# Modeling of the Interaction between New Ethidium Derivatives and TAR RNA of HIV-1

R. Terreux,[†] D. Cabrol-Bass,*[,†] V. Peytou,[‡] R. Condom,[‡] and R. Guedj[‡]

GRECFO−Equipe LARTIC and Laboratoire de Chimie Bio-organique CNRS ESA 6001, Université de Nice Sophia Antipolis, F 06108 Nice Cedex 2, France

Received May 26, 1998

During the HIV-1 replication process, interactions between the first sequence of RNA synthesized named TAR RNA and a viral protein named Tat permit a fast and efficient transcription of viral DNA in RNA. Based on the NMR structure of TAR RNA found on the PDB, new derivatives of ethidium were designed by molecular modeling to inhibit this interaction. The studied molecules are composed of three domains: an arginine, a linker, and an ethidium. Three linkers of different lengths were considered in the first step, with the TAR RNA−arginine interaction and the intercalation of the ethidium simulated by docking methods. In a second step, the structure of the TAR RNA was completed to obtain a whole ethidium interaction site and docking of the whole studied molecules was investigated. Molecules were synthesized and tested on infected cells. The predicted models and activity are in good agreement with the reported experimental results.

## 1. INTRODUCTION

During its replication in the infected cell, the human immunodeficiency virus (HIV) uses several viral and cellular proteins. The transcription of DNA, one important step of this process, is accomplished by a number of cell transcription factors that activate polymerase II.[1,2] This set of proteins and DNA forms a complex structure that is still nowadays an object of discussion. The first synthesized sequence of bases called transactivation responsive element (TAR) RNA adopts a folded duplex structure.[3] It has been observed that the initial folding of nascent TAR RNA reduces the stability of the polymerase II−DNA complex.[1] The transactivator of transcription (Tat) protein plays a determinant role at this stage, by interacting with TAR RNA.[4,5] The new complex, Tat-TAR, stabilizes the complex polymerase II−DNA and allows the transcription of DNA to progress efficiently.[6−9] Therefore, inhibiting the formation of the Tat−TAR complex seems to be a promising target for reducing HIV replication.[10] With this goal in mind, we designed new ethidium derivatives as potential inhibitors of the interaction between the Tat protein and TAR RNA. Interaction of the proposed new molecules with TAR RNA was investigated by molecular modeling. These molecules were synthesized and tested on HIV-1 infected cells. The obtained results are in good agreement with our model.

## 2. STRUCTURES

The Tat protein is composed of a sequence of 86 amino acid residues for HIV-1 and 130 for HIV-2.[11,12] This sequence is divided into two parts. The first one, made of 19 residues, is variable depending upon the strain and has no bearing on the activity of the protein. The second one, made of 67 residues, can be divided further into 4 parts coded I, II, III, and IV. The basic domain included in part IV, made of residues 48−57 (GlyArgLysLysArgArgGlnArgArgArg), is directly involved in the interaction with TAR RNA.[1,13,14] While lysine residues contribute by electrostatic interactions, arginine 52 interacts more closely with the TAR RNA structure.[15]

HIV-1 TAR RNA is made of a sequence of 59 nucleotides forming an hairpin stem−loop. Aboul-ela et al.[3,16] studied by NMR a portion of TAR RNA made of 27 nucleotides including the stem−loop of 6 nucleotides. The sequence comprises two base pair regions, made respectively of bases $x + 1$ to $x + 5$ and bases $x + 9$ to 12, $x$ being the number of unstacked adenosyne A17 (see Figure 1). Between these two duplex regions are three free bases, UCU in most strains (UUU in rare cases), which create a bulge. The angle between the two symmetry duplex axes is directly induced by the conformation of the bulge. On the NMR structure of TAR RNA free of arginine, this angle of 91° creates a pocket in the RNA structure. This pocket is the interaction site for the basic domain of the Tat protein.

The three-dimensional structure of part IV of the Tat protein in interaction with TAR RNA has been determined experimentally by 2D NMR spectroscopy by Puglisi et al.[13,17,18] Examination of this model structure shows that part IV of the Tat protein interacts with the major groove of TAR RNA between bases $C_{x+1}$ and $A_{x+5}$. Also, the guadininium group of arginine 52 occupies the same location in the model of the TAR−arginine complex as proposed by Aboul-ela.[3,16]

Design of new molecules was made on the hypothesis that locking the TAR RNA major groove and the specific interaction site of arginine should inhibit the Tat−TAR complex formation. The molecules constructed following this idea are made of three parts, each having a specific function, namely, an arginine residue, a linker, and an intercalator.

The arginine residue can lead the molecule to its specific interaction site on the TAR RNA. The function of the linker is 2-fold: first, it should take position in the TAR RNA major groove and block it in order to prevent its interaction with the Tat protein; second, it connects the arginine residue and

* To whom correspondence should be addressed.
[†] GRECFO−Equipe LARTIC.
[‡] Laboratoire de Chimie Bio-organique CNRS ESA 6001.

