

Dynamics of an Adenine-Adenine RNA Conformational Switch From Discrete Path Sampling

Debayan Chakraborty

Department of Chemistry, The University of Texas at Austin, Austin, TX 78712, USA.^{a)}

David J. Wales

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK^{b)}

The study of ‘rare event’ dynamics can be challenging despite continuing advances in computer hardware. A wide variety of methods based on the master equation approach have been developed to tackle such problems, where the focus is on Markovian dynamics between appropriately defined states. In this contribution we employ the discrete path sampling approach to characterize pathways and rates for an adenine-adenine RNA conformational switch. The underlying free energy landscape supports competing structures separated by relatively high barriers, with the two principal funnels leading to the major and minor conformations identified by NMR experiments. The interconversion time scale is predicted to be a few hundred seconds, consistent with the experimental lower bound estimates. We find that conformational switching occurs via stacked intermediates, through a sliding mechanism, in agreement with a previous simulation study. By retaining full dimensionality, and avoiding low-dimensional projections, the mechanism can be described at an atomistic level of detail.

^{a)}Electronic mail: debayan.chakraborty@utexas.edu

^{b)}Electronic mail: dw34@cam.ac.uk

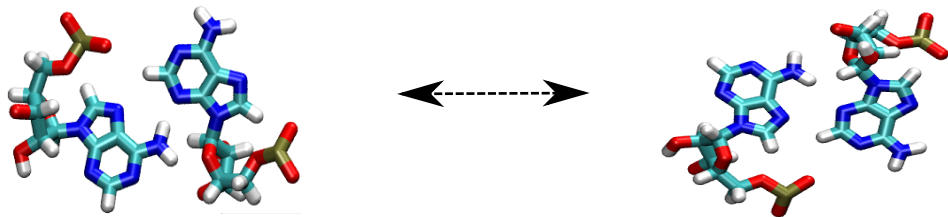
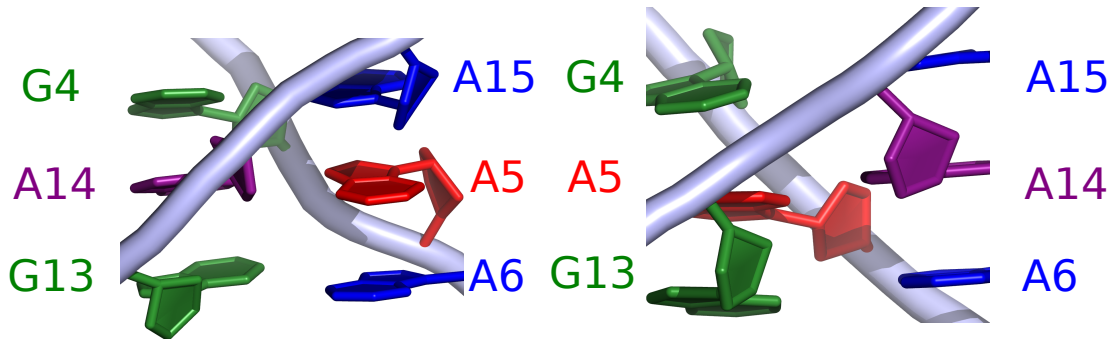
I. INTRODUCTION

Structural plasticity plays a key role in endowing RNA with unique capabilities, and is critical to its function, both as a catalyst and a binding partner.^{1,2} Conformational transitions in RNA occur over a hierarchy of timescales,² and are often associated with long waiting times, corresponding to kinetic trapping in metastable states.³⁻⁵ The manifestation of such rare event dynamics makes it difficult to obtain detailed insight into conformational transitions of biological relevance using conventional simulation techniques.

Considerable effort has been directed towards advancing enhanced sampling methods, many of which are now quite routinely used in unraveling various facets of RNA dynamics and thermodynamics. For example, different implementations of replica-exchange molecular dynamics (REMD) and parallel tempering,⁶⁻⁸ as well as metadynamics,^{9,10} have been used to explore the free energy landscapes of prototypical RNA hairpins, and hence suggest possible folding pathways. Applications of these thermodynamic sampling techniques to more complex RNA molecules are also emerging.^{11,12} However, extracting meaningful kinetic information is not straightforward for these methods, although various recipes for recovering dynamical information have been suggested.^{13,14}

Alternative efforts focused upon dynamics have employed the master equation framework^{15,16} to construct kinetic transition networks.¹⁷⁻¹⁹ This approach has a long history in the context of chemical kinetics, and was probably first applied to relaxation dynamics in atomic clusters by Kunz and coworkers.²⁰⁻²³ There have now been numerous applications to dynamical transitions in proteins, and more recently to RNA folding and conformational switching.^{24,25} Two distinct approaches have been employed to construct the corresponding networks. Methods based on geometry optimization characterize stationary points of the potential energy surface, namely local minima and the transition states that connect them via steepest-descent paths, with post-processing using the tools of statistical mechanics and unimolecular rate theory.^{19,26,27} Alternatively, if the conformations of interest can be connected using explicit dynamics, then a network can be constructed by defining suitable states and their inter-conversion rates; this approach is often called Markov state modeling.²⁸⁻³⁵ A variety of rare events methods based upon explicit dynamics have now been developed that focus on enhanced sampling in some form.³⁶⁻⁴⁸ The geometry optimization and explicit dynamics methodologies are quite complementary: geometry optimization schemes do not suffer from

Adenine-Adenine Conformational Switch



pair RNA duplex (Figure 1). Experiments by Turner and coworkers⁵⁹ indicate that the two adenines can exchange positions and switch between major and minor forms, which differ only in the base-stacking pattern along the minor groove. Based on NMR chemical shifts, the authors estimated a lower bound for the exchange rate of (300 s^{-1}). In a subsequent study,⁶⁰ Mathews and coworkers employed targeted molecular dynamics (TMD), as well as the nudged-elastic-band (NEB)^{61,62} method, to map out pathways between the major and the minor conformations. Their results suggested that the conformational switch preferentially occurs via intermediates where the two adenines are stacked, rather than hydrogen-bonded. Surprisingly, various combinations of force fields and simulation strategies predicted the minor form to be thermodynamically more stable compared to the major form, which seems inconsistent with experiment.

