ELSEVIER

Interaction of new PNA-based molecules with TAR RNA of HIV-1: Molecular modelling and biological evaluation

R. Terreux,* S. Pairot,† D. Cabrol-Bass,* N. Patino,† R. Condom†

**Laboratoire Arômes, Synthèses et Interactions. Equipe Chimiométrie et Modélisation; and †Laboratoire de Chimie Bio-organique CNRS ESA 6001, Université de Nice Sophia-Antipolis, F 06108 Nice Cedex 2, France*

During the HIV-1 replication process, interactions between the RNA sequence, named TAR RNA, and the viral protein, Tat, permit a fast and efficient transcription of viral DNA into RNA. Based on the NMR structure of TAR RNA from the PDB, two Peptidic Nucleic Analog- (PNA) based molecules were designed by molecular modelling, the first one targeting G_{32} U_{31} and the second targeting U_{31} C_{30} free loop bases. Before designing the molecules, the flexibility of the TAR RNA was evaluated by molecular dynamics (MD). The molecules studied are composed of three domains: an arginine, a linker, and two PNA bases. First, molecules were designed and the linker length was optimized to fit the TAR RNA; second, a MD simulation on the TAR RNA molecule complex was performed to validate the molecular structure. Optimal molecules were synthesized and tested on infected cells. The experimental results support the choices made in the design of the molecules. © *2001 by Elsevier Science Inc.*

INTRODUCTION

The mortality due to AIDS has dramatically decreased with polytherapy, an important advance in the fight against the Human Imunodeficiency Virus (HIV). Two, three, or more drugs are used simultaneously against two viral enzymes: reverse transcriptase and protease. Nevertheless, nucleoside analogues or nonnucleosidic inhibitors and antiproteases are responsible for mutations of these two enzymes, which become resistant to drugs.[1] The development of this viral resistance poses a major threat for drug efficacy. Thus, new compounds against other HIV targets are needed. Transcription of HIV viral DNA in infected cells proceeds through a complex mechanism that includes several specific viral protein–RNA interactions. The viral RNA genome contains a long terminal repeat domain (LTR) within which lies the transactivation response element (TAR) that is recognized by an HIV regulatory protein, Tat. The interaction of Tat with TAR triggers a transactivation process, which involves many host cellular proteins, and leads to efficient production of full-length viral transcripts.[2] This interaction between TAR RNA and Tat is essential for virus replication. Therefore, inhibiting the Tat–TAR complex formation by targeting the TAR RNA seems to be a promising target for inhibiting virus multiplication. We have designed and synthesized new conjugates of L-arginine with Polyamide Nucleic Acids (PNA) as potential inhibitors of the interaction between the Tat protein, the viral and cellular proteins, and the TAR RNA. Interaction of the proposed new molecules with TAR RNA was investigated by molecular modelling. Two compounds of the series show micromolar anti-HIV-1 activities and present no cellular toxicity up to 200 micromolar.

STRUCTURES AND OBJECTIVES

The TAR RNA element of the HIV-1 (nt +1 to +59) is present at the 5′ end of every HIV transcript.[3] TAR RNA forms a stable hairpin structure with six apical nucleotides (CUGGGA) and two stem regions, upper and lower, separated by a bulge made of three unpaired nucleotides (UCU). The three-nucleotide bulge induces a kink of 91° between the lower and upper stems. The pocket created by this kink forms the Tat interaction site. Indeed, Tat binds in this region of the three-base bulge and recognizes both the

Color Plates for this article are on pages 614–615.

Corresponding author: D. Cabrol-Bass, Laboratoire Arômes, Synthèses et Interactions. Equipe Chimiomêtrie et Modélisation, Université de Nice Sophia-Antipolis, F 06108 Nice Cedex 2, France.

