Nucleic Acids Research, 1998, Vol. 26, No. 8 **1991–1995**

NMR evidence for a base triple in the HIV-2 TAR C-G·C$^+$ mutant–argininamide complex

Alexander S. Brodsky[+], Heidi A. Erlacher[§] and James R. Williamson[*,§]

MIT Department of Chemistry, Building 56-546, Cambridge, MA 02139, USA

Received November 18, 1997; Revised and Accepted February 17, 1998

ABSTRACT

Formation of a specific complex between the HIV Tat protein and the small RNA element TAR is critical for activation of viral transcription. A model complex for this interaction composed of HIV-2 TAR and the amide derivative of arginine has been developed to study how Tat and TAR interact specifically. We have previously determined a high resolution NMR structure of the HIV-2 TAR–argininamide complex. The argininamide guanidium group hydrogen bonds to the major groove face of G26 and is stacked between U23 and A22, forming an arginine sandwich. This structure also provided evidence for formation of a U38-A27·U38 base triple, as U23 is positioned in the major groove within hydrogen bonding distance to A27. However, the expected U23 imino proton was not observed, preventing unambiguous identification of the base triple. Previous work on an isomorphic C38-G27·C23$^+$ base triple mutant of the three base bulge HIV-1 TAR–argininamide complex demonstrated that the base triple is required for specific argininamide binding. Here we investigate the same C38-G27·C23$^+$ base triple mutant in the context of two base bulge HIV-2 TAR. The improved NMR spectral properties of HIV-2 TAR allowed observation of the C23 amino and imino protons for the first time, providing direct evidence that a hydrogen bonding interaction is occurring. The NOEs observed correspond to those observed in the high resolution structure of the HIV-2 TAR–argininamide complex, confirming that a base triple is an important feature of the TAR–argininamide interaction.

INTRODUCTION

The Tat protein–HIV TAR RNA interaction is critical for viral replication (1,2) because binding of Tat to TAR activates transcription of the viral genome. Using both *in vivo* and *in vitro* biochemical approaches the Tat binding site has been localized to the bulge region of TAR located at the 5′-end of the viral mRNA (3–7). Specifically, the A27-U38 and G26-C39 base pairs in the upper stem and a single nucleotide in the bulge, U23, were identified as important for specific Tat binding and for *in vivo* activity (8,9). Small basic peptides, encompassing the basic region of Tat, were found to bind TAR specifically (4,8–10). Surprisingly, the Tat–TAR interaction can be further simplified to an arginine–TAR complex, as shown by a series of careful binding studies (11). Mutagenesis and modification interference studies identified the same nucleotides as important for binding to both the Tat peptides and a single arginine (11,12). These results suggest that TAR is recognizing arginine and the Tat protein in the same manner.

A number of NMR structural studies of the HIV TAR–argininamide complex have been performed. In the first study by Puglisi *et al*. a model was proposed where the argininamide guanidium group binds in the major groove near G26 and that U23 is near A27, suggesting formation of a U38-A27·U23 base triple upon argininamide binding to HIV-1 TAR (13). However, no direct NOE evidence for the proposed base triple was obtained. The second NMR study by Aboul-ela *et al.* of the HIV-1 TAR–argininamide complex included a much larger NOE data set (14). No evidence was found in this study for the proposed base triple, even though the global conformation of the TAR bulge was similar to the model of Puglisi *et al*. and U23 was positioned in the major groove near A27. The most recent high resolution NMR study focused on the two base bulge HIV-2 TAR–argininamide complex, which exhibited higher quality spectra than three base bulge HIV-1 TAR, allowing for an even larger NOE data set to be obtained (15). The improved spectral quality is presumably due to the different dynamic properties of two base bulge HIV-2 TAR. The HIV-2 TAR structure is consistent with the original Puglisi *et al.* model with U23 positioned in the major groove and within hydrogen bonding distance to A27, suggesting base triple formation. A series of modeling calculations demonstrated that idealized planar hydrogen bonds which enforced the base triple structure were consistent with all the NMR restraints. Even though the NMR spectra of the two base bulge HIV-2 TAR–argininamide complex are superior to those of HIV-1 TAR, the effects of dynamics are still present, limiting the number of weak NOEs observed and preventing unambiguous identification of the proposed base triple. The expected U23 imino proton may simply exchange too rapidly to be observed, as is frequently observed for terminal base pairs.

