

RCSB Protein Data Bank Advisory Committee

Report of April 21, 2020 Annual Meeting

Teleconference

Chair: Paul Adams

Membership:

Present: Peter Andolfatto, Judith Blake, Andy Byrd, Bridget Carragher, Wah Chiu, Kirk Clark, Paul Craig, Roland Dunbrack, Paul Falkowski, Thomas Ferrin, Mandë Holford, Cathy Peishoff, Sue Rhee, Torsten Schwede

Absent: Robert B. Darnell, Jill Trehwella

RCSB PDB AC E-mail Addresses:

PAdams@LBL.gov, pa2543@columbia.edu, judith.blake@jax.org, bcarr@nysbc.org, byrda@mail.nih.gov, wahc@stanford.edu, kirk.clark@novartis.com, paul.craig@rit.edu, Robert.Darnell@Rockefeller.edu, roland.dunbrack@fcc.edu, falko@marine.rutgers.edu, tef@cgl.ucsf.edu, mholford@hunter.cuny.edu, peishoffc@gmail.com, srhee@carnegiescience.edu, torsten.schwede@unibas.ch, jill.trehwella@sydney.edu.au

RCSB PDB Leadership: Stephen Burley (Director), Helen Berman (Director Emerita), Andrej Sali (UCSF Site Head)

RCSB PDB Leadership E-mail Addresses:

sburley@proteomics.rutgers.edu, berman@rcsb.rutgers.edu, sali@salilab.org

Executive Summary

The Advisory Committee (AC) to the Research Collaboratory for Structural Bioinformatics (RCSB) held a virtual meeting on April 21st, 2020 to review recent progress and provide feedback on specific questions.

Agenda items included

- Welcome and Introductions
- Response to COVID-19
- 2019 Overview
- Deposition/Biocuration
- New and Improved RCSB.org
- Outreach/Education
- Operations and Funding
- PDB50 and AC Meetings
- Discussion: PDB Format

The meeting was opened by Dr. Stephen Burley. Other RCSB PDB participants were Helen M. Berman, Robert Lowe, John Westbrook, Jasmine Young, Christine Zardecki (Rutgers); Andrej Sali (UCSF); and Jose Duarte (UCSD). Appendix 1 provides a summary of the RCSB responses to the 2019 Advisory Panel meeting recommendations. Appendix 2 provides a summary of global PDB deposition and data access statistics in 2019.

Overall Comments from the Advisory Panel

The team is congratulated on an excellent set of focused presentations, and their collective leadership of the RCSB and their individual projects over the last year. The committee applauds the continued efforts of the RCSB Director, Stephen Burley, and his colleagues to seek additional funding for the RCSB activities. The committee also recognizes that everyone is having to respond to an unprecedented global situation in the form of the COVID-19 pandemic.

Recommendations for future meetings

- Keep up the great work!
- A brief report in the next 2 months on the impact of the COVID-19 pandemic on the RCSB operations and PDB depositions would be very helpful for the committee.

Detailed Advisory Panel Comments and Feedback

COVID-19: Are there other projects in this area we could develop to support research and education?

The current COVID-19 pandemic presents many challenges but also some opportunities for the RCSB. There is a significant interest from researchers and also the general public to learn more about the SARS-CoV-2 virus and pandemics in general. Given the important role that structural biology is playing in understanding the virus, and the role it plays in developing therapeutics, there is a great opportunity for the RCSB to provide material to the community. They have already created the COVID-19/SARS-CoV-2 Resources website (<http://rcsb.org/covid19>), with links to relevant structures, a general coronavirus education page, educational videos, and many images. The committee also was pleased to hear about plans to host a summer boot camp around the topic of coronavirus in 3D, and the evolution of SARS-CoV-2 proteins. Beyond their COVID-19 current efforts, the RCSB should consider additional activities. The committee identified a number of opportunities, including:

- Creating PDB101s on viral infection processes, immunity, virus-mediated acute respiratory syndrome.
- Extending the current site to provide structural information and links to other material (experiments, recent news, etc) about each protein from the viral genome.
- Enabling or more directly supporting the collection of revisions of structures from the community, which could eventually lead to new version uploads by the original authors.
- Reaching out to the local community to provide information about the basic structure of COVID proteins, viral RNAs, interacting cellular proteins, virus and pathogen in relationship to human diseases through TV news stations, school districts or public health departments.

