Published in final edited form as:
Bioinformatics. 2008 September 1; 24(17): 1951–1952. doi:10.1093/bioinformatics/btn328.

iFoldRNA: Three-dimensional RNA Structure Prediction and Folding

Shantanu Sharma[1], **Feng Ding**[1], and **Nikolay V. Dokholyan**[1]

1 *Department of Biochemistry & Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599 USA*

Summary

Three-dimensional RNA structure prediction and folding is of significant interest in the biological research community. Here, we present iFoldRNA, a novel web-based methodology for RNA structure prediction with near atomic resolution accuracy and analysis of RNA folding thermodynamics. iFoldRNA rapidly explores RNA conformations using discrete molecular dynamics simulations of input RNA sequences. Starting from simplified linear-chain conformations, RNA molecules (<50 nucleotides) fold to native-like structures within half an hour of simulation, facilitating rapid RNA structure prediction. All-atom reconstruction of energetically stable conformations generates iFoldRNA predicted RNA structures. The predicted RNA structures are within 2–5 Angstrom root mean square deviations from corresponding experimentally derived structures. RNA folding parameters including specific heat, contact maps, simulation trajectories, gyration radii, root mean square deviations from native state, fraction of native-like contacts are accessible from iFoldRNA. We expect iFoldRNA will serve as a useful resource for RNA structure prediction and folding thermodynamic analyses.

1 INTRODUCTION

The central dogma of molecular biology presented RNA as the fundamental ingredient in genetic translational machinery. However, recent discoveries of RNAi, ribozymes and aptamers have extended the scope of RNA function beyond the central dogma. We now understand that RNA molecules serve diverse structural, catalytic and regulatory function in eukaryotic cells. The tertiary structure of RNA molecules plays a crucial role in determining RNA function. However, accurate prediction of three-dimensional (3D) structure and folding kinetics of RNA presents a significant challenge in molecular biotechnology. The necessity of large quantities of pure RNA samples and technical limitations hinder the applications of X-ray crystallography, and nuclear magnetic resonance for high-throughput structure elucidation. These challenges have lead to a newfound interest in computational prediction of RNA tertiary structure (Shapiro et al., 2007) and investigating thermodynamics and mechanism of RNA folding.

A majority of computational methods for probing RNA structure and folding dynamics are limited to secondary structure elucidation, while stochastic models have been used to study RNA folding kinetics. Although these computational approaches have demonstrated a significant utility in predicting the RNA secondary structure, they are largely inadequate for predicting 3D RNA structures. Recently, fragment assembly Monte Carlo (Das and Baker,

Contact: dokh@med.unc.edu.
Availability: http://iFoldRNA.dokhlab.org.
Conflict of interest: none declared.

2007), nucleotide cyclic motifs (Parisien and Major, 2008) and discrete molecular dynamics (DMD) simulations (Ding et al., 2008) have been proposed for 3D RNA structure prediction. Here, we present iFoldRNA (http://iFoldRNA.dokhlab.org), a web-resource for rapid and accurate predictions of 3D RNA structures and probing folding thermodynamics. iFoldRNA performs folding simulations using the DMD engine (Ding et al., 2008; Dokholyan et al., 1998) and Medusa force field (Ding and Dokholyan, 2006) to simulate RNA folding dynamics.

2 METHODS

A simplified three-bead per nucleotide model of RNA and replica-exchange DMD simulation protocol with eight replicas is used to sample the RNA conformational space (Ding et al., 2008). An estimate of the free energies of RNA loop regions is explicitly included in the force-field to model the entropic contributions from RNA loop formation. The simulation is followed by a reconstruction protocol to generate atomic resolution structures. 3D structures corresponding to the lowest free energy states in DMD scale are ascribed as the near-native conformations.

A 520-processor Topsail Linux cluster from the University of North Carolina is used for performing replica-exchange DMD simulations of RNA folding (Ding et al., 2008). The backend of iFoldRNA distributes simulation tasks from the iFoldRNA website to compute nodes of the Topsail cluster using a queue scheduler and a Java-based network communication (Sharma et al., 2006). Once a DMD simulation completes, the compute node generates the putative native-like structures having least relative free-energy in DMD scale and user-specified simulation thermodynamic outputs. These outputs are dispatched back to the scheduler and subsequently the user is notified of simulation results via email.

