Nucleic Acids Research, 2003, Vol. 31, No. 13 ***3311–3315***
DOI: 10.1093/nar/gkg619

EVA: evaluation of protein structure prediction servers

Ingrid Y. Y. Koh[1,*], Volker A. Eyrich[2], Marc A. Marti-Renom[3], Dariusz Przybylski[2,4], Mallur S. Madhusudhan[3], Narayanan Eswar[3], Osvaldo Graña[5], Florencio Pazos[5], Alfonso Valencia[5], Andrej Sali[3] and Burkhard Rost[1,2,6]

[1]Columbia University Center for Computational Biology and Bioinformatics (C2B2), Russ Berrie Pavilion, 1150 St Nicholas Avenue, [2]CUBIC, Department of Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street BB217, New York, NY 10032, USA, [3]Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, and California Institute for Quantitative Biomedical Research, University of California, San Francisco, CA 94143, USA, [4]Department of Physics, Columbia University, 538 West 120th Street, New York, NY 10027, USA, [5]Protein Design Group, Centro Nacional de Biotecnologia (CNB-CSIC), Cantoblanco, Madrid 28049, Spain and [6]North East Structural Genomics Consortium (NESG), Department of Biochemistry and Molecular Biophysics, Columbia University, 650 West 168th Street BB217, New York, NY 10032, USA

Received February 15, 2003; Revised and Accepted April 14, 2003

ABSTRACT

EVA (http://cubic.bioc.columbia.edu/eva/) is a web server for evaluation of the accuracy of automated protein structure prediction methods. The evaluation is updated automatically each week, to cope with the large number of existing prediction servers and the constant changes in the prediction methods. EVA currently assesses servers for secondary structure prediction, contact prediction, comparative protein structure modelling and threading/fold recognition. Every day, sequences of newly available protein structures in the Protein Data Bank (PDB) are sent to the servers and their predictions are collected. The predictions are then compared to the experimental structures once a week; the results are published on the EVA web pages. Over time, EVA has accumulated prediction results for a large number of proteins, ranging from hundreds to thousands, depending on the prediction method. This large sample assures that methods are compared reliably. As a result, EVA provides useful information to developers as well as users of prediction methods.

INTRODUCTION

Continuous, automated, large data sets, statistical significance. The goal of EVA is to evaluate the sustained performance of protein structure prediction servers through a battery of objective measures for prediction accuracy. While the bi-annual CASP (Critical Assessment of Techniques for Protein Structure Prediction) meetings address the question 'how well can experts predict protein structures with the help of machines?', the question addressed by EVA is 'how well can automatic servers predict protein structures?'. Conceptually, this is similar to CAFASP (Critical Assessment of Fully Automated Structure Prediction), but there is a major difference: EVA provides a continuous, fully automatic and statistically more significant analysis of structure prediction servers, whereas CAFASP only covers a limited number of proteins determined in a period of about 4 months in every 2 years: fewer than 10 proteins were available for the non-homology category at CAFASP3 in 2002. This implies that it is—at best—extremely difficult to infer differences of statistical significance from the CAFASP/CASP data sets. For example, the assessor for secondary structure prediction in 2002 concluded that there was no improvement in secondary structure predictions with respect to the CAFASP/CASP in 2000 although the numerical values differed by over six percentage points.

A tool for developers of prediction methods. EVA facilitates developers of structure prediction methods to improve their approaches and users of prediction servers to apply methods judiciously. The ranking of each prediction method is analysed and updated on the web every week. Ranking is a non-trivial task because of the non-uniformity in data sets and in the measures for accuracy. Another complication is that methods are compared most reliably when they are tested under identical conditions, i.e. with identical sets of proteins (1–3). Here, we sketch the EVA mechanisms that enable

*To whom correspondence should be addressed. Tel: +1 2123054018; Fax: +1 2123057932; Email: koh@cubic.bioc.columbia.edu

such large-scale assessment of prediction servers automatically and continuously.

DESIGN AND IMPLEMENTATION

Five steps from sequence to results

The analysis of prediction methods involves the following steps (Fig. 1): (i) select a set of suitable test sequences; (ii) apply prediction methods to those sequences; (iii) assess prediction methods by measuring prediction quality using certain scoring functions; (iv) determine criteria for statistically significant differences, and rank the methods accordingly; (v) merge results of the current week with those accumulated in the past, publish results on the web, and communicate with and gather results from the EVA satellites [at Centro Nacional de Biotecnologia (CNB) in Madrid, Spain and at University of California, San Francisco (UCSF)].

