Worldwide Protein Data Bank Advisory Committee (wwPDB-AC)
Report of September 30th and October 1st, 2020 Meeting
BMRB organized online meeting

Chair: Peter Rosenthal Co-Chair: Art Edison

PDB Site Representatives (Nominated by wwPDB partner): Paul Adams (RCSB-PDB), Kirk Clark (RCSB-PDB), Angela Gronenborn (BMRB), Arthur Edison (BMRB), Susan Lea (PDBe), Arwen Pearson (PDBe), Masaki Yamamoto (PDBj), Daisuke Kohda (PDBj), Corinne Smith (EMDB) and Juha Huiskonen (EMDB)

***Ex Officio* Community Stakeholder Representatives:** Edward N. Baker (IUCr), R. Andrew Byrd (ICMRBS/ISMAR), Peter Rosenthal (Macromolecular EM)

wwPDB Members: Stephen K. Burley (RCSB-PDB), Sameer Velankar (PDBe), Jeffrey C. Hoch (BMRB), Genji Kurisu (PDBj)

wwPDB Member Designate: Ardan Patwardhan (EMDB;apologies)

wwPDB Regional Representatives: Debasisa Mohanty (India), Wenqing Xu (China)

Institutional Representative: Gerard Kleywegt (EMBL-EBI)

wwPDB-AC E-mail Addresses:
PDAdams@lbl.gov, kirk.clark@novartis.com, amg100@pitt.edu, aedison@uga.edu, susan.lea@path.ox.ac.uk, arwen.pearson@cfel.de, yamamoto@riken.jp, kohda@bioreg.kyushu-u.ac.jp, corinne.smith@warwick.ac.uk, juha.huiskonen@helsinki.fi

***Ex Officio* Community Stakeholder E-mail Addresses:**
ted.baker@auckland.ac.nz, byrdra@mail.nih.gov, peter.rosenthal@crick.ac.uk

wwPDB Member E-mail Addresses:
sburley@proteomics.rutgers.edu, sameer@ebi.ac.uk, hoch@uchc.edu, gkurisu@proein.osaka-u.ac, ardan@ebi.ac.uk

wwPDB-AC Chair E-mail Address:
peter.rosenthal@crick.ac.uk

wwPDB Regional Representatives E-mail Addresses:
deb@nii.ac.in, xu.wq@zjlab.org.cn

Institutional Representative E-mail Addresses:
gerard@ebi.ac.uk

wwPDB AC Meeting, September 30-October 1, 2020:

The Worldwide Protein Data Bank Advisory Committee (wwPDB-AC) and the leadership of the Research Collaboratory for Structural Bioinformatics (RCSB-PDB), the BioMagResBank (BMRB), the Protein Data Bank in Europe (PDBe), and the Protein Data Bank Japan (PDBj) met for an online meeting hosted by BMRB on September 30th and October 1st 2020.

Agenda

September 30
Greeting, Introductions of AC members [Hoch]
State of wwPDB [Hoch]
State of NMR [Hoch]
State of MX [Burley]
State of EM [Velankar for Patwardhan]
Additional Discussion
AC Executive session

October 1
wwPDB Outreach and Expansion [Kurisu]
Questions for the AC and Discussion [Burley, AC]
AC Executive Session
Wrap-up

Introduction and Overview of the wwPDB (Presenter: Hoch)

Members of the AC were welcomed by the wwPDB PIs.

Summary: The wwPDB has had an extremely productive year with important progress on all fronts. This is a considerable achievement against the background of the Covid-19 pandemic and has led to new modes of interaction within teams and between partners. Furthermore, the importance of the wwPDB partnership in disseminating data essential to understanding Covid-19 was highlighted. A new Memorandum of Understanding (MOU) will be signed in 2021. BMRB has relocated to the University of Connecticut and remains under the direction of Jeff Hoch. Core archive funding is in an overall stable condition with some exceptions. There is strong collaboration across the partners and agreement on how to advance the mission of the wwPDB. Development of the Associate Members located in China and India continues.

