RNA Recognition

Evidence that Electrostatic Interactions Dictate the Ligand-Induced Arrest of RNA Global Flexibility**

*Stephen W. Pitt, Qi Zhang, Dinshaw J. Patel, and Hashim M. Al-Hashimi**

Arresting flexibility to prevent RNA from undergoing functionally important conformational transitions is an established strategy for developing RNA-targeting therapeutics.[1] To understand the role of RNA flexibility in adaptive recognition[2] is also important for the rational design of small molecules that bind their RNA target with high affinity and specificity.[3] Yet few studies have quantitatively examined how RNA-binding therapeutics affect the flexibility of their RNA targets. As a result, little is known about the RNA–ligand interactions that are important for arresting different types of RNA functional flexibility. Here, we provide direct evidence by using NMR residual dipolar couplings (RDCs)[4] that electrostatic interactions play a primary role in dictating the degree to which small molecules arrest global motions in RNA.

The transactivation response element (TAR) RNA from HIV-1 (Figure 1a) is a primary RNA target for developing anti-HIV therapeutics[5] and its molecular flexibility is implicated in its function. Upon binding to its cognate target, the transactivator protein (Tat), TAR undergoes large conformational changes that involve reorientation of two helical domains from a bent to a coaxially aligned state.[6] By using RDCs, we previously reported evidence for significant interdomain flexibility in free TAR which may play a direct role in mediating the conformational changes in TAR that
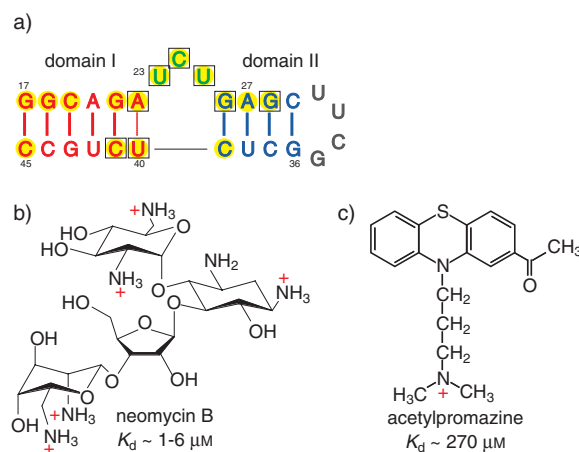
[*] Q. Zhang, Prof. H. M. Al-Hashimi
Department of Chemistry and Biophysics Research Division
University of Michigan
Ann Arbor, MI 48109 (USA)
Fax: (+1) 734-647-4865
E-mail: hashimi@umich.edu

Dr. S. W. Pitt
Department of Pharmacology
Weil Medical College of Cornell University
New York, NY 10021 (USA)

Prof. D. J. Patel
Cellular Biochemistry and Biophysics Program
Memorial Sloan-Kettering Cancer Center
1275 York Avenue, New York, NY 10021 (USA)

Figure 1. a) Secondary structure of TAR (transactivation response element) RNA in which the wild-type loop is replaced by a UUCG counterpart. The highlighted residues undergo $\Delta\delta > 0.1$ ppm changes in ¹H NMR chemical shifts upon binding to neomycin B (yellow) and acetylpromazine (□). b, c) Structures and TAR-dissociation constants (K_d) for neomycin B (b; NeoB)[9, 11] and acetylpromazine (c; AcP).[14]

accompany Tat recognition.[7a,b] We also demonstrated that Mg^{2+} or the ligand mimic of Tat, argininamide, can bind TAR and completely arrest these interdomain motions.[7c,d]

The RDC methodology and TAR provide a unique opportunity to examine RNA–ligand interactions that are particularly important for arresting a general class of RNA motions, which involve global reorientation of helical domains. Numerous studies on RNA–aminoglycoside recognition have demonstrated that electrostatic interactions between cationic groups on the ligand and regions of high negative-charge density in the RNA can contribute significantly to aminoglycoside-binding affinities and specificities.[8] As global motions in RNA originate from local mobility in the negatively charged backbone and as functional groups from many residues can be involved in forming pockets of high negative-charge density in RNA, we reasoned that similar electrostatic interactions are also likely to be important for arresting RNA global motions.

