Refolded HIV-1 tat Protein Protects both Bulge and Loop Nucleotides in TAR RNA from Ribonucleolytic Cleavage†

J. Wade Harper* and Naomi J. Logsdon

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030

Received March 14, 1991; Revised Manuscript Received May 29, 1991

ABSTRACT: Substantial evidence indicates that HIV-1 trans-activation by tat protein is mediated through the TAR RNA element. This RNA forms a stem–loop structure containing a three-nucleotide bulge and a six-nucleotide loop. Previous mutagenic analysis of TAR indicates that the bulge residues and a 4 bp segment of the stem constitute, in part, the tat binding site. However, there appears to be no sequence-specific contribution of the six-base loop. We have employed a ribonuclease protection technique to explore the interaction of tat with single-stranded regions of TAR. The results indicate that tat interacts with both the bulge and loop regions of TAR. Treatment of TAR RNA with RNase A results in cleavage at U^{23} and U^{31}, located in the bulge and loop regions, respectively. High concentrations (~ 2 μM) of *Escherichia coli* derived tat protein, prepared by standard procedures, gave complete protection of TAR RNA from RNase A cleavage. However, under these conditions, truncated TAR derivatives in which no stem–loop structure is expected to form were also protected, indicating nonspecific binding. In order to obtain a tat preparation with enhanced specificity toward TAR RNA, methods were developed for refolding the recombinant protein. This treatment enhanced the affinity of tat for TAR by $\sim$30-fold [K_d(apparent) $<$ 25 nM] and markedly increased its specificity for the TAR. Again, tat protected TAR RNA from RNase A cleavage at both U^{23} and U^{31}. Protection was also observed with RNase T1 which cleaves TAR RNA at three G residues in the six-base loop. Taken together with mutagenic studies, the data suggest that in addition to making sequence-specific interactions with bulge nucleotides, tat also interacts with components in the six-base loop but in a largely sequence-independent fashion.

There is currently widespread interest in the structure and mechanism of action of the HIV-1 trans-activator protein tat.[1] This basic protein of 86 amino acids can increase the expression of HIV LTR directed genes by as much as 1000-fold and is required for efficient virus production [for recent reviews, see Cullen (1990), Rosen and Pavlakis (1990), Cullen and Greene (1989), and Sharp and Marciniak (1989)]. While the precise mode of action of tat is not established, studies both in vitro (Marciniak et al., 1990) and in tissue culture cells (Cullen, 1986; Laspia et al., 1989; Rice & Mathews, 1988) indicate that the major effect of tat is to increase the efficiency of transcription elongation from the HIV promoter. This may lead secondarily to an increase in the apparent rate of transcription initiation (Cullen, 1990).

Initial experiments designed to elaborate the mechanism of trans-activation employed genetic approaches to identify cis-acting regulatory sequences in the HIV long terminal repeat (Hauber & Cullen, 1988; Selby et al., 1989; Feng & Holland, 1988; Berkhout & Jeang, 1989; Roy et al., 1990). It is now generally accepted that tat function is mediated in part through its interactions with an RNA element (termed TAR) located at the 5′ termini of HIV mRNA transcripts (nucleotides +19 to +42 relative to the site of initiation) [see Cullen (1990) for a recent review]. TAR RNA adopts a stem–loop structure (Muesing et al., 1987) containing a three-base bulge and a six-base loop (Figure 1). More recent studies employing chimeric tat proteins containing heterologous RNA (Selby & Peterlin, 1990; Southgate et al., 1990) or DNA (Berkhout et al., 1990) binding domains suggest that the tat–TAR interaction is not itself required for trans-activation but that this interaction may serve to position that tat protein such that it can interact with other proteins in the transcription complex.

Gel mobility assays have been used to map the nucleotides in TAR RNA important for recognition by bacterially produced tat protein and its fragments (Dingwall et al., 1989; Cordingley et al., 1990; Roy et al., 1990; Weeks et al., 1990; Calnan et al., 1991). On the basis of these studies, TAR RNA can be operationally divided into three functional elements (Figure 1). The 4 bp helical region linking the three-base bulge and the six-base loop appears to provide a structural framework for proper presentation of nucleotides involved in sequence-specific contacts with tat (Cullen, 1990). Single-base mutations in this region reduce the apparent affinity of tat for TAR RNA, but binding is rescued by compensatory mutations which regenerate base pairing (Dingwall et al., 1989; Roy et al., 1990). Presumably, interactions with tat in this region are largely sequence-independent. The major sequence-specific binding determinant identified thus far is U^{23}, which is located in the three-base bulge; mutation to A or C diminishes binding substantially (Roy et al., 1990; Cordingley et al., 1990; Calnan et al., 1991). While a variety of mutations in the six-base loop affect trans-activation in cells (Berkhout & Jeang, 1989; Feng & Holland, 1988), nucleotide replacements have little effect on tat binding in vitro (Roy et al., 1990; Cordingley et al.,

† This work was supported by an American Cancer Society Junior Faculty Award (J.W.H.) and by NIH Grant AI29833.

* Address correspondence to this author.

[1] Abbreviations: TAR, trans-activator responsive element; tat, trans-activator of HIV gene expression; RNase, ribonuclease; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane; HPLC, high-performance liquid chromatography; DEAE, diethylaminoethyl; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid.

0006-2960/91/0430-8060$02.50/0

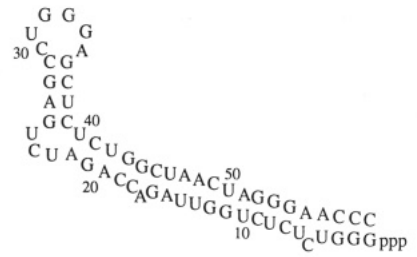

FIGURE 1: Proposed secondary structure of the TAR RNA element from HIV-1. The four structural elements are (1) the double-stranded region encompassing nucleotides 1–22 and 40–59, (2) the three-base bulge, (3) the six-base loop, and (4) the four-base helix linking the bulge and loop.

