Dimerization of the SP1 Region of HIV-1 Gag Induces a Helical Conformation and Association into Helical Bundles: Implications for Particle Assembly

Siddhartha A. K. Datta,^a Patrick K. Clark,^b Lixin Fan,^c Buyong Ma,^d Demetria P. Harvin,^a Raymond C. Sowder II,^e Ruth Nussinov,^{d,f} Yun-Xing Wang,^g Alan Rein^a

HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute—Frederick, Frederick, Maryland, USA^a; Retroviral Replication Laboratory, Basic Science Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA^b; Structural Biophysics Laboratory, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA^c; Cancer and Inflammation Program, Basic Science Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA^d; AIDS and Cancer Virus Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA^e; Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel^f; Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute—Frederick, Frederick, Maryland, USA^g

ABSTRACT

HIV-1 immature particle (virus-like particle [VLP]) assembly is mediated largely by interactions between the capsid (CA) domains of Gag molecules but is facilitated by binding of the nucleocapsid (NC) domain to nucleic acid. We previously investigated the role of SP1, a “spacer” between CA and NC, in VLP assembly. We found that small changes in SP1 drastically disrupt assembly and that a peptide representing the sequence around the CA-SP1 junction is helical at high but not low concentrations. We suggested that by virtue of such a concentration-dependent change, this region could act as a molecular switch to activate HIV-1 Gag for VLP assembly. A leucine zipper domain can replace NC in Gag and still lead to the efficient assembly of VLPs. We find that SP1 mutants also disrupt assembly by these Gag-Zip proteins and have now studied a small fragment of this Gag-Zip protein, i.e., the CA-SP1 junction region fused to a leucine zipper. Dimerization of the zipper places SP1 at a high local concentration, even at low total concentrations. In this context, the CA-SP1 junction region spontaneously adopts a helical conformation, and the proteins associate into tetramers. Tetramerization requires residues from both CA and SP1. The data suggest that once this region becomes helical, its propensity to self-associate could contribute to Gag-Gag interactions and thus to particle assembly. There is complete congruence between CA/SP1 sequences that promote tetramerization when fused to zippers and those that permit the proper assembly of full-length Gag; thus, equivalent interactions apparently participate in VLP assembly and in SP1-Zip tetramerization.

IMPORTANCE

Assembly of HIV-1 Gag into virus-like particles (VLPs) appears to require an interaction with nucleic acid, but replacement of its principal nucleic acid-binding domain with a dimerizing leucine zipper domain leads to the assembly of RNA-free VLPs. It has not been clear how dimerization triggers assembly. Results here show that the SP1 region spontaneously switches to a helical state when fused to a leucine zipper and that these helical molecules further associate into tetramers, mediated by interactions between hydrophobic faces of the helices. Thus, the correct juxtaposition of the SP1 region makes it “association competent.” Residues from both capsid and SP1 contribute to tetramerization, while mutations disrupting proper assembly in Gag also prevent tetramerization. Thus, this region is part of an associating interface within Gag, and its intermolecular interactions evidently help stabilize the immature Gag lattice. These interactions are disrupted by proteolysis of the CA-SP1 junction during virus maturation.

Immature retrovirus particles are assembled from a single protein, the virus-coded Gag protein. Gag is a multidomain protein in which the N-terminal matrix (MA) domain functions in the association of the protein with the plasma membrane of the virus-producing cell, the capsid (CA) domain contributes most or all of the Gag-Gag interactions in particle assembly, and the nucleocapsid (NC) domain is principally responsible for interactions of the protein with RNA (1). Assembly of immature particles *in vivo* is tightly coupled with nucleic acid binding and incorporation. The vast majority of HIV-1 particles contain a dimeric genome (2); in the absence of a packageable genome, cellular mRNA is incorporated, most of it nonselectively (3).

The HIV-1 Gag protein (lacking an N-terminal fatty acid modification and the C-terminal “p6” domain) can be produced in

Escherichia coli and purified as a soluble protein. Interestingly, the addition of virtually any single-stranded nucleic acid to the pro-

Received 13 August 2015 Accepted 25 November 2015

Accepted manuscript posted online 4 December 2015

Citation Datta SAK, Clark PK, Fan L, Ma B, Harvin DP, Sowder RC, II, Nussinov R, Wang Y-X, Rein A. 2016. Dimerization of the SP1 region of HIV-1 Gag induces a helical conformation and association into helical bundles: implications for particle assembly. *J Virol* 90:1773–1787. doi:10.1128/JVI.02061-15.

Editor: W. I. Sundquist

Address correspondence to Siddhartha A. K. Datta, dattasi@mail.nih.gov, or Alan Rein, reina@mail.nih.gov.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

tein leads to its assembly into virus-like particles (VLPs) (4–7). How does nucleic acid facilitate VLP assembly? It is striking that chimeric proteins, in which the NC domain of Gag has been replaced with a leucine zipper domain, assemble *in vivo* into VLPs (8). These particles are nearly identical in overall morphology to those assembled from full-length Gag protein, but unlike the authentic particles, they do not contain detectable RNA (9). As a leucine zipper promotes oligomerization, this observation raises the possibility that oligomerization of Gag is an essential prerequisite for assembly. Oligomerization might be induced either by cooperative binding of two or more Gag molecules to a single RNA molecule or, alternatively, by fusion to a dimerization domain such as a leucine zipper.

We and others have previously analyzed the properties of SP1, a 14-residue “linker” between the CA and NC regions in HIV-1 Gag (7, 10–15). This general region of Gag has been proposed to form a helical assembly motif: it is crucial in particle assembly, as many relatively subtle changes in its sequence drastically disrupt proper assembly (7, 10–15). We found that a peptide representing the last 8 residues of CA and the first 10 residues of SP1 formed an α -helix in aqueous solution at high but not low concentrations (10). Thus, this region of Gag is responsive to its own concentration. The sequence of this peptide indicates that it will form an amphipathic helix. It seemed likely that when the Gag concentration is high enough, the SP1 region within it could fold into a helix, and the juxtaposition of Gag molecules would permit the burial of hydrophobic residues in bundles of SP1 helices. In turn, this change in SP1 conformation, coupled with the association of multiple SP1 regions, might be “propagated” into the CA domain, triggering the exposure of new interfaces for Gag-Gag interaction as required for the assembly of spherical, immature particles. This might be the mechanism by which oligomerization of Gag leads to the assembly of immature virus.

In the present work, we dissect the mechanism underlying the switch-like behavior of the SP1 region. Much of the data presented here were obtained from analyses of very small proteins consisting simply of SP1 (plus the extreme C terminus of CA) fused to a leucine zipper, referred to as SP1-Zip. These proteins are small fragments of the assembly-competent Gag-zipper chimeric proteins (9) discussed above. Within SP1-Zip proteins, SP1 spontaneously assumes a helical conformation, even under dilute conditions where free SP1 is unstructured. While leucine zippers cause these proteins to dimerize, their helical SP1 moieties induce further association into discrete tetrameric forms. Mutagenesis shows that tetramerization of these proteins is driven by interactions between helical SP1 regions. Moreover, there is a near-perfect correlation between SP1 mutants that impair assembly *in vivo* and those that prevent tetramer formation of SP1-Zip chimeras. A number of features of these proteins are fully consistent with the general hypothesis that helix formation in SP1, once induced by a high local Gag concentration and the correct juxtaposition of the C termini, can render Gag assembly competent; moreover, they suggest for the first time that the interaction between SP1 regions is strong enough to contribute directly to the association of Gag proteins in particle assembly.

MATERIALS AND METHODS

Plasmids. Except where otherwise noted, all experiments were performed with derivatives of pCMV55M1-10, the plasmid directing the Rev-independent expression of HXB2 Gag in mammalian cells (16). This plasmid

was a kind gift of Barbara Felber (National Cancer Institute). We also used the Gag-Z plasmid, in which NC, SP2, and p6 of pCMV55M1-10 have been replaced by a leucine zipper motif (9). This chimeric protein is referred to as “Gag-Zip” here. The Gag-Z plasmid was used to construct the “SP1-Zip” plasmid, in which the C terminus of CA, together with all of SP1 (GHKARVLA EAMSQVTNSATIM), was fused to a leucine zipper motif. It was expressed in *E. coli* from a pET15b construct (Novagen). Each construct had an N-terminal His₆ tag followed by a thrombin cleavage tag, contributed by the pET15b vector. All deletions and point mutations were generated by using the QuikChange technique as directed by the manufacturer (Agilent Technologies). These deletions included the deletion of the entire CA-SP1 sequence, resulting in “LZip.”

Cells and viruses. All experiments involving mammalian cells were conducted with HEK 293T cells. Twenty-four-h harvests of culture fluid were collected 48 and 72 h after transfection. The fluids were first filtered through 0.45- μ m-pore-size filters, and VLPs were then collected by centrifugation at 25,000 rpm in an SW28 rotor (Beckman) for 1 h through a cushion of 20% (wt/wt) sucrose. Transfections, immunoblotting, and electron microscopy (EM) of thin sections of transfected cells were performed as described previously by Crist et al. (9). At least 20 fields were examined in each EM analysis.

