IPR seminar, 2016-02-19

Search and view 3DEM structures in databanks

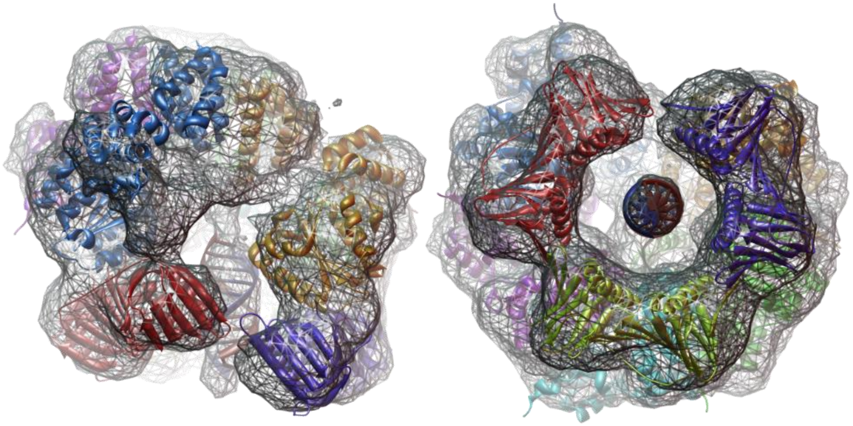
Hirofumi Suzuki
PDBj / IPR, Osaka univ.

I am ...

Once upon a time...

- Namba-lab@IIAR, Morikawa-lab@BERI, etc.
- Single particle analyses research

Flagellar motor base
2004, Suzuki *et al.*



DNA clamp loading
2005, Miyata *et al.*

Present:

- in PDBj
- EM Navigator, Yorodumi, Omokage search ...

We are …

PDBj

- @IPR, Osaka-univ.
- a member of wwPDB

wwPDB

- = RCSB-PDB + PDBe + PDBj + BMRB
- runs PDB
- helps EMDB & BMRB

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-NBDC and Institute for Protein Research IPR, Osaka University.

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.

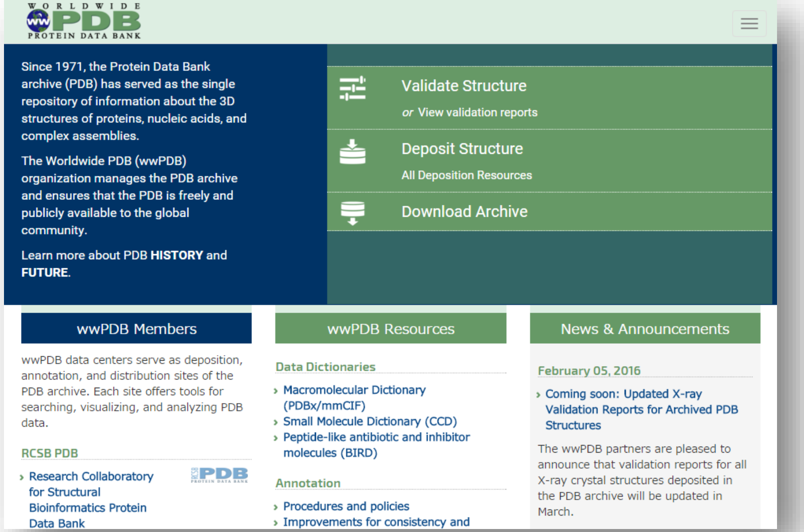
To further improve our **new web page**, any comments and suggestions are welcome. The legacy PDBj website continues to be available at http://legacy.pdbj.org/.

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.

http://pdbj.org



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 data centers serve as deposition, annotation, and distribution sites of the PDB archive. Each site offers tools for searching, visualizing, and analyzing PDB data.

- Research Collaboratory for Structural Bioinformatics Protein Data Bank

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

- Procedures and policies
- Improvements for consistency and

February 05, 2016

- Coming soon: Updated X-ray Validation Reports for Archived PDB Structures

The wwPDB partners are pleased to announce that validation reports for all X-ray crystal structures deposited in the PDB archive will be updated in March.

http://wwpdb.org

Questions

Q1: Are you using EM?

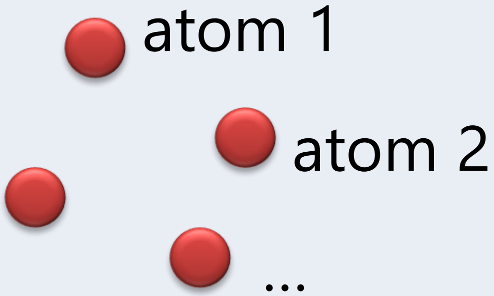
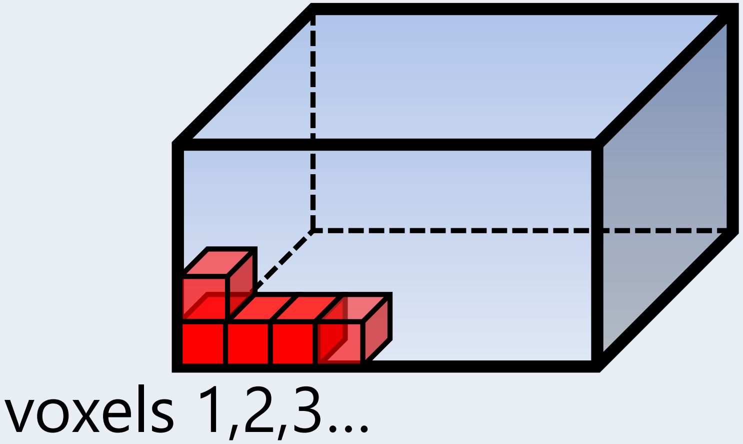
(including EM sample, EM data, etc)

Questions

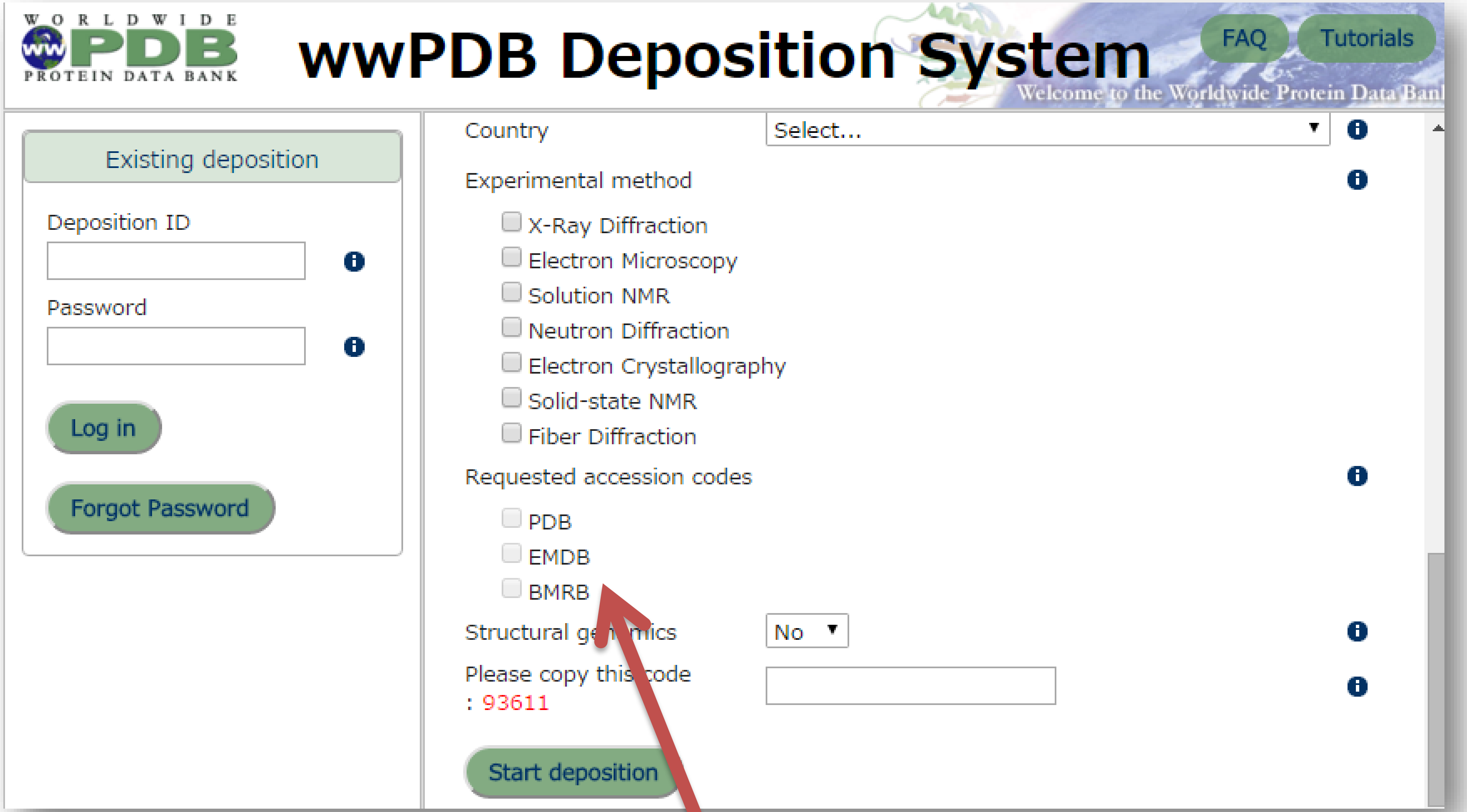
Q2: Do you know “EM Navigator”?

PDB & EMDB

PDB & EMDB

	PDB	EMDB
method	various	EM
main data	atomic model 	3D map（≒ED map） 
num. of entries	~110,000	~4,000

New Deposition System



Deposition ID

Password

Country: Select...

Experimental method

- ☐ X-Ray Diffraction
- ☐ Electron Microscopy
- ☐ Solution NMR
- ☐ Neutron Diffraction
- ☐ Electron Crystallography
- ☐ Solid-state NMR
- ☐ Fiber Diffraction

Requested accession codes

- ☐ PDB
- ☐ EMDB
- ☐ BMRB

Structural genomics: No

Please copy this code : 93611

EMDB/BMRB checkbox

EMDataBank.org

2016-02-17 : 3468 EMDB map entries, 1025 PDB coordinate entries

Unified Data Resource for 3-Dimensional Electron Microscopy

EMDataBank is a unified global portal for deposition and retrieval of 3DEM density maps, atomic models, and associated metadata, as well as a resource for news, events, software tools, data standards, validation methods for the 3DEM community.

For up-to-date information about map and model challenges, visit challenges.emdatabank.org.

EMD-2817

Nov. 14, 2014 RELEASED ON Feb. 17, 2016 singleParticle 24.0 Å NEW

Electron cryoEM structure of lactococcal siphophage 1358 virion

Spinelli S, Bebeacua C, Orlov I,

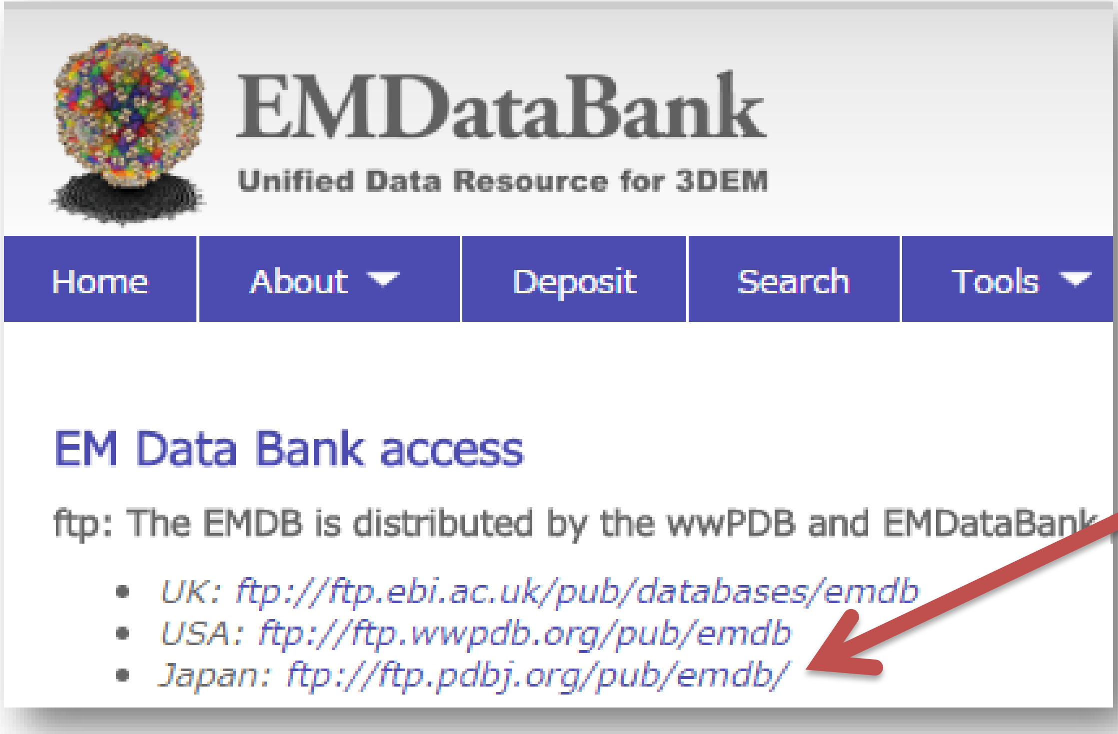
wwPDB Deposition & Ann
System Now Available fo
Structures

The wwPDB partners and
EMDataBank project are p
announce the launch of a
Deposition & Annotation s
supports structures detern
3DEM, NMR, and X-ray, ne
electron crystallography.

