PROTEINS: Structure, Function, and Genetics Suppl 3:22–29 (1999)

# Processing and Analysis of CASP3 Protein Structure Predictions

**Adam Zemla,[1] Česlovas Venclovas,[1] John Moult,[2] and Krzysztof Fidelis[1]***
[1]*Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California*
[2]*Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland*

***ABSTRACT*** **Livermore Prediction Center provides basic infrastructure for the CASP (Critical Assessment of Structure Prediction) experiments, including prediction processing and verification servers, a system of prediction evaluation tools, and interactive numerical and graphical displays. Here we outline the essentials of our approach, with discussion of the superposition procedures, definitions of basic measures, and descriptions of new methods developed to analyze predictions. Our primary focus is on the evaluation of three-dimensional models and secondary structure predictions. To put the results of the three prediction experiments held to date on the same footing, the latest CASP3 evaluation criteria were retrospectively applied to both CASP1 and CASP2 predictions. Finally, we give an overview of our website (*http://PredictionCenter.llnl.gov*), which makes the target structures, predictions, and the evaluation system accessible to the community. Proteins Suppl 1999;3:22–29.** **Published 1999 Wiley-Liss, Inc.†**

**Key words: prediction evaluation; prediction assessment; comparative modeling; fold recognition; ab initio prediction**

## INTRODUCTION

With the significant expansion of activity in the structure prediction field, processing and subsequent analysis of predictions has become increasingly complex. Livermore Prediction Center addresses this problem in three ways: first, by soliciting and verifying prediction targets and by overseeing that their release status has not been compromised; second, by processing predictions, that is, format verification, submission updates, and compliance with specific target deadlines; and third, by evaluation of predictions, including development and implementation of evaluation methods, necessary calculations, and publication of the evaluation results.

One of the lessons learned from CASP is that analyzing the effectiveness of prediction methods is not a trivial matter. The relatively simple comparative tools we had at our disposal at the beginning of the process in 1994 proved cumbersome to use and in the final consideration not fully satisfactory. To evaluate predictions, first we need an analytical approach to identify what in a prediction worked and what failed. Second, we need a comparative approach, using both general and specialized techniques, to identify which methods work best, and which address a specific aspect of prediction most successfully.

Since the first CASP prediction experiment, a number of new measures and methods aimed at both the analytical and comparative aspects of evaluation have been developed. As the final assessment of predictions in the CASP process rests with the independent assessors, new evaluation criteria have often been introduced at that stage. To date, ten independent assessors have contributed to the process. Taken together with the methods suggested by the organizers, by consultancy groups, and by predictors, the number of evaluation techniques tried at CASP became considerable. Nevertheless, the methods we have at hand today still do not address all the evaluation problems to our complete satisfaction. Discussions of these matters continue in the community and new techniques are being considered.

This paper gives an outline of the evaluation methods presently implemented at the Prediction Center. We hope that together with other methods described in this issue it will provide a basis for discussion of which methods to use in the next round, the CASP4 prediction experiment.

## APPROACH

### Processing of Predictions

All predictions are accepted electronically. Only in the first CASP were other than electronic submissions considered. A need for a consistent and complete form of submissions, as well as their large number, quickly necessitated development of an efficient processing system. Each submitted model is automatically tested by the format verification server. Models that conform to the format and deadline requirements are assigned an accession code. A unique accession code is composed of the number of the prediction target, format category, submitting group number, and model index (a number assigned by prediction groups to rank their submissions).

Recommendation by the consultancy groups of putting major emphasis on 3D coordinate models and to eliminate pre-classification of targets into specific prediction categories resulted in a simplification of prediction formats. In



Česlovas Venclovas' permanent address is Institute of Biotechnology, Graičiūno 8, 2028 Vilnius, Lithuania.
*Correspondence to: Krzysztof Fidelis, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, CA 94550. E-mail: fidelis@llnl.gov
Received 8 June 1999; Accepted 15 June 1999

CASP3, predictions were accepted in a form of atomic coordinates, alignments to known, publicly available structures, secondary structure assignments, and residue-residue distances (TS, AL, SS, and RR respectively). Up to five models were accepted from a prediction group per each target, with one of them designated as the primary prediction (model 1). Submission of a duplicate model (same target, group, model index) would replace a previously accepted model, provided it was received before target's prediction deadline.

All AL type predictions were converted to 3D backbone coordinates prior to evaluation. Alternatively, a facility to translate sequence-structure alignments (AL format) into standard PDB atom records (TS format) was made available at the Prediction Center website[1]. There were relatively few RR type predictions submitted to CASP3. Although they were accepted and format-verified at Livermore, their evaluation has been carried out by Tim Hubbard and Christine Orengo's group—a clearly defined and community accepted assessment system has yet to be agreed upon and implemented. Here we discuss evaluation of the 3D and secondary structure predictions.

