J Biomol NMR (2011) 51:347–356
DOI 10.1007/s10858-011-9573-6

ARTICLE

Interactions of protein side chains with RNA defined with REDOR solid state NMR

Wei Huang · Gabriele Varani · Gary P. Drobny

Received: 29 March 2011 / Accepted: 11 August 2011 / Published online: 25 September 2011

Abstract Formation of the complex between human immunodeficiency virus type-1 Tat protein and the transactivation response region (TAR) RNA is vital for transcriptional elongation, yet the structure of the Tat-TAR complex remains to be established. The NMR structures of free TAR, and TAR bound to Tat-derived peptides have been obtained by solution NMR, but only a small number of intermolecular NOEs could be identified unambiguously, preventing the determination of a complete structure. Here we show that a combination of multiple solid state NMR REDOR experiments can be used to obtain multiple distance constraints from ^{15}N to ^{13}C spins within the backbone and side chain guanidinium groups of arginine in a Tat-derived peptide, using ^{19}F spins incorporated into the base of U23 in TAR and ^{31}P spins in the P22 and P23 phosphate groups. Distances between the side chain of Arg52 and the base and phosphodiester backbone near U23 measured by REDOR NMR are comparable to distances observed in solution NMR–derived structural models, indicating that interactions of TAR RNA with key amino acid side chains in Tat are the same in the amorphous solid state as in solution. This method is generally applicable to other protein-RNA complexes where crystallization or solution NMR has failed to provide high resolution structural information.

Keywords Solid state NMR · REDOR · RNA · Tat-TAR

W. Huang · G. Varani · G. P. Drobny (✉)
Department of Chemistry, University of Washington,
Box 351700, Seattle, Washington, DC 98195, USA
e-mail: drobny@chem.washington.edu

G. Varani
Department of Biochemistry, University of Washington,
Box 357350, Seattle, Washington, DC 98195, USA

Introduction

The human immunodeficiency virus type-1 (HIV-1) Tat protein binds specifically to the transactivation response region (TAR) RNA stem-loop structure, located at the 5′ end of the HIV-1 mRNA, and stimulates transcriptional elongation (Dingwall et al. 1989, 1990; Roy et al. 1990a, b; Calnan et al. 1991a, b; Muesing et al. 1987; Harper and Logsdon 1991; Delling et al. 1992; Churcher et al. 1993). The secondary structure of TAR (Fig. 1) contains a three-nucleotide bulge (UCU) linking two stem regions and a loop of six nucleotides where critical cellular co-factors bind. U23 in the bulge and two base-pairs in the stem above, G26·C39 and A27·U38, are required for the specific recognition by Tat protein (Puglisi et al. 1992; Aboul-ela et al. 1995; Aboul-ela et al. 1996; Brodsky and Williamson 1997; Long and Crothers 1999; Puglisi et al.1993; Tao et al. 1997). Several arginine side chains within the basic region of Tat have been shown to be essential for this process (Roy et al. 1990a, b; Calnan et al. 1991a, b; Weeks et al. 1990; Cordingley et al. 1990). A conformational change in TAR occurs upon the binding of Tat, and makes functional groups in the major groove more accessible to the protein, including the critical phosphates between G21 and A22 (P22) and A22 and U23 (P23) (Churcher et al. 1993; Aboul-ela et al. 1995; Calnan et al. 1991a, b; Hamy et al. 1993; Pritchard et al. 1994; Tao and Frankel 1992).

The interaction of Tat with TAR has been widely studied because it is critical for viral replication and therefore could be targeted by new antivirals (Aboul-ela et al. 1995; Brodsky and Williamson 1997; Karn 1999; Rana and Jeang 1999; Davidson et al. 2009). In the past 15 years, considerable effort has been put into the synthesis and evaluation of small molecules and peptidic inhibitors of the Tat-TAR interaction (Gallego and Varani 2001; Huq

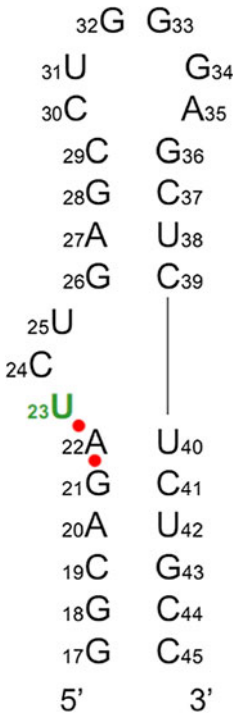

Fig. 1 Secondary structure of the 29-mer TAR RNA construct used in the present work; nucleotides and phosphate groups involved in Tat recognition and studied in the present work are highlighted in *color*

et al. 1999; Murchie et al. 2004; Baba 2006), but none have advanced to preclinical studies and many unknowns still remain. One of the difficulties in the design of inhibitors of the Tat-TAR interaction is the fact that the detailed nature of RNA–protein interactions within the Tat-TAR complex is not fully elucidated. Early studies of TAR complexes with Tat peptides aimed at deriving models for specific RNA–protein interactions. For example, using ethylation interference techniques, Calnan et al. (1991a, b) proposed an "arginine fork" model to describe the network of hydrogen bonds formed between an arginine guanidinium group in Tat and pairs of adjacent phosphate groups in the bulged loop of TAR RNA. Using solution NMR to study the complex of argininamide with TAR RNA, Puglisi et al. (1992) observed NOEs from argininamide (δ) protons to protons on A22, U23 and A27, further suggesting the close proximity between the arginine guanidinium group and the bulged loop region of TAR RNA. Aboul-ela et al. (1995) applied multi-dimensional NMR techniques to uniformly ^{15}N and ^{13}C labeled TAR RNA in complexes with argininamide and a Tat-derived 37-mer peptide, and determined the structure of TAR in the bound form. However, insufficient intermolecular NOEs were observed that would have allowed establishing the exact nature of TAR-Tat interactions. Thus, this very important structure remains to be determined.

In principle, solid-state NMR (ssNMR) dipolar recoupling techniques could be used to address the interaction between peptides and RNA. Since the distance range accessible to experiments such as REDOR (Gullion and Schaefer 1989a, b) generally exceeds that of nuclear Overhauser effects (NOEs), many structural features obtained indirectly by short range distance NOE measurements can be directly observed using REDOR. To prove the potential of these techniques, Jehle et al. (2010) have shown that solid-state NMR spectroscopy can be used to probe intermolecular interactions at other protein-RNA interfaces. Distances between the amide ^{15}N nuclei of the protein backbone and the ^{31}P nuclei in the RNA backbone of an RNA–protein complex were measured with TEDOR experiments, and the accuracy of the TEDOR distance measurements was demonstrated by comparison to the crystal structure.

