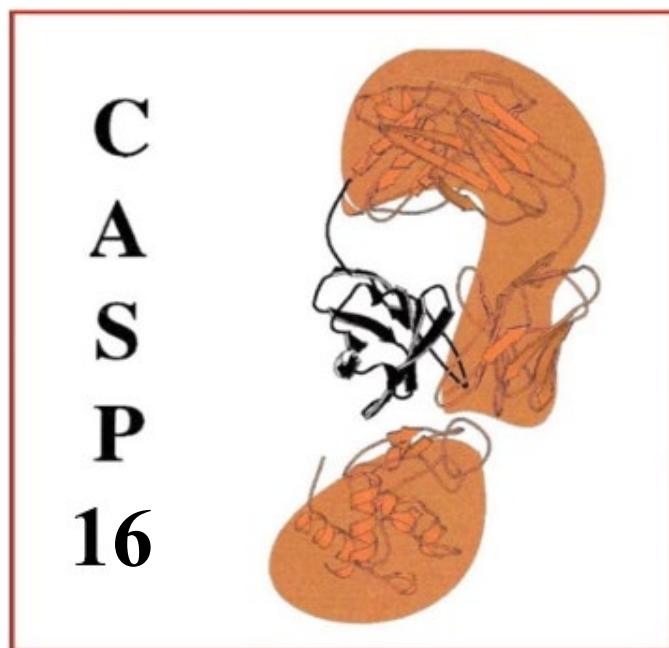


CRITICAL ASSESSMENT OF TECHNIQUES FOR PROTEIN STRUCTURE PREDICTION



ABSTRACT BOOK

Sixteenth round
May-August 2024

TABLE OF CONTENTS

2Vinardo.....	11
A classical empirical scoring function (2Vinardo) trained using multi-task learning with deep-learning-inspired strategies	11
AF_unmasked.....	13
Quality assessment using AF_unmasked	13
AF3-server, elofsson, pDockQ, pDockQ2, Pcons	14
AlphaFold3 in CASP16.....	14
aicb.....	18
Perturbation of AlphaFold to predict protein structure ensembles	18
APOLLO (QA).....	20
Quality Estimation of Single-Model Individual Interface Residues using Graph Neural Networks	20
ARC (QA, TS)	21
ARC: Single-Model Quality Assessment of Interface Residues Using an Ensemble of Graph Neural Networks	21
arosko	23
Protein-Small Molecule Complex Prediction in CASP16	23
Ayush	25
Protein Assembly Refinement with Variational Autoencoders	25
Bhattacharya (TS).....	27
Interplay of AlphaFold2 and AlphaFold3 for Protein Structure Prediction.....	27
Bhattacharya (TS, RNA, Protein-RNA)	28
Scoring-guided RNA Structure and Integrative Protein-RNA Complex Prediction.....	28
BRIQX.....	30
RNA Structure modeling with BRIQ2	30
bussilab_plain_md	31
Predicting RNA solvation shell using molecular dynamics simulations	31
bussilab_replex	34
Predicting RNA solvation shell using enhanced sampling molecular dynamics simulations.....	34
ChaePred	37

Protein Complex Structure Prediction and Accuracy Estimation Using Machine Learning-based Docking Potential.....	37
Cheatham-lab.....	39
Tetrahymena Ribozyme RNA Ion and Water Placement Prediction with Comparative Divalent Ion Models.....	39
ClusPro and Kozakov/Vajda	41
Prediction of protein assemblies and ligand binding modes using a combination of ClusPro, AlphaFold and LigTBM diffusion	41
COAST (QA, TS), APOLLO (TS).....	43
COAST: Protein Complex Accuracy Estimation Transformer.....	43
ColabFold.....	46
Benchmarking ColabFold in CASP16	46
Coogs2.....	48
Prediction of Local Solvent Shell Structure of Tetrahymena Ribozyme RNA Using Molecular Dynamics.....	48
Coogs3.....	51
Prediction of Local Solvent Shell Structure of Tetrahymena Ribozyme RNA Using Molecular Dynamics.....	51
Cool-PSP.....	54
Structure Prediction and Selection Method of Cool-PSP.....	54
CSSB_experimental (TS, Assembly).....	57
Improved MSA and New Deep Learning Model for Protein Structure Prediction in CASP16.....	57
CSSB_experimental (NA).....	60
Improving RNA and Protein-Nucleic Acid Complex Structure Predictions with Deep Learning-Based Diverse Structure Sampling	60
CSSB_FAKER (TS, Assembly)	62
Generating MSA for Protein Structure Prediction via Structure Search and Sequence Design	62
CSSB-Human (TS, Assembly)	65
Guiding Structure Prediction with Human Intuition	65
CSSB_experimental (Antibody).....	68
Epitope-guided predictions of antibody-antigen complex structures.....	68
DeepFold, DeepFold-server.....	70

Protein 3D Structure Prediction with DeepFold in CASP16	70
DeepFold-interact.....	72
Deep-interact: A Graph Transformer-Based Protein-Ligand Interaction Site Prediction.....	72
DeepFold-refine	75
Refining Protein Structures with Molecular Dynamics and Targeted Restraints	75
DELCLAB	77
DELCLAB in CASP16	77
DIMAIO	79
Method description for modeling ligand targets from CASP16	79
dNAfold	81
Ab initio RNA structure prediction with RNA composite likelihood maximized language model and deep end-to-end potential	81
Drugit	83
Crowdsourcing Small Molecule Structure Prediction with Foldit Drug Discovery	83
elofsson	85
Methods for Improved Protein Structure and Protein-Protein Interaction Prediction	85
FALCON	88
Fernandez-Recio	89
Modeling of protein assemblies by applying pyDock energy-based scoring to AlphaFold models....	89
FTBiot0119 (TS)	91
Template-Based Structure Prediction by using deep learning	91
FTBiot0119 (assembly)	92
Assembly prediction with ab initio docking and template-based modeling	92
GeneSilico-Human, SimRNA-server, GeneSilico-server.....	93
Computational modeling of nucleic acid 3D structures and interactions.....	93
GHZ-ISM, GHZ-MAN	96
UniFold, AlphaFold2, AlphaFold2.3, AlphaFold3 and updated DMFold-based Optimization for Protein Structure Prediction.....	96
GromihaLab	97
Biomolecular structure prediction and accuracy estimation methodology for human prediction using different modeling techniques	97

GruLab; Convex-PL-R; KORP-PL-W	100
Predicting protein-ligand binding with Convex-PL and KORP-PL tools.....	100
GuangzhouRNA-human	101
Enhancing RNA 3D Structure Prediction: Integrating Computational Tools with Expert Refinement for Functional Insights	101
GuangzhouRNA-meta	103
Integrating multiple tools for RNA 3D structure prediction in CASP16.....	103
GuijunLab-Complex	105
Protein multimer structure modeling using deep learning-based paired MSAs.....	105
GuijunLab-Human	107
Protein complex modeling based on inter-chain interactions obtained by deep learning and traditional docking methods	107
GuijunLab-PAthreader	109
Improving protein structure prediction with enhanced MSA and model quality assessment.....	109
GuijunLab-QA	111
Protein model quality assessment and selection	111
Haiping	114
SGraph_RMSD/affinity: predicting RMSD or affinity of protein-ligand complex by graph transformer with single graphic to represent the protein-ligand interface	114
Huang-HUST (Lig)	116
Protein-ligand Structure Prediction by Template-guided Ensemble Docking	116
Huang-HUST (QS).....	118
Modeling Multimeric Structures through Enhanced Sampling of AlphaFold-Multimer and HDOCK	118
HYU_MLLAB	120
Enhancing Protein Embeddings via Contrastive Learning	120
isyslab-hust	123
Optimizing protein structure prediction using deep learning and Rosetta	123
isyslab-hust (LIG)	125
Molecular Docking of Small Molecule Ligands with Targets Using AutoDock Vina	125
isyslab-hust (RNA)	126
Structure Prediction of RNA Monomers and Complexes Using Rosetta and AlphaFold3	126

KiharaLab.....	128
Integrating Advanced Computational Techniques for Biomolecular Structure Prediction	128
Koes	132
Predicting Ligand Binding Poses and Affinities for Four Target-Ligand Sets Using	132
Open-ComBind, GNINA, Dynamicbind and GNINA Rescoring of DynamicBind Predictions	132
KUMC	135
Template based Protein-Ligand Complex Structure Prediction in CASP16.....	135
LCBio	137
Comprehensive RNA and RNA-Complex Structural Modeling for CASP16	137
LCDD-team.....	140
Ligand binding pose prediction via the combination of geometric deep-learning and physics-based scoring models.....	140
Lindorff-Larsen.....	142
CALVADOS Predictions of the Conformational Ensemble of the Two-Domain Protein ZLBT-C ..	142
Lindorff-Larsen.....	144
Martini 3 simulations with increased protein-water interactions	144
MassiveFold.....	146
Massive sampling with MassiveFold in CASP16-CAPRI.....	146
McGuffin.....	149
Manual Prediction of Protein Tertiary and Quaternary Structures, Model Quality, Protein-Ligand Interactions and Ensembles	149
MIEensembls-Server	153
Protein and Nucleic Acid Ensemble Structure Prediction in CASP16	153
ModFOLDdock2, ModFOLDdock2R, ModFOLDdock2S.....	156
Automated Quality Assessment of Protein Quaternary Structure Models using the ModFOLDdock2 Server	156
MQA_base, MQA_server, MQA.....	159
Deep Learning based protein complex model accuracy estimation and model selection	159
MRAFold	161
Enhanced Protein Structure Prediction through MSA Realignment.....	161
MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human (TS)	163

Enhancing AlphaFold2/AlphaFold3-based Protein Tertiary Structure Prediction with Extensive Modeling Sampling and Ranking.....	163
MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human (QS).....	167
Improving AlphaFold2/AlphaFold3-based Protein Quaternary Structure Prediction with Large-Scale Modeling Sampling and Ranking.....	167
MULTICOM_ligand (LG)	171
Geometric Deep Learning, Generative Modeling, and Structural Consensus Ranking for Protein-Ligand Structure Prediction and Binding Affinity Estimation.....	171
MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human, MULTICOM_ligand (QA)	174
Protein Model Evaluation and Selection Using Graph Transformers and Multiple Quality Metrics in CASP16.....	174
MultiFOLD2	176
Automated Sampling, Quality Assessment and Refinement of Tertiary and Quaternary Structure Models using the MultiFOLD2 Server.....	176
NEC_Compute	180
Low-Rank Adaptation for Folding tools	180
NKRNA-s	181
Protein-Nucleic Acid Complex Structure Prediction Using Deep Learning and Language Models ..	181
OpenComplex.....	184
OpenComplex-2: A Biomolecular Foundation Model.....	184
orangeballs	188
AlphaFlow: flow-based generative models from single-structure predictors	188
Pascal_Auffinger	189
Completion of the solvation shell of 7ez0 based on MD simulations at cryo-EM temperatures.....	189
PerezLab_Gators	192
MELD-RNA: RNA structure prediction from ambiguous data sets	192
PerezLab_Gators	194
Improving protein and RNA structure prediction via sampling of AlphaFold and MELD simulations	194
PEZYFoldings	196
PEZYFoldings pipeline in CASP16.....	196

PEZYFoldings (EpitopeFinder)	199
Epitope Finding Protocol with AlphaFold-Multimer.....	199
PEZYFoldings (Extended).....	201
PEZYFoldings Extended Pipeline for CASP16.....	201
Pfaender.....	203
Graph Neural Network Ensemble Prediction of Protein Ligand Complex.....	203
PFSC_PFVM.....	204
Predict Conformation Ensembles of Protein Structures with PFSC-PFVM Approach	204
PIEFold_human	207
PIEFold: an Improved Template-based Protein Structure Prediction Pipeline with PBEscore	207
PLMFold	209
Enhancing Protein Structure Prediction through MSA Quality Improvement with PLM-MSA	209
PocketTracer	211
Advancing Protein-Ligand Binding Site Prediction with 3D Instance Segmentation Techniques....	211
RNA_Dojo	213
RNA 3D Structure Prediction with Human Intervention: Combining MXfold2, IPknot and FARFAR2	213
RNAFOLDX	216
<i>Ab initio</i> RNA structure prediction with AI-based inter-and physics-based potential.....	216
RNApolis	218
Modeling RNA and RNA-containing 3D structures in CASP16 by the RNApolis group	218
Schneidman.....	220
Advancing structure prediction: modeling large complexes, protein-ligand, and antibody-antigen interactions beyond AlphaFold	220
Seder2024	222
SHORTLE	223
Quality Assessment of QA1 and QA3 Models by Ranking a Set of Score for Parameters Based on PDB Statistics Covering Diverse Structural Features.....	223
SNU-CHEM-lig, SNU-CHEM-aff (Ligand)	225
Enhancing Structure and Affinity Prediction of Receptor-Ligand Interactions with Variable Receptor Flexibilities Using Deep Learning Methods	225
SoutheRNA	228

An Improved Pipeline for RNA Structure Prediction: Using Coarse-Grained Simulation and Refinement to Enhance AF3 Predictions	228
thermomaps.....	229
In Silico Generation and Thermodynamic Ranking of RNA Tertiary Structures.....	229
Unicorn.....	231
Updated DMFold-based Optimization for Protein Structure Prediction.....	231
UNRES.....	232
Protein and RNA/DNA structure prediction with the UNRES and NARES-2P coarse grained models assisted by AlphaFold	232
Vakser	235
Modeling Protein-Protein Targets by Combining and Analyzing AlphaFold Predictions	235
Vfold	238
RNA Structure Prediction in CASP16 Using Hybrid Methods	238
VifChartreuse, VifChartreuseJaune	240
Interface-focused Scoring of Protein Assemblies in CASP16	240
VoroAffinity, VoroAffinityB	242
Protein-Ligand Binding Affinity Prediction Using Descriptors Derived From Voronoi Tessellation	242
Wallner	243
Improved Massive Sampling using AFsample2.....	243
XGroup, XGroup-server	246
A protein-protein complex structure prediction model via MSA and deep learning	246
Yang-Server, Yang-Multimer, Yang	247
Protein and RNA structure prediction with trRosetta, trRosettaRNA and AlphaFold	247
Zheng	248
Protein and Nucleic Acid Structure Prediction by Zheng Group in CASP16.....	248
Zheng-Multimer	251
Protein Complex Structure Prediction Using Deep Learning and Multi-MSA Strategy in CASP16.	251
Zheng-Server	254
Prediction of Protein Tertiary Structures Utilizing Deep Learning Spatial Restraints and Knowledge-Based Potentials in CASP16	254
Zou_lab	257

Integration of template guidance, docking, and deep-learning for protein-ligand binding mode and affinity prediction.....	257
OTHER.....	260
How Physics can Improve AI Generated TCR-pMHC Complexes	260
MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human (Flexible targets)	262
Predicting Conformation Ensembles for Flexible Protein Targets with MULTICOM4	262
MoMAteam1	265
Exhaustive conformational sampling of domain-linker-domain proteins.....	265
JFK-THG.....	268
Using IDPConformerGenerator and Molecular Dynamics Strategies to Predict Conformational Ensembles of Staph. Aureus ZLBT-C.....	268
KiharaLab.....	272
Integrating Advanced Computational Techniques for Biomolecular Structure Prediction	272
CoDock.....	273
Strategies for ligand binding prediction in CASP16.....	273

A classical empirical scoring function (2Vinardo) trained using multi-task learning with deep-learning-inspired strategies

R. Quiroga and M.A. Villarreal

Departamento de Química Teórica y Computacional, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba.

Instituto de Investigaciones en Físico-Química de Córdoba (INFIQC-CONICET). ARGENTINA.

mvillarreal@unc.edu.ar

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:N; MD:N

Structure-based drug discovery methodologies such as molecular docking and virtual screening have proved to be invaluable tools in developing novel drugs. Central to these methods are scoring functions, which predict the binding affinity between ligands and protein targets. Machine-learning based docking has proven to improve upon classical empirical scoring functions for single tasks such as scoring and virtual screening, although it has been shown that in some cases these results can be due to overfitting, memorization, and learning the inherent bias of training sets. In recent years, several innovative deep learning-based docking methods were introduced, however, it was shown that on carefully curated datasets, these methods did not outperform classical docking software and scoring functions such as Autodock Vina and Gold¹. We propose that applying deep learning strategies to train novel empirical scoring functions may offer a promising solution to outperform current applications while displaying generalizability and avoiding overfitting and memorization issues.

Methods

We have developed a novel but classical empirical scoring function denominated 2Vinardo. This scoring function is the second generation of scoring functions developed in our lab, with the first being Vinardo², based on the Autodock Vina scoring function³. 2Vinardo consists of a linear combination of short-range, long-range, non-directional-hydrogen-bond and repulsion terms. Each of these terms are calculated on an atom-pair bases, for which we define 21 different atom types using a custom version of Openbabel. For this empirical scoring function, we developed a parameter training framework programmed in Julia, which uses a procedure which combines what we call batch-annealing, automatic differentiation and gradient descent to optimize a loss function which consist of linearly combined multi-task objectives. These objectives aim to simultaneously optimize the scoring, docking and virtual screening capabilities of the 2Vinardo scoring function, by tuning the approximately 250 parameters that define 2Vinardo.

We used the 2Vinardo scoring function to perform rigid-flexible protein-ligand docking for the L1000 CASP16 supertarget. To do this, we performed docking using the 2vinardo software, based on idock⁴ and Autodock Vina³ code, which uses a Monte Carlo procedure deemed “iterated local search global optimizer” with BFGS as the method for local optimization and the Metropolis criterion as the decision mechanism to decide to accept or reject each step³. We used all protein structures for chymase available in the PDB database, performed docking for each ligand, and selected the structure for the best scoring protein-ligand complex.

Availability

2Vinardo software is at early development stages, and is available from authors upon request. Please contact Marcos Villarreal at mvillarreal@unc.edu.ar

1. Buttenschoen M, Morris GM, Deane CM. (2024). PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chem Sci.* **15**(9), 3130–3139.
2. Quiroga R, Villarreal MA. (2016). Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. Sticht H, editor. *PLOS ONE*. **11**(5):e0155183.
3. Trott O, Olson AJ. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* **31**(2):455–461.
4. Li H, Leung KS, Wong MH. (2012). idock: A multithreaded virtual screening tool for flexible ligand docking. *2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE **2012**:77–84.

AF_unmasked

Quality assessment using AF_unmasked

Claudio Mirabello^{1,2}, Yogesh Kalakoti¹ and Björn Wallner¹

¹ Division of Bioinformatics, IFM, Linköping University

² National Bioinformatics Infrastructure Sweden

bjorn.wallner@liu.se

Key: Auto:Y; CASP_serv:Y; TempL:N; MSA:N; DeepL:Y; EMA:Y; MD:N

The default version of AlphaFold-Multimer¹ only uses templates for the monomers of a multimer, completely ignoring the relative orientation of subunits. AF_unmasked² enables the use of templates for multimeric proteins, making it possible to predict interactions with large complexes without remodelling everything from scratch. In the quality assessment category, we utilized AF_unmasked as a scoring function by inputting the multimeric models as templates with no MSA information to the AF2 inference system.

Methods

The rescoring using AF_unmasked uses inference with the ‘multimer_5_v2’ neural network with the input model as a template and no MSA information. If the length of the input model was shorter than 1,400 amino acids, the input model was relaxed using the standard AF2 Amber relaxation protocol before rescoring. Five models were generated, and the best model confidence was used as the score for the input model. For models larger than 1,400 amino acids, the unrelaxed model was used, and only one model was generated to save time.

The local probabilities for interaction between chains were simply the pLDDT scores for the residues within 8 angstroms CB of another chain.

Availability

AF_unmasked is available at https://github.com/clami66/AF_unmasked

1. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *Biorxiv* 2021.10.04.463034 (2021) doi:10.1101/2021.10.04.463034.
2. Mirabello, C., Wallner, B., Nystedt, B., Azinas, S. & Carroni, M. Unmasking AlphaFold: integration of experiments and predictions in multimeric complexes. *Nature Communication* (2024).

AlphaFold3 in CASP16

Arne Elofsson

Dept of Biochemistry and Biophysics and Science for Life Laboratory, Stockholm University

arne@bioinfo.se

Key: Auto:Y; CASP_serv:Y; TempL:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:Y

We submitted manual predictions as group Elofsson (#241) and server predictions as AF3-server (#304). All predictions were made using the AlphaFold3 server¹. The five models from one target were submitted for the server predictions. In contrast, more models were generated, and good models with some structural variability were submitted for the manual predictions. For some models, additional changes were made; see methods.

Quality estimations were performed using pDockQ²(#080), pDockQ2³(#446) and Pcons⁴(#471) whenever the methods worked and managed to submit the results in time. These methods are not discussed here as they have been described before, and we do not believe the results should be perfect for CASP16 targets.

Methods

Here, we will describe the manual interventions we have performed for different categories of targets.

Unknown stoichiometry

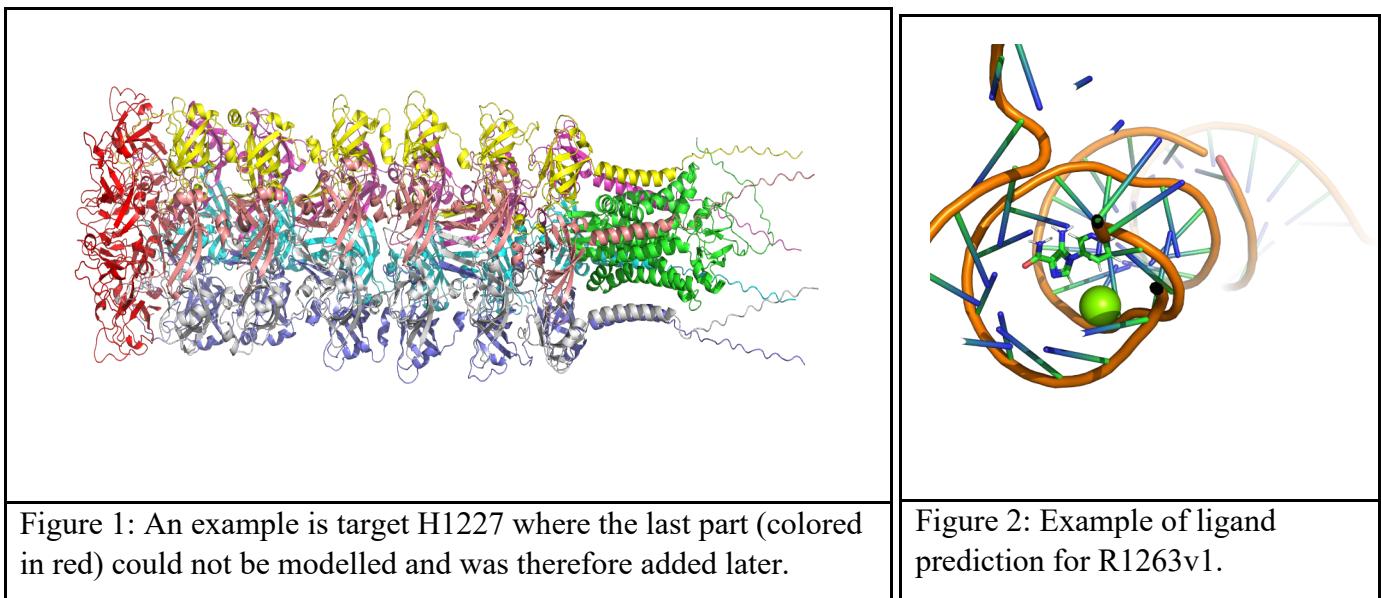
For targets with unknown stoichiometry, we examined the possibility of predicting stoichiometry by generating models with possible stoichiometries. For computational reasons, it was not possible to generate all possible stoichiometries. Therefore, we used some manual intervention to select which ones to try. We usually ran monomers to hexamers for a start - but this was impossible for all H targets as the number of combinations was too big. After an initial screen, we used the ranking confidence and evaluation of the PAE map to predict the stoichiometry.

Too big targets

Some targets (e.g., H1227, H1257) were too big to run on the server (the maximum is 5000 residues/nucleotides). Here, we used the strategy from MolPC5, i.e., building different types of overlapping fragments, superposing the shared parts, and thereby generating complete models, see Figure 1.

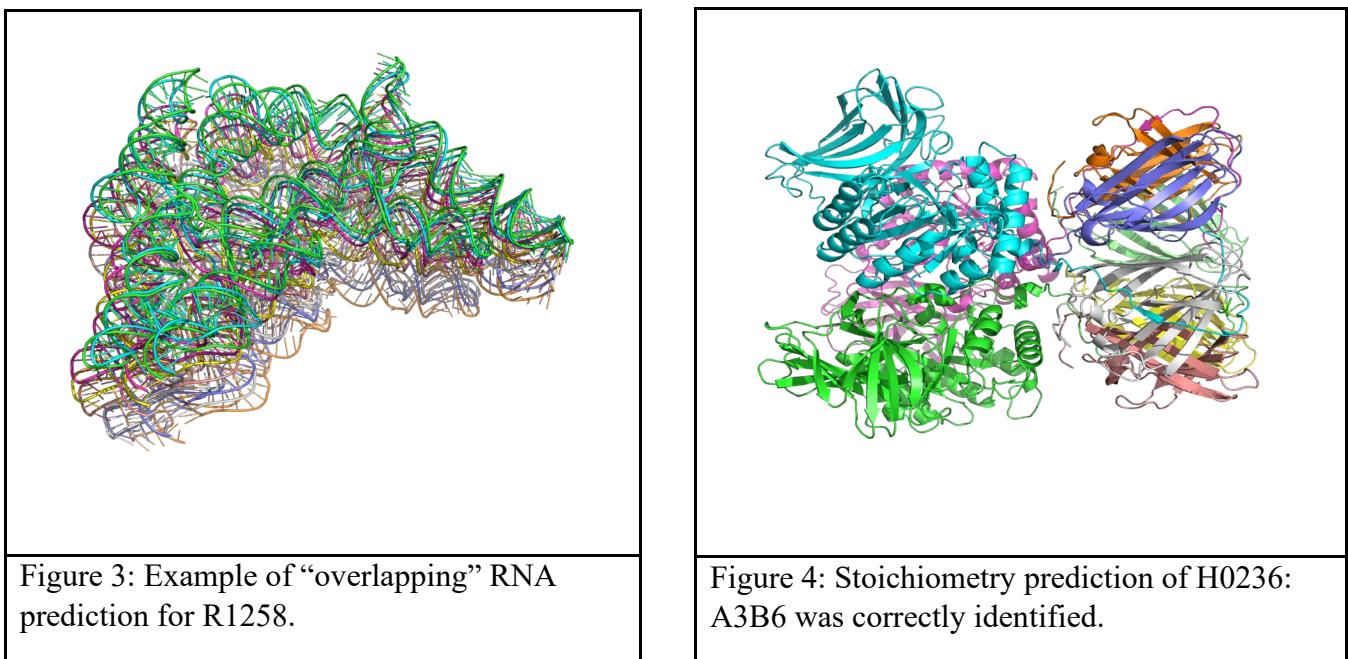
Ligand predictions

For ligand targets, we (1) generated a structural model from the smiles using online tools and (2) ran the AlphaFold3 Server with the available ligand most resembling the ligand that should be docked. Finally, the CASP ligand was superposed on the AlphaFold3, and a new PDB/Ligand file was generated, Figure 2.



RNA/DNA targets

It was noted that for many RNA-containing targets, the server's resulting models put the chains on top of each other, Figure 3. For these targets, we generated many models with the server and used internal scripts to identify models with no (or minimal) overlap. This was also apparent from the internal AlphaFold3 scores.



Ensemble targets

Different strategies were used for different ensemble targets.

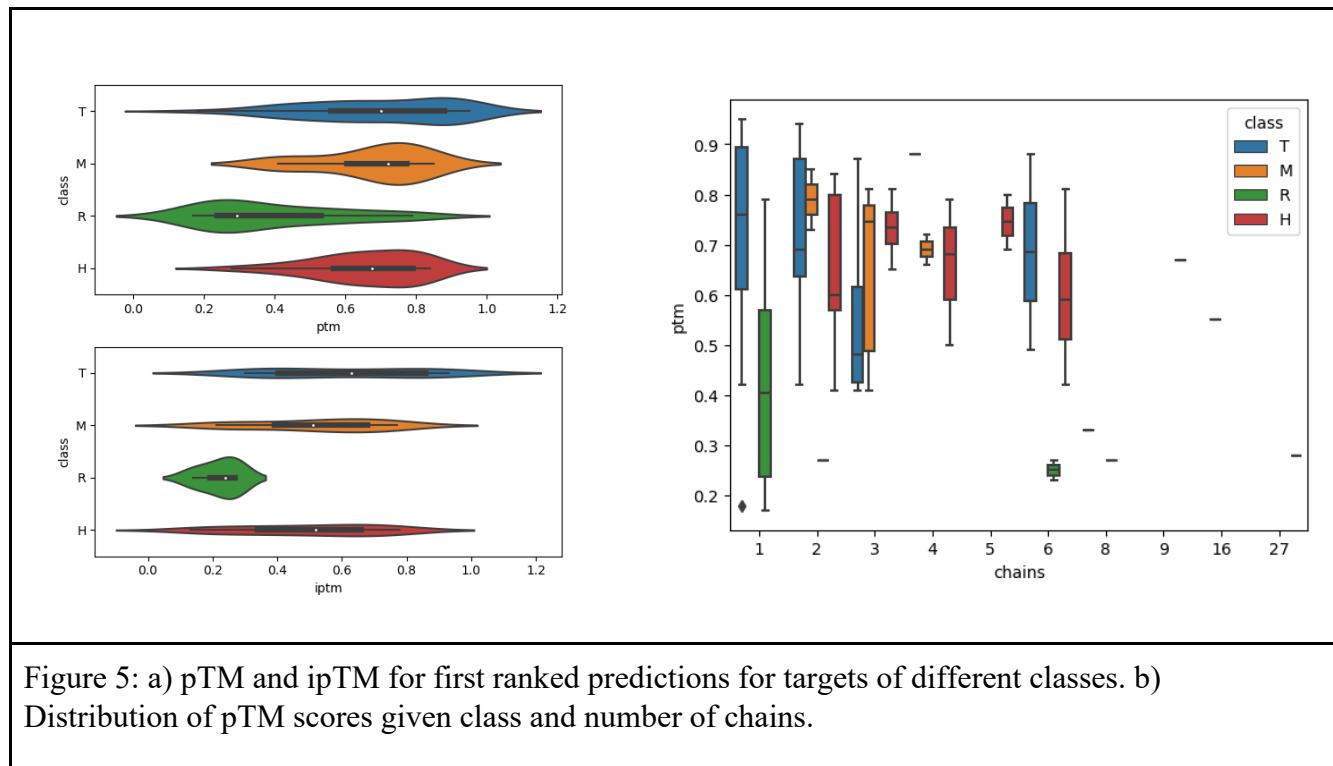
- R1203, T1214, R0283 - Standard predictions using AlphaFold3
- T1294, T1200/T1300, M1239, T1249, M1228, T2249, R1253 - Generated many models and submitted ensembles of them after clustering
- R1260 (solvation shell) - For AF3, an ensemble of models was generated and hydrated using standard protocols. Elofsson took the best models and ran a 47 ns long MD simulation.

Results

Here, we report some interesting observations.

Stoichiometry

For the monomeric target, we submitted the correct model (as first ranked) for 9 (and for lower rank for the other targets), i.e. it seems quite possible to predict the stoichiometry for the monomeric target. However, it should be noted that these predictions were not made with a single score, as it was observed that the monomer's score was often higher than for multimeric models, even when it was a multimer. Most models with only one copy of each chain were correctly predicted for multimeric models. However, only two models that did not have this case were predicted correctly. For mixed “M” models - no single guess of stoichiometry was correct.



Predictions

It was observed that predictions of RNA targets were significantly less reliable than other type of targets, see Figure 5a. For all other types of targets the pTM were 0.6-0.7 on average and ipTM scores 0.5 to 0.6. For RNA targets average pTM was 0.3 and ipTM 0.2.

Availability

AlphaFold3 is only available as a web server for non-commercial entities. Code under unknown licensing rules (likely CC BY-NC-SA 4.0) has been promised to be released before the meeting.

Details about all predictions can be found here

<https://github.com/ElofssonLab/casp16/blob/master/Targets.MD>

1. Abramson J., et al. Accurate structure prediction of biomolecular interactions with AlphaFold. *Nature* 630, 493–500 (2024)
2. Bryant, P, Pozzati, G, Elofsson, A Improved prediction of protein-protein interactions using AlphaFold2, *Nat Commun* 2022 Mar 10;13(1):1265.
3. Zhu, W, Shenoy, A, Kundrotas, P , Elofsson, A Evaluation of AlphaFold-Multimer prediction on multi-chain protein complexes, *Bioinformatics*. 2023 Jul 1;39(7):btad424.
4. Lundström, L, Rychlewski, L, Bujnicki,, J. Elofsson, A. Pcons: a neural-network-based consensus predictor that improves fold recognition. *Protein Sci* . 2001 Nov;10(11):2354-62.
5. Bryant, P, Pozzati, G, Zhu, W, Shenoy, A, Kundrotas, P, Elofsson, A, Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search. *Nature Comm* 13,: 6028 (2022)

Perturbation of AlphaFold to predict protein structure ensembles

Ziheng Zou, Shuya Li and Jianyang Zeng^{*}

School of Engineering, Westlake University

* corresponding author: zengjy@westlake.edu.cn

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; EMA:N; MD:Y

We participate in CASP16 focusing on the protein ensemble tasks T1200 and T1300. We developed perturbation strategies based on OpenFold, which is a torch-based reproduction of AlphaFold2. By introducing physical perturbations into the structure modules, we obtained a series of protein structural variants as the output. Then, we calculated the probabilities of perturbed conformers with the help of the Amber energy function. We utilized our pipelines *pert-linker* and *pert-plddt* to produce protein ensembles with and without the help of prior knowledge of the linker and functional domains.

Methods

We modified the inference process of OpenFold to introduce more flexibility to the output structures without changing the feature generation and transformations^{1,2}. Without specific descriptions, our method shared the same procedures as OpenFold, including the MSA generation and templates searching. Here, we introduced the detailed modifications to predict the protein structure ensembles with the prior knowledge of PDB records 2LR2 and 4NDP by *pert-linker*, and to use *pert-plddt* to infer the structures automatically.

Reference rigid generation

Based on the original OpenFold pipeline, we first obtained the raw predicted structure. We aligned the 2LR2 and 4NDP structures to the raw structure and calculated the rigids of structures. According to sequence alignments, we replaced parts of the rigids of the raw predicted structure by those of 2LR2 and 4NDP. The merged rigids would serve as the reference in the perturbation step for *pert-linker*. And *pert-plddt* requires no reference rigids.

Reinitialization and Perturbation

To modify the predicted structures based on our prior knowledge, we introduced two novel operations into the structure module of AF2, that is, reinitialization and perturbation. Here we applied the operations in one of the eight blocks in the structure module. To be specific, reinitialization reads the given reference rigids to rebuild the initial conformer, and perturbation adds a random rotation to specific residues. The random rotations were sampled from uniform distributions with user-defined parameters. Based on the requirements of the T1200 and T1300 targets, we set the linker residues in *pert-linker* as the perturbed residues manually. And *pert-plddt* chose low plddt-value residues as the perturbation sites.

Structure prediction and scoring

With ten weight parameter files published by AF2, we generated 300 random conformers for each parameter file with diverse perturbation parameters and different random seeds. The output conformers were relaxed by the standard OpenFold pipeline, and the resulting Amber energies were recorded³.

Population distribution estimation

Inspired by the conformer searching steps in classical molecular dynamics methods, we clustered the predicted conformers with MaxCluster and selected represents in each cluster⁴. Then, the probabilities of filtered conformers were calculated based on the Boltzmann distribution.

Availability

The related code and manuscript will be released after publication.

1. Ahdritz, G., Bouatta, N., Floristean, C., Kadyan, S., Xia, Q., Gerecke, W., ... & AlQuraishi, M. OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods*. 21 (2024) 1514-1524.
2. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596 (2021) 583-589.
3. Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., ... & Merz Jr, K. M. AmberTools. *Journal of chemical information and modeling*. 63 (2023) 6183-6191.
4. Herbert, A., & Sternberg, M. J. E. MaxCluster: a tool for protein structure comparison and clustering. (2008) <http://www.sbg.bio.ic.ac.uk/~maxcluster/index.html>

Apollo (QA)

Quality Estimation of Single-Model Individual Interface Residues using Graph Neural Networks

Gabriel Huang, Andrew Jordan Siciliano, Bishal Shrestha, and Zheng Wang

Department of Computer Science, University of Miami

zheng.wang@miami.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:N; MD:Y

Apollo is a single-model quality assessment method that uses graph neural networks (GNNs) to predict the quality of interface residues within protein-complex models. Our server participated in QA (QMODE-2 and QMODE-3) categories. QMODE-2 was largely done through automated scripts with minimal human intervention, while QMODE-3 targets required human intervention.

Methods

For QMODE-2 we directly predicted the local qualities of interface residues. To predict the global interface score we average the predicted quality scores for interface residues, which was used for both QMODE-2 and the ranking of relaxed models in QMODE-3. QMODE-3 was skipped for targets that could not be relaxed in time due to computational issues (approx. > 1000 residues or heavy server usage). Relaxation was performed using OpenMM¹ and the Amber99sb force field² as described and implemented by AlphaFold2³⁻⁴.

Apollo first represents the protein as a graph in which each node corresponds to a residue, and the edges represent spatial relationships between residues. The GNN performs message passing between connected residues to learn meaningful representations of the protein and its features. This message-passing operation is implemented through a MetaLayer⁵, where multi-layer perceptron (MLP) blocks are used to aggregate and transform information from neighboring residues. PatchQS⁶ was calculated using OpenStructure⁷ and was used as the ground truth. Information on node and edge features, as well as dataset generation, can be found within the abstract for COAST.

1. Eastman, P. et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Computational Biology* 13, e1005659.
2. Hornak, V. et al. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3), 712-725.
3. Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589
4. Evans, Richard, et al. "Protein complex prediction with AlphaFold-Multimer." *biorxiv*: 2021-10.
5. Battaglia, P.W., et al., Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261, 2018.
6. Studer, G., Tauriello, G., & Schwede, T. (2023). Assessment of the assessment—All about complexes. *Proteins: Structure, Function, and Bioinformatics*, 91(12), 1850-1860.
7. Biasini, Marco, et al. "OpenStructure: an integrated software framework for computational structural biology." *Acta Crystallographica Section D: Biological Crystallography* 69.5 (2013): 701-709

ARC (QA, TS)

ARC: Single-Model Quality Assessment of Interface Residues Using an Ensemble of Graph Neural Networks

Bishal Shrestha, Andrew Jordan Siciliano, and Zheng Wang^{*}

Department of Computer Science, University of Miami

^{*}Corresponding Author: zheng.wang@miami.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:Y; MD:Y

We developed ARC, a single-model quality assessment method that predicts interface residue quality scores for protein-complex models. Our server participated in both assembly (TS) and Quality Assessment (QA) categories, specifically QMODE-2 and QMODE-3. QMODE-2 was largely done through automated scripts with minimal human intervention, while TS and QMODE-3 targets required human intervention.

Methods

For QMODE-2, local interface qualities were predicted directly from ARC, and the global interface prediction (QSCORE) was the average of the predicted local interface scores. QSCORE was used as the metric for ranking the top 5 relaxed models for QMODE-3. QMODE-3 was skipped for targets that could not be relaxed in time due to computational issues (approx. > 1000 residues or heavy server usage). For TS predictions, we sampled multimeric structures using dropout-enabled ESM-Fold¹⁻² with varying gap and linker sizes. Structures with varying stoichiometry were generated for Phase 0 TS targets. For both Phase 0 and Phase 1 targets, the top 5 sampled models ranked by ARC's predicted global quality score were relaxed and submitted. Relaxation was performed using OpenMM³ and the Amber99sb force field⁴ as described and implemented by AlphaFold2⁵⁻⁶.

Model Architecture. ARC implements an ensemble of state-of-the-art graph-based deep learning algorithms to predict per-residue quality scores. These include Generalized Graph Convolution (GENConv)⁷, Graph Attention Networks (GATv2)⁸, GeneralConv⁹, Pathfinder Discovery Networks (PDNConv)¹⁰, Graph Isomorphism Network with Edge Features (GINEConv)¹¹, Residual Gated Graph ConvNets (ResGatedGraphConv)¹², and Graph Transformer (TransformerConv)¹³. Each base model processes the protein-complex model as a graph where the nodes correspond to the amino acid residues and edges represent spatial relationships within 14Å. The node and edge features consist of a comprehensive set of attributes derived from both the sequence and structure of the protein complex model. The ground truth (PatchQS)¹⁵ was calculated using OpenStructure¹⁴. Details regarding dataset generation, node features, and edge features are located within the abstract for COAST.

Each GNN processes the graph using specialized convolutional layers and message-passing mechanisms that leverage node and edge attributes to capture complex structural patterns and interactions. The ensemble integrates diverse techniques such as dynamic attention mechanisms⁸, edge convolutions^{11,13}, learnable aggregation functions⁷, message normalization⁷, residual connections¹², gating mechanisms¹², flexible multi-head operations⁹, and pathfinder discovery networks¹⁰. These mechanisms collectively enhance the models' ability to represent intricate spatial relationships, capture long-range dependencies, and distinguish between different types of residue-residue interactions.

After processing through successive layers of message passing and feature transformation, each model outputs per-residue quality predictions, which are passed through a sigmoid activation function to produce scores between 0 and 1. The ensemble aggregates the outputs from all base models by averaging the per-residue scores, providing the final per-residue prediction.

1. Lin, Zeming, et al. "Evolutionary-scale prediction of atomic-level protein structure with a language model." *Science* 379.6637 (2023): 1123-1130.
2. Rives, A., et al., Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 2021. 118(15): p. e2016239118.
3. Eastman, P. et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Computational Biology* 13, e1005659.
4. Hornak, V. et al. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3), 712-725.
5. Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
6. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *BioRxiv*, 2021.
7. Li, Guohao, et al. "Deepgcn: All you need to train deeper gcns." *arXiv preprint arXiv:2006.07739* (2020).
8. Brody, S., Alon, U., & Yahav, E. (2021). How attentive are graph attention networks?. *arXiv preprint arXiv:2105.14491*.
9. You, J., Ying, Z., & Leskovec, J. (2020). Design space for graph neural networks. *Advances in Neural Information Processing Systems*, 33, 17009-17021.
10. Rozemberczki, B., Englert, P., Kapoor, A., Blais, M., & Perozzi, B. (2021, April). Pathfinder discovery networks for neural message passing. In *Proceedings of the Web Conference 2021* (pp. 2547-2558).
11. Hu, W., Liu, B., Gomes, J., Zitnik, M., Liang, P., Pande, V., & Leskovec, J. (2019). Strategies for pre-training graph neural networks. *arXiv preprint arXiv:1905.12265*.
12. Bresson, X., & Laurent, T. (2017). Residual gated graph convnets. *arXiv preprint arXiv:1711.07553*.
13. Shi, Y., Huang, Z., Feng, S., Zhong, H., Wang, W., & Sun, Y. (2020). Masked label prediction: Unified message passing model for semi-supervised classification. *arXiv preprint arXiv:2009.03509*.
14. Biasini,Marco,et al. "OpenStructure: an integrated software framework for computational structural biology." *Acta Crystallographica Section D: Biological Crystallography* 69.5 (2013): 701-709
15. Studer, G., Tauriello, G., & Schwede, T. (2023). Assessment of the assessment—All about complexes. *Proteins: Structure, Function, and Bioinformatics*, 91(12), 1850-1860.

Protein-Small Molecule Complex Prediction in CASP16

Andrew Rosko

Buck Institute for Research on Aging

arosko@buckinstitute.org

Key: *Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:N; MD:N.*

Seventy years ago, physical model building using metal rods allowed Watson, Crick, and Franklin to propose the now famous double-helical, hydrogen bonded structure for DNA based on low-resolution data from the then novel technique of macromolecular crystallography¹, and more recently, the game Foldit² allowed players from around the world to contribute to predicting the three-dimensional structures of novel proteins, in some instances matching or even ouperforming the conformational sampling algorithms available at the time, and has since expanded into small-molecule ligand prediction³.

I have participated in the 16th CASP experiment to test to what extent human manual model builders can contribute to ligand complex prediction in the current state of development of the field, an area that lags behind the availability of accurate target structures at close to whole-genome scale⁴. One of the drawbacks of such model building is its low throughput, which did not allow pose prediction for the autotaxin data set consisting of over 200 ligands. This is somewhat unfortunate given the fact that this set contained the largest diversity of chemotypes and binding modes of the six protein-ligand prediction targets in CASP16—however there were five other protein targets whose ligand poses have been submitted.

If I am successful in CASP16, I hope to attempt to apply whatever insights I have developed toward developing the next generation of sampling algorithms and ligand interaction force fields, as well as toward prediction of endogenous ligands for proteins of currently unknown function.

Methods

When predicting ligand poses and binding affinities, I utilized existing structures of the targets and their homologues from the PDB to guide assessments of the dynamic conformational variability of the proteins, and existing ligand complexes (when available) to inform possible modes of interaction and as benchmarks for ligand absolute affinities (through data from PDBBind⁵)—however I did not restrict poses or side chain conformations to those previously observed.

For the three enzyme-inhibitor data sets L1000, L2000, and L4000, I assumed that the ligands were competitive active site inhibitors and thus needed to at least partially overlap the substrate envelope, an assumption that seems, based on preliminarily released poses, to have been correct for all molecules except for one of the coronaviral MPro ligands. For the PQQ transporter model, the conformational changes undergone by other TonB-dependent transporters upon binding of their cognate substrates were used to predict the hinge points for the closure motion of the extracellular loops.

Ligand poses were submitted directly as predicted manually—no energy minimization was performed, except to generate physically plausible bond lengths and angles for the ligands themselves prior to model building. In some cases, where multiple plausible poses were found, more than one model was submitted.

1. Watson,J.D.,Crick,F.H.(1953) A Structure for Deoxyribose Nucleic Acid, *Nature* 171, 737-738.
2. Cooper,S., Khatib,F., Treuille,A.,Barbero,J., Lee,J., Beenen,M., Leaver-Fay,A., Baker,D., Popović,Z., Players,F.(2010) Predicting protein structures with a multiplayer online game, *Nature* 466, 756-760.
3. Eiben,C.B., Siegel,J.B., Bale,J.B., Cooper,S., Khatib,F., Shen,B.W., Players,F., Stoddard,B.L., Popovic,Z., Baker,D.(2012) Increased Diels-Alderase activity through backbone remodeling guided by Foldit players, *Nat. Biotechnol.* 30, 190-192.
4. Robin,X., Studer,G., Durairaj,J., Eberhardt,J., Schwede,T.(2023) Walters,W.P., Assessment of protein-ligand complexes in CASP15, *Proteins* 91, 1811-1821.
5. Wang,R., Fang,X., Lu,Y., Wang,S.(2004) The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures, *J. Med. Chem.* 47, 2977-2980.

Protein Assembly Refinement with Variational Autoencoders

Sailesh Prabhu

Independent Researcher

sailash.n.prabhu@gmail.com

Key: *Auto:N; CASP_serv:Y; Templ:N; MSA:N; DeepL:Y; AF:AF2; EMA:N; MD:Y*

Our solution utilizes a variational autoencoder¹ to refine predictions of protein assembly structure. The encoder embeds the solution from an upstream solver and samples several embeddings in the neighborhood. The decoder generates candidate protein structures from the sampled embeddings, and the protein structure with lowest RMSD to other samples is selected.

Methods

Model Architecture: The model utilizes an Encoder-Decoder architecture. The encoder is parameterized by a Signal-to-Noise ratio. This ratio adapts during the model training. When a Protein Structure is passed through the encoder, it generates another sample with the specified Signal-to-Noise Ratio.

The Decoder generates a predicted structure from the encoder's noisy sample. Each atomic coordinate is modeled as a Gaussian distribution. Atomic-level features are passed through multiple Sparse Transformer layers² to predict the mean and variance of the atomic coordinate's Gaussian Distribution. The features employed by the Decoder include: 1) the noisy samples from the Encoder and 2) the sequence of atoms and residues.

Model Training: The Model is trained on protein assemblies in the RCSB PDB³. The training pipeline utilizes MMseqs2⁴ to assign individual protein sequences to clusters. On each training step, the training pipeline randomly selects a cluster, and it randomly selects assemblies with sequences in that cluster. Each training step takes a gradient step to maximize the Evidence Lower Bound on the selected assemblies.

Inference: The inference pipeline utilizes the variational autoencoder to refine protein structure predictions from AlphaFold2⁵. The pipeline utilizes the encoder to sample 10 noisy versions of the protein structure. It passes the 10 noisy protein structures through the decoder to generate predicted solutions. The predicted solution with the lowest RMSD to all other predicted solutions is chosen as the most likely structure. The chosen structure is passed through an OpenMM⁶ minimization using the Amber Force Field⁷ to remove steric violations.

1. Kingma,D.P. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114, 2013.
2. Child R, Gray S, Radford A, Sutskever I. Generating long sequences with sparse transformers. arXiv preprint arXiv:1904.10509, 2013
3. Burley,S.K., Berman,H.M., Bhikadiya,C., Bi,C., Chen,L., Di Costanzo,L., Christie,C., Dalenberg,K., Duarte,J.M., Dutta,S., Feng,Z., Ghosh,S., Goodsell,D.S., Green,R.K., Guranoví,V., Guzenko,D., Hudson,B.P., Kalro,T., Liang,Y., Lowe,R., Namkoong,H., Peisach,E., Periskova,I., Prlí,A., Randle,C., Rose,A., Rose,P., Sala,R., Sekharan,M., Shao,C., Tan,L., Tao,Y.-P., Valasatava,Y., Voigt,M., Westbrook,J., Woo,J., Yang,H., Young,J., Zhuravleva,M., and Zardecki,C. (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Research, 47.
4. Steinegger, M. & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nature Biotechnology 35, 1026–1028.
5. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. Nat Commun. 2022 Mar 10;13(1):1265.
6. Eastman, P. et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. Plos Comput Biol 13.
7. Case, D. A., et al. (2021). Amber 2021. University of California, San Francisco.

Interplay of AlphaFold2 and AlphaFold3 for Protein Structure Prediction

Debswapna Bhattacharya

Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA
dbhattacharya@vt.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF:Y; EMA:Y; MD:N

This method tests an integrated protein structure prediction system based on an interplay of AlphaFold2¹ and AlphaFold3² for both monomeric and multimeric protein targets. The objective is to investigate the ability of an integrated version of AlphaFold2 and AlphaFold3 to improve the predictive modeling accuracy compared to the baseline AlphaFold versions. Since AlphaFold3 is only available as an online server with limited number of job submissions per day, the method participated in CASP16 as a human group to allow for more computational time.

Methods

First, AlphaFold3² (*alphafoldserver.com*) was used to predict 5 structural models given the CASP target sequences for monomeric inputs or for multimeric inputs of known stoichiometry. All AlphaFold3 predictions are then fed into AlphaFold2¹ as custom templates together with custom multiple sequence alignments (MSAs) generated by DeepMSA2³ and/or MMseqs2⁴. The predicted structural models from the customized AlphaFold2 run were submitted as the final predictions by selecting models with the highest confidence score.

ColabFold⁵ was used to run custom version of AlphaFold v2.3.0 that takes the AlphaFold3 predicted structural models as templates and custom MSAs generated by DeepMSA2 and/or MMseqs2. For the monomeric targets, *alphaFold2_ptm* weights were used with the maximum number of recycles set to 3, whereas for *alphaFold2_multimer_v3* weights were used for multimeric inputs with the maximum number of recycles increased to 20.

Since *alphafoldserver.com* only allows the submission of jobs with a total length less than 5,000 residues, the default version of AlphaFold2 was employed for large multimeric targets having length more than 5,000 residues. In addition, for some of the large multimeric targets the maximum number of recycles was reduced during the customized AlphaFold2 run to save computational time.

1. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
2. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 1–3 (2024).
3. Zheng, W. *et al.* Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods* **21**, 279–289 (2024).
4. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology* **35**, 1026–1028 (2017).
5. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature methods* **19**, 679–682 (2022).

Scoring-guided RNA Structure and Integrative Protein-RNA Complex Prediction

Sumit Tarafder, Rahmatullah Roche, and Debswapna Bhattacharya

Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA

dbhattacharya@vt.edu

Key: *Auto:N; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; AF:Y; EMA:Y; MD:N*

In CASP16, we tested a fully automated scoring-guided RNA 3D structure prediction system to investigate the effectiveness of our newly developed RNA scoring function for 3D structural model selection and ranking. In addition, we tested an integrative protein-RNA complex prediction system by combining our newly developed single-sequence method based on geometric attention-enabled pairing of biological language models with other existing methods.

Methods

Scoring-guided RNA Structure Prediction

We used several deep learning-based RNA structure prediction methods to generate an ensemble of RNA 3D structural models from the CASP16 target sequence. To generate the structural ensemble, we used trRosettaRNA¹, a modified version of DeepFoldRNA², DRfold³, NuFold⁴, our customized version of NuFold with fine-tuned weights, E2EFold-3D⁵, RoseTTAFold2NA⁶, AlphaFold3⁷ webserver, and AlphaFold3 predictions relaxed by restrained optimization using PyRosetta⁸. We used the default rMSA⁹ pipeline to generate the MSA for a given RNA sequence as an input to the MSA-based RNA structure prediction methods.

To select the top 5 models from this structural ensemble, we used our newly developed locality-aware invariant point attention model for scoring RNA 3D structures, lociPARSE¹⁰, to score the structural models. The top 4 submitted unique models are selected based on four different variants of lociPARSE trained independently (i.e., with varying sets of weights and loss functions). The 5th model is chosen based on consensus having the highest aggregated rank based on all four lociPARSE models.

Integrative Protein-RNA Complex Prediction

Our integrative protein-RNA complex prediction pipeline is powered by our newly developed single-sequence method ProRNA3D-single¹¹ based on geometric attention-enabled pairing of biological language models. We employed multiple strategies for the 5 submissions per target. One submission was purely based on biological language models, with embeddings generated from protein and RNA language models, informed by single-sequence-based ESMFold¹² and E2EFold-3D⁵ for predicting the component protein and RNA structures, ultimately predicting the protein-RNA interactions using ProRNA3D-single, followed by restrained optimization using PyRosetta⁸ to generate protein-RNA complex structural models. For the other submissions, we used a customized version of ProRNA3D-single, where AlphaFold3⁷, RoseTTAFold2NA⁶, and RoseTTAFold All-Atom¹³ were employed to predict protein-RNA complex structures. For these predictions, we extracted the component protein and RNA structures and used

ProRNA3D-single for predicting the protein-RNA interactions followed by restrained optimization using PyRosetta⁸ to generate protein-RNA complex structural models by updating the relative coordinates of the components. Additionally, we submitted the AlphaFold3-predicted protein-RNA complex structure as a baseline.

Availability

lociPARSE is freely available at: <https://github.com/Bhattacharya-Lab/lociPARSE>.

ProRNA3D-single is freely available at <https://github.com/Bhattacharya-Lab/ProRNA3D-single>.

1. Wang,W., Feng,C., Han,R., Wang,Z., Ye,L., Du,Z., Wei,H., Zhang,F., Peng,Z. and Yang,J. (2023) trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat Commun*, 14, 7266.
2. Pearce,R., Omenn,G.S. and Zhang,Y. (2022) De Novo RNA Tertiary Structure Prediction at Atomic Resolution Using Geometric Potentials from Deep Learning. 10.1101/2022.05.15.491755.
3. Li,Y., Zhang,C., Feng,C., Pearce,R., Lydia Freddolino,P. and Zhang,Y. (2023) Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nat Commun*, 14, 5745.
4. Kagaya,Y., Zhang,Z., Ibtehaz,N., Wang,X., Nakamura,T., Huang,D. and Kihara,D. (2023) NuFold: A Novel Tertiary RNA Structure Prediction Method Using Deep Learning with Flexible Nucleobase Center Representation. 10.1101/2023.09.20.558715.
5. Shen,T., Hu,Z., Peng,Z., Chen,J., Xiong,P., Hong,L., Zheng,L., Wang,Y., King,I., Wang,S., et al. (2022) E2Efold-3D: End-to-End Deep Learning Method for accurate de novo RNA 3D Structure Prediction. 10.48550/arXiv.2207.01586.
6. Baek,M., McHugh,R., Anishchenko,I., Baker,D. and DiMaio,F. (2022) Accurate prediction of nucleic acid and protein-nucleic acid complexes using RoseTTAFoldNA. 10.1101/2022.09.09.507333.
7. Abramson,J., Adler,J., Dunger,J., Evans,R., Green,T., Pritzel,A., Ronneberger,O., Willmore,L., Ballard,A.J., Bambrick,J., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 10.1038/s41586-024-07487-w.
8. Chaudhury,S., Lyskov,S. and Gray,J.J. (2010) PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics*, 26, 689–691.
9. Zhang,C., Zhang,Y. and Pyle,A.M. (2023) rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *Journal of Molecular Biology*, 435, 167904.
10. Tarafder,S. and Bhattacharya,D. (2024) lociPARSE: a locality-aware invariant point attention model for scoring RNA 3D structures. 10.1101/2023.11.04.565599.
11. Roche,R., Tarafder,S. and Bhattacharya,D. (2024) Single-sequence protein-RNA complex structure prediction by geometric attention-enabled pairing of biological language models. 10.1101/2024.07.27.605468.
12. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., et al. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379, 1123–1130.
13. Krishna,R., Wang,J., Ahern,W., Sturmels,P., Venkatesh,P., Kalvet,I., Lee,G.R., Morey-Burrows,F.S., Anishchenko,I., Humphreys,I.R., et al. (2024) Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*, 384, eadl2528.

RNA Structure modeling with BRIQ2

Heqin Zhu, Ke Chen, Peng Xiong

Suzhou Institute for Advanced Research, University of Science and Technology of China

xiongxp@ustc.edu.cn

Key: *Auto:N; CASP_serv:N; Templ:N; MSA:N; Fragm:N; Cont:N; Dist:N; Tors:N; DeepL:N; EMA:N; MD:Y.*

In CASP16, we tested our RNA tertiary structure prediction method, BRIQ2, using Monte Carlo sampling with a statistical force field. We developed a coarse-grained version of the BRIQ1 force field and optimized its all-atomic counterpart. The sampling process involves two steps: coarse-grained sampling and all-atomic sampling, both using the Monte Carlo method. Before the coarse-grained sampling, we apply various secondary structure prediction tools, such as SPOT-RNA, RNAStructure, UFold, and our newly developed method, bpFold, to obtain the secondary structure as input. However, this pipeline is only effective for small RNAs with sequence lengths of less than 200. For larger RNAs, we use the AlphaFold3 webserver to generate an initial model, followed by structure refinement with BRIQ2 to produce the final model.

1. Xiong, P. et al.(2021). Pairing a high-resolution statistical potential with a nucleobase-centric sampling algorithm for improving RNA model refinement. *Nature Communications* **12**(2777).

Predicting RNA solvation shell using molecular dynamics simulations

E. Posani, O. Languin-Cattoën, and G. Bussi

¹Scuola Internazionale di Studi Avanzati (SISSA) - Trieste, Italy.

bussi@sissa.it

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; EMA:N; MD:Y

RNA structure and dynamics are heavily dependent on the properties of the surrounding solvent, in particular its ionic composition. Most notably, RNA establishes strong interactions with Mg²⁺ cations, for which accurate attribution within published PDB structures is a long-standing challenge.¹ In order to reach maximum accuracy, RNA structure prediction tasks must incorporate implicitly or explicitly the presence of Mg²⁺. We here used molecular dynamics (MD) simulations to predict the solvation shell of the Apo L-21 Scal Tetrahymena ribozyme. We present two closely related predictions. In this one, we used plain MD simulations. In the related one, we used enhanced sampling techniques.

Methods

We tackle the task of local buffer structure ensemble prediction with molecular dynamics (MD) simulations. Simulations were done with GROMACS² 2022.3 patched with PLUMED³ 2.8.1.

Preliminary analysis. Before starting the simulations, we inspected the Cryo-EM structure of the Apo L-21 Scal Tetrahymena ribozyme (PDB: 7EZ0)⁴, focusing on the ions already present in the PDB. Since ions in Cryo-EM experiments are often misplaced/misidentified,¹ we tested the accordance of each ion with the Cryo-EM potential density map (EMBD: emd_31385) related to this structure. We tried to change atom types and occupancies, finding the best accordance with all Mg²⁺ ions, as was deposited initially. Since all 27 ions were already in the optimal site, according to the potential density map, we decided to restrain their positions in the structure during the following simulations.

Topology and force field. The structure of the Apo L-21 Scal Tetrahymena ribozyme (PDB: 7EZ0)⁴ was then sanitized with PDBFixer to add missing heavy atoms. All 27 Mg²⁺ ions present in the original structure were kept. We used the TIP3P⁵ water model and a modified version of the amber14sb.ff.tar.gz force field parameters, corresponding to the standard AMBER force field for RNA⁶⁻⁸ with “microMg” parameters for magnesium from Grotz and Schwierz (2021),⁹ and chloride/sodium parameters from Mamatkulov and Schwierz (2018).¹⁰

Buffer preparation. The structure was solvated in a rectangular box ($\sim 17 \times 12 \times 13 \text{ nm}^3$) and the RNA was neutralized with Na⁺ and Mg²⁺ in a 3-to-1 ratio (including the already present 27 Mg²⁺) so as to match the ion competition observed in experiments.¹¹ Ions were added in the bulk to reach experimental buffer conditions (10 mM MgCl₂, and 50 mM Na-HEPES treated as NaCl). All extra ions (124 Mg,

121 Na, 91 Cl) were added with *gmx genion*, with independent random seeds for each replicate mentioned hereafter.

Equilibration. Energy minimization was done with steepest descent algorithm and the system was equilibrated for 5 ns in the NPT ensemble (1 bar, 300 K, C-rescale barostat¹² and V-rescale thermostat¹³) with positional restraints on the RNA heavy atoms and the 27 PDB Mg²⁺ ions.

Plain MD simulation. Considering that all the Mg²⁺ ions included in PDB 7EZ0 were important to match the density map reported in previous experiments, we decided to perform extensive sampling with the same setup used in the *Equilibration* phase. Specifically, we independently solvated and prepared 16 systems and ran these as replicates for 165 ns each.

Final ensemble. For each of the 16 replicates, 62 equally spaced frames were extracted with a pace of 2.5 ns, skipping the initial 10 ns. This resulted in a total of 992 frames submitted for ensemble assessment.

Availability

Input files and scripts will be made available as part of an upcoming publication of the method's details.

1. Auffinger,P., Ennifar,E., d'Ascenzo,L. (2021) Deflating the RNA Mg²⁺ bubble: stereochemistry to the rescue! *RNA*, 27(3), 243-252.
2. Abraham,M.J., Murtola,T., Schulz,R., Páll,S., Smith,J.C., Hess,B., Lindahl,E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1, 19–25.
3. Tribello,G.A., Bonomi,M., Branduardi,D., Camilloni,C., Bussi,G. (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* 185(2), 604–613.
4. Su,Z., Zhang,K., Kappel,K., Li,S., Palo,M.Z., Pintilie,G.D., Rangan,R., Luo,B., Wei,Y., Das,R., Chiu,W. (2021) Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 angstrom resolution. *Nature* 596, 603–607.
5. Jorgensen,W. L., Chandrasekhar,J., Madura,J.D., Impey,R.W., Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926–935.
6. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W., Kollman,P.A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117(19), 5179–5197.
7. Pérez,A., Marchán,I., Svozil,D., Šponer,J., Cheatham,T.E., Laughton,C.A., Orozco.M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* 92(11), 3817–3829.
8. Zgarbová,M., Otyepka,M., Šponer,J., Mládek,A., Banáš,P., Cheatham,T.E., Jurečka,P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *Chem. Theory Comput.* 7(9), 2886–2902.

9. Grotz,K.K., Schwierz,N. (2021) Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties in SPC/E, TIP3P-fb, TIP4P/2005, TIP4P-Ew, and TIP4P-D. *J. Chem. Theory Comput.* 18(1), 526–537.
10. Mamatkulov,S., Schwierz,N. (2018) Force fields for monovalent and divalent metal cations in TIP3P water based on thermodynamic and kinetic properties. *J. Chem. Phys.* 148(7), 074504.
11. Gebala,M., Herschlag,D. (2019) Quantitative Studies of an RNA Duplex Electrostatics by Ion Counting. *Biophys. J.* 117(6), 1116–1124.
12. Bernetti,M., Bussi,G. (2020) Pressure control using stochastic cell rescaling. *J. Chem. Phys.* 153(11), 114107.
13. Bussi,G., Donadio,D., Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126(1), 014101.

Predicting RNA solvation shell using enhanced sampling molecular dynamics simulations

O. Languin-Cattoën, E. Posani, and G. Bussi

Scuola Internazionale di Studi Avanzati (SISSA) - Trieste, Italy.

bussi@sissa.it

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; EMA:N; MD:Y

RNA structure and dynamics are heavily dependent on the properties of the surrounding solvent, in particular its ionic composition. Most notably, RNA establishes strong interactions with Mg²⁺ cations, for which accurate attribution within published PDB structures is a long-standing challenge.¹ In order to reach maximum accuracy, RNA structure prediction tasks must incorporate implicitly or explicitly the presence of Mg²⁺. We here used molecular dynamics (MD) simulations to predict the solvation shell of the Apo L-21 Scal Tetrahymena ribozyme. We present two closely related predictions. In this one, we used enhanced sampling techniques. In the related one, we used plain MD simulations.

Methods

We tackle the task of local buffer structure ensemble prediction with molecular dynamics (MD) simulations, complemented with enhanced sampling techniques to accelerate the rearrangement of Mg²⁺ coordination shells. Simulations were done with GROMACS² 2022.3 patched with PLUMED³ 2.8.1.

Topology and force field. The structure of the Apo L-21 Scal Tetrahymena ribozyme (PDB: 7EZ0)⁴ was sanitized with PDBFixer to add missing heavy atoms. All 27 Mg²⁺ ions present in the original structure were kept. We used the TIP3P⁵ water model and a modified version of the amber14sb.ff.tar.gz force field parameters, corresponding to the standard AMBER force field for RNA⁶⁻⁸ with “microMg” parameters for magnesium from Grotz and Schwierz (2021),⁹ and chloride/sodium parameters from Mamatkulov and Schwierz (2018).¹⁰

Buffer preparation. The structure was solvated in a rectangular box ($\sim 17 \times 12 \times 13 \text{ nm}^3$) and the RNA was neutralized with Na⁺ and Mg²⁺ in a 3-to-1 ratio (including the already present 27 Mg²⁺) so as to match the ion competition observed in experiments.¹¹ Ions were added in the bulk to reach experimental buffer conditions (10 mM MgCl₂, and 50 mM Na-HEPES treated as NaCl). All extra ions (124 Mg, 121 Na, 91 Cl) were added with *gmx genion*, with independent random seeds for each replica mentioned hereafter.

Equilibration. Energy minimization was done with steepest descent algorithm and the system was equilibrated for 5 ns in the NPT ensemble (1 bar, 300 K, C-rescale barostat¹² and V-rescale thermostat¹³) with positional restraints on the RNA heavy atoms and the 27 PDB Mg²⁺ ions.

RMSD restraint. To restrain RNA orientation in the non-cubic simulation box and to limit RNA conformational dynamics, a soft harmonic bias was applied to the RMSD of the RNA to its native coordinates, computed on a subset of each nucleotide atoms (C1', C2, P) and with simple translational

alignment. A harmonic constant of $200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ was empirically selected to obtain an average RMSD around 4 Å.

Accelerated Mg-binding dynamics. Mg-binding dynamics to the backbone phosphate free oxygen atoms was accelerated using ad hoc biases. Specifically, two-dimensional bias potentials were applied to the space of two collective variables for each Mg atom: its distance from the closest phosphate oxygen atom (O1P, O2P) and the coordination number of Mg to water oxygen atoms (OW). The bias functional form was parameterized on a simple system composed of a diuridine molecule and a single magnesium ion in water. Parameters were chosen by hand to compromise between the following goals: minimize barrier height on the observed reaction pathway; minimize bias away from the barrier and hence maximize overlap with the original ensemble; avoid the generation of new, spurious metastable states.

Replica-exchange phase. Sampling of the canonical distribution of ions was enhanced by coupling the bias-accelerated approach to a Hamiltonian Replica Exchange setup (HREx).^{14,15} A ladder of 16 replica of the system with various bias amplitudes were simulated in parallel, and configurational exchanges between adjacent replicas were accepted or rejected with a Metropolis Monte Carlo criterion. The first 5 replicas were kept bias-free (original force-field) while the following 11 got increasing scaling factors for the bias potential. Spacing between the non-zero scaling factors was optimized to enforce homogeneous acceptance rates.

Relaxation phase. As a trade-off between the limited throughput of the Replica-exchange simulation and the slow diffusion of free ions in bulk water, we interspersed Hamiltonian replica-exchange phases with ordinary Molecular Dynamics for the first 5 (bias-free) replicas. Our method thus alternates costly sampling with accelerated Mg-binding kinetics, and cheaper sampling with traditional MD.

Production run. One sampling round comprised 5 ns of HREx followed by 10 ns of ordinary MD on the unbiased replicas. A total of 12 sampling rounds were simulated on CINECA's Leonardo supercomputer (Bologna, Italy). Each replica used one node with 4 NVidia Ampere GPUs. Two independent replicates of the simulation setup were made, totaling 32 (16 + 16) replicas (all with different initial configurations for bulk ions). Total accumulated simulation time was $2 \times 16 \times 12 \times 15 \text{ ns} = 5760 \text{ ns}$, corresponding to about 300 000 CPU hours.

Final ensemble. For each of the 5 unbiased replicas of the two replicates, 100 equally spaced frames were extracted from the concatenated HREx phases after discarding the first two rounds. This resulted in a total of 1000 frames submitted for ensemble assessment.

Results

Ionic concentrations in the bulk ($> 2 \text{ nm}$ from RNA) quickly equilibrate ($\sim 10 \text{ ns}$) close to their target nominal values. The number of RNA phosphate oxygen atoms bound to magnesium increases with sampling time from 35–40 in the original structure and plateaus at 70–75 after 60 ns of cumulated replica exchange sampling, with similar dynamics in both replicate. For comparison, 170 ns of plain MD did not significantly change that number, with only a tenth of binding/unbinding events.

Availability

Input files and scripts will be made available as part of an upcoming publication of the method.

1. Auffinger,P., Ennifar,E., d'Ascenzo,L. (2021) Deflating the RNA Mg²⁺ bubble: stereochemistry to the rescue! *RNA*, 27(3), 243–252.
2. Abraham,M.J., Murtola,T., Schulz,R., Páll,S., Smith,J.C., Hess,B., Lindahl,E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1, 19–25.
3. Tribello,G.A., Bonomi,M., Branduardi,D., Camilloni,C., Bussi,G. (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* 185(2), 604–613.
4. Su,Z., Zhang,K., Kappel,K., Li,S., Palo,M.Z., Pintilie,G.D., Rangan,R., Luo,B., Wei,Y., Das,R., Chiu,W. (2021) Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 angstrom resolution. *Nature* 596, 603–607.
5. Jorgensen,W. L., Chandrasekhar,J., Madura,J.D., Impey,R.W., Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926–935.
6. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W., Kollman,P.A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117(19), 5179–5197.
7. Pérez,A., Marchán,I., Svozil,D., Šponer,J., Cheatham,T.E., Laughton,C.A., Orozco,M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* 92(11), 3817–3829.
8. Zgarbová,M., Otyepka,M., Šponer,J., Mládek,A., Banáš,P., Cheatham,T.E., Jurečka,P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *Chem. Theory Comput.* 7(9), 2886–2902.
9. Grotz,K.K., Schwierz,N. (2021) Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties in SPC/E, TIP3P-fb, TIP4P/2005, TIP4P-Ew, and TIP4P-D. *J. Chem. Theory Comput.* 18(1), 526–537.
10. Mamatkulov,S., Schwierz,N. (2018) Force fields for monovalent and divalent metal cations in TIP3P water based on thermodynamic and kinetic properties. *J. Chem. Phys.* 148(7), 074504.
11. Gebala,M., Herschlag,D. (2019) Quantitative Studies of an RNA Duplex Electrostatics by Ion Counting. *Biophys. J.* 117(6), 1116–1124.
12. Bernetti,M., Bussi,G. (2020) Pressure control using stochastic cell rescaling. *J. Chem. Phys.* 153(11), 114107.
13. Bussi,G., Donadio,D., Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126(1), 014101.
14. Curuksu,J., Zacharias,M. (2009) Enhanced conformational sampling of nucleic acids by a new Hamiltonian replica exchange molecular dynamics approach. *J. Chem. Phys.* 130(10), 104110.
15. Gil-Ley,A., Bussi,G. (2015) Enhanced Conformational Sampling Using Replica Exchange with Collective-Variable Tempering. *J. Chem. Theory Comput.* 11(3), 1077–1085.

Protein Complex Structure Prediction and Accuracy Estimation Using Machine Learning-based Docking Potential

Myong-Ho Chae

Department of Life Science, University of Sciences, Unjon-District, Pyongyang, DPR Korea

cmh1971@star-co.net.kp

Key: Auto:Y; CASP_serv:Y; TempL:N; MSA:N; DeepL:N; AF:Y; EMA:Y; MD:N

Our protein complex structure prediction and accuracy estimation is based on an optimized protein docking potential (ODP) derived using a single-layer perceptron and docking decoys.

Methods

Docking Potential for complex structure scoring

ODP potential is an upgraded version of neural network-based distance-dependent atom-pair potential for protein docking as described in ref¹. The potential was trained on a protein complex dataset(472 protein complexes) from ref². To train neural network, 4000 decoys of even distance(iRMSD) distribution generated by in-house docking decoy generation program were used for each of 472 protein complexes in the training set.

We used atom-pair distance distributions to evaluate docking geometries. The neural network input information uses type specific atom-pairs (with atoms belonging to different proteins of the complex), which are assigned to different distance classes (bins). The complete distance range that we take into account extends from 0.0 to 8Å and is divided into 12 distance bins.

As a target function we used F_{nat} which is the fraction of native interfacial contacts preserved in the decoy.

Assembly prediction

For phase 1 oligomeric targets (T1xxx homooligomers and H1xxx heterooligomers), phase 0 oligomer prediction models were ranked using the ODP potential and top 5 models were submitted. For phase 2 oligomeric targets (T2xxx homooligomers and H2xxx heterooligomers), phase 1 oligomer prediction models were ranked using the ODP potential and top 5 models were submitted.

Multimer structure model accuracy estimation

For each predicted quaternary structure model, ODP score was calculated to estimate the overall interface accuracy.

For QMODE1/QMODE2, all prediction models of a target oligomer submitted to CASP16 are ranked according to their ODP interface accuracy scores, and a reference model set is constructed from top-scoring N models. Then, the pair-wise similarity score is computed between each model and all models of the reference set using TM-score³ to produce N TM scores. The consensus-based quality score is calculated as the mean of N TM scores. Overall fold accuracy score of a model is the weighted sum of consensus-based quality score and overall interface accuracy score. The size of reference model pool N was set to 21.

Interface reliability score of an interface residue was predicted by a multilayer perceptron using the following features as input.

1. One-hot encoding of a residue.
2. Per-residue ODP potentials and ODP potentials averaged on residues within 8-, 12-, 15- and 20- \AA sphere from a specific residue.
3. Residue contact frequencies between a residue and neighboring residues within 8 \AA distance.
4. Nine amino acid features.

