Introductions and Overview of the wwPDB

Sameer Velankar
Our vision is to:

Sustain freely accessible, interoperating Core Archives of structure data and metadata for biological macromolecules as an enduring public good to promote basic and applied research and education across the sciences.
wwPDB Mission Statement

- Manage the wwPDB Core Archives as a public good according to the **FAIR** Principles.
- Provide expert deposition, validation, biocuration and remediation services at no charge to Data Depositors worldwide.
- Ensure universal open access to public domain structural biology data with no limitations on usage.
- Develop and promote community-endorsed data standards for archiving and exchange of global structural biology data.
Developments since 2017 - Meeting I

- Very positive peer review of BMRB grant for 2019-2024.
- Plan for orderly leadership change at BMRB.
- Continued enhancement of the OneDep system for deposition/validation/biocuration of MX, NMR, and 3DEM structures.
- Continued growth in 3DEM structure depositions and engagement with the 3DEM community.
- Implementation of PDB archive versioning.
Developments since 2017- Meeting II

- BMRB designated as wwPDB Core Archive.
- EMDB to join as wwPDB Core Archive and as wwPDB Core Member.
- Ardan Patwardhan, EMDB Team leader, will represent EMDB.
- Continued depositions to PDB-Dev for archiving of other I/HM structures.
PDB Core Archive Depositions

- On track for ~12,500 depositions in 2018.
- 3DEM growth continuing in 2018.

<table>
<thead>
<tr>
<th>Method</th>
<th>2017 Depositions</th>
<th>2016 Depositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MX</td>
<td>11,889 (91.2%)</td>
<td>10,583</td>
</tr>
<tr>
<td>NMR</td>
<td>460 (3.5%)</td>
<td>474</td>
</tr>
<tr>
<td>3DEM</td>
<td>658 (5.0%)</td>
<td>531</td>
</tr>
<tr>
<td>Other</td>
<td>44 (0.3%)</td>
<td>27</td>
</tr>
</tbody>
</table>

2017 Processing Sites

- PDBj 21%
- RCSB 48%
- PDB 31%
- PDBe 31%

2017 Depositor Locations

- North America 30%
- Europe 38%
- Asia 1%
- South America <1%
- Oceania 7%
- Africa 1%
- Commercial 3%
PDB Core Archive Growth

* As of 7 September 2018
## PDB Core Archive Downloads

<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>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>
<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>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>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>
</tbody>
</table>

More than 1.8 million/day!

N.B.: Some 2018 data lost due to GDPR. Hope to be back on track for 2019→.
- BMRB is on track for ~870 new entries released for 2018.
- Total released entries estimated to reach ~12,700 by the end of 2018.
## BMRB Core Archive Growth

### Total Released Entries

<table>
<thead>
<tr>
<th>Year</th>
<th>Total released</th>
<th>Yearly increase</th>
<th>Structures</th>
<th>Yearly increase</th>
<th>Non-structures</th>
<th>Yearly increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>8075</td>
<td>814</td>
<td>4852</td>
<td>770</td>
<td>3223</td>
<td>44</td>
</tr>
<tr>
<td>2013</td>
<td>8895</td>
<td>820</td>
<td>5647</td>
<td>795</td>
<td>3248</td>
<td>25</td>
</tr>
<tr>
<td>2014</td>
<td>9884</td>
<td>989</td>
<td>6614</td>
<td>967</td>
<td>3270</td>
<td>22</td>
</tr>
<tr>
<td>2015</td>
<td>10343</td>
<td>459</td>
<td>7025</td>
<td>411</td>
<td>3318</td>
<td>48</td>
</tr>
<tr>
<td>2016</td>
<td>11136</td>
<td>793</td>
<td>7649</td>
<td>624</td>
<td>3487</td>
<td>169</td>
</tr>
<tr>
<td>2017</td>
<td>11851</td>
<td>715</td>
<td>8125</td>
<td>476</td>
<td>3726</td>
<td>239</td>
</tr>
<tr>
<td>2018</td>
<td>12430</td>
<td>579</td>
<td>8411</td>
<td>286</td>
<td>4019</td>
<td>293</td>
</tr>
<tr>
<td>2018 est.</td>
<td>12720</td>
<td>869</td>
<td>8554</td>
<td>429</td>
<td>4166</td>
<td>440</td>
</tr>
<tr>
<td>Year</td>
<td>Server requests</td>
<td>Page requests</td>
<td>File requests</td>
<td>Distinct hosts served</td>
<td>Total data transferred</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>---------------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>12,703,408</td>
<td>3,290,436</td>
<td>1,730,857</td>
<td>310,074</td>
<td>8.56 TB</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>10,521,902</td>
<td>2,176,005</td>
<td>1,508,510</td>
<td>342,291</td>
<td>4.85 TB</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>11,314,520</td>
<td>2,193,844</td>
<td>1,561,558</td>
<td>383,052</td>
<td>6.99 TB</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>12,869,845</td>
<td>2,348,759</td>
<td>1,991,915</td>
<td>444,873</td>
<td>12.6 TB</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>15,986,881</td>
<td>2,766,144</td>
<td>2,496,663</td>
<td>445,679</td>
<td>12.0 TB</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>42,071,167</td>
<td>10,541,522</td>
<td>6,633,980</td>
<td>460,550</td>
<td>13.1 TB</td>
<td></td>
</tr>
<tr>
<td>2018 (by 9/12)</td>
<td>40,014,144</td>
<td>10,303,509</td>
<td>3,949,570</td>
<td>344,560</td>
<td>7.61 TB</td>
<td></td>
</tr>
</tbody>
</table>