The committee also recognizes that there may be opportunities to combine education and outreach activities around COVID-19 with fund raising activities, especially with the PDB50 celebrations next year.

Recommendations

- Develop an action plan for expanding the RCSB role in educating the community about COVID-19 and other related pandemics, and the role of structural biology. These plans would ideally be integrated with current and future fundraising activities.
- Track access to the RCSB maintained COVID-19 materials; this would be very helpful for future efforts to highlight the impact of the resource.

Deposition/Biocuration: Any concerns about the Deposition/Biocuration work underway with our wwPDB partners?

The committee heard from Jasmine Young about deposition and biocuration activities. We continue to be very impressed with Jasmine's leadership in this area, and the report on last year's activities was very positive. The load balancing across the wwPDB locations appears to be working well, and new features such as GroupDep are making it easier for some groups to make their depositions. The introduction of mandatory mmCIF deposition for crystallographic models is also hopefully improving the workload on the annotators. A number of other activities across the wwPDB are also likely to improve deposition and biocuration activities in the future, including improved ligand validation and biological assembly annotations, author initiated coordinate replacement, EM map validation, and the remediation of the carbohydrates.

It is also clear that the wwPDB will need to accommodate significant growth in the deposition of atomic resolution models from cryo-EM in the next 5 years. At the same time new XFEL approaches are gaining in popularity. It also currently looks unlikely that there will be dramatic reduction in the number of crystallographic structures deposited each year. This increased volume of structures will need to be processed without a backlog developing. The committee was very pleased to see that an analysis had been performed to provide projections of depositions from 2020 to 2024 for all of the experimental techniques. However, there was a concern that the projections for cryo-EM might be underestimated and not reflective of the current exponential growth.

Recommendations

- Continue to monitor the growth of cryo-EM depositions, and be prepared to prioritize the implementation of deposition standards and tools to help respond to the increased load.
- Continue to track the time taken for depositions, to both measure the load on annotators, and to provide metrics about how process improvements are increasing deposition throughput. This information will be helpful for funding justifications in the future.

New and Improved RCSB.org: Additional Site and Search Functionality requests?

The committee heard from John Westbrook on the ongoing efforts to improve the infrastructure for the rcsb.org website and associated backends. We were impressed how quickly this has been implemented without any substantial interruptions to providing services. Demonstrations of the new search functionality highlighted useful new features. However, there were some concerns about the complexity of the search system for many users, and the loss of important features (such as refining a search to provide a non-redundant set of results). One suggestion was the creation of question-driven functionalities and workflows for popular search activities. The committee recognizes that the RCSB has undertaken community outreach to get user feedback, but these efforts might need to be extended. The new Mol* 3D visualization system was also presented, and clearly shows great potential for interactive display of molecules and maps. However, the committee feels that further development, and in some cases simplification, of the interface would benefit many of the RCSB users.

Recommendations

- Seek further community input, maybe through the creation of focus groups or targeted outreach, to refine the search functionality and the Mol* visualization services.

Outreach/Education: Suggestions for new materials and virtual venues for celebrating PDB50 throughout 2021?

