PROTEINS: Structure, Function, and Genetics 53:334–339 (2003)

Critical Assessment of Methods of Protein Structure Prediction (CASP)-Round V

John Moult,[1] Krzysztof Fidelis,[2] Adam Zemla,[2] and Tim Hubbard[3]
[1]*Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland*
[2]*Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California*
[3]*Sanger Institute, Wellcome Trust Genome Campus, Cambridgeshire, United Kingdom*

ABSTRACT **This article provides an introduction to the special issue of the journal *Proteins* dedicated to the fifth CASP experiment to assess the state of the art in protein structure prediction. The article describes the conduct, the categories of prediction, and the evaluation and assessment procedures of the experiment. A brief summary of progress over the five CASP experiments is provided. Related developments in the field are also described. Proteins 2003;53:334–339.** © **2003 Wiley-Liss, Inc.**

Key words: protein structure prediction; communitywide experiment; CASP

INTRODUCTION

This issue of *Proteins* is devoted to articles reporting the outcome of the fifth communitywide experiment to assess methods of protein structure prediction (CASP5) and related activities. Four previous CASP experiments in 1994, 1996, 1998, and 2000 were reported in previous special issues of Proteins[1–4] as well as elsewhere.[5–13] Independent discussions of CASP5 have also appeared.[14,15]

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. With ~10 years of effort now recorded, these latter factors—progress or the lack of it—have assumed increasing importance. 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 a description of the CASP5 target proteins. There are sections of the special issue for each of the main CASP prediction categories: Comparative Modeling, Fold Recognition, and New Fold methods. These sections begin with an article by the assessment team in that area and continue with contributions from the prediction groups the assessors considered to have done the most interesting work. The number of predictor articles in each category varies—there are only three in comparative modeling, four in the fold recognition, and seven in the new folds category. The small number of articles in comparative modeling reflects the fact that there again appears to have been little progress in this area since the last CASP. The assessors' articles are probably the most important in the whole issue and describe the state of the art as they found it in CASP5.

The role and importance of automated servers in the structure prediction field continue to grow. Another main section of the issue deals with this topic. The first of these articles describes the CAFASP3 experiment. The goal of CAFASP is to assess the state of the art in automatic methods of structure prediction.[16] Whereas CASP allows any combination of computational and human methods, CAFASP captures predictions directly from fully automatic servers. CAFASP makes use of the CASP target distribution and prediction collection infrastructure, but is otherwise independent. The results of the CAFASP3 experiment were also evaluated by the CASP assessors, providing a comparison of fully automatic and hybrid methods. Full information is available at the Web site (http://www.cs.bgu.ac.il/~dfischer/CAFASP3/). This first article is followed by three that report some of the more interesting CAFASP results.

Large-scale benchmarking of prediction server performance is reported in the following two articles: one for Livebench,[16] and one for EVA.[17] In contrast to CASP and CAFASP, the benchmarking experiments run continuously. Both Livebench and EVA operate by sending the sequences of just released PDB entries to automatic prediction servers and collating and analyzing the results over time. Livebench focuses primarily on fold recognition and EVA primarily on secondary structure predictions. Livebench and EVA are entirely independent of CASP, and we are grateful to the organizers for their participation in the CASP5 meeting and in contributing to this issue of *Proteins.* Benchmarking experiments complement CASP, particularly by clarifying issues of the statistical significance of the results.

Prediction of disorder in protein structures was included for the first time in CASP5. An assumption behind the prediction of protein structure is that, under specified environmental conditions, every protein molecule has es-

Grant sponsor: Lawrence Livermore Laboratory; Grant sponsor: Department of Energy; Grant sponsor: National Library of Medicine; Grant number: LM07085; Grant sponsor: National Institutes of Health; Grant number: GM/DK61967; Grant sponsor: HP; Grant sponsor: Amgen; Grant sponsor: GlaxoSmithKline; Grant sponsor: IBM.

*Correspondence to: John Moult, Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Dr., Rockville, MD 208.50, E-mail: moult@umbi.umd.edu

Received: 21 June 2003; Accepted 24 June 2003

sentially a single functional three-dimensional structure. A number of experimental studies have established that this does not always appear to be the case.[18] Thus, the ability to predict disorder is of considerable importance. The section on this topic contains three articles: one providing an evaluation of the predictions and two describing some of the more interesting results.

