RNA Biology

ISSN: 1547-6286 (Print) 1555-8584 (Online) Journal homepage: http://www.tandfonline.com/loi/krnb20

The C-terminal p6 domain of the HIV-1 Pr55^{Gag} precursor is required for specific binding to the genomic RNA

Noé Dubois, Keith K. Khoo, Shannon Ghossein, Tanja Seissler, Philippe Wolff, William J. McKinstry, Johnson Mak, Jean-Christophe Paillart, Roland Marquet & Serena Bernacchi

To cite this article: Noé Dubois, Keith K. Khoo, Shannon Ghossein, Tanja Seissler, Philippe Wolff, William J. McKinstry, Johnson Mak, Jean-Christophe Paillart, Roland Marquet & Serena Bernacchi (2018): The C-terminal p6 domain of the HIV-1 Pr55^{Gag} precursor is required for specific binding to the genomic RNA, RNA Biology, DOI: 10.1080/15476286.2018.1481696

To link to this article: https://doi.org/10.1080/15476286.2018.1481696

RNA BIOLOGY
https://doi.org/10.1080/15476286.2018.1481696

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

The C-terminal p6 domain of the HIV-1 Pr55^{Gag} precursor is required for specific binding to the genomic RNA

Noé Dubois[a], Keith K. Khoo [b,c], Shannon Ghossein[b,c], Tanja Seissler[a], Philippe Wolff[a,d], William J. McKinstry [c], Johnson Mak [b,e], Jean-Christophe Paillart [a], Roland Marquet [a], and Serena Bernacchi [a]

[a]Architecture et Réactivité de l'ARN, UPR 9002, IBMC, CNRS, Université de Strasbourg, Strasbourg, France; [b]School of Medicine, Deakin University, Geelong, Australia; [c]CSIRO Manufacturing, Parkville, Australia; [d]Plateforme protéomique Strasbourg-Esplanade, IBMC, CNRS, Université de Strasbourg, Strasbourg, France; [e]Institute for Glycomics, Griffith University, Southport, Australia

ABSTRACT

The Pr55^{Gag} precursor specifically selects the HIV-1 genomic RNA (gRNA) from a large excess of cellular and partially or fully spliced viral RNAs and drives the virus assembly at the plasma membrane. During these processes, the NC domain of Pr55^{Gag} interacts with the gRNA, while its C-terminal p6 domain binds cellular and viral factors and orchestrates viral particle release. GagΔp6 is a truncated form of Pr55^{Gag} lacking the p6 domain usually used as a default surrogate for wild type Pr55^{Gag} for *in vitro* analysis. With recent advance in production of full-length recombinant Pr55^{Gag}, here, we tested whether the p6 domain also contributes to the RNA binding specificity of Pr55^{Gag} by systematically comparing binding of Pr55^{Gag} and GagΔp6 to a panel of viral and cellular RNAs. Unexpectedly, our fluorescence data reveal that the p6 domain is absolutely required for specific binding of Pr55^{Gag} to the HIV-1 gRNA. Its deletion resulted not only in a decreased affinity for gRNA, but also in an increased affinity for spliced viral and cellular RNAs. In contrast GagΔp6 displayed a similar affinity for all tested RNAs. Removal of the C-terminal His-tag from Pr55^{Gag} and GagΔp6 uniformly increased the Kd values of the RNA-protein complexes by ~ 2.5 fold but did not affect the binding specificities of these proteins. Altogether, our results demonstrate a novel role of the p6 domain in the specificity of Pr55^{Gag}-RNA interactions, and strongly suggest that the p6 domain contributes to the discrimination of HIV-1 gRNA from cellular and spliced viral mRNAs, which is necessary for its selective encapsidation.

ARTICLE HISTORY
Received 1 February 2018
Accepted 22 May 2018

KEYWORDS
HIV-1; Pr55^{Gag}; p6 domain; genomic RNA; RNA-protein binding specificity; fluorescence spectroscopy

1. Introduction

Retroviral genomes are specifically selected for packaging from an excess of cellular and spliced viral mRNAs (for reviews see [1–4]). The human immunodeficiency virus type 1 (HIV-1) Pr55^{Gag} precursor drives specific selection of the genomic RNA (gRNA) by interacting with packaging signals (Psi) located within the 5′ untranslated region (UTR) and the beginning of the *gag* gene of gRNA (Fig. 1A) [1,3–6]. The Psi is composed of four stem-loops (SL1 to SL4). SL1 constitutes the gRNA Dimerization Initiation Site (DIS) [7–11]; gRNA dimerization is mediated by a 6 nucleotides palindromic sequence located in the apical loop of SL1 [12]. Interestingly recent findings also showed that SL1 contains the main Pr55^{Gag} recognition signal [13,14]. SL2 contains the major splice donor (SD) site, and SL3 contributes to gRNA packaging [15–19]. Finally, the folding of SL4 is most likely in equilibrium with the so-called U5-AUG long-range interaction [20–23], which may regulate the last stages of the packaging process (Fig. 1A) [24,25]. Upstream of Psi, highly structured hairpins including the trans-activating responsive element (TAR), the poly(A) hairpin and the Primary Binding Site (PBS) domain have also been proposed to be involved in gRNA packaging [26–28]. Although packaging of gRNA is a highly selective process, spliced viral RNAs [5,29,30], and cellular RNAs [30,31] are also found in viral particles. For instance, 7SL RNA, a component of the signal recognition particle [32,33], tRNAs, U6 spliceosomal RNA [30,34] are enriched in retroviral particles. However, the encapsidation of spliced viral and cellular RNAs follow different mechanisms [30] (for review see [35]): indeed, while SL1 is crucial for packaging of gRNA, it does not seem to be involved in the packaging of spliced viral RNAs [30].

The 55-kDa Pr55^{Gag} precursor is composed of several domains, starting with the matrix (MA) at the N-terminus, capsid (CA), nucleocapsid (NC) flanked by the two small peptides SP1 and SP2, and p6 at its C-terminus (Fig. 1B) (for review see [36]). The MA domain mediates Pr55^{Gag} membrane binding through a bipartite signal consisting of the N-terminal myristoylated glycine and the highly basic region (HBR) [37,38]. The CA domain drives Pr55^{Gag} multimerization thus leading to the formation of the structural viral core, and the NC contains two zinc finger motifs, a major determinant for gRNA recognition and packaging [39–43]. Interestingly, the HBR has been shown to promote the interaction of the MA domain with RNA *in vitro* [43–45], as well as in the cytosol [46]. Finally, the C-terminal p6 domain

CONTACT Serena Bernacchi s.bernacchi@ibmc-cnrs.unistra.fr CNRS, Architecture et Réactivité de l'ARN, Université de Strasbourg, UPR 9002, IBMC-15 rue René Descartes, F-67000, Strasbourg, France
Supplemental data for this article can be accessed here.

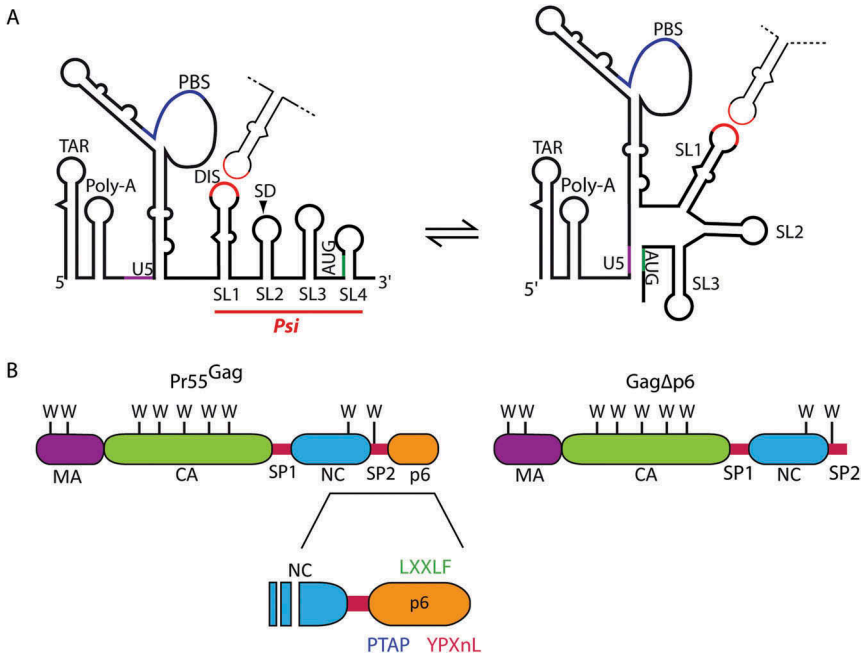

Figure 1. Domains involved in the interaction between the HIV-1 RNA genome and the Gag polyprotein precursor. (A) Schematic representation of the secondary structure of the 5′-end of HIV-1 genome. The red line delimits the packaging signal (Psi) region comprising SL1-SL4. SL1 contains in its apical loop the Dimerization Initiation Site (DIS). The major Splice Donor (SD) site and *gag* AUG initiation codon are represented. The secondary structure model of the dimer with the long-range U5-AUG base pairing is represented on the right side. **(B)** Schematic representation of the full-length $Pr55^{Gag}$ and GagΔp6 proteins used in this study. Their different domains are indicated: the matrix (MA), capsid (CA), nucleocapsid (NC) flanked by the two spacer peptides SP1 and SP2, and finally the C-terminal p6 domain. Tryptophan residues W and p6 late domains are represented.

regulates budding of nascent virions at the plasma membrane [47]. This process involves the interaction of its two late domains PTAP and YPXnL with two host factors associated to the ESCRT (Endosomal Sorting Complex Required for Transport) machinery, namely TSG101 (Tumor Susceptibility Gene 101) [48–50] and ALIX (ALG-2 interaction protein X) [51–54] (Fig. 1B). The LXXLF p6 domain also binds the viral Vpr protein, ensuring its encapsidation [55–57] and it is excluded from the capsid core during viral maturation [58,59] (Fig. 1B).

Production of full-length recombinant $Pr55^{Gag}$ has been hindered by the instability of the C-terminus of $Pr55^{Gag}$ [60], hence a non-physiological C-terminus truncated version of this precursor named GagΔp6 (also named $Pr50^{GagΔp6}$) (Fig. 1B) is often used as a surrogate for biochemical analysis of $Pr55^{Gag}$ function. It is indeed usually assumed that the p6 domain does not affect RNA binding, and hence the vast majority of *in vitro* studies using recombinant Gag were conducted with GagΔp6 [43,61–66]. However, recent studies comparing the RNA binding properties of NCp7 (i.e. the mature NC domain) and NCp15 (i.e. NC-SP2-p6), suggested that the p6 domain might indirectly affect RNA binding [67]. Similarly, the p6 domain was recently shown to affect binding of Gag proteins to short oligoribonucleotides [68]. We thus decided to test the role of the p6 domain in the specific recognition of $Pr55^{Gag}$ to the HIV-1 gRNA. We performed a systematic comparison of $Pr55^{Gag}$ and GagΔp6 binding to a series of viral and cellular RNA species under strictly identical conditions using fluorescence spectroscopy. Unexpectedly, our results demonstrate that p6 deletion resulted not only in a decreased affinity for gRNA, but also in an increased affinity for spliced viral and cellular RNAs. Indeed, all tested RNAs bound GagΔp6 with similar affinity, revealing a complete lack of binding specificity. Altogether our findings demonstrate a novel role for the p6 domain as a regulator of the binding specificity of $Pr55^{Gag}$ to HIV-1 gRNA.

