Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly

Jamil S. Saad, Jaime Miller, Janet Tai, Andrew Kim, Ruba H. Ghanam, and Michael F. Summers*

Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250

Edited by John M. Coffin, Tufts University School of Medicine, Boston, MA, and approved June 5, 2006 (received for review April 6, 2006)

During the late phase of HIV type 1 (HIV-1) replication, newly synthesized retroviral Gag proteins are targeted to the plasma membrane of most hematopoietic cell types, where they colocalize at lipid rafts and assemble into immature virions. Membrane binding is mediated by the matrix (MA) domain of Gag, a 132-residue polypeptide containing an N-terminal myristyl group that can adopt sequestered and exposed conformations. Although exposure is known to promote membrane binding, the mechanism by which Gag is targeted to specific membranes has yet to be established. Recent studies have shown that phosphatidylinositol (PI) 4,5-bisphosphate [PI(4,5)P_2], a factor that regulates localization of cellular proteins to the plasma membrane, also regulates Gag localization and assembly [Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. (2004) Proc. Natl. Acad. Sci. USA 101, 14889–14894]. Here we show that PI(4,5)P_2 binds directly to HIV-1 MA, inducing a conformational change that triggers myristate exposure. Related phosphatidylinositides PI, PI(3)P, PI(4)P, PI(5)P, and PI(3,5)P_2 do not bind MA with significant affinity or trigger myristate exposure. Structural studies reveal that PI(4,5)P_2 adopts an “extended lipid” conformation, in which the inositol head group and 2'-fatty acid chain bind to a hydrophobic cleft, and the 1'-fatty acid and exposed myristyl group bracket a conserved basic surface patch previously implicated in membrane binding. Our findings indicate that PI(4,5)P_2 acts as both a trigger of the myristyl switch and a membrane anchor and suggest a potential mechanism for targeting Gag to membrane rafts.

matrix protein | membrane targeting | NMR | lipid rafts

Retroviral genomes encode a polypeptide called Gag that contains all of the viral elements required for virus assembly (1). Subsequent to ribosomal synthesis, the Gag proteins are directed to punctate sites on plasma and/or endosomal membranes, where they assemble and bud to form immature virions (2–4). Approximately 1,500–5,000 copies of Gag contribute to the formation of a mature virion (5). Gag may also be transiently routed through the nucleus before assembly (26–28). In primary macrophages, budding occurs mainly in multivesicular body (MVB) compartments (14, 22). Gag may also be transiently routed through the nucleus before assembly (26–28). In primary macrophages, budding occurs mainly in MVBs (22–25). Recent studies indicate that the ultimate localization of Gag at virus assembly sites depends on phosphatidylinositol (PI) 4,5-bisphosphate [PI(4,5)P_2; ref. 21], a member of a family of differentially phosphorylated phosphatidylinositides that serve as membrane markers for specific cellular proteins (29–31). PI(4,5)P_2 is normally associated with the inner leaflet of the plasma membrane (PM) (30). Depletion of PI(4,5)P_2 inhibits virus assembly and leads to accumulation of Gag at the membranes of late endosomes and multivesicular bodies (21). Induction of PI(4,5)P_2-enriched endosomes also retargets Gag to endosome/MVBs and induces intravesicle budding (21). In both cases, virus production is severely attenuated (21). Substitution of MA by the membrane-binding N terminus of Fyn kinase reduces the sensitivity of virus assembly to PI(4,5)P_2 manipulation, suggesting that PI(4,5)P_2-dependent membrane selection is mediated by the MA domain of Gag (21). Consistent with this hypothesis, Reins and coworkers recently showed that phosphoinositides are capable of binding to the MA domain of unmyristoylated Gag fragments and promoting their assembly in vitro into virus-like particles (32, 33). To gain insight into the structural basis for PI(4,5)P_2-dependent membrane targeting, we conducted NMR studies of phosphatidylinositol interactions with the HIV-1 MA protein.

