Stabilizing non-native excited states as a therapeutic strategy for targeting RNA

Authors

Laura R. Ganser[1+], Chia-Chieh Chu[1+], Hal P. Bogerd[2], Megan L. Kelly[1], Bryan R. Cullen[2*], Hashim M. Al-Hashimi[1*]

[1]Department of Biochemistry, Duke University Medical Center, Durham, NC, USA

[2]Department of Molecular Genetics and Microbiology, Center for Virology, Duke University Medical Center, Durham, NC, USA

*Correspondence to hashim.al.hashimi@duke.edu and bryan.cullen@duke.edu

[+] Equal Contributions

Abstract

Regulatory RNAs transiently morph into non-native high-energy ‘excited’ conformational states that remodel the native ground state secondary structure needed for function. For two HIV-1 regulatory RNAs, we show that stabilizing excited states using point substitution mutations leads to potent conformation-dependent inhibition of RNA cellular activity to a degree that correlates with the extent of excited state stabilization. Stabilizing non-native excited states potentially provides a general therapeutic strategy for targeting RNA.

Main Text

Regulatory RNAs are increasingly recognized as a new class of drug targets for treating infectious diseases, cancers, and genetic disorders[1]. RNAs are highly flexible macromolecules and their regulatory functions often require specific changes in their three-dimensional (3D) structure[2]. Such structural dynamics can optimize interactions with proteins, provide a basis for RNA-based molecular switches, or allow one RNA to satisfy the distinct structural requirements of multi-step catalytic cycles[2]. Given the importance of structural dynamics to RNA function, it should not come as a surprise that certain aminoglycoside antibiotics act by impairing dynamics in ribosomal RNAs that is important for codon recognition[3]. Because this is one of few proven strategies for targeting RNA therapeutically, there is considerable interest in pursuing approaches for inhibiting RNA activity by interfering with, or exploiting, RNA structural dynamics[4,5].

Recently, regulatory RNAs have been shown to transiently morph into non-native secondary structures, termed 'excited states' (ESs), by reshuffling base pairs (bps) in and around non-canonical motifs such as bulges and apical loops[6–11]. Despite their low abundance, ESs have been shown in some instances to play functional roles as intermediates during folding or regulation[6,9,10]. Among the many conformations sampled by a given RNA, ESs are unique in that they represent the most stable conformations that have lost key characteristics of the native ground state (GS) secondary structure typically required for normal function. They therefore represent an interesting class of conformational states that, if stabilized with a therapeutic agent, could lead to potent inhibition of RNA cellular activity. For example, this could provide a means by which to inhibit RNA-protein interactions by disrupting the RNA structural elements recognized by proteins. However, it has yet to be shown for any RNA that its non-native excited-state does not support cellular activity as is expected for functions that critically depend on RNA secondary structure (Fig 1a). It also remains unclear whether these non-native conformations form readily within the cellular environment as observed *in vitro* (Fig 1a). The cellular environment could energetically bias the ensemble toward the native GS making it

more difficult to stabilize the ES through interactions with a therapeutic agent *in vivo* compared to *in vitro*.

To address the stability and biological activity of ESs *in vivo*, we developed a stabilize-and-rescue approach that uses point substitution mutations to invert the equilibrium so that the ES becomes the dominant (>50% population) conformation and secondary rescue mutations to restore the GS conformation and control for sequence effects (Fig 1a). The mutations are chosen to minimally interfere with structural features of the GS known to be important for function. The degree to which a given mutation stabilizes the ES relative to the GS is measured *in vitro* using Nuclear Magnetic Resonance (NMR) spectroscopy. We then examine the degree to which increasing the population of the ES at the expense of the GS leads to a corresponding decrease in cellular activity (Fig 1a). If an ES forms in cells with abundances comparable to those measured *in vitro*, and if the ES conformation does not support cellular activity, mutations stabilizing the ES should inhibit cellular activity to a degree commensurate with the degree to which they decrease the GS population *in vitro* (Fig 1b). This in turn would imply that ESs are viable cellular drug targets. If, on the other hand, the cellular environment preferentially biases the ensemble away from the ES in favor of the GS, and/or if the ES supports cellular activity, the impact of the mutations on activity would be attenuated in cells (Fig 1b). This would in turn imply that the ES is not a viable drug target.

We applied the above stabilize-and-rescue approach (Fig 1) to two regulatory RNAs in the HIV-1 virus, trans-activation response element (TAR) (Fig 2a) and Rev response element (RRE) (Fig 2b). The structure of these highly conserved RNAs has previously been shown to be critical for viral functions that involve binding to viral and host proteins[12–20]. In both cases, the ES remodels key structural features in the native GS required for protein recognition[6–8].

HIV-1 TAR is located at the 5' end of the viral genome and promotes transcription elongation by binding to the viral protein Tat and human super elongation complex (SEC). This interaction is required to release RNA polymerase II from promoter-proximal pausing, allowing for efficient transcription of the viral genome[21]. The TAR GS structure has been shown to be critical for its

biological activity and Tat and SEC form extensive interactions with the TAR bulge, upper stem, and apical loop[22,23]. Prior studies showed that TAR adopts two non-native ESs, termed 'ES1' and 'ES2' (Fig 2a)[6,7]. ES1 involves conformational changes within the apical loop, which zip up the hairpin by sequestering nucleotides into two non-canonical bps (Fig 2a)[6]. On the other hand, ES2 remodels the entire secondary structure from the bulge motif to the apical loop, creating a 1-nucleotide bulge, a dinucleotide apical loop, and four non-canonical bps (Fig 2a)[7]. Because ES2 greatly remodels the TAR structure, it is highly unlikely to support transcriptional trans-activation. Therefore, mutations stabilizing ES2 are expected to significantly inhibit viral transcription unless the cellular environment substantially biases the equilibrium to favor the GS.

