Virus Research 155 (2011) 98–105

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Structural changes in the SL5 and SL6 leader sequences of HIV-1 RNA following interactions with the viral mGag protein

Ariel Roldan, Chen Liang, Mark A. Wainberg*

McGill University AIDS Center, Lady Davis Institute for Medical Research, Sir Mortimer B. Davis Jewish General Hospital, 3755 Chemin-Côte-Ste-Catherine, Montréal, QC H3T 1E2, Canada

ARTICLE INFO

Article history:
Received 9 June 2010
Received in revised form 31 August 2010
Accepted 7 September 2010
Available online 22 September 2010

Keywords:
RNA–protein interaction
Footprinting
HIV assembly

ABSTRACT

We previously identified sequences downstream of the SL4 region of HIV-1 RNA that are involved in the recognition of the 5′ leader region of HIV-1 RNA by a minimal version of the HIV-1 Gag protein (mGag). These sequences increase the affinity of this interaction, promote Gag multimerization, and enhance formation of an early mGag–RNA complex. Now, we provide protein footprinting results on the +350 to +500 nucleotide region of viral RNA, based on use of different single-stranded and base-paired ribonucleases. Use of the *mfold* program confirmed the existence of both a stem-loop 5 (SL5), downstream of SL4, and a more complex multi-stem-loop structure (SL6). Footprinting analysis using mGag and single-stranded-specific nucleases showed almost complete protection of the single-stranded region. In contrast, results obtained with RNase V1, a double-stranded-specific nuclease, suggest that the RNA structure is changed upon protein binding, presumably because of formation of novel and longer stems. Furthermore, RNA footprinting, using viral nucleocapsid protein (NC) and RNase VI, indicate a highly double-stranded structure in several regions. These findings show that viral RNA structure is modified by interaction with proteins, and that NC may possess different chaperone activity in the context of the Gag precursor than in its mature form.

1. Introduction

The HIV-1 Gag polyprotein is responsible for viral particle formation and genomic RNA packaging during the assembly process (Coffin et al., 1997; Frankel and Young, 1998). Each of the smaller proteins that result from cleavage of the Gag polyprotein (Hellen et al., 1989) (matrix, MA; capsid, CA; nucleocapsid, NC; p6 and two spacer peptides, SP1 and SP2) also have multiple functions (Göttlinger, 2001; Scarlata and Carter, 2003). It is now well understood that certain polypeptides may not behave similarly, depending whether they are present within Gag or within cleaved products. In particular, the C-terminal domain of CA (CTD-CA), along with SP1 and NC, forms an assembly domain in the immature polyprotein (Accola et al., 1998; Liang, 2002), whereas the same peptides constitute three independent domains that possess different functions when part of mature proteins. We wished to study the interactions of this assembly domain with genomic RNA and to compare such interactions with those that occur between genomic RNA and the mature NC protein.

The strongest evidence of an assembly domain constituted from the C-terminal portion of CA and extending through NC comes from studies showing that these regions represent the minimum requirements for the generation of virus-like particles (VLPs) (Accola et al., 2000; Krausslich et al., 1995; Morikawa et al., 2000). Interestingly, the highest association constant for HIV-1 Gag protein–protein interaction is found at the dimer interface of amino acids W317 and M318, located in the CTD-CA, suggesting that this dimer interface plays a crucial role in the assembly process (Gamble et al., 1997). In *in vitro* assembly experiments, deletion of SP1 from Gag or the creation of SP1 mutants yielded cones or tubules instead of the spheres obtained with wild-type Gag, suggesting that the amino acids in this small peptide were able to modulate protein–protein interaction (Gross et al., 2000; Keller et al., 2008). *In vivo* evidence for the participation of SP1 in Gag multimerization has also been obtained (Liang, 2002; Morikawa et al., 2000). This small peptide is also involved in genomic RNA packaging (Kaye and Lever, 1998; Russell et al., 2003), and restriction of the packaging of spliced viral RNA is thought to be determined by a single amino acid located within SP1 (Roy et al., 2006). Although NC can be replaced by other nucleic acid binding motifs for VLP formation, its role in viral assembly is very well documented (Rein et al., 1998) and it is crucial for the maintenance of viral structural stability (Wang et al., 1998).

* Corresponding author. Tel.: +1 514 340 8260; fax: +1 514 340 7537.
E-mail address: mark.wainberg@mcgill.ca (M.A. Wainberg).

0168-1702/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2010.09.005

The first 500 bases of the HIV genome are highly structured and are important in RNA recognition during the viral life cycle (Berkowitz et al., 1995; Clever et al., 1995). The untranslated leader segment of this RNA is important in regulation of gene expression and in multiple virion-associated processes including dimerization, encapsidation and reverse transcription (Harrison et al., 1998; Huthoff and Berkhout, 2002). The overall structure of the first 300 bases has been well defined, even though some issues remain in regard to exact conformation and whether this sequence exists as one or as alternative conformers (Huthoff and Berkhout, 2001; Ooms et al., 2004). Most of the RNA structures predicted *in vitro* for this region, e.g. TAR, poly (A), PBS, SL1–4 have been confirmed in *ex vivo* structural studies (Paillart et al., 2002). In contrast, little is known about RNA structures downstream of SL4.

