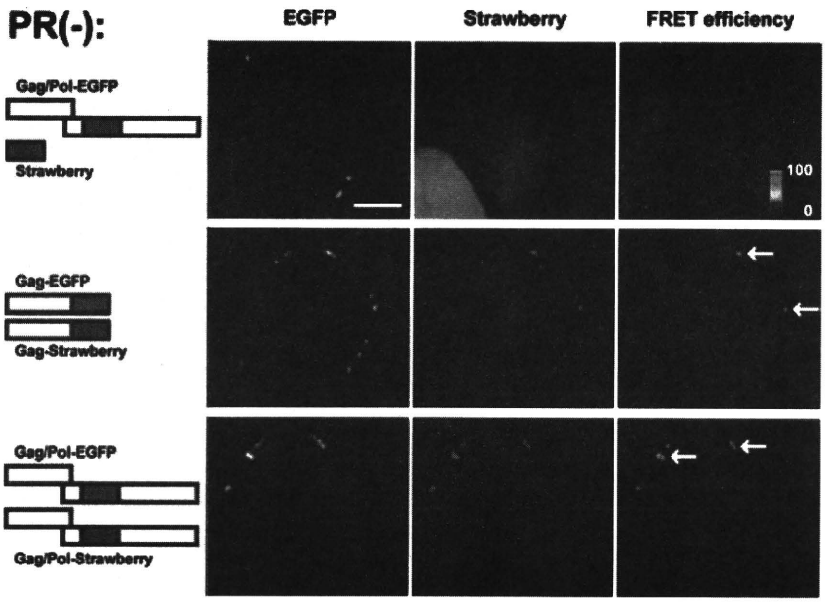


Fig. 4. Intracellular sites of GagPol dimerization. HeLa cells were cotransfected with Gag/Pol-EGFP and Gag/Pol-Strawberry constructs, both of which have an inactive form of PR. After 24 hr, cells were fixed and subjected to FRET analysis by confocal microscopy. A combination of Gag-EGFP and Gag-Strawberry constructs and a combination of Gag/Pol-EGFP construct and Strawberry (not fused to any other protein) plasmid were used as positive and negative controls, respectively. FRET efficiencies were calculated from images with three filter combinations. Arrows indicate FRET-positive puncta.

takes place at the plasma membrane during co-assembly of Gag and GagPol.

GagPol dimerization occurs predominantly at the plasma membrane

To define the sites of GagPol dimerization, we inserted EGFP and Strawberry fluorescent protein into the PR/RT junction of GagPol in the context of HIV molecular clone containing inactive PR (referred to as Gag/Pol-EGFP and Gag/Pol-Strawberry, respectively), and carried out FRET assays (Fig. 4). The Gags fused with EGFP and Strawberry at its C-termini (referred to as Gag-EGFP and Gag-Strawberry, respectively) were used as positive controls. A combination of the Gag-EGFP and Gag-Strawberry constructs exhibited robust FRET at the puncta on the plasma membrane, consistent with previous FRET analysis of Gag assembly (30). In contrast, no FRET was detected when Gag/Pol-EGFP was co-expressed with cytosolic Strawberry fluorescent protein. When FRET analysis was carried out using a combination of the Gag/Pol-EGFP and Gag/Pol-Strawberry constructs, FRET signals were only detected at the puncta on the plasma membrane (Fig. 4). To our knowledge, this is the first demonstration of GagPol-GagPol interactions. Together, these data suggest that GagPol dimerization occurs predominantly at the plasma membrane and not in the cytoplasm.

Alteration of the Gag-to-GagPol ratio has significant impacts on intracellular localization of p17/MA and viral particle production

Retroviral Gag and GagPol proteins are generated at a ratio of 10–20:1 (1). Previous studies have shown that the Gag-to-GagPol ratio has impacts on intracellular processing and viral particle production when PR is active (13). However, the localization of the processing products produced by an altered Gag-to-GagPol ratio has not been studied. We first confirmed that: (i) cotransfection of Gag-FLAG and GagPol-HA at a Gag-to-GagPol ratio of 1:0.1 produced similar amounts of virus particles to those produced by Gag-FLAG/Pol-HA; but ii) cotransfection at a Gag-to-GagPol ratio of 1:1 showed little or no particle production, although in both experiments the amount of DNA for Gag-FLAG was constant (Fig. 5a). These findings were further confirmed by electron microscopy (data not shown).

Next, we examined the intracellular distribution of the processing products by confocal microscopy (Fig. 5b) and, based on antigen distribution, the observed cells were sorted into three categories (Fig. 5c). As expected, when cells were cotransfected with Gag-FLAG and GagPol-HA at a Gag-to-GagPol DNA ratio of 1:0.1, both p17/MA and GagPol-HA (plus IN-HA) accumulated at the plasma membrane (97.8 and 79.8% of antigen-positive cells, respectively), these being similar distributions to those found in Gag-FLAG/Pol-HA-transfected cells (100 and

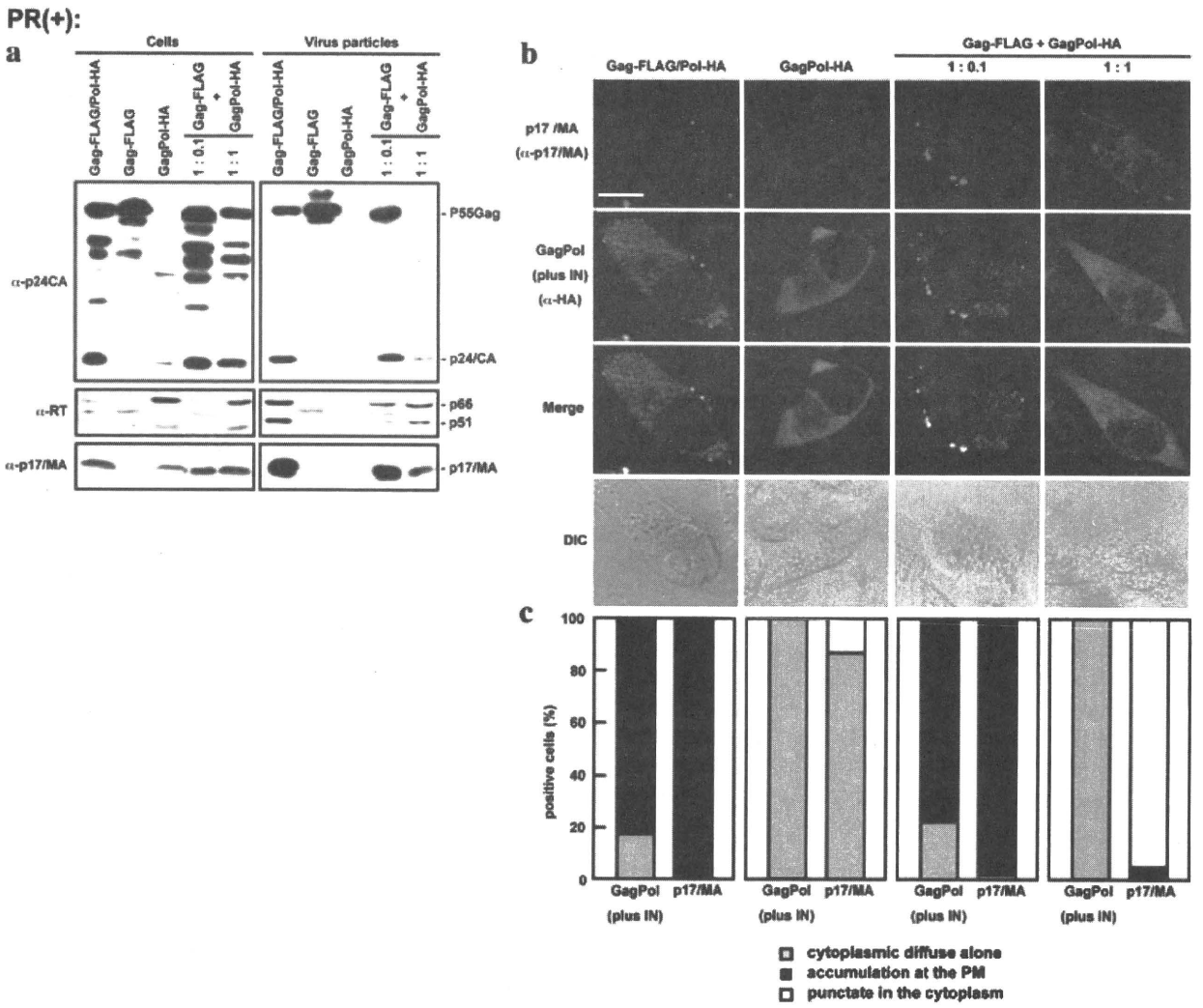


Fig. 5. Intracellular localization of mature MA and viral particle production in overexpression of GagPol containing active PR. HeLa cells were singly transfected with the Gag-FLAG/Pol-HA, GagPol-HA (both of which contain active PR), and Gag-FLAG construct, or doubly transfected with a combination of the Gag-FLAG and GagPol-HA constructs at a Gag-to-GagPol DNA ratio of 1:0.1 and 1:1. Total DNA amounts were normalized to 1.5 μg with pUC plasmid. (a) Viral particle production. Equivalent volumes of cell samples and of virus particle samples were subjected to Western blotting using anti-HIV-1 p24/CA, anti-RT, and p17/MA antibodies. (b) Intracellular localization of the mature form of p17/MA and GagPol products. At 24 hr post-transfection, cells were fixed and immunostained with anti-p17/MA (for mature MA) and anti-HA (for GagPol and IN) antibodies. Nuclei were stained with TOPRO-3. All micrographs are shown at the same magnification. Bar, 10 μm. DIC, differential interference contrast. (c) Semiquantification of localization of GagPol (plus IN), and mature MA. Approximately 50–80 antigen-positive cells were observed in each sample and the number of cells with each pattern of antigen distribution was counted. PM, plasma membrane.

