

A comprehensive evaluation of RNA secondary structure prediction methods

Xinlong Chen(陈昕龙), En Lou(娄恩), Zouchenyu Zhou(周邹辰毓), Ya-Lan Tan(谭雅岚), and Zhi-Jie Tan(谭志杰)
Citation: Chin. Phys. B, 2025, 34 (**8**): 088710. DOI: 10.1088/1674-1056/adea9c
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Analysis of spatiotemporal dynamic patterns of gene expression during mouse embryonic development based on Moran's I and spatial transcriptomics

Qi-Chao Li(李啟超), Hai Lin(林海), Peng Wang(王鹏), Qiutong Dong(董秋彤), Kun Wang(王坤), Jian-Wei Shuai(帅建伟), and Fang-Fu Ye(叶方富)
Chin. Phys. B, 2025, 34 (**8**): 088703. DOI: 10.1088/1674-1056/ade427

Multiscale structural complexity analysis of neuronal activity in suprachiasmatic nucleus: Insights from tetrodotoxin-induced disruptions

Ping Wang(王萍), Changgui Gu(顾长贵), and Huijie Yang(杨会杰)
Chin. Phys. B, 2025, 34 (**4**): 048701. DOI: 10.1088/1674-1056/adaccd

Heterogeneous hydration patterns of G-quadruplex DNA

Cong-Min Ji(祭聪敏), Yusong Tu(涂育松), and Yuan-Yan Wu(吴园燕)
Chin. Phys. B, 2023, 32 (**2**): 028702. DOI: 10.1088/1674-1056/ac7452

RNAGCN: RNA tertiary structure assessment with a graph convolutional network

Chengwei Deng(邓成伟), Yunxin Tang(唐蕴芯), Jian Zhang(张建), Wenfei Li(李文飞), Jun Wang(王骏), and Wei Wang(王炜)
Chin. Phys. B, 2022, 31 (**11**): 118702. DOI: 10.1088/1674-1056/ac8ce3

Improving RNA secondary structure prediction using direct coupling analysis

Xiaoling He(何小玲), Jun Wang(王军), Jian Wang(王剑), Yi Xiao(肖奕)
Chin. Phys. B, 2020, 29 (**7**): 078702. DOI: 10.1088/1674-1056/ab889d

---

SPECIAL TOPIC — A celebration of the 90th Anniversary of the Birth of Bolin Hao

A comprehensive evaluation of RNA secondary structure prediction methods

Xinlong Chen(陈昕龙)[1], En Lou(娄恩)[1], Zouchenyu Zhou(周邹辰毓)[1], Ya-Lan Tan(谭雅岚)[2,‡], and Zhi-Jie Tan(谭志杰)[1,†]

[1] *Department of Physics and Key Laboratory of Artificial Micro & Nano-structures of Ministry of Education, School of Physics and Technology, Wuhan University, Wuhan* 430072, *China*

[2] *School of Bioengineering and Health & Research Center of Nonlinear Science, Wuhan Textile University, Wuhan* 430200, *China*

(Received 11 June 2025; revised manuscript received 26 June 2025; accepted manuscript online 2 July 2025)

RNAs have important biological functions and the functions of RNAs are generally coupled to their structures, especially their secondary structures. In this work, we have made a comprehensive evaluation of the performances of existing top RNA secondary structure prediction methods, including five deep-learning (DL) based methods and five minimum free energy (MFE) based methods. First, we made a brief overview of these RNA secondary structure prediction methods. Afterwards, we built two rigorous test datasets consisting of RNAs with non-redundant sequences and comprehensively examined the performances of the RNA secondary structure prediction methods through classifying the RNAs into different length ranges and different types. Our examination shows that the DL-based methods generally perform better than the MFE-based methods for RNAs with long lengths and complex structures, while the MFE-based methods can achieve good performance for small RNAs and some specialized MFE-based methods can achieve good prediction accuracy for pseudoknots. Finally, we provided some insights and perspectives in modeling RNA secondary structures.

Keywords: RNA secondary structure prediction, computational methods, comprehensive evaluation, traditional methods, deep-learning-based methods

PACS: 87.10.Vg, 87.15.bg, 87.14.gn

DOI: 10.1088/1674-1056/adea9c **CSTR:** 32038.14.CPB.adea9c

1. Introduction

Ribonucleic acids (RNAs) are a kind of biological macromolecules with important functions. An RNA is composed of a sequence of four nucleotides — adenine (A), uracil (U), cytosine (C), and guanine (G), and it generally folds into its specific secondary and tertiary structures through base pairing/stacking and tertiary interactions,[1,2] generally with the aid of the binding of ions or ligands.[3–5] The structures of RNAs enable their diverse functions, including protein synthesis,[6] catalysis,[7] and gene regulation.[8]

RNA structure folding is generally hierarchical and RNA secondary structures can exist stably due to the relatively strong base pairing/stacking interactions.[4,9–11] RNA secondary structures are especially important since some RNAs execute their biological functions at the secondary level such as RNA silencing and RNA interference.[12] Furthermore, RNA secondary structures can form a substantial basis for accurately modeling RNA 3-dimensional (3D) structures based on either physical principles[13–24] or fragment assembly.[25–31] Therefore, the knowledge of RNA secondary structures is crucial for understanding RNA functions and related RNA medical applications as well as RNA 3D structure modeling.[32–35]

Although experimental methods such x-ray crystallography, nuclear magnetic resonance (NMR), and cryo-EM have been employed to determine RNA 3D structures which can be conveniently converted to secondary structures,[4,34] the determined 3D structures of RNAs are rather limited possibly due to the high time and facility cost.[36,37] For example, until May 23, 2025, there are only 1981 3D structures of isolated RNAs and 5381 structures of RNAs complexed with other molecules.[36] Consequently, it is highly desired to develop computational models for predicting RNA secondary structures at high precision and low cost.

During the past decades, various computational methods have been developed to predict RNA secondary structures through various principles, including traditional methods[38–54] and recently developed machine learning (ML)/deep learning (DL)-based methods.[55–66] The traditional methods can be classified into minimum free energy (MFE)-based methods and multiple sequence alignment (MSA)-based ones. The MFE-based methods predict RNA secondary structures through calculating the free energies of possible conformations generated from dynamic programing or other sampling algorithms and the predicted structure is that

†Corresponding author. E-mail: zjtan@whu.edu.cn

‡Corresponding author. E-mail: yltan@wtu.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

with the lowest free energy.[43,47,49] The MSA-based methods are based on the evolution principle that homologous sequences tend to conserve structural features during evolution, and the accuracy of the MSA-based methods relies on the involved sequence databases for alignment.[67–69] Very recently, ML/DL-based methods such as SPOT-RNA,[55] MXfold2,[56] and BPfold[64] have been developed through involving large datasets and various networks to learn RNA structural features including base-pairing patterns and contact maps, and the prediction accuracy of these methods for RNA secondary structure prediction has been greatly improved by recent advances in deep learning technologies.[70–75] However, there is still a lack of a comprehensive understanding on the prediction performance of existing methods for RNA secondary structures.

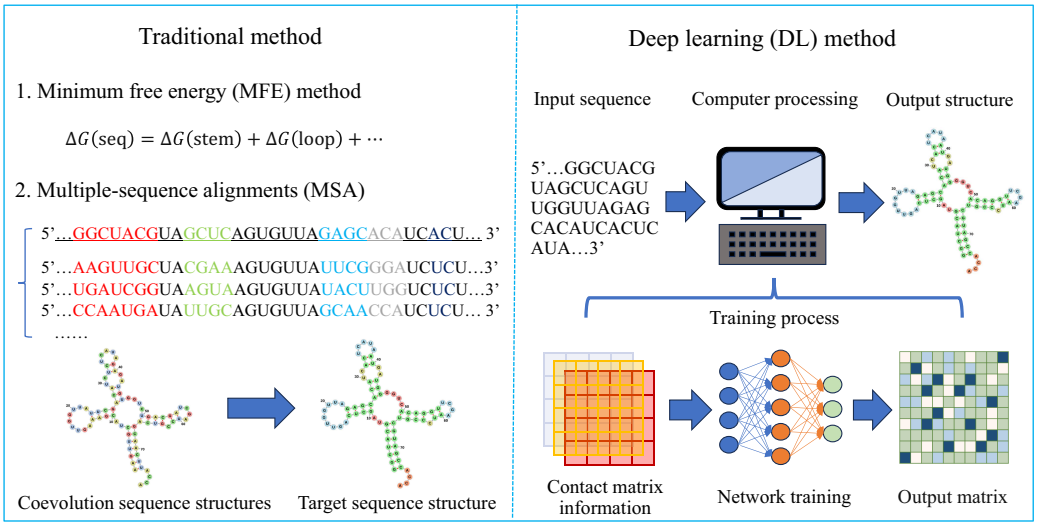


Fig. 1. An illustration for two types of RNA secondary structure prediction methods: traditional methods including minimum free energy (MFE)-based methods and multiple-sequence alignment (MSA)-based methods, and deep learning (DL)-based methods. The secondary structures were displayed by ViennaRNA web server.[49]

In this work, we made a comprehensive evaluation on the performance of several top RNA secondary structure prediction methods in different types. Specifically, we selected Mfold,[39,40,43] ViennaRNA,[41,49] RNAstructure,[47,76] IPknot,[48] and Knotty[51] as the typical ones of MFE-based methods, and we selected SPOT-RNA,[55] MXfold2,[56] REDfold,[60] and BPfold[64] as the typical ones of DL-based methods. Additionally, we also included AlphaFold3[62] for an evaluation through dissolving predicted tertiary structures to secondary ones.

