Expression of an Mg$^{2+}$-Dependent HIV-1 RNase H Construct for Drug Screening$^\dagger$

Richard V. Farias,$^2$‡ Deborah A. Vargas,$^1$ Andres E. Castillo,$^1$ Beatriz Valenzuela,$^1$§ Marie L. Côté,$^2$¶ Monica J. Roth,$^2$ and Oscar Leon$^1$$^*$

Programa de Virología, Instituto de Ciencias Biomédicas, Facultad de Medicina Universidad de Chile, Independencia 1027, Santiago, Chile,$^1$ and Department of Biochemistry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, New Jersey 08854$^2$

Received 12 May 2011/Returned for modification 21 June 2011/Accepted 11 July 2011

A single polypeptide of the HIV-1 reverse transcriptase that reconstituted Mg$^{2+}$-dependent RNase H activity has been made. Using molecular modeling, the construct was designed to encode the p51 subunit joined by a linker to the thumb (T), connection (C), and RNase H (R) domains of p66. This p51-G-TCR construct was purified from the soluble fraction of an Escherichia coli strain, MIC2067(DE3), lacking endogenous RNase HI and HII. The p51-G-TCR RNase H construct displayed Mg$^{2+}$-dependent activity using a fluorescent nonspecific assay and showed the same cleavage pattern as HIV-1 reverse transcriptase (RT) on substrates that mimic the tRNA removal required for second-strand transfer reactions. The mutant E706Q (E478Q in RT) was purified under similar conditions and was not active. The RNase H of the p51-G-TCR RNase H construct and wild type HIV-1 RT had similar $K_m$'s for an RNA-DNA hybrid substrate and showed similar inhibition kinetics to two known inhibitors of the HIV-1 RT RNase H.

The HIV reverse transcriptase (RT) has been a major target for antiviral drug development. The enzyme contains three activities: RNA- and DNA-dependent DNA polymerase activities and RNase H activity. Currently, all clinically approved RT inhibitors target the polymerase function. RNase H cleaves RNA within a RNA-DNA hybrid and is essential for viral replication (7, 32, 33, 36). RNase H inhibitors are currently being developed as antiviral agents (35). Several classes of HIV RNase H inhibitors, which act through distinct mechanisms, have been identified, including tropolones (5), N-hydroxyimides (16, 24), dihydroxy benzoyl naphthyl hydrazones (DHBNH) (17), diketo acids (4, 38, 45), and dimeric lactones (11).

The HIV-1 RT is heterodimer consisting of a p66 subunit and a p51 subunit (2, 26). The RNase H (R) domain is unique to the C terminus of p66. The shape of the asymmetric heterodimer has been related to a right hand with subdomains fingers (F), palm (P), thumb (T), and connection (C), in addition to RNase H. RNase H cleavages have been further classified as polymerase dependent or polymerase independent, where polymerase-dependent cleavages are defined as those that are coupled approximately 18 nucleotides from the positioning of the 3’OH in the polymerase active site (14, 46).

Two structures of RNase H molecules complexed with RNA-DNA have been obtained. The first is with the Bacillus halodurans RNase H (31); the second is with HIV RT in complex with a noncleavable 3’ polyuridine tract (PPT) hybrid (37). Through structural and biochemical studies, limited regions of RT have been identified to be in close contact with nucleic acid substrates. Within RT, both DNA-DNA (19) and RNA-DNA (37) substrates display a bend of approximately 40° between the polymerase active site and the p66 thumb. With respect to RNase H, two specific regions are of interest. The first is the minor groove binding tract (MGBT) within the p66 thumb domain (residues 258 to 266 of helix H) (3, 27). The second is the RNase H primer grip domain, which contacts the DNA primer strand (37). Biochemical studies have suggested the possibility of alternative binding orientations of the RNA-DNA hybrid with RT (13). Cross-linking studies on an HIV-1 RNase H domain also agree with this hypothesis (15). Recently, single-molecule fluorescence resonance energy transfer was used to examine the interactions between RT and nucleic acid substrates in real time. These studies show that the RNA-DNA hybrid can bind in two orientations (28) and that the orientation of the RT depends on the composition of the substrate (RNA-DNA versus DNA-DNA) (1).

