Has AlphaFold3 achieved success for RNAs?

Clément Bernard, Guillaume Postic, Sahar Ghannay, Fariza Tahi

Has AlphaFold3 achieved success for RNAs?

Clément Bernard[1,2], Guillaume Postic[1], Sahar Ghannay[2], and Fariza Tahi[1]

[1]Université Paris Saclay, Univ Evry, IBISC, 91020, Evry-Courcouronnes, France

[2]LISN - CNRS/Université Paris-Saclay, France, 91400, Orsay, France

Predicting the 3D structure of RNA is a significant challenge despite ongoing advancements in the field. Although AlphaFold has successfully addressed this problem for proteins, RNA structure prediction raises difficulties due to fundamental differences between proteins and RNAs, which hinder direct adaptation. The latest release of AlphaFold, AlphaFold 3, has broadened its scope to include multiple different molecules like DNA, ligands and RNA. While the article discusses the results of the last CASP-RNA dataset, the scope of performances and the limitations for RNAs are unclear. In this article, we provide a comprehensive analysis of the performances of AlphaFold 3 in the prediction of RNA 3D structures. Through an extensive benchmark over five different test sets, we discuss the performances and limitations of AlphaFold 3. We also compare its performances with ten existing state-of-the-art *ab initio*, template-based and deep-learning approaches. Our results are freely available on the EvryRNA platform: `https://evryrna.ibisc.univ-evry.fr/evryrna/alphafold3/`.

RNA 3D structure | AlphaFold3 | Deep learning | Benchmark

Correspondence: ***fariza.tahi@univ-evry.fr***

Introduction

Ribonucleic acids (RNAs) are fundamental molecules crucial to cellular activities. While their functions are directly linked to their structures, the prediction of the latter remains an open challenge to be addressed. Knowing the structure of RNA could be of great interest for drug design or the comprehension of biological processes like cancer (1). While experimental methods like X-ray crystallography, NMR, or cryo-EM can determine the RNA 3D structures, their use is costly (in terms of time and resources) and hardly scalable for the number of RNAs in the living. Computational approaches have emerged with *ab initio*, template-based and, more recently, deep learning methods. *Ab initio* methods (2–11) tend to reproduce the physics of the system, with force field applied to a coarse-grained representation (low-resolution where a nucleotide is replaced by some of its atoms). The template-based approaches (12–21) create a mapping between sequences and fragments of structure before refining the assembled structures.

With the recent success of AlphaFold for proteins (22, 23), approaches have been made to replicate its success to RNAs. Directly using protein methods to infer RNA 3D structures is impossible, as RNAs and proteins are chemically and physically different molecules. Current methods try to adapt what exists for proteins to RNAs like DeepFoldRNA (24), RhoFold (25), DrFold (26), NuFold (27), and trRosettaRNA (28). They consider coarse-grained representation and predict Euclidean transformation before reconstructing the full-atom structure. The use of torsional angles is also adapted to RNAs, with either the standard torsional angles (25, 27) or angles from their coarse-grained representations (24, 26).

While being better than existing template-based or *ab initio* methods, deep learning approaches do not solve the prediction of RNA structures yet, as shown in CASP-RNA (29) and in our recent benchmark State-of-the-RNArt (30). Recently, a critical review (31) explains the reasons why the AlphaFold for RNA has not yet happened, and might not arrive for the next decades. However, AlphaFold has released its latest version, named AlphaFold 3 (32), that extends its predictions to different molecules, including RNAs. In this work, we aim to provide a response to (31) in order to know if AlphaFold 3 meets its success for RNAs.

To extend its range of molecules, AlphaFold 3 has made changes in its architecture to better adapt to the variety of available inputs. It no longer relies on torsional angles to prevent the restriction to specific molecules, as was the case in AlphaFold 2 (23). It directly predicts atom coordinates with the use of a multi-cross diffusion model. Through a benchmark on CASP-RNA (32), the authors mentioned good results, but AlphaFold 3 did not outperform human-helped methods. Furthermore, it is not clear what the current limitations are and how well it performs compared to state-of-the-art solutions.

This article aims to provide a comprehensive extension on the evaluation and benchmark of AlphaFold 3 for RNAs. We first describe the main differences between RNAs and proteins to highlight the challenges of RNA 3D structure prediction and we describe the AlphaFold 3 solution before discussing the benchmark we did. Then, we evaluate AlphaFold 3 and comment on the results of AlphaFold 3 and the current limitations of the model. Our benchmark also compares the performances with state-of-the-art solutions to provide a complete comparison. The results and the data are freely available and usable in the EvryRNA platform: `https://evryrna.ibisc.univ-evry.fr/evryrna/alphafold3/`.

RNAs vs proteins

RNAs and proteins are both molecules that play crucial roles in the living. They share the characteristic of having a 3D structure that directly defines their function. However, it is

important to acknowledge that dynamics, transient structures, and the role of unstructured proteins also play a significant role in protein function, making this relationship more complex. This section discusses the differences between RNAs and proteins, highlighting the reasons why adapting existing protein models has been challenging.

RNAs comprise four nucleotides (A, C, G and U), whereas proteins comprise 20 amino acids. This difference has a high consequence on the adaptation of protein algorithms to RNA. The vocabulary available for RNA is limited to four unique elements, making the use of protein vocabulary not directly adaptable. The sequence length of RNA molecules also has a high variability (from a dozen to thousands of nucleotides) compared to proteins (from a dozen to hundreds of amino acids).

A major difference between RNAs and proteins lies in the folding stabilisation. RNA structure is maintained by base pairing and base stacking, while protein structure is supported by hydrogen interactions in the skeleton. The protein backbone is also modelled by torsion angles (Φ and Ψ) for each amino acid because the peptide bond is planar. This is not the case for RNA, where each nucleotide can be described by seven torsion angles (α, β, γ, δ, ϵ, ξ and χ) and the sugar-pucker pseudorotation phase P. An approximation usually involves pseudo-torsion η and θ (33). However, the complexity of the RNA backbone arises not only from the number of torsional degrees of freedom but also from their intricate correlations. Specifically, the structural divergence at the phosphodiester linkage is influenced by the sugar pucker and glycosidic bond orientation of both nucleotides connected to the phosphate group. This interdependence often necessitates describing RNA structure using dinucleotide-like fragments to accurately capture the backbone geometry (34). Protein models, therefore, learn a conformational mechanism fundamentally different from the RNA folding process, where such adaptations and structural dependencies must be carefully accounted for.

The nature of pairwise interactions of RNA 3D molecules differ from those of proteins. The RNA interactions can be made through three different edges of the RNA base: WC edge, Hoogsteen edge and sugar edge (35) shown in Figure 1. In addition, the orientation of the glycosidic bonds gives another property to an interaction: *cis* or *trans*. The combination of edge and orientation gives 12 possibilities of interaction between bases. The standard Watson-Crick (WC) base pair corresponds to the cis WC/WC pairing. Given the orientations (*cis* or *trans*), the edges and the base pairing, there are more than 200 possible base pairs. Only the standard WC pairs (cis WC/WC) of AU and CG (and also GU wobble pair) are used for the 2D structure representation. RNA bases also have common patterns of interactions, where a base stacks on another one. The base-stacking (36, 37) refers to the four base-stacking types from relative orientations (upward, downward, outward and inward) (38). The extended secondary and tertiary interactions (long-range base pairs) play a crucial role in the overall topology of the RNA folding process. They help stabilise the structure and can not be ignored when working on RNA 3D structures.

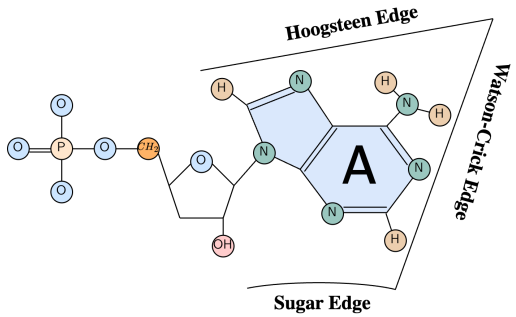


Figure 1. Description of the three different edges of the Adenine RNA nucleotide: Watson-Crick edge, Hoogsteen edge and Sugar edge. The three other nucleotides share similar edges.

The stability of the RNA and protein structures is different. More than five decades ago, the Nobel Prize-winning work of Christian B. Anfinsen established that, under physiological conditions, the protein chain spontaneously folds into its native structure, which is the conformation corresponding to a minimum of the Gibbs free energy that is both a kinetically accessible and thermodynamically stable (39). This native structure of the protein is also characterized by its uniqueness—although it may be altered by dynamic behaviours, such as domain motions, the global fold of the protein remains the same. On the contrary, RNA molecules often have a more rugged Gibbs free energy landscape, thus populating multiple conformational states (40). The switching between these conformations supports some RNA functions, such as riboswitches or ribozymes, and may be driven by environmental changes, such as ions (notably Mg2+), pH, temperature, or ligand binding (41, 42).

There is a huge disparity in the protein and RNA data. Even if there is a higher proportion of RNAs than proteins in the living, this is not reflected in the available data: only a small amount of 3D RNA structures are known. Up to June 2024, 7,759 RNA structures were deposited in the Protein Data Bank (43), compared to 216,212 protein structures. The quality and diversity of data are also different: a huge proportion of RNAs come from the same families. It implies several redundant structures that could prevent a model from being generalized to other families. In addition, a huge amount of RNA families have not yet solved structures in the PDB, usually long RNAs. This means there is no balanced and representative proportion of RNA families through the known structures.

Finally, no standard dataset has been used through the community for RNAs. Each research group uses its dataset with different preprocessing associated. It prevents using deep learning methods, as a lot of work is needed for a clean dataset. While the community agrees to use RNA-Puzzles (44–48) or the newly CASP-RNA (29) to test the generalization of proposed models, no clear training set is available. The first solution has been RNANet (49), developed in our lab to solve this issue. It is a database that uses MySQL and gathers diverse RNA information to train deep learning methods. A new approach, RNA3DB (50), creates indepen-

dent datasets for deep learning approaches where clustering is done based on sequence and structure disparity.

AlphaFold 3

Building on the recent success of AlphaFold 2 (23) in protein structure prediction, AlphaFold 3 (32) expands its predictions to structures from all molecules available in the PDB (43). The authors highlight several differences from the previous architecture, contributing to successful predictions of a wide range of molecules. One key difference is the introduction of a diffusion model that reconstructs coordinates from the residue level to the atom level. AlphaFold 3 also outputs directly the coordinate atom positions compared to the prediction of rotation/translation vectors (and torsional angles) in the previous version. It also weights less the multiple sequence alignment (MSA) in the overall model. In the case of RNA, AlphaFold 3 has been evaluated on the CASP-RNA dataset (29), demonstrating improved predictions compared to RosettaFold2NA (51) and AIchemy_RNA (25) (the best AI-based submission in the competition). Despite these advancements, AlphaFold 3's performance lags behind AIchemy_RNA2 (52) (the top human-expert-aided submission). More details on the architecture, the training procedure and the differences between AlphaFold 2 and AlphaFold 3 are provided in the Supplementary file.