2. Test dataset and metrics for evaluation

2.1. Test dataset

For a comprehensive evaluation, we constructed a benchmark dataset based on RNA 3D structures downloaded from Protein Data Bank (PDB) database,[36] and the secondary structures for the RNAs were obtained through dissolving the 3D structures with the software of DSSR.[77] Practically, similar to our previous works,[78–80] we downloaded all the PDB files of isolated RNAs (until February 27, 2025), and converted the 3D structures into the corresponding secondary structures with DSSR.[77] Afterwards, we removed the sequences with low base-pairing density of $< 20\%$ and lengths of < 20 nts. Moreover, we removed redundant sequences with sequence similarity $> 80\%$ by using CD-HIT-EST.[81] Finally, we obtained a dataset containing 514 RNA structures with length ≥ 20 nts and ≤ 3119 nts, sequence similarity $< 80\%$ and base-pairing density $> 20\%$, and the dataset was called as test set I. To test the prediction robustness of DL-based methods, especially Alphafold3 which has exceptionally excellent performance for test set I, we built test set II through choosing the RNAs whose 3D structures were released after the release date of AlphaFold3 (May 8, 2024) from test set I, and test set II is composed of 11 newly released RNAs. The PDBs of test set I and test set II are listed in Table S1 in the supplementary material.

For a comprehensive evaluation, we examined the performances of the secondary structure prediction methods along two dimensions: RNA length and RNA structure type. First, we categorized test set I into six length ranges: 20–50 nts (195 RNAs), 51–100 nts (198 RNAs), 101–200 nts (79 RNAs), 201–500 nts (25 RNAs), 501–1000 nts (7 RNAs), and >1000 nts (10 RNAs). Second, we also classified test set I into different structure types: RNAs with 3-way junction (42 RNAs), RNAs with 4-way junction (61 RNAs), RNAs with 5-way junction (10 RNAs), and RNAs with pseudoknot (197 RNAs). The length distributions for the different length and structure type categories have been shown in Figs. S1 and S2 in the supplementary material.

2.2. Evaluation metrics

We employed two complementary metrics: F1-score and Matthews correlation coefficient (MCC) to quantify the pre-

diction performance of various methods, and both of F1-score and MCC can be derived from the confusion matrix of base-pair predictions by[82]

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}; \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}; \tag{2}$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}; \tag{3}$$

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}}. \tag{4}$$

Here, TP and TN represent the numbers of correctly predicted pairing and un-pairing bases, while FP and FN denote the numbers of incorrectly predicted pairing and un-pairing bases. Thus, precision and recall describe the percentage of correctly predicted pairing bases in the total predicted ones and the percentage of correctly predicted pairing bases in the native ones, and F1-score serves as an integrated metric of precision and recall. MCC provides another important metric by integrating all four confusion matrix categories. Thus, F1-score emphasizes detection of true base-pairing interactions, and MCC better reflects overall structure facticity. This dual-metric approach enables robust evaluation of both base-pair identification accuracy and global structure prediction quality.

3. Overview of representative RNA secondary structure prediction methods

In the following, we briefly introduce the representative MFE-based methods and DL-based methods used in our evaluation. The former methods include Mfold,[39,40,43] ViennaRNA,[41,49] RNAstructure,[47,76] IPknot,[48] and Knotty,[51] and the latter methods include SPOT-RNA,[55] MXfold2,[56] REDfold,[60] and BPfold[64] and AlphaFold3.[62] Please see Table 1 for the features of these methods as well as the websites for online use and for package download.

Table 1. Existing major RNA secondary structure prediction methods.

Methods	Features[a)]	Year[b)]	Package link	Web server	Ref.
Mfold	MFE	2021	https://www.unafold.org/mfold/software/download-mfold.php	https://www.unafold.org/	[43]
ViennaRNA	MFE	2024	https://www.tbi.univie.ac.at/RNA/#	http://rna.tbi.univie.ac.at/	[49]
RNAstructure	MFE	2024	https://rna.urmc.rochester.edu/RNAstructure.html	https://rna.urmc.rochester.edu/RNAstructureWeb/	[47]
IPknot	MFE	2023	https://github.com/satoken/ipknot	https://ws.sato-lab.org/rtips/ipknot++/	[48]
Knotty	MFE	2024	https://github.com/TheCOBRALab/Knotty	/	[51]
SPOT-RNA	DL	2019	https://github.com/jaswindersingh2/SPOT-RNA/	https://sparks-lab.org/server/spot-rna/	[55]
SPOT-RNA2	DL	2021	https://github.com/jaswindersingh2/SPOT-RNA2/	https://sparks-lab.org/server/spot-rna2/	[57]
MXfold2	DL	2021	https://github.com/mxfold/mxfold2	http://www.dna.bio.keio.ac.jp/mxfold2/	[56]
REDfold	DL	2023	https://github.com/aky3100/REDfold	https://redfold.ee.ncyu.edu.tw	[60]
BPfold	DL	2024	https://github.com/heqin-zhu/BPfold	/	[64]
AlphaFold3	DL	2024	/	https://alphafoldserver.com/	[62]

[a)] Here, MFE and DL are the abbreviations of minimum free energy and deep learning, respectively.
[b)] This indicates the release years of the versions of these methods involved for the evaluation, although the early versions of some MFE-based models have been released for a long time.

3.1. MFE-based methods

Mfold Mfold was originally developed by Zuker *et al.* in the 1980s, and remains a widely-used tool for RNA secondary structure prediction based on free energy minimization principles.[39,40] The current web server of Mfold provides a comprehensive prediction for secondary structures and thermodynamics. Moreover, the server generates detailed outputs such as structure displaying, energy dot plots, and single-strand frequency plots, and supports user-defined constraints to guide the folding process.[43]

ViennaRNA ViennaRNA was originally introduced by Lorenz *et al.* in 1994 by implementing dynamic programming algorithms for RNA secondary structure prediction.[41] ViennaRNA calculates minimum free energies and equilibrium partition functions, providing detailed basepairing probabilities. ViennaRNA includes optimized implementations for both sequential and parallel computing environments, and uniquely addresses the inverse-folding problem for designing RNA sequences with specific structural properties.[49]

RNAstructure RNAstructure was initially developed by Mathews *et al.* in 1997, which incorporates thermodynamic parameters from the Turner group to predict RNA secondary structures through multiple approaches, including minimum free energy, maximum expected accuracy, and partition function calculations.[76] RNAstructure provides detailed energy analysis and structure annotation features, supporting both a single sequence analysis and comparative approaches using multiple sequences.[47]

IPknot IPknot was presented by Sato *et al.* in 2011, which is a novel method for predicting RNA secondary structures with pseudoknots using integer programming.[48] IPknot optimizes structure prediction by maximizing expected accuracy and includes an iterative refinement algorithm to enhance

prediction accuracy. IPknot supports both single sequence and aligned sequences as input, providing robust comparative analysis capabilities. IPknot may be a reliable model with efficiency and accuracy, especially for predicting complex pseudoknot structures.

Knotty Knotty was developed by Jabbari *et al*. in 2018, which is an advanced algorithm for predicting RNA secondary structures with pseudoknots.[51] Knotty builds on the CCJ algorithm[83] and employs sparsification techniques[84,85] to significantly reduce space complexity while maintaining high prediction accuracy. Knotty uses the HotKnots DP09 energy model[86] to improve prediction quality and supports complex pseudoknot structures, including kissing hairpins. Similar to the MFE-based methods mentioned above, Knotty also offers cross-platform support and detailed energy analysis. However, Knotty is inapplicable to sequences of length $>$ 500 nts.

3.2. DL-based methods

SPOT-RNA SPOT-RNA was proposed by Singh *et al*. in 2019, which utilizes an ensemble of two-dimensional deep neural networks, i.e., ResNets and long short-term memory (LSTM) networks.[55] SPOT-RNA employs transfer learning, where the model is initially trained on a large dataset bpRNA[87] and then fine-tuned with high-resolution RNA structures. The combination of these techniques allows SPOT-RNA to achieve a significant improvement in predicting all base pairs for RNAs, including non-canonical and non-nested base pairs. SPOT-RNA is not applicable for sequences $>$1500 nts.

SPOT-RNA2 SPOT-RNA2 was introduced by Singh *et al*. in 2021 based on SPOT-RNA, through integrating evolutionary profiles, mutational coupling, and transfer learning for improving RNA secondary structure prediction.[57] SPOT-RNA2 predicts RNA contact maps by direct mutational coupling technique GREMLIN.[88] Compared to SPOT-RNA, it leverages position specific score matrix (PSSM) and direct coupling analysis (DCA) to improve accuracy for complex structures like non-canonical base pairs and pseudoknots. However, since SPOT-RNA2 involves the calculation of MSA information, SPOT-RNA2 is significantly less efficient than SPOT-RNA and requires a lot of storage space (more than 2 TB). Thus, we did not examine SPOT-RNA2 in this work.

MXfold2 MXfold2 was introduced by Sato *et al*. in 2021, which integrates deep learning predictions with Turner's nearest-neighbor free energy parameters.[56] This method employs a max-margin framework with thermodynamic regularization to ensure that the folding scores are close to the free energy calculations. By combining these approaches, MXfold2 minimizes overfitting while maintaining computational efficiency, resulting in robust and accurate predictions.

REDfold REDfold was developed by Chen and Chan in 2023, which involves a residual encoder–decoder network based on convolutional neural networks to learn both short and long-range dependencies within RNA sequences.[60] REDfold utilizes symmetric skip connections to efficiently propagate activation information across layers, enhancing the ability to capture complex interactions. REDfold is computationally efficient and can handle RNAs with pseudoknots.

BPfold BPfold was presented by Zhu *et al*. in 2024, which constructs a base pair motif library and uses a hybrid transformer and convolutional neural network architecture.[64] BPfold integrates base pair motif energy and base pair attention to improve generalizability and accuracy. Additionally, BPfold provides a confidence index to measure the reliability of predicted structures, and thus offers users a way to assess prediction quality.

