Synthetic RNA Evolution Enables Accurate Alignment-Free 3D Structure Prediction

Jörg K.H. Franke[1,4,5,*] **Iris K. Tennie**[2,*] **Michael Uhl**[1] **Ryan Köksal**[1]
Dominika Matus[1] **Frank Hutter**[3,4,1] **Rolf Backofen**[1] **Frederic Runge**[1]

[1]University of Freiburg, Germany [2]FernUniversität in Hagen, Germany
[3]Prior Labs, Germany [4]ELLIS Institute Tübingen, Germany [5] LAION, Germany

Correspondence to runget@cs.uni-freiburg.de

Abstract

Accurate RNA 3D structure prediction remains a bottleneck in computational biology. Although protein structures can now be predicted with near-experimental fidelity, RNA 3D prediction still lags behind. One reason is that state-of-the-art tools such as AlphaFold 3 require deep multiple sequence alignments (MSAs) that are considerably more difficult to obtain for RNA than for proteins. Here, we remove this bottleneck by synthesizing homologous RNAs. Starting from a single RNA sequence, our novel deep learning model RNAformer predicts secondary structure with high fidelity, which we use to generate structurally consistent synthetic homologs through lightweight, evolution-inspired mutation rules. This process produces deep, MSA-like sequence ensembles in seconds, without reliance on natural sequence databases while bypassing alignment steps. When supplied to AlphaFold 3, synthetic homologs substantially improve local RNA structural accuracy and rescue predictions for orphan RNAs where no natural alignment exists. Secondary structure-driven synthetic evolution, therefore, unlocks deep alignment benefits for RNA 3D structure prediction.

1 Introduction

RNA molecules play central roles in gene regulation, epigenetics, and cellular control, with function tightly linked to three-dimensional (3D) structure (Morris & Mattick, 2014; Statello et al., 2021; Mangiavacchi et al., 2023; Oksuz et al., 2023). Despite major advances in protein structure prediction (Jumper et al., 2021), accurate RNA 3D modeling remains challenging, particularly for orphan RNAs that lack detectable homologs (Schneider et al., 2023; Das et al., 2023; Bernard et al., 2024; Kwon, 2025). A primary reason is the dependence of modern predictors on deep multiple sequence alignments (MSAs), which are often unavailable or shallow for RNA (Schneider et al., 2023; Szikszai et al., 2025). Covariation signals extracted from MSAs encode conserved structural constraints and are a cornerstone of state-of-the-art 3D structure prediction. However, RNA homologs are sparse, biased, and slow to retrieve, leaving most RNAs without reliable alignment-derived signal (Schneider et al., 2023; Szikszai et al., 2025). This limitation persists even as structure prediction models themselves continue to improve, making the availability of MSAs a fundamental bottleneck rather than a modeling issue.

Here, we propose a conceptual inversion of the traditional sequence-to-structure pipeline. Instead of inferring structure from evolutionary covariation, we generate evolutionarily plausible sequence variation directly from structural constraints. RNA folding provides a natural basis for this inversion: secondary structure forms early (Tinoco Jr & Bustamante, 1999), is strongly conserved, and tightly constrains allowable sequence evolution. Therefore, we developed a fast and accurate deep learning model to predict secondary structure with high fidelity, the RNAformer. Treating predicted secondary structure as a prior on mutation,

*Equal Contribution.

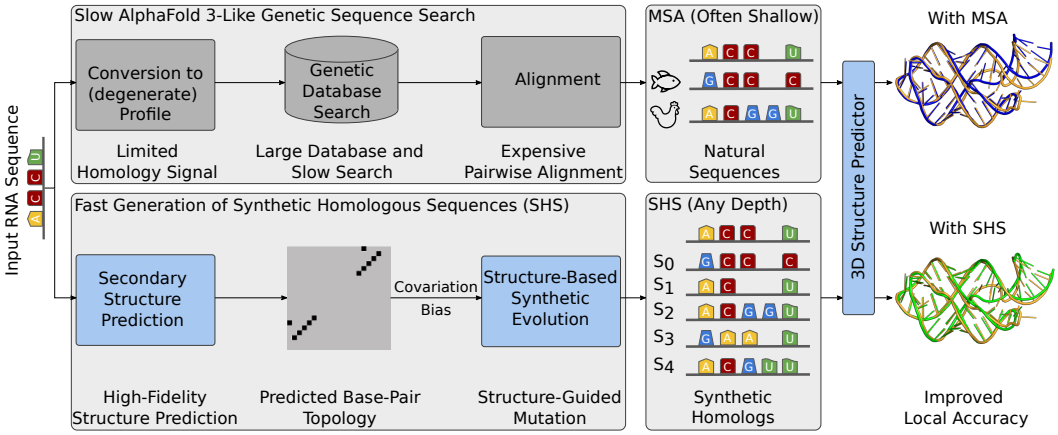

Figure 1: **Comparison of natural and synthetic homology pipelines for RNA structure prediction.** AlphaFold-3-like pipeline (top) using a (degenerate) profile to search a large genetic sequence database and subsequent expensive HMM-based pairwise alignments. The resulting multiple sequence alignment (MSA) of natural homologs is often shallow. Our synthetic homologous sequence (SHS) pipeline (bottom) replaces database search with fast, structure-guided synthetic evolution. Starting from a predicted RNA secondary structure, SHS generates alignment-like ensembles of synthetic homologs of arbitrary depth within seconds, requires no natural sequences, and entirely bypasses pairwise alignment, leading to improved local structural accuracy in downstream prediction. Prediction: 3FO4.

insertion, and deletion, we can synthesize covariation patterns consistent with RNA folding physics. Building on this idea, we introduce secondary structure-guided *synthetic homologous sequences* (SHS; see Figure 1), a fast and scalable protocol that inflates a single RNA sequence into deep, MSA-like sequence ensembles without querying natural databases or performing alignment. SHS applies structure-preserving mutations at paired positions while allowing greater flexibility in unpaired regions, reflecting established evolutionary patterns in natural RNAs. The resulting synthetic ensembles encode realistic covariation signals and can be generated in seconds for arbitrary RNA sequences.

We demonstrate that SHS provides a drop-in replacement for natural MSAs in downstream structure prediction. When supplied to AlphaFold 3 (AF3) (Abramson et al., 2024), synthetic homologs substantially improve RNA 3D local accuracy and consistently rescue predictions for orphan RNAs. These gains are achieved without modifying the predictor itself, highlighting that AlphaFold 3's RNA structure prediction remains strongly dependent on secondary structure topology and the covariation it induces. Our results establish secondary structure-driven synthetic evolution as a conceptually distinct alternative to natural MSAs.

We summarize our main contributions as follows:

- We introduce synthetic homologous sequences (SHS), a fast and scalable alternative to natural MSAs that requires neither database search nor sequence alignment (Section 2).
- We demonstrate that SHS substantially improves AlphaFold 3 RNA 3D local accuracy across a broad benchmark, including RNA-protein complexes (Section 3.1).
- We show that SHS consistently rescues RNA structure prediction in cases where no natural homologs are available (Section 3.2).
- We provide evidence that AlphaFold 3 predictions remain strongly dependent on secondary structure topology and the covariation signals it induces (Section 3.3).
- We validate that SHS introduces biologically meaningful covariation patterns using independent case studies with thermodynamics-based models (Section 3.4).

We discuss related work in Appendix A and limitations in Appendix B.

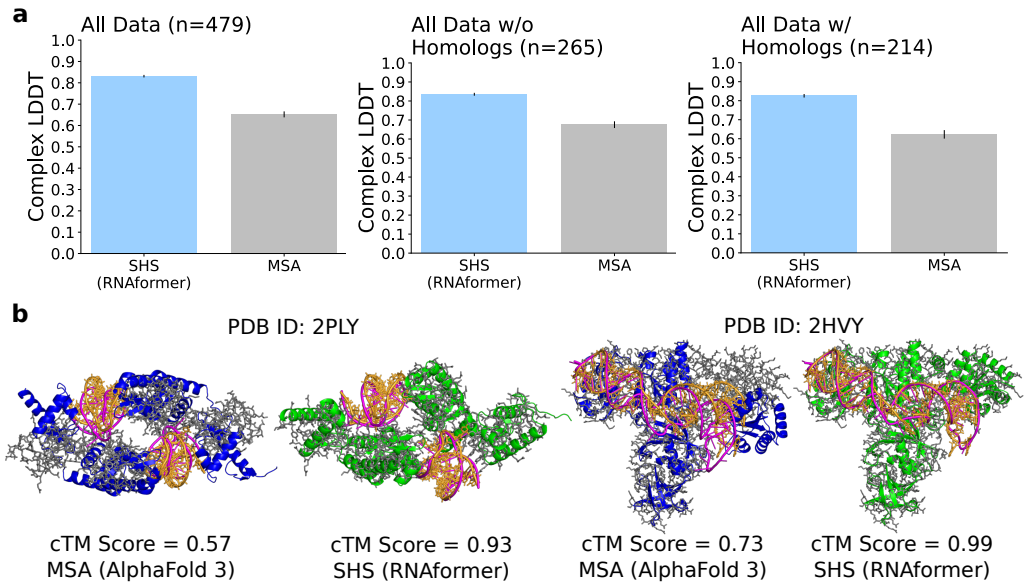

Figure 2: **Synthetic homologs improve local structural accuracy a,** Comparison of AF3 with and without RNAformer-guided SHS across all RNAs and subsets with/without natural MSAs. SHS consistently achieves higher mean LDDT values compared to vanila AF3. **b,** Example 3D predictions with improved local RNA structure via RNAformer-guided SHS for complexes with two RNA chains (2PLY; orphan) and one RNA chain (2HVY; non-orphan). Error bars show the s.e.m.

2 Secondary Structure Based Synthetic Evolution

We aim at generating synthetic homologous sequences (SHS) by inflating a single RNA input sequence into a set of synthetic homologs for downstream structural modeling. Building on the observation that covariation in natural MSAs concentrates in paired positions whereas unpaired regions mutate more freely (Holmes, 2004), we leverage predicted nucleotide interactions to partition an input RNA sequence into paired and unpaired regions. To do so, our novel SHS pipeline requires two ingredients: High-fidelity predictions of nucleotide interactions, and fast and controllable sequence generation to support subsequent predictions.

2.1 Accurate Nucleotide Interaction Prediction

We develop RNAformer, a secondary structure predictor based on axial-attention (Ho et al., 2019) trained on carefully created non-homologous data splits (see Appendix C and D for details of the data pipeline and architecture), and fine-tune two versions using experimentally validated data from the Protein Data Bank (PDB) (Burley et al., 2023): One version for benchmarking against eleven state-of-the-art secondary structure predictors (see Appendix F.1), and one version to compare against secondary structures obtained from AlphaFold 3 using DSSR (Lu et al., 2015), when training with the same cutoff-date as AlphaFold 3. On a commonly used secondary structure prediction benchmark (originally provided by Singh et al. (2021); see Appendix C), we find that in both settings, RNAformer clearly achieves the best performance (see Table 7).

2.2 Fast Secondary Structure-Guided Homolog-Like Ensemble Generation

To obtain alignment-like information for downstream RNA predictions, particularly in cases where natural MSAs are shallow or absent, we generate synthetic homologous sequences (SHS) directly from predicted nucleotide interaction maps (see Figure 1): Given an input RNA sequence $\mathbf{s} = (s_1, \ldots, s_l)$ of length l, and a nucleotide interaction map P that partitions positions into paired and unpaired sites, we apply a single-pass stochastic mutation

Table 1: **RNAformer-guided synthetic homologs improve 3D predictions for orphan RNAs.** We present mean results with standard deviation on single-chain RNA structure prediction without available natural homologs (left) and with existing MSA (right). RNAformer generated SHS clearly improves the predictions for RNA structures without available natural homologs. Blue, best performance; gray, equal best performance.

Model	Single-chain RNA w/o Homologs (n=20)			Single-chain RNA w/ Homologs (n=71)		
	RMSD (↓)	LDDT (↑)	TM (↑)	RMSD (↓)	LDDT (↑)	TM (↑)
AlphaFold 3 (MSA Search)	4.723 ±5.549	0.755 ±0.219	0.480 ±0.250	3.177 ±4.376	0.825 ±0.122	0.723 ±0.227
AlphaFold 3 (SHS; ours)	4.593 ±5.518	0.770 ±0.217	0.480 ±0.227	4.203 ±5.475	0.779 ±0.114	0.647 ±0.206

protocol that treats these two classes differently. Unpaired nucleotides are subject to local substitutions, insertions, and deletions that introduce loop variability, whereas paired nucleotides are mutated in a coordinated, structure-preserving manner: with high probability the original base pair is retained, and otherwise it is replaced by a compensatory pair drawn from a curated mutation dictionary. Insertions and deletions are length-limited and scale with l, allowing us to model realistic loop flexibility without destabilizing the global pairing topology. For an input RNA chain, we generate an ensemble $\mathcal{S} = \{\mathbf{s}^{(0)}, \ldots, \mathbf{s}^{(N-1)}\}$ with $\mathbf{s}^{(0)}$ equal to the original sequence and $\mathbf{s}^{(r)}$ for $r > 0$ obtained by one stochastic mutation sweep. Using these simple, evolution-inspired rules, our approach yields thousands of diverse homolog-like sequences within seconds. Their covariance patterns resemble those of natural MSAs: paired columns show concentrated compensatory variation, whereas loops exhibit high entropy and length heterogeneity (See Figure 5). For details on SHS, the mutation parameters, and their optimization, we refer to Appendix E.