### Superposition Procedures

Before we describe the measures used in evaluation of predictions, let us consider the superposition procedures designed to spatially align models and targets. The first, most obvious approach is to align residues in both model and target in a 1:1 correspondence, or sequence-dependent manner. An advantage of this procedure, for all residues in a protein or for a specified subset, is that a unique and optimal superposition can easily be obtained.[2]

In many predictions, some regions of structure are often less well modeled than others. In such cases an iterative procedure can be applied to exclude the outlying residues from the superposition. The problem does not lend itself to optimal solutions easily; in practice only approximate results are produced. When most of the structure can be superimposed unambiguously, as in comparative modeling, we have used a single iterative procedure. For ab initio predictions we have generated a very large number of different superpositions to assure a very good approximation of the optimal one. This procedure (GDT) is described in more detail later in this article.

In some instances it is necessary to generate superpositions in a sequence-independent manner, unrestricted by the 1:1 sequence correspondence requirement. These are designed to recognize structural similarity even in the presence of model-target alignment errors. Sequence-independent superpositions present a considerably more involved computational problem. Procedures developed to date are suboptimal in nature and, therefore, at least to some degree arbitrary. In our calculations we have used results from a well established DALI method.[3] DALI has also been used to identify, whenever possible, close structural homologues and produce corresponding target-parent structural alignments (described below).

## EVALUATION OF 3D MODELS

### Basic Measures

The basic measure implemented in our evaluation software is the RMSD (root-mean-square-deviation) between model and target. It is used to measure differences between atomic coordinates, with results dependent on structural superposition, and between dihedral angles, independent of superposition. In the case of RMSD over coordinates, results are calculated for all atomic positions or subsets including Cα's, main chain, and side chain atoms. RMSD over dihedral angles is calculated separately for $\phi/\psi$, and for $\chi$ angles. In some cases, cutoff parameters are used to define subsets of structure. These are based on experience or generally accepted values and when performing evaluation they may be modified, if desired. Completeness of a prediction determines how many atomic positions or dihedral angles may be included in the evaluation. In each case we report:

NP—Number of Predicted atomic positions or dihedral angles

TN—corresponding Total Number in the target structure

PP—Percent Predicted (NP/TN)

Additionally, for dihedral angles we report:

PC—Percent of Correctly predicted (within a 30-degree cutoff).

To assess the quality of a prediction in more detail, measures described above may be applied to specific subsets of structure. Taking averages of quantities such as RMSD over the whole structure smoothes out the most interesting differences, so a range of subsets of the data have been defined to allow separate evaluations and statistics to be calculated for each. These have been designed to single out elements of protein structure and to eliminate the effect of possible experimental uncertainties (for brevity, only qualitative descriptions are given here, more details may be obtained from the Prediction Center web page):

"ALL" All atoms or dihedral angles possible to evaluate are considered.

"SECONDARY STRUCTURE" Secondary structure elements in the target structure. Helices and strands, as defined by DSSP,[4] with lower bounds of six residues for helix and three residues for strand, are included.

"SURFACE" Surface residues specified by calculating accessibility according to Lee and Richards[5] and fractional values relative to Shrake and Rupley's Gly-X-Gly standards.[6] Residues with values greater than cutoff (20% accessibility) are included.

"BURIED" This subset is complementary to "SURFACE."

"WELL ORDERED" In the case of crystallographically determined target structures, only parts of structure that are not affected by the uncertainty associated with thermal motion or disorder are considered.

"RELIABLE SIDE CHAINS" Segments of side chains deemed unreliable crystallographically (e.g., a rotation of 180 degrees could be undetectable) are excluded from Cartesian and angular RMSD calculations.