10.1021/ci980094k CCC: $18.00 
Published on Web 03/02/1999

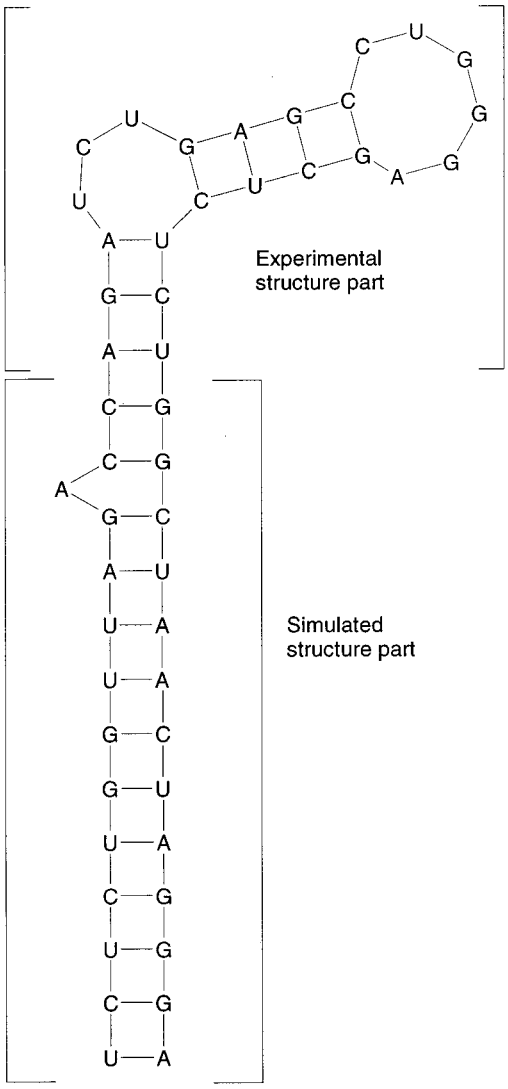


**Figure 1.** Scheme of the investigated TAR RNA model.

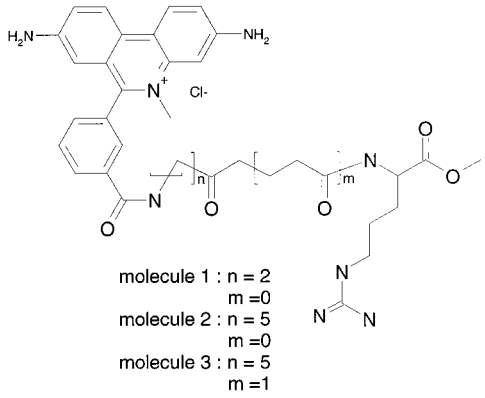


**Figure 2.** Structural formulas of molecules 1−3.

the intercalator. Linkers of three different lengths were considered. An ethidium derivative was used as the intercalator; its function is to keep the linker tightly in the major groove. The structural formulas of the three molecules, coded respectively 1, 2, and 3, are given in Figure 2.

## 3. MOLECULAR MODELING STUDY

To evaluate the designed molecules and to choose the best linker length, the interaction between TAR RNA and molecules 1−3 was investigated by molecular modeling. The study was conducted along the following steps: approach of the arginine residue and interaction with TAR RNA, intercalation of ethidium between bases pairs, and location of the linker in the major grove.

**3.1. Simulation of the RNA structure.** The structures of TAR, with and without an associated arginine group, were determined experimentally in order to define the 3D structures of the loop and the bulge.[3,16] The studied sequence extends from $C_{x+1}$ to the complementary base $G_{x+27}$ (see Figure 1). The possible insertion site of the ethidium of the considered molecules is located at the very beginning of this sequence. Furthermore, the base $A_x$, which is known to be unstacked, is not included in the published experimental structure. Therefore, the reported geometry for the insertion site is not useful for the present study.

Clearly, it is not possible to obtain a reliable 3D structure only by modeling RNA with a large number of bases. However, starting from the consistent experimental data, it seems reasonable to extend the known structure by about a dozen base pairs assuming they form a standard RNA duplex with the exception of the base $A_x$. The angle between the duplex sections and the geometry of this exteriorized base remains unknown. The retained experimental and modeled parts of the investigated TAR RNA are indicated in Figure 1.

In the modeling, a set of geometric constraints has been imposed by defining cutoff values above which a multiplying factor of 10 is applied to the force-field parameters.

The following distance constraints were set up: between complementary bases (cut off = 2.1 Å); between the phosphorus atom of a given base and the phosphorus atom of the next one (cutoff = 5 Å); between the phosphorus atom of a given base and the phosphorus atom of bases up to one turn away (cutoff = 10 Å).

The entire geometry of the modeled part was energetically optimized with this set of constraints by an annealing method. The retained experimental part was added to the modeled one, and energetic minimization was applied to the junction between the two. Figure 3 shows the obtained structure with arginine.