3 RESULTS

Multiple native-like RNA topologies and the corresponding relative free energy values are accessible from the iFoldRNA server. Our recent work has demonstrated the efficacy of the DMD conformational sampling engine in rapid simulations of RNA folding dynamics (Ding et al., 2008). The iFoldRNA resource enables world-wide access to rapid tertiary structure prediction and folding thermodynamics of RNA molecules using the DMD engine. Folding parameters including inter-nucleotide contact maps, simulation trajectories, gyration radii, root mean square deviations from native state, and fraction of native-like contacts (Q-value) are accessible from the iFoldRNA server. Secondary structures generated by iFoldRNA are consistent with Mfold and ViennaRNA predictions.

Low root mean square deviations (2–3 Å) are observed in 3D super-positions of iFoldRNA predictions against experimental structures, demonstrating the accuracy of iFoldRNA in structure prediction (Fig. 1a, b). Typical iFoldRNA folding simulations and analyses are performed within an hour (Fig. 1b) as compared to months to years spent on conventional molecular dynamics simulations to explore conformational space. Fast conformational sampling ability of DMD enables rapid structure prediction of putative RNA sequences using iFoldRNA. We have also developed a post-simulation analysis tool, iFoldRNA-Analysis available at the iFoldRNA website for user-specified analyses of RNA folding using the weighted histogram analysis method (http://www.mmtsb.org). Sample simulation outputs obtained from iFoldRNA and iFoldRNA-Analysis are elucidated in Fig. 1c, d. Folding transition temperatures obtained from specific heat graph (Fig. 1c) and fractions of native base-pairs (Fig. 1d) can be directly compared across different RNA sequence.

Large RNA molecules having >50 nucleotides (e.g. ribosomal RNA, NDB: 2il9, 142 nucleotides) require significantly longer time scales to sample the exponentially increasing

conformational space. This limits the accuracy of the iFoldRNA structure prediction to intermediate-length RNA molecules (<50 nucleotides). In future, experimental constraints, e.g. using SHAPE (Wilkinson et al., 2005) may be integrated with iFoldRNA to overcome such size limitations. We anticipate that the iFoldRNA server will gather significant attention in the research community interested in predicting 3D structures and probing folding mechanisms of RNA molecules. The iFoldRNA server is freely accessible at http://iFoldRNA.dokhlab.org for academic and non-profit users.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by the National Institutes of Health grant R01CA084480-07. National Institutes of Health Fellowship 1T90DA022857-01 supported S. S. We thank the University of North Carolina IT Services for providing hardware support.

References

Das R, Baker D. Automated de novo prediction of native-like RNA tertiary structures. Proc Natl Acad Sci U S A. 2007

Ding F, Dokholyan NV. Emergence of protein fold families through rational design. PLoS Comput Biol 2006;2:e85. [PubMed: 16839198]

Ding F, et al. Ab initio RNA folding by discrete molecular dynamics: From structure prediction to folding mechanisms. RNA. 2008

Dokholyan NV, et al. Discrete molecular dynamics studies of the folding of a protein-like model. Fold Des 1998;3:577–587. [PubMed: 9889167]

Parisien M, Major F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. Nature 2008;452:51–55. [PubMed: 18322526]

Shapiro BA, et al. Bridging the gap in RNA structure prediction. Curr Opin Struct Biol 2007;17:157–165. [PubMed: 17383172]

Sharma S, et al. iFold: a platform for interactive folding simulations of proteins. Bioinformatics 2006;22:2693–2694. [PubMed: 16940324]

Wilkinson KA, Merino EJ, Weeks KM. RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNA(Asp) transcripts. J Am Chem Soc 2005;127:4659–4667. [PubMed: 15796531]