In this paper, the 29 nucleotide TAR RNA construct of Fig. 1 was bound to an 11 amino acid peptide 47YGRKKRRQRRR57 corresponding to the arginine-rich 47–57 region of the Tat protein that provides direct contacts with the TAR bulge region upon binding (Roy et al. 1990a, b; Calnan et al. 1991a, b; Weeks et al. 1990; Cordingley et al. 1990). It has been shown that the conformational change induced in TAR by the binding of this peptide mimics that of the full Tat protein (Dingwall et al. 1989, 1990; Roy et al. 1990a, b; Calnan et al. 1991a, b; Churcher et al. 1993; Puglisi et al. 1993; Weeks et al. 1990, Cordingley et al. 1990; Tao and Frankel 1992; Loret et al. 1992; Olsen et al. 2005; Weeks and Crothers 1991; Davis et al. 2004; Sumner-Smith et al. 1991; Weeks and Crothers 1992). Here we use a suite of solid state NMR dipolar recoupling experiments to determine multiple distances from ^{13}C to ^{15}N nuclei in the guanidinium group of the side chain of Arg52 to the P22 and P23 phosphate groups and the base of U23. We use a combination of REDOR experiments including ^{13}C{^{31}P}, ^{13}C{^{19}F}, ^{15}N{^{31}P}, and ^{15}N{^{19}F}, while the conformational change of TAR itself upon binding to Tat peptide is further assayed using ^{31}P{^{19}F} REDOR. A structural model of the interaction between the side chain of Arg52 with the phosphodiester backbone of the bulged loop of TAR is obtained from the protein-RNA distances obtained by quantitative simulations of the REDOR dephasing curves. The degree to which solid state NMR-derived model of Tat-TAR interactions conforms qualitatively and quantitatively to comparable models obtained previously with solution NMR is discussed.

Materials and methods

Sample preparation

The TAR RNA 29mer 5′-GGCAGA-U[5F]-CUGA*GC CUGGGAGCU(pS)CUC-U[5F]-GCC-3′ (U[5F] indicates a 5-F base-labeled Uridine, A* indicates a 1′ deuterium

substitution, and pS shows a phosphorothioate label) was purchased from Dharmacon. Previous work has demonstrated that phosphorothioate and 5-fluorine substitutions at those positions do not perturb the structure of TAR (Gonzalez et al. 1995; Bachelin et al. 1998; Merritt et al. 1999; Puffer et al. 2009). The A* and pS labels are used for purposes other than those described in this work. The RNA oligonucleotide was converted to the 2′-hydroxyl form and desalted by the manufacturer, and no further purification was performed. The sample was checked for homogeneity by analytical denaturing polyacrylamide gel electrophoresis. 4.04 μmol sample was dissolved in buffer (50 mM NaCl, 10 mM sodium cacodylate, pH 5.5), then frozen by liquid nitrogen and lyophilized. The final sample contains 10% NaCl and 4.7% cacodylate, respectively, by weight upon lyophilization.

Peptide synthesis

The 11-mer tat peptide 47YGRKKRRQRRR57 was synthesized on a third generation Rainin PS3 solid phase peptide synthesizer using standard Fmoc chemistry. Arg52 was uniformly ^{13}C and ^{15}N labeled and obtained in protected form from Isotec (Sigma-Aldrich). The sample was cleaved and deprotected by stirring in a 95% TFA, 2.5% triisopropylsilane and 2.5% water mixture for 2.5 h. The resin was filtered from the product, and TFA/product solution was reduced to 2 ml. The peptide was precipitated by dripping the TFA/product solution into cold tert-butyl methyl ether, and collected by centrifugation. The sample was rinsed three times by cold tert-butyl methyl ether and recentrifuged, then dried under vacuum for 2 days. The crude peptide was purified by HPLC. The purity was confirmed using mass spectrometry and no further purification was required.

Complex formation

The lyophilized TAR RNA sample was re-dissolved in 1,530 μl sterile water. 4.85 μmol (1.2 equivalents) of purified 11mer Tat peptide were needed to ensure the saturation of this complex (Bardaro et al. 2009). The peptide was dissolved in 1,530 μl of the same buffer (50 mM NaCl, 10 mM sodium cacodylate, pH 5.5). Both solutions were heated at 37°C for 10 min, then the peptide solution was added to the TAR RNA solution drop by drop to avoid the formation of any precipitate. The final mixture was kept at 37°C for 50 min with continued gentle vortexing. The solution was cooled to room temperature, then frozen by liquid nitrogen, lyophilized, and transferred to the MAS rotor.

Solid-state NMR experiments

Experiments were performed on a home-built spectrometer at a 11.74 T field operating at Larmor frequencies of 500.00 MHz for ^{1}H, 470.47 MHz for ^{19}F, 202.53 MHz for ^{31}P, 125.76 MHz for ^{13}C and 50.98 MHz for ^{15}N using a 4 mm HFX and HXY triple-resonance Varian/Chemagnetics magic angle spinning probe. All measurements were performed at −32°C with a sample spinning speed of 6,000 Hz which was regulated to ± 2 Hz, except for ^{15}N–^{19}F REDOR which was done at 8,000 Hz spinning rate. REDOR experiments were performed using XY-8 phase cycling with alternating π-pulses on both the observed and dephasing channels (Fig. 2; here ^{13}C or ^{15}N is the observed channel, ^{19}F or ^{31}P is the dephasing channel).

In ^{13}C–^{19}F REDOR, the ^{13}C NMR signal was enhanced using ramped cross-polarization with a ^{1}H–^{13}C contact time of 2.6 ms. Pulse lengths were 3.6 μs for ^{1}H $\pi/2$, 6.0 μs for ^{13}C π pulses, and 8.2 μs for ^{19}F π. REDOR points were collected at a dephasing time of 1.33 ms with 20 k scans, 2.67 ms with 30 k scans, 4.00 ms with 36 k scans, and 5.33 ms with 40 k scans (23–45 h each point, total experiment time 140 h). In ^{13}C–^{31}P REDOR, the ^{13}C NMR signal was enhanced using ramped cross-polarization with a ^{1}H–^{13}C contact time of 2.0 ms. Pulse lengths were 2.8 μs for ^{1}H $\pi/2$, 4.0 μs for ^{13}C π, and 5.5 μs for ^{31}P π. REDOR points were collected at a dephasing time of 1.33 ms with 16 k scans, 4.00 ms with 22 k scans, 6.67 ms with 34 k scans, and 8.00 ms with 34 k scans (18–38 h each point, total experiment time 118 h). In ^{15}N–^{19}F REDOR, the ^{15}N NMR signal was enhanced using ramped cross-polarization with a ^{1}H–^{15}N contact time of 2.6 ms. Pulse lengths were 3.4 μs for ^{1}H $\pi/2$, 6.0 μs for ^{15}N π and 6.5 μs for ^{19}F π. REDOR points were collected at dephasing times of 1 ms with 22 k scans, 3 ms with 26 k scans, 5 ms with 50 k scans and 6 ms with 56 k scans (25–63 h each point, total experiment time 172 h). In ^{15}N–^{31}P REDOR, the ^{15}N NMR

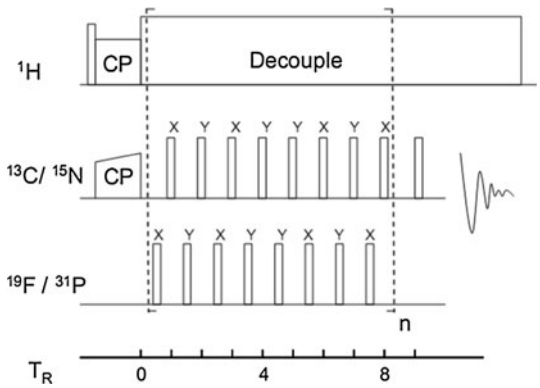

Fig. 2 XY-8 phase cycling REDOR pulse sequence with alternating π-pulses on both the observed and dephasing channels

signal was enhanced using ramped cross-polarization with a ^{1}H–^{15}N contact time of 2.0 ms. Pulse lengths were 2.8 μs for ^{1}H $\pi/2$, 6.5 μs for ^{15}N π, and 7.0 μs for ^{31}P π. REDOR points were collected at a dephasing time of 1.33 ms with 20 k scans, 5.33 ms with 30 k scans, 6.67 ms with 50 k scans, and 8.00 ms with 60 k scans (23–67 h each point, total experiment time 178 h).