AlphaFold3 AlphaFold3 was released by Google DeepMind in 2024, which represents a significant advancement in predicting 3D structures of biomolecules including proteins, nucleic acids, small molecules, ions, and modified residues.[62] AlphaFold3 employs a diffusion-based architecture with pairfromer module that predicts the joint structure of these complexes with high accuracy. AlphaFold3 represents a significant contribution to structural biology. AlphaFold3 was involved in the work through dissolving predicted 3D structures to secondary ones with DSSR.[77]

4. Results and discussion

In this work, first, we provided the overall performance of the above-introduced methods for RNA secondary structure prediction on test set I consisting of all redundant RNAs and test set II containing newly released redundant RNAs from the PDB database. Afterwards, we made a comprehensive evaluation on the methods along two dimensions of RNA length and RNA structure types for the RNAs in test set I. Our quantitative evaluations were based on the above-introduced metrics including F1-score and MCC. Since Knotty, REDfold, and SPOT-RNA are inapplicable to sequences of $>$ 500 nts, those of $>$ 1000 nts and those of $>$ 1500 nts respectively, the calculated metrics are the averages over those applicable RNAs.

4.1. Overall prediction performance

The average F1-scores and MCCs for all RNAs in test set I from the various methods have been shown in Fig. 2(a) and Table 2. The performances of the methods reflected by F1-score follow the order of AlphaFold3 (0.93) $\gg$ Knotty (0.80) $>$ MXfold2 (0.79) $>$ SPOT-RNA (0.77) $>$ IPknot (0.74) $>$ Mfold (0.71) $>$ ViennaRNA (0.70) $\sim$ RNAstructure (0.70) $\sim$ BPfold (0.70) $>$ REDfold (0.64). Simultaneously, the performances of the methods reflected by MCC follow the order of AlphaFold3 (0.82) $\gg$ Knotty (0.52) $\sim$ MXfold2 (0.52) $>$ BPfold (0.48) $>$ SPOT-RNA (0.47) $>$ IPknot (0.45) $>$ REDfold (0.39) $>$ Mfold (0.36) $>$ ViennaRNA (0.34) $\sim$ RNAstructure

(0.34). Thus, AlphaFold3 exhibits an apparently superior performance with an F1-score of 0.93 and an MCC of 0.82, compared with other MFE and DL-based methods. This outstanding performance of AlphaFold3 is also reflected by its high precision (0.90) and recall (0.95), indicating its exceptional capability in predicting true base pairs. Following AlphaFold3, Knotty and MXfold2 emerge as the second-class methods with visibly lower F1-scores ($\sim$ 0.80) and lower MCCs (0.52), respectively. Next, SPOT-RNA and IPknot have slightly lower performance than Knotty and MXfold2, while other methods such as BPfold, Mfold and RNAstructure have slightly lower performance than SPOT-RNA and IPknot. Here, it is noted that Knotty, REDfold and SPOT-RNA are applicable for RNAs with lengths of $\leq$ 500 nts, $\leq$ 1000 nts and $\leq$ 1500 nts, respectively.

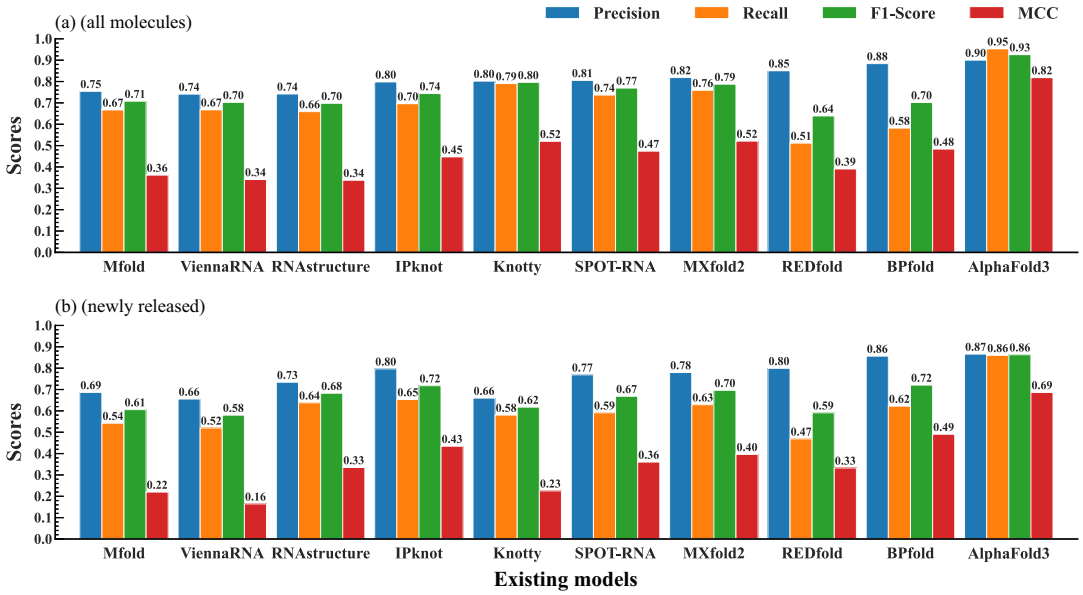


Fig. 2. The performances of different RNA secondary structure prediction methods on (a) test set I, and (b) test set II, which consisted of newly released RNAs after May 8, 2024 (the date when AlphaFold3 was published[62]). Since Knotty, REDfold, and SPOT-RNA are inapplicable for sequences of > 500 nts, those of > 1000 nts, and those of > 1500 nts, respectively, the shown values are the averages over those applicable RNAs.

Table 2. Performances of different methods on test set I. The metrics are bolded for the best values and underlined for the next-best ones by different methods.

Methods	Precision	Recall	MCC	F1-score
Mfold	0.7541	0.667	0.3617	0.7079
ViennaRNA	0.7410	0.6677	0.3411	0.7024
RNAstructure	0.7418	0.6591	0.3377	0.698
IPknot	0.7982	0.6962	0.4472	0.7437
Knotty	0.8019	<u>0.7904</u>	0.5200	<u>0.7961</u>
SPOT-RNA	0.8053	0.7367	0.4739	0.7694
MXfold2	0.8185	0.7589	<u>0.5210</u>	0.7876
REDfold	0.8504	0.5116	0.3907	0.6389
BPfold	<u>0.8839</u>	0.5819	0.4838	0.7018
AlphaFold3	**0.9001**	**0.9531**	**0.8182**	**0.9259**

For another interesting evaluation especially for AlphaFold3, we used test set II containing newly released structures after AlphaFold3.[62] The F1-scores and MCCs from various methods are shown in Fig. 2(b) and Table 3. The F1-scores from the different methods follow the order of AlphaFold3 (0.86) > BPfold (0.72) $\sim$ IPknot (0.72) > MXfold2 (0.70) > RNAstructure (0.68) > SPOT-RNA (0.67) > Knotty (0.62) > Mfold (0.61) > REDfold (0.59) > ViennaRNA (0.58), and the MCCs follow the order of AlphaFold3 (0.69) > BPfold (0.49) > IPknot (0.43) > MXfold2 (0.40) > SPOT-RNA (0.36) > RNAstructure (0.33) $\sim$ REDfold (0.33) > Knotty (0.23) > Mfold (0.22) > ViennaRNA (0.16). While AlphaFold3 maintains its leading position, the performance of AlphaFold3 decreases compared with that of test set I, and the performance superiority of AlphaFold3 also decreases compared with other methods. For example, when test set II is used instead of test set I, the F1-score and MCC of AlphaFold3 decrease from 0.93 to 0.86 and from 0.82 to 0.69 respectively, and simultaneously, the MCC difference between AlphaFold3 and the second-ranked method decreases from 0.3 to 0.2. Such decrease in the performance of AlphaFold3 may suggest that test set II may contain the structure features not fully captured in the training set of AlphaFold3. It is noted that BPfold keeps its consistent performance for test set I and test set II, and emerges as the second-ranked method. It is encouraging that several MFE-based methods such as IPknot and RNAstructure keep their consistent performance, reflected by their consistent F1-scores for test set I (0.74 and 0.70) and test set II (0.72 and 0.68). This may suggest the traditional principles can still play some special roles in RNA secondary structure predictions, in addition to modeling RNA thermodynamics and kinetics.[23,89–93]

Table 3. Performances of different methods on test set II with newly released RNA structures. The metrics are bolded for the best values and underlined for the next-best ones by different methods.

Methods	Precision	Recall	MCC	F1-score
Mfold	0.6864	0.5422	0.2198	0.6058
ViennaRNA	0.6555	0.5200	0.1645	0.5799
RNAstructure	0.7340	0.6378	0.3348	0.6825
IPknot	0.7967	<u>0.6533</u>	0.4344	0.7179
Knotty	0.6598	0.5803	0.2255	0.6175
SPOT-RNA	0.7688	0.5911	0.3603	0.6683
MXfold2	0.7796	0.6289	0.3962	0.6962
REDfold	0.7992	0.4689	0.3327	0.5910
BPfold	<u>0.8563</u>	0.6222	<u>0.4904</u>	<u>0.7207</u>
AlphaFold3	**0.8658**	**0.8600**	**0.6862**	**0.8629**

The computational efficiency is also important for a prediction method. We examined the computational costs for different methods on test set II except for AlphaFold3 which can only be used through the web server; see Table 4. Overall, the MFE-based methods except Knotty are more efficient than the DL-based methods. Here, it needs to be emphasized that the run time of the methods for an RNA is dependent on not only its sequence length but also its structure complexity, in addition to the computing machine.

Table 4. The total run time of different methods on test set II. The methods were examined on a PC machine with AMD Ryzen 9 7950X @ 4.5 GHz CPU.

Methods	Run time (s)
Mfold	7.5
ViennaRNA	0.2
RNAstructure	3.4
IPknot	1.6
Knotty	1292.7
SPOT-RNA	236.2
MXfold2	10.5
REDfold	53.4
BPfold	41.0

Therefore, based on the examinations on test set I with a large number of redundant RNAs and on test set II with a few newly released RNAs, AlphaFold3 exhibits an exceptionally outstanding performance compared with other methods. Moreover, MXfold2, SPOT-RNA and BPfold exhibit relatively reliable performance with mean F1-scores > 0.7 in the DL-based methods, and IPknot, Knotty, and RNAstructure also show relatively reliable performance with mean F1-scores 0.7 in the MFE-based methods. In the following, we would make a comprehensive analysis of the performances of the various methods in two dimensions: RNA length and RNA structure type.