Previous attempts to express the HIV-1 RNase H as a catalytically active domain have been possible only in the presence of Mn$^{2+}$, and many of these attempts required an alternative substrate binding domain (16, 21, 39, 40, 42, 47). The role of metal binding in the RNase H domain of HIV has been a controversial area. Reports of one and two metal binding sites have been described both biochemically and structurally (10, 12, 29, 34). The presence of Mn$^{2+}$ has been shown to induce conformational changes within the RNase H as well as activate known RNase H catalytic site mutations (E478Q) (8, 15, 29, 34). In addition, many of the candidate drugs identified have the potential to bind to metals (5, 24, 38). It has been...
postulated that the Mn$^{2+}$-dependent activity might not reflect the catalytic activity of the native protein (9).

Mg$^{2+}$-dependent RNase H activity could be reconstituted using specific combinations of RT subdomains, with optimal activity obtained with p51 plus a TCR construct initiating at Q222 (40). In the study described in this report, this combination of p51 plus TCR has been generated within a single polypeptide (p51-G-TCR). Expression and purification of the p51-G-TCR RNase H construct indicated that Mg$^{2+}$-dependent RNase H activity was maintained. This HIV-1 RNase H construct has the potential to greatly assist in mechanistic, structural, as well as drug discovery studies aimed at targeting the RNase H function of HIV-1 RT.

**MATERIALS AND METHODS**

**Materials.** pRT4, encoding HIV-IIIB HXB2 p66, was a gift of Bradley Preston, University of Washington, Seattle, WA. HIV-1 RT was purchased from Calbiochem, and Escherichia coli RNase H was purchased from New England BioLabs. The E. coli strain MIC2067(DE3) (F+ mbi A339:cat mbi B716:lam) was a gift of Shigenori Kanaya of the Osaka University Graduate School of Engineering, Department of Material and Life Sciences.

**Generation of p51-linker-TCR.** The region encoding p51 was PCR amplified using pRT4 as a template, 5’ EX primer (5’-GGAGATTCCTATGGGATAGATACTGA-3’) incorporating an NdeI restriction site (underlined), 3’ EX primer (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCG-3’) incorporating a BamHI restriction site (underlined), and an inserted stop (TAA) codon. The 1,344-bp product was digested with NdeI and BamHI, extracted with a QiAquick gel extraction kit (Qiagen), and ligated into the 6HisT-pET11 vector, yielding the p51-G-TCR construct has the potential to greatly assist in mechanistic, structural, as well as drug discovery studies aimed at targeting the RNase H function of HIV-1 RT.

**RESULTS**

**Incorporation of linker region E.** The linker region E was made by overlapping PCR. Codons were optimized for expression in E. coli. A 10:1 mixture of Taq and Pfu (Invitrogen) was used for the PCR amplification. DNA primers for construct E included unit A (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), unit B (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), and unit C (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’). Individual PCR products included unit A-G1 and G2-unit B, yielding AE1 and E2B, respectively. The products AE1 and E2B were denatured, annealed at 57°C, and extended for 8 cycles prior to the addition of primers LEXE (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), LEXS (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), and LEXE (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), and unit B (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), and unit B (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’), and unit B (5’-GGGGTACCTGTTAAGAGGGTGTCTCTTATTTGCGA-3’).

Individual PCR products were ligated in a 1:1 ratio of p51 and BamHI, extracted with a QiAquick gel extraction kit (Qiagen), and ligated into the 6HisT-pET11 vector, yielding the pHTPS1.

