University of Reading

# *ModFOLD9: a web server for independent estimates of 3D protein model quality*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

McGuffin, L. J. ORCID: https://orcid.org/0000-0003-4501-4767 and Alharbi, S. M. A. (2024) ModFOLD9: a web server for independent estimates of 3D protein model quality. Journal of Molecular Biology, 436 (17). 168531. ISSN 1089-8638 doi: 10.1016/j.jmb.2024.168531 Available at https://centaur.reading.ac.uk/115716/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.jmb.2024.168531

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

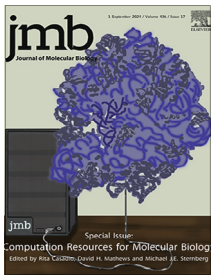
www.reading.ac.uk/centaur

**CentAUR**

University of Reading

Central Archive at the University of Reading

Reading's research outputs online



# ModFOLD9: A Web Server for Independent Estimates of 3D Protein Model Quality

**Liam J. McGuffin** [*,†] **and Shuaa M. A. Alharbi** [†]

***School of Biological Sciences,*** *University of Reading, UK*

***Correspondence to Liam J. McGuffin:*** *l.j.mcguffin@reading.ac.uk (L.J. McGuffin)*
https://doi.org/10.1016/j.jmb.2024.168531
***Edited by Rita Casadio***

## Abstract

Accurate models of protein tertiary structures are now available from numerous advanced prediction methods, although the accuracy of each method often varies depending on the specific protein target. Additionally, many models may still contain significant local errors. Therefore, reliable, independent model quality estimates are essential both for identifying errors and selecting the very best models for further biological investigations. ModFOLD9 is a leading independent server for detecting the local errors in models produced by any method, and it can accurately discriminate between high-quality models from multiple alternative approaches. ModFOLD9 incorporates several new scores from deep learning-based approaches, leading to greatly improved prediction accuracy compared with earlier versions of the server. ModFOLD9 is continuously independently benchmarked, and it is shown to be highly competitive with other public servers. ModFOLD9 is freely available at https://www.reading.ac.uk/bioinf/ModFOLD/.
© **2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).**

## Introduction

In a post-AlphaFold2 world, there are numerous alternative methods for accurately modelling the 3D structures of proteins from amino acid sequences. While this progress in modelling is a major achievement, it is essential to note that many predicted 3D models still contain significant local errors. Here, we describe major updates to ModFOLD (https://www.reading.ac.uk/bioinf/ModFOLD/), which has maintained its position as a leading server for the prediction of global and local quality of 3D protein models over the past decade. ModFOLD9 can successfully detect local modeling errors, including those in high-quality models that are close to experimental structures, such as those from AlphaFold2,[1] RoseTTAFold[2] and ColabFold.[3] Furthermore, ModFOLD9 can be used to discriminate between multiple alternative high-quality 3D models by ranking them in order of global quality based on consistent independent scores. This independent scoring information is essential for the successful utility and application of the very best 3D models for further biological investigations, and it will allow their wider acceptance by the bioscience community. As diverse protein modelling methods continue to emerge, it is vital that general biologists can compare them and build trust in them through the use of freely available unbiased model quality assessment (QA) methods, such as ModFOLD9.

Successive versions of the ModFOLD server have been maintained and made freely available to users since 2008.[4–7] The ModFOLD server itself has served >7,000 unique external users with over 250,000 jobs completed. ModFOLD predictions are also built into the IntFOLD server results, which have served >26,000 unique external users with over >340,000 jobs completed. Each version of the ModFOLD server has been independently blind-tested in the biennial CASP experiments[8–12] as well as the CAMEO project.[13]

0022-2836/© 2024 The Author(s). Published by Elsevier Ltd.This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