4.2. Prediction performance versus RNA sequence length

To systematically evaluate the length-dependent performance of RNA structure prediction methods, we classified the test set I into six categories according to sequence lengths: 20–50 nts, 51–100 nts, 101–200 nts, 201–500 nts, 501–1000 nts, and $>$ 1000 nts. The number of RNAs varies in the length range of [20 nts, 3119 nts]; see the details in Subsection 2.1. The mean metrics from the ten methods for the categories in the different length ranges are shown in Figs. 3 and 4.

For the category with 195 RNAs in the range of 20–50 nts, all the ten methods demonstrated competitive performance; see Fig. 3(a) and Table S2 in the supplementary material. Specifically, the F1-scores from the different methods follow the order of BPfold (0.90) $\sim$ AlphaFold3 (0.90) $>$ MXfold2 (0.88) $\sim$ SPOT-RNA (0.88) $\sim$ Mfold (0.88) $\sim$ ViennaRNA (0.88) $\sim$ IPknot (0.88) $>$ RNAstructure (0.87) $\sim$ Knotty (0.87) $>$ REDfold (0.77), and the MCCs follow the order of BPfold (0.78) $>$ AlphaFold3 (0.77) $>$ MXfold2 (0.73) $>$ IPknot (0.72) $\sim$ Mfold (0.72) $\sim$ ViennaRNA (0.72) $>$ RNAstructure (0.71) $\sim$ SPOT-RNA (0.71) $>$ Knotty (0.70) $>$ REDfold (0.57). The MFE-based methods (Mfold, ViennaRNA, and RNAstructure) remain competitive for RNAs in this length range, suggesting their ability for small RNA structures and thermodynamics analyses.

For the category with 198 RNAs in the range of 51–100 nts, MFE-based methods began to show gradually declined performances, compared to DL-based methods and compared with those for the RNA in the range of 20–50 nts; see Fig. 3(b). Specifically, the F1-scores from the different methods follow the order of AlphaFold3 (0.92) $>$ BPfold (0.88) $>$ MXfold2 (0.86) $>$ SPOT-RNA (0.84) $>$ Knotty (0.81) $>$ IPknot (0.79) $\sim$ ViennaRNA (0.79) $>$ RNAstructure (0.78) $>$ Mfold (0.77) $>$ REDfold (0.75), and the MCCs follow the order of AlphaFold3 (0.81) $>$ BPfold (0.73) $>$ MXfold2 (0.69) $>$ SPOT-RNA (0.60) $>$ Knotty (0.57) $>$ REDfold (0.54) $\sim$ IPknot (0.54) $>$ RNAstructure (0.51) $\sim$ ViennaRNA (0.51) $>$ Mfold (0.49); see Fig. 3(b) and Table S3 in the supplementary material. It is noted that the DL-based methods began to overwhelm the MFE-based methods and AlphaFold3 began to exhibit the exceptional performance for RNAs in the length range.

For the category with 79 RNAs in the range of 101–200 nts, the F1-scores from the ten methods follow the order of AlphaFold3 (0.89) $>$ BPfold (0.83) $>$ MXfold2 (0.80) $>$ IPknot (0.77) $>$ Mfold (0.76) $\sim$ RNAstructure (0.76) $\sim$ ViennaRNA (0.76) $>$ SPOT-RNA (0.75) $>$ Knotty (0.73) $>$ REDfold (0.58), and the MCCs follow the order of AlphaFold3 (0.74) $>$ BPfold (0.63) $>$ MXfold2 (0.56) $>$ IPknot (0.51) $>$ Mfold (0.47) $>$ RNAstructure $>$ (0.46) $>$ SPOT-RNA (0.45) $>$ ViennaRNA (0.44) $>$ Knotty (0.39) $>$ REDfold (0.34); see Fig. 3(c) and Table S4 in the supplementary material. For the category with 25 RNAs in the range of 201–500 nts, the F1-scores from the ten methods follow the order of AlphaFold3 (0.88) $\gg$ MXfold2 (0.77) $\sim$ BPfold (0.77) $>$ Knotty (0.75) $>$ IPknot (0.73) $>$ RNAstructure (0.71) $>$ ViennaRNA (0.70) $>$ Mfold (0.68) $>$ SPOT-RNA (0.66) $>$ REDfold (0.45), and the MCCs follow the order of AlphaFold3 (0.69) $\gg$ BPfold (0.48)

> MXfold2 (0.46) > Knotty (0.38) > IPknot (0.37) > SPOT-RNA (0.32) > RNAstructure (0.31)> ViennaRNA (0.30) > Mfold (0.26) > REDfold (0.14); see Fig. 4(a) and Table S5 in the supplementary material. For intermediate-length RNAs (in the length range of 101–500 nts), AlphaFold3 exhibited visibly superior prediction accuracy, while BPfold and MXfold2 showed reliable but marginally lower performance than AlphaFold3. The MFE-based methods showed comparable performances to each other but are consistently outperformed by the DL-based methods.

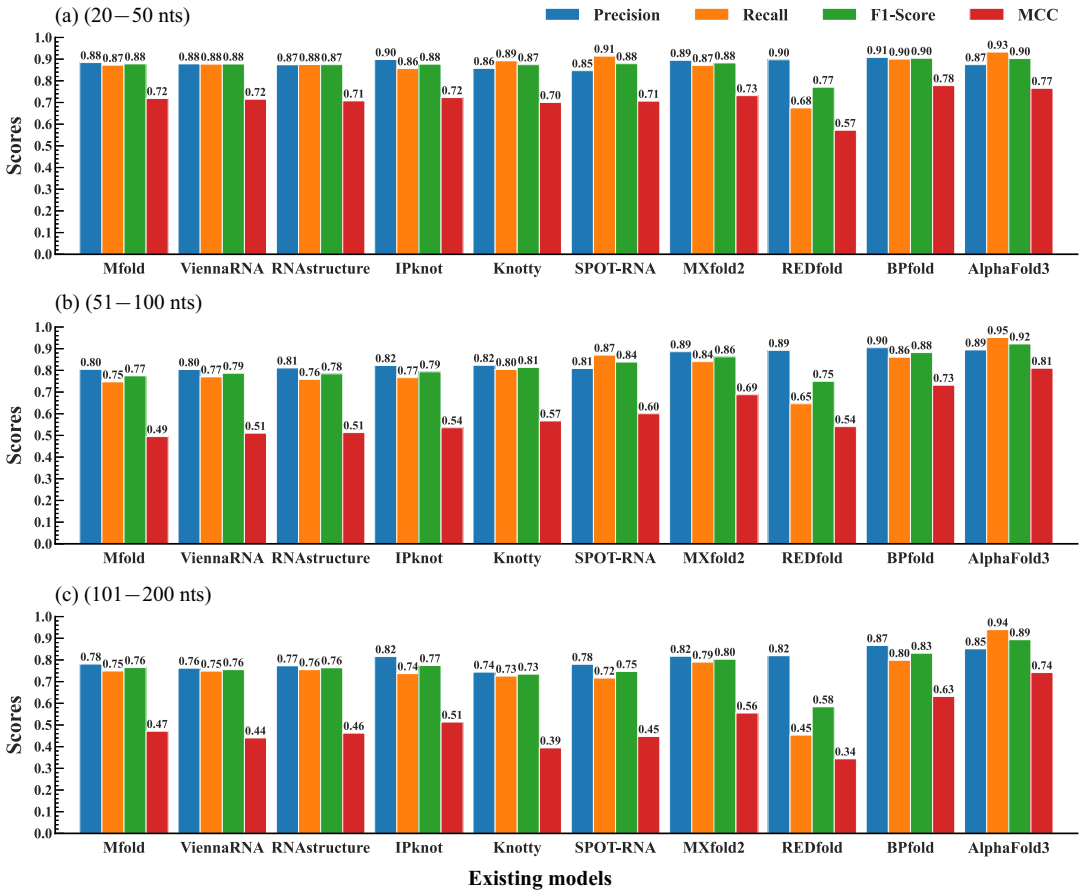


Fig. 3. The performances of different secondary structure prediction methods on short RNAs with sequence length ≤ 200 nts in test set I: (a) for RNA length range of 20–50 nts; (b) for RNA length range of 51–100 nts; (c) for RNA length range of 101–200 nts. Since Knotty, REDfold, and SPOT-RNA are inapplicable for sequences of > 500 nts, those of > 1000 nts, and those of > 1500 nts, respectively, the shown values are the averages over those applicable RNAs.

For the category with 7 large RNAs in the range of 501–1000 nts, the F1-scores from the different methods follow the order of AlphaFold3 (0.93) ≫ IPknot (0.83) ∼ MXfold2 (0.83) > Mfold (0.81) ∼ ViennaRNA (0.81) > RNAstructure (0.80) > SPOT-RNA (0.73) > BPfold (0.67) > REDfold (0.43), and the MCCs follow the order of AlphaFold3 (0.76) ≫ IPknot (0.54) > MXfold2 (0.52) > Mfold (0.46) > ViennaRNA (0.44) > RNAstructure (0.42) > SPOT-RNA (0.37) ∼ BPfold (0.37) > REDfold (0.14); see Fig. 4(b) and Table S6 in the supplementary material. It is noted that Knotty cannot predict RNAs with length > 500 nts and thus was not involved in the evaluation. For the category, it is interesting that the MFE-based methods outperformed most of the DL-based methods except for AlphaFold3 and MXfold2, which is different from the relative performance between the MFE and DL-based methods for other length ranges. Our structural analyses revealed the RNAs in this category predominantly adopted simple or nested topologies with minimal pseudoknots or ≤3-way junctions and the features of such structures may be particularly amenable to dynamic programming algorithms involved in the traditional models. Another possible reason is that most of the DL-based methods did not involve large RNAs during their training process.