New Publications

New open access articles a
EMDataBank and EMDB ac
now available online, in ac
publication in the upcomin
2016 Nucleic Acids Resear
Database Issue.

PDBj contribution for EMDB



EMDataBank

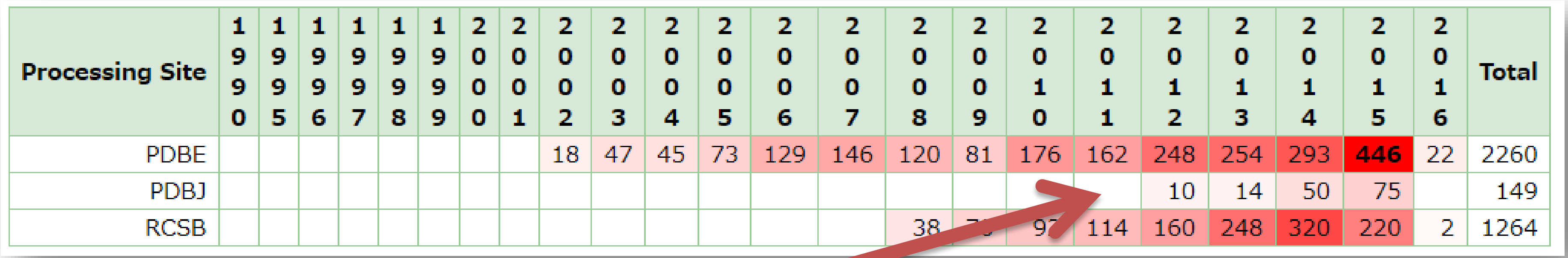
Unified Data Resource for 3DEM

EM Data Bank access

ftp: The EMDB is distributed by the wwPDB and EMDataBank

- *UK: ftp://ftp.ebi.ac.uk/pub/databases/emdb*
- *USA: ftp://ftp.wwpdb.org/pub/emdb*
- *Japan: ftp://ftp.pdbj.org/pub/emdb/*

Japanese server

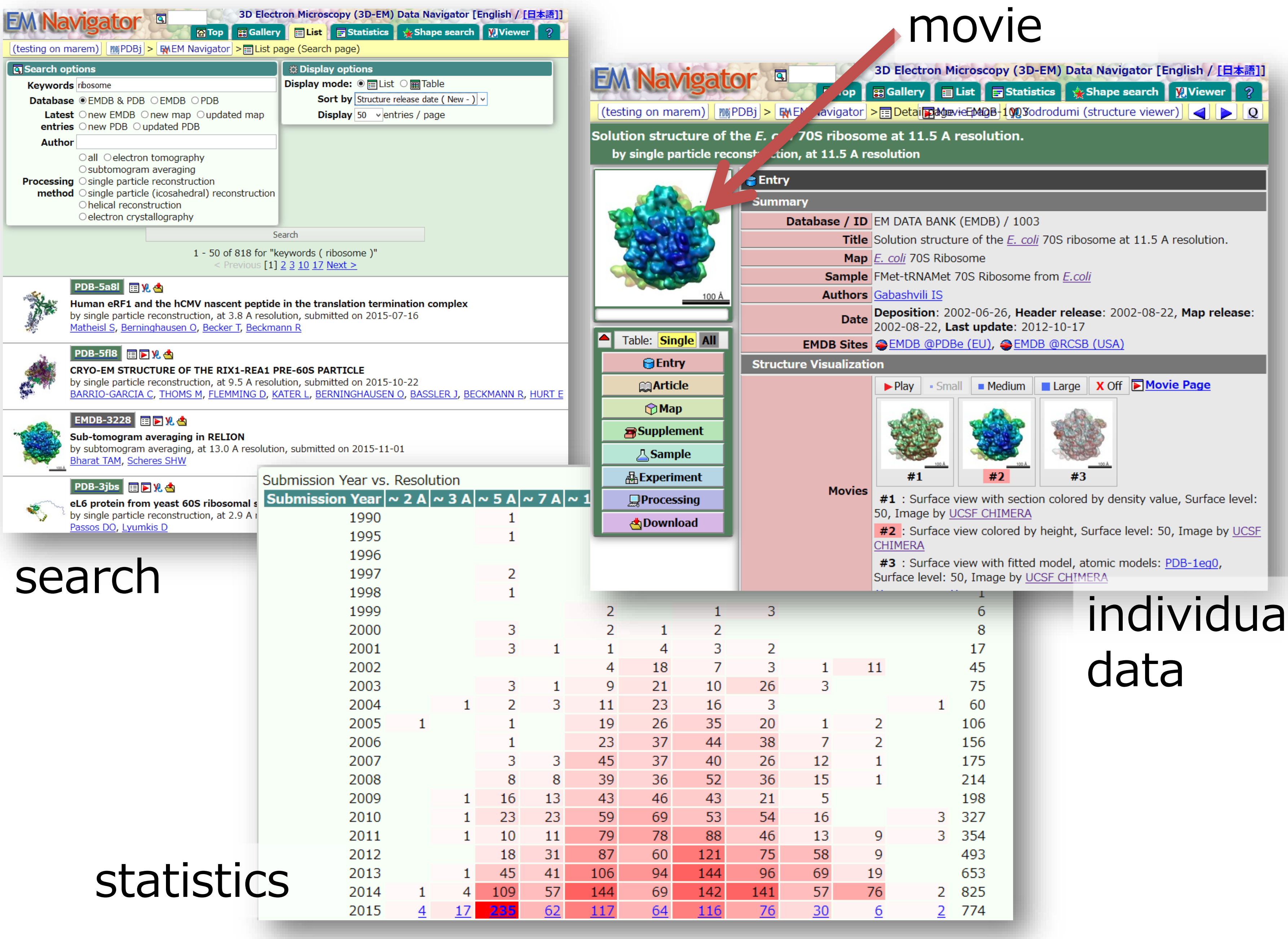


Processing Site	**1990**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**Total**
PDBE									18	47	45	73	129	146	120	81	176	162	248	254	293	**446**	22	2260
PDBJ																			10	14	50	75		149
RCSB															38	[illegible]	[illegible]	114	160	248	320	220	2	1264

data annotated by Japanese staff

New EM Navigator & Yorodumi

EM Navigator



movie

EM Navigator

3D Electron Microscopy (3D-EM) Data Navigator [English / [日本語]]

(testing on marem) PDBj > EM Navigator > List page (Search page)

Search options

Keywords ribosome

Database ● EMDB & PDB ○ EMDB ○ PDB

Latest entries ○ new EMDB ○ new map ○ updated map ○ new PDB ○ updated PDB

Author

Processing method ○ all ○ electron tomography ○ subtomogram averaging ○ single particle reconstruction ○ single particle (icosahedral) reconstruction ○ helical reconstruction ○ electron crystallography

Display options

Display mode: ● List ○ Table

Sort by Structure release date (New -)

Display 50 entries / page

1 - 50 of 818 for "keywords (ribosome)"

< Previous [1] 2 3 10 17 Next >

PDB-5a8l

Human eRF1 and the hCMV nascent peptide in the translation termination complex
by single particle reconstruction, at 3.8 A resolution, submitted on 2015-07-16
Matheisl S, Berninghausen O, Becker T, Beckmann R

PDB-5fl8

CRYO-EM STRUCTURE OF THE RIX1-REA1 PRE-60S PARTICLE
by single particle reconstruction, at 9.5 A resolution, submitted on 2015-10-22
BARRIO-GARCIA C, THOMS M, FLEMMING D, KATER L, BERNINGHAUSEN O, BASSLER J, BECKMANN R, HURT E

EMDB-3228

Sub-tomogram averaging in RELION
by subtomogram averaging, at 13.0 A resolution, submitted on 2015-11-01
Bharat TAM, Scheres SHW

PDB-3jbs

eL6 protein from yeast 60S ribosomal s
by single particle reconstruction, at 2.9 A r
Passos DO, Lyumkis D

search

EM Navigator

3D Electron Microscopy (3D-EM) Data Navigator [English / [日本語]]

(testing on marem) PDBj > EM Navigator > Detail page - EMDB-1003 Movie page Yodrodumi (structure viewer)

Solution structure of the *E. c* 70S ribosome at 11.5 A resolution.
by single particle reconstruction, at 11.5 A resolution

Table: **Single** **All**

Entry

Summary

Database / ID	EM DATA BANK (EMDB) / 1003
Title	Solution structure of the *E. coli* 70S ribosome at 11.5 A resolution.
Map	*E. coli* 70S Ribosome
Sample	FMet-tRNAMet 70S Ribosome from *E.coli*
Authors	Gabashvili IS
Date	**Deposition**: 2002-06-26, **Header release**: 2002-08-22, **Map release**: 2002-08-22, **Last update**: 2012-10-17
EMDB Sites	EMDB @PDBe (EU), EMDB @RCSB (USA)

Structure Visualization

Movies

#1 : Surface view with section colored by density value, Surface level: 50, Image by UCSF CHIMERA

#2 : Surface view colored by height, Surface level: 50, Image by UCSF CHIMERA

#3 : Surface view with fitted model, atomic models: PDB-1eg0, Surface level: 50, Image by UCSF CHIMERA

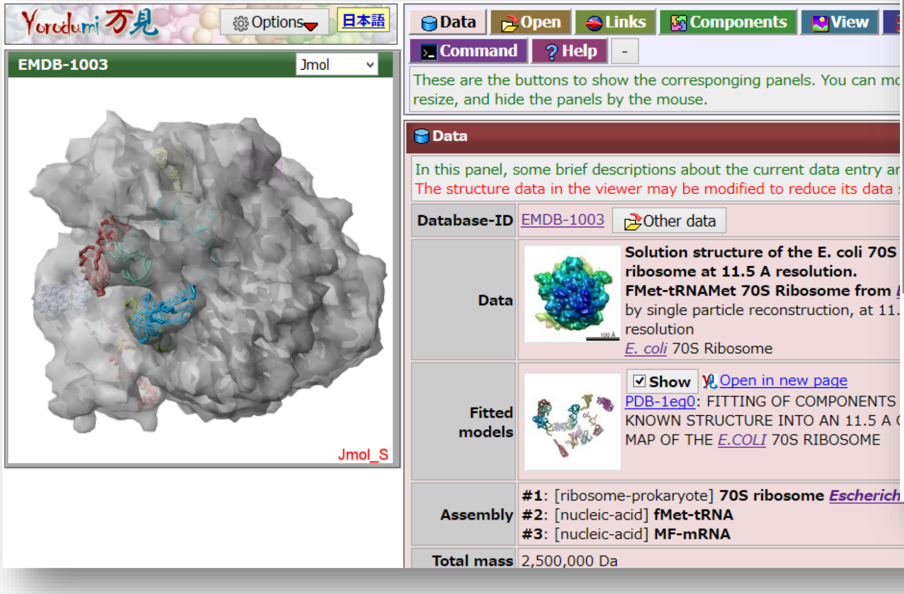
individual data

Submission Year vs. Resolution

Submission Year	~ 2 A	~ 3 A	~ 5 A	~ 7 A	~ 1							
1990			1									
1995			1									
1996												
1997			2									
1998			1									1
1999					2		1	3				6
2000			3		2	1	2					8
2001			3	1	1	4	3	2				17
2002					4	18	7	3	1	11		45
2003			3	1	9	21	10	26	3			75
2004		1	2	3	11	23	16	3			1	60
2005	1		1		19	26	35	20	1	2		106
2006			1		23	37	44	38	7	2		156
2007			3	3	45	37	40	26	12	1		175
2008			8	8	39	36	52	36	15	1		214
2009		1	16	13	43	46	43	21	5			198
2010		1	23	23	59	69	53	54	16		3	327
2011		1	10	11	79	78	88	46	13	9	3	354
2012			18	31	87	60	121	75	58	9		493
2013		1	45	41	106	94	144	96	69	19		653
2014	1	4	109	57	144	69	142	141	57	76	2	825
2015	4	17	235	62	117	64	116	76	30	6	2	774

statistics

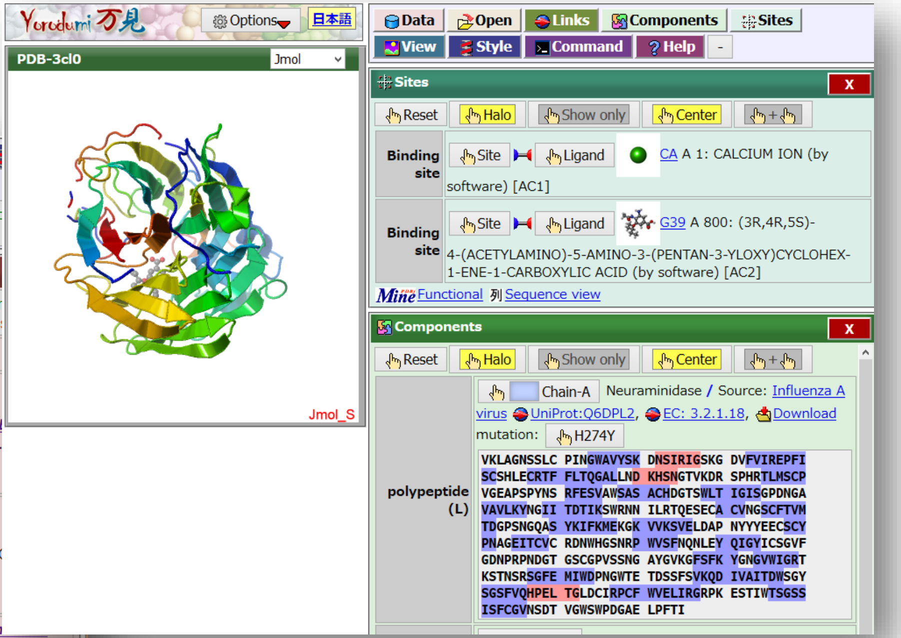
Yorodumi



These are the buttons to show the corresponging panels. You can mo
resize, and hide the panels by the mouse.

