

2009 EM DataBank (EMDB) Advisory Committee Meeting Baylor College of Medicine, April 23, 2009

The EM DataBank (EMDB) Advisory Committee met at Baylor College of Medicine from 9 am – 3:30 pm, Friday March 20, 2009. The members in attendance were Joachim Frank (Columbia Univ.), Robert M. Glaeser (Chair, LBNL), B. S. Manjunath (UC Santa Barbara), and Maryann Martone (UC San Diego).

Dr. Wah Chiu (Baylor) opened the meeting with an overview of the objectives of the NIH grant (1R01 GM079429) that currently supports the establishment of the EMDB, which is a United Data Resource for CryoEM. His overview included a review of the deliverables for the end of the second year of this grant, which were described in a letter to Dr. Deatherage on June 20, 2007. In addition, Dr. Chiu showed examples of the types of specimens (single particles, filaments, 2-D crystals, and objects that can only be reconstructed by EM tomography) for which structures are obtained by electron microscopy, and the types of computational tools for segmentation and fitting that are needed in the analysis of these structures. Dr. Chiu next described the growth of EM depositions prior to 2006, the content of the database that had been developed at the European Bioinformatics Institute (EBI), and the independent systems for deposition, archiving and retrieval of EM data that had been developed at the EBI and at the RCSB PDB as of 2007. With this background, Dr. Chiu then gave an overview of the proposed EMDB system for data deposition, archiving and web services.

Dr. Kim Hendrick (EBI) then gave a further overview of the creation of the EM Data Bank by EBI in 2002, the structure for deposition and retrieval that existed in 2007, and the structure for deposition and retrieval that now exists, including identical portals in Europe and in the US. Dr. Hendrick also gave further information about current deposition statistics and some of the new tools for search and visualization.

Dr. Helen Berman (Rutgers Univ.) then moved on to address each of the deliverables that had been proposed for the end of the second year of the grant, which were defined in Dr. Chiu's June 2007 letter to Dr. Deatherage at NIH. Dr. Berman reported that, in addition to creating suitable management and communication protocols (deliverable "0"), the status for each of these is:

1. The web-based deposition system, EMDEP, is now open at both sites (EBI and PDB), as scheduled.
2. Goals have been met for extending the search capabilities of the retrieval engine, EMSearch.
3. 3DEM xml web services have been produced, as scheduled.
4. Concurrent deposition and retrieval of atomic coordinates and cryoEM maps became operational in Feb. 2009, as scheduled.
5. Development of a portable version of EMDEP, which can be downloaded for in-house use and that is capable for exporting complete depositions, is still underway, with on-time completion still expected by June 2009.
6. EMDB and PDBx dictionaries have been integrated, as planned.
7. The building of community consensus on expansion of this dictionary is ongoing, as is appropriate.

8. A technical report has been prepared on EM map “formats” – in this case the word refers to the wrapper and the content. This deliverable is not, in a sense, really completed, as Dr. Berman was not certain what would happen next, other than that one cannot keep discussing this issue forever.
9. A collaboration has been initiated with the developers of EMAN to automate the harvesting of metadata. Discussion on this topic emphasized, however, that the tone of this deliverable was undesirably “Baylor-centric”, and that the intent of the deliverable was indeed to open such collaborations to developers of any software package that would be the source of depositions to the EMDB. The Advisory Committee asks Dr. Chiu to appropriately reformulate (rephrase) this deliverable in future presentations.
10. The first web-based visualization tools were implemented on schedule in Feb. 2009.
11. Standard procedures have been established for data processing by the staff of EMDB.

Following Dr. Berman’s report on the outstanding progress that has been made so far in completing the proposed deliverables, future work that is now planned for this grant was described by three members of the staff, Cathy Lawson, Christoph Best, and Matt Baker.

The Advisory Committee is pleased to express how impressed they are with the excellence and the timeliness with which the deliverables proposed for the first two years of the grant have been completed. In addition, the Committee fully endorses the plans for future work that were described in the presentation (and in the booklet of slides that accompanied the presentation).

Additional items that the Advisory Committee discussed with Dr. Chiu and his colleagues included:

1. The Advisory Committee considers it to be an excellent idea for EMDB to be a source of “benchmark data” that various laboratories may be willing to deposit. Such data can be used by software developers, and individuals can use such data to become familiar with the use of data-analysis software. The Committee encourages the EMDB to publicize its capability to accept deposits of such data sets, and perhaps to even recruit such data sets from research groups who are likely to have ones that could be shared. The Committee also encourages the EMDB to publicize the availability of such benchmark data, not just to the EM community but also more widely to the computer vision community. In the context of outreach to the latter community, the EMDB might consult with a few representatives for advice on how to make the website easier to navigate and the nature of the data more easily understandable by a more general audience.
2. The Advisory Committee commends the participants in this grant for the amount of their effort to interact with the EM community and to solicit their input through workshops such as the ones that were held at Rutgers and at San Francisco. The Committee believes that it would be a good thing to continue to be proactive in its outreach at meetings such as the Biophysical Society or the 3-D EM Gordon Conference, as well as through organizing additional community input workshops.

3. Dr. Berman addressed the issue that work needs to be done to simplify and speed up the deposition process, which at present is too cumbersome and time-consuming. As she explained, this is an issue that is being actively worked on at the moment for depositions to the PDB, and she expects that whatever is developed for the PDB will also translate to and benefit depositions to the EMDB.

Submitted on behalf of the Advisory Committee,

Robert M. Glaeser, Chair

Cc: Joachim Frank
B. S. Manjunath
Maryann Martone