



The HIV-1 Gag Protein Displays Extensive Functional and Structural Roles in Virus Replication and Infectivity

Veronna Marie * D and Michelle Lucille Gordon

KwaZulu-Natal Research, Innovation and Sequencing Platform, University of KwaZulu-Natal, Durban 4041, South Africa; tarinm@ukzn.ac.za

* Correspondence: vmarie@randwater.co.za

Abstract: Once merely thought of as the protein responsible for the overall physical nature of the human immunodeficiency virus type 1 (HIV-1), the Gag polyprotein has since been elucidated to have several roles in viral replication and functionality. Over the years, extensive research into the polyproteins' structure has revealed that Gag can mediate its own trafficking to the plasma membrane, it can interact with several host factors and can even aid in viral genome packaging. Not surprisingly, Gag has also been associated with HIV-1 drug resistance and even treatment failure. Therefore, this review provides an extensive overview of the structural and functional roles of the HIV-1 Gag domains in virion integrity, functionality and infectivity.

Keywords: Gag molecular structure; p17 Gag; p24 Gag; p7 Gag; p6 Gag; p2 and p1 Gag



Citation: Marie, V.; Gordon, M.L. The HIV-1 Gag Protein Displays Extensive Functional and Structural Roles in Virus Replication and Infectivity. *Int. J. Mol. Sci.* **2022**, 23, 7569. https://doi.org/10.3390/ ijms23147569

Academic Editors: Manuela Grimaldi and Jean-Marc Lancelin

Received: 28 April 2022 Accepted: 19 June 2022 Published: 8 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

In the human immunodeficiency virus type 1 (HIV-1), the autocatalytic cleavage of the aspartyl protease enzyme from Gag–Pol (PR160) prompts the cleavage of both Gag-Pol and Gag (PR55) at predefined sites in the polyprotein [1]. Group specific antigen, commonly referred to as Gag, is the HIV-1 polyprotein associated with viral structure and infectivity [2]. The 500 amino acid protein comprises four main structural domains (matrix, capsid, nucleocapsid, p6) and two small spacer peptides (SP1 and SP2) that aid in virion formation [2] as shown theoretically in Figure 1 below. For successful maturation to occur, a systematic proteolytic cleavage of Gag–Pol and Gag [3] is imperative. Furthermore, RNA dimerization as well as accurate RNA–Gag interactions are vital for the proteolytic cleavage of the substrate and enzyme [4].

Upon translation from the 9 kb transcript in the cytoplasm [5], Gag is trafficked to the plasma membrane where the membrane-bound polyprotein recruits and packages two copies of viral RNA per virion [6]. At the site of assembly, Gag binds cell membranes, undergoes multimerization and finally buds. Host factors employed by Gag facilitate scission and release of the immature viral particles from the cell membrane [6]. As a result, Gag and thereby its domains have several functional roles in viral replication and infectivity as summarized in Table 1.

Structural studies focusing on Gag and its assembly have furthered our understanding of Gag domain interactions with the host as well as with each other [4]. In particular, studies on individual Gag domains such as matrix, capsid and nucleocapsid in both mature and immature virions have been evaluated [7]. While it has been postulated that amino acid variation is observed in HIV-1 non-B vs. B subtypes [8], sequence analysis and structural studies indicate that strategic amino acid substitutions in Gag may also contribute to drug resistance [9–11].

Therefore, this paper will review the various structural and functional roles of the Gag domains in the viral maturation process, and thereby infectivity.

Gag Domain	Summarized Functional Roles
Matrix; p17	 Facilitates targeting of Gag–Pol and Gag to the plasma membrane Incorporation of Env into virions Associated with reverse transcription and the pre-integration complex during early stages of replication Binds tRNA (transfer RNA), rRNA (ribosomal RNA) and heterologous DNA Acts as a viral cytokine to promote a favourable network of virus- and host-derived stimulatory factors for optimal infection and replication
Capsid; p24	 Houses viral RNA and core-associated proteins Interaction with several cellular host factors and/or metabolites including cyclophilin A and inositol hexaphosphate (IP6) Interacts with cellular nucleoporins to allow import of the PIC The precursor PR55 Gag capsid domain and the mature p24 capsid display several functional roles in the viral replication cycle, including assembly, reverse transcription, integration and infectivity
Nucleocapsid; p7	 Packages viral RNA into rapidly assembling, immature viral particles Ability to renature nucleic acids Stimulates tRNALys binding to the primer binding site, initiates reverse transcription from the bound tRNALys and partakes in strand transfer Possesses nucleic binding and chaperone activity in the conversion of gRNA (genomic RNA) to viral dsDNA (double-stranded DNA)
p6	 Flexible docking site for cellular host factors, specifically ALIX and TSG101 Contributes to RNA binding specificity of gRNA of Gag PR55 Incorporation of viral accessory protein, Vpr Primary phosphoprotein Substrate for virus-associated kinases such as Erk-2 and Elk tyrosine kinases Contributes to RNA binding specificity of PR55 Gag
SP1	 May have active roles in viral assembly and possibly infectivity Associated with Gag–Gag interactions in budding of immature virus particles
SP2	 Plays a role in the incorporation of Pol and Gag in the immature virus particles Effect on the stability of nucleocapsid-RNA (NC–RNA) complexes and thus viral infectivity May have an effect on proviral integration when present as the p9 NC SP2 peptide

Table 1. Summary of known functional roles employed by the HIV-1 Gag domains in viral replication and infectivity.



Figure 1. Schematic representation of a rod-shaped HIV-1 Gag polyprotein obtained by the concatenation of previously resolved structures of individual Gag domains. PDB IDs: matrix (2h3i; [12]); capsid (3nte; [13]); SP1 (1u57; [14]); nucleocapsid (1a1t; [15]); p6 (2c55; [16]). Note: The matrix-capsid linker and SP2 are hypothetical representations.

2. Matrix

The matrix (MA; p17) protein is a 132 amino acid polypeptide that lines the inner surface of the virion lipid bilayer [4]. Structurally, MA consists of an amino-terminal (N-terminal) globular head and a flexible carboxyl-terminal (C-terminal) tail [17,18]. In MA, the globular head is composed of a 3_{10} helix and four α -helices that are capped by a three-stranded β -sheet. A fifth α -helix connects the capsid and MA domains which project away from the β -sheet, exposing the C-terminal residues and forming the tail [19]. Based on the ionic strength and concentration of protein buffers, MA may exist as either monomers, dimers and trimers or oligomers of monomers, dimers and trimers [20].

2.1. MA and Myristoylation

Functionally, the flexible nature of MA [21] allows for the targeting of the precursor Gag–Pol and the precursor PR55 Gag to the plasma membrane [22]. This process requires two important chemical signals, namely (i) myristoylation and (ii) the plasma membranespecific phosphatidylinositol-4,5-bisphosphate (PIP₂) [6]. Firstly, myristoylation occurs when N-methyltransferase catalyses a co-translational covalent bond between myristic acid and the G2 amino acid residue of the N-terminal domain (within the MGXXX{S/T}XX consensus sequence) of MA (Figure 2A) [4]. Since myristoylation increases the affinity of MA to the plasma membrane, it is not only critical for membrane targeting but for the production of immature, non-infectious virions [23]. Disruption of the myristoylation signal by mutations, such as G2A, can affect membrane targeting and virion production [24]. During myristoylation, myristic acid undergoes two reversible conformational transitions, known as the "myristoylation switch". These transitions are classified as exposed or sequestered within a hydrophobic pocket (Figure 2B) [25]. The hydrophobic pocket comprising the sequestered myristic acid is formed by the α -helices of MA [26]. As the myristoylation switch is dependent on several amino acid residues, particularly V7 and L8 [27], mutations at these positions can affect membrane targeting of Gag as well as inhibition of viral assembly and release [28]. In addition to these, the capsid domain, pH and the rate of calmodulin-binding may also impact the exposure of myristic acid [29].

The second signal involving PIP₂ and a cluster of basic residues localised around the cationic loop connecting β -strands one and two, otherwise known as the highly basic region (HBR; amino acids 17–31) [30], stabilizes membrane binding to the inner leaflet of the plasma membrane [31]. The HBR facilitates an electrostatic interaction between the protein and the negatively charged PIP₂ in the plasma membrane as seen in Figure 2C [30]. The exposure of myristic acid, and thereby its insertion into the lipid bilayer, is influenced by the HBR and the PIP₂ head group [25]. Upon transition from sequestered to exposed, the PIP₂ ³⁰KLKH³⁴ recognition site sequesters the PIP₂ fatty acid into the hydrophobic pocket (Figure 2C). The interaction between the plasma membrane and myristic acid thereby promotes the formation of a bi-directional lipid bilayer [4].

