PROTEINS: Structure, Function, and Genetics 53:389–394 (2003)

Modeling Three-Dimensional Protein Structures for CASP5 Using the 3D-SHOTGUN Meta-Predictors

Iris Sasson and Daniel Fischer*
Bioinformatics, Department of Computer Science, Ben Gurion University, Beer-Sheva, Israel

ABSTRACT **Full-atom models were generated for all CASP5 targets by using the fully automated 3D-SHOTGUN fold recognition meta-predictors (Fischer D, Proteins 2003;51:434–441). The 3D-SHOTGUN meta-predictors assemble hybrid 3D models by combining structural information of a number of independently generated, fold recognition models. At the time CASP5 took place, the 3D-SHOTGUN servers generated unrefined C_α-only models. Fischer's participation in CASP had three main goals. The first was to test the value of using 3D-SHOTGUN models as input to a refinement procedure. The second goal was to test whether human intervention could result in a better performance than that of the automated servers. The third goal was to evaluate which human procedures, not yet implemented within the 3D-SHOTGUN servers, can be implemented in the future. For CASP5, our group's predictions applied a very simple approach using the multiple parent option of the Modeller program (Sali and Blundell, J Mol Biol 1993;234:779–815). The input to Modeller was different combinations of the unrefined 3D-SHOTGUN models and the sequence-template alignments used by 3D-SHOTGUN's assembly step. Our evaluation of the accuracies of the refined versus the SHOTGUN models shows that the refined models were consistently slightly more accurate than SHOTGUN's. For a few targets, the manual use of the information from the CAFASP servers resulted in better human models. This manual intervention was particularly valuable in the identification of domains, still a difficult feature for automated servers. The CASP5 results indicate that 3D-SHOTGUN's hybrid models can be a valuable starting point for full-atom refinement and that the resulting refined models are, on average, more accurate than those produced by the servers. Thus, we conclude that our three goals were achieved. A preliminary automated version of the refinement procedure, named SHGUM, is now available. Proteins 2003;53:389–394.** © 2003 Wiley-Liss, Inc.

Key words: homology modeling; protein fold recognition; protein structure prediction; critical assessment of protein structure prediction; 3D-SHOTGUN meta-predictor

INTRODUCTION

The need for accurate automated protein structure-modeling tools has become increasingly evident.[1,2] However, experiments, such as CASP, have shown that current homology modeling and fold recognition methods continue to require improvements in accuracy at all levels of homology distances. When a close homologue of known structure exists, the sequence-template alignments produced by current methods are in many cases accurate, but the fine details of the generated models (e.g., loops, and side-chains) are still below desireable levels of accuracy. When only distant homologues exist, inaccuracies, such as shifts and gap placement in the sequence-template alignments, continue to be a limiting factor. Furthermore, for targets having only very distant homologues of known structure, the sequence-template alignments are often of relatively poor quality, and even the identification of the correct homologue(s) can become a difficult problem.

Generating full-atom models using existing structures as templates has traditionally been applied only to those cases of clear homology. However, as fold recognition methods have improved, it seems that the time is ripe to aim at generating relatively accurate, full-atom models at larger homology distances. Thus, the boundaries between classical comparative or homology-modeling methods and fold recognition methods are fading. Modern fold recognition methods, which are aimed at producing accurate sequence-template alignments even with little or insignificant sequence similarity, have thus become essential components of distant homology modeling.

Our group's participation in CASP had three main goals. The first was to test the expected accuracy of full-atom refined models generated by using as input the results of the fully automated fold recognition models produced by the 3D-SHOTGUN meta-predictors.[3] During the prediction session of CASP5, the 3D-SHOTGUN meta-predictors did not include a refinement step and produced C_α-only unrefined models. Consequently, for the hardest targets, the resulting unrefined models often contained a number of non-protein-like regions (e.g., fragmentation and colli-

Grant sponsor: United States-Israel Binational Science Foundation, Jerusalem, Israel.

