Design and synthesis of RNA miniduplexes via a synthetic linker approach. 2. Generation of covalently closed, double-stranded cyclic HIV-1 TAR RNA analogs with high Tat-binding affinity

Michael Y.-X. Ma, Kirk McCallum, Shane C. Climie, Raya Kuperman, Wing C. Lin, Martin Sumner-Smith and Richard W. Barnett
Allelix Biopharmaceuticals Inc., 6850 Goreway Drive, Mississauga, Ontario L4V 1P1, Canada

Received March 24, 1993; Revised and Accepted May 4, 1993

ABSTRACT

We recently developed an approach which allows rapid generation of short, double-stranded oligonucleotides whereby one end of the duplex was joined and stabilized by a synthetic linker of specific design (miniduplexes)(6). Model miniduplexes based on the HIV-1 TAR RNA hairpin were shown to be thermodynamically stable and good substrates for binding by the HIV-1 Tat protein which normally bind to natural TAR (6). In this study, we have extended our studies to the design, synthesis and analysis of the binding properties of covalently closed, double-stranded, cyclic RNA miniduplexes. A strategy using automated chemical synthesis and T4 RNA ligase-catalyzed cyclization was employed to generate cyclic oligoribonucleotides. When both ends of a shortened, wild-type TAR RNA stem (9 bp) were covalently linked through either nucleotidic loops (4–6 nt) or synthetic linkers (derivatized from hexaethylene glycol), the resulting cyclic TAR RNA analogs were good substrates for binding by both Tat-derived peptide or full-length Tat protein. Interestingly, the cyclic TAR analogs failed to show any binding if the synthetic linker was reduced in length (e.g. derivatized from triethylene glycol), although such linkers are acceptable in the hairpin-shaped miniduplexes series (6). This implies that RNA conformational changes are required for Tat binding and that these changes are restricted in certain cyclic variants. Our findings suggest that covalently-closed nucleic acid miniduplexes may be useful both to study nucleic acid-protein interactions as well as to provide a basis for therapeutic intervention as transcription decoys.

INTRODUCTION

Oligonucleotide-based therapeutics are of considerable current interest. In principle, one of their potential benefits, once practical chemical forms have been devised, would be the ability to rapidly develop therapeutics for a wide range of diseases through simple modifications of sequence. Most current research efforts in this emerging field have been devoted to the application of antisense and antigen approaches (1–2), or ribozyme technology (3). In these approaches, the targets are either single-stranded (ss) messenger RNA or double-stranded (ds) genomic DNA. Another oligonucleotide-based therapeutic approach which has received less attention is the development of decoys targeting essential, nucleic acid-binding, regulatory proteins (e.g. transcription factors) (4). This approach, which has been termed the sense approach, shares similar simplicity of the antisense approach in the design of sequence-specific inhibitors, but it will also encounter the same, or even greater, challenges as every other oligonucleotide-base strategy, i.e. difficulties in economic large-scale production, cellular uptake, in vivo stability, efficacy, etc.

One advantage of the sense approach is that ds nucleic acids are generally more stable than their ss homologs (5). Therefore, minor, chemical modifications may be sufficient to stabilize these molecules inside cells. A major limitation of the sense approach is that two relatively long oligonucleotides have to be hybridized together to form a stable double-stranded structure. This will substantially increase synthetic cost and create possible pharmacokinetic hurdles. In order to reduce the size of sense oligonucleotides we have recently developed a 'synthetic linker approach' whereby smaller duplexes (miniduplexes) can be generated directly from standard oligonucleotide synthesis (6, 7). Using this approach, any nucleic acid duplex of reasonable length (e.g. <40 bp) can be generated by spontaneous folding of a single oligonucleotide instead of the traditional method of hybridizing two separate strands. In addition, the thermal stability of the resulting hairpin-shaped duplexes was dramatically increased, while specific recognition by binding protein was maintained (6).

