ELSEVIER

Servers for protein structure prediction

Daniel Fischer[1,2]

The 1990s cultivated a generation of protein structure human predictors. As a result of structural genomics and genome sequencing projects, and significant improvements in the performance of protein structure prediction methods, a generation of automated servers has evolved in the past few years. Servers for close and distant homology modeling are now routinely used by many biologists, and have already been applied to the experimental structure determination process itself, and to the interpretation and annotation of genome sequences. Because dozens of servers are currently available, it is hard for a biologist to know which server(s) to use; however, the state of the art of these methods is now assessed through the LiveBench and CAFASP experiments. Meta-servers — servers that use the results of other autonomous servers to produce a consensus prediction — have proven to be the best performers, and are already challenging all but a handful of expert human predictors. The difference in performance of the top ten autonomous (non-meta) servers is small and hard to assess using relatively small test sets. Recent experiments suggest that servers will soon free humans from most of the burden of protein structure prediction.

Addresses
[1] Buffalo Center of Excellence in Bioinformatics, and Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260, USA
[2] Department of Computer Science, Ben Gurion University, Beer-Sheva 84105, Israel

Corresponding author: Fischer, Daniel (df33@cse.buffalo.edu)

Current Opinion in Structural Biology 2006, **16**:178–182

This review comes from a themed issue on
Theory and simulation
Edited by Joel Janin and Michael Levitt

Available online 20th March 2006

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2006.03.004

Introduction

The availability of automated methods for protein structure prediction as computer servers has proliferated in the past dozen years. This is evident from a PubMed query using the terms ''protein structure prediction AND server''. This query returns 175 articles, of which half were published between 1993 and 2004, and the other half in the past two years alone. This proliferation has motivated *Nucleic Acids Research* to publish a dedicated web server issue every year, which covers servers aimed at the many aspects of computational structure prediction [1•]. These include, among others, servers for secondary structure prediction, contacts prediction, docking, close homology modeling and distant homology modeling. Here, I review recent work addressing the latter two aspects of computational structure prediction, focusing on selected available and operational servers that can be accessed without restrictions or fees by any user, and that have been evaluated by the LiveBench [2,3••] or CAFASP (Critical Assessment of Fully Automated Structure Prediction) [4,5,6••] experiments. A more extensive overview of recently published servers, including other aspects of structure prediction, is given in [1•] and three recent independent reviews are [7•,8,9].

Close and distant homology modeling servers accept a protein's amino acid sequence as input, search known 3D structures for appropriate template(s) and automatically generate a 3D model containing the coordinates of the atoms of the protein. When sequence similarity between the target sequence and a protein of known structure is significant, this process is referred to as (close) homology modeling. In homology modeling, relatively simple sequence comparison methods are applied (e.g. BLAST or PSI-BLAST [10]) in order to find a template, and to generate the alignment between target and template. When there exists no known structure with significant sequence similarity to the target (e.g. PSI-BLAST fails to find a significant hit or finds a hit that aligns only a fraction of the query sequence or contains many gaps), more sophisticated methods are needed to find appropriate templates (if any) and to generate more accurate alignments. This process is referred to as distant homology modeling, fold recognition or threading, but the end result is also a 3D model of the target protein. Methods that do not directly use known templates are referred to as *ab initio*, but they are not yet widely available in the form of servers.

Recent proliferation and value of structure prediction servers

The proliferation of structure prediction servers in the past few years may be due to the fact that, recently, the performance of automated methods has improved significantly and has already produced valuable predictions, many of which have been verified experimentally. Thus, demand from the biology community for such servers is increasing. Although the goal of generating models that match (or improve) experimental accuracy has not yet been achieved, an increasing number of cases have been reported whereby automatically generated models have helped solve and improve the experimental ones. Thus, automated methods are already playing an increasingly

important role as complements of both the various experimental structural genomics projects [11] and the experimental structure determination process itself [12–14]. As methods continue to improve, some researchers believe today that it is no longer unrealistic to expect that the accuracy of automated (close homology) models will soon begin to rival that of low to medium resolution experimental models.

