nature machine intelligence

Article

https://doi.org/10.1038/s42256-026-01223-x

Predicting RNA 3D structure and conformers using a pre-trained secondary structure model and structure-aware attention

Received: 26 May 2025

Accepted: 23 March 2026

Published online: 21 April 2026

Check for updates

Wenkai Wang, Zhenling Peng & Jianyi Yang

Determining RNA three-dimensional (3D) structure and conformers remains a grand challenge in structural biology, primarily owing to the scarcity of experimental data, the intrinsic flexibility of RNA molecules, and the limitations of current experimental and computational methods. Here we propose trRosettaRNA2, a deep learning-based end-to-end approach to this problem. Considering the scarcity of RNA 3D structure data, trRosettaRNA2 integrates an auxiliary secondary structure (SS) prior module, pre-trained on extensive SS data, to generate informative base-pairing priors. This module also serves as an independent RNA SS prediction method, trRNA2-SS, and achieves state-of-the-art performance. To enable end-to-end prediction, trRosettaRNA2 uses SS-aware attention to generate RNA 3D structure and conformers (distinct 3D spatial arrangements of the same molecule resulting from its intrinsic flexibility). Rigorous benchmarks demonstrate that trRosettaRNA2 outperforms other RNA 3D structure prediction methods, despite using substantially fewer parameters and computational resources. Notably, its flexibility in leveraging diverse secondary structure inputs provides a pathway to generate accurate 3D structure and explore the RNA conformers. Based on trRosettaRNA2, our group, Yang-Server, was the top automated server for RNA structure prediction in the CASP16 blind test, surpassing AlphaFold 3. This performance highlights that trRosettaRNA2 represents a solid step forward for RNA structure prediction. Application to the ribonuclease P RNA demonstrates that trRosettaRNA2 successfully captures its structural heterogeneity even without requiring experimental data, showing its potential to predict RNA conformational ensembles.

Determining the structures and conformations of RNA is crucial for elucidating biological functions fundamental to life, yet remains experimentally challenging owing to the intrinsic flexibility and instability of RNA[1–3]. To address this challenge, decades of effort have been dedicated to developing computational methods for RNA structure prediction[4–10]. In recent years, deep learning has been introduced to improve performance[11–18]. For example, trRosettaRNA[11] and DeepFoldRNA[12] use transformers to predict geometry for energy minimization; RhoFold+[14] introduces RNA language models; while RoseTTAFoldNA[16], RoseTTAFold All-Atom[17] and AlphaFold 3[18] target broader biomolecular systems. A recent method, GraphaRNA, utilizes graph neural networks and diffusion models to predict RNA substructures[19]. Benchmarks and blind tests have confirmed that deep learning enhances automated modelling accuracy[11,12,14].

Nevertheless, RNA-Puzzles and Critical Assessment of Structure Prediction (CASP) competitions, the leading community-wide

MOE Frontiers Science Center for Nonlinear Expectations, State Key Laboratory of Cryptography and Digital Economy Security, Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, China. e-mail: yangjy@sdu.edu.cn

experiments in the field, highlighted that RNA structure prediction is far from being solved[20–23]. The challenge stems from several aspects. First, the scarcity of experimental three-dimensional (3D) data in the Protein Data Bank (PDB)[24] (only ~4% of entries contain RNAs) hinders the performance of data-driven methods. Second, the quality of multiple sequence alignment (MSA), which is essential for current structure prediction paradigms, is often suboptimal for RNA. Specifically, the alignment specificity and sensitivity of sequence-based tools (for example, BLASTN[25]) are limited by RNA's simple four-base alphabet and their failure to account for structural conservation that often outweighs sequence identity[23,26]; advanced structure-aware methods[27–29] are hampered by the prohibitive computational cost of iterative covariance model-based searches and the quality of the initial secondary structure (SS) priors. Third, RNA is inherently more flexible than proteins[1–3], a characteristic largely overlooked by existing methods.

Recently, a few studies have begun to predict RNA conformers by integrating computational methods with experimental data. For example, HORNET, a computational framework, was proposed to determine RNA 3D conformations from atomic force microscopy (AFM) images[30,31]. However, the de novo modelling of RNA structural heterogeneity without relying on experimental data remains a major challenge.

To address 3D data scarcity and capture RNA dynamics, we leverage the abundance of SS data. Given that RNA folding is largely hierarchical[32–34], SS information serves as a critical foundation for scaffolding the 3D structure. This relationship formed the basis of early RNA 3D prediction methods, as originally introduced by the MC-Fold/MC-Sym pipeline[35]. Notably, databases such as bpRNA[36] contain ~12 times more entries than the RNA structures in PDB, providing a rich resource to mitigate data scarcity. Moreover, SS variations often link to 3D conformational heterogeneity. Therefore, we argue that an effective solution for predicting RNA 3D structures and conformers can come from a close connection between RNA 2D and 3D structures.

Following this rationale, we propose trRosettaRNA2, an end-to-end (E2E) method to predict RNA 3D structure and conformers. Distinct from traditional approaches that treat SS merely as external static constraints, trRosettaRNA2 incorporates an internal SS module, which is pre-trained to learn fundamental RNA folding rules from massive SS data. Evaluated as an independent SS predictor (designated as trRNA2-SS), it achieves state-of-the-art performance. To facilitate E2E prediction, we introduce an SS-aware structure module to directly predict the 3D structure. Rigorous benchmarks and CASP16 blind assessments[37,38] demonstrate that trRosettaRNA2 is competitive with AlphaFold 3, despite requiring substantially fewer parameters and training resources, while consistently outperforming other automated methods. Notably, trRosettaRNA2 successfully recapitulates the heterogeneous conformational landscape of ribonuclease P (RNase P) RNA as observed in AFM experiments.

Results

Overview of trRosettaRNA2

As depicted in Fig. 1, trRosettaRNA2 uses an E2E pipeline for RNA 3D structure prediction (Supplementary Methods 2 provides the detailed pseudocode). The pipeline comprises four primary components: an SS prior module for base-pairing initialization, an RNAformer module for feature encoding and updating, an SS-aware structure module for coordinate prediction, and a post-processing refinement procedure.

Specifically, starting from the target RNA sequence, an MSA is generated and encoded into feature embeddings (Algorithm 2 in Supplementary Methods 2). Subsequently, the SS prior module (Algorithm 3 in Supplementary Methods 2), pre-trained on the extensive bpRNA database[36], generates SS priors in the form of base-pairing probabilities. These priors, integrated with the MSA embeddings, are encoded into one-dimensional (1D) and two-dimensional (2D) representations and updated by RNAformer. As a two-track transformer, RNAformer, iteratively updates the bimodal representations through axial attention and cross-modal interactions (pair-biased attention and outer-product expansion; Algorithm 4 in Supplementary Methods 2). Finally, the SS-aware structure module predicts the all-atom Cartesian coordinates from the updated representations.

To ensure physical realism, the predicted coordinates undergo a rapid relaxation protocol, involving bond idealization and energy-guided clash minimization. This protocol is implemented using PyRosetta[39]. PyRosetta was selected for its physically grounded energy functions and mature refinement protocols, which are community standards for fine-grained molecular modelling. Its Python-based interface also allows for seamless integration of Rosetta's versatile optimization algorithms into our workflow. We refer to this integrated workflow as trRosettaRNA2 (E2E), standing for end-to-end. Alternatively, the final 3D structure can be reconstructed via PyRosetta energy minimization based on the geometric constraints derived from the predicted 2D representations, bypassing the structure module. This latter approach is termed trRosettaRNA2 (PyRosetta).

Notably, the SS prior module, beyond providing critical base-pairing priors, functions as a standalone RNA SS predictor termed trRNA2-SS. Comprehensive benchmarks demonstrate that it achieves state-of-the-art (SOTA) performance (Supplementary Results 1, Supplementary Fig. 1 and Supplementary Tables 1–4).

Performance of trRosettaRNA2 on RNA 3D structure prediction

In addition to the SS prior module, another major advancement of trRosettaRNA2 over trRosettaRNA lies in its E2E architecture powered by an SS-aware structure module. To assess the 3D structure prediction capabilities of trRosettaRNA2, we constructed a benchmark test set, TS28, which was carefully curated to ensure no redundancy with the training data (Methods). This set provides a standard, unbiased benchmark, characterized by internal non-redundancy and the absence of interactions with non-RNA macromolecules. For a comprehensive evaluation of performance, we assess multiple metrics, including global topology measures—root-mean-square deviation (RMSD) and template modeling score (TM-score)—and local accuracy metrics, namely the local distance difference test (lDDT) score and interaction network fidelity (INF) (see Methods for details).

To directly compare trRosettaRNA2 with the original trRosettaRNA, which takes SPOT-RNA-predicted SSs as input, we initially evaluate trRosettaRNA2 using SS information from SPOT-RNA. This is achieved by bypassing the internal SS prior module and directly feeding the base-pairing probabilities predicted by SPOT-RNA into the RNAformer module. As shown in Fig. 2a and Table 1, trRosettaRNA2 (E2E) achieves an average RMSD of 9.08 Å on the TS28 dataset, an ~20% improvement (lower RMSD) compared with trRosettaRNA. With the incorporation of the SS prior module, that is, using the enhanced SS predictor trRNA2-SS, all four accuracy metrics are further improved. This demonstrates the effectiveness of the E2E architecture adopted in trRosettaRNA2.

Interestingly, we observe that the PyRosetta version of trRosettaRNA2 is slightly more accurate than the E2E version, probably owing to the enhanced structural plausibility achieved through energy minimization. Considering that the energy minimization procedure is primarily guided by the 2D geometries predicted by the neural network, this result suggests that the E2E training also boosts the 2D geometries prediction. Consequently, the final trRosettaRNA2 achieves an average RMSD of 8.66 Å, representing a 24% improvement over trRosettaRNA (Fig. 2a and Table 1), and outperforms its predecessor on 71.4% of targets (Fig. 2b). The impact of individual model components is discussed in Supplementary Results 2.