E-mail address: cabrol@unice.fr (D. Cabrol-Bass)

Journal of Molecular Graphics and Modelling 19, 579–585, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(01)00095-X

identity of adjacent Watson-Crick base pairs and the positions of surrounding phosphate groups. Aboul-ela and colleagues[4] have studied a 27-nucleotide sequence of TAR RNA by 2D NMR, including the loop made of six nucleotides. The sequence of bases from the base C_{18} to the base G_{44} comprises two duplex regions. This part of the TAR RNA interacts with the Tat protein.[5] A view of the 3D representation of this sequence following Aboul-ela[5] with the schematic representation of the sequence[5] is presented in Color Plate 1.

HIV-1 Tat[6] is a protein with 86 amino-acid residues, which can be subdivided into six regions. Of these, a basic region that is particularly rich in arginine is responsible for specific binding to TAR RNA. Indeed, upon binding to TAR RNA, the basic region of Tat promotes a conformational rearrangement in TAR, which places the functional groups recognized by the protein in a specific spatial arrangement. It has been shown that arginine 52 is largely responsible for the binding specificity and thus for the conformational change in TAR RNA.[7,8] Nonetheless, a single arginine residue (or arginine amide) binds specifically to TAR and induces a change in RNA conformation that largely mimics the conformation of a portion of Tat peptides-TAR complexes.[4,9–11] NMR studies show that ligands are bound in the major groove of the RNA. Indeed, the RNA bulge structure allows ligands to access the major groove of TAR, which induces folding in the bulge and formation of unusual base-triplets.

The observation that a single arginine amide induces similar conformational changes in Tat or Tat peptides suggests that one key arginine residue (of the basic region of Tat) might be sufficient to trigger the conformational change in TAR. The six nucleotides of the apical loop of TAR are implicated in the binding of cellular proteins. From these two observations, the design of new molecules was made based on the hypothesis that, in the native TAR RNA, this loop could be free of interactions with cellular proteins and Tat. Consequently, the six nucleotides of the loop could form Watson-Crick base pairs with complementary nucleotides or analogues, and an arginine residue could guide the molecule to the bulge target, inhibiting the Tat–TAR complex formation.

New molecules designed following this idea are made of dinucleotide analogues (PNA) linked to an arginine residue. In the PNA structure, the sugar backbone is replaced by a repeating N-(2-aminoethyl)glycine unit with a nucleobase linked through a methylene carbonyl linker.[12,13]

We have chosen PNA analogues of nucleotides for three main reasons: (1) the affinity of PNA for RNA is higher than that of oligonucleotide homologues; (2) the absence of negative charges on PNA increases the lipophilicity, which could permit better cellular penetration; and (3) PNA is stable toward nucleases, proteases, and peptidases.

These new molecules are composed of four parts (Figure 1):

- An arginine residue that should play an important role in guiding the molecule to its interaction site
- A PNA dimer, whose residues (AG or CA) are complementary to the loop nucleotides (UC or GU, respectively)
- A benzyloxycarbonyl group (Z) that has been introduced at the terminal amine group of the PNA dimer both to further the formation of π interactions with the nucleic bases of TAR RNA and to increase the lipophilicity of the molecules and facilitate their cellular penetration
- A linker that connects the arginine residue to the PNA dimer.

MOLECULAR MODELLING

The design of molecules was carried out in three separate phases. First, the arginine and the PNA domains were positioned in interaction with the TAR RNA. Second, the linker was designed and geometrically optimized; in the last phase, a molecular dynamics study was performed to validate the whole design. The loop flexibility is different than the flexibility of the double helix domain. An analysis of the nucleotides positions in the loop domain in the twenty RMN-2D structures[4] shows that the bases could change faces and the loop flexibility in such a way that bases can easily change orientation to accommodate favorable interactions with complementary bases. This conclusion was supported by a molecular dynamic study, not reported here, performed as a preliminary investigation to this work.