Our results indicate that the major and minor forms predicted by the NMR experiments lie at the bottom of two principal funnels on the energy landscape. The most kinetically relevant transition pathway between the two conformations features stacked intermediates, in agreement with Mathews and coworkers.⁶⁰ Our calculated conformational switching rate from the minor to the major form is 165.9 s^{-1} , in quantitative agreement with the experimental lower bound estimate.⁶⁰ Furthermore, we obtain the correct ordering of free energies for the major and the minor forms. It is possible that force field inaccuracies, or problems associated with the projections associated with low-dimensional reaction coordinates, may have resulted in a reversal of thermodynamic stabilities in the earlier study.

II. METHODOLOGY

System Setup: The initial coordinates for the major and minor forms of the RNA duplex were taken from structures deposited in the PDB database (PDB ID: 2DD2 and 2DD3, respectively).⁵⁹ Following Mathews and coworkers,⁶⁰ we truncated the system by removing the dangling ends consisting of unpaired uracil and purine. The RNA molecule was modeled using a properly symmetrized version⁶³ of the AMBER99bsc0 force field,⁶⁴ employing the χOL3 torsional corrections.⁶⁵ Solvent effects were treated implicitly using a generalized Born model,^{66,67} and an effective monovalent salt concentration of 1.0 M was maintained using the Debye-Hückel approximation.⁶⁸

Molecular Dynamics Simulations: The GPU-enabled version of the AMBER12 code was used to carry out implicit solvent molecular dynamics simulations of length 50 ns starting from local potential energy minima corresponding to the major and minor forms. An infinite cutoff was employed for the non-bonded interactions. The simulation temperature was maintained at 300 K using a Langevin thermostat.⁶⁹

Discrete Path Sampling: DPS simulations were carried out to survey the underlying energy landscape, and characterize transition pathways for the adenine-adenine conformational switch in full dimensionality.

The DPS procedure provides a systematic framework based on geometry optimization for building a transition network, which encodes the global thermodynamics and kinetics.^{50,51,70} A sequence of minima connected by intervening transition states constitutes a discrete path. A local minimum is defined as a stationary point (vanishing gradient) for which all the nonzero normal mode frequencies are positive. Transition states are defined geometrically as saddle points of index one,⁴⁹ with a single imaginary normal mode frequency corresponding to a local reaction coordinate. Steepest-descent paths from the transition state define the connected local minima. Here, we employed a modified version of the L-BFGS algorithm⁷¹ for local minimizations to characterize approximate steepest-descent paths. The doubly-nudged⁷² elastic band^{61,62} (DNEB) method was used to find candidate transition state structures between intervening minima. These transition state candidates were further optimized using hybrid-eigenvector following.^{73,74} We note that it is important to establish the connectivity of local minima by checking the steepest-descent pathways, since the original DNEB interpolation often misses some of the intervening transition states and local minima, especially for lengthy gaps when a small number of images are used. It is also noteworthy that the kinetic transition network usually requires substantial additional sampling to identify kinetically relevant pathways once an initial connected path has been found.

A convergence criterion of 10^{-6} kcal/mol/Å for the root-mean-square gradient was employed for all the geometry optimizations. The OPTIM code⁷⁵ interfaced with the AMBER9 package⁷⁶ was used to carry out the local minimizations, the transition state searches, and pathway analysis.

Representative structures corresponding to the major and minor conformations identified from the MD simulations were used as the initial endpoints for the DPS calculations. After a

fully connected initial discrete path was found between the two structures, further refinement of the resulting kinetic transition network was carried out using several schemes available within the PATHSAMPLE code.⁷⁷ These schemes have been described in detail elsewhere, and for brevity we will only outline the key features. To locate pathways with lower energy barriers we used the SHORTCUT BARRIER scheme,^{78–80} which selects pairs of minima on either side of, and an equal number of steps away from, the largest barriers for additional connection attempts. The SHORTCUT scheme,^{78–80} which attempts to connect minima that are close together in configuration space, but are separated by a minimum number of steps on the discrete paths, was used to locate shorter paths between the endpoints. Finally, the UNTRAP scheme,⁷⁹ which selects minima for connection-making attempts based on the ratio of the potential energy barrier to the potential energy difference, was used to remove artificial frustration from the transition network, which provides a more faithful representation of the global dynamics. The three refinement schemes were used in sequence until the rate constant for the transition between the major and minor forms was converged to within an order of magnitude. Given the systematic sources of error in the empirical potential, the sampling, the use of harmonic densities of states, and the assumption of Markovian dynamics between appropriately regrouped states,⁸¹ we would not expect better than order of magnitude agreement with experiment.

Estimating free energies and interconversion rates: The vibrational partition functions associated with the minima and the transition states in the network were estimated using a harmonic approximation.^{82,83}

$$Z_i(T) = \frac{n_i e^{-V_i/kT}}{(h\bar{\nu}_i/kT)^\kappa}. \quad (1)$$

In equation (1), V_i denotes the potential energy of minimum i , n_i is the number of distinct permutational isomers of i , $\bar{\nu}_i$ denotes the geometric mean normal mode frequency associated with minimum i , and $\kappa = 3N - 6$ is the number of vibrational degrees of freedom, where N is the number of atoms.

The full canonical partition function, $Z(T)$, can be written as a sum of all the contributions from the different catchment basins corresponding to each local minimum.

$$Z(T) = \sum_i^N Z_i(T). \quad (2)$$

The local free energy, $F_i(T)$, and the equilibrium occupation probability, $p_i^{eq}(T)$, of each minimum are:

$$F_i(T) = kT \ln Z_i(T), \quad (3)$$

$$\text{and} \quad p_i^{eq}(T) = \frac{Z_i(T)}{Z(T)}. \quad (4)$$

The partition functions and the free energies for the transition states are defined in an analogous way, except that the normal mode frequency corresponding to the unique negative Hessian eigenvalue (imaginary normal mode frequency) is excluded.

