Visualizing transient low-populated structures of RNA

Elizabeth A. Dethoff*, Katja Petzold*, Jeetender Chugh*, Anette Casiano-Negron† & Hashim M. Al-Hashimi†

The visualization of RNA conformational changes has provided fundamental insights into how regulatory RNAs carry out their biological functions. The RNA structural transitions that have been characterized so far involve long-lived species that can be captured by structure characterization techniques. Here we report the nuclear magnetic resonance visualization of RNA transitions towards ‘invisible’ excited states (ESs), which exist in too little abundance (2–13%) and for too short a duration (45–250 μs) to allow structural characterization by conventional techniques. Transitions towards ESs result in localized rearrangements in base-pairing that alter building block elements of RNA architecture, including helix–junction–helix motifs and apical loops. The ES can inhibit function by sequestering residues involved in recognition and signalling or promote ATP–independent strand exchange. Thus, RNAs do not adopt a single conformation, but rather exist in rapid equilibrium with alternative ESs, which can be stabilized by cellular cues to affect functional outcomes.

Nuclear magnetic resonance (NMR) relaxation dispersion methods8–11, which measure microsecond-to-millisecond conformational exchange, have made it possible to characterize the transient, low-populated excited state (ES) structures of proteins8–12 and to establish their importance in catalysis13, folding14–16, signalling17 and recognition18. These ESs exist in too little abundance (typically with populations <5%) and for too short a duration (lifetime < milliseconds) to allow structural characterization by conventional techniques. Recent advances that extend the timescale sensitivity of rotating frame (R1) carbon relaxation dispersion experiments have made it possible to characterize fully exchange processes in nucleic acids19–21, culminating in the discovery of ES Hoogsteen base pairs in DNA12. Although evidence for RNA ESs has been reported for decades, their structure and role in function have remained elusive13–15.

Here we report a strategy for characterizing the ES structures of RNA that combines R1ρ NMR experiments, mutagenesis and secondary structure prediction. With this approach, we visualized ES structures for three distinct RNAs and obtained insights into their biological functions.

**ES structure of the HIV TAR apical loop**

We used a low spin-lock field R1ρ NMR experiment6–11 to measure microsecond-to-millisecond conformational exchange at sugar (C1′) and nucleobase (C8 and C6) carbon sites in the well studied hexanucleotide apical loop of the transcriptional response element (TAR)16 from the human immunodeficiency virus type-1 (HIV-1). The TAR apical loop is a flexible recognition site that allows adaptive binding to a variety of proteins17. We observed conformational exchange (Fig. 1b and Supplementary Fig. 1) at carbon sites spread throughout the entire TAR apical loop (Fig. 1a). The R1ρ data could be collectively fitted to a two-state (GS ↔ ES) exchange process (where GS indicates ground state) that is directed towards an ES with population pES ≈ 13% and lifetime (t1ES ≈ 1/k−1ES) ≈ 45 μs (Supplementary Table 1). A slower exchange process is observed at C1′ of G33 (G33-C1′), G33-C8 and A35-C8 (pES <1% and t1ES = 1.9–2.3 ms), which can be assigned to a distinct higher energy ES that will not be discussed further (Supplementary Discussion and Supplementary Fig. 5).

In the ground state (GS), apical loop residues exist in equilibrium between C2′-endo and C3′-endo sugar puckers, G34 forms a flexible cross-loop C30•G34 Watson–Crick (WC) base pair, whereas the bases of U31, G32 and A35 are flexible18,19. To gain insights into the ES structure, we examined the sugar and base ES carbon chemical shifts (δC) obtained from the two-state analysis of the R1ρ data, which are sensitive reporters of base stacking, sugar pucker and syn versus anti glycosidic angles20. The downfield-shifted sugar ES C30•C1′, U31-C1′ and A35-C1′ chemical shifts strongly suggest that in the ES these residues adopt a pure C3′-endo sugar pucker characteristic of a helical conformation (Fig. 1a and Supplementary Table 1). The downfield-shifted base ES G34•C8 can unambiguously be assigned to a syn base21 (Supplementary Discussion) and has a chemical shift that is highly characteristic of a UUCG tetraloop, which features a trans-wobble G•U base pair (underlined) and a syn base (italics)22. Notably, TAR can accommodate a similar U31G32G33C8 tetraloop. This places G34 in a syn position, where it can base pair with U31, thus explaining exchange at U31-C6. It also leads to the formation of C30•A35 and U31•G34 non-canonical closing base pairs, explaining the helical conformation observed for these residues in the ES. Transitions towards this ES require disruption of the cross-strand C30•G34 base pair, explaining the measured activation free energy (12.6 kcal mol−1) (Supplementary Fig. 2), which is at the low end of the free energy range required to open RNA WC base pairs (13–16 kcal mol−1)23. This ES is also predicted to be the second most energetically favourable conformation using the secondary structure prediction program MC-Fold24 (Supplementary Fig. 3).

