JOURNAL OF VIROLOGY, Oct. 1991, p. 5196–5202
0022-538X/91/105196-07$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 65, No. 10

Critical Chemical Features in *trans*-Acting-Responsive RNA Are Required for Interaction with Human Immunodeficiency Virus Type 1 Tat Protein

MARTIN SUMNER-SMITH,[1] SOPHIE ROY,[2] RICHARD BARNETT,[1] LORNE S. REID,[1] RAYA KUPERMAN,[1] ULRIKE DELLING,[2] AND NAHUM SONENBERG[2,3]*

Allelix Biopharmaceuticals Inc., 6850 Goreway Drive, Mississauga, Ontario L4V 1P1,[1] and Department of Biochemistry[2] and McGill Cancer Center,[3] McGill University, Montreal, Quebec H3G 1Y6, Canada

Received 24 April 1991/Accepted 8 July 1991

The human immunodeficiency virus type 1 Tat protein binds to an RNA stem-loop structure called TAR which is present at the 5′ end of all human immunodeficiency virus type 1 transcripts. This binding is centered on a bulge within the stem of TAR and is an essential step in the *trans*-activation process which results in a dramatic increase in viral gene expression. By analysis of a series of TAR derivatives produced by transcription or direct chemical synthesis, we determined the structural and chemical requirements for Tat binding. Tat binds well to structures which have a bulge of two to at least five unpaired bases bounded on both sides by a double-stranded RNA stem. This apparent flexibility in bulge size is in contrast to an absolute requirement for an unpaired uridine (U) in the 5′-most position of the bulge (+23). Substitution of the U with either natural bases or chemical analogs demonstrated that the imido group at the N-3 position and, possibly, the carbonyl group at the C-4 position of U are critical for Tat binding. Cytosine (C), which differs from U at only these positions, is not an acceptable substitute. Furthermore, methylation at N-3 abolishes binding. While methylation of U at the C-5 position has little effect on binding, fluorination reduces it, possibly because of its effects on relative tautomer stability at the N-3 and C-4 positions. Thus, we have identified key moieties in the U residue that are of importance for the binding of Tat to TAR RNA. We hypothesize that the invariant U is involved in hydrogen bond interactions with either another part of TAR or the TAR-binding domain in Tat.

trans activation of human immunodeficiency virus type 1 (HIV-1) gene expression (8, 35) requires the viral Tat protein (1, 9, 16, 40, 45). Tat function is dependent on another viral component—namely, a *trans*-acting-responsive (TAR) RNA stem-loop structure (30, 36)—found at the 5′ end of all viral transcripts (see Fig. 2). The presence of a stable TAR secondary structure in vitro has been demonstrated by nuclease mapping (30), and the importance of its structural integrity in vivo has been shown by mutagenesis studies (3, 15, 17, 21, 25, 39, 43).

Recently, it has been shown that Tat is an RNA-binding protein with specificity for TAR RNA (7, 11, 12, 31, 38, 48). The presence of a bulge structure of two or three bases in the TAR stem is required, and in particular, the 5′-most invariant U of this bulge is especially important (38). Modifications to the invariant U in the bulge and to three base pairs below and two above the bulge affect Tat binding (48). While the sequence of the TAR loop is important for *trans* activation, it is not required for binding of Tat (7, 12, 38). However, host factors that bind specifically to the loop have been identified (18, 19, 27, 28).

We wished to determine the structural and chemical basis for the specific binding of Tat to TAR. In general, nucleic acid-binding proteins have a nonspecific affinity for nucleic acids and an enhanced affinity for a specific target. This enhancement can be achieved by interaction between the protein and certain bases (sequence specificity) but can also depend on precise structural features (structural specificity). Indeed, while the primary sequence of most of the stem of TAR is apparently not important for activity, disruption of base pairing, and therefore structure, drastically reduces *trans* activation (3, 15, 17, 21, 25, 39, 43) and Tat binding (12, 38). However, the relative contribution of particular structural features to RNA folding per se or to specific binding should be distinguished. To define this, we analyzed the binding of Tat to a number of TAR structural variants. Furthermore, we analyzed the binding of Tat to a number of TAR chemical variants in which the invariant U at position +23, previously shown to be required for nucleoprotein complex formation (11, 38), was replaced by nonnatural base analogs. These studies establish the structural and chemical features of TAR RNA that are required for specific interaction with Tat.

MATERIALS AND METHODS

RNA secondary-structure prediction. The program LRNA (linear RNA folding) was obtained from M. Zuker (National Research Council, Ottawa, Ontario, Canada) to run in a VAX host environment (23, 24, 49). The calculations of optimal and suboptimal RNA folding patterns were based upon the standard Zuker folding rules as calculated for 37°C. No alterations to the folding energies were used. The NEWTEMP option was used to calculate parameters at temperatures other than the default 37°C.

