Single-sequence protein-RNA complex structure prediction by geometric attention-enabled pairing of biological language models

Rahmatullah Roche[1], Sumit Tarafder[1], Debswapna Bhattacharya[1,*]

[1] Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, United States of America.

* dbhattacharya@vt.edu

Abstract

Ground-breaking progress has been made in structure prediction of biomolecular assemblies, including the recent breakthrough of AlphaFold 3. However, it remains challenging for AlphaFold 3 and other state-of-the-art deep learning-based methods to accurately predict protein-RNA complex structures, in part due to the limited availability of evolutionary and structural information related to protein-RNA interactions that are used as inputs to the existing approaches. Here, we introduce ProRNA3D-single, a new deep-learning framework for protein-RNA complex structure prediction with only single-sequence input. Using a novel geometric attention-enabled pairing of biological language models of protein and RNA, a previously unexplored avenue, ProRNA3D-single enables the prediction of interatomic protein-RNA interaction maps, which are then transformed into multi-scale geometric restraints for modeling 3D structures of protein-RNA complexes via geometry optimization. Benchmark tests show that ProRNA3D-single convincingly outperforms current state-of-the-art methods including AlphaFold 3, particularly when evolutionary information is limited; and exhibits remarkable robustness and performance resilience by attaining better accuracy with only single-sequence input than what most methods can achieve even with explicit evolutionary information. Freely available at https://github.com/Bhattacharya-Lab/ProRNA3D-single, ProRNA3D-single should be broadly useful for modeling 3D structures of protein-RNA complexes at scale, regardless of the availability of evolutionary information.

Introduction

The interactions between protein and ribonucleic acid (RNA) underpin a wide range of cellular processes from gene regulation to transcription to protein synthesis (1-6), with enormous implications in understanding the molecular basis of disease (7-10). Therefore, knowledge of protein-RNA interactions in atomic detail is critically important. However, experimental determination of the atomic-level protein-RNA complex structures is not always feasible or practical (11,12). To address the challenges associated with experimental characterization of protein-RNA complexes, several computational methods for predicting the interaction between protein and RNA have been developed, including traditional methods based on template-based modeling (13,14) and protein-RNA docking (15-19). However, their predictive modeling performance remains limited by the availability of the homologous template information or the deficiencies of docking sampling and scoring.

Following the breakthrough of deep learning-based protein structure prediction methods AlphaFold2 (20) and RoseTTAFold (21), a handful of deep learning-based methods have been

recently developed for structure prediction of biomolecular interactions including protein-RNA complex structure prediction. For example, Baker and co-workers extended the RoseTTAFold approach for the prediction of protein-nucleic acid complex structures by developing RoseTTAFold2NA (RF2NA) (22) and subsequently introduced a generalized biomolecular modeling and design framework by developing RoseTTAFold All-Atom (RF2AA) (23) that is capable of protein-RNA complex structure prediction. Very recently, Google DeepMind and Isomorphic Labs developed a new generation of the AlphaFold system, called AlphaFold 3 (AF3) (24), a single unified deep learning framework capable of predicting the joint structure of complexes of various biomolecular interactions including protein-RNA complexes. However, the prediction accuracies of even these state-of-the-art deep learning-based methods remain low for protein-RNA complex structure prediction. For example, the average iLDDT score (25) of AF3 is 39.4 for protein-RNA complex structure prediction task on the test set, whereas the average iLDDT of RF2NA is 19.0 (see Extended Data Table 1 of (24)). While AF3 clearly outperforms RF2NA, the overall accuracy of AF3 is far from accurate (<40%), indicating that there is a large room for improvement in predicting structures of protein-RNA complexes.

The limited successes in protein-RNA complex structure prediction are mainly due to: 1) lack of structural data in the Protein Data Bank (PDB) (26) beyond the most common families of RNA-binding proteins (RBPs) (27); and 2) even basic amino acid and nucleic acid joint information such as the evolutionary sequences of joint protein-RNA multiple sequence alignments (MSAs) are often shallow and lack sufficient coverage. Thus, it is worthwhile to develop single-sequence protein-RNA complex structure prediction methods.

Promisingly, protein language models (pLM) (28-32) pretrained on a large corpus of protein sequences using attention-based transformer neural network architectures (33) have shown unprecedented scalability and broad generalizability in a wide range of predictive modeling tasks—from protein structure prediction (32) to protein engineering (34,35). Alongside the progress in pLMs, significant advances have been made in pretrained RNA language models (36,37), showing promising performance in various downstream prediction tasks. These biological language models, pretrained with an unsupervised learning objective across a large corpus of sequence database, are highly effective for downstream structure-related tasks even if they are trained purely with sequence information or subsequently fine-tuned with small-scale structural data. Importantly, biological language models implicitly capture evolutionary patterns present in biological sequences by means of transformer's attention mechanism, without the need to supply explicit evolutionary data in the form of MSAs. That is, an effective coupling of protein and RNA language models not only has the potential to compensate for the lack of structural data that only exist for the most common families of RBPs, but also reduces the dependence on the availability of explicit evolutionary information in the form of joint protein-RNA MSAs that are often scarce and noisy, paving the way to single-sequence prediction of protein-RNA complex structure. Given the progress, a natural question arises: can we develop an integrative modeling approach to combine a pair of biological language models – one pretrained using protein sequences and the other pretrained using RNA sequences – for single-sequence protein-RNA complex structure prediction with improved accuracy and robustness?

Here, we present ProRNA3D-single, a single-sequence protein-RNA complex structure prediction method by geometric attention-enabled pairing of biological language models. Instead of relying on explicit evolutionary information by means of joint protein-RNA MSAs, our single-sequence method directly leverages the biological language model embeddings for the input protein and RNA sequences together with single-sequence protein and RNA structure prediction using pretrained biological language models. By converting protein and RNA language model embeddings into a structure-aware graph representation, ProRNA3D-single employs symmetry-aware graph convolutions followed by a pairwise combination of the individual component embeddings to generate protein-RNA pair embedding. The pair embedding is then fed into a ResNet-Inception module to capture the inter-component interactions between the residue-nucleotide pairs, followed by a geometric attention module to account for the many-body effect in protein-RNA interactions while satisfying geometric consistency through an attention mechanism (33), which can predict the inter-component interactions, resulting in protein-RNA interaction maps. Finally, the predicted protein-RNA interactions are transformed into multi-scale geometric restraints to optimize the relative spatial position and orientation of the protein and RNA components, leading to protein-RNA complex structure prediction.

Our method ProRNA3D-single convincingly outperforms the state-of-the-art deep learning-based protein-RNA complex structure prediction methods, including RoseTTAFold2NA, RoseTTAFold All-Atom, and AlphaFold 3. Despite being a single-sequence method, our method attains improved accuracy compared to the MSA-dependent methods when MSA information is limited, and exhibits remarkable robustness and performance resilience by attaining better accuracy with only single-sequence input than what most state-of-the-art MSA-dependent methods can achieve even with explicit MSA information. We directly verify that the structural-level predictive modeling accuracy of our method is closely connected to the quality of the protein-RNA interaction maps predicted through geometric attention-enabled pairing of biological language models. The improved accuracy and robustness of ProRNA3D-single make it broadly useful for modeling 3D structures of protein-RNA complexes at scale, regardless of the availability of explicit evolutionary information. An open-source software implementation of ProRNA3D-single, licensed under the GNU General Public License v3, is freely available at https://github.com/Bhattacharya-Lab/ProRNA3D-single.