2. Results

2.1. DLS and SLS analysis of GagΔp6 and $Pr55^{Gag}$ proteins

Our recombinant GagΔp6 and $Pr55^{Gag}$ proteins were tested and compared using Diffusion Light Scattering (DLS) and Static Light Scattering (SLS) (Fig. 2). The DLS intensity distribution of GagΔp6 and $Pr55^{Gag}$ in the storage buffer (50 mM Tris-HCl pH 8, 1 M NaCl, 5 mM DTT,) was unimodal and fairly monodisperse (P_d index ~ 18.3% and 11.8%, for GagΔp6 and $Pr55^{Gag}$, respectively). The higher molecular weight species observed in the DLS intensity distribution (Fig. 2A, left panel) correspond to very minor populations, and were not observed in the DLS distribution by number (Fig. 2A, right panel). These profiles correspond to mean hydrodynamic radii (Rh) of 4–5.1 nm and 6.5–7.3 nm for GagΔp6 and $Pr55^{Gag}$, respectively (Fig. 2A). Using the Stokes-Einstein equation and assuming spherical proteins (see Methods), our analysis suggests that in solution GagΔp6 corresponds to a dimer-trimer, while $Pr55^{Gag}$ to tetramer-pentamer. On the other hand, SLS data on GagΔp6 and $Pr55^{Gag}$ in the storage buffer provided molecular weight (MW)

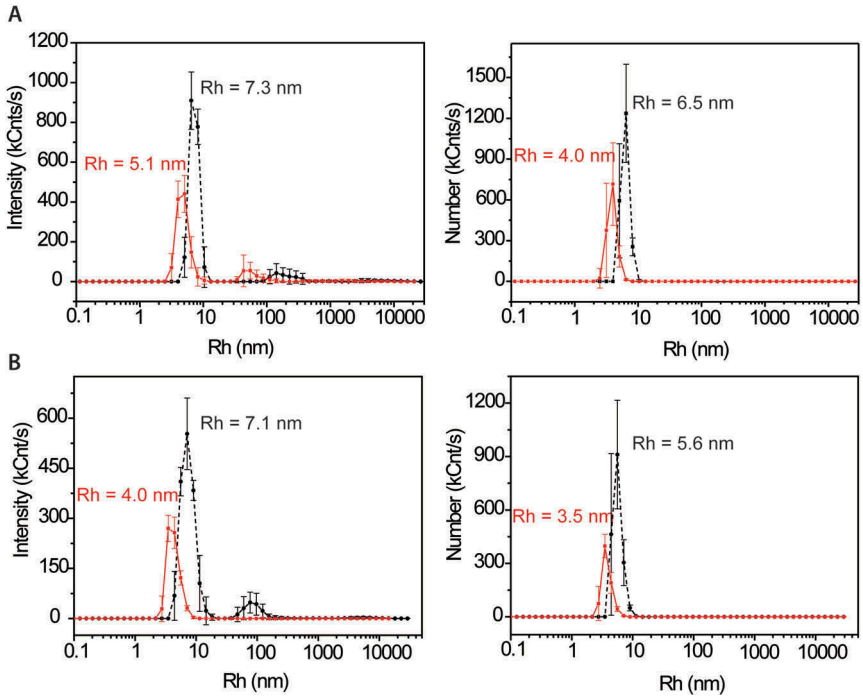

Figure 2. DLS analysis of GagΔp6 and Pr55^{Gag}. DLS analysis of GagΔp6 (red continuous line) and Pr55^{Gag} (black discontinuous line) **(A)** in the storage buffer (1 M NaCl, 5 mM DTT, 50 mM Tris-HCl pH 8) **(B)** and in the binding buffer (30 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) are represented. The intensity (left) and number (right) distributions are represented. The polydispersity index (PdI) and hydrodynamic radius (Rh) of each protein are indicated close to its corresponding peak. Mean ± SD of ten measurements.

estimations of 46.9 ± 1.8 kDa and 114.0 ± 4.0 kDa, respectively, corresponding to a monomer state for GagΔp6 and a dimer for Pr55^{Gag} in a good agreement with previous analysis [69]. The discrepancy between the oligomeric states determined by DLS and SLS is probably due to the limiting hypothesis imposed by the Stokes-Einstein model, which assumes the proteins to be spherical. Nevertheless, DLS and SLS both indicated that the oligomeric state of the GagΔp6 and Pr55^{Gag} proteins is different. The DLS profiles and SLS analysis of GagΔp6 and Pr55^{Gag} were also performed in the binding buffer (30 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) and were shown to be similar to the storage conditions (Fig. 2B).

2.2. GagΔp6 and Pr55^{Gag} binding to the 5' region of HIV-1 gRNA

We exploited the intrinsic fluorescent signal of GagΔp6 and Pr55^{Gag}, which harbor 9 Trp residues in their MA, CA, SP2 and NC domains [70] (Fig. 1B), to determine the GagΔp6 binding parameters (the dissociation constants K_d, and the binding stoichiometry n) for RNA fragments corresponding to the first 600/615 of the HIV-1 gRNA of NL4.3 (N1-600 WT) and MAL (M1-615 WT) isolates (Fig. 3, which were previously shown to bind Pr55^{Gag} with the same affinity [13,14]. In addition, RNA binding assays with GagΔp6 and Pr55^{Gag} were performed under the exact same conditions (buffer, temperature, refolding and incubation protocols) [14].

The experimental binding curves for both RNAs to GagΔp6 were fitted with one-binding site model (Fig. 4, Equation 7, see Material and Methods). Accordingly, the Scatchard plots of GagΔp6 to N1-600 WT and M1-615 WT RNAs confirmed the presence of only one class of binding sites of very similar affinity (K_d ~ 11–12 nM, Table 1, Fig. 4). This observation strongly contrasts with the binding of these RNAs to the full-length Pr55^{Gag} that displayed two classes of binding sites for those RNA fragments: a very high affinity binding site (K_{d1} ~ 2–3 nM, Table 1, Fig. 4) and a lower one (K_{d2} ~ 15 nM, Table 1, Fig. 4) [14] whose affinity was similar to the one observed for GagΔp6 (Table 1). Note that while the data in Tables 1–3 concerning Pr55^{Gag} are from our previous publication [14], all curves in Figs. 4 to 7 and in Supplementary Figures 1 and 2 were previously unpublished and allow direct comparison between GagΔp6 and Pr55^{Gag}. Analysis of the stoichiometry showed about 6–7 Pr55^{Gag} proteins bound to the high affinity binding site, while only 3 to the lower one. By comparison, about 3–4 GagΔp6 proteins bound to N1-600 WT and M1-615 WT RNAs (Table 1). Altogether, our data show that GagΔp6 does not display a very high affinity binding site for the 5′-end region of the HIV-1 gRNA, which is the hallmark of the specific binding of Pr55^{Gag} to this RNA.

2.3. GagΔp6 and Pr55^{Gag} binding to SL1 RNA mutants

We previously identified two major determinants for the interaction between Pr55^{Gag} and gRNA located in SL1 [13,14,71]. They correspond to gRNA dimerization, which is ensured by a 6 nucleotides palindromic sequence in the apical loop [7–11], and to the purine-rich internal loop (Fig. 1A). To compare the role of gRNA dimerization on GagΔp6- and Pr55^{Gag}-RNA binding, we analyzed RNA fragments in which the apical loop of SL1 was mutated to prevent gRNA dimerization (M1-615SL1sAL and N1-600SL1sAL, Fig. 3) [7,9].

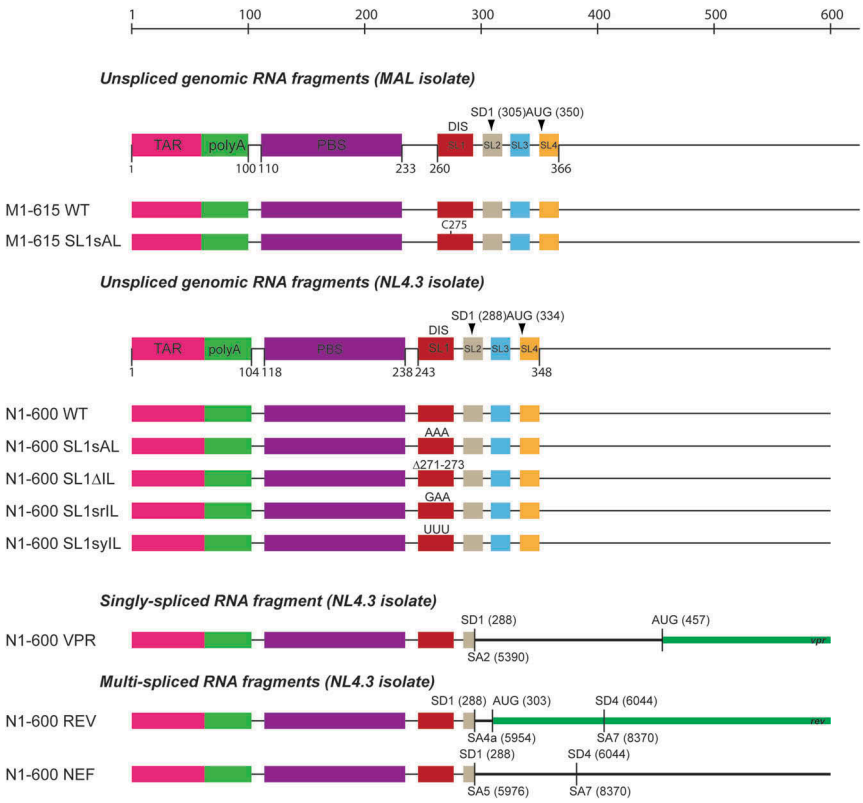

Figure 3. Schematic representation of the RNA fragments used in this study. The fragments whose names start with M and N are derived from the MAL and NL4-3 strains, respectively [13,14].

