

Has *AlphaFold3* achieved success for RNA?

Clément Bernard,^{a,b} Guillaume Postic,^{a*} Sahar Ghannay^b and Fariza Tahia^{a*}

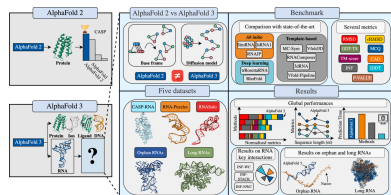
^aUniversité Paris-Saclay, Université Evry, IBISC, 91020 Evry-Courcouronnes, France, and ^bLISN – CNRS/Université Paris-Saclay, 91400 Orsay, France. *Correspondence e-mail: guillaume.postic@universite-paris-saclay.fr, fariza.tahia@univ-evry.fr

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 the fundamental differences between proteins and RNA, which hinder its direct adaptation. The latest release of *AlphaFold*, *AlphaFold3*, has broadened its scope to include multiple different molecules such as DNA, ligands and RNA. While the *AlphaFold3* article discussed the results for the last CASP-RNA data set, the scope of its performance and the limitations for RNA are unclear. In this article, we provide a comprehensive analysis of the performance of *AlphaFold3* in the prediction of 3D structures of RNA. Through an extensive benchmark over five different test sets, we discuss the performance and limitations of *AlphaFold3*. We also compare its performance with ten existing state-of-the-art *ab initio*, template-based and deep-learning approaches. Our results are freely available on the EvryRNA platform at <https://evryrna.ibisc.univ-evry.fr/evryrna/alphafold3/>.

1. Introduction

Ribonucleic acids (RNA) are fundamental molecules that are crucial to cellular activities. While their functions are directly linked to their structures, prediction of the latter remains an open challenge to be addressed. Knowledge of the structure of RNA could be of great interest for drug design or for the comprehension of biological processes such as cancer (Zhu *et al.*, 2022). While experimental methods such as X-ray crystallography, NMR and cryo-EM can determine 3D structures of RNA, their use is costly (in terms of time and resources) and is hardly scalable to the number of RNA molecules that are found in life. Computational approaches have emerged using *ab initio*, template-based and, more recently, deep-learning methods. *Ab initio* methods (Li & Chen, 2023; Krokhotin *et al.*, 2015; Zhang, Chen *et al.*, 2021; Zhang, Li *et al.*, 2021; Boniecki *et al.*, 2016; Cragolini *et al.*, 2015; Kerpedjiev *et al.*, 2015; Šulc *et al.*, 2014; Jonikas *et al.*, 2009; Frelsen *et al.*, 2009) tend to reproduce the physics of the system, with a force field applied to a coarse-grained representation (low resolution, in which a nucleotide is replaced by some of its atoms). Template-based approaches (Li *et al.*, 2022; Zhou *et al.*, 2022; Watkins *et al.*, 2020; Xu & Chen, 2017; Zhang, Wang *et al.*, 2022; Popenda *et al.*, 2012; Cao & Chen, 2011; Rother *et al.*, 2011; Flores *et al.*, 2010; Das & Baker, 2007) create a mapping between sequences and fragments of structure before refining the assembled structures.

With the recent success of *AlphaFold* for proteins (Senior *et al.*, 2020; Jumper *et al.*, 2021), approaches have been made to replicate its success with RNA. The direct use of protein methods to infer 3D structures of RNA is impossible, as RNA and proteins are chemically and physically different molecules. Current methods, such as *DeepFoldRNA* (Pearce *et al.*, 2022), *RhoFold* (Shen *et al.*, 2022), *DrFold* (Li *et al.*, 2023), *NuFold*



(Kagaya *et al.*, 2025) and *trRosettaRNA* (Wang *et al.*, 2023), try to adapt what already exists for proteins to RNA. They consider coarse-grained representations and predict Euclidean transformations before reconstructing the full-atom structure. The use of torsion angles is also adapted to RNA, using either the standard torsion angles (Shen *et al.*, 2022; Kagaya *et al.*, 2025) or angles from their coarse-grained representations (Pearce *et al.*, 2022; Li *et al.*, 2023).

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 (Das *et al.*, 2023) and in our recent benchmark State-of-the-RNART (Bernard *et al.*, 2024b). Recently, a critical review (Schneider *et al.*, 2023) explained the reasons why *AlphaFold* for RNA has not yet arrived and might not arrive for some decades. However, *AlphaFold* has released its latest version, named *AlphaFold3* (Abramson *et al.*, 2024), that extends its predictions to different molecules, including RNA. In this work, we aim to provide a response to Schneider *et al.* (2023) in order to determine whether *AlphaFold3* achieves success for RNA.

To extend its range of molecules, *AlphaFold3* has made changes in its architecture in order to better adapt to the variety of available inputs. It no longer relies on torsion angles, which restricted it to specific molecules, as was the case in *AlphaFold2* (Jumper *et al.*, 2021). It directly predicts atom coordinates with the use of a multi-cross-diffusion model. The authors mentioned good results through a benchmark on CASP-RNA (Abramson *et al.*, 2024), but *AlphaFold3* did not outperform human-assisted methods. Furthermore, it is not clear what the current limitations are and how well it performs compared with state-of-the-art solutions.

This article aims to provide a comprehensive extension of the evaluation and benchmarking of *AlphaFold3* for RNA. We first describe the main differences between RNA and proteins to highlight the challenges of RNA 3D structure prediction and describe the *AlphaFold3* solution before discussing the benchmark that we performed. We then evaluate *AlphaFold3* and comment on the results of *AlphaFold3* 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 at <https://evryrna.ibisc.univ-evry.fr/evryrna/alphafold3/>.

2. RNA versus proteins

RNA and proteins are both molecules that play crucial roles in life. 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 unstructured proteins also play a significant role in protein function, making this relationship more complex. This section discusses the differences between RNA and proteins, highlighting the reasons why adapting existing protein models has been challenging.

RNA comprises four nucleotides (A, C, G and U), whereas proteins comprise 20 amino acids. This difference has a large

consequence for the adaptation of protein algorithms to RNA. The vocabulary available for RNA is limited to four unique elements, making the protein vocabulary not directly adaptable. The sequence length of RNA molecules also has a high variability (from a dozen to thousands of nucleotides) compared with proteins (from a dozen to hundreds of amino acids).

A major difference between RNA and proteins lies in the folding stabilization. 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-torsions η and θ (Wadley *et al.*, 2007). 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 the description of RNA structure using dinucleotide-like fragments to accurately capture the backbone geometry (Černý *et al.*, 2020). Protein models therefore have a conformational mechanism that is fundamentally different from the RNA folding process, where such adaptations and structural dependencies must be carefully accounted for.

The nature of pairwise interactions in 3D RNA molecules differs from those in proteins. RNA interactions can be made through three different edges of the RNA base: the WC edge, Hoogsteen edge and sugar edge (Westhof & Fritsch, 2000), as shown in Fig. 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 for interaction between bases. The standard Watson–Crick (WC) base pair corresponds to a *cis* WC/WC pairing. Given

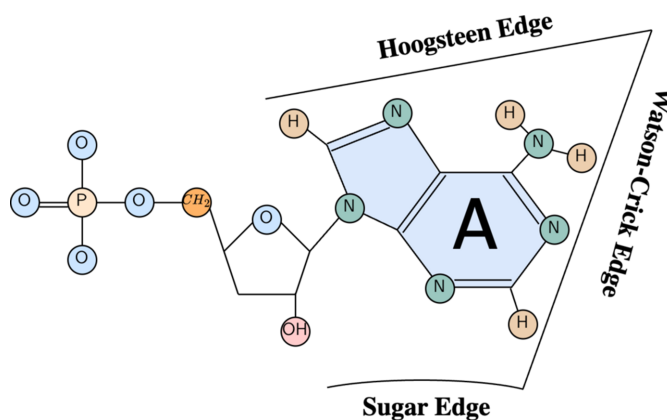


Figure 1
Description of the three different edges of the adenine RNA nucleotide: Watson–Crick edge, Hoogsteen edge and sugar edge. The three other nucleotides share similar edges.

the 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) AU and CG (and also the GU wobble pair) are used in the 2D structure representation. RNA bases also have common patterns of interactions, where a base stacks on another one. The base stacking (Gendron *et al.*, 2001; Gabb *et al.*, 1996) refers to the four base-stacking types in relative orientations (upwards, downwards, outwards and inwards; Parisien *et al.*, 2009). 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 to stabilize the structure and cannot be ignored when working on 3D structures of RNA.

The stability of 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 kinetically accessible and thermodynamically stable (Anfinsen, 1973). 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. In contrast, RNA molecules often have a more rugged Gibbs free-energy landscape, thus populating multiple conformational states (Jang *et al.*, 2023). Switching between these conformations supports some functions of RNA, such as riboswitches or ribozymes, and may be driven by environmental changes, such as ions (notably Mg^{2+}), pH, temperature or ligand binding (Yamagami *et al.*, 2021; Chheda *et al.*, 2024).

There is a huge disparity between protein and RNA data. Even if there is a higher proportion of RNA than proteins in life, this is not reflected in the available data: only a small number of 3D structures of RNA are known. Up to June 2024, 7759 RNA structures had been deposited in the Protein Data Bank (PDB; Berman *et al.*, 2000), compared with 216 212 protein structures. The quality and diversity of the data are also different: a huge proportion of RNA structures come from the same families. This implies several redundant structures that could prevent a model from being generalized to other families. In addition, a huge amount of RNA families do not yet have solved structures in the PDB, usually those of long RNA. This means there is no balanced and representative proportion of RNA families among the known structures.

