Global Changes in the RNA Binding Specificity of HIV-1 Gag Regulate Virion Genesis

Sebla B. Kutluay,1 Trinity Zang,1,2 Daniel Blanco-Melo,1 Chelsea Powell,1 David Jannain,1 Manel Errando,3 and Paul D. Bieniasz1,2,*
1Laboratory of Retrovirology
2Howard Hughes Medical Institute
3Department of Physics and Astronomy, Barnard College, Columbia University, New York, NY 10027, USA
*Correspondence: pbienias@adarc.org
http://dx.doi.org/10.1016/j.cell.2014.09.057

SUMMARY

The HIV-1 Gag protein orchestrates all steps of virion genesis, including membrane targeting and RNA recruitment into virions. Using crosslinking-immuno-precipitation (CLIP) sequencing, we uncover several dramatic changes in the RNA-binding properties of Gag that occur during virion genesis, coincident with membrane binding, multimerization, and proteolytic maturation. Prior to assembly, and after virion assembly and maturation, the nucleocapsid domain of Gag preferentially binds to psi and Rev Response elements in the viral genome, and GU-rich mRNA sequences. However, during virion genesis, this specificity transiently changes in a manner that facilitates genome packaging; nucleocapsid binds to many sites on the HIV-1 genome and to mRNA sequences with a HIV-1-like, A-rich nucleotide composition. Additionally, we find that the matrix domain of Gag binds almost exclusively to specific tRNAs in the cytosol, and this association regulates Gag binding to cellular membranes.

INTRODUCTION

The HIV-1 Gag protein coordinates all major steps in virion assembly. In so doing, it changes subcellular localization and multimerization state and becomes proteolytically processed (Bell and Lever, 2013; Sundquist and Kräusslich, 2012). One function of Gag is to selectively package a dimeric, unspliced viral RNA genome selected from a pool of excess cellular RNAs and spliced viral mRNAs (Kuzembayeva et al., 2014; Lu et al., 2011b; Rein et al., 2011). Genome packaging requires binding of the nucleocapsid (NC) domain to viral genomic RNA (Aldovini and Young, 1990; Berkowitz et al., 1993; Gorelick et al., 1990). Selection of the viral genome is thought to be governed by a cis-acting packaging element, psi (ψ), within the 5′ leader of the viral genome, composed of sequences in the unique 5′ region (U5) and between the tRNA primer binding site (PBS) and the 5′ portion of the Gag open reading frame (ORF) (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Lever et al., 1989; Luban and Goff, 1994). This element is highly structured (Clever et al., 1995; Harrison and Lever, 1992; McBride and Panganiban, 1996) and may exist in two conformations that favor translation versus dimerization and packaging (Lu et al., 2011a).

Knowledge of the viral RNA sequences that are directly bound by Gag is largely inferred from determinations of functional packaging signals in genetic studies, complemented by limited in vitro data. No assay has yet demonstrated a direct interaction between Gag and ψ in a biologically relevant setting, i.e., in cells or virions. Additionally, some findings suggest that sequences outside ψ might facilitate genome packaging. First, disruption of ψ does not eliminate specific RNA encapsidation (Clever and Parslow, 1997; Laham-Karam and Bacharach, 2007; McBride and Panganiban, 1997). Second, sequences outside ψ can increase HIV-1 vector titers or virion RNA levels (Berkowitz et al., 1995; Chamanian et al., 2013; Das et al., 1997; McBride et al., 1997; Richardson et al., 1993). Third, virions can package cellular RNAs (Muriaux et al., 2001; Rulli et al., 2007), lacking a ψ sequence, particularly in the absence of viral RNA. The RNA properties that underlie these findings are unknown.

HIV-1 Gag molecules exist as monomers or low-order multimers in the cell cytosol and form higher-order multimers only after binding to the plasma membrane (Kutluay and Bieniasz, 2010). Imaging studies indicate that small numbers of Gag molecules recruit a single viral RNA dimer to the plasma membrane, nucleating the assembly of thousands of Gag molecules into an immature virion (Jouvenet et al., 2009). Thereafter, Gag proteolysis liberates NC and other Gag domains, triggering virion maturation. Whether these changes in Gag/NC configuration affect its RNA-binding properties is unknown, but there is clear potential for the RNA binding specificity of Gag to change during virion genesis.

Although NC is thought to be the primary Gag domain that binds to viral RNA, the matrix (MA) domain can also bind RNA in vitro (Alfadhli et al., 2009; Chukkapalli et al., 2010; Cimarelli and Luban, 1999; Levin et al., 2010; Ott et al., 2005; Ramalingam et al., 2011). N-terminal basic amino acids in MA that mediate membrane binding also drive in vitro RNA binding (Chukkapalli et al., 2010, 2013; Hill et al., 1996; Saad et al., 2006; Shkriabai et al., 2006; Zhou et al., 1994). Because RNA is better able to block MA binding to membranes that are devoid of phosphatidylinositol-(4,5)-bisphosphate PI(4,5)P2 (Chukkapalli et al., 2010,
RNA might help to target particle assembly to the plasma membrane. However, RNA binding by MA is not thought to be specific, and whether it actually occurs in cells is unknown.

To obtain a complete account of the RNA sequences bound by Gag during virion genesis, we employed crosslinking-immunoprecipitation-sequencing PAR-CLIP and HITS-CLIP techniques (Hafner et al., 2010; Licatalosi et al., 2008). We find that cytosolic Gag binds to three sequence elements within the 5' leader of the viral genome, which are brought together in a secondary structure that defines a minimal Ψ element. We also find that cytosolic Gag binds to additional discrete sites on the viral genome, including the Rev Response Element (RRE). Gag association with the plasma membrane and its assembly into immature virions trigger a profound change in RNA binding specificity that favors genome packaging. Subsequently, particle maturation largely reverses this change. Finally, we find that MA is a bona fide RNA binding domain that selects a subset of tRNAs in the cytosol, which regulate Gag-membrane binding. Overall, these studies provide a dynamic, quantitative, and high-resolution account of the global changes in Gag-RNA binding during HIV-1 virion genesis.

RESULTS

CLIP Assay for HIV-1 Gag-RNA Binding

We employed recently developed CLIP approaches (Hafner et al., 2010; Licatalosi et al., 2008) to identify RNA molecules bound by HIV-1 Gag protein during particle genesis. To facilitate the purification of Gag-RNA adducts, we generated HIV-1NL4-3 (subtype B) and HIV-1NDK (subtype D) proviral clones carrying an inactivating mutation in the viral protease and three consecutive copies of a HA-tag within the stalk region of MA (MA-3xHA/PR"). The MA modification did not affect Gag expression or assembly and had only a small effect on the infectiousness of a PR" virus (Figures S1A and S1B available online). Cells transfected with HIV-1NL4-3(MA-3xHA/PR") proviral plasmids were grown in the presence of ribonucleoside analogs (4SU or 6SG), which also had minimal effects on infectious virion yield (Figures S1C and S1D).

Cells and virions were UV-irradiated, lysed, and digested with ribonuclease A. Then, Gag-RNA adducts were immunopurified, end-labeled with γ-32P-ATP, and visualized after SDS-PAGE and transfer to nitrocellulose membranes. Gag-RNA adduct formation was dependent on UV irradiation in cells and virions (Figure 1A). We primarily used 4SU-based CLIP thereafter, because it efficiently generated crosslinks, whereas 6SG and unmodified RNA was used for confirmatory purposes. Gag-crosslinked RNA oligonucleotides were purified, sequenced, and mapped to the HIV-1 and human genomes (see Extended Experimental Procedures). Reads derived from the terminal repeat (R) region of the HIV-1 genome ambiguously map to 5' and 3' ends but are displayed at 5' end of the viral genome, and cautiously interpreted, in our analyses.

In six independent Gag-CLIP libraries prepared from HIV-1NL4-3(MA-3xHA/PR") or HIV-1NDK(MA-3xHA/PR") expressing cells, 2.5–7% of the total reads were HIV-1 derived, whereas ~60% were from host cell RNA (Figure 1B; Table S1). In comparison, RNA sequencing (RNA-seq) libraries, which measure the abundance of cellular and viral RNAs, contained 0.3%–1.6% of reads derived from HIV-1, and ~75% were from cellular RNA (Figure 1B; Table S1). In immature virions, ~50% and 20% of CLIP reads were from viral and host RNAs, respectively (Figure 1B; Table S1), broadly similar to RNA abundance as determined by RNA-seq (Figure 1B). Thus, viral RNA sequences were somewhat selectively bound by Gag in cells but were enriched to a far greater extent in virions.

HIV-1 RNA Sequences Bound by Gag in Cells and Virions

We plotted the frequencies with which each nucleotide in the HIV-1NL4-3 genome was represented in reads from cell- and virion-derived HIV-1NL4-3(MA-3xHA/PR") Gag-CLIP libraries ("read density", Figure 1C). In cells, a major proportion of the Gag-linked reads were derived from discrete sites in the viral genome. As might be expected, the 5' leader was a frequent site of Gag binding. However, additional sites of frequent Gag binding included the RRE, sequences overlapping the Nef start codon, and untranslated sequences in U3 (Figure 1C). Several other sites on the viral genome were bound by Gag, but at lower frequencies. The distribution of Gag-crosslinked viral RNA reads was highly reproducible in HIV-1NL4-3(MA-3xHA/PR") Gag-CLIP-seq libraries as indicated by the nearly perfect correlation between independent experiments (Figure 1D). However, when a divergent HIV-1NDK strain was included in these analyses, the only prominent Gag binding sites that were present in both strains were the 5' leader and RRE (Figure 1C).

The frequency with which Gag was crosslinked to sites across the viral genome was starkly different in immature virions. Sites of frequent Gag binding on HIV-1NL4-3 RNA were far more numerous in immature virions than in cells. Moreover, specific Gag binding to the 5' leader and RRE was not evident in immature virions (Figure 1C). The pattern of Gag binding frequency across the viral genome was highly reproducible (Figure 1D) and was unlikely to be generated by methodological bias during CLIP-seq library generation, because it was largely unaffected by the choice of ribonucleoside analog (4SU versus 6SG; Figures S1E and S1F), ribonuclease (RNase A versus RNase T1; Figures S1G and S1H), or the immunoprecipitating antibody (anti-HA versus anti-NC; Figures S1I and S1J) used to generate the CLIP libraries. HIV-1NDK Gag similarly bound to sites throughout the viral genome (Figure 1C). Despite overall similarity, there were some clear discrepancies in viral RNA sites that were frequently occupied by HIV-1NL4-3 and HIV-1NDK Gag (Figures 1E and S1K), presumably due to differences in target RNA sequence, or subtle differences in the RNA binding specificities of the two Gag proteins.