Christine Zardecki presented many of the great efforts in the area of outreach and education over the last year. The PDB-101 resource currently has two-thirds of a million users, and over 2 million page views per year. Significant effort has been spent on developing new content, in particular different aspects of

human health. The committee also was very encouraged to see the results of a community survey, and that 3D print files have been made available for several molecule of the month topics. The community outreach by the team is excellent and clearly very important for educating the research community about the RCSB and structural biology. Of particular importance are the efforts to educate other educators and students. Clearly, the current pandemic will have an impact on the team's ability to perform outreach in person and will require them to develop new approaches to communications. The committee had several suggestions for ways to engage the community in the current circumstances and leverage the importance of structure in the COVID-19 response. One idea was the creation of virtual reality resources, perhaps using ChimeraX. This might provide a platform to propose something similar to Folding@Home where people would be in VR or AR space and attempting to design drugs for COVID-19 proteins. Another suggestion was a competition for creating protein structures from found objects around the home.

The PDB50 celebration in 2021 provides a great opportunity to promote the RCSB widely and emphasize the impact of structural biology. The committee suggested addressing this in multiple forums. Large conferences provide an opportunity for outreach, and in some cases these may be well organized as virtual conferences - the Intelligent Systems for Molecular Biology in 2021 was one example. Museums and other public facing organizations may also provide a great opportunity for engaging a broader audience. The American Museum of Natural History was put forward as an organization looking for online content. Ultimately, the committee feels that there is an opportunity for either local or national recognition through mainstream media, such as NPR and network TV channels. Science Friday at NPR would be a great target for a PDB50 piece, as would a NOVA documentary.

Recommendations

- Create a plan for online outreach and communications for the next 12 months, which incorporates some PDB50 celebration activities
- Develop a PDB50 media communications strategy targeted both locally and nationally.

Next Advisory Committee Meeting location?

While the committee looks forward to the next meeting in person, it seems unlikely that it will come to pass in the first half of 2021. We expect that the next committee meeting will be held virtually. If an in person meeting is possible, the option of coordinating with the PDB50 event at the ACA in Baltimore seems reasonable.

Recommendations

- Plan for a virtual meeting in the first half of 2021.

Discussion: What advice do you have regarding PDB Legacy Format sunseting?

The committee engaged in a discussion with the RCSB staff on the topic of sunseting the PDB legacy format, in favor of the mmCIF/PDBx format. The continued support required to serve legacy format files presents an additional burden for the RCSB, and the wwPDB more broadly. **The committee agrees that a timeline for a permanent transition needs to be established, communicated and enacted.** The details of that timeline will require some careful consideration. Firstly, depositions for all structure types will need to move to mmCIF/PDBx before the legacy format can be dropped. Currently, there are plans for cryo-EM submissions to move to mmCIF/PDBx later this year. The committee is still uncertain as to the plans for NMR-based structure depositions. Provided that the milestone of technique-wide mmCIF/PDBx deposition is either reached or scheduled, it seems reasonable to set a firm deadline for dropping support for serving the legacy format files to the RCSB user community. Clearly, extensive outreach and communication should be performed prior to this.

A general transitional approach to the legacy format was also raised by the committee, where the legacy format versions of structures currently available are frozen, and they are not changed any further. This would send the message that mmCIF is the only format capable of handling complex structures but wouldn't remove access to all structures on a specific date, which could present problems for some. This approach also minimizes the work necessary for the RCSB to that of creating a static repository of legacy format structures. The committee did see that there are some additional remaining issues. For example, biological assemblies are currently not available in mmCIF format from RCSB, although they are available from PDBe. If the legacy format is going to be dropped, these files need to be provided sooner rather than later. The mmCIF format is much better than the legacy format for assemblies, which use the MODEL-ENDMDL format to show multiple copies of the ASU in larger assemblies. There is also a good case for the RCSB to generate additional material to help educators know how to best use mmCIF format files. The PDB 101 has a site about beginners using mmCIF/PDBx, which is very helpful. However, it would be even better to have one or two videos for educators that explain how to teach with structures in this file format.

Recommendations

- Develop a timeline for sunseting the PDB format, which can be widely socialized for feedback. The timeline will need to be consistent with the termination of legacy format model submission for all experimental methods. Consideration could be given to providing an unsupported set of legacy version files to ease the transition for some researchers.
- Make it a priority to address cases where mmCIF files are currently not available, such as assemblies.

