

Reconstitution of selective HIV-1 RNA packaging *in vitro* by membrane-bound Gag assemblies

Lars-Anders Carlson^{1,2}, Yun Bai^{1,6}, Sarah C. Keane⁴, Jennifer A. Doudna^{1,2,3,5}, James H. Hurley^{1,2,5}

¹Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA

²California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, CA 94720, USA

³Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA;

⁴Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland Baltimore County (UMBC), 1000 Hilltop Circle, Baltimore, MD 21250, USA.

⁵Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA

⁶Current address: School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, China

Corresponding authors: L.-A. C. lacarlson@berkeley.edu or J. H. H. jimhurley@berkeley.edu

17

18 **Abstract**

19 HIV-1 Gag selects and packages a dimeric, unspliced viral RNA in the context of a large excess of
20 cytosolic human RNAs. As Gag assembles on the plasma membrane, the HIV-1 genome is enriched
21 relative to cellular RNAs by an unknown mechanism. We used a minimal system consisting of purified
22 RNAs, recombinant HIV-1 Gag and giant unilamellar vesicles to recapitulate the selective packaging of
23 the 5' untranslated region of the HIV-1 genome in the presence of excess competitor RNA. Mutations in
24 the CA-CTD domain of Gag which subtly affect the self-assembly of Gag abrogated RNA selectivity. We
25 further found that tRNA suppresses Gag membrane binding less when Gag has bound viral RNA. The
26 ability of HIV-1 Gag to selectively package its RNA genome and its self-assembly on membranes are thus
27 interdependent on one another.

28

Introduction

Packaging the correct genetic material is a key event in the assembly of an infectious virus. In the case of HIV-1, both the recognition of the genomic RNA, its targeting to the plasma membrane, and assembly of a membrane-enveloped virus around it is carried out by the viral protein Gag (Rein et al. 2011, Sundquist et al. 2012, Freed 2015). The nucleocapsid (NC) domain of Gag has two zinc knuckle motifs which together with adjacent residues interact with the viral RNA (Berkowitz et al. 1993, Lu et al. 2011). Mutations in NC that interfere with RNA binding lead to non-infectious viruses (Aldovini et al. 1990, Gorelick et al. 1990). Transcription of the integrated HIV-1 provirus results in both spliced and unspliced transcripts, coding for different viral proteins (Coffin et al. 1997). Gag selectively packages unspliced, full-length genomic RNA through interactions with its 5' untranslated region (5'UTR), which is different in spliced and unspliced constructs (Lu et al. 2011, Kuzembayeva et al. 2014). In addition to sequences that mediate the packaging of the genomic RNA, the 5'UTR also has sequence elements related to transcription processivity (TAR), binding of tRNA^{Lys-3} as primer for reverse transcription (PBS), and genome dimerization (DIS) (Johnson et al. 2010, Lu et al. 2011, Kuzembayeva et al. 2014).

Single-virion fluorescence assays indicate that >90% of released HIV-1 particles have packaged genomic RNA (Chen et al. 2009), and packaging of a dimeric genome appears to be strongly preferred (Sakuragi et al. 2003, Chen et al. 2009, Nikolaitchik et al. 2013). The packaging of a single dimeric genome is independent of RNA mass, since genomes with altered lengths are always packaged as single dimers (Nikolaitchik et al. 2013). Genomic RNAs with sequence differences are efficiently packaged as heterodimers as long as they have identical DIS sequences (Chen et al. 2009). The ψ /SL3 stem loop region in the 5'UTR was initially identified as a sequence required for efficient packaging of genomic RNA into HIV-1 virions (Lever et al. 1989). Later work characterized a more extended core encapsidation signal (CES) which also includes the U5:AUG and PBS stems, the dimerization signal (DIS) and the splice donor (SD) region (Heng et al. 2012). Unpaired and weakly paired guanosines in several regions of the CES, including but not limited to ψ /SL3 were identified as important for packaging (Abd El-Wahab et al. 2014, Keane et al. 2015, Kenyon et al. 2015). Structures of NC complexed with ψ /SL3 (De Guzman et al. 1998) and other smaller fragments from the 5'UTR (Lu et al. 2011) have provided atomic details of how the two zinc knuckles and adjacent motifs of NC interact with RNA. A crystal structure of the DIS:DIS dimerization complex (Ennifar et al. 2001) along with an NMR structure of the entire CES (Keane et al. 2015) put these structures in the context of a larger RNA. Thus, we now have a detailed understanding of RNA structural determinants for packaging at the level of individual 5'UTR dimers.

A single released HIV-1 virus contains one dimeric genomic RNA (Chen et al. 2009) and 2400 ± 700 Gag molecules (Carlson et al. 2008), indicating that the selective genome packaging may depend on more than just the binary NC-CES interaction. Indeed, cross-linking-immunoprecipitation sequencing (CLIP-seq) revealed a strikingly different RNA preference of cytosolic and membrane-bound Gag (Kutluay et al. 2014). Cytosolic Gag binds tRNA through its otherwise membrane-binding matrix (MA) domain, and associates mainly with cellular RNAs through its NC domain. The minor fraction of cytosolic Gag which binds viral RNA has a strong preference for the 5'UTR but also binds to the Rev-responsive element (RRE), whose main function is thought to be in nuclear export (Kuzembayeva et al. 2014). On the other hand, Gag in the membrane fraction has lost its tRNA association and is strongly enriched for

viral RNA, interacting with equal probability with all parts of the genomic RNA. Existing structural and biochemical data are not sufficient to explain how a pool of cytosolic Gag, binding mainly cellular RNA, assembles into virus particles which package dimeric viral genomes with 90% probability.

Here, we have reconstituted the selective genome packaging of HIV-1 in a minimal system using giant unilamellar vesicles (GUVs), recombinant myristoylated HIV-1 Gag and purified RNAs. Gag assembling on GUV membranes selectively packages the HIV-1 5'UTR at subnanomolar concentrations in the presence of excess competitor RNA. We find that the RNA selectivity of Gag is abolished by mutations that subtly affect its ability to cluster on a membrane. Using SHAPE analysis, we find that membrane recruitment of the 5'UTR by Gag does not lead to drastic changes in its secondary structure, and we further demonstrate that the association of soluble Gag with tRNA confers additional RNA packaging selectivity.

Results

In vitro reconstitution of selective RNA packaging by membrane-bound HIV-1 Gag

We previously reported an in vitro reconstitution of HIV-1 assembly using recombinant myristoylated Gag protein and PI(4,5)P₂-containing giant unilamellar vesicles (GUVs) (Carlson et al. 2012). We showed that GUV-bound Gag clusters mimic the behavior of cellular HIV-1 budding sites in their lipid requirements for assembly, nucleic acid- and ESCRT (endosomal sorting complex required for transport) protein binding. We reasoned that this system could also be used to study the RNA selection of an assembling HIV-1 virus. The 5' untranslated region (5'UTR) and the Rev-responsive element (RRE) of the HIV-1 genome (nucleotides 1-345 and 7259-7612, of the genomic RNA of HIV-1 strain NL4-3 and ARV-2/SF2, respectively), along with a nominally unstructured 378 nt single-stranded control RNA (RNA378), were produced by in vitro transcription. The 1-345 construct was previously shown to dimerize even in the absence of Gag (Heng et al. 2012). RNAs were labeled with the fluorophore Alexa488 on random guanosines, at measured average labeling efficiencies of ~1.1-1.2 fluorophores per RNA molecule. 100 nM Gag-ATTO594 was premixed with 0.5 nM fluorescent RNA, and added to GUVs which were imaged by confocal microscopy starting 10 minutes after addition of protein and RNA. Due to the reported Mg²⁺-dependence of the 5'UTR dimerization (Wilkinson et al. 2008), 2 mM MgCl₂ was included in the final buffer. The addition of Mg²⁺ to the imaging buffer made the GUV morphology more variable, but this increased morphological heterogeneity did not hamper the experiments nor the image quantification.

In order to quantitate the RNA occupancy of assemblies, ten confocal z-stacks were recorded at random positions on each sample, with each data set containing an average of 50-100 GUVs. Strong RNA fluorescence was consistently observed at Gag clusters on GUV membranes, regardless of which RNA was used (Figure 1a-c,g). This RNA recruitment was dependent on the NC domain of Gag since neither a ΔNC Gag nor Gag treated with the Zn-finger disrupting drug AT-2 (Rossio et al. 1998) was able to recruit fluorescent RNA to GUV membranes (Figure 1 - figure supplement 1). To study the RNA selectivity of assembling Gag in the reconstituted system, we included a tenfold excess (5 nM) of non-fluorescent

RNA378 when premixing RNA and Gag. After assembly on GUVs, the fluorescent 5'UTR was still visible at most Gag clusters in the presence of tenfold excess competitor RNA, whereas RRE was present at fewer Gag clusters and at lower fluorescence intensities, and fluorescent RNA378 was virtually absent (Figure 1d-f, g). Image analysis based on automatic segmentation of Gag clusters and measurement of their average RNA fluorescence confirmed the visual impressions (Figure 1g).

To further characterize the sensitivity of the in vitro system, we created a set of altered 5'UTR constructs (Figure 2a). 5'UTR 3way1 has several packaging-critical guanines around the three-way-1 junction in the CES mutated to adenines (G116A, G318A, G320A, G328A, G329A, G331A, G333A). This set of mutations was previously found to reduce packaging in a cellular packaging assay (Keane et al. 2015). 5'UTR mono has the GCGCGC sequence in DIS replaced by GAGA, creating a monomeric 5'UTR. 5'UTR 3way1-mono combines the 3way1 and mono mutations. Lastly, 5'UTR env-1 represents the first 345 nt of a spliced viral RNA, with the nucleotides after splice donor site G290 replaced by the nucleotides after splice acceptor site A5 (Stoltzfus 2009). In a competition assay similar to the one reported in Figure 1d, 100 nM Gag-ATTO594 was premixed with 0.5 nM 5'UTR-Alexa488 and 5 nM of either non-fluorescent 5'UTR construct, and added to GUVs (Figure 2b-f). In the presence of 5 nM 5'UTR very little fluorescent 5'UTR could be seen at Gag clusters on GUVs (figure 2b,g). However, in the presence of either of the altered constructs, more fluorescent 5'UTR was visible at Gag clusters (Figure 2c-f,g). Thus, the in vitro packaging assay is able to detect the lower affinity of assembling Gag for monomeric, 3way1-mutated, and spliced 5'UTRs.

In summary, a minimal system consisting of purified RNAs, Gag, and PI(4,5)P₂-containing GUVs recapitulates the selective RNA packaging of an assembling HIV-1 particle at RNA concentrations far below reported solution affinities of NC for 5'UTR. The 5'UTR is strongly selected for in a competition assay, whereas the RRE is only partially selected for after addition of competitor. 5'UTRs that are either monomerized, have packaging-critical guanines mutated to adenines, or represent a spliced viral RNA are less favored in the competition assay. Given the low concentrations at which this selectivity is observed, below the reported 17 nM affinity constant for 5'UTR-NC association (Heng et al. 2012), we turned our attention to potential avidity effects due to Gag multimerization.