(i) Selection of test sequences. Every day, EVA downloads the sequences for the newest experimentally determined protein structures from the Protein Data Bank (PDB) (4) web site. Sequences are dissected into protein chains that constitute the basic units for EVA. Very short sequences (<30 residues) and proteins containing a significant number of unresolved residues are excluded. The remaining sequences are sent by META-PredictProtein (5,6) (META-PP) to prediction servers that consented to the evaluation by EVA. Threading/fold recognition servers constitute an exception to this 'send-all' rule: in order to reduce the load on these servers, we submit only sequences without clearly homologous structures (i.e. novel proteins) (7,8). These novel sequences have no hits in the previous version of the PDB below a PSI-BLAST (9) E-value of 10^{-3} and/or an HSSP-distance <0 (8). Over the last 3 years, this filtering step reduced the number of chains to about 8%; threading servers therefore have to handle <10 submissions from EVA per week. While secondary structure prediction methods handle all proteins, currently EVA publishes results only for the subset of the novel proteins every week. For contact predictions, proteins with homologous structures are considered separately from proteins without structurally defined homologues. Obviously, most results analysed in the comparative modelling category (EVA-CM) are based on proteins that are not novel. However, EVA-CM currently does not apply any particular threshold: all models are evaluated.

(ii) Collection of predictions. Once a day, META-PP (5,6) submits sequences to prediction servers and collects the results. Once a week, these results are sent to EVA satellites for evaluation, namely to Columbia University for secondary structure prediction and fold recognition/threading, to UCSF for comparative modelling and to CNB for inter-residue distances/contacts.

(iii) Evaluation of sustained performance. Prediction quality is evaluated using a battery of scoring functions sketched below for all four categories.

(iv) Ranking prediction methods. Ranking is most reliable when prediction methods are tested under identical circumstances. The best way to rank two methods is to assess their performance based on the identical test sets. Two ranking methods are currently available in EVA. The first one is based on sub-sets of all proteins that are common to all methods. The limitations of this approach are that: (i) not all methods exist at the same time; and (ii) not all sequences are predicted by all methods at any given time due to server downtime and errors. In practice, these two effects reduce the size of the common sub-sets dramatically. The second ranking approach relies on pairwise method comparisons that depend on the sub-set of proteins common to the two compared methods (3). This pairwise ranking approach determines for each pair of participating servers whether or not it is possible to discriminate their accuracies, given the size of the test set and the particular accuracy measure used. The downside of this approach is that the overall ranking list obtained by averaging the pairwise results may be 'frustrated' due to the different testing sets for the different pairs of methods.

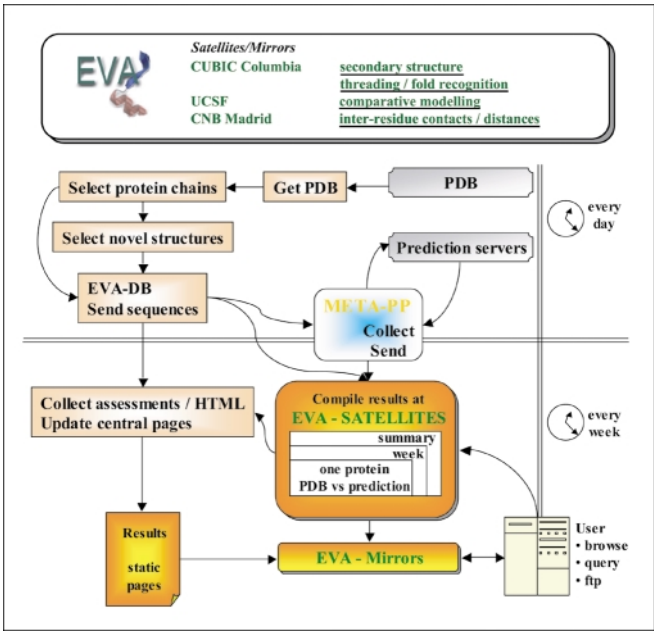

Figure 1. Flowchart of EVA. Every day, EVA downloads the newest protein structures from PDB (4). The structures are added to mySQL databases, sequences are extracted for every protein chain and are sent to each prediction server by META-PredictProtein (5) (except for threading in which only novel structures are sent). META-PP collects the results and sends them to EVA. Every week, predictions of secondary structure, threading/fold recognition, comparative modelling and inter-residue contacts are evaluated at the EVA satellites at Columbia University, University of California, San Francisco, and CNB Madrid. The central EVA site at Columbia collects all the assessments from the satellites and the results from the database searches, and publishes the updated web pages. Finally, all web pages are mirrored at the satellites.

