6278 *Biochemistry* **2009,** *48,* 6278–6284
DOI: 10.1021/bi802373d

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

In Vitro Selection of RNA Aptamers Derived from a Genomic Human Library against the TAR RNA Element of HIV-1

Marguerite Watrin,[‡,§] Frederike Von Pelchrzim,[∥] Eric Dausse,[‡,§] Renée Schroeder,[∥] and Jean-Jacques Toulmé*[,‡,§]

[‡]*Inserm U869, European Institute of Chemistry and Biology, Pessac, France,* [§]*University of Bordeaux, Bordeaux, France, and* [∥]*Max F. Perutz Laboratories, University of Vienna, Vienna, Austria*

Received December 31, 2008; Revised Manuscript Received May 29, 2009

ABSTRACT: The transactivating responsive (TAR) element is a RNA hairpin located in the 5′ untranslated region of HIV-1 mRNA. It is essential for full-length transcription of the retroviral genome and therefore for HIV-1 replication. Hairpin aptamers that generate highly stable and specific complexes with TAR were previously identified, thus decreasing the level of TAR-dependent expression in cultured cells [Kolb, G., et al. (2006) *RNA Biol. 3*, 150−156]. We performed genomic SELEX against TAR using a human RNA library to identify human transcripts that might interact with the retroviral genome through loop−loop interactions and potentially contribute to the regulation of TAR-mediated processes. We identified a genomic aptamer termed a1 that folds as a hairpin with an apical loop complementary to five nucleotides of the TAR hexanucleotide loop. Surface plasmon resonance experiments performed on a truncated or mutated version of the a1 aptamer, in the presence of the Rop protein of *Escherichia coli*, indicate the formation of a highly stable a1−TAR kissing complex. The 5′ ACCCAG loop of a1 constitutes a new motif of interaction with the TAR loop.

In prokaryotes, loop−loop interactions between RNA hairpins, so-called kissing complexes, participate in the regulation of numerous cellular processes (*1*−*3*). As an example, the replication of the ColE1 plasmid in *Escherichia coli* is regulated by the formation of a natural kissing complex between two RNA hairpins displaying complementary loops (*4*−*7*). A limited number of examples are available in eukaryotes: it has been demonstrated that the dimerization of the HIV-1[1] RNA is also initiated by the formation of an RNA loop−loop complex involving the DIS (dimerization initiation site) hairpin loop structure (*8*−*10*).

Numerous aptamers were generated against RNA structures such as the yeast $tRNA^{Phe}$, the dimerization initiation site of HIV-1 mRNA, or regions of the hepatitis C virus mRNA (*11*−*16*). All these aptamers adopt hairpin structures that bind to their cognate target hairpin through loop−loop interactions. Kissing aptamers such as RNA aptamer R06 were also obtained against the transactivation responsive (TAR) RNA element of HIV-1 (*17*−*20*), an imperfect 59-nucleotide RNA haipin located at the 5′ end of the 5′ untranslated leader region of HIV-1 mRNA. Upon binding to the viral protein Tat and to host proteins, TAR allows the efficient full-length transcription of the retroviral genome and is therefore essential for HIV-1 replication (*21*−*27*). Anti-TAR RNA aptamer R06 was shown to reduce the level of TAR-dependent expression of reporter genes in cultured cells (*20*).

*To whom correspondence should be addressed: Inserm U869, European Institute of Chemistry and Biology, University of Bordeaux, 2 rue Robert Escarpit, 33607 Pessac, France. Telephone: 33 (0)5 5757 1014. Fax: 33 (0)5 5757 1015. E-mail: jean-jacques.toulme@inserm.fr.

[1]Abbreviations: TAR, transactivating responsive; HIV, human immunodeficiency virus; SPR, surface plasmon resonance.

It was previously reported that the HIV genome contains a hairpin named TAR* susceptible to giving rise to kissing interaction with TAR (*28*−*30*), but no function was described for this viral RNA element. As kissing interactions result in strong RNA−RNA association, we wondered whether transcript (s) from the host cell could be identified that might generate stable complexes with the TAR RNA through loop−loop interaction, thus constituting potential regulatory element(s) for the viral infection. To address this question, we conducted a genomic SELEX experiment (*31*, *32*). Genomic SELEX allows the identification of RNA fragments that might reflect natural interactions with the chosen target, very generally a protein. We extended this approach to the identification of RNA−RNA interactions using the TAR hairpin as a target. Indeed, genomic SELEX previously enabled the discovery of novel noncoding RNAs (*33*). We identified genomic RNA aptamers able to interact with the TAR RNA element through loop−loop interactions. We characterized an aptamer, forming a stable kissing complex with TAR and containing a motif that was not fished out in previous SELEX experiments conducted with synthetic RNA libraries (*18*, *19*).

MATERIALS AND METHODS

Library and Oligonucleotide Synthesis. The starting genomic DNA library has been generated from double-stranded fragments of genomic DNA from human placenta. Genomic DNA was incubated with random primers containing nine-nucleotide random sequence and a fixed forward primer (5′ AGGGGAATTCGGAGCGGGGCAGCN$_9$) or a fixed reverse primer (5′ CGGGATCCTCGGGGCTGGGATGN$_9$). After a

 Published on Web 06/04/2009

Klenow extension step, PCR amplification and size fractionation were conducted. Genomic sequences have been flanked by 5′ and 3′ additional primer regions to enable the PCR amplification of the library, the T7 transcription reaction, the reverse transcription, and the PCR amplification of selected sequences. The two oligonucleotides used, Fixfor containing the T7 transcription promotor (underlined) (5′ CCAAGTAATACGACTCACTATAGGGGAATTCGGAGCGGG) and Fixrev (5′ CGGGATCCTCGGGGCTG), have been synthetized by Eurogentec. Variants of RNA aptamers were synthesized in our laboratory on an Expedite 8909 synthesizer (Applied Biosystems) as well as the TAR sequence 5′ CGCCAGAUUUGAGCCUGGGAGCUCUCUGGCG. All oligonucleotides were purified by electrophoresis on denaturing 20% polyacrylamide, 7 M urea gels.