RNA selectivity depends upon intact CA domain lattice contacts

We next sought to determine whether there was a connection between RNA selectivity by Gag, and its assembly into the membrane-bound hexagonal "immature" lattice found in viral budding sites and released immature virions (Ganser-Pornillos et al. 2012, Sundquist et al. 2012, Freed 2015) (Figure 3a). To this end we produced Gag proteins mutated at symmetry contacts in its lattice-forming CA domain. The W316D,M317D mutation ("Gag 2-fold") alters a motif at the twofold symmetry contact in the CA C-terminal domain (CA-CTD), which is described in literature to be important for virus assembly (Gamble et al. 1997, von Schwedler et al. 2003) (Figure 3a-b). When this protein was added to GUVs at 100 nM, it bound in a strikingly different way from wild-type Gag, often resulting in a diffuse Gag fluorescence covering the entire GUV membrane (Figure 3f,g). For comparison, wild type Gag essentially only bound

to GUVs in clusters even in the absence of RNA (Figure 3e,g). Based on the structure of the immature CA lattice in intact HIV-1 virions (Schur et al. 2015), we further produced a triple mutant ("Gag 3-fold": P170E, A174E, E177R) designed to alter a three-helix interaction at the threefold interface in CA-NTD (Figure 3a,c). This protein bound GUVs similarly to wild-type Gag, albeit with a moderate increase in the diffuse binding first observed for Gag 2-fold (Figure 3g). A Gag protein mutated at the sixfold interface in CA-CTD (Figure 3a,d) with the single-residue replacement K290A ("Gag 6-fold") reported to affect virus assembly (von Schwedler et al. 2003, Robinson et al. 2014), also showed an increase in diffuse Gag binding over wild-type Gag (Figure 3g). Combining the 2-,3-, and 6-fold mutations in one protein ("Gag 2-,3-,6-fold") resulted in an assembly phenotype similar to that of Gag 2-fold (Figure 3g). Taken together, mutations in the CA-CTD of either the W316,M317 motif at the 2-fold symmetry contact, or of K290 at the 6-fold symmetry contact, both described as leading to virus release defects (von Schwedler et al. 2003), affect the ability of Gag to cluster on membranes.

Having characterized the membrane binding of the mutant Gag proteins, we proceeded to study their RNA binding and selectivity. We first compared the membrane binding of Gag proteins in the presence of RNA to Gag proteins alone as characterized above. We observed that the presence of RNA increased the clustering of all Gag proteins, and increased the fraction of GUVs which had bound Gag (Figure 4a-j, Figure 4 - figure supplement 1). First, the capacity of the Gag proteins to recruit fluorescent RNA in the absence of competitor was studied. Each Gag construct was premixed with 0.5 nM fluorescent 5'UTR and added to GUVs. Recruitment of fluorescent 5'UTR to the GUV membrane was observed for all Gag constructs (Figure 4a-e), with a quantitation showing a slight decrease in recruitment for the Gag 2-fold and Gag 2-,3-,6-fold mutants (Figure 4k). However, when a tenfold excess of non-fluorescent competitor RNA (5 nM RNA378) was added, the Gag proteins behaved differently from each other. Wild type Gag retained as much 5'UTR fluorescence in the presence of competitor RNA as in its absence (Figure 4f,k) and the Gag 3-fold mutant showed only a minor decrease (Figure 4h,k). On the other hand, GUV-bound clusters of the Gag 2-fold, Gag 6-fold and Gag 2-,3-,6-fold mutants lost most of their 5'UTR fluorescence in the presence of competitor RNA (Figure 4g,i,j,k). Thus, RNA selectivity in this assay closely correlates with clustered binding of Gag to GUVs in the absence of RNA (compare Figures 3g and 4k). In summary, mutations which cause subtle changes in the ability of Gag to multimerize on a membrane drastically decrease the RNA selectivity of assembled Gag clusters, without greatly affecting their ability to bind RNA non-specifically.

SHAPE-based mapping of interactions between the HIV-1 5'UTR and Gag in solution and on membranes

To further understand how membrane association of Gag affects its interactions with the HIV 5'UTR, we performed selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) analysis (Merino et al. 2005) on the 5'UTR and determined its SHAPE profiles in the presence and absence of Gag constructs either in solution or in association with GUVs. First, we examined the SHAPE profile of the 5'UTR RNA in solution by itself. The RNA-only SHAPE profile is highly consistent with the recently reported solution structure of the HIV-1 RNA packaging signal (Keane et al. 2015), with the tandem three-way junctions

exhibiting modest SHAPE reactivities, reflecting the loosely base-paired nature of this region. On the other hand, well based-paired stem regions in the DIS, PBS, U5:AUG, and ψ /SL3 sequences exhibit low or no SHAPE reactivities (Figure 5a). The dimer-promoting DIS loop (G257-C262) had no SHAPE reactivity. This is not per se an unequivocal indication of the 5' UTR being dimeric since DIS has been reported to partially base-pair with U5 in the monomeric conformation of the 5'UTR (Lu et al. 2011). However, the SHAPE reactivity of the SD region G289-G292 (highlighted with a gray shaded box in Figure 5a) is consistent with the reported 3D structure of the dimeric conformation (Keane et al. 2015), not with the completely unprotected loop proposed for the monomeric 5'UTR (Lu et al. 2011). A longer 356 nt 5'UTR able to assume both monomer- and a dimer-favoring conformations (Heng et al. 2012) had a similar SHAPE reactivity to the 345 nt 5'UTR (Figure 5 - figure supplement 1b), indicating that it is present mainly as a dimer under the conditions of the study. We next compared the SHAPE profiles of isolated RNA, and RNA in complexes mimicking steps in the genome packaging. Isolated NC domain was selected to represent the initial interaction of a single Gag molecule with the 5'UTR, Δ MA(1-119)-Gag for a potential soluble Gag-RNA complex involving CA-mediated Gag-Gag interactions, and Gag clusters on GUVs as a model of viral budding sites on the plasma membrane. In all measurements, protein was kept in excess (solution protein:RNA ratio 8:1, GUV protein:RNA ratio 20:1). For SHAPE analysis on GUV-bound RNA, reactions were set up as for confocal imaging, and GUVs were found to be intact during the treatment (Figure 5 - figure supplement 2). The amount of free RNA in the GUV supernatant was typically 10-20% of the total RNA, dependent on the presence of both GUVs and Gag. The SHAPE profiles for 5'UTR were very similar in the presence of NC, Δ MA-Gag, and when bound to Gag clusters on GUVs (Figure 5b, Figure 5 - figure supplement 1). Upon Gag binding, major SHAPE changes were observed in the following clusters: U117-U122, G234-A242, A271-G273, and A302-G309 (Figure 5b). Among them, the two clusters exhibiting decreased SHAPE reactivity (G234-C242 and A271-G273) surround the two guanine-rich bulges on the DIS stem. This Gag-mediated decrease of RNA flexibility is consistent with, but does not necessarily reflect, direct Gag-RNA contacts (Kenyon et al. 2015). These clusters show a more pronounced decrease in SHAPE reactivity than the ψ loop G317-G320. The other two clusters showing major SHAPE reactivity changes, U117-U122 and A302-G309, are located around the tandem three-way junctions. Both regions exhibit higher SHAPE reactivities upon Gag binding, indicating that protein interactions lead to rearrangement of the local structure that results in higher RNA flexibility. However, GUV recruitment of the complex did not trigger additional SHAPE changes (Figure 5c, Figure 5 - figure supplement 1), indicating that the overall structure of the 5'UTR is not further altered between the states of initial Gag binding in solution and recruitment to assembled Gag on a membrane. Noteworthy, the SHAPE reactivity of the SD region G289-G292 is similar under all condition studied (Fig 5a,c, Figure 5 - figure supplement 1), indicating that the 5'UTR is maintained in the dimeric conformation over the assembly process. Taken together, we see no evidence for large changes in secondary structures in the 5'UTR when interacting with either Gag in solution or assembled Gag on a membrane, and it appears dimeric under the conditions studied.

The effects of tRNA and viral RNAs on Gag membrane association

The membrane-binding matrix (MA) domain of cytosolic Gag has been shown to bind primarily tRNA in cells (Kutluay et al. 2014). Several reports have shown that RNA binding to the basic patch on the MA domain of HIV-1 Gag can compete with its ability to bind acidic phospholipids such as PI(4,5)P₂ (Alfadhli et al. 2009, Chukkapalli et al. 2010, Alfadhli et al. 2011, Chukkapalli et al. 2013, Dick et al. 2013, Dick et al. 2014). We wanted to investigate whether Gag-tRNA interactions also play a role in selective RNA packaging. Premixing 100 nM Gag-ATTO594 with 1 μM yeast tRNA before addition to GUVs led to a virtually complete block in Gag assembly on GUVs (Figure 6a-b). Since the amount of Gag on individual GUVs is more variable than the relative RNA fluorescence at Gag clusters, a membrane flotation assay was devised to quantitate the amount of Gag assembled on membranes. 100 nM Gag-ATTO488 was mixed with 0-20 μM tRNA, and incubated for 10 minutes with 1.0 μm large unilamellar vesicles (LUVs) containing 5 mol% PI(4,5)P₂. Vesicles were subjected to flotation on a density step gradient and Gag fluorescence in the top fraction was measured. For the tRNA titration, a 50% inhibition of Gag membrane association was evident at 200 nM tRNA (Figure 6c). At 20 μM tRNA, the amount of Gag found in the vesicle fraction dropped to ~10% of the value seen in the absence of RNA (Figure 6c). Next, long viral or unspecific RNAs at 5 nM concentration were added to Gag with or without tRNA. In addition to the above studied 5'UTR, RRE, and RNA378, an RNA corresponding to the first 1947 nucleotides of the HIV-1 strain NL4-3 genomic RNA was studied (HIV1947), along with a single-stranded control RNA of similar length (RNA2266). In the absence of tRNA, all these RNAs slightly increased the amount of Gag associated with vesicles (Figure 6d). In the presence of 20 μM tRNA however, the 5'UTR and the longer HIV1947 RNA (which contains the 5'UTR sequence) enhanced membrane association by Gag more than the other RNAs (Figure 6e). RNA378 had no effect on Gag membrane association in the presence of 20 μM tRNA, whereas the effect of RRE or the longer RNA2266 was slight (Figure 6e). In summary, tRNA at submicromolar concentrations potently inhibits membrane association of Gag. Under such conditions, Gag that encounters viral RNA has a higher probability of overcoming that inhibition and associating with membrane.