Chemical synthesis of oligonucleotides and verification. Oligodeoxyribonucleotides (DNA) were prepared by the β-cyanoethyl–phosphoramidite method (6) on an Applied Biosystems 391EP synthesizer (0.15-μmol scale). Cleavage and deprotection were effected by standard ammonia treatment (44). Oligoribonucleotides (RNA) were prepared by the method of Ogilvie et al. (32) by using 5′-dimethoxytrityl-2′-*t*-butyldimethylsilyl (TBDMS) ribonucleoside-3′-cyanoethyl-

* Corresponding author.

phosphoramidites (Peninsula Laboratories or ChemGenes Corp.). Cyanoethyl-phosphoramidites of the modified 2'-deoxynucleosides were purchased from commercial sources (Peninsula, ChemGenes, or Applied Biosystems Canada) and used as recommended by the suppliers. Syntheses were carried out on an Applied Biosystems 380B synthesizer using a modified 0.2-μmol cycle. Cleavage from the support, base and phosphate deprotection, and removal of the 2'-TBDMS groups were done by established procedures (42). The crude oligonucleotides in tetra-*n*-butylammonium fluoride solution were desalted on a C_{18} Sep-Pak cartridge prior to purification by standard electrophoresis methods (2) using 15 to 20% polyacrylamide–7 M urea gels. The eluted DNA or RNA was finally desalted on a C_{18} Sep-Pak and quantitated by optical density at 260 nm. Each oligonucleotide was checked for homogeneity and sized by ^{32}P labelling-analytical polyacrylamide gel electrophoresis-autoradiography versus the crude starting material and oligonucleotide markers. Some RNA oligonucleotides containing only ribonucleosides, e.g., ST20R1 and ST20R3, were sequence verified by the procedure of Donis-Keller (13) with a commercially available kit (BRL). Oligoribonucleotides containing modified deoxyribonucleosides, e.g., 5-fluorodeoxyuridine and *N*-methyldeoxythymidine, were further characterized by base composition analysis (14).

In vitro transcriptions and RNA gel purification. Full-length wild-type and mutant (B-UAU, B-UCU, and AGA-Ins) TAR RNAs (nucleotides +1 to +80) were synthesized by SP6 polymerase from SP64TARCAT constructs (33) as previously described (39). RNAs from oligodeoxynucleotide templates were prepared by transcription using T7 polymerase as described by Meerovitch et al. (29), except that the GTP concentration was raised to 0.5 mM. The transcripts were gel purified and extracted as previously described (38). Chemically synthesized oligoribonucleotides were 5' end labelled with [γ-^{32}P]ATP by T4 polynucleotide kinase by using standard methods (41). The RNAs were then passed through a G-50 spin column to remove unincorporated radiolabel. The specific activity of internally labelled RNAs produced by in vitro transcription and of 5'-end-labelled RNAs produced by direct chemical synthesis was 2.94×10^6 cpm/μg.

RNA mobility shift assay. Tat protein was prepared and purified as previously described (20) and kindly supplied by Craig Rosen (Roche Institute for Molecular Biology, Nutley, N.J.). The mobility shift assay was performed as previously described (38).

RESULTS

Primary sequence requirements in the trinucleotide bulge for Tat binding are limited to the U at position +23. In most HIV-1 isolates, TAR consists of 59 bases arranged in a stem-loop structure with a two- or three-nucleotide bulge in the stem, positioned 5 bp below the loop (see Fig. 2). TAR is likely to be a fairly lengthy structure because of the relatively rigid nature of the double-stranded RNA helices which make up much of it. Since the U at +23 is absolutely required for Tat binding (38), Tat probably interacts with that part of the bulge. We wished to determine whether there are additional sequence requirements in the bulge besides the U at +23. Natural isolates have either a U or a C at the second position of the bulge (+24). We tested TAR RNAs (nucleotides +1 to +80) containing an A or a G at +24 in the RNA mobility shift assay. These mutants bound to Tat as efficiently as wild-type TAR RNA (Fig. 1, compare lane 2

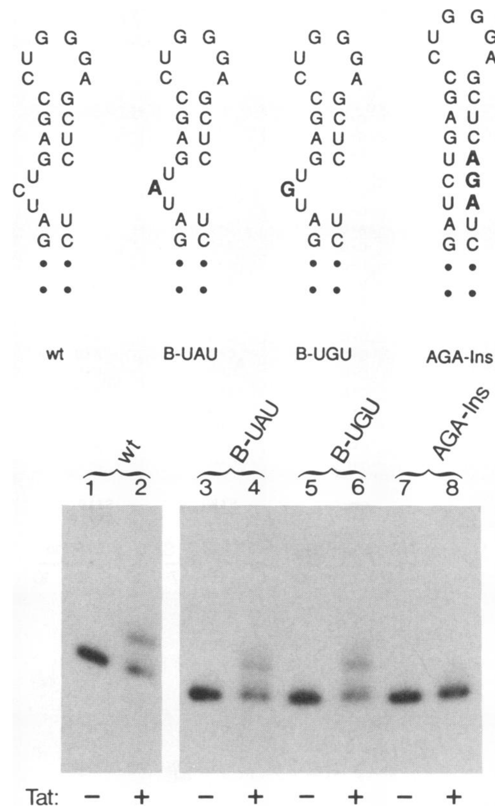

FIG. 1. Effects of nucleotide substitutions and insertions in TAR RNA on Tat binding in vitro. Mutations were introduced by site-directed mutagenesis and subcloned in pSP64TARCAT as described in Materials and Methods. The predicted effects of these mutations (indicated in boldface type) on the TAR stem-loop structure are shown. ^{32}P-labelled wild-type (wt) and mutant RNAs transcribed by SP6 polymerase were incubated in the absence (−) or presence (+) of 100 ng of purified Tat, and nucleoprotein complexes were analyzed on a nondenaturing polyacrylamide gel as described in Materials and Methods.

with lanes 4 and 6). The corresponding mutants in expression vector pU3RIII (which contains the chloramphenicol acetyltransferase gene under control of the HIV-1 long terminal repeat; 45) were responsive to Tat *trans* activation in vivo (data not shown). Thus, there does not appear to be any sequence specificity at the second position. Likewise, the third position seems to be unimportant both in vitro and in vivo, since it may be deleted (38).