3 Experiments

We evaluate SHS for downstream RNA structure prediction using AlphaFold 3 (AF3) on a benchmark of 479 RNA-containing PDB samples (maximum length 200 nt), comprising 91 single-chain RNAs and 388 RNA-protein complexes (Appendix C). For each RNA chain, we generate an SHS ensemble of $N = 100$ homologs (Figure 6 and Appendix E) and insert it into the AF3 input by replacing only the corresponding RNA `unpairedMsa` field, leaving all other inputs unchanged. Performance is assessed using LDDT (Local Distance Difference Test), TM Score (Template Modeling Score), and RMSD (Root-Mean-Square Deviation) (Appendix F), with SHS seeded by RNAformer trained on the same cutoff date as AF3 unless stated otherwise; additional results are reported in Appendix J.

3.1 Synthetic Homologs Improve RNA 3D Local Accuracy

As shown in Figure 2a, SHS markedly increase AF3's local RNA accuracy across the full benchmark, improving mean complex LDDT from 0.652 ± 0.306 with AF3 to 0.831 ± 0.129 when natural MSAs were replaced with SHS. These local improvements hold robustly across samples with and without available natural homologs (Figure 2a), are independent of the complex composition (Figure 2b; Figure 7), and are consistent across RNA length bins, including the longest RNAs in our dataset, though these are less frequent in the PDB (Szikszai et al., 2024) (Figure 8). Thus, improvements reflect enhanced local geometry rather than sample-specific effects. Across alternative global metrics, family-level stratifications, and a post-cutoff test set (see Appendix J), SHS largely preserve AF3's global fold accuracy while primarily improving local geometry, with the largest benefits observed when natural homology is sparse. This confirms that RNAformer-guided SHS act as a fast, database-independent source of evolutionary-like constraints that robustly enhance local RNA structure prediction without altering overall topology.

3.2 Synthetic Homologs Rescue Predictions for Orphan RNAs

To understand how SHS influences AF3's prediction quality, we analyze the predictions for single-chain RNAs without proteins. Shown in Table 1, SHS improves over vanilla AF3 for

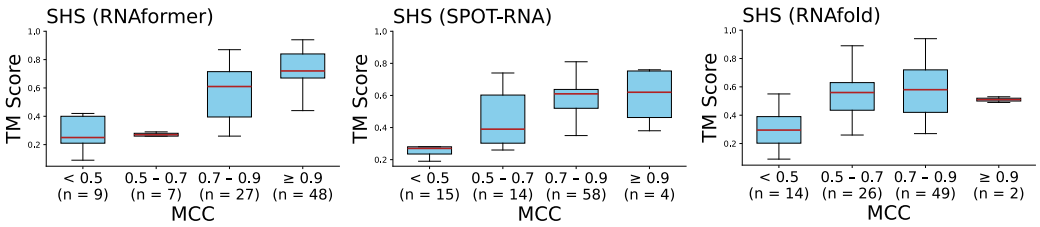

Figure 3: **Quality of synthetic homologs is set by the input 2D prediction.** TM Score of AF3 models as a function of the MCC of the 2D prediction used to seed SHS; higher 2D accuracy yields higher 3D accuracy for all three predictors.

orphan samples across performance measures, including equal TM Scores. However, where natural MSA is available, AF3 leverages the natural covariation signal, outperforming SHS on single-chain RNAs with natural MSA available, reinforcing that AF3 RNA prediction quality relies strongly on RNA-derived evolutionary signal even in the absence of protein context. Our findings establish SHS as a practical and effective strategy for enhancing RNA structure prediction for orphan RNAs, where conventional MSA-based approaches provide little or no evolutionary signal.

3.3 Synthetic Homolog Quality Depends on the Accuracy of the Underlying 2D Predictions

The effectiveness of SHS is governed by the accuracy of the underlying secondary structure prediction. Repeating the SHS workflow with interaction maps from SPOT-RNA and RNAfold in place of RNAformer yields progressively lower complex LDDT values across all benchmarks, consistent with the reduced fidelity of their 2D inputs (Figure 9a). This dependence is further reflected in a monotonic relationship between 2D prediction accuracy (MCC) and resulting 3D quality (TM Score; Figure 3).

Qualitative examples illustrate this effect: for PDB 1SJ4, SHS derived from RNAformer or ground-truth secondary structure (via DSSR (Lu et al., 2015)) produce near-native folds, whereas SPOT-RNA- and RNAfold-based SHS propagate base-pairing errors into AF3 predictions (Figure 9b). A rescue experiment for PDB 4P8Z further isolates this dependency: replacing an erroneous RNAformer-derived interaction map with the DSSR-derived ground truth restores the correct topology and reduces RMSD from 20.16 Å to 2.33 Å (Figure 10), demonstrating that SHS quality is determined by the 2D interaction pattern alone.

3.4 Synthetic Homologs Yield Coherent Consensus Structures and Increasing Covariation Signal

To assess the biological plausibility of SHS, we examined whether synthetic ensembles preserve consensus secondary structure while accumulating covariation. For three representative RNAs (3A3A, 1Q9A, 3GAO), we subsampled $N = 20$ SHS, aligned them with LocARNA (Will et al., 2012), and computed consensus structures using RNAalifold (Lorenz et al., 2011). In all cases, RNAalifold recovered the major stem architecture predicted by RNAformer, with discrepancies largely confined to non-canonical or tertiary contacts beyond the scope of thermodynamic models, while AF3 contact maps closely followed the same topology (Figure 11). We further quantified covariation using R-scape (Rivas et al., 2017) for 1Q9A, observing significant compensatory base pairs already at $N = 20$, increasing with ensemble depth ($N = 500$). Together, consensus folding and emerging covariation provide independent evidence that SHS encode a coherent and biologically plausible synthetic mutation landscape.

4 Conclusion

We show that deep evolutionary signal for RNA 3D prediction can be constructed rather than mined: accurate secondary structure alone is sufficient to synthesize alignment-like information that modern predictors such as AlphaFold 3 can effectively exploit. By inverting the classical sequence-to-structure paradigm, SHS remove database dependence, rescue orphan RNAs, and reveal that secondary structure topology, not natural homology, is the dominant driver of RNA 3D accuracy. This reframing positions synthetic evolution as a powerful, scalable representation for RNA structure prediction and a foundation for future RNA-centric biotechnology.

Acknowledgments

This research was funded by the German Research Foundation (DFG) under SFB 1597 (SmallData), grant no. 499552394, and through grant no. 417962828. The authors further acknowledge the Gauss Centre for Supercomputing e.V. for funding this project by providing computing time on the GCS JUWELS Cluster at Jülich Supercomputing Centre as well as support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 39/963-1 FUGG (bwForCluster NEMO) and grant INST 35/1597-1 FUGG (bwForCluster Helix). Frank Hutter acknowledges the financial support of the Hector Foundation. This research was funded by the European Union (via ERC Consolidator Grant DeepLearning 2.0, grant no. 101045765). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, pp. 1–3, 2024.

Stephen F Altschul, Thomas L Madden, Alejandro A Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J Lipman. Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic acids research*, 25(17):3389–3402, 1997.

Mirela Andronescu, Zhi Chuan Zhang, and Anne Condon. Secondary structure prediction of interacting rna molecules. *Journal of molecular biology*, 345(5):987–1001, 2005.

Mirela Andronescu, Vera Bereg, Holger H Hoos, and Anne Condon. RNA STRAND: the RNA secondary structure and statistical analysis database. *BMC bioinformatics*, 9:1–10, 2008.

Minkyung Baek, Ryan McHugh, Ivan Anishchenko, Hanlun Jiang, David Baker, and Frank DiMaio. Accurate prediction of protein–nucleic acid complexes using rosettafoldna. *Nature methods*, 21(1):117–121, 2024.

Clément Bernard, Guillaume Postic, Sahar Ghannay, and Fariza Tahi. Has alphafold 3 reached its success for rnas? *bioRxiv*, pp. 2024–06, 2024.

M Biasini, T Schmidt, S Bienert, V Mariani, G Studer, J Haas, N Johner, A Schenk, A Philippsen, and T Schwede. ¡i¿openstructure¡/i¿: an integrated software framework for computational structural biology. *Acta Crystallographica Section D Biological Crystallography*, 69(5):701–709, 4 2013. doi: 10.1107/s0907444913007051.

Stephen K Burley, Charmi Bhikadiya, Chunxiao Bi, Sebastian Bittrich, Henry Chao, Li Chen, Paul A Craig, Gregg V Crichlow, Kenneth Dalenberg, Jose M Duarte, et al. Rcsb protein data bank (rcsb. org): delivery of experimentally-determined pdb structures alongside one million computed structure models of proteins from artificial intelligence/-machine learning. *Nucleic acids research*, 51(D1):D488–D508, 2023.

Reed A Cartwright. Dna assembly with gaps (dawg): simulating sequence evolution. *Bioinformatics*, 21:iii31–iii38, 2005.

Jiayang Chen, Zhihang Hu, Siqi Sun, Qingxiong Tan, Yixuan Wang, Qinze Yu, Licheng Zong, Liang Hong, Jin Xiao, Irwin King, et al. Interpretable rna foundation model from unannotated data for highly accurate rna structure and function predictions. *arXiv preprint arXiv:2204.00300*, 2022.

Xinshi Chen, Yu Li, Ramzan Umarov, Xin Gao, and Le Song. RNA secondary structure prediction by learning unrolled algorithms. International Conference on Learning Representations, 2020. URL `https://openreview.net/forum?id=S1eALyrYDH`.

Padideh Danaee, Mason Rouches, Michelle Wiley, Dezhong Deng, Liang Huang, and David Hendrix. bpRNA: large-scale automated annotation and analysis of RNA secondary structure. *Nucleic acids research*, 46(11):5381–5394, 2018.

Kévin Darty, Alain Denise, and Yann Ponty. Varna: Interactive drawing and editing of the rna secondary structure. *Bioinformatics*, 25(15):1974, 2009.

Rhiju Das, Rachael C Kretsch, Adam J Simpkin, Thomas Mulvaney, Phillip Pham, Ramya Rangan, Fan Bu, Ronan M Keegan, Maya Topf, Daniel J Rigden, et al. Assessment of three-dimensional rna structure prediction in casp15. *Proteins: Structure, Function, and Bioinformatics*, 91(12):1747–1770, 2023.

Warren L DeLano et al. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. protein crystallogr*, 40(1):82–92, 2002.

Chuong B Do, Daniel A Woods, and Serafim Batzoglou. CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics*, 22(14):e90–e98, 2006.

Sean R Eddy. A probabilistic model of local sequence alignment that simplifies statistical significance estimation. *PLoS computational biology*, 4(5):e1000069, 2008.

Sean R Eddy. A new generation of homology search tools based on probabilistic inference. In *Genome Informatics 2009: Genome Informatics Series Vol. 23*, pages 205–211, 2009.

Sean R Eddy. Accelerated profile hmm searches. *PLoS computational biology*, 7(10): e1002195, 2011.

Stefan Elfwing, Eiji Uchibe, and Kenji Doya. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural networks*, 107:3–11, 2018. Special issue on deep reinforcement learning.

Matteo Figliuzzi, Pierre Barrat-Charlaix, and Martin Weigt. How pairwise coevolutionary models capture the collective residue variability in proteins? *Molecular biology and evolution*, 35(4):1018–1027, 2018.

Christoph Flamm, Julia Wielach, Michael T Wolfinger, Stefan Badelt, Ronny Lorenz, and Ivo L Hofacker. Caveats to deep learning approaches to RNA secondary structure prediction. *Biorxiv*, pp. 2021–12, 2021.

William Fletcher and Ziheng Yang. Indelible: a flexible simulator of biological sequence evolution. *Molecular biology and evolution*, 26(8):1879–1888, 2009.

Jörg Franke, Frederic Runge, and Frank Hutter. Probabilistic transformer: Modelling ambiguities and distributions for RNA folding and molecule design. *Advances in Neural Information Processing Systems*, 35:26856–26873, 2022.

Laiyi Fu, Yingxin Cao, Jie Wu, Qinke Peng, Qing Nie, and Xiaohui Xie. Ufold: fast and accurate RNA secondary structure prediction with deep learning. *Nucleic acids research*, 50(3):e14–e14, 2022.

Limin Fu, Beifang Niu, Zhengwei Zhu, Sitao Wu, and Weizhong Li. Cd-hit: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23):3150–3152, 2012.

Jonathan Ho, Nal Kalchbrenner, Dirk Weissenborn, and Tim Salimans. Axial attention in multidimensional transformers. *arXiv preprint arXiv:1912.12180*, 2019.

Ian Holmes. A probabilistic model for the evolution of rna structure. *Bmc Bioinformatics*, 5(1):166, 2004.