Another factor in the recent proliferation of servers is the 'open source' approach of modern bioinformaticians, who aim to meet the demand from the community of users by providing a 'service' (hence the name 'server'). This frees a biologist from the burden of implementing and/or maintaining complicated and resource-demanding software, who, in most cases, wishes to have a fully automated, easy-to-use internet service. In addition, and not negligibly, developers wish to make their servers available, because this provides them with one of their utmost sources of scientific satisfaction: their methods are being used by 'real' experimentalists.

Possibly one of the most important factors spurring the recent proliferation of structure prediction servers is the availability of the complete genome sequences of hundreds of organisms. In addition to being able to automatically generate homology models for a fraction of the proteins encoded in these genomes, structure prediction servers aimed at distant homology detection have found a new application: establishing distant evolutionary relationships when standard methods, such as PSI-BLAST, fail. In this application, users are interested in finding possible distant relationships and are not always interested in the (relatively inexact) 3D coordinates of each of the atoms. The complete genome sequences have revealed that the fraction of open reading frames (ORFs) lacking significant sequence similarity to other proteins is surprisingly high and, thus, a large fraction of the new ORFs remain of unknown function; these are usually annotated as 'hypotheticals' [15]. Thus, better methods for distant homology detection are increasingly playing a critical role in the interpretation and functional annotation of genome sequences [16,17].

The LiveBench and CAFASP experiments

This proliferation of servers has become a curse as well as a blessing. A biologist today frequently asks: which server should I use; as a server always returns an answer, how do I know if I can trust the result; what does the server output mean; how fast is it; should I use more than one server? To attempt to address these questions, Rychlewski and Fischer have created the LiveBench (LB) experiment. LB continuously assesses the capabilities of automated servers using a relatively large number of prediction targets, compiled every week from newly released protein structures, and provides an evaluation of the servers' capabilities approximately every half year. The CASP (Critical Assessment of Structure Prediction) and CAFASP experiments, held every two years, use a significantly smaller number of prediction targets; EVA [18] is another evaluation project, which focuses on other aspects of structure prediction.

The main findings of the latest LB experiment (LB-8; see http://bioinfo.pl/meta/livebench.pl for details) confirmed what previous evaluations have indicated, namely that the so-called meta-servers outperform all the individual autonomous servers. We distinguish meta-servers from autonomous servers by the type of input required: a meta-server cannot run independently, explicitly requiring as input the predictions of at least one other participating server [19•]. Meta-servers attempt to automate the process that many expert human predictors have successfully applied: instead of relying on a single structure prediction method, they utilize diverse sources of information, including the top models predicted by a number of servers. This has been a useful approach, because often a correct prediction can be obtained by one server but not by the others. Furthermore, for the hardest prediction targets, a server often generates a correct prediction among its top results, but this prediction either has a below-threshold confidence score or is not the top-ranked prediction. Given the success of this 'meta-prediction' approach, a number of meta-servers have been developed, including higher order 'meta-meta-servers', which use as input information from other meta-servers.

In LB-8, three series of reliable, highly accessible meta-servers were assessed: the PCONS/PMOD series [20], the 3D-SHOTGUN series [21,22] and the newer 3D-JURY series [23]. After the meta-servers, the difference in performance of the 5–10 best autonomous servers is small, and thus any ranking is highly dependent on the particular evaluation method and test set used. This is even more pronounced for smaller evaluation test sets, such as that used in CAFASP. For example, removal of one single target from the CAFASP test set results in significant changes in the rankings [6••]; application of slightly different evaluation methods and scoring systems also results in significant ranking changes (e.g. see the ''Preliminary server standings'' at the CAFASP web site at http://fischerlab.cse.buffalo.edu/CAFASP/ and two alternative independent evaluations at http://www.forcasp.org and in [20]). Thus, the following list should be considered as the 'pack' of top-performing autonomous servers, without placing significance on the order.