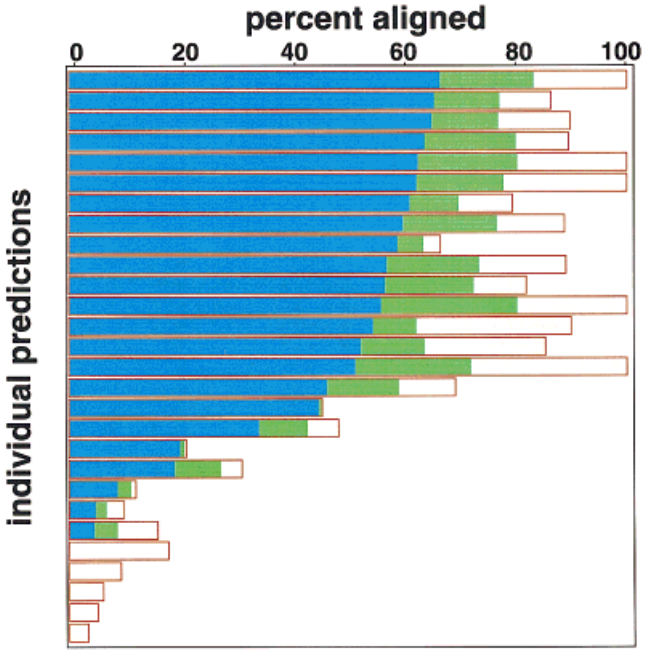


Fig. 1. Summary of alignment accuracy plots for all 3D predictions on target T0049 (EstB, P. marginata). The number of correctly aligned residues (see text) is shown in blue as a percentage of the total number of residues in this target. The additional fraction of residues aligned within +/-4 residues is shown in green, and the remainder of submitted prediction in red. Such alignment plots provide an immediate measure of the overall quality of predictions on a given target.

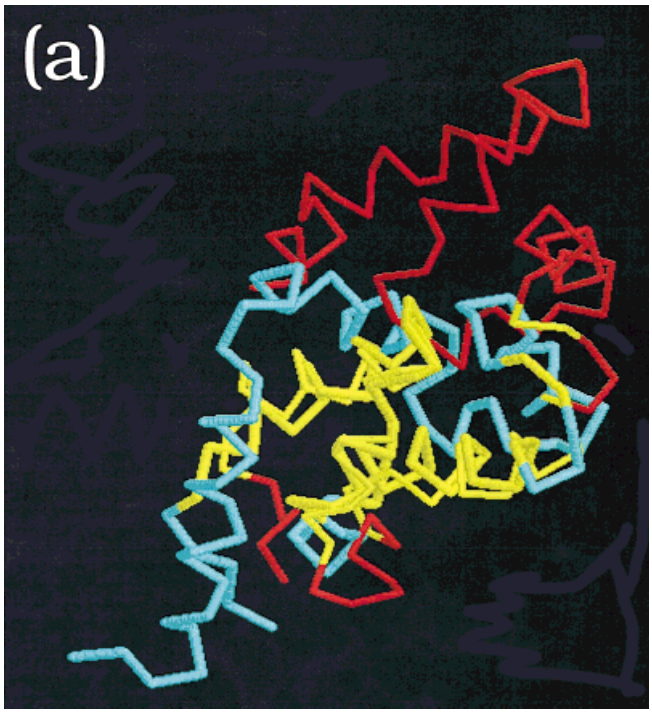


Fig. 2. An illustration of three representations of prediction quality: 3D plots, identification of longest continuous high quality segments (LCS), and largest superpositions (largest superimposable sets of residues) under a specified cutoff (GDT) for a single prediction on target T0056 (DnaB helicase N-terminal domain, E. coli, PDB code 1jwe). **A**: RASMOL Cα trace plot of the sequence-dependent model-target superposition (using only the Cα atom pairs closer than 2.5 Å to obtain the superposition). Target and model structures are shown in cyan and red, respectively. Corresponding atoms closer than 6 Å are shown in yellow. **B**: Longest continuous segments expressed as percentage of the total number of residues in the target structure. The three blue lines correspond to segments superimposable under 1, 2, and 5 Å RMS deviation respectively, with the uppermost line reflecting the largest cutoff value. DSSP assignments of secondary structure in the target and model are shown at the top of the plot, with helices in purple and strands in green. **C**: Largest superimposable sets of residues. Three blue lines correspond to the best fit for 5, 10, and 50 percent of the modeled structure, respectively. Secondary structure assignments as in B. A group of three correctly predicted helices (A, yellow) correspond to segments positioned approximately between residues 30 and 55 (two helices), and 65 and 80 (one helix). Both LCS and GDT show that under the most generous cutoff in either plot an additional helix (residues 83–93) may be fitted as well. Local secondary structure is predicted well throughout the main chain, even in the region where tertiary structure prediction is incorrect. Comparing LCS and GDT plots allows to immediately identify if the observed structural similarity is localized or if it extends to multiple regions in protein's sequence.

"NO INTERMOL CONTACTS" Parts of target structure that are not affected by interactions with neighboring molecules (crystal contacts).