Results and Discussion

PI(4,5)P_2 Binds HIV-1 MA and Triggers Myristate Exposure. PI phosphates contain two long-chain fatty acids that promote micelle formation in aqueous solution (34). The most abundant cellular form of PI(4,5)P_2 contains stearate and arachidonate at the 1'- and 2'-positions of the glycerol group, respectively (35). Addition of substoichiometric amounts of this PI(4,5)P_2 species to unmyristoylated [myr(−)]MA and myrMA led to severe broadening in the 1H–31N heteronuclear single-quantum coherence (HSQC) NMR spectra, and at 1:1 PI(4,5)P_2:MA stoichiometries,

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: HSQC, heteronuclear single-quantum coherence; MA, matrix; myr(−), unmyristoylated; myr(s), myristate sequestered state; myr(e), myristate exposed state; PI, phosphatidylinositol; PI(4,5)P_2, PI(4,5)-bisphosphate; NOE, nuclear Overhauser effect; PM, plasma membrane; MVB, multivesicular body.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 2H3F, 2H3I, 2H3Q, 2H3V, and 2H3Z for myr(−)-MA, myrMA, di-C4-PI(4,5)P_2:myr(−)-MA, di-C4-PI(4,5)P_2:myr(−)-MA, and di-C4-PI(4,5)P_2:myr(−)-MA, respectively].

*To whom correspondence should be addressed. E-mail: summers@hhmi.umbc.edu.

© 2006 by The National Academy of Sciences of the USA
signals for backbone NH groups were broadened beyond detection. Attempts to study native PI(4,5)P₂ binding in the presence of dodecylphosphocholine detergent micelles, which were successfully used to study PI-3–phosphate interactions with the FYVE membrane-binding domain (36), were precluded by detergent-induced protein unfolding. Studies were therefore conducted with soluble forms of PI(4,5)P₂ containing truncated lipids.

Representative ¹H-¹⁵N HSQC NMR data obtained upon titration of myrMA with di-butyryl-PI(4,5)P₂ [di-C₄-PI(4,5)P₂] are shown in Fig. 1. Although a majority of the signals were relatively insensitive to titrations, a subset of signals corresponding to residues Arg-22, Lys-26, Lys-27, His-33, Glu-73, Leu-75, and Ser-77 (group 1) exhibited significant chemical-shift changes upon addition of di-C₄-PI(4,5)P₂ (Δδ¹⁵N ((Δδ¹⁵N)² + (Δδ¹H)²)¹/²) = 0.1–0.8 ppm; see Fig. 1). These residues reside on helices II and V and a meandering β-hairpin and contribute to a hydrophobic cleft (β-II-V cleft; refs. 37 and 38). Nonlinear least-squares fits of the titration data afforded a dissociation constant (K_d) of 150 ± 30 μM (mean ± SD; Fig. 1). Similar results were obtained upon titration of myr(−)MA with di-C₄-PI(4,5)P₂ (K_d = 240 ± 60 μM; Fig. 6, which is published as supporting information on the PNAS web site). Titration of myrMA with di-octyl-PI(4,5)P₂ [di-C₈-PI(4,5)P₂], which contains longer fatty acid chains, resulted in similar chemical-shift changes at low di-C₈-PI(4,5)P₂:myrMA ratios (<0.5:1), but aggregation and signal broadening at higher ratios precluded quantitative determination of the binding constant. Signal broadening was not observed upon titration of myr(−)MA with di-C₄-PI(4,5)P₂, which binds with ~3-fold greater affinity than di-C₄-PI(4,5)P₂ (K_d = 83 ± 4 μM; see Fig. 7, which is published as supporting information on the PNAS web site).

For the myristoylated MA protein, a second subset of residues (group 2: Gly-2–Ser-9, Glu-12–Asp-15, Arg-39, and His-89) that are well removed from the group 1 residues in the MA protein structure (37, 38) exhibit di-C₄-PI(4,5)P₂- and di-C₈-PI(4,5)P₂-dependent ¹H-¹⁵N HSQC signals that shift progressively toward values observed for myr(−)MA. Similar changes were observed previously for these residues upon concentration-dependent protein trimerization (19) and are indicative of a shift in the myristyl switch equilibrium from a predominantly myr(s) to a myr(e) state. Structural studies (below) confirmed that these spectral changes reflect di-C₄-PI(4,5)P₂-dependent exposure of the myristyl group.