For the category with 10 large RNAs in the range of > 1000 nts, the F1-scores from the different methods follow the order of AlphaFold3 (0.97) ≫ MXfold2 (0.68) > IPknot (0.62) > Mfold (0.54) > ViennaRNA (0.51) > RNAstructure (0.49) > BPfold (0.21), and the MCCs follow the order of AlphaFold3 (0.94) ≫ MXfold2 (0.34) > IPknot (0.26) > BPfold (0.21) > Mfold (0.11) > ViennaRNA (0.04) > RNAstructure (0.02); see Fig. 4(c) and Table S7 in the supplementary material. It is noted that REDfold and SPOT-RNA cannot make predictions for RNAs with length > 1000 nts and those with length > 1500 nts, respectively, and consequently, we did not include them in the evaluation for the category. AlphaFold3 showed remarkably high and superior performance,

while other methods showed severe performance reduction. Additionally, MFE-based methods except IPknot exhibited dramatically reduced performance, possibly suggesting their limitations in handling genomic-scale RNAs.

The above comprehensive length-dependent analyses revealed several critical insights. First, while prediction performance generally exhibited an inverse relationship with sequence length, structure complexity can serve as an important factor. Second, AlphaFold3 demonstrated uniquely consistent accuracy across all length ranges, establishing it as the current benchmark for RNA secondary structure prediction. Third, MFE-based methods remained viable for short (< 100 nts) and simple structures but were consistently outperformed by DL-based methods for longer or more complex RNAs. Fourth, it is interesting that the prediction accuracy of AlphaFold3 increases with RNA length, which may be attributed to the sparser data for larger RNAs and consequently more specific features captured through training. Finally, almost all of the existing methods for RNA secondary structure prediction are still limited to small and middle-size RNAs, which significantly constrain genomic-scale applications. These findings collectively highlight both the progress and persistent challenges in computational RNA secondary structure prediction, particularly for RNAs with long sequences or complex architectures.

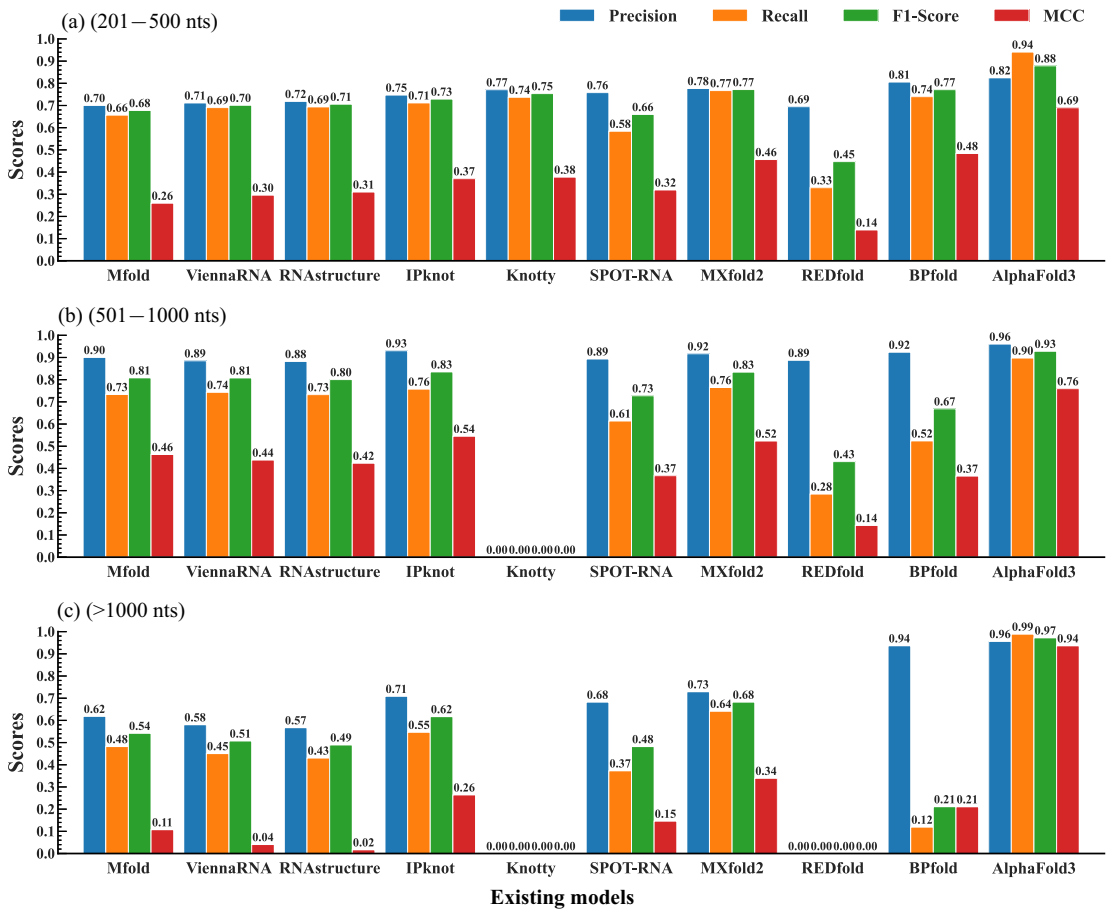


Fig. 4. The performances of different secondary structure prediction methods on long RNAs with sequence length > 200 nts in test set I: (a) for RNA length range of 201–500 nts; (b) for RNA length range of 501–1000 nts; (c) for RNA length longer than 1000 nts. Since Knotty, REDfold, and SPOT-RNA are inapplicable for sequences of > 500 nts, those of > 1000 nts, and those of > 1500 nts, respectively, the shown values are the averages over those applicable RNAs.

4.3. Prediction performance versus RNA structure types

We conducted a further evaluation of the above-introduced methods for RNA secondary structure prediction by examining their performance across several important structure motifs. Specifically, we chose the RNAs from test set I and classified them into four categories: RNAs with 3-way junction, those with 4-way junction, those with 5-way junction, and those with pseudoknot. The mean metrics from the ten methods for the categories with different types have been shown in Fig. 5.

For the category consisting of 42 RNAs with 3-way junction, the F1-scores from the different methods follow the order of AlphaFold3 (0.88) $>$ BPfold (0.83) $>$ MXfold2 (0.80) $>$ Knotty (0.78) $>$ SPOT-RNA (0.76) $>$ Mfold (0.73) $>$ IPknot (0.72) $>$ RNAstructure (0.71) $>$ ViennaRNA (0.69) $>$ REDfold (0.64), and the MCCs follow the order of AlphaFold3 (0.73) $>$ BPfold (0.62) $>$ MXfold2 (0.56) $>$ Knotty (0.51) $>$ SPOT-RNA (0.47) $>$ Mfold (0.41) $\sim$ IPknot (0.41) $>$ RNAs-

tructure (0.38) > REDfold (0.37) > ViennaRNA (0.33); see Fig. 5(a) and Table S8 in the supplementary material. While AlphaFold3 exhibits its leading performance, such leading performance is relatively marginal over other methods, particularly compared to BPfold. The MFE-based methods consistently underperformed in this category except for Knotty which has a slightly lower performance than BPfold and MXfold2.

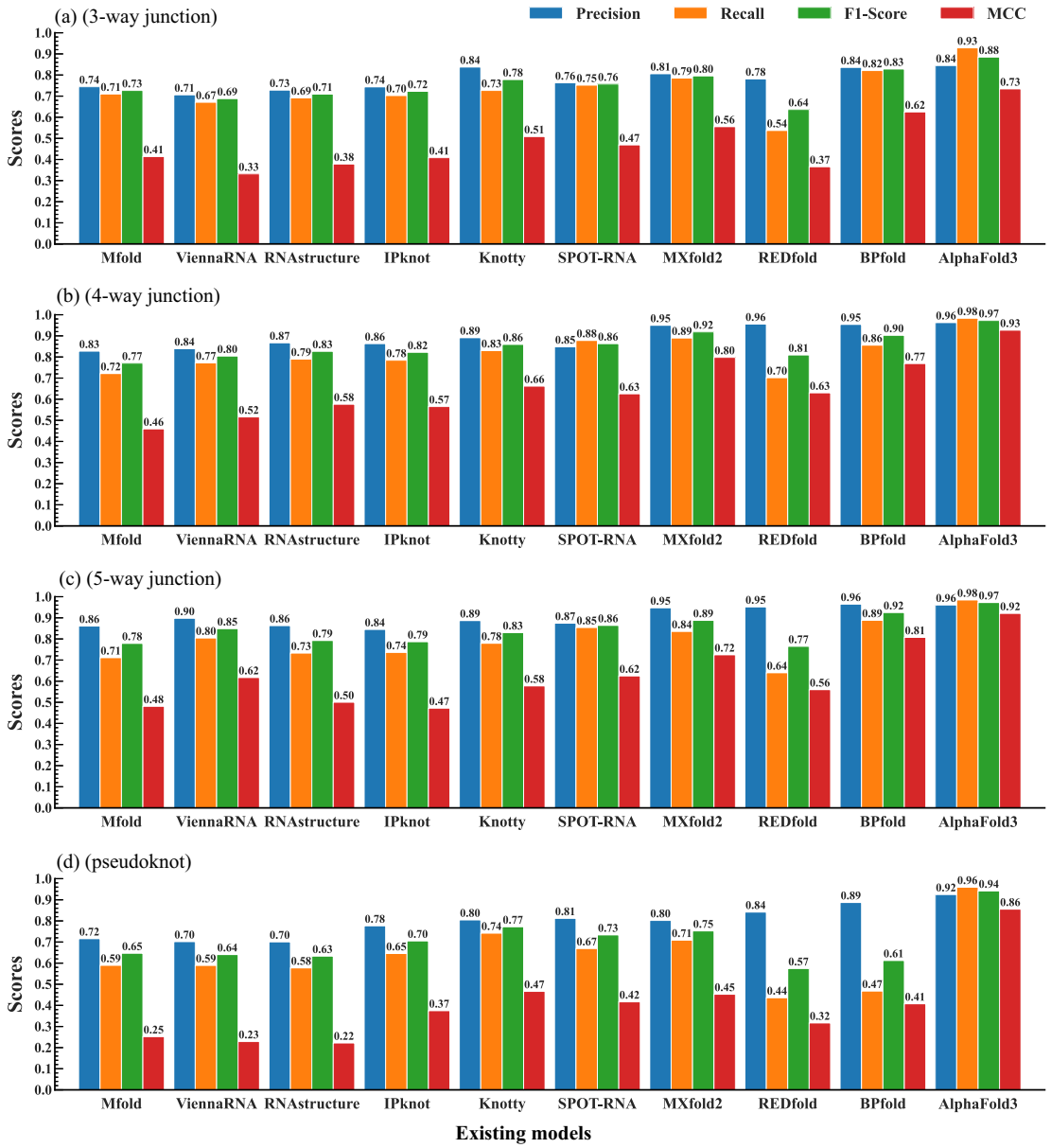


Fig. 5. The performances of different secondary structure prediction methods on different types of RNAs in test set I: (a) RNAs with 3-way junction; (b) RNAs with 4-way junction; (c) RNAs with 5-way junction; (d) RNAs with pseudoknots. Since Knotty, REDfold, and SPOT-RNA are inapplicable for sequences of > 500 nts, those of > 1000 nts, and those of > 1500 nts, respectively, the shown values are the averages over those applicable RNAs.