Discussion

Here, we used biochemical reconstitution methods to show that a minimal system consisting of the HIV-1 Gag protein, the HIV-1 5'UTR, and PI(4,5)P₂-containing GUVs is capable of selective RNA packaging in the presence of an excess of non-specific RNA. We find a strong correlation between the ability of Gag to multimerize on a membrane and its selective packaging of viral RNA. Mutations in the lattice-forming CA domain that lessen Gag clustering on membranes (W316D, M317D and K290A, Figure 3b,d,f,g) drastically lowered its RNA selectivity (Figure 4g,i,k), without affecting the total capacity of Gag to recruit RNA to a membrane (Figure 4b,d,k). Mutations of these residues in the C-terminal part of CA (CA-CTD) have previously been described to lead to virus release defects related to aberrant virus assembly (Gamble et al. 1997, von Schwedler et al. 2003). In line with our findings, it has been reported that the amount of viral RNA coprecipitated with cytoplasmic Gag is reduced fivefold when its CA-CTD is deleted, whereas the amount of cellular RNA coprecipitated remains the same (Kutluay et al. 2010). On the other hand, in experiments which did not address selectivity, a ΔCA-CTD Gag was still able to tether genomic RNA to the plasma membrane in cells (Jouvenet et al. 2009, Sardo et al. 2015), which is consistent with our finding that the mutant Gag proteins can still recruit 5'UTR although being less selective in the

presence of competitor. Mutations at the 3-fold symmetry contact in CA-NTD did not alter Gag clustering on membranes (Figure 3g), concurring with observations that a Δ CA-NTD Gag still supports particle release (Borsetti et al. 1998, Accola et al. 2000).

Live-cell fluorescence microscopy has delivered important constraints on the stoichiometry and location of the RNA selection event. Genomic RNA is present at the onset of HIV-1 assembly at the plasma membrane, at which point the number of Gag molecules was reported to be below the detection limit of ~ 12 molecules (Jouvenet et al. 2009). Cytosolic Gag does not form large oligomers (Robinson et al. 2014, Hendrix et al. 2015) and mainly associates with cellular RNAs (Kutluay et al. 2014). A small fraction of cytosolic Gag is found to be slowly diffusing and consist of oligomers of up to ~ 5 molecules (Hendrix et al. 2015). In a recent study, TIRF microscopy was used to study the monomer-dimer state of genomic HIV-1 RNAs (Chen et al. 2015). It was reported that the large majority of plasma membrane-proximal genomic HIV-1 RNAs are monomeric in the absence of HIV-1 Gag. In cells expressing Gag, monomeric genomic RNAs are seen to coalesce to dimers on the membrane, but were never observed to arrive to the membrane as dimers. That study and our results underscore the importance of membrane-localized Gag assembly events in HIV-1 genome selection and packaging.

The SHAPE analysis focused on the secondary structure of the 5'UTR RNA in situations representing different stages of packaging (Figure 5a,b). When comparing SHAPE reactivities of isolated 5'UTR, 5'UTR interacting with NC domain in solution, Δ MA-Gag in solution, or recruited to GUUV membranes by clusters of full-length Gag, the changes in SHAPE reactivities were consistent with a conservation of secondary structure between all these states. Based on these results, there is no gross unfolding of the dimeric 5'UTR in the early steps of genomic RNA selection and packaging for HIV-1. This contrasts with reports that the packaging signal of the related retrovirus Murine Leukemia Virus (MLV) is partially unfolded in released immature virions (Grohman et al. 2014). The SHAPE reactivity pattern indicated that the 5'UTR was, as expected, dimeric in solution. This was observed both with the shorter 345 nt 5'UTR, which has been reported to be locked in a dimer-favoring conformation, and the longer 356 nt 5'UTR which is also capable of assuming a second conformation (Lu et al. 2011, Heng et al. 2012) (Figure 5 - figure supplement 1). In the light of a recent report that most genomic HIV-1 RNAs in cells are monomeric until interacting with Gag on a membrane (Chen et al. 2015), the isolated 5'UTR would thus be representative of the genome after engagement with Gag. It is possible that the RNA chaperone activity of Gag may be necessary to expose the 5'UTR and its DIS in the context of a full-length genome. The strongest decrease in SHAPE reactivity upon Gag interaction was found in the guanosine-rich bulges on the DIS stem (Figure 5c), consistent with a recent study using full-length Gag in solution (Abd El-Wahab et al. 2014). The increased SHAPE reactivity in central regions of the tandem three-way junction suggests that Gag engagement, while not leading to a drastic rearrangement of the 5'UTR secondary structure, may lead to alterations its 3D structure. Overall, our results show that RNA conformational changes upon binding to membrane-associated Gag are minor in magnitude and thus suggest that conformation rearrangements are likely to be a minor contributor to selective packaging.

Cytosolic RNA has been reported to regulate Gag membrane association, effectively increasing its selectivity for the plasma membrane lipid PI(4,5)P₂ (Alfadhli et al. 2009, Chukkapalli et al. 2010, Alfadhli et al. 2011, Chukkapalli et al. 2013, Dick et al. 2013, Dick et al. 2014). The RNA species that binds

the Gag MA domain in cells was identified by deep sequencing as primarily consisting of tRNA (Kutluay et al. 2014). Our reconstituted system allowed us to disentangle the effect that different RNAs have on Gag membrane association (Figure 6). We show that Gag-tRNA interactions may also serve to increase Gag fidelity for viral RNA. In the absence of tRNA, any long RNA caused a slight increase in Gag membrane association. But in the presence of 20 μ M tRNA, 5'UTR and a longer 1947 nt fragment of the HIV-1 genome increased Gag membrane association more than other RNAs. We hypothesize that a viral RNA with several high-affinity binding sites for Gag can recruit several Gag molecules to the membrane if only one of them dissociates from its membrane-occluding tRNA.

In summary, our study leads to a model for HIV-1 genome selection where the NC domain of Gag is necessary but not sufficient for the selective packaging of the HIV-1 genomic RNA (Figure 7). Lattice contacts conferring Gag ability to multimerize on a membrane are crucial for the selective packaging of viral RNA. The 5'UTR is recruited to membranes by multimerized Gag without any gross changes to its secondary structure. And the association of Gag with tRNA serves to inhibit genome-less virus assembly.

Acknowledgements

We thank Michael F. Summers (University of Maryland, Baltimore County) and Johannes Schöneberg (UC Berkeley) for fruitful discussions and A. Yeremenko for drawing Figure 7. This work was supported by the National Institutes of Health under award numbers R01AI112442 (J. H. H), P50GM082250 (J. A. D.) and P50GM103297 (M. F. S.).

323

324 **Materials and methods**

325 **Protein expression and purification**

326 Full-length, N-terminally myristoylated HIV-1 Gag was purified as previously described (Carlson et al.
327 2012). Gag was labelled with ATTO488-maleimide or ATTO594-maleimide (Sigma-Aldrich) on an
328 engineered cysteine (A120C) in the MA-CA linker region. The resulting fluorophore-labeled protein
329 eluted as a monodisperse peak from the final size exclusion column. Δ NC Gag was cloned from the
330 expression plasmid for full-length Gag by deleting the sequence corresponding to the two Zn fingers of
331 NC (residues 393-448 of full-length Gag, with the initiator Met being residue 1), and expressed and
332 purified analogously to full-length Gag. The nucleocapsid (NC) domain of Gag (residues 378-432 of full-
333 length Gag) was expressed in *E. coli* as a C-terminal, TEV-protease cleavable, maltose-binding
334 protein(MBP)-(His)₆ fusion protein and purified analogously to full-length Gag, using NiNTA- and Heparin
335 affinity columns followed by TEV protease digestion and size exclusion chromatography on a Superdex
336 75 16/60 column. Δ MA Gag (residues 120-500) was based on the same synthetic gene as used for full-
337 length Gag, expressed in *E. coli* as a TEV-cleavable (His)₆-MBP- Δ MA Gag-MBP construct, and purified as
338 full-length Gag.

339 **RNA generation and fluorophore labeling**

340 Yeast tRNA was purchased from Invitrogen. All other RNAs were generated by in vitro transcription
341 using T7 polymerase, and purified either by sequential LiCl and ethanol precipitation, or polyacrylamide
342 gel electrophoresis. The 5'UTR and RRE RNAs consisted of nucleotides 1-345 and 7259-7612 of the NL4-3
343 (GenBank: AF324493.2) and ARV-2/SF2 (GenBank: K02007.1) unspliced RNAs, respectively (Heng et al.
344 2012, Bai et al. 2014). RNA378 and RNA2266 were transcribed from the first 378 nucleotides of a
345 synthetic gene encoding human VPS25 (Im et al. 2008), and the first 2266 nucleotides of the gene
346 encoding *H. sapiens* AIP1/ALIX (GenBank: NM_013374.5), respectively. The integrity of the RNAs was
347 monitored by denaturing formaldehyde-agarose gel electrophoresis. Fluorophore labeling of RNA was
348 carried out using ULYSIS Alexa Fluor 488 (Molecular Probes) at 90°C for 10 minutes in the presence of 3
349 mM MgCl₂ followed by rapid cooling to 4°C to refold the labeled RNAs. Labeled RNAs were buffer-
350 exchanged using Bio-Spin 6 spin columns (BIO-RAD), and the final degree of labeling was measured using
351 UV-VIS spectrometry.

352 **Vesicle preparations**

353 Lipids were acquired from Avanti polar lipids, except the fluorophore ATTO647-DOPE which was
354 acquired from ATTO-TEC GmbH. GUVs with a molar composition of 69.9% palmitoyl-oleoyl-
355 phosphatidylcholine (POPC), 25% cholesterol, 5% brain-extracted phosphatidylinositol-4,5-bisphosphate
356 (PI(4,5)P₂) and 0.1% ATTO647-dioleoyl-phosphatidylethanolamine were electroformed as described
357 previously (Carlson et al. 2012). For LUVs, lipids with a molar composition of 70% POPC, 25% cholesterol,
358 and 5% PI(4,5)P₂ were dried over night in a desiccator and resuspended in LUV buffer (20 mM Tris, 150

mM NaCl, pH 7.4). They were subjected to ten freeze-thaw cycles on liquid nitrogen, followed by eleven times extrusion through a Nuclepore Track-Etched Membrane with 1.0 μ m pore size (Whatman).

Reconstitution on GUVs and confocal microscopy

GUVs electroformed in 600 mM sucrose were stored at room temperature and imaged within 24 hours. Protein and RNAs were preincubated for 5 minutes at room temperature in GUV dilution buffer (20 mM Tris, 300 mM NaCl, 4 mM $MgCl_2$, pH=7.4) and then diluted 1:1 with GUVs in a Lab-Tek II chambered coverglass (Fisher Scientific). Ten minutes after mixing, the reaction was imaged using an LSM 5 LIVE confocal microscope (Carl Zeiss, Jena). The membrane fluorophore was imaged using the 635 nm laser and a 650 nm long-pass emission filter, Gag using the 532 nm laser and a 550-615 nm emission filter and RNA using the 488 nm laser and a 500-525 nm emission filter. Membrane, protein, and RNA were imaged using separate tracks, and cross-excitation was found to be minimal with the chosen imaging parameters. On each reaction, ten z-stacks (each consisting of ten slices with 1.0 μ m spacing) were acquired at positions selected solely on the basis of containing GUVs, without observing the fluorescence channels prior to acquisition. All experiments shown in the same figure were done within 24h using the same GUV batch for comparability. Each such experiment series was repeated on three separate occasions with different GUV preparations.