*Correspondence to: Daniel Fischer, Bioinformatics, Department of Computer Science, Ben Gurion University, Beer-Sheva 84015, Israel. E-mail: dfischer@cs.bgu.ac.il

Received 13 February 2003; Accepted 6 April 2003

sions), requiring subsequent refinement. The second goal was to test whether human intervention could result in a better performance than that of the automated servers. The third goal was to evaluate which human procedures, not yet implemented within the 3D-SHOTGUN servers, can be implemented in the future. The comparative modeling and the fold recognition CASP5 assessors ranked Fischer's predictions among the top CASP groups (and above the 3D-SHOTGUN meta-predictors), suggesting that the 3D-SHOTGUN models can be valuable starting points for full-atom distant homology modeling and that the performance of the automated servers could be improved if some of the human procedures applied are incorporated.

MATERIALS AND METHODS

Two approaches were used. In the first approach, a fully automated procedure, developed during the CASP5 prediction session, was applied. This procedure makes use of the information from 3D-SHOTGUN's version SHGU only. SHGU is an independent meta-predictor that runs locally by using the output from the bioinbgu server (http://www.cs.bgu.ac.il/bioinbgu[4]) without additional information from other servers. We have named this automated procedure "SHGUM," for 3D-SHotGUn-Modeller. In the second approach, a semi-automated procedure was applied, in which the information from other servers was also used. The latter included mainly information from the other two versions of 3D-SHOTGUN, 3DS3, and 3DS5.[3] In addition to the information from bioinbgu, 3DS3 and 3DS5 use the information from two and four other external servers, respectively. The predictions using the latter approach were filed to CASP5 under the group name "Fischer," whereas those of the former approach (i.e., those obtained by the "SHGUM" procedure) were filed under the group name "Sasson."

Both procedures generate full-atom refined models by applying Sali's Modeller[5] program using multiple templates. In comparative modeling, it has become clear that the use of (good) multiple templates usually results in better models. Templates were selected from among those used to assemble the rank-1 hybrid model computed by the 3D-SHOTGUN servers. The sequence-structure alignments corresponding to the models used in the 3D-SHOTGUN's rank-1 model (stored in an internal list generated by the servers) were fed as input to Modeller (for a detailed illustration of the procedure see the T0195 example below).

For some targets filed by Fischer, the actual C_α-only hybrid produced by SHOTGUN was also given as input to Modeller as an additional template. Manual browsing of the results from other CAFASP servers helped select additional templates in two scenarios. The first was when none of the SHOTGUN servers had an above-confidence-thresholds prediction. The second scenario was when domain partitioning, different from that compiled in CAFASP, was apparent from the various servers' results.

One problem in our use of Modeller arose because the assembly conducted by the SHOTGUN servers can often include significantly different models based on structurally inconsistent parents.[3] Consequently, Modeller often failed to generate a model due to restraints violations, such as collisions between the different models. We iteratively removed sequence-template alignments from the input until Modeller succeeded to produce a refined model. The order of alignment removal was determined according to their similarity to SHOTGUN's rank-1 hybrid, beginning with the most dissimilar (see the T0195 example below). This removal procedure usually terminated after few iterations and always included at least one alignment (Modeller almost never fails for a single template).

RESULTS

Models for all CASP5 targets were submitted, without regard to their level of homology to proteins of known structure. To quantify the quality of our models, here we use the MaxSub[6] sequence-dependent measure, which assigns scores in the range from 0 (models with unsignificant structural similarity to the native structure) to 1.0 (perfect match). According to MaxSub, the accuracies of the Fischer refined, full-atom models were, in general, very similar to the accuracies of the unrefined 3DS3 and 3DS5 models. Similarly, the accuracy difference between Sasson's models and that of SHGU's models were, in general, very slight. For most targets, if the 3D-SHOTGUN models corresponded to a "correct" model (a positive MaxSub score), the human models were also correct, and with very similar scores, and conversely, if the 3D-SHOTGUN models were "incorrect" (zero MaxSub score), then the human models were also incorrect. This result is consistent with a benchmark test of SHGUM conducted at Leszek Rychlewski's toolshop at http://bioinfo.pl using LiveBench-4 targets.[7] In this test, SHGUM's performance was only slightly higher ($<5\%$ cumulative MaxSub scores) than that of SHGU.