Comparison of trRosettaRNA2 and other methods

In our previous work trRosettaRNA, we have demonstrated that deep learning-based methods outperform traditional automated

https://doi.org/10.1038/s42256-026-01223-x

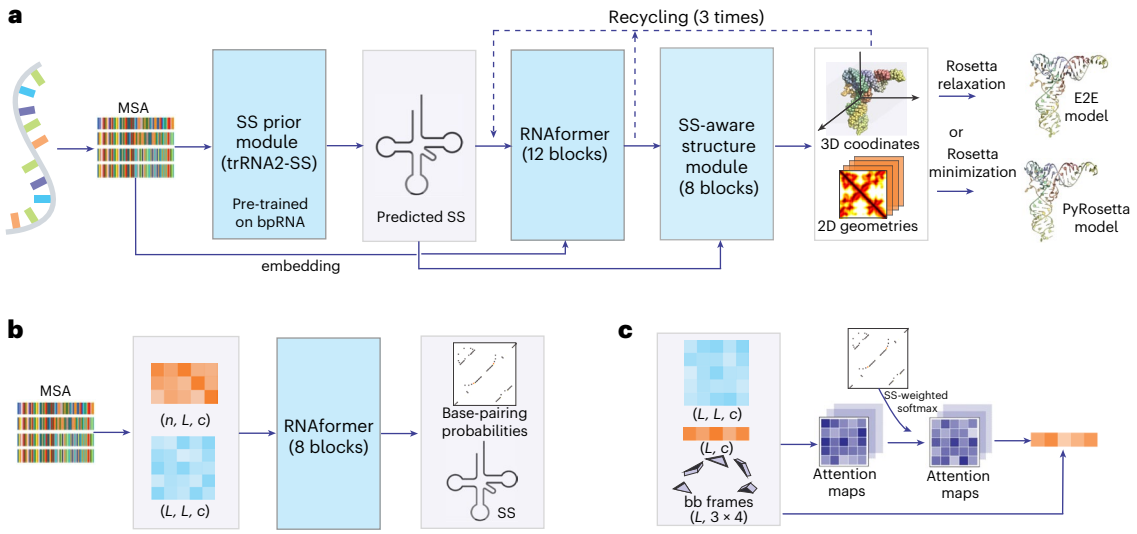

Fig. 1 | Overview of the trRosettaRNA2 pipeline. a, Diagram for RNA 3D structure prediction. The blocks with blue backgrounds represent the neural networks utilized, and those with grey backgrounds indicate the data embeddings used. If the E2E model predicts substantial steric clashes, Rosetta energy minimization will be applied to fold the RNA structure, guided by deep learning constraints. **b**, Architecture of the pre-trained SS prior module, trRNA2-SS. The input MSA is transformed into two distinct representations: MSA representation (highlighted in orange) and pair representation (highlighted in blue). The variables *n*, *L* and *c* are the number of sequences in the MSA, the length of the RNA sequence and the number of channels, respectively. RNAformer is a transformer-based neural network similar to those used in trRosettaRNA. **c**, Architecture of an SS-aware invariant point attention block from the structure module. It utilizes predicted SS information, specifically base-pairing probabilities, to optimize the attention weights between any two nucleotides in the query RNA. This adjustment enhances the model's ability to accurately capture the interactions and relationships between nucleotides based on their predicted SS. bb, RNA backbone.

methods relying on physical energy and/or fragment assembly. Thus, in this study, we only compare trRosettaRNA2 with other SOTA deep learning-based methods, including DeepFoldRNA and RoseTTAFoldNA (published around the same time as trRosettaRNA), as well as more recently published methods RhoFold+ and AlphaFold 3. Unless otherwise specified, all methods were run locally using identical inputs (MSA and/or SS).

As shown in Table 1 and Fig. 2a, trRosettaRNA2 shows competitive performance with AlphaFold 3 and outperforms other methods on TS28. Specifically, trRosettaRNA2 achieves better RMSD and lDDT than AlphaFold 3, although with a slightly lower TM-score and INF, while consistently outperforming all other representative methods across all metrics. Notably, as illustrated in Fig. 2c, trRosettaRNA2 shows a performance profile complementary to that of AlphaFold 3: trRosettaRNA2 outperforms AlphaFold 3 by >1 Å on 32.1% of targets, while AlphaFold 3 outperforms trRosettaRNA2 by the same margin on 25%. Furthermore, trRosettaRNA2 models show substantially fewer steric clashes (~2 versus >10 for AlphaFold 3) and are devoid of geometric entanglements[40–42] (verified via RNAspider[43]; Supplementary Table 5), highlighting the efficacy of its violation-based fine-tuning strategy (see ablation study in Supplementary Fig. 3) and energy-based refinement protocols. A similar result is observed in the CASP15 benchmark (Fig. 2d), where trRosettaRNA2 achieves an average RMSD second only to the top human group, AIchemy_RNA2 (see Supplementary Results 3 for a detailed discussion). It also outperforms two recently released methods DRfold2[44] and RNAbpFlow[45] (Supplementary Results 4 and Supplementary Tables 9 and 10).

One of the key advantages of trRosettaRNA2 lies in its parameter efficiency. As shown in Fig. 2e, trRosettaRNA2 achieves competitive performance with only ~30 million parameters. This represents a substantial reduction compared with other E2E methods: only 1/2 of RoseTTAFoldNA, 1/4 of RhoFold+ and 1/12 of AlphaFold 3. This efficiency is also observed in training cost (Fig. 2f), as the training of trRosettaRNA2 required only ~33 days on a single A100/A800 graphics processing unit (GPU; including the SS-based pre-training), substantially less than RhoFold+ (~1 week on 8 A100 80GB GPUs), RoseTTAFoldNA (~4 weeks on 64 GPUs; the specific GPU type was not mentioned in its paper), and AlphaFold 3 (~20 days on 256 A100 GPUs). These findings highlight the remarkable efficiency of trRosettaRNA2.

Figure 2g,h presents an example RNA: a NAD-II riboswitch with PDB ID 8I3Z, comprising a 5′-triple helix and a 3′-pseudoknot. Although most methods (excluding RoseTTAFoldNA) roughly reconstructed these local motifs, they struggled to predict the correct spatial arrangement of the 3′ terminus relative to the P1a helix (highlighted in Fig. 2g,h). This misorientation led to poor global accuracy (RMSD, 11.8–19.3 Å). Conversely, trRosettaRNA2 accurately resolved the tertiary packing of these domains, achieving a lowest RMSD of 9.3 Å. In addition, the trRosettaRNA2 model showed substantially fewer steric clashes (clashscore, 3.3) compared with other methods (~7.6–80.6).

Blind test results in CASP16

The design of the SS prior module allows trRosettaRNA2 to enhance performance by exploring diverse prior information (for example, the combination of the SS from SPOT-RNA and trRNA2-SS further improves accuracy; Table 1 and Supplementary Results 2), whereas other SOTA methods such as AlphaFold 3 and DRfold2 rely solely on the co-evolutionary signal extracted by MSA or language model. Based on trRosettaRNA2, we participated in the blind tests of the CASP16 RNA prediction experiment as two automated server groups, Yang-Server (TS052) and Yang-Multimer (TS456). Yang-Multimer used the default trRosettaRNA2 configuration (that is, using the trRNA2-SS prior) to compare with the other four deep learning predictors. Yang-Server was further optimized for submitting the best predictions using diverse SS resources[38], which are trivial to obtain (see Methods for details).

Yang-Server results. According to the official ranking for the 36 RNA monomer targets (https://predictioncenter.org/casp16/zscores_rna.cgi), Yang-Server is the top-performing server group (1/16), either ranked by the 'first' models or the 'best' models, and is surpassed