Five-layer perceptron was trained using above 54 features as input to predict the local lddt score for each residue in the model in the training set.

For QMODE3, 8040 MassiveFold models of a target oligomer were ranked using the ODP potential and top 5 models were selected.

Availability

ChaePred is still under development.

1. Chae, M.H., Krull, F., Lorenzen, S., Knapp, E.W.(2010) Predicting protein complex geometries with a neural network. *Proteins*, 78, 1026-1039.
2. Ravikant, D.V., Elber, R. (2010). PIE-efficient filters and coarse grained potentials for unbound protein-protein docking. *Proteins*, 78, 400–419
3. Zhang,Y. Skolnick,J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins*, 57, 702-710.

Tetrahymena Ribozyme RNA Ion and Water Placement Prediction with Comparative Divalent Ion Models

Olivia Fisher, Thomas Cheatham III

Department of Medicinal Chemistry, University of Utah

tec3@utah.edu

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; EMA:N; MD:N

RNA structure is notoriously sensitive to its environment- namely, the presence of solvent, ions, and other biomolecules in the system. Water and ions play an especially important role in RNA folding, dynamics, and structure. Molecular dynamics (MD) simulation of these interactions at an atomistic level has proved difficult without preexisting experimental data. Past work has shown that the selection of the ion parameters plays a critical role in representing accurate dynamics and structural changes of RNA.¹ We selected two commonly used divalent ion models to represent Mg²⁺ in our CASP16 simulations, the Li Merz and the Villa ion parameters.^{2,3} These models were selected because they have previously shown low levels of direct magnesium chelation that often happen during MD simulation, causing structural alterations and trapping the RNA for the length of the trajectory.

Methods

The experimental cryo-EM structure (PDB 7EZ0)⁴ was used for the initial Tetrahymena Ribozyme RNA conformation. The hydrogen atoms and initial Mg²⁺ ions were removed from the file before parametrizing the system. The RNA was described with the Amber OL3 RNA force field parameters⁵ and solvated with a truncated octahedron 10.0 Å box using the OPC water model⁶. The system was neutralized using 194 Mg²⁺ ions and 2 Cl⁻ ions and solvated with 82,494 water molecules. To generate the 10mM MgCl₂ experimental conditions, an additional 18 Mg²⁺ and 35 Cl⁻ ions were added to the system based on the initial volume. Two different parameter sets were used to model the MgCl₂ atoms: the Li Merz divalent ion parameters and the Villa divalent ion parameters. The van der Waal parameters were modified in the resulting topology files according to the Leonard-Jones backbone parameters which increased the radii of the OP1, OP2, O5', and O3' oxygen atoms by 0.0884 Å.^{7,8} Four replicas were run for each system to ensure reproducibility.

Each system was minimized with 1000 steps steepest descent minimization with strong restraints on heavy atoms followed by NTV MD at 300K and SHAKE on hydrogens for 15 ps. Steepest descent minimization was done again first with relaxed restraints on heavy atoms, then minimal restraints, and then no restraints. Lastly, NTP MD was done with SHAKE, first with low restraints and then with minimal restraints on heavy atoms for 5 ps, then with minimal backbone restraints for 10 ps, and finally with no heavy atom restraints for 10 ps. After minimization, hydrogen mass repartitioning was done for both systems, which increased non-solvent hydrogen masses to 3.02 Da by redistributing weight from adjacent heavy atoms, allowing for an increased time step in production (2 fs to 4 fs).⁹

MD production was run with constant pressure and volume, with the temperature set to 300K. The Villa systems ran for a combined total of 2.12 μs, and the Li Merz systems ran for 2.2 μs. Following production,

trajectories from each replica were combined, and then the trajectories from each divalent ion model were clustered according to the K-means algorithm using CPPTRAJ.¹⁰ The clusters with over 1000 frames were selected (6 clusters from Li Merz and 4 from Villa) to represent the most populated conformational states. For CASP submission, 50 frames were extracted evenly across each cluster and saved as PDBs, the hydrogens were removed to save space, and all 500 PDBs were submitted.

1. Bergonzo, C.; Hall, K. B.; Cheatham, T. E. I. Divalent Ion Dependent Conformational Changes in an RNA Stem-Loop Observed by Molecular Dynamics. *J. Chem. Theory Comput.* 2016, 12 (7), 3382–3389. <https://doi.org/10.1021/acs.jctc.6b00173>.
2. Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* 2013, 9 (6), 2733–2748. <https://doi.org/10.1021/ct400146w>.
3. Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion–Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* 2012, 8 (4), 1493–1502. <https://doi.org/10.1021/ct3000734>.
4. Avila, J. R.; Lee, J. S.; Torii, K. U. Co-Immunoprecipitation of Membrane-Bound Receptors. *Arab. Book Am. Soc. Plant Biol.* 2015, 13, e0180. <https://doi.org/10.1199/tab.0180>.
5. Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E. I.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* 2011, 7 (9), 2886–2902. <https://doi.org/10.1021/ct200162x>.
6. Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* 2014, 5 (21), 3863–3871. <https://doi.org/10.1021/jz501780a>.
7. Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.* 2012, 8 (11), 4405–4412. <https://doi.org/10.1021/ct300613v>.
8. Bergonzo, C.; Cheatham, T. E. I. Improved Force Field Parameters Lead to a Better Description of RNA Structure. *J. Chem. Theory Comput.* 2015, 11 (9), 3969–3972. <https://doi.org/10.1021/acs.jctc.5b00444>.
9. Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* 2015, 11 (4), 1864–1874. <https://doi.org/10.1021/ct5010406>.
10. Roe, D. R.; Cheatham, T. E. I. PTraj and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 2013, 9 (7), 3084–3095. <https://doi.org/10.1021/ct400341p>.

Prediction of protein assemblies and ligand binding modes using a combination of ClusPro, AlphaFold and LigTBM diffusion

Ryota Ashizawa², Omeir Khan¹, Sergei Kotelnikov², Maria Lazou, Xiaognang Li, UsmanGhani¹, Dzmitry Padhorny², Dmitri Beglov¹, Sandor Vajda¹, Dima Kozakov^{1,2}

1 - Boston University, 2 - Stony Brook University

midas@laufercenter.org

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

In the latest CASP-CAPRI round, our group generated models of protein assemblies using a combination of baseline and customly trained AlphaFold derived architectures, and docking using the ClusPro webserver.¹⁻³ Here, we will describe the methods used for both generating and ranking ensembles of protein—protein complexes. For ligand modeling we have used ClusPro LigTBM, and diffusion based generalisaiton approach.⁴

Methods

Assembly Prediction

For assembly prediction, our group utilized a two-stage methodology in which we first generate an ensemble of initial models using various AlphaFold (v1,v2,v3) based architectures, including customly trained ones. Briefly, the protocols used for initial model generation are described below:

Docking Alphafold2 Models with Cluspro (AF2+ClusPro): The structure of each chain in the assembly is independently predicted using most confident prediction among different AF architectures. These single-chain predictions are then ranked by the predicted LDDT (pLDDT). The top ranked model for each chain is selected, and low confidence residues ($p\text{LDDT} < 0.50$) are cut from the termini. The trimmed models are then docked using the ClusPro web server.³ All models generated using the “Electrostatic-favored” coefficient set are downloaded and retained for further processing. For antibody and nanobody targets, ClusPro was run in antibody mode.⁵ For homomeric complexes additional models were generated using ClusPro’s multimer docking mode.

The models generated using each of the aforementioned approaches are then refined with Alphafold-Multimer. The refinement stage is dual purpose, as it can not only improve the quality of template models, but also produce a confidence score for each model that can be used for ranking. For refinement, MSAs were prepared for each subunit using the AFMMseqs2 API.⁶

Ligand Docking

We applied the template-based small-molecule docking algorithm ClusPro LigTBM to build the model of the ligand. If no global template was found, LigTBM was extended to consider local templates of binding pockets in PDB structures containing fragments of the candidate metabolite. Instead of searching for fully homologous receptor-ligand pairs, our approach identifies ligand substructures and matching binding

pockets on the target protein surface. In addition to original protocol we have used its torsional diffusion generalization.

Availability

ClusPro and ClusPro LigTBM are available as web servers that are free for academic and governmental use.

1. Evans,R., O'Neill,M., Pritzel,A., Antropova,N., Senior,A., Green,T., Zidek,A., Bates,R., Blackwell,S., Yim,J., Ronnenberger,O., Bodenstein,S., Zielinski,M., Bridgland,A., Potapenko,A., Cowie,A., Tunyasuvunakool,K., Jain,R., Clancy,E., Kohli,P., Jumper,J. & Hassabis,D. (2022). Protein complex prediction with AlphaFold-Multimer. Preprint at Biorxiv.
2. Jumper, J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronnenberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., Back,T., Petersen,S., Reiman,D., Clancy,E., Zielinski,M., Steinegger,M., Pacholska,M., Berghammer,T., Bodenstein,S., Silver,D., Vinyals,O., Senior,A.W., Kavukcuoglu,K., Kogli,P. & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*. 596, 583-589
3. Kozakov,D., Hall,D.R., Xia,B., Porter,K.A., Padhorny,D., Yueh,C., Beglov,D. & Vajda,S (2017). The ClusPro web server for protein-protein docking. *Nat. Protoc.* 12, 255-278.
4. Alekseenko,A., Kotelnikov,S., Ignatov,M., Egbert,M., Kholodov,Y., Vajda,S. & Kozakov,D. (2020). ClusPro LigTBM: Automated Template-based Small Molecule Docking. *J. Mol. Biol.* 432, 3404-3410.
5. Brenke,R., Hall,D.R., Chuang,G-Y., Comeau,S.R., Bohnuud,T., Beglov,D., Schuler-Furman,O., Vajda,S., & Kozakov,D. (2012). Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics*. 28, 2608-2614.
6. Steinegger,M. & Soding,J. (2017). MMseqs2 enabled sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol.* 35, 1026-1028.
7. Mirdita,M., von den Driesch,L., Galiez,C., Martin,M.J., Soding,J. & Steinegger,M. (2016) UniClust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* 45, D170-D176.
8. Hildebrand,A., Remmert,M., Biegert,A. & Soding,J., F (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins*. 77, 128-132.
9. Padhorny,D., Porter,K.A., Ignatov,M., Alekseenko,A., Beglov,D., Kotelnikov,S., Ashizawa,R., Desta,I., Alam,N., Sun,Z., Brini,E., Dill,K., Schueler-Furman,O., Vajda, S. & Kozakov, D. (2020) Cluspro in rounds 38-45 of CAPRI: Toward combining template-based methods with free docking. *Proteins*. 88, 1082-1090

COAST (QA, TS), APOLLO (TS)

COAST: Protein Complex Accuracy Estimation Transformer

Andrew Jordan Siciliano and Zheng Wang

Department of Computer Science, University of Miami

zheng.wang@miami.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:Y; MD:Y

We developed a graph-based deep learning algorithm that predicts qualities of single-model protein-complex structures. Our models predict scores for both the overall fold and overall interface of an input structure. Our server participated in both TS and QA categories. QMODE-1 was largely done through automated scripts with minimal human intervention, while QMODE-3 and TS targets required human intervention.

Methods

For QA we directly predicted the global qualities (QMODE-1 and QMODE-3). We ranked relaxed QMODE-3 models using the predicted overall fold scores from. QMODE-3 was skipped for targets that could not be relaxed in time due to computational issues (approx. > 1000 residues or heavy server usage). For TS predictions we sampled multimeric structures using dropout enabled ESM-Fold¹⁻² with varying gap and linker sizes. For Phase 0 TS targets, we generated structures with different stoichiometry information and submitted the top 5 ranked models according to the predicted overall interface scores from COAST. For Phase 1 TS targets the top 5 ranked models according to predicted overall fold scores from COAST were submitted. Targets that could not be predicted in time (approx. > 1600 residues or heavy server usage) were also skipped. Relaxation was performed using OpenMM³ and the Amber99sb force field⁴ as described and implemented by AlphaFold2⁵⁻⁶.

Our server APOLLO was used in the TS category solely for comparison with ESMFold's internal QA estimates. For Phase 0 TS targets, we generated structures with different stoichiometry information and submitted the top 5 ranked models according to the average pLDTT of interface residues predicted by ESM-Fold. For Phase 1 TS targets we submitted the top 5 ranked models according to the average pLDTT of all residues predicted by ESM-Fold.

Dataset. A total of 3,722 reference native protein complex structures with a resolution < 3Å were fetched from the protein data bank⁷. For each of these structures we predicted models using state of the art tools^{1-2; 6; 8-15}. We used PyRosetta¹⁶ to perturb both the predicted and native structures. Our comprehensive dataset contained a total of 152,512 computationally predicted and perturbed structures. All structures were relaxed. TM-score¹⁷ and QS-best¹⁸ were both calculated using OpenStructure¹⁹, with the respective native structures as reference, and designated as the ground truth values for overall fold and overall interface accuracy, respectively.

Features. Protein complex models are parsed and mapped to a graph which preserves relevant topological and geometrical attributes of the quaternary structure. Each node corresponds to a single amino acid residue. Edges are defined between any two nodes (residues) that have a CB-CB (CA for Glycine) distance $\leq 14\text{\AA}$. Node features include: predictions from NetSurfP3.0²⁰, DisoFlag²¹, and Seqinsite²², energy scores from Rosetta^{16;23}, self-interaction scores calculated using Surfaces²⁴, DSSP²⁵⁻²⁶ annotations

of both the model and predicted AlphaFold2^{5; 9} tertiary structures, and graph kernel scores²⁷⁻³⁰ which compare the model structure to predicted intra and inter-chain structures from AlphaFold2^{5; 9} and CDPred³¹, respectively. AlphaFold2^{5; 9} tertiary structures were predicted using ColabFold⁹ and were used along with the MSA's produced by ColabFold's MMSeqs2 Server^{8; 13} as the input to CDPred³¹. We also developed custom positional node features based upon the relative distances and angles from each residue to the center of mass of the protein and respective chain. Edge features include: a binary variable indicating intra vs. inter-chain edges, distances and angles parsed from the model structure, distance and angle predictions from AlphaFold2^{5; 9} (tertiary), distance and contact predictions from CDPred³¹ (dimeric), and energy scores from Rosetta^{16; 23} and Surfaces²⁴.

Model Architecture. Our model architecture pipeline starts by processing the entire input graph (protein model) using a global attention layer, with each node having access to every other node's features irrespective of the edges. As information progresses through the model nodes are dropped via learned projection scores and updated via neighborhood aggregations. The input graph (protein model) is reduced, refined, and aggregated to a single output prediction that reflects the quality of the input protein model. To achieve this our model utilized a GPSConv³² layer, TransformerConv³³ layers, Top-K Pooling³⁴ layers, and a Global Attention Pooling³⁵ layer.

1. Lin, Zeming, et al. "Evolutionary-scale prediction of atomic-level protein structure with a language model." *Science* 379.6637 (2023): 1123-1130.
2. Rives, A., et al., Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 2021. 118(15): p. e2016239118.
3. Eastman, P. et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Computational Biology* 13, e1005659.
4. Hornak, V. et al. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3), 712-725.
5. Jumper, J. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
6. Evans, Richard, et al. "Protein complex prediction with AlphaFold-Multimer." *biorxiv*: 2021-10.
7. Berman, H. M. (2000). The Protein Data Bank. In *Nucleic Acids Research* (Vol. 28, Issue 1, pp. 235–242). Oxford University Press (OUP).
8. Mirdita, M., Steinegger, M., & Söding, J. (2019). MMseqs2 desktop and local web server app for fast, interactive sequence searches. In J. Hancock (Ed.), *Bioinformatics* (Vol. 35, Issue 16, pp. 2856–2858). Oxford University Press (OUP).
9. Mirdita, Milot, et al. "ColabFold: making protein folding accessible to all." *Nature methods* 19.6 (2022): 679-682.
10. Li, Z., Liu, X., Chen, W., Shen, F., Bi, H., Ke, G., & Zhang, L. (2022). Uni-Fold: An Open-Source Platform for Developing Protein Folding Models beyond AlphaFold.
11. Li, Z., Yang, S., Liu, X., Chen, W., Wen, H., Shen, F., Ke, G., & Zhang, L. (2022). Uni-Fold Symmetry: Harnessing Symmetry in Folding Large Protein Complexes.
12. Baek, M., Anishchenko, I., Humphreys, I. R., Cong, Q., Baker, D., & DiMaio, F. (2023). Efficient and accurate prediction of protein structure using RoseTTAFold2. Cold Spring Harbor Laboratory.
13. Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, 35(11), 1026-1028.

14. Villegas-Morcillo, Amelia, et al. "ManyFold: an efficient and flexible library for training and validating protein folding models." *Bioinformatics* 39.1 (2023): btac773.
15. Wu, Ruidong, et al. "High-resolution de novo structure prediction from primary sequence." *BioRxiv* (2022): 2022-07.
16. Chaudhury, S., Lyskov, S., & Gray, J. J. (2010). PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics*, 26(5), 689-691
17. Zhang, Yang, and Jeffrey Skolnick. "Scoring function for automated assessment of protein structure template quality." *Proteins: Structure, Function, and Bioinformatics* 57.4 (2004): 702-710.
18. Bertoni, Martino, et al. "Modeling protein quaternary structure of homo-and hetero-oligomers beyond binary interactions by homology." *Scientific reports* 7.1 (2017): 10480.
19. Biasini, Marco, et al. "OpenStructure: an integrated software framework for computational structural biology." *Acta Crystallographica Section D: Biological Crystallography* 69.5 (2013): 701-709
20. Høie, Magnus Haraldson, et al. "NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning." *Nucleic acids research* 50.W1 (2022): W510-W515
21. Pang, Y., & Liu, B. (2024). DisoFLAG: accurate prediction of protein intrinsic disorder and its functions using graph-based interaction protein language model. *BMC biology*, 22(1), 3.
22. Hosseini, S., Golding, G. B., & Ilie, L. (2024). Seq-InSite: sequence supersedes structure for protein interaction site prediction. *Bioinformatics*, 40(1), btad738
23. Alford, Rebecca F., et al. "The Rosetta all-atom energy function for macromolecular modeling and design." *Journal of chemical theory and computation* 13.6 (2017): 3031-3048.
24. Teruel, N., Borges, V. M., & Najmanovich, R. (2023). Surfaces: a software to quantify and visualize interactions within and between proteins and ligands. *Bioinformatics*, 39(10), btad608
25. Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* (Vol.22, Issue 12, pp. 2577–2637). Wiley.
26. Touw, Wouter G., et al. "A series of PDB-related databanks for everyday needs." *Nucleic acids research* 43.D1 (2015): D364-D368.
27. Borgwardt, Karsten M., and Hans-Peter Kriegel. "Shortest-path kernels on graphs." *Fifth IEEE international conference on data mining (ICDM'05)*. IEEE, 2005.
28. Siglidis, G., Nikolenz, G., Limnios, S., Giatsidis, C., Skianis, K., & Vazirgiannis, M. (2020). Grakel: A graph kernel library in python. *Journal of Machine Learning Research*, 21(54), 1-5.
29. Weisfeiler, B., & Leman, A. (1968). The reduction of a graph to canonical form and the algebra which appears therein. *nti, Series*, 2(9), 12-16.
30. Shervashidze, N., Schweitzer, P., Van Leeuwen, E. J., Mehlhorn, K., & Borgwardt, K. M. (2011). Weisfeiler-lehman graph kernels. *Journal of Machine Learning Research*, 12(9).
31. Guo, Z., Liu, J., Skolnick, J., & Cheng, J. (2022). Prediction of inter-chain distance maps of protein complexes with 2D attention-based deep neural networks. In *Nature Communications* (Vol. 13, Issue 1). Springer Science and Business Media LLC.
32. Rampášek, Ladislav, et al. "Recipe for a general, powerful, scalable graph transformer." *Advances in Neural Information Processing Systems* 35 (2022): 14501-14515.
33. Shi, Yunsheng, et al. "Masked label prediction: Unified message passing model for semi-supervised classification." *arXiv preprint arXiv:2009.03509* (2020).
34. Gao, H., & Ji, S. (2019, May). Graph u-nets. In *international conference on machine learning* (pp. 2083-2092). PMLR.
35. Li, Y., Tarlow, D., Brockschmidt, M., & Zemel, R. (2015). Gated graph sequence neural networks. *arXiv preprint arXiv:1511.05493*.

Benchmarking ColabFold in CASP16

Dongwook Kim^{1,2}, Sergey Ovchinnikov², Martin Steinegger^{1,2,4,5} and Milot Mirdita²

1 - Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, 08826, South Korea, 2 - School of Biological Sciences, Seoul National University, Seoul, 08826, South Korea, 3 - Massachusetts Institute of Technology, Cambridge, MA 02139, USA, 4 - Artificial Intelligence Institute, Seoul National University, Seoul, 08826, South Korea, 5 - Institute of Molecular Biology and Genetics, Seoul National University, Seoul, 08826, South Korea

so3@mit.edu, martin.steinegger@snu.ac.kr, mmirdit@snu.ac.kr

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:Y; AF:Y

ColabFold [1] is a protein structure prediction platform that exposes accessible and powerful notebooks and command line tools for leading structure prediction models. ColabFold-AlphaFold2 is the most widely used instance that extends the groundbreaking AlphaFold2 and AlphaFold-multimer models with performance enhancements and exposes various advanced parameters. Its largest diversion from the base AlphaFold2 is the revamped MMseqs2-based [2] multiple sequence alignment (MSA) stage offering orders of magnitude acceleration at matching prediction quality. We previously found [3] that MSAs enriched with SRA-mined protein sequences contribute to a large increase in prediction quality for difficult targets. However, searching through petabytes of data is time- and resource-consuming, thus we built large clustered metagenomic databases for monomer- as well as for the first time a metagenomic database suited for taxonomic pairing necessary for multimer-predictions.

Methods

We registered two servers for CASP16: colabfold_baseline and colabfold. Both used identical parameters and software versions (ColabFold commit 57b220e) except as highlighted in the following. For monomer prediction, we used the AlphaFold2-ptm model [4] and for multimeric protein structure prediction we used the AlphaFold-multimer-v2 model [5]. The MSAs were generated following the protocol described in [1]. The sequence and template databases we used were UniRef100 2302, ColabFoldDB 202108, and PDB70 230517. The colabfold server replaces the ColabFoldDB with a newer version (colabfold_envdb_2023) and further includes an additional metagenomic pairing database (spire_ctg10_2401_db; described below). For the AlphaFold runs, we used 12 recycles, enabled templates and GPU-based relaxation. For the longest sequences whose memory requirements exceeded the VRAM available on Nvidia A5000 GPUs, we manually transferred the automatically generated job script to an Nvidia H100 based system and executed it with automatic prediction submission upon completion. Manual parameter tweaks were additionally required for some of these large models to fit them into system memory or within the time limit (reduced recycles, or disabled relaxation). We further registered a manual-intervention group (colabfold_human), where we submitted predictions for only one target. Here, we exceeded the time-limit for the automatic submission by approximately a day. The predictions were ranked using the default prediction confidence metrics. For monomers, this was the predicted LDDT (pLDDT) and for multimers a combination of predicted interface and overall TMscore (ipTM and pTM).

Metagenomic databases update for MSAs. We updated the ColabFoldDB to include clustered representatives from over 35 billion protein sequences by integrating Serratus [6] and SPIRE [7]. We removed all clusters with two or fewer members from the resulting database, followed by redundancy filtering to keep only the 10 most diverse sequences for each cluster. This database was converted to a ColabFold expandable search database, resulting in 169 million representatives and over a 1.3 billion members, thus nearly doubling sequence count and average cluster size over the previous ColabFoldDB. The SPIRE database includes information from which contig each protein originates, thus offering us the possibility to retrieve additional metagenomic co-occurrence information for multimeric predictions. We kept all proteins from contigs containing at least 10 proteins and created a ColabFold expandable search database, resulting in 100 million representatives and over 3.5 billion members in total.

Results

colabfold_baseline: Mean pLDDT of all predicted targets (rank1 model) is 81.36 (1st Quantile: 77.13, Median: 84.4, 3rd Quantile 88.6). For heteromeric targets (H prefix) the mean ranking metric ($0.8 * \text{ipTM} + 0.2 * \text{pTM}$) was 0.622 (1st Quantile: 0.475, Median: 0.631; 3rd Quantile: 0.796);

colabfold: Mean pLDDT of all predicted targets is 81.4 (1st Quantile: 75.5, Median: 84.1, 3rd Quantile 88.6). For heteromeric targets the mean ranking metric was 0.62 (1st Quantile: 0.48, Median: 0.64; 3rd Quantile: 0.80).

Availability

ColabFold is free and open source software that can be installed locally from <https://github.com/sokrypton/ColabFold> or used online with a web browser through Google Colab at <https://colabfold.com>. The ColabFold databases can be found at <https://colabfold.mmsegs.com>. Submitted predictions, including MSAs and confidence metrics, were uploaded immediately after submission to CASP16 to <https://casp16.colabfold.com>.

1. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, 19(6), 679-682.
2. Steinegger, M. & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* 35, 1026–1028.
3. Lee, S., Kim, G., Karin, E.L., Mirdita, M., Park, S., Chikhi, R., Babaian, A., Kryshtafovych, A., and Steinegger, M. (2024) Petabase-Scale Homology Search for Structure Prediction. *Cold Spring Harbor Perspectives in Biology*, 16, a041465.
4. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
5. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A. W., Green, T., ... & Hassabis, D. (2021). Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.10.04.463034v2.
6. Edgar, R.C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, ... & Babaian, A. (2022) Petabase-scale sequence alignment catalyses viral discovery. *Nature*, 602, 142–147.
7. Schmidt, T.S.B., Fullam, A., Ferretti, P., Orakov, A., Maistrenko, O.M., Ruscheweyh, H.-J., ... & Bork, P. (2024) SPIRE: a Searchable, Planetary-scale mIcrobiome REsource. *Nucleic Acids Research*, 52, D777–D783.

Prediction of Local Solvent Shell Structure of Tetrahymena Ribozyme RNA Using Molecular Dynamics

Karim Malekzadeh, Ayush Gupta, Güл H. Zerze*

*William A. Brookshire Department of Chemical and Biomolecular Engineering,
University of Houston, Houston, Texas 77204, USA*

gzerze@uh.edu

Keys: Auto:N; CASP_serv:N; Templ:N; MSA:N; AF:N; DeepL:N; EMA:N; MD:Y.

Molecular dynamics simulations are widely used to study various biological phenomena on the length and timescales often inaccessible by experiments.¹⁻³ An all-atom molecular dynamics (MD) simulation treats atoms as classical particles, where the position and velocity of every atom in the system evolve according to Newton's equations of motion. The forces acting on the atoms are computed using an energy function, known as a force field model, which is typically designed using a combination of first-principles physics and empirical fitting to quantum mechanical and experimental data. In this study, we used all-atom MD simulation to predict water and ions' structure and local motion around the Tetrahymena ribozyme RNA molecule.

Method

The starting structure for the Tetrahymena ribozyme RNA was taken from the Protein Data Bank (PDB) entry 7EZ0. This PDB contained 27 Mg²⁺ ions around the RNA molecule. We maintained Mg²⁺ ions at positions given in the PDB file and added the corresponding number of Cl⁻ ions (54) to adjust the salt as MgCl₂. We modeled the RNA sequence using the nucleic acid force field DESRES⁴ and combined it with the TIP4PD⁶ water model using the library files provided by Kührova et al.⁵ A single copy of the RNA molecule was solvated in a cuboidal box of volume 2060 nm³. Simulation input files were generated using AmberTools⁸ and then converted to a format compatible with the simulation suite GROMACS¹⁰. 386 Na⁺ ions were added to maintain electroneutrality. The ions were modeled using CHARMM22 parameters⁷. After solvation and salt addition, we minimized the energy of the system using the steepest descent algorithm. The simulation box was then equilibrated with 100 ps NVT simulation (T = 300 K) followed by 2 ns NPT simulation (T = 300 K, P = 1 bar). Atmospheric pressure (1 bar) was maintained using an isotropic Parrinello-Rahman barostat⁹ with a time constant of 2 ps, while the temperature (300 K) was maintained using v-rescale thermostat with a time constant of 1 ps. In both the equilibration steps, the positions of all the heavy atoms of RNA were restrained using a force constant of 500 kJ/(mol nm²). We used GROMACS 2023.5¹⁰ to perform all the production runs at 300 K and 1 bar, with a time step of 2 fs. Electrostatic interactions were calculated using the particle-mesh Ewald method¹¹ with a real space cutoff distance of 1 nm. A cutoff distance of 1 nm was also used for the van der Waals interactions. In the production steps, the positions of all the heavy atoms of RNA were restrained using a force constant of 250 kJ/(mol nm²). The simulation was run on 4 nodes having 2 GPUs and 28 CPU cores per node.

Results

After running an MD simulation for 247 ns, we computed the root-mean-square deviation (RMSD) of the heavy atoms in the RNA structure. All the sampled frames had an RMSD value of less than 0.4 nm from the native structure, indicating structural stability throughout the simulation. In total, 248 distinct models were reported, corresponding to each nanosecond from 0 to 247 ns of simulation time.

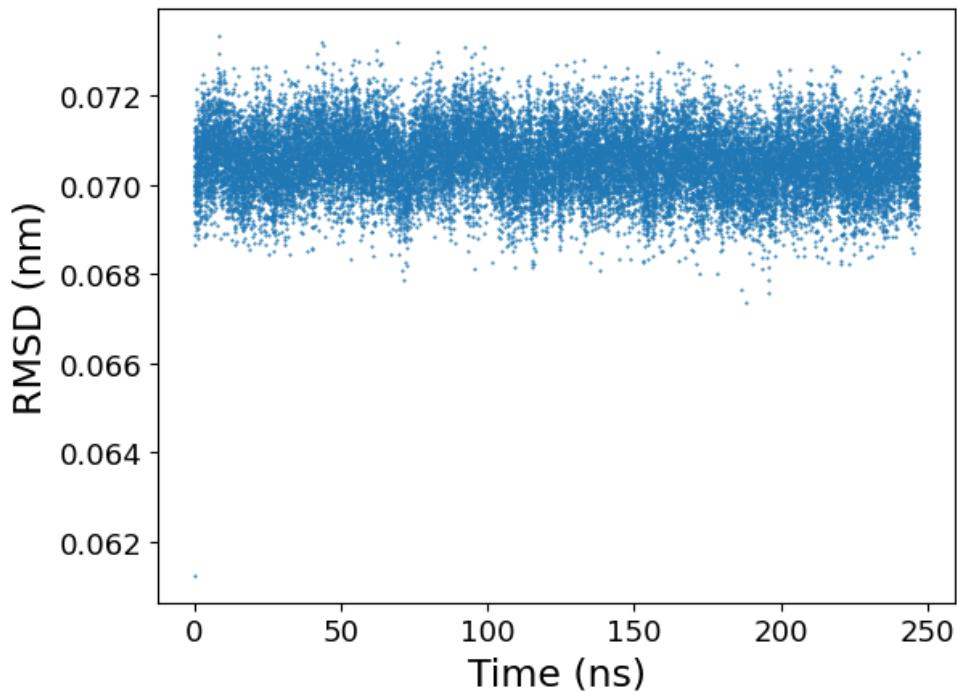


Figure 1: Root-mean-square deviation (RMSD) of the heavy atoms of the RNA molecule as a function of time during production run. The positions of all the heavy atoms of RNA were restrained using a force constant of 250 kJ/(mol nm²) in this simulation.

Availability

The input files used to perform the simulation are provided in this zip file: [simulation077.zip](#)

1. Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular simulation: a computational microscope for molecular biology. *Annual review of biophysics*, 41(1), 429-452.
2. Bottaro, S., & Lindorff-Larsen, K. (2018). Biophysical experiments and biomolecular simulations: A perfect match?. *Science*, 361(6400), 355-360.
3. Lane, T. J., Shukla, D., Beauchamp, K. A., & Pande, V. S. (2013). To milliseconds and beyond: challenges in the simulation of protein folding. *Current opinion in structural biology*, 23(1), 58-65.

4. Tan, D., Piana, S., Dirks, R. M., & Shaw, D. E. (2018). RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proceedings of the National Academy of Sciences*, 115(7), E1346-E1355.
5. Kuhrova, P., Mlynsky, V., Zgarbová, M., Krepl, M., Bussi, G., Best, R. B., ... & Banas, P. (2019). Improving the performance of the amber RNA force field by tuning the hydrogen-bonding interactions. *Journal of chemical theory and computation*, 15(5), 3288-3305.
6. Piana, S., Donchev, A. G., Robustelli, P., & Shaw, D. E. (2015). Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *The journal of physical chemistry B*, 119(16), 5113-5123.
7. MacKerell Jr, A. D., Bashford, D., Bellott, M. L. D. R., Dunbrack Jr, R. L., Evanseck, J. D., Field, M. J., ... & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*, 102(18), 3586-3616.
8. Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., ... & Merz Jr, K. M. (2023). AmberTools. *Journal of chemical information and modeling*, 63(20), 6183-6191.
9. Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12), 7182-7190.
10. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., ... & Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7), 845-854.
11. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of chemical physics*, 103(19), 8577-8593.

Prediction of Local Solvent Shell Structure of Tetrahymena Ribozyme RNA Using Molecular Dynamics

Karim Malekzadeh, Ayush Gupta, Güл H. Zerze*

*William A. Brookshire Department of Chemical and Biomolecular Engineering,
University of Houston, Houston, Texas 77204, USA*

gzerze@uh.edu

Keys: Auto:N; CASP_serv:N; Templ:N; MSA:N; AF:N; DeepL:N; EMA:N; MD:Y.

Molecular dynamics simulations are widely used to study various biological phenomena on the length and timescales often inaccessible by experiments.¹⁻³ An all-atom molecular dynamics (MD) simulation treats atoms as classical particles, where the position and velocity of every atom in the system evolve according to Newton's equations of motion. The forces acting on the atoms are computed using an energy function, known as a force field model, which is typically designed using a combination of first-principles physics and empirical fitting to quantum mechanical and experimental data. In this study, we used all-atom MD simulation to predict water and ions' structure and local motion around the Tetrahymena ribozyme RNA molecule.

Method

The starting structure for the Tetrahymena ribozyme RNA was taken from the Protein Data Bank (PDB) entry 7EZ0. This PDB contained 27 Mg²⁺ ions around the RNA molecule. We maintained Mg²⁺ ions at positions given in the PDB file and added the corresponding number of Cl⁻ ions (54) to adjust the salt as MgCl₂. We modeled the RNA sequence using the nucleic acid force field DESRES⁴ and combined it with the TIP4PD⁶ water model using the library files provided by Kührova et al.⁵ A single copy of the RNA molecule was solvated in a cuboidal box of volume 2060 nm³. Simulation input files were generated using AmberTools⁸ and then converted to a format compatible with the simulation suite GROMACS¹⁰. 386 Na⁺ ions were added to maintain electroneutrality. The ions were modeled using CHARMM22 parameters⁷. After solvation and salt addition, we minimized the energy of the system using the steepest descent algorithm. The simulation box was then equilibrated with 100 ps NVT simulation (T = 300 K) followed by 2 ns NPT simulation (T = 300 K, P = 1 bar). Atmospheric pressure (1 bar) was maintained using an isotropic Parrinello-Rahman barostat⁹ with a time constant of 2 ps, while the temperature (300 K) was maintained using v-rescale thermostat with a time constant of 1 ps. In both the equilibration steps, the positions of all the heavy atoms of RNA were restrained using a force constant of 500 kJ/(mol nm²). We used GROMACS 2023.5¹⁰ to perform all the production runs at 300 K and 1 bar, with a time step of 2 fs. Electrostatic interactions were calculated using the particle-mesh Ewald method¹¹ with a real space cutoff distance of 1 nm. A cutoff distance of 1 nm was also used for the van der Waals interactions. The simulation was run on 4 nodes having 2 GPUs and 28 CPU cores per node.

Results

After running a 193 ns simulation, we calculated the root-mean-square deviation (RMSD) of the heavy atoms in the RNA. Frames with an RMSD above 0.4 nm were excluded, and the first 571 out of 2176 frames(models) with RMSD values below 0.4 nm were submitted. The remaining frames were discarded due to the extensive translational movement, which rendered them unsuitable for predicting the surrounding solvent shell.

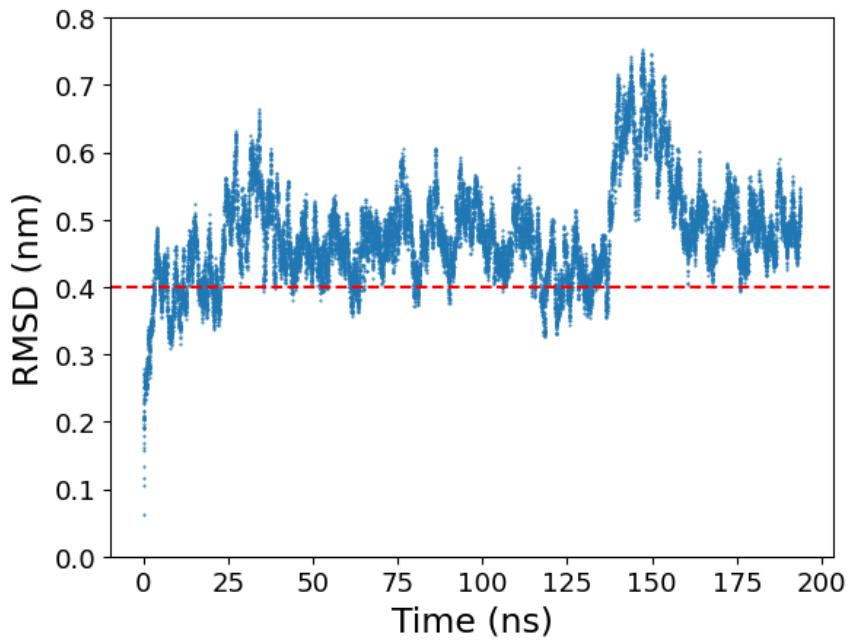


Figure 2: Root-mean-square deviation (RMSD) of the heavy atoms of the RNA molecule as a function of time during production run. The horizontal dashed line at RMSD = 0.4 nm marks the cutoff above which the frames are excluded from submission.

Availability

The input files used to perform the simulation are provided in this zip file: [simulation466.zip](#)

1. Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular simulation: a computational microscope for molecular biology. *Annual review of biophysics*, 41(1), 429-452.
2. Bottaro, S., & Lindorff-Larsen, K. (2018). Biophysical experiments and biomolecular simulations: A perfect match?. *Science*, 361(6400), 355-360.
3. Lane, T. J., Shukla, D., Beauchamp, K. A., & Pande, V. S. (2013). To milliseconds and beyond: challenges in the simulation of protein folding. *Current opinion in structural biology*, 23(1), 58-65.
4. Tan, D., Piana, S., Dirks, R. M., & Shaw, D. E. (2018). RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proceedings of the National Academy of Sciences*, 115(7), E1346-E1355.
5. Kuhrova, P., Mlynsky, V., Zgarbová, M., Krepl, M., Bussi, G., Best, R. B., ... & Banas, P. (2019).

- Improving the performance of the amber RNA force field by tuning the hydrogen-bonding interactions. *Journal of chemical theory and computation*, 15(5), 3288-3305.
- 6. Piana, S., Donchev, A. G., Robustelli, P., & Shaw, D. E. (2015). Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *The journal of physical chemistry B*, 119(16), 5113-5123.
 - 7. MacKerell Jr, A. D., Bashford, D., Bellott, M. L. D. R., Dunbrack Jr, R. L., Evanseck, J. D., Field, M. J., ... & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*, 102(18), 3586-3616.
 - 8. Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., ... & Merz Jr, K. M. (2023). AmberTools. *Journal of chemical information and modeling*, 63(20), 6183-6191.
 - 9. Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12), 7182-7190.
 - 10. Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., ... & Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7), 845-854.
 - 11. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of chemical physics*, 103(19), 8577-8593.

Structure Prediction and Selection Method of Cool-PSP

Panpan Wang

Individual group

wangpanpan2992@163.com

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaSource.TARA; Cont:Y; DeepL:Y; EMA:Y; MD:Y.

Our model is trained by feeding self-prepared data into the framework of AF2 with precisely adjusted hyperparameters [1]. Under most conditions, we need only take the target sequence as the input, then, the structure prediction is performed simultaneously, also providing us with a ranked result by the built-in scoring function, meanwhile, the structure is refined based on the Amber force field [2]. While, in order to further improve the prediction and selection accuracy, we also made some modifications, including replacing the default MSA with that is derived from self-developed multiple sequence alignment algorithm, assigning the specific template, offering binding sites by literature reviewing, optimizing the score function and model selection strategy.

Methods

1. MSA

In order to guarantee the diversity and integrity of MSA, we developed a new algorithm to perform sequence alignment, which aims to find the best alignment sequence to query sequence through multiple sequence search tools for cascading search. Based on this framework, we further introduced self-built metagenomic clustering data, like MetaClust, TARA, MetaSource and JGI. Once targeting the sequence for individual chain, the next step is pairing them and forming an intact complex, while assembling, the species information is the priority, then, the sequence similarity.

2. Template

As the template is identified based on the result of MSA, the precise MSA determines the high quality of template. Although this could improve the result of monomeric template search significantly, it doesn't perform well when encountering multimeric template. Owing to the template is determined separately for each chain according to the sequence similarity, the interaction between chains is ignored. To rescue the information loss and capture the inter-chain interactions, we developed a new strategy to find the multimeric template. Firstly, each chain is aligned, among this, the pdb id is recorded, then, by calculating the distance between each two chains in the same pdb to estimate if these chains are indeed within a complex, finally, only the actual complex with the highest average similarity is maintained and treated as the final template.

3. Model training

Our model is trained based on the framework of AF2 but increasing the training data by 200%, which is achieved by data augmentation. We built quantities of pseudo-multimer data by modifying the multi-domain proteins from both native monomer and multimer protein structures. According to the annotation in CATH, SCOP and ECOD, the position of each domain is assigned and these domains with freeSASA larger than 1000 Å² are assembled into a pseudo-multimer. Besides, we also optimize the loss function and some specific hyperparameters while training.

4. Model selection

In our pipeline, more than 300 models are generated for each target, both the default score and diversity are the criteria for model selection. In order to distinguish models in different conformation, we applied the cluster method before ranking with the values of score function. Briefly, every two models are paired, the TMscore between them is calculated, then, all of these models are divided into groups according to the TMscore, finally, the model with highest score is selected.

5. Recycle

To ensure the model selected is indeed in high accuracy, we could take it as the assigned template and perform structure prediction for three cycles. In each cycle, the template is composed of the original one and the one with highest score in the last round. If the input is indeed a model in high quality, we supposed that it wouldn't be changed largely.

6. Manual intervention

For some cases, especially those with barely homologs, it's difficult for the prediction model to provide a credible result. Under this circumstance, some extra information is introduced manually, like the experimental determined binding sites, symmetric parameters or native stoichiometry, etc.

Results

To evaluate the result of our self-trained model, we test it on the CASP15 and experimentally solved 100 multimer cases, which are out of the training dataset. Statistically, the overall TMscore increased by 10% in monomer and 3% in multimer compared to AF2 and AF3 (Figure 1a). Here is a showcase of our model (Figure 1b), as shown in the figure, the Cool-PSP model surpasses AF2 and AF3 greatly, which is mainly a result of accurate improvement in the multimer interface identification.

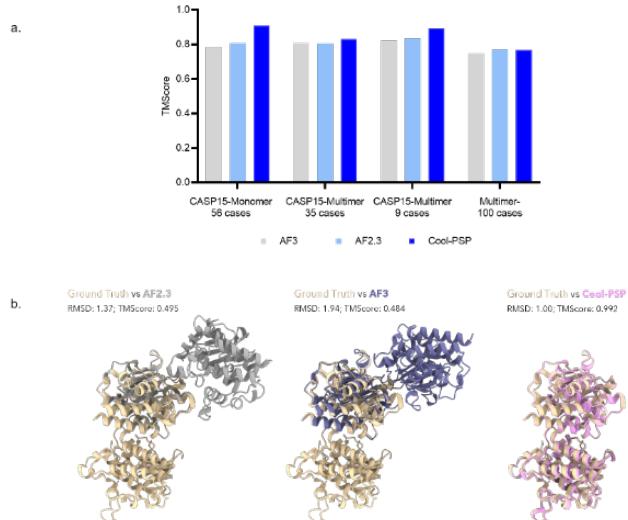


Figure 1. The result of Cool-PSP model. a. The average TMscore of three models on CASP15 monomer, CASP15-multimer and multimer-100 test databases. b. Showcase of three models on PDB 7XDY.

Availability

By now, our method is still under development and we plan to integrate each module mentioned above into an end-to end model, which will not be released until the completion.

1. Z. Yang, X. Zeng, Y. Zhao, R. Chen, AlphaFold2 and its applications in the fields of biology and medicine, *Signal Transduction and Targeted Therapy* 8 (2023) 115.
2. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins* 65 (2006) 712-725.

CSSB_experimental (TS, Assembly)

Improved MSA and New Deep Learning Model for Protein Structure Prediction in CASP16

Hyunho Shin[#], Sanggeun Park[#], Heesoo Ki and Minkyung Baek

School of Biological Sciences, Seoul National University, Seoul, 08826, Republic of Korea
- equal contribution

minkbaek@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

In CASP16, we developed UltimateMSA that generates multiple sequence alignment (MSA) by integrating multiple sequence databases, and an in-house deep learning protein structure prediction model combining elements of RoseTTAFold2¹ and AlphaFold-multimer². The overall pipeline is illustrated in Figure 1, and it was applied to all targets except antibodies. For antibody predictions, please refer to the CSSB_experimental (Antibody) abstract.

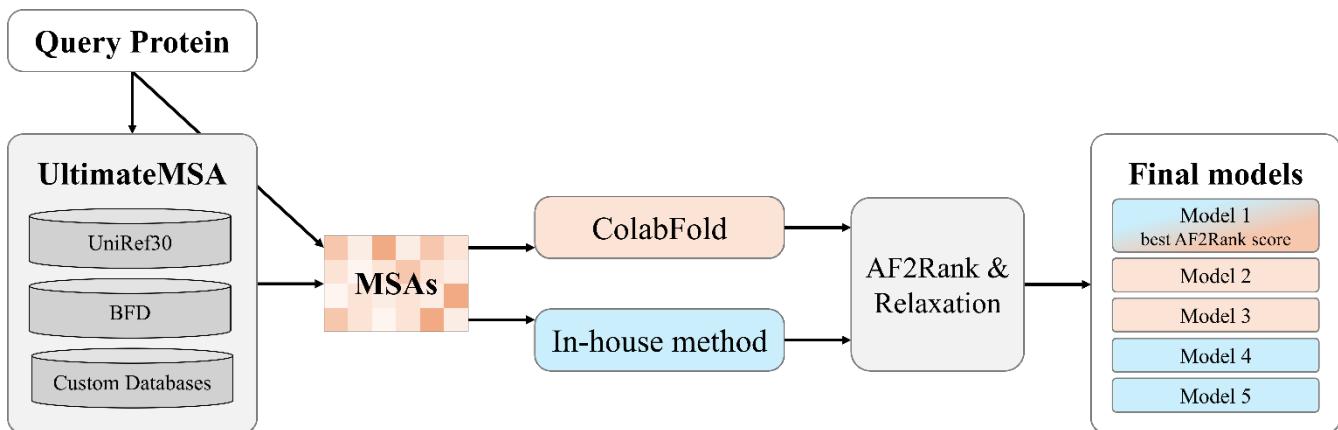


Figure 1. Schematic diagram of overall pipeline.

Methods

Phase 0 and 1. We developed UltimateMSA, a tool that generates MSAs by integrating data from UniRef30³, BFD^{4,5}, and custom databases built from NCBI BLAST⁶ searches against target sequences across multiple databases (nr, tsa_nr, and env_nr). MSAs were constructed using HHblits⁷ and MMseqs2⁸, followed by structural predictions using ColabFold⁹ or our in-house protein structure prediction method. For multi-domain proteins, we searched MSAs for each domain and integrated them with the original MSAs. For viral proteins, we used sequences from the UniProtKB/Swiss-Prot viral protein database (uniprot_sprot_vir70) and the NCBI virus database¹⁰ to improve MSA quality.

We also developed an in-house protein structure prediction model, a 3-track AI model incorporating elements from RoseTTAFold2 (RF2) and AlphaFold-Multimer. This model retains the core 3-track architecture of RF2. In the MSA module, we implemented more efficient column-wise attention to mitigate memory limitations. Additionally, we replaced the SE(3)-Transformer¹¹ with a modified Invariant Point Attention (IPA) mechanism¹², which calculates attention based on pair embeddings only. We refined the model's training objective by adding the Locally Aligned Displacement Error (LADE) and

Chain Aligned Displacement Error (ChADE), smoother alternatives to the FAPE loss used in AlphaFold and RoseTTAFold2 training^{1,12}. To better utilize experimental data found in literature, we integrated an interaction embedding module that incorporates both positive and negative interaction data.

With a few exceptions, we submitted structures generated by UltimateMSA using ColabFold for models 2 and 3, while for models 4 and 5, we submitted models from our in-house model. For model 1, the structure with the highest AF2Rank¹³ score from these two methods was submitted.

Phase 2. Based on various metrics provided with the MassiveFold structures, we filtered out low-confidence structures. The remaining structures were ranked using AF2Rank. From the top-ranked structures, 10 structures with meaningful structural diversity, as determined by TM-score¹⁴, were selected. Each of these selected structures were used as a template, from which five structures were sampled using our in-house protein structure prediction model in single-sequence mode. Out of the resulting 50 structures, 20 were chosen based on our model's confidence score. These structures were then relaxed using Rosetta FastRelax¹⁵ and re-ranked with AF2Rank. For assembly targets, we applied a composite score of AF2Rank multiplied by iPTM, while for monomer targets, only the AF2Rank composite score was used.

Acknowledgement

This work was supported by IITP/MSIT (RS-2023-00220628), NRF/MSIT (RS-2023-00210147).

1. Baek, M., Anishchenko, I., Humphreys, I.R., Cong, Q., Baker, D., DiMaio, F. (2023). Efficient and accurate prediction of protein structure using RoseTTAFold2. *BioRxiv*, 2023-05.
2. Evans, Richard, et al. (2021). Protein complex prediction with AlphaFold-Multimer. *BioRxiv*, 2021-10.
3. Suzek,B.E., Wang,Y., Huang,H., McGarvey,P.B., Wu,C.H., Consortium,U. (2014) UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926932.
4. Steinegger, M., Mirdita, M., Söding, J. (2019). Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods* **16**(7), 603-606.
5. Steinegger, M., and Söding, J. (2018). Clustering huge protein sequence sets in linear time. *Nature Communications* **9**(1), 1-8.
6. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W., Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
7. Remmert, M., Biegert, A., Hauser, A. (2012) HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment, *Nature Methods* **9**(2), 173-175.
8. Steinegger, M., Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* **35**, 1026-1028.
9. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods* **19**(6), 679-682.
10. NCBI Virus [Internet]. (Bethesda, MD: National Library of Medicine, US), National Center for Biotechnology Information; [cited 2024 Sep 17]. Available from: <https://www.ncbi.nlm.nih.gov/labs/virus/>.
11. Fuchs F, Worrall D, Fischer V, et al. (2020) SE(3)-transformers: 3D roto-translation equivariant attention networks[J]. *Advances in Neural Information Processing Systems* **33**, 19701981.
12. Jumper, J., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873), 583-590.

13. Roney J P, Ovchinnikov S. (2022) State-of-the-Art estimation of protein model accuracy using AlphaFold. *BioRxiv*, 2022-03.
14. Zhang,Y. & Skolnick,J. (2005) TM-align: a protein structure alignment algorithm based on the TMscore. *Nucleic Acids Res* **33**, 2302-2309.
15. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Science: A Publication of the Protein Society* **23(1)**, 47–55.

CSSB_experimental (NA)

Improving RNA and Protein-Nucleic Acid Complex Structure Predictions with Deep Learning-Based Diverse Structure Sampling

Sojung Myung¹, Howon Lee², Minkyung Baek¹

1 - School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

2- Department of Life Sciences, College of Natural Science, Hanyang University, Seoul 04763, Republic of Korea

minkbaek@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3; EMA:Y; MD:N

In CASP16, we predicted protein-nucleic acid complex structures and RNA structures using RoseTTAFoldNA¹, AlphaFold3², and our in-house RNA structure generation model with human interventions. Since neither RoseTTAFoldNA¹ nor AlphaFold3² shows high-accuracy modeling performance for nucleic acids, we focused on sampling diverse structures and incorporating experimental data reported in literature.

Methods

Sampling diverse structures with various deep learning models

For both RNA and protein-nucleic acid complexes, a variety of structures were sampled using RoseTTAFoldNA¹ and AlphaFold3² by providing different random seeds. Template structures identified with BLAST³ and HHsearch⁴ were used for RoseTTAFoldNA-based prediction. With AlphaFold3, we also modeled structures with additional metal ions, such as magnesium, sodium, and potassium, as nucleic acids commonly interact with these ions. By modeling structures with different types or quantities of metal ions, we were able to sample more diverse structures using AlphaFold3.

For RNA monomers, we employed our in-house conditional RNA structure generation model to sample diverse RNA structures in addition to RoseTTAFoldNA and AlphaFold3. This model uses a flow matching algorithm⁵ to generate the tertiary structure of RNA based on the given sequence and base-pairing information. We provided base-pairing data, either predicted or reported in the literature, to guide structure sampling with this in-house model.

For complexes that were too large to predict as a whole, models were generated in segments and then combined. Subunits known to be adjacent were grouped and modeled together. In the case of protein-RNA complexes, since RNA structure is highly dependent on the surrounding proteins, RNA was always modeled with its interacting protein subunits to improve structural accuracy.

Ranking predicted structures with confidence metrics and experimental data

For RNA structure modeling, given the variety of methods used, it was challenging to directly compare

confidence metrics across different approaches. To address this, we ranked the predicted structures using pLDDT scores from RoseTTAFoldNA, treating each predicted structure as a template. When RNA secondary structure data were available in the literature, models that best matched to this information were selected as the final models.

For protein-nucleic acid complexes, we primarily relied on information extracted from relevant literature. The predicted structures were manually inspected and ranked based on their agreement with experimental data, such as nucleic acid binding domains or residues. In cases where no additional data were available, we used confidence metrics from AlphaFold3 to select the final models.

For both RNA and protein-nucleic acid complex modeling, the final models were optimized using Rosetta FastRelax⁶ before submission.

Availability

RoseTTAFoldNA is freely available at <https://github.com/uw-ipd/RoseTTAFold2NA> (source code). The in-house RNA structure generation model is currently under development. The code will be made publicly accessible upon completion.

Acknowledgement

This work was supported by the AI-Bio Research Grant through Seoul National University and the POSCO Science Fellowship of POSCO TJ Park Foundation.

1. Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D., & DiMaio, F. (2024). Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature methods*, 21(1), 117-121.
2. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3.
3. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410.
4. Söding, J. (2005). Protein homology detection by HMM–HMM comparison. *Bioinformatics*, 21(7), 951-960.
5. Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M., & Le, M. (2022). Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*.
6. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., & Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Science: A Publication of the Protein Society*, 23(1), 47–55.

CSSB_FAKER (TS, Assembly)

Generating MSA for Protein Structure Prediction via Structure Search and Sequence Design

Soohyun Jo, Hayoung Lee, Heesoo Ki[#], Min Su Yoon[#], Jeonghoon Park, and Minkyung Baek

School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

- equal contribution

minkbaek@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF2:Y; EMA:Y; MD:N

FAKER was employed to predict protein structures except huge complexes and DNA/RNA-protein complexes. This method focuses on improving multiple sequence alignments (MSA) for protein structure prediction by collecting sequences based on structural information. The overall pipeline is described in **Figure 1**.

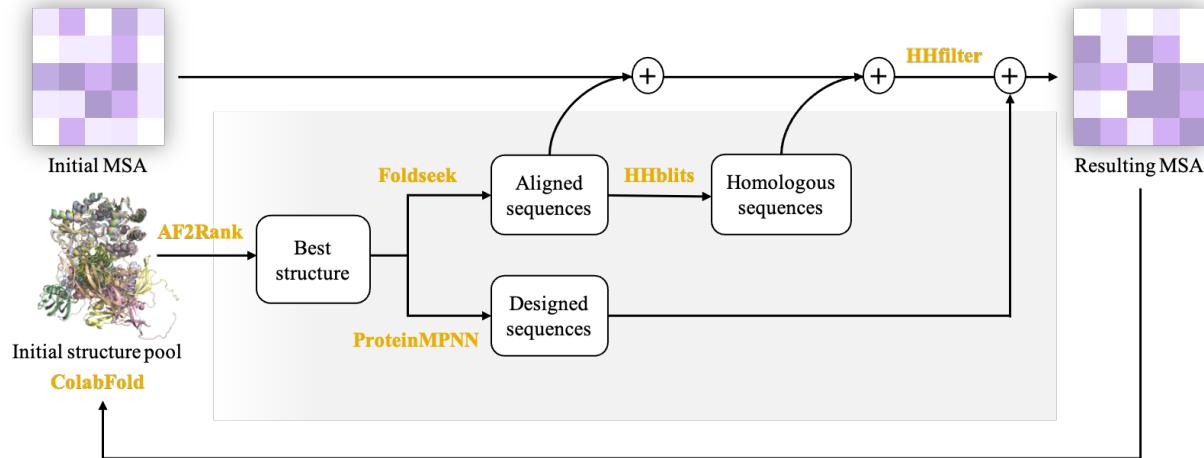


Figure 1. Schematic diagram of FAKER pipeline.

Methods

Initial structure prediction

FAKER requires initial predicted structures to improve MSA. For *phase 0*, initial structures were generated using ColabFold (MMseqs2-based MSA generation followed by AlphaFold protein structure prediction). Diverse structures were generated with and without template information, and AF2Rank then utilized to select the most confident structure. Composite scores of AF2Rank were used to measure

confidence for monomers, while, for multimers, we developed custom scores by combining metrics from AF2Rank. The custom score is the product of AF2Rank outputs including pLDDT, pTM, and ipTM. For *phase 1*, the initial structure was sourced from the top-ranked model submitted in *phase 0* from the CSSB-Human group.

Additional MSA generation with structural information

The selected initial structure was subsequently used as a query in Foldseek to identify structural homologs, retaining those that covered at least 50% of the query sequence. Each structurally homologous sequence was subjected to further sequence searching using HHblits and the resulting sequences were combined to the original MSA, which was then filtered using HHfilter³.

Additionally, the structure was used as input for ProteinMPNN whose noise-adding region was modified from all to low pLDDT residues only. ProteinMPNN designed multiple sequences under conditions of high noise and low temperature, and these were integrated into the filtered MSA. As FAKER has been developed until now, ProteinMPNN was conditionally applied during the later stages of CASP16 if the addition of designed sequences led to improvement in AF2Rank scores.

Structure refinement with iterative MSA generation

The refined protein structures were predicted using ColabFold with the extended MSA as described in the previous section. This MSA reconstruction and refinement process was repeated five times, and the top-scoring decoys, based on the AF2Rank metric, were submitted as final models after energy relaxation⁷. However, manual intervention was applied if there was evidence suggesting that a low-scored structure was more likely to be correct.

Notes on phase 2 submissions

For phase 2, FAKER algorithm was not applied. Instead, MassiveFold structures were filtered based on provided confidence metrics, and the remaining structures were evaluated using AF2Rank. Starting from the top-scoring decoys, a pairwise comparison of the structures was performed using TM-score to identify ten models as diverse as possible. These models were then clustered into five distinct groups, from which five representative structures were selected. After energy relaxation, the five representatives were submitted.

Acknowledgement

This work was supported by IITP/MSIT (RS-2023-00220628), NRF/MSIT (RS-2023-00210147). We thank the National Supercomputing Center in Korea with supercomputing resources including technical support (KSC-2022-CRE-0472, KSC-2023-CRE-0133).

Availability

FAKER is currently under development. The code will be made publicly accessible upon completion.

1. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature methods*, 19(6), 679-682.
2. Van Kempen, M., et. al. (2024). Fast and accurate protein structure search with Foldseek. *Nature biotechnology*, 42(2), 243-246.
3. Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., Söding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC bioinformatics*, 20, 1-15.
4. Dauparas, J., et. al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 378(6615), 49-56.
5. Roney, J. P., Ovchinnikov, S. (2022). State-of-the-art estimation of protein model accuracy using AlphaFold. *Physical Review Letters*, 129(23), 238101.
6. Zhang, Y., Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4), 702-710.
7. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein science*, 23(1), 47-55.

Guiding Structure Prediction with Human Intuition

Soohyun Jo^{1,#}, Heesoo Ki^{1,#}, Kunhee Kim^{1,#}, Yeajin Kim^{1,#}, Sojung Myung^{1,#}, Sanggeun Park^{1,#}, Hyunho Shin^{1,#}, Min Su Yoon^{1,#}, Minji Kang¹, Jeonghoon Park¹ and Minkyung Baek¹

¹ - School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

#These authors contributed equally to this work and are listed in alphabetical order

minkbaek@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

The CSSB-Human group performed structure sampling with various structure prediction methods, ranked the structures with AF2Rank scores, and selected the final models based on a mixture of human analysis and literature information.

Methods

Initially, proteins related to the target query protein were identified using BLAST¹, FoldSeek², and relevant literature. Literature information was primarily used to find clues about the tertiary structures, as well as to find potential interfaces for protein-protein interactions and their stoichiometries. Based on the initial structure and confidence metrics given by our in-house automated server (CSSB-server), we categorized the prediction process into three main branches, depending on the confidence level of the structure predictions. Special cases that require a deviation from the primary prediction process are described separately below.

High confidence predictions: Confidence metrics such as the pLDDT, pTM, or PAE scores provided by protein structure prediction methods are known to be reliable indicators of prediction quality. Therefore, when both the pLDDT and pTM were high and the PAE was low, no further interventions were made to improve the prediction quality. This typically happened when coevolutionary information in the MSA was sufficient or when clear templates were available. A literature search was also conducted in parallel, and if the predictions did not align with experimental results reported in literature, the case was treated as a special one.

Medium confidence predictions: For cases where the server predictions showed medium confidence, such as when the PAE of the interdomain regions was high, or when the pLDDT and pTM scores ranged between 0.6 and 0.8, additional interventions were made to improve structure quality. For example, we filtered out sequences based on coverage and identity when the MSA was too deep and diverse. When the MSA depth was insufficient, we reconstructed the MSA with additional sequence databases. For proteins where only specific regions had low sequence coverage, the MSA for “under-covered” regions were separately generated and then concatenated back into the original MSA, potentially providing additional coevolutionary signals. A parallel literature search was performed to identify potential homologues to use as structural templates, with human intuition playing a role in the final structure selection process.

Low confidence predictions: For cases where the confidence metrics were low, we performed more extensive literature search, as the methods above were difficult to apply. These targets therefore had to be handled on a case-by-case basis. Compared to the other confidence levels, targets with extremely low confidence metrics were relatively uncommon.

Special cases: When complex template structures were identified by BLAST¹, HHsearch³, or Foldseek², we used a modified version of AlphaFold2⁴ to incorporate complex template information, including interchain orientation. In cases where experimental data, such as interface residue information, were available to guide structure modeling, we employed various docking methods, including GalaxyTongDock⁵ and HADDOCK⁶, to sample a diverse range of complex structures that satisfied the experimental data. If the query protein was too large for our computational resources, models were predicted in segments and combined. For antibody cases, we sampled thousands of antibody-antigen complex structures using ColabFold⁷ and AlphaFold3⁸, selecting the top-ranked models with AF2Rank⁹.

Final selection process: After running the processes above, we ranked the candidate structures with AF2Rank⁹. Complexes were ranked with a newly defined custom score, designed to put the interfaces more into consideration. When experimental data was available, human intervention was made accordingly. We submitted the top 5 models after performing energy optimization using Rosetta FastRelax¹⁰.

Handling the New Phases of CASP16

Phase 0: In phase 0, we first referenced templates identified by the HHsearch³ or FoldSeek² to approximate the possible stoichiometries. We then examined the confidence metrics provided by ColabFold⁷, AlphaFold3⁸, and RoseTTAFold2¹¹. By incorporating additional information found in the literature, we made a final decision on the most likely stoichiometry.

Phase 2: Based on the various confidence metrics provided with the MassiveFold structures, we select the most likely structures using AF2Rank⁹ after removing structures below a predefined quality threshold. Up to 10 of the highest ranked, structurally diverse models were selected by clustering using the TM-score¹². These models were then further clustered into four groups, and the high-confidence regions of representative structures from each cluster were given as custom templates for ColabFold⁷ modeling. Among the ColabFold⁷ outputs, we selected the structures based on the confidence metrics and submitted the top 5 “unique” structures after energy minimization using Rosetta FastRelax¹⁰.

Acknowledgement

This work was supported by IITP/MSIT (RS-2023-00220628), NRF/MSIT (RS-2023-00210147), and the New Faculty Startup Fund from Seoul National University.

1. Altschul,S.F., Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. & Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.
2. van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J., & Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. Nature

Biotechnology, 42(2), 243–246.

3. Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., & Söding, J. (2019). HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics*, 20(1), 1–15.
4. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.
5. Park, T., Baek, M., Lee, H., & Seok, C. (2019). GalaxyTongDock: Symmetric and asymmetric ab initio protein-protein docking web server with improved energy parameters. *Journal of Computational Chemistry*, 40(27), 2413–2417.
6. Dominguez, C., Rolf Boelens, A., & Bonvin*, A. M. J. (2003). HADDOCK: A Protein–Protein Docking Approach Based on Biochemical or Biophysical Information. <https://doi.org/10.1021/ja026939x>
7. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, 19(6), 679–682.
8. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500.
9. Roney, J. P., & Ovchinnikov, S. (2022). State-of-the-Art Estimation of Protein Model Accuracy Using AlphaFold. *Physical Review Letters*, 129(23), 238101.
10. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., & Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Science: A Publication of the Protein Society*, 23(1), 47–55.
11. Baek, M., Anishchenko, I., Humphreys, I. R., Cong, Q., Baker, D., & DiMaio, F. (2023). Efficient and accurate prediction of protein structure using RoseTTAFold2. In bioRxiv (p. 2023.05.24.542179). <https://doi.org/10.1101/2023.05.24.542179>
12. Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, 33(7), 2302–2309.

CSSB_experimental (Antibody)

Epitope-guided predictions of antibody-antigen complex structures

Kunhee Kim¹, Yubeen Kim², Minkyung Baek¹

1 - School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

2 - Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea

minkbaek@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:N; EMA:Y; MD:N

We present RF-AbAg, a docking model based on RoseTTAFold2¹, designed for the unbound docking of antibody-antigen complexes. RF-AbAg constructs antibody-antigen orientations by utilizing epitope information, multiple sequence alignments (MSA), and individual antibody and antigen structures. It was developed by fine-tuning RoseTTAFold2 on antibody-antigen complex structures curated from SabDab² as well as general loop-mediated protein-protein complex structures. We applied RF-AbAg for antibody complex structures during CASP16.

Methods

Phase 0 & 1: Epitope scanning for antibody-antigen complex structure predictions

In phase 0, antigen stoichiometry was predicted through literature or template searches. Individual antibody and antigen structures were modeled using ColabFold³. When epitope information was unavailable, we scanned the antigen surface to generate potential epitope candidates. For each surface residue, all nearby antigen residues within 5 Å were grouped into a single epitope candidate. RF-AbAg then sampled antibody-antigen complex structures based on these epitope candidates. To enable the modeling of CDR loops, we used CDR-removed antibody structures as inputs, allowing RF-AbAg to reconstruct the CDR regions during interface modeling. Structures with low interface Predicted Aligned Error (iPAE) scores were selected and underwent further relaxation⁴. The final models were re-ranked using AF2Rank⁵.

Phase 2: Utilizing predicted epitopes from MassiveFold

The complex structures provided by MassiveFold were first sorted based on confidence metrics, primarily the iPAE score from AF2Rank. Epitope candidates were then extracted from the top-scored MassiveFold complexes. For each of these epitope candidates, the antibody-antigen complexes were re-modeled using RF-AbAg, employing the same individual antibody and antigen structures from the previous phases. The resulting complexes were ranked using the iPAE metric from RF-AbAg. From these, a structurally diverse set of five models was selected for additional energy optimization using FastRelax⁴. The final models were re-ranked using AF2Rank⁵.

Availability

RF-AbAg is currently under development. The code will be made publicly accessible upon completion.

Acknowledgement

This work was supported by IITP/MSIT (RS-2023-00220628), NRF/MSIT (RS-2023-00210147, RS-2024-00397865). We thank the National Supercomputing Center in Korea with supercomputing resources including technical support (KSC-2022-CRE-0472, KSC-2023-CRE-0133).

1. Baek, M. et al. (2021) Efficient and accurate prediction of protein structure using RoseTTAFold2. *Biorxiv*, 2023-05.
2. Dunbar, J. et al. (2014) SabDab: The structural antibody database. *Nucleic Acids Res.* **42**, D1140-6.
3. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods* **19**(6), 679-682.
4. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., & Baker, D. (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Science: A Publication of the Protein Society*, 23(1), 47–55.
5. Roney J P, Ovchinnikov S. (2022) State-of-the-Art estimation of protein model accuracy using AlphaFold. BioRxiv: 2022.03.11.484043, 2022.

Protein 3D Structure Prediction with DeepFold in CASP16

Minsoo Kim¹, Hanjin Bae¹, Gyeonpil Jo¹, Kunwoo Kim¹, Seonggwang Jeon², Iljung Kim², Sehoon Park³, Seokjun On², Hosung Kim⁵, Jinwoo Lee⁵, Jaeoh Shin⁴, Sung Jong Lee⁶, Yung-Kyun Noh², Eun-Sol Kim², Eunok Paek², Jejoong Yoo^{1*}, and Keehyoung Joo^{4*}

1 - Department of Physics, Sungkyunkwan University, Korea, 2 - Department of Computer Science, Hanyang University, Korea, 3 - Department of Artificial Intelligence, Hanyang University, Korea, 4 - Center for Advanced Computation, Korea Institute for Advanced Study, Korea, 5 - Department of Mathematics, Kwangwoon University, Korea, 6 - Basic Science Research Institute, Changwon National University, Changwon, Korea

* Corresponding authors: Keehyoung Joo (newton@kias.re.kr); Jejoong Yoo (jejoong@skku.edu)

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:N EMA:Y; MD:Y

In the CASP16 experiment, we developed DeepFold2, an upgraded version of DeepFold ¹. The DeepFold2 pipeline (DeepFold-server) introduces new features: (1) a refactored baseline framework offering greater flexibility in handling diverse inputs, (2) fast vector embedding-based databases, which significantly accelerate sequence alignment searches by leveraging protein language models (PLM), and (3) a customized multimer pipeline that utilizes monomer templates generated by our monomer protocol to enhance multimer modeling. Furthermore, as a human protocol, we refined the structure predicted by DeepFold-server through a conformational space annealing (CSA) method ², which improved both the accuracy and reliability of the prediction (DeepFold).

Methods

Refactored baseline framework: We refactored the entire DeepFold2 pipeline using PyTorch, introducing modern AI techniques such as multi-GPU inference, flash attention for the Evoformer network, and reduced precision (bfloating16) to decrease memory usage for the embedder and Evoformer stacks. This transition allowed us to take advantage of PyTorch's dynamic computation graph for fast inference and handling of large protein complexes, enabling easier debugging and offering greater flexibility for future model development.

MSA feature: To improve the quality of MSA, DeepFold2 introduced new approaches to the conventional methods, which relied on sequence databases searches using HHblits and JackHMMER. Our new approach integrated vector database MSA searches, which utilizes PLMs. Specifically, newly-constructed offline sequence embedding databases enabled quick searches for similar sequences by comparing vector embeddings rather than raw sequences. Potential sequences were then aligned using PLMalign.

Template features: In template-based structure prediction, the DeepFold2 pipeline combines several complementary techniques to improve accuracy. Initially, our pipeline performs an hhsearch against the PDB70 database to identify potential templates. Additionally, the pool of viable templates is expanded by using Foldseek to find structures similar to those generated by AF2 baseline predictions ³. Finally, DeepFold2 applies the CRFalign method ⁴, based on conditional random fields, to realign templates, optimizing selection and ensuring high-quality alignments.

Structure prediction: To predict the three-dimensional structure of proteins, we utilized both AF2 network and DeepFold2 network together. The pipeline produced approximately 35 distinct structures for each query sequence, which were subsequently evaluated and ranked.

Ranking and Clustering: After generating structure predictions, these models are subsequently clustered using hierarchical clustering based on TM-scores to group similar structures together. From these structures, we select the top five models according to their pLDDT scores, ensuring that only the highest quality predictions are considered for further analysis.

Multimer pipeline: Our multimer pipeline utilized AlphaFold-Multimer version 2.3 (AFM) network with customized input features. This pipeline incorporates both paired MSA features based on taxonomy annotation and updated template derived from the monomer prediction process. The multimer predictions are clustered using US-align⁵, and the best structures are selected based on their pTM+ipTM scores, enabling accurate prediction of multimeric assemblies.

DeepFold as Human protocol: We applied the conformational space annealing (CSA) method to globally optimize the predicted structures generated by our DeepFold-server pipeline. Using distogram restraint potentials and the ref2015_cart score function from the PyRosetta package, the CSA method produces refined structures that satisfy the distogram restraints obtained from DeepFold-server while balancing the Rosetta forcefield.

Availability

The code for DeepFold will soon be available on GitHub. Additionally, we will be launching (<https://www.deepfold.org/>), which will be accessible upon request.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2023-00220628).

- 1 Lee, J.-W. *et al.* DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function. *Bioinformatics* **39** (2023).
- 2 Lee, J., Scheraga, H. A. & Rackovsky, S. New optimization method for conformational energy calculations on polypeptides: Conformational space annealing. *Journal of Computational Chemistry* **18**, 1222-1232 (1997).
- 3 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
- 4 Lee, S. J. *et al.* CRFalign: A Sequence-Structure Alignment of Proteins Based on a Combination of HMM-HMM Comparison and Conditional Random Fields. *Molecules* **27**, 3711 (2022).
- 5 Zhang, C., Shine, M., Pyle, A. M. & Zhang, Y. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nature Methods* **19**, 1109-1115 (2022).

DeepFold-interact

Deep-interact: A Graph Transformer-Based Protein-Ligand Interaction Site Prediction

San Kim¹, Seungjun Lee², Sichan Oh³, Minsoo Kim⁴, Jejoong Yoo⁴, Jaekwang Kim^{1,2,†}, and Keehyoung Joo^{5,†}

¹ - Department of Computer Science and Engineering, Sungkyunkwan University, Suwon, Republic of Korea;

² - Department of Immersive Media Engineering / Convergence Program for Social Innovation, Sungkyunkwan University, Suwon, Republic of Korea; ³ - Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Republic of Korea;

⁴ - Department of Physics, Sungkyunkwan University, Suwon, Republic of Korea;

⁵ - Korea Institute for Advanced Study (KIAS), Seoul, South Korea

†Jaekwang Kim (linux@skku.edu); Keehyoung Joo (newton@kias.re.kr)

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; EMA:N; MD:N

We developed Deep-interact, a graph transformer-based GNN (Graph Neural Network) architecture for protein-interaction site prediction¹. We predict the interaction site of the protein structure using a regression approach rather than classification. We designed the model to learn interactions by using both protein structural information and ligand data together as inputs to the GNN. To enhance the model performance, we used residual connections, global multi-head attention, and deep graph transformer layers in the model. Using this approach, we performed inference and prediction on the L5001 target protein-ligand interaction data from CASP16. The L5001 protein structures were generated using the DeepFold protocol².