We previously showed that a RNA fragment spanning the +350 to +500 nucleotide region of Gag is important for protein–RNA binding and Gag multimerization (Roldan et al., 2004). This RNA fragment remarkably increases protein–RNA binding association forces involving a minimal Gag (mGag) protein and genomic RNA. Moreover, this RNA fragment was crucial for multimerization patterns observed in gel shift assays in interactions with mGag. Therefore, we were interested in further studying interactions between mGag and this important RNA fragment. The objectives of the present work were: (1) to study the RNA structure of the fragment spanning the +350 to +500 nt positions of HIV-1 genomic RNA; (2) to analyze protein:RNA interactions in terms of RNA protection and conformational changes of this RNA fragment upon binding with mGag; (3) to compare binding results obtained with mGag with those obtained with viral NC protein. We now report the structural analysis of this RNA stretch together with relevant protein footprinting results. Our data are in agreement with previously reported findings on the structure of the HIV-1 leader RNA and also show that this viral RNA structure is modified by interaction with proteins. Furthermore, the viral NC protein has different chaperone activity in the context of its presence within the Gag precursor than in its mature form.

2. Materials and methods

2.1. Protein purification

Proteins were purified as previously published (Roldan et al., 2004). Briefly, proteins were expressed in BL21 (DE3) cells (Novagen, Madison, WI) and purified under denaturing conditions by affinity with nickel agarose. Further purification was performed by anion exchange chromatography (Q-sepharose) in 7 M urea, 50 mM Tris–HCl, pH 8.5. Proteins were collected from the unbound material and redissolved to 0.5–2 mg/ml. Protein was refolded by dialysis against 3 M urea, 500 mM Tris–HCl pH 8, 50 mM NaCl, 10 μM ZnCl, 5/0.5 mM reduced/oxidized glutathione and 0.01% CHAPS at 4 °C for 16 h overnight. Thereafter, two buffer exchanges were performed with the same buffer but without urea for 3 h at 4 °C, and this was followed by a final dialysis step against storage buffer (same buffer but 50 mM Tris–HCl, pH 8, and no CHAPS). Non-soluble proteins were pelleted by ultracentrifugation at 100,000 × *g* for 1 h. Proteins from different batches were employed in each of the different experiments and no differences among these various preparations were observed in regard to results obtained.

2.2. Synthesis and folding of RNA

A DNA PCR product spanning +223 to +560 of HIV proviral DNA (GenBank™ K03455) was used as template to *in vitro* synthesize a RNA probe spanning the desired genomic RNA fragment. For this purpose, the T7 promoter was included at the 5′ end of the forward primer. PCR products were agarose-purified and used as templates in T7 transcription reactions with the T7 Megashortscript Kit (Ambion, Texas) following manufacturer's recommendations. Reactions were stopped by adding 2 units of DNase I and further incubation at 37 °C for 15 min. RNA transcripts were purified on denaturing polyacrylamide gels and eluted in 600 mM NaOAc, 1 mM EDTA, and 0.1% SDS. RNA templates were precipitated with 2.5 volumes of ethanol and resuspended in 100 mM NaCl, 20 mM Tris, pH 7.5. RNA molecules were refolded by heating at 85 °C for 5 min and slowly cooled down to room temperature. RNA was then aliquoted and stored at −80 °C.

2.3. RNA structural analyses and footprinting assays

RNase A, T1 and V1 were purchased from Ambion, Texas; and RNase One from Promega, Madison, WI. Nuclease treatments were performed in 10 μl containing 200 ng of RNA probe, 10 mM Tris, pH 7, 100 mM KCl, 10 mM $MgCl_2$, 1 μg of yeast RNA. The RNase concentration employed was 0.01 U/μl for RNase T1, 0.001 U/μl for RNase V1 and 0.01 μg/ml for RNase A. After incubation for 15 min on ice, reactions were stopped by proteinase K digestion at 37 °C for 30 min and then RNA was phenol:chloroform extracted and ethanol precipitated. For footprinting assays, protein and RNA were preincubated for 15 min on ice in 9 μl. Then, 1 μl containing the above specified amount of RNase was added and after incubation for another 15 min, reactions were processed as described above for RNA structural analyses.

2.4. Reverse transcription elongation assays

Reverse transcription was performed using the Superscript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA) following manufacturer's recommendations. Briefly, RNA pellets were resuspended directly in 10 μl of mix A (2 μmol of radiolabelled primer and 20 mM of dNTPs) and then samples were incubated at 65 °C for 5 min and then chilled on ice. Primer elongation was performed at 55 °C for 50 min in 20 mM Tris–HCl, pH 8.4, 50 mM KCl, 5 mM $MgCl_2$, 10 mM DTT, 2 units of RNase OUT and 10 units of Superscript III RT enzyme. Reactions were terminated by incubation at 85 °C for 5 min and then RNA was removed from samples by adding 2 units of RNase H and further incubation for 20 min at 37 °C.

3. Results

3.1. RNA structure of SL5 and SL6

First, we wished to better understand the nature of the RNA fragment downstream of SL4, particularly that spanning positions +350 to +500. Toward this end, we first used a fragment spanning nucleotide positions +223 to +350 in order to probe important RNA structures known to interact with the viral proteins (mGag and NC) that govern protein–RNA association events. Several commercially available nucleases were tested (not shown), but RNases A, One, T1 and V1 were found to be the more informative. After nuclease treatment, the RNA probe was retrotranscribed using a primer complementary to sequences located 3′ to the predicted SL6 structure (Fig. 1).

As a result, we were able to reconstruct the physical map of the RNA fragment spanning positions +375 to +500 of HIV-1 genomic RNA (Fig. 2). It can be seen that the four different RNase enzymes, i.e. RNase A (A), One (O), VI (V) and TI (T) each yielded a distinct pattern of digestion of the viral RNA probe that was used that spanned positions +223 to +560. It can also be seen that the most apparent cleavages by each of these enzymes occurred in the region between +375 and +484, as is illustrated by the band at the left side of the