

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

MARTIN ZACHARIAS AND PAUL J. HAGERMAN*

Department of Biochemistry, Biophysics, and Genetics, University of Colorado Health Sciences Center, Denver, CO 80262

Communicated by Robert L. Baldwin, Stanford University Medical Center, Stanford, CA, March 27, 1995

ABSTRACT The trans-activation response element (TAR) found near the 5' end of the viral RNA of the human immunodeficiency virus contains a 3-nt bulge that is recognized by the virally encoded trans-activator protein (Tat), an important mediator of transcriptional activation. Insertion of the TAR bulge into double-stranded RNA is known to result in reduced electrophoretic mobility, suggestive of a bulge-induced bend. Furthermore, NMR studies indicate that Arg causes a change in the structure of the TAR bulge, possibly reducing the bulge angle. However, neither of these effects has been quantified, nor have they been compared with the effects of the TAR–Tat interaction. Recently, an approach for the quantification of bulge-induced bends has been described in which hydrodynamic measurements, employing the method of transient electric birefringence, have yielded precise estimates for the angles of a series of RNA bulges, with the angles ranging from 7° to 93°. In the current study, transient electric birefringence measurements indicate that the TAR bulge introduces a bend of $50^\circ \pm 5^\circ$ in the absence of Mg^{2+} . Addition of Arg leads to essentially complete straightening of the helix (to $<10^\circ$) with a transition midpoint in the 1 mM range. This transition demonstrates specificity for the TAR bulge: no comparable transition was observed for U_3 or A_3 (control) bulges with differing flanking sequences. An essentially identical structural transition is observed for the Tat-derived peptide, although the transition midpoint for the latter is near 1 μ M. Finally, low concentrations of Mg^{2+} alone reduce the bend angle by $\approx 50\%$, consistent with the effects of Mg^{2+} on other pyrimidine bulges. This last observation is important in view of the fact that most previous structural/binding studies were performed in the absence of Mg^{2+} .

Trans-activation of the viral mRNA of the human immunodeficiency virus by the virally encoded trans-activator protein (Tat) requires a short stretch of RNA, designated the trans-activator response element (TAR), located near the 5' end of the viral RNA (1–3). The Tat–TAR interaction is essential for activation of transcription, prevention of premature transcription termination, and viral replication (3–6). A combination of mutational (3–14), chemical protection (14, 15), and NMR (16, 17) studies have defined the TAR element as a helical stem interrupted by a 3-base bulge and capped by a 6-base loop. The bulge and flanking base pairs are essential for Tat binding (3), whereas the loop region is important for trans-activation but not for Tat binding (9, 11, 18). Peptides containing the basic Arg-rich region of Tat are able to bind specifically and with high affinity to the TAR bulge region (11, 12, 19–21); replacement of all but one Arg within the binding domain still facilitates specific binding and a functional Tat protein (22). Finally, Arg is able to bind specifically, albeit with low affinity, to the TAR bulge (16, 23, 24). These observations, in agree-

ment, suggest that Arg constitutes an element of specificity for the Tat–TAR interaction. However, since Tat (or Tat-derived peptides) bind to the TAR bulge several orders of magnitude more tightly than does Arg, it is unclear whether additional specific contacts are made between the peptide and the TAR bulge.

One approach for evaluating the participation of Arg in the specificity of the Tat–TAR interaction is to ask whether the global conformational changes in TAR are similar for the Arg–TAR and Tat–peptide–TAR interactions. The presence of the TAR bulge in otherwise double-stranded (ds)RNA results in reduced electrophoretic mobility of the bulge-containing molecules (25), indicative of a bulge-induced bend. Moreover, NMR studies of the free TAR bulge and its complex with Arg suggest that the bulge undergoes a conformational shift upon Arg binding that corresponds to a looping out of the bulge nucleotides and straightening of the flanking helices (16, 17). Alternative models have been proposed in which Tat facilitates a widening of the major groove near the bulge to accommodate the bulge bases (14, 26). However, it is not known whether changes in the bulge-induced bend are similar for Arg and Tat peptide.