~150K/day Server and page requests
# BMRB Core Archive Growth

## Internet Servers Traffic (FTP servers) – All Mirrors

<table>
<thead>
<tr>
<th>Year</th>
<th>Server requests</th>
<th>Distinct files requested</th>
<th>Distinct hosts served</th>
<th>Total data transferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>2,304,687</td>
<td>1,820,240</td>
<td>4,788</td>
<td>1.15 TB</td>
</tr>
<tr>
<td>2013</td>
<td>1,992,706</td>
<td>1,481,995</td>
<td>5,253</td>
<td>1.40 TB</td>
</tr>
<tr>
<td>2014</td>
<td>1,994,184</td>
<td>1,488,431</td>
<td>4,645</td>
<td>1.59 TB</td>
</tr>
<tr>
<td>2015</td>
<td>2,197,159</td>
<td>1,666,353</td>
<td>3,742</td>
<td>0.905 TB</td>
</tr>
<tr>
<td>2016</td>
<td>5,459,161</td>
<td>1,655,031</td>
<td>5,774</td>
<td>1.66 TB</td>
</tr>
<tr>
<td>2017</td>
<td>5,502,929</td>
<td>2,461,117</td>
<td>4,126</td>
<td>5.02 TB</td>
</tr>
<tr>
<td>2018 (by 9/12)</td>
<td>8,168,965</td>
<td>2,950,997</td>
<td>3,778</td>
<td>4.72 TB</td>
</tr>
</tbody>
</table>
EMDB Core Archive Depositions

- On track for ~1400 3DEM depositions in 2018

<table>
<thead>
<tr>
<th>Processing Site</th>
<th>2018 Depositions</th>
<th>2017 Depositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDBj</td>
<td>90</td>
<td>162</td>
</tr>
<tr>
<td>PDBe</td>
<td>336</td>
<td>516</td>
</tr>
<tr>
<td>RCSB</td>
<td>828</td>
<td>667</td>
</tr>
</tbody>
</table>

2017 Processing Sites:
- RCSB 50%
- PDBj 12%
- PDBe 38%
EMDB Core Archive Growth

* As of 2 October 2018
Core Member Funding Status

- BMRB: NIH NIGMS funding → 03/31/2019; renewal pending.
- PDBe: EMBL-EBI, Wellcome Trust → 01/01/2020.
- PDBj: NBDC-JST and AMED funding → 03/31/2022.
- EMDB: EMBL-EBI, Wellcome Trust → 01/01/2024.
wwPDB Outreach

2018 OneDep Developer Summit

IUCr

ACA

ECM

NMR

3DEM

AsCA –
2-5 December 2018
PDB NAR Database Issue Paper published on 24th October 2018

https://doi.org/10.1093/nar/gky949
Website improved.

Fundraising ongoing.

2017 events:
- WT meeting on Sustainability.

