The p12 Domain Is Unstructured in a Murine Leukemia Virus p12-CA^N Gag Construct

Sampson K. Kyere, Prem Raj B. Joseph, Michael F. Summers*

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Maryland, Baltimore, Maryland, United States of America

Abstract

The Gag polyproteins of gammaretroviruses contain a conserved p12 domain between MA and CA that plays critical roles in virus assembly, reverse transcription and nuclear integration. Here we show using nuclear magnetic resonance, that p12 is unstructured in a Moloney murine leukemia virus (MMLV) Gag fragment that includes the N-terminal domain of CA (p12-CA^N). Furthermore, no long range interactions were observed between the domains, as has been previously predicted. Flexibility appears to be a common feature of Gag "late" domains required for virus release during budding. Residues near the N-terminus of CA^N that form a β-hairpin in the mature CA protein are unfolded in p12-CA^N, consistent with proposals that hairpin formation helps trigger capsid assembly.

Introduction

All retroviruses encode a polyprotein called Gag that serves as the major structural protein of the virus and is capable of assembling into virus-like particles in the absence of any other viral constituent. Gag proteins contain three major domains: an N-terminal matrix (MA) domain that regulates intracellular trafficking and membrane targeting, a capsid (CA) domain (that consists of N- and C-terminal subdomains) that promotes virus assembly and forms the capsid shell of the viral core during proteolytic maturation, and a nucleocapsid (NC) domain that is responsible for genome selection and encapsidation. Atomic level structures have been determined for several MA, CA and NC domains, and in general, the domain structures do not vary significantly among different genera of retroviruses [1,2].

Gag proteins contain additional polypeptide elements and spacers that are often critical for proper virus assembly. In particular, short, proline-rich segments have been identified that facilitate viral budding at late stages of the viral lifecycle, and mutations within these segments typically give rise to elongated or partially formed virus-like buds that are not released from the plasma membrane [3–11]. These "late domains" are typically located in disparate regions of the Gag proteins of different retroviruses [9]. The late budding activity of the human immunodeficiency virus type-1 (HIV-1) is mediated by a conserved PT/SAP segment that resides within the C-terminal p6 region of Gag [3,6]. In contrast, the late budding activity of the gamma-retroviruses is mediated by a conserved PPPY motif found within an 84 residue polypeptide called p12, which is located between the N-terminal MA and CA domains [12]. Disruption of the Moloney murine leukemia virus (MMLV) PPPY element gives rise to budding defects similar to those observed for HIV-1 PT/SAP mutants [12].

Proteolytic cleavage of the p12-CA junction is critical for viral infectivity and precedes cleavage of both the MA-p12 and CA-NC

sites [13]. The N-terminal residues of mature N-tropic murine leukemia virus (N-MLV) CA (identical with MMLV CA except for five amino acid substitutions) adopt a β-hairpin that makes intermolecular CA-CA contacts upon hexamer formation [14]. It is likely that cleavage of the p12-CA junction is required for both hairpin formation and capsid assembly, as appears to be the case for HIV-1 capsid assembly [15,16]. Cryo-electron microscopy (EM) images of immature MMLV virions revealed a low density zone of 25–60 Å between the MA and CA domains, suggesting that p12 is either unfolded or folded but highly mobile prior to proteolysis [17]. In addition to its role in virus assembly, the mature p12 peptide facilitates reverse transcription and the delivery of the pre-integration complex (PIC) to the nucleus for integration of the viral DNA into the host genome during the early phase of infection [18–23]. To further understand the structure and diverse roles of p12 as well as the cooperative nature of p12 and CA, we have characterized the dynamic properties and solution structure of a recombinant construct composed of p12 and the N-terminal domain of CA (CA^N).