(v) Results presented on the EVA web sites. The central EVA site at Columbia University collects either the assessments or the html pages with assessments from the satellites every week and presents them on the web. The central EVA site is mirrored at all EVA satellites (Fig. 1).

CATEGORIES OF EVALUATION

EVA currently addresses the following protein structure prediction categories (Table 1): comparative modelling

Table 1. Prediction methods evaluated by EVA

Method	URL	Main developer(s)	References
Comparative modeling			
3D-Jigsaw	http://www.bmm.icnet.uk/servers/3djigsaw/	PA Bates, P Fitzjohn and BC Moreira	(26,27)
CPHModels	http://www.cbs.dtu.dk/services/CPHmodels/	S Brunak *et al.*	(28)
ESyPred3D	http://www.fundp.ac.be/urbm/bioinfo/esypred/	C Lambert	(29)
SDSC1	http://cl.sdsc.edu/hm.html	IN Shindyalov and PE Bourne	(24)
SwissModel	http://www.expasy.org/swissmod/	T Schwede, MC Peitsch and N Guex	(30)
Threading/fold recognition			
3D-PSSM	http://www.sbg.bio.ic.ac.uk/~3dpssm	L Kelley, B Maccallum and M Sternberg	(31)
BLAST	http://www.ncbi.nlm.nih.gov/BLAST	S Karlin and S Altschul	(32)
FUGUE	http://www-cryst.bioc.cam.ac.uk/~fugue/	K Mizuguchi	—
Libellula	http://www.pdg.cnb.uam.es:8081/libellula.html	—	—
Prospect	http://www.aber.ac.uk/~phiwww/prof/	Y Xu	(33)
PSI-BLAST	http://www.ncbi.nlm.nih.gov/BLAST/	S Altschul *et al*	(9)
SAMt99	http://www.cse.ucsc.edu/research/compbio/HMM-apps/model-library-search.html	K Karplus, C Barrett and R Hughey	(34,35)
Superfamily	http://supfam.mrc-lmb.cam.ac.uk/SUPERFAMILY/	J Gough	(36)
Inter-residue contacts			
CORNET	http://prion.biocomp.unibo.it/cornet.html	P Fariselli, O Olmea and A Valencia, R Casadio	(10,37,38)
PDGCON	http://www.pdg.cnb.uam.es:8081/	F Pazos, O Olmea and A Valencia	(13)
CONcons/CONhydro	http://www.pdg.cnb.uam.es:8081/	F Pazos, O Olmea and A Valencia	(37–39)
Secondary structure			
APSSP2	http://www.imtech.res.in/raghava/apssp2/	G Raghava	(40)
Jpred	http://jura.ebi.ac.uk:8888/	JA Cuff and GJ Barton	(41)
PHDsec	http://cubic.bioc.columbia.edu/predictprotein	B Rost and C Sander	(42)
PHDpsi	http://cubic.bioc.columbia.edu/predictprotein	D Przybylski and B Rost	(43)
PROF_king	http://www.aber.ac.uk/~phiwww/prof/	M Ouali and R King	(44)
PROFsec	http://cubic.bioc.columbia.edu/predictprotein	B Rost	(45)
PSIpred	http://insulin.brunel.ac.uk/psiform.html	D Jones	(46,47)
SAM-T99sec	http://www.cse.ucsc.edu/research/compbio/HMM-apps/T99-query.html	K Karplus, C Barrett and R Hughey	(34,48)
SSpro2	http://promoter.ics.uci.edu/BRNN-PRED/	G Pollastri and P Baldi	(49)

(EVA-CM); inter-residue contact prediction (EVA-con); secondary structure prediction (EVA-sec); and threading (EVA-FR). In the following, we sketch the measures for accuracy employed for each category. Note that the detailed definitions of the scores are available through the EVA web sites.

EVA-CM. Implements a small number of criteria—arranged hierarchically from coarser to finer—that measure the accuracy of a comparative model. The assessed aspects of a model include fold type, alignment, whole structure, core structure, loops and side-chains. Final ranking is reported using the 'pairwise' comparison of prediction servers (3). From May 2000 to January 2003, predictions were collected from five different servers, resulting in 20 957 submitted models for 9050 different PDB chains. On average, 2.3 models were predicted per chain.