Those RNA fragments are defective in RNA dimerization under the buffer conditions and at the RNA concentration used in this study [13,14]. Mutations in the SL1 apical loop had little or no effect on the binding of GagΔp6, as our analysis displayed one class of binding sites, with binding affinity (K_d ~ 11–12 nM, Table 1) and stoichiometry (about 4 proteins, Table 1) very similar to gRNA. This is in contrast with the effect of these mutations on $Pr55^{Gag}$ binding, which resulted in loss of the class of high-affinity binding sites and the appearance of a of low affinity binding class (K_{d3} ~ 65 nM, Table 1) in addition to the moderate affinity binding class (K_{d2} ~ 12 nM, Table 1) [14]. Furthermore, mutations in the SL1 apical loop did not significantly alter the GagΔp6 binding stoichiometry (3–4 GagΔp6/RNA, Table 1), while it reduced the number of $Pr55^{Gag}$ molecules bound per RNA molecules from 9–10 to 3 (Table 1).

Next, we compared the interaction of GagΔp6 and $Pr55^{Gag}$ with RNA mutants in which purines of the internal loop of SL1 were substituted (AGG -> GAA, N1-600 SL1srIL, or AGG-> UUU, N1-600 SL1syIL) or deleted (N1-600 SL1ΔIL) (Figs. 3 and 5A). Importantly, these mutations do not affect the RNA secondary structure, nor abrogate the RNA dimerization under our experimental conditions [13,14]. Mutations in the SL1 internal loop had little or no effect on GagΔp6 binding, which displayed one class of binding sites (Fig. 5, left panels) with an affinity and a stoichiometry similar to gRNA (Fig. 5, left panels and Table 1). At the opposite, no binding of $Pr55^{Gag}$ could be detected when the purines of the internal loop were deleted (N1-600 SL1ΔIL RNA) or substituted by pyrimidines (N1-600 SL1syIL RNA) (Fig. 5, right panels). Similarly, substituting GAA for AGG in this loop (N1-600 SL1srIL RNA) dramatically reduced binding of the full-length $Pr55^{Gag}$ bound with moderate affinity (Fig. 5, right panels and Table 1) [14]. Thus, while GagΔp6 has a lower affinity than $Pr55^{Gag}$ for the wild type 5′-end region of HIV-1 gRNA, it has a higher affinity than the full-length Gag precursor for the SL1 internal loop mutants.

In addition, to gain a better understanding of the interaction between GagΔp6 and Psi, we compared the interactions between GagΔp6 and $Pr55^{Gag}$ with truncated viral RNA fragments partially including Psi (Supplementary Fig. 1A). In contrast with $Pr55^{Gag}$, truncations in the 5′-end region of the HIV-1 gRNA did not impact the GagΔp6 binding affinity, implying that the interaction between GagΔp6 and this gRNA portion is non-specific (Supplementary Table 1 and Supplementary Fig. 1).

2.4. GagΔp6 and $Pr55^{Gag}$ binding to the Psi individual elements

We next compared binding of GagΔp6 and $Pr55^{Gag}$ to each of the four stem-loops located in the Psi region (Fig. 1A). Our data with these short chemically synthesized RNA fragments (14 to 35 nucleotides) revealed that GagΔp6 bound these stem-loops with a very similar affinity (K_d ~ 21–29 nM) and stoichiometry (n = ~ 3) (Supplementary Fig. 2 and Table 2). The full-length precursor also displayed only one class of binding sites for these individual stem-loops (Table 2). However, $Pr55^{Gag}$ displayed a preferential binding to SL1, since the affinity for this motif was found to be 3- to 8-fold higher compared to the affinity determined for SL2 to SL4 [14]. Altogether these results further confirm that while $Pr55^{Gag}$

Table 1. GagΔp6 and Pr55^{Gag} binding to the first 600/615 nts of the genomic RNA and SL1 mutant RNAs. On the left, the binding parameters derived from the single binding site model [82] and from the stoichiometry analysis [14] (see Methods) for GagΔp6 interacting with RNA fragments corresponding to the first 600/615 nts of gRNA and SL1 mutant RNAs are indicated. On the right, the binding parameters determined for Pr55^{Gag} in interaction with the same RNA fragments [14]. K_{di} (i = 1, 2, 3) correspond to the three different classes of binding affinity. n.d. stands for not determined values. Mean ± SD of at least three independent experiments.Gag.

	GagΔp6		Pr55^{Gag}					
		Stoichiometry				Stoichiometry		
RNA	Kd (nM)		Kd1 (nM)	Kd2 (nM)	Kd3 (nM)	Site 1	Site 2	Site 3
N1-600 WT	12.1 ± 0.8	3.7 ± 0.5	1.8 ± 0.4	14.6 ± 3.8		6.9 ± 0.7	3.1 ± 0.8	
M1-615 WT	12.0 ± 1.9	3.5 ± 0.5	2.9 ± 0.6	15.4 ± 3.4		6.7 ± 1.1	3.0 ± 0.9	
N1-600 sAL	12.1 ± 1.3	3.8 ± 0.3		9.5 ± 2.4	62.5 ± 2.4		2.1 ± 0.4	1.0 ± 0.3
M1-615 sAL	11.3 ± 1.6	4.0 ± 1.2		12.5 ± 2.7	65.5 ± 6.6		1.9 ± 0.2	1.1 ± 0.2
N1-600 srIL	11.1 ± 1.9	3.8 ± 0.2			53.6 ± 9.1			2.4 ± 0.2
N1-600 syIL	12.1 ± 1.9	3.6 ± 0.5	n.d.	n.d.	n.d.			
N1-600 ΔIL	11.3 ± 1.8	3.8 ± 0.6	n.d.	n.d.	n.d.			

Table 2. GagΔp6 and Pr55^{Gag} binding to the individual stem-loops of the Psi region. On the left, binding parameters derived from the single binding site model [82] and from the stoichiometry analysis [14] (see Methods) for GagΔp6 in interaction with individual stem-loops of Psi. On the right, the binding parameters determined for Pr55^{Gag} in interaction with the same RNA fragments [14]. Mean ± SD of at least three independent experiments.

		GagΔp6		Pr55^{Gag}	
RNA	Length (nts)	Kd (nM)	Stoichiometry	Kd (nM)	Stoichiometry
N35 SL1	35	21 ± 2	3.2 ± 0.4	5 ± 2	3.50 ± 0.65
NSL2	17	22 ± 4	3.1 ± 0.6	44 ± 4	2.17 ± 0.49
NSL3	14	29 ± 1	2.8 ± 0.5	21 ± 2	2.33 ± 0.67
NSL4	24	25 ± 3	2.9 ± 0.4	18 ± 4	2.35 ± 0.78

preferentially binds SL1, GagΔp6 does not discriminate between the four hairpins located in Psi.

2.5. GagΔp6 and Pr55^{Gag} interaction with spliced viral RNAs and non-viral RNAs

Finally, we compared GagΔp6 and Pr55^{Gag} binding to several spliced viral RNAs (N1-600 NEF, N1-600 VPR, and N1-600 REV) (Fig. 3) [72]. GagΔp6 showed one class of binding sites of moderate affinity (K_d ~ 11–14 nM, Table 3, Fig. 6) for those RNAs with a stoichiometry of about 3 proteins per RNA molecule (Table 3). Of note, Pr55^{Gag} bound the same RNA fragments with similar affinity and stoichiometry (K_d ~ 12–20 nM, and n ~ 3, Table 3) [14].

Finally, we analyzed the binding parameters of GagΔp6 towards non-viral RNAs such as 7SL RNA which is known to be packaged into HIV-1 viral particle [32,33], and the 3′UTR of APOBEC3G mRNA [73]. GagΔp6 displayed only one class of binding sites for those non-viral RNAs (Fig. 7) with a K_d of about 14 nM and 10.6 nM for the 7SL RNA and the 3′UTR of APOBEC3G mRNA, respectively (Table 3). Both RNAs were bound by 3 GagΔp6 proteins (Table 3). Pr55^{Gag} bound these RNAs similarly (K_d ~ 18–20 nM) with a stoichiometry of 2–3 proteins (Table 3) [14]. These data show that deletion of p6 leads to very similar binding parameters for viral and non-viral RNA fragments, thus resulting in non-specific RNA binding properties.

2.6. The C-terminal Hist-tag does not affect the binding specificity of Pr55^{Gag} and GagΔp6

To test whether the presence of a His-tag in the Gag proteins affects their binding specificity to RNAs, we compared binding of His tagged Pr55^{Gag} and GagΔp6 proteins with Pr55^{Gag} and GagΔp6 proteins from which the His-tag had been removed by TEV (Tobacco Etch Virus) protease (Supplementary Figs. 3 and 4) to a representative panel of RNA fragments, including N1-600WT RNA, SL1 mutants N1-600 sAL and N1-600 srIL RNAs, and spliced viral N1-600 VPR RNA (Table 4 and Fig. 8). The impact of cleavage of the C-terminal His-tag on GagΔp6 binding to the RNA fragments resulted in increased Kd values (~ 2.5-fold) compared to the ones obtained with the His-tagged GagΔp6. Similarly, the TEV-cleaved Pr55^{Gag} displayed ~ 2.5-fold increased Kd values for the two classes of binding sites to N1-600 WT RNA, compared to the His-tagged Pr55^{Gag} protein (Fig. 8). Binding to N1-600 sAL RNA showed a single class of binding sites corresponding to the lower affinity component previously observed for the His-tagged Pr55^{Gag} (Kd ~ 60 nM,

Table 3. GagΔp6 and Pr55^{Gag} binding to the spliced viral and cellular RNA species. On the left, binding parameters derived from the single binding site model [82] and from the stoichiometry analysis [14] (see Methods) for GagΔp6 in interaction with RNA fragments corresponding to spliced viral and cellular RNAs. On the right, binding parameters determined for Pr55^{Gag} in interaction with the same RNAs fragments [14]. K_{di} (i = 1, 2) correspond to the two different classes of binding affinity. Mean ± SD of at least three independent experiments.

