HIV-1 Sequence Necessary and Sufficient to Package Non-viral RNAs into HIV-1 Particles

Yang Liu [1], Olga A. Nikolaitchik [1], Sheikh Abdul Rahman [1], Jianbo Chen [1], Vinay K. Pathak [2] and Wei-Shau Hu [1]

*1 - **Viral Recombination Section,** HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702, USA*
*2 - **Viral Mutation Section,** HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702, USA*

***Correspondence to Wei-Shau Hu:** National Cancer Institute, 1050 Boyles Street, Building 535, Room 336, Frederick, MD 21702, USA. Wei-Shau.Hu@nih.gov*
http://dx.doi.org/10.1016/j.jmb.2017.06.018
Edited by M.F. Summers

Abstract

Genome packaging is an essential step to generate infectious HIV-1 virions and is mediated by interactions between the viral protein Gag and *cis*-acting elements in the full-length RNA. The sequence necessary and sufficient to allow RNA genome packaging into an HIV-1 particle has not been defined. Here, we used two distinct reporter systems to determine the HIV-1 sequence required for heterologous, non-viral RNAs to be packaged into viral particles. Although the 5′ untranslated region (UTR) of the HIV-1 RNA is known to be important for RNA packaging, we found that its ability to mediate packaging relies heavily on the context of the downstream sequences. Insertion of the 5′ UTR and the first 32-nt of *gag* into two different reporter RNAs is not sufficient to mediate the packaging of these RNA into HIV-1 particles. However, adding the 5′ half of the *gag* gene to the 5′ UTR strongly facilitates the packaging of two reporter RNAs; such RNAs can be packaged at >50% of the efficiencies of an HIV-1 near full-length vector. To further examine the role of the *gag* sequence in RNA packaging, we replaced the 5′ *gag* sequence in the HIV-1 genome with two codon-optimized *gag* sequences and found that such substitutions only resulted in a moderate decrease of RNA packaging efficiencies. Taken together, these results indicated that both HIV-1 5′ UTR and the 5′ *gag* sequence are required for efficient packaging of non-viral RNA into HIV-1 particles, although the *gag* sequence likely plays an indirect role in genome packaging.

Published by Elsevier Ltd.

Introduction

Most HIV-1 particles contain two copies of full-length viral RNA [1,2]; this efficient genome packaging is mediated by the interactions between HIV-1 Gag polyprotein and *cis*-acting elements in full-length viral RNA [3–5]. All orthoretroviral Gag polyproteins contain matrix (MA), capsid (CA), and nucleocapsid (NC) domains. In addition, the HIV-1 Gag polyprotein contains the p6 domain and two spacer peptides, SP1 and SP2, which are located between CA-NC and NC-p6 junctions, respectively. Of the Gag domains, NC plays an important role in retroviral RNA genome packaging [6–13]. Mutations in the HIV-1 NC domain, including those that alter the two CCHC zinc-chelating motifs, can reduce virion RNA packaging efficiency [10,11,13]. In addition, complementation experiments have demonstrated that of the thousands of Gag molecules assembled into each particle, a significant portion of the Gag (>17%) needs to contain functional NC before RNA can be packaged efficiently [14].

In the full-length retroviral RNAs, there are *cis*-acting sequences important for the encapsidation of the genome; such sequences are often referred to as the packaging signal. The packaging signal has been studied in multiple retroviruses including murine leukemia virus (MLV), spleen necrosis virus, Rous sarcoma virus (RSV), HIV-1, HIV-2, bovine leukemia virus (BLV), Mason Pfizer monkey virus (MPMV), and mouse mammary tumor virus [15–26]. In these retroviruses, sequences in the 5′ untranslated region (UTR) and often 5′ end of the *gag* gene are important for packaging. In MLV and RSV, the sequence necessary and sufficient for RNA packaging has

been defined [15,17]; when present in heterologous RNA, these sequences can mediate the encapsidation of heterologous RNA into MLV and RSV particles, respectively. Furthermore, the heterologous RNA containing the MLV packaging signal was encapsidated into viral particles as a dimer [27]. However, the sequence necessary and sufficient for efficient packaging of non-viral RNA into HIV-1 particles has not been defined.

