Meeting Participants

- Advisory Committee
  - Participating: Cynthia Wolberger (Chair), Paul Adams, Peter Andolfatto, Judy Blake, Andy Byrd, Bridget Carragher, Wah Chiu, Kirk Clark, Paul Craig, Roland Dunbrack, Cathy Peishoff, Sue Rhee, Torsten Schwede, Jill Trewhella
  - Absent: Robert B. Darnell, Paul Falkowski, Thomas Ferrin, Andrej Sali*

- RCSB PDB
  - Rutgers: Stephen K. Burley, Helen M. Berman, John Westbrook, Jasmine Young, Christine Zardecki
  - UCSD: Cole H. Christie

*Sali/UCSF will formally join RCSB PDB in 2019
Highlights: 2017 - Present
Year in the Life of the RCSB PDB Community

13,049 structures deposited into the PDB
New structures added to the archive for a total of 136,472 entries
Over 1 million unique users served
>679 million data files downloaded from wwPDB web and FTP sites

- Molecule of the Month on Biodegradable Plastic
- Annual IQB Boot Camp Single Particle Cryo-Electron Microscopy
- Rutgers Undergraduate Course on Antimicrobial Resistance
- ISB International Society for Biocuration
- ACS Chemistry for Life
- IUCN Hydropower
- ISMB ISMB 2017 Prague July 21-26
- ECCB
- 4th Annual Video Challenge Results Molecular View of Diabetes Treatment and Management
- RCSB PDB AC Meeting
- wwPDB Summit
## Responses to 2017 RCSB PDB AC Report

<table>
<thead>
<tr>
<th>Committee strongly encourages the RCSB leadership to use the renewal as an opportunity to explain the significance of each activity and how as an integrated whole they address the needs of the research, industry and education communities.</th>
<th>Series of posters/flyers documenting RCSB PDB impact and support for federal funding agency goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCSB PDB impact analyses published in <em>Protein Science, Scientific Data</em></td>
<td>PDB Impact on Recent US FDA Drug Approvals now in press at <em>Structure</em></td>
</tr>
</tbody>
</table>

| The Committee also encourages the RCSB PDB to aggressively pursue new sources of support by approaching private foundations, pharmaceutical companies and NIH institutes that utilize RCSB PDB resources but do not currently provide funding. | Ongoing; Conversations initiated with
- HHMI
- NCI
- Science Philanthropy Alliance
- Science Gateways Community Institute |
RCSB PDB: Four Interoperating Services

1. Deposition/Biocuration
   - Deposition
   - Validation
   - Biocuration

2. Archive Management/Access
   - Data standards
   - Data integration
   - Data storage
   - Data access

3. Data Exploration
   - Portal
   - Search
   - Browse
   - 3D visualization

4. Outreach/Education
   - PDB-101

Customer Service Help Desk and IT Support
1. **Deposition/Biocuration in 2017**

- On track for ~12,100 depositions in 2018
- 3DEM growth continuing in 2018

### 2017 Processing Sites

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

### 2017 Depositor Locations

- North America: 30%
- Europe: 38%
- Asia: 21%
- Oceania: <1%
- South America: 7%
- Africa: 1%
- Commercial: 1%

### Method 2017 Depositions

<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>
1. Deposition/Biocuration in 2018

- **OneDep**
  - ORCiD now mandatory
  - Biocuration more efficient
  - Supports SFX/XFEL entries
  - Better software management via GitHub

- **Carbohydrate Remediation**
  - Collaboration with glycoscience community
  - Project announced at wwpdb.org
  - PDBx/mmCIF Dictionary extension and example files available via GitHub
1. *Deposition/Biocuration in 2019*

- Ligand Validation enhancement
- NMR Restraint Validation implementation
- Author-initiated Coordinate replacement
- Carbohydrate remediation
- Chemical Component versioning
- Ongoing Deposition/Biocuration efficiency improvement
RCSB PDB Data Architecture Redesign

PDBx/mmCIF Data Schema Throughout!
2. Archive Management/Access in 2018

- Extended PDBx/mmCIF data schema across all four RCSB PDB services
- Integrated Archive Management/Access and Data Exploration by developing new APIs (Application Program Interface) and Web Services
- Legacy search and data delivery infrastructure replaced by cloud friendly technologies (in beta)
  - Search indexing and suggestions (Apache Solr)
  - Archiving services/updates transitioned to a distributed object store (MongoDB)
  - Data Access services transitioned to GraphQL API
  - Specialized search (Sequence & 3D) features re-packaged as independent Web Services
2. Archive Management/Access in 2019

- Upgrade Archive Management data storage system
- Continue to productionize new service architecture in support of the new RCSB.org website design and expanded programmatic data access
- Continue targeted remediation (carbohydrates) and extended data integration (PubChem & CARD)
- Continue cloud migration of the weekly update operations
- Migrate services to a more portable packaging using Docker
3. Data Exploration in 2017

RCSB.org Users

- >395,000 monthly, >1 million annually
- 3% annual growth in non-bounce unique users

RCSB.org Sessions

- 35% growth since 2010
- High average session duration (~6 minutes)
- Low fraction of 0-second “bounce” sessions