84.5% of antigen-positive cells, respectively). In contrast, both antigens were diffusely distributed in the cytoplasm in cells expressing GagPol-HA alone (82.4 and 100% of antigen-positive cells, respectively). When the Gag-to-GagPol ratio was altered from 1:0.1 to 1:1, a fraction of the p17/MA population was uniformly distributed at the plasma membrane and a large population of p17/MA was observed in the cytoplasm along with punctate accumulation (95.4% of antigen-positive cells). The HA signals were also diffusely distributed in the cytoplasm (99.8% of antigen-positive cells). These patterns of antigen distribution were similarly observed at 12 hr post-transfection (data not shown). These data indicate that when PR is active, the ratio of Gag to GagPol has a significant impact on intracellular localization of their processing products.

The Gag-to-GagPol ratio has impacts on particle morphology when PR is inactive

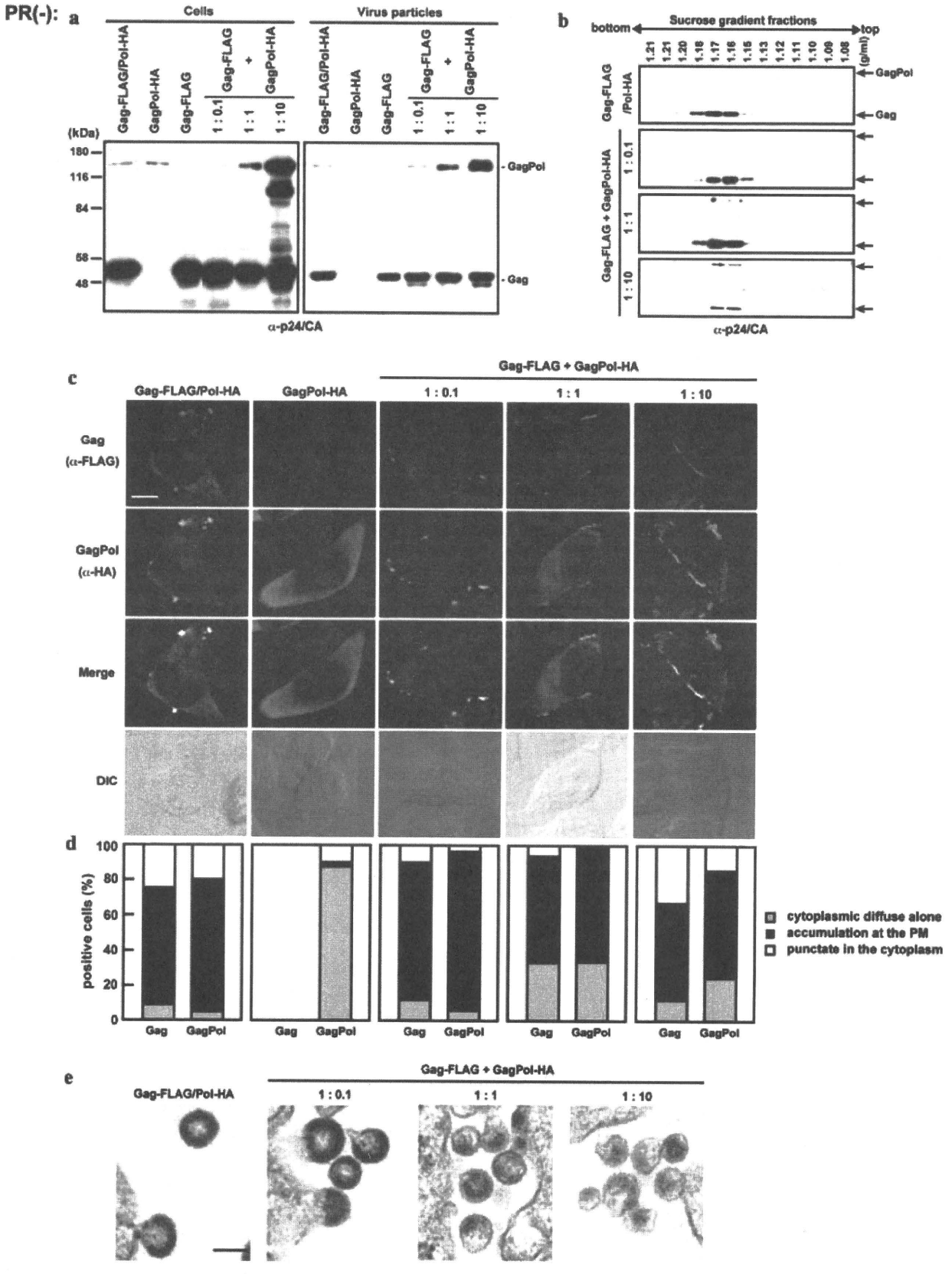
Previous studies have suggested that when PR is inactive, the Gag-to-GagPol ratio is less critical for viral particle production, although it is still critical for virion RNA dimerization (13). We examined the impacts of the Gag-to-GagPol ratio on viral particle production using an HIV-1 molecular clone containing inactive PR. HeLa cells were cotransfected with Gag-FLAG and GagPol-HA constructs at a Gag-to-GagPol DNA ratio of 1:0.1, 1:1, and 1:10 (with a fixed amount of DNA for Gag-FLAG). Western blotting using anti-p24/CA antibody confirmed that the level of expression of Gag was constant whereas that of GagPol was increased in a DNA dose-dependent manner (Fig. 6a). As expected, the Gag-FLAG/Pol-HA construct produced viral particles containing Gag and GagPol precursors, an immature form. A number of studies have indicated that expression of the Gag region alone, or without active PR, produces an immature form of virus-like particles (2, 3, 5), consistent with our Gag-FLAG construct. In contrast, expression of GagPol-HA alone did not produce viral particles, similar to GagPol-HA containing active PR, consistent with previous studies (4, 5). When GagPol-HA was co-expressed with Gag-FLAG, we observed efficient particle production at any Gag-to-GagPol ratio that we tested here (Gag:GagPol from 1:0.1 to 1:10). Surprisingly, Western blotting using particle fractions revealed that the amounts of GagPol-HA incorporated into viral particles were increased in a DNA dose-dependent manner (Fig. 6a), indicating that a considerable number of GagPol molecules are incorporated into viral particles when its PR is inactive (compare Figs 5a and 6a). When the virus particles produced by cotransfection at a Gag-to-GagPol ratio of 1:10 were subjected to equilibrium centrifugation on sucrose gradients, they were detected at a density of 1.16–1.18 g/mL, whereas the particles obtained by the Gag-FLAG/Pol-HA construct were also sedimented (Fig. 6b). Consistent with these findings, confocal microscopy revealed that both Gag and GagPol accumulated at the plasma membrane in cells cotransfected with Gag-FLAG and GagPol-HA (Fig. 6c, d). However, when these cells were observed by electron microscopy, virus particles produced by cotransfection at a Gag-to-GagPol ratio of 1:1 and 1:10 displayed aberrant morphology: not hollow but encasing electron-dense amorphous materials. In contrast, cotransfection at a ratio of 1:0.1 produced doughnut-like particles, characteristic of the immature form (Fig. 6e). Together, these data suggest that the Gag-to-GagPol ratio has substantial impacts on particle morphology even when PR is inactive.

DISCUSSION

Gag processing and GagPol dimerization at sites on the plasma membrane where Gag and GagPol are assembled

It is generally accepted that retroviral GagPol protein is incorporated into virus particles through co-assembly with Gag protein, since GagPol protein alone does not produce viral particles but it is incorporated into viral particles when co-expressed with Gag (4, 5). We confirmed these observations and further found that both Gag and GagPol are relocated from the cytoplasm to the plasma membrane and subsequently accumulate at certain sites on the plasma membrane (Figs 2, 3). These sites are most likely to be lipid raft microdomains because they have been suggested to be sites of virus particle assembly for many viruses including HIV (31–34). In this study, we used anti-HIV-1 p17/MA monoclonal antibody that only recognizes mature cleaved MA and demonstrated that Gag processing, at least cleavage of the MA/CA junction, occurred predominantly during assembly at the plasma membrane. Previous studies of *in vitro* processing have described the order of Gag processing: Gag protein is initially cleaved at the p2/NC junction, followed by the MA/CA junction, and finally at the CA/p2 junction (p2/NC > NC/p6 ≥ MA/CA > CA/p2) (35, 36). Although these data raise the possibility that initial cleavage of the p2/NC junction might start during membrane trafficking, we suggest this possibility is unlikely because FRET signals with GagPol dimerization were predominantly observed at the plasma membrane (Fig. 4) even when the constructs contained an active PR (data not shown).

We employed confocal microscopy-based FRET assays to detect the sites of GagPol dimerization and found that it essentially occurs at sites of the plasma membrane where Gag and GagPol have accumulated (Fig. 4). Similarly, many studies have indicated that HIV Gag is multimerized after membrane targeting, suggesting that Gag oligomers formed in the cytosol before membrane targeting are possibly dead-end products (37, 38). It is tempting to speculate that in expression at a normal Gag-to-GagPol ratio, efficient dimerization of Gag-Pol might be suppressed during plasma membrane trafficking, resulting in no PR activity. However, because the FRET system used here was based on EGFP-to-Strawberry (green-to-red) FRET, which is specific with lower background but less sensitive than cyan fluorescent protein-to-yellow fluorescent protein FRET, we cannot rule out that a fraction of GagPol dimers below our detection limitation might have been present in the cytoplasm.



