

Tertiary Element Interaction in HIV-1 TAR

Konrad Krawczyk, Adelene Y. L. Sim, Bernhard Knapp, Charlotte M. Deane, Peter Minary†

AUTHOR ADDRESS KK, PM, Department of Computer Science, Oxford University, Parks Rd, OX1 3QD, Oxford, UK; BK, CMD, Department of Statistics, Oxford University, St Giles, OX1 3LB, Oxford, UK; AYL, Bioinformatics Institute (A*STAR), 30 Biopolis Street #07-01, Matrix, Singapore 138671

KEYWORDS. HIV, TAR, RNA, Molecular Dynamics, Molecular Simulation

ABSTRACT . HIV-1 replication requires binding to occur between Trans-activation Response Element (TAR) RNA and the TAT protein. This TAR-TAT binding depends on the conformation of TAR and therapeutic development has attempted to exploit this dynamic behavior. Here we simulate TAR dynamics in the context of mutations inhibiting TAR binding. We find that two tertiary elements, the apical loop and the bulge can interact directly and this interaction may be linked to the affinity of TAR for TAT.

1. Introduction

The Trans-activation Response RNA (TAR) is a viral transcript of HIV-1. It is essential for the virus to duplicate itself^{1,2}. TAR mediates the transactivation of the viral promoter by binding to the virus-transcribed TAT protein and human cyclin T1. Disrupting the binding between TAR and TAT – thus reducing the ability of the virus to replicate – is a viable anti-HIV therapeutic strategy.^{3–5}

Development of successful drugs against TAR requires a structural understanding of the conformational variability of TAR and the binding between the TAR and TAT molecules^{2,6–8} (see Fig. 1

for a structural description of the TAR molecule). Nuclear Magnetic Resonance (NMR) and Molecular Dynamics (MD) simulations have demonstrated that TAR in its unbound form adopts many conformational states beyond those observed when TAR is bound to ligands.⁹⁻¹² Furthermore NMR studies have shown that transient structural states of the TAR molecule exist, some of which are not favorable for TAT binding^{13,14} while others are.¹⁵ Recently microsecond-long MD simulations of TAR RNA in its bound and unbound forms have shed light on some functionally relevant conformational changes of the molecule.¹⁶

Conformational variability of TAR can be described by motions of its tertiary elements with respect to one another which might allow the molecule to efficiently probe various possible bound states.^{9,13,16} The loop and the bulge (Figure 1) tertiary elements are particularly important as many peptides and small molecules bind in the groove between them¹⁷ and both loop and bulge interact with TAT and cyclin.^{18,19} Thus, elucidating possible loop-bulge interactions not only builds our fundamental understanding of TAR function, but could also be important in TAR-targeted drug design.

Previously, direct interactions between the bulge and nearby basepairs were proposed.²⁰ However, NMR²¹ and MD^{9,11} studies suggest that any motional correlation between the loop and the bulge are long-range rather than direct. The nature of such long-range interaction between these elements remains unclear. In several of the NMR models of the unbound form of TAR, the loop and the bulge are within hydrogen bonding distance (See Supporting Figure S1). Hydrogen bonding between the loop and bulge were also transiently present in the most recent and longest MD simulation of TAR RNA¹⁶ (Supporting Figure S2). In this manuscript we investigate the structural nature of this interaction and show that it can be linked to TAR-TAT binding propensity.

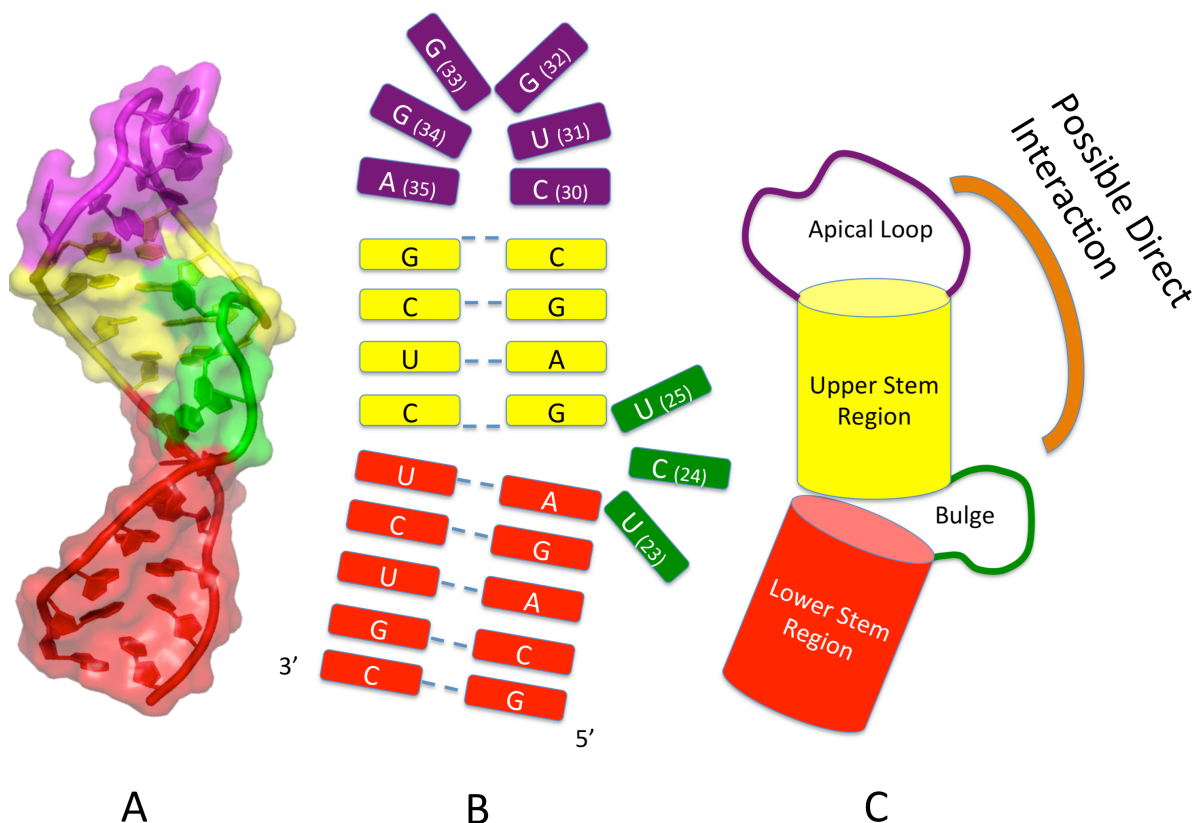


Figure 1. Topology of the TAR molecule. **A.** TAR from NMR structure (PDB code 1ANR). **B.** Diagram showing the secondary structure of TAR. **C.** The four different colours denote the Quasi-rigid Domains proposed by Musiani *et al.*.¹⁶ These domains, define our partitioning into regions of collective motion for Hierarchical Natural Moves.²²

2. Results and Discussion

2.1 Quantifying the tertiary element interactions via MOSAICS simulations

We investigated the possibility of direct loop-bulge interactions using Hierarchical Natural Move Monte Carlo (HNMMC)²² as implemented in MOSAICS.²³ We define the loop and the bulge to be interacting if hydrogen bond donor and acceptor atoms from loop and bulge are separated by less than 3.2 Å.

HNMMC method has been demonstrated to be able to efficiently sample the structural states of nucleic acids.²² Rather than using Cartesian degrees of freedom, as is often the case in MD, we employed

natural moves of nucleic acid residues with internal torsional degrees of freedom.²⁴ Furthermore, we introduce hierarchical natural moves,²² where residues may be grouped (*e.g.* based on their natural organization into helices) and assigned collective motions (degrees of freedom) on top of individual motions. This technique allows more efficient exploration of RNA tertiary structure where experimental data supporting such partitioning is available (see Materials and Methods).