To examine this hypothesis, we employed RDC NMR methodology to characterize the conformational dynamics of TAR when it is bound to two molecules that have different electrostatic charges: the aminoglycoside neomycin B (NeoB),[9] which bears five positively charged ammonium groups, and the small organic molecule acetylpromazine (AcP),[10] which carries a single positive charge (Figure 1 b,c). Both NeoB and AcP bind TAR and inhibit its interaction with Tat.[9,10] A previously reported NOE-based NMR spectroscopic study of the structure of the TAR–NeoB complex[11] showed that NeoB binds TAR in the minor groove and stabilizes a coaxially stacked conformation similar to that observed for TAR bound to Tat peptides and Arg (arginine).[6a,b,d] In contrast, the NMR-determined structure of the TAR–AcP complex indicates that AcP induces minor changes in the conformation of TAR and that it binds a cavity at the interdomain interface.[12] Both NMR studies did not report on the global dynamics of TAR in the ligand-bound state.

Shown in Figure 2 are the RDC-derived best-fit generalized degrees of order (ϑ)[13] determined for each domain in the two TAR complexes. The generalized degree of order describes the degree of alignment experienced by each domain which should be identical if the domains are held rigid relative to one another.[13] Results are shown when different input domain structures were used in the order tensor calculations, including idealized A-form helices generated by using Insight II (Molecular Simulations Inc.) and the previous NOE-based NMR spectroscopically determined structures of TAR–NeoB (pdb 1QD3)[11] and TAR–AcP (pdb 1LVJ).[12] Also shown are the values of the root-mean-square deviation (rmsd) between measured and back-calculated RDC values which provide a measure of agreement between a domain structure and measured RDC values.

In the structure determined from NMR spectroscopic studies, NeoB induces local distortions in domain I.[11] Accordingly, most NMR models for domain I in TAR–NeoB yield rmsd values that are lower than the idealized A-form helix (Figure 2a). More importantly, the ϑ value determined for the two domains in TAR–NeoB are similar, particularly for domains that yield the lowest rmsd values (Figure 2a). The best-fit ϑ_{int} value $(\vartheta_{domain\,I}/\vartheta_{domain\,II})$,[13] which provides a measure of interdomain motional amplitudes, is close to unity $(\vartheta_{int} = 0.94)$ and thus indicates that NeoB arrests interdomain motions in TAR (Figure 2c). Remarkably, this is not the case for TAR bound to AcP. As previously observed for free TAR,[7a] the ϑ value for domain I is significantly smaller than that for domain II (Figure 2b), and the best-fit ϑ_{int} value (0.69) for TAR–AcP is only slightly higher than that previously reported for free TAR (0.59; Figure 2c). The low ϑ_{int} value for TAR–AcP cannot be attributed to errors in the input domain structures because the best-fit domains yield rmsd values that are equal to or smaller than the experimental

uncertainty in the RDC values (see Figure 2b and Supporting Information) and because low ϑ_{int} values are observed independent of both input domain structure (Figure 2b) and RDCs (see Supporting Information). The low ϑ_{int} value for TAR–AcP could arise from fast exchange between free and AcP-bound TAR states. On the basis of the concentrations of TAR (≈ 1 mM) and AcP (≈ 2 mM) used in the RDC measurements and the binding constant (K_a) of $K_a \approx 270$ μM,[14] approximately 81 % of TAR is computed to be in the bound state. If a population-weighted average of free and bound ϑ_{int} values is assumed (although, strictly, averages over RDCs should be considered), then a similar population-corrected ϑ_{int} value for TAR–AcP of 0.71 is obtained. In agreement with this analysis, an ϑ_{int} value of 0.71 was determined for molar ratios of about 1:5 for TAR/AcP in which the bound state is approximately 94 % populated (data not shown). Rather, the low ϑ_{int} value determined for TAR–AcP argues that unlike NeoB, AcP only marginally arrests global motions in TAR.