To examine this issue, we have employed a combination of gel electrophoretic and transient electric birefringence (TEB) measurements to first quantify the bend angle introduced between flanking helices by the TAR bulge, followed by an examination of the influence of Arg (and Arg derivatives) and various Tat-derived peptides on this angle. These measurements indicate that the angle introduced by the TAR bulge is substantial ($50^\circ \pm 5^\circ$) and that both Arg and Tat-derived peptide lead to essentially complete straightening of the helix in a manner that displays both sequence specificity for the bulge nucleotides and Arg specificity.

MATERIALS AND METHODS

Preparation of RNA Expression Plasmids. Two synthetic dsDNA oligomers, T1 (5'-AGCTGCCAGAGAGCTC/5'-AGCTGAGCTCTCTGGC) and T2 (5'-AGCTGAGCTCAGATCTGGC/5'-AGCTGCCAGATCTGAGCTC) with 5'-AGCT (*Hind*III compatible) overhangs, were inserted into the *Hind*III sites of the RNA expression plasmids pGJ122A and pGJ122B (27, 28). The resulting plasmids were used to transform *Escherichia coli* DH5 α cells (Promega). Plasmids containing inserts were sequenced by using a DNA primer, 5'-GACGATCCCGGAGAT, that binds upstream of the T7 promoter, thus allowing the entire region between the T7 promoter and the *Sma* I site (defining the 3' terminus of the

Abbreviations: TAR, trans-activation response element; TEB, transient electric birefringence; ArgN, argininamide; Gdn, guanidine; ds, double stranded.

*To whom reprint requests should be addressed at: Department of Biochemistry, Biophysics, and Genetics B-121, University of Colorado Health Sciences Center, 4200 East 9th Avenue, Denver, CO 80262.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

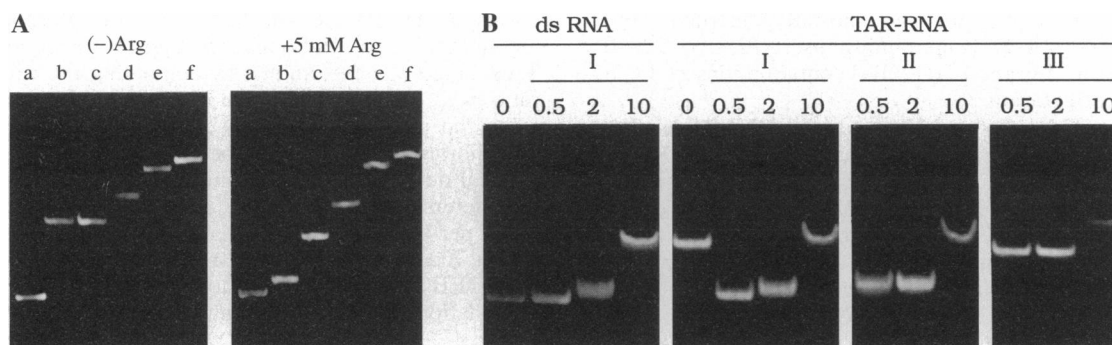


FIG. 2. (A) Gel electrophoresis of bulge-containing RNA molecules in the absence or presence of 5 mM Arg in the gel running buffer. Lanes: a, linear (152 bp) dsRNA; b, TAR bulge-containing RNA (TAR RNA); c, U₃ RNA; d, U₄ RNA; e, U₅ RNA; f, U₆ RNA. (B) Gel electrophoresis of dsRNA or TAR RNA after preincubation of the RNAs with peptide I (Tat-derived peptide), II (K₃RK₅), or III (K₉), at the indicated peptide concentrations (in μM). All four parts of B are from contiguous sections on a single gel, without vertical repositioning; the photograph has been sectioned to emphasize individual peptide-RNA combinations.

instance, the bend-containing molecule rotates more rapidly than does its linear counterpart, and the bend angle is extracted from the ratio of the rotational decay times of the linear and bent species. In the current instance, the birefringence decay curve for the TAR RNA is substantially different from that of the linear control RNA (Fig. 3A). In particular, whereas the terminal decay component of the linear species dominates the decay process ($\approx 90\%$), characteristic of nonbent dsRNA 150–180 bp long, the terminal (slower) phase represents only $\approx 50\%$ of the decay process for the TAR RNA, characteristic of interhelix bends of $40\text{--}60^\circ$ (27). Indeed, direct analysis of the ratio of the terminal decay times (0.82 ± 0.03) yields a bend angle for the TAR bulge of $50^\circ \pm 5^\circ$ in the absence of Arg or peptide (and also Mg^{2+} ; see below). This is comparable to the value of 55° obtained previously for a U₃ bulge (27).