### Additional Measures Used When a Homologous Structure was Available at Time of Prediction (Comparative Modeling)

To evaluate comparative modeling predictions, additional criteria have been developed based on the presence of an alignment with structural homologue(s). These, taken together with the basic measures just described, were designed first to isolate the different stages in comparative modeling, so that performance of methods addressing each can be assessed, and second to allow focus on segments of structure that are either relatively easy or more difficult to model. The design considerations have been previously discussed[7] and here we only briefly summarize the additional measures and subsets of structure. Alignment between target and structural homologue that is closest by sequence (principal parent) and with other potential template structures was done with DALI.

The additional comparative modeling measures include RMSD over Cα's calculated for three different sets of atoms:

(a) alignable region between target and principal parent;
(b) the same region for target and model;
(c) remaining region for target and model (loops).

Additional subsets of structure for comparative modeling include:

"CHANGED ANGLES" Angles that are rotamerically different from the corresponding ones in the parent structure.

"SHIFTED CHAIN" Segments of target structure that differ in position in the global alignment with parent by more than a cutoff (1 Å).

"ALTERNATIVE PARENT" Segments of target structure for which selection of a parent other than the one closest by sequence is preferred (better by more than 1 Å).

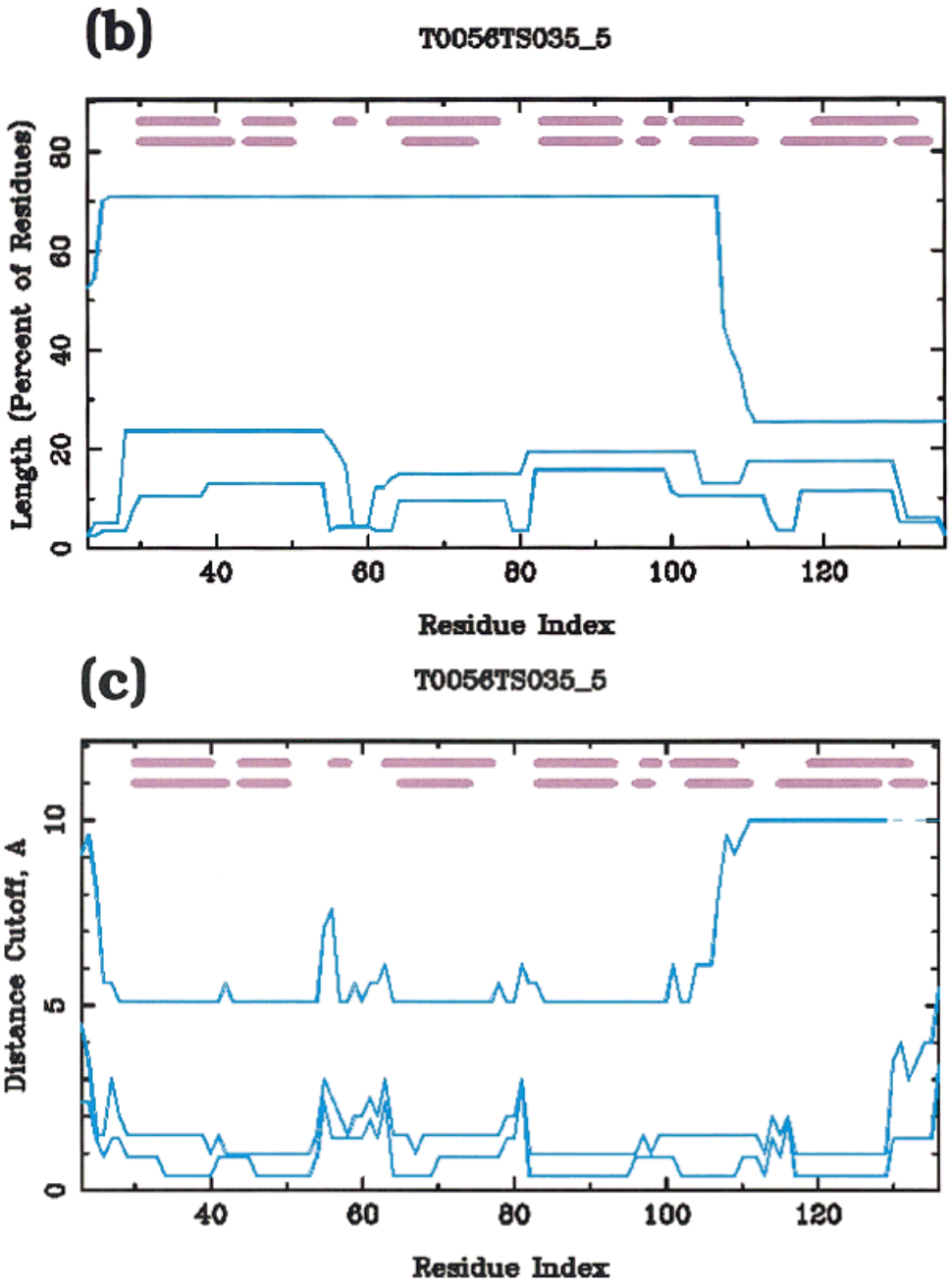


Figure 2. (Continued.)

