Worldwide Protein Data Bank Advisory Committee (wwPDB-AC)
Report of October 10th 2014 Meeting
PDBe, EMBL-EBI, Hinxton, UK

Chair: Soichi Wakatsuki

PDB Site Representatives: Wayne Hendrickson (RCSB-PDB), Paul Adams (RCSB-PDB), Angela Gronenborn (BMRB), Randy Read (PDBe), Helen Saibil (PDBe), Gaetano Montelione (BMRB), Genji Kurisu (PDBj), and Soichi Wakatsuki (PDBj)

***Ex Officio* Community Stakeholder Representatives:** Edward N. Baker (IUCr), R. Andrew Byrd (ICMRBS), and Wah Chiu (Macromolecular EM)

wwPDB Members: Stephen K. Burley (RCSB-PDB), Helen M. Berman (RCSB-PDB), Gerard J. Kleywegt (PDBe), John Markley (BMRB), and Haruki Nakamura (PDBj)

wwPDB Associate Members: Manju Bansal (India), Jianping Ding (China)

wwPDB-AC E-mail Addresses:
wayne@xtl.cumc.columbia.edu, PDAdams@lbl.gov, amg100+@pitt.edu, rjr27@cam.ac.uk, h.saibil@mail.cryst.bbk.ac.uk, guy@cabm.rutgers.edu, gkurisu@protein.osaka-u.ac.jp, soichi.wakatsuki@stanford.edu

***Ex Officio* Community Stakeholder E-mail Addresses:**
ted.baker@auckland.ac.nz, byrdra@mail.nih.gov, wah@bcm.edu

wwPDB Member E-mail Addresses:
sburley@proteomics.rutgers.edu, berman@rcsb.rutgers.edu, gerard@ebi.ac.uk, harukin@protein.osaka-u.ac.jp, markley@nmrfam.wisc.edu

wwPDB Associate Member E-mail Addresses:
mb@mbu.iisc.ernet.in, jpding@sibs.ac.cn

wwPDB-AC Mission Statement
To help ensure that the Protein Data Bank is maintained for the public good as a secure, singular global archive for experimental structural biology data that is freely accessible in perpetuity.

Meeting Summary
The Worldwide Protein Data Bank Advisory Committee (wwPDB-AC) to 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 at PDBe, EBI, Hinxton, UK, on October 10th 2014.

The agenda included

(1) Overview;
(2) Common Tool for D&A;
(3) NMR;
(4) Outreach;
(5) Looking ahead, questions, discussion topics; and
(6) Executive session and feedback

Following a welcome talk by Janet Thornton, Helen Berman explained the RCSB PDB leadership transition from her to Stephen K. Burley. Overview of the state of the wwPDB was presented by G. Kleywegt, followed by summaries of recent activities from S.K. Burley, J. Markley, and H. Nakamura.

(1) Overview (presented by Kleywegt)

G. Kleywegt provided an overview of the past year.

The year 2014 has marked a special PDB milestone: 100,000 entries on May 14. With continuing and collaborative efforts in developing and maintaining the state-of-the-art archive of macromolecular structures, the wwPDB continues to serve broad science communities including structural biology, biochemistry, bioinformatics, cell and molecular biology, translational medicine, bioenergy and environmental biology, as well as various areas of industries. Data depositions continue to grow and the daily downloads exceed one million. At the same time, the wwPDB is preparing new directions toward integrative structural biology through coordinated approaches such as hybrid task force and workshops.

wwPDB organization

The wwPDB-AC was impressed with the collaboration between the four centers and with the excellent synergy and collegiality among all members of all centers as evidenced in development and implementation of the common deposition and annotation system (D&A), PDBx format, and various (validation) task forces involving key experts in the relevant areas, X-ray crystallography, NMR, 3DEM, small-angle scattering, and hybrid methods. It is particularly noteworthy that the global workload balance is ensured by the D&A implementation.