The Bend Created by the TAR Bulge Is Essentially Eliminated by Arg or the Tat Peptide. Upon addition of Arg or Gdn (≈ 1 mM) to TAR RNA, a conformational transition is observed in which the birefringence decay behavior of the bulge-containing heteroduplex becomes indistinguishable from that of the linear control (Fig. 3B). In particular, the terminal decay component now dominates, and the terminal decay time is characteristic of fully dsRNA. Thus, within the limits of uncertainty of the TEB method, any residual bulge angle would be quite small ($<10^\circ$), and the bulge region would be relatively rigid. The same effect is observed for ArgN, albeit at much lower concentrations (≈ 0.02 mM). The observation (Fig. 2A) that the gel mobility of TAR RNA in the presence of Arg is still slightly lower than that of the linear species could

be indicative of a residual fixed distortion or of a bulge region that is more easily distortable than the surrounding helix. Finally, addition of the Arg-containing peptides (I or II) effected the same change as observed for Arg, namely, the conversion of TAR RNA to a form that is indistinguishable from the linear species (residual angle $<10^\circ$; Fig. 3C). Thus, within the limits of the current methods, Arg and Arg-containing basic peptides induce the same structural transition with respect to the relative positioning of the helix segments that flank the bulge.

The Conformational Transition Demonstrates Specificity for Bulge Sequence and Arg. Since both the terminal decay times and fractional amplitudes are distinctly different for the two bulge conformations, both can be used to monitor the conformational transition in TAR RNA as a function of ligand concentration. In Fig. 4, a series of binding isotherms are displayed for the conformational transition in TAR RNA induced by various ligands. The transition midpoints of these curves were used to estimate the (equilibrium) dissociation constants (K_d) for Arg and Gdn, both ≈ 1 mM. ArgN binds more tightly ($K_d \approx 0.02$ mM; TAR RNA concentration ≈ 1 μM). These effects are not simply a consequence of charge, since neither Lys nor ammonium ion induced any conformational change at concentrations up to 10 mM.

A corresponding set of titrations for the basic peptides is also displayed in Fig. 4. For the Arg-containing peptides (I and II), nearly identical transitions are observed, with midpoints near 1 μM . Since these binding studies require TAR RNA in the 1 μM range for adequate precision, the K_d values for peptides I