Two additional articles complete the issue. The first was chosen by a vote on the FORCASP Web site (see below) as representing work done with particularly interesting methods. It is hoped that in future CASPs it will be possible to include more methods-related material. The last article is an evaluation of progress since CASP4 using similar methods to those for progress evaluation after CASPs 3[19] and 4.[20]

The CASP5 Experiment

The structure of the experiment was very similar to that of the earlier ones and consisted of three steps:

1. Information about "soon to be solved" structures was collected from the experimental community and passed on to the prediction community. Target information was made available through the CASP Web site and sent directly to registered CAFASP servers.
2. Prediction teams deposited models of the structures before the experimental results were public. For CASP, deposition was required by a specified deadline. For CAFASP, servers were required to respond within 48 hours
3. The models were compared with experiment by using numerical evaluation techniques and human assessment, and a meeting was held to discuss the significance of the results.

Management and Organization

CASP has a multilevel structure, intended to ensure substantial input from the prediction community:

A. Organizers. The authors of this article, responsible for all aspects of the organization of the experiment and meeting.
B. Consultancy groups. Three groups of ~10 veteran CASP predictors each, one for each of the three primary prediction categories. These groups, first introduced in CASP3, are involved in the selection of the independent assessors, are influential in the choice of numerical evaluation methods, and provide advice on other aspects of the experiment.
C. Predictors' meeting at Asilomar. During each CASP conference, there is a predictors' meeting with votes on issues of CASP policy, particularly the timing of the next experiment, the organization team for the next experiment, and 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 commenting on the current state of the art. Assessors are provided with numerical analysis data generated with approved methods, and may also add their own numerical methods.
E. Protein Structure Prediction Center at Lawrence Livermore Laboratory. 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 of predictions, collection of numerical evaluations from other workers, and maintenance of a Web site where all data are available. Details of these aspects of the experiment are described in Ref. 21.
F. The FORCASP Web site (www.FORCASP.org). As discussed below, FORCASP provides a forum where members of the prediction community may discuss aspects of the CASP experiment.

Collection of Targets

X-ray crystallographers and NMR spectroscopists were solicited to provide information about structures that were either expected to be solved shortly or that were already solved but had not yet been discussed in public. U.S. structural genomics projects were also asked to contribute prediction targets, and more than half of the targets came from that source. Target information was made available to predictors through a Web interface. Details of 67 structures were obtained, and of these, 60 were solved in time to be included in the assessment. A number of these targets were divided into two or more domains for assessment purposes. The total again falls short of what we had hoped for (~100), but because of the participation of the structural genomics projects, it is about 50% larger than in CASP4.

Categories of Prediction

The quality of a structure model depends on how much information from already known structures can be used—at one extreme, models competitive with experiment can be produced for proteins with sequences very similar to that of a known structure. At the other, models for proteins with no detectable sequence or structure relationship to one of known structure are still at best very approximate. In all the CASPs so far, targets have been divided into three broad categories, reflecting how extensively models could be based on knowledge of other structures. There is general agreement among the organizers and assessors that changes in the nature of structure modeling have made these broad categories outdated, and a major revision will be introduced in CASP6. In CASP5, as in CASP4,[22] this issue was partly addressed by the use of seven finer grained categories. For CASP5, each assessor was still asked to focus on one of three original categories, recognizing there would be some overlap in coverage of the targets. The three broad categories are as follows:

1. Comparative or homology modeling. When the sequence of the target structure is clearly related to that of one or more structures, the structures will also be similar. Thus, an approximate model can be created simply by copying related regions of polypeptide from

the parent structures and changing the side-chains where necessary. A total 37 target domains were considered by the assessors to be in the comparative modeling category. These domains were divided into two finer categories: the 22 that could be related to known structures using a simple BLAST search (high-sequence identity) and the 15 where a relationship to a known structure could be identified by using moderately sophisticated PSI-BLAST searches (low-sequence identity). Models for the high-sequence identity set were analyzed in more detail than the rest, considering the accuracy of side-chains, the construction of regions not present in available template structures, and whether the overall backbone accuracy is higher than that obtained by simply copying the best template. The boundary between the comparative modeling category and fold recognition has become blurred, because continually improving sequence comparison techniques have made it possible to reliably identify folds where the sequence relationship would once have been regarded as being in the twilight zone. A further six targets were considered to be in the overlap (CM/FR) region between comparative modeling and fold recognition.