The 3xHA tag did not affect the pattern of Gag binding at sites proximal or distal to the insertion site (Figure S1L). Moreover, autocorrelation analysis of Gag binding frequency revealed no peaks other than at a separation of s = 0 (Figure S1M). Thus, the frequency with which a given site in the viral genome was bound by Gag in immature virions appeared unaffected by its position relative to the Ψ sequence or other Gag binding sites. Rather, Gag binding appeared to be a function of local nucleotide sequence or structure.
Analysis of Gag Binding to 5’ Leader and RRE

In cells, the most prominent Gag binding sites on the viral genome coincided with the most prominently structured elements in the HIV-1 genome, namely, the 5’ leader and the RRE (Figure 1C). However, Gag was not bound across the 5’ leader and the RRE with uniform frequency, but selectively associated with small determinants within these structures. In the case of the 5’ leader, Gag was most frequently bound to three distinct sequences, including one at the 5’ end of U5, a second site between the PBS and the major splice donor, and a third site 3’ to the major splice donor (Figure 2A). These sites are separated from each other in linear sequence by approximately 100 nucleotides. Strikingly, however, a nuclear-magnetic-resonance-based analysis of the structure of a Ψ sequence (Lu et al., 2011a) predicts that these three Gag-binding sequences would be in close proximity, and partly base paired with each other upon RNA folding into a structure that favors genome packaging (Figure 2B).

Frequent Gag binding to the RRE occurred in stem I, a site clearly distinct from the primary site of Rev-RRE interaction in stem-loop IIB (Malim et al., 1990) (Figures 2C and 2D). Importantly, the Gag-bound reads derived from both the 5’ leader and RRE contained high rates of T-to-C substitution, identifying individual nucleotide bases that were in close proximity to Gag.
Additionally, CLIP-seq experiments performed using 4SU or 6SG yielded a similar footprint of Gag on viral RNA, with 5' leader and RRE binding prominently featured (Figures S2A and S2B). Thus, the apparently specific binding of Gag to these sequence elements was not an artifact of the particular crosslinking nucleotide used.

Changes in Gag RNA Binding to Viral RNA during HIV-1 Virion Assembly

Given the stark differences in the interaction of Gag with viral RNA in cells and immature virions, we attempted to determine what triggers this dramatic change. To this end, cells were fractionated after UV irradiation (Figure 3A), and Gag-RNA adducts...
were immunoprecipitated from cytosol, in which Gag is primarily monomeric, and from a membrane fraction where multimerized assembly intermediates form (Kutluay and Bieniasz, 2010) (Figure 3B). The RNA binding profile of cytosolic Gag was nearly indistinguishable from that described above for total cell lysates, with Ψ and RRE representing specific binding sites (Figures 3C and 3D). Similarly, a myristoylation-defective Gag mutant (Gag-G2A), which does not bind efficiently to membranes, was also restricted to discrete binding sites on viral RNA (Figure S3A). Conversely, Gag that was immunoprecipitated from membrane fractions was bound to many sites on the viral genome, and, although binding to Ψ and RRE remained prominent, the overall

Figure 3. Changes in Gag Binding to Viral RNA during Virion Assembly and Maturation
(A) Western blot analysis of Gag, and markers of cytosol (LC3A/B) and plasma membrane (Na-K ATPase) in fractionated, 4SU-fed 293T cells transfected with a HIV-1NL4-3(MA-3xHA/PR/NC) proviral plasmid.
(B) Autoradiogram of Gag-RNA complexes recovered from fractions in CLIP assays.
(C) Read density distribution on viral RNA from CLIP experiments in which WT and mutant Gag proteins were immunoprecipitated from cell fractions.
(D) Correlation analysis of CLIP data from cell fractions and immature virions.
(E) Autoradiogram of NC-RNA and Gag-RNA complexes recovered from mature and immature virions using an anti-NC antibody.
(F) Read density distribution on viral RNA from CLIP experiments in which Gag and NC proteins were immunoprecipitated from immature and mature virions using an anti-NC antibody.
(G) Correlation analysis of CLIP data from cells and mature and immature virions.
See also Figure S3.
RNA binding profile of cell-membrane-associated Gag otherwise resembled that of immature virion-associated Gag (Figures 3C, 3D, and S3B). A late-domain mutant Gag protein (Δp6) was similarly bound to sites throughout the viral genome (Figure 3C) indicating that the completion of budding was not required for this apparent shift in RNA binding specificity. Notably, a Gag CAΔCTD mutant that is fully competent to localize to the plasma membrane but is significantly impaired in the formation of high-order multimers (Kutluay and Bieniasz, 2010) exhibited a binding profile that was more reminiscent of cytosolic Gag, although additional sites in the viral genome were bound with some prominence (Figure 3C). Together, these data indicate that high-order oligomerization of Gag at the plasma membrane drives a profound change in Gag’s RNA-binding properties (Figure 3C, 3D, and S3B), dramatically increasing the extent to which Gag and viral RNA interact with each other.

Changes in Gag-Viral RNA Binding Triggered by Virion Maturation

Gag undergoes proteolytic cleavage during virion maturation, liberating NC, which is then thought to condense with the viral RNA inside a remodeled conical core (Sundquist and Kräusslich, 2012). To determine whether virion maturation affects Gag-viral RNA interactions, we did CLIP experiments using an antibody that specifically recognizes NC (Figures S3D and S3E) and could efficiently immunoprecipitate Gag-RNA and NC-RNA adducts from immature and mature virions, respectively (Figure 3E). Comparison of RNA sequences bound by Gag in immature virions and NC in mature virions revealed profound changes accompanying virion maturation (Figures 3F and 3G). Unlike intact Gag in immature virions, NC in mature virions preferentially occupied discrete sequences on the viral genome, that coincided in large part with the major sites of Gag-viral RNA interaction in the cytosol (i.e., Ψ and RRE) (Figures 3F and 3G). Many of the additional prominent NC-binding sites in mature virion RNA were also bound by intact Gag in the cytosol at lower frequencies (Figure 3F). Thus, the RNA-binding properties of NC in mature virions resemble that of unassembled Gag in the cytosol to a surprising extent, exhibiting statistically significant correlation (Figure 3G), possibly reflecting the monomeric state of cytosolic Gag and mature NC.

Gag Binding to Viral RNA and Viral RNA Packaging in the Absence of Ψ or RRE

To test the importance of the Gag/NC-binding sites in Ψ and RRE, we generated viruses carrying deletions of these sequences. First, we determined whether the Gag binding to additional sites on the viral genome in cells and in virions required initial binding to Ψ. To this end, we performed CLIP experiments using viral constructs carrying deletions of two of the three Gag-binding sequences (Δ105-278), or all three Gag binding sequences (Δ105-278/Δ301-332) within Ψ. These deletions left sequences surrounding the major 5′ splice donor intact, and only modestly reduced Gag expression (Figure 4A). Surprisingly, in cells, neither the Δ105-278 nor the Δ105-278/Δ301-332
mutation affected the fraction of Gag-crosslinked reads derived from viral versus human genomes (Figure 4B). Moreover, in immature virions, these ψ-deletions caused only a modest ~3-fold decrease in the fraction of reads derived from the viral genome, with a corresponding increase in the fraction of reads from cellular RNAs (Figure 4B). The pattern of Gag binding to viral RNA in cells was not greatly affected by the ψ deletions (Figures 4C and 4D), although there was an apparent reduction in the frequency with which RRE was bound. Similarly, the overall pattern of Gag binding to viral RNA in immature virions was also relatively unperturbed by ψ deletions (Figures 4C, 4D, and S4A). However, there was a tendency for read densities to increase toward the 3′ end of the viral genome, perhaps reflecting the incorporation of spliced viral RNAs into virions generated by ψ-deleted genomes (Clever and Parslow, 1997; Houzet et al., 2007; Russell et al., 2003). Importantly, these results indicate that multiple sites on the viral genome bind to Gag independently of ψ and facilitate HIV-1 genome packaging.

To analyze the effects of Gag-RRE interaction on HIV-1 infectivity, we generated viruses in which Rev was deleted and the nuclear export of unspliced viral RNAs was mediated by the Mason-Pfizer monkey virus constitutive transport element (CTE) to circumvent effects of RRE on RNA nuclear export (Figure S4B). In this setting, the presence or absence of RRE had no discernible effect on Gag protein levels or infectious virus yield (Figure S4C). Although it is formally possible that the MPMV CTE might recapitulate the Gag-binding properties of RRE, this analysis suggests that Gag RRE interactions do not regulate Gag expression and are not required for RNA packaging.

Changes in the Sequence Specificity of Gag RNA Binding during Virion Genesis

The aforementioned results suggest a degree of redundancy, and contributions from multiple domains of the viral genome to RNA packaging. In an attempt to explain the factors driving the selectivity with which Gag bound to and packaged viral RNA, we determined the identities of cellular mRNA sequences that were most frequently bound to Gag in cells and in virions. Reads that aligned to the human genome were clustered using PARalyzer (Corcoran et al., 2011), which defines a cluster, or binding site, based on the occurrence of a minimum number overlapping reads proximal to a T-to-C substituted crosslinking site in 4SU-based CLIP assays (Table S2). Note that a single cluster is counted once for the analysis shown in Figure 5A, irrespective of the number of reads associated with it. In cells, ~90% of clusters (Gag binding sites) were within genes, and ~80% of these were derived from mRNAs (Figure 5A) with an overlap of >95% between HIV-1NL4-3 and HIV-1NDK Gag-bound mRNA clusters (Figure S5A). A similar analysis of immature virions generated a collection of clusters that were also mostly derived from mRNAs (Figure 5A) and overlapped by >70% in HIV-1NL4-3 and HIV-1NDK libraries. Notably, greater discrepancies were observed when the Gag-binding clusters in cells versus virions were compared (Figure S5A), indicating that many mRNA sequences that are preferred Gag binding sites in cells are not preferred binding sites in virions and vice versa.