https://doi.org/10.1038/s42256-026-01223-x

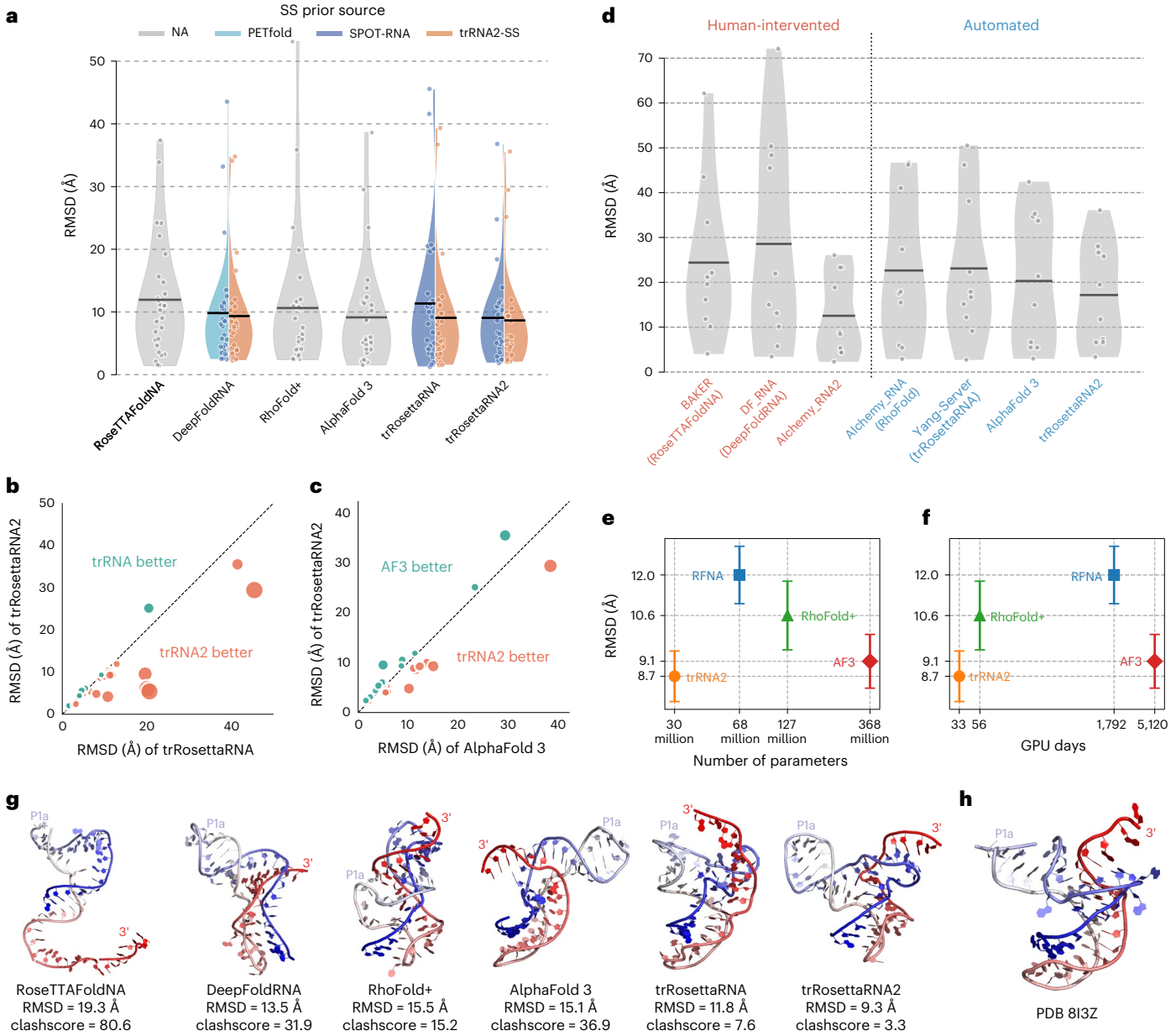

Fig. 2 | Performance of trRosettaRNA2 and other methods for RNA 3D structure prediction on the benchmark datasets. a,**d**, RMSD distributions of different methods on the TS28 dataset ($n = 28$; **a**) and the CASP15 dataset ($n = 10$; **d**). Horizontal black lines indicate the mean RMSD values. In **d**, methods are categorized into human-intervened (red labels) and automated server (blue labels) groups. **b**,**c**, Head-to-head RMSD comparisons of trRosettaRNA2 (trRNA2) against trRosettaRNA (trRNA; **b**) and AlphaFold 3 (AF3; **c**) on the TS28 set. Orange points below the diagonal line indicate better performance by trRosettaRNA2. **e**,**f**, Computational efficiency analysis comparing the RMSD on TS28 ($n = 28$) against the number of parameters (**e**) and training cost in GPU days (that is, GPU count × training days; **f**) for E2E methods. Error bars represent the mean ± 0.1 s.d. Note that the metrics for trRNA2 include its SS module. The x axes in **e** and **f** are logarithmically scaled for enhanced visualization. RFNA, RoseTTAFoldNA. **g**,**h**, Visual comparison of 3D structures predicted by different methods (**g**) against the experimental reference structure (**h**) for an example target (PDB ID 8I3Z).

by only 3 human groups (that is, 4th among 64 participating groups). In contrast, AF3-server (based on AlphaFold 3; TS304) ranks 9th among the 64 groups, and dNAfold (using DRfold2; TS448) ranks 24th.

Owing to computational resource constraints on the training crop size, we incorporated AlphaFold 3 predictions for targets exceeding 400 nucleotides that lacked homologous templates (Methods). Nevertheless, for the remaining 23 targets, Yang-Server still ranks as the best server group (Fig. 3a). As shown in Supplementary Table 11, the 'best' models for these 23 targets from Yang-Server outperform those from AF3-server across global metrics (RMSD, TM-score and global distance test total score (GDT-TS)) and show competitive performance in local metrics (lDDT and INF), while achieving a much lower clashscore. In addition, Yang-Server surpasses dNAfold (based on DRfold2) across all the metrics, again with a substantially lower clashscore.

Yang-Multimer results. We use the Yang-Multimer group to perform blind benchmark tests between the default trRosettaRNA2 and other deep learning methods (trRosettaRNA, RoseTTAFoldNA, DeepFoldRNA and RhoFold; see Methods for details). According to the CASP16 official rankings, Yang-Multimer places 5th among the 16 server groups ranked by the summed Z-score (>0) of the 'first' models. Note that only 23 targets with Yang-Multimer submissions are included in the

https://doi.org/10.1038/s42256-026-01223-x

Table 1 | Comparison between trRosettaRNA2 and other representative methods on the TS28 dataset

Method	SS predictor	RMSD (Å) ↓	lDDT ↑	TM-score ↑	INF ↑	clashscore ↓
RoseTTAFoldNA	–	11.95	0.610	0.328	0.722	71.7
DeepFoldRNA	PETfold	9.56	0.671	0.349	0.752	15.5
	trRNA2-SS	8.98	0.672	0.353	0.762	14.3
RhoFold+	–	10.63	0.640	0.333	0.717	11.3
AlphaFold 3	–	9.14	**0.701**	**0.387**	**0.814**	11.6
trRosettaRNA	SPOT-RNA	11.35	0.641	0.336	0.723	3.4
	trRNA2-SS	9.07	0.685	0.360	0.734	3.4
trRosettaRNA2 (E2E)	SPOT-RNA	9.08	0.665	0.358	0.765	3.6
	trRNA2-SS	9.07	0.682	0.361	0.770	5.1
	pLDDT top1	8.53	0.686	0.368	<u>0.781</u>	4.5
trRosettaRNA2 (PyRosetta)	SPOT-RNA	9.06	0.677	0.360	0.744	2.9
	trRNA2-SS	8.66	0.695	0.364	0.765	**2.0**
	pLDDT top1	**8.37**	<u>0.698</u>	<u>0.371</u>	0.769	<u>2.2</u>

The downwards arrow indicates that lower is better The upwards arrow indicates that higher is better. The best results are shown in bold and the second-best results are underlined.

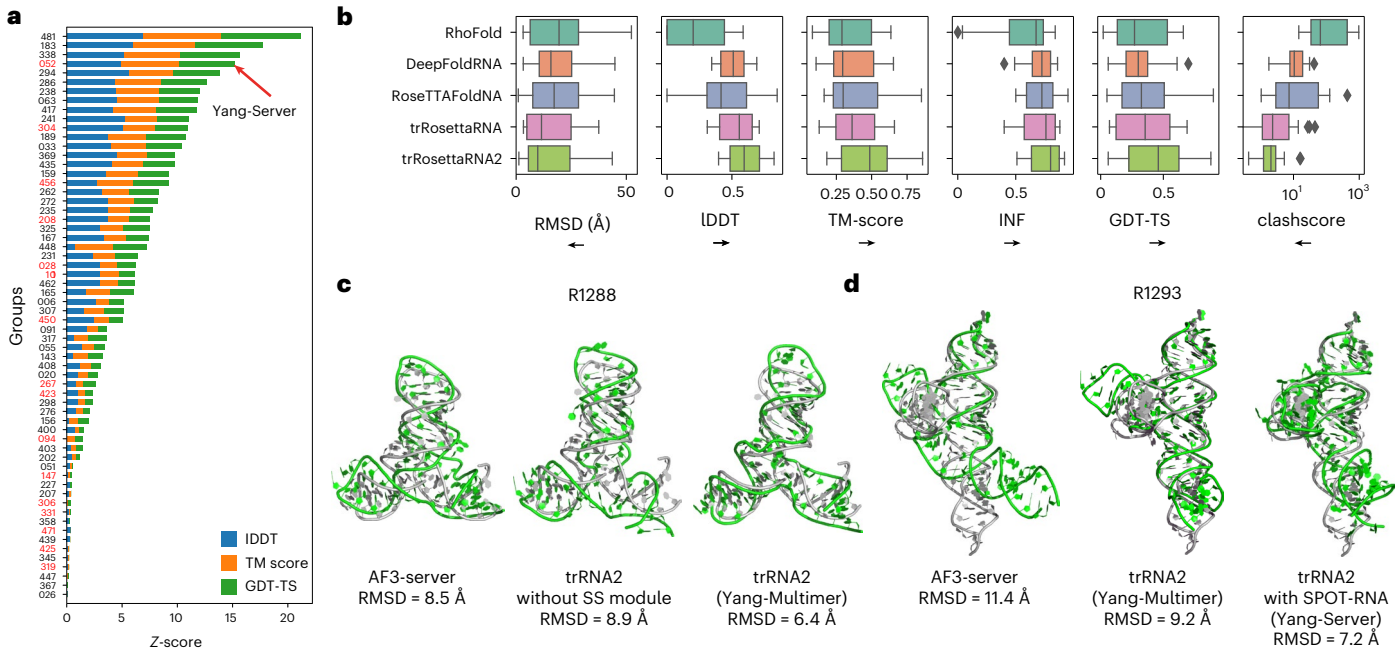

Fig. 3 | CASP16 blind test results. a, Ranking of the summed *Z*-score (>0; calculated according to CASP16 official guidelines) for the 23 RNA targets shorter than 400 nucleotides or with templates ($n = 23$). Automated server groups are highlighted in red on the *y*-axis tick labels. **b**, Results for blind benchmark test performed by Yang-Multimer submission ($n = 26$ targets with Yang-Multimer submissions; note that the RhoFold and RoseTTAFoldNA predictions were not relaxed). Arrows below the *x* axes indicate the direction of better performance. For visualization clarity, the RMSD *x* axis is truncated at 55 Å to exclude sporadic outliers (from DeepFoldRNA and trRosettaRNA), and the clashscore *x* axis is logarithmically scaled. The box in each plot spans the interquartile range, with the median marked by a line. Whiskers extend to 1.5 times the interquartile range, and any data points outside this range are plotted as outliers. **c**,**d**, Two examples to illustrate the comparison between trRosettaRNA2 (trRNA2) and AlphaFold 3 (AF3-server). Predicted structures (in green) are superimposed onto the experimental structures (in grey).

official ranking, which means that the *Z*-scores for the remaining 13 targets are both 0. If ranked by the average *Z*-score, Yang-Multimer is the second-best server group (behind Yang-Server) among the 37 groups (10 server groups) with more than 20 submissions, reflecting the superior performance of trRosettaRNA2.