Benchmark

To assess the extent of AlphaFold 3 performances, we have evaluated and compared it with other state-of-the-art methods on five datasets. This section describes the datasets, the methods as well as the metrics used to evaluate AlphaFold 3.

Datasets

To evaluate the prediction of RNA structures, we considered the following five test sets, with the first three from our previous work (30).

- **RNA-Puzzles**: the first dataset is composed of the single-stranded structures from RNA-Puzzles (44–48), a community initiative to benchmark RNA structures. We considered only single-stranded solutions to have a fair comparison between the benchmarked models. It is composed of 22 RNAs of length between 27 and 188 nt (with a mean of 83 nt).

- **CASP-RNA**: the second test set is CASP-RNA (29) structures, which is a collaboration between the CASP team and RNA-Puzzles. It is composed of 12 RNAs with wide-range sequences, from 30 to 720 nt (with a mean of 209 nt).

- **RNASolo**: the third test set is a custom test set composed of independent structures from RNAsolo (53). We downloaded the representative RNA molecules from RNAsolo (53) with a resolution below 4 Å and removed the structures with sequence identity higher than 80%. Then, we considered only the structures with a unique Rfam family ID (54), leading to 25 non-redundant RNA molecules, with a sequence between 45 and 298 nt (and a mean of 100 nt). It can not be ensured that the structures from this dataset were not used in the training set of the different models. We keep this dataset for comparison, as we already have the results for the benchmarked methods.

- **RNA3DB_0**: this dataset is composed of a non-redundant set of structurally and sequentially independent structures from RNA3DB (50). It comprises the component #0, which is composed of orphan structures that are advised to be used as a test set. These structures do not have Rfam families (54) and include synthetic RNAs, small messenger RNAs crystallized as part of larger complexes. After removing structures with sequences below ten nucleotides and sequence identity below 80% (using CD-HIT (55)), we ended up with a dataset of 224 structures from 10 to 339 nt (with a mean of 55 nt). Nonetheless, these structures come from complexes, meaning they do not behave well in isolation, and thus their experimentally observed conformations depend on other chains. To account for this, in evaluating models, we considered 113 structures with their full context, and we predicted the structures with AlphaFold 3 (the other structures have a too large context, and we failed in predicting them using AlphaFold 3). We name this subset RNA3DB_0 (Context).

- **RNA3DB_Long**: the last dataset comprises long RNA structures from RNA3DB (50). We considered structures with a release date after January 2023 to avoid any structure leakage for fair comparison. We considered structures with sequences between 800 nt (800 nt being the limit from previous test sets) and 5,000 nt, as we wanted to study the performances of long RNAs. It leads to 58 structures with a sequence between 828 and 3,619 nt (with a mean of 2005 nt). It comprises 57 ribosomal RNAs and one structure of a Group II intron.

We have also ensured that all the datasets (except RNA-Puzzles and CASP-RNA) have a sequence identity below 80% to have non-redundant structures for robust evaluation.

To comprehend and detail the predictions of AlphaFold 3, we studied in detail three main interactions in the folding of RNA 3D structures: Watson-Crick (WC), non-Watson-Crick (nWC) and stacking (STACK). The proportion of these interactions is presented in Table **??**. All datasets have the same proportion of stacking (around 75%), except for the RNA3DB_0 dataset (around 56%). As RNA3DB_0 contains orphan structures, it implies structures with less common folding, reflected by the lower proportion of stacking interactions. For all the datasets, there is a higher proportion of stacking interactions, followed by Watson-Crick and non-Watson-Crick interactions. The number of non-Watson-Crick interactions ranges from 5 to 10%, meaning that these interactions would be challenging for predictive models as they are rare in the original structure. When comparing

Table 1. Proportion of key RNA interactions for the five test sets (and the subset of RNA3DB with context). The interactions are normalised by the number of residues. Interactions are either stacking (STACK), Watson-Crick (WC) or non-Watson-Crick (nWC), extracted from MC-Annotate (36). Interactions for RNA3DB_0 are computed without context, while RNA3DB_0 (C) includes the context.

Interaction type	STACK	WC	nWC
RNA-Puzzles	0.78	0.33	0.10
CASP-RNA	0.75	0.35	0.05
RNASolo	0.77	0.31	0.09
RNA3DB_0	0.56	0.14	0.04
RNA3DB_0 (C)	0.61	0.16	0.04
RNA3DB_Long	0.74	0.29	0.10

RNA3DB_0 with or without the context, we observe a greater proportion of stacking and Watson-Crick interactions in the presence of context. However, the number of non-Watson-Crick interactions remains unchanged.

State-of-the-art methods

Existing solutions for the prediction of RNA 3D structures are based on three main types of methods: *ab initio*, template-based and deep-learning ones. As discussed previously in our work (30), *ab initio* methods (2, 4, 6) integrate the physics of the system by usually simplifying the representation of nucleotide (coarse-grained). Instead of using all the atoms for one nucleotide, they create a low-resolution representation that simplifies the computation time while losing information. They use approaches like molecular dynamics (56) or Monte Carlo (57) to perform sampling in the conformational space and use a force field to simulate real environment conditions. On the other hand, template-based methods (12, 16–18, 58) create a mapping between sequences and known motifs with, for instance, secondary structure trees (SSEs) before reconstructing the full structure from its subfragments. Finally, the recent methods tend to incorporate deep learning methods (24–28) by using attention-based architectures with self-distillation and recycling as done in AlphaFold 2 (23).

To compare the performances of AlphaFold 3 (32), we benchmarked ten approaches, the ones used in our previous work (30). For the *ab initio* methods, we benchmarked SimRNA (6), IsRNA1 (4) and RNAJP (2). Only RNAJP was used locally. For the template-based approaches, we benchmarked MC-Sym (58), Vfold3D (18), RNAComposer (17), 3dRNA (16) and Vfold-Pipeline (12). For the deep learning methods, we benchmark trRosettaRNA (28) and RhoFold (25). More details on each method are provided in our previous article (30). For RNA-Puzzles and CASP-RNA, we included in the benchmark the predictions from the official results of the competitions. We refer to them as "Challenge best" and correspond to different methods for each RNA. We normalised each prediction using RNA-tools (59) to have a standard format for all structures. It gives standardised names for chains, residues and atoms and removes ions and water.

We used the web servers with default parameters to compare available models fairly, where each user could reproduce our experiments. As we made most of the predictions using web servers, the predictions on RNA3DB_0 were hardly applicable to all the methods. Therefore, we benchmarked the RNA3DB_0 dataset with one method per approach (the quickest method per approach): RhoFold (25) for deep learning, RNAComposer (17) for template-based and RNAJP (2) for *ab initio*. For the RNA3DB_Long dataset, only AlphaFold 3 could predict structures with sequences up to 3000 nt. For this dataset, we only considered the predictions from AlphaFold 3.

Evaluation metrics

To compare the predictions, we used the RNAdvisor tool (60) developed by our team, which enables the computation of a wide range of existing metrics in one command line. For the evaluation of RNA 3D structures, a general assessment of the folding of the structure can be done with either the root-mean-square deviation (RMSD) or its extension adding RNA features ϵRMSD (61). Protein-inspired metrics can also be adapted to assess structure quality like the TM-score (62, 63) of the GDT-TS (64) (counts the number of superimposed atoms). There are also the CAD-score (65) (which measures the structural similarity in a contact-area difference-based function) and the lDDT (66) (which assesses the interatomic distance differences between a reference structure and a predicted one). Finally, RNA-specific metrics have been developed, like the P-VALUE (67) (which assesses the non-randomness of a given prediction). The INF-ALL (38) and DI (38) have been developed to consider RNA-specific interactions. The INF score incorporates canonical and non-canonical pairing with Watson-Crick (INF-WC), non-Watson-Crick (INF-NWC), and stacking (INF-STACK) interactions. The consideration of torsional angles has been developed with the mean of circular quantities (MCQ) (68) and LCS-TA (longest continuous segments in torsion angle space) (69). As discussed in (60), all these metrics are complementary and can infer different aspects of RNA 3D structure behaviour. For the rest of the article, we will discuss the RMSD, INF-ALL, lDDT, TM-score and MCQ and let the results on the other metrics in the Supplementary file. Indeed, the RMSD is the most used metric in the literature, and the INF-ALL incorporates key RNA interactions. The lDDT and TM-score allow for evaluating global conformations (widely used in AlphaFold 3), and MCQ gives the torsional deviation. We only mention all the metrics when comparing the different models to ensure a complete evaluation.

Results

This section presents the results of AlphaFold 3 predictions on the discussed test sets. We start by comparing the results of AlphaFold 3 with existing solutions and then discuss in detail the link between the performances relative to sequence length. Next, we discuss the results of AlphaFold 3 on ribosomal structures (RNA3DB_Long dataset) and orphan structures (RNA3DB_0 dataset). Then, we discuss the results of specific RNA key interactions in detail before shedding light on the computation time.