Finally, no standard data set has been used for RNA throughout the community. Each research group uses a data set with different associated preprocessing. This prevents the use of deep-learning methods, as a lot of work is needed to obtain a clean data set. While the community has agreed to use RNA-Puzzles (Cruz *et al.*, 2012; Miao *et al.*, 2015, 2017, 2020) or the new CASP-RNA (Das *et al.*, 2023) to test the generalization of proposed models, no clear training set is available. The first solution was RNANet (Becquey *et al.*, 2021), which was developed in our laboratory 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 (Szikszai *et al.*, 2024), creates independent data sets

for deep-learning approaches, where clustering is performed based on sequence and structural disparity.

3. AlphaFold3

Building on the recent success of *AlphaFold2* (Jumper *et al.*, 2021) in protein structure prediction, *AlphaFold3* (Abramson *et al.*, 2024) has expanded its predictions to the structures of all molecules available in the PDB (Berman *et al.*, 2000). The authors highlight several differences from the previous architecture that contribute 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 atomic level. *AlphaFold3* also directly outputs the coordinate atom positions, compared with the prediction of rotation/translation vectors (and torsion angles) in the previous version. It also weights the multiple sequence alignment (MSA) less in the overall model. In the case of RNA, *AlphaFold3* has been evaluated on the CASP-RNA data set (Das *et al.*, 2023), demonstrating improved predictions compared with *RosettaFold2NA* (Baek *et al.*, 2024) and *AIchemy_RNA* (the best AI-based submission in the competition; Shen *et al.*, 2022). Despite these advancements, the performance of *AlphaFold3* lags behind that of *AIchemy_RNA2* (the top human-expert-aided submission; Chen *et al.*, 2023). Further details of the architecture, the training procedure and the differences between *AlphaFold2* and *AlphaFold3* are provided in the supporting information.

4. Benchmark

To assess the performance of *AlphaFold3*, we have evaluated it and compared it with other state-of-the-art methods on five data sets. This section describes the data sets and the methods as well as the metrics used to evaluate *AlphaFold3*.

4.1. Data sets

To evaluate the prediction of RNA structures, we considered the following five test sets, with the first three being used in our previous work (Bernard *et al.*, 2024b).

(i) *RNA-Puzzles*. The first data set is composed of the single-stranded structures from RNA-Puzzles (Cruz *et al.*, 2012; Miao *et al.*, 2015, 2017, 2020), a community initiative to benchmark RNA structures. We considered only single-stranded solutions in order to have a fair comparison between the benchmarked models. It is composed of 22 RNAs with lengths between 27 and 188 nt (with a mean of 83 nt).

(ii) *CASP-RNA*. The second test set is composed of the CASP-RNA (Das *et al.*, 2023) structures from a collaboration between the CASP team and RNA-Puzzles. It is composed of 12 RNAs with a wide range of sequences from 30 to 720 nt (with a mean of 209 nt).

(iii) *RNASolo*. The third test set is a custom test set composed of independent structures from RNAsolo (Adamczyk *et al.*, 2022). We downloaded representative RNA molecules from RNAsolo (Adamczyk *et al.*, 2022) with resolutions below 4 Å and removed structures with a sequence identity of higher than 80%. We then considered only the

structures with a unique Rfam family ID (Kalvari *et al.*, 2020), leading to 25 nonredundant RNA molecules with a sequence of between 45 and 298 nt (and a mean of 100 nt). It cannot be ensured that the structures from this data set were not used in the training sets for the different models. We keep this data set for comparison, as we already have the results for the benchmarked methods.

(iv) *RNA3DB_0*. This data set is composed of a non-redundant set of structurally and sequentially independent structures from RNA3DB (Szikszai *et al.*, 2024). 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 belong to Rfam families (Kalvari *et al.*, 2021) and include synthetic RNAs and 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*; Fu *et al.*, 2012), we ended up with a data set of 224 structures from 10 to 339 nt (with a mean of 55 nt). Nonetheless, these structures come from complexes, meaning that they do not behave well in isolation, and thus their experimentally observed conformations depend on other chains. To account for this, in evaluating the models we considered 113 structures with their full context and predicted the structures with *AlphaFold3* (the other structures have too large a context and we failed to predict them using *AlphaFold3*). We name this subset *RNA3DB_0* (Context).

(v) *RNA3DB_Long*. The last data set comprises long RNA structures from RNA3DB (Szikszai *et al.*, 2024). 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 5000 nt, as we wanted to study the performance of long RNAs. This led to 58 structures with a sequence of between 828 and 3619 nt (with a mean of 2005 nt). They comprise 57 ribosomal RNAs and one structure of a group II intron.

We have also ensured that all of the data sets (except RNA-Puzzles and CASP-RNA) have a sequence identity below 80% in order to have nonredundant structures for robust evaluation.

To comprehend and detail the predictions by *AlphaFold3*, we studied the three main interactions in the folding of 3D structures of RNA in detail: Watson–Crick (WC), non-Watson–Crick (nWC) and stacking (STACK). The proportion of these interactions is presented in Table 1. All data sets have the same proportion of stacking (around 75%), except for the *RNA3DB_0* data set (around 56%). As *RNA3DB_0* contains orphan structures, this implies structures with less common folding, as reflected by the lower proportion of stacking interactions. For all of the data sets 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 structures. When comparing *RNA3DB_0* with or without context, we observe a greater proportion of stacking and Watson–Crick interactions in the

Table 1

Proportion of key RNA interactions in the five test sets (and the subset of RNA3DB with context).

The interactions are normalized by the number of residues. Interactions are either stacking (STACK), Watson–Crick (WC) or non-Watson–Crick (nWC), as extracted from MC-Annotate (Gendron *et al.*, 2001). Interactions for *RNA3DB_0* are computed without context, while *RNA3DB_0* (C) includes context.

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

presence of context. However, the number of non-Watson–Crick interactions remains unchanged.

4.2. State-of-the-art methods

Existing solutions for the prediction of 3D structures of RNA are based on three main types of methods: *ab initio*, template-based and deep-learning methods. As discussed previously in our work (Bernard *et al.*, 2024b), *ab initio* methods (Boniecki *et al.*, 2016; Zhang, Li *et al.*, 2021; Li & Chen, 2023) usually integrate the physics of the system by simplifying the representation of nucleotides (coarse-grained). Instead of using all of the atoms for one nucleotide, they create a low-resolution representation that simplifies the computation time while losing information. They use approaches such as molecular dynamics (Qiang *et al.*, 2022) or Monte Carlo (Liu & Ou-Yang, 2005) to perform sampling in conformational space and use a force field to simulate real environmental conditions. On the other hand, template-based methods (Parisien & Major, 2008; Cao & Chen, 2011; Popenda *et al.*, 2012; Zhang, Wang *et al.*, 2022; Li *et al.*, 2022) create a mapping between sequences and known motifs with, for instance, secondary-structure trees (SSEs) before reconstructing the full structure from its subfragments. Finally, recent methods tend to incorporate deep-learning methods (Wang *et al.*, 2023; Kagaya *et al.*, 2025; Li *et al.*, 2023; Pearce *et al.*, 2022; Shen *et al.*, 2022) by using attention-based architectures with self-distillation and recycling, as performed in *AlphaFold2* (Jumper *et al.*, 2021).

To compare the performance of *AlphaFold3* (Abramson *et al.*, 2024), we benchmarked ten approaches, those used in our previous work (Bernard *et al.*, 2024b). For the *ab initio* methods, we benchmarked *SimRNA* (Boniecki *et al.*, 2016), *IsRNA1* (Zhang, Li *et al.*, 2021) and *RNAJP* (Li & Chen, 2023). Only *RNAJP* was used locally. For the template-based approaches, we benchmarked *MC-Sym* (Parisien & Major, 2008), *Vfold3D* (Cao & Chen, 2011), *RNAComposer* (Popenda *et al.*, 2012), *3dRNA* (Zhang, Wang *et al.*, 2022) and *Vfold-Pipeline* (Li *et al.*, 2022). For the deep-learning methods, we benchmarked *trRosettaRNA* (Wang *et al.*, 2023) and *RhoFold* (Shen *et al.*, 2022). Further details of each method have been provided in our previous article (Bernard *et al.*, 2024b). For RNA-Puzzles and CASP-RNA, we included the predictions from the official results of the competitions in the

benchmark. We refer to these as ‘challenge best’ and they correspond to different methods for each RNA. We normalized each prediction using *RNA-tools* (Magnus *et al.*, 2020) to give a standard format for all structures. It gives standardized names for chains, residues and atoms and removes ions and water.

We used the web servers with default parameters to compare available models fairly, so that users could reproduce our experiments. As we made most of the predictions using web servers, the predictions for RNA3DB_0 were hardly applicable to all of the methods. Therefore, we benchmarked the RNA3DB_0 data set with one method per approach (the quickest method per approach): *RhoFold* (Shen *et al.*, 2022) for deep learning, *RNAComposer* (Popenda *et al.*, 2012) for template-based and *RNAJP* (Li & Chen, 2023) for *ab initio*. For the RNA3DB_Long data set, only *AlphaFold3* could predict structures with sequences up to 3000 nt. For this data set, we only considered the predictions from *AlphaFold3*.