Materials and methods

Overview of ProRNA3D-single framework

Fig 1 illustrates our ProRNA3D-single method for protein-RNA complex structure prediction. Different from the recent deep learning-based approaches for protein-RNA complex structure prediction that require explicit evolutionary information in the form of joint protein-RNA MSAs, our single-sequence method directly leverages the biological language model embeddings for the input protein and RNA sequences together with single-sequence protein and RNA structure prediction using pretrained biological language models. ProRNA3D-single first converts the protein and RNA language model embeddings into a structure-aware graph representation and then employ symmetry-aware graph convolutions followed by a pairwise combination of the individual component embeddings to generate protein-RNA pair embedding. The pair embedding is then fed into a ResNet-Inception module to capture the inter-component interactions between the residue-nucleotide pairs, followed by a geometric attention module to account for the many-body effect in

protein-RNA interactions while satisfying geometric consistency through an attention mechanism (33), which can predict the inter-component interactions, resulting in protein-RNA interaction maps. Finally, the predicted protein-RNA interactions are transformed into multi-scale geometric restraints to optimize the relative spatial position and orientation of the protein and RNA components, leading to protein-RNA complex structure prediction.
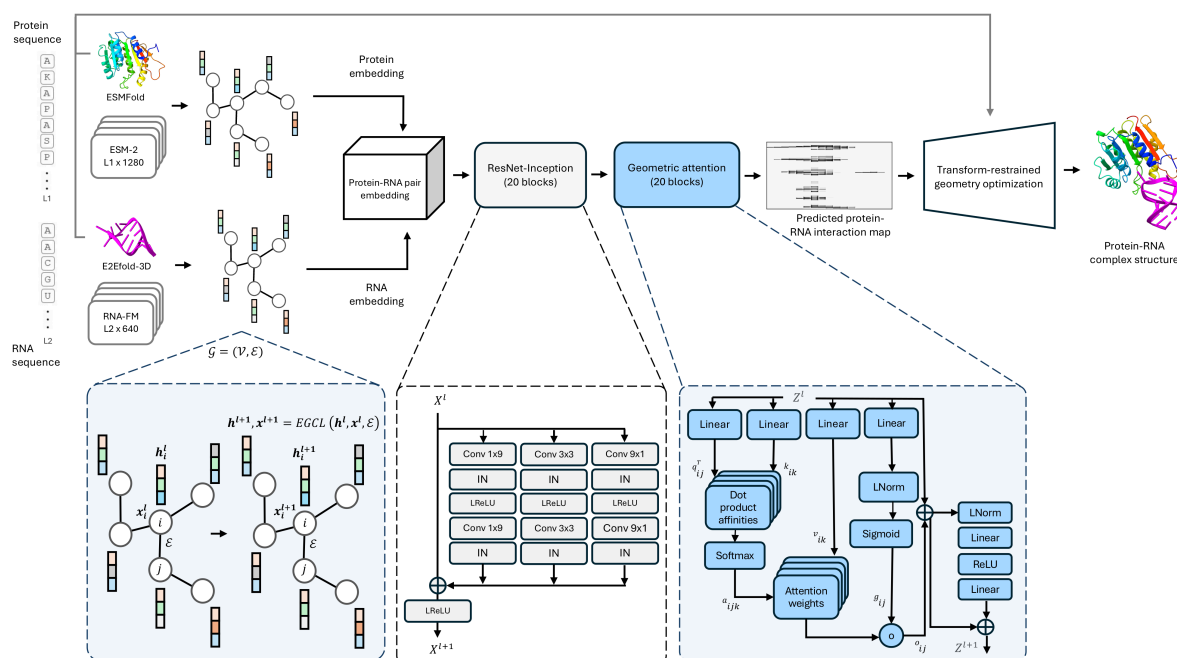

Fig 1. Illustration of ProRNA3D-single method for single-sequence protein-RNA complex structure prediction, consisting of three major modules. First, protein-RNA pair embedding generation using structure-aware graph representation of the biological language model embeddings for the input protein and RNA sequences, symmetry-aware graph convolutions, followed by a pairwise combination of the individual component embeddings. Second, protein-RNA interaction map prediction using a ResNet-Inception module followed by a geometric attention module. Third, protein-RNA complex structure prediction using transform-restrained geometry optimization.

Protein-RNA pair embedding generation

Given the input protein and RNA sequences, we use ESMFold (32) to predict the structure of the input protein and E2EFold-3D (36) with 'single sequence mode' to predict the structure of the RNA. Both ESMFold and E2EFold-3D leverage pretrained biological language models for single-sequence structure prediction, without any explicit evolutionary information such as multiple sequence alignments (MSAs). In addition to single-sequence structure prediction, we generate language model embeddings with the input protein and RNA sequences using ESM-2 (32) and RNA-FM (37) models, respectively, with the protein embedding dimension of L1x1280 where L1 represents the length of the protein, and RNA embedding dimension of L2x640 where L2 represents the length of the RNA. The protein and RNA language model embeddings are then converted into two independent structure-aware graph representations $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where a node $v \in \mathcal{V}$ contains the residue-level embeddings (for the protein graph) or the nucleotide-level embeddings (for the RNA graph) as the

node features, and an edge $e \in \mathcal{E}$ represents an interacting residue pair (for the protein graph) or a nucleotide pair (for the RNA graph). A residue pair is considered to be interacting if the Euclidean distance between the corresponding C_α atoms is within 14Å; and a nucleotide pair is considered to be interacting if the Euclidean distance between the corresponding C'_4 atoms is within 20Å. The distance thresholds are chosen following the previous studies (38-40). For the edge features in both protein and RNA graph representations, we utilize the ratio of $log |i\text{-}j|$, and $||i\text{-}j||^2$, where $|i\text{-}j|$ is the absolute sequence difference between two residues (for protein) or two nucleotides (for RNA), and $||i\text{-}j||^2$ is the Euclidean distance between the residues (or nucleotides). In addition, we use coordinate information, extracted directly from the Cartesian coordinates of the C_α atoms of the predicted protein structures and the C'_4 atoms of the predicted RNA structures.

The node and edge features together with the coordinate information are then fed into separate protein- and RNA-specific E(3)-equivariant graph neural networks (EGNNs) (41) to employ symmetry-aware graph convolutions. Both the protein- and RNA-specific EGNNs consist of a series of equivariant graph convolution layers (EGCL) to perform a series of transformations to the inputs. The EGCL updates the node and coordinate embeddings based on the node, coordinate, and edge embeddings from the previous layer, eventually resulting in enriched protein and RNA embeddings (h_i^d) by applying a linear transformation to the last layer of the protein- and RNA-specific EGNNs. Formally, the EGCL operations are defined as follows:

$$m_{ij} = \phi_e(h_i^l, h_j^l, \left\| x_i^l - x_j^l \right\|^2, a_{ij})$$

$$x_i^{l+1} = x_i^l + C \sum_{j \neq i} (x_i^l - x_j^l) \phi_x(m_{ij})$$

$$m_i = \sum_{j \neq i} m_{ij}$$

$$h_i^{l+1} = \phi_h(h_i^l, m_i)$$

$$h_i^d = \phi_{linear}(h_i^L, m_i)$$

where ϕ_e, ϕ_x and ϕ_h are non-linear transformations for edge (m_{ij}), coordinate (x_i^l) and node (h_i^l) embeddings at node i in layer l; m_i is the aggregated message at node i, and ϕ_{linear} represents the linear transformation on the final layer node embeddings (h_i^L) to obtain the enriched embeddings (h_i^d). Both the protein- and RNA-specific EGNNs are built on 4 layers of EGCL with a hidden dimension of 128. The number of layers and hidden dimensions are determined using an independent validation set (see **Supplementary Table 1**). Finally, a pairwise combination is performed between the enriched protein embeddings having a dimension of (L1xD) where L1 represents the length of the protein, and the enriched RNA embeddings having a dimension of (L2xD) where L2 represents the length of the RNA, resulting in protein-RNA pair embedding having a dimension of (L1xL2x2D).

