

日本生物物理学会 年会2012年、名古屋大学、
PDBjランチョンセミナー 2012年9月23日

Roles of wwPDB and PDBj towards Big Data Era

Big Data時代に向けたwwPDBとPDBjの役割

Haruki Nakamura

中村 春木



Institute for Protein Research

大阪大学蛋白質研究所

<http://pdbj.org/>

<http://wwpdb.org/>

Big Data



Office of Science and Technology Policy
Executive Office of the President
New Executive Office Building
Washington, DC 20502

FOR IMMEDIATE RELEASE
March 29, 2012

Contact: Rick Weiss 202 456-6037 rweiss@ostp.eop.gov
Lisa-Joy Zgorski 703 292-8311 lisaioy@nsf.gov

OBAMA ADMINISTRATION UNVEILS "BIG DATA" INITIATIVE: ANNOUNCES \$200 MILLION IN NEW R&D INVESTMENTS

Aiming to make the most of the fast-growing volume of digital data, the Obama Administration today announced a "Big Data Research and Development Initiative." By improving our ability to extract knowledge and insights from large and complex collections of digital data, the initiative promises to help solve some of the Nation's most pressing challenges.

To launch the initiative, six Federal departments and agencies today announced more than \$200 million in new commitments that, together, promise to greatly improve the tools and techniques needed to access, organize, and glean discoveries from huge volumes of digital data.

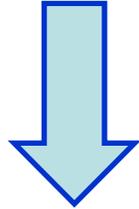
- Advance state-of-the-art core technologies needed to **collect, store, preserve, manage, analyze, and share huge quantities of data.**
- Harness these technologies to accelerate the pace of discovery in science and engineering, strengthen our national security, and transform teaching and learning; and
- Expand the workforce needed to develop and use Big Data technologies.

NIH is particularly interested in imaging, molecular, cellular, electrophysiological, chemical, behavioral, epidemiological, clinical, and other data sets related to health and disease.

(In the life-science fields, **Genome** data, **Image** data, and **Medical** data are typical Big Data. For example, 200 TB Data in "1000 Genomes Project" will be released from Amazon Web Services, AWS).

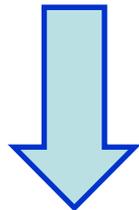
Paradigm in Science

Experiment, Theory



Sir Isaac Newton
(1642-1727)

Third Paradigm:
Computational Science

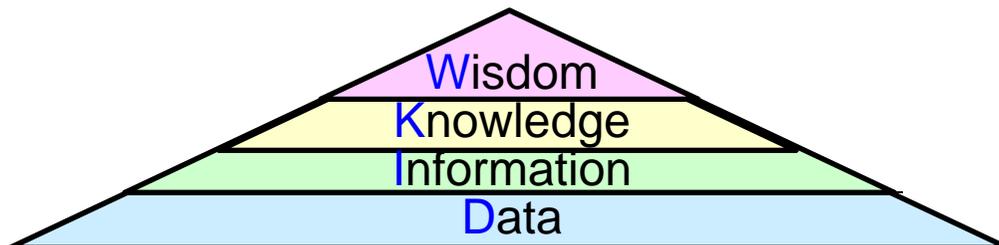


10 PF (京)



Fourth Paradigm:
Data-intensive Science

16 → 20 PF
(Sequoia)



DIKW pyramid

(J. Rowley: The wisdom hierarchy, 2007)

“The Fourth Paradigm: Data-intensive Scientific Discovery”

James P. Collins, *SCIENCE* 327, 19 MARCH 2010, page 1455

Researchers in genomics, astronomy, and many other active areas of science face a different challenge: Gathering data is so easy and quick that it exceeds our capacity to validate, analyze, visualize, store, and curate the information. The Fourth Paradigm addresses this challenge - and the opportunity it presents.



At SDSC by Prof. Hideo Matsuda (Osaka Univ.)

What Super-Computer is expected for the big-data era?

GORDON: Designed for data-intensive study **to store 100,000 entire human genomes in flash memory** by Appro & SDSC using Intel Xeon E5 (Sandy Bridge) **16384 cores, 341 TF, 64 TB DRAM, 300 TB Flash Memory, 4 PB HDD,** QDR InfiniBand, 3D Torus interconnect

Integration of DBs for Data-intensive Science

NBDC (バイオサイエンスデータベースセンター)

The screenshot shows the NBDC website interface. At the top, the NBDC logo and name are displayed, along with the JST logo and text: 独立行政法人 科学技術振興機構 (JST Japan Science and Technology Agency). Navigation links include Home, About Us, Contact Us, and Link. A search bar is present with the text "Japanese | sitemap | Site search".

The main content area features a navigation menu on the left with options: Welcome from the Director-General, Projects and Activities, Organization, and List of Members. The central focus is the "Organization" section, which contains a detailed organizational chart for the National Bioscience Database Center (NBDC).

The organizational chart is structured as follows:

- Council for Science and Technology Policy (CSTP)**
 - Life Science Project Team
 - DB Integration Promotion Task Force
- National Bioscience Database Center (NBDC)**
 - Director-General Michio Oishi, Ph.D.
 - Deputy Director-General Toshihisa Takagi, Ph.D.
 - Steering Committee
 - Department of Planning and Management
 - Researchers
 - Program Concerning Technology Development for Database Integration
 - Research Supervisor: Takeshi Nagase, Ph.D.
 - Research Subject
 - Research Advisor (3)
 - Program for Coordination Toward Integration of Related Databases
 - Research Supervisor: Toshihisa Takagi, Ph.D.
 - Research Subject (3)
 - Research Advisor (3)

Below the organizational chart, there is a section titled "About the Director-General (Michio Oishi, Ph.D.)" which includes a small portrait photo of Michio Oishi and his educational background:

Education :

- 1958 B. S. Dept. of Biology, Faculty of Sciences, The University of Tokyo
- 1963 Ph.D. Dept. of Biochemistry and Biophysics, Faculty of Sciences, The University of Tokyo



データベースのカタログ

[Inteqbioデータベースカタログβ版](#)

([生命科学系 データベース カタログ](#) とWINGproが含まれています)



コンテンツ

[生命科学系 学協会 カタログ](#)

[生命科学系主要プロジェクト一覧](#)

[生物アイコン](#)

[Webリソースポータルサイト](#)

[ゲノム解析ツールリンク集](#)

[ライフサイエンス 新着論文レビュー](#)

[ライフサイエンス 領域融合レビュー](#) **New!!**

[統合TV](#)

[MDeR](#)



検索

[生命科学データベース横断検索](#)

[TogoProt](#)

[GenLibi](#)

[Allie](#)

[inMeXes](#)

[SRAs \(β版\)](#)

[鎖鎖 \(β版\)](#)

[統合遺伝子検索 GGRNA \(β版\)](#)

[HOWDY](#)

[RefEx \(β版\)](#)



広報

[NBDC広報サイト](#)

[パンフレット\(PDF:1.59MB\)](#)



アーカイブ

[生命科学系データベースアーカイブ](#)



開発ツール

[TogoDB](#)

[TogoWS](#)

[DBCLS Galaxy](#)

[アナトモグラフィ-/BodyParts3D](#)

[TogoDoc](#)



データベース

[疾患・医薬品・環境物質関連データベース](#)

[植物関連データベース](#) **New!!**

[蛋白質構造関連データベース](#)

[糖鎖関連データベース](#)