The TAR RNA stem−loop structure used in this study was restricted to nucleotides C18−G44 of the HIV-1 TAR element. The hairpin stem was lengthened by two GC pairs at the bottom and was biotinylated at the 3′ end.

Transcription of the Genomic Library. The DNA genomic library (6 μg) was transcribed at 37 °C for 2 h in a final volume of 40 μL using the Ampliscribe T7 high-yield transcription kit (Epicentre Technologie). DNase I treatment was performed (2 μL at 1 unit/μL) at 37 °C for 30 min. Transcription products were purified by phenol (pH 4.3)/chloroform extraction and ethanol precipitated. The resuspended pellet was purified on Sephadex G-25 Fine (GE Healthcare). The RNA amount was quantified by UV absorption at 260 nm. Genomic RNA candidates were then ready for the first round of selection in R buffer [20 mM HEPES, 20 mM sodium acetate, 140 mM potassium acetate, and 3 mM magnesium acetate (pH 7.4)].

Genomic SELEX Procedure. *(i) Counterselection*. To eliminate nonspecific binders, the genomic RNA pool was submitted to a counterselection step prior to each round of selection. The RNA pool was heated at 95 °C for 1 min in water, cooled to 4 °C for 1 min, and left at room temperature for 5 min in R buffer. Then, 250 pmol of the genomic RNA library (or decreasing amounts of the selected RNA pool in the successive rounds) was incubated in 100 μL of R buffer for 15 min at room temperature (23 °C) with 100 μg of streptavidin MagneSphere Paramagnetic beads (Promega) (or 1 mg of Dynabeads M-280) previously equilibrated in R buffer for 10 min. The supernatant was recovered for proceeding to the positive selection step. Bound RNA candidates were eluted from the beads and quantified by UV absorption at 260 nm.

(ii) Positive Selection. The TAR RNA target was heated at 65 °C for 3 min, chilled in ice for 1 min, and folded at room temperature for 5 min in R buffer. Then, 10 pmol of the target was mixed in a final volume of 100 μL of R buffer, with 100 μg of streptavidin magnetic beads from Promega (or 1 mg of Dynabeads M-280) at room temperature for 10 min. After the free target had been discarded, 250 pmol of the counterselected library was added to the beads and incubated at room temperature in a final volume of 100 μL of R buffer for 15 min. Nonselected candidates were removed, and the beads were washed once with 100 μL of R buffer. The target-bound candidates were eluted in 80 μL of water by being heated at 80 °C for 45 s. Selected RNAs were copied into cDNA with 240 units of M-MLV reverse transcriptase RNase H^- point mutant (Promega) for 50 min at 50 °C in a final volume of 20 μL. Then, PCR amplification was performed for 10 cycles in a final volume of 1 mL, using 40 units of Ampli Taq gold (Applied Biosystems) and Fixfor and Fixrev primers. After phenol (pH 8)/chloroform extraction and ethanol precipitation, PCR products were purified with nucleospin Extract II (Macherey-Nagel), transcribed, and submitted to a DNase I treatment as described for the genomic library. RNA candidates were then ready for the next round of selection. The stringency of the selection increased from one round to the other (see Table S1 of the Supporting Information).

Electrophoretic Mobility Assay on Nondenaturing Gels. The genomic RNA library and RNA pools from each round of selection have been treated as described previously in TBM buffer [89 mM Tris-borate (pH 8.3) and 5 mM $MgCl_2$] containing 5% glycerol. One microgram of each sample was loaded on a 12% (w/v), 75:1 acrylamide/bisacrylamide gel, in TBM buffer, and run at 15 W for 3−4 h (16.5 cm × 21.5 cm). The gel was stained with ethidium bromide (0.5 μg/mL) for 30 min and washed twice with water. The RNA species were analyzed by UV detection.

Surface Plasmon Resonance (SPR) Measurements. SPR experiments were performed with a Biacore 3000 apparatus as described previously (*34*, *35*). All experiments were conducted in R buffer. For the analysis of a1 RNA variants, 350 ru (resonance units) of biotinylated TAR (50 nM) was immobilized on an activated SA sensor ship coated with streptavidin. As a negative control, a nonrelevant RNA hairpin has been immobilized on another channel. Aptamer samples were prepared in R buffer as described previously, at different concentrations ranging from 500 nM to 4 μM; 80 μL of each sample was injected at a flow rate of 20 μL/min, at 23 °C. The dissociation was observed for 600 s. Between two injections, the target was regenerated with 20 mM EDTA, 100 mM EDTA, or 40% formamide. Then the target as well as the needle and the integrated fluidic cartridge were washed with R buffer. To evaluate the binding of Rop to the kissing complex, 20 μL of the aptamer (2 μM) was coinjected with 20 μL of Rop from 0.5 to 40 μM, at a rate of 20 μL/min, at 5 °C. The sensorgrams were analyzed with BIAeval version 2.2.4.

Cloning and Sequencing. An additional elongation step (10 min, 72 °C) was performed to enable the addition of a single deoxyadenosine at the 3′ end of the PCR products so they could be cloned into the TOPO TA cloning vector (Invitrogen). *E. coli* TOP10 One Shot cells were transformed according to the manufacturer's instructions. Clones from the sixth round were sequenced with the BigDye Terminator version 3.1 Cycle Sequencing Kit (PE Applied Biosystems). Sequences were analyzed via 4Peaks version 1.7.2 (4Peaks by A. Griekspoor and T. Groothuis, mekentosj.com)

Rop Expression. Rop mutagenesis was performed using the p2R plasmid containing the Rop sequence, kindly provided by B. Bishop and L. Regan, using the Quickchange TM XL site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. The Phe14 residue essential to the kissing RNA−RNA−Rop complex was substituted with an Ala residue that abolishes the interaction between Rop and loop−loop complexes (*36*). The purification of the mutated Rop protein was performed as described by Predki et al. (*36*) using BL21 (DE3) *E. coli* cells. The Rop protein with a six-histidine tag was expressed and purified as previously described by Di Primo and Lebars (*37*).

