Active site binding and sequence requirements for inhibition of HIV-1 reverse transcriptase by the RT1 family of single-stranded DNA aptamers

Jay D. Kissel1, Daniel M. Held1,2,3, Richard W. Hardy1 and Donald H. Burke3,*

1Department of Biology, Indiana University, Bloomington, IN 47405, 2Biosciences Division, SRI International, Menlo Park, CA 94025 and 3Department of Molecular Microbiology & Immunology and Department of Biochemistry, 471h Life Sciences Center, University of Missouri School of Medicine, Columbia, MO 65211, USA

Received April 9, 2007; Revised May 4, 2007; Accepted May 8, 2007

ABSTRACT

Nucleic acid aptamers can potentially be developed as broad-spectrum antiviral agents. Single-stranded DNA (ssDNA) aptamer RT1t49 inhibits reverse transcriptases (RT) from HIV-1 and diverse lentiviral subtypes with low nanomolar values of Kd and IC50. To dissect the structural requirements for inhibition, RT-catalyzed DNA polymerization was measured in the presence of RT1t49 variants. Three structural domains were found to be essential for RT inhibition by RT1t49: a 5’ stem (stem I), a connector and a 3’ stem (stem II) capable of forming multiple secondary structures. Stem I tolerates considerable sequence plasticity, suggesting that it is recognized by RT more by structure than by sequence-specific contacts. Truncating five nucleotides from the 3’ end prevents formation of the most stable stem II structure, yet has little effect on IC50 across diverse HIV-1, HIV-2 and SIVcpz RT. When bound to wild-type RT or an RNase H active site mutant, site-specifically generated hydroxyl radicals cleave after nucleotide A32. Cleavage is eliminated by either of two polymerase (pol)-active site mutants, strongly suggesting that A32 lies within the RT pol-active site. These data suggest a model of ssDNA aptamer–RT interactions and provide an improved molecular understanding of a potent, broad-spectrum ssDNA aptamer.

INTRODUCTION

HIV-1 reverse transcriptase (RT) copies the viral genomic RNA into double-stranded DNA (dsDNA). Due to its essential role in viral replication and to the early availability of RT inhibitors, RT has been a leading target for anti-retroviral treatments. Currently, over half of the US Food and Drug Administration (FDA) approved anti-retroviral drugs target RT. These drugs fall into two categories: nucleoside analog RT inhibitors (NRTI) block extension of the template DNA upon incorporation into the replicating genome, and non-nucleoside RT inhibitors (NNRTI) bind a hydrophobic pocket near the RT active site resulting in allosteric inhibition (1,2). Although these small-molecule inhibitors have helped slow the progression of AIDS, their long-term utility can be compromised by cellular toxicity and the emergence of drug resistant HIV-1 strains (3–6). The proven effectiveness of anti-RT therapeutics validates the push for new molecular inhibitors of RT. Antagonists that utilize novel inhibition mechanisms are especially attractive in that they may be less cytotoxic and may avoid the current escape mutations associated with NRTIs and NNRTIs.

High-affinity DNA and RNA aptamers have been selected to bind RT. These aptamers inhibit both the polymerase (pol) and RNase H functions of the protein in vitro (7–12) and have the potential to inhibit all steps of reverse transcription, including RNA- and DNA-primed extensions on either RNA- or DNA-templates, strand displacement and RNA cleavage by RNase H (12). Half-maximal inhibition is observed in the picomolar to low micromolar range (7–14), with RNA-primed reactions showing the greatest susceptibility to aptamer inhibition (12). Aptamers appear to compete with primer/template for binding to RT (8,9,15,16), and have accordingly been referred to by some authors as template-primer analog RT inhibitors (TRTIs) (16). Biochemical probing (17) and crystallographic studies (18) have shown that a canonical RNA aptamer folds into a pseudoknot structure and binds to RT in the primer-template binding cleft. Because aptamers exploit inhibitory mechanisms that are distinct from those utilized by small-molecule inhibitors, they offer a unique opportunity in combating HIV.

Several studies have shown that intracellular expression of RNA aptamers to RT protects these cells from HIV-1...
challenge and HIV-1 gene expression (19–23), and that virus produced in cells expressing RNA aptamers are less infectious when applied to aptamer-naive cells (22). This protection extended across multiple HIV-1 subtypes and several drug-resistant viruses (22). Other studies have identified ssDNA aptamers and double-stranded, sulfur-containing thioaptamers that bind the RNase H domain of RT (14). Although the affinity of these aptamers for RT is much weaker than that of ssDNA aptamer RT1t49 (described below), these aptamers also afford protection to cells when administered prior to challenge with low to moderate levels of virus (7,14). The demonstrated antiviral efficacy of aptamers in three distinct modes—expression within target cells, co-packaging into nascent virus within producer cells and exogenous delivery to target cells—motivate further analysis of the molecular basis of RT inhibition by aptamers.