The set of constraints used leaves the RNA duplex structures almost unchanged while the angle between the axes of these two duplex sections remains low. However, the exteriorized base increases the volume of the major groove pocket. This conclusion is in good agreement with the work of Yao.[19] The resulting model was used in the subsequent investigation of the interaction between molecules 1−3 and TAR RNA.

**3.2. Approach of the Arginine Group to TAR RNA.** The aim of this preliminary study is 2-fold: First, to explore if the arginine could interact with TAR RNA on other sites than the one identified in the experimental structure; second, to identify possible paths for arginine to approach its interaction site. A better understanding of the approach of arginine is critical to evaluate if the studied molecules, which have larger sizes, might be able to approach and interact with TAR RNA following a similar path.

The three-dimensional structures of TAR RNA, in its free form or in interaction with arginine, used for this study were taken from the Brookhaven Protein Databank 1ANR. A comparison of these structures shows that interaction with

**Figure 3.** Three-dimensionnal structure of the TAR RNA in interaction with arginine.

**Table 1.** Energetic Interaction Components with Respect to the Guanidinium−$C_{x+7}$ Distance

| | 30 Å | 15 Å | 10 Å | 6 Å | 4 Å | 2 Å |
|---|---|---|---|---|---|---|
| van der Waals, kcal·mol$^{-1}$ | −0.05 | −1.26 | −1.84 | −3.93 | −1.21 | −7.32 |
| electrostatic, kcal·mol$^{-1}$ | 0.90 | 1.06 | 1.11 | 0.02 | −7.81 | −6.02 |
| total, kcal·mol$^{-1}$ | 0.85 | −0.20 | −0.72 | −3.91 | −9.02 | −13.34 |

arginine induces two important structural modifications of the RNA, i.e., affecting the sequence $C_{x+12}$, $G_{x+19}$ nucleotides hairpin, on one hand, and the $U_{x+6}$, $C_{x+7}$, $U_{x+8}$ nucleotides bulge, on the other hand.

The study of the arginine approach was conducted in two successive steps. First, the interaction energy between TAR RNA and arginine was calculated by moving manually the arginine molecule around the TAR RNA in order to discover all favorable energetic areas. Second, the approach of arginine following different paths was simulated in order to evaluate which one of the areas thus determined is the most favorable.

The arginine residue interacts with the basic parts of the TAR RNA molecule by its guanidinium group, which is partially positively charged. The most favorable area is located in a pocket created by the bulge and the two RNA duplex strands. During the approach, the guanidinium group is oriented toward TAR RNA, and arginine is moved along the linear pathway which strains away from the $C_{x+7}$ base in the less hindered direction. Table 1 gives the values of the energetic interaction components with respect to the distance between the guanidinium group of the arginine and the center of the aromatic cycle of the $C_{x+7}$ base. The distance was manually varied from 30 down to 2 Å. The stabilization energy is approximately estimated to be 14 kcal·mol$^{-1}$. See the Procedure section for technical details.

For closer distances, it is necessary to take into account the local structural changes of TAR RNA. Therefore, the docking of arginine at the interaction site was simulated from a distance of 2 Å, leaving the three bases free to adopt the most favorable orientation. These simulations confirm the high affinity of the arginine toward the two bases $U_{x+6}$ and $C_{x+7}$ of the bulge. The interaction of arginine with these bases induces their rotation, leading to their "unstacking". Figure 4 shows the first step of this interaction, before the rotation of the bases. Yellow straight lines represent the interactions between two atoms. The resulting repulsive forces, represented on this figure by red vectors, can explain the rotation of the three bases of the bulge. Once this rotation is initiated, the arginine interacts more closely with bases pair $G_{x+9}C_{x+22}$. The final location of arginine in the interaction with TAR RNA obtained by this docking procedure is in agreement with the experimental results of Aboul-ela.[3,16]

The rotation of the bulge bases is concomitant with two major structural changes in the TAR RNA conformation. The first one concerns the folding of the whole structure, which results in a reduction of the angle between the two RNA duplex strands from 91° to 85°. This folding up leads to an important reduction of the internal cavity volume, locking arginine in the interaction site. The second major structural modification is related to the orientation of the loop made of six nucleotides ($C_{x+13}$, $A_{x+18}$). While in the absence of arginine, the bases of these nucleotides are oriented toward the bulge UCU; this is no longer the case in the folded structure with arginine.