Here we divide the molecule into four natural move regions: two helical stems, the loop and the bulge (Figure 1). This choice was inferred from the latest MD results and from a large-scale investigation of the dynamics of A-stem-loops.^{16,25} Following the study of Musiani *et al.*,¹⁶ we used the Amber parmbsc0 potential as it has been shown to be suited to simulating nucleic acids. Further details of the simulation protocols can be found in the Materials and Methods. Simulations were performed at 300K. We performed 10 independent HNMMC simulations for each of the 20 models of the unbound TAR NMR structure (Protein Data Bank identifier 1ANR).

Of the 20 sets of 10 simulations each starting from an NMR model of 1ANR, we observed direct loop-bulge interaction in seven cases (see Figure 2 and Supporting Table S1). In models 7 and 8 the interaction is only transient (~1% of the trajectory) and in models 2, 3 and 7 the loop and the bulge are within weak hydrogen bond distance to begin with. In three out of those seven cases, the initial separation of loop and the bulge was greater than 5 Å, (models 1, 6 and 11) and a loop-bulge interaction occurs for more than 20% of the trajectories.

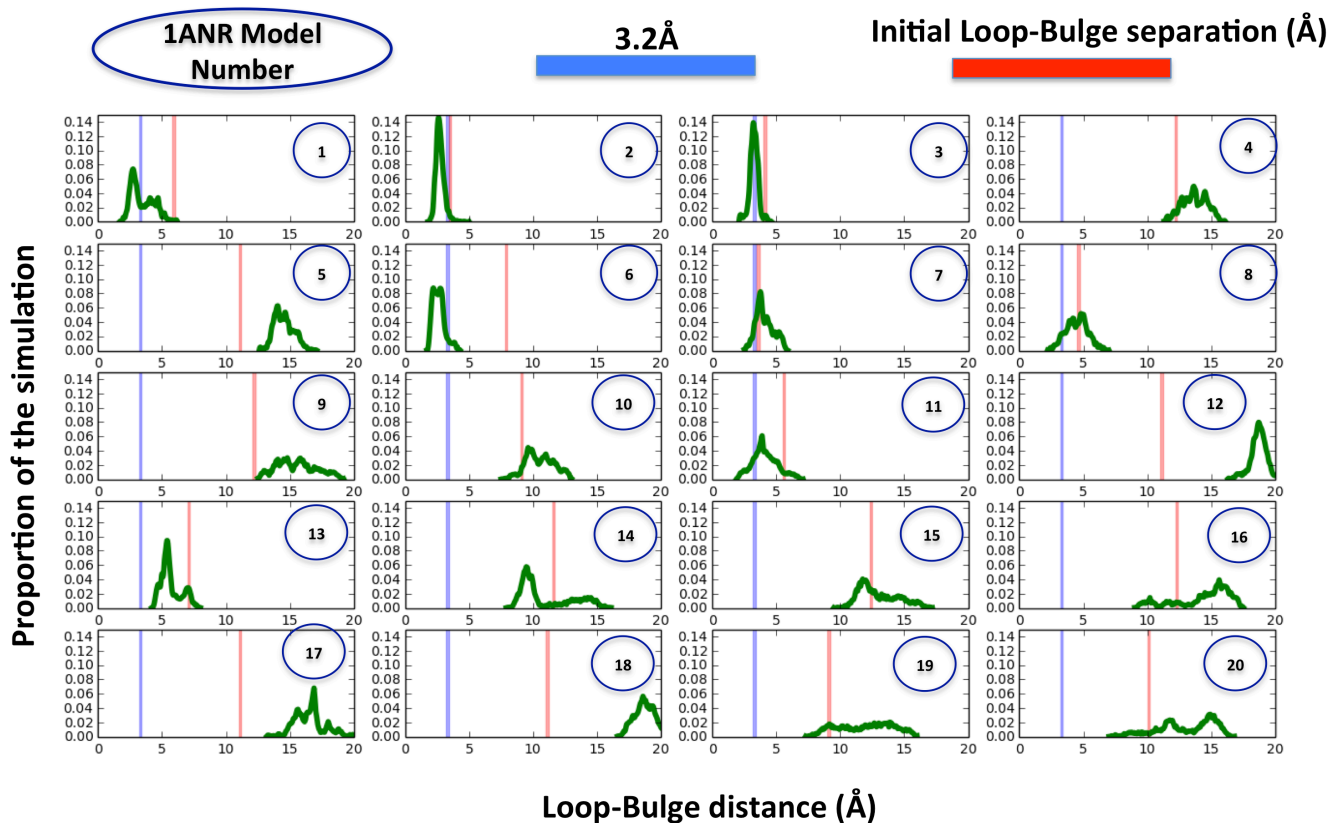


Figure 2. Loop-bulge distance distributions based on finite temperature (300K) conformational trajectories started from the twenty 1ANR NMR models. The blue line indicates the cutoff for a moderately strong hydrogen bond (3.2 Å) that we use as an indicator of loop-bulge interaction. The red line denotes the initial loop-bulge separation. We observe a loop-bulge interaction (<3.2 Å) in seven out of the 20 models. In three (1ANR model number 1, 6, 11) out of those seven models, the initial loop-bulge separation was greater than 5 Å and we observe the formation of non-transient loop-bulge interactions (for more than 20% of the trajectory, the loop and the bulge are less than 3.2 Å apart).

2.2 Validation of the Simulation results

In order to validate our MOSAICS setup using our initial trajectories of the 20 1ANR models, we followed the protocol of Musiani *et al.*,¹⁶ comparing our conformations with available Residual Dipolar Coupling (RDC) data²⁶ using PALES.²⁷ Our trajectories showed good correlation with the experimental data (Fig. 3), supporting the accuracy of our simulations. We have further checked the stability of the loop-bulge interaction by running MD simulations following the protocol of Musiani *et al.*¹⁶ from a randomly chosen set of frames from our simulations where loop and the bulge are less than 3.2 Å apart. These controls indicated that the loop-bulge interaction is transient in nature but the states we found are not inherently unstable, supporting the existence of a loop-bulge interaction (data not shown).