"SHIFTED SS UNITS" Secondary structure elements that have significantly moved relative to the parent structure (by more than 1 Å).

"LARGE SHIFTS/INSERTIONS" "Loop" segments. In a global alignment (DALI) between structures of target and parent, residues with corresponding Cα distances greater than cutoff (2.5 Å) are included. If fewer than three residues exist between such segments, they are included in a merged segment.

"CORE" This subset is complementary to "LARGE SHIFTS/INSERTIONS."

"LIGAND CONTACTS" Regions of structure that are in contact with the ligand molecule(s). A cutoff of 6 Å defines protein neighborhood for local model/target structural alignment. Subsequently, protein atoms in contact with ligand (4 Å cutoff) are included in this subset.

### *RMSD details of loops*

To specifically address modeling performance on individual loops, Cartesian RMSD in both global and local superposition is calculated on Cα, main chain, and all atoms for each loop that contains at least three residues.

The loops are defined as above ("LARGE SHIFTS/INSERTIONS"). For a given loop, the completeness of a prediction is also reported (NP, TN, PP).

### *Model refinement*

To evaluate the success of refinement procedures, such as energy minimization or molecular dynamics, parallel submissions of refined models are accepted and fully assessed. Relative change in the evaluated parameters may be easily monitored in the prediction database.

### *Evaluation of confidence assessments*

Estimation of a position-specific reliability is an important factor contributing to the potential usefulness of a model, especially in comparative modeling. It is unfortunate that this aspect of prediction often is omitted. Submission formats encourage confidence estimation at atomic level with results reported for Cα's, main chain, side chain, and all atoms. The measures used to evaluate accuracy of error estimates are:

(a) D-E—mean deviation between actual distance in atomic positions (D) and the corresponding estimate (E), between model and target in Angstroms;
(b) D-E / D+E—mean value of the normalized deviation.

Exact definitions are given on the Prediction Center web page.

## New Analytical Tools and Graphical Representations

### *Alignment accuracy*

Evaluation of alignment accuracy has been redefined for CASP3. Previously, predictors' defined alignments of target sequence with a template structure were used to assess the correctness of alignment, both in comparative modeling and in fold recognition. With abandoning of specific prediction categories and corresponding simplification of submission formats, alignment evaluation was based on direct comparison between model and target structures. The added benefit is that comparisons with homologous structures do not need to be invoked. Furthermore, alignment accuracy defined in such a way is applicable to all 3D predictions, including models generated ab initio.

New alignment accuracy measures are based on the lowest RMSD sequence-independent superposition currently generated by the DALI server. For each residue in the model structure the closest residue in the target is identified. The number of correctly aligned residues is defined as the number of residues in the model for which the closest residue in the target is the correct one, and the distance between them is less than 3.8 Angstroms (Cα-Cα distances). The number of positions for which the model residue is closest to a residue in the target within +/-4 residues, and the distance is less than 3.8 Angstroms, is also reported (Fig. 1).

Such defined alignment accuracy is also an excellent general measure of prediction quality, applicable in a wide range of prediction difficulty, ranging from comparative modeling to ab initio targets (cf. discussion in Venclovas et al., this volume[8]).

### *RASMOL plots*

We have used the RASMOL package (written by Roger Sayle) to couple an approximate representation of prediction quality with 3D visualization of model and target structures. Superpositions are made in both sequence-dependent and sequence-independent (DALI) manner. Displayed Cα-trace structures may be colored by prediction quality with a user-specified cutoff (Fig. 2a).

### *Šali plots*

Plots directly comparing deviations between target and model with deviations between target and principal parent proved to be a very useful estimate of prediction quality in comparative modeling. First suggested in the CASP process by Andrej Šali,[9] they illustrate how a particular prediction fares when compared with a "straightforward" copying from the nearest homologous structure. It has to be emphasized that correct copying of structure can only be achieved after a correct alignment with the parent has been generated, and this remains a significant difficulty, even in comparative modeling. We generate Šali plots for predictions on all comparative modeling targets.