EVA-con. Evaluates inter-residue contact/distance predictions. A number of servers predict contacts directly, using neural networks of different kinds trained on contact maps (10,11). There are also predictions of contacts based on assembled structures (12). The current evaluation criteria implemented in EVA-con include: (i) accuracy—the number of the correctly predicted contacts divided by the total number of predicted contacts (13); (ii) improvement over random—the calculated accuracy divided by the random accuracy (13); (iii) distance distribution of the predicted contacts—the weighted harmonic average difference between the predicted contact distance distribution and the all-pairs distance distribution (14); and (iv) delta evaluation—the percentage of correctly predicted contacts that are within a certain number (delta) of residues of the experimental contact, measured along the sequence (15). EVA-con may also be used to evaluate *ab initio*, fold recognition and comparative modelling servers by transforming models into intra-molecular contacts between the corresponding C-beta atoms (C-alpha for Gly) with a 8 Å cut-off.

EVA-sec. Evaluates protein secondary structure predictions. Secondary structures are assigned from 3D structures through DSSP (16) and STRIDE (17). EVA-sec measures accuracy by: (i) per-residue accuracy (18) (Q_3)—percentage of residues correctly predicted in one of the three states (helix, strand or other); (ii) per-segment accuracy (18,19) (*SOV*)—average overlap between segments (methods that get most of the segment cores right are generally more useful than those that get some of the entire segments right); and (iii) accuracy of predicting structural class—percentage of proteins correctly predicted in one of the following classes: all-alpha, all-beta, alpha/beta and others (20,21). Rankings are presented using both the 'common subset' and 'pairwise' comparison approaches.

EVA-FR. Currently evaluates models only for novel sequences (i.e. proteins for which PSI-BLAST searches do not reveal similarity to a known structure). Since there is no

single measure that can comprehensively assess the quality of threading models, EVA-FR employs an array of alignment dependent and alignment independent measures (22–24). For most of the measures used, two aspects of server performance are considered: (i) the ability to produce good models for each target (rank analysis); and (ii) the ability to assign reliable scores to its models, measured through Receiver Operator Characteristics curves (ROC; note this aspect is often referred to with 'fold recognition'). Methods are ranked through both the 'common subset' and 'pairwise' comparison approaches.

DISCUSSION

EVA provides an automated and continuous evaluation. Every week, test sequences are automatically submitted to prediction servers and results are evaluated and posted on the EVA web sites. The test sets are constructed so that methods could not have been trained based on the sequences in the test sets. Moreover, the test sets are as large as possible. In addition, the reliability of the comparisons between methods is maximised by using only test sets common to the methods assessed.

EVA provides supplemental information to CASP. Since 1994, the development of structure prediction methods has been influenced by the CASP meetings. While EVA uses well-defined numerical criteria to evaluate sustained performance, expert evaluations are still needed to learn what measures are most useful. However, human assessors are not likely to be able to handle many more test sequences than those at CASP. At the same time, there are problems with ranking methods based on test sets that are too small (1–3). EVA rankings are statistically more significant than those at CASP, because EVA assesses prediction methods continuously on as many proteins every month as CASP in 2 years (1). We believe that CASP needs to be supplemented by a large-scale, automated and continuous assessment, such as that by LiveBench (25) (assessment for threading methods only) and EVA. In fact, EVA may replace certain CASP categories in the future. For example, it was proposed at the last 2002 CASP meeting to eliminate secondary structure predictions from CASP. Instead, EVA-sec will replace CASP/CAFASP for users interested in those methods. This decision was partially influenced by the fact that the evaluation of secondary structure prediction methods has matured and this matured analysis has demonstrated beyond doubt that the set of proteins at CASP5 (2002) was not representative and too small.

EVA allows developers to focus on developing better methods. The best secondary structure prediction methods have reached a sustained level of 76% accuracy for the last 2 years (2) which indicates a substantial improvement in secondary structure prediction over the last 4 years. While it is always difficult to choose an appropriate set of measures, EVA uses standard criteria that have been largely used by experts in the area. For secondary structure prediction, these criteria are well established. For all other categories, we are currently experimenting with new criteria, others will be incorporated into EVA upon request from users. The precise definitions of the criteria are available on the web. While we can make our original scripts available upon request, we currently do not have the resources to cast the whole EVA code into a form that guarantees portability or ease-of-use. Overall, EVA allows developers to focus on the development of better methods, rather than on the generally time-consuming evaluation.