1ANR model number

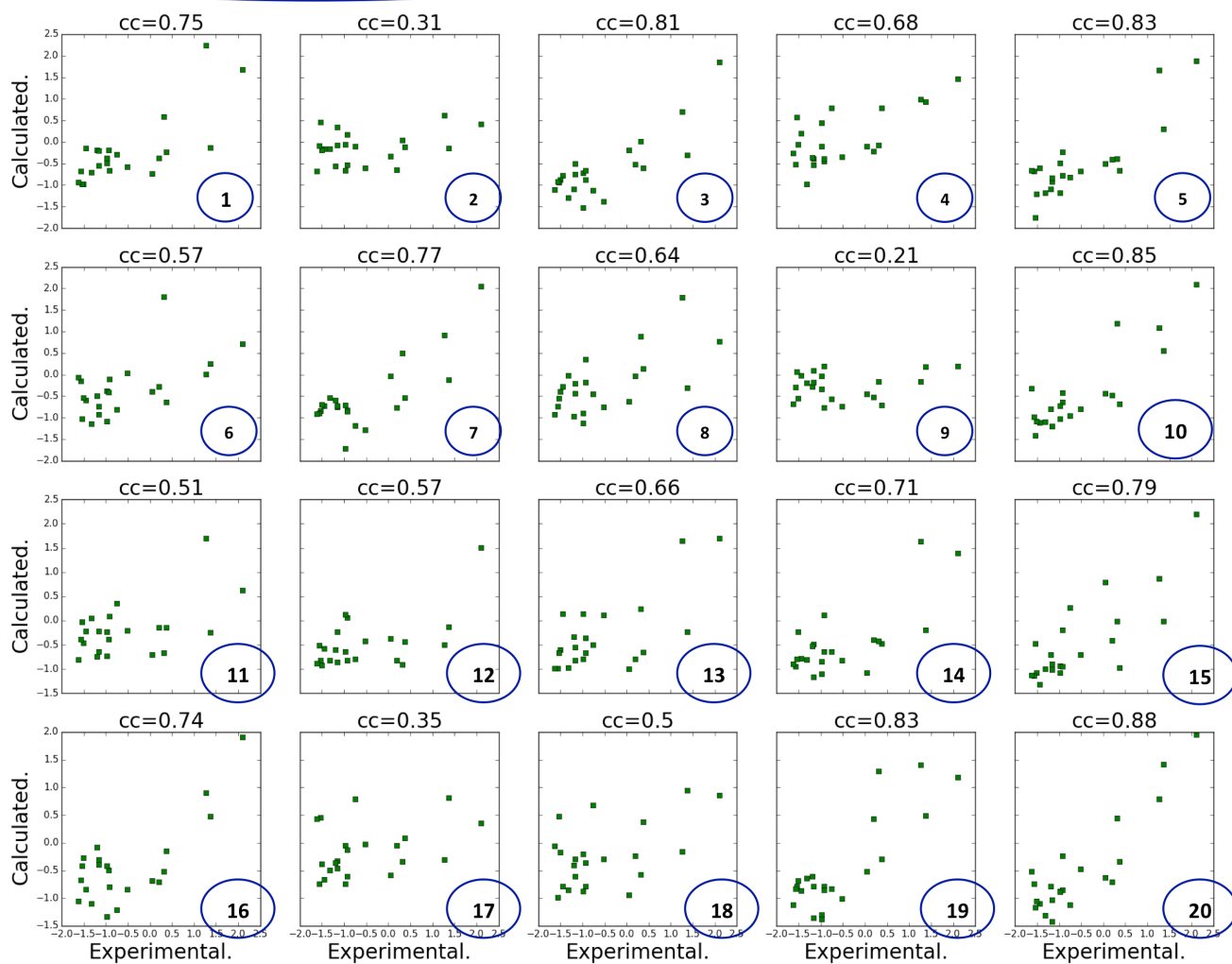


Figure 3. Residual Dipolar Coupling correlations for the 20 1ANR models. The trajectories of each independent repeat for each model were pooled together (concatenated trajectories). RDC analysis was performed following previous studies into the dynamics of TAR.¹⁶ The 1ANR model number is given in the blue circle. The correlation coefficient is given as ‘cc’ above the each graph.

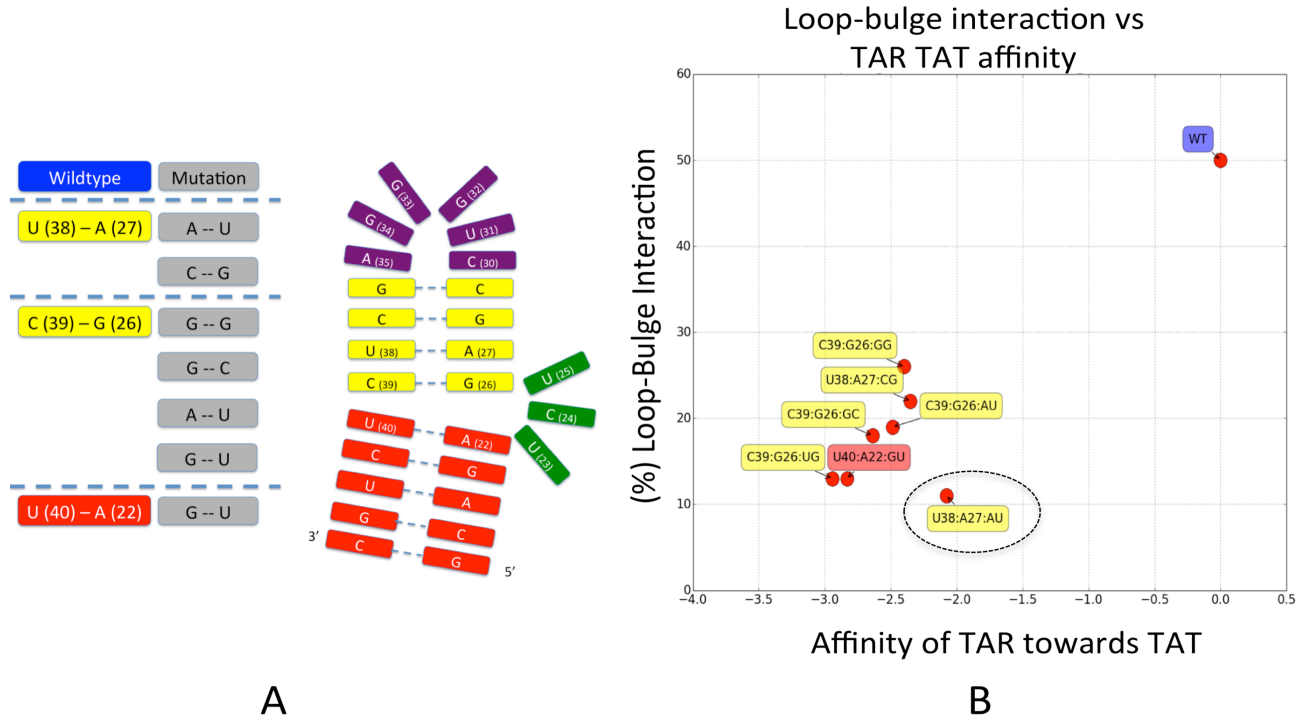


Figure 4. A: For our mutations study, we have picked 7 basepair mutations with a broad range of TAR-TAT affinities under the same experimental conditions.²⁸ The authors²⁸ studied UUU bulge variant rather than UCU which is the variant in 1ANR. However, this mutation does not have a drastic effect on TAR-TAT affinity.²⁸ **B:** Plot of proportion of trajectory where loop-bulge interaction is observed against TAR-TAT affinity for a given mutant (log-scaled) (correlation 0.71). The strongest mutant outlier (U38:A27:AU) is highlighted with the dotted ellipse. This outlier might not allow A27 to engage in triple formation which interacts with the TAT protein;^{29–31} U38:A27:AU may therefore be interacting with TAT via a different mechanism that is independent of loop-bulge interaction.