## Identifying Longest Continuous Superimposable Segments of Residues (LCS)

The next two measures are to some extent complementary and have been designed to facilitate the detection of good and bad regions in the modeled structure. Our LCS algorithm identifies all the continuous segments of residues in the prediction deviating from the target by no more than a specified Cα RMSD cutoff. For a given residue, we consider all segments containing that residue and assign it to the longest one. The measure may be used to evaluate all 3D predictions. For CASP3 we introduced graphical rendering of this data as a function of residue position along the main chain and plotting segment lengths as percentage of target structure. The plots provide an intuitive representation of sets of residues in the model that are characterized by similar spatial shifts relative to the target structure, typically reflecting the elements of secondary structure in that protein. Taken together with information on chain topology, they provide a reasonably complete representation of the quality of a prediction (Fig. 2b).

### *Identifying Largest Superimposable Sets of Residues (GDT)*

For each residue in the prediction our algorithm identifies the largest set of residues containing that residue and deviating from the target by no more than a specified Cα distance cutoff (Global Distance Test). In comparison with LCS, which provides numerically exact results, generation of maximal sets that are not necessarily continuous along the main chain is only approximate. The algorithm however uses a very large number of different superpositions providing consistently reliable results. Details of the implementation are described at the Prediction Center website. As with LCS, this measure may be used to evaluate any 3D prediction, regardless of the extent of relatedness with a