4.3. Evaluation metrics

To compare the predictions, we used the *RNAAdvisor* tool developed by our team (Bernard *et al.*, 2024a), which enables the computation of a wide range of existing metrics on one command line. For the evaluation of 3D structures of RNA, a general assessment of the folding of the structure can be performed with either the root-mean-square deviation (RMSD) or its extension adding RNA features ϵ RMSD (Bottaro *et al.*, 2014). Protein-inspired metrics can also be adapted to assess structure quality, such as the TM-score (Zhang & Skolnick, 2004; Gong *et al.*, 2019) or the GDT-TS (which counts the number of superimposed atoms; Zemla *et al.*, 1999). There are also the CAD-score (which measures the structural similarity in a contact-area difference-based function; Olechnovič *et al.*, 2013) and the IDDT (which assesses the interatomic distance differences between a reference structure and a predicted structure; Mariani *et al.*, 2013). Finally, RNA-specific metrics have been developed, such as the *P*-value (which assesses the non-randomness of a given prediction; Hajdin *et al.*, 2010). The INF-ALL (Parisien *et al.*, 2009) and DI (Parisien *et al.*, 2009) have been developed to consider RNA-specific interactions. The INF score incorporates canonical and noncanonical pairing with Watson–Crick (INF-WC), non-Watson–Crick (INF-NWC) and stacking (INF-STACK) interactions. The consideration of torsion angles has been developed with the mean of circular quantities (MCQ; Zok *et al.*, 2014) and LCS-TA (longest continuous segments in torsion angle space; Wiedemann *et al.*, 2017). As discussed in Bernard *et al.* (2024a), all of 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, IDDT, TM-score and MCQ; the results for the other metrics are given in the supporting information. Indeed, the RMSD is the most used metric in the literature, and the INF-ALL incorporates key RNA interactions. The IDDT and TM-score allow the evaluation of global conformations (widely used in *AlphaFold3*), and MCQ gives the torsional deviation. We only

mention all of the metrics when comparing the different models to ensure a complete evaluation.

5. Results

This section presents the results of *AlphaFold3* predictions on the discussed test sets. We start by comparing the results of *AlphaFold3* with existing solutions and then discuss in detail the link between the performance and the sequence length. Next, we discuss the results of *AlphaFold3* on ribosomal structures (RNA3DB_Long data set) and orphan structures (RNA3DB_0 data set). We then discuss the results of specific RNA key interactions in detail before shedding light on the computation time.

5.1. *AlphaFold3* compared with the state of the art

We compare the predictions of the ten existing methods presented above and *AlphaFold3* on our different test sets. Fig. 2 presents the different normalized metrics computed for the prediction of the different models over the five test sets. We included all metrics to show the cumulative performance. The RNA3DB_Long data set only has predictions from *AlphaFold3*, which is the only method that is capable of processing long sequences. All of the metrics are normalized by the maximum values and converted to be better when near to 1 and worse when near to 0. Real values for each metric for the five test sets are reported in Supplementary Tables S1, S2, S3, S4 and S5.

The best models from the CASP-RNA competition, which are human-guided, outperform *AlphaFold3* (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 data set. On the other hand, *AlphaFold3* 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 with the more recent CASP-RNA solutions. For the RNA3DB_0 data set, the performance of *AlphaFold3* is slightly better compared with *RhoFold*, which gives a better RMSD but a worse MCQ and LCS-TA. *AlphaFold3* always has a high MCQ value, indicating that it returns structures which are more physically plausible than *ab initio* methods (which use physics properties in their predictions). Nonetheless, it does not always have the best RMSD (outperformed in CASP-RNA and RNA3DB_0), suggesting that *AlphaFold3* does not always have the best alignment (in terms of all atoms) compared with the reference structure.

To compare the global performance of each type of approach, in Fig. 3 we report the averaged metrics over the different types of approach depending on the sequence length. We grouped the results for structures in 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 of the other metrics are shown in Supplementary Fig. S2. None of the benchmarked *ab initio* methods successfully predicted structures for sequences

exceeding 200 nt, particularly when using web servers. The best results from the CASP-RNA and RNA-Puzzles challenges outperform *AlphaFold3* across most metrics, except for sequences between 150 and 250 nt, where *AlphaFold3* showed comparable results. The values of RMSD, TM-score, MCQ and IDDT 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 a global alignment that is worse than the other methods (due to the high simulation time, which is a bottleneck for web-server usage). Deep-learning approaches, in particular, produced worse MCQ scores than traditional methods. *AlphaFold3* demonstrated an especially strong MCQ performance, with comparative results for the best solutions of challenges for sequences greater than 250 nt.

These results suggest that *AlphaFold3* achieves a competitive performance, particularly in capturing more realistic

torsion 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 of more than 200 nt.

5.2. The performance of *AlphaFold3* relative to sequence length

As seen previously, the prediction of 3D structures of RNA usually becomes harder when the sequence length increases. Indeed, the *ab initio* methods fail to predict long interactions as the computation time increases greatly with sequence length. The template-based approaches, as well as the deep-learning methods, are limited by the small number of long RNA structures, as shown in Bernard *et al.* (2024b). To observe the relation between sequence length and *AlphaFold3* performance in more detail, in Fig. 4 we report the RMSD, MCQ, TM-score, IDDT and INF-ALL metrics depending on sequence length (for the five test sets). The links between

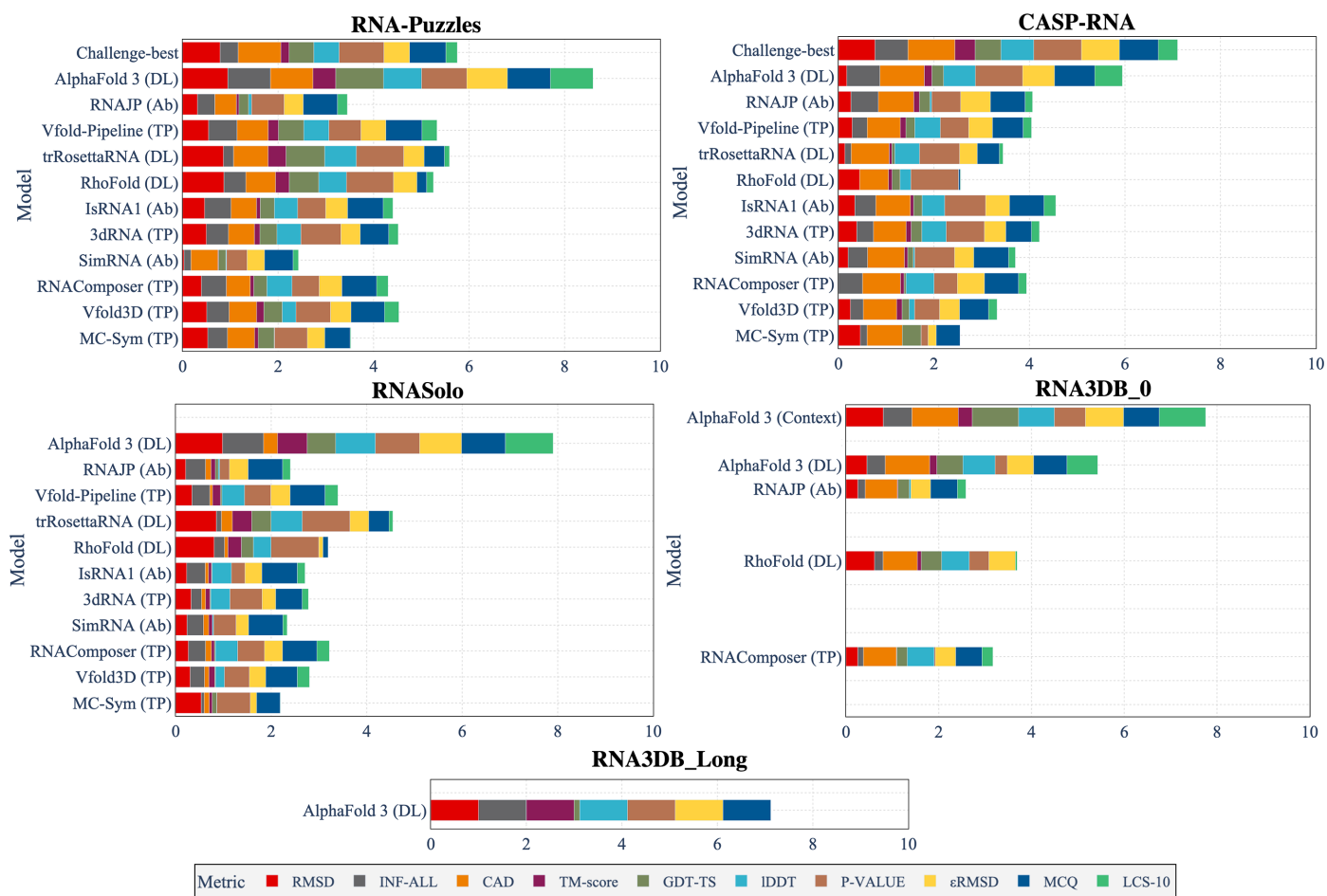


Figure 2

Cumulative normalized metrics (the higher the better) for each of the benchmarked methods for our five test sets. Each metric is normalized by the maximum value over the five test sets, and the decreased metrics are inverted to have better values close to 1. Challenge-best means the best solutions from the RNA-Puzzles and CASP-RNA competitions (and corresponds to different solutions for each challenge). The types of methods are also mentioned with the abbreviations DL for deep learning, TP for template-based and Ab for *ab initio*. Methods are sorted by release time (except for challenge-best). *AlphaFold3* (Context) represents the predictions of *AlphaFold3* for 113 structures of the RNA3DB_0 data set with the context of the structures added as input.

the other metrics and the sequence length are available in Supplementary Fig. S3.

Fig. 4 indicates that, except for the RNA3DB_0 data set, the RMSD becomes worse for sequences between 0 and 1000 nt. For the RNA3DB_Long data set with sequences longer than 1000 nt the predictions have good results for every metric. We also observe a tendency for degradation in the IDDT, TM-score and INF-ALL (smaller decrease) when the structures have sequences of longer than 100 nt (and below 1000 nt). For every metric, the predictions for the RNA3DB_0 (with or without context) data set 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 performance for the *AlphaFold3* predictions.

5.3. AlphaFold3 results on long RNA

Current methods for the prediction of RNA 3D structures are limited for long RNA and hardly predict structures with sequences longer than 200 nt. *AlphaFold3* is, to the best of our knowledge, the only method that can predict long RNA structures (with sequences longer than 1000 nt). Its predic-

tions on RNA3DB_Long show a good performance, as shown in Fig. 2. The only metrics for which the results are not good are GDT-TS, CAD-score and LCS-TA (threshold of 10°), which might be due to an error in computation. For LCS-TA, the low 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 data set.