Target T0165, the crystal structure of *B. subtilis's* cephalosporin C deacetylase,[8] illustrates a typical example. Figure 1 shows that the core of the protein was correctly modeled by 3D-SHOTGUN and that a number of surface loops were not only incorrect but were also modeled as unconnected fragments. Figure 1 shows that the Fischer model removed the fragmentation and produced a "protein-like" model for all regions but, unfortunately, did not produce a more accurate model (the MaxSub scores of the 3D-SHOTGUN and Fischer models were 5.0 and 5.1, respectively). In general, the Fischer models relieved most of the "non-protein-like" regions of the 3D-SHOTGUN models but, unfortunately, these regions generally remained non-native-like. This is so because most of the fragmentation in the 3D-SHOTGUN models occurs at the "difficult" to model regions where little structural consistency is found among the initial models. However, the most important feature of the human models is that they were refined, protein-like, full-atom models, lacking the fragmentation present in some of the SHOTGUN models. Thus, obtaining similarly accurate models lacking the non-protein-like features of the 3D-SHOTGUN models can be considered a positive feature of our approach.

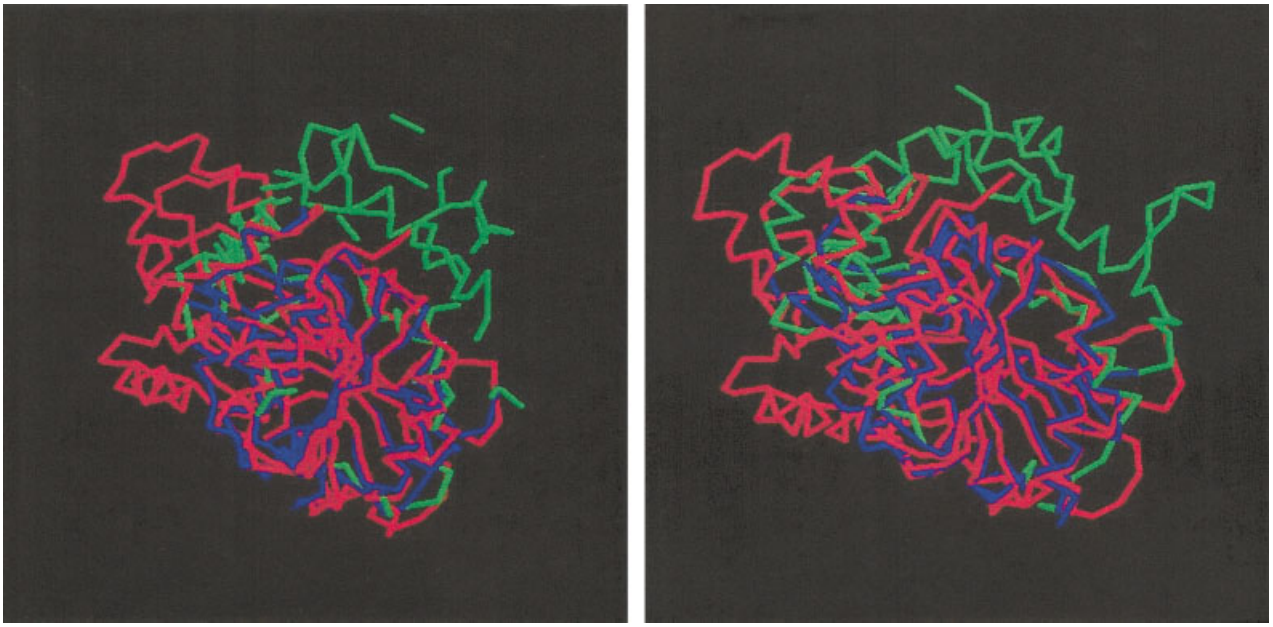

Fig. 1. Our human models were only slightly more accurate than those produced by the 3D-SHOTGUN servers. **a:** MaxSub superposition of the native structure of T0165 (red: 318 residues) with the unrefined hybrid 3D-SHOTGUN model (blue: correctly predicted regions; green: wrongly predicted regions). The sequence similarity of T0165 to its closest structural template is 18%. The MaxSub score of this model was 5.0, with a "well-predicted" subset of size 194 superimposing to native with an RMSD of 1.8 Å. **b:** MaxSub superposition of the native structure of T0165 with Fischer's refined model [colors in this and all other figures are as in Fig. 1(a)]. The MaxSub score of this model was 5.1, with a "well-predicted" subset of the same size as the 3D-SHOTGUN model (194) superimposing to native with an RMSD of 1.7 Å. Notice that the correctly model regions are almost identical in both models, with most differences occurring at the difficult to model regions, in this case corresponding to a number of non-core regions. Most of the fragmentation of the 3D-SHOTGUN models occur at regions where little structural consistency is found among the initial models used to build the assembled hybrid model. In general, these regions were not corrected in Fischer's models. Nevertheless, relieving most of the "non-protein-like" regions of the 3D-SHOTGUN models, without sacrificing accuracy, is a positive feature of our approach.