**3.3. Intercalation of Ethidium between Base Pairs.** The ethidium group is well-known to be an effective nucleotide intercalator. Its structure is made of a phenanthrene substituted by a nitrogen atom and a fourth α aromatic ring. As shown experimentally by White and Draper,[20,21] the ethidium intercalation is preferential between the C−G and G−C base

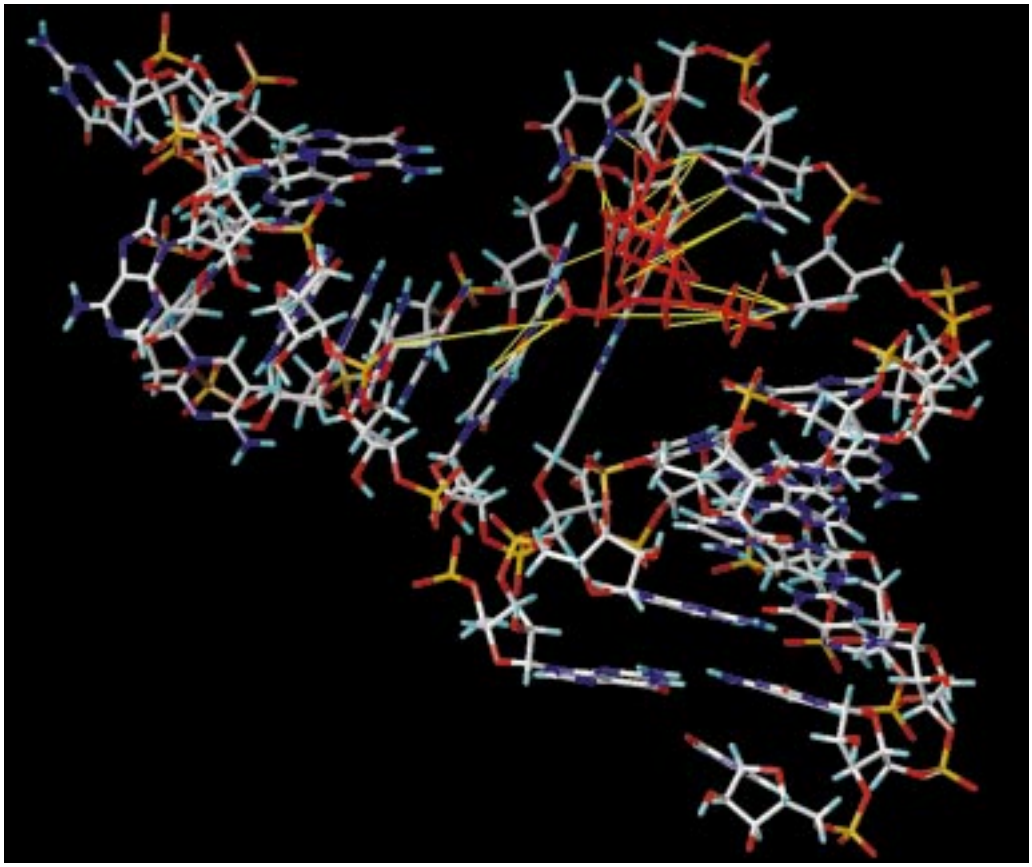

**Figure 4.** Visualization of the interaction between arginine and TAR RNA.

pairs. Energetics calculations of ethidium intercalation between the various base pairs were performed in order to compare their energies. During ethidium intercalation, the RNA structure may undergo local deformations. Therefore, the approach of ethidium along the two groves of RNA was simulated to investigate the possible resulting deformations.

**3.3.1. Energetic Study of the Intercalation between Various Base Pairs.** Energy calculations were conducted using four base pairs in typical RNA duplex conformation, the active site being defined as the two central base pairs XY. All combinations of XY base pairs were considered. These two base pairs have been widened apart by 3.6 Å to accommodate the ethidium group that was manually positioned between them. Table 2 contains the results of the energetic interaction evaluation computed after minimization of the complex geometry by the conjugate gradient method using the CFF force field.

The values in Table 2 confirm the finding results obtained by Yao,[19] and they can be clustered into three cases:

(1) Intercalation of ethidium is the most favorable between the CG and GC base pairs. The energetic differences for the various combinations are too weak to conclude a significant preference between them.

(2) A slightly less favorable situation corresponds to CG or GC base pairs in the location YY′ and UA or AU in the location XX′.

(3) Intercalation between base pairs UA in the YY′ location is noticeably less favorable.

**3.3.2. Insertion Simulation.** There are two main pathways for the approach of ethidium to RNA, namely, by the major or by the minor grooves. The same RNA model of four base pairs with the UCCA sequence was used for this intercalation simulation. Several constraints were fixed to simulate the RNA continuity. For this purpose, the atom coordinates of the sugar cycles of the terminal bases were frozen. The