2. Fold recognition. Increasingly, new structures deposited in the protein data bank turn out to have folds that have been seen before, even though there is no obvious sequence relationship between the related structures. Thus, methods of identifying folds from sequence information continue to grow in importance. There are two main questions to be asked: how successful are the different methods at identifying fold relationships, and when successful, what is the quality of the models produced? Techniques for fold recognition include advanced sequence comparison methods, comparison of predicted secondary structure strings with those of known folds, tests of the compatibility of sequences with three-dimensional folds (threading), and the use of human expert knowledge. Evaluation of the quality of the models produced has common components with comparative modeling, specifically alignment accuracy, and with new fold methods, specifically recognizing correct architecture, even in cases where the topology is incorrect. Targets in this category are subdivided into domains that are considered to have diverged from a common ancestor of known structure— homologous folds [FR(H)] and domains that are considered more likely to resemble known structures as a result of convergent evolution–analogous folds [FR(A)]. There are six FR(H) targets and seven FR(A) targets. In addition, the six target domains in the overlap region with comparative modeling (CM/FR targets) were also considered in this category, as were a further six targets that overlap with the new fold category (NF/FR). In all, a total of 25 target domains were evaluated by the fold recognition assessor.
3. New fold methods. In early CASPs, targets where there was no relationship to an already known complete structure were described as ab initio. This name implies that there is no reliance on known structures in building models. In practice, most of the methods used for such targets do make extensive use of available structural information, both in devising scoring functions to distinguish between correct and incorrect predictions and in choosing fragments to incorporate in the model. For this reason, the category was renamed, starting in CASP4. A wide range of knowledge-based techniques are used: well-established secondary structure prediction tools, sequence-based identification of sets of possible conformations for short fragments of chain, methods that assemble three-dimensional folds from candidate fragments and predicted secondary structure; prediction of which residues are in contact in the structure; "mini-threading" methods that identify supersecondary structure motifs; and full-domain fold recognition methods that may establish an approximate or partial topology. These approaches are sometimes combined with numerical search methods such as molecular dynamics, Monte Carlo, and genetic algorithms. There are a few pure ab initio methods, usually based on some form of numerical simulation techniques together with more traditional "empirical potentials."

Important evaluation criteria in the new fold category are the fraction of the structure that is predicted below a specified error level and recognition of success in identifying general architecture. Some targets are somewhat similar to a known fold and are classified as midway between the fold recognition and new fold categories. In all, six target domains in the analogous fold/new fold (NF/FR) category and six new fold domains were considered.

Level of Participation

A high level of participation from the prediction community is critical to the success of the experiment. As usual, participation was solicited through announcements in published articles and news groups, a Web site, and direct approaches to known prediction groups. Overall participation has steadily increased over the CASPs from 34 groups in CASP1, then 70, 98, 163, and in CASP5, 216. Figures for the last three CASPs include CAFASP participants.

Collecting and Validating Predictions

The total number of models received in CASP5 exceeds 28,000. This is again an increase from the previous CASP, where there were approximately 11,000. As before, all predictions were required to be in a machine readable format. Submission through a Web interface was encouraged to allow for predeposit checking of format conformity and prediction completeness. All submissions were processed by the Prediction Center at the Lawrence Livermore Laboratory.[21] Accepted submissions were issued an accession number that served as the record that a prediction had been made by a particular group on a particular target. A final acceptance time was established for predictions on each target, determined by the expected release date of the experimental structure. The prediction season

ran from spring until mid-September 2002. As previously, predictors 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 (often referred to as model 1).

Numerical Evaluation of Predictions

Numerical evaluation of model structures remains an imperfect science. Evaluation must not only provide an overall picture of model quality but must provide information about specific areas, so that the bottlenecks to progress can be precisely pinpointed. Furthermore, different categories of modeling achieve different levels of accuracy and provide different levels of detail, making it hard to devise universal measures. The evaluation methods also provide one of the primary means of comparing the performance of different groups, and so must be agreed and respected by the community.

CASP evaluation is based on comparison of each model with the corresponding experimental structure. Numerical evaluation criteria have been moderately stable for the last two or three CASPs. In CASP5, the GDT_TS[23] measure has been used by all three assessors as the principal metric of main-chain accuracy. In comparative modeling, alignment accuracy is also of primary importance, and for the high-sequence identity targets, side-chain accuracy and the accuracy of "loop" region main-chain were also considered. In fold recognition, the assessors experimented with a number of additional measures, but found that GDT_TS was a reasonable consensus. There, sequence alignment accuracy was also an important measure. The New Fold assessors found that GDT_TS was the best single measure, but GDT_TS rankings of accuracy were sometimes modified by visual inspection. The new fold assessor for CASP4 reached similar conclusions,[24] and it is clear that better measures are still needed in this category.