Methods

The pNCA MMLV proviral plasmid [24] was used to subclone MMLV p12CA^N, with a C-terminal His-tag, into pET11a (Novagen) using NdeI and BamHI restriction sites, and subsequently transformed into BL21 codon plus RP (DE3) cells (Stratagene). Cells were grown in LB or M9 minimal media supplemented with 99.9 % enriched ^{15}N-ammonium chloride and/or 99.8 % enriched ^{13}C-glucose as the sole nitrogen and/or carbon sources. Protein expression was induced in shake flasks with 1 mM IPTG. The cells were harvested and lysed with a microfluidizer (Micorfluidics), clarified by centrifugation, and the target protein was purified to homogeneity using cobalt affinity (Talon), cation exchange and size exclusion chromatographies
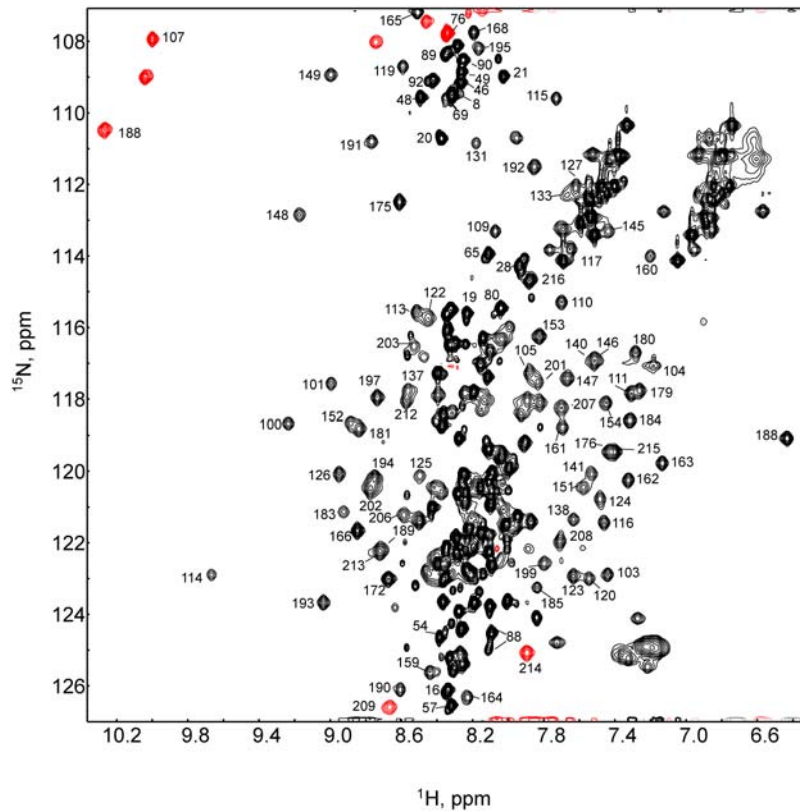
Figure 1. ¹H-¹⁵N correlation (HSQC) spectrum obtained for MMLV p12CA^N. Assignments are shown for signals in less-crowded regions of the spectrum. Red peaks represent signals folded in the ¹⁵N dimension.
doi:10.1371/journal.pone.0001902.g001

(Amersham). Fractions containing pure protein were concentrated using Centripreps (Amicon, MWCO = 3,500 Da) and Centricons (Amicon, MWCO = 3,500 Da). The mass of p12CA^N was confirmed by mass spectroscopy.

NMR data were collected at 35°C with a Bruker AVANCE 600 MHz spectrometer equipped with a cryoprobe using samples of 0.8–1.0 mM p12CA^N in buffer containing 50 mM sodium phosphate pH 5.5, 100 mM NaCl, 5 mM DTT and 10% D_2O. Backbone assignments were obtained using 2D ¹H,¹⁵N-HSQC, 3D ¹⁵N-edited TOCSY, and 3D HNCA and HN(CO)CA experiments [25–29]. 3D ¹⁵N-edited NOESY-HSQC, 4D ¹³C,¹⁵N-edited NOESY and 4D ¹³C,¹³C-edited NOESY experiments were used for side chain assignments and to identify long range intra- and interdomain [¹H,¹H] NOEs [30].

