doi:10.1016/j.jmb.2011.04.024 J. Mol. Biol. (2011) 410, 863–874

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Secondary Structure of the HIV Reverse Transcription Initiation Complex by NMR

Elisabetta Viani Puglisi* and Joseph D. Puglisi

Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305-5126, USA

Received 18 February 2011; received in revised form 6 April 2011; accepted 11 April 2011

Edited by M. F. Summers

Keywords:
HIV;
reverse transcriptase;
initiation complex;
RNA;
NMR

Initiation of reverse transcription of genomic RNA is a key early step in replication of the human immunodeficiency virus (HIV) upon infection of a host cell. Viral reverse transcriptase initiates from a specific RNA–RNA complex formed between a host transfer RNA ($tRNA^{Lys}_3$) and a region at the 5′ end of genomic RNA; the 3′ end of the tRNA acts as a primer for reverse transcription of genomic RNA. We report here the secondary structure of the HIV genomic RNA–human $tRNA^{Lys}_3$ initiation complex using heteronuclear nuclear magnetic resonance methods. We show that both RNAs undergo large-scale conformational changes upon complex formation. Formation of the 18-bp primer helix with the 3′ end of $tRNA^{Lys}_3$ drives large conformational rearrangements of the tRNA at the 5′ end while maintaining the anticodon loop for potential loop–loop interactions. HIV RNA forms an intramolecular helix adjacent to the intermolecular primer helix. This helix, which must be broken by reverse transcription, likely acts as a kinetic block to reverse transcription.

Introduction

RNA structure guides key functions in human immunodeficiency virus (HIV) infection and replication. The viral genome is RNA, and RNA–RNA and RNA–protein interactions pervade the viral life cycle. The first step in viral replication after entry into an infected cell is reverse transcription, which is catalyzed by a viral enzyme, reverse transcriptase (RT).[1] RT must initiate from a specific RNA assembly and then navigate the complex secondary structure of genomic RNA while copying RNA to DNA.[2] RT is a major target of therapeutic intervention in treatment of acquired immune deficiency syndrome.

HIV RT is a virally encoded RNA/DNA-directed DNA polymerase, which consists of two subunits, p66 and p51. RT has both polymerase activity, as well as an RNase H domain, which cleaves RNA–DNA hybrids. The general mechanism of reverse transcription has been determined.[3] RT initiates reverse transcription from a specific RNA complex formed between a host $tRNA^{Lys}_3$[4,5] and a specific region [primer binding site (PBS)] of the HIV genome near the 5′ end of HIV RNA[6–8] (shown schematically in Fig. 1a). This initiation complex is preassembled in the virion, and initiation occurs by the addition of deoxynucleotides to the free 3′ OH of the $tRNA^{Ly3}_3$ (nucleotides).[9,10] RT copies the 5′ end of the viral RNA into DNA, using its RNase H activity to digest the RNA template. The DNA product is complementary to the 3′ long terminal repeat region of genomic RNA; interaction of the two strands allows reverse transcription to continue through the viral RNA 3′ to 5′. Short stretches of genomic RNA that are resistant to RNase H cleavage remain, which then prime synthesis by RT of positive-strand DNA. Complementary regions of the single-stranded

*Corresponding author. E-mail address: epuglisi@stanford.edu.

Abbreviations used: HIV, human immunodeficiency virus; RT, reverse transcriptase; PBS, primer binding site; TROSY, transverse relaxation optimized spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; NOE, nuclear Overhauser enhancement; COSY, correlated spectroscopy; DMS, dimethyl sulfate.