In this panel, some brief descriptions about the current data entry ar
The structure data in the viewer may be modified to reduce its data

Database-ID	EMDB-1003 Other data
Data	**Solution structure of the E. coli 70S ribosome at 11.5 A resolution. FMet-tRNAMet 70S Ribosome from** by single particle reconstruction, at 11. resolution *E. coli* 70S Ribosome
Fitted models	☑ **Show** Open in new page PDB-1eg0: FITTING OF COMPONENTS KNOWN STRUCTURE INTO AN 11.5 A MAP OF THE *E.COLI* 70S RIBOSOME
Assembly	**#1**: [ribosome-prokaryote] **70S ribosome *Escherich*** **#2**: [nucleic-acid] **fMet-tRNA** **#3**: [nucleic-acid] **MF-mRNA**
Total mass	2,500,000 Da



Binding site	Site Ligand CA A 1: CALCIUM ION (by software) [AC1]
Binding site	Site Ligand G39 A 800: (3R,4R,5S)-4-(ACETYLAMINO)-5-AMINO-3-(PENTAN-3-YLOXY)CYCLOHEX-1-ENE-1-CARBOXYLIC ACID (by software) [AC2]

Mine Functional 列 Sequence view

polypeptide (L)

Chain-A Neuraminidase / Source: Influenza A virus UniProt:Q6DPL2, EC: 3.2.1.18, Download

mutation: H274Y

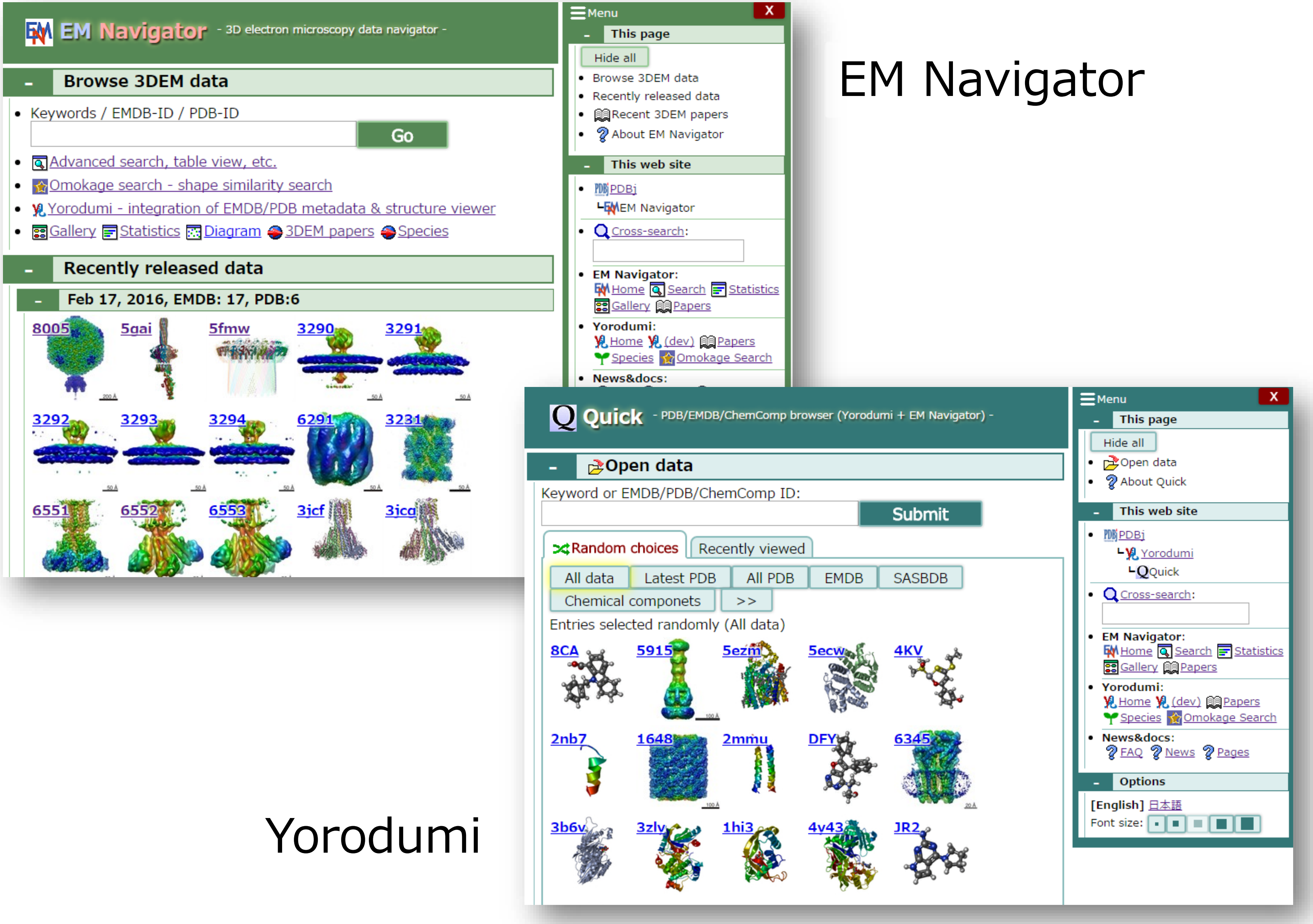
```
VKLAGNSSLC PINGWAVYSK DNSIRIGSKG DVFVIREPFI
SCSHLECRTF FLTQGALLND KHSNGTVKDR SPHRTLMSCP
VGEAPSPYNS RFESVAWSAS ACHDGTSWLT IGISGPDNGA
VAVLKYNGII TDTIKSWRNN ILRTQESECA CVNGSCFTVM
TDGPSNGQAS YKIFKMEKGK VVKSVELDAP NYYYEECSCY
PNAGEITCVC RDNWHGSNRP WVSFNQNLEY QIGYICSGVF
GDNPRPNDGT GSCGPVSSNG AYGVKGFSFK YGNGVWIGRT
KSTNSRSGFE MIWDPNGWTE TDSSFSVKQD IVAITDWSGY
SGSFVQHPEL TGLDCIRPCF WVELIRGRPK ESTIWTSGSS
ISFCGVNSDT VGWSWPDGAE LPFTI
```

EMDB map + PDB model

active site & entity information

- EM map + atomic model
- Sites (ligand binding, mutation, etc)
- Biological assembly, crystal packing, NMR ensemble & many others

New versions



EM Navigator

EM Navigator - 3D electron microscopy data navigator -

Browse 3DEM data

- Keywords / EMDB-ID / PDB-ID
- Advanced search, table view, etc.
- Omokage search - shape similarity search
- Yorodumi - integration of EMDB/PDB metadata & structure viewer
- Gallery Statistics Diagram 3DEM papers Species

Recently released data

Feb 17, 2016, EMDB: 17, PDB:6

Quick - PDB/EMDB/ChemComp browser (Yorodumi + EM Navigator) -

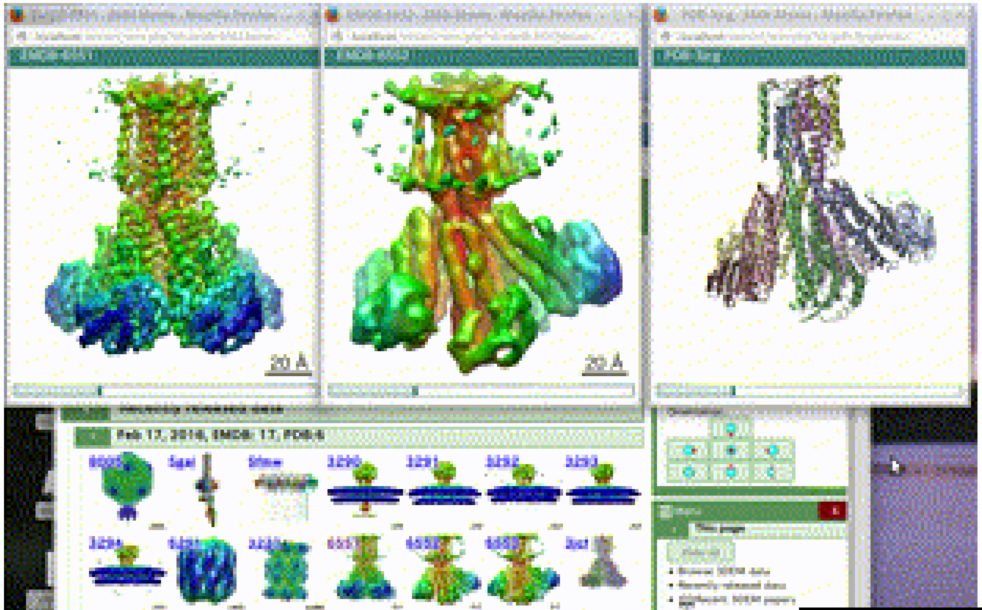
Open data

Keyword or EMDB/PDB/ChemComp ID:

Entries selected randomly (All data)

