PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

CASP INTRODUCTION—ROUND VIII

Critical assessment of methods of protein structure prediction—Round VIII

John Moult,[1]* Krzysztof Fidelis,[2] Andriy Kryshtafovych,[2] Burkhard Rost,[3] and Anna Tramontano[4]

[1] Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850
[2] Genome Center, University of California, Davis, California 95616
[3] Department of Biochemistry and Mol Biophysics, CUBIC, C2B2, NESG, Columbia University, New York, New York 10032
[4] Istituto Pasteur-Fondazione Cenci Bolognetti, University of Rome "La Sapienza", P.le Aldo Moro 5, Rome 00185, Italy

ABSTRACT

This article is an introduction to the special issue of the journal *Proteins*, dedicated to the eighth CASP experiment to assess the state of the art in protein structure prediction. The article describes the conduct of the experiment, the categories of prediction included, and outlines the evaluation and assessment procedures. Highlights are the first blind assessment of model refinement methods showing that under some circumstances substantial model improvements are possible; improvements in the performance of methods for determining the accuracy of a model; and some progress in the accuracy of comparative models in regions not present in a principal template. Against these advances must be stacked the fact that there is no detectable progress in model quality compared with CASP7 in either template-based or template free modeling, using the established CASP measures.

Proteins 2009; 77(Suppl 9):1–4.

Key words: protein structure prediction; community wide experiment; CASP.

INTRODUCTION

This issue of *Proteins* is devoted to articles reporting the outcome of the eighth community wide experiment to assess methods of protein structure prediction (CASP8) and related activities. There have been seven previous CASP experiments at 2 years intervals from 1994 through 2006, and these were reported in previous special issues of *Proteins*.[1–7]

The primary goals of CASP are to establish the capabilities and limitations of current methods of modeling protein structure from sequence, to determine where progress is being made, and to determine where the field is held back by specific bottlenecks. Methods are assessed on the basis of the analysis of a large number of blind predictions of protein structure.

This article outlines the structure and conduct of the experiment and is followed by descriptions of the data handling procedures and numerical analysis methods[8] and the CASP8 target proteins.[9] For the first time, there is an article describing numerical analysis of the results using the now standard set of measures.[10] In previous CASPs, these data have been presented in the assessors' articles, but the methods are now sufficiently stable that the analysis has become routine. It is still left to the assessors to interpret the significance of these data. There are articles by the assessment teams in each of the three-dimensional prediction categories: template-based modeling,[11] template free modeling,[12] and model refinement.[13] There are four papers assessing other aspects of structure modeling: domain and three-dimensional (3D) contact identification,[14] assessment of disorder,[15] prediction of model quality,[16] and prediction of ligands binding sites.[17] All these categories have been included in

The authors state no conflict of interest.
Grant sponsor: National Library of Medicine; Grant number: LM07085; Grant sponsor: NIH Institute of General medical Sciences; Grant number: GM072354; Grant sponsor: EMBO.
*Correspondence to: John Moult, Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, MD 20850. E-mail: moult@umbi.umd.edu
Received 10 August 2009; Accepted 14 August 2009
Published online 21 August 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22589

one or more previous CASPs. Changes since last time include an increasing emphasis on ability to predict the quality of a model and restriction of function prediction to just prediction of which residues of a model are in contact with a known bound ligand. The assessment articles are followed by 14 articles from some of the more successful prediction teams, spanning the prediction categories. Finally, there is a article putting the results of this CASP in the context of the previous ones and highlighting the areas where progress has been made.[18] As always, the assessors' articles are probably the most important in the issue and describe the state of the art as found in CASP8.