[メタボローム関連データベース](#)

[微生物関連データベース](#)

[ヒトゲノム多様性関連データベース \(旧ヒトゲノムバリエーションデータベース\)](#)

[フェノーム関連データベース](#)

[データベース統合化基盤技術](#)



連携

[inteqbio.jp](#)

[BioHackathon](#)

<http://biosciencedbc.jp/>

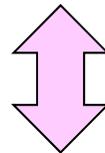
Organization of National Bioscience Database Center (NBDC)

CSTP in Cabinet Office of Japanese Government

Life Science Project
Team in CSTP

Headquarter of NBDC
in CSTP

Advisory
Committee



Director of National Bioscience Database Center

JST

National Bioscience Database Center

DBs

DB

DB

DB

PDBj

DB

DB

DB

Protein Data Bank Japan

日本蛋白質構造データバンク

<http://pdbj.org/>

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



PDBj / PDBj-BMRB staffs (April 2012)

- **Head**
 - Nakamura, Haruki, Ph. D. (Prof., IPR, Osaka Univ.)
- **Group for PDB Database Curation**
 - Nakagawa, Atsushi, Ph. D. (Group Leader, Prof., IPR, Osaka Univ.)
 - Matsuda, Makoto (IPR, Osaka Univ.)
 - Igarashi, Reiko (IPR, Osaka Univ.)
 - Kengaku, Yumiko (IPR, Osaka Univ.)
 - Cho, Hasumi (IPR, Osaka Univ.)
 - Ikegawa, Yasuyo (IPR, Osaka Univ.)
 - Sato, Junko (IPR, Osaka Univ.)
- **Group for Development of new tools and services**
 - Kinjo, Akira, Ph. D. (IPR, Osaka Univ.)
 - Iwasaki, Kenji, Ph. D. (IPR, Osaka Univ.)
 - Suzuki, Hirofumi, Ph. D. (IPR, Osaka Univ.)
 - Yamashita, Reiko (IPR, Osaka Univ.)
 - Kudou, Takahiro (IPR, Osaka Univ.)
 - Nishikawa, Ken, Ph. D. (Guest Prof., IPR, Osaka Univ.)
 - Bekker, Gert-Jan (IPR, Osaka Univ.)
- **Group for BMRB**
 - Fujiwara, Toshimichi, Ph. D. (Group Leader, Prof. Osaka Univ.)
 - Akutsu, Hideo, Ph. D. (Prof., IPR, Osaka Univ.)
 - Kojima, Chojiro (IPR, Osaka Univ.)
 - Kobayashi, Naohiro (IPR, Osaka Univ.)
 - Iwata, Takeshi (IPR, Osaka Univ.)
 - Takahashi, Ami (IPR, Osaka Univ.)
- **Collaboratory Researchers**
 - Wako, Hiroshi, Ph. D. (Prof., Waseda Univ.) (**for Pro Mode**)
 - Ito, Nobutoshi, Ph. D. (Prof., Tokyo Medical and Dental Univ.)
 - Kinoshita, Kengo, Ph.D. (Prof., Tohoku Univ.) (**for eF-site**)
 - Standley, Daron, Ph. D. (IFReC, Osaka Univ.) (**for SeqNavi, StructNavi, SeSAW, and ASH**)
 - Katoh, Kazutaka, Ph. D. (IFReC, Osaka Univ.) (**for ASH**)
- **Secretary**
 - Haruki, Nahoko (IPR, Osaka Univ.)

WORLDWIDE
wwPDB
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

NSF, NIGMS, DOE,
NLM, NCI, NINDS,
NIDDK



BioMagResBank
www.bmrwisc.edu

NLM

EMBL-EBI **PDB^e**
PROTEIN DATA BANK EUROPE

Protein Data Bank in Europe

pdbe.org

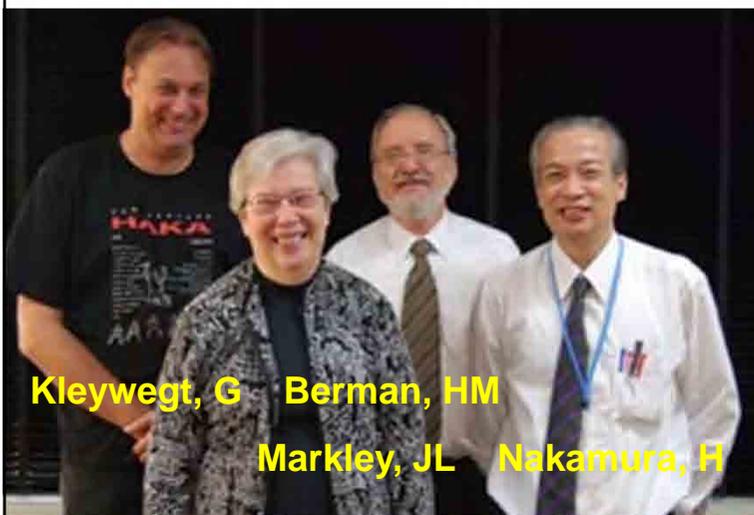
EMBL-EBI,
Wellcome Trust,
BBSRC, NIGMS, EU

PDB^j
Protein Data Bank Japan

Protein Data Bank Japan

www.pdbj.org

NBDC-JST



Kleywegt, G Berman, HM
Markley, JL Nakamura, H

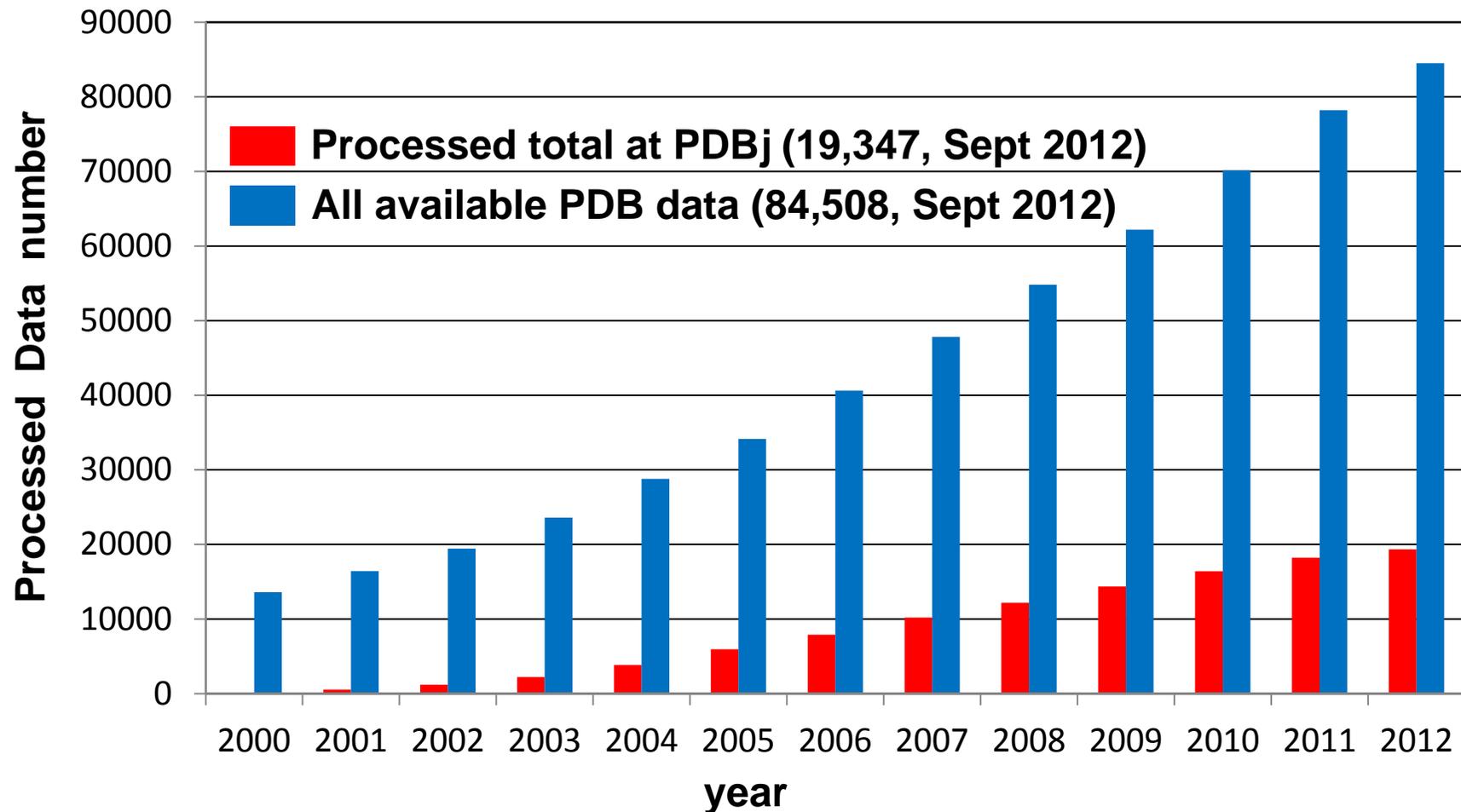
**wwPDB and wwPDBAC members
at EBI, Hinxton, on Sept. 30 2011**