2018 events:
- Tokyo meeting on Sustainability.

http://foundation.wwpdb.org/
## wwPDB Collaboration Resource Commitments
### October 2017 - September 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></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.0</td>
<td>13.8</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></td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>7.7</td>
<td>7.7</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>
<tr>
<td></td>
<td>6.65*</td>
<td></td>
<td></td>
<td></td>
<td>14.3*</td>
<td>29.05*</td>
</tr>
</tbody>
</table>

* loss of committed resource
# OneDep 2017/2018 Progress vs. Goals

<table>
<thead>
<tr>
<th>Projects</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2017 Q4</td>
</tr>
<tr>
<td><strong>1. Validation</strong></td>
<td></td>
</tr>
<tr>
<td>1.1: Annual recalculation of validation reports</td>
<td></td>
</tr>
<tr>
<td>1.2: NMR restraint validation</td>
<td></td>
</tr>
<tr>
<td>1.3: Provide ligand 2D geometrical quality depiction</td>
<td></td>
</tr>
<tr>
<td>1.4: Provide ligand ED maps</td>
<td></td>
</tr>
<tr>
<td><strong>2. Backend Stabilization</strong></td>
<td></td>
</tr>
<tr>
<td>2.1: Towards automated testing framework</td>
<td></td>
</tr>
<tr>
<td>2.2: Improve session management</td>
<td></td>
</tr>
<tr>
<td>2.3: User account management</td>
<td></td>
</tr>
<tr>
<td><strong>3. OneDep public facing</strong></td>
<td></td>
</tr>
<tr>
<td>3.1: Retain Previous Annotation During Coordinate Replacement Post Submission</td>
<td></td>
</tr>
<tr>
<td>3.2: Enable NEF upload for NMR restraint validation</td>
<td></td>
</tr>
<tr>
<td>3.3: Extend metadata collection to support SFX/XFEL</td>
<td></td>
</tr>
<tr>
<td>3.4: Mandatory ORCID for contact authors</td>
<td></td>
</tr>
<tr>
<td>3.5: Enable depositor-initiated coordinate replacement post release</td>
<td></td>
</tr>
<tr>
<td>3.6: Enhance cross validation and simplify DepJUI</td>
<td></td>
</tr>
<tr>
<td><strong>4. Biocuration</strong></td>
<td></td>
</tr>
<tr>
<td>4.1: Retain previous annotation for entry reprocessing</td>
<td></td>
</tr>
<tr>
<td>4.2: Improve overall efficiency</td>
<td></td>
</tr>
<tr>
<td>4.3: Enable WF to use external computing resources</td>
<td></td>
</tr>
<tr>
<td>4.4: Apply same annotation to identified multi-related entries</td>
<td></td>
</tr>
<tr>
<td><strong>5. Archive Improvements</strong></td>
<td></td>
</tr>
<tr>
<td>5.1: XFEL remediation</td>
<td></td>
</tr>
<tr>
<td>5.2: Carbohydrate remediation</td>
<td></td>
</tr>
<tr>
<td>5.3: Protein Modification remediation</td>
<td></td>
</tr>
<tr>
<td>5.4: Expand content at versioned FTP</td>
<td></td>
</tr>
<tr>
<td>5.5: Improve release process</td>
<td></td>
</tr>
<tr>
<td><strong>6. EMDB related projects</strong></td>
<td></td>
</tr>
<tr>
<td>6.1 Improve data field validation with new EMDB schema</td>
<td></td>
</tr>
<tr>
<td>6.2 Extend EMDB accession code</td>
<td></td>
</tr>
<tr>
<td><strong>7. Software Upgrade</strong></td>
<td></td>
</tr>
<tr>
<td>7.1 Deposition Software</td>
<td></td>
</tr>
</tbody>
</table>

*Delivered, Work in progress, Bold: Re-forecast to 2019/2020, Added to original plan*
The wwPDB has recently produced an extension to the PDBx/mmCIF dictionary to incorporate multiple crystal data collection techniques used in serial femtosecond crystallography (SFX) and X-ray free electron laser (XFEL) experiments.

An agreed process for dictionary updates.

All wwPDB sites involved in dictionary updates.

Invitation to EMDB.

Restructuring of wwPDB agreement.
Biocuration Performance

New Structures/wwPDB Biocurator

** launch of OneDep
2017/2018 Progress vs. Goals I

OneDep projects re-forecasted to 2018/2019

- Implement Ligand Validation Workshop recommendations (2D depictions and ED maps).
- Implement NMR restraint validation.
- Implement EM map validation.
- Implement Author-initiated Coordinate Replacement.
- Use ORCiD for User Authentication.