The wwPDB Architecture (**Figure 1**) and Organization (**Figure 2**) was described, consisting of core archives, core members, associate members, and federated resources. EMDB is still in process of becoming a core member/core archive.

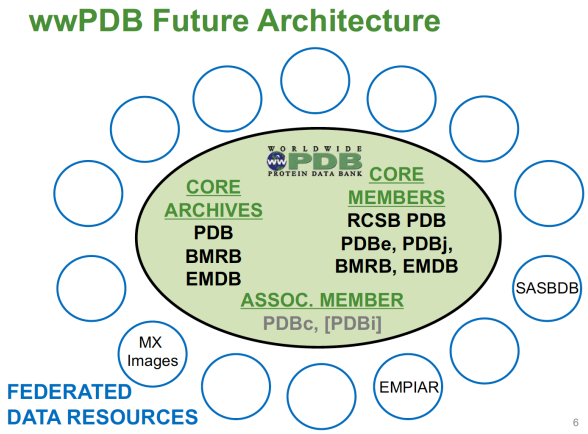

Figure 1

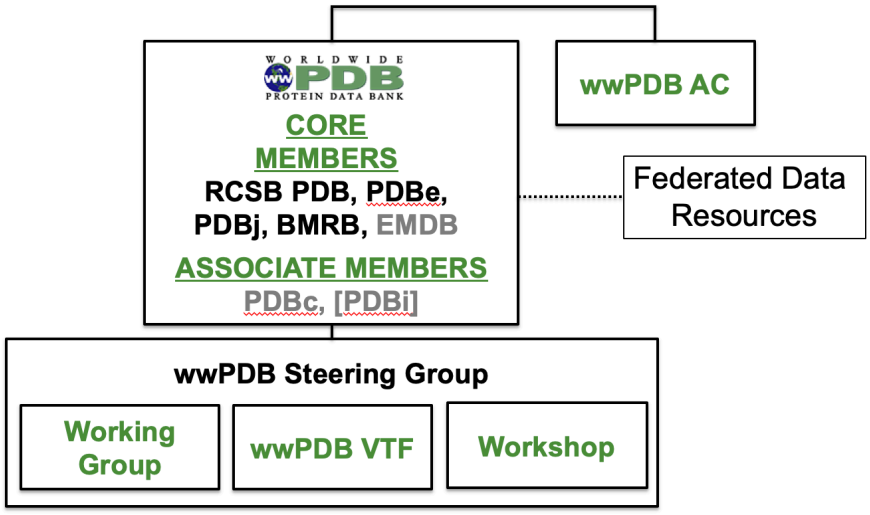

Figure 2

wwPDB Vision Statement

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 2019 AC Meeting

wwPDB

- Continued enhancement of OneDep system for deposition/validation/biocuration of MX, NMR, and 3DEM
- Continued growth in 3DEM structure depositions and engagement with the 3DEM community
- Continued depositions to PDB-Dev for I/HM structures
- Presented at the Biophysical Society I/HM workshop (March 2019). Manuscript submitted
- Workshop on improving deposition and validation of single-particle EM data (January 2020)
- Finalizing the new MOU including EMDB
- PDBe/RCSB PDB Mol* collaboration continues

PDB Core Archive

- OneDep upgraded to support remote operation
- There has been increased activity likely resulting from the Covid-19 pandemic: Projecting 15,224 depositions for calendar 2020 (13,377 depositions in 2019) and increased communication from depositors.
- Near 100% compliance on voluntary immediate release of Covid-19 entries on deposition. 371 Covid-19 related entries as of Sept 8; 838 coronavirus-related entries
- Downloads are about 2 million a day over all sites.

The AC recognizes considerable progress and efforts of the wwPDB partners and their teams in achievements during this year.

The AC anticipates signing of the MOU in 2021 and requests that the MOU be circulated to the AC.