		GagΔp6		Pr55^{Gag}			
			Stoichiometry			Stoichiometry	
RNA	Length (nts)	Kd (nM)		Kd1 (nM)	Kd2 (nM)	Site 1	Site 2
N1-600 WT	600	12.1 ± 0.8	3.7 ± 0.5	1.8 ± 0.4	14.6 ± 3.8	6.9 ± 0.7	3.1 ± 0.8
N1-600 NEF	600	11.0 ± 0.5	2.8 ± 0.4		16.3 ± 4.2		2.7 ± 0.6
N1-600 REV	600	13.7 ± 1.7	3.4 ± 0.5		11.7 ± 2.0		3.5 ± 0.8
N1-600 VPR	600	12.4 ± 0.9	3.5 ± 0.1		18.1 ± 3.9		3.0 ± 0.2
A3G 3′UTR	325	10.6 ± 0.4	3.8 ± 0.4		18.5 ± 4.1		3.3 ± 0.6
7SL	300	14.5 ± 1.7	3.1 ± 0.2		20.5 ± 1.1		2.2 ± 0.1

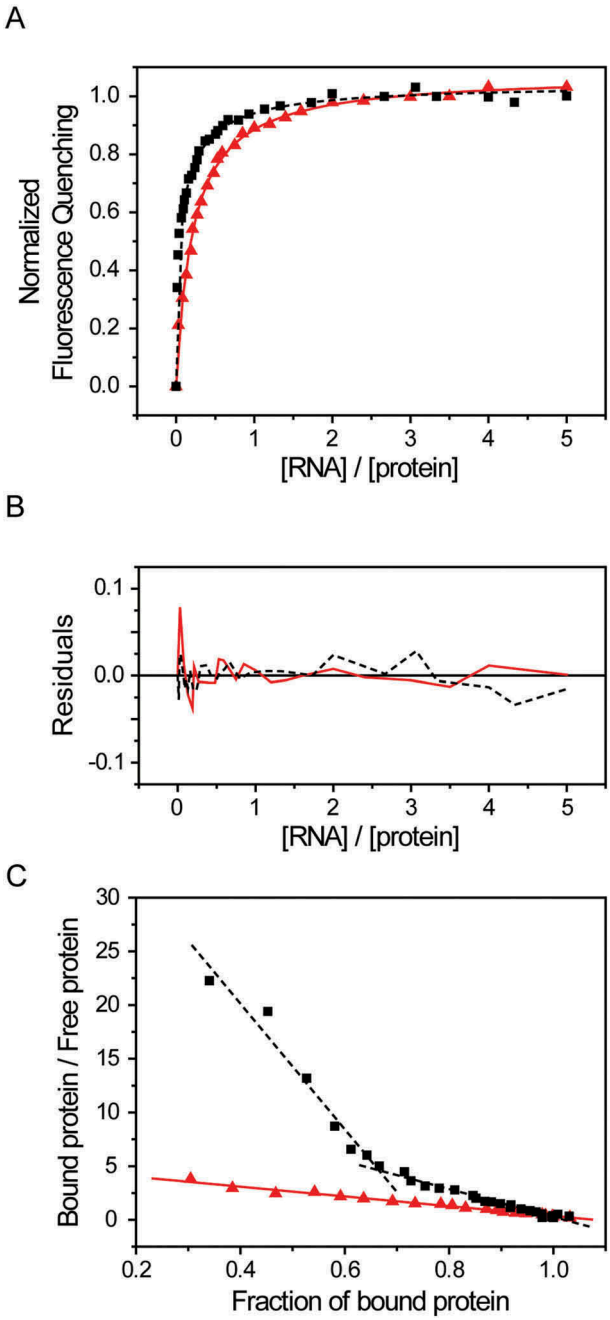

Figure 4. Representative experiments of GagΔp6 and Pr55^{Gag} binding to the first 600 nucleotides of gRNA. (A) Increasing concentrations of RNA were added to 50 nM of protein. The binding curves corresponding to GagΔp6 binding to N1-600 WT (red triangles) were best fitted according to a single binding site Scatchard model [82]. The interaction of Pr55^{Gag} with N1-600 WT RNA (black squares) was fitted with a Scatchard-like equation corresponding to a two binding sites-model as previously described [14]. **(B)** The residual plots for each curve fitted in **A** are represented. **(C)** On the left the Scatchard plot of GagΔp6 interaction with N1-600 WT RNA fragment (red triangles) yielded a single linear pattern. Conversely, on the right Pr55^{Gag} interaction with this same RNA yielded two linear patterns (black squares).

Table 4). Finally, TEV-cleaved Pr55^{Gag} displayed rather similar values for N1-600 srIL and N1-600 VPR RNAs compared to the His-tagged Pr55^{Gag} (Table 4). Altogether these results demonstrate that even though removal of the His-tag results in a general increase of the Kd values, this tag has a significant effect neither on the specific binding of Pr55^{Gag} to N1-600 WT RNA nor on the complete lack of RNA binding specificity of GagΔp6.

Finally, to assess the impact of ionic strength on RNA binding, we also tested His-tagged and TEV-cleaved GagΔp6 and Pr55^{Gag} proteins binding to N1-600 WT under more physiological ionic conditions (30 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$ and 4.8 mM spermidine). The resulting binding parameters were found to be similar to the ones obtained in our standard binding buffer conditions (30 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) (Supplementary Table 2).

3. Discussion

It is usually assumed that the p6 domain of Pr55^{Gag} is not involved in its RNA binding properties, and GagΔp6, which is easier to produce than the full-length precursor is frequently used as a surrogate for this protein in RNA binding studies [60–66]. However, an NMR study of NCp15 [67] and a recent ITC analysis of Pr55^{Gag} and GagΔp6 binding to short oligoribonucleotides [68] suggested a different picture. To dissect the possible impact of the C-terminal p6 domain of Pr55^{Gag} on the specific recognition of HIV-1 gRNA, we systematically compared the binding of GagΔp6, and full-length Pr55^{Gag} with a panel of wild type and mutant RNAs, also including spliced viral and cellular RNAs, which Pr55^{Gag} must discriminate during the specific gRNA selection and packaging [27,30,32,33].

Unlike full-length Pr55^{Gag}, which bound HIV-1 gRNA with a ten-fold higher affinity than spliced viral and cellular RNAs (Tables 1 and 3), GagΔp6 displayed a very narrow range of affinities (K_d ~ 11–14 nM; Tables 1 and 3) for all tested RNA fragments, therefore revealing a complete lack of binding specificity. In addition, while about 3–4 GagΔp6 molecules bound to any of these RNAs, significantly more Pr55^{Gag} molecules bound to the 5′-end region of the HIV-1 gRNA than to spliced viral or cellular RNAs (~ 10 versus ~ 2–3; Tables 1 and 3). Importantly, our present findings are in complete agreement with previous publications. On the one hand, preferential binding of Pr55^{Gag} to HIV-1 gRNA was reported using fluorescence spectroscopy (this work and Bernacchi et al. [14]), as well as filter binding, band-shift, and footprinting assays [13]. On the other hand, fluorescence anisotropy [43] and FCS [74] studies showed that binding of GagΔp6 is non-specific, except in the presence of very high salt concentrations or competitors. The discrepancies between these studies could have various origins: the methods used, the experimental protocols, the RNA fragments, as well as the proteins used (Pr55^{Gag} *vs.* GagΔp6, and His-tagged *vs.* non-tagged proteins); hence, the importance of conducting a comparative study under strictly identical conditions.

Altogether, our results indicate that Pr55^{Gag} and GagΔp6 bind RNAs in very different ways, suggesting that GagΔp6 is not an adequate surrogate for Pr55^{Gag} when looking at specific interaction with HIV-1 gRNA. Of note, the Pr55^{Gag} and GagΔp6 proteins used in most of our experiments were His-tagged at their C-termini (Tables 1–3 and Figs. 2–7); however our experiments with TEV-cleaved Gag proteins showed that the His-tag does not impact the specificity of Pr55^{Gag} and GagΔp6 binding to viral RNAs (Table 4 and Fig. 8); therefore, the differences observed cannot be attributed to this tag. In

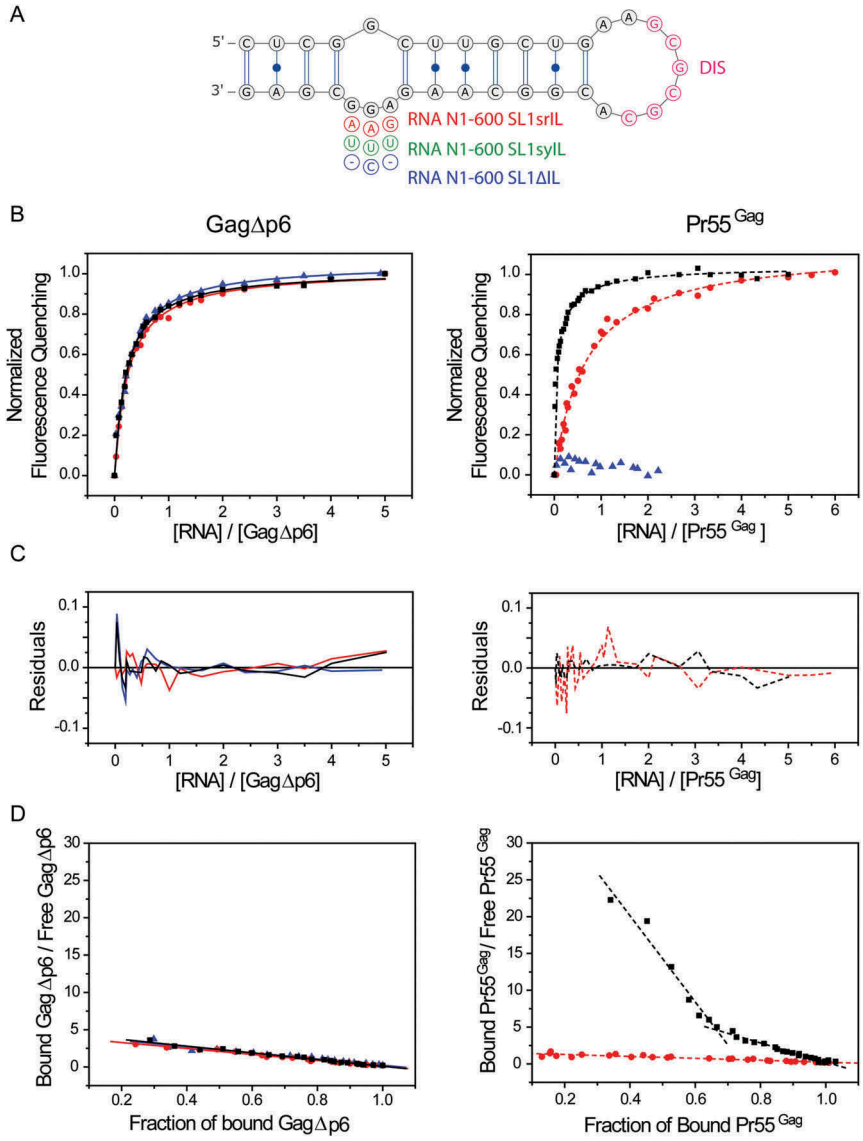

Figure 5. Representative experiments of GagΔp6 and Pr55^{Gag} binding to SL1 internal loop mutants. (A) The internal loop mutants of SL1 used in this study. The DIS sequence within SL1 apical loop is highlighted in magenta. Increasing concentrations of RNA were added to 50 nM of protein. **(B)** On the left hand side, the data corresponding to GagΔp6 binding to N1-600 WT (black squares), N1-600 SL1srIL (red circles), and N1-600 SL1ΔIL RNAs (blue triangles) were best fitted according to the single binding site model [82]. On the right hand side, the binding curves corresponding to Pr55^{Gag} interaction with N1-600 WT RNA (black squares) were fitted with a two binding sites-model as previously described [14]. N1-600 SL1srIL (red circles) was best fitted with the single binding site model, while data corresponding to N1-600 SL1DIL (blue triangles) could not be fitted. **(C)** The corresponding residual plots for each curve fitted in **B** are represented. **(D)** On the left, we observed for GagΔp6 interaction with N1-600 WT (black squares), N1-600 SL1srIL (red circles), and N1-600 SL1ΔIL (blue triangles) RNAs single linear patterns. On the right, the Scatchard plots of Pr55^{Gag} interaction with N1-600 WT RNA yielded two linear patterns (black squares), while a single linear pattern for N1-600 SL1srIL RNA (red circles) was observed.

line with these results, recent reports showed that while a His-tag can affect the RNA binding properties of Pr55^{Gag} at low ionic strength, it has negligible effect under the ionic conditions used in our study [75] and the same pattern of Pr55^{Gag} binding to short nucleic acids was observed with untagged and His-tagged Pr55^{Gag} [68].