*To whom correspondence should be addressed at present address. Tel: +1 619 784 8740; Fax: +1 619 784 2199; Email: jrwill@scripps.edu

Present addresses: [+]Dana Farber Cancer Institute, Department of Cancer Biology, 44 Binney Street, Boston, MA 02115, USA and [§]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

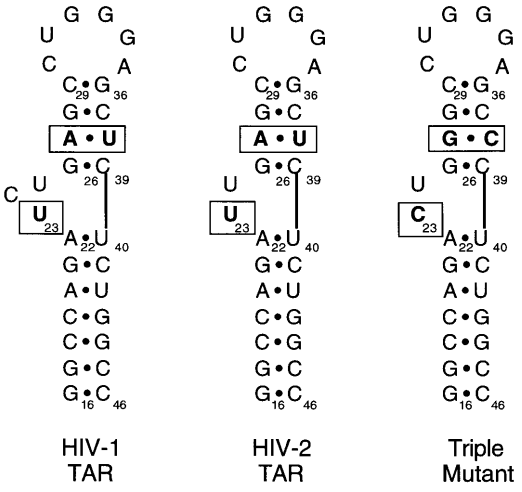

Figure 1. Sequence and secondary structure of wild-type three base bulge HIV-1 TAR, two base bulge HIV-2 TAR and the HIV-2 TAR CGC triple mutant. The critical nucleotides required for arginine and Tat binding in the bulge region are highlighted in bold.

To determine if the base triple is required for argininamide binding, an isomorphic C-G·C^+ base triple mutant was introduced in three base bulge HIV-1 TAR and was found to bind argininamide in the expected pH-dependent manner (16). However, the proton resonances at the interface of the base triple were unobservable and no NOEs between C23 and the G27-C38 base pair were observed due to the dynamic nature of the complex. Here, we present NMR studies of the HIV-2 TAR C-G·C^+ base triple mutant. The same pH-dependent argininamide binding was observed, however, additional proton resonances due to hydrogen bonding at the interface of the proposed base triple were observed for the first time. These protons exhibit a number of inter-residue NOEs that position C23 and G27 near each other, consistent with base triple formation. This work provides direct NMR evidence of the base triple structure in the HIV-2 TAR–argininamide complex.

MATERIALS AND METHODS

Sample preparation

The HIV-2 TAR CGC RNA was synthesized by *in vitro* transcription by T7 RNA polymerase and purified by denaturing gel electrophoresis (17). Samples were dialyzed for >48 h against NMR buffer (10 mM sodium phosphate, pH 6.4 at 25°C, 50 mM NaCl and 0.1 mM EDTA). The final concentration in 600 µl was ~1.5 mM. The sample was then titrated with argininamide (Sigma) in the NMR tube to a final total concentration of 6 mM and the pH was then lowered to 5.4 with HCl.

NMR spectroscopy

NMR experiments were recorded on a Varian INOVA 600 MHz spectrometer, Varian Unity Plus 750 MHz spectrometer or a 500 MHz spectrometer constructed at the Francis Bitter Magnet Laboratory. All spectra were processed and analyzed using NMRPipe and PIPP on a Silicon Graphics Indy workstation (18,19). Proton NMR spectra in H_2O were recorded at a variety of temperatures and mixing times, including 50, 100 and 200 ms. NOESY spectra used a WATERGATE 3-9-19 water suppression