BMRB Core Archive

- NMR-STAR dictionary enlarged with tags for unassigned coupling constants, updated enumerations for experiments including SSNMR
- Testing of pipeline to calculate structures using X-PLOR NIH with NMR-STAR as input file complete
- A source code and NMR-STAR dictionary migrated to GitHub
- New data visualizations added to entry summary pages
- BMRBdep now in production mode (449 depositions)
- OneDep now employing PyNMR-Star to parse depositions(link with BMRBdep)
- ADIT-NMR decommissioned
- BMRbig conceived and beta deployed
- Graphic design for website redesign completed
- New API endpoints developed to support UNIPROT links
- Restraint validation package integrated into OneDep and testing underway
- Refactoring and containerization of multiple services improves efficiency and robustness
- NIH R01 grant migrated to UConn

- NIH U24 proposal submitted
- Visits to BMRB Eminent Community Champions: Julie Forman-Kay, Lewis Kay, Mitsu Ikura, Cheryl Arrowsmith, Jane Dyson, Peter Wright

The AC applauds these developments at BMRB and supports directions taken by Hoch and acknowledges excellent support of wwPDB PI's to BMRB transition. The AC has requested that BMRB provide a policy and procedure document for the new MOU and this has been provided.

BMRB transition

- Encountered hurdles that resulted in funding hiatus from 4/1/20 to 7/21/20
- NIH grant transferred in entirety from UW to UConn 7/21 with Hoch as sole PI
- UConn Vice President for Research (Radenka Maric) commits resources to aid transition : ($80K in hardware capital costs to recapitulate UW infrastructure for BMRB operations, 20% effort of IT/Bioinformatics Analyst, 20% effort of IT Project Manager to assist with WBS for U24 grant proposal. Challenge commitment: 50% of PM if BMRB can raise 50% through grant(s)
- 12 40-core Dell servers installed
- Services and personnel move update.
- Services migrated as of 8/31/20: ETS completely moved, Database, website, API move imminent, Deposition system ready to switch upon annotation workflow move.

The AC understands that it has been an eventful year for BMRB but supports the conceptual and physical move as an important step for the future and is pleased to hear of support from University of Connecticut.

EMDB Core Archive. (Velanker for Patwardhan)

EMDB policy created and publicly released

- https://www.ebi.ac.uk/pdbe/emdb/policies.html

EMDB release policies are consistent with PDB policies

EMDB XML header no longer released upon entry approval

The AC is pleased to see that the policy and procedures document is now released.

Development of EMDB validation reports
Validation report improvements for EM entries now including map-only, map vs. model, and map with multiple models
Soon the full validation reports will be included for everything, not just the more recent entries.

EM Data Management Workshop January 2020:

- Recommendations about improved data capture by wwPDB/EMDB, validation reports. To be published as white paper.

- Recommend to make PDBx/mmCif mandatory for model deposition from 1st July 2021
- EMDB to implement and make available methods where recommendations cannot yet be made to enable archive-wide analysis and expert assessment of performance in individual cases
- Higher level grouping of related entries, requires discussion.
- Single particle half-map deposition is recommended to be mandatory.

The AC is pleased to see continued community consultation on EM validation and deposition and looks forward to the white paper release.

Update on Macromolecular Crystallography (Stephen K. Burley)

MX depositions continue notably at high resolution, with many more higher MW and multi-chains structures deposited, and including new small molecules, i.e. expanded CCD.

Biocurator productivity increase is reflected in MX.
Structure versioning has led to 100 coordinate replacements for OneDep and legacy pre-OneDep structures.
PDBx/mmCIF deposition mandatory.
More documentation on mmCIF needed for depositors.

AC is pleased to see that MX depositions remain strong. We recognize the work of the partners and the working group (Chair: Paul Adams) on mmCIF adoption. The AC considers versioning to be an important subject for the future and encourages continued discussion to incentivize versioning. The AC notes that versioning is a time-consuming activity.