We used a ‘mutate-and-chemical-shift-fingerprint’ (MCSF) strategy to test the proposed TAR ES. Here, a mutation or chemical modification is introduced to stabilize (or destabilize) a candidate ES, and the mutant’s NMR carbon chemical shift fingerprints are compared with those of the ES (or GS). We stabilized the proposed TAR ES using two point mutations, C30U (TAR(C30U)) and A35G (TAR(A35G)), that replace the ES C30•A35 non-canonical base pair with more stable WC U30•A35 and C30•G35 base pairs, respectively (Fig. 1c). Both mutants adopted the proposed ES structure, as confirmed by NMR
The functional significance of the TAR ES. The ES sequesters the U31-C1′ and G34-C8 points to a larger structural rearrangement. The downfield-shifted base carbon chemical shift for U95 indicates the opening of C07 and Supplementary Table 1). These data can be explained by an alternative structure in which U95 bulges out while A93 is partially flipped out and flexible31 (Fig. 2a). An ES involving the flipping out of A92 and A93, as observed in several X-ray and NMR structures of drug-bound A-site25, can be ruled out based on the observation of exchange below the internal loop, ES chemical shift fingerprints that suggest increased stacking for A93 (Supplementary Fig. 5), and by comparison of ES chemical shifts with those of drug-bound A-site (Supplementary Fig. 7).

Rather, the breadth of exchange across many different residues points to a larger structural rearrangement. The downfield-shifted base ES chemical shift for U95-C6 indicates looping out of U95, whereas the upfield-shifted base carbon ES chemical shifts indicate increased stacking for A92, A93, G94 and C96 (Fig. 2a and Supplementary Table 1). A transition towards such an ES requires the opening of C07•G94, explaining the sizable free-energy barrier of ~14.8 kcal mol⁻¹ (Supplementary Fig. 2)35. This ES is predicted by MC-Fold to be the second most energetically favourable secondary structure (Supplementary Fig. 3) and has previously been observed in molecular dynamics simulations33.

ES structure of the ribosomal A-site

We used our strategy to characterize the ES structure of the ribosomal A-site internal loop29 (Fig. 2a). The A-site has essential roles in decoding messenger RNA by flipping out two internal-loop adenines (A1492 and A1493, referred to hereafter as A92 and A93), which interact with and stabilize the codon–anticodon mini-helix formed between the cognate aminoacyl tRNA and mRNA29,30 (see Fig. 2e). We observed extensive carbon chemical exchange at seven residues within and below the A-site internal loop (Fig. 2a, b and Supplementary Fig. 1). A two-state analysis of the R₁ data revealed a global exchange process directed towards an ES with population ρES ≃ 2.5% and lifetime τES = 1/κ₁ ≃ 248 μs (Supplementary Table 1).