We counted the number of reads associated with each cluster and determined the 100 sites in cellular mRNAs that were most frequently bound by Gag in cells and immature virions. The nucleotide composition of these Gag binding sites revealed a striking change in the RNA binding preference of Gag during virion genesis (Figure 5B). In cells, preferred Gag binding sites had a strong tendency to be G-rich (mean G-content of ~40%). In contrast, in immature virions, preferred Gag binding sites displayed a tendency to be A-rich. Remarkably, the nucleotide content of the preferred Gag binding sites in cellular mRNAs associated with immature virions was strikingly similar to the nucleotide composition of the HIV-1 genome. We used cERMIT (Georgiev et al., 2010), to search for sequence motifs that were enriched in Gag-bound cellular mRNA clusters recovered at various stages of virion genesis. G/U-rich sequence motifs were the most often present in host mRNA sequences bound to Gag in the cytoplasm (Figure 5C). At the plasma membrane, and particularly in immature virions, there was a clear change in binding specificity. Here, Gag bound more frequently to A-rich and A/G-rich cellular mRNA motifs (Figure 5C). Inspection of the sequences of the 12 clusters most frequently bound by Gag in the viral genome, in immature virions, revealed multiple instances of similar A/G-rich motifs scattered in the viral genome (Figure 5D). In mature virions, RNA binding specificity reverted, and the sequence motifs favored by NC were again G/U-rich. These findings reinforce the notion that Gag binds preferentially to RNAs that have an A-rich nucleotide composition, and particularly to A-rich motifs that are present in the viral genome, transiently during immature virion assembly.

tRNAs Are the Most Frequently Gag-Bound RNA in Cells, but Not in Virions

Although the aforementioned analysis focused on Gag binding to mRNA sequences, when the number of individual Gag-bound reads associated with each cluster or binding site was counted, it was evident that mRNAs were responsible for only a fraction (~12%) of the cellular RNA reads crosslinked to Gag. In fact, tRNAs were the dominant RNA species bound by Gag in cells (~60%–70% of reads), with 7SL RNA, a known component of HIV-1 virions (Onafuwa-Nuga et al., 2006) constituting the bulk of the remaining reads (Figure 6A). Strikingly, the fraction tRNA-derived reads was greatly reduced in CLIP-seq libraries prepared from virions as compared to cells. This was not the case for other RNA classes, suggesting that tRNAs were specifically excluded from Gag-bound RNAs in virions (Figure 6A).

MA Binds to Particular tRNAs

Although the aforementioned Gag-RNA binding events were assumed to involve the NC domain of Gag, there are numerous reports that MA can interact with viral or cellular RNAs (reviewed in Alfadhli and Barklis, 2014). However, these interactions, generally assayed in vitro, are of uncertain significance. To test whether MA binds to RNA in cells, we modified the aforementioned HIV-1NL4-3(3xHA/PR-3HA) clones to include a FXa protease cleavage site immediately N- or C-terminal to the 3xHA-tag (Figures 6B and S6A). After UV crosslinking, cell lysates could be treated with FXa protease to generate two Gag fragments (MA and CA-NC-p6) only one of which would be tagged and immuno-precipitated (Figures 6B and 6C).
Like full-length Gag, the immunoprecipitated CA-NC-p6 contained crosslinked RNA species that caused adducts to migrate ~0–10 kDa above the expected molecular weight of the protein. Notably, the isolated MA domain also generated prominent crosslinked species, indicating RNA binding. Unusually, these MA-RNA adducts migrated at a discrete molecular weight of ~35 kDa, suggesting that MA bound to a unique RNA species (Figure 6D).

Although the CA-NC-p6-RNA adducts contained viral RNA sequences with a very similar distribution to those recovered from full-length Gag, MA-RNA adducts were nearly devoid of viral RNA-derived sequence (Figure 6E). In fact, the RNA molecules crosslinked to MA were almost exclusively host tRNAs (Figure 6F). When RNase treatment was omitted from the CLIP procedure prior to immunoprecipitation, MA-RNA adducts migrated at a larger, but still discrete, molecular weight (Figure 6D) matching the size of one tRNA molecule (~20 kDa) crosslinked to MA, suggesting that MA binds to full-length tRNAs.

Although tRNAs constituted the majority of cellular RNA molecules bound by intact Gag in cells, deletion of MA caused a major
reduction in the fraction of reads that were tRNA derived but did not inhibit binding to mRNAs or 7SL RNA (Figures 6F and S6B).

Conversely, a mutant Gag protein lacking the NC domain did not bind mRNA or 7SL, but bound nearly exclusively to tRNAs (Figures 6F and S6B). Progressive and specific loss of tRNA binding occurred as more lysines in the N-terminal MA basic domain were substituted, and mutation of six lysines (MAK6T) reduced Gag-tRNA binding to the same degree as deletion of MA (Figures 6F and S6B). Progressive and specific loss of tRNA binding occurred as more lysines in the N-terminal MA basic domain were substituted, and mutation of six lysines (MAK6T) reduced Gag-tRNA binding to the same degree as deletion of MA (Figures 6F and S6B).

See also Figure S6.
6F and S6B). As expected, MAK4T and MAK6T mutations caused decreases in the levels of released particles (Figure S6B), consistent with a plasma membrane targeting defect in these mutants.

MA binding to tRNAs was highly selective. Indeed, GluCTC, GluTTC, GlyGCC, GlyCCC, LysCTT, LysTTT, ValAAC, and ValCAC tRNAs were bound up to 100-fold more frequently than the majority of tRNAs (Figure 6G). No such enrichment occurred in CLIP experiments done with GagD MA. Although intact tRNA molecules were bound by MA (Figure 6D), binding apparently involved the 5′ half of the tRNA molecules, and especially the dihydrouridine loop, as indicated by the very high rates of T-to-C conversions at this site (Figure 6H). Binding was not induced by 4SU incorporation into tRNAs, because the same tRNAs were selectively MA-bound in CLIP experiments where crosslinking was induced by UV irradiation at 254 nm in the absence of modified nucleotides (Figure S6C).

Regulation of Gag Membrane Binding by tRNA
The finding that the same lysine residues that mediate plasma membrane binding also mediate tRNA binding, suggested that tRNAs might regulate Gag localization. Indeed, previous work has indicated that exogenous RNA can inhibit Gag binding to liposomes in vitro (Chukkapalli et al., 2010, 2013; Dick et al., 2013). We took a simple cell-based approach (without addition of exogenous RNA or liposomes) and tested whether RNase treatment of cell lysates increased Gag binding to endogenous cellular membranes. Lysates of HeLa and HOS cells stably expressing Gag-CFP were treated with ribonucleases and subjected to membrane flotation analysis. Strikingly, RNase treatment caused a significant redistribution of Gag from cytosol to membrane fractions (Figures 7A and 7B), whereas a control protein, β-tubulin, was largely unaffected. Consistent with a model in which tRNAs compete with membranes for Gag binding, tRNAs comprised a significantly smaller fraction of Gag-bound RNAs at the plasma membrane than in the cytoplasm (Figure 7C). This difference was more pronounced when particle budding was blocked at the plasma membrane by deletion of the late budding domain of Gag (Figure 7C). Thus, these results indicate that MA can bind to tRNA or cell membranes, but not both simultaneously, and strongly suggest that tRNAs regulate Gag localization by binding to basic amino acids in MA.

DISCUSSION
Two central conclusions of these studies are (1) the HIV-1 Gag protein has two RNA binding domains (NC and MA) with very different specificities and (2) dramatic changes in RNA binding regulate Gag localization and genome packaging during virion genesis (Figure 7D).

Prior to virion assembly, Gag exists as a diffuse pool of monomers or low-order multimers in the cytoplasm with its NC domain
bound primarily to mRNA, with some binding to 7SL and tRNA (Figure 7D). A fraction of NC-RNA interactions are with particular sites on the HIV-1 genome, including theΨsequence. Our data reveal the specific RNA sequences withinΨthat are in proximity to Gag, in a physiological setting. Satisfyingly, the three noncontiguous RNA elements that are most frequently crosslinked to Gag (nucleotides 100–126, 195–260, and 300–350) coincide nearly precisely with a minimalΨelement, which adopts a secondary structure that putatively favors genome packaging (Lu et al., 2011a).

Surprisingly, cytoplasmic Gag bound to additional discrete elements on the viral RNA, including RRE. Previously, env sequences, including RRE, were shown to facilitate packaging (Kaye et al., 1995; Richardson et al., 1993), but a discrete packaging sequence within env could not be mapped. Other reports suggest that Rev enhances packaging (Brandt et al., 2007), although effects of Rev/RRE on viral RNA export are potential confounders in packaging experiments. Despite strong evidence for Gag-RRE binding, the RRE did not affect infectious virion yield when its nuclear export function was replaced, suggesting that Gag-RRE binding is not required for genome packaging. Although it is possible that Gag/NC-RRE interaction plays a redundant role in packaging, other plausible functions include (1) shielding double-stranded stem I RNA from cytoplasmic sensors, (2) coupling RNA-export with packaging, (3) displacement of Rev from the RRE for recycling, (4) regulation of Env translation, and (5) enhancement of reverse transcription via NC’s unwinding/chaperone function (Levin et al., 2010). Further work will be required to elucidate the functional role, if any, of Gag/NC-RRE binding.