Protein-RNA interaction map prediction

The protein-RNA pair embedding is passed as input to a two-stage deep learning architecture consisting of a ResNet-Inception module followed by a geometric attention module to predict inter-protein-RNA interaction maps. The ResNet-Inception module aims to capture the inter-component interactions between the residue-nucleotide pairs through a neural architecture comprising of a stack of residual neural networks (ResNet) having three parallel branches of convolutional layers (42,43). The three parallel branches each consists of two convolutional layers with kernel size of (1x9), (3x3), and (9,1), respectively. In each branch, an instance normalization followed by a Leaky ReLU (44) is applied to the output of the first convolutional layer, and the output of the activation layer is passed to the second convolutional layer. A shortcut connection is applied from the input to the output of all the three branches, and the final output followed by a Leaky ReLU activation is passed to the next ResNet-Inception block as input. Formally, the protein-RNA pair embedding is updated through the ResNet-Inception block as follows:

$$X_{1\times9}^{l+1} = \phi_{IN}(f_{1\times9}(\phi_{LReLU}(\phi_{IN}(f_{1\times9}(X^l)))))$$

$$X_{3\times3}^{l+1} = \phi_{IN}(f_{3\times3}(\phi_{LReLU}(\phi_{IN}(f_{3\times3}(X^l)))))$$

$$X_{9\times1}^{l+1} = \phi_{IN}(f_{9\times1}(\phi_{LReLU}(\phi_{IN}(f_{9\times1}(X^l)))))$$

$$X^{l+1} = \phi_{LReLU}(X^l + (X_{1\times9}^{l+1} + X_{3\times3}^{l+1} + X_{9\times1}^{l+1}))$$

where, X^l and X^{l+1} are the pair representation at layer l and $l+1$, respectively, ϕ_{IN} is the instance normalization, ϕ_{LReLU} is the Leaky ReLU activation operation. The 2D convolutions with kernel size of (1x9), (3x3), and (9x1) are represented by $f_{1\times9}$, $f_{3\times3}$, and $f_{9\times1}$, respectively. The network architecture of the ResNet-Inception module consists of 20 consecutive blocks, determined using an independent validation set (See **Supplementary Table 1**).

To account for the many-body effect in protein-RNA interactions while satisfying geometric consistency, a geometric attention module is subsequently employed that performs triangle-aware self-attention mechanism similar to a recent study (43). Formally, the triangle-aware self-attention mechanism is defined as follows:

$$q_{ij} = \phi_L(z_{ij}^l)$$

$$k_{ij} = \phi_L(z_{ij}^l)$$

$$v_{ij} = \phi_L(z_{ij}^l)$$

$$g_{ij} = \phi_{Sigmoid}(\phi_L(z_{ij}^l))$$

$$a_{ijk} = \phi_{Softmax}(\frac{1}{\sqrt{c}}q_{ij}^T k_{ik})$$

$$o_{ij} = g_{ij} \odot \Sigma_k a_{ijk} v_{ik}$$

$$z_{ij}^{l+1} = z_{ij}^l + (\phi_f(o_{ij}))$$

$$z_{ij}^{l+1} = z_{ij}^{l+1} + \phi_t(\phi_{ReLU}(\phi_t(\phi_{LNorm}(z_{ij}^{l+1}))))$$

where, ϕ_L is the linear operation that transforms the triangle-aware pair representation (z_{ij}^l) at layer l into query (q_{ij}), key (k_{ij}), and value (v_{ij}) representations for a pair (i, j). An additional gate (g_{ij}) representation is obtained from a linear transformation followed by a non-linear sigmoidal operation. The multi-head attention weight (a_{ijk}) between the three points (i, j, k) in the triangle is utilized to obtain the updated representation (o_{ij}), which is transformed through a final linear operation (ϕ_f). Finally, a non-linear transition transformation involving a layer normalization (ϕ_{LNorm}), ReLU Activision (ϕ_{ReLU}), and two linear operations (ϕ_t) is applied along with short-cut connection to obtain the updated triangle-aware pair representation (z_{ij}^{l+1}). The network architecture of the geometric attention module consists of 20 triange-aware self-attention blocks, determined using an independent validation set (See **Supplementary Table 1**).

The output from the geometric attention module undergoes a layer normalization, followed by a 1x1 convolutional layer for predicting the pairwise distance between the C_α atom of a protein residue and the C4' atom of an RNA nucleotide, resulting in a protein-RNA interaction map. The predicted distances are distributed across 37 bins, ranging from 2.5 Å to 20 Å with a step size of 0.5 Å, resulting in 36 bins. The final (37th) bin is considered for the distance exceeding 20 Å.

Training details

Our method is implemented using Pytorch 1.12.0 (45). During model training, we use cross entropy loss function and ADAM (46) optimizer with a learning rate of 1e-4. Additionally, cosine annealing from SGDR (47) and a weight decay of 1e-16 are used. The model training runs for at most 100 epochs on a single NVIDIA-A100 GPU. In addition to the full-fledged version of our method, we separately train multiple baseline models by discarding equivariant updates (thus making the model invariant), by replacing the protein- and RNA-specific EGNNs with convolutional blocks for protein-RNA pair embedding generation; and by discarding geometric attention module employing triangle-aware self-attention.

Transform-restrained geometry optimization

The predicted inter-protein-RNA interaction maps are transformed into multi-scale geometric restrains to optimize the relative spatial position and orientation of the protein and RNA components using PyRosetta-based relaxation algorithm (48). First, we transform the predicted distance maps into hybrid interaction maps with 10-level thresholding, similar to previous studies (49,50). Specifically, we obtain a hybrid interaction map with upper limits set at 2.5 Å, 4 Å, 6 Å, 8 Å, 10 Å, 12 Å, 14 Å, 16 Å, 18 Å, and 20 Å by summing the predicted likelihood values up to the respective distance thresholds. Subsequently, we generate 'BOUNDED' 'AtomPair' constraints in PyRosetta for the interactions having likelihood values ≥ 0.5 and utilize scaler-weighted function where the predicted likelihood values are incorporated as the scaler weights the lower bounds of the restraints are set to 2 Å, and the upper bounds are derived from the hybrid interaction maps. Finally, we feed the restraints to the Rosetta FastRelax algorithm (51) implemented in PyRosetta (48), which employs gradient-based optimization to generate 3D complex protein-RNA structural models. The initial pose originates from the previously described single-sequence protein and RNA

monomers generated by language model-based ESMFold and E2Efold-3D, respectively. We set the following parameters for initialization '-hb_cen_soft -constant_seed -relax:default_repeats 5 -default_max_cycles 200 -out:level 100', and utilize 'ref2015' as the scoring function and a maximum FastRelax iteration of 500.

Datasets, competing methods, and performance evaluation

To train our new method, ProRNA3D-single, we collect X-ray crystal structures of protein-RNA complexes having resolution up to 3.5 Å that are deposited to the PDB on and before October 2022. Furthermore, the redundancy among the collected structures is removed following previous studies (52,53). We also discard structures with a) proteins chain length>1500, b) RNA chain length>500, and c) RNA chain length<10. Finally, we obtain a total of 798 non-redundant interacting protein-RNA complex structures, from which we consider 750 targets for training and 48 targets for validation set.

Using protein-RNA complexes published more recently than any training- and validation-set targets, we evaluate the predictive modeling accuracy of our method. Our test set consists of targets released on and after November 2022 till November 2023. Following a recent study (54), we further filter the test targets ensuring that test protein chains do not share over 50% sequence identity with the train and validation set using CD-HIT (55) to reduce their overlap, resulting in a total of 39 protein-RNA complexes for the test set, denoted by Test_39 in our performance benchmarking.