Assessment

All CASP experiments so far have placed the primary responsibility for assessing the significance of the results in the hands of independent assessors. The CASP5 assessors are Anna Tramontano assisted by Veronica Morea for comparative modeling; Nick Grishin assisted by Lisa Kinch for fold recognition; and Rob Russell assisted by Patrick Aloy for the new fold category. Articles by each of the assessment teams are included in the issue and constitute the most thorough and authoritative analysis available. As usual, the identities of the prediction teams were not known to assessors until they had completed an analysis and ranking of the results.

Statistical Significance of the Results

The primary goal of the CASP experiments is to assess the state of the art in protein structure prediction. In general, with a large number of prediction teams taking part, and an increased number of prediction targets, the results do provide a sound basis for drawing conclusions concerning the accuracy of models in particular prediction categories and for determining where significant progress has or has not been made. In addition, in general, there are enough data to indicate which prediction teams are producing the most accurate models in each category. However, there are some circumstances where reliable ranking of the performance of prediction teams is not possible. Although ranking is not an objective of CASP, understandably, predictors are very sensitive to any perceived misranking. As a consequence, arguments of ranking reliability continue to threaten the future of CASP. In CASP5, all three assessors have taken considerable care to evaluate the reliability of rankings and to address this issue frontally in their articles and in the choice of prediction groups invited to submit articles to the special issue. As in CASP4, there are few predictor articles in the comparative modeling category, because a number of groups perform roughly equally well, and there are no successful original approaches. It is clear that ranking distinctions in secondary structure prediction are very difficult to make reliably, and a decision was taken at the CASP5 users meeting to exclude that category in future CASPs.

In general, the performance of methods converges when there is a bottleneck—a group of competent predictors are all bumping up against a performance ceiling. In comparative modeling, there has been no significant progress in improving alignment quality or in refining initial models. These problems were identified way back in CASP2 and appear to be solvable. Yet, remarkably little effort in the field seems to be directed to the task. The issue with secondary structure prediction is somewhat different. It is likely that these methods are as close to accurate as they will ever be, given that secondary structure is partially determined by tertiary factors. Very small improvements continue to be made but probably only as a consequence of increased sequence database size. In these circumstances, it is not surprising that a number of methods have converged to approximately equivalent performance.

Meetings, Web Site, and Publications

Following the closing of the prediction season, two planning meetings involving the assessment teams and the organizers were held at the Sanger Institute; one before any assessment of the predictions and one when a full assessment was complete. The first of these meetings was also attended by several assessors from earlier CASPs, and the primary aim was to provide guidance to the CASP5 assessors. At the second meeting, the assessors presented the results of their work, including a full ranking of prediction teams, and these were extensively discussed. Only then were the names of the prediction teams made known to the assessors.

The meeting to discuss the outcome of the experiment was held at the customary place, in Asilomar, California, in December 2002. The assessors selected those prediction teams they considered had done the most significant work to talk at the meeting and also those invited to write prediction report articles. Both at the meeting and in the articles, participants have been 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.llnl.gov) provides extensive details of the targets, the predictions, and the numerical analyses. There are many possible views that may be taken of the results, and the interested reader is encouraged to consult other sources (e.g., Ref. 15) for alternative points of view.

