Coarse Grained Models Reveal Essential Contributions of Topological Constraints to the Conformational Free Energy of RNA Bulges

Anthony M. Mustoe[a], Hashim M. Al-Hashimi[a,b,‡], and Charles L. Brooks III[a,b,*]

Departments of Biophysics[a] and Chemistry[b], University of Michigan, 930 North University Avenue, Ann Arbor, MI 48109

[‡]Present address: Department of Biochemistry, Duke University Medical Center, Durham, NC 27710

*To whom correspondence should be addressed:

Email: brookscl@umich.edu

Phone: (734) 647-6682

Fax: (734) 647-1604

Supporting Methods

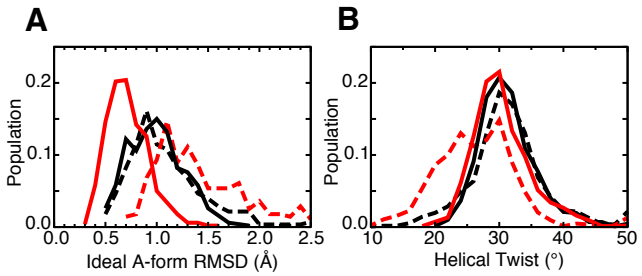
Energy Analysis of Rigid-Body Model Predicted Conformations

2-nt, 3-nt, and 4-nt TOPRNA bulge systems were constructed according to same methods as that used for the equilibrium simulations, using sequences of 5'GCG$(U)^X$CGC and 5'GCGCGC for the two strands. These systems share the same base pairs as the idealized helix used in our previous rigid-body calculations, with $(U)^X$ indicating a poly-U bulge of X-nts. Molecular scaffolds possessing each of the rigid-body predicted (α_h, β_h, γ_h) conformations were also built for each bulge system by performing the necessary rotations to 6-bps of a TOPRNA representation of an idealized A-form helix, as described previously.[1-2] The bulge systems were then targeted to the scaffolds by applying RMSD restraints to the P and S atoms of the base-paired residues with a force constant of 200 kcal/mol/Å^2 and performing 2 ns of dynamics at 300 K. After minimization, the energy of the bulge systems was evaluated, excluding the energetic contributions of the harmonic restraints. We note that (α_h, β_h, γ_h) of the rigid-body topologically allowed spaces that were added as error padding or to account for the intrinsic flexibility of A-form helices were not considered in this analysis given their speculative nature.[1]

Table S1
Sequences of the simulated two-helix junction motifs. Bolded Bs are used to indicate where the 'bulged' single stranded nucleotides were inserted into the sequences, with the inserted bulges consisting of randomized sequences of lengths varying from 1 to 7 residues.

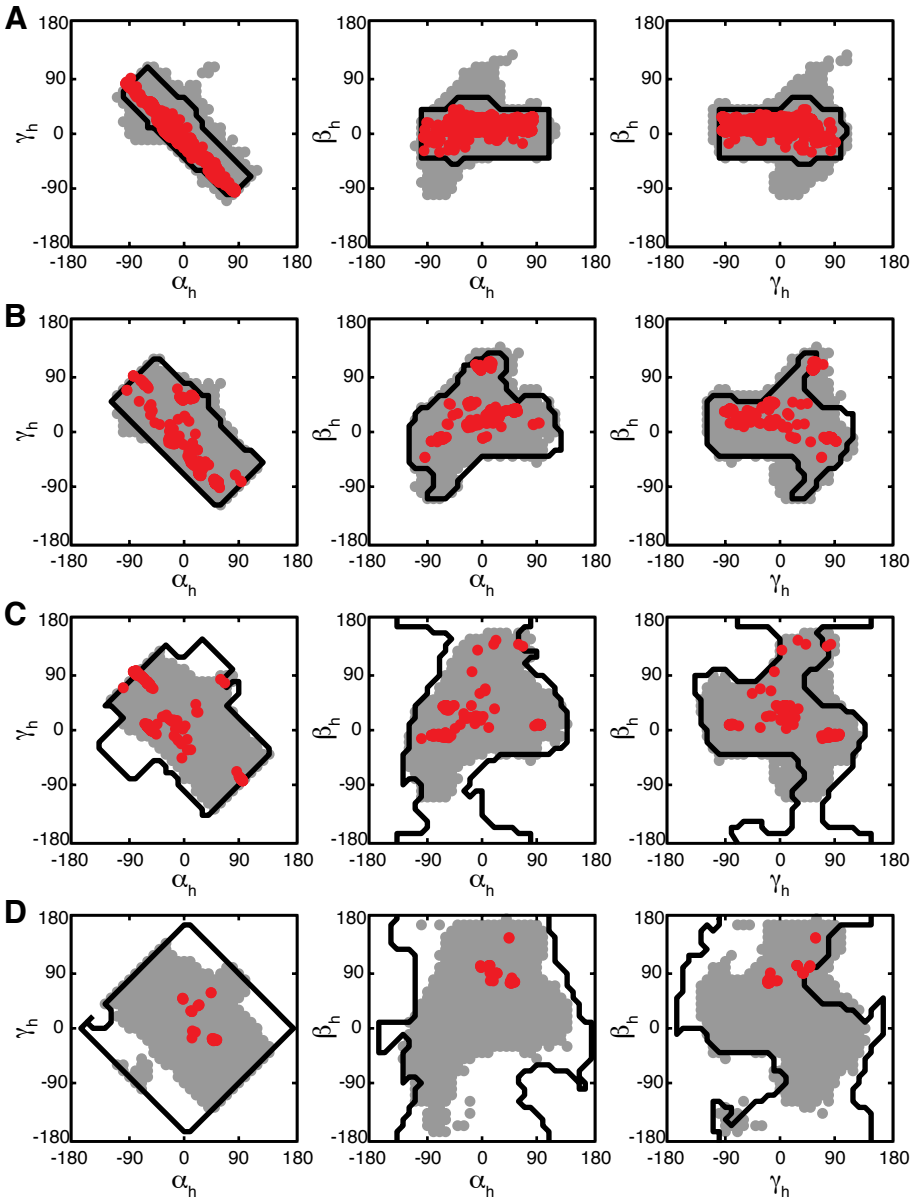
ID	Sequence
1	$^{5'}$GGU**B**CCG CCA GGC$^{5'}$
2	$^{5'}$CAU**B**GCG GUA CGC$^{5'}$
3	$^{5'}$AGC**B**UUU UCG GAA$^{5'}$
4	$^{5'}$AGA**B**AUC UCU UAG$^{5'}$
5	$^{5'}$UUG**B**UCA AAU AGU$^{5'}$
6	$^{5'}$ACU**B**GUG UGA CAC$^{5'}$
7	$^{5'}$GGC**B**CGU CCG GCA$^{5'}$
8	$^{5'}$UGU**B**CCA ACA GGU$^{5'}$
9	$^{5'}$AGC**B**CCG UCG GGC$^{5'}$
10	$^{5'}$CGA**B**GCC GCU UGG$^{5'}$