Figure 3b and the top portion of Table 2 summarize the results of Yang-Multimer. Among the Yang-Multimer submissions, trRosettaRNA2 outperforms the other four methods in terms of both accuracy and physical plausibility. We also extended our analysis to the subsequently published RhoFold+[14], finding its performance to be on par with RhoFold. These findings align with our earlier observations on the TS28 dataset.

Case study for CASP16. The competitive performance of trRosettaRNA2 over AlphaFold 3 can be attributed to the incorporation of prior information via the SS module, leveraging the fact that SSs are generally more readily predictable than complex 3D structures. Figure 3c,d presents two examples to illustrate this.

Table 2 | Performance on CASP16 RNA monomer targets with Yang-Multimer submissions

Group	Method	RMSD (Å) ↓	lDDT ↑	TM-score ↑	GDT-TS ↑	INF ↑	clashscore ↓
Yang-Multimer	RhoFold[b]	19.9	0.224	0.328	0.298	0.562	223.9
	DeepFoldRNA	20.1	0.509	0.345	0.312	0.701	14.8
	RoseTTAFoldNA[b]	18.8	0.428	0.387	0.355	0.718	44.4
	trRosettaRNA	20.1	0.530	0.385	0.351	0.708	6.9
	trRosettaRNA2	**15.9**	**0.598**	**0.464**	**0.432**	**0.759**	**3.2**
Yang-Server	First model	13.8	0.635	0.510	0.464	0.790	3.7
	Best model	**10.2**	0.663	**0.548**	**0.496**	0.815	**1.4**
AF3-server	First model	12.4	0.659	0.504	0.457	0.809	16.3
	Best model	11.3	**0.675**	0.527	0.478	**0.818**	11.5
Reduced set[a]	RhoFold[b]	19.6	0.496	0.317	0.313	0.522	199.6
	RhoFold+	25.6	0.512	0.317	0.320	0.636	65.7
	trRosettaRNA2	**14.3**	**0.636**	**0.443**	**0.429**	**0.758**	**2.1**

All metrics are sourced from the official CASP16 evaluation unless specified otherwise. The best results in each method group are shown in bold, and the second-best results are underlined. [a]Twenty targets with available experimental structures. The metric values on this set are slightly different from the official results as we manually calculated them. [b]RhoFold and RoseTTAFoldNA predictions were not relaxed.

The first example is R1288, a stabilized S-adenosylmethionine analogue-utilizing ribozyme (SAMURI; Fig. 3c) that binds to S-adenosyl-L-homocysteine (SAH) and two Mg^{2+} ions. For this target, trRosettaRNA2 achieved similar accuracy to the best model from AF3-server when no SS prior information was used (8.9 Å and 8.5 Å, respectively). With the SS prior module incorporated, the RMSD of the trRosettaRNA2 model was improved to 6.4 Å. This example illustrates the advantage of leveraging SS prior information, leading to improved results compared with the methods that do not support this, such as AlphaFold 3.

The second example is R1293, which comes from a translation enhancer motif M1293, a complex composed of two protein subunits and one RNA subunit. For this target, while the default trRosettaRNA2 model outperforms AlphaFold 3 (that is, Yang-Multimer versus AF3-server), the performance can be further enhanced when the SPOT-RNA SS is used as input, resulting in an even more accurate model for Yang-Server compared with Yang-Multimer. This example highlights the potential of trRosettaRNA2 to explore diverse SS prompts to yield more accurate predictions.

Note that the above two RNA targets are both involved in complexes with other molecules: R1288 is bound to ligands and R1293 interacts with proteins. The prediction of these entire complex structures falls outside the scope of our method, which is a limitation of our approach compared with AlphaFold 3. However, we observe that incorporating binding partner information does not improve the performance of AlphaFold 3 for these two RNA subunits (Supplementary Fig. 11a). For instance, despite AlphaFold 3 predicting a nearly correct binding site for R1293, the overall bound conformation shows substantial deviations from the experimental structure, with a notable mis-twisting of the junction region (Supplementary Fig. 11b). This reflects that accurate prediction of intermolecular interactions, particularly those involving RNA, remains a challenging problem, possibly owing to a scarcity of relevant training data.

Exploring the conformational landscape of RNase P RNA

RNA structures are typically more flexible and dynamic than proteins, often existing as a conformational ensemble rather than a single static structure[46–48]. However, modelling this inherent dynamic nature remains a challenge for current computational methods. Building on our findings that SS inputs modulate trRosettaRNA2's output, we apply trRosettaRNA2 to predict the heterogeneous structures of the RNase P RNA. This RNA consists of a highly- conserved catalytic domain (C-domain) and a specificity domain (S-domain) known for substantial structural heterogeneity across different species. Elucidating the S-domain's structural variability is crucial for understanding the catalytic activity and mechanism of RNase P. Recently, 158 distinct RNase P RNA conformations were determined via deep learning analysis of AFM images[31]. This experimentally observed ensemble shows substantial diversity, with an average root mean square fluctuation (RMSF) reaching 19.1 Å (Fig. 4a). The variability is substantially more pronounced in the S-domain (average RMSF 31.1 Å) than in the C-domain (average RMSF 11.7 Å), a trend clearly evident in the per-residue RMSF plot (Fig. 4b).

To sample heterogeneous conformations using trRosettaRNA2, we generate predictions using 11 distinct SSs as input. These SSs were obtained from various tools: trRNA2-SS (along with its three models), SPOT-RNA, SPOT-RNA2, EternaFold, MXfold2, RNAfold, RNAstructure and a template SS detected by R2DT. For each SS input, four structures were predicted using the E2E version of trRosettaRNA2 (corresponding to its four internal cycles), while the input MSA was kept fixed. As a comparison, we also run AlphaFold 3 with 20 different random seeds (5 models per seed). In total, we generate 44 trRosettaRNA2 predictions and 100 AlphaFold 3 models for RNase P RNA.

As shown in Fig. 4a, the trRosettaRNA2 predictions show substantial dynamics (overall RMSF = 12.5 Å), particularly within the S-domain (RMSF > 20 Å). Furthermore, its per-residue RMSF profile aligns well with that derived from the experimental conformations (Fig. 4b), achieving a Pearson correlation coefficient (PCC) of 0.894. Notably, trRosettaRNA2 accurately captures the high flexibility observed in the two most dynamic stem loops within the S-domain (residues ~150–179 and ~205–235). In contrast, the AlphaFold 3 predictions are highly homogeneous across both domains, with an overall RMSF of only ~5 Å, and substantially underestimate the conformational fluctuation for almost all residues (Fig. 4b). These results highlight the advantage of trRosettaRNA2 over AlphaFold 3 for predicting conformational ensembles.

This advantage is further underscored by the conformational landscape visualization in Fig. 4c. Despite generating twice as many structures, the predictions of AlphaFold 3 are highly concentrated within a small area (red stars), whereas the predictions of trRosettaRNA2 show greater diversity and spread (blue stars). For instance, for the three structurally distinct conformations highlighted in Fig. 4c, trRosettaRNA2 successfully generates accurate predictions via different SS inputs. As a specific example, for the P58 conformation, the closest AlphaFold 3 prediction remains distant in the landscape, yielding

