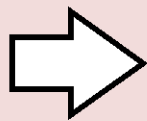
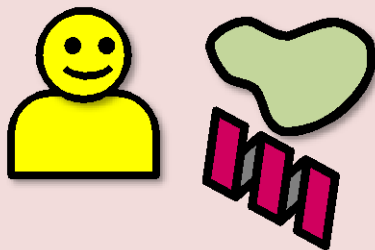
wwPDB collaboration



Deposition

PDB, EMDB,
& BMRB

Archive
distribution



Search



Visualization



Atomic coordinates

```

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.Cartn_x_esd
_atom_site.Cartn_y_esd
_atom_site.Cartn_z_esd
_atom_site.occupancy_esd
_atom_site.B_iso_or_equiv_esd
_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 . ALA A 1 1 ? 38.840 0.236 1.012 1.00 34.65 ? ? ? ? ? ? 1 ALA A N 1
ATOM 2 C CA . ALA A 1 1 ? 38.356 -0.999 0.357 1.00 42.26 ? ? ? ? ? ? 1 ALA A CA 1
ATOM 3 C C . ALA A 1 1 ? 37.098 -1.547 1.056 1.00 41.25 ? ? ? ? ? ? 1 ALA A C 1
ATOM 4 O O . ALA A 1 1 ? 36.619 -0.946 2.028 1.00 29.44 ? ? ? ? ? ? 1 ALA A O 1
ATOM 5 C CB . ALA A 1 1 ? 39.398 -2.114 0.379 1.00 40.70 ? ? ? ? ? ? 1 ALA A CB 1
ATOM 6 N N . SER A 1 2 ? 36.610 -2.666 0.495 1.00 32.67 ? ? ? ? ? ? 2 SER A N 1
ATOM 7 C CA . SER A 1 2 ? 35.411 -3.244 1.202 1.00 34.90 ? ? ? ? ? ? 2 SER A CA 1
ATOM 8 C C . SER A 1 2 ? 35.683 -4.740 1.081 1.00 38.30 ? ? ? ? ? ? 2 SER A C 1
ATOM 9 O O . SER A 1 2 ? 36.827 -5.147 0.747 1.00 28.59 ? ? ? ? ? ? 2 SER A O 1
ATOM 10 C CB . SER A 1 2 ? 34.063 -2.660 0.823 1.00 24.49 ? ? ? ? ? ? 2 SER A CB 1
ATOM 11 O OG . SER A 1 2 ? 33.031 -3.308 1.686 1.00 20.37 ? ? ? ? ? ? 2 SER A OG 1

```

Activities/Services of each member of the wwPDB

- **“Data-in”** activity, common in all the wwPDB members with high quality control. For that purpose, new format, data deposition, and validation system are developed
- **“Data-out”** services, common archive as the ftp site and the characteristic services by each wwPDB member

File Versioning: Planning Process

Current Issues:

- Loss of connection between PDB ID and Publication under current wwPDB Obsolete/Supersede Policy
- Enable revisions to entries updated by the Depositor of Record (e.g., Version 1-0 → 1-1; 1-0 → 2-0)
 - wwPDB will NOT assign a new PDB ID going forward (for Depositor of Record revision only)
- **Introduce new PDB ID code format**
 - With PDB prefix and extension of 4 characters (e.g., from “1ABC” to “**PDB_00001ABC**”)
- Example: PDB_00001ABC_XYZ_V2-2.cif.gz

DOI Landing Page Layout (Planned)

About

- Contact Us
- Publications
- Advisory Committee
- Agreement
- Outreach
- FAQ

Data Access Options for 10.2210/pdb1kip/pdb

This page presents data access options for PDB entry **10.2210/pdb1kip/pdb**.
Questions about this page may be sent to info@wwpdb.org.

Data download options:

- [Structure coordinates \(PDBx/mmCIF\)](#)
- [Structure coordinates \(PDBML\)](#)
- [Structure coordinates \(PDB\)](#)
- [Structure coordinates \(RDF\)](#)
- [X-ray diffraction data \(PDBx/mmCIF\)](#)

Further resources for entry 10.2210/pdb1kip/pdb at: [PDBe](#) [PDBj](#) [RCSB PDB](#)



News & Announcements

Members:



Download Archive

[RCSB PDB ftp](#) | [PDBe ftp](#) | [PDBj ftp](#)
[Instructions](#)

Archive Snapshots

[RCSB PDB](#) | [PDBj](#)

Cite wwPDB:

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

Worldwide Protein Data Bank (wwPDB)

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

RCSB PDB An Information Portal to 132905 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligand

Advanced Search | Browse by Annotations

EMDataBank EMBL-EBI PDB-DE

Welcome


Deposit Search Visualize Analyze Download Learn

A Structural View of Biology

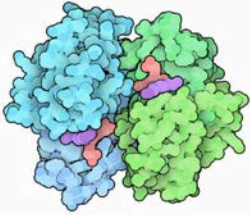
This resource is powered by the Protein Data Bank archive—information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Video: How Enzymes Work



August Molecule of the Month



Glutathione Transferases

EMBL-EBI Protein Data Bank in Europe

Bringing Structure to Biology

Search: hemoglobin, BRCA1_HUMAN

PDBe home Deposition PDBe services PDBe training Documentation About PDBe

PDBe is the European resource for the collection, organisation and dissemination of data on biological macromolecular structures. Read more about PDBe.

Featured structure

Turning cartwheels

1st August 2017

The image from our 2017 calendar for August is part of a molecular cartwheel essential for the mobility of single celled parasites called trypanosomes. Trypanosomes are the causative agents of several potentially fatal human diseases, including Leishmaniasis, Chagas disease and Sleeping Sickness.

Read more...