Activities/Services of each member of the wwPDB

- **“Data-in”** activity, common in all the wwPDB members with high quality control.
- **“Data-out”** services, characteristic at each wwPDB member site except common ftp-site data. Concerted but sometimes competing.

Data-in at PDBj and wwPDB



PDBj curates and processes about **a Quarter of the deposited data**, mainly from Asian and Oceania regions.

Validation of PDBx and PDBML description

The program package will be distributed to each user



Data type validation tool for PDBML/PDBx

PDBML is validated against PDBML Scheme (pdbx-v40.xsd)

PDBx is validated against dictionary definitions (mmcif_pdbx_v40.dic)



Validation Results

Validated file: ./data2/4g44_2.xml

No.	XPath	Message	line
1	/datablock/refineCategory/refine[@entry_id="4344"]/ls_d_res_low	Error: value '47.x' does not match any member types of the union	96042

% dictionary.version 4.024

number of categories: 151 -> 82

Options for Error messages:

Errors for public cif (default) **All errors**

Errors for public cif files are displayed as default and all errors for internal cif files can be displayed when selected.

Category name	data type errors	mandatory errors	primary key errors	enumeration errors	data in updated cif file (lines from 429 to 436)
._refine	data type: 0	mandatory: 0	primary key: 0	enumeration: 0	all
._struct_conf_type	data type: 0	mandatory: 0	primary key: 0	enumeration: 0	all
._database_PDB_rev	data type: 0	mandatory: 1	primary key: 1	enumeration: 0	all
._struct_biol	data type: 0	mandatory: 0	primary key: 0	enumeration: 0	all
._refine	data type: 1	mandatory: 0	primary key: 0	enumeration: 0	alias: 18

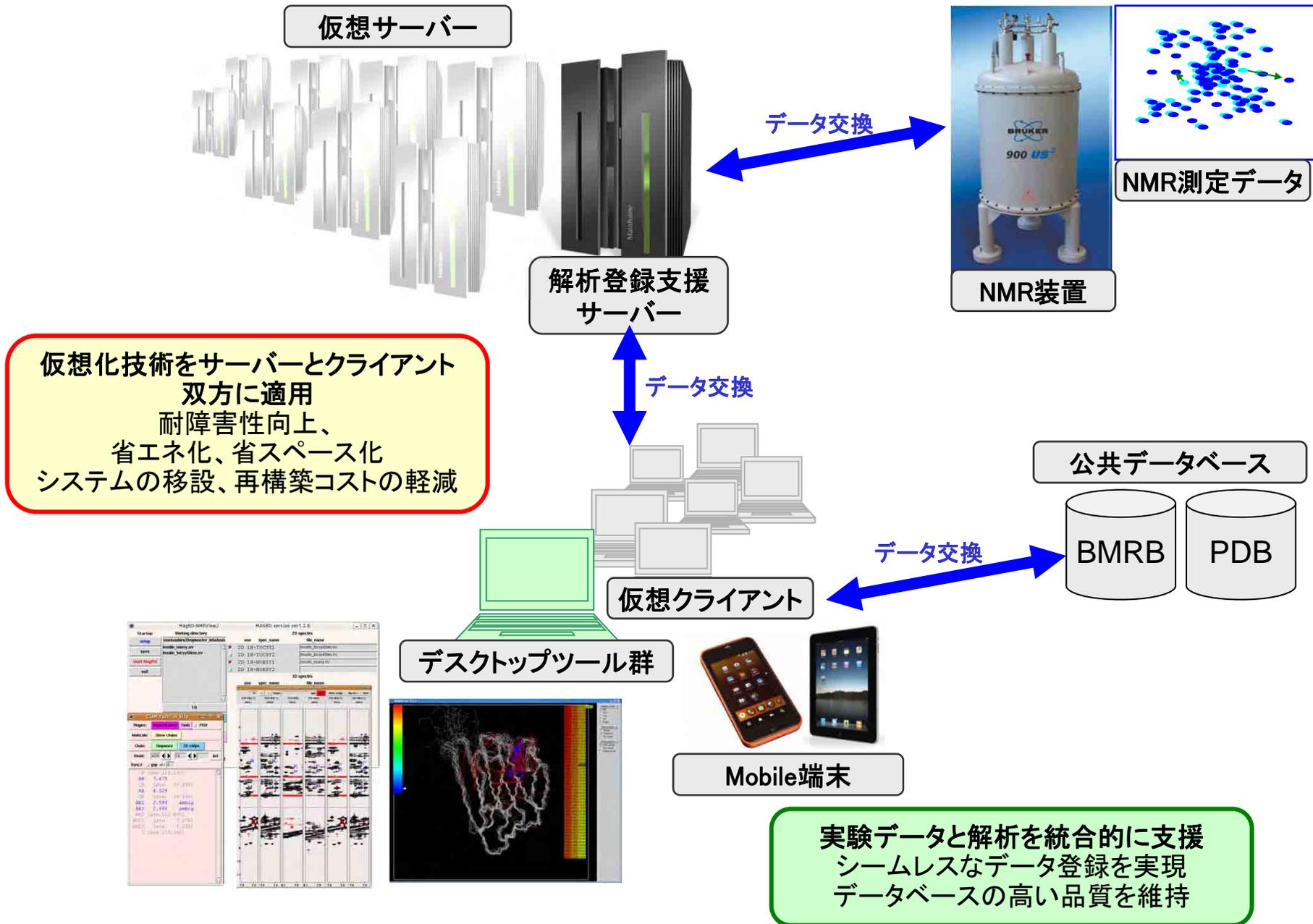
< Error type >	< Item name >	< error message >	< line >
Error(data_types)	._refine.ls_d_res_low	{ row 1, col 9 } = <<47>>; float /-?((0-9)+)[.](0-9)+([0-9]+)X(0-9)+D)X([E][+-]?(0-9)+)?/	439 data