https://doi.org/10.1038/s42256-026-01223-x

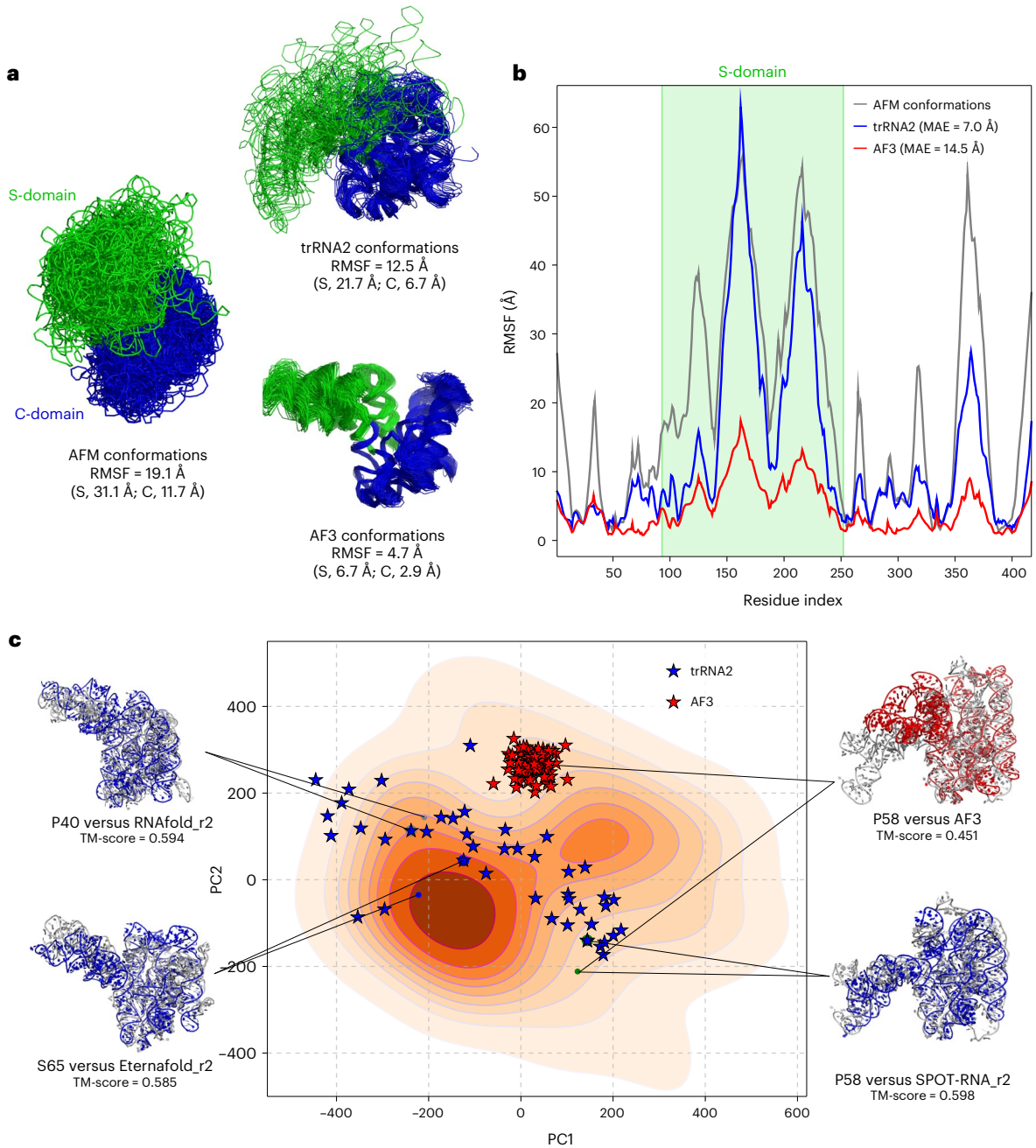

Fig. 4 | Exploring the RNase P RNA conformational landscape. a, Comparison of conformation ensembles determined by AFM versus those predicted by trRosettaRNA2 (trRNA2) and AlphaFold 3. The conformations have been superimposed onto a reference structure (PDB 2A64). Parentheses show average S- and C-domain RMSFs. **b**, Per-residue RMSF plot, with the S-domain shaded in green. **c**, Visualization of the conformational landscape via principal component (PC) analysis based on C1′ atom positions. For each highlighted example, predicted structures (trRNA2, blue; AlphaFold 3, red) are superimposed onto the corresponding experimental conformation (grey). The trRNA2 predictions are named using the format '{SS predictor}_r{recycling times}'.

a TM-score of only 0.451, considerably lower than the trRosettaRNA2 prediction (~0.6). These findings illustrate the potential of trRosettaRNA2 for applications in studying RNA conformational landscapes, a crucial and challenging area in RNA biology.

To elucidate the source of heterogeneity, we quantified the SS dynamics of each ensemble by calculating the standard deviation of base-pairing states (0/1) for each residue pair, and subsequently averaging these values over all pairs (denoted as SD_{SS}). Supplementary Fig. 12 shows that the SS variability of AFM conformations mirrors the 3D landscape, with the S-domain being substantially more dynamic. This trend is roughly captured by the ensemble of 11 predicted SS inputs, whose SD_{SS} correlates weakly with the AFM-derived value (PCC > 0.4). While the network's recycling mechanism alone (1 SS × 4 cycles) offers moderate heterogeneity (PCC ≈ 0.4), using diverse inputs (11 SSs × 1 cycle) provides better dynamic resolution, especially for the S-domain (PCC ≈ 0.5). By integrating input diversity and inference recycling (11 SSs × 4 cycles), trRosettaRNA2 produces an ensemble that closely matches experimental AFM dynamics, especially for the S-domain (SD_{SS}, 1.38 versus 1.60; PCC = 0.7). Conversely, the AlphaFold 3 ensemble shows high homogeneity (SD_{SS} < 0.1). Collectively, these findings provide a mechanistic explanation for the conformational heterogeneity observed in trRosettaRNA2 ensembles.

Discussion

We develop trRosettaRNA2, an automated E2E algorithm for RNA 3D structure prediction. Rigorous benchmark tests, along with blind assessments during CASP16, demonstrate that trRosettaRNA2 outperforms its predecessor, trRosettaRNA, and other representative deep learning-based methods. Notably, trRosettaRNA2 is competitive with AlphaFold 3, despite utilizing considerably fewer parameters and training resources. Crucially, trRosettaRNA2 shows a unique advantage in predicting the heterogeneous conformations of RNase P RNA, probably stemming from its effective use of diverse SS inputs. This capability positions trRosettaRNA2 as a promising tool for investigating RNA conformational landscapes, a fundamental and challenging area within RNA biology.

According to the ablation study, the most substantial contribution stems from the SS module, while the MSA embedding and specialized fine-tuning stages are also beneficial. The SS module provides crucial prior information about base-pairing, which helps mitigate the scarcity of 3D structure data. Furthermore, this module can also be used as an independent, SOTA SS predictor (trRNA2-SS). As a practical application, trRNA2-SS is used to predict the SS for nearly 100,000 RNAcentral sequences lacking known SS, yielding high-confidence predictions across various RNA types, which can provide valuable structural reference for related research (Supplementary Results 1).

The analysis in Supplementary Results 5 demonstrates the robust generalization capabilities of trRosettaRNA2 and trRNA2-SS. However, the challenges in predicting synthetic RNAs highlight the need for further methodological advancements (Supplementary Table 12). Considering the inherent difficulties in modelling RNA flexibility and RNA-protein/ligand complexes, future breakthroughs may come from integrating deep learning methods with experimental or template information. Indeed, our CASP16 report indicates that incorporating 3D homologous template information effectively improves accuracy for synthetic RNAs[38]. Moreover, integrating experimental data (for example, cryogenic electron microscopy (cryo-EM), AFM) is highly promising[30,49–51]. For example, EMRNA[49] achieves a median accuracy of ~2 Å starting from the cryo-EM density maps with high-to-medium resolutions (~2–6 Å). In this direction, trRosettaRNA2 can be coupled with SOTA cryo-EM modelling algorithms, such as CryoAtom[52] from our lab. Specifically, trRosettaRNA2 can provide SS priors and de novo models to facilitate CryoAtom's model building, which in turn generates high-quality structural templates to further enhance trRosettaRNA2's accuracy.

Methods

trRosettaRNA2 algorithm

Different from trRosettaRNA, trRosettaRNA2 uses an E2E architecture with three modules: the SS module, the RNAformer module and the SS-aware structure module. Supplementary Methods 2 provides pseudocodes of the inference procedure and core modules of trRosettaRNA2.

SS prior module. The SS prior module generates a base-pairing probability map from the input MSA (Algorithm 3 in Supplementary Methods 2). The MSA is initially embedded into MSA and pair representations, which are subsequently updated through 8 RNAformer blocks. Finally, a 2D map containing the base-pairing probabilities for all residue pairs is predicted from the updated pair representation. For threshold-dependent evaluations (F1 score and Matthews correlation coefficient (MCC); Supplementary Tables 1 and 2), the probability maps were binarized into 2D base-pairing maps using an unadjusted probability threshold of 0.5.

RNAformer module. The RNAformer module largely inherits from trRosettaRNA, that is, iteratively updating MSA and pair representations and predicting the 2D geometries using a Res2Net[53]-enhanced two-track transformer network (Algorithm 4 in Supplementary Methods 2). The primary distinction lies in the reduced block number: RNAformer in trRosettaRNA2 uses only 12 blocks, down from 48 in trRosettaRNA, as we observed minimal performance impact from this reduction (Supplementary Table 13). This module outputs the updated representations, which are subsequently used to generate the full-atom 3D structure by the structure module (described later). Meanwhile, this module also predicts the 2D geometries. Different from trRosettaRNA, trRosettaRNA2 predicts only inter-residue distances and contacts. Inter-residue angles were excluded owing to their marginal contribution to trRosettaRNA2's accuracy and the substantial increase in computational cost for energy minimization.

SS-aware structure module. The SS-aware structure module generates a 3D structure from the updated representations, assisted by SS information (line 8 in Algorithm 1 in Supplementary Methods 2). This module is primarily driven by invariant point attention, a mechanism adopted in AlphaFold2[54], which updates the single representation using multi-source attention maps. Specifically, starting with the single representation (the first row of the updated MSA representation) and the pair representation, this module iteratively updates the single representation via invariant point attention and then refines the backbone pose (position and orientation, which are initialized by a graph transformer block[55]) of each nucleotide. This iterative process is repeated eight times with shared parameters. Finally, side-chain torsion angles are predicted and used to reconstruct full-atom structures. In trRosettaRNA2, the N1(pyrimidine)/N9(purine)–C1′–C4′ frame defines the backbone of each nucleotide, with remaining atom positions reconstructed from predicted torsion angles.

Considering the unique importance of SS for RNA 3D structure, we attempt to integrate SS into the structure module. Compared with the vanilla softmax operation used in self-attention[56] (equation (1)), the SS-aware structure module adopts an SS-weighted softmax. This involves multiplying the base-pair probability map with the attention maps during softmax (that is, the SS-weighted softmax; equation (2) and Fig. 1c). To prevent over-attenuation of the original attention values, base-pair probabilities are processed using an annealing-smoothing function, with temperature T gradually decreasing across network blocks (Supplementary Fig. 16). Formally, the vanilla attention softmax operation and the SS-weighted softmax operation can be written as:

$$\text{Vanilla attention softmax}: \ a_{ij} = \frac{e^{\mathbf{q}_i \mathbf{k}_j/\sqrt{d}}}{\sum_{l=1}^{L} e^{\mathbf{q}_i \mathbf{k}_l/\sqrt{d}}} \tag{1}$$

$$\text{SS-weighted softmax}: \ a_{ij} = \frac{\left(e^{p_{ij}/T} - 0.99\right) e^{\mathbf{q}_i \mathbf{k}_j/\sqrt{d}}}{\sum_{l=1}^{L} \left(e^{p_{il}/T} - 0.99\right) e^{\mathbf{q}_i \mathbf{k}_l/\sqrt{d}}} \tag{2}$$

where a_{ij} is the attention weight assigned by the ith residue to the jth residue; $\mathbf{q}_i$ and $\mathbf{k}_j$ refer to the query and key vectors of the ith and jth residues, respectively; d represents the dimensionality of the query and key vectors; L is the sequence length; p_{ij} refers to the base-pairing probability between the ith and jth residues; T is the temperature factor of the annealing-smoothing function, which shows a base-2 exponential dependence on the network block count.

