The Sampling Performance of Multiple Independent Molecular Dynamics Simulations of an RNA Aptamer

Shuting Yan[1], Jason M. Peck[1], Muslum Ilgu[2,3,4], Marit Nilsen-Hamilton[2,3] and Monica H. Lamm[1]*

[1]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA, USA 50011-2140

[2]Roy J Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, USA 50011-2140

[3]Aptalogic Inc., Ames, IA, USA 50011

[4]Department of Biological Sciences, Middle East Technical University, Ankara, Ankara, TR 06800

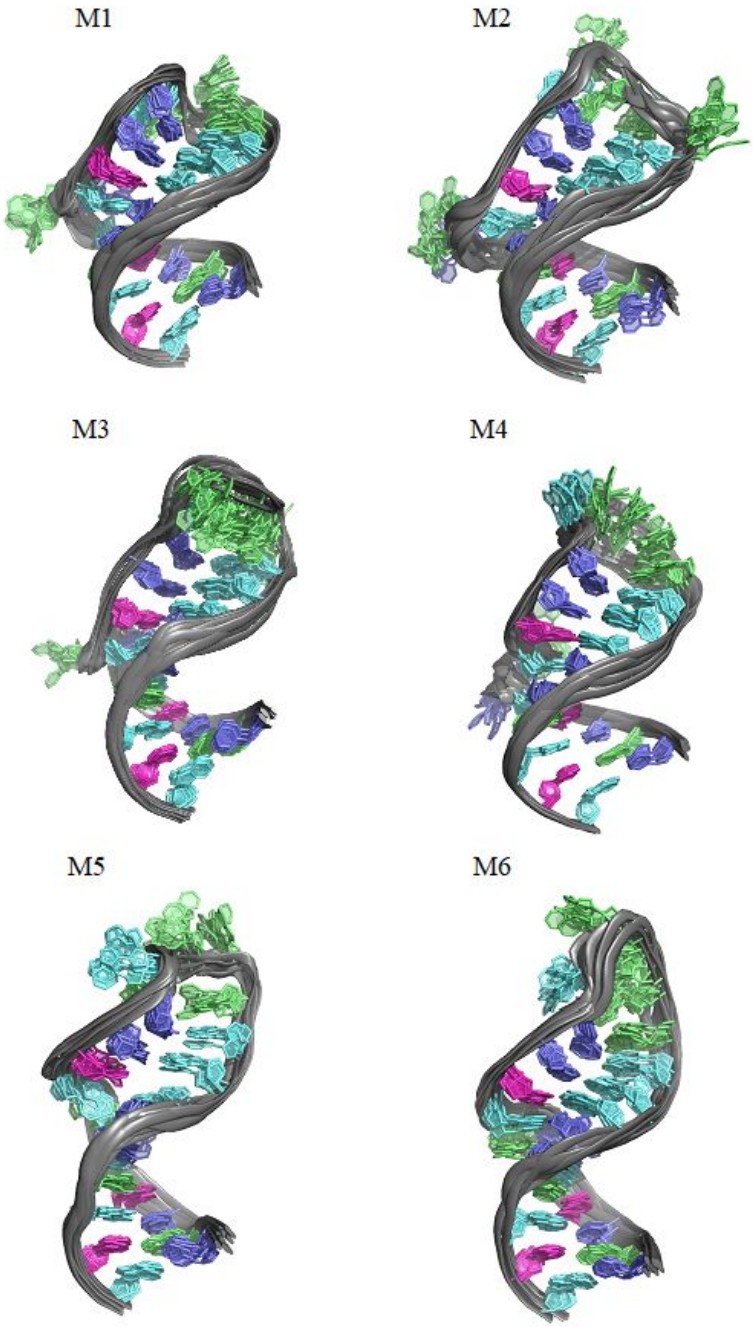

Figure S1. Screenshot of 60 initial structures of MD simulations in this study. All the structures within a group are superimposed.