We compare our method with three state-of-the-art methods for predicting biomolecular interactions in general and protein-RNA complex structures in particular, including RosettaFold2NA (RF2NA) (22), RosettaFold All Atom (RF2AA) (23), and AlphaFold 3 (AF3) (24). RF2NA predicts multimeric protein, RNA, DNA and their complexes by utilizing MSA information for protein, RNA as well as the paired protein/RNA using a 3-track architecture introduced in RoseTTAFold (21) with 1D-track utilizing multiple sequence alignments (MSAs), 2D-track utilizing pairwise distances, and 3D-track utilizing structural information. In RF2NA, the 1D-track is extended from RoseTTAFold by including the additional input tokens for nucleotides, 2D-track is extended by including the interaction between nucleotides and amino acids, and the 3D-track is extended by including the coordinate frames and 10 torsion angles for backbone and side chains for nucleic acids. RF2AA is a generalized biomolecular modeling and design system that includes the prediction of protein-RNA complex structures. RF2AA system represents arbitrary molecules as atom-bond graphs that initially are disconnected gas of amino acid residues, nucleic acid bases, and freely moving atoms, which are successively transformed through multiple blocks of the 3-track network architecture into physically plausible assembled structures. AlphaFold 3 is a unified deep learning framework for predicting biomolecular interactions that is substantially updated from AlphaFold2 (20) by introducing a lightweight Pairformer module for MSA processing and a diffusion-based architecture for directly predicting atomic coordinates starting from the single chain to complex structures including protein, nucleic acids, and small ligands. While RF2AA and RF2NA are publicly available methods, the code and model weights of the AlphaFold 3 are not yet publicly available, except for a webserver (alphafoldserver.com) with limited job submission and restrictions on its usage. Therefore, we obtain the predictions by AlphaFold 3 (AF3) by directly submitting the jobs in the alphafoldserver.com webserver and for the comparison, we consider the best model (model 0) out of 5 predictions based on its default self-estimation scores. The RF2NA and RF2AA predictions are obtained by running

them locally, after their inference code and model weights are downloaded in October 2023, and in April 2024, respectively. It is worth noting that all three competing state-of-the-art methods utilizes MSA and template information for prediction, whereas our new method, ProRNA3D-single is a single-sequence method. Therefore, for a fair comparison with our template- and MSA-free method ProRNA3D-single, we additionally run a customized version of RF2NA without MSA and template information, naming it RF2NA-single.

To evaluate the accuracy of the protein-RNA 3D complex structures, we use fnat score, a widely used CAPRI metric (56) for evaluating the interface of predicted biomolecular interactions. The metric fnat denotes the fraction of native contacts that are present in the predicted structure. In the case of protein-RNA interactions, a residue-nucleotide pair is considered to be in contact if their heavy atoms are within 5 Å. The fnat scores lie in range of [0-1], and the higher score represents better accuracy of the predicted complex structural model. We use two additional metrics named iRMS and lRMS for the evaluation of the predicted protein-RNA complex structures. The score for iRMS (interface root mean square) is relatively relaxed metric, where the interface is defined by an atomic contact cut-off of 10 Å, and the root mean square deviations of the backbone atoms are calculated after optimally superposing them with the native. The third metric lRMS (ligand root mean square) is measured by calculating the root mean square deviations of the shorter chain backbone atoms after optimally superposing the longer chain with the native. For both iRMS and lRMS, lower scores represent better accuracy of the predicted structural models. We use DockQ program (57) (downloaded in December 2023) for measuring the fnat, iRMS, and lRMS metrics.

Results

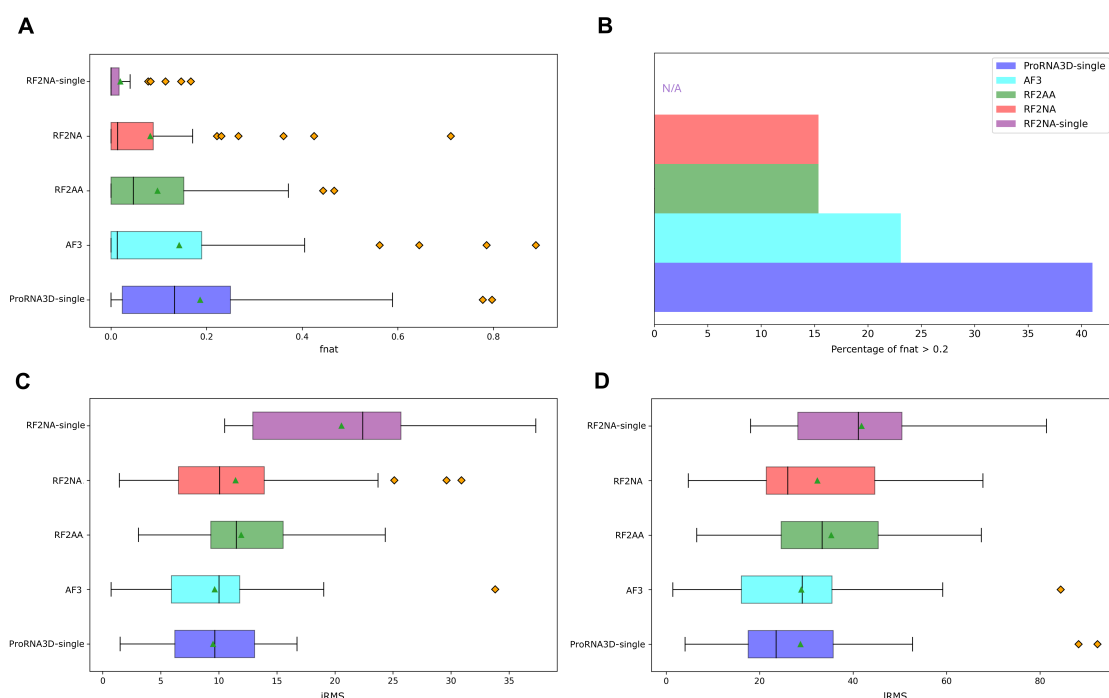### Test Set Performance

Fig 2. Test set performance of ProRNA3D-single and the competing methods. (**A**) The distributions of the fnat scores. (**B**) Percentage of "successful predictions" having fnat>0.2. (**C**) The distributions of the iRMS scores. (**D**) The distributions of the lRMS scores.

9

Fig 2 shows the predictive modeling accuracy of our new method ProRNA3D-single compared to the state-of-the-art deep learning-based methods AF3, RF2AA and RF2NA on the test set, Test_39. ProRNA3D-single attains the best fnat scores compared to the state-of-the-art methods, attaining 41.03% "successful prediction" having fnat>0.2, which is 25.64%, 25.64%, and 17.95% higher than RF2NA, RF2AA, and AF3, respectively (**Fig 2 B**), where a prediction is considered "successful" if the predicted structural model has fnat>0.2 following CAPRI criteria and previous studies (56,58). ProRNA3D-single also attains better performance compared to the competing methods in terms of iRMS (**Fig 2 C**) and lRMS (**Fig 2 D**). Of note, ProRNA3D-single is the only single-sequence method, whereas all competing methods take advantage of explicit evolutionary information in the form of MSAs as well as homologous templates whenever available. That is, ProRNA3D-single attains state-of-the-art performance despite being a single-sequence method. When compared head-to-head against the customized version of RF2NA without MSA and template information, RF2NA-single, which fails to predict any model with fnat>0.2, our method ProRNA3D-single outperforms across all metrics by a large margin.
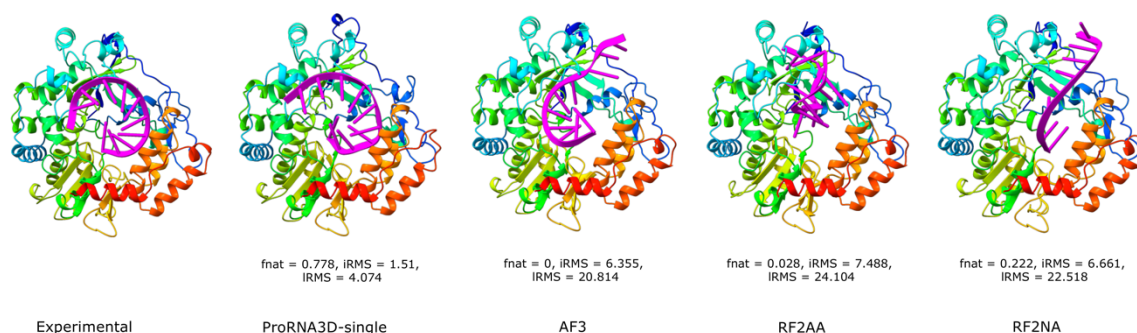

Experimental	ProRNA3D-single	AF3	RF2AA	RF2NA
	fnat = 0.778, iRMS = 1.51, lRMS = 4.074	fnat = 0, iRMS = 6.355, lRMS = 20.814	fnat = 0.028, iRMS = 7.488, lRMS = 24.104	fnat = 0.222, iRMS = 6.661, lRMS = 22.518

Fig 3. A representative example of protein-RNA complex structure prediction using ProRNA3D-single and the competing methods for PDB ID 7w9s, protein chain A, RNA chain C, representing TRANSFERASE/RNA of Enterovirus A71. Protein molecules are rainbow colored from blue to red from the N- to C-termini and RNA molecules are colored in magenta.