Appendix 1: Responses to 2019 RCSB PDB AC Recommendations

RCSB PDB thanks the Advisory Committee for their participation and thoughtful meeting report. Our responses to recommendations bulleted in the report follow (from July 2019) with updates noted as of March 2020.

AC: Recommendations for future meetings

- *Limit presentations to 15-20 minutes each, leaving adequate time for discussion. The content of the presentations was very thorough and helpful, but presenters need to be succinct in speaking to the content.*
- *It would be helpful for the advisory committee if the speakers could each provide a concise set of questions or topics for discussion at the end of their presentation. This would help move the discussions in directions most beneficial to the RCSB.*

RCSB PDB (2019): We plan to follow this model during the 2020 Advisory Committee teleconference (April 21) and subsequent meetings.

2020 Update: We are also providing updates to last year's response in advance of the meeting.

Deposition/Biocuration

- *AC: The team is encouraged to factor potential growth in experimental techniques, particularly cryo-EM into their deposition projections*

RCSB PDB (2019): We currently track and analyze these data as part of our review of Key Performance Indicators (KPIs).

2020 Update: [The KPI Report for 2019 is available for review.](#)

We have successfully coped with the 2019 increase in 3DEM depositions. To support continued success, we are preparing a proposal to extend the current BioSync resource that supports synchrotrons (<http://biosync.rcsb.org/>) to include EM facilities. This will enable more efficient data curation.

- *AC: The transition to mmCIF for crystallographic structure deposition in July 2019 is an excellent opportunity for the wwPDB as a whole. If successfully managed it can be expected to lead to a further increase in annotator efficiency. The team is encouraged to closely monitor the transition and work with community software developers to ensure that users have the tools for effective mmCIF deposition, and that the RCSB staff are prepared for the first wave in July.*

RCSB PDB (2019): We have been collaborating with the software developers represented in the PDBx/mmCIF Working Group, and working with the wwPDB organization to update related documentation and improve PDB to PDBx/mmCIF conversion tools. We plan to monitor this transition intensively and present the results of our analysis at the next Advisory Committee meeting.

2020 Update: As an ongoing collaboration effort with PDBx/mmCIF Working Group (*Acta Cryst* (2019. D75: 451-454 doi: [10.1107/S2059798319004522](https://doi.org/10.1107/S2059798319004522)), reported issues are either addressed or being addressed by refinement software developers. PDB to PDBx/mmCIF conversion tools have been improved with two releases and better documentation has been included by the software developers:

(<https://www.wwpdb.org/deposition/preparing-pdbx-mmCIF-files>).

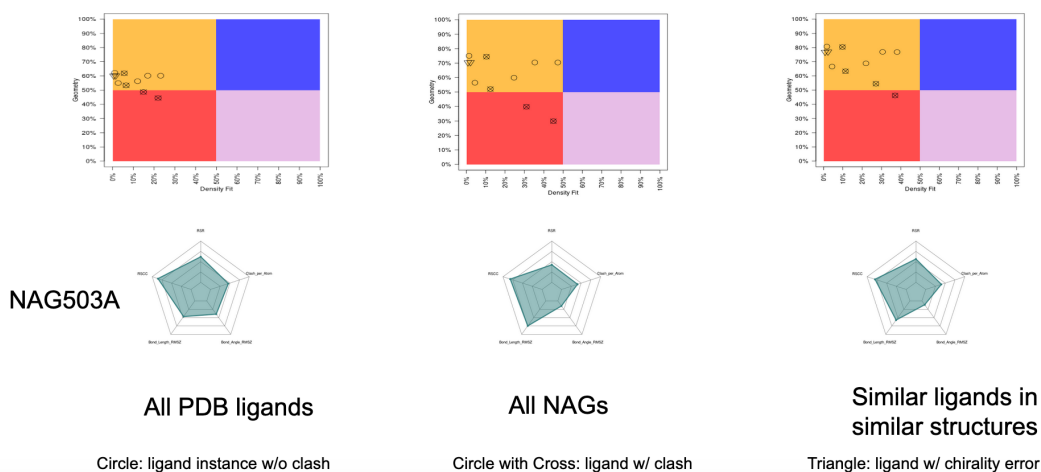
- *AC: The RCSB should consider providing visual views of the ligand validation information analogous to the wwPDB validation summary graphic for macromolecular structures on the structure summary pages - so that less knowledgeable users can easily see if there are issues with ligands.*