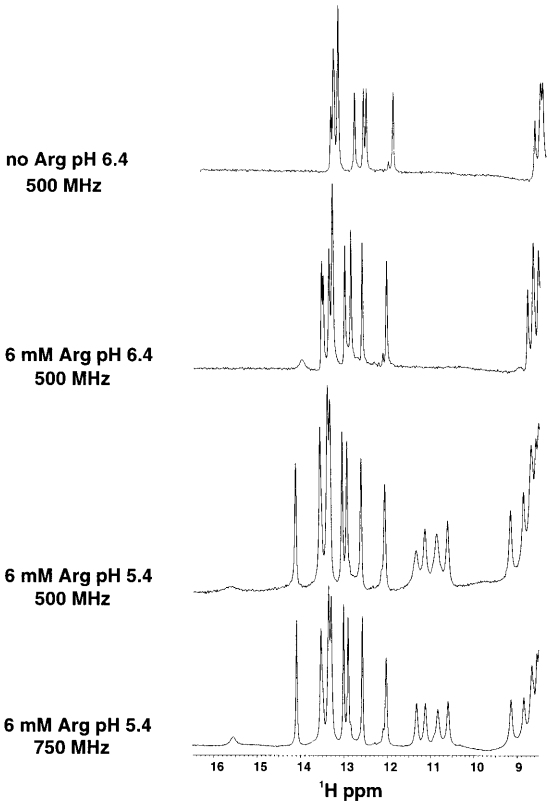

Figure 2. The imino–amino proton region of HIV-2 TAR CGC shows a number of new resonances upon addition of argininamide and lowering of the pH to 5.4. At pH 6.4 a few changes are observed upon addition of argininamide. At pH 5.4 the U40 H3 sharpens significantly and the C23 H3 resonance appears at 15.55 p.p.m. In addition, the C23 H42/H41 are also observable at 10.61 and 9.17 p.p.m. The chemical shifts of the C23 amino and imino protons are in the range observed for protonated cytidines found in C-G·C^+ base triples.

scheme with a selective E-BURP-1 flip-back pulse (20–22). Sweep widths were 12 000, 14 000 and 16 000 Hz at 500, 600 and 750 MHz respectively, with 4096 × 512 points collected. Some spectra were recorded with sweep widths of 14 000 × 8000 Hz, which increased the digital resolution. Data were zero filled to 4k × 2k real points and apodized using Gaussian–Lorentzian functions in both dimensions before Fourier transformations.

Standard NOESY, TOWNY (23) and DQ-COSY spectra in 99.996% ^2H_2O were recorded at 25°C. To monitor NOE build-ups, spectra were recorded at 500, 600 and 750 MHz with mixing times of 50, 100, 200 and 400 ms. Sweep widths were 5500, 6000 and 8000 Hz at 500, 600 and 750 MHz respectively. Some spectra were obtained with 4096 × 1024 real points, which aided assignment of some crowded resonances.

RESULTS AND DISCUSSION

Previous studies showed that an isomorphic base triple mutant TAR can be designed that binds argininamide. In our work with two base bulge HIV-2 TAR we found improved dynamic properties of the TAR–argininamide complex allowing for high resolution structure determination of this complex. Although this work provided more indirect evidence for base triple formation upon argininamide binding, no direct NOE evidence indicative of

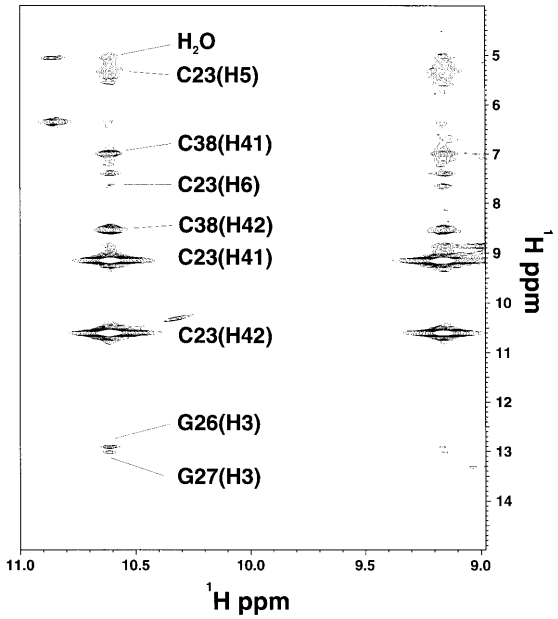

Figure 3. The internucleotide NOEs to the C23 amino protons are shown in a H_2O NOESY at 2°C. A number of NOEs between C23 and G26, G27 and U38 position C23 near G27, consistent with a base triple structure.

the presence of a base triple in the HIV-2 TAR–argininamide complex was observed. We reasoned that replacement of the U38-A27·U23 base triple with an isomorphic C38-G27·C23$^+$ base triple in the context of HIV-2 TAR may allow us to observe critical resonances indicative of base triple formation in the argininamide complex. Typically, the cytidine amino proton resonances are sharper than uridine imino proton resonances and, therefore, may be observable upon hydrogen bonding in the base triple of TAR. Also, the extra G-C base pair in the upper stem may improve overall stability of the complex. Thus, we constructed the two base bulge HIV-2 TAR CGC triple mutant as shown in Figure 1.