A representative example from the test dataset shown in **Fig 3** illustrates the advantage of ProRNA3D-single over the competing methods. This example, representing TRANSFERASE/RNA of Enterovirus A71 (PDB ID 7W9S, protein chain A, RNA chain C), falls in a distinct group of nucleic acid polymerases encoded by RNA viruses, which are crucial for viral genome replication and transcription, making them prime candidate for the development of antiviral drugs (59). Our method yields an accurate prediction attaining fnat score of 0.778 that is much better than AF3, RF2AA, and RF2NA. It is interesting to note that for this target, all competing methods except RF2NA fail to predict structures with fnat>0.2, with AF3 delivering the lowest performance. The performance advantage of ProRNA3D-single is striking with 0.778, 0.75, and 0.556 in fnat score points gain for ProRNA3D compared to AF3, RF2AA, and RF2NA, respectively.

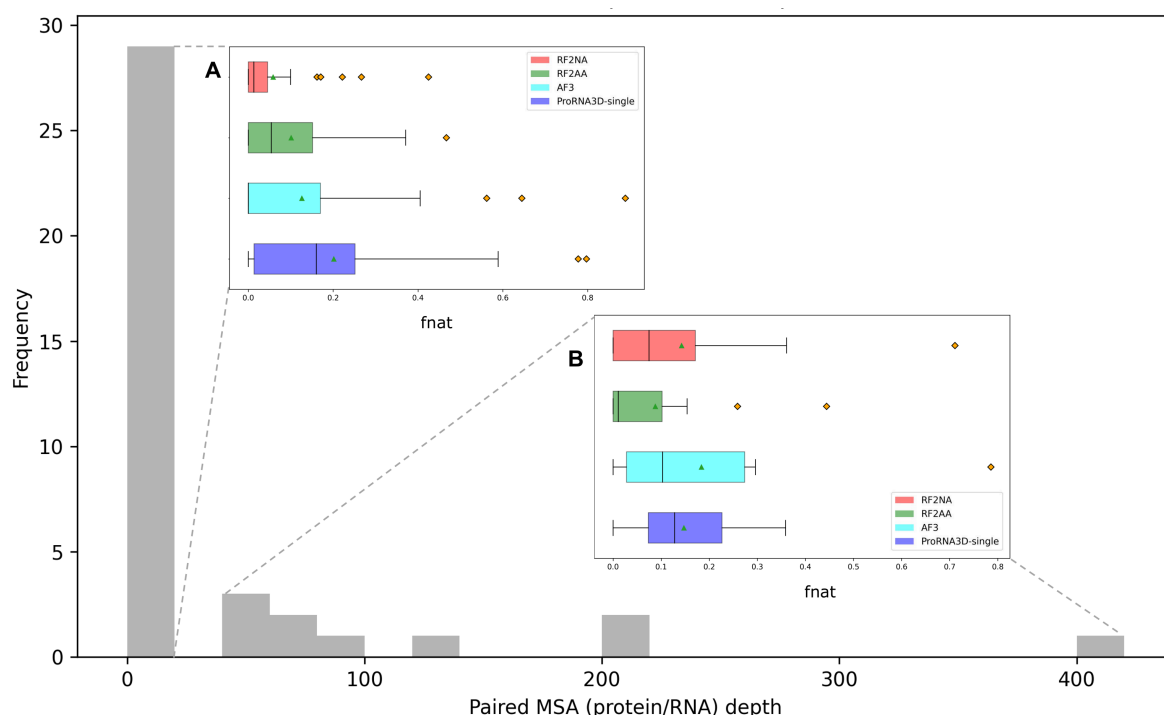## Effect of evolutionary sequence information

Fig 4. The distribution of evolutionary sequences in the Test_39 set is represented as the histogram of protein/RNA paired MSAs generated using the evolutionary sequence search engine of RF2NA. (**Inset A**) Performance of ProRNA3D-single and the competing methods when there is no sequence in paired MSA except for the query sequence (paired MSA depth = 1). (**Inset B**) Performance of ProRNA3D-single and the competing methods when there exist evolutionary sequences in paired MSA (paired MSA depth > 1).

While our new method, ProRNA3D-single, is a single-sequence method harnessing the power of biological language models, without utilizing any explicit evolutionary sequence information in the form of MSAs (or homologous templates), the competing methods rely on the availability of explicit MSA information (and homologous templates whenever available). To investigate how the availability of explicit evolutionary sequences in the form of protein/RNA paired MSAs affects the prediction accuracy of the competing methods compared to the MSA-agnostic ProRNA3D-single, we split the Test_39 set into two subgroups based on the paired MSAs generated using the MSA search engine of RF2NA : one where the MSA search is unsuccessful and there are no evolutionary sequences in paired MSA except for the query sequence (paired MSA depth = 1), and the other when there exists evolutionary sequences in paired MSA (paired MSA depth > 1). As shown in **Fig 4,** 28 out of 39 targets in the test set, the MSA search is unsuccessful where the protein/RNA paired MSA depth is 1. That is, for the majority of targets, there is no protein/RNA paired MSA, highlighting the scarcity of explicit evolutionary information and the urgent need to develop single-sequence methods for protein-RNA complex structure prediction. For the subset of targets where the paired MSA depth = 1, our single-sequence method ProRNA3D-single attains remarkably better accuracy than AF3, RF2NA, and RF2AA (**Fig 4 inset A, Supplementary Table 2**). For the subset of targets when

there exist evolutionary sequences (paired MSA depth > 1), ProRNA3D-single still attains state-of-the-art performance, second only to AF3 in terms of fnat score, while outperforming RF2AA and RF2NA (**Fig 4 inset B, Supplementary Table 2**), thereby exhibiting performance resilience. It is interesting to note that AF3 and RF2NA attain higher overall accuracy for the subset of targets having explicit evolutionary information (paired MSA depth > 1) compared to the subset where there are no evolutionary sequences in paired MSA except for the query sequence (paired MSA depth = 1), demonstrating that the availability of (or lack thereof) evolutionary sequences affects the predictive modeling accuracy of the MSA-dependent methods. In contrast, the prediction accuracy of our single-sequence method ProRNA3D-single, remains unaffected by explicit evolutionary information, demonstrating its robustness.
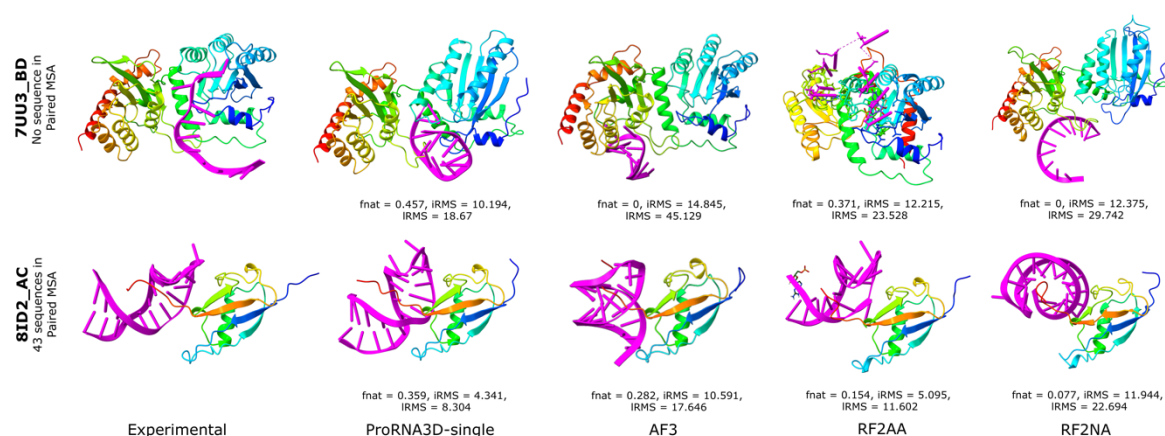

Fig 5. Protein-RNA complex structure prediction by ProRNA3D-single and the competing methods for two representative examples: HYDROLEASE/RNA for Macaca mulatta (PDB ID 7UU3, protein chain B and RNA chain D), and RNA Binding Protein/RNA for Homo Sapiens (PDB ID 8ID2, protein chain A, RNA chain C).