Biophysical studies show that in the GS, A92 is looped inside, probably forming a base pair with A08, whereas A93 is partially flipped out and flexible31 (Fig. 2a). An ES involving the flipping out of A92 and A93, as observed in several X-ray and NMR structures of drug-bound A-site25, can be ruled out based on the observation of exchange below the internal loop, ES chemical shift fingerprints that suggest increased stacking for A93 (Supplementary Fig. 5), and by comparison of ES chemical shifts with those of drug-bound A-site (Supplementary Fig. 7).
We confirmed the proposed A-site ES using MCSF analysis. We were able to block transitions towards the ES by replacing U06•U95 with a more stable WC G06•C95 base pair (A-site(U06G/U95C)) (Fig. 2c). This locked the A-site into the GS as judged by the GS-like chemical shifts (Fig. 2d and Supplementary Fig. 5) and absence of chemical exchange, including at sites (for example, A92 and A93) that are distant from the site of mutation (Supplementary Fig. 1). This also confirmed that all sites experience a common global exchange process. We then stabilized the proposed ES by deleting U95, which bulges out in the ES (A-site(U95-N3M)), and by introducing a methyl group at U95-N3 (A-site(U95-N3M)), which is expected to disrupt the GS U06•U95 non-canonical base pair in favour of the bulged-out ES conformation (Fig. 2c). The A-site(U95) mutant adopted the proposed ES structure as confirmed by NMR (Supplementary Fig. 4) and resulted in large changes in the carbon chemical shifts specifically at sites showing exchange that are directed towards the ES chemical exchange (Fig. 2d and Supplementary Fig. 5). More dramatically, the A-site(U95-N3M) mutant exhibited two equally populated sets of resonances in slow exchange on the NMR timescale (Supplementary Fig. 4), with one set corresponding to the GS and the other in near-perfect agreement with the ES (Fig. 2d and Supplementary Fig. 5).

The A-site ES sequesters A92 and A93 into base pairs, making them unavailable to decode mRNA. It also affects the structural presentation of A-site residues involved in protein recognition and formation of the B2a intersubunit crossbridge (Fig. 2a, e). Thus, we analysed previous mutational data in light of the ES A-site structure determined here. Interestingly, mutants that are predicted to stabilize the A-site ES are distant from the site of mutation (Supplementary Table 1). Rather, at least two distinct ESs (ES1 and ES2) need to be invoked that are sensed by residues above (ES1, ΔES1 = 0.1 kHz) and below (ES2, ΔES2 = 0.2 kHz) the ES carbon chemical shift fingerprints and MCSF analysis, with the ES carbon chemical shifts specific to the ES, GS, and mutant mimics.

**Two ES structures in HIV-1 stem loop 1**

Finally, we used our strategy to study the ES structure of the HIV-1 stem loop 1 (SL1) (Fig. 3a). SL1 spontaneously forms kissing dimers, which isomerize during viral maturation into more stable duplex dimers through mechanisms that remain poorly understood. We confirmed the ES structure determined here (data not shown). The ES ES invites reassessment of the A-site region in current ribosome structures and suggests a new route for targeting the A-site in the development of antibiotics.
In ES1, the bulge migrates upward by 3 base pairs. Here, G29-G28 swap base-pairing partners with A27-G26, A27 swaps with A25, and G26-A25 are bulged out (Fig. 3a). We stabilized ES1 using two point mutants (SL1m(G8C) and SL1m(G8U)) that replace the ES G8×G9 mismatch with the more stable C8×G9 and U8×G9 base pairs (Fig. 3c). Both mutants adopted the ES1 structure as verified by NMR (Supplementary Fig. 4), and relative to wild-type SL1m resulted in large changes in carbon chemical shifts for residues within (G28 and G29) and above (C30 and G31) the internal loop that are directed towards the ES chemical shifts (Fig. 3d and Supplementary Fig. 5). In ES2, the bulge migrates downward by 2 base pairs. Here, G28-G29 swap base-pairing partners with C30-G31, which are now bulged out (Fig. 3a). We stabilized ES2 by replacing the ES2 G7×G8/ G8×A27 mismatch with C7×G8/U8×A27 WC base pairs (Fig. 3c). This double mutant (SL1m(G7C/G8U)) adopted the proposed ES2 structure as verified by NMR (Supplementary Fig. 4) and resulted in large changes in the carbon chemical shifts for residues within (G28 and G29) and below (C30 and G31) the internal loop that are directed towards the ES chemical shifts (Fig. 3d). mutant mimics of ES1 and ES2 induce similar chemical shift perturbations for G28 (C8 and C1′) and G29 (C8), as expected given that they form base pairs in the two cases (Fig. 3d, Supplementary Fig. 5 and Supplementary Discussion). Notably, mutants that stabilize residues above the bulge in their ES conformation also stabilize residues below the bulge in their GS conformation and vice versa (Fig. 3d and Supplementary Fig. 5). This supports the mutual exclusivity of ES1 and ES2 (Fig. 3a); ‘trapping’ the bulge in the upper (or lower) helix prevents downward (or upward) migration and therefore traps residues in the lower (or upper) helix in their GS.