The minimum-to-minimum rate constants are estimated using harmonic transition state theory (TST).

$$k_i^\ddagger(T) = \frac{kT}{h} \frac{Z^\ddagger(T)}{Z_i(T)} e^{-\beta\Delta V}. \quad (5)$$

In equation (5), $Z^\ddagger(T)$ denotes the partition function of the transition state; $Z_i(T)$ is the partition function of minimum i ; ΔV is the potential energy difference between the transition state and minimum i . The total rate constant $k_{ji}(T)$ for an elementary transition from minimum i to minimum j is obtained by summing the $k_i^\ddagger(T)$ values for all transition states that connect the two minima.

The minimum-to-minimum rate constants are used in calculating global dynamical properties. Within the steady-state approximation for intervening minima, the rate constants, k_{AB}^{SS} and k_{BA}^{SS} between reactant (A) and product (B) states, can be expressed as weighted sums over all discrete paths in the network, assuming that the dynamics between adjacent minima or lumped states⁸¹ is Markovian:⁵⁰

$$k_{AB}^{SS} = \frac{1}{p_B^{eq}} \sum_{a \leftarrow b} \frac{k_{ai_1} k_{i_1 i_2} k_{i_2 i_3} \dots k_{i_n b} p_b^{eq}}{\sum_{j_1} k_{j_1 i_1} \sum_{j_2} k_{j_2 i_2} \sum_{j_3} k_{j_3 i_3} \dots \sum_{j_n} k_{j_n i_n}}. \quad (6)$$

In terms of transition probabilities, $P_{\gamma\alpha}$ between directly connected minima γ and α , equation (6) can be rewritten as:⁵⁰

$$k_{AB}^{SS} = \frac{1}{p_B^{eq}} \sum_{a \leftarrow b} P_{ai_1} P_{i_1 i_2} P_{i_2 i_3} \dots P_{i_n b} p_b^{eq} \tau_b^{-1}. \quad (7)$$

The individual sums in the denominators of equation (6) consist of the unimolecular rate constants for all direct transitions from minimum j_k to i_k . The discrete path that makes the largest contribution to the steady-state rate constant is termed the ‘fastest path’, and can be extracted from the network using Dijkstra’s shortest path algorithm with appropriate edge weights^{78,84} $-\ln P_{\alpha\beta}$, providing access to the product of transition probabilities in equation

(7). To characterize additional paths in order of their contribution to k^{SS} we employ the recursive enumeration algorithm⁸⁵ within the same framework,⁸⁶ and these pathways were examined to deduce mechanistic details of the adenine conformational switch. The steady-state approximation for the intervening minima can be relaxed to yield rate constants k_{AB} and k_{BA} that correspond to the mean first passage times between reactants and products.⁸⁷ To extract these values we employ a graph transformation technique (NGT)^{87,88} where minima in the intervening region are progressively removed, and the transition probabilities as well as the waiting times are renormalized to conserve the average mean first passage time (MFPT).^{87,89}

We compute the rate constants in conjunction with a recursive free energy regrouping scheme,⁸¹ which lumps together structures separated by free energy barriers below a certain threshold into a single macrostate. This approach is similar in spirit to the kinetic clustering schemes employed in methods based upon explicit dynamics.^{26,33} In the regrouping, the original reactant and product states are expanded into ensembles of conformations assumed to be in local equilibrium, and hence the global dynamics can be directly compared to the observation time scale of experiments.⁹⁰ After regrouping, the equilibrium occupation probability and the free energy associated with group J are

$$p_J^{eq}(T) = \sum_{j \in J} p_j^{eq}(T), \quad (8)$$

$$F_J = -kT \ln \sum_{j \in J} Z_j(T). \quad (9)$$

where minimum j is a member of group J . The free energy of the group of transition states linking J and K is:⁸¹

$$F_{KJ}^\ddagger = -kT \ln \sum_{k \leftarrow j} Z_{kj}^\ddagger(T) \equiv -kT \ln Z_{KJ}^\ddagger(T), \quad (10)$$

To analyze global dynamics corresponding to regrouped databases, the rate constants corresponding to transitions between different free energy groups are required, which can then be used in the appropriate expressions for rate constants and committor probabilities.⁸⁷ The intergroup rate constant from J to K is:⁸¹

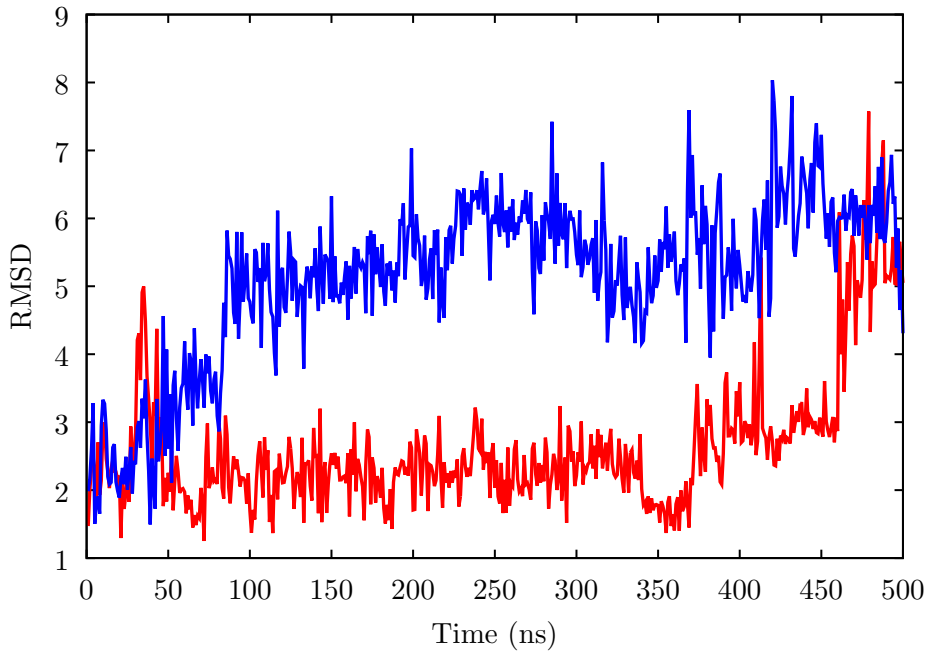
$$\begin{aligned} k_{KJ} &= \sum_{k \leftarrow j} \frac{p_j^{eq}(T)}{p_J^{eq}(T)} k_{kj}(T) = \sum_{k \leftarrow j} \frac{Z_j(T)}{Z_J(T)} \frac{kT}{h} \frac{Z_{kj}^\ddagger(T)}{Z_j(T)}, \\ &= \frac{kT}{h} \frac{Z_{KJ}^\ddagger(T)}{Z_J(T)} = \frac{kT}{h} e^{-[F_{KJ}^\ddagger(T) - F_J(T)]/kT}. \end{aligned} \quad (11)$$