**Table 2.** Interaction Energies of Ethidium Intercalation between Various RNA Base Pairs

| type | | | energy, kcal·mol$^{-1}$ | | type | | | energy, kcal·mol$^{-1}$ | |
|---|---|---|---|---|---|---|---|---|---|
| C | | G | vdW | −38.66 | G | | C | vdW | −38.62 |
| | Et | | elect | −18.41 | | Et | | elect | −18.49 |
| C | | G | total | −57.07 | C | | G | total | −57.11 |
| C | | G | vdW | −39.65 | A | | U | vdW | −34.50 |
| | Et | | elect | −17.48 | | Et | | elect | −11.94 |
| G | | C | total | −57.13 | C | | G | total | −46.44 |
| C | | G | vdW | −38.03 | U | | A | vdW | −34.501 |
| | Et | | elect | −17.48 | | Et | | elect | −11.94 |
| A | | U | total | −55.52 | C | | G | total | −46.44 |
| C | | G | vdW | −39.50 | U | | A | vdW | −35.31 |
| | Et | | elect | −15.96 | | Et | | elect | −12.79 |
| U | | A | total | −55.48 | A | | U | total | −48.11 |
| G | | C | vdW | −38.71 | U | | A | vdW | −34.50 |
| | Et | | elect | −18.41 | | Et | | elect | −11.61 |
| G | | C | total | −57.12 | U | | A | total | −46.44 |
| C | | G | vdW | −36.38 | A | | U | vdW | −30.82 |
| | Et | | elect | −17.86 | | Et | | elect | −12.61 |
| A | | U | total | −54.24 | G | | C | total | −43.43 |
| G | | C | vdW | −44.6 | A | | U | vdW | −32.26 |
| | Et | | elect | −9.2 | | Et | | elect | −16.02 |
| U | | A | total | −53.8 | A | | U | total | −48.28 |
| U | | A | vdW | −34.43 | A | | U | vdW | −33.25 |
| | Et | | elect | −12.43 | | Et | | elect | −16.61 |
| G | | C | total | −46.88 | U | | A | total | −49.86 |

**Table 3.** Significant Energetic Term Variations before and after the Insertion of the Ethidium by the Major Groove

| energetic term, kcal·mol$^{-1}$ | before | after | diff |
|---|---|---|---|
| total potential energy | 272.9 | 237.6 | −35.3 |
| nonbond | −209.6 | −245.3 | −35.7 |
| vdW | 10.8 | −15.4 | −26.2 |
| vdW repulsive | 547.2 | 599.2 | 52.0 |
| vdW dispersive | −536.4 | −614.6 | −78.2 |
| electrostatic | −220.4 | −229.9 | −9.5 |

adopted simulation procedure was the same for both pathways; technical details are given in the Procedure section.

**3.3.2.1. Approach and Insertion of Ethidium by the Major Groove.** The deformation of the RNA from its standard duplex conformation in order to allow insertion of ethidium requires some energy, which is compensated by the stabilizing interaction of ethidium with the base pairs. The deformation energy is estimated by computing the difference in energy between the standard RNA duplex conformation and the deformed one. A value of 9.97 kcal·mol$^{-1}$ was found. The same procedure was used to evaluate the energy of the system before and after the insertion of ethidium. The significant energetic terms are reported in Table 3.

The total energetic stabilization gained by insertion of the ethidium is 35.3 kcal·mol$^{-1}$. Close inspection of the resulting model shows that ethidium can occupy a position that is favorable to a $\pi-\pi$ interaction with the guanidine−cytosine base pair. As a matter of fact, analysis of the energetic terms reveals that this is the most important stabilizing factor. The formation of a hydrogen bond between the NH of the ethidium side chain and one of the two bases is also noticeable. This hydrogen bond contributes (as an additional factor) to the complex stabilization and maintains ethidium in its inserted position.

**3.3.2.2. Approach and Insertion of Ethidium by the Minor Groove.** The same procedure, models, and analysis were used to simulate the approach and the insertion of ethidium by the minor groove. Although the substituant of the model molecule is located on the opposite side of the RNA strand (relative to the previous case), the same values of energetic terms are obtained (see Table 3).

Thus, it can be concluded from this study that there is no preferential pathway for ethidium intercalation between RNA base pairs. The insertion is the most favorable with any guanine and cytosine combination of base pairs, but the difference in energy is not large enough to prevent insertion of ethidium between other base pair combinations.

**3.3.3. Identification of the Intercalation Site.** Looking at the TAR RNA sequence, one can identify two GC−GC base pairs not too far away from the arginine interaction site, namely, $C_{x+1}-G_{x+27}/C_{x+2}-G_{x+26}$ and $G_{x+11}-C_{x+20}/C_{x+12}-G_{x+19}$. The latter pair occupies a position in the upper part of the TAR RNA that is deformed by the six free nucleotides in the curve of the hairpin. Moreover, as pointed out above, the free bases of these nucleotides interact with the bulge. Therefore, it was concluded that this site is not favorable for ethidium intercalation. The other potential site of intercalation is located in a pocket composed of the major groove extended by the exteriorized base $A_x$. Subsequent work focuses on this site.

**3.4. Interaction Mechanism Hypothesis.** The data and conclusions drawn in the above sections were used to propose a hypothesis for the interaction mechanism of the studied molecules with TAR.