**RNase H fluorescence assay.** The RNase H fluorescence assay was carried out essentially as previously described (35). The reaction was carried out at 37°C in a 96-well plate in a final volume of 100 μl. The reaction mixture contained 0.25 μM DNA-RNA substrate, 50 mM Tris, pH 8.0, 60 mM KCl, and 5 mM MgCl$_2$. RNase H activity was detected by the increase of fluorescence at 525 nm due to substrate hydrolysis using an excitation wavelength of 490 nm. Fluorescence was measured using a Fluoroskan II apparatus, and data analysis was carried out using the program Spectrosoft (MTX Lab Systems).

**RNase H inhibition assays.** To determine the susceptibility of RNase H construct to known HIV-1 RNase H inhibitors, we tested the tropane derivative β-thujaplicin over a concentration range from 6.9 × 10$^{-8}$ M to 6.9 × 10$^{-4}$ M and an N-hydroxymide, ARK-2456, over a concentration range of from 7.05 × 10$^{-10}$ M to 7.05 × 10$^{-4}$ M. Both compounds were dissolved in dimethyl sulfoxide (DMSO) and added to the RNase H reaction mixtures. The DMSO concentration in the assay was 5%. The fluorescence assay was used under the conditions described above. Fifty percent inhibitory concentrations (IC$_{50}$) were determined by nonlinear regression using GraphPad Prism software, version 5. β-Thujapli- cin was obtained from Molecular Diversity Preservation International (Basel, Switzerland), and ARK-2456 was obtained from the Florida Center for Heterocyclic Compounds.

**iRNA synthesis.** The activity of the collected fractions was determined in a reaction (10 μl) containing 50 mM Tris-HCl, 50 mM KCl, 8 mM MgCl$_2$, and 1 mM MnCl$_2$, 1 μM of substrate, and 1 to 2 μM of the enzyme. After 30 min at 37°C, 1 μl of 10-fold diluted stop buffer was added and the products of the reaction were separated by electrophoresis on a 20% acrylamide–7 M urea gel.
RESULTS

An important tool for studying the HIV RT RNase H as well as for drug screening is an RNase H construct that maintains the specificity and biochemical properties of the wild-type (WT) enzyme. Reconstitution studies using reverse transcriptase subdomains had indicated that Mg\(^{2+}\)-dependent RNase H activity could be robustly restored in the presence of two components: p51 plus a p66 C-terminal construct encoding TCR (Q222) (40). Molecular modeling was utilized to design constructs which incorporated this combination of RT subdomains within a single polypeptide.

**Design criteria.** Figure 1 outlines the constructs that were developed and that joined the p51 polypeptide with the p66 TCR domains through a linker region. Deletion and biochemical studies of p51 had indicated that the C terminus influences RNase H cleavage and RT properties (6). The C terminus of p51 terminates on the back face of the protein, on the opposite side of the polymerase and RNase H active sites. This greatly facilitated the modeling of the linker, as it avoided generating a linker that would interfere with the pathway of substrate binding. The linker region between the p66 thumb and p51 connection spans a direct linear distance of 62 Å. The linker was designed to maintain a solvent-accessible hydrophilic face to improve solubility. Linker G consists of 33 amino acids derived from residues 85 to 102 of HIV-1 p66, a flexible region. Codons for the linker were optimized for expression in E. coli, and the construct contains an N-terminal His\(_6\) tag, followed by the thrombin cleavage site. Linker E is 2 amino acids shorter than linker G.

**Molecular modeling and surface potential of p51-linker-TCR constructs.** Figure 2B compares the molecular model of the p51-E-TCR constructs with that of the WT HIV-1 RT heterodimer (Fig. 2A). In the p51-E-TCR model, the linker is flexible and the side chains of the amino acids have no interaction with any of the residues of the p51-TCR backbone.