Methods

Deep-interact utilizes a dual-channel architecture comprising a protein channel and a ligand channel, predicting interaction sites by calculating a dot product between their representation vectors.

Protein representation: It is processed using Ankh, a pre-trained language model for protein residues, and an adjacency matrix was constructed using Euclidean distances between residues to capture spatial relationships³. This channel uses multiple layers of graph transformers with residual connections. The transformer layers apply global multi-head self-attention and adjust the attention scores based on the distance matrix between residues. With the residual connection and attention value adjustment, the model reduces the problem of over-smoothing.

Ligand representation: It is generated by Molformer, a transformer-based molecular model³. The model architecture consists of five layers of graph transformers in the protein channel with an input dimension of 1,536, and three layers of FFN in the ligand channel with an input dimension of 768.

Outputs from both channels are reduced to 48 dimensions before the final dot product. This representation is projected to match the protein channel's output size through a feed-forward network (FFN) employing

Exponential Linear Unit (ELU) activation functions.

To obtain a predicted interaction score for each residue-ligand pair, the model calculates a scalar product between the protein residue and ligand representations. The ground truth score is pre-processed from the ligand-residue distance. During pre-processing, the distance is normalized, with emphasis placed on values below 3Å. These values are then converted to their reciprocals to assign higher scores to interaction residues, and the data is smoothed to ensure stable learning.

Our regression-based method outputs continuous values representing the degree of interaction likelihood for each residue. This approach addresses class imbalance by using both positive (interacting) residues and negative (non-interacting) samples. As a result, the model's generalization capabilities improve, offering a more comprehensive understanding of protein-ligand interactions.

Data preprocessing: The preprocessing pipeline accounts for the three-dimensional nature of proteins and ligands by calculating distances between the ligand's center of mass and each residue's alpha carbon. We adjusted the raw distance values to account for different protein sizes and to highlight important interaction distances. We did this by applying transformations like logarithmic scaling, non-linear adjustments, and limiting the values within a specific range (clamping).

Training: We trained the model on the PDBBind2021 dataset. Training was performed on an NVIDIA RTX 4090 GPU⁵. The training process took approximately five hours over 35 epochs, using a batch size of 200 to balance efficiency with memory constraints.

Pose and affinity prediction: Based on the predicted interaction sites, we calculated the ligand center by taking the weighted mean of the coordinates of residues involved in interactions. We then employed QuickVina2 to predict the ligand poses⁶.

Availability

The code and model weights for Deep-interact will soon be available on GitHub.

Acknowledgment

This research was supported by the Ministry of Science and ICT (MSIT), Korea, through the ICT Challenge and Advanced Network of HRD (ICAN) program (IITP-2024-RS-2023-00259497), the Graduate School of Metaverse Convergence at Sungkyunkwan University (No.RS-2023-00254129), the Basic Science Research Program (No.RS-2024-00346737), and an IITP grant (No.RS-2023-00220628), all funded by the Korean government (MSIT).

1. Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., ... & Liu, T. Y. (2021) Do transformers really perform badly for graph representation?. *Advances in neural information processing systems*. **34**, 28877-28888.
2. Lee, J. W., Won, J. H., Jeon, S., Choo, Y., Yeon, Y., Oh, J. S., ... & Joo, K. (2023) DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function. *Bioinformatics*. **39**(12), btad712.
3. Elnaggar, A., Essam, H., Salah-Eldin, W., Moustafa, W., Elkerdawy, M., Rochereau, C., & Rost, B. (2023) Ankh: Optimized protein language model unlocks general-purpose modelling. *arXiv preprint arXiv:2301.06568*.

4. Wu, F., Radev, D., & Li, S. Z. (2023, June) Molformer: Motif-based transformer on 3d heterogeneous molecular graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 4, pp. 5312-5320).
5. Liu, Z., Su, M., Han, L., Liu, J., Yang, Q., Li, Y., & Wang, R. (2017) Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of chemical research*. **50**(2), 302-309.
6. Alhossary, A., Handoko, S. D., Mu, Y., & Kwoh, C. K. (2015) Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*. **31**(13), 2214-2216.

DeepFold-refine

Refining Protein Structures with Molecular Dynamics and Targeted Restraints

Jaeoh Shin¹, Minsoo Kim², Keehyoung Joo^{1,*}, Jejoong Yoo^{2,*}

¹ Korea Institute for Advanced Study (KIAS), Seoul, South Korea,

² Department of Physics, Sungkyunkwan University, Suwon, South Korea

newton@kias.re.kr; jejoong@skku.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:N; EMA:N; MD:Y

We participated in CASP16 with the DeepFold-refine protocol, aimed at improving structures predicted by our DeepFold-server¹ using molecular dynamics (MD) simulations. To enhance the detailed quality of the predictions without significantly altering the original structures, several restraint potentials were applied during the MD simulations. In most cases, our protocol successfully improved the clash score and stereochemistry of both the backbone and side chains.

Methods

The main challenge of the refinement protocol lies in selecting optimal conditions that allow for local conformational optimization while preserving the global structure. To achieve this, we incorporated modifications to the force field and applied several restraint potentials within the standard MD simulation protocol. These modifications and restraints were implemented in the OpenMM program².

Force field: We employed the AMBER ff19SB force field, refined by CUFIX corrections³, which improve inter-molecular interactions for charged residues, especially enhancing conformational sampling for loops and inter-domain contacts.

Restraint potentials: We utilized a combination of three restraint potentials:

- (a) Distogram-based restraints: The distogram information generated by the DeepFold-server was converted into a restraint energy function to improve agreement between the predicted structure and the distogram.
- (b) Side-chain torsion restraints: A flat-bottom Lorentzian-type potential energy function was applied to further improve the agreement of the side-chain torsions with the prediction made by DeepFold.
- (c) Position restraints: A restraint potential was applied to prevent significant positional changes in the C-alpha atoms of the protein.

Production for refinement: First, using the DeepFold-predicted structure as an initial input, we performed restraint MD simulations, typically running for 50 ns. Structures were saved every 10 ns, followed by clustering using the agglomerative clustering algorithm. Finally, the centroid of the largest cluster was selected as the refined structure. This structure then underwent final energy minimization for submission.

Production for ensemble: For R1260 (solvent shell) and T1200/T1300 (inter-domain ensemble)

targets, we performed the AMBER-based MD simulations with CUFIX corrections for 40 nanoseconds and 20 microseconds, respectively, to obtain thermodynamic conformational ensemble at room temperature. The ensembles generated from these MD simulations were submitted without further refinement.

Results

The TM-score comparison between the initial and refined structures shows that the global structure remains largely unchanged. However, the side-chain geometry exhibited improvements, as evidenced by an enhanced MolProbity score.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2023-00220628).

1. Lee, J.-W. et al. DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function. *Bioinformatics* 39 (2023). <https://doi.org/10.1093/bioinformatics/btad712>
2. Eastman, P. et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Computational Biology* 13, e1005659 (2017). <https://doi.org/10.1371/journal.pcbi.1005659>
3. Yoo, J. & Aksimentiev, A. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Physical Chemistry Chemical Physics* 20, 8432-8449 (2018). <https://doi.org/10.1039/c7cp08185e>

DELCLAB in CASP16

Carlos A.Del Carpio Munoz

Choju-Medical Institute, Fukushima Hospital. Noyori-cho, Yamanaka 19-14. Toyohashi-City, Aichi-ken 441-8124, Japan

We present an enhanced version of our bioinformatics platform designed for the prediction of the 3D and 4D structures of proteins and other biomolecules, incorporating advanced machine learning algorithms and innovative methodologies. While our previous work laid the groundwork for predicting protein folding patterns along with local substructures and domains, we recognize the significant opportunity for improvement in modeling long-range amino acid interactions through the integration of structural insights from protein structural knowledge bases.

In prior CASP rounds, our approach combined traditional homology modeling with our unique method involving spectral analysis of amino acid sequences based on their physicochemical properties [1,2]. These methodologies have been instrumental in identifying the overarching folding families of target sequences; however, accurately predicting amino acid and atomic long-range interactions remains crucial for achieving near-native structural fidelity. Notably, results from CASP-15 demonstrated that AlphaFold's[3] sophisticated stereochemical treatment of proteins yields remarkable accuracy. As a consequence, we introduced AlphaFold-predicted structures in CASP-16 to enhance our predictions.

This approach has proven especially beneficial in predicting protein quaternary structures and complexes. Whenever feasible, we incorporated AlphaFold-generated structures alongside those derived from our original methodology to model these higher-order assemblies. We employed energy minimization and molecular dynamics simulations to refine atom placements and rank the final structural predictions.

Methods

Our multi-platform automated system initiates with the identification of optimal homologs using established methodologies. In instances where no homologs are available, we pivot to a spectral analysis of sequences, producing outputs that are analyzed in conjunction with the target sequence. The construction of the requisite 3D structure follows, facilitated by our platform, which subsequently conducts loop and structural stability analyses. The most promising candidate structures undergo molecular dynamics simulations and minimization processes, culminating in energetic ranking.

For predicting protein assemblies, we leverage MIAX[4,5], our system for protein interaction assessment. This component encompasses binding site prediction and the docking of predicted structures. For hetero-multimer structure predictions, we have developed a novel approach to assess the interaction order among subunits, enhancing the accuracy of our multimer configurations.

1. Del Carpio, C. A. & Yoshimori, A. (2002). Fully automated protein tertiary structure prediction using Fourier transform spectral methods. Protein Structure Prediction: Bioinformatics, University of California, International University Line.
2. Del Carpio, C. A. & Carbajal, J. C. (2002). Folding pattern recognition in proteins using spectral analysis methods. *Genome Inform* 13, 163-72.
3. Jumper J. et all (2021). Highly Accurate Protein Structure Prediction with AlphaFold, *Nature* 596(2021).
4. Del Carpio, C.A., Ichiishi E. (2017). Inference of Protein Multimeric Complex Dynamic Order of Formation: An Active Region Recognition Based Approach. *International Journal of Genomics and Data Mining* 2017, 1.
5. Del Carpio, C. A., Ichiishi, E., Yoshimori, A. & Yoshikawa, T. (2002). A new paradigm for modeling biomacromolecular interactions and complex formation in condensed pahses. *Proteins: Structure, Function, and Genetics* 48, 696-732.

Method description for modeling ligand targets from CASP16

Guangfeng Zhou, Frank DiMaio

Institute for Protein Design, Department of Biochemistry, University of Washington

dimaio@uw.edu

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; EMA:Y; MD:Y.

Here we describe the methods we used for modeling ligand targets from CASP16. We heavily rely on the physics-based ligand docking protocol, Rosetta GALigandDock^{1,2}, for predicting protein-ligand complex structures and ranking ligands.

Methods

Small molecule ligand preparation: Initial ligand conformation is generated from SMILES strings using UCSF Chimera³. Ideal bond geometry and partial charges are computed using the MMFF94⁴ force field in OpenBabel 3.1.0⁵ for small molecules. Small molecule parameter files for Rosetta GALigandDock^{1,2} are generated using a utility script “mol2genparams.py” distributed with Rosetta.

Protein receptor preparation:

Crystal structures with 100% sequence identity to the provided target sequence are downloaded from the RCSB database (<http://www.rcsb.org>). Solvent molecules, ions, and ligands are removed if they exist in the crystal structure. The apo structures are then relaxed with coordinate constraints added to all heavy atoms to remove potential clashes.

Pocket selection:

For L1000-L4000 targets, the pocket is selected based on the known ligand binding site from the crystal structure. For target L5000, we use the Chai-1⁶ model to first predict the protein-ligand complex structure, and the predicted binding site is used as the pocket for ligand docking using GALigandDock.

Ligand Docking:

Docking of small molecule ligands is performed using Rosetta GALigandDock, which employs a physical energy model with genetic algorithm optimization. The flexible docking mode “dockflex” is used to simultaneously model the ligand and pocket side chain conformations. Binding affinity of the docked structure is estimated in GALigandDock with the “Simple” entropy model². Sixteen independent docking runs are performed for each putative receptor structure. All the docked structures of the same ligand are first ranked by the predicted binding affinity (dG), and the top five structures are manually examined in PyMol⁷ for submission.

Binding affinity estimation:

We use Rosetta GALigandDock “eval” mode² to estimate the binding affinities on the provided experimental determined protein-ligand complexes in stage two. Binding affinities are estimated with the “Simple” entropy model, and we convert the binding affinity value from GALigandDock to relative ranking for submission.

Availability

Rosetta is freely available at <https://github.com/RosettaCommons/rosetta> for non-commercial usage. Chai-1 model is freely available at <https://github.com/chaidiscovery/chai-lab> for non-commercial usage.

1. Park, H.; Zhou, G.; Baek, M.; Baker, D.; DiMaio, F. Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein–Ligand Docking. *J. Chem. Theory Comput.* 2021, 17 (3), 2000–2010. <https://doi.org/10.1021/acs.jctc.0c01184>.
2. Zhou, G.; Rusnac, D.-V.; Park, H.; Canzani, D.; Nguyen, H. M.; Stewart, L.; Bush, M. F.; Nguyen, P. T.; Wulff, H.; Yarov-Yarovoy, V.; Zheng, N.; DiMaio, F. An Artificial Intelligence Accelerated Virtual Screening Platform for Drug Discovery. *Nat. Commun.* 2024, 15 (1), 7761. <https://doi.org/10.1038/s41467-024-52061-7>.
3. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* 2004, 25 (13), 1605–1612. <https://doi.org/10.1002/jcc.20084>.
4. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* 1996, 17 (5-6), 490–519. [https://doi.org/10.1002/\(sici\)1096-987x\(199604\)17:5/6<490::aid-jcc1>3.0.co;2-p](https://doi.org/10.1002/(sici)1096-987x(199604)17:5/6<490::aid-jcc1>3.0.co;2-p).
5. O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* 2011, 3 (1), 33. <https://doi.org/10.1186/1758-2946-3-33>.
6. ChaiDiscovery. Chai-1 Technical Report. https://chaiassets.com/chai-1/paper/technical_report_v1.pdf (accessed 2024-09-16).
7. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Ab initio RNA structure prediction with RNA composite likelihood maximized language model and deep end-to-end potential

Yang Li¹, Chenjie Feng² and Yang Zhang¹

¹ - Cancer Science Institute of Singapore, National University of Singapore,

² - School of Science, Ningxia Medical University

liyangum@nus.edu.sg

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:Y; AF:N; EMA:Y; MD:Y

RNAs play fundamental roles in living cells, with the biological functions largely determined by their tertiary structures. However, accurately modeling 3D RNA structures remains a significant challenge. To address this, we propose a novel method, DRfold2, that combines an RNA language model, pre-trained using composite likelihood maximization, with end-to-end RNA structure learning. The proposed method demonstrates improved structural modeling ability based solely on single-sequence information.

Methods

The proposed method consists of three main steps: (1) RNA sequence embedding through the RNA Composite Language Model (RCLM), (2) end-to-end structure and geometry prediction, and (3) structure selection, optimization, and refinement.

The traditional Masked Language Model (MLM) typically uses cross-entropy loss summed over masked tokens to approximate the full negative log-likelihood, which may ignore higher-order interactions between masked tokens. To address this limitation, we introduce a pairwise composite likelihood¹ to model RNA sequences. This approach can better approximate the full likelihood by considering pairs of observations, while avoiding the high computational cost of calculating the full joint likelihood. We have designed a transformer-based module that simultaneously models both sequential and pairwise representations. The predictions consist of two components: token-level (marginal) distributions and pairwise (joint) distributions. Given an input sequence, RCLM generates both sequential and pairwise representations, which are then used as input embeddings for structure modeling,

The end-to-end folding model is similar to the previous version², but with a modified FAPE loss function, which assigns greater weight to nucleotide contacts. Additionally, a reference term is introduced to encourage frames to move outward from their initial positions. The deep learning model also predicts inter-nucleotide geometries for the construction of a hybrid energy function, as FAPE does not perfectly correlate with structural evaluation indices and may require further optimizations.

The predicted conformations are selected using a hybrid energy function that integrates end-to-end frames and predicted geometries. These selected conformations are further optimized using a similar hybrid potential, where only the conformations and corresponding geometries from the top 5 selected models are used for the hybrid potential construction. Conformation optimization is performed using the L-BFGS algorithm with a differentiable potential function implemented in PyTorch. The resulting structures are then refined using Arena³ and OpenMM⁴ to correct possible incorrect bond lengths and angles, base and base-pair conformations, and atom clashes.

To objectively assess our approach, we predicted all monomer RNA structures in CASP16 solely based on DRfold2. For Target R1281, an RNA 6-helix bundle dimer, we constructed the dimer structure by incorporating some conformations from other group predictions of Target R0281.

Availability

The web server will be made available at <https://zhanggroup.org/>

1. Lindsay, B. G. (1988). Composite likelihood methods. *Contemporary Mathematics*, **80**, 221-239.
2. Li, Y., Zhang, C., Feng, C., Pearce, R., Lydia Freddolino, P., & Zhang, Y. (2023). Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nature Communications*, **14**, 5745.
3. Perry, Z. R., Pyle, A. M., & Zhang, C. (2023). Arena: rapid and accurate reconstruction of full atomic RNA structures from coarse-grained models. *Journal of Molecular Biology*, **435**, 168210.
4. Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., ... & Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, **13**, e1005659.

Crowdsourcing Small Molecule Structure Prediction with Foldit Drug Discovery

T. Scott¹, J. Meiler^{1,2}, Foldit Players³ and R. Moretti¹

¹ –Department of Chemistry, Vanderbilt University, Nashville, TN, USA, ² –Institute for Drug Discovery, University of Leipzig, Leipzig, Germany, ³ –Worldwide consortium

rocco.moretti@vanderbilt.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:N; AF:N; EMA:N; MD:N

Foldit¹ is a citizen science game where people from around the world can log on and contribute to scientific discovery of structural biology tasks. Recently the Foldit interface has been expanded to include drug discovery tools (“Drugit”), allowing Foldit players to diversify and redock molecules. For CASP16 these tools have been repurposed to allow Foldit players to “design” within a limited subset of molecules, with the aim of finding the best binding position of a suite of small molecules to a given protein target.

Methods

Structures of full-length target proteins were created with RosettaCM² using pre-existing structures of the protein as targets. In cases where multiple template structures were available, each was modeled independently. A relaxation with restraints³ was then performed to reduce the deviation of the RosettaCM generated structure to the input structure. Low energy structures thus obtained are used as the protein model for puzzle setup.

Ligand docking was performed by Foldit Players through the Foldit game interface. Players were presented with a series of puzzles which allowed them to manipulate the protein-ligand complex, searching for the best scoring structure. In addition to the standard tools for optimization (gradient descent, conformer sampling, internal bond rotation, and global rotation/translation of the ligand), players were allowed to change the chemical identity of the existing ligand to one of the other molecules. Throughout the week-long puzzle, players alternated between optimizing the current ligand and switching the identity of the ligand to a different molecule.

Initial ligand placement were selected from existing crystalized ligands. (In general, the starting molecule was not one of the experimental ligands.) Only the “primary” ligand was provided to the players. Ligands deemed to be crystallization adducts were omitted. Experimental ligands were provided to the users through the “Ligand Queue”, “Reaction Design” and “Compound Library” tools. Ligand Queue provides a small set of compounds to the local machine which players can cycle through to replace the current working molecule, and was used for most puzzles. This is limited to a small number of compounds (< 20), so the Reaction Design tool was repurposed for larger compound sets. Normally intended for combinatorial libraries, for CASP16 it was populated with sets of single-component “reactions” which yielded the experimental compounds. The size of target L3000 induced us to also make the Compound Library search tool available, which allowed players to search for similar compounds. Using the Ligand Queue and subsequent Compound Library search access to all ligands was possible.

Full backbone and sidechain flexibility was permitted to players, although strong C alpha restraints were included to keep the backbone from moving too far. Scoring was primarily the Rosetta energy function, though the protein-ligand contribution was upweighted by a factor of 20 to favor results with good protein-ligand interactions. A torsional quality objective (based on the Rarey torsional library⁴) was added to ensure that the ligand conformations were not improperly strained. Foldit is typically based around coming up with a single “best” structure per puzzle, however structures with different compounds identities are captured due to “shares” by the players as well as regular intermediate updates. To better capture the top results for the different small molecules, facilities were added to better capture the top-scoring structures on a per-compound rather than per-puzzle basis.

All structures submitted by the Foldit players to the Foldit servers were collected and separated out on a per-ligand basis. Post processing was deliberately constructed to prioritize Foldit player results and minimize judgement calls and discretion from the scientists involved in post-processing. For each chemical identity, structures were ranked based on Foldit puzzle score. Iterating through the list, the best three structures that were (a) not more than 2 Ang protein C alpha rmsd from the input backbone and (b) more than 2 Ang automorphic ligand rmsd from any previously selected structure were selected. The global list of structures was re-ordered based on predicted protein-ligand interaction energy, and the process repeated to obtain up to two additional structures. As absolute Foldit score is not necessarily comparable between puzzles with different settings, for targets with multiple puzzle rounds the lists of five selected structures from each round were combined by prioritizing those puzzles with highest ligand interaction scores in their set of three total score results. The two interaction score candidates were selected based on absolute interaction score. Selected structures were made full length (where size considerations required us to trim the structures before presenting it to the players) and converted to CASP format prior to being submitted.

Availability

The Foldit client is downloadable from <https://fold.it>. Rosetta can be obtained from <https://rosettacommons.org/> or through <https://github.com/RosettaCommons/rosetta>.

1. Cooper.S., Khatib.F., Treuille.A., Barbero.J., Lee.J., Beenan.M., Leaver-Fay.A., Baker.D., Popović.Z., & Players.F. (2010). Predicting protein structures with a multiplayer online game. *Nature* **466(7307)**, 756-60. doi: 10.1038/nature09304.
2. Song.Y., DiMaio.F., Wang.R.Y., Kim.D., Miles.C., Brunette.T., Thompson.J. & Baker.D. (2013) High-resolution comparative modeling with RosettaCM. *Structure* **21(10)**, 1735-42. doi: 10.1016/j.str.2013.08.005.
3. Nivón.L.G., Moretti.R., & Baker.D. (2013) A Pareto-optimal refinement method for protein design scaffolds. *PLoS One*. **8(4)**, e59004. doi: 10.1371/journal.pone.0059004.
4. Guba.W., Meyder.A., Rarey.M. & Hert.J. (2016) Torsion Library Reloaded: A New Version of Expert-Derived SMARTS Rules for Assessing Conformations of Small Molecules. *J. Chem. Inf. Model.* **56(1)**, 1-5. doi: 10.1021/acs.jcim.5b00522

Methods for Improved Protein Structure and Protein-Protein Interaction Prediction

Marko Ludaic, Sarah Narowe Danielsson, Matteo Tadiello and Arne Elofsson

Department of Biochemistry and Biophysics, Stockholm University, Science for Life Laboratory, Box 1031, 17121 Solna, Sweden

sarah.narowe@scilifelab.se

Proteins are the workforce of all living organisms and thereby play a crucial role in a large number of biological processes. In order to understand how proteins function, it is vital to understand their structure. Additionally, proteins rarely act alone. Instead, they interact with each other in order to fulfill their tasks. In the Elofsson lab, we focus on using machine learning methods for protein structure and protein-protein interaction (PPI) prediction, in order to uncover the mysteries of proteins. In addition, we have recently started exploring methods for RNA structure prediction. Compared to models for protein structure prediction, methods for RNA still have a long way to go until they become well-established, mostly due to the lack of RNA data and limited information obtained from RNA sequences¹.

Methods

The highest-performing machine learning model for predicting protein structure and protein-protein interactions is AlphaFold2 as shown in previous CASP competitions as well as previous benchmark studies²⁻⁴. However, despite its high performance, there are still parts of the AlphaFold2 model that can be optimized. One aspect that the lab is looking at is the time consumption. Due to the time it may take to predict a single protein-protein interaction it is currently not feasible to utilize AlphaFold2 to look at millions of pairwise interactions. Therefore, the Elofsson group is investigating several ways to optimize AlphaFold2 using the trainable open-source version called OpenFold⁵.

AlphaFold2 relies on co-evolutionary information in the form of multiple sequence alignments (MSAs) for its structure prediction⁶. However, other methods have tried to replace the MSA generation step and instead rely on protein language models. There are also methods that try to predict protein-protein interactions based on sequences instead of structures. The methods included in this benchmark are AlphaFold2⁶, OmegaFold⁷, ESMFold2⁸, and 3 different versions of D-Script⁹. These methods cover both protein language models with OmegaFold and ESMFold2 and sequence-based methods with the different versions of D-Script. The benchmark study only includes dimers and not multimers of higher stoichiometry due to the fact that the different D-Script methods are only applicable to dimers.

In order to make AlphaFold2 more efficient, we located the parts of AlphaFold that are the most time-consuming. Disregarding MSA generation, the most time-consuming part of AlphaFold2 is the Evoformer. In order to speed up the entire AlphaFold2 model, we are therefore looking at ways to make the Evoformer more efficient. One approach that is being tested in the lab is to retrain AlphaFold2 with a smaller version of the EvoFormer and investigate whether the model is still able to separate interacting and non-interacting pairs. Inspired by a recent 12-block release by AlphaFlow, the lab is testing whether

a monomer 12-layer version of OpenFold is able to perform similarly to a monomer version of AlphaFold2 on PPI prediction. In order to predict multimers with a monomer version of OpenFold the input is changed in the same fashion as in FoldDock, released by the Elofsson lab in 2021¹⁰.

Other than changing the amount of blocks in the Evoformer, we are also working on other optimizations. At the core of the Evoformer, we find transformers¹¹, a popular neural network component that has shown great results in various domains thanks to its ability to extrapolate valuable relationships in long sequences. A major drawback is that this component scales quadratically with the sequence length making it difficult to perform efficiently for long sequences. AlphaFold2 uses a modified version of transformers, called Invariant Point Attention (IPA) that scales cubically with the sequence length. A recent alternative to transformers has been State Space Models (SSM) and in particular, the Mamba architecture¹² has shown promising results obtaining a linear scaling with the sequence length while keeping similar performances. The lab is now testing a modified version of Openfold, where the transformer layers in the Evoformer are substituted with Mamba layers in order to try to improve the Openfold computational efficiency, especially for proteins with longer sequences.

The question of predicting RNA structure is often divided in two parts consisting of inferring RNA secondary structure and predicting RNA 3D structure. Recent Ribonanza challenge at Kaggle enabled participation, training, and evaluation of different deep-learning models and led to state-of-the-art performance in RNA secondary structure prediction¹³. Therefore, we aim to utilize secondary structure information obtained from this model in combination with diverse machine learning models such as AlphaFold3¹⁴, HelixFold3¹⁵, and RosettaFoldNA¹⁶ for predicting 3D RNA structure to benchmark and evaluate results across different models.

Results

The unpublished results of the PPI benchmark study show that AlphaFold2.3 is the highest-performing model. It also shows that the structure-based models perform better than the sequence-based models. Interestingly, it also shows that there is a clear difference in performance of AlphaFold2 between homodimers and heterodimers. Future research in the lab will try to understand what could be behind this. It may be that the method used for the benchmarking was flawed or AlphaFold2 for unknown reasons simply does not perform equally well on heterodimers and homodimers.

Availability

This research is currently unpublished but all code related to the project will be up on Github or likewise.

1. Jun Zhang, Mei Lang, Yaoqi Zhou, Yang Zhang. Predicting RNA structures and functions by artificial intelligence. *Trends in Genetics*. 2024; 94-107.
2. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins*. 2021; 89(12): 1607-1617.
3. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. Critical assessment of methods of protein structure prediction (CASP)-Round XV. *Proteins*. 2023 Dec;91(12):1539-1549.

4. Aditi Zhu W, Shenoy A, Kundrotas P, Elofsson A. Evaluation of AlphaFold-Multimer prediction on multi-chain protein complexes. *Bioinformatics*. 2023 Jul 1;39(7):btad424. doi: 10.1093/bioinformatics/btad424. PMID: 37405868; PMCID: PMC10348836.
5. Ahdritz, G., Bouatta, N., Floristean, C. *et al.* OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nat Methods* 21, 1514–1524 (2024). <https://doi.org/10.1038/s41592-024-02272-z>
6. Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).
7. High-resolution de novo structure prediction from primary sequence Ruidong Wu, Fan Ding, Rui Wang, Rui Shen, Xiwen Zhang, Shitong Luo, Chenpeng Su, Zuofan Wu, Qi Xie, Bonnie Berger, Jianzhu Ma, Jian Peng *bioRxiv* 2022.07.21.500999
8. Zeming Lin *et al.*, Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379,1123-1130(2023).
9. Sledzieski, Samuel *et al.* D-SCRIPT translates genome to genome with sequence-based, structure-aware, genome-scale predictions of protein-protein interactions *Cell Systems*, Volume 12, Issue 10, 969 - 982.e6
10. Bryant, P., Pozzati, G. & Elofsson, A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat Commun* 13, 1265 (2022).
11. Vaswani, A., *et al.* (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
12. Dao, T., & Gu, A. (2024). Transformers are SSMs: Generalized Models and Efficient Algorithms Through Structured State Space Duality. *International Conference on Machine Learning (ICML)*.
13. Ribonanza: deep learning of RNA structure through dual crowdsourcing Shujun He, Rui Huang, Jill Townley *et al.* *bioRxiv* 2024.02.24.581671
14. Abramson, J., Adler, J., Dunger, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500 (2024).
15. Liu, Lihang & Zhang, Shanzhuo & Xue, Yang & Ye, Xianbin & Zhu, Kunrui & Li, Yuxin & Liu, Yang & Zhang, Xiaonan & Fang, Xiaomin. (2024). Technical Report of HelixFold3 for Biomolecular Structure Prediction.
16. Baek, M., McHugh, R., Anishchenko, I. *et al.* Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldDNA. *Nat Methods* 21, 117–121 (2024).

FALCON

Dongbo Bu, Chungong Yu, Haicang Zhang, Jianquan Zhao, Tian Zhu, Milong Ren, Yue Yu, Zhiyuan Wang, Qian Tan, Siyuan Tao

We use AlphaFold 3 server, FoldSeek, and ChatGPT as primary tools in our pipeline.

For regular monomers and complexes, we submit prediction tasks to AlphaFold3 server with at least 4 different seeds and select the best 5 prediction results as the final predictions based on the ranking score.

For multi-conformations, run at least 8 seeds for each target. Select the top prediction as conformation 1. Cluster the remaining predictions into two clusters based on TM-score relative to conformation 1. Then, select the top 5 predictions from each cluster as the final predictions.

For stage 2 targets:

- Use AlphaFold 3 to predict the target structure.
- Use FoldSeek to search the PDB template library.
- Use the protein ID to query GPT-4 for relevant structural information.

Based on the retrieved templates and GPT-4 responses, infer possible assembly strategies.

For each assembly strategy, use AlphaFold 3 to predict the structure and select the top 5 rank scores as the final prediction results.

Modeling of protein assemblies by applying pyDock energy-based scoring to AlphaFold models

L.A. Rodríguez-Lumbreras¹, V. Monteagudo¹, A. Jiménez-Panizo¹, F. Glaser² and J. Fernández-Recio¹

¹ - Instituto de Ciencias de la Vid y del Vino (ICVV), CSIC-UR-Gobierno de La Rioja, Logroño, Spain,

² - Technion-Israel Institute of Technology, Haifa, Israel

juan.fernandezrecio@icvv.es

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:N; DeepL:N; AF:Y; EMA:Y; MD:Y

We participated as predictors and scorers in the 70 targets of the 6th joint CASP-CAPRI Assembly Prediction challenge (CAPRI Round 57), as part of the CASP16 Assembly category. As human predictors, we participated in all targets (39 distinct targets plus 31 ones where the same target was released after providing MassiveFold dataset). We also participated in all 31 proposed targets for the scorers CAPRI experiment. We explored here new protocols to score the models generated by different versions of AlphaFold¹ with the energy-based scoring function in pyDock².

Methods

In most of the targets, the models of each proposed assembly submitted as predictors were built by AlphaFold2 (AF2)-Multimer³ and AlphaFold3 (AF3)⁴.

AF2-Multimer models: In the case of AF2-Multimer, for each of the three available versions (Version 1: v2.1.0; Version 2: v2.2.0; Version 3: v2.3.0) we run five trajectories. For each trajectory we kept the intermediate models generated in 21 recycling steps ($r0$ to $r20$) plus the top model selected and minimized by AlphaFold (22 models x 5 trajectories x 3 versions: 330 models in total per target). All recycling steps were further minimized using OpenMM. During the first targets, we observed that AF2 Version 1 produced an excessive number of clashes. To ensure consistency in the number of final models, in most targets after H1217/T258), Version 3 was run twice with different seeds, while Version 1 was not used.

AF3 models: In the case of AF3, we ran one job in the web server (<https://alphafoldserver.com>) to generate five models (by default). These models were added to the pool of models generated by AF2-Multimer as above described. In some targets, where AF2 produced significant clashes (e.g. M1239v1/T264, M1239v2/T265, M1282/T304, M1276/T306, M1287/T308), we used only AF3, generating a higher number of AF3 models by running several jobs (between 72 and 120) on the web server. In target H1227/T272, we used AF3 to generate partial models, which yielded a global model by superimposition. In the case of protein-DNA targets M1228v1/T254, M1228v2/T255, protein models generated by AF2 were superimposed onto the global AF3 protein-DNA model.

Checking structural quality: The pool of generated models (including the ones from MassiveFold dataset) were evaluated with different *ad-hoc* quality measures, like number of unstructured regions in the protein, maximum length of these regions, total number of clashes and their distribution across chains, percentage of residues in each chain with a pLDDT score below 50, cyclic symmetry between subunits, or presence of knots between chains. This filtering step aimed to discard models with significant symmetry discrepancies or critical unstructured regions, ensuring that the final models were of high quality.

Scoring: Then, all the generated models were sorted based on a new scoring function that combined AF model confidence (computed as $0.8*ipTM + 0.2*pTM$, hereinafter called AF-MC) and pyDock-IVDW score (which uses 100% van der Waals, instead of 10% van der Waals by default). To

combine these two heterogeneous values, we normalized them first by computing their z-scores and then summed them up, after changing the sign of the pyDock scores to have them in the same scale:

$$Z_{AF/pyDock-1VDW} = Z_{AF-MC} - Z_{pyDock-1VDW}$$

In the targets released after providing the MassiveFold dataset, all the models were sorted based only on *pyDock-1VDW* score. The same strategy was applied for the scorers CAPRI experiment to evaluate the scorer set plus the models from MassiveFold dataset.

Docking: In targets H1204/T240, M1282/T304, the AF2-Multimer/AF3 models were complemented with docking models generated by pyDock², using FTdock (electrostatics on; 0.7 Å grid resolution) on AF2 models of the subunits (for the nanobodies in T240 we used different MSA depths to increase conformational variability), to generate 10,000 rigid-body docking poses. For submission, we combined (in alternative order) the top 5 docking models from AF (AF/pyDock-1VDW scoring) and the top 5 models from pyDock (default scoring).

Clustering: Finally, a clustering strategy was developed during the first targets and was applied after target H1204/T240 (but discontinued after T1237/T282 due to lack of time). We grouped the models according to ligand RMSD similarity, using a 2 Å cutoff. If any of the top 5 ranked models were part of a large-sized cluster (> 10 models), the models were submitted without grouping them. Otherwise, we removed redundant models within 2 Å ligand RMSD. This approach ensured that, if the models built by AF clearly converged, the top 5 submitted models should be very similar, whereas if the AF models did not clearly converge, the 5 submitted models would be sufficiently varied.

Availability

The pyDock 3.0 program is available for academic use as a GNU/Linux binary and as a web server (<https://life.bsc.es/pid/pydock/>).

1. Jumper,J., Evans,R., Pritzel,A. et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**:583–589.
2. Cheng,T.M.-K., Blundell.T.L. & Fernandez-Recio,J. (2007) pyDock: electrostatics and desolvation for effective rigid-body protein-protein docking. *Proteins* **68**, 503-515.
3. Evans,R., O'Neill,M., Pritzel,A. et al. (2021) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*.
4. Abramson,J., Adler,J., Dunger,J., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500.

Template-Based Structure Prediction by using deep learning

Gwang So¹, Ryo-Yon Ri², Yong-Gong Sin¹, Si-Nae Kim³, Hyon-Gyong Jong³, Gyong-Jin Pang¹, Chung-Hyok Kang¹, Chang-Hyok Ri¹ and Gwang-Hyok Kim¹

1- Department of Bioinformatics, Branch of Biotechnology, State Academy of Sciences, Daedonggang-District, Pyongyang, DPR Korea, 2- Department of Agriculture and Life Science, Pyongyang University of Science and technology, Rakrang-District, Pyongyang, DPR Korea, 3- Department of Drug Analysis, Namsam Hospital, Daedonggang-District, Pyongyang, DPR Korea

sg1992@star-co.net.kp

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

In CASP16, we used the dPPAS alignment method based on PSI-blast¹ profile information and some structural information. And the database that consist of log-odds profile, relative solvent accessibility, and three-state secondary structure information of templates were constructed, and alignment between the target and the template was performed based on the dPPAS alignment method. Then, the quaternary structure of the target was modeled by MODELLER².

Methods

To obtain the log-odds profile of a target sequence, we used PSI-blast against the lastest Uniref50 database. To predict the three-state secondary structure information and relative solvent accessibility value of a target sequence, we used FTBiot method that is based a deep learning model.

Then, we used the profile, the three-statue secondary structure information, relative solvent accessibility value of target and template sequence with the dPPAS alignment method, to obtain pair alignment between a target and template sequence. The weight coefficients of this dPPAS alignment method were optimized through RNN models.

In our 3D-model construction step, wo used MODELLER² with the result of the pair alignment between a target and template sequence and PDB entry file of templates.

1. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. & Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.
2. Webb,B. and Sali,A. (2016) Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinforma., 2016, 5.6.1-5.6.37.

Assembly prediction with ab initio docking and template-based modeling

Gwang So¹, Ryo-Yon Ri², Yong-Gong Sin¹, Si-Nae Kim³, Hyon-Gyong Jong³, Gyong-Jin Pang¹, Chung-Hyok Kang¹, Chang-Hyok Ri¹ and Gwang-Hyok Kim¹

1- Department of Bioinformatics, Branch of Biotechnology, State Academy of Sciences, Daedonggang-District, Pyongyang, DPR Korea, 2- Department of Agriculture and Life Science, Pyongyang University of Science and technology, Rakrang-District, Pyongyang, DPR Korea, 3- Department of Drug Analysis, Namsam Hospital, Daedonggang-District, Pyongyang, DPR Korea

sg1992@star-co.net.kp

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

In CASP16, we used ab initio docking(zdock³), template-based modeling(MODELLER⁵), symmetry constraints.

Methods

For each assembly, the models of the individual subunits were taken from server models of the AlphaFold and Baker servers.

First, we searched template of target sequence against the PDB database² by using PSI-blast¹. If the target-like template is identified in PDB, we simulated the quaternary structure of the target using the multi-chain method of MODELLER. Here we use homo-types.

And for targets with a small number of subunits (2 ~ 3), initial models were obtained using zdock³, and excellent models were selected from them.

In addition, for targets in which subunits are in symmetrical conformations, symmdock⁴ was used to assemble the structures of subunits presented in the target, and excellent models were selected from the results.

In the final result model selection stage, the three-dimensional structure of the templates from the PSI-blast search results for the PDB database² was analyzed and final models were selected from the various docking results.

1. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. & Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
2. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res* 28, 235-242 (2000).
3. Chen,R., Li,L., Weng,Z. (2003) ZDOCK: An initial-stage protein-docking algorithm. *Proteins*. 52, 80–87.
4. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *NAR*, 33: W363-W367, 2005.
5. Webb,B. and Sali,A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.*, 2016, 5.6.1-5.6.37.

Computational modeling of nucleic acid 3D structures and interactions

E.F. Baulin, S. Mukherjee, S.N. Moafinejad, M.A. Farsani, R.J. Loureiro, K. Mondal, G.I. Nikolaev,
F. Stefaniak, and J.M. Bujnicki*

International Institute of Molecular and Cell Biology in Warsaw

*janusz@iimcb.gov.pl

Key: *Templ:Y; MSA:Y; DeepL:Y; AF:AF3; EMA:Y; MD:Y*

Structured RNA and DNA molecules are the master regulators of cells. They are involved in many molecular processes: They can transmit genetic information, sense cellular signals, relay responses, and catalyze chemical reactions. Their function, particularly their ability to interact with other molecules, is encoded in the sequence. Understanding how structured nucleic acid molecules perform their biological tasks requires detailed knowledge of spatial structure and dynamics, which determine how RNA/DNA folds and interacts in the cellular environment.

Methods

Our workflow for computational modeling of RNA and DNA 3D structures and their interactions with other molecules is based on a suite of methods developed in our laboratory, including PARNASSUS for remote homology detection, MeSSPredRNA and SQUARNA¹ for the prediction of canonical and non-canonical base pairs, ARTEM² and ARTEMIS³ for 3D structure alignment and tertiary motif search, ModeRNA⁴ for template-based RNA 3D structure modeling, the SimRNA-family⁵ of programs for the modeling of RNA or DNA 3D structure and its complexes with other molecules, and QRNAS⁶ for structure refinement.

Our automated SimRNA-server method involved RNA secondary structure prediction with SQUARNA and using the predicted structures as restraints for 3D simulations with SimRNA. Our automated GeneSilico-server method involved initial RNA 3D structure predictions with third-party programs, e.g., trRosettaRNA⁷, Nufold⁸, and Vfold⁹, that were then used to derive consensus distance restraints for SimRNA simulations. The results of the restrained simulations were submitted as the predicted models.

The workflow of the GeneSilico-Human group involved a substantial level of human intervention. Here we describe a general approach for typical RNA targets. Once we had a target RNA sequence, we started with a sequence search among known 3D structures in the RCSB database¹⁰ and well-studied RNA families in the Rfam database¹¹. If no homologs were found, we performed RNA secondary structure prediction using SQUARNA¹ and MeSSPredRNA. Additionally, we tried to identify known sequence motifs in the target sequence, e.g., U1A-binding sites, kink-turns, etc., and use them to guide the secondary

structure predictions. Subsequently, we used the predicted structures for remote homology detection using PARNASSUS. Then we proceeded to RNA 3D modeling. If close homologs with known 3D structures were present, we first performed template-based modeling with ModeRNA⁴. We carried out conformational space sampling with the SimRNA-family⁵ of tools using restraints obtained at all the early stages of the analysis. Additionally, we also run third-party 3D structure prediction tools, e.g., RNAComposer¹², RNA-BRiQ¹³, RNAJP¹⁴, trRosettaRNA⁷, and AlphaFold3¹⁵. We used their results as starting models for our simulations and as alternative models.

The final set of models was ranked by scoring with tools such as cgRNASP¹⁶ and our in-house scoring functions, and based on discussions in our team. For visualizations and manipulations with structures we commonly used ChimeraX¹⁷. For the refinement of selected models, we used QRNAS⁶. If time permitted in the context of CASP deadlines, we optimized the models with RNA-BRiQ¹³. For formatting the final models we used rna-pdb-tools¹⁸ and ARTEMIS³ along with other custom scripts.

Availability

SimRNA is available as a standalone tool at <http://genesilico.pl/software/stand-alone/simrna>

SimRNA is available as a web server at <https://genesilico.pl/SimRNAweb>

SQUARNA is available as a standalone tool at <https://github.com/febos/SQUARNA>

ARTEMIS is available as a standalone tool at <https://github.com/david-bogdan-r/ARTEMIS>

ARTEM is available as a standalone tool at <https://github.com/david-bogdan-r/ARTEM>

ModeRNA is available as a web server at <https://www.genesilico.pl/moderna/>

QRNAS is available as a standalone tool at <https://genesilico.pl/software/stand-alone/qrnas>

Other elements of our computational workflow are experimental and are not yet available.

1. Bohdan, D. R., Nikolaev, G. I., Bujnicki, J. M., & Baulin, E. F. (2023). SQUARNA - an RNA secondary structure prediction method based on a greedy stem formation model. bioRxiv, 2023-08.
2. Bohdan, D. R., Voronina, V. V., Bujnicki, J. M., & Baulin, E. F. (2023). A comprehensive survey of long-range tertiary interactions and motifs in non-coding RNA structures. Nucleic Acids Research, 51(16), 8367-8382.
3. Bohdan, D. R., Bujnicki, J. M., & Baulin, E. F. (2024). ARTEMIS - a method for topology-independent superposition of RNA 3D structures and structure-based sequence alignment. Nucleic Acids Research, gkae758.
4. Rother, M., Rother, K., Puton, T., & Bujnicki, J. M. (2011). ModeRNA: a tool for comparative modeling of RNA 3D structure. Nucleic acids research, 39(10), 4007-4022.
5. Moafinejad, S. N., de Aquino, B. R., Boniecki, M. J., Pandaranadar Jeyeram, I. P., Nikolaev, G., Magnus, M., ... & Bujnicki, J. M. (2024). SimRNAweb v2. 0: a web server for RNA folding simulations and 3D structure modeling, with optional restraints and enhanced analysis of folding trajectories. Nucleic Acids Research, gkae356.

6. Stasiewicz, J., Mukherjee, S., Nithin, C., & Bujnicki, J. M. (2019). QRNAs: software tool for refinement of nucleic acid structures. *BMC structural biology*, 19, 1-11.
7. Wang, W., Feng, C., Han, R., Wang, Z., Ye, L., Du, Z., ... & Yang, J. (2023). trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nature Communications*, 14(1), 7266.
8. Kagaya, Y., Zhang, Z., Ibtehaz, N., Wang, X., Nakamura, T., Huang, D., & Kihara, D. (2023). NuFold: a novel tertiary RNA structure prediction method using deep learning with flexible nucleobase center representation. *bioRxiv*.
9. Li, J., Zhang, S., Zhang, D., & Chen, S. J. (2022). Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. *Bioinformatics*, 38(16), 4042-4043.
10. Burley, S. K., Bhikadiya, C., Bi, C., Bitrich, S., Chao, H., Chen, L., ... & Zardecki, C. (2023). RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic acids research*, 51(D1), D488-D508.
11. Kalvari, I., Nawrocki, E. P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., ... & Petrov, A. I. (2021). Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research*, 49(D1), D192-D200.
12. Biesiada, M., Purzycka, K. J., Szachniuk, M., Blazewicz, J., & Adamiak, R. W. (2016). Automated RNA 3D structure prediction with RNAComposer. *RNA Structure Determination: Methods and Protocols*, 199-215.
13. Xiong, P., Wu, R., Zhan, J., & Zhou, Y. (2021). Pairing a high-resolution statistical potential with a nucleobase-centric sampling algorithm for improving RNA model refinement. *Nature communications*, 12(1), 2777.
14. Li, J., & Chen, S. J. (2023). RNAJP: enhanced RNA 3D structure predictions with non-canonical interactions and global topology sampling. *Nucleic acids research*, 51(7), 3341-3356.
15. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3.
16. Tan, Y. L., Wang, X., Yu, S., Zhang, B., & Tan, Z. J. (2023). cgRNASP: coarse-grained statistical potentials with residue separation for RNA structure evaluation. *NAR Genomics and Bioinformatics*, 5(1), lqad016.
17. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., ... & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein science*, 30(1), 70-82.
18. Magnus, M., Antczak, M., Zok, T., Wiedemann, J., Lukasiak, P., Cao, Y., ... & Miao, Z. (2020). RNA-Puzzles toolkit: a computational resource of RNA 3D structure benchmark datasets, structure manipulation, and evaluation tools. *Nucleic acids research*, 48(2), 576-588.

GHZ-ISM, GHZ-MAN

UniFold, AlphaFold2, AlphaFold2.3, AlphaFold3 and updated DMFold-based Optimization for Protein Structure Prediction

Weikang Gong^{1, #}, Biao Zhang^{1, #}, Jun Hu^{1, #}, Jun Liu² and Yang Zhang^{1,2,3,4, *}

1- Center for AI and Computational Biology, Institute of Systems Medicine, Chinese Academy of Medical Sciences /Suzhou Institute of Systems Medicine;

2- Cancer Science Institute of Singapore, National University of Singapore

3- Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore

4- Department of Computer Science, School of Computing, National University of Singapore

#co-first authors, *corresponding author: zhang@zhanggroup.org

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3/AF2; EMA:Y; MD:Y.

The protein structure prediction of the GHZ-ISM and GHZ-MAN group in CASP16 is based on a pipeline combining UniFold, AlphaFold2, AlphaFold2.3, AlphaFold3 and updated DMFold. We selected the structures predicted by these methods, performed energy optimization scoring, and chose the top five models as the prediction results.

Methods

The full pipeline contains three steps: (1) the protein monomer and complex structure prediction through UniFold, AlphaFold2, AlphaFold2.3, AlphaFold3 and updated DMFold, (2) all structures were put together and near-native structure selection was performed using SPICKER, (3) the selected structures are optimized and scored.

Availability:

<https://github.com/dptech-corp/Uni-Fold>

<https://github.com/google-deepmind/alphafold>

<https://alphafoldserver.com/>

<https://zhanggroup.org/DMFold/>

<https://zhanggroup.org/SPICKER/>

1. Li Z, Liu X, Chen W, et al. Uni-Fold: an open-source platform for developing protein folding models beyond AlphaFold[J]. bioRxiv, 2022: 2022.08. 04.502811.
2. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. nature, 2021, 596(7873): 583-589.
3. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer[J]. biorxiv, 2021: 2021.10. 04.463034.
4. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. Nature, 2024: 1-3.
5. Zheng W, Wuyun Q, Freddolino P L, et al. Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15[J]. Proteins: Structure, Function, and Bioinformatics, 2023, 91(12): 1684-1703.
6. Zhang Y, Skolnick J. SPICKER: a clustering approach to identify near-native protein folds[J]. Journal of computational chemistry, 2004, 25(6): 865-871.

Biomolecular structure prediction and accuracy estimation methodology for human prediction using different modeling techniques

K.Harini, Sowmya Ramaswamy Krishnan, Divya Sharma, Fathima Ridha, Preethi P, Amit Phogat, P Ramakrishna Reddy, Suraj Kumar Shah, Sankaran Venkatachalam, S Lekshmi, Subramanian Bhoopathi, Kulandaisamy Arulsamy, Pragathi Sneha Nela, Prabakaran R, Anoosha Paruchuri, Dhanusha Yesudhas, N R Siva Shanmugam, Rahul Nikam, Ambuj Srivastava, Medha Pandey, M. Michael Gromiha*

Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

gromiha@iitm.ac.in

Key: Auto:N; CASP_serv:N; Templ:Y; DeepL:Y; AF3:Y; EMA:N; MD:Y

We have collected the available experimental data (domains, binding sites, surface residues, interactions etc.) about the target protein/complex from the literature. The structures were predicted using existing tools and the best-predicted models were selected by comparing experimental data, domain-knowledge and human expertise. The structures were tailored to fulfill the constraints and optimized using energy minimization. The best structures were selected based on the adaptability of physical interactions, disorderness, as well as predicted free energy of binding and stability.

Methods

Protein structure: Initially, detailed information about target proteins were obtained using BLASTp against the UniProt database, and homologous structures, if available, were identified from the Protein Data Bank (PDB) [1,2]. We used the template if similar structures were reported in PDB and utilized different prediction methods, such as AlphaFold3, I-TASSER, GalaxyWeb, ESM Fold, and trRosetta for generating initial structures [3-7]. Further, the structures were energy minimized and loops were refined using the MODELLER loop optimization option [8]. Finally, the top five models were selected based on information gathered from the literature and domain knowledge. In addition, the stability of the proteins was predicted using FoldX and other in-house methods such as TMH Stab-pred, and TMB Stab-pred for membrane proteins [9,10].

RNA structure: Initially, we collected secondary structures from literature if available, or predicted using IPknot, RNAFold, and CentroidFold [11-13]. Utilizing this information, tertiary structural models were obtained using trRosetta, and FARFAR2 methods [14,15]. In addition, tertiary structures were also generated using AlphaFold3 [16]. Final models were selected based on structural homologs or evidence from the literature, such as the presence of higher-order structures such as pseudoknots gathered from experiments.

Protein-protein complexes: In the case of protein-protein complexes, we initially identified the binding site and stoichiometry (for the Phase 0 targets), based on either homologous PDB structures or the experimental evidence available in the literature. In addition, for the antigen-antibody binding site, we predicted the CDR regions from the ANARCI server [17]. Further, the monomers were modeled using

similar procedures mentioned above, and the complexes were docked using docking platforms such as GRAMM, LZeroD, SymmDock, SAM, and ClusPro with the binding constraints obtained from the literature [18-22]. Final models were selected based on binding free energy, binding site residues, and acceptable physical interactions.

Protein-ligand docking: For the ligand targets, we identified the ligand binding sites from the literature and performed rigid docking at the site using AutoDock [23]. Multiple poses were generated for each ligand, and the final models were selected based on their binding affinities and interactions with the target.

RNA/DNA-ligand docking: For the nucleic acid ligand targets, the ligand binding sites were identified from the literature, and rigid docking was performed on the site using rDock [24]. When the binding site was unknown, such as in the case of ZTP riboswitch, the position of catalytic magnesium ion in the template PDB structure was used to select the binding site through the rDock blind docking protocol. Multiple poses were generated for each ligand, and the final models were selected based on their binding affinities and interactions with the target.

Quality assessment: For the quality assessment, we sorted the high-quality structures based on the percentage of residues with prediction confidence above 0.75 and the highest average confidence score. Models with the highest average scores and percentage of residues with confidence above 0.75 were selected to be the top models.

Results

We have followed the methodology discussed above to predict the structures of proteins, RNA as well as protein-protein, protein-DNA/RNA, protein-ligand and RNA-ligand complexes. It is mainly based on experimental data available in the literature, domain-knowledge, predicted structures from existing software, adaptability of physical interactions and energetic contributions. The top five models obtained were submitted to CASP16. Based on the released PDB structures, our predicted structures for H1202 and R1203 targets showed RMSD values of 0.81 Å and 2.96 Å, respectively.

1. Ye,J., McGinnis,S., & Madden,T. L. (2006). BLAST: improvements for better sequence analysis. *Nucleic acids research*, 34(suppl_2), W6-W9.
2. Bittrich,S., Bhikadiya,C., Bi,C., Chao,H., Duarte,J. M., Dutta,S., Fayazi,M., Henry,J., Khokhriakov,I., Lowe,R., Piehl,D. W., Segura,J., Vallat,B., Voigt,M., Westbrook,J. D., Burley,S. K., & Rose,Y. (2023). RCSB Protein Data Bank: Efficient Searching and Simultaneous Access to One Million Computed Structure Models Alongside the PDB Structures Enabled by Architectural Advances. *Journal of molecular biology*, 435(14), 167994.
3. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *nature*, 596(7873), 583-589.
4. Zhang,Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC bioinformatics*, 9, 1-8.
5. Blankenberg,D., Kuster,G. V., Coraor,N., Ananda,G., Lazarus,R., Mangan,M., & Taylor,J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. *Current protocols in molecular biology*, 89(1), 19-10.
6. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637), 1123-1130.

7. Du,Z., Su,H., Wang,W., Ye,L., Wei,H., Peng,Z., & Yang,J. (2021). The trRosetta server for fast and accurate protein structure prediction. *Nature protocols*, 16(12), 5634–5651.
8. Fiser,A., Do,R. K. G., & Šali,A. (2000). Modeling of loops in protein structures. *Protein science*, 9(9), 1753–1773.
9. Schymkowitz,J., Borg,J., Stricher,F., Nys,R., Rousseau,F., & Serrano,L. (2005). The FoldX web server: an online force field. *Nucleic acids research*, 33(suppl_2), W382-W388.
10. Reddy,P. R., Kulandaisamy,A., & Gromiha,M. M. (2023). TMH Stab-pred: Predicting the stability of α -helical membrane proteins using sequence and structural features. *Methods*, 218, 118–124.
11. Sato,K., Kato,Y., Hamada,M., Akutsu,T., & Asai,K. (2011). IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics* (Oxford, England), 27(13), i85–i93.
12. Gruber,A. R., Lorenz,R., Bernhart,S. H., Neuböck,R., & Hofacker,I. L. (2008). The Vienna RNA websuite. *Nucleic acids research*, 36(Web Server issue), W70–W74.
13. Sato,K., Hamada,M., Asai,K., & Mituyama,T. (2009). CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic acids research*, 37(Web Server issue), W277–W280.
14. Du,Z., Su,H., Wang,W., Ye,L., Wei,H., Peng,Z., Anishchenko,I., Baker,D., & Yang,J. (2021). The trRosetta server for fast and accurate protein structure prediction. *Nature protocols*, 16(12), 5634–5651.
15. Watkins,A. M., Rangan,R., & Das,R. (2020). FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure* (London, England : 1993), 28(8), 963–976.e6.
16. Abramson,J., Adler,J., Dunger,J., Evans,R., Green,T., et. al., (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500.
17. Dunbar,J., & Deane,C. M. (2016). ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* (Oxford, England), 32(2), 298–300.
18. Singh,A., Copeland,M. M., Kundrotas,P. J., & Vakser,I. A. (2024). GRAMM Web Server for Protein Docking. *Methods in molecular biology* (Clifton, N.J.), 2714, 101–112.
19. Christoffer,C., Chen,S., Bharadwaj,V., Aderinwale,T., Kumar,V., Hormati,M., & Kihara,D. (2021). LZerD webserver for pairwise and multiple protein-protein docking. *Nucleic acids research*, 49(W1), W359–W365.
20. Schneidman-Duhovny,D., Inbar,Y., Nussinov,R., & Wolfson,H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research*, 33(Web Server issue), W363–W367.
21. Kozakov,D., Hall,D. R., Xia,B., Porter,K. A., Padhorny,D., Yueh,C., Beglov,D., & Vajda,S. (2017). The ClusPro web server for protein-protein docking. *Nature protocols*, 12(2), 255–278.
22. Ritchie,D.W. and Grudinin,S. (2016) Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry, *Journal of Applied Crystallography*, 49(1), 158–167.
23. Goodsell,D. S., Sanner,M. F., Olson,A. J., & Forli,S. (2021). The AutoDock suite at 30. *Protein science* : a publication of the Protein Society, 30(1), 31–43.
24. Ruiz-Carmona,S., Alvarez-Garcia,D., Foloppe,N., Garmendia-Doval,A. B., Juhos,S., Schmidtke,P., Barril,X., Hubbard,R. E., & Morley,S. D. (2014). rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS computational biology*, 10(4), e1003571.

Predicting protein-ligand binding with Convex-PL and KORP-PL tools

S. Grudinin

Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

sergei.grudinin@univ-grenoble-alpes.fr

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:N; DeepL:N; AF:N, EMA:N; MD:N

In CASP16, we assessed our scoring techniques specifically developed for small-molecule docking. These include docking scoring functions Convex-PL^R^{1,2}, KORP-PL^W, and KORP-PL³.

Methods

We have submitted predictions for Stage 2 affinity prediction experiment only, targets L1000_exper, L2000_exper, L3000_exper, and L4000_exper. We used experimentally-determined models of protein-ligand complexes. For all targets, we computed docking scores with Convex-PL^R¹ (Convex-PL-R team), KORP-PL^W³ (KORP-PL-W team) and KORP-PL³ (GruLab team). Parametrization of the molecules was done using the built-in Knodle atom types⁴. We then converted the scores, based on our previous analysis^{1,3}, into binding affinities.

Availability

Our methods are available on our website at <https://grulab.imag.fr/#software>.

1. Kadukova,M., Chupin,V., & Grudinin,S. (2021). Convex-PLR-Revisiting affinity predictions and virtual screening using physics-informed machine learning. *bioRxiv*.
2. Kadukova,M., & Grudinin,S. (2017). Convex-PL: a novel knowledge-based potential for protein-ligand interactions deduced from structural databases using convex optimization. *J. Comput. Aided Mol. Des.* **31**, 943-958.
3. Kadukova,M., Machado,K.D.S., Chacón,P., & Grudinin,S. (2021). KORP-PL: a coarse-grained knowledge-based scoring function for protein-ligand interactions. *Bioinformatics* **37**, 943-950.
4. Kadukova,M. & Grudinin,S. (2016) Knodle, a Support Vector Machines-based automatic perception of organic molecules from 3D coordinates, *J. Chem. Inf. Model.*, **56**, 8, 1410–1419.

Enhancing RNA 3D Structure Prediction: Integrating Computational Tools with Expert Refinement for Functional Insights

Bowen Xiao^{1#}, Yaohuang Shi¹, Jiajun Lin¹, Jiaxin Zhao¹, Lin Huang^{2#}

1- Guangzhou National Laboratory, Guangzhou International Bio Island, Guangzhou 510005, China;

2- Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangdong-Hong Kong Joint Laboratory for RNA Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, China

xiao_bowen@gzlab.ac.cn, huanglin36@mail.sysu.edu.cn

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3; EMA:N; MD:Y.

The accurate prediction of RNA three-dimensional (3D) structures is crucial for deciphering their biological functions. Here, we present the method we used in CASP16. For targets with homologous templates identified through literature review and sequence search, we employ MSA-based homology modelling for prediction. Conversely, for those without identifiable homologous templates, we refine predictions from open-access RNA 3D prediction software by incorporating human expertise in functional RNA 3D structures, adjusting the topology to better reflect a presumed functional state.

Methods

Targets with homologous templates

For targets where homologous templates can be identified through comprehensive literature review and sequence searches, we utilize multiple sequence alignment (MSA)-based homology modeling to predict RNA 3D structures. This approach involves aligning the target sequence with known homologous sequences to construct a consensus model. The structural information from the homologous templates is then used to guide the modeling process, allowing for the accurate prediction of the target's 3D conformation. The strategy leverages evolutionary conservation and structural similarities to enhance the reliability of the predicted models. Based on high-accuracy manual multiple sequence alignment, we edit the homologous templates which download from PDB(Burley et al. 2022) while maintaining the global topology unchanged. Additionally, refinement techniques such as energy minimization and BRIQ(Xiong et al. 2021) have been employed to further improve the accuracy and stability of the predicted structures.

Targets without homologous templates

For targets lacking homologous templates, we employ multiple RNA structure 3D prediction software, such as AlphaFold3(Abramson et al. 2024), to generate initial models. Based on the expertise of structural biologists, we acknowledge that these preliminary models, while aligning with the secondary structures defined by Rfam⁴, their global topology often lacks the characteristics of functional RNA 3D structures. Therefore, it is necessary to refine the global topology of these initial structures to achieve a more organized and functionally relevant conformation. Tools such as PyMOL(DeLano 2002) and Coot(Emsley et al. 2010) are utilized in this refinement process. This approach enhances the accuracy and reliability of RNA 3D structure predictions, particularly in scenarios where homologous templates are unavailable. By

integrating computational predictions with expert-guided structural adjustments, we aim to produce models that more accurately reflect the functional states of RNA molecules.

1. Abramson, Josh, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, et al. 2024. “Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3.” *Nature* 630 (8016): 493–500.
2. Burley, Stephen K., Charmi Bhikadiya, Chunxiao Bi, Sebastian Bittrich, Henry Chao, Li Chen, Paul A. Craig, et al. 2022. “RCSB Protein Data Bank (RCSB.org): Delivery of Experimentally-Determined PDB Structures alongside One Million Computed Structure Models of Proteins from Artificial Intelligence/machine Learning.” *Nucleic Acids Research* 51 (D1): D488–508.
3. DeLano, W. L. 2002. “Pymol: An Open-Source Molecular Graphics Tool.” *CCP4 Newslett. Protein Crystallogr.* <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.231.5879&rep=rep1&type=pdf#page=44>.
4. Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. “Features and Development of Coot.” *Acta Crystallographica. Section D, Biological Crystallography* 66 (Pt 4): 486–501.
5. Kalvari, Ioanna, Eric P. Nawrocki, Nancy Ontiveros-Palacios, Joanna Argasinska, Kevin Lamkiewicz, Manja Marz, Sam Griffiths-Jones, et al. 2021. “Rfam 14: Expanded Coverage of Metagenomic, Viral and microRNA Families.” *Nucleic Acids Research* 49 (D1): D192–200.
6. Xiong, Peng, Ruibo Wu, Jian Zhan, and Yaoqi Zhou. 2021. “Pairing a High-Resolution Statistical Potential with a Nucleobase-Centric Sampling Algorithm for Improving RNA Model Refinement.” *Nature Communications* 12 (1): 2777.

Integrating multiple tools for RNA 3D structure prediction in CASP16

Yaohuang Shi¹, Bowen Xiao^{1#}, Jiajun Lin¹, Jiaxin Zhao¹, Lin Huang^{2#}

1- Guangzhou National Laboratory, Guangzhou International Bio Island, Guangzhou 510005, China;

2- Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangdong-Hong Kong Joint Laboratory for RNA Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, China

shi_yaohuang@gzlab.ac.cn, huanglin36@mail.sysu.edu.cn

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:AF3; EMA:N; MD:Y.

Accurately predicting RNA three-dimensional (3D) structures is essential for elucidating their biological functions. Our strategy for the CASP16 experiment involved a synergistic application of various techniques to predict the structure of RNA molecules. We utilized a comprehensive set of structural constraints to discern and select the top five most plausible 3D models."

Methods

Collect structure information: We first collect constraint information that can be used to judge the prediction of RNA targets, such as the template structure, secondary structure prediction information, or multiple sequence alignments of homologous RNA sequences.

Generation prediction structure: Multiple advanced RNA structure prediction tools (DeepfoldRNA(Pearce, Omenn, and Zhang 2022), trRosettaRNA(Wang et al. 2023), AlphaFold3(Abramson et al. 2024), FARFAR2(Watkins, Rangan, and Das 2020), RNAComposer(Biesiada et al. 2016), etc.) were run on a given sequence, and then the prediction results were screened based on the collected structural restriction information and functional description.

Filter Submission structure: The selection criteria primarily focused on thermodynamic stability, evolutionary conservation, and recognized structural motifs. We then identified the most plausible structure or integrated results from various predictive algorithms to formulate the five submitted models. For DeepfoldRNA and FARFAR2, local installations were used with default parameters for predictions. Other tools are accessed through online servers, as detailed below:

Availability

trRosettaRNA online server: <https://yanglab.qd.sdu.edu.cn/trRosettaRNA/>

AlphaFold3 online server: <https://alphafoldserver.com/>

RNAComposer online server:<https://rnacomposer.cs.put.poznan.pl/>

1. Abramson, Josh, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, et al. 2024. “Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3.” *Nature* 630 (8016): 493–500.
2. Biesiada, Marcin, Katarzyna J. Purzycka, Marta Szachniuk, Jacek Blazewicz, and Ryszard W. Adamiak. 2016. “Automated RNA 3D Structure Prediction with RNAComposer.” *RNA Structure Determination*, 199–215.
3. Pearce, Robin, Gilbert S. Omenn, and Yang Zhang. 2022. “De Novo RNA Tertiary Structure Prediction at Atomic Resolution Using Geometric Potentials from Deep Learning.” *bioRxiv*. <https://doi.org/10.1101/2022.05.15.491755>.
4. Wang, Wenkai, Chenjie Feng, Renmin Han, Ziyi Wang, Lisha Ye, Zongyang Du, Hong Wei, Fa Zhang, Zhenling Peng, and Jianyi Yang. 2023. “trRosettaRNA: Automated Prediction of RNA 3D Structure with Transformer Network.” *Nature Communications* 14 (1): 7266.
5. Watkins, Andrew Martin, Ramya Rangan, and Rhiju Das. 2020. “FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds.” *Structure* 28 (8): 963–76.e6.

Protein multimer structure modeling using deep learning-based paired MSAs

Minghua Hou, Pengcheng Wang, Yuhao Xia, Xuanfeng Zhao, Dong Liu, Xinyue Cui, Kailong Zhao,
Suhui Wang, Fang Liang and Guijun Zhang*

¹ - College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

zgj@zjut.edu.cn

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:[AF3/AF2]; EMA:Y; MD:N

In CASP16 experiment, we developed a pipeline for protein multimer structure modeling focusing on the paired MSAs construction based on deep learning. We first developed a sequence-based deep learning model to predict protein-protein structural similarity (pSS-score) and inter-action probability (pIA-score). These predicted scores were used to efficiently construct diverse paired MSAs for multimer prediction by integrating them with species information. Subsequently, we fed the paired MSAs into AlphaFold-Multimer^{1,2} and our previously developed DeepAssembly³ to generate a series of multimer structures. Finally, our in-house multi-model quality assessment methods were used for model scoring and selection.

Methods

The key components of our method include: 1) a sequence-based deep learning model for predicting protein-protein structural similarity (pSS-score) and inter-action probability (pIA-score); 2) MSA sampling and pair-MSA construction using pSS-score and pIA-score; and 3) model quality assessment for model scoring and selection.

structural similarity and inter-action probability prediction

Pair-MSA is crucial for protein multimer structure modeling. However, the construction method that only uses traditional sequence similarity combined with existing complex structure information may still face difficulties in providing high-quality paired MSAs. In order to efficiently construct paired MSAs, we developed a sequence-based deep learning model to capture the relationships between sequences. Firstly, it extracts sequence-based physicochemical features and combines them with the feature obtained from a protein language model (PLM)⁴. Then, the input features are encoded into sequence representations through a self-attention module. Finally, the sequence representations of two query sequences are decoded through a cross-attention module to predict their structural similarity (pSS-score) and interaction probability (pIA-score).

MSA sampling and pair-MSA construction

In order to obtain more diverse MSAs, we first search multiple sequence databases (UniRef30, UniRef90, UniProt⁵, BFD⁶, MGnify⁷, and the ColabFold DB⁸) to obtain the MSAs of each monomer. Then, we use the predicted pSS-score as a supplement to the traditional sequence similarity in the ranking and selection of the searched MSAs. For trimers, we use the developed deep learning model to predict the pIA-scores for each pair of sequence alignments from different subunit MSAs. The subunit MSAs are then concatenated based on these interaction probabilities to construct paired MSAs. Additionally, we use information from multiple sources, such as species annotations, UniProt accession number, and protein complexes from the Protein Data Bank (PDB), to further construct a series of paired MSAs. The paired

MSAs constructed above are used in the structure predictor to enhance diversity and prevent falling into local optimum.

Model quality assessment

We use our in-house developed model quality assessment methods for scoring and selection, as detailed in the method QA3 of GuijunLab-QA.

Availability

DeepAssembly is available at <http://zhanglab-bioinf.com/DeepAssembly/>

1. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. Nature, 2021, 596(7873): 583-589.
2. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer[J]. biorxiv, 2021: 2021.10. 04.463034.
3. Xia Y, Zhao K, Liu D, et al. Domain-based multi-domain protein and complex structure prediction using inter-domain interactions from deep learning[J]. bioRxiv, 2023: 2023.05. 05.539574.
4. Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model[J]. Science, 2023, 379(6637): 1123-1130.
5. Consortium U. UniProt: a worldwide hub of protein knowledge. Nucleic acids research 2019;47:D506-D515.
6. Steinegger M, Mirdita M, Söding J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. Nature methods 2019;16:603-606.
7. Mitchell AL, Almeida A, Beracochea M, et al. MGnify: the microbiome analysis resource in 2020. Nucleic acids research 2020;48:D570-D578.
8. Mirdita M, Schütze K, Moriwaki Y, et al. ColabFold: making protein folding accessible to all. Nature Methods 2022:1-4.

Protein complex modeling based on inter-chain interactions obtained by deep learning and traditional docking methods

Xinyue Cui, Yuhao Xia, Minghua Hou, Dong Liu, Kailong Zhao, Suhui Wang, Xuanfeng Zhao, Fang Liang and Guijun Zhang

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

zgj@zjut.edu.cn

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; Dist:Y; DeepL:Y; EMA:Y; MD:N

In this CASP16, we developed a protein complex assembly method based on the predicting inter-chain interactions using deep learning and traditional docking methods. First, the monomer and the corresponding complex structure are predicted using the Guijunlab-complex, AlphaFold2¹, AlphaFold3², and HDOCK³. Then the inter-chain distances of the complex structure were extracted and input into DeepAssembly⁴, together with the monomer structures to predict the complex structure. Finally, the models were selected using our in-house model quality assessment method.

Methods

The key components of our method include: 1) protein complex structure assembly, and 2) model quality assessment.

Protein complex structure assembly

In this method, we first predicted the complex structures using Guijunlab-complex, AlphaFold2¹, AlphaFold3². To further ensure the diversity of the complex structures, we employed the traditional HDOCK³ (monomer structures generated by AlphaFold3) method to generate 100 additional complex structures. Subsequently, the predicted complex structures from these methods were ranked using our in-house model quality assessment (GuijunLab-QA:QA4). From the ranked results, the top 50 structures were selected, split into monomers as the basic units and extracted the inter-chain distance as the interaction constraints for assembly. During the protein modeling process, the monomer structures are fed into the DeepAssembly⁴, and the protein monomers are rotated and translated using an evolutionary algorithm under the guidance of different interaction constraints to generate a series of complex structures. Finally, the model quality assessment (GuijunLab-QA:QA4) used to select the six best structures from the multiple structures generated.

Model quality assessment

We use our in-house developed model quality assessment methods for scoring and selection, as detailed in the method QA4 of GuijunLab-QA.

Availability

DeepAssembly is available at <http://zhanglab-bioinf.com/DeepAssembly/>

1. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. *nature*, 2021, 596(7873): 583-589.
2. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. *Nature*, 2024: 1-3.
3. Yan, Yumeng & Zhang, Di & Zhou, Pei & Li, Botong & Huang, Sheng-You. (2017). HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic acids research*. 45. 10.1093/nar/gkx407.
4. Xia Y, Zhao K, Liu D, et al. Multi-domain and complex protein structure prediction using inter-domain interactions from deep learning[J]. *Communications Biology*, 2023, 6(1): 1221.

Improving protein structure prediction with enhanced MSA and model quality assessment

Kailong Zhao, Dong Liu, Minghua Hou, Yuhao Xia, Xinyue Cui, Suhui Wang, Fang Liang,
Guijun Zhang*

¹ College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

* zgj@zjut.edu.cn

Key: Auto:Y; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; Fragm:Y; Cont:N; Dist:Y; Tors:Y; DeepL:Y; EMA:Y; MD:N

In CASP16, we used enhanced MSA and model quality assessment techniques to improve the accuracy of protein structure prediction. Starting from the query sequence of the target protein, the three-dimensional structure is first modeled by AlphaFold2, and remote homologous structures of the target are searched from the AlphaFold DB library through the Foldseek. Sequence information was then extracted from the searched homologous structures to construct high-quality MSA. At the same time, 10 MSAs were obtained with various cutoffs for e-value and coverage, and high-quality homologous templates were searched using our proposed PAthreader. Finally, multiple models were predicted by AlphaFold2 using enhanced MSA and homologous template, and the best model was selected through our in-house single-model quality assessment method and AlphaFold2's self-assessment.

Methods

MSA generation

In the case of highly distant evolutionary relationships, sequence similarity may become difficult to detect. However, the three-dimensional (3D) structure of proteins is often constrained by their functional requirements and remains more conserved. Even if the sequence changes greatly, the overall 3D shape of the protein may remain relatively stable. Therefore, detecting similarities between protein through 3D structure alignment can provide higher sensitivity.

We used two different methods to generate MSAs of targets. The first one used the structure predicted by AlphaFold2 as the input of Foldseek, and searched the AlphaFold DB database of 214 million structures to obtain a large number of remote homologous structures [1]. Then, sequence information was extracted from them to construct a high-quality MSA. The second one was to search unclust30, uniref90, MGnify and BFD data. 10 MSAs were obtained from this search with various e-value cutoffs (10^{-40} , 10^{-20} , 10^{-10} , 10^{-3} , 1) and coverage (50%, 80%). A total of 11 MSAs were generated by the two approaches.