To gain insight into the dynamical properties of MMLV p12, {¹H}–¹⁵N steady-state heteronuclear NOE (XNOE) data were obtained for the backbone ¹⁵N nuclei [31]. The XNOE experiment consists of a Reference and NOE experiment collected in an interleaved mode. The reference experiment contained a recovery delay of 8 sec and the NOE experiment applied proton saturation during the last 3 sec of the 8 sec recovery delay [31]. The XNOE value for a given residue is derived from the intensity ratio (I/I_0) of its ¹⁵N/¹H correlation peak in the presence of 3 sec proton saturation (I) and in the absence of proton saturation (I_0). Peak intensities and the XNOE value were calculated using the HetNOE Analysis module in NMRVIEW. Errors were estimated from the baseline noise in the two spectra. All NMR data were processed using NMRPipe [32] and analyzed using NMRView [33]. Backbone chemical shifts were deposited in the Biological Magnetic Resonance Bank; Accession number 15672.

Results and Discussion

Initial examination of the 2D ¹H,¹⁵N-HSQC spectra for p12CA^N revealed a large subset of signals with poor chemical shift dispersion in the proton dimension, suggestive of a largely unstructured domain within the protein (Figure 1). Backbone resonance assignments established that these signals corresponded to residues in the p12 domain. The high degree of spectra overlap made backbone assignments difficult for this domain, yielding 60 % of the resonances assigned for p12. However, backbone assignments were made for 95 % of the non-proline residues of CA^N.

Backbone Cα NMR chemical shift indices (CSI, the difference between the observed chemical shift and shifts observed for random coil structures) provide information on local secondary structure [34,35], with downfield shifts (positive deviations) reflecting α-helical conformations and upfield shifts reflecting extended (β-strand) conformations. Zero or near zero CSIs were observed for all assigned residues of the p12 domain of p12CA^N, indicating that the p12 domain is unstructured, Figure 2. In contrast, most residues of the CA^N domain of p12CA^N exhibit positive deviations (Figure 2), in agreement with the crystal structure of MMLV CA^N hexamer [14].

Atomic level structures have been determined for the mature CA^N domains of several retroviruses, and in all cases, the first ~15 residues form a β-hairpin stabilized by a salt bridge between Pro 1 and the carboxyl group of an Asp residue subsequent to proteolytic induced maturation [14,15,36–39]. These residues are unstructured in an HIV-1 MA-CA Gag construct, indicating that β-hairpin formation occurs subsequent to proteolytic cleavage of the MA-CA junction [40]. Near-zero CSIs observed here for assigned N-terminal
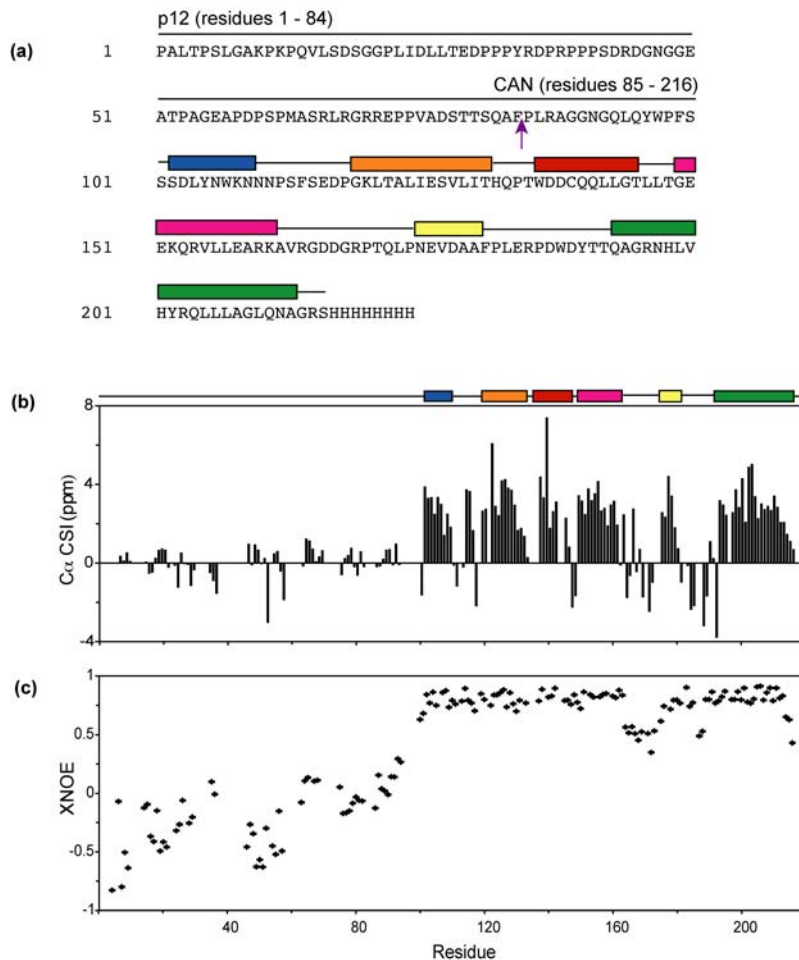
Figure 2. NMR chemical shift and relaxation data that identify regions of structure and mobility in p12CA^N. (A) Amino acid sequence of p12CA^N (arrow denotes proteolytic cleavage site). Residues of CA^N that adopt α-helical conformations in the N-MLV CA^N crystal structure are denoted by colored rectangles. **(B)** NMR chemical shift indices for the backbone Cα atoms of p12CA^N. Positive values denote helical regions, negative values denote regions of β-structure, and stretches of residues with near-zero values denote random coil conformations. For comparison, α-helical segments observed in the N-MLV CA^N crystal structure are aligned at the top of the panel. **(C)** ^{15}N{^{1}H} heteronuclear NOE (XNOE) data obtained for p12CA^N. Values near 1.0 reflect reduced molecular motion, and smaller or negative values reflect motion on a fast (ps-ns) timescale.
doi:10.1371/journal.pone.0001902.g002