Liang Huang, He Zhang, Dezhong Deng, Kai Zhao, Kaibo Liu, David A Hendrix, and David H Mathews. Linearfold: linear-time approximate RNA folding by 5'-to-3'dynamic programming and beam search. *Bioinformatics*, 35(14):i295–i304, 2019.

Stefan Janssen and Robert Giegerich. The RNA shapes studio. *Bioinformatics*, 31(3): 423–425, 2015.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

Andrew J Jung, Leo J Lee, Alice J Gao, and Brendan J Frey. RTfold: RNA secondary structure prediction using deep learning with domain inductive bias. The 2022 ICML Workshop on Computational Biology. Baltimore, Maryland, USA, 2022.

Ioanna Kalvari, Eric P Nawrocki, Nancy Ontiveros-Palacios, Joanna Argasinska, Kevin Lamkiewicz, Manja Marz, Sam Griffiths-Jones, Claire Toffano-Nioche, Daniel Gautheret, Zasha Weinberg, Elena Rivas, Sean R Eddy, Robert D Finn, Alex Bateman, and Anton I Petrov. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research*, 49(D1):D192–D200, 11 2020. ISSN 0305-1048.

Jinpyo Kim, Mingi Kwon, and Jishen Zhao. Alphafold3 workload characterization: A comprehensive analysis of bottlenecks and performance scaling.

Diana Kwon. Rna function follows form-why is it so hard to predict? *Nature*, 639(8056): 1106–1108, 2025.

Sita J Lange, Daniel Maticzka, Mathias Möhl, Joshua N Gagnon, Chris M Brown, and Rolf Backofen. Global or local? predicting secondary structure and accessibility in mrnas. *Nucleic acids research*, 40(12):5215–5226, 2012.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637): 1123–1130, 2023.

Ronny Lorenz, Stephan H. Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F. Stadler, and Ivo L. Hofacker. ViennaRNA package 2.0. *Algorithms for Molecular Biology*, 6(1):26, Nov 2011. ISSN 1748-7188.

Xiang-Jun Lu, Harmen J Bussemaker, and Wilma K Olson. Dssr: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic acids research*, 43(21):e142–e142, 2015.

Nhan Ly-Trong, Suha Naser-Khdour, Robert Lanfear, and Bui Quang Minh. Alisim: a fast and versatile phylogenetic sequence simulator for the genomic era. *Molecular biology and evolution*, 39(5):msac092, 2022.

Arianna Mangiavacchi, Gabriele Morelli, and Valerio Orlando. Behind the scenes: How rna orchestrates the epigenetic regulation of gene expression. *Frontiers in Cell and Developmental Biology*, 11:1123975, 2023.

David H Mathews. How to benchmark RNA secondary structure prediction accuracy. *Methods*, 162:60–67, 2019.

Francisco McGee, Sandro Hauri, Quentin Novinger, Slobodan Vucetic, Ronald M Levy, Vincenzo Carnevale, and Allan Haldane. The generative capacity of probabilistic protein sequence models. *Nature communications*, 12(1):6302, 2021.

Kevin V Morris and John S Mattick. The rise of regulatory RNA. *Nature Reviews Genetics*, 15(6):423–437, 2014.

Eric P Nawrocki and Sean R Eddy. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, 29(22):2933–2935, 2013.

Ozgur Oksuz, Jonathan E Henninger, Robert Warneford-Thomson, Ming M Zheng, Hailey Erb, Adrienne Vancura, Kalon J Overholt, Susana Wilson Hawken, Salman F Banani, Richard Lauman, et al. Transcription factors interact with rna to regulate genes. *Molecular Cell*, 83(14):2449–2463, 2023.

Andrew Rambaut and Nicholas C Grass. Seq-gen: an application for the monte carlo simulation of dna sequence evolution along phylogenetic trees. *Bioinformatics*, 13(3): 235–238, 1997.

Roshan M Rao, Jason Liu, Robert Verkuil, Joshua Meier, John Canny, Pieter Abbeel, Tom Sercu, and Alexander Rives. Msa transformer. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8844–8856. PMLR, 2021.

Jessica S Reuter and David H Mathews. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC bioinformatics*, 11(1):1–9, 2010.

Adam J Riesselman, John B Ingraham, and Debora S Marks. Deep generative models of genetic variation capture the effects of mutations. *Nature methods*, 15(10):816–822, 2018.

Elena Rivas, Jody Clements, and Sean R Eddy. A statistical test for conserved rna structure shows lack of evidence for structure in lncrnas. *Nature methods*, 14(1):45–48, 2017.

Frederic Runge, Karim Farid, Jorg KH Franke, and Frank Hutter. RnaBench: a comprehensive library for in silico RNA modelling. *bioRxiv*, pp. 2024–01, 2024.

William P Russ, Matteo Figliuzzi, Christian Stocker, Pierre Barrat-Charlaix, Michael Socolich, Peter Kast, Donald Hilvert, Remi Monasson, Simona Cocco, Martin Weigt, et al. An evolution-based model for designing chorismate mutase enzymes. *Science*, 369(6502): 440–445, 2020.

Kengo Sato, Yuki Kato, Michiaki Hamada, Tatsuya Akutsu, and Kiyoshi Asai. Ipknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics*, 27(13):i85–i93, 2011.

Kengo Sato, Manato Akiyama, and Yasubumi Sakakibara. RNA secondary structure prediction using deep learning with thermodynamic integration. *Nature communications*, 12 (1):1–9, 2021.

Bohdan Schneider, Blake Alexander Sweeney, Alex Bateman, Jiri Cerny, Tomasz Zok, and Marta Szachniuk. When will rna get its alphafold moment? *Nucleic Acids Research*, 51 (18):9522–9532, 2023.

Jaswinder Singh, Jack Hanson, Kuldip Paliwal, and Yaoqi Zhou. RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nature communications*, 10(1):1–13, 2019.

Jaswinder Singh, Kuldip Paliwal, Tongchuan Zhang, Jaspreet Singh, Thomas Litfin, and Yaoqi Zhou. Improved RNA secondary structure and tertiary base-pairing prediction using evolutionary profile, mutational coupling and two-dimensional transfer learning. *Bioinformatics*, 37, 2021.

Michael F Sloma and David H Mathews. Exact calculation of loop formation probability identifies folding motifs in RNA secondary structures. *RNA*, 22(12):1808–1818, 2016.

Luisa Statello, Chun-Jie Guo, Ling-Ling Chen, and Maite Huarte. Gene regulation by long non-coding rnas and its biological functions. *Nature reviews Molecular cell biology*, 22(2): 96–118, 2021.

Jens Stoye, Dirk Evers, and Folker Meyer. Rose: generating sequence families. *Bioinformatics (Oxford, England)*, 14(2):157–163, 1998.

Jianlin Su, Murtadha Ahmed, Yu Lu, Shengfeng Pan, Wen Bo, and Yunfeng Liu. Roformer: Enhanced transformer with rotary position embedding. *Neurocomputing*, 568:127063, 2024.

Marcell Szikszai, Marcin Magnus, Siddhant Sanghi, Sachin Kadyan, Nazim Bouatta, and Elena Rivas. Rna3db: A structurally-dissimilar dataset split for training and benchmarking deep learning models for rna structure prediction. *Journal of Molecular Biology*, pp. 168552, 2024.

Marcell Szikszai, Marcin Magnus, Sachin Kadyan, and Elena Rivas. On inputs to deep learning for rna 3d structure prediction. *bioRxiv*, pp. 2025–02, 2025.

Zhen Tan, Yinghan Fu, Gaurav Sharma, and David H Mathews. Turbofold ii: RNA structural alignment and secondary structure prediction informed by multiple homologs. *Nucleic acids research*, 45(20):11570–11581, 2017.

Ignacio Tinoco Jr and Carlos Bustamante. How RNA folds. *Journal of molecular biology*, 293(2):271–281, 1999.

Johanna Trost, Julia Haag, Dimitri Höhler, Laurent Jacob, Alexandros Stamatakis, and Bastien Boussau. Simulations of sequence evolution: how (un) realistic they are and why. *Molecular biology and evolution*, 41(1):msad277, 2024.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

Sebastian Will, Tejal Joshi, Ivo L Hofacker, Peter F Stadler, and Rolf Backofen. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. *RNA*, 18 (5):900–914, 2012.

Chengxin Zhang, Morgan Shine, Anna Marie Pyle, and Yang Zhang. Us-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nature methods*, 19(9):1109–1115, 2022.

Chengxin Zhang, Yang Zhang, and Anna Marie Pyle. rmsa: a sequence search and alignment algorithm to improve rna structure modeling. *Journal of Molecular Biology*, 435(14): 167904, 2023.

Le Zhang, Jiayang Chen, Tao Shen, Yu Li, and Siqi Sun. Msa generation with seqs2seqs pretraining: Advancing protein structure predictions. *Advances in Neural Information Processing Systems*, 37:57324–57348, 2024a.

Yikun Zhang, Mei Lang, Jiuhong Jiang, Zhiqiang Gao, Fan Xu, Thomas Litfin, Ke Chen, Jaswinder Singh, Xiansong Huang, Guoli Song, et al. Multiple sequence alignment-based rna language model and its application to structural inference. *Nucleic Acids Research*, 52(1):e3–e3, 2024b.

A Related Work

Database Mining Typically, MSAs are obtained by mining large genomic databases, and recent approaches often use rMSA (Zhang et al., 2023) (e.g., RoseTTAFoldNA (Baek et al., 2024)) or a HMMER-based (Eddy, 2008; 2009; 2011) approach (e.g., AlphaFold 3) to build MSAs. Although RNA evolution is widely recognized as being structure-driven, these methods fail to fully incorporate structural features during homology search (Szikszai et al., 2025). Therefore, the usage of covariance models of RNA families from the RNA Family Database (Rfam) (Kalvari et al., 2020) was proposed to integrate consensus secondary structure information (Szikszai et al., 2025). While beneficial for known RNAs, this approach cannot be applied to one third of known 3D structures, as only 67% of the $1,869$ RNA representative chains in the Protein Data Bank (PDB) (Burley et al., 2023) (according to the 2024-12-04 full release of RNA3DB (Szikszai et al., 2024)) have homology to at least one Rfam family with a reasonable E-value cutoff of $1e^{-5}$ (Szikszai et al., 2025). Moreover, RNA alignments are often shallow, biased (Schneider et al., 2023), and expensive to compute (Kim et al.). The shortage of suitable MSAs is therefore a major barrier to accurate RNA 3D prediction.

Learning-Based MSA Generators For proteins, Potts models (Figliuzzi et al., 2018; Russ et al., 2020), masked language modeling (Rao et al., 2021), variational autoencoders (Riesselman et al., 2018; McGee et al., 2021), and the recent seqs2seqs framework can create deep MSAs and improve AlphaFold 2's folding accuracy (Zhang et al., 2024a). These methods, however, depend on large curated training sets that are scarce for RNA and may not generalize to novel sequence space. Foundation models trained on individual sequences, such as ESMFold (protein) can predict protein folds without alignments (Lin et al., 2023), but comparable RNA models have not yet reached the required accuracy (Das et al., 2023; Chen et al., 2022; Zhang et al., 2024b).

Our SHS pipeline is orthogonal to both approaches as it does not require any databases for expensive search and alignment steps nor training MSAs while at the same time does not suffer from generalization considerations as it is not based on learning a generative model but generation is rule-based. This makes SHS a fast and scalable alternative to existing pipelines.

Synthetic Evolution Existing approaches to generating synthetic multiple sequence alignments are almost exclusively rooted in phylogenetic sequence simulation. Classical simulators such as Seq-Gen (Rambaut & Grass, 1997), ROSE (Stoye et al., 1998), DAWG (Cartwright, 2005), INDELible (Fletcher & Yang, 2009), and more recently AliSim (Ly-Trong et al., 2022), model substitutions and indels along explicit evolutionary trees, often inferring model parameters from empirical MSAs and prioritizing biological realism and scalability for phylogenetic benchmarking. As a result, these methods require either an input MSA or a phylogenetic tree, incur substantial computational cost at scale, and are inherently inapplicable to orphan RNAs for which no homologous sequences exist. Recent analyses further show that even highly sophisticated simulators fail to reproduce key properties of empirical MSAs, underscoring that generic evolutionary realism does not necessarily translate into useful downstream signal (Trost et al., 2024). In contrast, SHS intentionally departs from realistic evolutionary simulation and reframes synthetic evolution as a representation-generation problem: starting from a single RNA sequence, it directly synthesizes deep, MSA-like ensembles from explicit structural priors in seconds, including non-canonical base pairs, pseudoknots, and tertiary constraints. By inverting the traditional use of MSAs (from inferring structure via covariation to generating covariation from structure) SHS provides a fast, database-free mechanism to inject structurally meaningful signal into modern RNA predictors.

To our knowledge, no existing method generates alignment-like RNA sequence ensembles from a single sequence without large scale database search, alignment, or model training.