Template recognition

Remote homologous templates were recognized by our proposed PAthreader [2][3]. First, multi-peak distance profiles are obtained by our in-house DeepMDisPre, which may predict multiple possible distances for flexible protein regions. Structural profiles were extracted from PAcluster80, a master structure database constructed by clustering PDB and AlphaFold DB with a structural similarity threshold of 80%. Then, the query sequence was aligned with each cluster seed of PAcluster80 by a three-track alignment algorithm, where the protein-specific score matrix was first calculated by residue alignment and profile alignment, and then the optimal sequence alignment was searched by dynamic programming and the maximum alignment score (alignScore) was obtained. Subsequently, physical and geometric

features were extracted from the aligned structure and input into a convolutional network with self-attention to predict the DMScore, a global structural scoring metric that is complementary to the alignScore and linearly weighted with the alignScore for template ranking.

Model quality assessment

The final 5 models were selected by linearly weighting the scores of our in-house single model quality assessment [4] and AlphaFold2's self-assessment [1]. The details of the model quality assessment method are described in the GuijunLab-QA group.

Availability

PAthreader is available at <http://zhanglab-bioinf.com/PAthreader>

1. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589 (2021).
2. Zhao K, Xia Y, Zhang F, Zhou X, Li SZ, Zhang G. Protein structure and folding pathway prediction based on remote homologs recognition using PAthreader. *Commun Biol*. 2023 Mar 4;6(1):243.
3. Zhao K, Zhao P, Wang S, Xia Y, Zhang G. FoldPAthreader: predicting protein folding pathway using a novel folding force field model derived from known protein universe. *Genome Biol*. 2024 Jun 11;25(1):152.
4. Liu D, Zhang B, Liu J, et al. Assessing protein model quality based on deep graph coupled networks using protein language model[J]. *Briefings in bioinformatics*, 2024, 25(1): bbad420.

Protein model quality assessment and selection

Dong Liu, Xuanfeng Zhao, Haodong Wang, Fang Liang, Minghua Hou, Xinyue Cui, Suhui Wang, Kailong Zhao, Yuhao Xia, and Guijun Zhang*

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

zgi@zjut.edu.cn

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; EMA: Y; MD: N;

In this CASP, we developed different protein model quality assessment and selection methods for five GuijunLab groups. This series of methods uses model quality assessment and structure prediction techniques including deep learning networks, structural consensus, protein language models, AlphaFold, ultrafast shape recognition, and combined methods.

Methods

Model quality assessment and selection of GuijunLab-QA (QA1)

For complex model quality assessment (MQA), we develop a multi-model approach based on global and local views using structural consensus to predict quality scores (Method A). For global view, we uses different scales to represent proteins and inputs a multi-level graph neural network to predict the overall quality of the complex (see Method E of QA3). For local view, we uses the embedding representation of the protein language models and AlphaFold2, supplemented by sequence and structural features, to predict the quality of the local interface score by an attention-based graph encoding and decoding module (see Method B of QA2). These methods evaluate all protein models to jointly select high-quality structures, which are used as references to predict the quality of complex models through structural alignment.

For MassiveFold¹ model selection, we developed a two-stage model evaluation approach. In the first stage, for eight structure prediction methods, we screened out eight high-quality candidate pools (Pool-Eight) through their self-assessed confidence score. In the second stage, the eight structure pools are aggregated into one model pool, and Method A is used to select the top 5 structures in this pool.

For complex structure prediction, we collect models from the structure prediction groups and AlphaFold^{2,3} to build a candidate structure pool. For the models in the structure prediction group, we selected structures submitted by the server group and the human group in phase 0/1. For models of AlphaFold, we predicted multiple structures using Alphafold2 and online server Alphafold3. Finally, we use Method A to evaluate all models and select the best five models.

Model quality assessment and selection of GuijunLab-PATHreader (QA2)

For complex MQA, we developed three single-model quality assessment methods for SCORE, QSCORE, Local evaluation metrics. For SCORE, based on our in-house MQA method GraphCPLMQA4, we improved the network architecture of the decoding module to capture the relationship between topological

structure and global quality, and used transfer learning way to train the network model (Method B). For QSCORE, we intercept interfaces on protein models and extract their sequence, structural, and physicochemical features, which are input into a transformer-based network to predict the quality scores of interactions (Method C). For Local, we introduced the representation of residue nodes and edges between residues of AlphaFold2 in GraphCPLMQA and improved the network architecture to combine protein language model (PLM)5 embedding and AlphaFold2 representation (Method D).

For MassiveFold model selection, inspired by VoroIF-jury6, we develop a multi-model evaluation method based on iterative structural consensus. First, we predict the quality of the models in Pool-Eight by these three MQA methods to obtain three different quality rankings for each pool. Secondly, in every pool, these differently ranked models were used to calculate the quality scores through iterative structural alignment, which is building scalable model pools based on ranking scores for structural consensus. Finally, between each model pool, we iteratively align different models to calculate the scores and select the top 5 models.

Model quality assessment and selection of GuijunLab-Complex (QA3)

For complex MQA, we developed a single model quality assessment method based on multi-scale protein representation using hierarchical networks (Method E). To represent the protein, we extract physical and geometric properties of residues at the 1D scale, interaction relationships of residues at the 2D scale, and structural topology at the 3D scale. These representations at different scales are fed into a multi-layer network module, including graph attention, transformer, and convolutional network, to predict model quality, where the output of the network can be changed according to the SCORE、QSCORE and Local evaluation of the downstream task.

For MassiveFold model selection, we develop a multi-model evaluation method based on structural consensus using Method E. First, this method uses AlphaFold2 and AlphaFold3 to generate multiple candidate conformations and selects the Top N predicted models as the high-quality model pool (Pool-N) through Method E. Secondly, all prediction models are aligned with the high-quality model pool to calculate the quality of each model, and the Top 5 high-quality models are selected.

Model quality assessment and selection of GuijunLab-Human (QA4)

For complex MQA, we use the Method A variant of QA1, only replacing the global and local view single model methods and some hyperparameters.

For MassiveFold model selection, we developed a multi-model quality assessment method based on structural consensus using GraphCPLMQA to select Top5 models. We first structurally clustered all models to select the structure of the cluster center. Secondly, we used our in-house modeling methods, which generated many candidate structures. Then, the structure of the cluster center and the candidate structure are input into the improved methods based on GraphCPLMQA to predict the quality score, where a high-quality model pool is constructed according to the score. Finally, the structure of the cluster center is aligned with the high-quality model pool to calculate the score, and the five models with the best quality are selected.

Model quality assessment and selection of GuijunLab-Assembly (QA5)

For complex MQA, we deployed two different quality assessment methods to evaluate the accuracy of models in SCORE, Local metrics. For SCORE, we used the MSA from Guijunlab-Complex as input of AF2, generating node and edge representations. These embeddings, combined with ESM2 embeddings, are then fed into a graph encoder-decoder module to predict the global quality of the protein model. For Local, we employed Voronoi diagrams⁶ to characterize atomic interaction interfaces for physical interactions, while utilizing distance and orientation between residues for geometric structures. Furthermore, these features are combined with the PLM embedding and feeding into a graph neural network to predict residue level quality. In addition, for QSCORE, we weight the outputs of Method D and Method F as the interface quality score.

For model selection, all structures in the Pool-Eight were evaluated by Method B of QA2 and GraphCPLMQA to jointly screen out high-quality prediction models. These models were ranked by weighted global quality and local interface quality, and the five models with the best scores were submitted.

1. Brysbaert G, Raouraoua N, Mirabello C, et al. MassiveFold: unveiling AlphaFold’s hidden potential with optimized and parallelized massive sampling[J]. 2024.
2. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. nature, 2021, 596(7873): 583-589.
3. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. Nature, 2024: 1-3.
4. Liu D, Zhang B, Liu J, et al. Assessing protein model quality based on deep graph coupled networks using protein language model[J]. Briefings in bioinformatics, 2024, 25(1): bbad420.
5. Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model[J]. Science, 2023, 379(6637): 1123-1130.
6. Olechnovič K, Venclovas Č. VoroIF-GNN: Voronoi tessellation-derived protein–protein interface assessment using a graph neural network[J]. Proteins: Structure, Function, and Bioinformatics, 2023, 91(12): 1879-1888.

SGraph_RMSD/affinity: predicting RMSD or affinity of protein-ligand complex by graph transformer with single graphic to represent the protein-ligand interface

Haiping Zhang

Faculty of Synthetic Biology and Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

hp.zhang@siat.ac.cn

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; EMA:Y; MD:N

Correctly identifying protein-ligand conformations that resemble natural states is crucial for obtaining reliable protein-ligand complexes. Accurate protein-ligand conformations are essential for subsequent target-based drug design and predicting structural protein-ligand affinity. Currently, effectively representing protein-ligand interaction space and physicochemical information as inputs for deep learning models is key. Many previous input representations suffer from information loss, sparsity, high dimensionality, and difficulty in ensuring rotational invariance.

In this CASP competition, we developed three models: SGraph_RMSD (for predicting the RMSD of docked protein-ligand complexes), Graph_RG (for predicting affinity when no complex is available, using separate graphs for the pocket and ligand), and SGraph_affinity (for predicting affinity based on the given protein-ligand complex interface). These models innovatively use a single graph to comprehensively represent the entire protein-ligand interface, aiming to preserve both the physicochemical properties of amino acids/atoms and spatial information. By leveraging powerful graph transformers, we efficiently identify critical interaction information. This results in SGraph_RMSD identifying near-natural binding conformations among a large set of UniDock-generated conformations. These models are expected to significantly enhance structure-based drug design screening.

SGraph_RMSD was used to predict RMSD values. Graph_RG was employed to predict affinity when no complex was available. SGraph_affinity was utilized to predict protein-ligand affinity after the protein-ligand complex was provided.

Methods

SGraph_RMSD

Data Preparation

We downloaded the PDBbind2020 database and used UniDock1 for docking protein-ligand pairs. Each pair generated up to 100 conformations, with DockRMSD computing RMSD values against the native ligand to serve as labels. Conformations resulting in docking failures or with RMSD values exceeding 5 Å were excluded, leaving a total of 990,000 conformations for training and 14,896 conformations for testing.

Input Data Representation

We represented interface amino acids and small molecule atoms as nodes in our graph. Edges were formed based on adjacency: an edge was defined between amino acids if the distance between C_α atoms was less than 5 Å, and between amino acids and small molecule atoms if the distance was less than 9 Å. For

ligands, edges were based on bond information. Amino acids were represented with 30-dimensional word vectors retrained using mol2vec, while small molecules were encoded similarly to one-hot encoding. To integrate amino acid and small molecule nodes into a unified graph, we standardized their vector lengths to 75 dimensions: amino acid vectors were padded from the back and small molecule vectors were padded with zeros from the front.

Model architecture

We adopt a Graph to represent the interface amino acids and small molecule atoms as nodes, using their adjacency relationships as edges. The model was built using PyTorch and PyTorch Geometric. The architecture includes two Transformer Convolution layers (TransformerConv) to capture complex interactions between nodes. The first Transformer Convolution layer has 5 attention heads, with both input and output dimensions set to num_features_xd. The second Transformer Convolution layer has 10 attention heads, with an input dimension of 5 times num_features_xd and an output dimension of 10 times num_features_xd. Each Transformer Convolution layer incorporates a dropout rate of 0.1 to prevent overfitting and accounts for edge dimensions (edge_dim_xd).

SGraph_affinity is an affinity prediction model based on interface conformation with single graph representation.

We developed an affinity prediction model called SGraph_affinity using the PDBbind V2020 database². This model employs the same architecture and input representation as the previously described SGraph_RMSD model. The affinity labels are represented as $-\log(P \cdot 10^{-9})$, where P represents the KD, KI, or IC₅₀ values in nM. The native protein-ligand interface was used to construct the graph representation for the input. We constructed the input graph representation based on the native protein-ligand interface. During application we use: affinity= $10^{(-\text{float}(\text{predict_value})) \cdot (10^9)}$ formula to convert the prediction back to nM unit.

Graph_RG is an affinity prediction model not dependent on protein-ligand interface.

Graph_RG uses training data similar to that of SGraph_affinity but operates independently of protein-ligand complex information. It separately uses the protein pocket and ligand as inputs. This approach is similar to our previously published model, DeepBindGCN_RG3, but it employs a graph transformer instead. The label unit is same as SGraph_affinity. During this competition, we have use: affinity= $10^{(-\text{float}(\text{predict_value})) \cdot (10^9)}$ formula to convert the prediction back to nM unit.

Availability

The proposed models and the scripts are available in GitHub public repositories (https://github.com/haiping1010/haiping_methods).

1. Yu, Y. et al. Uni-Dock: GPU-Accelerated Docking Enables Ultralarge Virtual Screening. *Journal of Chemical Theory and Computation* (2023) doi:10.1021/acs.jctc.2c01145.
2. Liu, Z. et al. PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics* 31, 405–412 (2015).
3. Zhang, H., Saravanan, K. M. & Zhang, J. Z. H. DeepBindGCN: Integrating Molecular Vector Representation with Graph Convolutional Neural Networks for Protein–Ligand Interaction Prediction. *Molecules* 28, (2023).

Protein-ligand Structure Prediction by Template-guided Ensemble Docking

Keqiong Zhang, Qilong Wu, and Sheng-You Huang*

School of Physics, Huazhong University of Science and Technology, Wuhan, China 430074

huangsy@hust.edu.cn

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:Y; EMA:Y; MD:N

We applied a template-guided ensemble docking strategy for LG tasks in CASP 16. AlphaFold3 and AlphaFold-Multimer were used for de novo structure modeling of proteins and nucleic acids. Modeller was used for homology modeling of proteins, resulting in a structure ensemble of the receptor in each target. If a reliable template complex exists, LSalign was used to generate aligned poses of small molecules. For targets where no suitable template can be found, the complex structure was generated directly by XDock or Autodock Vina. A knowledge-based scoring function, ITScore, was used to score ligand poses against the structure ensemble of the receptor. After scoring and sorting, combined with the manual selection, five optimal models were submitted.

Methods

For the tertiary structure prediction of the receptor in each target, we used AlphaFold¹ for monomer proteins and AlphaFold-Multimer² for the assembly proteins. For nucleic acid receptors, we used AlphaFold3³ for structure modeling. Considering the phosphate groups of nucleic acids make them negatively charged, and the metal ions have a certain contribution to the stability of the 3D structure of nucleic acids, we also added 2~6 magnesium or potassium ions together with the nucleic acids during the modeling process to improve the modeling accuracy. We selected the top-ranked models based on the confidence score produced by AlphaFold. Modeller⁴ was used for homology modeling of proteins. Proteins with sequence identity higher than 30% in the Protein Data Bank (PDB)⁵ were selected as template structures. A study of the receptor protein in T1214 shows that it shares structural similarity with the ferric citrate transporter FecA (sequence identity < 30%), including a compact positively charged substrate-binding site⁶. Thus, we also used structures of FecA as templates for modeling. For each target, the protein or nucleic acid structures modeled by the above methods were used to construct a 3D structural ensemble of the receptor.

In the complex prediction stage, we combined a template-guided strategy with an energy-based evaluation process to generate potential binding poses of each ligand. Firstly, up to 500 conformations of each ligand were generated by RDKit⁷. We filtered protein-ligand (rna-ligand) complexes in PDB according to three rules: the sequence identity is higher than 30% (60%), the similarity of the ligand is higher than 0.3 and the maximum common substructure (MCS) coverage score of the ligand is over 0.5. Similarity and MCS were calculated and detected by RDKit. The MCS score was calculated by dividing the sum of the number of atoms and bonds in the searched MCS by the total number of atoms and bonds of the target ligand. Once a qualified template was detected, a flexible alignment of the target ligand and the template ligand was done by LSalign⁸. A knowledge-based scoring function ITScore^{9,10} was used to score these aligned poses on the receptor structure ensemble, which resulted in our first set of predicted structures of the complex (Set1_align). These conformations were optimized by MDock (ligand-rigid

minimization) and XDock¹¹ (ligand-flexible minimization) on the receptor ensemble, resulting in the second set of candidate conformations (Set2_min). If no acceptable template complexes were found, MDock and XDock were used for rigid and flexible ensemble docking, respectively. Thus, the third set of potential structures of the target complex has been generated (Set3_dock). Finally, ITScore was used to evaluate each predicted structure in the three sets, and the top 20 complexes with the lowest score were selected from each set. After manual inspection, we submitted the best five models as our final results. If the binding affinity is requested, the lowest score among the five models has been used for affinity ranking. For covalently binding ligands in the L4000 target, another docking program HCovDock is used to predict binding poses¹². The top 100 conformations from HCovDock were rescored based on the pharmacophore similarity on acceptable template ligands using an in-house script. Then, the final models were manually selected based on docking and pharmacophore scores. For RNA targets, we used a similar process for complex structure prediction. The only difference is that MDock is not involved in the docking and optimization process of RNA targets, because MDock is a docking program developed only for protein-small molecule systems.

4. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.
5. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *BioRxiv*, 2022: 2021.10.04.463034.
6. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630(8016):493-500.
7. Webb B, Sali A. Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics*, 2016, 54(1): 5.6. 1-5.6. 37.
8. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28(1):235-242.
9. Grinter R, Lithgow T. The crystal structure of the TonB-dependent transporter YncD reveals a positively charged substrate-binding site. *Acta Crystallogr D Struct Biol*. 2020;76(Pt 5):484-495.
10. RDKit: Open-source cheminformatics. <https://www.rdkit.org>
11. Hu J, Liu Z, Yu DJ, Zhang Y. LS-align: an atom-level, flexible ligand structural alignment algorithm for high-throughput virtual screening. *Bioinformatics*. 2018;34(13):2209-2218.
12. Huang S-Y, Zou X. An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials. *J Comput Chem*. 2006;27(15):1866-1875.
13. Huang S-Y, Zou X. An iterative knowledge-based scoring function to predict protein-ligand interactions: II. Validation of the scoring function. *J Comput Chem*. 2006;27(15):1876-1882.
14. Wu Q, Huang S-Y. XDock: a general docking method for modeling protein-ligand and nucleic acid-ligand interactions. Submitted, 2024.
15. Wu Q, Huang S-Y. HCovDock: an efficient docking method for modeling covalent protein-ligand interactions. *Brief Bioinform*. 2023;24(1):bbac559.

Modeling Multimeric Structures through Enhanced Sampling of AlphaFold-Multimer and HDOCK

Hao Li, Yurui Li, and Sheng-You Huang

¹School of Physics, Huazhong University of Science and Technology, Wuhan, China 430074

huangsy@hust.edu.cn

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; EMA:Y; MD:Y

In this CASP experiment, we enhanced AlphaFold-Multimer's sampling by using different versions and inputting various MSAs. Additionally, for targets where a high ipTM model could not be achieved with num_recycle set to 3, we increased num_recycle to 20. For targets where AlphaFold-Multimer struggled to make predictions, we utilized docking algorithm HDOCK and the AlphaFold3 server to assist with complex structure prediction.

Methods

Inspired by the participants of CASP15¹⁻³, we tested the impact of various parameters on the quality of AlphaFold-Multimer models⁴. The results indicated that differences in versions and the input of multiple sequence alignments (MSAs) significantly affect the quality of the AlphaFold-Multimer models. Therefore, we enhanced the sampling by using different versions of AlphaFold-Multimer (v1, v2, v3) and inputting different MSAs. For the input MSAs, in addition to using the default MSAs of AlphaFold-Multimer (comprising block-diagonalized MSAs and paired MSAs), we also utilized MSAs with only paired sequences and those with only block-diagonalized MSAs. Furthermore, considering that many sequences in the input MSAs are highly similar, we employed MMseqs2⁵ to reduce redundancy in the MSAs for different chains based on sequence identity thresholds of 30%, 40%, 50%, 60%, 70%, 80%, and 90%. Subsequently, we paired the MSAs for different chains according to their sequence similarity to the target sequence, generating a total of seven distinct paired MSAs.

To save time, the default num_recycle for AlphaFold-Multimer is set to 3. However, if no model with an ipTM greater than 0.8 is generated with num_recycle set to 3, we increase num_recycle to 20. Due to the high structural redundancy among AlphaFold-Multimer-generated models, we applied a redundancy reduction process to these models, using an interface rmsd threshold of 2 Å. If fewer than five models remained after this process, we applied a more stringent threshold of 1 Å. If more than one model with an ipTM greater than 0.8 remained after redundancy reduction, we re-ranked them using our HITScorePP scoring function⁶. For targets where AlphaFold-Multimer could not generate high-ipTM models, we utilized HDOCK2 to obtain docking models⁶. For targets with too many amino acids to be processed by AlphaFold-Multimer, we used the AlphaFold3 server to generate structures⁷. We submitted five tasks

under different random seeds, yielding a total of 25 models, which were then ranked by considering both ipTM and HITScorePP scores.

For protein-nucleic acid complexes, we first predicted their structures using the AlphaFold3 server, submitting five separate tasks and selecting the highest ipTM model from each, yielding a total of five models. Next, the protein and nucleic acid components of the complexes were separated, and the five proteins and five nucleic acids were recombined in pairs. Using HDOCK^{8,9}, we generated 25 docking results. All docking results were ranked using the HITScorePR scoring function, which incorporates a desolvation term. Finally, both the AlphaFold3-generated models and the docking models were submitted together.

Availability

HDOCK is freely available for academic users at <http://hdock.phys.hust.edu.cn/>.

1. Wallner B. Improved multimer prediction using massive sampling with AlphaFold in CASP15. *Proteins*. 2023; 91(12):1734-1746.
2. Olechnovič K, Valančauskas L, Dapkūnas J, Venclovas Č. Prediction of protein assemblies by structure sampling followed by interface-focused scoring. *Proteins*. 2023; 91(12):1724-1733.
3. Zheng W, Wuyun Q, Freddolino PL, Zhang Y. Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15. *Proteins*. 2023; 91(12):1684-1703.
4. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, Ronneberger O. Protein complex prediction with AlphaFold-Multimer. *biorxiv*. 2021; 2021-10.
5. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol*. 2017; 35(11):1026-1028.
6. Li H, Yan Y, Zhao X, Huang S-Y. Inclusion of Desolvation Energy into Protein-Protein Docking through Atomic Contact Potentials. *J Chem Inf Model*. 2022; 62(3):740-750.
7. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung CC, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024; 630(8016):493-500.
8. Yan Y, Tao H, He J, Huang S-Y. The HDOCK server for integrated protein-protein docking. *Nat Protoc*. 2020; 15(5):1829-1852.
9. Li H, Huang E, Zhang Y, Huang S-Y, Xiao Y. HDOCK update for modeling protein-RNA/DNA complex structures. *Protein Sci*. 2022; 31(11):e4441.

Enhancing Protein Embeddings via Contrastive Learning

SeokJun On¹, Minseo Kim², Jeongyong Shim³ and Eun-Sol Kim^{1,2,3}

¹ – Department of Artificial Intelligence, Hanyang University, Korea, ² – Department of Computer Science, Hanyang University, Korea, ³ – Department of Artificial Intelligence Application, Hanyang University, Korea

onseokjun@hanyang.ac.kr

Key: Auto:N; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; AF:N; EMA:N; MD:N

We developed a novel MSA(Multiple Sequence Alignment) search pipeline that leverages protein embeddings to efficiently perform MSA searches without the need to examine each amino acid individually. Traditional MSA search techniques often rely on comparing individual amino acids, which can be computationally expensive and time-consuming.

Protein embeddings typically represent protein sequences in a fixed-dimension vector, focusing on patterns within the primary amino acid sequence. However, this approach does not consider MSA, which contains crucial coevolutionary information important for structure prediction. Due to this limitation, we propose enhancing protein embeddings through contrastive learning to improve them, allowing the embeddings to include the information from MSAs.

In this work, we introduced a method for enhancing protein embeddings through contrastive learning and described how this enables a simple similarity-based approach for MSA search.

Methods

We propose a novel method for finding MSAs based on the similarity between the representations of individual proteins, as an alternative to the traditional MSA construction approach based on sequence search. To achieve this, we introduce a method for learning sequence representations for proteins. Our approach is based on contrastive learning to enhance the similarity between sequences within an MSA. Figure 1 illustrates the overall structure of the proposed method.

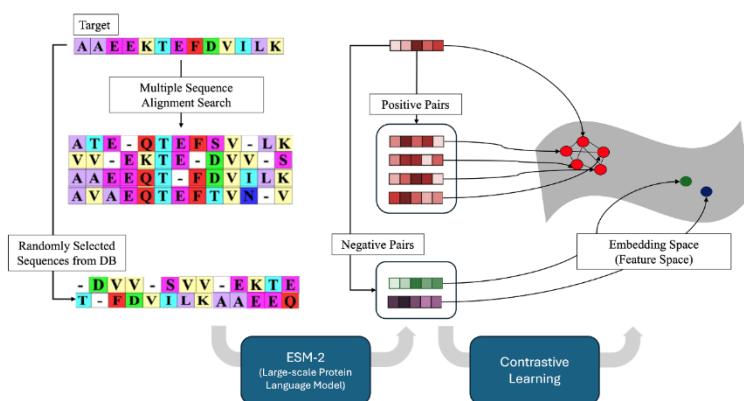


Figure 3. This is the overall structure for enhancing embeddings through contrastive learning. Through this approach, we can obtain protein embeddings that consider MSA, allowing us to perform MSA search based on the similarity within the embedding space

Enhance embedding & MSA search pipeline

For the baseline of protein embedding, we used a standard BERT model trained with a masked language model¹ to capture the patterns and features of each protein sequence. By passing a protein sequence through this model, we obtain amino acid-level embeddings, which can then be averaged to produce the overall representation for the whole protein.

Through this approach, the protein embedding was able to aggregate and encapsulate the features of individual proteins, which is composed of amino acids, and capturing unique characteristic information. In addition to this, we can obtain embeddings that consider the MSA in the embedding space of proteins by applying a few layers of MLP and contrastive loss. To define the loss, let $I = \{1, 2, \dots, B\}$ which represent the index of MSAs within the batch and use $J = \{1, 2, \dots, N_p\}$ to index the proteins belonging to each MSA. The data can be expressed as $\{x_{(m,n)} | m \in I, n \in J\}$ where $\{x_{(m,n)} | n \in J\}$ is a set of positive data with respect to each other. The loss function² for contrastive learning is defined L as follows, where $z_{(m,n)}$ is the baseline embedding vector of protein $x_{(m,n)}$.

$$L = - \sum_{i \in I} \sum_{j \in J} \log \left(\frac{\exp(\text{sim}(z_{(i,j)}, z_{(i,l)}))}{\sum_{k \in I} \sum_{l \in J} \exp(\text{sim}(z_{(i,j)}, z_{(k,l)}))} \right)$$

We used the OpenProteinSet³ dataset, which consists of approximately 270,000 MSAs generated by HHblits⁴, to perform contrastive learning. Using this trained network that produces enhanced embedding, we precomputed and stored the embeddings for the protein sequences in UniRef50⁵, which contains around 60 million sequences. Finally, for the query sequences provided in CASP 16, we obtained their representation and searched the precomputed database for the most similar sequences, using these sequences as the MSA.

Training (Positive and Negative)

To apply Contrastive Learning to protein representation learning, we first need to define positive and negative pairs. To incorporate MSA information, we define sequences belonging to the same MSA as positive pairs and sequences belonging to different MSAs within the same batch as negative pairs. In other words, we train the model so that the representation of proteins within the same MSA are represented as similar vectors, while the embeddings of proteins from different MSAs have low similarity.

Therefore, we constructed the batches as follows. Each of the B proteins has an associated MSA, and from each MSA, we randomly select $N_p - 1$ sequences, excluding the query sequence. In other words, within a single batch, N_p sequences from the same MSA are defined as positive pairs, and the remaining $N_p \times B - N_p$ sequences are considered as negative pairs.

Results

For CASP 16, we developed protein embeddings that incorporate MSA information through contrastive learning. By enhancing the embeddings in this way, we aimed to capture the evolutionary relationships between proteins more effectively. Using these enhanced embeddings, we proposed a novel MSA search method that leverages the improved representation to identify similar sequences.

Acknowledgements

This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) (No.II212068, Artificial Intelligence Innovation Hub, No.00220628, Artificial

intelligence for prediction of structure-based protein interaction reflecting physicochemical principles, No.II201373, Department of Graduate School Program, Hanyang University)

Availability

A github repository for contrastive learning and search pipeline is getting ready.

1. Zeming Lin et al., Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123–1130 (2023). DOI: 10.1126/science.adc2574
2. Khosla, Prannay, et al. "Supervised contrastive learning." *Advances in neural information processing systems* 33 (2020): 18661-18673.
3. Ahdritz, Gustaf, et al. "OpenProteinSet: Training data for structural biology at scale." *Advances in Neural Information Processing Systems* 36 (2024).
4. Remmert, M., Biegert, A., Hauser, A. et al. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods* 9, 173–175 (2012). <https://doi.org/10.1038/nmeth.1818>
5. Baris E. Suzek, Hongzhan Huang, Peter McGarvey, Raja Mazumder, Cathy H. Wu, UniRef: comprehensive and non-redundant UniProt reference clusters, *Bioinformatics*, Volume 23, Issue 10, May 2007, Pages 1282–1288, <https://doi.org/10.1093/bioinformatics/btm098>

Optimizing protein structure prediction using deep learning and Rosetta

Lei Wang^{1,#}, Jinyi Wang^{1,#}, Han Zhang^{1,#}, Xudong Li^{1,#}, Yan Wang² and Zhidong Xue¹

1 - School of Software Engineering, Huazhong University of Science and Technology

2 - School of Life Science and Technology, Huazhong University of Science and Technology

- equal contribution

Corresponding author: wanglei94@hust.edu.cn, yanw@hust.edu.cn, zdxue@hust.edu.cn

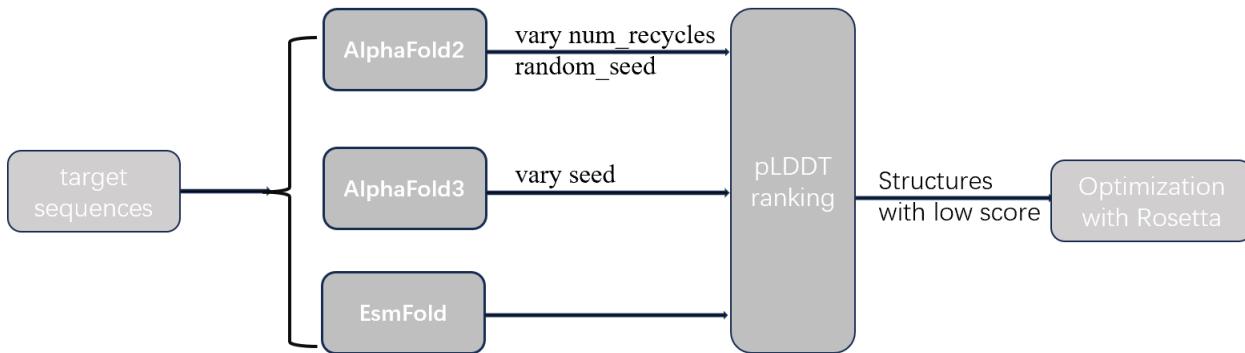
Key: Auto:Y; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

To predict the structures of proteins, we experimented with three models. For structures with lower pLDDT scores, we optimized them using Rosetta. Finally, we compared the TM-scores of structures with higher pLDDT scores and selected those with significant differences for submission.

Methods

To predict the structure of proteins, we employed three protein structure prediction models: AlphaFold2, AlphaFold3, and EsmFold. Each model was tested with specific parameters, especially AlphaFold2, we varied the num_recycles parameter using values of 6, 12, 24, 40, 70, and 80. The random seed parameter was tested with 100 randomly generated values to explore diverse conformations.

Then all the predicted structures were ranked based on their pLDDT scores. For structures with low pLDDT scores, we applied Rosetta to refine the predictions and enhance structural accuracy. We focused on structures with higher pLDDT values and calculated the TM-score for each. Structures with significant TM-score differences were selected for submission.



Results

The models produced multiple predicted structures for each protein, ranked by their pLDDT scores. The use of Rosetta improved the confidence and quality of structures with initially low pLDDT scores, leading to a final set of optimized structures for submission.

Finally, one EsmFold-generated structure, the top three structures generated by AlphaFold2, and one of the highest-scoring structures generated by AlphaFold3 were submitted.

Availability

1. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
2. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
3. Lin, Z. *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
4. GitHub - RosettaCommons/trRosetta2: Repository for publicly available deep learning models developed in Rosetta community

Molecular Docking of Small Molecule Ligands with Targets Using AutoDock Vina

Lei Wang^{1,#}, Han Zhang^{1,#}, Xudong Li^{1,#}, Jinyi Wang^{1,#}, Yan Wang² and Zhidong Xue¹

1 - School of Software Engineering, Huazhong University of Science and Technology

2 - School of Life Science and Technology, Huazhong University of Science and Technology

- equal contribution

Corresponding author: wanglei94@hust.edu.cn, yanw@hust.edu.cn, zdxue@hust.edu.cn

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; EMA:N; MD:Y

To explore the interactions between small molecule ligands and targets such as RNA or proteins, we performed computationally assisted molecular docking predictions. Docking analysis is a commonly used tool in drug design, which simulates the binding of small molecule ligands to their targets, predicting binding sites and interaction energies. We utilized RDKit to generate the structure of the small molecule, employed Open Babel for file format conversion, and finally carried out molecular docking using AutoDock Vina.

Methods

In predicting the docking of small molecule ligands with targets, we first utilized RDKit to convert the given SMILES string into the corresponding molecular object. Subsequently, Open Babel was used to convert the molecule's .mol file into a format compatible with AutoDock Vina (.pdbqt), while hydrogen atoms were added to the target's .pdb structure file and converted to the corresponding .pdbqt format. During the docking simulation, AutoDock Vina was employed to model the interactions between the small molecule ligand and the target. Vina utilizes a global search algorithm based on energy scoring to identify the optimal binding conformation. The docking parameters, including the coordinate range and grid center for the ligand and receptor, were specified in a configuration file. Finally, after the docking was completed, the pdbqt file was further converted into mol format for subsequent structural analysis and visualization.

Results

We successfully generated the structural files for small molecule ligands and completed molecular docking simulations with the target using AutoDock Vina. During the docking process, Vina provided a detailed simulation of the interactions between the small molecule ligand and the target, producing predicted binding conformations and their corresponding energy scores. These predictions not only reveal potential binding modes between the ligand and the target but also offer quantitative information regarding the binding strength.

1. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature protocols*, 11(5), 905-919.

Structure Prediction of RNA Monomers and Complexes Using Rosetta and AlphaFold3

Lei Wang^{1,#}, Xudong Li^{1,#}, Han Zhang^{1,#}, Jinyi Wang^{1,#}, Yan Wang² and Zhidong Xue¹

1 - School of Software Engineering, Huazhong University of Science and Technology

2 - School of Life Science and Technology, Huazhong University of Science and Technology

- equal contribution

Corresponding author: wanglei94@hust.edu.cn, yanw@hust.edu.cn, zdxue@hust.edu.cn

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; EMA:N; MD:Y

To predict the structures of RNA monomers and complexes, we employed two advanced computational methods: Rosetta and AlphaFold3. Each method has its own strengths: Rosetta excels in energy calculation and conformation sampling, while AlphaFold3 offers unique advantages in deep learning and large-scale data training. Ultimately, we selected structures with significant differences from the results for submission.

Methods

To achieve structural prediction of RNA monomers and complexes, we first analyzed the RNA sequences. We employed two advanced RNA secondary structure prediction tools, RFold and RNAFormer. These tools, based on different algorithms, accurately predicted the base pairing and secondary structure formation of the RNA sequences, generating high-quality RNA secondary structure models.

After obtaining the RNA secondary structures, we conducted a search for multiple sequence alignments (MSA). MSA can reveal co-evolutionary relationships and sequence conservation during the evolutionary process, providing crucial constraint information for structure prediction. By searching relevant databases (NR, RNACentral) and selecting data with similarity ranges of $30\pm5\%$ or $90\pm5\%$, we identified homologous sequences related to the target RNA sequence and constructed an MSA suitable for subsequent structural prediction.

Next, we used Rosetta to predict the RNA structures. Rosetta generated constraint conditions based on the previously obtained MSA and secondary structures, which were essential for ensuring the accuracy and plausibility of the three-dimensional structure. Utilizing these constraints, Rosetta carried out a comprehensive prediction of the RNA three-dimensional structures, ultimately producing several potential structure models.

Results

We utilized two advanced computational methods, Rosetta and AlphaFold3, to generate multiple three-dimensional RNA structures. First, we employed Rosetta to perform conformation sampling and energy optimization, producing a variety of potential RNA structures. Simultaneously, we conducted supplementary predictions using AlphaFold3, leveraging its deep learning algorithm to model the RNA sequence, further enhancing the prediction coverage and accuracy. After obtaining these structures, we performed a detailed comparative analysis of the results generated by the two methods, with a focus on identifying significant differences between them. Ultimately, we selected the structures with the most pronounced differences as the final results for submission.

1. Du, Z., Su, H., Wang, W., Ye, L., Wei, H., Peng, Z., ... & Yang, J. (2021). The trRosetta server for fast and accurate protein structure prediction. *Nature protocols*, 16(12), 5634-5651.

2. Su, H., Wang, W., Du, Z., Peng, Z., Gao, S. H., Cheng, M. M., & Yang, J. (2021). Improved protein structure prediction using a new multi-scale network and homologous templates. *Advanced Science*, 8(24), 2102592.
3. Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., & Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences*, 117(3), 1496-1503.
4. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3.

Integrating Advanced Computational Techniques for Biomolecular Structure Prediction

Yuki Kagaya¹, Tsukasa Nakamura¹, Jacob Verburgt¹, Anika Jain¹, Genki Terashi¹, Pranav Punuru¹, Emilia Tugolukova¹, Joon Hong Park², Anouka Saha³, David Huang¹, and Daisuke Kihara^{1,2}

¹ -Department of Biological Sciences, Purdue University, West Lafayette, IN, USA,

² - Department of Computer Science, Purdue University, West Lafayette, IN, USA,

³ -Department of Mathematics, The University of Texas at Austin, Austin, TX, USA

dkihara@purdue.edu

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3/2; EMA:Y; MD:N

Our group participated in the TS (Tertiary Structure), H (Heteromeric assembly), M (Protein-Nucleic Acids complex), and LG (Ligand) prediction categories for both protein and nucleic acid targets.

Methods

Protein structure modeling: For protein-only targets, we used AlphaFold2¹ as the core of the pipeline. Multiple sequence alignments (MSAs) were generated using the standard AlphaFold2 pipeline, utilizing UniRef30 (2023_03), BFD (latest)², UniRef90 (2024_02)³, UniProt (2024_02)⁴, and MGnify NR (2023_02)⁵ as the sequence databases. For the template database, we used the Protein Data Bank (PDB)⁶ as of April 17, 2024. Additionally, we integrated an enhanced in-house MSA, incorporating metagenomic and viral sequence data from Logan contigs⁷, and metagenome assemblies from NCBI⁸. DNA sequences were translated into protein sequences using prodigal⁹ as preprocess for those databases, and JackHMMER¹⁰ was employed for sequence searching and MSA construction.

Both the default and enhanced MSAs were used for AlphaFold2 inference. For multi-chain complex targets, we employed three published versions of AlphaFold2 model weights, each with their default settings. For monomeric protein targets, the AlphaFold2 monomer model was used with standard configurations. After May 9, 2024, the AlphaFold3¹¹ web server was also integrated, with the generated models added to our structure pool.

Various ad hoc strategies were applied to optimize the results for each target. If the MSA was sparse, we reinforced it by combining structure-based MSAs from Foldseek¹², or incorporating mutation databases to construct MSAs manually. If the inference process was computationally impossible due to the size of the target, or if a reasonable model could not be generated by normal inference, we reduced stoichiometry or divided the target into more manageable domains or chunks. The resulting models were combined using Modeller¹³ or PyMOL to ensure it matches to target sequence.

Finally, all the available models were pooled, and ranked using the VoroIF-jury (Voronoi-based InterFace jury) scoring pipeline¹⁴. We supplemented this pipeline with GOAP¹⁵, DFIRE¹⁶, and ITSscorePro¹⁷, which were components of the ranksum score used in our LZerD method^{18,19}, incorporating these as external scores. After the initial rank was generated, each target was meticulously reviewed against relevant literature, and the final models for submission were selected by human experts through thorough visual inspection.

RNA structure modeling: For monomeric RNA targets, we used our RNA structure prediction method, NuFold²⁰. Additionally, we employed DeepFoldRNA²¹, DRfold²², RosettaFold2NA²³, RhoFold²⁴, and trRosettaRNA²⁵. The baseline MSA input was generated using the rMSA pipeline²⁶, supplemented by three variants of metagenome MSAs used in NuFold. After AlphaFold3 became available, AlphaFold3 models were also generated and included as part of the structure pool. These structures were ranked by combining Rosetta score²⁷ and ARES²⁸ score. As the same as protein targets, the final submission models were selected through visual inspection and comparison with literature information.

RNA complex and protein-NA complex modeling: For multi-chain RNA targets and protein-nucleic acid complex targets, RosettaFold2NA and AlphaFold3 were used to generate structures. For very large complexes, various manual modeling techniques were used as needed. These included breaking down the complex into smaller sets of chains, predicting only part of a chain, or applying symmetry operations to monomer structures. Model selection was based on both the scores generated by each method and manual review by human experts.

Ligand docking prediction: For small molecule ligands, GLIDE XP²⁹ and Induced Fit Docking (IFD)³⁰ tools from the Schrödinger software suite were used for ligand docking. For docked ligands, pocket locations were identified via SiteMap³¹. For some targets, the pocket location and docking pose were estimated using homologue information from the PDB as a template or by inputting ligand analogues into AlphaFold3. To refine and diversify submitted structures, some models were subjected to MD relaxation via Desmond³² using the OPLS4 forcefield³².

Automatic server: Our kiharalab_server participated in protein targets (T/H), complex targets (M), and RNA targets (R), but not ligands (L). The server group used the same components as our human group, but used an automated pipeline without any human intervention to generate and submit predictions. The server group used minimum stoichiometry for phase 0 targets.

Acknowledgements

We are grateful for ITaP Research Computing at Purdue University for providing us computational resources for this project. This work is partially supported by the National Institutes of Health (R01GM133840), the National Science Foundation (IIS2211598, DMS2151678, DBI2003635, and DBI2146026). J.V. were supported by a National Institute of General Medical Sciences-funded predoctoral fellowship to J.V. (T32 GM132024).

1. Jumper,J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).
2. Steinegger,M., Mirdita,M. & Söding,J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods* 2019 16:7 **16**, 603–606 (2019).
3. Suzek,B.E., Wang,Y., Huang,H., McGarvey,P.B. & Wu,C.H. UniRef clusters: A comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 31, 926–932 (2015).
4. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49, D480–D489 (2021).
5. Mitchell,A.L. et al. MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res* 48, D570–D578 (2020).

6. Burley,S.K. et al. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res* 51, D488–D508 (2023).
7. Chikhi,R., Raffestin,B., Korobeynikov,A., Edgar,R. & Babaian,A. Logan: Planetary-Scale Genome Assembly Surveys Life's Diversity. *bioRxiv* 2024.07.30.605881 (2024) doi:10.1101/2024.07.30.605881.
8. Kitts,P.A. et al. Assembly: a resource for assembled genomes at NCBI. *Nucleic Acids Res* 44, D73–D80 (2016).
9. Hyatt,D. et al. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 1–11 (2010).
10. Johnson,L.S., Eddy,S.R. & Portugaly,E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11, 1–8 (2010).
11. Abramson,J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024 630:8016 630, 493–500 (2024).
12. van Kempen,M. et al. Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* 2023 42:2 42, 243–246 (2023).
13. Webb,B. & Sali,A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* 54, 5.6.1-5.6.37 (2016).
14. Olechnovič,K., Valančauskas,L., Dapkūnas,J. & Venclovas,Č. Prediction of protein assemblies by structure sampling followed by interface-focused scoring. *Proteins: Structure, Function, and Bioinformatics* 91, 1724–1733 (2023).
15. Zhou,H. & Skolnick,J. GOAP: A Generalized Orientation-Dependent, All-Atom Statistical Potential for Protein Structure Prediction. *Biophys J* 101, 2043 (2011).
16. Zhou,H. & Zhou,Y. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Science* 11, 2714–2726 (2002).
17. Huang,S.Y. & Zou,X. Statistical mechanics-based method to extract atomic distance-dependent potentials from protein structures. *Proteins: Structure, Function, and Bioinformatics* 79, 2648–2661 (2011).
18. Christoffer,C. et al. LZerD webserver for pairwise and multiple protein-protein docking. *Nucleic Acids Res* 49, W359–W365 (2021).
19. Venkatraman,V., Yang,Y.D., Sael,L. & Kihara,D. Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinformatics* 10, (2009).
20. Kagaya,Y. et al. NuFold: A Novel Tertiary RNA Structure Prediction Method Using Deep Learning with Flexible Nucleobase Center Representation. *bioRxiv* 2023.09.20.558715 (2023) doi:10.1101/2023.09.20.558715.
21. Pearce,R., Omenn,G.S. & Zhang,Y. De Novo RNA Tertiary Structure Prediction at Atomic Resolution Using Geometric Potentials from Deep Learning. *bioRxiv* 2022.05.15.491755 (2022) doi:10.1101/2022.05.15.491755.
22. Li,Y. et al. Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nature Communications* 2023 14:1 14, 1–13 (2023).
23. Baek,M. et al. Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature Methods* 2023 21:1 21, 117–121 (2023).

24. Shen,T. et al. E2Efold-3D: End-to-End Deep Learning Method for accurate de novo RNA 3D Structure Prediction. ArXiv (2022).
25. Wang,W. et al. trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nature Communications* 2023 14:1 14, 1–13 (2023).
26. Zhang,C., Zhang,Y. & Pyle,A. M. rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *J Mol Biol* 435, 167904 (2023).
27. Alford,R.F. et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J Chem Theory Comput* 13, 3031–3048 (2017).
28. Townshend,R.J.L. et al. Geometric deep learning of RNA structure. *Science* (1979) 373, 1047–1051 (2021).
29. Friesner,R.A. et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* 49, 6177–6196 (2006).
30. Sherman,W., Day,T., Jacobson,M.P., Friesner,R.A. & Farid,R. Novel procedure for modeling ligand/receptor induced fit effects. *J Med Chem* 49, 534–553 (2006).
31. Halgren,T.A. Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model* 49, 377–389 (2009).
32. Lu,C. et al. OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J Chem Theory Comput* 17, 4291–4300 (2021).

Koes

Predicting Ligand Binding Poses and Affinities for Four Target-Ligand Sets Using Open-ComBind, GNINA, Dynamicbind and GNINA Rescoring of DynamicBind Predictions

S. Pirhadi, A.T. McNutt and D.R. Koes*

Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA, USA

dkoes@pitt.edu

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:N; Fragm:N; Cont:N; Dist:N; Tors:N; DeepL:Y; EMA:N; MD:N

We investigated ligand target affinity and pose predictions of four distinct target-ligand sets, including L1000: Chymase, L2000: Cathepsin_G, L3000:Autotaxin, and L4000:Mpro using model1:Open-ComBind, model2:GNINA, model3: re-scoring of Dynamicbind predicted ligand poses by GNINA and model4: Dynamicbind.

GNINA¹ is a deep learning docking method that utilizes an ensemble of convolutional neural networks (CNNs) as a scoring function on top of conventional Monte Carlo search.

Open-ComBind² is an open-source pose prediction pipeline leveraging the understanding that different ligands will bind to a receptor in similar ways. GNINA is used to generate a set of highly likely poses for a set of ligands binding to the same receptor. Then, interaction fingerprints are calculated for the set of poses. The Tanimoto similarity of interaction fingerprints and the RMSD of the maximum common substructure (MCSS) is calculated between all pairs of poses. Poses are selected for all ligands such that they maximize the Open-ComBind objective which simultaneously ensures high similarity between poses and biases the predicted poses towards higher GNINA scores.

DynamicBind³ is an advanced geometric deep generative model tailored for "dynamic docking." Unlike conventional docking methods that often consider proteins as rigid structures, DynamicBind dynamically refines the protein conformation, transitioning it from its initial AlphaFold-predicted state to a more realistic, holo-like state, enabling more accurate interaction modeling.

Methods

For all models, we utilized a single, holo receptor structure; where possible an experimentally determined structure was used. We utilized the experimentally determined structures 1T32, 1KLT, and 7D3I for Cathepsin_G, Chymase, and MPro, respectively. Since Autotaxin lacked an experimentally determined holo structure, we ran ColabFold to create an apo structure then ran DynamicBind with a single ligand to morph the apo structure into a holo structure and determine the pocket.

Model1: Open-ComBind: we generated a set of 10 conformations for each ligand SMILES, using RDKit to generate a set of 50 conformations and then minimizing each conformation with UFF to get the 10 lowest energy conformations. We then docked each conformation against the static receptor structure,

using the cognate ligand to determine the pocket location in the case of experimentally determined structures. The docking predictions of each conformation were pooled and sorted by the CNNscore. Ligands for each target set were clustered on the Hamming distance of their Morgan fingerprint using average-linkage agglomerative clustering with a maximum cluster size of 15. Any clusters with less than 5 members were combined with their nearest cluster until no clusters with less than 5 members remained. We then calculated the hydrophobic contacts of the top 100 predicted poses. The Tanimoto similarity of the hydrophobic contacts and the RMSD of the MCSS was calculated between all pairs of poses in the cluster. The poses were selected by maximizing the Open-ComBind objective for each cluster of ligands. Affinity values were determined by converting the reported CNNaffinity value of the selected ligand conformation from pK to K_d . In the case of multiple ligands in the binding site, we followed an autoregressive docking procedure. In complexes with 2 identical ligands in the binding pocket, we followed the Open-ComBind procedure above and then docked the second ligand (with knowledge of the placement of the first ligand) using Open-ComBind on the set. For complexes containing other ligands in the binding site, the additional ligands were placed with knowledge of all larger ligands using the top ranked GNINA pose.

Model2: GNINA: we generated a set of 10 conformations for each ligand SMILES, using RDKit to generate a set of 50 conformations and then minimizing each conformation with UFF to get the 10 lowest energy conformations. We then docked each conformation against the static receptor structure, using the cognate ligand to determine the pocket location in the case of experimentally determined structures. The docking predictions of each conformation were pooled and sorted by the CNNscore. For complexes containing multiple ligands in the binding site, we performed an auto-regressive docking procedure in which the ligands were docked in order of their size. Affinity values were determined by converting the reported CNNaffinity value of the top ranked ligand conformation from pK to K_d .

Model3: DynamicBind + GNINA: The protein structures (in PDB format) and ligand structures (in SMI format) were input into the Dynamicbind. For complexes containing multiple ligands in the binding site, predictions were performed for only the first ligand. The ligand-target complexes generated by DynamicBind were processed using the GNINA. Utilizing GNINA's --minimize option, all ligand poses generated by DynamicBind were energy-minimized. The top-scoring conformations, sorted by pose_sort_order based on the CNNscore, were reported as the predicted binding poses and the conversion of their CNNaffinity values to molar concentration values were used as affinities for the intended ligand-target pairs.

Model4: Dynamicbind: The top-ranked complex predicted by DynamicBind was selected to determine the binding pose, and the predicted affinity of the best pose was converted to molar concentration, which was then reported as the predicted affinity.

Results

The success rates of different approaches in predicting ligand poses with an RMSD below 2 Å across various ligand-target pairs were evaluated. The model1: Open-ComBind showed low success rates, achieving 5.9% for Chymase, 100% for Cathepsin_G, 2.1% for Autotaxin, and 4% for MPro. The model2:

GNINA docking yielded slightly varied results, with success rates of 5.9% for Chymase, 100% for Cathepsin_G, 2.6% for Autotaxin, and 8% for MPro ligands.

The model3: DynamicBind + GNINA approach demonstrated improved performance, particularly for some targets. For Chymase, the success rate was 47.1%, while for Cathepsin G, it was 50%. The success rate for Autotaxin was 6.3%, and for MPro, it was 8%. The model4: DynamicBind achieved the best success rate of 64.7%, and 14.8% among all four models for Chymase and Autotaxin, respectively. For Cathepsin_G, it successfully predicted poses for both ligands but it dropped to 0% for MPro.

1. McNutt,A.T., Francoeur,P., Aggarwal,R. et al. (2021). GNINA 1.0: molecular docking with deep learning. *J. Cheminform.* **13**, 43.
2. McNutt,A.T., Koes, D.R. (2024). Open-ComBind: harnessing unlabeled data for improved binding pose prediction. *J. Comput. Aided. Mol. Des.* **38**, 3.
3. Lu,W., Zhang,J., Huang,W., Zhang,Z., Jia,X., Wang,Z., Shi,L., Li,C., Wolynes,P.G., Zheng,S., (2024). DynamicBind: predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model. *Nat. Commun.* **15**, 1071.

Template based Protein-Ligand Complex Structure Prediction in CASP16

Won-Kyeong Jang, Jin-Hyeok Yoo, Woong-Hee Shin*

Department of Biomedical Informatics, Korea University School of Medicine

*Corresponding author: whshin@korea.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; Cont:Y; Dist:Y; Tors:Y; DeepL:Y; EMA:Y; MD:Y

In CASP16, we proposed a new protein-ligand complex structure prediction protocol. This method uses a modified version of AlphaFold2 (AF2)¹ to predict the optimal structure of protein-ligand complexes by searching experimental data with more than 70% identical amino acid sequences in the Protein Data Bank (PDB)². From the structure list, we use the protein structure as a template if the ligand is similar to the given SMILES. After the receptor structure prediction, a protein-ligand docking is performed, and then rescoring the docked poses was followed.

Methods

Template Selection

We used PSI-Blast³ to search for similar protein sequences among the experimentally determined structures in the PDB. The sequence similar proteins were further filtered by the compound similarity, between the co-crystallized ligand and the query, using SHAFTS⁴. The selected protein structure was provided as a template for AF2, instead of giving multiple sequence alignment⁵.

Protein Structure Prediction

We used AF2 to generate the protein structure based on the template. Additionally, we used OpenMM⁶ to relax the generated structure. Both the original AF2-generated structure and the relaxed structure were used in the subsequent steps.

Protein-Ligand Docking

We performed protein-ligand docking using the generated protein structure in the previous step and the ligand's three-dimensional coordinates obtained from the template if available. To determine the binding site, we used TM-align⁷ to align the template and the model. After the alignment, the binding ligands from template were copied to the modeled structure, and the binding site of the model was determined from the copied molecules. If the ligands lie within 5 Å, their binding sites were classified as identical pockets. The average coordinates of these ligands was used to determine the binding site. To the determined binding pockets, Autodock-GPU⁸ was employed to dock the query molecules.

Complex Scoring and Refinement

AutoDock-GPU generated 1000 conformations per ligand. The poses were rescored using AK-Score2⁹. We selected the top 5 complexes based on AK-Score2 and minimized them using UCSF Chimera¹⁰ with Amber ff14SB¹¹ to obtain the final structures.

1. Jumper, J., et al. (2021) "Highly accurate protein structure prediction with AlphaFold" *Nature* 596: 583-589.
2. Bank, P. D. (1971). "Protein data bank." *Nature New Biol* 233(223): 10-1038.
3. Altschul, S. F., et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic acids research* 25(17): 3389-3402.
4. Liu, X., et al. (2011). "SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening." *J Chem Info Model* 51(9): 2372-2385.
5. Song, J., et al. (2024) "Applying multi-state modeling using AlphaFold2 for kinases and its application for ensemble screening" *bioRxiv*, doi:10.1101/2024.04.04.588044.
6. Eastman, P., et al. (2017). "OpenMM 7: Rapid development of high performance algorithms for molecular dynamics." *PLoS computational biology* 13(7): e1005659.
7. Zhang, Y. and J. Skolnick (2005). "TM-align: a protein structure alignment algorithm based on the TM-score." *Nucleic acids research* 33(7): 2302-2309.
8. Santos-Martins, D., et al. (2021). "Accelerating AutoDock4 with GPUs and gradient-based local search." *Journal of chemical theory and computation* 17(2): 1060-1073.
9. Kwon, Y., et al. (2020). "AK-score: accurate protein-ligand binding affinity prediction using an ensemble of 3D-convolutional neural networks." *International journal of molecular sciences* 21(22): 8424.
10. Pettersen, E. F., et al. (2004). "UCSF Chimera—a visualization system for exploratory research and analysis." *Journal of computational chemistry* 25(13): 1605-1612.
11. Maier, J. A., et al. (2015). "ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB." *Journal of chemical theory and computation* 11(8): 3696-3713.

Comprehensive RNA and RNA-Complex Structural Modeling for CASP16

C. Nithin, S.P. Pilla and S. Kmiecik

¹ - Laboratory of Computational Biology, Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Zwirki i Wigury 101, 02-089 Warsaw, Poland.

sekmi@chem.uw.edu.pl

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; EMA:Y; MD:Y

RNA molecules play crucial roles in various cellular processes, often interacting with proteins, small molecules, and other RNAs to form complex functional units. Accurate modeling of RNA and its complexes is essential for understanding RNA structure-function relationships. The CASP16 experiment provided an opportunity to test and refine methods for RNA structure prediction and RNA-protein/ligand complex modeling on a set of blinded targets. In this study, we applied a comprehensive workflow that integrates multiple computational tools to predict RNA secondary and tertiary structures and RNA complexes based on the target sequences provided by the CASP organizers.

Methods

As a first step, databases such as Rfam¹ and RNACentral² were utilized to search for sequences or structures similar to the target RNAs. Multiple sequence alignments and family information were retrieved where applicable, enhancing the context for structural predictions. We employed a suite of tools—ViennaRNA³, RNAStructure⁴, CentroidFold⁵, ContraFold⁶, and IPknot⁷—to predict RNA secondary structure, leveraging distinct algorithms and statistical models. Conservation analysis, a crucial step in RNA modeling, was performed using alignments generated by rMSA2⁸, with conservation scoring from RNAalifold, alifoldz, and RNAz. This process allowed us to identify conserved and functionally critical regions, which helped refine the secondary structure and provide valuable restraints for 3D modeling.

For 3D structure predictions, we utilized various methods, including SimRNA⁹, DeepFoldRNA¹⁰, FARFAR2¹¹, RhoFold¹², Vfold2¹³, trRosettaRNA¹⁴, NuFold¹⁵, and AlphaFold 3¹⁶ (AF3). Both restrained and unrestrained simulations were conducted to explore the conformational landscape of the RNA molecules. Models were evaluated for convergence to ensure structural fidelity and high-resolution refinements were applied to top-scoring models to reduce potential errors introduced during coarse-grain modeling.

For long RNA sequences or RNA targets with complex stoichiometry, we explored the performance of AF3, alongside other methods. Protein-DNA complexes were modeled using AF3 with different stoichiometries, followed by manual post-processing to ensure correct interactions. Additionally, small molecules (e.g., ATP, NADPH, and metal ions) were included in the AF3 simulations to explore possible conformations. For the RNA target with the local solvent shell (water and ions), the structures were modeled based on homology, and molecular dynamics simulations were performed with the amber force field.

For RNA-protein complexes, the interfaces were aligned with experimentally derived structures when available, ensuring biologically relevant interaction sites. For RNA-ligand complexes, we used

RLDOCK¹⁷ and AutoDock Vina¹⁸ programs to predict potential docking poses by identifying conserved pockets based on structural homology and secondary structure conservation. The top-ranked models from these predictions were selected for further optimization. Ligand positions within the complexes were refined through energy minimization, ensuring accurate binding interactions and enhancing the overall accuracy of the RNA-ligand complex models.

Results

We participated in the TS and LG categories of CASP16 as group LCBio, with group number 189. We submitted a total of 292 models for TS and LG targets. For the ribozyme target, an ensemble of 1,000 conformations of the RNA, including the solvent shell, was submitted.

Availability

The various methods used in this study are publicly available.

Acknowledgments

We acknowledge funding from the National Science Centre, Poland [OPUS 2020/39/B/NZ2/01301]; and Polish high-performance computing infrastructure PLGrid (HPC Center: ACK Cyfronet AGH) for providing computer facilities and support within computational grant no. PLG/2024/016931.

1. Kalvari, I., Nawrocki, E.P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., Griffiths-Jones, S., Toffano-Nioche, C., Gautheret, D., Weinberg, Z., Rivas, E., Eddy, S.R., Finn, R.D., Bateman, A. & Petrov, A.I. (2021). Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res.* 49, D192–D200
2. RNACentral Consortium (2021). RNACentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.* 49, D212–D220
3. Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. & Hofacker, I.L. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6, 26
4. Reuter, J.S. & Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 11, 129
5. Sato, K., Hamada, M., Asai, K. & Mituyama, T. (2009). CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic Acids Res.* 37, W277–80
6. Do, C.B., Woods, D.A. & Batzoglou, S. (2006). CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics* 22, e90–8
7. Sato, K., Kato, Y., Hamada, M., Akutsu, T. & Asai, K. (2011). IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics* 27, i85–93
8. Zhang, C., Zhang, Y. & Pyle, A.M. (2023). rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *J. Mol. Biol.* 435, 167904
9. Wirecki, T.K., Nithin, C., Mukherjee, S., Bujnicki, J.M. & Boniecki, M.J. (2020). Modeling of Three-Dimensional RNA Structures Using SimRNA. *Methods Mol. Biol.* 2165, 103–125
10. Pearce, R., Omenn, G.S. & Zhang, Y. (2022). DE NovoRNA tertiary structure prediction at atomic resolution using geometric potentials from deep learning. *bioRxiv* doi:10.1101/2022.05.15.491755
11. Watkins, A.M. & Das, R. (2023). RNA 3D Modeling with FARFAR2, Online. *Methods Mol. Biol.* 2586, 233–249

12. Shen, T., Hu, Z., Peng, Z., Chen, J., Xiong, P., Hong, L., Zheng, L., Wang, Y., King, I., Wang, S., Sun, S. & Li, Y. (2022). E2Efold-3D: End-to-End Deep Learning Method for accurate de novo RNA 3D Structure Prediction. doi:10.48550/ARXIV.2207.01586
13. Li, J., Zhang, S., Zhang, D. & Chen, S.-J. (2022). Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. *Bioinformatics* 38, 4042–4043
14. Wang, W., Feng, C., Han, R., Wang, Z., Ye, L., Du, Z., Wei, H., Zhang, F., Peng, Z. & Yang, J. (2023). trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat. Commun.* 14, 7266
15. Kagaya, Y., Zhang, Z., Ibtehaz, N., Wang, X., Nakamura, T., Huang, D. & Kihara, D. (2023). NuFold: A Novel Tertiary RNA Structure Prediction Method Using Deep Learning with Flexible Nucleobase Center Representation. *bioRxiv* doi:10.1101/2023.09.20.558715
16. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D. & Jumper, J.M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500
17. Jiang, Y. & Chen, S.-J. (2022). RLDOCK method for predicting RNA-small molecule binding modes. *Methods* 197, 97–105
18. Eberhardt, J., Santos-Martins, D., Tillack, A.F. & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* 61, 3891–3898

Ligand binding pose prediction via the combination of geometric deep-learning and physics-based scoring models

Dongwoo Kim^{1*}, Jaemin Sim^{2*}, Soowon Lee^{1*}, Sangmin Lee^{2*}, Kyunghwan Yeo^{2*}, Juyong Lee^{1,2†}

¹ - College of Pharmacy, Seoul National University, ² - Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, and College of Medicine, Seoul National University

*Equal contribution

†Corresponding author: Juyong Lee (nicole23@snu.ac.kr)

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF:AF3; EMA:Y; MD:N

For a CASP16 ligand target, we employed a dual approach starting with a template search. If templates were found, we performed the template-based docking protocol. When templates are not detected, we docked ligands into predicted ligand binding pockets using the ensemble of docking tools. All docked poses were subsequently re-evaluated using AI-based protein-ligand scoring methods. For nucleic acid complexes, we used AlphaFold3 (AF3) to generate initial complex structures. Using the AF3 results as initial templates, we docked a query ligand to the ligand docked site of AF3 results using Glide.

Template-based docking: For a given target-ligand pair, we searched for a template complex structure in the PDB, considering both target and ligand similarities. A template complex structure was identified if a protein has sequence identity higher than 70% or E-value is smaller than 10⁻³ to a target sequence. Ligand similarity was measured by the Tanimoto similarity between the query ligand and the template ligand using ECFP41 1024-bit fingerprints. When the similarity exceeded 0.3, template-based docking was performed using Glide's restrained docking and pharmacophore docking functionalities².

For covalently bonded ligands, Glide's CovDock was used to generate covalently docked poses.

Ensemble docking and AI-based re-scoring method: When no suitable template complex structure was found, we generated a large number of protein-ligand binding poses (> 10,000) using the ensemble of various docking tools such as AutoDock-GPU3, AutoDock-Vina4, PLANTS5, and LeDock6. To identify native-like bound poses, we combined four AI-based scoring functions—AK-score2⁷, RTMScore8, and two in-house protein-ligand RMSD prediction models—to evaluate and select the optimal binding pose for final submission. AK-score2 is a graph neural network (GNN) based scoring function designed to predict protein-ligand binding affinity and bound pose RMSD. RTMScore is a deep-learning based ligand pose scoring function using a mixture density network.

Furthermore, two in-house protein-ligand RMSD prediction models were integrated: one is a GNN that directly predicts the binding RMSD, while the other is a binary classification model that predicts whether the RMSD is within 2 Å. After combining these models into a final scoring function, we selected the binding pose with the best score. To ensure pose diversity, we selected additional poses that were more than 2 Å apart from the best pose. Once additional poses were selected, we ensured further diversity by

picking ligands that were more than 2 Å apart from all previously selected poses. For tasks requiring binding affinity prediction, we submitted predictions using a GNN model designed to directly predict protein-ligand binding affinity based on the 3D structure of the selected poses.

Nucleotide-specific docking: Ligand docking to nucleic acids was approached by first generating protein structures using AlphaFold3, which produced 500 models per protein. We targeted proteins with the most similar ligands among ligands supported by the AF3 server to a query ligand, resulting in initial guesses for the docked poses of CASP16 query ligands. Based on the AF3 predicted nucleic acids structures, query ligands were docked using Glide. Both induced-fit and rigid dockings were performed to consider the flexibility of nucleic acid. Following initial docking, all poses underwent redocking with Glide to refine their accuracy. Submission models were then selected based on Glide docking scores, ensuring that only the most reliable poses were advanced for final evaluation.

Availability

Binding RMSD prediction model parameters available at: <https://github.com/eightmm/BindingRMSD>.

Binding affinity prediction parameters available at: <https://github.com/eightmm/BAPred>.

1. Burley SK, Bhikadiya C, Bi C, et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* 2021; **49**: D437–D451.
2. Rogers D, Hahn M. Extended-Connectivity Fingerprints. *J Chem Inf Model* 2010; **50**: 742–754.
3. Friesner RA, Banks JL, Murphy RB, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* 2004; **47**: 1739–1749
4. Santos-Martins D, Solis-Vasquez L, Tillack AF, et al. Accelerating AutoDock4 with GPUs and Gradient-Based Local Search. *J Chem Theory Comput* 2021; **17**: 1060–1073.
5. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* 2010; **31**: 455–461.
6. Korb O, Stützle T, Exner TE. Empirical Scoring Functions for Advanced Protein–Ligand Docking with PLANTS. *J Chem Inf Model* 2009; **49**: 84–96.
7. Clemente CM, Prieto JM, Martí M. Unlocking Precision Docking for Metalloproteins. *J Chem Inf Model* 2024; **64**: 1581–1592.
8. Kempen M van, Kim SS, Tumescheit C, et al. Fast and accurate protein structure search with Foldseek. 2023; 2022.02.07.479398.
9. Hong Y, Ha J, Lim CJ, et al. Accurate Prediction of Protein-Ligand Interactions by Combining Physical Energy Functions and Graph-Neural Networks. Epub ahead of print January 25, 2024. DOI: 10.21203/rs.3.rs-3887850/v1.
10. Chen C, Zhang X, Deng Y, et al. Boosting Protein–Ligand Binding Pose Prediction and Virtual Screening Based on Residue–Atom Distance Likelihood Potential and Graph Transformer. *J Med Chem* 2022; **65**: 10691–10706.

CALVADOS Predictions of the Conformational Ensemble of the Two-Domain Protein ZLBT-C

Fan Cao¹, F. Emil Thomassen¹, Hamidreza Ghafouri², Giulio Tesei¹, Silvio C.E. Tosatto²
and Kresten Lindorff-Larsen¹

¹ - Structural Biology and NMR Laboratory & the Linderstrøm-Lang Centre for Protein Science, Department of Biology,
University of Copenhagen, Copenhagen, Denmark,

² - Department of Biomedical Sciences, University of Padova, Padova, Italy

lindorff@bio.ku.dk

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:Y; AF:AF2; EMA:N; MD:Y

We used coarse-grained molecular dynamics simulations to predict the conformational ensemble of ZLBT-C and capture the interdomain orientations between the ZLBT and C domains in challenges T1200 and T1300. Using the cg2all machine-learning model, we then reconstructed the all-atom structures from the coarse-grained simulation trajectories. The following section outlines the approach in detail.

Methods

We performed coarse-grained molecular dynamics simulations of the wild-type and mutant protein targets using CALVADOS 3¹. We used AlphaFold2 predictions² as the initial structures for simulations. Simulations were performed in openMM³ as described previously¹. We assigned a partial charge of -0.25 to the charged beads of D54, D58, E62 and E65 (PDB CODE: 2lr2) to compensate for the absence of an explicit Lanthanide ion with a +3 charge. We back-mapped our coarse-grained simulations to an all-atom representation using cg2all⁴.

To ensure that the backbone RMSD of the two folded regions fall within 0.5 Å of the experimental structures, we used the following approach:

1. Superpose the two experimental structures to each frame of the all-atom simulation trajectories.
2. Delete the folded domains from our original simulation trajectory.
3. Combine the remaining part of the simulation trajectory with the superposed experimental structures.
4. Fix the local geometry by relaxing each trajectory structure with position restraints on the folded domains using Charmm36 in GROMACS⁵.

Availability

Code to run CALVADOS3 simulations is available open-source at <https://github.com/KULL-Centre/CALVADOS>. openMM and cg2all are available open-source at <https://github.com/openmm/openmm> and <https://github.com/huhlim/cg2all>, respectively. References for the tools used in this task can be found below.

1. Cao, F., von Bülow, S., Tesei, G., & Lindorff-Larsen, K. (2024). A coarse-grained model for disordered and multi-domain proteins. *bioRxiv*, 2024-02.
2. Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
3. Eastman, P., Swails, J., Chodera J. D., et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics *PLoS Comput Biol*, 13(7):e1005659
4. Heo, L., & Feig, M. (2024). One bead per residue can describe all-atom protein structures. *Structure*, 32(1), 97-111.
5. Abraham, M. J., Murtola, T., Schulz, R., et al. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1, 19-25.

Martini 3 simulations with increased protein-water interactions

F. Emil Thomasen¹, Fan Cao¹, Hamidreza Ghafouri², Giulio Tesei¹, Silvio C.E. Tosatto²
and Kresten Lindorff-Larsen¹

¹ - Structural Biology and NMR Laboratory & the Linderstrøm-Lang Centre for Protein Science, Department of Biology,
University of Copenhagen, Copenhagen, Denmark,

² - Department of Biomedical Sciences, University of Padova, Padova, Italy

lindorff@bio.ku.dk

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:Y; AF:AF2; EMA:N; MD:Y

We used coarse-grained molecular dynamics simulations to predict the ensemble of structures and capture the interdomain orientations between the ZLBT and C domains in challenges T1200 and T1300. Using the cg2all algorithm, we then generated the all-atom structures. The following section outlines the approach in detail.

Methods

We performed coarse-grained molecular dynamics (MD) simulations of the wild-type and mutant protein targets using Martini 3¹ with the strength of protein-water interactions increased by 10%². We used AlphaFold2 predictions³ as the initial structures and performed MD simulations using Gromacs 2021.1⁴ as previously described². We assigned a partial charge of -0.25 to the charged beads of D54, D58, E62 and E65 (PDB CODE: 2lr2) to compensate for the absence of an explicit Lanthanide ion with a +3 charge. We back-mapped our coarse-grained simulations to an all-atom representation using cg2all⁵.

To ensure that the backbone RMSD of the two folded regions fall within 0.5 Å of the experimental structures, we used the following approach:

1. Superpose the two experimental structures to each frame of the all-atom simulation trajectory.
 2. Delete the folded domains from our simulation trajectory.
 3. Combine the remaining part of the simulation trajectory with the superposed experimental structures.
 4. Fix the local geometry by relaxing each trajectory structure with position restraints on the folded domains using Charmm36⁶ in Gromacs⁴.
-
1. Souza, P. C. T., Alessandri, R., Barnoud, J., Thallmair, S., Faustino, I., Grünewald, F., Patmanidis, I., Abdizadeh, H., Bruininks, B. M. H., Wassenaar, T. A., Kroon, P. C., Melcr, J., Nieto, V., Corradi, V., Khan, H. M., Domański, J., Javanainen, M., Martinez-Seara, H., Reuter, N., ... Marrink, S. J. (2021). Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature Methods*, 18(4), 382–388.

2. Thomasen, F. E., Skaalum, T., Kumar, A., Srinivasan, S., Vanni, S., & Lindorff-Larsen, K. (2024). Rescaling protein-protein interactions improves Martini 3 for flexible proteins in solution. *Nature Communications*, 15(1), 6645.
3. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
4. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1, 19-25.
5. Heo, L., & Feig, M. (2024). One bead per residue can describe all-atom protein structures. *Structure*, 32(1), 97-111.
6. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmüller, H., & MacKerell, A. D. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, 14(1), 71–73.

MassiveFold

Massive sampling with MassiveFold in CASP16-CAPRI

Nessim Raouraoua¹, Marc F. Lensink¹ and Guillaume Brysbaert¹

¹ Univ. Lille, CNRS, UMR 8576 – UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France

guillaume.brysbaert@univ-lille.fr

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:Y; MD:N

The results of CASP15 have shown that increasing the number of predictions while including diversity in the inference process led to a significant improvement for multimer predictions¹. However, this massive sampling strategy requires access to a large GPU infrastructure to be able to generate the predictions in a short period of time, and is therefore not accessible to all predictor groups.

For CASP16, we used MassiveFold² (<https://github.com/GBLille/MassiveFold>), which allows massively expanding the sampling of structure predictions by optimizing the computing of AlphaFold^{3,4} based predictions. It improves the parallelization of the structure inference by splitting the computing on CPU for alignments, running automated batches of structure prediction on GPU, and gathering the results in a single output directory, with a consolidated ranking and a variety of plots. MassiveFold uses AFmassive (<https://github.com/GBLille/AFmassive>) inference engine, an updated version of AFsample⁶ that offers additional diversity parameters for massive sampling. MassiveFold can also use ColabFold⁵.

We used a large GPU cluster to generate 8040 predictions for the majority of the targets, submitted our top 5 and made the ensemble of predictions available to all predictors for a CASP16 phase 2 prediction round where they could use our predictions in any way they wanted to submit an updated top 5.

Methods

In collaboration with the IDRIS national french GPU supercomputing center of GNCI for the structure inference and the PLBS/SINBIOS platform for data storage, we generated massive sampling data based on AlphaFold and ColabFold. The inference integrated diversity following mainly Björn Wallner's parameterization used in CASP15¹ but exploiting the three neural network models currently available for AlphaFold for multimers. In addition, the last CAPRI conference in Feb 2024 showed that generating a high amount of structures with default AlphaFold parameters in many instances also leads to an increase in good quality models⁷, thus we added this condition.