residues of the CA^N domain of p12CA^N indicate that the β-hairpin observed in the X-ray structure of the mature MMLV CA^N domain is similarly unfolded in the immature p12CA^N construct.

To gain insight into the dynamical properties of N-MLV p12, {^{1}H}–^{15}N steady-state heteronuclear NOE (XNOE) data were obtained for the backbone ^{15}N atom [31]. XNOEs provide information regarding high-frequency (psec-nsec) backbone motions and are therefore useful for identifying regions of the protein with high internal mobility. The maximum theoretical XNOE value of 0.86 reflects highly restricted internal motion, whereas values smaller than ~0.7 are indicative of substantial internal motion [31]. XNOE values were measured for 44 residues of p12 and 106 of CA^N that were well resolved in the 2D ^{1}H,^{15}N-HSQC spectra [31]. The residues of p12 exhibited primarily negative or near zero XNOEs, indicative of a high degree of internal flexibility and consistent with the random coil CSI analysis. The high degree of mobility extends through the first 16 residues of the CA^N domain, confirming that the β-hairpin is unfolded in the immature p12CA^N construct. In contrast, residues within the α-helices of the CA^N domain exhibit relatively large XNOE values, consistent with a regularly folded tertiary structure.

Multi-dimensional ^{1}H-^{1}H NOE data (from 3D ^{15}N-edited NOESY-HSQC, 4D ^{13}C,^{15}N-edited NOESY and 4D ^{13}C,^{13}C-edited NOESY experiments; data not shown) were also obtained to probe for intra- and potential inter-domain contacts. Except for the N-terminal residues of the CA^N domain, which did not exhibit long-range interactions, the intra-CA^N NOEs observed in all the NOE data were fully consistent with the X-ray structure of the mature CA^N hexamer. No long-range intra- or inter-domain interactions were observed for residues of p12.

The combined CSI, XNOE and ^{1}H-^{1}H NOE NMR data indicate that the p12 domain of MMLV p12CA^N is conformationally labile. The folded regions of the HIV-1 MA and CA domains are also separated by a stretch of flexible residues (~20 in the MA-CA^N NMR structure) [40], and flexibility between these domains may be important for allowing the CA-CA interactions to adjust during virus assembly and maturation. Flexibility in the p12 domain may also be required for interactions with host cell proteins involved in late assembly processes, including viral release. The PTAP element of HIV-1 Gag, which interacts directly with TSG101 [9,41–45] (a component of the cellular protein sorting machinery), is also highly flexible in solution [9].