Premature processing by overexpression of GagPol

Previous studies have shown that overexpression of GagPol alone results in absence of particle production (10, 12) and that coexpression of Gag and GagPol at a nearly equimolar ratio severely impairs particle production (13). In this study, we observed the sites of Gag processing under these overexpression conditions. Our confocal data revealed that cleaved p17/MA was diffusely distributed in the cytoplasm when GagPol was expressed alone, suggesting that Gag processing might have occurred before membrane targeting of GagPol (Fig. 5). Interestingly, the equimolar cotransfection of Gag and GagPol constructs exhibited unique distribution patterns of p17/MA: uniform distribution at the plasma membrane and punctate accumulation in the cytoplasm (Fig. 5). The former pattern may result from Gag processing after plasma membrane targeting but before Gag assembly, whereas the latter may represent cytoplasmic processing and/or retargeting to cellular membranous compartments, although these are not proven in this study. Previous studies have shown that cleavage of GagPol occurs even when GagPol lacks the myristoylation signal (5, 12), suggesting there is no relationship between N-terminal myristoylation and processing. Rather, it is conceivable that if the amount of Gag is higher than that of GagPol, GagPol preferentially interacts with Gag and is transported to the plasma membrane (by virtue of the membrane targeting signals of Gag), where it accumulates and forms a homodimer, resulting in PR activation and subsequent Gag/GagPol processing at the plasma membrane. In contrast, when the amount of GagPol is higher than that of Gag, GagPol may preferentially interact with GagPol and form a homodimer in the cytoplasm. This GagPol dimerization would result in aberrant Gag/GagPol processing in the cytoplasm. The resultant processing products would not reach the plasma membrane or participate in virion assembly. Altogether, overexpression of GagPol beyond the normal level results in aberrant distribution of mature cleaved p17/MA, which does not contribute to efficient particle production. Thus, the Gag-to-GagPol ratio has significant impacts on particle assembly when PR is active.

Incorporation of GagPol precursors into virus particles

The incorporation of GagPol precursors containing inactive PR into virus particles has been studied by expression of the *gag-pol* region containing an inactive PR domain and by cotransfection experiments with Gag and GagPol constructs containing inactive PR (4, 5). One study has suggested that nearly equimolar cotransfection of Gag and GagPol constructs produces virus particles as efficiently as expression of the Gag construct alone (5). However in another study, especially when the GagPol construct was cotransfected with the HIV molecular clone, virus particle production was slightly impaired, although the impacts were much less than those observed when PR was active (13). To understand to what extent GagPol molecules can be incorporated into virus particles, we altered the Gag-to-GagPol ratio from 1:0.1 to 1:10 and carried out cotransfection experiments. Surprisingly, we found that the amounts of GagPol incorporated into viral particles increased in a DNA dose-dependent manner (Fig. 6a). These data suggest that the incorporation limitation of GagPol molecules into virus particles is not strictly regulated when its PR is inactive. However, our electron microscopy data indicated that particle morphology became aberrant when the Gag-to-GagPol ratio was altered from 1:0.1 to 1:1 or 1:10 (Fig. 6e), implying that the Gag-to-GagPol ratio has substantial impacts on particle morphology even when PR is inactive. Careful observation of the aberrance in particle morphology, namely amorphous materials encased in particles, suggested that Gag and GagPol might not properly co-assemble or easily dissociate from the envelope membrane during/after particle budding (Fig. 7). We speculate

Fig. 6. Intracellular localization of Gag/GagPol precursors and viral particle production in overexpression of GagPol containing inactive PR. HeLa cells were singly transfected with the Gag-FLAG/Pol-HA, GagPol-HA, and Gag-FLAG construct, or doubly transfected with a combination of the Gag-FLAG and GagPol-HA constructs at a Gag-to-GagPol DNA ratio of 1:0.1, 1:1, and 1:10. Total DNA amounts were normalized to 1.5 μg with pUC plasmid. (a) Virus particle production. Equivalent volumes of cell samples and virus particle samples were subjected to Western blotting using anti-HIV-1 p24/CA antibody. (b) Equilibrium gradient centrifugation of virus particles. Produced particles were applied onto 20–60% (w/v) sucrose gradients and centrifuged at 4°C at 120,000 × *g* overnight. Gradient fractions were analyzed by Western blotting using anti-p24/CA antibody. (c) Intracellular localization of Gag and GagPol proteins. At 24 hr post-transfection, cells were fixed and immunostained with anti-FLAG (for Gag) and anti-HA (for GagPol) antibodies. Nuclei were stained with TOPRO-3. All micrographs are shown at the same magnification. Bar, 10 μm. DIC, differential interference contrast. (d) Semiquantification of Gag and GagPol localization. Approximately 100 antigen-positive cells were observed in each sample and the number of cells with each pattern of Gag and GagPol distribution was counted. PM, plasma membrane. (e) Electron microscopy. All micrographs are shown at the same magnification. Bar, 100 nm.

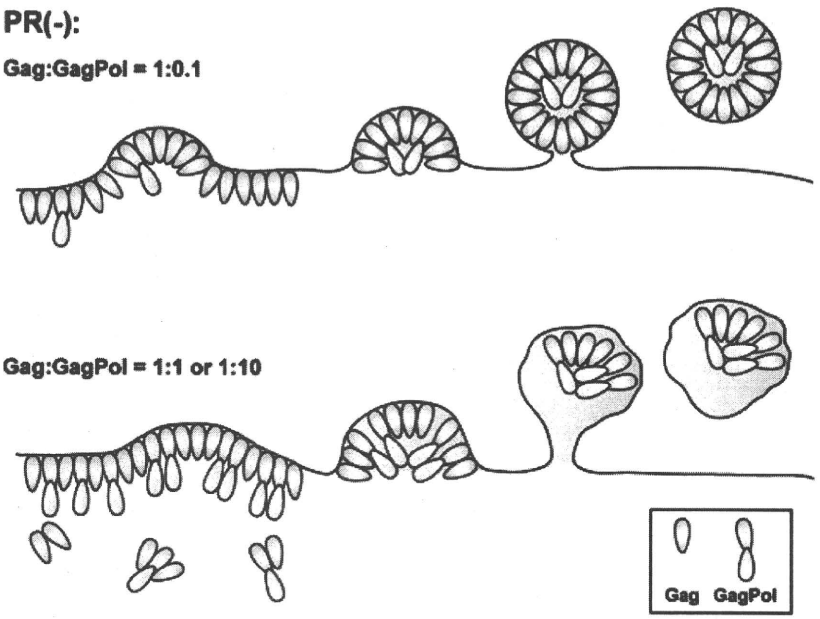


Fig. 7. A model for viral particle production in overexpression of GagPol containing an inactive PR. When the Gag-to-GagPol ratio is 1:0.1, both molecules target to the plasma membrane and assemble into spherical virus particles that exhibit doughnut-like morphology, characteristic of the immature form. When the Gag-to-GagPol ratio is 1:1 or 1:10, both molecules target to the plasma membrane but, because GagPol is considerably larger and less competent at membrane association than Gag, budding particles may exhibit aberrant assembly morphology with irregular membrane curvature. Budding particles may accommodate eccentric cores, possibly representing aggregation of Gag/GagPol molecules after dissociation from the particle membrane.

that GagPol may have intrinsically negative effects on its membrane binding and/or assembly.

ACKNOWLEDGMENTS

We thank Y. Tsunetsugu-Yokota for supply of anti-HIV-1 p24 mouse antibody. This work was supported by an AIDS grant from the Ministry of Health, Labor, and Welfare of Japan, by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science, in part by Grants-in-Aid for Specially Promoted Research and for Scientific Research, by Exploratory Research for Advanced Technology (Japan Science and Technology Agency), and by the National Institute of Allergy and Infectious Diseases Public Health Service research grants, USA.

REFERENCES

1. Jacks T., Power M.D., Masiarz F.R., Luciw P.A., Barr P.J., Varmus H.E. (1988) Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331:** 280–3.
2. Gheysen D., Jacobs E., de Foresta F., Thiriart C., Francotte M., Thines D., de Wilde M. (1989) Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59:** 103–12.
3. Morikawa Y. (2001) HIV capsid assembly. *Curr HIV Res* **1:** 1–14
4. Park J., Morrow C.D. (1992) The nonmyristylated Pr160gag-pol polyprotein of human immunodeficiency virus type 1 interacts with Pr55gag and is incorporated into virus-like particles. *J Virol* **66:** 6304–13.
5. Smith A.J., Srinivasakumar N., Hammarskjold M.L., Rekosh D. (1993) Requirements for incorporation of Pr160*gag-pol* from human immunodeficiency virus type 1 into virus-like particles. *J Virol* **67:** 2266–75.
6. Mervis R.J., Ahmad N., Lillehoj E.P., Raum M.G., Salazar F.H., Chan H.W., Venkatesan S. (1988) The *gag* gene products of human immunodeficiency virus type 1: alignment within the *gag* open reading frame, identification of posttranslational modifications, and evidence for alternative *gag* precursors. *J Virol* **62:** 3993–4002.
7. Gottlinger H.G., Sodroski J.G., Haseltine W.A. (1989) Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc Natl Acad Sci USA* **86:** 5781–5.
8. Rose J.R., Babe L.M., Craik, C.S. (1995) Defining the level of human immunodeficiency virus type 1 (HIV-1) protease activity required for HIV-1 particle maturation and infectivity. *J Virol* **69:** 2751–8.
9. Kaplan A.H., Manchester M., Swanstrom R. (1994) The activity of the protease of human immunodeficiency virus type 1 is initiated at the membrane of infected cells before the release of viral proteins and is required for release to occur with maximum efficiency. *J Virol* **68:** 6782–6.
10. Karacostas V., Wolffe E.J., Nagashima K., Gonda M.A., Moss B. (1993) Overexpression of the HIV-1 *gag-pol* polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology* **193:** 661–71.
11. Krausslich H.G. (1991) Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proc Natl Acad Sci USA* **88:** 3213–17.
12. Park J., Morrow C.D. (1991) Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. *J Virol* **65:** 5111–7.
13. Shehu-Xhilaga M., Crowe S.M., Mak J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J Virol* **75:** 1834–41.
14. Louis J.M., Clore G.M., Gronenborn A.M. (1999) Autoprocessing of HIV-1 protease is tightly coupled to protein folding. *Nat Struct Biol* **6:** 868–75.

