# Enhanced Sampling and Overfitting Analyses in Structural Refinement of Nucleic Acids into Electron Microscopy Maps

Harish Vashisth[a], Georgios Skiniotis[b], and Charles L. Brooks III[a,*]

[a]*Department of Chemistry and Biophysics Program, University of Michigan, Ann Arbor, MI, USA*
[b]*Life Sciences Institute, Department of Biological Chemistry, and Biophysics Program, University of Michigan, Ann Arbor, MI, USA*

* Corresponding author; Email: brookscl@umich.edu

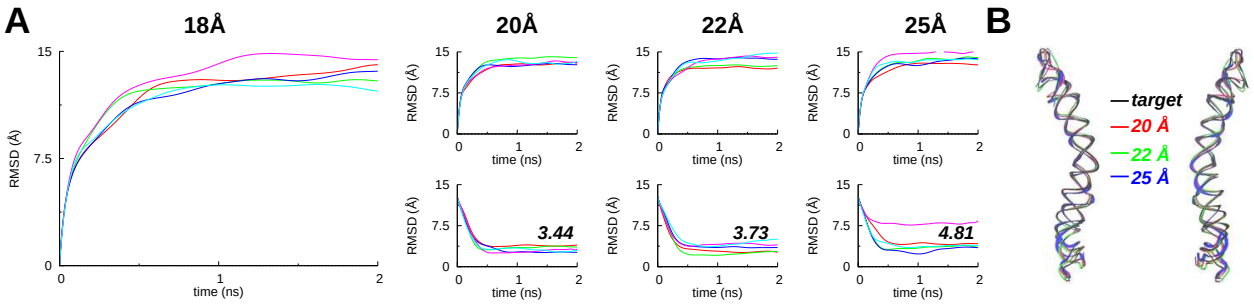


**Figure S1. MDFF fitting of H44-RNA with high structural restraints at different map resolutions.** (**A**) Top panels for each map resolution show the RMSD traces (backbone P-atoms) from the known initial structure and bottom panels show the RMSD traces from the final structure. Data from five runs for each map resolution are shown. Due to the absence of the known conformation of H44 in the immature map at 18 Å resolution, the RMSDs from the target conformation in EM maps filtered to 20 Å to 25 Å resolution were computed by considering as a basis the structure with the highest correlation-coefficient (among five runs) obtained from MDFF refinement in the 18 Å map. Therefore, RMSD traces are shown only from the initial conformation for 18 Å map. (**B**) Cartoon representations of two different views of the overlay of final conformations generated via MDFF fitting of each RNA at each map resolution is shown. Out of five independent runs for each map resolution, only the structure with the highest correlation-coefficient is rendered. Black cartoons correspond to the target conformation, while cartoons of other colors correspond to the fitting with the highest correlation-coefficient for each map resolution.