Errors exist in the following data

```

data in updated cif file (lines from 429 to 436)
%
_refine.entry_id          RCSB173713
_refine.ls_number_refine RCSB173713
_refine.ls_number_refine RCSB173713
_refine.ndb_ls_siana_F    ?
_refine.ndb_ls_siana_F    0
_refine.ndb_data_cutoff_high_abcF  ?
_refine.ndb_data_cutoff_low_abcF   ?
_refine.ndb_data_cutoff_high_abcF  ?
_refine.ls_d_res_low      I 47.x
    
```

Data-in at PDBj-BMRB (NMR experimental information)



Data-out: wwPDB FTP Traffic



30,627,357 files were downloaded during March 2012 from the wwPDB member sites (RCSB-PDB, EBI-PDBe, and PDBj).

Data-out from PDBj Top Page

The screenshot shows the PDBj website interface. At the top, there are language options: English, Japanese, simplified Chinese, traditional Chinese, and Korean. On the right, there are links for 統計情報, ヘルプ, FAQ, and お問い合わせ. The main content area is divided into several sections:

- トップページ**: Includes links for データ登録 >>, 検索 >>, サービス&ソフトウェア >>, タウンロード >>, and リンク集.
- データ登録**: Information about the Protein Data Bank Japan, mentioning JST-NBDC and Osaka University support. It includes a link for データ登録のご案内 >> and buttons for PDB登録 and NMRデータ登録.
- 検索**: A search section featuring the **PDB検索 Mine** logo. A red circle highlights the search input field labeled "PDB IDまたはキーワード" and the "検索" button. Below it is a link for "詳細条件検索 >>". To the right, there are radio buttons for "Accession number" (selected) and "Deposition code", along with a "Go" button.
- 最新情報**: A list of recent news items with dates and brief descriptions, such as "2012/9/13 2012年9月23日(日)に、第50回日本生物物理学会年会にて、ランチョンセミナーを開催いたします。" and "2012/8/16 2012年10月13日(土)に、大阪梅田にてwwPDB Outreach講演会が開催されます。"

On the right side of the page, there are several logos and statistics:

- A red circle highlights the text: **84508 entries available on 12 Sep., 2012** (00:00(UTC) / 09:00(JST)).
- Logos for PDB (Protein Data Bank), eProts (Encyclopedia of Protein Structures), Protein Globe, DBCLS (Database Center for Life Sciences), Tanpaku.org, and NBDC (National Institute of Advanced Industrial Science and Technology).

Get Entry Data from our browser

Access to <http://pdbj.org/>

The screenshot shows the PDBj homepage in English. The search bar is circled in red, showing the search criteria: "PDB IDまたはキーワード" and "トリプトファン合成酵素". The search button is labeled "検索".

The screenshot shows the PDBj search results page in Japanese. The search criteria are "トリプトファン合成酵素" and "PDB ID or Keyword". The results show two entries for "TRYPTOPHAN SYNTHASE ALPHA CHAIN".

2zlh	分子名	: TRYPTOPHAN SYNTHASE ALPHA CHAIN (E.C.4.2.1.20), TRYPTOPHAN SYNTHASE BETA CHAIN (E.C.4.2.1.20)
	タイトル	: TRYPTOPHAN SYNTHASE IN COMPLEX WITH (NAPHTHALENE-2'-SULFONYL)-2-AMINO-1-ETHYLPHOSPHATE (F15)
	著者	: Ngo, H., Harris, R., Kimmich, N., Casano, P., Nika, D., Blumenstein, L., Baranda, T.R., Kulik, V., Weyand, M., Schlichting, J., Dunn, M.F.
	実験手法	: X-RAY DIFFRACTION
	登録日	: 2008-04-27
	公開日	: 2007-08-12

2zlk	分子名	: TRYPTOPHAN SYNTHASE ALPHA CHAIN (E.C.4.2.1.20), TRYPTOPHAN SYNTHASE BETA CHAIN (E.C.4.2.1.20)
	タイトル	: TRYPTOPHAN SYNTHASE IN COMPLEX WITH D-GLYCERALDEHYDE 3-PHOSPHATE (G3P)
	著者	: Ngo, H., Harris, R., Kimmich, N., Casano, P., Nika, D., Blumenstein, L., Baranda, T.R., Kulik, V., Weyand, M., Schlichting, J., Dunn, M.F.
	実験手法	: X-RAY DIFFRACTION
	登録日	: 2008-04-27
	公開日	: 2007-08-12

日本語のキーワード
を入力

→ 検索結果ページ

Get Entry Data from our browser

Access to <http://pdbj.org/>

The screenshot shows the PDB Mine website interface in Japanese. The search bar is highlighted with a red circle, containing the text '12AS'. Below the search bar, there are buttons for '検索' (Search) and '詳細条件検索 >>' (Advanced Search). The page also features a sidebar with various navigation options and a main content area with search results and a 'Mine' logo.

The screenshot shows the summary page for PDB entry 12AS. The page includes a 3D molecular model of the protein structure and a table of entry information. The table contains the following data:

フィールド名	値
エントリ ID (PDB ID)	12as 配列情報 (FASTA形式) PDBファイルのダウンロード
分子名称	ASPARAGINE SYNTHETASE, L-ASPARAGINE, ADENOSINE MONOPHOSPHATE
タイトル	ASPARAGINE SYNTHETASE MUTANT C51A, C315A COMPLEXED WITH L-ASPARAGINE AND AMP
従属のキーワード	LIGASE, ASPARAGINE SYNTHETASE, NITROGEN FIXATION
由来する生物種	Escherichia coli K12
細胞内の位置	[UNP - P00963] Cytoplasm
ポリマー鎖の合計数	2
分子量の合計	74226 (詳細は 構造情報のページ)
著者	Nakatsu, T., Kato, H., Oda, J. (登録日: 1997-12-02, 公開日: 1998-12-30)
引用文献	Nakatsu, T., Kato, H., Oda, J. Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase. <i>Nature Struct. Biol.</i> , 5:15 - 19, 1998. (PubMed: 9437423) (DOI: 10.1038/nsb0198-15)
実験手法	X-RAY DIFFRACTION (2.2[Å])
他のデータベース情報	CATH, CE, FSSP, SCOP, VAST, UniProt (P00963), eF-site, KEGG (EC 6.3.1.1), GDB, E2CatDB, PISA, PQS

PDBID (e.g. 12as) is input in a box and GO



Summary for each PDBID is displayed.

Summary for each PDBID

<http://pd bj.org/>

Amino acid sequence (FASTA)

PDBj (Protein Data Bank Japan) maintains a centralized PDB archive of macromolecular structures and provides integrated tools, in collaboration with the RCSB, the BMRB in USA and the PDBe in EU. PDBj is supported by JST-BIRD.