15. Tang C., Louis J.M., Aniana A., Suh J-Y., Clore G.M. (2008) Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature* **455:** 693–6.
16. Suyama M., Daikoku E., Goto T., Sano K., Morikawa Y. (2009) Reactivation from latency displays HIV particle budding at plasma membrane, accompanying CD44 upregulation and recruitment. *Retrovirology* **6:** 63.
17. Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A., Martin M.A. (1986) Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* **59:** 284–91.
18. Huang M., Orenstein J.M., Martin M.A., Freed E.O. (1995) p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J Virol* **69:** 6810–18.
19. Tsunetsugu-Yokota Y., Ishige M., Murakami M. (2007) Oral attenuated *Salmonella enterica* serovar Typhimurium vaccine expressing codon-optimized HIV type 1 Gag enhanced intestinal immunity in mice. *AIDS Res Hum Retroviruses* **23:** 278–286.
20. Feige J.N., Sage D., Wahli W., Desvergne B., Gelman L. (2005) PixFRET, an ImageJ plug-in for FRET calculation that can accommodate variations in spectral bleed-throughs. *Microsc Res Tech* **68:** 51–8.
21. Demirov D.G., Ono A., Orenstein J.M., Freed E.O. (2002) Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *Proc Natl Acad Sci USA* **99:** 955–60.
22. Garrus J.E., von Schwedler U.K., Pornillos O.W., Morham S.G., Zavitz K.H., Wang H.E., Wettstein D.A., Stray K.M., Cote M., Rich R.L., Myszka D.G., Sundquist W.I. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107:** 55–65.
23. Strack B., Calistri A., Craig S., Popova E., Gottlinger H.G. (2003) AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* **114:** 689–99.
24. VerPlank L., Bouamr F., LaGrassa T.J., Agresta B., Kikonyogo A., Leis J., Carter C.A. (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:** 7724–9.
25. von Schwedler U.K., Stuchell M., Muller B., Ward D.M., Chung H.Y., Morita E., Wang H.E., Davis T., He G.P., Cimbora D.M., Scott A., Krausslich H.G., Kaplan J., Morham S.G., Sundquist W.I. (2003) The protein network of HIV budding. *Cell* **114:** 701–13.
26. Kondo E., Gottlinger H.G. (1996) A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr. *J Virol* **70:** 159–64.
27. Lu Y.L., Bennett R.P., Wills J.W., Gorelick R., Ratner L. (1995) A leucine triplet repeat sequence (LXX)4 in p6gag is important for Vpr incorporation into human immunodeficiency virus type 1 particles. *J Virol* **69:** 6873–9.
28. Jouvenet N., Neil S.J., Bess C., Johnson M.C., Virgen C.A., Simon S.M., Bieniasz P.D. (2006) Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol* **4:** e435.
29. Kawada S., Goto T., Haraguchi H., Ono A., Morikawa Y. (2008) Dominant negative inhibition of human immunodeficiency virus particle production by the nonmyristoylated form of gag. *J Virol* **82:** 4384–99.
30. Hogue I.B., Hoppe A., Ono A. (2009) Quantitative fluorescence resonance energy transfer microscopy analysis of the human immunodeficiency virus type 1 Gag-Gag interaction: relative contributions of the CA and NC domains and membrane binding. *J Virol* **83:** 7322–36.
31. Holm K., Weclewicz K., Hewson R., Suomalainen M. (2003) Human immunodeficiency virus type 1 assembly and lipid rafts: Pr55(gag) associates with membrane domains that are largely resistant to Brij98 but sensitive to Triton X-100. *J Virol* **77:** 4805–17.
32. Lindwasser O.W., Resh M.D. (2001) Multimerization of human immunodeficiency virus type 1 Gag promotes its localization to barges, raft-like membrane microdomains. *J Virol* **75:** 7913–24.
33. Nguyen D.H., Hildreth J.E. (2000) Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. *J Virol* **74:** 3264–72.
34. Ono A., Freed E.O. (2001) Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proc Natl Acad Sci USA* **98:** 13,925–30.
35. Pettit S.C., Moody M.D., Wehbie R.S., Kaplan A.H., Nantermet P.V., Klein C.A., Swanstrom R. (1994) The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J Virol* **68:** 8017–27.
36. Tritch R.J., Cheng Y.E., Yin F.H., Erickson-Viitanen S. (1991) Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 *gag* polyprotein. *J Virol* **65:** 922–30.
37. Li H., Jun D., Ding L., Spearman P. (2007) Myristoylation is required for human immunodeficiency virus type 1 Gag-Gag multimerization in mammalian cells. *J Virol* **81:** 12,899–910
38. Tritel M., Resh, M.D. (2000) Kinetic analysis of human immunodeficiency virus type 1 assembly reveals the presence of sequential intermediates. *J Virol* **74:** 5845–55.

Biochemical and Biophysical Research Communications 404 (2011) 822–828

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Distinct fucosylation of M cells and epithelial cells by Fut1 and Fut2, respectively, in response to intestinal environmental stress

Kazutaka Terahara [a], Tomonori Nochi [b], Masato Yoshida [b], Yuko Takahashi [b], Yoshiyuki Goto [b], Hirotsugu Hatai [b], Shiho Kurokawa [b], Myoung Ho Jang [c], Mi-Na Kweon [d], Steven E. Domino [e], Takachika Hiroi [f], Yoshikazu Yuki [b], Yasuko Tsunetsugu-Yokota [a], Kazuo Kobayashi [a], Hiroshi Kiyono [b,*]

[a] Department of Immunology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan
[b] Division of Mucosal Immunology, Department of Microbiology and Immunology, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[c] Laboratory of Gastrointestinal Immunology, World Premier International Immunology Frontier Research Center, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan
[d] Mucosal Immunology Section, International Vaccine Institute, Seoul 151-818, Republic of Korea
[e] Department of Obstetrics and Gynecology, The University of Michigan Medical School, 6428 Medical Science Bldg I, 1150 West Medical Center Drive, Ann Arbor, MI 48109-5617, USA
[f] Department of Allergy and Immunology, The Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya-ku, Tokyo 156-8506, Japan

ARTICLE INFO

Article history:
Received 13 December 2010
Available online 21 December 2010

Keywords:
Epithelial cell
Fucosyltransferase
Intestine
M cell
Peyer's patch

ABSTRACT

The intestinal epithelium contains columnar epithelial cells (ECs) and M cells, and fucosylation of the apical surface of ECs and M cells is involved in distinguishing the two populations and in their response to commensal flora and environmental stress. Here, we show that fucosylated ECs (F-ECs) were induced in the mouse small intestine by the pro-inflammatory agents dextran sodium sulfate and indomethacin, in addition to an enteropathogen derived cholera toxin. Although F-ECs showed specificity for the M cell-markers, lectin *Ulex europaeus* agglutinin-1 and our monoclonal antibody NKM 16-2-4, these cells also retained EC-phenotypes including an affinity for the EC-marker lectin wheat germ agglutinin. Interestingly, fucosylation of Peyer's patch M cells and F-ECs was distinctly regulated by α(1,2)fucosyltransferase Fut1 and Fut2, respectively. These results indicate that Fut2-mediated F-ECs share M cell-related fucosylated molecules but maintain distinctive EC characteristics, Fut1 is, therefore, a reliable marker for M cells.

1. Introduction

M cells are generally observed in the follicle-associated epithelium (FAE) of mucosa-associated lymphoid tissues including Peyer's patches (PPs) and isolated lymphoid follicles (ILFs) in the small intestine [1,2], and are morphologically and functionally distinct from their neighboring epithelial cells (ECs) by the presence of relatively short and irregular microvilli on their apical surface and of lymphocytes and antigen-presenting cells frequently enfolded within a pocket structure in their basolateral region [3–5]. In the small intestine of mice, the expression of α(1,2)fucose is believed to be a reliable marker of M cells, because lectin *Ulex europaeus* agglutinin-1 (UEA-1), which has an affinity for α(1,2)fucose, was found to bind exclusively to M cells in the PPs [6,7]. Subsequently, we could found M cells located within the non-FAE-associated small intestinal villous epithelium by utilizing an affinity of UEA-1 [8].

Interestingly, α(1,2)fucosylation is also induced in ileal villous ECs by a variety of intestinal environmental stresses (IES) such as colonization by commensal bacteria, weaning, mechanical injury or treatment with chemicals inhibiting protein synthesis [9–11]. When considering the possible involvement of IES in the development of (or conversion to) M cells, it is reasonable to postulate that such fucosylated ECs form a subset of M cells, because the number of PP M cells is increased rapidly and transiently by alteration from specific pathogen-free (SPF) conditions to a conventional environment [12], by interaction with bacteria such as *Salmonella typhimurium* [13] and *Streptococcus pneumoniae* [14] and during indomethacin-induced ileitis [15]. Like PP M cells, villous M cells might also be induced (or converted) by IES, because a higher frequency of villous M cells is observed in the terminal ileum, which is enriched for commensal bacteria [16]. Recently, we found that some ECs underwent α(1,2)fucosylation in the small intestinal villous epithelium when a mucosal adjuvant cholera toxin (CT) derived from a well known enteropathogen *Vibrio cholerae* was orally administered into mice, and that these cells, in part, shared the same gene expression profile as PP M cells; we previously designated them "M-like cells" [17].