Data processing

Data processing was performed using in-house written NMR processing software. All experimental REDOR data were simulated using the simulation software SIMPSON (Bak et al. 2000), with the assumption of ideal cross polarization and proton decoupling, and without considering relaxation. Powder averaging used a minimal set of 232 $\{\alpha, \beta\}$ Euler angles defined by Zaremba-Conroy-Wolfsberg scheme with 10 γ angles. The required experimental NMR parameters are described above. A χ^2 analysis was used to determine the best fit distances. The polynomial fittings in χ^2 plots (red lines) are used to calculate the inverse of the 2nd derivative of χ^2 at the minimum. The error bounds of distances can be estimated by $\sigma^2 = 2\left(\frac{\partial^2\chi^2}{\partial r^2}\right)^{-1}_{r=r\,\min}$ (Bevington and Robinson 1992).

Results and discussion

In the free TAR structure, the presence of the three bulged nucleotides generates a distortion of the RNA at the junction of the upper and lower helices. Solution NMR studies found the unpaired bases U23 and C24 to stack continuously above A22 (Puglisi et al. 1992; Aboul-ela et al. 1996). This distortion in the helix is accommodated by a looping out of U25 and a widening of the major groove (Aboul-ela et al. 1996; Weeks and Crothers 1991), and results in an approximately 50 degree average bend in the inter-helical angle (Aboul-ela et al. 1996; Al-Hashimi et al. 2002; Riordan et al. 1992; Zacharias and Hagerman 1995; Long 1997). Solution NMR, circular dichroism, and transient electric birefringence experiments indicate that binding of tat protein straightens and rigidifies TAR (Aboul-ela et al. 1995; Aboul-ela et al. 1996; Long and Crothers 1999; Zacharias and Hagerman 1995; Pitt et al. 2004; Tan and Frankel 1992). Binding is accompanied by the un-stacking of U23 and C24 and looping out of two of the three bulge nucleotides, allowing the upper and lower helices to stack coaxially, forming an arginine binding pocket (Hamy et al. 1993; Aboul-ela et al. 1995; Aboul-ela et al. 1996; Long and Crothers 1999; Puglisi et al. 1993; Davis et al. 2004; Zacharias and Hagerman 1995; Pitt et al. 2004).

We used two approaches to characterize the TAR-Tat complex with solid state NMR dipolar recoupling methods. First, we used heteronuclear dipolar recoupling to assay for the structural rearrangements of the TAR RNA, described above, which occur upon binding to Tat. Long distance $^{31}P\{^{19}F\}$ REDOR NMR measurements were used for this purpose, as demonstrated previously for DNA (Merritt et al. 1999) and for this same RNA (Olsen et al. 2005). Solid state NMR (ssNMR) dipolar recoupling was used to monitor internuclear distances which span the bulged loop of TAR RNA and which are expected to change markedly when the conformation of the loop changes in response to peptide binding. Specifically, Olsen et al. used $^{31}P\{^{19}F\}$ REDOR to monitor the distance from a ^{19}F spin attached to the 2′ position of nucleotide U23 to a phosphorothioate located between G26 and A27. The chemical shift of a ^{31}P spin in a phosphothioate is shifted by about 50 ppm relative to a ^{31}P spin in a corresponding phosphate group, thus enabling the detection of REDOR dephasing of a selected ^{31}P spin in the upper helix by a ^{19}F spin attached to a nucleotide stacked on the lower helix in TAR. $^{31}P\{^{19}F\}$ REDOR studies showed that the distance from the ^{19}F spin attached to the 2′ of U23 to the ^{31}P spin between G26 and A27 changes significantly, from 10.3 Å in the unbound RNA to 6.6 Å upon binding of the Tat peptide. This result is consistent with previously reported solution NMR studies for bound and unbound HIV-1 TAR constructs (Aboul-ela et al. 1995). In the 20 published models (PDB #1ANR), the average inter-nuclear separation corresponding to our label positions was 11.1 Å. In a set of 20 published model structures (PDB #1ARJ) for the TAR apical region in complex with a peptide containing both the tat basic binding domain and core region, or with argininimide, the average inter-nuclear separation corresponding to the label positions was 4.8 Å, while the separation ranged from 3.7 to 6.1 Å (std. dev. 0.74 Å.)

To further characterize the conformational changes in TAR that occur upon binding to Tat, $^{31}P\{^{19}F\}$ REDOR was used to measure the distance from a ^{19}F spin in 5-fluorouridine incorporated at U42 to a phosphorothioate between U38 and C39. This distance was observed to change from 10.9 Å in the unbound TAR RNA to 6.8 Å following binding of the peptide. Corresponding distances in the solution NMR structures are 12.1 Å in the unbound form (ranging from 7.5 to 16.0 Å, std. dev. 2.3 Å), and 10.9 Å in the bound form (ranging from 8.5 to 13.4 Å, std. dev. 1.2 Å).

In addition to using REDOR to assay for structural changes of TAR RNA in response to Tat binding, we used heteronuclear dipolar recoupling to assess the proximity between ^{13}C and ^{15}N spins in the Arg52 side chain and ^{31}P spin in the phosphodiester backbone of the RNA. However, to interpret $^{13}C\{^{31}P\}$ and $^{15}N\{^{31}P\}$ REDOR data unambiguously, information is required to determine which ^{31}P

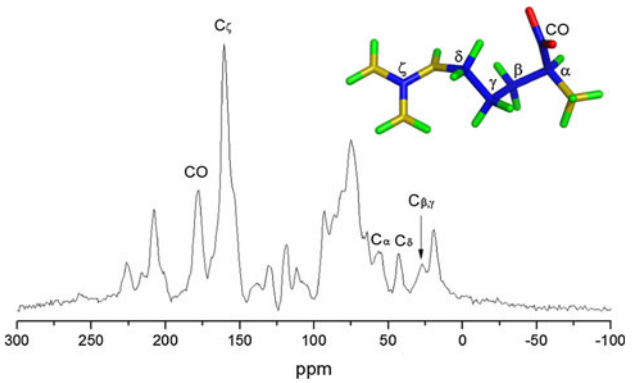

Fig. 3 ^{13}C-observed reference REDOR MAS spectrum (S_0) recorded at the initial dephasing time point (1.33 ms) along with spectral assignments (20,000 scans under spinning speed of 6,000 Hz). The inset shows the position of each carbon in the arginine amino acid

spins that are dephasing the ^{13}C or ^{15}N spins, i.e., whether or not the dephasing ^{31}P spins are in fact in the P22 and P23 phosphate groups. Thus, to determine if U23 is interacting with the side chain of Arg52, 5-fluoro-uridine was incorporated at U23. Distances between the ^{19}F spin in the base of U23 and the side chain spins of Arg52 were determined with ^{15}N{^{19}F} and ^{13}C{^{19}F} REDOR (Huang et al. 2010).