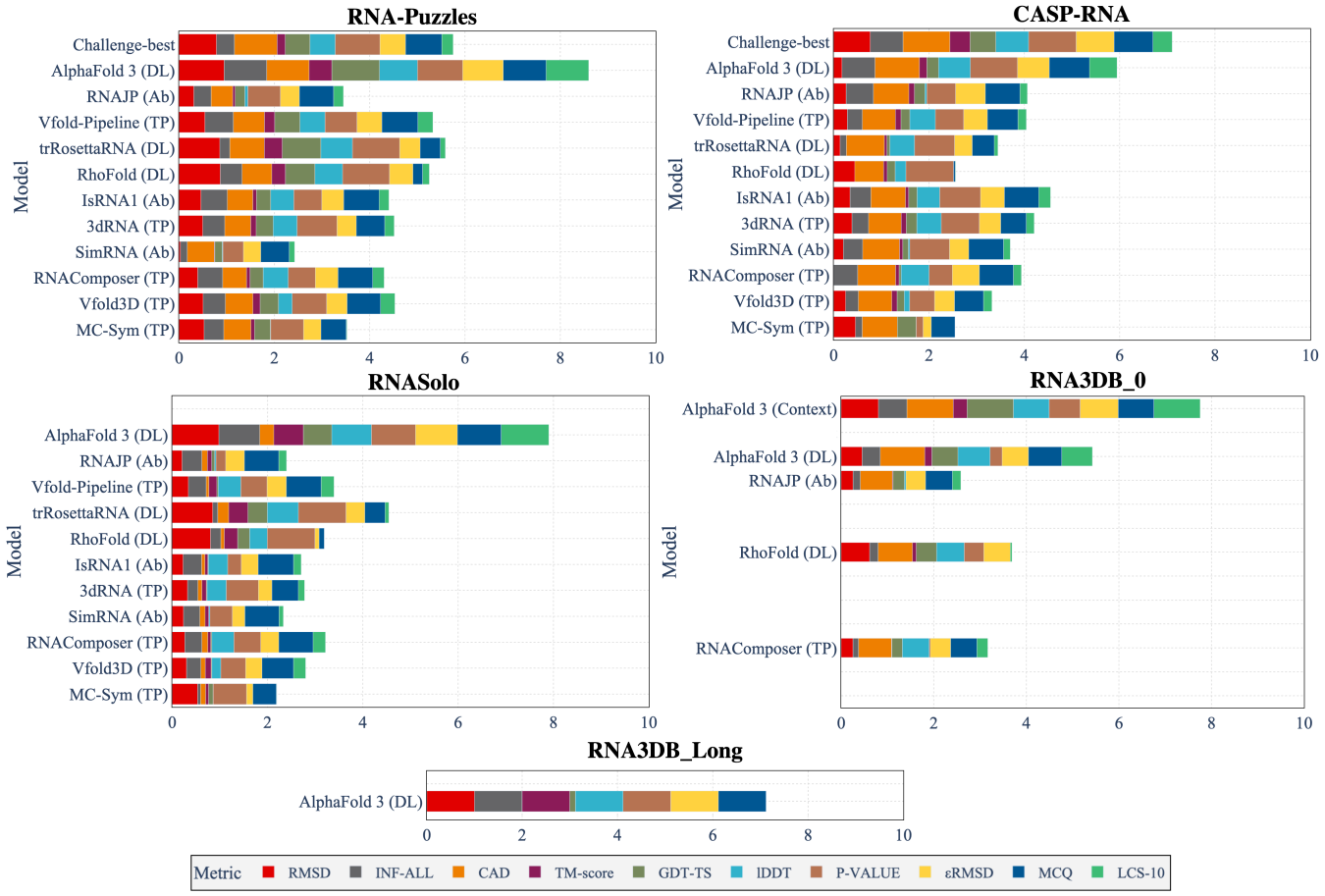


Figure 2. Cumulative normalised metrics (the higher, the better) for each of the benchmarked methods for our five test sets. Each metric is normalised by the maximum value over the five test sets, and the decreased metrics are inversed to have better values close to 1. The *Challenge-best* means the best solutions from the RNA-Puzzles and CASP-RNA competitions (and corresponds to different solutions for each challenge). Types of methods are also mentioned with the names: DL for deep learning, TP for template-based, and Ab for *ab initio*. Methods are sorted by release time (except for *Challenge-best*). AlphaFold 3 (Context) stands for the predictions of AlphaFold 3 for 113 structures of RNA3DB_0 dataset with the context of the structures added as inputs.

AlphaFold 3 compared to the state-of-the-art

We compare the predictions of ten existing methods presented above and AlphaFold 3 on our different test sets. Figure 2 presents the different normalised metrics computed for the prediction of the different models over the five test sets. We included all metrics to show the cumulative performances. The RNA3DB_Long dataset has only predictions from AlphaFold 3, which is the only method capable of processing long sequences. All the metrics are normalised by the maximum values and converted to be better where near to 1 and worst when near to 0. Real values for each metric for the five test sets are reported in Table S1, S2, S3, S4 and S5 of the Supplementary file.

The best models from the CASP-RNA competitions, which are human-guided, outperform AlphaFold 3 (p-value=0.007; Wilcoxon signed-rank test) for every metric (except for LCS-TA with a threshold of 10° and MCQ) for the CASP-RNA dataset. On the other hand, AlphaFold 3 shows a cumulative sum of metrics greater than the other methods for the other test sets (p-value$< 10^{-5}$ for RNA-Puzzles, p-value$< 10^{-4}$ for RNASolo). For RNA-Puzzles, the challenge-best solutions are from older solutions, with less advanced architectures compared to the more recent CASP-RNA solutions. For the RNA3DB_0 dataset, AlphaFold 3 performances are slightly better compared to RhoFold, which has a better RMSD but a worst MCQ and LCS-TA. AlphaFold 3 always has a high MCQ value, indicating it returns structures which are more physically plausible than *ab initio* methods (that use physics properties in their predictions). Nonetheless, it does not always have the best RMSD (outperformed in CASP-RNA and RNA3DB_0), suggesting that AlphaFold 3 does not always have the best alignment (in terms of all atoms) compared to the reference structure.

To compare the global performances of each type of approach, we report in Figure 3 the averaged metrics over the different types of approaches depending on the sequence length. We grouped the results for structures with a sequence length window of 25 nt (each point represents the mean computed on the best results per approach with sequence length from this 25 nucleotide window). Results on the other metrics are shown in Figure S2 in the Supplementary file. None of the benchmarked *ab initio* methods successfully predicted structures for sequences exceeding 200 nt, particularly when using web servers. The best results from CASP-RNA and RNA-Puzzles challenges outperform AlphaFold 3 across most metrics, except for sequences between 150 and 250 nt,

where AlphaFold 3 showed comparable results. Values for RMSD, TM-score, MCQ, and lDDT tend to worsen with sequence length, reflecting a general trend of loss of accuracy with longer RNA structures. For INF, there is no clear degradation tendency, meaning that the reproduction of the interactions does not have a strong link to the sequence length. *Ab initio* and template-based methods have competitive MCQ values, while *ab initio* methods tend to have global alignment worse than the other methods (due to high simulation time, which is a bottleneck for web server usage). Deep learning approaches, in particular, produced worse MCQ scores than traditional methods. AlphaFold 3 demonstrated especially strong MCQ performance, with comparative results for the best solutions of challenges for sequences greater than 250 nt.

These results suggest that AlphaFold 3 achieves competitive performance, particularly in capturing more realistic torsional angles through better MCQ scores (which is not the case for other existing deep learning methods), although it remains outperformed by global assessment for structures with more than 200 nt.

The performance of AlphaFold 3 relative to sequence length.

As seen previously, the prediction of RNA 3D structures usually becomes harder when the sequence length increases. Indeed, the *ab initio* methods fail to predict long interactions as the computation time highly increases with the sequence length. The template-based approaches are limited by the small number of long RNAs, as well as the deep learning methods, as shown in (30).
To observe more in detail the relation between sequence length and AlphaFold 3 performances, we report in Figure 4 the RMSD, MCQ, TM-score, lDDT and INF-ALL metrics depending on sequence length (for the five test sets). Links between the other metrics and the sequence length are available in Figure S3 of the Supplementary file.

Figure 4 indicates that, except for the RNA3DB_0 dataset, the RMSD becomes worse for sequences between 0 and 1000 nt. For the RNA3DB_Long dataset with sequences higher than 1000 nt, the predictions have good results for every metric. We also observe a tendency for degradation in the lDDT, TM-score, and INF-ALL (smaller decrease) when the structures have sequences higher than 100 nt (and below 1000 nt). For every metric, the predictions for the RNA3DB_0 (with or without context) dataset seem to have no clear dependence on the sequence length. For the other test sets with structures with sequences between 200 and 1000 nt, there is a common tendency to worsen in terms of performances for the AlphaFold 3 predictions.

AlphaFold 3 results on long RNAs

Current methods for the prediction of RNA 3D structures are limited for long RNAs and hardly predict structures with sequences longer than 200 nt. AlphaFold 3 is, to the best of our knowledge, the only method that can predict long RNAs (with sequences higher than 1000 nt). Its predictions on RNA3DB_Long are of good performance, as shown in Figure 2. The only metrics where the results are not good are the GDT-TS, the CAD-score and the LCS-TA (threshold of 10°), which might be due to an error in computation. For the LCS-TA, the small score could be explained by the difficulty of keeping a low MCQ for a high proportion of the structure, as the sequences are long for this dataset.

The good results for long RNAs can be explained by the types of structures used in RNA3DB_Long. Indeed, all of the structures (except for one) are ribosomal RNAs, and thus have a lot of redundancy. This might be reflected in the PDB, which has been memorised by AlphaFold 3 during its training. As AlphaFold 3 uses the MSA as inputs, it could find similarities with trained structures and thus return excellent predictions if the families are well known. Most of the long RNAs in the PDB share common structures in the ribosomal family. Therefore, these results show a good generalisation of already-seen families from AlphaFold 3.

We report the two worst predictions of AlphaFold 3 on the RNA3DB_Long dataset in Figure 5. The two worst predictions for the other test sets are provided in Figure S4 of the Supplementary file. The RMSD for the two structures is relatively high (more than 19Å). The second worst structure has a high TM-score (0.74), meaning that even for a long structure (1487 nt), the global alignment of atoms is well predicted. The INF-ALL is also high for these structures (higher than 0.68), meaning it returns a high proportion of key RNA interactions. In details, it is most likely no coincidence that the worst prediction (TM-score=0.38) corresponds to the only non-ribosomal RNA of the RNA3DB_Long dataset, while the overwhelming majority of available native structures for long RNA sequences belong to ribosomes. In addition, the lack of structural context did not help AlphaFold 3 either, as this Group II intron RNA can be found in complex with its large maturase/reverse transcriptase (PDB entry: 8FLI (71)). The medium-to-high quality of the second worst prediction (TM-score=0.74) can be explained by the fact that it occurred for the 15S mitochondrial ribosomal RNA (PDB entry: 8OM4 (72)). This RNA is analogous, yet evolutionarily distant, from its bacterial and eukaryotic counterparts (the 16S and 18S RNAs, respectively) and its 3D structure has rarely been studied, as it was reported in only three articles (73–75).

AlphaFold 3 results on orphan structures

The RNA3DB_0 dataset is mainly composed of structures without any hit in the Rfam family, which constitutes orphan structures. The results of AlphaFold 3 for this dataset, as presented in Figure 2 and Figure 4, show overall lower performances compared to the other datasets when there is no use of context. AlphaFold 3 performs slightly better than RhoFold for this dataset (p-value=0.015). When having the context, AlphaFold 3 produces improved results compared to without context (p-value$< 10^{-19}$).