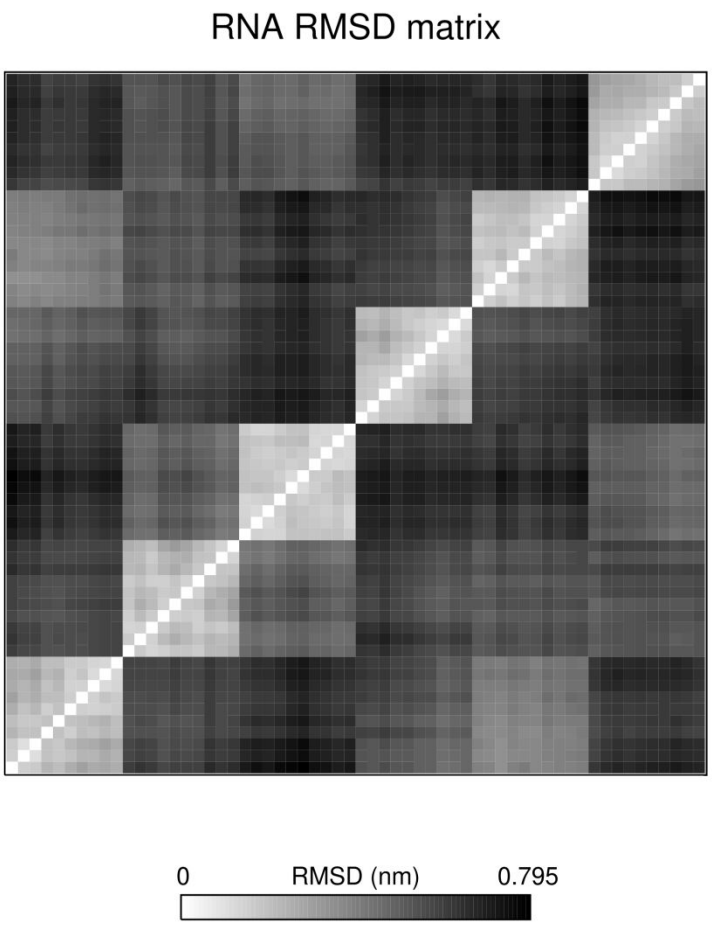

Figure S2. RMSD matrix of initial structures of 60 independent MD simulations. Each block is the group of 10 structures generated based on predicted model. RMSD values increases from white to black.

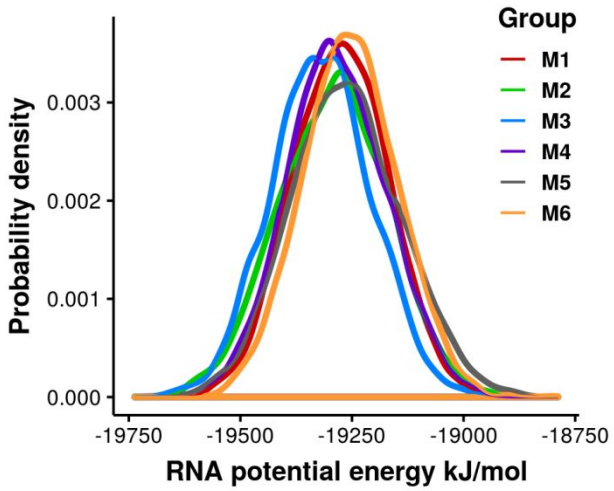

Figure S3. RNA potential energy distribution from 10 independent simulations. The curves are colored by group of simulations. Each group consists of 10 independent simulations starting from the conformations generated based on the same predicted model.