Figure S1



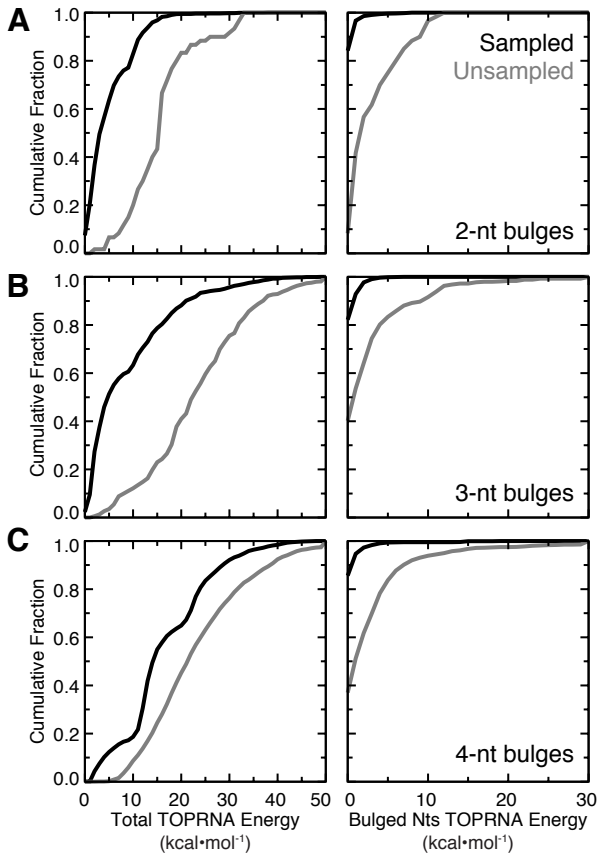
The distribution of helical parameters measured from TOPRNA simulations (solid lines) and NMR-MD dynamic ensemble of HIV-2 TAR (dashed lines).[3] Parameters were measured for the first five base pairs of the lower helix (black) and the four base pairs of the upper helix (red), and the populations of the helical twist parameters represent collections over both the different conformations and constituent base-pair steps. There is a transient shift the base-pair register of the upper helix in the NMR-MD ensemble which explains the divergence between the NMR-MD and TOPRNA helical parameters of this helix.

Figure S2



α_h-γ_h, α_h-β_h, and γ_h-β_h projections of the (α_h, β_h, γ_h) conformations sampled by TOPRNA (gray points), measured in the PDB (red points), and predicted to be allowed by the rigid-body models[1] (black outlines) for 1-nt (A), 2-nt (B), 3-nt (C), and 4-nt (D) bulges.

Figure S3



Cumulative TOPRNA energy distributions of rigid-body-predicted allowed conformations that are sampled (black) and unsampled (gray) during equilibrium TOPRNA simulations for 2-nt (A), 3-nt (B), and 4-nt (C) bulge systems. On the left are distributions of the total system energy and on the right distributions of the bulge-comprising nucleotides' energy.

REFERENCES

1. Mustoe, A. M.; Bailor, M. H.; Teixeira, R. M.; Brooks, C. L., III; Al-Hashimi, H. M. New Insights into the Fundamental Role of Topological Constraints as a Determinant of Two-Way Junction Conformation. *Nucleic Acids Res.* **2012,** *40*, 892-904.
2. Bailor, M. H.; Sun, X. Y.; Al-Hashimi, H. M. Topology Links RNA Secondary Structure with Global Conformation, Dynamics, and Adaptation. *Science* **2010,** *327*, 202-206.
3. Frank, A. T.; Stelzer, A. C.; Al-Hashimi, H. M.; Andricioaei, I. Constructing RNA Dynamical Ensembles by Combining MD and Motionally Decoupled NMR RDCs: New Insights into RNA Dynamics and Adaptive Ligand Recognition. *Nucleic Acids Res.* **2009,** *37*, 3670-9.