In general, nucleic acid miniduplexes can be divided into two groups: the hairpin-shaped miniduplexes in which one end of the stem is connected by a nucleotidic loop or a synthetic linker (Fig. 1, structures 1 and 3); and the dumbbell-shaped miniduplexes in which both ends of the stem are covalently closed by nucleotidic loops or synthetic linkers (Fig. 1, structure 2, and
structures 4–6). The dumbbell-shaped miniduplexes, which were explored in the work described in this paper, may provide several advantages over their hairpin-shaped counterparts. First, they should have enhanced thermal stability since all termini have been constrained. This additional thermal stability is useful when modified oligonucleotides (e.g., ds phosphorothioates) are considered as potential transcription decoys (8) or very small duplexes (6 to 8 bp) are required. Second, they should have enhanced enzymatic stability against various exonucleases and single-strand-specific endonucleases when synthetic linkers are used as the bridging chains (5) (Fig. 1, structures 4–6). A third benefit of dumbbell-shaped miniduplexes is that they should retain their double-stranded, monomeric duplex structures even at very high oligonucleotide concentrations, thus avoiding the tendency of hairpin-shaped miniduplexes to form alternate structures, such as bulged dimeric duplexes (9, 10). We report herein our recent studies on the design, synthesis and binding properties of covalently closed, double-stranded cyclic RNA miniduplexes.

MATERIALS AND METHODS

Chemical synthesis of oligoribonucleotides, preparation of linker-derivatized 3-cyanoethyl phosphoramidites, RNA gel retardation assays, and measurements of melting temperatures ($T_m$) were carried out as previously described (6).

Preparation of internally $^{32}$P-labeled cyclic TAR RNA analogs (Fig. 1, oligomers 2 and 4–6) was performed as follows: 5 pmol of gel-purified RNA precursor (the respective ligation precursors were derived from ligation site b, Fig. 2) was dissolved in 14 µl of sterile water. 2 µl of [γ-$^{32}$P]ATP (30 µCi) (Amersham) and 1 µl of T4 polynucleotide kinase (NEB) were then added for a total volume of 20 µl. Following incubation at 37°C for 1–1.5 h, the reaction mixture was then heated to 85°C for 3 min, and slowly cooled to room temperature. 10 µl of the mixture was kept as a control to determine the labeling yield and specific activity. 5 µl of sterilized water, 2 µl of DMSO, 2 µl of freshly prepared 10 mM cold ATP, and 1 µl of T4 RNA ligase (NEB) were added to the remaining 10 µl of mixture. The ligation reaction was performed either at 37°C for 4 h or at room temperature overnight. Once the ligation was complete, the reaction mixture was heated to 90°C for 5 min, rapidly cooled on ice and purified on 20% polyacrylamide/7M urea gels. The predominant, faster moving band (relative to the unligated starting material) was excised and eluted from the gel matrix with 0.3 M sodium acetate. The resulting oligonucleotide was phenol-chloroform extracted and recovered by ethanol precipitation. Larger amounts of ligated materials were purified by reverse-phase HPLC [Waters Model 510 HPLC system, Synchropak RP-P C$_{18}$ column (SynChrom, Inc.)] according to previously published procedures (11–14).

The binding dissociation constants ($K_D$) were determined by quantitative binding assays. Briefly, starting with a fixed amount of $^{32}$P-labeled TAR RNA analog, the binding reactions were followed by adding an increasing amount of unlabeled full-length Tat protein (American Bio-Technologies, Inc.) over a wide range of concentrations (0.5 to 1000 nM) (6). The amount of material in each spot was quantified and the data were fit to the equation for a simple bimolecular equilibrium.

RESULTS AND DISCUSSION

In this study, we examined shortened variants of the human immunodeficiency virus type-1 (HIV-1) trans-activation response RNA element (TAR) as a model, duplex system (for recent reviews see 15–19). A viral transactivating protein called Tat is known to bind to the loop proximal stem of TAR. Tat-TAR binding is a relatively specific event centered at a two or three-nucleotide (nt) bulge in the stem. The binding event is sensitive to structural changes in TAR RNA, as shown for example, by the finding that analogs with only a single nucleotide bulge are not good binding substrates unless a synthetic, ‘spacer’ linker is introduced in addition to the single nucleotide (20).