Disconnectivity Graphs: The potential and free energy landscapes constructed using DPS were visualized in the form of disconnectivity graphs.^{91–93} In these diagrams the landscape is segregated into disjoint sets of minima known as superbasins, at regular intervals of energy. Minima within each superbasin are mutually accessible via transition states lying below the energy threshold, while transitions out of superbasins require higher energy barriers to be surmounted. A key feature of disconnectivity graphs is that they should provide a faithful representation of the underlying kinetics,⁹⁴ avoiding low-dimensional projection, which may lump together states separated by significant energy barriers.

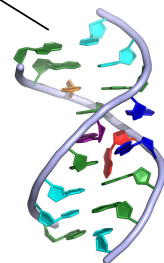
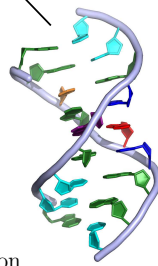
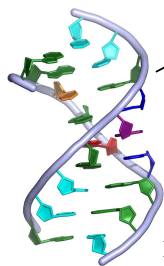
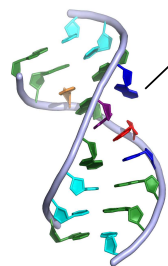
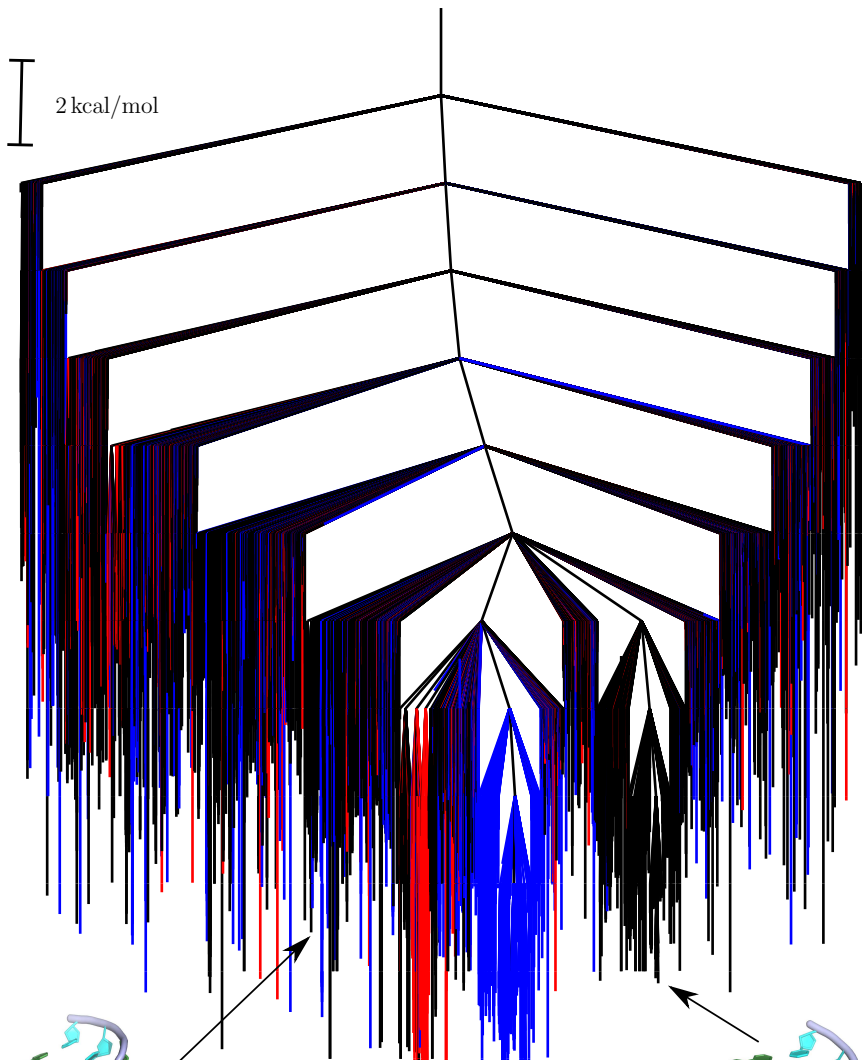
III. RESULTS AND DISCUSSION

Although the time scale of the MD simulations is not long enough to explore complete conformational switching between the major and minor forms, they provide some useful insight into the local fluctuations around the NMR structures. As illustrated by the time evolution of the RMSD (Figure 2), the minor conformation is quite flexible, whereas the major conformation seems to exhibit some degree of conformational rigidity over short time scales. The large jumps in the RMSD profile for both the major and the minor forms are associated with the flipping of the A5 nucleobase out of the helix, which in turn leads to the loss of the hydrogen-bonding interactions between the central adenine-adenine sheared base pair.