THE CASP8 EXPERIMENT

The structure of the experiment was very similar to that of the earlier ones, with a prediction season of about 3 months, and three main steps:

1. As in recent CASPs, participants were required to register for the experiment in one or both of two categories: as human teams, where a combination of computational methods and investigator expertise is used; and as servers, where methods are only computational and fully automated, so that a target sequence is sent directly to a machine. In some cases, the same investigators registered in both categories.
2. Information about "soon to be solved" structures was collected from the experimental community and passed on to the prediction community. As in CASP7, and in contrast to earlier experiments, nearly all targets were obtained from the structural genomics community, particularly the NIH PSI centers (http://www.nigms.nih.gov/Initiatives/PSI) and the structural genomics consortium (http://www.sgc.utoronto.ca/). The high throughput of structural genomics again allowed us to collect in excess of 100 targets.
   One hundred and twenty eight protein sequences were released for prediction. Of these, seven were cancelled, leaving 121 for assessment in the 3D and contact prediction categories. These were divided into domains, each of which was treated as a separate evaluation unit. In all, 164 domains were included. Two additional targets—T0484 and T0500—were assessed in the disorder prediction category only. With help from the PDB, information about appropriate targets was sent directly to the CASP prediction center. Target information was verified and appropriately formatted.
   For the first time, targets were divided into two sets. All targets were sent to registered servers, and a subset of 57 was identified as designated for human teams as well. Restriction of the number of targets for human teams was introduced in response to requests from the prediction community and reflects the fact that human teams cannot devote adequate attention to a large number of targets in such a short time. Human teams were required to deposit structure models and other predictions by a specified deadline, usually 3 weeks after release of the targets. Servers were required to return predictions within 72 hours of receiving target information. No experimental structures were released before the prediction deadlines. As in previous CASPs, prediction groups were limited to a maximum of five models per target and were instructed that most emphasis would be placed on the model they designated as the best (referred to as 'model 1'). The level of participation in the CASP experiment continues to be high, with 233 predictions groups, similar to that of CASP7 (253 groups) and representing a large fraction of the relevant community.
3. The models were compared with experiment, using numerical evaluation techniques and human assessment, and a meeting was held to discuss the significance of the results.

MANAGEMENT AND ORGANIZATION

CASP requires careful data management and security, and mechanisms to ensure that the prediction community is informed and consulted. The principal components are:

A. Organizers: The authors of this article, responsible for all aspects of the organization of the experiment and meeting. The organizers are the same as for CASP7, except that Tim Hubbard, a key organizer since CASP2 has stepped down.
B. The FORCASP web site (www.FORCASP.org): FORCASP provides a forum, where members of the prediction community may discuss aspects of the CASP experiment.
C. Predictors' meeting: During each CASP conference, there is a predictors' meeting with votes on issues of CASP policy, particularly major changes and extensions of the CASP process.
D. Independent assessors: The independent assessors have primary responsibility for judging the quality of the predictions received and for commenting on the current state of the art. Assessors are provided with numerical analysis data generated using approved procedures and may also add their own numerical methods.
E. Protein structure prediction center: The prediction center is responsible for all data management aspects of the experiment, including the distribution of target information, collection of predictions, generation of numerical evaluation data, developing tools for data analysis, data security, and maintenance of a web site where all data are available. Details of these aspects of the experiment are described in Ref. 8.

CATEGORIES OF PREDICTION

As in CASP7, modeling targets were divided into two categories: template-based modeling, where a relationship to one or more experimentally determined structures could be identified, so providing a modeling template and template free modeling; where there are either no usefully related structures or the relationship is so distant that it cannot be detected.

In CASP7, an additional high-accuracy modeling subcategory was also used, for targets where problems of alignment and template coverage were expected to be sufficiently small that the accuracy of resulting models should be competitive with experimental structures. For CASP8, more detailed analysis was performed on all template-based models, including packing, hydrogen bonding, side chain orientation, and usefulness of a model for molecular replacement.

COLLECTING AND VALIDATING PREDICTIONS

There were a total of 80,560 models deposited in CASP8 of which 55,130 are 3D co-ordinate sets. A further 1241 are alignments that were converted into co-ordinates for assessment. The remainder are residue–residue contacts (4154), domain assignments (2335), disorder predictions (4441), binding site predictions (5898), 3D model quality predictions (6466), and model refinements (892). All predictions were submitted to the prediction center in a machine readable format. Accepted submissions were issued an accession number, serving as the record that a prediction had been made by a particular group on a particular target.