Design of cyclic TAR RNA analogs

Full-length, viral TAR is a 59-nt hairpin, but a number of studies have demonstrated that hairpins as short as 27-nt are good substrates for Tat binding (17, 21, 22). As a starting structure for our covalent cyclization studies we used a 27-nt TAR-based oligoribonucleotide (Fig. 1, oligomer 1) which folds into a specific and stable stem-loop structure with a 3-nt bulge within the stem (21). We have previously shown that the 6-nt loop is completely dispensable for Tat binding, since it can be replaced by synthetic linkers of various lengths [e.g., the loop substituted variant oligomer 3 exhibits thermal stability ($T_m$) and binding affinity ($K_D$) similar to that of the wild-type sequence (1)](6). Two such linkers were used in this study; one was derived from triethylene glycol [linker 1 (L$_1$)] (Fig. 1, oligomers 3 and 4), and the other from hexaethylene glycol [linker 2 (L$_2$)] (Fig. 1, oligomers 5 and 6). The only difference between these two linkers is their respective chain length (e.g. the distance between the two terminal phosphates is approximately equivalent to a C$_6$ chain in oligomer 4 versus a C$_{17}$ chain in oligomer 5).

In designing dumbbell-shaped cyclic analogs we reasoned that connection of the two remaining free ends in TAR RNA should further stabilize the lower stem segment, thereby increasing the

---

**Figure 1. Design of covalently-closed, double-stranded cyclic TAR RNA miniduplexes.** The hairpin-shaped oligomers 1 and 3 were used as the 'starting' molecules since both are capable of binding to full-length Tat protein and Tat-derived peptide. The two remaining free ends of the TAR stem were covalently closed either through a nucleotidic loop (oligomer 2) or synthetic linkers of varying length (a triethylene glycol linker for oligomer 4 and hexaethylene glycol linker for oligomers 5–6). The cyclic oligomer 6 has a 8-bp stem instead of a 9-bp stem as in oligomer 5.
overall stability of the duplex. We chose to employ either a nucleotidic loop or synthetic linker. Thus, several cyclic miniduplexes were designed with either two nucleotide loops (Fig. 1, oligomer 2) or two synthetic linkers of varying length (Fig. 1, L1 in oligomer 4, and L2 in oligomers 5 and 6). The lower nucleotide loop employed in oligomer 2 was chosen on the basis of the previous finding by Cheong et al. (23) that a UUGC tetraloop found in naturally occurring RNAs possesses unusual stability.

Cyclization

In order to generate radio-labeled cyclic oligoribonucleotides for RNA binding assays, a chemo-enzymatic synthetic strategy using phosphoramidite chemistry and T4 RNA ligase was adapted (Fig. 2). Briefly, the ligation precursors were first synthesized by standard β-cyanoethyl phosphoramidite chemistry on a polymeric CPG support. After cleavage and depurination, the crude oligoribonucleotides were purified on 20% polyacrylamide denaturing gels. The purified oligomers were then 5'-end-labeled with [γ-32P]ATP and T4 polynucleotide kinase, and the resulting 5'-phosphorylated ligation precursors converted directly to the covalently-closed, cyclic TAR RNAs by T4 RNA ligase in the presence of ATP. No significant difference in ligation efficiency was observed if the 5'-end-labeled precursor was gel purified prior to ligation. On the other hand, addition of 10% DMSO to the ligation mixture sometimes increased the ligation yield (24).