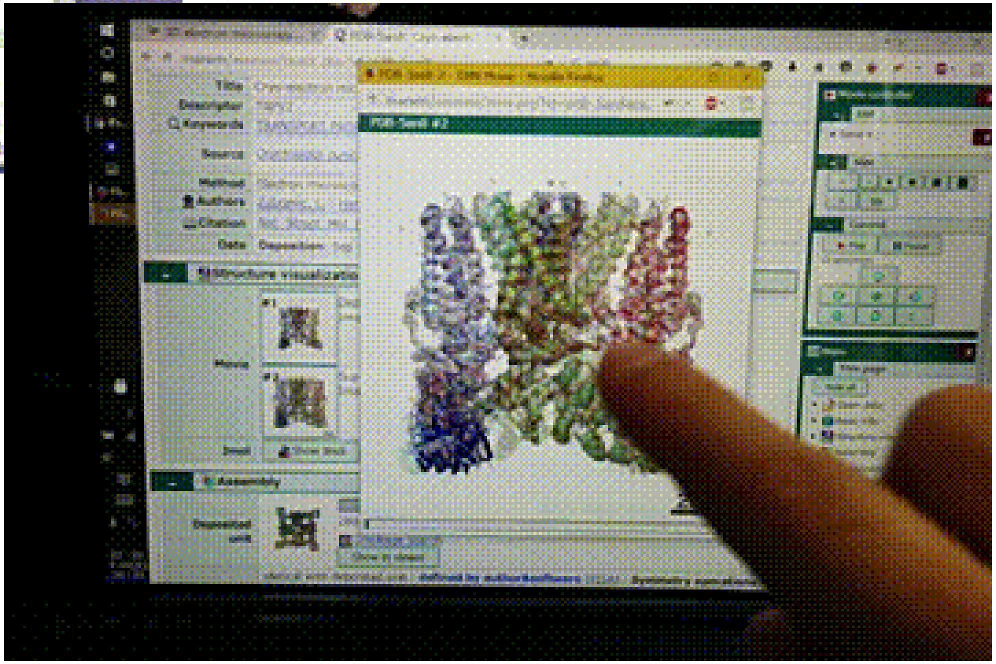
Yorodumi

New movie interface



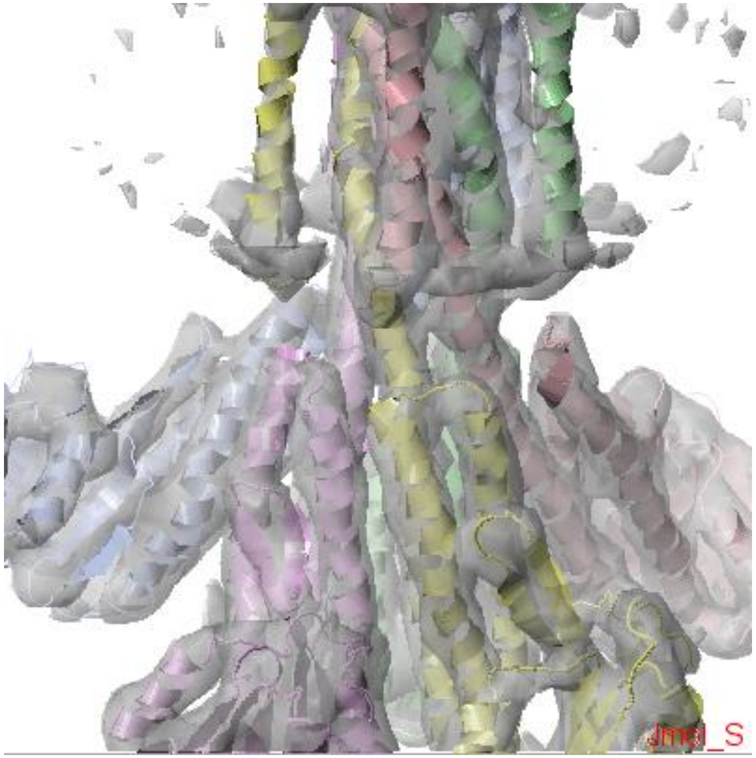
Rotation by mouse

Tiled & synchronized
multiple movies



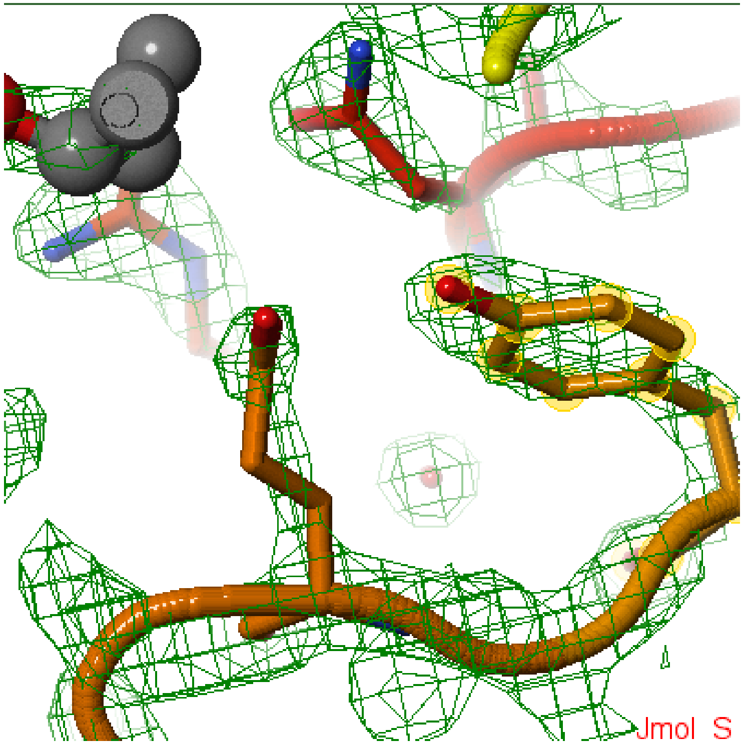
Rotation by finger
on a touch device

Molecular viewer interface



Simplified EM Map with PDB model

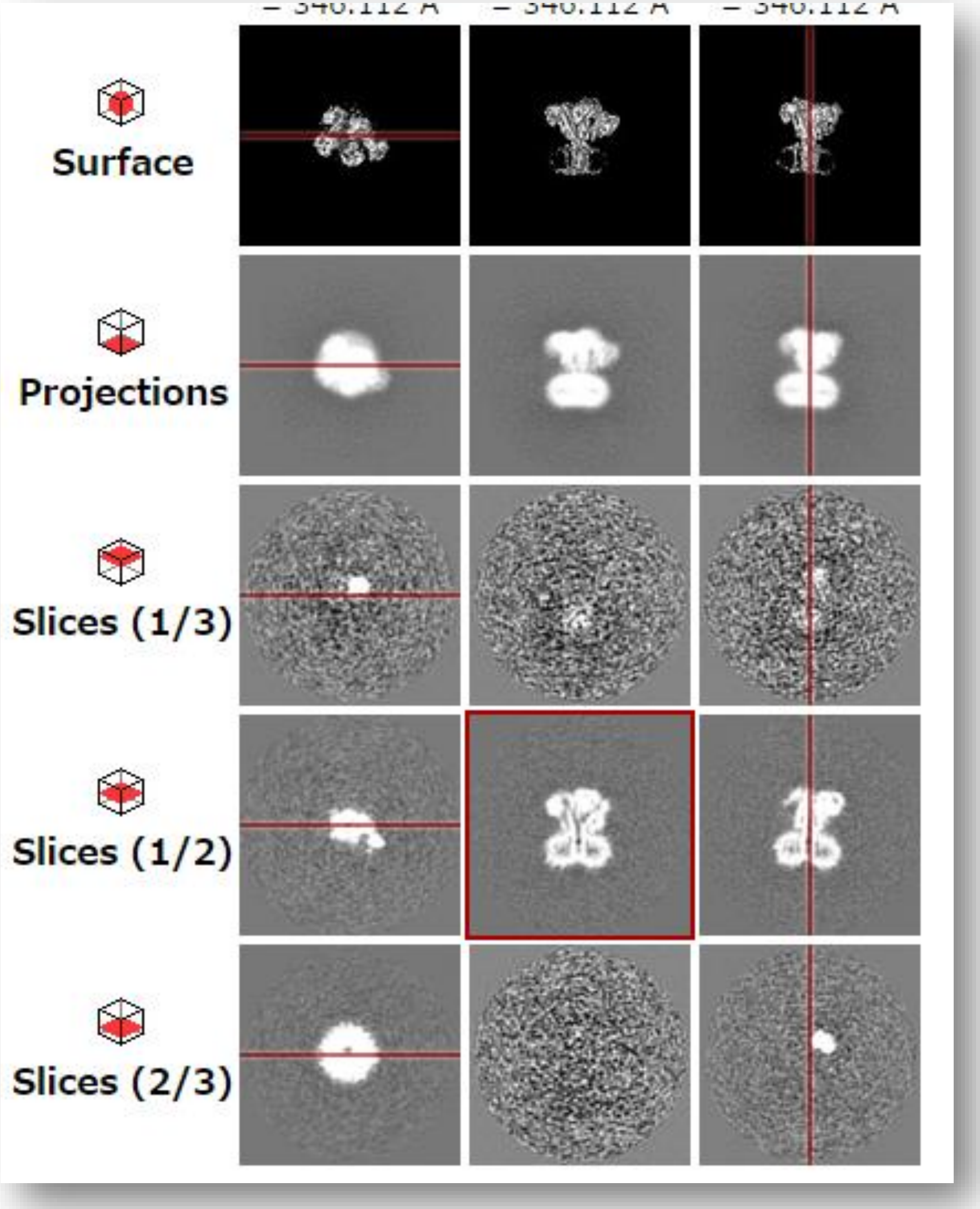
EMDB-6553 + PDB-3jch



PDB model + ED map (PDB-3cl0)

full-res. EM map interface, under developing

Projections & sections



For each the data

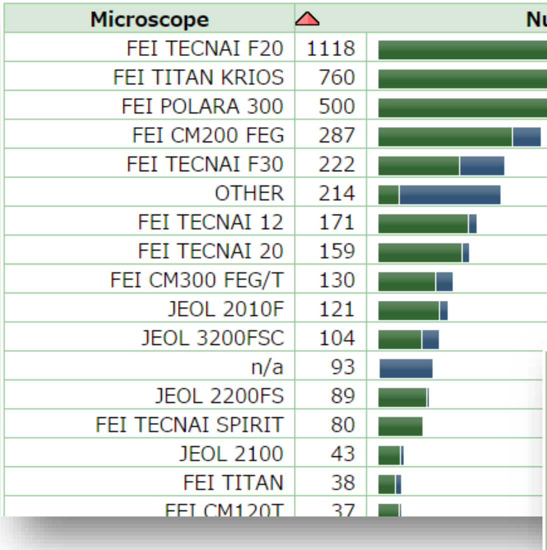
(except for some huge raw tomogram)

can see background noise in sections

can peep (盗み見)…

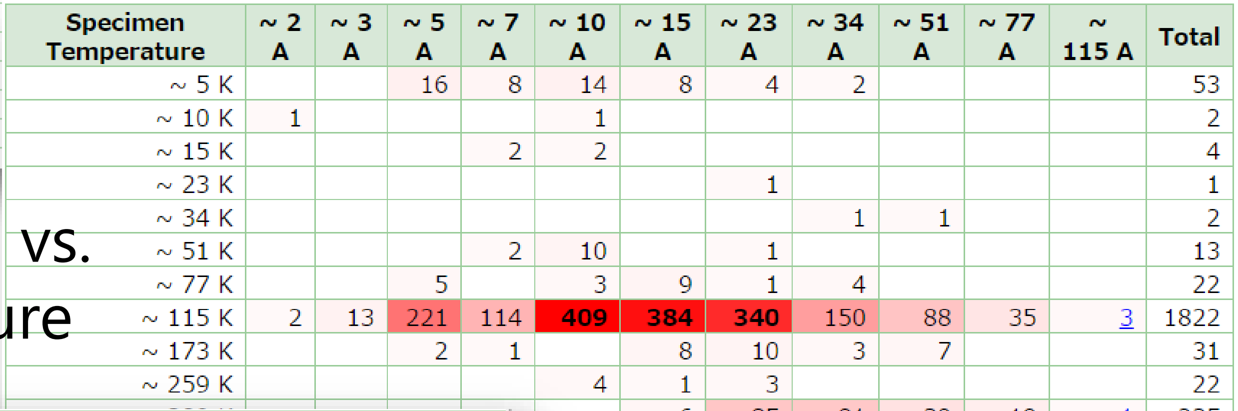
- type of real-space mask
- CTF amplitude correction
- etc.

Statistics



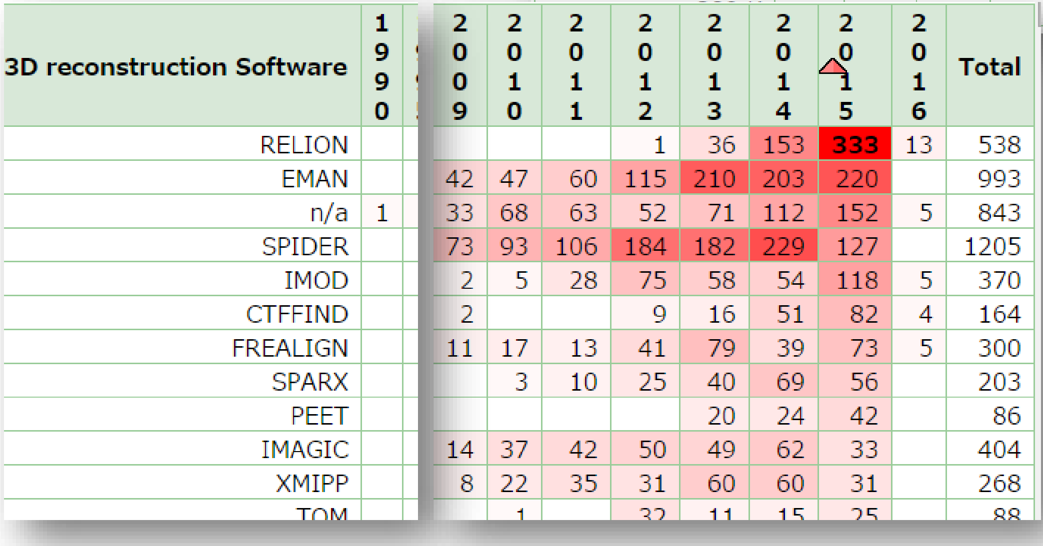
Microscope hit-chart

Microscope	Number of data entries
FEI TECNAI F20	1118
FEI TITAN KRIOS	760
FEI POLARA 300	500
FEI CM200 FEG	287
FEI TECNAI F30	222
OTHER	214
FEI TECNAI 12	171
FEI TECNAI 20	159
FEI CM300 FEG/T	130
JEOL 2010F	121
JEOL 3200FSC	104
n/a	93
JEOL 2200FS	89
FEI TECNAI SPIRIT	80
JEOL 2100	43
FEI TITAN	38
FEI CM120T	37



Resolution vs. temperature

Specimen Temperature	~ 2 A	~ 3 A	~ 5 A	~ 7 A	~ 10 A	~ 15 A	~ 23 A	~ 34 A	~ 51 A	~ 77 A	~ 115 A	Total
~ 5 K			16	8	14	8	4	2				53
~ 10 K	1				1							2
~ 15 K				2	2							4
~ 23 K							1					1
~ 34 K								1	1			2
~ 51 K				2	10		1					13
~ 77 K			5		3	9	1	4				22
~ 115 K	2	13	221	114	409	384	340	150	88	35	3	1822
~ 173 K			2	1		8	10	3	7			31
~ 259 K					4	1	3					22
						6	85	81	39	19	1	235



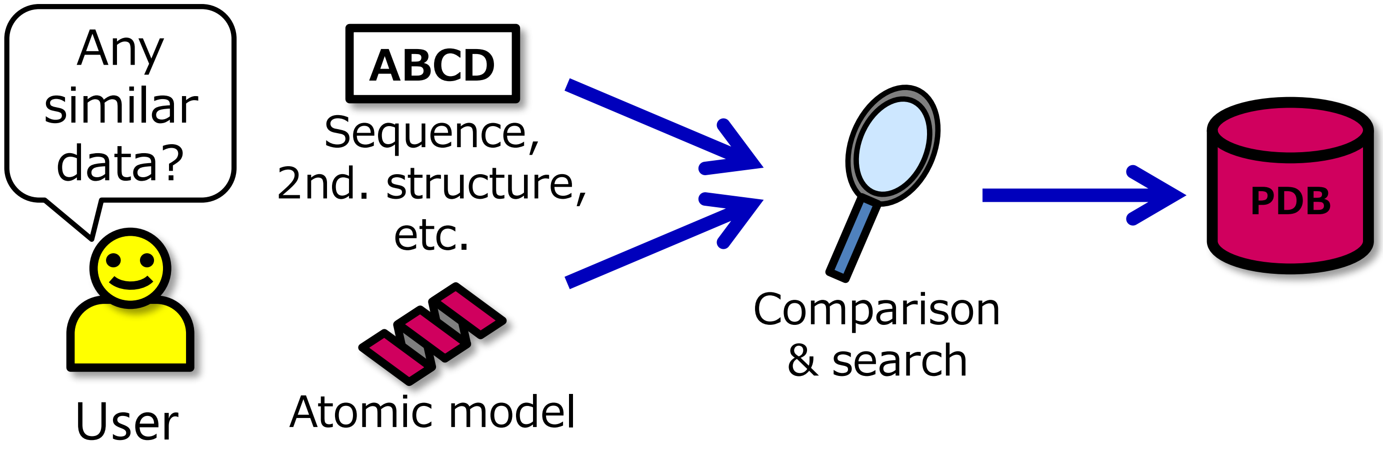
3D reconstruction Software	1990	2009	2010	2011	2012	2013	2014	2015	2016	Total
RELION					1	36	153	333	13	538
EMAN		42	47	60	115	210	203	220		993
n/a	1	33	68	63	52	71	112	152	5	843
SPIDER		73	93	106	184	182	229	127		1205
IMOD		2	5	28	75	58	54	118	5	370
CTFFIND		2			9	16	51	82	4	164
FREALIGN		11	17	13	41	79	39	73	5	300
SPARX			3	10	25	40	69	56		203
PEET						20	24	42		86
IMAGIC		14	37	42	50	49	62	33		404
XMIPP		8	22	35	31	60	60	31		268
TOM			1		32	11	15	25		88

Software yearly trends

Omokage search

Omokage search

Typical structure search (*e.g. Structure Navigator* of PDBj)



Any similar data?