AT-2 treatment of Gag

To Gag-ATTO594, AT-2 (aldrithiol-2, Sigma-Aldrich) dissolved in DMSO was added to a final concentration of 100 μ M, followed by a 15 min incubation at 37°C prior to imaging.

Image analysis

Statistics on RNA recruitment to Gag clusters was performed as described previously (Carlson et al. 2012). Briefly, a custom-written script for MATLAB (Mathworks) used the confocal z-stacks to create a binary mask excluding non-membrane voxels, then identified continuous areas with high Gag fluorescence within the membrane mask. Such an area (Gag cluster) was considered positive for RNA recruitment if its average RNA fluorescence was > 2.0 times higher than the average in the GUV membrane outside Gag puncta, except for Figure 2 where a lower threshold of 1.25 was used. To quantitate the Gag binding type (Figure 2g-h), all GUVs with a diameter > 5 μ m were counted. The GUVs were inspected for Gag binding in the 532 nm fluorescence channel, and those GUVs with Gag clusters but areas completely devoid of detectable Gag fluorescence were counted as “clustered only” The remaining GUVs either had no detectable Gag binding, or had an even fluorescence covering the entire membrane, sometimes with additional brighter areas. These GUVs were counted as “diffuse binding”.

SHAPE analysis of the interactions between the HIV-1 5'UTR and the Gag proteins

For SHAPE analysis, SHAPE handles were added to both ends of the 345-nt and 356-nt 5'UTR construct as previously reported (Bai et al. 2014). Purified RNA samples were annealed at 0.1 mg/ml in a buffer containing 50 mM HEPES-KOH pH 7.5, 200 mM KOAc, and 3 mM $MgCl_2$ by heating at 75 °C for 2 min and snap cooling on ice. Before SHAPE reactions, 9 μ l of annealed 5'UTR RNA at 0.1 mg/ml was mixed with 1

μl of variants of Gag protein to achieve a final RNA:protein stoichiometry of ~1:8 (0.8 μM:7.2 μM). Matching protein buffer containing 20 mM Tris, 150 mM NaCl, 0.1 mM TCEP, pH=7.4 was used as a control. The resulting mixtures were incubated for 15 min at room temperature. SHAPE probing was performed as previously reported (Berry et al. 2011) with 1-methyl-7-mitroisatoic anhydride (1M7) as the 2' hydroxyl-selective electrophile. For SHAPE analysis of GUV-bound RNA, reactions were set up in Lab-Tek II chambered coverglasses as for confocal imaging to ensure identical mixing kinetics and GUV settling as during imaging. Gag and 5'UTR were premixed with GUV dilution buffer to 150 μl total volume and after 5 minutes at room temperature were added to 150 μl GUVs in imaging chambers for final concentrations of 100 nM Gag and 5 nM 5'UTR. 10 minutes after mixing, 200 μl supernatant was gently aspirated. 15 μl 1M7 in DMSO was briefly premixed with 35 μl GUV dilution buffer to avoid exposing GUVs to undiluted DMSO, added to the remaining 100 μl GUV solution in the chamber, and mixed by gentle stirring. Eight chambers were used for each reaction. The control reaction was performed by adding DMSO only. For the test of GUV integrity after this treatment (Figure 5 - figure supplement 2), the control reaction was repeated with fluorescent Gag-ATTO594 and imaged by confocal microscopy as described above 5 minutes after DMSO addition. To measure the amount of free RNA in the GUV reactions, the supernatant which was removed prior to SHAPE reagent addition was centrifuged for 5 minutes in a tabletop centrifuge at 21,000*g to pellet any remaining vesicle-bound RNA. The RNA concentration in the supernatant of this sample was measured using a Qubit fluorimeter (ThermoFisher Scientific). Raw traces from fragment analysis was analyzed using ShapeFinder (Vasa et al. 2008).

Vesicle flotation assay

Gag and RNA were mixed in a total volume of 60 μl LUV buffer containing 1.67 mM MgCl₂ and incubated for 5 minutes at room temperature. 40 μl 1.0 mg/ml freshly prepared LUVs were added and the samples were incubated for 10 minutes at 37 °C. The reactions were mixed with OptiPrep solution (Sigma-Aldrich) and LUV buffer to 700 μl and 30 % (w/v) iodixanol. This solution was placed at the bottom of ultracentrifuge tubes and overlaid with 900 μl 18 % (w/v) iodixanol in LUV buffer and 700 μl LUV buffer. The gradients were subjected to ultracentrifugation in an Sw55.Ti rotor for 3h at 45,000 rpm at 4°C. At the end of the run the vesicles fractions were collected from the interface between the 0% and 18% iodixanol fraction, and the amount of Gag associated with the vesicles was measured in a GloMax-Multi microplate reader (Promega) using 490 nm excitation light and a 510-570 nm emission filter. All experiments shown in the same panel were conducted with the same preparation of LUVs and each experiment series was repeated on three separate occasions with different LUV preparations

Figure legends

Figure 1. In vitro reconstitution of selective RNA packaging by HIV-1 Gag. (a) Confocal fluorescence micrograph of GUVs containing 5% PI(4,5)P₂. 100 nM HIV-1 Gag-ATTO594 and 0.5 nM HIV-1 5'UTR-Alexa488 were premixed and added to the exterior of the GUVs which were imaged 10 min later. Upper panel, membrane in red and Gag in white/cyan. Lower panel, RNA. (b-c) As (a) but with fluorescent HIV-

1 RRE and a 378 nt control ssRNA (RNA378), respectively. (d-f) As (a-c) but with 5 nM non-fluorescent RNA378 added together with the fluorescent RNAs. (a-f) White arrows mark Gag clusters on GUVs, and their position on the corresponding RNA images. Scale bar for all images, 10 μ m. (g) Quantitation of fluorescent RNA binding to Gag clusters. Gag clusters on membranes were identified in 10 z-stack each containing ~10 GUVs, using an unsupervised script which calculated the average RNA fluorescence within the clusters. Gag clusters with an average RNA fluorescence > 2.0 times that of the surrounding membrane were counted as positive for colocalization. All measurements were conducted on the same batch of GUVs and error bars indicate standard deviation between three repeats on separate GUV preparations. n.s./*, not significant, and significant, respectively, at $p < 0.05$ level by Student's t-test.

Figure 2. Effect of 5'UTR alterations on packaging by HIV-1 Gag. (a) Schematic of the HIV-1 5'UTR indicating introduced mutations. Blue indicates the position of the GCGCGC to GAGA mutation in the dimerization motif creating the monomerized construct "5'UTR mono". Red indicates the position of the splice donor site G290. Replacement of nucleotides after G290 with the sequence from a spliced RNA created the construct "5'UTR env-1". Cyan indicates the position of packaging-critical unpaired and weakly paired guanines mutated to adenines in the construct "5'UTR 3way1". PBS, primer binding site. DIS, dimerization signal. SD, splice donor site. ψ , psi stem loop. The schematic shows one monomer of the dimerizing 5'UTR. (b-f) Confocal micrographs of GUVs, 10 minutes after adding 100 nM Gag-ATTO594, 0.5 nM 5'UTR -Alexa488, and 5 nM non-fluorescent competitor RNA as indicated. Upper panels, membrane in red and Gag in white/cyan. Lower panel, RNA. White arrows mark Gag clusters on GUVs, and their position on the corresponding RNA images. Scale bar for all images, 10 μ m. (g) Quantitation of fluorescent RNA binding to Gag clusters, performed as for Figure 1g but using a lower threshold of 1.25 for counting Gag clusters as positive for fluorescent RNA. n.s./*, not significant, and significant, respectively, at $p < 0.05$ level by Student's t-test.

Figure 3. CA domain mutants affect Gag clustering on membranes. (a) Schematic of a hexagonal lattice showing the Gag CA domain at the two-, three- and sixfold symmetry contacts, as arranged in the immature Gag lattice in HIV-1 assembly sites and immature virions (Schur et al. 2015) (PDB entry 4USN used in panel a-d). (b) Model of the CA domain at the twofold, with mutated residues in "Gag 2-fold" marked in red. (c-d) As (b), for Gag mutated at the 3- and 6-fold symmetry contacts, respectively. (e) Confocal micrograph of GUVs, 10 minutes after adding 100 nM Gag-ATTO594. Membrane in red and Gag in white/cyan. (f) As (e), using Gag-ATTO594 2-fold. GUVs with diffuse Gag binding are denoted with white asterisks. (e-f) White arrows mark Gag clusters on GUVs. Scale bar, 10 μ m. (g) Type of Gag binding to GUVs. For each protein, GUVs were counted in 10 confocal z-stacks, and classified according to having no Gag fluorescence, only clustered Gag fluorescence (black), or a diffuse Gag fluorescence covering the entire membrane (with or without additional brighter clusters, grey). All measurements were conducted on the same preparation of GUVs, with error bars indicating the standard deviation between three repeats conducted on separate GUV preparations. *, significant at $p < 0.05$ level by Student's t-test.

Figure 4. RNA selectivity of CA domain mutants. (a) Confocal micrograph of GUV, 10 minutes after adding 100 nM Gag-ATTO594 and 0.5 nM 5'UTR -Alexa488. Upper panel, membrane in red and Gag in white/cyan. Lower panel, RNA. (b-e) As (a) using Gag 2-fold, Gag 3-fold, Gag 6-fold and Gag 2-,3-, 6-fold, respectively. (f-j) As (a-j) with the addition of 5 nM non-fluorescent RNA378. (a-j) White arrows mark

473 Gag clusters on GUVs, and their position on the corresponding RNA images. Scale bar for all images, 10
 474 μm . (k) Quantitation of RNA binding to Gag clusters, performed as for Figure 1g. n.s./*, not significant,
 475 and significant, respectively, at $p < 0.05$ level by Student's t-test.

476 **Figure 5. Probing Gag induced structural changes on the HIV 5'UTR using SHAPE.** (a) Secondary
 477 structure of the HIV 5'UTR RNA with SHAPE reactivities labeled. The secondary structure is based on the
 478 reported NMR structure (Keane et al. 2015) (PDB 2N1Q). Nucleotides exhibiting high, medium, low, or
 479 no SHAPE reactivities are labeled in red, orange, cyan, and black, respectively. SHAPE handles are
 480 labeled by grey shadows. The 3D structure colored based on the SHAPE profile using the same color
 481 scheme is shown as insert. U5, U5 region of long terminal repeat. PBS, primer binding site. DIS,
 482 dimerization signal. SD, splice donor site. ψ , psi stem loop. AUG, start codon of gag gene. The
 483 nucleotides of SD that would be in a loop in the alternative conformation (289-292) are underlined with
 484 a grey box. The schematic shows one monomer of the dimerizing 5'UTR. (b) SHAPE profiles of the 5'UTR
 485 either alone (in grey) or in complex with the following Gag variants: the NC domain of Gag (in cyan),
 486 ΔMA -Gag (in blue), and Gag assembled on GUV (in red). The regions not showing significant SHAPE
 487 changes are masked. (c) SHAPE changes between the 5'UTR alone and the 5'UTR bound by the Gag on
 488 GUV. Nucleotides exhibiting reduced or increased SHAPE value upon complex formation are labeled in
 489 cyan and dark red, respectively. Corresponding nucleotide are also labeled on the NMR 3D structure
 490 model using the same color scheme.