RESULTS

Genomic Selection against the TAR RNA Element. Genomic selection was performed using a RNA library generated by in vitro transcription of the whole human genome (exonic, intronic, and intergenic sequences), against a chemically synthesized RNA hairpin derived from the TAR element.

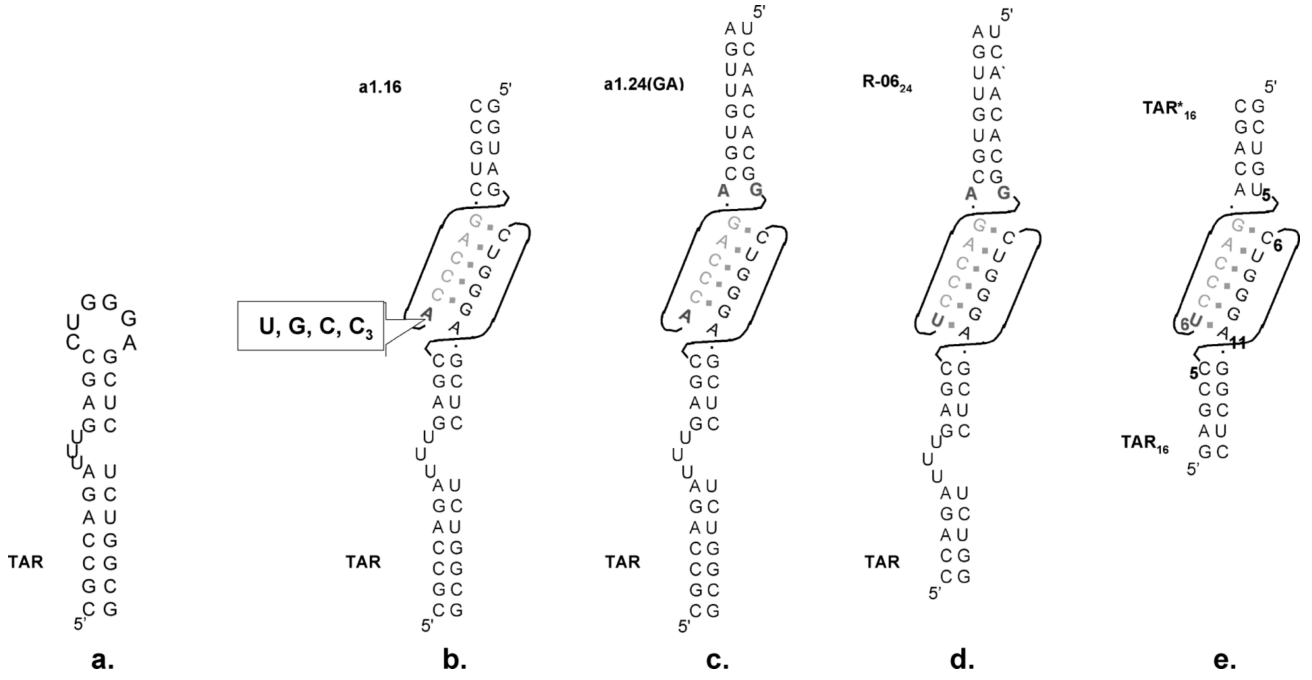

FIGURE 1: Secondary structure of the TAR RNA element of HIV-1 and of aptamer−TAR complexes. (a) Hairpin derivative corresponding to the top part of TAR, used for in vitro genomic selection. Two G·C pairs were added at the bottom of the stem. In addition, the 3′ end was biotinylated. (b−e) Kissing complexes between TAR and various hairpins. Mutations were introduced into truncated a1 genomic aptamers (b and c). C_3 (b) stands for a non-nucleotide spacer made of three methylene groups. Previously studied kissing anti-TAR hairpins R06 (d) and TAR* (*18*, *29*) are given for the sake of comparison. The consensus 5′ CCCAG sequence of the aptamer loop and critical positions are indicated. Potential Watson−Crick base pairs of the loop−loop helix are represented as gray dots.

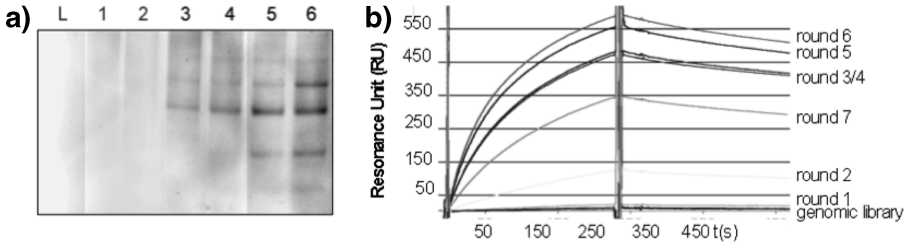

FIGURE 2: Evolution of the genomic selection against TAR. (a) Electrophoresis on a nondenaturing 12% polyacrylamide gel of the candidates contained in the genomic library (lane L) or in selected RNA pools (rounds 1−6 in lanes 1−6, respectively). (b) SPR analysis of complexes formed between biotinylated TAR (immobilized) and RNA sequences from either the library or the pools obtained after rounds 1−7 as indicated to the right of the panel. Samples (80 μL at 500 nM) were injected on the surface at a flow rate of 20 μL/min in R buffer (see Materials and Methods for details).

This truncated hairpin, 31 nucleotides in length, corresponding to the top part of the retroviral RNA structure retains the protein binding properties of the full-length element. It has been used throughout our study (Figure 1a). The stem is extended by two additional G-C pairs at the bottom for locking the hairpin in the folded conformation. In addition, the TAR derivative is biotinylated at the 3′ end for capture by streptavidin beads or immobilization on sensor chips.

The selection pressure has been adjusted all along the SELEX procedure, to select the genomic sequences that display the highest affinity for the TAR target. The concentration of the RNA candidates and of the RNA library has been decreased in every round. Moreover, the candidate:target ratio has been kept high to promote competition between candidates (Table S1 of the Supporting Information). Six selection rounds were performed in R buffer as described in Materials and Methods.