Two representative examples shown in **Fig 5** may help further shed light on the effect of evolutionary sequences. The first example, HYDROLEASE/RNA for Macaca mulatta (target ID 7UU3, protein chain B and RNA chain D), is from the subset of targets where there are no evolutionary sequences in paired MSA except for the query sequence (paired MSA depth = 1). ProRNA3D-single attains much better accuracy compared to the competing methods with 0.457, 0.457, and 0.086 in fnat score points gain for ProRNA3D-single compared AF3, RF2NA and RF2AA, respectively. The improved accuracy of ProRNA3D-single is visually noticeable, whereas AF3 and RF2NA fail to predict accurate protein-RNA interface and RF2AA suffers from unrealistic RNA conformation. The second example of RNA Binding Protein/RNA for Homo Sapiens (target ID 8ID2, protein chain A, RNA chain C) is from the subset having explicit evolutionary information (43 sequences in paired MSA). ProRNA3D-single still outperforms AF3, RF2NA and RF2AA by 0.077, 0.282, and 0.205 fnat score points, respectively. In both cases, our method consistently yields "successful prediction" having fnat>0.2, whereas only RF2AA in the first example and AF3 in the second example attain fnat>0.2 among the three competing methods.

Contribution of the interatomic protein-RNA interaction map prediction

A core component of our method is the prediction of interatomic (C_α - C'_4) protein-RNA interaction maps, which are subsequently transformed into geometric restraints to optimize the relative spatial position and orientation of the protein and RNA components, leading to protein-RNA complex structure prediction. Therefore, a natural question to ask is, does the quality of the predicted interaction maps translate into a good (or bad) complex structure prediction? To examine the contribution of interatomic protein-RNA interaction map prediction, we evaluate the precision of contact maps extracted from the hybrid interaction maps with 10-level thresholding in the Test_39 set. **Fig 6 A** represents the average precision of the predicted contacts at the terminal distance threshold of 20 Å. Considering the top1, top10, and top predicted contacts having likelihood values ≥ 0.5 at 20 Å threshold, our interatomic (C_α - C'_4) protein-RNA interaction predictions have the average precision of 43.59%, 35.385%, and 35.374% respectively (a full set of the average precision scores can be found in **Supplementary Table 3**). More importantly, the precision of the interaction prediction maintains a high correlation (Pearson r=0.75) with the fnat scores of the resulting complex structures of the corresponding targets (**Fig 6 B)**, verifying that the structural-level predictive modeling accuracy of our method is closely connected to the quality of the protein-RNA interaction maps.
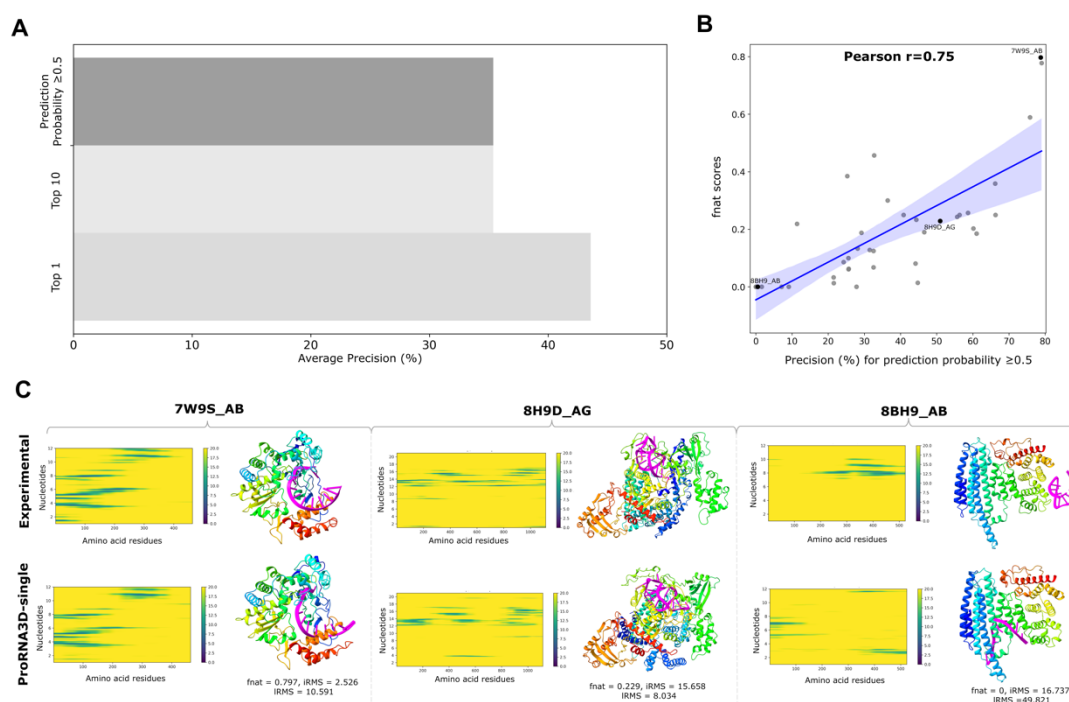

Fig 6. (**A**) The interatomic (C_α - C'_4) protein-RNA interaction prediction performance at 20 Å thresholds. Average precision for top 1, 10, and interactions with prediction probability ≥ 0.5 are reported. (**B**) Scatterplot of precision (%) for prediction probability ≥ 0.5 at 20 Å thresholds against the fnat scores in the Test_39 set. (**C**) Interatomic (C_α - C'_4) protein-RNA interaction maps predicted by our method compared to the native distance maps along with the corresponding predicted and native structures for three representative examples: TRANSFERASE/RNA of Enterovirus A71 (PDB ID 7W9S, protein chain A, RNA chain B), DNA binding protein of Lachnospiraceae bacterium (PDB ID 8H9D, protein chain A, RNA chain G), and RNA binding protein of Candida tropicalis MYA-3404 (PDB ID 8BH9, protein chain A, RNA chain B).

Fig 6 C shows three representative examples of our predicted interactions and the resulting 3D structural models, compared to the corresponding ground truth (experimental) distance maps and structures. For the TRANSFERASE/RNA of Enterovirus A71 (target ID 7W9S, protein chain A, RNA chain B), the interactions predicted by our method exhibit a clear similarity with the corresponding ground truth distance map, and the resulting 3D structural model demonstrates high accuracy, achieving an fnat score of 0.797. For the DNA binding protein of Lachnospiraceae bacterium (target ID 8H9D, protein chain A, RNA chain G), the predicted interactions are somewhat similar to the ground truth distance map, and the resulting structural model is deemed as an acceptable prediction with an fnat score of 0.229. Lastly, for RNA binding protein of Candida tropicalis MYA-3404 (target ID 8BH9, protein chain A, RNA chain B), our method yields inaccurate interaction prediction, showing significant dissimilarity to the corresponding ground truth distance map. Consequently, the resulting structural model is inaccurate. Overall, our analysis collectively demonstrates that the quality of predicted interactions is a key driver of the structural-level predictive modeling accuracy of our method, underscoring the importance of accurate inter-protein-RNA interaction prediction for improved protein-RNA complex structure prediction.