User

ABCD

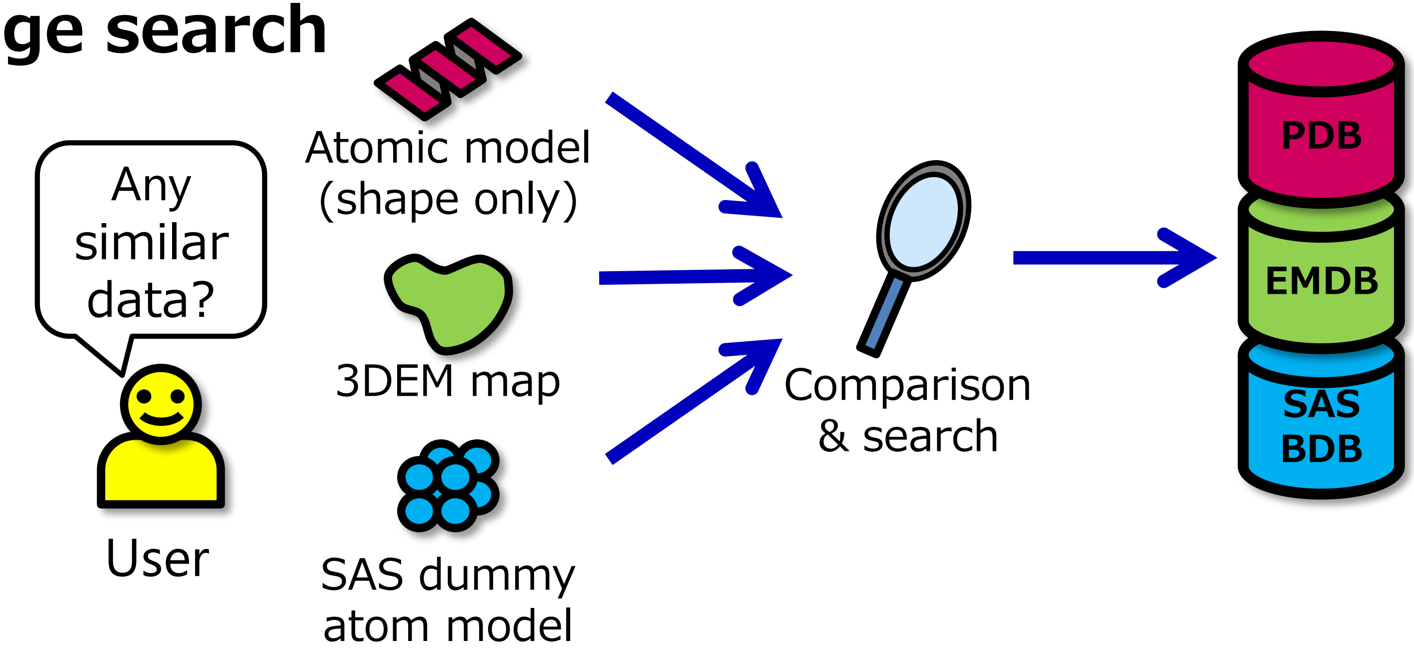
Sequence, 2nd. structure, etc.

Atomic model

Comparison & search

PDB

Omokage search



Any similar data?

User

Atomic model (shape only)

3DEM map

SAS dummy atom model

Comparison & search

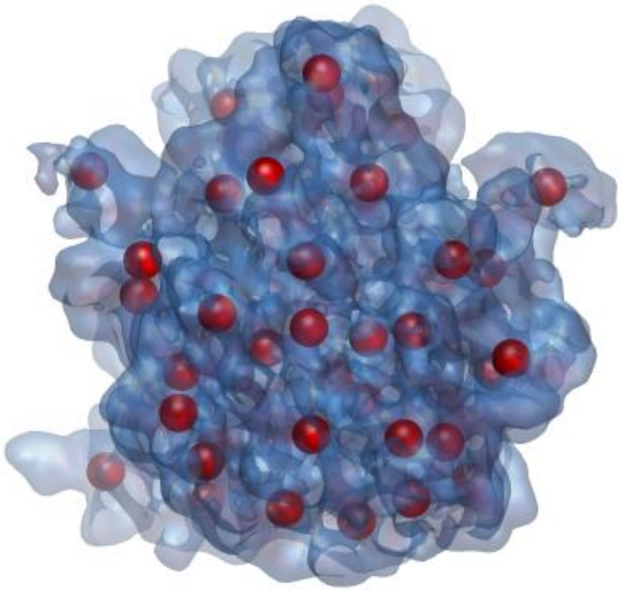
PDB

EMDB

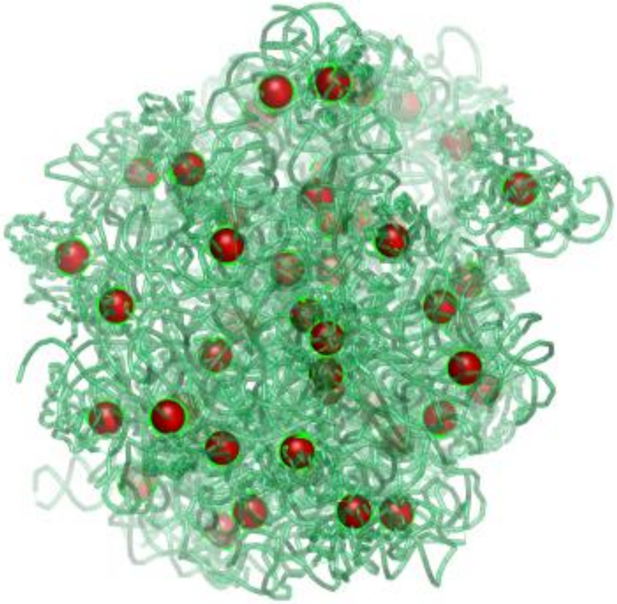
SAS BDB

Dot model by vector quantization

“dot model” by *qvol* & *qpdb* in *Situs* package
（*Situs*： famous software for “3D fitting” of EM data）



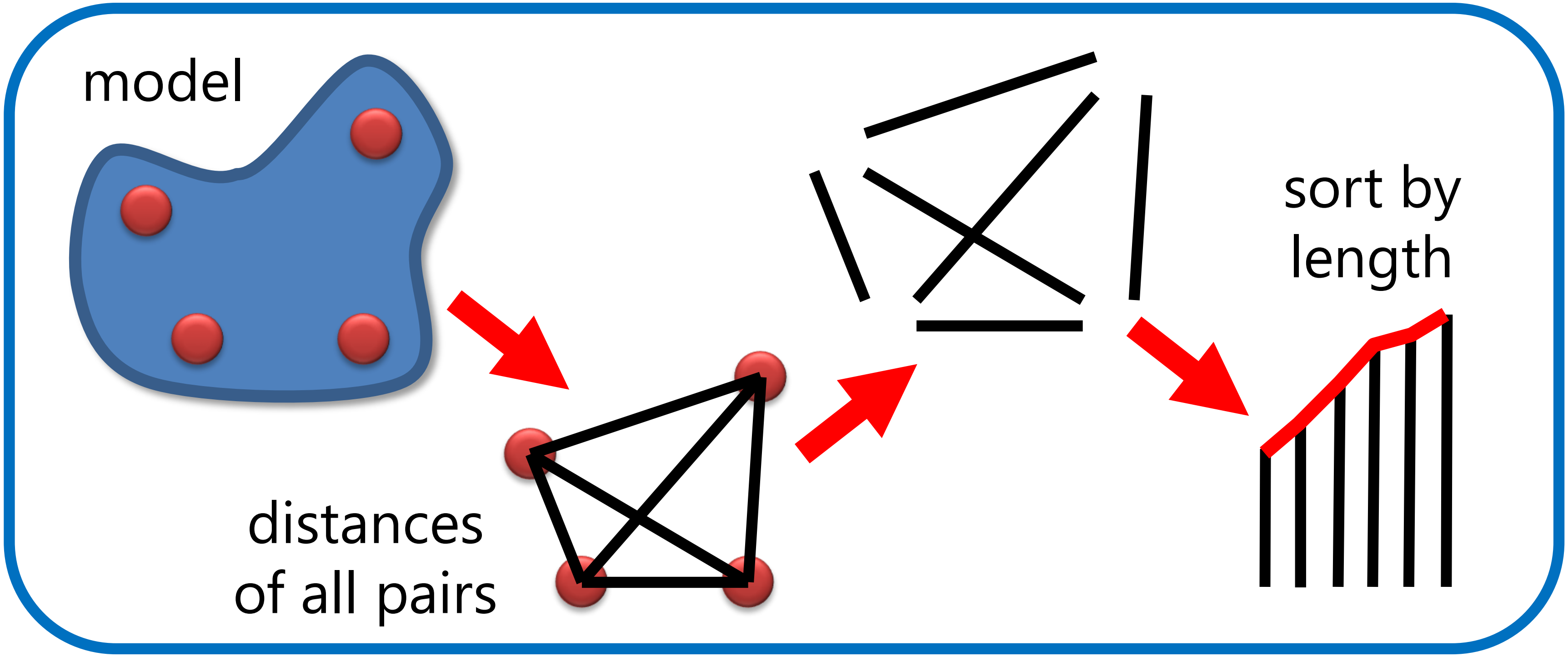
3D-map
EMDB-1003



Atomic model
PDB-1ml5

Each data -> represent with 50 dots
-> compare with “dots model” instead of structure