In the cytosol, Gag favored binding to the viral genome over cellular mRNAs by a few fold. This level of discrimination is insufficient to account for the selectivity with which viral genomes are packaged. Our data indicate that a dramatic change in Gag-RNA binding specificity, coincident with CA-CTD-dependent high-order multimerization at the plasma membrane, contributes to selective packaging (Figure 7D). GU-rich sequences in cellular mRNA were targeted by Gag in the cytosol, consistent with previous in vitro and structural studies indicating that the isolated, monomeric NC domain favors binding to such sequences (Berglund et al., 1997; De Guzman et al., 1998; Fisher et al., 1998). However, during assembly, Gag molecules become tightly packed in hexameric lattices (Briggs et al., 2009; Wright et al., 2007). By constraining thousands of NC domains into a pseudo-2D curved array, local NC concentration is dramatically elevated. Potentially, features of NC that govern RNA binding specificity (zinc knuckles and basic amino acids) might be differentially accessible in an assembled Gag lattice. Under these conditions, we found that A-rich mRNA sequences were preferentially bound by Gag (Figure 7D). Remarkably, the nucleotide composition of mRNA sequences bound by assembled Gag reflects an unusual, heretofore unexplained, property of the HIV-1 genome. Thus, our findings suggest that a need to selectively package viral RNA caused HIV-1 to evolve an unusually A-rich genome. Conversely, an A-rich genome may have evolved for other reasons and then drove Gag to gain a unique oligomerization-driven specificity for A-rich RNA. Notably,Ψdeletion caused only a 3-fold reduction in the fraction of virion-associated, Gag-bound, RNA sequences that were viral RNA derived. Moreover, the pattern of Gag binding to many sites in the viral genome was not solely a secondary effect of physical proximity to Gag-Ψinteractions. Rather, it appears that HIV-1 genome packaging is a two-step process, involving interactions between (1)Ψand monomeric Gag and (2) A-rich viral RNA and multimeric Gag. This scenario should selectively drive particle assembly on viral RNAs, and we speculate that the biases in nucleotide composition exhibited by HIV-1 and other retroviruses serves as a proofreading-like mechanism to enhance the fidelity of genome packaging following initial Gag-Ψinteraction.

An unexpected finding was that proteolytic cleavage of Gag caused NC to revert to a preference for GU-rich mRNAs and discrete viral RNA sequences. This result reinforces the notion that Gag/NC RNA binding specificity is multimerization dependent. By liberating the majority of viral RNA from NC, while maintaining interaction with structured elements (to enable NC’s chaperone activity), maturation-dependent changes in Gag/NC-RNA binding could facilitate reverse transcription (Levin et al., 2010).

Another surprising finding was that MA binding to specific tRNAs constitutes the most frequent binding events between Gag and RNA in cells. MA-tRNA binding was independent of NC and the PBS and is thus unlikely to involve the tRNA primer annealed to viral genome. Rather, we found that MA-tRNA interaction could regulate the binding of Gag to cellular membranes. MA specifies the location of virion assembly, and it was previously shown that RNA can block in vitro MA binding specifically to liposomes that lack acidic phospholipids (Chukkapalli et al., 2010, 2013). Thus, occlusion of MA basic residues by specific tRNAs might inhibit nonproductive assembly at most intracellular membranes and facilitate targeting to the plasma membrane where resident lipids have a high affinity for MA. Alternatively, MA-tRNA binding might provide a mechanism by which virion assembly is temporally regulated.

MA-tRNA interactions could serve additional purposes. Some degree of RNA binding may be an inevitable consequence of encoding a highly basic domain. Thus, specific MA binding in a 1:1 complex to small RNAs might be a mechanism to avoid the aggregation of a protein that has both two distinct RNA binding domains and a tendency to multimerize. MA-tRNA binding might also prevent nonproductive interaction of a viral genome with a Gag monomer whose NC domain has engaged viral RNA. Conceivably, MA may facilitate the selection of the RT primer, as tRNAlys3 is among the tRNAs bound by MA, but several other tRNAs constitute the most frequent binding events between MA and viral genome. Rather, we found that MA-tRNA interaction could regulate the binding of Gag to cellular membranes. MA specifies the location of virion assembly, and it was previously shown that RNA can block in vitro MA binding specifically to liposomes that lack acidic phospholipids (Chukkapalli et al., 2010, 2013). Thus, occlusion of MA basic residues by specific tRNAs might inhibit nonproductive assembly at most intracellular membranes and facilitate targeting to the plasma membrane where resident lipids have a high affinity for MA. Alternatively, MA-tRNA binding might provide a mechanism by which virion assembly is temporally regulated.

MA-tRNA interactions could serve additional purposes. Some degree of RNA binding may be an inevitable consequence of encoding a highly basic domain. Thus, specific MA binding in a 1:1 complex to small RNAs might be a mechanism to avoid the aggregation of a protein that has both two distinct RNA binding domains and a tendency to multimerize. MA-tRNA binding might also prevent nonproductive interaction of a viral genome with a Gag monomer whose NC domain has engaged viral RNA. Conceivably, MA may facilitate the selection of the RT primer, as tRNAlys3 is among the tRNAs bound by MA, but several other tRNAs are also bound by MA more frequently. Finally, MA-tRNA interaction could regulate viral and/or host translation. The A-richness of the HIV-1 genome results in suboptimal codon usage (Grantham and Perrin, 1986; Kypr and Mrázek, 1987; Sharp, 1986) and an elevated number of Ile, Lys, Glu, and Val codons in the Gag and Pol ORFs (Berkhout and van Hemert, 1994). Notably, Lys, Glu, and Val tRNAs are among those specifically bound by MA, providing a potential opportunity for Gag to regulate its own translation as it accumulates to high levels and sequesters tRNAs (perhaps facilitating packaging as a consequence). Similarly, MA could inhibit translation of host mRNAs whose products may be deleterious for viral replication. Indeed, interaction of MA with host elongation factors via a tRNA bridge...
has been reported to inhibit translation in vitro (Cimarelli and Luban, 1999).

Overall, our global survey reveals surprising ways in which the interaction between Gag, viral, and host RNAs can change and modulate the process of virion genesis and genome packaging.

EXPERIMENTAL PROCEDURES

Proviral Plasmids and Cells

HIV-1NL4-3-derived proviral plasmids containing a 3xHA tag in the stalk region of matrix (HIV-1NL4-3MA-3xHA) were constructed using overlap extension PCR. Various derivatives of this construct encoding a catalytically inactive viral protease (PR1), a Factor Xa cleavage site on either side of the HA tag (MA-FXa-3xHA and MA-3xHA-FXa), deletions of nucleocapsid domain (ΔNC), the globular head of MA (ΔMA), the CA CTD (CA ΔCTD), or the RRE signal (Δ105-278 and Δ105-278/Δ301-332) were constructed using PCR-based deletion mutagenesis. Constructs carrying mutations at binding sites for Tsg101 and Alix proteins in the p6 domain of Gag (Δp6), the Gag myristoylation signal (G2A), or at lysine residues in MA (MAK26,27T MAK4T and MAK6T) were constructed using PCR overlap extension-based mutagenesis. Proviral plasmids with deletions or mutations in Env, Rev, and RRE and encoding a Mason-Pfizer monkey virus constitutive transport element were similarly constructed. Details of the construction are described in Extended Experimental Procedures.

PAR-CLIP, HITS-CLIP, and RNA-Seq

For CLIP experiments, HEK293T cells were grown in 10 cm dishes and transfected with proviral plasmids using polyethyleneimine (PolySciences). Virions were harvested from filtered supernatant by ultracentrifugation through sucrose and UV irradiated, whereas cells were irradiated while adhered to culture dishes. Prior to UV crosslinking, a fraction of cells and virions were collected for RNA-seq analysis. After crosslinking, the CLIP procedure was performed on unfraccionated lysates (after removal of nucleic) or on membrane and cytoplasmic fractions. For the CLIP procedure, cell and virion lysates were treated with RNaseA or RNaseT1 and then incubated with Protein G-conjugated Dynabeads coated with anti-HA or anti-NC antibodies. After immunoprecipitation of RNA-protein adducts, beads were washed and treated sequentially with calf intestinal alkaline phosphatase and then polynucleotide kinase and γ-32P-ATP. RNA-protein adducts were eluted from the beads, separated by SDS-PAGE, blotted onto nitrocellulose, and digested with protease K. RNA oligonucleotides were then ligated to adapters, amplified by PCR and sequences determined using an Illumina HiSeq 2000 platform.

CLIP-seq experiments were performed two to six times on cells and virions. Further details of the method are in Extended Experimental Procedures.

Bioinformatic Analyses

The FASTX toolkit was used to process the reads from CLIP and RNA-seq libraries before mapping. Reads were separated based on their 5’ barcode sequences and collapsed to generate a set of unique sequences. Unique CLIP-seq and RNA-seq reads were mapped to the human (hg19) and HIV-1 genomes using the Bowtie. SAMtools, in house scripts, and GraphPad Prism were used to calculate and display read densities associated with viral and genomes using the Bowtie, SAMtools, in house scripts, and GraphPad Prism.

Membrane Flotation Assays

The membrane flotation assays were performed using HeLa and HOS cells stably expressing cyan fluorescent protein-tagged Gag proteins as described (Kutluay and Bieniasz, 2010) with modifications outlined in the Extended Experimental Procedures.

ACCESSION NUMBERS

CLIP-seq data were deposited to the GEO database with accession number GSE61508.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.09.057.

AUTHOR CONTRIBUTIONS

S.B.K. and P.D.B. conceived the study, designed the experiments, analyzed the data, and wrote the paper. S.B.K., T.Z., C.P., and D.J. executed the experiments. S.B.K., D.B.M., and M.E. analyzed the sequencing data. P.D.B. supervised the work.

ACKNOWLEDGMENTS

We thank Markus Hafner and members of the Darnell and Tuschl labs for advice and helpful discussions. We thank Michael Summers for the diagram of the RRE and RRE sequences in Figure 2. S.B.K. was supported in part by an AmFAR Mathilde Krim Postdoctoral Fellowship. This work was supported by NIH grants R01AI501111 and P50GM103297.

Received: August 3, 2014

Revised: September 17, 2014

Accepted: September 24, 2014

Published: November 6, 2014

REFERENCES


donor site are required for both packaging and dimerization of human immunodeficiency virus type 1 RNA. J. Virol. 77, 84–96.


Global Changes in the RNA Binding Specificity of HIV-1 Gag Regulate Virion Genesis

Sebä B. Kutluay,1 Trinity Zang,1,2 Daniel Blanco-Melo,1 Chelsea Powell,1 David Jannain,1 Manel Errando,3 and Paul D. Bieniasz1,2,*
1Laboratory of Retrovirology
2Howard Hughes Medical Institute
3Department of Physics and Astronomy, Barnard College, Columbia University, New York, NY 10027, USA
*Correspondence: pbienias@adorc.org
http://dx.doi.org/10.1016/j.cell.2014.09.057

SUMMARY

The HIV-1 Gag protein orchestrates all steps of virion genesis, including membrane targeting and RNA recruitment into virions. Using crosslinking-immunoprecipitation (CLIP) sequencing, we uncover several dramatic changes in the RNA-binding properties of Gag that occur during virion genesis, coincident with membrane binding, multimerization, and proteolytic maturation. Prior to assembly, and after virion assembly and maturation, the nucleocapsid domain of Gag preferentially binds to psi and Rev Response elements in the viral genome, and GURich mRNA sequences. However, during virion genesis, this specificity transiently changes in a manner that facilitates genome packaging; nucleocapsid binds to many sites on the HIV-1 genome and to mRNA sequences with a HIV-1-like, A-rich nucleotide composition. Additionally, we find that the matrix domain of Gag binds almost exclusively to specific tRNAs in the cytosol, and this association regulates Gag binding to cellular membranes.

