2017 wwPDB AC Meeting

Stephen K. Burley, Genji Kurisu, John L. Markley, and Sameer Velankar
Introductions and State of the wwPDB

Stephen K. Burley
Welcome

wwPDB Advisory Committee Chair and ICMRBS Representative

- R. Andrew Byrd

wwPDB Advisory Committee Members

- RCSB PDB: Paul Adams and Cynthia Wolberger
- PDBe: David Brown and Sarah Butcher
- PDBj: Tsuyoshi Inoue and Kei Yura
- BMRB: Gaetano Montelione and Arthur Edison
Welcome (cont.)

Regional Representatives
- China: Jianping Ding
- India: Manju Bansal

IUCr Representative
- Edward Baker

Macromolecular EM Community Representative
- Wah Chiu
Logistics/Support

In Case of Fire

- Emergency exits at both ends of corridor
- Rally point in front of medical school tower

Restrooms Across the Hall

RCSB PDB Administration Room 110

- Nicole Oorbeek (nicole.oorbeek@rcsb.org)
  Tel: 848-445-4903; Mobile: 732-859-9040
wwPDB Vision

Our Vision is to

Sustain a freely accessible, single global archive of experimentally determined structure data for biological macromolecules as an enduring public good.
Our Mission is to

1. Ensure open access to public domain experimentally determined structural biology data.
2. Provide expert deposition, validation, and biocuration services at no charge to Data Depositors.
3. Enable universal access for expert and non-expert Data Consumers with no limitations on usage.
4. Manage the PDB archive as a public good according to the FAIR Principles.
5. Lead the world in structural biology data representation, exchange, and visualization.
Developments since 2016 Meeting I

- Genji Kurisu succeeded Nakamura as PDBj Head
- Continued enhancement of the OneDep system for MX, NMR, 3DEM, and now SAS
- Rapid growth in 3DEM structure depositions and engagement with the 3DEM community
- Continued engagement with the NMR community
Developments since 2016 Meeting II

- Implementation of PDB archive versioning initiated

- Prototyping of a loosely coupled Federation of Structural Biology Data Archives with SASBDB

- Archival content and management improvement

- Collection of ligand of interest and experimental support of multimeric assemblies
2016 PDB Deposition Statistics

- On track for >12,000 depositions in 2017
- More 3DEM structures deposited in 2016 versus NMR structures
- On track for repetition of 3DEM>NMR in 2017

<table>
<thead>
<tr>
<th>Method</th>
<th>2016 Depositions</th>
<th>2015 Depositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MX</td>
<td>10583 (91%)</td>
<td>10167</td>
</tr>
<tr>
<td>NMR</td>
<td>473 (4%)</td>
<td>510</td>
</tr>
<tr>
<td>3DEM</td>
<td>531 (4.6%)</td>
<td>255</td>
</tr>
<tr>
<td>Other</td>
<td>27 (.2%)</td>
<td>25</td>
</tr>
</tbody>
</table>

**Processing Site**
- PDBe, 4051 (35%)
- PDB, 2240 (19%)
- RCSB, 5323 (46%)

**Depositor Location**
- North America 34%
- Europe 33%
- South America 19%
- Oceania 9%
- Asia 9%
- Africa 3%
- Industry <1%
- Other <1%
Growth in Annual PDB Depositions

- More than 1 billion atoms
- Total Number of Annual Depositions
- Projected Annual Depositions
Growth of the PDB Archive