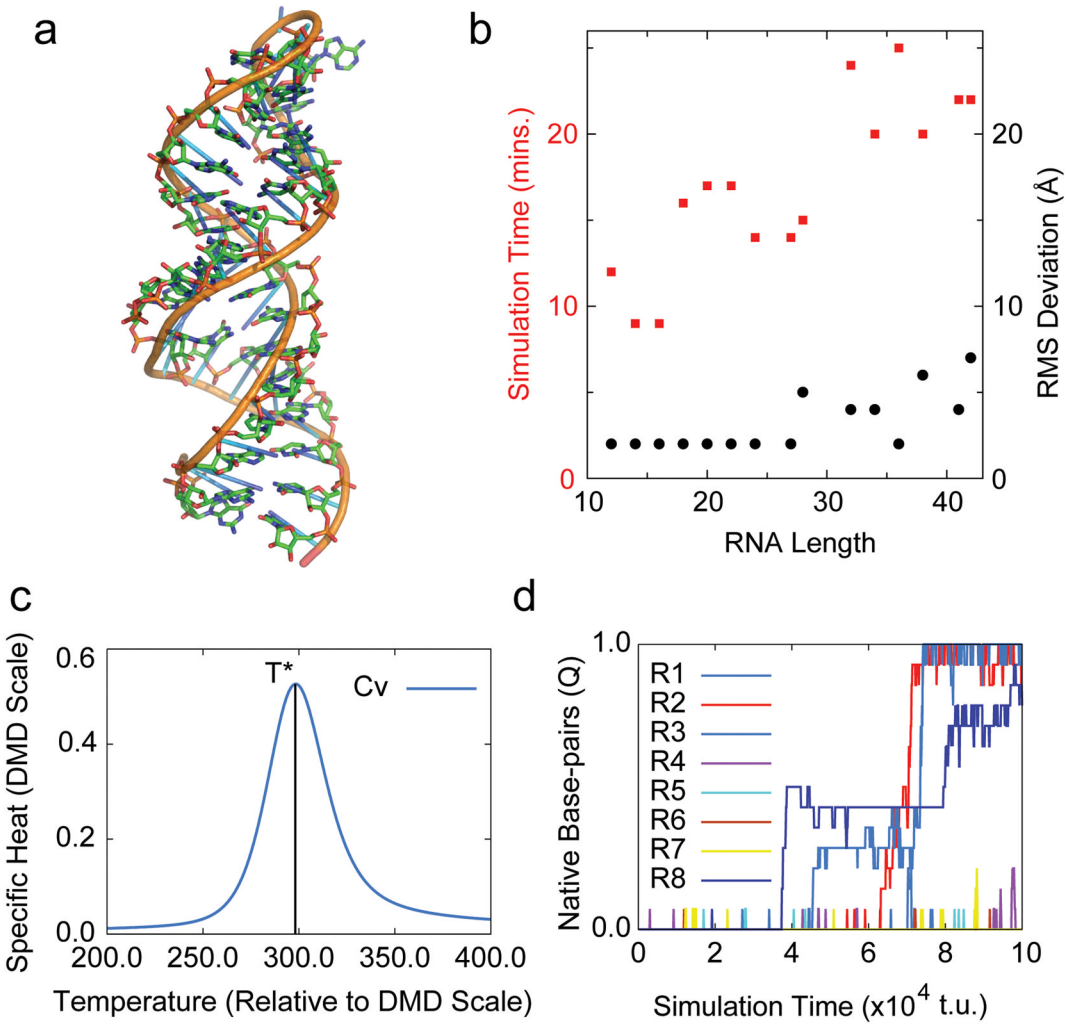

Fig. 1. iFoldRNA tertiary structure prediction and folding thermodynamics

(a) Superposition of iFoldRNA generated stem-loop (sticks) vs. corresponding NMR structure (cartoon; Nucleic Acid DataBank: 1n8x; all-atom RMSD: 2.65 Å). (b) Scaling of iFoldRNA simulation times (red squares) and root mean square deviations (black circles) with increasing RNA length. Turnover time of RNA simulations is plotted as a function of total number of nucleotides in RNA. For intermediate-length RNA sequences (<50 nucleotides), simulation turnover times are within 30 minutes yielding 2–5 Å RMSD between predicted and experimental structures. (c) Graph of specific heat of the stem-loop (Nucleic Acid DataBank: 1n8x) vs. temperature generated using iFoldRNA-Analysis. Weighted histogram analysis method is used to compute the two-dimensional potential of mean force from replica-exchange DMD simulations. T* denotes the conformational transition temperature in relative DMD units. Figure generated using iFoldRNA-Analysis (d) Fractions of native-like base-pairs (Q-values) for eight replicas (R1-R8) in the model 1n8x folding simulation. Replicas (R1, R2, R8) explore native-like conformations with Q-values ~1.0 within 1×10^5 time units of replica exchange DMD simulation. Graph generated by iFoldRNA using GNUPlot software.