Therefore, to integrate the maximum number of parameters in a massive sampling strategy but limiting the structures to a reasonable amount, we predicted, **for multimers**, 67 structures per AlphaFold neural network model (5 NN models x 3 NN versions = 15 for multimers) with AFmassive, resulting in 1005 structures per condition for the 6 first sets of parameters listed in **Table 1**, for a total of 6030 structures. To complement them with more diversity in particular with the ColabFold_DB and MMseqs2 approach for alignments, we computed an extra 1005 predictions with ColabFold for an additional two sets of parameters, adding 2010 ColabFold predictions to the total (Table 1).

This results in a total of **8040** structure models for each target, except for the largest ones which would require too many GPU hours. For these, we reduced the number of predictions according to the size of the target.

For monomers, because only one neural network version was published by DeepMind and not three like multimers, we compensated by computing $67 \times 3 = 201$ predictions per neural network model, therefore obtaining 8040 structure models as well for each target, again except for the largest targets for which we reduced the number of predictions according to their size.

Predicted structures were not relaxed.

All predictions were ranked based on the AlphaFold confidence score and the top 5 structures were submitted to CASP16-CAPRI phase 1 as a MassiveFold baseline. All our predictions were then made available to all the predictors for phase 2. For the specific case of targets presenting alternative conformations, human input was involved in structure selection for the top 5.

Setup	Dropout Evoformer	Dropout structure module	Templates	Recycles	Structure inference engine
afm_basic			X	21	AFmassive
afm_woTemplates				21	AFmassive
afm_dropout_full	X	X	X	21	AFmassive
afm_dropout_full_woTemplates	X	X		21	AFmassive
afm_dropout_full_woTemplates_r3	X	X		3	AFmassive
afm_dropout_noSM_woTemplates	X			21	AFmassive
cf_woTemplates				21	ColabFold
cf_dropout_full_woTemplates	X	X		21	ColabFold

Table 1: Sets of parameters used for massive sampling in CASP16-CAPRI

Results

We computed massive sampling with MassiveFold for 73 targets (34 monomers and 39 multimers). 8040 predictions were generated for each target, except for the following ones because of their large size:

- H1217 (5878 residues): 395 predictions
- H1227 (5689 residues): 45 structures were generated and the top 5 were submitted for phase 1, but for phase 2, the structure was trimmed to 2101 residues and 8040 predictions were generated
- H1258 (3092 residues); T1257 (3789 residues) and H1265 (3924 residues): 2040 predictions
- T1271 subunits: 2680 predictions each
- T1295 (3752 residues) and T1269 (2820 residues): 4080 predictions

In total, 510475 structures were produced.

Availability

MassiveFold is available here: <https://github.com/GBLille/MassiveFold>

Acknowledgements

This work was granted access to the HPC resources of IDRIS (under the allocations 2024-AD010715407, 2024-AD010715332 and 2024-AD010715333) made by GENCI (Grand Equipement National de Calcul Intensif).

1. Wallner B. 2023. Improved multimer prediction using massive sampling with AlphaFold in CASP15. *Proteins*.
2. Raouraoua N, Mirabello C, Véry T, Blanchet C, Wallner B, Lensink MF, et al. 2024. MassiveFold: unveiling AlphaFold's hidden potential with optimized and parallelized massive sampling. *Research Square*
3. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589.
4. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. 2021. Protein complex prediction with AlphaFold-Multimer.
5. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. *Nat. Methods* 19: 679–682.
6. Wallner B. 2023. AFsample: improving multimer prediction with AlphaFold using massive sampling. *Bioinforma. Oxf. Engl.* 39: btad573.
7. Raouraoua, Nessim, Lensink, Marc F, Brysbaert, Guillaume. 2024. Massive sampling strategy for antibody-antigen targets in CAPRI Round 55 with MassiveFold. Available from <https://doi.org/10.22541/au.172592104.47153431/v2>.

Manual Prediction of Protein Tertiary and Quaternary Structures, Model Quality, Protein-Ligand Interactions and Ensembles

L.J. McGuffin¹, N.S. Edmunds¹, A.G. Genc¹, S.N. Alhaddad¹, R. Adiyaman¹

1 - School of Biological Sciences, University of Reading, Reading, UK

l.j.mcguffin@reading.ac.uk

Key: Auto:Y; CASP_serv:Y; Templ:Y; MSA:Y; DeepL:Y; AF:AF2/AF3; EMA:Y; MD:N

For our manual predictions, we used several components from our recent IntFOLD¹ and ModFOLD² servers and the newest versions of our quaternary structure modelling and scoring servers, MultiFOLD¹ and ModFOLDdock³ (see our MultiFOLD2 and ModFOLDdock2 abstracts). For our tertiary and quaternary structure predictions (TS format), we made use of our MultiFOLD2 models, the CASP-hosted 3D server models and AlphaFold3⁴ models, which we ranked using either our ModFOLD9² or ModFOLDdock2 methods and then refined with our AlphaFold2-Multimer_Refine (AF2MR)⁵ method. The likely ligand binding sites (LG format) were predicted with our latest version of FunFOLD⁶ and ensembles were generated using our latest ReFOLD⁵ method.

Methods

Tertiary and Quaternary Structure Predictions:

For Phase 0 (T0/H0 targets), no stoichiometry information was provided for the multimers, so this was predicted by MultiFOLD2 and manually checked (see our MultiFOLD2 abstract). The quaternary structure server models were ranked using our ModFOLDdock2R method (see our ModFOLDdock2 abstract). The top 10 models were then selected and used as input templates for the AF2MR⁵ pipeline (see our MultiFOLD2 abstract), which used the LocalColabFold⁷ 1.5.2 method with the “--custom-template-path” option. We used 12 recycles and relaxation for targets <=2000 residues or 3 recycles without relaxation for larger targets.

For Phase 1 (T1/H1 targets), the stoichiometry information for each target was provided. If the stoichiometry was incorrectly predicted at Phase 0, then the MultiFOLD2 protocol was rerun with the corrected stoichiometry and additional server models were collected. The tertiary structure Phase 1 server models and the quaternary structure Phase 1 server models were ranked using the ModFOLD9² and ModFOLDdockR methods respectively (see our ModFOLDdock2 abstract). The top 10 models were then selected and used as input templates for the AF2MR⁵ pipeline.

For Phase 2 (T2/H2 targets), the stoichiometry information and models from all groups and MassiveFold⁸ were available. For the multimer targets, ModFOLDdock2R (see our ModFOLDdock2 server abstract) was used to select the top 5 ranked models from all groups. For the monomer (A1) targets

a quick version of our ModFOLD9² method (ModFOLD9Q) was used to select the top 5 ranked models from the MassiveFold set (see below). Again, the top 10 models were selected and used as input templates for the AF2MR⁵ pipeline.

For each TS format prediction, the final model rankings and the predicted per-residue quality scores (pLDDT*100) from LocalColabFold were added to the B-factor column for each set of atom records. For some of the very large complexes (>6000 residues), due to our limited GPU resources, we had to divide sequences up into overlapping fragments for submission to MultiFOLD2. The resulting modelled fragments were then manually assembled, using structural superposition in PyMOL (<https://www.pymol.org>), to form larger, more complete models.

Model Quality Assessment (QA QMODE3): Due to the lack of time for processing the MassiveFold data, quicker versions of our ModFOLD9² method (ModFOLD9Q) and ModFOLDdock2 method (ModFOLDdock2Q) were used to manually score and rank the monomeric and multimeric MassiveFold models respectively for QA QMODE3. The ModFOLD9Q method used the top 40 ModFOLD9 ranked server models for Phase 1 (T1) targets as reference sets for comparison against the MassiveFold models. OpenStructure⁹ version 2.7 was used to obtain the global scores for each pairwise model comparison (using the “ost compare-structures” action). MassiveFold models were ranked by the mean of the Oligo-GDTJury and IDDTJury scores (see our ModFOLDdock2 server abstract) and the top 5 models were returned in the QMODE3 file. The ModFOLDdock2Q method used the same approach as the ModFOLDdock2R method for ranking (see our ModFOLDdock2 server abstract) however, no local scores were generated and MassiveFold models were ranked by the mean of the IDDTJury, DockQ-waveJury and VoroIF(weighted_average_pcadscore) scores. Additionally, in the first stage the top n models were selected using VoroIF(sum_of_gnn_scores) and these models served as a reference set against which all other models were compared (where n=40 for targets with a total sequence length of <=500 amino acids (aa), n=30 for lengths >500aa & <=1000aa, n=20 for lengths >1000aa & <=1500aa, n=10 for lengths >1500aa & <=2000aa, n=5 for lengths >2000aa).

Ligand binding predictions (FunFOLD5): Our top five manually selected TS models in the human prediction category were used to locate the target ligands in the individual subunits by using the template-based and blind docking pipelines. For the template-based pipeline, the relevant template list was generated using Foldseek¹⁰ for the related MultiFOLD2 models. Then FunFOLD⁶ was run to find the biologically relevant binding sites and ligands based on the template list. The chemical properties of the ligands predicted by FunFOLD5 were then compared to those of the CASP target ligand. For the blind docking method, AutoSite¹¹, Fpocket¹², P2rank¹³ were run to identify potential binding sites and Gnina¹⁴, DiffDock¹⁵ RoseTTAFold-All-Atom¹⁶ and AutoDock Vina¹⁷ were run to identify potential binding poses for the target ligand without relying on any specific templates. These potential binding sites and poses were also interpreted with the FunFOLD predictions, and then the most common binding sites were selected to re-dock the target ligands. For each target, the ligand pose, LScore and affinity data were generated by redocking with Gnina¹⁴ and AutoDock Vina¹⁷ for the top five TS models (the corresponding protein receptor), which were then submitted in LG format.

Ensembles: The 3D models produced by MultiFOLD2 for T1200 and T1300 underwent further refinement through our ReFOLD4 pipeline to capture conformational shifts in the associated models. Subsequently, ModFOLD9² scored all models generated by ReFOLD4⁵. The ModFOLD9 scores were then normalized by the total score for all models and used as probabilities for the populations.txt file.

Availability

Server methods are available via: <https://www.reading.ac.uk/bioinf/>. Software is free to download via: <https://www.reading.ac.uk/bioinf/downloads/>.

1. 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. *Nucleic Acids Research*. **51**(W1), W274-W280. doi: 10.1093/nar/gkad297
2. McGuffin,L.J. & Alharbi,S.M.A. (2024) ModFOLD9: a web server for independent estimates of 3D protein model quality. *J. Mol. Biol.* **436**, 168531. doi:10.1016/j.jmb.2024.168531
3. Edmunds,N.S., Alharbi,S.M.A., Genc,A.G., Adiyaman,R. & McGuffin,L.J. (2023) Estimation of Model Accuracy in CASP15 Using the ModFOLDDock Server. *Proteins*. **91**, 1871-1878. doi: 10.1002/prot.26532
4. Abramson,J., Adler,J., Dunger,J. et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. **630**, 493–500. doi: 10.1038/s41586-024-07487-w
5. Adiyaman,R., Edmunds,N.S., Genc,A.G., Alharbi,S.M.A. & McGuffin,L.J. (2023) Improvement of protein tertiary and quaternary structure predictions using the ReFOLD refinement method and the AlphaFold2 recycling process. *Bioinformatics Advances*. vbad078. doi: 10.1093/bioadv/vbad078
6. Roche,D.B., Tetchner,S.J., & McGuffin,L.J. (2011) FunFOLD: an improved automated method for the prediction of ligand binding residues using 3D models of proteins. *BMC Bioinformatics*. **12**(1), 1-20. doi: 10.1093/nar/gkt498
7. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S., & Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nature Methods*. **19**(6), 679–682. doi: 10.1038/s41592-022-01488-1
8. Brysbaert,G., Raouraoua,N., Mirabello,C. et al. (2024) MassiveFold: unveiling AlphaFold's hidden potential with optimized and parallelized massive sampling. PREPRINT (Version 1) available at Research Square. doi: 10.21203/rs.3.rs-4319486/v1
9. Biasini,M., Schmidt,T., Bienert,S., Mariani,V., Studer,G., Haas,J., Johner,N., Schenk,A.D., Philippse,A. & Schwede,T. (2013) OpenStructure: an integrated software framework for computational structural biology. *Acta Crystallographica Section D: Biological Crystallography*. **69**(Pt 5), 701–709. doi: 10.1107/S0907444913007051
10. van Kempen,M., Kim,S.S., Tumescheit,C. et al. (2024) Fast and accurate protein structure search with Foldseek. *Nat Biotechnol*. **42**, 243–246. doi: 10.1038/s41587-023-01773-0
11. Ravindranath,P.A. & Sanner,M.F. (2016) AutoSite: an automated approach for pseudo-ligands prediction—from ligand-binding sites identification to predicting key ligand atoms. *Bioinformatics*. **32**(20), 3142-3149. doi: 10.1093/bioinformatics/btw367
12. Le Guilloux,V., Schmidtke,P. & Tuffery,P. (2009) Fpocket: an open source platform for ligand pocket

- detection. *BMC Bioinformatics*. **10**, 1-11. doi: 10.1186/1471-2105-10-168
- 13. Krivák,R. & Hoksza,D. (2018) P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J. Cheminform.* **10**, 1-12. doi: 10.1186/s13321-018-0285-8
 - 14. McNutt,A.T., Francoeur,P., Aggarwal,R. et al. (2021) GNINA 1.0: molecular docking with deep learning. *J. Cheminform.* **13**, 43. doi: 10.1186/s13321-021-00522-2
 - 15. Corso,G., Deng,A., Fry,B., Polizzi,N., Barzilay,R. & Jaakkola,T. (2024) Deep confident steps to new pockets: Strategies for docking generalization. *ArXiv*. 2402.18396. doi: 10.48550/arXiv.2402.18396
 - 16. Krishna,R., Wang,J., Ahern,W. et al. (2024). Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*. **384**(6693), eadl2528. doi: 10.1126/science.adl2528
 - 17. Trott,O., & Olson,A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. **31**(2), 455-461. doi: 10.1002/jcc.21334

Protein and Nucleic Acid Ensemble Structure Prediction in CASP16

Qiqige Wuyun¹, Wentao Ni², Chunxiang Peng^{3,4} and Wei Zheng^{2,3,4}

1 - Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

2 - School of Statistics and Data Science, Nankai University, 94 Weijin Road, 300071, Tianjin, China

3 - Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

4 - Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

zhengwei@umich.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:N; MD:N

The MIEnsembles-Server group in CASP16 is designed for modeling a wide range of targets, including protein monomers, nucleic acids (DNA or RNA), nucleic acid-nucleic acid complexes, protein-protein complexes, protein-nucleic acid complexes, ensemble targets, and model quality assessment (QA) targets. Nucleic acid-related targets are predicted using the newly developed deep learning method, DeepProtNA, which integrates pre-trained language model embeddings, multiple sequence alignments (MSA), predicted secondary structures, and structural templates as inputs for modified Evoformer blocks. Protein monomer and multimer targets are modeled using a modified version of DMFold¹. Ensemble targets are predicted using a clustering approach based on model structural similarity, and QA targets are assessed by selecting the best DMFold model as a reference.

Methods

Protein monomer and multimer modeling. Protein monomer and multimer targets are predicted using a modified DMFold approach, which involves three steps.

The first step is MSA construction, which builds on the previous DeepMSA2 pipeline. Compared to the version used in CASP15, this updated DeepMSA introduces two key improvements: (i) a larger in-house metagenomic sequence database, incorporating data from IMG/M, NCBI, and EBI, and (ii) a multi-domain MSA assembly method that merges domain-level MSAs into a full chain-level MSA. Similar to DeepMSA2, the new pipeline also contains three MSA construction sub-pipelines: dMSA, qMSA, and mMSA. These three sub-pipelines are iteratively used to collect homologous sequences from genomic and metagenomic databases, including UniClust30, UniRef90, MetaClust, Mgnify, BFD, and an in-house huge metagenomics database. MSAs generated from these sub-pipelines are input into AlphaFold2 (1-embedding) to predict a set of models. These MSAs are then ranked by their associated *pLDDT* scores from AlphaFold2. For multi-domain targets, the same MSA generation method is used to construct domain-level MSAs based on predicted domain boundaries, which are then assembled into full-length MSAs by linking sequences from the same species. The ranked MSAs are either directly used in protein monomer modeling or paired as multimer MSAs for protein complex modeling. For heteromeric complexes, an additional selection process generates an optimal set of paired MSAs by combining individual constituent MSAs. The top *N* ranked MSAs for each constituent protein are chosen to form

potential paired MSAs, and for a heteromeric complex with M constituent proteins, N^M distinct paired MSAs are generated and evaluated based on a combined score of the depth of the MSAs and pLDDT score of the monomer chains. To ensure the pipeline completion within three days, N is selected to satisfy $N^M \leq 64$.

The second step is template detection based on a new version of LOMETS (LOMETS4). Compared to LOMETS3², which was used in CASP15, the major update in LOMETS4 is its ability to handle protein complexes. For protein heteromers, templates are identified as follows: first, homologous templates for each constituent chain in the target complex are identified using LOMETS3, which includes six profile-based threading methods, five contact/distance-based threading methods, three protein language model-based threading methods, and one structure-based threading method. Notably, templates for individual chains that have already been considered in previous steps are excluded to prevent the similar query constituent chain from hitting identical templates. The templates are ranked by quality (e.g., *Z-score*). Finally, if at least two constituent chains share templates from the same protein complex and have a high sum *Z-score*, these complexes are considered potential templates.

The third step is structure model generation, which utilizes a modified AlphaFold2 modeling engine. The MSAs from the first step and the structure templates from the second step serve as input features for this modeling engine. Key modifications to the AlphaFold2 modeling engine include: (i) using templates or not, (ii) adjusting the dropout rate, (iii) applying different versions of AlphaFold2 pre-trained weights (v1-v3), (iv) generating a higher number of decoys than the default setting (25 models), (v) applying or omitting the early stop strategy in AlphaFold2 (v2.3), and (vi) extending the modeling iterations, (vii) MSA sampling for ensemble targets. The final models are ranked based on the *pLDDT* score for monomer targets, or by confidence scores ($0.8pTM + 0.2ipTM$) for complex targets.

For targets with unknown stoichiometry, a newly developed method is used to determine stoichiometry information. This method uses two pipelines: one derives stoichiometry from top ranked LOMETS4 templates, the other predicts stoichiometry directly by DMFold confidence scores, where the oligomeric state is predicted based on the model with the highest confidence.

Nucleic acid-related target modeling. DeepProtNA is an end-to-end deep learning algorithm designed for predicting protein-nucleic acid complex and nucleic acid complex structures. The method integrates pre-trained language model embeddings, multiple sequence alignment information, predicted secondary structure, and structural templates to directly generate three-dimensional coordinates of the complexes from input sequences through a set of modified Evoformer blocks and a structure module similar to AlphaFold2³. Protein and RNA sequences are input respectively into the pre-trained language models, ESM⁴ and RNA-FM⁵, to generate high-dimensional sequence embeddings. MSA is generated for both protein and nucleic acid sequences using the modified version DeepMSA2¹ and rMSA⁶. The MSA-derived features are combined with the language model embeddings to enhance the understanding of sequence interactions. Structural templates are selected from PDB by LOMETS4 for protein or BLASTn⁷ for RNA. Additionally, the predicted secondary structures of nucleic acids are also selected as inputs of the deep learning networks. The main trunk of DeepProtNA processes the embeddings for proteins and nucleic acids, leveraging self-attention mechanisms to capture long-range dependencies within each sequence. Cross-attention mechanisms are used to handle interactions between protein and nucleic acid sequences, allowing the model to focus on key interaction sites. A structure decoder network takes the embedded representation to generate the three-dimensional coordinates for the complex. This network translates

sequence information directly into spatial coordinates for each residue and nucleotide, producing the final structure. DeepProtNA also outputs a confidence score that measures the reliability of the predicted interactions, and the confidence score is also used for predicting the oligomer states for those RNA targets with unknown stoichiometry.

Ensembles modeling. For ensemble targets, the SPICKER⁸ method is used to cluster the decoys based on structural similarity for protein monomers, RNAs, protein-protein complexes, and protein-nucleic acid complexes. The centers of clusters with a large number of members and high confidence scores are selected as models representing potential alternative conformations.

Model accuracy estimation. For model accuracy estimation targets, the best model from DMFold is selected as the reference model, and the quality of decoys is assessed based on this reference. The TM-score⁹ between the reference model and each decoy is used to predict overall fold accuracy, while the DockQ¹⁰ score is used to predict overall interface accuracy.

1. Zheng,W., Wuyun,Q., Li,Y., Zhang,C., Freddolino,P.L. & Zhang,Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods* 21, 279-289.
2. Zheng,W., Wuyun,Q., Zhou,X., Li,Y., Freddolino,P.L. & Zhang,Y. (2022). LOMETS3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic Acids Research* 50, W454-W464.
3. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., Back,T., Petersen,S., Reiman,D., Clancy,E., Zielinski,M., Steinegger,M., Pacholska,M., Berghammer,T., Bodenstein,S., Silver,D., Vinyals,O., Senior,A.W., Kavukcuoglu,K., Kohli,P. & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
4. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., dos Santos Costa,A., Fazel-Zarandi,M., Sercu,T., Candido,S. & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123-1130.
5. Chen,J., Hu,Z., Sun,S., Tan,Q., Wang,Y., Yu,Q., Zong,L., Hong,L., Xiao,J., Shen,T., King,I. & Li,Y. (2022). Interpretable RNA Foundation Model from Unannotated Data for Highly Accurate RNA Structure and Function Predictions. *bioRxiv*, 2022.08.06.503062.
6. Zhang,C., Zhang,Y. & Pyle,A.M. (2023). rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *Journal of Molecular Biology* 435, 167904.
7. Zhang,Z., Schwartz,S., Wagner,L. & Miller,W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology* 7, 203-214.
8. Zhang,Y. & Skolnick,J. (2004). SPICKER: a clustering approach to identify near-native protein folds. *Journal of computational chemistry* 25, 865-871.
9. Zhang,Y. & Skolnick,J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics* 57, 702-710.
10. Basu,S. & Wallner,B. (2016). DockQ: A Quality Measure for Protein-Protein Docking Models. *PLOS One* 11.

ModFOLDdock2, ModFOLDdock2R, ModFOLDdock2S

Automated Quality Assessment of Protein Quaternary Structure Models using the ModFOLDdock2 Server

L.J. McGuffin¹, A.G. Genc¹, R. Adiyaman¹ and N.S. Edmunds¹

1 - School of Biological Sciences, University of Reading, Reading, UK

l.j.mcguffin@reading.ac.uk

Key: Auto:Y; CASP_serv:Y; Templ:Y; MSA:Y; DeepL:Y; AF:AF2; EMA:Y; MD:N

The ModFOLDdock2 server is our new web resource for the Quality Assessment (QA QMODE2) of protein quaternary structure models. Three variants of ModFOLDdock2 were tested at CASP16, which were optimised for the different facets of the quality estimation problem.

Methods

The ModFOLDdock2 server uses a hybrid consensus approach for producing both global and local (interface residue) quality scores for predicted quaternary structures. In brief, the main differences from the original ModFOLDdock¹ server were the addition of new scores and a neural network to predict local scores. The ModFOLDdock2 variants use various combinations of global and local scores (detailed in the sections below), which are calculated using the output from 12 individual scoring methods: QS-bestJury, DockQ-waveJury, TM-scoreJury, Oligo-GDTJury, IDDTJury, CADJury, PatchQSJury, PatchDockQJury, VoroMQA, VoroIF, CDA, and ModFOLDIA.

For the QS-bestJury, DockQ-waveJury, TM-scoreJury, Oligo-GDTJury, IDDTJury, CADJury, PatchQSJury, and PatchDockQJury scoring methods, selected pairwise comparisons were made between each quaternary structure model and every other model and then the mean scores were calculated. OpenStructure² version 2.7 was used to obtain the QS³, DockQ⁴, TM-score⁵, GDT⁶, IDDT⁶ and CAD⁷ scores for each pairwise comparison (using the “ost compare-structures” action). The VoroIF(VoroIF-GNN)⁸ and VoroMQA(voronota-js-voromqa)⁹ methods were used off-the-shelf, VoroIF with the “--assembly true” and “--local-column true” options, and VoroMQA with the “--inter-chain” and “--output-dark-scores” options.

The CDA score was based on the original Contact Distance Agreement (CDA) score^{10,11}, which relates to the agreement between the residue contacts predicted from the sequence and the measured Euclidean distance (in Å) between residues in the model. In this case, we used the contact prediction profiles that resulted from generating LocalColabFold¹² version 1.0.0 multimer models.

The ModFOLDIA method was also used to carry out structure-based comparisons of alternative oligomer models and it produced both global and local/per-residue interface scores. The first stage of the ModFOLDIA method was to identify the interface residues in the model to be scored (defined as $\leq 8\text{\AA}$ between C β atoms, or C α for GLY) and then obtain the minimum contact distance (D_{min}) for each contacting residue. The second stage was to locate the equivalent residues in all other models and then obtain the mean minimum distances of those residues in all other models ($MeanD_{min}$). The final Interface

Accuracy (IA) score for each of the interface residues in the model was the absolute difference in the S_i from the mean S_i : $IA = 1 - |S_i - MeanS_i|$, where $S_i = 1/(1 + (D_{min}/20)^2)$ and $MeanS_i = 1/(1 + (MeanD_{min}/20)^2)$. The global ModFOLDIA score for a model was then taken as the total interface score (sum of residue scores) normalised by the maximum of either the number of residues in the interface or the mean number of interface residues across all models for the same target.

ModFOLDdock2: This variant produced predicted scores optimised for positive linear correlations with the observed scores, i.e., the predicted quality scores correlated well with the observed quality scores, according to the assessors' formulae for CASP15 multimer models¹³. In the first stage the top n models were selected using VoroIF(sum_of_gnn_scores) and these models served as a reference set against which all other models were compared (where n=40 for targets with a total sequence length of <=2000 amino acids (aa), n=10 for lengths >2000aa & <=8000aa, n=5 for lengths >8000aa). The overall fold accuracy (column 2 in the QA file) was calculated from the mean of the TM-scoreJury and Oligo-GDTJury scores. The overall interface accuracy (column 3) was calculated from the mean of the QS-bestJury, DockQ-waveJury and VoroIF(weighted_average_pcadscore) scores. Additionally, confidence scores for the interface residues in each model were calculated from the mean of the IDDTJury(local), CADJury(local), PatchQSJury and PatchDockQJury scores.

ModFOLDdock2R: This variant produced predicted scores optimised for ranking, i.e. the top-ranked models (top 1) should have higher observed overall accuracy, but the relationship between predicted and observed scores may not be linear. The overall fold accuracy (column 2) was calculated from the mean of the IDDTJury, DockQ-waveJury and VoroIF(weighted_average_pcadscore) scores. The overall interface accuracy (column 3) was calculated from the mean of the VoroMQA, DockQ-waveJury and VoroIF(weighted_average_pcadscore) scores. The confidence scores for the interface residues in each model were also calculated from the mean of the IDDTJury(local), CADJury(local), PatchQSJury and PatchDockQJury scores.

ModFOLDdockS2: This variant used a quasi-single model approach to score models. Sets of reference multimer models were first generated from the input sequences using our MultiFOLD2 method (see our MultiFOLD2 abstract for details) then each model was compared individually against the reference set using the individual scoring methods described above. The local scores were fed as inputs into a neural network (NN), implemented using the MLP (multi-layer perceptron) from the RSNNNS package to predict the interface residue scores. The MLP was trained to predict the mean of the local IDDT, CAD, PatchQS and PatchDockQ scores as the target score. The MLP input data consisted of the local VoroMQA, VoroIF, IDDTJury, CADJury, PatchQSJury, PatchDockQJury, CDA, and ModFOLDIA scores for each interface residue as well the 5 nearest contacting (<=8.0Å) interface residues in order of their proximity, and the output was a single quality score (8 local scores * 6 residues = 48 input neurons, 6 hidden, 1 output). The MLP was trained and tested on the CASP15 model data using three-fold cross-validation. The overall fold accuracy (column 2) was calculated from the mean of the TM-scoreJury, VoroIF(weighted_average_pcadscore) and global NN output scores. The overall interface accuracy (column 3) was calculated from the mean of the QS-bestJury, VoroIF(weighted_average_pcadscore) and global NN output scores. Confidence scores for interface residues in each model were the local NN output scores.

Note: Quicker versions of our ModFOLD9¹⁴ method (ModFOLD9Q) and ModFOLDdock2 method (ModFOLDdock2Q) were used to manually score and rank the monomeric and multimeric MassiveFold models respectively for QA QMODE3 (see our McGuffin group abstract).

Availability

The ModFOLDdock server is available at:

https://www.reading.ac.uk/bioinf/ModFOLDdock/ModFOLDdock_form.html

1. Edmunds,N.S., Alharbi,S.M.A., Genc,A.G., Adiyaman,R. & McGuffin,L.J. (2023) Estimation of Model Accuracy in CASP15 Using the ModFOLDdock Server. *Proteins.* **91**, 1871-1878. doi: 10.1002/prot.26532
2. Biasini,M., Schmidt,T., Bienert,S., Mariani,V., Studer,G., Haas,J., Johner,N., Schenk,A.D., Philippse,A. & Schwede,T. (2013) OpenStructure: an integrated software framework for computational structural biology. *Acta Crystallographica Section D: Biological Crystallography.* **69**(Pt 5), 701–709. doi: 10.1107/S0907444913007051
3. Bertoni,M., Kiefer,F., Biasini,M., Bordoli,L. & Schwede,T. (2017) Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep.* **7**, 10480. doi: 10.1038/s41598-017-09654-8.
4. Basu,S. & Wallner,B. (2016) DockQ: A Quality Measure for Protein-Protein Docking Models. *PLoS One.* **11**, e0161879. doi: 10.1371/journal.pone.0161879.
5. Zhang, Y. & Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins.* **57**, 702-710. doi: 10.1002/prot.20264
6. Mariani,V., Biasini,M., Barbato,A. & Schwede,T. (2013) IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics.* **29**(21), 2722–2728. doi: 10.1093/bioinformatics/btt473
7. Olechnovic,K., Kulberkyte,E. & Venclovas C. (2013) CAD-score: A new contact area difference-based function for evaluation of protein structural models. *Proteins.* **81**, 149-162. doi: 10.1002/prot.24172
8. Olechnovic.,K & Venclovas,C. (2023) VoroIF-GNN: Voronoi tessellation-derived protein-protein interface assessment using a graph neural network. *Proteins.* **91**, 1879-1888. doi: 10.1002/prot.26554.
9. Olechnovič,K, & Venclovas,C. (2014) Voronota: A fast and reliable tool for computing the vertices of the Voronoi diagram of atomic balls. *Journal of Computational Chemistry.* **35**(8), 672–681. doi: 10.1002/jcc.23538
10. Maghrabi,A.H.A. & McGuffin,L.J. (2017) ModFOLD6: an accurate web server for the global and local quality estimation of 3D models of proteins. *Nucleic Acids Res.*, **45**, W416-W421. doi: 10.1093/nar/gkx332.
11. McGuffin,L.J., Aldowsari, F.M.F., Alharbi,S.M.A. & Adiyaman,R. (2021) ModFOLD8: accurate global and local quality estimates for 3D protein models. *Nucleic Acids Research.* **49**(W1), W425–W430. doi: 10.1093/nar/gkab321
12. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. & Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nature Methods.* **19**(6), 679–682. doi: 10.1038/s41592-022-01488-1
13. Studer,G., Tauriello,G., & Schwede,T. (2023) Assessment of the assessment—All about complexes. *Proteins.* **91**(12), 1850-1860. doi: 10.1002/prot.26612
14. McGuffin,L.J. & Alharbi,S.M.A. (2024) ModFOLD9: a web server for independent estimates of 3D protein model quality. *J. Mol. Biol.* **436**, 168531. doi: 0.1016/j.jmb.2024.168531

[MQA_base](#), [MQA_server](#), [MQA](#)

Deep Learning based protein complex model accuracy estimation and model selection

Jun Liu¹, Weikang Gong², Jun Hu², Biao Zhang², and Yang Zhang^{1,3,4}

1 - Cancer Science Institute of Singapore, National University of Singapore;

2 - Center for AI and Computational Biology, Institute of Systems Medicine, Chinese Academy of Medical Sciences /Suzhou Institute of Systems Medicine;

3 - Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore;

4 - Department of Computer Science, the School of Computing, National University of Singapore.

zhang@zhanggroup.org

Key: Auto:Y; CASP_serv:Y; Templ:N; MSA:N.MetaG; Fragm:N; Cont:N; Dist:N; Tors:N; DeepL:Y; EMA:Y; MD:N.

While remarkable progress has been recently achieved on protein tertiary structure prediction, accurate prediction of protein quaternary complex structures remains a significant unsolved problem. To address the challenge, development of effective methods, to assess the precision of the protein complex structural models and recognize the best from conformational decoys created by current structural prediction approaches, becomes increasingly important. In CASP16, we developed and tested three model accuracy estimation programs, named **MQA_base**, **MQA_server**, and **MQA**, respectively, to evaluate both the global and the interface realities of the predicted protein complex models.

Methods

We developed the three programs by extending the previous method of DeepUMQA3^{1,2}. These programs use the same input features and network architecture but have major differences in the training strategies and the evaluation score processing.

Training data collection: The training protein complexes were collected from the PDB database with the following criteria: a resolution of ≤ 9 Å, no more than 10 chains, each chain with a maximum length of 1500 amino acids, and a total complex length of ≤ 3000 amino acids. Complexes with the same number of protein chains were clustered based on 30% sequence similarity. One complex was selected from each cluster, and for each complex, multiple structural models were generated using AlphaFold-Multimer³ and Hdock⁴, respectively. These protein complexes were then combined with protein complex from CASP15 to form the final training dataset, where for the CASP15 complexes, the structural models predicted by the CASP15 predictors are used as the training models.

Feature collection: Features are extracted exclusively from the protein complex model itself to predict both per-residue accuracy and overall model accuracy. These features can be categorized into three groups: (1) sequence-based attributes, including amino acid sequence encoding and positional encoding; (2) structural attributes, including inter-residue constraints across the entire complex, residue-level local and topological structural information, and inter-monomer residue topology; (3) physicochemical attributes, including energy terms for individual residues and residue pairs within the monomer.

Accuracy prediction: A transformer-based network model was trained to predict both the IDDT⁵ for individual residues and the TM-score⁶ for the global protein complex. The network comprises three

components. The first component processes and integrates input features into pairwise representations. The second component updates these representations through transformer blocks. In the final component, residual networks are used to predict the inter-residue distance deviation map and a mask map with a 15 Å threshold to compute residue-level IDDT, while a multi-layer perceptron (MLP) is employed to predict the global TM-score.

Difference between the three programs: **MQA_base** computes residue-level IDDT and extracts interface residue scores utilizing the complete distance deviation map and mask map. Its training loss function comprises cross-entropy loss for the distance deviation map, binary cross-entropy loss for the mask map, L2 loss for IDDT, and L1 loss for the TM-score. **MQA_server** calculates residue-level IDDT using the inter-chain distance deviation map and mask map, with the distance deviation loss and mask map loss specifically considering only inter-chain residue pairs. Additionally, **MQA_server** evaluates protein models generated by the GHZ-ISM group and selects high-scoring models for submission. Finally, **MQA** utilizes **MQA_server** to evaluate all models, with the score for each interface residue defined as the average score of that residue across all models.

1. Liu J, Liu D, Zhang G J. DeepUMQA3: a web server for accurate assessment of interface residue accuracy in protein complexes. *Bioinformatics*, 2023, 39(10): btad591.
2. Liu J, Liu D, He G, et al. Estimating protein complex model accuracy based on ultrafast shape recognition and deep learning in CASP15. *Proteins: Structure, Function, and Bioinformatics*, 2023, 91(12): 1861-1870.
3. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021: 2021.10. 04.463034.
4. Yan Y, Tao H, He J, et al. The HDOCK server for integrated protein–protein docking. *Nature protocols*, 2020, 15(5): 1829-1852.
5. Mariani V, Biasini M, Barbato A, et al. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 2013, 29(21): 2722-2728.
6. Zhang C, Shine M, Pyle A M, et al. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nature methods*, 2022, 19(9): 1109-1115.

Enhanced Protein Structure Prediction through MSA Realignment
Lupeng Kong

Beijing Changping Lab, Beijing, China,

lupengkong@cpl.ac.cn

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

MRAFold performed fully automated protein tertiary structure predictions for protein monomer and complex targets. MRAFold aims to improve protein structure prediction performance by realigning MSAs, with the goal of enhancing their quality.

Methods

The overall structure prediction pipeline includes the following steps:

1) MSA Generation. We utilized the default MSA generation process of AlphaFold2¹⁻² to search for MSA, in BFD³, updated UniRef30⁴, updated UniRef90⁵, and updated MGnify⁶ databases. Simultaneously using Jackhmmer⁷ to search JGI metagenomic database⁸ to address the situation where relatively few MSAs constructed by AlphaFold2.

2) MSA Realignment. We realigned the MSAs generated in step 1 using an in-house sequence alignment tool. This tool accepts protein sequences as input and employs neural networks to learn the similarity between these sequences, thereby outputting higher-quality sequence alignments.

3) Protein Structure Prediction. The MSA from steps 1 and 2 were fed into the AlphaFold2 model to predict protein structures. The predicted structures were sorted based on the confidence score provided by AlphaFold2 and the consistency score of the predicted structures. The top 5 structures were then selected for submission.

For protein complexes longer than 3000 residues, we divided the protein into protein components smaller than 3000 residues, ensuring that one chain was identical between each component. Finally, we used TM-align⁹ to reassemble these components into a complete complex.

1. Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
2. Evans, R., O'Neill, M., Pritzel, A., et al. (2021). Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021-10.
3. Steinegger, M., Mirdita, M., & Söding, J. (2019). Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods*, 16(7), 603-606.
4. Remmert, M., Biegert, A., Hauser, A., et al. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods*, 9(2), 173-175.
5. Suzek, B. E., Huang, H., McGarvey, P., et al. (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, 23(10), 1282-1288.
6. Richardson, L., Allen, B., Baldi, G., et al. (2023). MGnify: the microbiome sequence data analysis resource in 2023. *Nucleic Acids Research*, 51(D1), D753-D759.

7. Eddy, S. R. (2011). Accelerated profile HMM searches. PLoS Computational Biology, 7(10), e1002195.
8. Chen, I. M. A., Chu, K., Palaniappan, K., et al. (2019). IMG/M v. 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Research, 47(D1), D666-D677.9.
9. Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Research, 33(7), 2302-2309.

MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human
(TS)

Enhancing AlphaFold2/AlphaFold3-based Protein Tertiary Structure Prediction with Extensive Modeling Sampling and Ranking

Jian Liu, Pawan Neupane, Jianlin Cheng*

University of Missouri, Columbia, MO 65211, USA

*Corresponding author: chengji@missouri.edu

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3/AF2; EMA:Y; MD:N;

During CASP16, we applied our latest MULTICOM4 protein structure prediction system (an upgrade of MULTICOM^{1,2}) built on top of AlphaFold2³ and AlphaFold-Multimer⁴ as well as AlphaFold3⁵ to generate a large number of tertiary structural models for monomer targets. The structural models were evaluated using multiple quality assessment (QA) methods, including global pLDDT scores, GATE (a Graph Transformers for Estimating Protein Model Accuracy), GCPNet-EMA⁶, EnQA⁷, average pairwise similarity scores, and DeepRank3⁸. MULTICOM4 was implemented as three server predictors (MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM) and two human predictors (MULTICOM, MULTICOM_human), participating in the tertiary structure prediction in CASP16.

Methods

1. Phase 1 monomer structure prediction

Tertiary structural model generation for a single protein chain (monomer)

Multiple sequence alignment (MSA) sampling. The sequence of a monomer was searched against various sequence databases, including UniRef30, UniRef90⁹, BFD^{10,11}, MGnify clusters¹², UniProt⁹, JGIclust¹³, TaraDB¹⁴ and MetaSourceDB metagenome databases¹⁵ using HHblits^{16,17}, JackHMMER¹⁸ to generate a diverse set of multiple sequence alignments (MSAs).

Template identification. In addition to using the structural templates identified by the default AlphaFold2, the MSA generated from UniRef90 was used to search our inhouse template database curated from Protein Data Bank (PDB)¹⁹ to identify alternative templates.

Structure generation with AlphaFold2 and AlphaFold3. A customized version of AlphaFold2 was used to generate structural models using the MSAs and templates generated from the previous steps. Each combination of a MSA and a set of templates is used to generate 25 ~ 1000 predictions. Multiple combinations of MSAs and templates are usually used to 750 - 300, 000 models generated for each target in total. In addition, the AlphaFold3 server was also used to generate hundreds (or thousands) of models for each monomer target.

The difference between the human predictors (MULTICOM and MULTICOM_human) and the server predictors (MULTICOM_AI, MULTICOM_GATE, and MULTICOM_LLM) is that the former generated many more models in the human prediction window on top of the models already generated by the latter in the server prediction window.

Model ranking

A set of quality assessment methods, including the global pLDDT score, the quality score predicted by GATE, GCPNet-EMA, EnQA, average pairwise similarity score and DeepRank3 methods (DeepRank3_Cluster, DeepRank3_SingleQA, DeepRank3_SingleQA_lite) are used to rank the AlphaFold2 and AlphaFold3 models generated above as follows.

- *MULTICOM_AI* ranked structural models using the global pLDDT score.
- *MULTICOM_GATE* ranked structural models based on the quality scores predicted by GATE.
- *MULTICOM_LLM* used the average score of GATE and global pLDDT to rank the structural models. It also used AlphaFold3 ranking scores to rank AlphaFold3 models.
- *MULTICOM_human* ranked structural predictions using the global pLDDT score. The ranking was subjected to human intervention.
- *MULTICOM* used AlphaFold3 ranking score, global pLDDT score, GATE, GCPNet-EMA, EnQA and the average pairwise TM-scores to rank structural models separately. The ranking was subjected to human intervention.

Different strategies were then applied to select the final 5 models for submission as follows.

- For *MULTICOM_AI* and *MULTICOM_human*, the no. 1 ranked model was automatically selected, with the remaining four chosen from top ranked models that have less than TM-score of 0.8 with the already selected models. If fewer than five models met this criterion, the remaining one were selected by the model ranking without considering the similarity between the selected models. If none of the Top 5 selected models came from AlphaFold3, the No. 5 model was replaced with a top-ranked AlphaFold3 model.
- *MULTICOM_GATE* used K-means clustering to group structural models into five clusters and selecting a top-ranked model from each cluster. If none of the Top 5 selected models came from AlphaFold3, the No. 5 model was replaced with a top-ranked AlphaFold3 model.
- For *MULTICOM_LLM*, the no. 1 model was no. 1 AlphaFold3 model selected by the AlphaFold3 ranking score. Its other 4 models were selected from the top-ranked models from model clusters.
- *MULTICOM*'s final 5 models were selected in sequential order from the no. 1 models ranked by the average of GATE and AlphaFold3 ranking scores, average of GATE and global pLDDT scores, GATE scores, GCPNet-EMA score, EnQA scores and average pairwise TM-scores, with identical models removed.

Tertiary structure prediction for monomer targets that are a part of a multimer target

The tertiary structural models for monomer targets that are a part of a multimer target were extracted from the top ranked structural models predicted for the multimer target. The method for multimer (quaternary) structure prediction is described in the MULTICOM QS abstract in the CASP16 abstract book.

2. Phase 2 monomer structure prediction

Only our two human predictors (MULTICOM and MULTICOM_human) participated in the Phase 2 monomer structure prediction. They did not generate additional models. Instead, they combined our in-house AlphaFold3 models, top MassiveFold models, and CASP16 server models together and then ranked and refined them as the final predictions.

Model collection and redundancy reduction. Our top 20 in-house AlphaFold3 models ranked by AlphaFold3 ranking score, top 10 MassiveFold models selected by the global pLDDT score, and CASP16 server models were collected and filtered to remove highly similar structures with TM-score equal to 1.

Model ranking. The models were evaluated and ranked using multiple quality assessment (QA) methods, including the global pLDDT, GATE, GCPNet-EMA, EnQA, average pairwise similarity score, and DeepRank3 methods.

Model combination. To further enhance the quality of the top-ranked predictions, each of the top-ranked model was combined with similar structures in the model pool with a TM-score higher than 0.95 to generate a refined model by Modeller. If the TM-score between each top ranked model and the refined model is ≥ 0.95 , the refined model is kept. Otherwise, only the top ranked model is used as a template to generate a refined model.

Final model selection. For *MULTICOM*, the models were ranked based on the average scores of pairwise GDT-scores, global pLDDT scores, GCPNet-EMA scores, DeepRank3_SingleQA, and DeepRank3_SingleQA_lite first. The in-house AlphaFold3 prediction most similar (e.g., with a TM-score higher than 0.95) to the top-ranked predictions was used as no. 1 model, while the remaining refined models were chosen according to the ranking.

For *MULTICOM_human*, the no. 1 model was selected according to the average score of GATE, GCPNet-EMA, and average pairwise IDDT scores. The remaining four models were selected sequentially from the no. 1 refined models ranked by the average score of three methods (GATE, GCPNet-EMA, and average pairwise IDDT scores), GATE, EnQA, DeepRank3_Cluster, GCPNet-EMA, and average pairwise TM-scores, with identical models removed.

Tertiary structure prediction for monomer targets that are a part of a multimer target

Similarly, as in Phase 1, the tertiary structural models for monomer targets that are a part of a multimer target were extracted from the top ranked Phase 2 structural models predicted for the multimer target. The method for Phase 1 multimer (quaternary) structure prediction is described in the *MULTICOM* QS abstract in the CASP16 abstract book.

Availability

We will make the source code of the *MULTICOM4* prediction system available at GitHub upon publication of our CASP16 results.

1. Liu J, Guo Z, Wu T, et al. Enhancing alphafold-multimer-based protein complex structure prediction with *MULTICOM* in CASP15. *Communications biology* 2023;6:1140.
2. Liu J, Guo Z, Wu T, et al. Improving AlphaFold2-based protein tertiary structure prediction with *MULTICOM* in CASP15. *Communications chemistry* 2023;6:188.
3. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *nature* 2021;596:583-589.
4. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *biorxiv* 2021:2021.2010. 2004.463034.
5. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024:1-3.
6. Morehead A, Liu J, Cheng J. Protein structure accuracy estimation using geometry - complete perceptron networks. *Protein Science* 2024;33:e4932.
7. Chen C, Chen X, Morehead A, et al. 3D-equivariant graph neural networks for protein model quality assessment. *Bioinformatics* 2023;39:btad030.
8. Liu J, Wu T, Guo Z, et al. Improving protein tertiary structure prediction by deep learning and distance

- prediction in CASP14. *Proteins: Structure, Function, and Bioinformatics* 2022;90:58-72.
- 9. Consortium U. UniProt: a worldwide hub of protein knowledge. *Nucleic acids research* 2019;47:D506-D515.
 - 10. Steinegger M, Söding J. Clustering huge protein sequence sets in linear time. *Nature communications* 2018;9:2542.
 - 11. Steinegger M, Mirdita M, Söding J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature methods* 2019;16:603-606.
 - 12. Mitchell AL, Almeida A, Beracochea M, et al. MGnify: the microbiome analysis resource in 2020. *Nucleic acids research* 2020;48:D570-D578.
 - 13. Nordberg H, Cantor M, Dusheyko S, et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic acids research* 2014;42:D26-D31.
 - 14. Wang Y, Shi Q, Yang P, et al. Fueling ab initio folding with marine metagenomics enables structure and function predictions of new protein families. *Genome biology* 2019;20:1-14.
 - 15. Yang P, Zheng W, Ning K, et al. Decoding the link of microbiome niches with homologous sequences enables accurately targeted protein structure prediction. *Proceedings of the National Academy of Sciences* 2021;118:e2110828118.
 - 16. Steinegger M, Meier M, Mirdita M, et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC bioinformatics* 2019;20:1-15.
 - 17. Remmert M, Biegert A, Hauser A, et al. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature methods* 2012;9:173-175.
 - 18. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC bioinformatics* 2010;11:1-8.
 - 19. Sussman JL, Lin D, Jiang J, et al. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography* 1998;54:1078-1084.

MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human
(QS)

Improving AlphaFold2/AlphaFold3-based Protein Quaternary Structure Prediction with Large-Scale Modeling Sampling and Ranking

Jian Liu, Pawan Neupane, Jianlin Cheng*

University of Missouri, Columbia, MO 65211, USA

*Corresponding author: chengji@missouri.edu

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3/AF2; EMA:Y; MD:N;

During CASP16, we used our MULTICOM4 system (an enhanced version of MULTICOM3^{1,2} built on top of AlphaFold2³ and AlphaFold-Multimer⁴) as well as AlphaFold3⁵ web server to generate a large number of quaternary structural models for multimer targets. The structural models were then ranked by various model quality assessment methods with our three server predictors (MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM) and two human predictors (MULTICOM and MULTICOM_human). The top ranked models were selected as final predictions.

Methods

Model Generation

During CASP16 Phase 0 and Phase 1 experiments, we used both AlphaFold2 and AlphaFold3 to generate structural models for a multimer target with predicted or known stoichiometry as follows.

MULTCOM4 pipeline for AlphaFold2-based model generation. The pipeline has the following four main steps:

(1) protein monomer structure generation for each chain of an assembly target (see the detailed description in the MULTICOM TS abstract in this abstract book)

(2) monomer alignments concatenation. The MSAs of the subunits of the multimer target are concatenated using potential protein-protein interactions extracted from the species information, UniProt accession IDs, the protein-protein interactions in the STRING database⁶, the protein complexes in the Protein Data Bank (PDB)⁷, and the DeepMSA2 pairing protocol, resulting in a series of MSAs for the multimer.

(3) monomer templates concatenation. The sequence of each subunit in the multimer target is searched against PDB70 and an inhouse complex template database built from PDB by HHsearch⁸ to identify the structural templates. The templates for each subunit are concatenated together if they share the same PDB code. Moreover, the predicted tertiary structural model of each unit of the multimer is searched against the inhouse structure template database by FoldSeek⁹ to identify more templates, which are concatenated as multimer templates.

(4) multimer structure generation. Each combination of the concatenated MSAs and templates is fed for the customized AlphaFold-Multimer to generate multimer structural predictions, resulting in 1000 ~ 670,000 structural models per target for most multimer targets except for very large targets for which fewer models can be generated within a short period of time.

AlphaFold3 model generation: AlphaFold3 web server was used to generate hundreds or thousands of structural models for each multimer target.

The same model generation protocol above was used in both the Phase 0 and Phase 1 prediction.

Phase 0 prediction for multimer targets with unknown stoichiometry information

MULTICOM4 was run to generate templates for each subunit (chain) of a Phase 0 multimer target. Based on the template information, a set of candidate stoichiometries was proposed. AlphaFold3 was then used to generate 25 to 100 structural models for each candidate stoichiometry. The average and highest AlphaFold3 ranking scores of the models for each candidate were calculated. The scores were used to select one or a few predicted stoichiometries for the target. If necessary, the stoichiometry information from templates and/or related literature was also used to predict the stoichiometry of the target.

Once the stoichiometry was predicted, the MULTICOM4 model generation pipeline and AlphaFold3 web server were used to generate many (e.g., hundreds/thousands of) structural models for each selected candidate stoichiometry. The main difference between our sever and human predictions was that a lot of more models were generated for human prediction during a long two/three-week prediction window than for server prediction during a short three-day prediction window.

After the models were generated, a variety of multimer quality assessment methods (e.g., AlphaFold2 confidence scores, AlphaFold3 ranking scores, graph transformer (GATE) predicted quality scores (see our MULTICOM QA abstract in the CASP16 abstract book for details), VoroMQA scores¹⁰, GCPNet-EMA quality scores¹¹, EnQA quality scores¹², and pairwise model similarity scores are used separately or together to rank and select models.

MULTICOM_AI server predictor used the AlphaFold confidence score to rank both AF2 and AF3 models. *MULTICOM_GATE* server predictor used the quality score predicted by GATE to rank and select models. *MULTICOM_LLM* server predictor used the average score of GATE and AlphaFold confidence score to rank the AlphaFold2 models. It used AlphaFold3 ranking scores to select AlphaFold3 models.

MULTICOM (a human predictor) used AlphaFold3 ranking score, AlphaFold confidence score, GATE, GCPNet-EMA, EnQA and the average pairwise TM-scores calculated by MMAlign to rank structural models separately. The ranking was subjected to human intervention.

MULTICOM_human (a human predictor) used the AlphaFold confidence score to rank models. The model ranking was subjected to human intervention.

To increase the diversity of the top 5 selected models for hard targets, *MULTICOM_AI* and *MULTICOM_human* automatically used the no.1 ranked model as final no. 1 model for submission, while choosing the remaining models based on both the ranking and their similarity with the already selected models (i.e., TM-score with the selected models < 0.8). If fewer than five models meet this structural similarity criterion, the remaining models are selected according to the ranking alone. *MULTICOM_GATE* grouped structural models into five distinct clusters using K-means clustering. The top-ranked structure from each cluster was then selected as one of the Top 5 submissions. For *MULTICOM_AI* and *MULTICOM_GATE*, if none of the Top 5 submissions is predicted by AlphaFold3, replace the No. 5 submission with a top-ranked AlphaFold3 predicted structure. *MULTICOM_LLM* used Top 1 AlphaFold3 predicted structure ranked by the AlphaFold3 ranking score as no. 1 model, while choosing the remaining four models from top-ranked models of different model clusters. *MULTICOM*'s final Top 5 models were selected in sequential order from predictions ranked by the average of GATE and AlphaFold3 ranking scores, average of GATE and AlphaFold confidence scores, GATE, GCPNet-EMA, EnQA, GATE without using AlphaFold features, and average pairwise TM-scores, with identical structural models removed. Finally, if a target is predicted to have multiple possible stoichiometries, models for the different stoichiometries were included into the final top-5 models with the no. 1 model selected from the most likely stoichiometry.

Phase 1 prediction for multimers with known stoichiometry

The same model generation and ranking methods used in the Phase 0 prediction were applied to the Phase 1 prediction for MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM and MULTICOM_human. The main difference between Phase 1 and Phase 0 is that the Phase 1 prediction combines the new models generated in Phase 1 and the models with the correct stoichiometry generated in Phase 0 if available together for model ranking and selection. The number of models for each target in Phase 1 is usually more than double that for the corresponding target in Phase 0.

Phase 2 prediction: selecting and combining models in in-house model pool, CASP16 model pool, and MassiveFold model pool

Our two human predictors (MULTICOM and MULTICOM_human) participated in the Phase 2 prediction. They did not generate new models for most targets. Instead, they combined our in-house AlphaFold3 models, CASP16 structural models downloaded from the CASP16 website, and top ranked MassiveFold models together and applied different methods to rank and refine them to generate final five models. The protocols are described as follows.

Model collection and redundancy reduction. Our top 20 in-house AlphaFold3 models ranked by AlphaFold3 ranking score, top 10 MassiveFold models selected by AlphaFold confidence score, and CASP16 Phase 1 server models were collected. The highly similar models (TM-score = 1) were filtered out.

Initial model ranking. The filtered structural predictions were evaluated using multiple quality assessment methods to select the top-ranked predictions, including the GATE quality scores, GCPNet-EMA scores, EnQA scores, average pairwise similarity scores, global pLDDT scores, and VoroMQA scores. For *MULTICOM*, models were ranked based on the average scores of pairwise TM-scores, EnQA scores, VoroMQA scores, GCPNet-EMA scores. For *MULTICOM_human*, models were ranked using GATE quality scores, GCPNet-EMA scores, EnQA scores, average pairwise similarity scores, global pLDDT scores, and VoroMQA scores, separately.

Model combination. To further enhance the quality of the top-ranked predictions, each of the top-ranked model was combined with similar structures in the model pool with a TM-score higher than 0.95 to generate a refined model by Modeller. If the TM-score between each top ranked model and the refined model is ≥ 0.95 , the refined model is kept. Otherwise, only the top ranked model is used as a template to generate a refined model.

Final model selection. *MULTICOM* used an in-house AlphaFold3 model most similar (e.g., with a TM-score higher than 0.95) to the top-ranked prediction as no. 1 model, while choosing the remaining four refined models from the model pool according to the ranking generated above.

MULTICOM_human selected five models sequentially from the no. 1 refined models selected by the six quality assessment methods (average pairwise TM-scores, global pLDDT scores, EnQA scores, VoroMQA scores and the GATE scores), GATE, average pairwise TM-scores, VoroMQA scores, GCPNet-EMA scores and EnQA scores), with identical models removed.

Availability

We will make the source code of the MULTICOM4 prediction system available at GitHub upon publication of our CASP16 results.

1. Liu J, Guo Z, Wu T, et al. Enhancing alphafold-multimer-based protein complex structure prediction with MULTICOM in CASP15. *Communications biology* 2023;6:1140.
2. Liu J, Guo Z, Wu T, et al. Improving AlphaFold2-based protein tertiary structure prediction with MULTICOM in CASP15. *Communications chemistry* 2023;6:188.
3. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *nature* 2021;596:583-589.
4. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *biorxiv* 2021:2021.2010. 2004.463034.
5. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024:1-3.
6. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research* 2021;49:D605-D612.
7. Sussman JL, Lin D, Jiang J, et al. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography* 1998;54:1078-1084.
8. Steinegger M, Meier M, Mirdita M, et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC bioinformatics* 2019;20:1-15.
9. Van Kempen M, Kim SS, Tumescheit C, et al. Fast and accurate protein structure search with Foldseek. *Nature biotechnology* 2024;42:243-246.
10. Olechnovič K, Venclovas Č. VoroIF - GNN: Voronoi tessellation - derived protein–protein interface assessment using a graph neural network. *Proteins: Structure, Function, and Bioinformatics* 2023;91:1879-1888.
11. Morehead A, Liu J, Cheng J. Protein structure accuracy estimation using geometry - complete perceptron networks. *Protein Science* 2024;33:e4932.
12. Chen C, Chen X, Morehead A, et al. 3D-equivariant graph neural networks for protein model quality assessment. *Bioinformatics* 2023;39:btad030.

MULTICOM_ligand (LG)

Geometric Deep Learning, Generative Modeling, and Structural Consensus Ranking for Protein-Ligand Structure Prediction and Binding Affinity Estimation

Alex Morehead, Jian Liu, Pawan Neupane, Jianlin Cheng*

University of Missouri, Columbia, MO 65211, USA

*Corresponding author: chengji@missouri.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; AF:AF3; EMA:Y; MD:Y;

Significant progress in protein-ligand structure prediction (PLSP) has recently been made using deep learning (DL) methods, with the new AlphaFold 3¹ method serving as a hallmark example. Such progress has the potential to reshape the landscape of modern drug discovery by allowing researchers to precisely design drugs for important protein targets with reduced off-target effects. Currently, open-source DL algorithms for PLSP are numerous yet disparate, leading to lack of clarity in which methods perform best in certain molecular contexts. Furthermore, it remains unclear how to best compare each method's predictions to each other to perform model ranking and how to accurately estimate the binding affinity of their top-ranked predicted (multi-ligand) structural complexes. In CASP16, we debuted PoseBench², our recent PLSP benchmark for DL methods, as a standalone predictor in the CASP ligand (LG) prediction category. PoseBench enables researchers to predict the structure of protein-ligand complexes using several of the most recent DL docking algorithms such as DiffDock-L³ and NeuralPLexer⁴ using a unified software pipeline and to rank-order each predicted structure using a structural consensus heuristic for ligand binding modes inspired by previous CASP ranking methods for protein structures⁵. Furthermore, our PoseBench pipeline for CASP16 introduces a new confidence and binding affinity estimation method called FlowDock, a geometric generative DL algorithm pre-trained on the Protein Data Bank⁶ and fine-tuned on the PDDBind 2020 dataset⁷ to predict scalar IDDT and binding affinity values given the (multi-chain) protein sequences, (multi-ligand) SMILES strings, and optional (predicted) structure of a protein-ligand complex. Combining these top-ranked predicted structures and estimated confidence and affinity measures, PoseBench lastly applies a series of molecular sanity checks to its top-ranked predictions using the PoseBusters software suite⁸ to optimize the structural validity of its top-5 structural models.

Methods

Our PoseBench protein-ligand structure prediction pipeline consists of five sequential steps: **(1) protein-ligand structure generation, (2) structural consensus ranking, (3) protein-specific ligand structure relaxation, (4) chemical and structural sanity checking, and (5) structure-aware confidence and binding affinity estimation**. For the nucleic acid-ligand targets presented in the CASP16 competition, we modified each step above to instead use AlphaFold 3 and RoseTTAFold-All-Atom⁹ in a combined manner to predict these complex structures, a procedure we describe in detail in the following sections. Notably, both the structure generation as well as the confidence and binding affinity estimation components of our pipeline are driven by various geometric and generative DL algorithms, with the majority of these DL methods using protein structures predicted by AlphaFold 3 as a starting point for molecular docking or structure sampling. In the following sections, we describe in detail how we employed our PoseBench pipeline in each of the various prediction contexts available in CASP16.

1. Context 1: Protein-(super)ligand structure prediction

Protein-ligand structure generation. For each protein-(super)ligand prediction target, the PoseBench pipeline first generates 100 protein-binding ligand conformations with DiffDock-L³, 40 with DynamicBind¹⁰ and NeuralPLexer⁴ each, and 1 with RoseTTAFold-All-Atom⁹ (RFAA). The first three of these methods are provided with a protein structure predicted by AlphaFold 3¹ to use an initial protein state for docking or structure generation, whereas RFAA relies on its own template search and multiple sequence alignment procedures to perform structure prediction for each target. If not already predicted by one of these methods, AlphaFold 3’s predicted confidence (i.e., IDDT) scores for each protein atom are reported for the resulting complexes.

Structural consensus ranking. The pipeline then (re)rank-orders each method’s (intrinsically rank-ordered) top-3 predicted (multi-)ligand structures according to their structural similarity to each other (i.e., inter-consistency) using average pairwise (multi-)ligand root-mean-squared deviation as a proxy for ligand structural model accuracy. Two important caveats included here are that (1) if a top consensus-ranked prediction contains any severe atomic steric clashes between protein residues, it is assigned the lowest consensus ranking; and (2) if a target contains multiple ligands, the structures produced by NeuralPLexer for these targets are automatically designated as our top-3 consensus-ranked predictions based on the results of our previous benchmarking study² for DL docking methods.

Ligand structure relaxation. The top-5 predictions resulting from this structural consensus ranking procedure then have their ligand structures relaxed using protein-fixed molecular dynamics simulations driven by OpenMM¹¹.

Structure sanity checking. The relaxed predictions replace their unrelaxed prediction counterparts in rank-ordering if and only if a relaxed ligand prediction passes more of the protein-ligand complex validity sanity checks available in the PoseBusters software suite⁸ than the unrelaxed prediction and if the relaxed ligand prediction is less than 5 Å in centroid distance from the unrelaxed prediction.

Confidence and binding affinity estimation. Once this top-5 ranking of predictions is finalized, the PoseBench pipeline runs our new FlowDock model for each top-5 prediction to estimate its average (multi-)ligand confidence (i.e., IDDT) scores and binding affinity values.

2. Context 2: Nucleic acid-ligand structure prediction

Structure generation and combination. For the nucleic acid-ligand prediction targets presented in CASP16, in contrast to our protein-(super)ligand structure prediction procedure above, we first used AlphaFold 3¹ and RFAA⁹ to predict the nucleic acid and nucleic acid-ligand complex structures of each target, respectively. For each of these targets, we then optimally aligned AlphaFold 3’s predicted nucleic acid structure onto RFAA’s predicted nucleic acid-ligand complex structure using the chain alignment feature available in Mol*¹². Once aligned, we then copied RFAA’s corresponding (multi-)ligand structures to form a new nucleic acid-ligand complex using AlphaFold 3’s (aligned) predicted nucleic acid structure. When these “combined” AlphaFold3-based complexes did not contain any severe nucleic acid-ligand steric clashes, we selected them as our top-1 predictions for such targets and then placed RFAA’s predictions as our next best predictions. Otherwise, if severe clashes were present, we instead ranked RFAA’s structures as our best predictions. In both cases, this procedure produced 2 predicted complexes for each nucleic acid-ligand target in CASP16.

Confidence and binding affinity estimation. Once the rankings of our top nucleic acid-ligand predictions were finalized, we then employed a surrogate protein structure (PDB ID: 5S8I) to temporarily replace our predicted nucleic acid structures for each complex to enable us to predict confidence (i.e., IDDT) scores and binding affinity values for each predicted (multi-)ligand structure using our new FlowDock model for protein-ligand complexes.

Context 3: Protein-(multi-)ligand binding affinity estimation

Single and multi-ligand binding affinity estimation for the experimental protein-ligand complex structures provided later in the CASP16 competition was performed similarly to Step 5 of our protein-(super)ligand structure prediction procedure, with an important caveat being that, in this context, we provided FlowDock with the available crystal structure of each protein-ligand complex instead of a structure predicted by a DL method. We note that, as a generative model of protein-ligand complexes, FlowDock could produce multiple distinct affinity predictions for each given experimental complex, though in this setting, for simplicity we chose to report the first predicted affinity value for each ligand.

Availability

We will make our PoseBench CASP16 pipeline available upon publication.

1. Abramson J, Adler J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024.
2. Morehead A, Giri N, et al. Deep Learning for Protein-Ligand Docking: Are We There Yet? *International Conference on Machine Learning (ICML) AI4Science Workshop* 2024.
3. Corso Gabriele, Deng A, et al. Deep confident steps to new pockets: Strategies for docking generalization. *International Conference on Learning Representations (ICLR)* 2024.
4. Qiao Z, Nie W, et al. State-specific protein-ligand complex structure prediction with a multiscale deep generative model. *Nature Machine Intelligence* 2024.
5. Roy RS, Liu J, et al. Combining pairwise structural similarity and deep learning interface contact prediction to estimate protein complex model accuracy in CASP15. *Proteins: Structure, Function, and Bioinformatics* 2023.
6. Berman HM, Westbrook J, et al. The Protein Data Bank. *Nucleic Acids Research* 2000.
7. Liu Z, Su M, et al. Forging the Basis for Developing Protein-Ligand Interaction Scoring Functions. *Accounts of Chemical Research* 2017.
8. Buttenschoen M, Garrett MM, et al. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science* 2024.
9. Krishna R, Wang J, et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 2024.
10. Lu W, Zhang J, et al. DynamicBind: Predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model. *Nature Communications* 2024.
11. Eastman P, Swails J, et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology* 2017.
12. Sehnal D, Bittrich S, et al. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research* 2021.

MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human,
MULTICOM_ligand (QA)

Protein Model Evaluation and Selection Using Graph Transformers and Multiple Quality Metrics in CASP16

Jian Liu, Pawan Neupane, Alex Morehead, Jianlin Cheng*

University of Missouri, Columbia, MO 65211, USA

*Corresponding author: chengji@missouri.edu

Key: Auto:Y; CASP_serv:Y; Templ:N; MSA:N; DeepL:Y; AF:AF3/AF2; EMA:Y; MD:N;

In CASP16, we tested different deep learning-based and traditional model quality assessment (QA) techniques as six predictors, including MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human and MULTICOM_ligand in the regular QA category and/or the MassiveFold model selection category.

Methods

1. MULTICOM Predictors for Regular QA

MULTICOM_GATE. Given a pool of models for a target, it first samples a subset of models from the pool either randomly or evenly from different model clusters. The subset of models is then used to build a model similarity graph, in which a node denotes a model and an edge connects two models that are similar to each other. A graph transformer for estimating model accuracy (called GATE) is used to predict the quality of each model in the graph. The process is repeated multiple time so that each model in the pool is selected at least once. The average predicted quality for a model used as its final quality score. The same approach is applied to both tertiary structural models and quaternary structural models except for some difference in node features. For tertiary structural models, the node features include average pairwise model similarity score, AlphaFold global (average) pLDDT score, model quality score predicted by a single-model quality assessment method based on equivariant graph neural networks EnQA¹, model quality score predicted by a single-model quality assessment method based on geometric-complete graph neural network GCPNet-EMA², and model quality score predicted by a consensus deep learning method DeepRank3³. Additionally, the quality scores used in DeepRank3 are also included as node features. For quaternary structural models, the node features include the average pairwise similarity scores, AlphaFold global pLDDT score, ICPS score⁴, DProQA score⁵, EnQA score, GCPNet-EMA score and VoroMQA⁶ scores. The pairwise similarity between two similar model is used as the feature of the edge connecting them.

MULTICOM_AI uses the single-model QA method - GCPNet-EMA to predict the global quality of a model.

MULTICOM_LLM uses the average pairwise similarity score between a model and all other models calculated by MMAlign as the predicted quality score.

MULTICOM uses the average of GCPNet-EMA score and GATE score as the predicted quality score.

MULTICOM_human uses the average of EnQA score, GCPNet-EMA score, and GATE score as the predicted quality score.

2. MULTICOM predictors for MassiveFold model selection

For the phase 2 model selection for monomer targets, **MULTICOM_AI** uses GCPNet-EMA, **MULTICOM_GATE** uses GATE, **MULTICOM_LLM** employs the average pairwise model similarity score calculated by TMscore, **MULTICOM** uses EnQA, **MULTICOM_human** utilizes the average of the pairwise IDDT score, GCPNet-EMA score, and GATE score, while **MULTICOM_ligand** employs the average of the GCPNet-EMA score and global pIDDT score.

For the phase 2 model selection for multimer targets, **MULTICOM_AI** employs GCPNet-EMA, **MULTICOM_GATE** utilizes GATE with additional AlphaFold confidence scores⁷ as input features, **MULTICOM_LLM** selects models using the average pairwise similarity scores calculated by MMAAlign, **MULTICOM** applies EnQA, **MULTICOM_human** utilizes the average of pairwise CAD-scores, GCPNet-EMA, and GATE, while **MULTICOM_ligand** uses the average of GCPNet-EMA and AlphaFold confidence scores.

It is worth noting that the implementation of our predictors above evolved during the CASP16 experiment. For the simplicity and clarity, only the final version of each predictor is described above.

Availability

The source code of our methods will be made available at GitHub upon publication of our CASP16 results.

1. Chen C, Chen X, Morehead A, et al. 3D-equivariant graph neural networks for protein model quality assessment. *Bioinformatics* 2023;39:btad030.
2. Morehead A, Liu J, Cheng J. Protein structure accuracy estimation using geometry - complete perceptron networks. *Protein Science* 2024;33:e4932.
3. Liu J, Wu T, Guo Z, et al. Improving protein tertiary structure prediction by deep learning and distance prediction in CASP14. *Proteins: Structure, Function, and Bioinformatics* 2022;90:58-72.
4. Roy RS, Liu J, Giri N, et al. Combining pairwise structural similarity and deep learning interface contact prediction to estimate protein complex model accuracy in CASP15. *Proteins: Structure, Function, and Bioinformatics* 2023;91:1889-1902.
5. Chen X, Morehead A, Liu J, et al. A gated graph transformer for protein complex structure quality assessment and its performance in CASP15. *Bioinformatics* 2023;39:i308-i317.
6. Olechnovič K, Venclovas Č. VoroIF - GNN: Voronoi tessellation - derived protein–protein interface assessment using a graph neural network. *Proteins: Structure, Function, and Bioinformatics* 2023;91:1879-1888.
7. Yu D, Chojnowski G, Rosenthal M, et al. AlphaPulldown—a python package for protein–protein interaction screens using AlphaFold-Multimer. *Bioinformatics* 2023;39:btac749.

MultiFOLD2

Automated Sampling, Quality Assessment and Refinement of Tertiary and Quaternary Structure Models using the MultiFOLD2 Server

L.J. McGuffin¹, A.G. Genc¹, S.N. Alhaddad¹, N.S. Edmunds¹ and R. Adiyaman¹

1 - School of Biological Sciences, University of Reading, Reading, UK

l.j.mcguffin@reading.ac.uk

Key: Auto:Y; CASP_serv:Y; TempL:Y; MSA:Y; DeepL:Y; AF:AF2; EMA:Y; MD:N

The MultiFOLD2 server is our new integrated pipeline for producing tertiary and quaternary structure models of proteins via sampling, quality scoring and refinement.

Methods

The MultiFOLD2 protocol has 3 main steps: sampling, scoring, and refinement. In brief, the main differences from the original MultiFOLD¹ version were the inclusion of stoichiometry prediction and improved sampling with the addition of AlphaFold2-Multimer including dropout², RosseTTAFold2³ with and without dropout, and RosseTTAFold-All-Atom⁴. Additionally, we improved scoring using the latest version of our ModFOLDdock⁵ method (see our ModFOLDdock2 server abstract for further details). Finally, the refinement step was the same as in the original MultiFOLD method. Different approaches were used based on prediction phase and target type.

For Phase 0 (T0/H0 targets), no stoichiometry information was provided for the multimers, so this was predicted by MultiFOLD2 as follows. Firstly, initial 3D models were generated using LocalColabFold⁶ version 1.5.2 with templates. If templates could not be found from the target sequences, then Foldseek⁷ was used with each chain of the top initial 3D model to find templates for each subunit. Stoichiometry for each template was then determined using PISA⁸. If templates were found for all subunits, then the most frequent stoichiometries from all templates were assigned to the target sequences and then used in subsequent modelling. If no templates could be found, QUEEN⁹ assigned stoichiometry directly from the target sequences.

Following stoichiometry prediction, in the first step, 3D models of tertiary and quaternary structures were built using two different versions of LocalColabFold⁶ (<https://github.com/YoshitakaMo/localcolabfold>). Firstly, LocalColabFold version 1.0.0 was run without relaxation for targets with total sequence lengths of <=2500 amino acids. LocalColabFold version 1.0.0 is based on the ColabFold/AlphaFold2_advanced notebook integrating the AlphaFold2¹⁰ (AF2) weights (alphafold2_ptm) and modified to produce models for multimers. Secondly, LocalColabFold version 1.5.2 was run with relaxation, with and without dropout for targets with lengths <=6000. LocalColabFold version 1.5.2 is based on the ColabFold/AlphaFold2_mmseqs2 notebook integrating the official AlphaFold2-Multimer¹¹ (AF2M) weights (alphafold2_multimer_v3) and is specifically tuned for multimer prediction. Additionally, RosseTTAFold2³ (RF2) was run for targets with lengths <=1800 with

all possible symmetric and non-symmetric configurations, both with and without dropout. The top 10 models were selected from the scores in the JSON files. Finally, RosseTTAFold-All-Atom⁴ (RFAA) was run generating 10 models for targets with lengths <=1300. Thus, in the first step of MultiFOLD2 up to 45 initial 3D models (5xAF2, 20xAF2M, 10xRF2, 10xRFAA) were generated.

In the second step of the process, the models were scored and ranked using ModFOLDdock2S, which is a single-model approach for producing both global and local (interface residue) quality scores for predicted quaternary structures (see our ModFOLDdock2 server abstract).

In the final step, the top 5 ModFOLDdock2S selected models were reformatted to mmCIF files using MAXIT¹² and then used as input templates for our AlphaFold2-Multimer_Refine (AF2MR) protocol¹³. The AF2MR approach used the LocalColabFold 1.5.2 method with the “--custom-template-path” option, with 12 recycles and relaxation for targets <=2000 residues or 3 recycles without relaxation for larger targets. For each model, the model rankings and predicted per-residue quality scores (pLDDT*100) from LocalColabFold were added to the B-factor column for each set of atom records.