In mice, α(1,2)fucosyltransferase Fut1 and Fut2 are the enzymes responsible for catalyzing an α(1,2) linkage of fucose to terminal β-galactoside, and Fut2 is involved in the IES-associated fucosylation

* Corresponding author. Fax: +81 3 5449 5411.
E-mail address: kiyono@ims.u-tokyo.ac.jp (H. Kiyono).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.12.067

whereas little is known about Fut1 in the intestine [11,18–20]. In this study, we aimed to elucidate the biological characteristics of ECs that shared the α(1,2)fucose modification with M cells focusing on the fucosylation mechanism, in the hope of better understanding of whether ECs can be reprogrammed into M (or M-like) cells in response to IES.

2. Materials and methods

2.1. Mice

BALB/c and C57BL/6J mice were purchased from SLC (Shizuoka, Japan). Fut1-null and Fut2-null mice (C57BL/6J background) were generated as previously described [21]. These mice were maintained under SPF conditions and used in experiments at 6–9 weeks old. All animal experiments were approved by the Animal Care and Use Committee of The University of Tokyo.

2.2. Lectins and antibodies

The following lectins and antibodies were used for flow cytometry (FCM) and confocal laser scanning microscopy (CLSM): PE-conjugated UEA-1 (UEA-1-PE; Biogenesis, England, UK), tetramethylrhodamine isothiocyanate (TRITC)-conjugated UEA-1 (Vector Laboratories, Burlingame, CA), Alexa Fluor 633-conjugated wheat germ agglutinin (WGA-AF633; Molecular Probes, Eugene, OR), FITC-conjugated NKM 16-2-4 mAb (NKM 16-2-4-FITC) [22], APC-conjugated anti-mouse CD45 mAb (anti-CD45-APC; BD Biosciences, San Jose, CA).

2.3. Alteration of the intestinal environment

A mucosal adjuvant, CT (List Biologic Laboratories, Campbell, CA), and two pro-inflammatory agents, dextran sodium sulfate (DSS, m.w. 36,000–50,000; ICN Biomedicals, Irvine, CA) and indomethacin (Sigma–Aldrich, St. Louis, MO), were used as stress-inducing agents to alter the intestinal environment of mice as described previously [17,23,24] (see Supplementary information).

2.4. Preparation of intestinal epithelial cells for FCM

The small intestinal epithelium was dissociated by a mechanical procedure as described previously [17]. The mononuclear cells were stained with NKM 16-2-4-FITC, UEA-1-PE and anti-CD45-APC and dead cells were stained with 7-aminoactinomycin D (BD Biosciences). Fluorescently labeled cells were analyzed and, if necessary, sort-purified using a FACSAria flow cytometer (BD Biosciences) (see Supplementary information).

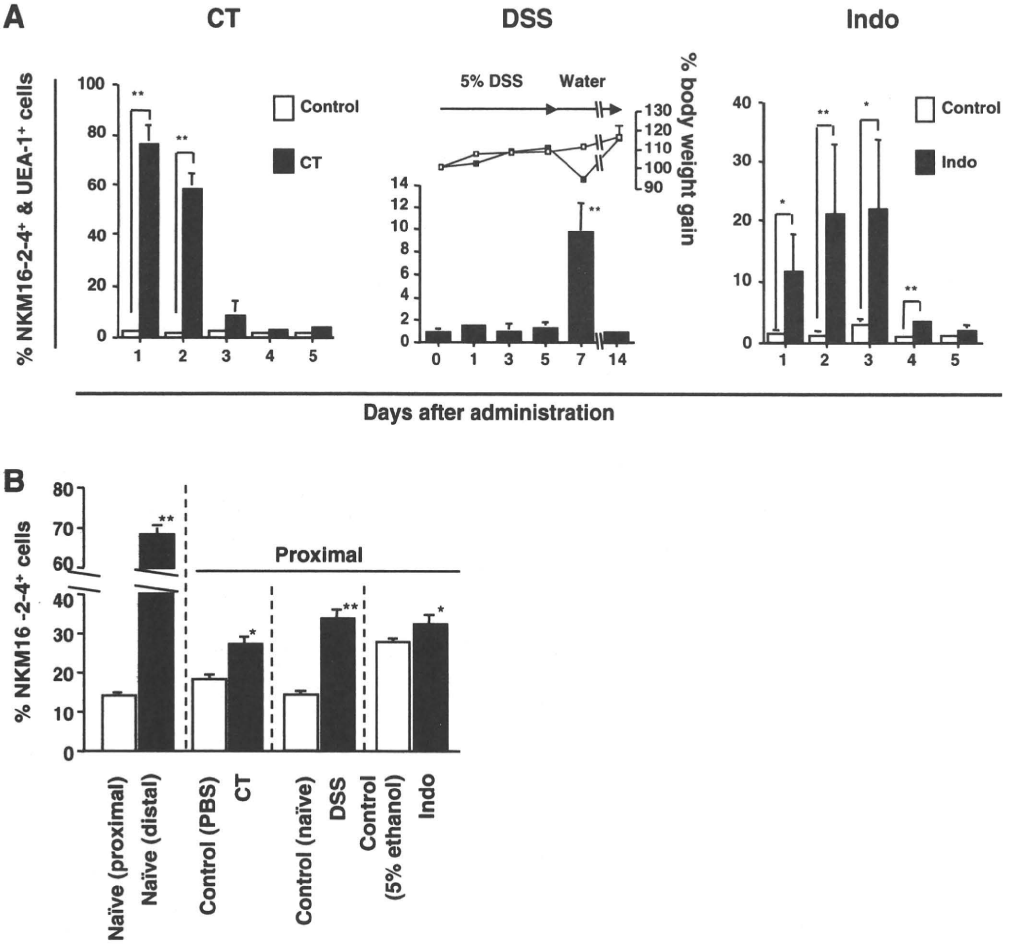


Fig. 1. The influence of IES on the induction of α(1,2)fucosylation in the small intestinal epithelia of BALB/c mice. (A) Daily analysis of the frequency of NKM 16-2-4$^+$/UEA-1$^+$ cells in the proximal villous epithelium of CT-, DSS- and indomethacin (Indo)-administered mice based on FCM. The ratio of NKM 16-2-4$^+$/UEA-1$^+$ cells was enumerated in cells, with 7-aminoactinomycin D$^+$ dead cells, CD45$^+$ leukocytes and small forward- and side-scattered lymphocytes gated out. The line graph in the middle panel shows the percentage body weight gain of control (open squares) and DSS-administered (filled squares) mice. Data are given as means ± SE (n = 3–7). Significant differences ($^*P < 0.05$, $^{**}P < 0.01$) were determined by t-test or Mann–Whitney's U test. (B) The proportions of NKM 16-2-4$^+$ cells in the PP domes based on histoplanimetrical analysis of CLSM images. Mice used were naïve, or were administered CT (day 1), DSS (day 7) or Indo (day 1). Data are given as means ± SE (n = 3, 19–76 domes). Significant differences ($^*P < 0.05$, $^{**}P < 0.01$) were determined by t-test.

2.5. Histological analysis

Fluorescently labeled whole-mount tissues were analyzed by CLSM as described previously [8,22]. Each area of NKM 16-2-4$^+$ cells and whole FAE in PPs was quantitated using Scion Image software (Scion Corporation, Frederick, MA) based on the data obtained by CLSM (see Supplementary information).

2.6. Quantitative real-time RT-PCR for Fut1 and Fut2 transcripts

The levels of the *Fut1* and *Fut2* transcripts were quantitated by real-time RT-PCR in the cDNA samples from the sorted cells with reference to the level of hypoxanthine guanine phosphoribosyl transferase (*Hprt*) transcripts (see Supplementary information).