The RCSB PDB Leadership transition from Helen Berman to Stephen Burley was well planned and seamlessly executed, and the new constellation is highly effective, and well integrated with the rest of wwPDB organization.

wwPDB funding

In addition, funding for the next 1-2 years is in place for all centers and a healthy level of support is evident. This is an outstanding achievement in the present funding climate. For all four centers the following timelines of funding stability were reported:

- RCSB: non-competitive renewal of NSF funding (2014-18); project grants from the NIH
- PDBe: core support from EMBL-EBI and the Wellcome Trust (new grant 2015-2019); project grants from MRC, BBSRC, EU, NIH and CCP4
- PDBj: new JST-NBDC funding started for 3 years from April 2014 to March 2017
- BMRB: NIGMS R01 funding started for 5 years, 2014- 2019

The four wwPDB-AC centers are commended for their efforts in securing these funds for the next several years. Cognizant of global funding climate in science the four centers are encouraged to collaborate on securing sustainable funding as well as explore new future possibilities.

wwPDB Foundation

The wwPDB Foundation seems to function efficiently and provides funds for otherwise unmet needs.

Hybrid Methods

The hybrid methods task force, which met at Hinxton in early October, began discussions for a global strategy for dealing with structural models derived from diverse experimental inputs, no one of which by itself may be definitive. Many structural representations need to be anticipated and many types of supporting data need to be accommodated in some manner. The plan to develop a federation of data systems seems sensible, and it is admirable that wwPDB is taking a leading role in this important new direction, which likely will take years to reach maturity. We look forward to seeing the developing white paper on hybrid methods take shape for publication.

(2) Common tools for Deposition and Annotation (D&A) (presented by Burley)

The wwPDB should be congratulated on their excellent progress in developing the Common Deposition & Annotation tool and making it available to depositors of crystallographic structures. However, current usage of the system means that many structures are still being deposited through the legacy systems at other sites. The Advisory Committee strongly recommends that crystallographic users be transitioned to the new common system now. This will enable load balancing across the wwPDB sites and provide researchers access to a significantly improved deposition tool. In concert, we recommend that the wwPDB engage users worldwide, but particularly in Asia and Oceania, to encourage adoption of the new Common Deposition tool. It was also noted it was not universally clear how users could obtain the current structure validation reports required for manuscript submission and review. The wwPDB is strongly encouraged to provide clear links to up-to-date validation reports where possible. Some committee members expressed a concern that there were still issues with the built-in user communication system which is how annotators now communicate with the depositing researchers. It seems to add more work than just exchanging emails. The committee was heartened to hear that the wwPDB Centers are working together to define a streamlined mechanism for the creation and acceptance of new mmCIF/PDBx data items as they are needed. Finally, it was noted that there are still cases where the internal chemical component dictionaries are not correct and this can make deposition of a structure problematic. Corrections should be made expeditiously to avoid holding up the process.

X-ray validation

The wwPDB has done an excellent job of implementing and rolling out the X-ray validation reports, and the wwPDB-AC was pleased to see that they are now highlighted on the entry pages at all individual partner web sites. It was also good to see that user feedback is being solicited, as the validation report requires some training to interpret and there may be ways of presenting some of the information more clearly. Currently, while the old and new deposition and annotation systems coexist, it is still possible for depositors to receive old-style validation reports, but it would be better if minor changes to the legacy deposition systems were made so that depositors were encouraged to download new-style reports, which are much more useful for referees and editors.

It is already three years since the X-ray validation task force report was published, and the advisory committee is pleased to hear that the wwPDB is looking forward to updating the validation criteria on a 5-year cycle with the plan to reconvene the task force in 2015.

EM validation

FSC and tilt pair validation are on the EMDB web site, but there are major scientific problems still to solve, to develop general validation tools for EM maps and fitted atomic models.

The case of high resolution maps and fitting is most straightforward, since they contain backbone and some side chain densities. Nevertheless, there is a need for further work on validation of models derived from 3-4.5 A maps. It was suggested that polyalanine chains should be used for backbone tracing, rather than alpha carbons. Lower resolution single particle maps are extremely difficult to validate, and there are no obvious general validation methods yet available.