The inherent flexibility allows p6 to bind to a cleft on TSG101 UEV domain via an induced-fit mechanism [46,47].

In summary, the MMLV p12 domain of p12CA^N exhibits structural/dynamical properties similar to those observed for HIV-1 p6, despite that fact that p12 is twice as large as p6, is located in a very different region of Gag, and has additional functions during the early phase of viral replication.

References

1. Turner BG, Summers MF (1999) Structural Biology of HIV. J Mol Biology 285: 1–32.
2. Freed EO (1998) HIV-1 Gag Proteins: Diverse Functions in the Virus Life Cycle. Virology 251: 1–15.
3. Göttlinger HG, Dorfman T, Sodroski JG, Haseltine WA (1991) Effect of mutations affecting the p6 *gag* protein on human immunodeficiency virus particle release. Proceedings of the National Academy of Sciences USA 88: 3195–3199.
4. Le Blanc I, Prevost MC, Dokhelar MC, Rosenberg AR (2002) The PPPY motif of human T-cell leukemia virus type 1 Gag protein is required early in the budding process. J Virol 76: 10024–10029.
5. Wills JW, Cameron CE, Wilson CB, YXiang Y, Bennett RP, et al. (1994) An assembly domain of the Rous sarcoma virus Gag protein required late in budding. J Virol 68: 6605–6618.
6. Huang M, Orenstein JM, Martin MA, Freed EO (1995) p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. J Virol 96: 6810–6818.
7. Puffer BA, Parent LJ, Wills JW, Montelaro RC (1997) Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. J Virol 72: 10218–10221.
8. Yasuda J, Hunter E (1998) A proline-rich motif (PPPY) in the Gag polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. J Virol 72: 4095–4103.
9. Morita E, Sundquist WI (2004) Retrovirus Budding. Annual Review of Cell and Develpomental Biology 20: 395–425.
10. Yuan B, Fassati A, Yueh A, Goff SP (2000) Infectivity of Moloney murine leukemia virus defective in late assembly events is restored by late assembly domains of other retroviruses. J Virol 74: 7250–7260.
11. Wang HE, Norris KM, Mansky LM (2002) Analysis of bovine leukemia virus gag membrane targeting and late domain function. J Virol 76: 8485–8493.
12. Yuan B, Li X, Goff SP (1999) Mutations altering the moloney murine leukemia virus p12 Gag protein affect virion production and early events of the virus life cycle. Embo J 18: 4700–4710.
13. Oshima M, Muriaux D, Mirro J, Nagashima K, Dryden K, et al. (2004) Effects of blocking individual maturation cleavages in murine leukemia virus Gag. J Virol 78: 1411–1420.
14. Mortuza G, Haire LF, Stevens A, Smerdon SJ, Stoye JP, et al. (2004) High-resolution structure of a retroviral capsid hexameric amino-terminal domain. nature 431: 481–485.
15. Gitti RK, Lee BM, Walker J, Summers MF, Yoo S, et al. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. Science 273: 231–235.
16. von Schwedler UK, Stemmler TL, Klishko VY, Li S, Albertine KH, et al. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. Embo J 17: 1555–1568.
17. Yeager M, Wilson-Kubalek EM, Weiner SG, Brown PO, Rein A (1998) Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. Proc Natl Acad Sci USA 95: 7299–7304.
18. Yuan B, Fassati A, yueh A, Goff SP (2002) Characterization of Moloney Leukemia Virus p12 Mutants Blocked during Early events of Infection. Journal of Virology 76: 10801–10810.
19. Auerbach MR, Shu C, Kaplan A, Singh IR (2003) Functional characterization of a portion of the Moloney murine leukemia virus gag gene by genetic footprinting. Proc Natl Acad Sci USA 100: 11678–11683.
20. Bowerman B, Brown PO, Bishop JM, Varmus HE (1989) A nucleoprotein complex mediates the integration of retroviral DNA. Genes & Development 3: 469–478.
21. Yuan B, Li X, Goff SP (1999) Mutations altering the Moloney murine leukemia virus p12 Gag affect virion production and early events of the virus life cycle. The EMBO journal 18: 4700–4710.