491 **Figure 6. Effect of tRNA and viral RNA on Gag membrane association.** (a) Confocal micrograph of GUVs,
 492 10 minutes after adding 100 nM Gag-ATTO594. (b) Confocal micrograph of GUVs, 10 minutes after
 493 adding 100 nM Gag-ATTO594, premixed with 1 μM tRNA. (a-b) Membrane in red and Gag in white/cyan.
 494 White arrows mark Gag clusters on GUVs. Scale bar, 10 μm . (c) Amount of Gag fluorescence in the
 495 vesicle fraction after a vesicle flotation assay with 100 nM Gag-ATTO488, tRNA at 0-20 μM , and 0.4
 496 mg/mg 1.0 μm LUVs. (d) Amount of Gag fluorescence in the vesicle fraction after a vesicle flotation assay
 497 with 100 nM Gag-ATTO488 and RNAs at 5 nM. (e) Amount of Gag fluorescence in the vesicle fraction
 498 after a vesicle flotation assay with 100 nM Gag-ATTO488, 20 μM tRNA, and other RNAs at 5 nM. (c-e) All
 499 measurements shown in the same panel were conducted on the same LUV preparation, with error bars
 500 indicating the standard deviation between three repeats conducted on separate LUV preparations.
 501 n.s./*, not significant, and significant, respectively, at $p < 0.05$ level by Student's t-test.

502 **Figure 7. A model for selective HIV-1 genome packaging.** Cytosolic Gag (light grey triangles) is inhibited
 503 from membrane association by its association with tRNA (yellow). Gag (triangles and hexagons, colored
 504 dark grey where forming specific contacts with the packaging signal) binding viral genomic RNA (red;
 505 thick portions correspond to the 5'UTR, thin portions to the remainder, not shown to scale) has a higher
 506 probability to overcome this inhibition and associate with the plasma membrane. The selectivity of
 507 membrane-bound Gag for the viral genome necessitates CA-CA immature lattice contacts.

508 **Figure 1 - figure supplement 1. RNA recruitment by clusters of Gag with altered NC domain.** (a)
 509 Confocal micrograph of GUV, 10 minutes after adding 100 nM ΔNC Gag-ATTO594 and 0.5 nM 5'UTR -
 510 Alexa488. Upper panel, membrane in red and Gag in white/cyan. Lower panel, RNA. (b) As (a), but with
 511 full-length Gag treated for 15 min at 37°C with 100 μM of the Zn-finger disrupting compound AT-2. (c) As

(b), but Gag treated with DMSO only as negative control for (b). (a-c) White arrows mark Gag clusters on GUVs, and their position on the corresponding RNA images. Scale bar for all images, 10 μ m.

Figure 4 - figure supplement 1. Type of Gag binding to GUVs in the presence of RNA. Statistics of Gag membrane binding was performed as for Figure 3g, using the data presented in Figure 4f-j (ten z-stacks per condition, in three independent repeats).

Figure 5 - figure supplement 1. SHAPE changes for protein addition in solution and a longer 5'UTR RNA. a. Changes in SHAPE reactivity are plotted as in figure 4c, for addition of Δ MA-Gag and NC, respectively, and between a longer 356 nt 5'UTR and the 345 nt 5'UTR used throughout the study. b. Secondary structure of the 356-nt HIV 5'UTR RNA with SHAPE reactivities labeled as in figure 4a.

Figure 5 - figure supplement 2. Gag binding to GUVs under conditions of the SHAPE analysis. Gag and GUVs were prepared as for SHAPE analysis. In the final step DMSO without SHAPE reagent was added and GUVs were imaged by confocal microscopy. A representative field of view is shown.

References

- Abd El-Wahab, E. W., R. P. Smyth, E. Mailler, S. Bernacchi, V. Vivet-Boudou, M. Hijnen, F. Jossinet, J. Mak, J. C. Paillart and R. Marquet (2014). "Specific recognition of the HIV-1 genomic RNA by the Gag precursor." *Nat Commun* **5**: 4304.
- Accola, M. A., B. Strack and H. G. Gottlinger (2000). "Efficient particle production by minimal Gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain." *J Virol* **74**(12): 5395-5402.
- Aldovini, A. and R. A. Young (1990). "Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus." *J Virol* **64**(5): 1920-1926.
- Alfadhli, A., H. McNett, S. Tsagli, H. P. Bachinger, D. H. Peyton and E. Barklis (2011). "HIV-1 matrix protein binding to RNA." *J Mol Biol* **410**(4): 653-666.
- Alfadhli, A., A. Still and E. Barklis (2009). "Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids." *J Virol* **83**(23): 12196-12203.
- Bai, Y., A. Tambe, K. Zhou and J. A. Doudna (2014). "RNA-guided assembly of Rev-RRE nuclear export complexes." *Elife* **3**: e03656.
- Berkowitz, R. D., J. Luban and S. P. Goff (1993). "Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays." *J Virol* **67**(12): 7190-7200.
- Berry, K. E., S. Waghay, S. A. Mortimer, Y. Bai and J. A. Doudna (2011). "Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning." *Structure* **19**(10): 1456-1466.
- Borsetti, A., A. Ohagen and H. G. Gottlinger (1998). "The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly." *J Virol* **72**(11): 9313-9317.
- Carlson, L. A., J. A. Briggs, B. Glass, J. D. Riches, M. N. Simon, M. C. Johnson, B. Müller, K. Grunewald and H. G. Kräusslich (2008). "Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis." *Cell Host Microbe* **4**(6): 592-599.

551 Carlson, L. A. and J. H. Hurley (2012). "In vitro reconstitution of the ordered assembly of the endosomal
 552 sorting complex required for transport at membrane-bound HIV-1 Gag clusters." Proc Natl Acad Sci U S A
 553 **109**(42): 16928-16933.
 554 Chen, J., O. Nikolaitchik, J. Singh, A. Wright, C. E. Bencsics, J. M. Coffin, N. Ni, S. Lockett, V. K. Pathak and
 555 W. S. Hu (2009). "High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed
 556 by single virion analysis." Proc Natl Acad Sci U S A **106**(32): 13535-13540.
 557 Chen, J., S. A. Rahman, O. A. Nikolaitchik, D. Grunwald, L. Sardo, R. C. Burdick, S. Plisov, E. Liang, S. Tai, V.
 558 K. Pathak and W. S. Hu (2015). "HIV-1 RNA genome dimerizes on the plasma membrane in the presence
 559 of Gag protein." Proc Natl Acad Sci U S A.
 560 Chukkapalli, V., J. Inlora, G. C. Todd and A. Ono (2013). "Evidence in support of RNA-mediated inhibition
 561 of phosphatidylserine-dependent HIV-1 Gag membrane binding in cells." J Virol **87**(12): 7155-7159.
 562 Chukkapalli, V., S. J. Oh and A. Ono (2010). "Opposing mechanisms involving RNA and lipids regulate
 563 HIV-1 Gag membrane binding through the highly basic region of the matrix domain." Proc Natl Acad Sci
 564 U S A **107**(4): 1600-1605.
 565 Coffin, J. M., S. H. Hughes and H. E. Varmus (1997). Retroviruses, Cold Spring Harbor Laboratory Press.
 566 De Guzman, R. N., Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer and M. F. Summers (1998).
 567 "Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element." Science
 568 **279**(5349): 384-388.
 569 Dick, R. A., E. Kamynina and V. M. Vogt (2013). "Effect of multimerization on membrane association of
 570 Rous sarcoma virus and HIV-1 matrix domain proteins." J Virol **87**(24): 13598-13608.
 571 Dick, R. A. and V. M. Vogt (2014). "Membrane interaction of retroviral Gag proteins." Front Microbiol **5**:
 572 187.
 573 Ennifar, E., P. Walter, B. Ehresmann, C. Ehresmann and P. Dumas (2001). "Crystal structures of coaxially
 574 stacked kissing complexes of the HIV-1 RNA dimerization initiation site." Nat Struct Biol **8**(12): 1064-
 575 1068.
 576 Freed, E. O. (2015). "HIV-1 assembly, release and maturation." Nat Rev Microbiol **13**(8): 484-496.
 577 Gamble, T. R., S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W.
 578 I. Sundquist and C. P. Hill (1997). "Structure of the carboxyl-terminal dimerization domain of the HIV-1
 579 capsid protein." Science **278**(5339): 849-853.
 580 Ganser-Pornillos, B. K., M. Yeager and O. Pornillos (2012). "Assembly and architecture of HIV." Adv Exp
 581 Med Biol **726**: 441-465.
 582 Gorelick, R. J., S. M. Nigida, Jr., J. W. Bess, Jr., L. O. Arthur, L. E. Henderson and A. Rein (1990).
 583 "Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA." J Virol **64**(7):
 584 3207-3211.
 585 Grohman, J. K., R. J. Gorelick, S. Kottegoda, N. L. Allbritton, A. Rein and K. M. Weeks (2014). "An
 586 immature retroviral RNA genome resembles a kinetically trapped intermediate state." J Virol **88**(11):
 587 6061-6068.
 588 Hendrix, J., V. Baumgartel, W. Schrimpf, S. Ivanchenko, M. A. Digman, E. Gratton, H. G. Krausslich, B.
 589 Muller and D. C. Lamb (2015). "Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-
 590 interacting Gag oligomers." J Cell Biol **210**(4): 629-646.
 591 Heng, X., S. Kharytonchyk, E. L. Garcia, K. Lu, S. S. Divakaruni, C. LaCotti, K. Edme, A. Telesnitsky and M.
 592 F. Summers (2012). "Identification of a minimal region of the HIV-1 5'-leader required for RNA
 593 dimerization, NC binding, and packaging." J Mol Biol **417**(3): 224-239.
 594 Im, Y. J. and J. H. Hurley (2008). "Integrated structural model and membrane targeting mechanism of the
 595 human ESCRT-II complex." Dev Cell **14**(6): 902-913.
 596 Johnson, S. F. and A. Telesnitsky (2010). "Retroviral RNA dimerization and packaging: the what, how,
 597 when, where, and why." PLoS Pathog **6**(10): e1001007.