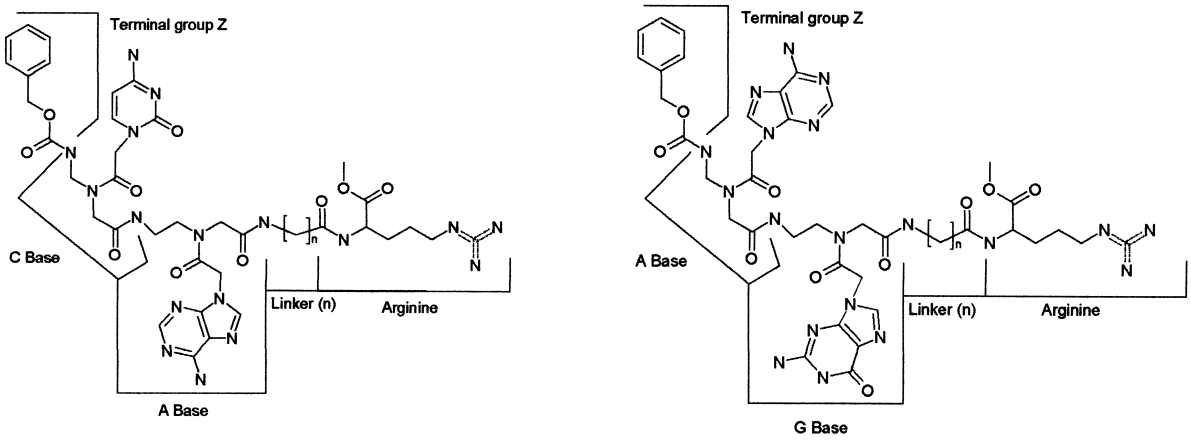

Figure 1. Structural formulas of molecules ZCA-n and ZAG-n.

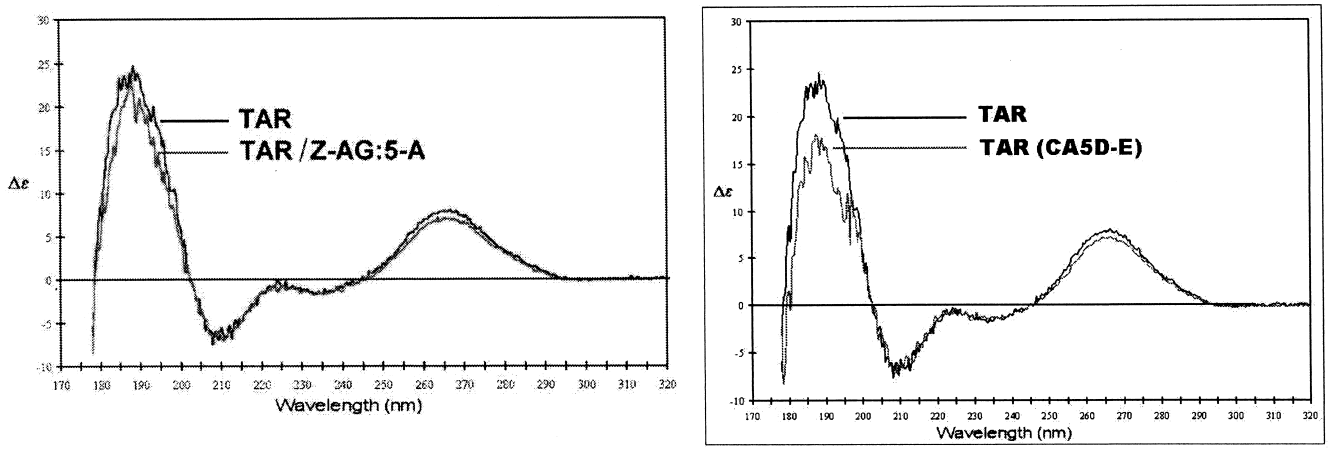

Figure 2. VUV-CD spectra of TAR (C45-G17) and TAR / ZAG-5 complex (left) and TAR (C45-G17) and TAR / ZCA-5 complex (right).