Of key importance was whether the RNase H constructs maintained surface charges that mimicked the WT protein. To analyze this, the surface potential of the p51-E-TCR (Fig. 2D) construct was compared with that of the WT HIV-1 RT heterodimer (Fig. 2C). The surface potential of the p51-E-TCR construct mimics rather well the homologous region of the HIV heterodimer. Of interest is the deep, highly positive region (blue) prior to the RNase H, indicating that the primer grip and minor groove binding track of the p66 thumb are likely to be maintained. Changes in the individual construct compared with the WT RT are due to the energy minimizations performed on each molecular model. In addition, docking experiments with a 23-mer–22-mer RNA-DNA hybrid

![FIG. 1. Schematic diagram of single-polypeptide HIV-1 RNase H constructs. (A) The structure of the HIV-1 reverse transcriptase p66/p51 heterodimer is shown on the top, where subdomains are labeled F (fingers), P (palm) T (thumb), C (connection), and R (RNase H). The boundaries of the subdomains are marked above the sequence and are based on HXB2. Domains of p66 and p51 are color coded, where p51 FP domains are gray, p66 FP domains are white, the p66 thumb domain is blue, the p51 thumb domain is turquoise, the p66 connection is yellow, the p51 connection is orange, and RNase H is green. The general scheme for the p51-linker-TCR construct is outlined at the bottom using the color scheme defined for p66/p51. The construct encodes an N-terminal His\(_6\) tag, followed by a thrombin cleavage site. Key restriction sites used in the assembly of the DNA components are indicated in their relative coding positions. (B) Sequences of linkers E and G. The sequence of HXB2 from position Y427 is included. The linker sequences are positioned between p51 F440 and p66 P247 and are indicated in bold.

![FIG. 2. Molecular modeling and surface potential of p51-linker-TCR RNase H constructs. (A) Ribbon diagram of WT HIV-1 RT (1RTD); (B) ribbon diagram of p51-E-TCR. The coloring of the domains follows that defined in the legend to Fig. 1; the linker region is indicated in purple. (C) 2D Grasp analysis of the surface charge potentials, where blue is a positive charge and red is a negative charge. (C) WT HIV-1 RT; (D) p51-E-TCR.]
(1HYS) indicated that the substrate would maintain access to the MGBP of the p66 thumb as well as the RNase H primer grip (data not shown).

**Complementation studies.** *E. coli* strain MIC2067(DE3) (F\(^{−}\) rnhA339::cat rnhB716::kam) lacks RNase H1 and HII and displays temperature-sensitive (ts) growth at 42°C that can be relieved by the introduction of plasmids carrying an RNase H gene. The ability of the p51-E-TCR construct (Fig. 1) to complement MIC2067 growth at 30°C and 42°C was tested (Fig. 3). At 30°C (Fig. 3B), growth was detected in all strains, as predicted. At 42°C (Fig. 3C), the growth of BL21(DE3) cells was normal; however, MIC2067(DE3) was ts for growth. With introduction of HIV p51, lacking the RNase H domain, the MIC2067(DE3) remained ts. Considerable growth was observed when the p51-E-TCR construct was introduced into the MIC2067(DE3) cells. This indicates that the construct provides sufficient RNase H activity to complement *E. coli* growth at 42°C. These results are extremely supportive of the suggestion that molecularly designed HIV-1 RNase H single polypeptide constructs provide a soluble, functional RNase H.

**Protein expression.** Plasmid encoding p51-E-TCR was assembled through a generalized scheme. Briefly, the linker regions were assembled using overlapping PCR and inserted into the p51 construct through the KpnI-BamHI sites (Fig. 1A). The TCR region was then inserted into linker-positive isolates through the PvuII-BamHI sites, generating p51-E-TCR. p51-G-TCR was obtained by adding 2 amino acids in the linker by site-directed mutagenesis (Fig. 1B). Plasmids were introduced into *E. coli* MIC6067(DE3) and induced for expression by isopropyl-/H11001/D-thiogalactopyranoside. Initially, purification of the constructs indicated that p51-G-TCR was more stable than p51-E-TCR (see Fig. S1 in the supplemental material); therefore, construct p51-G-TCR was used in the following studies.

Whole-cell extracts of individual colonies from the control pET15B plasmid, p51-G-TCR, and p51-G-TCR E706Q were analyzed on SDS-polyacrylamide gels by Coomassie blue staining and Western blotting using a His\(_6\) tag detection reagent. Comparison of the control pET15B extract with the p51-G-TCR and p51-G-TCR E706Q extracts showed a protein product which corresponded with the predicted full-length 93-kDa protein. This observation was confirmed by Western blotting (see Fig. S2 in the supplemental material).