In contrast to short time scale MD simulations, discrete path sampling provides a viable route to global exploration of the underlying energy landscape, and facilitates the detailed study of transition pathways of varying complexity. To identify suitable endpoints for DPS simulations, snapshots along the MD trajectories were saved every 10 ps, and subsequently locally minimized. The lowest potential energy minima corresponding to the minor and major forms were selected as the representative reactant and product states, respectively. After an initial discrete pathway was found between these endpoints, connection attempts were made for the rest of the local minima (identified from the MD simulations) to either the reactant or the product state in a pairwise fashion, based on the lowest distance in configuration space after optimal alignment. The transition network was subsequently refined employing the strategies described in the Methodology section.



2 kcal/mol



Stacked conformation

Major conformation

Minor conformation

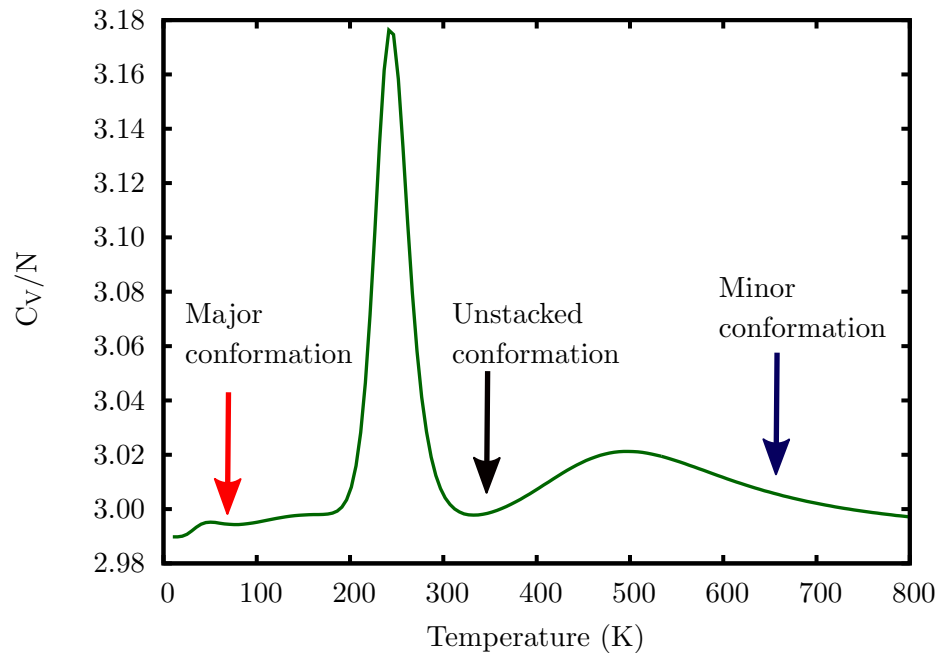
Unstacked conformation

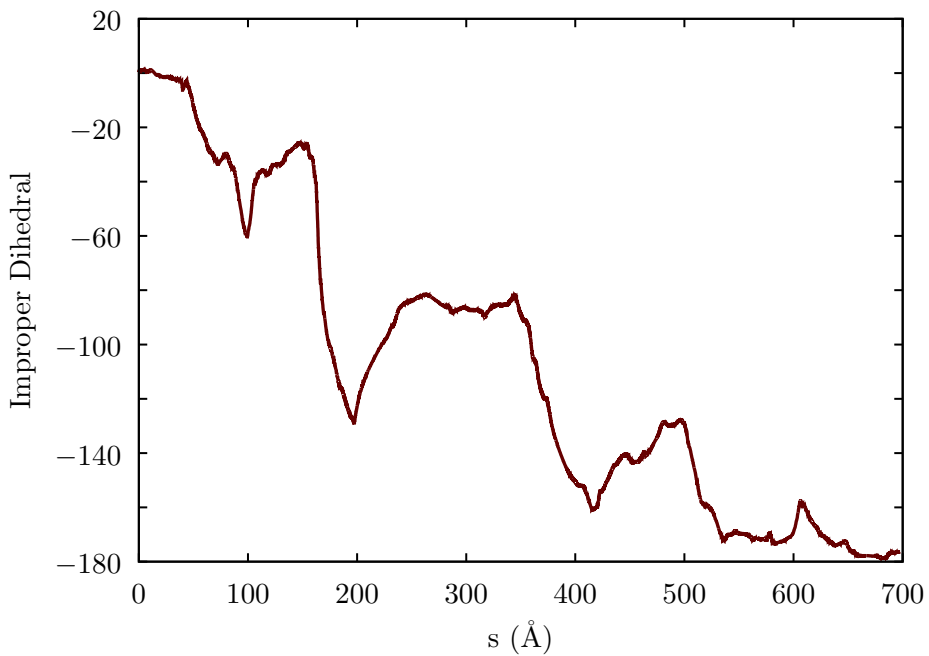
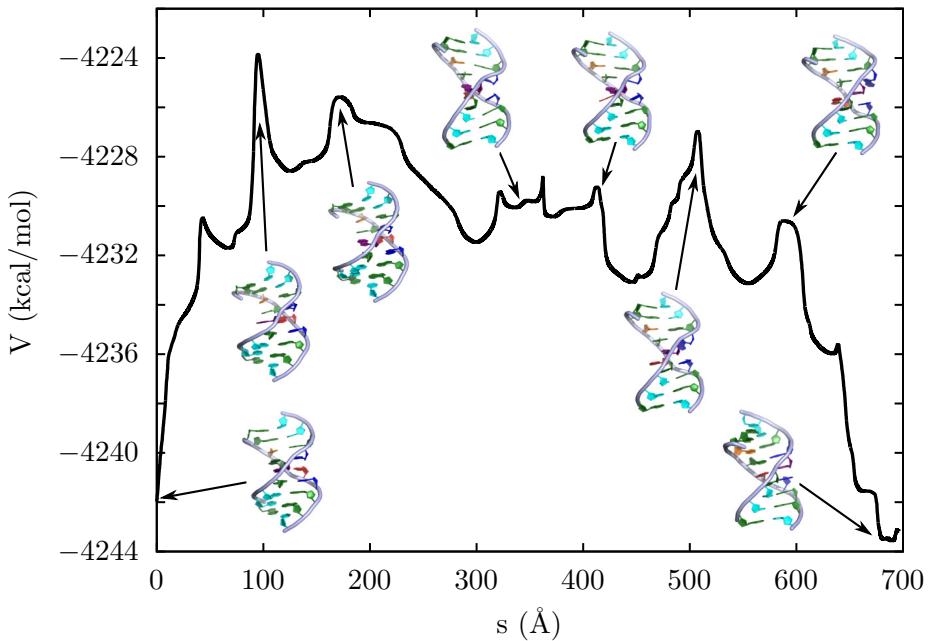
by about 1.8 kcal/mol with respect to the minor form. This estimate is in good agreement with the upper bound of the experimental free energy difference (around 1.3 kcal/mol), calculated based on the population analysis of the major and the minor forms. However, our results differ from the earlier study by Mathews and coworkers,⁵⁹ who predicted that the minor form is stabilized by around 7.04 kcal/mol. In that work the authors estimated the free energy difference based on a one-dimensional representation of the free energy surface along a predefined reaction coordinate. It is possible that the low-dimensional projection could have resulted in the opposite stability that was reported.