1990; Calnan et al., 1991). This has led to the conclusion that TAR loop sequences play no role in tat binding. While a cellular protein has been identified which specifically binds the TAR loop (Marciniak et al., 1990), its role in the transactivation process remains unknown. Efforts to probe the structure of the tat–TAR complex in greater detail have been hindered by the fact that tat protein expressed in bacteria tends to form aggregates which bind nucleic acids nonspecifically; typically less than 1% of the expressed protein binds TAR tightly (Cullen, 1990; Roy et al., 1990; Cordingley et al., 1990; Dingwall et al., 1989). This is largely a result of the seven cysteine residues in tat which can form intermolecular disulfide bonds and bind divalent metals (Frankel et al., 1988).

In the present study, we have used an alternative approach—RNase protection—to examine the interaction of tat and TAR. One advantage of this approach over those previously employed is that discrimination between bound and unbound RNA occurs in solution, thereby avoiding the intrusive influence inherent in gel mobility or filter binding assays. The results indicate that in addition to interacting with the bulge region, tat also interacts with components in the six-nucleotide loop, but apparently in a sequence-independent fashion. During the course of this work, a method was developed for producing tat preparations in which a much larger proportion of the protein is capable of binding TAR RNA specifically, as compared with methods employed previously (Frankel et al., 1988; Dingwall et al., 1989).

Materials and Methods

Materials. Ribonucleotide triphosphates were from Pharmacia/LKB. T7 RNA polymerase was prepared from *Escherichia coli* strain BL21 containing plasmid pAR1219 (Davanloo et al., 1985) (generously provided by Dr. John Dunn and Dr. William Studier) by using standard procedures (King et al, 1986). Plasmid pLTR-tatIII (Vogel et al., 1988) was provided by Dr. Gilbert Jay. Synthetic oligonucleotides were purified by reverse-phase HPLC using a Vydac C4 column and triethylammonium acetate/acetonitrile gradients. All aqueous solutions for RNA and tat protein were prepared by using DEPC-treated water or were passed through Sep-Pak C18 cartridges (Waters Associates) to remove adventitious RNases. Ultrapure urea (U.S. Biochemicals) was deionized with AG501-X8(D) ion-exchange resin (Bio-Rad) prior to use. [5′-^{32}P]pGp suitable for 3′ modification of RNA was prepared from 3′-GMP and [γ-^{32}P]ATP by using T4 polynucleotide kinase as described (Owens & Haley, 1984) and purified on a Waters Sep-pak C18 cartridge using 50 mM triethylammonium acetate (pH 7)/acetonitrile solvent systems.

Construction of T7–TAR Transcription Vectors. The transcription template for synthesis of TAR RNA (Figure 2A) was generated from pLTR-tatIII DNA by using the polymerase chain reaction. The 44 bp sense-strand primer contained

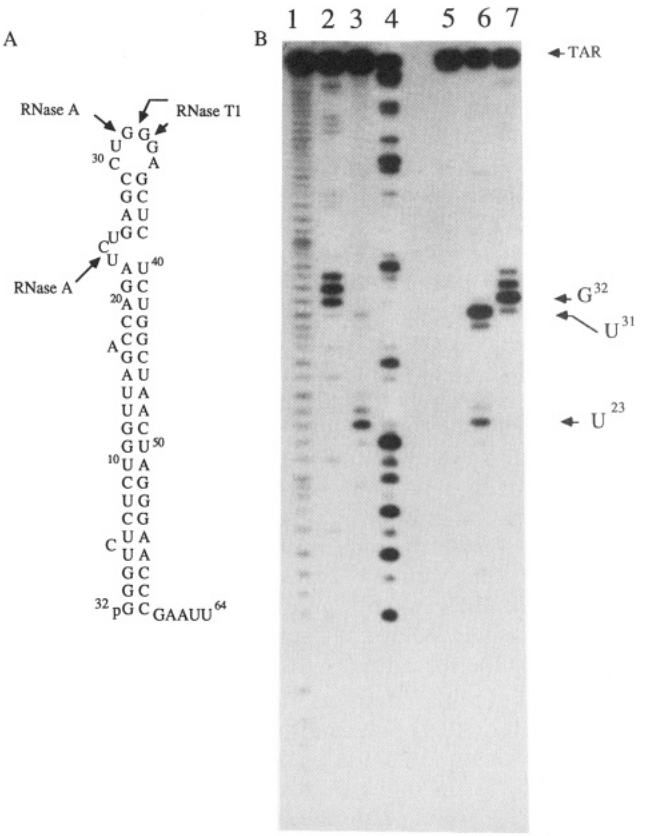

FIGURE 2: (A) Structure of the TAR RNA element used in this study. Nucleotides 1–59 correspond to the TAR sequence from HIV-1, while the 3′-extension GAAUU is derived from vector sequences. The major sites of cleavage by RNase A and RNase T1 are indicated. (B) Identification of RNase cleavage sites in TAR RNA. 5′-Labeled TAR RNA was incubated either with 1 pM RNase A (lane 6), with 25 units/mL RNase T1 (lane 7), or without added enzyme (lane 5) as described under Materials and Methods. Enzymatic sequencing reactions were performed in parallel under denaturing conditions: (lane 1) alkaline treatment; (lane 2) RNase T1; (lane 3) RNase A; (lane 4) RNase U2. Samples [(1–2) × 10^4 cpm] were subjected to electrophoresis on a 10% polyacrylamide/urea gel, and the gel was autoradiographed at −80 °C with an intensifying screen. Reactions performed using 0.1 pM RNase A gave TAR cleavage patterns which were indistinguishable from that observed with 1 pM RNase, indicating that the protection assays were performed under "single-hit" conditions.

a 17 bp T7 RNA polymerase promoter linked to the first 18 nucleotides of the TAR sequence, while the 26 bp anti-sense primer was designed to hybridize to the last 18 bp of DNA sequence encoding the TAR stem–loop. Both oligonucleotides were equipped on their 5′ termini with recognition sequences for *Eco*RI. Amplification was accomplished by using AmpliTaq DNA polymerase (Cetus/Perkin Elmer) employing conditions recommended by the supplier. Amplified DNA was digested with *Eco*RI and cloned into the *Eco*RI site of pUC19. The sequences of plasmids used for transcription were confirmed by double-stranded sequencing using T7 DNA polymerase (U.S. Biochemicals). Prior to use in transcription reactions, plasmid DNA was digested with *Eco*RI to release the T7–TAR transcription unit, and the DNA was purified by phenol extraction and NH_4OAc/ethanol precipitation.