Previous featured structures

News

Enriched Model Files Conforming to OneDep Data Standards Now Available in the PDB FTP Archive

Events

24th Congress & General Assembly of the International Union of Crystallography

Popular: EMsearch, PDBFold, PDBePISA, Sequence search, PDBe REST API, EM resources, NMR resources, EMPiAR, Coordinate Server, PDB Component Library

Latest archive statistics

As of 16 August 2017 the PDB contains 132905 entries (latest PDB entries, chemistry, biology) and EMBL contains 5077 entries (latest map releases, latest header releases, latest updates).

Connect with us

PDBj 132905 entries (2017-08-16)

English 日本語 简体中文 繁體中文 한국어

Search: pdbj.org

Menu

日本蛋白質構造データバンク (PDBj: Protein Data Bank Japan), 由JST-NBDC 和大阪大学蛋白質研究所 提供支援, 与美国的 RCSB PDB, EMBL EBI 和欧洲 PDBe 共同合作, 在对国际性的生物大分子立体结构数据库的PDB档案文件进行维护的同时, 还提供各种结构解析的应用工具。

Guide for first time visitors

For an introduction to the new web interface, please read [Using PDBj's web interface](#). An introduction to the customization features offered by the new PDBj web interface can be found [here](#). To get a more in-depth explanation on the various features of the PDBj website, please take a look at the [Interactive tutorial series](#).

The [legacy PDBj website](#) will no longer be updated after July 12, 2017 and will be closed at the end of August, 2017.

Find the service you need

Choose a keyword listed below or input keywords into the textbox at the right of the keyword list. The brief explanation of the matched services will be displayed.

- Click the 'Show all services' button to display the explanation for all services.
- Input some keywords into the 'Word Search Box' to narrow down the search results.

PDB
 BMRB
 EMBL

search
 deposition
 viewer

NMR
 electron microscopy
 secondary structure

[WORLDWIDE PDB PROTEIN DATA BANK](#)
[wwPDB](#)
 wwPDB Events at JUCR (August 21-28)
[Better Management of PDB Archive with File Versioning and Revision History](#)
[Enriched Model Files Conforming to OneDep Data Standards Now Available in the PDB FTP Archive](#)
[5 Easy Steps to PDB Deposition](#)
[Revise Your Structure Without Changing the PDB Accession Code and Related Changes to the FTP Archive](#)

Biological Magnetic Resonance Data Bank

Member of WORLDWIDE PDB PROTEIN DATA BANK

A Repository for Data from NMR Spectroscopy on Proteins, Peptides, Nucleic Acids, and other Biomolecules

Instant entry access: Searches all entries on many criteria: Title, Author, Entity, Organism, Database code, etc. Hover over a result for more information.

Google Cl

Home

- About BMRB
- Search archive
- BMRB data by type
- Validation Tools
- Deposit Data
- NMR Statistics
- Spectroscopists' Corner
- Programmers' Corner
- Metabolomics
- Structural genomics
- Educational Outreach
- NMR Data Formats
- Links to External

Search macromolecule database

If you have a query you would like to run on the BMRB database, please e-mail bmrhelp@bmrw.wisc.edu

Field	Value to search for	Display
<input type="checkbox"/> Entry ID (entry or accession number)	<input type="text"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> PDB	<input type="text"/>	<input type="checkbox"/>
<input type="checkbox"/> Title	<input type="text"/>	<input type="checkbox"/>
<input type="checkbox"/> Author (family name)	<input type="text"/>	<input type="checkbox"/>
<input type="checkbox"/> Molecule name	<input type="text"/>	<input type="checkbox"/>

Output HTML CSV Inline

[Restrains Search](#) [Metabolomics Advanced Search](#) [Help](#)


Deposit data: ADIT-NMR data deposition system.

Please look at the data accepted before depositing.

CS-Rosetta server.

Submit your chemical shifts to run CS-Rosetta.

File Formats provided by PDBj

- mmCIF
 - The canonical format of the wwPDB.
 - Ver. 5 released this year!
 - PDBML
 - “direct translation” of mmCIF into XML.
 - (Legacy) PDB format
 - *NOT RECOMMENDED!*
 - PDB/RDF
 - Translation of PDBML into RDF/XML (the standard format for the Semantic Web).
- 
- Experimental Data (X-ray, NMR, 3DEM=EMDB)
 - **Validation Report (PDF, XML, RDF)**

Where is the detailed validation data?

▼ クイックリンク

- ヘルプ
- 巨大構造エントリー
- グループ登録エントリー
- 化合物一覧
- 最新エントリー

▼ 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
- Structure-Navigator
- EM Navigator
- Omokage検索
- wwPDB/RDF
- SeSAW
- Ligand Binding Sites (GIRAF)
- 未公開エントリーのステータス