Interaction between TAR RNA and Molecule ZCA

The molecule has been specifically designed to interact with TAR RNA. On the one hand, the arginine interacts with the bulge bases. On the other hand, the two PNA analogue bases interact with the G_{32} and U_{31} bases of the free loop. The benzyloxy group (Z) was selected as an N-terminal group of the molecule. In previous work[17] we investigated the role of arginine to lead the molecule toward the TAR RNA bulge region. In the present study we are dealing with the interactions between the PNA domain, the linker molecule, and the TAR RNA.

To locate roughly the PNA domain facing the TAR RNA stem loop, a molecular model was constructed by replacing the linker and the arginine group by an ethyl group. This model molecule (ZCA-Et) was manually located in front of the corresponding bases, and distance constraints between heavy atoms involved in Watson-Crick hydrogen bonds were set as follows. A distance constraint equal to 50 kcal if the distance was longer than the equilibrium distances. Between the bases C and G, three distance constraints were set with an equilibrium distance parameter fixed to: 2.72 Å from O6 of the C bases to N4 of the G bases, and from N1 of the C bases to O2 of the G bases, and to 2.91 Å from N1 of the C bases to N3 of the G bases. Between bases A and U, two distance constraints were set with an equilibrium distance parameter fixed to: 2.83 Å from atoms N6 of the A bases to the O4 atoms of the U bases, and 2.91 Å from N1 of the C bases to N3 of the G bases.

Energy minimization was performed using the following protocol. First, 500 steps of a steepest descent were done, and then conjugated gradient minimization was conducted until the convergence criteria reached 0.01 kcal; finally, the BFGS algorithm,[18] which is a variant of the Newton-Raphson algorithm, was used with a convergence criteria defined to 0.001 kcal. During energy minimization, the Z group moves spontaneously between the two PNA bases and contributes to a stacking interaction with the bases π system. This interaction induces an angle of 27° between the two PNA bases, which moves them away. This structural change yields an optimized interaction between the two PNA bases and their complementary bases on the TAR RNA. The angle induced by the Z group in front of the TAR RNA sequence is shown in Color Plate 2.

To investigate the stability of the π stacking of the two PNA groups and the Z terminal group, A 200-ps dynamic simulation at constant pressure and temperature was performed using the Insight II/Discover molecular modelling package.[19] The temperature was set constant at 300 K, and the pressure at 1 atm. The Verlet-leapfrog algorithm was used with a time step of 1 fs. The Nosé-Hoover thermostat was used. The TAR RNA structure was cut at the G_{26}-C_{39} level, and only the upper part of the structure was kept to perform the simulation. The coordinates of the three last base pairs were frozen to simulate the continuation of the RNA sequence. The PNA domain of the molecule was located in the favorable stacking interaction as determined by energy minimization. A water box with 1,885 water molecules was created using the Insight II / Soak command. The previous additional distance constraints were added between heavy atoms of the PNA bases involved in a H bonding interaction with the bases of the free loop. Analysis of the dynamic trajectory shows that the π stacking of the Z group is maintained during the whole simulation and contributes by stabilizing the interaction of the PNA bases with the free loop TAR RNA bases.

To link the arginine residue, which is located in its interaction site, to the PNA domain of the molecule, in an interaction with the bases of the free loop, the model shows that it is necessary to insert the linker into the minor groove. Several linkers of different lengths where considered. Simple geometrical considerations led to the hypothesis that a chain made of five methylene groups connecting the arginine residue by a carbonyl group, and the PNA by an amino group could be optimal. To test this hypothesis, models with x = 4, 5, and 6 methylene groups were studied.

The interaction energy between the three studied ZCA-x molecules and TAR RNA was calculated with the Hager force field CFF93. The interaction energy of the three domains of the molecule are reported in Table 1.