22. Yueh A, Goff SP (2003) Phosphorylated serine residues and an arginine-rich domain of the moloney murine leukemia virus p12 protein are required for early events of viral infection. J Virol 77: 1820–1829.
23. Lee SK, Nagashima K, Hu WS (2005) Cooperative effect of gag proteins p12 and capsid during early events of murine leukemia virus replication. J Virol 79: 4159–4169.
24. Colicelli J, Goff SP (1988) Sequence and spacing requirements of a retrovirus integration site. J Mol Biol 199: 47–59.
25. Wüthrich K (1986) NMR of Proteins and Nucleic Acids. New York: John Wiley & Sons.
26. Grzesiek S, Anglister J, Bax A (1993) Correlation of backbone amide and aliphatic side-chain resonances in ^{13}C/^{15}N-enriched proteins by isotropic mixing of ^{13}C magnetization. J Magn Reson 101: 114–119.
27. Grzesiek S, Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. Journal of Magnetic Resonance 96: 432–440.
28. Grzesiek S, Bax A (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. Journal of Magnetic Resonance 99: 201–207.
29. Grzesiek S, Bax A (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. J Am Chem Soc 114: 6291–6293.
30. Kay LE, Clore GM, Bax A, Gronenborn AM (1990) Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 beta in solution. Science 249: 411–414.
31. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 28: 8972–8979.
32. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, et al. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6: 277–293.
33. Johnson BA, Blevins RA (1994) NMRview: a Computer Program for the Visualization and Analysis For NMR Data. J Biomol NMR 4: 603–614.
34. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR 5: 67–81.
35. Wishart DS, Sykes BD (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. J Biomol NMR 4: 171–180.
36. Gamble TR, Vajdos F, Yoo S, Worthylake DK, Houseweart SM, et al. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell 87: 1285–1294.
37. Momany C, Kovari LC, Prongay AJ, Keller W, Gitti RK, et al. (1996) Crystal structure of dimeric HIV-1 capsid protein. Nature Struct Biol 9: 763–770.
38. Khorasanizadeh S, Campos-Olivas R, Summers MF (1999) Solution structure of the capsid protein from the human T-cell leukemia virus Type-I. J Mol Biol 291: 491–505.
39. Campos-Olivas R, Newman JL, Summers MF (2000) Solution structure and dynamics of the Rous Sarcoma Virus capsid protein and comparison with capsid proteins of other retroviruses. J Mol Biol 296: 633–649.
40. Tang C, Ndassa Y, Summers MF (2002) Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nature Struct Biol 9: 537–543.
41. Demirov DG, Ono A, Orenstein JM, Freed EO (2002) Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. Proc Natl Acad Sci USA 99: 955–960.
42. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, et al. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell 107: 55–65.
43. Martin-Serrano J, Zang T, Bieniasz PD (2001) HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat Med 7: 1313–1319.
44. VerPlank L, Bouamr F, LaGrassa TJ, Agresta BE, Kikonyogo A, et al. (2001) Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55^{Gag}. Proc Natl Acad Sci USA 98: 7724–7729.
45. Demirov D, Freed EO (2004) Retrovirus budding. Virus Research 106: 87–102.
46. Pornillos O, Alam SL, Davis DR, Sundquist WI (2002) Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. Nat Struct Mol Biol 9: 812–817.
47. Sundquist WI, Schubert HL, Kelly BN, Hill G, Hill CP (2004) Ubiquitin recognition by the human TSG101 protein. Mol Cell 13: 783–789.

Author Contributions

Conceived and designed the experiments: MS SK PJ. Performed the experiments: SK PJ. Analyzed the data: MS SK PJ. Contributed reagents/materials/analysis tools: MS. Wrote the paper: MS SK.