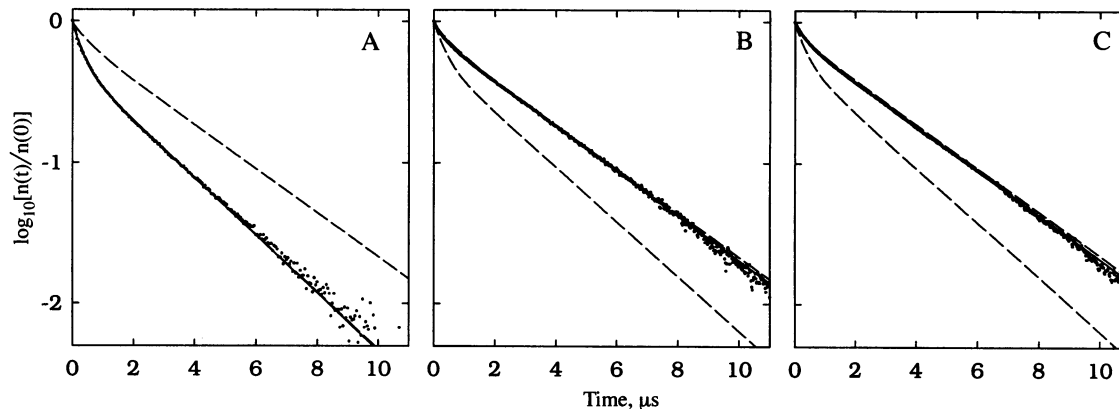


FIG. 3. TEB decay curves of the TAR RNA molecule in the absence of ligand (A), in the presence of 5 mM Arg (B), or in the presence of 2 μM (Tat-derived) peptide I (C). Upper dashed curve, position of the decay curve for the linear control RNA; lower dashed curve, position of the decay curve for the U₃ RNA; solid lines, best-fit curves to the data points for TAR RNA. In A, the solid and lower dashed lines are essentially superimposed. Note that the presence of Arg or peptide I results in a slight reduction in the bulge angle for the U₃ RNA (B and C), consistent with the gel results. $n(t)$ is the sample birefringence at time t , and $n(0)$ is the birefringence at $t = 0$.

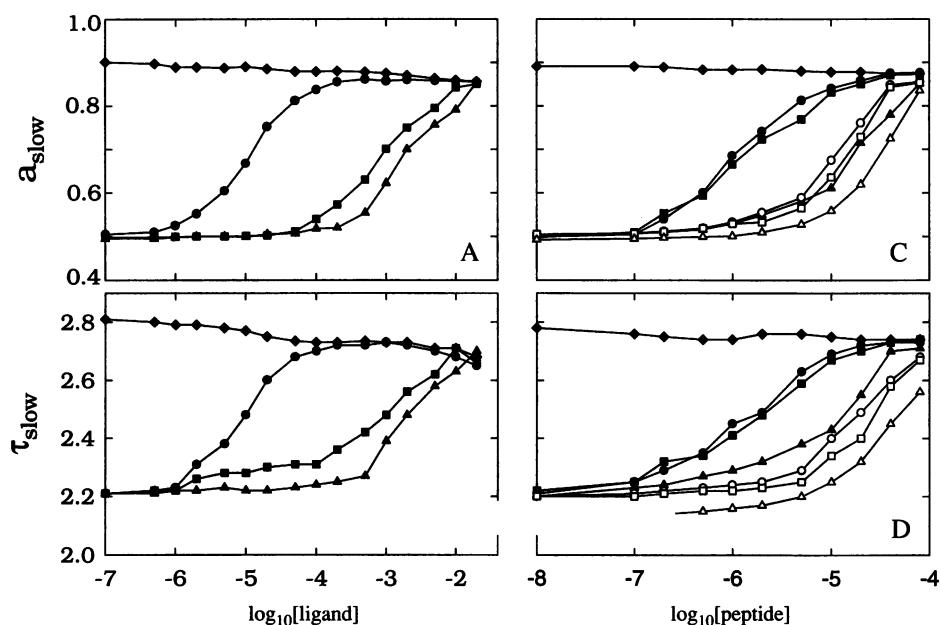


FIG. 4. Birefringence decay data for the linear and TAR RNA species as a function of the logarithm of the ligand or peptide concentration (M). (A and C) Fractional amplitudes of the terminal decay component (a_{slow}). (B and D) Corresponding decay times (μs). (A and B) \blacklozenge , Linear RNA + Arg; \blacksquare , TAR RNA + Arg; \bullet , TAR RNA + ArgN; \blacktriangle , TAR RNA + Gdn. (C and D) \blacklozenge , Linear RNA + peptide I; \bullet , TAR RNA + peptide I; \blacksquare , TAR RNA + peptide II; \blacktriangle , TAR RNA + peptide III; \circ , U₃ RNA + peptide I; \square , U₃ RNA + peptide III; \triangle , A₃ RNA + peptide I.

and II could be much smaller than 1 μM . In fact, it has been demonstrated by gel-shift methods that the K_d values for those peptides are in the nanomolar range (922). Thus, for peptides I and II, the transition midpoints represent a strict upper bound for K_d . For peptide III (Lys₉), straightening of the TAR RNA is also observed, albeit at 10–20 μM peptide. This latter effect is nonspecific, since it occurs for U₃ and A₃ bulges in different sequence contexts (Fig. 1) at the same peptide concentration and at approximately the same concentration range in which peptides I and II induce a conformational shift in the U₃ and A₃ bulges (Fig. 4). Thus, under the relatively low salt conditions used for the TEB measurements, the guanidinium moiety appears to decrease K_d by at least 10- to 15-fold over nonspecific ammonium ion (charge) effects, whether or not it is in the peptide.