We designed a series of mutations that stabilize ES2 and corresponding secondary mutations that restore the GS conformation to control for effects arising purely due to changes in sequence (Fig 2c, Supplementary note 1, and Supplementary Figs 1-3). The degree to which a given mutation biases the ensemble toward the GS or ES conformation was measured using NMR spectroscopy (see Methods, Fig 2c, Supplementary Fig 2). The mutants were then tested in a Tat-dependent HIV-1 LTR-driven luciferase reporter assay (Fig 3a-b).

Based on NMR relaxation dispersion experiments[24], TAR-G28U stabilizes ES2 to a population of ~99.8% (Fig 2c, Supplementary Fig 4, Supplementary Tables 1 and 2). The experiments indicate that TAR-G28U back-exchanges with a GS-like conformation that has an equilibrium population of ~0.2% (Supplementary Fig 4 and Supplementary Tables 1 and 2). NMR analysis reveals that the TAR-G28U/C37A rescue mutation restores a GS-like conformation, although a small population for ES2 is likely given line broadening in NMR spectra (Fig 2c, Supplementary Fig 2a). These results highlight how mutations do not induce all-or-nothing behavior (*i.e.* GS or ES) but rather redistribute the ensemble along a continuous scale (Fig 1).

If TAR-G28U forms the ES2 conformation within cells with an abundance comparable to that measured *in vitro*, and if the TAR ES2 conformation is indeed inactive, we would expect the mutant to show at least a 1000-fold reduction in activity relative to WT (an even greater level of inhibition is in

principle possible if the exchange kinetics between the GS and ES are slow relative to protein binding). As positive controls, we first tested double mutants TAR-G36C/C29G and TAR-G28C/C37G, which adopt a predominantly GS-like conformation based on NMR (Fig 2c, Supplementary Fig 2c). As expected, the mutants did not have statistically different trans-activation activity relative to WT (two-way ANOVA, $P > 0.05$) (Fig 3b). In contrast, the trans-activation activity of the ES2-stabilizing TAR-G28U mutant fell below detectable levels with indistinguishable activity from the -Tat negative control (two-sided unpaired Student's t-test, $P > 0.05$), corresponding to at least 100-fold decrease in activity (Fig 3b). The activity increased significantly (two-way ANOVA, $P < 0.05$) and was restored to WT TAR levels (two-way ANOVA, $P > 0.05$) upon introduction of a second rescue mutation (TAR-G28U/C37A) (Fig 3b). Although not statistically significant, the lower mean activity observed for TAR-G28U/C37A relative to WT could arise from incomplete restoration of the GS and/or because the sequence of the G28-C37 base pair is important for activity.

To test the robustness of these results, experiments were repeated for three other sets of mutants targeting different TAR residues (TAR-A27C, TAR-G36U, and TAR-U38A) and the results were similar to those obtained for TAR-G28U. All three ES2-stabilizing mutants drastically diminish TAR activity, while all but one GS-rescue mutation at least partially restores activity (two-way ANOVA, $P < 0.05$) (Fig 3b). For most cases, the effect of the mutation on activity *in vivo* correlates with the degree of ES stabilization measured *in vitro*. For example, NMR spectra of TAR-G36U and TAR-A27C indicate that these mutants bias the conformational equilibrium toward ES2 to a similar extent as TAR-G28U (Fig 2c), and correspondingly, their activity is indistinguishable from that observed for TAR-G28U (two-way Anova, $P > 0.05$) (Fig 3b). While TAR-U38A also stabilizes ES2, NMR spectra indicate that the bias is not as strong as the other mutants (Fig 2c). Correspondingly TAR-U38A has marginally higher activity (~6% of WT TAR) compared to other ES2-stabilizing mutants (Fig 3b). Likewise, the partial rescue of TAR-G36U/C29A in the transactivation assay mirrors the partial conformational rescue as assessed by NMR (Fig 2c and Fig 3b), although we cannot rule out that the identity of the closing G36-C29 bp is also important for protein recognition

(Fig 2c and Fig 3b). While the TAR-U38A/A27U and TAR-A27C/U38G mutants only partially restore activity despite NMR spectra showing a strong bias toward the GS, this is not surprising given that the mutations disrupt an U23•U38-A27 base triple known to be important for Tat binding (Fig 2c and Fig 3b)[25,26]. Taken together, the agreement between the conformations determined *in vitro* by NMR and the cell-based trans-activation activity suggests that the energetic stability of ES2 relative to the GS are not substantially altered in the cell relative to the *in vitro* environment. The results also confirm that the ES2 conformation is inhibitory, as is expected given prior studies showing the importance of the TAR GS secondary structure for transactivation[12–16,22].

TAR ES1 does not change the secondary structure as significantly as ES2, and two ES1-stabilizing mutants[6] (TAR-C30U and TAR-A35G) does not significantly decrease trans-activation relative to WT activity (two-way ANOVA, $P > 0.5$) (Fig 3b). Therefore, stabilizing ES1 is unlikely to represent a favorable therapeutic strategy.