OneDep 2019/2020 Progress *v*ersus Goals

- Provided wwPDB DOI resolution
- Enabled author-initiated coordinate replacement (Legacy entries, phase II)
- Enabled single NMR data file deposition in NEF or NMR-STAR format
- Completed carbohydrate remediation
- Improved biocuration processes on entity transformation for BIRD molecules
- Streamlined weekly update- enabled per-site generation of validation reports
- Updated archive validation reports with enhancements for ligands and 3DEM maps and provided ED map coefficient files

Re-forecasted
- Implement NMR restraint validation
- Depositor-annotated assembly classification
- Post-Translational Modification project planning

Mitigation
- Actively engaged NMR community in 2020
- Set clear requirements and phased plan for depositor-annotated assembly
- Follow carbohydrate remediation project template

wwPDB Biocurator Productivity

- Continuing increased efficiency since 2009
- Significant increase with OneDep system
- Remote biocuration since Mar. 2020
- Pandemic doesn't impact biocuration productivity

The wwPDB-AC recognizes continued development of OneDep and the remarkable flexibility and response to difficult circumstance arising from the Covid-19 pandemic leading to outstanding productivity. We thank Jasmine Young and the OneDep team for regular, positive updates. The AC supports efforts during the past year to engage with the NMR and EM communities and expects these to continue.

wwPDB DOI Resolution

- Provided for all on-hold and released PDB entries
- Accessed frequently > 335K times
- First coronavirus entry 6lu7 has top visit
- Some journals have adopted DOI links: Acta Cryst. D & F, FEBS J., JBC (Communication ongoing)

The AC is pleased to see this implemented. We still encourage work on the web landing page for clarity and a design that draws users to the resource.

wwPDB Core Member Funding Status

- RCSB PDB and PDBe received a joint NSF/BBSRC grant (3 years duration) to support development of the Next Generation PDB Archive (presented at 2019 wwPDB AC meeting)
- RCSB PDB: NSF/NIH/DOE funding renewed: 2019-2023
- BMRB: NIH NIGMS funding: 2019-2023: Inadequate budget: need to find additional support: NIH R01 transferred to UConn, NIH U24 submitted
- PDBe: EMBL-EBI, Wellcome Trust: 2021-2025
- PDBj: NBDC-JST and AMED funding: 2019-2022, Possible additional budget from S. Korea
- EMDB: EMBL-EBI, Wellcome Trust: 2019-2024
- wwPDB Collaboration Resources have light reduction through 2021.

The AC is positive on the overall funding of the wwPDB partners. We understand that BMRB needs additional funding and this also affects collaboration resources. The AC will appreciate periodic update on progress from the partners and also from associate members.

Scheduling of Future wwPDB AC Meetings

EMDB to Host 2021 wwPDB AC

Next wwPDB AC meeting: Tuesday, Oct. 19nd 2021

Host: EMDB Venue: Online

PDB50 Celebration (Europe) to follow immediately thereafter (Oct. 20th-22nd 2021) at EMBL-Heidelberg.

2022 wwPDB AC meeting: Friday, Oct. 14th or Friday Oct. 21nd

Host: RCSB PDB Venue: Rutgers University, Piscataway, NJ, USA

The AC looks forward to marking the PDB50 milestone.

BMRB Core Archive Plans (Presenters: Hoch)

Policy statements on OneDep/BMRBdep, NMR- STAR/NEF:

- As an essential partner in the OneDep Team, BMRB commits to ensuring that BMRBdep is fully integrated in OneDep
- While NMR-STAR remains the archive format for biomolecular NMR data hosted by BMRB, BMRB is fully committed to supporting NEF as an exchange and deposition format
- Explore expansion of small molecule data sharing with PDB (aligning with Chemical Component Dictionary)
- Complete overhaul of web site, new logo
- Work on documenting, strengthening, and streamlining internal systems and SOPs
- Finalize work on curated/normalized NMR-STAR schema
- Continue expansion of curated NMR data types
- Continue expansion of curated collections pertinent to specialized areas: SSNMR, Disordered systems, Metabolomics
- Continue rollout and testing of BMRbig: Develop tools to facilitate populating BMRB and PDB depositions from BMRbig uploads.
- Additional funding will be sought.