Finally, it should be noted that none of the peptides has any significant influence on the structure of the linear control. This observation demonstrates that the substantial reduction in mobility of all of the RNA species in the presence of 10 μM peptide is, in effect, an artifact of the gel system (one that is expected to be quite general), most likely due to partial charge neutralization of the RNA–peptide complexes. This observation underscores the need to exercise care when interpreting gel shifts for protein–nucleic acid complexes in terms of induced structural transitions in the RNA.

The Influence of Mg^{2+} on the Conformation of the TAR Bulge. In a recent investigation of a series of U_n and A_n bulges using TEB (27), it was observed that Mg^{2+} induced a partial straightening of the RNAs possessing U_n (but not A_n) bulges. A similar effect is observed for the TAR bulge (Fig. 5), where Mg^{2+} shifts the terminal amplitude from $\approx 45\%$ to $\approx 65\%$. Although we have not observed complete straightening of the TAR RNA for Mg^{2+} up to ≈ 2 mM ($\theta = 36^\circ \pm 5^\circ$ at 1 mM Mg^{2+}), additional straightening at higher Mg^{2+} concentrations cannot be ruled out. It is apparent from Fig. 5B that although Mg^{2+} does reduce the bend angle, the transition midpoint for Arg binding is not shifted significantly. Although we have not investigated this effect in detail, one would expect that the presence of Mg^{2+} would reduce the net (electrostatic) attraction between Arg and RNA. However, since Mg^{2+} also partially straightens TAR RNA, this effect may represent a compensatory (favorable) contribution to Arg binding.

DISCUSSION

The current study has demonstrated that the TAR bulge induces a bend in the helix axis of $\approx 50^\circ$ in the absence of Mg^{2+} .

Upon binding of Arg or Arg-containing basic peptides, the TAR bulge undergoes a conformational transition in which the bend angle is reduced to near zero ($<10^\circ$). These observations are consistent with an NMR-derived model (16, 17) in which Arg induces a straightening of the TAR bulge via looping out of the unpaired bases. In the bound form, the TAR-bulge RNA displays hydrodynamic behavior that is indistinguishable from that of a fully duplex RNA, suggesting that not only is the bend