Current research in the field is addressing these problems, with challenges for both maps and fitted models planned in the coming year. Working groups from the 3DEM community will be organized and supported by the EMDataBank in order to provide a critical evaluation of proposed validation methods and criteria. The EM validation task force will be reconvened after the challenges have taken place.

(3) NMR (presented by Markley)

The wwPDB NMR VTF has made considerable progress on the representation of NMR restraints. A workshop was held in September, and a draft report is in progress. Leadership by Professors Michael Nilges and Guy Montelione is invigorating the process and driving considerable progress. Major NMR software developers are engaged in the process, making a strong and impactful outcome very promising. The VTF has another meeting scheduled for January 2015. The wwPDB AC encourages that the NMR VTF establish a timeline for milestones on completing and implementing the restraints standards. It is hoped that progress with this effort will smoothly accompany the implementation of the Common D&A tool for NMR-based depositions, expected to begin in 2015.

Development of BioMagResBank and NMR Structure Validation

The PIs of the BioMagResBank are commended for organizing and successfully funding a new five-year RO1 grant from the National Institutes of Health. Although the new grant provides funding at a reduced level compared to the previous BioMagResBank funding from National Library of Medicine, it will provide the resources needed to maintain the core activities of BioMagResBank at Madison. The BioMagResBank has also made progress in developing NMR components of the Common D&A Tool. Implementation of these developments is currently in progress at PDBe, and full alpha testing at BioMagResBank is scheduled to begin soon. The advisory committee is anxious to see how well the Common D&A Tool works for deposition of data associated with NMR structures.

The recommendations of the NMR VTF are also being implemented at PDBe as part of a standardized NMR structure validation report. The draft NMR validation report looks very good, and incorporates most of the recommendations proposed for Phase 1, including validations of the chemical shift data themselves, and standardized methods for annotating well-defined regions of the NMR structure model. The outstanding exception is validation of structural models against constraint data. Although programs exist for carrying out these validations, they do not work robustly for all formats of constraint data. To address this problem for the entire community, wwPDB has organized a new effort to develop a common exchange format for NMR constraints and related data. A workshop to brainstorm the new NMR Exchange Format (NEF), organized primarily by G. Vuister, A. Gutmanas and leaders of the wwPDB, was held in November 2014 in Hinxton, and a follow-up workshop to discuss prototypes and specific implementations has been scheduled for January 2015 at Rutgers. The wwPDB-AC has requested a timeline for implementation of the new exchange format. We look forward to seeing the completion of Phase 1

of the work of the wwPDB NMR Validation Task Force, and release of the first wwPDB validation reports for NMR structures, during 2015. The NMR VTF will also be convening in January 2015 to discuss Phase 2 validation recommendations.

(4) Outreach (presented by Nakamura)

Events on 100,000+ entries and the International Year of Crystallography (IYCr), including regional activities and the 2014 wwPDB calendar, have been well managed. Publications are also good in the last 12 months.

Despite the many efforts by the four wwPDB centers on rebranding as wwPDB, the wwPDB is still relatively less known to the general public. Therefore, redesigning the wwPDB.org web is an important task. The β-version of the wwPDB website looks excellent and the wwPDB-AC strongly encourages launching it as soon as possible.

The wwPDB-AC encourages exploring different ways to present the accomplishments of the wwPDB, for example, number of atoms in the database, new tools becoming available on the wwPDB website, etc.).

Further announcements of new D&A system should be made to increase its use especially in the Asia-Pacific region. In addition, the wwPDB is recommended that a similar effort be initiated and carried through the introduction and implementation of the common D&A tool for NMR-based depositions. The NMR VTF can be consulted for appropriate forums to conduct this type of outreach in the community.

(5) Looking ahead, questions, discussion topics (presented by Kleywegt)

wwPDB-AC is pleased to learn of the events and activities are well planned for the next 3 years.

Question to wwPDB-AC #1: Release of all sequences 5 days prior to release of the corresponding structures as requested by the comparative modeling community?