Extension of the EVA framework to other prediction categories. In principle, the concepts implemented in EVA could and should be generalised to evaluating a larger variety of prediction methods. Often, the problem is the availability of new high-resolution data. We intend to explore extensions that cover the predictions of protein–protein interactions, membrane regions, signal peptides, cleavage sites, structural/functional motifs and sub-cellular localisation.

ACKNOWLEDGEMENTS

Thanks to Jinfeng Liu and Megan Restuccia (Columbia) for computer assistance. We are grateful to members of the Protein Design Group. The contribution of the PDG is supported in part by a grant from the Spanish Ministry of Science and Technology (PDG, CNB-CSIC). I.Y.Y.K. was supported by the grant 5-P20-LM7276 from the National Institute of Health (NIH); D.P. was supported by the NIH grant RO1-GM63029-01, A.S., M.A.M.R., M.S.M. and N.E. by the NIH grants R01 GM54762 and P50 GM62529, B.R. by the NIH grant 1-P50-GM62413-01 and the NSF grant DBI-0131168. Thanks to Phil Bourne (UCSD) and the RCBS crews for maintaining an excellent PDB and to all experimentalists who enabled this analysis by making their data publicly available. Last, but not least, thanks to all those developers who support EVA by going through the trouble of making their methods publicly available.

REFERENCES

1. Eyrich,V., Martí-Renom,M.A., Przybylski,D., Fiser,A., Pazos,F., Valencia,A., Sali,A. and Rost,B. (2001) EVA: continuous automatic evaluation of protein structure prediction servers. *Bioinformatics*, **17**, 1242–1243.
2. Rost,B. and Eyrich,V. (2001) EVA: large-scale analysis of secondary structure prediction. *Proteins*, **45** (Suppl. 5), S192–S199.
3. Marti-Renom,M.A., Madhusudhan,M.S., Fiser,A., Rost,B. and Sali,A. (2002) Reliability of assessment of protein structure prediction methods. *Structure*, **10**, 435–440.
4. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
5. Eyrich,V.A. and Rost,B. (2003) META-PP: single interface to selected web servers. *Nucleic Acids Res.*, **31**, 3308–3310.
6. Eyrich,V. and Rost,B. (2000). *CUBIC*. Columbia University, Department of Biochemistry and Molecular Biophysics.
7. Sander,C. and Schneider,R. (1991) Database of homology-derived structures and the structural meaning of sequence alignment. *Proteins*, **9**, 56–68.
8. Rost,B. (1999) Twilight zone of protein sequence alignments. *Protein Eng.*, **12**, 85–94.
9. Altschul,S., Madden,T., Shaffer,A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D. (1997) Gapped Blast and PSI-Blast: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
10. Fariselli,P., Olmea,O., Valencia,A. and Casadio,R. (2001) Prediction of contact maps with neural networks and correlated mutations. *Protein Eng.*, **14**, 835–843.
11. Pollastri,G. and Baldi,P. (2002) Prediction of contact maps by GIOHMMs and recurrent neural networks using lateral propagation from all four cardinal corners. *Bioinformatics*, **18**, S62–S70.

