June 21 meeting notes

Scope of the Challenge

different possibilities (Paul):

- reoptimize an existing model+map to make the best model
- fit and optimize an X-ray/NMR/EM-model/homology model
- ab initio building (manual or automated) and optimization : automated

Randy: Could be CASP-like, but how do we keep people from being influenced by a known answer

Klaus: How to evaluate is an important part of the decision

Method: must explicitly specify -- must be able to demonstrate method use.

Rules to keep people from “gaming”

Wah: What are the urgent needs of the community -- how to nurture participation.

Here we won’t have gold standard (unlike CASP).

New validation methods developed through the challenge can be discussed via a forum on the challenges website.

Special journal issue -- Nature Methods? Structure? Open access cost should be considered.

Biopolymers, Biophysics Progress (T. Blundell is an editor).

Randy: Have people who are doing validation also upload their analysis to the challenge site?

At the end there will be a meeting to discuss the validation results. Wah: possibly more meetings?

Torsten: CASP has been beneficial: -- “publication hygiene” (less junk) -- triggered development of methods/validation procedures. how to evaluate model quality? CASP has been running for 24 years. ~100 groups participate. CASP ranking can be problematic r.e. funding.

One possible outcome is new ways of combining methods.

Ligands in addition to backbone.

June 21 meeting notes

Map Datasets

ribosome significantly bigger modelling problem than the others.

multiple categories of targets?

Helen suggests maximum of 10

Resolution between 3-4 seems the most challenging.

Suggestion: 5 mandatory, 5 optional

work with the map depositors so that they will feel it is a collaboration

10 with variety of sizes, types of secondary structure.

target maps should be unsharpened, unmasked -- half-maps available too

Evaluation criteria

Jane -- CabLAM Validation
70S ribosome beta-strand issues detected
beta-galactosidase 2.2 : CabLAM picks up a number of issues
based on 2 C-alpha torsions + carbonyl pseudo-angle

2' vs 3' pucker is diagnostic for RNA

EMRinger is also a way to evaluate/correct a model

Local cross-correlation --
Z-scores (resolution dependent)

matrix of structures vs. methods

Logistics

File format: CIF vs. PDB…

PDB extract can be used to convert PDB to CIF

June 21 meeting notes

mmCIF working group -- help PDB w/ getting community to adopt CIF.

we could have people deposit PDB and then internally convert to CIF.

create a separate site based on D&A for the model challenge depositions.

template for order of chains? ground rules -- if you are re-refining a structure, use same order as published structure.

decide on test data in 1 month --by 1st week of August.

then need to go back to the target providers and ask for additional info as necessary (half-maps, unsharpened, unfiltered).

advocate for computing resources for the challenge -- rather than requiring individuals to seek resources themselves. They will want to know how much cpu is needed.

e.g. exceed program… get program officers on board. Paul will ask at LBL supercomputer center. NSF cyberinfrastructure--Helen can contact. Wah-PNNL.

September-October announce. How long ? ~~end of summer--Aug 2016. evaluations in fall 2016.~~ Better if offset from CASP competition? (Mar-end of July). So then ~April 2016. Validation period 3 months -- meeting next Sept. 2016.

spread out announcement of targets?

Blind assessment.

email list