Structure of the di-C₄-PI(4,5)P₂:myrMA Complex. The structure of the di-C₄-PI(4,5)P₂:myrMA complex was determined by using a combination of ¹⁵N-¹³C-, ¹³C/¹³C-, ¹⁵N/¹³C-, and ¹³C-edited/¹²C-double-half-filtered nuclear Overhauser effect (NOE) experiments (Fig. 1; refs. 39 and 40). NOEs observed between the myristyl group and core residues of myr(s)MA were absent in the di-C₄-PI(4,5)P₂:myrMA complex, confirming that di-C₄-PI(4,5)P₂ binding causes extrusion of the myristyl group. The NMR data are consistent with a single binding mode, in which the glycerol moiety packs against the side chains of His-33 and Trp-36, and the C₄-acyl chain attached to the 2'-position of the glycerol packs within the β-II-V cleft against the side chains of Leu-21, Lys-27, Tyr-29, His-33, Trp-36, and Ser-77. The phosphoinositide head group packs against Leu-21 and Lys-27, burying the 2'-fatty acid chain (Fig. 1). The 1'-acyl chain does not exhibit NOEs with the protein and appears exposed to solvent and disordered. In addition to the hydrophobic contacts, the 1'-phosphodiester is poised to make favorable electrostatic interactions with the positively charged side chains of His-33 and Lys-27, and the 4'- and 5'-phosphates are positioned to form salt bridges with the side chains of Arg-22 and Arg-76, respectively. No intermolecular
interactions were observed for Lys-30 or Lys-32, which were predicted to bind to PI(4,5)P2 on the basis of N-hydroxysuccin (NHS)-biotin protection experiments (32).

To determine whether the N-terminal myristyl group influences the binding mode, we determined the structure of the di-C4-PI(4,5)P2 complex with myr(−)MA. di-C4-PI(4,5)P2 binds myr(−)MA and myrMA in an essentially identical manner, as might be expected given the similar di-C4-PI(4,5)P2-dependent chemical-shift changes and binding constants (Fig. 8, which is published as supporting information on the PNAS web site). We also determined the structure of the di-C8-PI(4,5)P2 complex with myr(−)MA. This structure is very similar to that observed for di-C4-PI(4,5)P2:myr(−)MA, except that the additional residues of the 2'-fatty acid chain extend further into the β-II-V cleft (Figs. 2 and 7). The additional hydrophobic contacts appear to be responsible for the 3-fold enhanced affinity of di-C8-PI(4,5)P2 for myr(−)MA relative to di-C4-PI(4,5)P2. As observed in the di-C4-PI(4,5)P2:myrMA structure, the 1'-phosphodiester is poised to interact electrostatically with His-33 and Lys-27, and the 4'- and 5'-phosphates are poised to interact with Arg-22 and Arg-76, respectively (Fig. 3).

Myristate Exposure Is Induced by an Allosteric Mechanism. The binding of di-C4-PI(4,5)P2 to myrMA results in small local structural changes, as evidenced by very small changes in intraprotein NOE crosspeak patterns and intensities for residues Arg-20, Arg-22, Gly-25, and Gln-28 of the β-hairpin. No changes in intraprotein NOEs were detected for residues of helices II and V. However, significant NMR spectral and structural changes were observed for residues Myr-1–Lys-18 of myrMA. Residues Ser-9–Lys-18, which form an α-helix in the absence of di-C4-

![Fig. 2. Structure of the di-C4-PI(4,5)P2:myr(−)MA complex. (a) Interactions between di-C4-PI(4,5)P2 (sticks) and myr(−)MA (colored according to electrostatic surface potential) showing the 2'-fatty acid extending in a hydrophobic cleft and the inositol ring packing against a basic patch of the protein. (b) The structure is rotated ~90° relative to a, and the C4 acyl chains are shown in space-filling format.](image)