B Limitations

Our study has several limitations. First, our 3D benchmark is restricted to RNAs shorter than 200 nucleotides. This range reflects the dominant length regime among structured RNAs in the PDB (Szikszai et al., 2024) and the lengths typically used to evaluate modern secondary-structure predictors (Singh et al., 2019; Fu et al., 2022). Consistent with this, RNA structure prediction methods generally achieve their highest measured accuracy within this regime, both for 2D prediction (Andronescu et al., 2005; Lange et al., 2012; Singh et al., 2019; Flamm et al., 2021) and for recent 3D modeling approaches (Bernard et al., 2024). Although RNAformer was initially trained on sequences up to 500 nucleotides and while highly desirable, systematically evaluating substantially longer RNAs or multi-chain assemblies was not computationally feasible: AF3 inference is already resource-intensive (Bernard et al., 2024), and its MSA search can dominate runtime (up to 94%) and require hundreds of gigabytes of memory for long RNAs (Kim et al.). Importantly, the SHS framework is agnostic to the choice of 2D predictor and could in principle scale as soon as accurate long-range 2D models become available. Second, SHS quality is tied directly to the fidelity of the input interaction map; erroneous contacts propagate into the generated homologs and thus into AF3 predictions. Third, our mutation model is deliberately simple and does not attempt to model natural evolutionary processes, compositional biases, or high-order covariation. Finally, all 3D evaluations were performed with AF3; while the principles should generalize, empirical validation with alternative RNA-capable engines will be important.

C Data Details

C.1 RNA Secondary Structure Data

For our experiments assessing the performance of RNAformer, we carefully curate RNA secondary structure training, validation and test sets based on existing publicly available datasets. Our data pipeline consists of data collection, secondary structure extraction, homology filtering, and data splitting.

C.1.1 Data Collection

We collect a large training data pool from the following public sources: the bpRNA-1m meta-database (Danaee et al., 2018), the ArchiveII (Sloma & Mathews, 2016) and RNAStrAlign (Tan et al., 2017) datasets provided by (Chen et al., 2020), all data from RNAStrand (Andronescu et al., 2008), as well as all RNA-containing data from the Protein Data Bank (PDB) (Burley et al., 2023), downloaded on September 23, 2023. After removing redundant sequences, our initial data pool consists of 107 098 samples. We use the commonly used test sets TS1, TS2, TS3, and TS-Hard, and the sets VL0, VL1, with additional 50 randomly selected PDB samples for validation. All four test sets, as well as VL0 and VL1 are originally provided by (Singh et al., 2021). For our evaluations, we gather TS1, TS2, and TS3 into a single test set, TS-PDB ($n = 125$) similar to (Fu et al., 2022), and keep the TS-Hard set ($n = 28$) separate.

C.1.2 Secondary Structure Extraction

Secondary structures for PDB samples were derived from the 3D structure information using DSSR (Lu et al., 2015). For NMR-solved structures, model-1 structures were considered the reference structure.

C.1.3 Three-Stage Homology Filtering

To avoid homology contamination between training, validation, and test data, we applied the following steps:

1. **Sequence-identity filtering (CD-HIT-EST).** All datasets were clustered at 80% sequence identity using CD-HIT-EST (Fu et al., 2012) (version 4.8.1). Any clusters containing test-set members were removed from the training and validation

pools. This pipeline is commonly used in previous works (Singh et al., 2019; Sato et al., 2021; Fu et al., 2022; Chen et al., 2022; Franke et al., 2022).

2. **Sensitive sequence-based filtering (BLAST-N).** For every test sample, we performed BLAST-N Altschul et al. (1997) (ncbi-blast-2.12.0+) searches against the remaining training and validation sequences at a permissive E-value cutoff of 10. All hits were removed. This procedure follows SPOT-RNA and SPOT-RNA2 (Singh et al., 2019; 2021).
3. **Structure-aware homology removal (LocARNA-P + Infernal).** For each TS-PDB and TS-Hard sample, we retrieved homologous sequences by BLAST-N against NCBI's nt database, constructed sequence-structure-based alignments with LocARNA-P (Will et al., 2012) (version 2.0.0RC10), and built covariance models (CMs) with Infernal (Nawrocki & Eddy, 2013) (version 1.1.4). Any training or validation sequence with a CM hit at a high $E \leq 0.1$ was removed.

A maximum length cutoff of 500 nt was applied during training to reduce computational cost (Singh et al., 2019; Fu et al., 2022).

C.1.4 Data Splitting

According to the strictest data processing protocol used in previous work, we apply steps 1. and 2. of our three-stage pipeline with respect to TS-PDB, and the full three-stage pipeline with respect to TS-Hard (Singh et al., 2021). This yields a pre-training set of 66 242 samples, ensuring that RNAformer is comparable to all other methods after pre-training while further considering harder similarity criteria than any other competitor regarding TS-Hard samples.

We then create a fine-tuning set of 3 432 PDB-derived samples without homology to the train set by applying the three-stage pipeline between them and the test samples in TS-PDB and TS-Hard. We denote this set FT-Non-Homolog.

To compare against AlphaFold 3, we create another fine-tuning set, consisting of 4 244 samples drawn from all PDB-derived entries in the initial pool, using the same cutoff date for data selection as AlphaFold 3 (September 30, 2021), excluding sequences with an exact match in TS-PDB or TS-Hard. This set was denoted FT-Cutoff. We provide an overview of the datasets in Table 2.

Table 2: Overview of datasets derived from experimental structures and comparative sequence analysis.

Dataset	# Samples	Length			
		Minimum	Maximum	Mean	Median
Pre-training	66242	13	500	129.0	99.0
FT-Cutoff	4244	4	200	57.9	47.0
FT-Non-Homolog	3432	11	200	61.7	48.0
Valid (Pre-Training)	1302	33	497	131.0	105.0
Valid (FT-Cutoff)	105	33	189	68.0	58.0
Valid (FT-Non-Homolog)	35	33	159	76.4	64.0
TS-PDB	125	33	189	68.0	61.0
TS-Hard	28	34	189	65.6	50.5

C.2 3D Structure Data

For evaluating 3D accuracy, we downloaded a set of 533 RNA-containing structures (between 8 February and 6 May 2025) with a maximum number of four RNA chains (filtered for five; none of the samples contained five) and five protein chains and a maximum RNA chain length of 200 nucleotides from the PDB (Burley et al., 2023), filtered to a resolution ≤ 4.5 Å. We sampled a validation set of 50 data points from the subsets of single chain RNAs and RNA-protein complexes uniformly at random with a focus on single chain RNA data.

After running AlphaFold 3, we removed 7 samples from the validation set and 4 samples from the remaining evaluation set due to disconnected chains. We used 36 single-chain RNAs and 7 RNA-protein complexes for validation during SHS parameter optimization. The remaining 479 RNA-containing samples were used for evaluation, including 91 single-chain RNAs and 388 RNA–protein complexes. We further stratified samples into orphan and non-orphan categories based on whether AlphaFold 3 identified natural RNA homologs during its database search for any of the RNA chains, and separated structures deposited before or after the AlphaFold 3/RNAformer training cutoff date (30 September 2021).

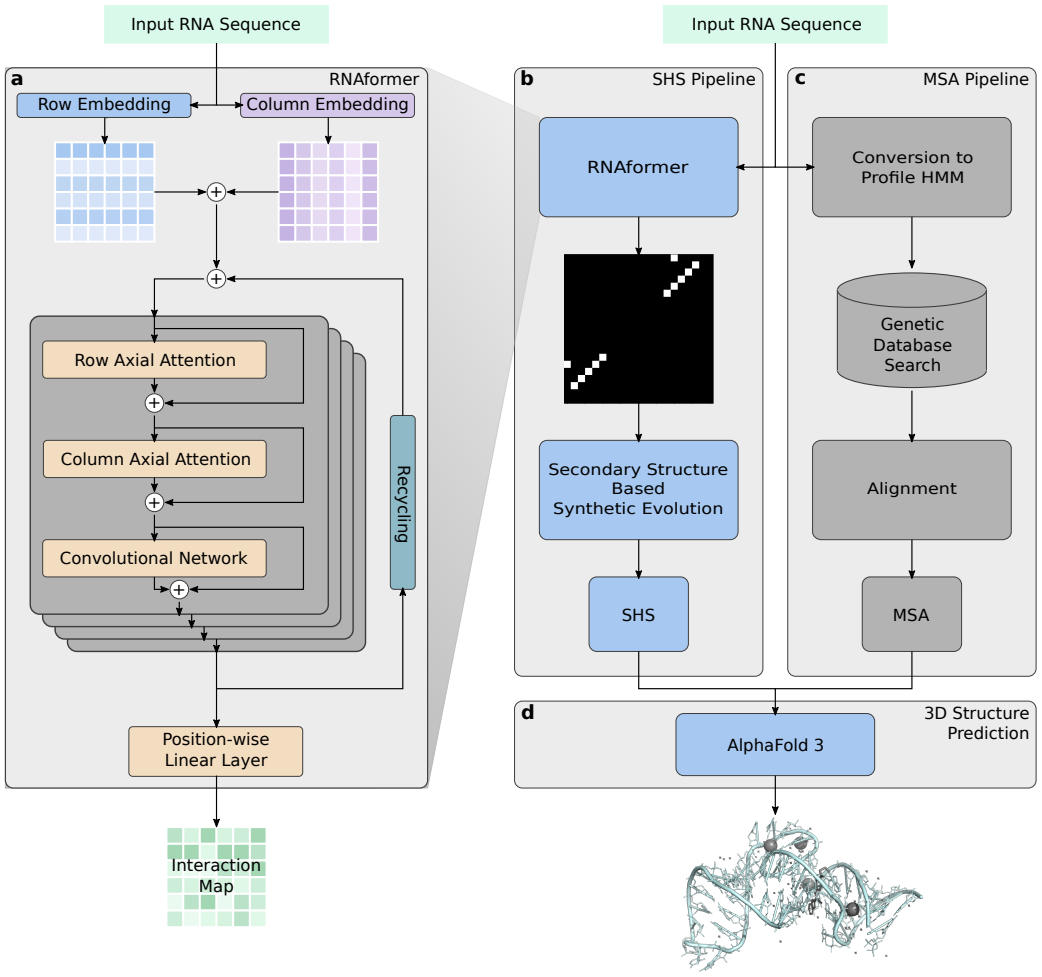

Figure 4: **RNAformer network architecture and workflow for the generation of synthetic homologous sequences. a,** After embedding the input RNA sequence into a 2D latent representation, the RNAformer successively refines this internal representation in subsequent blocks leveraging axial attention layers to finally predict a full nucleotide interaction map. **b,** The synthetic homologous sequences (SHS) generation pipeline uses RNAformer to predict an interaction map. The interaction information is subsequently used to inform the generation of SHS. **c,** For comparison with our novel SHS approach, we show a (non-iterative) standard pipeline for obtaining MSA using a profile HMM for database mining and subsequent sequence alignment as e.g. used in the HMMER-based (Eddy, 2011) genetic search of AlphaFold 3 (Abramson et al., 2024). **d,** The generated set of SHS (or the obtained MSA in the case of vanilla AlphaFold 3 predictions) serves as the interaction information for 3D Structure prediction of AlphaFold 3.

D RNAformer Details

RNAformer Figure 4 is a deep neural network for RNA secondary structure prediction that directly models the full nucleotide–nucleotide interaction map, enabling predictions of all kind of nucleotide interactions including non-canonical-, crossing-, and higher-order interactions. Inspired by AlphaFold 2 (Jumper et al., 2021), we model an interaction matrix in the latent space with the usage of axial attention (Ho et al., 2019). However, following the MSA-free design philosophy of ESMFold (Lin et al., 2023), RNAformer operates entirely on single sequences and avoids the use of natural MSAs, which are sparse and unreliable for RNA (Schneider et al., 2023; Szikszai et al., 2025).

Unlike previous RNA structure predictors (Singh et al., 2019; 2021; Chen et al., 2022; 2020; Franke et al., 2022; Jung et al., 2022), RNAformer constructs a 2D latent representation of size $l \times l$ by combining independent row and column embeddings of the input sequence with length l. This representation is processed by a stack of axial-attention blocks (Ho et al., 2019), which factorize 2D attention into efficient row-wise and column-wise operations, enabling a full-sequence receptive field independent of RNA length. A lightweight convolutional transition module models local RNA motifs such as stems and loops, while the axial attention captures long-range dependencies important for pseudoknots and tertiary-like interactions.

After M blocks, a linear projection followed by a sigmoid activation produces a binary interaction matrix $P \in \mathbb{R}^{l \times l}$, where l is the length of the input RNA sequence, allowing direct prediction of all canonical, non-canonical, crossing, and higher-order nucleotide interactions without relying on dot-bracket representations.

D.1 Overview

RNAformer predicts the full pairwise nucleotide–nucleotide interaction map of an RNA sequence by operating directly on a 2D latent representation. Given an input sequence $X \in \{A,\ C,\ G,\ U,\ N\}^l$ of length l, where N serves as a placeholder for unknown or modified nucleotides, the model embeds X independently along rows and columns to construct a tensor

$$L^{(0)} \in \mathbb{R}^{l \times l \times d}, \tag{1}$$

which is processed by M RNAformer blocks, consisting of axial-attention networks followed by a convolutional transition module. A final linear projection maps the internal representation $L^{(M)}$ to pairing logits.