More than 1 billion atoms
# PDB Data Download Statistics

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Total FTP Archive</th>
<th>Total Website</th>
<th>RCSB PDB FTP Archive</th>
<th>RCSB PDB Website</th>
<th>PDBe FTP Archive</th>
<th>PDBe Website</th>
<th>PDBj FTP Archive</th>
<th>PDBj Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>328,362,536</td>
<td>271,116,934</td>
<td>57,245,602</td>
<td>222,984,760</td>
<td>53,507,785</td>
<td>30,141,339</td>
<td>1,475,116</td>
<td>17,990,835</td>
<td>2,262,701</td>
</tr>
<tr>
<td>2010</td>
<td>294,326,976</td>
<td>213,180,966</td>
<td>81,146,010</td>
<td>159,248,214</td>
<td>64,569,658</td>
<td>34,383,219</td>
<td>14,017,349</td>
<td>19,549,533</td>
<td>2,559,003</td>
</tr>
<tr>
<td>2011</td>
<td>383,131,048</td>
<td>276,952,286</td>
<td>106,178,762</td>
<td>204,939,406</td>
<td>81,560,098</td>
<td>40,960,368</td>
<td>18,515,245</td>
<td>31,052,512</td>
<td>6,103,419</td>
</tr>
<tr>
<td>2013</td>
<td>441,262,210</td>
<td>296,176,290</td>
<td>145,085,920</td>
<td>215,331,908</td>
<td>97,549,580</td>
<td>43,684,850</td>
<td>37,762,496</td>
<td>37,159,532</td>
<td>9,773,844</td>
</tr>
<tr>
<td>2015</td>
<td>534,339,871</td>
<td>368,244,766</td>
<td>166,095,105</td>
<td>255,346,630</td>
<td>111,802,897</td>
<td>48,544,330</td>
<td>41,127,219</td>
<td>64,353,806</td>
<td>13,164,989</td>
</tr>
</tbody>
</table>

More than 1.5 million/day

Geographic origins of FTP downloads, 2012-2015
Funding Status Update

- RCSB PDB: NSF/NIH/DOE funding through 12/31/2018 (Competing renewal process)

- BMRB: NIH NIGMS funding through 03/31/2019 (Competing renewal process)

- PDBe: EMBL-EBI, Wellcome Trust through 01/01/2020 (Competing renewal process)

- PDBj: NBDC-JST and AMED funding through 03/31/2022 (Competing renewal process)
## wwPDB Collaboration Staffing Commitments Oct 2017-Sep 2018

<table>
<thead>
<tr>
<th>wwPDB Partner</th>
<th>Software Development</th>
<th>Production Maintenance/Production Management</th>
<th>Requirements Setting/Testing</th>
<th>Archive Keeping/Outreach</th>
<th>Biocuration/Remediation</th>
<th>Total FTE Commitments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCSB PDB</td>
<td>3.2</td>
<td>1.5</td>
<td>1.0</td>
<td>2.1</td>
<td>7.0</td>
<td>14.8</td>
</tr>
<tr>
<td>PDBe</td>
<td>2.9</td>
<td>1.0</td>
<td>0.5</td>
<td>0.1</td>
<td>4.0</td>
<td>8.5</td>
</tr>
<tr>
<td>PDBj</td>
<td>0.1</td>
<td>0.9</td>
<td>0.5</td>
<td>0.5</td>
<td>4.3</td>
<td>6.3</td>
</tr>
<tr>
<td>BMRB</td>
<td>1.25</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.25</td>
</tr>
<tr>
<td>wwPDB</td>
<td>7.45</td>
<td>3.4</td>
<td>2.0</td>
<td>2.7</td>
<td>15.3</td>
<td>30.85</td>
</tr>
</tbody>
</table>
# OneDep 2016/2017 Progress vs. Goals