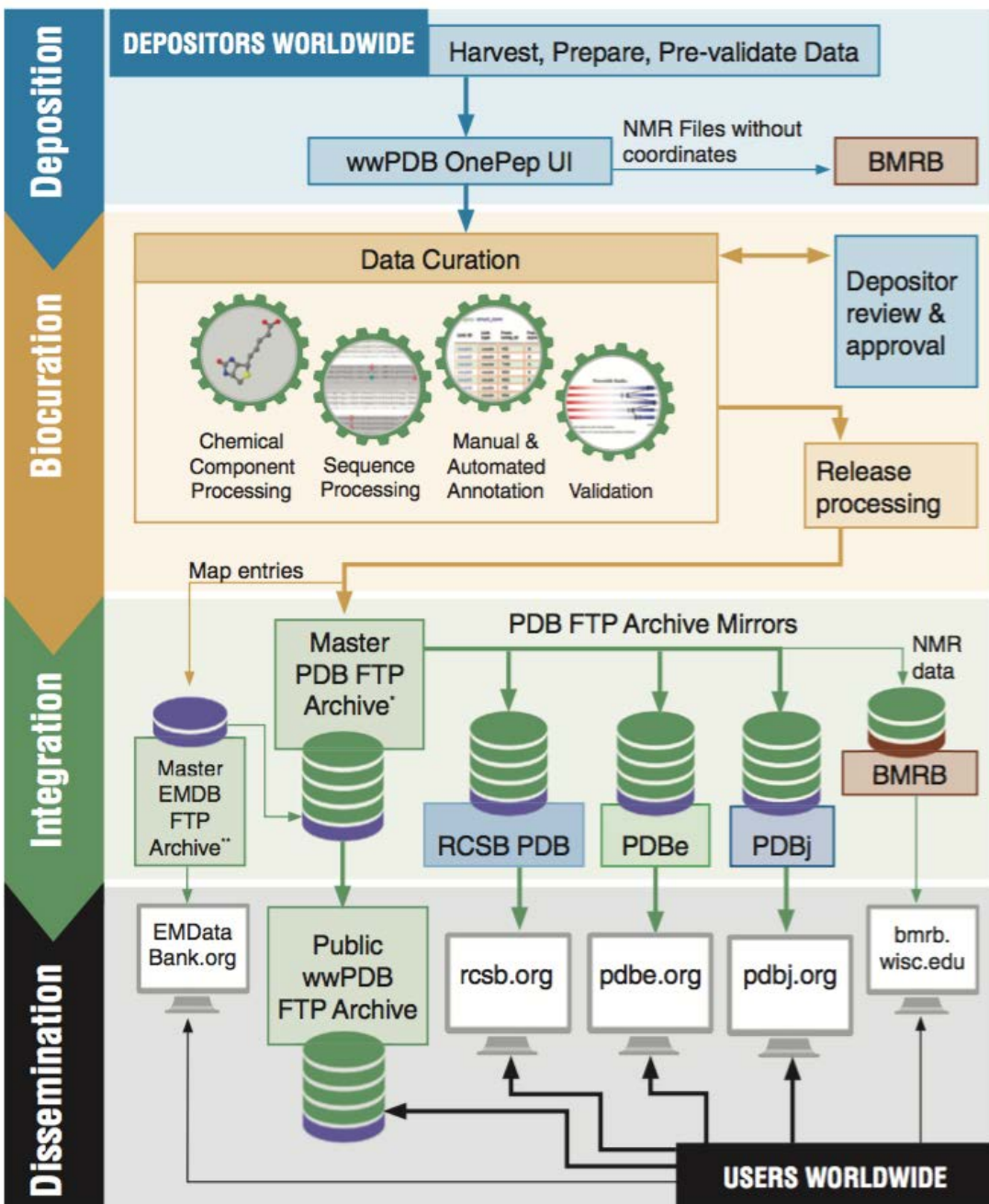
▶ 分子ビューア

▼ サービス&ソフトウェア

- ヘルプ
- 万見 (Yorodumi)
- ASH
- MAFFTash
- NMRToolBox
- gmfit
- CRNPRED
- Spanner
- SFAS
- HOMCOS

	PDBx/mmCIF	1cbs.cif.gz (36.71 KB)	画面表示
		全ての情報	画面表示
PDBML		ヘッダのみ	画面表示
		座標情報のみ	画面表示
		全ての情報	画面表示
PDBMLplus		ヘッダのみ	画面表示
		付加情報のみ	画面表示
		全ての情報	画面表示
	RDF	1cbs.rdf.gz (23.62 KB)	画面表示
	構造因子	r1cbsf.ent.gz (149.64 KB)	画面表示
	生物学的単位 (PDB形式)	1cbs.pdb1.gz (25.94 KB) (A) *author defined assembly, 1 molecule(s) (monomeric)	画面表示
	PDF	1cbs_validation.pdf.gz (411.76 KB)	画面表示
	PDF-full	1cbs_full_validation.pdf.gz (411.94 KB)	画面表示
	検証レポート	1cbs_validation.xml.gz (8.17 KB)	画面表示
	PNG	1cbs_multipercentile_validation.png.gz (140.86 KB)	画面表示
	SVG	1cbs_multipercentile_validation.svg.gz (904 B)	画面表示

- ▼
- FSSP
- SCOP
- VAST
- PISA
- UniProt
- PFam
- PF00061
- eF-site
- 1cbs-A
- 電子密度マップ (EDM) (molmil)
- wwPDB/RDF
- Promode Elastic



Overview of OneDep system



Full wwPDB X-ray Structure Validation Report ⓘ

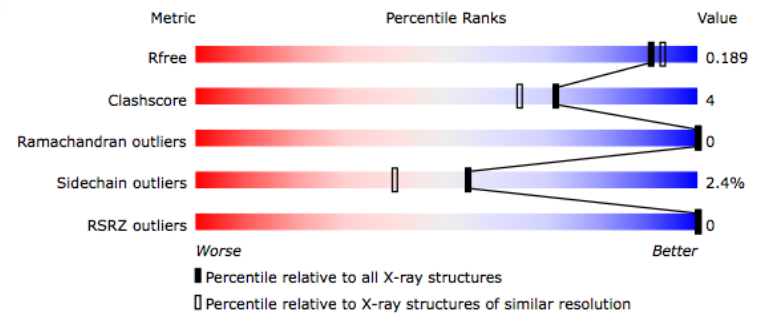
Feb 12, 2017 – 07:52 pm GMT

PDB ID : 1CBS
 Title : CRYSTAL STRUCTURE OF CELLULAR RETINOIC-ACID-BINDING PROTEINS I AND II IN COMPLEX WITH ALL-TRANS-RETINOIC ACID AND A SYNTHETIC RETINOID
 Authors : Kleywegt, G.J.; Bergfors, T.; Jones, T.A.
 Deposited on : 1994-09-28
 Resolution : 1.80 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org
 A user guide is available at <http://wwpdb.org/validation/2016/XrayValidationReportHelp> with specific help available everywhere you see the ⓘ symbol.