Among the top-performing autonomous servers is the recently developed series of Meta-BASIC servers [24], which are variants of a number of profile comparison methods. The top-ranked servers in LB-8 include those that have also ranked among the top performers in previous experiments: the autonomous SHOTGUN version [21], the ORFeus series [25] and FFAS03 [26]. Following

are older servers that had ranked among the top performers in previous experiments: 3D-PSSM [27], INBGU [28] (and its newer version Inub at http://inub.cse.buffalo.edu), MGENTHREADER [28], FUGUE [29] and SP3 [30••]. Notice that, although the difference in performance of the above servers is not large, they all performed significantly better than PSI-BLAST.

In CAFASP4, despite the significantly smaller set of targets, the set of top-performing servers was similar to that found in LB-8. Comparison of the CAFASP results with those generated by expert human predictors has indicated that the best-performing (meta) servers are already outperforming all but a handful of human predictors. Among the set of top-performing servers, we observe some differences in the LB and CAFASP rankings; the newer version of SP3 performed much better in CAFASP, and Meta-BASIC and ORFeus slightly worse. In addition, the meta-servers Robetta [31••], and ACE and RAPTOR [32], which did not participate in LB-8, ranked in the top ten in CAFASP4. Other recently published reports found using the above-mentioned PubMed query include the following servers that participated in CAFASP, but did not rank at the very top: HHpred [33], Arby [34], PROSPECT-PSPP [35], Wurst [36] and the meta-server PROTINFO [37]. Table 1 lists some of the best publicly available servers identified at CAFASP4 and LB-8 (other top-performing CAFASP servers were not available at the time of writing). We refer the reader to the LB and CAFASP web pages for detailed tables, which include the results of all participating servers, separate sensitivity and specificity analyses, division into 'easy' and 'hard' targets, and the option for the user to generate alternative rankings using different criteria (target subsets, evaluation methods, separate sensitivity and specificity analyses, server subsets, etc.). The specificity analyses are probably among the most valuable results of these experiments: they help users both understand and interpret a server's output, and determine when a server's result is reliable. Notice that CASP6 carried out a parallel, but limited, server evaluation, which also identified Robetta and SP3 among the top performers (see the upcoming reports in the CASP special issue of *Proteins: Structure, Function and Bioinformatics*). Nevertheless, the CASP server evaluation is not very useful, because some of the top-performing servers, including the best meta-servers, participated in CAFASP only. In addition, the CASP evaluations have not included specificity analyses, which are necessary for determining when a server's result is reliable.

Current bottlenecks and future prospects for structure prediction servers

Current structure prediction servers perform best when the target sequence is composed of a single structural domain. However, many proteins, especially eukaryotic, are often composed of more than one domain. Thus, for such cases, manual division of the target sequence followed by separate submission to the servers is required. This, of course, does not correspond to a fully automatic approach, and requires the human to know or guess where the domain boundaries are. This may be one of the reasons why a few expert predictors today still outperform servers; indeed, one of the main factors leading to human success in CASP has been the correct pre-identification of domains in the query sequence. Thus, the ability to automatically identify domains is likely to result in significantly better performance and has already been introduced in at least three servers. As a step towards improving domain prediction, the CAFASP4 experiment

Table 1

A partial list of currently available best-performing servers in CAFASP4.

3D-JURY [23] http://bioinfo.pl/meta/	A widely used interactive meta-server that allows the user to select both the set of servers (and meta-servers) and the selection procedure to be used for consensus building. It automatically compiles the results from the selected external and internal servers. Although 3D-JURY utilized a large number of participating servers in CAFASP and LB, currently only a dozen servers are integrated.
3D-SHOTGUN [21] http://inub.cse.buffalo.edu	A new, local meta-server named SHUB, which does not depend on external services; it locally generates results from three fold recognition methods (INUB, SP3 and PROSPECTOR) and computes a hybrid model by combining partial structures using the 3D-SHOTGUN algorithm. It includes a refinement module that generates full-atom models with correct geometry.
Pcons/Pmodel [20] http://www.sbc.su.se/~bjorn/Pcons5	A new improved version of Pcons, integrating a consensus analysis similar to that of 3D-JURY and a structural analysis using the ProQ MQAP. Currently available only as a downloadable standalone program accepting as input the results of various methods, which need to be generated separately.
PROTINFO [37] http://protinfo.compbio.washington.edu	A server for comparative and *ab initio* modeling. The results of other fold recognition servers can be used as input (e.g. 3D-JURY is suggested), in which case a refinement procedure is applied. The selection procedure is a consensus of 15 scoring functions (e.g. MQAPs).
SP3 [30••] http://theory.med.buffalo.edu	A local, autonomous server using sequence profiles from structural fragments. Available also for downloading.
Robetta [31••] http://robetta.bakerlab.org	A meta-meta-server for both *ab initio* and comparative modeling using the ROSETTA fragment insertion method. For parent detection, it uses FFAS or 3D-JURY.