Molecules 1, 2, and 3 comprise three domains, namely, the arginine, the linker, and the ethidium group. As a first step, arginine interacts with the bulge and initiates some major local conformational changes described in the previous section. As a result of this initial step, arginine is located in its preferential pocket of interaction. The whole molecule takes an extended conformation, with the ethidium part oriented toward the outside. This complex will be referred to as state 1. To study the next step, the position of the guanidinium function of the arginine part in its interaction site and the entire TAR RNA were frozen by applying a fixed constraint. Only the linker and the ethidium were left free to move during the following energetic minimization cycles. During these cycles, the linker folds up, allowing the ethidium to interact with bulge bases, resulting in state 2. During the next series of minimization cycles, the ethidium slips along the RNA ribbon, extending the linker, which is now roughly parallel to the major groove (state 3). In this model, the molecule cannot penetrate into the narrow major groove. To verify if this result originates from the use of a fixed TAR RNA conformation, the major groove width has been investigated by molecular dynamics studies in aqueous phases. This analysis shows that the major groove width can vary to a large extent. The various widths predicted by the dynamic studies are highly correlated with those reported for different conformers based on NMR studies.[3] It is observed that the simulation could not lead the linker to enter into the major groove when keeping the TAR RNA conformation rigid, probably due to the small width of the conformer of state 3. Thus, the linker was inserted manually in the major groove, and several cycles of energetic minimization were allowed to investigate its accommodation in this location (state 4). The obtained results show that the linker does not induce any bump contact with the TAR RNA. On the contrary, the heteroatoms of the linker create interactions between the linker and the ribbon of the TAR RNA, which contribute to lock the major groove.

In the next step, the linker length is of critical importance with respect to the possibility of ethidium intercalation. For molecule 1, the linker is too short to allow the ethidium group to access the pocket composed of the major groove and the exteriorized base. Therefore, the ethidium group remains locked in the major groove. For this molecule, the insertion mechanism stops at state 4. In the case of molecule 2, the linker is too short as well; although the ethidium reaches the entrance of the pocket, it is not well located to intercalate itself between GC base pairs. Intercalation between another combination of bases is impossible because the ethidium group is too close to the beginning of the pocket and this constraint reduces its mobility. Contrary to the two previous cases, the ethidium group of molecule 3 is properly located in the pocket to allow its intercalation. This hypothesis was verified by several simulations following a procedure similar to the one used in the study of ethidium intercalation reported above (see section 3.3). Although the entire molecule is in complex association with TAR RNA, the local deformations of the base pairs are very similar.