D.2 Input Embeddings

The row- ($E_{\text{row}} \in \mathbb{R}^{l \times d}$) and column-wise ($E_{\text{col}} \in \mathbb{R}^{l \times d}$) embeddings $E_{\text{row}} \in \mathbb{R}^{l \times d}$ and $E_{\text{col}} \in \mathbb{R}^{l \times d}$ are computed as

$$E_{\text{row}} = \text{Embed}_{\text{row}}(X), \quad E_{\text{col}} = \text{Embed}_{\text{col}}(X), \tag{2}$$

where both embeddings have the latent dimension d. A 2D latent representation is formed by broadcasting and summation:

$$L^{(0)} = E_{\text{row}} \oplus E_{\text{col}}^T, \tag{3}$$

where $L^{(0)} \in \mathbb{R}^{l \times l \times d}$ is the 2D representation in the d-dimensional latent space, and $\oplus$ denotes the broadcasting and addition operation; i.e., $L^{(0)}[i,j] = E_{\text{row}}[i] + E_{\text{col}}[j]$. We use rotary position embedding instead of positional encoding (Su et al., 2024).

D.3 AxialAttentionNet

Full attention over a $l \times l$ 2D latent representation would create a 3D attention tensor which is computationally infeasible. RNAformer, therefore, factorizes attention into independent row- and column-wise operations (Ho et al., 2019), enabling memory-efficient multi-dimensional attention. Each of these independent *AxialAttentionNet*s consists of a linear layer to generate the query, key, and value as well as an additional linear layer to project its output. The axial attention mechanism can be represented with indices for rows i and columns j for each 2-dimensional input to the attention mechanism (Vaswani et al., 2017): query $Q \in \mathbb{R}^{l \times l \times d}$, key $K \in \mathbb{R}^{l \times l \times d}$, and value $V \in \mathbb{R}^{l \times l \times d}$ for a sequence length of l and a latent dimension of d. We compute for each column $j = 1, \cdots, l$

$$\text{AxialAttention}_{\text{row}}(Q, K, V, j) = \text{softmax}\left(\frac{Q_{:,j,:} K_{:,j,:}^T}{\sqrt{d}}\right) V_{:,j,:} \tag{4}$$

and for each row $i = 1, \cdots, l$ the respective $\text{AxialAttention}_{\text{col}}(Q, K, V, i)$. In contrast to CNN-based architectures where the receptive field expands over multiple layers, RNAformer achieves a complete receptive field by applying attention consecutively along each axis.

A full RNAformer block applies a row-wise and column-wise axial attention network *'AxialAttentionNet'*, followed by a transition network *'TransitionConvNet'*:

$$\begin{aligned} L^{(i)\prime} &= L^{(i)} + \text{AxialAttentionNet}_{\text{row}}(L^{(i)}) \\ L^{(i)\prime\prime} &= L^{(i)\prime} + \text{AxialAttentionNet}_{\text{col}}(L^{(i)\prime}) \\ L^{(i+1)} &= L^{(i)\prime\prime} + \text{TransitionConvNet}(L^{(i)\prime\prime}). \end{aligned} \tag{5}$$

We apply residual connections, pre-layer norm, and dropout to all three layers.

The resulting latent representation is then processed by a stack of M RNAformer blocks

$$L^{(i)} = \text{RNAformerBlock}(L^{(i-1)}), \quad \text{for } i = 1, 2, \ldots, M. \tag{6}$$

This allows RNAformer to iteratively refine the structure prediction in subsequent blocks.

D.4 TransitionConvNet

While the axial attention layers capture long-range information across the entire input sequence, the transition module in RNAformer is a lightweight convolutional feed-forward network composed of two 3×3 convolutions with a SiLU activation (Elfwing et al., 2018), specifically engineered to capture local RNA motifs (stems, bulges, internal loops) that are difficult for axial attention to model alone.

D.5 Recycling

Similar to (Jumper et al., 2021), the output is normalized and added to the original latent representation after each full pass of M RNAformer blocks:

$$L^{(0)} \leftarrow \text{LayerNorm}(L^{(M)}) + L^{(0)}. \tag{7}$$

The stack is then re-run. During training, gradients are computed only for the final recycling iteration.

D.6 Output Projection

A final linear layer after the RNAformer stack maps the $L^{(M)}$ latent representation to a pairing probability matrix of logits, followed by a sigmoid to represent the binary nucleotide interaction map $P \in \mathbb{R}^{l \times l}$ directly:

$$\text{P} = \text{sigmoid}(\text{Linear}(L^{(M)})). \tag{8}$$

D.7 Sparse Adjacency Loss

RNA adjacency matrices are intrinsically sparse: at most $l/2$ interacting pairs exist among l^2 entries. To compensate for this imbalance, we apply a structured masking strategy:

1. All true interactions and their neighborhoods are kept.
2. A fixed proportion of zero entries is unmasked (40% during pre-training, 80% during fine-tuning).
3. Binary cross-entropy loss is computed only on unmasked entries:

$$\mathcal{L} = -\frac{1}{|\Omega|} \sum_{(i,j)\in\Omega} \big[y_{ij} \log P_{ij} + (1 - y_{ij}) \log(1 - P_{ij}) \big], \tag{9}$$

   where Ω is the set of unmasked positions.

D.8 Training Hyperparameters

The full list of hyperparameters for training RNAformer is shown in Supplementary Tables 3 and 4.

Table 3: RNAformer hyperparameters for pre-training.

Hyperparameter	**Homology-Aware Base Model**	**Biophysical Model**
Model Architecture		
Model Dimension	256	256
Number of RNAformer Blocks	6	6
Number of Attention Heads	4	4
ConvNet Dimension	1024	1024
ConvNet Kernel Size	3	3
Embedding Dropout	0.4	0.1
Residual Dropout	0.4	0.1
Layer Normalization Epsilon	1.0e-05	1.0e-05
Initializer Range	0.02	0.02
Maximum Sequence Length	500	200
Minimum Sequence Length	10	10
Optimizer (AdamW)		
Learning Rate	0.001	0.001
Weight Decay	0.1	0.1
Beta 1	0.9	0.9
Beta 2	0.98	0.98
Learning Rate Scheduler		
Schedule	Cosine	Cosine
Decay Factor	0.01	0.01
Warmup Steps	1000	2000
Total Training Steps	20000	100000
Training Configuration		
Batch Token Size	400	600
Latent Recycling	1	6
Gradient Accumulation Steps	8	0
Number of Devices	4	4
Gradient Clipping Value	1.0	1.0
Number of Nodes	2	2
Precision	BF16 Mixed	BF16 Mixed

Table 4: RNAformer hyperparameters for fine-tuning.

Hyperparameter	**Homology-Aware Finetuning**	**AF3-like Finetuning**
Batch Size	128	128
Effective Batch Size	4	4
Maximum Sequence Length	200	200
Maximum Training Steps	1200	4000
Warmup Steps	800	2000
Learning Rate	1.0e-06	1.0e-04
Learning Rate Scheduler	Constant	Constant
Gradient Clipping Value	0.1	0.1
Number of Devices	4	4
Precision	BF16 Mixed	BF16 Mixed
Cycling	8	8
CPR Initialization	Dependent	Dependent
CPR Parameter	0.8	0.8

E Synthetic Homologous Sequence Generation Details

To provide AlphaFold 3 with alignment-like information for RNA, particularly in cases where natural MSAs are shallow or absent, we generate synthetic homologous sequences (SHS) directly from predicted nucleotide interaction maps (see Fig. 1).

For the generation process, the exact protocol is as follows.

Let $\mathbf{s} = (s_1, \ldots, s_l) \in \{\mathrm{A}, \mathrm{U}, \mathrm{G}, \mathrm{C}\}^l$ be an RNA sequence of length l. RNAformer predicts a nucleotide pairing map

$$\pi : \{1, \ldots, l\} \to \{1, \ldots, l\}, \tag{10}$$

where $\pi(i) = j$ denotes a predicted base pair ($P_{i,j} = 1$) and any site without a partner satisfies $i \notin \mathrm{dom}(\pi)$. From $\mathbf{s}$ we generate an ensemble of size N

$$\mathcal{S} = \{\mathbf{s}^{(0)}, \mathbf{s}^{(1)}, \ldots, \mathbf{s}^{(N-1)}\}, \qquad \mathbf{s}^{(0)} = \mathbf{s}, \tag{11}$$

by applying a single stochastic mutation sweep to obtain each replicate $\mathbf{s}^{(r)}$ for $r = 1, \ldots, N-1$.

E.1 Mutation Rules

For each replicate $r > 0$, we initialize $\mathbf{s}^{(r)} \leftarrow \mathbf{s}$ and scan positions $i = 1, \ldots, l$ in sequence order. Insertions and deletions act on the working copy and shift downstream indices; the partner index $\pi(i)$ is updated accordingly when necessary.

Unpaired sites ($i \notin \mathrm{dom}(\pi)$). At unpaired positions, we model loop and bulge variability using a combination of insertions, deletions, and substitutions:

- **Insertions**
  - With probability $p_{\text{long\_ins}}$, we insert a random nucleotide block of length
    $$k \sim \mathrm{Uniform}\big(2, \lfloor L_{\max}^{\text{ins}} \rfloor\big) \tag{12}$$
    immediately after position i.
  - Otherwise, with probability $p_{\text{loop\_ins}}$, we insert a single random nucleotide.
- **Deletions**
  - With probability $p_{\text{long\_del}}$, we delete a contiguous block of size
    $$k \sim \mathrm{Uniform}\big(2, \lfloor L_{\max}^{\text{del}} \rfloor\big) \tag{13}$$
    starting at position i.
  - Otherwise, with probability $p_{\text{loop\_del}}$, we delete the nucleotide s_i.
- **Substitutions** If no insertion or deletion was applied at i, we replace the nucleotide by a uniformly random draw
  $$s_i \leftarrow \mathrm{Uniform}\{\mathrm{A}, \mathrm{U}, \mathrm{G}, \mathrm{C}\}. \tag{14}$$

Paired sites ($i < \pi(i)$). For each base pair $(i, \pi(i))$, we update the pair jointly in a structure-aware fashion:

- With probability $p_{\text{stem\_keep}}$, we retain the original pair $(s_i, s_{\pi(i)})$.
- Otherwise, we sample a new pair according to a pair-specific mutation dictionary:
  $$(s_i, s_{\pi(i)}) \leftarrow \texttt{mutate\_pair}(s_i, s_{\pi(i)}), \tag{15}$$
  where with probability p_{wobble} we enforce a GU/UG wobble and, otherwise, draw uniformly from a predefined set of compensatory or near-canonical pairs.

The mapping used by `mutate_pair` is summarized in Table 5. Any pair not listed defaults to the set of standard Watson–Crick pairs $\{\mathrm{AU}, \mathrm{UA}, \mathrm{GC}, \mathrm{CG}\}$.

Table 5: **Base-pair mutation mapping used in the SHS generation pipeline.** If a pair is selected for mutation, the new pair is sampled uniformly from the corresponding set.

Original pair	Allowed mutated pairs
CU	GU, AU, CG
CA	UA, CG
GA	UA, GC, GU
CC	CG, GC
AA	AU, UA
UU	AU, UA
GG	GC, CG, GU, UG
GC	AU, CG, GC
CG	GC, AU
AU	UA, GC
UA	AU, CG
GU	GC, AU, UG
UG	GC, AU, GU
Other	AU, UA, GC, CG

E.2 Control parameters

Mutation diversity and indel length are controlled by

$$\{N,\ p_{\text{loop\_ins}},\ p_{\text{long\_ins}},\ p_{\text{loop\_del}},\ p_{\text{long\_del}},\ p_{\text{wobble}},\ p_{\text{stem\_keep}},\ \alpha_{\text{ins}},\ \alpha_{\text{del}}\},$$

with maximum insertion and deletion block lengths

$$L_{\max}^{\text{ins}} = \lfloor \alpha_{\text{ins}} \cdot L \rfloor, \qquad L_{\max}^{\text{del}} = \lfloor \alpha_{\text{del}} \cdot L \rfloor.$$

We conduct a small grid search over 569 configurations and a scaling analysis of the SHS ensemble size (see Supplementary Table 6 and Extended Data Fig. 3). For our hyperparameter search, we use AlphaFold 3 predictions with RNAformer-SHS inputs and select configurations based on validation performance. As shown in Supplementary Table 6, our search covers only a fraction of the possible hyperparameters space to keep the optimization computationally feasible. We expect further gains in performance for SHS with evaluations of more configurations. The scaling experiment was performed on the full benchmark ($n = 479$) after selection of an appropriate configuration. The finally selected configuration for our SHS generation is shown in Supplementary Table 6. This was used as our default for AlphaFold 3 experiments (unless stated otherwise in the main text).

Table 6: **SHS mutation hyperparameters and final configuration.** The values were chosen based on the hyperparameter optimization and SHS-depth analysis.