Summary [1gof]

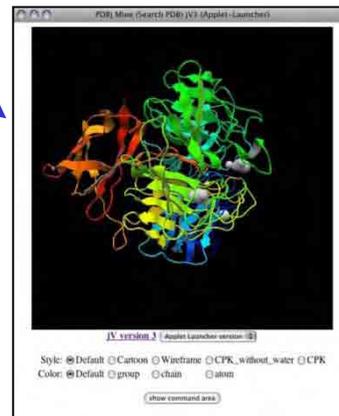
PDB ID	1gof	sequence information (FASTA format)	download PDB format file
Descriptor	GALACTOSE OXIDASE (E.C.1.1.3.9) (PH 4.5)		
Title	NOVEL THIOETHER BOND REVEALED BY A 1.7 ANGSTROMS CRYSTAL STRUCTURE OF GALACTOSE OXIDASE		
Functional Keywords	OXIDOREDUCTASE(OXYGEN(A))		
Biological source	Hypomyces rosellus		
Cellular location	[UNP - Q01745] Secreted		
Total number of polymer chains	1		
Total molecular weight	68785.9 (the details in Structural Details Page)		
Authors	Ito, N., Phillips, S.E.V., Knowles, P.F. (deposition date : 1993-09-30, release date : 1994-01-31)		
Primary citation	Ito, N., Phillips, S.E., Stevens, C., Ogel, Z.B., McPherson, M.J., Keen, J.N., Yadav, K.D., Knowles, P.F. Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. <i>Nature</i> , 350:87 - 90, 1991.(PubMed: 2002850) (DOI: 10.1038/350087a0)		
Experimental method	X-RAY DIFFRACTION (1.7[Å])		
Other Database Information	CATH, CE, FSSP, SCOP, VAST, UniProt (Q01745), eF-site, KEGG (EC 1.1.3.9), EzCatDB, PISA, PQS		

```
>1GOF:GALACTOSE OXIDASE
ASAFIGSALSRNNNAVTCDSAQSGNECKAIDGKEDTFPHITFYGANGDFKPHYTYIDMK
TTQVUNGLMPLPQDGHQGWIGRHEVYLSGDCRWGSPVAGSHFADSTTKYSNFETFP
ARVRLVAITZANQPWTSIAEIMVFGASSTYAPQGLGRWPTIDLFVPAARAEIPTS
GRVLNHSYRNDAPGSGPGEITLSSWDFSGIVSDNVTYVTRDMFCPIIINDGNGQIV
VTGNDAAKTLSDYSSDSWIPGDMQVARGYSATSDGRVFTIGGSMGSGVFEKNGE
VYSPSSKTFSLFNAKRVNMLFADKQGLYRSDHAWLFGKKGSVFQAGPFTAMNWTYS
GSGDYSAGRRKSNVQVAFDRCQVADNVRGKILTFGSGFQDSKATNARIITLG
EPGTSTVTFASNGLYVARTFESVLLPDCGTFITGQKRGILPFEDSTPFTPEIIVPEQ
DTFFKQNRSLVRYVHESILLPDCRVFNOGGLCGDCTNHFDAQITPHYLVSNGML
ATREKIRTSSTQSVKVGGRITISDSEISKALIRYGTATHTVMDQRRIPLTLTNGGN
SYSFVDFSDGVALFGYHMLFVHNSAGVFSVASTIRVTO
```

eF-site **1gof-A**

Functional site	focus & details	sequence	description	source
1) A:495	on off	Y	Proton acceptor.	Swiss-Prot: 1
2) A:172	on off			
3) A:495	on off			
4) A:496	on off			
5) A:581	on off			
6) A:228	on off			
7) A:290	on off			
8) A:581	on off			
9) A:499	on off			
10) A:75-87	on off			
11) A:194	on off			
12) A:227-228	on off			
13) A:272	on off			

Data viewer at PDBj



Graphic viewer: jV
<http://pd bj.org/jV/>

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

PDB/RDF format for Semantic Web

Service from wwPDB by Akira R. Kinjo (PDBj) & Tom Oldfield (PDBe)

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

WORLDWIDE PDB PROTEIN DATA BANK

Welcome to the Worldwide Protein Data Bank

About PDB/RDF
PDB/RDF , chem_comp/RDF

PDB ID: (e.g., '7RSA') PDB ID
property: (e.g., 'PDBo:entity.pdbx_description')
keywords: (e.g., 'alcohol')

Download XSLT stylesheet for converting PDBML to RDF: [PDBML2rdf.xsl.gz](#) (gzipped 22KB)

Published online 5 October 2011

Nucleic Acids Research, 2012, Vol. 40, Database issue D453–D460
doi:10.1093/nar/gkr811

Protein Data Bank Japan (PDBj): maintaining a structural data archive and resource description framework format

Akira R. Kinjo¹, Hirofumi Suzuki¹, Reiko Yamashita¹, Yasuyo Ikegawa¹, Takahiro Kudou¹, Reiko Igarashi¹, Yumiko Kengaku¹, Hasumi Cho¹, Daron M. Standley², Atsushi Nakagawa¹ and Haruki Nakamura^{1*}

¹Institute for Protein Research and ²Immunology Frontier Research Center, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565-0871, Japan

Received August 8, 2011; Accepted September 15, 2011

Kinjo et al. (2012) Nucl. Acids Res. 40, D453-D460.

Subject: <http://rdf.wwpdb.org/pdb/1GOF>

Predicate	Object
PDBo:datablockName	1GOF-noatom
PDBo:has_atom_sitesCategory	PDBr:1GOF/atom_sitesCategory
PDBo:has_atom_sites_footnoteCategory	PDBr:1GOF/atom_sites_footnoteCategory
PDBo:has_atom_typeCategory	PDBr:1GOF/atom_typeCategory
PDBo:has_audit_authorCategory	PDBr:1GOF/audit_authorCategory
PDBo:has_audit_conformCategory	PDBr:1GOF/audit_conformCategory
PDBo:has_cellCategory	PDBr:1GOF/cellCategory
PDBo:has_chem_compCategory	PDBr:1GOF/chem_compCategory
PDBo:has_citationCategory	PDBr:1GOF/citationCategory
PDBo:has_citation_authorCategory	PDBr:1GOF/citation_authorCategory
PDBo:has_computingCategory	PDBr:1GOF/computingCategory
PDBo:has_database_2Category	PDBr:1GOF/database_2Category
PDBo:has_database_PDB_matrixCategory	PDBr:1GOF/database_PDB_matrixCategory
PDBo:has_database_PDB_revCategory	PDBr:1GOF/database_PDB_revCategory
PDBo:has_database_PDB_rev_recordCategory	PDBr:1GOF/database_PDB_rev_recordCategory
PDBo:has_diffractionCategory	PDBr:1GOF/diffractionCategory
PDBo:has_diffraction_radiationCategory	PDBr:1GOF/diffraction_radiationCategory
PDBo:has_diffraction_radiation_wavelengthCategory	PDBr:1GOF/diffraction_radiation_wavelengthCategory
PDBo:has_entityCategory	PDBr:1GOF/entityCategory
PDBo:has_entity_keywordsCategory	PDBr:1GOF/entity_keywordsCategory
PDBo:has_entity_polyCategory	PDBr:1GOF/entity_polyCategory

PDB/RDF example

By accessing <http://rdf.wwpdb.org/pdb/1GOF>, a list of category holders for the PDB entry 1GOF can be retrieved in the RDF/XML format. Then, a list of category elements can be retrieved (again in the RDF/XML format). Finally, a particular category element, the list of properties of that element is retrieved.