We previously showed that SL1 plays a key role in the specific binding of Pr55^{Gag} to HIV-1 gRNA [13,14,71], with the internal loop and the apical loop which mediates gRNA dimerization, both contributing to specificity. However, substituting or deleting the SL1 internal loop or preventing RNA dimerization by mutating the SL1 apical loop had no effect on GagΔp6 binding, while the same mutations drastically reduced or even abolished binding of Pr55^{Gag} (Table 1). Indeed, GagΔp6 bound the four hairpins present in Psi equally well (Table 2), whereas Pr55^{Gag} has a higher affinity for SL1 and requires the complete 5′-end region of the HIV-1 gRNA for optimal binding. Interestingly, the p6 domain increased the specificity of Pr55^{Gag} not only by increasing its affinity for gRNA, but also by decreasing its affinity for mutant RNAs (especially those in the SL1 internal loop (Table 1) and cellular RNAs (Table 3).

Nevertheless, it seems unlikely that the Pr55^{Gag} binding specificity is regulated by a direct interaction of p6 with gRNA, as this domain contains seven conserved Glu residues, conferring to p6 a high negative charge density. Interestingly, NMR studies on the first protease-induced maturation product, NCp15 (NC-sp2-p6) (Fig. 1B), suggested that the acidic

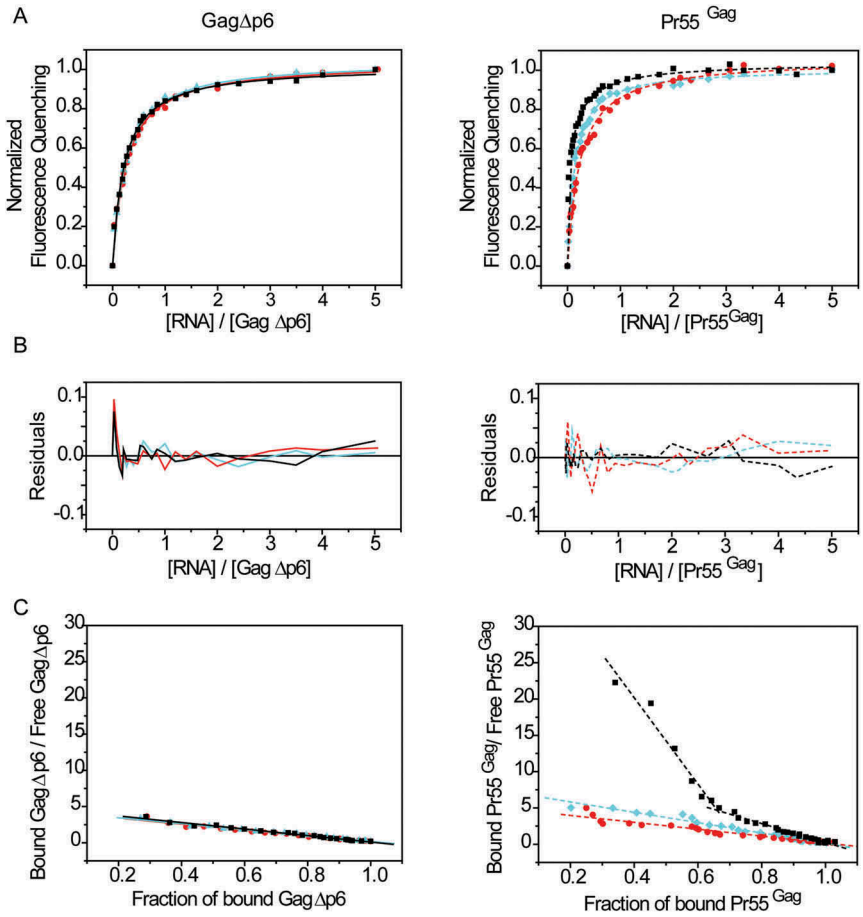

Figure 6. Representative experiments of GagΔp6 and Pr55^{Gag} binding to genomic and spliced viral RNA fragments. Increasing concentrations of RNA were added to 50 nM of protein. **(A)** On the left, the binding curves corresponding to GagΔp6 binding to N1-600 WT (black squares), N1-600 VPR (red circles) and N1-600 REV (cyan circles) RNAs were fitted according to the single binding site model. On the right, the data of Pr55^{Gag} in interaction with N1-600 WT RNA (black squares) were fitted with a two binding sites-model as previously described [14], while data corresponding to Pr55^{Gag} interaction with N1-600 VPR (red circles) and N1-600 TAT (cyan circles) RNAs were best fitted with a single binding site model [82]. **(B)** The corresponding residual plots for each curve fitted in **A** are represented. **(C)** On the left, single linear patterns were observed for GagΔp6 interaction with N1-600 WT (black squared), N1-600 VPR (red circles) and N1-600 REV (cyan circles) RNAs. On the right, the Scatchard plots of Pr55^{Gag} interaction with N1-600 WT RNA yielded two linear patterns (black squares), while single linear patterns were observed for N1-600 VPR (red circles) and N1-600 TAT (cyan circles) RNAs.

p6 domain folds back and interacts with the basic zinc fingers motifs of the NC domain [67]. We propose that the same interaction may take place in the context of full-length Pr55^{Gag} and increase the RNA binding specificity by partially masking the NC positive charges, and likely increasing steric selection (Fig. 9). Thus, this comparative analysis of Pr55^{Gag} and GagΔp6 proteins performed under strictly identical conditions unveiled a new role for the p6 domain in the selective binding of Pr55^{Gag} to the 5′-end region of the HIV-1 gRNA. Our study thus strongly suggests that the p6 domain contributes to the specific selection/discrimination of gRNA from spliced viral RNAs and cellular RNAs by Pr55^{Gag} required for its selective packaging into viral particles.

4. Material and methods

4.1. Pr55^{gag} and GagΔp6 protein expression and purification

The full-length Pr55^{Gag} and the GagΔp6 proteins were expressed and purified as described previously [76]. Protein samples were ultra-centrifuged for 1 h at 100,000 g at 4°C immediately prior use and the upper half of the samples was then carefully transferred into a protein low-binding microtube (Eppendorf) for further experiments.

The plasmid containing the full-length wild type Pr55^{Gag} gene with a non-cleavable hexa-histidine sequence at the C-terminus of the gene in an engineered pET28a vector was used as the template to generate the TEV cleavable GagΔp6-TEV-His and Pr55^{Gag}-TEV-His constructs. These proteins containing the TEV sequence were digested with 1:25 (w/w) TEV protease (His-tagged produced in-house) at 4°C for 14 hrs. The TEV digested proteins were applied to a 1 ml Ni-NTA column and the flow through was collected and passed over the column two more times. The cleaved His-Tag, His-TEV and uncut fusion proteins were eluted from the column using the elution buffer (1.0 M NaCl, 50 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 250 mM imidazole, 1% (v/v) Tween-20, 10% (v/v) glycerol, 5 mM DTT). His-tag cleavage from Pr55^{Gag} and GagΔp6 proteins was checked by western blot and mass spectrometry analyses (Supplementary Figs. 3 and 4).

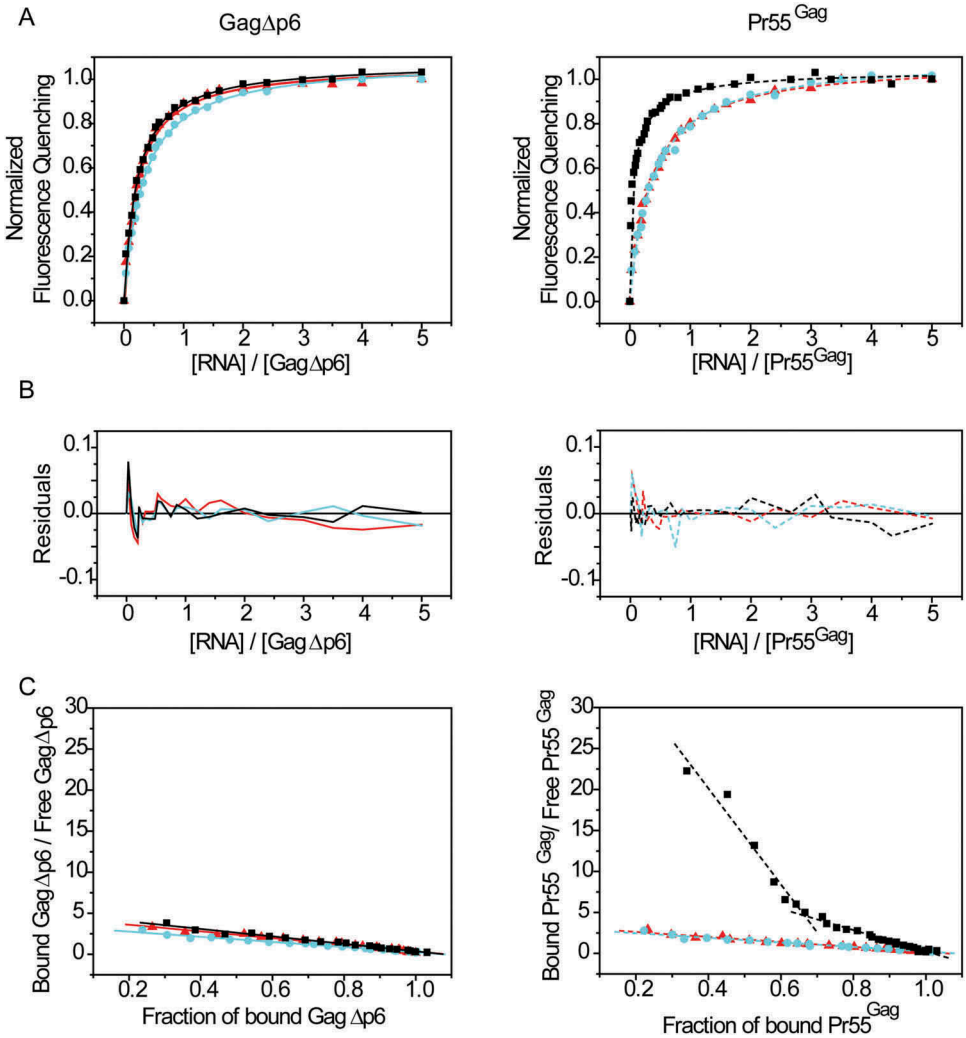

Figure 7. Representative experiments of GagΔp6 and Pr55^{Gag} binding to gRNA and non-viral RNA species. Increasing concentrations of RNA were added to 50 nM of protein. **(A)** On the left, the binding curves corresponding to GagΔp6 binding to N1-600 WT (black squares), APOBEC3G 3'UTR (red triangles) and 7SL (cyan circles) RNAs were fitted according to the single binding site model [82]. On the right, the data of Pr55^{Gag} in interaction with N1-600 WT RNA (black squares) were fitted with a two binding sites-model as previously described [14]. APOBEC3G 3'UTR (red triangles) and 7SL (cyan circles) RNAs were best fitted with the single binding site model. **(B)** The corresponding residual plots for each curve fitted in **A** are represented. **(C)** On the left, single linear patterns were observed for GagΔp6 in interaction with N1-600 WT (black squares), APOBEC3G 3'UTR (red triangles) and 7SL (cyan circles) RNAs. On the right, the Scatchard plots of Pr55^{Gag} interaction with N1-600 WT RNA yielded two linear patterns (black squares), while single linear patterns were observed for in interaction with APOBEC3G 3'UTR (red triangles) and 7SL (cyan circles) RNAs.