<table>
<thead>
<tr>
<th>Projects</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Projects</strong></td>
<td><strong>Timeline</strong></td>
</tr>
<tr>
<td>Archive content improvement: V5</td>
<td>Q4</td>
</tr>
<tr>
<td>Phase 1: Migrate and remediate legacy EM entries to V5 files</td>
<td></td>
</tr>
<tr>
<td>Phase 2: Migrate and remediate legacy NMR and X-ray entries to V5</td>
<td></td>
</tr>
<tr>
<td>Phase 3: Update FTP archive with remediated V5 files</td>
<td></td>
</tr>
<tr>
<td>Archive content improvement: Author of record file versioning (Parallel versioned ftp tree)</td>
<td></td>
</tr>
<tr>
<td>Phase 0: Provide better history revision record in the archive files</td>
<td></td>
</tr>
<tr>
<td>Phase 1: Provide versioned FTP tree for model files</td>
<td></td>
</tr>
<tr>
<td><strong>Phase 2: Enable depositor initiated coordinate replacement</strong></td>
<td></td>
</tr>
<tr>
<td>Archive content improvement: Carbohydrate and Protein modification remediation</td>
<td></td>
</tr>
<tr>
<td>Phase 1: Finalize representation requirements</td>
<td></td>
</tr>
<tr>
<td>Backend stabilization</td>
<td></td>
</tr>
<tr>
<td>Managing the lifecycle of inactive un-submitted sessions</td>
<td></td>
</tr>
<tr>
<td>Ensure all traffic is encrypted (https)</td>
<td></td>
</tr>
<tr>
<td>Separate user account from DepUI to support distributed access</td>
<td></td>
</tr>
<tr>
<td><strong>Enable WF to use external computing resources</strong></td>
<td></td>
</tr>
<tr>
<td>OneDep public facing</td>
<td></td>
</tr>
<tr>
<td>Extend collection of ORCID ID</td>
<td></td>
</tr>
<tr>
<td>Better collection of exp. evidence for depositor’s assemblies</td>
<td></td>
</tr>
<tr>
<td>Collection of ligand info that the focus of depositor’s research</td>
<td></td>
</tr>
<tr>
<td>Inclusion of SAXS/NMR hybrid</td>
<td></td>
</tr>
<tr>
<td>Re-use previous annotation during coordinate replacement post submission</td>
<td></td>
</tr>
<tr>
<td><strong>Enable upload of NEF format file for NMR restraint data</strong></td>
<td></td>
</tr>
<tr>
<td>Validation report</td>
<td></td>
</tr>
<tr>
<td>Recalculation of archive reports w/ new stats</td>
<td></td>
</tr>
<tr>
<td><strong>Implement EM MAP validation</strong></td>
<td></td>
</tr>
<tr>
<td>Validation report: Improve ligand validation</td>
<td></td>
</tr>
<tr>
<td>Phase 1: Improve ligand handling in the preliminary validation</td>
<td></td>
</tr>
<tr>
<td>Phase 2: Provide quality indicator in ligand 2D depiction</td>
<td></td>
</tr>
<tr>
<td><strong>Phase 3: Provide ligand density map in the validation report</strong></td>
<td></td>
</tr>
<tr>
<td>Improve biocuration pipeline</td>
<td></td>
</tr>
<tr>
<td>Provide method specific view for CIF Editor</td>
<td></td>
</tr>
<tr>
<td>Re-use previous annotation for re-processing</td>
<td></td>
</tr>
<tr>
<td><strong>Improve WF to increase efficiency on processing large structures</strong></td>
<td></td>
</tr>
<tr>
<td>Publication</td>
<td></td>
</tr>
<tr>
<td>Validation paper</td>
<td></td>
</tr>
<tr>
<td>Biocuration paper</td>
<td></td>
</tr>
</tbody>
</table>