Although myristic acid exposure has been almost exclusively associated with MA trimerization (Figure 2B) [32,33], Valentine et al. [34] showed that after encapsidation into reverse micelles, myristic acid exposure could occur in monomer forms. Interestingly, molecular dynamic simulations conducted by Monje-Galvan and Voth [26] showed that while MA monomers did not form trimers, monomer interactions increased the concentration of PIP₂ and recruited it to the binding site, thereby enhancing viral assembly. Apart from the interactions between myristic acid and PIP₂, un-cleaved PR55 Gag multimerization may enhance its membrane binding activity through MA, indicating complex interactions between MA, capsid and the nucleocapsid domains (extensively reviewed in [35]). Furthermore, Wen et al. [36] showed that PR55 Gag can selectively target pre-enriched PIP₂ membranes for viral assembly of the immature viral particles, while the Gag multimerization, it is clear that the role of MA in membrane targeting is a complex process that is affected by numerous factors.



Figure 2. Proposed HIV-1 myristoylation at the molecular level. (**A**) Full length Gag polyprotein with a myristoylated MA at the N-terminal domain. (**B**) Sequestered (S) and Exposed (E) transition states in different MA conformations [37]. (**C**) The sequestered myristic acid switches to exposed upon PIP₂ binding at the plasma membrane. The PIP₂ fatty acid is then sequestered into the vacant hydrophobic pocket (adapted from [38]).

2.2. Other MA Functions

During the early stages of viral replication, MA is also associated with reverse transcription as well as the pre-integration complex (PIC) [39]. The PIC is a large nucleoprotein composed of several viral and cellular (host) proteins that allow integration during replication [40]. In addition to viral heterologous DNA and ribosomal RNA (rRNA), MA also binds to host transfer RNAs (tRNAs) [41]. A study conducted by Bou-Nadar et al. [41] demonstrated that aromatic and basic residues present within MA are preconfigured to recognize tRNA elbows. Though the authors indicate that it is highly improbable that tRNA and membrane binding occurs simultaneously, point substitutions that abolished MA-tRNA binding resulted in large scale redistribution of Gag to the plasma membrane and reduced HIV-1 replication [41]. Interestingly, MA also has cytokine-like functions that act on pre-activated human T cells encouraging proliferation, viral replication after p17R host-cellular receptor binding and pro-inflammatory cytokine release [42,43]. Therefore, MA may have several important known and/or unknown roles in the virus- and hostderived factors that contribute to a favourable HIV-1 infection and replication cycle [44]. Furthermore, it has been suggested that Env incorporation into the budding viral particles is mediated through interactions between the precursor PR55 Gag MA and the cytoplasmic tail of Env (gp41) [6]. This was substantiated by point mutations and deletions in MA that adversely affected Env incorporation without influencing particle formation [45,46]. In addition, biochemical analyses conducted by Tedbury et al. [47] showed that Env incorporation was negatively impacted by MA's inability to form trimers, thus indicating a correlation between MA trimerization and Env incorporation. However, the direct mechanism on Env incorporation remains unclear [6].

3. Capsid

In a mature virion, the HIV-1 capsid (CA, p24) protein oligomerizes to form a shell around the virus's RNA and core-associated proteins [48]. All morphological variances of the CA domain are derived from a collection of hexameric (rings) CA monomers [49] that have the same tertiary structure [50]. CA comprises two separate domains: the N-

terminal domain (CA_{NTD 1-145}) and the C-terminal domain (CA_{CTD 151-231}) [4], connected by a flexible inter-domain linker (residues 146–150) [51,52]. Structural studies revealed that CA is predominantly composed of α -helices, with seven and four α -helices in CA_{NTD} and CA_{CTD}, respectively [53]. Moreover, the NTD of CA contains an amino-terminal β -hairpin and resembles an arrowhead, whereas the globular CTD comprises one 3_{10} helix [53]. Extensive intrahexameric NTD–NTD and NTD–CTD interactions stabilize the central CA rings [53]. Accordingly, the first three helices of CA comprising the NTD–NTD interacting residues form a loose 18-helix central barrel that encompasses a hydrophobic core of aliphatic sidechains. In addition to a hydrophobic component, NTD-CTD interactions are primarily formed by sidechains extending from helix 8 of the CTD to helix 4 of the NTD. Further contacts are also made between helix 7 (C-terminal end) and helix 11 of the CTD [54]. Flexible CTD–CTD interactions link the adjacent hexameric rings, thus allowing for the formation of the curved lattice [55]. Additionally, it has also been shown that the inositol hexaphosphate (IP6) metabolite has a significant role in stabilizing the CA hexamers upon incorporation [56]. Recently, Renner et al. [57] showed that the coordination of IP6 to promote the assembly of mature CAs capable of reverse transcription and infection was mediated by both arginine and lysine ring binding. Furthermore, the asymmetrical structure infers that no two CAs are identical in capsid [58]. In addition to hexamers, formation of the mature "fullerene cone" of capsids is also dependent on the incorporation of 12 CA protein pentamers. Zhao et al. [58] showed that pentamer incorporation results in proximal trimer contacts by way of sharp bite angles (approximately 137° between pentamer-hexamer subunits), thereby inducing acute surface curvature. Noteworthy, Mattei et al. [59] suggest that either the angle of curvature determines pentamer position or the pentamer positions are determined by the angle of curvature, as pentamers were present at sites of great curvature. Therefore, the flexibility and complex molecular interactions of CA support a variety of conformations that ultimately contribute to structural variation within the protein itself [60].

3.1. CA, Cyclophilin A and Other Host Factors

The flexible nature of CA also allows for the interaction of a range of host factors such as CPSF6, NUP153, NUP358, TRIM5 α and MxB [61]. One of the more notable host factors that binds to the CA domain of the precursor PR55 Gag is the proline isomerase cyclophilin A (CypA) [62]. An extended proline-rich loop between the fourth and fifth helices in the NTD compose the binding site for CypA [63]. The CypA–CA stoichiometry is approximately 1:10 per particle [64] with residue P90 being the most important amino acid for CypA binding [65]. While the role of CypA is still debatable [66], it has been suggested that the addition of CypA into CA may be an important defence mechanism against antiviral innate immune responses after viral entry into the host cells [67]. A structural study conducted by Liu et al. [68] showed that CypA is not only recruited to facilitate HIV-1 infection but to also stabilize CA itself. Interestingly, although CypA is imperative for group M HIV-1 replication, group O variants are insensitive to CypA-inhibiting drugs and replicate efficiently in CypA-deficient cells [63]. Consequently, the role of CypA in specific group M HIV-1 replication versus its absence in other groups is still largely unknown.

Since reverse transcription is carried out inside the mature CA, dynamic, charged pores positioned at the centre of each CA multimer allows for the import of nucleotides [57]. In addition to IP6, Xie et al. [69] demonstrated that CA interacts with cellular nucleoporins, including NUP358, to allow import of the PIC into the nucleus. Additionally, it has been shown that partially(intact) viral cores enter the nucleus by directly interacting with CPSF6, the disruption of which results in altered target sites for HIV-1 integration [70].

3.2. CA and the Major Homology Region

Each CA domain has a different function in HIV-1 morphogenesis. Although the NTD of CA is not necessary for immature virion assembly, it is indispensable for mature core formation [71]. Contrastingly, the CTD of CA is important for both the formation

of the core in mature particles and in virion assembly in immature particles [72]. In all retroviruses, a conserved sequence known as the major homology region (MHR) is located in the CTD of the precursor PR55 Gag CA [73]. The entire MHR region forms a compacted strand-turn-helix stabilized by a network of salt bridges and hydrogen bonds [74]. While this region (284 DIRQGPKEPFRDYVDRFYKTL 304) is not clearly understood [4,75], its presence has been suggested to have a role in CTD dimerization [23], thus contributing to the stability of the viral shell [76]. Although the MHR does not form part of the dimerization interface, it incorporates the binding region for some compounds that bind at this site [53]. It may therefore serve to interact with a cellular factor or viral component [74]. However, a structural study conducted on the SCAN domain (a dimerization element found in mammalian zinc-finger proteins with evolutionary and structural similarity to retroviral CTD of CA) suggests that the MHR contributes to the intertwined dimer interface via domain swapping [76]. The exchange of identical protein domains has been suggested as a mechanism for the evolution of oligomer interfaces as well as protein refolding and synthesis [77]. Ivanov et al. [78] showed that dimerization of a CA-swapped domain was perturbed via an alanine deletion. Mutations in this region may therefore affect viral assembly of the immature virus particles. Furthermore, the authors also suggest that a network of conserved hydrogen bonds may be implicated in the structural stability of the MHR and prevent protein unfolding [78]. In contrast, thermodynamic and kinetic experiments of domain-swapped CTDs by Bocanegra et al. [79] demonstrated that no correlation was observed between the MHR residues and protein domain swapping. Instead, the authors conclude that the MHR residues were alternatively imperative for stable CTD folding and therefore have important roles in immature CA assembly as well as other infectivity processes [79]. Though much research has been conducted on this region, it is evident that the true purpose of the MHR region is not yet known.

To conclude, as CA is involved in several stages of the viral replication cycle it is an attractive target for the development of inhibitory drugs. Consequently, therapeutic agents, such as Lenacapavir and PF74, have been shown to exhibit substantial anti-viral activity in both clinical [80] and in silico [81] studies, respectively.