Recycling. After each network cycle, the first row of MSA representation, the pair representation and the distance map extracted from the predicted structure are recycled as input for the next cycle (Algorithm 5 in Supplementary Methods 2). The whole network undergoes a total of four iterations, that is, with three recycling steps.

Relaxation. Although the training objective of the E2E network includes structural violation loss (see 'Training procedure'), minor violations can persist in the raw network-generated structures.

To address this issue, we implement a rapid relaxation protocol in PyRosetta. This protocol idealizes bond geometries and reduces steric clashes through a fast MinMover procedure utilizing the 'ref2015_cart' score function within Rosetta. This rapid, yet effective, relaxation step resolves the majority of structural violations, allowing for fast and accurate structure prediction.

Energy minimization. Beyond the E2E prediction, trRosettaRNA2 retains the option for energy minimization using predicted 2D geometries. This procedure is analogous to that used in trRosettaRNA, except that we exclusively utilize inter-residue distance restraints. Orientation restraints were excluded as their incorporation did not improve trRosettaRNA2's performance and substantially increased the minimization time. The specific energy function utilized is as follows:

$$E = E_{\text{dist}} + E_{\text{ros}} \tag{3}$$

where E_{dist} and E_{ros} represent the distance-based restraints and Rosetta's internal energy terms, respectively.

Training data preparation

Training set from PDB. Our primary training dataset was derived from the PDB. We collected all RNA-containing entries released in PDB before January 2024 and isolated the RNA components. The multi-chain entries were systematically split and subsequently reassembled into individual chains according to inter-chain base-pairing interactions. The chains shorter than ten nucleotides were removed. The resulting dataset comprised 10,699 RNA chains, which were used to train both trRosettaRNA2 and trRNA2-SS. To ensure rigorous benchmark evaluations and avoid data leakage, a subset of 8,598 RNAs released before 2022 was designated as the training set for benchmarking. To prevent sampling bias from sequence redundancy, we clustered the training targets to ensure that near-identical sequences (defined as 100% identity and >90% coverage; using CD-HIT-EST[57] with version 4.8.1) were not repeatedly sampled during the initial epochs (see Supplementary Table 14 for details).

Training set from bpRNA. To facilitate transfer learning for SS prediction in trRNA2-SS, we incorporated data from the bpRNA database. In detail, we collected the bpRNA sequences that are also annotated in the Rfam database, allowing for the direct use of established Rfam MSAs. To reduce redundancy, sequences were filtered at an 80% sequence identity threshold. The final bpRNA training set consists of 14,648 RNA chains.

Test data preparation

Test sets for RNA 3D structure prediction. We use two test sets, TS28 and CASP15, to benchmark trRosettaRNA2's performance in RNA 3D structure prediction. TS28 was constructed by collecting RNA-only entries in PDB released after January 2022, splitting and reassembling multi-chain entries according to base-pairing, and removing chains shorter than 20 nucleotides or longer than 400 nucleotides. Redundant sequences within TS28 were removed at an 80% sequence identity cut-off. Finally, RNAs showing substantial homology to either the PDB training set (2022-01 version) or the bpRNA training set were excluded based on BLASTN[25] (version 2.7.1; *E*-value < 10) and CD-HIT-EST (sequence identity >80%), yielding a final dataset of 28 RNAs. The CASP15 set comprised ten RNAs from CASP15 experiments, excluding R1189 and R1190 owing to their complicated protein interactions. Among these ten RNAs, six were naturally occurring and four were synthetic.

Test sets for RNA SS prediction. We evaluated trRNA2-SS using two test sets: ArchiveII[58] and PDB2D. The ArchiveII dataset[58], a standard benchmark, was strictly filtered to exclude RNAs homologous to our training data (BLASTN *E*-value < 10 or CD-HIT-EST identity >80%), yielding 54 RNAs. We further removed sequences containing ambiguous bases ('N') or exceeding 600 nucleotides in length, resulting in a final set of 50 RNAs. Notably, these constraints regarding ambiguous bases and sequence length were imposed solely owing to the limitations of compared methods (for example, UFold), rather than trRNA2-SS itself. Indeed, trRNA2-SS maintains SOTA performance on CASP16 targets exceeding 600 nucleotides (Supplementary Table 15).

Nevertheless, the dataset ArchiveII has limitations: its structures are derived from comparative sequence analysis rather than experimental determination; and potential overlap with the training sets of other methods may compromise the reliability of the benchmark. To complement this, we constructed the PDB2D set, comprising 19 RNAs from the TS28 dataset (excluding sequences with undefined 'N' bases). These structures were experimentally validated (PDB-derived) and were released after January 2022, later than the training cut-offs of all compared methods, thereby ensuring a strict temporal split to prevent data leakage.

For enhanced clarity, a comprehensive summary of all datasets utilized in this study is provided in Supplementary Table 16.

Training procedure

Pre-training of SS prior module. The SS prior module (that is, trRNA2-SS) was pre-trained in a transfer learning manner: initial pre-training on the bpRNA dataset followed by fine-tuning on the PDB dataset. This strategy is motivated by two factors. First, the vast scale of the bpRNA dataset (>100,000 entries) compensates for the limited number of RNA 3D structures in PDB (~9,000 entries), providing a solid foundation for supervised pre-training. Second, while extensive, bpRNA relies partially on predicted labels and overlooks higher-order base pairs (base multiplets) that are critical for stabilizing 3D structure. Therefore, fine-tuning on PDB-derived data is indispensable to learn the higher-order interaction information preserved in these experimentally determined structures. Ablation studies in Supplementary Table 4 confirm that both stages of this training process are effective in boosting performance.

For the bpRNA samples, the training labels are the base pairs extracted from the dot-bracket notations given by the bpRNA database. For the PDB samples, the labels are the base pairs extracted from the PDB files using DSSR[59]. The loss function is defined as the cross-entropy between the predicted base-pairing probability and the one-hot encoding of true labels. The training was conducted on an A100 40GB GPU, using an Adam optimizer with a learning rate of 0.0002. We trained three models with identical configurations. The final trRNA2-SS prediction is the ensemble (average) of the predictions from these three models.

Training of trRosettaRNA2. The training of trRosettaRNA2 consists of three stages (Supplementary Table 14). In the first stage, trRosettaRNA2 is primarily trained to produce structures with high accuracy. The loss items include the frame-aligned-position-error (FAPE) loss $\mathcal{L}_{\text{FAPE}}$, the 2D geometries loss $\mathcal{L}_{\text{geo2D}}$, the torsion loss $\mathcal{L}_{\text{torsion}}$, the predicted local distance difference test (pLDDT) loss $\mathcal{L}_{\text{pLDDT}}$, and an extra loss item $\mathcal{L}_{\text{BP}}$ to optimize the inter-atom distances of the predicted structure according to the native base-pairing information. Specifically, for each native base pair, $\mathcal{L}_{\text{BP}}$ forces all the inter-atom distances between this pair of nucleotides in the predicted structure to fall within the range of [1.5 Å, d + 3 Å], where d is the corresponding inter-atom distance in the experimental structure. In total, the loss function used in the first stage can be written as:

$$\mathcal{L}_{\text{init}} = \mathcal{L}_{\text{FAPE}} + 0.5\mathcal{L}_{\text{geo2D}} + \mathcal{L}_{\text{torsion}} + 0.05\mathcal{L}_{\text{pLDDT}} + 0.05\mathcal{L}_{\text{BP}} \tag{4}$$

The second stage is the violation fine-tuning, which fine-tunes the model to improve the physicochemical plausibility of predicted

structures. Specifically, this stage introduces two new loss items. The first item, $\mathcal{L}_{\text{bond}}$, enforces the O3′–P bond length and P–P distance, as well as the O3′–P–O5′ angle between neighbouring residues, to be within ideal ranges, which were statistically derived from the RNAs in the training set. The second is the clash loss $\mathcal{L}_{\text{clash}}$, which penalizes steric clashes in the predicted structures. In detail, $\mathcal{L}_{\text{clash}}$ penalizes the inter-atom distances shorter than $\min(d - 0.1\,\text{Å}, 2\,\text{Å})$, where d is the corresponding distance in the experimental structure. Furthermore, the weights for $\mathcal{L}_{\text{pLDDT}}$ and $\mathcal{L}_{\text{BP}}$ are increased to 0.1 and 0.5 in this stage, respectively. In total, the loss function used in the second stage can be written as:

$$\begin{aligned}\mathcal{L}_{\text{vioFT}} = \mathcal{L}_{\text{FAPE}} + 0.5\mathcal{L}_{\text{geo2D}} + \mathcal{L}_{\text{torsion}} \\ +0.1\mathcal{L}_{\text{pLDDT}} + 0.5\mathcal{L}_{\text{BP}} + 0.3\mathcal{L}_{\text{bond}} + \mathcal{L}_{\text{clash}}\end{aligned} \tag{5}$$

For the first two stages, the training RNAs are clustered based on their sequence identities. During each training epoch, we iterate through all the clusters, sampling one RNA from the current cluster at each step.

The third stage, termed resampled fine-tuning, further fine-tunes the model to improve performance on challenging targets. This is achieved by using a dynamic training sampler that increases the sampling weights for targets exhibiting low lDDT scores. Specifically, each training epoch involves sampling 8,000 training examples, where the sampling weight for each RNA is defined as the reciprocal of its lDDT value in the previous epoch. The loss function of this stage is the same as the second stage, that is

$$\begin{aligned}\mathcal{L}_{\text{resFT}} = \mathcal{L}_{\text{FAPE}} + 0.5\mathcal{L}_{\text{geo2D}} + \mathcal{L}_{\text{torsion}} \\ +0.1\mathcal{L}_{\text{pLDDT}} + 0.5\mathcal{L}_{\text{BP}} + 0.3\mathcal{L}_{\text{bond}} + \mathcal{L}_{\text{clash}}\end{aligned} \tag{6}$$

A detailed description of the specific loss terms is provided in Supplementary Methods 3. To address GPU memory limitations and augment the training sample, we perform random subsampling during training. Specifically, for targets longer than the maximum crop size (256 or 384), a subsequence is randomly cropped for training. In addition, the MSA is randomly subsampled to a depth of fewer than 200 sequences per target. These strategies ensure that the same RNA is presented as distinct training samples across different epochs. More training details, including learning rates, crop size and training times, are shown in Supplementary Table 14.