As previously observed in studies of HIV-1 TAR CGC, argininamide does not bind significantly at pH 6.4. However, at pH 5.4 the U40 imino proton resonance significantly sharpens and the G26 and G21 imino resonances shift, as is observed for argininamide binding to wild-type TAR. In addition, a number of new peaks appear in the imino and amino proton regions, as shown in Figure 2. These exchangeable resonances are generally not observable under these conditions unless they are involved in a hydrogen bonding interaction. The peak at 15.55 p.p.m. is too broad to be observed in 2D ^{1}H-^{1}H NOESY experiments but its chemical shift is in the exact range observed for protonated cytidine imino protons found in base triple structures in DNA triplex NMR studies (24–26). The remaining new resonances are easily identified as cytidine amino protons in 750 MHz NOESY spectra. The cytidine amino protons are also observed at chemical shifts typical for a base triple structure (24–26).

The NOEs observed to C23 amino protons in a H_2O NOESY are shown in Figure 3. Since the chemical shifts for most of the molecule are very similar to wild-type HIV-2 TAR, resonance assignments were straightforward. Most of the exchangeable imino protons and amino protons observed in the high resolution HIV-2 TAR–argininamide structure were assigned, along with most of the H8/H6/H5 and H1′ resonances, as shown in Table 1.

Complete assignments were not possible without isotopic labeling, but the key resonances in the bulge region were easily assigned, since most of the resonances have almost identical chemical shifts to the wild-type HIV-2 TAR–argininamide complex. The C38 and G27 resonances were readily assigned by observation of standard A-form sequential NOEs. The C23 amino protons were assigned based on very strong NOEs to C23 H5, which in turn gives NOEs to the C23 base and ribose. The C23 sugar pucker is C2′-*endo*, as in wild-type HIV-2 TAR, which allows assignment via ^{1}H-^{1}H TOCSY and DQ-COSY experiments (data not shown; 27).

Table 1. Proton chemical shifts (p.p.m.) for HIV-2 TAR CGC RNA bound to argininamide

Residue	H1/H3	H42/H41	H8/H6	H5/H2	H1′
G16		–	8.09	–	5.78
G17	13.51	–	7.69	–	5.93
C18		8.70[a]	7.72	5.27	
C19		8.44,6.95	7.74	5.19	5.55
A20		–	8.00	6.91	5.92
G21	12.60	–	7.10	–	5.56
A22	–		7.76	7.15	5.99
C23	15.55	10.61,9.17	7.76	5.77	5.85
U25		–	7.89	5.99	6.05
G26	12.93	–	7.79	–	5.94
G27	13.02	–	7.58	–	5.95
G28	13.56	–	7.43	–	5.87
C29		8.37	7.59	5.17	5.54
C30			7.69	5.55	
U31		–	7.74	5.84	
G32		–		–	
G33		–		–	
G34		–	7.94	–	
A35		–	8.42	8.20	5.98
G36	13.25	–	7.45	–	5.55
C37		8.88,7.03	7.89	5.23	5.55
C38		8.56,7.01	7.88	5.45	5.63
C39		8.51,7.42	7.41	5.58	5.45
U40	14.11	–	8.02	5.52	5.57
C41		8.36,7.10[a]	7.94	5.67	5.50
U42	13.35	–	7.86	5.34	5.46
G43	12.05	–	7.86	–	5.77
G44	13.31	–	7.35	–	5.70
C45		8.66,7.00	7.66	5.18	
C46					

The errors for the chemical shift are ±0.01. The exchangeable proton chemical shifts were assigned at 2°C, while all the non-exchangeable data were assigned at 25°C.