Specific and nonspecific interactions of candidates with beads either naked or bearing the TAR hairpin have been quantified by UV absorbance at 260 nm for each round. When the background was increasing, beads were exchanged (see Materials and Methods). Alternating different types of streptavidin beads allowed us to keep the background stable and low during the selection procedure. No significant binding was detected with the population issued from the first five rounds of selection, whereas limited interaction was observed in the sixth round (not shown). The evolution of the RNA pool was monitored by electrophoretic mobility on a nondenaturing polyacrylamide gel (Figure 2a). Rounds 1 and 2 showed the same migration pattern as the genomic library; i.e., no discrete band was observed. The presence of a smear is explained by the size of the fragments in the library, ranging from 100 to 800 nucleotides (*32*). From round 3, distinct bands were detected corresonding to over-represented species, indicating that the population evolved.

The affinity of the different RNA pools selected against TAR was measured by SPR (see Materials and Methods). Sensorgrams revealed a marked evolution of the binding of the candidates to the TAR hairpin immobilized on the chip. The magnitude of the signal increased progressively from round 2 to 6 and then decreased for the population of the seventh round (Figure 2b). Such an evolution is frequently observed and might be related to

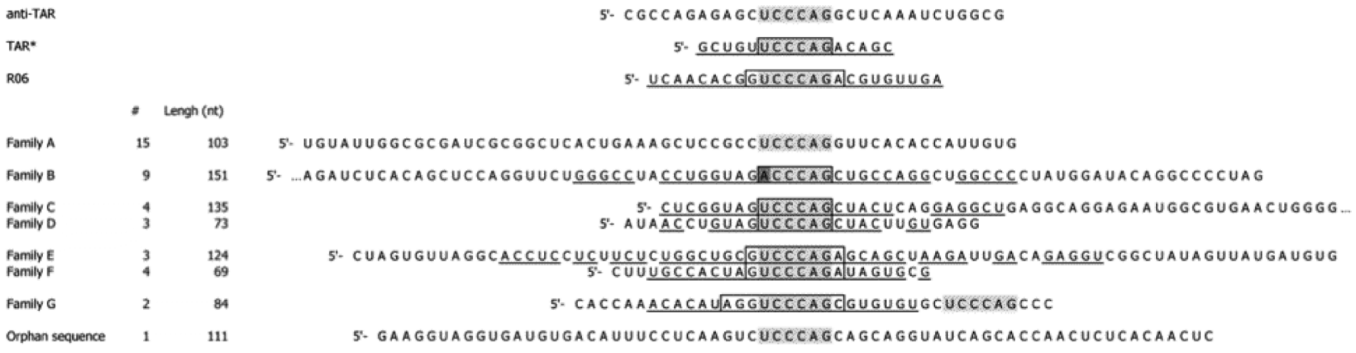

FIGURE 3: Genomic RNA sequences selected against TAR RNA. The number of identical sequences selected (#) and the lengh of the sequences are given at the left. The sequences fully complementary to TAR, TAR* (*29*, *30*), and R06 (*18*, *34*) are given at the top for direct comparaison with selected sequences. Structure prediction was conducted using Mfold. Predicted loops are boxed, whereas double-stranded stems are underlined. All sequences termed A−G as well as an orphan sequence displayed a 5′ CCCAG consensus motif complementary to part of the TAR apical loop sequence. The sequence complementary to the TAR loop is indicated (light gray). The A residue (dark gray) in sequence B is characteristic of a new kissing motif (see the text).

the ability of a given sequence to be amplified whatever its binding efficiency. No signal was observed on a channel on which a nonrelevant RNA hairpin was imobilized, indicating that binding to TAR was specific. The candidates of the sixth round were cloned and sequenced.

Identification of Selected Genomic RNAs against TAR. The sequences of 41 candidates have been compared. Whereas the starting genomic library consisted of sequences ranging from 100 to 800 nucleotides, selected sequences ranged from 69 to 151 nucleotides in length. The shortening of the selected candidates has been previously noticed and is likely due to the reverse transcription step (*32*). The sequence alignement shows that the selection has generated seven different sequences (A−G) and one orphan sequence. Selected genomic sequences are presented in Figure 3. Interestingly, all sequences displayed a 5′ CCCAG consensus complementary to five nucleotides of the TAR apical loop (5′ CUGGG). Moreover, all sequences except sequence A and the orphan sequence are predicted to fold into a hairpin structure, the loop of which corresponds to the sequence complementary to the TAR loop (*38*). The predicted secondary structures are given in Figure S1 of the Supporting Information.

Sequence A is the most represented (15 identical candidates, 103 nucleotides in length). The candidate includes a stretch of 12 nucleotides complementary to TAR (except the second U of the TAR bulge). Sequence B, 151 nucleotides in length, is represented nine times. Candidates C and D show very similar sequences of different lengths (73−135 nucleotides). All of them display a 5′ UCCCAG loop sequence fully complementary to the TAR one. This loop is identical to that of the rationally designed TAR* hairpin described previously (*29*, *30*). Candidates E and F were 124 and 69 nucleotides long, respectively. Both sequences show a 5′ GUCCCAGA loop sequence, which is characteristic of the previously identified R06 anti-TAR aptamer, i.e., a loop complementary to the TAR one framed by a G,A combination (*18*, *34*). Lastly, the two G sequences display two times the sequence complementary to the TAR loop (Figure 3). Therefore, the selection of candidates B−G might have been driven by the formation of kissing complexes with TAR, as demonstrated previously for other aptamers targeted to RNA hairpins (*11*, *18*).

Genomic RNA Aptamer Properties. The binding properties of selected genomic sequences A−G were evaluated by SPR. Equilibrium dissociation constants (K_d) ranged from low to high nanomolar: B (K_d = 4.4 ± 1.6 nM) > F (K_d = 9.0 ± 0.9 nM) > G (K_d = 14.1 ± 1.2 nM) > A (K_d = 17.4 ± 1.9 nM) > E (K_d = 40.9 ± 8.6 nM) > D (K_d = 136 ± 8 nM) > C (K_d = 297 ± 11 nM). The binding constants of sequences E and F are in fair agreement with the results previously obained for R06 [K_d = 17 nM (*34*)] which shares the same loop sequence.

