

蛋白質構造データバンクの 国際的な構築と統合化

(Protein Data Bank japan: PDBj, <http://pdbj.org>)

代表研究者: 中村 春木

分担研究者: 藤原 敏道

(大阪大学蛋白質研究所・教授)

1. 本研究開発課題の目的・要旨

国際協力のもと、蛋白質構造データバンク(PDB)の構築・運営を実施するとともに、NMR実験情報のデータバンク(BMRB)、電子顕微鏡やX線小角散乱による構造情報データベース(DB)を構築し公開する。オントロジー、データ管理システム、新たなデータ書式、利用者インターフェースや二次的DBおよび種々のサービスも開発して公開し、PDBの高度化・統合化を図り、人材育成も実施して、他の生命DBとの統合化による高次生命機能の理解に資する研究開発を行う。

2. 3年間の研究開発のスケジュール

	H23 年度	H24 年度	H25年度
(1) PDBデータおよびNMR実験情報の登録・編纂の国際連携による実施とデータ検証法の開発	蛋白質構造情報とNMR実験情報の登録		
	登録業務の標準化・自動化の開発		
	wwPDB活動の推進		
	電子顕微鏡およびX線小角散乱の構造情報DBの構築と公開		
	NBDCとの統合的データベース運営の仕組みを構築し実施		
(2) 統合化に向けたデータベースの高度化	構造情報のRDF形式等による高度化と配布		
	構造データのオントロジー化		
	構造データの新フォーマットの開発・改良		
	ユーザインターフェースの改良と高度化		
	NMR化学シフトデータの管理・配布システムの高度化 (STAR形式への変換)		(XML形式への変換)
	新たなNMR実験データのDB化と登録システムの開発		
(3) 機能情報と構造情報との統合化システムの開発	蛋白質相互作用部位データベースの改良とPDBデータとの統合化		
	配列情報を入力とした構造情報に基づく機能の解析・推定パイプラインの構築 (パイプライン化と公開) (GIRAF, eF-seekの組込) (システムの確立と公開)		
(4) 人材養成の実施	データ寄託者・登録者に対する教育・講習会の実施		
	データ利用者に対する教育・講習会の実施		
	キュレーター・アナテータの人材育成		

3. 本研究開発課題のH24年度の目標と当初計画

(3-1) PDBデータおよびNMR実験情報の登録・編纂の国際連携による実施とデータ検証法の開発

- ・Primary Annotatorsによる欧米との協力したデータ登録の継続的实施。
- ・登録業務の標準化・自動化(Common Annotation & Deposition)システム開発の国際協力(データ検証部分を担当)。
- ・wwPDB活動を推進。(2012年10月12日にwwPDBACを大阪で開催。wwPDBとしての教育面を重要視した第1回目のアウトリーチ活動も行う。)
- ・クラウドコンピューティング技術を用いたNMR実験データの統合的管理システムを構築。
- ・NBDCとの統合的データベース運営の仕組みを構築し実施。

(3-2) 統合化に向けたデータベースの高度化

- ・蛋白質構造データの管理・配布システムの高度化: RDF化したPDBデータをさらに高度化し、他のRDF化がなされたDBとの統合化環境を整備。
- ・wwPDBとしての新たなフォーマット開発とその記述検証法の開発(PDBj: 金城、池川)
- ・ユーザインターフェースの改良: 継続的にポータルの改良を行い、多言語でのweb page記述とグラフィックスビューアによるGUIの高度化を実施。
- ・NMR実験データの管理・配布システムの高度化: オントロジー工学に基づく標準的なXML記述法を開発し、外部DBとの統合的利用を可能として、薬剤や蛋白質間の相互作用、動力学解析データ等の登録を支援する仕組みを開発。H24年度はプロトタイプを作成。

(3-3) 機能情報と構造情報との統合化システムの開発

・配列情報を入力とした、構造情報に基づく機能の解析・推定パイプラインの構築：
ゲノム解析等によって得られた配列情報から、構造情報あるいは複合体を含むモデル
構造情報を経由し、高次生命機能情報を推定できる高速なパイプラインを改良・提供する。
H23年度に、*PreMAFFTash*, *MAFFTash*, *MSThread*, *Spanner*, *SeSAW*, *surFit* 等で構築
されたパイプラインを、複数のthreading手法(*HHPred*, *FFAS*, *FORTE*)が利用できる
ように改良して成熟させ、パイプラインによる機能推定が行えるシステムとする。