4. Nucleocapsid

The nucleocapsid (NC; p7) protein, consisting of 55 amino acids [82], packages two copies of the viral RNA into rapidly assembling virions [83]. A prominent feature of NC is the presence of two highly conserved Cys-X2-Cys-X4-His-X4-Cys (CCHC) signatures that resemble zinc-finger motifs [83]. The zinc fingers, each comprising an aromatic residue (F16 in the N-terminal zinc finger and W37 in the C-terminal zinc finger) [84], coordinate a zinc ion [85] and are divided by a ⁴⁰⁶RAPRKKG⁴¹² basic domain linker [4]. These separately folded zinc fingers resemble beads attached to a string [22,86]. Nuclear magnetic resonance (NMR) studies indicate that the zinc fingers, positioned at the globular centre, are noninteracting and independently folded domains [86], whereas the N- and C-terminals of NC comprise a disordered structural conformation [82]. The chelation of zinc ions to the zinc fingers folds the central domain whilst the disordered N- and C-termini remain unfolded [87]. The flexible RAPRKKG linker between the W37 and F16 residues in NC modulates the spatial proximity of the globular structure [88]. In addition to these amino acids, residues V13, I24 and A25 in the first zinc finger and Q45 and M46 in the second zinc finger form a hydrophobic plateau in the central globular domain [87]. The hydrophobic plateau is essential for the interaction of small oligonucleotides and NC. Oligonucleotide-NC structural complexes show that the hydrophobic plateau allows NC to bind nucleic acids via multiple phosphate backbone and nucleotide base contacts [15]. In addition, W37 is stringently stacked with guanine, contributing to nucleic acid–NC binding energy [15]. Moreover, complex stability is further maintained through electrostatic interactions with the base amino acids in the zinc finger linker and the disordered NTD of NC [89]. Mutations affecting the ordered zinc finger folding and the formation of the hydrophobic plateau [85] modify the structure of the virion core, alter Gag trafficking [82] and result in the loss of viral

infectivity [90]. For example, the substitution of P408 for D408 in the basic linker resulted in non-infectious, immature virions [91]. Although spatial closeness and orientation of the zinc fingers may be essential for successful virion production, only transient inter-zinc finger interactions were observed. Weak signal perturbations at the knuckle–knuckle (one turn α -helix followed by a β -hairpin) [92] interface of the zinc fingers were not consistent with tightly packed globular proteins [84,87,91,93].

4.1. NC and Genome Packaging

Genome packaging in HIV-1 is dependent on interactions between an un-cleaved precursor PR55 NC domain and a 110 nucleotide Psi (ψ ; RNA packaging signal) segment [94] in the 5' RNA leader sequence [83]. The ψ segment comprises four stem loops (SL1–SL4) that are connected by short linkers [95]. The stem loops are proposed to have independent but redundant functions [94] which include the facilitation of genomic RNA dimerization by the dimer initiation site (DIS) in SL1; the packaging of un-spliced mRNAs when the signal is overlapped with the ψ -site by the major splice donor (SD) in SL2; heterologous RNA packaging by SL3; and the adoption of alternate structures by SL4 [22]. Lu et al. [96] depicted SL2 and SL3 of the RNA packaging signal bound to NC. The second and fourth G residues in the SL3 tetra loop (677GGAG680) interact with both zinc fingers. The flexibility of the zinc finger linker allows the NC domain to interface with double or single stranded genomes [22]. Although both zinc fingers share similar conformation as well as hydrophobic and polar amino acid distributions, the two have completely different functions. The first zinc finger is essential for RNA packaging. The substitution of an amino acid sequence corresponding to the first zinc finger results in a 15% reduction of RNA packaging efficiency. Contrastingly, the second zinc finger has been highly associated with virion stabilization rather than encapsidation [89].

Although SL3 is important in RNA packaging, SL3-deficent ψ -sites only reduce the efficacy of genome packaging without complete elimination [97]. Similarly, mutations in SL3 and other SL ψ -sites do not result in a total loss of genomic packaging but reduce its ability to behave in this way [98].

Furthermore, it has been demonstrated that the nucleic binding and chaperone activity (NAC) associated with NC promotes thermodynamically stable conformations of RNA and DNA [96]. Consequently, NAC activity plays an important role in the conversion of genomic RNA (gRNA) to viral double-stranded DNA (dsDNA) [84]. Recently, Jiang et al. [82] showed that NC plays an integral role in the compaction of viral dsDNA within mature HIV-1 CAs, suggesting a pathway for de-condensation upon un-coating and NC loss.

4.2. Other NC Functions

Other functions of NC include the ability to renature nucleic acids with catalytic rates of about four orders of magnitude [99]. In reverse transcription, NC can also stimulate tRNA^{Lys} binding to the primer binding site found at the N-terminal region of the genome, initiate reverse transcription from the bound tRNA^{Lys} and partake in strand transfer [100]. Additionally, NC was shown to modulate RNA G-quadruplex stability through binding and unfolding of the RNA G-quadruplex as well as the promotion of DNA–RNA duplex formation, thereby allowing reverse transcription to proceed successfully [101].

5. p6

The C-terminal region of Gag comprises a proline-rich, 52 amino acid p6 domain [102] where two amino acid sequences are translated: the -1 frameshift Gag–Pol p6 and the in-frame Gag p6 domain [4]. Although p6 is largely considered to have a random structure in aqueous solution displaying little, if any secondary structure [67], Solbak et al. [103] showed that in a 100 mM dodecylphosphocholine micelle solution at pH 7, p6 forms defined N- and C-terminal helices that are connected by a flexible hinge. In spite of its debatable structure, it is widely accepted that p6 serves as a flexible docking site for cellular host factors [102–104]. HIV-1 p6 comprises two unique late domains [102]. As part

of the precursor PR55 Gag polyprotein, these domains recruit and bind the endosomal sorting complex required for transport (ESCRT) cellular factor to promote the abscission of immature virions from the host cell [95]. The first late domain, YPXL (Tyr-Pro-X-Leu), binds cellular factor ALG2-interacting protein X (ALIX) [105]. Although the YPXL late domain varies between retroviruses, the conserved tyrosine always binds within a pocket on the ALIX V domain, and hydrophobic residues contact ALIX along a shallow groove [106]. Apart from p6, ALIX and its isolated Bro domain also interact with NC [107]. While this interaction could be RNA dependent, it may be non-specific [4]. The second late domain, PTAP binds the tumour susceptibility gene 101 (TSG101) cellular protein [67]. The four PTAP residues make contact with the N-terminal ubiquitin E2 variant (UEV) domain of TSG101 [108]. Not only are PTAP motifs found in related proteins that recruit ESCRT but also in the human HRS protein. Therefore, p6 can be regarded as a viral mimic of the

cellular ESCRT-1 recruiting motif in HIV-1 [109].
Furthermore, while it is known that the precursor PR55 Gag NC domain mediates the selection of gRNA by interacting with the beginning of the *gag* gene gRNA and the ψ segment, the p6 domain also contributes to the RNA binding specificity of PR55 [110]. Using a truncated form of PR55 lacking the p6 domain, Dubois et al. [110] demonstrated a significant reduction in affinity for gRNA and high affinity for spliced cellular and viral RNAs. This suggests that the p6 domain has an essential role in the discrimination of viral and cellular RNAs from the selective gRNA required for encapsidation during immature particle assembly.

Interestingly, Yu et al. [111] demonstrated that point mutations at highly conserved residues, specifically at position F8 in the transframe p6 region of Gag–Pol, result in insufficient viral processing due to impaired protease and reverse transcriptase activity.

5.1. p6 and Vpr

During the late stage of replication, the p6 domain, as part of the un-cleaved precursor PR55 Gag, is also necessary for the incorporation of viral accessory protein, Vpr, into the virions [112]. The 96 amino acid protein is comprised of three amphipathic α -helices specifically arranged to form a hydrophobic centre surrounded by flexible sequences [113]. Propagation in HeLa cells showed that Vpr oligomer formation and its targeting to the nuclear envelope is mediated by this hydrophobic centre [114]. Vpr is associated with trans-activation of long terminal repeats (LTR), import of the PIC in non-dividing cells, the induction of apoptosis (programmed cell death) and cellular halt at the G_2/M transition pathway [115]. The G_2/M transition is a point in a cell's lifecycle where, after the second growth phase (G2) and DNA replication (S phase), mitosis occurs (M phase) to separate the cell into identical daughter cells. In addition to translation being repressed, many of the cell's components undergo a dramatic structural change that occurs during the G2/M transition pathway. This is imperative if identical daughter cells are to be produced [116]. Vpr is therefore critical in viral pathogenesis and must be incorporated in viral particles [117]. The ⁴¹LXXLF⁴⁵ motif in p6 has been frequently associated with Vpr binding [102]. In contrast, the 15 FRFG¹⁸ and 34 ELY³⁶ p6 motifs were found to be associated with Vpr packaging [104]. Wanaguru and Bishop [104] further demonstrated that removal of either the FRFG or LXXLF binding motifs coupled with the disruption of Vpr oligomerization significantly reduced Vpr incorporation by approximately 25- to 50-fold. Additionally, the authors also observed that although Vpr is concomitantly lost with p6 during infection, remnants of Vpr remain associated with CA for several hours, implicating Vpr functionality in early replication.