Together, the GS, ES1 and ES2 define a moving zipper in which bulge residues invade base pairs in the upper or lower helix. Remarkably, an analogous process, if carried out in an intermolecular manner between two SL1 monomers, naturally leads to isomerization and duplex formation most probably through a previously proposed quadruplex-like intermediate39 (Fig. 3e). Here, bulged-out G28 and G29 can invade base pairs in the upper or lower helix in another monomer to generate ES1- or ES2-like intermolecular base pairs (Fig. 3e). The bulged-out G26 and A25 or C30 and G31 can, in turn, carry out further intermolecular strand invasions, and this process can be repeated to generate a duplex dimer (Fig. 3e). In support of this important role for ES1 and ES2 in SL1 isomerization, mutations that trap ES1 or inhibit formation of ES2 significantly diminish the rate of isomerization, whereas control sequences that preserve the stability of the stem-loop without disrupting conformational exchange show little to no effect (Fig. 3f and Supplementary Fig. 8). Thus, transitions between the GS and ES can promote ATP-independent changes in RNA secondary structure without disrupting the structural integrity of entire hairpins, which may be required for other functions, such as the formation of kissing dimers in SL1.

Compared to secondary structural transitions observed in many regulatory RNA switches43−46, transitions between the ground and excited states uncovered here involve much more localized changes in RNA structure, occur at rates that are two-to-four orders of magnitude faster, and do not require assistance from external factors. Thus, they can meet unique demands in biological circuits and molecular machines. The ESs also present new drug targets and offer opportunities for mutating RNA-based devices. Line-broadening indicative of ESs is routinely observed in NMR spectra of RNA and we therefore predict that RNA ESs exist in great abundance throughout the transcriptome. By combining NMR data with structure prediction tools, it should be possible to determine the three-dimensional structures of RNA ESs at atomic resolution.

**METHODS SUMMARY**

Detailed methods on RNA sample preparation and assignment, NMR relaxation dispersion data collection and analysis, and isomerization assays can be found in Methods.

©2012 Macmillan Publishers Limited. All rights reserved
Full Methods and any associated references are available in the online version of the paper.

Received 13 June; accepted 10 August 2012.

Published online 7 October 2012.

1. Palmer, A. G. & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. Chem. Rev. 106, 1706–1719 (2006).
2. Baldwin, A. J., Hansen, D. F., Vallurupalli, P. & Kay, L. E. Measurement of methyl axis orientations in invisible, excited states of proteins by relaxation dispersion NMR spectroscopy. J. Am. Chem. Soc. 131, 11939–11948 (2009).
3. Neudecker, P. et al. Structure of an intermediate state in protein folding and aggregation. Science 336, 362–366 (2012).
4. Hengerl-Wildman, K. & Kern, D. Dynamic personalities of proteins. Nature 450, 964–972 (2007).
5. Sugase, K., Dyson, H. J. & Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature 447, 1021–1025 (2007).
6. Kozhnhev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R. & Kay, L. E. A transient and low-populated protein-folding intermediate at atomic resolution. Science 329, 1312–1316 (2010).
7. Li, P. et al. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. Chem. Rev. 106, 1706–1719 (2006).
8. Boehr, D. D., McElhaney, D., Dyson, H. J. & Wright, P. E. The dynamic energy landscape of enzyme catalysis. Science 313, 1638–1642 (2006).
9. Hansen, A. L., Nikolova, E. N., Casiano-Negron, A. & Al-Hashimi, H. M. Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R(1rho) NMR spectroscopy. J. Am. Chem. Soc. 131, 3818–3819 (2009).
10. Massi, F., Johnson, E., Wang, C., Rance, M. & Palmer, A. G. NMR R(1rho) rotating-frame relaxation with weak radio frequency fields. J. Am. Chem. Soc. 126, 2247–2256 (2004).
11. Ghosh, D. M., Orekhov, Y. V. & 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. J. Am. Chem. Soc. 127, 713–721 (2005).
12. Nikolova, E. N. et al. Transient Hoogsteen base pairs in canonical duplex DNA. Nature 470, 499–502 (2011).
13. Hoogenstraat, G. K., Rose, M. K. Internal dynamic control activation and activity of the autoinhibited Vah DH domain. Nature Struct. Biol. 15, 613–618 (2008).
METHODS