There was considerable discussion on the pros and cons as well as possibility of violation of journal embargos. It was clarified that the release of sequences just 5 days prior to release/publication of the structure will not necessarily violate journal norms. The wwPDB-AC endorses the proposal of releasing sequence information followed by structures 5 days later. Additionally the depositor will be provided an option of suppression of all information that could be used to infer the identity of a structural target prior to release of sequence information 5 days before the final, full release of the structure.

Question to wwPDB-AC #2: Reduce computed items in the core archive?

There was general consensus that in view of the fast expanding size of the PDB archive and the corresponding secondary information on assemblies, sites, secondary structure, etc., it would be pragmatic to limit this. However, wwPDB-AC recommends that each of the wwPDB centers provide up-to-date tools for computing items such as assemblies, sites, secondary structures etc. to the viewer of their website.

Question to wwPDB-AC #3: Versioning of PDB entries?

The wwPDB-AC endorses introduction of a versioning system. However, considering the complexity of this issue, the committee encourages wwPDB to establish a detailed operational and manageable mechanism and will welcome further discussion on this issue in the next meeting.

Question to wwPDB-AC #4: Adopt unambiguous author identifiers?

The wwPDB-AC endorses the general idea of adopting an unambiguous author identifier system and agrees that ORCID could be a likely choice. The committee encourages the wwPDB centers to continue exploring if there are any other stable, internationally recognized mechanisms or systems and will welcome further discussion on this issue in the next meeting.

EM general points

The wwPDB-AC recommends EMDB be represented equivalently to BMRB on the wwPDB web site and at wwPDB meetings. It is also urgent to develop and implement the new D&A system for EM maps and fits.

The activities undertaken by the RCSB PDB and PDBe to interact with the cryoEM communities are very impressive. They have already received a NIGMS grant to support the 3DEM activities in the last five years. Their current grant support will expire in May of 2017. Their effort in leading the cryoEM community to understand and reach a consensus on map and model validation criteria is timely and important. They are making headway to incorporate the EM deposition into the new D&A system. The panel feels strongly that this incorporation of one stop shopping for map and model deposition into D&A should be completed by early 2015. The panel recommends that wwPDB should add cryoEM as one of its organizational components as done for NMR. A cryoEM representative to report the 3DEM activities in the Advisory Committee meeting is necessary to keep the panel informed so that we can advise on such an important area of development in structural biology today.

International Union of Crystallography (IUCr) relations

The relationship between the wwPDB and the IUCr is of fundamental importance to both organizations, given the large proportion of depositions (~90%) that come from X-ray crystallography, and the long-standing importance of the IUCr in helping with policies such as those related to deposition, validation, release, and journal requirements. We are confident that this relationship is strong and in good shape. Reports are provided annually to the IUCr, *via* its representative on the wwPDB-AC, and the IUCr for its part continues to express its willingness to help in any way it can.

The wwPDB-AC appreciates the efforts made by the wwPDB site heads to make the presentation and background materials available one week before the meeting which had been very helpful for the committee members to prepare for the meeting and discussions.

The next wwPDB-AC meeting will be held in Osaka, Japan on Friday October 2, 2015 to be followed by a Symposium for "Integrative Structural Biology with Hybrid Methods (title to be confirmed)" which will be held in Osaka on Saturday October 3, 2015.

Appendix: PDB Metrics

In aggregate, 10566 (10,500*) depositions were processed between January 1st and December 31st 2013 with a two-week average turnaround (*2014 projection).

Breakdown of depositions by discipline in calendar year 2013 was as follows:
X-ray: 9697 (92%, up from 9269 in 2012)
NMR: 590 (6%, up from 585 in 2012)
EM: 234 (2%, up from 100 in 2012)
Other: 45 (0.4%, up from 12 in 2012)

Breakdown of depositions by wwPDB processing site in calendar year 2013 was as follows:
RCSB PDB: 6652 (63%)
PDBj: 2128 (20%)
PDBe: 1786 (17%)

NMR structures
BMRB: 466 (92%)
PDBe: 12 (2%)
PDBj-BMRB: 28 (6%)

Breakdown of depositors by location in calendar year 2013 was as follows:
North America 40%
Europe 30%
Asia 18%
Industry 7%
South America 0.6%
Australasia 4%
Africa 0.1%