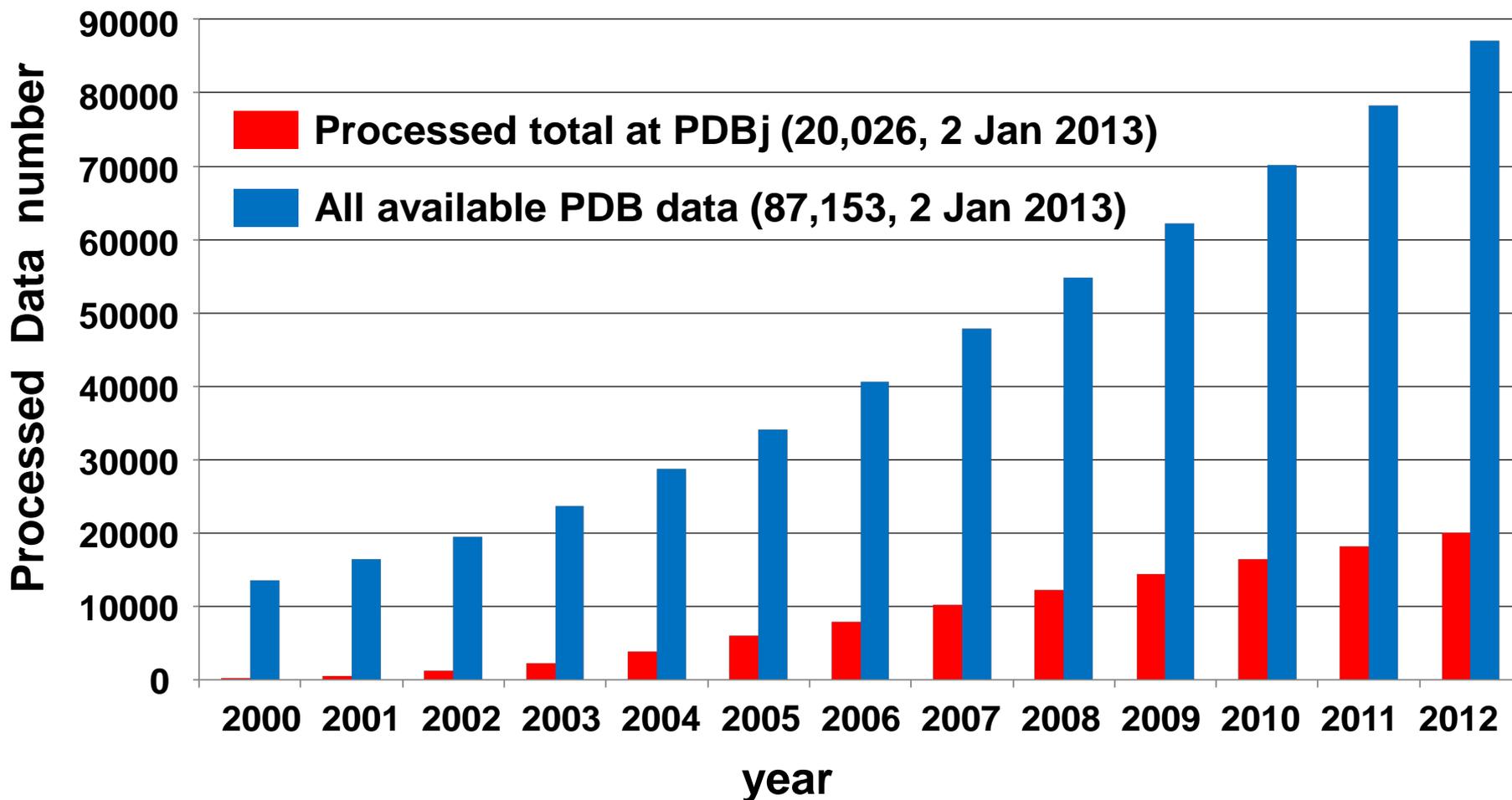
(3-4) 人材養成の継続的实施

・データ寄託・登録者に対する高精度のデータ整理と登録についての教育：構造生物学
の研究者・学生に対し、データの品質管理について講習会やセミナーで紹介する。
・データ利用者に対する初歩的および高度な利用法についての教育：初歩的な利用法
と、高度なweb serviceの利用法の双方について、講習会やセミナーで紹介する。
・キュレータ、アノテータの人材養成：国際的なon-the-job trainingを実施し、国際感覚を
身に付けたキュレータ、アノテータを育成する。

4. 本研究開発課題のH24年度の進捗状況

(4-1) PDBデータおよびNMR実験情報の登録・編纂の国際連携による実施とデータ検証法の開発

- Primary Annotatorsによる欧米との協力したデータ登録の継続的实施。



・EMデータと登録・公開と、SAXS/SANSデータの受付のタスクフォースを開始



WORLDWIDE
wwPDB
PROTEIN DATA BANK

Welcome to the Worldwide Protein Data Bank

Home | wwPDB Agreement | Statistics | FAQ | News | About Us | f | RSS

Access the PDB FTP:
RCSB PDB | PDBe | PDBj
Archive Download
Chemical Component Dictionary

Deposit Data to the PDB:
RCSB PDB | PDBe
PDBj | BMRB

Search for Structures:
RCSB PDB | PDBe
PDBj | BMRB

PDB Archive Snapshots:
RCSB PDB | PDBj

Instructions to Journals

Documentation
Format
Annotation and Policies
Remediation

Validation Reports

Workshops and Task Forces
X-ray Validation
NMR Validation
EM Validation
SAS Task Force
PDBx/mmCIF Working Group

EM Data Bank part of PDB archive

The EM Data Bank (EMDB), the primary archive for experimentally-determined maps obtained using three-dimensional electron microscopy methods, has joined the PDB archive (<ftp://ftp.wwpdb.org>), as announced previously.

The merger makes 3DEM results available in a single archive, including over 1300 electron microscopy derived maps from EMDB and 400 coordinates for EM map-derived models in PDB. It is also an essential step in the wwPDB's development of a Common Deposition & Annotation Tool that will cover all experimental methods, including hybrid methods.

With the addition of EMDB data, the physical size of the complete wwPDB archive jumps to roughly 180 GB (from its previous 130 GB). Sites that mirror the full wwPDB archive will need to increase storage capacity accordingly.

The main impact for 3DEM data users will be the change in download links for maps:

- <ftp://ftp.wwpdb.org/pub/emdb>
- <ftp://ftp.ebi.ac.uk/pub/databases/emdb>
- <ftp://ftp.pdbj.org/pub/emdb>

3DEM depositors will not be affected--map and model deposition sessions will continue at **EMDataBank**. All 3DEM map depositions will continue to receive EMDB IDs and all 3DEM coordinate entries will continue to receive PDB IDs.

The wwPDB partners are organizations that act as deposition, data processing and distribution centers for PDB data. Members are RCSB PDB (USA), PDBe (Europe) PDBj (Japan), and BMRB (USA).

The EMDB was established in 2002 by PDBe at the European Bioinformatics Institute, and is currently operated by **EMDataBank, Unified Data Resource for 3DEM**, established in 2007 as a joint effort of PDBe, RCSB PDB, and NCMI. EMDataBank will continue to serve the 3DEM and wider scientific communities with its 3DEM-specific portal for deposition and access to EM-derived structures with links to software, validation standards and conventions.