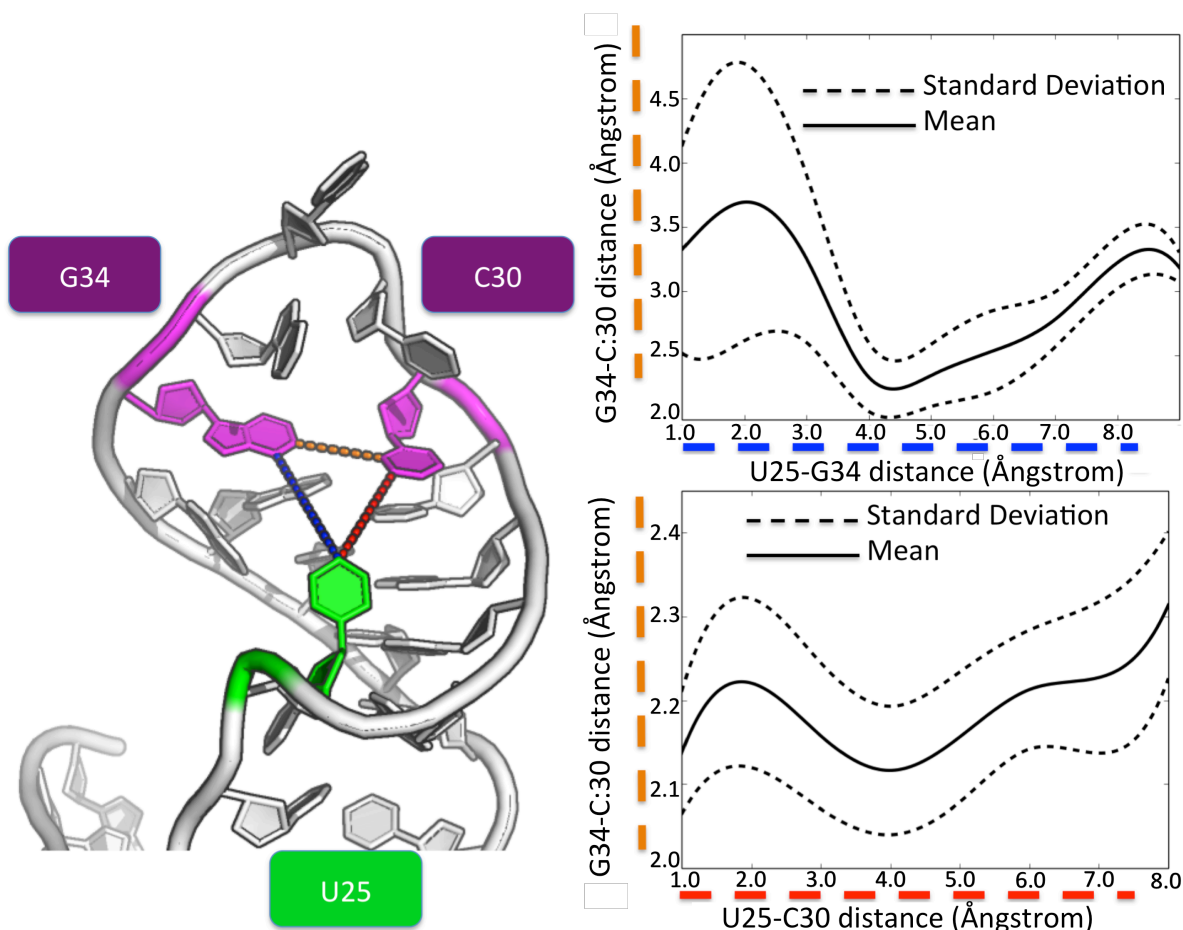


Figure 5. The C30:G34 apical loop basepair separation plotted as a function of the distance from residue U25 of bulge (wildtype). The C30:G34 distance is taken as the closest nucleobase atom between the two nucleotides. The distance from U25 to either G34 or C30 is taken as closest distance between hydrogen bond donor and acceptor. The apical basepair is colored in magenta and U25 in green. It appears that there is a loop-U25 distance that achieves a minimum basepair separation, with closer or farther separation, pulling the loop basepair (C30:G34) apart. At the same time, U25:G34 and U25:C30 distances are moderately negatively correlated (-0.68).

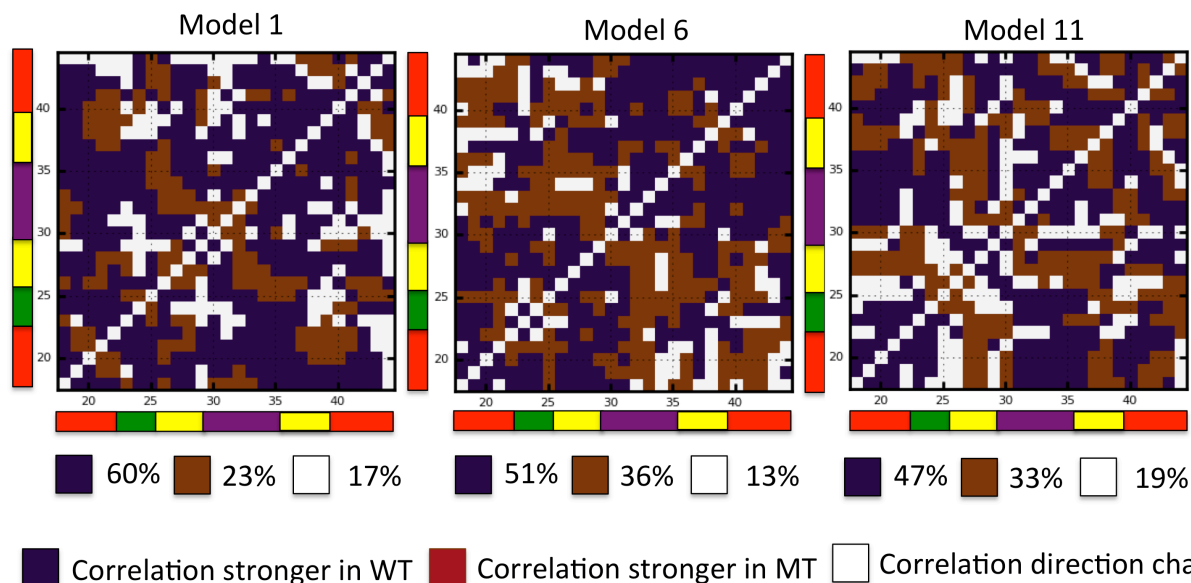


Figure 6: Differential correlated motions for mutation U40:A22:=GU. The blue patches correspond to positions where the sign of correlated motions (negative or positive) remained the same and the correlation is stronger in the wildtype. Brown patches indicate positions where the sign of correlated motions remained the same but the correlation is stronger in the mutant. The white patches indicate positions where the sign of the correlated motions changed. The color-codes next to residue numbers are the same as secondary structure annotation in Figure 1.