The results confirm that the optimal linker should be made of 5 methylene groups. In comparison to ZCA-5, the loss in interaction energy is 9.09 kcal/mol^{-1} for ZCA-4 and 13.43

Table 1. Interaction energy between ZCA-x to TAR RNA

Molecule	Energy Terms	Molecule ZCA-x/TAR	PNA domain/TAR	Linker/TAR	Arginine/TAR
ZCA-4	Van der Waals energy	−36.03	−20.19	−3.21	−12.63
	Coulomb energy	−125.49	−90.72	−9.26	−25.51
	Total Energy	−161.52	−110.91	−12.47	−38.14
ZCA-5	Van der Waals energy	−41.75	−21.19	−7.70	−12.86
	Coulomb energy	−128.86	−90.02	−13.23	−25.61
	Total Energy	−170.61	−111.21	−20.93	−38.47
ZCA-6	Van der Waals energy	−34.84	−20.19	−1.78	−12.87
	Coulomb energy	−122.34	−90.13	−6.45	−25.76
	Total Energy	−157.18	−110.32	−8.23	−38.63

kcal/mol^{-1} for the ZCA-6 molecules. An analysis of Table 1 shows that this loss in interaction energy results entirely from a poorer adaptation of the linker domain in the minor grove of TAR RNA. The linker with 4 methylene groups is too short and its linear extended conformation is unfavorable. The linker with 6 methylene groups is too long and methylene groups enter in steric interaction with the sugar backbone of TAR RNA. For the optimal molecule, an analysis of the energetic terms shows that the major contribution to the interaction energy is provided by the Coulombic interaction terms. Furthermore, the major interaction contribution is supported by the PNA domain, which interacts with the TAR RNA by five different hydrogen bonds. This interaction increases the electrostatic energy interaction term. The arginine residue interacts by an average contribution of approximately 38 kcal, with a major electrostatic contribution due to the two hydrogen bonds from the guanidinium group to the G_{26} C_{39} TAR RNA. The linker domain poorly contributes to the complex stabilization because it is made of five hydrophobic methylene groups and the minor groove is covered by charged phosphate groups. The optimum final structure of the complex ZCA-5 / TAR RNA is shown in Color Plate 3.

To evaluate the stability of the TAR RNA / ZCA-5 complex, a molecular dynamic simulation of 200 ps at constant temperature and pressure was performed. The same set of parameters used for the study of ZCA-Et model molecule was retained. During this dynamic, the atoms coordinates of TAR RNA were frozen and the dielectric constant was set to 80 with a distance factor to simulate the solvent. Analysis of the dynamic trajectory supports the hypothesis that the ZCA-5 molecule keeps in close interaction with the structure of the TAR RNA.

Interaction between TAR RNA and Molecule ZAG-5

The ZAG molecule has been designed on the basis of the same hypothesis stated for the ZCA molecule, but the target of the PNA domain is the U_{31} C_{30} base of the free loop instead of G_{32} U_{31}.

Additional constraints were set between heavy atoms involved in Watson - Crick hydrogen bonds, between the PNA analog bases, and the corresponding bases in the free loop of TAR RNA. Constraints similar to those for the ZCA molecule were used. The PNA domain was positioned in front of the two free bases, with energetic minimization and a molecular dynamics study performed after the same procedure used for the ZCA molecule. The Z group remains on the side of the molecule, making a π stack with the adjacent A base. This result contrasts with the case of the ZCA molecule in which the Z group goes between the two PNA bases. Considering the position and orientation of the ZAG molecule with respect to TAR RNA, it is impossible to conceive of a linker with the arginine residue located along the minor groove, as it is for the ZCA molecule. Close examination of the 3D model reveals that it might be possible to connect the arginine residue to the C-terminal end of the PNA domain by a linker running along the upper face of the major groove. Surprisingly, although the hypothetical location of the linkers in the two cases are completely different for ZCA and ZAG, their optimal length has been found to be the same, i.e., five methylene groups separating a carbonyl and an amino group.

The same strategy as for ZCA-x molecule was used. Several models with different numbers of methylene groups were built to evaluate the gain and confirm the optimal length of the linker. The interaction energy was evaluated between each domain of the ZAG-5 molecule and the TAR RNA, with the CFF93 force field after minimization. The results are reported in Table 2.