**Protein purification.** MIC2067(DE3) cells harboring a plasmid containing the p51-G-TCR construct were autoinduced (43) to increase production of soluble proteins. Furthermore, to make sure that the RNase H activity is due to the expression of the p51-G-TCR protein, we replaced the amino acid residue E706 in p51-G-TCR that corresponds to E478 in HIV-1 RT with glutamine. E478Q HIV-1 RT lacks RNase H activity in Mg\(^{2+}\) (8). p51-G-TCR and p51-G-TCR E478Q were purified from the soluble fraction and chromatographed on nickel nitriiotriacetic acid (Ni\(^{2+}\)-NTA) resin. The purity of the proteins was greater than 90%, as judged by SDS-PAGE (see Fig. S3 in the supplemental material). Analysis of the p51-G-TCR protein solutions at 0.25 and 2.0 mg/ml by dynamic light scattering indicated that more than 98% of the protein is a monomer (results not shown).

**Characterization of magnesium-dependent RNase H activity of p51-G-TCR.** The RNase H activity of the RNase H constructs was determined in the presence of Mg\(^{2+}\) by the fluorescence assay as described in Materials and Methods. As shown in Fig. 4, the RNase H activity of p51-G-TCR was similar to that of HIV-1 RT, whereas E706Q was inactive, as expected, ruling out potential contamination with *E. coli* RNase H. To compare the kinetic properties of the p51-G-TCR protein and HIV-1 RT, we determined the \(K_m\) using the...
fluorescence assay. Under our conditions, the $K_m$ for the RNase H activity of p51-G-TCR and HIV-1 RT were 0.036 ± 0.020 and 0.025 ± 0.012 μM, respectively, suggesting that the structure of the substrate binding site was maintained in p51-G-TCR.

**tRNA removal assay.** Fractions eluting from the Ni$^{2+}$-NTA column were assayed for Mg$^{2+}$- and Mn$^{2+}$-dependent RNase H activity using a specific cleavage reaction. Figure 5B compares the activities of *E. coli* RNase H, HIV-1 RT, and p51-G-TCR on the RNA-DNA substrate shown in Fig. 5A. The specific cleavage patterns of the substrate that were detected for HIV-RT and p51-G-TCR in Mg$^{2+}$ were distinct from the cleavage pattern of *E. coli* RNase H. The cleavage patterns for the HIV-1 RT and p51-G-TCR construct correspond to the release of the 11-mer RNA, corresponding with the authentic cleavage identified *in vivo* and *in vitro* (39). (B and C) Time course of RNase H cleavages using the substrate defined in panel A. (B) RNase H activity in 10 mM Mg$^{2+}$. Results obtained with 2 pmol of either the *E. coli* RNase H, HIV-1 RT, or p51-G-TCR construct are shown. Times (t) are indicated. The position of the initial cleavage product is indicated. (C) RNase H activity of p51-G-TCR in the presence of 8 mM Mn$^{2+}$.

**DISCUSSION**

This study defines the first HIV-1-based construct to display Mg$^{2+}$-dependent RNase H activity in the absence of a polymerase domain. Multiple hypotheses as to why the minimal isolated HIV-1 RNase H constructs do not maintain RNase H activity in Mg$^{2+}$ have previously been proposed. These include limitations in the substrate binding, the presence of disordered loops, extensive mobility and flexibility of the domain, and metal-induced conformations. Within isolated HIV-1 RNase H domains, cross-linking studies have indicated that the Mn$^{2+}$ plays a structural role as well as a catalytic role within RNase H. Nuclear magnetic resonance analysis of the isolated domain...
indicated that divalent metals stabilize the structure of the isolated RNase H domain. Incorporation of the p51 thumb and the substrate binding domain of the two connections (TCCR) also was not able to sustain Mg\(^{2+}\)-dependent RNase H activity. Activity remained Mn\(^{2+}\)-dependent (47).

