RNA-Puzzles - the evaluation and automation of RNA 3D structure prediction

Zhichao Miao[1,2]

[1]*European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK*
[2]*Translational Research Institute of Brain and Brain-Like Intelligence and Department of Anesthesiology, Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine, Shanghai 200081, China;*

zmiao@ebi.ac.uk

RNA 3D structure prediction is a key bioinformatics topic, which is related to RNA design and RNA therapy. RNA-Puzzles is a collective endeavour dedicated to the advancement and improvement of RNA 3D structure prediction. With agreement from crystallographers, the RNA structures are predicted by various groups before the publication of the crystal structures. Systematic protocols for comparing models and crystal structures are described and analyzed. In RNA-Puzzles, we discuss a) the capabilities and limitations of current methods of 3D RNA structure based on sequences; b) the progress in RNA structure prediction; c) the possible bottlenecks that hold back the field; d) the comparison between the automated web server and human experts; e) the prediction rules, such as coaxial stacking; f) the prediction of structural details and ligand binding; g) the development of novel prediction methods; and h) the potential improvements to be made.

Till now, 31 RNAs with crystal structures have been predicted, while many of them have achieved high accuracy in comparison with the crystal structures. We have summarized part of our results in three papers and two community-wide meetings. With the results in RNA-Puzzles, we illustrate that the current bottlenecks in the field may lie in the prediction of non-Watson-Crick interactions and the reconstruction of the global topology. Correct coaxial stacking and tertiary contacts are key for the prediction of RNA architecture, while ligand binding modes can only be predicted with low resolution.

We now further extend the prediction to RNA sequences in the Rfam families. We have predicted structures for 20 RNA families, while some of the predictions have been confirmed by crystal or cryo-EM structures, indicating the possibility to use predicted models for functional inference. The predicted models also helped in 'Molecular Replacement' for crystal structures.

For the model evaluation, we present RNA-Puzzles toolkit, an RNA computational resource including (i) decoy sets generated by different RNA 3D structure prediction methods (raw, for-evaluation and standardized datasets), (ii) 3D structure normalization, analysis, manipulation, visualization tools (RNA_format, RNA_normalizer, rna-tools) and (iii) 3D structure comparison metric tools (RNAQUA, MCQ4Structures).

With the increasing number of RNA structures being solved as well as the high-throughput biochemical experiments, RNA 3D structure prediction is becoming routine and accurate. Experimental data-aided structure modelling may effectively help in understanding the noncoding RNA function, especially the viral RNAs.

Keywords: RNA, structure prediction, RNA-Puzzles, Assessment, viral RNA