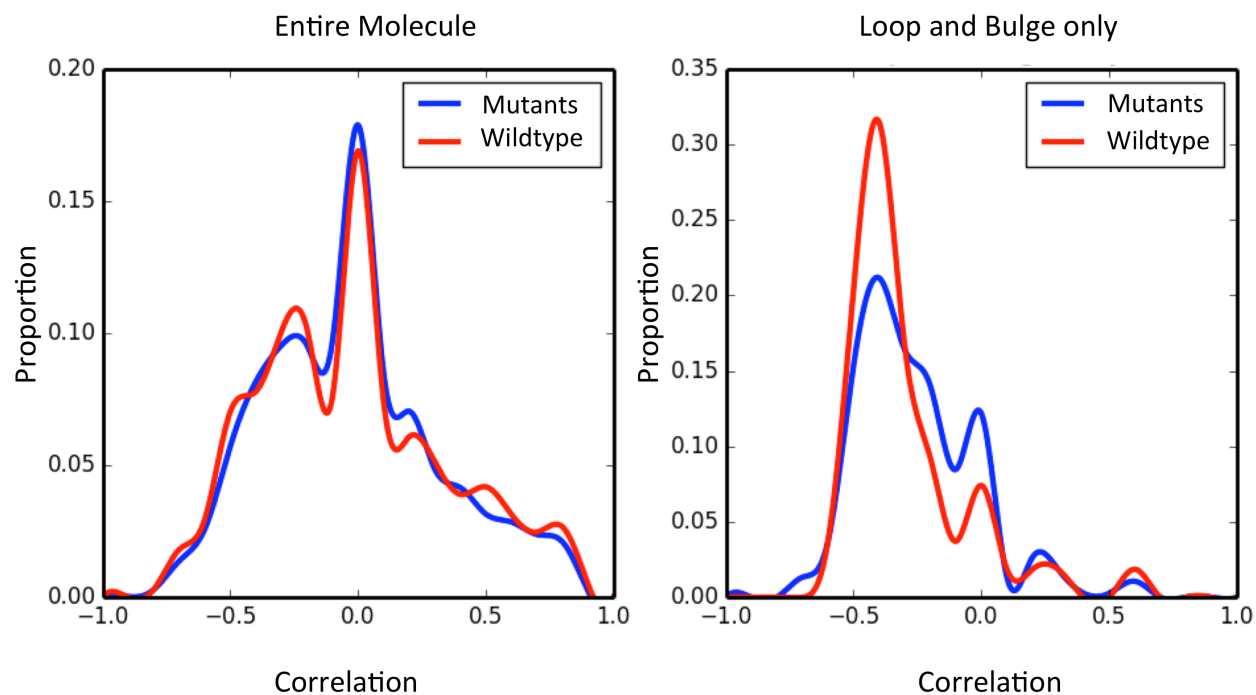


Figure 7: Correlated motions for wildtype and mutant TAR molecules were calculated using Ambertools (Amber14). **LEFT:** Histogram of the correlated motions for each nucleotide pair in the entire molecule for mutants and wildtype. The distributions are not statistically dissimilar according to a Kolmogorov-Smirnov procedure (p-value 0.09). **RIGHT:** Histogram of the correlated motions for pairs of nucleotides from the bulge and loop only. The distributions are dissimilar according to the Kolmogorov-Smirnov procedure (p-value < 0.01) and mutants have lower proportion of negatively correlated pairs (around -0.5) and higher proportion of weakly or not correlated pairs (around 0).

2.3 Tertiary element interaction in the context of TAR mutations indicates correlation to TAT binding.

In order to gain further insight into the mechanics of the loop-bulge interaction in TAR, we studied it in the context of mutations. We used the seven basepair mutations reported by Churcher et al. in 1993. These show a broad range of effects on TAR-TAT binding,²⁸ ranging in reducing affinity between 8 and 19-fold under the same experimental conditions (Figure 4A). The mutations were introduced to models 1, 6 and 11 as these had initial loop-bulge separation greater than 5 Å and the identified loop-bulge interaction in the wildtype simulations is not transient here. We performed MOSAICS simulations on the 21 configurations (7 mutations, 3 starting models). We still observed direct loop-bulge interactions in the mutants, following the same character as their parent wildtype models (Supporting Figure S3). Nevertheless the proportion of such interactions is smaller than in their wildtype equivalents. For each of the seven mutations, we calculated the average percentage of a trajectory where we observe a direct loop-bulge interaction. These percentages show moderate positive correlation (correlation coefficient of 0.71) with the affinity of TAR towards TAT²⁸ (See Figure 4B).

To account for the paucity of experimental mutational points we investigated whether the above relation between the frequency of loop bulge interaction and the affinity of TAT towards TAR can be predicted solely from the mutant group. Although we found that the resulting correlation is weak (0.23), further investigation on how this correlation changes by removal of individual mutation points revealed convincing evidence on the presence of a correlation between the propensity of loop-bulge interaction and the affinity of TAT towards TAR. We used jack-knife resampling to investigate the contribution of individual mutations (since we have a

small number of mutations to source from, some correlations might arise at random). We have calculated the correlation of TAT-affinity against loop-bulge interaction after systematically removing one of the seven mutations from our dataset. This procedure revealed that there exists one strong correlation (.92), which is unlikely to have arisen by chance (Supporting Table S2). This correlation arises when we remove mutation U38:A27:AU, which is an outlier with respect to the other mutants and thus is affecting the correlation when it is present (mutation in the dashed circle, Figure 4B, Supporting Table S2). Removing this outlier from Figure 4B, leads to a correlation coefficient (among the wild type and six remaining mutant points) of 0.98 (p-value < 0.01). We then further investigated whether the underlying mutation (U38:A27:AU) causes any unique characteristic change (in structural or chemico-physical properties) not present in the cases of other mutations. Residue A27 is known to be involved in TAT binding by forming a triple that interacts with the protein.^{29–31} Changing the character of A27 (purine to pyrimidine) would lead to three pyrimidines and prevent the formation of any canonical³² base triple (even the formation of any novel base triples³³ would require at least one purine among the three bases). Thus the mutation of A27 to U would prevent the TAT binding-induced base triple formation, thereby altering the TAT-TAR binding mode. As a consequence, the dependence of TAT binding on loop-bulge interaction is no longer valid for this particular mutation. In contrast, the other mutation to A27, that we used (U38:A27:CG) is not an outlier with respect to the loop-bulge TAT-affinity correlation. Here, the pyrimidine/purine character of the original residues is not changed, and likely the base triple that interacts with TAR can still be formed; i.e. TAT binds to this mutated TAR in a similar way to the wildtype TAR. These results suggest that the loop-

bulge interaction could be gating the formation of the triple that interacts directly with TAT, and should be investigated further in the future.

2.4 Possible relationships between the loop-bulge interaction and intra-loop basepair formation.

We have investigated the possible effects of loop-bulge interaction on the structure of TAR by looking at existing structural features.

Two of the most commonly observed loop-bulge interactions in models 1, 6 and 11 are U25:C30 and U25:G34. There is a transient intra-loop Watson-Crick basepair formed between C30 and G34 which is important for TAR-TAT recognition¹⁵ (Figure 5). We find that the distance of bulge U25 to either of the basepair residues appears to be correlated with the basepair distance (Figure 5). In addition, U25:G34 and U25:C30 distances are moderately negatively correlated ($cc=-0.68$). Also, when the U25-basepair distance distribution is different between wildtype and mutants, so is the basepair distance (Supporting Figure S4). Thus ability to maintain a correct loop-bulge distance appears to impact the formation of the C30:G34 basepair or vice versa.

Furthermore, the differences between wildtype and mutants might be the result of weaker correlated motions in mutants. We contrasted the overall correlated motions between wildtype and mutants. We note that on average, wildtype trajectories had a higher proportion of stronger correlated motions as compared to mutants (using the ptraj command in AmberTools14;³⁴ see Figures 6 & 7). This difference is particularly pronounced in the loop and bulge regions (Figure

7). From Figure 7 one may also infer that the mutations could affect collective motions involving distant residues

3. Discussion.

RNAs are highly malleable molecules whose dynamic states might be viable drug targets.^{12,35–39} Studying such conformational ensembles has been previously demonstrated to provide docking targets for small molecules.^{8,17} Nevertheless, charting the conformational space covered by these molecules is expensive experimentally and computationally.¹² Being able to design better drugs against RNA-targets relies in part on our ability to efficiently explore the structural space of the target to identify transient binding-pockets and interactions.⁴⁰ Among RNA targets, TAR is especially important, as development of ligands that disrupt its functions can pave a way for a viable anti-HIV therapy. Furthermore, different dynamic conformational states adopted by TAR subsequently expose a multitude of contact sites on this molecule (see Supporting Figure S5): these dynamic states open diverse pockets that may be druggable. Distinguishing between these conformational states, and their druggability is an important problem that needs to be addressed.

Here, building on prior experimental knowledge, we were able to perform simulations on a wide range of mutational TAR variants. Our simulations indicate that the direct loop-bulge interaction has a significant probability of occurrence. We further indicate that this phenomenon might be linked to the propensity of TAR against TAT via correlated motions and/or the simultaneous formation of apical intra-loop basepair and loop-bulge basepairs. The apical base-pair has been demonstrated previously to be important for TAR function¹⁵ and the existence of distant loop-bulge interaction has been previously proposed⁹.