RRE RNA mediates nuclear export of incompletely spliced viral RNAs by cooperatively binding multiple copies of the viral Rev protein[17]. This assembly is initiated through binding of Rev to the purine rich region of RRE stem IIB[17]. Our recent work[8] showed that RRE stem IIB forms two non-native ESs (ES1 and ES2) both of which disrupt key structural elements recognized by Rev, including the G47-A73 and G48-G71 mismatches and U72 bulge (Fig 2b). ES1 reshuffles the lower bulge while ES2 remodels both the lower and upper bulge (Fig 2b)[8].

As was done for TAR, point substitution mutations stabilizing the RRE ESs and corresponding rescue mutations that restore the GS were designed and verified using NMR (Fig 2c, Supplementary note 2, Supplementary Figs 5-6). The mutants were then tested in an RRE-dependent RNA export assay to monitor RRE activity (Fig 3c-d). Due to the localized nature of the conformational change, it was challenging to design mutants that stabilize the RRE ESs to >99% population. Rather, extreme broadening of the NMR spectra indicate that RRE-A68C only partially stabilizes ES2 to a population ranging between 20% and 80%. The GS conformation is largely rescued by RRE-A68C/G50A/C69U based on the restoration of WT chemical shifts. As expected since it only partially stabilizes the ES,

RRE-A68C showed only a 2-fold reduction in RRE-mediated export relative to WT while the rescue mutant RRE-A68C/G50A/C69U restores activity to within ~80% of WT (Fig 3d). Likewise, RRE-U72C partially stabilizes the cumulative ES1 + ES2 population to 20%-80%, and decreases export activity to ~70% of WT, while the rescue mutant RRE-U72C/G48A/G71A restores activity to ~100% of WT. Another double mutant, RRE-A68C/U72C, which also partially stabilizes ES2 to 20%-80% based on NMR, also shows a 2-fold reduction in RRE-mediated export to ~50% of WT while the rescue mutant RRE-U72C/G48A/G71A/A68C/G50A/C69U restores activity to ~70% of WT (Fig 3c). Similar results were also observed for Rev-RRE binding in cells using a previously described TAR-RRE hybrid trans-activation assay[27] (Supplementary Fig 7), confirming that inhibition of RRE export is indeed due to inhibition of Rev-RRE binding.

Although the effects are smaller than those observed for TAR, the robust correlation between the impact on conformation and on RRE activity for different ES-stabilizing and rescue mutants indicates that stabilizing the RRE ES conformation is indeed inhibitory (Fig 3d and Supplementary Figs 5-7). Since the mutants only stabilize the RRE ES to ~50% population, the 2-fold reduction in activity is consistent with the ES conformation being fully inhibitory and the observed remaining activity can be attributed to a residual GS population. Once again, if the cellular environment preferentially stabilizes the GS relative to the ES, we would have expected the mutations to have a smaller effect within cells than predicted from their equilibrium populations, which is not observed.

The results presented here indicate that stabilizing non-native RNA ESs could be a viable strategy for inhibiting RNA-protein interactions and cellular activity. For two RNAs, the cellular environment does not seem to significantly alter the energetic stabilities of the ES relative to the GS compared to *in vitro* conditions. Because the ES stabilize-and-rescue approach presented here maximally changes RNA structure while minimally impacting sequence, it provides a powerful approach by which to test the structure-dependence of RNA function within cells.

Online Methods

RNA sample preparation

Unlabeled RNA samples (TAR-G28U, TAR-G28U/C37A, TAR-G36U, TAR-G36U/C29A, TAR-U38A, TAR-U38A/A27U, TAR-A27C, TAR-A27C/U38G, WT TAR, and uucgES2TAR) were prepared using a MerMade 6 DNA/RNA synthesizer (BioAutomation) for solid-phase oligonucleotide synthesis using standard phosphoramidite RNA chemistry and deprotection protocols. $^{13}C/^{15}N$ labeled RNA samples (RRE-A68C/G50A/C69U/U72C/G48A/G71A, RRE-A68C/U72C, RRE-U72C/G48A/G71A, RRE-U72C, RRE-A68C/G50A/C69U, RRE-A68C and TAR-G28U) used for NMR characterization were prepared by *in vitro* synthesis using T7 RNA polymerase (New England BioLabs), DNA template (Integrated DNA Technologies) containing the T7 promoter, and uniformly labeled $^{13}C/^{15}N$- labeled nucleotide triphosphates (Cambridge Isotope Laboratories, Inc.). The transcription reaction was carried out at 37 °C for 12 hours and then filtered with a 0.2 μm filter.

All RNA samples were purified using a 20% (w/v) polyacrylamide gel with 8 M urea and 1 X Tris/borate/EDTA. The RNA was removed from the excised gel by electro-elution in 1X Tris/acetic acid/EDTA followed by ethanol precipitation. The RNA was annealed in water at a concentration of 50 μM by heating at 95 °C for 5 min followed by cooling on ice for 30 min. The RNA was then buffer exchanged using an Amicon Ultra-15 centrifugal filter (EMD Milipore) with a 3 KDa cutoff into NMR buffer [15 mM sodium phosphate, 25 mM sodium chloride, and 0.1 mM EDTA, pH 6.4]. Finally 10% (v/v) D_2O was added to the sample before NMR data were collected. The final concentration of RNA in all samples was between 0.6 and 1.1 mM.

