Available online at www.sciencedirect.com
ScienceDirect
Virus Research 134 (2008) 86–103
Virus Research
www.elsevier.com/locate/virusres

# RNase H activity: Structure, specificity, and function in reverse transcription

Sharon J. Schultz, James J. Champoux *

*Department of Microbiology, School of Medicine, Box 357242, University of Washington, Seattle, WA 98195, USA*

Available online 7 February 2008

**Abstract**

This review compares the well-studied RNase H activities of human immunodeficiency virus, type 1 (HIV-1) and Moloney murine leukemia virus (MoMLV) reverse transcriptases. The RNase H domains of HIV-1 and MoMLV are structurally very similar, with functions assigned to conserved subregions like the RNase H primer grip and the connection subdomain, as well as to distinct features like the C-helix and loop in MoMLV RNase H. Like cellular RNases H, catalysis by the retroviral enzymes appears to involve a two-metal ion mechanism. Unlike cellular RNases H, the retroviral RNases H display three different modes of cleavage: internal, DNA 3′ end-directed, and RNA 5′ end-directed. All three modes of cleavage appear to have roles in reverse transcription. Nucleotide sequence is an important determinant of cleavage specificity with both enzymes exhibiting a preference for specific nucleotides at discrete positions flanking an internal cleavage site as well as during tRNA primer removal and plus-strand primer generation. RNA 5′ end-directed and DNA 3′ end-directed cleavages show similar sequence preferences at the positions closest to a cleavage site. A model for how RNase H selects cleavage sites is presented that incorporates both sequence preferences and the concept of a defined window for allowable cleavage from a recessed end. Finally, the RNase H activity of HIV-1 is considered as a target for anti-virals as well as a participant in drug resistance.


*Keywords:* RNase H; Reverse transcriptase; Human immunodeficiency virus, type 1 (HIV-1); Moloney murine leukemia virus (MoMLV); Reverse transcription; Polypurine tract (PPT)

## 1. Introduction

Reverse transcription is performed by the retroviral enzyme called reverse transcriptase. This multifunctional enzyme carries out RNA-dependent DNA polymerization, DNA-dependent DNA polymerization, strand displacement synthesis, strand transfers, and degrades the RNA strand in RNA/DNA hybrids. To perform these diverse functions, reverse transcriptase uniquely combines two distinct enzymatic activities, a DNA polymerase activity that uses RNA or DNA as a template, and an RNase H activity that cleaves the RNA strand of an RNA/DNA hybrid (Gilboa et al., 1979). These activities are localized in two separate protein domains. The polymerase domain comprises the N-terminal two-thirds of reverse transcriptase, while the RNase H domain is the C-terminal one-third (reviewed in Champoux, 1993; Telesnitsky and Goff, 1993a). Mutations that inactivate the functions of either domain result in a retrovirus incapable of replication (Tanese and Goff, 1988; Telesnitsky and Goff, 1993b).

The utilization of an RNase H activity during retroviral replication represents a unique strategy to copy a single-stranded RNA genome into a double-stranded DNA, in part since the minus-strand DNA remains base-paired to the retrovirus genome in the first cycle of DNA synthesis. The RNase H activity is essential in several aspects of reverse transcription (reviewed in Champoux, 1993; Arts and Le Grice, 1998; Rausch and Le Grice, 2004). This review examines and compares the structures, activities, and functions of the human immunodeficiency virus, type 1 (HIV-1) and Moloney murine leukemia virus (MoMLV) RNases H, and considers the prospects of targeting the RNase H activity of HIV-1 by anti-viral drugs, as well as the role of RNase H in drug resistance.

## 2. RNase H in reverse transcription

Polymerization-dependent RNase H activity occurs during minus-strand synthesis and initiates the degradation of the RNA genome. Secondary structures in the template such as hairpins

* Corresponding author. Tel.: +1 206 543 8574; fax: +1 206 543 8297.
*E-mail address:* champoux@u.washington.edu (J.J. Champoux).

0168-1702/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2007.12.007

can cause pausing by reverse transcriptase, and such pauses promote RNase H cleavages and facilitate strand transfers (Luo and Taylor, 1990; DeStefano et al., 1992, 1994a; Peliska and Benkovic, 1992; Lanciault and Champoux, 2006). However, the polymerization rate of reverse transcriptase is greater than the hydrolysis rate of the enzyme and the polymerization-dependent RNase H activity is insufficient to completely degrade the genomic template (DeStefano et al., 1991b, 1994b; Dudding et al., 1991; Kati et al., 1992; Götte et al., 1995; Kelleher and Champoux, 2000).

The polymerization-independent RNase H activity participates in removal of genomic RNA that remains annealed to the minus-strand DNA. This synthesis-independent degradation of the RNA genome, along with displacement synthesis, enables efficient synthesis of the plus-strand DNA (DeStefano et al., 1991b, 1992, 1994b; Gopalakrishnan et al., 1992; Fuentes et al., 1996; Smith et al., 1999; Kelleher and Champoux, 2000; Schultz et al., 2004). Polymerization-independent cleavages by RNase H can also generate the polypurine tract (PPT) primer and remove the extended tRNA and PPT primers. Numerous studies have shown that these cleavages require recognition of specific sequences by the enzyme (for example, see Finston and Champoux, 1984; Rattray and Champoux, 1987; Huber and Richardson, 1990; Luo et al., 1990; Pullen and Champoux, 1990; Furfine and Reardon, 1991a; Pullen et al., 1992; Smith and Roth, 1992; Fuentes et al., 1995; Schultz et al., 1995; Powell and Levin, 1996). In addition, the RNase H activity assists the initiation of plus-strand synthesis, since the PPT primer is not extended efficiently unless a gap is generated by cleavage at several sites that are located downstream of the PPT (Schultz et al., 2003).