RNA Synthesis and Modification. As TAR RNA initiated with GTP is not amenable to standard dephosphorylation/kinase strategies for 5′-labeling, a modified procedure was required for synthesis of 5′-^{32}P-labeled TAR RNAs. The modification makes use of the fact that guanosine (at [S] ≫ K_m) can effectively compete with GTP in the initiation of transcription by T7 RNA polymerase [Martin & Coleman, 1989; also see Sampson & Uhlenbeck (1988)]. Thus, the

predominant RNA product (>90%) contains a 5′-hydroxyl which can be phosphorylated directly by T4 polynucleotide kinase. Transcription reactions were performed in 40 mM Tris-HCl (pH 8.0) containing 8 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 50 mM NaCl, 3 mM ATP/CTP/UTP, 250 μM GTP, 2.4 mM guanosine (Sigma), 73 nM human placental ribonuclease inhibitor (Promega), 700 nM template DNA, and 1.2 μM T7 RNA polymerase at 42 °C (100–200-μL total volume). After 4 h, reaction mixtures were adjusted to 50 mM in EDTA, centrifuged 10 min at 14000*g*, and purified to apparent homogeneity by anion-exchange HPLC on a Waters DEAE-PAK column using a sodium acetate/NaCl gradient system at pH 7. Fractions (0.2 mL) containing RNA were immediately precipitated by addition of 2.5 volumes of ethanol (−20 °C), and the dried pellets were subsequently dissolved in 10 mM Tris, pH 8. Full-length TAR RNA eluted as a single peak and was well resolved from plasmid DNA as assessed by agarose gel electrophoresis (3% NuSieve). Enzymatic sequencing of HPLC-purified RNA (after 3′-labeling with [5′-^{32}P]pGp and RNA ligase) demonstrated that the run-off transcripts generated from these templates contain the first 59 nucleotides of the TAR element followed by the sequence GAAUU.

5′-Labeling with ^{32}P was carried out as follows: HPLC-purified RNAs (0.02–0.2 nmol) were treated with T4 polynucleotide kinase (30 units) [in 50 mM Tris·HCl (pH 9.5), 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT, 0.1 mM spermidine, 5% glycerol, and 2.2 μM [γ-^{32}P]ATP (3500 Ci/mmol)] for 30 min at 37 °C. RNA was subjected to two cycles of NaOAc/ethanol precipitation and dissolved in 200 μL of 50 mM NaOAc (pH 7). The mixture was heated for 45 s at 98 °C, cooled for 3 min on ice, and then subjected to purification by HPLC as described above. Analysis of aliquots from peak fractions on denaturing 10% polyacrylamide gels indicated that the full-length transcript was largely resolved from smaller fragments generated during the kinase reaction. RNA was stored at −20 °C in 10 mM Tris, pH 8. The truncated TAR RNAs were synthesized as described above for TAR by using templates digested with either *Sst*I (TAR^{1-37}) or *Bst*NI (TAR^{1-31}). The products were purified on 10% polyacrylamide/urea gels by using standard procedures.

Preparation of tat Protein. HIV-1 tat protein was prepared by using the bacterial expression system described by Aldovini et al. (1986) (provided by Dr. Allan Shatzman; Smith, Kline and French) and was purified by using a combination of established procedures (Dingwall et al., 1989; Aldovini et al., 1986; Frankel et al., 1988) with modifications. The major steps included sonication of tat-containing AR120 cells, poly(ethyleneimine) precipitation (Frankel et al., 1988), solubilization of the poly(ethylenimine) pellet in 8 M urea/300 mM DTT, clarification by centrifugation (15000*g*, 20 min), chromatography on a CM-Sephadex column (Dingwall et al., 1989), dialysis against 25 mM Tris, pH 8, 2 mM DTT, 1 M NaCl, and 1 mM EDTA, and chromatography on a Vydac C4 reverse-phase HPLC column (Frankel et al., 1988; Dingwall et al., 1989) using gradients of acetonitrile/2-propanol/0.1% TFA. HPLC solvents were removed in vacuo; the protein was resuspended in 10 mM Tris (pH 8) containing 10 mM DTT and stored at −20 °C. Fractions employed for further study were >95% pure as assessed by SDS-PAGE under reducing conditions. The amino acid composition of tat (100–200 pmol) [determined by PicoTag analysis (Waters Associates)] was in excellent agreement with that expected on the basis of the primary sequence of the protein. Methods employed for refolding of tat protein are described under Results. Stock solutions of tat protein were quantitated by amino acid analysis.

Tat protein preparations were biologically active as assessed by trans-activation assays performed as described (Frankel & Pabo, 1988). Briefly, HeLa HL3T1 cells containing an integrated copy of HIV LTR-CAT (Wright et al., 1986) (50% confluent in DMEM/10% fetal bovine serum) were incubated in the presence of 50 nM tat, 0.1 mM chloroquine, and 50 μM 2-mercaptoethanol for 24 h and the media then replaced with DMEM/10% fetal bovine serum. After an additional 48 h, cells were harvested and chloramphenicol acetyltransferase assays performed as described (Frankel & Pabo, 1988). At 50 nM tat, ~500-fold trans-activation was observed compared to controls lacking tat protein, consistent with previous studies (Frankel & Pabo, 1988).

Tat–TAR Complex Formation and RNase Cleavage. Unless noted otherwise, tat–TAR complexes were formed in 20 mM Hepes (pH 7.0) containing 200 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, and 2 mM spermidine with ~1 nM 5′-^{32}P-labeled TAR RNA and the indicated quantity of tat protein (total volume was 22.5 μL). After 20 min at 23 °C, 2.5 μL of the appropriate dilution of either bovine RNase A (U.S. Biochemicals) or RNase T1 (Boehringer-Mannheim) was added, and the mixtures were incubated for 15 min at 23 °C. The concentration of RNases required to obtain partial (<10%) cleavage was determined empirically. Reactions were terminated by phenol/chloroform extractions in the presence of 25 μg of carrier tRNA (type X, Sigma). RNAs were quantitatively precipitated with NH_4OAc/ethanol, and the dried pellet was dissolved in gel loading buffer (50 mM Tris-HCl, 10 M urea, 10 mM EDTA, and 0.02% bromophenol blue/xylene cylenol) prior to analysis on 10% polyacrylamide/8 M urea gels. Electrophoresis was carried out for 2 h at 55 W prior to autoradiography. RNA sequencing ladders were generated by using an RNA sequencing kit from Pharmacia/LKB.