For the category consisting of 61 RNAs with 4-way junction, the F1-scores from the different methods follow the order of AlphaFold3 (0.97) > MXfold2 (0.92) > BPfold (0.90) > SPOT-RNA (0.86) ∼ Knotty (0.86) > RNAstructure (0.83) > IPknot (0.82) > REDfold (0.81) > ViennaRNA (0.80) > Mfold (0.77), and the MCCs follow the order of AlphaFold3 (0.93) > MXfold2 (0.80) > BPfold (0.77) > Knotty (0.66) > REDfold (0.63) ∼ SPOT-RNA (0.63) > RNAstructure (0.58) > IPknot (0.57) > ViennaRNA (0.52) > Mfold (0.46); see Fig. 5(b) and Table S9 in the supplementary material. AlphaFold3 achieved near-perfect recall and MXfold2 demonstrated remarkable improvement compared with the RNAs with three-way junction. Overall, the MFE-based methods have lower performance than the DL-based ones, and Knotty has leading performance among the MFE-based methods.

For the category consisting of 10 RNAs with 5-way junctions, the F1-scores from the different methods follow the order of AlphaFold3 (0.97) > BPfold (0.92) > MXfold2 (0.89) > SPOT-RNA (0.86) > ViennaRNA (0.85) > Knotty (0.83) > RNAstructure (0.79) ∼ IPknot (0.79) > Mfold (0.78) > RED-

fold (0.77), and the MCCs follow the order of AlphaFold3 (0.92) > BPfold (0.81) > MXfold2 (0.72) > SPOT-RNA (0.62) ∼ ViennaRNA (0.62) > Knotty (0.58) > REDfold (0.56) > RNAstructure (0.50) > Mfold (0.48) > IPknot (0.47); see Fig. 5(c) and Table S10 in the supplementary material. For this category, the methods showed similar performances to the RNAs with 4-way junctions, while ViennaRNA becomes the leading method in the MFE-based ones.

For the category consisting of 197 RNAs with pseudoknots, the F1-scores from the different methods follow the order of AlphaFold3 (0.94) ≫ Knotty (0.77) > MXfold2 (0.75) > SPOT-RNA (0.73) > IPknot (0.70) > Mfold (0.65) > ViennaRNA (0.64) > RNAstructure (0.63) > BPfold (0.61) > REDfold (0.57), and the MCCs follow the order of AlphaFold3 (0.86) ≫ Knotty (0.47) > MXfold2 (0.45) > SPOT-RNA (0.42) > BPfold (0.41) > IPknot (0.37) > REDfold (0.32) > Mfold (0.25) > ViennaRNA (0.23) > RNAstructure (0.22); see Fig. 5(d) and Table S11 in the supplementary material. While AlphaFold3 still kept its leading performance, Knotty showed the superior capability compared to all other methods. Besides, IPknot outperformed other MFE-based methods, reflected by higher MCC and F1-score, indicating that even early pseudoknot-specific methods captured essential features missing in conventional methods. MXfold2 and SPOT-RNA also performed well, suggesting that these DL-based methods are also effective for pseudoknot structures. It is worth noting that Mfold, ViennaRNA and RNAstructure cannot make reliable prediction pseudoknots, and they exhibit visibly lower performance for the category consisting of RNA pseudoknots.

The above comprehensive structure type-dependent analyses revealed the following insights. First, for the RNAs without pseudoknots, the DL-based methods overwhelm the MFE-based methods, while for those RNAs with pseudoknots, the newly developed MFE-based methods such as Knotty can overwhelm the DL-based ones except for AlphaFold3. Second, overall, the involved methods except for AlphaFold3 become less reliable for more complex structures, such as multi-way junctions and pseudoknots. Third, AlphaFold3 exhibits higher prediction accuracy for more complex structures, which may be attributed to the specific structure features from sparse data involved in the training process. Finally, MFE-based methods can become competitive for RNAs with certain complex motifs such as pseudoknots.

5. Conclusion

Accurate prediction of RNA secondary structures is crucial for understanding RNA functions. In this review, we have made a comprehensive evaluation of ten popular RNA secondary structure prediction methods, through assessing their performance across two dimensions of RNA length and structure types. We built a test set containing a large number of nonredundant RNAs and used four typical metrics to provide quantitative insights into the existing computational methods of RNA secondary structure prediction. The main conclusions are listed in the following:

1. AlphaFold3 consistently achieves the highest prediction accuracy (highest F1-scores & MCC) across almost all RNA lengths and structure types, as compared with all other MFE and DL-based methods. The DL-based methods of BPfold and MXfold2 can be strong alternatives to AlphaFold3 for RNAs with medium lengths (100–500 nts) and multi-way junctions.

2. Overall, MFE-based methods have lower performance than the DL-based methods in RNA secondary structure predictions, while they can still be a practical choice for small RNAs (< 100 nts) due to their computational portability and interpretability. But the poor performance of MFE-based methods on large RNAs with complex structures limits their applicability for predicting RNA secondary structures.

3. For RNAs with pseudoknots, besides AlphaFold3, MXfold2 and Knotty can be reliable candidates for predicting secondary structures, despite that the application of Knotty is limited to the sequence length of < 500 nts.

4. For almost all the methods except for AlphaFold3, there is an inverse relationship between RNA length and prediction accuracy. While AlphaFold3 demonstrates even unprecedentedly higher accuracy for larger RNAs, all other examined methods exhibit severe accuracy reduction when RNA length exceeds 500 nts, suggesting a pressing need for the high prediction accuracy across genomic-scale RNAs.

Although we have obtained the above major conclusions, there are several limitations in this work. First, the test set was built based on the RNA 3D structures deposited in the PDB database, which may be already involved in the training process of the DL-based methods especially AlphaFold3. This may contribute to the overall superior performance of the DL-based methods over the MFE-based methods and the higher accuracy of AlphaFold3 for larger RNAs. A more comprehensive evaluation would require a large number of new RNA structures as a "blind" test set. Second, this evaluation did not include SPOT-RNA2, an improved version of SPOT-RNA involving MSA information, since our test set contains 514 RNAs and SPOT-RNA2 is computationally very slow. However, it has been shown that SPOT-RNA2 has a visibly improved performance over SPOT-RNA.[2] Thus, an evaluation for SPOT-RNA2 would require a small test set with redundant RNAs. Moreover, our evaluations were performed on a CPU machine and we did not include the GPU-based methods such as Ufold.[58] Third, our evaluation only involved the single-stranded RNAs since most of the existing methods

cannot be directly applicable to double-stranded or multiple-stranded RNAs, and an evaluation on the methods for predicting secondary structures of double-/multiple-stranded RNAs would be very valuable and is deserved to be performed in the near future. Nevertheless, the comprehensive evaluation in this work would be helpful for users to choose the proper secondary structure prediction methods and can form a solid basis for improving RNA secondary structure predictions in the near future.

6. Challenges and perspectives

Although extensive computational methods have been developed for predicting RNA secondary structures, the problem of modeling RNA secondary structures is still far from the final answer. In the following, we will discuss the major challenges and perspectives in modeling RNA secondary structures in several aspects.

6.1. On the robustness of DL-based methods

Overall, the DL-based methods have generally achieved superior performances in predicting RNA secondary structures, especially AlphaFold3, in comparisons with the MFE-based methods. However, the high accuracy of DL-based methods may face the serious robustness in "blind" predictions.[61] For example, the prediction accuracy of AlphaFold3 increases with the RNA length, e.g., F1-score increases from 0.89 for the RNA length range of 101–200 nts to 0.97 for the RNA length range of $>$ 1000 nts, and simultaneously, the overall F1-score of AlphaFold3 decreases visibly from 0.93 to 0.86 when test set I is replaced by test set II with newly released RNA structures. This suggests that the DL-based methods may exhibit excellent prediction accuracy for the RNAs in the structure spectrum involved in the training process, while such accuracy may fall significantly for the RNAs outside the training structure spectrum. Thus, achieving a reliable and robust accuracy for "blind" predictions remains a challenge for DL-based methods. For the issue, a hybrid strategy with combining MFE-based methods and DL-based ones can be an optional way to achieve high prediction accuracy for the RNAs in the training structure spectrum through DL-based methods while achieve acceptable predictions through MFE-based methods for the RNAs significantly outside the training structure spectrum.[65] Another possible approach is to reasonably involve physical principles or experimental parameters in the training process for a DL-based method, such as the thermodynamics-integration technique[56] or the thermodynamics parameters.[64] Of course, with the increase of the number of experimentally determined RNA structures, the DL-based methods would become more robust and more reliable in prediction accuracy.