RCSB PDB (2019): We thank the Advisors for this suggestion. We are developing a value for an overall score and graphic that would provide a quantitative assessment of the ligand/s of interest. We also plan to provide an interactive 3D view of ligand validation. Both should be in production by the end of 2019.

2020 Update: The timeline for this project was moved to accommodate development on RCSB.org searching and reporting. Our 3D visualization viewer Mol* has enabled a validation coloring scheme that is accessible from a button next to the Validation Slider graphic on Structure Summary pages. This 3D view uses a coloring scheme to indicate if a part of the structure has no issues (blue), one issue (yellow), two issues (orange) or more (red).

A draft representation of the overall score has been developed (see below) and should be implemented in the new RCSB.org system Q3 of 2020.

1CQD NAGs, 4-Quadrant and One Radar Instance



Ligand NAG quality plots in entry 1CQD. Top row: quality percentile plots of geometry vs density fit for all NAG instances against all ligands (1st plot), all NAGs (2nd plot), and similar ligands (3rd plot) in the PDB archive. Circle indicates no clash; circle with cross: clash; and triangle: chirality error. Higher percentile indicates better quality.

Bottom row: quality radar plots of RSR, RSCC, bond length RMSZ, bond angle RMSZ, and atom clashes for ligand instance NAG503A against all ligands (1st plot), all NAGs (2nd plot), and similar ligands (3rd plot) in the PDB archive. Center indicates the best quality, hence a smaller area covered indicates better overall quality.

- *AC: The RCSB is encouraged to work with the carbohydrate community to provide open source software tools to handle the new carbohydrate description mechanisms - this will greatly increase the likelihood of other community software packages adopting these mechanisms.*

RCSB PDB (2019): We are actively collaborating with leadership of two relevant communities: computational bioinformatics (Robert Woods, University of Georgia) and carbohydrates (Martin Frank, Biognos AB; Thomas Lutteke, Justus-Liebig University Giessen; and Hans Vliegthart, Utrecht University). The project has, is, and will be publicized at glycoscience community conferences, and via

online announcements and documentation. The computational bioinformatics software used in this project is hosted publicly by The University of Georgia; we will publish software APIs for additional public use. Most importantly, we will continue efforts with the PDBx/mmCIF Working Group on carbohydrate representation support.

2020 Update: The project is on target for release of data in July 2020.

Remediated files were tested across wwPDB partner sites. News was [announced in February](#); the project [described in more detail at wwPDB.org](#); and [large example files published at GitHub](#) for community software adaption.

Archive Management/Access and Data Exploration

- *AC: The RCSB is encouraged to think about how new search features, and/or modification of existing approaches could help serve the broader community of visitors to the site. In particular ways to provide data out that are more meaningful to those not familiar with 3D structures.*

RCSB PDB (2019): We plan to adopt two approaches, both enabled by the new data out architecture. First, we will improve tracking of website usage to better gauge which search features are more actively used. Second, we will broaden the number of integrated data resources over the next four years (e.g., PubChem, Comprehensive Antibiotic Resistance Database) to expand the RCSB PDB user base and expose new communities to PDB data.

2020 Update: Our legacy system did not work well with the Google Analytics system. The first phase of improved tracking of website usage is implemented within the new search system. By using more meaningful URLs and establishing certain user “events,” we anticipate having a better understanding of user searching within Google Analytics. We are planning additional development to continue in this direction.