© wwPDB



SAXS/SANS VTF (held on July 12-13, 2012, at Rutgers U)

[Members]

Jill Trewhella (Chair, U of Sydney)
Dmitri Svergun (EMBL-Hamburg)
John Tainer (The Scripps Res Inst)
Wayne Hendrickson (Columbia U)
Mamoru Sato (Yokohama City U)
Torsten Schwede (U of Basel)

・登録業務の標準化・自動化 (Common Annotation & Deposition) システム開発の国際協力 (データ検証部分を担当し、他のメンバーが利用を開始)。



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

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

Options for Error messages:
 Errors for public cif (default) All errors

Category name	data type errors	mandatory errors	primary key errors	enumeration errors
Data check results				
._refine	data type: 0	mandatory: 0	primary key: 0	enumeration: 0
._struct_conf_type	data type: 0	mandatory: 0	primary key: 0	enumeration: 0
._database_PDB_rev	data type: 0	mandatory: 1	primary key: 1	enumeration: 0
._struct_biol	data type: 0	mandatory: 0	primary key: 0	enumeration: 0
._refine	data type: 1	mandatory: 0	primary key: 0	enumeration: 0

< 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]+)[X]([0-9]+D)?([eE][+-]?[0-9]+)?/?	409 data

Errors exist in the following data

data in uploaded cif file (lines: from 429 to 438)

```

#
_refine.entry_id          RCSB073713
_refine.ls_number_ref_ins_obs  121874
_refine.ls_number_ref_ins_all  121874
_refine.ndb_ls_siana_f      ?
_refine.ndb_ls_siana_F      0
_refine.ndb_data_cutoff_high_absF  ?
_refine.ndb_data_cutoff_low_absF  ?
_refine.rcsb_data_cutoff_high_rms_absF  ?
_refine.ls_d_res_low       I  47.x
    
```

■ **wwPDB活動を推進。**

wwPDBAC meeting (2012年10月12日阪大蛋白研で開催)

アウトリーチ活動 (2012年10月13日大阪梅田で開催)



13:30-14:00
中村春木 (大阪大学 蛋白質研究所) / 中村春木らによる蛋白質構造データベース (PDB) の役割

14:00-15:00
Stephen Kevin Bailey (カリフォルニア大学サンディエゴ校) / 日本への PDB データバンクのインパクト (日本語解説)

15:00-15:20 休憩

15:20-16:20
中村春木 (大阪大学 生命情報研究所) / 蛋白質を支える地球分子機構 - アノテーションを基にした構造解析の展望 -

16:20-16:30 質疑応答、総括

一般社会人・学生 (高校生以上) 向け講演会
PDB データバンク:
タンパク質のかたちが支える生命科学と創薬への応用

日時 **2012年10月13日 (土)**
13:30-16:30 (13:00 受付開始)

会場 **ハーブホール毎日新聞ビル B1**
〒530-0001 大阪市北区梅田 3-4-5

主催 **wwPDB Foundation**
(国際蛋白質構造データベース財団)

協賛 (左) 科学技術振興機構 / 大阪サイエンスデータセンター
(中) 大阪大学蛋白質研究所 / 大阪大学生命情報研究所
(右) 日本生物物理学会 / 日本蛋白質科学会 / 大阪医薬品協会

■参加無料・事前申込不要

Protein Data Bank Japan 事務局 (東京)
Tel: 03-5289-4211
Email: info@pdb.jp / info@wwpdb.org
http://pdb.org/pdb_contact.html

**wwPDB and wwPDBAC members
at IPR, Osaka U, on Oct. 12 2012**



wwPDBとしての第1回アウトリーチ活動の実施

2012年10月13日に、大阪・梅田において教育面を重要視したwwPDBアウトリーチ活動
講演者：中村春木(阪大蛋白研), S.K. Burley (UCSD), 難波啓一(阪大生命機能)

wwPDB Outreach講演会 in 大阪
(2012年10月13日(土)、梅田・ハートンホール毎日新聞ビル)

- 「生命科学における蛋白質構造データバンク(PDB)の役割」(pdf 4.8MB) Video
講師：中村 春木(大阪大学蛋白質研究所)



- 「創薬へのPDBデータバンクのインパクト」(pdf 4.2MB) [日本語版(pdf 4.3MB)] Video
講師：Stephen K. Burley(カリフォルニア大学サンディエゴ校)



NBDCとの統合的データベース運営の仕組みを構築し実施。

NBDCのデータセンターに、ポータルを設置し、運営を開始(2012年12月から本格運用)



NBDCの札幌データセンター内に設置したPDBjのフロントエンドサーバ



NBDC札幌データセンターと阪大蛋白研との冗長的運営

NBDCと阪大蛋白研とでデータベースを同期



(4-2) 統合化に向けたデータベースの高度化

PDB/RDF format for Semantic Web (wwPDBの標準仕様に採用)

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

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

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

WORLDWIDE
PDB
PROTEIN DATA BANK

Welcome to the Worldwide Protein Data Bank

PDB/RDF [About PDB/RDF](#)
[PDB/RDF , chem_comp/RDF](#)

PDB ID: (e.g., '7RSA')

property: (e.g., 'PDBo:entity.pdbx_description')

keywords: (e.g., 'alcohol')

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

In UniProt RDF:

```
<rdf:Description rdf:about="http://rdf.wwpdb.org/pdb/1BY4">
  <rdf:type rdf:resource="http://purl.uniprot.org/core/Structure_Resource"/>
  <database rdf:resource="http://purl.uniprot.org/database/PDB"/>
  <method rdf:resource="http://purl.uniprot.org/core/X-Ray_Crystallography"/>
  <resolution rdf:datatype="http://www.w3.org/2001/XMLSchema#float">2.10</resolution>
</rdf:Description>
```

PDB new format in the near future:

PDBx

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



- Preserve **backward compatibility** where possible
- Web service to **create current PDB format data**
- **Delivery target early 2013**