![Fig. 3. Structure of the di-C4-PI(4,5)P2- myr(−)MA complex showing the electrostatic interactions implicated in binding.](image)

![Fig. 4. Comparison of the myrMA structures before (gray) and after (slate) binding to di-C4-PI(4,5)P2. (a) Superposition showing that di-C4-PI(4,5)P2 binding induces only minor structural changes in the loop connecting helices I and II and much larger changes in the structure and orientation of helix I. (b) View of the N terminus of myr(s)MA showing the orientation of helix I and relative location of Glu-12 associated with sequestration of the myristyl group. (c) View of the N-terminal portion of the di-C4-PI(4,5)P2:myr(e)MA complex showing the packing and hydrogen-bonding interactions that stabilize the myr(e) conformation.](image)
tightly against the side chain of Trp-16 and occupy space that was formerly occupied by the terminal CH3 group of the myristate in the myr(s)MA structure (19). These structural elements are very similar to those observed in the myr(H11002/H9252/MA crystal structure (38). Myristate exposure thus appears to be triggered by an allosteric mechanism, in which PI(4,5)P2 binding induces small conformational changes in the /H9252-hairpin that in turn lead to more significant changes in the structure and orientation of helix I. These changes reposition hydrophobic residues near the N terminus of helix I in a manner that displaces the myristyl group and stabilizes the myr(e)MA species (Fig. 4).

Specificity of Phosphatidylinositide Binding. The above structural studies suggest that both hydrophobic and electrostatic interactions contribute to binding. To determine the specificity of binding, 1H-15N HSQC titration experiments were conducted with di-C4-PI containing other combinations of phosphate groups at biologically relevant positions. No detectable changes were observed in the NMR spectra of myrMA upon titration with PI(4,5)P2 binding induces small conformational changes in the /Hbeta-hairpin that in turn lead to more significant changes in the structure and orientation of helix I. These changes reposition hydrophobic residues near the N terminus of helix I in a manner that displaces the myristyl group and stabilizes the myr(e)MA species (Fig. 4).

Implications for the Mechanism of Membrane Targeting. Phosphatidylinositides comprise a class of differentially phosphorylated lipids that facilitate intracellular trafficking by establishing the identity of organelles. At least five phosphatidylinositides that differ in the number and position of attached phosphates are associated preferentially with Golgi [PI(4)P], early endosome [PI(3)P], late endosome [PI(3,5)P2], and plasma [PI(4)P, PI(4,5)P2, and PI(3,4,5)P3] membranes (30). Cellular proteins that interact differentially with these species can thus be targeted to specific membranes, enabling spatial distribution of subcellular activities (30). PI(4,5)P2 is considered a major landmark for proteins that need to associate with the PM (30), and it functions in the regulation of a variety of activities, including endocytosis, exocytosis, synaptic vesicle trafficking, and enzyme activation (31). Recent studies by Freed and coworkers (21) indicate that HIV-1 hijacks the phosphatidylinositide signaling system, and that PI(4,5)P2 [and possibly PI(3,4,5)P3] plays a critical role in targeting Gag to the PM.

The present studies demonstrate that PI(4,5)P2 can function as both an allosteric trigger for myristate exposure and a direct membrane anchor, providing a simple mechanism for targeting Gag to membranes enriched in PI(4,5)P2. That PI(3)P, PI(4)P, and PI(3,5)P2 do not bind MA with significant affinity or trigger myristate exposure is consistent with observations that Gag localizes and assembles at the PMs of most infected cell types. Membrane discrimination by this thermodynamic mechanism does not necessarily require additional cellular trafficking machinery, although we cannot rule out the possibility that other factors may be involved in trafficking events such as transient nuclear import/export (26, 28) or MVB targeting (22).