ModFOLD9 is the latest version of the server, which combines the strengths of multiple pure-single and quasi-single model methods. As well as updates to the web interface, ModFOLD9 has increased prediction accuracy compared with previous versions, which was achieved through the integration of newly developed scoring methods and advanced deep learning-based approaches. The ModFOLD9 protocol builds on that of ModFOLD8[7] by including 6 new integrated scoring methods: 3 new Contact Distance Agreement (CDA) scores and the 3 variants of the DeepAccNet[14] methods (DeepAccNet, DeepAccNet-Bert and DeepAccNet-MSA). Our CDA scores measure the agreement between the residue contacts predicted from the target sequence and the measured Euclidean distance (in Å) between residues in the predicted 3D model. The contact predictions from trRosetta2,[2] DeepDist[15] and TripletRes[16] were used for the three new CDA scores, CDA_trR2 and CDA_DD and CDA_TR, respectively. As in previous versions of ModFOLD, neural networks were then used to combine the component local quality scores from each of the scoring methods, resulting in a final consensus of per-residue quality scores for each model. For each model submitted in CASP15 experiment and CAMEO benchmark, the predicted per-residue quality scores (plDDT*100) from ModFOLD9 were added to the B-factor column for each set of atom records.

ModFOLD9 played a key part in the high performance of our group in the CASP15 experiment, and it is continuously benchmarked using the CAMEO resource.[13] According to the independent benchmarking results, our new version of ModFOLD shows improved performance over our previous versions of the server, and it is competitive with the other public servers.

## Results and discussion

### ModFOLD9 inputs and outputs

ModFOLD9 requires two inputs: the amino acid sequence of the target protein and a single 3D model in PDB format for evaluation. Additionally, users have the option to upload multiple alternative models as a gzipped tar file, provide name their protein sequence, add their email addresses for the result delivery, and select their global score optimisation preference (see Materials and Methods). Users are required to submit a reference sequence both to ensure a fair comparison for all submitted models using the same input data calculated from the sequence and to maintain consistent residue numbering for all models. The analysis time using ModFOLD9 depends on factors such as the required sequence length, the number of submitted models, and the server's capacity at the time of submission. Generally, users receive results within the same day for a new run on a single model, but for larger proteins, or if there are several hundred alternative models for a single target, then the analysis time may exceed several days. However, if a model for the same target has already been submitted within the same week, then its reference model library will be available to the server, and so the results can be delivered within a few hours.

The ModFOLD9 web server offers a user-friendly interface for straightforward interaction. Figure 1A shows a screenshot of the main results page as a summary table of quality scores for each submitted model, along with plots indicating local errors and annotated 3D model images. Each row in the table represents a specific model with estimation results, including the rank and ID of the model, overall quality scores, a confidence score and P-value, and thumbnails of the graphical results. The error plots and images of 3D models are colour-coded to highlight local quality, enabling users to easily identify regions of high and low accuracy in the model. The main results page provides push-buttons to show these graphics in full size on separate pages, allowing users to compare plots and images separately for the best visualisation. The per-residue plots, shown in Figure 1B, can be downloaded in PDF format, and users can view 'b-factor' annotated models interactively in 3D directly within the browser (Figure 1C), providing a tangible understanding of the quality estimation results. Additionally, users have the option to download raw, machine-readable files encapsulating the quality estimation data in CASP format and compressed archives for all annotated models, facilitating storage, accessibility, and further independent analysis.

### Independent benchmarking and cross-validation

***Performance on the CASP14 and CASP15 model data.*** ModFOLD9 was initially trained and cross-validated on the CASP14 model data, and subsequently, it was used to assess models during the CASP15 prediction season. Post CASP15, the method was also 3-fold cross-validated using the CASP15 model data. The data in Table 1 shows ModFOLD9 outperforming all other tested methods on the CASP15 model data in terms of the predicted local lDDT scores, while the retrained version demonstrates further improvements. Additional CASP14 and CASP15 performance data are shown in Supplementary Figures 1 and 2, and Supplementary Tables 1–3. The data show that ModFOLD9 outperforms all other tested approaches in terms of its local scores measured by both the lDDT scores and the superposition-dependent S-scores on both the CASP14 and CASP15 model data (Table 1,