A ^{13}C observed magic angle spinning REDOR spectrum (Fig. 3) was obtained for labeled bound Tat-TAR RNA complex. The chemical shifts of labeled $^{13}C_\zeta$ and ^{13}CO in the complex are not changed relative to free arginine within the experimental linewidth, which is supported by BioMagResBank (Ulrich et al. 2008) showing a C_ζ chemical shift distribution of only 0.5 ppm (FWHM). As indicated in Fig. 3, $^{13}C_\zeta$ and ^{13}CO resonances in Arg52 are well resolved. Therefore, distances from the ^{19}F spin in the base of U23 or phosphates in the backbone of TAR RNA to C_ζ and CO can be observed by monitoring decays in the ratio of the signal S (with the rf pulse in dephasing channel) to a reference signal S_0 (without the rf pulse in dephasing channel) as a function of the number of REDOR cycles (dephasing time). The resonances of other labeled ^{13}C in Arg52 are too weak to provide usable information. Other resonances visible in the spectrum come from the RNA backbone and other amino acids in the peptide at natural abundance. Those signals could be overlapped with C_ζ and CO peaks, which may generate unpredictable decays.

Figure 4a shows the dephasing curves of ^{13}C–^{19}F REDOR experiments to obtain distances from C_ζ to U23 (5F) and from CO to U23 (5F). Homonuclear ^{13}C–^{13}C couplings are only considered for CO with the closest Cα carbon; all other carbons are ignored because they are distant from the label. Therefore, in the simulation of C_ζ, the spin system can be simplified to a single C–F pair; in the simulation of CO, the system can be simplified to two C and one F atoms. The homonuclear CO–Cα coupling is assumed to be 2080.45 Hz (1.54 Å); CO-F and Cα-F pairs are assumed to have similar distances since they are directly connected. U42 (5F) in TAR RNA is also labeled in the sample, but it is ignored here since it is far removed from both U23 (5F) and arginine, and does not have any effect on the current experiments. A χ^2 analysis (Fig. 4b) found the simulations have best agreement with the experimental data when the U23(5F)–C_ζ inter-label distance is 5.6 ± 0.1 Å, and U23(5F)–CO inter-label distance is 6.6 ± 0.4 Å. Both the U23(5F)–C_ζ and the slightly larger U23(5F)–CO distances indicate close contacts between the base of U23 within the bulged loop region of TAR RNA and the guanidinium group and carbonyl of Arg52, an observation quantitatively in agreement with previous reported solution NMR models (Aboul-ela et al. 1995). In the 20 published models (PDB #1 ARJ), the U23(5F)–C_ζ distance ranges from 3.1 to 5.9 Å, with an average inter-label distance of 4.2 Å and a standard deviation of 0.8 Å, while the U23(5F)–CO distance ranges from 5.1 to 8.5 Å, with an average inter-label distance of 6.6 Å and a standard deviation of 0.9 Å. In both cases, the inter-label distances from REDOR NMR measurement lie within the range of inter-label distances observed in solution NMR structures.

The ^{13}C{^{19}F} REDOR experiments having established the proximity of the guanidinium group of Arg52 to the base of U23 in the solid peptide-RNA complex, and we used ^{13}C{^{31}P} REDOR to orient the Arg52 guanidinium group relative to the phosphodiester backbone of the bulged loop. Figure 5 shows the results of ^{13}C{^{31}P} REDOR experiments applied to the $^{13}C_\zeta$ and ^{13}CO spins of Arg52. A significant decay of S/S_0 for $^{13}C_\zeta$ in the guanidinium group is observed, while the S/S_0 of ^{13}CO does not decay, indicating that the guanidinium group in the side chain has detectable contacts with one or more phosphate groups, but the ^{13}CO group in the main chain of the tat peptide is not located close enough to a phosphate group to display any significant ^{13}C{^{31}P} REDOR dephasing. To simulate the $^{13}C_\zeta$ {^{31}P} REDOR dephasing curve in Fig. 5a, we assume the $^{13}C_\zeta$ may be located close to at least two phosphate groups. Based on the proximity of $^{13}C_\zeta$ to the base of U23 and prior solution NMR results, we presume to be the P22 and P23 phosphates (Churcher et al. 1993; Aboul-ela et al. 1995; Calnan et al. 1991a, b; Hamy et al. 1993; Pritchard et al. 1994; Tao and Frankel 1992). In the simulation, all homonuclear ^{31}P–^{31}P couplings from outside the P22–P23 pair were neglected and the ^{31}P–^{31}P dipolar coupling of P22–P23 pair was assumed to be 72 Hz (6.5 Å).

Multiple solutions to the fitting of the ^{13}C{^{31}P} REDOR data exist, corresponding to different orientations of the guanidinium $^{13}C_\zeta$ relative to the P22–P23 ^{31}P spin pair, are shown as a χ^2 plot in Fig. 5b. Specifically, 6 sets of

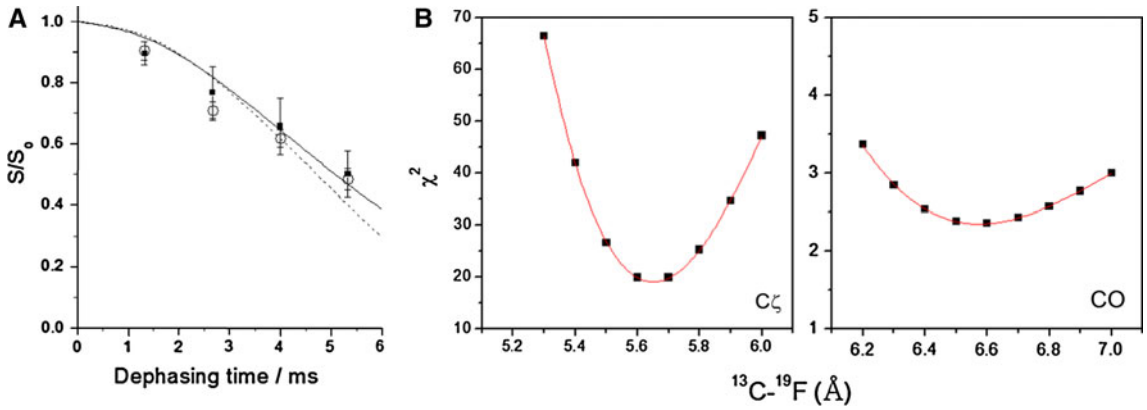

Fig. 4 **a** ^{13}C–^{19}F REDOR dephasing curves for the Tat peptide-TAR RNA complex, along with the results of simulations performed with SIMPSON. Empty cycles and *dashed line* denote the decay of C_ζ in Arg52. *Solid squares* and *solid line* denote the decay of CO. **b** The graphs of the χ^2 function identify distances for which the REDOR data are best fit by the simulations