We detail the two worst predictions for RNA3DB_0 and RNA3DB_0 (Context) from AlphaFold 3, in Figure 5. We observe bad results in terms of metrics (high RMSD and

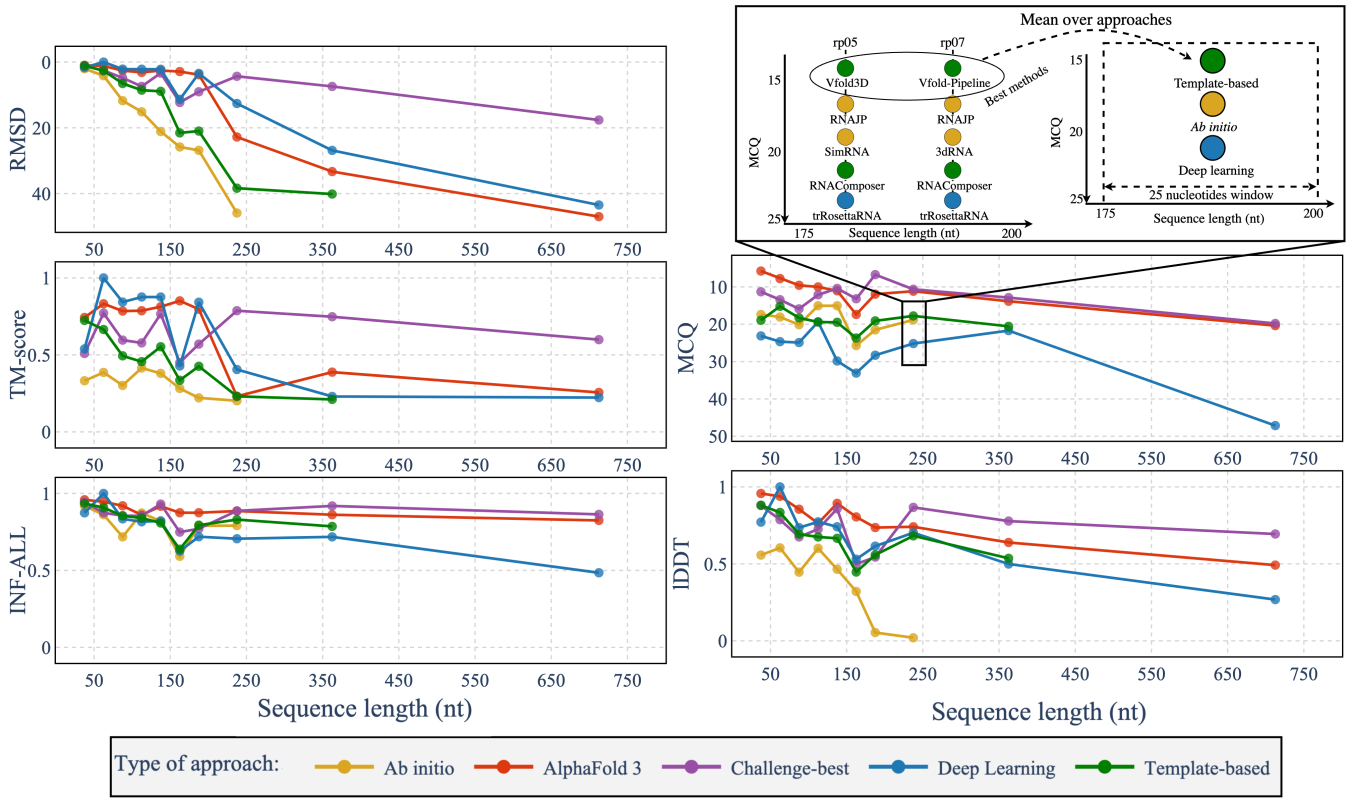


Figure 3. Averaged metrics depending on the sequence length for the different approaches (AlphaFold 3, *ab initio*, deep-learning, template-based and *challenge-best*). Each point represents the metric averaged over the best models of each approach for a window of 25 nt, from 25 to 750 nt. *Ab initio* methods group RNAJP, IsRNA1 and SimRNA while template-based methods group Vfold-Pipeline, 3dRNA, RNAComposer, Vfold3D and MC-Sym. Deep learning methods group trRosettaRNA and RhoFold. Metrics are computed for the RNA-Puzzles, CASP-RNA and RNASolo datasets. The *Challenge-best* corresponds to the best results from either RNA-Puzzles or CASP-RNA competitions but does not appear for the RNASolo dataset. The metrics are RMSD, MCQ, TM-score, lDDT and INF-ALL. RMSD and MCQ are reversed to have the best values near the top and the worst values at the bottom.

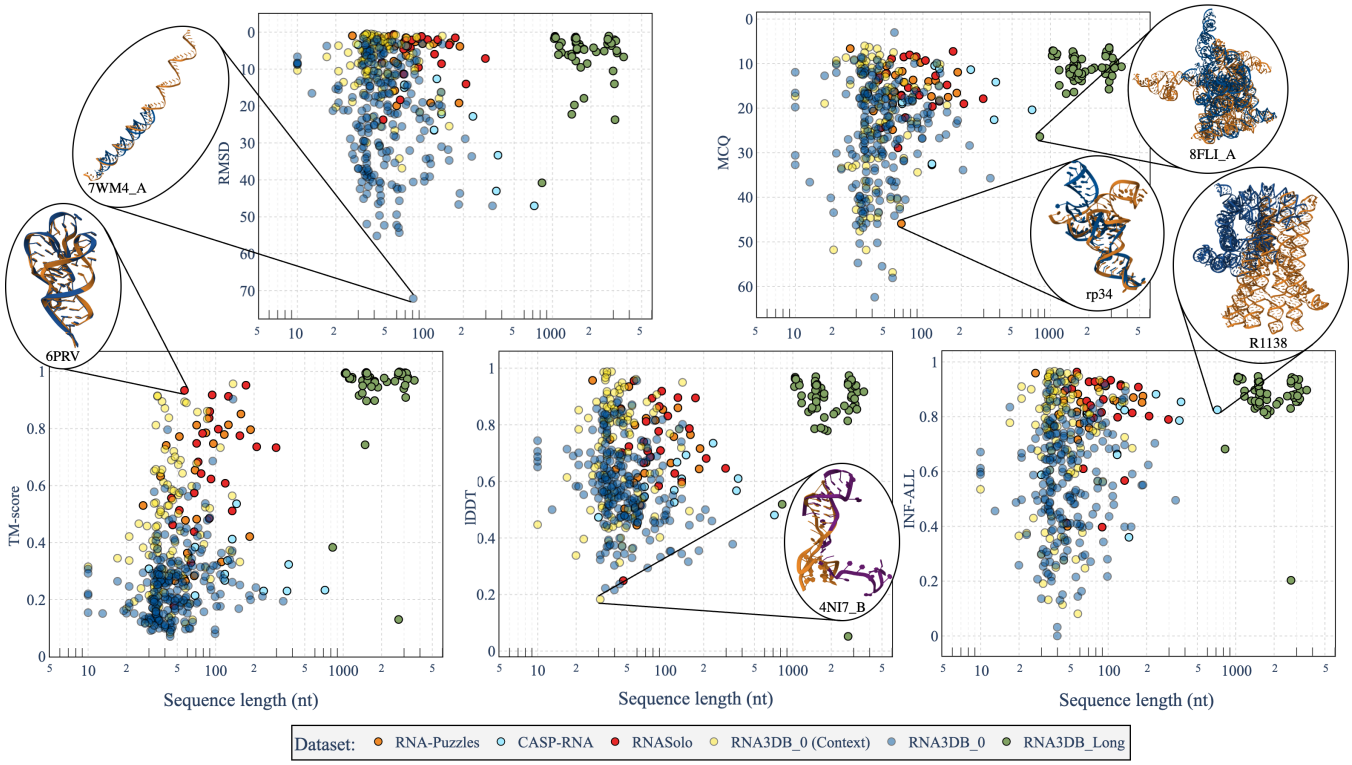


Figure 4. Dependence of metrics with the sequence length on the prediction of AlphaFold 3 (32) on the five test sets. For some of the predictions, we show the predicted structure (in blue or purple if predicted using context) aligned with the native one (in orange) using US-Align (70). Metrics are RMSD, MCQ (68), TM-score (62, 70), lDDT (66) and INF-ALL (38). RMSD and MCQ are reversed to have the best values near the top and the worst values at the bottom.

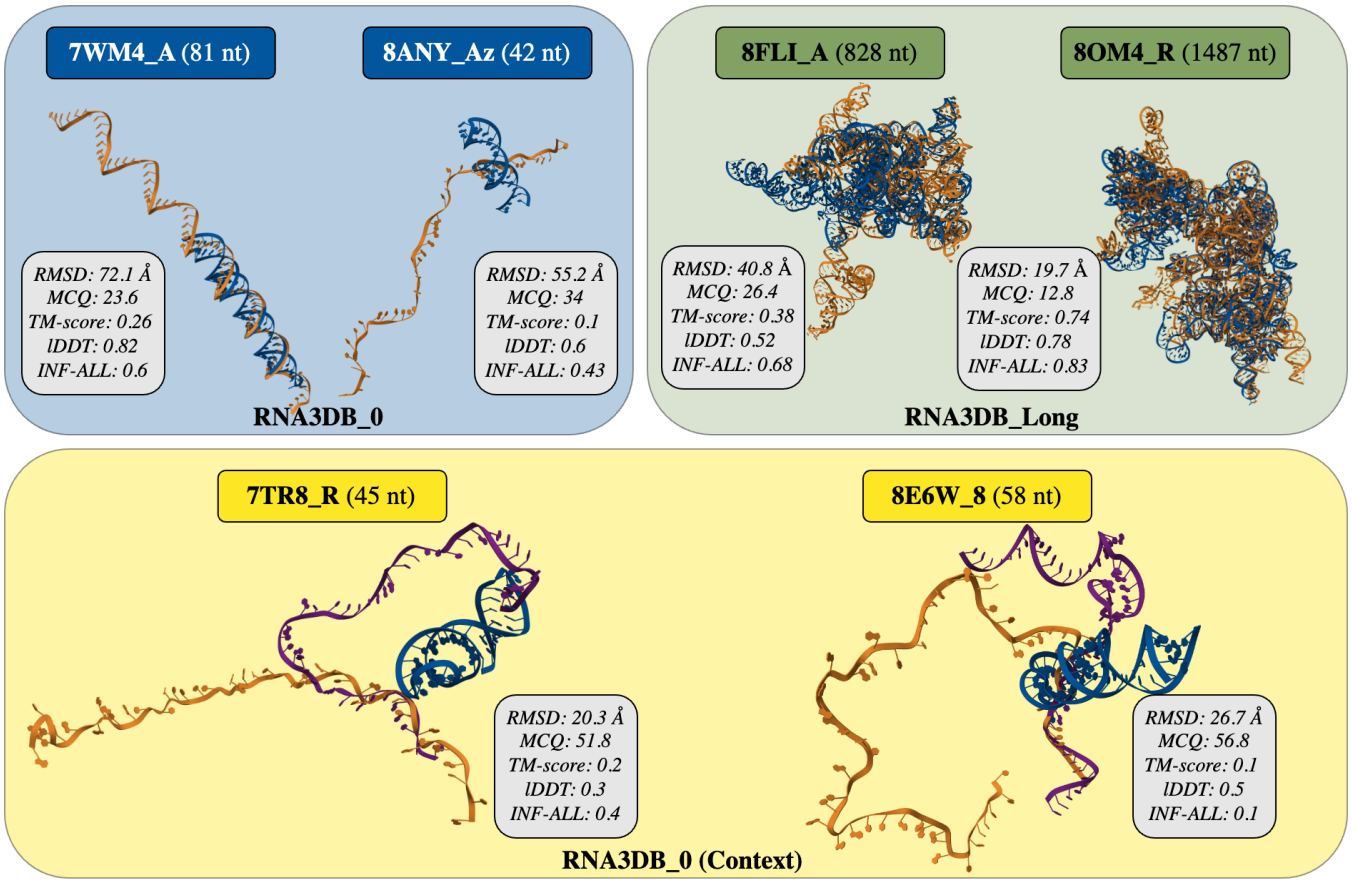


Figure 5. Worst two predicted structures (based on cumulative sum of metrics) from AlphaFold 3 (32) for RNA3DB_0 (left), RNA3DB_Long (right) and RNA3DB_0 (Context) (bottom) datasets. The RMSD, MCQ (68), TM-score (62, 70), lDDT (66) and INF-ALL(38) are provided for each structure. The predictions from AlphaFold 3 (in blue) are aligned with the native ones (in orange) using US-align (70). The predictions of AlphaFold 3 with context (only for RNA3DB_0 (Context)) are provided in purple.