Preparation and NMR resonance assignment of labelled and unlabelled RNA. RNA samples were prepared by in vitro transcription using T7 RNA polymerase (Takara Mirus Bio, Inc.), uniformly $^{13}C^{15}$N-labelled nucleotide triphosphates (ISOTEC, Inc., Cambridge Isotope Labs) or unlabelled (Sigma-Aldrich) nucleotide triphosphates, and synthetic DNA templates (Integrated DNA Technologies, Inc.) containing the T7 promoter and sequence of interest. All RNAs were purified by 20% (w/v) denaturing polyacrylamide gel electrophoresis, using 8 M urea and 1× TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA). The RNA was electro-eluted from the gel in 20 mM Tris pH 8 buffer followed by ethanol precipitation. The RNA pellet was dissolved in water, annealed by heating to 95 °C for 10 min and rapid cooling on ice and exchanged into NMR buffer (15 mM sodium phosphate, 0.1 mM EDTA, and 25 mM NaCl at pH 6.4) multiple times using an Amicon Ultra-4 Centrifugal Filter Unit (Millipore Corp.). Unlabelled RNA samples (TAR(C90U), TAR(A35-MDA), TAR(A35G), A-site(AU95), A-site(U95-C31)) were purchased from Dharmaco (Thermo Fisher Scientific) and Integrated DNA Technologies and dissolved in NMR buffer (15 mM sodium phosphate, 0.1 mM EDTA, 25 mM NaCl) at pH 6.4. The TR pH studies employed the following NMR buffers: pH 7.4 (15 mM sodium phosphate, 0.1 mM EDTA and 25 mM NaCl) and pH 4.6 (15 mM acetate-d$_3$, 0.1 mM EDTA, and 25 mM NaCl). Resonance assignments of wild-type RNA samples were obtained from prior studies$^{49,50,60}$ and confirmed using standard resonance assignment experiments.

Carbon $R_{1p}$ relaxation dispersion. All relaxation dispersion NMR experiments were performed on a Bruker Avance 600 MHz NMR spectrometer equipped with a 5-mm triple-resonance cryogenic probe. Experiments were performed at 25 °C, 25°/15 °C, and 15 °C for TAR, A-site and SL1, respectively, using uniformly $^{13}C^{15}$N-labelled RNA constructs shown in Supplementary Fig. 1. For TAR, we used a second construct lacking the bulge (EII-TAR, Supplementary Fig. 1) to measure dispersion data for U31-C6 resonance, which is otherwise overlapped. For A-site, all data were measured at 25 °C with the exception of A92-C2 and A93-C8, which was measured at 15 °C to push the system into slower exchange and obtain more reliable data. Frame carbon $R_{1p}$ relaxation dispersion data were measured using a 1D acquisition scheme that extends the sensitivity to chemical exchange into millisecond timescales relative to conventional 2D relaxation dispersion methods$^{51,52}$.