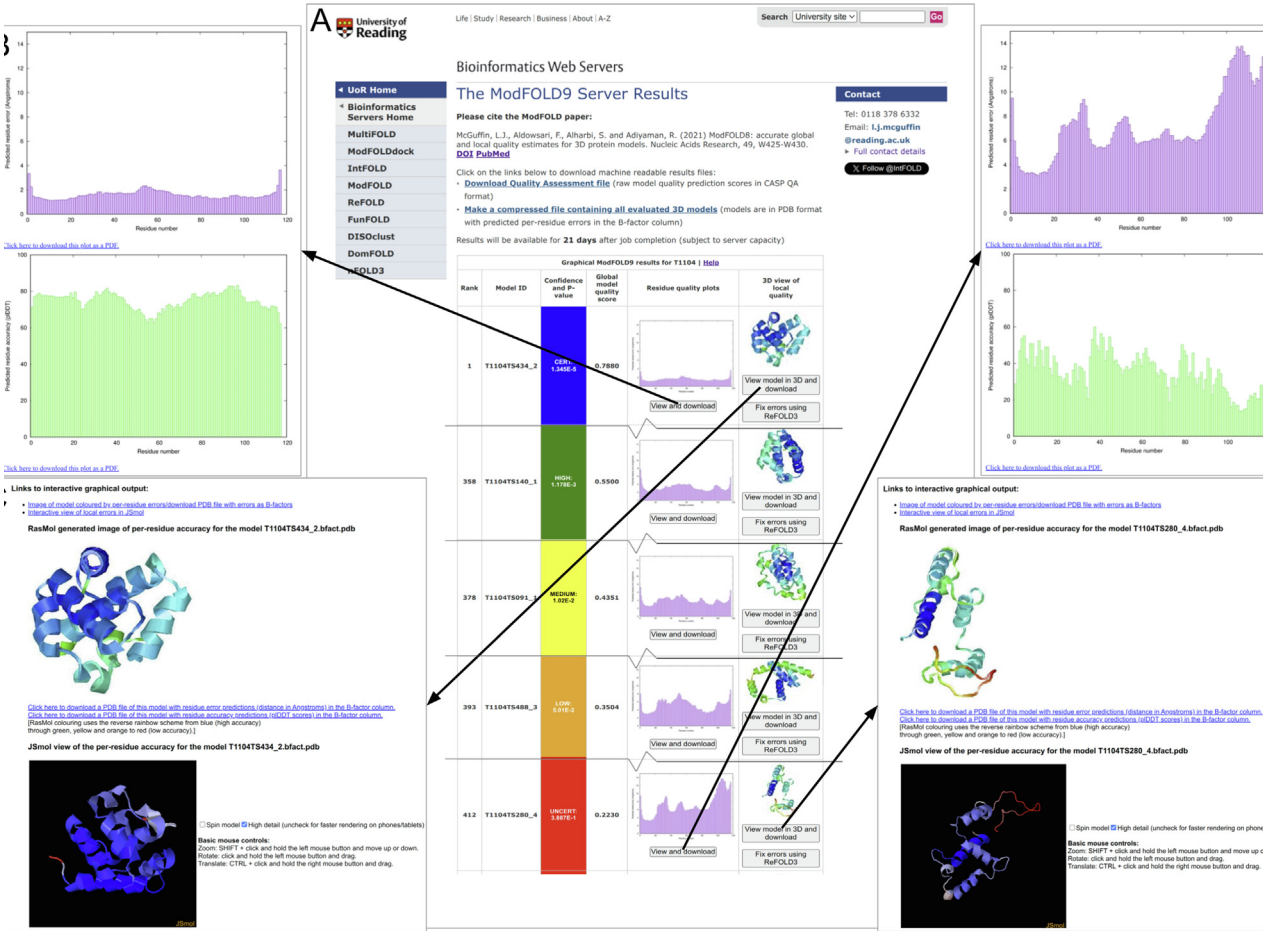



**Figure 1.** ModFOLD9 server results for the CASP15 target T1104. (A) The main results page shows a summary of the graphical output for each model (the table is truncated to show a range of model quality data). The arrows point to additional graphical results that are accessed via buttons on the main page. The 'Fix errors using ReFOLD3' buttons allow users to submit their 3D models to the ReFOLD server (19) for refinement guided by the local quality scores. (B) An example of the plots for a high-accuracy model showing the per-residue errors (predicted distance in Å of each Cα atom from the native structure) and accuracy scores (predicted lDDT scores) in the model, which can be downloaded as PDFs. (C) The interactive JSmol view of the high-accuracy model. Users can also download their models in PDB format with the predicted residue errors shown in the b-factor column. (D) and (E) show examples of plots and interactive output for an example of a low-quality model.