Aptamer RT1 is an 81-nucleotide ssDNA that was selected from a degenerate library containing 35 random positions. It has a reported Kd value of 1 nM and IC50 (reported as Ki) value of <0.3 nM, reflecting an RT-binding affinity that is more than 1000 times greater than that of the library from which it originated (9). The authors introduced random mutations into RT1 and re-selected molecules that retained a high affinity for HIV-1 RT (2 nM after six SELEX cycles versus ~1500 nM Kd for the partially randomized library). Comparative sequence analysis of the reselected species enabled truncation of the original 81-nucleotide aptamer to a 49-nucleotide version, denoted RT1t49, with similar affinity for HIV-1 RT (Kd≈4 nM) (9). A recent study using capillary electrophoresis has shown that RT1t49 binds to HIV-1 RT with a 1:1 stoichiometry, whereas two other DNA aptamers from the same selection (RT12 and RT26) appear to form complexes with two aptamers per RT (24). Using in vitro enzymatic extension assays, we have previously shown that the potent inhibition of RT polymerization and RNase H activities by RT1t49 extends across multiple subtypes of HIV-1, HIV-2 and SIVCPZ RTs, whereas another ssDNA aptamer (RT8) inhibited only the HIV-1 RT subtype utilized in the original selection (36). Fisher et al. (25) have demonstrated inhibition of multiple drug-resistant forms of HIV-1 RT by RT1t49, and observed that mutations in RT that give rise to biochemical resistance to RT1t49 come at a price of a severe replication deficiency in the virus (16). Each of these observations highlights the importance of defining the molecular contacts that determine aptamer specificity.

There is little information on the specific nucleic acid structural requirements of RT1t49 for RT binding. Therefore, we analyzed a panel of RT1t49 variants to dissect the secondary structural requirements of RT1t49. We found that a stable, stem-loop structure with considerable sequence plasticity is important at the 5'-end, as is a shorter stem near the 3'-end, and that multiple regions of the molecule tolerate alterations from the primary sequence. The results of our mutational and deletion analysis of RT1t49 support a model wherein several alternative secondary structures contribute to the high-potency inhibition of HIV-1 RT by RT1t49. In addition, both solution and site-specifically generated hydroxyl radical probing enabled us to identify a specific interaction between RT1t49 position A32 and the pol-active site of HIV-1 RT. Finally, we show that a truncated version of RT1t49 that is unable to form the previously proposed stem II structure nevertheless retains an ability to inhibit RTs across multiple subtypes of HIV-1, HIV-2 and SIVCPZ. The implications of these results for drug resistance and for potential drug development are discussed.

**MATERIALS AND METHODS**

**Materials**

Cy3 fluorophore-labeled and unlabeled DNA oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Nuclease S1 was purchased from Roche and used as recommended by the manufacturer at a dilution of 1:2000 in S1 nuclease buffer. Radioactive nucleotides were obtained from Amersham Biosciences (Piscataway, NJ, USA). RT from HIV-1 strain HXB2 (used in most experiments) and from other sources (used in Table 1) were expressed in Escherichia coli and purified as described (37).

**DNA-dependent DNA polymerization assays**

For RT1t49 mutational analysis, DNA-dependent DNA polymerization was monitored by single nucleotide (ddCTP) extension of a Cy3-labeled 18 nt primer (5' Cy3-GTCCCCGTTCGGCGCGACA 3') on a 40-nt template (5' CAGTGTTGAAAATCTCTGACCTGCGCCGAACACCGGGG 3'). A 1:1 ratio of primer to template (25 nM primer, 37.5 nM template) resulted in nearly complete annealing of primer with template (data not shown). Therefore, reactions were performed using 25 nM primer, 37.5 nM template and 100 ng/μl bovine serum albumin (BSA) in 1x reaction buffer (75 mM KCl, 5 mM MgCl2, 50 mM Tris–HCl pH 8.3, 10 mM DTT) and 0.2 mM ddCTP. Primer and template strands were heated together at 95°C for 2 min and allowed to anneal at room temperature for 5 min. Synthetic ssDNA aptamers were re-suspended in water and added to the primer/template mix (final concentrations of 0.3 to 729 nM, in 3-fold serial dilution increments) prior to the addition of RT. Extension reactions were initiated by addition of HIV-1 RT (0.4 nM final active concentration) and incubated at 37°C. Reactions were quenched after 3 min with two volumes of stop buffer (urea-saturated 95% formamide, 50 mM EDTA and bromophenol blue). Reaction products were separated on 15% denaturing (7.5 M urea) polyacrylamide gels and scanned for Cy3 fluorescence using a Molecular Dynamics Typhoon Phosphorimager. Image analysis was performed using ImageQuant 5.2 software.