NMR spectroscopy

Assessing impact of mutations using chemical shift mapping experiments

The impact of the mutations on RNA structure was initially assessed by recording 1D [^{1}H] and 2D [^{13}C-^{1}H] SOFAST-HMQC NMR experiments[28] using unlabeled RNA samples. All NMR

experiments were carried out at 298 K on a 600 MHz Bruker spectrometer equipped with HCN cryogenic probes. Data were processed using NMRpipe[29] and analyzed using SPARKY[30].

TAR: Comparison of NMR spectra for each TAR mutant with corresponding spectra of WT TAR and a validated TAR ES2 mutant (uucgES2 TAR)[31,32] was used to assess the degree of ES- versus GS-stabilization. Although the entire spectrum was evaluated (Supplementary Fig 2), key residues used in this analysis include A22 C8-H8/C2-H2, U23 C6-H6, C24 C6-H6, A35 C8-H8, U40 C6-H6, and C41 C6-H6. To determine the degree of stabilization of ES-and GS- stabilizing mutants, the extent of broadening for these spectra was investigated. The broadening of TAR-U38A is apparent in many resonances, but particularly in the peak tentatively assigned to U23, which is not broadened out in the other ES2 traps. The broadening of TAR-G28U/C37A and TAR-G36U/C29A relative to the other rescue mutants is also particularly visible in U23.

RRE: Likewise, comparison of NMR spectra for each RRE mutant with corresponding spectra of WT RRE was used to assess the degree ES- versus GS-stabilization. The entire spectra were evaluated (Supplementary Fig 5), and the key resonances used in this analysis include A68 C8-H8, U72 C6-H6, G47 C8-H8, G48 C8-H8, and G71 C8-H8. The broadening of these resonances, which experience significant changes in chemical shift between the GS and ESs were used to estimate the relative stability of ESs versus GS.

^{15}N $R_{1\rho}$ relaxation dispersion

^{15}N $R_{1\rho}$ relaxation dispersion (RD) NMR experiments were carried out using a 700 MHz Bruker Avance III spectrometer equipped with an HCN cryogenic probe. Measurements were carried out at 25°C on the 0.7 mM uniformly $^{13}C/^{15}N$-labeled TAR-G28U sample. $R_{1\rho}$ RD for N3 spins were measured using a 1D acquisition scheme[33] using spinlock powers ($\omega 2\pi^{-1}$ Hz) and offsets ($\Omega 2\pi^{-1}$ Hz) listed in Supplementary Table 1 and using 7 time delay points between 0 and 120 ms. For each delay

time point, peak intensities were obtained using NMRPipe and $R_{1\rho}$ values were calculated by fitting to a monoexponential decay function using an in-house python script[34]. Errors in $R_{1\rho}$ were estimated using Monte Carlo simulations with 500 iterations as previously described[35]. The RD data was analyzed to obtain exchange parameters through numerical solutions of the Bloch-McConnell equations[36] using an in-house python script[37]. As expected the TAR-U42-N3 negative control showed no dispersion. The TAR-U38-N3 data was fit with an individual two-state model using ground state effective field alignment. Model selection was preformed as previously described[37] using the Akaike's (*w*AIC) and Bayesian information criterion (*w*BIC) weights to select the model with the highest relative probability[38,39].

Plasmid construction

The Tat and Rev expression plasmids, pcTat[27] and pcRev[40], pcTat/Rev[27], and pBC12-CMV[27] have been described previously. The Renilla luciferase (RLuc) internal control was generated by PCR amplification of the RLuc ORF, which was then cloned into pcDNA3 (Invitrogen) using unique HindIII and XhoI restriction sites. The Firefly luciferase (FLuc) reporter plasmids with TAR and RRE, both wild-type and mutants, were constructed as follows. The pFLuc-TAR plasmid was generated by cloning the FLuc gene into pcDNA3 (Invitrogen) using NotI and XhoI restriction sites to generate pcDNA-FLuc. The HIV-LTR was next cloned as a Bam HI/Hind III fragment into pcDNA-FLuc digested with Bgl II and Hind III, which removes the entire CMV promoter. TAR mutants were introduced at Bgl II and Hind III restriction sites using annealed DNA oligos (IDT) that were designed to have overhangs complimentary to Bgl II and Hind III. The pFLuc-TAR/RRE plasmid was generated with the same Bgl II and Hind III restriction sites using annealed DNA oligos (IDT) containing RRE stem II sequence (66nt). The pFLuc-RRE was derived from pDM128-PL[41]. The CAT gene was replaced with the FLuc gene using Sal I and Bgl II restriction sites (pMP-1). The RRE and variants were generated by ligating in the whole RRE (~350 nt) inserts amplified using two-step recombinant

PCR, using the HIV-1 proviral clone pNL4-3 as the template, (AIDS reagent cat #114). The PCR products of RRE and variants were then inserted to pMP-1 at Kpn I/Xba I restriction sites.

Trans-activation assays

HeLa cells were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 0.1% gentamycin at 37 °C and 5% CO_2. Cells were plated to 3 X 10^5 cells per well in 12-well plates and transfected using PEI with 250 ng pFLuc-TAR reporter plasmid, 20 ng RLuc internal control, 20 ng pcTat, and pBC12-CMV filler DNA up to 1750 ng total DNA. Media was replaced at 24 h post-transfection and cells were lysed at 48 h post-transfection using 200 μl passive lysis buffer (Promega) followed by freezing at -80 °C. FLuc and RLuc activity was measured using a dual-luciferase reporter assay system (Promega). Reported values are the quotient of FLuc activity and RLuc activity with values normalized to WT TAR for every replicate (Fig 3b).