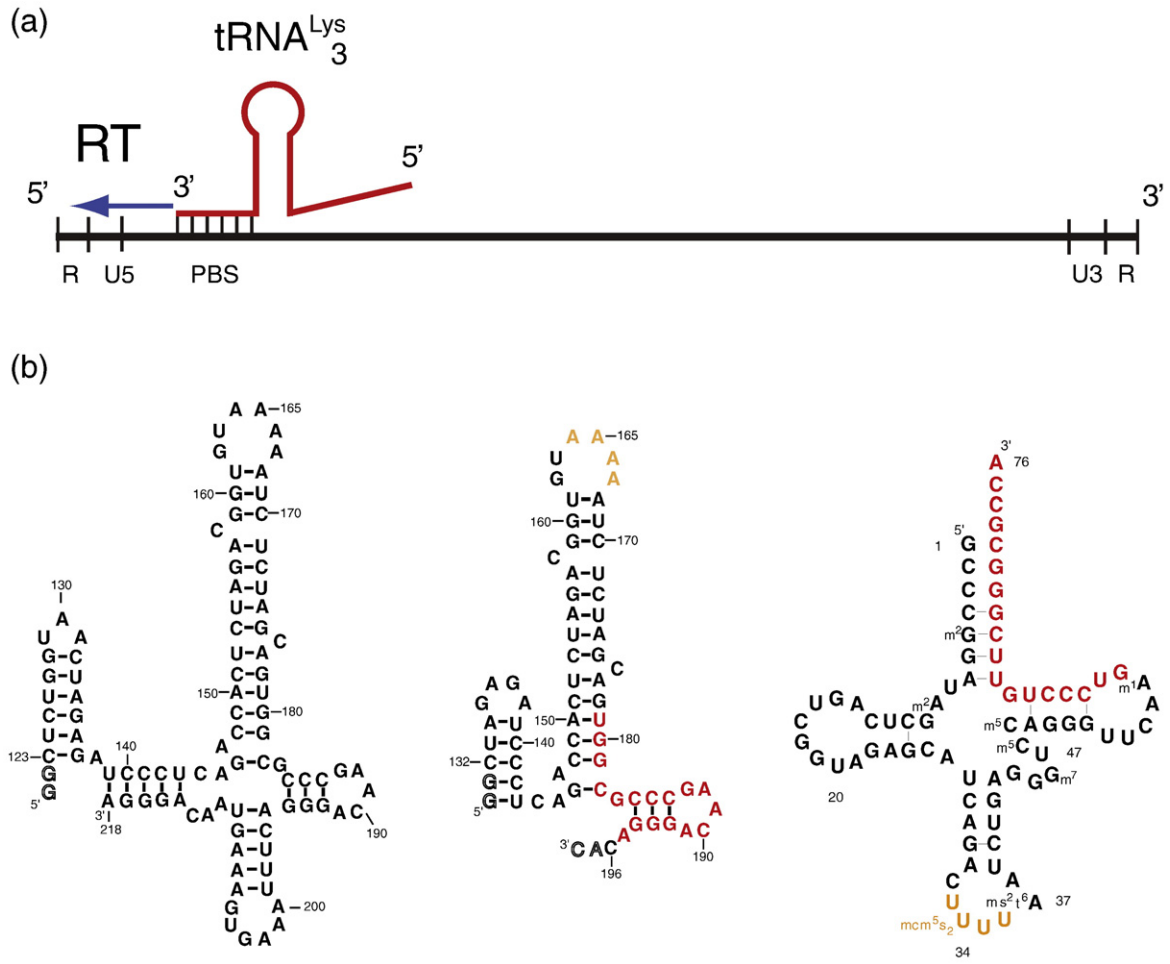

Fig. 1. (a) Schematic of the reverse transcription initiation complex within the HIV genome. The RNA genome of HIV RNA has the standard structure for retroviral genomes, with long terminal repeat regions (U5, U3, and R), and the PBS near the 5′ end. Host $tRNA^{Lys}_3$ interacts with the PBS through an 18-bp interaction, providing the 3′ OH group to initiate reverse transcription by viral RT. (b) RNA oligonucleotides corresponding to HIV-1 genomic RNAs used in the current study, (left) a 99-nt RNA and (right) a 69-nt RNA with numbering according to the Mal isolate. Additional nucleotides added at the 5′ end for transcription with T7 RNA polymerase are indicated as outlines. The PBS region of 18 bp of complementarity with tRNA is highlighted in red, whereas the A-rich loop is highlighted in yellow. (c) Secondary structure of human $tRNA^{Lys}_3$ with modified nucleotides; coloring as in (b).

DNA hybridize and lead to eventual synthesis of a fully double-stranded DNA that is competent for integration into the host genome.

Initiation of reverse transcription in HIV is directed from a specific RNA–RNA complex formed between host $tRNA^{Lys}_3$ and HIV genomic RNA.[11] The PBS forms an 18-bp duplex with the 3′ end of $tRNA^{Lys}_3$, but sequences outside this region of tRNA–HIV RNA complementarity are also critical for viral replication. Attempts to force HIV to use other tRNAs by making the PBS sequence complementary to the 3′-most 18 nt of the tRNA lead to rapid reversion of the virus back to using $tRNA^{Lys}_3$ by mutation of the PBS.[12–16] Thus, sequences outside the simple PBS–tRNA base-pairing region are required for efficient initiation. Chemical probing of the viral RNA focused on a ca 100-nt region whose reactivity to chemical probes changed upon binding of the viral RNA.[6,7] Mutations within this region affect both the efficiency of tRNA–viral RNA complex formation and the rate of initiation of reverse transcription *in vitro*[8,17] and *in vivo*.[18,19] Different viral isolates have sequence changes in this region.[20–22]

Biochemical experiments have defined the general RNA–RNA interactions that direct the formation of the initiation complex. The RNA–RNA complex that forms between host $tRNA^{Lys}_3$ and HIV PBS is >40 kDa in size, and global features of its secondary structure have been determined by chemical and enzymatic probing.[7,8,23] Modified nucleotides in $tRNA^{Lys}_3$ facilitate proper initiation,[6] and a specific