Parameter	Selected Value	Evaluated Parameter Range
Ensemble size N	100	$\{100, 5\,000, 10\,000, 20\,000\}$
$p_{\text{loop\_ins}}$	0.2	$[0.2, 0.8]$; step 0.1
$p_{\text{loop\_del}}$	0.8	$[0.2, 0.8]$; step 0.1
$p_{\text{long\_ins}}$	0.1	$\{0.05, 0.1\}$
$p_{\text{long\_del}}$	0.05	$\{0.05, 0.1\}$
p_{wobble}	0.1	$\{0.1\}$
$p_{\text{stem\_keep}}$	0.99	$\{0.8, 0.9, 0.99\}$
α_{ins}	0.1	$\{0.1\}$
α_{del}	0.1	$\{0.1\}$

E.3 Integration into AlphaFold 3

For each RNA-containing complex described, we first ran the original AlphaFold 3 pipeline, including database and template searches, to obtain reference JSON input files. For each RNA chain, we then replaced only the RNA MSA entry (`"unpairedMsa"`) with the SHS

ensemble described above. All other entries, including protein MSAs, protein templates, seeds, or modifications, were left unchanged. The modified JSON files were subsequently passed to the standard AlphaFold 3 inference code, ensuring that differences in prediction accuracy arise solely from the substitution of natural RNA MSAs with SHS.

F Evaluation

F.1 Secondary Structure Prediction

Learning-Based Algorithms We use the following learning-based competitors: MXFold2 (Sato et al., 2021) (version: 0.1.2), SPOT-RNA (Singh et al., 2019) (default parameters), UFold (Fu et al., 2022) (default parameters), ContraFold (Do et al., 2006) (version 2.02; command: `contrafold predict <fast file>`), RNA-FM (Chen et al., 2022) (default parameters).

Non-Learning-Based Algorithms We use the following non-learning-based competitors: RNAfold (Lorenz et al., 2011) (version 2.6.3), LinearFold (Huang et al., 2019) (version 1.0; variants -C and -V), ipknot (Sato et al., 2011) (0.0.4), pkiss (Janssen & Giegerich, 2015) (version 2.2.14), RNAStructure (Reuter & Mathews, 2010) (version 6.3). All evaluations were performed using RnaBench (Runge et al., 2024).

F.1.1 Secondary Structure Evaluation Measures

For the evaluation of secondary structures of 3D RNA predictions, we extract nucleotide information directly from the 3D predictions using DSSR (Lu et al., 2015), and use the commonly used F1-Score and the Matthews Correlation Coefficient (MCC) as previously proposed (Mathews, 2019). The MCC is based on a confusion matrix, which describes the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) of a given prediction. The MCC is then calculated as follows.

$$MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \quad . \tag{16}$$

The F1 score describes the harmonic mean of precision ($PR = TP/(TP + FP)$) and recall ($RC = TP/(TP + FN)$ written as $F1 = 2 \cdot TP/(2 \cdot TP + FP + FN)$.

For all RNA secondary structure evaluations, we use our recently proposed benchmark RnaBench (Runge et al., 2024).

F.2 3D Structure Prediction

F.2.1 AlphaFold 3 Predictions

All AlphaFold 3 predictions were performed on a single NVIDIA A100 GPU (80GB). For all evaluations, we select model 0.

F.2.2 3D Structure Evaluation Measures

For the 3D structural evaluation of RNA and RNA-protein complexes, we used PyMOL (DeLano et al., 2002) (version 3.1.0) for structural alignment and coordinate extraction. All complexes were cleaned by removing solvents, ions, and other non-polymeric components, and were subsequently separated. The predicted structures of the extracted RNA-protein and single chain RNAs were then aligned to their corresponding crystallographic reference structures using C^{α} atoms for proteins (polymer.protein and name CA) and phosphorus (P) atoms for RNA (polymer.nucleic and name P). The aligned structures were saved in separate PDB files. These files were then used to calculate both the RNA and RNA-protein complex RMSD, LDDT and TM values, denoted RNA-X or cX, respectively, where X represents RMSD, LDDT or TM.

Following structural superposition in PyMOL, RMSD was computed as the Root-Mean-Square Deviation of aligned backbone atoms providing a direct measure of coordinate accuracy in Ångströms. In Cartesian coordinates, the RMS distance $d_{i,j}$ between predicted and reference conformations i and j of a molecule, respectively, is defined as:

$$RMSD = \sqrt{d_{i,j}} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (r_k^{(i)} - r_k^{(j)})^2}, \tag{17}$$

where N is the number of backbone atoms in the sum, k is the atom index, and $r_k^{(i)}, r_k^{(j)}$ are the Cartesian coordinates of atom k in conformations i and j. The minimum value is obtained by an optimal superposition of the two structures.

The TM-score, which is based on global superposition and uses backbone coordinates (C^α atoms for proteins and C3' atoms for nucleic acids), was calculated using the USalign tool (Zhang et al., 2022) (version 20241108) with RNA-specific parameters (-mol RNA -mm 0 -ter 2) for RNA structures and protein-specific parameters (-mol prot -mm 0 -ter 2) for protein moieties, providing a normalized, length-independent metric that, unlike RMSD, is less sensitive to local errors and focuses on the overall topology of the structures (Bernard et al., 2024).

$$TM = \max \left(\frac{1}{L_{\text{ref}}} \sum_{i=1}^{L_{\text{ali}}} \frac{1}{1 + \left(\frac{d_i}{d_0(L_{\text{ref}})} \right)^2} \right), \tag{18}$$

where L_{ref} is the length of the reference structure, L_{ali} is the number of aligned residues between the predicted and reference structures, d_i is the distance between the i-th pair of aligned residues and $d_0(L_{\text{ref}})$ is a scaling factor to ensure the score of random RNA pairs is independent of RNA sequence length ($d_0(L_{\text{ref}}) = 0.6\sqrt{L_{\text{ref}} - 0.5} - 2.5$).

In addition, an adapted code snippet from AlphaFold 2 (Jumper et al., 2021) was used for LDDT calculation, which computes an approximate Local Distance Difference Test score (Biasini et al., 2013) by comparing inter-atomic distances between predicted and reference structures using backbone atoms (C^α atoms for proteins and C3' atoms for nucleic acids) with a 15 Å cutoff and standard LDDT δ thresholds (0.5, 1.0, 2.0, and 4.0 Å) to capture varying levels of distance accuracy.

$$\text{LDDT} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{L_i} \sum_{j \in \text{neighbours}(i)} \Theta(\delta - |d_{ij}^{\text{ref}} - d_{ij}^{\text{pred}}|), \tag{19}$$

where N is the total number of atoms considered, L_i is the number of neighboring atoms of atom i within the distance cutoff, d_{ij}^{ref} is the distance between atoms i and j in the reference structure, d_{ij}^{pred} is the distance between atoms i and j in the predicted structure, Θ is a step function that returns 1 if $|d_{ij}^{\text{ref}} - d_{ij}^{\text{pred}}| \leq \delta$, and 0 otherwise.

G Alignments and Thermodynamic Consensus Structures

We use LocARNA-P (`mlocarna --probabilistic`) to obtain sequence and structure-based alignments in Stockholm format from our synthetic homologous sequences. RNAalifold (version 2.7.0) was run with "-p"-option on these alignments. We show the resulting dot-plots produced by RNAalifold .ps files directly.

H Structure Visualization

Secondary structure arc-plots were produced using VARNA (Darty et al., 2009) (version 3.93). 3D structures were visualized using PyMOL (version 3.1.0).

I Statistical Analysis

We generally report statistical results in terms of paired Wilcoxon signed-rank test. The reported significance levels are $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$. We use the web service of R-scape available at `http://eddylab.org/R-scape/` to analyze LocARNA alignments of our synthetic homologous sequences.

Table 7: The mean F1-score on the base pair prediction of three RNAformer runs with different randomly chosen seeds in comparison to other methods on the TS-PDB and TS-Hard benchmarks. We observe that the RNAformer outperforms existing methods, despite having a stricter data pipeline.

Model	TS-PDB	TS-Hard
RNAformer	0.764	0.679
RNAformer pretrain	0.723	0.601
UFold (Fu et al., 2022)	0.738	0.587
SPOT-RNA (Singh et al., 2019)	0.734	0.663
RNA-FM (Chen et al., 2022)	0.729	0.665
MXFold2 (Sato et al., 2021)	0.691	0.667
ContraFold (Do et al., 2006)	0.669	0.625
RNAFold (Lorenz et al., 2011)	0.659	0.636
LinearFold-V (Huang et al., 2019)	0.657	0.633
IPknot (Sato et al., 2011)	0.652	0.611
RNAstructure (Reuter & Mathews, 2010)	0.642	0.606
LinearFold-C (Huang et al., 2019)	0.632	0.610
PKiss (Janssen & Giegerich, 2015)	0.615	0.613
RNAformer cutoff-date	0.855	0.845
Alphafold 3 (Abramson et al., 2024)	0.817	0.688

J Additional Results

J.1 Secondary Structure Prediction Benchmarking

J.1.1 Synthetic Homologous Sequences

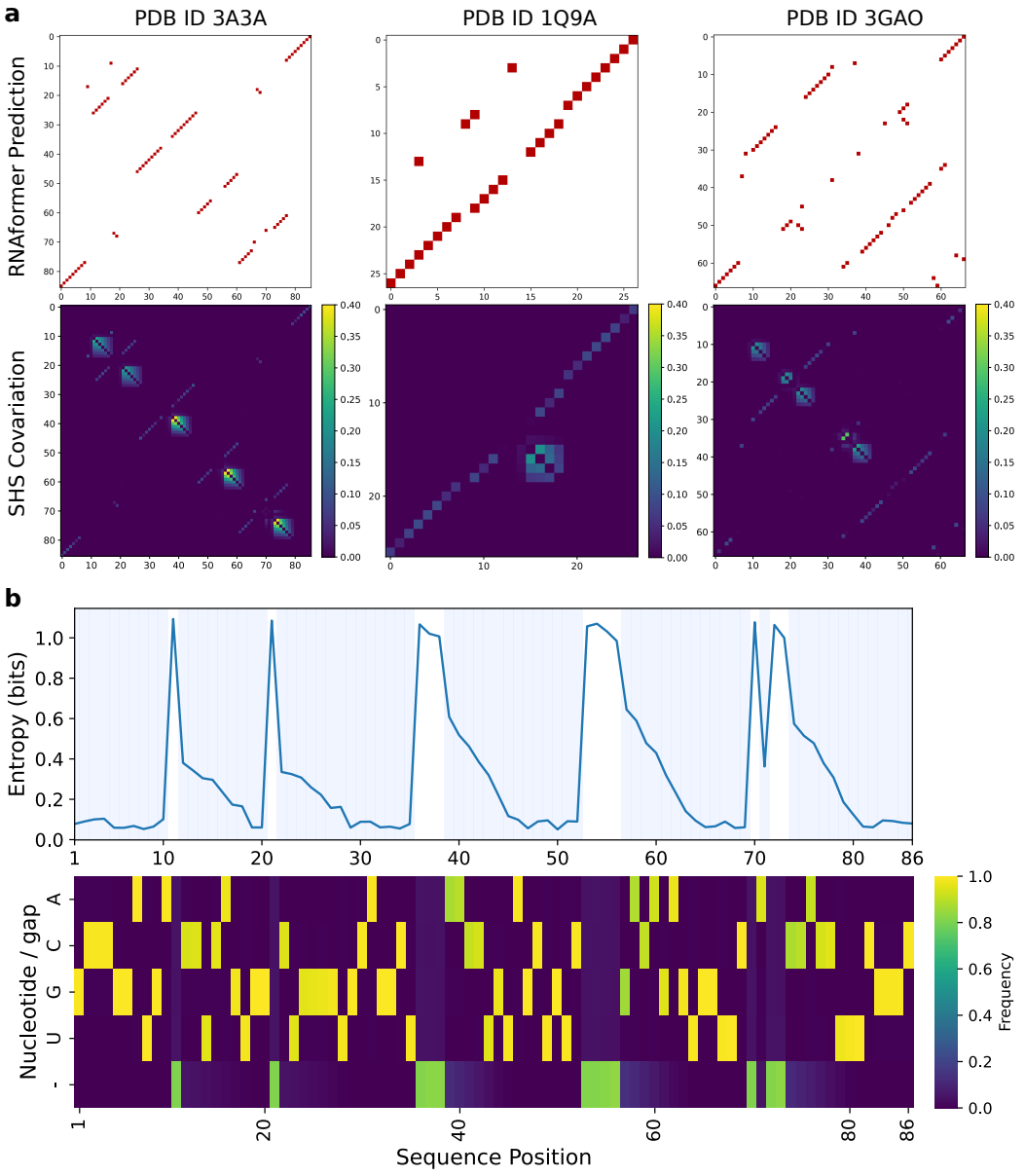

Figure 5: **Synthetic homologous sequences reproduce predicted covariation patterns and structured position-wise diversity. a,** For three representative RNAs (PDB Ids: 3A3A, 1Q9A, 3GAO; same as in Fig 11), we compare the RNAformer-predicted contact maps (top) with mutual information (MI) computed from the SHS (bottom). High-MI off-diagonal signal aligns with predicted stems and tertiary contacts. **b,** Per-position entropy (top) and nucleotide/gap (bottom row) frequencies (bottom) along the sequence of 3A3A reveal reduced diversity in paired positions (light blue) and increased variability in loop regions, indicating that SHS inject structured rather than random noise. Results are shown for 10 000 generated SHS for each RNA.