Protein expression and purification. BL21(DE3) RIPL cells carrying the pET15b expression plasmids described above were induced with 1 mM isopropylthio- β -galactoside at 37°C for 5 h. Cells were lysed in a solution containing 20 mM HEPES (pH 7.5), 0.3 M NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP), and 6 M guanidinium chloride (GdmCl). After clarification by centrifugation at 20,000 $\times$ g for 30 min, the lysate was subjected to nickel affinity chromatography in 6 M GdmCl. Fractions containing protein were pooled and purified further by reverse-phase high-pressure liquid chromatography on a C₁₈ matrix (Waters Corp.) using a water-acetonitrile-trifluoroacetic acid solvent system. Fractions containing pure protein were pooled and lyophilized to dryness. The identity and chemical homogeneity of the proteins were confirmed by electrospray mass spectroscopy (model number 6130; Agilent). Subsequently, as needed, lyophilized proteins were denatured in a solution containing 20 mM Tris HCl (pH 7.3), 0.2 M NaCl, and 7 M urea and refolded by rapid dilution to $\sim$ 50 μ g/ml into a solution containing 20 mM Tris HCl (pH 8.0), 0.5 M NaCl, and 2 mM TCEP. The protein was dialyzed further against the same buffer and concentrated on Amicon Ultra centrifugal concentrators before use. Protein concentrations were determined by amino acid analysis and corroborated by UV spectroscopy using extinction coefficients of an optical density at 280 nm (OD₂₈₀) of 0.177 for 1 mg/ml of wild-type (WT) SP1-Zip and mutants, except for the A364W, A366W, and M367W mutants, where an OD₂₈₀ of 0.87 was assumed. Except where otherwise noted, all the data presented here were obtained with proteins containing the His₆ tag. In addition, however, size exclusion chromatography (SEC)-multiangle light scattering (MALS) and quasi-elastic light scattering (QELS) were also done on proteins from which the tags were removed, as indicated in Table 1. All of the circular dichroism (CD) data are for proteins without tags. In no case did the removal of the tag significantly affect the association properties of the proteins.

Protein cross-linking. Purified proteins in a solution containing 20 mM HEPES (pH 7.5) and 0.5 M NaCl were cross-linked in the presence of 2 mM Tris (2,2'-bipyridyl)dichlororuthenium(II) hexahydrate (Rubpy) and ammonium persulfate by controlled exposure to white light. Proteins in thin-walled PCR tubes at a distance of 3 to 5 mm from a planar 15-W cold fluorescent light source were exposed to light for various periods of time. Cross-linking was terminated by stopping further light exposure and adding loading dye containing 5% (vol/vol) β -mercaptoethanol. Cross-linked protein bands were resolved on SDS-PAGE gels and visualized by Coomassie brilliant blue staining.

Circular dichroism. The peptide spanning P356 to T373 of HIV-1 Gag (peptide P356-T373 [PGHKARVLA EAMSQVTNT]) was obtained from New England Peptide LLC (Gardner, MA) and further purified by reverse-phase high-performance liquid chromatography on a C₁₈ column

TABLE 1 Properties of SP1-Zip constructs^a

Construct	Theoretical molecular mass (Da)	MALS molecular mass (Da)	N-mer ^b	R_h (SEC)	R_h (QELS)	Assembly (NC)	Assembly (Zip)
WT	8,442.6	33,830; (8×) 31,070; (20×) 30,850 (24,730)	4	2.73/2.65/2.6	3.06	+	+
A364V	8,470.6	33,170 (24,660)	4	2.65	3.18	+	+
A364W	8,557.7	33,800; (12×) 31,090 (25,170)	4	2.66/2.60	2.85	+	+/-
A364P	8,468.6	19,710	2	2.27	2.45	—	ND
E365D	8,428.6	32,520	4	2.62	3.12	+	+
E365G	8,370.5	20,089	2	2.13	2.42	—	ND
E365Q	8,441.6	33,210 (24,520)	4	2.64	3.05	+	+
A366G	8,428.6	18,330	2	2.24	2.4	—	—
A366W	8,557.7	27,500 (UV)	3 ^c	2.32	— ^c	+/-	—
M367W	8,497.6	31,460; (5×) 26,130; (20×) 24,720	4 ^{c,d}	2.64/2.38/2.0	3.1	+/-	ND
ΔCTD	7,811.8	16,970 (12,710)	2	2.1	2.26	—	ND
Δ367	8,311.4	18,880	2	2.22	2.51	—	ND
Ins-GG	8,556.7	17,350 (13,830)	2	2.14	2.31	ND	ND
L-Zip	6,245.0	13,830 (9,656)	2	1.68	2.0	NA	NA

^a For each SP1-Zip construct, the molecular mass of the monomer; the molecular mass in solution as determined by MALS; the number of monomers in the oligomer present in solution; the R_h as determined by SEC, as described in the legend of Fig. 7A and B; the R_h as determined by QELS, as described in the legend of Fig. 6A; and the ability of full-length Gag ("NC") and Gag-Zip ("Zip") containing the respective SP1 sequence to assemble into ~100-nm spherical particles in mammalian cells are shown: +, assembles correctly, as in Fig. 3D; —, misassembles, as in Fig. 3F. The dilution factor and mass for dilutions of several SP1-Zip constructs are also shown. In some cases, the His₆ tag was removed with thrombin, reducing the monomeric molecular mass by 1,750.9 Da; underlined values in parentheses represent the molecular masses of these constructs. R_h values of the dilutions, as determined by SEC, are also shown. Elution of all constructs by SEC was measured by determining the refractive index, except for the A366W mutant, which was monitored by UV absorbance. ND, not determined; NA, not applicable.

^b N-mer, MALS molecular mass divided by theoretical monomeric mass, rounded to the nearest integer.

^c A366W SP1-Zip interacts with the column matrix and is polydisperse.

^d M367W SP1-Zip dissociates upon dilution.

with a linear gradient of water and acetonitrile containing 0.1% trifluoroacetic acid. After confirmation of homogeneity by matrix-assisted laser desorption ionization mass spectroscopy, the purified peptide was lyophilized to a powder. The lyophilized peptide was rehydrated overnight in deionized water at 4°C to make a stock solution at 20 mg/ml (10.4 mM). The peptide concentration was determined by UV absorbance spectroscopy at 205 nm (17) and confirmed by amino acid analysis. CD spectra of samples at various temperatures ($\pm 0.1^\circ\text{C}$) were acquired by using an Aviv 202 CD spectrometer (Aviv Instruments). Beginning with the 10.4 mM stock, the peptide was diluted into 0.5 mM sodium borate (pH 8.0). CD spectrum acquisition was initiated within 1 min of peptide dilution. CD spectra of SP1-Zip and LZip, at 0.06 and 0.03 mM, respectively, were acquired with a solution containing 10 mM sodium borate (pH 8.0) and 0.5 M NaCl, after thrombin cleavage of the His₆ tag. Buffer was supplemented with 30% (vol/vol) trifluoroethanol (TFE) as needed. Quartz cells (Hellma) with 1-mm, 0.2-mm, or 0.1-mm path lengths were used to acquire CD spectra in the region of 180 to 260 nm or 190 to 260 nm (wavelength step, 0.5 nm; averaging time, 1.000 s; settling time, 0.330 s).

MW determination and hydrodynamic measurements. Hydrodynamic properties and molecular weights (MWs) of SP1-Zip proteins, at concentrations of approximately ≤ 400 to $450 \mu\text{M}$, were measured on a Superdex 75 (GE Healthcare) SEC column, using a Rainin Dynamax UV-1 detector (Agilent), a Wyatt Systems Dawn Helios MALS and QELS detector (Wyatt Technology), and an Optilab T-rex interferometric refractometer (Wyatt). Columns were equilibrated in a solution containing 20 mM Tris HCl (pH 7.4), 0.5 M NaCl, and 0.5 mM TCEP. Proteins (0.2 to 3.0 mg/ml) were filtered through 300-kDa centrifugal filtration devices (Pall Corporation) before injection onto the column, using a flow rate of 0.3 ml/min. Bovine serum albumin (BSA; Sigma) was used to normalize detectors and establish detector train delay times. Data were collected simultaneously from light detectors 5 through 18, with the exception of detector 13, and used for MALS, while detector 13 was modified for QELS measurements. A 681-nm laser was used, and refractive index values were assumed to be 1.33, with 0.185 ml g^{-1} for the refractive index increment.

Molecular weights were determined from aligned elution profiles by using ASTRA for Windows software, version 6.1.2.84 (Wyatt).

Protein standards (GE Healthcare), including BSA (hydrodynamic radius [R_h] of 35.5 Å), ovalbumin (R_h of 30.5 Å), carbonic anhydrase (R_h of 23.5 Å), and cytochrome *c* (R_h of 17.2 Å), were used for the determination of R_h from retention time. Retention time was used to calculate the distribution coefficient K_D and the parameter $(K_D)^{1/3}$, which has a linear relationship to the R_h of the eluting species (18). Eluting peaks from the column were collected and concentrated on Amicon Ultra 3K centrifugal devices before confirmatory R_h measurements on a Wyatt Titan QELS instrument.