Supplementary Tables 1–3). Furthermore, the data demonstrate that models can be accurately assessed by the ModFOLD9 global scores, both in terms of their absolute score correlations with observed scores and when they are used to rank models by their relative quality (Supplementary Figures 1 and 2). Indeed, in terms of the Pearson correlation coefficients (Supplementary Figure 2), the ModFOLD9_cor global scores correlate better with the observed GDT_HA, GDT_TS, MaxSub and TM-scores (the R values are 0.8112, 0.8328, 0.8237, and 0.8294, respectively) than the observed lDDT scores correlate with the observed GDT_HA, GDT_TS, MaxSub and TM-scores (the R values are 0.78560, 0.7754, 0.7753, and 0.7527, respectively).

In CASP15, each participating group was required to provide their own Accuracy Self Estimates (ASE) in the "b-factor" column for each generated model file (e.g., the plDDT scores generated by AlphaFold2 and other methods). Table 1 and Supplementary Figure 2 show the performance that can be gained by considering the local and global ASE scores contained within each model file. The data show that the local ASE scores are outperformed by ModFOLD9 and many other QA methods according to all measures (Table 1). Likewise, the global ASE scores are greatly outperformed by ModFOLD9 and other QA methods (Supplementary Figure 2).

While the individual ASE local scores from different methods can be compared to some

Table 1 The performance of ModFOLD9 local model quality estimates on CASP15 models (65 targets, 29,915 models, 7,926,950 residues). The ModFOLD9 predicted local lDDT scores versus the observed local lDDT scores. AUC = Area Under the ROC Curve. AUC 0–0.1 = Area Under the ROC curve with False Positive Rate $\leq$ 0.1. A true positive (TP) is defined as a residue correctly identified to be low quality, with local lDDT <= 0.6. The table is sorted by the AUC score. ModFOLD9 – the original full version trained on CASP14 data, ModFOLD9_pure – only uses inputs from the pure-single model methods, ModFOLD9_retrained – the latest full version cross-validated using CASP15 data.

| **Evaluated on lDDT** | Pearson v lDDT | Spearman v lDDT | AUC (TP lDDT <= 0.6) | AUC 0.0–0.1 (TP lDDT <= 0.6) |
|---|---|---|---|---|
| ModFOLD9_retrained | **0.7800** | **0.6599** | **0.9107** | **0.0501** |
| ModFOLD9 | **0.7252** | **0.6113** | **0.8870** | **0.0443** |
| ModFOLD9_pure | **0.7340** | **0.5916** | **0.8702** | **0.0307** |
| DeepAccNet | 0.6738 | 0.5886 | 0.8591 | 0.0338 |
| DeepAccNet_MSA | 0.6807 | 0.5541 | 0.8441 | 0.0247 |
| DeepAccNet_Bert | 0.6467 | 0.5666 | 0.8385 | 0.0273 |
| ProQ2D | 0.5335 | 0.4701 | 0.8141 | 0.0316 |
| ProQ2 | 0.5471 | 0.4671 | 0.8094 | 0.0309 |
| ProQ3D | 0.5433 | 0.4963 | 0.8087 | 0.0304 |
| ProQ4 | 0.5130 | 0.4693 | 0.8087 | 0.0365 |
| ASE | 0.4592 | 0.5296 | 0.7742 | 0.0066 |
| CDA_SC | 0.3750 | 0.3325 | 0.7492 | 0.0201 |
| ModFOLD5_single | 0.3874 | 0.4088 | 0.7467 | 0.0276 |
| VoroMQA | 0.4080 | 0.3860 | 0.7411 | 0.0213 |
| ModFOLDclustQ_single | 0.3821 | 0.3951 | 0.7390 | 0.0278 |
| DBA | 0.3569 | 0.3622 | 0.7134 | 0.0210 |
| CDA_trR2 | 0.3603 | 0.3490 | 0.7111 | 0.0266 |
| ResQ | 0.3004 | 0.2782 | 0.6752 | 0.0239 |
| CDA_TR | 0.2345 | 0.2215 | 0.6639 | 0.0201 |
| CDA | 0.2492 | 0.2694 | 0.6418 | 0.0189 |
| CDA_DMP | 0.1889 | 0.1495 | 0.6404 | 0.0063 |
| SSA | 0.1523 | 0.1394 | 0.5987 | 0.0098 |
| CDA_DD | 0.0546 | 0.0476 | 0.5501 | 0.0037 |

extent, they are inconsistent. If the ASE scores are considered in a global context, they are shown to correlate quite poorly with the observed scores and are less effective than using standard QA methods for ranking. Therefore, independent model quality estimates are superior for comparing multiple models from different methods. The scores from ModFOLD9 are more consistent and can be used to directly compare models ranging in quality from a variety of different methods more accurately.