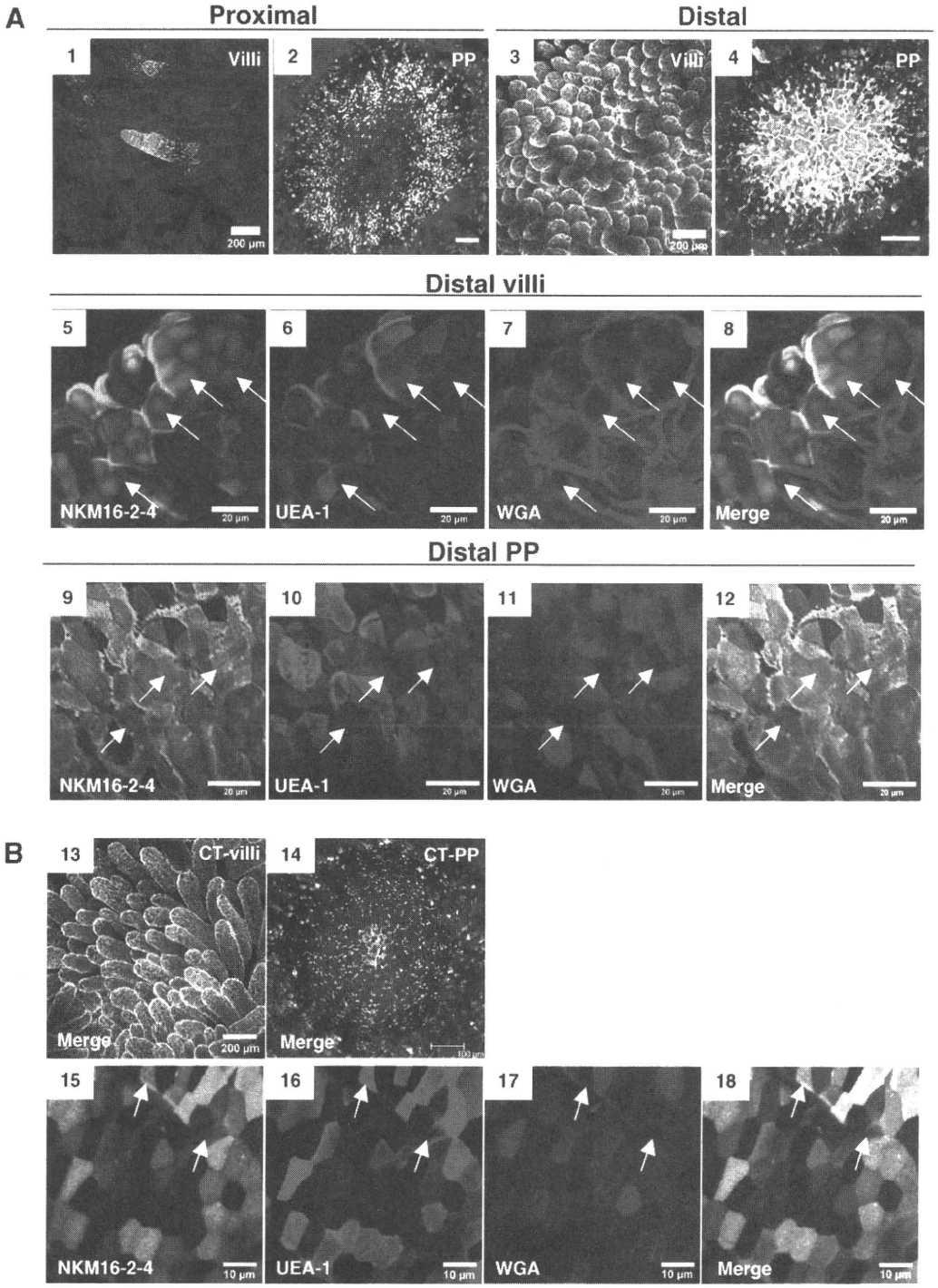


Fig. 2. CLSM analysis of whole-mount small intestinal epithelia of BALB/c mice. Confocal images stained with NKM 16-2-4-FITC, UEA-1-TRITC and WGA-AF633 are shown by green, red and blue, respectively. (A) Proximal villi (1) and PP (2), and distal villi (3, 5–8) and PP (4, 9–12) from naïve mice. (B) Proximal villi (13, 15–18) and PP (14), 1 day after oral CT administration. Arrows show villous M cells (NKM 16-2-4$^+$/UEA-1$^+$/WGA$^-$). Scale bars are 200 μm (1, 3, 13), 100 μm (2, 4, 14), 20 μm (5–12) or 10 μm (15–18). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.7. Statistical analysis

The significance of the data was evaluated by the unpaired *t*-test, Mann–Whitney's U test, Tukey's or Scheffé's multiple comparison test based on the normality and variance of the data compared using Statcel2 software (OMS Publishing Inc., Saitama, Japan). $P < 0.05$ was considered statistically significant.

3. Results

3.1. Induction of α(1,2)fucosylation in the small intestinal epithelium by IES

To examine the influence of IES on M cell-associated α(1,2)fucosylation (NKM 16-2-4$^+$/UEA-1$^+$), we focused on the proximal epithelium, where NKM 16-2-4$^+$/UEA-1$^+$ cells were rarely found in naïve mice (Supplementary Fig. S1). When CT was orally administered to BALB/c mice, NKM 16-2-4$^+$/UEA-1$^+$ cells were dramatically increased in the proximal villous epithelium, with an average of 75.9% double-positive cells one day post-inoculation (Fig. 1A). The proportion of NKM 16-2-4$^+$/UEA-1$^+$ cells returned to the control level (approximately 2%) at 3 days post-inoculation. Similarly, a significant increase in NKM 16-2-4$^+$/UEA-1$^+$ cells was observed when pro-inflammatory agents, such as DSS or indomethacin, were administered (Fig. 1A).

A similar tendency was also seen in the FAE of PPs. We next performed histoplanimetrical analysis based on single NKM 16-2-4 signals obtained by CLSM. We previously demonstrated that NKM 16-2-4$^+$ cells included UEA-1$^+$ M cells but not goblet cells [22]. However, similar to UEA-1, because NKM 16-2-4 reacts to Paneth cells (Supplementary Fig. S2), NKM 16-2-4$^+$ cells were enumerated upward the crypts where Paneth cells locally exist. Therefore, goblet cells and Paneth cells were excluded in this analysis. When the proportions of NKM 16-2-4$^+$ cells were compared between the proximal and distal PP FAE of naïve BALB/c mice, a higher frequency of NKM 16-2-4$^+$ cells was observed in the distal (68.4%) than in the proximal (13.9%) PP FAE (Fig. 1B). Furthermore, a significant increase in NKM 16-2-4$^+$ cells was observed in the proximal PP FAE following CT-, DSS- or indomethacin-administration, averaging 27.2% (control: 18.1%), 33.8% (control: 13.9%) and 32.4% (control: 27.5%) positive cells, respectively (Fig. 1B). These results indicate that IES enhances α(1,2)fucosylation in both the PP FAE and the villous epithelium.

3.2. CLSM analysis of IES-induced NKM 16-2-4$^+$/UEA-1$^+$ cells

To assess qualitative cellular traits of IES-induced NKM 16-2-4$^+$/UEA-1$^+$ cells, we performed CLSM analysis using lectin WGA, which has an affinity for ECs and goblet cells but not M cells [6,8]. As indicated by FCM (Supplementary Fig. S1), a higher frequency of NKM 16-2-4$^+$/UEA-1$^+$ cells was observed in the distal (Fig. 2A; 3 and 4) than in the proximal villi and PPs (Fig. 2A; 1 and 2) in naïve BALB/c mice. In general, these NKM 16-2-4$^+$/UEA-1$^+$ cells were preferentially located at the tips of the villi (Fig. 2A; 1 and 3) and PP domes (Fig. 2A; 4) and a large proportion of them showed an affinity for WGA in both the villous epithelium (Fig. 2A; 7) and the PP FAE (Fig. 2A; 11), although a substantial number of villous M cells sharing the typical M cell hallmark (NKM 16-2-4$^+$/UEA-1$^+$/WGA$^-$) existed in the distal villi of naïve mice (Fig. 2A; 5–8: arrows).

The CLSM analysis further demonstrated that CT-induced NKM 16-2-4$^+$/UEA-1$^+$ cells also reacted with WGA (Fig. 2B). In contrast, villous M cells showing the M cell-phenotype (NKM 16-2-4$^+$/UEA-1$^+$/WGA$^-$) remained at a very low frequency irrespective of IES by CT (Fig. 2B; 15–18: arrows). A similar observation was made when DSS or indomethacin was administered (Supplementary Fig. S3). In the proximal PP FAE, whereas a substantial number of NKM 16-2-4$^+$/UEA-1$^+$ cells were negative for WGA and were radially distributed on the dome, indicating an M cell-phenotype (Fig. 2A; 2), triple-positive cells (NKM 16-2-4$^+$/UEA-1$^+$/WGA$^+$) were evident and were located on the tip of the dome after oral CT administration (Fig. 2B; 14). These results indicate that IES-induced NKM 16-2-4$^+$/UEA-1$^+$ cells share an affinity for WGA, a common trait of normal ECs [6], and hardly contain any villous M cells. We thus designated them fucosylated ECs (F-ECs), to distinguish them from typical M cells.

3.3. Different expression patterns of Fut1 and Fut2 transcripts in the small intestinal epithelium

To examine in more detail the mechanism of α(1,2)fucosylation between F-ECs and M cells, we performed quantitative real-time RT-PCR for *Fut1* and *Fut2* transcripts. Quantitative real-time RT-PCR demonstrated that high expression of *Fut1* transcripts was seen only in NKM 16-2-4$^+$/UEA-1$^+$ cells isolated from the naïve PP FAE where M cells predominantly exist (Fig. 3A). On the other hand, elevated expression of *Fut2* transcripts, but not *Fut1* transcripts, was detected in F-ECs located in the distal epithelia of naïve mice (Fig. 3A and B). Similarly, enhanced expression of *Fut2* transcripts, but not *Fut1* transcripts, was seen in CT-, DSS- and indomethacin-induced F-ECs of the proximal epithelia (Fig. 3A and B). These results indicate that α(1,2)fucosylation of F-ECs in the villous epithelium is induced by Fut2, and suggest that Fut1 is expressed in PP M cells irrespective of IES.