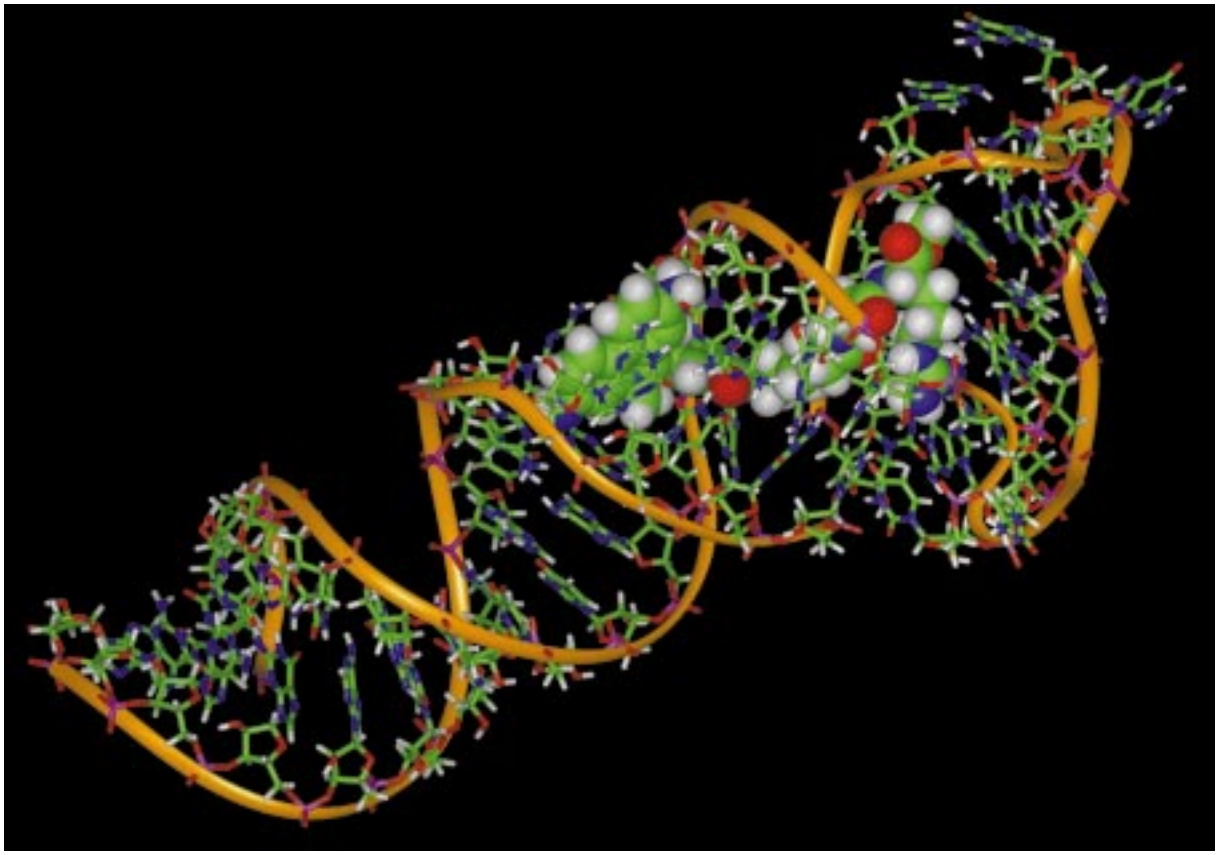

**Figure 5.** Final complex of molecule 3 with TAR RNA.

**Table 4.** Anti-HIV-1 Activity of Studied Molecules on CEM-SS Cells

| eval molec | IC 50, $\mu$M | CC 50, $\mu$M |
|---|---|---|
| 1, Eth ($n = 2$, $m = 0$) ArgOH | >100 | >100 |
| 2, Eth ($n = 5$, $m = 0$) ArgOH | >100 | >100 |
| 3, Eth ($n = 5$, $m = 1$) ArgOH | 2.4 | >100 |

**Table 5.** Anti-HIV-1 Activity of Studied Molecules on MT4 Cells

| eval molec | IC 50, $\mu$M | CC 50, $\mu$M |
|---|---|---|
| 1, Eth ($n = 2$, $m = 0$) ArgOH | >100 | 47 |
| 2, Eth ($n = 5$, $m = 0$) ArgOH | >100 | 59 |
| 3, Eth ($n = 5$, $m = 1$) ArgOH | 11 | >100 |

Figure 5 shows the final complex of molecule 3 with TAR RNA, as obtained from this interaction mechanism hypothesis. The major groove is locked by interaction with the molecule 3 linker heteroatoms. As mentioned above, it is expected that these structural modifications can inhibit the Tat protein−TAR RNA interaction.

## 4. BIOLOGICAL RESULTS

On evaluation *in vitro*, molecules 1 and 2 were inactive against HIV-1 at concentrations up to 100 $\mu$M (or at nontoxic concentrations), on cultures of both CEM-SS or MT4 cells. Under the same conditions, on cultures of CEM-SS cells acutely infected by HIV-1, compound 3 reduced viral proliferation by 50% at a micromolar concentration, without apparent cellular toxicity (Table 4). On the other hand, on cultures of MT4 cells, anti-HIV-1 activity of molecule 3 was strongly reduced (Table 5). This could be explained by the presence of Tat protein in MT4 cells, and these results suggested that the TAR−Tat interaction site is the targeted site of compound 3. This observation strongly suggested that there was no problem of cellular penetration of the acid form of the potential inhibitors.

Furthermore, experiments by fusion temperature show an increase of 12 °C of the dissociation temperature of wild TAR RNA with molecule 3, confirming the strong association.

## 5. PROCEDURE

The Sybyl 6.2 software package was used for the study of the approach of arginine to TAR RNA. The interaction energy was calculated with the MM2 force field, and minimization was performed using the conjugate gradient algorithm with 5000 iteration steps. Visualization of different interactions between arginine and TAR RNA were made with the Sybyl program (Figure 4). All other molecular modeling studies were performed using the Insight II software package. The energy was computed with the CFF force field, and minimizations were performed using the conjugate gradient algorithm with 9000 iteration steps or until convergence (defined as an energy gradient of 0.0001 kcal/mol or less). The annealing was defined by 50 cycles of 3000 minimization steps followed by 5000 molecular dynamic steps of 1 fs with an initial temperature of 300 K and an gradual increment to 800 K. All calculations were carried out on the IRIS indigo 2 R10.000 workstation of the "Centre de Modélisation et d'Imagerie Moléculaire" of UNSA.

## 6. CONCLUSIONS

The observed differences in activities between molecules can be explained by the last model presented in Figure 5.

Molecules 1 and 2 do not exhibit biological activities, while molecule 3 shows anti-HIV-1 activity at micromolar concentrations. The linkers of molecules 1 and 2 are too short to allow favorable interactions of ethidium with any base pairs. The superior part of the TAR RNA is too deformed and obstructed by the free hairpin to make ethidium intercalation between $G_{x+11}-C_{x+20}/C_{x+12}-G_{x+19}$ possible.

Our results lead to insightful remarks concerning the interaction site. Contrary to proteins, which have one or a few interaction sites, RNA has as many potential interaction sites as the base pairs. One could think that molecules 1−3 should be toxic because there are a large number of potential interaction sites on the RNAs, but experiments show this deduction to be wrong (CC 50 > 100 $\mu$M). This can be explained by the specificity of the studied molecules toward TAR RNA, which originate from the presence of the arginine end; therefore, the combination arginine−linker−ethidium is rather specific for TAR RNA. This observation supports the hypothesis that the arginine plays an important role in "piloting" the molecule to its interaction site. The hypothesis that the superior part of the TAR RNA is too deformed and obstructed is reinforced by the fact that for molecule 1 and 2 the only possible interaction site is located in this region.