The good results for long RNA 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 they have a high redundancy. This might be reflected in the PDB, which has been memorized by *AlphaFold3* during its training. As *AlphaFold3* 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 generalization of previously observed families from *AlphaFold3*.

We report the two worst predictions of *AlphaFold3* on the RNA3DB_Long data set in Fig. 5. The two worst predictions for the other test sets are provided in Supplementary Fig. S4. The RMSD for the two structures is relatively high (greater than 19 Å). The second worst structure has a high TM-score

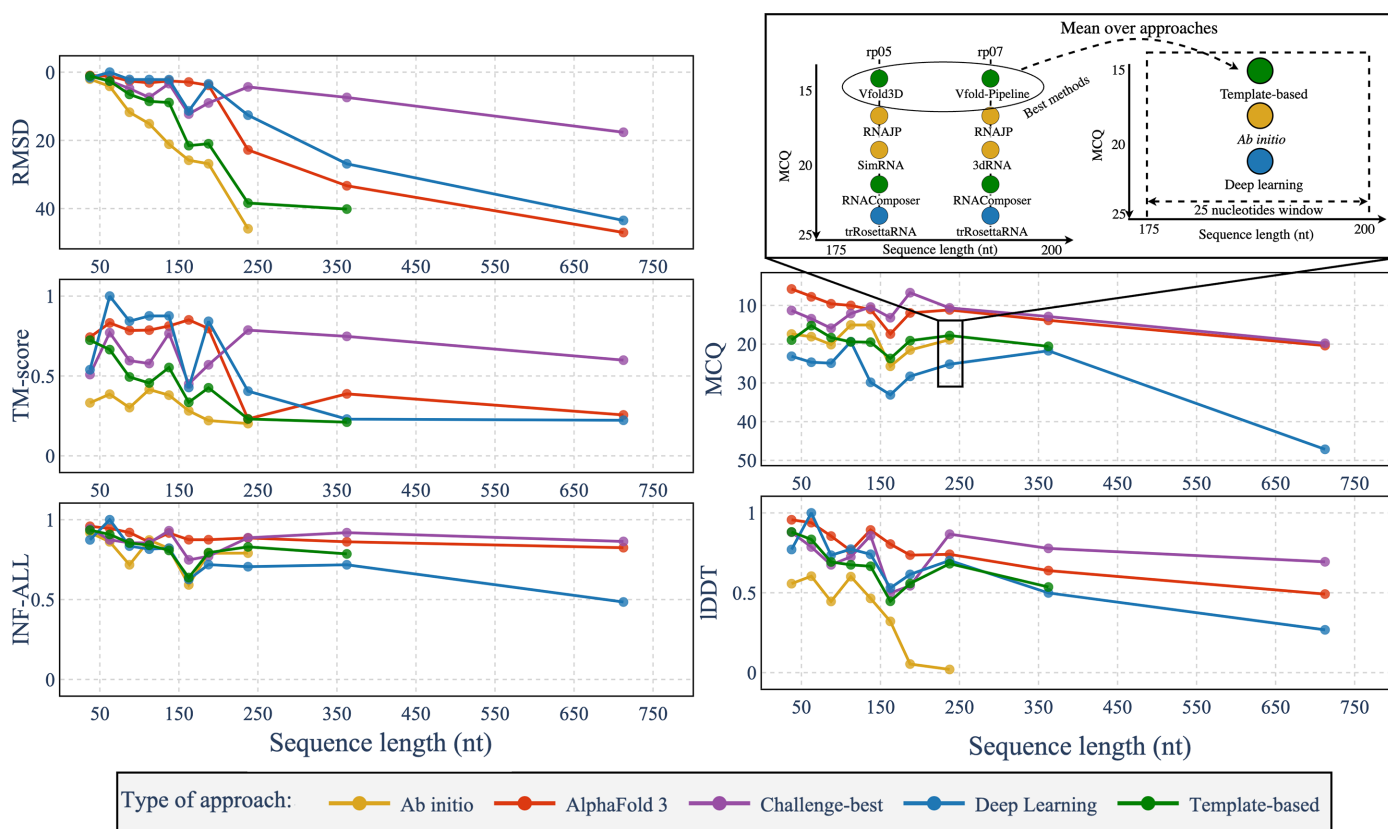


Figure 3

Averaged metrics depending on the sequence length for the different approaches (*AlphaFold3*, *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 data sets. Challenge-best corresponds to the best results from either the RNA-Puzzles or CASP-RNA competitions but does not appear for the RNASolo data set. The metrics are RMSD, MCQ, TM-score, IDDT and INF-ALL. RMSD and MCQ are reversed to have the best values near the top and the worst values at the bottom.

(0.74), meaning that even for a long structure (1487 nt) the global alignment of atoms is well predicted. INF-ALL is also high for these structures (higher than 0.68), meaning that it returns a high proportion of key RNA interactions. In detail, it is most likely to be no coincidence that the worst prediction (TM-score = 0.38) corresponds to the only nonribosomal RNA in the RNA3DB_Long data set, 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 *AlphaFold3* either, as this group II intron RNA can be found in complex with its large maturase/reverse transcriptase (PDB entry 8fli; Haack *et al.*, 2024). 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; Ast *et al.*, 2024). 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; it has been reported in only three articles (Desai *et al.*, 2017; Harper *et al.*, 2023; Ast *et al.*, 2024).

5.4. AlphaFold3 results on orphan structures

The RNA3DB_0 data set is mainly composed of structures without any hit in the Rfam family, and thus contains orphan

structures. The results of *AlphaFold3* for this data set, as presented in Figs. 2 and 4, show an overall lower performance compared with the other data sets when there is no use of context. *AlphaFold3* performs slightly better than *RhoFold* for this data set (p -value = 0.015). When using context, *AlphaFold3* produces improved results compared with those without context (p -value < 10^{-19}).

We detail the two worst predictions for RNA3DB_0 and RNA3DB_0 (Context) from *AlphaFold3* in Fig. 5. We observe poor results in terms of metrics (high RMSD and MCQ values and low TM-score and INF-ALL) for the two structures without context. With context, *AlphaFold3* seems to understand that the predictions are not only helices but still fails in these two worst examples to predict the complex non-common folding of these RNAs. These structures also have a small number of nucleotides (81, 42, 45 and 58 nt), meaning that *AlphaFold3* might not fail because of long-range interactions. Instead, these structures do not have a known family and rely on a complex environment of other molecules. With context, *AlphaFold3* has a better chance of predicting the structural folding well, but the generalization is not always robust for structures without known families, even with small structures (as shown by the mean value of TM-score, which is less than 0.5, in Supplementary Table S4).

To further study the impact of context for the prediction of RNA structures, in Fig. 6 we report the differences per metric

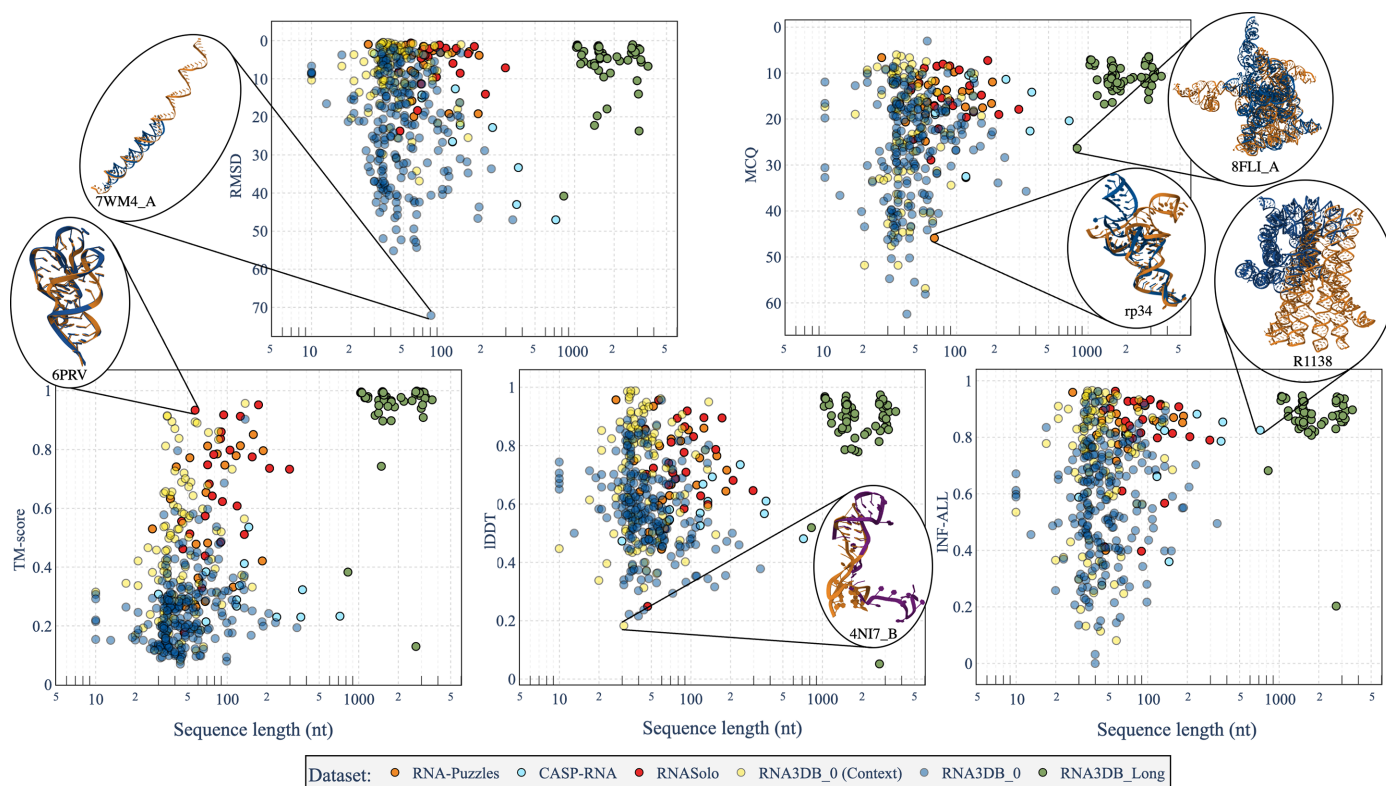


Figure 4 Dependence of metrics on the sequence length in predictions of *AlphaFold3* (Abramson *et al.*, 2024) on the five test sets. For some of the predictions, we show the predicted structure (in blue or purple if predicted using context) aligned with the native structure (in orange) using *US-align* (Zhang, Shine *et al.*, 2022). The metrics are RMSD, MCQ (Zok *et al.*, 2014), TM-score (Zhang & Skolnick, 2004; Zhang, Shine *et al.*, 2022), IDDT (Mariani *et al.*, 2013) and INF-ALL (Parisien *et al.*, 2009). RMSD and MCQ are reversed to have the best values near the top and the worst values at the bottom.