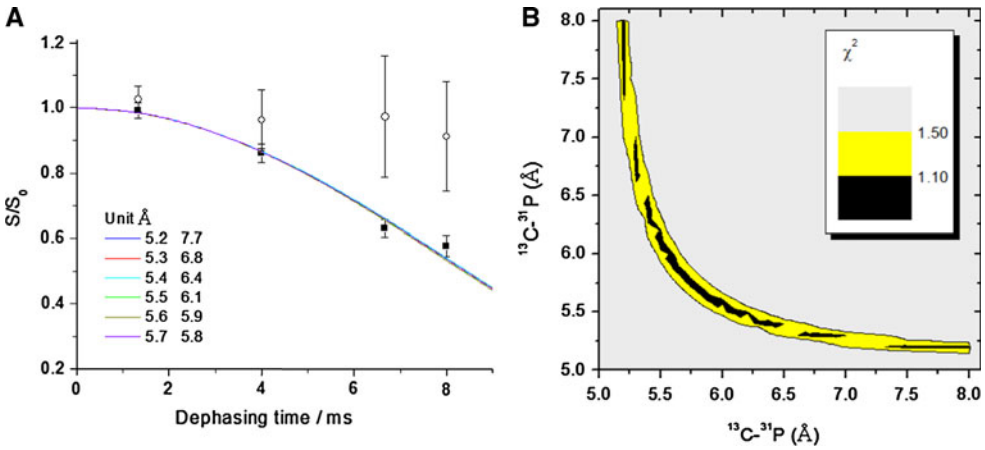

Fig. 5 **a** ^{13}C–^{31}P REDOR dephasing *curves* for the Tat peptide-TAR RNA complex, along with the simulations performed with SIMPSON. *Empty circles* denote the decay of CO in Arg52. *Solid squares* and *lines* denote the decay of $^{13}C_\zeta$ signal, and multiple solutions are shown in the legend. **b** Contour plot of normalized χ^2; each axis shows the distance of the ^{31}P to the ^{13}C

solutions are consistent with the normalized χ^2 values. These solutions are expressed as two distances from the $^{13}C_\zeta$ spin to the two ^{31}P spins of P22 and P23: (5.7, 5.8 Å), (5.6, 5.9 Å), (5.5, 6.1 Å), (5.4, 6.4 Å), (5.3, 6.8 Å) and (5.2, 7.7 Å), see the legend to Fig. 5a. By way of comparison, 20 solution NMR model structures (Aboul-ela et al. 1995) were examined for agreement with REDOR-measured distances between $^{13}C_\zeta$ of Arg52 and the ^{31}P spins of the P22 and P23 phosphate groups. Because these distances were not measured directly but deduced from NOE data, there was some variation in the $^{13}C_\zeta$–^{31}P22 and $^{13}C_\zeta$–^{31}P23 distances in the 20 models. Best agreement to REDOR-measured distances are: 5.0 Å (C_ζ-P22) and 5.5 Å (C_ζ-P23) in model 2 (1ARJ-2), and 6.9 Å (C_ζ-P22) and 5.2 Å (C_ζ-P23) in model 8 (1ARJ-8). We conclude that the $^{13}C\{^{31}P\}$ data place the gaunidinium group of Arg52 in close proximity to the ^{31}P spins of P22 and P23 in agreement with solution NMR structures.

Besides the ^{13}C observed REDOR experiments, ^{15}N–^{19}F and ^{15}N–^{31}P REDOR were also performed in order to obtain the additional information on the proximity of the guanidinium group to the base of U23, and to observe direct interactions between the ^{15}N spins of the guanidinium group of Arg52 and the ^{31}P spins of the P22 and P23 groups. As indicated in Fig. 6, the labeled ^{15}N resonances are also well resolved, except the $\eta1$ and $\eta2$ nitrogens. The Nε and N$\eta1$,$\eta2$ signals are partially overlapped, but they can be separated well into two Gaussian-shaped peaks. The resonances of labeled ^{15}N are not shifted in the complex compared with free arginine within the experimental linewidth. BioMagResBank (Ulrich et al. 2008) also shows a Nε chemical shift distribution of 1 ppm (FWHM), which is in agreement with our observation. According to the well assigned and resolved ^{15}N resonances, the designed measurements are practical.

The 15N–19F and 15N–31P REDOR dephasing curves are shown in Figs. 7a and 8, respectively. Because the 15N$\eta1$ and 15N$\eta2$ are not resolved, only one dephasing curve is obtained for those two 15N spins. This results in multiple solutions for N$\eta1$ and N$\eta2$ distances in the simulations. In the upper panel of Fig. 7a, a slower decay is observed for the amide 15N of Arg52 in the main chain of

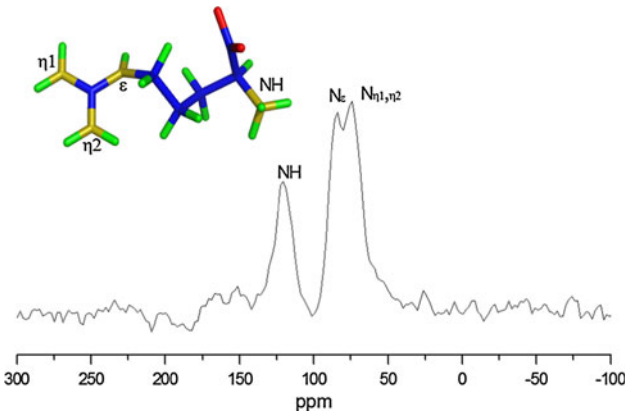

Fig. 6 ^{15}N observed reference REDOR MAS spectrum (S_0) recorded at the initial dephasing time point (1.00 ms) along with spectral assignments (22,000 scans under spinning speed of 8,000 Hz). The inset shows the position of each nitrogen in arginine

the Tat peptide compared with Nε, confirming that the guanidinium group has closer contacts with the bulge region. In the simulation, homonuclear 15N-15N couplings are only considered within the guanidinium group, and are assumed to be 111 Hz (2.23 Å). Best fits to the 15N–19F REDOR data (Fig. 7b) shows the U23(519F)–15NH distance is 5.2 ± 0.2 Å, U23(519F)–15Nε distance is 4.6 ± 0.1 Å. Multiple solutions exist for the U23(519F)–15Nη1 and U23(519F)–15Nη2 distances. As shown in contour plot of normalized χ2 in Fig. 7c, there are 6 regions ((a, b) and (b, a) are considered to be the same region) with the relative small χ2 compared with surroundings. The best fits in each region are (4.3, 6.7 Å), (4.4, 5.9 Å), (4.5, 5.5 Å), (4.6, 5.2 Å), (4.7, 5.1 Å) and (4.8, 4.9 Å), see also the legend of Fig. 7a.