The bulge sequence must not be paired for Tat binding. Deletion of the trinucleotide bulge or disruption of the stem structure in its vicinity was previously shown to abolish Tat binding (11, 12, 38). These results suggested that the conformation of the bulge structure per se is required for recognition by Tat. To provide direct evidence for this, we inserted a sequence complementary to the bulge on the opposite strand (AGA-Ins; Fig. 1). This mutant RNA failed to bind to Tat (lane 8), indicating that Tat does not recognize the bulge sequence within the TAR stem context unless it is unpaired. Accordingly, the corresponding mutation in expression vector pU3RIII (45) abolished Tat *trans* activation in vivo (data not shown).

Predicted RNA structures. Because small basic peptides derived from Tat have the same binding specificity as the whole molecule (4, 7, 10, 27, 38, 48), we reasoned that only

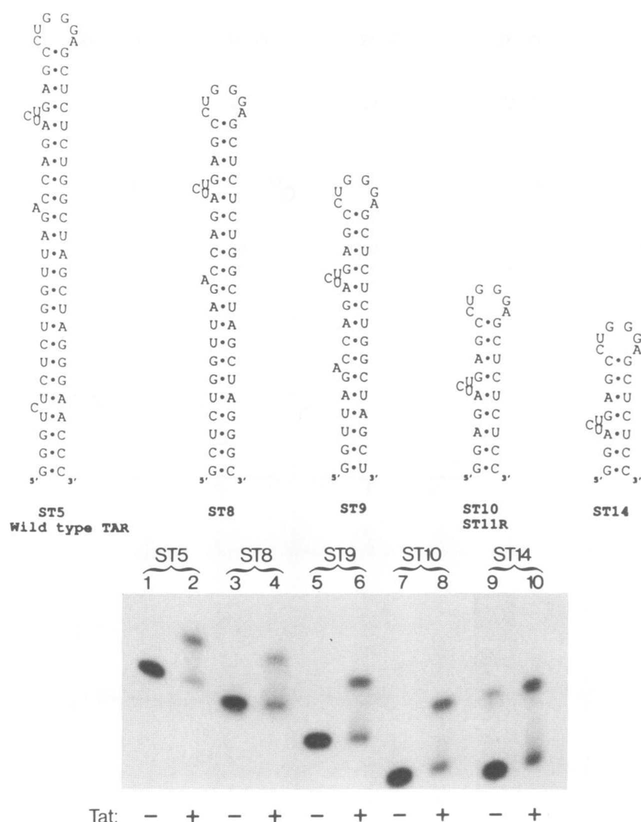

FIG. 2. Effects of truncations of the TAR stem on Tat binding in vitro. The predicted stem-loop structures of RNAs transcribed by T7 polymerase from oligodeoxynucleotide templates (ST5, ST8, ST9, ST10, and ST14) and an RNA derived from direct chemical synthesis (ST11R) are shown. ^{32}P-labelled RNAs were incubated in the absence (−) or presence (+) of 100 ng of purified Tat, and nucleoprotein complexes were analyzed on a nondenaturing polyacrylamide gel as described in Materials and Methods.

a relatively restricted region around the bulge is contacted by Tat. We made a series of derivatives which deleted defined TAR elements to map this region of contact. To ensure that the desired structures of these shorter RNAs would form under experimental conditions, we examined the predicted stability of the structures. The secondary structure of RNA is commonly predicted to the 70% accuracy level from empirically derived folding rules (23). However, those structures which are also possible but have slightly less favorable energies (i.e., the suboptimal folded structures) are generally not considered. Examination of these structures is necessary to ensure that the intended structure will predominate in solution. Ideally, the prediction should indicate a considerable difference in free energies so that there is confidence that the desired structure forms, even if there is some error in estimation of the free energies.

The wild-type canonical stem-loop structure (Fig. 2) is favored over the next suboptimal structure (which is identical except that there is an additional interaction between the single C and the third G in the loop) by an average of $\Delta G = 1.4 \pm 0.8$ kcal (1 cal = 4.184 J)/mol (Table 1). At this level (1.4 kcal/mol), the optimal structure would compose 99.8% of the population. Other alternate folded structures are even less favorable. For each derivative studied, the calculated free energies of the desired and next-most-preferred structures were determined (Table 1). As predicted, except where noted in the text, only one RNA structure was seen on native gels for each derivative in the absence of Tat.

TABLE 1. Predicted RNA structures

RNA	No. of residues	Free energy (kcal/mol)	
		Canonical	Alternate
ST5	59	−27.7	−26.3
ST8	47	−19.2	−17.8
ST9	39	−13.8	−12.4
ST10-ST11R	27	−9.0	−7.6
ST14	23	−5.0	−3.6
ST19	26	−16.1	−12.4
ST20R1	14	NA[a]	−2.5
ST20R3	16	NA	−1.6
ST20R1, ST20R3	30	−21.4	−14.2
ST21	29	−16.5	NA
ST22	15	−1.7	−0.1
ST23	13	−7.6	−0.5
ST17	25	−10.3	−8.9
ST15	26	−9.4	−8.0
ST31	27	−9.0	−7.6
ST16	28	−8.3	−6.9
ST32	29	−7.7	−6.3

[a] NA, not applicable.