In addition to the major and the minor forms, we also identify another prominent funnel on the landscape, which primarily consists of structures in which the A5 nucleobase is flipped out of the RNA helix. In fact, these structures are quite similar to those that were encountered during the initial MD simulations. The stacked conformation, which was predicted by Mathews and coworkers to be a key intermediate along the conformational switching pathway,⁶⁰ lies at the bottom of a narrow subfunnel.

The overall topography of the landscape leads to multiple peaks in the heat capacity profile (Figure 4). Such features are characteristics of multifunnel landscapes associated with conformational switches,^{54,95,96} and were first observed for the solid-solid type transition between competing morphologies observed in an atomic cluster.^{97,98} At low temperature, where enthalpy dominates, the major conformation is thermodynamically favored, although kinetic trapping in the competing basins is likely to impede relaxation. At intermediate temperatures the competing effects of enthalpy and entropy cause the free energy global minimum to switch to the unstacked conformation. At high temperatures, the minor conformation dominates the equilibrium population. It is likely that the transition temperatures associated with the switching between the different funnels may be overestimated here due to the inherent limitations of the harmonic superposition approximation.⁷⁰ Nonetheless, the relative positions of the peaks in the heat capacity curve provide additional insight into the contrasting flexibilities of the major and the minor forms observed during the MD simulations. The major conformation, which is stabilized by enthalpy, undergoes only restricted motions for much of the trajectory. On the other hand, the minor form is more flexible and exhibits substantially larger fluctuations.

The large free energy barrier separating the major and the minor conformations suggests that conformational switching between the two forms is a slow process. Using the NGT





contrast to the relatively smooth profile reported before. This discrepancy suggests that there are additional degrees of freedom coupled to the dihedral degree of freedom, which could have important implications in the analysis of thermodynamics and kinetics. Interestingly, we do not find any evidence for a 3' facing mechanism among the kinetically relevant set of pathways.

IV. CONCLUSION

In this study, we have provided atomistic insight into the transition mechanism between two conformations of an adenine-adenine RNA conformational switch using discrete path sampling. The interconversion kinetics are relatively slow, making this a challenging problem for conventional simulation techniques. Conformational switching between the two forms occurs via stacked intermediates, along a preferred sliding direction. Overall, the generic features of the transition pathway are found to be consistent with a previous study by Mathews and coworkers.⁵⁹

In agreement with the experimental findings,⁵⁹ we find that the major form is stabilized by around 1.8 kcal/mol with respect to the minor form in terms of free energy. The interplay between enthalpy and entropy leads to multiple features in the heat capacity profile. Such signatures are likely to be characteristic of conformational switches, and suggest that a subtle modulation of environmental conditions could result in population shifts to favor one particular form. In the future, it would be interesting to see how the landscape evolves as function of tunable parameters, which could provide further insight into the remarkable functionality of more complex conformational switches.

V. ACKNOWLEDGEMENTS

This research was funded by the EPSRC and the ERC. DC would like to thank the Cambridge Commonwealth, European and International Trusts for providing a scholarship while he was a PhD student at the University of Cambridge.

REFERENCES

- ¹A. M. Mustoe, C. L. Brooks, and H. M. Al-Hashimi, *Annu. Rev. Biochem.* **83**, 441 (2014).
- ²E. A. Dethoff, J. Chugh, A. M. Mustoe, and H. M. Al-Hashimi, *Nature* **482**, 323 (2012).
- ³D. Thirumalai and S. A. Woodson, *Acc. Chem. Res.* **29**, 433 (1996).
- ⁴D. Thirumalai and C. Hyeon, *Biochemistry* **44**, 4957 (2005).
- ⁵J. C. Lin, C. Hyeon, and D. Thirumalai, *J. Phys. Chem. Lett.* **3**, 3616 (2012).
- ⁶A. E. Garcia and D. Paschek, *J. Am. Chem. Soc.* **130**, 815 (2008).
- ⁷A. A. Chen and A. E. Garcia, *Proc. Natl. Acad. Sci. USA* **110**, 16820 (2013).
- ⁸P. Kührová, P. Banáš, R. B. Best, J. Šponer, and M. Otyepka, *J. Chem. Theory Comput.* **9**, 2115 (2013).
- ⁹S. Haldar, P. Kührová, P. Banáš, V. Spiwok, J. Šponer, P. Hobza, and M. Otyepka, *J. Chem. Theory Comput.* **11**, 3866 (2015).
- ¹⁰S. Bottaro, P. Banáš, J. Šponer, and G. Bussi, *J. Phys. Chem. Lett.* **7**, 4032 (2016).
- ¹¹X. Xue, W. Yongjun, and L. Zhihong, *J. Theor. Biol.* **365**, 265 (2015).
- ¹²Y. Bian, J. Zhang, J. Wang, J. Wang, and W. Wang, *PloS One* **10**, e0129089 (2015).
- ¹³N. V. Buchete and G. Hummer, *Phys. Rev. E* **77**, 030902 (2008).
- ¹⁴P. Tiwary and M. Parrinello, *Phys. Rev. Lett.* **111**, 230602 (2013).
- ¹⁵N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
- ¹⁶R. E. Kunz, *Dynamics of First-Order Phase Transitions* (Deutsch, Thun, 1995).
- ¹⁷F. Noé and S. Fischer, *Curr. Opin. Struct. Biol.* **18**, 154 (2008).
- ¹⁸D. Prada-Gracia, J. Gómez-Gardenes, P. Echenique, and F. Fernando, *PLoS Comput. Biol.* **5**, e1000415 (2009).
- ¹⁹D. J. Wales, *Curr. Opin. Struct. Biol.* **20**, 3 (2010).
- ²⁰R. S. Berry and R. Breitengraser-Kunz, *Phys. Rev. Lett.* **74**, 3951 (1995).
- ²¹R. E. Kunz and R. S. Berry, *J. Chem. Phys.* **103**, 1904 (1995).
- ²²B. Kunz, R. S. Berry, and T. Astakhova, *Surf. Rev. Lett.* **3**, 307 (1996).
- ²³R. E. Kunz, P. Blaudeck, K. H. Hoffmann, and R. S. Berry, *J. Chem. Phys.* **108**, 2576 (1998).
- ²⁴G. Pinamonti, J. Zhao, D. E. Condon, F. Paul, F. Noé, D. H. Turner, and G. Bussi, *J. Chem. Theory Comput.* **13**, 926 (2017).