SAXS. Small-angle X-ray scattering (SAXS) measurements were carried out at room temperature at the 12-ID-B beamline of the Advanced Photon Source (APS), Argonne National Laboratory. A 14-keV X-ray beam was used as the photon source, with a sample-to-detector distance of 2 m for the SAXS setup. Thirty two-dimensional (2D) images were recorded for each buffer and sample solution using a flow cell, with an exposure time of 1 to 2 s to minimize radiation damage and obtain good statistics in the data. The 2D images were reduced to one-dimensional (1D) scattering profiles by using the Matlab software package at the beamlines. The 1D SAXS profiles were grouped by sample and averaged, followed by buffer background subtraction. WT SP1-Zip and E365G samples were measured at three different concentrations, 0.8, 1.25, and 2 mg/ml for WT SP1-Zip and 1, 2, and 3 mg/ml for the E365G mutant, in a final buffer containing 20 mM HEPES, 0.05% Na₃N, 500 mM NaCl, 1 mM TCEP, and 3% (vol/vol) glycerol (pH 7.3). The SAXS intensity is extrapolated to infinite dilution and zero scattering angles to remove the scattering contribution due to interparticle interactions. The procedures for data collection, processing, and analysis are similar to those previously described (19).

Fluorescence measurements. Tryptophan fluorescence emission spectra of proteins at the above-mentioned concentrations in a solution containing 20 mM Tris HCl (pH 7.4), 500 mM NaCl, and 0.5 mM TCEP were acquired at 25°C on a FluoroMax 2 instrument (Horiba Scientific).

Samples were excited at 280 nm, and the emission between 300 and 400 nm was recorded, with both excitation and emission band pass at 2.5 nm. For quenching experiments, the tryptophan fluorescence of proteins at 70 to 100 μ M was acquired after the sequential addition of small aliquots of freshly prepared 2 M potassium iodide (KI) in the same buffer. Spectra were recorded after a 1-min equilibration time after KI addition, using the parameters mentioned above. The intensity of fluorescence at 344 nm was used to generate the Stern-Volmer plots.

Molecular dynamics simulations. The protein sequence GSHMGHKARVLAAMSQVTNSATIMLRMKQLEDKVEELLSKNYHLENEVARLKKLVGE was used to search templates and build a homology structure using SWISS-MODEL (20). Several possible structures were tested. The C4 symmetry tetramer (tetramer 1) was based on the structure reported under PDB accession number 1GCL, and an antiparallel tetramer (tetramer 2) was based on the structure reported under PDB accession number 3NWH. The dimer and tetramer structures were built by using the structures reported under PDB accession numbers 3BAT and 2VKY, respectively, as the templates. The corresponding structures of the shortened peptide sequence AEAMSQVTNSATIMLRMKQLEDKVEELLSKNYHLENEVARLKKLVGE were based on the molecular dynamics (MD)-equilibrated structures by removing the GSHMGHKARVL residues and performing 60 ns of additional simulation on the remaining peptides. The SP1-Zip hexamer was built by using the following procedure. First, a monomer structure of HIV-1 CA-SP1 (N-terminal domain [NTD], C-terminal domain [CTD], and SP1) was constructed by merging known structures of NTD, CTD, and SP1 fragments (PDB accession numbers 3J4F, 4IPY, and 1U57, respectively). The monomeric CA-SP1 structure was then used to generate a hexameric CA-SP1 model, using the structure reported under PDB accession number 4D1K as a template (21). Finally, the monomer structure of SP1-Zip, used to generate the other models, was superimposed on the SP1 hexamer portion of the CA-SP1 hexamer assembly to obtain the starting structure of the SP1-Zip hexamer used in our MD simulations.

MD simulations of the solvated variant models were performed in NPT Ensemble using the NAMD program (22) with the CHARMM27 force field (23, 24) for 60 ns. The models were explicitly solvated with TIP3P water molecules (25, 26), with a minimum distance of 10 Å from any edge of the box to any protein atom. Sodium and chloride ions were added at random locations to neutralize the peptide charge and to maintain the NaCl concentration at 0.5 M.

The Langevin piston method (22, 27, 28) with a decay period of 100 fs and a damping time of 50 fs was used to maintain a constant pressure of 1 atm. The temperature (300 K) was controlled by a Langevin thermostat with a damping coefficient of 10 ps^{-1} (22). The short-range van der Waals interactions were calculated by using the switching function, with twin-range cutoffs of 10.0 and 12.0 Å. Long-range electrostatic interactions were calculated by using the particle mesh Ewald method with a cutoff of 12.0 Å for all simulations (29, 30). The equations of motion were integrated by using the leapfrog integrator with a step of 2 fs. The hydrogen atoms were constrained to the equilibrium bond length by using the SHAKE algorithm (31).

To obtain the relative structural stability of the variant models, the protein structure trajectories of the last 4-ns MD simulations were first extracted to exclude water molecules. The solvation energies of all systems were calculated by using the generalized Born method with molecular volume (GBMV) (32, 33). In the GBMV calculations, the dielectric constant of water was set to 80.0. The hydrophobic solvent-accessible surface area term factor was set to $0.00592 \text{ kcal/mol} \cdot \text{\AA}^2$. Each variant is minimized for 400 cycles, and the conformation energy is evaluated by the grid-based GBMV. A total of 1,600 conformations (400 conformations for each of the 4 examined conformers) were used to construct the free energy landscape of the conformers and to evaluate the conformer probabilities by using Monte Carlo simulations (34).

FIG 1 SP1 domain of HIV-1 Gag protein. (A) Schematic of the domain organization of HIV-1 Gag showing the location and sequence of SP1. (B) Helical wheel representation of the CA-SP1 junction sequence showing polar (black) and hydrophobic (gray) residues. (Republished from reference 10.)

RESULTS

Temperature dependence of helix formation by SP1. We and others have previously characterized the role of SP1 in HIV assembly. The sequence of the stretch of Gag investigated and its location within Gag are shown in Fig. 1A. The sequence spanning the last 8 amino acids of CA, together with SP1, has been shown to have a helical propensity (10, 11, 35, 36). Upon helix formation, the resulting helix would have a strong amphipathic character, with a polar face and a hydrophobic face, as shown in Fig. 1B. We previously found that an SP1 peptide (the last 8 residues of CA and the first 10 of SP1) undergoes a concentration-dependent structural transition in aqueous buffers (10). At low concentrations ($\sim 0.05 \text{ mM}$), the peptide exists as a coil, while at a far-higher concentration ($\sim 5 \text{ mM}$), it displays helical character. At an intermediate concentration of 2.7 mM , it has mixed character. The amphipathic nature of the resulting helix could explain why the secondary structure of this peptide is affected (10) by its own concentration. At high concentrations, helix formation could be driven and stabilized by the burial of these hydrophobic faces by bundle formation. To test this possibility, we determined the temperature dependence of helix formation. At an intermediate peptide concentration of $\sim 3.0 \text{ mM}$, raising the temperature from 10°C to 70°C caused the CD spectrum of the peptide (Fig. 2A) to increasingly adopt the high-concentration profile, resembling that seen at $>5 \text{ mM}$ (10), indicative of higher helical content. This temperature-dependent change in the CD spectrum is reversible (data not shown). However, at a low concentration (0.1 mM), where the peptide is unstructured, raising the temperature from 8°C to 70°C did not change the CD spectrum (Fig. 2B). The ratio of ellipticities at 222 and 208 nm, an indicator of helicity, is plotted

FIG 2 Temperature dependence of the secondary structure of SP1. (A and B) CD spectra of the SP1 peptide at 3.0 mM (A) and 0.08 mM (B) were obtained at different temperatures. (C) Ratio of ellipticity at 222 nm to that at 208 nm from panels A and B, plotted as a function of temperature.

as a function of temperature for the two concentrations in Fig. 2C. The fact that the helicity of the peptide increases with temperature at the higher concentration shows that entropic, rather than enthalpic, gains are primarily responsible for its adoption of a helical

conformation under these solution conditions, while the linkage of this effect to the concentration of the peptide implies that it involves intermolecular interactions.

Role of SP1 in VLP assembly of Gag-Zip proteins. To dissect further the mechanism of the structural transition of SP1, we have used an alternative approach. It is known that replacing the NC domain of Gag with leucine zippers allows for robust assembly of VLPs (8, 9, 37). These immature VLPs are similar in size and shape to authentic VLPs (9). Does SP1 contribute significantly to the assembly of these VLPs, as it does in WT Gag? We have examined the effects of mutations in SP1 in the context of Gag-Zip particle assembly. Gag-Zip proteins with mutations in SP1 that are innocuous (E365D, E365Q, A364W, and A366W) or deleterious (A366G) in a wild-type Gag background (10, 11) were tested. While the expression levels of these Gag-Zip proteins in cells were comparable (Fig. 3A), the presence of pelletable Gag in the culture fluid varied considerably (Fig. 3B). (Some degradation of Gag-Zip in cell extracts and dimerization in virus pellets were also evident.) We found that Gag-Zip proteins containing the SP1 mutations E365Q and E365D (not shown) assembled as efficiently as Gag-Zip itself. However, in the A364W, A366G, and A366W mutants, the release of pelletable material was sharply reduced.