[a]The chemical shift was assigned at 15°C as the peaks overlapped at 2°C.

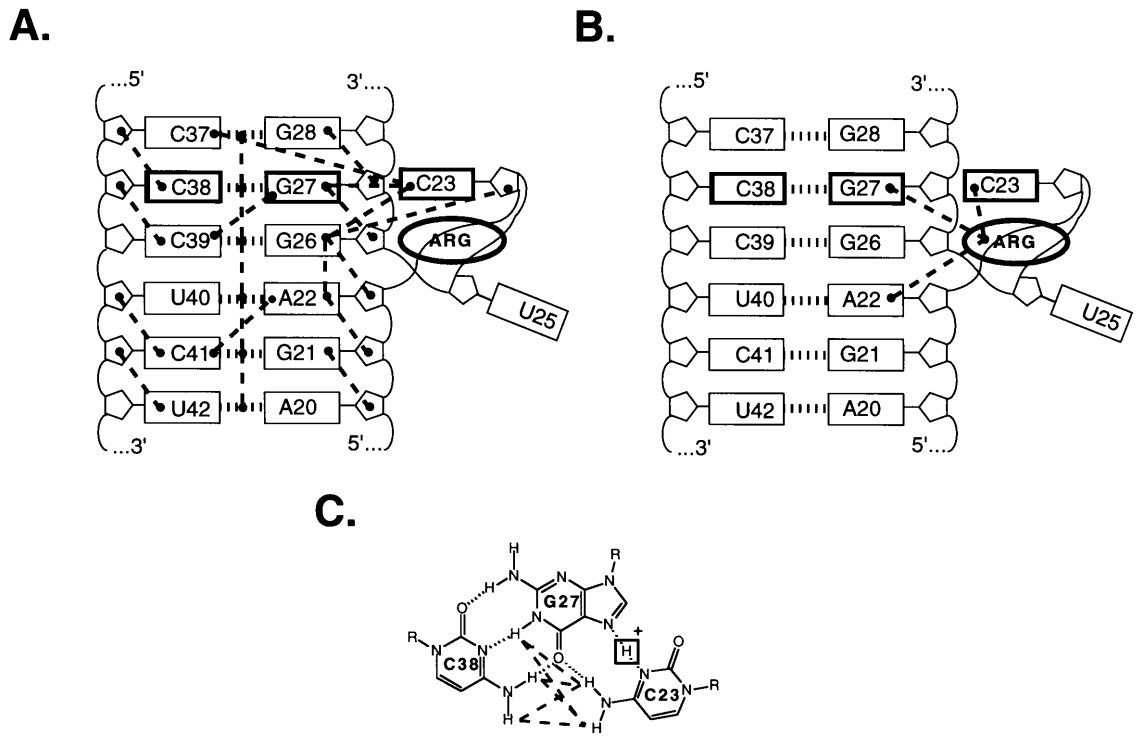

Figure 4. Summary of NMR data for the HIV-2 TAR CGC triple mutant bound to argininamide. (**A**) Internucleotide NOEs in the bulge region. Bases are indicated by rectangles and ribose sugars by pentagons. Ribose, base H8/H6 and imino protons are represented by dots within pentagons or on the outside of the base respectively. Base pair hydrogen bonding is shown by dashed lines between bases. Observed internucleotide NOEs are indicated by dashed lines. The nucleotides that form the base triple are highlighted. (**B**) The same schematic shows the intermolecular NOEs to the argininamide, indicated by dashed lines. (**C**) C38-G27·C23^{+} is shown with the key internucleotide NOEs that are observed that position C23 next to G27, shown by the dashed lines. The C23 imino proton is highlighted by the box.

Some weaker NOEs that were observed in the HIV-2 TAR–argininamide complex were not observed in the triple mutant spectra. Most of these differences can be attributed to the use of multi-dimensional heteronuclear NMR for the HIV-2 TAR assignments, which was not replicated here for the base triple mutant. Intermolecular NOEs to the argininamide guanidium protons were not detected because these resonances were broader in the HIV-2 TAR CGC spectra. Also, the weak C39(H42/H41)–Arg(Hδ) intermolecular NOEs were not observed. Spectral crowding inhibited unambiguous assignment of intermolecular NOEs between A22(H2′) and A22(H3′) and Arg(Hε), even though candidate peaks were observed.