The AC is pleased to hear BMRB's full commitment to OneDep in its plans. BMRB and OneDep are working well together and the direction looks positive. Congratulations are in order for all that has been achieved during the last year, and hopes the team will continue to thrive. A white paper describing these future directions may now be timely. The AC understands the value in supporting both NMRstar and NEF, and recognizes BMRB's effort to endeavor to this end. AC advisory board is positive about BMRB transition but concerned about funding and notes loss of one position.

EMDB Core Archive Plans (Presenter: Gerard Kleywegt for Ardan Patwardhan)

Publishing recommendations, incl. as preprint (white paper in progress)

Implement easy-to-do recommendations (on-going); then:

- Update OneDep and validation server
- Calculate and release validation reports for all current EM entries in EMDB and PDB

Later:

- Add many new methods to EMDB Validation Analysis (VA) web pages to enable evaluation of robustness, reliability, information, usefulness, etc. (on-going)

- Implement additional recommendations in validation pipeline/reports
- Wait for the field to do additional work and review in a few years' time
- Get recommendations for EM modalities other than single particle analysis

The AC supports the progress on EM validation and deposition and encourages publishing of the white paper from the 2020 EM validation workshop during 2021. This should be provided on the wwPDB websites following publication.

As suggested previously, "the wwPDB should consult with software developers and detector manufacturers to identify important fields to capture in supporting mmCIF format files for annotation and validation of electron microscopy data. For microED, policies need to be set on what critical data fields are required to accept electron diffraction data, and where and how to deposit PDB and mmCIF data. The wwPDB should continue to develop ways of capturing and linking structural data for biological assemblies in different states or by different methods across the archives"(AC report 2019).

wwPDB Outreach and Expansion (Presenter: Genji Kurisu)

wwPDB Outreach
Outreach activities have included ECM2019, the 2020 OneDep Developer Summit, AsCa2019 Singapore, and Covid-19-related outreach. Several wwPDB joint publications from the last year are available to the AC on request.

The AC notes that Outreach activities have still managed to flourish despite the conditions of the pandemic. Furthermore, the wwPDB has responded in a positive way through Covid-19 outreach activities in addition to the basic science role it has played during the pandemic.

wwPDB Foundation
http://foundation.wwpdb.org
The wwPDB Foundation is a private entity that works with the wwPDB and acts to support outreach activities of the wwPDB that cannot be supported by the individual partner funding sources. Chair is Celia A. Schiffer, U. of Massachusetts Medical School. Fundraising continues and PDB50 celebrations planned

wwPDB Associate Members

The associate members (See Figure 1. wwPDB Future Architecture) are as follows:

PDB China

National Facility for Protein Science in Shanghai (NFPS) and iHuman Institute and SIAIS, Shanghai Tech University, Pudong, Shanghai, China

- Director, Wenqing Xu

PDB India

Molecular Biophysics Unit, Indian Institute of Science,

Bangalore, India

- PI, Manju Bansal
- Co-Investigator, Debasisa Mohanty and K. Sekar

Implementation plans for the associate members were described.

Implementation plan for new "data out" activities at both PDBc and PDBi

- Background Training (conducted remotely)
- Hardware setup (local with remote support)
- wwPDB authorized ftp service
- Set-up of new pdbc.org or pdbi.org websites

Implementation plan for "data in" activity at PDBc (remote/onsite)

- Remote Training of PDBc biocurators by RCSB/PDBe/PDBj
- Scientific Training and OneDep system education and system training
- Onsite Training at PDBj by PDBj

Note: Invitation of PDBj members to visit China and curators to visit Japan were postponed and Onsite OneDep system training (postponed)

Implementation plan for OneDep at PDBc

PDBc's OneDep system setup at Osaka, PDBc biocurators to the workflow manager, PDBj will assign the depositions with PROC status in PDBc's

OneDep@Osaka

PDBj biocurators will process depositions from China that are beyond PDBc capacity. PDBc will start with x-ray entries only.