Ribonuclease Assays. The activity of RNase A was determined under k_{cat}/K_m conditions (Harper et al., 1988) using 0.1 mM uridylyl(3′,5′)adenosine in 20 mM Hepes (pH 7.0) containing 200 mM NaCl, 1 mM DTT, and 10 mM $MgCl_2$ in the presence and absence of 3 μM tat protein. The activity of RNase T1 on a dansyl-pppGpCpApU substrate was determined by using an HPLC assay analogous to that described previously (Shapiro et al., 1986). Details are available upon request.

Results

Probing the Structure of the tat–TAR Complex with Single-Strand-Specific RNases. RNase A (U/CpN specific) and RNase T1 (GpN specific) were used to investigate the accessibility of bulge and loop nucleotides in TAR RNA toward ribonucleolytic cleavage in the presence and absence of recombinant tat protein. On the basis of the putative structure of TAR RNA (Muesing et al., 1987), RNase A was expected to cleave the RNA in both the bulge and loop regions. In contrast, RNase T1 was expected to cleave only in the loop (Figure 2A). As shown in Figure 2B, limited digestion of HPLC-purified 5′-^{32}P-labeled TAR RNA with 1 pM RNase A (lane 6) result in the generation of two major fragments not found in control samples lacking RNase A treatment (lane 5). On the basis of RNA sequencing ladders generated in parallel (lanes 1–4), these cleavage fragments correspond to $G^1 \rightarrow U^{31}$ and $G^1 \rightarrow U^{23}$, respectively. The major site of cleavage by RNase T1 is G^{32}pG, although substantial cleavage is observed at G^{33}pG and G^{34}pA (Figure 2B, lane 7). These

results are consistent with the secondary structure expected and indicate that these enzymes can be used to examine the interaction of tat with both the bulge and loop regions of TAR RNA.

In order to assess the capacity of tat protein to protect TAR RNA from nucleolytic cleavage, TAR RNA (Figure 3A, lane 2) was incubated with varying concentrations of tat protein (0–2.5 μM) to allow for complex formation and the mixture then treated with RNase A under conditions which give only partial RNA cleavage (Figure 3A, lanes 3–9). In these initial experiments, the tat protein employed was prepared by procedures analogous to those previously described (see Materials and Methods). At the lower concentrations of tat protein employed (0.5–1.0 μM), no significant protection from cleavage was observed (lanes 4–6) compared to samples lacking tat (lane 3). However, substantial protection from cleavage at both $U^{31}pG$ and $U^{23}pC$ dinucleotides was observed at 1.5 μM tat, and virtually complete protection was observed with 2.0 and 2.5 μM tat protein (Figure 3A, lanes 7–9).[2] Similar concentration-dependent protection was observed with *Bacillus cereus* RNase. This enzyme cleaves TAR RNA at three sites: $U^{23}pC$, $C^{30}pU$, and $U^{31}pG$ (data not shown).

In order to examine the specificity of protection in greater detail, additional experiments were carried out using a truncated RNA (TAR^{1-37}). This RNA contains many of the nucleotides implicated in tat binding but is not expected to have extended regions of base-pairing. RNase A cleavage of 5′-^{32}P-labeled TAR^{1-37} (Figure 3B, lane 3) results primarily in the production of fragments terminating at U^{31}, C^{30}, and C^{29} (Figure 3B, lane 4). Even in the absence of a complete TAR stem–loop, substantial protection from cleavage was observed in the presence of 2 μM tat (Figure 3B, lane 5), the concentration of tat near that required for protection of TAR RNA (Figure 3A).

The results of several studies (Dingwall et al., 1989; Roy et al., 1990) have suggested that only a small percentage (<1%) of bacterially produced tat is capable of binding TAR RNA with high specificity. In fact, gel-shift experiments typically require a 10^3–10^4-fold molar excess of tat protein (~0.5–2 μM) over TAR RNA in order to observe partial binding. This being the case, it seemed plausible that the apparent lack of specificity for intact TAR RNA in the RNase protection experiments described above was due to the presence of substantial quantities of misfolded or aggregated tat in the preparations employed. An experiment was performed to estimate the percentage of tat protein displaying high specificity toward TAR RNA. 5′-^{32}P-labeled TAR RNA (~1 nM) was mixed with either 10 or 400 nM unlabeled HPLC-purified TAR RNA and then allowed to form complexes with 2 μM tat prior to treatment with 2 pM RNase A. With 10 nM unlabled TAR RNA, protection of 5′-^{32}P TAR was virtually complete, indicating that >11 nM tat was capable of binding TAR with high affinity. However, with 400 nM unlabeled TAR RNA, no protection was observed. Thus, between 0.5 and 20% of the highly purified tat preparation displayed high affinity toward TAR RNA.

Refolding of tat Protein. Since the presence of substantial quantities of "misfolded" tat protein prohibits a structural analysis of the tat–TAR complex, methods were sought for refolding tat protein. Fluorescence spectroscopy was employed to monitor the folded state of the protein. tat gave a weak but detectable tryptophan fluorescence signal (λ_{em} = 348 nm, 250 nM in 25 mM Tris-HCl, pH 8, and 10 mM DTT), consistent with the presence of a single Trp residue (Trp-11). Under denaturing and reducing conditions (6 M urea, 10 mM DTT), however, Trp fluoresence increased 3.5-fold, consistent with a change in the local Trp environment upon unfolding. These data indicated that tat, as isolated, adopts a folded structure in which Trp-11 is effectively quenched.