We have identified 21 targets where either Fischer's models or Sasson's models significantly differed from those obtained by the 3D-SHOTGUN servers (Table I). Asterisks indicate predictions that had MaxSub scores higher than any other CAFASP server (and in a number of cases also higher than any other CASP prediction). Target T0157, and in particular target T0170, illustrate the case in which the human predictions were inferior to those of the 3D-SHOTGUN servers. Sasson's predictions for targets T0134_2, T0183, and T0195 and Fischer's predictions for targets T0138, T0148_2, and T0172_1 are examples in which the human predictions were superior to those of the 3D-SHOTGUN servers.

Figure 2 shows the superposition of Fischer's T0148_2 model with the native structure (protein HI1034 from *H. influenzae*). This example illustrates a case in which Fischer's model excelled. Fischer's model had a MaxSub score of 3.4, which corresponds to a relatively good prediction for a target in the FR(A) category. The figure shows that the major secondary structure elements were well predicted with a subset of 57 residues (out of a total of 91 residues in the target) that superimpose into the native structure with an root-mean-square deviation (RMSD) of 3.0 Å. This model was produced by human observation of the domain structure of the target and by analyzing the results of other fold recognition tests, in addition to the results reported by the SHOTGUN servers.

Target T0172_1, corresponding to the crystal structure of *T. maritima's* protein Tm0872, a putative Sam-depen-

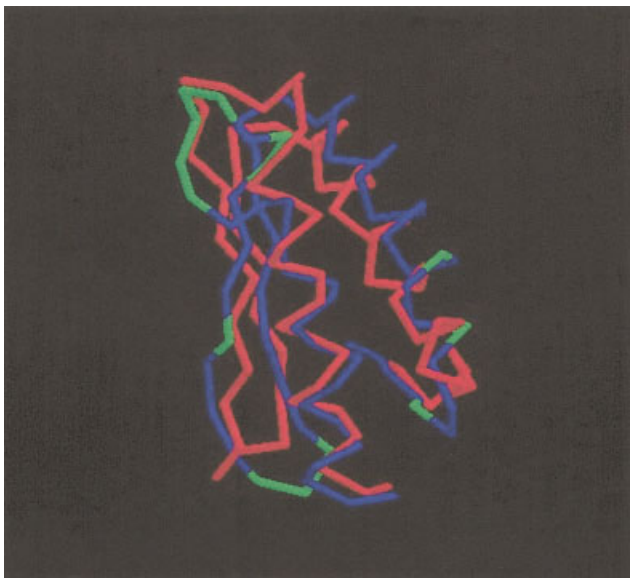

Fig. 2. Fischer's prediction for a difficult target: T0148_2. MaxSub superposition of the native structure of T0148_2 (red: 91 residues; for clarity, only residues 10–80 are shown) with the refined model submitted by Fischer (blue: correctly predicted regions; green: wrongly predicted regions). The sequence similarity of T0148 to its closest structural template is only 15%. This model had a MaxSub score of 3.4. MaxSub (with a distance threshold of 5.0 Å) identified a "well-predicted" subset of size 57, which superimposes to native with an RMSD of 3.0 Å.