5.2. p6 and Phosphorylation

The p6 domain is also regarded as the primary phosphoprotein in HIV-1 and serves as a substrate for virus associated kinases, such as the extracellular signal regulated kinase 2 (Erk-2) targeting threonine; the atypical protein kinase C (aPKC) targeting serine [102]; and Elk tyrosine kinases [118]. The highly conserved S40 residue associated with phosphorylation has been related to core formation and capsid manipulation in infectious particles.

However, the functional relevance of phosphorylation in p6 residues is still debatable [103]. Radestock et al. [119] attempted to elucidate the role of p6 phosphorylation by using a comprehensive mutational analysis. The analysis included a wild type proviral genetically uncoupled (unc) *gag* and *pol* plasmid (pNL4-3_{unc}); a derivate with non-phosphorylation residues but with similar chemical structures (pNL4-3_{unc}FL); and a derivative with wild type S40 and all other substitutions (pNL4-3_{unc}FL-N40S). Controls included a derivate with a substituted alanine in FRFG and diminished Vpr incorporation; pNL4-3 Vpr(-), a derivate with a defective late domain; and pNL4-3 late(-) and a derivate that does not express Vpr (pNL4-3 Δ Vpr). The virus (NL4-3_{unc}FL, NL4-3_{unc}FL-N40S and S40N) displayed no significant change in replication capacity in comparison to the wild type virus (NL4-3_{unc}). Consequently, the authors suggested that phosphorylation of p6 residues may not be essential for viral replication and morphogenesis, but may be structurally important as the small, hydrophilic serine is favoured over the hydrophobic, bulky phenylalanine [119]. Other authors have alternatively shown that the S40F substitution may diminish cleavage at the CA |SP1 site [120,121].

6. SP1 and SP2

The 14 amino acid SP1 (³⁶⁴AEAMSQVTNSATIM³⁷⁷), known interchangeably as p2, is found wedged between the N-terminal CA and C-terminal NC proteins [4]. The C- and N-termini of the CA and SP1 domains, respectively, were shown to form two parts of an α -helix [122]. Mutations in residues M367 and A366 within SP1 and K359 and H358 just outside SP1 were depicted to alter the conformation of the α -helix, subsequently inhibiting infectivity [123].

Although the structure of SP1 has not been resolved independently of the other Gag domains [67], current molecular studies show that SP1 has a propensity to form an α -helix with residues beginning from L343 in CA to V390 in SP1 directly linked to helix formation [14,124]. Interestingly, a circular dichroism analysis revealed that this helical conformation was considerably enhanced when 30% trifluoroethanol, a helix-promoting solvent, was subsequently added to the peptide solution after it had been purified. This conformational switch may imply the active role of SP1 in viral assembly [124]. However, as the primary assembly unit of the Gag lattice remains largely unknown [125], maintenance of this helical conformation in full length Gag is debatable. Electron cryotomography analyses were consistent with SP1 forming a six α -helical bundle that stabilized Gag hexamers in immature virus particles [126]. As a result, cleavage of SP1 from NC is necessary for the formation of ribonucleoprotein within RNA and condensation of the CA domain [127]. The late cleavage of CA | SP1 allows morphogenesis through CA–CA interactions. The importance of the integrity of the CA | SP1 segment in formation of immature CA was shown in an electron microscopic study [128].

In addition, the SP1 | NC and CA | SP1 regions have been associated with Gag–Gag interactions as deletion of the peptide can lead to arrested ribonucleoprotein budding [123]. Atomic-level resolution structural studies on the CA | SP1 cleavage segment of Gag are necessary since this region not only contributes to virion assembly but may also have an important role in the HIV-1 maturation inhibitor binding site [129]. At present, inhibitors such as Bevirimat and PF-96 have been developed to block CA | SP1 cleavage during maturation [122]. However, natural polymorphisms, predominantly the V370A substitution in SP1, produce high level drug resistance to these inhibitors [130,131].

The 16 amino acid SP2 (⁴³³FLGKIWPSYKGRPGNF⁴⁴⁸), also known interchangeably as p1, lies between the CTD of NC and the NTD of p6. This peptide contains two highly conserved proline residues, namely P445 and P439 [4]. Hill et al. [132] implicated the importance of SP2 for Pol and Gag incorporation into immature virus particles. The "slippery site" involved in the Gag–Pol ribosomal frameshift overlaps with the end of SP2 leading to potentially complex effects of mutations at this site [4]. The substitution of either proline (P445 and P439) by leucines results in lower stability of the NC–RNA complex and rescinds infectivity [133]. Moreover, mutations at the cleavage sites at either end of SP2 have discrepant effects. For example, mutations in the p9 peptide (NC|SP2) have no effect on proviral integration, whereas mutations at the C-terminal that produces the p15 (NC|SP2|p6) or p8 (SP2|p6) peptides decrease the amount of integrated provirus on subsequent infection. This differential effect may suggest the role of p9 in proviral integration as opposed to p15 and p8 [134].

7. Conclusions and Future Perspectives

The Gag polyprotein and its individual subunits have significant roles in the lifecycle and perseverance of HIV-1. While full length Gag structures have not yet been resolved, which is partly due to their large molecular size, increasing our understanding of Gag at the structural level will further our knowledge on the elucidation of pathways to resistance and ultimately advancements in drug therapy. The complex, yet intricate roles that the Gag polyprotein and its individual subunits play in viral replication and infectivity, as reviewed in this article, indicate that while our understanding is not lacking, it is also not yet complete. Therefore, the following important aspects are still far from clear and requires extensive research:

- As MA plays a crucial role in membrane targeting, as well as Env incorporation, it is a constantly evolving area of research. Extensive mapping of these independent yet linked pathways, especially where myristoylation is concerned, can possibly aid in the development of drug inhibitors targeting this stage in the viral lifecycle.
- 2. As CA proves to be one of the most dynamic Gag domains, it is imperative that the role of cellular host factors, such as CypA as well as the MHR region is understood to further research in the area of therapeutic agents. In addition, HIV-1 subtype-associated discrepancies in the role of CypA must be further explored.
- The NC domain is an important component for nucleic acid packaging in immature viral particles. Its role in promoting stable nucleic acid conformations in the conversion–condensation–decondensation process requires extensive research to understand this mechanistic approach to replication.
- 4. p6 serves several functional roles in the overall viral replication process. One of the most intriguing aspects of p6 is its highly conserved residues and their role in interacting with viral and host cellular factors. Further understanding of these residues and their affinity for being phosphorylated can aid in the development of therapeutic targets for this conserved site.
- 5. The spacer peptides (SP1 and SP2) remain elusive. Though some important work has been conducted on these short segments, their role in viral replication and the maintenance of structural integrity is lacking. As with the main Gag domains, complex interactions are constantly being observed. However, what is evident is that the spacer peptides may not only serve supportive roles to the Gag domains, but may also have independent roles in viral assembly and infectivity on their own.

Author Contributions: V.M. and M.L.G. both contributed equally to the writing, review and editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the South African Medical Research Council (SIR grant), Poliomyelitis Research Foundation (grant # 16/67) and the University of KwaZulu-Natal's College of Health Sciences Ph.D. grant.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the National Research Foundation, South Africa for providing the Ph.D. scholarship.

Conflicts of Interest: The authors declare they have no competing interest.