In Figure 2d, we compare the average domain–domain orientations determined for TAR–NeoB and TAR–AcP by using RDCs with counterparts determined previously for free TAR[7a] and TAR bound to Mg^{2+}[7c] and Arg.[7d] The more coaxially stacked TAR–NeoB conformation is similar to those of TAR–Arg and TAR–Mg, whereas TAR–AcP is very similar to free TAR. The RDC-derived global conformations for TAR–NeoB and TAR–AcP are also in excellent agreement with the corresponding results of NOE-based structural NMR studies (Figure 2e).[11,12]

Our results allude to a correlation between the average domain–domain orientation of TAR and dynamics with bent interdomain alignments being more globally flexible than their coaxially stacked counterparts (Figure 2c,d). Furthermore, all of the ligands that arrest global motions in TAR
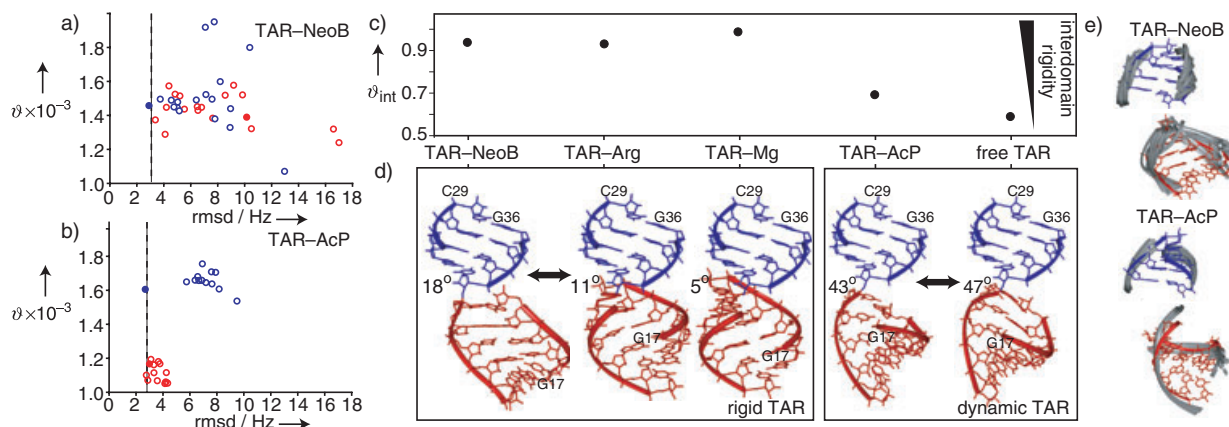

Figure 2. The global structural dynamics of TAR–NeoB and TAR–AcP. The best-fit generalized degree of order (ϑ) for domains I and II in a) TAR–NeoB and b) TAR–AcP as a function of the root-mean-square deviation (rmsd) between measured residual dipolar couplings (RDCs) and values back-calculated by using the best-fit order tensor (red unfilled circle: domain I NMR models; red filled circle: domain I idealized; blue unfilled circle: domain II NMR models; blue filled circle: domain II idealized). The experimental uncertainty (rmsd) in RDCs (see Supporting Information) is indicated by a dashed vertical line. c) The internal generalized degree of order, $\vartheta_{int} = \vartheta_{domain\,I}/\vartheta_{domain\,II}$, determined for TAR–NeoB and TAR–AcP and a comparison with previous values determined for free TAR,[7a] TAR–Mg,[7c] and TAR–Arg.[7d] d) The best-fit relative domain orientation in TAR–NeoB and TAR–AcP determined by superimposing order tensor frames and a comparison with free TAR,[7a] TAR—Mg,[7c] and TAR–Arg.[7d] Domain II (blue) is superimposed and the best-fit interhelical angle is shown next to each structure. The residues for the bulge are not shown. The estimated uncertainty in the interhelical angle is less than ±6°. e) Comparison of the RDC-derived interdomain alignment (red and blue) with structures as determined by NOE-based NMR spectroscopy (in gray) of TAR–NeoB (pdb 1QD3) and TAR–AcP (pdb 1LVJ).