Cells expressing these proteins were also examined by electron microscopy to determine whether the decreased efficiency of assembly was reflected in the disruption of VLP architecture. Gag-Zip VLPs with the WT SP1 sequence (Fig. 3D) are very similar in morphology to those formed by Gag (Fig. 3C) (9). We found that mutations in SP1 which do not impact assembly, such as E365Q (Fig. 3E) and E365D (not shown), are also not deleterious in a Gag-Zip background. Conversely, the A366G mutation, which disrupts assembly by WT Gag (10), also leads to misassembly of Gag-Zip VLPs, with the formation of blobs and tubes (Fig. 3F). A364W and A366W (Fig. 3G and H), which assemble well in the WT background (10), yield mixed phenotypes in Gag-Zip, characterized by some VLPs which are well formed and of the correct size but also many misshapen blebs formed at the membrane. Taken together, these results show that SP1 is critical for proper assembly of Gag-Zip proteins into VLPs, just as it is with WT Gag; moreover, the requirements for SP1 function are apparently similar in the two proteins, although they may be more stringent in Gag-Zip than in Gag VLP assembly. Therefore, the presence of the zipper domain does not circumvent the need for a functional SP1 for proper assembly. These correlations also suggest that SP1 adopts its functional state in the Gag-Zip particles, even in the absence of nucleic acids, by virtue of being fused with the leucine zipper domains.

Properties of SP1-Zip proteins. To shed some light on the contribution of SP1 to VLP assembly by WT Gag and Gag-Zip proteins, we analyzed the properties of small proteins consisting only of SP1 (plus the last 7 residues of CA) fused to a leucine zipper motif. These proteins are actually fragments of the assembly-competent Gag-Zip protein. Fusing it to the leucine zipper motif, which dimerizes with very high affinity, should place SP1 at a high local concentration, even at very low absolute concentrations.

The organization of and sequence contributions from the various Gag domains to these constructs are shown in Fig. 4. The constructs were made with an N-terminal His₆ tag, which can be removed with thrombin as necessary (Fig. 4A). The proteins were purified as described in Materials and Methods. Once checked for purity and chemical integrity, they were lyophilized to dryness and

FIG 3 Expression and assembly of SP1 mutant constructs in a Gag-Zip background. (A and B) Western blot detection of Gag-Zip proteins (arrows) using anti-p24^{CA} antiserum in cell lysates (A) and culture fluids (B). Lanes: 1, MW markers (in thousands); 2, A364W; 3, E365Q; 4, A366G; 5, A366W; 6, WT Gag-Zip; 7, pBluescript negative control. (C to H) Transmission electron micrographs of cells expressing various constructs: WT Gag (C) and WT and mutant Gag-Zip proteins (D to H).

then refolded from 7 M urea as described above. As needed, the N-terminal His₆ tag was cleaved off by using thrombin. As shown in the representative SDS-PAGE gel in Fig. 4B, they were free of detectable contaminating proteins after this purification. The properties of these proteins were then studied by a variety of techniques.

The conformation of WT SP1-Zip was first analyzed by CD spectroscopy. As shown in Fig. 5A, it appears to be completely helical. This is observed even under dilute conditions (~0.1 mM) where free SP1 peptide is unstructured (10). The CD spectrum is unaffected by the presence of 30% (vol/vol) TFE (Fig. 5A), a solvent which promotes helix formation (10, 35). These results sug-

FIG 4 SP1-Zip constructs. (A) Schematic and sequence of the SP1-Zip construct; (B) SDS-PAGE of purified SP1-Zip visualized by Coomassie brilliant blue staining. Lanes: 1, MW marker (in thousands); 2, uncleaved SP1-Zip; 3, thrombin-cleaved SP1-Zip.

FIG 5 Circular dichroism of SP1-Zip proteins. (A) Far-UV CD spectra of WT SP1-Zip at 0.06 mM in 10 mM Na borate, with and without 30% (vol/vol) TFE. (B) Spectra of LZip and of an equimolar mixture of LZip with SP1 peptide, in the absence and presence of 30% (vol/vol) TFE. See the text for details.

gest that SP1 is already helical in these fusion constructs. As a further control, we compared the CD spectrum of the SP1-Zip protein with that of a 1:1 mixture of the purified Zip domain and free SP1 under similar, dilute conditions. As shown in Fig. 5B, the spectrum of this mixture is quite different from that of the SP1-Zip fusion; presumably, in this mixture, the Zip domains form coiled coils, but the free SP1 remains unstructured. Moreover, the spectrum of the mixture is equivalent to the sum of spectra of free SP1 alone and the Zip domain alone (data not shown). The only difference between SP1-Zip shown in Fig. 5A and the 1:1 mixture shown in Fig. 5B (red squares) is that there is no peptide bond connecting SP1 to the Zip domain in Fig. 5B. The addition of TFE converted the free SP1 portion of the mixture to a helical conformation (Fig. 5B, purple x's) (10, 35); the resulting increase in helicity per residue of the mixture was easily detected by CD spectroscopy (Fig. 5B). As expected, the CD spectrum for the isolated Zip domain was insensitive to TFE. Taken together, the results show that when SP1 is fused to a zipper domain, it becomes helical, even under dilute conditions, as a consequence of being held close together by the zipper domains.

What are the oligomeric properties of the SP1-Zip protein? Analysis by dynamic light scattering (QELS) showed that SP1-Zip

is a homogeneous species, with an R_h of ~ 30 Å (Fig. 6A). This dimension seems too large for a monomer of an ~ 8 -kDa protein (for comparison, the R_h measured for BSA, 66.5 kDa, was 35 Å), suggesting the presence of an oligomer. We tested this possibility using ruthenium-mediated photo-cross-linking to probe the associative properties of the SP1-Zip protein. In the absence of a cross-linker, SP1-Zip migrates on a 4 to 12% reducing SDS-PAGE gel as a monomeric ~ 7 -kDa band, consistent with monomers of this 8.4-kDa protein (Fig. 6B, lane 2). However, upon photo-cross-linking in the presence of 2 mM Rubpy, light, multiple bands are seen. Cross-linking for ~ 20 s shows the clear presence of a dimer and traces of larger oligomers (Fig. 6B, lane 3). Further cross-linking for 2 min showed a clear ladder of oligomeric species (Fig. 6B, lane 4). The bands representing higher oligomeric species have decreasing intensities, but the drop of intensity does not occur linearly. Rather, the first 4 bands, representing the monomer through the tetramer, appear to be of higher intensity, followed by a sudden drop in intensity for the fifth and higher bands. These results suggest that the predominant species in the solution is a tetramer of SP1-Zip; higher cross-linked species (fifth band and higher) might be formed by cross-linking between tetramers and other species. The principal effect of increasing the Rubpy concentration with 2 min of irradiation was to intensify the dimer through tetramer bands (Fig. 6B, lane 5).

To further probe the association of SP1-Zip under non-cross-linked, native conditions, we used SEC-MALS. When the SP1-Zip construct was injected onto a Superdex 75 column, it eluted as a symmetrical peak, after ovalbumin (~ 43 kDa) and before carbonic anhydrase (~ 29 kDa). The mass of the eluting species, as determined by MALS, was ~ 33 kDa (Table 1 and Fig. 6C); thus, it is a tetramer, in full agreement with the cross-linking data shown in Fig. 6B.

What regions of the SP1-Zip proteins are involved in the observed tetramerization? To dissect this, we also analyzed the oligomeric status of SP1-Zip proteins with mutations in the SP1 sequence. The $\Delta M367$, ΔCTD (lacking the 7 residues from CA in the SP1-Zip construct), A366G (not shown), and A364P proteins all eluted from the SEC column significantly later than WT SP1-Zip. MALS showed that these proteins elute as dimers, with molecular masses of ~ 18 to 20 kDa (Fig. 7A and Table 1). In contrast, the A364V mutant, which assembles correctly in the WT Gag background (38, 39), elutes as a tetramer. Several other mutants were queried, and their oligomeric statuses are reported in Table 1. In contrast, free zipper was entirely dimeric, even at concentrations as high as 240 μ M (Table 1 and data not shown). The impact of the removal of the His₆ tag by thrombin on the oligomeric properties of several of the constructs was also tested: this had no effect on the oligomeric status of any of the proteins, as tested by SEC-MALS (summarized in Table 1) or by batch QELS (not shown).