References

- Lin, Y.-R.; Chu, S.-M.; Yu, F.-H.; Huang, K.-J.; Wang, C.-T. Effects of reduced gag cleavage efficiency on HIV-1 Gag-Pol package. BMC Microbiol. 2022, 22, 94. [CrossRef]
- 2. Samsudin, F.; Gan, S.K.-E.; Bond, P.J. The impact of Gag non-cleavage site mutations on HIV-1 viral fitness from integrative modelling and simulations. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 330–342. [CrossRef]
- Gilmer, O.; Mailler, E.; Paillart, J.-C.; Mouhand, A.; Tisné, C.; Mak, J.; Smyth, R.P.; Marquet, R.; Vivet-Boudou, V. Structural maturation of the HIV-1 RNA 5' untranslated region by Pr55^{Gag} and its maturation products. *RNA Biol.* 2022, 19, 191–205. [CrossRef]
- 4. Bell, N.M.; Lever, A.M.L. HIV Gag polyprotein: Processing and early viral particle assembly. *Trends Microbiol.* **2013**, *21*, 136–144. [CrossRef]
- 5. Sertznig, H.; Hillebrand, F.; Erkelenz, S.; Schaal, H.; Widera, M. Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet. *Virology* **2018**, *516*, 176–188. [CrossRef]
- 6. Murphy, R.E.; Saad, J.S. The interplay between HIV-1 Gag binding to the plasma membrane and Env incorporation. *Viruses* **2020**, 12, 548. [CrossRef]
- 7. Dick, A.; Cocklin, S. Recent advances in HIV-1 Gag inhibitor design and development. Molecules 2020, 25, 1687. [CrossRef]
- Teto, G.; Nka, A.D.; Fokam, J.; Bouba, Y.; Takou, D.; Fabeni, L.; Carioti, L.; Armenia, D.; Semengue, E.N.J.; Dambaya, B.; et al. Detection of Gag C-terminal mutations among HIV-1 non-B subtypes in a subset of Cameroonian patients. *Sci. Rep.* 2022, *12*, 1374. [CrossRef]
- Li, G.; Verheyen, J.; Rhee, S.-Y.; Voet, A.; Vandamme, A.-M.; Theys, K. Functional conservation of HIV-1 Gag: Implications for rational drug design. *Retrovirology* 2013, 10, 126. [CrossRef]
- 10. Marie, V.; Gordon, M. Gag-protease coevolution shapes the outcome of lopinavir-inclusive treatment regimens in chronically infected HIV-1 subtype C patients. *Bioinformatics* **2019**, *35*, 3219–3223. [CrossRef]
- 11. Marie, V.; Gordon, M. Understanding the co-evolutionary molecular mechanisms of resistance in the HIV-1 Gag and protease. *J. Biomol. Struct.* **2021**, 1–10, *in press.* [CrossRef] [PubMed]
- 12. Saad, J.S.; Miller, J.; Tai, J.; Kim, A.; Ghanam, R.H.; Summers, M.F. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Nacl. Acad. Sci. USA* **2006**, *103*, 11364–11369. [CrossRef] [PubMed]
- Du, S.; Betts, L.; Yang, R.; Shi, H.; Concel, J.; Ahn, J.; Aiken, C.; Zhang, P.; Yeh, J.I. Structure of the HIV-1 full-length capsid protein in a conformationally trapped unassembled state induced by small-molecule binding. *J. Mol. Biol.* 2011, 406, 371–386. [CrossRef] [PubMed]
- Morellet, N.; Druillennec, S.; Lenoir, C.; Bouaziz, S.; Roques, B.P. Helical structure determined by NMR of the HIV-1 (345–392) Gag sequence, surrounding p2: Implications for particle assembly and RNA packaging. *Protein Sci.* 2005, 14, 375–386. [CrossRef] [PubMed]
- 15. De Guzman, R.N.; Wu, Z.R.; Stalling, C.C.; Pappalardo, L.; Borer, P.N.; Summers, M.F. Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science* **1998**, 279, 384–388. [CrossRef]
- 16. Fossen, T.; Wray, V.; Bruns, K.; Rachmat, J.; Henklein, P.; Tessmer, U.; Maczurek, A.; Klinger, P.; Schubert, U. Solution structure of the human immunodeficiency virus type 1 p6 protein. *J. Biol. Chem.* **2005**, *280*, 42515–42527. [CrossRef]
- Hill, C.P.; Worthlake, D.; Bancroft, D.P.; Christensen, A.M.; Sundquist, W.I. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: Implications for membrane association and assembly. *Proc. Natl. Acad. Sci. USA* 1996, 93, 3099–3104. [CrossRef]
- 18. Fiorentini, S.; Marini, E.; Caracciolo, S.; Caruso, A. Functions of the HIV-1 matrix protein p17. New Microbiol. 2006, 29, 1–10.
- 19. Massiah, M.A.; Starich, M.R.; Paschall, C.; Summers, M.F.; Christensen, A.M.; Sundquist, W.I. Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. *J. Mol. Biol.* **1994**, 244, 198–223. [CrossRef]
- Forster, M.J.; Mulloy, B.; Nermut, M.V. Molecular modelling study of HIV p17gag (MA) protein shell utilising data from electron microscopy and X-ray crystallography. J. Mol. Biol. 2000, 298, 841–857. [CrossRef]
- Ohori, Y.; Okazaki, H.; Watanabe, S.; Tochio, N.; Arai, M.; Kigawa, T.; Nishimura, C. Flexible and rigid structures in HIV-1 p17 matrix protein monitored by relaxation and amide proton exchange with NMR. *Biochim. Biophys. Acta* 2014, 1844, 520–526. [CrossRef] [PubMed]
- 22. Boyd, P.S.; Brown, J.B.; Brown, J.D.; Catazaro, J.; Chaudry, I.; Ding, P.; Dong, X.; Marchant, J.; O'Hern, C.T.; Singh, K.; et al. NMR studies of retroviral genome packaging. *Viruses* 2020, 12, 1115. [CrossRef] [PubMed]
- 23. Lerner, G.; Weaver, N.; Anokhin, B.; Spearman, P. Advances in HIV-1 assembly. Viruses 2022, 14, 478. [CrossRef] [PubMed]
- 24. Lingappa, J.R.; Reed, J.C.; Tanaka, M.; Chutiraka, K.; Robinson, B.A. How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle. *Virus Res.* **2014**, *139*, 89–107. [CrossRef] [PubMed]
- 25. Qu, K.; Ke, Z.; Zila, V.; Anders-Össwein, M.; Glass, B.; Mücksch, F.; Müller, R.; Schultz, C.; Müller, B.; Kräusslich, H.-G.; et al. Maturation of the matrix and viral membrane of HIV-1. *Science* **2021**, *373*, 700–704. [CrossRef] [PubMed]
- 26. Monje-Galvan, V.; Voth, G.A. Binding mechanism of the matrix domain of HIV-1 *gag* on lipid membranes. *eLife* **2020**, *9*, e58621. [CrossRef]
- 27. Ono, A.; Orenstein, J.M.; Freed, E.O. Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J. Virol.* **2000**, *74*, 2855–2866. [CrossRef]
- Eastep, G.N.; Ghanam, R.H.; Green, T.J.; Saad, J.S. Structural characterization of HIV-1 matrix mutants implicated in envelope incorporation. J. Biol. Chem. 2021, 296, 100321. [CrossRef]