Rev-RRE binding assay

For Rev-RRE binding assay, HeLa cells (2 X10^5 cells per well) were transfected using PEI with 350 ng pFLuc-TAR/RRE reporter plasmid, 20 ng RLuc internal control, and 100 ng pcTat/Rev or 100 ng pcTat, and pBC12-CMV filler DNA up to 1370 ng total DNA. The cell lysate preparation and luciferase measurement are as described above for the trans-activation assay. Reported values are the quotient of FLuc activity and RLuc activity with values normalized to WT RREII for every replicate (Supplementary Fig 7).

RRE-mediated RNA export assay

293T cells were maintained in DMEM supplemented with 10% FBS and 0.1% gentamycin at 37 °C and 5% CO_2. Cells were plated to 2 X10^5 cells per well in 12-well plates and transfected using PEI with 5 ng pFLuc-RRE reporter plasmid, 5 ng RLuc internal control, 1 ng pcRev, and pBC12-CMV

filler DNA up to 1010 ng total DNA. Media was replaced at 24 h post-transfection and cells were lysed at 48 h post-transfection using 200 μl passive lysis buffer (Promega). FLuc and RLuc activity was measured using a dual-luciferase reporter assay system (Promega). Reported values are the quotient of FLuc activity and RLuc activity with values normalized to WT RREII activity for every replicate (Fig 3d).

Statistical analysis

Statistical analysis was done using the program JMP (JMP Pro, Version 14, SAS Institute Inc., Cary, NC, 1989-2019). The data was not normalized to WT TAR for the trans-activation assay because there was not a significant systematic effect of replicate on the data. For both RRE assays, the data was normalized to WT RREII because there was a significant systematic effect of replicate on the data. For all three assays, the FLuc/RLuc quotient was log transformed for statistical analysis. A two-sided unpaired Student's t-test was used to compare the +Tat and –Tat distributions for the trans-activation assays and a two-sided ANOVA was used to compare the difference between distributions (+Tat and –Tat; +Rev and –Rev; or +Tat/Rev and +Tat) for different samples for all three assays.

Acknowledgements

We would like to thank the Duke Magnetic Resonance Spectroscopy Center for NMR resources. This work was supported by the US National Institutes of Health (P50 GM103297 to B.R.C. and H.M.A.-H and F31 GM119306 to L.R.G).

Author Contributions

The study was designed by L.R.G, C-C.C. and H.M.A.-H. Experiments were performed by L.R.G. C-C.C., H.P.B. and M.L.K. The manuscript was written by L.R.G., C-C.C., and H.M.A.-H. with input from all other co-authors.

Competing Interest

H.M.A.–H is an advisor to and holds an ownership interest in Nymirum Inc, an RNA-based drug discovery company.

References

1. Connelly, C. M., Moon, M. H. & Schneekloth, J. S. The emerging role of RNA as a therapeutic target for small molecules. *Cell Chem. Biol.* **23,** 1077–1090 (2016).
2. Dethoff, E. A., Chugh, J., Mustoe, A. M. & Al-Hashimi, H. M. Functional complexity and regulation through RNA dynamics. *Nature* **482,** 322–330 (2012).
3. Walter, F., Vicens, Q. & Westhof, E. Aminoglycoside – RNA interactions. *Curr.Opin. Chem. Biol.* **3,** 694–704 (1999).
4. Hermann, T. Rational ligand design for RNA: The role of static structure and conformational flexibility in target recognition. *Biochimie* **84,** 869–875 (2002).
5. Dibrov, S. M. *et al.* Hepatitis C virus translation inhibitors targeting the internal ribosomal entry site. *J Med Chem* **57,** 1694–1707 (2014).
6. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491,** 724–728 (2012).
7. Lee, J., Dethoff, E. A. & Al-Hashimi, H. M. Invisible RNA state dynamically couples distant motifs. *PNAS* **111,** 9485–9490 (2014).
8. Chu, C.-C., Plangger, R., Kreutz, C. & Hashimi, H. M. A.-. Dynamic ensemble of HIV-1 RRE stem IIB reveals non-native conformations that disrupt the Rev binding site. *bioRxiv doi:https://doi.org/10.1101/498907* (2018).
9. Xue, Y., Gracia, B., Herschlag, D., Russell, R. & Al-Hashimi, H. M. Visualizing the formation of an RNA folding intermediate through a fast highly modular secondary structure switch. *Nat. Commun.* **7,** 11768 (2016).
10. Zhao, B., Guffy, S. L., Williams, B. & Zhang, Q. An excited state underlies gene regulation of a transcriptional riboswitch. *Nat. Chem. Biol.* **13,** 968–974 (2017).
11. Longhini, A. P. *et al.* Chemo-enzymatic synthesis of site-specific isotopically labeled nucleotides for use in NMR resonance assignment , dynamics and structural characterizations. *Nucleic Acids Res.* **44,** e52 (2016).

12. Feng, S. & Holland, E. C. HIV-1 tat trans-activation requires the loop sequene within tar. *Nature* **334,** 165–167 (1988).