INTRODUCTION

The HIV-1 Gag protein coordinates all major steps in virion assembly. In so doing, it changes subcellular localization and multimerization state and becomes proteolytically processed (Bell and Lever, 2013; Sundquist and Kräusslich, 2012). One function of Gag is to selectively package a dimeric, unspliced viral RNA genome selected from a pool of excess cellular RNAs and spliced viral mRNAs (Kuzembayeva et al., 2014; Lu et al., 2011b; Rein et al., 2011). Genome packaging requires binding of the nucleocapsid (NC) domain to viral genomic RNA (Aldovini and Young, 1990; Berkowitz et al., 1993; Gobrelick et al., 1990). Selection of the viral genome is thought to be governed by a cis-acting packaging element, psi (Ψ), within the 5′ leader of the viral genome, composed of sequences in the unique 5′ region (U5) and between the tRNA primer binding site (PBS) and the 5′ portion of the Gag open reading frame (ORF) (Aldovini and Young, 1990; Clavel and Orenstein, 1990; Lever et al., 1989; Luban and Goff, 1994). This element is highly structured (Clever et al., 1995; Harrison and Lever, 1992; McBride and Panganiban, 1996) and may exist in two conformations that favor translation versus dimerization and packaging (Lu et al., 2011a).

Knowledge of the viral RNA sequences that are directly bound by Gag is largely inferred from determinations of functional packaging signals in genetic studies, complemented by limited in vitro data. No assay has yet demonstrated a direct interaction between Gag and Ψ in a biologically relevant setting, i.e., in cells or virions. Additionally, some findings suggest that sequences outside Ψ might facilitate genome packaging. First, disruption of Ψ does not eliminate specific RNA encapsidation (Clever and Parslow, 1997; Laham-Karam and Bacharach, 2007; McBride and Panganiban, 1997). Second, sequences outside Ψ can increase HIV-1 vector titers or virion RNA levels (Berkowitz et al., 1995; Chamanian et al., 2013; Das et al., 1997; McBride et al., 1997; Richardson et al., 1993). Third, virions can package cellular RNAs (Muraiya et al., 2001; Rulli et al., 2007), lacking a Ψ sequence, particularly in the absence of viral RNA. The RNA properties that underlie these findings are unknown.

HIV-1 Gag molecules exist as monomers or low-order multimers in the cell cytosol and form higher-order multimers only after binding to the plasma membrane (Kutluay and Bieniasz, 2010). Imaging studies indicate that small numbers of Gag molecules recruit a single viral RNA dimer to the plasma membrane, nucleating the assembly of thousands of Gag molecules into an immature virion (Jouvenet et al., 2009). Thereafter, Gag proteolysis liberates NC and other Gag domains, triggering virion maturation. Whether these changes in Gag/NC configuration affect its RNA-binding properties is unknown, but there is clear potential for the RNA binding specificity of Gag to change during virion genesis.

Although NC is thought to be the primary Gag domain that binds to viral RNA, the matrix (MA) domain can also bind RNA in vitro (Alfadhli et al., 2009; Chukkapalli et al., 2010; Cimarelli and Luban, 1999; Levin et al., 2010; Ott et al., 2005; Ramalingam et al., 2011). N-terminal basic amino acids in MA that mediate membrane binding also drive in vitro RNA binding (Chukkapalli et al., 2010, 2013; Hill et al., 1996; Saad et al., 2006; Shkriabai et al., 2006; Zhou et al., 1994). Because RNA is better able to block MA binding to membranes that are devoid of phosphatidylinositol-(4,5)-bisphosphate PI(4,5)P2 (Chukkapalli et al., 2010,
RNA might help to target particle assembly to the plasma membrane. However, RNA binding by MA is not thought to be specific, and whether it actually occurs in cells is unknown. To obtain a complete account of the RNA sequences bound by Gag during virion genesis, we employed crosslinking-immuno-precipitation-sequencing PAR-CLIP and HITS-CLIP techniques (Hafner et al., 2010; Licatalosi et al., 2008). We find that cytosolic Gag binds to three sequence elements within the 5' leader of the viral genome, which are brought together in a secondary structure that defines a minimal Ψ element. We also find that cytosolic Gag binds to additional discrete sites on the viral genome, including the Rev Response Element (RRE). Gag association with the plasma membrane and its assembly into immature virions trigger a profound change in RNA binding specificity that favors genome packaging. Subsequently, particle maturation largely reverses this change. Finally, we find that MA is a bona fide RNA binding domain that selects a subset of tRNAs in the cytosol, which regulate Gag-membrane binding. Overall, these studies provide a dynamic, quantitative, and high-resolution account of the global changes in Gag-RNA binding during HIV-1 virion genesis.

RESULTS

**CLIP Assay for HIV-1 Gag-RNA Binding**

We employed recently developed CLIP approaches (Hafner et al., 2010; Licatalosi et al., 2008) to identify RNA molecules bound by HIV-1 Gag protein during particle genesis. To facilitate the purification of Gag-RNA adducts, we generated HIV-1NL4-3 (subtype B) and HIV-1NDK (subtype D) proviral clones carrying MA-3xHA/PR. The MA modification did not affect Gag expression or assembly and had only a small effect on the infectiousness of a PR virus (Figures S1A and S1B available online). Cells transfected with HIV-1NL4-3(MA-3xHA/PR) proviral plasmids were grown in the presence of ribonucleoside analogs (4SU or 6SG), which also had minimal effects on infectious virion yield (Figures S1C and S1D).

Cells and virions were UV-irradiated, lysed, and digested with ribonuclease A. Then, Gag-RNA adducts were immunopurified, end-labeled with γ-32P-ATP, and visualized after SDS-PAGE and transfer to nitrocellulose membranes. Gag-RNA adduct formation was dependent on UV irradiation in cells and virions (Figure 1A). We primarily used 4SU-based CLIP thereafter, because it efficiently generated crosslinks, whereas 6SG and unmodified RNA was used for confirmatory purposes. Gag-crosslinked RNA oligonucleotides were purified, sequenced, and mapped to the HIV-1 and human genomes (see Extended Experimental Procedures). Reads derived from the terminal repeat (R) region of the HIV-1 genome ambiguously map to 5' and 3' ends but are displayed at 5' end of the viral genome, and cautiously interpreted, in our analyses.

In six independent Gag-CLIP libraries prepared from HIV-1NL4-3(MA-3xHA/PR) or HIV-1NDK(MA-3xHA/PR) expressing cells, 2.5%–7% of the total reads were HIV-1 derived, whereas ~60% were from host cell RNA (Figure 1B; Table S1). In comparison, RNA sequencing (RNA-seq) libraries, which measure the abundance of cellular and viral RNAs, contained 0.3%–1.6% of reads derived from HIV-1, and ~75% were from cellular RNA (Figure 1B; Table S1). In immature virions, ~50% and 20% of CLIP reads were from viral and host RNAs, respectively (Figure 1B; Table S1), broadly similar to RNA abundance as determined by RNA-seq (Figure 1B). Thus, viral RNA sequences were somewhat selectively bound by Gag in cells but were enriched to a far greater extent in virions.

**HIV-1 RNA Sequences Bound by Gag in Cells and Virions**

We plotted the frequencies with which each nucleotide in the HIV-1NL4-3 genome was represented in reads from cell- and virion-derived HIV-1NL4-3(MA-3xHA/PR) Gag-CLIP libraries (“read density”, Figure 1C). In cells, a major proportion of the Gag-linked reads were derived from discrete sites in the viral genome. As might be expected, the 5' leader was a frequent site of Gag binding. However, additional sites of frequent Gag binding included the RRE, sequences overlapping the Nef start codon, and untranslated sequences in U3 (Figure 1C). Several other sites on the viral genome were bound by Gag, but at lower frequencies. The distribution of Gag-crosslinked viral RNA reads was highly reproducible in HIV-1NL4-3(MA-3xHA/PR) Gag-CLIP-seq libraries as indicated by the nearly perfect correlation between independent experiments (Figure 1D).

However, when a divergent HIV-1NDK strain was included in these analyses, the only prominent Gag binding sites that were present in both strains were the 5' leader and RRE (Figure 1C).

The frequency with which Gag was crosslinked to sites across the viral genome was starkly different in immature virions. Sites of frequent Gag binding on HIV-1NL4-3 RNA were far more numerous in immature virions than in cells. Moreover, specific Gag binding to the 5' leader and RRE was not evident in immature virions (Figure 1C). The pattern of Gag binding frequency across the viral genome was highly reproducible (Figure 1D) and was unlikely to be generated by methodological bias during CLIP-seq library generation, because it was largely unaffected by the choice of ribonucleoside analog (4SU versus 6SG; Figures S1E and S1F), ribonuclease (RNase A versus RNase T1; Figures S1G and S1H), or the immunoprecipitating antibody (anti-HA versus anti-NC; Figures S1I and S1J) used to generate the CLIP libraries. HIV-1NDK Gag similarly bound to sites throughout the viral genome (Figure 1C). Despite overall similarity, there were some clear discrepancies in viral RNA sites that were frequently occupied by HIV-1NL4-3 and HIV-1NDK Gag (Figures 1E and S1K), presumably due to differences in target RNA sequence, or subtle differences in the RNA binding specificities of the two Gag proteins.