- 29. Vlach, J.; Saad, J.S. Structural and molecular determinants of HIV-1 Gag binding to the plasma membrane. *Front. Microbiol.* **2015**, *6*, 232. [CrossRef]
- Bussienne, C.; Marquet, R.; Paillart, J.-C.; Bernacchi, S. Post-Translational modifications of retroviral HIV-1 Gag precursors: An overview of their biological role. Int. J. Mol. Sci. 2021, 22, 2871. [CrossRef]
- Resh, M.D. Trafficking and signaling by fatty-acylated and prenylated proteins. *Nat. Chem. Biol.* 2006, 2, 584–590. [CrossRef]
 [PubMed]
- Tang, C.; Loeliger, E.; Luncsford, P.; Kinde, I.; Beckett, D.; Summers, M.F. Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proc. Natl. Acad. Sci. USA* 2004, 101, 517–522. [CrossRef]
- Fledderman, E.L.; Fujii, K.; Ghanam, R.H.; Waki, K.; Prevelige, P.E.; Freed, E.O.; Saad, J.S. Myristate exposure in the HIV-1 matrix protein is modulated by pH. *Biochemistry* 2010, 49, 9551–9562. [CrossRef] [PubMed]
- Valentine, K.G.; Peterson, R.; Saad, J.S.; Summers, M.F.; Xu, X.; Ames, J.B.; Wand, A.J. Reverse micelle encapsulation of membrane anchored proteins for solution NMR studies. *Structure* 2010, 18, 9–16. [CrossRef] [PubMed]
- Sumner, C.; Ono, A. Relationship between HIV-1 Gag multimerization and membrane binding. *Viruses* 2022, 14, 622. [CrossRef] [PubMed]
- Wen, Y.; Feigenson, G.W.; Vogt, V.M.; Dick, R.A. Mechanisms of PI(4,5)P2 enrichment in HIV-1 viral membranes. J. Mol. Biol. 2020, 432, 5343–5364. [CrossRef]
- Resh, M.D. A myristoyl switch regulates membrane binding of HIV-1 Gag. Proc. Natl. Acad. Sci. USA 2004, 101, 417–418. [CrossRef]
- Mariani, C.; Desdouits, M.; Favard, C.; Benaroch, P.; Muriaux, D.M. Role of Gag and lipids during HIV-1 assembly in CD4+ T cells and macrophages. *Front. Microbiol.* 2014, *5*, 312. [CrossRef]
- 39. Li, Y.; Frederick, K.M.; Haverland, N.A.; Ciborowski, P.; Belshan, M. Investigation of the HIV-1 Matrix interactome during virus replication. *Proteom. Clin. Appl.* **2016**, *10*, 156–163. [CrossRef]
- Ramdas, P.; Sahu, A.K.; Mishra, T.; Bhardwaj, V.; Chande, A. From entry to egress: Strategic exploitation of the cellular processes by HIV-1. Front. Microbiol. 2020, 11, 559792. [CrossRef]
- Bou-Nader, C.; Muecksch, F.; Brown, J.B.; Gordon, J.M.; York, A.; Peng, C.; Ghirlando, R.; Summers, M.F.; Bieniasz, P.D.; Zhang, J. HIV-1 matrix-tRNA complex structure reveals basis for host control of Gag localization. *Cell Host Microbe* 2021, 29, 1421–1436. [CrossRef] [PubMed]
- 42. De Francesco, M.A.; Caruso, A.; Fallacara, F.; Canaris, A.D.; Dima, F.; Poiesi, C.; Licenziati, S.; Corulli, M.; Martinelli, F.; Fiorentini, S.; et al. HIV p17 enhances lymphocyte proliferation and HIV-1 replication after binding to a human serum factor. *AIDS* **1998**, *12*, 245–252. [CrossRef] [PubMed]
- 43. De Francesco, M.A.; Baronio, M.; Fiorentini, S.; Signorini, C.; Bonfanti, C.; Poiesi, C.; Popovic, M.; Grassi, M.; Garrafa, E.; Bozzo, L.; et al. HIV-1 matrix protein p17 increases the production of proinflammatory cytokines and counteracts IL-4 activity by binding to a cellular receptor. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 9972–9987. [CrossRef] [PubMed]
- Caccuri, F.; D'Ursi, P.; Uggeri, M.; Bugatti, A.; Mazzuca, P.; Zani, A.; Filippini, F.; Salmonad, M.; Ribattie, D.; Slevin, M.; et al. Evolution toward beta common chain receptor usage links the matrix proteins of HIV-1 and its ancestors to human erythropoietin. *Proc. Natl. Acad. Sci. USA* 2021, 118, e2021366118. [CrossRef] [PubMed]
- 45. Yu, X.; Yuan, X.; Matsuda, Z.; Lee, T.-H.; Essex, M. The matrix protein of human immunodeficiency virus type I is required for incorporation of viral envelope protein into mature virions. *J. Virol.* **1992**, *66*, 4966–4971. [CrossRef] [PubMed]
- 46. Cosson, P. Direct interaction between the envelope and matrix proteins of HIV-1. EMBO J. 1996, 15, 5783–5788. [CrossRef]
- 47. Tedbury, P.R.; Novikova, M.; Ablan, S.D.; Freed, E.O. Biochemical evidence of a role for matrix trimerization in HIV-1 envelope glycoprotein incorporation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 182–190. [CrossRef]
- Troyano-Hernáez, P.; Reinosa, R.; Holguín, A. HIV capsid protein genetic diversity across HIV-1 variants and impact on new capsid-inhibitor Lenacapavir. *Front. Microbiol.* 2022, 13, 854974. [CrossRef]
- Maillard, P.V.; Zoete, V.; Michielin, O.; Trono, D. Homology-based identification of capsid determinants that protect HIV1 from human TRIM5α restriction. *J. Biol. Chem.* 2011, 286, 8128–8140. [CrossRef]
- 50. Ganser-Pornillos, B.K.; Cheng, A.; Yeager, M. Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell* 2007, 131, 70–79. [CrossRef]
- Koma, T.; Kotani, O.; Miyakawa, K.; Ryo, A.; Yokoyama, M.; Doi, N.; Adachi, A.; Sato, H.; Nomaguchia, M. Allosteric regulation of HIV-1 capsid structure for Gag assembly, virion production, and viral infectivity by a disordered interdomain linker. *J. Virol.* 2019, 93, e00381-19. [CrossRef] [PubMed]
- 52. Krebs, A.-S.; Mendonça, L.M.; Zhang, P. Structural analysis of retrovirus assembly and maturation. Viruses 2022, 14, 54. [CrossRef]
- 53. McFadden, W.M.; Snyder, A.A.; Kirby, K.A.; Tedbury, P.R.; Raj, M.; Wang, Z.; Sarafianos, S.G. Rotten to the core: Antivirals targeting the HIV-1 capsid core. *Retrovirology* **2021**, *18*, 41. [CrossRef] [PubMed]
- 54. Pornillos, O.; Ganser-Pornillos, B.K.; Kelly, B.N.; Hua, Y.; Whitby, F.G.; Stout, C.D.; Sundquist, W.I.; Hill, C.P.; Yeager, M. X-ray structures of the hexameric building block of the HIV capsid. *Cell* **2009**, *137*, 1282–1292. [CrossRef] [PubMed]
- Novikova, M.; Zhang, Y.; Freed, E.O.; Peng, K. Multiple roles of HIV-1 capsid during the virus replication cycle. *Virol. Sin.* 2019, 34, 119–134. [CrossRef] [PubMed]

- Mallery, D.L.; Marquez, C.L.; McEwan, W.A.; Dickson, C.F.; Jacques, D.A.; Anandapadamanaban, M.; Bichel, K.; Towers, G.J.; Saiardi, A.; Bocking, T.; et al. IP6 is an HIV pocket factor that prevents capsid collapse and promotes DNA synthesis. *eLife* 2018, 7, e35335. [CrossRef] [PubMed]
- 57. Renner, N.; Mallery, D.L.; Faysal, K.M.R.; Peng, W.; Jacques, D.A.; Bocking, T.; James, L.C. A lysine ring in HIV capsid pores coordinates IP6 to drive mature capsid assembly. *PLoS Pathog.* **2021**, *17*, e1009164. [CrossRef]
- Zhao, G.; Perilla, J.R.; Yufenyuy, E.L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A.M.; Schulten, K.; Aiken, C.; et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* 2013, 497, 643–646. [CrossRef]
- 59. Mattei, S.; Glass, B.; Hagen, W.J.; Krausslich, H.G.; Briggs, J.A. The structure and flexibility of conical HIV-1 capsids determined within intact virions. *Science* 2016, 354, 1434–1437. [CrossRef]
- 60. Gres, A.T.; Kirby, K.A.; KewalRamani, V.N.; Tanner, J.J.; Pornillos, O.; Sarafianos, S.G. X-ray structures of native HIV-1 capsid protein reveal conformational variability. *Science* 2015, 349, 99–103. [CrossRef]
- 61. Selyutina, A.; Bulnes-Ramos, A.; Diaz-Griffero, F. Binding of host factors to stabilized HIV-1 capsid tubes. *Virology* **2018**, 523, 1–5. [CrossRef] [PubMed]
- 62. AlBurtamani, N.; Paul, A.; Fassati, A. The role of capsid in the early steps of HIV-1 infection: New insights into the core of the matter. *Viruses* **2021**, *13*, 1161. [CrossRef] [PubMed]
- 63. Saito, A.; Yamashita, M. HIV-1 capsid variability: Viral exploitation and evasion of capsid-binding molecules. *Retrovirology* **2021**, *18*, 32. [CrossRef] [PubMed]
- 64. Cui, M.; Huang, X.; Luo, X.; Briggs, J.M.; Ji, R.; Chen, K.; Shen, J.; Jiang, H. Molecular docking and 3D-QSAR studies on Gag peptide analogue inhibitors interacting with human cyclophilin A. *J. Med. Chem.* **2002**, *45*, 5249–5259. [CrossRef]
- 65. Takemura, T.; Murakami, T. Functional constraints on HIV-1 capsids: Their impact on the viral immune escape potency. *Front. Microbiol.* **2012**, *3*, 369. [CrossRef] [PubMed]
- 66. Guedán, A.; Donaldson, C.D.; Caroe, E.R.; Cosnefroy, O.; Taylor, I.A.; Bishop, K.N. HIV-1 requires capsid remodelling at the nuclear pore for nuclear entry and integration. *PLoS Pathog.* **2021**, *17*, e1009484. [CrossRef]
- 67. Freed, E.O. HIV-1 assembly, release and maturation. Nat. Rev. Microbiol. 2015, 13, 484–496. [CrossRef]
- 68. Liu, C.; Perilla, J.R.; Ning, J.; Lu, M.; Hou, G.; Ramalho, R.; Himes, B.A.; Zhao, G.; Bedwell, G.J.; Byeon, I.-J.; et al. Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat. Commun.* **2016**, *7*, 10714. [CrossRef]
- 69. Xie, L.; Chen, L.; Zhong, C.; Yu, T.; Ju, Z.; Wang, M.; Xiong, H.; Zeng, Y.; Wang, J.; Hu, H.; et al. MxB impedes the NUP358-mediated HIV-1 pre-integration complex nuclear import and viral replication cooperatively with CPSF6. *Retrovirology* **2020**, *17*, 16. [CrossRef]
- Burdick, R.C.; Li, C.; Munshi, M.; Rawson, J.M.O.; Nagashimac, K.; Hub, W.-S.; Pathak, V.K. HIV-1 uncoats in the nucleus near sites of integration. *Proc. Natl. Acad. Sci. USA* 2020, 117, 5486–5493. [CrossRef]
- López, C.S.; Tsagli, S.M.; Sloan, R.; Eccles, J.; Barklis, E. Second site reversion of a mutation near the amino terminus of the HIV-1 capsid protein. *Virology* 2013, 447, 95–103. [CrossRef] [PubMed]
- 72. McDermott, J.; Farrell, L.; Ross, R.; Barklis, E. Structural analysis of human immunodeficiency virus type 1 Gag protein interactions, using cysteine-specific reagents. *J. Virol.* **1996**, *70*, 5106–5114. [CrossRef] [PubMed]
- 73. Yang, H.; Talledge, N.; Arndt, W.G.; Zhang, W.; Mansky, L.M. Human immunodeficiency virus type 2 capsid protein mutagenesis defines the determinants for Gag-Gag interactions. *bioRxiv* 2022. [CrossRef]
- 74. Gamble, T.R.; Yoo, S.; Vajdos, F.F.; von Schwedler, U.K.; Worthylake, D.K.; Wang, H.; McCutcheon, J.P.; Sundquist, W.I.; Hill, C.P. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **1997**, *278*, 849–853. [CrossRef] [PubMed]
- Tanaka, M.; Robinson, B.A.; Chutiraka, K.; Geary, C.D.; Reed, J.C.; Lingappa, J.R. Mutations of conserved residues in the major homology region arrest assembling HIV-1 Gag as a membrane-targeted intermediate containing genomic RNA and cellular proteins. J. Virol. 2016, 90, 1944–1963. [CrossRef] [PubMed]
- Ivanov, D.; Stone, J.R.; Maki, J.L.; Collins, T.; Wagner, G. Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. *Mol. Cell* 2005, 17, 137–143. [CrossRef]
- Nandwani, N.; Surana, P.; Negi, H.; Mascarenhas, N.M.; Udgaonkar, J.B.; Das, R.; Gosavi, S. A five-residue motif for the design of domain swapping in proteins. *Nat. Commun.* 2019, 10, 452. [CrossRef]
- 78. Ivanov, D.; Tsodikov, O.V.; Kasanov, J.; Ellenberger, T.; Wagner, G.; Collins, T. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. *Proc. Nati. Acad. Sci. USA* **2007**, *104*, 4353–4358. [CrossRef]
- Bocanegra, R.; Fuertes, M.A.; Rodriguez-Huete, A.; Neira, J.L.; Mateu, M.G. Biophysical analysis of the MHR motif in folding and domain swapping of the HIV capsid protein C-terminal domain. *Biophysical. J.* 2015, 108, 338–349. [CrossRef]
- Segal-Maurer, S.; DeJesus, E.; Stellbrink, H.-J.; Castagna, A.; Richmond, G.J.; Sinclair, G.I.; Siripassorn, K.; Ruane, P.J.; Berhe, M.; Wang, H.; et al. Capsid inhibition with Lenacapavir in multidrug-resistant HIV-1 infection. N. Engl. J. Med. 2022, 386, 1793–1803. [CrossRef]
- 81. Vernekar, S.K.; Sahani, R.L.; Casey, M.C.; Kankanala, J.; Wang, L.; Kirby, K.A.; Du, H.; Zhang, H.; Tedbury, P.R.; Xie, J.; et al. Toward structurally novel and metabolically stable HIV-1 capsid-targeting small molecules. *Viruses* **2020**, *12*, 452. [CrossRef] [PubMed]
- Jiang, K.; Humbert, M.; Sriram, K.K.; Rouzina, L.; Mely, Y.; Westerlund, F. The HIV-1 nucleocapsid chaperone protein forms locally compacted globules on long double-stranded DNA. *Nucleic Acids Res.* 2021, 49, 4550–4563. [CrossRef] [PubMed]
- Wang, Y.; Guo, C.; Wang, X.; Xu, L.; Li, R.; Wang, J. The Zinc content of HIV-1 NCp7 affects its selectivity for packaging signal and affinity for stem-loop 3. *Viruses* 2021, 13, 1922. [CrossRef]
- Darlix, J.L.; Godet, J.; Ivanyi-Nagy, R.; Fossé, P.; Mauffret, O.; Mély, Y. Flexible nature and specific functions of the HIV-1 nucleocapsid protein. J. Mol. Biol. 2011, 410, 565–581. [CrossRef]