***CAMEO QE (Quality Estimates) performance.*** ModFOLD9 is continuously benchmarked using the independent CAMEO resource.[13] Figure 2 shows ROC analysis of the recent performance of ModFOLD9 according to the CAMEO QE benchmarking data compared with other methods on common subsets. At the time of writing, the CAMEO QE data show that ModFOLD9 is the leading public QA method for producing local (per-residue) quality scores, according to the lDDT scores, outperforming all other public methods when compared on common sets of models over the 1 week, 1 month, 3 month and 6 month time frames (Supplementary Table 4, data obtained on 02/12/23). Further ROC data are shown in Supplementary Figures 3–5 and Supplementary Table 5, which also show that ModFOLD9 greatly outperforms our previous versions of ModFOLD. These data further demonstrate that ModFOLD9 can accurately estimate the quality of models, offering an unbiased comparison of models from different methods regardless of the modelling approach used.

## Materials and methods

Our principal aim with the ModFOLD9 development was to further increase the predictive accuracy of per-residue quality estimates, given the prevalence of localised errors in otherwise high-quality models. The ModFOLD9 server was built on the strengths of our previously successful versions, which included various scores from different pure- and quasi-single model quality methods.[7]

The pure-single model methods were the ProQ family of methods, ProQ2,[17] ProQ2D,[18] ProQ3D,[18] and ProQ4,[12] in addition to the VoroMQA method,[19] the Secondary Structure Agreement (SSA) score and three different Contact Distance Agreement (CDA) scores.[6,10] The established CDA scores were CDA_DMP CDA_SC, which were computed according to DeepMetaPSICOV[20] and SPOT-Contact,[21] respectively, in addition to the original CDA score derived from MetaPSICOV.[22] Building on these scores, six new scores were added. Three scores were computed using the DeepAccNet variants: DeepAccNet, DeepAccNet_Bert and DeepAccNet_MSA.[14] The other three new scores