The percent extension was calculated by dividing the fluorescent signal from the N + 1 extended product by the sum of the signals from the unextended and extended primer bands. These data were normalized for extension without aptamer (100% relative extension) and without RT (0% relative extension). Normalized data were fit to
a sigmoidal dose-response curve with a constant slope of 1 with GraphPad Prism software using Equation (1):

\[ Y = \frac{100}{1 + 10^{(\log IC_{50} - x)}} \]

Where \( X \) is the log of inhibitor concentration and \( Y \) is measured relative percent extended at a given inhibitor concentration.

Aptamer inhibition of primer extension by RTs from across HIV-1, HIV-2 and SIVcpz clades was performed as previously described (36). Briefly, the Cy3-labeled 18 nt primer was extended on a 103 nt DNA template corresponding to the HXB2 5' LTR U5 and PBS sequences (5'-AAGTAGTGTTGACCCTCTGGTG TGACTCTGTTAACAGATCCCACTACAGCTTT TTATCGGTGTGGAATATCTTCAGCTGGCGC CCGAACAGGGAC-3'). The PBS sequence is underlined. Reactions were assembled in the 1x RT reaction buffer with 30 nM primer, 45 nM DNA template and 0.2 mM dNTPs. Aptamers were added to final concentrations of 0.3 to 243 nM, in 3-fold serial dilution increments. To maximize the dynamic range of polymerization reactions, extension proceeded until just prior to reaching the plateau phase of the product formation curve in unchallenged assays (10 min, data not shown). Reactions were initiated by the addition of enzyme to 0.4 nM final active site concentration, quenched after 10 min in 37°C with two volumes of 95% formamide, 50 mM EDTA and then analyzed as above. Three independent datasets were averaged to calculate the reported IC\(_{50}\) values; uncertainties reflect the standard deviations among IC\(_{50}\) values.

**S1 nuclease digestion of RT1t49**

For S1 nuclease digestion, [\(^{32}\)P] 5'-end-labeled DNA aptamer (>10\(^{6}\) c.p.m.) was unfolded at 95°C for 2 min with 1x S1 nuclease buffer (Roche) in the presence of BSA (100 ng/μl), allowed to fold at room temperature for 5 min, and then kept on ice. Reactions were initiated by adding an equal volume of S1 nuclease (0.4 U/μl in 1x S1 nuclease buffer.) After 15 min on ice, reactions were quenched with two volumes of stop buffer (urea-saturated 95% formamide, 50 mM EDTA and bromophenol blue) and electrophoresed on a 12% denaturing (7.5 M urea) polyacrylamide sequencing gel. The gel was dried, exposed to a storage phosphor screen and analyzed with the Molecular Dynamics Typhoon Phosphorimager.

**Structure probing by site-specifically generated and bulk hydroxyl radicals**

Previously, Gotte et al. (26) showed that by replacing the Mg(II) in the RNase H active site with Fe(II) allows the site-specific generation of OH- radicals that cleave the closest nucleotide in the DNA. Using the same methodology, we monitored both Fe(II)-dependent hydroxyl radical cleavage due to pol- and RNase H-specific cleavage as well as hydroxyl radical cleavage from bulk solution. Site-specifically generated hydroxyl radicals were generated as described (26,27). Briefly, 2.5 μl of [\(^{32}\)P] 5'-end-labeled DNA aptamer (>10\(^{6}\) c.p.m.) was incubated in 10 nM sodium cacodylate, 0.1 mM EDTA, 50 mM KCl, 1 mM CaCl\(_2\), 5 mM MgCl\(_2\) and 100 μg/ml BSA in a total volume of 16 μl. The sample was heated to 95°C for 2 min and cooled to room temperature for 10 min to facilitate aptamer folding. To a final concentration of 560 nM, 2 μl HIV-1 RT (or buffer) was added to the bottom of a 1.5 ml eppendorf tube. The hydroxyl radical solution was made by adding 1 μl of a freshly prepared Fe(II)-EDTA solution to the wall of the tube. A 1 μl drop containing 0.6% hydrogen peroxide (freshly diluted from a 30% stock) was added directly to the Fe(II)-EDTA solution, and finally 1 μl of 20 mM sodium ascorbate was added to the Fe(II)-EDTA/H\(_2\)O\(_2\) solution. The reaction was initiated by ‘flicking’ the hydroxyl radical solution into the DNA aptamer solution in the bottom and was quenched with stop solution (2.5 mM thioura, 0.1 M sodium acetate and 20 μg glycogen as carrier) 2 min later. Reaction products were precipitated with three volumes of 100% ethanol, and the DNA pellet was re-suspended in 20 μl of loading buffer containing 95% formamide, 50 mM EDTA, bromophenol blue and xylene cyanole and electrophoresed on a 12% denaturing (7.5 M Urea) polyacrylamide sequencing gel and analyzed as above.