*bold items were re-forecast to 2018*
OneDep 2017/2018 Goal Setting

<table>
<thead>
<tr>
<th>Projects</th>
<th>Estimated Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Archive content improvement:</strong></td>
<td></td>
</tr>
<tr>
<td>Author of record file versioning (Parallel versioned ftp tree)</td>
<td></td>
</tr>
<tr>
<td>Phase 2: Enable depositor initiated coordinate replacement</td>
<td>Q4</td>
</tr>
<tr>
<td>Phase 3: Expand versioned FTP tree to include remaining content</td>
<td>Q1</td>
</tr>
<tr>
<td><strong>Archive content improvement:</strong></td>
<td></td>
</tr>
<tr>
<td>Carbohydrate and Protein modification remediation</td>
<td></td>
</tr>
<tr>
<td>Phase 2: Software development and testing</td>
<td>Q2</td>
</tr>
<tr>
<td>Phase 3: Produce test data set for internal and friendly testing</td>
<td>Q3</td>
</tr>
<tr>
<td><strong>Archive content improvement</strong></td>
<td></td>
</tr>
<tr>
<td>RemEDIATE existing XFEL entries according to new mmCIF schema</td>
<td>Q4</td>
</tr>
<tr>
<td>Plan update of archival SF files with new data organization</td>
<td>Q1</td>
</tr>
<tr>
<td><strong>Backend stabilization</strong></td>
<td></td>
</tr>
<tr>
<td>Provide more automated testing (re-usable session test cases)</td>
<td>Q2</td>
</tr>
<tr>
<td>Separate user account from DepUI to support distributed access</td>
<td>Q3</td>
</tr>
<tr>
<td>Validation: Develop wrapper for Mogul API</td>
<td>Q4</td>
</tr>
<tr>
<td>Enable WF to use external computing resources</td>
<td>Q1</td>
</tr>
<tr>
<td>Validation: Modularize reporting, wrapper for RDKit/OpenEye</td>
<td>Q2</td>
</tr>
<tr>
<td><strong>OneDep public facing</strong></td>
<td></td>
</tr>
<tr>
<td>Re-use previous annotation during coordinate replacement post submission</td>
<td>Q3</td>
</tr>
<tr>
<td>Make collection of ORCID ID mandatory</td>
<td>Q4</td>
</tr>
<tr>
<td>Enable upload of NEF format file for NMR restraint data</td>
<td>Q1</td>
</tr>
<tr>
<td>Support XFEL- collect data related to XFEL technique</td>
<td>Q2</td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td></td>
</tr>
<tr>
<td>Ligand Phase 2: Provide geometry quality in 2D depiction</td>
<td>Q3</td>
</tr>
<tr>
<td>Ligand Phase 3: Provide ligand density map in the validation report</td>
<td>Q4</td>
</tr>
<tr>
<td>Ligand Phase 4: Provide Mogul standards in the CCD files</td>
<td>Q1</td>
</tr>
<tr>
<td>Ligand Phase 5: Use Mogul standards in the CCD files for validation</td>
<td>Q2</td>
</tr>
<tr>
<td>Recalculation of archive reports w/ new stats</td>
<td>Q3</td>
</tr>
<tr>
<td><strong>Implement EM map validation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Implement NMR restraint validation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Implement NMR/SAS validation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Improve biocrutation pipeline</strong></td>
<td></td>
</tr>
<tr>
<td>Re-use previous annotation for re-processing</td>
<td>Q4</td>
</tr>
<tr>
<td>Improve miscellaneous UI fixes in Sequence Module</td>
<td>Q1</td>
</tr>
<tr>
<td>Increase efficiency on processing large structures</td>
<td>Q2</td>
</tr>
<tr>
<td>Re-use previous annotation for batch processing</td>
<td>Q3</td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td></td>
</tr>
<tr>
<td>Versioning paper</td>
<td>Q4</td>
</tr>
<tr>
<td>NMR and/or EM validation paper</td>
<td>Q1</td>
</tr>
</tbody>
</table>

* Timeline will be refined further after wwPDB Developer Summit in November 2017
wwPDB Policy Proposals/Discussion

- Resolution of PDB Entry DOIs (Appendix 1)
- wwPDB AC Chair/Co-Chair Restructuring (Appendix 2)
- Invitation to EMDB to Join the wwPDB (Appendix 3)
- Addition of wwPDB Regional Partners (Appendix 4)
- Provision of Mogul Geometry for Ligands (Appendix 5)
- Individual wwPDB Partner AC Reports (Appendix 6)
- Any Requests for Additional Discussion Topics?
Remaining Agenda Items

- Partner Meetings and Outreach: Genji Kurisu
- Macromolecular Crystallography: Stephen K. Burley
- 3D Electron Microscopy: Sameer Velankar
- NMR Spectroscopy: John L. Markley
- Looking Ahead: Sameer Velankar
- Lunch and Executive Session (Noon-1:00pm)
- Questions for the AC: Stephen K. Burley
- Executive Session
- Departure for Dinner in New Brunswick (6:00pm)
Partner Meetings and Outreach

Genji Kurisu
Partner Meetings and Outreach

wwPDB Summit: May EMBL-EBI

OneDep Posters: ACA, ECM, AsCA, APPA, Biocuration Society Meetings

wwPDB PDBx/mmCIF Meeting: July EMBL-EBI

wwPDB Booth: IUCr India
I/HM Book Progress

- Title: "Integrative Structural Biology with Hybrid Methods"
- Publisher: Springer Japan
- Series: Advances in Experimental Medicine and Biology
- Manuscript Submission Deadline: Late 2017
Recent Publications

OneDep Validation Paper In Press Sept 12, 2017
OneDep Biocuration Paper Submitted Oct 11, 2017
World Data System Accreditation
Application for PDB Filed Oct 9 2017

- World Data System (WDS) is an Interdisciplinary Body of the International Council for Science (ICSU; http://www.icsu-wds.org)
- Promotes long-term stewardship of, and universal and equitable access to, quality-assured scientific data and data services, products, and information
- Copies of WDS application available on side table
wwPDB Foundation Progress

- 501(c)(3) entity exclusively for scientific, literary, charitable, and educational purposes
- Ongoing solicitations of Corporate donations
- Individual Membership program launched with limited success

http://foundation.wwpdb.org/
Macromolecular Crystallography

Stephen K. Burley
Agenda

- MX Data Deposition Metrics
- 2015 X-ray VTF Meeting Follow Up
- 2017 PDBx/mmCIF Working Group Meeting Outcome
Growth of Released MX Entries