between predictions of *AlphaFold3* with and without context depending on the sequence length. Details of each metric value for each RNA are provided in Supplementary Fig. S7. For all metrics, there is an improvement on using context: 91.1% of structures with context have a 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 PDB entry 7wm4 (Sakaniwa *et al.*, 2023), the context effectively facilitates the identification of the correct scale for one half of the double helix. Similarly, for PDB entry 8bvj (Dendooven *et al.*, 2023), which features a discontinuity, the context enables *AlphaFold3* to accurately detect the discontinuities. However, this does not result in better alignment in terms of the IDDT metric. Incorporating contextual information significantly enhances the global alignment performance, as reflected by improvements in the RMSD, TM-score and IDDT metrics. This is followed by moderately smaller, but still notable, improvements in reproducing key RNA interactions (INF metric) and torsion angles (MCQ metric). Among the benchmarked models, the possibility of using

context in the prediction is only available with *AlphaFold3*. The other models are specialized for RNA and are not designed to process different molecules.

5.5. *AlphaFold3* results on key RNA interactions

To evaluate the ability of *AlphaFold3* to predict noncanonical interactions, we depict the scatter plots between non-Watson–Crick INF (INF-NWC) and Watson–Crick INF (INF-WC) in Fig. 7. The size of the points is proportional to the RMSD of the structures and thus to their global atom 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 *AlphaFold3* does not predict any of the non-Watson–Crick interactions (mostly for the RNA3DB data set). Examples of successful and missing non-Watson–Crick interactions are shown in the figure. For the results on stacking interactions, there are predictions where *AlphaFold3* does not predict the Watson–Crick interactions well but still predicts the stacking interactions. This can be explained by good

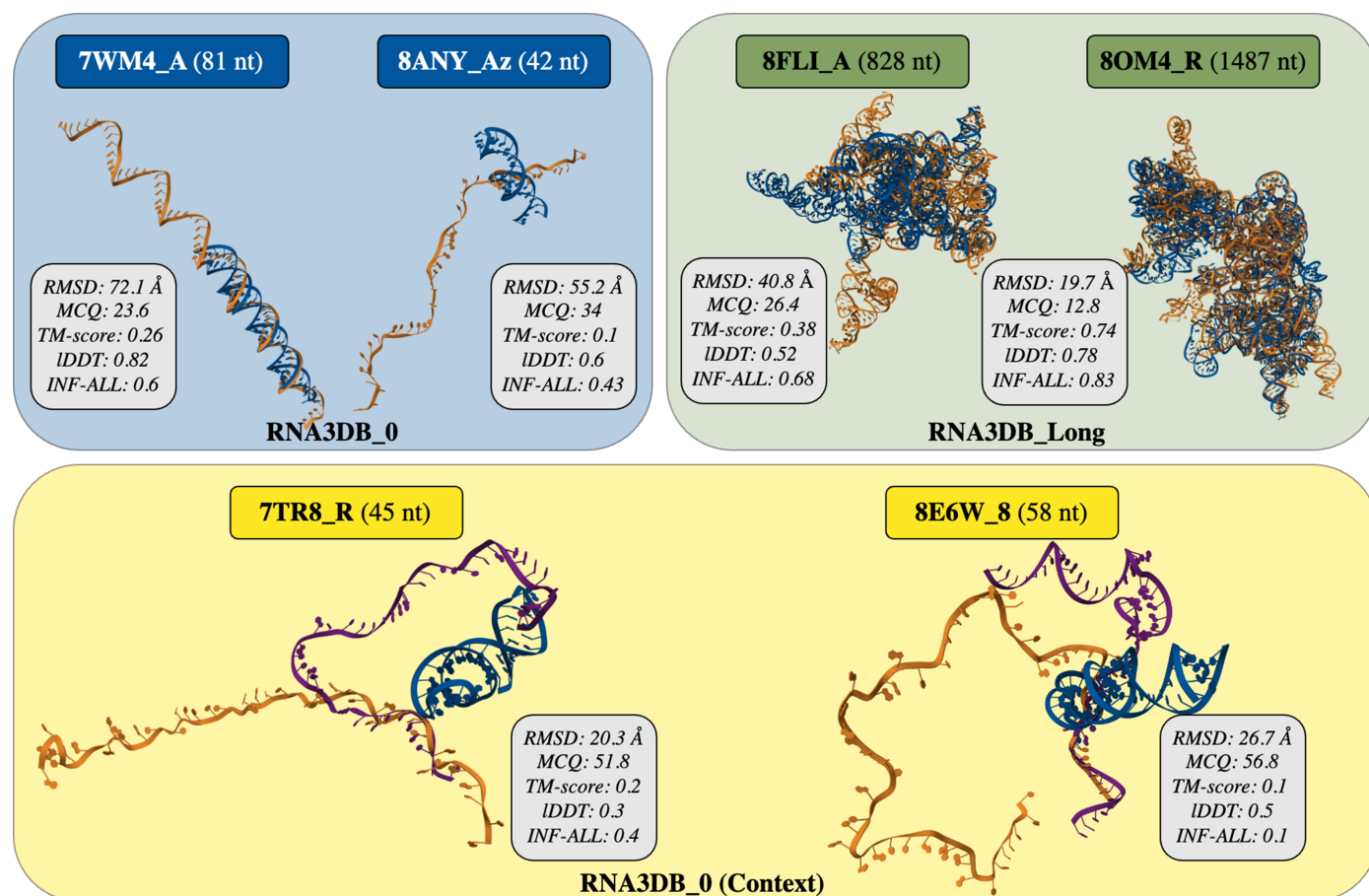


Figure 5

The worst two predicted structures (based on a cumulative sum of metrics) from *AlphaFold3* (Abramson *et al.*, 2024) for the RNA3DB_0 (left), RNA3DB_Long (right) and RNA3DB_0 (Context) (bottom) data sets. The RMSD, MCQ (Zok *et al.*, 2014), TM-score (Zhang & Skolnick, 2004; Zhang, Shine *et al.*, 2022), IDDT (Mariani *et al.*, 2013) and INF-ALL (Parisien *et al.*, 2009) are provided for each structure. The predictions from *AlphaFold3* (in blue) are aligned with the native structures (in orange) using *US-align* (Zhang, Shine *et al.*, 2022). The predictions of *AlphaFold3* with context [only for RNA3DB_0 (Context)] are provided in purple.

Table 2

Computation time for sequences of 27 nt (PDB entry 6y0y; E. Ennifar & E. Westhof, unpublished work) and 434 nt (PDB entry 7xd6; Luo *et al.*, 2023).

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 (27 nt)	Time (434 nt)
<i>RNAComposer</i> (Popenda <i>et al.</i> , 2012)	TP	1	3
<i>RhoFold</i> (Shen <i>et al.</i> , 2022)	DL	1	10
<i>trRosettaRNA</i> (Wang <i>et al.</i> , 2023)	DL	1	600
<i>RNAJP</i> † (Li & Chen, 2023)	Ab	120	900
<i>AlphaFold3</i> (Abramson <i>et al.</i> , 2024)	DL	2	5

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

skeleton predictions while lacking the base conformations that produce the WC interactions. Secondly, there is an increased correlation between INF-STACK and INF-WC: when *AlphaFold3* 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 to be influenced by whether the sequence can fold into the observed conformation. For instance, in Fig. 7, parts of PDB entry 8x9 chain B can fold, whereas others cannot.

To compare the key RNA interactions predicted from *AlphaFold3* with existing solutions, in Fig. 8 we present 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 data set are provided in Supplementary Table S6. We only show the results on these data sets as we only had complete predictions for each model for these three data sets. *AlphaFold3* has better values for each INF metric compared with 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 poor results in terms of key RNA interactions. Even if *AlphaFold3* outperforms other solutions for all of the INF metrics, the results for nWC interactions remain low (below 0.5), meaning that progress is still needed to reproduce RNA-specific interactions well.

5.6. Computation time

AlphaFold3 is a deep-learning method that has a complex architecture. Compared with 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 molecule (434 nt) for *RNAComposer* (Popenda *et al.*, 2012), *RhoFold* (Shen *et al.*, 2022), *trRosettaRNA* (Wang *et al.*, 2023), *RNAJP* (Li & Chen, 2023) and *AlphaFold3* (Abramson *et al.*, 2024) in Table 2. We report the computation time for the fastest methods, while the times for the rest of the methods are available in our previous work (Bernard *et al.*, 2024b). As we could only run *RNAJP* locally and each web server has different configurations, there is a bias in the comparison.