The most important factor affecting ligation efficiency was the ligation site. Of the five ligation sites tested (Fig. 2, sites a-e), ligation site b gave the cleanest ligation product (Fig. 3). This result is consistent with the previously published observation that adenosine is the preferred phosphate donating residue for T4 RNA ligase catalyzed-ligation (25). Sites a and e also produced the desired ligated oligomer, but with lower yield and more side products. In contrast, ligation sites c and d gave virtually no ligated product. Ligation sites c and d were initially chosen due to the expectation that these two precursors would bind to Tat protein prior to ligation. Therefore, the effect of ‘locking’ an initially bioactive molecule into a more constrained structure could be examined directly. However, it appears that T4 RNA ligase requires several single-stranded nucleotide residues flanking the putative ligation site in order to achieve efficient ligation. As a result, all the cyclic TAR RNAs used in this study were prepared using ligation site b. Under the optimum ligation conditions, conversion of the ligation precursor to ligated product was almost quantitative, with some minor dimeric contaminants. The ligated RNAs were gel purified and characterized by their faster migration patterns on denaturing polyacrylamide gels (Fig. 3, lanes 2, 6 and 10) and resistance to enzyme action. The ligation products were resistant to 5'-dephosphorylation by calf intestinal alkaline phosphatase (Fig. 3, lanes 4, 8 and 12) and to degradation by calf spleen phosphodiesterase (data not shown) (26).

Thermal stability

Recently, Ashley and Kushlan reported thermal stability studies for a series of DNA dumbbells (27). They observed that the melting temperatures of DNA dumbbells were significantly higher (28–41°C) than their nicked, ligation precursors. A similar increase in thermal stability (25–50°C) was also reported by Eri et al. (11). We have performed similar comparative thermal denaturation experiments with some of the cyclic TAR RNA analogs.

Figure 2. Chemo-enzymatic synthesis of double-stranded cyclic TAR RNA analogs. The chemically synthesized ligation precursors were 5'-end-labeled with [γ-32P]ATP and T4 polynucleotide kinase. The resulting oligomers were then cyclized with T4 RNA ligase using ligation sites a–e. Ligation site b gave the most efficient ligation, whereas ligation sites c and d gave virtually no ligated products. Ligation sites a and e produced the correct ligated oligomer but with more side products (see Figure 3).

Figure 3. Electrophoretic analysis of the ligation reaction and characterization of the ligated product(s) by alkaline phosphatase treatment. The ligation precursors were derived from either ligation site site b [e.g., 5'-AUCUGAACCUCUGGAGCUGUCUCUCUCUUCGGAG-3' for oligomer 2 (lane 1), 5'-AUCUGAGCL2GCCUGUCUCUCUCUGAG3'-1 for oligomer 5 (lane 5) or site e [e.g., 5'-UCUCCL2GGAGAUCCUGAGCL2GCUC-3' for oligomer 5 (lane 9)]. Lanes 2, 6 and 10 are the respective cyclic oligomers after ligation. Treatment of the ligation precursors (lanes 3, 7 and 11) and ligated oligomers (lanes 4, 8 and 12) with calf intestine alkaline phosphatase (37°C, 10 h) revealed that the cyclic oligomers become resistant to 5'-dephosphorylation, suggesting they are internally-labeled. Since these oligomers migrate faster than their ligation precursors on denaturing polyacrylamide gels, the presence of internal labeling is the result of cyclization and not dimerization. Comparison between the lanes 6 and 10 indicates that site b is a better ligation site than the site e.
Table I. Thermal stabilities ($T_m$) and binding affinities ($K_D$) of the hairpin and double-stranded cyclic TAR RNA analogs

<table>
<thead>
<tr>
<th>Oligomer</th>
<th>1</th>
<th>2*</th>
<th>2</th>
<th>3</th>
<th>5*</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_m$(°C)</td>
<td>60</td>
<td>71</td>
<td>&gt;85</td>
<td>63</td>
<td>62</td>
<td>82</td>
</tr>
<tr>
<td>$K_D$(nM)</td>
<td>+ (6.40)</td>
<td>–</td>
<td>+ (5.76)</td>
<td>+ (1.81)</td>
<td>–</td>
<td>+ (6.88)</td>
</tr>
</tbody>
</table>

The thermal denaturation experiments were carried out in 100 mM NaCl/10 mM sodium phosphate buffer (pH 7.0) at an oligomer concentration of 3 nM. The samples were heated from 25 to 90°C in 1°C increments. The RNA binding assays were performed at room temperature, using a wide range of Tat protein concentration. The complexes were analyzed on 5% native polyacrylamide gels at 4°C.