>120,000 Total Released MX Entries Projected for End 2017
MX Deposition Complexity

Annual Distribution for High Resolution Limit

Total Number of New CCD Entries

Annual Released Structures With AU MW > 500,000

Annual Released Large Structures (chains > 62 & atoms > 99999)
X-ray VTF Meeting Update

- Still waiting for 2015 wwPDB VTF Meeting report and recommendations

- wwPDB is in the process of implementing recommendations from LVW for ligand validation (verbally endorsed by the wwPDB VTF)

VTF Members: Paul Adams, Gérard Bricogne, Dave Brown, Paul Emsley, Richard Henderson, Nobutoshi Ito, Robbie Joosten, Thomas Lütteke, Michael Nilges, Arwen Pearson, Tassos Perrakis, Randy Read (Chair), Jane Richardson, Janet Smith, Tom Terwilliger, Ian Tickle, Gert Vriend

wwPDB Attendees: Burley, Feng, Gutmanas, Velankar, Westbrook
2017 PDBx/mmCIF Working Group

- PDBx/mmCIF is the archival data standard for the repository
- wwPDB together with the PDBx/mmCIF Working Group of community experts and methods developers oversee the evolution of the standard
- Working Group ensures that the standard is well supported by key community software tools.
- PDB hosts community workshops and maintains mmcif.wwpdb.org serving PDBx/mmCIF data dictionaries, schema and software tools
- 2017 PDBx/mmCIF Working Group meeting finalized new content recommendations for diffraction data and ligand refinement restraint data
3D Electron Microscopy

Sameer Velankar
Agenda

- Archiving 3DEM Data in PDB/EMDB
- 3DEM Data Deposition Metrics
- Status of 3DEM Validation
- Engaging the 3DEM Community and Software Providers
Archiving 3DEM Data in PDB/EMDB

- Integration of 3DEM into OneDep was a coordinated and collaborative effort between wwPDB and EMDataBank
- All 3DEM deposition, biocuration and validation services provided by RCSB PDB, PDBe, and PDBj (OneDep)
- 3DEM Structure Atomic Coordinates archived in PDB
- 3DEM Mass Density Maps archived in EMDB
- EMBL-EBI reorganization yielded independent PDBe (Velankar) and EMDB (Patwardhan) Teams

N.B.: At present, there is no formal wwPDB/EMDB agreement re data sharing, security, and release.
Growth of 3DEM Structures in PDB

As of Oct 1, 2017, >1600 3DEM Structures in PDB Archive
371 new structures released January 1 – September 1, 2017
As of Oct 1, 2017, >5100 3DEM Maps in EMDB archive
688 new entries released January 1 – September 1, 2017
3DEM Data Deposition with OneDep

wwPDB Deposition: D_8000220045 -- Requested IDs: PDB, EMDB

This page contains a summary of the uploaded data. Please check that the data content here is correct before proceeding. Data problems that require new data to be uploaded may result in the loss of information entered on subsequent pages.

You uploaded 3 files to the system.

<table>
<thead>
<tr>
<th>Number</th>
<th>Used</th>
<th>File name</th>
<th>Size</th>
<th>File type</th>
<th>File header check</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*</td>
<td>3ivd.pdb</td>
<td>2985903</td>
<td>Coordinates</td>
<td>The file has correct format</td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>EMD-5127.map</td>
<td>2049024</td>
<td>Main volume data</td>
<td>The file has correct format</td>
</tr>
<tr>
<td>3</td>
<td>*</td>
<td>3dem3.jpg</td>
<td>263064</td>
<td>Image</td>
<td>The file has correct format</td>
</tr>
</tbody>
</table>

Model coordinates

Format OK!