**RESULTS AND DISCUSSION**

**Inhibition of HIV-1 pol activity by RT1t49 as measured by single nucleotide incorporation**

Single-stranded DNA aptamer RT1t49 has previously been shown to be an effective inhibitor of HIV-1 reverse transcriptase (9). In the present study, we set out to determine the sequence and structural constraints associated with this inhibition. We have previously shown that RNA-dependent DNA polymerization is more susceptible to aptamer inhibition when compared with DNA-dependent DNA polymerization, therefore we have chosen to evaluate the inhibition of our aptamer constructs using RNA-dependent DNA polymerization assays (12). Single nucleotide incorporation assays offer a convenient and readily quantifiable means of assessing DNA polymerization activity (12). For these experiments we used a Cy3-labeled 18-nucleotide DNA primer corresponding to the

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Viral isolate</th>
<th>IC(_{50}) (nM, Mean ± SD)</th>
<th>RT1t49</th>
<th>RT1t49(-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DDDP</td>
<td>DDDP(-5)</td>
</tr>
<tr>
<td>B</td>
<td>HXB2</td>
<td>9.1 ± 4.2</td>
<td>7.6 ± 3.7</td>
<td></td>
</tr>
<tr>
<td>A/D</td>
<td>92UG021</td>
<td>9.7 ± 2.3</td>
<td>9.7 ± 2.2</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>98CN008</td>
<td>4.0 ± 2.2</td>
<td>1.6 ± 0.6</td>
<td></td>
</tr>
<tr>
<td>A/E</td>
<td>93TH233.3</td>
<td>18.4 ± 7.5</td>
<td>25.3 ± 6.6</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>94CY017.41</td>
<td>12.0 ± 5.3</td>
<td>8.6 ± 2.9</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>RU570</td>
<td>16.9 ± 7.9</td>
<td>7.6 ± 2.8</td>
<td></td>
</tr>
<tr>
<td>CPZ-US</td>
<td></td>
<td>12.3 ± 5.0</td>
<td>8.7 ± 1.3</td>
<td></td>
</tr>
<tr>
<td>MVP5180</td>
<td></td>
<td>9.7 ± 3.9</td>
<td>7.9 ± 2.3</td>
<td></td>
</tr>
<tr>
<td>CPZ-TAN1</td>
<td></td>
<td>9.7 ± 4.4</td>
<td>5.5 ± 1.6</td>
<td></td>
</tr>
<tr>
<td>EHO-287</td>
<td></td>
<td>90 ± 36</td>
<td>110 ± 14</td>
<td></td>
</tr>
</tbody>
</table>

IC\(_{50}\) values derived from full extension of an 18 nt primer on a103 nt template.

---

Downloaded from https://academic.oup.com/nar/article/35/15/5039/1074181 by guest on 22 June 2022
independent experiments was used to determine the IC₅₀ value. 243 nM and a no aptamer control. (Figure 1) The calculated IC₅₀ of 3.6 nM was used throughout the rest of this study as the reference against which all others were compared.

Evidence supporting the possibility of multiple RT1t49 conformations

The MFOLD algorithm (28) on the Integrated DNA Technologies web server (www.idtdna.com) identified three potential secondary structures for RT1t49 (Figure 2). The most stable structure (ΔG = -4.7) is also the one proposed by Schneider et al. (9) (Figure 2A, hereon referred to as ‘structure A’). All three conformations share an identical stem I at the 5′ end, followed by a single-stranded connector and a variable stem II at the 3′ end. The nucleotides responsible for the alternative stem II structures are highlighted in Figure 2 according to base-pairing potential. Overall, the structure of stem II contributes significantly to the stability of RT1t49 as the reference IC₅₀ against which all others were compared.

Multiple conformations of stem II contribute to RT inhibition

Three separate mutations were introduced into stem II of RT1t49 (Figure 3). The first of these was designed to address the importance of the three-nucleotide loop at the end of stem II, or to a decrease in the ability of RT1t49 to assume alternative stem II structures, the terminal pair of stem II was changed from G–T to G–C Watson–Crick base pair. This mutation stabilized the proposed stem by −0.6 kcal/mol and had little consequence on the IC₅₀ (5.5 ± 1.4 nM). While T45 is not predicted to be base-paired in the alternate stem II

Figure 1. RT1t49 inhibition of single nucleotide incorporation by HIV-1 RT. (A) Schematic of the single nucleotide incorporation assay. Extension was performed using a Cy3-labeled 18 nt DNA primer annealed with an unlabeled 40 nt template. (B) Representative gel of a standard assay with final aptamer concentrations ranging from 0.3 to 243 nM and a no aptamer control. (C) The plot of normalized percent extension as a function of aptamer concentration from three independent experiments was used to determine the IC₅₀ value. Normalized data were fit to a sigmoidal dose–response curve using a constant slope. The calculated IC₅₀ of 3.6 ± 0.9 nM was used throughout the rest of this study as the reference against which all others were compared.