* Ligation precursor using ligation site b
+ Strong binding
– No binding

Figure 4. Qualitative RNA gel-shift assays of the hairpin-shaped and dumbbell-shaped TAR RNA analogs. The 32P-labeled oligomers 1–6 were incubated in the absence (–) or presence (+) of 100-fold molar excess of Tat-derived peptide or full-length Tat protein. The complexes were analyzed on a 5% nondenaturing polyacrylamide gel at 4°C. While both cyclic oligomers 2 and 5 showed efficient binding to either Tat-derived peptide or Tat protein, the cyclic oligomer 4 (derivatized from triethylene glycol linker) failed to show any binding. The oligomer 6 displayed weak binding to Tat-derived peptide but no binding to full-length Tat protein.

analogs and their ligation precursors (Table I). The results, presented in Table I, showed an increase in melting temperature of approximately 20°C for the cyclic oligomer 5 ($T_m = 82°C$) relative to its ligation precursor ($T_m = 62°C$), and of at least 14°C for the cyclic oligomer 2. The precise determination of $T_m$ in the latter case was complicated by the fact that an upper temperature plateau was not reached even at 90°C (data not shown). Assuming the ligation precursors and cyclic oligomers exhibit similar hyperchromic effects, the estimated melting temperature is above 85°C. It is noteworthy that although it appears that the ligation precursors may have folded into nicked dumbbell structures prior to ligation since they have similar $R_f$ values as the ligated oligomers on native polyacrylamide gels (Fig. 4), none of them showed any binding to either Tat protein or Tat-derived peptide (see Table I), suggesting structural and/or conformational differences between the unligated and ligated oligomers.

Binding properties
Having demonstrated that it is possible to make stabilized, cyclic, RNA miniduplexes we wished to determine which variants could serve as good substrates for Tat binding. For this analysis we employed both comparative, qualitative binding assays (Fig. 4) as well as quantitative binding methods (Table I).

Figure 4 illustrates the results of an RNA gel retardation assay when a 100-fold molar excess of Tat-derived, TAR-binding peptide [Tat 49−62: RKKRRQRRRPPQGS (28)] or full-length Tat protein were added to each of the cyclic TAR RNAs under binding assay conditions. Interestingly, cyclic TAR analogs 2 and 5 showed efficient binding to both Tat-derived peptide and full-length Tat protein, while cyclic oligomer 4 failed to show any binding. As mentioned previously, the only difference between cyclic analogs 4 and 5 is the length of their synthetic linkers (the synthetic linker in oligomer 5 is twice as long as that in oligomer 4). This result differs from that previously observed in the hairpin-shaped TAR series (6), where there was virtually no difference in Tat binding ability between TAR analogs in which the loop region was replaced by either a triethylene or hexaethylene glycol linker. This observation suggests that the interaction between TAR RNA and Tat protein may entail a significant structural and/or conformational change in the TAR RNA which requires structural flexibility. Since the triethylene glycol linker may not provide sufficient length for such conformational changes if both duplex ends are closed with this linker (L₁), the corresponding cyclic TAR analog 4 is inactive for Tat-binding. Curving and bending of nucleic acids is a common phenomenon in nucleic acid-protein interactions (29). Indeed, a conformational change has been shown to be induced in TAR on binding of Tat (30, 31) or even arginines to TAR (32).