Much of the TAR stem is dispensable for Tat binding. Full-length TAR (ST5 [59-mer]) and serial deletion derivatives (ST8 [47-mer], ST9 [39-mer], and ST10 [27-mer]) were produced by transcription from DNA templates. Each of these derivatives retained full binding potential (Fig. 2, lanes 2, 4, 6, and 8). The smallest of these derivatives (ST10) has only five base pairs below the bulge. The same derivative was also synthesized chemically (ST11R) and gave equivalent results (see Fig. 5, lane 2). Substitution of the invariant U in the 27-mer structure with either A or C abolished binding (data not shown), as it did with full-length TAR (38), which confirms that even in this smaller RNA molecule, the invariant U plays a critical role in binding.

Deletion to within 3 bp of the bulge (ST14 [23-mer]) resulted in the appearance of a minor, more slowly migrating RNA species in the absence of Tat (Fig. 2, lane 9). Since the stability of the alternate folded structure is very weak (Table 1), annealing between two oligonucleotides is more likely to account for this minor species. Nevertheless, there is evidence for the formation of a complex between the folded structure and Tat which comigrates with the minor RNA species (Fig. 2, lane 10).

The loop is not required for Tat binding. Although the sequence of the loop is known not to be important for binding (7, 12, 38), the loop structure per se may be an essential feature which is specifically recognized by Tat. To examine this possibility, we designed two oligonucleotides (ST20R1 and ST20R3) which were predicted to pair to form a stem structure lacking a loop but containing a UU bulge (Fig. 3). When both oligonucleotides were treated under conditions which favor annealing, a stable complex to which Tat could bind was formed (Fig. 3, lane 4). Similar results were obtained with an oligonucleotide pair which annealed to produce a bulge with the sequence UCU (data not shown). Therefore, either the loop does not present an essential structural feature, or if it does, an additional stem sequence of not more than 3 bp can fully substitute.

A stem is required on both sides of the bulge for Tat binding. We wished to determine whether Tat binding re-

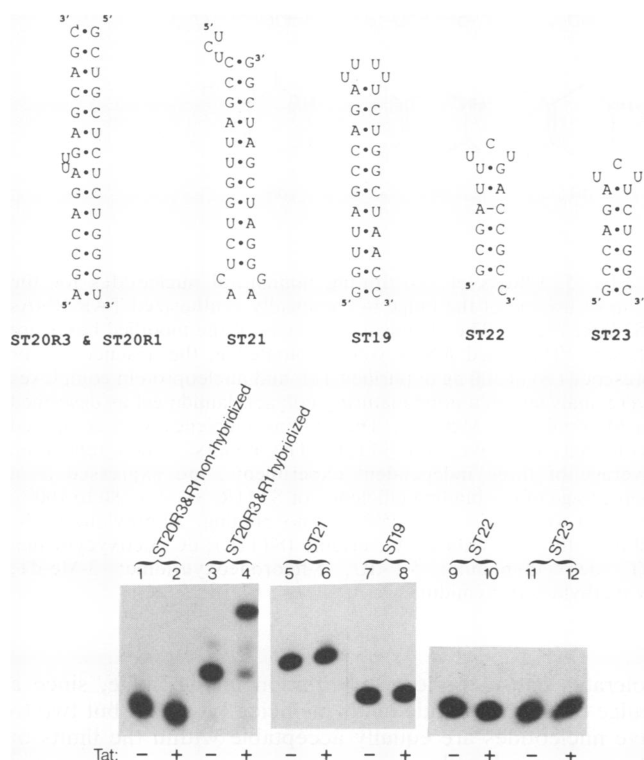

FIG. 3. Binding of non-TAR RNAs to Tat. The predicted stem-loop structures of RNAs transcribed by T7 polymerase from oligodeoxynucleotide templates (ST21, ST19, ST22, and ST23) or of RNAs derived from direct chemical synthesis (ST20R3 and ST20R1) are shown. ^{32}P-labelled RNAs were incubated in the absence (−) or presence (+) of 100 ng of purified Tat, and nucleoprotein complexes were analyzed on a nondenaturing polyacrylamide gel as described in Materials and Methods. Hybridization of ST20R3 and ST20R1 was performed by heating the RNAs at 90°C for 5 min and slow cooling in the presence of 20 mM Tris-HCl (pH 7.6)–3 mM magnesium acetate–400 mM NaCl–1 mM dithiothreitol prior to incubation with Tat.

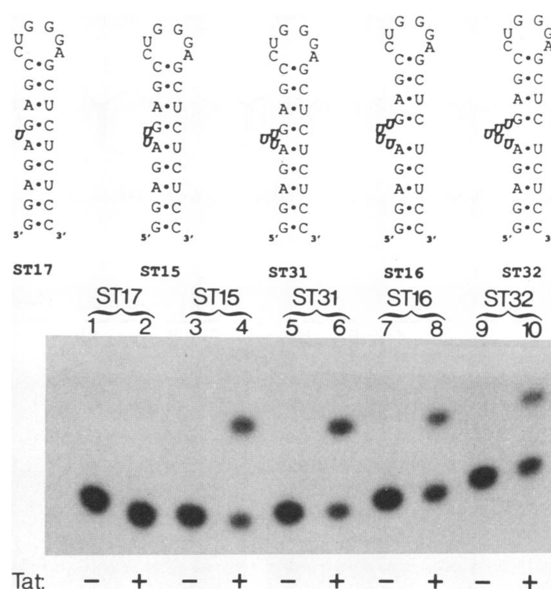

FIG. 4. Effect of bulge size in TAR on Tat binding in vitro. The predicted stem-loop structures of RNAs transcribed by T7 polymerase from oligodeoxynucleotide templates are shown. Nucleotides in the bulge are indicated in boldface type. ^{32}P-labelled RNAs were incubated in the absence (−) or presence (+) of 100 ng of purified Tat, and nucleoprotein complexes were analyzed on a nondenaturing polyacrylamide gel as described in Materials and Methods.