3′-end of human tRNA¹⁵yr,³ annealed to a 40-nucleotide DNA template containing the HIV-1 primer binding site (PBS) (Figure 1A). Annealed primer/template was present in excess of primer/template over RT to allow multiple-turnover conditions. Increasing amounts of inhibitor were added to the reaction mixture prior to the addition of RT. Fitting the data for the RT1t49 titration to a sigmoidal dose-response curve gives half-maximal inhibition (IC₅₀) at 3.6 ± 0.9 nM (Figure 1B and C). This equation assumes both a 1:1 aptamer:RT stoichiometry and that the aptamer is in excess of RT near the IC₅₀ value. Because the total RT concentration is comparable with this value of IC₅₀, these data were also fit to the quadratic form of the binding equation and to a sigmoidal dose–response curve with a variable slope function. For each fit, the reported IC₅₀ value fell within the error of the original fit (data not shown). Therefore, 3.6 nM was used throughout the rest of this study as the reference IC₅₀ against which all others were compared.

Evidence supporting the possibility of multiple RT1t49 conformations

The MFOLD algorithm (28) on the Integrated DNA Technologies web server (www.idtdna.com) identified three potential secondary structures for RT1t49 (Figure 2). The most stable structure (ΔG = -4.7) is also the one proposed by Schneider et al. (9) (Figure 2A, hereon referred to as ‘structure A’). All three conformations share an identical stem I at the 5′ end, followed by a single-stranded connector and a variable stem II at the 3′ end. The nucleotides responsible for the alternative stem II structures are highlighted in Figure 2 according to base-pairing potential. Overall, the structure of stem II contributes significantly to the stability of RT1t49 as the reference IC₅₀ against which all others were compared.

Multiple conformations of stem II contribute to RT inhibition

Three separate mutations were introduced into stem II of RT1t49 (Figure 3). The first of these was designed to address the importance of the three-nucleotide loop capping the end of stem II in structure A (Figure 3A). Converting the predicted tri-loop to an unrelated stable tetra-loop (AAC → TGCG) increased IC₅₀ nearly 20-fold to 65 ± 10 nM. To determine whether this perturbation was due to loss of specific contacts between RT and the tri-loop at the end of stem 2, or to a decrease in the ability of RT1t49 to assume alternative stem II structures, the terminal pair of stem II was changed from G–T to G–C Watson–Crick base pair. This mutation stabilized the proposed stem by −0.6 kcal/mol and had little consequence on the IC₅₀ (5.5 ± 1.4 nM). While T45 is not predicted to be base-paired in the alternate stem II
structures, the increased stability in structure A suggests a more-stable stem II is well tolerated. The third alteration disrupted the stem II in structure A by simultaneously introducing mutations G47A and G48A. Surprisingly, this mutation had no effect on the observed IC₅₀ (3.1 ± 0.5 nM).

The results described above suggest that either the stem II associated with structure A is not an important structural element for RT inhibition or that alternative structures play a role in the high-affinity inhibition by RT1t49 and can be accommodated by these mutations. To differentiate between these possibilities, a series of RT1t49 variants were created, each with successive single nucleotide deletions from the 3′-end until the entire stem II of structure A was completely deleted. Up to five nucleotides (−1 through −5) could be deleted (creating a 44-nt species terminating with C44) without significantly affecting the IC₅₀ values (Figure 3C), which ranged from 5.1 to 8.1 nM. Clearly RT1t49 variants are still capable of a high level of inhibition even when the stem II of structure A is unable to form. However, removal of one additional nucleotide (−6) caused a dramatic jump in IC₅₀ to 138 ± 68 nM. Further deletions of nucleotides −7 through −10 all resulted in IC₅₀ values between 215 and 285 nM, while additional deletions (−11 through −13) eliminated nearly all inhibition and yielded IC₅₀ values near 1 μM. These results strongly suggest that structures B or C (or both) contribute significantly to the inhibition of HIV-1 RT.

**The connector region and its role in RT inhibition**

Several mutations and deletions were made to address the importance of the length, sequence and single-stranded character of the short connector region separating the two stems of RT1t49 (Figure 4). Shortening this connector from six nucleotides to five by deleting A31 had only a minor effect, increasing IC₅₀ to 8.3 ± 2.1 nM. However, deletion of both A31 and A32 raised the overall IC₅₀ value nearly 40-fold to 127 ± 29 nM, and deletion of A31, A32 and A33 produced a 100-fold increase in IC₅₀ (361 ± 81 nM) relative to that of RT1t49. Deletion of nucleotides C34 and T35 separately increased IC₅₀ to 38 ± 18 nM and 58 ± 6 nM, respectively, but the simultaneous deletion of both nearly eliminated inhibitory function at all aptamer concentrations surveyed, yielding a calculated IC₅₀ that approached 1 μM. Several base-substitution mutations were introduced to test the importance of individual nucleotides within the connection region. Mutation C34A increased the IC₅₀ to 100 ± 35 nM, mutation T35A increased the IC₅₀ to