Progress Over the CASPs

One of the main objectives of CASP is to measure progress in structure prediction. The assessors' articles address specific advancement areas, and the more general progress article adds further information. Between CASP1 and CASP2, there was a detectable improvement in prediction quality in many areas.[2] In retrospect, it seems a large component of that may have been the community adjusting to the nature of the experiment. Since then, the story has been different in each prediction category. In the new fold category, models in CASP1 were of very poor quality, but there was continuous progress from CASP1 through CASP4. In CASP5, there is little evidence of additional improvement, but hopefully this may be only a pause—there are many promising methods under development. In the fold recognition category, there was little evidence of progress after CASP1 through CASP4. However, considerable progress is apparent in this category for CASP5. The primary reason appears to be the emergence of meta-servers[25] as part of the Livebench experiment. A number of groups have developed meta-servers that send target sequences to other servers, collect the predictions, and generate some kind of consensus model from that input. In CASP5, fold recognition category results from the best meta-servers were competitive with the best humans, and some of the best human performances were obtained by starting from meta-server output. This is a very interesting development in the prediction field, but it is not without its problems. As has been noted,[26] recognition of success tends to go to the meta-servers, not to the groups who developed the contributing servers, likely making funding of new methods more difficult. In addition, this progress does not represent any improvement in the ability of any single method to produce a good model. Nevertheless, it is to be hoped that lessons from meta-servers will lead to an improvement in single methods. Comparative modeling, particularly based on relatively high-sequence identity, continues to be a stagnant area of modeling. There has been no detectable improvement in model quality since CASP2, and few methods seem very original. This is surprising because the problems of improving alignment at least to some degree and of improving the accuracy of a model over that obtained by simply copying a template do not seem that intractable. As more and more of structure space is explored experimentally, both by conventional structural biology, and more aggressively by structural genomics, more and more modeling will fall into this category.

Related Developments

FORCASP: Five CASP experiments have resulted in the development of a close knit prediction community. FORCASP (www.FORCASP.org) provides an on-line forum for the community. The primary aim is to provide a medium in which a broad range of CASP-related activities can be easily reported and discussed. The site provides a mechanism by which any registered user may publish an article. The articles are unrefereed, may be in any format, and are deposited directly by the authors. In this way, there are minimum barriers to the distribution of information. Instead of relying on referring, quality of articles is controlled by a commentary system. Any registered user may comment on a FORCASP article, in an unmoderated manner. In turn, commentary quality is monitored by a karma system, similar to that of Slashdot.org and other sites. Every user carries a karma, resulting from ranking of their contributions by other site users. The higher the karma rank, the more prominence a user's contributions receive.

FORCASP opened at the time of CASP5 meeting and has so been used for two purposes: some post-CASP discussions and to choose a predictor team considered to have used interesting methods in CASP5. This team was then invited to submit an article to the special issue on the basis of FORCASP voting. Future uses will include a journal club for the discussion of new articles in the prediction field, and most importantly in CASP6, as a means for any predictor who so wishes to publish their CASP work to the community.

Ten Most Wanted (TMW): TMW is an idea that arose during discussions at the CASP4 meeting. It is an organization to use the skill and knowledge of the CASP community to build useful structural models of biologically important proteins. There are many proteins for which it has so far not been possible to obtain structural information experimentally and where such information would be of immediate use in analyzing and guiding experiments. TMW is intended to be collaborative, with many predictors submitting models of a target protein, resulting in a consensus structure. An initial set of 10 target proteins (http://s2f.umbi.umd.edu/families.php?list_ac=12) were identified by appeals to the experimental community,[27] and members of the CASP community were invited to submit models of these, through the Livermore infrastructure. An initial analysis was reported at the CASP5 meeting, and it is clear that in a number of cases, useful models have already been produced. It is hoped that TMW will become an important spin-off from CASP.

CAPRI: CAPRI (Critical Assessment of Predicted Interactions) is the latest CASP-like experiment in a different prediction area. CAPRI focuses on the prediction of the detailed docking of pairs of protein molecules. A special issue of *Proteins*[28] has just been published on the first rounds of results, with participation of 20 groups, making predictions on seven targets. A great deal of insight into

the strengths and weaknesses of current methods has already been obtained.[29]

Future Developments

There will likely be a CASP6 experiment, running from the spring of 2004 and culminating in a meeting in December of that year. Those interested should check the CASP Web site for further announcements. It is expected that CAFASP, Livebench, and EVA will also play a role. The nature of the prediction field continues to change rapidly, with additional sources of prediction targets, increasing importance of automatic prediction servers, and a growing role of benchmarking experiments. All these factors will have a major influence on the future of CASP.