We then examined whether renaturation from urea yielded a protein preparation with increased specificity toward TAR RNA. tat protein (7 μM) was denatured in 8 M urea and 12 mM Tris-HCl (pH 8), containing 10 mM DTT and dialyzed for 16 h against 1000 volumes of buffer lacking urea at 4 °C without stirring. After this period, further dialysis against fresh buffer (2 h) was performed with stirring, and the resulting tat solution was stored at −20 °C. Under these conditions, tat protein was recovered in 30% yield as assessed by amino acid analysis. SDS–PAGE analysis, immediately after dialysis and without addition of reducing agents or heat treatment, demonstrated that >90% of the protein was monomeric.[3]

Refolded tat Protein Specifically Protects TAR RNA from Ribonucleolytic Cleavage. tat protein obtained by using this refolding procedure was much more effective than nonrefolded tat in protecting TAR RNA from ribonucleolytic cleavage.[4] Complexes of TAR RNA (~1 nM) and refolded tat protein (0–200 nM) were subjected to RNase A cleavage, and the products were separated on denaturing polyacrylamide gels (Figure 4). In the absence of tat, the expected cleavage fragments ($G^1 \rightarrow U^{31}$ and $G^1 \rightarrow U^{23}$) were apparent (Figure 4, lane 7). However, with 50 nM tat (lane 6), extensive (>80%) protection from ribonucleolytic cleavage at both sites was observed, and with 100 nM tat (lane 5), protection was virtually complete. Incubation in the presence of 10 μM Zn^{2+} or 0.1 mM EDTA had no effect on the extent of protection with 50 nM tat (not shown). This increase in specific activity (~30-fold) may reflect a larger proportion of tat being properly folded.

In order to determine whether refolded tat displayed enhanced specificity as well as affinity for TAR, protection assays were carried out with two truncated TAR RNAs. The concentration of refolded tat employed was 200 nM, at least 2-fold higher than that required to give complete protection of full-length TAR. Under these conditions, refolded tat had no effect on the extent of cleavage of TAR^{1-37} (Figure 4A, lane 8–10). Similar results were obtained with TAR^{1-32} (Figure 4B, lanes 1–3).

It has recently been shown that tat peptides encompassing the basic domain (YGRKKRRQRRRP) bind to TAR RNA

[2] Of course it was possible that tat itself, or trace impurities in the tat preparation, inhibited RNase A directly. Therefore, conventional spectrophotometric RNase assays employing UpA as substrate were performed in the presence and absence of tat protein under reaction conditions identical with those used for the protection assay (see Materials and Methods). Values for k_{cat}/K_m in the presence and absence of tat differed by less than 5% in this assay, indicating that the observed protection of TAR from cleavage was not due to a direct interaction of tat with RNase A. RNase T1 assays gave analogous results.

[3] Prolonged storage at 4 °C in dialysis buffer results in the formation of multimers, as assessed by SDS–PAGE, consistent with the results of others (Dingwall et al., 1989; Frankel et al., 1988). Addition of 25 mM DTT followed by incubation at room temperature (1 h) results in conversion of multimers to monomers, suggesting that multimerization reflects intermolecular disulfide bond formation.

[4] Several alternative refolding procedures were evaluated including (1) initial dialysis against 25 mM Tris (pH 8) containing 100 mM 2-mercaptoethanol and 1 mM EDTA followed by dialysis against 10 mM Tris (pH 8) containing 1 mM EDTA and 40 mM DTT, (2) dialysis against buffers containing 100 μM zinc chloride, and (3) dialysis at much higher concentrations of tat (>40 μM). While each of these treatments resulted in an increase in affinity for TAR (typically 2-fold based on the quantity of tat required for complete protection), tat preparations still displayed substantial nonspecific binding to truncated TAR RNAs.