598 Jouvenet, N., S. M. Simon and P. D. Bieniasz (2009). "Imaging the interaction of HIV-1 genomes and Gag
 599 during assembly of individual viral particles." Proc Natl Acad Sci U S A **106**(45): 19114-19119.
 600 Keane, S. C., X. Heng, K. Lu, S. Kharytonchyk, V. Ramakrishnan, G. Carter, S. Barton, A. Hoscic, A. Florwick,
 601 J. Santos, N. C. Bolden, S. McCowin, D. A. Case, B. A. Johnson, M. Salemi, A. Telesnitsky and M. F.
 602 Summers (2015). "RNA structure. Structure of the HIV-1 RNA packaging signal." Science **348**(6237): 917-
 603 921.
 604 Kenyon, J. C., L. J. Prestwood and A. M. Lever (2015). "A novel combined RNA-protein interaction
 605 analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader." Sci
 606 Rep **5**: 14369.
 607 Kutluay, S. B. and P. D. Bieniasz (2010). "Analysis of the initiating events in HIV-1 particle assembly and
 608 genome packaging." PLoS Pathog **6**(11): e1001200.
 609 Kutluay, S. B., T. Zang, D. Blanco-Melo, C. Powell, D. Jannain, M. Errando and P. D. Bieniasz (2014).
 610 "Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis." Cell **159**(5): 1096-
 611 1109.
 612 Kuzembayeva, M., K. Dilley, L. Sardo and W. S. Hu (2014). "Life of psi: how full-length HIV-1 RNAs
 613 become packaged genomes in the viral particles." Virology **454-455**: 362-370.
 614 Lever, A., H. Gottlinger, W. Haseltine and J. Sodroski (1989). "Identification of a sequence required for
 615 efficient packaging of human immunodeficiency virus type 1 RNA into virions." J Virol **63**(9): 4085-4087.
 616 Lu, K., X. Heng, L. Garyu, S. Monti, E. L. Garcia, S. Kharytonchyk, B. Dorjsuren, G. Kulandaivel, S. Jones, A.
 617 Hiremath, S. S. Divakaruni, C. LaCotti, S. Barton, D. Tummlillo, A. Hoscic, K. Edme, S. Albrecht, A.
 618 Telesnitsky and M. F. Summers (2011). "NMR detection of structures in the HIV-1 5'-leader RNA that
 619 regulate genome packaging." Science **334**(6053): 242-245.
 620 Lu, K., X. Heng and M. F. Summers (2011). "Structural determinants and mechanism of HIV-1 genome
 621 packaging." J Mol Biol **410**(4): 609-633.
 622 Merino, E. J., K. A. Wilkinson, J. L. Coughlan and K. M. Weeks (2005). "RNA structure analysis at single
 623 nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE)." J Am Chem Soc
 624 **127**(12): 4223-4231.
 625 Nikolaitchik, O. A., K. A. Dilley, W. Fu, R. J. Gorelick, S. H. Tai, F. Soheilian, R. G. Ptak, K. Nagashima, V. K.
 626 Pathak and W. S. Hu (2013). "Dimeric RNA recognition regulates HIV-1 genome packaging." PLoS Pathog
 627 **9**(3): e1003249.
 628 Rein, A., S. A. Datta, C. P. Jones and K. Musier-Forsyth (2011). "Diverse interactions of retroviral Gag
 629 proteins with RNAs." Trends Biochem Sci **36**(7): 373-380.
 630 Robinson, B. A., J. C. Reed, C. D. Geary, J. V. Swain and J. R. Lingappa (2014). "A temporospatial map that
 631 defines specific steps at which critical surfaces in the Gag MA and CA domains act during immature HIV-
 632 1 capsid assembly in cells." J Virol **88**(10): 5718-5741.
 633 Rossio, J. L., M. T. Esser, K. Suryanarayana, D. K. Schneider, J. W. Bess, G. M. Vasquez, T. A. Wiltout, E.
 634 Chertova, M. K. Grimes, Q. Sattentau, L. O. Arthur, L. E. Henderson and J. D. Lifson (1998). "Inactivation
 635 of human immunodeficiency virus type 1 infectivity with preservation of conformational and functional
 636 integrity of virion surface proteins." Journal of Virology **72**(10): 7992-8001.
 637 Sakuragi, J., S. Ueda, A. Iwamoto and T. Shioda (2003). "Possible role of dimerization in human
 638 immunodeficiency virus type 1 genome RNA packaging." J Virol **77**(7): 4060-4069.
 639 Sardo, L., S. C. Hatch, J. Chen, O. Nikolaitchik, R. C. Burdick, D. Chen, C. J. Westlake, S. Lockett, V. K.
 640 Pathak and W. S. Hu (2015). "Dynamics of HIV-1 RNA Near the Plasma Membrane during Virus
 641 Assembly." J Virol **89**(21): 10832-10840.
 642 Schur, F. K., W. J. Hagen, M. Rumlova, T. Ruml, B. Müller, H. G. Kräusslich and J. A. Briggs (2015).
 643 "Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution." Nature **517**(7535):
 644 505-508.

645 Stoltzfus, C. M. (2009). "Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus
646 replication." Adv Virus Res **74**: 1-40.

647 Sundquist, W. I. and H. G. Krausslich (2012). "HIV-1 Assembly, Budding, and Maturation." Cold Spring
648 Harb Perspect Med **2**(7): a006924.

649 Vasa, S. M., N. Guex, K. A. Wilkinson, K. M. Weeks and M. C. Giddings (2008). "ShapeFinder: a software
650 system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by
651 capillary electrophoresis." RNA **14**(10): 1979-1990.

652 von Schwedler, U. K., K. M. Stray, J. E. Garrus and W. I. Sundquist (2003). "Functional surfaces of the
653 human immunodeficiency virus type 1 capsid protein." Journal of Virology **77**(9): 5439-5450.

654 Wilkinson, K. A., R. J. Gorelick, S. M. Vasa, N. Guex, A. Rein, D. H. Mathews, M. C. Giddings and K. M.
655 Weeks (2008). "High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly
656 conserved across distinct biological states." PLoS Biol **6**(4): e96.