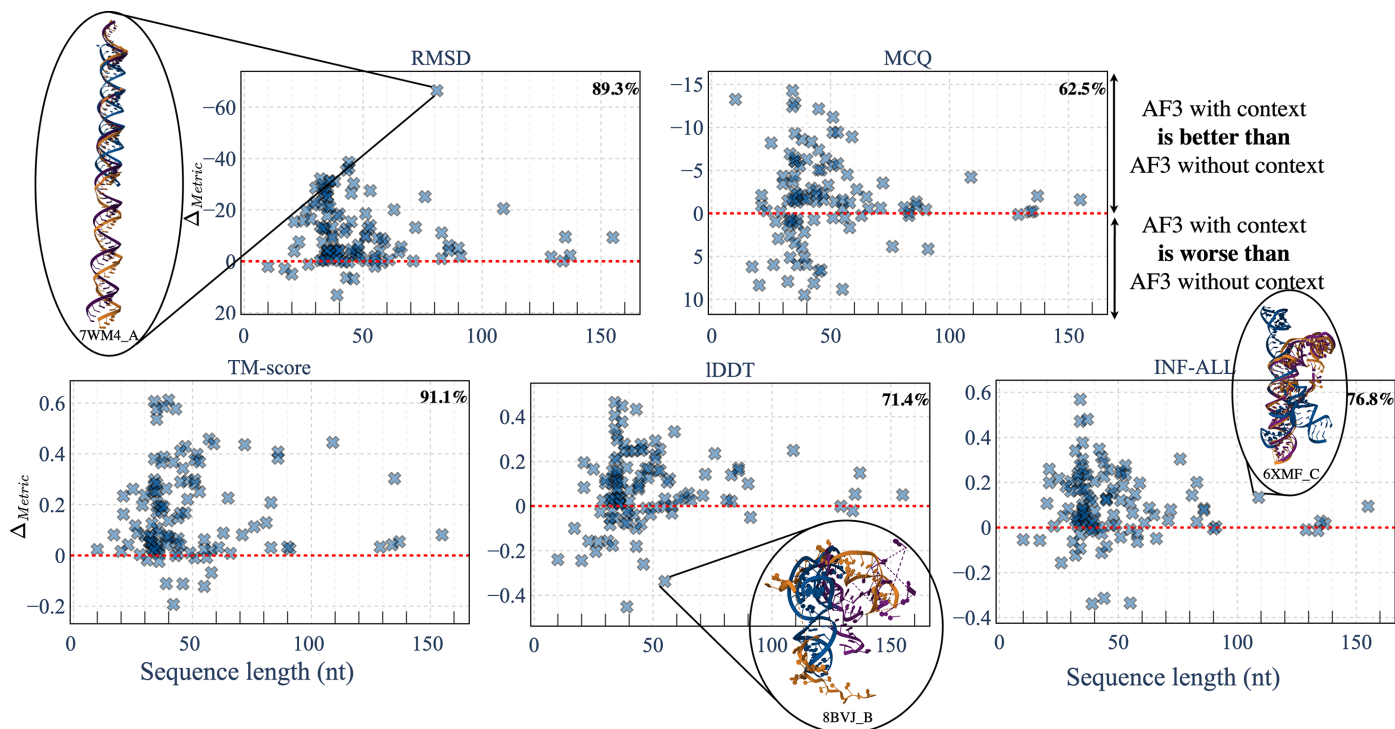


Figure 6

Difference per metric between results from *AlphaFold3* with context and without context for the common structures of the RNA3DB_0 data set depending on the sequence length. Regions above the red line correspond to structures where the results from *AlphaFold3* with context are better than those without context. We reversed the RMSD and MCQ metrics so that higher regions always depict the same behaviour. The percentage of cases where *AlphaFold3* with context outperforms predictions without context is reported in the top-right corner of each plot. Structures are reported with the native structure in orange, predictions with *AlphaFold3* without context in blue and with context in purple.

RNAComposer, *RhoFold* and *trRosettaRNA* all predict small RNA very quickly (in less than a minute), while *RNAJP* takes 2 h (with default parameters). For a structure with a longer sequence, it is *RNAComposer* that has the fastest computation time (around 3 min). The *ab initio* method *RNAJP* takes 15 h.

AlphaFold3 returns a prediction in around 5 min, which shows fast inference. For RNA with very long sequences (around 3000 nt), *AlphaFold3* take multiple hours to predict (and sometimes returns errors and needs to be run multiple times to obtain results).

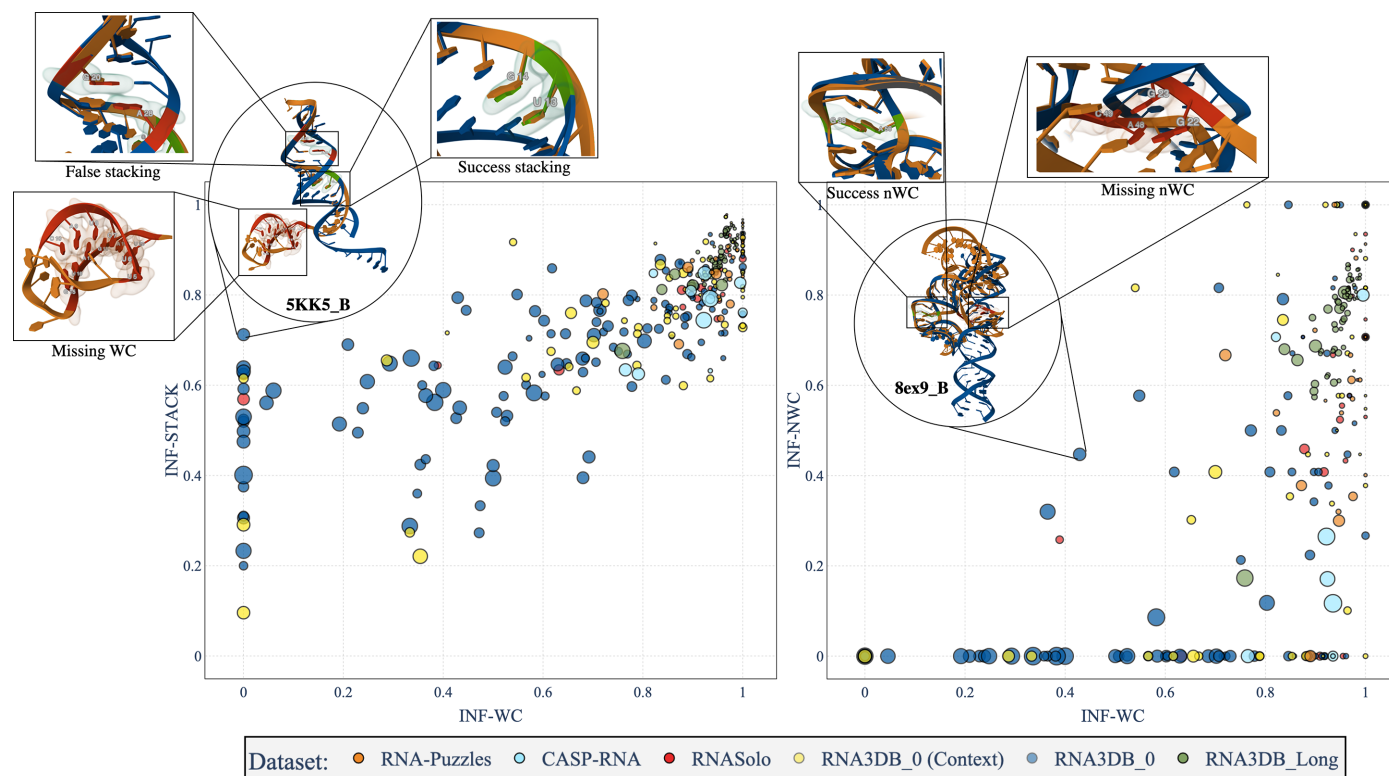


Figure 7 Link between INF Watson–Crick (WC) and non-Watson–Crick (nWC) and stacking (STACK) interactions in the predictions of *AlphaFold3* for our five test sets. The area of each point is proportional to the RMSD: the lower the better. Only structures with at least one non-Watson–Crick interaction are shown in the figures. An INF (Parisien *et al.*, 2009) value of 1 means accurate reproduction of key RNA interactions, while a value near 0 means that the structure does not reproduce the interactions. Left: INF stacking (INF-STOCK) depending on INF Watson–Crick (INF-WC) interactions. Right: INF non-Watson–Crick (INF-NWC) depending on INF Watson–Crick (INF-WC) interactions.

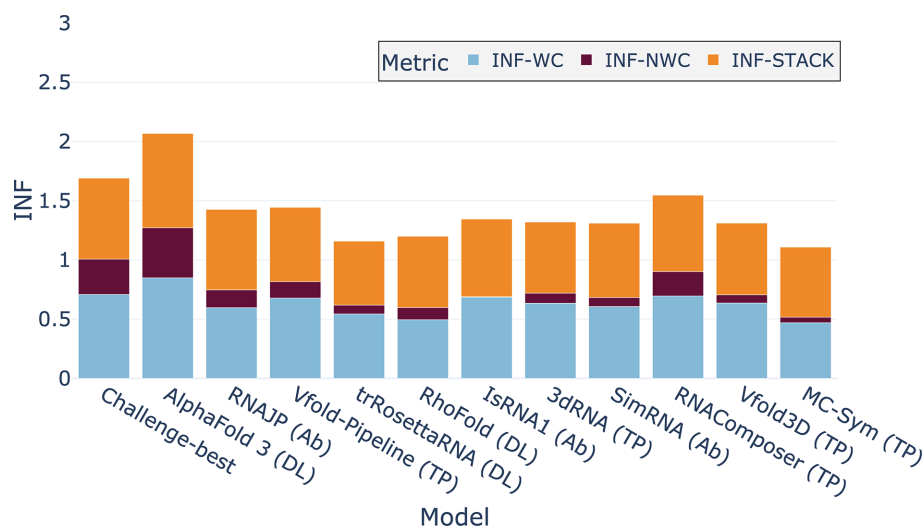


Figure 8 INF metrics for the different benchmarked models averaged over three test sets: RNA-Puzzles, CASP-RNA and RNASolo. INF metrics consider Watson–Crick (INF-WC), non-Watson–Crick (INF-NWC) and stacking (INF-STOCK) interactions.