Multiple studies have been performed to examine RNA elements important for HIV-1 RNA packaging ([18,19,28–32] and summarized in [33–38]). These studies showed that mutations in the 5′ UTR and *gag*, as well as other elements, affect HIV-1 RNA packaging. The 5′ UTR region of the viral RNA is highly structured and forms multiple stem–loop structures [39–44]; several studies suggest that many but not all of the 5′ UTR RNA structures are essential for genome packaging. For example, the trans-activation region (TAR), a stem–loop structure located at the very beginning of the HIV-1 RNA, was suggested to be important for RNA packaging. However, later studies showed that another RNA stem–loop structure can partially replace the TAR element function, and it was suggested that TAR mainly provides stability to the RNA structure to facilitate genome packaging [32,45–47]. Although it is generally agreed that sequences beyond the AUG of the *gag* gene can affect RNA packaging, the precise extent of the *gag* gene required for packaging has varied in different studies [32,48]. Furthermore, translation of the *gag* gene is not required for its effect in enhancing RNA packaging [49]. The Gag-Pol ribosomal frameshift signal and the Rev response elements (RRE) were also suggested to be important for RNA packaging. Viral protein Rev binds to RRE and mediates the export of full-length and partially spliced HIV-1 RNAs [50–54]; Rev–RRE was suggested to play a role in HIV-1 RNA packaging as the lack of Rev or mutation of RRE affects genome packaging [55,56]. However, other studies showed that the function of Rev–RRE can be replaced by the constitutive transport element (CTE) from MPMV [32,57,58]. The CTE from MPMV does not contain sequence or structural similarity with HIV-1 RRE; in addition, CTE mediates RNA export via the NXF1 pathway, whereas the Rev–RRE complex mediates the RNA export via the CRM1 pathway [59–62]. Hence, the major effect of the Rev–RRE mutations is the lack of proper RNA export that leads to defects in genome packaging and the Rev–RRE complex does not directly participate in genome packaging [57,58]. Similarly, the ribosomal frameshift signal that mediates the expression of Gag-Pol polyprotein was originally hypothesized to play an important role in RNA packaging but was later demonstrated to have no effect on RNA packaging [63,64].

At this time, the minimal sequence necessary and sufficient to allow non-viral RNAs to be efficiently packaged into HIV-1 particles has not been defined. In this report, we sought to define the sequences required for non-viral reporter RNAs to be efficiently packaged into HIV-1 particles. We used a previously described single-virion analysis system [1] to visualize viral RNA in individual particles to determine the efficiency of HIV-1 RNA genome packaging. We performed systematic deletion/replacement of portions of the viral genome to define regions dispensable for RNA packaging. We then inserted HIV-1 sequences into reporter RNAs encoding either a firefly luciferase gene or a *Renilla* luciferase gene and examined the packaging efficiency of these RNAs. We found that including the 5′ UTR and the 5′ half of the *gag* gene sequence allows the non-viral reporter RNAs to be packaged efficiently into viral particles. However, the role of the *gag* gene sequences is likely to be to stabilize the RNA structure rather than contain specific Gag:RNA recognition sites. These results define the minimum HIV-1 RNA packaging signal and provide a better understanding of RNA genome encapsidation, a process essential for the generation of infectious HIV-1.

Results

Experimental system used to examine HIV-1 RNA genome packaging efficiency

We used two previously described, NL4-3-based constructs to perform single-virion analysis. The general structure of one of the clones, 1Gag, is shown in Fig. 1a. This construct contains all of the *cis*-acting elements required for viral replication and expresses Gag fused to cerulean fluorescent protein (Gag–CeFP), Tat, Rev, and Nef, whereas portions of the *pol*, *env*, *vif*, *vpr*, and *vpu* gene were deleted. In addition, a set of stem–loop sequences (BSLs) recognized by bacterial protein BglG is located in the *pol* gene; thus, only the full-length, unspliced viral RNA contains the BSL. Construct 1GagΔCeFP has the same structure as 1Gag but expresses an untagged Gag. For each of the constructs described below, a pair of plasmids was used, one that expresses Gag–CeFP and another that expresses untagged Gag; for brevity, only the Gag–CeFP version of the constructs is shown in all figures.

Single-virion analysis was performed by transfecting 293T cells with 1Gag, 1GagΔCeFP, and Bgl–YFP, which encodes a truncated BglG fused to yellow fluorescent protein (YFP). HIV-1 particles were harvested and visualized using fluorescent microscopy (Fig. 1b). The coexpression of Gag and Gag–CeFP allows for the formation of morphologically normal HIV-1 particles [1,2,65,66] and the detection of viral particles by their CeFP signals. The BSL in the viral RNA allows for specific binding of Bgl–YFP to the