We have started to expand the data resources integration. Notably, connections to data in NIH Common Fund Data Resources (GTEX, IMPC, and Pharos) are now available from Advanced Search and our Structure Summary pages. A project to integrate PDB data with PubChem data is underway.

- *AC: The team should develop a firm timeline for delivery of the new data out technology to the user community.*

RCSB PDB (2019): Public beta testing will begin in December 2019, with the new site going into full production in 2020.

2020 Update: The new website went into production on April 8, 2020 at <http://RCSB.org>.

Outreach/Education

- *AC: The RCSB should consider making links to 3D printer files, along with instructions for printing more readily available from the RCSB website, as this might be of increasing interest to the community.*

RCSB PDB (2019): We thank the community for the suggestion. In response, we will provide curated 3D printer files to accompany the top accessed Molecule of the Month articles, starting in Q4 of 2019.

2020 Update: PDB-101 has launched a [curated collection](#) of files for 3D printing models of [alpha-amylase](#), [ferritin](#), [GFP](#), and [hemoglobin](#). Each 3D model file highlights a special feature of the

molecule and its function. Suggested modeling additions, such as making heme molecules, can help tell a molecular story.

An [overview](#) is provided to start to create models for any PDB structure.

Integrative Methods

- *AC: It was encouraging to see that there is now NSF support for some of these activities (development of the pipeline), with core RCSB support covering the remainder (implementation in OneDep). The proposed approach for hybrid methods validation stems from a recent workshop held in Baltimore. The team is applauded for seeking community input to help guide these efforts. Helen Berman is also recognized for her fund raising efforts to support the workshop, and her continued engagement in hybrid methods development through the NSF funding. Her continued participation in this project is viewed as key to its success. The committee expressed some concern about how some of the hybrid methods data would be validated, with cryo-electron tomography reconstructions being a particular issue. The RCSB approach of developing a pipeline that can run a broad range of hybrid method models/data, albeit imperfectly at the beginning, is appropriate - recognizing that “perfect is the enemy of good”. The proposed use of Bayesian approaches for validation in the future is very exciting, and has a good statistical basis. The RCSB is encouraged to consider this a topic for future fund raising.*

RCSB PDB (2019): Some of the work reported at the meeting was funded by an existing NSF grant to Helen Berman and Andrej Sali and we will develop future proposals that will be out of scope of core RCSB PDB award.

We acknowledge that validation of integrative methods data represents a significant challenge for the organization and we are very fortunate that there is separate NSF funding allocated to Andrej and Helen to address these issues prior to adoption by RCSB PDB. For the validation of the data themselves, we shall rely on the advice from the communities generating the data. We are currently engaged with Jill Trehwella to elaborate on this plan, using the SAS community as the first example, justified by their recent progress in organizing SASBDB. The EMDB group may provide a similar opportunity for obtaining data quality definitions for cryo-electron tomography reconstructions. We recognize the importance of Helen’s continuing efforts in organizing the other communities who are generating the data needed for integrative modeling.

2020 Update: The workshop White Paper has been published: Federating Structural Models and Data: Outcomes from A Workshop on Archiving Integrative Structures (2019) *Structure* 27: 1745-1759 doi: [10.1016/j.str.2019.11.002](https://doi.org/10.1016/j.str.2019.11.002).

Visualization

- *AC: The NGL library has been used by a few groups to develop new tools. The RCSB are encouraged to develop guidelines to help those groups transition to the Mol* system.*

RCSB PDB (2019): We will initially begin by helping our website users to use Mol* instead of NGL on the 3D View option of our Structure Summary pages. Financial support for additional software developer-power will be included in our proposal currently under development for submission to the Chan-Zuckerberg Institute in August 2019. We plan to develop additional resources in Q4 2019 to start to help developers make the transition from NGL to Mol*.

2020 Update: Mol* was released in Fall 2019 at RCSB.org and PDB.org. The source code is available on [Github](#). Individual components (including the Viewer) of Mol* can be readily used in

3rd party applications. Additionally, the viewer is available as an easily embeddable web component through the PDB [PDB component library](#).