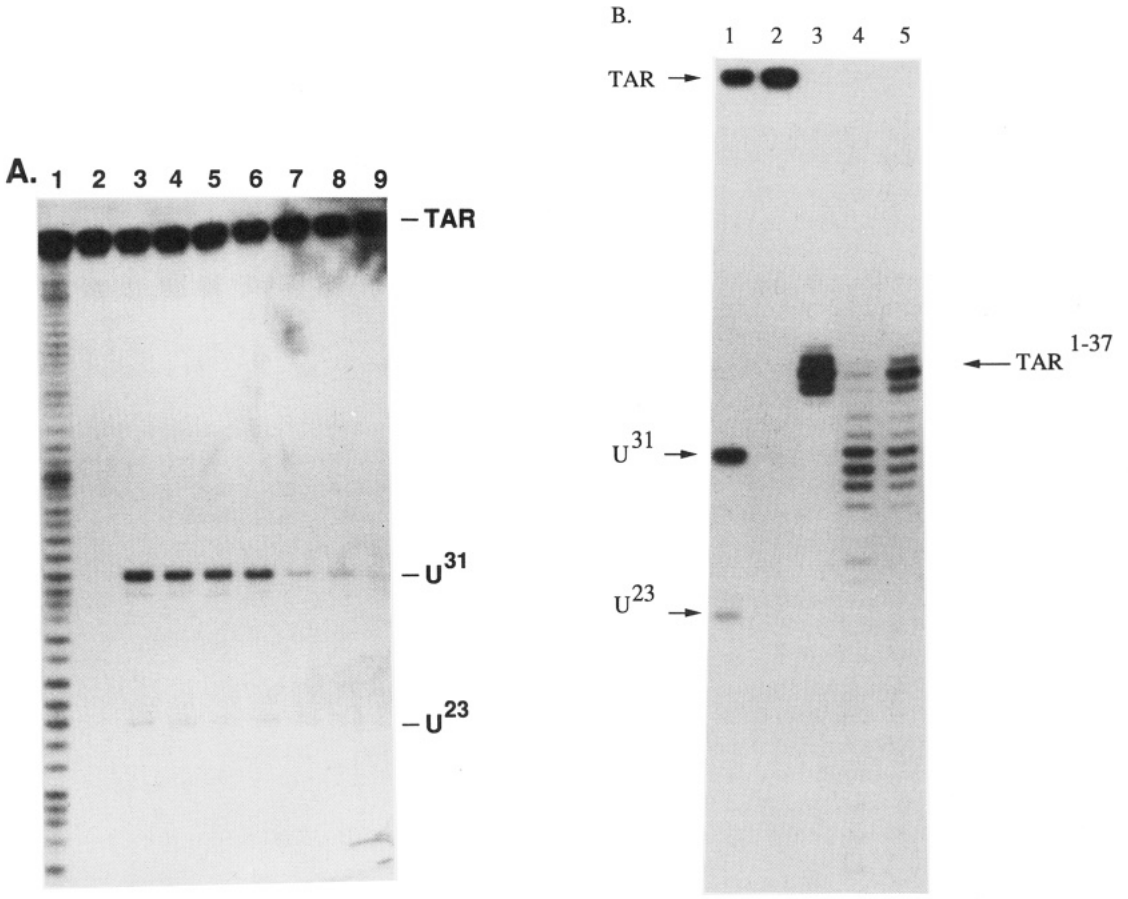

FIGURE 3: (A) Protection of TAR RNA from RNase A cleavage by tat protein. 5′-Labeled TAR RNA was incubated in the presence or absence of tat protein and then treated with RNase A as described under Materials and Methods. After phenol/chloroform extraction and ethanol precipitation, ~10000 cpm from each sample was subjected to electrophoresis as described in the legend to Figure 2B. (Lane 1) Products from alkaline cleavage of TAR RNA; (lane 2) RNA incubated in the presence of 3 μM tat without RNase A treatment; (lane 3) TAR RNA incubated with RNase A in the absence of tat protein; (lanes 4–9) TAR RNA incubated with 0.5, 0.75, 1.0, 1.5, 2.0, and 2.5 μM tat protein, respectively, prior to treatment with RNase A. The positions of RNase A cleavage fragments are indicated by arrows. (B) Effect of tat protein on RNase A cleavage of TAR^{1-37}. 5′-Labeled TAR^{1-37} was incubated in the presence or absence of tat followed by RNase A treatment and gel electrophoresis (see Materials and Methods). Since *Sst*I generates a 3′ overhang, TAR^{1-37} RNA generated from the transcription template is heterogeneous and contains up to 3 additional nucleotide residues at the 3′ end of the RNA, the major species being 36 nucleotides in length. Previous studies (Milligan et al., 1987) have shown that T7 RNA polymerase can add nontemplate nucleotides at the 3′ end of some transcripts. The actual TAR sequence extends through G^{34}. (Lane 3) TAR^{1-37} incubated with 2 μM tat without RNase A treatment; (lane 4) TAR^{1-37} treated with RNase A in the absence of tat; (lane 5) TAR^{1-37} incubated with 2 μM tat prior to RNase A treatment. In order to allow for direct comparison, reactions were repeated with 5′-labeled TAR RNA in the absence of tat (lane 1) and in the presence of 2 μM tat (lane 2).

with K_d values of ~6 nM (Calnan et al., 1991). Such peptides lack the cysteine-rich region, and thus problems with aggregation and nonspecific binding are avoided. Because RNase binding may perturb the tat–TAR equilibrium, we have made no attempt to determine a true tat–TAR binding constant using the RNase protection assay. However, the apparent K_d is below 25 nM, well within the range of binding constants for tat peptides. This would imply that the majority of the refolded tat protein is capable of binding TAR with high affinity and specificity.

Refolded tat Protects G^{32}–G^{34} from Ribonucleolytic Cleavage. As noted above, RNase T1 cleavage of TAR RNA occurs at GpN dinucleotides located in the six-base loop. As with RNase A, cleavage by RNase T1 is greatly reduced in the presence of 50 nM tat and is virtually undetectable with 200 nM tat (Figure 4A, lanes 11–13).[2] Upon long exposures of these gels, trace levels of RNase T1 cleavage at G^{60} can be discerned. Interestingly, cleavage at this position is significantly enhanced by tat (Figure 4A, lanes 12 and 13). Taken together, these data suggest that tat interacts with multiple residues contained in the loop region as well as those located in the three-base bulge.

Discussion

In vitro studies of tat–TAR complex formation have been hindered by the fact that only a fraction (<1%) of tat produced in *E. coli* has high affinity and specificity for TAR (Frankel et al., 1988; Dingwall et al., 1989; Roy et al., 1990). This is presumably due to formation of misfolded species which both aggregate and form disulfide or metal-linked multimers. We have found that a simple refolding procedure carried out under reducing conditions results in a tat preparation which displays ~30-fold higher apparent affinity for TAR RNA as well as diminished "non-specific" binding to truncated TAR RNAs. Thus, many of the problems associated with studying tat in vitro (Frankel et al., 1988; Dingwall et al., 1989; Roy et al., 1990; Weeks et al., 1990) are alleviated.

A variety of studies both in vitro and in vivo [Feng & Holland, 1988; Roy et al., 1990; Cordingley et al., 1990; Dingwall et al., 1989; Calnan et al., 1991; for a review, see Cullen (1990)] indicate that the tat binding site is confined to nucleotides 19–42 in TAR (Figure 1). Because the double-stranded regions of TAR appear to play largely a structural role, the search for sequence-specific interactions has focused on single-stranded regions in TAR (Roy et al., 1990; Cord-