For the very large complexes (>6000 residues), due to our limited GPU resources, we had to divide sequences up into overlapping fragments for submission to MultiFOLD2. The resulting modelled fragments were then manually assembled, using structural superposition in PyMOL (<https://www.pymol.org>), to form larger, more complete models.

For Phase 1 (T1/H1 targets), the stoichiometry information for each target was provided. If the stoichiometry was correctly predicted for the equivalent targets in Phase 0, then the top 5 McGuffin group Phase 0 models were returned (see our McGuffin group abstract). If the stoichiometry was incorrectly predicted at Phase 0, then the MultiFOLD2 protocol described above was rerun with the corrected stoichiometry.

For Phase 2 (T2/H2 targets), the stoichiometry information and models from all groups and MassiveFold¹⁴ were available. For the multimer targets, ModFOLDdock2R (see our ModFOLDdock2 server abstract) was used to select the top 5 ranked models from all groups. For the monomer (A1) targets a quick version of our ModFOLD9¹⁵ method (ModFOLD9Q) was used to select the top 5 ranked models from the MassiveFold set (see our McGuffin group abstract).

Results

MultiFOLD2 is continuously benchmarked using the CAMEO-BETA resource¹⁶ where it has been shown to significantly outperform other tested methods on multimer modelling according to Wilcoxon tests on common subsets (MultiFOLD2 versus MultiFOLD Oligo-IDDT scores, p=9.18E-05; versus Server76, p=9.69E-10; versus Server 994(AF3), p=0.01034).

Availability

The MultiFOLD server is available at:

https://www.reading.ac.uk/bioinf/MultiFOLD/MultiFOLD2_form.html

- 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. *Nucleic Acids Research*. **51**(W1), W274-W280. doi: 10.1093/nar/gkad297
- Wallner,B. (2023). Improved multimer prediction using massive sampling with AlphaFold in CASP15. *Proteins*. **91**, 1734-1746. doi: 10.1002/prot.26562
- Baek,M., Anishchenko,I., Humphreys,I.R., Cong,Q., Baker,D. & DiMaio,F. (2023). Efficient and accurate prediction of protein structure using RoseTTAFold2. *bioRxiv*. 2023.2005.2024.542179. doi: 10.1101/2023.05.24.542179
- Krishna,R., Wang,J., Ahern,W., Sturmels,P., Venkatesh,P., Kalvet,I., Lee, G. R., Morey-Burrows,F. S., Anishchenko,I., Humphreys,I.R., McHugh,R., Vafeados,D., Li,X., Sutherland,G.A., Hitchcock,A., Hunter,C.N., Kang,A., Brackenbrough,E., Bera,A.K.,...Baker,D. (2024). Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*. **384**(6693), eadl2528. doi: 10.1126/science.adl2528
- Edmunds,N.S., Alharbi,S.M.A., Genc,A.G., Adiyaman,R. & McGuffin,L.J. (2023) Estimation of Model Accuracy in CASP15 Using the ModFOLDdock Server. *Proteins*. **91**, 1871-1878. doi: 10.1002/prot.26532
- Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. & Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nature Methods*. **19**(6), 679–682. doi: 10.1038/s41592-022-01488-1
- van Kempen,M., Kim,S.S., Tumescheit,C. et al. (2024) Fast and accurate protein structure search with Foldseek. *Nat Biotechnol*. **42**, 243–246. doi: 10.1038/s41587-023-01773-0
- Krissinel,E. & Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-797. doi: 10.1016/j.jmb.2007.05.022
- Avraham,O., Tsaban,T., Ben-Aharon,Z. et al. (2023) Protein language models can capture protein quaternary state. *BMC Bioinformatics*. **24**, 433. doi:10.1186/s12859-023-05549-w
- Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., ... Hassabis,D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**(7873), 583–589. doi: 10.1038/s41586-021-03819-2
- Evans,R., O'Neill,M., Pritzel,A., Antropova,N., Senior,A., Green,T., Žídek,A., Bates,R., Blackwell,S., Yim,J., Ronneberger,O., Bodenstein,S., Zielinski,M., Bridgland,A., Potapenko,A., Cowie,A., Tunyasuvunakool,K., Jain,R., Clancy,E., ... Hassabis,D. (2021) Protein complex prediction with AlphaFold-Multimer. *In bioRxiv*. doi: 10.1101/2021.10.04.463034
- Adams,P.D., Afonine,P.V., Baskaran,K., Berman,H.M., Berrisford,J., Bricogne,G., Brown,D.G., Burley,S.K., Chen,M., Feng,Z., Flensburg,C., Gutmanas,A., Hoch,J.C., Ikegawa,Y., Kengaku,Y., Krissinel,E., Kurisu,G., Liang,Y., Liebschner,D., ... Young,J.Y. (2019) Announcing mandatory submission of PDBx/mmCIF format files for crystallographic depositions to the Protein Data Bank (PDB). *Acta Crystallographica Section D: Structural Biology*. **75**(Pt 4), 451–454. doi: 10.1107/S2059798319004522
- Adiyaman,R., Edmunds,N.S., Genc,A.G., Alharbi,S.M.A. & McGuffin,L.J. (2023) Improvement of protein tertiary and quaternary structure predictions using the ReFOLD refinement method and the AlphaFold2 recycling process. *Bioinformatics Advances*. vbad078. doi: 10.1093/bioadv/vbad078
- Brysbaert,G., Raouraoua,N., Mirabello,C. et al. (2024) MassiveFold: unveiling AlphaFold's hidden potential with optimized and parallelized massive sampling. PREPRINT (Version 1) available at Research Square. doi: 10.21203/rs.3.rs-4319486/v1

15. McGuffin,L.J. & Alharbi,S.M.A. (2024) ModFOLD9: a web server for independent estimates of 3D protein model quality. *J. Mol. Biol.* **436**, 168531. doi:10.1016/j.jmb.2024.168531
16. Leemann,M., Sagasta,S., Eberhardt,J., Schwede,T., Robin,X. & Durairaj,J. (2023) Automated benchmarking of combined protein structure and ligand conformation prediction. *Proteins*. **91** 1912–1924. doi: 10.1002/prot.26605

Low-Rank Adaptation for Folding tools

Tatsuya Kobayashi and Yusuke Kobayashi

NEC Corporation

t.kobayashi07@nec.com and kobayashi-yusuke@nec.com

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:N; MD:N

In this proposal, we attempt to provide a folding model suitable for a specific superfamily based on the AlphaFold2(*1) pipeline. To achieve this, we apply LoRA (Low-Rank Adaptation) to the attention layer of the AlphaFold2 learning model and input the PDB data set of a specific superfamily into the modified learning model. This allows us to develop an AlphaFold2 folding model and a trained model suitable for the specific superfamily. This abstract presents the specific procedures targeting T1300 and T1200.

Methods

In our method, we applied LoRA to the Attention Layer of AlphaFold2. This implementation will be made available after the CASP16 Conference, as mentioned in the Availability section. Using the modified AlphaFold2 with LoRA, we input the PDB data set of the superfamily to which T1300 and T1200 belong and created a trained model suitable for predicting the protein structures of this superfamily. This trained model is used to predict the structures of T1300 and T1200. The PDB data set of the superfamily to which T1300 and T1200 belong was obtained by searching for the amino acid sequence ZLBT-C, common to T1300 and T1200, in the SuperFamily database provided by InterPro(*2) at EMBL-EBI.

Availability

Our method and code will be available at the GitLab.

1. Jumper, J. et al. “Highly accurate protein structure prediction with AlphaFold.” *Nature*, 596, pages 583–589 (2021). DOI: 10.1038/s41586-021-03819-2.
2. InterPro database, EMBL-EBI. Available at: <https://www.ebi.ac.uk/interpro>.

NKRNA-s

Protein-Nucleic Acid Complex Structure Prediction Using Deep Learning and Language Models

Wentao Ni¹, Qiqige Wuyun², Gang Hu¹, and Wei Zheng^{1,3,4}

1 - School of Statistics and Data Science, Nankai University, 94 Weijin Road, 300071, Tianjin, China

2 - Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

3 - Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

4 - Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

zhengwei@umich.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:N; MD:N

The NKRNA-s server group participated in CASP16 with the aim of modeling nucleic acids (DNA or RNA), nucleic acid – nucleic acid complexes, protein – protein complexes, and protein – nucleic acid complexes. For nucleic acid-related targets, predictions were made using our newly developed deep learning method, DeepProtNA, which integrates pre-trained language model embeddings, multiple sequence alignments (MSAs), predicted secondary structures, and structural templates as inputs for modified Evoformer blocks. Protein – protein complex targets were predicted using a modified version of DMFold-Multimer¹.

Methods

Protein-protein complex modeling pipeline

Our protein–protein complex structures were predicted using a modified DMFold-Multimer approach, which involves three primary steps.

First, MSA construction builds upon the previous DeepMSA2 pipeline. Compared to the version used in CASP15, this updated DeepMSA introduces two key improvements: (i) a larger in-house metagenomic sequence database that incorporates data from IMG/M, NCBI, and EBI; and (ii) a multi-domain MSA assembly method that merges domain-level MSAs into full chain-level MSAs. Similar to DeepMSA2, the new pipeline contains three MSA construction sub-pipelines: dMSA, qMSA, and mMSA. These sub-pipelines are iteratively used to collect homologous sequences from genomic and metagenomic databases, including UniClust30², UniRef90³, Metaclust⁴, Mgnify⁵, BFD⁶ and an in-house huge metagenomics database. To speed up the search, the in-house database is clustered with a 30% sequence identity cutoff using MMseqs2⁷. MSAs generated from these sub-pipelines are input into AlphaFold2 (1-embedding) to predict a set of models. These MSAs are then ranked by their associated pLDDT scores from AlphaFold2. For multi-domain targets, the same MSA generation method is used to construct domain-level MSAs based on predicted domain boundaries, which are then assembled into full-length MSAs by linking sequences from the same species. The ranked MSAs are either directly used in protein homomer modeling or paired as multimer MSAs for protein heteromer modeling. For heteromeric complexes, an additional selection process generates an optimal set of paired MSAs by combining individual constituent MSAs. The top N ranked MSAs for each constituent protein are chosen to form potential paired MSAs, and for a

heteromeric complex with M constituent proteins, N^M distinct paired MSAs are generated and evaluated based on a combined score of the depth of the MSAs and pLDDT score of the monomer chains. To ensure the pipeline completes within three days, N is selected to satisfy $N^M \leq 64$.

Second, template detection is based on a new version of LOMETS (LOMETS4). Compared to LOMETS3⁸, which was used in CASP15, the major update in LOMETS4 is its ability to handle protein complexes. For protein heteromers, templates are identified as follows: first, homologous templates for each constituent chain in the target complex are identified using LOMETS3, which includes six profile-based threading methods, five contact/distance-based threading methods, three protein language model-based threading methods, and one structure-based threading method. Notably, templates for individual chains that have already been considered in previous steps are excluded to prevent the similar query constituent chain from hitting identical templates. The templates are ranked by quality (e.g., Z-score). Finally, if at least two constituent chains share templates from the same protein complex and have a high sum Z-score, these complexes are considered potential templates.

Third, structure model generation utilizes a modified AlphaFold2 modeling engine. The MSAs from the first step and the structure templates from the second step serve as input features for this modeling engine. Key modifications to the AlphaFold2 modeling engine include: (i) using templates or not, (ii) adjusting the dropout rate, (iii) applying different versions of AlphaFold2 pre-trained weights (v1-v3), (iv) generating a higher number of decoys than the default setting (25 models), (v) applying or omitting the early stop strategy in AlphaFold2 (v2.3), and (vi) extending the modeling iterations. The final models are ranked by confidence scores (e.g., $0.8pTM + 0.2ipTM$).

For targets with unknown stoichiometry, we developed a new method to determine stoichiometry information. This method uses three pipelines: one derives stoichiometry from top-ranked LOMETS4 templates; the second predicts stoichiometry using a deep learning method combining sequence data and embeddings from protein language models⁹ and text-based language models using UniProt descriptions; and the third pipeline predicts directly from DMFold confidence scores, where the oligomeric state is predicted based on the model with the highest confidence.

Nucleic acid-related target modeling pipeline

DeepProtNA is an end-to-end deep learning algorithm designed for predicting protein-nucleic acid complex and nucleic acid complex structures. The method integrates pre-trained language model embeddings, multiple sequence alignment information, predicted secondary structure and structural templates to directly generate three-dimensional coordinates of the complexes from input sequences through a set of modified Evoformer blocks and a structure module similar in AlphaFold2¹⁰. Protein and RNA sequences are input respectively into the pre-trained language models, ESM⁹ and RNA-FM¹¹, to generate high-dimensional sequence embeddings. MSA is generated for both protein and nucleic acid sequences using the modified version DeepMSA2¹ and rMSA¹². The MSA-derived features are combined with the language model embeddings to enhance the understanding of sequence interactions. Structural templates are selected from PDB by LOMETS4 for protein or BLASTn¹³ for RNA. Additionally, the predicted secondary structure of nucleic acids is also selected as inputs of the deep learning networks. The core architecture of DeepProtNA processes embeddings for proteins and nucleic acids, utilizing self-attention mechanisms to capture long-range relationships within individual sequences. To model interactions between protein and nucleic acid sequences, cross-attention mechanisms are employed, enabling the model to concentrate on key interaction sites. A structural decoder network uses these embedded representations to generate the three-dimensional coordinates of the complex. This network translates sequence information directly into spatial coordinates for each residue and nucleotide, resulting in the final structure. DeepProtNA also provides a confidence score that assesses the reliability of the

predicted interactions, which is additionally used to predict oligomeric states for RNA targets with unknown stoichiometry.

1. Zheng,W., Wuyun,Q., Li,Y., Zhang,C., Freddolino,P.L. & Zhang,Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods* 21, 279-289.
2. Mirdita,M., von den Driesch,L., Galiez,C., Martin,M.J., Söding,J. & Steinegger,M. (2017). UniClust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Research* 45, D170-D176.
3. Suzek,B.E., Wang,Y., Huang,H., McGarvey,P.B., Wu,C.H. & UniProt,C. (2015). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* (Oxford, England) 31, 926-932.
4. Steinegger,M. & Söding,J. (2018). Clustering huge protein sequence sets in linear time. *Nature Communications* 9, 2542.
5. Mitchell,A.L., Almeida,A., Beracochea,M., Boland,M., Burgin,J., Cochrane,G., Crusoe,M.R., Kale,V., Potter,S.C., Richardson,L.J., Sakharova,E., Scheremetjew,M., Korobeynikov,A., Shlemov,A., Kunyavskaya,O., Lapidus,A. & Finn,R.D. (2020). MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Research* 48, D570-D578.
6. Steinegger,M., Mirdita,M. & Söding,J. (2019). Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods* 16, 603-606.
7. Steinegger,M. & Söding,J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* 35, 1026-1028.
8. Zheng,W., Wuyun,Q., Zhou,X., Li,Y., Freddolino,P.L. & Zhang,Y. (2022). LOMETS3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic Acids Research* 50, W454-W464.
9. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., dos Santos Costa,A., Fazel-Zarandi,M., Sercu,T., Candido,S. & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123-1130.
10. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., Back,T., Petersen,S., Reiman,D., Clancy,E., Zielinski,M., Steinegger,M., Pacholska,M., Berghammer,T., Bodenstein,S., Silver,D., Vinyals,O., Senior,A.W., Kavukcuoglu,K., Kohli,P. & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
11. Chen,J., Hu,Z., Sun,S., Tan,Q., Wang,Y., Yu,Q., Zong,L., Hong,L., Xiao,J., Shen,T., King,I. & Li,Y. (2022). Interpretable RNA Foundation Model from Unannotated Data for Highly Accurate RNA Structure and Function Predictions. *bioRxiv*, 2022.08.06.503062.
12. Zhang,C., Zhang,Y. & Pyle,A.M. (2023). rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *Journal of Molecular Biology* 435, 167904.
13. Zhang,Z., Schwartz,S., Wagner,L. & Miller,W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology* 7, 203-214.

OpenComplex-2: A Biomolecular Foundation Model

Mingliang Zeng^{1,#}, Jingcheng Yu^{1,#}, Zhaoming Chen¹, Huang He¹, Xi Chen¹, Yaqing Zhang¹, Kai Chen¹, and Qiwei Ye^{1,#}

I - BAAI, # - Equal Contribution.

Chivee.ye@gmail.com

Key: Auto:Y; CASP_serv:Y; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:N; EMA:Y; MD:Y

OpenComplex-2 is a state-of-the-art biomolecular foundation model designed to predict the structures of proteins, DNA, RNA, small molecules, and their complexes. By leveraging graph-based molecular representations and advanced generative techniques, such as diffusion models, OpenComplex-2 surpasses traditional structural prediction approaches and enables the prediction of **ensemble behaviors** of molecular complexes. This capability allows the model to estimate probabilistic distributions of conformations, offering enhanced potential for downstream tasks (e.g., binding affinity prediction, mutation prediction). With an all-atom representation, OpenComplex-2 excels at capturing dynamics of flexible regions within biological macromolecules. A detailed technical report on OpenComplex-2 is forthcoming.

Methods

The training of OpenComplex-2 utilizes all available Protein Data Bank (PDB) structural data up to 2023, encompassing DNA, RNA, proteins, small molecules, and ions. A visual overview of the model architecture is provided in Figure 1. The training process consists of two phases:

Pre-training phase: Each molecule is represented as a molecular graph, encoding atomic-level information. The model undergoes diffusion-based pre-training to capture spatial relationships between atomic pairs, allowing it to learn possible structural configurations.

Fine-tuning phase: In this phase, additional information, such as multiple sequence alignments⁶ (MSA) and binding affinities, is incorporated to further refine the model for specific tasks.

CASP Variation

Server Group: For the Server group, we selected results based on the model's highest confidence score. These results underwent molecular dynamics (MD) relaxation⁷ to ensure no atomic clashes.

Human Group: The Human group saw several methodological adjustments based on the server version: 1) We conducted more extensive sampling, approximately 10 times that of the Server group. 2) A novel sampling selection strategy was employed, based on the cluster centroids of the sampling distribution, and refined using bond-level details. 3) For certain tasks, we performed manual cross-validation to select samples that aligned more closely with the literature, without intervening in the model input.

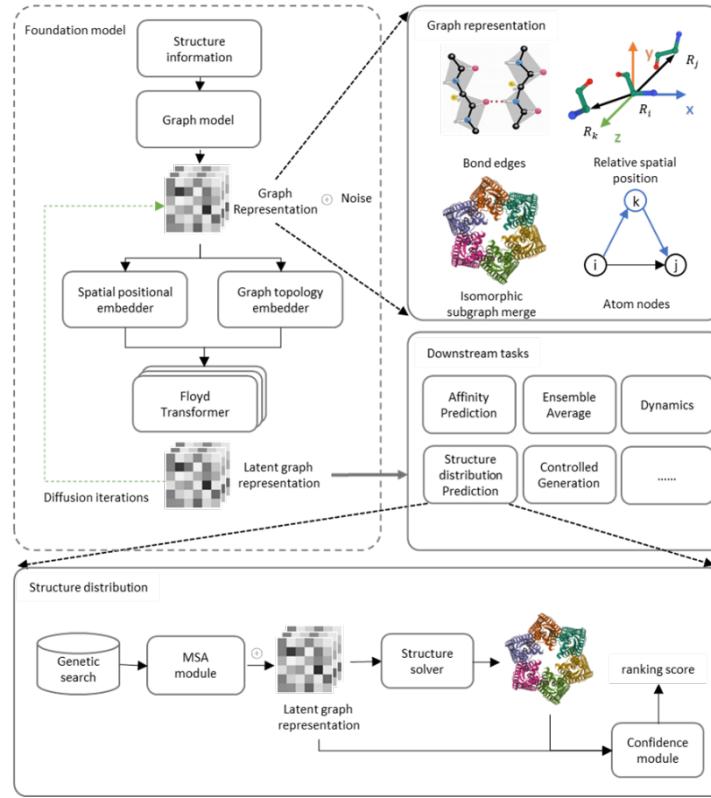


Figure 1. Overview of OpenComplex-2 architecture and training flow. OpenComplex-2 consists of two training stages: in the pre-training stage, the model focuses on structural relationships between atoms; in the fine-tuning stage, it incorporates additional information for specific tasks.

Results

We employed a single OpenComplex-2 model to predict all tasks for CASP16. The model's ability to understand biomolecular structures spans multiple dimensions, particularly in terms of ensemble prediction: structural ensemble, docking position ensemble, and stoichiometry ensemble.

- **Structural Ensemble:** OpenComplex-2 demonstrates an exceptional ability to predict secondary structures of proteins and nucleic acids without relying on templates. It then folds these into tertiary structures, depending on the availability of MSA information (Figure 2A). This allows the model to account for conformational flexibility and structural variance, especially in flexible regions like protein loops and RNA.
- **Complex Ensemble:** OpenComplex-2 excels at identifying intermolecular interaction sites, accommodating diverse docking configurations. For instance, in the case of the L1000 series protein, which contains a loop at its binding pocket, OpenComplex-2 successfully predicts the small molecule binding site and provides plausible ligand positions based on structural variation (Figure 2B). Furthermore, due to the fine-tuning phase, OpenComplex-2 can directly predict the binding affinity associated with multiple ligand distributions.
- **Stoichiometry Ensemble:** Due to no dependency on fixed number of chains during modeling (by merging all isomorphic subgraphs to one), OpenComplex-2 gains the ability to predict

stoichiometry directly (without finetuning). The predicted number of chains is extracted from the latent graph representation by the structure solver. Out of 31 predictions related to protein stoichiometry, OpenComplex-2 made correct predictions in 28 cases (Figure 2C). This underscores the model’s capability to predict complex macromolecular interactions and assemblies.

For the structure prediction track, OpenComplex-2 generated predictions that reflect the possible distributions of molecular structures and their complexes. The submission to CASP included the sampling cluster centroids of these predicted distributions, providing a robust representation of the most likely configurations.

For the SuperTarget track, we further fine-tuned the model using the LBA dataset. In this track, we modified the selection strategy for submitted predictions by incorporating an ensemble of scoring functions, including OpenComplex-2’s binding affinity, Cyscore², LEGO⁴, and Schrödinger Glide’s Gscore³. This multi-faceted approach allowed us to enhance the accuracy and confidence of our binding predictions.

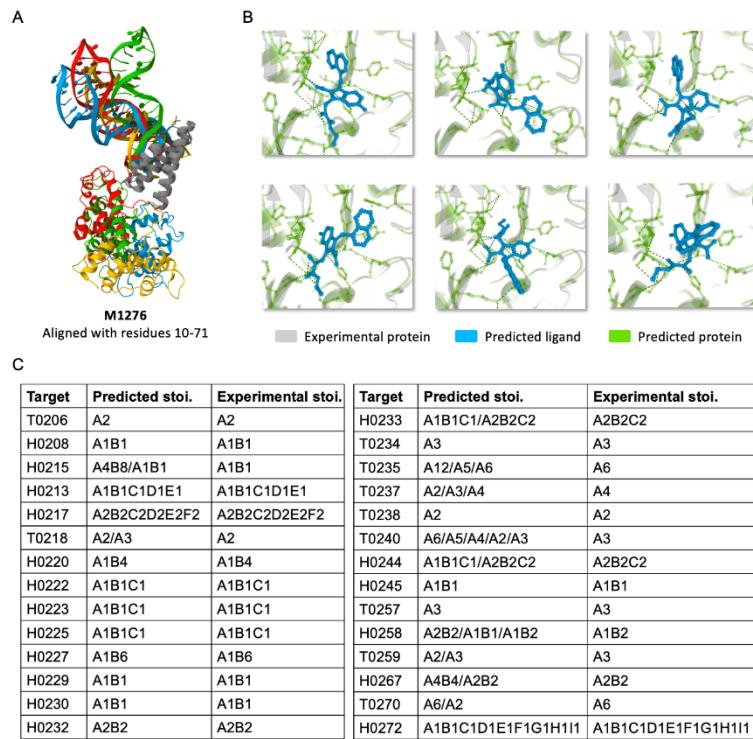


Figure 2. A illustrative example of of OpenComplex-2’s results across different tasks. A) shows structural ensemble predictions of M1276. B) demonstrates small molecule docking ensemble of L1000 series. C) illustrates predictions related to protein stoichiometry.

Availability

The OpenComplex-2 code and model will be released soon at <https://github.com/baaihealth/OpenComplex>. Please stay tuned.

1. Abramson, Josh, et al. "Accurate structure prediction of biomolecular interactions with AlphaFold 3." *Nature* (2024): 1-3.
2. Improved protein-ligand binding affinity prediction by using a curvature dependent surface area model. Yang Cao and Lei Li, *Bioinformatics*, 30(12):1674-1680, 2014
3. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., "Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening", *J. Med. Chem.*, 2004, 47, 1750–1759
4. Yuancheng S U N, Chen K, Liu K, et al. 3D Molecular Pretraining via Localized Geometric Generation[J].
5. Ye Q, Chen X, Huang H, et al. Sable: Bridging the Gap in Protein Structure Understanding with an Empowering and Versatile Pre-training Paradigm[J]. 2024.
6. Hauser M, Steinegger M, Söding J. MMseqs software suite for fast and deep clustering and searching of large protein sequence sets[J]. *Bioinformatics*, 2016, 32(9): 1323-1330.
7. Case D A, Darden T A, Cheatham T E, et al. Amber 10[J]. 2008.
8. Townshend, R. J., Vögele, M., Suriana, P., Derry, A., Powers, A., Laloudakis, Y., ... & Dror, R. O. (2020). Atom3d: Tasks on molecules in three dimensions. *arXiv preprint arXiv:2012.04035*.

AlphaFlow: flow-based generative models from single-structure predictors

Bowen Jing

Massachusetts Institute of Technology, Cambridge, MA, USA

bjing@mit.edu

Key: *Auto:N; CASP_serve:N; Templ:N; MSA:Y.MetaG; DeepL:Y; AF:Y; EMA:N; MD:N*

We submit predictions for the protein conformational ensemble targets from a derivative version of AlphaFold2 fine-tuned under a generative modeling objective, as described in previously published work.¹ These predictions demonstrate significant variability that aims to capture the structural heterogeneity of the protein target in solution.

Methods

AlphaFlow is a generative model of protein structures conditioned on sequence and MSA information. It is designed for a distributional modeling setting, i.e., to predict a conformational ensemble for a given protein sequence, rather than the single-structure prediction paradigm adopted by AlphaFold2 and successor models. Specifically, a deep denoising network parameterizes a continuous flow over protein structure space that transports a noisy prior distribution to the conformational ensemble for a given sequence; this network is trained under the flow-matching generative modeling paradigm.² The deep denoising network is heavily based on the neural network architecture of AlphaFold2,³ except that the template input is replaced with the noisy input; the output structure prediction is regarded as the direction of the denoising flow. The denoising network is initialized with the pretrained weights of AlphaFold2 and fine-tuned on the PDB for 1.3M training examples with OpenFold.⁴

Availability

The source code and parameters of the model are readily accessible at <https://github.com/bjing2016/alphaflow>.

1. Jing, B., Berger, B. and Jaakkola, T., AlphaFold Meets Flow Matching for Generating Protein Ensembles. In Forty-first International Conference on Machine Learning.
2. Lipman, Y., Chen, R.T., Ben-Hamu, H., Nickel, M. and Le, M., Flow Matching for Generative Modeling. In The Eleventh International Conference on Learning Representations.
3. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. and Bridgland, A., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), pp.583-589.
4. Ahdritz, G., Bouatta, N., Floristean, C., Kadyan, S., Xia, Q., Gerecke, W., O'Donnell, T.J., Berenberg, D., Fisk, I., Zanichelli, N. and Zhang, B., 2024. OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods*, pp.1-11.

Completion of the solvation shell of 7ez0 based on MD simulations at cryo-EM temperatures

Anja Henning-Knechtel¹, Naleem Nawavi¹ Pascal Auffinger²

¹Chemistry Program, Science Division, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates;

²Université de Strasbourg, Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, 2 Allée Konrad Roentgen, 67084 Strasbourg, France.

p.auffinger@ibmc-cnrs.unistra.fr

Key: Auto:N; CASP_serv:N; Temp:N; MSA:N; DeepL:N; EMA:N; MD:Y.

This CASP R1260 task is devoted to modeling the solvation shell of a full-length Apo L-21 ScaI *Tetrahymena* ribozyme structure which was deposited in the PDB in 2021 with the PDBid:7ez0 (1). Our approach primarily relied on prior-knowledge from several publications of our group, along with molecular dynamics (MD) simulations (2-7).

We opted to generate a single structure that includes both the frozen cryo-EM RNA structure and its solvation shell. This structure was derived from MD simulations performed at cryo-EM temperatures (77 K). We believe such a model offers a good alignment with experimental data. We assumed that the RNA particles used to generate the cryo-EM structure were already sorted and relatively homogeneous, which is especially true for the new ribozyme structure, solved at 2.3 Å (CASP target). Based on stereochemical insights, we placed several ions in unassigned density spots visible in the 7ez0 map. The MD simulations were particularly valuable for positioning water molecules, most of which displayed excellent stereochemical characteristics. The solvation shell extends to 5 Å around the RNA as requested.

Note: due to time constraints, the model we submitted is not as complete or fully refined as we initially planned.

Methods

Several decisions were made to model the RNA solvation shell, all based on the 7ez0 structure. First, we inspected the structure using MolProbity and Coot tools. Overall, the structural modeling appeared to us to be of good quality and only a few changes were made, primarily in regions with poor atomic B-factors that are excluded from the CASP quality assessment process.

Next, we examined the 27 assigned Mg²⁺ ions. Their positions were found to be compatible with octahedral coordination, as determined using a custom Coot script written by Dr. Naleem Nawavi, one of our team members. However, based on previous studies, we reassigned two ions, (ions 506 and 507), as Na⁺ instead of Mg²⁺. Additional 36 Mg²⁺ ions were placed at unassigned density peak locations and we checked their potential for octahedral coordination. Around 50 Na⁺ ions were placed at lower-density peak locations based on stereochemical intuition. We tentatively assigned one Cl⁻ ion in our model that contacts the amino group of G91.

Once this partial model was constructed, we used the Coot script written by Dr. Naleem Nawavi, to place the water molecules required to complete the Mg²⁺ ion octahedral coordination shells. This model was then subjected to MD simulations following the protocol described below.

We performed the MD simulations using the GROMACS 2028.8 package with the HB-CUFIX force field (8). The model structure was placed in a 16.5 x 16.5 x 16.5 nm box containing 141,947 water molecules, along with 145 Mg²⁺, 356 Na⁺ and 260 Cl⁻ ions to approximate the experimental conditions. The 50 mM HEPES molecules were not considered in our simulation.

To equilibrate the system, we chose to freeze the structure of the atoms present in the 7ez0 PDB structure by using the appropriate options in the Gromacs program. We also froze in a first step the water molecules that are part of the cryo-EM Mg²⁺ ion solvation shells. The additional ions and water molecules were allowed to move freely. Given limitations related to the Gromacs “freeze group” option, we used an NVT thermodynamic ensemble and ensured the absence of any volume artifacts in the simulation box.

After equilibration, we generated 4 ns of MD trajectories at room temperature to further equilibrate the solvent particles. This was followed by an additional 2 ns of MD trajectories at cryo-EM temperatures (77 K). At this point, we assumed the solvent particles were equilibrated and had settled into positions compared to those in the experimental structure. Based on these MD trajectory at cryo-EM temperatures, we proceeded to generate atomic densities for the solvent particle. This was accomplished with a custom MDAnalysis script developed by Dr. Anja Henning-Knechtel, which averages the snapshots from the cryo-EM temperature MD simulations to produce the final densities. Water molecules were then placed into the density peaks using the Phenix ”DOUSE” tool. After a final inspection and adjustment of certain solvation sites, and due to time constraints, we decided to submit this model to CASP.

Results

The final model consists in a single structure comprising 79 Mg²⁺ ions, 50 Na⁺ ions, 1 Cl⁻ ion and 6665 water molecules.

We are indebted to Prof. Serdal Kirmizialtin for discussions and support in this undertaking.

1. Su,Z., Zhang,K., Kappel,K., Li,S., Palo,M.Z., Pintilie,G.D., Rangan,R., Luo,B., Wei,Y., Das,R. & Chiu,W. (2021) Cryo-EM structures of full-length *Tetrahymena* ribozyme at 3.1 Å resolution. *Nature*, **596**, 603-7.
2. Auffinger,P., Louise-May,S. & Westhof,E. (1999) Molecular dynamics simulations of the solvated yeast tRNA(Asp). *Biophys. J.*, **76**, 50-64.
3. Auffinger,P. & Westhof,E. (2001) Water and ion binding around r(UpA)₁₂ and d(TpA)₁₂ oligomers - Comparison with RNA and DNA (CpG)₁₂ duplexes. *J. Mol. Biol.*, **305**, 1057-72.
4. Auffinger,P. & Westhof,E. (2000) Water and ion binding around RNA and DNA (C,G)-oligomers. *J. Mol. Biol.*, **300**, 1113-31.
5. Leonarski,F., D'Ascenzo,L. & Auffinger,P. (2019) Nucleobase carbonyl groups are poor Mg²⁺ inner-sphere binders but excellent monovalent ion binders - A critical PDB survey. *RNA*, **25**, 173-92.
6. Leonarski,F., D'Ascenzo,L. & Auffinger, P. (2017) Mg²⁺ ions: do they bind to nucleobase nitrogens?

Nucleic Acids Res., **45**, 987-1004.

7. Leonarski,F., Henning-Knechtel,A., Kirmizialtin,S., Ennifar,E. & Auffinger,P. (2024) Principles of ion binding to RNA inferred from the analysis of a 1.55 Å resolution bacterial ribosome structure – Part I: Mg²⁺. *Nucleic acids Res.*, **in press**.
8. He,W., Naleem,N., Kleiman,D. & Kirmizialtin,S. (2022) Refining the RNA force field with small-angle X-ray scattering of helix-junction-helix RNA. *J. Phys. Chem. Lett.*, **13**, 3400-8.

MELD-RNA: RNA structure prediction from ambiguous data sets

J. Gaza, R. Esmaeeli and A. Perez

University of Florida

perez@chem.ufl.edu

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N.MetaG; DeepL:N; EMA:N; MD:Y

We previously used MELD to predict the structures of single-stranded RNA sequences using server predicted structures as initial templates, and base pair distances as restraints. For one of the CASP 15 targets, our method surprisingly provided the best prediction among all participants. In CASP 16, we applied the same methodology to predict the structures of 14 single-stranded RNA sequences. Our criteria for choosing the targets were based on the limitations of the web servers we used for the initial structures and base pairing predictions. All of our simulations were performed using the HiPerGator supercomputer at the University of Florida.

Methods

MELD (Modeling Employing Limited Data) is an enhanced sampling method that uses a Bayesian approach to combine physics-based models (i.e. molecular dynamics force fields) with ambiguous data sets.¹ For proteins, we have previously used NMR chemical shifts², cryo-EM data sets³, and core insights from secondary structure predictions⁴ to predict protein structures. These data sets are transformed to spatial information such as distance and dihedral restraints for the MD simulations. During the simulations, MELD monitors whether the restraints are satisfied, applying energy penalties if they are not. Thus, this approach converts the rugged energy landscape of the simulations into a more focused one, prioritizing conformations that are compatible with the provided data sets. Due to the ambiguous nature of the data sets, different interpretations of the data can arise to different minima. To escape these minima and sample all possible interpretations, MELD employs a Hamiltonian and Temperature Replica Exchange (H, T-REMD) protocol, where the strength of the restraints is zero in the higher-temperature replicas and gradually increases, reaching full strength in the middle replica, and remains constant throughout the lowest-temperature replicas. The simulations were performed for 1.5 μ s using 30 replicas with a temperature range of 300 to 550 K. We used GBneck2 for the implicit solvent⁵, and the OL3 force field for the RNA.⁶

For the RNA targets, we used RNAfold⁷ and RNAComposer⁸ predictions of the secondary structures as the data set. We converted these predictions as distance restraints between the predicted base pairs, and satisfied only 80% of these restraints during the simulations. In some targets, restraints that correspond to tertiary structures (i.e. pseudoknot predictions from IPknot⁹ and ProbKnot¹⁰) were also implemented. For all the targets, our simulations started with template structures using webserver predicted structures. Most of the targets used three sources for the templates: (1) RNAComposer⁸, (2) trRosettaRNA¹¹, and (3) AlphaFold3.¹² These templates were then distributed periodically along the 30 replica ladder.

After the simulations, we performed hierarchical clustering on the first five replicas using an epsilon value of 5.0 Å. We then visualized the top clusters, and submitted the structures with the best base pairing

quality. Along with coordinates, we also reported the RMS fluctuations within the clusters. These values were provided in the temperature column of the PDB files.

Availability

MELD is available as a plugin for OpenMM. All MD engines and force fields are available and free to use. Trajectories and clusters are available upon request from the authors. The MELD plugin can be accessed at: <https://github.com/maccallumlab/meld>

1. MacCallum,J., Perez,A., & Dill,K. (2015). Determining protein structures by combining semireliable data with atomistic physical models by Bayesian inference. *Proceedings of the National Academy of Sciences*, 112(22), 6985–6990.
2. Mondal,A., Swapna,G., Lopez,M., Klang,L., Hao,J., Ma,L., Roth,M., Montelione,G., & Perez,A. (2023). Structure Determination of Challenging Protein–Peptide Complexes Combining NMR Chemical Shift Data and Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 63(7), 2058–2072.
3. Chang,L., Mondal,A., MacCallum,J., & Perez,A. (2023). CryoFold 2.0: Cryo-EM Structure Determination with MELD. *The Journal of Physical Chemistry A*, 127(17), 3906–3913.
4. Perez,A., Morrone,J., Brini,E., MacCallum,J., & Dill,K. (2016). Blind protein structure prediction using accelerated free-energy simulations. *Science Advances*, 2(11).
5. Nguyen,H., Roe,D., & Simmerling,C. (2013). Improved Generalized Born Solvent Model Parameters for Protein Simulations. *Journal of Chemical Theory and Computation*, 9(4), 2020–2034.
6. Zgarbová,M., Otyepka,M., Šponer,J., Mládek,A., Banáš,P., Cheatham,T., & Jurečka,P. (2011). Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *Journal of Chemical Theory and Computation*, 7(9), 2886–2902.
7. Gruber,A., Lorenz,R., Bernhart,S., Neubock,R., & Hofacker,I. (2008). The Vienna RNA Websuite. *Nucleic Acids Research*, 36(Web Server), W70–W74.
8. Popenda,M., Szachniuk,M., Antczak,M., Purzycka,K., Lukasiak,P., Bartol,N., Blazewicz,J., & Adamiak,R. (2012). Automated 3D structure composition for large RNAs. *Nucleic Acids Research*, 40(14), e112–e112.
9. Sato,K., Kato,Y., Hamada,M., Akutsu,T., & Asai,K. (2011). IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics*, 27(13), i85–i93.
10. Bellaousov,S., & Mathews,D. (2010). ProbKnot: Fast prediction of RNA secondary structure including pseudoknots. *RNA*, 16(10), 1870–1880.
11. Wang,W., Feng,C., Han,R., Wang,Z., Ye,L., Du,Z., Wei,H., Zhang,F., Peng,Z., & Yang,J. (2023). trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nature Communications*, 14(1).
12. Abramson,J., Adler,J., Dunger,J., Evans,R., Green,T., Pritzel,A., Ronneberger,O., Willmore,L., Ballard,A., Bambrick,J., Bodenstein,S., Evans,D., Hung,C.C., O'Neill,M., Reiman,D., Tunyasuvunakool,K., Wu,Z., Žemgulytė,A., Arvaniti,E., Beattie,C., Bertolli,O., Bridgland,A., Cherepanov,A., Congreve,M., Cowen-Rivers,A., Cowie,A., Figurnov,M., Fuchs,F., Gladman,H., Jain,R., Khan,Y., Low,C., Perlin,K., Potapenko,A., Savy,P., Singh,S., Stecula,A., Thillaisundaram,A., Tong,C., Yakneen,S., Zhong,E., Zielinski,M., Žídek,A., Bapst,V., Kohli,P., Jaderberg,M., Hassabis,D., & Jumper,J. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500.

Improving protein and RNA structure prediction via sampling of AlphaFold and MELD simulations

L. Chang, B. Singh, J.T. Gaza, Q. Liu, N. R. De Silva, A. Perez

Chemistry Department and Quantum Theory Project, University of Florida

perez@chem.ufl.edu

Key Protein: Auto:N; CASP_serv:N; Templ:N; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:Y

Our group is engaged in various aspects of method development for biomolecular simulations. In the current CASP, we focused on structure prediction for single-domain proteins, following the pipeline discussed here.

For protein structure prediction, we rely on the following hypotheses: (1) AlphaFold2 (AF2) has learned a type of "biophysical scoring function" that can accurately rank structure predictions¹⁻²; (2) tweaking AF2's input can guide the model towards different regions of its encoded conformational space. This approach has been successfully applied to predict more accurate protein complex structures in previous CASP challenge³ and to even identify folding intermediate states that resemble experimental observations⁴. In essence, we first sample structures from AF models using different inputs and then select the structure with the highest prediction confidence. The accuracy of these predictions heavily depends on the quality of the multiple sequence alignments (MSAs), with novel MSA generation pipelines significantly improving prediction accuracy⁵. However, regions with poor MSA coverage tend to be predicted less accurately. In such cases, MELD offers the potential to sample alternative conformations using a physics-based approach to further refine the structures⁶⁻⁷. We applied this pipeline to 57 protein sequences from CASP16, performing all calculations on the HiPerGator supercomputer at the University of Florida.

Method

Our pipeline starts by providing three different inputs to AF2: (1) the query sequence with full and clustered MSAs (generated from ColabFold and DeepMSA2 options), (2) running each protocol both with and without template information (if available), and (3) utilizing AlphaFold2-ptm, AlphaFold2-multimer models⁸, and DeepFold⁹, running 50 recycles with 10 randomly generated seeds. We then select the top 10–20 predictions based on their pLDDT scores. If any of the pLDDT scores exceed 95, the pipeline concludes, and the top 5 structures are submitted. If not, we proceed with further refinement to enhance the quality of the predictions.

We have three possible strategies to improve predictions by using the previous model as a template for the next round of predictions. Our first approach examines per-residue pLDDT scores to identify regions

from independent AF2 predictions that exhibit locally high pLDDT scores. We then create "Frankenstein templates" by stitching together high pLDDT regions from different predictions, which are subsequently provided to AF2 as templates for further structure refinement.

The second approach uses the AF2 iterative approach from each of the best templates. We run 50 iterations under five different recycling conditions (0, 1, 2, 3, 4), generating 250 predictions for each original template. These predictions are then ranked by their pLDDT scores.

If neither of the previous two methods leads to significant improvement, we turn to MELD to enhance conformational sampling. MELD employs a physics-based approach combined with ambiguous and noisy data. In this case, we begin with the template structures and enforce contacts between residue pairs with pLDDT scores above 85. The simulations are clustered, and the centroids of these clusters are used as templates for AF2. Finally, we select the highest pLDDT-scoring structures.

Through these pipelines, we aim to sample more broadly the “biophysical energy function” that AF2 has learned, in search of higher pLDDT score structures.

Availability

MELD is available as a plugin for OpenMM. All molecular dynamics (MD) engines and force fields are freely accessible. Trajectories and cluster data can be requested from the authors. The MELD plugin can be found at: <https://github.com/maccallumlab/meld>.

1. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *nature*, 596(7873), 583-589.
2. Roney, J. P., & Ovchinnikov, S. (2022). State-of-the-art estimation of protein model accuracy using AlphaFold. *Physical Review Letters*, 129(23), 238101.
3. Wallner, B. (2023). AFsample: improving multimer prediction with AlphaFold using massive sampling. *Bioinformatics*, 39(9), btad573.
4. Chang, L., & Perez, A. (2024). AlphaFold2 knows some protein folding principles. *bioRxiv*, 2024-08.
5. Zheng, W., Wuyun, Q., Li, Y., Zhang, C., Freddolino, P. L., & Zhang, Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods*, 21(2), 279-289.
6. MacCallum, J. L., Perez, A., & Dill, K. A. (2015). Determining protein structures by combining semireliable data with atomistic physical models by Bayesian inference. *Proceedings of the National Academy of Sciences*, 112(22), 6985-6990.
7. Perez, A., MacCallum, J. L., & Dill, K. A. (2015). Accelerating molecular simulations of proteins using Bayesian inference on weak information. *Proceedings of the National Academy of Sciences*, 112(38), 11846-11851.
8. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., ... & Hassabis, D. (2021). Protein complex prediction with AlphaFold-Multimer. *biorxiv*, 2021-10.
9. Lee, J. W., Won, J. H., Jeon, S., Choo, Y., Yeon, Y., Oh, J. S., ... & Joo, K. (2023). DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function. *Bioinformatics*, 39(12), btad712.

PEZYFoldings

PEZYFoldings pipeline in CASP16

Toshiyuki Oda and Hitoshi Ishikawa

PEZY Computing K.K.

oda@pezy.co.jp

Key: *Auto:N; CASP_serv:N; Templ:Y; MSA:YMetaG; DeepL:Y; AF:AF2; EMA:Y; MD: Y*

Our team, PEZYFoldings, demonstrated a remarkable performance at CASP15¹. However, it was clear that there was still room for improvement². In light of what we learned after the competition, we made strategic updates to our pipeline for CASP16.

Methods

Our pipeline used in CASP16 was largely similar to that used in CASP15². It included extended sequence similarity searches, structure prediction using AlphaFold2³ (AF2) or AlphaFold-Multimer⁴ (AFM), refinement using a fine-tuned version of AFM⁵, and relaxation using OpenMM⁶. Due to space limitations, we will focus on the novel aspects introduced in CASP16 that we consider particularly important.

	Identity Threshold for Sequence Filtering	Template Search Method	Gap-Rich Region Masking	MSA Construction Method
A	95%*	hmmsearch ⁷	No	default
B	95%	hmmsearch	Yes	default
C	95%	hmmsearch	Yes	plm_based
D	95%	plm_based	No	default
E	95%	plm_based	Yes	default
F	95%	plm_based	Yes	plm_based
G	None	hmmsearch	No	default
I	None	hmmsearch	Yes	default
J	None	hmmsearch	Yes	plm_based
K	None	plm_based	No	default
L	None	plm_based	Yes	default
M	None	plm_based	Yes	plm_based

Table 1. Input feature construction methods. *The sequence filtering was done with hhfilter⁸.

Sequence similarity search: Sequence similarity searches were performed using the same tools and databases that we used in CASP15, but with more recent versions. In addition, we extended our approach to include searches against Sequence Read Archive⁹ (SRA) entries when initial searches did not yield sufficient hits. We selected SRA entries according to the results of Pebblescout¹⁰ or taxonomic information obtained from literature searches via Google¹¹ or NCBI Entrez⁹. SRA reads were searched directly using blastn, tblastn or tblast¹². Blastp¹² or PSI-BLASTeX¹³ searches were also performed against translated reads using prodigal¹⁴.

Multiple Sequence Alignment (MSA) Construction: In addition to the default MSA, we extended our strategy by implementing a protein-language-model (PLM)-representation-based sequence alignment tool¹⁵ using sequence representations generated by ESM2-650M-UR50D¹⁶. Notably, this PLM representation-based aligner was also used to search for templates and generate pairwise alignments between query and template.

Post-processing of MSAs: Prior to input feature construction, we filtered MSAs with a simple script to avoid highly skewed MSA depths, one of the problems we mentioned in our CASP15 paper. In addition, we created a simple script to mask gap-rich regions in MSAs that were assumed to be low-quality alignments. As a result, we constructed several sorts of input features (Table 1).

Model building: We primarily used AFM v2.3 for model building. When time and computational resources allowed, we also used AF2 and AFM v2.2. We also used RoseTTAFold All-Atom¹⁷ for T1276 MODEL 5.

Model refinement and rescoring: Our model selection step remained identical to the one we used in CASP15. The selected models were processed with our refiner. In cases where the refined structure had a higher sum of per-residue pLDDTs higher than 70 (we now refer to this metric as 'sum of pLDDTs > 70' for brevity) than the input structure, the refined structure was used. Otherwise, the pLDDTs of the input structures were modified using the pLDDTs of the refined structure, and the refined structures were discarded. For monomeric targets, the rankings were updated based on these adjusted scores. For multimeric targets, the ranking was revised if the sum of pLDDTs > 70 of a lower ranked structure exceeded that of a higher ranked structure by a factor of 1.1 or greater.

Large targets: For targets too large to process with our system, we broke input features into small pieces and randomly picked up the pieces and modelled partial structures. We then assembled them, taking into account interacting residue pairs of partially modelled structures. Model selection was performed based on sum of pLDDTs > 70. The assembled structures were again split into small pieces and processed with our refiner. The refined substructures were reassembled and assembled structure was treated as a 'refined structure' described in the 'Model refinement and rescoring' section.

Manual intervention: While our CASP15 paper emphasized the minimization of manual intervention, the process described in the 'Large targets' section was in many cases too costly and required manual interventions similar to those used in CASP15. It included visual inspection using PyMOL¹⁸ v0.99rc6, identification of interacting regions in large complexes, construction of partial models and their subsequent concatenation. In addition, where CASP organizers provided additional information on targets (e.g. filaments), a model was selected to match this information.

It is important to note that our methodology was adaptively refined for individual targets throughout the competition period. As a result, the specific processing details varied among targets, with each being treated according to its unique characteristics and challenges. We will provide detailed information on the processing of specific targets upon request from interested readers.

Acknowledgements

We thank all the members of PEZY Computing K.K. for managing the supercomputing system. We also express our gratitude to the National Institute of Genetics and RIKEN for supporting our metagenomic projects and PZLAST^{19,20}, which is a key component in implementing our extended sequence similarity search methodology. We are grateful to Google DeepMind for AlphaFold2 and AlphaFold-Multimer, and to DeLano Scientific, LLC and Schrödinger, LLC for PyMOL. We sincerely acknowledge other developers and maintainers for providing useful tools and databases under generous licenses. Finally, we appreciate the assistance of Claude 3.5 and DeepL for their insightful suggestions to improve our English text.

1. Simpkin, A.J. et al. (2023) Tertiary structure assessment at CASP15. *Proteins* 91, 1616-1635.
2. Oda, T. (2023) Improving protein structure prediction with extended sequence similarity searches and deep-learning-based refinement in CASP15. *Proteins* 91, 1712-1723.
3. Jumper, J. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
4. Evans, R. et al. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034.
5. Oda, T. (2023) Refinement of AlphaFold-Multimer structures with single sequence input. *bioRxiv*, 2022.2012.2027.521991.
6. Eastman, P. et al. (2024) OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. *J Phys Chem B* 128, 109-116.
7. Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* 7, e1002195.
8. Steinegger, M. et al. (2019) HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* 20, 473.
9. Sayers, E.W. et al. (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res* 50, D20-D26.
10. Shiryev, S.A. & Agarwala, R. (2024) Indexing and searching petabase-scale nucleotide resources. *Nat Methods* 21, 994-1002.
11. <https://www.google.com/>
12. Camacho, C. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.
13. Oda, T., Lim, K. & Tomii, K. (2017) Simple adjustment of the sequence weight algorithm remarkably enhances PSI-BLAST performance. *BMC Bioinformatics* 18, 288.
14. Hyatt, D. et al. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119.
15. https://github.com/yamule/matrix_align
16. Rives, A. et al. (2021) Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc Natl Acad Sci U S A* 118.
17. Krishna, R. et al. (2024) Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 384, eadl2528.
18. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.2r3pre.
19. Mori, H. et al. (2021) PZLAST: an ultra-fast amino acid sequence similarity search server against public metagenomes. *Bioinformatics*.
20. Ishikawa, H. et al. (2022) PZLAST: an ultra-fast sequence similarity search tool implemented on a MIMD processor. *International Journal of Networking and Computing* 12, 446-466.

Epitope Finding Protocol with AlphaFold-Multimer

Toshiyuki Oda and Hitoshi Ishikawa

PEZY Computing K.K.

oda@pezy.co.jp

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:YMetaG; DeepL:Y; AF:AF2; EMA:Y; MD: Y

Antibody-antigen complex prediction presented a significant challenge, with CASP15 results demonstrating reduced performance of predictions¹ for these targets. In response to this challenge, we developed a specialized pipeline tailored for antibody-antigen complexes. The summary of our core pipeline is described in the PEZYFoldings abstract without parentheses in this issue, hereafter referred to as 'our main abstract'.

Methods

Our methodology consisted of the following steps:

1. **Sequence Collection:** The sequence collection procedures were identical to our main abstract.
2. **Initial Structure Prediction:** We performed structure predictions for antibodies and antigens separately with AlphaFold-Multimer² (AFM) v2.3. The features named B and G in Table 1 of our main abstract were used. The top two models, selected based on criteria consistent with our core pipeline, were passed to subsequent stages of the pipeline.
3. **Strategic feature cropping:** A critical difference between our core pipeline and this specialized pipeline was the strategic cropping of input features based on the initial predicted structures: For antigens, a surface residue identified by HSSP³ was randomly selected along with surrounding surface residues and non-surface residues that interact with them. For antibodies, only the residues of the variable domain were selected for extraction. The input features of these selected residues were concatenated and stored. This process was repeated until all antigen surface residues were covered. The stored partial features were passed to the next step.
4. **Partial complex structure prediction:** Using the partial features constructed in step 3, we predicted partial complex structures using AFM v2.3. Our motivation for this approach is based on the assumption that restricting the search space could improve prediction and that AFM (and AlphaFold2⁵) trained on cropped data would have robust performance in predicting partial structures. When computational resources allowed, AFM v2.2 or AFM-based refinement tool¹⁴ was also used. The AFM-based refinement tool was applied to extracted residues in step 3 to obtain refined structures.
5. **Model selection:** The predicted structures were ranked according to one of the self-confidence metrics of AFM, iptm, between antibody and antigen. Structures with iptm values above a certain threshold were selected. We typically set this threshold between 0.4 and 0.7, adjusting it to balance the number of selected models with our available time and computational resources. Then TM-scores⁶ between predicted models were calculated using US-align⁷ or MM-align⁸ and highly similar models were filtered out to remove redundancy. If the antigen was too small, DockQ⁹ was used. Note that prior to calculating the TM-score or DockQ score, full length structures were constructed using a simple structure alignment technique.
6. **Second stage residue selection and structure prediction:** We performed a second residue selection process focusing on the interfaces present in the selected structures from step 5. This was done

under the assumption that the epitope found in step 5 might be incomplete and a complete epitope might improve prediction accuracy. As in steps 3 and 4, features of selected antigen residues and antibody variable domains were concatenated and used to predict partial complex structures.

7. **Second stage model selection:** Full-length structures were constructed as in step 5. We then selected models from both the initial and this second prediction stage, again using iptom between antibody and antigen as the primary criterion, with filtering to remove redundancy using US-align, MM-align or DockQ.

8. **Final refinement and rescoreing:** These structures underwent a final round of refinement and rescoreing, following the same procedure described in our main abstract, except that we used per-residue pLDDTs only at the interface between antibody and antigen instead of whole structure pLDDTs. In cases where the sum of interface pLDDT greater than 70 of a lower ranked structure exceeded that of a higher ranked structure by a factor of 1.1 or greater, we updated the rankings.

9. **Relax:** The same as our main pipeline, the selected structures were processed with OpenMM¹⁰ before submission.

It is important to note that there are various targets in the competition (e.g. antigen of H1222, H1223, and H1225 was relatively small compared with ordinary targets), therefore we modified some details of the protocol. Detailed information on the processing of specific targets upon request from interested readers.

Acknowledgements

We appreciate the assistance of Claude 3.5 and DeepL for their insightful suggestions to improve our English text.

1. Ozden, B., Kryshtafovych, A. & Karaca, E. (2023) The impact of AI-based modeling on the accuracy of protein assembly prediction: Insights from CASP15. *Proteins* 91, 1636-1657.
2. Evans, R. et al. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034.
3. Joosten, R.P. et al. (2011) A series of PDB related databases for everyday needs. *Nucleic Acids Res* 39, D411-419.
4. Oda, T. (2023) Refinement of AlphaFold-Multimer structures with single sequence input. *bioRxiv*, 2022.2012.2027.521991.
5. Jumper, J. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
6. Zhang, Y. & Skolnick, J. (2004) Scoring function for automated assessment of protein structure template quality. *Proteins* 57, 702-710.
7. Zhang, C., Shine, M., Pyle, A.M. & Zhang, Y. (2022) US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nat Methods* 19, 1109-1115.
8. Mukherjee, S. & Zhang, Y. (2009) MM-align: a quick algorithm for aligning multiple-chain protein complex structures using iterative dynamic programming. *Nucleic Acids Res* 37, e83.
9. Basu, S. & Wallner, B. (2016) DockQ: A Quality Measure for Protein-Protein Docking Models. *PLoS One* 11, e0161879.
10. Eastman, P. et al. (2024) OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. *J Phys Chem B* 128, 109-116.

PEZYFoldings (Extended)

PEZYFoldings Extended Pipeline for CASP16

Toshiyuki Oda and Hitoshi Ishikawa

PEZY Computing K.K.

oda@pezy.co.jp

Key: *Auto:N; CASP_serv:N; Templ:Y; MSA:YMetaG; DeepL:Y; AF:AF2; EMA:Y; MD: Y*

PEZYFoldings team, who showed outstanding performance in CASP15¹, joined CASP16 with updated pipeline. The summary of our core pipeline is described in the PEZYFoldings abstract without parentheses in this issue, hereafter referred to as 'our main abstract'. In this abstract, we introduce the outline of our additional pipeline for challenges newly introduced in CASP16 with restricted or additional resources: Phase 0, MODEL 6 in Phase 1, and Phase 2.

Methods

Please refer to the official CASP16 website for detailed information on the challenges.

Phase 0: For predictions where stoichiometry information was not available, we inferred stoichiometry based on search results from the RCSB PDB² website or PLM-representation-based similarity searches (see our main abstract) against our template database. The remaining procedures were the same as the standard pipeline, but used a limited number of input features (B and G in Table 1 of our main abstract).

MODEL 6 in Phase 1: We re-aligned the ColabFold³ baseline Multiple Sequence Alignments (MSAs) using the PLM-representation-based sequence aligner (see our main abstract). These MSAs were used to predict models with AlphaFold-Multimer⁴ v2.3 without templates, generating 5 models per target (one model per parameter). The top-ranked model was then relaxed using OpenMM⁵ prior to submission.

Phase 2: We applied the model selection step described in our main abstract to the combined pools of MassiveFold⁶ and our own models. We also extracted monomers from all models to construct monomer pools. Where structurally similar monomers had a higher sum of pLDDTs > 70, the corresponding subunits of the selected models were replaced with the monomers. The refinement and rescoring procedures described in our main abstract were applied, followed by relaxation using OpenMM⁵ prior to submission.

It is important to note that our methodology was adaptively refined for individual targets throughout the competition period. As a result, the specific processing details varied among targets, with each being treated according to its unique characteristics and challenges. We will provide detailed information on the processing of specific targets upon request from interested readers.

Acknowledgements

We appreciate the assistance of Claude 3.5 and DeepL for their insightful suggestions to improve our English text.

1. Simpkin, A.J. et al. (2023) Tertiary structure assessment at CASP15. Proteins 91, 1616-1635.
2. Berman, H.M. et al. (2000) The Protein Data Bank. Nucleic Acids Res 28, 235-242.
3. Mirdita, M. et al. (2022) ColabFold: making protein folding accessible to all. Nat Methods 19, 679-682.

4. Evans, R. et al. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034.
5. Eastman, P. et al. (2024) OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials. *J Phys Chem B* 128, 109-116.
6. Brysbaert, G. et al. (2024) MassiveFold: unveiling AlphaFold's hidden potential with optimized and parallelized massive sampling.

Graph Neural Network Ensemble Prediction of Protein Ligand Complex

Isaac Pfaender

University College London

Isaac.pfaender.22@ucl.ac.uk

Key: *Auto:N; CASP_serv:N; Templ:N; MSA:N.MetaG; DeepL:Y; EMA:N; MD:N*

Protein-ligand bound complex structure prediction is crucial to drug design, enabling more robust in silico screening of molecules and development cost reductions. Here in CASP16, we explore using graph-based approaches to predict the ligand coordinates within a folded structure.

Methods

The structure prediction pipeline consists of the following steps: querying PDB for the folded protein structure and a bound ligand, generating a conformation of the target ligand, converting the three structures into a graphical representation which feeds into 3 separate graph neural networks which then get processed by fully connected neural network layers with the final prediction output. However, the model size was limited due to hardware constraints.

Results

A prediction was made on L5001. A future direction could include scaling model size, incorporating energy functions for refinement and training on full protein-ligand complexes rather than isolated ligand coordinates within a predetermined structure.

Predict Conformation Ensembles of Protein Structures with PFSC-PFVM Approach

Jiaan Yang^{1,2}, Wen Xiang Cheng¹ and Peng Zhang¹

¹ - Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China,

² - Micro Biotech, Ltd., Shanghai, 200123, China

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; EMA:N; MD:N

The PFSC-PFVM (Protein Folding Shape Code¹ - Protein Folding Variation Matrix²) approach in protein structure fingerprint technology is a single sequence method to predict the conformational ensembles of protein structures. First, a set of PFSC is established to cover full folding space of five free connected points for five amino acid residues. Then, a database, 5AAPFSC, is established to collect all possible folding patterns in PFSC letters for any combination of five amino acids. Subsequently, the PFVM is built up, which assembled all possible local folding variations in PFSC along sequence. Based on PFVM, a massive number of folding conformations are obtained by PFSC strings for the entire protein with combination of various local folding variations. Furthermore, according to these PFSC strings the protein multiple conformational structures are constructed as the predicted results. Starting from a protein sequence, the PFSC-PFVM provided an effective algorithm with both biological and physical meaningful process to predict the conformational ensembles of protein structures.

Methods

The PFSC-PFVM approach is composed of four modules to predict an ensemble of multiple conformation protein structures. The process is presented in Figure 1.

PFSC: A protein conformation is able to be completely described by a PFSC alphabetic letter string. Mathematically, starting successive five points connection without any constrain, a set of folding shapes for five of amino acid residues is obtained completely to cover folding space, which are represented by 27 letters including “\$” as PFSC. The 27 PFSC letters well character protein folding features, including alpha-helix, beta-strand, irregular folds and mixture in various degrees. The PFSC string is able fully to describe the folding conformation without gap along sequence, covering secondary structure fragments as well as tertiary structure fragments.

5AAPFSC: All possible folding patterns for each five amino acid fragment in PFSC letters are assembled into the 5AAPFSC database. For 20 amino acids, 3,200,000 permutations of five amino acids exist mathematically. The structural data about two-thirds of permutations of five amino acids are available in PDB, so their folding patterns are able to be collected. The structures of remnant one-third of permutations of five amino acids were computed by MD simulation with CHARMM (Chemistry at Harvard Macromolecular Mechanics)³. Then, all folding patterns for five amino acids are converted into the PFSC letters and stored in 5AAPFSC database.

PFVM: The local folding variations of a protein are fully presented by its PFVM. A PFVM can be assembled by extracting the folding shapes for five amino acid residues from the 5AAPFSC database according to its sequence. In PFVM, the protein sequence from N-terminus to C-terminus is horizontally displayed on top, and the possible folding shapes in PFSC letters for five amino acid residues are listed in a column below corresponding five residues. Based on a PFVM, an astronomical number of folding conformations in PFSC strings can be explicitly obtained for the protein without ambiguity. A set of most possible folding conformations in PFSC are able to be obtained. The PFSC string on first row in the PFVM

is one of most possible conformation for protein, and more possible conformations can be formed by optimization with replacement of PFSC letters in PFVM.

Construct an Ensemble of Protein 3D Structures: An ensemble of multiple conformational protein structures can be predicted according to the PFSC strings from PFVM. Based on each PFSC string, its protein 3D structure can be constructed by a high throughput screening the PDB database with homologous conformation process. For homologous conformation search, a protein is usually divided into multiple fragments with length of PFSC string about 50-100 residues. And then fragments are connected back for the whole protein. Thus, the protein 3D structure is constructed according to each PFSC string, and an ensemble of protein 3D structures for multiple conformations is predicted.

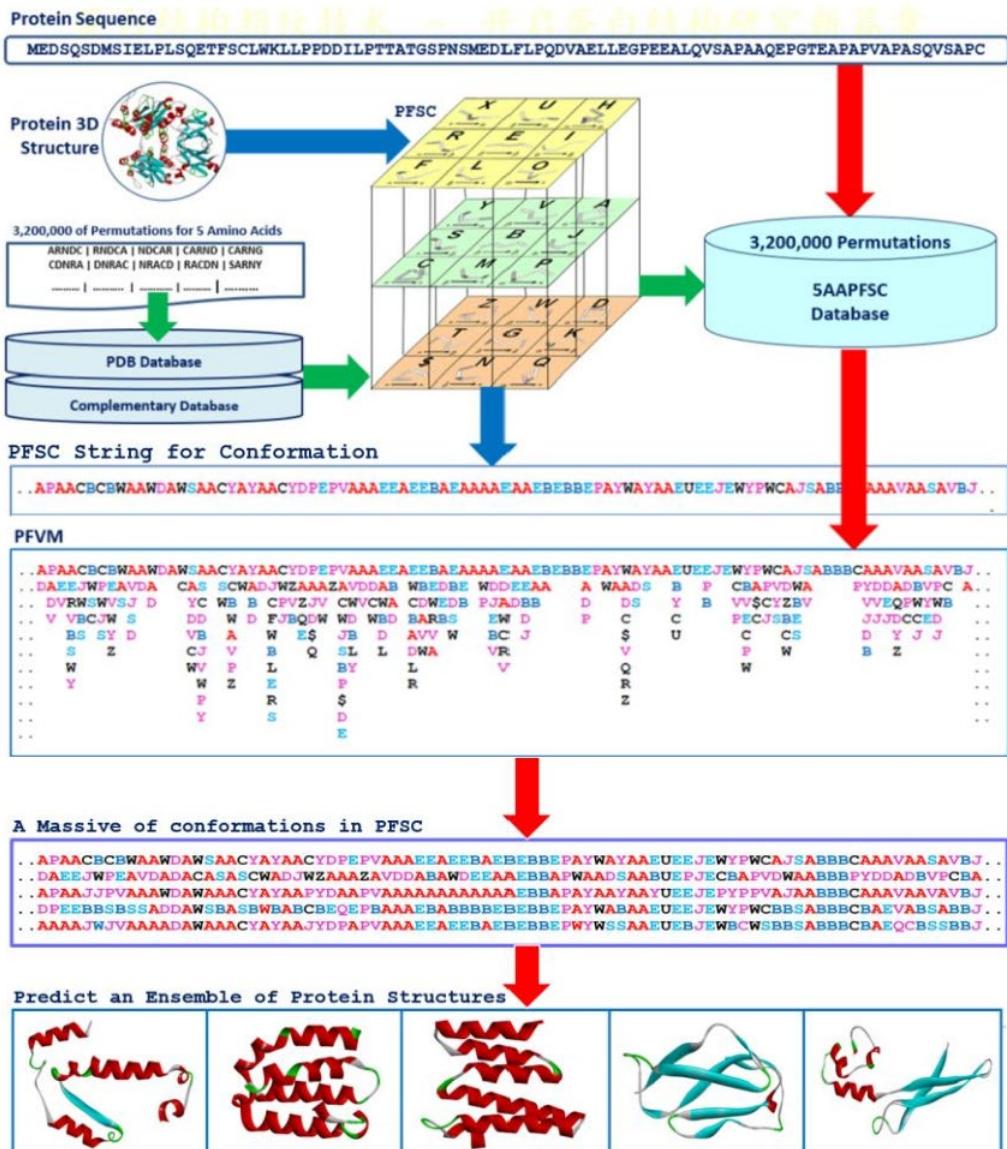


Figure 1. Protein structure prediction for conformational ensembles by PFSC-PFVM approach. The cubic contains a set of 27 PFSC as folding patterns. The blue arrows indicated the process the PFSC string for a protein with known 3D structure. The green arrows indicated the process to construct 5AAPFSC database, which contained all folding shapes in PFSC letters for 3,200,000 of permutations of five amino acids. The red arrows indicated the process how to obtain the PFVM and to predict an ensemble of protein structures from a sequence. The PFSC

letters with red and pink colors are for typical helix and alike-helix local folds; the blue and light blue colors for beta strand and alike-beta strand and block color for irregular folds.

Results

We predicted structures for 98 proteins. For some proteins, we provided 3D protein structures for five of multiple conformations.

Availability

The PFVM can be generated at a link in <http://www.microph.com/>. Click “Protein Structure Fingerprint Technology” in green bar, then click “Login” with Username = public and Password = public. Then select “Prediction” in top menu. User can enter any amino acid sequence as input, and the PFVM will be output as result on screen or save into a file as output.

1. Jiaan Yang. Comprehensive description of protein structures using protein folding shape code. *Proteins*;71.3:1497-1518 (2008)
2. Jiaan. Yang, Wen Xiang Cheng, Xiao Fei Zhao, Gang Wu, Si Tong Sheng, Qi Yue Hu and Peng Zhang. Comprehensive folding variations for protein folding. *Proteins*. 2022, 90(11), 1851-1872.
3. Brooks BR, Brooks CL, Mackerell AD, et al. CHARMM: The biomolecular simulation program. *J Comput Chem*.2009;30(10):1545-1614

PIEFold_human

PIEFold: an Improved Template-based Protein Structure Prediction Pipeline with PBEscore

Shun-Jie Xiong, Wei Cao, Lu-Yun Wu, Zhi-Xin Wang and Xian-Ming Pan*

Key Laboratory of Ministry of Education for Protein Science, School of Life Sciences, Tsinghua University, Beijing 100084, China

*Corresponding author: Dr. Xian-Ming Pan (pan-xm@mail.tsinghua.edu.cn)

Key: Auto:Y; CASP_serv:Y; Templ:Y; MSA:Y; Dist:Y; Tors:N; DeepL:N; AF:AF2; EMA:Y; MD:Y

In CASP16, we updated our pipeline mainly based on our workflow in CASP15. The first step of our pipeline is remote homologous protein search with our alignment-based method. Then, RosettaCM¹ is used to build the 3D protein structures based on both the outputs of the first step, and ColabFold² predicted models released on CASP site. Then, 3DRobot³ is used to generate decoys which are near top-ranked structures. Then, PBEscore, a knowledge-based energy scoring function developed by us, is applied to rank and pick plausible structures. This step is repeated for three times. In the end, the picked structures are refined by molecular dynamics software OpenMM⁴ to obtain final predicted models.

Methods

Given a target sequence, our pipeline has five steps as follows:

1. Detecting remote protein homology by alignment-based PairThreading. The quantity and quality of homology protein structures, especially remote homology proteins', largely influence the performance of our pipeline. Although there are many alignment methods for remote protein homology detection, most of them are based on the assumption that the types of residues at different positions are independent of each other. We abolished this assumption in our proposed method, PairThreading, which is based on residue pair substitution information. PairThreading obtains position-specific residue pair substitution information indirectly from the position-specific score matrices (PSSMs) rather than directly from the multiple sequence alignments (MSAs) to avoid statistical non-convergence problem. Thus, PairThreading can detect more remote homologous proteins and can generate more accurate alignments. For targets with sequences longer than 1000, our method has a dynamic strategy to detect domain regions and generate fragments for further 3D modeling and template hybridization.

2. Constructing 3D-models by RosettaCM¹. We use RosettaCM¹ to construct 3D-models based on the single or multiple templates that are selected by PairThreading along with corresponding alignments. In this step, we also include the ColabFold² predicted structures of the target as the templates for hybridization.

3. Generating decoys using 3DRobot³. The constructed 3D models are ranked and selected by our energy function PBEscore, and the top-ranked models are used to generate protein structure decoys by 3DRobot³, many of which have enhanced hydrogen-bondings and compactness.

4. Ranking 3D-models by PBEscore. PBEscore is a novel knowledge-based energy scoring function, simply considering the interactions of peptide bonds, rather than the conventional residues or atoms to be the most important energy contributors. Compared to our work in CASP15, the energy function has been retrained on an in-house dataset and it shows improved performance on several independent benchmark datasets. We apply PBEscore in every ranking step in our pipeline.

5. Refining 3D-models by OpenMM⁴. In the end, we run the molecular dynamics simulation to refine the final top 5 models. The program PDBFixer is used to add hydrogen atoms, N- and C-terminal patches to the selected models. All simulations are run using OpenMM⁴ under AMBER14⁵ force field.

1. Song, Y., DiMaio, F., Wang, R. Y. R., Kim, D., Miles, C., Brunette, T. J., ... & Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure*, 21(10), 1735-1742.
2. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature methods*, 19(6), 679-682.
3. Deng, H., Jia, Y., & Zhang, Y. (2016). 3DRobot: automated generation of diverse and well-packed protein structure decoys. *Bioinformatics*, 32(3), 378-387.
4. Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., ... & Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, 13(7), e1005659.
5. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8), 3696-3713.

Enhancing Protein Structure Prediction through MSA Quality Improvement with PLM-MSA

Wei Liu¹, Shanfeng Zhu^{2*}

¹ Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, USA;

² Institute of Science and Technology for Brain-Inspired Intelligence and MOE Frontiers Center for Brain Science, Fudan University, Shanghai, 200433, China

Accurate protein structure prediction is crucial for understanding biological functions and mechanisms. Multiple Sequence Alignment (MSA) quality significantly influences the performance of structure prediction methods like AlphaFold2 [1]. We introduce PLMFold, a novel approach that enhances protein structure prediction by improving MSA quality using a Protein Language Model-based MSA (PLM-MSA). PLMFold integrates advanced homology search techniques, a bucket-filter algorithm for sequence selection, and a query-centered MSA construction method, ultimately leading to more accurate structure predictions.

Methods

Homology Protein Search

We perform extensive homology searches against two comprehensive protein databases to obtain a diverse set of homologous sequences. First, we utilize the PLMSearch server (<https://dmiip.sjtu.edu.cn/PLMSearch/>) [2] to search the UniRef50 database [3], capturing well-annotated sequences with high coverage. Second, recognizing the limitations of environmental sequences in UniRef50, we employ the ColabFold framework to search the ColabFoldDB [4], which includes sequences from metagenomic sources, thus enriching our dataset with diverse environmental sequences. By merging the search results from both databases, we maximize sequence diversity for subsequent processing.

Bucket-Filter Algorithm

To address redundancy and over-representation of similar sequences, we implement a bucket-filter algorithm [4]. Homologous sequences are grouped into identity buckets defined by sequence identity ranges (e.g., (0.2–0.3], (0.3–0.4], ..., (0.9–1.0]). In each bucket, we retain up to 1,000 of the most diverse sequences, ensuring a balanced representation across different similarity levels. For buckets with fewer than 1,000 sequences, all available sequences are included without filtering. Clustering is performed using Linclust to select representative sequences, reducing redundancy while maintaining coverage [5]. This approach enhances the MSA's capacity to reflect evolutionary diversity and avoids the inclusion of highly similar sequences that could bias the alignment.

PLM-MSA Pipeline

Traditional MSA construction methods often rely on phylogenetic trees and may not adequately capture the evolutionary relationships pertinent to a specific query protein. In PLMFold, we develop PLM-MSA, a query-centered MSA construction pipeline. We leverage the PLMAAlign algorithm, which utilizes embeddings from protein language models to perform pairwise sequence alignments with improved accuracy over classical methods like Needleman-Wunsch [6]. The PLM-MSA pipeline aligns target sequences sequentially to the query sequence, sorted by decreasing similarity, ensuring that the alignment

focuses on accurately representing the query's evolutionary context. Importantly, we avoid introducing insertions (gaps) in the query sequence, as these gaps do not contribute to understanding the query's structural features.