6. Discussion

AlphaFold3 is a deep-learning method that has widened its scope to predict RNA structures (as well as other molecules) compared with its previous approach. Through our benchmark, we showed that *AlphaFold3* is a competitive method that outperforms most of the existing solutions. It yields better results for RNA-Puzzles and RNASolo, but remains outperformed by the best solutions from the CASP-RNA challenge.

AlphaFold3 has achieved good generalization properties for ribosomal structures (RNA3DB_Long data set). This shows bias from the existing data for RNA: most of the long RNA structures available in the PDB are of ribosomal-related RNA.

AlphaFold3 returns results with an overall good reproduction of key RNA interactions compared with existing solutions. It is also the best method to reproduce RNA torsion angles (best results in terms of MCQ), which was lacking in the existing deep-learning methods (Bernard *et al.*, 2024b).

There remain limitations that need to be addressed regarding the RNA folding problem. *AlphaFold3* does not reproduce all of the non-Watson–Crick interactions, which is essential for the stability of 3D RNA structures. Furthermore, *AlphaFold3* fails to predict structures from orphan families (RNA3DB_0 data set) without 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. *AlphaFold3* achieves better results when providing the context, but there remains a limitation of generalization for these orphan RNAs in 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 proposed by *AlphaFold3*, whereas multiple conformations are possible for RNA. *AlphaFold3*, while reducing the impact of the 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.

7. Conclusion

AlphaFold2 has had 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 *AlphaFold3*, has extended the model to predict all molecules in the PDB, such as ions, ligands, DNA and RNA.

Through an extensive benchmark on five different test sets, we have evaluated the quality of predictions of *AlphaFold3* 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 *AlphaFold3* 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 data

set while better reproducing key RNA interactions and torsion angles. It also returns predictions very quickly compared with *ab initio* or current template-based approaches [but does not exceed *RNAComposer* (Popenda *et al.*, 2012) for inference time].

For ribosomal long RNAs, *AlphaFold3* returns highly accurate predictions. This could be explained by its capability to generate structures from known families which have been seen in its training data. As there are not a lot of data available, it is difficult to find complex structures without any homologs to evaluate performances fairly.

Nonetheless, *AlphaFold3* has not yet covered RNA with the same success as it has proteins. Its new architecture allows the prediction of a wide range of molecules but remains limited and hardly predicts non-Watson–Crick interactions. It does not generalize well to orphan structures which are not related to any known RNA families. Prediction of these structures requires knowledge of the context, which it is possible to integrate with *AlphaFold3*.

The prediction of atom coordinates instead of base frames, as performed in *AlphaFold2*, allows the extension of predictions to a wide range of molecules but prevents the generalization of RNA-specific interactions. The lack of data is also a limitation that prevents the robustness of deep-learning methods in general, including *AlphaFold3*.

8. Related literature

The following references are cited in the supporting information for this article: Evans *et al.* (2021), Ji *et al.* (2023), RNA Consortium (2021), Sayers *et al.* (2023) and Sha *et al.* (2023).

Funding information

The following funding is acknowledged: Udopia (bursary No. UDOPIA-ANR-20-THIA-0013); Labex Digicosme (bursary No. ANR11LABEX0045DIGICOSME); Idex ParisSaclay (bursary No. ANR11IDEX000302); GENCI/IDRIS (grant No. AD011014250).

References

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Zémgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A. I., Cowie, A., Figurnov, M., Fuchs, F. B., Gladman, H., Jain, R., Khan, Y. A., Low, C. M. R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E. D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D. & Jumper, J. M. (2024). *Nature*, **630**, 493–500.
- Adamczyk, B., Antczak, M. & Szachniuk, M. (2022). *Bioinformatics*, **38**, 3668–3670.
- Anfinsen, C. B. (1973). *Science*, **181**, 223–230.
- Ast, T., Itoh, Y., Sadre, S., McCoy, J. G., Namkoong, G., Wengrod, J. C., Chicherin, I., Joshi, P. R., Kamenski, P., Suess, D. L., Amunts, A. & Mootha, V. K. (2024). *Mol. Cell*, **84**, 359–374.

- Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D. & DiMaio, F. (2024). *Nat. Methods*, **21**, 117–121.
- Becquey, L., Angel, E. & Tahi, F. (2021). *Bioinformatics*, **37**, 1218–1224.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Bernard, C., Postic, G., Ghannay, S. & Tahi, F. (2024a). *Brief. Bioinform.* **25**, bbae064.
- Bernard, C., Postic, G., Ghannay, S. & Tahi, F. (2024b). *NAR Genom. Bioinform.* **6**, lqae048.
- Boniecki, M. J., Lach, G., Dawson, W. K., Tomala, K., Lukasz, P., Soltysinski, T., Rother, K. M. & Bujnicki, J. M. (2016). *Nucleic Acids Res.* **44**, e63.
- Bottaro, S., Di Palma, F. & Bussi, G. (2014). *Nucleic Acids Res.* **42**, 13306–13314.
- Cao, S. & Chen, S.-J. (2011). *J. Phys. Chem. B*, **115**, 4216–4226.
- Černý, J., Božiková, P., Svoboda, J. & Schneider, B. (2020). *Nucleic Acids Res.* **48**, 6367–6381.
- Chen, K., Zhou, Y., Wang, S. & Xiong, P. (2023). *Proteins*, **91**, 1771–1778.
- Chheda, U., Pradeepan, S., Esposito, E., Streszak, S., Fernandez-Delgado, O. & Kranz, J. (2024). *J. Pharm. Sci.* **113**, 377–385.
- Cragolini, T., Laurin, Y., Derreumaux, P. & Pasquali, S. (2015). *J. Chem. Theory Comput.* **11**, 3510–3522.
- Cruz, J. A., Blanchet, M.-F., Boniecki, M., Bujnicki, J. M., Chen, S.-J., Cao, S., Das, R., Ding, F., Dokholyan, N. V., Flores, S. C., Huang, L., Lavender, C. A., Lisi, V., Major, F., Mikolajczak, K., Patel, D. J., Philips, A., Puton, T., Santalucia, J., Sijenyi, F., Hermann, T., Rother, K., Rother, M., Serganov, A., Skorupski, M., Soltysinski, T., Sripakdeevong, P., Tuszynska, I., Weeks, K. M., Waldsich, C., Wildauer, M., Leontis, N. B. & Westhof, E. (2012). *RNA*, **18**, 610–625.
- Das, R. & Baker, D. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 14664–14669.
- Das, R., Kretsch, R. C., Simpkin, A. J., Mulvaney, T., Pham, P., Rangan, R., Bu, F., Keegan, R. M., Topf, M., Rigden, D. J., Miao, Z. & Westhof, E. (2023). *Proteins*, **91**, 1747–1770.
- Dendooven, T., Sonnleitner, E., Bläsi, U. & Luisi, B. F. (2023). *EMBO J.* **42**, e111129.
- Desai, N., Brown, A., Amunts, A. & Ramakrishnan, V. (2017). *Science*, **355**, 528–531.
- Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Židek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J. & Hassabis, D. (2021). *bioRxiv*, 2021.10.04.463034.
- Flores, S. C., Wan, Y., Russell, R. & Altman, R. B. (2010). *Pac. Symp. Biocomput.*, pp. 216–227.
- Frellsen, J., Moltke, I., Thiim, M., Mardia, K. V., Ferkinghoff-Borg, J. & Hamelryck, T. (2009). *PLoS Comput. Biol.* **5**, e1000406.
- Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. (2012). *Bioinformatics*, **28**, 3150–3152.
- Gabb, H. A., Sanghani, S. R., Robert, C. H. & Prévost, C. (1996). *J. Mol. Graph.* **14**, 6–11.
- Gendron, P., Lemieux, S. & Major, F. (2001). *J. Mol. Biol.* **308**, 919–936.
- Gong, S., Zhang, C. & Zhang, Y. (2019). *Bioinformatics*, **35**, 4459–4461.
- Haack, D. B., Rudolfs, B., Zhang, C., Lyumkis, D. & Toor, N. (2024). *Nat. Struct. Mol. Biol.* **31**, 179–189.
- Hajdin, C., Ding, F., Dokholyan, N. & Weeks, K. (2010). *RNA*, **16**, 1340–1349.
- Harper, N. J., Burnside, C. & Klinge, S. (2023). *Nature*, **614**, 175–181.
- Jang, S. S., Dubnik, S., Hon, J., Hellenkamp, B., Lynall, D. G., Shepard, K. L., Nuckolls, C. & Gonzalez, R. L. J. (2023). *J. Am. Chem. Soc.* **145**, 402–412.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A. & Fung, P. (2023). *ACM Comput. Surv.* **55**, 248.
- Jonikas, M. A., Radmer, R. J., Laederach, A., Das, R., Pearlman, P., Herschlag, D. & Altman, R. B. (2009). *RNA*, **15**, 189–199.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohli, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. (2021). *Nature*, **596**, 583–589.
- Kagaya, Y., Zhang, Z., Ibtehaz, N., Wang, X., Nakamura, T., Punuru, P. D. & Kihara, D. (2025). *Nat. Commun.* **16**, 881.
- Kalvari, I., Nawrocki, E. P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., Griffiths-Jones, S., Toffano-Nioche, C., Gautheret, D., Weinberg, Z., Rivas, E., Eddy, S. R., Finn, R., Bateman, A. & Petrov, A. I. (2021). *Nucleic Acids Res.* **49**, D192–D200.
- Kerpedjiev, P., Höner zu Siederdisen, C. & Hofacker, I. L. (2015). *RNA*, **21**, 1110–1121.
- Krokhotin, A., Houlihan, K. & Dokholyan, N. V. (2015). *Bioinformatics*, **31**, 2891–2893.
- Li, J. & Chen, S.-J. (2023). *Nucleic Acids Res.* **51**, 3341–3356.
- Li, J., Zhang, S., Zhang, D. & Chen, S.-J. (2022). *Bioinformatics*, **38**, 4042–4043.
- Li, Y., Zhang, C., Feng, C., Pearce, R., Freddolino, P. L. & Zhang, Y. (2023). *Nat. Commun.* **14**, 5745.
- Liu, F. & Ou-Yang, Z.-C. (2005). *Biophys. J.* **88**, 76–84.
- Luo, B., Zhang, C., Ling, X., Mukherjee, S., Jia, G., Xie, J., Jia, X., Liu, L., Baulin, E. F., Luo, Y., Jiang, L., Dong, H., Wei, X., Bujnicki, J. M. & Su, Z. (2023). *Nat. Catal.* **6**, 298–309.
- Magnus, M., Antczak, M., Zok, T., Wiedemann, J., Łukasiak, P., Cao, Y., Bujnicki, J. M., Westhof, E., Szachniuk, M. & Miao, Z. (2020). *Nucleic Acids Res.* **48**, 576–588.
- Mariani, V., Biasini, M., Barbato, A. & Schwede, T. (2013). *Bioinformatics*, **29**, 2722–2728.
- Miao, Z., Adamiak, R. W., Antczak, M., Batey, R. T., Becka, A. J., Biesiada, M., Boniecki, M. J., Bujnicki, J. M., Chen, S.-J., Cheng, C. Y., Chou, F.-C., Ferré-D'Amaré, A. R., Das, R., Dawson, W. K., Ding, F., Dokholyan, N. V., Dunin-Horkawicz, S., Geniesse, C., Kappel, K., Kladwang, W., Krokhotin, A., Łach, G. E., Major, F., Mann, T. H., Magnus, M., Pachulska-Wieczorek, K., Patel, D. J., Piccirilli, J. A., Popena, M., Purzycka, K. J., Ren, A., Rice, G. M., Santalucia, J., Sarzynska, J., Szachniuk, M., Tandon, A., Trausch, J. J., Tian, S., Wang, J., Weeks, K. M., Williams, B., Xiao, Y., Xu, X., Zhang, D., Zok, T. & Westhof, E. (2017). *RNA*, **23**, 655–672.
- Miao, Z., Adamiak, R. W., Antczak, M., Boniecki, M. J., Bujnicki, J., Chen, S.-J., Cheng, C. Y., Cheng, Y., Chou, F.-C., Das, R., Dokholyan, N. V., Ding, F., Geniesse, C., Jiang, Y., Joshi, A., Krokhotin, A., Magnus, M., Mailhot, O., Major, F., Mann, T. H., Piątkowski, P., Pluta, R., Popena, M., Sarzynska, J., Sun, L., Szachniuk, M., Tian, S., Wang, J., Wang, J., Watkins, A. M., Wiedemann, J., Xiao, Y., Xu, X., Yesselman, J. D., Zhang, D., Zhang, Y., Zhang, Z., Zhao, C., Zhao, P., Zhou, Y., Zok, T., Żyła, A., Ren, A., Batey, R. T., Golden, B. L., Huang, L., Lilley, D. M., Liu, Y., Patel, D. J. & Westhof, E. (2020). *RNA*, **26**, 982–995.
- Miao, Z., Adamiak, R. W., Blanchet, M. F., Boniecki, M., Bujnicki, J. M., Chen, S.-J., Cheng, C., Chojnowski, G., Chou, F.-C., Cordero, P., Cruz, J. A., Ferré-D'Amaré, A. R., Das, R., Ding, F., Dokholyan, N. V., Dunin-Horkawicz, S., Kladwang, W., Krokhotin, A., Łach, G., Magnus, M., Major, F., Mann, T. H., Masquida, B., Matelska, D., Meyer, M., Peselis, A., Popena, M., Purzycka, K. J., Serganov, A., Stasiewicz, J., Szachniuk, M., Tandon, A., Tian, S., Wang, J., Xiao, Y., Xu, X., Zhang, J., Zhao, P., Zok, T. & Westhof, E. (2015). *RNA*, **21**, 1066–1084.
- Olechnovič, K., Kulberkytė, E. & Venclovas, C. (2013). *Proteins*, **81**, 149–162.
- Parisien, M., Cruz, J., Westhof, E. & Major, F. (2009). *RNA*, **15**, 1875–1885.