---

**Figure 3.** Mutations to stem II of RT1t49. Individual and combination mutations were created within stem II of structure A. (A) The mutation (A42T, A43G, +G) converted the tri-loop (AAC) to a stabilizing tetra-loop (TGCG) (35). The mutation T45C allows for G41 to base-pair more strongly with C45. Mutations G47A and G48A eliminated stem II in structure A. (B) Stepwise deletions were introduced at the 3′-end ranging from a single nucleotide (A49, deletion 1) to the entire stem II (deletion −13). Reactions using aptamers (−11) through (−13) did not reach 50% inhibition within the range of aptamer concentrations tested; therefore, IC₅₀ values were estimated from the curve fit by extrapolation beyond the data and are reported as >250 nM (calculated IC₅₀ values between 250 nM and 500 nM) and ≥250 (>500 nM). The dotted line represents the reference IC₅₀ of RT1t49. Similar denotations are used in subsequent figures.

**Figure 4.** Mutational analysis of the connector region. Deletions and mutations to the connector region of RT1t49 were analyzed for RT inhibition. The original connector, as defined in structure A, was replaced with the sequences shown below graph (5′-3′, left to right).
228 ± 96 nM and the combination mutation of C34A/T35A increased the IC50 to >500 nM. These observations clearly indicate that the single-stranded connector between stems I and II plays an important role in RT inhibition by RT1t49. Several individual nucleotides are implicated as either making individual, specific contacts with the protein or in forming the alternative stem II structures found in structures B and C.

Nucleotides A31–A33 are not predicted by MFOLD to be paired in any of the structures. These nucleotides may serve to establish the overall spacing of the two stems of the aptamer. Subsets of the nucleotides C34–T37 are base-paired in structures B and C. Mutations to both C34 and T35 have detrimental effects on the ability of RT1t49 to inhibit RT, consistent with at least two models for RT inhibition: one in which RT contacts these two nucleotides, and one in which they are involved in the alternative stem II conformation shown in structure B.

### Structural and sequence requirements for stem I

Three sets of mutations were introduced into RT1t49 to determine the consequences of weakening, strengthening or altering the sequence of stem I (Figure 5). The C4A and T11A/T12A mutations weakened the base-pairing potential of stem I, and caused a nearly 5-fold increase in the IC50 value relative to that of RT1t49 (IC50 = 17 ± 1.2 nM and 21 ± 3.3 nM, respectively.) Disruption of T8–A23 and G9–C22 bp near the center of stem I prevented formation of stem I altogether (as predicted using MFOLD) and yielded a more substantial loss of inhibition (IC50 = 70 ± 29 nM).

Mutations that increased stem I stability had only modest effects on RT inhibition. Three of the four mutants surveyed gave IC50 values within 2-fold of the parental RT1t49. Especially interesting is the absence of significant effect on IC50 for the two constructs that fully pair the internal bulges. Crystallographic analysis has shown that the primer-template binding cleft between the pol and RNase H active sites of RT can accommodate between 17 and 18 bp of dsDNA (29–32), which takes on an A-form/B-form mixed helix and has a bend of 40° to 45° at the junction between the A-form and B-form segments (31). It has been suggested that aptamers with pre-existing bends may be recognized more easily than a simple uninterrupted helix by RT (8,9). In particular, the unpaired nucleotides in stem I of RT1t49 have been proposed to serve this role (9). However, the results above suggest that the helical nature of stem I is more important for RT recognition than bending associated with the bulged positions within the stem.

To exclude the possibility of aptamer recognition through sequence-specific interactions, several mutations were introduced wherein nucleotides in one strand were switched with those in the other strand in various parts of stem I (Figure 5B and C). Inverting the base-paired region in the beginning segment of stem I (RT1t49-Inv A) resulted in a moderate increase in IC50 (16 ± 4.9 nM). When the middle segment of stem I was inverted (RT1t49-Inv B) there was only a small increase in IC50 (10 ± 3.2 nM). The two alterations were not additive,
as introduction of the two inversions simultaneously (RT1t49-Inv A + B) resulted in an IC$_{50}$ of 9.3 ± 2.6 nM. Furthermore, changing most of the nucleotides in both stems to other Watson–Crick pairs (variant RT1t49-Change Sequences) yielded an IC$_{50}$ value (8.8 ± 3.1 nM) that is in line with the other variants in this series. Thus, RT recognizes stem I of RT1t49 in a fashion that is predominantly independent of sequence.