An analysis of this table shows that the minimal interaction energy is obtained for the ZAG-5 molecule. Compared with ZAG-5, the loss in energy is 9.44 kcal.mol^{-1} for ZAG-4 and 12.28 kcal.mol^{-1} for ZAG-6. Analysis of interaction energy terms shows that the major interaction contribution is supported by the electrostatic term, as for the previous molecule. Besides the hydrogen bonding network of the linker region, the ZAG-5 molecule interacts with at least seven other hydrogen bonds with TAR RNA: five supported by the PNA domain and two from the arginine residue. The interaction energy of the arginine domain is roughly equal for ZAG-5 and ZCA-5 because the arginine position is the same in both cases. The difference between the interaction energy of the ZCA-5 and the ZAG-5 PNA domain may be explained by the different location of the Z group that contributes to the ZAG-5 / TAR RNA complex stabilization of ZCA-5. The final structure of the complex is shown in Color Plate 4.

For each ZAG-x model, we performed a constant energy molecular dynamic simulation of 200 ps, using the same parameters as for the ZCA molecule. Globally, the simulation shows that the two domains of the molecule maintain close

Table 2. Interaction energy between ZAG-x to TAR RNA

Molecule	Energy Term	Molecule ZAG-x/TAR	PNA domain/TAR	Linker/TAR	Arginine/TAR
ZAG-4	Van der Waals energy	−30,44	−15,21	−2,21	−13,02
	Coulomb energy	−111,98	−80,23	−6,30	−25,85
	Total Energy	−142,42	−95,44	−8,51	−38,47
ZAG-5	Van der Waals energy	−35.56	−17.06	−12.75	−13.28
	Coulomb energy	−116.29	−78.03	−5.80	−24.92
	Total Energy	−151.86	−95.09	−18.56	−38.20
ZAG-6	Van der Waals energy	−28,14	−14,23	−1,02	−12,89
	Coulomb energy	−111,44	−80,78	−5,34	−25,32
	Total Energy	−139,58	−95,01	−6,36	−38,21

interaction with their respective targets; however, analysis of the trajectory highlights several specific interactions:

- The amino group of the G_{33} loop base participates in a hydrogen bond with the G base of the PNA domain. This hydrogen bond contributes to the stabilization of the interaction between the molecule and the TAR RNA.
- The C_{30} TAR RNA base makes a π stack with the C_{29} G_{36} base pair. This stacking creates an extension of the duplex domain and contributes to the stabilization of the complex ZAG-5 / TAR RNA.
- A network of hydrogen bonds is formed, involving:
- (i) the amide group located at the beginning of the linker; (ii) the acid group of the arginine residue; (iii) the phosphate group of the A_{35} base; and (iv) the NH_2 of the G_{34} base. This network of hydrogen bonds stabilizes the linker domain of the molecule.

This interaction seems critical because the linker occupies the Tat protein interaction site.

CIRCULAR DICHROISM AND BIOLOGICAL RESULTS

The strategy for the preparation of the bi-PNA parts was discussed in a previous article.[20] The synthesis of the specific compounds considered in this study will be described elsewhere.

Circular Dichroism Measurements

The TAR RNA (C45-G17) was synthesized on a 1-μmol scale using controlled pore glass (CPG) by standard phosphoramidite (obtained from Glen Research) synthesis[21] using a Perkin Elmer ABI 392 DNA/RNA synthesizer (Perkin Elmer Instruments, Shelton, CT USA).

TAR and TAR-PNA molecule complexes for vacuum UV circular dichroism (VUV-CD) measurements were prepared in 40 mM sodium phosphate buffer (pH 7). Molar ratios for the TAR-PNA molecule complex preparations were 1:3 (TAR: PNA molecule). CD spectra were measured in 50μm pathlength cells from 320 nm to 178 nm with a JOBIN-YVON VUV-CD spectrophotometer (MARK VI Long-Jumeau, France). Measurements were performed at 5°C for TAR and TAR-PNA molecules complexes (5°C was required to stabilize the complexes). CD spectra are reported as the difference in extinction coefficient, $\Delta\epsilon$, per nucleotide or per amide (the TAR extinction coefficient is 229500 M^{-1} cm^{-1} at 260 nm per TAR molecule). The VUV-CD data were analyzed for secondary structure by the method of Manavalan and Johnson.[22] The CD spectra of TAR (C45-G17) and TAR-PNA molecule complexes are presented in Figure 2.