- Parisien, M. & Major, F. (2008). *Nature*, **452**, 51–55.
- Pearce, R., Omenn, G. S. & Zhang, Y. (2022). *bioRxiv*, 2022.05.15.491755.
- Popenda, M., Szachniuk, M., Antczak, M., Purzycka, K. J., Lukasiak, P., Bartol, N., Blazewicz, J. & Adamiak, R. W. (2012). *Nucleic Acids Res.* **40**, e112.
- Qiang, X.-W., Zhang, C., Dong, H.-L., Tian, F.-J., Fu, H., Yang, Y.-J., Dai, L., Zhang, X.-H. & Tan, Z.-J. (2022). *Phys. Rev. Lett.* **128**, 108103.
- RNAcentral Consortium (2021). *Nucleic Acids Res.* **49**, D212–D220.
- Rother, M., Rother, K., Puton, T. & Bujnicki, J. M. (2011). *Nucleic Acids Res.* **39**, 4007–4022.
- Sakaniwa, K., Fujimura, A., Shibata, T., Shigematsu, H., Ekimoto, T., Yamamoto, M., Ikeguchi, M., Miyake, K., Ohto, U. & Shimizu, T. (2023). *Nat. Commun.* **14**, 164.
- Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Farrell, C. M., Feldgarden, M., Fine, A. M., Funk, K., Hatcher, E., Kannan, S., Kelly, C., Kim, S., Klimke, W., Landrum, M. J., Lathrop, S., Lu, Z., Madden, T. L., Malheiro, A., Marchler-Bauer, A., Murphy, T. D., Phan, L., Pujar, S., Rangwala, S. H., Schneider, V. A., Tse, T., Wang, J., Ye, J., Trawick, B. W., Pruitt, K. D. & Sherry, S. T. (2023). *Nucleic Acids Res.* **51**, D29–D38.
- Schneider, B., Sweeney, B. A., Bateman, A., Cerny, J., Zok, T. & Szachniuk, M. (2023). *Nucleic Acids Res.* **51**, 9522–9532.
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K. & Hassabis, D. (2020). *Nature*, **577**, 706–710.
- Sha, C. M., Wang, J. & Dokholyan, N. V. (2023). *Biophys. J.* **122**, 444a.
- Shen, T., Hu, Z., Peng, Z., Chen, J., Xiong, P., Hong, L., Zheng, L., Wang, Y., King, I., Wang, S., Sun, S. & Li, Y. (2022). *arXiv:2207.01586*.
- Šulc, P., Romano, F., Ouldrige, T. E., Doye, J. P. K. & Louis, A. A. (2014). *J. Chem. Phys.* **140**, 235102.
- Szikszai, M., Magnus, M., Sanghi, S., Kadyan, S., Bouatta, N. & Rivas, E. (2024). *J. Mol. Biol.* **436**, 168552.
- Wadley, L. M., Keating, K. S., Duarte, C. M. & Pyle, A. M. (2007). *J. Mol. Biol.* **372**, 942–957.
- Wang, W., Feng, C., Han, R., Wang, Z., Ye, L., Du, Z., Wei, H., Zhang, F., Peng, Z. & Yang, J. (2023). *Nat. Commun.* **14**, 7266.
- Watkins, A. M., Rangan, R. & Das, R. (2020). *Structure*, **28**, 963–976.
- Westhof, E. & Fritsch, V. (2000). *Structure*, **8**, R55–R65.
- Wiedemann, J., Zok, T., Milostan, M. & Szachniuk, M. (2017). *BMC Bioinformatics*, **18**, 456.
- Xu, X. & Chen, S.-J. (2017). *J. Phys. Chem. B*, **122**, 5327–5335.
- Yamagami, R., Sieg, J. P. & Bevilacqua, P. C. (2021). *Biochemistry*, **60**, 2374–2386.
- Zemla, A., Venclovas, C., Moulton, J. & Fidelis, K. (1999). *Proteins*, **37**, 22–29.
- Zhang, C., Shine, M., Pyle, A. M. & Zhang, Y. (2022). *Nat. Methods*, **19**, 1109–1115.
- Zhang, D., Chen, S.-J. & Zhou, R. (2021). *J. Phys. Chem. B*, **125**, 11907–11915.
- Zhang, D., Li, J. & Chen, S.-J. (2021). *J. Chem. Theory Comput.* **17**, 1842–1857.
- Zhang, Y. & Skolnick, J. (2004). *Proteins*, **57**, 702–710.
- Zhang, Y., Wang, J. & Xiao, Y. (2022). *J. Mol. Biol.* **434**, 167452.
- Zhou, L., Wang, X., Yu, S., Tan, Y.-L. & Tan, Z.-J. (2022). *Biophys. J.* **121**, 3381–3392.
- Zhu, Y., Zhu, L., Wang, X. & Jin, H. (2022). *Cell Death Dis.* **13**, 644.
- Zok, T., Popenda, M. & Szachniuk, M. (2014). *Cent. Eur. J. Oper. Res.* **22**, 457–473.