Secondary structural determination by enzymatic probing

To define more precisely the secondary structure of the folded aptamer, two versions of RT1t49 were subjected to enzymatic digestion with S1 nuclease (Figure 6). For the full-length RT1t49, major cleavages are observed at nucleotides G5/G6 and from A31–G36, with secondary cleavages occurring at nucleotides C3, T12–T21, A40–G41 and A43–T45. These weaker cleavages could indicate that the dominant solution structure is in equilibrium with one or more minor, alternative folds. When higher molecular weight fractions of these same samples were separated by running the samples longer, no S1 cleavage was observed at nucleotides T46 through A49 of RT1t49, indicating that these nucleotides are base-paired or otherwise structured in the dominant conformation (data not shown). Thus, the S1 nuclease digestion of RT1t49 supports the structural prediction of structure A. For aptamer RT1t49(-5), which lacks the 3′-most 5 nucleotides, the cleavage patterns are essentially identical at G5/G6 and from A31–A36, although the secondary cleavages sites at nucleotides A17–T21 appear to be cleaved less efficiently than in RT1t49. Although both aptamers appear to form identical stem I and possess similar connection regions, the RT1t49(-5) deletion mutant lacks the capability of folding into a structure with the proposed stem II in structure A. Interestingly, the secondary cleavages observed beyond nucleotide T36 in RT1t49 are absent from the digestion pattern for RT1t49(-5). These data suggest that the 3′-terminal segment of RT1t49(-5) is base-paired or structured (as in structure B and C), and they support our hypothesis that alternative structures within stem II contribute to inhibition of HIV-1 RT.

RT1t49(-5) inhibits RT from HIV-1, HIV-2 and SIV$_{CPZ}$ RT

We have recently shown that RT1t49 inhibits pol and RNase H activities of RTs from diverse clades of HIV-1, HIV-2 and SIV$_{CPZ}$ (36). The present work demonstrates that multiple structures may contribute to inhibition of HXB2 RT, raising the possibility that multi-clade inhibition may have been dependent upon access to multiple structures. For example, some RTs may have been inhibited only by the dominant solution structure (structure A) while others were inhibited only by one of the alternatives (structure B or C). We therefore evaluated the effect of RT1t49(-5) on cross-clade inhibition of DNA-dependent DNA pol extension of an 18-nt primer on a 103-nt template. Inhibition by the full-length aptamer (RT1t49) is essentially identical to inhibition by RT1t49(-5) (Table I). The difference of the IC$_{50}$ values between pairs was not significant, and we found that RT’s from HIV-1 subtypes C and G were more than 2-fold more sensitive to inhibition by RT1t49(-5) than by RT1t49. These data support our hypothesis that the stem II associated with structure A is not essential for the overall inhibition of HIV-1 RT and that the alternative structures associated with stem II must contribute significantly to the inhibition of lentiviral RT.
Site-specifically generated hydroxyl radicals identify point of contact

Previously Götte et al. (26) showed the RNase H domain of HIV-1 RT could utilize two distinct mechanisms to cleave both the RNA template of DNA/RNA hybrids and dsDNA (26). Replacing the Mg(II) in the RNase H active site with Fe(II) allows the site-specific generation of OH⁻/C₁ radicals that cleave the closest nucleotide in the dsDNA. Building on this concept, we evaluated whether RT1t49 and RT1t49(-5) bound with HIV-1 RT could react with OH⁻ radicals generated from Fe(II) bound at the active sites of HIV-1 RT (Figure 7A). Both aptamers show uniform, low-level reactivity to OH⁻ radicals in the absence of protein (lanes 2 and 7). Upon the addition of wild-type RT, new major cleavages are observed at nucleotides 11 and 32 (lanes 3 and 8). To determine whether these sites of hyperactivity were due to Fe(II) bound in the pol or RNase H active sites, wild-type RT was replaced with mutants that selectively disrupt metal ion binding in either the pol (D185N or D186N) or the RNase H (E478Q) active sites. Although RNase H catalytic activity can be rescued in the last mutant in the presence of Mn(II) (33), this replacement abolishes RNase H-directed, Fe(II)-mediated radical cleavage of dsDNA (26). Cleavage at position 11 remained strong for all three mutants, supporting aptamer being bound to the RT mutants. No change in the cleavage pattern of either aptamer was observed for the RNase H mutant (data not shown), ruling out radical cleavage mediated by the RNase H active site. In contrast, cleavage associated

Figure 7. Site-specifically generated hydroxyl radical hyperactivity at the RT pol active site. (A) The cleavage patterns of uncomplexed aptamers RT1t49(-5) and RT1t49 exposed to OH⁻ radicals (lanes 2 and 7, respectively), or these same aptamers when bound with RT from HXB2 (lanes 3 and 8), with pol active-site mutant D185N (lanes 4 and 9) or with pol active-site mutant D186N (lanes 5 and 10). Lanes 1 and 6 show non-reacted aptamers. (B) Representative graph plotting the intensity of bands from RT1t49(-5) from nucleotides 30 to 35 for HXB2 (continuous line), D185N (dotted line) or D186N (dashed line). (C) The site of hypercleavage is mapped to the RT1t49 structure C.
with nucleotide A32 was essentially eliminated in both of the pol-active site mutants for both aptamers. A plot of the intensity of cleavage from nucleotides 30–34 shows a dramatic increase in cleavage of wild-type versus pol-active site mutants at A32 (Figure 7B). These results strongly suggest that A32 is located very near the pol-active site.

