Slide 1

PDBj
Protein Data Bank Japan

Recent activities of PDBj and wwPDB

Genji Kurisu
Institute for Protein Research,
Osaka University

1

Slide 2

PDBj
Protein Data Bank Japan

Protein Data Bank Japan
http://pdbj.org/
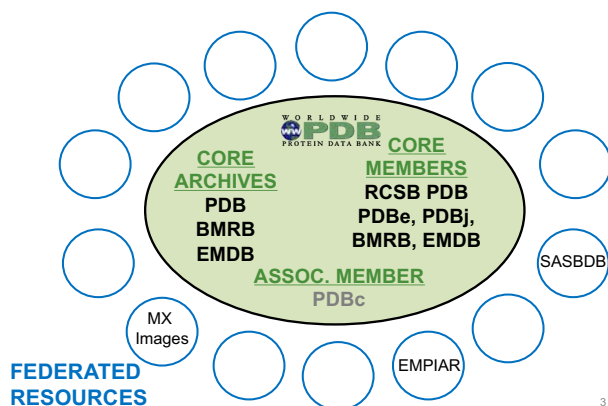
Since 2001, PDBj has been managed at **Institute for Protein Research, Osaka University** as a member of the **wwPDB**, to curate and process the deposited data for an open and single archive.
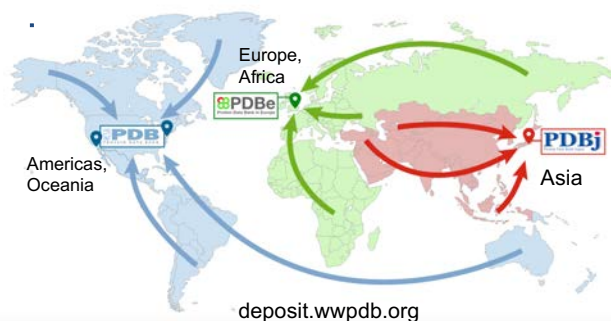
Slide 3

PDBj

wwPDB Architecture

CORE ARCHIVES
PDB
BMRB
EMDB

CORE MEMBERS
RCSB PDB
PDBe, PDBj,
BMRB, EMDB

ASSOC. MEMBER
PDBc

SASBDB

MX Images

EMPIAR

FEDERATED RESOURCES

3

Slide 4

PDBj

PDBj is in charge of processing and annotating the depositions from Asian and Middle-east regions

Europe, Africa

Americas, Oceania
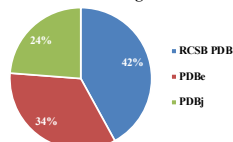
Asia

deposit.wwpdb.org

4

Slide 5

PDBj

PDB Core Archive Depositions
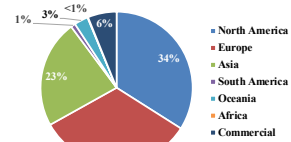
- 12,179 depositions in 2018.
- Rapid growth in 3DEM
  - Exceeded NMR depositons

Method	2017 Depositions	2018 Depositions
MX	11,889 (91.1%)	10594 (87.0%)
NMR	460 (3.5%)	418 (3.4%)
3DEM	674 (5.2%)	1140 (9.4%)
Other	26 (0.2%)	27 (0.2%)

2018 Processing Sites

- RCSB PDB 42%
- PDBe 34%
- PDBj 24%

2018 Depositor Locations

- North America 34%
- Europe 33%
- Asia 23%
- South America 3%
- Oceania 1%
- Africa <1%
- Commercial 6%

5

Slide 6

PDBj

PDB Core Archive Growth

of Structures (y-axis): 0 to 180000

Year (x-axis): 2000–2019
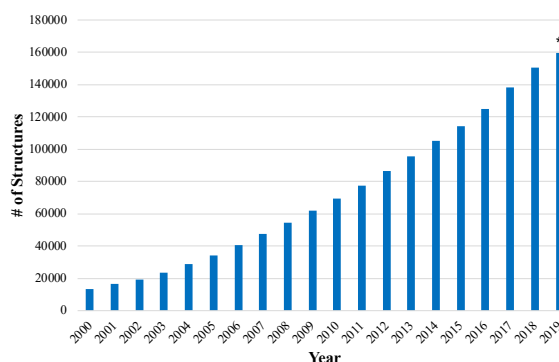
* As of 3 Sep 2019