Ablation study

Recognizing the importance of the interatomic (C_α - C'_4) protein-RNA interaction prediction of our method, we investigate the relative contribution of different components of our interaction prediction networks. For that purpose, we conduct an ablation study to separately train multiple baseline models on the same train set by discarding equivariant updates (w/o equivariance) making the model invariant, by replacing the protein- and RNA-specific EGNNs with convolutional blocks for protein-RNA pair embedding generation (w/o graph embedding); and by discarding geometric attention module employing triangle-aware self-attention (w/o geometric attention) and compare their performance head-to-head with the full-fledged version of ProRNA3D-single on the Test_39 set.
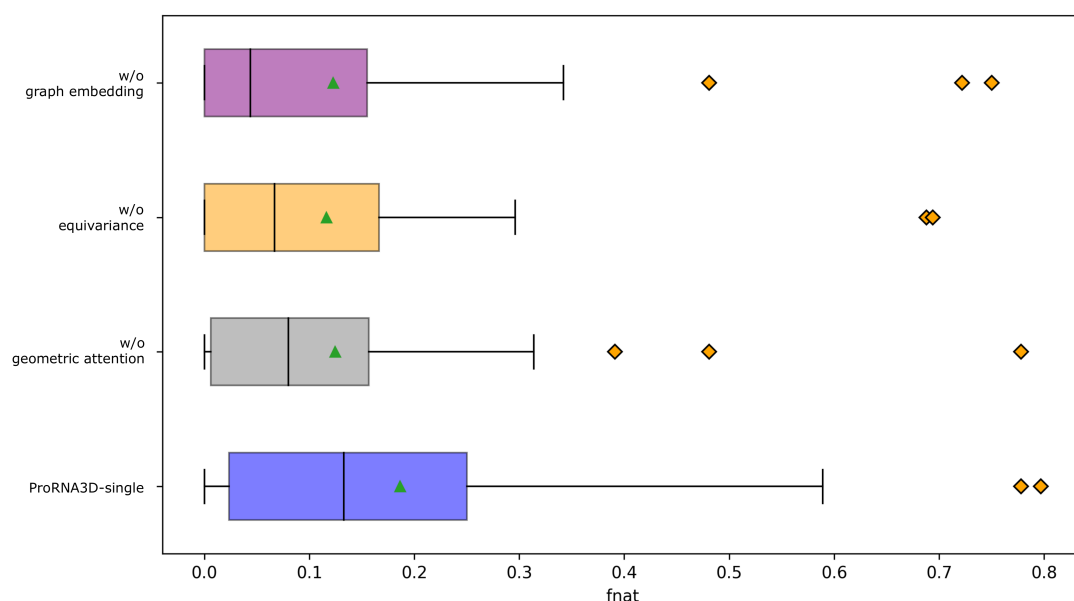

Fig 7. Ablation study showing the fnat distributions of ProRNA3D-single and the ablated baseline models on the Test_39 set.

Fig 7 shows the performance comparison of ProRNA3D-single full-fledged model and the ablated baseline models. The average fnat scores of model w/o geometric attention, model w/o equivariance, and w/o graph embedding are 0.125, 0.116, and 0.123, noticeably lower than that of ProRNA3D-single default (0.186). Furthermore, the ablated baseline models exhibit reduced consistency having 17.95%, 28.21%, and 17.95% lower "successful prediction" having fnat>0.2 than the full-fledged ProRNA3D. Overall, the results demonstrate the performance benefits of employing symmetry-aware graph convolutions and geometric attention module, which are the key modules of the neural architecture of ProRNA3D-single.

Discussion

This work introduces ProRNA3D-single, a new single-sequence method for predicting protein-RNA complex structures powered by biological language models. We demonstrate that ProRNA3D-single consistently outperforms the state-of-the-art deep learning-based protein-RNA complex structure prediction methods, including RoseTTAFold2NA, RoseTTAFold All-Atom, and AlphaFold 3. A major contribution of our work is the development of a generalized neural architecture for pairing biological language model embeddings, a previously unexplored avenue in the context of predicting biomolecular interactions. Experimental results show that our method attains substantially improved accuracy compared to the MSA-dependent methods when MSA information is limited, and exhibits remarkable robustness and performance resilience by attaining better accuracy in single-sequence protein-RNA complex structure prediction than what most state-of-the-art MSA-dependent methods can achieve even with explicit MSA information. Bypassing the dependence on the availability of explicit evolutionary information, which is not always abundant, our method paves the way to scalable and generalizable modeling of protein-RNA complex structures. Our ablation study reveals that the interplay of various modules in our neural network architecture cooperatively contributes to the improved accuracy, and the structural-level predictive modeling accuracy is closely connected to the quality of the protein-RNA interaction maps predicted through geometric attention-enabled pairing of biological language models.

Despite the improved performance of our method, there is still substantial room for improvement in protein-RNA complex structure prediction. A promising direction for future work is to investigate the possibility of incorporating biophysical "background knowledge" and inclusion of high-throughput experimental data in training to compensate for the lack of sufficient high-resolution structural data for protein-RNA complexes. Furthermore, we will study whether 3D modeling accuracy can be further improved through an integrative modeling approach that combines multimodal experimental data with deep learning-based predicted geometric restraints. Finally, while we find that ProRNA3D-single exhibits excellent predictive accuracy and remarkable robustness by geometric attention-enabled pairing of biological language models, an open challenge that remains is how to make the biological language models "partner-aware" specifically for biomolecular interaction prediction tasks. To this end, the possibility of fine-tuning biological language models using partner-specific knowledge may be explored. We expect our proposed method can be easily extended to predict the complex structures of other biomolecular interactions, including protein-DNA and protein-ligand as well as molecular assemblies containing proteins, small molecules, metals, and chemical modifications with improved accuracy and robustness.

Availability

An open-source software implementation of ProRNA3D-single, licensed under the GNU General Public License v3, is freely available at https://github.com/Bhattacharya-Lab/ProRNA3D-single.

Acknowledgements

Funding

References

1.   Gerstberger, S., Hafner, M. and Tuschl, T. (2014) A census of human RNA-binding proteins. *Nature Reviews Genetics*, **15**, 829-845.
2.   Glisovic, T., Bachorik, J.L., Yong, J. and Dreyfuss, G. (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, **582**, 1977-1986.
3.   Licatalosi, D.D. and Darnell, R.B. (2010) RNA processing and its regulation: global insights into biological networks. *Nature Reviews Genetics*, **11**, 75-87.
4.   Mittal, N., Roy, N., Babu, M.M. and Janga, S.C. (2009) Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. *Proceedings of the National Academy of Sciences*, **106**, 20300-20305.
5.   Noller, H.F. (2012) Evolution of protein synthesis from an RNA world. *Cold Spring Harbor perspectives in biology*, **4**, a003681.
6.   Cusack, S. (1999) RNA–protein complexes. *Current opinion in structural biology*, **9**, 66-73.
7.   Khalil, A.M. and Rinn, J.L. (2011), *Seminars in cell & developmental biology*. Elsevier, Vol. 22, pp. 359-365.
8.   Lukong, K.E., Chang, K.-w., Khandjian, E.W. and Richard, S. (2008) RNA-binding proteins in human genetic disease. *Trends in Genetics*, **24**, 416-425.
9.   Cooper, T.A., Wan, L. and Dreyfuss, G. (2009) RNA and disease. *Cell*, **136**, 777-793.
10. Modic, M., Ule, J. and Sibley, C.R. (2013) CLIPing the brain: Studies of protein–RNA interactions important for neurodegenerative disorders. *Molecular and Cellular Neuroscience*, **56**, 429-435.
11. König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M. and Ule, J. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*, **17**, 909-915.
12. Ke, A. and Doudna, J.A. (2004) Crystallization of RNA and RNA–protein complexes. *Methods*, **34**, 408-414.
13. Zheng, J., Kundrotas, P.J., Vakser, I.A. and Liu, S. (2016) Template-based modeling of protein-RNA interactions. *PLoS computational biology*, **12**, e1005120.
14. Zheng, J., Hong, X., Xie, J., Tong, X. and Liu, S. (2020) P3DOCK: a protein–RNA docking webserver based on template-based and template-free docking. *Bioinformatics*, **36**, 96-103.