The molecular dimensions of the eluting species were determined by both SEC, using a series of proteins with known R_h values as elution standards, and QELS. The results are shown in Table 1. The R_h values of the constructs fall into two broad groups: constructs forming tetramers have an R_h of ~ 2.8 to 3.1 nm, while dimeric constructs have an R_h closer to ~ 2.4 nm. The injection concentration of the proteins did not influence the mass of the eluting species in the constructs tested (Table 1), except for the M367W construct. This mutant eluted as a tetramer at higher concentrations (>0.25 mM), but upon 20-fold dilution, it eluted

FIG 6 Oligomeric status of SP1-Zip. (A) Estimation of R_h of SP1-Zip by QELS. Shown are the QELS fit, the raw autocorrelation data, and root mean square deviation of the fit. (B) Rubpy-mediated photo-cross-linking in SP1-Zip oligomers. SP1-Zip containing a His₆ tag was irradiated in the presence of Rubpy for different intervals, as described in Materials and Methods, and then analyzed by SDS-PAGE. Lanes: 1, MW markers (in thousands); 2, irradiation for 0 s; 3, irradiation for 20 s; 4, irradiation for 2 min; 5, irradiation for 2 min in the presence of 10× Rubpy. (C) SEC-MALS of SP1-Zip bundles. Shown are the elution chromatograms for SP1-Zip, ovalbumin (OVA) (~43 kDa), and carbonic anhydrase (CA) (~29 kDa).

later (not shown), with a smaller mass (Table 1). It thus appears that this mutation destabilized the tetramer. Sizes determined by using SEC standards showed a systematic difference (due to the globular nature of the SEC protein standards) from those determined by QELS, but the overall trend is the same for the two groups of constructs.

We also tested the role of the linkage between the SP1 and leucine zipper portions of the constructs in tetramerization. Interestingly, insertion of a pair of glycine residues between these two moieties (“InsGG”) prevented tetramerization (Fig. 7A and B and Table 1). This observation suggests that the precise, rigid apposition of the C termini of SP1 is essential for tetramer formation; it is also possible that there is a continuous α -helix spanning the SP1-zipper junction and that interruption of this helix with glycines destroys the structure required for tetramerization.

To obtain further information on the overall structure of the SP1-Zip tetramer, we also analyzed it by SAXS. The E365G SP1-Zip construct, which forms dimers rather than tetramers (Fig. 7B and Table 1), was analyzed as a control. The results are shown in Fig. 7C to F. The Guinier plot (Fig. 7D) shows that the radius of gyration (R_g) values of the two proteins are very similar to each other (26.58 versus 27.23 Å). Significantly, the Kratky plot (Fig. 7E), which reports on the overall fold of the molecule, indicates that the WT SP1-Zip construct shows substantial flexibility, while the E365G construct is largely unstructured. The pair distance distribution function [$P(r)$] plot on the WT construct (Fig. 7F) yields a $D_{\max}$ of ~94 Å; because of the lack of structure, it was not possible to construct the equivalent plot for the mutant. A second dimerizing mutant, A366G, was also analyzed by SAXS and gave results very similar to those found for the E365G mutant (data not shown).

Organization of SP1-Zip bundles. The organization of the SP1-Zip tetramer bundles was of considerable interest. We probed their structure by replacing individual SP1 residues with tryptophan and measuring the susceptibility of tryptophan fluorescence to quenching by KI, taking advantage of the fact that the action of this water-soluble diffusional quencher can be a measure of solvent accessibility of a tryptophan (40). The results of titration of KI into solutions of A364W, A366W, and M367W SP1-Zip proteins are shown as Stern-Volmer plots in Fig. 8A. The tryptophan in the A366W mutant is quenched most easily, while that of M367W is most protected and that at position 364 has intermediate accessibility to KI. This is reflected in Stern-Volmer quenching constants, K_{sw} , of 5.8, 4.0, and 2.15 M⁻¹ for A366W, A364W, and M367W, respectively.

To obtain further information on the organization of the bundles, we monitored the maximum emission wavelengths of the three tryptophan substitutions. Solvent-exposed tryptophans typically emit at ~355 nm, while tryptophans in a hydrophobic environment are blue shifted, emitting at lower wavelengths (40). As shown in Fig. 8B, the emission maxima of the A364W and A366W mutants are relatively constant at concentrations of between 2 and 100 μ M, at ~347 and 345 nm, respectively. In contrast, the emission wavelength of M367W increases continuously, from ~340 to ~351 nm, as the protein is diluted from 0.6 mM to 0.2 μ M. This increase suggests that the tryptophan at position 367 in this mutant is progressively solvent exposed as the protein is diluted, in complete harmony with the observed change in the apparent MW of the protein upon dilution (Table 1). As a control, we measured the emission maxima of the three proteins in 6 M GdmCl; they

FIG 7 Oligomeric properties of SP1-Zip mutants. (A) SEC-MALS analysis of WT SP1-Zip, A364V, Δ M367, InsGG, Δ CTD, and A364P. The mass of the eluting species as determined by MALS is shown as horizontal lines overlaid on the chromatograms. (B) R_h values for the various SP1-Zip bundles (labeled) and R_h standards (diamonds), as described in Materials and Methods. (C) SAXS curve for WT SP1-Zip at 0.8, 1.25, and 2 mg/ml. The graph also shows the results of extrapolation of the data to 0 mg/ml and to 0 mg/ml and a Q value of 0. a.u., arbitrary units. (D) Guinier fit of the data in panel C (WT) and of SAXS data for E365G SP1-Zip. (E) Kratky plot of SAXS data for WT and E365G SP1-Zip proteins. (F) $P(r)$ plot for WT SP1-Zip from extrapolation to 0 mg/ml and a Q value of 0, as shown in panel C.

were all ~ 355 nm (not shown), as expected. In summary, the results in Fig. 8A imply that the tryptophan at position 367 is in a more internal position in M367W helical bundles than those at position 364 or 366 are in their respective bundles. At the same time, it appears (Fig. 8B and Table 1) that the tryptophan at position 367 significantly destabilizes the bundles so that they have a strong tendency to dissociate upon dilution.

Molecular modeling. Under the experimental conditions used, why are tetramers of SP1-Zip observed? Using techniques described in Materials and Methods, we constructed models of SP1-Zip dimers, trimers, parallel and antiparallel tetramers, and hexamers (shown in Fig. 9) and compared their relative stabilities. The models were simulated at 0.5 M NaCl for 60 ns. As shown in

Table 2, we found that tetramer 1 with C4 symmetry was the most stable conformer, although the antiparallel tetramer 2 also had very comparable stability. The dimeric and trimeric structures modeled here were clearly less stable than the tetramers. The models depicted here have M367 shielded from the bulk solvent to a greater degree than A364 or A366, in conformity with the KI quenching data (Fig. 8A). The tested models also have overall dimensions and R_g parameters consistent with those determined for SP1-Zip by SAXS (Fig. 7). We also simulated possible SP1-Zip hexamer models. The initial model was built upon the hexameric immature lattice (21). This structure recapitulates the symmetry and the hollow ring of density attributed to SP1 in cryotomographic reconstructions (Fig. 9B, left). However, upon simulation,

FIG 8 Properties of tryptophan mutants of SP1-Zip. (A). Stern-Volmer plot of tryptophan fluorescence intensity as a function of KI concentration for A364W, A366W, and M367W SP1-Zip constructs. (B). Tryptophan fluorescence emission maxima as a function of protein concentration for A364W, A366W, and M367W SP1-Zip constructs.

these features are rapidly lost, reflecting the instability of the structure (Fig. 9B, right). In the final structure, the annulus is collapsed, and the six monomers are no longer interacting equivalently: the initial contacts are apparently too weak to maintain the symmetric hexameric configuration. Using the same conditions as those used for the models in Table 2, GBMV energy analysis indicated that the MD-optimized hexamer structures have much higher energy than even the dimer structure (the relative energy is 119 kcal/mol higher than that of the dimer). These results suggest one possible reason for our failure to detect SP1-Zip hexamers under our experimental conditions. As described above (Fig. 7 and Table 1), the “ Δ CTD” SP1-Zip construct is dimeric rather than tetrameric. We also modeled this construct; as indicated in Table 2, the antiparallel tetramer 2 is the most stable oligomer in these models, with the total tetramer conformers accounting for $\sim$ 55% of the population. However, these tetramers are thermodynamically less stable than those containing the CA residues, with the predicted dimer population increasing from 17.7% to 23.9% of the total.

DISCUSSION

In retroviruses such as HIV-1, expression of the viral Gag protein is sufficient for efficient assembly and release of virus-like particles in mammalian cells. All Gag proteins contain at least 3 domains (from the N to C termini): MA, CA, and NC. However, HIV-1 Gag also contains a 14-residue spacer, termed SP1, between CA and NC as well as an additional spacer and the p6 domain distal to NC.

We and others have previously shown that the SP1 region plays a crucial role in virus particle assembly, as even subtle changes in SP1 or in the last few residues of CA can drastically disrupt proper assembly (10–15, 41). This region may adopt a helical conformation in assembled virions, as cryotomographic analysis of immature HIV virions reveals a pillar of density, attributed to a hexameric bundle of SP1, beneath the lattice formed by the CA domain (42, 43).

The present work is aimed at further elucidating the contributions of the SP1 region (including the last few residues of CA) to particle assembly. It is striking that a chimeric protein in which the NC (RNA-binding) domain of Gag is replaced with a leucine zipper (dimerizing) motif assembles well in mammalian cells (8, 9, 37), forming particles that contain little or no RNA (9). The fact that the dimerizing zipper domain can supplant RNA binding suggests that assembly requires the juxtaposition of Gag molecules, and this association can be achieved either by cooperative binding to RNA or, alternatively, by attachment to a leucine zipper. SP1 is also important for proper assembly of this Gag-Zip protein, as mutations in SP1 interfere with proper assembly in mammalian cells (Fig. 3).