4.2. Western blots

One µg of His-tagged or TEV-cleaved Pr55^{Gag} and GagΔp6 proteins was loaded on a 4–12% Criterion TGX 4–15% gels (Bio-Rad) and transferred to a 22 µm PVDF membranes using the Trans-Blot® Turbo™ Transfer System (Bio-Rad). Gag proteins were detected with a HIV-positive patient serum and anti-His monoclonal antibody (sc-8036, Santa Cruz Biotechnology,) followed by horseradish peroxidase conjugated anti-human (NA933, GE Healthcares) or anti-mouse antibodies (170–6516, Bio-Rad), respectively. Proteins were then visualized by chemiluminescence using the ECL Prime Western blotting detection reagent (GE Healthcares) and analyzed with the ChemiDoc™ Touch Imaging System (Bio-Rad).

4.3. Mass spectrometry

Prior mass analysis, protein samples were dialyzed against 200 mM ammonium acetate in a Slide-A-Lyzer™ MINI Dialysis Device 7000 NMWL. Mass spectrometry was performed on an ESI-TOF mass spectrometer Synapt G2-s (Waters, MA, USA). Mass analysis was performed under denaturing conditions: the protein was diluted in 50/50 water acetonitrile (v/v) mixture acidified with 1% formic acid to achieve a final concentration of 0.3 µM. Data were acquired in the positive ionization mode from 500 to 2000 m/z. The ionization conditions were 30 V for the sample cone, 3.0 kV for the capillary and the source temperature was set to 130°C. Multiple charge spectrum were deconvoluted using Waters MassLynx MaxEnt1 software and protein molecular weights were calculated using the ProtParam tool (https://web.expasy.org/protparam/).

Table 4. His-tagged and TEV-cleaved GagΔp6 and Pr55^{Gag} proteins binding to the first 600/615 nts of the gRNA, SL1 mutant RNAs and N1-600 VPR spliced viral RNA. On the left side, binding parameters derived from the single binding site model[82] and from the stoichiometry analysis[14] (see Methods) for GagΔp6 His-tagged and TEV-cleaved in interaction with RNA fragments corresponding to the first 600/615 nts of the gRNA, N1-600 srlL and N1-600 sAL, SL1 mutant RNAs, and N1-600 VPR spliced viral RNA. On the right side, binding parameters determined for Pr55^{Gag} His-tagged[14], and for Pr55^{Gag} TEV-cleaved in interaction with the same RNA fragments. K_{di} (i = 1, 2) correspond to the two different classes of binding affinity. Mean ± SD of three independent experiments.

	GagΔp6				Pr55^{Gag}			
	His-tagged	TEV- cleaved			His-tagged	TEV-cleaved		
RNA	Kd (nM)	Kd (nM)	Kd1 (nM)	Kd2 (nM)	Kd3 (nM)	Kd1 (nM)	Kd2 (nM)	Kd3 (nM)
N1-600 WT	12.1 ± 0.8	33.7 ± 2.4	1.8 ± 0.4	14.6 ± 3.8		4.2 ± 0.9		32.5 ± 5.8
N1-600 sAL	12.1 ± 1.3	22.7 ± 2.5		9.5 ± 2.4	62.5 ± 2.4			60.8 ± 4.2
N1-600 srlL	11.1 ± 1.9	31.2 ± 3.6			53.6 ± 9.1			40.1 ± 4.7
N1-600 VPR	11.3 ± 1.8	28.6 ± 4.1		18.1 ± 3.9			24.3 ± 2.1	

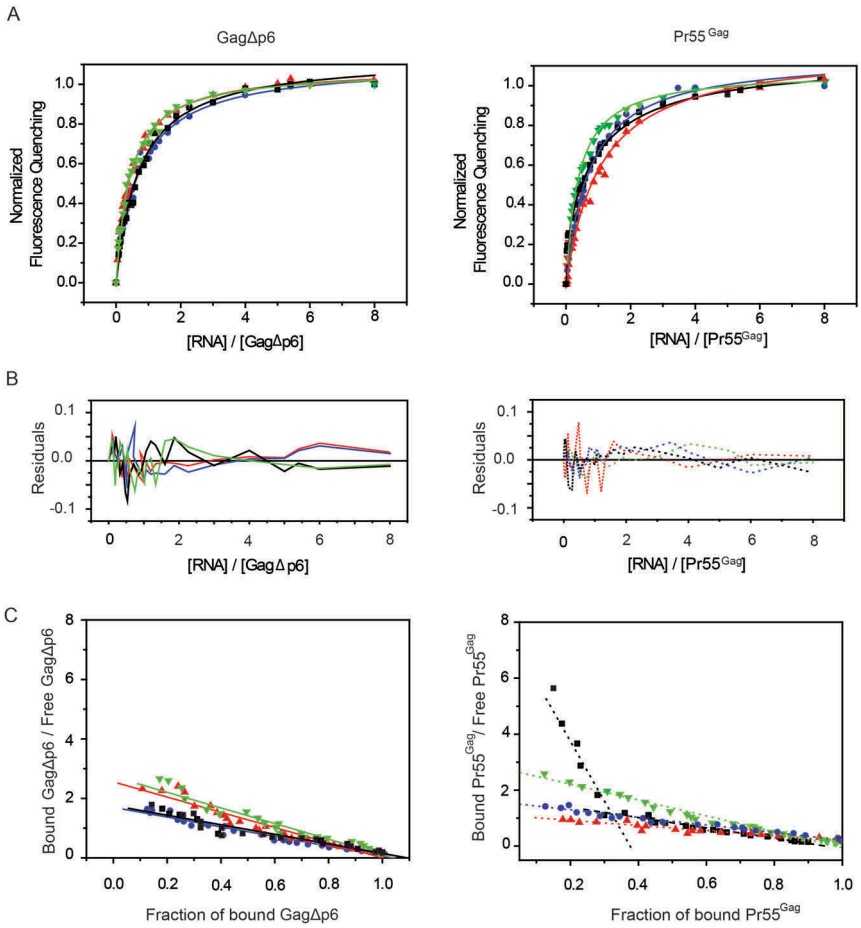

Figure 8. Representative experiments of TEV-cleaved GagΔp6 and Pr55^{Gag} proteins binding to gRNA, SL1 mutant RNAs and N1-600 VPR spliced viral RNA Increasing concentrations of RNA were added to 50 nM of protein. **(A)** On the left side, the binding curves corresponding to TEV-cleaved GagΔp6 binding to N1-600 WT (black squares), N1-600 sAL (red triangles), N1-600 srlL (blue circles), and N1-600 VPR (green triangles) RNAs were fitted according to the single binding site model [82]. On the right side, the data of TEV-cleaved Pr55^{Gag} in interaction with N1-600 WT RNA (black squares) were fitted with a two binding sites-model, as previously described [14]. N1-600 sAL (red triangles), N1-600 srlL (blue circles), and N1-600 VPR (green triangles) RNAs were best fitted with the single binding site model. **(B)** The corresponding residual plots for each curve fitted in **A** are represented. **(C)** On the left side, single linear patterns were observed for TEV-cleaved GagΔp6 in interaction with N1-600 WT (black squares), N1-600 sAL (red triangles), N1-600 srlL (blue circles), and N1-600 VPR (green triangles) RNAs. On the right side, the Scatchard plots of Pr55^{Gag} interaction with N1-600 WT RNA yielded two linear patterns (black squares), while single linear patterns were observed for in interaction with N1-600 sAL (red triangles), N1-600 srlL (blue circles), and N1-600 VPR (green triangles) RNAs.

4.4. Dynamic light scattering (DLS) analysis

GagΔp6 and Pr55^{Gag} samples were characterized by DLS. Proteins were diluted to a final concentration of 5 µM in the storage buffer (50 mM Tris-HCl pH 8, 1 M NaCl, 5 mM DTT,) or in the binding buffer (30 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$). Intensities of the scattered light and correlation times were measured using a Zetasizer^{TM} Nano S apparatus (4 mW He-Ne laser, λ_0 = 633 nm, scattering angle θ = 173°) (Malvern, UK). Measurements were performed at 20°C in a single 50 µl trUView cuvette (BioRad Laboratories, CA USA). Variations of the diffused light intensity were recorded at microsecond

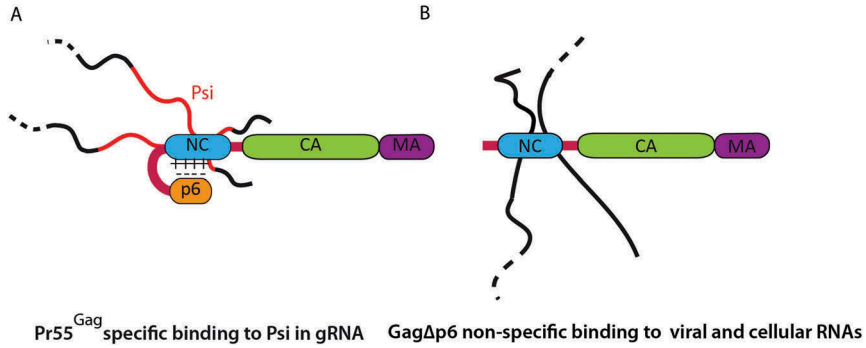

Figure 9. Model of Pr55^{Gag} binding specificity to gRNA. (A) The C-terminal acidic p6 domain of Pr55^{Gag} folds over the highly basic NC domain [67], thus partially masking the NC positive charges. According to our model this would promote the specific binding of Pr55^{Gag} to gRNA. **(B)** Our data show that the Pr55^{Gag} p6 domain deletion results in a non-specific binding to all the tested viral and cellular RNAs, since GagΔp6 displayed a similar affinity for all of them.

time intervals. An autocorrelation function was derived, allowing the determination of the translational diffusion coefficients (D). Assimilating the proteins in solution to spheres, the diffusion coefficients were related to the hydrodynamic radius (R_h) of the molecules populations present in solution, *via* the Stokes-Einstein equation:

$$D = \frac{kT}{R_h 6\pi\mu} \quad (1)$$

in which k is the Boltzmann constant, T is the temperature and μ is the viscosity of the solvent. All experimental data were corrected for solvent viscosity and refractive index. In our experimental settings, solvent viscosity was 1.104 cP and 0.9891 cP for storage and binding buffer, respectively. Solvent refractive index was 1.341 and 1.333 for storage and binding buffer, respectively, as estimated using Malvern Zetasizer Software calculator (Malvern, UK).