introduced a new category: assessment of domain prediction (DP) servers. The results of the DP assessment show that, despite being a very old problem and the availability of a dozen or so DP servers, performance is not very good in the lack of clear homology to known domains. Another finding was that, similar to what has been observed for structure prediction servers, DP meta-servers tend to be more robust and reliable [38]. For further details on CAFASP-DP, see the published results on the CAFASP web site, reference [38] and the upcoming CASP6 reports (see Update), which, following our CAFASP initiative, also included DP server evaluation.

Another open problem in automated structure prediction is the selection of alternative models using pseudo-energy potentials or empirical quality scores. This is critical for the refinement of homology models, where the aim is to obtain models that are closer to the native structure than to the template(s) used in their construction. Identification of 'better' models is also critical for the most difficult targets. To gain insights into the ability of model quality assessment programs (MQAPs), the CAFASP4 experiment also introduced the CAFASP-MQAP category. Surprisingly, the results show that several MQAPs are very successful in selecting the better models and their performance as 'selectors' is comparable to that of the best servers (see the published results on the CAFASP web site for details). One of the best MQAPs was Verify3D [39]. This is surprising because Verify3D is one of the earliest methods, developed over a dozen years ago, and is extraordinarily simple: it uses 18 discrete environmental classes for each amino acid (not surprisingly, many top-performing human predictors make extensive use of it when selecting models from the various servers). Again, as for the other types of servers, a 'meta-MQAP', developed in my laboratory and named MQAP-CONSENSUS, performed at the very top; this method (under the name MCon) was also one of the top ten predictors in the *ab initio* section of CASP6. Reports about two of the best-performing MQAPs in CAFASP have just been published: Victor/FRST [40] and MODCHECK [41].

Further incremental improvements in automated methods, including the incorporation into servers of automatic domain detection methods and MQAPs, together with the growth of sequence and structure databases, will result in significantly better server performance. As of today, only a handful of expert human predictors outperform the best of the (meta) servers, at the expense of significant human effort. As automated structure prediction servers improve, they will play an increasingly important role in the experimental structure determination process itself and in the annotation of genome sequences. Soon, as has happened, for example, in secondary structure prediction, humans will stop attempting to manually improve the servers' results. Instead, they will use the automatically generated models to study other more challenging and fundamental problems of modern biology, where automation has not yet been developed to the same extent [9,42], including automated protein–protein interaction prediction [43], protein design or automatic function prediction in the lack of homology.

Update

The CASP6 reports have now been published in *Proteins: Structure, Function and Bioinformatics* [44].

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • *Nucleic Acids Res*: **Web Server Issue.** *Nucleic Acids Res* 2005, **33**:W1-W786.
The yearly compilation of available web servers, including various protein structure prediction servers.

2. Bujnicki JM, Elofsson A, Fischer D, Rychlewski L: **LiveBench-1: continuous benchmarking of protein structure prediction servers**. *Protein Sci* 2001, **10**:352-361.

3. •• Rychlewski L, Fischer D: **LiveBench-8: the large-scale, continuous assessment of automated protein structure prediction**. *Protein Sci* 2005, **14**:240-245.
This paper reports the results of the latest LB experiment. 44 servers and meta-servers were evaluated using a test set of 172 newly released protein structures. The results corroborated the superior performance of meta-servers and identified a handful of top-performing, newly developed, autonomous (non-meta) servers.

4. Fischer D, Barret C, Bryson K, Elofsson A, Godzik A, Jones D, Karplus KJ, Kelley LA, MacCallum RM, Pawowski K *et al.*: **CAFASP-1: critical assessment of fully automated structure prediction methods**. *Proteins* **(suppl 3):1999:209-217**.