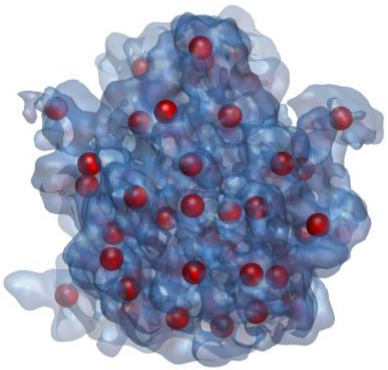
iDR profile comparison



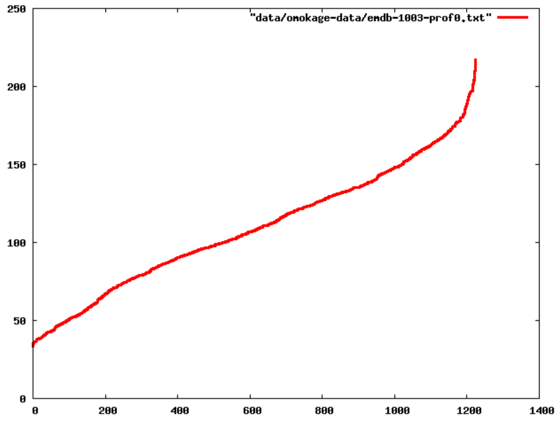
model

distances
of all pairs

sort by
length

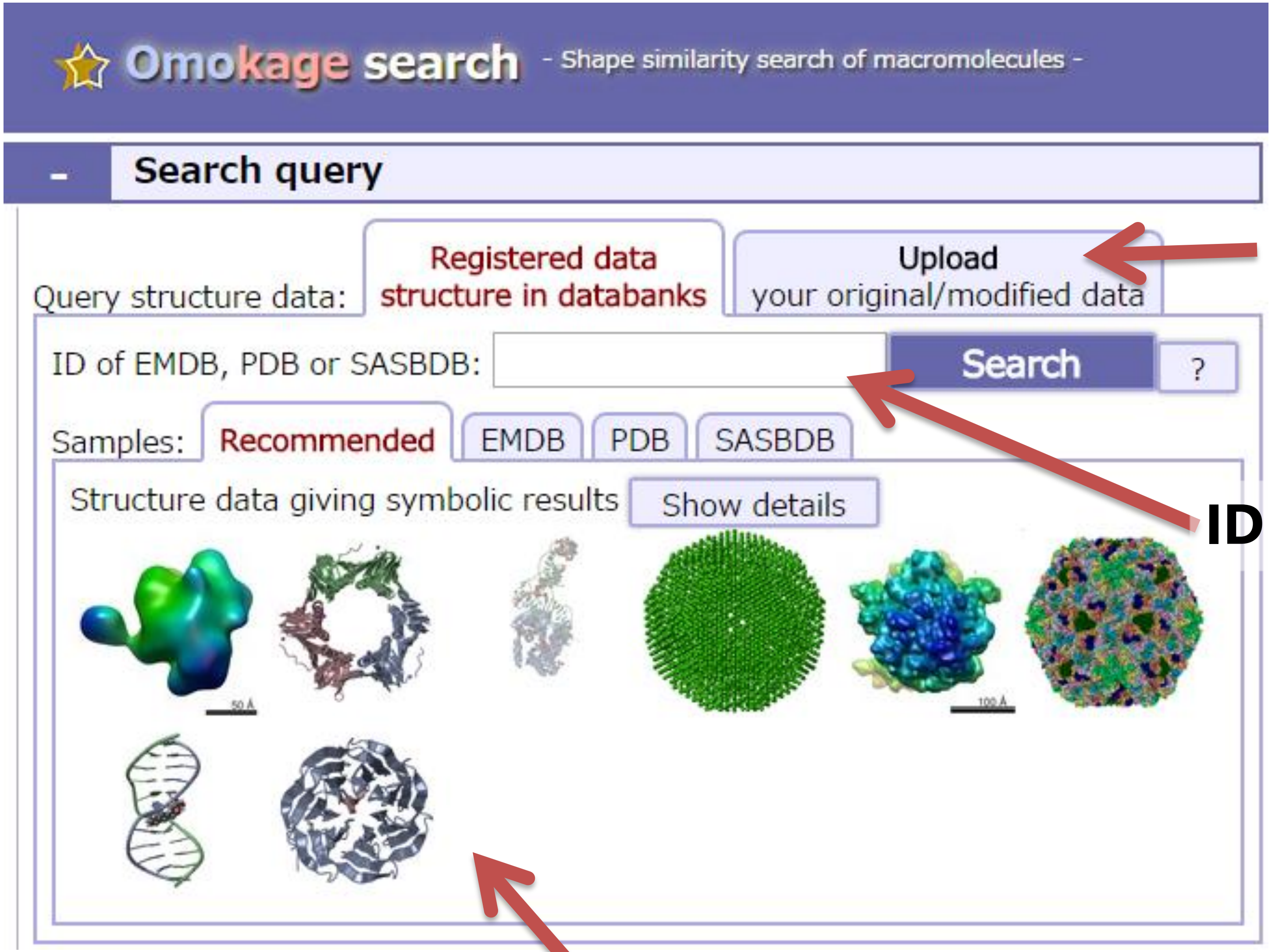


50 dots
1225 pairs



Suzuki *et al. bioinformatics*, 2015

Omokage search



file upload

ID input box

Sample data

Search results

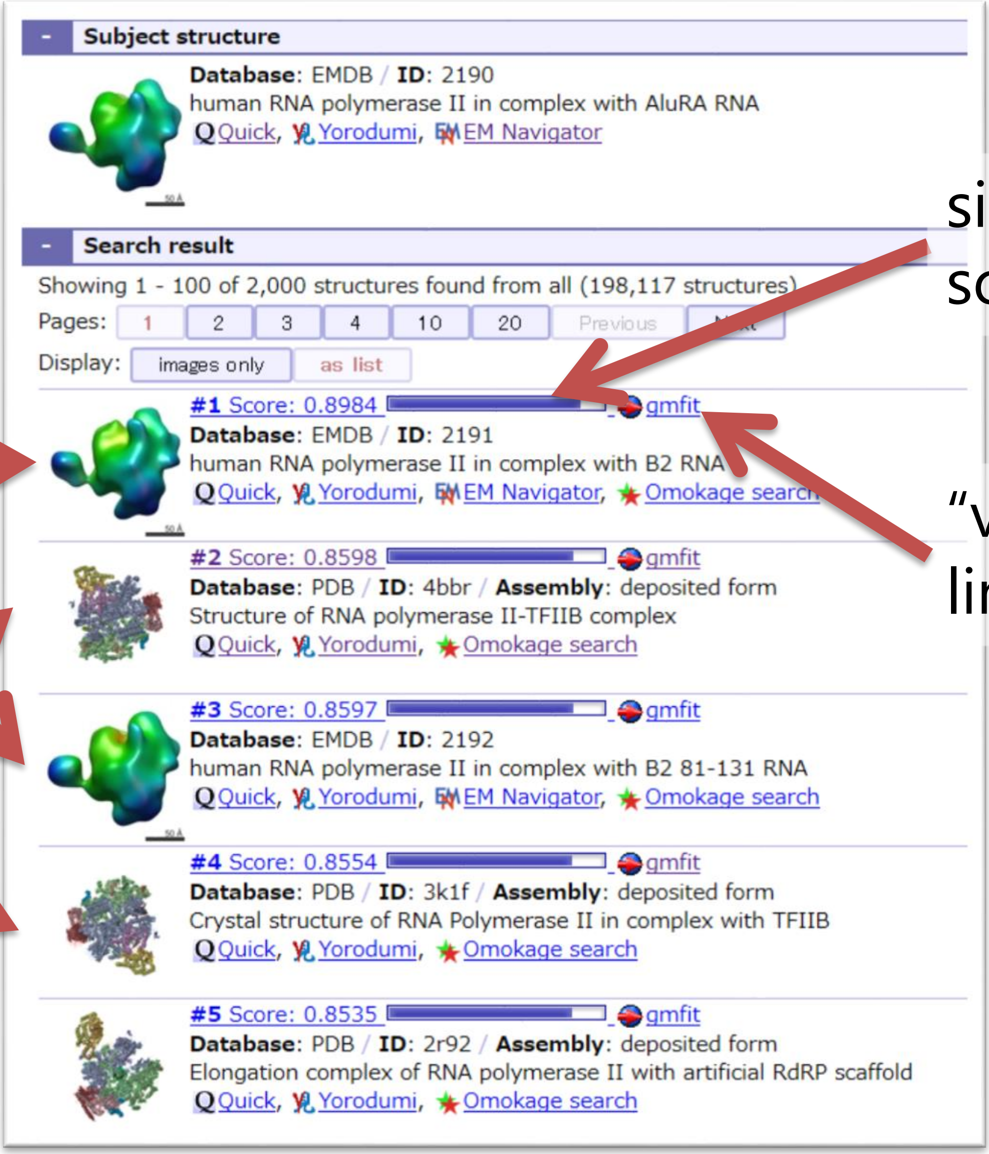
RNAP II
@25A res.

EMDB
entries

PDB
entries

similarity
score

“view fitting”
link to *gmfit*



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 2,000 structures found from all (198,117 structures)

Pages: 1 | 2 | 3 | 4 | 10 | 20 | Previous | Next

Display: images only | as list

#1 Score: 0.8984 gmfit
Database: EMDB / **ID**: 2191
human RNA polymerase II in complex with B2 RNA
Quick, Yorodumi, EM Navigator, Omokage searc

#2 Score: 0.8598 gmfit
Database: PDB / **ID**: 4bbr / **Assembly**: deposited form
Structure of RNA polymerase II-TFIIB complex
Quick, Yorodumi, Omokage search

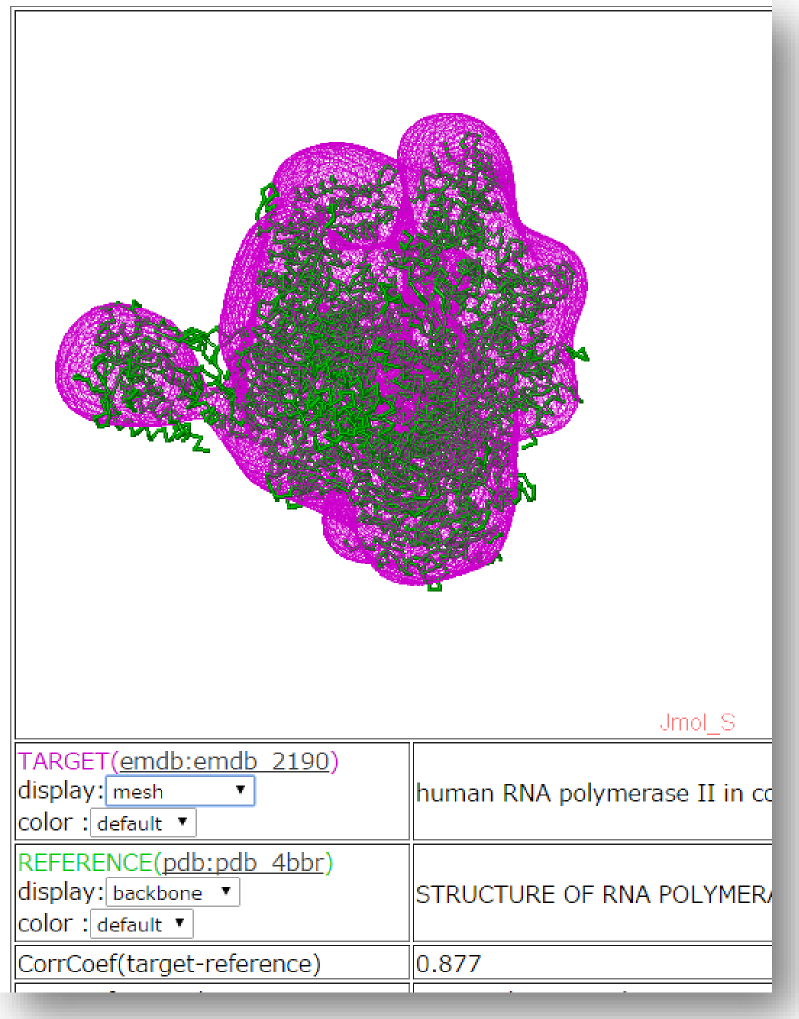
#3 Score: 0.8597 gmfit
Database: EMDB / **ID**: 2192
human RNA polymerase II in complex with B2 81-131 RNA
Quick, Yorodumi, EM Navigator, Omokage search

#4 Score: 0.8554 gmfit
Database: PDB / **ID**: 3k1f / **Assembly**: deposited form
Crystal structure of RNA Polymerase II in complex with TFIIB
Quick, Yorodumi, Omokage search

#5 Score: 0.8535 gmfit
Database: PDB / **ID**: 2r92 / **Assembly**: deposited form
Elongation complex of RNA polymerase II with artificial RdRP scaffold
Quick, Yorodumi, Omokage search

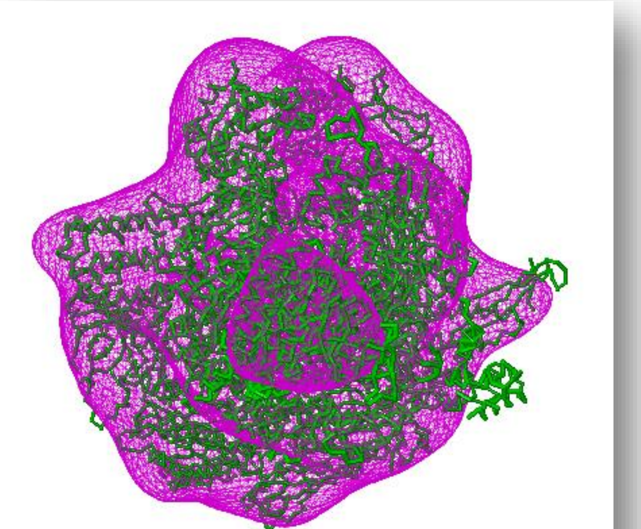
gmfit

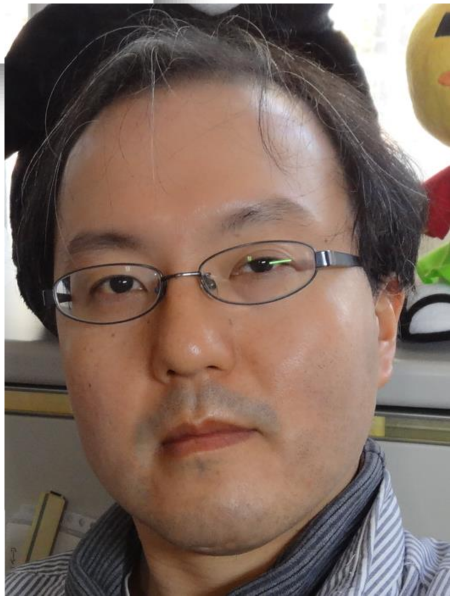
Fitting by *gmfit*



TARGET(emdb:emdb_2190) display: mesh ▼ color : default ▼	human RNA polymerase II in co
REFERENCE(pdb:pdb_4bbr) display: backbone ▼ color : default ▼	STRUCTURE OF RNA POLYMERA
CorrCoef(target-reference)	0.877