The VUV-CD spectra of double-stranded RNA are highly sensitive to the conformation of the polynucleotide helix. The spectrum of TAR RNA (C45-G17) alone is typical for RNA and is usually associated with the A-conformation of a polynucleotide helix, especially the 210nm band. The CD spectrum of the complex between TAR and ZAG-5 or ZCA-5 is, in these two cases, dominated by the RNA structure. The 265 nm band of the RNA is sensitive to base stacking. The decrease in $\Delta\epsilon$ upon forming the complex can be interpreted as a modification of the base stacking induced by the binding of ZAG-5 and ZCA-5. Moreover, a decrease in intensity for the 265nm band of RNA is always accompanied by a decrease in intensity for the shorter wavelength bands, here there is a decrease in intensity for the 190nm band. CD spectra indicates that both ZAG-5 and ZCA-5 PNA molecules do interact with TAR.

In Vitro Biological Evaluation

The anti-HIV-1 activity of ZAG-5, ZCA-5, PNA dimers (ZAGOH and ZCAOMe), and arginine was evaluated in infected PBMC, CEM-SS, and MT4 cells (Table 3).

Both ZAG-5 and ZCA-5 were active against HIV-1 at micromolar concentrations on cultures of PBMC and CEM-SS cells and without apparent cellular toxicity. On cultures of MT4 cells, the anti-HIV-1 activity of ZAG-5 was diminished. PNA dimers and arginine were devoid of anti-HIV-1 activity at concentrations up to 200μM in infected PBMC, CEM-SS, and MT4 cell lines. These results suggested that the activity is due to the entire structures of ZAG-5 or ZCA-5.

Binding assays are currently in progress to analyze the ligand capability to compete with the Tat protein for TAR RNA. The results of these assays and a more complete analysis of the antiviral activity of these molecules will be described in a future article in a specialized journal.

TECHNICAL DETAIL

The molecular dynamics of the whole structure of the TAR RNA was computed with the AMBER 5.0 molecular modelling package. The calculation of the molecular dynamics simulation

Table 3. Antiviral activities of molecules on cells

Evaluated molecules	PBMC		CEM-SS		MT4	
	IC50(μM)	CC50(μM)	IC50(μM)	CC50(μM)	IC50(μM)	CC50(μM)
ZAG-5	1.0	>200	1.4	>200	19	>200
ZCA-5	3.5	>200	3.8	>200	>200	>200
ZAGOH	>200	>200	>200	>200	>200	>200
ZCAOMe	>200	>200	>200	>200	>200	>200
H-Arg-OH	>200	>200	>200	>200	>200	>200

was performed on the "Centre Informatique National de l'Enseignement Supérieur (CINES)" 207 node IBM SP2 supercomputer. The short dynamics used for the drug design were computed on the SGI Indigo 2 R10.000 of the Centre d'Imagerie et de Modélisation Moléculaire (CIMM) of the university of Nice-Sophia Antipolis, using the Insight II molecular modelling package. The calculations were computed using Discover with the Hager CFF force-field.

CONCLUSION

The antiviral activities of the two molecules are roughly equal to the two interaction energies that were computed. Our results led to insightful remarks concerning the interaction site. Unlike proteins, which have one or a few interaction sites, RNA has many potential interaction sites. It is feasible that the molecules studied might be toxic because there are a large number of potential interaction sites on RNAs, but our experiments show this reasoning to be wrong (CC 50 $>$ 200 μM). The toxicity 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–PNA is rather specific to the TAR RNA. This observation supports the hypothesis that the arginine plays an important role in "piloting" the molecule to its interaction site. The TAR RNA has the same selectivity to these molecules as the proteases or reverse transcriptase proteins that are the main targets of drugs against AIDS. The results of biological tests validate the process of the RNA structure simulation upon which our interaction model is partly based. 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 that deserves further investigation along with the protease and reverse transcriptase strategies, for developing new therapeutic scheme against HIV-1.