Hydroxyl radical footprinting is an established method for determining dsDNA binding domains and has been used to evaluate helicity of dsDNA (27). Plotting the intensity of bands from nucleotides 6 through 36 in these OH· radical probing experiments provide further insight into the structure of bound aptamer (Figure 8A and B). Cleavage levels are above background for two regions (nucleotides 10–15 and 22–27) when exposed to OH· radicals in the presence of protein. These nucleotides are located within stem I and are arranged such that these regions pair with (or are opposite of) regions found within the background intensity (Figure 8C). Overlaying these sites with a crystal structure of HIV-1 RT bound to a dsDNA primer-template [2HMI and 1RTD, (30,34)] show this difference in intensities can be explained by the formation of a double helix and that these regions would be solvent exposed if they bound RT as primer-template mimics (Figure 9A). Interestingly, this assignment places A32 within the dNTP binding cleft and in direct contact with the Mg(II) binding sites (Figure 9B), supporting our results of hyperactive OH· radical cleavage associated at the pol-active site. The lack of hyperactive OH· radical cleavage at the RNase H active site can be readily explained by this model, as stem I is too short to reach the RNase H active site when A32 is in the pol active site. Our data ascribe no specific structural role for the 3′ stem II, although a reasonable extrapolation is that it

Figure 8. Hydroxyl radical footprint of RT1t49 bound to HXB2 RT maps to a double helix structure. The pattern of cleavage from a representative OH· radical experiment (A) was plotted from nucleotides 3 through 35 (B). (C) Peaks were mapped to structure C based on intensity compared with background (black circles indicate 5-fold or greater enhanced cleavage, dark gray circles indicate enhanced cleavage of 2-fold, and light gray circles are between background and an enhancement of 2-fold.) Diagonal-patterned circles represent cleavage within the background noise, and unfilled circles represent nucleotides that could not be resolved using PAGE.
extends beyond the active site to the ‘back side’ of RT, interacting either with the fingers (in analogy with the 5’ unpaired template) or thumb domain.

Further modeling suggests RT1t49-resistant mutations N255D and N265D (15,16) would be in direct contact with nucleotides 27 and 5, respectively, and place these amino acids in a position to influence interactions with RT1t49. Mutation R277K has been shown to be resistant to Type I RNA pseudoknot aptamers but not to RT1t49 [(37), data not shown]. Based on our modeling, this mutation faces away from the helix and may have only minimal effect on RT1t49 binding (data not shown).

We believe understanding the molecular interactions between RT and RT1t49 may lead to new insights into the development of alternative therapeutics that can inhibit broadly across the primate lentiviral family.

CONCLUSION

While a substantial body of work has been directed towards understanding the essential structural features of RNA aptamers (8,10,17) and their interactions with RT, including a low-resolution crystal structure (18), there has been much less work devoted to the overall structural requirements and interactions of ssDNA aptamers with RT. In light of the broad, cross-clade RT inhibition that has only been observed for ssDNA aptamers, it is essential to define the structural elements that contribute to this broad inhibition. The present work details the inhibitory potency of 42 variants of ssDNA aptamer RT1t49 designed to test specific structural features of the aptamer. The results from this study support the hypothesis that HIV-1 RT recognizes the long, base-paired structure of stem I without substantial nucleotide specificity, a second base-paired region (stem II) at the 3’ end of the aptamer and a single-stranded connector region that links stems I and II. Additional stem II structures that form at the 3’-most nucleotides of the connection domain may be responsible for most or all of the observed inhibition. Hydroxyl radical footprinting and site-specifically-generated hyperactive OH·/C1-radical cleavage support our model of RT1t49 binding to the primer-template binding cleft as a TRTI and shed light on molecular contacts between aptamer and RT, including specific interactions between nucleotide position 32 and the RT pol-active site. These results advance our understanding of how HIV-1 RT recognizes RT1t49 and lay a solid foundation for establishing the requirements for cross-clade inhibitory ssDNA aptamers.

ACKNOWLEDGEMENTS

The authors would like to thank Matthias Götte for the RT active site mutants and Stefan Sarafianos for careful reading of the manuscript. This work was supported by NIH grant AI62513 to DHB and by a Milton Taylor
graduate fellowship in virology to JDK. Funding to pay the Open Access publication charges for this article was provided by NIH grant A162513 to DHB.

Conflict of interest statement. None declared.

REFERENCES