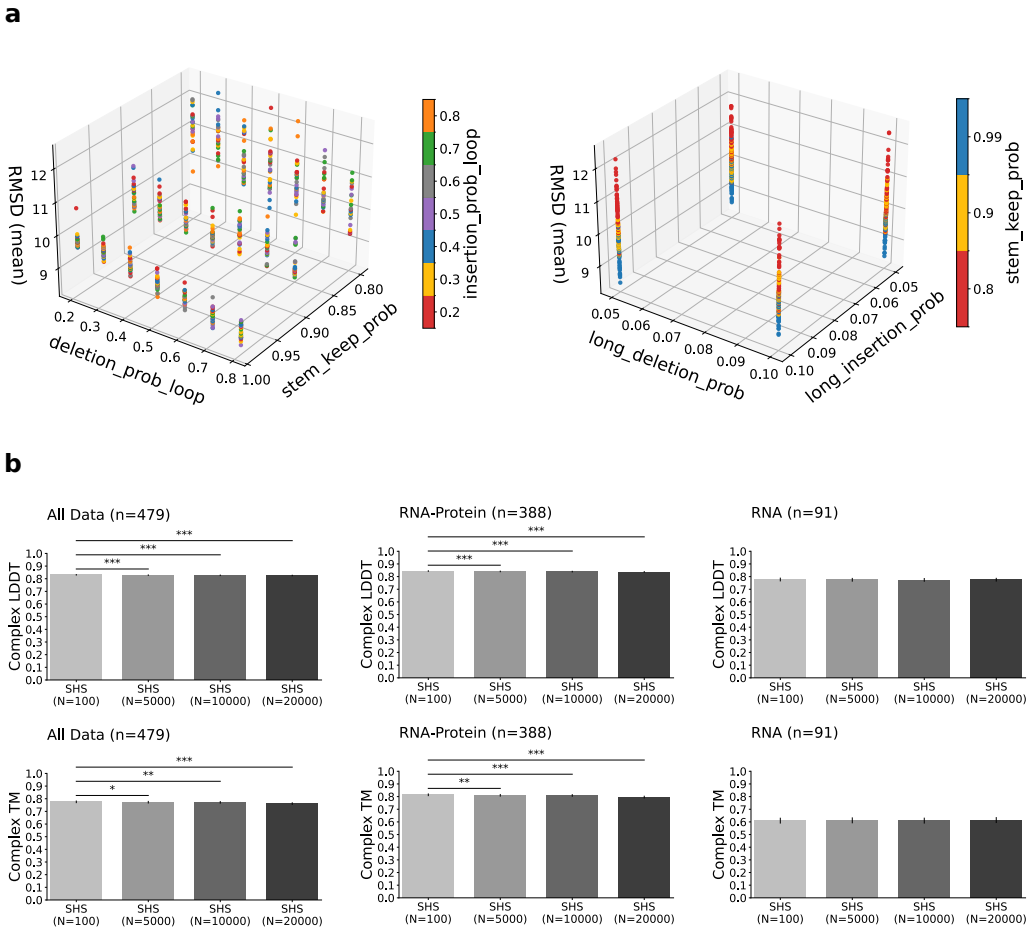

Figure 6: **Hyperparameter analysis for generation of synthetic homologs. a,** Mean RMSD for 569 configurations evaluated on the 36 single-chain RNAs of the validation set. The deletion probability for loop regions and the probability for keeping stem regions unchanged (left plot) have the largest impact on performance while the length of the insertions and deletions do not have a strong impact over the selected parameter ranges. **b,** Scaling analysis for the number of synthetic homologs (SHS) generated. We show LDDT and TM Score results on the test set for the generation of 100, 5 000, 10 000, and 20 000 homologs using RNAformer for initial structure predictions. SHS with 100 homologs is yields significantly better results compared to the generation of more homologs on RNA-Protein samples. For RNA monomers, the difference is not significant. Error bars show the s.e.m. Wilcoxon significance levels: $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

J.2 3D Benchmarking

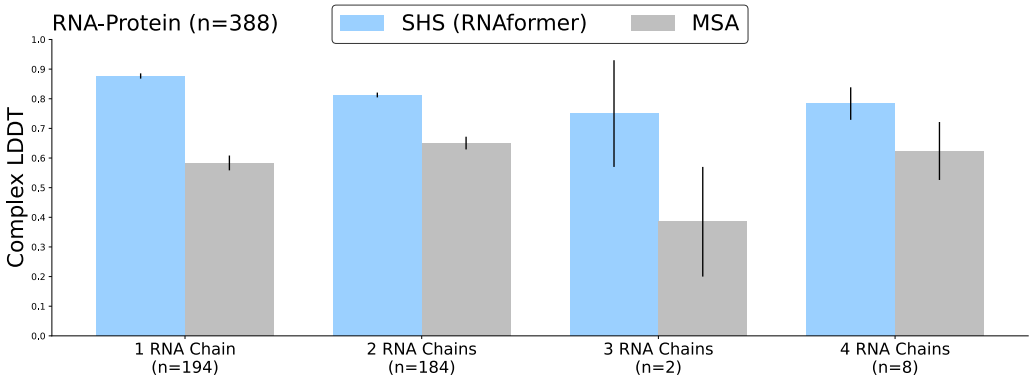

Figure 7: **Synthetic homologous sequences improve local 3D accuracy across different complex compositions.** For RNA-protein complexes, RNAformer-seeded SHS substantially improves local RNA 3D accuracy when supplied to AlphaFold 3 compared to vanilla AF3 predictions, independent of the number of RNA chains in the complex.

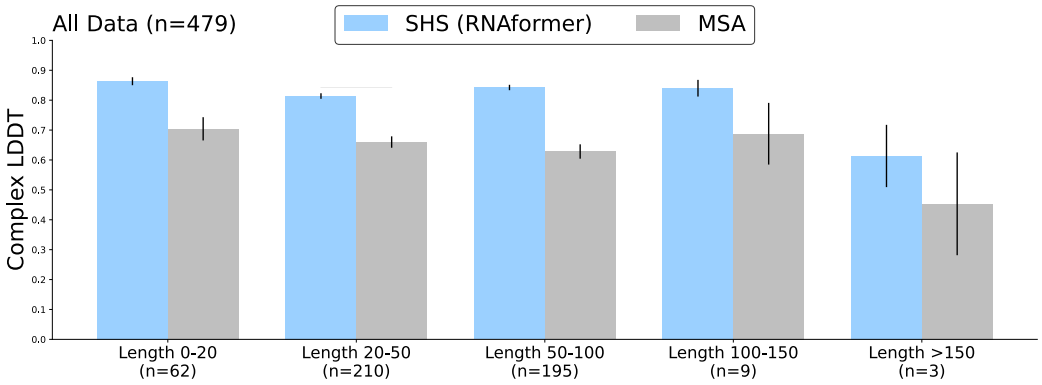

Figure 8: **Synthetic homologs improve RNA 3D accuracy across sequence-length regimes.** Mean complex LDDT of AlphaFold 3 predictions using either RNAformer-guided synthetic homologous sequences (SHS) or natural MSAs, stratified by RNA length. Across all length bins (0–20, 20–50, 50–100, 100–150, and > 150 nucleotides), SHS consistently yield higher local structural accuracy than natural MSA inputs, with the strongest improvements in the three most populated bins ($n = 62$, $n = 210$, and $n = 195$). For longer RNAs (100–150 nt and > 150 nt), SHS maintain comparable or improved accuracy despite limited sample sizes ($n = 9$ and $n = 3$). For multi-chain complexes, we use the length of the longest chain of the complex. Error bars denote the s.e.m.

Table 8: Comparison of AlphaFold 3 (AF3) with and without synthetic homologs (SHS; RNAformer, $N = 100$) across benchmark splits. Values are mean ± s.d.; p-values from Wilcoxon signed-rank tests. Significance levels: ns, not significant; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

LDDT

Subset	n	AF3	SHS ($N = 100$)	p-value	Sig.
All	479	0.652 ± 0.306	0.831 ± 0.129	8.21×10^{-20}	***
All w/o homologs	265	0.675 ± 0.287	0.835 ± 0.123	2.82×10^{-11}	***
All w/ homologs	214	0.623 ± 0.325	0.826 ± 0.135	2.99×10^{-7}	***
Post-2021 (all)	166	0.731 ± 0.187	0.772 ± 0.121	0.253	ns
Post-2021 w/o homologs	115	0.761 ± 0.156	0.787 ± 0.112	0.173	ns
Post-2021 w/ homologs	51	0.665 ± 0.230	0.737 ± 0.135	0.194	ns

TM-score

Subset	n	AF3	SHS ($N = 100$)	p-value	Sig.
All	479	0.793 ± 0.224	0.776 ± 0.225	0.0456	*
All w/o homologs	265	0.803 ± 0.222	0.794 ± 0.223	0.570	ns
All w/ homologs	214	0.781 ± 0.226	0.753 ± 0.226	0.0137	*
Post-2021 (all)	166	0.737 ± 0.240	0.732 ± 0.237	0.782	ns
Post-2021 w/o homologs	115	0.789 ± 0.214	0.793 ± 0.203	0.474	ns
Post-2021 w/ homologs	51	0.618 ± 0.253	0.593 ± 0.250	0.432	ns

RMSD (Å)

Subset	n	AF3	SHS ($N = 100$)	p-value	Sig.
All	479	12.144 ± 13.744	12.449 ± 13.421	0.301	ns
All w/o homologs	265	13.789 ± 13.518	13.813 ± 13.028	0.844	ns
All w/ homologs	214	10.108 ± 13.749	10.760 ± 13.706	0.0146	*
Post-2021 (all)	166	15.038 ± 11.320	15.527 ± 11.240	0.370	ns
Post-2021 w/o homologs	115	14.824 ± 10.830	14.480 ± 10.298	0.715	ns
Post-2021 w/ homologs	51	15.522 ± 12.341	17.888 ± 12.807	0.139	ns

Table 9: Family-level comparison of AlphaFold 3 with natural MSAs (MSA) and RNAformer-guided synthetic homologs (SHS). We report mean per-chain LDDT scores (MSA mean, SHS mean) and Wilcoxon signed-rank p-values for single-chain RNAs and RNA-protein complexes, split by subset and RNA-family annotation status. a2021 denotes data deposited in PDB after the training cutoff date of AlphaFold 3 and RNAformer, b2021 denotes data from before the training cutoff date that the models might have visited during training. Families were annotated by `cmscan` against Rfam using GA cutoffs (`--rfam --cut_ga --nohmmonly --oskip --fmt 2` with appropriate `clanin` and CM files). For complexes, an RNA family was assigned if any of the RNA chains had a hit with any covariance model. Significance levels: ns, not significant; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Subset	Molecule	Family status	n	MSA mean	SHS mean	p (Wilcoxon)	Sig.
All	RNA	all	91	0.810	0.774	3.567×10^{-9}	***
	RNA	has fam	58	0.872	0.811	1.172×10^{-10}	***
	RNA	no fam	33	0.701	0.708	7.272×10^{-1}	ns
	Complex	all	388	0.615	0.839	1.516×10^{-10}	***
	Complex	has fam	106	0.506	0.863	5.538×10^{-10}	***
	Complex	no fam	282	0.656	0.831	2.615×10^{-4}	***
a2021	RNA	all	31	0.658	0.657	9.763×10^{-1}	ns
	RNA	has fam	8	0.752	0.660	4.883×10^{-2}	*
	RNA	no fam	23	0.625	0.656	1.334×10^{-1}	ns
	Complex	all	135	0.748	0.798	3.758×10^{-2}	*
	Complex	has fam	18	0.796	0.831	8.880×10^{-1}	ns
	Complex	no fam	117	0.741	0.793	3.011×10^{-2}	*
b2021	RNA	all	60	0.888	0.839	9.931×10^{-11}	***
	RNA	has fam	50	0.891	0.836	1.140×10^{-9}	***
	RNA	no fam	10	0.875	0.853	1.311×10^{-1}	ns
	Complex	all	253	0.544	0.868	1.438×10^{-20}	***
	Complex	has fam	88	0.446	0.867	2.002×10^{-10}	***
	Complex	no fam	165	0.596	0.869	5.071×10^{-12}	***

Table 10: **Per-family counts and complex-level LDDT of AlphaFold 3 with natural MSAs and RNAformer-based synthetic homologs (SHS) for single-chain RNAs and RNA-protein complexes.** We removed one sample containing two tRNA chains to show per chain evaluations (cLDDT: MSA, 0.880; SHS, 0.950). "–" denotes undefined standard deviation (n=1) or absence of that family in the corresponding category. RNAformer-based SHS performance appears more consistent across all families compared to vanilla AlphaFold 3 using natural MSAs.