657

Figure 1

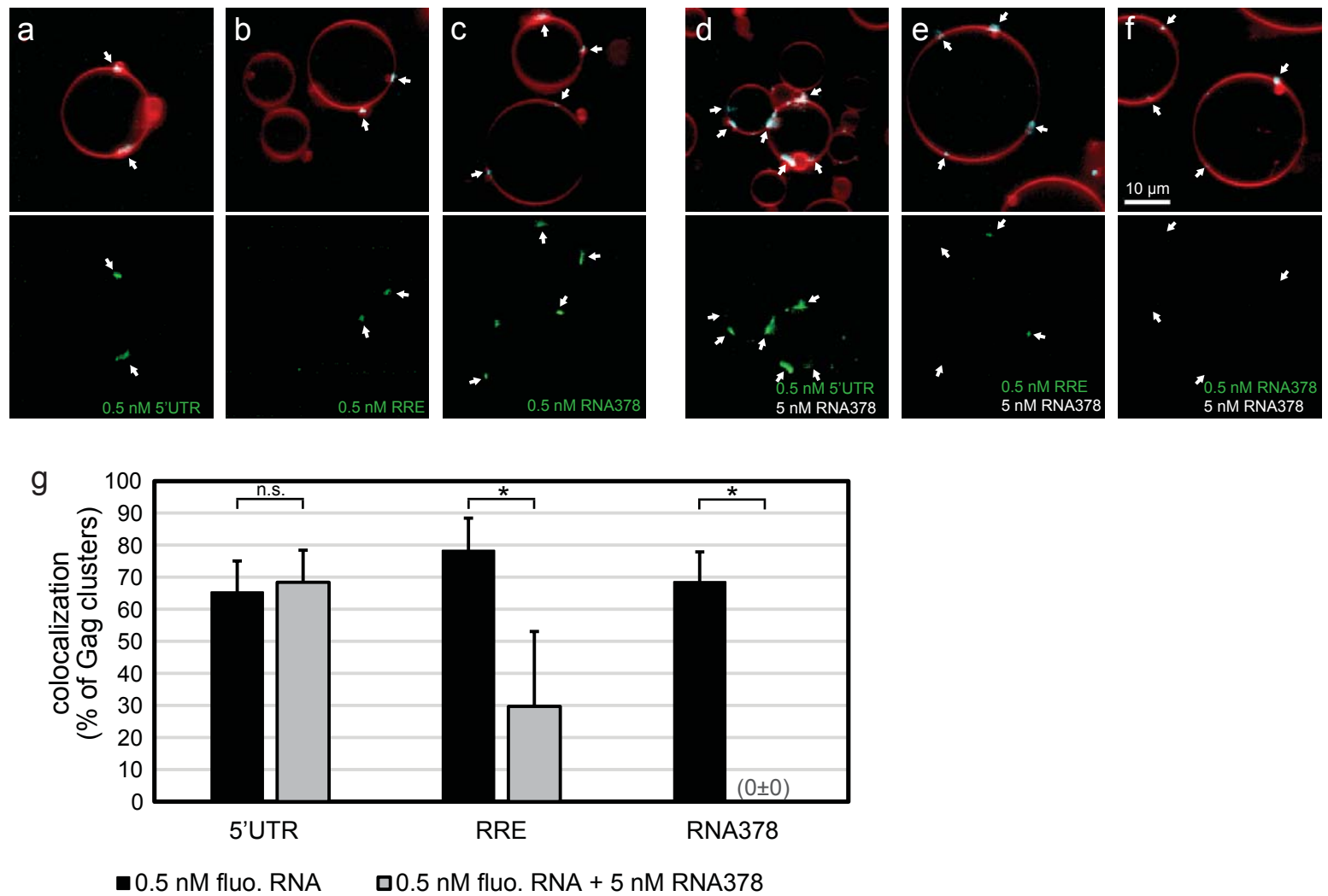


Figure 2

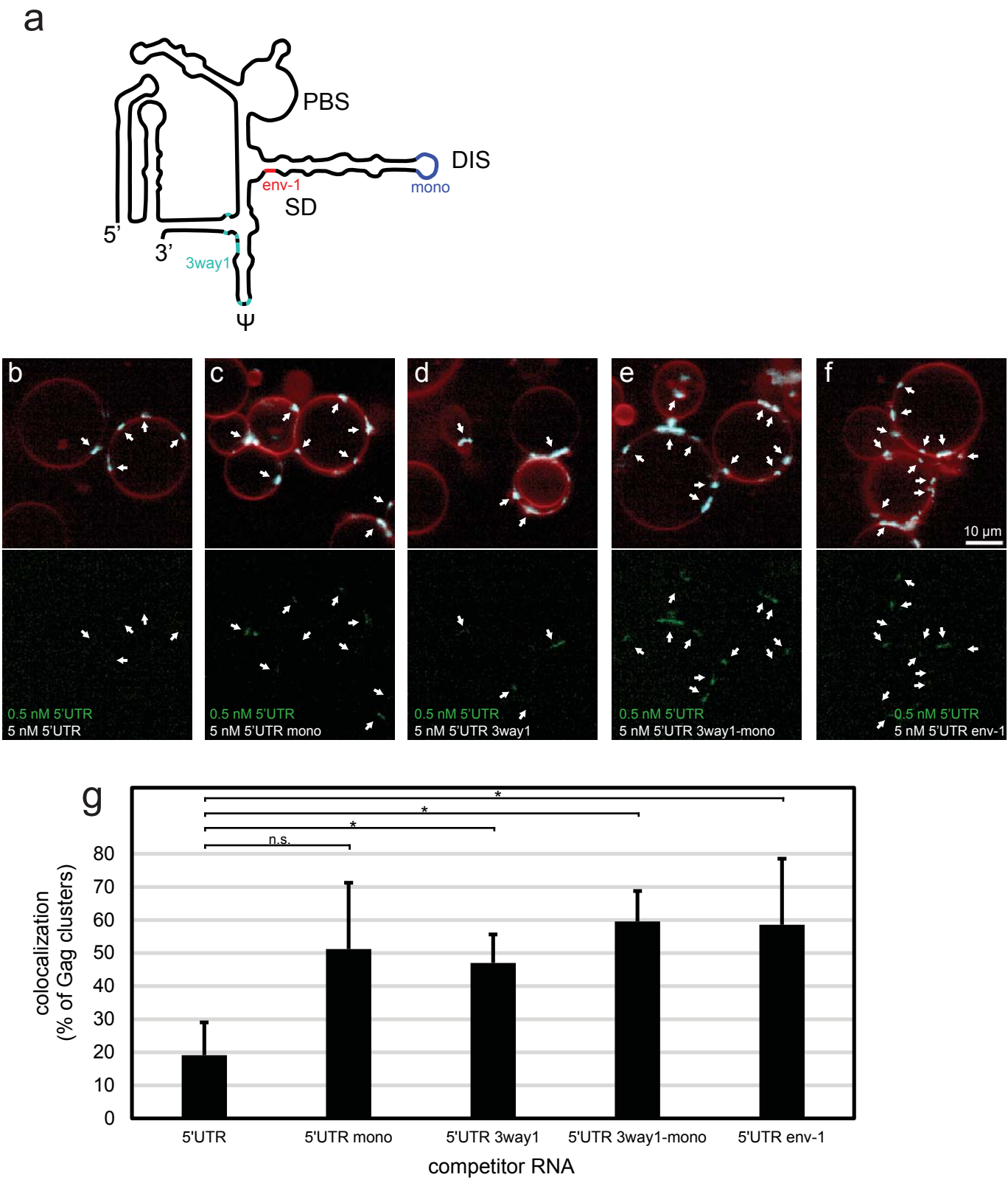


Figure 3

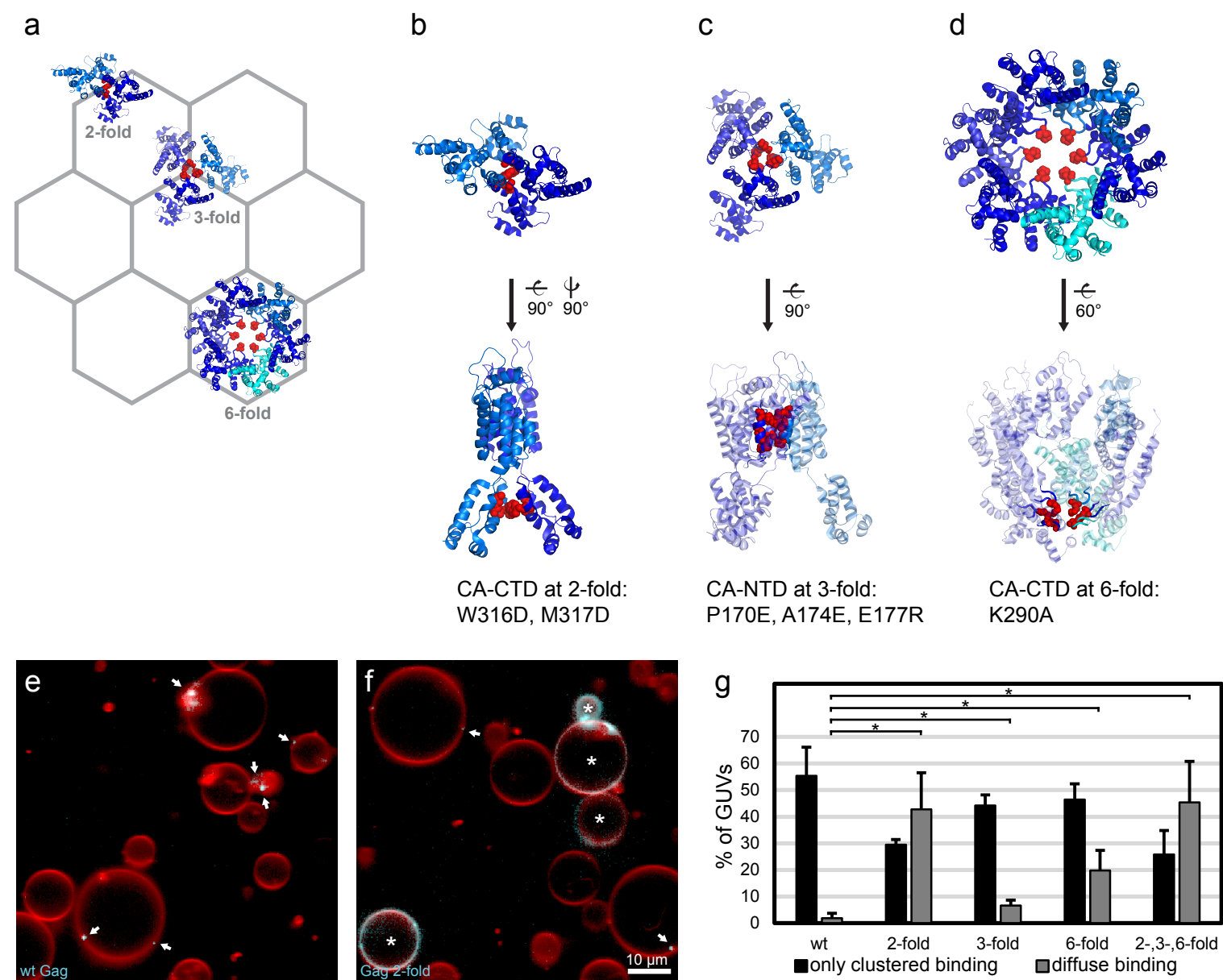
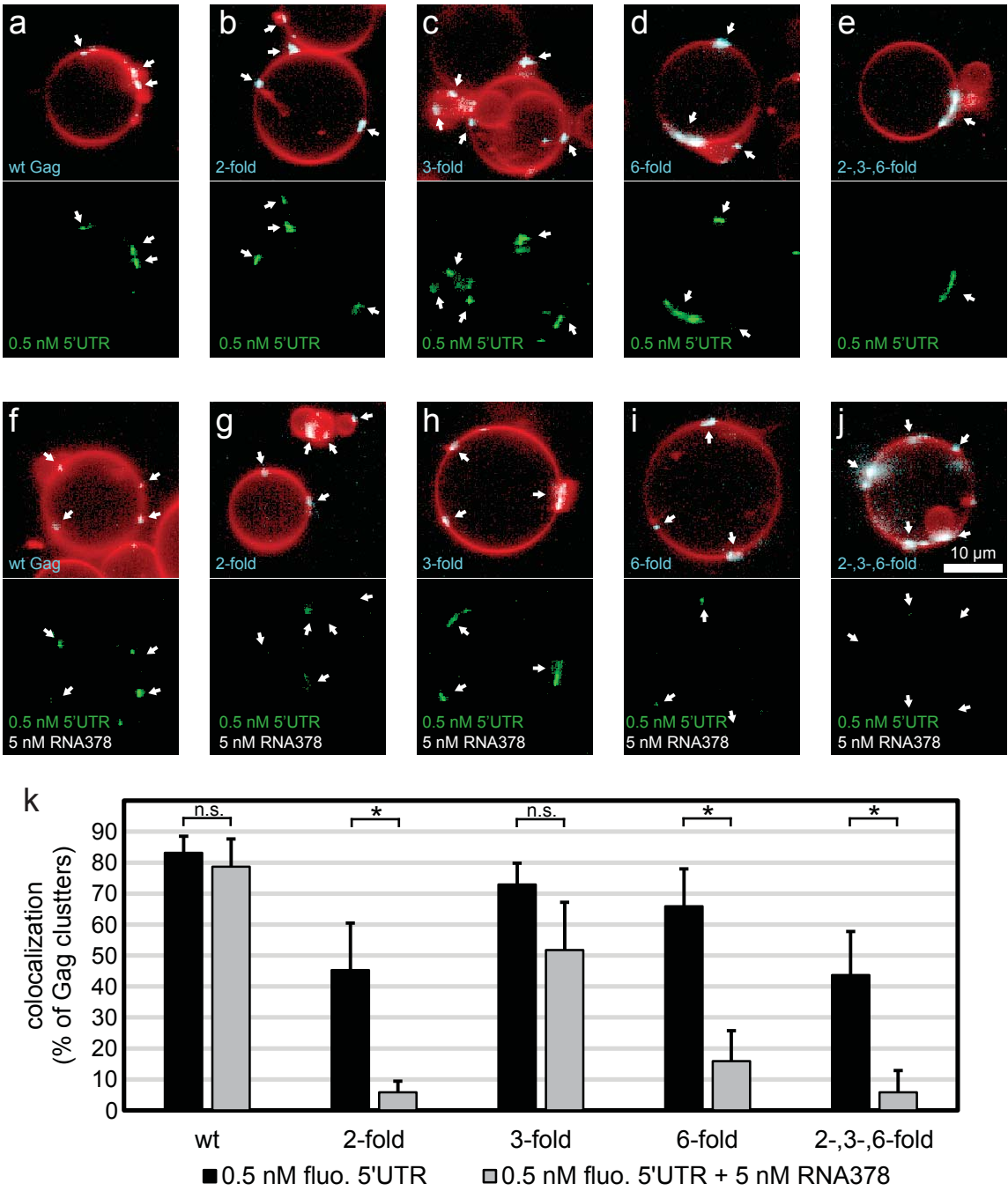
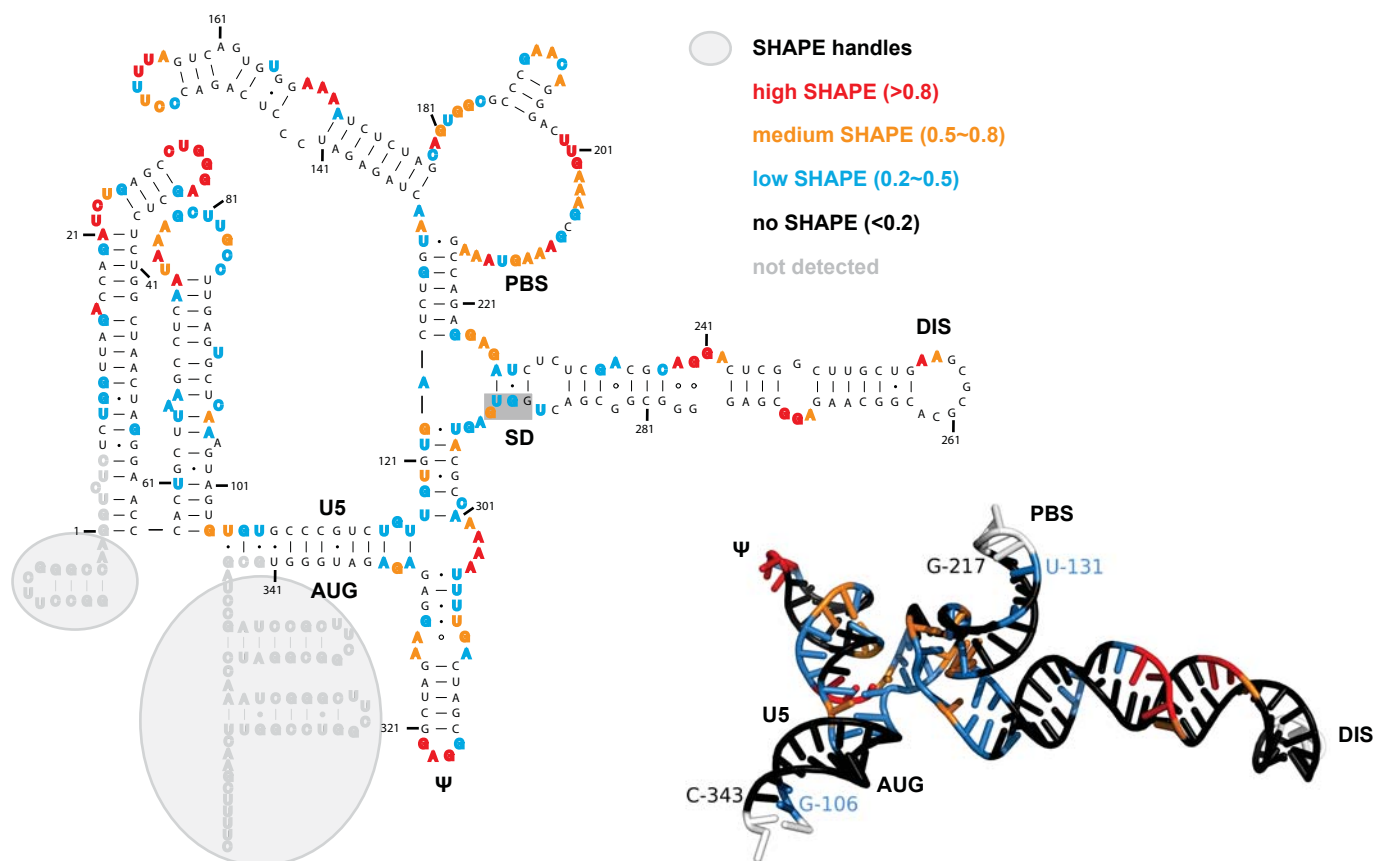