Our large-scale simulations suggest that loop-bulge interaction as well as the apical loop Watson-Crick basepair could be intertwined, and therefore inhibiting the loop-bulge interaction could be a potential therapeutic strategy. Some ligands and TAT-mimetic peptides bind in the pocket between the loop and the bulge. Therefore, it would be beneficial to identify novel ligands capable of binding in the pocket between the apical loop and the bulge, thus preventing the interaction between the two. Alternatively, ligands might be designed to disrupt one of the three most common interactions that we observe: U25:G34, U25:C30 or U25:G32. The residue U25, which is central to these interactions, is very flexible, and is usually solvent exposed⁸. Because U25 can be mutated or even deleted, without affecting the function of TAR,²⁸ U25 was previously regarded as uninteresting for further investigation.

In the light of our results, we propose that the role of U25, specifically in the context of loop-bulge interactions be investigated further. Other possible effects such as the role of solvent accessibility of A27 or the formation of base triple call for a thorough computational exploration of the conformational space of TAR. On the other hand, prior experimental basis exists for studying correlated motions¹⁰ as well as the role of the apical loop basepair¹⁵ on TAT binding which might readily shed light on the existence and role of the loop-bulge exploration as highlighted in this manuscript.

4. Materials and Methods.

4.1 Utilized Data.

We used the 1ANR⁴¹ NMR structure of TAR as the starting point of our MOSAICS simulations, motivated by the previous study into TAR dynamics.¹⁶ There are 20 NMR models in 1ANR structure and we have used each of these as a starting point for our simulations. Three of the 20 models (1, 6 and 11) have been modified to study the structural effects of mutations on the structure of TAR RNA. Mutations of these models given in Figure 4A were created using the *mutate_bases* utility in the X3DNA suite.⁴²

In order to prepare the structure for MOSAICS, missing atoms were modeled using Swiss PDB Viewer.⁴³ Missing hydrogen atoms were added using the *h_add* command in Pymol.⁴⁴ The added hydrogen atom names were converted to follow IUPAC standards using scripts available from the MOSAICS website.²³ All the initial models used in our simulations are available from the MOSAICS website,²³ following <Examples of Application> and then <Tertiary element interaction in HIV-1 TAR>.

4.2 MOSAICS Framework.

Conventional molecular simulation methodologies such as MD face the computational challenge of multiple degrees of freedom and the complexity of the energy function. MOSAICS is a molecular simulation package which addresses this problem by using reduced degrees of freedom and employing prior structural information to make sampling more efficient.

MOSAICS encompasses a range of structural molecular simulations methodologies, most notably Hierarchical Natural Move Monte Carlo (see section 4.3) or HNMMC. The use of HNMMC has expanded the boundaries of computational structural biology with key applications such as aiding the design of RNA nanotechnology,²² increasing the resolution of macromolecular cryo-electron microscopy,⁴⁵ atomistic prediction of primary chromatin structure and the epigenetic effects,⁴⁶ modeling diabodies that tune extracellular receptor signaling to counteract intracellular oncogenic mutations,⁴⁷ and exploring essential molecular processes that are part of the adaptive immune response.⁴⁸

MOSAICS allows for simulations of nucleic acids²² as well as proteins.⁴⁸ Users can perform all-atom or coarse-grained simulations choosing from a variety of popular sampling methods such as Parallel Tempering, Normal Modes and others. The software can use any of the well-established energy functions such as CHARMM⁴⁹ or Amber parmbsc0.⁵⁰ Users can also specify different degrees of freedom. For instance using torsional rather than Cartesian degrees of freedom reduces the computational search space for sampling. If prior information on the collective

dynamics of a molecule is available, it is possible to further improve sampling efficiency by using computationally or experimentally inferred collective degrees of freedom used by methods such as HNMMC.

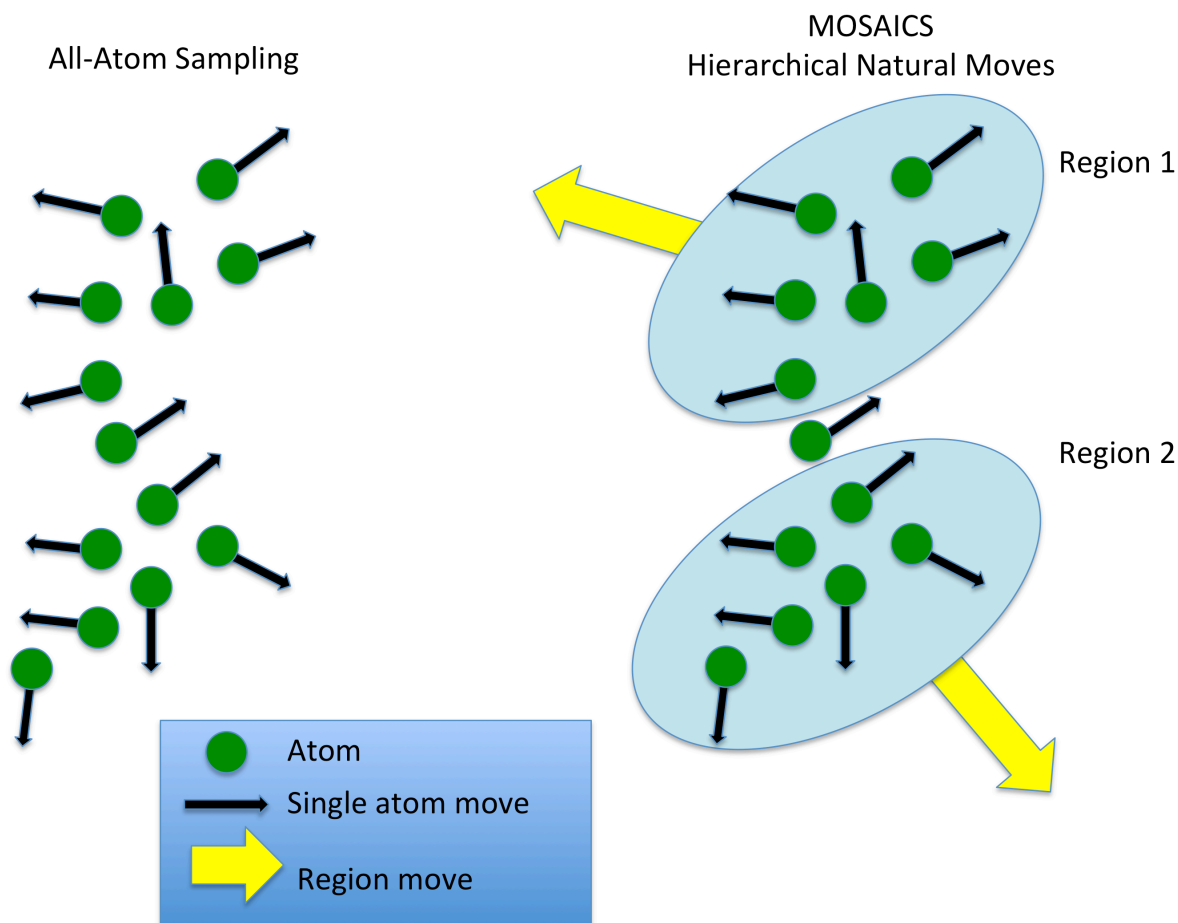


Figure 8. A simplified diagram illustrating Hierarchical Natural Move Monte Carlo, available in MOSAICS. Atoms are grouped into the natural move regions, motivated by experimental or computational studies. The groups are moved together, allowing for more efficient sampling of large moves. Natural moves may lead to chain breaks, which are solved by an efficient chain closure algorithm²⁴.