- Mori, M.; Kovalenko, L.; Lyonnais, S.; Antaki, D.; Torbett, B.E.; Botta, M.; Mirambeau, G.; Mély, Y. Nucleocapsid protein: A desirable target for future therapies against HIV-1. *Cur. Top. Microbiol. Immunol.* 2015, 389, 53–92.
- 86. Summers, M.F.; Henderson, L.E.; Chance, M.R.; Bess, J.W., Jr.; South, T.L.; Blake, P.R.; Sagi, I.; Perez-Alvarado, G.; Sowder, R.C., III; Hare, D.R.; et al. Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. *Protein Sci.* 1992, 1, 563–574. [CrossRef] [PubMed]
- Muriaux, D.; Darlix, J.L. Properties and functions of the nucleocapsid protein in virus assembly. *RNA Biol.* 2010, 7, 744–753. [CrossRef]
- Morellet, N.; de Rocquigny, H.; Mély, Y.; Jullian, N.; Déméné, H.; Ottmann, M.; Gérard, D.; Darlix, J.L.; Fournie-Zaluski, M.C.; Roques, B.P. Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. J. Mol. Biol. 1994, 235, 287–301. [CrossRef]
- 89. Thomas, J.A.; Gorelick, R.J. Nucleocapsid protein function in early infection processes. Virus Res. 2008, 134, 39–63. [CrossRef]
- 90. Vivanco-Lira, A.; Nieto-Saucedo, J.R. In silico model of infection of CD4(+) T-cell by a human immunodeficiency type 1 virus, and a mini-review of its molecular pathophysiology. *arXiv* 2021, arXiv:2102.03876.
- Godet, J.; Boudier, C.; Humbert, N.; Ivanyi-Nagy, R.; Darlix, J.-L.; Mély, Y. Comparative nucleic acid chaperone properties of the nucleocapsid protein NCp7 and Tat protein of HIV-1. *Virus Res.* 2012, *169*, 349–360. [CrossRef] [PubMed]
- Krishna, S.S.; Majumdar, I.; Grishin, N.V. Structural classification of zinc fingers. Nucleic Acids Res. 2003, 31, 532–550. [CrossRef] [PubMed]
- 93. Bell, N.M.; Kenyon, J.C.; Balasubramanian, S.; Lever, A.M.L. Comparative structural effects of HIV-1 Gag and nucleocapsid proteins in binding to and unwinding of the viral RNA packaging signal. *Biochemistry* **2012**, *21*, 3162–3169. [CrossRef] [PubMed]
- Amarasinghe, G.K.; Zhou, J.; Miskimon, M.; Chancellor, K.J.; McDonald, J.A.; Matthews, A.G.; Miller, R.R.; Rouse, M.D.; Summers, M.F. Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein structural studies and implications for genome recognition. J. Mol. Biol. 2001, 314, 961–970. [CrossRef]
- 95. Bernacchi, S. Visualization of retroviral Gag-genomic RNA cellular interactions leading to genome encapsidation and viral assembly: An overview. *Viruses* **2022**, *14*, 324. [CrossRef]
- 96. Lu, K.; Heng, X.; Summers, F. Structural determinants and mechanism of HIV-1 genome packaging. J. Mol. Biol. 2011, 410, 609–633. [CrossRef]
- 97. Gorelick, R.J.; Henderson, L.E.; Hanser, J.P.; Rein, A. Point mutants of Moloney murine leukemia virus that fail to package viral RNA: Evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Nat. Acad. Sci. USA* **1988**, *85*, 8420–8424. [CrossRef]
- Guo, C.; Yao, X.; Wang, K.; Wanga, J.; Wang, Y. Comparison of HIV-1 Gag and NCp7 in their selectivity for package signal, affinity for stem-loop 3, and Zn²⁺ content. *Biochimie* 2020, 179, 135–145. [CrossRef]
- 99. Dib-Hajj, F.; Khan, R.; Giedroc, D.P. Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. *Protein Sci.* **1993**, *2*, 231–243. [CrossRef]
- 100. Levin, J.G.; Mitra, M.; Mascarenhas, A.; Musier-Forsyth, K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* **2010**, *7*, 754–774. [CrossRef]
- 101. Butovskaya, E.; Soldà, P.; Scalabrin, M.; Nadai, M.; Richter, S.N. HIV-1 Nucleocapsid protein unfolds stable RNA G-quadruplexes in the viral genome and is inhibited by G-quadruplex ligands. *ACS Infect. Dis.* **2019**, *5*, 2127–2135. [CrossRef] [PubMed]
- 102. Klingler, J.; Anton, H.; Réal, E.; Zeiger, M.; Moog, C.; Mély, Y.; Boutant, E. How HIV-1 Gag manipulates its host cell proteins: A focus on interactors of the nucleocapsid domain. *Viruses* 2020, *12*, 888. [CrossRef] [PubMed]
- Solbak, S.M.O.; Reksten, T.R.; Hahn, F.; Wray, V.; Henklein, P.; Henklein, P.; Halskau, O.; Schubert, U.; Fossen, T. HIV-1 p6—A structured to flexible multifunctional membrane-interacting protein. *Biochim. Biophys. Acta* 2013, 1828, 816–823. [CrossRef] [PubMed]
- 104. Wanaguru, M.; Bishop, K.N. HIV-1 Gag recruits oligomeric Vpr via two binding sites in p6, but both mature p6 and Vpr are rapidly lost upon target cell entry. *J. Virol.* **2021**, *95*, e00554-21. [CrossRef]
- Usami, Y.; Popov, S.; Popova, E.; Inoue, M.; Weissenhorn, W.; Gottlinger, H.G. The ESCRT pathway and HIV-1 budding. *Biochem. Soc. Trans.* 2009, 37, 181–184. [CrossRef]
- Zhai, Q.; Landesman, M.B.; Chung, H.-Y.; Dierkers, A.; Jeffries, C.M.; Trewhella, J.; Hill, C.P.; Sundquist, W.I. Activation of the retroviral budding factor ALIX. J. Virol. 2011, 85, 9222–9226. [CrossRef]
- 107. Dussupt, V.; Javid, M.P.; Abou-Jaoude, G.; Jadwin, J.A.; de La Cruz, J.; Nagashima, K.; Bouamr, F. The nucleocapsid region of HIV-1 Gag cooperates with the PTAP and LYPXnL late domains to recruit the cellular machinery necessary for viral budding. *PLoS Pathog.* 2009, 5, e1000339. [CrossRef]
- Im, Y.J.; Kuo, L.; Ren, X.; Burgos, P.V.; Zhao, X.Z.; Liu, F.; Burke, T.R., Jr.; Bonifacino, J.S.; Freed, E.O.; Hurley, J.H. Crystallographic and functional analysis of the ESCRT-I/HIV-1 Gag PTAP interaction. *Structure* 2010, 18, 1536–1547. [CrossRef]
- Sundquist, W.I.; Kräusslich, H.-G. HIV-1 assembly, budding and maturation. Cold Spring Harb. Perspect. Med. 2012, 2, a006924.
 [CrossRef]
- Dubois, N.; Khoo, K.K.; Ghossein, S.; Seissler, S.; Wolff, P.; McKinstry, W.J.; Mak, J.; Paillart, J.-C.; Marquet, R.; Bernacchi, S. The C-terminal p6 domain of the HIV-1 Pr55^{Gag} precursor is required for specific binding to the genomic RNA. *RNA Biol.* 2018, 15, 923–936. [CrossRef]
- 111. Yu, F.-H.; Huang, K.-J.; Wang, C.-T. Amino acid substitutions at the HIV-1 transframe region significantly impair virus infectivity. *PLoS ONE* **2022**, *17*, e0262477. [CrossRef]