6.2. On the accuracy of MFE-based methods.

The MFE-based methods are overall outperformed by the top DL-based methods in prediction accuracy, while the methods can generally make the predictions for RNA folding thermodynamics and structure ensemble obeying the Boltzmann law.[38–41,67] Thus, to improve the structure prediction accuracy of MFE-based methods would greatly benefit the understanding of RNA-related functions and the RNA design. In fact, some MFE-based methods can give reliable predictions for the secondary structures of small RNAs ($\leq$ 50 nts), even compared with the DL-based methods such as AlphaFold3. Such reliable performance of MFE-based methods such as RNAstructure and Knotty may be attributed to the fact that the used thermodynamic parameters were determined for very small RNAs.[86,94] However, such parameters may become inaccurate for large RNAs due to the correlations between different structure domains of large RNAs. Thus, a possible way to improve the prediction accuracy of MFE-based methods is to develop or optimize more accurate and reliable parameters applicable for large RNAs, and such parameters may be RNA size-dependent. Such improved parameters would contribute not only to more accurate secondary structure prediction but also to predicting RNA structure ensemble and folding thermodynamics, which is far beyond the ability of DL-based methods in the near future.

6.3. On predicting RNA secondary structures at varying ion conditions.

Due to the highly polyanionic nature of RNAs,[95,96] RNA structures can be sensitive to the ion environments.[97–99] The existing methods for RNA secondary structure predictions were developed either based on the existing determined RNA structures or based on the experimentally derived thermodynamic parameters at a conventional ion condition (e.g., 1 M NaCl),[94,100–103] and the effects of ions were usually ignored in modeling RNA secondary structures. It is currently impossible for DL-based methods to account for the effects of ions due to the lack of abundant data on the effect of ions, while it is possible for a physics-based model to predict RNA secondary structures at various ion conditions. One possible way is to involve ion-dependent thermodynamic parameters in modeling RNA secondary structures, like Mfold and RNAstructure. For examples, the salt-extension of the thermodynamic parameters for helix and loop formation has been proposed for RNAs and DNAs, including the effect of monovalent and divalent ions.[104–111] Another possible way is to develop a RNA folding model in which an RNA is modeled as a polyelectrolyte chain with base pairing/stacking interactions,[10,23,112–115] and thus the effects of monovalent and divalent ions can be implicitly or explicitly included in the structure predictions.

6.4. On predicting RNA secondary structure in an *in vivo*-like environment

RNAs usually perform their biological functions in a crowded cellular environment which contains not only ions but also many other solute (macro-)molecules such as proteins, and such volume percentage of crowding macromolecules can reach $\sim$ 40%.[116] Thus, to predict RNA secondary structures in a crowded cellular environment is crucial for understanding RNA functions.[117–121] However, it would be a great challenge for the major existing MFE and DL-based methods which completely ignore the realistic spatial coordinates of RNA structures and consequently cannot account for the interactions with crowded macromolecules. The effect of crowded cellular environment can be taken into account in two possible ways: one is that an RNA can be modeled as one in a confined space;[122,123] the other way is to explicitly consider the existence of solute molecules including proteins or their mimic spheres.[116,124,125] Simultaneously, an RNA is modeled as a polyelectrolyte chain with base pairing/stacking interactions in such a crowding environment.[123] Therefore, a physics-based RNA folding model in a crowding environment can be developed and employed to predict RNA secondary structures in an *in vivo*-like environment.

In summary, great progress has been made in the last three decades in modeling RNA secondary structures, while some challenges still remain. Thus, great efforts are still required for modeling RNA secondary structures at high accuracy and to understand the associated RNA functions in vivo. With the increase in RNA structure data and the advances in physics-based modeling principles and computation technology, the exciting progress would be greatly anticipated on modeling RNA secondary structures in the near future.

Acknowledgements

We are grateful to Profs Shi-Jie Chen (University of Missouri) and Jian Zhang (Nanjing University) for valuable discussions. This work was supported by grants from the National Science Foundation of China (Grant Nos. 12375038 and 12075171 to ZJT, and 12205223 to YLT). Parts of the numerical calculations in this work were performed on the super computing system in the Supercomputing Center of Wuhan University.

References

[1] Tinoco I and Bustamante C 1999 *J. Mol. Biol.* **293** 271
[2] Qi F, Chen J, Chen Y, Sun J, Lin Y, Chen Z and Kapranov P 2024 *Genomics Proteomics Bioinformatics* **22** qzae043
[3] Montange R K and Batey R T 2008 *Annu. Rev. Biophys.* **37** 117
[4] Chen S J 2008 *Annu. Rev. Biophys.* **37** 197
[5] Ken M L, Roy R, Geng A, Ganser L R, Manghrani A, Cullen B R, Schulze-Gahmen U, Herschlag D and Al-Hashimi H M 2023 *Nature* **617** 835
[6] Moore P B and Steitz T A 2011 *Cold Spring Harb. Perspect. Biol.* **3** a003780
[7] Wilson T J and Lilley D M J 2021 *WIREs RNA* **12** e1651
[8] Roundtree I A, Evans M E, Pan T and He C 2017 *Cell* **169** 1187
[9] Woodson S A 2008 *Curr. Opin. Chem. Biol.* **12** 667
[10] Shi Y Z, Wang F H, Wu Y Y and Tan Z J 2014 *J. Chem. Phys.* **141** 105102
[11] Sun L Z, Zhang D and Chen S J 2017 *Annu. Rev. Biophys.* **46** 227
[12] Agrawal N, Dasaradhi P V N, Mohmmed A, Malhotra P, Bhatnagar R K and Mukherjee S K 2003 *Microbiol. Mol. Biol. Rev.* **67** 657
[13] Zhang J, Lin M, Chen R, Wang W and Liang J 2008 *J. Chem. Phys.* **128** 125107
[14] Zhang J, Dundas J, Lin M, Chen R, Wang W and Liang J 2009 *RNA* **15** 2248
[15] Zhang Y, Zhang J and Wang W 2011 *J. Am. Chem. Soc.* **133** 6882
[16] Cao S and Chen S J 2011 *J. Phys. Chem. B* **115** 4216
[17] Zhao Y, Huang Y, Gong Z, Wang Y, Man J and Xiao Y 2012 *Sci. Rep.* **2** 734
[18] Boniecki M J, Lach G, Dawson W K, Tomala K, Lukasz P, Soltysinski T, Rother K M and Bujnicki J M 2016 *Nucleic Acids Res.* **44** e63
[19] Wang J, Mao K, Zhao Y, Zeng C, Xiang J, Zhang Y and Xiao Y 2017 *Nucleic Acids Res.* **45** 6299
[20] Zhang D and Chen S J 2018 *J. Chem. Theory Comput.* **14** 2230
[21] Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, Anishchenko I, Baker D and Yang J 2021 *Nat. Protoc.* **16** 5634
[22] Li J, Zhang S, Zhang D and Chen S J 2022 *Bioinformatics* **38** 4042
[23] Wang X, Tan Y L, Yu S, Shi Y Z and Tan Z J 2023 *Biophys. J.* **122** 1503
[24] Tan Y L, Guo C, Xu J J, Shi Y Z and Tan Z J 2025 *Commun. Theor. Phys.* in Press
[25] Zhang J, Bian Y, Lin H and Wang W 2012 *Phys. Rev. E* **85** 021909
[26] Popenda M, Szachniuk M, Antczak M, Purzycka K J, Lukasiak P, Bartol N, Blazewicz J and Adamiak R W 2012 *Nucleic Acids Res.* **40** e112
[27] Biesiada M, Pachulska-Wieczorek K, Adamiak R W and Purzycka K J 2016 *Methods* **103** 120
[28] Wang J, Wang J, Huang Y and Xiao Y 2019 *Int. J. Mol. Sci.* **20** 4116
[29] Xiong P, Wu R, Zhan J and Zhou Y 2021 *Nat. Commun.* **12** 2777
[30] Zhou L, Wang X, Yu S, Tan Y L and Tan Z J 2022 *Biophys. J.* **121** 3381
[31] Zhang Y, Wang J and Xiao Y 2022 *J. Mol. Biol.* **434** 167452
[32] Tan Y L, Feng C J, Wang X, Zhang W and Tan Z J 2021 *Chin. Phys. B* **30** 028705
[33] Bugnon L A, Edera A A, Prochetto S, Gerard M, Raad J, Fenoy E, Rubiolo M, Chorostecki U, Gabaldón T, Ariel F, Di Persia L E, Milone D H and Stegmayer G 2022 *Brief. Bioinform.* **23** bbac205
[34] Wang X, Yu S, Lou E, Tan Y L and Tan Z J 2023 *Molecules* **28** 5532
[35] Binet T, Padiolleau-Lefèvre S, Octave S, Avalle B and Maffucci I 2023 *BMC Bioinformatics* **24** 422
[36] Rose P W, Prlić A, Altunkaya A, Bi C, Bradley A R, Christie C H, Costanzo L D, Duarte J M, Dutta S, Feng Z, Green R K, Goodsell D S, Hudson B, Kalro T, Lowe R, Peisach E, Randle C, Rose A S, Shao C, Tao Y P, Valasatava Y, Voigt M, Westbrook J D, Woo J, Yang H, Young J Y, Zardecki C, Berman H M, Burley S K 2017 *Nucleic Acids Res.* **45** D271
[37] Zhang J, Fei Y, Sun L and Zhang Q C 2022 *Nat. Methods* **19** 1193
[38] Nussinov R and Jacobson A B 1980 *Proc. Natl. Acad. Sci. USA* **77** 6309
[39] Zuker M and Stiegler P 1981 *Nucleic Acids Res.* **9** 133
[40] Zuker M 1989 *Sci. New Ser.* **244** 48
[41] Hofacker I L, Fontana W, Stadler P F, Bonhoeffer L S, Tacker M and Schuster P 1994 *Monatshefte Chem. Chem. Mon.* **125** 167
[42] Knudsen B and Hein J 2003 *Nucleic Acids Res.* **31** 342
[43] Zuker M 2003 *Nucleic Acids Res.* **31** 3406
[44] Do C B, Woods D A and Batzoglou S 2006 *Bioinformatics* **22** e90
[45] Hamada M, Kiryu H, Sato K, Mituyama T and Asai K 2009 *Bioinformatics* **25** 465
[46] Lu Z J, Gloor J W and Mathews D H 2009 *RNA* **15** 1805
[47] Reuter J S and Mathews D H 2010 *BMC Bioinformatics* **11** 129
[48] Sato K, Kato Y, Hamada M, Akutsu T and Asai K 2011 *Bioinformatics* **27** i85
[49] Lorenz R, Bernhart S H, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler P F and Hofacker I L 2011 *Algorithms Mol. Biol.* **6** 26
[50] Deng F, Ledda M, Vaziri S and Aviran S 2016 *RNA* **22** 1109
[51] Jabbari H, Wark I, Montemagno C and Will S 2018 *Bioinformatics* **34** 3849