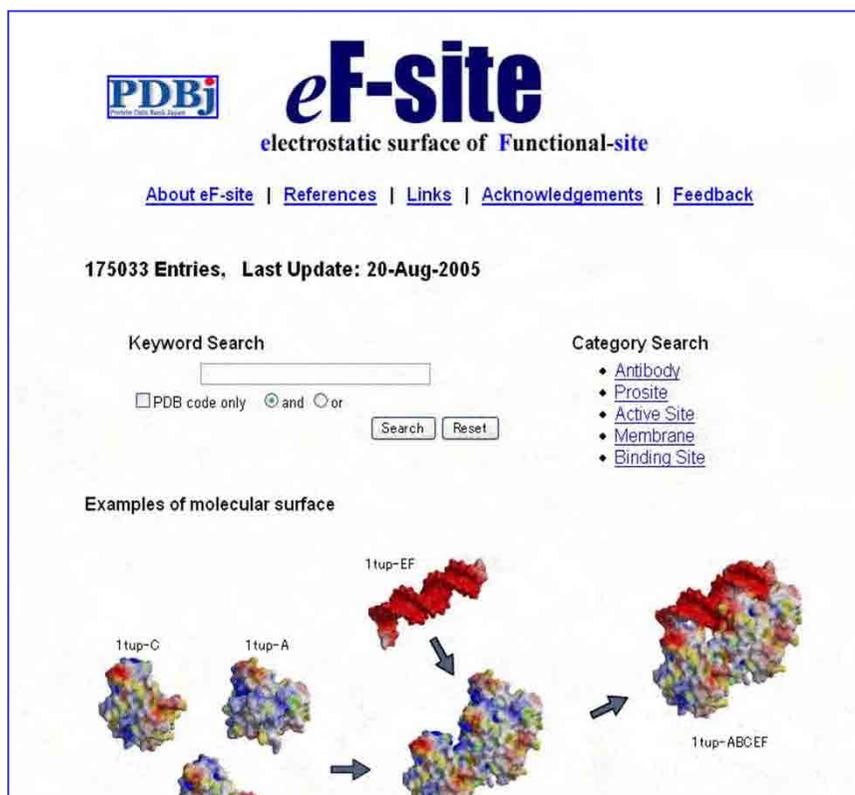
Subject: <http://rdf.wwpdb.org/pdb/1GOF/entityCategory>

Predicate	Object
PDBo:has_entity	PDBr:1GOF/entity/1
PDBo:has_entity	PDBr:1GOF/entity/2
PDBo:has_entity	PDBr:1GOF/entity/3
PDBo:has_entity	PDBr:1GOF/entity/4
PDBo:has_entity	PDBr:1GOF/entity/5
rdf:type	http://rdf.wwpdb.org/schema/pdbx-v40

Subject: <http://rdf.wwpdb.org/pdb/1GOF/entity/1>

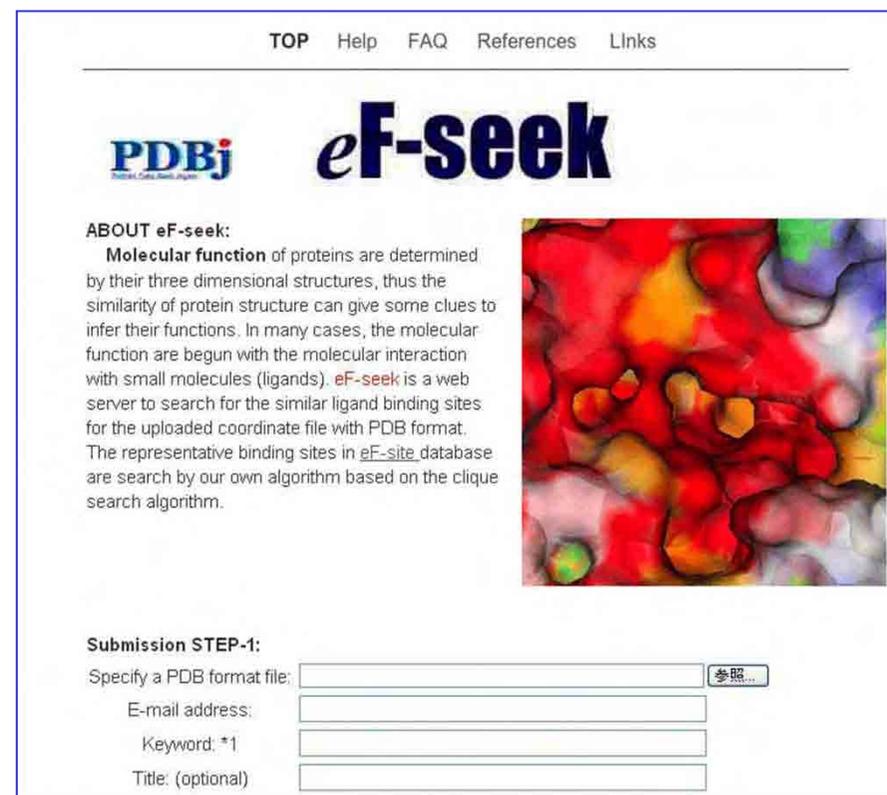
Predicate	Object
PDBo:of_datablock	PDBr:1GOF
PDBo:entity.formula_weight	68579.250
PDBo:entity.id	1
PDBo:entity.pdbx_description	GALACTOSE OXIDASE
PDBo:entity.pdbx_ec	1.1.3.9
PDBo:entity.pdbx_number_of_molecules	1
PDBo:entity.src_method	man
PDBo:entity.type	polymer
PDBo:link_to_enzyme	http://purl.uniprot.org/enzyme/1.1.3.9
PDBo:referenced_by_entity_keywords	PDBr:1GOF/entity_keywords/1
PDBo:referenced_by_entity_poly	PDBr:1GOF/entity_poly/1
PDBo:referenced_by_entity_src_gen	PDBr:1GOF/entity_src_gen/1
PDBo:referenced_by_struct_asym	PDBr:1GOF/struct_asym/A
PDBo:referenced_by_struct_ref	PDBr:1GOF/struct_ref/1
rdf:type	http://rdf.wwpdb.org/schema/pdbx-v40.owl#entity

Annotation of Protein Function from Molecular Surface Similarity: eF-site / eF-seek



The screenshot shows the eF-site website. At the top left is the PDB logo. The main title is "eF-site" with the subtitle "electrostatic surface of Functional-site". Below the title are navigation links: "About eF-site", "References", "Links", "Acknowledgements", and "Feedback". A status line indicates "175033 Entries, Last Update: 20-Aug-2005". There are two search sections: "Keyword Search" with a text input field and "and/or" radio buttons, and "Category Search" with a list of categories: Antibody, Prosite, Active Site, Membrane, and Binding Site. Below the search sections is a section titled "Examples of molecular surface" showing several protein surface models labeled 1tup-C, 1tup-A, 1tup-EF, and 1tup-ABC:EF, with arrows indicating relationships between them.

Protein Molecular Surface DB



The screenshot shows the eF-seek website. At the top right are navigation links: "TOP", "Help", "FAQ", "References", and "Links". The main title is "eF-seek" with the PDB logo to its left. Below the title is the "ABOUT eF-seek:" section, which explains that molecular function is determined by 3D structures and that eF-seek is a web server for finding similar ligand binding sites. To the right of the text is a colorful molecular surface visualization. Below the text is the "Submission STEP-1:" section, which includes input fields for "Specify a PDB format file:", "E-mail address:", "Keyword: *1", and "Title: (optional)", along with a "参照..." button.