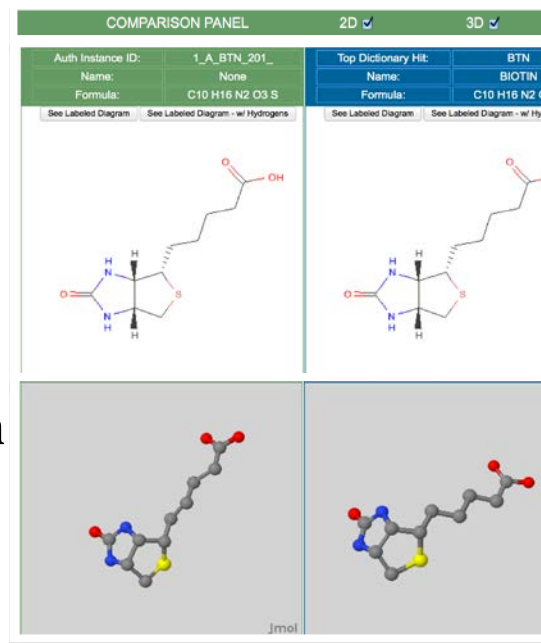
wwPDB validation report (PDF)



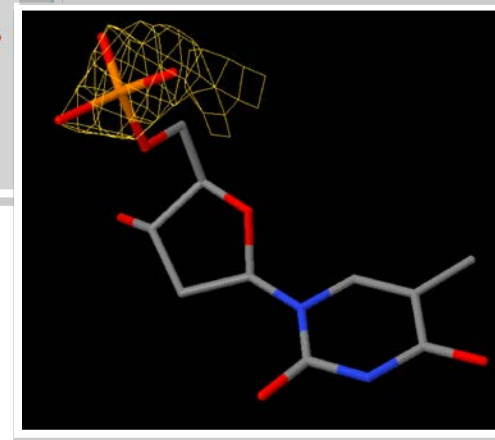
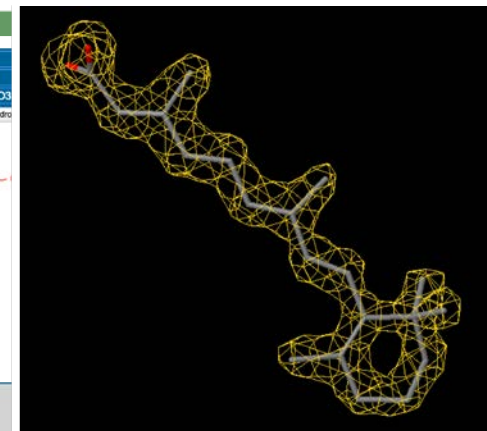
Percentile view of validation report

Improved Ligand Validation

- Batch search against Chemical Component Dictionary with automated CCD ID assignment
- Captures and displays author-provided chemical information
- Comparison panel
 - 2D and 3D views of ligand for review
 - ID assignment
- Display of local ligand electron density fit



Deposited instance from coordinates (left) and the closest match in the dictionary (right)



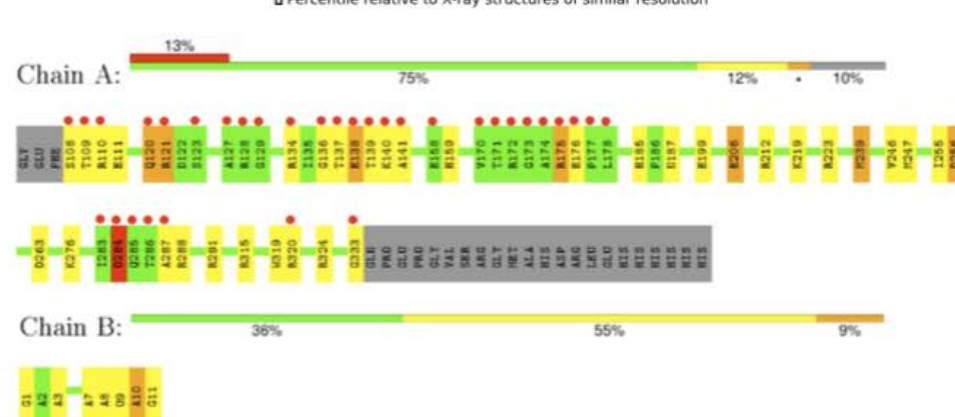
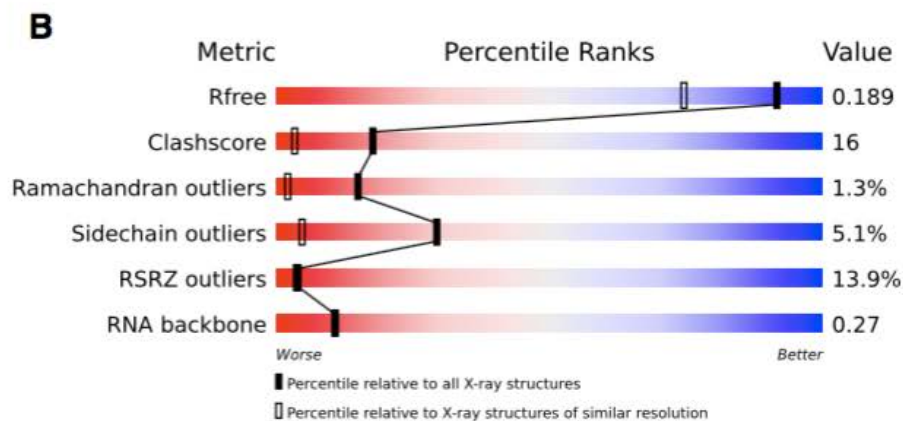
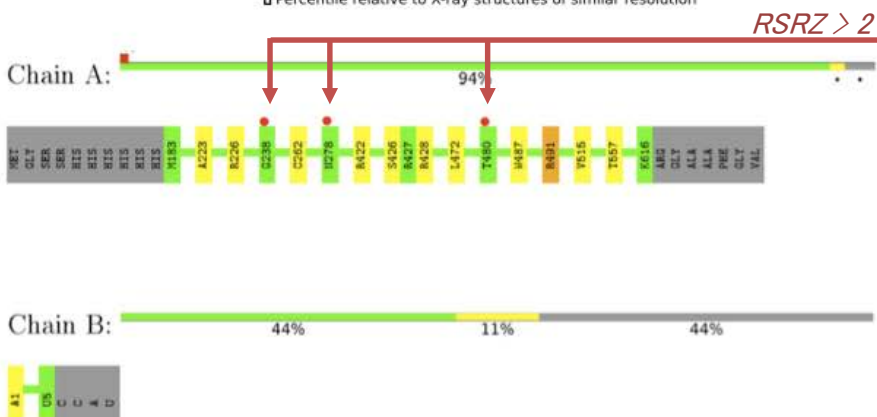
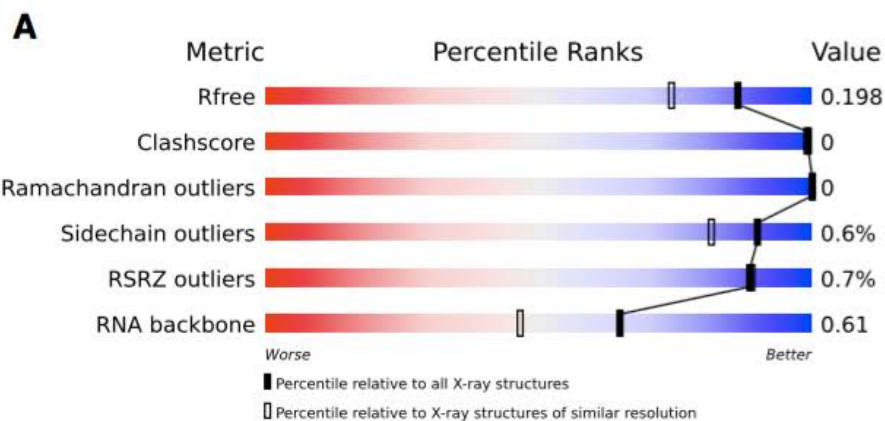
Local ligand density display (1.5 sigma omit map)