quires a double-stranded RNA stem on one or both sides of the bulge. ST21 has an unpaired UCU sequence adjacent to a stem of 11 bp, but it is not a suitable substrate for Tat binding (Fig. 3, lane 6). This result implies that a stem structure is required on both sides of the bulge. However, it is possible that the strict conformation to which the UCU bulge is constrained when it is between two stem structures is required for strong Tat binding. Therefore, three oligonucleotides which have the bulge sequence constrained as a loop were devised. Recently, Puglisi et al. (34) demonstrated by nuclear magnetic resonance analysis that an artificial RNA oligonucleotide containing a 3-base loop with the sequence UCU (i.e., the TAR bulge sequence) is able to form a stable structure although the stem has only six base pairs, of which one is an unusual A · C pair and another is a G · U pair. We tested this structure (ST22), as well as a derivative that has a sequence related to that in the stem below the bulge of TAR and is predicted to be much more stable (ST23). We also tested an RNA with a loop of four U's constrained by a stem of 11 bp, since four U's, in the context of a bulge, is a suitable substrate for Tat binding (see below). Tat did not bind to any of these non-TAR RNAs (Fig. 3, lanes 8, 10, and 12). These results suggest that Tat binding requires the presence of a stem structure on both sides of the trinucleotide bulge.

A functional bulge can vary significantly in size. Tat was previously shown to bind to TAR RNAs with bulges of two or three nucleotides where the 5′-most position is a uridine (38). We tested the limits to the size of the bulge with variants which had one to five U's (Fig. 4). As expected, UU (ST15) and UUU (ST31) were efficient substrates (lanes 4 and 6) but, surprisingly, Tat also bound to UUUU (ST16) and UUUUU (ST32) (lanes 8 and 10). These results indicate that the bulge can vary significantly in size without affecting Tat binding. However, very weak binding was seen with a single U (ST17, lane 2; other experiments showed somewhat better binding). No binding was seen under the same conditions in the absence of a bulge (38; Fig. 1, AGA-Ins). Since Tat bound to another four-base bulge (UAAC) with only the 5′-most base as a U (data not shown), it appears that the other U's in the large bulge structures do not satisfy the 5′-U requirement. Moreover, a variant with a CU bulge was a poor substrate for Tat binding (data not shown), also emphasizing the importance for a U in the 5′-most position of the bulge.

Modification of the invariant U. The invariant U in the 5′ position of the bulge cannot be replaced with either A, C, or G (results above; 9a, 38). The ability to make effective substrates by direct chemical synthesis allowed us to incorporate nonnatural nucleotides to study elements critical for this specific interaction (Fig. 5). RNAs containing nucleotide analogs were chemically synthesized, and their sequences were confirmed by sequencing and/or composition analysis, as described in Materials and Methods.

Deoxyuridine (dU) was found to be an acceptable substitute (compare lanes 8 and 2), suggesting that the 2′ hydroxyl of the ribose does not contribute significantly to the interaction. However, deoxycytosine, which differs from dU only

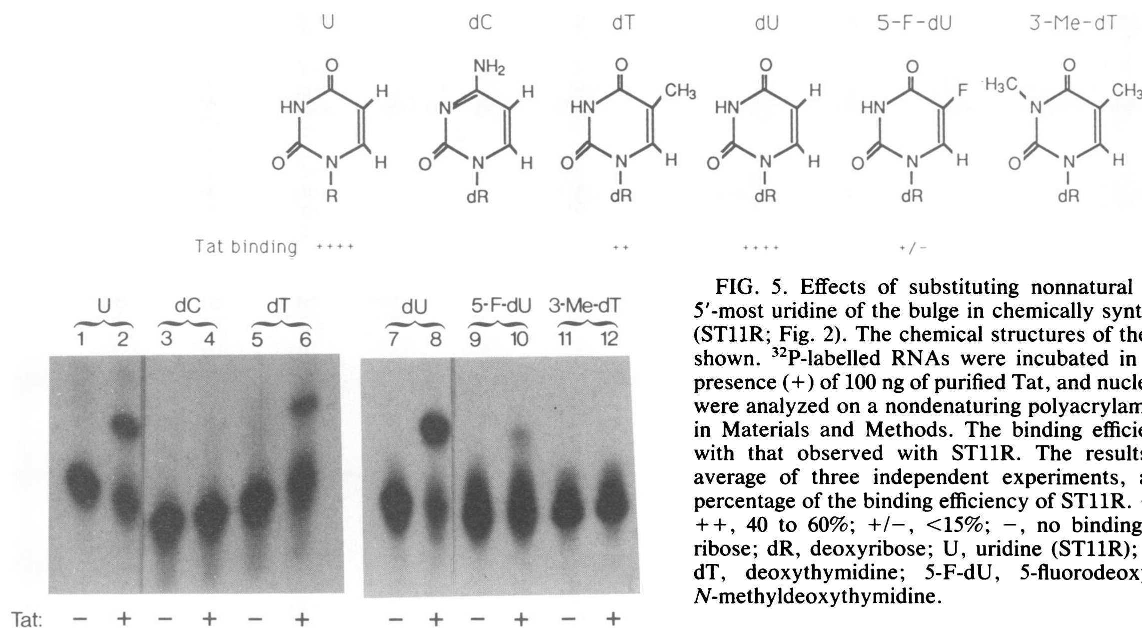

FIG. 5. Effects of substituting nonnatural nucleotides for the 5′-most uridine of the bulge in chemically synthesized TAR RNAs (ST11R; Fig. 2). The chemical structures of the modified bases are shown. ^{32}P-labelled RNAs were incubated in the absence (−) or presence (+) of 100 ng of purified Tat, and nucleoprotein complexes were analyzed on a nondenaturing polyacrylamide gel as described in Materials and Methods. The binding efficiency was compared with that observed with ST11R. The results, which reflect an average of three independent experiments, are expressed as a percentage of the binding efficiency of ST11R. ++++, 80 to 100%; ++, 40 to 60%; +/−, <15%; −, no binding. Abbreviations: R, ribose; dR, deoxyribose; U, uridine (ST11R); dC, deoxycytosine; dT, deoxythymidine; 5-F-dU, 5-fluorodeoxyuridine; 3-Me-dT, *N*-methyldeoxythymidine.