Blind tests in CASP16

We participated in the CASP16 RNA structure prediction experiment via two automated server groups: Yang-Server (TS052) and Yang-Multimer (TS456). Yang-Multimer was configured to perform a direct blind benchmark against other SOTA methods (using standard inputs), whereas Yang-Server was optimized to achieve maximal performance by systematically exploring diverse SS inputs.

Yang-Server submission. We prepared multiple input SSs for each target during CASP16. These SSs included predictions from various algorithms (trRNA2-SS, SPOT-RNA and EternaFold) as well as those derived from available templates (detected by RNAthreader[60] and R2DT[61]). In addition, we trained a variant of trRosettaRNA2 that does not use the SS prior module, providing an SS-free prediction. For several targets (R1221s2, R1224s2, R1281, R1289, R1291) with identifiable 3D templates (using RNAthreader), we also performed template-based modelling to expand the decoy set. Finally, for each target, we used an in-house RNA quality assessment method[62] to select the top-five models from the decoy set for submission as Yang-Server. Considering that trRosettaRNA2 was trained with a maximum crop size of 384 nucleotides owing to the limited computational resources, its applicability to larger RNAs (>400 nucleotides) may be hindered. Therefore, for the large RNA targets without available templates in CASP16 (R1241, R1248, R1250-R1254, R1283, R1285, R1286, R1290), we chose to submit the AlphaFold 3 predictions, as its training utilized a larger crop size of 768.

Yang-Multimer submission. To conduct a more direct blind benchmark test of trRosettaRNA2, we simultaneously submitted multiple predictions through another server group, Yang-Multimer (TS456). Specifically, we submitted the predictions from trRosettaRNA2, trRosettaRNA, RoseTTAFoldNA, DeepFoldRNA and RhoFold as models 1–5, respectively, within the Yang-Multimer group. All five methods used the same inputs (MSA and/or SS).

Experiment set-up

RNA 3D structure prediction. We benchmark trRosettaRNA2 against its predecessor, trRosettaRNA, and other representative deep learning methods: AlphaFold 3, RhoFold+, DeepFoldRNA and RoseTTAFoldNA. All methods utilize identical MSAs as input. For methods using SS input (trRosettaRNA2, trRosettaRNA and DeepFoldRNA), the specific SS sources are detailed in Table 1 and Fig. 2a. AlphaFold 3 was run using its standalone package with five seeds (0–4), generating 5 models per seed. The final AlphaFold 3 prediction was selected as the model with the highest ranking score (predicted template modeling score, or pTM) among the 25 generated structures. Owing to observed relaxation issues with RhoFold+ for several long RNAs (listed in Supplementary Table 17), unrelaxed predictions were evaluated in those instances for RhoFold+. All other methods were executed using their publicly available standalone packages with default settings. The 3D structures are presented using PyMOL (version 2.5.2).

RNA SS prediction. To assess the performance of RNA SS prediction, we benchmark trRNA2-SS against a diverse set of representative methods. This set included traditional energy- and/or evolution-based approaches (EternaFold, RNAfold, RNAstructure, PETfold) as well as advanced deep learning-based methods (RNA-FM, RNA-MSM, UFold, MXfold2, SPOT-RNA, SPOT-RNA2, KnotFold, BPfold). For methods that utilize MSAs as input (trRNA2-SS, PETfold, SPOT-RNA2 and RNA-MSM), we ensure consistency by providing them with the same MSA inputs. For deep learning-based methods that use transfer learning, that is, those pre-trained on SS datasets and fine-tuned on PDB data (including SPOT-RNA, SPOT-RNA2 and UFold), we evaluate their PDB-fine-tuned versions. The SSs are visualized using forna[63] (version 1.0).

Evaluation. To comprehensively assess the performance of trRosettaRNA2, we evaluated its predicted 3D structures using multiple metrics. Among them, RMSD and TM-score[64,65] focus on the global topology; the lDDT[66] to assess local deviations; and INF[67] to measure the accuracy of base-pairing interactions. Unless otherwise specified, the RMSD (based on all atoms) and INF are calculated using the RNA_assessment toolkit[68]; TM-score is calculated using the RNAalign program[65]; lDDT is calculated using the following equation:

$$\text{lDDT} = \frac{1}{4L}\sum_{\varepsilon\in\{1,2,4,6\}}\sum_{i=1}^{L}\frac{\sum_{j:D_{ij}<30} I\left(\left|d_{ij} - D_{ij}\right| < \varepsilon\right)}{\sum_{j:D_{ij}<30} 1} \tag{7}$$

where L is the sequence length; d_{ij} and D_{ij} are the C1′ distances between the ith and jth residues in the predicted and experimental structures, respectively; and I is the indicator function. To accommodate the typically larger scale of RNA inter-residue distances compared with proteins, we modified the standard lDDT calculation in two aspects. First, the local neighbourhood radius used for comparisons increased from 15 Å to 30 Å. Second, the set of distance error thresholds was changed from {0.5 Å, 1 Å, 2 Å, 4 Å} to {1 Å, 2 Å, 4 Å, 6 Å}. Compared with the original formulation, this modified lDDT exhibits a stronger correlation with RMSD (PCC, −0.55 versus −0.77; Supplementary Fig. 17).

For the evaluation of SS prediction, we primarily use the area under the precision–recall curve, which can provide a direct evaluation of the raw output (that is, the base-pairing probabilities) generated by various methods.

Application of trRNA2-SS to RNAcentral sequences. We apply trRNA2-SS to perform high-throughput SS predictions for RNAcentral sequences that lack known SSs. Specifically, we downloaded a list of RNAcentral entries meeting the following criteria: available genomic mapping, no known SS, and sequence length between 20 and 300 nucleotides. Using this list, we collected the corresponding sequences from our localized RNAcentral database, yielding 588,227 sequences. Redundant sequences were then removed using the CD-HIT-EST program with a sequence identity cut-off of 80%. The sequences with a substantial similarity to training sets (BLASTN *E*-value < 10 or CD-HIT-EST sequence identity >80%) were also removed. The final set contains 92,370 non-redundant RNA sequences without known SSs. We then used trRNA2-SS to predict the SSs for these sequences. Binary base pairs were derived from the predicted probabilities using a simple cut-off of 0.5.

To facilitate the evaluation of the SS predicted by trRNA2-SS for these RNAcentral sequences, we trained a confidence score (C-Score$_{ss}$) on 1,000 RNAs randomly sampled from the training set. This score correlates well with the actual F1-score on the test sets (PCC = 0.83; Supplementary Fig. 18). Specifically, C-Score$_{ss}$ can be written as:

$$\text{C-Score}_{ss} = 0.271\text{prob}_{top} + 0.906\text{prob}_{>0.2} + 0.01\text{std}_{>0.2} - 0.22 \quad (8)$$

where prob$_{top}$ is the average probability of the top $L/10$ predicted base pairs (L is sequence length) that have a sequence separation greater than 12; prob$_{>0.2}$ is the average probability of the pairs with probabilities >0.2; std$_{>0.2}$ is the average standard deviation of the probabilities predicted by the 3 individual trRNA2-SS models, calculated only for those base pairs where the final predicted probability is greater than 0.2.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the benchmark datasets (ArchiveII, PDB2D, TS28, CASP15, and CASP16) are available at https://yanglab.qd.sdu.edu.cn/trRosettaRNA/benchmark/.

Code availability

The web server is available at https://yanglab.qd.sdu.edu.cn/trRosettaRNA/. The source code is available via GitHub at https://github.com/YangLab-SDU/trRosettaRNA2/ and via Zenodo at https://doi.org/10.5281/zenodo.18873019 (ref. 69).

References

1. Zhang, J., Fei, Y., Sun, L. & Zhang, Q. C. Advances and opportunities in RNA structure experimental determination and computational modeling. *Nat. Methods* **19**, 1193–1207 (2022).
2. Wang, W., Su, B., Peng, Z. & Yang, J. Integrated experimental and AI innovations for RNA structure determination. *Nat. Biotechnol.* **44**, 205–214 (2026).
3. Kwon, D. RNA function follows form—why is it so hard to predict? *Nature* **639**, 1106–1108 (2025).
4. Sharma, S., Ding, F. & Dokholyan, N. V. iFoldRNA: three-dimensional RNA structure prediction and folding. *Bioinformatics* **24**, 1951–1952 (2008).
5. Das, R. & Baker, D. Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl Acad. Sci. USA* **104**, 14664–14669 (2007).
6. Das, R., Karanicolas, J. & Baker, D. Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **7**, 291–294 (2010).
7. Boniecki, M. J. et al. SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.* **44**, e63 (2016).
8. Popenda, M. et al. Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* **40**, e112 (2012).
9. Zhao, Y. et al. Automated and fast building of three-dimensional RNA structures. *Sci. Rep.* **2**, 734 (2012).
10. Zhang, Y., Wang, J. & Xiao, Y. 3dRNA: 3D structure prediction from linear to circular RNAs. *J. Mol. Biol.* **434**, 167452 (2022).
11. Wang, W. et al. trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat. Commun.* **14**, 7266 (2023).
12. Pearce, R., Omenn, G. S. & Zhang, Y. De novo RNA tertiary structure prediction at atomic resolution using geometric potentials from deep learning. Preprint at *bioRxiv* https://doi.org/10.1101/2022.05.15.491755 (2022).
13. Li, Y. et al. Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nat. Commun.* **14**, 5745 (2023).
14. Shen, T. et al. Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nat. Methods* **21**, 2287–2298 (2024).
15. Kagaya, Y. et al. NuFold: end-to-end approach for RNA tertiary structure prediction with flexible nucleobase center representation. *Nat. Commun.* **16**, 881 (2025).
16. Baek, M. et al. Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods* **21**, 117–121 (2024).
17. Krishna, R. et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **384**, eadl2528 (2024).
18. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
19. Justyna, M., Zirbel, C., Antczak, M. & Szachniuk, M. Graph neural network and diffusion model for modeling RNA interatomic interactions. *Bioinformatics* **41**, btaf515 (2025).
20. Cruz, J. A. et al. RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *RNA* **18**, 610–625 (2012).
21. Miao, Z. et al. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA* **23**, 655–672 (2017).
22. Das, R. et al. Assessment of three-dimensional RNA structure prediction in CASP15. *Proteins Struct. Funct. Bioinf.* **91**, 1747–1770 (2023).
23. Schneider, B. et al. When will RNA get its AlphaFold moment? *Nucleic Acids Res.* **51**, 9522–9532 (2023).
24. Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
25. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinform.* **10**, 421 (2009).
26. Zhang, C. et al. The historical evolution and significance of multiple sequence alignment in molecular structure and function prediction. *Biomolecules* **14**, 1531 (2024).
27. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935 (2013).
28. Zhang, T. et al. RNAcmap: a fully automatic pipeline for predicting contact maps of RNAs by evolutionary coupling analysis. *Bioinformatics* **37**, 3494–3500 (2021).
29. 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).
30. Degenhardt, M. F. S. et al. Determining structures of RNA conformers using AFM and deep neural networks. *Nature* **637**, 1234–1243 (2025).
31. Lee, Y.-T. et al. The conformational space of RNase P RNA in solution. *Nature* **637**, 1244–1251 (2025).