Top: REA in entry 1CBS with LLDF=1.31 (RSR=0.10, CC=0.95)

Bottom: TMP in entry 3HW4 with LLDF=6.77 (RSR=0.41, CC=0.70)

wwPDB validation report PDFs

Summary quality metrics in wwPDB validation reports



wwPDB validation PDFs are easily reviewed and shared an assessment of structure quality.

Validation metrics in wwPDB validation reports

X-ray/EM/NMR

- Geometric & conformational
 - bond, angle, planarity
 - protein backbone conformation
 - protein side-chain conformation
- Atomic & molecular interaction
 - all-atom contacts
 - under packing
 - hydrogen bond quality
- Non-protein
 - nucleic acids (RNA pucker, suite)
 - carbohydrates (N-glycan core)
 - ligands (CSD)
 - ions & other solvent
- Incomplete model (e.g. CA_ONLY)

Caveat:

LLDF (Local Ligand Density Fit) has been replaced by a combination of RSR (Real-space R factor) > 0.4 and RSCC (Real-space correlations coefficient) < 0.8 since this March.

X-ray

- Structure factor & electron density
 - Wilson plot, outliers, tNCS
 - wrong symmetry
 - twinning
 - agreement (R_{free} , RSR, RSCC)

NMR

- Chemical shifts
 - completeness
 - outliers
 - estimated reference error
 - random coil index
- Structure ensembles
 - representative model (medoid)
 - domain detection

wwPDB validation report PDFs

• Standard geometry

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.47	0/1107	0.71	0/1491

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

• Too close contacts

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1091	0	1106	7	0
2	A	22	0	27	2	0
3	A	100	0	0	2	0
All	All	1213	0	1133	9	0

• Protein backbones

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	135/137 (98%)	132 (98%)	3 (2%)	0	100 100

• Protein sidechains

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	123/123 (100%)	120 (98%)	3 (2%)	52 38

• Ligand geometry

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	REA	A	200	-	19,22,22	1.05	1 (5%)	26,30,30	1.02	2 (7%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	200	REA	C1-C6	2.25	1.56	1.53

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	200	REA	C11-C10-C9	-2.40	123.89	127.31
2	A	200	REA	C18-C5-C6	2.08	126.83	124.51