ACKNOWLEDGMENTS

As always, this CASP experiment depends on the cooperation, hard work, and support of a large number of people. We are grateful to the members of the experimental community who agreed to provide targets. In CASP5, structural genomics groups contributed more than half the targets, and we particularly appreciate their cooperation. Taking part required courage and hard work on the part of all the predictors. Without the careful and critical assessment of the results by the assessment teams, the experiment would have failed. We are very grateful to the organizers of CAFASP for their participation and allowing us to make full use of the data. We also thank the organizers of the large scale benchmarking experiments, Livebench and EVA, for their participation in the CASP meeting. We thank the editor of this journal, Ed Lattman, for once again providing a mechanism for peer-reviewed publication of the results. Thanks once again to Ceslovas Venclovas for his role in the conduct of the experiment. James Cluff provided essential computing support at the Sanger Institute computing resource, which was used to identify all PDB structural relatives for all the targets. We thank the agencies that have provided financial support. Lawrence Livermore Laboratory and the Department of Energy (OHER) have provided long-term funding for the Prediction Center at Livermore. The center is now also supported by a grant from the National Library of Medicine (LM07085). The NIH (GM/DK61967) provided generous meeting support. Thanks also to HP, Amgen, GlaxoSmithKine, and IBM for additional meeting support.

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. Pennisi E. Teams tackle protein prediction. Science 1996;273: 426–428.
6. Eisenberg D. Into the black of night. Nat Struct Biol 1997;4: 95–97.
7. Jones DT. Progress in protein structure prediction. Curr Opin Struct Biol 1997;7: 377–387.
8. Marchler-Bauer A, Bryant SH. A measure of success in fold recognition. Trends Biochem Sci 1997;22: 236–240.
9. Sternberg MJ, Bates PA, Kelley LA, MacCallum RM. Progress in protein structure prediction: assessment of CASP3. Curr Opin Struct Biol 1999;9:368–373.
10. Shortle D. Prediction of protein structure. Curr Biol 2000;10:R49–51.
11. Murzin AG. Progress in protein structure prediction. Nat Struct Biol 2000;8:110–112.
12. Al-Lazikani B, Jung J, Xiang Z, Honig B. Protein structure prediction. Curr Opin Chem Biol 2001;5:51–56.
13. Bonneau R, Baker D. Ab initio protein structure prediction: progress and prospects. Annu Rev Biophys Biomol Struct 2001;30: 173–189.
14. Tramontano A. Of men and machines. Nat Struct Biol 2003;10:87–90.
15. Bourne PE. CASP and CAFASP experiments and their findings. Methods Biochem Anal 2003;44:501–507.
16. Fischer D, Elofsson A, Rychlewski L. The 2000 olympic games of protein structure prediction; fully automated programs are being evaluated vis-a-vis human teams in the protein structure prediction experiment CAFASP2. Protein Eng 2000;13:667–670.
17. Eyrich VA, Marti-Renom MA, Przybylski D, Madhusudhan MS, Fiser A, Pazos F, Valencia A, Sali A, Rost B. EVA: continuous automatic evaluation of protein structure prediction servers. Bioinformatics 2001;17:1242–1243.
18. Dyson HJ, Wright PE. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. Adv Protein Chem 2002;62:311–340.
19. Venclovas C, Zemla A, Fidelis K, Moult J. Some measures of comparative performance in the three CASPs. Proteins 1999; Suppl 3:231–237.
20. Venclovas C, Zemla A, Fidelis K, Moult J. Comparison of performance in successive CASP experiments. Proteins 2001;Suppl 5:163–170.
21. Zemla A, Venclovas C, Moult J, Fidelis K. Processing and evaluation of predictions in CASP4. Proteins 2001;Suppl 5:13–21.
22. Sippl MJ, Lackner P, Domingues FS, Prlic A, Malik R, Andreeva A, Wiederstein M. Assessment of the CASP4 fold recognition category. Proteins 2001;Suppl 5:55–67.
23. Zemla A. LGA—a method for finding 3D similarities in protein structures. Nucleic Acids Res 2003;31:3370–3374.
24. Lesk A, Lo Conte L, Hubbard TJ. Assessment of novel fold targets in CASP4: predictions of three-dimensional structures, secondary structures, and inter-residue contacts. Proteins 2001;Suppl 5:98–118.
25. Bujnicki JM, Elofsson A, Fischer D, Rychlewski L. Structure prediction meta server. Bioinformatics 2001;17:750–751.
26. Valencia A. Meta, meta(n) and cyber servers. Bioinformatics 2003;19:795.
27. Abbott A. Computer modellers seek out “ten most wanted” proteins. Nature 2001;409:4.
28. Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJ, Vajda S, Vakser I, Wodak SJ. CAPRI: a critical assessment of predicted interactions. Proteins 2003;52:2–9.
29. Mendez R, Leplae R, De Maria L, Wodak SJ. Assessment of blind predictions of protein-protein interactions: current status of docking methods. Proteins 2003;52:51–67.