[52] Wang L, Liu Y, Zhong X, Liu H, Lu C, Li C and Zhang H 2019 *Front. Genet.* **10** 143
[53] Cheng Y, Zhang S, Xu X and Chen S J 2021 *J. Phys. Chem. B* **125** 10108
[54] Andrikos C, Makris E, Kolaitis A, Rassias G, Pavlatos C and Tsanakas P 2022 *Methods Protoc.* **5** 14
[55] Singh J, Hanson J, Paliwal K and Zhou Y 2019 *Nat. Commun.* **10** 5407
[56] Sato K, Akiyama M and Sakakibara Y 2021 *Nat. Commun.* **12** 941
[57] Singh J, Paliwal K, Zhang T, Singh J, Litfin T and Zhou Y 2021 *Bioinformatics* **37** 2589
[58] Fu L, Cao Y, Wu J, Peng Q, Nie Q and Xie X 2022 *Nucleic Acids Res.* **50** e14
[59] Mao K, Wang J and Xiao Y 2022 *Molecules* **27** 1030
[60] Chen C C and Chan Y M 2023 *BMC Bioinformatics* **24** 122
[61] Qiu X 2023 *PLOS Comput. Biol.* **19** e1011047
[62] Abramson J, Adler J, Dunger J, *et al.* 2024 *Nature* **630** 493
[63] Zhou Y, Pedrielli G, Zhang F and Wu T 2024 *BMC Bioinformatics* **25** 316
[64] Zhou S K, Zhu H and Xiong P 2024 preprint
[65] Qiu X 2025 *Biol. Methods Protoc.* **10** bpae097
[66] Yang C, Wu H, Shen T, Zou K and Sun S 2025 *Proc. AAAI Conf. Artif. Intell.* **39** 950
[67] Seetin M G and Mathews D H 2012 *Bacterial Regulatory RNA* (Totowa, NJ: Humana Press) p. 99
[68] Zhang C, Zhang Y and Pyle A M 2023 *J. Mol. Biol.* **435** 167904
[69] Zhang Y, Lang M, Jiang J, Gao Z, Xu F, Litfin T, Chen K, Singh J, Huang X, Song G, Tian Y, Zhan J, Chen J and Zhou Y 2024 *Nucleic Acids Res.* **52** e3
[70] Calonaci N, Jones A, Cuturello F, Sattler M and Bussi G 2020 *NAR Genomics Bioinforma.* **2** lqaa090
[71] Zhao Q, Zhao Z, Fan X, Yuan Z, Mao Q and Yao Y 2021 *PLoS Comput. Biol.* **17** e1009291
[72] Yu H, Qi Y and Ding Y 2022 *Front. Mol. Biosci.* **9** 869601
[73] Justyna M, Antczak M and Szachniuk M 2023 *Brief. Bioinform.* **24** bbad153
[74] Wu K E, Zou J Y and Chang H 2023 *Brief. Bioinform.* **24** bbad210
[75] Zhang S, Li J and Chen S J 2024 *Biophys. J.* **123** 2647
[76] Mathews D H, Andre T C, Kim J, Turner D H and Zuker M 1997 *Molecular Modeling of Nucleic Acids* (Washington, DC: American Chemical Society) pp. 246–257
[77] Lu X J, Bussemaker H J and Olson W K 2015 *Nucleic Acids Res.* **43** e142
[78] Tan Y L, Feng C J, Jin L, Shi Y Z, Zhang W and Tan Z J 2019 *RNA* **25** 793
[79] Tan Y L, Wang X, Shi Y Z, Zhang W and Tan Z J 2022 *Biophys. J.* **121** 142
[80] Tan Y L, Wang X, Yu S, Zhang B and Tan Z J 2023 *NAR Genomics Bioinforma.* **5** lqad016
[81] Fu L, Niu B, Zhu Z, Wu S and Li W 2012 *Bioinformatics* **28** 3150
[82] Matthews B W 1975 *Biochim. Biophys. Acta BBA - Protein Struct.* **405** 442
[83] Chen H L, Condon A and Jabbari H 2009 *J. Comput. Biol.* **16** 803
[84] Wexler Y, Zilberstein C and Ziv-Ukelson M 2007 *J. Comput. Biol.* **14** 856
[85] Backofen R, Tsur D, Zakov S and Ziv-Ukelson M 2011 *J. Discrete Algorithms* **9** 12
[86] Andronescu M S, Pop C and Condon A E 2010 *RNA* **16** 26
[87] Danaee P, Rouches M, Wiley M, Deng D, Huang L and Hendrix D 2018 *Nucleic Acids Res.* **46** 5381
[88] Kamisetty H, Ovchinnikov S and Baker D 2013 *Proc. Natl. Acad. Sci. USA* **110** 15674
[89] McCaskill J S 1990 *Biopolymers* **29** 1105
[90] Rauzan B, McMichael E, Cave R, Sevcik L R, Ostrosky K, Whitman E, Stegemann R, Sinclair A L, Serra M J and Deckert A A 2013 *Biochemistry* **52** 765
[91] Weber G 2015 *Bioinformatics* **31** 871
[92] Wang Y, Wang Z, Wang Y, Liu T and Zhang W 2018 *J. Chem. Phys.* **148** 045101
[93] Matthies M C, Krueger R, Torda A E and Ward M 2023 *Nucleic Acids Res.* **52** e14
[94] Mathews D H, Sabina J, Zuker M and Turner D H 1999 *J. Mol. Biol.* **288** 911
[95] Tan Z J and Chen S J 2010 *Biophys. J.* **99** 1565
[96] Tan Z J and Chen S J 2011 *Biophys. J.* **101** 176
[97] Xi K, Wang F H, Xiong G, Zhang Z L and Tan Z J 2018 *Biophys. J.* **114** 1776
[98] Dong H L, Zhang C, Dai L, Zhang Y, Zhang X H and Tan Z J 2024 *Nucleic Acids Res.* **52** 2519
[99] Qiang X W, Zhang C, Dong H L, Tian F J, Fu H, Yang Y J, Dai L, Zhang X H and Tan Z J 2022 *Phys. Rev. Lett.* **128** 108103
[100] Sugimoto N, Nakano S, Yoneyama M and Honda K 1996 *Nucleic Acids Res.* **24** 4501
[101] SantaLucia, J, Allawi H T and Seneviratne P A 1996 *Biochemistry* **35** 3555
[102] Xia T, SantaLucia J, Burkard M E, Kierzek R, Schroeder S J, Jiao X, Cox C and Turner D H 1998 *Biochemistry* **37** 14719
[103] Zuber J, Schroeder S J, Sun H, Turner D H and Mathews D H 2022 *Nucleic Acids Res.* **50** 5251
[104] SantaLucia J 1998 *Proc. Natl. Acad. Sci. USA* **95** 1460
[105] SantaLucia J and Hicks D 2004 *Annu. Rev. Biophys. Biomol. Struct.* **33** 415
[106] Owczarzy R, You Y, Moreira B G, Manthey J A, Huang L, Behlke M A and Walder J A 2004 *Biochemistry* **43** 3537
[107] Tan Z J and Chen S J 2006 *Biophys. J.* **90** 1175
[108] Tan Z J and Chen S J 2007 *Biophys. J.* **92** 3615
[109] Tan Z J and Chen S J 2008 *Biophys. J.* **95** 738
[110] Owczarzy R, Moreira B G, You Y, Behlke M A and Walder J A 2008 *Biochemistry* **47** 5336
[111] Chen Z and Znosko B M 2013 *Biochemistry* **52** 7477
[112] Jin L, Shi Y Z, Feng C J, Tan Y L and Tan Z J 2018 *Biophys. J.* **115** 1403
[113] Shi Y Z, Jin L, Feng C J, Tan Y L and Tan Z J 2018 *PLoS Comput. Biol.* **14** e1006222
[114] Nguyen H T, Hori N and Thirumalai D 2019 *Proc. Natl. Acad. Sci. USA* **116** 21022
[115] Jin L, Tan Y L, Wu Y, Wang X, Shi Y Z and Tan Z J 2019 *RNA* **25** 1532
[116] Zhou H X, Rivas G and Minton A P 2008 *Annu. Rev. Biophys.* **37** 375
[117] Qin S and Zhou H X 2009 *Biophys. J.* **97** 12
[118] Knowles D B, LaCroix A S, Deines N F, Shkel I and Record M T 2011 *Proc. Natl. Acad. Sci. USA* **108** 12699
[119] Denesyuk N A and Thirumalai D 2011 *J. Am. Chem. Soc.* **133** 11858
[120] Adams M S and Znosko B M 2019 *Nucleic Acids Res.* **47** 3658
[121] Ghosh S, Takahashi S, Banerjee D, Ohyama T, Endoh T, Tateishi-Karimata H and Sugimoto N 2023 *Nucleic Acids Res.* **51** 4101
[122] Tan Z J and Chen S J 2012 *Biophys. J.* **103** 827
[123] Feng C, Tan Y L, Cheng Y X, Shi Y Z and Tan Z J 2021 *Front. Mol. Biosci.* **8** 666369
[124] Banks A, Qin S, Weiss K L, Stanley C B and Zhou H X 2018 *Biophys. J.* **114** 1067
[125] Zheng C C, Chen Y L, Dong H L, Zhang X H and Tan Z J 2024 *J. Chem. Phys.* **161** 075101