32. Tinoco, I. & Bustamante, C. How RNA folds. *J. Mol. Biol.* **293**, 271–281 (1999).
33. Brion, P. & Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys.* **26**, 113–137 (1997).
34. Herschlag, D. RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**, 20871–20874 (1995).
35. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).
36. Danaee, P. et al. bpRNA: large-scale automated annotation and analysis of RNA secondary structure. *Nucleic Acids Res.* **46**, 5381–5394 (2018).
37. Kretsch, R. C. et al. Assessment of nucleic acid structure prediction in CASP16. *Proteins Struct. Funct. Bioinform.* **94**, 192–217 (2026).
38. Wang, W., Luo, Y., Peng, Z. & Yang, J. Accurate biomolecular structure prediction in CASP16 with optimized inputs to state-of-the-art predictors. *Proteins Struct. Funct. Bioinform.* **94**, 142–153 (2026).
39. Chaudhury, S., Lyskov, S. & Gray, J. J. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **26**, 689–691 (2010).
40. Dabrowski-Tumanski, P., Rubach, P., Niemyska, W., Gren, B. A. & Sulkowska, J. I. Topoly: Python package to analyze topology of polymers. *Brief. Bioinform.* **22**, bbaa196 (2021).
41. Gren, B. A., Antczak, M., Zok, T., Sulkowska, J. I. & Szachniuk, M. Knotted artifacts in predicted 3D RNA structures. *PLoS Comput. Biol.* **20**, e1011959 (2024).
42. Poblete, S., Mlynarczyk, M. & Szachniuk, M. Unknotting RNA: a method to resolve computational artifacts. *PLoS Comput. Biol.* **21**, e1012843 (2025).
43. Luwanski, K. et al. RNAspider: a webserver to analyze entanglements in RNA 3D structures. *Nucleic Acids Res.* **50**, W663–W669 (2022).
44. Li, Y. et al. DRfold2 is a deep learning-based tool that enables efficient and accurate RNA structure prediction. *PLoS Biol.* **24**, e3003659 (2026).
45. Tarafder, S. & Bhattacharya, D. RNAbpFlow: base pair-augmented SE(3)-flow matching for conditional RNA 3D structure generation. Preprint at *bioRxiv* https://doi.org/10.1101/2025.01.24.634669 (2025).
46. Cruz, J. A. & Westhof, E. The dynamic landscapes of RNA architecture. *Cell* **136**, 604–609 (2009).
47. Vicens, Q. & Kieft, J. S. Thoughts on how to think (and talk) about RNA structure. *Proc. Natl Acad. Sci. USA* **119**, e2112677119 (2022).
48. Ganser, L. R., Kelly, M. L., Herschlag, D. & Al-Hashimi, H. M. The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **20**, 474–489 (2019).
49. Li, T. et al. All-atom RNA structure determination from cryo-EM maps. *Nat. Biotechnol.* **43**, 97–105 (2025).
50. Li, T., Cao, H., He, J. & Huang, S.-Y. Automated detection and de novo structure modeling of nucleic acids from cryo-EM maps. *Nat. Commun.* **15**, 9367 (2024).
51. Jamali, K. et al. Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450–457 (2024).
52. Su, B., Huang, K., Peng, Z., Amunts, A. & Yang, J. CryoAtom improves model building for cryo-EM. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-025-01713-3 (2025).
53. Gao, S.-H. et al. Res2Net: a new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 652–662 (2021).
54. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
55. Shi, Y. et al. Masked label prediction: unified message passing model for semi-supervised classification. In *Proc. Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21* (ed. Zhou, Z.-H.) 1548–1554 (International Joint Conferences on Artificial Intelligence Organization, 2021).
56. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30**, 5998–6008 (2017).
57. Li, W. & Godzik, A. CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
58. Sloma, M. F. & Mathews, D. H. Exact calculation of loop formation probability identifies folding motifs in RNA secondary structures. *RNA* **22**, 1808–1818 (2016).
59. Lu, X.-J., Bussemaker, H. J. & Olson, W. K. DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.* **43**, e142 (2015).
60. Du, Z., Peng, Z. & Yang, J. RNA threading with secondary structure and sequence profile. *Bioinformatics* **40**, btae080 (2024).
61. Sweeney, B. A. et al. R2DT is a framework for predicting and visualising RNA secondary structure using templates. *Nat. Commun.* **12**, 3494 (2021).
62. Liu, X. et al. Quality assessment of RNA 3D structure models using deep learning and intermediate 2D maps. *Commun. Biol.* **9**, 293 (2026).
63. Kerpedjiev, P., Hammer, S. & Hofacker, I. L. Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. *Bioinformatics* **31**, 3377–3379 (2015).
64. Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins Struct. Funct. Bioinform.* **57**, 702–710 (2004).
65. Gong, S., Zhang, C. & Zhang, Y. RNA-align: quick and accurate alignment of RNA 3D structures based on size-independent TM-scoreRNA. *Bioinformatics* **35**, 4459–4461 (2019).
66. Mariani, V., Biasini, M., Barbato, A. & Schwede, T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728 (2013).
67. Parisien, M., Cruz, J. A., Westhof, É & Major, F. New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA* **15**, 1875–1885 (2009).
68. Bu, F. et al. RNA-Puzzles round V: blind predictions of 23 RNA structures. *Nat. Methods* **22**, 399–411 (2025).
69. Wang, W., Peng, Z. & Yang, J. Source code for trRosettaRNA2. *Zenodo* https://doi.org/10.5281/zenodo.18873019 (2026).

Acknowledgements

This work is supported by the following funding sources: National Natural Science Foundation of China (NSFC T2225007 and 32430063 to J.Y., T2222012 to Z.P., 62501364 to W.W.), Postdoctoral Fellowship Program and China Postdoctoral Science Foundation (BX20240212 and 2025M783122 to W.W.), and Fundamental Research Funds for the Central Universities.

Author contributions

J.Y. conceptualized and administered the study. W.W. designed and implemented the network. W.W. and J.Y. implemented energy minimization. W.W., Z.P. and J.Y. prepared the training data. All authors wrote and revised the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42256-026-01223-x.

Correspondence and requests for materials should be addressed to Jianyi Yang.

Peer review information *Nature Machine Intelligence* thanks Walter N. Moss, Marta Szachniuk and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Jianyi Yang

Last updated by author(s): Mar 11, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	CD-HIT (version 4.8.1) was used to remove redundant sequences in RNAcentral sequences. BLAST+ (version 2.7.1) was used to remove the tested RNAs redundant to the training sets, as well as generate MSA. Infernal (version 1.1.4) was also used to generate MSA. The source code of trRosettaRNA2 are available at Github (https://github.com/YangLab-SDU/trRosettaRNA2).
Data analysis	PyRosetta4 (version 2019.23) was used to perform energy minimization. RNA-Puzzles toolkit (version 1.0) and RNAalign (version 1.0) were used to evaluate the predicted structures. PyMOL (version 2.5.2) was used to display the 3D structures. forna (version 1.0) was used to visualize RNA secondary structures.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All the benchmark datasets (ArchiveII, PDB2D, TS28, CASP15, and CASP16) are available at: https://yanglab.qd.sdu.edu.cn/trRosettaRNA/benchmark/.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical mmethods were used to predetermine sample sizes. The number of datasets was determined by the availability of high-quality, non-redundant RNA structures that ensure a rigorous evauation.Speciafically, the TS28 set (n=28) was constructed from all available RNA-only entries released after 2022-01 in PDB to ensure a strict temporal separation from the training data. The CASP15 set (n=10) was included as it represents the community-standard blind test for SOTA methods. For secondary structure prediction, 50 RNAs from ArchiveII and 19 from TS28 (filtered for excluding "N" residues) were selected. Both datasets were filtered to exclude sequences similar to the training sets. These datasets are sufficient as they cover a diverse range of RNA classes and represent the most recent and chllengig targets, providing a statistically representative and unbiased benchmark for current structure prediction methods.
Data exclusions	The 2 CASP15 targets featured as their complex interaction with over 4 protein chains were excluded, as our methods are mainly for RNA-only modeling.
Replication	We have run the programs 264 (=(50+28+10)*3) times with each RNA running three replications. All replications were successful. Please follow our instructions on the web server page, or instructions in the standalone package.
Randomization	The datasets are divided into two groups according to their label types (i.e., secondary structure or 3D structure).
Blinding	There was no blinding group or analysis in this manuscript as all RNAs in our test have the deposited models in the public databases.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A