NUMERICAL EVALUATION OF PREDICTIONS

As noted earlier, a number of numerical evaluation techniques have become widely accepted by the prediction community and so are regarded as standards. The accuracy of 3D structure models is primarily evaluated using two metrics: GDT_TS, a multi threshold measure of the difference in main chain Cα atoms between a model and the corresponding experimental structure[19] and AL0, a measurement of alignment accuracy showing how well the assigned amino acid positions accord with those in the experimental structure. An alternative measure of alignment accuracy based on a dynamic programming procedure (SWALI[20]) is also used to establish maximum possible alignability between the target and a single template. The prediction center also provided results from DAL, MAMMOTH, and ACE software to the assessors to facilitate their structural analysis. As in previous CASPs, the assessors for the template free category found that GDT_TS is useful for shortlisting the most noteworthy models, but that visual inspection is necessary to obtain a final ranking.[12] The assessors also used their own measures and approaches to complement the established CASP ones. Evaluation methods remained the same as in CASP7 for disorder, domain, and contact prediction.

The numerical evaluation metrics, though critical, are not generally sufficient to draw final conclusions about the quality and usefulness of modeling methods. A key principle of CASP is that primary responsibility for assessing the significance of the results is placed in the hands of independent assessors. This continues to be a major source of insight in CASP, as well as ensuring that organizer biases are not imposed on the outcome.

MEETINGS, WEB SITE AND PUBLICATIONS

A 1 day "Between CASPs" public meeting was held in Madrid in April 2008, co-organized with Alfonso Valencia. The aim of these meetings is to bring the CASP results to a less specialized audience than attend the regular workshops. The first CASP8 planning meeting, attended by the assessment teams for CASP8 and the previous assessors, was held in association with the Madrid event. Following standard CASP procedure, a second planning meeting was held after the close of the prediction season at which the assessors presented their results to each other and to the organizers. As always, prediction team identities were hidden from the assessors until after those presentations, to avoid ranking bias.

The meeting to discuss the outcome of the experiment was held in Cagliari, Sardinia in December. The format of the meeting was once again slightly modified. The first day was devoted to template-based modeling and the related topic of model quality prediction. The second day covered template free modeling and a session organized by junior participants. The third day dealt with the non 3D categories of disorder, domains, and contacts as well as a discussion on the nature of protein structure space. Finally, the last morning of the meeting was devoted to the new topic of model refinement. There were a number of other sessions and group meetings. The full program can be found on the prediction center web site.

This issue of *Proteins* is the official report of the CASP8 experiment. As usual, predictors submitting articles were urged to concentrate on what went right, what went wrong, and where possible, to explain why, and what they learned as a result. Because of space limitations, details of the methods are often absent, and readers are requested to turn to the references for more information. All the prediction and assessment articles in this issue have been peer reviewed. The CASP web site (http://predictioncenter.org) provides extensive details of the targets, the predictions, and the numerical analyses.

Discussions of some issues can also be found on the FORCASP site (www.FORCASP.org).

PROGRESS IN CASP8

In contrast to some earlier CASPs, progress between CASP7 and CASP8 was modest. Notable, though minor progress was found in two areas: First, improved added value in template models (inclusion of information not available from a single template) was evident by two measures—greater accuracy of the regions of template model not covered by the single best template and a higher number of models where the number of residues correctly aligned is greater than could be obtained by copying a single best template.[18] Second, there was an improvement in ability to judge the quality of a model, as evidenced by a better ranking by predictors of their own submissions,[18] and improved ability to predict the quality of CASP models overall.[16] In addition, and most noteworthy, significant refinements over originally submitted best models were achieved for a number of targets by a number of groups. There is considerable variation in performance of the refinement methods, and it is not yet clear what factors are important. But overall, these first results are very encouraging.[13]

FUTURE DEVELOPMENTS

There will be a CASP9 experiment running from the Spring of 2010 and culminating in a meeting in December of that year. The meeting is planned to take place at Asilomar Conference Center, in the USA. There will likely also be a "Between CASPs" meeting early in 2010 aimed at a broader audience. Those interested in any of these areas should check the CASP web site for further announcements.