5. Fischer D, Rychlewski L, Dunbrack RL Jr, Ortiz AR, Elofsson A: **CAFASP3: the third critical assessment of fully automated structure prediction methods**. *Proteins* 2003, **53(suppl 6)**: 503-516.

6. •• Fischer D, Rychlewski L: **CAFASP4.** Available on the World Wide Web at: http://fischerlab.cse.buffalo.edu/CAFASP/.
An extensive web site describing the latest CAFASP experiment. Over 90 servers and meta-servers participated in the various categories: structure prediction, domain prediction (CAFASP-DP) and quality assessment programs (CAFASP-MQAP). The site includes all data and programs, and an interactive evaluation facility that allows the user to analyze subsets of the data using different evaluation methods. The structure prediction section of CAFASP corroborated the findings of previous experiments — meta-servers are the best performers, already outperforming all but a handful of expert human predictors.

7. • Ginalski K, Grishin NV, Godzik A, Rychlewski L: **Practical lessons from protein structure prediction**. *Nucleic Acids Res* 2005, **33**:1874-1891.
An excellent, timely overview of current practical approaches, including sections on methods, evaluation analysis, lessons learnt and the usefulness of protein structure prediction.

8. Wan XF, Xu D: **Computational methods for remote homolog identification**. *Curr Protein Pept Sci* 2005, **6**:527-546.

9. Petrey D, Honig B: **Protein structure prediction: inroads to biology**. *Mol Cell* 2005, **20**:811-819.

10. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25**:3389-3402.

11. Fischer D, Baker D, Moult J: **We need both computer models and experiments**. *Nature* 2001, **409**:558.

12. Bujnicki J, Rychlewski L, Fischer D: **Fold-recognition detects an error in the Protein Data Bank**. *Bioinformatics* 2002, **18**:1391-1395.

13. Claude JB, Suhre K, Notredame C, Claverie JM, Abergel C: **CaspR: a web server for automated molecular replacement using homology modelling**. *Nucleic Acids Res* 2004, **32**:W606-W609.

14. Bujnicki JM, Feder M, Rychlewski L, Fischer D: **Errors in the *D. radiodurans* large ribosomal subunit structure detected by protein fold-recognition and structure validation tools**. *FEBS Lett* 2002, **525**:174-175.

15. Doolittle RF: **Evolutionary aspects of whole-genome biology**. *Curr Opin Struct Biol* 2005, **15**:248-253.

16. Siew N, Fischer D: **Twenty thousand ORFan microbial protein families for the biologist?** *Structure* 2003, **11**:7-9.

17. Siew N, Saini HK, Fischer D: **A putative novel alpha/beta hydrolase ORFan family in Bacillus**. *FEBS Lett* 2005, **579**:3175-3182.

18. Koh IY, Eyrich VA, Marti-Renom MA, Przybylski D, Madhusudhan MS, Eswar N, Grana O, Pazos F, Valencia A, Sali A *et al.*: **EVA: Evaluation of protein structure prediction servers**. *Nucleic Acids Res* 2003, **31**:3311-3315.

19. • Bujnicki JM, Fischer D: **'Meta' approaches to protein structure prediction**. In *Nucleic Acids and Molecular Biology series: Practical Bioinformatics.*. Edited by Bujnicki JM. Springer-Verlag; 2004:23-34.
A review of the birth and development of meta-predictors in protein structure prediction.

20. Wallner B, Elofsson A: **Pcons5: combining consensus, structural evaluation and fold recognition scores**. *Bioinformatics* 2005, **21**:4248-4254.

21. Fischer D: **3D-SHOTGUN: a novel, cooperative, fold-recognition meta-predictor**. *Proteins* 2003, **51**:434-441.

22. Fischer D: **3DS3 and 3DS5 3D-SHOTGUN meta-predictors in CAFASP3**. *Proteins* 2003, **53(suppl 6)**:517-523.

23. Ginalski K, Elofsson A, Fischer D, Rychlewski L: **3D-Jury: a simple approach to improve protein structure predictions**. *Bioinformatics* 2003, **19**:1015-1018.