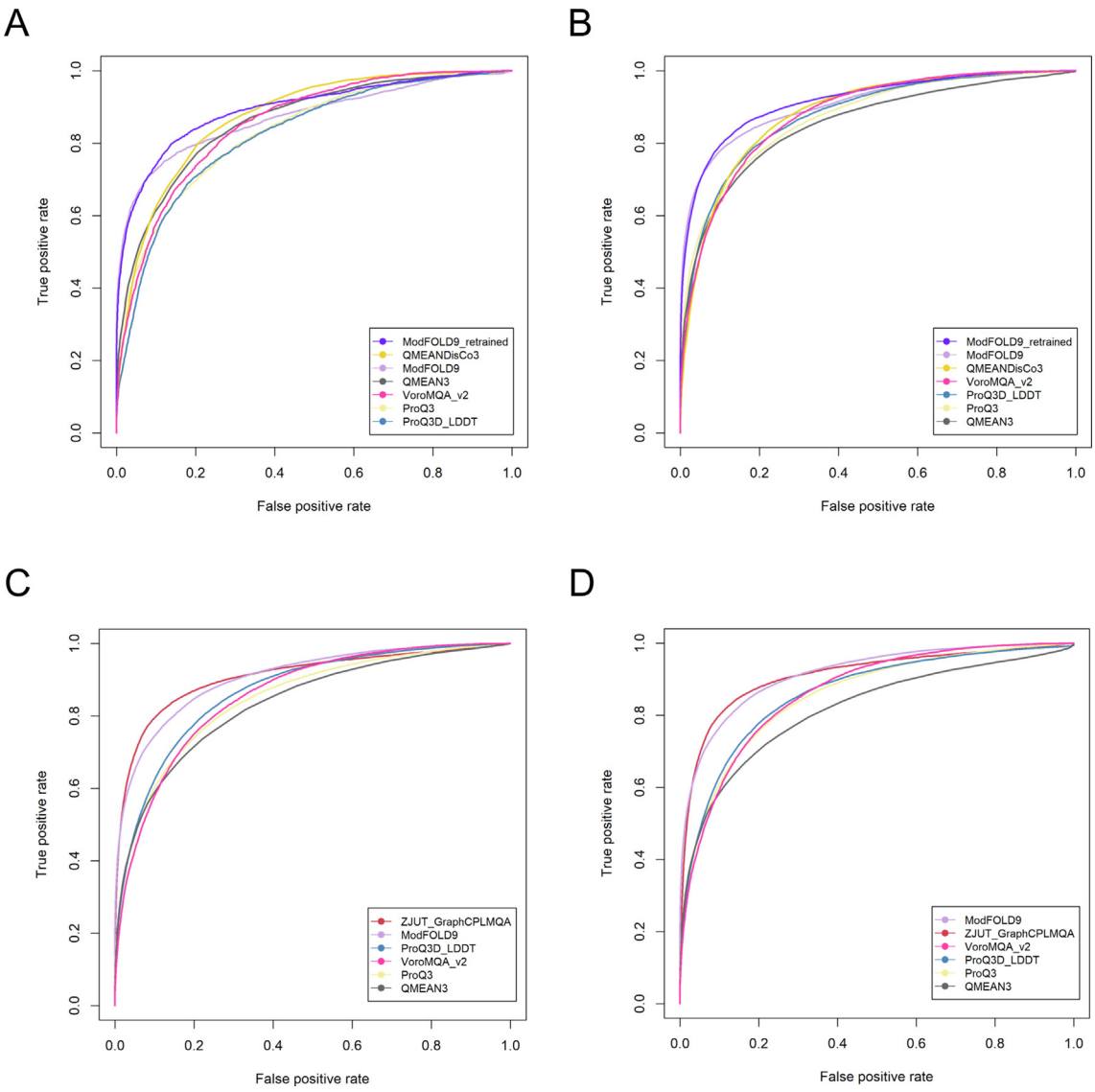



**Figure 2.** The performance of ModFOLD9 local model quality estimates on CAMEO independent benchmarking data – a comparison with the top 5 publicly available methods for each time frame. ROC analysis using common subsets of CAMEO quality estimates (QE) data up until 02/12/23. A true positive is defined as a residue correctly identified to be low quality, with local lDDT <= 60. Full ROC plots are shown for common subsets, including data for the top 5 publicly available methods at the time of writing. (A) 1 week of data, 95 targets, 665 models. (B) 1 month of data, 393 targets, 2751 models. (C) 3 months of data, 577 targets, 3462 models. (D) 6 months of data, 699 targets, and 4194 models.

were CDA_TR, CDA_trR2 and CDA_DD. These new CDA scores were computed according to three deep-learning contact prediction methods: TripletRes,[16] trRosetta2,[2] and DeepDist,[15] respectively.

The quasi-single model methods involved ResQ,[23] Disorder B-factor Agreement (DBA), ModFOLDclust_single (MF5s) and ModFOLDclustQ_single (MFcQs).[6,10] Our group developed the latter three methods, which assess the single model in the context of multiple reference models from tertiary structure prediction servers. ResQ compares the single model to reference models predicted from the LOMETS method,[24] whereas our three scores used the IntFOLD7 server to generate 135 reference models to perform the comparison approach.[25]

The scores were fed as inputs into two versions of a neural network, implemented using the MLP (multi-layer perceptron) from the RSNNS package

to predict the local quality scores. The input of each MLP version contains 90 neurons with a sliding window size of five from 18 quality scores for each residue (5*18 = 90). The two MLP versions were trained to output the predicted residue quality score, either based on the S-score[6,10] or the lDDT score.[26] The MLP was trained and tested on both the CASP14 and CASP15 model data using a three-fold cross-validation procedure.

As with our previous versions of ModFOLD, we produced three variants of ModFOLD9, each optimised for the different facets of the quality estimation problem. Firstly, ModFOLD9_rank was developed to provide an optimised ranking of models, with the highest-ranked models being closer to the highest observed accuracy. However, the relationship between the predicted and observed scores may not be linear. Secondly, ModFOLD9_cor was optimised to produce correlations with observed scores that were closer to linear. Finally, ModFOLD9 provides balanced performance in terms of both correlations of predicted and observed scores and rankings of top models.