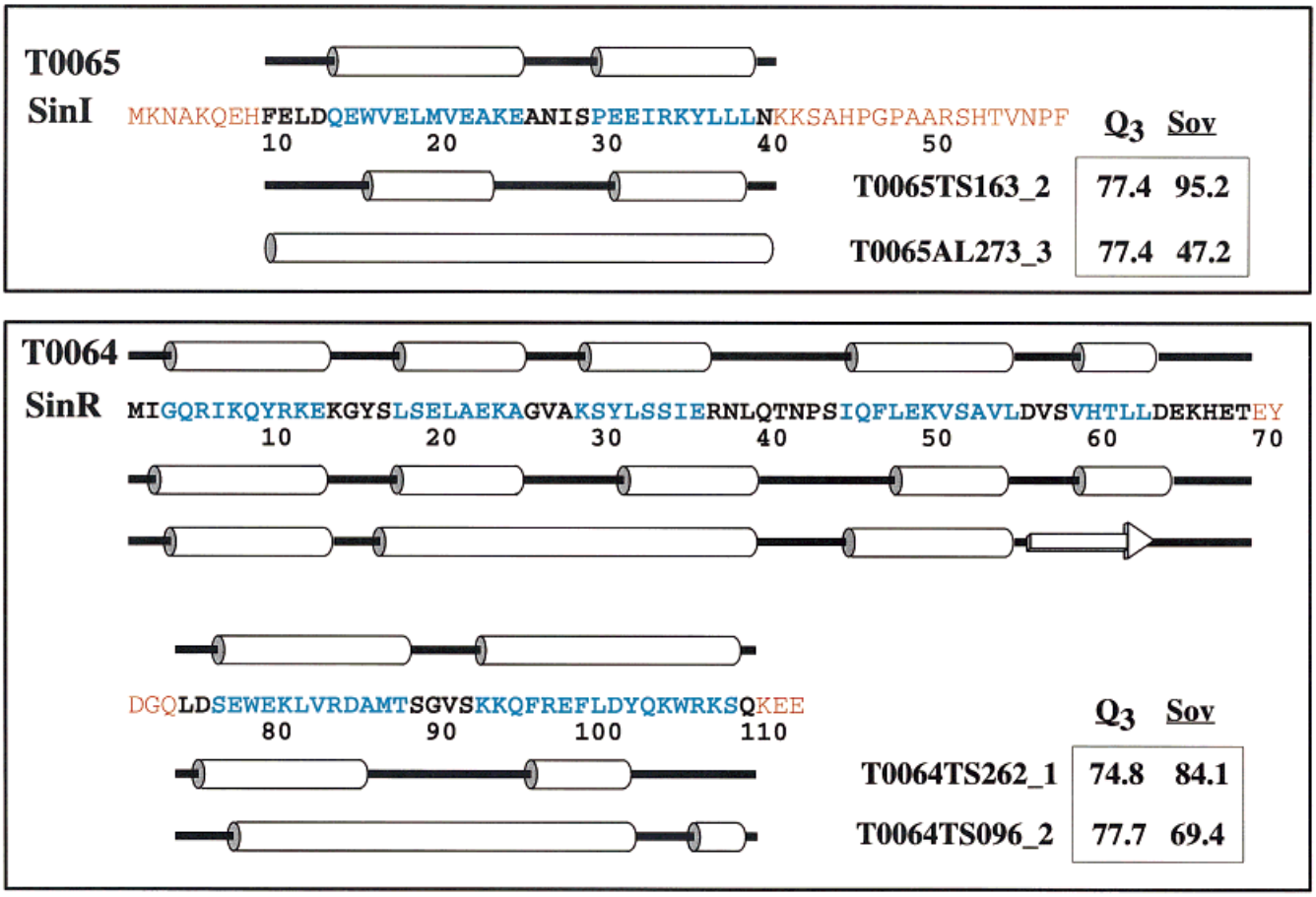


Fig. 3. A comparison of DSSP-generated secondary structure assignments for Bacillus subtilis SinI and SinR (target structures T0065 and T0064), and several predictions submitted during CASP3. Both Q3 and Sov scores are reported. Results for sequence fragments for which crystallographic data is missing could not be compared. In the case of SinI, Q3 assigns equal scores to an essentially correct and a structurally misleading prediction. In SinR, prediction which correctly identifies and places all of the secondary structure elements is scored by Q3 lower than the assignment which predicts helices across joining loops and in one case a strand instead of a helix. Assignments of secondary structure in the four examples shown have been extracted from predictions of 3D structure. These examples do not represent extremes in prediction quality—both better and worse predictions have been submitted on these two targets. Prediction groups 96 and 163 used ab initio, while groups 262 and 273 used fold recognition methods.

known structure. Having generated the largest superimposable sets for all the residues in the prediction, a variety of result presentation schemes may be considered. At present we provide summaries for all predictions on a given target with plots of distance cutoffs versus the set size. For a particular prediction, an option specifying the distance cutoff under which 5, 10, and 50 percent of the modeled structure can be fitted is also used for graphical representations (Fig. 2c).

### *Evaluations of Secondary Structure Predictions*

We calculate scores according to both Q3 and Sov,[10] and also report fractions of the target sequence submitted in each prediction. Q3 is a simple and probably the most commonly used measure for the evaluation of secondary structure predictions. With Q3, evaluations are made residue-by-residue, often providing a misleading assessment of the predicted secondary structure segments. Rather than making a per residue assessment of conformational state, Sov places the emphasis on correct prediction of secondary structure segments, i.e., their type and location. We have evaluated and discussed the performance of Q3 and Sov using both the designed test cases as well as complete secondary structure prediction data accumulated during CASP2.[10,11] Here we demonstrate the difference between the two measures using several examples of predictions taken from CASP3 (Fig. 3).

### Organization of the Website

Our website, located at http://PredictionCenter.llnl.gov, provides comprehensive access to prediction targets, files containing original predictions, and evaluations made using the criteria and methods discussed in this paper. The data are available via an adjustable interface that allows generation of user-defined comparison tables. Result tables may be specified by selecting targets, prediction groups, and measures to be displayed. The site also allows access to visualization tools described above. Data for all three

CASP prediction experiments are available, including earlier CASP predictions evaluated with CASP3 criteria. For comparison and methods development purposes, it is possible to submit and evaluate models on all targets that were part of the CASP process to date.

The website provides four main modes of access to evaluation data:

1. User-defined comparison tables with links to graphical representations. This option allows specification of:
   A) Prediction experiment (CASP1, 2, or 3)
   B) Category of prediction data
      i) 3D models (Ab Initio and Fold Recognition)
      ii) Extended evaluation for Comparative Modeling
      iii) Evaluation of Secondary Structure predictions
   C) Targets and Prediction Groups
   D) Evaluation criteria and subsets of structure

   Generated results tables then provide links to graphical tools as follows:
   A) Plots of longest continuous segments (LCS) for each prediction
   B) Plots of largest superpositions (GDT) between predicted and target structures
   C) Šali plots
   D) RASMOL plots of predicted and target structures in
      i) all Cα sequence-dependent superposition
      ii) iterative sequence-dependent superposition
      iii) DALI sequence-independent superposition
   E) RASMOL plots of superimposed target and parent structures (for Comparative Modeling targets)
   F) Comparisons of residue by residue secondary structure assignments between predicted and target structures.

   Links to summary plots for a given target (see 3. below) are also provided.

2. Tables with default evaluation data for each prediction.

   To simplify navigation through evaluation results, a default set of measures and subsets of structure are used to generate the comparison tables.

3. Summary graphics organized by prediction target.

   This option allows quick comparisons of all predictions submitted on a given target structure. Three types of graphs are presently available:
   A) Alignment accuracy plots
   B) GDT plots
   C) Sov and Q3.

4. Links to results generated by other evaluation methods.
   A) Structural alignment summary data generated with the ProSup software (sequence-independent superposition,[12]) by Manfred Sippl's group.
   B) RMSD versus coverage plots generated by Tim Hubbard.