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

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

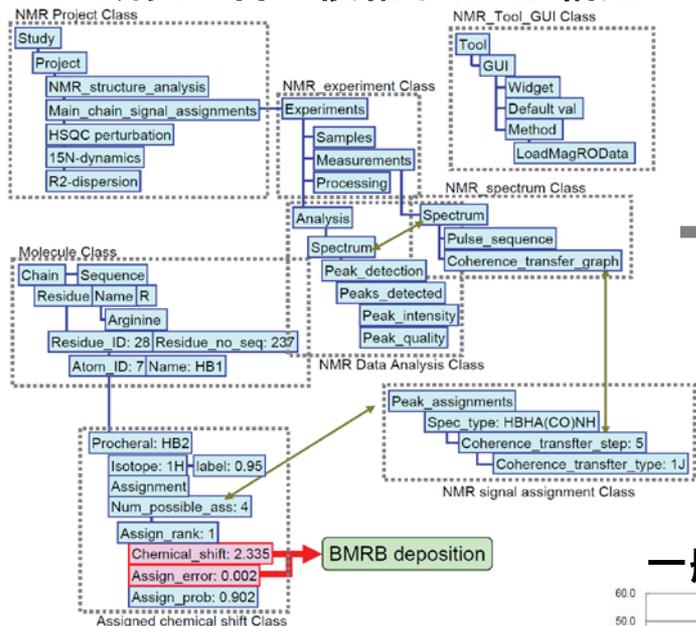
```

PDBx

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

NMRデータ構造の管理配布システムの高度化

NMR研究が持つ複雑なデータ構造



STAR形式として表現

```
#
_cyana_input.id 201110101
_cyana_input.user 'Naohiro Kobayashi'
_cyana_input.mag ?
#

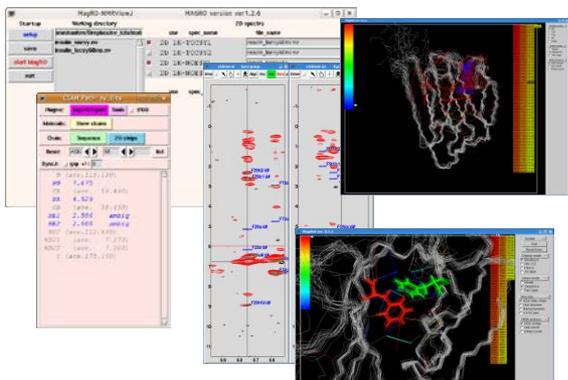
loop_
_hierachy_cyana_input.tag
_hierachy_cyana_input.item
_hierachy_cyana_input.type
_hierachy_cyana_input.item_up

cyana_input_gui_mother_top item top none
cyana_input_gui_mother_top widget norm item
cyana_input_gui_mother_top geom end widget
cyana_input_gui_mother_top color end widget
cyana_input_gui_mother_frame item top none