3.4. Distinct requirements for Fut1 or Fut2 for α(1,2)fucosylation of M cells or F-ECs, respectively

To clarify the distinct requirements for the Fut isoforms in F-ECs and M cells, Fut1-null and Fut2-null mice were employed for FCM

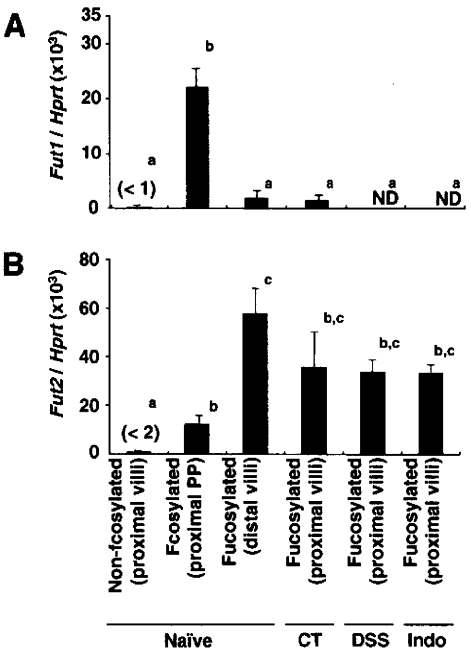


Fig. 3. Quantification of the expression levels of *Fut1* and *Fut2* transcripts relative to levels of *Hprt* transcripts. Fucosylated (NKM 16-2-4$^+$/UEA-1$^+$) and non-fucosylated (UEA-1$^-$) cells were purified from the proximal or distal small intestinal epithelia using a cell-sorter. Naïve, or CT- (Day 2), DSS- (Day 7) or indomethacin (Indo)- (Day 2) treated BALB/c mice were used. (A) *Fut1* transcripts. (B) *Fut2* transcripts. Data are given as means ± SE ($n = 3$–4). Different letters indicate significant differences ($P < 0.05$) determined by Tukey's multiple comparison test.

and CLSM analyses. Because these mice are on a C57BL/6J background [21], wild-type (WT) C57BL/6J mice were employed as a control group. Like BALB/c mice, WT C57BL/6J mice showed a higher frequency of F-ECs (NKM 16-2-4$^+$/UEA-1$^+$ cells) in the distal villous epithelium than in the proximal villous epithelium, and that the frequency of F-ECs in the latter increased after oral CT administration (Fig. 4A). This was also observed when Fut1-null mice were orally exposed to CT (Fig. 4A and C; 2 and 4). On the

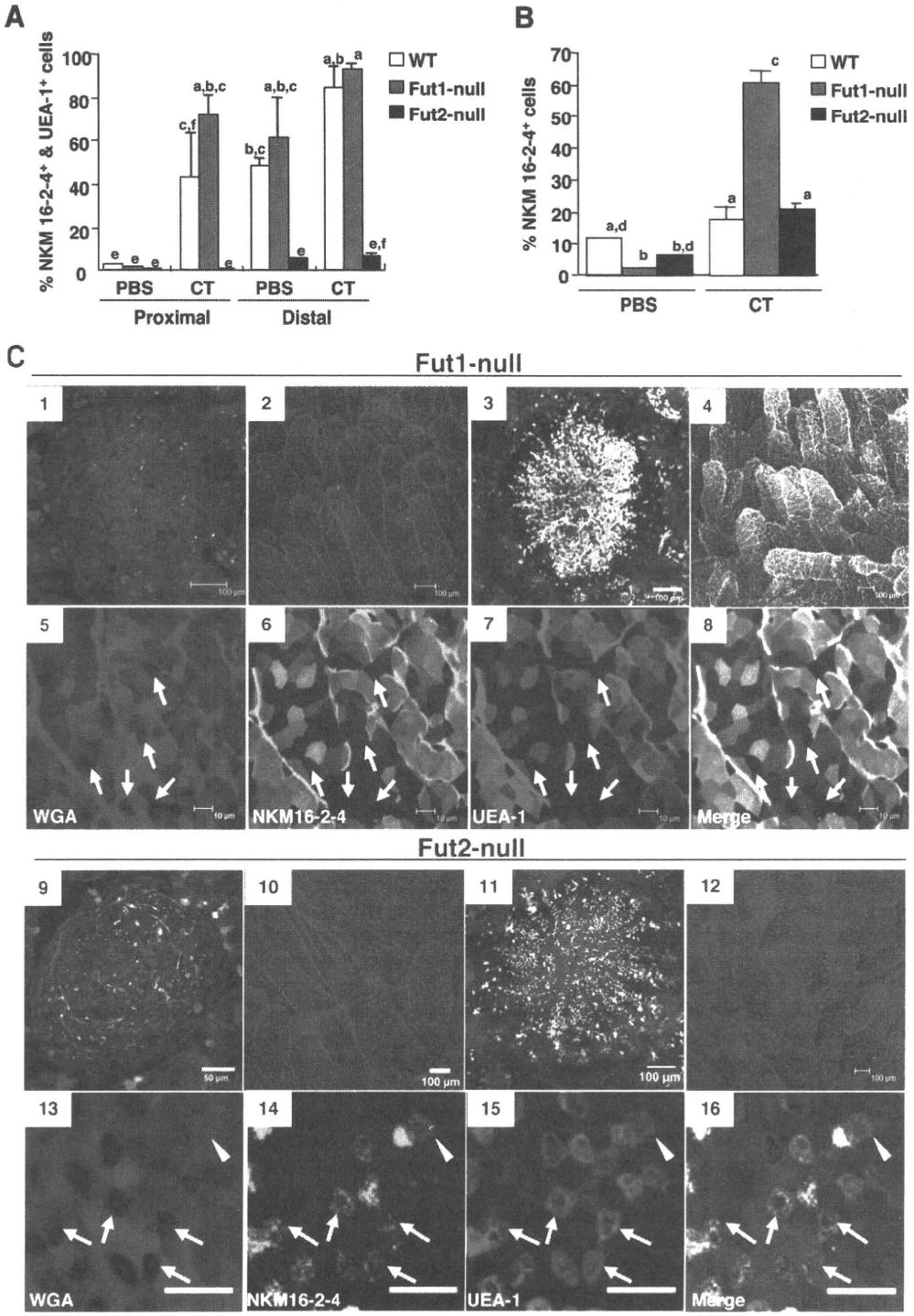


Fig. 4. Fut1- and Fut2-dependent α(1,2)fucosylation in PP M cells and F-ECs, respectively. (A) The proportions of NKM 16-2-4$^+$/UEA-1$^+$ cells in the proximal and distal villous epithelia of either day 1 PBS- or CT-administered (CT) WT C57BL/6J, Fut1-null and Fut2-null mice based on FCM as described in the Fig. 1 legend. Data are given as means ± SE (n = 3). Different letters indicate significant differences ($P < 0.05$) determined by Tukey's multiple comparison test. (B) The proportions of NKM 16-2-4$^+$ cells in the proximal PP domes based on histoplanimetrical analysis of CLSM images. WT, Fut1-null and Fut2-null mice were used after oral administration of either PBS or CT (day 1). Data are given as means ± SE (n = 3, 12–25 domes). Different letters indicate significant differences ($P < 0.05$) determined by Scheffé's multiple comparison test. (C) CLSM analysis for the whole-mount small intestinal epithelia of either naïve or day 1 CT-treated Fut1-null (naïve; 1, 2, CT; 3–8) and Fut2-null mice (naïve; 9, 10, CT; 11–16). Confocal images stained with NKM 16-2-4-FITC, UEA-1-TRITC and WGA-AF633 are shown by green, red and blue, respectively. Arrows and arrowheads indicate WGA$^-$ PP M cells and Fut1-dependent WGA$^+$ cells, respectively. Scale bars are 100 μm (1–4, 10–12), 50 μm (9), 20 μm (13–16) or 10 μm (5–8). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

other hand, Fut2-null mice possessed few NKM 16-2-4$^+$/UEA-1$^+$ cells in the villous epithelia and the number of these cells was not increased upon oral CT administration (Fig. 4A and C; 10 and 12). PBS-administered Fut2-null mice showed 0.1% and 4.9% of NKM 16-2-4$^+$/UEA-1$^+$ cells in the proximal and distal villous epithelia, respectively, and CT-administered Fut2-null mice showed 0.3% and 5.3% in the proximal and distal villous epithelia, respectively (Fig. 4A).

In contrast to the villous epithelia, the PP FAE contained both types of fucosylated cells, dependent on either Fut1 or Fut2. In PBS-administered conditions, 2.0% and 6.3% of NKM 16-2-4$^+$ cells were observed in the proximal PP FAE of Fut1-null and Fut2-null mice, respectively (Fig. 4B). Oral CT administration did not notably induce α(1,2)fucosylation in the PP FAE of WT mice; CT-administered and PBS-administered WT mice contained 17.8% and 12.0% of NKM 16-2-4$^+$ cells, respectively, and there was no statistical difference between the two groups (Fig. 4B). However, both Fut1-null and Fut2-null mice showed significant induction of NKM 16-2-4$^+$ cells in the proximal PP FAE, with the proportion of positive cells being 60.5% and 21.0%, respectively (Fig. 4B). CLSM analysis further revealed that the fucosylation of M cells was dependent on Fut1, because WGA$^-$ M cells did not induce α(1,2)fucosylation and thus did not react with either NKM 16-2-4 or UEA-1 in Fut1-null mice (Fig. 4C; 1, 3 and 5–8: arrows) whereas they did react with these markers in Fut2-null mice (Fig. 4C; 9, 11 and 13–16: arrows) irrespective of oral CT administration. In addition, the fucosylation of F-ECs in the PP FAE was dependent on Fut2 because triple-positive cells (NKM 16-2-4$^+$/UEA-1$^+$/WGA$^+$ cells) were preferentially observed in the CT-administered Fut1-null mice (Fig. 4C; 3 and 5–8). Although a small number of Fut1-dependent WGA$^+$ cells were observed in the PP FAE of CT-administered Fut2-null mice (Fig. 4C; 13–16: arrowheads), they were distributed radially, accompanying abundant WGA$^-$ M cells on the PP dome (Fig. 4C; 11), and were distinguishable from the Fut2-dependent F-ECs that were distributed all over the dome (Fig. 4C; 3).

Taken together, these results indicate that α(1,2)fucosylation of PP M cells is dependent on Fut1 irrespective of IES, and that Fut2 is involved in α(1,2)fucosylation of F-ECs residing in both the PP FAE and villous epithelium in response to IES.