4.5. Static light scattering (SLS) analysis

The molecular mass of GagΔp6 and Pr55^{Gag} in solution, in storage or in binding buffer, was determined by SLS. The intensities of scattered light were measured using a DynaPro Nanostar^{TM} (100 mW He-Ne laser) (Wyatt Technologies,) in a 1 μl quartz cuvette (JC-006, Wyatt Technologies). The intensity of the scattered light I_{SL} is a function of the particles size and is proportional to the particles concentration and their mass. Thus I_{SL} can be used to derive the mass of the particles in solution through Zimm's development of Rayleigh equation [77,78]:

$$\frac{KC}{R(\theta, C)} = \frac{1}{M_w P(\theta)} + 2A_2 C \quad (2)$$

in which R(θ,C) is the excess Rayleigh ratio of the solution as a function of scattering angle θ (with θ = 90° in our setup) and concentration C. C is the solute concentration, M_w is the weight-averaged solute molar mass, A_2 is the second virial coefficient, P(θ) is the angular dependence of sample scattering and K a constant defined as:

$$K = \frac{2\pi^2}{{\lambda_0}^4 N_A} \left(\frac{d_n}{d_c} n_0\right)^2 \quad (3)$$

where λ_0 is the laser wavelength (λ_0 = 633 nm), N_A is Avogadro's number, n_0 is the refractive index of the solvent and dn/dc the increment refractive index with sample concentration. The refractive index of each solvent was set as mentioned above in the DLS section. Before sample acquisition the offset of the solvent was measured for subsequent sample data treatment. All buffers were filtered using 0.02 μm filters (Millex ®) before analysis or sample dilution.

4.6. Plasmids, in vitro RNA transcription and purification

All plasmids used for the synthesis of wild type and mutant HIV-1 RNA fragments were previously described [7,9,13,14,72,79], as well as plasmids used for *in vitro* transcription of 7SL RNA and of the 3′UTR of APOBEC3G mRNA [14,73]. Linearized plasmids were used as templates for the synthesis of RNA fragments by *in vitro* run-off transcription using T7 bacteriophage RNA polymerase, followed by purification on a size-exclusion chromatography column [80]. The purified transcripts were folded as described previously [14]. Briefly, RNA fragments were prepared in Milli-Q (Millipore), denatured for 2 min at 90°C and snap-cooled on ice for 2 min. Proper folding was achieved by the addition of the binding buffer (30 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$) and incubation for 15 min at 37°C.

4.7. Synthetic HIV-1 RNA oligonucleotides

RNAs corresponding to the individual stem-loops of the Psi region were chemically synthesized and purified by reverse-phase HPLC and polyacrylamide gel electrophoresis (Integrated DNA Technologies Inc.). RNA oligonucleotides were then folded as described above.

4.8. Steady-state fluorescence spectroscopy

Prior to protein binding analysis, 1 μM RNA was folded as described above. Fluorescence measurements were performed in quartz cells at 20 ± 0.5°C on a Fluoromax-4 fluorimeter (HORIBA Jobin-Yvon Inc., NJ., USA). The excitation wavelength was set at 295 nm for selective excitation of tryptophan residues and the emission wavelength was scanned from 305

to 450 nm. The integration time was set on 0.1 s and the excitation and emission bandwidth on 5 nm. Increasing amounts of RNA were added to 50 or 100 nM GagΔp6 or $Pr55^{Gag}$ in the binding buffer, so that the RNA/protein ratio varied from 0 to 5. To assess the impact of ionic strength in the binding experiments, we also tested GagΔp6 and $Pr55^{Gag}$ binding to N1-600 WT under more standard physiological salt conditions (30 mM Tris-HCl pH 7.5, 1 mM $MgCl_2$, 150 mM NaCl, and 4.8 mM spermidine [81]) (see Results). After addition of each RNA aliquot, the quartz cell was rapidly homogenized and the fluorescence emission measured.

The emission spectra of each titration were integrated and the fluorescence intensities corrected for buffer fluorescence and dilution effects. In order to determine the binding parameters of GagΔp6 and $Pr55^{Gag}$ to each RNA fragment, the corrected fluorescence intensity I measured for any added RNA concentration was converted into binding density ν which corresponds to the protein bound, P_b, to nucleic acid concentration, A_t, ratio:

$$= \frac{P_b}{A_t} = \frac{(I_0 - I)}{(I_0 - I_F)} \frac{P_t}{A_t} \quad (4)$$

With I_0 corresponding to the protein fluorescence intensity in absence of nucleic acids, I_F to the fluorescence intensity at the end of the titration, and P_t to the total protein concentration.

Since:

$$\frac{P_b}{P_t} = \frac{(I_0 - I)}{(I_0 - I_F)} \quad (5)$$

$$P_f = P_t - P_b = P_t - A_t \quad (6)$$

The concentration of bound and free protein (P_b and P_f, respectively) can be calculated as functions of I, I_0, I_f, P_t and A_t using Equations (5) and (6). The experimental observed affinity constant K_{obs}, and thus the dissociation constant K_d, mathematically corresponding to its inverse, was then computed by fitting the experimental data to equation [82]:

$$= \frac{(K_{obs}\ P_f)}{(1 + K_{obs}\ P_f)} \quad (7)$$

Plots were fitted with equations corresponding to models with a single class or two classes of binding sites. To confirm the number of $Pr55^{Gag}$ and GagΔp6 binding site classes for viral and cellular RNAs, we then plotted the fraction of bound protein *vs*. the unbound protein fraction (Scatchard plots). The presence of one or two linear pattern(s) in these plots indicated the presence of one or two classes of RNA binding sites, respectively.

The analysis of fluorescence binding curves also allowed determination of the GagΔp6 and $Pr55^{Gag}$ binding stoichiometry (n), corresponding to the average number of proteins bound to one RNA molecule. The experimental data expressed as normalized fluorescence quenching were reported *vs*. the molar ratio of total [RNA] expressed in strands to [GagΔp6] or [$Pr55^{Gag}$]. The stoichiometry of the complexes could then be graphically recovered by the intersection of the initial slope at low [RNA]/[protein] ratio with the fluorescence plateau at the end of the titration, as previously described [14,83].

Acknowledgments

We thank Dr Bernard Lorber (UPR9002 CNRS, Université de Strasbourg) for assistance with DLS and SLS analyses. We thank Dr Han Tanwar (Deakin University) who has contributed to the production of recombinant $Pr55^{Gag}$ and GagΔp6 proteins.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by grants from: the
Agence Nationale de la Recherche sur le SIDA et les Hepatites Virales (ANRS) to SB, and a fellowship of the French Ministry of Higher Education and Research to ND.

Notes on contributor

ND performed DLS, SLS and fluorescence spectroscopy experiments, analyzed results, and generated the figures. KK, SG, WM and JM produced $Pr55^{Gag}$ and GagΔp6 proteins. TS performed western blots, and PW contributed with mass spectrometry analysis. JCP and RM contributed to the design of the study and the data analysis and revised the manuscript. JM contributed to revise the manuscript. SB designed and supervised the study, analyzed the data and drafted the manuscript.

ORCID

Keith K. Khoo http://orcid.org/0000-0002-3334-6564
William J. McKinstry http://orcid.org/0000-0001-9668-9364
Johnson Mak http://orcid.org/0000-0002-5229-5707
Jean-Christophe Paillart http://orcid.org/0000-0003-1647-8917
Roland Marquet http://orcid.org/0000-0002-4209-3976
Serena Bernacchi http://orcid.org/0000-0003-2824-8326

References

1. Mailler E, Bernacchi S, Marquet R, et al. The life-cycle of the HIV-1 Gag-RNA complex. Viruses. 2016;8(9). pii: E248. doi:10.3390/v8090248.
2. Comas-Garcia M, Davis SR, Rein A. On the selective packaging of Genomic RNA by HIV-1. Viruses. 2016;8(9). pii: E246. doi:10.3390/v8090246.
3. D'Souza V, Summers MF. How retroviruses select their genomes. Nat Rev Microbiol. 2005;3:643–655.
4. Kuzembayeva M, Dilley K, Sardo L, et al. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology. 2014;454:362–370.
5. Berkowitz R, Fisher J, Goff SP. RNA packaging. Curr Top Microbiol Immunol. 1996;214:177–218.
6. Russell RS, Liang C, Wainberg MA. Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? Retrovirology. 2004;1:23.
7. Paillart JC, Marquet R, Skripkin E, et al. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J Biol Chem. 1994;269:27486–27493.
8. Skripkin E, Paillart JC, Marquet R, et al. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc Natl Acad Sci U S A. 1994;91:4945–4949.