Search for Similar Surface

Viewing Folds and Dynamics @ PDBj

PDBj Protein Globe (pre-beta)
A graphical user interface (GUI) to the PDBj database. [Description | Tutorial]

[-] <-zoom-> [+]

xPSSS SeqNavi StructNavi ef-site JV SCOP CATH reset

Select Display
 1. globe.xml.g
 2. latlong.xml.i
 3. cartoons.xtr
 4. sphere0.xm

PDB ID: 1d13
Chain: B
Domain: 0
xPSSS

This domain:

In PDB:

```
jv> Welcome to PDBj Globe!  
jv>  
jv> 1 Atom Selected.  
jv> Atom: 1d13 1647 Group: sco 1647 Chain: B Model: 0 File: 1  
jv>
```

Protein Globe: Protein Folds Browser

PDBj ProMode Elastic ProMode ProMode oligomer

Database of normal mode analysis of PDB data using elastic network model in torsional angle space

Home | What is ProMode-Elastic | Help Japanese

No. of entries 13

PDB code (4 chars) Find Example 1abc

Select from a list of entries

Submission of your data to be analyzed. (In preparation)

Download of

ProMode-Elastic using the pro Torsional A atoms in PD ligand molec molecules a

Reference: I

TOPICS
2009.12.12

PDB id: 7rsa Name: Hydrolase (phosphoric diester) Title: Structure of phosphate-free ribonuclease a refined at 1.26 angstroms Structure: Ribonuclease A Chain: A Ec

*Click on image for an enlarged image and more information.

PageTop | Back
(Latest update 2009.12.21)

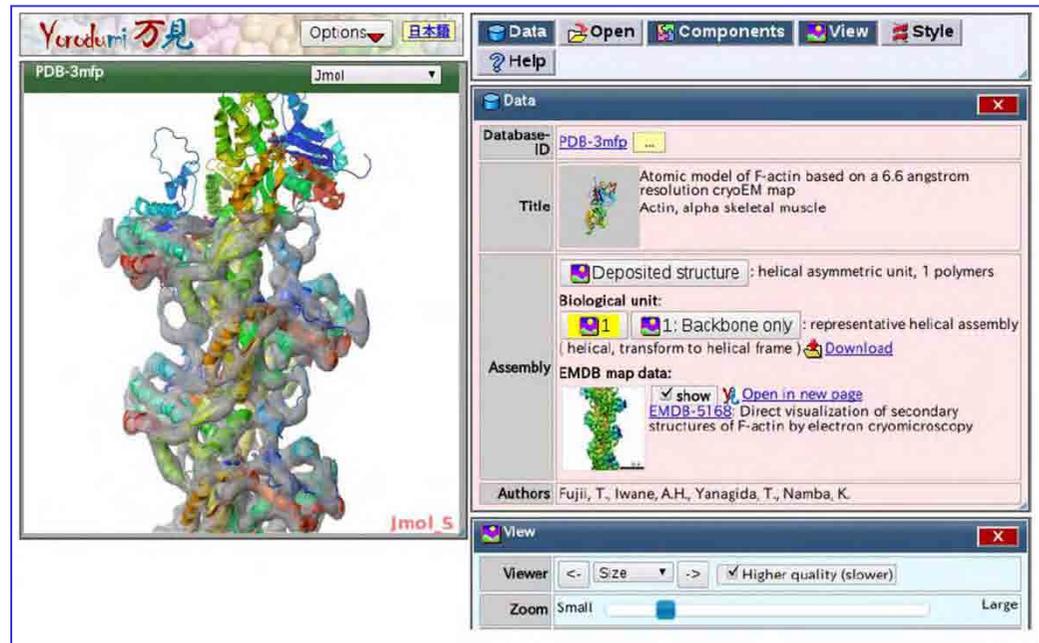
Copyright © WASEDA Univ.Japan. All rights reserved. Email: promode@list.waseda.jp

ProMode: Protein Dynamics Database by NMA

Viewing both EM Image and Atomic Structure



EM Navigator: Viewer of Images of EM-DB

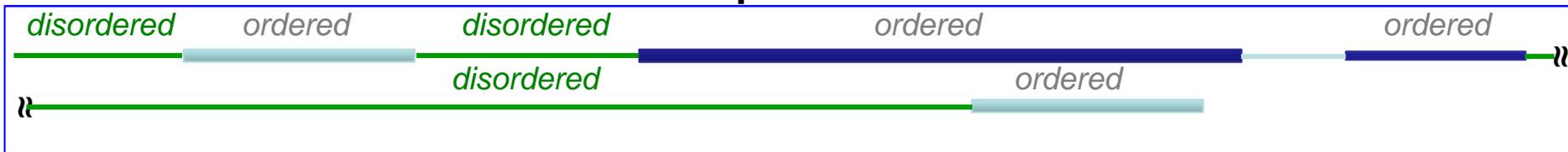


Yorodumi: Viewer of both Image and Atomic Structure

Analysis and Prediction of Protein Function from Sequence through Structure (D. M. Standley et al.)

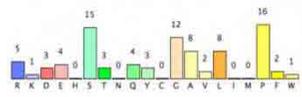
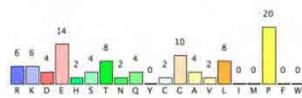
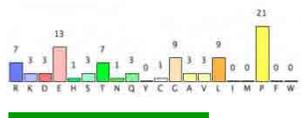
- **Structure-informed multiple sequence alignment using MAFFTash:**
<http://sysimm.ifrec.osaka-u.ac.jp/MAFFTash/>
- **Function prediction for ordered and disordered domains using SFAS:**
<http://sysimm.ifrec.osaka-u.ac.jp/sfas2/>

Example: zc3h12a



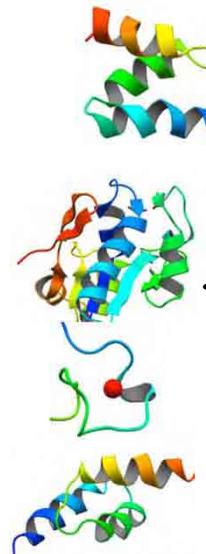
disordered Function Prediction *ordered*

Composition-based Function Prediction (IDD Navigator)



- RNA Processing
- Endonuclease
- Zinc binding
- Metal ion binding
- Intracellular/cytoplasm
- Zinc finger

Structure-based Function Prediction (Spanner, OSCAR & SeSAW)