RCSB PDB is currently developing tutorials and resources to support users.

We have not yet been successful in obtaining funding for Mol* development.

Management

- *AC: The RCSB leadership should develop a management plan to deal with increased activity that will require funding and additional wwPDB centers.*

RCSB PDB (2019): We plan to develop a proposal to submit to the NSF to provide funding for RCSB PDB activities necessary to help with establishing and supporting wwPDB centers in India and China.

2020 Update: A proposal to support establishing and training PDB China was submitted to the NSF.

- *AC: The RCSB should provide information to the committee at the next advisory meeting about how projects are prioritized across the center.*

RCSB PDB: We will provide this information for RCSB PDB projects at the next Advisory Committee meeting.

2020 Update: Under Jasmine Young's direction, the wwPDB software development and biocuration teams meet regularly to collaboratively plan and prioritize projects. These plans are reviewed by the wwPDB PIs. The plan is then presented for endorsement annually at the wwPDB Advisory Committee. The report should be made available shortly.

Funding

AC: There were many options for fundraising that were discussed during the meeting:

- *Although HHMI hasn't been successful, some committee members wondered if Janelia Farm might provide some opportunities given their recent call for ideas about their future directions.*
- *Philanthropic support is very challenging to obtain, but maybe it would be good to target prior supporters of science like Jim Simmons.*
- *The recent NCI ITCR program might provide some opportunities.*

RCSB PDB (2019): We will continue to target public and private funding, and are in the process of developing a proposal for the NCI ITCR program.

2020 Update: Submitted proposals are highlighted in the presentation; [summaries are available for review](#).

Recommendations

- *AC: A direct discussion with Susan Gregurick at NIH might help understand the landscape for upcoming data-related opportunities.*

RCSB PDB (2019): We will contact Susan Gregurick to set up a meeting during Stephen's planned Washington, DC visit in September 2019.

2020 Update: We are in regular contact with Susan Gregurick and are waiting for guidance on when to initiate submission of proposals under the new NIH-wide data resource/knowledgebase funding mechanisms.

- *AC: RCSB should develop, immediately, 1-2 pages documents describing potential projects (outside their core mission) that will enhance the RCSB activities. These will be invaluable for seeking funding from industrial sources.*

RCSB PDB (2019): We will develop summary documents as we create new proposals and share these with the Advisory Committee.

2020 Update: Submitted proposals are highlighted in the presentation; [summaries are available for review](#).

Expanding the User Base/Measuring Impact

- *AC: The RCSB should develop a concise justification for the need for an expanded user base, and the general topic of demonstrating impact, with clear goals.*

RCSB PDB (2019): We acknowledge that the imperative for expanding our user base could have been better articulated and will revisit this matter at the next Advisory Committee Meeting.

In addition, we will review the Key Performance Indicators enumerated in the RCSB PDB Project Execution Plan submitted in 2018, including

- Number of website visits/year
- Number of unique IP addresses served/year
- Website feature usage during year
- Number of Mobile/Tablet device sessions/year
- Number of laptop/desktop computer sessions/year
- Number of scientific publications citing RCSB PDB publications/year
- Number of scientific publications citing PDB structures/year
- Number of newly issued US patents and in process applications citing PDB/year
- Number of User issues addressed/year
- Number of Users of new tools for complex structure data

2020 Update: [The KPI Report for 2019 is available for review](#). All areas show positive growth trends from 2018 to 2019.

Continued federal funding depends critically on showing that utilization of the PDB archive and RCSB PDB websites increases year on year. The easiest way to accomplish these ends is to accumulate more users across a wider range of scientific and educational disciplines (as opposed to convincing existing users to make more intensive use of our resource offerings).

- *AC: The 50th anniversary of the PDB in 2021 provides a great opportunity for community outreach. The RCSB is encouraged to plan for high profile coverage of the event, including the national press. Examples discussed included special issues of journals, highlight articles in Science or Nature, CBS Sunday Morning, family days at NIH, a BBC special (online or broadcast), and engagement with educational networks, e.g. a Nova special.*

RCSB PDB (2019): We look forward to working with the Advisory Committee and the wwPDB partners on promotion of various PDB50 events worldwide.