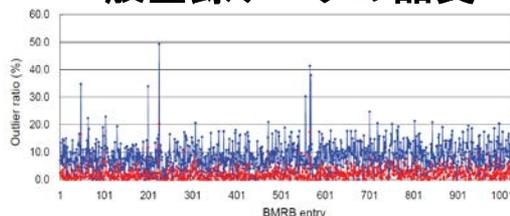
cyana_input_gui_mother_frame widget norm item
cyana_input_gui_mother_frame row end widget
cyana_input_gui_mother_frame col end widget
cyana_input_gui_mother_frame anchor end widget
cyana_input_gui_mother_frame parent end widget
#
```

RDB,
XML,
RDF化へ

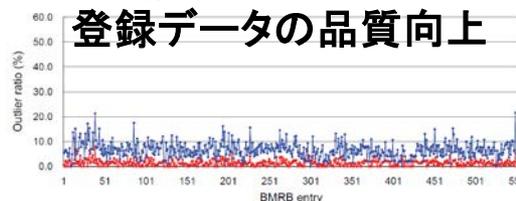
NMRデータの解析と登録支援



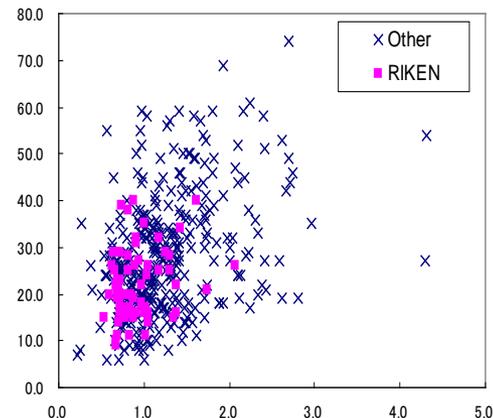
一般登録データの品質



支援システムによる 登録データの品質向上

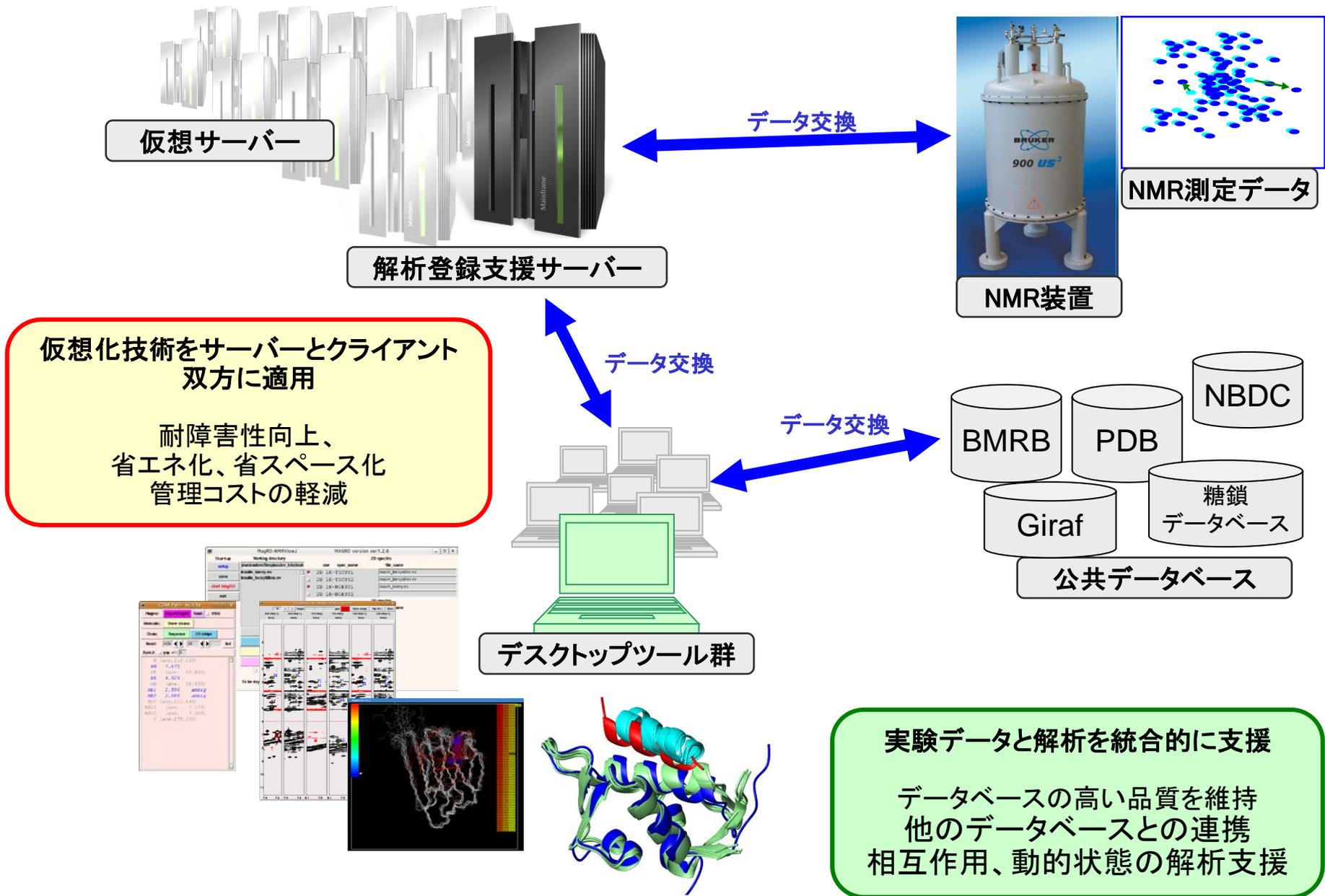


NMRデータと立体構造 の決定精度の再評価



NMR構造から逆算される化学シフトと登録された値との差 (赤:Z>3.0, 青:Z>2.0)

統合を推進する生体系NMRデータ群の網羅的登録



・新たなポータル (PDBj-BMRB) の開発(2012年12月に公開)

[English](#)



PDBj-BMRBでは生体高分子のNMR実験データの登録受付・公開をPDBjグループの一員として行っています。

BMRB Search

Retrieve entry by BMRB accession No.

 Accession number Deposition code

登録後の処理状況もAcc.No.で検索すると確認できます。

Retrieve by keywords [Google]

 BMRB Database www

Deposit Data

ADIT-NMR



生体高分子のNMR実験データの登録はこちら。

SMS-Dep



低分子生体分子のNMR実験データと原子座標フォーマット変換

BMRB Database

Osaka mirror



BMRBウェブサイトはこちら。





Contents

Dep Manual



ADIT-NMRへの登録方法の解説は
説はこちら。

FAQ



登録時によく寄せられるご質問を
FAQ形式でまとめました。

NMR-STAR Format

BMRBで使用しているNMR-STAR
ファイルフォーマット情報。

File Converter



各種データフォーマットからNMR-
STARファイルへの変換。

NMR Tool BOX

登録やデータの評価などに役立つ
ツールを公開しています。

Other Tools



その他ツールやWebブラウザで
動作するツールへのリンク。

Contact us



お問い合わせ。






Information

Dec. 6 PDBj-BMRBのポータルサイトを更新、利用が多い項目へのアクセスが簡単にできるように改善しました。
Oct. 30 NMR-ToolBOXへFit robotが追加されました。
Sep. 24 2012年9月23日(日)に行われました第50回日本生物物理学会年会におけるランチョンセミナーの発表内容を掲載しました。
[ご興味のある方は是非ご覧ください。\(外部リンク\)](#)

(4-3) 機能情報と構造情報との統合化システム (SFAS) の開発

Sequence to Function Annotation Service

<http://sysimm.ifrec.osaka-u.ac.jp/sfas2/>

Intrinsic Disorder (DISOPRED, IUPred)

Secondary Structure (PSIPred)

Domain boundaries (based on above)

Composition-based Function prediction (IDD Navigator [1-2])

Threading (HHPred, FFAS, Forte, PSI-BLAST, BLAST)

3D Modeling (Spanner [3])

Structure-based function prediction (SeSAW [4])

1. Teraguchi S, Patil A, Standley D, BMC Bioinformatics 2010, 11:S7
2. Patil A, Teraguchi S, Dinh H, Nakai K, Standley DM, Pac Symp Biocomput 2011, 17:164-175.
3. Lis M, Kim T, Sarmiento J, Kuroda D, Dinh H, Kinjo AR, Devadas S, Nakamura H, Standley DM, Immunome Research 2011, 7
4. Standley DM, Yamashita R, Kinjo AR, Toh H, Nakamura H: Bioinformatics 2010, 26:1258-1259.

Structure & Function Prediction Pipeline

This web service is currently in *BETA* testing, your [feedback](#) is important

Input Progress 2D Predictions Alignments View Structure Spanner Summary About

Query Sequence