Critical Data Values

Cis peptides

The following cis-peptides were detected in your coordinates. Please check these are expected:

| A TRP321 | A PRO322 |
| A PHE57  | A PRO58  |
| A ARG371 | A PRO372 |
| D LEU120 | D PRO121 |
| D ALA426 | D PRO427 |
| F GLU503 | F PRO504 |

-0.8
-1.50
-1.70
-0.75
-1.26
-2.85
### 3DEM Deposition Metrics via OneDep

<table>
<thead>
<tr>
<th>Structures</th>
<th>Total Depositions</th>
<th>Processed by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>2016</td>
<td>513</td>
<td>214</td>
</tr>
<tr>
<td>2017</td>
<td>*389</td>
<td>189</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maps</th>
<th>Total Depositions</th>
<th>Processed by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>2016</td>
<td>1097</td>
<td>477</td>
</tr>
<tr>
<td>2017</td>
<td>*633</td>
<td>307</td>
</tr>
</tbody>
</table>

*As of September 7, 2017
Status of 3DEM Validation

- OneDep manages deposition, biocuration, and limited validation (structures) of all incoming 3DEM data
- EMDB performs limited validation (maps, fit of structures to maps) after data transfer from OneDep
- Wrap Up Workshop for current round of EMDataBank Map and Model Challenges was held in early Oct 2017
- Validation methods development funding sources
  - NIGMS – EMDataBank (Stanford/RCSB)
  - WT – broadly collaborative UK effort (incl. EMDB)
  - EU – INSTRUCT collaboration
Engaging the EM Community I

- Pressing need to formalize coordination of deposition, biocuration, and validation efforts between wwPDB and EMDB

- 2017 PDBx/mmCIF Working Group Meeting Day 2
  - Discussion on EM model deposition and validation requirements
  - Participants included – Paul Adams, Maya Topf, Wah Chiu, Corey Hryc, Garib Murshudov, Carsten Sachse, Martyn Winn, Cathy Lawson, John Westbrook, Stephen Burley, Ardan Patwardhan, Sameer Velankar, Gerard Kleywegt
Engaging the EM Community II

- 2017 PDBx/mmCIF Working Group made progress with software developers on 3DEM-related data matters

- The wwPDB PDBx/mmCIF Working Group will include a 3DEM subcommittee going forward
  - Currently recruiting others involved in development of 3DEM model building software

- wwPDB will collaborate with EM stakeholders to reconvene the EM Validation Task Force to review present status and determine consensus path forward
NMR Spectroscopy

John L. Markley
Agenda

- Archiving NMR Data in PDB/BMRB
- NMR Data Deposition Metrics
- BMRB Activities Related to OneDep
- NMR/SAS Hybrid Method Progress
Biomolecular NMR Data Archiving

- NMR model coordinates and related data are archived in PDB and backed up in BMRB
- NMR data pertaining to structures, but not archived by PDB, are solicited by BMRB
- NMR data not pertaining to structures are archived at BMRB: assignments without structure, information on dynamics, conformational changes, ligand binding, solvent accessibility, etc.
Archiving NMR Data in PDB/BMRB II

- Biocuration and validation services for NMR structures are provided through the OneDep system by RCSB PDB, PDBj, and PDBj
  - Current data content: coordinates, assigned chemical shifts, peak lists (optional), metadata, restraints in native format
  - Current validation: coordinates as for X-ray, chemical shift outliers (software provided by BMRB)
- Entries deposited via OneDep are sent to BMRB, where they are checked by BMRB software and biocurated as needed
## NMR Data Depositions (2014-2017)

<table>
<thead>
<tr>
<th>Year</th>
<th>NMR Depositions to PDB Archive</th>
<th>Depositions of NMR data without Structures</th>
<th>Total BMRB Depositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>515</td>
<td>240</td>
<td>755</td>
</tr>
<tr>
<td>2015</td>
<td>510</td>
<td>333</td>
<td>843</td>
</tr>
<tr>
<td>2016</td>
<td>473</td>
<td>276</td>
<td>749</td>
</tr>
<tr>
<td>2017 (to 9/1)</td>
<td>280</td>
<td>253</td>
<td>533</td>
</tr>
</tbody>
</table>
BMRB Activities Related to OneDep

- BMRB staff members participate in OneDep data development
- NMR Exchange Format (NEF)
  - NEF was developed as a means for exchanging data among software developers and for restraint deposition
  - NMR-STAR files containing restraints and peak lists are available
  - A NEF to NMR-STAR converter has been implemented by BMRB for incorporation into OneDep
- wwPDB now accepts NEF as an alternative deposition format for structural restraints
- BMRB has worked with NEF developers on NEF v1.0 and incorporated feedback into NMR-STAR
  - NEF restraint data must be accompanied by a NMR-STAR v3.1 or later chemical shift assignment file and mmCIF coordinate file with NEF atom nomenclature mapping
Discussed: NEF progress and next steps (6-month, 1-year); validation as part of OneDep; SAS-NMR validation; CryoEM-NMR validation.
Kumaran Baskaran (BMRB-Wisconsin) participates in the Integrative/Hybrid Methods (I/HM) Model Validation Developer Group.