All the observed NOEs for the triple mutant are consistent with the base triple structure of the wild-type HIV-2 TAR–argininamide complex. In fact, a similar set of intermolecular and internucleotide NOEs are observed in both HIV-2 TAR CGC and wild-type HIV-2 TAR complexes, indicating that argininamide is binding to both molecules in the same manner. Table 2 lists the critical NOEs in the bulge region that are observed in the HIV-2 TAR CGC–argininamide complex. Important intermolecular NOEs that are found in the wild-type HIV-2 TAR–argininamide complex, including NOEs between C23 H5 and argininamide Hγ and Hδ, are also observed in the triple mutant. Significantly, NOEs from C23(H42/H41) to G27(H1) and C38(H42/H41) are observed, positioning C23 near G27, as shown schematically in Figure 4.

Table 2. Intermolecular and internucleotide NOEs

Arg(Hβ)	–C39	H42	Y[a]	C23	H42/H41	–G26	H1	N
		H41	Y			–G27	H1	N
						–C37	H42	N
						–C38	H42	N
Arg(Hγ)	–A22	H8	Y				H41	N
	-C23	H5	Y			–C39	H42	N
		H6	Y				H41	N
Arg(Hδ)	–A22	H8	Y	G26	H8	–A22	H1′	Y
	–C23	H5	Y			–C23	H1′	Y
	–G27	H8	Y					Y
	–C39	H42	Y	G26	H1′	–A22	H2	Y
		H41	Y					
Arg(Hε)	–C23	H5	Y					
		H6	Y					

[a]The Y or N indicates whether the NOE was observed in the wild-type HIV-2 TAR–argininamide complex.

CONCLUSIONS

The arginine binding motif originally identified in the HIV Tat–TAR interaction has recently been observed in other contexts. A TAR bulge motif was identified in the selection of high affinity Rev binding aptamers, which gave the same chemical modification interference signature as TAR (28,29). The ensuing solution structure of the Rev peptide–aptamer complex provided unambiguous evidence for a U-A·U base triple forming an arginine binding site (30), very similar to the HIV-2 TAR structure. The r.m.s. deviation between the Rev–aptamer complex and HIV-2 TAR for the bases involved in the base triple, U-A·U, the arginine and the G-C base pair contacting the arginine is 1.12 Å. The architecture of the arginine binding site in these two different contexts is the same. A U-A·U base triple also forms an arginine binding site in the BIV Tat–TAR complex (31). The TAR bulge motif has also been identified in the Rex response element (RexRE) RNA–Rex peptide interaction by SELEX (32). Because of the sequence similarity, we expect this structure to resemble the HIV TAR–arginine interaction. Finally, small molecules that bind to the HIV TAR bulge mimic the arginine guanidium group and bind in a very similar manner as arginine-rich peptides (33). Thus the base triple in the TAR bulge may prove to be a common structural motif for arginine guanidium group recognition.

ACKNOWLEDGEMENTS

We thank J.Puglisi for inspiration and initial studies which guided this project. Some experiments were performed at the NMR center at the Francis Bitter Magnet Laboratory, which is supported by NIH grant PYRR0095. This work was also supported by a grant from the National Institutes of Health (GM39489).