## DISCUSSION

The CASP process has turned out to be a strong catalyst for the development of techniques for prediction evaluation. A number of new and specialized methods have been suggested, developed, and tested in practice. These have greatly enhanced our ability to extract specific information from predictions. Ultimately, a set of selected, agreed upon, and to a large extent automatic evaluation tools are sought**.** In the discussion, we focus primarily on one particular aspect of the evaluation methods: the superposition techniques behind the large majority of evaluative approaches in structure prediction.

To effectively guide the development of new and better prediction techniques, we need assessment tools applicable to a wide range of prediction problems. While evaluation issues in comparative modeling seem to be adequately addressed with presently available methods, fold recognition and ab initio are less finalized. In general, current evaluation methods tend to work worse as the dissimilarity between target and predicted structures increases. The most radical example is when only the general architecture of the target structure is correctly identified. In CASP3, both threading and ab initio techniques were sometimes successful in this respect, while failing in others. To date, there are no automatic procedures that can reliably classify structures in terms of architecture.

For difficult targets, the question to ask is whether at least part of the structure is predicted correctly. One answer is to map out all possible areas of local similarity. The LCS method accomplishes this goal in terms of segments identified along the main chain of the target structure. An important property of this measure is that the answers are by definition numerically optimal, i.e., no approximations are used to arrive at the results. The shortcoming of this method is that it is able to evaluate the quality of a set of predicted segments only when they are adjacent in sequence. The GDT algorithm, on the other hand, addresses the problem of identifying the largest, not necessarily continuous sets of superimposable residues. Comparing LCS and GDT plots allows one to determine whether correct tertiary associations in the model structure extend beyond regions neighboring in sequence. Both of these algorithms, as well as the one implemented by Tim Hubbard, also discussed in this volume, use sequence-dependent superpositions. In the case for comparative modeling and ab initio predictions, this is usually appropriate. There may be predictions, however, where segments of structure are modeled correctly but the sequence-to-structure alignment is flawed. This is often the case in fold recognition where detecting similar but incorrectly aligned structures requires a sequence-independent model-to-target alignment technique.

Finally, all of the structure superposition methods discussed here so far are of the rigid body type. These, when not carefully applied, are likely to produce misleading results for multi-domain structures, where one domain is shifted relative to another. Also, similarity between mod-

els characterized by a gradual deformation of one structure versus another is not very well described by this approach. In such cases, a method such as SSAP,[13] which scores the similarity of the structural environment at each residue position, may be more effective.

Important aspects of prediction analyses are the methods of displaying the results. Effective graphics are extremely helpful when multiple features of a prediction have to be analyzed or many predictions have to be compared. In this paper we have presented some of the options we think are particularly effective. Other useful ways of displaying evaluation results may be found elsewhere in this volume.

As a community, over the last five years we have moved the methods of prediction evaluation forward considerably. Although a lot still remains to be done, we hope the methods described here and elsewhere in this volume will provide essential data for the next round of discussion regarding evaluation criteria to be held prior to CASP4.

## REFERENCES

1. Zemla A. 1998. AL2TS program: translate sequence-structure alignment (AL) to tertiary structure (TS). Accessed at http://PredictionCenter.llnl.gov/local/al2ts.
2. McLachlan AD. Gene duplications in the structural evolution of chymotrypsin. J Mol Biol 1979;128:49–79.
3. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993;233:123–138.
4. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 1983;22:2577–2637.
5. Lee B, Richards FM. Interpretation of protein structure: estimation of static accessibility. J Mol Biol 55:379–400.
6. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J Mol Biol 1973;79:351–371.
7. Venclovas Č, Zemla A, Fidelis K, Moult J. Criteria for evaluating protein structures derived from comparative modeling. Proteins Suppl 1997;1:7–13.
8. Venclovas Č, Zemla A, Fidelis K, Moult J. Some measures of comparative performance in the three CASPs. Proteins Suppl 1999;3:231–237.
9. Sanchez R, Šali A. Evaluation of comparative protein structure modeling by MODELLER-3. Proteins Suppl 1997;1:50–58.
10. Zemla A, Venclovas Č, Fidelis K, Rost B. A modified definition of Sov, a segment-based measure for protein secondary structure prediction assessment. Proteins 1999;34:220–223.
11. Zemla A, Venclovas Č, Reinhardt A, Fidelis K, Hubbard TJ. Numerical criteria for the evaluation of ab initio predictions of protein structure. Proteins Suppl 1997;1:140–150.
12. Zu-Kang F, Sippl M. Optimum superimposition of protein structures: ambiguities and implications. Folding & Design 1996;1:123–132.
13. Taylor WR, Orengo C. Protein structure alignment. J Mol Biol 1989;208:1–22.