4.3 Hierarchical Natural Move Monte Carlo

In this study we employ the HNMMC²² methodology as implemented in MOSAICS²³. This technique aims to increase the efficiency of sampling by grouping the atoms into regions of collective motion, also called “natural move regions”. Such natural move regions can correspond to rigid or semi-rigid elements such as alpha helices, basepaired DNA strands or entire protein domains (Figure 8). Atoms grouped into such natural move regions are moved together but they can also retain a degree of movement of their own. Such large movements can break the bonds holding the chain together, which is amended by a loop closure algorithm implemented in MOSAICS²⁴. This approach increases the efficiency of sampling since it allows for big collective moves, which would otherwise be very unlikely using any of the conventional simulation techniques. Nevertheless, elements can be grouped into natural move regions only if suitable data supporting such partitioning is available, otherwise artificial moves may be observed.

4.4 Defining Natural Moves for TAR RNA

Since there exists a lot of data on the dynamic nature of the TAR RNA, we have employed it to define the natural move regions for our simulations. Our partitioning into natural move regions was motivated by the study of Musiani *et al.*¹⁶. The authors performed long (one microsecond) MD study simulations of TAR, which revealed the Quasi-rigid Domains of the molecule. We have defined two basepaired regions (lower and upper stem) and two non-basepaired regions (loop and the bulge), which are given in Figure 1. The relevant region file

used in the MOSAICS simulations is available from the MOSAICS website²³ following <Examples of Application> and then <Tertiary element interaction in HIV-1 TAR>.

4.5 MOSAICS Setup

MOSAICS version 3.9 was compiled for the purposes of this study. Simulations were performed using HNMMC sampling at a temperature of 300K. Following the protocol of Musiani *et al.* we have employed the Amber parmbsc0 energy function for our simulations⁵⁰. We used implicit counterions to neutralize the molecule and a distance-dependent dielectric as implicit solvent, a protocol that has been used in previous studies⁴⁶. Each run was set to four million Monte Carlo steps after initial test runs indicated that loop-bulge interaction can be observed after circa one million steps. Each run was repeated 20 times with different random seeds. The simulations were performed on the Supercomputing Cluster available via the Hartree Centre of the Science and Technology Facilities Council.

Instructions on how to reproduce the simulations in this paper are available from the MOSAICS website²³ following <Examples of Application> and then <Tertiary element interaction in HIV-1 TAR>. An example is provided which should act as a suitable starting point to perform an arbitrary run by changing the PDB file used as the initial conformation of the simulation (e.g. one of the 1ANR models).

ASSOCIATED CONTENT

Supporting Information. Supporting information are available via the journal website as well as from the MOSAICS website MOSAICS website²³ following <Examples of Application> and then <Tertiary element interaction in HIV-1 TAR>.

AUTHOR INFORMATION

Corresponding Author

†To whom the correspondence should be addressed: peter.minary@cs.ox.ac.uk

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by the 2020 Science Programme (EPSRC Cross Discipline Interface Programme, EP/I017909/1) to C.M.D and the Agency for Science, Technology and Research (A*STAR) Joint Council Office (JCO) Career Development Award (project number 15302FG145) to A.Y.L.S.

ACKNOWLEDGMENT

We would like to thank Dr Musiani for kindly sharing his work with us. We acknowledge computational support from the 1st Intel Xeon Phi Access Programme in the Hartree Centre of the Science and Technology Facilities Council.

REFERENCES

- (1) Weeks, K. M.; Ampe, C.; Schultz, S. C.; Steitz, T. A.; Crothers, D. M. Fragments of the HIV-1 Tat Protein Specifically Bind TAR RNA. *Science* **1990**, *249*, 1281–1285.
- (2) Dingwall, C.; Ernberg, I.; Gait, M. J.; Green, S. M.; Heaphy, S.; Karn, J.; Lowe, A. D.; Singh, M.; Skinner, M. A. HIV-1 Tat Protein Stimulates Transcription by Binding to a U-Rich Bulge in the Stem of the TAR RNA Structure. *EMBO J.* **1990**, *9*, 4145–4153.
- (3) Esteller, M. Non-Coding RNAs in Human Disease. *Nature Reviews Genetics*. 2011, pp 861–874.
- (4) Davidson, A.; Begley, D. W.; Lau, C.; Varani, G. A Small-Molecule Probe Induces a Conformation in HIV TAR RNA Capable of Binding Drug-like Fragments. *J. Mol. Biol.* **2011**, *410*, 984–996.
- (5) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. Discovery of Selective, Small-Molecule Inhibitors of RNA Complexes - I. The Tat protein/TAR RNA Complexes Required for HIV-1 Transcription. In *Bioorganic and Medicinal Chemistry*; 1997; Vol. 5, pp 1173–1184.
- (6) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. Role of RNA Structure in Arginine Recognition of TAR RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3680–3684.
- (7) Aboul-ela, F.; Karn, J.; Varani, G. The Structure of the Human Immunodeficiency Virus Type-1 TAR RNA Reveals Principles of RNA Recognition by Tat Protein. *J Mol Biol* **1995**, *253*, 313–332.
- (8) Bardaro, M. F.; Shajani, Z.; Patora-Komisarska, K.; Robinson, J. A.; Varani, G. How Binding of Small Molecule and Peptide Ligands to HIV-1 TAR Alters the RNA Motional Landscape. *Nucleic Acids Res.* **2009**, *37*, 1529–1540.
- (9) Fulle, S.; Christ, N. A.; Kestner, E.; Gohlke, H. HIV-1 TAR RNA Spontaneously Undergoes Relevant Apo-to-Holo Conformational Transitions in Molecular Dynamics and Constrained Geometrical Simulations. *J. Chem. Inf. Model.* **2010**, *50*, 1489–1501.
- (10) Al-Hashimi, H. M.; Gosser, Y.; Gorin, A.; Hu, W.; Majumdar, A.; Patel, D. J. Concerted Motions in HIV-1 TAR RNA May Allow Access to Bound State Conformations: RNA Dynamics from NMR Residual Dipolar Couplings. *J. Mol. Biol.* **2002**, *315*, 95–102.

- (11) Musselman, C.; Al-Hashimi, H. M.; Andricioaei, I. iRED Analysis of TAR RNA Reveals Motional Coupling, Long-Range Correlations, and a Dynamical Hinge. *Biophys. J.* **2007**, *93*, 411–422.
- (12) Stelzer, A. C.; Frank, A. T.; Bajor, M. H.; Andricioaei, I.; Al, H. M. Constructing Atomic-Resolution RNA Structural Ensembles Using MD and Motionally Decoupled NMR RDCs. *Methods* **2009**, *49*, 167–173.
- (13) Dethoff, E. A.; Petzold, K.; Chugh, J.; Casiano-Negroni, A.; Al-Hashimi, H. M. Visualizing Transient Low-Populated Structures of RNA. *Nature*. 2012.
- (14) Lee, J.; Dethoff, E. a; Al-Hashimi, H. M. Invisible RNA State Dynamically Couples Distant Motifs. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 9485–9490.
- (15) Kulinski, T.; Olejniczak, M.; Huthoff, H.; Bielecki, L.; Pachulska-Wieczorek, K.; Das, A. T.; Berkhout, B.; Adamiak, R. W. The Apical Loop of the HIV-1 TAR RNA Hairpin Is Stabilized by a Cross-Loop Base Pair. *J. Biol. Chem.* **2003**, *278*, 38892–38901.
- (16) Musiani F1, Rossetti G, Capece L, Gerger TM, Micheletti C, Varani G, C. P. Molecular Dynamics Simulations Identify Time Scale of Conformational Changes Responsible for Conformational Selection in Molecular Recognition of HIV-1 Transactivation Responsive RNA. *J Am Chem Soc.* **2014**, *136*, 15631–15637.
- (17) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nat. Chem. Biol.* **2011**, *7*, 553–559.
- (18) Richter, S.; Ping, Y.-H.; Rana, T. M. TAR RNA Loop: A Scaffold for the Assembly of a Regulatory Switch in HIV Replication. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 7928–7933.
- (19) Braddock, M.; Powell, R.; Sutton, J.; Kingsman, A. J.; Kingsman, S. M. Orientation-Specific Cis Complementation by Bulge- and Loop-Mutated Human Immunodeficiency Virus Type 1 TAR RNAs. *J. Virol.* **1994**, *68*, 8396–8400.
- (20) Huthoff, H.; Girard, F.; Wijmenga, S. S.; Berkhout, B. Evidence for a Base Triple in the Free HIV-1 TAR RNA. *RNA* **2004**, *10*, 412–423.
- (21) Dethoff, E. A.; Hansen, A. L.; Musselman, C.; Watt, E. D.; Andricioaei, I.; Al-Hashimi, H. M. Characterizing Complex Dynamics in the Transactivation Response Element Apical Loop and Motional Correlations with the Bulge by NMR, Molecular Dynamics, and Mutagenesis. *Biophys. J.* **2008**, *95*, 3906–3915.
- (22) Sim, A. Y. L.; Levitt, M.; Minary, P. Modeling and Design by Hierarchical Natural Moves. *Proceedings of the National Academy of Sciences*. 2012, pp 2890–2895.
- (23) Peter Minary. MOSAICS v 3.9 <http://www.cs.ox.ac.uk/mosaics/>.
- (24) Minary, P.; Levitt, M. Conformational Optimization with Natural Degrees of Freedom: A Novel Stochastic Chain Closure Algorithm. *J. Comput. Biol.* **2010**, *17*, 993–1010.