(NeoB, Arg, and Mg²⁺) are known to be involved in intermolecular electrostatic interactions. In the structure determined from NMR studies,[11] the five positively charged amino groups in NeoB are in close proximity to many backbone phosphate groups of TAR and other electronegative groups that belong to domains I and II and the bulge (U-C-U). The positively charged guanidinium group of Arg interacts with phosphate groups in domain I (A22) and the bulge (U23) as well as basic electronegative groups in domain II (G26).[6a,b,d] Recent studies indicate that two Arg ligands may be involved in electrostatic interactions with two regions of negative-charge density in TAR.[15] The crystal structure of TAR in the presence of Mg²⁺ and Ca²⁺ shows a network of inner- and outer-sphere interactions between three divalent ions and functional groups that span domains I (A22) and II (G26, A27, and G28) as well as the bulge (U23, C24, U25).[6e] In sharp contrast, the structure determined from NMR studies of TAR–AcP indicates that the RNA–ligand intermolecular contacts primarily involve stacking and hydrophobic interactions between the three-membered ring and base moieties in the bulge and neighboring residues.[12] There is also evidence that the aliphatic side chain that harbors the single positive charge is flexible and protrudes in and out of the TAR binding pocket.[14]

Our results can therefore be rationalized in part on the basis of electrostatic interactions between negatively charged pockets in TAR that can be formed by backbone phosphate groups as well as other sugar/base electronegative groups and cationic groups in the ligand. Such interactions can uniquely stabilize the negatively charged backbone that is responsible for activating global motions in TAR and/or allow simultaneous clamping interactions with electronegative groups that belong to residues that span the interdomain interface. In TAR, formation of such negatively charged pockets appears to require a degree of coaxial stacking which likely serves to bring backbone phosphates and other electronegative groups in the bulge and neighboring residues into spatial proximity.[15] In this manner, electrostatic interactions may simultaneously arrest global motions and stabilize coaxially aligned conformations of TAR.

To our knowledge, the results reported here for TAR–AcP represent the first example in which global motions are observed in an RNA complex. This illustrates how ligand binding in itself is not sufficient for arresting RNA global motions, even if the RNA-binding site is at a critical junction that intersperses helical domains. Our results also indicate that electrostatic interactions offer one approach for stabilizing the global alignment of RNA domains that are separated by flexible linkers. The RDC approach presented in this work provides insight into such dynamical features underlying RNA–small-molecule recognition that are critical to rational drug design and that are not available from static three-dimensional structures.

Experimental Section

Uniformly ¹⁵N/¹³C-labeled TAR was prepared by in vitro transcription. Samples for NMR studies contained TAR (≈ 1.0 mM), sodium phosphate (15 mM), sodium sulfate (25 mM), and EDTA (ethylenediaminetetraacetic acid, 0.1 mM) at pH 6.0–6.2. The samples of TAR–NeoB and TAR–AcP also contained NeoB (2 mM, MP Biomedical Inc.) or AcP (2 mM, Research Diagnostics), respectively. Identical samples that contained Pf1 phage (≈ 25 mg mL⁻¹, ASLA Ltd) were also prepared for NMR studies.[16] NMR spectroscopy experiments were recorded at 600 MHz at 25 °C. One-bond RDCs for C1′–H1′, C2–H2, C5–H5, C6–H6, C8–H8, N1–H1, and N3–H3 were measured twice by using experiments that yield splittings along either the direct (¹H) and indirect (¹³C/¹⁵N) dimension, as previously described.[7d] Average RDC values were used whenever possible. Order tensors were computed using ORDERTEN SVD[17a] and REDCAT.[17b] A total of 20/18 and 21/19 RDCs were measured in domains I/II in TAR–NeoB and TAR–AcP, respectively. In all cases, RDCs from terminal residues G17 and C45 in domain I were omitted from analysis to avoid complications resulting from end-fraying effects. For NMR models of TAR–AcP, five RDC interactions from G18 and C44 had to be omitted owing to differences between the secondary structure of domain I in the two TAR constructs.

[1] a) B. T. Porse, I. Leviev, A. S. Mankin, R. A. Garrett, *J. Mol. Biol.* **1998**, *276*, 391–404; b) T. Pape, W. Wintermeyer, M. V. Rodnina, *Nat. Struct. Biol.* **2000**, *7*, 104–107.

[2] a) J. R. Williamson, *Nat. Struct. Biol.* **2000**, *7*, 834–837; b) N. Leulliot, G. Varani, *Biochemistry* **2001**, *40*, 7947–7956.

[3] a) T. Hermann, *Biochimie* **2002**, *84*, 869–875; b) Q. Vicens, E. Westhof, *ChemBioChem* **2003**, *4*, 1018–1023.