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	200	REA	2	0

Detailed wwPDB validation reports (XML)

```
<?xml version="1.0" ?>
<wwPDB-validation-information xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="http://wwpdb.org/validation/schema/wwpdb_validation_v002.xsd">
  <Entry B_factor_type="FULL" CCP4version="7.0 (Gargrove)" DCC_R="0.18" DCC_Rfree="0.19" DCC_refinement_program="CNS" DataAnisotropy="0.434"
    DataCompleteness="90.54" EDS_R="0.18" EDS_resolution="1.80" EDS_resolution_low="14.93" Fo_Fc_correlation="0.956" IoverSigma="3.77(1.79A)" PDB-R="0.20"
    PDB-Rfree="0.24" PDB-deposition-date="1994-09-28" PDB-resolution="1.80" PDB-resolution-low="8.00" PDB-revision-date="2011-07-13" PDB-revision-number="3"
    RefmacVersion="5.8.0158" RestypesNotCheckedForBondAngleGeometry="REA" TransNCS="The largest off-origin peak in the Patterson function is 9.26% of the
    height of the origin peak. No significant pseudotranslation is detected." TwinFraction="k,h,-l:0.027" TwinL="0.515" TwinL2="0.357"
    WilsonBano="16.802,17.606,11.032,0.000,0.000,0.000" WilsonBestimate="14.785" XMLcreationDate="Mar 10, 2018 -- 04:41 pm GMT" absolute-percentile-
    DCC_Rfree="90.4" absolute-percentile-clashscore="69.2" absolute-percentile-percent-RSRZ-outliers="100.0" absolute-percentile-percent-rama-
    outliers="100.0" absolute-percentile-percent-rotam-outliers="51.8" acentric_outliers="1" angles_rmsz="0.71" attemptedValidationSteps="mogul,molprobit,
    validation-pack,xtriage,eds,percentiles.writexml" babinet b="141.456" babinet k="0.156" bonds_rmsz="0.47" bulk solvent b="72.956" bulk_solvent_k="0.401"
    centric_outliers="0" clashscore
    relative-percentile-percent-RSR
    outliers="1.8" low-resol-relati
    num-free-reflections="1496" num
    numPDBids-absolute-percentile-percent-RSRZ-outliers="108989" numPDBids-absolute-percentile-percent-rama-outliers="120053" numPDBids-absolute-percentile-
    percent-rotam-outliers="120020" numPDBids-relative-percentile-DCC_Rfree="5253" numPDBids-relative-percentile-clashscore="6077" numPDBids-relative-
    percentile-percent-RSRZ-outliers="5157" numPDBids-relative-percentile-percent-rama-outliers="6011" numPDBids-relative-percentile-percent-rotam-
    outliers="6010" num_angles_rmsz="1491" num_bonds_rmsz="1107" pdbid="1CBS" percent-RSRZ-outliers="0.00" percent-free-reflections="10.19" percent-rama-
    outliers="0.00" percent-rotam-outliers="2.44" percentilebins="all,1.8,xray" protein-DNA-RNA-entities="1" relative-percentile-DCC_Rfree="92.5" relative-
    percentile-clashscore="62.5" relative-percentile-percent-RSRZ-outliers="100.0" relative-percentile-percent-rama-outliers="100.0" relative-percentile-
    percent-rotam-outliers="38.4" xtriage_input_columns="F_meas_au,F_meas_sigma_au"/>
  <ModelledSubgroup NatomsEDS="7" altcode=" " avgoccu="1.000" chain="A" ent="1" icode=" " model="1" num-H-reduce="9" owab="28.790" resname="PRO" resnum="1"
    rota="Cg_endo" rsccl="0.908" rsr="0.143" rsrz="0.706" said="A" seq="1">
    <clash atom="HB2" cid="4" clashmag="0.47" dist="1.97"/>
  </ModelledSubgroup>
  <ModelledSubgroup NatomsEDS="8" altcode=" " avgoccu="1.000" chain="A" ent="1" icode=" " model="1" num-H-reduce="6" owab="21.720" phi="-124.5" psi="100.5"
    rama="Favored" resname="ASN" resnum="2" rota="t30" rsccl="0.906" rsr="0.144" rsrz="0.485" said="A" seq="2"/>
  <ModelledSubgroup NatomsEDS="11" altcode=" " avgoccu="1.000" chain="A" ent="1" icode=" " model="1" num-H-reduce="9" owab="11.800" phi="-80.0" psi="-9.9"
    rama="Favored" resname="PHE" resnum="3" rota="m-85" rsccl="0.966" rsr="0.084" rsrz="-0.490" said="A" seq="3"/>
  <ModelledSubgroup NatomsEDS="6" altcode=" " avgoccu="1.000" chain="A" ent="1" icode=" " model="1" num-H-reduce="5" owab="12.240" phi="-55.6" psi="140.2"
    rama="Favored" resname="SER" re
  <ModelledSubgroup NatomsEDS="4" ali
    rama="Favored" resname="GLY" re
  </ModelledSubgroup>
  <ModelledEntityInstance absolute_RSRZ_percentile="100.00" absolute_rama_percentile="100.00" absolute_sidechain_percentile="51.82" angles_rmsz="0.71"
    bonds_rmsz="0.47" chain="A" ent="1" model="1" num_angles_rmsz="1491" num_bonds_rmsz="1107" relative_RSRZ_percentile="100.00"
    relative_rama_percentile="100.00" relative_sidechain_percentile="38.37" said="A"/>
</wwPDB-validation-information>
```

Entry-level validation information:

Percentiles, overall validation metrics (e.g. Rfree), statistics.

Residue-level validation information:

Geometric outliers, torsion angles, RSR, RSRZ, clash score, occupancy

validation report of modeled subgroups (residues) repeat...

Entity-level validation information: Percentiles

validation report of modeled entities repeat...

PDBML-validation and wwPDB/RDF-validation

PDBML-validation:

Note that PDBML-validation is an experimental archive and may be changed or replaced in the future.

```
% rsync -av --delete rsync://bmrpub.pdbj.org/pdbml-valid .
```

wwPDB/RDF-validation:

Note that wwPDB/RDF-validation is an experimental archive and may be changed or replaced in the future.

```
% rsync -av --delete rsync://bmrpub.pdbj.org/wwpdb-rdf-valid .
```

PostgreSQL dump image:

category	description	size (GB)
pdbx_dcc_map	output of MAPMAN used by DCC (RSR, RSCC, LLDF)	10
pdbx_poly_seq_scheme	residue nomenclature mapping for polymer entities	8
struct_mon_prot	structure properties of a protein	6
entity_poly_seq	sequence of monomers in a polymer	2
pdbx_validate_close_contact	close contact with regard to the distance expected	2

SPARQL endpoint contains wwPDB/RDF-validation graph

<https://bmrpub.pdbj.org>

PDBj-BMRB Data Server:

common open representations of BMRB NMR-STAR data in XML, RDF and JSON formats

[Home](#) [Search](#) [Examples](#) [Download](#) [Resources](#) [NEWS](#)