15.   Arnautova, Y.A., Abagyan, R. and Totrov, M. (2018) Protein-RNA docking using ICM. *Journal of Chemical Theory and Computation*, **14**, 4971-4984.

16.   Tuszynska, I., Magnus, M., Jonak, K., Dawson, W. and Bujnicki, J.M. (2015) NPDock: a web server for protein–nucleic acid docking. *Nucleic acids research*, **43**, W425-W430.

17.   Huang, Y., Liu, S., Guo, D., Li, L. and Xiao, Y. (2013) A novel protocol for three-dimensional structure prediction of RNA-protein complexes. *Scientific reports*, **3**, 1887.

18.   Setny, P. and Zacharias, M. (2011) A coarse-grained force field for Protein–RNA docking. *Nucleic acids research*, **39**, 9118-9129.

19.   Yan, Y., Zhang, D., Zhou, P., Li, B. and Huang, S.-Y. (2017) HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic acids research*, **45**, W365-W373.

20.   Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A. and Potapenko, A. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.

21.   Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D. *et al.* (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, **373**, 871-876.

22.   Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D. and DiMaio, F. (2024) Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature Methods*, **21**, 117-121.

23.   Krishna, R., Wang, J., Ahern, W., Sturmfels, 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.

24.   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*.

25.   Mariani, V., Biasini, M., Barbato, A. and Schwede, T. (2013) lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, **29**, 2722-2728.

26.   Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The protein data bank. *Nucleic acids research*, **28**, 235-242.

27.   Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R. and Weirauch, M.T. (2018) The human transcription factors. *Cell*, **172**, 650-665.

28.   Brandes, N., Ofer, D., Peleg, Y., Rappoport, N. and Linial, M. (2022) ProteinBERT: a universal deep-learning model of protein sequence and function. *Bioinformatics*, **38**, 2102-2110.

29.   Ferruz, N., Schmidt, S. and Höcker, B. (2022) ProtGPT2 is a deep unsupervised language model for protein design. *Nature communications*, **13**, 4348.

30.   Chowdhury, R., Bouatta, N., Biswas, S., Floristean, C., Kharkar, A., Roy, K., Rochereau, C., Ahdritz, G., Zhang, J. and Church, G.M. (2022) Single-sequence protein structure prediction using a language model and deep learning. *Nature Biotechnology*, **40**, 1617-1623.

31.   Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., Guo, D., Ott, M., Zitnick, C.L. and Ma, J. (2021) Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, **118**, e2016239118.

32.   Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O. and Shmueli, Y. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, **379**, 1123-1130.

33.   Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) Attention is all you need. *Advances in neural information processing systems*, **30**.

34. Horne, J. and Shukla, D. (2022) Recent advances in machine learning variant effect prediction tools for protein engineering. *Industrial & engineering chemistry research*, **61**, 6235-6245.

35. Madani, A., Krause, B., Greene, E.R., Subramanian, S., Mohr, B.P., Holton, J.M., Olmos, J.L., Xiong, C., Sun, Z.Z. and Socher, R. (2023) Large language models generate functional protein sequences across diverse families. *Nature Biotechnology*, **41**, 1099-1106.

36. Shen, T., Hu, Z., Peng, Z., Chen, J., Xiong, P., Hong, L., Zheng, L., Wang, Y., King, I. and Wang, S. (2022) E2Efold-3D: end-to-end deep learning method for accurate de novo RNA 3D structure prediction. *arXiv preprint arXiv:2207.01586*.

37. Chen, J., Hu, Z., Sun, S., Tan, Q., Wang, Y., Yu, Q., Zong, L., Hong, L., Xiao, J., Shen, T. *et al.* (2022) Interpretable RNA Foundation Model from Unannotated Data for Highly Accurate RNA Structure and Function Predictions. 10.48550/arXiv.2204.00300.

38. Roche, R., Moussad, B., Shuvo, M.H. and Bhattacharya, D. (2023) E (3) equivariant graph neural networks for robust and accurate protein-protein interaction site prediction. *PLoS Computational Biology*, **19**, e1011435.

39. Roche, R., Moussad, B., Shuvo, M.H., Tarafder, S. and Bhattacharya, D. (2024) EquiPNAS: improved protein–nucleic acid binding site prediction using protein-language-model-informed equivariant deep graph neural networks. *Nucleic Acids Research*, **52**, e27-e27.

40. 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. *Nature Communications*, **14**, 7266.

41. Satorras, V.c.G., Hoogeboom, E. and Welling, M. (2021) In Marina, M. and Tong, Z. (eds.), *Proceedings of the 38th International Conference on Machine Learning*. PMLR, Proceedings of Machine Learning Research, Vol. 139, pp. 9323--9332.

42. Si, Y. and Yan, C. (2021) Improved protein contact prediction using dimensional hybrid residual networks and singularity enhanced loss function. *Briefings in Bioinformatics*, **22**.

43. Lin, P., Tao, H., Li, H. and Huang, S.-Y. (2023) Protein–protein contact prediction by geometric triangle-aware protein language models. *Nature Machine Intelligence*, **5**, 1275-1284.

44. Xu, J., Li, Z., Du, B., Zhang, M. and Liu, J. (2020), *2020 IEEE Symposium on Computers and communications (ISCC)*. IEEE, pp. 1-7.

45. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. *et al.* (2019) PyTorch: An Imperative Style, High-Performance Deep Learning Library. 10.48550/arXiv.1912.01703.

46. Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. 10.48550/arXiv.1412.6980.

47. Loshchilov, I. and Hutter, F. (2016) SGDR: Stochastic Gradient Descent with Warm Restarts. 10.48550/arXiv.1608.03983.

48. 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.

49. Roche, R., Bhattacharya, S. and Bhattacharya, D. (2021) Hybridized distance-and contact-based hierarchical structure modeling for folding soluble and membrane proteins. *PLOS Computational Biology*, **17**, e1008753.

50. Bhattacharya, S., Roche, R., Moussad, B. and Bhattacharya, D. (2022) DisCovER: distance-and orientation-based covariational threading for weakly homologous proteins. *Proteins: Structure, Function, and Bioinformatics*, **90**, 579-588.

51. Khatib, F., Cooper, S., Tyka, M.D., Xu, K., Makedon, I., Popović, Z., Baker, D. and Players, F. (2011) Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, **108**, 18949-18953.

52. Zeng, C., Jian, Y., Vosoughi, S., Zeng, C. and Zhao, Y. (2023) Evaluating native-like structures of RNA-protein complexes through the deep learning method. *Nature Communications*, **14**, 1060.

53. Zhao, H., Yang, Y. and Zhou, Y. (2011) Structure-based prediction of RNA-binding domains and RNA-binding sites and application to structural genomics targets. *Nucleic acids research*, **39**, 3017-3025.

54. Huang, T., Song, Z., Ying, R. and Jin, W. FAFormer: Frame Averaging Transformer for Predicting Nucleic Acid-Protein Interactions.

55. Huang, Y., Niu, B., Gao, Y., Fu, L. and Li, W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics*, **26**, 680-682.

56. Lensink, M.F., Velankar, S. and Wodak, S.J. (2017) Modeling protein–protein and protein–peptide complexes: CAPRI 6th edition. *Proteins: Structure, Function, and Bioinformatics*, **85**, 359-377.

57. Basu, S. and Wallner, B. (2016) DockQ: a quality measure for protein-protein docking models. *PloS one*, **11**, e0161879.

58. Huang, S.-Y. and Zou, X. (2014) A knowledge-based scoring function for protein-RNA interactions derived from a statistical mechanics-based iterative method. *Nucleic Acids Research*, **42**, e55-e55.

59. Li, R., Wang, M. and Gong, P. (2022) Crystal structure of a pre-chemistry viral RNA-dependent RNA polymerase suggests participation of two basic residues in catalysis. *Nucleic Acids Research*, **50**, 12389-12399.