Mitigations

- Upgrade third party software in 2018 to enable new feature development in 2019.
- Cross-site resource sharing for validation projects.
2017/2018 Progress vs. Goals II

- Remediation work continuing
  - Carbohydrates (Lead: RCSB PDB).
  - Post-translational modifications (Lead: PDBe).

- wwPDB Partnership (re-forecasted to 2018/2019)
  - Weekly release process auditing and automation.
  - Resolution of DOI to a single web page on wwPDB site with links to all wwPDB partner sites.
  - Exploring possible new Regional wwPDB Partners.
Updating the wwPDB
Vision, Mission, and Scope

Stephen K. Burley
Drivers for Updating the wwPDB Vision, Mission, and Scope

- Structural biology is much more than just MX/NMR.
- Structural biology techniques are no longer being used one at a time.
- Contributions from Integrative/Hybrid Methods to PDB-Dev are increasing.
- Data archiving to multiple specialized structural biology data resources needs to be coordinated.
- Role of BMRB archive in wwPDB was never explicitly defined when wwPDB was expanded in 2005.
- wwPDB and EMDB are exchanging data without a formal agreement that protects their Depositors.
- wwPDB and SASBDB are exchanging data without a formal agreement that protects their Depositors.
wwPDB Guiding Principles

- Structural biology is an international science.
- Secure storage and preservation of global structural biology data is essential to the long-term success of the discipline.
- Open access to global structural biology data without limitations on usage is essential to basic and applied research and education across the sciences.
- Global structural biology data represent public goods that require joint international management.
- wwPDB partnership provides an effective mechanism for joint international management of global structural biology data.
- The multilateral wwPDB Agreement is periodically reviewed, revised, and reaffirmed.
The New Vision of the wwPDB is to

Sustain freely accessible, interoperating Core Archives of structure data and metadata for biological macromolecules as an enduring public good to promote basic and applied research and education across the sciences.
The New Mission of the wwPDB is to

- Manage the wwPDB Core Archives as a public good according to the **FAIR** Principles.
- Provide expert deposition, validation, biocuration, and remediation services at no charge to Data Depositors worldwide.
- Ensure universal open access to public domain structural biology data with no limitations on usage.
- Develop and promote community-endorsed data standards for archiving and exchange of global structural biology data.
wwPDB Aspirational Architecture

CORE ARCHIVES
- PDB
- BMRB
- EMDB

CORE MEMBERS
- RCSB PDB
- PDBe, PDBj, BMRB, EMDB

FEDERATED RESOURCES
- MX Images
- EMPIAR
- SASBDB
wwPDB Core Archives

Definition: A wwPDB “Core Archive” is a global structural biology data resource jointly managed by wwPDB Core Members, with one member designated as the “Archive Keeper”.

- Current wwPDB Core Archives:
  - PDB Core Archive: **3D Structure Data Resource** housing multiscale/atomic structural models plus molecular data and metadata, MX experimental data and metadata, and other experimental data.
  - BMRB Core Archive: **Biomolecular NMR Data Resource** housing molecular data and metadata, NMR experimental data and metadata, and other experimental data.

- Next Core Archive expected to join wwPDB:
  - EMDB Core Archive: **Molecular and Cellular EM Data Resource** housing molecular/biological data and metadata, experimental electric potential map data, and other experimental data.
wwPDB Archive Keepers

Role: A wwPDB “Archive Keeper” is a wwPDB Core Member appointed, according to the terms of the wwPDB Agreement, as having primary responsibility for data storage, preservation, composition, representation, and maintenance of a particular Core Archive, carried out in concert with other wwPDB Core Members.

- **Current wwPDB Archive Keepers**
  - PDB Archive Keeper: RCSB PDB
  - BMRB Archive Keeper: BMRB

- **Expected addition to wwPDB**
  - EMDB Archive Keeper: EMDB
wwPDB Core Members

- Role: A wwPDB “Core Member” contributes to all wwPDB partnership activities according to the terms of the wwPDB Agreement.

- Current wwPDB Core Members:
  - RCSB PDB
  - PDBe
  - PDBj
  - BMRB

- Expected addition to wwPDB:
  - EMDB
wwPDB Associate Members

- Role: A wwPDB “Associate Member” contributes to some of the wwPDB partnership activities according to the terms of the wwPDB Agreement, with the goal of becoming a wwPDB Core Member.