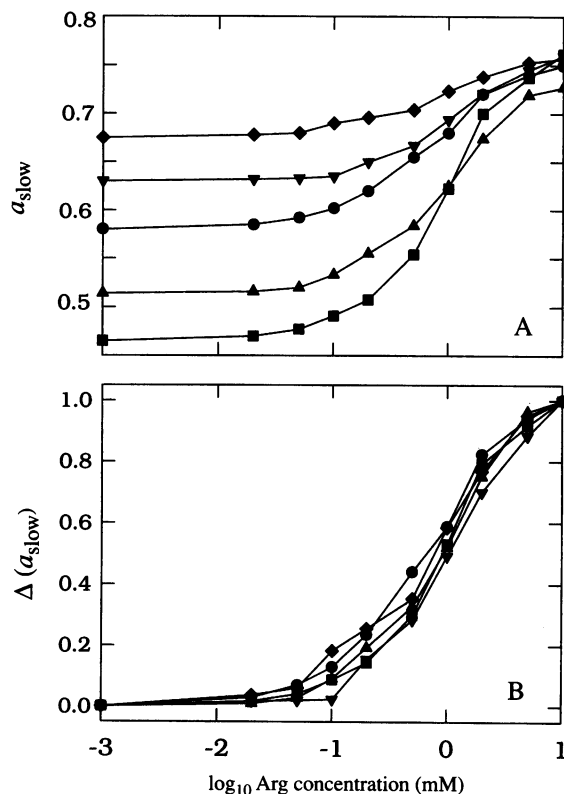


FIG. 5. Plots of a_{slow} (A) as a function of the logarithm of Arg concentration (mM) as an indicator of Arg binding, for the following MgCl_2 concentrations: \blacksquare , 0.0 mM; \blacktriangle , 0.075 mM; \bullet , 0.15 mM; \blacktriangledown , 0.30 mM; \blacklozenge , 0.90 mM. (B) Normalized fractional change in a_{slow} for the curves in A: $\Delta a_{\text{slow}} = \{a_{\text{slow}}([\text{Arg}]) - a_{\text{slow}}(10 \text{ mM})\} / \{a_{\text{slow}}(0) - a_{\text{slow}}(10 \text{ mM})\}$.

angle largely eliminated but also the complexed TAR bulge is not substantially more flexible than the surrounding helix. It should be noted that while we cannot exclude an increase in flexibility as a component of the apparent bulge-induced angle, the strong dependence of the mobility of RNA molecules possessing the TAR bulge and a second bulge on the phasing between the bulges (25) indicates that the TAR bulge is relatively rigid.

The electrophoretic mobility of TAR RNA is increased (toward the value of the linear control) in the presence of Arg (or its derivatives) or Arg-containing peptides (I or II). The effect is specific for the TAR bulge at peptide concentrations up to 2 μ M. For higher peptide concentrations, the influence of peptide binding on mobility becomes nonspecific and of opposite sign: linear and bulge-containing RNAs are all retarded, and peptides containing or lacking Arg are equally effective in reducing mobility. This latter effect is, therefore, likely due to partial charge neutralization of the RNA by the bound basic peptide. In this regard, it should be noted that in previous gel studies of the binding of Tat-derived peptides to RNA, complex formation was always associated with reductions in mobility (10–12, 20–23). Since the previous studies employed much smaller RNA molecules (\approx 30 nt), it is likely that the partial charge neutralization upon peptide binding was the dominant factor in the mobility shift, even for specific binding. In contrast, the current investigation has employed much larger RNAs (\approx 300 nt), where the conformational shift would be expected to outweigh the charge-neutralization effect of the single specifically bound peptide. Only at higher peptide concentrations, where nonspecific (multiple) binding is occurring, is the charge effect apparent. This result illustrates the difficulties inherent in the interpretation of direction or magnitude of a mobility shift in a gel electrophoresis experiment, where changes in both net charge and conformation are important. In contrast, since the rotational decay times are largely independent of the changes in net charge under conditions of approximately constant bulk ion concentration (e.g., Fig. 4 C and D; linear RNA), the TEB approach can be used to characterize the conformational change, *per se*.

The current study has provided additional evidence that Arg represents a significant element of specificity in the interaction between Tat and TAR. The fact that the K_d of Arg for TAR is in the millimolar range, whereas Tat (and Tat-derived peptides) are in the nanomolar range, may reflect (in part) the additional nonspecific (electrostatic) attraction of the basic peptide for RNA. Thus, Arg could be thought of as contributing at least 1 or 2 kcal (1 cal = 4.184 J) of free energy toward the specificity of interaction. Of course, the current observations do not address the possibility of other specific interactions in the Tat–TAR interaction; extensive mutational and chemical modification/interference studies have identified additional specific contacts between Tat and TAR (8–15). However, it is clear that the global conformational transition (helix repositioning) induced by Tat-derived peptide is nearly identical to the transition induced by Arg alone or by Arg in a Lys₈ context. Thus, Arg could be a principal participant in the repositioning of the TAR hairpin, perhaps as an antecedent to Tat-mediated trans-activation.

We thank E. Vacano for carrying out the hydrodynamic computations using DIFFROT and J. Mills for synthesizing the DNA oligomers employed in this work. This work was supported by a grant from the National Institutes of Health (GM 35305 to P.J.H.).