RNA Family	n (RNA)	MSA	SHS	n (complexes)	MSA	SHS
5S_rRNA	0	–	–	1	0.150 ± –	0.810 ± –
Archaea_SRP	0	–	–	1	0.900 ± –	0.900 ± –
Bacteria_small_SRP	1	0.870 ± –	0.840 ± –	3	0.283 ± 0.293	0.820 ± 0.171
CPEB3_ribozyme	0	–	–	2	0.875 ± 0.092	0.880 ± 0.085
DENV_SLA	0	–	–	1	0.750 ± –	0.750 ± –
FMN	1	0.880 ± –	0.850 ± –	0	–	–
HDV_ribozyme	0	–	–	11	0.108 ± 0.004	0.977 ± 0.005
LSU_rRNA_archaea	0	–	–	2	0.505 ± 0.488	0.875 ± 0.035
LSU_rRNA_bacteria	0	–	–	7	0.720 ± 0.414	0.913 ± 0.082
Metazoa_SRP	0	–	–	1	0.160 ± –	0.890 ± –
Purine	23	0.871 ± 0.041	0.794 ± 0.030	0	–	–
SAM	13	0.943 ± 0.012	0.913 ± 0.014	1	0.760 ± –	0.760 ± –
SAM-I-IV	1	0.850 ± –	0.720 ± –	0	–	–
SSU_rRNA_eukarya	0	–	–	1	0.830 ± –	0.810 ± –
THF	6	0.893 ± 0.008	0.818 ± 0.004	0	–	–
TLS-PK3	1	0.640 ± –	0.590 ± –	0	–	–
TPP	1	0.830 ± –	0.610 ± –	1	0.800 ± –	0.800 ± –
Xanthine	1	0.720 ± –	0.710 ± –	0	–	–
c-di-GMP-I-GGC	0	–	–	2	0.085 ± 0.007	0.495 ± 0.092
group-II-D1D4-3	1	0.840 ± –	0.650 ± –	0	–	–
mir-TAR	1	0.730 ± –	0.750 ± –	0	–	–
nadA	2	0.795 ± 0.007	0.755 ± 0.007	1	0.110 ± –	0.540 ± –
raiA	1	0.610 ± –	0.430 ± –	0	–	–
tRNA	2	0.900 ± 0.014	0.900 ± 0.014	69	0.552 ± 0.336	0.859 ± 0.122
tRNA-Sec	3	0.850 ± 0.044	0.850 ± 0.052	0	–	–
tmRNA	0	–	–	1	0.150 ± –	0.850 ± –

LDDT-Bins Color Scheme

0.0 - 0.2 | 0.2 - 0.4 | 0.4 - 0.6 | 0.6 - 0.8 | 0.8 - 1.0

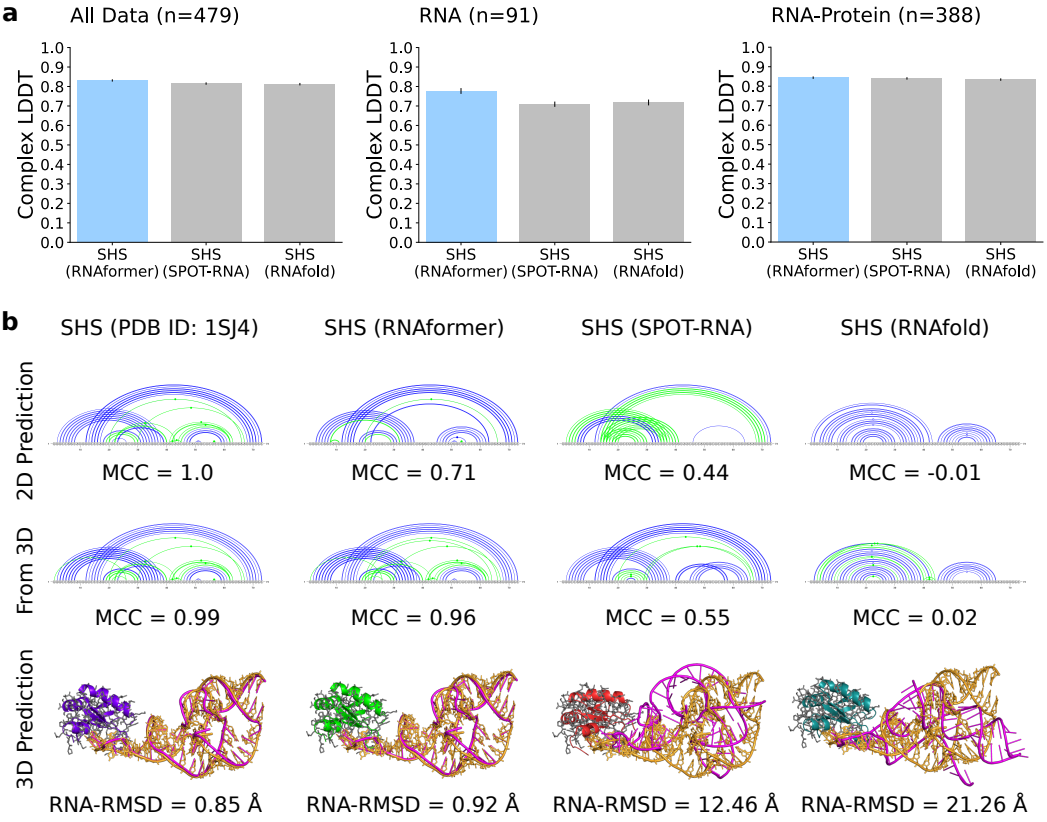

Figure 9: **Quality of synthetic homologs is set by the input 2D prediction. a,** Complex LDDT of AF3 models when SHS are generated from RNAformer, SPOT-RNA or RNAfold contact maps for all RNAs ($n = 479$), single-chain RNAs ($n = 91$) and RNA-protein complexes ($n = 388$). RNAformer-based SHS give the highest mean cLDDT in all subsets. **b,** Example (PDB ID 1SJ4): Nucleotide interactions from PDB or predicted by RNAformer, SPOT-RNA, or RNAfold (top). 2D interactions (middle) derived from corresponding 3D structures (bottom) generated by AF3 using SHS informed by ground-truth 2D (DSSR), RNAformer, SPOT-RNA and RNAfold. True RNA, orange; predicted RNA, magenta. Error bars show s.e.m.

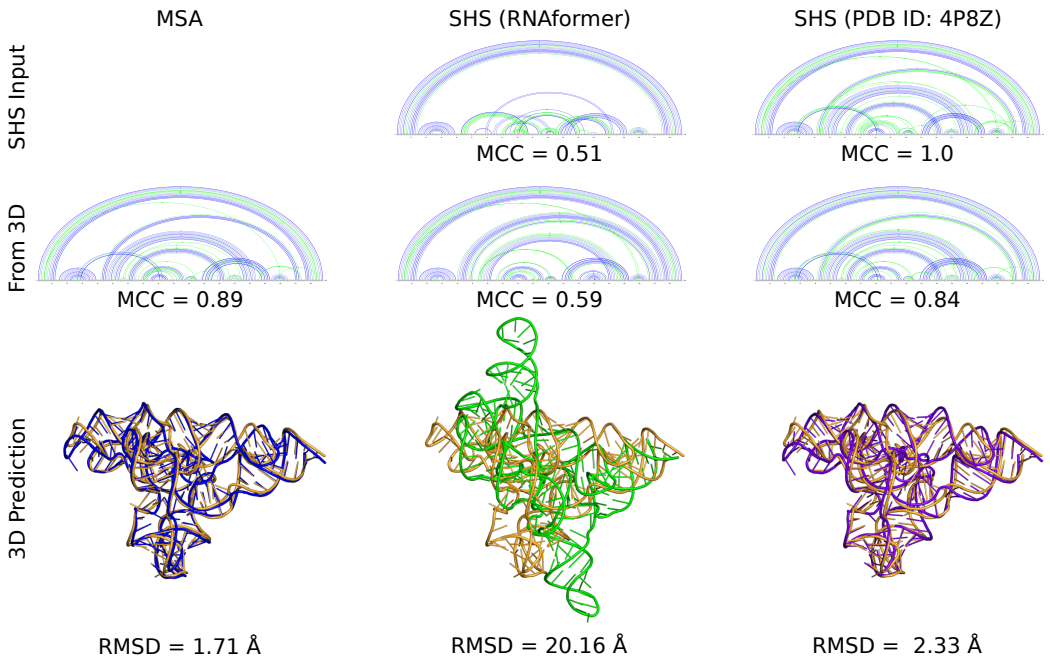

Figure 10: **SHS-guided rescue of 4P8Z.** For 4P8Z, we show a rescue experiment for a mispredicted 2D nucleotide interaction map of RNAformer (top, middle). We show the nucleotide contacts used as input for our SHS generation (top row), the nucleotide contacts derived from the respective 3D prediction (middle row) and the corresponding 3D predictions (bottom row) for AlphaFold 3 (AF3; left), AF3 using SHS with RNAformer predictions (middle) and AF3 using SHS based on ground truth contacts derived via DSSR (right). Note that the input to AF3 is an MSA obtained via database mining and that there thus is no SHS input for AF3 (top, left). Using the ground truth 2D contacts (top, right) as input for SHS restores the AF3 3D structure (bottom right), showing that AF3 closely follows SHS-derived 2D signals.

J.2.1 Case Study

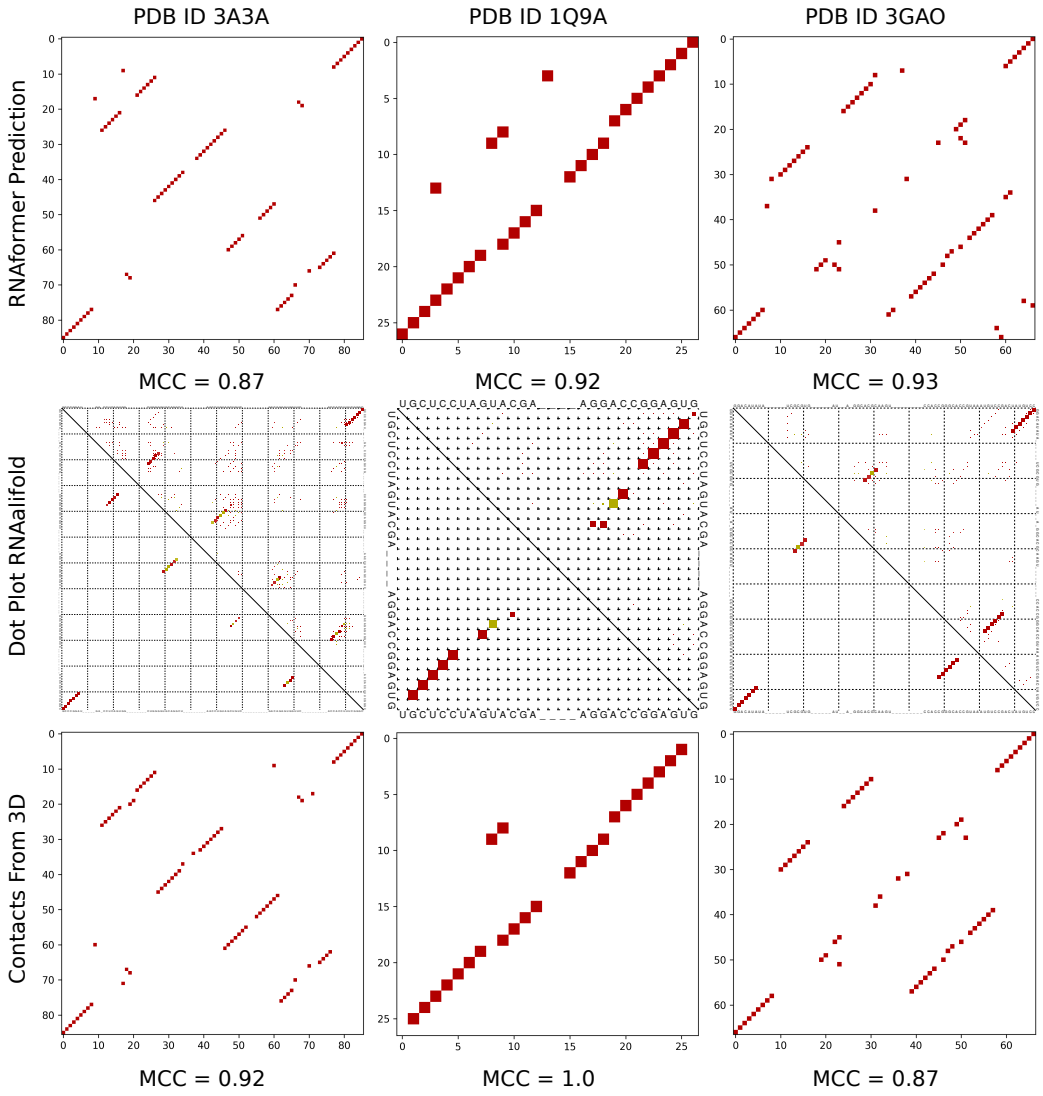

Figure 11: **Synthetic homologs inform a thermodynamics-based consensus structure model.** For PDB IDs 3A3A, 1Q9A, and 3GAO, we show (top to bottom): the nucleotide interaction map predicted by RNAformer, the consensus dot plot obtained by applying RNAalifold to a LocARNA alignment of synthetic homologs (SHS), and the nucleotide interactions extracted from the corresponding AlphaFold 3 predictions using SHS (via DSSR). Across all examples, RNAalifold recovers a consensus structure that closely matches the RNAformer prediction, aside from tertiary and non-canonical interactions that thermodynamic models cannot capture. MCC values denote Matthews Correlation Coefficients compared to the ground-truth. RNAalifold was run on an alignment of 20 SHS, subsampled from the full set of 100 generated sequences, to ensure computational efficiency and clearer consensus dot-plots.