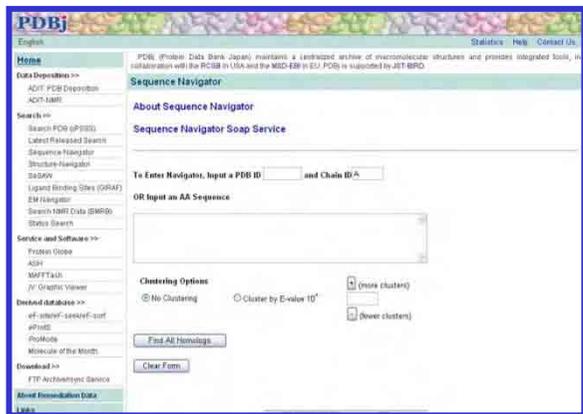
Ubq binding-like

PIN-like Rnase

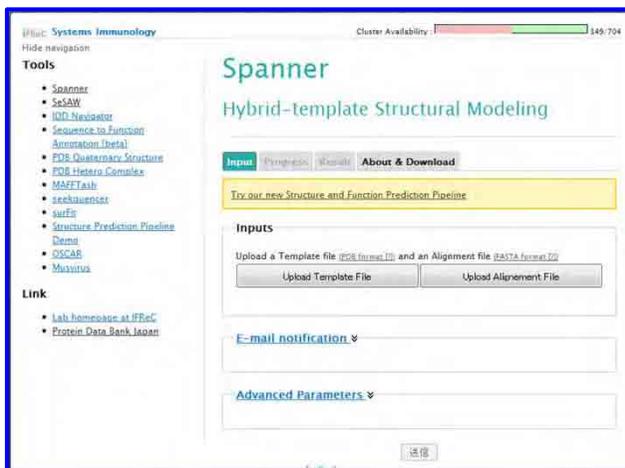
CCCH ZF

Ubq binding-like

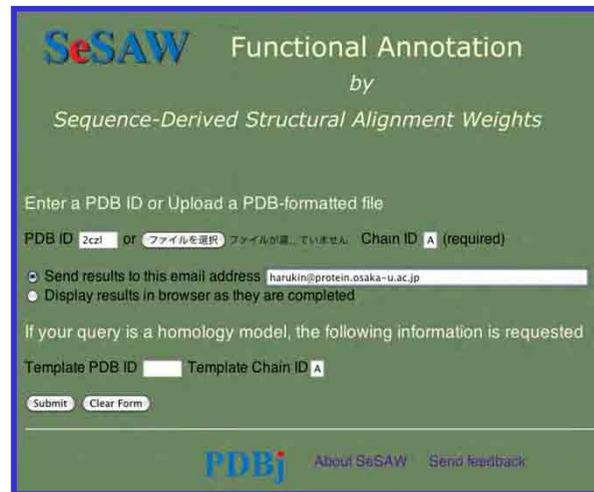
Development of other Databases and Services



Homolog protein search,
Sequence Navigator
(Standley)



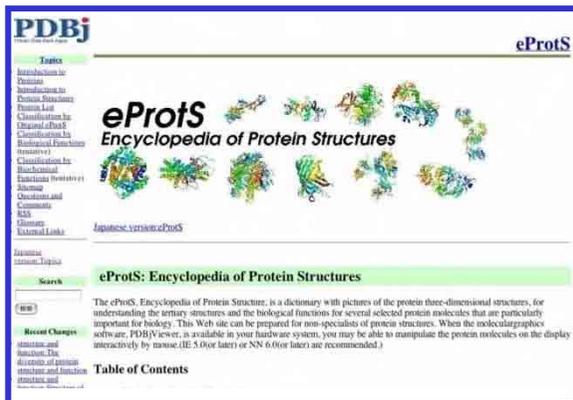
Homology modeling,
Spanner
(Lis, Standley & Nakamura)



Function Annotation from
Folds and Sequences,
SeSAW (Standley)



Alignment of Sequence and
Structures. MAFFTash
(Kato, Toh & Standley)



Encyclopedia of Protein
Structures, eProtS
(Kinjyo, Kudo, & Ito)



Molecule of the Month, MoM
(Goodsell & Kudo)

Home

Data Deposition >>

- ADIT: PDB Deposition
- ADIT-NMR

Search >>

- Search PDB (Mine/xPSSS)
- PDB/RDF, chem_comp/RDF
- Latest Release Search
- Sequence-Navigator
- Structure-Navigator
- SeSAW
- Ligand Binding Sites (GIRAF)
- EM Navigator
- Search NMR Data (BMRB)
- Status Search

Service and Software >>

- JV: Graphic Viewer
- Yorodumi
- Protein Globe
- ASH
- MAFFTash
- SEALA
- Structure Prediction >>
- CRNPRED
- Spanner
- SFAS

Derived database >>

- eF-site/eF-seek/eF-surf
- ePrintS

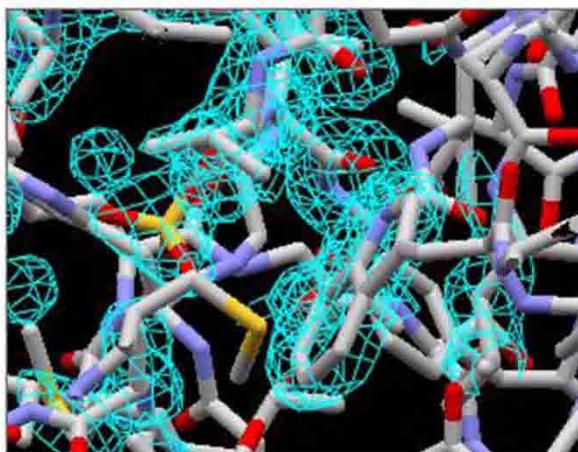
PDBj (Protein Data Bank Japan) maintains a centralized archive of macromolecular structures and provides integrated tools, in collaboration with the RCSB, the BMRB in USA and the PDBe in EU. PDBj is supported by JST-NBDC and Osaka University.



Electron Density Map [2xyj]

[About PDBj Mine Update Information](#)

About the PDBj Electron Density Map Viewer



```
jv>
Atom: S 2075 Group: S04 1272 Chain: A
File: 1
jv>
```

jv version 4

Applet Launcher (new Java Plug-in) version

- Style: Cartoon Backbone Wireframe
 Color: group chain atom

When the mouse right button is clicked in the jv4 window, t
 For a trouble to display a picture, please retry selecting
 "Applet Launcher (new Java Plug-in) version" or
 "Applet Launcher (classic Java Plug-in) version" from the above pull-down menu.
 You may also refer the known problems page.

Here,
 Applet Launcher (classic Java Plug-in) version: for Mac OS X or for Windows
 Applet Launcher (new Java Plug-in) version: for Windows or Linux with JRE 1.6

Parameters for Electron Density Map

Type of the map:

- contour mesh iso surface

Map position:

- atom nearest to the center of the map

Chain Residue Atom

- Atom ID:

(you can select "Atom ID" by clicking in the viewer)

- coordinates: (X , Y , Z)

mapped area: 10 A

(this is the length of edge of a cube)

contour level: 1 σ

color: R 0.0 G 1.0 B 1.0

isosurface transparency level: 0.5

Electron Density Map Download/Delete ...

service.pdbj.org/mine_test/jsp/download_map2.jsp

Electron Density Map Download/Delete Page

file format	filename	download
structure factor	r2xyysfent.gz	<input type="button" value="download"/>
refinement file	2xyy.reftar.gz	<input type="button" value="download"/>
edmap file	20111122111327_2xyy.c.xml.gz	<input type="button" value="delete"/> <input type="button" value="download"/>

- **統括責任者**
 - 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)
 - 松田 真 (大阪大学蛋白質研究所・特任研究員)
 - 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 恭代 (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 憲治 (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)
 - 西川 健 (大阪大学蛋白質研究所・客員教授)
 - 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**
 - 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・招聘教授)
 - 児島 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)
 - 高橋 あみ (大阪大学蛋白質研究所・特任研究員)
- **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)
 - 伊藤 暢聡 (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 (大阪大学免疫学フロンティア研究センター・准教授) (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・特任准教授) (for ASH)
- **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)