ModFOLD9 was used as the self-estimation method for IntFOLD7,[25] our integrated server for modelling tertiary protein structures and functions. The IntFOLD7 tertiary structure predictions were blind tested during the recent CASP15 experiment. ModFOLD9 also helped us greatly in the CASP15 experiment with our manual (McGuffin group) predictions of regular targets. As a result, we were the top-performing academic group from the UK and we ranked 6th out of all groups on regular targets (https://predictioncenter.org/casp15/zscores_final.cgi). Moreover, both IntFOLD7 and ModFOLD9 are continuously benchmarked using the CAMEO resource.[13] According to the 3D and QE benchmark results, both of our new servers show improved performance over our previous versions of those methods, and they are competitive with the other public servers in their respective categories.

## CRediT authorship contribution statement

**Liam J. McGuffin:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Shuaa M.A. Alharbi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

## DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/T018496/1 to Liam J. McGuffin) and the Saudi Arabia Cultural Bureau (to Shuaa M. A. Alharbi).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2024.168531.

*Received 21 January 2024;*
*Accepted 6 March 2024;*
*Available online 11 March 2024*

***Keywords***:
protein structure prediction;
model quality assessment (QA);
model quality estimates (QE);
estimates of model accuracy (EMA);
structural bioinformatics

† Both authors contributed equally to this work as first authors.

## References

1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., et al., (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
2. Anishchenko, I., Baek, M., Park, H., Hiranuma, N., Kim, D. E., Dauparas, J., Mansoor, S., Humphreys, I.R., et al., (2021). Protein tertiary structure prediction and refinement using deep learning and Rosetta in CASP14. *Proteins* **89**, 1722–1733. https://doi.org/10.1002/prot.26194.
3. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M., (2022). ColabFold: making protein folding accessible to all. *Nature Methods* **19**, 679–682. https://doi.org/10.1038/s41592-022-01488-1.
4. McGuffin, L.J., (2008). The ModFOLD server for the quality assessment of protein structural models. *Bioinformatics* **24**, 586–587. https://doi.org/10.1093/bioinformatics/btn014.
5. McGuffin, L.J., Buenavista, M.T., Roche, D.B., (2013). The ModFOLD4 server for the quality assessment of 3D protein models. *Nucl. Acids Res.* **41**, W368–W372. https://doi.org/10.1093/nar/gkt294.