- Current wwPDB Associate Members:
  - None

- Expected additions to wwPDB:
  - PDB-China
  - PDB-India
wwPDB Advisory Committee

- **Role**
  - Existing Terms of Reference under revision (Appendix 4)

- **Current wwPDB AC Roster**
  - Core Member Representatives of RCSB PDB, PDBe, PDBj, and BMRB (2 each)
  - Organizational Representatives of IUCr, ICMRBS, and EM-TBN (1 each)
  - National Representatives: China and India (1 each)
  - Institutional Representatives (1 per organization)

- **Expected changes to wwPDB AC**
  - Core Member Representatives of EMDB (2)
  - Observers from PDB-China, PDB-India (1 or 2 each, replacing National Representatives)
Definition: A wwPDB “Federated Resource” is a structural biology data resource that participates in data exchange with one or more wwPDB Core Archives, under a formal agreement governing data exchange and confidentiality, etc.

Current Federated Resources:
- None

Federated Resources expected to align with wwPDB:
- SASBDB
- EMPIAR
wwPDB Federated Members

- Role: A wwPDB Federated Member manages one or more wwPDB Federated Resources and collaborates with the wwPDB Core Members in developing and maintaining data exchange infrastructure.

- Current Federated Members:
  - None

- Federated Members expected to align with wwPDB:
  - SASBDB
  - EMPIAR
wwPDB Community Engagement

Sameer Velankar
wwPDB Community Engagement: Historical Drivers

- Establishing data standards and improving data quality of the archived data.
- Consulting with community experts
  - PDBx/mmCIF Working Group – Developing data standards for representation of structure data.
  - Validation and other Task Forces – Recommendations from community experts on improving quality of experimental data, metadata, and structures.
  - Workshops – Recommendations from community experts on improving co-crystal structures.
Impact of Changes to wwPDB

- Multiple Core Archives – PDB, BMRB and EMDB
- Each Core Archive will require community input on data standards and validation.
- Number of structures using multiple data types is increasing.
- Data standards and data deposition, validation, and biocuration are all interconnected.
- Need an integrated approach to data standards and data deposition, validation, and biocuration across all wwPDB Core Archives.
Plan to establish a wwPDB Steering Group responsible for oversight of data standards and data deposition, validation, and biocuration requirements across all Core Archives.

- wwPDB Core Member PIs will join the steering group
- Each wwPDB Core Member will appoint one qualified external individual and one staff member to the Steering Group.
- wwPDB Core Members will jointly appoint a Chair for a 3-year term (renewable).
- One wwPDB Core Member PI will serve as rotating Co-chair (1 year term)
Steering Group and wwPDB Core Members will agree on priorities and convene teams of qualified experts (Expert Advisory Groups) from the community to address issues and provide recommendations as needed:

- The Expert Advisory Groups (EAGs) will publish recommendations in a peer reviewed journal.
- Steering Group will meet quarterly and publish meeting minutes on the wwPDB website.
- Steering Group will report annually to the wwPDB Advisory Committee on progress from all EAGs.
PDB Core Archive plans

Genji Kurisu
wwPDB Collaboration Resource Commitments
November 2018-October 2019

<table>
<thead>
<tr>
<th>wwpDB Partner</th>
<th>Software Development</th>
<th>Production Maintenance/Management</th>
<th>Requirements Setting/Testing</th>
<th>*Core Archive Keeping</th>
<th>Outreach</th>
<th>Biocuration/Remediation</th>
<th>Total FTE Commitments</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCSB PDB</td>
<td>3.0</td>
<td>1.8</td>
<td>0.35/0.35</td>
<td>2.0</td>
<td>0.2</td>
<td>6.0</td>
<td>13.7</td>
</tr>
<tr>
<td>PDBj</td>
<td>1.6</td>
<td>1.0</td>
<td>0.35/0.35</td>
<td>-</td>
<td>0.2</td>
<td>4.2</td>
<td>7.7</td>
</tr>
<tr>
<td>PDBj</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2/0.2</td>
<td>-</td>
<td>0.2</td>
<td>4.2</td>
<td>5.6</td>
</tr>
<tr>
<td>BMRB</td>
<td>1.5</td>
<td>-</td>
<td>-/-</td>
<td>1.5</td>
<td>0.2</td>
<td>0.2</td>
<td>3.9</td>
</tr>
<tr>
<td>EMDB</td>
<td>1</td>
<td>0.1</td>
<td>-/-</td>
<td>0.9</td>
<td>-</td>
<td>0.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Total wwpDB</td>
<td>7.5</td>
<td>3.3</td>
<td>0.9/0.9</td>
<td>4.4</td>
<td>0.8</td>
<td>15.1</td>
<td>33.4</td>
</tr>
</tbody>
</table>