12. Bonneau,R., Ruczinski,I., Tsai,J. and Baker,D. (2002) Contact order and ab initio protein structure prediction. *Protein Sci.*, **11**, 1937–1944.
13. Goebel,U., Sander,C., Schneider,R. and Valencia,A. (1994) Correlated mutations and residue contacts in proteins. *Proteins*, **18**, 309–317.
14. Pazos,F., Helmer-Citterich,M., Ausiello,G. and Valencia,A. (1997) Correlated mutations contain information about protein–protein interaction. *J. Mol. Biol.*, **271**, 511–523.
15. Ortiz,A.R., Kolinski,A., Rotkiewicz,P., Ilkowski,B. and Skolnick,J. (1999) *Ab initio* folding of proteins using restraints derived from evolutionary information. *Proteins*, **37** (Suppl. 3), 177–185.
16. Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
17. Frishman,D. and Argos,P. (1995) Knowledge-based protein secondary structure assignment. *Proteins*, **23**, 566–579.
18. Rost,B., Sander,C. and Schneider,R. (1994) Redefining the goals of protein secondary structure prediction. *J. Mol. Biol.*, **235**, 13–26.
19. Zemla,A., Venclovas,C., Fidelis,K. and Rost,B. (1999) A modified definition of SOV, a segment-based measure for protein secondary structure prediction assessment. *Proteins*, **34**, 220–223.
20. Levitt,M. and Chothia,C. (1976) Structural patterns in globular proteins. *Nature, London*, **261**, 552–558.
21. Levitt,M. (1976) A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.*, **104**, 59–107.
22. Ortiz,A.R., Strauss,C.E. and Olmea,O. (2002) MAMMOTH (Matching molecular models obtained from theory): an automated method for model comparison. *Protein Sci.*, **11**, 2606–2621.
23. Cristobal,S., Zemla,A., Fischer,D., Rychlewski,L. and Elofsson,A. (2001) A study of quality measures for protein threading models. *BMC Bioinformatics*, **2**, 5.
24. Shindyalov,I.N. and Bourne,P.E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.*, **11**, 739–747.
25. Bujnicki,J.M., Elofsson,A., Fischer,D. and Rychlewski,L. (2001) LiveBench-1: continuous benchmarking of protein structure prediction servers. *Protein Sci.*, **10**, 352–361.
26. Bates,P.A., Kelley,L.A., MacCallum,R.M. and Sternberg,M.J. (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins*, **45** (suppl.), 39–46.
27. Bates,P.A. and Sternberg,M.J. (1999) Model building by comparison at CASP3: Using expert knowledge and computer automation. *Proteins*, **37**, 47–54.
28. Lund,O., Hansen,J.E., Brunak,S. and Bohr,J. (1996) Relationship between protein structure and geometrical constraints. *Protein Sci.*, **5**, 2217–2225.
29. Lambert,C., Leonard,N., De Bolle,X. and Depiereux,E. (2002) ESyPred3D: Prediction of proteins 3D structure. *Bioinformatics*, **18**, 1250–1256.
30. Guex,N., Diemand,A. and Peitsch,M.C. (1999) Protein modelling for all. *Trends Biochem. Sci.*, **24**, 364–367.
31. Kelley,L.A., MacCallum,R.M. and Sternberg,M.J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.*, **299**, 499–520.
32. Altschul,S.F. and Gish,W. (1996) Local alignment statistics. *Methods Enzymol.*, **266**, 460–480.
33. Xu,Y. and Xu,D. (2000) Protein threading using PROSPECT: design and evaluation. *Proteins*, **40**, 343–354.
34. Karplus,K., Barrett,C., Cline,M., Diekhans,M., Grate,L. and Hughey,R. (1999) Predicting protein structure using only sequence information. *Proteins*, **S3**, 121–125.
35. Karplus,K., Karchin,R., Barrett,C., Tu,S., Cline,M., Diekhans,M., Grate,L., Casper,J. and Hughey,R. (2001) What is the value added by human intervention in protein structure prediction? *Proteins*, **45**, 86–91.
36. Gough,J. and Chothia,C. (2002) SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments. *Nucleic Acids Res.*, **30**, 268–272.
37. Olmea,O. and Valencia,A. (1997) Improving contact predictions by the combination of correlated mutations and other sources of sequence information. *Fold. Design*, **2**, S25–S32.
38. Olmea,O., Rost,B. and Valencia,A. (1999) Effective use of sequence correlation and conservation in fold recognition. *J. Mol. Biol.*, **293**, 1221–1239.
39. Pazos,F., Olmea,O. and Valencia,A. (1997) A graphical interface for correlated mutations and other protein structure prediction methods. *Comp. Appl. Biol. Sci.*, **13**, 319–321.
40. Raghava,G.P.S. (2000) Protein secondary structure prediction using nearest neighbor and neural network approach. *CASP4*, 75–76.
41. Cuff,J.A., Clamp,M.E., Siddiqui,A.S., Finlay,M. and Barton,G.J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics*, **14**, 892–893.
42. Rost,B. (1996) PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.*, **266**, 525–539.
43. Przybylski,D. and Rost,B. (2002) Alignments grow, secondary structure prediction improves. *Proteins*, **46**, 195–205.
44. Ouali,M. and King,R.D. (2000) Cascaded multiple classifiers for secondary structure prediction. *Protein Sci.*, **9**, 1162–1176.
45. Rost,B. (2001) Protein secondary structure prediction continues to rise. *J. Struct. Biol.*, **134**, 204–218.
46. Jones,D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, **292**, 195–202.
47. McGuffin,L.J., Bryson,K. and Jones,D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404–405.
48. Karplus,K., Barrett,C. and Hughey,R. (1998) Hidden Markov models for detecting remote protein homologies. *Bioinformatics*, **14**, 846–856.
49. Pollastri,G., Przybylski,D., Rost,B. and Baldi,P. (2002) Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins*, **47**, 228–235.