6. Maghrabi, A.H.A., McGuffin, L.J., (2017). ModFOLD6: an accurate web server for the global and local quality estimation of 3D protein models. *Nucl. Acids Res.* **45**, W416–W421. https://doi.org/10.1093/nar/gkx332.
7. McGuffin, L.J., Aldowsari, F.M.F., Alharbi, S.M.A., R., (2021). Adiyaman ModFOLD8: accurate global and local quality estimates for 3D protein models. *Nucl. Acids Res.* **49**, W425–W430. https://doi.org/10.1093/nar/gkab321.
8. McGuffin, L.J., (2009). Prediction of global and local model quality in CASP8 using the ModFOLD server. *Proteins* **77**, 185–190. https://doi.org/10.1002/prot.22491.
9. McGuffin, L.J., Roche, D.B., (2011). Automated tertiary structure prediction with accurate local model quality assessment using the Intfold-TS method. *Proteins* **79**, 137–146. https://doi.org/10.1002/prot.23120.
10. McGuffin, L.J., Shuid, A.N., Kempster, R., Maghrabi, A.H. A., Nealon, J.O., Salehe, B.R., Atkins, J.D., Roche, D.B., (2018). Accurate template-based modeling in CASP12 using the IntFOLD4-TS, ModFOLD6, and ReFOLD methods. *Proteins* **86**, 335–344. https://doi.org/10.1002/prot.25360.
11. Elofsson, A., Joo, K., Keasar, C., Lee, J., Maghrabi, A.H.A., Manavalan, B., McGuffin, L.J., Ménendez Hurtado, D., et al., (2018). Methods for estimation of model accuracy in CASP12. *Proteins* **86**, 361–373. https://doi.org/10.1002/prot.25395.
12. Cheng, J., Choe, M., Elofsson, A., Han, K., Hou, J., Maghrabi, A.H.A., McGuffin, L.J., Menéndez-Hurtado, D., et al., (2019). Estimation of model accuracy in CASP13. *Proteins* **87**, 1361–1377. https://doi.org/10.1002/prot.25767.
13. Robin, X., Haas, J., Gumienny, R., Smolinski, A., Tauriello, G., Schwede, T., (2021). Continuous Automated Model EvaluatiOn (CAMEO)—Perspectives on the future of fully automated evaluation of structure prediction methods. *Proteins* **89**, 1977–1986. https://doi.org/10.1002/prot.26213.
14. Hiranuma, N., Park, H., Baek, M., Anishchenko, I., Dauparas, J., Baker, D., (2021). Improved protein structure refinement guided by deep learning based accuracy estimation. *Nature Commun.* **12**, 1340. https://doi.org/10.1038/s41467-021-21511-x.
15. Wu, T., Guo, Z., Hou, J., Cheng, J., (2021). DeepDist: real-value inter-residue distance prediction with deep residual convolutional network. *BMC Bioinformat.* **22**, 30. https://doi.org/10.1186/s12859-021-03960-9.
16. Li, Y., Zhang, C., Bell, E.W., Zheng, W., Zhou, X., Yu, D.-J., Zhang, Y., (2021). Deducing high-accuracy protein contact-maps from a triplet of coevolutionary matrices through deep residual convolutional networks. *PLoS Comput. Biol.* **17**, e1008865. https://doi.org/10.1371/journal.pcbi.1008865.
17. Ray, A., Lindahl, E., Wallner, B., (2012). Improved model quality assessment using ProQ2. *BMC Bioinformat.* **13**, 224. https://doi.org/10.1186/1471-2105-13-224.
18. Uziela, K., Menéndez Hurtado, D., Shu, N., Wallner, B., Elofsson, A., (2017). ProQ3D: improved model quality assessments using deep learning. *Bioinformatics* **33**, 1578–1580. https://doi.org/10.1093/bioinformatics/btw819.
19. Olechnovič, K., Venclovas, Č., (2017). VoroMQA: assessment of protein structure quality using interatomic contact areas: contact area-based protein structure assessment. *Proteins* **85**, 1131–1145. https://doi.org/10.1002/prot.25278.
20. Kandathil, S.M., Greener, J.G., Jones, D.T., (2019). Prediction of interresidue contacts with DeepMetaPSICOV in CASP13. *Proteins* **87**, 1092–1099. https://doi.org/10.1002/prot.25779.
21. Hanson, J., Paliwal, K., Litfin, T., Yang, Y., Zhou, Y., (2018). Accurate prediction of protein contact maps by coupling residual two-dimensional bidirectional long short-term memory with convolutional neural networks. *Bioinformatics* **34**, 4039–4045. https://doi.org/10.1093/bioinformatics/bty481.
22. Jones, D.T., Singh, T., Kosciolek, T., Tetchner, S., (2015). MetaPSICOV: combining coevolution methods for accurate prediction of contacts and long range hydrogen bonding in proteins. *Bioinformatics.* **31**, 999–1006. https://doi.org/10.1093/bioinformatics/btu791.
23. Yang, J., Wang, Y., Zhang, Y., (2016). ResQ: An approach to unified estimation of B-factor and residue-specific error in protein structure prediction. *J. Mol. Biol.* **428**, 693–701. https://doi.org/10.1016/j.jmb.2015.09.024.
24. Wu, S., Zhang, Y., (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucl. Acids Res.* **35**, 3375–3382. https://doi.org/10.1093/nar/gkm251.
25. McGuffin, L.J., Edmunds, N.S., Genc, A.G., Alharbi, S.M. A., Salehe, B.R., Adiyaman, R., (2023). Prediction of protein structures, functions and interactions using the IntFOLD7, MultiFOLD and ModFOLDdock servers. *Nucl. Acids Res.*. https://doi.org/10.1093/nar/gkad297 gkad297.
26. Mariani, V., Biasini, M., Barbato, A., Schwede, T., (2013). lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728. https://doi.org/10.1093/bioinformatics/btt473.