Table 1: Rate (k_{on} and k_{off}) and Equilibrium Dissociation (K_d) Constants of the R06−TAR Complex and Various Aptamer a1.16−TAR Kissing Complexes in R Buffer at 23 °C[a]

	k_{on} ($\times 10^4$ M^{-1} s^{-1})	k_{off} ($\times 10^{-3}$ s^{-1})	K_d (nM)
R06	6.30 ± 0.60	1.10 ± 0.10	17.00 ± 3.00
a1.16	4.23 ± 0.12	0.57 ± 0.01	13.60 ± 0.39
a1.16C_3	6.78 ± 0.03	19.15 ± 1.35	282.00 ± 21.00
a1.16U	10.83 ± 0.20	2.88 ± 0.22	26.60 ± 2.16
a1.16G	3.88 ± 0.40	4.30 ± 0.24	111.67 ± 5.04
a1.16C	6.01 ± 0.30	94.66 ± 7.45	> 1000
a1.24(GA)	3.34 ± 0.13	11.60 ± 0.70	348.50 ± 34.50

[a] see Figure 1 for variant sequences.

Sequence B is the strongest TAR binder even though the formation of at most 5 bp in the loop−loop helix is expected, due to the presence of an A at the 5′ position of the aptamer loop. We focused on this 5′ ACCCAG loop sequence and characterized in more detail the binding of candidate a1.

In contrast to aptamer R06 identified previously which has an eight-nucleotide loop, genomic aptamer a1 is characterized by a predicted six-nucleotide loop, five nucleotides of which are complementary to the TAR one. We investigated the role of the A residue located on the 5′ side by monitoring the affinity of different a1 variants (Figure 1b,c). Toward this end, a truncated version of the genomic RNA aptamer has been synthetized, called a1.16. This shortening slightly reduced the stability of the complex (K_d = 13.6 ± 0.4 nM) compared to that of the parent molecule (Table 1). The replacement of the 5′ A with a non-nucleotide linker (three carbons) led to a K_d value of 282 ± 21 nM. The substitution of A at the 5′ position of the loop with either U, G, or C led to a destabilization of the complex, the amplitude of which increased in this order (Table 1). The a1.16 (A→U)−TAR complex was the most stable of the mutated complexes; this might be due to the formation of an additional A-U pair at the end of the loop−loop helix. In contrast, the presence of C was detrimental to the aptamer−TAR association (K_d > 1 μM).

On and off rate constants for a1-TAR (k_{on} = 4.23 × 10^4 M^{-1} s^{-1}, and k_{off} = 0.57 × 10^{-3} s^{-1}) were approximatively 2 times lower

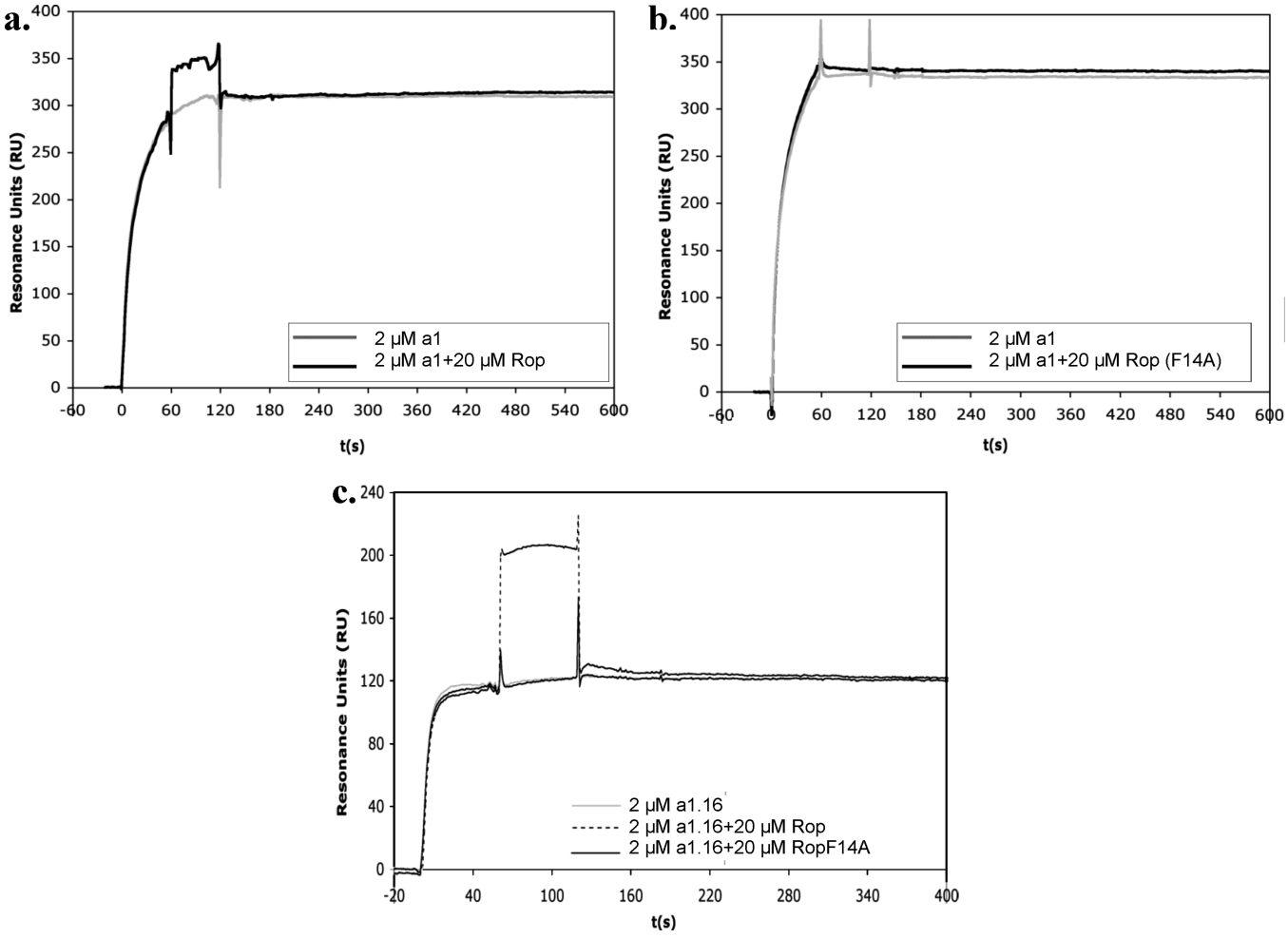

FIGURE 4: Rop binding to TAR−aptamer complexes. Biotinylated TAR was immobilized; aptamers and Rop protein were passed over the SPR instrument on the surface at a rate of 20 μL/min in R buffer (see Materials and Methods for details). Sensorgrams were obtained with either full-length a1 (a and b) or the truncated a1.16 aptamer at 2 μM (panel c; see Figure 1b). Wild-type Rop (a and c) or mutated Rop F14A (b and c) was used at the indicated concentration.