TABLE I. Selected Human Predictions With Significant Differences to Those Produced by the 3D-SHOTGUN Meta-Predictors

			MaxSub Scores				
Target	Cat	Id%	Fis	Sas	SHGU	3DS3	3DS5
T0134_2	FR(H)	13	6.0	6.5	5.6	5.5	5.9
T0136_2	CM/FR	17	3.5*	3.2	0.0	0.0	2.5
T0137	CM	43	9.5*	9.5	9.0	9.4	9.4
T0138	FR(H)	19	6.1*	5.6	5.4	5.3	5.6
T0142	CM	26	6.7	6.6	5.3	5.2	4.8
T0143	CM	28	8.1*	7.3	6.4	6.6	6.5
T0148_2	FR(A)	15	3.4*	0.0	0.0	0.0	0.0
T0150	CM	34	7.6	8.4*	7.9	8.1	7.9
T0152	CM/FR	12	4.1*	4.0	3.2	3.4	3.7
T0157	FR(H)	15	4.2	4.1	3.5	4.5	4.7*
T0165	CM/FR	18	5.1*	3.7	4.6	5.0	5.0
T0169	CM/FR	9	5.7	6.7*	6.3	6.3	6.5
T0170	NF/FR	n.a.	0.0	3.0	3.0	2.9	3.1
T0172_1	CM/FR	16	4.1	3.1	3.8	3.8	3.8
T0177_1	CM	30	9.2*	9.2	8.0	8.5	8.7
T0177_2	CM	30	9.1*	8.9	8.8	8.6	8.5
T0183	CM	30	7.1	7.6*	6.7	6.6	6.9
T0185_3	CM/FR	22	4.0	5.4*	5.4	4.6	5.3
T0188	CM	31	7.6*	7.5	7.3	7.5	6.9
T0189	CM/FR	14	4.7	5.3*	4.9	5.0	5.1
T0195	CM/FR	18	5.1	6.0*	5.5	5.6	5.5

Asterisks indicate predictions with MaxSub scores higher than any other CAFASP server (and in a number of cases also higher than any other CASP prediction). Cat, the targets category as classified by the CASP5 assessors; Id% the sequence identity percentage of the target to its closest protein of known structure; Fis, Fischer's models; Sas, Sasson's models; SHGU, 3DS3, 3DS5. The three versions of the 3D-SHOTGUN meta-predictors (see text).

dent Methyltransferase,[9] illustrates a case in which human intervention allowed for the identification of a multidomain target. Figure 3 shows that T0172_1 is composed of two separate segments of the sequence. Current automatic servers are not very efficient in identifying such multidomain proteins. However, human observation allowed us to identify this particular domain, obtaining a model of higher accuracy than those produced by the servers.