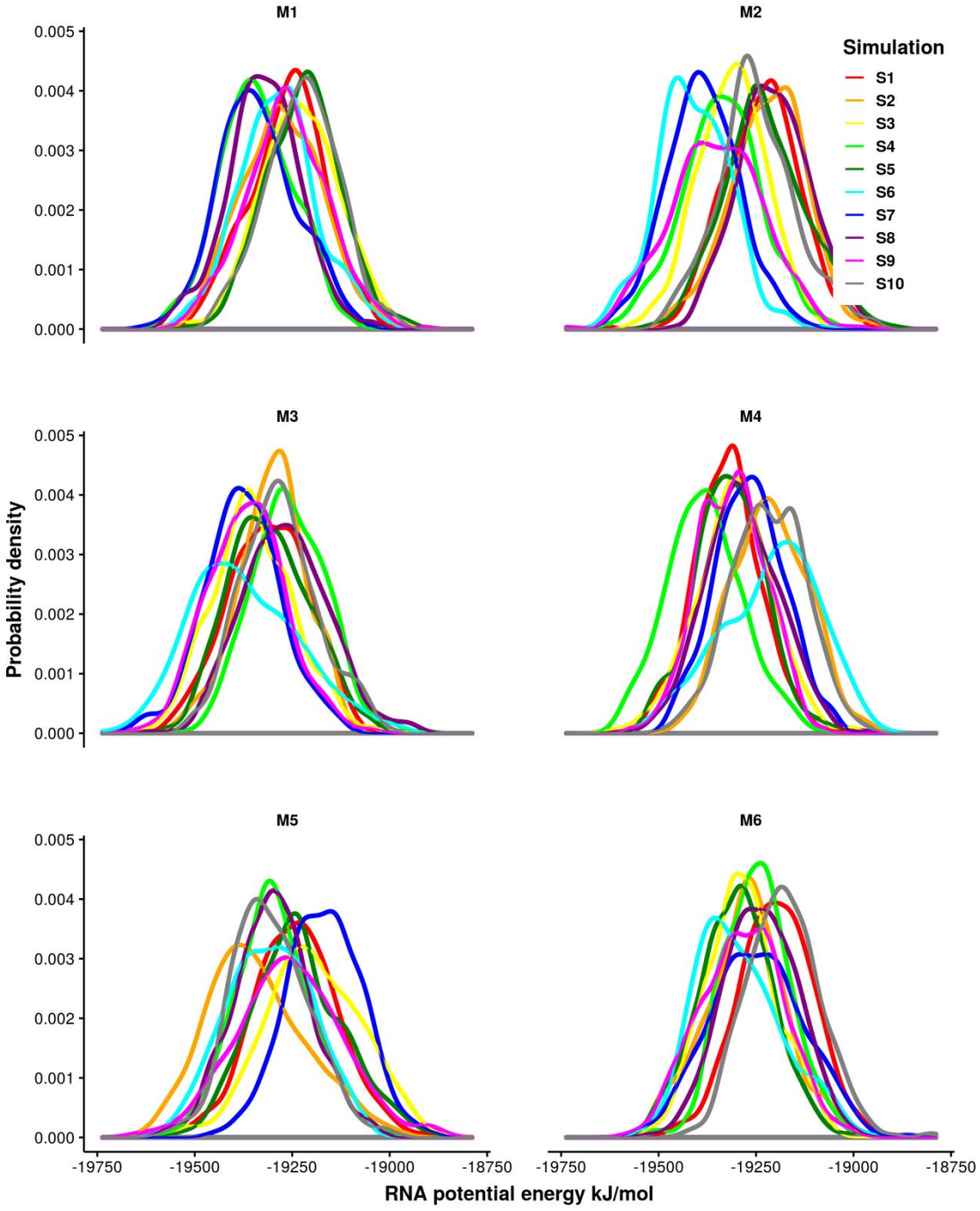

Figure S4. Probability distribution of RNA potential energy calculated from each independent simulation.

Table S1. Top 5 PCs from 1-D clustering assessment.

Rank	PC(Purity)	PC(NMI)	PC(RI)
1st	4 (0.374)	4 (0.201)	5 (0.635)
2nd	1 (0.353)	1 (0.180)	13 (0.617)
3rd	13 (0.321)	9 (0.140)	4 (0.613)
4th	9 (0.317)	13 (0.116)	2243 (0.613)
5th	2 (0.312)	2 (0.111)	2245 (0.612)