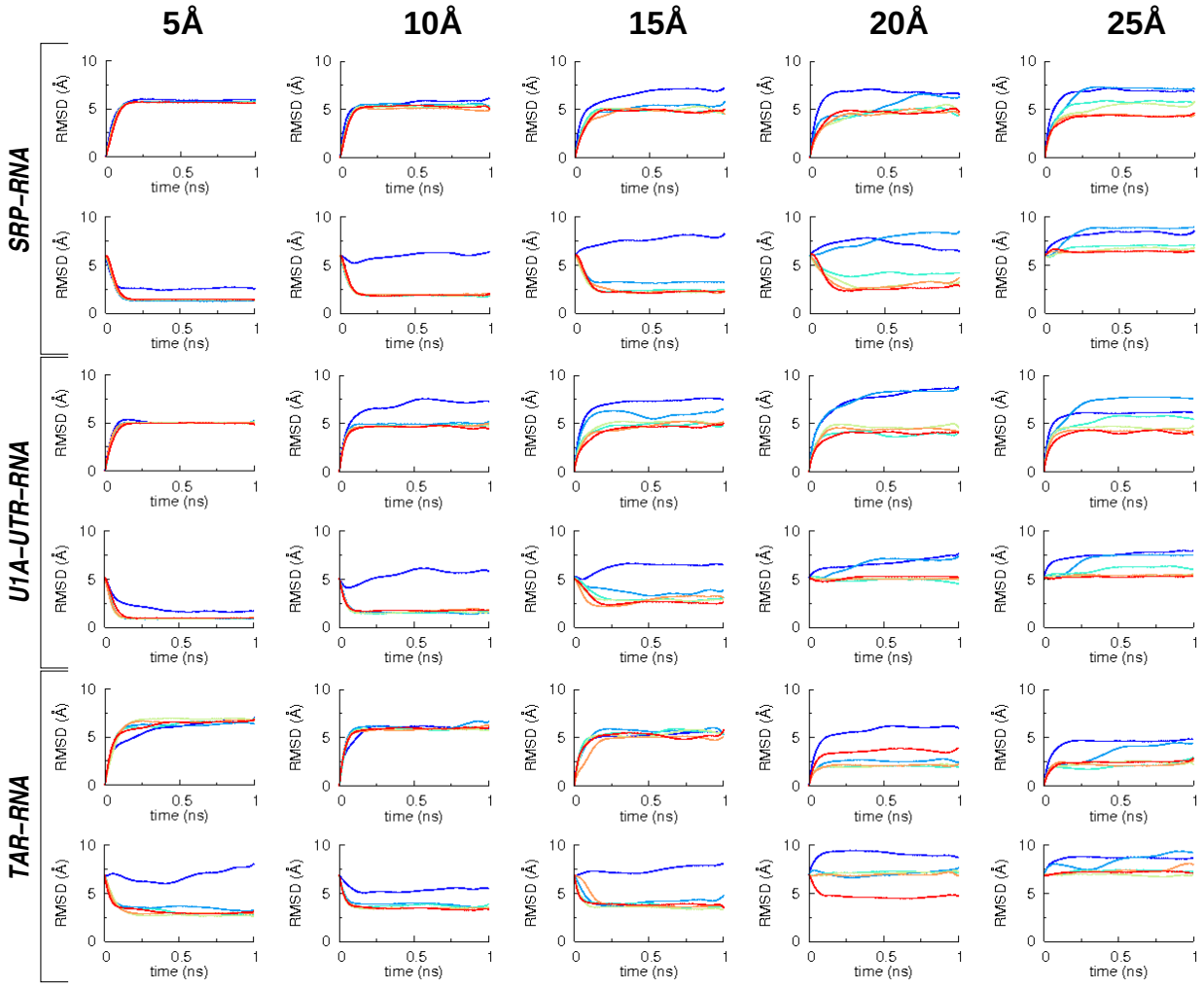


**Figure S2. MDFF fitting for different nucleic acids with varying degree of structural restraints.** Top panels for each nucleic acids show the RMSD traces (backbone P-atoms) from the initial structures and bottom panels show the RMSD traces from the final structures. Blue to red indicates the lowest to highest restraining spring constants ($k$; units of kcal mol$^{-1}$ rad$^{-2}$ or kcal mol$^{-1}$ Å$^{-2}$) used. Six different colors of RMSD traces correspond to $k$ = 0 (dark blue), 10, 50, 100, 300, and 500 (dark red). Note that better final models are obtained with higher structural restraints at any map resolution (red traces in lower RMSD panels).

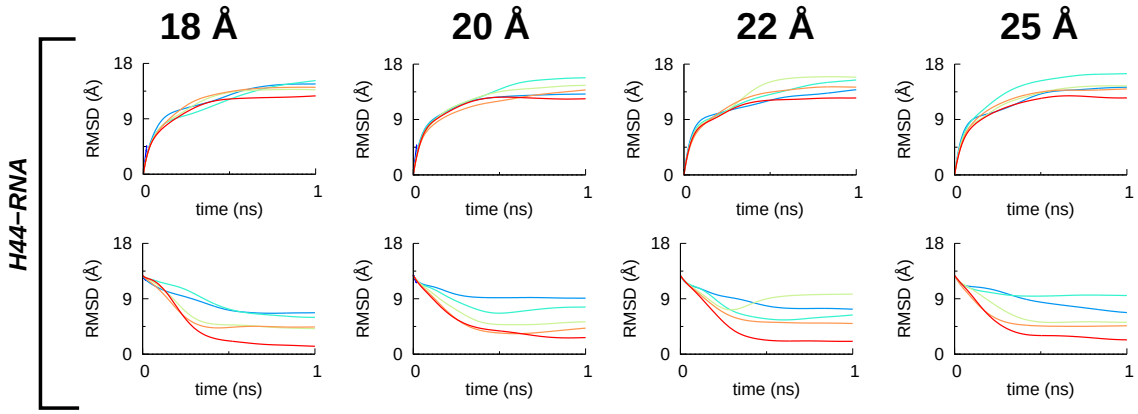


**Figure S3. MDFF fitting of H44-RNA in experimental maps with varying degree of structural restraints.** See the caption of Figure S2 for details. Absence of RMSD traces at k = 0 corresponding to unstable simulations.