- ²⁵B. M. Warfield and P. C. Andersen, *PloS One* **12**, e0176229 (2017).
- ²⁶F. Noé and S. Fischer, *Curr. Opin. Struct. Biol.* **18**, 154 (2008).
- ²⁷J. A. Joseph, K. Rder, D. Chakraborty, R. G. Mantell, and D. J. Wales, *Chem. Commun.* **53**, 6974 (2017).
- ²⁸C. Schütte, A. Fischer, W. Huisinga, and P. Deuffhard, *J. Comput. Phys.* **151**, 146 (1999).
- ²⁹M. Shirts and V. S. Pande, *Science* **290**, 1903 (2000).
- ³⁰N. Singhal, C. D. Snow, and V. S. Pande, *J. Chem. Phys.* **121**, 415 (2004).
- ³¹W. C. Swope, J. W. Pitera, and F. Suits, *J. Phys. Chem. B* **108**, 6571 (2004).
- ³²J. D. Chodera, K. A. Dill, N. Singhal, V. S. Pande, W. C. Swope, and J. W. Pitera, *J. Chem. Phys.* **126**, 155101 (2007).
- ³³V. S. Pande, K. Beauchamp, and G. R. Bowman, *Methods* **52**, 99 (2010).
- ³⁴J. H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schutte, and F. Née, *J. Chem. Phys.* **134**, 174105 (2011).
- ³⁵B. E. Husic and V. S. Pande, *J. Am. Chem. Soc.* **140**, 2386 (2018).
- ³⁶G. A. Huber and S. Kim, *Biophys. J.* **70**, 97 (1996).
- ³⁷C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *J. Chem. Phys.* **108**, 1964 (1998).
- ³⁸P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annu. Rev. Phys. Chem.* **53**, 291 (2002).
- ³⁹T. S. van Erp, D. Moroni, and P. G. Bolhuis, *J. Chem. Phys.* **118**, 7762 (2003).
- ⁴⁰A. K. Faradjian and R. Elber, *J. Chem. Phys.* **120**, 10880 (2004).
- ⁴¹A. M. A. West, R. Elber, and D. Shalloway, *J. Chem. Phys.* **126**, 145104 (2007).
- ⁴²G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello, *J. Amer. Chem. Soc.* **128**, 13435 (2006).
- ⁴³R. J. Allen, D. Frenkel, and P. R. t. Wolde, *J. Chem. Phys.* **124**, 024102 (2006).
- ⁴⁴E. Vanden-Eijnden, M. Venturoli, G. Ciccotti, and R. Elber, *J. Chem. Phys.* **129**, 174102 (2008).
- ⁴⁵K. Kuczera, G. S. Jas, and R. Elber, *J. Phys. Chem. A* **113**, 7461 (2009).
- ⁴⁶B. W. Zhang, D. Jasnow, and D. M. Zuckerman, *J. Chem. Phys.* **132**, 054107 (2010).
- ⁴⁷E. Vanden-Eijnden and M. Venturoli, *J. Chem. Phys.* **130**, 194101 (2009).
- ⁴⁸R. Elber, J. Bello-Rivas, P. Ma, A. Cardenas, and A. Fathizadeh, *Entropy* **19**, 219 (2017).
- ⁴⁹J. N. Murrell and K. J. Laidler, *Trans. Faraday. Soc.* **64**, 371 (1968).