```
>sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus GN=N4bp1 PE=1 SV=2
MAARVVLDEFTAPAEGAALLERSRGRIEALFGVGLAVLALGAEELPARIWLQLRGAQE
AVHSAKEYIKGICEPELEEKECYPKAMHCIFVGAQSLFLKSLIQDTCADLCVLDLDTGLLGI
RGSAAEAVVMARSHIQQFVKLFESNENLPSNQRESEIKREFRQFVEAHADSYTMDLLILPT
SLKKEKLLSLTQGEESLFETDDDVTIVGDVDPPEYQSAATGPSSARDEVVVQEDSRNKAR
TPVSELTKHMDTVFSSSPDVLFPVNGLSPDEDALSKDRVCHKRRSSDTEERHTKKQFSL
ENVPEGELLPDGKGSAGNEVIDLSDPASNSTNLSPDGKDTTEEMEYNILVNFCKTMGYSQ
EIVEKVIREYGPSTEPLLLLLEEIEKENKRLQEDRDFPCTVYPDASQSRNAGVGSSTNEL
TADSTPKKAQSHTQSMVERFSQLPFKSKHCTSNCKVNSFRTVPVGQKQEIWGSKQNSS
CTVDLETDGHSASAASASPKDISFVSRGASGHQQRNPAFPENGFQQQTEPLLPNNTKPAC
EKRSKSSPQPKPNYPPLSPPLPQLLPVTEARLGGSSDHIDSSVTGVQRFRTDLKI
PYKLELKNPGRADLKHIVIDGSNVAITHGLKFFSCRGIAIAVEYFWKLGNRNITVFPV
QWRTRRDPNITEQHF LTQLQELGILSLTPARMVFGERIASSHDRFLHLADKTGGIIVTN
```

Job id (optional)

Action **2D Prediction**

If your query contains multiple domains, or if you are unsure of the domain boundaries, we recommend performing a **2D prediction** at the first step

2D prediction

- IUPred disorder domain prediction
- disopred2 disorder domain prediction
- PSIPRED secondary structure prediction
- Jpred 3 secondary structure prediction

Run

2D View

Intrinsic Disorder

Predicted Secondary Structure

Structure & Function Prediction Pipeline

This web service is currently in "BETA" testing, your [feedback](#) is important

Input Progress **2D Predictions** Alignments View Structure Designer Summary About

Available Results [sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus CN=N4bp1 PE=1 SV=2][disopred] disordered domain prediction

[sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus CN=N4bp1 PE=1 SV=2][disopred] disordered domain prediction

Disorder domain prediction score and secondary structure prediction
Select your sequence range by dragging your mouse over the graph, you can also drag the borders to re-adjust your selection

Quick selection (according to disorder prediction data / click to select)

Overview

Range [615 - 778]

Selected Sequence

LKHIVIDGSNVAITHGLKKFFSCRGIAIAVEYFVKLGRRNITVFPVQWRTRRDPNI TEQHFLTQLQELGILSLTPARMVFERIASHDDRFLHLADKTTGGII

submit your query to

Sequence alignment

Alignment methods

- Hhpred
- Itas
- forte
- psi-blast
- blast

[Feedback](#)

Close-up view of selected region

Global view of selected region

Threading Methods

2D Alignment View

Select an alignment to view or build

Structure & Function Prediction Pipeline

This web service is currently in *BETA* testing, your [feedback](#) is important

Input Progress 2D Predictions **Alignments** View Structure Spanner Summary About

[sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus GN=N4bp1 PE=1 SV=2:615-778] HHPred Alignment calculation

Score	Template	%id	Align view	Actions