We have now analyzed the properties of a small fragment of this Gag-zipper chimeric protein, consisting of only the last 7 residues of CA, all of SP1, and the leucine zipper. We found that this protein forms a stable tetramer in solution (Fig. 6 and Table 1). Moreover, specific mutations in the SP1 moiety (or deletion of the CA segment) cause the formation of dimers rather than tetramers. This shows that interactions in the zipper portion lead to dimerization, as expected, while interactions between the CA-SP1 regions cause the dimers themselves to dimerize, forming tetramers. In the discussion that follows, the term “SP1” refers to the last few residues of CA as well as SP1 itself.

We also found that SP1 within the SP1-zipper fusion is helical at low concentrations in aqueous solution, while free SP1 is not (Fig. 5). Inspection of the SP1 sequence shows that it could form an amphipathic α -helix, with a polar face and a hydrophobic face. We previously reported that free SP1 peptide is helical at high but not at low concentrations. We therefore believe that when it is present at high local concentrations, it folds into a helix, and the hydrophobic faces of the monomers are buried in helical bundles. In support of this hypothesis, we show here that this conversion to helicity is entropically driven, as it is promoted by increasing temperature, though only at intermediate or higher concentrations (Fig. 2). However, we do not believe that these SP1 helices are in discrete oligomers, since QELS measurements show significant polydispersity (data not shown). The helicity of SP1-Zip in dilute solution (Fig. 5) shows that close association of only two SP1 molecules is sufficient to induce this conformational change. Taken together, these results raise the possibility that this region acts as a switch, undergoing a concentration-dependent shift in conformation, and that this shift is an essential early step in particle assembly. This general hypothesis is depicted in Fig. 10.

We introduced a number of mutations into the SP1 moiety of the SP1-Zip fusions. We observed a striking near-absolute correlation between mutations preventing tetramerization and mutations interfering with spherical particle assembly (Table 1). Tetramerization was blocked by mutations inserting helix breakers, such as A364P, E365G, and A366G. CD data showed that helicity was reduced in the E365G and A366G mutants (data not shown), while Kratky analysis of SAXS data (Fig. 7E and data not shown)

FIG 9 Molecular modeling of SP1-Zip oligomers. (A) Two orthogonal views of molecular models of dimers, trimers, parallel tetramers, and antiparallel tetramers. The right-hand view shows the positions of A364 (red), A366 (blue), and M367 (yellow) in the models. The SP1 sequence is highlighted in red. (B) Molecular model of a hexamer of SP1-Zip, built as detailed in Materials and Methods. Shown are initial (left) and final (right) structures upon simulation.

indicated that these constructs have greater flexibility than WT SP1-Zip. These data suggest that even upon close juxtaposition of SP1 peptides, helix formation is necessary for further association into tetramers. Other mutants that failed to tetramerize included Δ M367, which presumably permits helix formation but alters the register and amphipathic character of the helix, and InsGG, which alters the register, “decouples” SP1 from the C-terminal leucine zippers, and presumably increases the flexibility in the SP1 moiety. In contrast, SP1 mutants consistent with assembly in the Gag con-

text also tetramerize in the SP1-Zip context. This correlation suggests that the ability to form helices that associate into bundles is significant in normal particle assembly. While the predominant Gag-Gag interactions in particle assembly are presumably localized to the CA domain, the present results further suggest that SP1-SP1 interactions may also contribute significant binding energy in assembly. The constructs which formed tetramers did so under a range of concentrations tested, except for M367W. This construct forms a labile tetramer that falls apart upon dilution on

TABLE 2 Evaluation of relative stabilities of modeled structures from MD simulation and GBMV calculations

Structure	With CA sequence				Without CA sequence		
	Oligomer energy (kcal/mol)	Relative energy (kcal/mol) ± SD ^a	Relative population (%)	R _g (Å) ± SD	Oligomer energy (kcal/mol)	Relative energy (kcal/mol) ± SD ^a	Relative population (%)
Dimer	−4,769	602 ± 43	17.7	25.1 ± 0.1	−4,133	122 ± 18	23.9
Trimer	−7,244	240 ± 20	21.9	25.3 ± 0.1	−6,189	163 ± 17	21.8
Tetramer 1 (C4 symmetry)	−9,739	0	30.4	25.7 ± 0.1	−8,276	91 ± 16	25.0
Tetramer 2 (antiparallel)	−9,736	8.3 ± 19	30.0	27.8 ± 0.1	−8,307	0	29.3

^a Calculated based on equivalent dodecamer molecules (6 dimer, 4 trimer, and 3 tetramer molecules).

a size exclusion column (Table 1). A366W interacts with the column matrix, leading to a loss of recovered protein and an atypical retention time.

How might SP1-Zip bundles be organized? We do not have direct structural information, but the structure of individual SP1 helices in these bundles could well be the same as that of monomeric helical SP1 in 30% TFE (35). We attempted to gauge the solvent exposure of specific positions on the hydrophobic surface of helical SP1 within tetrameric SP1-Zip bundles. This was done by replacing individual amino acids with tryptophan and then monitoring the sensitivity to quenching of tryptophan fluorescence by the water-soluble quencher potassium iodide. These experiments showed that tryptophan at position 367 was relatively well shielded from the solvent, while position 366 was far more exposed; residue 364 showed intermediate accessibility (Fig. 8A). These results are very consistent with the idea that the hydrophobic faces of the SP1 helices are directed toward the interior of the tetrameric bundles, as position 367 is centrally placed within this face, while positions 364 and 366 are peripheral (Fig. 1B).

The fluorescence emission maxima (λ_{max}) of these tryptophan-substituted SP1-Zip constructs also provided some infor-

mation on their organization. In general, the λ_{max} of tryptophan fluorescence is reduced in a hydrophobic environment. We found (Fig. 8B) that the λ_{max} of A366W SP1-Zip was ~ 345 nm, and that of A364W SP1-Zip was ~ 347 nm; thus, both of these tryptophans are in environments with significant hydrophobic character. In contrast, the λ_{max} of M367W SP1-Zip was strongly affected by its concentration, being low (~ 340 nm) at high concentrations but higher (~ 351 nm), indicative of solvent exposure, in dilute solution. This conforms with data from the MW analysis, which showed (Table 1) that the apparent MW of the M367W mutant decreased upon dilution. It is clear that placement of tryptophan at position 367, in the middle of the hydrophobic face, destabilizes the tetrameric bundles. Although A366W SP1-Zip is somewhat polydisperse (Table 1), we believe that it is largely tetrameric; thus, the exposed tryptophan residue in these tetramers can evidently drive both nonspecific associations of this protein at high concentrations and nonspecific interactions with the SEC matrix, as discussed above. Interestingly, we previously reported that in a full-length virus genome, both M367W and A366W retain some assembly competence and infectivity, but A366W is significantly more impaired than M367W (10).

We also used MD simulation to explore the thermodynamic stability of different oligomeric forms of the SP1-Zip constructs. SP1-Zip models with dimeric, trimeric, as well as tetrameric arrangements (both parallel and antiparallel) were generated (Fig. 9), as described in Materials and Methods. These models have overall dimensions comparable to those determined experimentally by SAXS (Fig. 7F). As shown in Fig. 9, the models also place M367 into the hydrophobic interior of the structure, while A366 and A364 are more peripheral and solvent accessible, consistent with KI quenching data (Fig. 8). MD simulations predict that these tetramers, both parallel and antiparallel, are somewhat more stable than hypothetical dimers or trimers (Table 2). If the seven residues from the CA domain were to be removed from the chimeric protein models, the stability of the tetramers would be reduced somewhat, but according to these calculations, the tetramers would still be more stable than dimers, in contrast with our experimental results (Fig. 7A). The reason for this discrepancy is not clear. It is conceivable that the block to tetramer formation in this case is kinetic rather than thermodynamic.

It was somewhat surprising to find that the SP1-zipper proteins form tetrameric bundles, as the Gag protein forms a hexameric lattice in immature HIV-1 particles, and a “pillar” of density, suggested to be a 6-helix bundle of SP1 residues, is seen beneath the CA layer in cryotomographic reconstructions of these particles (42, 43). However, in Gag, SP1 is linked at its N terminus to the CA domain. CA is known to form hexamers under many conditions,

FIG 10 General model of the mechanism of the SP1 conformational switch in assembly. (A) When multiple Gag molecules bind side by side to an RNA molecule, their local concentration increases, and the SP1 region adopts a helical conformation (I). This change is propagated into the CA domain, leading to spherical, immature particle assembly (II). (B) When the SP1 region is linked to a leucine zipper, dimerization of the zipper brings two SP1 regions into correct and close proximity. This causes the SP1 region to adopt a helical conformation; interactions between these SP1 helices cause these constructs to self-associate into tetramers.

in both the immature and mature lattices (1, 43–45). Presumably, the presence of the CA domain in full-length Gag “overrides” the tendency of oligomeric SP1 to form tetramers. We also attempted to assess the stability of putative SP1-Zip hexamers by MD simulation; the results indicated that these hexamers would be less stable than dimers, trimers, or tetramers (Fig. 9). However, because of the strong correlation between tetramerization in SP1-Zip and proper assembly by Gag, we speculate that the structures of SP1 in tetrameric and hexameric bundles are very similar to each other, perhaps in analogy to the quasiequivalent structures of capsid proteins in the pentamers and hexamers found in simple icosahedral viruses. In both tetramers and hexamers, each SP1 monomer is in contact with two neighboring monomers; however, in a hexameric Gag lattice, the N terminus of SP1 is anchored to CA, while in an SP1-Zip tetramer, the N terminus is unconstrained.