- ⁵⁰D. J. Wales, *Mol. Phys.* **100**, 3285 (2002).
- ⁵¹D. J. Wales, *Mol. Phys.* **102**, 891 (2004).
- ⁵²D. Chakraborty, R. Collepardo-Guevara, and D. J. Wales, *J. Am. Chem. Soc.* **136**, 18052 (2014).
- ⁵³D. Chakraborty and D. J. Wales, *Phys. Chem. Chem. Phys.* **19**, 878 (2016).
- ⁵⁴D. Chakraborty and D. J. Wales, *J. Phys. Chem. Lett.* **9**, 229 (2018).
- ⁵⁵I. Yildirim, D. Chakraborty, M. Disney, D. J. Wales, and G. C. Schatz, *J. Chem. Theory Comput.* **11**, 4943 (2015).
- ⁵⁶T. Cragolini, D. Chakraborty, J. Šponer, P. Derreumaux, S. Pasquali, and D. J. Wales, *J. Chem. Phys.* **147** (2017), 10.1063/1.4997377.
- ⁵⁷J. S. Klimavicz, K. Röder, and D. J. Wales, *J. Chem Theory Comput.* **14**, 3870 (2018).
- ⁵⁸N. B. Leontis and E. Westhof, *RNA* **7**, 499 (2001).
- ⁵⁹G. Chen, S. D. Kennedy, J. Qiao, T. R. Krugh, and D. H. Turner, *Biochemistry* **45**, 6889 (2006).
- ⁶⁰K. P. Van Nostrand, S. D. Kennedy, D. H. Turner, and D. H. Mathews, *J. Chem. Theory Comput.* **7**, 3779 (2011).
- ⁶¹G. Henkelman, B. P. Uberuaga, and H. Jönsson, *J. Chem. Phys.* **113**, 9901 (2000).
- ⁶²G. Henkelman and H. Jónsson, *J. Chem. Phys.* **113**, 9978 (2000).
- ⁶³E. Malolepsza, B. Strodel, M. Khalili, S. Trygubenko, S. N. Fejer, and D. J. Wales, *J. Comput. Chem.* **31**, 1402 (2010).
- ⁶⁴A. Pérez, I. Marchán, D. Svozil, J. Šponer, T. E. Cheatham, C. A. Loughton, and M. Orozco, *Biophys. J.* **92**, 3817 (2007).
- ⁶⁵M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham, and P. Jurečka, *J. Chem. Theory Comput.* **7**, 2886 (2011).
- ⁶⁶A. Onufriev, D. Bashford, and D. A. Case, *Proteins* **55**, 383 (2004).
- ⁶⁷A. Onufriev, D. Bashford, and D. A. Case, *J. Phys. Chem. B.* **104**, 3712 (2000).
- ⁶⁸J. Srinivasan, M. W. Trevathan, P. Beroza, and D. A. Case, *Theor. Chem. Acc.* **101**, 426 (1999).
- ⁶⁹R. J. Loncharich, B. R. Brooks, and R. W. Pastor, *Biopolymers* **32**, 523 (1992).
- ⁷⁰D. J. Wales, *Energy Landscapes* (Cambridge University Press, U.K., 2003).
- ⁷¹D. Liu and J. Nocedal, *Math. Program.* **45**, 503 (1989).
- ⁷²S. A. Trygubenko and D. J. Wales, *J. Chem. Phys.* **120**, 2082 (2004).

- ⁷³L. J. Munro and D. J. Wales, *Phys. Rev. B* **59**, 3969 (1999).
- ⁷⁴Y. Zheng, P. Xiao, and G. Henkelman, *J. Chem. Phys.* **140**, 044115 (2014).
- ⁷⁵D. J. Wales, “OPTIM: A program for geometry optimisation and pathway calculations,” <http://www-wales.ch.cam.ac.uk/software.html> ().
- ⁷⁶D. A. Case, T. A. Darden, T. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, and I. Kolossváry, “Amber 9,” (2006).
- ⁷⁷D. J. Wales, “PATHSAMPLE: A program for generating connected stationary point databases and extracting global kinetics,” <http://www-wales.ch.cam.ac.uk/software.html> ().
- ⁷⁸J. M. Carr, S. A. Trygubenko, and D. J. Wales, *J. Chem. Phys.* **122**, 234903 (2005).
- ⁷⁹B. Strodel, C. S. Whittleston, and D. J. Wales, *J. Amer. Chem. Soc.* **129**, 16005 (2007).
- ⁸⁰D. J. Wales, J. M. Carr, M. Khalili, V. K. de Souza, B. Strodel, and C. S. Whittleston, in *Proteins: Energy, Heat and Signal Flow*, Computation in chemistry (CRC Press, 2009) p. 315.
- ⁸¹J. M. Carr and D. J. Wales, *J. Phys. Chem. B* **112**, 8760 (2008).
- ⁸²M. R. Hoare and J. J. McInnes, *Faraday Discuss. Chem. Soc.* **61**, 12 (1976).
- ⁸³M. R. Hoare, “Advances in chemical physics,” (John Wiley and Sons, USA, 1979) pp. 49–129.
- ⁸⁴J. M. Carr and D. J. Wales, in *Latest Advances in Atomic Cluster Collisions: Structure and Dynamics from the Nuclear to the Biological Scale*, edited by J.-P. Connerade and A. Solov’yov (Imperial College Press, London, 2008) pp. 321–330.
- ⁸⁵V. M. Jiménez and A. Marzal, in *Algorithm Engineering: 3rd International Workshop, WAE’99, London, UK, July 1999.*, Vol. 1668, edited by J. S. Vitter and C. D. Zaroliagis (Springer Berlin / Heidelberg, 1999) pp. 15–29.
- ⁸⁶J. M. Carr and D. J. Wales, *Phys. Chem. Chem. Phys.* **11**, 3341 (2009).
- ⁸⁷D. J. Wales, *J. Chem. Phys.* **130**, 204111 (2009).
- ⁸⁸J. D. Stevenson and D. J. Wales, *J. Chem. Phys.* **141**, 041104(1) (2014).
- ⁸⁹R. S. MacKay and J. D. Robinson, *Phil. Trans. R. Soc. A* **376**, 20170232 (2018).
- ⁹⁰D. J. Wales and P. Salamon, *Proc. Natl. Acad. Sci. USA* **111**, 617 (2014).
- ⁹¹O. M. Becker and M. Karplus, *J. Chem. Phys.* **106**, 1495 (1997).
- ⁹²D. J. Wales, M. A. Miller, and T. R. Walsh, *Nature* **394**, 758 (1998).
- ⁹³S. V. Krivov and M. Karplus, *J. Chem. Phys.* **117**, 10894 (2002).

- ⁹⁴S. V. Krivov and M. Karplus, Proc. Natl. Acad. Sci. USA **101**, 14766 (2004).
- ⁹⁵Y. Chebaro, A. J. Ballard, D. Chakraborty, and D. J. Wales, Sci. Rep. **5**, 10386 (2015).
- ⁹⁶K. Röder and D. J. Wales, J. Chem. Theory Comput. **13**, 1468 (2017).
- ⁹⁷D. J. Wales, M. A. Miller, and T. R. Walsh, Nature **394**, 758 (1998).
- ⁹⁸J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **110**, 6896 (1999).