[4] a) J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279–9283; b) N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111–1114.

[5] a) A. Krebs, V. Ludwig, O. Boden, M. W. Gobel, *ChemBioChem* **2003**, *4*, 972–978; b) M. Froeyen, P. Herdewijn, *Curr. Top. Med. Chem.* **2002**, *2*, 1123–1145.

[6] a) J. D. Puglisi, R. Tan, B. J. Calnan, A. D. Frankel, J. R. Williamson, *Science* **1992**, *257*, 76–80; b) F. Aboul-ela, J. Karn, G. Varani, *J. Mol. Biol.* **1995**, *253*, 313–332; c) F. Aboul-ela, J. Karn, G. Varani, *Nucleic Acids Res.* **1996**, *24*, 3974–3981; d) A. S. Brodsky, J. R. Williamson, *J. Mol. Biol.* **1997**, *267*, 624–639; e) J. A. Ippolito, T. A. Steitz, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9819–9824.

[7] a) H. M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, D. J. Patel, *J. Mol. Biol.* **2002**, *315*, 95–102; b) Q. Zhang, R. Throolin, S. W. Pitt, A. Serganov, H. M. Al-Hashimi, *J. Am. Chem. Soc.* **2003**, *125*, 10530–10531; c) H. M. Al-Hashimi, S. W. Pitt, A. Majumdar, W. Xu, D. J. Patel, *J. Mol. Biol.* **2003**, *329*, 867–873; d) S. W. Pitt, A. Majumdar, A. Serganov, D. J. Patel, H. M. Al-Hashimi, *J. Mol. Biol.* **2004**, *338*, 7–16.

[8] a) Y. Tor, T. Hermann, E. Westhof, *Chem. Biol.* **1998**, *5*, 277–283; b) T. Hermann, E. Westhof, *J. Mol. Biol.* **1998**, *276*, 903–912; c) T. Hermann, E. Westhof, *J. Med. Chem.* **1999**, *42*, 1250–1261; d) Y. Tor, *ChemBioChem* **2003**, *4*, 998–1007.

[9] a) S. Wang, P. W. Huber, M. Cui, A. W. Czarnik, H. Y. Mei, *Biochemistry* **1998**, *37*, 5549–5557; H.-Y. Mei, A. A. Galan, N. S. Halim, D. P. Mack, D. W. Moreland, K. B. Sanders, H. N. Truong, A. W. Czarnik, *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2755–2760.

[10] a) A. V. Filikov, V. Mohan, T. A. Vickers, R. H. Griffey, P. D. Cook, R. A. Abagyan, T. L. James, *J. Comput.-Aided Mol. Des.* **2000**, *14*, 593–610; b) K. E. Lind, Z. Du, K. Fujinaga, B. M. Peterlin, T. L. James, *Chem. Biol.* **2002**, *9*, 185–193.

[11] C. Faber, H. Sticht, K. Schweimer, P. Rösch, *J. Biol. Chem.* **2000**, *275*, 20660–20666.

[12] Z. Du, K. E. Lind, T. L. James, *Chem. Biol.* **2002**, *9*, 707–712.

[13] J. R. Tolman, H. M. Al-Hashimi, L. E. Kay, J. H. Prestegard, *J. Am. Chem. Soc.* **2001**, *123*, 1416–1424.

[14] M. Mayer, T. L. James, *J. Am. Chem. Soc.* **2004**, *126*, 4453–4460.

[15] B. Davis, M. Afshar, G. Varani, A. I. Murchie, J. Karn, G. Lentzen, M. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. Swarbrick, F. Aboul-ela, *J. Mol. Biol.* **2004**, *336*, 343–356.

[16] a) M. R. Hansen, L. Mueller, A. Pardi, *Nat. Struct. Biol.* **1998**, *5*, 1065–1074; b) G. M. Clore, M. R. Starich, A. M. Gronenborn, *J. Am. Chem. Soc.* **1998**, *120*, 10571–10572.

[17] a) J. A. Losonczi, M. Andrec, M. W. F. Fischer, J. H. Prestegard, *J. Magn. Reson.* **1999**, *138*, 334–342; b) H. Valafar, J. H. Prestegard, *J. Magn. Reson.* **2004**, *167*, 228–241.