To explore the potential of smaller, cyclic RNA miniduplexes, a series of cyclic TAR analogs with shorter stems (8 to 6 bp) were synthesized using the hexaethylene glycol linker. As assayed by RNA mobility-shift, analog 6 displayed weak binding to Tat-derived peptide, and no binding to full-length Tat protein (Fig. 4, lanes 17–18). None of the other shorter analogs showed any binding to either Tat-peptide or full-length protein (data not shown).

The qualitative RNA binding assays described above demonstrate that when nucleotidic loops or synthetic linkers of appropriate length are used to covalently close both ends of the wild-type TAR RNA stem (9 bp), the resulting double-stranded, cyclic RNAs are capable of binding to their protein target. In order to address the question of whether or not the introduction of such structural constraints would reduce the binding affinity of the cyclic analogs relative to their hairpin homologs, quantitative binding assays with full-length Tat protein were performed and the binding dissociation constants ($K_D$) determined. The results, summarized in Table I, revealed that the cyclic TAR RNA analogs 2 and 5 had binding affinities similar to that of the wild-type hairpin sequence 1, while the linker-derivatized hairpin TAR RNA oligomer 3 showed a moderately better binding affinity than the other derivatives. It is noteworthy to mention that the RNA binding experiments performed in this study address primarily the question of whether a covalently closed, double-stranded, cyclic TAR RNA dumbbell would maintain its high binding affinity for full-length Tat protein and Tat-derived peptide. The issues of binding specificity and exact binding site(s) were not investigated in this study. However, we have previously demonstrated that a hairpin-shaped linker-derivatized TAR analog (e.g. oligomer 3) can effectively compete with wild-type TAR for binding to Tat protein (6). This suggests that the linker-derivatized TAR analogs may bind to Tat protein in a manner similar to that of the wild-type TAR RNA. Understanding of the mode of Tat-TAR interaction, especially the possible interactions between the synthetic linker and Tat protein is of great interest. Such an understanding may lead to the development of better synthetic linkers that further enhance the interaction between nucleic acid miniduplexes and their protein targets.
Enzymatic stability and cellular uptake

Preliminary in vitro stability and cellular uptake studies indicated that the double-stranded cyclic TAR RNAs are much more resistant to degradation in HEp-2 nuclear extracts than the hairpin RNAs (more than 80% of cyclic TAR oligomer 5 remained intact after 8 h while the control oligomer 1 was completely degraded), and that they are taken up by Hut 78 cells (about 10 to 15% were cell-associated after 24 h). Full-length, internally-labeled cyclic TAR RNA 5 was still detectable in the cellular extracts after 48 h. Currently, we are developing a scalable chemical ligation procedure in order to assess the in vivo activity of these double-stranded cyclic RNAs.

CONCLUSION

Results presented in this study have demonstrated that, like their hairpin homologs, covalently closed, double-stranded cyclic nucleic acid miniduplexes can form defined secondary structures. Model cyclic TAR RNA analogs, derived respectively from nucleotidic loops (oligomer 2) or synthetic linkers (oligomer 5), exhibited Tat-binding affinities similar to that of the control hairpin TAR molecule (oligomer 1). However, the linker-derivatized cyclic TAR RNA analogs failed to show any binding if the synthetic linkers were shortened (e.g. derivatized from triethylene glycol), implying that such changes conformationally restrict TAR RNA and thus affect Tat binding. Replacement of the two nucleotidic loops by synthetic linkers reduced the size of the cyclic TAR RNAs from 31 to 21 nt. To our knowledge, these findings represent the first successful application of synthetic linkers to generate double-stranded, cyclic nucleic acid duplexes with high binding affinity for their protein targets. These results are of interest not only for general understanding of nucleic acid-protein interactions, but also for the development of nucleic acid miniduplexes as possible transcription decoys for therapeutic gene regulation.

ACKNOWLEDGEMENTS

We would like to thank Dr Michael Twist for his continuing interest and support for this project, Barbara Dabek for performing preliminary cellular uptake studies, and Suzanne Presseault for isolating HEp-2 nuclear extracts.

REFERENCES