than those of the R06−TAR kissing complex (k_{on} = 6.30 × 10^4 M^{-1} s^{-1}, and k_{off} = 1.10 × 10^{-3} s^{-1}) (Table 1) (*34*). Interestingly, affinity variation was generally due to off rate differences, the k_{off} for the weakest binder (a1.16C) being 2 orders of magnitude higher than that of the parent aptamer (Table 1). For the A sequence, the on and off rate constants were 9.52 × 10^3 M^{-1} s^{-1} and 1.65 × 10^{-4} s^{-1}, respectively.

We wondered whether the 5′ ACCCAG motif could generate strong interaction in the context of the previously identified R06 kissing aptamer. This hexanucleotide has been substituted for the original sequence in the eight-nucleotide loop closed by G and A residues. The resulting variant, a1.24, was still predicted to fold into a hairpin. The dissociation constant of the complex it formed with TAR was ~350 nM, i.e., more than 1 order of magnitude higher than that of the R06−TAR complex. These results demonstrated that both the loop and the top of the stem of the genomic aptamer a1 are crucial for strong interaction with TAR.

Rop Recognition of the a1−TAR Complex. Rop (repression of primer) protein from *E. coli* interacts specifically with and stabilizes the loop−loop interaction of the RNAI−RNAII regulatory complex. It is known to recognize natural and artificial kissing complexes (*4*, *39*, *40*). We investigated the interaction between Rop and the genomic aptamer−TAR complex by SPR. As shown in Figure 4a, a resonance signal was detected when Rop was injected in the presence of a1 on a chip on which TAR was immobilized. Such a signal was not observed when Rop was added in the absence of a1, indicating that the protein interacted with the a1−TAR complex. No signal was detected either when the mutated Rop (F14A) protein which does not bind to kissing complexes was used (Figure 4b). Similar results were obtained with the truncated a1.16 derivative (Figure 4c). This strongly suggests that the a1−TAR complex involves loop−loop (kissing) interaction. The equilibrium dissociation constant of Rop for the a1.16−TAR complex is rather high (K_d > 40 μM).

The a1−TAR Kissing Complex Does Not Correspond to a Natural Interaction. Genomic SELEX has been performed to identify genomic RNA aptamers that might reflect a natural interaction between the TAR RNA element of HIV-1 and human transcripts. As the starting library was generated by in vitro transcription of the entire genome (*32*), selected sequences might actually not correspond to natural transcripts in the cell. We have localized selected sequences on the human genome using the humanBLAT software (http://genome.ucsc.edu/cgi-bin/hgBlat). The a1 sequence was localized on chromosome 16(p12.2) and corresponds to the complementary sequence of one intron of the BC056679 transcript. The a1 sequence did not correspond to any identified human transcript, taking into account available databases. Moreover, RT-PCR experiments conducted with RNA extracted from different human cell lines (AGS, H9, HEK-293,

Hela, Jurkat, THP-1, and U937) did not allow the detection of a natural transcript containing a1. Investigations are under way for other selected sequences.

DISCUSSION

Genomic SELEX allows the identification of putative functional interactions involving RNA motifs. Because of the way the library is built, the starting RNA pool does not depend on the level of differentially expressed RNAs (with respect to different developmental stages or growth conditions), or RNAs expressed in low abundance (*32*). In addition, it may allow the detection of interactions involving noncoding RNAs. Conversely, in vitro synthesized RNAs derived from the entire genome do not represent celullar RNA species. Genomic SELEX carried out against the TAR element of HIV-1, using a human RNA pool, led to the identification of several aptamer sequences. Even though a natural transcript containing the selected candidates could not be identified, these aptamers which display a strong affinity for the target RNA hairpin are artificial ligands of interest.

All selected sequences possess a 5′ CCCAG consensus motif complementary to part of the 5′ CUGGGA TAR hexaloop sequence. All of them but sequence A were predicted to form hairpin structures, the 5′-CCCAG motif being included in the loops. Interestingly, the two main families, A and B, exhibit similar binding affinity. In vitro selection shows that different binding modes may give rise to stable RNA−TAR complexes. The results obtained for the B−G aptamers suggest that these oligonucleotides interact with TAR through loop−loop association in agreement with previous studies. Interestingly, the strong binder a1 aptamer is characterized by a 5′ ACCCAG loop sequence, different from what has been previously identified (*18*). The a1 sequence has been localized in the human genome. Disappointingly enough, the a1 sequence does not correspond to any known human mRNA, miRNA, or snoRNA in data banks, and we failed to identify it experimentally in different cell lines.

The a1 genomic aptamer was found to form a complex with TAR that is recognized by the protein Rop, strongly suggesting the formation of a kissing interaction. The affinity of a1 for TAR was ∼4 times higher than that of the previously identified aptamer R06 (K_d = 4 and 17 nM, respectively) using an RNA library (*18*).

Kissing complexes are restricted to a limited set of hairpin loops. Comparison of sense−antisense RNA complexes showed that the most common size of kissing loops is six or seven nucleotides (*41*). Moreover G-C pairs are crucial for nucleating the kissing interaction (*2*, *12*). Lastly, purine residues have been shown to be essential at both the first and last positions of kissing loops. This holds true for natural dimerization elements of the HIV-1 genome (*12*) and for aptamer R06 (*34*). These flanking purines were proposed to be involved in noncanonical interactions (*42*). In the case of the TAR−R06 aptamer complex, molecular modeling (*43*), NMR studies (*44*, *45*), and the crystal structure (*35*) pointed out the key role played by the G-A pair closing the aptamer loop in the stabilization of the kissing complex. The a1 aptamer selected in the frame of our genomic SELEX fulfils all the criteria of loop size and sequence for generating a stable kissing complex with the TAR element.