Biomolecular NMR is itself already a “hybrid method” (i.e., structures are derived from multiple experiments).

NMR-STAR dictionary accommodates data from a number of experimental methods complementing NMR:
- SAXS
- H/D exchange
- FRET
- MS
NMR/SAS Hybrid Methods Progress II

- SASCIF extensions have been added to PDBx/mmCIF

- SAS data deposition now supported by OneDep via API calls to SASBDB Deposition User Interface

- SAS data retrieval from SASBDB currently limited SASBDB accession code (to be expanded)

- Validation of NMR/SAS Hybrid structures pending recommendations from NMR and SAS VTFs
Looking Ahead

Sameer Velankar
Plans for the Coming Years I

2017/2018 (OneDep Team)

- Implement Ligand Validation Workshop recommendations (2D depictions and ED maps)
- Enable Data Depositor initiated coordinate update post release (versioned file)
- Collect new data related to XFEL/SFX techniques
- Automate testing of updated OneDep system
- Increase biocuration efficiency
- Implement NMR restraint validation
- Implement EM map validation
Plans for the Coming Years II

2017/2018

- Remediation work continuing
  - Carbohydrates (Lead: RCSB PDB)
  - Post-translational modifications (Lead: PDBe)
- wwPDB Partnership
  - PDBx/mmCIF dictionary management
  - Weekly release process auditing and automation
  - Invitation to EMDB (subject AC concurrence)
  - Negotiations with new Regional wwPDB Partners
- wwPDB AC meeting at PDBe EMBL-EBI
Plans for the Coming Years III

2018/2019

- wwPDB AC meeting at PDBj Osaka University
  - Friday October 18 2019 *versus*
  - Friday October 25 2019

- Planning for PDB 50th Anniversary in 2021 with celebratory scientific meetings and outreach events by the wwPDB and individual wwPDB partners
PDBe to Host 2018 wwPDB AC

- Date: Friday November 2\textsuperscript{nd} 2018

- Location:
  Madingley Hall Cambridge Conference Center
  University of Cambridge, Madingley
  Cambridge CB23 8AQ
  United Kingdom
Lunch and Executive Session
Questions for the wwPDB AC

Stephen K. Burley
Questions for the wwPDB AC

1. Does the wwPDB AC concur with recommendation from wwPDB Partners re DOI resolution of PDB Entries as outlined in Appendix 1?

2. Does the wwPDB AC concur with recommendation from wwPDB Partners re wwPDB Chair/Co-Chair restructuring as outlined in Appendix 2?

3. Does the wwPDB AC concur with recommendation from wwPDB Partners re inviting EMDB to join the wwPDB organization as outlined in Appendix 3?
Questions for the wwPDB AC (cont.)

4. Does the wwPDB AC concur with recommendation from wwPDB Partners re the process for adding new Regional wwPDB Partners as outlined in Appendix 4?

5. Does the wwPDB AC concur with recommendation from wwPDB Partners re providing Mogul geometry standards for ligands as outlined in Appendix 5?

6. Does the wwPDB AC have any questions or concerns regarding the individual RCSB PDB, PDBe, PDBj, or BMRB Advisory Committee reports provided in Appendix 6?
Acknowledgements and Closing Remarks

Stephen K. Burley and R. Andrew Byrd
Slide Preparation and Organization

- Jasmine Young
- Christine Zardecki
- Nicole Oorbeek (nee Malkiewicz)
- RCSB Team Members
- IQB Support Staff
- Pauline Haslam
- Tomoko Shimizu
Meeting Funding

- NSF/NIH/DoE grant to RCSB PDB
- Institute for Quantitative Biomedicine and Rutgers, The State University of New Jersey
- wwPDB Foundation
- NIGMS grant to BMRB
- EMBL-EBI and Wellcome Trust grant to PDBe
- IPR and JST and AMED grants to PDBj
Closing Remarks

Thank you for your enduring support of the Worldwide Protein Data Bank partnership and for taking the time to attend the annual meeting

Safe travels home