MCQ values and low TM-score and INF-ALL) for the two structures without context. With the context, AlphaFold 3 seems to understand that the predictions are not only helices but still fail in these two worst examples to predict the complex non-common folding of these RNAs. These structures also have a small number of nucleotides (81nt, 42nt, 45nt and 58nt), meaning that AlphaFold 3 might not fail because of the long-range interactions. Instead, these structures do not have a known family and rely on a complex environment of other molecules. With the context, AlphaFold 3 has a better chance to predict well the structural folding, but the generalisation is not always robust to structures without known families, even with small structures (as shown by the mean value of TM-score which is less than 0.5 in Table S4).

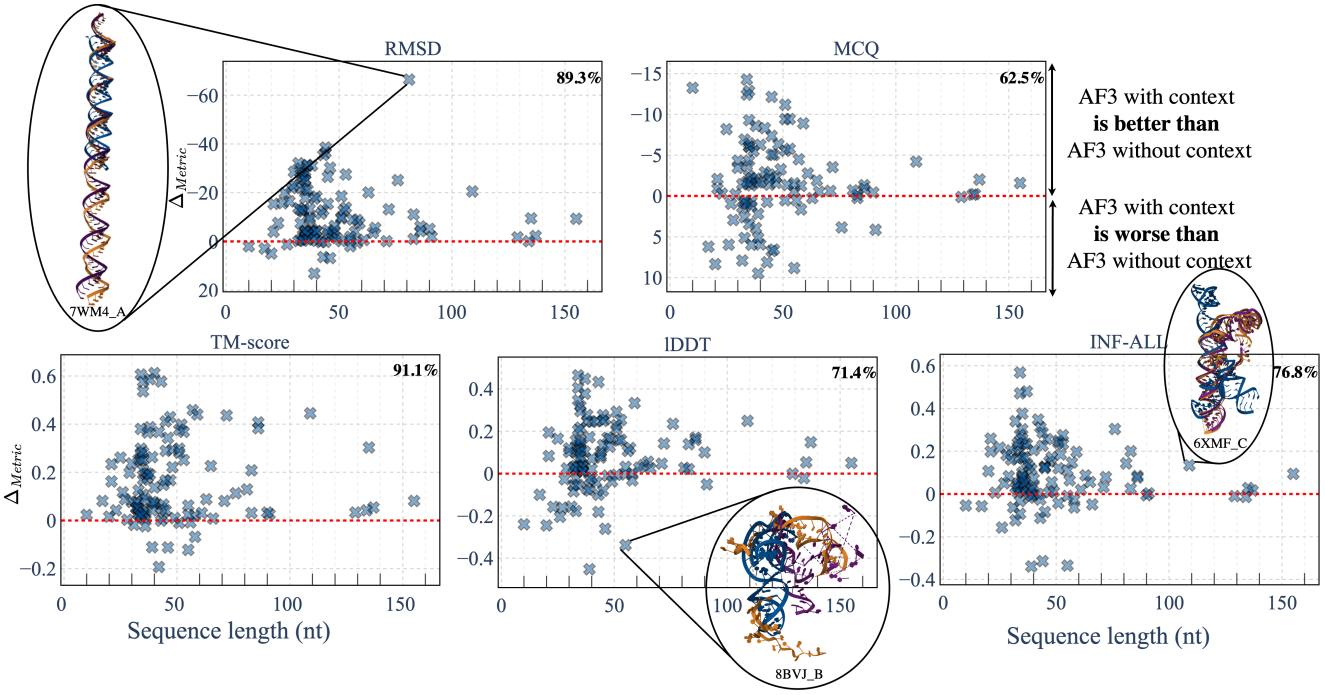
To further study the impact of the context for the prediction of RNA structures, we report in Figure 6 the differences per metric between predictions of AlphaFold 3 with and without context depending on the sequence length. Details of each metric value for each RNA are provided in Figure S7 of the Supplementary file. For all metrics, there is an improvement in using the context: 91.1% of structures with context have better TM-Score than those without context. For the MCQ metric, 62.5% of structures with context outperform those without context, which is less dominant than for the other metrics. For example, in the case of structure 7WM4 (76), the context effectively facilitates identifying the correct scale for one half of the double helix. Similarly, for structure 8BVJ (77), which features a discontinuity, the context enables AlphaFold 3 to accurately detect the discontinuities. However, this does not result in better alignment in terms of the lDDT metric.

Incorporating contextual information significantly enhances global alignment performance, as reflected by improvements in RMSD, TM-Score, and lDDT metrics. This is followed by moderately smaller, but still notable, improvements in reproducing key RNA interactions (INF metric) and torsional angles (MCQ metric). Among the benchmarked models, the possibility to use the context for the prediction is only available with AlphaFold 3. The other models are specialised on RNA and are not designed to process different molecules.

AlphaFold 3 results on key RNA interactions

To evaluate the ability of AlphaFold 3 to predict non-canonical interactions, we depict the scatter plots between non-Watson-Crick INF (INF-NWC) and Watson-Crick INF (INF-WC) in Table 7. The size of the points is proportional to the RMSD of structures and, thus, to their global atoms alignment. We observe a tendency to have a low RMSD (small points) whenever the INF-WC and INF-NWC are high. There are also many structures with an INF-NWC of 0, suggesting that AlphaFold 3 does not predict any of the non-Watson-Crick interactions (mostly for the RNA3DB dataset).

Figure 6. Difference per metric between results from AlphaFold 3 with context and without context for the common structures of the RNA3DB_0 dataset depending on the sequence length. Regions higher than the red line correspond to structures where the results from AlphaFold 3 with context are better than those without the context. We reversed the RMSD and MCQ metrics so that higher regions always depict the same behavior. The percentage of cases where AlphaFold 3 with context outperforms predictions without context is reported in the top right corner of each plot. Structures are reported with the native one in orange, predictions with AlphaFold 3 without context in blue and with context in purple.



Examples of successful and missing non-Watson-Crick interactions are shown in the Figure. For the results on stacking interactions, there are predictions where AlphaFold 3 does not predict the Watson-Crick interactions well but still predicts the stacking ones. It can be explained by good skeleton predictions while lacking the base conformations that produce the WC interactions. Secondly, there is an increased tendency between the INF-STACK and INF-WC: when AlphaFold 3 predicts the WC interactions well, it also tends to estimate the stacking well. Indeed, the stacking interactions tend to align with the correct base pairing, but the correlation is likely influenced by whether the sequence can fold into the observed conformation. For instance, in Figure 7, parts of 8ex9_B can fold, whereas others cannot.

To compare the key RNA interactions predicted from AlphaFold 3 with existing solutions, we present in Figure 8 the mean INF metrics (INF-WC, INF-NWC and INF-STACK) over RNA-Puzzles, CASP-RNA and RNASolo for the ten benchmarked models. Details for each dataset are provided in Table S6 of the Supplementary file. We show the results only on these datasets as we had complete predictions for each model only for these three datasets. AlphaFold 3 has better values for each INF metric compared to the other methods. The second best method to reproduce RNA key interactions is RNAComposer. While having good overall results in terms of cumulative metrics, trRosettaRNA shows bad results in terms of key RNA interactions. Even if AlphaFold 3 outperforms other solutions for all the INF metrics, the results for nWC interactions remain low (below 0.5), meaning there is still progress to reproduce RNA-specific interactions well.