9. Paillart JC, Skripkin E, Ehresmann B, et al. A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. Proc Natl Acad Sci U S A. 1996;93:5572–5577.
10. Paillart JC, Marquet R, Skripkin E, et al. Dimerization of retroviral genomic RNAs: structural and functional implications. Biochimie. 1996;78:639–653.
11. Berkhout B, Ooms M, Beerens N, et al. In Vitro evidence that the untranslated leader of the HIV-1 Genome is an RNA checkpoint that regulates multiple functions through conformational changes. J Biol Chem. 2002;277:19967–19975.
12. Dubois N, Marquet R, J-C P, et al. Retroviral RNA dimerization: from structure to functions. Front Microbiol. 2018;9:527.
13. Abd El-Wahab EW, Smyth RP, Mailler E, et al. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. Nat Commun. 2014;5:4304.
14. Bernacchi S, Abd El-Wahab EW, Dubois N, et al. HIV-1 Pr55 (Gag) binds genomic and spliced RNAs with different affinity and stoichiometry. RNA Biol. 2017;14:90–103.
15. Lever A, Gottlinger H, Haseltine W, et al. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J Virol. 1989;63:4085–4087.
16. Aldovini A, Young RA. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J Virol. 1990;64:1920–1926.
17. Clever JL, Parslow TG. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J Virol. 1997;71:3407–3414.
18. McBride MS, Panganiban AT. Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. J Virol. 1997;71:2050–2058.
19. Russell RS, Hu J, Bériault V, et al. Sequences downstream of the 5′ splice donor site are required for both packaging and dimerization of human immunodeficiency Virus Type 1 RNA. J Virol. 2003;77:84–96.
20. Abbink TEM, Berkhout B. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. J Biol Chem. 2003;278:11601–11611.
21. Lu K, Heng X, Garyu L, et al. NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. Science. 2011;334:242–245.
22. Ooms M, Huthoff H, Russell R, et al. RNA dimerization and packaging in human immunodeficiency Virus Type 1 Virions. J Virol. 2004;78:10814–10819.
23. Keane SC, Summers MF. NMR studies of the structure and function of the HIV-1 5′-leader. Viruses. 2016;8(12). pii: E338. doi:10.3390/v8120338.
24. Wilkinson KA, Gorelick RJ, Vasa SM, et al. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLOS Biol. 2008;6:e96.
25. Heng X, Kharytonchyk S, Garcia EL, et al. Identification of a minimal region of the HIV-1 5′-leader required for RNA dimerization, NC binding and packaging. J Mol Biol. 2012;417:224–239.
26. Helga-Maria C, Hammarskjöld M-L RD. An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency Virus Type 1 Genomic RNA. J Virol. 1999;73:4127–4135.
27. Didierlaurent L, Racine PJ, Houzet L, et al. Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. Nucleic Acids Res. 2011;39:8915–8927.
28. Clever JL, Miranda D Jr, Tg P. RNA structure and packaging signals in the 5′ leader region of the human immunodeficiency Virus Type 1 Genome. J Virol. 2002;76:12381–12387.
29. Liang C, Hu J, Russell RS, et al. Spliced human immunodeficiency Virus Type 1 RNA is reverse transcribed into cDNA within infected cells. AIDS Res Hum Retroviruses. 2004;20:203–211.
30. Houzet L, Paillart JC, Smagulova F, et al. HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res. 2007;35:2695–2704.
31. Rulli SJ, Hibbert CS, Mirro J, et al. Selective and nonselective packaging of cellular RNAs in retrovirus particles. J Virol. 2007;81:6623–6631.
32. Keene SE, Telesnitsky A. cis-Acting determinants of 7SL RNA packaging by HIV-1. J Virol. 2012;86:7934–7942.
33. Onafuwa-Nuga AA, Telesnitsky A, King SR. 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles. RNA. 2006;12:542–546.
34. Muriaux D, Mirro J, Harvin D, et al. RNA is a structural element in retrovirus particles. Proc Natl Acad Sci U S A. 2001;98:5246–5251.
35. Telesnitsky A, Wolin SL. The host RNAs in retroviral particles. Viruses. 2016;8:235.
36. Bell NM, Lever AML HIV. Gag polyprotein: processing and early viral particle assembly. Trends Microbiol. 2013;21:136–144.
37. Chukkapalli V, Ono A. Molecular determinants that regulate plasma membrane association of HIV-1 Gag. J Mol Biol. 2011;410:512–524.
38. Hogue IB, Llewellyn GN, Ono A. Dynamic association between HIV-1 Gag and membrane domains. Mol Biol Int. 2012;2012:979765. doi:10.1155/2012/979765.
39. Cimarelli A, Sandin S, Höglund S, et al. Basic residues in human immunodeficiency Virus Type 1 Nucleocapsid promote virion assembly via interaction with RNA. J Virol. 2000;74:3046–3057.
40. Zhang Y, Qian H, Love Z, et al. Analysis of the assembly function of the human immunodeficiency virus Type 1 Gag protein nucleocapsid domain. J Virol. 1998;72:1782–1789.
41. Dannull J, Surovoy A, Jung G, et al. Specific binding of HIV-1 nucleocapsid protein to PSI RNA in vitro requires N-terminal zinc finger and flanking basic amino acid residues. EMBO J. 1994;13:1525–1533.
42. Geigenmüller U, Linial ML. Specific binding of human immunodeficiency virus type 1 (HIV-1) Gag-derived proteins to a 5′ HIV-1 genomic RNA sequence. J Virol. 1996;70:667–671.
43. Webb JA, Jones CP, Parent LJ, et al. Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging. RNA. 2013;19:1078–1088.
44. Chukkapalli V, Oh SJ, Ono A. Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. Proc Natl Acad Sci U S A. 2010;107:1600–1605.
45. Alfadhli A, McNett H, Tsagli S, et al. matrix protein binding to RNA. J Mol Biol. 2011;410:653–666.
46. Kutluay SB, Zang T, Blanco-Melo D, et al. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell. 2014;159:1096–1109.
47. Friedrich M, Setz C, Hahn F, et al. Glutamic acid residues in HIV-1 p6 regulate virus budding and membrane association of Gag. Viruses. 2016;8:117.
48. Garrus JE, von Schwedler UK, Pornillos OW, et al. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell. 2001;107:55–65.
49. Martin-Serrano J, Zang T, Pd B. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat Med. 2001;7:1313–1319.
50. Demirov DG, Ono A, Orenstein JM, et al. Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. Proc Natl Acad Sci U S A. 2002;99:955–960.
51. Strack B, Calistri A, Craig S, et al. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell. 2003;114:689–699.
52. von Schwedler UK1, Stuchell M, Müller B, et al. The protein network of HIV budding. Cell. 2003;114:701–713.
53. Fisher RD, Chung H-Y, Zhai Q, et al. Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. Cell. 2007;128:841–852.
54. Zhai Q, Fisher RD, Chung H-Y, et al. Structural and functional studies of ALIX interactions with YPX(n)L late domains of HIV-1 and EIAV. Nat Struct Mol Biol. 2008;15:43–49.

55. Kondo E, Mammano F, Cohen EA, et al. The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. J Virol. 1995;69:2759–2764.
56. Salgado GF, Marquant R, Vogel A, et al. Structural Studies of HIV-1 Gag p6ct and its interaction with Vpr determined by solution nuclear magnetic resonance. Biochemistry (Mosc). 2009;48:2355–2367.
57. Kondo E, Göttlinger HG. A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency Virus Type 1 Vpr. J Virol. 1996;70:159–164.
58. Welker R, Hohenberg H, Tessmer U, et al. Biochemical and structural analysis of isolated mature cores of human immunodeficiency Virus Type 1. J Virol. 2000;74:1168–1177.
59. Accola MA, Å Ö, Göttlinger HG. Isolation of human immunodeficiency virus type 1 cores: retention of Vpr in the absence of p6gag. J Virol. 2000;74:6198–6202.
60. Campbell S, Rein A. In Vitro assembly properties of human immunodeficiency Virus Type 1 Gag protein lacking the p6 domain. J Virol. 1999;73:2270–2279.
61. Wu T, Datta SAK, Mitra M, et al. Fundamental differences between the nucleic acid chaperone activities of HIV-1 nucleocapsid protein and Gag or Gag-derived proteins: biological implications. Virology. 2010;405:556–567.
62. Jones CP, Datta SAK, Rein A, et al. Matrix domain modulates HIV-1 Gag's nucleic acid chaperone activity via inositol phosphate binding. J Virol. 2011;85:1594–1603.
63. Munro JB, Nath A, Färber M, et al. Observed in single HIV-1 Gag molecules during In Vitro assembly of Virus-like particles. J Virol. 2014;88:3577–3585.
64. Datta SAK, Curtis JE, Ratcliff W, et al. Conformation of the HIV-1 Gag protein in solution. J Mol Biol. 2007;365:812–824.
65. Datta SAK, Heinrich F, Raghunandan S, et al. Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. J Mol Biol. 2011;406:205–214.
66. Kenyon JC, Prestwood LJ, Lever AML. A novel combined RNA-protein interaction analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader. Sci Rep. 2015;5:14369.
67. Wang W, Naiyer N, Mitra M, et al. Distinct nucleic acid interaction properties of HIV-1 nucleocapsid protein precursor NCp15 explain reduced viral infectivity. Nucleic Acids Res. 2014;42:7145–7159.
68. Tanwar HS, Khoo KK, Garvey M, et al. The thermodynamics of Pr55Gag-RNA interaction regulate the assembly of HIV. PLOS Pathog. 2017;13:e1006221.
69. Datta SAK, Zhao Z, Clark PK, et al. Interactions between HIV-1 Gag molecules in solution: an inositol phosphate-mediated switch. J Mol Biol. 2007;365:799–811.
70. Chen Y, Barkley MD. Toward understanding tryptophan fluorescence in proteins. Biochemistry (Mosc). 1998;37:9976–9982.
71. Smyth RP, Despons L, Huili G, et al. Mutational interference mapping experiment (MIME) for studying RNA structure and function. Nat Methods. 2015;12:866–872.
72. Sinck L, Richer D, Howard J, et al. In vitro dimerization of human immunodeficiency virus type 1 (HIV-1) spliced RNAs. RNA. 2007;13:2141–2150.
73. Mercenne G, Bernacchi S, Richer D, et al. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. Nucleic Acids Res. 2010;38:633–646.
74. Comas-Garcia M, Datta SA, Baker L, et al. Dissection of specific binding of HIV-1 Gag to the "packaging signal" in viral RNA. eLife. 2017;6:e27055.
75. Bewley MC, Reinhart L, Stake MS, et al. A non-cleavable hexahistidine affinity tag at the carboxyl-terminus of the HIV-1 Pr55 (Gag) polyprotein alters nucleic acid binding properties. Protein Expr Purif. 2017;130:137–145.
76. McKinstry WJ, Hijnen M, Tanwar HS, et al. Expression and purification of soluble recombinant full length HIV-1 Pr55Gag protein in Escherichia coli. Protein Expr Purif. 2014;100:10–18.
77. Zimm BH. The scattering of light and the radial distribution function of high polymer solutions. J Chem Phys. 1948;16:1093–1099.
78. Wyatt PJ. Light scattering and the absolute characterization of macromolecules. Anal Chim Acta. 1993;272:1–40.
79. Lodmell JS, Paillart JC, Mignot D, et al. Oligonucleotide-mediated inhibition of genomic RNA dimerization of HIV-1 strains MAL and LAI: a comparative analysis. Antisense Nucleic Acid Drug Dev. 1998;8:517–529.
80. Paillart J-C, Skripkin E, Ehresmann B, et al. In Vitro evidence for a long range pseudoknot in the 5′-Untranslated and matrix coding regions of HIV-1 Genomic RNA. J Biol Chem. 2002;277:5995–6004.
81. Igarashi K, Kashiwagi K. Polyamines: mysterious modulators of cellular functions. Biochem Biophys Res Commun. 2000;271:559–564.
82. Scatchard G. Attractions of proteins for small molecules and ions. Ann N Y Acad Sci. 1949;51:660–672.
83. Bernacchi S, Henriet S, Dumas P, et al. RNA and DNA binding properties of HIV-1 Vif Protein A fluorescence study. J Biol Chem. 2007;282:26361–26368.