It should be emphasized again that the sequence that we have studied here, both as part of the SP1-Zip fusions and as the free SP1 peptide (10), actually includes residues from the C-terminal end of CA as well as SP1. These CA residues are essential for tetramerization of the SP1-zipper fusions (Fig. 7 and Table 1) and for proper particle assembly (41). In other words, the functional unit under study here is partly in CA and partly in SP1; this implies that it is destroyed by cleavage during virus maturation. It seems possible that its destruction is a prerequisite for the changes in CA leading to the formation of the mature viral core.

Several laboratories have described a promising class of antiviral compounds termed “maturation inhibitors” (46–48). These compounds block cleavage between CA and SP1 *in vivo*. They appear to bind to the CA-SP1 cleavage site in immature virus particles (38, 49) but not to this site in unassembled Gag protein. We are currently testing whether they will bind to the SP1-Zip tetramers; if so, the tetramers would be extremely useful in screening for more members of this class of inhibitors.

It is intriguing to note the similarities and differences between the phenomena studied here in HIV-1 and those observed for other retroviral genera. In alpharetroviruses, betaretroviruses, and gammaretroviruses, as well as lentiviruses, point mutations near the C terminus of CA drastically interfere with immature particle assembly. There are many parallels between alpharetroviruses (e.g., Rous sarcoma virus) and lentiviruses like HIV-1: in both cases, there is more than one cleavage site between CA and NC, and a peptide representing the end of CA and the spacer between CA and NC becomes helical at high concentrations, apparently by forming 6-helix bundles and burying hydrophobic residues in the bundles (50). The suggestion that dimerization is a trigger for immature assembly originated in studies on alpharetroviruses (51–53). On the other hand, in gammaretroviruses such as Moloney murine leukemia virus, there is only a single cleavage between CA and NC. Near the C terminus of CA is the charged assembly helix motif or “electric wire,” a remarkable run of charged residues (with little or no net charge). If these were placed into a helix, the helix would have positively charged residues on one side and negatively charged residues on the other; it seems possible that gammaretroviral immature assembly is initiated by the association of these helices but that this association is driven by electrostatic forces rather than burial of hydrophobic residues. Sizable deletions within the electric wire retained proper function as long as they maintained helical register (54). Finally, in the

betaretrovirus Mason-Pfizer monkey virus (MPMV), there is only one cleavage between CA and NC. While the sequence of this region of Gag appears consistent with a helical conformation, careful studies found no evidence for a helix here (55). It is intriguing to note that a region of MPMV Gag on the N-terminal side of CA exhibits behavior somewhat analogous to that of HIV-1 SP1 (56). These differences highlight the diversity of mechanisms in retroviral assembly.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research; in part with funds from the Intramural AIDS Targeted Antiviral Therapy program; and in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN26120080001E. For the SAXS experiments, we gratefully acknowledge use of the SAXS core facility of Center for Cancer Research, National Cancer Institute (NCI). The SAXS data were collected at beamline 12-ID-B. The shared scattering beamline 12-ID-B resource is allocated under the PUP-24152 agreement between the National Cancer Institute and Argonne National Laboratory (ANL). The Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility, was operated for the DOE Office of Science by Argonne National Laboratory under contract number DE-AC02-06CH11357. Some measurements were performed in the Biophysics Core, Structural Biophysics Laboratory, NCI—Frederick.

We thank Bridget Heeney and Sofia Ryan for invaluable help with experiments, Mauricio Comas-Garcia for a critical reading of the manuscript, Raul Cachau for his interest in the project, and James David Roser for amino acid analyses. We gratefully acknowledge the help of Sergey Tarasov and Marzena Dyba of the Biophysics Core. We also thank Xiaobing Zuo for expert support at ANL.

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

FUNDING INFORMATION

Intramural AIDS Targeted Antiviral Therapy Program provided funding to Siddhartha A. K. Datta, Demetria Harvin, and Alan Rein. HHS | NIH | National Cancer Institute (NCI) provided funding to Patrick K. Clark, Lixin Fan, Buyong Ma, Raymond C. Sowder, and Ruth Nussinov under grant number HHSN26120080001E. HHS | NIH | National Cancer Institute (NCI) provided funding to Siddhartha A. K. Datta, Demetria Harvin, Yun-Xing Wang, and Alan Rein.

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

REFERENCES

1. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* 2:a006924. <http://dx.doi.org/10.1101/cshperspect.a006924>.
2. Moore MD, Hu WS. 2009. HIV-1 RNA dimerization: it takes two to tango. *AIDS Rev* 11:91–102.
3. Rulli SJ, Jr, Hibbert CS, Mirro J, Pederson T, Biswal S, Rein A. 2007. Selective and nonselective packaging of cellular RNAs in retrovirus particles. *J Virol* 81:6623–6631. <http://dx.doi.org/10.1128/JVI.02833-06>.
4. Campbell S, Fisher RJ, Towler EM, Fox S, Issaq HJ, Wolfe T, Phillips LR, Rein A. 2001. Modulation of HIV-like particle assembly in vitro by inositol phosphates. *Proc Natl Acad Sci U S A* 98:10875–10879. <http://dx.doi.org/10.1073/pnas.191224698>.