41.6	1bgx:T	23.36	LKHIVIDGSNVAITHGLKFFS... RVLVLDGHLAYFHALKGLT... SICLIDFSQIALSTALVRHL... NLMIVDGTNLGFKHNSK... RKVAIDASMSIYQPEGETT... KIIVIDGYNALYQFNGRIT... KRVSIDGYNALYQFGRVTS... TESEGDPFLIEAAEKYGC... VTPEETIVTVG-----G... IVKDNVVLCSG-----G... ILVDVQNVYITCREAYRSN... VVSAKAYAIASNDPKQRQ... DEKLLFLARLEAALVTDH... YGADNIVLTNG-----G... QVVAVDTYCWLHKKRGEPT... LKAENVTVTTG-----A... AEERGVLAERSKEMGAKAF... IDSQTVVYGPS-----V... VIYMSELIRQWSETGEG-V... VHTPAYDAFYDGNFCMDG... DLWIAVAASRALPVITQD... DDFAA... SYTDAISEVVAEELKKLIS... SRFSL... SIPDLLIAATAEISGLT... KDFDA... ADPKTEIMVLLG-----... ANQAFMLGLSAFLKDGEE... VLIPTPAFVSYADEFRNL... VDELKKYVTDKALINSPC... NPTGAULTKKLLEEIADF... VVEHDLVIVSDEVYEHFY... DDA... FLTDDLKVTWDAKKGIMR... FMTAFPYYEMEGRLPDHG... EPVDLQARGILLDGGSV... EGDKRLLQI... SKADIEVLALAYELKGEI... FSDDY... LNTGDIVETVRTMRAAGVQ... FLDTPDSYYDTGEWVGDT... RVPVDTLRELKILA... FNSNDAVRAVKALSERGVE... FLKTPGAYYDLGERITIL... QTHSLDLDRATNVLA... IQTDWIINTAG-----... VVPVAVNAVREFTKPGDG... VIIITPVYYPFFGYTIDF... QKLERLSKDRALLFCSP... HNPVGRVWKKDELQIKI... KDIVLKSIDLMLWSDEIH... FDLIMPGY... VKPEQVLVSRG-----... ADEGIELLIRAFCEPKDA... ILYCFPPYGYMSDNWQLD... LQGISDKLDGKVVVYVCS... PNPTGQLINPQDFRTLE... LTRGKAVVADEAYIEFC... PQLA... GPFQMIAGHARSRGLIV... TNN... GNDLLIASHAIAENATLV... TNN... NNDYICPHYATYAG... QYALYEDL... SUSABE... LACMXYK... ITNGYQ... TAEQ... RQW... MGG... GQ... Q... M... NE... Y... E... M... N... M...	View Build
40.2	3h7i:A	16.82	...	View Build
38.3	1exn:A	14.02	...	View Build
35.4	3q8k:A	15.09	...	View Build
34.9	3ory:A	15.09	...	View Build
34.7	3i8o:A	17.65	...	View Build
34.5	2izo:A	15.09	...	View Build
33.2	1o4w:A	40.00	...	View Build
33.1	1b5p:A	9.65	...	View Build
33.0	1bw0:A	15.45	...	View Build
32.6	2qip:A	15.38	...	View Build
32.3	3ix7:A	45.00	...	View Build
31.9	1j32:A	11.94	...	View Build
31.9	3qe9:Y	15.09	...	View Build
31.2	3b46:A	12.69	...	View Build
31.1	3qi7:A	7.95	...	View Build
30.9	1d2f:A	11.82	...	View Build
30.7	3dbo:B	19.23	...	View Build
30.6	1w8i:A	17.24	...	View Build
30.0	3h87:A	20.69	...	View Build
29.7	1qd9:A	10.43	...	View Build
29.6	1c1j:A	20.00	...	View Build
29.4	2lca:A	34.78	...	View Build
29.2	1t47:A	15.38	...	View Build
28.8	2r5w:A	15.38	...	View Build
28.8	1c7n:A	12.39	...	View Build
28.4	1fg7:A	12.69	...	View Build
28.2	3tnd:A	36.36	...	View Build
28.1	3zvk:A	22.73	...	View Build
28.1	1yje:A	12.79	...	View Build

Results : 41

[feedback](#)

3D Alignment View

Structure & Function Prediction Pipeline

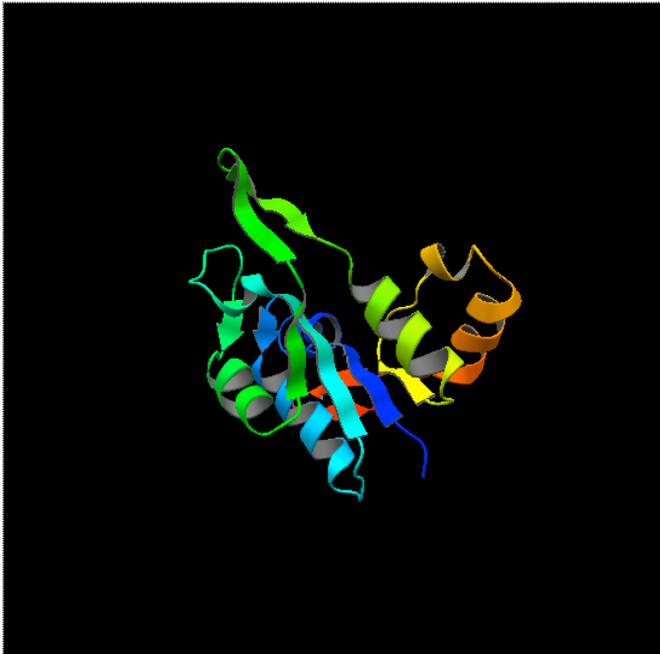
Build the model with Spanner

This web service is currently in *BETA* testing, your [feedback](#) is important

[Input](#) [Progress](#) [2D Predictions](#) [Alignments](#) [View Structure](#) [Spanner](#) [Summary](#) [About](#)

[[sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus GN=N4bp1 PE=1 SV=2:615-778] FFAS Alignment calculation][[3v33_A][Alignment]

[switch to Preview](#) [Build model with Spanner](#)



[Download template file](#)