13. Berkhout, B. & Jeang, K.-T. trans activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. *J. Virol.* **63,** 5501–4 (1989).

14. Roy, S., Delling, U., Chen, C. H., Rosen, C. A. & Sonenberg, N. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev.* **4,** 1365–73 (1990).

15. Dingwall, C. *et al.* HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J.* **9,** 4145–4153 (1990).

16. Churcher, M. J. *et al.* 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.* **230,** 90–110 (1993).

17. Bartel, D. P., Zapp, M. L., Green, M. R. & Szostak, J. W. HIV-1 Rev Regulation Involves Recognition of Non-Watson-Crick Base Pairs in Viral RNA. *Cell* **67,** 529–536 (1991).

18. Jain, C. & Belasco, J. G. A Structural Model for the HIV-1 Rev – RRE Complex Deduced from Altered-Specificity Rev Variants Isolated by a Rapid Genetic Strategy. *Cell* **87,** 115–125 (1996).

19. Daugherty, M. D., D'Orso, I. & Frankel, A. D. A Solution to Limited Genomic Capacity : Using Adaptable Binding Surfaces to Assemble the Functional HIV Rev Oligomer on RNA. *Mol. Cell* **31,** 824–834 (2008).

20. Malim, M. H. *et al.* HIV-1 Structural Gene Expression of the Rev Trans-Activator to Its RNA Target Sequence Requires Binding to its RNA target sequence. *Cell* **60,** 675–683 (1990).

21. He, N. *et al.* HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol. Cell* **38,** 428–438 (2010).

22. Schulze-Gahmen, U. *et al.* Insights into HIV-1 proviral transcription from integrative structure

and dynamics of the Tat:AFF4:P-TEFb:TAR complex. *Elife* **5,** e15910 (2016).

23. Schulze-Gahmen, U. & Hurley, J. H. Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex. *PNAS* **115,** 12973–12978 (2018).

24. Xue, Y. *et al. Methods in Enzymology* 558, 39–73 (Elsevier Inc., 2015).

25. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D. & Williamson, J. R. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science.* **257,** 76–80 (1992).

26. Puglisi, J. D., Chen, L., Frankel, A. D. & Williamson, J. R. Role of RNA structure in arginine recognition of TAR RNA. *PNAS* **90,** 3680–3684 (1993).

27. Tiley, L. S., Madore, S. J., Malim, M. H. & Cullen, B. R. The VP16 transcription activation domain is functional when targeted to a promoter-proximal RNA sequence. *Genes Dev.* **6,** 2077–2087 (1992).

28. Sathyamoorthy, B., Lee, J., Kimsey, I., Ganser, L. R. & Al-Hashimi, H. Development and application of aromatic [13C, 1H] SOFAST-HMQC NMR experiment for nucleic acids. *J. Biomol. NMR* **60,** 77–83 (2014).

29. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–93 (1995).

30. Goddard, T. D. & Kneller, D. G. SPARKY 3.

31. Merriman, D. K. *et al.* Shortening the HIV-1 TAR RNA bulge by a single nucleotide preserves motional modes over a broad range of time scales. *Biochemistry* **55,** 4445–4456 (2016).

32. Clay, M. C., Ganser, L. R., Merriman, D. K. & Al-Hashimi, H. M. Resolving sugar puckers in RNA excited states exposes slow modes of repuckering dynamics. *Nucleic Acids Res.* **45,** e134 (2017).

33. Nikolova, E. N., Gottardo, F. L. & Al-Hashimi, H. M. Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *JACS* **134,** 3667–3670 (2012).

34. Kimsey, I. J., Petzold, K., Sathyamoorthy, B., Stein, Z. W. & Al-Hashimi, H. M. Visualizing

transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature* **519,** 315–320 (2015).

35. Bothe, J. R., Stein, Z. W. & Al-Hashimi, H. M. Evaluating the uncertainty in exchange parameters determined from off-resonance R1ρ relaxation dispersion for systems in fast exchange. *J. Magn. Reson.* **244,** 18–29 (2014).

36. Korzhnev, D. M., Orekhov, V. Y. & Kay, L. E. Off-resonance R(1rho) NMR studies of exchange dynamics in proteins with low spin-lock fields: an application to a Fyn SH3 domain. *JACS* **127,** 713–21 (2005).

37. Kimsey, I. J. *et al.* Dynamic basis for dG • dT misincorporation via tautomerization and ionization. *Nature* **554,** 195–201 (2018).

38. Wagenmakers, E. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11,** 192–196 (2004).

39. Burnham, K. P. & Anderson, D. R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33,** 261–304 (2004).

40. Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335,** 181–183 (1988).

41. Fridell, R. A., Partin, K. M., Carpenter, S. & Cullen, B. R. Identification of the Activation Domain of Equine Infectious Anemia Virus Rev. **67,** 7317–7323 (1993).

Figure Legends

Figure 1. Evaluating a therapeutic strategy involving targeted stabilization of RNA ESs in cells. **a.** Substitution mutations (shown as star) redistribute the WT RNA ensemble *in vitro* increasing the abundance of an ES over the GS to various degrees measured by NMR spectroscopy. The impact of the mutations on *in vivo* activity is then measured using an appropriate assay. **b.** A decrease in cellar activity relative to WT (blue circle) commensurate with the fold-increase in ES abundance (red circles) implies that the ES forms in cells with energetics comparable to those measured *in vitro* and that the ES conformation is fully inhibitory. Weaker levels of inhibition of cellular activity not commensurate with the fold-increase in ES population (green triangle) imply that the cellular environment favors the GS over the ES relative to *in vitro* and/or that the ES is not fully inhibitory. Rescue mutations (open red circles) are used to restore the GS conformation and to control for the impact of changes in sequence on activity.