Structure Prediction with AlphaFold2

The refined MSA generated by PLM-MSA serves as input to AlphaFold2 [1], utilizing the ColabFold team's modified backend for efficient computation [4]. By providing a high-quality, diverse, and appropriately filtered MSA, we enhance AlphaFold2's ability to generate accurate structural models of the query protein.

Accuracy Estimation using pLDDT

For local structure confidence estimation, we adopt the predicted Local Distance Difference Test (pLDDT) scores calculated by AlphaFold2 [1]. The pLDDT score provides residue-level confidence measurements, essential for assessing the reliability of specific regions within the predicted structure. By improving MSA quality, PLMFold indirectly enhances the accuracy of pLDDT scores, offering more reliable confidence estimations.

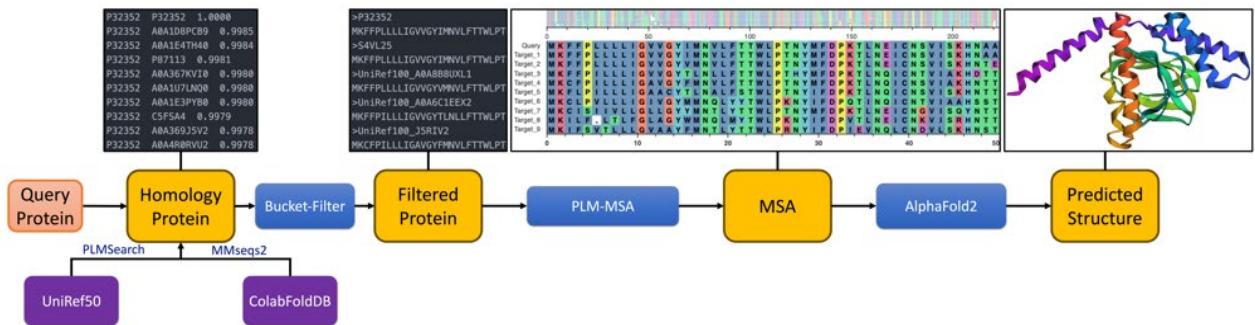


Figure 1. PLMFold Pipeline

1. Jumper, J., et al., Highly accurate protein structure prediction with AlphaFold. *Nature*. 596(7873): p. 583-589.
2. Liu, W., et al., PLMSearch: Protein language model powers accurate and fast sequence search for remote homology. *Nature Communications*, 2024. 15(1): p. 2775.
3. The UniProt, C., UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*. 51(D1): p. D523-D531.
4. Mirdita, M., et al., ColabFold: making protein folding accessible to all. *Nature Methods*, 2022. 19(6): p. 679-682.
5. Steinegger, M. and J. Söding, Clustering huge protein sequence sets in linear time. *Nature Communications*, 2018. 9(1): p. 2542.
6. Needleman, S.B. and C.D. Wunsch, A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, 1970. 48(3): p. 443-453.

Advancing Protein-Ligand Binding Site Prediction with 3D Instance Segmentation Techniques

Esha Rajesh Gavali¹, Angel Zarate¹, Rui Ding², Yang Wang², Renzhi Cao³, Jie Hou^{4*}, Dong Si^{1*}

1 - Division of Computing and Software Systems, University of Washington Bothell, Bothell, WA, USA;

2 -Information Materials and Intelligent Sensing Laboratory of Anhui Province, School of Computer Science and Technology, Anhui University, Hefei, China, 3 - Department of Computer Science, Pacific Lutheran University, Tacoma, WA, USA; 4 -Department of Computer Science, Saint Louis University, Saint Louis, MO, USA

jhou4@slu.edu, dongs@uw.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; AF:Y; EMA:N; MD:N

In CASP16, we evaluated our new ligand binding site prediction (LBSP) method, ISBNet-Pocket, on protein-ligand complex structure prediction. This method leverages advanced 3D point cloud instance segmentation, enhancing LBSP accuracy by representing proteins as 3D point clouds. Employing deep-learning-based instance segmentation, ISBNet-Pocket identifies and differentiates specific binding pockets, delivering accurate predictions of ligand binding sites. This approach also eliminates the need for post-processing clustering of predicted atoms and effectively handles the often proximate or overlapping nature of binding sites.

For CASP16, we initially predicted the 3D structure of target sequences using AlphaFold2¹ for the ligand binding prediction targets. These predicted structures were input into ISBNet-Pocket to derive predicted pockets at the residue level. Subsequently, the provided smile string of the ligand was converted into PDBQT file through the Meeko ligand preparation program (<https://github.com/forlilab/Meeko>). The final ligand binding poses were generated using AutoDock Vina², based on the predicted protein structure and the predicted pocket. We defined the center of the AutoDock grid as the geometric center of the predicted pockets, and the grid dimensions were manually specified based on the pocket size within the protein structure, as visualized in PyMOL³. We also used the online SwissDock site for prediction validation and visualization⁴. We evaluated the quality of all predicted structures based on the default binding affinity scores, and the top-ranked models were selected for submission to CASP16.

Methods

Dataset: The ISBNet-Pocket method was trained on sc-PDB dataset⁵ and evaluated on three datasets commonly employed to benchmark ligand binding site prediction (LBSP) models: BU48⁶, COACH420⁷, and COACH420(mlig)⁷. These datasets include both protein structures and their corresponding ligand structures. We applied feature engineering strategies akin to those utilized by GrASP⁸, enriching our dataset with both chemical and physical attributes of residues known to influence ligand binding. Such features, traditionally used to identify ligand-binding pockets, were logically incorporated as inputs to our model. We employed 13 physiochemical features, including coordinates, residue names, atom types, solvent accessible surface area, and RDKit features such as the number of bonds for heavy atoms, formal charges, ring structures, aromaticity, hybridization, and chemical hydrophobic properties. The model is designed to predict whether an atom is part of a binding pocket (1) or not (0).

Model Training: We utilized ISBNet⁹, a new deep learning model with cutting-edge techniques for efficient instance segmentation, for protein data training. The method adopts a cluster-free framework using Instance-aware Farthest Point Sampling (IA-FPS) and Box-aware Dynamic Convolution to generate

high-recall kernels and leverage geometric cues from axis aligned bounding boxes. Initially designed for dense point clouds representing 3D scenes, ISBNet was adapted to manage the segmentation of sparse point cloud data from protein atoms. During hyper-parameter tuning, particular attention was given to class imbalance issues, considering the relatively small size of ligand binding pockets compared to the entire protein. Key configuration adjustments included setting weights for cross-entropy loss, defining the radius value in the aggregation layer of instance segmentation, and applying standard non-max suppression (NMS). The training process for ISBNet-Pocket is a 2-step process: initially focusing on semantic segmentation of protein atoms, followed by fine-tuning to specifically identify individual instances of ligand binding pockets.

Inferences: Post-processing was essential to isolate atoms within the pockets identified by ISBNet-Pocket. The model produces distinct binary masks for each instance, requiring scripts to read and apply these masks to the original PDB files to facilitate pocket retrieval. Only masks with a confidence level above 0.65 were considered valid. The extracted pockets are subsequently saved in .pdb format, ready for further analysis or application.

1. Cramer, P. AlphaFold2 and the future of structural biology. *Nat. Struct. Mol. Biol.* 28, 704–705 (2021).
2. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461 (2010).
3. Seeliger, D. & de Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.* 24, 417–422 (2010).
4. Grosdidier, A., Zoete, V. & Michelin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 39, W270–W277 (2011).
5. Desaphy, J., Bret, G., Rognan, D. & Kellenberger, E. sc-PDB: a 3D-database of ligandable binding sites—10 years on. *Nucleic Acids Res.* 43, D399–D404 (2014).
6. Lu, C. et al. Protein-Ligand Binding Site Prediction and De Novo Ligand Generation from Cryo-EM Maps. *bioRxiv* 2023.11.16.567458 (2023) doi:10.1101/2023.11.16.567458.
7. Krivák, R. & Hoksza, D. P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J. Cheminform.* 10, 1–12 (2018).
8. Smith, Z., Strobel, M., Vani, B. P. & Tiwary, P. Graph Attention Site Prediction (GrASP): Identifying Druggable Binding Sites Using Graph Neural Networks with Attention. *J. Chem. Inf. Model.* (2024) doi:10.1021/acs.jcim.3c01698.
9. Ngo, T. D., Hua, B.-S. & Nguyen, K. ISBNet: A 3D Point Cloud Instance Segmentation Network With Instance-Aware Sampling and Box-Aware Dynamic Convolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 13550–13559 (2023).

[RNA_Dojo](#)

RNA 3D Structure Prediction with Human Intervention: Combining MXfold2, IPknot and FARFAR2

Junichi Iwakiri¹, Takumi Otagaki¹, Kazuteru Yamamura¹, Shunsuke Sumi^{2,3}, Ikuo Kurisaki³, Michiaki Hamada³, Jiro Kondo⁴ and Kengo Sato⁵

1 Graduate School of Frontier Sciences, The University of Tokyo, 2 Institute for Quantitative Biosciences (IQB), The University of Tokyo, 3 Graduate School of Advanced Science and Engineering, Waseda University, 4 Department of Materials and Life Sciences, Sophia University, 5 School of Life Science and Technology, Tokyo Institute of Technology

iwakiri@edu.k.u-tokyo.ac.jp

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N.MetaG; DeepL:Y; EMA:Y; MD:Y

The RNA three-dimensional structure predictions by the RNA_Dojo group in CASP16 primarily consisted of a two-step prediction process: RNA secondary structure prediction using a customized MXfold2, followed by three-dimensional structure prediction using FARFAR2. Since this was our group's first participation in CASP, instead of using a fully automated prediction pipeline, we approached each target with human intervention, where team members discussed and explored methods best suited to each specific target.

Methods

We performed RNA secondary structure prediction using customized MXfold2 [1] and IPknot [2]. MXfold2 was re-trained with approximately 16 times larger training data than the original version. This training data includes the bpRNA-1m dataset and the Ribonanza dataset, which contains extensive RNA structure probing data. The base-pairing probability matrix calculated with the re-trained MXfold2 was input into IPknot to predict RNA secondary structures including pseudoknots.

Our baseline method for RNA tertiary structure prediction involves predicting the RNA secondary structure using the new MXfold2 and IPknot, and using the predicted secondary structure as input to FARFAR2 [3] to generate 1000 candidate 3D structures (Figure 1). These 1000 predicted structures are then clustered using the DBSCAN method, and we select the top-scoring structure within each cluster, based on RNA-BRiQ [4] scores, as the structure to be submitted. If a homologous RNA structure is available, we calculate the RMSD between the predicted structure and the homologous structure, and select the structure with the smallest RMSD for submission. In the case of RNA-protein complexes, the RNA tertiary structure is predicted using the method described above, the protein structure is predicted using AlphaFold3 [5], and the complex structure is predicted using ZDOCK [6] with the predicted RNA and protein structures as input. The team members discussed the initially predicted structures and explored methods suited to each target, including performing secondary structure analysis using RNAPdbee [7] and executing FARFAR2 with fixed chunks, where partial structures were predicted using AlphaFold3.

Ion's distributions around the RNA are simulated by employing a three-dimensional reference interaction site model (3D RISM) [8]. Radial distribution functions are for aqueous solution under 140 mM KCl and 10 mM MgCl₂. Kovalent-Hirata closure is employed to solve RISM equations.

RLDOCK is employed to perform global docking search with multiple ligand conformations [9]. A ligand structure is generated from a given SMILE format file and used to prepare multiple conformations (200, at maximum) by using Open Babel [10]. RNA-ligand complexes are ranked by AnnapuRNA potential energy [11] and the energy-minimum structure is selected for submission.

Atomistic molecular dynamics simulations were performed by using Amber 22 MD simulation packages with biomolecule's force field parameters [12-15] and MD trajectories were analyzed by using AmberTools 22 [12].

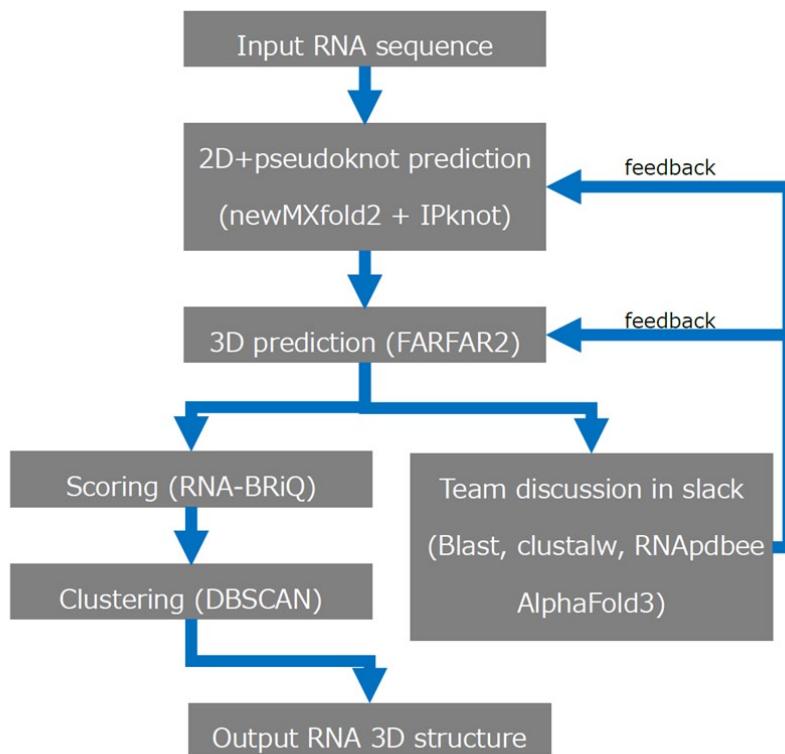


Figure 1. Overview of our RNA 3D structure prediction in CASP16

1. Sato,K., Akiyama,M. & Sakakibara,Y. (2021). RNA secondary structure prediction using deep learning with thermodynamic integration. *Nat. Commun.* **12**, 941.
2. Sato,K. & Kato,Y. (2022). Prediction of RNA secondary structure including pseudoknots for long sequences. *Brief. Bioinform.* **23**, bbab395.

3. Watkins,A.M., Rangan,R., & Das,R. (2020). FARFAR2: improved de novo rosetta prediction of complex global RNA folds. *Structure*, **28**(8), 963-976.
4. Xiong,P., Wu,R., Zhan,J., & Zhou,Y. (2021). Pairing a high-resolution statistical potential with a nucleobase-centric sampling algorithm for improving RNA model refinement. *Nat. Commun.*, **12**(1), 2777.
5. Abramson,J., Adler,J., Dunger,J., Evans,R., Green,T., Pritzel,A., ... & Jumper,J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**, 493–500.
6. Iwakiri,J., Hamada,M., Asai,K., & Kameda,T. (2016). Improved accuracy in RNA–protein rigid body docking by incorporating force field for molecular dynamics simulation into the scoring function. *J. Chem. Theory Comput.*, **12**(9), 4688-4697.
7. Zok,T., Antczak,M., Zurkowski,M., Popenda,M., Blazewicz,J., Adamiak,R.W., & Szachniuk, M. (2018). RNAPdbe 2.0: multifunctional tool for RNA structure annotation. *Nucleic Acids Res.*, **46**(W1), W30-W35.
8. Imai,T., Kovalenko,A. & Hirata,F. (2006) Hydration structure, thermodynamics, and functions of protein studied by the 3D-RISM theory. *Mol. Sim.* **32**, 817-824
9. Sun,L., Jiang,Y., Zhou,Y. & Chen,S. (2020). RLDOCK: a new method for predicting RNA-ligand interactions. *J. Chem. Theory Comput.* **16**, 7173-7183.
10. O'Boyle,N.M., Banck,M., James,C.A., Morley,C., Vandermeersch,T. & Hutchison,G.R. (2011) Open Babel: An open chemical toolbox. *J. Cheminf.* **3**, 33.
11. Stefaniak,F. & Bujnicki,J.M. (2021) AnnapuRNA: A scoring function for predicting RNA-small molecule binding poses. *PLoS. Comput. Biol.* **17**, e1008309..5. D.A. Case, H.M. Aktulga, K. Belfon, I.Y. et. al. Amber 2024, University of California, San Francisco, **2024**.
12. Case,D.A., Aktulga,H.M., Belfon,K. et. al. Amber 2024, University of California, San Francisco, **2024**.
13. Steinbrecher,T.J., Latzer; Case,D.A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.*, **2012**, 8, 4405–4412
14. Jorgensen,W.L., Chandrasekhar,J., MaduraJ., Klein,M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **1983**, 79, 926–935.
15. Joung,I.S., Cheatham,T.E.III. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B*, 2009, **113**, 13279 13290.

***Ab initio* RNA structure prediction with AI-based inter-and physics-based potential**

Chenjie Feng¹, Zhang Tian², Yang Li³, and Yang Zhang³

¹ -School of Science, Ningxia Medical University

² -Research Center for Mathematics and Interdisciplinary Sciences, Shandong University

³ -Cancer Science Institute of Singapore, National University of Singapore,

chenjief@nxmu.edu.cn

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF:N; EMA:N; MD:N

The biological importance of RNAs has been increasingly recognized over the past years. As is the case with proteins, the function of an RNA molecule is encoded in its tertiary structure, which is in turn determined by the nucleotide sequence. The tertiary structures of RNAs are usually less stable and therefore more difficult to predict compared to proteins. Here, we propose a hierarchical approach for fully automated RNA tertiary structure prediction, which combines the structural templates by threading alignments and inter-nucleotide geometries predicted by a deep residual convolutional neural network.

Methods

The proposed method consists of four main steps: (1) RNA templates detection by threading method, (2) (2) multiple sequence alignment generation, (3) inter-nucleotide geometries prediction, and (3) structure selection, optimization, and refinement.

Starting from an RNA sequence, the approach first collects structural templates using a meta-threading method. Meanwhile, the MSA is composed of four stages that perform blastn, nhmmer, infernal searches against NCBI's nt, Rfam, RNACentral¹ and MARS² database, respectively. They are then converted into an MSA representation and a pair representation, which are fed into a transformer network to predict inter-nucleotide geometries³. It then assembles full-length coarse-grained models, which are specified by the phosphate group (P), sugar ring (C) and base (N), through replica-exchange Monte Carlo (REMC) simulations under the guidance of a composite energy force field consisting of knowledge-based terms and template- and deep learning-based spatial restraints. Finally, the resulting structures are then refined using Arena⁴ for the full-atomic reconstruction of RNA coarse-grained models. To objectively assess our approach, we predicted all monomer RNA structures in CASP16 based on RNAFOLDX.

Availability

The web server will be made available at <https://zhanggroup.org/>

1. Consortium, R. (2021) RNACentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Res.*, **49**, 212–220.
2. Chen, K., Litfin, T., Singh, J., Zhan, J., Zhou, Y. (2024). MARS and RNACmap3: the master database of all possible RNA sequences integrated with RNACmap for RNA homology search. *Genomics, Proteomics & Bioinformatics*, **22**, 1.

3. Li, Y., Zhang, C., Feng, C., Pearce, R., Lydia Freddolino, P., & Zhang, Y. (2023). Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nature Communications*, **14**, 5745.
4. Perry, Z. R., Pyle, A. M., & Zhang, C. (2023). Arena: rapid and accurate reconstruction of full atomic RNA structures from coarse-grained models. *Journal of Molecular Biology*, **435**, 168210.

Modeling RNA and RNA-containing 3D structures in CASP16 by the RNApolis group

M. Popenda², J. Sarzynska², N. Dutta³, L. Popenda⁴, T. Zok¹,
M. Antczak^{1,2}, M. Szachniuk^{1,2}

1- Institute of Computing Science, Poznan University of Technology, Poland

2- Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poland

3 - University of Calcutta, India

4 - NanoBioMedical Centre, Adam Mickiewicz University, Poland

mszachniuk@cs.put.poznan.pl

Building on our group's success in modeling RNA tertiary structures in CASP15 [1], we applied a similar approach to predict the CASP16 targets for RNA, RNA-protein, and RNA-ligand complexes in the human category. We used RNAComposer, a tool that assembles RNA 3D structures from fragments derived from high-quality experimental data. Given the increased complexity of the CASP16 targets, we adapted and refined our methodology accordingly, although the overall pipeline remained unchanged. The workflow consisted of three steps: (1) determining the RNA secondary structure, (2) building an ensemble of RNA 3D structures from fragments derived from high-quality experimental structures, and (3) evaluating the models and selecting the best ones for submission.

In the first step, secondary structures were primarily determined using consensus models from Rfam [2] and the literature. For the second step, 3D structure components were automatically searched in an associated structure repository (an unreleased version of RNA FRABASE) or selected manually. If homologous structures were available for the target domains, structural components were derived from these homologs. However, for cases such as large noncoding RNAs — R1242 (RF03072), R1248, R0250, R0251, R0254 (RF02032), R0252, R0253, R0283, R0285, R1286 (RF03087), R0285 (RF01071) — no homologs existed. For these targets, we employed new in-house tools (RNAPony) that enabled us to search for extended structure elements, compute the percentage of topology identity (PTI) and sequence identity (PSI), and cluster the results. This approach helped us rationally select 3D structure elements. Among the structural elements found, those with high sequence similarity were clustered, allowing the best candidates to be chosen. For large, multi-domain targets, individual domains were built and then assembled. We extensively utilized the capability to modify the dot-bracket-encoded secondary structure to fragment it according to the modeler's needs.

The complexes were built using structural alignment with reference structures or structural elements using the selected protein and RNA models. In the assembly step, we were prompted to update the repository of experimentally-determined structure elements (RNA FRABASE). We also increased the extent of the modeler's intervention in the modeling process. First, we introduced the ability to indicate from which experimental structure we want to retrieve structural elements as a priority, and which ones we want to exclude (preferPDB, filterPDB). Secondly, we made it possible for the user to modify the minimization protocol, in particular freezing the selected residues and providing distance restraints at the stage of minimization in the torsion space. Extended user control over the modeling process was helpful, especially in modeling targets for which homologous structures for the targets, individual domains, or their fragments were available. For RNA-protein complexes (M1209, M1282, M1293, and M1296), the RNA components were modeled as described above. The protein 3D structures were obtained from the Protein Data Bank (PDB) [3], and any missing residues were modeled using Modeller [4]. The complexes

were built through structural alignment with reference structures or structural elements, using the selected protein and RNA models.

In this step, we updated the repository of experimentally-determined structure elements. Additionally, we increased the level of modeler intervention by introducing new features. First, we allowed modelers to prioritize specific experimental structures for retrieving structural elements (preferPDB) and exclude others (filterPDB). Second, we made it possible to modify the minimization protocol, including freezing selected residues and applying distance restraints during torsion space minimization. These extended user controls were particularly useful for modeling targets with available homologous structures for domains or fragments.

In the third step, RNA 3D models were ranked according to the total energy coefficient calculated by XPLOR [5]. Promising models, with total energy below a threshold value of -20 kcal/mol per residue, were selected for further processing using RNAspider [6]. This tool allowed us to eliminate models with entanglements, which were especially prevalent in those generated for long non-coding RNAs. The post-refinement total energy was verified for models that required additional refinement via RNAComposer. We used RNApdbee [7] to identify non-canonical interactions in the predicted models. Finally, RMSD-based clustering was performed using the OC program [8], and centroids from groups that were consistent with available literature and expert knowledge were selected for submission.

Availability

The methods we developed for RNA 3D structure prediction and used in CASP16 are available at <https://rnapolis.pl/>. The updated version of RNA FRABASE, which incorporates the RNAPony algorithm, will be available soon.

1. Sarzynska J, Popenda M, Antczak M, Szachniuk M (2023) RNA tertiary structure prediction using RNAComposer in CASP15, PROTEINS: Structure, Function, and Bioinformatics 91(12):1790-1799.
2. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: an RNA family database, Nucleic Acids Research 31(1):439-441.
3. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank, Nucleic Acids Research 28(1):235-242.
4. Webb B, Sali A (2016) Comparative Protein Structure Modeling Using Modeller, Current Protocols in Bioinformatics 54.
5. Schwieters CD, Bermejo GA, Clore GM (2018) Xplor-NIH for molecular structure determination from NMR and other data sources, Protein Science 27(1):26-40.
6. Luwanski K, Hlushchenko V, Popenda M, Zok T, Sarzynska J, Martsich D, Szachniuk M, Antczak M (2022) RNAspider: a webserver to analyze entanglements in RNA 3D structures, Nucleic Acids Research 50(W1):W663-W669.
7. Zok T, Antczak M, Zurkowski M, Popenda M, Blazewicz J, Adamiak RW, Szachniuk M (2018) RNApdbee 2.0: multifunctional tool for RNA structure annotation, Nucleic Acids Research 46(W1):W30-W35.
8. Barton GJ (1993, 2002, 2004) OC - A cluster analysis program, University of Dundee, UK.

Advancing structure prediction: modeling large complexes, protein-ligand, and antibody-antigen interactions beyond AlphaFold

Ben Shor, Alon Aronson, Dina Schneidman

The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

Keys: Auto:N; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; AF:AF2/AF3; EMA:N; MD:N.

Recognizing that achieving better results than AlphaFold2 [1] on monomeric structures would be challenging, we have decided to concentrate on areas where AlphaFold2 still exhibits limitations. We have focused our prediction efforts on large complexes, protein-ligand interactions, and nanobody interactions.

Protein complexes. To model large complexes comprising multiple subunits, we utilized a variation of the CombFold [2] method that incorporated manual assembly of predictions from AlphaFold2 [1] and AlphaFold3 [3]. We divided long polypeptide chains into subunits based on their domains, as inferred from the predicted monomer structures. For complexes with unknown stoichiometry, we began by predicting each subunit's dimers, trimers, and tetramers with copies of itself. We also predicted every possible pair of different subunits. Next, we examined the confidence metrics, specifically the Predicted Aligned Error (PAE). Out of all predicted subcomplexes, we grouped subcomplexes that are likely to interact and then predicted them together, and again, according to the confidence measures, we assessed what subcomplexes are most likely to be accurate. Finally, we identified a set of subcomplexes, each with high confidence, that together formed the complete complex, and merged them either manually using Biopython or automatically using CombFold with constraints.

Protein-ligand interactions. We developed and applied a deep learning model called EvoDocker or DockFormer. The model's architecture is based on AlphaFold2 but includes several significant modifications to support ligands and improve efficiency. The main difference is the replacement of MSA and template inputs with a reference input structure. This change simplifies the Evoformer, making it more similar to the PairFormer architecture presented in AlphaFold3. The number of layers in the Evoformer portion of the model is reduced to 8, instead of 48 in AlphaFold2. Additionally, we incorporated tokens representing ligand atoms alongside those representing amino acids. Lastly, we added an affinity module, enabling the model to be trained as a multi-task model and to predict affinity for predicted bound complexes.

Antibody-antigen complexes. We utilized Fold&Dock, a model previously developed in our lab, to predict and score protein complexes [4]. Following the prediction and scoring process, we employed Modeller [5] and Amber relaxation [6] to add missing side-chain atoms and resolve any structural clashes. For symmetric complexes, we first predicted a single antibody and then calculated the transformation between the symmetric antigen units, applying this transformation to the predicted antibody unit.

1. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold*. Nature, 2021.
2. Shor, B. and D. Schneidman-Duhovny, CombFold: predicting structures of large protein assemblies using a combinatorial assembly algorithm and AlphaFold2. Nat Methods, 2024. 21(3): p. 477-487.
3. Abramson, J., et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 2024. 630(8016): p. 493-500.
4. Cohen, T. and D. Schneidman-Duhovny, End-to-end accurate and high-throughput modeling of antibody-antigen complexes. NeurIPS MLSB, 2022.
5. Webb, B. and A. Sali, Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics, 2016. 54: p. 5 6 1-5 6 37.
6. Hornak, V., et al., Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins, 2006. 65(3): p. 712-25.

Seder2024

Eshel Faraggi¹, Robert Jernigan², Andrzej Kloczkowski³

1 - Physics Department, Indiana University, Indianapolis, IN

2 - Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA

3 - The Steve and Cindy Rasmussen Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH

This server is an updated version of the Seder2022 server that participated in CASP15. The main differences in this version are a new selection of optimized weights to achieve prediction, and obtaining sequence profiles from the Swissprot dataset instead of the complete non-redundant dataset of deposited protein sequences. The Seder server attempts to predict the TM score of protein models based on the sequence of the protein and the structure of its model using an iterative deep neural network. The cycle of Seder is unique in that it uses a balanced set of hybrid back-propagation/Levenberg-Marquardt and standard back-propagation neural networks. The hybrid networks are useful in cases where a crucial part of the network is of a limited extent, while other parts may be very large. This is only partially the case here, since there is not enough data to support larger networks. Both networks use associative memory. The hard/easy flavors of the server are distinguished by the training set used for them. For the hard set only proteins with low identity to templates, less than 40% sequence identity as judged by three iterations of PSI-BLAST.

1. Faraggi, Eshel, and Andrzej Kloczkowski. A global machine learning based scoring function for protein structure prediction. *Proteins: Structure, Function, and Bioinformatics* 82, no. 5 (2014): 752-759.
2. Faraggi, Eshel, Robert L. Jernigan, and Andrzej Kloczkowski. A Hybrid Levenberg–Marquardt Algorithm on a Recursive Neural Network for Scoring Protein Models. In *Artificial Neural Networks*, pp. 307-316. Humana, New York, NY, 2021.
3. The UniProt Consortium. UniProt: the Universal Protein Knowledge base in 2023. *Nucleic Acids Res.* 51:D523–D531 (2023)
4. Altschul, Stephen F., et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25.17 (1997): 3389-3402.

SHORTLE

Quality Assessment of QA1 and QA3 Models by Ranking a Set of Score for Parameters Based on PDB Statistics Covering Diverse Structural Features

D. Shortle

The Johns Hopkins University School of Medicine, Baltimore, MD 21205

dshortl1@jhmi.edu

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N.MetaN; DeepL:N; AF:N; EMA:Y; MD:N

Models provided for QA1 and QA2 assessment were scored with a collection of statistical parameters and potentials accumulated from earlier CASP competitions. Comparing model scores with the scores of 6000 high resolution PDB x-ray structures allowed ranking of a model's quality.

Methods

The statistics-based scores used in this work are outlined below. Values of these parameters for residues in turn/loop segments were scored separately, as were two residues at junctions of turn/loops with helix/strand segments.

0.) sum of linear and squared atom-atom overlap calculated for separations of i to i+1; i to i +2-5; i to i+6 and longer

1.) several phi/psi potentials with and without chi angles for single residues¹, for residue pairs in helix, strand, turn/loop and junctions, and for triplets of residues in turn/loop segments.

2.) probabilities of distances for backbone NH to O and O to NH atoms as a function of amino acid types for single and pairs of residues

3.) a conventional distance pair potential for 80-atom types and 4 secondary structure types

4.) sum of the values of the distance pair potential of neighboring atom parameters within less than 1.5 angstroms distance.

5.) atom solvation (surface exposure) calculated by the EFF1 method²

6.) hydrogen bond energies for the backbone³; for side chains with polar or charged atoms, the 80-atom statistical potential was used

7.) for the 50% most buried atoms, the count of neighboring atoms less than 1.5A distant for each type-specific atom in a model was calculated. The mean and standard deviation of these count numbers was converted to a score of packing homogeneity.

8.) For each atom of 30 atom types⁴, the number of instances in which an atom had 3 neighbors less than 1.5A distant was counted. At least 2 of the 3 neighbors were required to be from a residue different than the residue of the scored atom.

As would be expected, there is a range of correlations among these different parameters, and their success in decoy discrimination tests varies greatly.

6000 PDB structures with R-factors of 1.4A or less were scored for each parameter and these values were sorted from lowest energy value to highest energy value. The full range of these values was divided into 20 equal bins. The parameter values for each model in a CASP set were then compared to this list, allowing a percentile ranking score to be derived from 0 lowest energy to 19 highest energy. A score of 20 was assigned if the model's parameter value did not reach the values of the highest energy PDB structure. These percentile scores for 3 sets of 8 to 10 parameters were then added with weights from 1, 2 or 3, giving a combination score.

When a majority of models had values of 20 for a particular parameter, values of all models in the set were sorted and a rank percentile score was derived from binning these scores into 10 equal bins.

The five submitted models were selected by hand based on the sum of these 3 combined scores.

Availability

Any inquiries about the details of this scoring strategy will receive a detailed response.

1. Fang, Q. & Shortle, D. (2005). A consistent set of statistical potentials for quantifying local side-chain and backbone interactions. *Proteins* 60: 90-96.
2. Lazaridis, T. & Karplus, M. (1999). Discrimination of the native from mis-folded protein models with an energy function including implicit solvation. *J.Mol.Biol.* 288, 477-487.
3. Kortemme, T., Morozov, A.V. & Baker, D (2003). An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J.Mol.Biol.* 326, 1239-1259.
4. Fang, Q & Shortle, D (2006). Protein refolding in silico with atom-based statistical potentials and conformational search using a simple genetic algorithm. *J. Mol. Biol.* 359:1456-1467.

SNU-CHEM-lig, SNU-CHEM-aff (Ligand)

Enhancing Structure and Affinity Prediction of Receptor-Ligand Interactions with Variable Receptor Flexibilities Using Deep Learning Methods

S. Lee^{1†}, J. Choi^{1†}, J. Kim^{1†}, C. Lee^{1†}, S. Lee^{1†}, J. Sim^{1†}, J. Jun², B. Bae¹, N. Jung¹,
M. Baek³, H. Park⁴ and C. Seok^{1*}

¹Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea

²College of Medicine, Seoul National University, Seoul 03080, Republic of Korea

³School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea

⁴Brain Science Institute, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea

†Equal contribution; *Corresponding author: chaok@snu.ac.kr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF:Y; EMA:Y; MD:N

In CASP16, we submitted structure and affinity predictions for all ligand category targets. These ligand targets exhibited a range of receptor conformational changes upon ligand binding, from a few side-chain rotations in pocket residues to large-scale backbone movements. To address this variability, we employed a diverse set of receptor-flexible docking, receptor conformational sampling, and re-ranking tools. When possible, we performed target-specific benchmark tests to determine the most suitable tools for each prediction. Many of these tools leverage novel deep learning algorithms, which demonstrated significant improvement over traditional template-based or energy-based methods.

Methods

Super-targets (L1000, L2000, L3000, L4000)

For all super-targets, complex structures of the target receptors bound to various ligands were available in the PDB and were used as input receptor structures for ligand docking.

For L1000 and L2000, receptor conformational changes upon ligand binding were limited to a few pocket residues in the PDB structures. Therefore, we utilized our in-house deep learning-based docking tool, GalaxyDock4, which incorporates pocket side-chain flexibility, as the primary docking method. Model 1 was obtained by first docking the ligand to the receptor structures in the PDB using GalaxyDock4 and then re-ranking the binding poses using the confidence score. Four additional models were selected from those generated using alternative docking tools, including GalaxyDock-DL¹, GALigandDock², and CSAlign-Dock³.

For L3000, the potential binding pocket is quite large, and significant receptor conformational changes upon ligand binding were observed in the available complex structures. To address this, global docking with DiffDock-L⁴ and flexible-receptor docking with GALigandDock were employed for ligands without available PDB structures containing similar ligands (Tanimoto similarity > 0.55). For ligands with PDB structures involving similar ligands, template-based docking with CSAlign-Dock and rigid-receptor docking with GalaxyDock-DL were used. Additionally, re-ranking was performed using our in-house deep learning-based energy function, DENOISer, originally developed to evaluate complex conformations generated by ligand docking on predicted receptor structures.

For L4000, receptor conformational changes upon ligand binding were minimal in the available structures. For non-covalent ligand targets, template-based docking (CSAlign-Dock), rigid-receptor docking (GalaxyDock-DL), and flexible-receptor docking (GALigandDock) were used to generate complex structures, depending on the availability of PDB structures with high ligand similarity. Model 1 was selected based on the consensus results from multiple docking tools and expert evaluation. For covalent ligand targets, our in-house covalent docking tool, C-Dock, was used to generate complex structures, with Model 1 selected using the neural network energy score from GalaxyDock-DL.

All final poses were locally optimized using the GALigandDock energy function.

Regular protein-ligand targets (T1214, L5001)

For T1214 and L5001, receptor structures were sampled using our in-house ensemble generator which leverages AlphaFold2 and 3Di. For T1214, we anticipated a significant conformational change in the ligand binding loop, which was considered when selecting receptor structures for docking. GALigandDock was used to generate binding poses, and DENOISer was applied to rank the generated poses.

NA-ligand complexes (RNA: R1261, R1262, R1263, R1264, and R1288, DNA: D1273)

Template searches for nucleic acid receptors were conducted using BLAST and RCSB. Receptor modeling was performed with RosettaFold2NA⁵ and AlphaFold3⁶, and the identified templates were cross-verified through a literature survey. Discrepancies were addressed by adjusting base orientation (R1261, R1262) or applying conditional flow matching, incorporating secondary structures identified from literature (R1263, R1264, D1273). For D1273, 3dRNA/DNA⁷ was used to enforce model conformity with the expected secondary structures. Binding sites were selected based on the detected templates and supporting literature. For each nucleic acid receptor structure, 100 ligand poses were generated via rigid-receptor docking with GALigandDock, and the poses were rescored using Rosetta energy and AnnapuRNA⁸. Metal ions were aligned based on the templates where applicable.

Binding affinity prediction (L1000, L3000)

For protein-ligand binding affinity predictions for targets L1000 and L3000, we used an in-house model, BG. BG was developed by adding additional network layers to the re-ranking score of GalaxyDock-DL¹ and trained on a binding affinity dataset curated from BindingDB using predicted protein-ligand complex structures. The model was further fine-tuned using target-specific data collected from BindingDB. In the first submission round, affinity predictions were based on our Model 1 structure, while in the second round, they were based on the crystal structures.

Results

To evaluate protein-ligand complex structure predictions, the ligand RMSD of the submitted models was calculated relative to the experimentally resolved ligand structures after superposing the model and experimental receptor structures. Based on the PDB structure of T1214 (PDB ID: 9C4O), our Model 1 for this target achieved a ligand RMSD of 0.8 Å, indicating highly accurate predictions for both receptor flexibility and ligand binding. For super-targets L1000 and L2000, the percentage of predictions with ligand RMSD within 2 Å and 1 Å was compared with AutoDock Vina⁹ in Table 1, demonstrating improved performance by our protocol. Further analysis is ongoing.

Table 1. Success rate of our method (SNU-CHEM-lig) and AutoDock Vina with RMSD cutoff of 2 Å and 1 Å. AutoDock Vina was run using the PDB receptor structure containing the ligand with the highest similarity to the target.

L1000 and L2000 (n = 19)	SNU-CHEM-lig		AutoDock Vina	
	RMSD < 2 Å	RMSD < 1 Å	RMSD < 2 Å	RMSD < 1 Å
Model 1	73.7%	36.8%	57.9%	36.8%
Best of 5 Models	89.5%	63.2%	89.5%	42.1%

Availability

Manuscripts detailing the in-house tools mentioned above—GalaxyDock4, DENOISer, C-Dock, and BG—are currently in preparation, and the programs will be made available in the future.

1. Lee, C., Won, J., Ryu, S., Yang., J. Jung, N. Park, H., & Seok, C. (2024). GalaxyDock-DL: Protein-ligand docking by global optimization and neural network energy. *Journal of Chemical Theory and Computation*, **20**, 7370-7382.
2. Park, H., Zhou, G., Baek, M., Baker D. and DiMaio, F. (2021). Force field optimization guided by small molecule crystal lattice data enables consistent sub-angstrom protein-ligand docking. *Journal of Chemical Theory and Computation*, **17**(3), 200-2010.
3. Kwon, S., & Seok, C. (2023). CSAlign and CSAlign-Dock: Structure alignment of ligands considering full flexibility and application to protein–ligand docking. *Computational and Structural Biotechnology Journal*, **21**, 1-10.
4. Corso, G., Deng, A., Polizzi, N., Barzilay, R., & Jaakkola, T. S. (2024). Deep Confident Steps to New Pockets: Strategies for Docking Generalization. *The Twelfth International Conference on Learning Representations*.
5. Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D., & DiMaio, F. (2024). Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature Methods*, **21**(1), 117-121.
6. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3.
7. Zhang, Y., Xiong, Y., Yang, C., & Xiao, Y. (2024). 3dRNA/DNA: 3D Structure Prediction from RNA to DNA. *Journal of Molecular Biology*, **436**(17), 168742.
8. Stefaniak, F., & Bujnicki, J. M. (2021). AnnappuRNA: A scoring function for predicting RNA-small molecule binding poses. *PLoS computational biology*, **17**(2), e1008309.
9. Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling*, **61**(8), 3891-3898.

An Improved Pipeline for RNA Structure Prediction: Using Coarse-Grained Simulation and Refinement to Enhance AF3 Predictions

F. Pucci¹, S. Poblete²

¹Computational Biology and Bioinformatics, Université Libre de Bruxelles, Belgium

²Centro BASAL Ciencia & Vida, Santiago, Chile

fabrizio.pucci@ulb.be

Key: RNA structure prediction, Combining deep learning and physics-based methods

RNA molecules play critical roles in various biological processes, including gene expression, regulation, and protein synthesis. The structure and function of RNA are closely intertwined, driving significant efforts within the scientific community to resolve RNA structures and better understand their biological roles. However, compared to proteins, the number of RNA structures available in standard databases, such as the Protein Data Bank (PDB) [1], is relatively limited. Over the past two decades, computational tools based on physics-based models and deep learning approaches have been developed to complement experimental efforts in solving RNA structures. Despite these advances, RNA structure prediction remains a significant challenge in bioinformatics.

Methods

Recently, many promising deep learning approaches have been developed. In particular, AlphaFold3 (AF3) [2] has demonstrated good performance compared to other state-of-the-art methods. However, we are still far from achieving highly accurate predictions, as RNA structures predicted by AF3 often have low LDDT scores, indicating low confidence in the predictions.

To enhance AF3 predictions, we designed a simple pipeline for the CASP16 round that integrates the output of AF3 with coarse-grained simulations using SimRNA. For each sequence submitted to AF3, all five structures generated by it are used as starting points for SimRNA simulations [3]. SimRNA is a simulated annealing-based method that employs coarse-grained physical potentials. We performed 10 simulations with 10 replicas each for every structure predicted by AF3. From all the generated configurations, we selected only the top 1% with the lowest energy and clustered them using a 5 Å RMSD threshold. The representative of the most populated cluster was then selected. In the final stage of the pipeline, we refined the proposed all-atom models using QRNAs and/or molecular dynamics (MD) relaxation, using AMBER force field under a restrained protocol [4].

1. Berman, Helen M., et al. "The protein data bank." Nucleic acids research 28.1 (2000): 235-242.
2. Abramson, Josh, et al. "Accurate structure prediction of biomolecular interactions with AlphaFold 3" Nature (2024): 1-3.
3. Boniecki, Michal J., et al. "SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction." Nucleic acids research 44.7 (2016): e63-e63
4. Poblete, Simon, et al. "Effects and limitations of a nucleobase-driven backmapping procedure for nucleic acids using steered molecular dynamics". Biochem. Biophys. Res. Comm. 498.2 (2018): 352-358.

In Silico Generation and Thermodynamic Ranking of RNA Tertiary Structures

L. Herron¹²⁶, V. Adury¹³, R. John¹⁴, S. Lee¹²⁶, S. Mehdi¹², D. Sanwal⁵, A. Verma¹², P. Tiwary¹⁵⁶

1 - Institute for Physical Science and Technology, University of Maryland, College Park, MD, 20742, USA

2 - Biophysics Program, University of Maryland, College Park, MD, 20742, USA

3 - Chemical Physics Program, University of Maryland, College Park, MD, 20742, USA

4 -Department of Physics, University of Maryland, College Park, MD, 20742, USA

5 - Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, 20742, USA

6 - University of Maryland Institute for Health Computing, Bethesda 20852, USA

ptiwary@umd.edu

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; EMA:Y; MD:Y

Predicting RNA tertiary structures is a significant challenge in computational biology due to the flexible, dynamic nature of RNA molecules and the scarcity of experimentally resolved RNA structures. Unlike proteins, which have extensive structural data available, RNA structures are underrepresented in databases like the Protein Data Bank (PDB), making data-driven prediction methods less effective¹.

Moreover, RNA molecules exhibit a high degree of conformational variability, challenging traditional structure prediction methods that often assume a single stable structure. Indeed, there is now a growing recognition that biomolecules should be viewed not as static entities but as dynamic ensembles of structures²⁻³. Therefore, effective RNA structure prediction methods need to generate not just a single static structure but an ensemble of possible structures, ranked by their thermodynamic stability and Boltzmann weights at physiological temperatures⁴.

In this work, we introduce RNAnneal, a purely computational two-stage pipeline designed to predict Boltzmann-scored ensembles of RNA tertiary structures starting from the primary RNA sequence. Our approach addresses the challenges of limited structural data and the need for ensemble predictions by combining bioinformatics methods, molecular dynamics simulations, and generative artificial intelligence models without pretraining on existing RNA structures.

Methods

In the first stage of RNAnneal, we generate diverse candidate tertiary structures by predicting RNA secondary structures using two tools: **EternaFold (EF)** for pseudoknot-free structures⁵ and **pKiss (PK)** for structures containing pseudoknots⁶. The predicted secondary structures are then assembled into tertiary structures with Rosetta's FARFAR2 algorithm⁷, which constructs models by assembling RNA structural motifs in a low-resolution potential.

To enhance accuracy, the FARFAR2-generated models are further refined in a high-resolution molecular dynamics RNA force field⁸. High-energy structures are discarded, and the lowest-energy secondary structures are passed back to FARFAR2 for another round of tertiary structure assembly. This iterative process steers the ensemble toward low-energy regions, yielding hundreds of candidate structures for subsequent ranking.

In the second stage, we score the candidates based on their thermodynamic properties using thermodynamic maps – a physics-based generative AI framework that learns the equilibrium distribution of structures and their temperature dependence⁹. Thermodynamic maps parameterize an invertible mapping between RNA structural models and a generative system via a diffusion process¹⁰. This allows us to generate samples at any temperature by sampling from the generative system at the corresponding temperature and applying the inverse map, which implicitly accounts for the effect of temperature on the ensemble.

Finally, we integrate thermodynamic maps with MD simulations to actively steer simulations launched from FARFAR2 models toward equilibrium. Once convergence is achieved, the free energy estimates from the thermodynamic map are used to assign Boltzmann scores to the candidate structures. This approach enables us to select a representative ensemble of RNA tertiary structures that reflects their thermodynamic stability and abundance at physiological temperatures. Importantly, our method does not require any pretraining on existing RNA structures, sequence information, or use of templates, standing in contrast to other AI approaches.

Availability

The thermodynamic map framework and integration with molecular dynamics simulation are outlined in detail in Herron et. al.⁹.

1. H. M. Berman, The protein data bank, *Nucleic Acids Research* 28, 235–242 (2000).
2. Q. Vicens and J. S. Kieft (2022), Thoughts on how to think (and talk) about rna structure, *Proceedings of the National Academy of Sciences* 119, e2112677119.
3. X.-C. Bai, G. McMullan, and S. H. Scheres (2015), How cryo-em is revolutionizing structural biology, *Trends in bio-chemical sciences* 40, 49.
4. G. R. Bowman (2024), Alphafold and protein folding: Not dead yet! the frontier is conformational ensembles, *Annual Review of Biomedical Data Science* 10.1146/annurev-biodatasci-102423-011435
5. H. K. Wayment-Steele, W. Kladwang, A. I. Strom, J. Lee, A. Treuille, A. Becka, and R. Das (2022), Rna secondary structure packages evaluated and improved by high-throughput experiments, *Nature Methods* 19, 1234–1242.
6. S. Janssen and R. Giegerich (2014), The rna shapes studio, *Bioinformatics* 31, 423–425.
7. A. M. Watkins, R. Rangan, and R. Das (2020), Farfar2: Improved de novo rosetta prediction of complex global RNA folds, *Structure* 28, 963.
8. D. Tan, S. Piana, R. M. Dirks, and D. E. Shaw (2018), Rna force field with accuracy comparable to state-of-the-art protein force fields, *Proceedings of the National Academy of Sciences* 115, E1346.
9. L. Herron, K. Mondal, J. S. Schneekloth, and P. Tiwary (2023), Inferring phase transitions and critical exponents from limited observations with thermodynamic maps, [arXiv:2308.14885](https://arxiv.org/abs/2308.14885).
10. Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole (2020), Score-based generative modeling through stochastic differential equations, arXiv preprint arXiv:2011.13456.

Updated DMFold-based Optimization for Protein Structure Prediction

Weikang Gong

5- Center for AI and Computational Biology, Institute of Systems Medicine, Chinese Academy of Medical Sciences /Suzhou Institute of Systems Medicine

wkgongwork@gmail.com

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:N; MD:N.

The protein structure prediction of the Unicorn group in CASP16 is based on a pipeline combining updated DeepMSA2 and AlphaFold2-Multimer. I have updated the database for calculating MSA, and after testing, the performance of the updated DMFold has improved.

Methods

The full pipeline contains three steps: (1) multiple sequence alignment (MSA) generation for the individual constituent proteins of the complexes by updated DeepMSA2, (2) MSA selection for each constituent, and (3) complex model construction and ranking by AlphaFold2-multimer pipeline with updated DeepMSA2 constituent MSAs as input.

Availability:

<https://zhanggroup.org/DMFold/>

1. Li Z, Liu X, Chen W, et al. Uni-Fold: an open-source platform for developing protein folding models beyond AlphaFold[J]. bioRxiv, 2022: 2022.08. 04.502811.
2. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. nature, 2021, 596(7873): 583-589.
3. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer[J]. biorxiv, 2021: 2021.10. 04.463034.
4. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. Nature, 2024: 1-3.
5. Zheng W, Wuyun Q, Freddolino P L, et al. Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15[J]. Proteins: Structure, Function, and Bioinformatics, 2023, 91(12): 1684-1703.

Protein and RNA/DNA structure prediction with the UNRES and NARES-2P coarse grained models assisted by AlphaFold

Cezary Czaplewski^{1,2}, Adam K. Sieradzan¹, Marta Pągielska¹, Jakub Susoł¹, Magdalena J. Ślusarz¹, Rafał Ślusarz¹, Łukasz Dziadek¹, Annemarie Danielsson¹, Truong Co Nguyen¹, Mateusz Leśniewski¹, Emilia A. Lubecka³, Sümeyye Atmaca¹, Krzysztof K. Bojarski⁴, Alexandra Leinberger⁵, Sergey A. Samsonov¹, Łukasz Golon¹, Adam Liwo^{1*}

¹ – Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland, ² – School of Computational Sciences, Korea Institute for Advanced Study, Seoul 02455, Republic of Korea, ³ – Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, G. Narutowicza 11/12, 80-233 Gdańsk, Poland, ⁴ – Faculty of Chemistry, Gdańsk University of Technology, G. Narutowicza 11/12, 80-233 Gdańsk, Poland, ⁵ – Horace Greeley High School Chappaqua, NY, 10514

adam.liwo@ug.edu.pl

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:N; DeepL:N; EMA:Y; MD:Y

We tested our physics-based approach for protein- and DNA/RNA-structure prediction, whose key components are the coarse-grained UNRES model of polypeptide chains¹ and the NARES-2P model of nucleic acids², respectively, with CASP16 targets. The approach is based on extensive conformational search by means of the multiplexed replica exchange molecular dynamics (MREMD)³. As opposed to most of the other approaches, the prediction candidates are selected according to the probabilities of the conformational ensembles they belong to. With respect to the last CASP, the program code has been heavily optimized⁴, thus enabling us to handle very large targets and our methodology was extended to treat RNA/DNA molecules.

Methods

The prediction procedure was similar to that developed in our earlier work⁵ and consisted of the following four stages (i) running restrained MREMD simulations of the targets with UNRES or NARES-2P, depending on target type, (ii) determining the probabilities of the conformations by using the binless weighted histogram analysis method (WHAM)⁶, (iii) dissecting the simulated conformational ensemble into 5 (CASP) or 10 (CAPRI) families by using the minimum-variance clustering and selecting the conformations closest to the cluster mean for further processing, the ranking following cluster free energy and (iv) conversion of the coarse-grained structures to all-atom structures. For proteins, the cg2all⁷ and for nucleic acids our in-house NARall⁸ backmapping softwares, respectively, were used to execute stage (iv). The final all-atom structures were submitted to CASP/CAPRI. The same protocol was used to process the ENSEMBLES targets except that 100 structures were finally generated.

For proteins, distance and angular restraints derived from the AlphaFold2⁹ (through ColabFold¹⁰), AlphaFold3¹¹ (through the AlphaFold3 server), and iTASSER¹² (in-house installation) models were imposed on MREMD simulations of stage (i) and these simulations were started from models, which were distributed to all MREMD trajectories. We also used the models from MassiveFold, which were initially clustered to extract sufficiently distinct structures, symmetry taken into account for oligomers. Additionally, disulfide-bond formation was considered in the simulations by means of the dynamic disulfide-bond handling functionality of UNRES¹³.

For RNA/DNA targets, the secondary structure was predicted initially, by using the RNAdorf web server¹⁴. The predictions with the highest-probability were then used to build the initial structures, with the RNA Composer web server¹⁵. Distance restraints derived from secondary-structure-prediction information were imposed during the MREMD simulation.

Availability

The standalone version of UNRES is available from <https://unres.pl/> and the web server version is available at <https://unres-server.chem.ug.edu.pl/>.

1. Sieradzan,A.K., Czaplewski,C., Krupa,P., Mozolewska,M.A., Karczyńska,A.S., Lipska,A.G., Lubecka,E.A., Gołaś,E., Wirecki,T., Makowski,M., Ołdziej,S. & Liwo,A. (2022) Modeling the structure, dynamics, and transformations of proteins with the UNRES force field. In: *Protein Folding: Methods and Protocols*, Springer US, New York, NY, 2022, pp. 399–416.
2. He,Y., Liwo,A. & Scheraga,H.A. (2015) Optimization of a nucleic acids united-residue 2-point model (NARES-2P) with a maximum-likelihood approach, *J. Chem. Phys.*, **143**, 243111.
3. Sieradzan,A.K., Sans-Duñó,J., Lubecka,E.A., C. Czaplewski, Lipska,A.G., Leszczyński,H., Ocetkiewicz,K.M., Proficz,J., Czarnul,J., Krawczyk,H. & Liwo,A. (2023) Optimization of parallel implementation of UNRES package for coarse-grained simulations to treat large proteins. *J. Comput. Chem.*, **44**, 602–625.
4. Czaplewski,C., Kalinowski,S., Liwo,A. & Scheraga,H.A. (2009) Application of multiplexed replica exchange molecular dynamics to the UNRES force field: Tests with α and $\alpha+\beta$ proteins. *J Chem. Theory Comput.*, **5**, 627-640.
5. Antoniak,A., Biskupek,I., Bojarski,K.K. et al. (2021) Modeling protein structures with the coarse-grained UNRES force field in the CASP14 experiment, *J. Mol. Graph. Model.*, **108**, 108008.
6. Kumar,S., Bouzida,D., Swendsen,R.H., Kollman,P.A. & Rosenberg,J.M. (1992) The weighted histogram analysis method for free energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, **13**, 1011-1021.
7. Heo,L. & Feig,M. (2024) One bead per residue can describe all-atom protein structures, *Structure*, **32**, P97-111.E6.
8. Golon,Ł. & Sieradzan,A.K. (2023) NARall: a novel tool for reconstruction of the all-atom structure of nucleic acids from heavily coarse-grained model. *Chem. Pap.* **77**, 2437–2445.
9. Jumper, J., Evans, R., Pritzel, A. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
10. Mirdita,M., Schütze,K., Moriwaki,Y. et al. (2022) ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682.
11. Thompson, & Howe,N.P. (2024) Alphafold 3.0: the AI protein predictor gets an upgrade, *Nature Podcast*, May 8, 2024, <https://www.nature.com/articles/d41586-024-01385-x>

12. Zhou,X., Zheng,W., Li,Y., Pearce,R., Zhang,C., Bell,E.W., Zhang,G. & Zhang,Y. (2022) I-TASSER-MTD: a deep-learning-based platform for multi-domain protein structure and function prediction. *Nat. Protocols*, **17**, 2326-2353.
13. Chinchio,M., Czaplewski,C., Liwo,A., Ołdziej,S. & Scheraga,H.A. (2007) Dynamic formation and breaking of disulfide bonds in molecular dynamics simulations with the UNRES force field. *J. Chem. Theory Comput.*, **3**, 1236–1248.
14. Gruber,A.R., Lorenz,R., Bernhart,S.H., Neuböck,R. & Hofacker,I.L. (2008) The Vienna RNA websuite, *Nucleic Acids Res.* **36**, W70–W74.
15. Popenda,M., Szachniuk,M., Antczak,M., Purzycka,K.J., Lukasiak,P., Bartol,N., Blazewicz, J. & Adamiak,R.W. (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res.*, **40**, e112.

Modeling Protein-Protein Targets by Combining and Analyzing AlphaFold Predictions

P.J. Kundrotas¹, M.M. Copeland¹ and I.A. Vakser^{1,2}

¹ - Computational Biology Program, ² - Department of Molecular Biosciences, The University of Kansas, Lawrence, KS, USA

vakser@ku.edu and pkundro@ku.edu

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF2; EMA:Y; MD:Y

The introduction of Deep Learning (DL) methods, particularly AlphaFold2 (AF2)¹ with its off-spring AlphaFold2-Multimer² and RoseTTAFold,³ led to significant increase in accuracy of modeling protein complexes, setting a new standard in structure prediction. However, these methods still face significant limitations, particularly in accounting for the physiological environment in which proteins operate, and capturing the dynamic and flexible nature of proteins, which are important factors in protein interactions. DL-based methods are also computationally demanding, requiring high-end hardware (Graphical Units Processors - GPU, fast memory and storage modules, etc.), with high computational costs of large-scale modeling of protein interactions. The CASP-CAPRI prediction rounds provide a unique opportunity to test DL methods expansion and combination with other docking approaches. The current round involved 31 protein-protein assemblies of different sizes and oligomeric states.

Methods

For most targets, we generated 100 - 500 models ranked by the AF2 confidence score (80% of predicted TM-score, and 20% of interface predicted TM-score),² except for the larger targets where AF2 failed to generate such number of models (Table 1). The AF2 pipeline was run on computational nodes with A100 and Q600 GPUs and nodes with Tesla T4 GPU. The models were clustered by highly connected subgraphs⁴ with 0.8 MM-score⁵ threshold. For each model, we analyzed predicted aligned errors (PAE) and predicted LDDT values of each residue, extracted from the AF2 pickle file and plotted by a Python script using standard *pickle*, *numpy*, *panda* and *matplotlib* packages. Docking models selected for submission were minimized by TINKER⁶ (CPU version 8 and GPU version 9) with CHARMM22 forcefield.⁷

Results

Our focus was only on modeling protein-protein assemblies. We did not use AF3 procedure⁸ available only as a server with limited capacity. Models for the submission were typically selected based on occupancy of the AF2 models clusters and quality of the inter-chain contacts assessed from the PAE diagrams. When time constraints permitted, AF2 modeling was also performed for the assemblies with lower, and sometimes higher, oligomeric states and some submissions were built from such models. This will allow us, after the release of the results assessment, to analyze the ability of AF2 to predict correct oligomeric states, and to correlate the quality of the models with the number of chains in the assembly. For several larger targets, AF2 modeling was also carried out for overlapping fragments with subsequent re-assembly of the modeled parts into the full structure. Manual curation involved analysis of available PDB structures similar to the target, using UNIPROT database.⁹

Table 1. Statistics on targets.

CAPRI target	CASP target	Oligomeric state	Total number of residues	Number of AF2 models ^a	Number of clusters ^a	Partial modeling ^b
T236	T1201	A2	420	100	5	yes
T238	H1202	A2B2	380	200	1	yes
T240	H1204	A2B2C2	858	200	18	yes
T242	T1206	A2	474	300	1	no
T244	H1208	A1B1	646	300	1	no
T248	H1215	A1B1	369	300	4	no
T250	T1249v1	A3	1464	20	N/A	no
T252 ^c	T1249v2	A3	1464	-	-	-
T256	H1213	A1B1C1D1E1	1373	N/A	N/A	yes
T258	H1217	A2B2C2D2E2F2	5878	N/A	N/A	yes
T260	T1218	A2	2328	22	1	no
T262	H1220	A1B4	2515	N/A	N/A	yes
T266	H1222	A1B1C1	485	500	1	no
T268	H1223	A1B1C1	486	500	2	no
T270	H1225	A1B1C1	483	500	1	no
T272	H1227	A1B6	5689	N/A	N/A	yes
T274	H1229	A1B1	987	500	1	no
T276	H1230	A1B1	984	500	1	no
T278	H1232	A2B2	924	457	43	yes
T280	H1233	A2B2C2	1316	109	28	yes
T282	T1237	A4	1952	339	3	yes
T284	T1238	A2	654	500	31	no
T286	T1240	A3	1959	347	50	no
T288	T1234	A3	1239	322	36	no
T290	T1235	A6	690	500	18	yes
T292	H1236	A3B6	1929	367	48	yes
T294	H1244	A2B2C2	850	500	60	yes
T296	H1245	A1B1	317	500	172	no
T298	T1259	A3	729	500	6	yes
T300	H1267	A2B2	1852	500	12	yes
T302	T1270	A6	2622	415	8	yes

^a Number of AF2 models and clusters are for given oligomeric state.

^b Submitted structures were built from models with reduced oligomeric state or from models containing only parts of the protein chains.

^c Modeled together with T250 as both consist of the same proteins in two distinct conformations.

- Varadi, M., Anyango, S., Deshpande, M., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 50, D439-D444.
- Evans, R., O'Neill, M., Pritzel, A., et al. (2021). Protein complex prediction with AlphaFold-Multimer. bioRxiv. doi:10.1101/2021.10.04.463034.

3. Krishna, R., Wang, J., Ahern, W., et al. (2024). Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*. 384, eadl2528.
4. Hartuv, E., Shamir, R. (2000). A clustering algorithm based on graph connectivity. *Inform. Process Lett.* 76, 175-181.
5. Mukherjee, S., Zhang, Y. (2009). MM-align: a quick algorithm for aligning multiple-chain protein complex structures using iterative dynamic programming. *Nucleic Acids Res.* 37, e83.
6. Rackers, J.A., Wang, Z., Lu, C., et al. (2018). Tinker 8: software tools for molecular design. *J. Chem. Theory Comput.* 14, 5273-5289.
7. MacKerell, A.D., Banavali, N., Foloppe, N. (2000). Development and current status of the CHARMM force field for nucleic acids. *Biopolymers*. 56, 257-265.
8. Abramson, J., Adler, J., Dunger, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 630, 493-500.
9. Bateman, A., Martin, M.J., Orchard, S., et al. (2023). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* 51, D523-D531.

RNA Structure Prediction in CASP16 Using Hybrid Methods

Sicheng Zhang¹, Jun Li², Yuanzhe Zhou¹, and Shi-Jie Chen^{1*}

1- Department of Physics, Department of Biochemistry, and MU Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65211, USA; 2- School of Sciences, Great Bay University, China

chenshi@missouri.edu

We use various methods, including Vfold-Pipeline¹, IsRNA^{2,3,4}, and RNAJP⁵, to generate RNA 3D structures from the sequences. If structural templates for specific motifs in the RNA targets are identified in the PDB database, we incorporate them as folding constraints. For protein/RNA/DNA complexes, RNA with ligands, and RNA with solvent shell, we employ AlphaFold3⁶, ITScore^{7,8}, and AMBER⁹ to generate structural candidates.

Methods

RNA structures:

- **Vfold-Pipeline**¹ predicts RNA 2D and 3D structures from the sequence. In this pipeline, 2D structures are predicted using Vfold2D¹⁰, a physics-based model capable of predicting 2D structures that include a variety of RNA loop motifs including pseudoknots. 3D structures are generated using the template-based models Vfold3D¹¹ and VfoldLA¹².
- **IsRNA**^{2,3,4} is a coarse-grained model for RNA 3D structure prediction based on the sequence and the 2D structure. It conducts replica exchange molecular dynamics (MD) simulations using a coarse-grained force field, which can account for correlations between various structural parameters. The simulated low-energy structures are clustered, and the centroid structures of these clusters are selected as the predicted structures.
- **RNAJP**⁵ is a coarse-grained model for RNA 3D structure prediction, with a primary focus on junction structures. Given the RNA 2D structure, it conducts Monte Carlo/MD hybrid sampling of helix arrangements with non-canonical base pairing and stacking interactions and long-range loop-loop interactions.

Protein/RNA/DNA complexes: We use AlphaFold3⁶ to generate structural candidates, subsequently employing our models to evaluate the structure candidates and using available experimental data in literature to further select the models. Additionally, we employ the ITScore^{7,8} model to dock proteins and RNAs to evaluate the complexes.

RNA with ligands: Firstly, we model the RNA-only structures; subsequently, we employ AMBER⁹ to conduct all-atom molecular simulations using the candidate RNA structures and the provided ligands to generate potential candidates. The final models are selected according to docking models and the AMBER energies and available literature information.

RNA with solvent shell: The initial RNA structure is derived from a PDB template, while the initial ion placements are based on the template, and the predictions from MCTBI¹³ and MgNet¹⁴. We subsequently conduct 550 ns of AMBER molecular simulations to sample the water and ions.

Model ranking: We rank the 3D structure candidates based on their scores and the clustering information from the simulations. Literature information, if available, was also considered for model selection and ranking.

Availability

Vfold-Pipeline is available at <http://rna.physics.missouri.edu/vfoldPipeline/index.html>

IsRNA is available at <http://rna.physics.missouri.edu/IsRNA/index.html>

RNAJP is available at <http://rna.physics.missouri.edu/RNAJP/index.html>

1. Li, J., Zhang, S., Zhang, D., & Chen, S. J. (2022). Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. *Bioinformatics*, 38(16), 4042-4043.
2. Zhang, D., & Chen, S. J. (2018). IsRNA: An iterative simulated reference state approach to modeling correlated interactions in RNA folding. *Journal of chemical theory and computation*, 14(4), 2230-2239.
3. Zhang, D., Li, J., & Chen, S. J. (2021). IsRNA1: de novo prediction and blind screening of RNA 3D structures. *Journal of chemical theory and computation*, 17(3), 1842-1857.
4. Zhang, D., Chen, S. J., & Zhou, R. (2021). Modeling noncanonical RNA base pairs by a coarse-grained IsRNA2 model. *The Journal of Physical Chemistry B*, 125(43), 11907-11915.
5. Li, J., & Chen, S. J. (2023). RNAJP: enhanced RNA 3D structure predictions with non-canonical interactions and global topology sampling. *Nucleic acids research*, 51(7), 3341-3356.
6. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 1-3.
7. Huang, S. Y., & Zou, X. (2006). An iterative knowledge-based scoring function to predict protein-ligand interactions: II. Validation of the scoring function. *Journal of computational chemistry*, 27(15), 1876-1882.
8. Huang, S. Y., & Zou, X. (2014). A knowledge-based scoring function for protein-RNA interactions derived from a statistical mechanics-based iterative method. *Nucleic acids research*, 42(7), e55-e55.
9. Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I., Brozell, S. R., Cerutti, D. S., ... & Kollman, P. A. (2021). *Amber 2021*. University of California, San Francisco.
10. Cheng, Y., Zhang, S., Xu, X., & Chen, S. J. (2021). Vfold2D-MC: a physics-based hybrid model for predicting RNA secondary structure folding. *The Journal of Physical Chemistry B*, 125(36), 10108-10118.
11. Cao, S., & Chen, S. J. (2011). Physics-based de novo prediction of RNA 3D structures. *The journal of physical chemistry B*, 115(14), 4216-4226.
12. Xu, X., & Chen, S. J. (2017). Hierarchical assembly of RNA three-dimensional structures based on loop templates. *The Journal of Physical Chemistry B*, 122(21), 5327-5335.
13. Sun, L. Z., & Chen, S. J. (2016). Monte Carlo tightly bound ion model: predicting ion-binding properties of RNA with ion correlations and fluctuations. *Journal of chemical theory and computation*, 12(7), 3370-3381.
14. Zhou, Y., & Chen, S. J. (2022). Graph deep learning locates magnesium ions in RNA. *QRB discovery*, 3, e20.

Interface-focused Scoring of Protein Assemblies in CASP16

K. Olechnovič^{1,2}, S. Grudinin¹

1 - Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France,

2 - Institute of Biotechnology, Life Sciences Center, Vilnius University, 10257 Vilnius, Lithuania

klement.olechnovic@univ-grenoble-alpes.fr, klement.olechnovic@btu.vu.lt

Key: Auto:Y; CASP_serv:Y; Templ:N; MSA:N; DeepL:Y; AF:N; EMA:Y; MD:N

In CASP16 we participated in the accuracy estimation (also known as the quality assessment - QA) category. We also participated in the MassiveFold scoring experiment.

Methods

Our server group, VifChartreuse, was running the VoroIF-jury (Voronoi tessellation-based InterFace jury) algorithm^{1,2} that utilized the new version of VoroIF-GNN³ method - VoroIF-GNN-v2. The core idea of VoroIF-GNN remained the same - to train and apply a graph attention-based graph neural network that predicts CAD-score-based⁴ local scores for interface residue-residue contacts that can be summed into global, interface-level scores. The unique aspect of the VoroIF-GNN approach is that it considers the graph of interface contacts and their adjacencies, not the graph of atoms or residues. The first version of VoroIF-GNN described contacts using the Voronoi tessellation-derived contact areas and the VoroMQA⁵ pseudoenergy potential values. For VoroIF-GNN-v2, new tessellation contact area-based pseudoenergy descriptors were derived. The original VoroMQA descriptors, derived from experimentally determined structures from the Protein Data Bank⁶ (PDB), were based on the observed and expected probabilities of contacts to occur in folded conformations. For new descriptors, additional empirically-derived probabilities were used - the observed and expected probabilities of contacts to persist in folded conformations. Those probabilities were derived from distributions of tessellation-derived contact areas in ensembles of protein structures. The main set of ensembles was derived from PDB by grouping protein chains according to 90% sequence identity clustering - this resulted in 38807 ensembles formed from 429945 chains. For an additional set of empirical probabilities, an alternative set of ensembles was taken from the IDRome dataset that was published as a part of the recent study of human intrinsically disordered proteome⁷ - the authors of that study generated ensembles by running coarse-grained simulations using the CALVADOS method. We took 16774 ensembles for chains of 60 to 600 residues in length, every ensemble contained 1000 conformations. We utilized our recently developed Voronota-LT⁸ software to rapidly compute tessellation-derived contact areas for both the PDB-derived and the IDRome-derived ensembles. We then derived the contact occurrence and persistence probabilities and transformed them into pseudoenergies (by simply applying the log function to them, so that summing transformed values can be interpreted as multiplying the corresponding probabilities). Those pseudoenergy coefficients were used to describe contact nodes in interface graphs for VoroIF-GNN-v2.

VoroIF-GNN-v2 was trained on diverse docking models of 1567 protein-protein heterodimers to predict interface CAD-score, in a fashion similar to the first version of VoroIF-GNN. The VoroIF-GNN-v2 primary output is residue-residue contact scores, which can be converted into per-residue and per-assembly scores. We applied VoroIF-GNN-v2 to every assembly model twice - before and after rebuilding side-chains using FASPR⁹. Several variants of VoroIF-GNN-v2 per-assembly scores were used to produce rankings of structural models, then those rankings were combined by the VoroIF-jury algorithm to produce the primary global scores in the [0;1] interval. For every structural model, the secondary global score was

computed without involvement of VoroIF-jury, by converting the raw VoroIF-GNN-v2 global scores into a CAD-score estimate. Similarly, for the local scores, per-residue VoroIF-GNN-v2 scores (computed before rebuilding side-chains) were converted to per-residue CAD-score estimates.

Our human group, VifChartreuseJaune, was not more human than our server group - we just swapped the primary and the secondary global scores in the output lines and used the per-residue scores computed after rebuilding side-chains. Both the VifChartreuse and the VifChartreuseJaune predictions were generated by the single pipeline implemented in the FTDMP framework.

For the MassiveFold experiment involving assembly models, the VifChartreuse results were produced using the VoroIF-jury global scores, while the VifChartreuseJaune results were produced using the VoroIF-GNN-v2 global scores. We did not expect the MassiveFold experiment to be extended to predicting non-oligomeric targets. We did not have time to develop any special new methods for assessing monomeric conformations - we simply used VoroMQA-dark¹⁰ and VoroMQA-light⁵ global scores to produce results for the VifChartreuse and VifChartreuseJaune groups, respectively.

Availability

The FTDMP, VoroMQA and Voronota-LT software is freely available on GitHub at <https://github.com/kliment-olechnovic>.

1. Olechnovič,K., Valančauskas,L., Dapkūnas,J., Venclovas,Č. (2023) Prediction of protein assemblies by structure sampling followed by interface-focused scoring. *Proteins* 91(12):1724-1733.
2. Olechnovič,K., Banciu,R., Dapkūnas,J., Venclovas,Č. (2024) FTDMP: a framework for protein-protein, protein-DNA and protein-RNA docking and scoring. Submitted to *Proteins* (CAPRI special issue).
3. Olechnovič,K., Venclovas,Č. (2023) VoroIF-GNN: Voronoi tessellation-derived protein-protein interface assessment using a graph neural network. *Proteins* 91(12):1879-1888.
4. Olechnovič,K., Venclovas,Č. (2020) Contact Area-Based Structural Analysis of Proteins and Their Complexes Using CAD-Score. *Methods Mol Biol* 2112, 75–90.
5. Olechnovič,K., Venclovas,Č. (2017) VoroMQA: Assessment of protein structure quality using interatomic contact areas. *Proteins* 85, 1131–1145.
6. Burley,S.K., Bhikadiya,C., Bi.C, et al. (2023) RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res.* 2023;51(D1):D488-D508.
7. Tesei,G., Trolle,A.I., Jonsson,N. et al. (2024) Conformational ensembles of the human intrinsically disordered proteome. *Nature* 626(8000):897-904.
8. Olechnovič,K., Grudinin,S. (2024) Voronota-LT: efficient, flexible and solvent-aware tessellation-based analysis of atomic interactions. Preprint in bioRxiv.
9. Huang,X., Pearce,R., Zhang,Y. (2020) FASPR: an open-source tool for fast and accurate protein side-chain packing. *Bioinformatics* 36:3758-3765.
10. Dapkūnas,J., Olechnovič,K., Venclovas,Č. (2021) Modeling of protein complexes in CASP14 with emphasis on the interaction interface prediction. *Proteins* 89(12):1834-1843.

Protein-Ligand Binding Affinity Prediction Using Descriptors Derived From Voronoi Tessellation

A. Kudrevceva, J. Dapkūnas

Institute of Biotechnology, Life Sciences Center, Vilnius University, Lithuania

justas.dapkus@bt.vu.lt

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N, EMA:N; MD:N

In CASP16 we tested our protein-ligand binding affinity predictors. We submitted predictions for CASP targets L1000 (chymase, 17 protein-ligand complexes) and L3000 (autotaxin, 93 protein-ligand complexes) obtained using the experimental structures provided by CASP as input.

Methods

Voronoi tessellation-based affinity predictors. We developed models for prediction of protein-ligand binding affinity using descriptors derived from Voronoi tessellation¹ of protein-ligand complexes: interatomic contact areas, molecular volume change upon complex formation and number of formed hydrogen bonds. To reveal the relationship between computed descriptors and binding affinity we used experimentally determined protein-ligand complex structures with known affinity collected in the PDBbind database² and different machine learning algorithms. For CASP16 we selected the best model based on evaluation on diverse test sets and submitted its predictions as VoroAffinity group.