The p51-G-TCR Mg\(^{2+}\)-dependent RNase H construct differs from the prior TCCR construct by incorporating the p66 thumb and MGBT as well as the p51 finger-palm domain. Complementation studies (40) indicated that restoration of Mg\(^{2+}\) activity required the presence of subdomains of p51 and the T subdomain of p66. Accordingly, our results with the p51-G-TCR construct confirm these observations, indicating that the RNase H active site is stabilized by these RT subdomains, providing the conformation needed for catalysis with Mg\(^{2+}\). The inclusion of the p66 thumb and MGBT assists with the positioning and binding of substrate. Motifs within the finger-palm have been identified to be involved in RNA-DNA binding, in particular, p51 Lys 22, which is reported to interact only with the RNA-DNA PPT duplex and not DNA-DNA. Mutants with mutations within the MGBT have shown differential utilization of DNA and RNA templates and have been shown to function as protein sensors of the substrate minor groove.

Nonnucleoside RT inhibitors (NNRTIs) bind to the base of the thumb and have been shown to enhance RNase H cleavage and increase the angle between the polymerase active site-NNRTI binding site and the RNase H active site by 10°. In the RT-PPT crystal structure, it was noted that minor alterations in the substrate trajectory of 2 Å could reposition an RNA-DNA hybrid out of the RNase H catalytic core. Thus, the overall positioning of the substrate within the RNase H active site appears to be critical for the maintenance of Mg\(^{2+}\)-dependent RNase H activity.

The mechanism by which the substrate binding regulates the reverse transcriptase activity remains unclear; apparently, the activities of RT, DNA synthesis, and RNA hydrolysis are determined by its binding orientation on substrates. It has been reported that RT adopted opposite binding orientations on duplexes containing DNA or RNA primers (1, 13). The RNase H construct that has Mg\(^{2+}\)-dependent activity in the absence of polymerization will be useful to determine the substrate orientation during tRNA removal.

The RNase H activity within the viral RT is being developed as a target for drug development. Initial screens have utilized the full-length RT. Due to the complexity of the enzymes, inhibitors can affect RNase H activity directly through the RNase H active site or indirectly through positioning of the RNA-DNA template. The use of truncated RT constructs lacking the polymerase active site and NNRTI binding sites could assist in localizing the effects of lead compounds in secondary screens. The crystal structures of the complexes of HIV-1 RT with β-thujaplicinol (18), a naphthyridine derivative (44), and pyrimidol carboxylic acids (23) have been solved. The p51-G-TCR construct could be used for similar structural analysis of RNase H inhibitors in the absence of a polymerase active site.

### TABLE 1. IC\(_{50}\)s for RNase H activity of p51-G-TCR and HIV-1 RT

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>IC(_{50}) (M) ± SD*</th>
<th>p51-G-TCR</th>
<th>HIV-1 RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-Thujaplicinol</td>
<td>1.76 × 10^{-7} ± 0.086 × 10^{-7}</td>
<td>2.57 × 10^{-7} ± 0.027 × 10^{-7}</td>
<td></td>
</tr>
<tr>
<td>ARK-2456</td>
<td>1.86 × 10^{-7} ± 0.023 × 10^{-7}</td>
<td>4.70 × 10^{-7} ± 0.026 × 10^{-7}</td>
<td></td>
</tr>
</tbody>
</table>

*Data are the means ± standard deviations for four independent experiments.
In conclusion, the p51-G-TCR construct should be valuable for drug screening and a useful model for basic studies on RNA-DNA recognition and polymerase-independent RNase H cleavage.

ACKNOWLEDGMENTS

We thank Gaurav Patel for generating pTCR.

This work was supported by grant FONDECYT 1080137, awarded to O.L., and National Institutes of Health grant R01 GM070837, awarded to M.J.R. A Conicyt fellowship was awarded to D.A.V.

REFERENCES