Virtuoso SPARQL Query Editor

[About](#) | [Namespace Prefixes](#) | [Inference rules](#)

Default Data Set Name (Graph IRI)

Query Text

```
select distinct ?Concept where {[ ] a ?Concept} LIMIT 100
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

Execution timeout:

milliseconds *(values less than 1000 are ignored)*

Options:

Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see [details](#))

Query examples

Category holders

1. Select all category holders of datablock class of BMRB entry 15400:
2. Select all category holders of datablock class of Metabolomics entry bmse000400:

Entry statistics

3. Count entries per submission year and experimental method (subtype):

Assembly descriptions

4. Select all assembly names, asym IDs, entity IDs, polymer types, formula weights and functions in a assembly:

Entity descriptions


5. Select all entity names and sequences of polymer entities expressed using one-letter code:
6. Select all original source information of molecular entities and external links to NCBI Taxonomy:
7. Select all biological systems to produce molecular entities and external links to NCBI Taxonomy:

Citation information

8. Select citation information of all entries together with external links to PubMed and DOI, if available:

Example #1: Search wwPDB/RDF-validation with SPARQL


Search all enzyme–ligand complexes of which real space R–factor (RSR) of ligand is less than 10%. (showing only essential part of about 30 line–SPARQL query)

```
PREFIX PDBov: <https://rdf.wwpdb.org/schema/pdbx-validation-v1.owl#>
SELECT ?PDB_ID ?enzyme ?ligand ?comp_id MIN(?RSR AS ?minRSR)
FROM <http://rdf.wwpdb.org/pdb-validation>
WHERE {
  ?entity PDBov:link_to_enzyme ?link_to_enzyme ;
    PDBov:entity.pdbx_description ?enzyme ;
    PDBov:of_datablock ?datablock .
   selection of enzyme


  BIND (SUBSTR(STR(?datablock),38,4) AS ?PDB_ID)

  ...

  FILTER (?ligand != "water" && !STRENDS(?ligand, "ION"))

  ...
   ligand selection: non–polymer,
  not water, not ion
  ?dcc_map PDBov:pdbx_dcc_map.auth_asym_id ?asym_id ;
    PDBov:pdbx_dcc_map.auth_comp_id ?comp_id ;
    PDBov:pdbx_dcc_map.RSR ?RSR .

  FILTER (xsd:float(?RSR) < 0.1)

} GROUP BY ?PDB_ID ?enzyme ?ligand ?comp_id  RSR < 0.1
```

Example #1: Search wwPDB/RDF-validation with SPARQL

Found 15k pairs of enzyme–ligand complexes of which real space R–factor (RSR) of ligand is less than 10%.

PDB ID, enzyme name, ligand name, ligand (3–letters code), minimum RSR value of the ligand

"4CK1", "INTEGRASE", "(4-CARBOXY-1,3-BENZODIOXOL-5-YL)METHYL-[[2-[(4-METHOXYPHENYL)METHYLCARBAMOYL]PHENYL]METHYL]AZANIUM", "OM1", "0.081"
"2IOD", "Dihydroflavonol 4-reductase", "NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE", "NAP", "0.091"
"5BYR", "Iron hydrogenase 1", "FE2/S2 (INORGANIC) CLUSTER", "FES", "0.096"
"2PU0", "Enolase", "PHOSPHONOACETOHYDROXAMIC ACID", "PAH", "0.075"
"4LV2", "Beta-lactamase", "[1-(6-chloropyrimidin-4-yl)-1H-pyrazol-4-yl]boronic acid", "N95", "0.083"
"3OLE", "Pancreatic alpha-amylase", "ALPHA-D-GLUCOSE", "GLC", "0.084"
"4MOR", "Pyranose 2-oxidase", "DODECAETHYLENE GLYCOL", "12P", "0.093"
"4FKX", "Nucleoside diphosphate kinase", "CYTIDINE-5'-DIPHOSPHATE", "CDP", "0.073"
"4JPU", "Cytochrome c peroxidase", "PROTOPORPHYRIN IX CONTAINING FE", "HEM", "0.094"
"1GTV", "THYMIDYLATE KINASE", "THYMIDINE-5'-DIPHOSPHATE", "TYD", "0.090"
"5IA2", "7-(5-hydroxy-2-methylphenyl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[2,1-f]purine-2,4(3H,8H)-dione", "7-(5-hydroxy-2-methylphenyl)-8-(2-methoxyphenyl)-1-methyl-1H-imidazo[2,1-f]purine-2,4(3H,8H)-dione", "L66", "0.077"
"1NFQ", "Putative oxidoreductase Rv2002", "1,4-DIHYDRONICOTINAMIDE ADENINE DINUCLEOTIDE", "NAI", "0.090"

Example #2: Search wwPDB/RDF-validation with SPARQL

Search all enzyme–ligand complexes of which percentage of outlier in real space R–factor, defined by Z–score (RSRZ) is larger than 2, of enzyme is less than 1%. (showing only essential part of about 20 line–SPARQL query)

```
PREFIX PDBov: <https://rdf.wwpdb.org/schema/pdbx-validation-v1.owl#>
SELECT ?PDB_ID ?Enzyme (GROUP_CONCAT(?Ligand; SEPARATOR=",") AS ?Ligands)
?RSRZ_outliers_percent
FROM <http://rdf.wwpdb.org/pdb-validation>
WHERE {
  ?map_overall PDBov:pdbx_dcc_map_overall.entry_id ?PDB_ID ;
    PDBov:pdbx_dcc_map_overall.RSRZ_outliers_percent ?RSRZ_outliers_percent .


  FILTER (xsd:float(?RSRZ_outliers_percent) < 0.01)
  BIND (IRI(CONCAT("https://rdf.wwpdb.org/pdb-validation/", ?PDB_ID, "/entityCategory")) AS
?entity_category)


  ?entity_category PDBov:has_entity ?entity .