```
>3v33A
LRPVVIDGSNVAMSHGNKEVFSRCRGILLAVNWFLERGHTDITVFVPSWRKEQPRPDVPI TDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIVKLAYE
SDGIVVSNDTYRDLQGERQEWKRFIEERLLMYSFVNDKFMP-----
>sp|Q6A037|N4BP1_MOUSE
LKHIVIDGSNVAITHGLKKFFSCRGIAIAVEYFWKLGNRNITVFVPQWRTRRDP---
NITEQHFLTQLQELGILSLTPARMVFGERIAASHDDRFLLHLADKGGIIVTNDNFREFVTEVSWSREIITKRLLQYTFVGDIFMVPDDPLGRNGPRLEEF
RKEA
```

Query residues mapped on Template

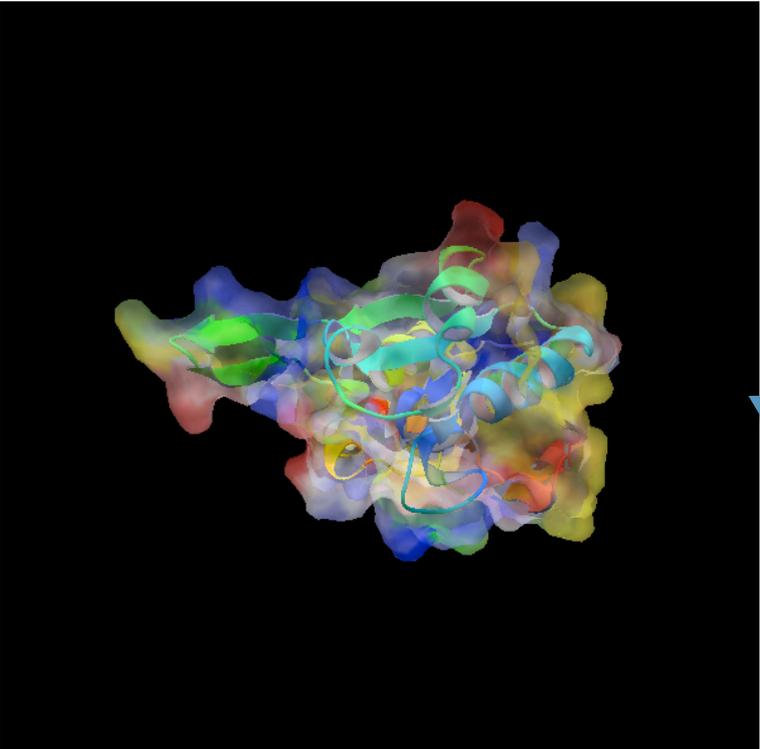
3D Model View

Structure & Function Prediction Pipeline

This web service is currently in *BETA* testing, your [feedback](#) is important

[Input](#) [Progress](#) [2D Predictions](#) [Alignments](#) [View Structure](#) **[Spanner](#)** [Summary](#) [About](#)

[sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus GN=N4bp1 PE=1 SV=2][ffas]
[3v33A] Model Creation



Spanner model with electrostatic potential mapped on molecular surface

[Close Interactive Applet](#)

Download

- [alignment](#)
- [template pdb](#)
- [output](#)
- [computed electrostatic surface](#)

[Submit to SeSAW](#)

Structure & Function Prediction Pipeline

This web service is currently in *BETA* testing, your [feedback](#) is important

[Input](#)
[Progress](#)
[2D Predictions](#)
[Alignments](#)
[View Structure](#)
[Spanner](#)
[SeSAW](#)
[Summary](#)
[About](#)

[616-779] [sp|Q6A037|N4BP1_MOUSE NEDD4-binding protein 1 OS=Mus musculus GN=N4bp1 PE= load

template	score	identity	coverage	domain	alignment	superposition
3v33A_PDP_A	617.942	49	100	-	view	view
3v33B_PDP_A	617.614	49	99	-	view	view
3v34B_PDP_A	591.746	49	96	-	view	view
3v34A_PDP_A	588.199	49	96	-	view	view
3v32B_PDP_A	583.877	50	95	-	view	view
3v32A_PDP_A	583.841	50	95	-	view	view
1tauA_SCOP_B	114.599	17	73	SCOP:c.120.1.2	view	view
1taqA_SCOP_B	92.843	16	72	SCOP:c.120.1.2	view	view
1xo1A_SCOP_B	90.241	9	70	SCOP:c.120.1.2	view	view
1xo1B_SCOP_B	90.005	9	71	SCOP:c.120.1.2	view	view
1o4wA_SCOP_A	86.553	19	66	SCOP:c.120.1.1	view	view
3r5hB_PDP_A	77.799	17	54	-	view	view
1ut5B_SCOP_B	75.760	10	71	SCOP:c.120.1.2	view	view
2goyC_PDP_A	75.568	20	59	-	view	view
1exnA_SCOP_B	74.332	11	70	SCOP:c.120.1.2	view	view
3dfzB_PDP_A	73.891	14	56	-	view	view
3fcaA_PDP_B	71.523	19	52	-	view	view
3qffA_PDP_A	71.137	17	56	-	view	view
1ut5A_SCOP_B	71.133	9	69	SCOP:c.120.1.2	view	view

(4-4) 人材養成の実施

- ・データ寄託・登録者に対する高精度のデータ整理と登録についての教育: 構造生物学の研究者・学生に対し、データの品質管理について講習会やセミナーで紹介する。

2012年日本結晶学会年会 ランチョンセミナー

(2012年10月26日(金)、**東北大学 片平キャンパス**)

「PDBj (PDB Japan) の活動とwwPDB」 講師: 中村 春木

「PDB (Protein Data Bank) への登録について」 講師: 中川 敦史

- ・データ利用者に対する初歩的および高度な利用法についての教育: 初歩的な利用法と、高度なweb serviceの利用法の双方について、講習会やセミナーで紹介する。

第12回日本蛋白質科学会年会ランチョンセミナー

(2012年6月20日(水)、**名古屋国際会議場**)

「PDBj のデータベース高度化・統合化について」 講師: 中村 春木

「PDBj のウェブサービス」 講師: 金城 玲

第50回生物物理学会年会 ランチョンセミナー

(2012年9月23日(日)、**名古屋大学東山キャンパス**)

「Big Data時代へ向けたwwPDBとPDBjの役割」 講師: 中村 春木

「仮想化技術を用いたNMRデータの統合的解析環境」 講師: 小林 直宏

5. 本研究開発課題のH25年度の実施計画

(5-1) PDBデータおよびNMR実験情報の登録・編纂の国際連携による実施とデータ検証法の開発

- ・Primary Annotatorsによる欧米との協力したデータ登録の継続的实施。
- ・登録業務の標準化・自動化 (Common Annotation & Deposition) システム開発の国際協力 (データ検証部分を担当)。
- ・wwPDB活動を推進。(2013年秋にwwPDBACを米国で開催。)
- ・クラウドコンピューティング技術を用いたNMR実験データの統合的管理システムを構築。
- ・NBDCとの統合的データベース運営の仕組みを継続的に実施。

(5-2) 統合化に向けたデータベースの高度化

- ・蛋白質構造データの管理・配布システムの高度化: RDF化したPDBデータをさらに高度化し、他のRDF化がなされたDBとの統合化環境を整備。
- ・wwPDBとしての新たなフォーマット開発とその記述検証法の開発
- ・ユーザインターフェースの改良: 継続的にポータルの改良を行い、多言語でのweb page記述とグラフィックスビューアによるGUIの高度化を実施。
- ・NMR実験データの管理・配布システムの高度化: オントロジー工学に基づく標準的なXML記述法を開発し、外部DBとの統合的利用を可能として、薬剤や蛋白質間の相互作用、動力学解析データ等の登録を支援する仕組みを開発。H24年度はプロトタイプを作成。

(5-3) 機能情報と構造情報との統合化システムの開発

- ・**蛋白質相互作用部位データベース**の改良とPDBデータとの統合化: 蛋白質構造解析用の種々のサービス・ツールに対してもSemantic Web技術を適用したサービスとし、原子レベル→残基レベル→分子レベル→分子複合体レベルへつなげて、高次生命機能情報へ、原子レベルの情報を統合化できる環境を整える。
- ・**配列情報を入力とした、構造情報に基づく機能の解析・推定パイプライン**の構築: ゲノム解析等によって得られた配列情報から、構造情報あるいは複合体を含むモデル構造情報を経由し、高次生命機能情報を推定できる高速なパイプラインを改良・提供する。

(5-4) 人材養成の継続的实施

- ・**データ寄託・登録者に対する高精度のデータ整理と登録についての教育**: 構造生物学の研究者・学生に対し、データの品質管理について講習会やセミナーで紹介する。
- ・**データ利用者に対する初歩的および高度な利用法についての教育**: 初歩的な利用法と、高度なweb serviceの利用法の双方について、講習会やセミナーで紹介する。
- ・**キュレータ、アノテータの人材養成**: 国際的なon-the-job trainingを実施し、国際感覚を身に付けたキュレータ、アノテータを育成する。

5. 実施体制