The AC continues to be impressed with developments at PDBc and PDBi and recognizes contributions of wwPDB partners in training. We understand the unusual circumstances and support continued training remotely through online methods and hope for in-person visits when they become possible. There is now more experience in online training and this may play an increased role in the future. The AC is pleased to hear that consistent training and quality control is a goal through common software and practices and co-operation between sites. The AC recognizes the importance generally of internet communication between partners. The AC suggests that timelines be prepared for the next AC meeting that include plans for funding and other criteria for associate membership and how regional capacity will meet demand. We look forward to hearing about extended funding for PDBi which was disrupted due to covid-19.

The following Items were raised for discussion

1. PDBc will invite all wwPDB PIs to Shanghai to check their status. After approval by the wwPDB PIs, an official announcement that the data processing at Shanghai is starting gradually will be announced to Chinese depositors.

2. PDBj-BMRB will keep covering all BMRB deposition mainly from Asia.

The AC supports these proposed actions.

Specific Questions/Actions for the wwPDB AC

1. Does the wwPDB AC concur with proposed plans for making Principal Investigator Contact Information available for new structures deposited to the PDB Core Archive (Appendix 1)?

   - *The plans are unanimously supported by the AC.*

2. Does the wwPDB AC concur with proposed plans regarding deprecation of Legacy PDB File Format (Appendix 2)?

   - *The plans to consult the user community regarding current reliance on legacy PDB File Format and how to support transition to PDBx/mmCIF are unanimously endorsed by the AC. The AC anticipates updates regarding the progress of consultation, dates of deprecation, and notification of problems as they arise.*

3. Does the wwPDB AC concur with proposed plans for addition of CAVEAT records in the model coordinate files when there are serious problems highlighted in validation report (Appendix 3)?

   - *The AC supports the proposed plans to communicate this new policy and implement these CAVEAT additions according to objective criteria. Metrics for both MX and EM require evaluation. The AC will anticipate a report on the fully developed policy following the consultation.*

4. Does the wwPDB AC concur with proposed plans for encouraging deposition of raw data to curated Raw Data Archives for new structures deposited to the PDB Core Archive (Appendix 4)?

   - *The AC supports the proposed plans to encourage raw data depositions including forming collaborations with curated archives and seeking funding to achieve it. The AC supports EMPIAR being federated with wwPDB. Raw data ideally will include ID linkages to archive data and may use the same API with OneDep for upload provided that the task should not become onerous on the user at the time of deposition.*

5. Does the wwPDB AC concur with proposed plans for training and expansion of the wwPDB Franchise (Appendix 5)?

    - *The AC concurs with proposed plans.*

6. Does the wwPDB AC have any questions or concerns regarding the OneDep Software Development Team Quarterly Reports for 2019-2020 (Appendix O)?

    - *The AC thinks the reports are fantastic and shows how well the team is working. We have accepted the reports from all of the partner ACs and thanked the wwPDB for providing this information. The AC requests that individual PDBc and PDBi AC reports be submitted to the wwPDB and made available to the wwPDB AC.*

7. Does the wwPDB AC have any questions or concerns regarding the individual RCSB PDB, PDBe/EMDB, PDBj, or BMRB Advisory Committee reports (Appendix R)?

    - *The AC accepted the reports from all of the partner ACs and thanked the wwPDB for providing this information. The AC requests that individual PDBc and PDBi AC reports be submitted to the wwPDB and made available to the wwPDB AC.*

8. Does the wwPDB AC have any additional questionsor concerns?

    - *The AC suggests that community discussion is needed about the mandatory release of coordinates reported in BioRxiv manuscript depositions which are considered publications. Depositors should be clear on a uniform and transparent policy and author agreement on coordinate and data release policy set by BioRxiv.*
    - *The AC believes there needs to be community consultation on whether circumstances require priority release of coordinates such as recently supported for Covid-19-related structures.*
    - *The AC understands that the relationship between BMRB and OneDep is moving in the right direction but will discuss the progress including discussion with the greater NMR community. We understand that depositions with coordinates will go through OneDep as an entry point, but in the absence of coordinates the deposition may go through BMRBdep.*

Conclusion

The wwPDB AC meeting concluded with a general discussion with the wwPDB PIs following the executive session.

The next wwPDB AC meeting will be held online on Tuesday 19 October 19, 2021.