24. Ginalski K, von Grotthuss M, Grishin NV, Rychlewski L: **Detecting distant homology with Meta-BASIC**. *Nucleic Acids Res* 2004, **32**:W576-W581.

25. Ginalski K, Pas J, Wyrwicz LS, von Grotthuss M, Bujnicki JM, Rychlewski L: **ORFeus: detection of distant homology using sequence profiles and predicted secondary structure**. *Nucleic Acids Res* 2003, **31**:3804-3807.

26. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A: **FFAS03: a server for profile–profile sequence alignments**. *Nucleic Acids Res* 2005, **33**:W284-W288.

27. Kelley LA, MacCallum RM, Sternberg MJ: **Enhanced genome annotation using structural profiles in the program 3D-PSSM**. *J Mol Biol* 2000, **299**:499-520.

28. Fischer D: **Hybrid fold recognition: combining sequence derived properties with evolutionary information**. *Pac Symp Biocomput* 2000:119-130.

29. Shi J, Blundell TL, Mizuguchi K: **FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties**. *J Mol Biol* 2001, **310**:243-257.

30. •• Zhou H, Zhou Y: **Fold recognition by combining sequence profiles derived from evolution and from depth-dependent structural alignment of fragments**. *Proteins* 2005, **58**:321-328.
See annotation to [31••].

31. •• Kim DE, Chivian D, Baker D: **Protein structure prediction and analysis using the Robetta server**. *Nucleic Acids Res* 2004, **32**:W526-W531.
Two servers [30••,31••] were among the best performers in recent evaluation experiments. Other top performers discussed in the text and listed in Table 1 are not annotated in the reference list because they were published before 2004.

32. Xu J, Li M, Kim D, Xu Y: **RAPTOR: optimal protein threading by linear programming**. *J Bioinform Comput Biol* 2003, **1**:95-117.

33. Soding J, Biegert A, Lupas AN: **The HHpred interactive server for protein homology detection and structure prediction**. *Nucleic Acids Res* 2005, **33**:W244-W248.

34. von Ohsen N, Sommer I, Zimmer R, Lengauer T: **Arby: automatic protein structure prediction using profile-profile alignment and confidence measures**. *Bioinformatics* 2004, **20**:2228-2235.

35. Guo JT, Ellrott K, Chung WJ, Xu D, Passovets S, Xu Y: **PROSPECT-PSPP: an automatic computational pipeline for protein structure prediction**. *Nucleic Acids Res* 2004, **32**:W522-W525.

36. Torda AE, Procter JB, Huber T: **Wurst: a protein threading server with a structural scoring function, sequence profiles and optimized substitution matrices**. *Nucleic Acids Res* 2004, **32**:W532-W535.

37. Hung LH, Ngan SC, Liu T, Samudrala R: **PROTINFO: new algorithms for enhanced protein structure predictions**. *Nucleic Acids Res* 2005, **33**:W77-W80.

38. Saini HK, Fischer D: **Meta-DP: domain prediction meta-server**. *Bioinformatics* 2005, **21**:2917-2920.

39. Eisenberg D, Luthy R, Bowie JU: **VERIFY3D: assessment of protein models with three-dimensional profiles**. *Methods Enzymol* 1997, **277**:396-404.

40. Tosatto SC: **The Victor/FRST function for model quality estimation**. *J Comput Biol* 2005, **12**:1316-1327.

41. Pettitt CS, McGuffin LJ, Jones DT: **Improving sequence-based fold recognition by using 3D model quality assessment**. *Bioinformatics* 2005, **21**:3509-3515.

42. Siew N, Fischer D: **Convergent evolution of protein structure prediction and computer chess tournaments: CASP, Kasparov and CAFASP**. *IBM Systems Journal* 2001, **40**:410-425.

43. Wodak SJ, Mendez R: **Prediction of protein-protein interactions: the CAPRI experiment, its evaluation and implications**. *Curr Opin Struct Biol* 2004, **14**:242-249.

44. Lattman EE (Ed): **Sixth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction**. *Proteins* 2005, **61(suppl S7)**:1-236.