Given the rigid conformation of the guanidinium group, the REDOR-derived ^{15}N–^{19}F distances are consistent with the U23(5F)–C_ζ distance, 5.6 Å, which is obtained in ^{13}C–^{19}F REDOR. By way of comparison, in 20 solution NMR models of the TAR-Tat complex (Aboul-ela et al. 1995), the U23 (5^{19}F)–^{15}Nε distance ranges from 2.8 to 5.3 Å with an average of 3.7 Å and a standard deviation of 0.7 Å. The U23 (5^{19}F)–^{15}Nη1 distance similarly ranges from 2.4 to 7.1 Å with an average of 4.7 Å and a standard deviation of 1.1 Å, and the U23 (5^{19}F)–^{15}Nη2 distance ranges from 3.3 to 7.1 Å with an 4.5 Å in average and standard deviation 0.9 Å. Thus, a quantitative agreement with the REDOR distances is observed.

Distances from ^{15}Nε to ^{15}Nη1, η2 to nearby phosphate groups were measured using ^{15}N{^{31}P} REDOR. No visible decay of S/S_0 for ^{15}Nε is observed in the time scale of 8 ms, which suggests a distance > 5.4 Å from the ^{15}Nε to the nearest phosphate group. Possible solutions for distances between Nη1/Nη2 and two close phosphate groups are much more complicated because four distances (Nη–P22, Nη1–P23, Nη2–P22 and Nη2–P23) are involved in the

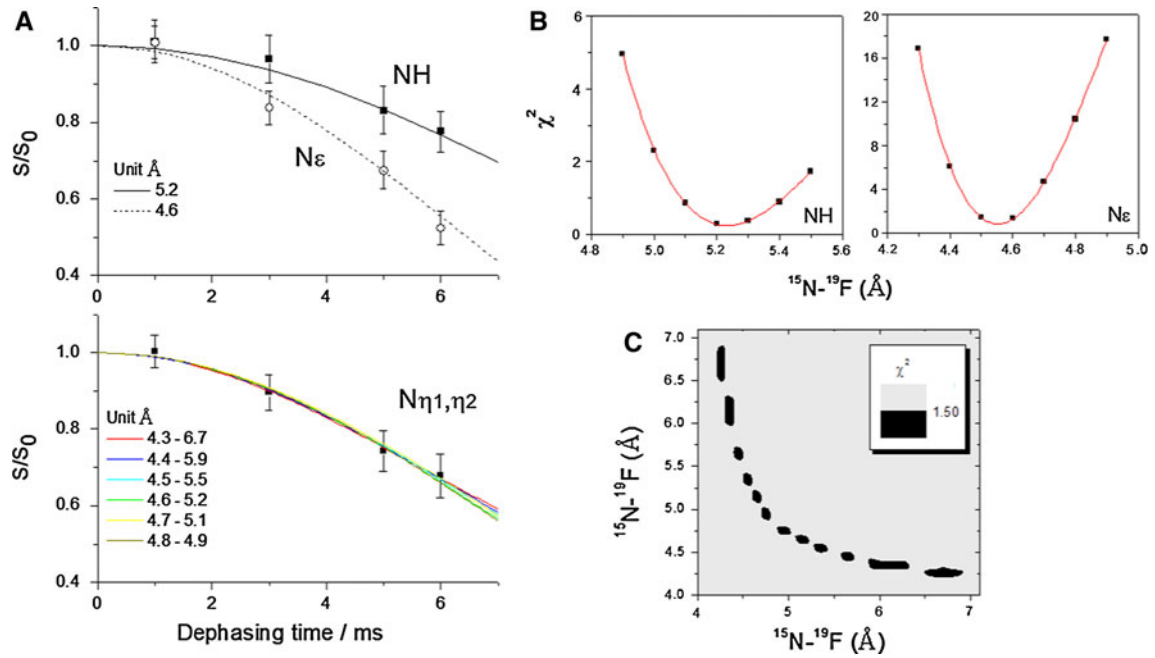

Fig. 7 **a** ^{15}N–^{19}F REDOR dephasing curves for the Tat peptide-TAR RNA complex, along with the simulations performed with SIMPSON. Upper panel shows the decay of NH and Nε in Arg52. Lower panel shows the decay of Nη1 and Nη2, and multiple solutions are shown in the legend. **b** The graphs of the χ^2 function regarding NH and Nε simulations, identify distances for which the REDOR data are best fit by the simulations. **c** Contour plot of normalized χ^2 regarding Nη1 and Nη2 simulations; each axis shows the distance of a ^{19}F and the ^{15}N

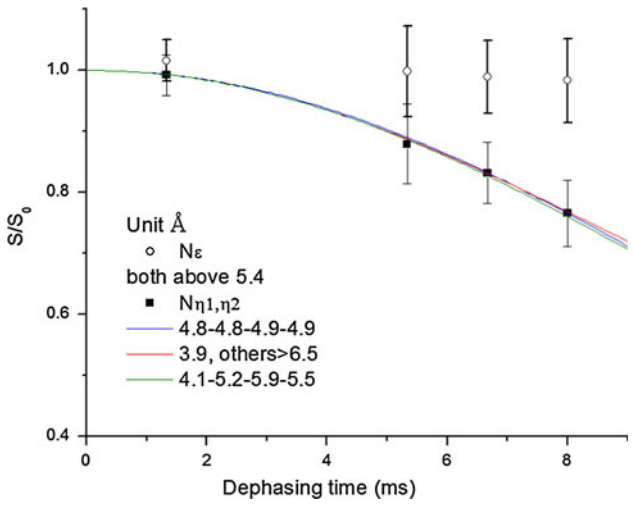

Fig. 8 ^{15}N–^{31}P REDOR dephasing curves for the Tat peptide-TAR RNA complex, along with the simulations performed with SIMPSON. *Empty circles* denote the decay of Nε in Arg52; *solid squares* and solid line denote the decay of Nη1 and Nη2