Figure 2. Excited states of HIV-1 TAR and RRE stem IIB. a. HIV-1 TAR RNA has two excited states that alter base pairing in the bulge (orange), upper stem (blue), and apical loop (green). Weak base pairs are indicated with an open dot and non-Watson-Crick base pairs are indicated with a red dot. **b.** HIV-1 RRE RNA has two excited states that reshuffle the non-canonical base pairs at the purine-rich region (blue), U72 (yellow), and A68 (green) bulge. Weak base pairs are indicated with an open dot and non-Watson-Crick base pairs are indicated with a red dot. **c.** ES-stabilizing mutants and corresponding rescue mutations. The extent of stabilization was estimated based on NMR line broadening (Supplementary Figs 2 and 5). (-) minimal line broadening, (+) partial line broadening, and (++) extensive line broadening.

Figure 3. Cell-based activity assays for TAR and RRE. **a. and c.** Examples of the stabilize-and-rescue approach for trapping TAR ES2 (**a**) and RRE ES2 (**c**) using an ES-stabilizing mutant (orange) and corresponding rescue mutation (blue). Weak base pairs are indicated with an open dot and non-

Watson-Crick base pairs are indicated with a red dot. **b.** Tat-dependent trans-activation assays of TAR mutants. HeLa cells were transfected with pFLuc-TAR reporter plasmid and an RLuc internal control in the presence (+ Tat) or absence (-Tat) of a Tat expression plasmid. At 48 hr post transfection, cell lysates were tested for luciferase activity. Reported values are the quotient of FLuc and RLuc activities with values normalized to WT TAR for every independent replicate. **d.** Rev-RRE dependent RNA export assays of RRE mutants. 293T cell were transfected with pFLuc-RRE reporter plasmid and an RLuc internal control in the presence (+ Rev) or absence (-Rev) of a Rev expression plasmid. At 48 hr post transfection, cell lysates were tested for luciferase activity. Reported values are the quotient of FLuc and RLuc activities with values normalized to WT RRE for every independent replicate. MP-1 is the control plasmid (no RRE). For **b** and **d**, statistical significance was determined after log-transformation of the data; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ (two-sided ANOVA).

Figures

Figure 1.

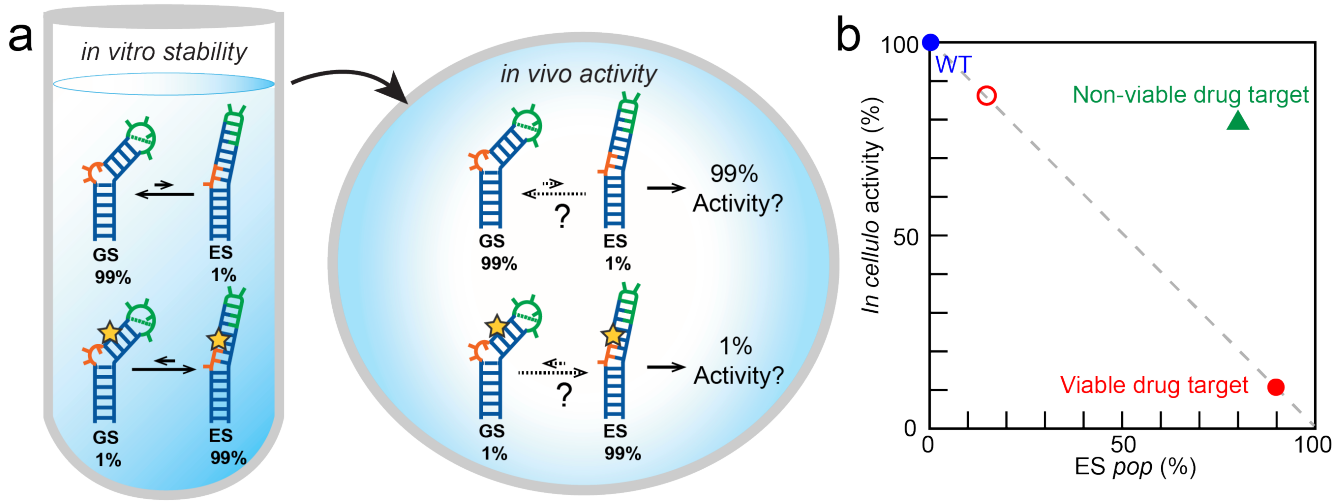
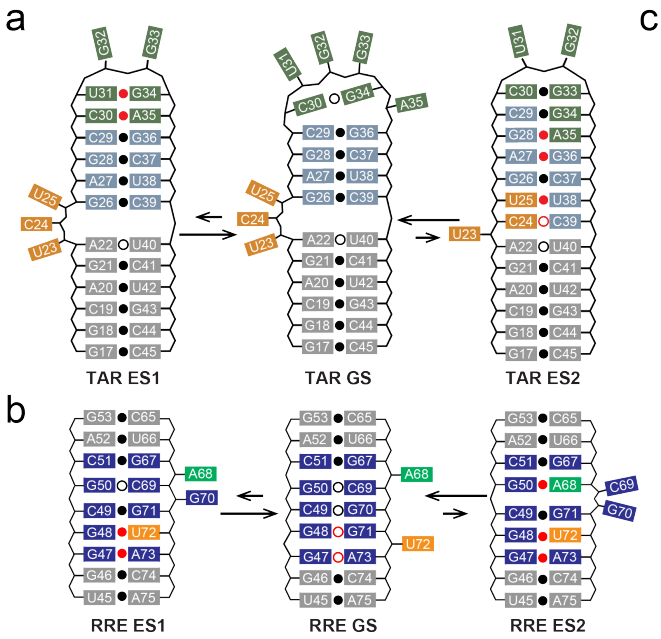