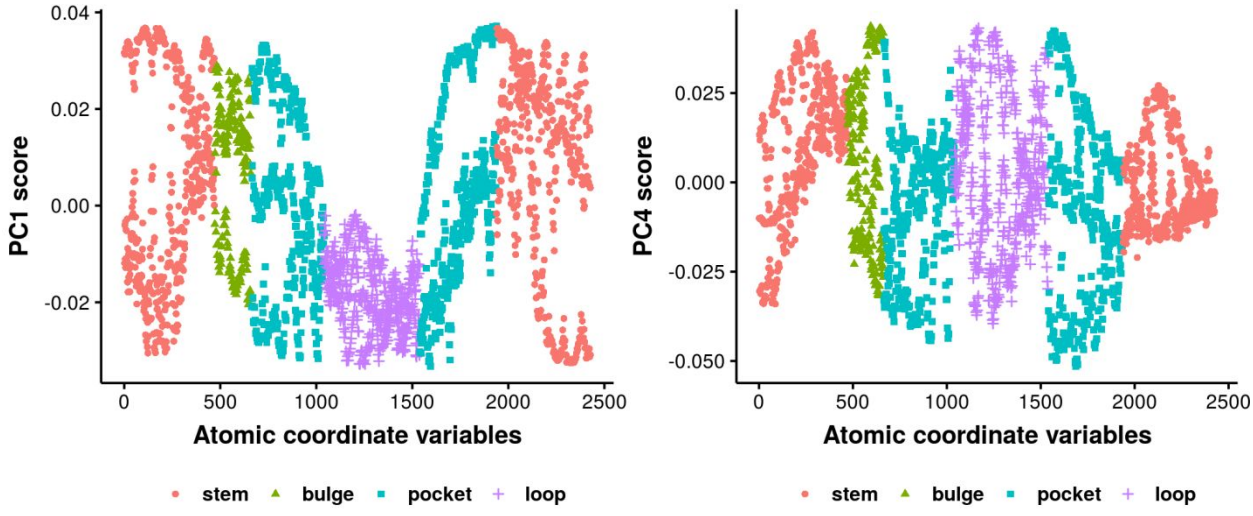

Figure S5. Loadings of PC1 (left) and PC4 (right), colored by different regions of RNA aptamer. The loading of a variable represents the variable's contribution to the projection on the PC. The difference in the average or spread of the loadings from a

specific group of variables shows the importance of this group in defining the PC. PC1 loadings reflect that the pentaloop contributes the most to PC1's ability to separate the groups of simulations. PC4 loadings reflect the stem region's contribution in separating the groups of simulations.

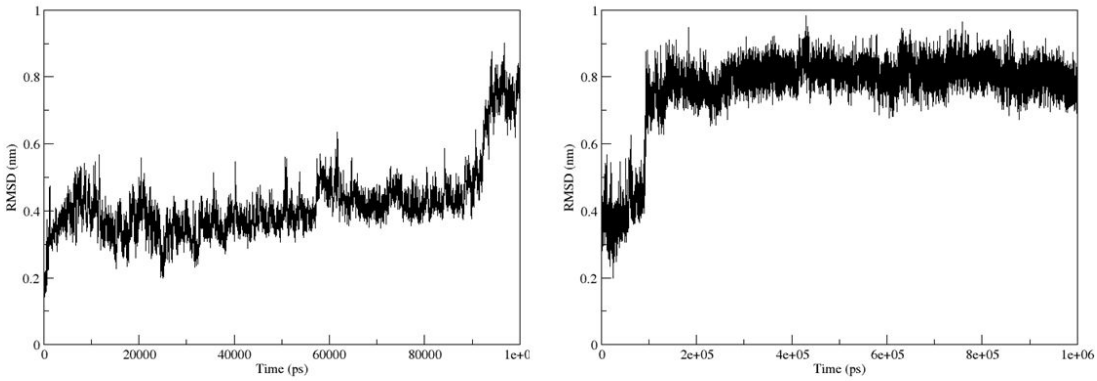

Figure S6. RMSD of 100 ns simulation M1 S1 (left) and 1 μs simulation extending M1 S1 (right).

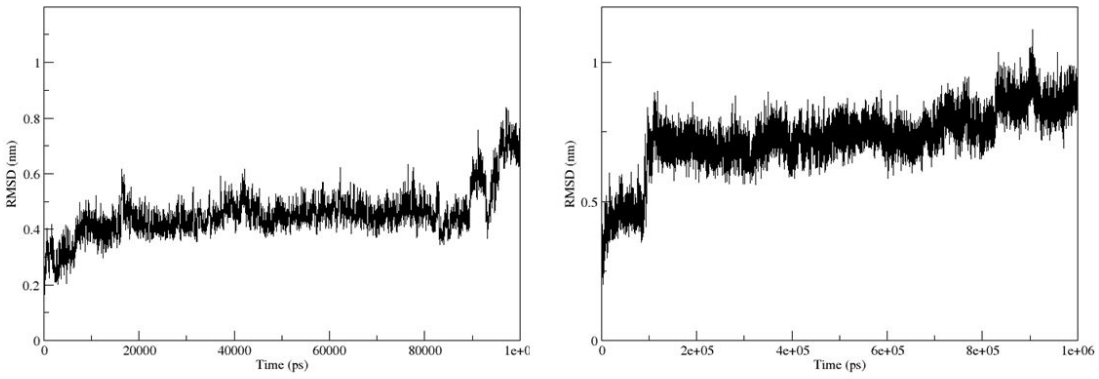

Figure S7. RMSD of 100 ns simulation M2 S7 (left) and 1 μs simulation extending M2 S7 (right).