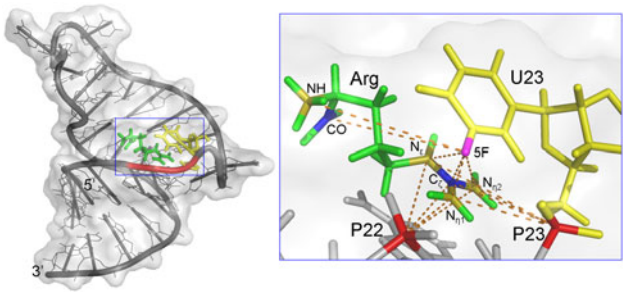

Fig. 9 Three dimensional structure of the complex between TAR and a Tat-derived 37-mer peptide (Aboul-ela et al. 1995); the Tat-binding region is zoomed in on the right. Distances between TAR RNA and arginineamide are indicated by *dash lines*

simulation, leading to numerous possibilities. However, ^{15}N–^{31}P distance ranges can be obtained that satisfy the REDOR data. There is no visible effect observed in simulations of REDOR dephasing curves when the distance of ^{15}N–^{31}P pair is above 6.5 Å. If only one ^{15}N–^{31}P pair is close enough to produce observable REDOR recoupling, and the distances between the other three are above 6.5 Å, the best fit short distance is 3.9 Å. If the distances of all four N–P pairs are about equal, the best fitting is (4.8, 4.8, 4.9 and 4.9 Å), as shown in the fitting lines in Fig. 8. Considering the position of the guanidinium group relative to the P22 and P23 phosphates indicated in Fig. 9 and the above simulation results, the Nη1/Nη2–P22/P23 distances are most likely to be in the range from 4 to 6 Å. No decay has been observed for the amide ^{15}N of Arg52 in the main chain of the Tat peptide, confirming that the guanidinium group directly contacts the important phosphates in the bulge region. To make a quantitative comparison, corresponding distances obtained from model 2 (1ARJ-2) of the 20 solution NMR models (Aboul-ela et al. 1995) are listed here, 5.6 and 6.7 Å (Nε–P22/P23), 3.7 and 4.9 Å (Nη1–P22/P23), and 5.9 and 5.2 Å (Nη2–P22/P23). According to these numbers, one possible solution of Nη1/Nη2–P22/P23 pair is (4.1, 5.2, 5.9 and 5.5 Å), as shown in Fig. 8. Thus, a good agreement between solid state and solution NMR is observed. Although a single solution is not obtained, the close contact with a distance at a range of about 5 Å between Nη1/Nη2 and phosphate groups is consistent with all observations. Based on ^{15}N{^{31}P} and ^{13}C{^{31}P] REDOR data, it is reasonable to conclude that a close interaction exists between the guanidinium group of Arg52 and the P22 and P23 phosphate groups.

Conclusions

The structure of Tat-TAR complex has been widely studied using different biophysical and biochemical techniques because of its important role in promoting HIV transcriptions. However, previously reported structures were incomplete, because insufficient intermolecular NOEs were available to define the structure of the complex. Our present work reports the measurement of several intermolecular distances between Arg52, especially its guanidinium group, and the TAR RNA bulge region. The close contacts from the guanidinium group to U23 and phosphate groups confirm the formation of the complex in solid state and provide distances consistent with existing partial solution NMR models of the complex, see Table 1. Clearly, solid-state REDOR NMR measurements have the ability to directly obtain distances between protein side chains and RNA, as shown schematically in Fig. 9. By measuring more distances using these methods, it may become possible to establish the complete structure of this paradigmatic complex. This structure would have

Table 1 Summary of comparison between solid state REDOR distances and the corresponding distances from the solution NMR model

	REDOR, Å	Solution NMR model, Å
U23(5F)–CO	6.6	6.6[b]
U23(5F)–C_ζ	5.6	4.2[b]
U23(5F)–Nε	4.6	3.7[b]
U23(5F)–Nη1/Nη2	(4.6, 5.2)[a]	(4.7, 4.5)[b]
P22/P23–C_ζ	(5.6, 5.9)[a]	(5.0, 5.5)[c]
P22/P23–Nε	>5.4	(5.6, 6.7)[c]
P22/P23–Nη1/Nη2	(4.1, 5.2, 5.9, 5.5)[a]	(3.7, 4.9, 5.9, 5.2)[c]

[a] One of possible solutions

[b] Average distance of 20 solution NMR models (PDB #1 ARJ)

[c] Model 2 (1ARJ-2) of the 20 solution NMR models

important applications for the discovery of small molecules inhibitors, and to understand the molecular basis for this interaction.

Acknowledgments The authors wish to thank Dr. Dirk Stueber for preparing the labeled TAR, and thank Dr. Jason Ash and Dr. Greg Olsen for the helpful discussions. This work was supported by the National Institutes of Health [RO1-EB03152]; and the National Science Foundation [MCB-0642253].

References

Aboul-ela F, Karn J, Varani G (1995) The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. J Mol Biol 253:313–332

Aboul-ela F, Karn J, Varani G (1996) Structure of HIV-1TARRNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. Nucl Acids Res 24:3974–3981

Al-Hashimi HM, Gosser Y, Gorin A, Hu W, Majumdar A, Patel DJ (2002) Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. J Mol Biol 315:95–102

Baba M (2006) Recent status of HIV-1 gene expression inhibitors. Antivir Res 71:301–306

Bachelin M, Hessler G, Kurz G, Hacia JG, Dervan PB, Kessler H (1998) Structure of a stereoregular phosphorothioate DNA/RNA duplex. Nat Struct Biol 5:271–276

Bak M, Rasmussen JT, Nielsen NC (2000) SIMPSON: a general simulation program for solid-state NMR spectroscopy. J Magn Reson 147:296–330

Bardaro MF, Shajani Z, Patora-Komisarska K, Robinson JA, Varani G (2009) How binding of small molecule and peptide ligands to HIV-1 TAR alters the RNA motional landscape. Nucl Acids Res 37:1529–1540

Bevington PR, Robinson DK (1992) Data reduction and error analysis for the physical sciences, 2nd edn. McGraw-Hill, New York

Brodsky AS, Williamson JR (1997) Solution structure of the HIV-2 TAR-argininamide complex. J Mol Biol 267:624–639

Calnan BJ, Biancalana S, Hudson D, Frankel AD (1991a) Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA–protein recognition. Genes Dev 5:201–210

Calnan BJ, Tidor B, Biancalana S, Hudson D, Frankel AD (1991b) Arginine-mediated RNA recognition: the arginine fork. Science 252:1167–1171

Churcher MJ, Lamont C, Hamy F, Dingwall C, Green SM, Lowe AD, Butler PJG, Gait MJ, Karn J (1993) High affinity binding of TAR RNA by the human immunodeficiency 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

Cordingley MG, LaFemina RL, Callahan PL, Condra JH, Sardana VV, Graham DJ, Nguyen TM, LeGrow K, Gotlib L, Schlabach AJ, Colonno RJ (1990) Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive sequence element of human immunodeficiency virus type 1 in vitro. Proc Natl Acad Sci USA 87:8985–8989

Davidson A, Leeper TC, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson JA, Varani G (2009) Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. Proc Natl Acad Sci USA 106:11931–11936

Davis B, Afshar M, Varani G, Murchie AI, Karn J, Lentzen G, Drysdale M, Bower J, Potter AJ, Starkey ID, Swarbrick T, Aboul-ela F (2004) Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic 'hot spots'. J Mol Biol 336:343–356

Delling U, Reid LS, Barnett RW, Ma MY, Climie S, Sumner-Smith M, Sonenberg N (1992) Conserved nucleotides in the TAR RNA stem of human immunodeficiency virus type 1 are critical for tat binding and trans-activation: model for TAR RNA tertiary structure. J Virol 66:3018–3025