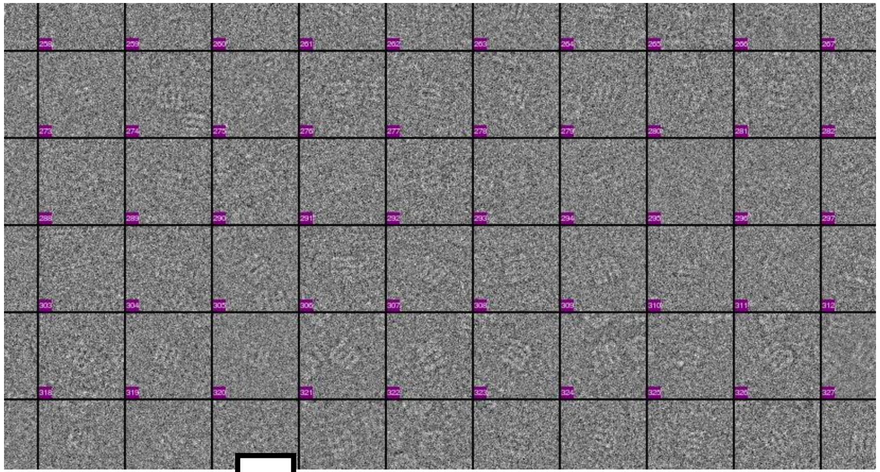
Side view



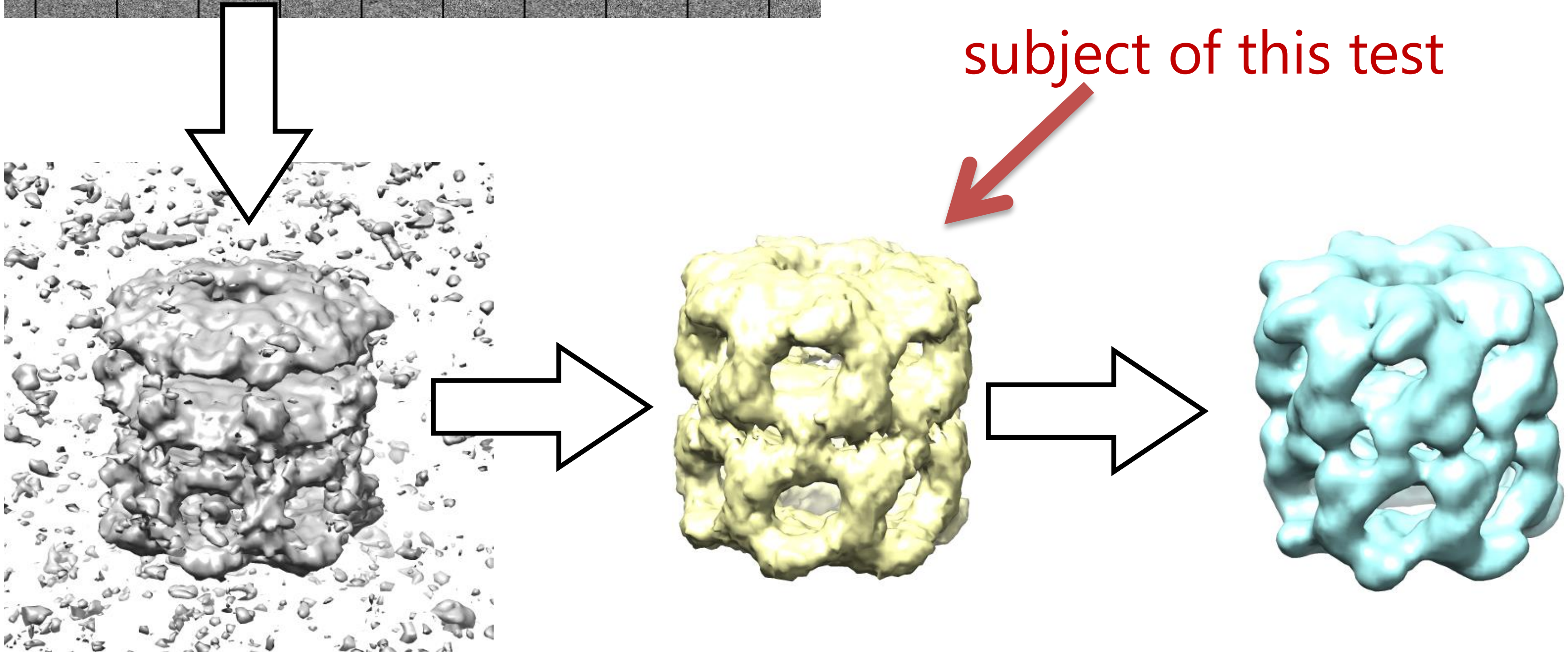


Dr. T. Kawabata

Test case with GroEL reconstruction



particle images
from “EMDB test image data”
http://www.ebi.ac.uk/pdbe/emdb/test_data.html



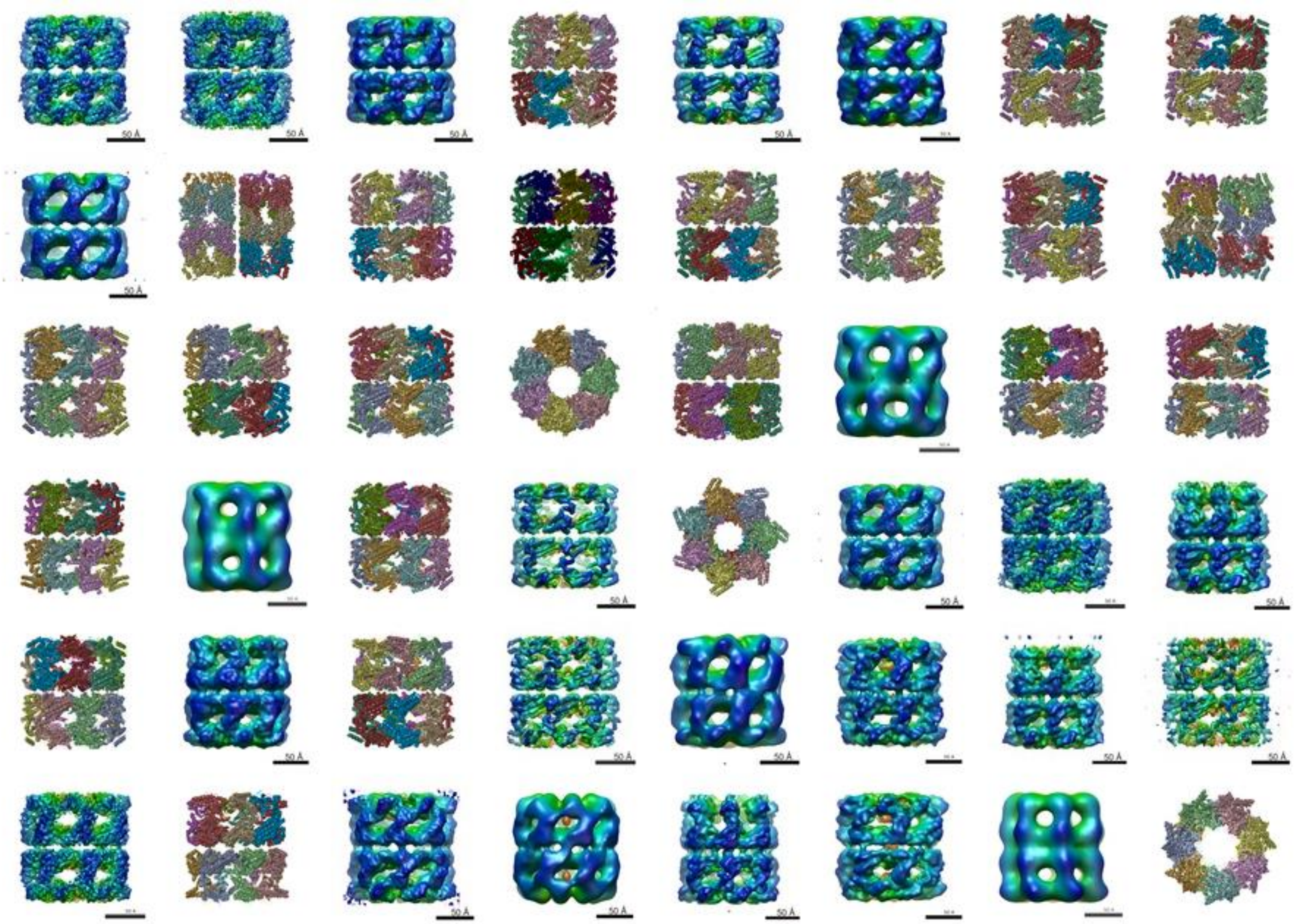
subject of this test

initial model

refinement #1

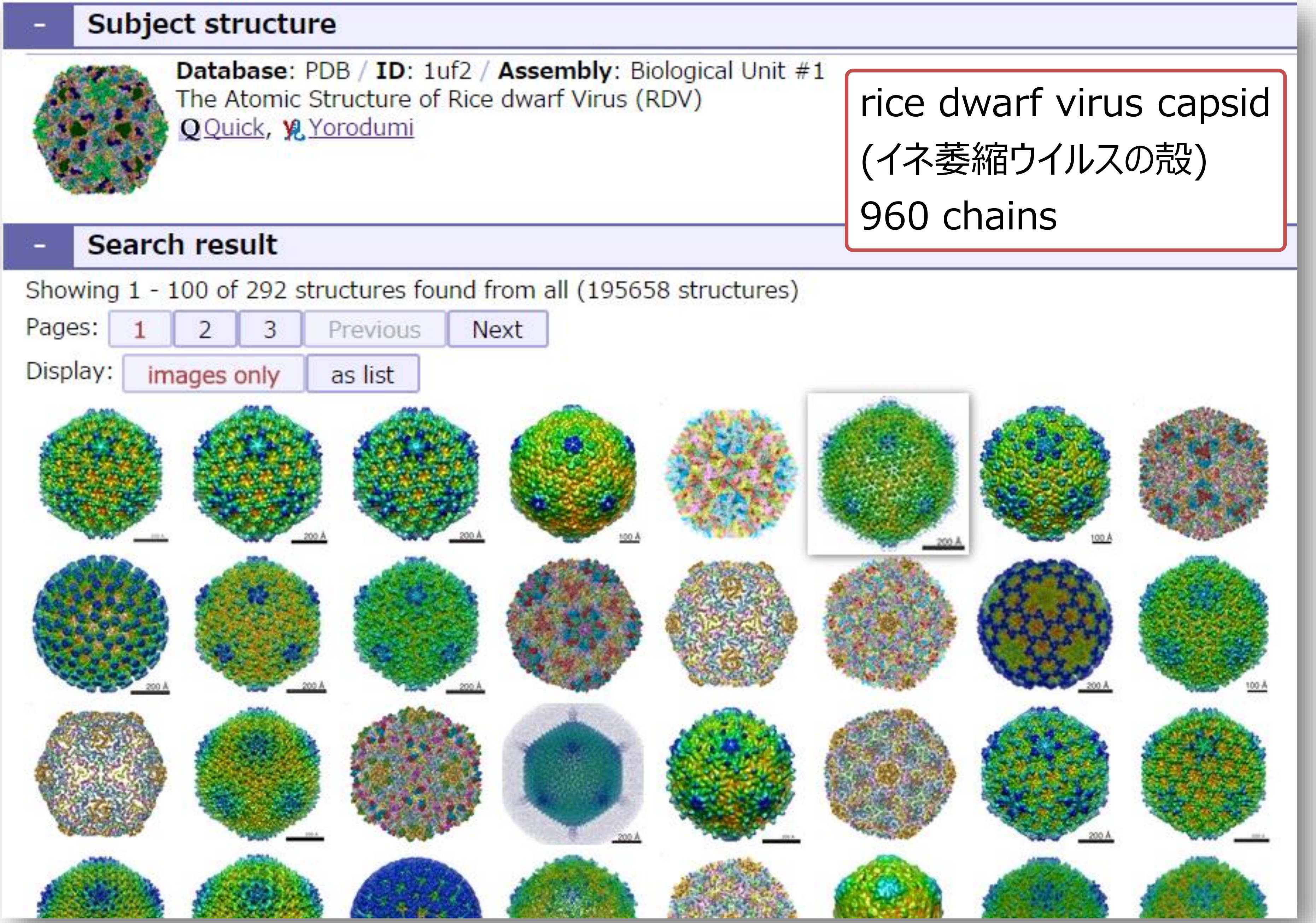
refinement #13

Search result



Top >50 data, GroEL only

Virus capsid



Subject structure

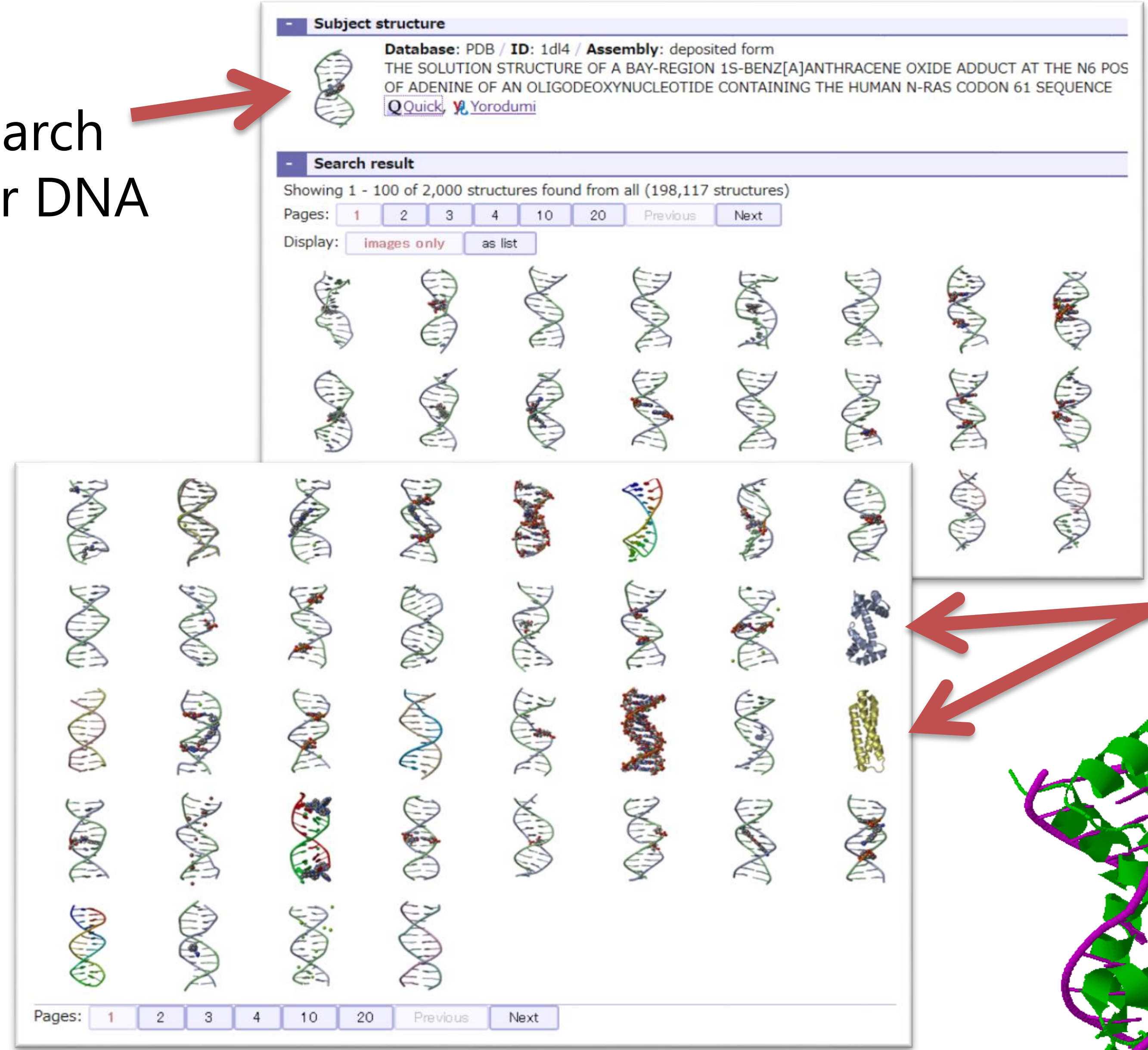
Database: PDB / **ID**: 1uf2 / **Assembly**: Biological Unit #1
The Atomic Structure of Rice dwarf Virus (RDV)
Quick, Yorodumi