The 3xHA tag did not affect the pattern of Gag binding at sites proximal or distal to the insertion site (Figure S1L). Moreover, autocorrelation analysis of Gag binding frequency revealed no peaks other than at a separation of s = 0 (Figure S1M). Thus, the frequency with which a given site in the viral genome was bound by Gag in immature virions appeared unaffected by its position relative to the Ψ sequence or other Gag binding sites. Rather, Gag binding appeared to be a function of local nucleotide sequence or structure.
Analysis of Gag Binding to 5’ Leader and RRE
In cells, the most prominent Gag binding sites on the viral genome coincided with the most prominently structured elements in the HIV-1 genome, namely, the 5’ leader and the RRE (Figure 1C). However, Gag was not bound across the 5’ leader and the RRE with uniform frequency, but selectively associated with small determinants within these structures. In the case of the 5’ leader, Gag was most frequently bound to three distinct sequences, including one at the 5’ end of U5, a second site between the PBS and the major splice donor, and a third site 3’ to the major splice donor (Figure 2A). These sites are separated from each other in linear sequence by approximately 100 nucleotides. Strikingly, however, a nuclear-magnetic-resonance-based analysis of the structure of a Ψ sequence (Lu et al., 2011a) predicts that these three Gag-binding sequences would be in close proximity, and partly base paired with each other upon RNA folding into a structure that favors genome packaging (Figure 2B).

Frequent Gag binding to the RRE occurred in stem I, a site clearly distinct from the primary site of Rev-RRE interaction in stem-loop IIB (Malim et al., 1990) (Figures 2C and 2D). Importantly, the Gag-bound reads derived from both the 5’ leader and RRE contained high rates of T-to-C substitution, identifying individual nucleotide bases that were in close proximity to Gag.
Additionally, CLIP-seq experiments performed using 4SU or 6SG yielded a similar footprint of Gag on viral RNA, with 5' leader and RRE binding prominently featured (Figures S2A and S2B). Thus, the apparently specific binding of Gag to these sequence elements was not an artifact of the particular crosslinking nucleotide used.

Changes in Gag RNA Binding to Viral RNA during HIV-1 Virion Assembly

Given the stark differences in the interaction of Gag with viral RNA in cells and immature virions, we attempted to determine what triggers this dramatic change. To this end, cells were fractionated after UV irradiation (Figure 3A), and Gag-RNA adducts
were immunoprecipitated from cytosol, in which Gag is primarily monomeric, and from a membrane fraction where multimerized assembly intermediates form (Kutluay and Bieniasz, 2010) (Figure 3B). The RNA binding profile of cytosolic Gag was nearly indistinguishable from that described above for total cell lysates, with Ψ and RRE representing specific binding sites (Figures 3C and 3D). Similarly, a myristoylation-defective Gag mutant (Gag-G2A), which does not bind efficiently to membranes, was also restricted to discrete binding sites on viral RNA (Figure S3A). Conversely, Gag that was immunoprecipitated from membrane fractions was bound to many sites on the viral genome, and, although binding to Ψ and RRE remained prominent, the overall

Figure 3. Changes in Gag Binding to Viral RNA during Virion Assembly and Maturation
(A) Western blot analysis of Gag, and markers of cytosol (LC3A/B) and plasma membrane (Na-K ATPase) in fractionated, 4SU-fed 293T cells transfected with a HIV-1NL4-3(MA-3xHA/PR) proviral plasmid.
(B) Autoradiogram of Gag-RNA complexes recovered from fractions in CLIP assays.
(C) Read density distribution on viral RNA from CLIP experiments in which WT and mutant Gag proteins were immunoprecipitated from cell fractions.
(D) Correlation analysis of CLIP data from cell fractions and immature virions.
(E) Autoradiogram of NC-RNA and Gag-RNA complexes recovered from mature and immature virions using an anti-NC antibody.
(F) Read density distribution on viral RNA from CLIP experiments in which Gag and NC proteins were immunoprecipitated from immature and mature virions using an anti-NC antibody.
(G) Correlation analysis of CLIP data from cells and mature and immature virions.
See also Figure S3.
RNA binding profile of cell-membrane-associated Gag otherwise resembled that of immature virion-associated Gag (Figures 3C, 3D, and S3B). A late-domain mutant Gag protein (Δp6) was similarly bound to sites throughout the viral genome (Figure 3C) indicating that the completion of budding was not required for this apparent shift in RNA binding specificity. Notably, a Gag CAΔCTD mutant that is fully competent to localize to the plasma membrane but is significantly impaired in the formation of high-order multimers (Kutluay and Bieniasz, 2010) exhibited a binding profile that was more reminiscent of cytosolic Gag, although additional sites in the viral genome were bound with some prominence (Figure 3C). Together, these data indicate that high-order oligomerization of Gag at the plasma membrane drives a profound change in Gag’s RNA-binding properties (Figure 3C, 3D, and S3B), dramatically increasing the extent to which Gag and viral RNA interact with each other.

Changes in Gag-Viral RNA Binding Triggered by Virion Maturation

Gag undergoes proteolytic cleavage during virion maturation, liberating NC, which is then thought to condense with the viral RNA inside a remodeled conical core (Sundquist and Kräusslich, 2012). To determine whether virion maturation affects Gag–viral RNA interactions, we did CLIP experiments using an antibody that specifically recognizes NC (Figures S3D and S3E) and could efficiently immunoprecipitate Gag-RNA and NC-RNA adducts from immature and mature virions, respectively (Figure 3E). Comparison of RNA sequences bound by Gag in immature virions and NC in mature virions revealed profound changes accompanying virion maturation (Figures 3F and 3G). Unlike intact Gag in immature virions, NC in mature virions preferentially occupied discrete sequences on the viral genome, that coincided in large part with the major sites of Gag-viral RNA interaction in the cytosol (i.e., ΔΨ and RRE) (Figures 3F and 3G). Many of the additional prominent NC-binding sites in mature virion RNA were also bound by intact Gag in the cytosol at lower frequencies (Figure S3F). Thus, the RNA-binding properties of NC in mature virions resemble that of unassembled Gag in the cytosol to a surprising extent, exhibiting statistically significant correlation (Figure 3G), possibly reflecting the monomeric state of cytosolic Gag and mature NC.

Gag Binding to Viral RNA and Viral RNA Packaging in the Absence of ΔΨ or RRE

To test the importance of the Gag/NC-binding sites in ΔΨ and RRE, we generated viruses carrying deletions of these sequences. First, we determined whether the Gag binding to additional sites on the viral genome required initial binding to ΔΨ. To this end, we performed CLIP experiments using viral constructs carrying deletions of two of the three Gag-binding sequences (Δ105-278), or all three Gag binding sequences (Δ105-278/Δ301-332) within ΔΨ. These deletions left sequences surrounding the major 5′ splice donor intact, and only modestly reduced Gag expression (Figure 4A). Surprisingly, in cells, neither the Δ105-278 nor the Δ105-278/Δ301-332
mutation affected the fraction of Gag-crosslinked reads derived from viral versus human genomes (Figure 4B). Moreover, in immature virions, these Ψ-deletions caused only a modest ∼3-fold decrease in the fraction of reads derived from the viral genome, with a corresponding increase in the fraction of reads from cellular RNAs (Figure 4B). The pattern of Gag binding to viral RNA in cells was not greatly affected by the Ψ deletions (Figures 4C and 4D), although there was an apparent reduction in the frequency with which RRE was bound. Similarly, the overall pattern of Gag binding to viral RNA in immature virions was also relatively unperturbed by Ψ deletions (Figures 4C, 4D, and S4A). However, there was a tendency for read densities to increase toward the 3’ end of the viral genome, perhaps reflecting the incorporation of spliced viral RNAs into virions generated by Ψ-deleted genomes (Cleaver and Parslow, 1997; Houzet et al., 2007; Russell et al., 2003). Importantly, these results indicate that multiple sites on the viral genome bind to Gag independently of Ψ and facilitate HIV-1 genome packaging.

To analyze the effects of Gag-RRE interaction on HIV-1 infectivity, we generated viruses in which Rev was deleted and the nuclear export of unspliced viral RNAs was mediated by the Mason-Pfizer monkey virus constitutive transport element (CTE) to circumvent effects of RRE on RNA nuclear export (Figure S4B). In this setting, the presence or absence of RRE had no discernible effect on Gag protein levels or infectious virus yield (Figure S4C). Although it is formally possible that the MPMV CTE might recapitulate the Gag-binding properties of RRE, this analysis suggests that Gag RRE interactions do not regulate Gag expression and are not required for RNA packaging.

Changes in the Sequence Specificity of Gag RNA Binding during Virion Genesis

The aforementioned results suggest a degree of redundancy, and contributions from multiple domains of the viral genome to RNA packaging. In an attempt to explain the factors driving the selectivity with which Gag bound to and packaged viral RNA, we determined the identities of cellular mRNA sequences that were most frequently bound to Gag in cells and virions. Reads that aligned to the human genome were clustered using PARalyzer (Corcoran et al., 2011), which defines a cluster, or binding site, based on the occurrence of a minimum number overlapping reads proximal to a T-to-C substituted crosslinking site in 4SU-based CLIP assays (Table S2). Note that a single cluster is counted once for the analysis shown in Figure 5A, irrespective of the number of reads associated with it. In cells, >90% of clusters (Gag binding sites) were within genes, and ∼80% of these were derived from mRNAs (Figure 5A) with an overlap of >95% between HIV-1NL4-3 and HIV-1NL4-3 CA-NC-p6) only one of which would be tagged and immunoprecipitated (Figures 6B and S6A). After UV crosslinking, cell lysates could be treated with FXa protease cleavage site immediately N- or C-terminal to the 3xHA-tag (Figures 6B and S6A). Although the aforementioned analysis focused on Gag binding to mRNA sequences, when the number of individual Gag-bound reads associated with each cluster or binding site was counted, it was evident that mRNAs were responsible for only a fraction (∼12%) of the cellular RNA reads crosslinked to Gag. In fact, tRNAs were the dominant RNA species bound by Gag in cells (As-60%–70% of reads), with 7SL RNA, a known component of HIV-1 virions (Onafuwa-Nuga et al., 2006) constituting the bulk of the remaining reads (Figure 6A). Strikingly, the fraction of Gag-bound RNA in immature virions, RNA binding specificity reverted, and the sequence motifs favored by NC were again G/U-rich. These findings reinforce the notion that Gag binds preferentially to RNAs that have an A-rich nucleotide composition, and particularly to A-rich motifs that are present in the viral genome, transiently during immature virion assembly.

trRNAs Are the Most Frequently Gag-Bound RNA in Cells, But Not in Virions

Although the aforementioned analysis focused on Gag binding to mRNA sequences, when the number of individual Gag-bound reads associated with each cluster or binding site was counted, it was evident that mRNAs were responsible for only a fraction (∼12%) of the cellular RNA reads crosslinked to Gag. In fact, tRNAs were the dominant RNA species bound by Gag in cells (As-60%–70% of reads), with 7SL RNA, a known component of HIV-1 virions (Onafuwa-Nuga et al., 2006) constituting the bulk of the remaining reads (Figure 6A). Strikingly, the fraction of tRNAs derived reads was greatly reduced in CLIP-seq libraries prepared from virions as compared to cells. This was not the case for other RNA classes, suggesting that tRNAs were specifically excluded from Gag-bound RNAs in virions (Figure 6A).