- 112. Bredbenner, K.; Simon, S.M. Vpr co-assembles with Gag during HIV-1 assembly. bioRxiv 2020. [CrossRef]
- 113. Morellet, N.; Bouaziz, S.; Petitjean, P.; Roques, B.P. NMR structure of the HIV-1 regulatory protein VPR. *J. Mol. Biol.* 2003, 327, 215–227. [CrossRef]
- 114. Fritz, J.V.; Didier, P.; Clamme, J.-P.; Schaub, E.; Muriaux, D.; Cabanne, C.; Morellet, N.; Bouaziz, S.; Darlix, J.-L.; Mély, Y.; et al. Direct Vpr-Vpr interaction in cells monitored by two photon fluorescence correlation spectroscopy and fluorescence lifetime imaging. *Retrovirology* 2008, 5, 87. [CrossRef]
- 115. Kogan, M.; Rapport, J. HIV-1 accessory protein Vpr: Relevance in the pathogenesis of HIV and potential for therapeutic intervention. *Retrovirology* **2011**, *8*, 25. [CrossRef]
- 116. Jackman, M. G2/M Transition. In Encyclopedia of Cancer; Schwab, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2011.
- 117. Fabryova, H.; Strebel, K. Vpr and its cellular interaction partners: Are we there yet? Cells 2019, 8, 1310. [CrossRef]
- 118. Müller, B.; Patschinsky, T.; Kräusslich, H.-G. The late-domain-containing protein p6 is the predominant phosphoprotein of human immunodeficiency virus type 1 particles. *J. Virol.* **2002**, *76*, 1015–1024. [CrossRef]
- Radestock, B.; Burk, R.; Müller, B.; Kräusslich, H.-G. Re-visiting the functional Relevance of the highly conserved Serine 40 Residue within HIV-1 p6^{Gag}. *Retrovirology* 2014, 11, 114. [CrossRef]
- 120. Votteler, J.; Neumann, L.; Hahn, S.; Hahn, F.; Rauch, P.; Schmidt, K.; Studtrucker, N.; Solbak, S.M.O.; Fossen, T.; Henklein, P.; et al. Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly. *Retrovirology* 2011, *8*, 11. [CrossRef]
- 121. Watanabe, S.M.; Chen, M.-H.; Khan, M.; Ehrlich, L.; Kemal, K.S.; Weiser, B.; Shi, B.; Chen, C.; Powell, M.; Anastos, K.; et al. The S40 residue in HIV-1 Gag p6 impacts local and distal budding determinants, revealing additional late domain activities. *Retrovirology* 2013, 10, 143. [CrossRef]
- 122. Mendonça, L.; Sun, D.; Ning, J.; Liu, J.; Kotecha, A.; Olek, M.; Frosio, T.; Fu, X.; Himes, B.A.; Kleinpeter, A.B.; et al. CryoET structures of immature HIV Gag reveal six-helix bundle. *Commun. Biol.* **2021**, *4*, 481. [CrossRef]
- 123. Briggs, J.A.G.; Riches, J.D.; Glass, B.; Bartonova, V.; Zanetti, G.; Kräusslich, H.G. Structure and assembly of immature HIV. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11090–11095. [CrossRef]
- 124. Datta, S.A.K.; Temeselew, L.G.; Crist, R.M.; Soheilian, F.; Kamata, A.; Mirro, J.; Harvin, D.; Nagashima, K.; Cachau, R.E.; Rein, A. On the role of the SP1 domain in HIV-1 particle assembly: A molecular switch? *J. Virol.* **2011**, *85*, 4111–4121. [CrossRef]
- 125. Tan, A.; Pak, A.; Morado, D.R.; Voth, G.A.; Briggs, J.A.G. Immature HIV-1 assembles from Gag dimers leaving partial hexamers at lattice edges as potential substrates for proteolytic maturation. *Proc. Natl. Acad. Sci. USA* 2021, 118, e2020054118. [CrossRef]
- 126. Wright, E.R.; Schooler, J.B.; Ding, H.J.; Kieffer, C.; Fillmore, C.; Sundquist, W.I.; Jensen, G.J. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J.* **2007**, *26*, 2218–2226. [CrossRef]
- 127. Wiegers, K.; Rutter, G.; Kottler, H.; Tessmer, U.; Hohenberg, H.; Krausslich, H.G. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J. Virol.* **1998**, *72*, 2846–2854. [CrossRef]
- Gross, I.; Hohenberg, H.; Wilk, T.; Wiegers, K.; Grättinger, M.; Müller, B.; Fuller, S.; Kräusslich, H.-G. A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* 2000, 19, 103–113. [CrossRef]
- 129. Kucharska, I.; Ding, P.; Zadrozny, K.K.; Dick, R.A.; Summers, M.F.; Ganser-Pornillos, B.K.; Pornillos, O. Biochemical reconstitution of HIV-1 assembly and maturation. J. Virol. 2020, 94, e01844-19. [CrossRef]
- 130. Adamson, C.S.; Sakalian, M.; Salzwedel, K.; Freed, E.O. Polymorphisms in Gag spacer peptide 1 confer varying levels of resistance to the HIV-1 maturation inhibitor bevirimat. *Retrovirology* **2010**, *7*, 36. [CrossRef]
- 131. Waki, K.; Durell, S.R.; Soheilian, F.; Nagashima, K.; Butler, S.L.; Freed, E.O. Structural and functional insights into the HIV-1 maturation inhibitor binding pocket. *PLoS Pathog.* **2012**, *8*, e1002997. [CrossRef]
- 132. Hill, M.K.; Bellamy-McIntyre, A.; Vella, L.J.; Campbell, S.M.; Marshall, J.A.; Tachedjian, G.; Mak, J. Alteration of the proline at position 7 of the HIV-1 spacer peptide p1 suppresses viral infectivity in a strain dependent manner. *Curr. HIV Res.* 2007, *5*, 69–78. [CrossRef] [PubMed]
- Hill, M.K.; Shehu-Xhilaga, M.; Crowe, S.M.; Mak, J. Proline residues within spacer peptide p1 are important for human immunodeficiency virus type 1 infectivity, protein processing, and genomic RNA dimer stability. J. Virol. 2002, 76, 11245–11253. [CrossRef] [PubMed]
- 134. Coren, L.V.; Thomas, J.A.; Chertova, E.; Sowder II, R.C.; Gagliardi, T.D.; Gorelick, R.J.; Ott, D.E. Mutational analysis of the C-terminal Gag cleavage sites in human immunodeficiency virus type 1. *J. Virol.* **2007**, *81*, 10047–10054. [CrossRef]