Finally, T0195, a yeast hypothetical esterase, illustrates a case in which Sasson's model excelled. This model was not only superior to that submitted by SHGU and the other 3D-SHOTGUN meta-predictions but also superior to all the other models submitted to CASP. We also use this example to illustrate how our automated refinement method SHGUM works. The SHGU meta-predictor produced a hybrid model using 25 of the alignments reported by the bioingu server (all five top models from each of the five methods run by bioinbgu). These 25 models generally corresponded to different sequence-structure alignments using 9 different templates corresponding to 5 different families of the α/β-hydrolase fold in SCOP.[10] SHGU's hybrid, C_α-only unrefined model had a MaxSub score of 5.5 with a subset of 190 residues that superimposed well over the native structure. Thus, to generate a refined full-atom model, SHGUM used the same 25 alignments as input into Modeller. However, 11 of these alignments were iteratively removed from the input to Modeller because too many violations were present (see Materials and Methods). These could be due either to significant structural differences between these templates, to significant differences in the sequence-template alignments, or both. In this case, the 11 rejected alignments corresponded to models less similar to the native structure than the 14 accepted ones. The templates used in these 14 alignments included two members of the carboxylesterase/thioesterase family (1gkl and 1jjf) and a member of the mycobacterial antigen family (1dqy). By analyzing these 14 alignments, it can be observed that generally they differ in the loop regions. SHGUM's resulting model had a MaxSub score of 6.0, with a core of 200 residues at an RMSD to native of only 0.58 Å (Fig. 4). It is worthwhile mentioning that all of the initial models used as input had a significantly lower accuracy. This finding shows that the use of a number of carefully selected templates can result in a significantly better model. It seems that the use of SHGU's information contained in these particular three templates, in combination with the 14 alternative alignments, allowed us to generate a model closer to native than any other model submitted to CASP.

DISCUSSION

Although a detailed analysis of the reasons why some of our refined models had lower accuracies than those obtained by the unrefined ones is difficult to conduct, one possible reason may be the following. Some of the unrefined models contain structurally conflicting regions. Parts of these regions may correspond to native-like features, which contribute to a higher score of the unrefined models. The refinement process needs to choose among the various alternatives in such regions and does not always succeed

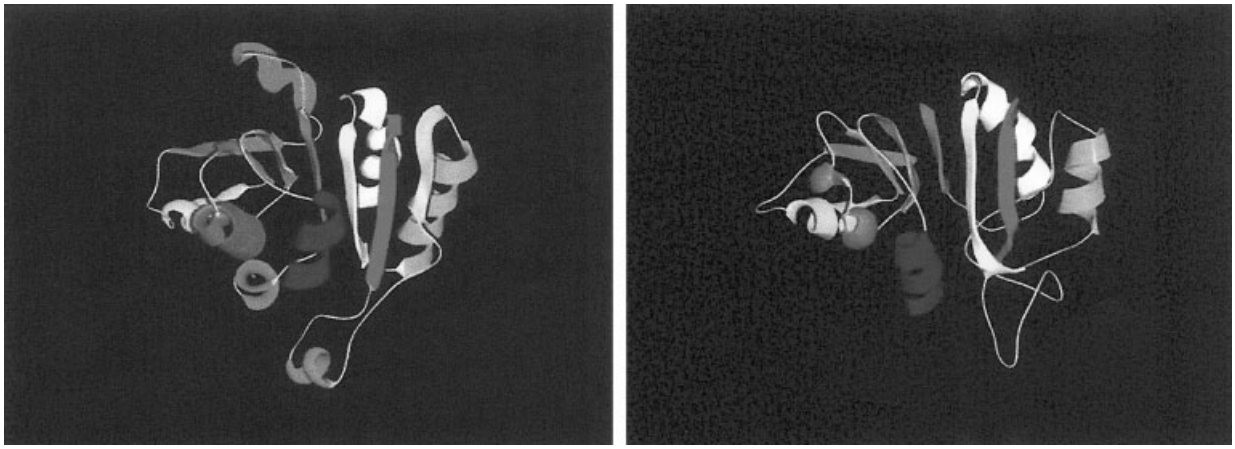

Fig. 3. Fischer's multidomain prediction for target T0172_1. The sequence similarity of T0172 to its closest structural template is only 16%. Fischer's model (**a**) had a MaxSub score of 4.1, with a subset of 132 residues superimposing to the native structure (**b**) with an RMSD of 2.7 Å. The models are shown in the orientation corresponding to the optimal superposition of the well-predicted subset identified by MaxSub.

Fig. 4. Sasson's superior prediction for target T0195. The sequence similarity of T0195 to its closest structural template is 18%. Sasson's model (**a**) had a MaxSub score of 6.0, with a core of 200 residues superimposing to native (**b**) with an RMSD of 0.58 Å.

to choose the native-like features. Thus, some of the resulting refined models may obtain lower scores than those obtained by the unrefined model. Efforts toward correcting such cases are part of our ongoing research.