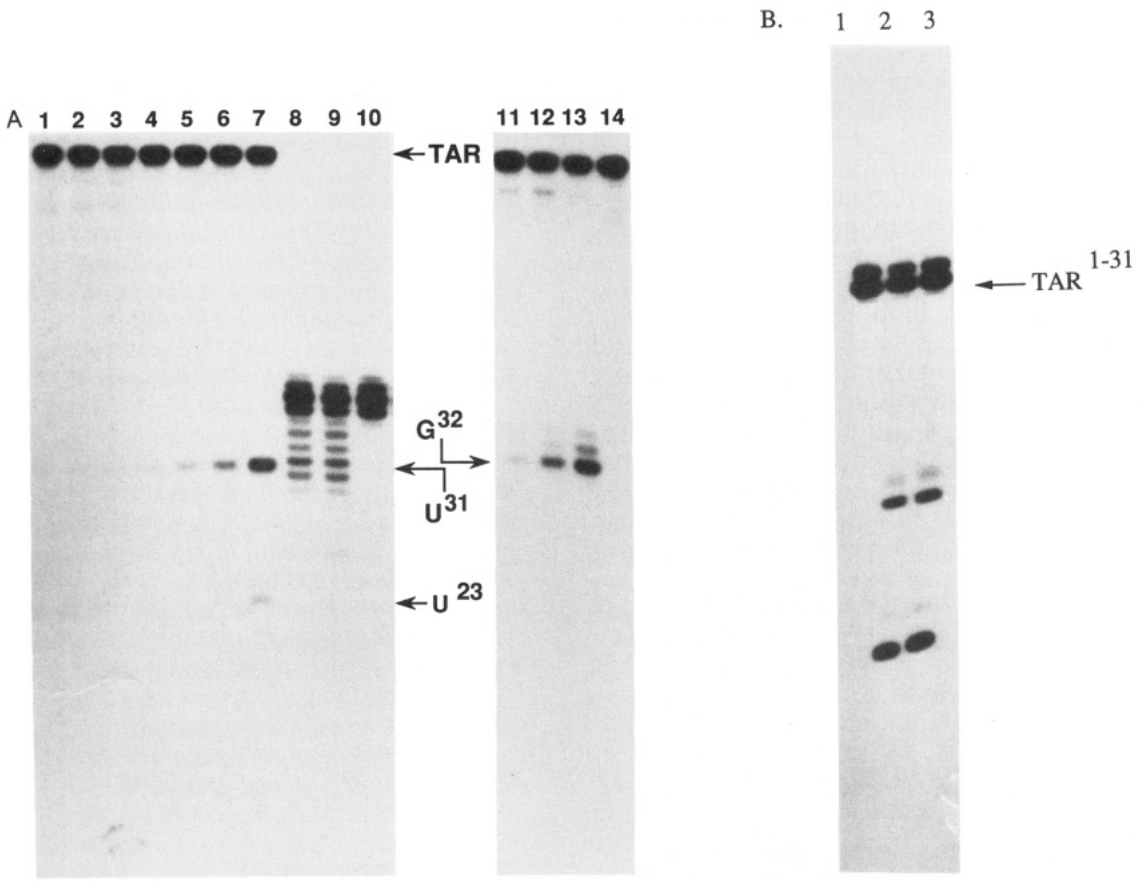

FIGURE 4: (A) Refolded tat protein specifically protects TAR RNA from ribonucleolytic cleavage by RNase A and RNase T1. tat protein was subjected to a refolding procedure as indicated under Results. Incubations were carried out in the presence or absence of refolded tat as described under Materials and Methods. (Lanes 1–6) TAR RNA incubated with 200, 175, 150, 125, 100, and 50 nM tat, respectively, prior to treatment with 1 pM RNase A; (lane 7) TAR RNA treated with 1 pM RNase A in the absence of tat; (lane 8) TAR^{1-37} incubated with 200 nM tat prior to treatment with 0.5 pM RNase A; (lane 9) TAR^{1-37} treated with 0.5 pM RNase A in the absence of tat; (lane 10) TAR^{1-37} incubated in the presence of 200 nM tat without RNase A treatment; (lanes 11 and 12) TAR RNA incubated with 200 and 50 nM, respectively, prior to treatment with RNase T1 (25 units/mL); (lane 13) TAR RNA treated with RNase T1 (25 units/mL) in the absence of tat; (lane 14) TAR RNA incubated with 200 nM tat without RNase treatment. (B) Effect of tat protein on RNase A cleavage of TAR^{1-31} RNA. Universally labeled TAR1–31 was incubated in the presence or absence of 200 nM refolded tat prior to treatment with RNase A and analysis by gel electrophoresis. (Lane 1) TAR^{1-31} RNA incubated with 200 nM tat without RNase treatment; (lane 2) TAR^{1-31} RNA treated with 0.5 pM RNase A in the absence of tat; (lane 3) TAR^{1-31} RNA treated with 0.5 pM RNase A in the presence of 200 nM tat.

ingley et al., 1990; Dingwall et al., 1989; Calnan et al., 1991). The predominant sequence-specific interaction identified by extensive mutagenic analysis of TAR is with the bulge nucleotide U^{23} (Roy et al., 1990; Cordingley et al., 1990); replacement by C increases the apparent dissociation constant ~20-fold (Calnan et al., 1991). In contrast, mutation of individual or multiple nucleotides in the loop has a more limited effect on tat binding in vitro. Mutation of G^{32}, G^{33}, or G^{34} to U was reported to cause no significant reduction in tat binding using gel retardation assays (Dingwall et al., 1989). Similar results were obtained when multiple mutations were made in the loop (Cordingly et al., 1990; Roy et al., 1990). However, because a single high concentration of tat (1–2 μM) was used in these studies, changes in affinity of 10-fold or less would have been difficult to observe. Another study in which titration experiments were carried out showed that peptides containing the basic domain of tat bind TAR containing a modified loop structure about 5-fold more weakly than the unmodified RNA (Calnan et al., 1991).

The approach employed in the present work, protection from ribonucleolytic cleavage by single-strand-specific RNases, complements the gel retardation techniques employed previously and can provide information not obtainable in other ways. This type of approach has been employed extensively for mapping protein binding sites on RNA [for examples, see Pelka and Schulman (1986) and Heaphy et al. (1990)]. RNases A and T1 are small enzymes (M_r ~ 13K). The crystal structure of RNase A bound to substrates and inhibitors (Richards & Wyckoff, 1973) revealed that the components of a dinucleotide (e.g., UpG) fit into the shallow active site but there is little evidence for interactions with adjacent nucleotide components. Thus, protection from cleavage at a particular site within a larger RNA molecule indicates that the bound protein is located near the protected dinucleotide. A major advantage of RNase protection over gel retardation assays with mutated RNAs is that potentially important sequence-independent interactions can be identified. In addition, the actual discrimination between bound and unbound RNA in the RNase protection assay takes place in solution and thus avoids intrusive effects inherent in gel retardation and filter binding techniques.

Limited treatment of TAR RNA with RNase A leads to cleavage at two sites: U^{23} in the bulge and U^{31} in the loop (Figure 2A). As expected on the basis of previous mutagenesis experiments (Roy et al., 1990; Cordingley et al., 1990; Dingwall et al., 1989; Calnan et al., 1991), refolded tat protein effectively protects TAR from cleavage at U^{23} (Figure 4). Presumably, base-specific interactions with this residue by tat shield the RNA from cleavage. More surprising was the finding that tat effectively protects U^{31} from RNase A cleavage (Figure 4). The extent of cleavage at U^{31} with varying tat concentration closely paralleled that observed at U^{23}, indicating

a similar protein–RNA binding constant at both sites. These results led us to examine whether other nucleotides in the loop are also protected by tat (Figure 4). In the absence of tat, RNase T1 cleaves TAR RNA at all three GpN bonds in the loop, the cleavage at $G^{32}pG$ being predominant. In the presence of refolded tat protein, however, all three sites are protected while cleavage at G^{60} is enhanced. The extent of cleavage at varying tat concentrations again paralleled that found with RNase A.