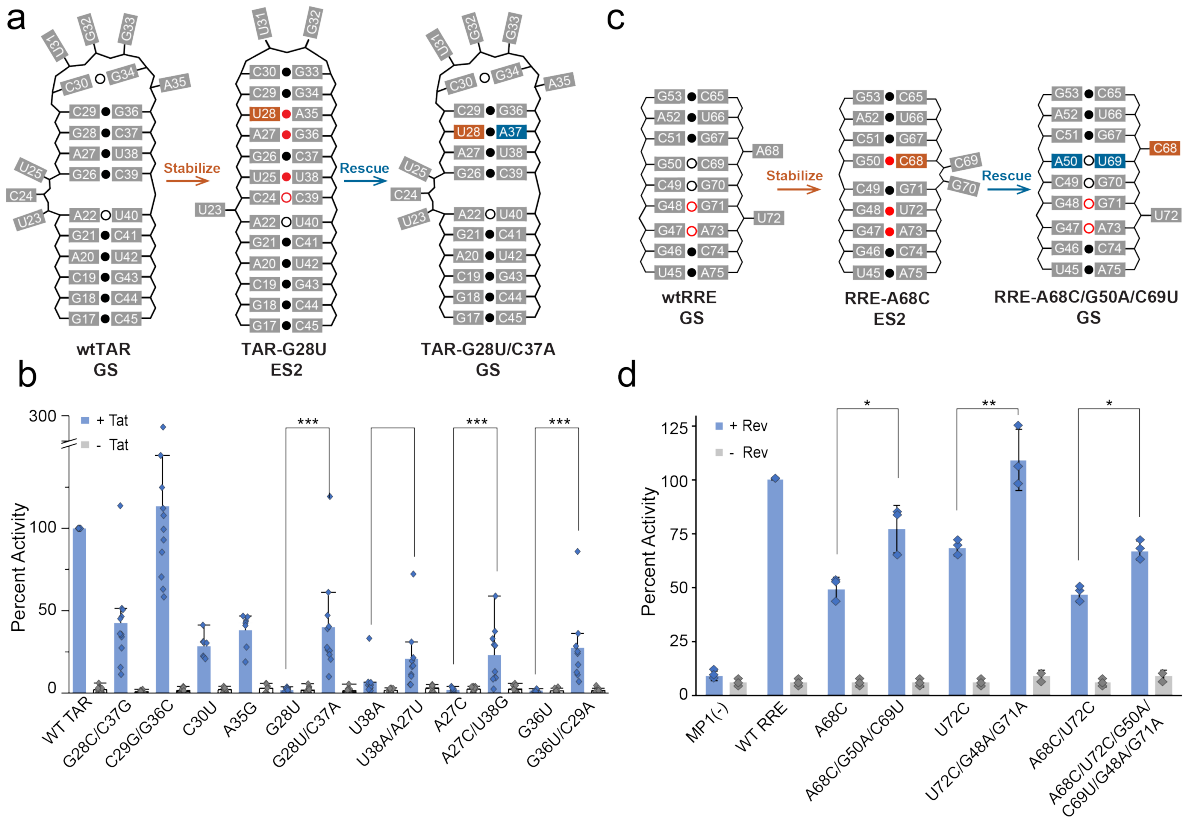
Figure 2.



c

	ES-stabilizing mutants	Extent of ES stabilization	Rescue mutants	Extent of GS stabilization
TAR ES2	G28U	-	G28U/C37A	+
	G36U	-	G36U/C29A	+
	A27C	-	A27C/U38G	-
	U38A	+	U38A/A27U	
TAR ES1	A35G	-	NA	
	C30U		NA	-
TAR GS	NA		G28C/C37G	-
	NA	-	C29G/G36C	-
RRE ES2	A68C	++	A68C/G50A/C69U	+
	A68C/U72C	++	A68C/G50A/C69U /U72C/G48A/G71A	+
RRE ES1	U72C	++	U72C/G48A/G71A	+

Figure 3.


a
wtTAR
GS
Stabilize
TAR-G28U
ES2
Rescue
TAR-G28U/C37A
GS
c
wtRRE
GS
Stabilize
RRE-A68C
ES2
Rescue
RRE-A68C/G50A/C69U
GS
b
+ Tat
- Tat
Percent Activity
WT TAR
G28C/C37G
C29G/G36C
C30U
A35G
G28U
G28U/C37A
U38A
U38A/A27U
A27C
A27C/U38G
G36U
G36U/C29A
d
+ Rev
- Rev
Percent Activity
MP1(-)
WT RRE
A68C
A68C/G50A/C69U
U72C
U72C/G48A/G71A
A68C/U72C
A68C/U72C/G50A/
C69U/G48A/G71A