  ?entity PDBov:link_to_enzyme ?link_to_enzume ;
    PDBov:entity.pdbx_description ?Enzyme .

  ...

}
```

 % of outliers in RSRZ < 1%

 selection of enzyme

 ligand selection (omission)

Example #2: Search wwPDB/RDF-validation with SPARQL

Found 5k pairs of enzyme–ligand complexes of which percentage of outlier in real space R-factor (RSRZ) of enzyme is less than 1%, 1k pairs for 0%.

PDB ID, enzyme name, ligand name, percentage of outliers in RSR value of the enzyme

"1BUL", "NMC-A BETA-LACTAMASE", "2-(1-CARBOXY-2-HYDROXY-2-METHYL-PROPYL)-5,5-DIMETHYL-THIAZOLIDINE-4-CARBOXYLIC ACID, 2-(N-MORPHOLINO)-ETHANESULFONIC ACID", "0.00"

"5A1G", "S-ADENOSYLMETHIONINE SYNTHASE ISOFORM TYPE-2", "(DIPHOSPHONO)AMINOPHOSPHONIC ACID, [(3S)-3-amino-3-carboxypropyl]{[(2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl]methyl}ethylsulfonium, (4S)-2-METHYL-2,4-PENTANEDIOL, IMIDAZOLE", "0.00"

"2DRS", "Xylanase Y", "GLYCEROL", "0.00"

"2AS1", "Cytochrome c peroxidase, mitochondrial", "PROTOPORPHYRIN IX CONTAINING FE, THIOPHENE-3-CARBOXIMIDAMIDE", "0.00"

"1H4W", "TRYPSIN IVA", "BENZAMIDINE", "0.00"

"142L", "T4 LYSOZYME", "BETA-MERCAPTOETHANOL", "0.00"

"4CIK", "PLASMINOGEN", "5-[(2R,4S)-2-(phenylmethyl)piperidin-4-yl]-1,2-oxazol-3-one", "0.00"

"4L4O", "Endo-1,4-beta-xylanase", "TRIS-HYDROXYMETHYL-METHYL-AMMONIUM", "0.00"

"4G5P", "Epidermal growth factor receptor", "N-{4-[(3-chloro-4-fluorophenyl)amino]-7-[(3S)-tetrahydrofuran-3-yloxy]quinazolin-6-yl}-4-(dimethylamino)butanamide", "0.00"

"3GA6", "Exodeoxyribonuclease", "GLYCEROL", "0.00"

SIFTS: Structure Integration with Function, Taxonomy and Sequences resource

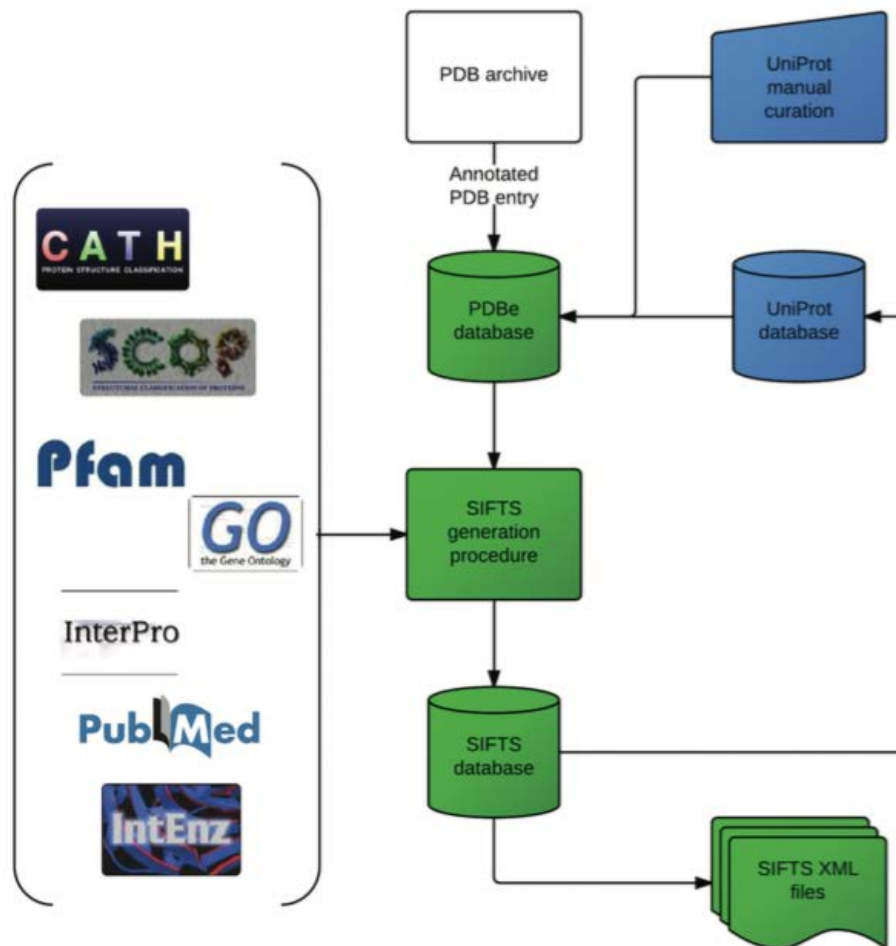


Figure 1. The SIFTS pipeline combines manual and automated processes to produce up-to-date residue-level mappings between proteins in the PDB and their corresponding UniProtKB entry. The pipeline also enriches the annotations of proteins in the PDB by adding data from other biological resources. The SIFTS data are distributed in XML format.

If you want to do complicated queries,
we may be able to help!

(質問はお気軽にPDBjまで)

Feel free to ask any questions at:

<https://pdbj.org/contact?tab=PDBjmaster>