6

PDB Core Archive Downloads

Year	Total	Total FTP Archive	Total Website
2019 ★	373,392,752	236,377,588	137,015,164
2018	749,356,769*	N/A	N/A
2017	679,421,200	454,723,083	224,698,117
2016	591,876,087	366,677,897	225,198,190
2015	534,339,871	368,244,766	166,095,105
2014	512,227,251	339,193,721	173,033,530
2013	441,262,210	296,176,290	145,085,920
2012	376,944,070	255,837,735	121,106,335
2011	383,131,048	276,952,286	106,178,762
2010	294,326,976	213,180,966	81,146,010
2009	326,362,536	271,116,934	57,245,602

More than 1.8 million/day!

N.B.: Some 2018 data lost due to GDPR. Hope to be back on track for 2019→.
* As of 3 Sep 2019

Geographic Origins of FTP downloads; 2012-2015

PDB Core Archive Update I

Ligand Validation
- Adapted software from Global Phasing Ltd. under a formal agreement
- Provides geometrical quality in 2D depiction
- Provides electron density fit for X-ray
- Now mandatory at deposition: identification of Ligand/s Of Interest (LOI, author's research focus)

8

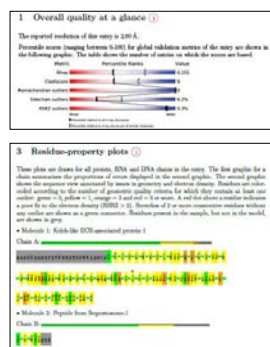## Official wwPDB Validation Report is requested for peer review

Example of Validation report

Ligand Validation- Examples on NADP Ligands

PDB entry 5zix (Better data quality)   PDB entry 1z4k (Worse data quality)
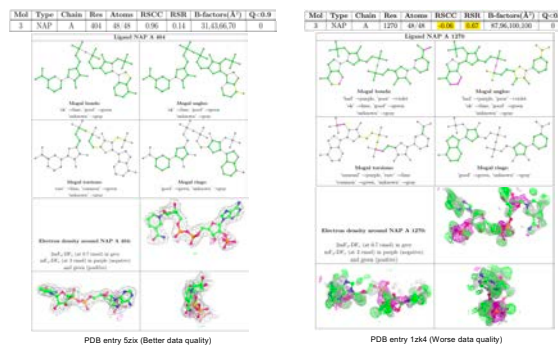
10

PDB Core Archive Update II

Coordinate Versioning
- **To improve data quality in the PDB archive**
- Depositors can now make corrections to existing structures in the PDB Core Archive by updating the atomic coordinates while preserving the original PDB identifier.
- Deployed on July 27th 2019 for all structures that were originally deposited via OneDep (Phase one)
- First coordinate replacement (PDB entry 5T26) versioned at FTP on August 7th 2019 following reviewers' comments based on the wwPDB validation reports
  - With PDB prefix and extension of 4 characters (e.g., from "1ABC" to "PDB_00001ABC")
  - Example: PDB_00001ABC_XYZ_V2-2.cif.gz

11

PDB Core Archive Update III

Use of ORCiD
- To enable better management of incoming data
- To credit PDB entries by author unambiguously
- 25% of the unique depositors provided ORCiD ids
- 3342 unique PIs with ORCiD ids

ORCID
Connecting Research and Researchers

12

Slide 13

PDB Core Archive Plans

- **ED map coefficients**
  - ED map coefficients have been provided to depositors during biocuration for new entries
  - As part of annual recalculation of validation reports, the ED map coefficients will be provided for existing entries at the wwPDB FTP site for better interpretation of the ligand validation by user community.
- **Coordinate Versioning (Legacy entries)**
  - Create OneDep deposition sessions
  - Enable authors access and login
  - Entries will be versioned at versioned FTP

13

Slide 14

wwPDB 2018-2019 Publications

Mandatory mmCIF deposition announced on 3rd April 2019
https://doi.org/10.1107/S2059798319004522
I/HM manuscript was submitted.

14

Slide 15

Atomic coordinates in PDBx/mmCIF

```
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
```
15

Slide 16

Acknowledgements

16