* Resource from Archive Keeper

Additional resources for software development made available by PDBj, EMDB and BMRB.
## OneDep 2018/2019 Goal Setting I

<table>
<thead>
<tr>
<th>Major Projects</th>
<th>Primary resource*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supporting NMR data in NEF format within OneDep</td>
<td>BMRB</td>
</tr>
<tr>
<td>NMR restraint validation (Includes modularizing validation software with unit tests)</td>
<td>BMRB</td>
</tr>
<tr>
<td>Content improvement in the EM validation report</td>
<td>EMDB</td>
</tr>
<tr>
<td>PTM remediation (Setting remediation plan)</td>
<td>PDBe</td>
</tr>
<tr>
<td><strong>Coord replacement post release (Includes User Authentication with ORCiD)</strong></td>
<td>PDBe: DepUI</td>
</tr>
<tr>
<td></td>
<td>RCSB PDB: API</td>
</tr>
<tr>
<td>Annual validation recalculation</td>
<td>PDBe</td>
</tr>
<tr>
<td>ED map coefficients</td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>Simplify DepUI backend software code and add unit tests</td>
<td>PDBe</td>
</tr>
<tr>
<td>Carbohydrate remediation (Software development)</td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>Biocuration software upgrade (python 3, unit tests, etc.)</td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>CCD versioning</td>
<td>RCSB PDB</td>
</tr>
<tr>
<td><strong>Enhance ligand validation (Includes modularizing validation software with unit tests)</strong></td>
<td>RCSB PDB</td>
</tr>
<tr>
<td>Biocuration and DepUI improvements (Includes Biocuration software upgrade and unit tests)</td>
<td>RCSB PDB: Biocuration</td>
</tr>
<tr>
<td></td>
<td>PDBe: DepUI</td>
</tr>
</tbody>
</table>

**Bold:** re-forecasted from 2017/2018

* Cross-site resource sharing identified for primary and supporting resource
* Timeline will be further refined after requirement setting.
PDB Core Archive Plans Specific to MX

- **ED map coefficients**: As part of ligand validation development, the ED map coefficients will be provided at the wwPDB FTP site, which will be implemented in 2019 for better interpretation of the ligand validation by user community.
Remediation of PDB Archive

- Remediation is a continuous process to ensure better consistency and searchability.

- Representation of carbohydrates led by RCSB PDB, and post-translational modifications (PTM) led by PDBe will be the focus of PDB Core Archive plans in 2019 and 2020.
  - 2019: Requirement setting for PTM and software development for carbohydrates.
  - 2020: Software development for PTM.
Immediate Objectives for Versioning (1)

- **Ligand improvements**: Recently, validation of ligands in the PDB Core Archive was improved by the adoption of a more robust way of flagging those molecules that do not fit electron density well. LLDF (Local Ligand Density Fit) has been replaced by a combination of RSR and RSCC since March 2018.

- **Global Phasing ligand validation code**: Their Ligand Validation Buster Report code is being evaluated by the wwPDB OneDep Software Development Team.
  - The code review has determined that the code is suitable for use in OneDep.
Immediate Objectives for Versioning (2)

- **Depositors’ corrections:** Depositors will be able to make corrections to existing structures in the PDB Core Archive by updating the atomic coordinates while preserving the original PDB identifier. The recent introduction of versioning makes this long desired opportunity possible.
  - In the first year, we will restrict this to improving co-crystal structures to understand the impact.
  - Only original depositor of an entry can provide updates.
  - Number of updates for each depositor restricted:
    - 1 coordinate replacement per entry per year.
    - 3 entries update per PI per year.
Use of ORCiD in OneDep

- To enable the depositors’ corrections, ORCiD ID(s) for contact author(s) became mandatory in 2018. To further improve the depositor experience and enable better management of incoming data, OneDep protocols will be changed to allow login using ORCiD in 2019.
- Receiving requests from depositors to back-populate ORCiD for released entries (updates underway).
BMRB Core Archive Plans

John Markley
BMRB Archive to Become a wwPDB Core Archive

- The BMRB will work with the other wwPDB partners to develop cross-links to data in BMRB associated with structures in the PDB archive.