MA Binds to Particular tRNAs

Although the aforementioned Gag-RNA binding events were assumed to involve the NC domain of Gag, there are numerous reports that MA can interact with viral or cellular RNAs (reviewed in Alfadhli and Barklis, 2014). However, these interactions, generally assayed in vitro, are of uncertain significance. To test whether MA binds to RNA in cells, we modified the aforementioned HIV-1NL4-3(MA-3xHA/PR) clones to include a FXa protease cleavage site immediately N- or C-terminal to the 3xHA-tag (Figures 6B and S6A). After UV crosslinking, cell lysates could be treated with FXa protease to generate two Gag fragments (MA and CA-NC-p6) only one of which would be tagged and immunoprecipitated (Figures 6B and 6C).
Like full-length Gag, the immunoprecipitated CA-NC-p6 contained crosslinked RNA species that caused adducts to migrate ~0–10 kDa above the expected molecular weight of the protein. Notably, the isolated MA domain also generated prominent crosslinked species, indicating RNA binding. Unusually, these MA-RNA adducts migrated at a discrete molecular weight of ~35 kDa, suggesting that MA bound to a unique RNA species (Figure 6D).

Although the CA-NC-p6-RNA adducts contained viral RNA sequences with a very similar distribution to those recovered from full-length Gag, MA-RNA adducts were nearly devoid of viral RNA-derived sequence (Figure 6E). In fact, the RNA molecules crosslinked to MA were almost exclusively host tRNAs (Figure 6F). When RNase treatment was omitted from the CLIP procedure prior to immunoprecipitation, MA-RNA adducts migrated at a larger, but still discrete, molecular weight (Figure 6D) matching the size of one tRNA molecule (~20 kDa) crosslinked to MA, suggesting that MA binds to full-length tRNAs.

Although tRNAs constituted the majority of cellular RNA molecules bound by intact Gag in cells, deletion of MA caused a major...
reduction in the fraction of reads that were tRNA derived but did not inhibit binding to mRNAs or 7SL RNA (Figures 6F and S6B). Conversely, a mutant Gag protein lacking the NC domain did not bind mRNA or 7SL, but bound nearly exclusively to tRNAs (Figures 6F and S6B). Progressive and specific loss of tRNA binding occurred as more lysines in the N-terminal MA basic domain were substituted, and mutation of six lysines (MAK6T) reduced Gag-tRNA binding to the same degree as deletion of MA (Figures 6F and S6B).

**Figure 6. HIV-1 MA Binds to Specific tRNAs, but Not Viral RNA**

(A) Classification of individual reads from cell and immature virion Gag-CLIP experiments that map to genes in the host cell genome.

(B) Schematic representation of FXa-cleavable Gag proteins.

(C and D) Western blot analysis (anti-HA, C) and autoradiography (D) of immunoprecipitated (anti-HA) proteins from HIV-1NL4-3 MA-FXa-3xHA/PR and MA-3xHA-FXa/PR-transfected cells that were subjected to the modified CLIP procedure. Cell lysates were treated with Factor Xa protease, with or without RNase A prior to immunoprecipitation. *Breakdown products of the 3xHA-CA-NC-p6 protein that retain the HA tag but are not crosslinked to RNA.

(E) Read density distribution on viral RNA following CLIP analysis of FXa-liberated 3xHA-CA-NC-p6 and MA-3xHA proteins.

(F) Classification of individual reads that map to cellular genes from cell associated Gag-CLIP experiments employing the FXa-liberated MA-3xHA protein or WT and mutant, uncleaved Gag proteins. Total number of reads is indicated below.

(G) Numbers of reads crosslinked to FXa-liberated MA-3xHA or uncleaved Gag proteins that lack NC (∆NC) and MA (∆MA) that map to each tRNA gene.

(H) Two examples of read density distributions (red) and T-to-C substitutions (blue) at crosslinking sites on tRNAs Gly[GCC] and Lys[TTT] following CLIP analysis of FXa-liberated MA-3xHA.

See also Figure S6.
As expected, MAK4T and MAK6T mutations caused decreases in the levels of released particles (Figure S6B), consistent with a plasma membrane targeting defect in these mutants.

MA binding to tRNAs was highly selective. Indeed, GluCTC, GluTTC, GlyGCC, GlyCCC, LysCTT, LysTTT, ValAAC, and ValCAC tRNAs were bound up to 100-fold more frequently than the majority of tRNAs (Figure 6G). No such enrichment occurred in CLIP experiments done with GagD. Although intact tRNA molecules were bound by MA (Figure 6D), binding apparently involved the 5′ half of the tRNA molecules, and especially the dihydrouridine loop, as indicated by the very high rates of T-to-C conversions at this site (Figure 6H). Binding was not induced by 4SU incorporation into tRNAs, because the same tRNAs were selectively MA-bound in CLIP experiments where crosslinking was induced by UV irradiation at 254 nm in the absence of modified nucleotides (Figure S6C).

Regulation of Gag Membrane Binding by tRNA
The finding that the same lysine residues that mediate plasma membrane binding also mediate tRNA binding, suggested that tRNAs might regulate Gag localization. Indeed, previous work has indicated that exogenous RNA can inhibit Gag binding to liposomes in vitro (Chukkapalli et al., 2010, 2013; Dick et al., 2013). We took a simple cell-based approach (without addition of exogenous RNA or liposomes) and tested whether RNase treatment of cell lysates increased Gag binding to endogenous cellular membranes. Lysates of HeLa and HOS cells stably expressing Gag-CFP were treated with ribonucleases and subjected to membrane flotation analysis. Strikingly, RNase treatment caused a significant redistribution of Gag from cytosol to membrane fractions (Figures 7A and 7B), whereas a control protein, β-tubulin, was largely unaffected. Consistent with a model in which tRNAs compete with membranes for Gag binding, tRNAs comprised a significantly smaller fraction of Gag-bound RNAs at the plasma membrane than in the cytoplasm (Figure 7C). This difference was more pronounced when particle budding was blocked at the plasma membrane by deletion of the late budding domain of Gag (Figure 7C). Thus, these results indicate that MA can bind to tRNA or cell membranes, but not both simultaneously, and strongly suggest that tRNAs regulate Gag localization by binding to basic amino acids in MA.

DISCUSSION
Two central conclusions of these studies are (1) the HIV-1 Gag protein has two RNA binding domains (NC and MA) with very different specificities and (2) dramatic changes in RNA binding regulate Gag localization and genome packaging during virion genesis (Figure 7D).

Prior to virion assembly, Gag exists as a diffuse pool of monomers or low-order multimers in the cytoplasm with its NC domain.

6F and S6B). As expected, MAK4T and MAK6T mutations caused decreases in the levels of released particles (Figure S6B), consistent with a plasma membrane targeting defect in these mutants.

MA binding to tRNAs was highly selective. Indeed, GluCTC, GluTTC, GlyGCC, GlyCCC, LysCTT, LysTTT, ValAAC, and ValCAC tRNAs were bound up to 100-fold more frequently than the majority of tRNAs (Figure 6G). No such enrichment occurred in CLIP experiments done with GagD. Although intact tRNA molecules were bound by MA (Figure 6D), binding apparently involved the 5′ half of the tRNA molecules, and especially the dihydrouridine loop, as indicated by the very high rates of T-to-C conversions at this site (Figure 6H). Binding was not induced by 4SU incorporation into tRNAs, because the same tRNAs were selectively MA-bound in CLIP experiments where crosslinking was induced by UV irradiation at 254 nm in the absence of modified nucleotides (Figure S6C).

Regulation of Gag Membrane Binding by tRNA
The finding that the same lysine residues that mediate plasma membrane binding also mediate tRNA binding, suggested that tRNAs might regulate Gag localization. Indeed, previous work has indicated that exogenous RNA can inhibit Gag binding to liposomes in vitro (Chukkapalli et al., 2010, 2013; Dick et al., 2013). We took a simple cell-based approach (without addition of exogenous RNA or liposomes) and tested whether RNase treatment of cell lysates increased Gag binding to endogenous cellular membranes. Lysates of HeLa and HOS cells stably expressing Gag-CFP were treated with ribonucleases and subjected to membrane flotation analysis. Strikingly, RNase treatment caused a significant redistribution of Gag from cytosol to membrane fractions (Figures 7A and 7B), whereas a control protein, β-tubulin, was largely unaffected. Consistent with a model in which tRNAs compete with membranes for Gag binding, tRNAs comprised a significantly smaller fraction of Gag-bound RNAs at the plasma membrane than in the cytoplasm (Figure 7C). This difference was more pronounced when particle budding was blocked at the plasma membrane by deletion of the late budding domain of Gag (Figure 7C). Thus, these results indicate that MA can bind to tRNA or cell membranes, but not both simultaneously, and strongly suggest that tRNAs regulate Gag localization by binding to basic amino acids in MA.

DISCUSSION
Two central conclusions of these studies are (1) the HIV-1 Gag protein has two RNA binding domains (NC and MA) with very different specificities and (2) dramatic changes in RNA binding regulate Gag localization and genome packaging during virion genesis (Figure 7D).

Prior to virion assembly, Gag exists as a diffuse pool of monomers or low-order multimers in the cytoplasm with its NC domain.
bound primarily to mRNA, with some binding to 7SL and tRNA (Figure 7D). A fraction of NC-RNA interactions are with particular sites on the HIV-1 genome, including the ψ sequence. Our data reveal the specific RNA sequences within ψ that are in proximity to Gag, in a physiological setting. Satisfyingly, the three noncontiguous RNA elements that are most frequently crosslinked to Gag (nucleotides 100–126, 195–260, and 300–350) coincide nearly precisely with a minimal ψ element, which adopts a secondary structure that putatively favors genome packaging (Lu et al., 2011a).