Search result

Showing 1 - 100 of 292 structures found from all (195658 structures)

rice dwarf virus capsid
(イネ萎縮ウイルスの殻)
960 chains

Small DNA fragment



Subject structure

Database: PDB / **ID**: 1dl4 / **Assembly**: deposited form
THE SOLUTION STRUCTURE OF A BAY-REGION 1S-BENZ[A]ANTHRACENE OXIDE ADDUCT AT THE N6 POS
OF ADENINE OF AN OLIGODEOXYNUCLEOTIDE CONTAINING THE HUMAN N-RAS CODON 61 SEQUENCE

Search result

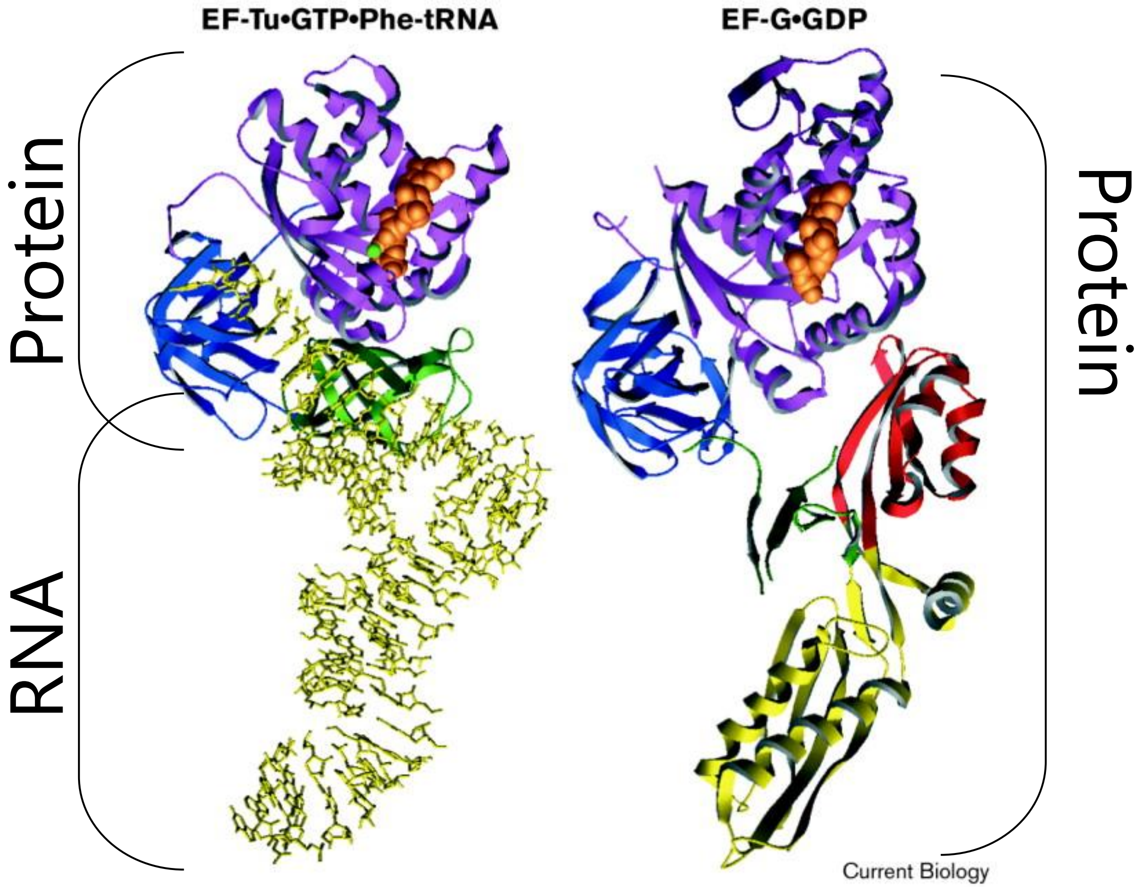
Showing 1 - 100 of 2,000 structures found from all (198,117 structures)

search
for DNA

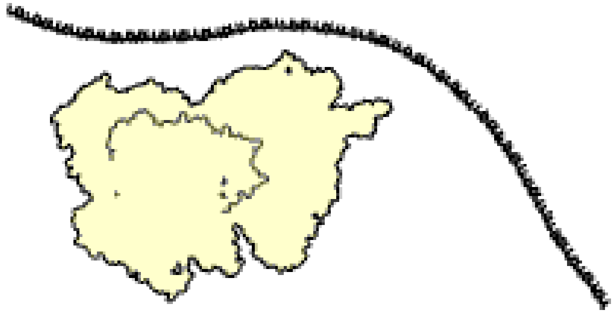
proteins
found

fitting
by gmfit

Function within shapes



Green, *Curr Biol.*, 2000

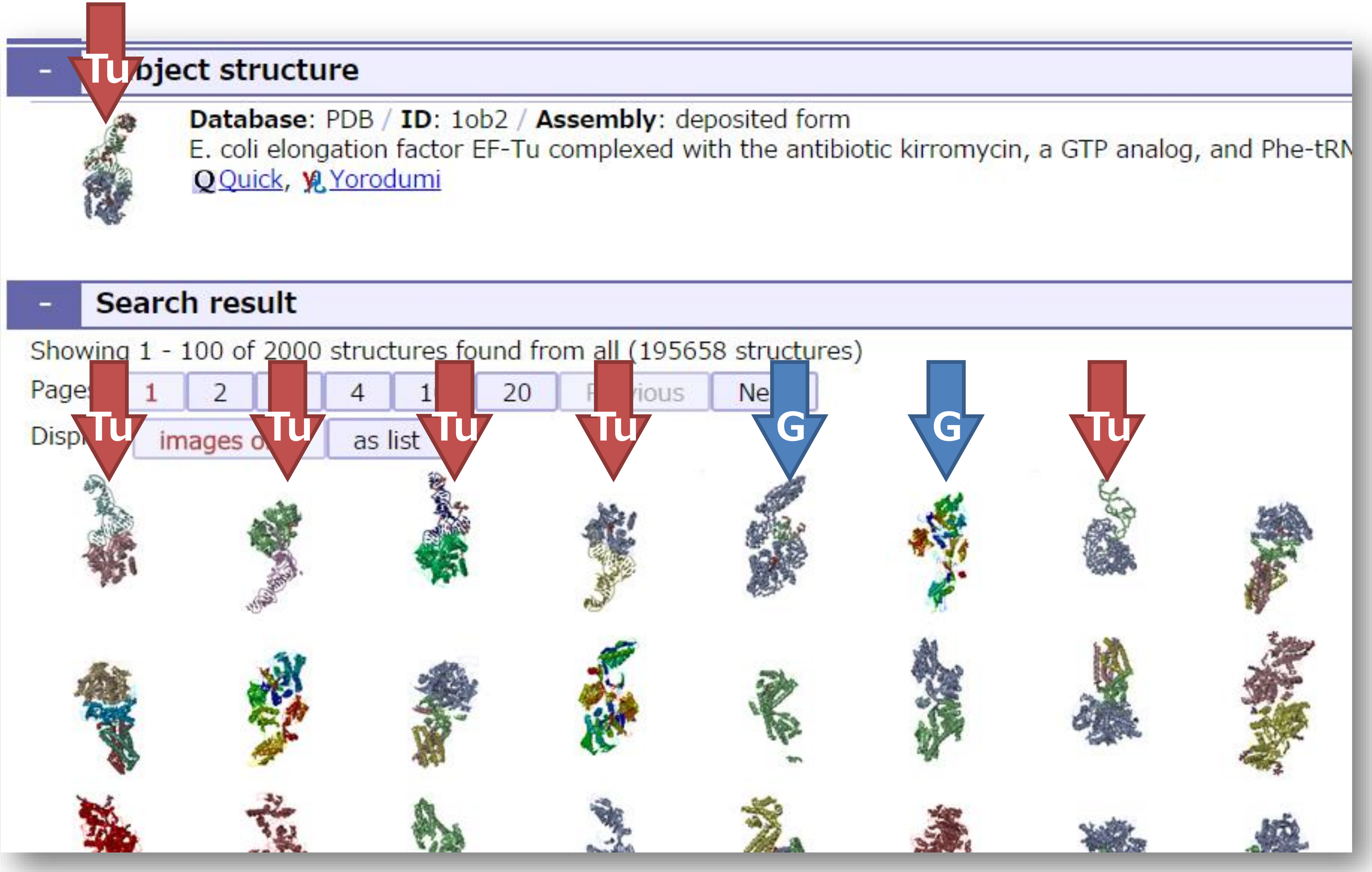


protein translation

Bensaccount at en. wikipedia / CC-BY-3.0.

Molecular mimicry (分子擬態)

Find molecular mimicry



bject structure

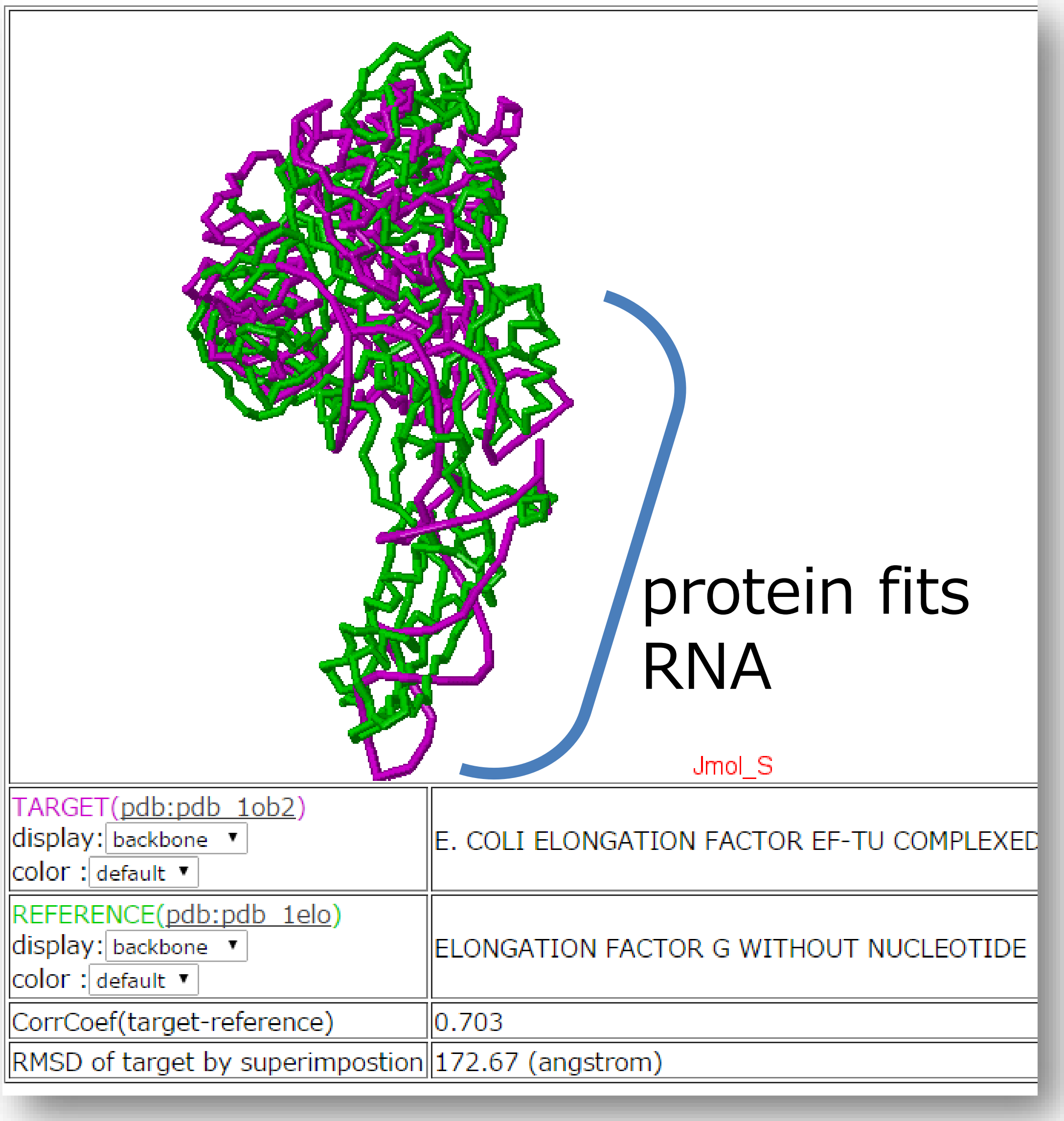
Database: PDB / **ID**: 1ob2 / **Assembly**: deposited form
E. coli elongation factor EF-Tu complexed with the antibiotic kirromycin, a GTP analog, and Phe-tRN

Search result

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

EF-G structures found by search for **EF-Tu** complex

Fit protein on RNA



protein fits RNA

TARGET(pdb:pdb_1ob2) display: backbone color : default	E. COLI ELONGATION FACTOR EF-TU COMPLEXED
REFERENCE(pdb:pdb_1elo) display: backbone color : default	ELONGATION FACTOR G WITHOUT NUCLEOTIDE
CorrCoef(target-reference)	0.703
RMSD of target by superimpostion	172.67 (angstrom)

Fitting of PDB 1ob2 + PDB-1elo by *gmfit*

Enjoy 3DEM data
& shape similarity

Special thanks to

- **all DB staffs**
- **all DB depositors**
- **all users**