at the N-3 and C-4 positions, is not an acceptable substitute (lane 4), consistent with previous results obtained by mutational analysis of the TAR region (38). We also tested deoxythymidine, which differs from dU only by methylation of C-5. Deoxythymidine substituted for dU, albeit with reduced efficiency (about twofold; compare lanes 6 and 8). This indicates that the C-5 hydrogen of uridine is not directly involved in the critical interaction required for Tat binding. The twofold-reduced binding may be due to steric hindrance from the methyl group. Substitution of the hydrogen at C-5 with fluorine (5-fluorodeoxyuridine), which is more similar in size to the hydrogen atom, resulted in a dramatic decrease in Tat binding (~10-fold; compare lanes 10 and 8). The strongly nucleophilic nature of fluorine is predicted to influence the tautomerization of the base at N-3 and C-4 and increase the prevalence of the enol form (22). Fluorine substitution has also been shown to affect the configuration of the base relative to the sugar residue (22). Significantly, methylation of N-3 in deoxythymidine completely abolished Tat binding (lane 12). These results indicate that the imido group at the N-3 position and possibly the carbonyl group at the C-4 position of U are important for efficient binding. The hydrogen and oxygen at these positions are capable of participating in hydrogen bonding, the former as a donor and the latter as an acceptor.

DISCUSSION

In this report, we provide further understanding of the chemical basis of the interaction between the HIV-1 Tat protein and TAR RNA. We show that efficient nucleoprotein complex formation requires flanking of the trinucleotide bulge of TAR by a stem structure. Moreover, we demonstrate that the strict requirement for a U at the 5′-most position in the bulge involves the imido group at the N-3 position and possibly the carbonyl group at the C-4 position of the base and may entail the formation of hydrogen bonds with either another part of TAR or the TAR-binding domain of Tat. Finally, we show that Tat binding to TAR RNA tolerates considerable size variation in the bulge, since a bulge of one nucleotide binds inefficiently to Tat, but two to five nucleotides are equally acceptable within the limits of the assay employed.

The TAR loop is completely dispensable for Tat binding, since two complementary oligonucleotides which form a UU or UCU bulge with seven base pairs on either side after annealing provide an effective substrate. It is possible that Tat could bind to such a structure containing even shorter stems. However, the absence of a loop that connects the two strands is predicted to diminish the stability of the smaller duplex. In the context of a stem-and-loop structure, our deletion experiments show that not more than five, and perhaps as few as three, base pairs are required below the bulge. These findings are consistent with a recent report showing that modification of bases in the three base pairs below and two base pairs above the bulge affects binding (48).

The variability in the size of the bulge which allows Tat binding was surprising at first. However, Tang and Draper (47) demonstrated that UUU and UUUUU bulges may induce similar effects on helix conformation, consistent with our finding that both bulges have equivalent Tat-binding potentials. In contrast, a bulge of a single U probably stacks or intercalates within the RNA helix (47). We have found such a structure to be a poor substrate for Tat binding, which suggests that the U must be "bulged out" from the helix to generate a functional TAR.

Peptides containing or limited to nine amino acids derived from Tat (Arg, Lys, Lys, Arg, Arg, Gln, Arg, Arg, and Arg) retain the binding specificity of full-length Tat (4, 7, 10, 27, 48). Given the highly basic nature of this sequence, much of the interaction between Tat and TAR may be electrostatic. Indeed, we have found that binding is adversely affected by increasing salt concentrations (37). The interaction of the eight basic amino acids with the phosphates in the backbone of the RNA stem is sufficient to explain the general affinity of Tat for nucleic acids but not its greater specificity for TAR. Our results suggest that the unpaired U at position +23 provides a basis for the specificity. Recognition of tRNA by tRNA synthetases also depends on a specific U (26, 46). The

latter interaction entails formation of a transient covalent bond (Michael adduct) between C-6 of the base and cysteine and is enhanced by the electron-withdrawing properties of bromine at C-5, which renders C-6 more susceptible to nucleophilic attack by cysteine (26). It is unlikely that a similar mechanism is responsible for the interaction between Tat and TAR RNA, since the basic region of Tat does not contain a nucleophile able to attack C-6 in U, and the presence of fluorine at C-5 reduces rather than enhances nucleoprotein complex formation.

Results presented in this study suggest that the N-3 and, possibly, the C-4 positions in the unpaired U are required for the specificity of the interaction. The imido group at N-3 and the carbonyl group at C-4 are exposed when the base is not paired and thus may act as a hydrogen bond donor and acceptor, respectively. The failure of *N*-methylthymidine to substitute for uridine in Tat binding implicates the N-3 imido group in a critical interaction. Although we have not ruled out the possibility that the base is involved in hydrogen bond interactions with one of the amino acids in the TAR-binding domain of Tat, we favor the hypothesis that it is hydrogen bonding with the TAR RNA backbone in the vicinity of the bulge. This tertiary interaction may alter the degree of kinking introduced by the bulged uridine. Recently, Calnan et al. (5) have postulated a complex hydrogen bond arrangement between a single arginine side chain in the basic domain of Tat and two phosphates on either side of the invariant U. The hypothesized tertiary interaction between the bulged U and the TAR RNA backbone may alter the spacing and orientation of the phosphates and allow hydrogen bonding with the arginine side chain in Tat.

Use of small TAR derivatives has proven to be very useful in elucidation of the chemical basis for the specificity of the interaction between Tat and TAR RNA. Moreover, these derivatives can be synthesized in sufficient quantities for structural analysis of this interaction. These studies constitute important groundwork for the development of therapeutic agents targeted against Tat.

ACKNOWLEDGMENTS

The two first authors contributed equally to this work.

We thank Craig Rosen (Roche Institute of Molecular Biology, Nutley, N.J.) for his generous gift of purified Tat protein, Heather Erfle and Yan Kwok for their contribution to the analysis of RNA products, and John Silvius for helpful comments throughout the course of this work.

This work was supported in part by a subcontract under NIH grant AI27221 to M.S.-S. and grants to N.S. from the American Foundation for AIDS Research (000946-7-RGR) and the Cancer Research Society (Montreal). N.S. is the recipient of a Medical Research Council Scientist award from the Medical Research Council of Canada. S.R. is a recipient of a studentship from the Cancer Research Society (Montreal), and U.D. is a recipient of a studentship from the Medical Research Council of Canada.

REFERENCES

1. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. *Trans*-activator gene of human T-lymphotropic virus type III (HTLV-III). Science **229**:69–73.
2. **Atkinson, T., and M. Smith.** 1984. Solid phase synthesis of oligodeoxynucleotides by the phosphite-triester method, p. 35–81. *In* M. J. Gait (ed.), Oligonucleotide synthesis: a practical approach. IRL Press, Cambridge.
3. **Berkhout, B., and K.-T. Jeang.** 1989. *trans* activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the *trans*-acting-responsive hairpin: a quantitative analysis. J. Virol. **63**:5501–5504.
4. **Calnan, B. J., S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. Genes Dev. **5**:201–210.
5. **Calnan, B. J., B. Tidor, S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Arginine-mediated RNA recognition: the arginine fork. Science **252**:1167–1171.
6. **Caruthers, M. H., A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. Matteuci, Z. Stabinsky, and J.-Y. Tang.** 1987. Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. Methods Enzymol. **154**:287–313.
7. **Cordingley, M. G., R. L. LaFemina, P. L. Callahan, J. H. Condra, V. V. Sardana, D. J. Graham, T. M. Nguyen, K. LeGrow, L. Gotlib, A. J. Schlabach, and R. J. Colonno.** 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.
8. **Cullen, B. R.** 1990. The HIV-1 Tat protein: an RNA sequence-specific processivity factor? Cell **63**:655–657.
9. **Dayton, A. I., J. G. Sodroski, C. A. Rosen, W. C. Goh, and W. A. Haseltine.** 1986. The *trans*-activator gene of the human T cell lymphotropic virus type III is required for replication. Cell **44**:941–947.

9a. **Delling, U.** Unpublished data.

10. **Delling, U., S. Roy, M. Sumner-Smith, R. Barnett, L. Reid, C. A. Rosen, and N. Sonenberg.** 1991. The number of positively charged amino acids in the basic domain of Tat is critical for trans-activation and complex formation with TAR RNA. Proc. Natl. Acad. Sci. USA **88**:6234–6238.
11. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 *tat* protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. EMBO J. **9**:4145–4153.
12. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, M. A. Skinner, and R. Valerio.** 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA *in vitro*. Proc. Natl. Acad. Sci. USA **86**:6925–6969.
13. **Donis-Keller, H.** 1980. Phy M: an RNase activity specific for U and A residues useful in RNA sequence analysis. Nucleic Acids Res. **8**:3133–3142.
14. **Eadie, J. S., L. J. McBride, J. W. Efcavitch, L. B. Hoff, and R. Cathcart.** 1987. High performance liquid chromatographic analysis of oligodeoxyribonucleotide base composition. Anal. Biochem. **165**:442–447.
15. **Feng, S., and E. C. Holland.** 1988. HIV-1 *tat trans*-activation requires the loop sequence within *tar*. Nature (London) **334**:165–168.
16. **Fisher, A. G., M. B. Feinberg, S. F. Josephs, M. E. Harper, L. M. Marselle, G. Reyes, M. A. Gonda, A. Aldovini, C. Debouk, R. C. Gallo, and F. Wong-Staal.** 1986. The *trans*-activator gene of HTLV-III is essential for virus replication. Nature (London) **320**:367–371.
17. **Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor.** 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. EMBO J. **8**:765–778.
18. **Gatignol, A., A. Kumar, A. Rabson, and K.-T. Jeang.** 1989. Identification of cellular proteins that bind to the human immunodeficiency virus type 1 trans-activation-responsive *TAR* element RNA. Proc. Natl. Acad. Sci. USA **86**:7828–7832.
19. **Gaynor, R., E. Soultanakis, M. Kuwabara, J. Garcia, and D. S. Sigman.** 1989. Specific binding of a HeLa cell nuclear protein to RNA sequences in the human immunodeficiency virus transactivating region. Proc. Natl. Acad. Sci. USA **86**:4858–4862.
20. **Gentz, R., C.-H. Chen, and C. A. Rosen.** 1989. Bioassay for trans-activation using purified human immunodeficiency virus *tat*-encoded protein: trans-activation requires mRNA synthesis. Proc. Natl. Acad. Sci. USA **86**:821–824.
21. **Hauber, J., and B. R. Cullen.** 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. **62**:673–679.

22. **Heidelberger, C.** 1965. Fluorinated pyrimidines, p. 1–50. *In* J. N. Davidson and W. E. Cohn (ed.), Progress in nucleic acid research and molecular biology. Academic Press, Inc., New York.
23. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86:**7706–7710.
24. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1990. Predicting optimal and suboptimal secondary structure for RNA. Methods Enzymol. **183:**281–306.
25. **Jakobovits, A., D. H. Smith, E. B. Jakobovits, and D. J. Capon.** 1988. A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV *trans* activator. Mol. Cell. Biol. **8:**2555–2561.
26. **Koontz, S. W., and P. R. Schimmel.** 1979. Aminoacyl-tRNA synthetase-catalyzed cleavage of the glycosidic bond of 5-halogenated uridines. J. Biol. Chem. **254:**12277–12280.
27. **Marciniak, R. A., B. J. Calnan, A. D. Frankel, and P. A. Sharp.** 1990. HIV-1 Tat protein *trans*-activates transcription in vitro. Cell **63:**791–802.
28. **Marciniak, R. A., M. A. Garcia-Blanco, and P. A. Sharp.** 1990. Identification and characterization of a HeLa nuclear protein that specifically binds to the trans-activation-response (TAR) element of human immunodeficiency virus. Proc. Natl. Acad. Sci. USA **87:**3624–3628.
29. **Meerovitch, K., J. Pelletier, and N. Sonenberg.** 1989. A cellular protein that binds to the 5′-noncoding region of poliovirus RNA: implications for internal translation initiation. Genes Dev. **3:**1026–1034.
30. **Muesing, M. A., D. H. Smith, and D. J. Capon.** 1987. Regulation of mRNA accumulation by a human immunodeficiency virus *trans*-activator protein. Cell **48:**691–701.
31. **Müller, W. E. G., T. Okamoto, P. Reuter, D. Ugarkovic, and H. C. Schröder.** 1990. Functional characterization of Tat protein from human immunodeficiency virus. Evidence that Tat links viral RNAs to nuclear matrix. J. Biol. Chem. **265:**3803–3808.
32. **Ogilvie, K. K., N. Usman, K. Nicoghosian, and R. J. Cedergren.** 1988. Total chemical synthesis of a 77-nucleotide-long RNA sequence having methionine acceptance activity. Proc. Natl. Acad. Sci. USA **85:**5764–5768.
33. **Parkin, N. T., E. A. Cohen, A. Darveau, C. Rosen, W. Haseltine, and N. Sonenberg.** 1988. Mutational analysis of the 5′ noncoding region of human immunodeficiency virus type 1: effects of secondary structure on translation. EMBO J. **7:**2831–2837.
34. **Puglisi, J. D., J. R. Wyatt, and I. Tinoco.** 1990. Solution conformation of an RNA hairpin loop. Biochemistry **29:**4215–4226.
35. **Rosen, C. A., and G. N. Pavlakis.** 1990. Tat and Rev: positive regulators of HIV gene expression. AIDS **4:**499–509.
36. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of *cis*-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41:**813–823.
37. **Roy, S.** Unpublished data.
38. **Roy, S., U. Delling, C.-H. Chen, C. A. Rosen, and N. Sonenberg.** 1990. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated *trans*-activation. Genes Dev. **4:**1365–1373.
39. **Roy, S., N. T. Parkin, C. Rosen, J. Itovitch, and N. Sonenberg.** 1990. 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.
40. **Sadaie, M. R., T. Benter, and F. Wong-Staal.** 1988. Site-directed mutagenesis of two trans-regulatory genes (*tat-III*, *trs*) of HIV-1. Science **239:**910–913.
41. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Scaringe, S. A., C. Franckyl, and N. Usman.** 1990. Chemical synthesis of biologically active oligoribonucleotides using β-cyanoethyl protected ribonucleoside phosphoramidites. Nucleic Acids Res. **18:**5433–5441.
43. **Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop in *tar* determine transcriptional elongation by *tat* through the HIV-1 long terminal repeat. Genes Dev. **3:**547–558.
44. **Sinha, N. D., J. Biernat, J. McManus, and H. Köster.** 1984. Polymer support oligonucleotide synthesis XVIII: use of β-cyanoethyl-N,N-dialkylamino-/N-morpholino phosphoramidite of deoxynucleosides for the synthesis of DNA fragments simplifying deprotection and isolation of the final product. Nucleic Acids Res. **12:**4539–4557.
45. **Sodroski, J., C. Rosen, F. Wong-Staal, S. Z. Salahuddin, M. Popovic, S. Arya, R. C. Gallo, and W. A. Haseltine.** 1985. *Trans*-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. Science **227:**171–173.
46. **Starzyk, R. M., S. W. Koontz, and P. Schimmel.** 1982. A covalent adduct between the uracil ring and the active site of an aminoacyl tRNA synthetase. Nature (London) **298:**136–140.
47. **Tang, R. S., and D. E. Draper.** 1990. Bulge loops used to measure the helical twist of RNA in solution. Biochemistry **29:**5232–5237.
48. **Weeks, K. M., C. Ampe, S. C. Schultz, T. A. Steitz, and D. M. Crothers.** 1990. Fragments of the HIV-1 Tat protein specifically bind TAR RNA. Science **249:**1281–1285.
49. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.