The activity and specificity of RNase H must be finely tuned during reverse transcription (for example, see Kotewicz et al., 1988; Julias et al., 2002; Rausch et al., 2002; Purohit et al., 2005). Excessive uncontrolled degradation of the template RNA could cause the primer strand to dissociate from the template strand and terminate synthesis. Insufficient RNase H activity could allow remaining RNA fragments to slow synthesis of the plus-strand DNA, or interfere with specific cleavages like PPT primer generation or primer removal to generate improper LTR ends. Because of its multiple essential roles in reverse transcription, the RNase H activity of HIV-1 reverse transcriptase is an excellent target for anti-virals.

## 3. Structure of HIV-1 and MoMLV RNase H domains

While the first reverse transcriptase to be purified and biochemically characterized was avian (Mölling et al., 1971), the majority of structure-function studies have focused on the reverse transcriptases of mammalian retroviruses. Of these, the two most prominently studied enzymes are the reverse transcriptases of HIV-1 and MoMLV. HIV-1 is heterodimer of p66 and p51 (Di Marzo Veronese et al., 1986). The p51 subunit is proteolytically derived from the p66 subunit and lacks the C-terminus containing the RNase H domain. MoMLV reverse transcriptase is found as a 76 kDa monomer (Moelling, 1974; Roth et al., 1985; Das and Georgiadis, 2004). Co-crystal structures of HIV-1 reverse transcriptase with duplex substrate have indicated

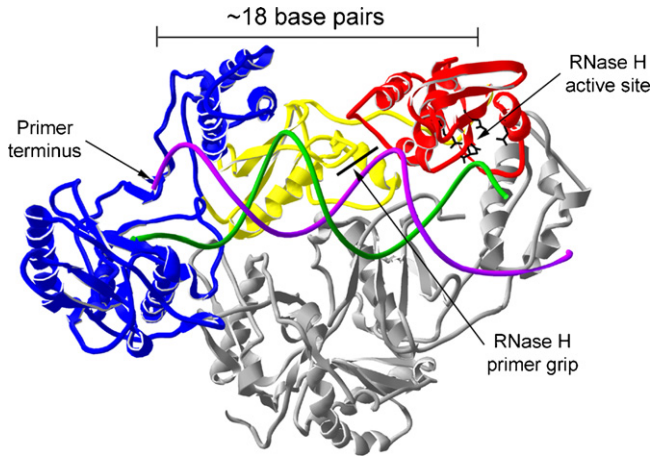


Fig. 1. Co-crystal structure of HIV-1 reverse transcriptase and an RNA/DNA substrate. A ribbon diagram of the HIV-1 reverse transcriptase associated with a PPT RNA/DNA hybrid, based on the report by Sarafianos et al. (2001) (pdb entry 1HYS). The polymerase, connection, and RNase H domains of p66 are shown in blue, yellow, and red, respectively; p51 is shown in gray. The RNA template and DNA primer strands are shown in green and purple, respectively. The location of the RNase H active site is indicated with the four key catalytic residues shown as ball and stick structures (Asp443, Glu478, Asp498 and Asp549). The approximate locations of the primer terminus and the RNase H primer grip are also indicated, as is the approximate distance in base pairs between the polymerase and RNase H active sites on the hybrid substrate. Diagram generated with Swiss PDB Viewer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

that the distance between the polymerase and RNase H active sites is 17 nucleotides for the DNA/DNA substrate (Jacobo-Molina et al., 1993; Huang et al., 1998) and 18 nucleotides for the RNA/DNA hybrid (Sarafianos et al., 2001). The substrate-dependent difference in the distances between the active sites of reverse transcriptase is likely nucleic acid-dependent, which matches with biochemical observations (Götte et al., 1998). Fig. 1 shows the co-crystal structure of HIV-1 reverse transcriptase with a PPT-containing RNA/DNA primer template (Sarafianos et al., 2001).

The RNases H of HIV-1 and MoMLV, both alone as a domain and in the holoenzyme of reverse transcriptase, exhibit very comparable tertiary folding (Davies et al., 1991; Kohlstaedt et al., 1992; Lim et al., 2006). An excellent comparison of the three-dimensional structures of these two RNase H domains is presented in an accompanying article of this special issue (see panel B, Fig. 6, in Coté and Roth, 2008). Also, comparisons to the structurally similar *Escherichia coli*, *Bacillus halodurans*, and most recently, human RNases H have proven highly informative regarding the structures and functions of the HIV-1 and MoMLV RNase H domains (Yang et al., 1990; Jacobo-Molina et al., 1993; Huang et al., 1998; Sarafianos et al., 2001; Das and Georgiadis, 2004; Nowotny et al., 2005, 2007; Lim et al., 2006). The MoMLV RNase H domain consists of 5 β-strands and 5 α-helices, whereas the HIV-1 RNase H domain lacks one of the α-helices (the positively charged C-helix; see Section 3.2), but is otherwise similar. Notably, the MoMLV RNase H has a network of hydrophobic residues that interacts in the N-terminal portion of the enzyme while the RNases H of HIV-1 and *E. coli* have less hydrophobic residues in this region and are