Surprisingly, cytoplasmic Gag bound to additional discrete elements on the viral RNA, including RRE. Previously, env sequences, including RRE, were shown to facilitate packaging (Kaye et al., 1995; Richardson et al., 1993), but a discrete packaging sequence within env could not be mapped. Other reports suggest that Rev enhances packaging (Brandt et al., 2007), although effects of Rev/RRE on viral RNA nuclear export are potential confounders in packaging experiments. Despite strong evidence for Gag-RRE binding, the RRE did not affect infectious virion yield when its nuclear export function was replaced, suggesting that Gag-RRE binding is not required for genome packaging. Although it is possible that Gag/NC-RRE interaction plays a redundant role in packaging, other plausible functions include (1) shielding double-stranded stem I RNA from cytoplasmic sensors, (2) coupling RNA-export with packaging, (3) displacement of Rev from the RRE for recycling, (4) regulation of Env translation, and (5) enhancement of reverse transcription via NC’s unwinding/chaperone function (Levin et al., 2010). Further work will be required to elucidate the functional role, if any, of Gag/NC-RRE binding.

In the cytosol, Gag favored binding to the viral genome over cellular mRNAs by a few fold. This level of discrimination is insufficient to account for the selectivity with which viral genomes are packaged. Our data indicate that a dramatic change in Gag-RNA binding specificity, coincident with CA-CTD-dependent high-order multimerization at the plasma membrane, contributes to selective packaging (Figure 7D). GU-rich sequences in cellular mRNA were targeted by Gag in the cytosol, consistent with previous in vitro and structural studies indicating that the isolated, monomeric NC domain favors binding to such sequences (Berglund et al., 1997; De Guzman et al., 1998; Fisher et al., 1998). However, during assembly, Gag molecules become tightly packed in hexameric lattices (Briggs et al., 2009; Wright et al., 2007). By constraining thousands of NC domains into a pseudo-2D curved array, local NC concentration is dramatically elevated. Potentially, features of NC that govern RNA binding specificity (Zinc knuckles and basic amino acids) might be differently accessible in an assembled Gag lattice. Under these conditions, we found that A-rich mRNA sequences were preferentially bound by Gag (Figure 7D). Remarkably, the nucleotide composition of mRNA sequences bound by assembled Gag reflects an unusual, heretofore unexplained, property of the HIV-1 genome. Thus, our findings suggest that a need to selectively package viral RNA caused HIV-1 to evolve an unusually A-rich genome. Conversely, an A-rich genome may have evolved for other reasons and then drove Gag to gain a unique oligomerization-driven specificity for A-rich RNA. Notably, ψ-deletion caused only a 3-fold reduction in the fraction of virion-associated, Gag-bound, RNA sequences that were viral RNA derived. Moreover, the pattern of Gag binding to many sites in the viral genome was not solely a secondary effect of physical proximity to Gag-ψ interactions. Rather, it appears that HIV-1 genome packaging is a two-step process, involving interactions between (1) ψ and monomeric Gag and (2) A-rich viral RNA and multimeric Gag. This scenario should selectively drive particle assembly on viral RNAs, and we speculate that the biases in nucleotide composition exhibited by HIV-1 and other retroviruses serves as a proofreading-like mechanism to enhance the fidelity of genome packaging following initial Gag-ψ interaction.

An unexpected finding was that proteolytic cleavage of Gag caused NC to revert to a preference for GU-rich mRNAs and discrete viral RNA sequences. This result reinforces the notion that Gag/NC RNA binding specificity is multimerization dependent. By liberating the majority of viral RNA from NC, while maintaining interaction with structured elements (to enable NC’s chaperone activity), maturation-dependent changes in Gag/NC-RNA binding could facilitate reverse transcription (Levin et al., 2010).

Another surprising finding was that MA binding to specific tRNAs constitutes the most frequent binding events between Gag and RNA in cells. MA-tRNA binding was independent of NC and the PBS and is thus unlikely to involve the tRNA primer annealed to viral genome. Rather, we found that MA-tRNA interaction could regulate the binding of Gag to cellular membranes. MA specifies the location of virion assembly, and it was previously shown that RNA can block in vitro MA binding specifically to liposomes that lack acidic phospholipids (Chukkapalli et al., 2010, 2013). Thus, occlusion of MA basic residues by specific tRNAs might inhibit nonproductive assembly at most intracellular membranes and facilitate targeting to the plasma membrane where resident lipids have a high affinity for MA. Alternatively, MA-tRNA binding might provide a mechanism by which virion assembly is temporally regulated.

MA-tRNA interactions could serve additional purposes. Some degree of RNA binding may be an inevitable consequence of encoding a highly basic domain. Thus, specific MA binding in a 1:1 complex to small RNAs might be a mechanism to avoid the aggregation of a protein that has both two distinct RNA binding domains and a tendency to multimerize. MA-tRNA binding might also prevent nonproductive interaction of a viral genome with a Gag monomer whose NC domain has engaged viral RNA. Conceivably, MA may facilitate the selection of the RT primer, as tRNAlys3 is among the tRNAs bound by MA, but several other tRNAs are also bound by MA more frequently. Finally, MA-tRNA interaction could regulate viral and/or host translation. The richness of the HIV-1 genome results in suboptimal codon usage (Grantham and Perrin, 1986; Kypr and Mářek, 1987; Sharp, 1986) and an elevated number of Ile, Lys, Glu, and Val codons in the Gag and Pol ORFs (Berkhout and van Hemert, 1994). Notably, Lys, Glu, and Val tRNAs are among those specifically bound by MA, providing a potential opportunity for Gag to regulate its own translation as it accumulates to high levels and sequesters tRNAs (perhaps facilitating packaging as a consequence). Similarly, MA could inhibit translation of host mRNAs whose products may be deleterious for viral replication. Indeed, interaction of MA with host elongation factors via a tRNA bridge.
has been reported to inhibit translation in vitro (Cimarelli and Luban, 1999).

Overall, our global survey reveals surprising ways in which the interaction between Gag, viral, and host RNAs can change and modulate the process of virion genesis and genome packaging.

**EXPERIMENTAL PROCEDURES**

**Proviral Plasmids and Cells**

HIV-1NL4-3-derived proviral plasmids containing a 3xHA tag in the stalk region of matrix (HIV-1NL4-3MA-3xHA) were constructed using overlap extension PCR. Various derivatives of this construct encoding a catalytically inactive viral protease (PR-), a Factor Xa cleavage site on either side of the HA tag (MA-FXa-3xHA and MA-3xHA-FXa), deletions of nucleocapsid domain (ΔNC), the globular head of MA (ΔMA), the CA CTD (CA ΔCTD), or the Ψ signal (Δ105-278 and Δ105-278/Δ301-332) were constructed using PCR-based deletion mutagenesis. Constructs carrying mutations at binding sites for Tag101 and ALIP proteins in the p6 domain of Gag (Δp6), the Gag myristoylation signal (G2A), or at tyrosine residues in MA (MAK26,27T, MAK4T, and MAK6T) were constructed using PCR overlap extension-based mutagenesis. Proviral plasmids with deletions or mutations in Env, Rev, and RRE and encoding a Mason-Pfizer monkey virus constitutive transport element were similarly constructed. Details of the construction are described in Extended Experimental Procedures.

**PAR-CLIP, HTS-CLIP, and RNA-Seq**

For CLIP experiments, HEK293T cells were grown in 10 cm dishes and transfected with proviral plasmids using polyethylenimine (PolySciences). Virions were harvested from filtered supernatant by ultracentrifugation through sucrose and UV irradiated, whereas cells were irradiated while adhered to culture dishes. Prior to UV crosslinking, a fraction of cells and virions were collected for RNA-seq analysis. After crosslinking, the CLIP procedure was performed on unfraccionated lysates (after removal of nuclei) or on membrane and cytoplasmic fractions.

For the CLIP procedure, cell and virion lysates were treated with RNaseA or RNaseT1 and then incubated with Protein G-conjugated Dynabeads coated with anti-HA or anti-NC antibodies. After immunoprecipitation of RNA-protein adducts, beads were washed and treated sequentially with calf intestinal alkaline phosphatase and then polynucleotide kinase and γ[32P]-ATP. RNA-protein adducts were eluted from the beads, separated by SDS-PAGE, blotted onto nitrocellulose, and digested with proteinase K. RNA oligonucleotides were then ligated to adaptors, amplified by PCR and sequenced determined using an Illumina HiSeq 2000 platform.

CLIP-seq experiments were performed two to six times on cells and virions. Further details of the method are in the Extended Experimental Procedures.

**Bioinformatic Analyses**

The FASTX toolkit was used to process the reads from CLIP and RNA-seq libraries before mapping. Reads were separated based on their 5' barcode sequences and collapsed to generate a set of unique sequences. Unique CLIP-seq and RNA-seq reads were mapped to the human (hg19) and HIV-1 genomes using the Bowtie. SAMtools, in house scripts, and GraphPad Prism were used to calculate and display read densities associated with viral and genomes using the Bowtie. SAMtools, in house scripts, and GraphPad Prism.

**Membrane Flotation Assays**

The membrane flotation assays were performed using HeLa and HOS cells stably expressing cyan fluorescent protein-tagged Gag proteins as described (Kutluay and Bieniasz, 2010) with modifications outlined in the Extended Experimental Procedures.

**ACCESSION NUMBERS**

CLIP-seq data were deposited to the GEO database with accession number GSE61508.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.09.057.

**AUTHOR CONTRIBUTIONS**

S.B.K. and P.D.B. conceived the study, designed the experiments, analyzed the data, and wrote the paper. S.B.K., T.Z., C.P., and D.J. executed the experiments. S.B.K., D.B.M., and M.E. analyzed the sequencing data. P.D.B. supervised the work.

**ACKNOWLEDGMENTS**

We thank Markus Hafner and members of the Darnell and Tuschi labs for advice and helpful discussions. We thank Michael Summers for the diagram of the Ψ and RRE sequences in Figure 2. S.B.K. was supported in part by an AmFAR Mathilde Krim Postdoctoral Fellowship. This work was supported by NIH grants R01AI501111 and P50GM103297.

Received: August 3, 2014

Revised: September 17, 2014

Accepted: September 24, 2014

Published: November 6, 2014

**REFERENCES**


donor site are required for both packaging and dimerization of human immuno deficiency virus type 1 RNA. J. Virol. 77, 84–96.