Binding energy in protein–RNA complexes is typically derived from a combination of two classes of interactions: (1) sequence-specific hydrogen bonding between protein side chains and nucleotide rings and (2) sequence-independent interactions which make use of phosphate or ribose groups on the RNA. Clearly, ^{23}U in TAR provides an important sequence-specific interaction for tat binding, but even in the absence of this interaction, tat peptides bind TAR with K_d values approaching 100 nM (Calnan et al., 1991). This implies that sequence-independent interactions contribute substantially to the overall binding energy of the complex.

Our data provide the first indication that tat physically interacts with the TAR loop. The fact that the extent of protection upon varying tat concentration is indistinguishable for each of four major cleavage sites in both the bulge and loop implies that tat interacts with both regions simultaneously. Thus, these two structural motifs may constitute a single binding site. In addition, the data suggest that tat may interact with multiple components within the loop since cleavage at all sites is blocked. In this regard, preliminary molecular modeling of TAR and the tat basic domain[5] indicates that, in the extended conformation, Arg and Lys side chains in the peptide may interact with loop and bulge residues simultaneously. Therefore, it is plausible that the interaction of an as yet unidentified residue(s) in the tat basic region (or perhaps residues outside this region) with the structurally constrained loop of TAR serves to provide binding energy as well as to align that tat basic domain so that sequence-specific interactions are made with the bulge.

ACKNOWLEDGMENTS

We thank Dr. Stephen Elledge and Dr. Theodore Wensel (Baylor College of Medicine) for valuable advice and Dr. Garry King (Rice University) for numerous helpful discussions and for molecular modeling.

REFERENCES

Aldovini, A., DeBouck, C., Feinberg, M. B., Rosenberg, M., Arya, S. K., & Wong Staal, F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6672–6676.

Berkhout, B., & Jeang, K.-T. (1989) *J. Virol. 63*, 5501–5504.

Berkhout, B., Gatignol, A., Rabson, A. B., & Jeang, K.-T. (1990) *Cell 62*, 757–767.

Calnan, B. J., Biancalana, S., Hudson, D., & Frankel, A. D. (1991) *Genes Dev. 5*, 201–210.

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., & Colonno, R. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8985–8989.

Cullen, B. R. (1986) *Cell 46*, 973–982.

Cullen, B. R. (1990) *Cell 63*, 655–657.

Cullen, B. R., & Greene, W. C. (1989) *Cell 58*, 423–426.

Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.

Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A., & Valerio, R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6925–6929.

Feng, S., & Holland, E. C. (1988) *Nature 334*, 165–167.

Frankel, A. D., & Pabo, C. O. (1988) *Cell 55*, 1189–1194.

Frankel, A. D., Bredt, D. S., & Pabo, C. O. (1988) *Science 240*, 70–73.

Harper, J. W., Auld, D. S., Riordan, J. F., & Vallee, B. L. (1988) *Biochemistry 27*, 219–226.

Hauber, J., & Cullen, B. R. (1988) *J. Virol. 62*, 673–679.

Heaphy, S., Dingwall, C., Ernberg, I., Gait, M. J., Greene, S. M., Karn, J., Lowe, A. D., Singh, M., & Skinner, M. A. (1990) *Cell 60*, 685–693.

King, G. C., Martin, C. T., Pham, T. T., & Coleman, J. E. (1986) *Biochemistry 25*, 36–40.

Laspia, M. F., Rice, A. P., & Mathews, M. B. (1989) *Cell 59*, 283–292.

Marciniak, R. A., Calnan, B. J., Frankel, A. D., & Sharp, P. A. (1990a) *Cell 63*, 791–802.

Marciniak, R. A., Garcia-Blanco, M. A., & Sharp, P. A. (1990b) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3624–3628.

Martin, C. T., & Coleman, J. E. (1989) *Biochemistry 28*, 2760–2762.

Muesing, M. A., Smith, D. H., & Capon, D. J. (1987) *Cell 48*, 691–701.

Owens, J. R., & Haley, B. E. (1984) *J. Biol. Chem. 259*, 14843–14848.

Pelka, H., & Schulman, L. H. (1986) *Biochemistry 25*, 4450–4456.

Rice, A. P., & Mathews, M. B. (1988) *Nature 332*, 551–553.

Richards, F. M., & Wyckoff, H. W. (1973) *Atlas of Molecular Structures in Biology* (Phillips, D. C., & Richards, F. M., Eds.) Vol. 1, Oxford University Press, London.

Rosen, C. A., & Pavlakis, G. N. (1990) *AIDS 4*, 499–509.

Roy, S., Delling, U., Chen, C.-H., Rosen, C., & Sonenberg, N. (1990) *Genes Dev. 4*, 1365–1373.

Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033–1037.

Selby, M. J., & Peterlin, B. M. (1990) *Cell 62*, 769–776.

Selby, M. J., Bain, E. S., Luciw, P. A., & Peterlin, B. M. (1989) *Genes Dev. 3*, 547–558.

Shapiro, R., Riordan, J. F., & Vallee, B. L. (1986) *Biochemistry 25*, 3527–3532.

Sharp, P. A., & Marciniak, R. A. (1989) *Cell 59*, 229–230.

Southgate, C., Zapp, M. L., & Green, M. R. (1990) *Nature 345*, 640–642.

Vogel, J., Hinrichs, S. H., Reynolds, R. K., Luciw, P. A., & Jay, G. (1988) *Nature 335*, 606–611.

Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., & Crothers, D. M. (1990) *Science 249*, 1281–1285.

Wright, C. M., Felber, B. K., Paskalis, H., & Pavlakis, G. N. (1986) *Science 234*, 988–992.

[5] Dr. Garry King (Rice University), personal communication.