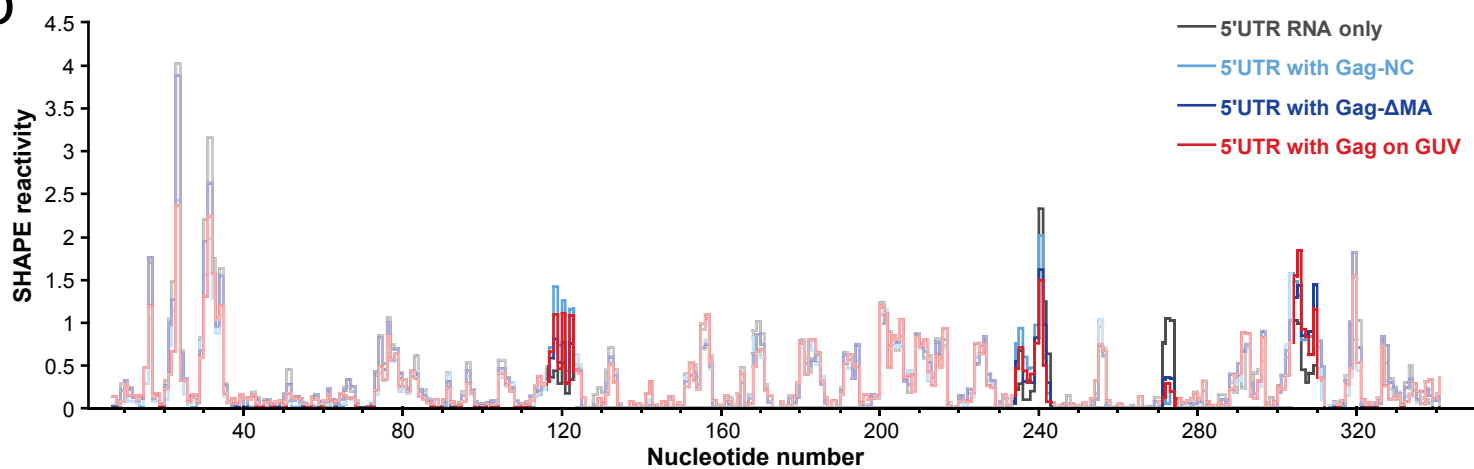
Figure 4



a



b



c

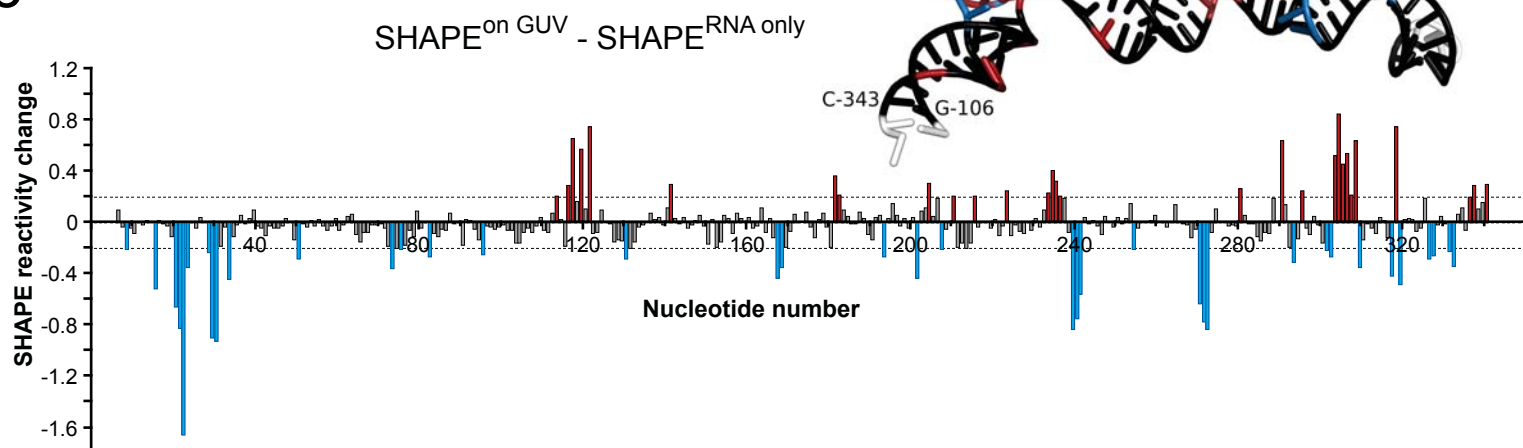


Figure 6

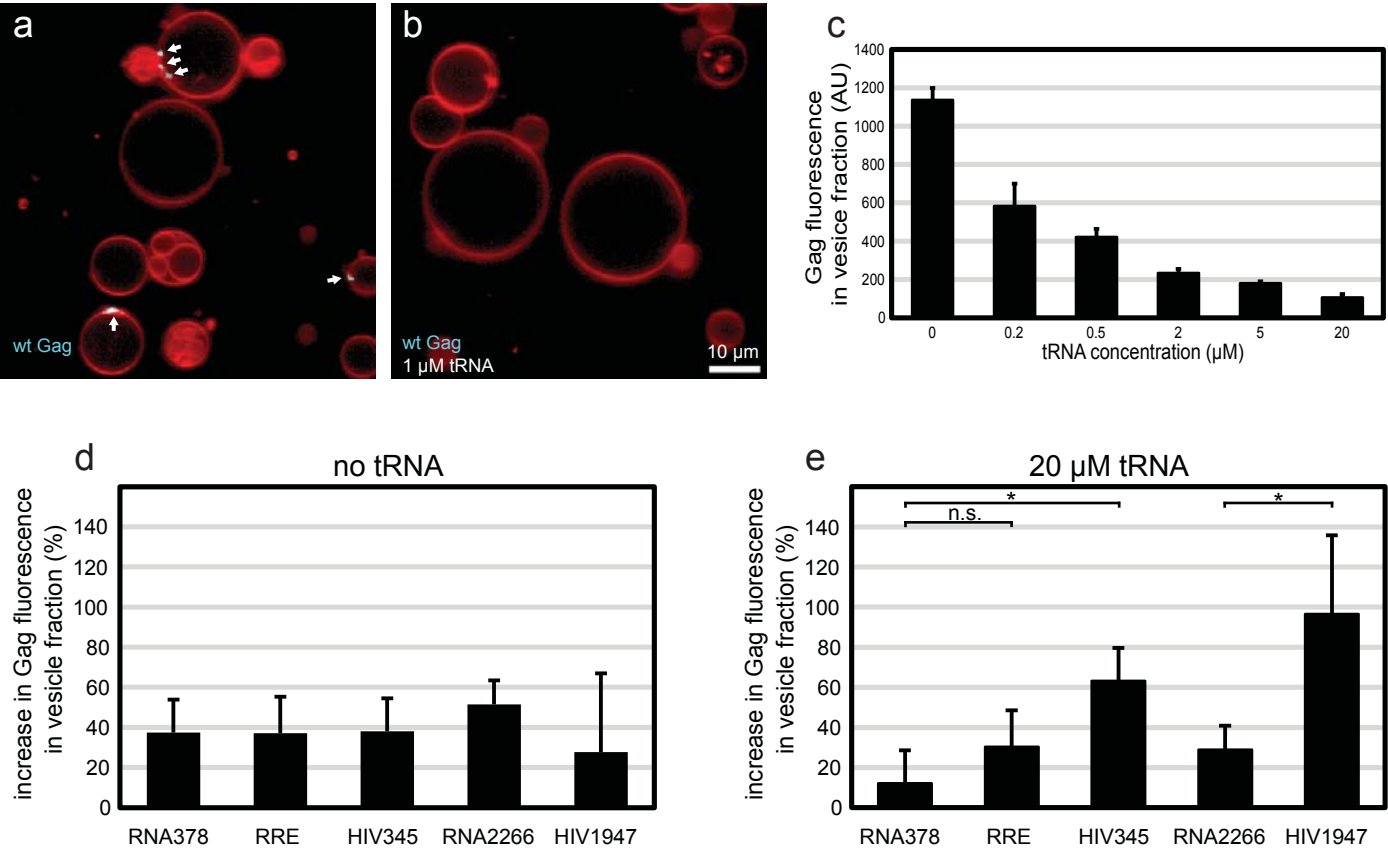


Figure 7

