Structural bioinformatics

The ModFOLD server for the quality assessment of protein structural models

Liam J. McGuffin
School of Biological Sciences, University of Reading, Whiteknights, Reading RG6 6AS, UK

Received on November 14, 2007; revised and accepted on January 7, 2008
Advance Access publication January 9, 2008
Associate Editor: Anna Tramontano

1 INTRODUCTION
The selection of the highest quality 3D model of a protein structure from a number of alternatives is a challenging problem in structural biology. Currently, there are two main approaches to the problem of ranking models. Firstly, there are the Model Quality Assessment Programs (MQAPs), which are able to provide a useful score by considering only a single model. Secondly, there are clustering or consensus-based approaches, which attempt to rank a population of alternative models based on pair-wise structural comparisons, and may also include additional information about the confidence of models from fold recognition servers or from sequence alignments. Both techniques are useful in different situations. Clustering-based methods have recently been shown to provide a more accurate assessment of model quality if many models are available from many different resources. However, MQAPs are arguably more useful than clustering when only a few models are available from a few different sources, or if many thousands of models are available, but there is insufficient CPU time available in order to carry out clustering (McGuffin, 2007).

Several MQAP servers exist for carrying out predictions on single models, for example ProQ (Wallner and Elofsson, 2003) and ProSA-web (Wiederstein and Sippl, 2007). Additionally, a few fold recognition meta-servers are now available which carry out clustering or a consensus of scores in order to rank predictions (Kajan and Rychlewski, 2007; Wallner and Elofsson, 2006). However, no servers have been implemented that allow users to accurately compare multiple 3D models of their own choice from any source in terms of predicted quality. The ModFOLD server provides a unified resource for obtaining accurate model quality information based on either individual or multiple uploaded models. The server incorporates two competitive methods: the latest version of the ModFOLD method and a clustering-based method called ModFOLDclust. Through the combination of methods the server can provide: (i) a single score and a P-value relating to the predicted quality of a single 3D model of a protein structure; (ii) rankings for multiple 3D models for the same protein target according to predicted model quality; (iii) predictions of the local quality (per-residue errors) within multiple models.

2 THE ModFOLD METHOD
The original version of the ModFOLD method was recently benchmarked against several widely used top performing MQAPs and clustering-based methods. ModFOLD was shown to be competitive with clustering methods and significantly outperformed the individual MQAPs tested (McGuffin, 2007). The original ModFOLD protocol combined scores obtained from the ModSSEA method (McGuffin, 2007), the MODCHECK method (Pettitt et al., 2005) and the two ProQ methods (Wallner and Elofsson, 2003) using a neural network trained with the TM-score (Zhang and Skolnick, 2004). The latest server implementation of the method has been retrained and includes two additional secondary structure scores, similar to those used by Eramian and colleagues (Eramian et al., 2006), as inputs to the neural network.

The ModFOLD method is an MQAP, which is capable of producing a consistent score based on the analysis of a single model. The consistency of the ModFOLD score allows us to calculate a P-value, which represents a quantitative measure of the confidence in a model. P-values for the ModFOLD server were calculated using a similar approach to that previously adopted for measuring the coverage of genomic scale fold recognition (McGuffin et al., 2004, 2007). Thus, for a given predicted model quality score, the P-value reported by the server is the proportion of models with that score that do not share any similarity with the native structure (TM-score < 0.2).

3 THE ModFOLDCLUST METHOD
The server also includes the option of clustering multiple models using the ModFOLDclust method. The method carries
out pairwise comparisons of models in order to produce both global and local predictions of model accuracy. The global clustering score is based on the 3D-Jury method (Ginalska et al., 2003), whereby each model is compared to every other model and the average structural similarity score is calculated. However, in this application, the TM-score is used for pairwise comparisons, with the cut-off score of >0.2. This emulation of the 3D-Jury score has been previously benchmarked on the set of CASP7 server models and was shown to significantly outperform every method tested for the selection of the highest quality models (McGuffin, 2007). Unlike the 3D-Jury server itself however, where users can evaluate a single model through the comparison with a few available fold recognition models (Kajan and Rychlewski, 2007), the ModFOLD server implementation allows users to directly upload multiple models of their own choice from any source.

In addition to the global clustering score, the ModFOLDclust method incorporates the scoring of local model quality on a per-residue basis. The local model quality is evaluated by using a score similar to the average S-score (Levitt and Gerstein, 1998), which was originally used for model evaluation in the 3D-SHOTGUN method (Fischer, 2003) and was more recently benchmarked using the Pcons server (Wallner and Elofsson, 2006). The idea in this implementation is to reuse each pairwise model superposition, carried out in the calculation of the global score, in order to evaluate the local structural conservation of each residue. Here, the S-score is used to evaluate residues that are within 3.9 Å according to pairwise TM-score superpositions, where the TM-scores >0.2. The S-score is defined as: $S_i = 1/(1 + (d_i/d_0)^2)$, where $S_i$ ranges from 0 to 1, $d_i$ is the distance between structurally aligned residues and $d_0$ is the distance threshold (3.9). An $S_i$ score of 0 is given if $d_i > 3.9$ Å. The S-scores for each residue are then summed and the mean score is taken. The mean S-score for each residue is then converted to the predicted distance from the native structure, by simply rearranging the equation: $d_i = d_0 \sqrt{(1/S_i) - 1}$.

4 SERVER IMPLEMENTATION

The ModFOLD server provides an intuitive web interface for job submission, whereby either single or multiple PDB files may be uploaded for evaluation with a selected method, and the results are returned via email. The results email contains a URL, which links to HTML formatted graphical results, as well as an attached file, which provides results in a machine readable format [CASP7 QMODE1 format for ModFOLD and QMODE2 for ModFOLDclust (Cozzetto et al., 2007)]. The ModFOLD results page simply consists of a table of model data sorted by decreasing predicted model quality score. Each row in the table also includes the calculated $P$-value and the associated colour coded confidence level (e.g. green indicates high confidence in a model, red indicates a random model). The ModFOLDclust results table differs as it is ranked according to the global clustering score. Each row also includes a small JPEG image of a plot depicting the residue error versus the residue number. Each small image links to a page that displays a larger version of the plot and a link to download a PostScript version. Each row in the results table also displays a small 3D cartoon view of the model that is colour coded with the residue error according to the temperature colouring scheme (Figure 1).

AGN. Each small image links to a page that shows a larger image of the 3D view and contains a link to download a PDB file of the model with the predicted residue accuracy (A) in the B-factor column. Figure 1 shows an example of the ModFOLDclust results for a model of CASP7 target T0284 compared with the observed error according to a superposition with the native structure.

ACKNOWLEDGEMENTS

This work was supported by an RCUK Academic Fellowship.

Conflict of Interest: none declared.

REFERENCES