On- and off-resonance relaxation dispersion data were recorded at various offset frequencies ($2$) and spinlock powers ($v$) (see Supplementary Table 2). The following relaxation delays were used: TAR: C90-U31, U95-C31, G34 and A35 C1 [0, 6 (2 $k_3$)]; G33 C1 [0, 12 (2 $k_3$)]; A32 C8, U34C6 A32 C2, G29 C1 [0, 6, 8 (2 $k_3$)]; G29 C8 [0, 6, 3.3 (2 $k_3$)]; A27 C2 [0, 4, 10, 17, 25 (2 $k_3$)]; G26 C8 [0, 12 (2 $k_3$)]; C8, A25-C8 and A25-C2 were combined in a global fit to characterize ES1, and G26-C8, C30-C6 and G7-C8 were included in a global fit to characterize ES2. The free energy difference between the GS and ES ($\Delta G_{ES}$) and between the GS and transition state ($\Delta G_{ES}^{t}$) with other data measured at 25 °C, 15 °C and 3 °C. The same $K_{ex}$ and $K_{rel}$ were used for dispersion profiles with $R_{ex} < 5$ Hz or that yielded ambiguous signs for $\Delta G$ during the Monte Carlo error analysis were deemed unreliable (these include A-site: G05-C8, G06-C8, G07-C8, G08-C8). The sign of $\Delta G_{ES}$ was confirmed using the $\chi^2$-test and the remaining resonances (G05-C8, A92-C1, U95-C8, C96-C8) can be included into the global fit without affecting the resulting fitted parameters (values are true when globally fitting four or seven resonances). A-site data measured at 15 °C were fitted individually although similar $\Delta G_{ES}$ values were obtained when these data were included in global fits with other data measured at 25 °C. The $\Delta G_{ES}$ values obtained from both individual and global fits are shown for A92-C2 in Supplementary Fig. 5. For SL1, the G26-C8, A25-C8 and A25-C2 were combined in a global fit to characterize ES1, and G31-C8, C30-C6 and G7-C8 were combined in a global fit to characterize ES2.

The free energy difference between the GS and ES ($\Delta G^{ES}$) and between the GS and transition state ($\Delta G^{ES}^{t}$) was computed using (with $\Delta G^{CS} = 0$):

$$\Delta G^{ES} = -\ln \left( \frac{k_{h1} R T}{k_{b1}} \right) - \left( - \ln \frac{k_{b1} R T}{k_{h1}} \right)$$

$$\Delta G^{ES}^{t} = -\ln \left( \frac{k_{h1} R T}{k_{b1}} \right)$$

where $k_{h1}$, $k_{b1}$ are rate constants, $h$ is Planck’s constant, $k_{b}$ is Boltzmann’s constant, $R$ is the gas constant and $T$ is temperature.

SL1 isomerization assay. SL1 isomerization assays were performed closely following the procedure described previously$^{63}$. Briefly, SL1 RNA samples (SL1, SL1(G8C), SL1(tGC), SL1(eGC)) (Fig. 1e and Supplementary Fig. 8) containing the wild-type apical loop were purchased from Integrated DNA Technologies, Inc. SL1 RNA samples were dissolved to a concentration of 5 μM, heated to 95 °C for 3 min and placed on ice for 30 min. Subsequently, 50% (v/v) 2× dimerization buffer (20 mM sodium phosphate, pH 6.4, 100 mM NaCl, and 0.2 mM MgCl2) was added to produce a final RNA concentration of 2.5 μM, and the sample incubated at 55 °C or on ice for a variable amount of time. Native gels
were run using TBE buffer and control with TBM (TBE with no EDTA but 10 mM MgCl₂) as previously described⁴⁰ and detected with ethidium bromide staining. **MC-fold predictions of RNA secondary structure.** All RNA secondary structures were predicted based on sequence using the program MC-Fold²⁴ (http://www.major.irc.ca/MC-Fold/) with standard input options.

⁴⁹. Fourmy, D., Yoshizawa, S. & Puglisi, J. D. Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA. *J. Mol. Biol.* **277**, 333–345 (1998).

⁵⁰. Delaglio, F. et al. Nmrpipe—a multidimensional spectral processing system based on Unix Pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

⁵¹. Spyracopoulos, L. A suite of Mathematica notebooks for the analysis of protein main chain 15N NMR relaxation data. *J. Biomol. NMR* **36**, 215–224 (2006).

⁵². Meinhold, D. W. & Wright, P. E. Measurement of protein unfolding/refolding kinetics and structural characterization of hidden intermediates by NMR relaxation dispersion. *Proc. Natl Acad. Sci. USA* **108**, 9078–9083 (2011).

⁵³. Vallurupalli, P., Bouvignes, G. & Kay, L. E. Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR. *J. Phys. Chem. B* **115**, 14891–14900 (2011).