Target-specific affinity predictors. Knowing the structures of CASP16 targets, we developed target-specific affinity predictors. We used the same modeling approach with training sets supplemented by data on proteins structurally similar (TM-score > 0.8) to the target. We extracted these additional protein-ligand complexes with known binding affinity from the PDBbind database² and from the Protein Data Bank. We used a dataset of protein-ligand complexes similar to L1000 to train a chymase-specific model. Because of the lack of data for proteins similar to autotaxin we trained L3000-specific model using the standard training set enriched by 23 autotaxin structures with different ligands.

To estimate the accuracy of developed models we performed leave-one-out training and testing using the structures of chymase and autotaxin protein-ligand complexes. The accuracy of chymase-specific model was additionally evaluated using external trypsin dataset³. We observed improved prediction accuracy of the target-specific models in comparison to models trained on standard dataset and submitted their predictions for CASP16 targets as VoroAffinityB group.

Prediction for CASP targets. For both VoroAffinity and VoroAffinityB groups we selected ligand bound closer to the catalytic site or weighted the predicted affinity based on occupancy when more than one ligand was present in the provided experimental structure.

1. Olechnovič,K., & Venclovas,Č. (2014). Voronota: A fast and reliable tool for computing the vertices of the Voronoi diagram of atomic balls. *J Comput Chem.* **35**, 672–681.
2. Liu,Z., Li,Y., Han,L., Li,J., Liu,J., Zhao,Z., Nie,W., Liu,Y., & Wang,R. (2015). PDB-wide collection of binding data: Current status of the PDBbind database. *Bioinformatics.* **31**, 405–412.
3. Durant,G., Boyles,F., Birchall,K., Marsden,B., & Deane,C.M. (2023). Robustly interrogating machine learning-based scoring functions: What are they learning? *BioRxiv.* 2023.10.30.564251 doi: <https://doi.org/10.1101/2023.10.30.564251>

Improved Massive Sampling using AFsample2

Yogesh Kalakoti and Björn Wallner

Division of Bioinformatics, IFM, Linköping University

bjorn.wallner@liu.se

Key: Auto:Y; CASP_serv:N; Templ:N; MSA:Y; DeepL:Y; EMA:Y; MD:N

We participated in CASP16 using an improved MSA sampling protocol called AFsample2¹. AFsample2 masks a random fraction of the MSA columns to perturb the input MSA before AlphaFold2^{2,3} inference with dropout⁴ activated. AFsample2 was initially developed to predict multi-state proteins, but it also works very well to improve sampling for difficult single-state proteins. The method is completely automated but was run as a manual server to allow for more computational time. We participated in the single proteins, complexes, and ensembles categories.

Methods

To improve sampling, we used AFsample2 with 0, 0.1, 0.2, or 0.3 fractions of the MSA columns masked. The 0 fraction of the MSA column is identical to running AF2 with dropout activated, which was so successful in CASP15⁵, while increasing the fraction will result in sampling with less information. In our benchmarks, going beyond 0.3 deteriorates the predictions too much. For each fraction, all available versions of the AF2 neural networks were used; for monomers, five networks, both the networks with and without ptm head were used, as well as all three versions of the multimer networks (a monomer is a multimer with one chain); for multimers, all three multimer networks were used. In addition, we also run with the pipeline with 3 and 20 recycles. Templates were not used at all.

For Phase 1 targets (T1*/H1*), the goal was to generate at least 100 models per network per fraction of column masking; since we run with four different fractions, there are five networks per version, and we used two recycle settings, this resulted in a minimum of 20,000 models (100x4x5x5) for monomers and 12,000 (100x4x5x3x2) models for multimers. In reality, the number of generated models was between 1,416 and 200,307, with a median of 36,395 models, and only ten targets had less than 12,000 models.

When AFsample2 perturbs the MSA, the model confidence drops. This drop is not necessarily coupled to worse models but rather an effect of the perturbation. To be able to compare the model confidences from all generated models, the top 10 models by model confidence from each MSA column masking fraction and neural network version were rescored using AF_unmasked⁶. The rescored uses inference with the ‘multimer_5_v2’ neural network with the input model as a template and no MSA information. If the length of the input model was shorter than 1,400 amino acids, the model was relaxed using the standard AF2 Amber relaxation protocol before rescored. Five models were generated, and the best model confidence was used as the score for the input model. For models larger than 1,400 amino acids, the unrelaxed model was used, and only one model was generated.

Finally, a ranking-based clustering was performed to avoid submitting too similar models: The model with the best score was always submitted as rank 1; lower ranks were only submitted if it was more distant than 2 angstroms for monomer and DockQ<0.5 for multimers to any higher-ranking models.

For Phase 2 targets (T2*/H2*), no new models were generated. Instead, all models submitted as in Phase 1 (T1*/H1*) to CASP were downloaded from the CASP website, rescored and refined using AF_unmasked. The model with the highest model confidence was submitted as TS1, followed by the same ranking-based clustering described above, to avoid submitting too similar models.

For targets with two different conformations, all models were compared to the first-ranked model using a reference-free plot¹ in which the TMscore⁷ or DockQ⁸ is plotted against the model confidence from AF_unmasked. The first-ranked model was submitted as one conformation, and the second conformation was selected from the reference-free plot as a model with high confidence and a significant structural difference from the first-ranked model.

For the ensemble targets, models were generated in the same as above. All models were clustered using Rosetta Clustering⁹ with a 5 Angstrom threshold. For each cluster, the highest model confidence, passing the *ad hoc* and very strict similarity filters, was taken as representative. The probability for each representative was taken as the relative size of the respective cluster.

The MSAs were generated using the default AF2 settings against the default BFD database, Uniref90 and Uniprot downloaded 20240521, and Mgnify version 2022_05. No templates were used.

Availability

AFsample2 is available at <http://wallnerlab.org/AFsample2>

1. Kalakoti, Y. & Wallner, B. AFsample2: Prediction of conformational states and transitions using improved sampling with AlphaFold. *bioRxiv* 2024.03.20.536385 (2024) doi:10.1101/2024.03.20.536385.
2. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 1–11 (2021) doi:10.1038/s41586-021-03819-2.
3. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *Biorxiv* 2021.10.04.463034 (2021) doi:10.1101/2021.10.04.463034.
4. Wallner, B. AFsample: improving multimer prediction with AlphaFold using massive sampling. *Bioinformatics* **39**, btad573 (2023).
5. Wallner, B. Improved multimer prediction using massive sampling with AlphaFold in CASP15. *Proteins: Struct., Funct., Bioinform.* (2023) doi:10.1002/prot.26562.
6. Mirabello, C., Wallner, B., Nystedt, B., Azinas, S. & Carroni, M. Unmasking AlphaFold: integration of experiments and predictions in multimeric complexes. *bioRxiv* 2023.09.20.558579 (2023) doi:10.1101/2023.09.20.558579.
7. Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins Struct Funct Bioinform* **57**, 702–710 (2004).

8. Mirabello, C. & Wallner, B. DockQ v2: Improved automatic quality measure for protein multimers, nucleic acids, and small molecules. *bioRxiv* 2024.05.28.596225 (2024) doi:10.1101/2024.05.28.596225.
9. Leaver-Fay, A. *et al.* ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* **487**, 545–574 (2011).

A protein-protein complex structure prediction model via MSA and deep learning

Chunqiu Xia, Yutao Mi, Chen Wang, Hongning Zhang, Xi Chen, Zongquan Li, Leilei Hou, Liu Jing Wang, Hongmei Li, Jingming Zhang, Sheng lin, and Yunfei Long

Ailux Biologics of XtalPi, Shanghai, China

chunqiu.xia@xtalpi.com, sheng.lin@xtalpi.com, yunfei.long@xtalpi.com

Key: Auto:N; CASP_serv:Y; TempL:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:Y

In CASP16, our group participated in the protein-protein complex structure modeling category. Deep learning methods have demonstrated a marked enhancement in the predictive accuracy of protein complex structures. AlphaFold-Multimer¹(AFM) has set a good start for deep learning methods to predict the structure of protein complexes. Inspired by many previous methods, we developed a new method for predicting protein-protein complex structures that organically integrates the advantages of previous methods.

Methods

Our method adopts the multiple sequence alignment (MSA) and structure templates as its input. MSA is retrieved by MMSeqs2² and another method against UniRef30³ and several metagenomic databases, including BFD and MGnify^{4,5}. Then, templates are retrieved against PDB100 by HHSearch⁶. Similar to AFM, only 20 structure templates will be employed for the purpose of modeling. After deriving enough diverse conformations, we cluster all the structures based on their structural similarity and structural confidence and submit the top-K structures with high confidence, where K is constrained to a maximum of five. For large protein complexes, our base method is additionally fused with automated assembly algorithmic modules and expert knowledge.

Availability

The source codes and models are not publicly available.

1. Evans,R. et al. (2022) Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.10.04.463034.
2. Steinegger,M. and Söding,J. (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nature Biotechnology, 35, 1026–1028.
3. Suzek,B.E. et al. (2015) UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics, 31, 926–932.
4. Steinegger,M. et al. (2019) Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. Nature Methods, 16, 603–606.
5. Mitchell,A.L. et al. (2020) MGnify: the microbiome analysis resource in 2020. Nucleic Acids Research, 48, D570–D578.
6. Steinegger,M. et al. (2019) HH-suite3 for fast remote homology detection and deep protein annotation. BMC Bioinformatics, 20, 473.

Protein and RNA structure prediction with trRosetta, trRosettaRNA and AlphaFold

Wenkai Wang¹, Zhenling Peng¹ and Jianyi Yang^{1,*}

¹ MOE Science Center for Nonlinear Expectations, Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao 266237, China.

*corresponding author: yangjy@sdu.edu.cn

Key: *Templ:Y; MSA:Y.MetaG; Dist:Y; Tors:Y; DeepL:Y; MD:N; AF:Y; EMA:N; CASP_serv:N*

Methods

Protein monomers. Monomeric protein targets were predicted with similar methods we used in CASP15¹. Multiple sequence databases (unclust30_2018, uniref30, bfd and manually collected sequences) were searched by HHblits and MMseqs2, generating a set of candidate multiple sequence alignments (MSAs). Optimal MSAs are selected by trRosettaX2 according to the probability of the top residue pairs in the predicted inter-residue distance matrix. For monomeric targets derived from multimeric targets, their structures are derived from the complex models predicted by AlphaFold-Multimer² and/or AlphaFold3³.

Protein multimers. AlphaFold-Multimer² and/or AlphaFold3³ are used to predict the structures for protein multimers. For AlphaFold-Multimer, optimal monomeric MSAs are prepared using the pipeline used for monomer structure prediction. For AlphaFold3, we submitted the sequences to its web server to predict the complex structure models. Top models are then selected for submission based on the predicted confidence scores.

RNAs. trRosettaRNA⁴ and its improved version are used to predict RNA structures. For the improved trRosettaRNA, an end-to-end version is used to predict an initial structure and inter-nucleotide distance and orientation matrices. The unrealistic geometry (e.g., broken bond, steric clashes, etc) of the end-to-end model is fixed by energy minimization, similar to the original trRosettaRNA.

Availability:

<https://yanglab.nankai.edu.cn/trRosetta/>,
<https://yanglab.nankai.edu.cn/trRosettaRNA>

1. Peng, Z., Wang, W., Wei, H., Li, X. & Yang, J. Improved protein structure prediction with trRosettaX2, AlphaFold2, and optimized MSAs in CASP15. *Proteins* 91, 1704-1711 (2023).
2. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *BioRxiv*, 2021.2010.2004.463034 (2022).
3. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493-500 (2024).
4. Wang, W. et al. trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat Commun* 14, 7266 (2023).

Protein and Nucleic Acid Structure Prediction by Zheng Group in CASP16

Wei Zheng^{1,2,3}

1 - Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

2 - Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

3 - School of Statistics and Data Science, Nankai University, 94 Weijin Road, 300071, Tianjin, China

zhengwei@umich.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:N; MD:Y

The ‘Zheng’ human group in CASP16 made use of a flexible set of workflows for modeling various targets, including protein monomers, nucleic acids (DNA or RNA), nucleic acid-nucleic acid complexes, protein-protein complexes, protein-nucleic acid complexes, and ensemble targets. Protein monomer targets are predicted using an updated version of the D-I-TASSER algorithm¹, similar to the ‘Zheng-Server’ but with extended simulation time. Protein-protein complex targets are modeled using a modified version of DMFold-Multimer², similar to ‘Zheng-Multimer’ but with more diverse MSA inputs. Nucleic acid-containing targets (with or without protein) are predicted using the newly developed deep learning method, DeepProtNA, which integrates pre-trained language model embeddings, multiple sequence alignments (MSA), predicted secondary structures, and structural templates as inputs for a set of modified Evoformer blocks.

Methods

Domain boundary prediction. To address the challenge of modeling multi-domain targets, we introduced a newly developed domain boundary prediction method. The domain boundary predictor contains a set of modified Evoformer blocks, with input features including threading template information from LOMETS3, contact/distance maps predicted by DeepPotential³, domain boundary scores from ThreaDom⁴, residue-level embeddings from several pre-trained protein language models (pLM), continuous fragment information obtained from MSA, and folding unit scores from FUpred⁵, all of which are combined by the model to predict domain boundary information.

MSA construction. The MSA construction pipeline used in the ‘Zheng’ workflow is built on DeepMSA2 with two key improvements: (i) a larger in-house metagenomic sequence database, incorporating expansive curated sequence data from IMG/M, NCBI, and EBI, and (ii) a multi-domain MSA assembly method that merges domain-level MSAs into a full chain-level MSA. Similar to DeepMSA2, the new pipeline also contains three MSA construction sub-pipelines: dMSA, qMSA, and mMSA, and diverse genomic and metagenomic databases, including Uniclust30, UniRef90, Metaclust, Mgnify, BFD and an in-house huge metagenomics database. MSAs generated from these sub-pipelines are input into AlphaFold2 to predict a set of models, and the MSAs are then ranked by their associated *pLDL* scores. For multi-domain targets, the same MSA generation method is used to construct domain-level MSAs based on predicted domain boundaries, which are then assembled into full-length MSAs by linking

sequences from the same species. The ranked MSAs are either directly used in protein homomer modeling (top-1 MSA) or paired as multimer MSAs for protein complex modeling. For heteromeric complexes, an additional selection process generates an optimal set of paired MSAs by combining individual constituent MSAs. The top N ranked MSAs for each constituent protein are chosen to form potential paired MSAs, and for a heteromeric complex with M constituent proteins, N^M distinct paired MSAs are generated and evaluated based on a combined score of the depth of the MSAs and pLDDT score of the monomer chains. To ensure the pipeline completes within two weeks, N is selected to satisfy $N^M \leq 150$.

Template detection. The template detection method is based on a new version of LOMETS (LOMETs4). Compared to LOMETs3⁶, which was used in CASP15, the major update in LOMETs4 is its ability to handle protein complexes. For protein heteromers, templates are identified as follows: first, homologous templates for each constituent chain in the target complex are identified using LOMETs3, which includes six profile-based threading methods, five contact/distance-based threading methods, three protein language model-based threading methods, and one structure-based threading method. Notably, templates for individual chains that have already been considered in previous steps are excluded to prevent the similar query constituent chain from hitting identical templates. The templates are ranked by quality (e.g., Z-score). Finally, if at least two constituent chains share templates from the same protein complex and have a high summed Z-score, these complexes are considered potential templates.

Protein monomer modeling. Protein monomer models are generated by D-I-TASSER using Replica-Exchange Monte Carlo (REMC) simulations, guided by knowledge-based potentials and deep learning-predicted residue-residue contact maps, distance distributions, inter-residue torsion angles, and hydrogen-bond networks, as well as the similar type restraints calculated from LOMETs4 threading templates.

Protein-protein complex modeling. Protein complex models are predicted using a modified AlphaFold2 modeling engine with our MSAs and structure templates as input features. Key modifications to the modeling engine include: (i) using templates or not, (ii) adjusting the dropout rate, (iii) applying different versions of AlphaFold2 pre-trained weights (v1-v3), (iv) generating a higher number of decoys than the default setting (25 models), (v) applying or omitting the early stop strategy in AlphaFold2 (v2.3), and (vi) extending the modeling iterations. The final models are ranked by confidence scores (e.g., $0.8pTM + 0.2ipTM$). For targets with unknown stoichiometry, a newly developed method is used to determine stoichiometry information. This method uses three pipelines: one derives stoichiometry from top-ranked LOMETs4 templates, another predicts stoichiometry using a deep learning method combining sequence data and embeddings from protein language models⁷ and text-based language models using UniProt descriptions, and the last pipeline is predicted directly by DMFold confidence scores, where the oligomeric state is predicted based on the model with the highest confidence.

Nucleic acid-related target modeling. DeepProtNA is an end-to-end deep learning algorithm designed for predicting oligomeric structures. The method integrates pre-trained language model embeddings, MSA information, predicted secondary structure, and structural templates to directly generate three-dimensional coordinates of the complexes from input sequences through a set of modified Evoformer blocks and a structure module similar to AlphaFold2. Protein and RNA sequences are input into pre-trained language models, ESM⁷ and RNA-FM⁸ (respectively), to generate high-dimensional sequence embeddings. MSAs are generated for both protein and nucleic acid sequences using the modified version DeepMSA2² and rMSA⁹. The MSA-derived features are combined with the language model embeddings to enhance the

information content available for predicting interactions between the sure sequences of sequence interactions. Structural templates are selected from the PDB by LOMETS4 for protein or BLASTn for RNA. Additionally, the predicted secondary structure of nucleic acids is also used as an input of the deep learning networks. The main trunk of the DeepProtNA processes the embeddings for proteins and nucleic acids, leveraging self-attention mechanisms to capture long-range dependencies within each sequence. Cross-attention mechanisms are used to handle interactions between protein and nucleic acid sequences or between multiple chains of the same type. A structure decoder network takes the embedded representation to generate the three-dimensional coordinates for the complex. This network translates sequence information directly into spatial coordinates for each residue and nucleotide, producing the final structure. DeepProtNA also outputs a confidence score that measures the reliability of the predicted interactions, and the confidence score is also used for predicting the oligomer states for those RNA targets with unknown stoichiometry.

Ensemble modeling. For ensemble targets, after initial generation of a series of models (decoy) using the workflows described above, the SPICKER¹⁰ method is used to cluster the decoys based on structural similarity for protein monomers, RNA, and protein complexes. The centers of clusters with a large number of members and high confidence scores are selected as models representing potential alternative conformations.

1. Zheng,W., Wuyun,Q., Freddolino,P.L. & Zhang,Y. (2023). Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15. *Proteins: Structure, Function, and Bioinformatics* 91, 1684-1703.
2. Zheng,W., Wuyun,Q., Li,Y., Zhang,C., Freddolino,P.L. & Zhang,Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods* 21, 279-289.
3. Li,Y., Zhang,C., Zheng,W., Zhou,X., Bell,E.W., Yu,D.-J. & Zhang,Y. (2021). Protein inter-residue contact and distance prediction by coupling complementary coevolution features with deep residual networks in CASP14. *Proteins: Structure, Function, and Bioinformatics* 89, 1911-1921.
4. Xue,Z., Xu,D., Wang,Y. & Zhang,Y. (2013). ThreaDom: extracting protein domain boundary information from multiple threading alignments. *Bioinformatics* 29, i247-i256.
5. Zheng,W., Zhou,X., Wuyun,Q., Pearce,R., Li,Y. & Zhang,Y. (2020). FUpred: detecting protein domains through deep-learning-based contact map prediction. *Bioinformatics* 36, 3749-3757.
6. Zheng,W., Wuyun,Q., Zhou,X., Li,Y., Freddolino,P.L. & Zhang,Y. (2022). LOMETs3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic Acids Research* 50, W454-W464.
7. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., dos Santos Costa,A., Fazel-Zarandi,M., Sercu,T., Candido,S. & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123-1130.
8. Chen,J., Hu,Z., Sun,S., Tan,Q., Wang,Y., Yu,Q., Zong,L., Hong,L., Xiao,J., Shen,T., King,I. & Li,Y. (2022). Interpretable RNA Foundation Model from Unannotated Data for Highly Accurate RNA Structure and Function Predictions. *bioRxiv*, 2022.08.06.503062.
9. Zhang,C., Zhang,Y. & Pyle,A.M. (2023). rMSA: A Sequence Search and Alignment Algorithm to Improve RNA Structure Modeling. *Journal of Molecular Biology* 435, 167904.
10. Zhang,Y. & Skolnick,J. (2004). SPICKER: a clustering approach to identify near-native protein folds. *Journal of computational chemistry* 25, 865-871.

Zheng-Multimer

Protein Complex Structure Prediction Using Deep Learning and Multi-MSA Strategy in CASP16

Wei Zheng^{1,2}, Qiqige Wuyun³, Quancheng Liu^{1,2}, Ziying Zhang⁴, and Lydia Freddolino^{1,2}

1 - Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

2 - Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

3 - Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

4 - College of Information Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China

zhengwei@umich.edu, lydsf@umich.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:N; MD:N

The protein structure prediction approach of the Zheng-Multimer server group in CASP16 is an extension of the DMFold¹ pipeline. It consists of two sub-pipelines: DMFold-Monomer for protein monomer modeling and DMFold-Multimer for protein complex structure prediction. Compared to the DMFold pipeline used in CASP15, the new method includes four major updates: (i) a novel template detection method, LOMETS4, for collecting templates for both protein monomers and multimers, (ii) an updated DeepMSA pipeline (including substantially expanded databases) for constructing MSAs, (iii) a more comprehensive structure sampling engine to generate diverse structure models using a modified AlphaFold2² structure module, and (iv) a stoichiometry determination module for predicting stoichiometry information.

Methods

The first step of Zheng-Multimer is MSA construction, which builds on the previous DeepMSA2 pipeline. Compared to the version used in CASP15, this updated DeepMSA introduces two key improvements: (i) a larger metagenomic sequence database, incorporating data from IMG/M, NCBI, and EBI, and (ii) a multi-domain MSA assembly method that assembles domain-level MSAs into a full chain-level MSA. Similar to DeepMSA2, the new pipeline also contains three MSA construction sub-methods: dMSA, qMSA, and mMSA. In dMSA, HHblits³, Jackhmmer⁴ and HMMsearch⁴ are used to search the query sequence against the UniClust30⁵, UniRef90⁶ and Metaclust⁷ databases in three stages (labeled stage 1-3 in the order listed above). qMSA is an extended version of dMSA with a new search added between stage 2 and stage 3 of dMSA, where HHblits is used to search the BFD⁸ metagenomic database. In addition, a new iteration stage (stage 4) is added in qMSA to search the query through the Mgnify⁹ metagenomic database. In mMSA, MSA from qMSA stage3 is used as the starting point for HMMsearch to search through the huge in-house metagenome database mentioned above. MSAs generated from dMSA, qMSA, and mMSA are input into AlphaFold2 (1-embedding) to predict a set of models. Those MSAs are then ranked by the associated *pLDDT* scores from AlphaFold2. For multi-domain targets, the same MSA generation method is applied to construct domain-level MSAs based on the predicted domain boundaries. These domain-level MSAs are subsequently assembled into full-length MSAs by linking sequences from the same species. The ranked MSAs are either directly used in protein monomer modeling or paired as

multimer MSAs for protein complex modeling. For heteromeric complexes, an additional selection procedure is employed to generate an optimal set of paired MSAs by combining the individual constituent MSAs. The top N ranked MSAs for each constituent protein are chosen to form potential paired MSAs. Each selected MSA for one constituent protein is paired with the MSA of another constituent. For a heteromeric complex containing M different constituent proteins, N^M distinct paired MSAs are generated and evaluated based on a combined score of the depth of the MSAs and pLDDT score of the monomer chains. To guarantee the pipeline could be completed within three days, N is selected as the maximal value to satisfy $N^M \leq 64$.

The second step of Zheng-Multimer is template detection, which is based on a new version of LOMETs. Compared to LOMETs3¹⁰, which was used in CASP15, the major update in LOMETs4 is its ability to handle protein complexes. Specifically, for protein heteromers, templates are identified as follows: first, for each constituent chain in the target complex, homologous templates are identified using LOMETs3, which includes six profile-based threading methods, five contact/distance-based threading methods, three protein language model-based threading methods, and one structure-based threading method. Notably, for each pipeline, templates for individual chains that have already been considered in previous steps are excluded to prevent the similar query constituent chain from hitting identical templates. Next, the templates for each chain are ranked based on their quality (e.g., Z-score). Finally, if at least two constituent chains share templates originating from the same protein complex, those complexes are considered potential complex templates. For protein monomers or homomeric complexes, LOMETs4 monomer templates are directly output and used in the structure model generation step.

The third step in Zheng-Multimer is model generation, which utilizes a modified version of the AlphaFold2 modeling engine. The MSAs from the first step and the structure templates from the second step serve as input features for this modeling engine. Key modifications to the AlphaFold2 modeling engine include: (i) using templates or not, (ii) adjusting the dropout rate, (iii) applying different versions of AlphaFold2 pre-trained weights (v1-v3), (iv) generating a higher number of decoys than the default setting (25 models), (v) applying or omitting the early stop strategy in AlphaFold2 (v2.3), and (vi) increasing the number of modeling iterations. The final models are ranked based on the *pLDDT* score for monomer targets, or by confidence scores ($0.8pTM + 0.2ipTM$) for complex targets.

For targets with unknown stoichiometry, a newly developed method is used to predict stoichiometry information and prioritize modeling efforts. This method incorporates two fast pipelines to detect potential stoichiometry states: one set of stoichiometry is derived from the top-ranked LOMETs4 templates, and another set of stoichiometry is predicted from a deep learning method that combines sequence data, embedding from protein language models¹¹, and embedding from text-based language models utilizing UniProt descriptions as input. DMFold is then used to model all highly ranked stoichiometry states predicted by these two pipelines, with the models and associated stoichiometry states selected based on the highest confidence scores.

1. Zheng,W., Wuyun,Q., Freddolino,P.L. & Zhang,Y. (2023). Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15. *Proteins: Structure, Function, and Bioinformatics* 91, 1684-1703.
2. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., Back,T., Petersen,S., Reiman,D., Clancy,E., Zielinski,M., Steinegger,M., Pacholska,M., Berghammer,T., Bodenstein,S., Silver,D., Vinyals,O.,

- Senior,A.W., Kavukcuoglu,K., Kohli,P. & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
3. Remmert,M., Biegert,A., Hauser,A. & Söding,J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods* 9, 173-175.
 4. Potter,S.C., Luciani,A., Eddy,S.R., Park,Y., Lopez,R. & Finn,R.D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research* 46, W200-W204.
 5. Mirdita,M., von den Driesch,L., Galiez,C., Martin,M.J., Söding,J. & Steinegger,M. (2017). UniClust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Research* 45, D170-D176.
 6. Suzek,B.E., Wang,Y., Huang,H., McGarvey,P.B., Wu,C.H. & UniProt,C. (2015). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* (Oxford, England) 31, 926-932.
 7. Steinegger,M. & Söding,J. (2018). Clustering huge protein sequence sets in linear time. *Nature Communications* 9, 2542.
 8. Steinegger,M., Mirdita,M. & Söding,J. (2019). Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nature Methods* 16, 603-606.
 9. Mitchell,A.L., Almeida,A., Beracochea,M., Boland,M., Burgin,J., Cochrane,G., Crusoe,M.R., Kale,V., Potter,S.C., Richardson,L.J., Sakharova,E., Scheremetjew,M., Korobeynikov,A., Shlemov,A., Kunyavskaya,O., Lapidus,A. & Finn,R.D. (2020). MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Research* 48, D570-D578.
 10. Zheng,W., Wuyun,Q., Zhou,X., Li,Y., Freddolino,P.L. & Zhang,Y. (2022). LOMETS3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic Acids Research* 50, W454-W464.
 11. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., dos Santos Costa,A., Fazel-Zarandi,M., Sercu,T., Candido,S. & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123-1130.

Zheng-Server

Prediction of Protein Tertiary Structures Utilizing Deep Learning Spatial Restraints and Knowledge-Based Potentials in CASP16

Wei Zheng^{1,2}, Qiqige Wuyun³, Chunxiang Peng^{1,2}, Xiaogen Zhou⁴, and Lydia Freddolino^{1,2}

1 - Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

2 - Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

3 - Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

4 - College of Information Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China

zhengwei@umich.edu, lydsf@umich.edu

Key: Auto:Y; CASP_serv:N; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:N; MD:Y

The Zheng-Server utilized in CASP16 is designed for protein tertiary structure prediction through a fully automated pipeline, expanded from D-I-TASSER¹. The pipeline consists of six key modules: (i) domain boundary prediction using a novel deep learning framework, (ii) multiple sequence alignment (MSA) construction through the newly developed and expanded DeepMSA2² pipeline, (iii) template detection through the recently updated LOMETS³, (iv) initial model generation and spatial geometric restraint prediction using AlphaFold2 and advanced deep learning predictors, (v) conformation generation through I-TASSER Replica-Exchange Monte Carlo (REMC) simulation guided by deep learning constraints, and (vi) atomic-level model refinement using molecular dynamics (MD) simulation.

Methods

Zheng-Server follows the general workflow of recent versions of D-I-TASSER, with several additional refinements to address especially difficult targets. To address the challenge of modeling multi-domain targets, Zheng-Server incorporates a newly developed domain boundary prediction method. The input features for the modified Evoformer blocks include threading template information from LOMETS3, contact/distance maps predicted by DeepPotential⁴, domain boundary scores from ThreaDom⁵, residue-level embeddings from pre-trained protein language models (pLM), continuous fragment information obtained from MSA, and folding unit scores from FUpred⁶, all of which are utilized to predict domain boundaries, which are then used to perform domain-level modeling prior to final assembly.

The MSA generation pipeline in Zheng-Server is an extension of DeepMSA2. Compared to the version used in CASP15, this updated DeepMSA introduces two key improvements: (i) a larger metagenomic sequence database, incorporating data from IMG/M, NCBI, and EBI, and (ii) a multi-domain MSA assembly method that assembles domain-level MSAs into a full chain-level MSA. Similar to DeepMSA2, the new pipeline also contains three MSA construction sub-methods: dMSA, qMSA, and mMSA. In dMSA, HHblits, Jackhmmer and HMMsearch are used to search the query sequence against the UniClust30, UniRef90 and Metaclust databases in three stages (labeled stage 1 – stage 3 in the order listed above). qMSA is an extended version of dMSA with a new search added between stage 2 and stage 3 of dMSA, where HHblits is used to search the BFD metagenomic database. In addition, a new iteration stage (stage 4) is added in qMSA to search the query through the Mgnify metagenomic database. In mMSA,

MSA from qMSA stage3 is used as the starting point for HMMsearch to search through the huge in-house metagenome database mentioned above. MSAs generated from dMSA, qMSA, and mMSA are input into AlphaFold2 (1-embedding) to predict a set of models. The MSA associated with the highest pLDDT score from the AlphaFold2 models is selected as the final output. For multi-domain targets, the same MSA generation method is applied to construct and select domain-level MSAs based on the predicted domain boundaries. These domain-level MSAs are then assembled into a full-length MSA by linking sequences from the same species.

The final MSA is used as input for AlphaFold⁷ (8-embedding), OpenFold⁸, UniFold⁹, ColabFold¹⁰, RosettaFold¹¹, ESMFold¹², OmegaFold¹³, AttentionPotential, and DeepPotential⁴ for the predictions of residue-residue contact maps, distance distributions, inter-residue torsion angles, and hydrogen-bond networks. Those deep learning-predicted restraints are utilized to guide the REMC folding simulation with the same set of restraints calculated from templates detected by new LOMETS. AttentionPotential is an extended pipeline from DeepPotential, which utilizes an MSA transformer architecture. The full sets of predicted restraints from AttentionPotential and DeepPotential are then fed into DeepFold, an L-BFGS folding system, to produce ten full-length models. Those ten models, as well as five models generated by each of the other structure prediction methods noted above, and full-chain level threading templates from the new LOMETS (see below), are used as initial conformations in the REMC folding simulation.

The MSA generated from the new DeepMSA method is also used to produce sequence profiles or profile Hidden Markov Models (HMMs) to be utilized by six profile-based threading methods employed by the new version of LOMETS. Additionally, the contact maps and distance distributions predicted from the deep learning predictors are used by five contact- and distance-based threading methods. In addition, with six profile-based threading methods and five distance-based threading methods, three pLM-based threading methods are also introduced into the new LOMETS. Finally, 140 full-chain level templates (10 templates from each component threading method) are collected by LOMETS, and then used as initial conformations in the REMC simulation as noted above.

For target proteins with lengths of less than 300 residues, an I-TASSER-based REMC simulation is utilized for generating 10,000 decoy conformations. The REMC simulation is guided by knowledge-based potentials and residue-residue contact maps, distance distributions, inter-residue torsion angles, and hydrogen-bond networks that are predicted by deep learning predictors and calculated from LOMETS threading templates (as noted above). The decoys are then clustered using SPICKER¹⁴ to obtain five clusters for final model selection. For the targets with lengths of greater than 300 residues, the top five ranked models by deep learning predictors are directly used in the next MD refinement, without the REMC stage.

The five cluster centroids (for target with length<300AA) or the five top ranked deep learning models (for target with length≥300AA) are further refined by FG-MD¹⁵ to remove steric clashes and refine the local structure packing, resulting in the final models.

1. Zheng,W., Wuyun,Q., Freddolino,P.L. & Zhang,Y. (2023). Integrating deep learning, threading alignments, and a multi-MSA strategy for high-quality protein monomer and complex structure prediction in CASP15. *Proteins: Structure, Function, and Bioinformatics* 91, 1684-1703.

2. Zheng,W., Wuyun,Q., Li,Y., Zhang,C., Freddolino,P.L. & Zhang,Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods* 21, 279-289.
3. Zheng,W., Wuyun,Q., Zhou,X., Li,Y., Freddolino,P.L. & Zhang,Y. (2022). LOMETS3: integrating deep learning and profile alignment for advanced protein template recognition and function annotation. *Nucleic Acids Research* 50, W454-W464.
4. Li,Y., Zhang,C., Zheng,W., Zhou,X., Bell,E.W., Yu,D.-J. & Zhang,Y. (2021). Protein inter-residue contact and distance prediction by coupling complementary coevolution features with deep residual networks in CASP14. *Proteins: Structure, Function, and Bioinformatics* 89, 1911-1921.
5. Xue,Z., Xu,D., Wang,Y. & Zhang,Y. (2013). ThreaDom: extracting protein domain boundary information from multiple threading alignments. *Bioinformatics* 29, i247-i256.
6. Zheng,W., Zhou,X., Wuyun,Q., Pearce,R., Li,Y. & Zhang,Y. (2020). FUpred: detecting protein domains through deep-learning-based contact map prediction. *Bioinformatics* 36, 3749-3757.
7. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., Bridgland,A., Meyer,C., Kohl,S.A.A., Ballard,A.J., Cowie,A., Romera-Paredes,B., Nikolov,S., Jain,R., Adler,J., Back,T., Petersen,S., Reiman,D., Clancy,E., Zielinski,M., Steinegger,M., Pacholska,M., Berghammer,T., Bodenstein,S., Silver,D., Vinyals,O., Senior,A.W., Kavukcuoglu,K., Kohli,P. & Hassabis,D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589.
8. Ahdritz,G., Bouatta,N., Floristean,C., Kadyan,S., Xia,Q., Gerecke,W., O'Donnell,T.J., Berenberg,D., Fisk,I., Zanichelli,N., Zhang,B., Nowaczynski,A., Wang,B., Stepniewska-Dziubinska,M.M., Zhang,S., Ojewole,A., Guney,M.E., Biderman,S., Watkins,A.M., Ra,S., Lorenzo,P.R., Nivon,L., Weitzner,B., Ban,Y.-E.A., Chen,S., Zhang,M., Li,C., Song,S.L., He,Y., Sorger,P.K., Mostaque,E., Zhang,Z., Bonneau,R. & AlQuraishi,M. (2024). OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods* 21, 1514-1524.
9. Li,Z., Liu,X., Chen,W., Shen,F., Bi,H., Ke,G. & Zhang,L. (2022). Uni-Fold: An Open-Source Platform for Developing Protein Folding Models beyond AlphaFold. *bioRxiv*, 2022.08.04.502811.
10. 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.
11. Baek,M., DiMaio,F., Anishchenko,I., Dauparas,J., Ovchinnikov,S., Lee,G.R., Wang,J., Cong,Q., Kinch,L.N., Schaeffer,R.D., Millán,C., Park,H., Adams,C., Glassman,C.R., DeGiovanni,A., Pereira, J.H., Rodrigues, A.V., van Dijk,A.A., Ebrecht,A.C., Opperman,D.J., Sagmeister,T., Buhlheller,C., Pavkov-Keller,T., Rathinaswamy,M.K., Dalwadi,U., Yip,C.K., Burke,J.E., Garcia,K.C., Grishin,N.V., Adams,P.D., Read,R.J. & Baker,D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373, 871-876.
12. Lin,Z., Akin,H., Rao,R., Hie,B., Zhu,Z., Lu,W., Smetanin,N., Verkuil,R., Kabeli,O., Shmueli,Y., dos Santos Costa,A., Fazel-Zarandi,M., Sercu,T., Candido,S. & Rives,A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 379, 1123-1130.
13. Wu,R., Ding,F., Wang,R., Shen,R., Zhang,X., Luo,S., Su,C., Wu,Z., Xie,Q., Berger,B., Ma,J. & Peng,J. (2022). High-resolution de novo structure prediction from primary sequence. *bioRxiv*, 2022.07.21.500999.
14. Zhang,Y. & Skolnick,J. (2004). SPICKER: a clustering approach to identify near-native protein folds. *Journal of computational chemistry* 25, 865-871.
15. Zhang,J., Liang,Y. & Zhang,Y. (2011). Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. *Structure* 19, 1784-1795.

Integration of template guidance, docking, and deep-learning for protein-ligand binding mode and affinity prediction

Xianjin Xu[†], Rui Duan[†], Xiaoqin Zou^{*}

Dalton Cardiovascular Research Center, Department of Physics and Astronomy, Department of Biochemistry, Institute for Data Science and Informatics, University of Missouri, Columbia, Missouri 65211, United States

[†]equal contribution; ^{*}corresponding author (zoux@missouri.edu)

Key: Auto:N; CASP_serv:N; Templ:Y

In CASP16, we submitted predicted binding modes and affinities for protein-ligand complexes. The binding mode predictions were based on a template-guided method and molecular docking. For binding energy calculations, we used both a deep learning-based scoring function and a docking scoring function.

Methods

Template-guiding method

An improved template-guiding strategy, based on our recent work [1-3], was employed for protein-ligand complex structure predictions in CASP16. This strategy allows the use of both similar and dissimilar ligands as templates through a newly developed intercomparison method, followed by local optimization and ranking with a hybrid scoring function, in addition to the standard docking protocol. Specifically, for each query target, an ensemble of ligand 3D conformers was generated from the SMILES string using the OMEGA2 program (Version 4.2.2.0, OpenEye Scientific Software, Santa Fe, NM, USA. <http://www.eyesopen.com>) [4,5]. Then, the Protein Data Bank [6] was then searched for template structures containing the target protein or its homologs. If a template structure contains one or more co-bound ligands, the target protein structure was superimposed onto the protein structure in the template with the MatchMaker tool in UCSF Chimera [7], and the conformers of the query ligand were matched to the co-bound ligands in the template using SHAFTS, a 3D molecular similarity measurement program [8]. The superimposed protein structure and the matched ligand conformers were combined, and local minimization was performed by AutoDock Vina [9] (with option “local_only”). The predicted complex structures were then ranked by a hybrid scoring function [1] that combines a protein-ligand binding score (AutoDock Vina score) and a 3D similarity score (SHAFTS score, characterizing the 3D similarity between the query ligand and the co-bound ligand in the template). The ranked binding modes were clustered using a ligand root-mean-square deviation (L-RMSD) cutoff 1.5 Å, where L-RMSD is the heavy atom RMSD of the ligand between two modes after structural superposition of the protein.

Molecular docking

Molecular docking was performed using Glide [10] within the Schrödinger suite [11]. The protein structures were generated through the AlphaFold3 server [12] and subsequently prepared using the Protein Preparation Workflow in Maestro [11] with default settings. This preparation involved determining the protonation states of residues, assigning hydrogen bonds, and minimizing the protein structure. Minimization was carried out until the RMSD reached 0.3 Å, using the OPLS4 force field [13]. Ligand structures, provided as SMILES strings by CASP, were converted into 3D conformations using LigPrep [11] with default parameters. Docking grids were generated based on the preprocessed protein structure

by selecting residues within the identified binding pocket. The binding pocket location was determined through the above template-based approach with protein-ligand complexes from related templates. All docking simulations were conducted in Glide, utilizing the standard precision (SP) mode and default settings.

Model submission

For the submission of Group *Zou*, ligands with similarity scores (SHAFTS scores) above 1.2 in the template-based approach had their top 3 binding modes submitted as predictions 1, 3, and 5, while the top 2 modes from the docking-based approach were submitted as predictions 2 and 4. For ligands with similarity scores below 1.2, the top 3 binding modes from the docking-based approach were submitted as predictions 1, 3, and 5, with predictions 2 and 4 derived from the template-based approach. For the submission of Group *Zou_aff2*, all five models were generated using the template-based approach.

Binding energy calculation

For energy prediction of targets L1000 and L3000, we used the method developed by Wei and co-workers to derive the TopBP-DL model [14]. This method incorporates persistent homology-based features and was employed in their machine learning models for energy prediction in the D3R Grand Challenge 3 and 4 [15,16]. Here, the TopBP-DL model was trained on the PDBbind-v2016 dataset [17] and fine-tuned with target-specific ligands from the BindingBD [18] database. The top binding mode predicted by Glide docking served as the input structure for model training and energy prediction for the CASP targets. The predicted energies were submitted to Group *Zou*. For Group *Zou_aff2*, the predicted energies were based on the Glide score of the top binding mode identified by Glide docking.

1. Xu, X., Duan, R., Zou, X. (2023) Template-guided method for protein–ligand complex structure prediction: Application to CASP15 protein–ligand studies. *Proteins*, 2023, 91: 1829-1836.
2. Xu, X., Zou, X. (2021) Dissimilar Ligands Bind in a Similar Fashion: A Guide to Ligand Binding-Mode Prediction with Application to CELPP Studies. *Int. J. Mol. Sci.* 22: 12320.
3. Xu, X., Ma, Z., Duan, R., Zou, X. (2019) Predicting protein–ligand binding modes for CELPP and GC3: Workflows and insight. *J. Comput.-Aided Mol. Des.* 33: 367-374.
4. Hawkins, P.C., Skillman, A.G., Warren, G.L., Ellingson, B.A., Stahl, M.T. (2010) Conformer generation with omega: Algorithm and validation using high quality structures from the protein databank and Cambridge structural database. *J. Chem. Inf. Model.* 50: 572-584.
5. Hawkins, P.C., Nicholls, A. (2012) Conformer generation with OMEGA: Learning from the data set and the analysis of failures. *J. Chem. Inf. Model.* 52: 2919-2936.
6. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2000) The protein data bank. *Nucleic Acids Res.* 28: 235-242.
7. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. (2004) UCSF chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605-1612.
8. Liu, X., Jiang, H., Li, H. (2011) SHAFTS: A hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *J. Chem. Inf. Model.* 51: 2372-2385.
9. Trott, O., Olson, A.J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31: 455-461.
10. Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: A new

- approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J. Med. Chem.* 47, 1739-1749.
11. Schrödinger Release 2024-2, Schrödinger, LLC, New York, NY, 2024.
 12. Abramson, J., et al. (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature* 630, 493-500.
 13. Lu, C., Wu, C., Ghoreishi, D., Chen, W., Wang, L., Damm, W., Ross, G. A., Dahlgren, M. K., Russell, E., Von Bargen, C. D., Abel, R., Friesner, R. A., and Harder, E. D. (2021) OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space, *J. Chem. Theory Comput.* 17, 4291-4300.
 14. Cang, Z., Mu, L., and Wei, G. W. (2018) Representability of algebraic topology for biomolecules in machine learning based scoring and virtual screening, *PLoS Comput. Biol.* 14, e1005929.
 15. Nguyen, D. D., Cang, Z., Wu, K., Wang, M., Cao, Y., and Wei, G. W. (2019) Mathematical deep learning for pose and binding affinity prediction and ranking in D3R Grand Challenges, *J. Comput. Aided Mol. Des.* 33, 71-82.
 16. Nguyen, D. D., Gao, K., Wang, M., and Wei, G. W. (2020) MathDL: mathematical deep learning for D3R Grand Challenge 4, *J. Comput. Aided Mol. Des.* 34, 131-147.
 17. Liu, Z., Su, M., Han, L., Liu, J., Yang, Q., Li, Y., and Wang, R. (2017) Forging the Basis for Developing Protein-Ligand Interaction Scoring Functions, *Acc. Chem. Res.* 50, 302-309.
 18. Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., and Chong, J. (2016) BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology, *Nucleic Acids Res.* 44, D1045-1053.

OTHER

How Physics can Improve AI Generated TCR-pMHC Complexes

Matthias Peter¹, Thomas Rösler¹, Floris J. van Eerden¹, Martin Löwer¹

¹TRON – Translationale Onkologie an der Universitätsmedizin der Johannes Gutenberg-Universität Mainz

matthias.peter@tron-mainz.de

Key: Auto:N; CASP_serv:N; Templ:N; MSA:N; DeepL:Y; EMA:N; MD:Y

Developing a reliable method to model high-quality models of the T cell receptor bound to the peptide presented by the major histocompatibility complex (TCR-pMHC) is of great interest in the field of personalized immunotherapies. The knowledge gained from analyzing these interactions contributes to a better understanding of T-cell activation, which can ultimately help optimize target selection for individualized therapies ^{1,2}.

The interactions of the TCR with the pMHC determine how strongly a T cell reacts to an antigen presented by a tumor or an infected cell. The advent of artificial intelligence has revolutionized structural biology, making it possible to predict the structure of any protein without the need for time-consuming laboratory methods³. Recently developed *AlphaFold*-based tools like *TCRdock*⁴ and *TCRmodel2*⁵ specialize in modeling the TCR-pMHC complex. However, TCR-pMHC complexes pose an additional challenge for evolution-based methods like *AlphaFold* due to the presence of hypervariable CDR3 regions, which do not follow conventional evolutionary patterns⁶.

Methods

In our benchmark, these tools produced models with overall good architecture, but the accuracy of the hypervariable CDR3 regions, which binds the presented peptide, was often insufficient. A direct comparison of the contacts in the binding interfaces of the models with the crystal structures showed that only about 20% to 40% of the native contacts were represented correctly.

Result

However, in some cases, a short molecular dynamics simulation increased the correctly predicted contacts to 60%–80%. We show that the prediction of TCR-pMHC complexes can be improved by combining AI methods with a physics-based refinement step. The aim is to create a method that enhances the quality of imperfect complexes with a certain degree of reliability, without impairing those that already exhibit good overall quality. Ultimately, this method could be used to improve *AlphaFold* generated TCR-pMHC complexes for which no crystal structure is available, significantly increasing the number of analyzable examples.

The synergy between AI-based structural modeling and physics-based molecular dynamics simulations could thus help clarify the still unresolved questions of T-cell activation and contribute to further improvements in personalized immunotherapies.

1. Corse,E., Gottschalk,R.A., Allison,J.P. Strength of TCR–Peptide/MHC Interactions and In Vivo T Cell Responses. *J Immunol.* 2011;186(9):5039-5045. doi:10.4049/JIMMUNOL.1003650
2. Shevyrev,D.V., Tereshchenko,V.P., Sennikov,S.V. The Enigmatic Nature of the TCR-pMHC Interaction: Implications for CAR-T and TCR-T Engineering. *Int J Mol Sci* 2022, Vol 23, Page 14728. 2022;23(23):14728. doi:10.3390/IJMS232314728
3. Jumper,J., Evans,R., Pritzel,A., et al. Highly accurate protein structure prediction with AlphaFold. *Nat* 2021 5967873. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2
4. Bradley,P. Structure-based prediction of T cell receptor:peptide-MHC interactions. *Elife.* 2023;12. doi:10.7554/ELIFE.82813
5. Yin,R., Ribeiro-Filho,H.V., Lin,V., Gowthaman,R., Cheung,M., Pierce,B.G. TCRmodel2: high-resolution modeling of T cell receptor recognition using deep learning. *Nucleic Acids Res.* 2023;51(W1):W569-W576. doi:10.1093/NAR/GKAD356
6. Zhang,H., LanJ., Wang, H., et al. AlphaFold2 in biomedical research: facilitating the development of diagnostic strategies for disease. *Front Mol Biosci.* 2024;11:1414916. doi:10.3389/FMOLB.2024.1414916/BIBTEX

MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, MULTICOM_human
(Flexible targets)

Predicting Conformation Ensembles for Flexible Protein Targets with MULTICOM4

Jian Liu, Pawan Neupane, Jianlin Cheng*

University of Missouri, Columbia, MO 65211, USA

*Corresponding author: chengji@missouri.edu

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; EMA:Y; MD:N

During CASP16, we employed the MULTICOM4 system - an enhanced version of MULTICOM^{3,2} developed using AlphaFold2³ and AlphaFold3⁴ - to generate numerous tertiary structural models for two flexible targets with multiple conformations (T1200 and T1300). These models were filtered to create a subset that met the criterion of having the four key helices within 0.5 Å RMSD (Root Mean Square Deviation) from the corresponding helices in the template structures (PDB codes: 2LR2 and 4NPD) in the Protein Data Bank (PDB). Subsequently, we used five different methods to calculate population scores (probability scores) and uncertainty scores for the selected models. These calculations were based on the global pLDDT scores predicted by AlphaFold, along with RMSD scores relative to the centers of model clusters. This effort resulted in five predictors - MULTICOM_AI, MULTICOM_GATE, MULTICOM_LLM, MULTICOM, and MULTICOM_human - participating in the flexible target prediction task

Methods

During CASP16 experiment, MULTICOM4 predictors made predictions for the two flexible targets as follows.

1) Model Generation

MULTICOM4 used a monomer structure prediction pipeline built on top of AlphaFold2 to generate thousands of structural models for each of the two targets (see the detailed description in the MULTICOM TS abstract in this abstract book). We also used AlphaFold3 web server to generate thousands of structural models for each of them.

2) Model Filtering

The structural models generated were filtered out using the following steps:

- Based on the information provided in the *Specifications* document for the two targets, we extracted 4 helices (named predicted_helix_1, predicted_helix_2, predicted_helix_3, predicted_helix_4) from the structural models and the corresponding 4 helices from the template structures (2 helices from 2LR2 (named 2LR2_helix_1 and 2LR2_helix_2), 2 helices from 4NPD (named 4NPD_helix_1 and 4NPD_helix_2)).

- Superimpose the 4 template helices with the corresponding 4 predicted helices of each model one by one.
- Replace the 4 helices in the models with the corresponding template helices according to the superimposition.
- Remove a structural model with the helix replacement that met any of the following conditions:
 - Contains severe clashes
 - Substantially deviated from its original counterpart without the helix replacement according to TM-score between them (i.e. TM-score < 0.99).
 - The combination of predicted_helix_1 and predicted_helix_2 in the model is not within 0.5 Å RMSD with the combination of 2LR2_helix_1 and 2LR2_helix_2.
 - The combination of predicted_helix_3 and predicted_helix_4 is not within 0.5 Å RMSD with the combination of 4NPD_helix_1 and 4NPD_helix_2.

3) Calculation of Population and Uncertainty Scores

Out of the remaining models after the filtering, 2000 models (1000 generated by AlphaFold2 and 1000 generated by AlphaFold3) were then selected based on their global pLDDT score ranking from high to low. For a AlphaFold2-based models, its global pLDDT score was provided by AlphaFold2 directly. For a AlphaFold3-based model, its pLDDT score was computed using the average of the per-residue confidence scores for alpha-carbon (Ca) atoms. To get the final population score for each model, the following two approaches were used:

Approach 1: The pLDDT score of each model was first scaled into the range [0,1]. Then the Sigmoid function was used to calculate the population score for each model with its scaled pLDDT score as input.

Approach 2: 1000 AlphaFold2 (or AlphaFold3) models were clustered into two clusters based on their similarity. Then for each model, its RMSD was calculated against its cluster center (the model with the highest average pairwise similarity score (TM-score) with all the other models in the cluster). The standard deviation of the RMSDs of 1000 models is calculated as σ . The population score for a model with an RMSD (r) is then calculated according to the Gaussian density function $N(r | \text{mean} = 0 \text{ and standard deviation} = \sigma)$.

The population scores for 1000 AlphaFold2 or AlphaFold3 models calculated by the two approaches above were then normalized by their sum so that they add up to 1. The standard deviation of the normalized population scores was calculated as the uncertainty score. If the uncertainty score for any model was greater than its population score, the uncertainty score was replaced by the population score for that model.

4) Implementation of Five MULTICOM Predictors

MULTICOM_AI server predictor used 1000 AlphaFold2 models with Approach 1. *MULTICOM_GATE* server predictor used 1000 AlphaFold2 models with Approach 2. *MULTICOM_LLM* server predictor used 1000 AlphaFold3 models with Approach 1. *MULTICOM* (a human predictor) used 1000 AlphaFold3 models with Approach 2. *MULTICOM_human* (a human predictor) used 2000 models (1000 AlphaFold2 models with Approach 1 and 1000 AlphaFold3 models with Approach 1). The population

scores for the 2000 models were then renormalized and the standard deviation was computed from the renormalized population scores as the uncertainty score.

1. Liu J, Guo Z, Wu T, et al. Improving AlphaFold2-based protein tertiary structure prediction with MULTICOM in CASP15. *Communications chemistry* 2023;6:188.
2. Liu J, Guo Z, Wu T, et al. Enhancing alphafold-multimer-based protein complex structure prediction with MULTICOM in CASP15. *Communications biology* 2023;6:1140.
3. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *nature* 2021;596:583-589.
4. Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024:1-3.

Exhaustive conformational sampling of domain-linker-domain proteins

Y. Bouchiba¹, S. Rodriguez², S. Bartels^{3,4}, S. Barbe¹, P. Bernadó² and J. Cortés⁴

1 - Toulouse Biotechnology Institute, CNRS, INRAE, INSA, Toulouse, France,

2 - Centre de Biologie Structurale, Université de Montpellier, INSERM, CNRS, Montpellier, France,

3 - Institut de Mathématiques de Toulouse, Université de Toulouse, CNRS, Toulouse, France,

4 - LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France

juan.cortes@laas.fr

Key: Auto:N; CASP_serv:N; Templ:Y; MSA:Y; DeepL:Y; AF:AF2; EMA:N; MD:Y

We used a combination of stochastic sampling algorithms, statistical analysis techniques and all-atom molecular dynamics simulations to predict conformational ensembles for targets T1200 and T1300 in CASP16. The pipeline is presented below.

Methods

Structure of the rigid domains

T1200 and T1300 targets are both constituted of two globular domains connected by a flexible linker. Both targets have an identical sequence length, and the sole difference lies in the amino acid composition of the linker. The N-ter globular domain (residues 1-70) is ZLBT, a variant of the B domain of the staphylococcal protein A, while the C-ter domain (residues 77-129) corresponds to the C domain of the same protein. The linker spans residues 71-76. Initial models were generated using ColabFold¹ with the automated MSA construction pipeline and by allowing the usage of structural templates from the PDB. The resulting models were then relaxed using AmberTools23² and the top pLDDT-ranked model was retained for subsequent operations. The ions observed in the C domain structure (4NPD) were not considered in our models.

Conformations sampling of the flexible linker

According to the specifications provided by the challenge organizers, the two residues preceding the linker, Ala-69 and Pro-70, can be considered part of the flexible region. Therefore, conformations of the fragment 69-76 were sampled using a method tailored for Intrinsically Disordered Regions (IDRs)³. The two globular domains were considered as rigid bodies at this stage. A total of 10000 conformations per target were generated.

Selection of initial states for MD simulations

The sampled conformations were grouped by binning the relative positions between the two domains into voxels of 10 Å. From each group, two conformations were selected as starting configurations for MD simulations: one representing the average domain orientation within the voxel and another with high probability density that differed significantly from the first. Following this protocol, 334 conformations were selected for the T1200 target and 640 for T1300.

MD simulations

Each initial state for all-atom MD simulations was prepared using the tleap module from AmberTools²³, solvating the proteins with approximately 6,700 OPC⁴ water molecules and counterions, and using the amber ff19SB force field⁵. The topology and coordinate files were converted for Gromacs 2022.5⁶ simulations using the amb2gmx.py script.

Energy minimization was performed with the steepest descent algorithm, followed by a 100 ps equilibration in a NVT ensemble (300 K, Berendsen thermostat⁷) and a 100 ps equilibration in a NPT ensemble (1 bar, Parrinello-Rahman barostat⁸), both with position restraints applied to the heavy atoms. Each conformation was submitted to a 20 ns unrestrained all-atom MD production run. Both equilibration and production steps were carried out with a 2 fs time step.

Selection of conformers and their populations

Frames from all MD simulations were pooled into a single trajectory and aligned to the ZBLT domain. The trajectory was then clustered using a hierarchical algorithm with average linkage and an RMSD metric focusing on residues 70 to 129. 1000 clusters were requested. Then, all conformations were aligned to both ZLBT (2LR2) and C domain (4NPD) and the backbone RMSD was computed. All conformations with high RMSD values of the globular domains ($>3.5\text{\AA}$) were discarded. As some clusters were removed from the initial 1000 requested clusters, the populations were defined as the fraction of the cluster population divided by the sum of all retained cluster populations. The uncertainties were arbitrarily set to 10% of the cluster populations.

Availability

Binaries and scripts will be made available upon publication.

1. Mirdita, M., et al. (2022). ColabFold: making protein folding accessible to all. *Nature Methods* 19, 679–682.
2. Case, D. A., et al. (2023). AmberTools. *Journal of Chemical Information and Modeling* 63 (20), 6183–6191.
3. Estaña A., Sibille N., Delaforge E., Vaisset M., Cortés J. and Bernadó P. (2019). Realistic ensemble models of intrinsically disordered proteins using a structure-encoding coil database. *Structure* 27(2), 381–391.
4. Xiong, Y., Shabane, P. S. and Onufriev, A. V. (2020). Melting Points of OPC and OPC3 Water Models. *ACS Omega*, 5 (39), 25087–25094.
5. Tian C., et al. (2020). ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation*, 16, 528–552.

6. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1, 19-25.
7. Bussi, G., Donadio, D. and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126 (1), 014101.
8. Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52 (12), 7182–7190.

Using IDPConformerGenerator and Molecular Dynamics Strategies to Predict Conformational Ensembles of *Staph. Aureus* ZLBT-C

Zi Hao Liu^{1,2,†}, Kunyang Sun^{3,4,†}, Teresa Head-Gordon^{3,4,5,6*}, Julie Forman-Kay^{1,2,*}

¹*Molecular Medicine Program, Hospital for Sick Children, Toronto, Ontario M5G 0A4, Canada*

²*Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada*

³*Pitzer Center for Theoretical Chemistry, University of California, Berkeley, California 94720, USA*

⁴*Department of Chemistry, University of California, Berkeley, California 94720-1460 USA*

⁵*Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720-1462, USA*

⁶*Department of Bioengineering, University of California, Berkeley, California 94720-1762, USA*

*forman@sickkids.ca, thg@berkeley.edu. †Authors having contributed equally to the work.

The ZLBT-C target for CASP16 consists of two experimentally solved folded domains connected by a 6-residue long (WT_T1200: KADNKF, Gly6_T1300: GGGGGG) flexible linker connecting the N-terminal ZLBT domain (PDB ID 2LR2 [1]) to the C domain (PDB ID 4NPD [2]). The linker conformations can be modeled using methods for the *in silico* modeling of intrinsically disordered regions (IDRs). To address the targets T1200 and T1300, we have chosen IDPConformerGenerator [3,4] to construct initial conformations for all-atom molecular dynamics (MD) processing. IDPConformerGenerator uses sequence-dependent statistical sampling of ϕ , ψ , and ω backbone torsion angles derived from the RCSB PDB [5] to construct all-atom conformations of the linker IDRs given PDB templates of folded domains to create ensembles of full proteins. We used this approach with PDB templates of the ZLBT and C regions to create initial ensembles of the ZLBT-C protein.

After selecting representative conformations from the IDPConformerGenerator ensemble, we used iMiner [6] workflow and OpenMM [7], respectively, to perform MD simulations using AMBER14SB/TIP3P [8] and CHARMM36/CHARMM36m-water [9] protein forcefields/water parameters. The simulation system is first solvated and neutralized with sodium and chloride ions before the parametrization. Then, the system is energy minimized and equilibrated with 1-ns simulation each under the NVT and NPT ensemble before the real production runs. For each system and forcefield combination, we run a 10-ms production and save structural snapshots every 2-ns to collect around 5000 structures for each trajectory.

The energy landscape visualization method (ELViM) [10] was then used on all of the ensembles to weight clusters of structurally similar conformations for final submission. The ELViM strategy uses a multidimensional reduction technique that analyzes internal distances between pairs of structural conformations on the entire dataset.

Methods

(1) Submission: JFK-THG-IDPCONFGEN

Auto:N; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:Y; MD:N.

IDPConformerGenerator [3] was used to generate 100,000 possible backbone configurations of the 6-residue long linker for each of the T1200 and T1300 targets (KADNKF and GGGGGG respectively). These linkers were then appended onto the ZLBT (2LR2) domain using the local disordered region sampling (LDRS) module in IDPConformerGenerator [4]. The C domain (4NPD) was then appended onto the ZLBT-linker construct using LDRS to avoid steric clashes with the ZLBT domain. A final 1705 conformations for T1200 and 1635 conformations for T1300 had sidechains added by FASPR via IDPConformerGenerator and subsequently hydrogenated using PDBFixer [11], weighting normalized to 1000 conformations each and submitted. The energy landscape for the ensembles were calculated using ELViM [10] for further MD (see below, submissions 2-5).

(2) *Submission: JFK-THG-AMBER*

Auto:Y; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:Y; MD:Y.

Two representative conformations from the IDPConformerGenerator [3] pools, one from each of the high density areas in the ELViM [10] clustering plots, were used as starting structures for MD simulation with the AMBER14SB/TIP3P [8] forcefield. After obtaining the 10-ms trajectory, we used ELViM to take 1000 representative structures based on the clustering density.

(3) *Submission: JFK-THG-AMBERstable*

Auto:Y; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:Y; MD:Y.

Representative conformations from IDPConformerGenerator [3] pools based on ELViM [10] clustering density were used as starting structures for MD simulation with AMBER14SB/TIP3P [8] forcefield. After obtaining the 10-ms trajectory, we used ELViM to take 1000 representative structures based on the clustering density as shown in Figure 1Di and 1Dii.

(4) *Submission: JFK-THG-CHARMM*

Auto:Y; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:Y; MD:Y.

Representative conformations from IDPConformerGenerator [3] pools, , one from each of the high density areas in the ELViM [10] clustering plots, were used as starting structures for MD simulation with CHARMM36/CHARMM36m-water [9] forcefield. After obtaining the 10-ms trajectory, we used ELViM to take 1000 representative structures based on the clustering density as shown in Figure 1Ci and Cii.

(5) *Submission: JFK-THG-CHARMMstable*

Auto:Y; CASP_serv:N; Templ:Y; MSA:N; DeepL:N; AF:N; EMA:Y; MD:Y.

Representative conformations from IDPConformerGenerator [3] based on ELViM [10] clustering desnsity were used as starting structure of the MD simulation with CHARMM36/CHARMM36m-water [9] forcefield. After obtaining the 10-ms trajectory, we used ELViM to take 1000 representative structures based on the clustering density as shown in Figure 1Ei and 1Eii.

Results

We find that different force fields suggest different orientations of the two domains. Although the wild type linker (T1200) provides significant flexibility, the poly-glycine variant (T1300) shows more flexibility as the simulation takes more timesteps to find a stable pose. Overall, the helical bundles in the two domains have a high tendency to stick together rather than separate during the simulations. Further downstream reweighting with experimental measurements is required to identify bound and unbound fractions between the two domains.

The JFK-THG-IDPCONFGEN ensemble has been submitted as “T1200TS097_IDPCONFGEN” and “T1300TS097_IDPCONFGEN” for the two targets, respectively. The JFK-THG-AMBER ensemble has been submitted as “T1200TS097_AMBER” and “T1300TS097_AMBER”. The JFK-THG-CHARMM ensemble has been submitted as “T1200TS097_CHARMM” and “T1300TS097_CHARMM”. The JFK-THG-AMBERstable and JFK-THG-CHARMMstable ensembles have been submitted as “T1200TS097_AMBER_STABLE”, “T1300TS097_AMBER_STABLE”, and “T1200TS097_CHARMM_STABLE”, “T1300TS097_CHARMM_STABLE”, respectively.

Availability

IDPConformerGenerator [3,4] is an open-source software and is available on GitHub at github.com/julie-forman-kay-lab/IDPConformerGenerator. ELViM [10] is available on GitHub at github.com/VLeiteGroup/ELViM. The GROMACS [12] engine was used for the MD simulations which can be found here manual.gromacs.org. The CHARMM36/CHARMM36m-water force field [9] is available on mackerell.umaryland.edu/charmm_ff.shtml. The AMBER14SB/TIP3P force field [8] is available on ambermd.org/AmberModels_proteins.php.

1. A.W. Barb, T.G. Ho, H. Flanagan-Steet, J.H. Prestegard, Lanthanide binding and IgG affinity construct: Potential applications in solution NMR, MRI, and luminescence microscopy, *Protein Science* 21 (2012) 1456–1466. <https://doi.org/10.1002/pro.2133>.
2. L.N. Deis, C.W. Pemble, Y. Qi, A. Hagarman, D.C. Richardson, J.S. Richardson, T.G. Oas, Multiscale Conformational Heterogeneity in Staphylococcal Protein A: Possible Determinant of Functional Plasticity, *Structure* 22 (2014) 1467–1477. <https://doi.org/10.1016/j.str.2014.08.014>.
3. J.M.C. Teixeira, Z.H. Liu, A. Namini, J. Li, R.M. Vernon, M. Krzeminski, A.A. Shamandy, O. Zhang, M. Haghighatlari, L. Yu, T. Head-Gordon, J.D. Forman-Kay, IDPConformerGenerator: A Flexible Software Suite for Sampling the Conformational Space of Disordered Protein States, *J. Phys. Chem. A* 126 (2022) 5985–6003. <https://doi.org/10.1021/acs.jpca.2c03726>.
4. Z.H. Liu, J.M.C. Teixeira, O. Zhang, T.E. Tsangaris, J. Li, C.C. Grdinaru, T. Head-Gordon, J.D. Forman-Kay, Local Disordered Region Sampling (LDRS) for ensemble modeling of proteins with experimentally undetermined or low confidence prediction segments, *Bioinformatics* 39 (2023) btad739. <https://doi.org/10.1093/bioinformatics/btad739>.
5. S.K. Burley, H.M. Berman, J.M. Duarte, Z. Feng, J.W. Flatt, B.P. Hudson, R. Lowe, E. Peisach, D.W. Piehl, Y. Rose, A. Sali, M. Sekharan, C. Shao, B. Vallat, M. Voigt, J.D. Westbrook, J.Y. Young, C. Zardecki, Protein Data Bank: A Comprehensive Review of 3D Structure Holdings and Worldwide

- Utilization by Researchers, Educators, and Students, *Biomolecules* 12 (2022) 1425. <https://doi.org/10.3390/biom12101425>.
- 6. J. Li, O. Zhang, K. Sun, Y. Wang, X. Guan, D. Bagni, M. Haghightlari, F.L. Kearns, C. Parks, R.E. Amaro, T. Head-Gordon, Mining for Potent Inhibitors through Artificial Intelligence and Physics: A Unified Methodology for Ligand Based and Structure Based Drug Design, *J. Chem. Inf. Model.* (2024). <https://doi.org/10.1021/acs.jcim.4c00634>.
 - 7. P. Eastman, R. Galvelis, R.P. Peláez, C.R.A. Abreu, S.E. Farr, E. Gallicchio, A. Gorenko, M.M. Henry, F. Hu, J. Huang, A. Krämer, J. Michel, J.A. Mitchell, V.S. Pande, J.P. Rodrigues, J. Rodriguez-Guerra, A.C. Simmonett, S. Singh, J. Swails, P. Turner, Y. Wang, I. Zhang, J.D. Chodera, G. De Fabritiis, T.E. Markland, OpenMM 8: Molecular Dynamics Simulation with Machine Learning Potentials, *J. Phys. Chem. B* 128 (2024) 109–116. <https://doi.org/10.1021/acs.jpcb.3c06662>.
 - 8. J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.* 11 (2015) 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
 - 9. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, *Nat Methods* 14 (2017) 71–73. <https://doi.org/10.1038/nmeth.4067>.
 - 10. R.G. Viegas, I.B.S. Martins, M.N. Sanches, A.B. Oliveira Junior, J.B. de Camargo, F.V. Paulovich, V.B.P. Leite, ELViM: Exploring Biomolecular Energy Landscapes through Multidimensional Visualization, *J. Chem. Inf. Model.* 64 (2024) 3443–3450. <https://doi.org/10.1021/acs.jcim.4c00034>.
 - 11. P. Eastman, J. Swails, J.D. Chodera, R.T. McGibbon, Y. Zhao, K.A. Beauchamp, L.-P. Wang, A.C. Simmonett, M.P. Harrigan, C.D. Stern, R.P. Wiewiora, B.R. Brooks, V.S. Pande, OpenMM 7: Rapid development of high performance algorithms for molecular dynamics, *PLOS Computational Biology* 13 (2017) e1005659. <https://doi.org/10.1371/journal.pcbi.1005659>.
 - 12. S. Páll, A. Zhmurov, P. Bauer, M. Abraham, M. Lundborg, A. Gray, B. Hess, E. Lindahl, Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS, *The Journal of Chemical Physics* 153 (2020) 134110. <https://doi.org/10.1063/5.0018516>.

Integrating Advanced Computational Techniques for Biomolecular Structure Prediction

Yuki Kagaya¹, Tsukasa Nakamura¹, Jacob Verburgt¹, Anika Jain¹, Genki Terashi¹, Pranav Punuru¹, Emilia Tugolukova¹, Joon Hong Park², Anouka Saha³, David Huang¹, and Daisuke Kihara^{1,2}

1 - Department of Biological Sciences, Purdue University, West Lafayette, IN, USA,

2 - Department of Computer Science, Purdue University, West Lafayette, IN, USA,

3 - Department of Mathematics, The University of Texas at Austin, Austin, TX, USA

dkihara@purdue.edu

Key: Auto:N; CASP_serv:Y; Templ:Y; MSA:Y.MetaG; DeepL:Y; AF:AF3/2; EMA:Y; MD:N

This section provides additional content to our previously submitted abstract, located earlier in this volume.

Methods

Ensemble modeling: For targets T1200 and T1300, we used MODELLER¹ to generate initial structures (using PDB IDs 4NPD and 2LR2 as templates respectively) and subsequently carried out 500 ns MD simulations on each structure using Desmond² and the OPLS4 forcefield³, resulting in 1000 MD frames for each target. To ensure the RMSDs of the helical bundles remained within the required 0.5 Å RMSD to the reference structures, we superimposed the helical bundles from the reference structures into each MD frame and cleared any steric clashes via PyMOL 2.4 [4]. All 1000 frames of each target were weighed equally in our final submissions.

1. Webb,B. & Sali,A. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics 54, 5.6.1-5.6.37 (2016).
2. Bowers,K.J. et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06, 84-es (2006)
3. Lu,C. et al. OPLS4: Improving force field accuracy on challenging regimes of chemical space. J Chem Theory Comput 17, 4291–4300 (2021).
4. Schrödinger, LLC, “The PyMOL Molecular Graphics System, Version 1.8,” Nov. 2015.

Strategies for ligand binding prediction in CASP16

Ren Kong^{1*}, Xufeng Lu², Zunyun Jiang³, Shan Chang^{1*}

¹*Institute of Bioinformatics and Medical Engineering, School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, Jiangsu, China; 2-Primary Biotechnology Inc., No. 3 Meishan Avenue, Changzhou 213125, Jiangsu, China; 3-Jiangsu Key Laboratory of Pesticide Science, College of Sciences, Nanjing Agricultural University, Nanjing 210095, China*

. E-mail: rkong@jsut.edu.cn, schang@jsut.edu.cn

In CASP16, we participated in the ligand-binding prediction category. A total of 11 targets, including both proteins and RNA/DNAs, were provided for prediction, with only the target sequences and the SMILES representations of the ligands available. For certain targets, such as L3000, more than 200 chemically diverse ligands required prediction of their binding poses and affinities. This scenario closely resembles the hit discovery process in drug development. To address this challenge, we developed a workflow applied for ligand prediction across different types of targets.

Method

First, we searched the target sequences in the RCSB Protein Data Bank (PDB) to determine whether any experimental structural information was available for the targets or their homologues. For targets L1000, L2000, L3000, and L4000, numerous PDB structures were identified, each containing different compounds in complex with the proteins or their homologues. Native ligands were retrieved from these PDB files and clustered with the ligands to be predicted based on their 2D chemical similarity. Codock-Ligand¹ and AutoDock Vina² were then employed to dock the ligands to the PDB structures within the same cluster, accounting for potential conformational changes induced by the ligand's chemical structure. For ligands without similar compounds available in complex with the protein, multiple receptor structures were used for docking. The resulting docking poses were clustered based on root-mean-square deviation (RMSD) and further evaluated by visual inspection. Additionally, an AI-based Convolutional Neural Network (CNN) scoring model³ was applied as a supplementary reference for pose selection and affinity prediction. For targets L2000 and L3000, affinity values were retrieved from the Binding Database⁴, and compounds with known activity values were clustered with the ligands to be predicted. Experimental activity values from the same cluster were used to adjust the predicted affinities based on docking scores.

For target L5001, although PDB structures of the protein was available, it represented the apo form with no ligand bound. Blind docking was utilized to predict potential binding poses. For both T1214 and L5001, manual adjustments were made to the orientation of aromatic side-chain residues in the predicted binding sites. For RNA/DNA targets, the AlphaFold3 server⁵ was used to generate initial receptor structures, and the ligands were docked using Codock-Ligand.

Acknowledgement

We sincerely thank Professor Jun Chen from Xi'an Jiaotong-Liverpool University and Professor Liangxu Xie from Jiangsu University of Technology for their valuable discussions and insightful feedback. We are also grateful to Primary Biotechnology Inc. for providing the computational resources essential to this project. Additionally, we appreciate the CASP16 organizers and the experimental specialists for generously providing the structural data used in the assessments.

1. M. Pang, W. He, X. Lu, et al., CoDock-Ligand: combined template-based docking and CNN-based scoring in ligand binding prediction. *BMC Bioinformatics*, 2023, 24 (1), 444.
2. J. Eberhardt, D. Santos-Martins, A. F. Tillack, et al., AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J Chem Inf Model*, 2021, 61 (8), 3891-3898.
3. M. Ragoza, J. Hochuli, E. Idrobo, et al., Protein-Ligand Scoring with Convolutional Neural Networks. *J Chem Inf Model*, 2017, 57 (4), 942-957.
4. M. K. Gilson, T. Liu, M. Baitaluk, et al., BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res*, 2016, 44 (D1), D1045-53.
5. J. Abramson, J. Adler, J. Dunger, et al., Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 2024, 630 (8016), 493-500.