Interestingly, an adenine residue is present at the 5′ position of the aptamer loop. In contrast to R06, this first residue does not allow the formation of a sixth Watson−Crick base pair with the TAR loop. Any substitution of this A resulted in a decrease in the stability of the aptamer−TAR complex, including the A → U substitution that might give rise to an AU pair. Therefore, this A nucleotide has been selected during the process. It is more than a spacer connecting the two ends of the loop−loop helix. It might contribute to the stability of the complex through stacking or hydrogen bond interactions as shown for other kissing complexes (*35*, *45*).

TAR*, a rationally designed RNA hairpin, possesses the same six-nucleotide 5′ UCCCAG loop sequence as the a1.16U aptamer (Figure 1e) (*29*). Previous SPR experiments have shown that the TAR−TAR* kissing complex was not sufficiently stable to determine its equilibrium association constant under the ionic conditions used for the a1−TAR complex (*34*). This is to be ascribed to the different stems of the two anti-TAR hairpins. Indeed, it was reported in the case of the R06 aptamer that the base pairs next to the loop have an impact on the binding (*34*). Therefore, genomic SELEX allowed the identification of a new kissing motif of high affinity. Both the loop and the upper part of the stem do contribute to the association with the target. Aptamer a1 constitutes a new TAR ligand that might prove to be useful in designing new molecules susceptible to interference with the function of the TAR element in the retroviral life cycle.

ACKNOWLEDGMENT

We thank Sébastien Fribourg for cloning Rop in *E. coli*, Noel Pinaud and Thierry Dakhli for purifying the Rop protein, Nathalie Pierre for synthesizing the biotinylated RNA, and Guillaume Drutel and Thierry Teste-Lasserre for the qRT-PCR experiments.

SUPPORTING INFORMATION AVAILABLE

Detailed conditions of the in vitro selection rounds as well as the predicted secondary srtuctures of the different sequences (see the text). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

(1) Wagner, E. G. H., and Simons, R. W. (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol. 48*, 713–742.
(2) Brunel, C., Marquet, R., Romby, P., and Ehresmann, C. (2002) RNA loop-loop interactions as dynamic functional motifs. *Biochimie 84*, 925–944.
(3) Brantl, S. (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol. 10*, 102–109.
(4) Eguchi, Y., and Tomizawa, J. (1991) Complexes formed by complementary RNA stem-loops. Their formations, structures and interaction with ColE1 Rom protein. *J. Mol. Biol. 220*, 831–842.
(5) Eguchi, Y., and Tomizawa, J. I. (1990) Complex formed by complementary RNA stem-loops and its stabilization by a protein: Function of ColE1 Rom protein. *Cell 60*, 199–209.
(6) Tomizawa, J. (1990) Control of ColE1 plasmid replication. Interaction of Rom protein with an unstable complex formed by RNA I and RNA II. *J. Mol. Biol. 212*, 695–708.
(7) Tomizawa, J. I. (1990) Control of ColE1 plasmid replication: Intermediates in the binding of RNA I and RNA II. *J. Mol. Biol. 212*, 683–694.
(8) Muriaux, D., Foose, P., and Paoletti, J. (1996) A kissing complex together with a stable dimer is involved in the HIV-1_{Lai} RNA dimerization process in vitro. *Biochemistry 35*, 5075–5082.
(9) Paillart, J. C., Westhof, E., Ehresmann, C., Ehresmann, B., and Marquet, R. (1997) Non-canonical interactions in a kissing loop complex: The dimerization initiation site of HIV-1 genomic RNA. *J. Mol. Biol. 270*, 36–49.
(10) Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) Identification of the primary site of the human

immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl. Acad. Sci. U.S.A. 91*, 4945–4949.
(11) Scarabino, D., Crisari, A., Lorenzini, S., Williams, K., and Tocchini-Valentini, G. P. (1999) tRNA prefers to kiss. *EMBO J. 18*, 4571–4578.
(12) Lodmell, J. S., Ehresmann, C., Ehresmann, B., and Marquet, R. (2000) Convergence of natural and artificial evolution on an RNA loop-loop interaction: The HIV-1 dimerization initiation site. *RNA 6*, 1267–1276.
(13) Lodmell, J. S., Ehresmann, C., Ehresmann, B., and Marquet, R. (2001) Structure and dimerization of HIV-1 kissing loop aptamers. *J. Mol. Biol. 311*, 475–490.
(14) Kikuchi, K., Umehara, T., Fukuda, K., Hwang, J., Kuno, A., Hasegawa, T., and Nishikawa, S. (2003) RNA Aptamers Targeted to Domain II of Hepatitis C Virus IRES That Bind to Its Apical Loop Region. *J. Biochem. 133*, 263–270.
(15) Aldaz-Carroll, L., Tallet, B., Dausse, E., Yurchenko, L., and Toulmé, J. J. (2002) Apical loop-internal loop interactions: A new RNA-RNA recognition motif identified through in vitro selection against RNA hairpins of the hepatitis C virus mRNA. *Biochemistry 41*, 5883–5893.
(16) Da Rocha Gomes, S., Dausse, E., and Toulmé, J. J. (2004) Determinants of apical loop-internal loop RNA-RNA interactions involving the HCV IRES. *Biochem. Biophys. Res. Commun. 322*, 820–826.
(17) Boiziau, C., Dausse, E., Yurchenko, L., and Toulmé, J. J. (1999) DNA aptamers selected against the HIV-1 TAR RNA element form RNA/DNA kissing complexes. *J. Biol. Chem. 274*, 12730–12737.
(18) Ducongé, F., and Toulmé, J. J. (1999) In vitro selection identifies key determinants for loop-loop interactions: RNA aptamers selective for the TAR RNA element of HIV-1. *RNA 5*, 1605–1614.
(19) Darfeuille, F., Sekkai, D., Dausse, E., Kolb, G., Yurchenko, L., Boiziau, C., and Toulmé, J.-J. (2002) Driving in vitro selection of anti-HIV-1 TAR aptamers by magnesium concentration and temperature. *Comb. Chem. High Throughput Screening 5*, 315–327.
(20) Kolb, G., Reigadas, S., Castanotto, D., Faure, A., Ventura, M., Rossi, J. J., and Toulmé, J. J. (2006) Endogenous expression of an anti-TAR aptamer reduces HIV-1 replication. *RNA Biol. 3*, 150–156.
(21) Dayton, A. I., Sodroski, J. G., Rosen, C. A., Goh, W. C., and Haseltine, W. A. (1986) The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell 44*, 941–947.
(22) Muesing, M. A., Smith, D. H., and Capon, D. J. (1987) Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. *Cell 48*, 691–701.
(23) Feng, S., and Holland, E. C. (1988) HIV-1 tat trans-activation requires the loop sequence within tar. *Nature 334*, 165–167.
(24) Berkhout, B., Silverman, R. H., and Jeang, K. T. (1989) Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell 59*, 273–282.
(25) Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., and Crothers, D. M. (1990) Fragments of the HIV-1 tat protein specifically bind TAR RNA. *Science 249*, 1281–1285.
(26) Gait, M. J., and Karn, J. (1993) RNA recognition by the human immunodeficiency virus Tat and Rev proteins. *Trends Biochem. Sci. 18*, 255–259.
(27) Zhou, Q., and Sharp, P. A. (1996) Tat-SF1: Co-factor for stimulation of transcriptional elongation by HIV-1 Tat. *Science 274*, 605–610.
(28) Jaeger, J. A., and Tinoco, I. J. (1993) An NMR study of the HIV-I TAR element hairpin. *Biochemistry 32*, 12522–12530.
(29) Chang, K. Y., and Tinoco, I. (1994) Characterization of a "kissing" hairpin complex derived from the human immunodeficiency virus genome. *Proc. Natl. Acad. Sci. U.S.A. 91*, 8705–8709.
(30) Chang, K. Y., and Tinoco, I. (1997) The structure of an RNA "kissing" hairpin complex of the HIV TAR hairpin loop and its complement. *J. Mol. Biol. 269*, 52–66.
(31) Gold, L., Brown, D., He, Y., Shtatland, T., Singer, B. S., and Wu, Y. (1997) From oligonucleotide shapes to genomic SELEX: Novel biological regulatory loops. *Proc. Natl. Acad. Sci. U.S.A. 94*, 59–64.
(32) Lorenz, C., von Pelchrzim, F., and Schroeder, R. (2006) Genomic systematic evolution of ligands by exponential enrichment (Genomic SELEX) for the identification of protein-binding RNAs independent of their expression levels. *Nat. Protoc. 1*, 2204–2212.
(33) Huttenhofer, A., and Vogel, J. (2006) Experimental approaches to identify non-coding RNAs. *Nucleic Acids Res. 34*, 635–646.
(34) Ducongé, F., Di Primo, C., and Toulmé, J. J. (2000) Is a closing "GA pair" a rule for stable loop-loop RNA complexes?. *J. Biol. Chem. 275*, 21287–21294.
(35) Lebars, I., Legrand, P., Aime, A., Pinaud, N., Fribourg, S., and Di Primo, C. (2008) Exploring TAR-RNA aptamer loop-loop interaction by X-ray crystallography, UV spectroscopy and surface plasmon resonance. *Nucleic Acids Res. 36*, 7146–7156.
(36) Predki, P. F., Nayak, L. M., Gottlieb, M. B., and Regan, L. (1995) Dissecting RNA-protein interactions: RNA-RNA recognition by Rop. *Cell 80*, 41–50.
(37) Di Primo, C., and Lebars, I. (2007) Determination of refractive index increment ratios for protein-nucleic acid complexes by surface plasmon resonance. *Anal. Biochem. 368*, 148–155.
(38) Zuker, M. (1989) Computer prediction of RNA structure. *Methods Enzymol. 180*, 262–288.
(39) Comolli, L. R., Pelton, J. G., and Tinoco, I. (1998) Mapping of a protein-RNA kissing hairpin interface: Rom and Tar-Tar. *Nucleic Acids Res. 26*, 4688–4695.
(40) Darfeuille, F., Cazenave, C., Gryaznov, S., Duconge, F., Di Primo, C., and Toulmé, J. J. (2001) RNA and N3′→P5′ kissing aptamers targeted to the trans-activation responsive (TAR) RNA of the human immunodeficiency virus-1. *Nucleosides, Nucleotides Nucleic Acids 20*, 441–449.
(41) Franch, T., and Gerdes, K. (2000) U-turns and regulatory RNAs. *Curr. Opin. Microbiol. 3*, 159–164.
(42) Jossinet, F., Paillart, J. C., Westhof, E., Hermann, T., Skripkin, E., Lodmell, J. S., Ehresmann, C., Ehresmann, B., and Marquet, R. (1999) Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. *RNA 5*, 1222–1234.
(43) Beaurain, F., Di Primo, C., Toulmé, J. J., and Laguerre, M. (2003) Molecular dynamics reveals the stabilizing role of loop closing residues in kissing interactions: Comparison between TAR-TAR* and TAR-aptamer. *Nucleic Acids Res. 31*, 4275–4284.
(44) Lebars, I., Richard, T., Di Primo, C., and Toulmé, J. J. (2007) NMR structure of a kissing complex formed between the TAR RNA element of HIV-1 and a LNA-modified aptamer. *Nucleic Acids Res. 35*, 6103–6114.
(45) Van Melckebeke, H., Devany, M., Di Primo, C., Beaurain, F., Toulmé, J. J., Bryce, D. L., and Boisbouvier, J. (2008) Liquid-crystal NMR structure of HIV TAR RNA bound to its SELEX RNA aptamer reveals the origins of the high stability of the complex. *Proc. Natl. Acad. Sci. U.S.A. 105*, 9210–9215.