2020 Update: Current plans are highlighted in the presentation.

In addition, Helen is working with *Journal of Biological Chemistry* on a special PDB50 issue.

- *AC: The RCSB should investigate opportunities for partnership with social sciences academic groups with expertise in researching user engagement and outreach. There may also be opportunities for funding projects to help in this area.*

RCSB PDB (2019): We will continue to collaborate with appropriate research teams at our host institutions, following on from the success of the [economic analyses](#) performed by the Rutgers Office of Research Analytics in 2017.

2020 Update: Collaborations initiated with Fred Ledley (Bentley University) and Carolyn Stein/Ryan Hill (MIT Economics Department), with the objective of publishing multiple economic/social impact papers before submission of the next renewal documents in early 2023 (for a January 1, 2024 new funding start).

- *AC: The RCSB has an opportunity to lead in the promotion of structural biology in science at a national level. This is a theme that could gain momentum if pursued in collaboration with other interested groups but will likely require engagement with federal officials and congress.*

RCSB PDB (2019): We will continue to study and publicize the value of structural biology, the PDB archive, and the RCSB PDB across the sciences.

2020 Update: Efforts are continuing as evidenced by our response to the COVID-19 pandemic online ([Resources page](#)) and in print (How to help the free market fight coronavirus (2020) *Nature* 580: 167 doi: [10.1038/d41586-020-00888-7](https://doi.org/10.1038/d41586-020-00888-7)).

Appendix 2: PDB 2019 Metrics

In aggregate, 13377 depositions were received and processed between January 1st and December 31st, 2019, with an average turnaround of two weeks. This represents an increase from the 12,179 entries deposited in 2018. Based upon the number of entries deposited in 2018 to date, it is estimated that an increased number of entries will be deposited this year.

Breakdown of depositions by discipline in calendar 2018 was as follows:

X-ray:	10952 (82% of entries deposited, up from 10,594 in 2018)
NMR:	403 (3%, down from 418 in 2017)
EM:	1996 (15%, up from 1,140 in 2017)
Other:	26 (.2%, down from 27 in 2017)

Breakdown of depositions by wwPDB processing site in calendar 2019 was as follows:

RCSB PDB:	5531 (41%)
PDBj:	3350 (25%)
PDBe-EBI:	4496 (34%)

Breakdown of depositors by location in calendar 2019 was as follows:

North America	31.7%
Europe	32.5%
Asia	24.6%
Commercial	6.0%
South America	1.0%
Oceania	4.1%
Africa	<1%

During 2019, RCSB PDB's website at <http://rcsb.org> was visited by millions of unique visitors.

During the same period, an estimated 838,269,170 million data files were downloaded from the PDB archive *via* the wwPDB member FTP and websites (RCSB PDB: 547,444,451; PDBj: 129,058,391; PDBe: 161,766,328).

Appendix 3: Post-Meeting Feedback from Jill Trehella

Sent via email May 16, 2020:

It is a very positive report indicating that things are progressing well. Very nice work promoting the impact and potential for structural biology in the current COVID-19 pandemic. The interview on NJTV was excellent and the PDB101 video on hand washing is something that can be extremely useful in educational contexts. I note the formal KPIs focus strongly on volume and quality of data, as well as utilization with funding as an implicit importance/significance metric. Would be good perhaps to think about how to incorporate elements of the COVID-19 response and pharmaceuticals generally into the KPIs.

I see funding continues as the challenge – I wonder if COVID-19 response will have positive or negative impact – I would be interested in your view of how funders' responses will balance clinical practice and advancing science.

It is hard to provide more insightful feedback having not been party to the discussion at the meeting, but I appreciate your making the materials available and I look forward to seeing PDB rise to the day's great challenge!