REFERENCES

1 Jones,K.A. and Peterlin,B.M. (1994) *Annu. Rev. Biochem.*, **63**, 717–743.
2 Frankel,A.D. (1992) *Curr. Opin. Genet. Dev.*, **2**, 293–298.
3 Dingwall,C., Ernberg,I., Gait,M.J., Green,S.M., Heaphy,S., Karn,J., Lowe,A.D., Singh,M., Skiner,M.A. and Valerio,R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6925–6929.
4 Cordingly,M.G., LaFemina,R.L., Callahan,P.L., Condra,J.H., Sardana,V.V., Graham,D.J., Nguyen,T.M., LeGrow,K., Gotlib,L., Schlabach,A.J. and Colonno,R.J. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8985–9090.
5 Dingwall,C., Ernberg,I., Gait,M.J., Green,S.M., Heaphy,S., Karn,J., Lowe,A.D., Singh,M. and Skinner,M.A. (1990) *EMBO J.*, **9**, 4145–4153.
6 Weeks,K.M., Ampe,C., Schultz,S.C., Steitz,T.A. and Crothers,D.M. (1990) *Science*, **249**, 1281–1285.
7 Roy,S., Delling,U., Chen,C.-H., Rosen,C.A. and Sonenberg,N. (1990) *Genes Dev.*, **4**, 1365–1373.
8 Weeks,K.M. and Crothers,D.M. (1991) *Cell*, **66**, 577–588.
9 Churcher,M.J., Lamont,C., Hamy,F., Dingwall,C., Green,S.M., Lowe,A.D., Butler,P.J.G., Gait,M.J. and Karn,J. (1993) *J. Mol. Biol.*, **230**, 90–110.
10 Long,K.S. and Crothers,D.M. (1995) *Biochemistry*, **34**, 8885–8895.
11 Tao,J. and Frankel,A.D. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 2723–2726.
12 Tao,J. and Frankel,A.D. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 1571–1575.
13 Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) *Science*, **257**, 76–80.
14 Aboul-ela,F., Karn,J. and Varani,G. (1995) *J. Mol. Biol.*, **253**, 313–332.
15 Brodsky,A.S. and Williamson,J.R. (1997) *J. Mol. Biol.*, **267**, 624–639.
16 Puglisi,J.D., Chen,L., Frankel,A.D. and Williamson,J.R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3680–3684.
17 Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.*, **15**, 8783–8798.
18 Garrett,D.S., Powers,R., Gronenborn,A.M. and Clore,G.M. (1991) *J. Magn. Resonance*, **95**, 214–220.
19 Delaglio,F., Grzesiek,S., Vuister,G., Zu,G., Pfeiffer,J. and Bax,A. (1995) *J. Biomol. NMR*, **6**, 277–293.
20 Green,H. and Freeman,R. (1991) *J. Magn. Resonance*, **93**, 93–141.
21 Lippens,G., Dhalluin,C. and Wieruszeki,J.-M. (1995) *J. Biomol. NMR*, **5**, 327–331.
22 Piotto,M., Saudek,V. and Skelnár,V. (1992) *J. Biomol. NMR*, **2**, 661–665.
23 Kadkhodaei,M., Hwang,T.-L., Tang,J. and Shaka,A.J. (1993) *J. Magn. Resonance A*, **105**, 104–107.
24 Dongen,M.v., Heus,H.A., Wymenga,S.S., Marel,G.A.v.d., Boom,J.H.v. and Hilbers,C.W. (1996) *Biochemistry*, **35**, 1733–39.
25 Macaya,R., Wang,E., Schultze,P., Sklenár,V. and Feigon,J. (1992) *J. Mol. Biol.*, **225**, 755–773.
26 Santos,C.d.l., Rosen,M. and Patel,D. (1989) *J. Mol. Biol.*, **28**, 7282–7289.
27 Varani,G. and Tinoco,I.J. (1991) *Q. Rev. Biophys.*, **24**, 479–532.
28 Giver,L., Bartel,D., Zapp,M., Pawul,A., Green,M. and Ellington,A.D. (1993) *Nucleic Acids Res.*, **21**, 5509–5516.
29 Jensen,K.B., Green,L., MacDougal-Waugh,S. and Tuerk,C. (1994) *J. Mol. Biol.*, **235**, 237–247.
30 Ye,X., Gorin,A., Ellington,A.D. and Patel,D.J. (1996) *Nature Struct. Biol.*, **3**, 1026–1033.
31 Ye,X., Kumar,R.A. and Patel,D.J. (1995) *Chem. Biol.*, **2**, 827–840.
32 Baskerville,S., Zapp,M. and Ellington,A. (1995) *J. Virol.*, **69**, 7559–69.
33 Hamy,F., Felder,E.R., Heizmann,G., Lazdins,J., Aboul-ela,F., Varani,G., Karn,J. and Klimkait,T. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 3548–3553.