1. Kao, S.-Y., Calman, A. F., Lueiw, P. A. & Peterlin, B. M. (1987) *Nature (London)* **330**, 489–493.
2. Berkhout, B., Silverman, R. H. & Jeang, K.-T. (1989) *Cell* **59**, 273–282.
3. Roy, S., Delling, U., Chen, C.-H., Rosen, A. & Sonenberg, N. (1990) *Genes Dev.* **4**, 1365–1373.
4. Selby, M. J., Bain, B. M., Lueiw, P. & Peterlin, B. M. (1989) *Genes Dev.* **3**, 547–558.
5. Laspia, M. F., Rice, A. P. & Mathews, M. B. (1989) *Cell* **59**, 283–292.
6. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heapy, S., Karn, J., Lowe, A. D., Singh, M. & Skinner, M. A. (1990) *EMBO J.* **9**, 4145–4153.
7. Muessing, M. A., Smith, D. H. & Capon, D. J. (1987) *Cell* **48**, 691–701.
8. Hauber, J. & Cullen, B. R. (1988) *J. Virol.* **62**, 673–679.
9. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heapy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A. & Vallerio, R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6925–6929.
10. Roy, S., Parkin, N. T., Rosen, C., Itovich, J. & Sonenberg, N. (1990) *J. Virol.* **64**, 1402–1406.
11. Cordingley, 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. & Colonna, R. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8985–8989.
12. Delling, U., Reid, L. S., Barnett, R. W., Ma, M. Y. X., Climie, S., Sumner-Smith, M. & Sonenberg, N. (1992) *J. Virol.* **66**, 3018–3025.
13. Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, J. G., Gait, M. J. & Karn, J. (1993) *J. Mol. Biol.* **230**, 90–110.
14. Weeks, K. M. & Crothers, D. M. (1991) *Cell* **66**, 577–588.
15. Hamy, F., Asseline, U., Grasby, J., Iawi, S., Pritchard, C., Slim, G., Butler, J. G., Karn, J. & Gait, M. J. (1993) *J. Mol. Biol.* **230**, 111–123.
16. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D. & Williamson, J. R. (1992) *Science* **257**, 76–80.
17. Puglisi, J. D., Chen, L., Frankel, A. D. & Williamson, J. R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3680–3684.
18. Feng, S. & Holland, E. C. (1988) *Nature (London)* **334**, 165–167.
19. Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A. & Crothers, D. M. (1990) *Science* **249**, 1281–1285.
20. Calnan, B. J., Biancalana, S., Hudson, D. & Frankel, A. D. (1990) *Genes Dev.* **5**, 201–210.
21. Tao, J. & Frankel, A. D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1571–1575.
22. Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D. & Frankel, A. (1991) *Science* **252**, 1167–1171.
23. Tao, J. & Frankel, A. D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2723–2726.
24. Tan, R. & Frankel, A. D. (1992) *Biochemistry* **31**, 10288–10294.
25. Riordan, F. A., Bhattacharyya, A., McAteer, S. & Lilley, D. M. J. (1992) *J. Mol. Biol.* **226**, 305–310.
26. Loret, E. P., Georgel, P., Johnson, W. C. & Ho, P. S. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9734–9738.
27. Zacharias, M. & Hagerman, P. J. (1995) *J. Mol. Biol.* **247**, 486–500.
28. Amiri, K. M. A. & Hagerman, P. J. (1994) *Biochemistry* **33**, 13172–13177.
29. Gast, F.-U. & Hagerman, P. J. (1991) *Biochemistry* **30**, 4268–4277.
30. Gast, F.-U., Amiri, K. M. A. & Hagerman, P. J. (1994) *Biochemistry* **33**, 1788–1796.
31. Shen, Z. & Hagerman, P. J. (1994) *J. Mol. Biol.* **241**, 415–430.
32. Press, W. H., Vetterling, W. T., Teukolsky, S. A. & Flannery, B. P. (1992) *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge, Univ. Press, Cambridge, U.K.), pp. 678–683.
33. Hagerman, P. J. & Zimm, B. (1981) *Biopolymers* **20**, 1481–1502.
34. Kebbekus, P., Draper, D. E. & Hagerman, P. J. (1995) *Biochemistry* **34**, 4354–4357.