Dingwall C, Ernberg I, Gait MJ, Green SM, Heaphy S, Karn J, Lowe AD, Singh M, Skinner MA, Valerio R (1989) Human immunodeficiency virus 1 Tat protein binds trans-activation-responsive region (TAR) RNA in vitro. Proc Natl Acad Sci USA 86:6925–6929

Dingwall C, Ernberg I, Gait MJ, Green SM, Heaphy S, Karn J, Lowe AD, Singh M, Skinner MA (1990) HIV-1 Tat protein stimulation transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. EMBO J 9:4145–4153

Gallego J, Varani G (2001) Targeting RNA with small-molecule drugs: Therapeutic promise and chemical challenges. Acc Chem Res 34:836–843

Gonzalez C, Stec W, Reynolds MA, James TL (1995) Structure and dynamics of a DNA.RNA hybrid duplex with a chiral phosphorothioate moiety: NMR and molecular dynamics with conventional and time-averaged restraints. Biochemistry 34:4969–4982

Gullion T, Schaefer J (1989a) Rotational-echo double-resonance NMR. J Magn Reson 81:196–200

Gullion T, Schaefer J (1989b) Detection of weak heteronuclear dipolar couplings by rotational-echo double-resonance nuclear magnetic resonance. Adv Magn Reson 13:57–83

Hamy F, Asseline V, Grasby J, Iwai S, Pritchard C, Slim G, Butler PJG, Karn J, Gait M (1993) Hydrogen bonding contacts in the major groove are required for human immunodeficiency virus type-1 Tat protein recognition of TAR RNA. J Mol Biol 230:111–123

Harper JW, Logsdon NJ (1991) Refolded HIV-1 Tat protein protects both bulge and loop nucleotides in TAR RNA from ribonucleolytic cleavage. Biochemistry 30:8060–8066

Huang W, Varani G, Drobny GP (2010) 13C/15 N–19F Intermolecular REDOR NMR Study of the Interaction of TAR RNA with Tat Peptides. J Am Chem Soc 132:17643–17645

Huq I, Ping YH, Tamilarasu N, Rana TM (1999) Controlling human immunodeficiency virus type 1 gene expression by unnatural peptides. Biochemistry 38:5172–5177

Jehle S, Falb M, Kirkpatrick JP, Oschkinat H, van Rossum BJ, Althoff G, Carlomagno T (2010) Intermolecular protein-RNA interactions revealed by 2D 31P–15 N magic angle spinning solid-state NMR spectroscopy. J Am Chem Soc 132:3842–3846

Karn J (1999) Tackling Tat. J Mol Biol 293:235–254

Long KS (1997) Characterization of a human immunodeficiency virus TAR RNA element and its complex with a Tat-derived peptide. Dissertation, Yale University

Long KS, Crothers DM (1999) Characterization of the solution conformations of unbound and Tat peptide-bound forms of HIV-1 TAR RNA. Biochemistry 38:10059–10069

Loret EP, Georgel P, Johnson WC, Ho PS (1992) Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator. Proc Natl Acad Sci USA 89:9734–9738

Merritt ME, Sigurdsson ST, Drobny GP (1999) Long-range measurements to the phosphodiester backbone of solid nucleic acids using ^{31}P–^{19}F REDOR NMR. J Am Chem Soc 121:6070–6071

Muesing MA, Smith DH, Capon DJ (1987) Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell 48:691–701

Murchie AIH, Davis B, Isel C, Afshar M, Drysdale MJ, Bower J, Potter AJ, Starkey ID, Swarbrick TM, Mirza S, Prescott CD, Vaglio P, Aboul-ela F, Karn J (2004) Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. J Mol Biol 336:625–638

Olsen GL, Edwards TE, Deka P, Varani G, Sigurdsson ST, Drobny GP (2005) Monitoring tat peptide binding to TAR RNA by solid-state ^{31}P–^{19}F REDOR NMR. Nucl Acids Res 33:3447–3454

Pitt SW, Majumdar A, Serganov A, Patel DJ, Al-Hashimi HM (2004) Argininamide binding arrests global motions in HIV-1TAR-RNA: comparison with Mg^{2+}-induced conformational stabilization. J Mol Biol 338:7–16

Pritchard CE, Grasby JA, Hamy F, Zacharech AM, Singh M, Karn J, Gait MJ (1994) Methylphosphonate mapping of phosphate contacts critical for RNA recognition by the human immunodeficiency virus Tat and Rev proteins. Nucl Acids Res 22: 2592–2600

Puffer B, Kreutz C, Rieder U, Ebert MO, Konrat R, Micura R (2009) 5-Fluoro pyrimidines: labels to probe DNA and RNA secondary structures by 1D ^{19}F NMR spectroscopy. Nucl Acids Res 37: 7728–7740

Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR (1992) Conformation of the TAR RNA–arginine complex by NMR spectroscopy. Science 257:76–80

Puglisi JD, Chen L, Frankel AD, Williamson JR (1993) Role of RNA structure in arginine recognition of TAR RNA. Proc Natl Acad Sci USA 90:3680–3684

Rana TM, Jeang KT (1999) Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. Arch Biochem Biophys 365:175–185

Riordan FA, Bhattacharyya A, McAteer S, Lilley DM (1992) Kinking of RNA helices by bulged bases, and the structure of the human immunodeficiency virus transactivator response element. J Mol Biol 226:305–310

Roy S, Delling U, Chen CH, Rosen CA, Sonenberg N (1990a) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tatmediated trans-activation. Genes Dev 4:1365–1373

Roy S, Parkin NT, Rosen C, Itovitch J, Sonenberg N (1990b) Structural requirements for trans activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by tat: importance of base pairing, loop sequence, and bulges in the tat-responsive sequence. J Virol 64:1402–1406

Sumner-Smith M, Roy S, Barnett R, Reid LS, Kuperman R, Delling U, Sonenberg N (1991) Critical chemical features in trans-acting-responsive RNA are required for interaction with human immunodeficiency virus type 1 Tat protein. J Virol 65: 5196–5202

Tan R, Frankel AD (1992) Circular dichroism studies suggest that TAR RNA changes conformation upon specific binding of arginine or guanidine. Biochemistry 31:10288–10294

Tao J, Frankel AD (1992) Specific binding of arginine to TAR RNA. Proc Natl Acad Sci USA 89:2723–2726

Tao J, Chen L, Frankel AD (1997) Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction. Biochemistry 36: 3491–3495

Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Wenger RK, Yao H, Markley JL (2008) BioMagResBank. Nucl Acids Res 36:402–408

Weeks KM, Crothers DM (1991) RNA recognition by Tat-derived peptides: interaction in the major groove? Cell 66:577–588

Weeks KM, Crothers DM (1992) RNA binding assays for Tat-derived peptides: implications for specificity. Biochemistry 31:10281–10287

Weeks KM, Ampe C, Schultz SC, Steitz TA, Crothers DM (1990) Fragments of the HIV-1 Tat protein specifically bind TAR RNA. Science 249:1281–1285

Zacharias M, Hagerman PJ (1995) The bend in RNA created by the trans-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat-derived peptide. Proc Natl Acad Sci USA 92:6052–6056