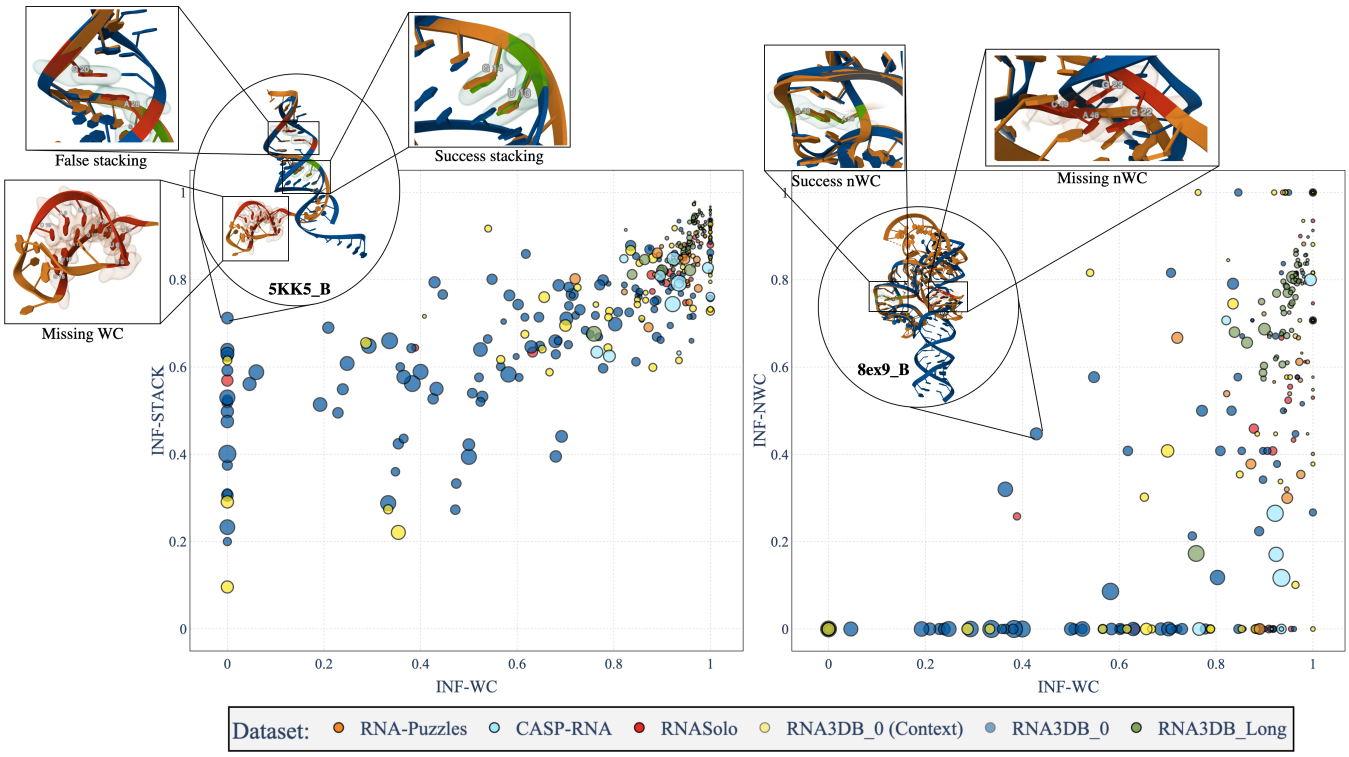
Computation time

AlphaFold 3 is a deep learning method that has a complex architecture. Compared to existing *ab initio* methods, deep learning methods tend to be faster for inference. We report the computation time for a small RNA molecule (27 nt) and a long RNA (434 nt) for RNAComposer (17), RhoFold (25), trRosettaRNA (28), RNAJP (2) and AlphaFold 3 (32) in Table 2. We report the computation time of the fastest methods, while the time for the rest of the methods is available in our previous work (30). As we could only run RNAJP locally and each web server has different configurations, there is a bias in the comparison. RNAComposer, RhoFold and trRosettaRNA all predict small RNA (in less than a minute) very quickly, while RNAJP takes two hours (with default parameters). For a structure with a longer sequence, this is RNAComposer that has the fastest computation time (around three minutes). The *ab initio* method, RNAJP, takes 15 hours. AlphaFold 3 returns prediction in around five minutes, which shows fast inference. For RNA of very long sequences (around 3000 nt), AlphaFold 3 take multiple hours to predict (and sometimes returns errors and needs to be run multiple times to get results).

Discussion

AlphaFold 3 is a deep learning method that has widened its scope to predict RNA structures (as well as other molecules) compared to its previous approach. Through our benchmark,

Figure 7. Link between INF Watson-Crick (WC) and non-Watson-Crick (nWC) and stacking (STACK) interactions for the predictions of AlphaFold 3 for our five test sets. The area of each point is proportional to the RMSD: the lower, the better. Only structures with at least one non-Watson-Crick interaction are shown in the figures. An INF (38) value of 1 means accurate reproduction of key RNA interactions, while a value near 0 means the structure does not reproduce the interactions. Left: INF non-Watson-Crick (INF-nWC) depending on INF Watson-Crick (INF-WC) interactions. Right: INF stacking (INF-STACK) depending on INF Watson-Crick (INF-WC) interactions.



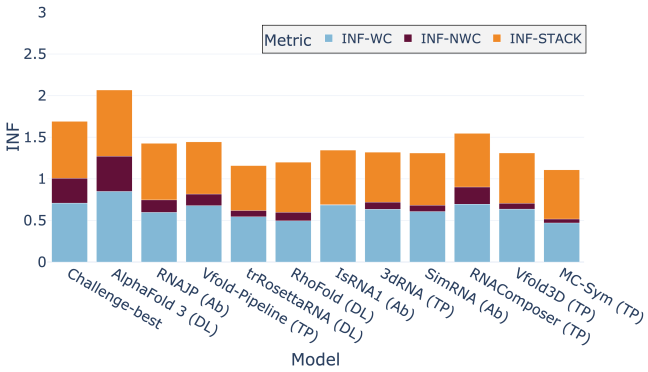


Figure 8. INF metrics for the different benchmarked models averaged over three test sets: RNA-Puzzles, CASP-RNA and RNASolo. INF metrics consider Watson-Crick (INF-WC), non-Watson-Crick (INF-NWC), and stacking (INF-STACK) interac-

Table 2. Computation time for a sequence of 27 nt (PDB ID: 6Y0Y) and 434 nt (PDB ID: 7XD6 (78)). Computation time is computed using web servers except for RNAJP. Methods are sorted by release time. The types of approaches are either template-based (TP), *ab initio* (Ab) or deep learning (DL).

Model	Approach	Time (27nt)	Time (434nt)
RNAComposer (17)	TP	**1**	**3**
RhoFold (25)	DL	**1**	10
trRosettaRNA (28)	DL	**1**	600
RNAJP* (2)	Ab	120	900
AlphaFold 3 (32)	DL	2	5

* RNAJP computation time is computed locally with a simulation time set to 50 × 10 6 steps on an NVIDIA P1000.

we showed that AlphaFold 3 is a competitive method that outperforms most of the existing solutions. It yields better results for the CASP-RNA challenge and RNASolo, but remains outperformed by the best solutions from CASP-RNA challenge.

AlphaFold 3 has achieved good generalisation properties for the ribosomal structures (RNA3DB_Long dataset). This shows bias from existing data for RNA: most of the long RNA available in the PDB is ribosomal-related RNA.

AlphaFold 3 returns results with an overall good reproduction of RNA key interactions compared to existing solutions. It is also the best method to reproduce RNA torsional angles (best results in terms of MCQ), which was lacking in the existing deep learning methods (30).

There remain limitations that need to be addressed regarding the RNA folding problem. AlphaFold 3 does not reproduce all the non-Watson-Crick interactions, which is essential for the stability of RNA 3D structures. Furthermore, AlphaFold 3 fail to predict structures from orphan families (RNA3DB_0 dataset) without the context. These structures are hard to predict as there is no hint in the available data, and reliable information is often supported by the context and the environment of the RNA. AlphaFold 3 achieves better results when having the context, but there remains a limitation of generalisation for these orphan RNAs for our evaluation. Evaluating orphan structures remains challenging, as environmental information or context is lacking. There is also no easy way to correctly evaluate the alternative solutions pro-

posed by AlphaFold 3, whereas multiple conformations are possible for RNAs. AlphaFold 3, while reducing the impact of MSA on its architecture, still uses it, restricting its scope for RNA (as there are still unknown families). The computation time for the inference is very fast but remains limited by its usage in web servers. The source code has been released but requires huge computational resources to be easily used.

Conclusion

AlphaFold 2 had a huge success in the prediction of protein folding and has changed the field by the quality of its predictions. The new release of AlphaFold, named AlphaFold 3, has extended the model to predict all molecules from the PDB, like ions, ligands, DNA or RNA.

Through an extensive benchmark on five different test sets, we have evaluated the quality of predictions of AlphaFold 3 for RNA molecules. We have also compared the results with ten existing methods, which are easily reproducible as their predictions are available using web servers.

Our results show that AlphaFold 3 is of competitive quality, as it outperforms most of the existing solutions. It returns more physically plausible structures than *ab initio* methods. It outclasses existing deep-learning approaches for every dataset while better reproducing key RNA interactions and torsional angles. It also returns predictions very quickly compared to *ab initio* or current template-based approaches (but does not exceed RNAComposer (17) for inference time).

For ribosomal long RNAs, AlphaFold 3 returns highly accurate predictions. It could be explained by its capability to generate structures from known families, which has been seen in its training data. As there is not a lot of data available, it is difficult to find complex structures without any homologs to evaluate performances fairly.

Nonetheless, AlphaFold 3 has not yet reached RNAs with the same success as proteins. Its new architecture allows the prediction of wide molecules but remains limited and hardly predicts non-Watson-Crick interactions. It does not generalize well on orphan structures, which are not related to any RNA-known families. Predictions of these structures require the knowledge of the context, which is possible to integrate with AlphaFold 3.

The prediction of atom coordinates instead of base frames, as done in AlphaFold 2, allows the extension of predictions for a wide range of molecules but prevents the generalisation of RNA-specific interactions. The lack of data is also a limitation that prevents the robustness of deep learning methods in general, and so is AlphaFold 3.

ACKNOWLEDGEMENTS

This work is supported in part by UDOPIA-ANR-20-THIA-0013 and performed using HPC resources from GENCI/IDRIS (grant AD011014250). It was also partially supported by Labex DigiCosme (project ANR11LABEX0045DIGICOSME), operated by ANR as part of the program "Investissement d'Avenir" Idex ParisSaclay (ANR11IDEX000302).