- As a step towards interoperability between the BMRB small molecule archive and the wwPDB Chemical Components Dictionary (CCD), BMRB has provided the OneDep developers an analysis of the CCD that provides standard InChI strings for compounds that lack them and full InChI compliant atom designators (as generated by ALATIS\(^1\)) as synonyms for current atom designators.

\(^1\)Sci Data. 2017 4:170073. doi: 10.1038/sdata.2017.73
BMRBdep will be Launched

- BMRBdep will replace ADIT-NMR at the Madison and Osaka branches of BMRB and will provide an API to facilitate the acquisition of NMR data.
- The goal will be development of seamless interaction between OneDep and BMRBdep for the benefit of depositors of structures utilizing NMR data.
- The BMRBdep API will guide the collection and validation of additional NMR data that is not currently collected in OneDep.
- BMRB will lead the efforts to improve validation and scope of data currently collected in OneDep.
- Validation of NMR data against coordinates will be carried out by OneDep using standards and software developed within the OneDep team.
- BMRBdep will continue to handle NMR-based structures of peptides, natural products, and small molecules that do not fit PDB criteria (e.g., small peptides and oligosaccharides).
Goals for BMRB Archive Re-Scoping

- BMRB is working with experts in a number of growth areas with the objective of capturing data from a higher percentage of NMR publications across a wider range of biomolecular applications:
  - Dynamics (Arthur Palmer)
  - Integrative/Hybrid Methods (wwPDB I/H Methods Task Force)
  - Intrinsically Disordered Proteins (Julie Forman-Kay)
  - Solid-state NMR (Chad Rienstra)
  - Membrane Proteins (Gianluigi Veglia)
  - Metabolomics (Teresa Fan)
  - Molecular Interactions (Giuseppe Melancini)
  - Natural Products (Arthur Edison)
  - NMR Software Standards (Wim Vranken)
  - Proteins (Ichio Shimada)
  - RNA (Hashim al Hashimi)
Facilitating NMR Data Deposition

- BMRB is working with the NMRbox projects on capturing information from biomolecular NMR workflows and studies in forms that can flow seamlessly into BMRBdep.
  - BMRB is working with NMRbox to promulgate the use of NMR-STAR throughout computational workflows, not merely at the end-stage.
  - BMRB and NMRbox will collaborate with other software developers as needed, to generate NMR-STAR output.
  - Restraint data from software packages that output NEF will be archived at BMRB and translated into NMR-STAR for deposition using BMRBdep and PDB.
The competitive grant renewal application, which was reviewed favorably, lists John Markley (UW-Madison) and Jeff Hoch (UConn Health) as Co-PIs and Co-Heads (Pedro Romero is the BMRB Director).

Grant will be administered through UW-Madison.

John Markley expects to become an Emeritus on 1 June 2020. At that point, Jeff Hoch will become the sole PI. The project will remain joint between UW-Madison and UConn Health with Pedro Romero continuing as BMRB Director.

John Markley will continue to be associated with BMRB as an Emeritus Professor on a voluntary basis and will provide advice and assistance as needed.
Questions for the wwPDB AC

Sameer Velankar
Questions for the wwPDB AC

1. Does the wwPDB AC concur with updated vision and mission statements as outlined in Appendix 1?

2. Does the wwPDB AC concur with our proposal to invite the SASBDB archive team at the EMBL-Hamburg to join the wwPDB organization as a Federated Member and for the SASBDB Archive to be recognized as a Federated Archive by establishing an agreement as outlined in Appendix 2?

3. Does the wwPDB AC concur with our proposal to establish a wwPDB Data Archiving Steering Group as outlined in Appendix 3?
Questions for the wwPDB AC (cont.)

4. Does the wwPDB AC concur with the updated Terms of Reference for the wwPDB Advisory Committee as outlined in Appendix 4?

5. What is the wwPDB AC members guidance regarding community representation for the 3DEM community (Appendix 5)?

6. What is the wwPDB AC members guidance on the comment period for the policy change of making PDBx/mmCIF files mandatory for deposition of X-ray structures as outlined in Appendix 6?
Questions for the wwPDB AC (cont.)

7. Does the wwPDB AC have any questions or concerns regarding the individual RCSB PDB, PDBe, PDBj, or BMRB Advisory Committee reports provided in Appendix 7?