The TAR RNA has the same selectivity to molecules as the proteases or reverse transcriptase proteins, which are the main targets of drugs against AIDS. The results of the biological tests validate the process of the RNA structure simulation on which our interaction model is partly based. The interaction model agrees with the one by Yao,[19] because the $A_x$ exteriorization increases the activity of active intercalator molecules. Molecule 3 is the only one to be active and to inhibit the interaction between Tat protein and the TAR RNA of HIV-1. Moreover, it has been demonstrated that molecule 3 is specific for TAR RNA, while only exhibiting a low level of cellular toxicity.

This study demonstrates that TAR RNA offers a new target for the design of inhibitors. The development of such inhibitors of the protein−RNA interaction is a promising strategy deserving further investigation, along with the protease and reverse transcriptase strategies, for combating HIV-1.

## ACKNOWLEDGMENT

The authors wish to acknowledge the association "Ensemble contre le SIDA" and ANRS (Agence Nationale de Recherche contre le SIDA" for supporting this work. We also thank Dr. A. Trapani for helping us in revising the manuscript.

## REFERENCES AND NOTES

(1) Karn, J.; Graeble, M. A. New insights into the mecanism of HIV-1 trans-activation. *Trends Genet.* **1992**, *8*, 365−368.

(2) Antoni, B. A.; Stein, S. B.; Rabson, A. B. Regulation of human immunodeficiency virus infection: Implications for pathogenesis. *Adv. Virus Res.* **1994**, *43*, 53−145.

(3) Aboul-ela, F.; Karn, J.; Varani, G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res.* **1996**, *24*, 3974−3981.

(4) Dingwall, C.; Ernberg, I.; Gait, M. J.; Green, S. M.; Heaphy, S.; Karn, J.; Lowe, A. D.; Singh, M.; Skinner, M. A. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *Embo J.* **1990**, *9*, 4125−4153.

(5) Kingsman, S. M.; Kingsman, A. J. The regulation of human immunodeficiency virus type-1 gene expression. *Eur. J. Biochem.* **1996**, *240*, 491−507.

(6) Graeble, M. A.; Churcher, M. J.; Lowe, A. D.; Gait, M. J.; Karn, J. Human immunodficiency virus type 1 transactivator protein, tat, stimulates transcriptional real-through of distal terminator sequences in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6184−6188.

(7) Laspia, M. F.; Rice, A.; Mathews, M. B. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* **1989**, *59*, 283−292.

(8) Jones, K. A.; Peterlin, B. M. Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* **1994**, *63*, 717−743.

(9) Jeang, K. T.; Chun, R.; Lin, N. H.; Gatignol, A.; Glabe, C. G.; Fan, H. In vitro and In vivo binding of human immunodeficiency virus type-1 Tat protein and Sp1 transcription factor. *J. Virol.* **1993**, *67*, 6224−6233.

(10) Sinet, M.; Kubar, J.; Condom, R.; Patino, N.; Guedj, R. The protein Tat of HIV−Potential target in antiretroviral chemotherapy. *MS, Méd. Sci.* **1993**, *9*, 1342−1351.

(11) Gregoire, J. C.; Loret, E. P. Conformational heterogeneity in two regions of Tat results in structural variations of this protein as a function of HIV-1 isolates. *J. Biol. Chem.* **1996**, *271*, 22641−22646.

(12) Bayer, P.; Kraft, M.; Ejchart, A.; Westendorp, M.; Frank, R.; Rosch, P. Structural studies of HIV-1 Tat protein. *J. Mol. Biol.* **1995**, *247*, 529−535.

(13) Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. Conformation of the TAR−RNA-arginine complex by NMR spectroscopy. *Science* **1992**, *257*, 76−80.

(14) Loret, E. P.; Georgel, P.; Johnson, W. C., Jr.; Ho, P. S. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9734−9738.

(15) Tao, J.; Frankel, A. D. Electrostatic interaction modulate the RNA-binding and trans-activation specificities of the HIV and SIV Tat proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90*, 1571−1575.

(16) Aboul-ela, F.; Karn, J.; Varani, G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* **1995**, *253*, 313−332.

(17) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. Role of RNA structure in arginine recognition of TAR RNA. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3680−3684.

(18) Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* **1995**, *270*, 1200−1203.

(19) Yao, S.; Wilson, W. D. A molecular mechanics investigation of RNA complexes. I. Ethidium intercalation in an HIV-1 TAR RNA sequence with an unpaired adenosine. *J. Biomol. Struct. Dyn.* **1992**, *10*, 367−387.

(20) White, S. A.; Draper, D. E. Single base bulges in small RNA hairpins enhance binding and promote an allosteric transition. *Nucleic Acid Res.* **1987**, 4049−4064.

(21) White, S. A.; Draper, D. E. Effects of single-base bulges on intercalator binding to small RNA and DNA hairpins and a ribosomal RNA fragment. *Biochemistry* **1989**, *28*, 1892−1897.

CI980094K