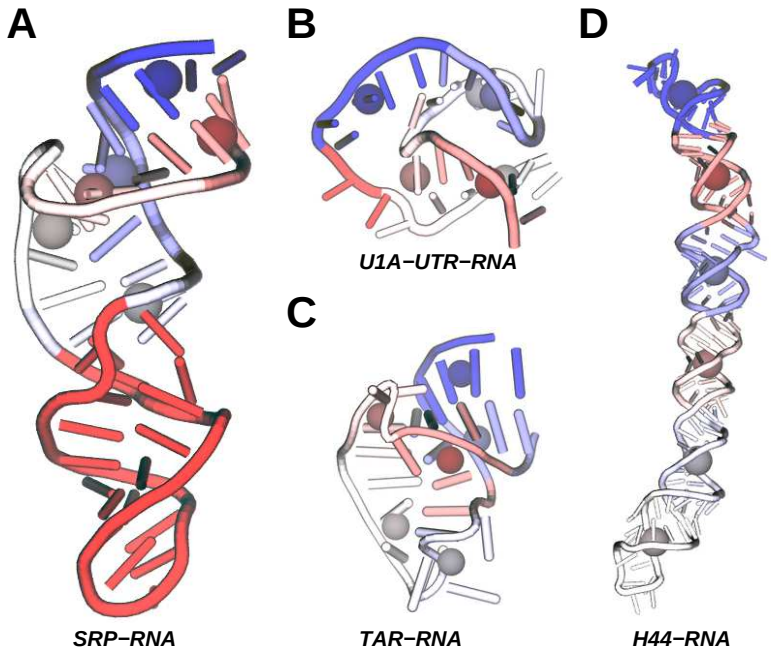


**Figure S4. Subdomain partitions for the TAMD-part of TAMDFF simulations of all nucleic acids.** (**A**) SRP-RNA, (**B**) U1A-UTR-RNA, (**C**) TAR-RNA, and (**D**) H44-RNA. Each sphere represents the center-of-mass (COM) of a mutually exclusive subdomain. Residue memberships for each subdomain are listed in Tables S1-S4. In some cases such as SRP-RNA, TAMD was not applied to the entire structure as indicated by the absence of subdomains in certain regions.

**Table S1. Subdomain Memberships in SRP-RNA**

| Subdomain | Mass (kDa) | Residues |
|---|---|---|
| 1 | 1.34 | 1-4 |
| 2 | 1.24 | 5-8 |
| 3 | 1.28 | 9-12 |
| 4 | 1.29 | 13, 33-35 |
| 5 | 1.34 | 36-39 |
| 6 | 1.25 | 40-43 |

**Table S2. Subdomain Memberships in U1A-UTR-RNA**

| Subdomain | Mass (kDa) | Residues |
|---|---|---|
| 1 | 1.67 | 19-23 |
| 2 | 1.28 | 24-25, 45-46 |
| 3 | 1.25 | 47-50 |
| 4 | 1.52 | 26-30 |
| 5 | 1.66 | 33-37 |
| 6 | 1.58 | 38-42 |

**Table S3. Subdomain Memberships in TAR-RNA**

| Subdomain | Mass (kDa) | Residues |
|---|---|---|
| 1 | 1.62 | 21-23, 36-37 |
| 2 | 1.28 | 24-27 |
| 3 | 1.26 | 28-31 |
| 4 | 1.36 | 32-35 |
| 5 | 1.24 | 19-20, 38-39 |
| 6 | 1.26 | 40-43 |

**Table S4. Subdomain Memberships in H44-RNA**

| Subdomain | Mass (kDa) | Residues |
|---|---|---|
| 1 | 7.45 | 1644-1655, 1744-1754 |
| 2 | 7.11 | 1667-1677, 1722-1732 |
| 3 | 7.14 | 1688-1709 |
| 4 | 7.16 | 1678-1687, 1710-1721 |
| 5 | 6.97 | 1656-1666, 1733-1743 |
| 6 | 7.03 | 1634-1643, 1755-1766 |