Global PDB Data Downloads

- Total: 679,421,200 total
  - FTP: 454,723,083
  - Websites: 224,698,117

Potassium Channel (PDB 1bl8)
Doyle et al. (1998) Science 280, 69-77

Frequently accessed structure—Structure data downloaded ~281K times since 2007
Cited >4700 times
3. **Data Exploration in 2018**

- Solr text search functionality implemented on RCSB.org (piloted on PDB101.RCSB.org)

- New NGL visualization features
  - Electron density maps
  - Ligand-protein interactions
  - Validation report in 3D

- New website architecture designed/developed
  - Improves speed and scaling of existing services
  - Accelerates software development of new services
3. Data Exploration in 2019

- New website design utilizing APIs for delivery of data to RCSB.org users
- Same APIs supporting programmatic access to RCSB.org data for power users, external resources
- New website capabilities supporting
  - Enhanced searching (Solr plus other data types)
  - Auto Suggest, Drill Down, and Advanced Search
  - Tabular Reporting
  - Batch Data Download
- Mol* (mol-star) community graphics library for increasing/extending NGL capabilities
4. Outreach/Education in 2017/2018/2019

- >620K PDB-101 Users in 2017
- Health Focus: Diabetes, Antibiotic Resistance
  - Video Challenge
  - Curricular materials
  - Global Health Resources

>313K views since 2017

>142K views since 2017

What is an Enzyme?
>142K views since 2017
Any Questions About Recent Milestones?
wwPDB AC Meeting November 2, 2018

- Introduced Chair-Elect (Peter Rosenthal, UK)
- Reviewed 2017 metrics
- Reviewed 2017/2018 progress versus goals
- Described new wwPDB organizational structure
- Explained new features of revised wwPDB Charter (to take effect January 1st 2019)
- Outlined 2018/2019 goals
- Obtained concurrence on various policy matters
- Thanked outgoing Chair (Andy Byrd, US)
New wwPDB Organizational Structure

**CORE ARCHIVES**
- PDB
- BMRB
- EMDB

**CORE MEMBERS**
- RCSB PDB
- PDBj

**FEDERATED RESOURCES**
- MX Images
- EMPIAR
- SASBDB
- EMDB
Definition: A wwPDB “Core Archive” is a global structural biology data resource jointly managed by wwPDB Core Members.

- 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. (Archive Keeper: RCSB PDB)
  - **BMRB Core Archive**: Biomolecular NMR Data Resource housing molecular data and metadata, NMR experimental data and metadata, and other experimental data. (Archive Keeper: BMRB)

- 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. (Archive Keeper: EMDB)
Any Questions About wwPDB AC?
Discussion Topics
Urgent Matters

- Fundraising: Other suggestions?

- Membership Transitions
  - Chair: 2019 - 2021
  - NSF Review Panel suggested *inclusion of additional dimensions of diversity, especially ... members from underrepresented communities, and experience in diverse organization types*

- RCSB PDB AC Meeting schedule 2019 and beyond
Membership Transitions

- **New Chair: Paul D. Adams**
  - Division Director
    Molecular Biophysics & Integrated Bioimaging
    Lawrence Berkeley National Laboratory

- **New Member: Mandë Holford**
  - Associate Professor
    Department of Chemistry and Biochemistry
    Hunter College Belfer Research Building and
    CUNY Graduate Center
  - Research Associate
    Sackler Institute for Comparative Genomics
    Invertebrate Zoology
    American Museum of Natural History
RCSB PDB AC Meeting Schedule

- Our Fall Meetings conflict with wwPDB AC

- Proposed meeting plan for 2019-2023
  - Spring 2019
    - Target Window Monday April 1 - Thursday April 4
  - Spring 2020
  - PDB50 in 2021
  - Repeat Washington, DC area meeting to enable program officer participation in Spring 2022
Planning ongoing for PDB 2021
Many Thanks to the RCSB PDB AC

- Comments on the renewal proposal were much appreciated
- Feedback on our May 2018 Site Visit presentations also provided significant benefit
- Look forward to your ongoing feedback on
  - New 2019 RCSB.org website design
  - SFX/XFEL
  - 3DEM (single-particle and tomography)
  - Integrative/Hybrid Methods

- Your help with fundraising activities going forward
Many Thanks to Cynthia Wolberger for 10 Years of Advice and Support

RCSB PDB AC member since 2009

RCSB PDB Chair, 2013-2018

RCSB PDB AC 2010

Celebration of Open Access in Structural Biology Symposium, 2013
Join the RCSB Protein Data Bank at University of California San Diego

Open Positions:
Postdoctoral Fellows

The Challenge:
Develop innovative analysis, integration, query, and visualization tools for 3D biomolecular structures to help accelerate research and training in biology, medicine, and related disciplines.
RCSB PDB Team

RCSB.ORG
info@rcsb.org

Funding

RCSB PDB is funded by a grant (DBI-1338415) from the National Science Foundation, the National Cancer Institute, the National Institute of General Medical Sciences, and the US Department of Energy

Management

RCSB PDB is hosted by:

Rutgers
UC San Diego
UCSF
SDSC

RCSB PDB is a member of the Worldwide Protein Data Bank partnership (wwPDB; wwpdb.org)

Follow us  🌐   🇺🇸   🇨🇳   🎥   🌍
Executive Session