Bibliography

1. Y. Zhu, L. Zhu, X. Wang, and H. Jin. RNA-based Therapeutics: An Overview and Prospectus. *Cell Death & Disease*, 13:644, 2022. doi: 10.1038/s41419-022-05075-2.
2. L. Jun and C. Shi-Jie. RNAJP: enhanced RNA 3D structure predictions with non-canonical interactions and global topology sampling. *Nucleic Acids Research*, 51:3341–3356, 4 2023. ISSN 0305-1048. doi: 10.1093/nar/gkad122.
3. A. Krokhotin, K. Houlihan, and N. V. Dokholyan. iFoldRNA v2: folding RNA with constraints. *Bioinformatics*, 31:2891–2893, 9 2015. ISSN 1367-4803. doi: 10.1093/bioinformatics/btv221.
4. D Zhang, J Li, and S-J. Chen. IsRNA1: De Novo Prediction and Blind Screening of RNA 3D Structures. *Journal of Chemical Theory and Computation*, 17:1842–1857, 3 2021. ISSN 1549-9618. doi: 10.1021/acs.jctc.0c01148.
5. D. Zhang, S-J. Chen, and R. Zhou. Modeling Noncanonical RNA Base Pairs by a Coarse-Grained IsRNA2 Model. *The journal of physical chemistry. B*, 2021.
6. M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, and J. M. Bujnicki. SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Research*, 44:e63–e63, 4 2016. ISSN 0305-1048. doi: 10.1093/nar/gkv1479.
7. T. Cragnolini, Y. Laurin, P. Derreumaux, and S. Pasquali. Coarse-Grained HiRE-RNA Model for ab Initio RNA Folding beyond Simple Molecules, Including Noncanonical and Multiple Base Pairings. *Journal of Chemical Theory and Computation*, 11(7):3510–3522, 2015. doi: 10.1021/acs.jctc.5b00200. PMID: 26575783.
8. P. Kerpedjiev, C. H. Zu Siederdissen, and I. L. Hofacker. Predicting RNA 3D Structure Using a Coarse-Grain Helix-Centered Model. *RNA*, 21:1110–1121, 2015. doi: 10.1261/rna.047522.114.
9. P. Sulc, F. Romano, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis. A Nucleotide-Level Coarse-Grained Model of RNA. *J. Chem. Phys.*, 140:235102, 2014. doi: 10.1063/1.4881424.
10. M. A. Jonikas, R. J. Radmer, A. Laederach, R. Das, P. Pearlman, D. Herschlag, and R. B. Altman. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, 15:189–199, 2 2009. ISSN 1355-8382. doi: 10.1261/rna.1270809.
11. J. Frellsen, I. Moltke, M. Thiim, K. V. Mardia, J. Ferkinghoff-Borg, and T. Hamelryck. A Probabilistic Model of RNA Conformational Space. *PLoS Computational Biology*, 5:e1000406, 6 2009. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000406.
12. J. Li, S. Zhang, D. Zhang, and S.-J. Chen. Vfold-Pipeline: a web server for RNA 3D structure prediction from sequences. *Bioinformatics*, 38(16):4042–4043, 2022. doi: 10.1093/bioinformatics/btab292.
13. L. Zhou, X. Wang, S. Yu, Y.-L. Tan, and Z.-J. Tan. FebRNA: an automated fragment-ensemble-based model for building RNA 3D structures. *Biophysical Journal*, 121:3381–3392, 2022.
14. A. M. Watkins, R. Rangan, and R. Das. FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure*, 28:963–976.e6, 8 2020. ISSN 09692126. doi: 10.1016/j.str.2020.05.011.
15. X. Xu and S.-J. Chen. Hierarchical Assembly of RNA Three-Dimensional Structures Based on Loop Templates. *Journal of Physical Chemistry B*, 122:5327–5335, 2017. doi: 10.1021/acs.jpcb.7b01933.
16. Y. Zhang, J. Wang, and Y. Xiao. 3dRNA: 3D Structure Prediction from Linear to Circular RNAs. *Journal of Molecular Biology*, 434(11):167452, 2022. ISSN 0022-2836. Computation Resources for Molecular Biology.
17. M. Popenda, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz, and R. W. Adamiak. Automated 3D structure composition for large RNAs. *Nucleic Acids Research*, 40:e112–e112, 8 2012. ISSN 1362-4962. doi: 10.1093/nar/gks339.
18. S. Cao and S.-J. Chen. Physics-Based De Novo Prediction of RNA 3D Structures. *Journal of Physical Chemistry B*, 115:4216–4226, 2011. doi: 10.1021/jp110370s.
19. M. Rother, K. Rother, T. Puton, and J. M. Bujnicki. ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Research*, 39:4007–4022, 5 2011. ISSN 1362-4962. doi: 10.1093/nar/gkq1320.
20. S. C. Flores, Y. Wan, R. Russell, and R. B. Altman. Predicting RNA structure by multiple template homology modeling. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, page 216—227, 2010. ISSN 2335-6928. doi: 10.1142/9789814295291_0024.
21. R. Das and D. Baker. Automated de novo prediction of native-like RNA tertiary structures. *Proceedings of the National Academy of Sciences of the United States of America*, 104:14664–9, 10 2007. doi: 10.1073/pnas.0703836104.
22. Andrew W. Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Žídek, Alexander W. R. Nelson, Alex Bridgland, Hugo Penedones, Stig Petersen, Karen Simonyan, Steve Crossan, Pushmeet Kohli, David T. Jones, David Silver, Koray Kavukcuoglu, and Demis Hassabis. Improved protein structure prediction using potentials from deep learning. *Nature*, 577:706–710, 1 2020. ISSN 0028-0836. doi: 10.1038/s41586-019-1923-7.
23. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596:583–589, 8 2021. ISSN 0028-0836. doi: 10.1038/s41586-021-03819-2.
24. R. Pearce, G. S. Omenn, and Y. Zhang. De Novo RNA Tertiary Structure Prediction at Atomic Resolution Using Geometric Potentials from Deep Learning. *bioRxiv*, May 2022. doi: 10.1101/2022.05.15.491755.
25. T. Shen, Z. Hu, Z. Peng, J. Chen, P. Xiong, L. Hong, L. Zheng, Y. Wang, I. King, S. Wang, S. Sun, and Y. Li. E2Efold-3D: End-to-End Deep Learning Method for Accurate de Novo RNA 3D Structure Prediction. *arXiv preprint arXiv:2207.01586*, 2022.
26. Y. Li, C. Zhang, C. Feng, R. Pearce, P. L. Freddolino, and Y. Zhang. Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nature Communications*, 14:5745, 2023. doi: 10.1038/s41467-023-41303-9.
27. Y. Kagaya, Z. Zhang, N. Ibtehaz, X. Wang, T. Nakamura, D. Huang, and D. Kihara. NuFold: A Novel Tertiary RNA Structure Prediction Method Using Deep Learning with Flexible Nucleobase Center Representation. *bioRxiv*, 2023. doi: 10.1101/2023.09.20.558715.
28. W. Wang, C. Feng, R. Han, Z. Wang, L. Ye, Z. Du, H. Wei, F. Zhang, Z. Peng, and J. Yang.

trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat Commun*, 14:7266, 2023. doi: 10.1038/s41467-023-42528-4.
29. R. Das, R. C. Kretsch, A. J. Simpkin, T. Mulvaney, P. Pham, R. Rangan, F. Bu, R. M. Keegan, M. Topf, D. J. Rigden, Z. Miao, and E. Westhof. Assessment of three-dimensional RNA structure prediction in CASP15. *Proteins*, 91(12):1747–1770, 2023. doi: 10.1002/prot.26602.
30. C. Bernard, G. Postic, S. Ghannay, and F. Tahi. State-of-the-RNArt: benchmarking current methods for RNA 3D structure prediction. *NAR Genomics and Bioinformatics*, 6(2):lqae048, June 2024. doi: 10.1093/nargab/lqae048.
31. B. Schneider, B. A. Sweeney, A. Bateman, J. Cerny, T. Zok, and M. Szachniuk. When will RNA get its AlphaFold moment? *Nucleic Acids Research*, page gkad726, 09 2023. ISSN 0305-1048. doi: 10.1093/nar/gkad726.
32. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, and D. Hassabis. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(493):493–500, 2024. doi: 10.1038/s41586-024-07487-w.
33. L. M. Wadley, K. S. Keating, C. M. Duarte, and A. M. Pyle. Evaluating and Learning from RNA Pseudotorsional Space: Quantitative Validation of a Reduced Representation for RNA Structure. *Journal of Molecular Biology*, 372(4):942–957, 2007. ISSN 0022-2836.
34. J. Černý, P. Božíková, J. Svoboda, and B. Schneider. A unified dinucleotide alphabet describing both RNA and DNA structures. *Nucleic Acids Research*, 48(11):6367–6381, Jun 2020. doi: 10.1093/nar/gkaa383.
35. Eric Westhof and Valérie Fritsch. RNA folding: beyond Watson–Crick pairs. *Structure*, 8(3):R55–R65, 2000. ISSN 0969-2126.
36. P. Gendron, S. Lemieux, and F. Major. Quantitative analysis of nucleic acid three-dimensional structures11edited by i. tinoco. *Journal of Molecular Biology*, 308(5):919–936, 2001. ISSN 0022-2836.
37. H. A. Gabb, S. R. Sanghani, C. H. Robert, and C. Prévost. Finding and visualizing nucleic acid base stacking. *Journal of Molecular Graphics*, 14(1):6–11, 1996. ISSN 0263-7855.
38. M. Parisien, J. Cruz, E. Westhof, and F. Major. New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA (New York, N.Y.)*, 15:1875–85, 09 2009. doi: 10.1261/rna.1700409.
39. Christian B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181 (4096):223–230, July 1973. doi: 10.1126/science.181.4096.223.
40. Seung S. Jang, Sara Dubnik, Jason Hon, Björn Hellenkamp, Daniel G. Lynall, Kenneth L. Shepard, Colin Nuckolls, and Ruben L. Jr. Gonzalez. Characterizing the conformational free-energy landscape of rna stem-loops using single-molecule field-effect transistors. *Journal of the American Chemical Society*, 145(1):402–412, Jan 2023. doi: 10.1021/jacs.2c10218.
41. Riku Yamagami, Jacob P. Sieg, and Philip C. Bevilacqua. Functional Roles of Chelated Magnesium Ions in RNA Folding and Function. *Biochemistry*, 60(31):2374–2386, Aug 2021. doi: 10.1021/acs.biochem.1c00012.
42. Urmi Chheda, Selvi Pradeepan, Edward Esposito, Steven Strezsak, Olivia Fernandez-Delgado, and James Kranz. Factors Affecting Stability of RNA – Temperature, Length, Concentration, pH, and Buffering Species. *Journal of Pharmaceutical Sciences*, 113(2):377–385, 2024. ISSN 0022-3549.
43. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucleic Acids Research*, 28(1):235–242, 01 2000. ISSN 0305-1048. doi: 10.1093/nar/28.1.235.
44. J. Cruz, M.-F. Blanchet, M. Boniecki, J. Bujnicki, S.-J. Chen, S. Cao, R. Das, F. Ding, N. Dokholyan, S. Flores, L. Huang, C. Lavender, V. Lisi, F. Major, K. Mikolajczak, D. Patel, A. Philips, T. Puton, J. Santalucia, and E. Westhof. RNA-Puzzles: A CASP-like evaluation of RNA three-dimensional structure prediction. *RNA (New York, N.Y.)*, 18:610–25, 02 2012. doi: 10.1261/rna.031054.111.
45. J. A. Cruz, M.-F. Blanchet, M. Boniecki, J. M. Bujnicki, S.-J. Chen, S. Cao, R. Das, F. Ding, N. V. Dokholyan, S. C. Flores, L. Huang, C. A. Lavender, V. Lisi, F. Major, K. Mikolajczak, D. J. Patel, A. Philips, T. Puton, J. Santalucia, F. Sijenyi, T. Hermann, K. Rother, M. Rother, A. Serganov, M. Skorupski, T. Soltysinski, P. Sripakdeevong, I. Tuszynska, K. M. Weeks, C. Waldsich, M. Wildauer, N. B. Leontis, and E. Westhof. RNA-Puzzles : A CASP-like evaluation of RNA three-dimensional structure prediction. *RNA*, 18:610–625, 4 2012. ISSN 1355-8382. doi: 10.1261/rna.031054.111.
46. Z. Miao, R. W. Adamiak, M. F. Blanchet, M. Boniecki, J. M. Bujnicki, S.-J. Chen, C. Cheng, G. Chojnowski, F.-C. Chou, P. Cordero, J. A. Cruz, A. R. Ferré-D'Amaré, R. Das, F. Ding, N. V. Dokholyan, S. Dunin-Horkawicz, W. Kladwang, A. Krokhotin, G. Łach, and M. Magnus. RNA-Puzzles Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA*, 21(6):1066–1084, 2015. doi: 10.1261/rna.049502.114.
47. Z. Miao, R. W. Adamiak, M. Antczak, R. T. Batey, A. J. Becka, M. Biesiada, M. J. Boniecki, J. M. Bujnicki, S.-J. Chen, C. Y. Cheng, F.-C. Chou, A. R. Ferré-D'Amaré, R. Das, W. K. Dawson, F. Ding, N. V. Dokholyan, S. Dunin-Horkawicz, C. Geniesse, K. Kappel, W. Kladwang, A. Krokhotin, G. E. Łach, F. Major, T. H. Mann, M. Magnus, K. Pachulska-Wieczorek, D. J. Patel, J. A. Piccirilli, M. Popenda, K. J. Purzycka, A. Ren, G. M. Rice, J. Santalucia, J. Sarzynska, M. Szachniuk, A. Tandon, J. J. Trausch, S. Tian, J. Wang, K. M. Weeks, B. Williams, Y. Xiao, X. Xu, D. Zhang, T. Zok, and E. Westhof. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA*, 23:655–672, 5 2017. ISSN 1355-8382. doi: 10.1261/rna.060368.116.
48. Z. Miao, R. W. Adamiak, M. Antczak, M. J. Boniecki, J. Bujnicki, S.-J. Chen, C. Y. Cheng, Y. Cheng, F.-C. Chou, R. Das, N. V. Dokholyan, F. Ding, C. Geniesse, Y. Jiang, A. Joshi, A. Krokhotin, M. Magnus, O. Mailhot, F. Major, T. H. Mann, P. Piatkowski, R. Pluta, M. Popenda, J. Sarzynska, L. Sun, M. Szachniuk, S. Tian, J. Wang, J. Wang, A. M. Watkins, J. Wiedemann, Y. Xiao, X. Xu, J. D. Yesselman, D. Zhang, Y. Zhang, Z. Zhang, C. Zhao, P. Zhao, Y. Zhou, T. Zok, A. Zyla, A. Ren, R. T. Batey, B. L. Golden, L. Huang, D. M. Lilley, Y. Liu, D. J. Patel, and E. Westhof. RNA-Puzzles Round IV: 3D structure predictions of four ribozymes and two aptamers. *RNA*, 26:982–995, 8 2020. ISSN 1355-8382. doi: 10.1261/rna.075341.120.
49. L. Becquey, E. Angel, and F. Tahi. RNANet: An Automatically Built Dual-Source Dataset Integrating Homologous Sequences and RNA Structures. *Bioinformatics*, 37(9):1218–1224, 2021. doi: 10.1093/bioinformatics/btaa944.
50. M. Szikszai, M. Magnus, S. Sanghi, S. Kadyan, N. Bouatta, and E. Rivas. RNA3DB: A structurally-dissimilar dataset split for training and benchmarking deep learning models for RNA structure prediction. *Journal of Molecular Biology*, March 2024. doi: 10.1016/j.jmb.2024.168552. Epub ahead of print.
51. M. Baek, R. McHugh, I. Anishchenko, H. Jiang, D. Baker, and F. DiMaio. Accurate prediction of protein-nucleic acid complexes using RoseTTAFoldNA. *Nat Methods*, 21:117–121, 2024. doi: 10.1038/s41592-023-02086-5.
52. K. Chen, Y. Zhou, S. Wang, and P. Xiong. Rna tertiary structure modeling with briq potential in casp15. *Proteins*, 91:1771–1778, 2023.
53. B. Adamczyk, M. Antczak, and M. Szachniuk. RNAsolo: a repository of clean, experimentally determined RNA 3D structures. *Bioinformatics*, 38(14):3668–3670, 2022. doi: 10.1093/bioinformatics/btac386.
54. I. Kalvari, E. P. Nawrocki, N. Ontiveros-Palacios, J. Argasinska, K. Lamkiewicz, M. Marz, S. Griffiths-Jones, C. Toffano-Nioche, D. Gautheret, Z. Weinberg, E. Rivas, S. R. Eddy, R. D. Finn, A. Bateman, and A. I. Petrov. Rfam 14: Expanded Coverage of Metagenomic, Viral and MicroRNA Families. *Nucleic Acids Research*, 2020. doi: 10.1093/nar/gkaa1047.
55. L. Fu, B. Niu, Z. Zhu, S. Wu, and W. Li. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23):3150–3152, 12 2012. ISSN 1367-4803. doi: 10.1093/bioinformatics/bts565.
56. X.-W. Qiang, C. Zhang, H.-L. Dong, F.-J. Tian, H. Fu, Y.-J. Yang, L. Dai, X.-H. Zhang, and Z.-J. Tan. Multivalent Cations Reverse the Twist-Stretch Coupling of RNA. *Phys. Rev. Lett.*, 128:108103, Mar 2022. doi: 10.1103/PhysRevLett.128.108103.
57. F. Liu and Z.-C. Ou-Yang. Monte Carlo Simulation for Single RNA Unfolding by Force. *Biophysical Journal*, 88(1):76–84, 2005. ISSN 0006-3495. doi: 10.1529/biophysj.104.049239.
58. M. Parisien and F. Major. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*, 452:51–55, 3 2008. ISSN 0028-0836. doi: 10.1038/nature06684.
59. M. Magnus, M. Antczak, T. Zok, J. Wiedemann, P. Łukasiak, Y. Cao, J. M. Bujnicki, E. Westhof, M. Szachniuk, and Z. Miao. RNA-Puzzles toolkit: a computational resource of RNA 3D structure benchmark datasets, structure manipulation, and evaluation tools. *Nucleic Acids Research*, 2019. doi: 10.1093/nar/gkz1108.
60. C. Bernard, G. Postic, S. Ghannay, and F. Tahi. RNAdvisor: a comprehensive benchmarking tool for the measure and prediction of RNA structural model quality. *Briefings in Bioinformatics*, 25:bbae064, March 2024. doi: 10.1093/bib/bbae064.
61. S. Bottaro, F. Di Palma, and G. Bussi. The Role of Nucleobase Interactions in RNA Structure and Dynamics. *Nucleic acids research*, 42, 10 2014. doi: 10.1093/nar/gku972.
62. Y. Zhang and J. Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins*, 57(4):702—710, December 2004. ISSN 0887-3585. doi: 10.1002/prot.20264.
63. S. Gong, C. Zhang, and Y. Zhang. RNA-align: quick and accurate alignment of RNA 3D structures based on size-independent TM-scoreRNA. *Bioinformatics (Oxford, England)*, 35 (21):4459–4461, 2019. doi: 10.1093/bioinformatics/btz282.
64. A. Zemla, C. Venclovas, J. Moult, and K. Fidelis. Processing and analysis of CASP3 protein structure predictions. *Proteins: Structure, Function, and Bioinformatics*, 37(S3):22–29, 1999.
65. K. Olechnovic, E. Kulberkyte, and C. Venclovas. CAD-score: A new contact area difference-based function for evaluation of protein structural models. *Proteins*, 81, 01 2013. doi: 10.1002/prot.24172.
66. V. Mariani, M. Biasini, A. Barbato, and T. Schwede. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics (Oxford, England)*, 29(21):2722–2728, 2013. doi: 10.1093/bioinformatics/btt473.
67. C. Hajdin, F. Ding, N. Dokholyan, and K. Weeks. On the significance of an RNA tertiary structure prediction. *RNA (New York, N.Y.)*, 16:1340–9, 07 2010. doi: 10.1261/rna.1837410.
68. T. Zok, M. Popenda, and M. Szachniuk. MCQ4Structures to compute similarity of molecule structures. *Central European Journal of Operations Research*, 22, 04 2013. doi: 10.1007/s10100-013-0296-5.
69. Jens Wiedemann, Thomas Zok, Martin Milostan, et al. LCS-TA to identify similar fragments in RNA 3d structures. *BMC Bioinformatics*, 18:456, 2017. doi: 10.1186/s12859-017-1867-6.
70. C. Zhang, M. Shine, A. M. Pyle, and Y. Zhang. US-align: Universal Structure Alignment of Proteins, Nucleic Acids and Macromolecular Complexes. *Nature Methods*, 19:1109–1115, 2022.
71. David B Haack, Benjamin Rudolfs, Cheng Zhang, Dmitry Lyumkis, and Navtej Toor. Structural basis of branching during RNA splicing. *Nat Struct Mol Biol*, 31(1):179–189, Jan 2024. ISSN 1545-9985. doi: 10.1038/s41594-023-01150-0.
72. Tali Ast, Yuki Itoh, Sophia Sadre, John G. McCoy, Grace Namkoong, Jacob C. Wengrod, Ivan Chicherin, Pramod R. Joshi, Pavel Kamenski, Daniel L. M. Suess, Alexey Amunts, and Vamsi K. Mootha. METTL17 is an Fe-S cluster checkpoint for mitochondrial translation. *Mol Cell*, 84(2):359–374.e8, Jan 2024. ISSN 1097-4164. doi: 10.1016/j.molcel.2023.12.016.
73. Nisarg Desai, Adam Brown, Alexey Amunts, and Venki Ramakrishnan. The structure of the yeast mitochondrial ribosome. *Science*, 355(6324):528–531, Feb 2017. doi: 10.1126/science.aal2415.
74. Nicholas J. Harper, Chloe Burnside, and Sebastian Klinge. Principles of mitoribosomal small subunit assembly in eukaryotes. *Nature*, 614:175–181, 2023. doi: 10.1038/s41586-022-05621-0.
75. Tslil Ast, Yuzuru Itoh, Shayan Sadre, Jason G. McCoy, Gil Namkoong, Jordan C. Wengrod, Ivan Chicherin, Pallavi R. Joshi, Piotr Kamenski, Daniel L.M. Suess, Alexey Amunts, and Vamsi K. Mootha. METTL17 is an Fe-S cluster checkpoint for mitochondrial translation. *Molecular Cell*, 84(2):359–374.e8, 2024. ISSN 1097-2765.
76. K. Sakaniwa, A. Fujimura, T. Shibata, H. Shigematsu, T. Ekimoto, M. Yamamoto, M. Ikeguchi, K. Miyake, U. Ohto, and T. Shimizu. TLR3 forms a laterally aligned multimeric complex along double-stranded RNA for efficient signal transduction. *Nature Communica-*

tions, 14(1):164, Jan 2023. doi: 10.1038/s41467-023-35844-2.

77. T. Dendooven, E. Sonnleitner, U. Bläsi, and B. F. Luisi. Translational regulation by Hfq-Crc assemblies emerges from polymorphic ribonucleoprotein folding. *EMBO Journal*, 42(3): e111129, Feb 2023. doi: 10.15252/embj.2022111129.
78. Bingnan Luo, Chong Zhang, Xiaobin Ling, Sunandan Mukherjee, Guowen Jia, Jiahao Xie, Xinyu Jia, Liu Liu, Eugene F. Baulin, Yongbo Luo, Longxing Jiang, Haohao Dong, Xiawei Wei, Janusz M. Bujnicki, and Zhaoming Su. Cryo-EM reveals dynamics of Tetrahymena group I intron self-splicing. *Nat Catal*, 6:298–309, 2023. ISSN 2520-1158. doi: 10.1038/s41929-023-00934-3.