The hydrophobic cleft that interacts with the 2'-acyl chain of di-C4-PI(4,5)P2 and di-C8-PI(4,5)P2 is capable of accommodating longer fatty acids without alteration of the protein structure, and a model of a trimeric PI(4,5)P2:myrMA–membrane complex
constructed by using 18- and 20-carbon 1'- and 2'-acyl chains, respectively, is shown in Fig. 5. It is noteworthy that the lipid chains of PI(4,5)P2 extend in opposite directions, with the 1'-chain inserted into the lipid bilayer and the 2'-chain sequestered by the protein. This conformation, and the predicted membrane structure that are strikingly similar to the predicted in "extended lipid" phospholipid-cytochrome c models (41–43) and could be used to anchor other proteins to membranes as well (44). Although extrusion of the 2'-chain from lamellar membranes might intuitively be considered energetically expensive, a number of studies suggest this can actually relieve conformational stress caused by lipids with propensities for negative membrane curvature (41). Conformational dynamic studies also suggest that the 2'-chain is specifically favored for extrusion from the bilayer (44), and fluorescence quenching experiments indicate that the 2'-acyl chain of a brominated phospholipid is sequestered by cytochrome c upon binding to liposomes (43). Thus, the PI(4,5)P2-binding mode shown in Fig. 5 is consistent with previous theoretical and experimental studies of other cellular membrane-binding proteins.

Confocal microscopy has shown that Gag molecules assemble at punctate sites on the PM (14), and there is considerable evidence that these sites comprise lipid raft microdomains (45–48). Lipid rafts contain elevated levels of cholesterol and sphingolipids with saturated fatty acids and form liquid-ordered membrane structures (49–51). Although PI(4,5)P2 molecules are homogeneously dispersed within the PM of quiescent cells, and that they colocalize with lipid rafts upon stimulation by a mechanism that has yet to be identified (52). Our finding that the 2'-acyl chain is sequestered by the protein suggests a potential mechanism for the lateral targeting of PI(4,5)P2:Gag complexes to lipid raft microdomains. It is well known that rafts interact preferentially with saturated fatty acids (49–51). In fact, substitution of the saturated myristoyl group of HIV-1 Gag by unsaturated lipids reduces the affinity of Gag for rafts, but not for membranes in general, and thereby inhibits particle assembly (53). Proteins that bind lipid rafts generally contain two saturated acyl chains or are anchored by adaptor molecules that contain two saturated chains (for example, glycosphosphatidylinositol-anchored proteins; refs. 48 and 54). Because cellular phosphatidylinositides generally contain stearyl, an 18-carbon saturated fatty acid, at the 1'-position, and arachidonate, a 20-carbon fatty acid with four nonconjugated double bonds, at the 2'-position, sequestration of the 2'-chain is likely to reduce the affinity of PI(4,5)P2 for fluid regions of the membrane and promote its association within rafts. Differential sequestration of the acyl chains could serve as a general mechanism for the lateral re-targeting of phosphatidylinositides within the membrane, such as that observed during PI(4,5)P2-dependent microtubule assembly (52).

We previously demonstrated that myristate exposure in myrMA and myrMA–CA protein constructs can be promoted by concentration-dependent protein self association. Such a myristyl switch mechanism is consistent with a number of in vitro experimental observations. For example, mutations that inhibit Gag assembly also disrupt membrane binding (55–57), and C-terminal truncations lead to progressive decreases in both Gag multimerization and membrane affinity (58). In addition, the binding of Gag to nucleic acid templates, which promotes Gag self association (59–62), also enhances Gag interactions with membranes (63–65). The relative influence of Gag self association and PI(4,5)P2 binding for triggering myristate exposure in vivo is not clear. It is possible that a fraction of viral Gag molecules interact with PI(4,5)P2 at lipid rafts, and that this initial complex serves as a nucleation site for additional Gag molecules that bind to the membrane in a PI(4,5)P2-independent (but Gag self-association-dependent) manner.

The PI(4,5)P2-binding site is highly conserved among the 454 published strains of HIV-1, with Ser-77, Asn-80, and Lys/Arg-22 being strictly conserved; Leu-21, Trp-36, and Thr-97 substituted once, and Lys-27 substituted twice. In contrast, other exposed residues that do not participate in structure stabilization or PI(4,5)P2 binding are often extensively substituted. For example, Lys-30 [which exhibited PI(4,5)P2-dependent protection in accessibility assays (32)] is substituted by a nonbasic residue in 40% of the known HIV-1 isolates. Such high conservation may be necessary for sites that interact with cellular constituents, which do not undergo evolutionary changes on the timescale of viral replication. This could make the PI(4,5)P2-binding site an attractive antiviral target.