- (25) Bajor, M. H.; Sun, X.; Al-Hashimi, H. M. Topology Links RNA Secondary Structure with Global Conformation, Dynamics, and Adaptation. *Science* **2010**, *327*, 202–206.
- (26) Zhang, Q.; Throolin, R.; Pitt, S. W.; Serganov, A.; Al-Hashimi, H. M. Probing Motions between Equivalent RNA Domains Using Magnetic Field Induced Residual Dipolar Couplings: Accounting for Correlations between Motions and Alignment. *J. Am. Chem. Soc.* **2003**, *125*, 10530–10531.
- (27) Zweckstetter, M. NMR: Prediction of Molecular Alignment from Structure Using the PALES Software. *Nat. Protoc.* **2008**, *3*, 679–690.
- (28) Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, J. G.; Gait, M. J.; Karn, J. High Affinity Binding of TAR RNA by the Human Immunodeficiency Virus Type-1 Tat Protein Requires Base-Pairs in the RNA Stem and Amino Acid Residues Flanking the Basic Region. *J. Mol. Biol.* **1993**, *230*, 90–110.
- (29) Davidson, A.; Patora-Komisarska, K.; Robinson, J. A.; Varani, G. Essential Structural Requirements for Specific Recognition of HIV TAR RNA by Peptide Mimetics of Tat Protein. *Nucleic Acids Res.* **2011**, *39*, 248–256.
- (30) Olsen, G. L.; Edwards, T. E.; Deka, P.; Varani, G.; Sigurdsson, S. T.; Drobny, G. P. Monitoring Tat Peptide Binding to TAR RNA by Solid-State ^{31}P - ^{19}F REDOR NMR. *Nucleic Acids Res.* **2005**, *33*, 3447–3454.
- (31) Davidson, A.; Leeper, T. C.; Athanassiou, Z.; Patora-Komisarska, K.; Karn, J.; Robinson, J. A.; Varani, G. Simultaneous Recognition of HIV-1 TAR RNA Bulge and Loop Sequences by Cyclic Peptide Mimics of Tat Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11931–11936.
- (32) Vasquez, K. M.; Glazer, P. M. Triplex-Forming Oligonucleotides: Principles and Applications. *Q. Rev. Biophys.* **2002**, *35*, 89–107.
- (33) Firdaus-Raih, M.; Harrison, A.-M.; Willett, P.; Artymiuk, P. J. Novel Base Triples in RNA Structures Revealed by Graph Theoretical Searching Methods. *BMC Bioinformatics* **2011**, *12 Suppl 1* (Suppl 13), S2.
- (34) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198–210.
- (35) Cooper, T. A.; Wan, L.; Dreyfuss, G. RNA and Disease. *Cell.* 2009, pp 777–793.
- (36) Zhang, Q.; Sun, X.; Watt, E. D.; Al-Hashimi, H. M. Resolving the Motional Modes That Code for RNA Adaptation. *Science* (80-.). **2006**, *311*, 653–656.
- (37) Tor, Y. Targeting RNA with Small Molecules. *ChemBioChem.* 2003, pp 998–1007.
- (38) Lu, J.; Nguyen, L.; Zhao, L.; Xia, T.; Qi, X. A Cyclic Mimic of HIV Tat Differentiates Similar TAR RNAs on the Basis of Distinct Dynamic Behaviors. *Biochemistry* **2015**, *54*, 3687–3693.

- (39) Lu, J.; Wong, V.; Zhang, Y.; Tran, T.; Zhao, L.; Xia, A.; Xia, T.; Qi, X. Distinct Conformational Transition Patterns of Noncoding 7SK snRNA and HIV TAR RNAs upon Tat Binding. *Biochemistry* **2014**, *53*, 675–681.
- (40) Fulle, S.; Gohlke, H. Molecular Recognition of RNA: Challenges for Modelling Interactions and Plasticity. *J. Mol. Recognit.* **2010**, *23*, 220–231.
- (41) Aboul-ela, F.; Karn, J.; Varani, G. Structure of HIV-1 TAR RNA in the Absence of Ligands Reveals a Novel Conformation of the Trinucleotide Bulge. *Nucleic Acids Res.* **1996**, *24*, 3974–3981.
- (42) Lu, X. J.; Olson, W. K. 3DNA: A Software Package for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic Acid Structures. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
- (43) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18*, 2714–2723.
- (44) DeLano, W. L. The PyMOL Molecular Graphics System. *Schrödinger LLC wwwpymol.org* **2002**, Version 1., <http://www.pymol.org>.
- (45) Zhang, J.; Minary, P.; Levitt, M. Multiscale Natural Moves Refine Macromolecules Using Single-Particle Electron Microscopy Projection Images. *Proceedings of the National Academy of Sciences.* **2012**, pp 9845–9850.
- (46) Minary, P.; Levitt, M. Training-Free Atomistic Prediction of Nucleosome Occupancy. *Proc. Natl. Acad. Sci.* **2014**, 201404475.
- (47) Moraga, I.; Wernig, G.; Wilmes, S.; Gryshkova, V.; Richter, C. P.; Hong, W. J.; Sinha, R.; Guo, F.; Fabionar, H.; Wehrman, T. S.; Krutzik, P.; Demharther, S.; Plo, I.; Weissman, I. L.; Minary, P.; Majeti, R.; Constantinescu, S. N.; Piehler, J.; Garcia, K. C. Tuning Cytokine Receptor Signaling by Re-Orienting Dimer Geometry with Surrogate Ligands. *Cell* **2015**, *160*, 1196–1208.
- (48) Knapp, B.; Demharther, S.; Deane, C. M.; Minary, P. Exploring peptide/MHC Detachment Processes Using Hierarchical Natural Move Monte Carlo. *Bioinformatics* **2016**, *32*, 181–186.
- (49) MacKerel Jr., A. D.; Brooks III, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and Its Parameterization with an Overview of the Program. In *The Encyclopedia of Computational Chemistry*; 1998; Vol. 1, pp 271–277.
- (50) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/gamma Conformers. *Biophys. J.* **2007**, *92*, 3817–3829.

Insert Table of Contents Graphic and Synopsis Here

TOC Graphic:

