10-1992

Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator

Erwann P. Loret

Philippe T. Georgel
Marshall University, georgel@marshall.edu

W. Curtis Johnson Jr.

Pui Shing Ho

Recommended Citation
Loret EP, Georgel P, Johnson WC, Ho PS. Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator. Proceedings of the National Academy of Sciences. 1992;89(20):9734-8.

Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 9734–9738, October 1992
Biophysics

Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator

(nucleic acid protein interaction/circular dichroism/energy minimization/AIDS)

ERWANN P. LORET†‡, PHILIPPE GEORGEL§, W. CURTIS JOHNSON, JR.§, AND PUI SHING HO§

†Laboratoire de Biochimie, Centre National de la Recherche Scientifique, Unite de Recherche Associée 1179, Faculté de Médecine, Secteur Nord, Boulevard Pierre Dramard, 13326 Marseille Cedex 15, France; and §Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331-6503

Communicated by Kensal E. van Holde, July 6, 1992 (received for review March 17, 1992)

ABSTRACT **Transcription in the human immunodeficiency virus type 1 (HIV-1) retrovirus is regulated by binding the viral Tat protein (trans-acting transcriptional activator) to the trans-activation response (TAR) RNA sequence. Here, vacuum UV circular dichroism (VUV-CD) is used to study the structure of TAR and its complex with two peptide fragments that are important for Tat binding to TAR. The VUV-CD spectrum of TAR is typical of A-form RNA and is minimally perturbed when bound to either the short or the long Tat peptide. The CD spectra of the complexes indicate an extended structure in the arginine-rich region of Tat from amino acid residue 47 through residue 58 and a short α-helix within the adjacent 59–72 region. Models of TAR and its peptide complexes are constructed to integrate these spectroscopic results with current biochemical data. The model suggests that (*i*) the arginine-rich 49–58 region is primarily responsible for electrostatic interactions with the phosphates of the RNA, (*ii*) the arginine side chains can additionally interact with substituent groups of the nucleotide bases to confer base recognition in the complex, (*iii*) the recognition of uracil-23 in TAR is facilitated by the peptide backbone, and (*iv*) the glutamine-rich face of an α-helix within the 59–72 region pairs to bases UGG at nucleotide positions 31–33 in the TAR loop and thus provides an additional motif in the Tat trans-activating protein to recognize TAR RNA.**

Human immunodeficiency virus type 1 (HIV-1) possesses regulatory genes not found in other retroviral genomes (1). One of these genes encodes the trans-acting transcriptional activator (Tat protein), which upon binding to the nucleotide trans-activation response (TAR) sequence, activates all viral mRNA transcripts (see ref. 2 for a recent review). The TAR sequence is present at the 5′ end of all HIV-1 mRNAs (3), suggesting that the binding of Tat to the RNA form of TAR is an important step in regulating transcription of HIV-1 genes (4–12), although the mechanism of this trans-activation remains unclear.

Tat is a small protein of 86 residues; however, only the 72 N-terminal amino acids are required for full activity (13). Mutations within the arginine-rich region from amino acid residue 47 through residue 58 yield a nonfunctional cytoplasmic Tat protein (14), and short peptides corresponding to this polybasic region show high affinity for the TAR sequence. A 9-residue peptide containing only arginine residues also binds TAR with high affinity, while an analogous peptide containing only lysine residues does not (12). Similarly, mutant forms of Tat in which residues 47–55 have been replaced by arginine residues are fully active, while lysine substitutions yield a marginally active protein (12). Arginine-52 appears to be particularly important in binding TAR. One model proposes that the imino nitrogens of the guanidino side chain form specific networks of hydrogen bonds to bridge adjacent phosphodiesters, linking nucleotides G21 to A22 and A22 to U23 in the TAR sequence (12). Recent molecular models of the arginine-rich region of Tat resulting from CD studies suggest that this basic region is an extended structure (15) or is unstructured in the absence of RNA (10). Upon binding to the TAR RNA, however, the peptide may become partially or fully structured and induce a conformational change in the RNA (10). Lengthening the peptides to include residues 47–72 enhances the stability but not the affinity of complexes with TAR (7).

The sequence in TAR that is recognized by the Tat protein spans nucleotides +19 to +42, and forms a stable hairpin that includes a six-nucleotide loop, a trinucleotide bulge, and an 8-base-pair (bp) stem (3, 16, 17). Mutations in the loop are deleterious for transactivation but have no effect on the binding of Tat protein or the polybasic peptides (5, 6, 9). Peptide binding studies show that the trinucleotide bulge and 5 bp adjacent to the bulge are important for Tat binding (8, 11). U23 in the bulge is invariant and is required for nucleoprotein complex formation (5, 6, 18). A recent model of TAR suggests that the major groove, where the bulge is located, is similar to that of a B-form double helix and that the Tat protein binds within this wider major groove (11).

MATERIALS AND METHODS

Synthesis. The RNA was the gift of Gabriele Varani, and Tat-(47–72)-hexacosapeptide was synthesized as described (15). Tat-(47–58)-dodecapeptide was synthesized with an Applied Biosystems 431A synthesizer using "Fast-moc" chemistry.

VUV-CD Measurements. TAR and TAR–Tat complexes for vacuum UV circular dichroism (VUV-CD) measurements were in 10 mM sodium phosphate buffer, pH 7.5/5% (vol/vol) glycerol/70 mM KF (10). Molar ratios for complex preparations were 1:1 as monitored by HPLC. Peptide samples were prepared in 10 mM sodium phosphate buffer (pH 7.5), 10 mM NaOH (pH 11), or 80% (vol/vol) CF_3CH_2OH. CD spectra were measured in 50-μm path-length cells from

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.

Abbreviations: HIV-1, human immunodeficiency virus type 1; TAR, trans-activation response element; Tat protein, trans-acting transcriptional activator of HIV-1; C_{18} HPLC, reverse-phase high-performance liquid chromatography on C_{18} columns; VUV, vacuum ultraviolet; Tat-(47–58), peptide matching Tat sequence 47-58; Tat-(47–72), peptide matching Tat sequence 47-72.

‡To whom reprint requests should be addressed.

320 nm to 178 nm with a VUV-CD spectrophotometer as described (15). Measurements were performed at 20°C for the peptides alone and 5°C for TAR and its peptide complexes (5°C was required to stabilize the complexes). CD spectra are reported as the difference in extinction coefficient, $\Delta\varepsilon$, per amide or per nucleotide. The VUV-CD data were analyzed for secondary structure by the method of Manavalan and Johnson (19). Purified peptides were analyzed for amino acid content and concentration on a Beckman 126AA system gold HPLC amino acid analyzer. The TAR extinction coefficient is 229,500 $M^{-1}\cdot cm^{-1}$ at 260 nm per TAR molecule.

Molecular Models of TAR and Its Tat Peptide Complexes. Initial models for the TAR RNA alone and TAR complexed with the Tat peptides were constructed by using the INSIGHTII program from Biosym Technologies, similar to previous work (15). Each model was subjected to energy minimization, followed by 1 ps of molecular dynamics at 500 K after equilibration, and then final energy minimization to a maximum derivative of 1.0 kcal per step by using the program DISCOVER from Biosym Technologies, with the AMBER (20) forcefields, running on a Silicon Graphics 4D30GTX Iris personal workstation. The structures were optimized in terms of the internal energies within each molecule and of the interaction energies between the molecules.

RESULTS AND DISCUSSION

CD Spectrum of TAR RNA. The VUV-CD spectra of double-stranded DNA and RNA are highly sensitive to the conformation of the polynucleotide helix. The spectrum of TAR RNA alone is typical for RNA and is usually associated with the A-conformation of a polynucleotide helix (Fig. 1).

CD Spectra of the Tat Peptides. The VUV-CD of Tat-(47–58) (Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg-Pro) at pH 7, pH 11, and in 80% CF_3CH_2OH is characterized by a negative band about 198 nm, which is normally attributed to a random coil structure (Fig. 2). Analysis of these spectra by the variable selection method gives antiparallel β-strand, β-turn, and other structure (Table 1). Random coil is not "other," but is a dynamic system with a propensity for the ϕ–Ψ angles associated with β-strand and β-turn. These analyses are typical for a random-coil CD.

The VUV-CD (Fig. 2) and subsequent secondary structure analyses (Table 1) of the longer Tat-(47–72) peptide (Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg-Pro-Pro-Gln-Gly-Ser-Gln-Thr-His-Gln-Val-Ser-Leu-Ser-Lys-Gln) at pH 7 and pH 11 are similar to those of Tat-(47–58) (Fig. 2). In 80% CF_3CH_2OH, however, there is a significant increase in α-helix structure, which was not observed with the short peptide (Table 1). This difference is well illustrated in the emergence of a broad negative band at long wavelength, a positive band at 187 nm, and a red shift from 197.5 to 201 nm of the original negative band. This same CD, showing a higher α-helical content, was obtained in the presence of 2 mM SDS (results not shown). This increase in α-helix structure is most likely located at residues 59–72, since Tat-(47–58) in 80% CF_3CH_2OH shows no α-helical structure.

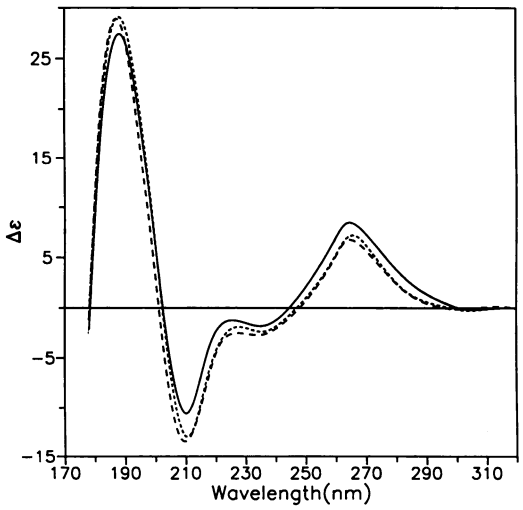

FIG. 1. VUV-CD spectra at 5°C of TAR (——), TAR–Tat-(47–58) complex (- - -), and TAR–Tat-(47–72) complex (— — —) in 10 mM phosphate buffer, pH 7.5/70 mM KF/5% glycerol, expressed as $M^{-1}\cdot cm^{-1}$ per base.

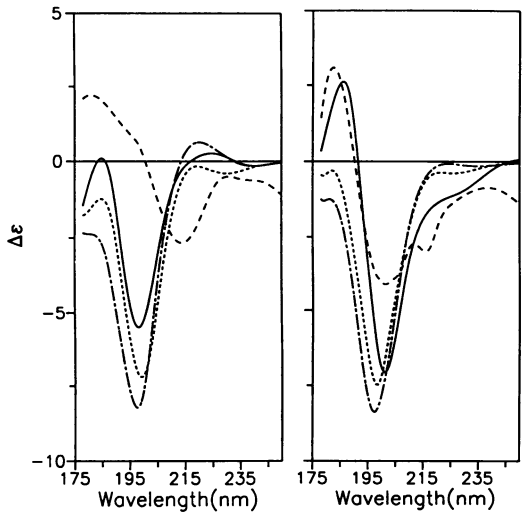

FIG. 2. VUV-CD spectra of the peptides Tat-(47–58) (*Left*) and Tat-(47–72) (*Right*) in 10 mM phosphate buffer (pH 7.5) (— · —), in 10 mM NaOH (pH 11) (- - -), and in 80% CF_3CH_2OH (——) as $M^{-1}\cdot cm^{-1}$ per amide basis. Difference spectra of the TAR–Tat peptide complexes minus the TAR RNA spectrum (– – –) are represented along with the corresponding peptide as $M^{-1}\cdot cm^{-1}$ per base.

CD Spectrum of the TAR–Tat-(47–58) Complex. The CD of the complex between TAR and the Tat-(47–58) peptide is dominated by the RNA (Fig. 1). Moreover, the modest CD due to the protein contributes only at a wavelength below 240 nm (Fig. 2). The 265-nm band of the RNA is sensitive to base stacking. The decrease in $\Delta\varepsilon$ upon forming the complex can be interpreted as a modification of the base stacking induced by the binding of Tat-(47–58). Overall, the spectrum, especially the 210-nm band, is still typical of the A-form; thus the TAR structure is only minimally perturbed in the complex.

A decrease in intensity for the 265-nm band of RNA is always accompanied by a decrease in intensity for the shorter wavelength bands, as shown in melting (21) and solvent perturbation studies (22). In these cases, the first crossover is virtually unchanged. Alternatively, when an A-form polynucleotide is perturbed toward B-form, the 265-nm band decreases and the 240-nm band increases in intensity with a

Table 1. Secondary structures predicted by variable selection for Tat-(47–58) and Tat-(47–72) in phosphate buffer (pH 7), sodium hydroxide (pH 11), or 80% CF_3CH_2OH

Peptide	Condition	H	A	P	T	O	Tot
Tat-(47–58)	pH 7	0.00	0.17	0.04	0.36	0.43	1.00
	pH 11	0.02	0.12	0.05	0.33	0.48	1.00
	80% TFE	−0.01	0.32	0.03	0.22	0.39	0.96
Tat-(47–72)	pH 7	0.04	0.11	0.01	0.37	0.48	1.00
	pH 11	0.06	0.16	0.00	0.31	0.46	1.00
	80% TFE	0.14	0.15	0.01	0.29	0.41	1.00

TFE, CF_3CH_2OH; H, helix; A, antiparallel β-sheet; P, parallel β-sheet; T, turn; O, other structure; Tot, total.

red shift of the crossover, while the shorter wavelength bands decrease in intensity (22). This observation is similar to the effect observed for the TAR complex (Fig. 1), including the increase in CD to the red-absorbing wavelengths of the protein bands. This suggests that upon binding, the TAR RNA assumes a somewhat more B-like character. In any event, the increases in the magnitude of the CD that we observe below 240 nm must be due to the protein. The difference CD shows a maximum at 185 nm and a minimum at 215 nm (Fig. 2). The negative band at 215 nm cannot be due to the RNA and matches very well with the CD of a β structure (21). If we scale the CD of the TAR RNA alone by the same percentage at all wavelengths to fit the 265-nm band of the complex (as suggested by solvent perturbations), we obtain a similar difference spectrum, but with intensities comparable to that of the uncomplexed peptides (results not shown). Together, these observations lead to the same interpretation—that Tat-(47–58) adopts an extended structure upon binding to TAR.

CD of the TAR–Tat-(47–72) Complex. The complex of TAR with the Tat-(47–72) peptide shows a CD that is very similar to that obtained for the TAR–Tat-(47–58) complex (Fig. 1). A clear difference occurs between 235 and 223 nm with a maximum at 225 nm. The intensity of the negative band near 210 nm is the same in both complexes but is slightly blue-shifted to 209 nm in the TAR–Tat-(47–72) complex. The crossover is blue-shifted from 203.5 to 201 nm. These results indicate that addition of the C-terminal sequences causes little additional modification of the TAR structure; therefore, the primary structural perturbations to TAR are associated with effects from the 47–58 region of the peptide.

The difference CD spectrum derived from the TAR–Tat-(47–72) complex and TAR alone shows two minima at 204 nm and 215 nm and a maximum at 185 nm (Fig. 2). Again, scaling the CD of TAR to account for solvent perturbation or B-form structure does not significantly change the extrema but does increase the intensity of the bands in the difference spectrum. The 215-nm band can be attributed to the β-structure of the polybasic region that is common between the short and the long peptide. However, the 204-nm band shows that there is an α-helix in the Tat peptide complexed to TAR.

The spectra in Fig. 2 are plotted on a per amide basis so that the bands associated with the β-structure would appear to be half the intensity in the longer oligomer as observed for the shorter peptide. The minimum at 204 nm is similar to the π–π^* transition at 205–208 nm of an α-helix (21). For short α-helices (seven to eight residues long), the characteristic CD bands are blue-shifted, with the positive band located at 185–187 nm and the negative band at 204–205 nm (23), as observed here. These results suggest the presence of a short α-helix, probably no longer than two turns, located at the C-terminal end of the Tat-(47–72) peptide when complexed to TAR. The n–π^* transition at 222 nm that is characteristic of an α-helix is reduced in intensity and broadened to a non-distinct band in the spectra of shorter peptides. In the present difference spectrum, the 215-nm band overlaps and obscures the n–π^* transition of the short α-helix, but this feature could account for the difference observed between 223 nm and 235 nm in the CD spectra of the TAR–Tat-(47–58) versus the TAR–Tat-(47–72) complexes (Fig. 1). The band at 204 nm is also observed for the Tat-(47–72) peptide alone in 80% CF_3CH_2OH and is associated with an increase in α-helix structure (Table 1). Thus, it appears that the 47–58 region is an extended structure, while a short α-helix of seven or eight residues is located in the C-terminal 59–72 region.

Model of TAR RNA. We used the crystal structure of residues 49–65 from phenylalanine t-RNA, plus the intervening G18 residue, to set guidelines for the structural features that are important in defining the conformation of the TAR RNA. The model for the TAR RNA follows the structural guidelines set by the t-RNA template (Fig. 3). The continuity of the A-helix is maintained, and base stacking and base pairing are maximized as much as possible. The base-paired regions in the TAR model adopt an A-helical conformation, as suggested by the CD. Residues C24 and U25 of the bulge extend away from the A-helix, while U23 is stacked between residues A22 and G26. This places a kink in the A-helix towards the major groove but otherwise leaves the base stacking and pairing of the stem intact. Interestingly, U25 appears to form a hydrogen bond to the 2′-hydroxy group of the ribose and to O-4 of the base at C24. This hydrogen-bonding scheme, plus the stacking of U23, may render the two uracil bases less susceptible to electrophilic reagents such as diethyl pyrocarbonate (11).

In the model it was logical to pair C30 and G34. Thus, a C·G base pair separates the UGG (positions 31–33) loop and a single-base bulge (A35). Residues G32, G33, and G34 are stacked, while the adenine of A35 forms a triple base pair with C30·G34.

Model of the TAR–Tat-(47–58) Complex. Weeks and Crothers (11) suggested that arginine-rich peptides sit in the major groove of the A-helix of TAR. Earlier studies suggested that this region of the Tat protein is extended (15) and may be induced to form a specific structure upon binding to the TAR RNA (10). The current CD results further suggest that the RNA is only minimally perturbed by binding of this polybasic region. A model for the complex was constructed by systematically adding amino acids to the major groove of the TAR model until the full peptide was generated. In building this model, we oriented the peptide with the C terminus towards the loop of the RNA. Studies have shown

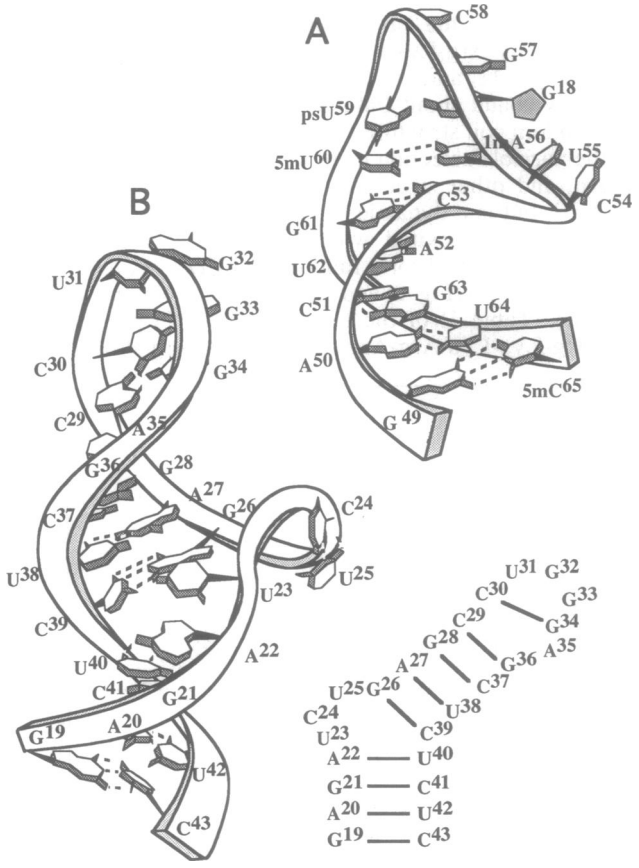

FIG. 3. Comparison of the phenylalanine t-RNA loop and bulge structure (*A*) with the TAR RNA model (*B*). Hydrogen bonds are drawn as broken lines. The base sequences are labeled along the ribbon backbone of each structure. The sequence of the TAR RNA is shown.

that U23 and at least one additional residue in the bulge of TAR and the G26·C39 and A27·U38 base pairs are required for specific binding by the Tat protein (11). In the peptide, Gln-54 is not crucial for binding to TAR; however, at least one arginine immediately preceding and at least two arginines immediately following Gln-54 are necessary. This suggests that Gln-54 straddles the bulge of the RNA, while one arginine contacts U23 and two arginines contact the G26·C39 and A27·U38 base pairs. The UCU bulge in the TAR sequence breaks the dyad symmetry of the A-helix. The peptide thus recognizes the RNA as an asymmetric structure and is skewed towards the UCU bulge. The role of the bulge in the TAR sequence, aside from U23, may therefore be primarily to orient the Tat peptide in the complex. In the final model of this complex, the peptide and the RNA appear to interact primarily through electrostatic attraction between the basic residues of the peptide and the negatively charged phosphates of the RNA (Fig. 4).

The arginines on either side of Gln-54 can also form hydrogen bonds with specific phosphate oxygens in a fork arrangement, as suggested by Calnan *et al.* (12). Unlike the previously proposed fork structure, however, the arginines in this model bridge the phosphate to nucleophilic substituents of the nucleotide bases in the major groove. In particular, Arg-55 bridges the phosphate between U25 and G26 with N-7 of G26, while Arg-56 bridges the phosphate between G26 and A27 to the N-7 of A27. This scheme is consistent with the observations that substituting the G26·C39 base pair with a U·A base pair or A27·U38 with either U·A or C·G base pairs diminishes the affinity of Tat for TAR (11).

A similar interaction is observed in this model between Arg-53 and the critical U23 residue in TAR. First, the base of U23 is pulled further into the major groove of TAR in this model of the complex. In this position, the side chain of Arg-53 bridges the phosphodiester between A22 and U23 and forms hydrogen bonds to the O-2 oxygen of U23. However, arginine side chains do not have the proper hydrogen bonding potential to recognize the acceptor–donor–acceptor hydrogen bonding pattern of a uracil base. The current model of the complex presents one possible solution to this problem by placing the amino nitrogen and carboxyl oxygen of the peptide backbone of Arg-53 3 Å from the extended uracil base. Thus, the backbone and side-chain functions of Arg-53 can provide the hydrogen bond donor–acceptor–donor pattern necessary to recognize the uracil base which is extended into the major groove.

Model of the TAR–Tat-(59–72) Complex. The difference CD of TAR complexed with Tat-(47–72) (Fig. 2) suggests that the C-terminal 14 amino acids adopt a predominantly α-helical conformation. The hydrogen bond donor–acceptor functions of the glutamine side chain are ideally suited for nucleotide base recognition and binding (24). The three glutamine residues between positions 59 and 66 are in an N+3 pattern, which defines a glutamine-rich face to an α-helix. In this arrangement, the glutamines are properly positioned to interact with the bases in the TAR RNA. A search through the Brookhaven crystallographic protein data base for this N+3 pattern of glutamines located four specific protein structural motifs. However, only the sequence Gln-Leu-Met-Gln-Ala-Ser-Gly of hemerythrin was fully an α-helix (25). This N+3 repeat of glutamines was located in the sequence but unfortunately not reported in the crystal structure of the coat protein from southern bean mosaic virus (26). Thus, although this peptide sequence may not have a distinct propensity to form an α-helix, the potential pairing of the glutamines with the exposed bases of the TAR RNA loop may work synergistically to induce such a structure.

The model of the Tat-(47–72) peptide complexed with TAR was built by extending the structure of Tat-(47–58) in the TAR complex with a 14-amino acid α-helix. The initial docking

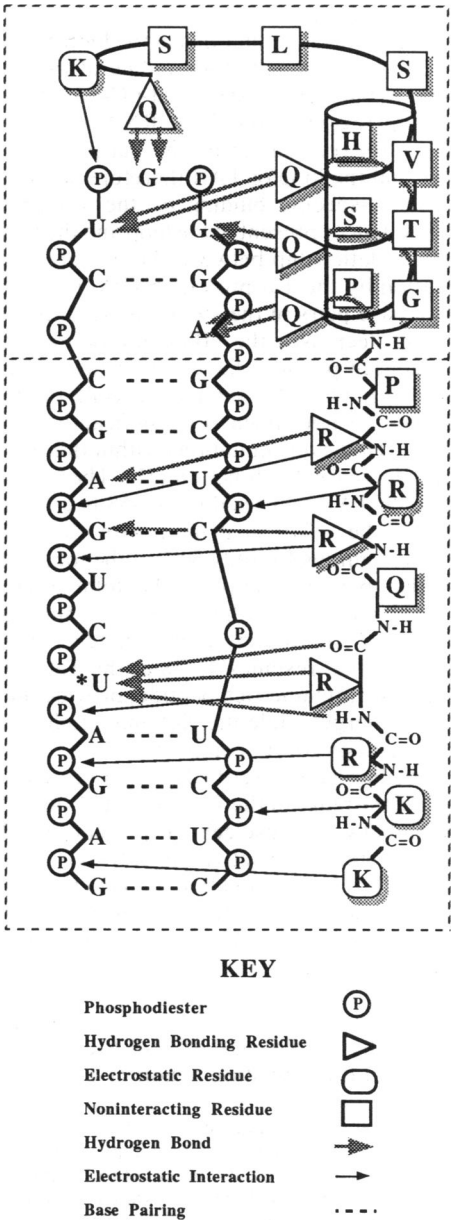

FIG. 4. Summary of interactions observed in the model of the TAR RNA complexed to Tat-(50–72). The diagram is divided by a dashed line into two segments corresponding to Tat-(59–72) (*Upper*) and Tat-(50–58) (*Lower*). The α-helix from Pro-59 to Val-67 is shown as a cylinder, with the glutamine-rich face oriented towards the loop of the TAR RNA.

studies showed that these glutamine residues in the α-helix could easily be positioned to interact with the non-base-paired residues U31, G32, G33, and A35 of TAR. Hydrogen bonds were thus specified between the peptide and the RNA residues for Gln-60 to A35, Gln-63 to G33, Gln-66 to U31, and Gln-72 to G32. These were given thermodynamic weights of 100 kcal/mol for the initial cycles of energy minimization and molecular dynamics, which were subsequently removed for the final cycles of molecular simulation. In the final model (Fig. 4), residues 59–66 remain α-helical, while the remainder of the C-terminal tail relaxes to form a loop structure. The additional unstacking of the now peptide-paired RNA bases is consistent with the observed perturbations to the CD of the RNA upon binding this longer peptide fragment.

This model is also consistent with the current biochemical data. RNase digestion studies have shown that the intact Tat

protein protects not only the bulge but also the loop region of TAR (27). Furthermore, mutations in the loop region of TAR have been shown to impede the binding of the longer Tat-(47–72) peptide but not the short Tat-(47–58) fragment (28). Thus, the loop region of the TAR RNA appears to be an additional site for recognition by the Tat protein.

The bulge and the loop of TAR therefore act in concert to confer sequence-specific binding by the Tat protein. The nucleotide sequences present in the loop and the bulge as well as the spatial relationship between the two are essential for trans-activation by the Tat protein in HIV-1 (16, 29). Insertion of extra base pairs in the stem that separates the bulge from the loop decreases the efficiency of trans-activation (30). Arginine-rich peptides have been shown to have very high affinities for the TAR sequence, while the C-terminal 59–72 residues do not appear to significantly enhance the affinity of Tat. However, mutations within this 59–72 region dramatically decrease Tat function (31). The intermolecular energy between TAR and residues 49–58 is calculated from our models to be −710 kcal/mol, with the electrostatic interactions accounting for fully 80% of the total energy. For the longer Tat peptide complex, the total intermolecular energy is −782 kcal/mol, with the four glutamine nucleotide pairs of our model contributing approximately −20 kcal/mol. Thus, it appears that residues 47–58 of the Tat protein, and indeed nearly any arginine-rich peptide, provide the majority of the binding energy, while the glutamine-rich-face confers additional specificity to the complex.

Previous studies have described cellular proteins that bind to the loop and the bulge of the TAR RNA (32–36). Therefore, a competition between these cellular proteins and Tat could be an important step in the activation of the virus. A host cell nuclear protein has been found that binds to the stem loop of TAR and promotes premature termination of transcripts from the HIV-1 promoter *in vivo* (37). Tat, with its higher affinity, could displace this protein from TAR and thus promote the elongation of the HIV-1 transcript, which is a Tat function (38). Another role of region 59–72 could be to facilitate the trans-activation by stabilizing the Tat–TAR complex, and to properly position the activating region of Tat.

We thank Dr. Gabriele Varani for providing us with the polynucleotide TAR and Dr. Eric Vives for providing us with the peptide Tat-(47–72). The authors are greatly indebted to Jeannine Lawrence and Scott Childs for their technical assistance. E.P.L. gratefully acknowledges Prof. Luc Montagnier and Dr. Clavel for their support of his research. This work was supported by grants from the Association pour la Recherche sur le Cancer, the National Institutes of Health (GM-21479 to W.C.J.), and the American Cancer Society (NP-740 to P.S.H.). P.S.H. is a recipient of a Junior Faculty Research Award (JFRA 306) from the American Cancer Society.

1. Arya, K. S., Guo, C., Josephs, S. F. & Wong-Staal, F. (1985) *Science* **229**, 69–73.
2. Cullen, B. (1990) *Cell* **63**, 655–657.
3. Muesing, M. A., Smith, D. H. & Capon, D. J. (1987) *Cell* **48**, 691–701.
4. Braddock, M., Chambers, A., Wilson, W., Esnouf, P. M., Adams, S. E., Kingsman, A. J. & Kingsman, S. M. (1989) *Cell* **58**, 269–279.
5. 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. USA* **86**, 6925–6929.
6. Roy, S., Delling, U., Chen, C.-H., Rosen, C. A. & Sonenberg, N. (1990) *Genes Dev.* **4**, 1365–1373.
7. Marciniak, R. A., Calnan, B., Frankel, A. D. & Sharp, P. A. (1990) *Cell* **63**, 791–802.
8. Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A. & Crothers, D. M. (1990) *Science* **249**, 1281–1285.
9. 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. USA* **87**, 8985–8989.
10. Calnan, J. B., Biancalana, S., Hudson, D. & Frankel, A. D. (1991) *Genes Dev.* **5**, 201–210.
11. Weeks, K. M. & Crothers, D. M. (1991) *Cell* **66**, 577–588.
12. Calnan, J. B., Tidor, B., Biancalana, S., Hudson, D. & Frankel, A. D. (1991) *Science* **252**, 1167–1171.
13. Cullen, B. R. (1986) *Cell* **46**, 973–982.
14. Hauber, J., Perkins, A., Heimer, E. P. & Cullen, B. R. (1989) *J. Virol.* **63**, 1181–1187.
15. Loret, E. P., Vives, E., Ho, P. S., Rochat, H., Van Rietschoten, J. & Johnson, W. C. (1991) *Biochemistry* **30**, 6014–6023.
16. Feng, S. & Holland, E. C. (1988) *Nature (London)* **334**, 165–167.
17. Garcia, J. A., Harrich, D., Soultanakis, E., Wu, F., Mitsuyasu, R. & Gaynor, R. B. (1989) *EMBO J.* **8**, 765–778.
18. Sumner-Smith, M., Roy, S., Barnett, R., Reid, L. S., Kuperman, R., Delling, U. & Sonenberg, N. (1991) *J. Virol.* **65**, 5196–5202.
19. Manavalan, P. & Johnson, W. C., Jr. (1987) *Anal. Biochem.* **167**, 76–85.
20. Weiner, S. J., Kollman, P. A. & Nguyen, D. T. (1986) *J. Comput. Chem.* **174**, 205–252.
21. Johnson, W. C., Jr. (1985) *Methods Biochem. Anal.* **31**, 61–163.
22. Goodman, M., Verdini, A. S., Toniolo, C., Phillips, W. D. & Bovey, F. A. (1969) *Proc. Natl. Acad. Sci. USA* **64**, 444–450.
23. Ivanov, V. I., Minchenkova, L. E., Schyolkina, A. K. & Poletayev, A. I. (1972) *Nucleic Acid Conf.* **5**, 90–110.
24. Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293–321.
25. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1983) *Acta Crystallogr. Sect. B* **39**, 697–703.
26. Silva, A. M. & Rossman, M. G. (1985) *Acta Crystallogr. Sect. B* **41**, 147–151.
27. Harper, J. W. & Logsdon, N. J. (1991) *Biochemistry* **30**, 8060–8066.
28. Kamine, J., Loewenstein, P. & Green, M. (1991) *Virology* **182**, 570–577.
29. Roy, S., Parkin, N. T., Rosen, C., Itovitch, J. & Sonenberg, N. (1990) *J. Virol.* **64**, 1402–1406.
30. Berkhout, B. & Jeang, K. T. (1991) *Nucleic Acids Res.* **22**, 6169–6176.
31. Kuppuswamy, M., Subramania, T. & Chinnadurai, G. (1989) *Nucleic Acids Res.* **17**, 3551–3561.
32. Nelbock, P., Dillon, P. J., Perkins, A. & Rosen, C. A. (1990) *Science* **248**, 1650–1653.
33. Gaynor, R., Soultanakis, E., Kuwabara, M., Garcia, J. & Sigman, D. S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4858–4862.
34. Marciniak, R. A., Garcia-Blanco, M. A. & Sharp, P. A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3624–3628.
35. Gatignol, A., Buckler-White, A., Berkuout, B. & Jeang, K. T. (1991) *Science* **251**, 1597–1600.
36. Sheline, C. T., Milocco, L. H. & Jones, K. A. (1991) *Genes Dev.* **5**, 2508–2520.
37. Rounseville, M. P. & Kumar, A. (1992) *J. Virol.* **66**, 1688–1694.
38. Marciniak, R. A. & Sharp, P. A. (1991) *EMBO J.* **13**, 4189–4196.