Materials and Methods

Sample Preparation. HIV-1 myrMA and myr(−)MA proteins were prepared as described (19, 37). Phosphoinositides PI(4,5)P2, di-C4-PI(4,5)P2, di-C8-PI(4,5)P2, di-Ca-PI, di-Ca-PI(3)P, di-Ca-PI(4)P, di-Ca-PI(5)P, di-Ca-PI(3,5)P2, di-Ca-PI(3,4,5)P3, and di-Ca-PI (Echelon, Salt Lake City, UT) and di-C4-phosphatidylcholine (Avanti Polar Lipids) were obtained commercially and used without further purification. Samples for all NMR experiments were prepared in 50 mM sodium phosphate at pH 5.5 and 5 mM DTT.

NMR Spectroscopy. NMR data [Bruker DRX (Billerica, MA); 800 MHz 1H and DIX 600 MHz 1H spectrometers equipped with cryoprobes] were obtained from a combination of 2D, 3D, and 4D NOESY data, for combinations of natural abundance; 15N-, 13C-, and 13C-labeled protein samples (35°C). Protein signals were assigned as described (19, 37). Phosphoinositide signals were assigned from 2D heteronuclear multiple-quantum coherence (HMQC), heteronuclear multiple bond correlation (HMBC), NOESY, total correlation spectroscopy (TOCSY), and COSY data (the 1H and 13C NMR signals of the 1- and 2'-acyl chains were resolved in the 2D spectra), and intra- and intermolecular 1H-1H NOEs were assigned from 2D (1H-1H), 3D (1H-15N-, 13C-edited) and 4D (15N- and 13C-edited) NOESY data (refs. 39 and 40 and references therein). Binding isotherms from 1H-15N NMR HSQC titration experiments were calculated with ORIGIN 7.0 software (Microcal Software, Northampton, MA).

Structure Calculations. Upper interproton distance bounds of 2.7, 3.3, and 5.0 Å (with appropriate corrections for pseudoatoms) were used for NOE crosspeaks of strong, medium, and weak intensity, respectively, which were qualitatively determined after intensity normalization of the different NOE data sets. No backbone hydrogen bond or chemical shift-based torsion angle restraints were used. 1H-13N residual dipolar couplings (RDCs) were measured for myr(s)MA, myr(−)MA, and PI(4,5)P2 complexes with myr(−)MA. Structures of myr(−)MA and myr(s)MA calculated with RDC restraints exhibited improved convergence but were in the range of structures calculated without these restraints (Tables 1 and 2, which are published as supporting information on the PNAS web site). Residual dipolar coupling (RDC) data obtained for myr(−)MA and myr(−)MA:di-C4-PI(4,5)P2 were very similar and, consistent with the NOE data, indicate that PI(4,5)P2 binding does not significantly alter the structure of the β-hairpin or the β-LV-binding site. RDCs could not be reliably measured for myristate-exposed species due to interactions with the alignment media and the tendency to form trimers (19). To avoid potential biases, comparisons with myr(e)MA were made by using structures calculated without the use of dipolar coupling restraints. Structures were calculated in torsion angle space with CYANA (L.A. Systems, Toshigi, Japan), starting from random initial angles. Statistical
information is provided in Tables 1 and 2. The lipid model shown in Fig. 5 is from ref. 66. Structure figures were generated with Pymol (http://pymol.sourceforge.net). We thank Eric Freed (National Institutes of Health, Bethesda) for critically reading our manuscript and alerting us to studies of fatty acid sequestration by other proteins. We thank the Howard Hughes Medical Institute (HHMI) staff and colleagues at the University of Maryland, Baltimore County, and David King (HHMI, University of California, Berkeley) for technical support. J.T. was supported by a MARC undergraduate fellowship (National Institutes of Health Grant GM08663). This work was supported by National Institutes of Health Grant AI30917.