The full-atom refined models generated by our methods were, in general, "plausible" models in the sense of correct, "protein-like" geometry, without the fragmentation present in some of the 3D-SHOTGUN models. In addition, our full-atom models were, in general, slightly more accurate than the 3D-SHOTGUN's models. On the one hand, solving the steric problems of 3D-SHOTGUN's models without sacrificing accuracy can be considered a success of our approach. On the other hand, the fact that the refined models were only slightly more accurate than the unrefined ones leaves much room for improvement. Further improvements may be achieved by using alternative refinement procedures, which are currently being investigated. A valuable feature to be implemented is to attach a "B-factor"-like reliability score to each modeled residue. This score is an indication of the amount of structural (in)consistency observed among the initial models.

The CASP5 results of our group indicate that there are a number of tasks that humans (us) still perform better than the servers (e.g., domain identification). Our next challenge is to identify as many successful human practices that are "computable," so that they can be incorporated as future improvements of the automated servers. Although the performance difference of humans and machines is narrowing, it is likely that humans will still outperform automated methods in the upcoming CASP6. Thus, CASP6 will continue to help identify further aspects of human expertise that need to be incorporated into automated methods.

ACKNOWLEDGMENTS

We thank the CASP5 organizers and assessors; 3DS3 and 3DS5 could not exist without the availability of the initial model-generating servers such as ffas,[11] 3D-PSSM,[12] GenThreader,[13] and Fugue[14]; our thanks to their developers. We especially thank Leszek Rychlewski for the availability of his meta-server at http://bioinfo.pl/meta.

REFERENCES

1. Fischer D, Baker D, Moult J. We need both computer models and experiments (correspondence). Nature 2001;409:558.
2. Fischer D, Elofsson A, Pazos F, Valencia A, Rost B, Ortiz AR, Dunbrack RL Jr. CAFASP2: The second critical assessment of fully automated structure prediction methods. Proteins 2001; Suppl 5:171–183. Special issue; see also http://www.cs.bgu.ac.il/~dfischer/CAFASP2.
3. Fischer D. 3D-SHOTGUN: a novel, cooperative, fold recognition meta-predictor. Proteins 2003;51:434–441.
4. Fischer D. Combining sequence derived properties with evolutionary information. Proceedings Pacific Symposium on Biocomputing. Jan 2000. p 119–130.
5. Sali A, Blundell T. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
6. Siew N, Elofsson A, Rychlewski L, Fischer D. MaxSub: an automated measure for the assessment of protein structure prediction quality. Bioinformatics 2000;16:776–785.
7. Fischer D, Rychlewski L. The 2002 olympic games of protein structure prediction. Protein Eng 2003. Forthcoming.
8. Zhang R, Koroleva O, Collert F, Joachimiak A. 1.5 Å crystal structure of the cephalosporin C deacetylase. 2003. Forthcoming.
9. Miller D, Ouellette N, Evdokimova E, Savchenko A, Edwards A, Anderson WF. Crystal complexes of a predicted Sam-dependent methyltransferase reveal a typical Sam domain and a substrate recognition domain. 2003. Forthcoming.
10. Murzin AG, Brenner SE, Hubbard T, Chothia C. Scop: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247:536–540.
11. Rychlewski L, Jaroszewski L, Li W, Godzik A. Comparison of sequence profiles. strategies for structural predictions using sequence information. Protein Sci 2000;9:232–241.
12. Kelley LA, MacCallum RM, Sternberg MJ. Recognition of remote protein homologies using three-dimensional information to generate a position specific scoring matrix in the program 3D-PSSM. RECOMB 99, Proceedings of the Third Annual Conference on Computational Molecular Biology, 1999. p 218–225.
13. Jones DT. Genthreader: an efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol 1999;287: 797–815.
14. 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.