ACKNOWLEDGMENTS

The authors are grateful to the members of the experimental community, particularly the structural genomics centers who agreed to provide targets. Taking part required courage and hard work on the part of all the predictors. Once again the assessment teams worked extremely hard and effectively to extract major insights from the results. The authors again thank *Proteins* for providing a mechanism for peer reviewed publication of the results. We are also grateful to *Wiley* and *Proteins* for agreeing to make it open access, so that all scientists may easily make use of the results. The authors thank Helen Berman and the PDB staff for their key role in structural genomics target processing. As noted earlier, Tim Hubbard has stepped down as CASP organizer. The authors have greatly appreciated Tim's many valuable contributions and unflagging hard work on the part of CASP.

REFERENCES

1. Moult J, Pedersen JT, Judson R, Fidelis K. A large-scale experiment to assess protein structure prediction methods. Proteins 1995; 23:ii–v.
2. Moult J, Hubbard T, Bryant SH, Fidelis K, Pedersen JT. Critical assessment of methods of protein structure prediction (CASP): round II. Proteins 1997; (Suppl 1):2–6.
3. Moult J, Hubbard T, Fidelis K, Pedersen JT. Critical assessment of methods of protein structure prediction (CASP): round III. Proteins 1999; (Suppl 3):2–6.
4. Moult J, Fidelis K, Zemla A, Hubbard T. Critical assessment of methods of protein structure prediction (CASP): round IV. Proteins 2001; (Suppl 5):2–7.
5. Moult J, Fidelis K, Zemla A, Hubbard T. Critical assessment of methods of protein structure prediction (CASP)-round V. Proteins 2003;53(Suppl 6):334–339.
6. Moult J, Fidelis K, Rost B, Hubbard T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP)–round 6. Proteins 2005;61(Suppl 7):3–7.
7. Moult J, Fidelis K, Kryshtafovych A, Rost B, Hubbard T, Tramontano A. Critical assessment of methods of protein structure prediction-Round VII. Proteins 2007;69(Suppl 8):3–9.
8. Kryshtafovych A, Krysko O, Daniluk P, Dmytriv Z, Fidelis K. Protein structure prediction center in CASP8. Proteins 2009;77(Suppl 9):5–9.
9. Tress ML, Ezkurdia I, Richardson JS. Target domain definition and classification in CASP8. Proteins 2009;77(Suppl 9):10–17.
10. Cozzetto D, Kryshtafovych A, Fidelis K, Moult J, Rost B, Tramontano A. Evaluation of template-based models in CASP8 with standard measures. Proteins 2009;77(Suppl 9):18–28.
11. Keedy DA, Williams CJ, Headd JJ, Arendall WB, III, Chen VB, Kapral GJ, Gillespie RA, Zemla A, Richardson DC, Richardson JS. The other 90% of the protein: Assessment beyond the C-alphas for CASP8 template-based models. Proteins 2009;77(Suppl 9):29–49.
12. Ben-David M, Noivirt O, Paz, A, Prilusky Jaime, Sussman J, Levy Y. Assessment of CASP8 structure predictions for template free targets. Proteins 2009;77(Suppl 9):50–65.
13. MacCallum J, Hua L, Schnieders M, Pande V, Jacobson M, Dill K. Assessment of the protein-structure refinement category in CASP8. Proteins 2009;77(Suppl 9):66–80.
14. Ezkurdia l, Gonzalez-Izarzugaza J, Graña O, Tress M. Assessment of domain boundary predictions and the prediction of intramolecular contacts in CASP8. Proteins 2009;77(Suppl 9):196–209.
15. Sussman J, Prilusky J, Noivirt O. Assessment of disorder predictions in CASP8. Proteins 2009;77(Suppl 9):210–216.
16. Cozzetto D, Kryshtafovych A, Tramontano A. Evaluation of CASP8 model quality predictions. Proteins 2009;77(Suppl 9):157–166.
17. Lopez G, Ezkurdia L, Tress M. Assessment of ligand binding residue predictions in CASP8. Proteins 2009;77(Suppl 9):138–146.
18. Kryshtafovych A, Fidelis K, Moult J. CASP8 results in context of previous experiments. Proteins 2009;77(Suppl 9):217–228.
19. Zemla A. LGA: A method for finding 3D similarities in protein structures. Nucleic Acids Res 2003;31:3370–3374.
20. Kryshtafovych A, Fidelis K, Moult J. Progress from CASP6 to CASP7. Proteins 2007;69:194–207.