ACKNOWLEDGEMENTS

We are grateful to Dr E. Loret for CD analysis and Dr A.M. Aubertin for biological evaluation of the molecules. This work was supported by a grant from Agence Nationale de Recherches sur le SIDA (ANRS).

REFERENCES

1 De-Clereq, E. Development of resistance of human immunodifiency virus (HIV) to anti-HIV agent: How to prevent the problem? *Internat. J. Antimicrob. Agents*. 1997, **9**, 21–36
2 Mareiniak, R.A., and Sharp, P.A. HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J.* 1991, **10**, 4188–4196
3 Rosen, C.A. Regulation of the HIV gene expression by RNA–protein interactions. *Trends Genet.* 1991, **7**, 7–14
4 Aboul-ela, F., Karn, J., and 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
5 Churcher, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Butler, P.J.G., Gait, M.J., and Karn, J. High affinity binding of TAR RNA by the human immunodeficiency virus type-1 Tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* 1993, **230**, 90–110
6 Culen, B.R. Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol. Rev.* 1992, **56**, 375–394
7 Tao, J., and Frankel, A.D. Specific binding of arginine to TAR RNA. *Proc. Natl. Acad. Sci. USA.* 1992, **89**, 2723–2726
8 Tao, J., Chen, L., and Frankel, A.D. Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction. *Biochemistry* 1997, **36**, 3491–3495
9 Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D., and Williamson, J.R. Conformation of the TAR-RNA-arginine complex by NMR spectroscopy. *Science* 1992, **257**, 76–80
10 Puglisi, J.D., Chen, L., Frankel, A.D., and Williamson, J.R. Role of RNA structure in arginine recognition of TAR RNA. *Proc. Natl. Acad. Sci. USA.* 1993, **90**, 3680–3684
11 Tan, R., and Frankel, A.D. Circular dichroism studies suggest that TAR RNA changes conformation upon specific binding of arginine or guanidine. *Biochemistry* 1992, **31**, 10288–10294
12 Nielsen, P.E., Egholun, M., Berg, R.H., and Buehardt, O. Sequence selective recognition by atomes displacement with a thymine substituted polyamide. *Science* 1991, **254**, 1497–1500
13 Egholen, M., Buchardt, O., Nielsen, P.E., and Berg, R.H. Peptide Nucleic Acids (PNA) oligonucleotide analogues with an achiral peptidic backbone. *J. Am. Chem. Soc.* 1992, **114**, 1895–1897
14 Cornell, W.D., Cieplack, P., Bayly, C.I., Gould, I.R.,

Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 1995, **117**, 5179–5197

15 van Gunsteren, W.F., and Berendsen, H.J.C. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 1977, **34**,1311–1327

16 Berendsen, H.J.C., Posma, J.P.M., Gunsteren, W.F., Dinola, A., and Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984, **81**, 3684–3690

17 Terreux, R., Cabrol-Bass, D., Peytou, V., Condom, R., and Guedj, R. Modeling of the interaction between new ethidium derivatives and TAR RNA of HIV-1. *J. Chem. Inf. Comput. Sci.* 1999, **39**, 413–419

18 Molecular Simulation Inc, San Diego, USA

19 Fetcher, R. Practical methods for optimization Vol 1, *Unconstrained Optimization;* New York: John Willey & Sons, 1980.

20 Farese, A., Pairot, S., Patino, N., Ravily, V., Condom, R., Aumelas, A., and Guedj, R. Liquid phase synthesis of peptide nucleic acid (or polyamide nucleic acid) dimers. *Nucleosides & Nucleotides,* 1997, **16**, (10&11), 1893–1906.

21 Wincott, F., Direnzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucl. Acids Res.* 1995, **23**, 2677–2684.

22 Manavalan, P., and Johnson, W.R. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* 1987, **167**, 76–85.