5. Campbell S, Rein A. 1999. In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J Virol* 73: 2270–2279.
6. Datta SAK, Rein A. 2008. Preparation of HIV-1 Gag protein and assembly of virus-like particles in vitro, p 197–208. *In* Prasad VR, Kalpana GV (ed), HIV protocols, 2nd ed, vol 485. Humana Press, Totowa, NJ.
7. Gross I, Hohenberg H, Wilk T, Wiegiers K, Grättinger, Müller B, Fuller S, Kräusslich HG. 2000. A conformational switch controlling HIV-1 morphogenesis. *EMBO J* 19:103–113. <http://dx.doi.org/10.1093/emboj/19.1.103>.
8. Zhang Y, Qian H, Love Z, Barklis E. 1998. Analysis of the assembly function of the human immunodeficiency virus type 1 Gag protein nucleocapsid domain. *J Virol* 72:1782–1789.
9. Crist RM, Datta SA, Stephen AG, Soheilian F, Mirro J, Fisher RJ, Nagashima K, Rein A. 2009. Assembly properties of human immunodeficiency virus type 1 Gag-leucine zipper chimeras: implications for retrovirus assembly. *J Virol* 83:2216–2225. <http://dx.doi.org/10.1128/JVI.02031-08>.
10. Datta SA, Temeselew LG, Crist RM, Soheilian F, Kamata A, Mirro J, Harvin D, Nagashima K, Cachau RE, Rein A. 2011. On the role of the SP1 domain in HIV-1 particle assembly: a molecular switch? *J Virol* 85: 4111–4121. <http://dx.doi.org/10.1128/JVI.00006-11>.
11. Accola MA, Hoglund S, Gottlinger HG. 1998. A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. *J Virol* 72:2072–2078.
12. Guo X, Hu J, Whitney JB, Russell RS, Liang C. 2004. Important role for the CA-NC spacer region in the assembly of bovine immunodeficiency virus Gag protein. *J Virol* 78:551–560. <http://dx.doi.org/10.1128/JVI.78.2.551-560.2004>.
13. Guo X, Roldan A, Hu J, Wainberg MA, Liang C. 2005. Mutation of the SP1 sequence impairs both multimerization and membrane-binding activities of human immunodeficiency virus type 1 Gag. *J Virol* 79:1803–1812. <http://dx.doi.org/10.1128/JVI.79.3.1803-1812.2005>.
14. Krausslich HG, Facke M, Heuser AM, Konvalinka J, Zentgraf H. 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. *J Virol* 69:3407–3419.
15. Liang C, Hu J, Russell RS, Roldan A, Kleiman L, Wainberg MA. 2002. Characterization of a putative alpha-helix across the capsid-SP1 boundary that is critical for the multimerization of human immunodeficiency virus type 1 gag. *J Virol* 76:11729–11737. <http://dx.doi.org/10.1128/JVI.76.22.11729-11737.2002>.
16. Schneider R, Campbell M, Nasioulas G, Felber BK, Pavlakis GN. 1997. Inactivation of the human immunodeficiency virus type 1 inhibitory elements allows Rev-independent expression of Gag and Gag/protease and particle formation. *J Virol* 71:4892–4903.
17. Scopes RK. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal Biochem* 59:277–282. [http://dx.doi.org/10.1016/0003-2697\(74\)90034-7](http://dx.doi.org/10.1016/0003-2697(74)90034-7).
18. Siegel LM, Monty KJ. 1966. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *Biochim Biophys Acta* 112:346–362.
19. Wang J, Zuo X, Yu P, Xu H, Starich MR, Tiede DM, Shapiro BA, Schwieters CD, Wang Y-X. 2009. A method for helical RNA global structure determination in solution using small-angle X-ray scattering and NMR measurements. *J Mol Biol* 393:717–734. <http://dx.doi.org/10.1016/j.jmb.2009.08.001>.
20. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42:W252–W258. <http://dx.doi.org/10.1093/nar/gku340>.
21. Bharat TA, Castillo Menendez LR, Hagen WJ, Lux V, Igonet S, Schorb M, Schur FK, Krausslich HG, Briggs JA. 2014. Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly. *Proc Natl Acad Sci U S A* 111:8233–8238. <http://dx.doi.org/10.1073/pnas.1401455111>.
22. Kalé L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K, Schulten K. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J Comput Phys* 151: 283–312. <http://dx.doi.org/10.1006/jcph.1999.6201>.
23. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217. <http://dx.doi.org/10.1002/jcc.540040211>.
24. MacKerell AD Jr, Brooks CL, III, Nilsson L, Roux B, Won Y, Karplus M. 1998. CHARMM: the energy function and its parameterization with an overview of the program, p 271–277. *In* Schleyer PVR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer HF III, Schreiner PR (ed), Encyclopedia of computational chemistry, vol 1. John Wiley & Sons, Chichester, United Kingdom. <http://dx.doi.org/10.1002/0470845015.cfa007>.
25. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935. <http://dx.doi.org/10.1063/1.445869>.
26. Mahoney MW, Jorgensen WL. 2000. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J Chem Phys* 112:8910–8922. <http://dx.doi.org/10.1063/1.481505>.
27. Feller SE, Zhang Y, Pastor RW, Brooks BR. 1995. Constant pressure molecular dynamics simulation: the Langevin piston method. *J Chem Phys* 103:4613–4621.
28. Martyna GJ, Tobias DJ, Klein ML. 1994. Constant-pressure molecular-dynamics algorithms. *J Chem Phys* 101:4177–4189. <http://dx.doi.org/10.1063/1.467468>.
29. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.
30. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. 1995. A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593. <http://dx.doi.org/10.1063/1.470117>.
31. Ryckaert J-P, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341. [http://dx.doi.org/10.1016/0021-9991\(77\)90098-5](http://dx.doi.org/10.1016/0021-9991(77)90098-5).
32. Lee MS, Feig M, Salsbury FR, Brooks CL. 2003. New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. *J Comput Chem* 24:1348–1356. <http://dx.doi.org/10.1002/jcc.10272>.
33. Lee MS, Salsbury FR, Brooks CL. 2002. Novel generalized Born methods. *J Chem Phys* 116:10606–10614. <http://dx.doi.org/10.1063/1.1480013>.
34. Ma B, Levine AJ. 2007. Probing potential binding modes of the p53 tetramer to DNA based on the symmetries encoded in p53 response elements. *Nucleic Acids Res* 35:7733–7747. <http://dx.doi.org/10.1093/nar/gkm890>.
35. Morelet N, Druillennec S, Lenoir C, Bouaziz S, Roques BP. 2005. Helical structure determined by NMR of the HIV-1(345–392)Gag sequence, surrounding p2: implications for particle assembly and RNA packaging. *Protein Sci* 14:375–386. <http://dx.doi.org/10.1110/ps.041087605>.
36. Newman JL, Butcher EW, Patel DT, Mikhaylenko Y, Summers MF. 2004. Flexibility in the P2 domain of the HIV-1 Gag polyprotein. *Protein Sci* 13:2101–2107. <http://dx.doi.org/10.1110/ps.04614804>.
37. Accola MA, Strack B, Gottlinger HG. 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* 74:5395–5402. <http://dx.doi.org/10.1128/JVI.74.12.5395-5402.2000>.
38. Adamson CS, Ablan SD, Boeras I, Goila-Gaur R, Soheilian F, Nagashima K, Li F, Salzwedel K, Sakalian M, Wild CT, Freed EO. 2006. In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (bevirimat). *J Virol* 80:10957–10971. <http://dx.doi.org/10.1128/JVI.01369-06>.
39. Adamson CS, Sakalian M, Salzwedel K, Freed EO. 2010. Polymorphisms in Gag spacer peptide 1 confer varying levels of resistance to the HIV-1 maturation inhibitor bevirimat. *Retrovirology* 7:36. <http://dx.doi.org/10.1186/1742-4690-7-36>.
40. Lakowicz JR. 1999. Principles of fluorescence spectroscopy, 2nd ed. Kluwer Academic/Plenum Press, New York, NY.
41. Melamed D, Mark-Danieli M, Kenan-Eichler M, Kraus O, Castiel A, Laham N, Pupko T, Glaser F, Ben-Tal N, Bacharach E. 2004. The conserved carboxy terminus of the capsid domain of human immunodeficiency virus type 1 Gag protein is important for virion assembly and release. *J Virol* 78:9675–9688. <http://dx.doi.org/10.1128/JVI.78.18.9675-9688.2004>.
42. Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, Krausslich HG.

2009. Structure and assembly of immature HIV. *Proc Natl Acad Sci U S A* 106:11090–11095. <http://dx.doi.org/10.1073/pnas.0903535106>.
43. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, Sundquist WI, Jensen GJ. 2007. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J* 26:2218–2226. <http://dx.doi.org/10.1038/sj.emboj.7601664>.
44. Gres AT, Kirby KA, KewalRamani VN, Tanner JJ, Pornillos O, Sarafianos SG. 2015. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science* 349:99–103. <http://dx.doi.org/10.1126/science.aaa5936>.
45. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M. 2009. X-ray structures of the hexameric building block of the HIV capsid. *Cell* 137:1282–1292. <http://dx.doi.org/10.1016/j.cell.2009.04.063>.
46. Kanamoto T, Kashiwada Y, Kanbara K, Gotoh K, Yoshimori M, Goto T, Sano K, Nakashima H. 2001. Anti-human immunodeficiency virus activity of YK-FH312 (a betulinic acid derivative), a novel compound blocking viral maturation. *Antimicrob Agents Chemother* 45:1225–1230. <http://dx.doi.org/10.1128/AAC.45.4.1225-1230.2001>.
47. Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, Matallana C, Castillo A, Zoumplis D, Martin DE, Orenstein JM, Allaway GP, Freed EO, Wild CT. 2003. PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc Natl Acad Sci U S A* 100:13555–13560. <http://dx.doi.org/10.1073/pnas.2234683100>.
48. Zhou J, Yuan X, Dismuke D, Forshey BM, Lundquist C, Lee KH, Aiken C, Chen CH. 2004. Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. *J Virol* 78:922–929. <http://dx.doi.org/10.1128/JVI.78.2.922-929.2004>.
49. Nguyen AT, Feasley CL, Jackson KW, Nitz TJ, Salzwedel K, Air GM, Sakalian M. 2011. The prototype HIV-1 maturation inhibitor, bevirimat, binds to the CA-SP1 cleavage site in immature Gag particles. *Retrovirology* 8:101. <http://dx.doi.org/10.1186/1742-4690-8-101>.
50. Bush DL, Vogt VM. 2014. In vitro assembly of retroviruses. *Annu Rev Virol* 1:561–580. <http://dx.doi.org/10.1146/annurev-virology-031413-085427>.
51. Johnson MC, Scobie HM, Ma YM, Vogt VM. 2002. Nucleic acid-independent retrovirus assembly can be driven by dimerization. *J Virol* 76:11177–11185. <http://dx.doi.org/10.1128/JVI.76.22.11177-11185.2002>.
52. Ma YM, Vogt VM. 2002. Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. *J Virol* 76:5452–5462. <http://dx.doi.org/10.1128/JVI.76.11.5452-5462.2002>.
53. Ma YM, Vogt VM. 2004. Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. *J Virol* 78:52–60. <http://dx.doi.org/10.1128/JVI.78.1.52-60.2004>.
54. Cheslock SR, Poon DT, Fu W, Rhodes TD, Henderson LE, Nagashima K, McGrath CF, Hu WS. 2003. Charged assembly helix motif in murine leukemia virus capsid: an important region for virus assembly and particle size determination. *J Virol* 77:7058–7066. <http://dx.doi.org/10.1128/JVI.77.12.7058-7066.2003>.
55. Strohalmová-Bohmová K, Spiwok V, Lepšík M, Hadravová R, Křížová I, Ulbrich P, Pichová I, Bednářová L, Ruml T, Rumlová M. 2014. Role of Mason-Pfizer monkey virus CA-NC spacer peptide-like domain in assembly of immature particles. *J Virol* 88:14148–14160. <http://dx.doi.org/10.1128/JVI.02286-14>.
56. Knejzlik Z, Ulbrich P, Strohalm M, Lastuvkova H, Kodicek M, Sakalian M, Ruml T. 2009. Conformational changes of the N-terminal part of Mason-Pfizer monkey virus p12 protein during multimerization. *Virology* 393:168–176. <http://dx.doi.org/10.1016/j.virol.2009.07.014>.