4. Discussion

In this study, we showed that intestinal environmental and biological stress induced F-ECs, which were recognized by NKM 16-2-4 and UEA-1, in both the PP FAE and the villous epithelium. However, such IES-induced F-ECs possessed a strong affinity for WGA (Fig. 2). In addition, F-ECs showed the same morphological characteristics as ECs such as columnar architecture, well-developed tall and dense microvilli (Supplementary Fig. S4), and did not possess *Salmonella* uptake-ability (Supplementary Fig. S5). Furthermore, F-ECs did not express glycoprotein 2 (Supplementary Fig. S6 and Table S1), recently identified as an M cell-specific molecule [17,25]. Therefore, F-ECs should be distinguished from typical M cells, and IES-induced α(1,2)fucosylation reflects only a phenotypic change of surface glycosylation pattern that is irrelevant to M-cell differentiation.

The requirements for different fucosylation-inducing enzymes clearly demonstrated a distinction between F-ECs and PP M cells: Fut1 is essential for the fucosylation associated with PP M cells while Fut2 is specifically involved in the fucosylation of IES-induced F-ECs in both the PP FAE and villous epithelium (Fig. 4). Although it has been reported that the expression of *Fut1* transcripts is rare and is not induced or altered in the small intestine by the transfer from germ-free to conventional conditions [11,18–20], these results are probably due to the low frequency of PP M cells throughout the small intestine. In contrast, it has been known that *Fut2* transcripts are induced in the small intestine, particularly in the ileum, of mice in response to colonization by commensal bacteria or treatment with a protein synthesis inhibitor [11,19]. Our present data, in which IES resulted in the induction of Fut2-dependent F-ECs, is consistent with and support these previous findings.

Fut1 and Fut2 provide insights into the involvement of IES in the development of not only F-ECs but also M cells. The PP dome epithelium consists of two cell lineages: one is derived from the dome-associated crypts and differentiates into either M cells or ECs, and the other is derived from villus-associated crypts and differentiates into ECs [26]. In addition, some studies have revealed a dynamic and plastic morphology of M cells; for example, the distinctive microfold and membranous structures occur transiently during the cell differentiation process, and M cell-lineage cells in their early and terminal development stages show the same morphological structure as ECs [27,28]. In this study, we showed a possibility that Fut1-dependent fucosylated cells are increased by IES (Fig. 4). These cells consisted of abundant PP M cells and a few WGA$^+$ EC-like cells, both of which were distributed radially on the dome. To this end, we suggest a possibility that Fut1-dependent cells are M cell-lineage cells derived from the dome-associated crypts that participate in the increase of M cells in response to IES, as described elsewhere [12–15]. In contrast to PP M cells, the fucosylation of villous M cells, like F-ECs, is regulated by Fut2 because α(1,2)fucosylation in the villi was not observed in Fut2-null but Fut1-null mice regardless of oral CT administration (Fig. 4 and Supplementary Table S1). However, IES alone would not influence the frequency of villous M cells because oral CT administration did not induce *Salmonella* uptake in the villi (Supplementary Fig. S5). It has recently been shown that receptor activator of nuclear factor-kappa B ligand (RANKL) is capable of the full development of both PP and villous M cells but RANKL-expressing inducer cells preferentially exist in the subepithelial dome of PPs [29]. Taken together, transient IES alone might be insufficient for the recruitment and/or induction of RANKL-expressing cells in the villi, and/or other factors might be required for the full development of villous M cells.

Although specific functions of F-ECs remain to be elucidated, our present study offers the possibility that Fut1-null and Fut2-null mice would provide a direct opportunity to examine *in vivo* the immuno-biological role of F-ECs and M cells, including their specific fucosylated glycans, towards a better understanding of the gut mucosal immune system.

Acknowledgments

We thank Dr. Osamu Igarashi for technical support and Dr. Rebecca Devon for editing the manuscript. This work was supported in part by Grants from the Ministry of Education, Science, Sports and Culture of Japan (H.K. and K.T.), and the Ministry of Health, Labour and Welfare of Japan (H.K.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.12.067.

References

[1] M.R. Neutra, A. Frey, J.P. Kraehenbuhl, Epithelial M cells: gateways for mucosal infection and immunization, Cell 86 (1996) 345–348.
[2] H. Hamada, T. Hiroi, Y. Nishiyama, et al., Identification of multiple isolated lymphoid follicles on the antimesenteric wall of the mouse small intestine, J. Immunol. 168 (2002) 57–64.
[3] R.L. Owen, A.L. Jones, Epithelial cell specialization within human Peyer's patches: an ultrastructural study of intestinal lymphoid follicles, Gastroenterology 66 (1974) 189–203.

[4] M.R. Neutra, N.J. Mantis, A. Frey, et al., The composition and function of M cell apical membranes: implications for microbial pathogenesis, Semin. Immunol. 11 (1999) 171–181.
[5] J.P. Kraehenbuhl, M.R. Neutra, Epithelial M cells: differentiation and function, Annu. Rev. Cell Dev. Biol. 16 (2000) 301–332.
[6] M.A. Clark, M.A. Jepson, N.L. Simmons, et al., Differential expression of lectin-binding sites defines mouse intestinal M-cells, J. Histochem. Cytochem. 41 (1993) 1679–1687.
[7] M.A. Clark, M.A. Jepson, N.L. Simmons, et al., Selective binding and transcytosis of Ulex europaeus 1 lectin by mouse Peyer's patch M-cells in vivo, Cell Tissue Res. 282 (1995) 455–461.
[8] M.H. Jang, M.N. Kweon, K. Iwatani, et al., Intestinal villous M cells: an antigen entry site in the mucosal epithelium, Proc. Natl. Acad. Sci. USA 101 (2004) 6110–6115.
[9] Y. Umesaki, M. Ohara, Factors regulating the expression of the neutral glycolipids in the mouse small intestinal mucosa, Biochim. Biophys. Acta 1001 (1989) 163–168.
[10] L. Bry, P.G. Falk, T. Midtvedt, et al., A model of host-microbial interactions in an open mammalian ecosystem, Science 273 (1996) 1380–1383.
[11] B. Lin, Y. Hayashi, M. Saito, et al., GDP-fucose: beta-galactoside alpha1,2-fucosyltransferase, MFUT-II, and not MFUT-I or -III is induced in a restricted region of the digestive tract of germ-free mice by host-microbe interactions and cycloheximide, Biochim. Biophys. Acta 1487 (2000) 275–285.
[12] M.W. Smith, P.S. James, D.R. Tivey, M cell numbers increase after transfer of SPF mice to a normal animal house environment, Am. J. Pathol. 128 (1987) 385–389.
[13] T.C. Savidge, M.W. Smith, P.S. James, et al., Salmonella-induced M-cell formation in germ-free mouse Peyer's patch tissue, Am. J. Pathol. 139 (1991) 177–184.
[14] C. Borghesi, M.J. Taussig, C. Nicoletti, Rapid appearance of M cells after microbial challenge is restricted at the periphery of the follicle-associated epithelium of Peyer's patch, Lab. Invest. 79 (1999) 1393–1401.
[15] T. Kucharzik, A. Lugering, N. Lugering, et al., Characterization of M cell development during indomethacin-induced ileitis in rats, Aliment. Pharmacol. Ther. 14 (2000) 247–256.
[16] J. Mach, T. Hshieh, D. Hsieh, et al., Development of intestinal M cells, Immunol. Rev. 206 (2005) 177–189.
[17] K. Terahara, M. Yoshida, O. Igarashi, et al., Comprehensive gene expression profiling of Peyer's patch M cells, villous M-like cells and intestinal epithelial cells, J. Immunol. 180 (2008) 7840–7846.
[18] S.E. Domino, N. Hiraiwa, J.B. Lowe, Molecular cloning, chromosomal assignment and tissue-specific expression of a murine alpha(1,2)-fucosyltransferase expressed in thymic and epididymal epithelial cells., Biochem. J. 327 (Pt 1) (1997) 105–115.
[19] M. Iwamori, S.E. Domino, Tissue-specific loss of fucosylated glycolipids in mice with targeted deletion of alpha(1, 2)fucosyltransferase genes, Biochem. J. 380 (2004) 75–81.
[20] B. Lin, M. Saito, Y. Sakakibara, et al., Characterization of three members of murine alpha1,2-fucosyltransferases: change in the expression of the Se gene in the intestine of mice after administration of microbes, Arch. Biochem. Biophys. 388 (2001) 207–215.
[21] S.E. Domino, L. Zhang, P.J. Gillespie, et al., Deficiency of reproductive tract alpha(1,2)fucosylated glycans and normal fertility in mice with targeted deletions of the FUT1 or FUT2 alpha(1,2)fucosyltransferase locus, Mol. Cell. Biol. 21 (2001) 8336–8345.
[22] T. Nochi, Y. Yuki, A. Matsumura, et al., A novel M cell-specific carbohydrate-targeted mucosal vaccine effectively induces antigen-specific immune responses, J. Exp. Med. 204 (2007) 2789–2796.
[23] I. Okayasu, S. Hatakeyama, M. Yamada, et al., A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice, Gastroenterology 98 (1990) 694–702.
[24] T. Kunikata, H. Araki, M. Takeeda, et al., Prostaglandin E prevents indomethacin-induced gastric and intestinal damage through different EP receptor subtypes, J. Physiol. Paris 95 (2001) 157–163.
[25] K. Hase, K. Kawano, T. Nochi, et al., Uptake through glycoprotein 2 of FimH(+) bacteria by M cells initiates mucosal immune response, Nature 462 (2009) 226–230.
[26] A. Gebert, S. Fassbender, K. Werner, et al., The development of M cells in Peyer's patches is restricted to specialized dome-associated crypts, Am. J. Pathol. 154 (1999) 1573–1582.
[27] S. Onishi, T. Yokoyama, K. Chin, et al., Ultrastructural study on the differentiation and the fate of M cells in follicle-associated epithelium of rat Peyer's patch, J. Vet. Med. Sci. 69 (2007) 501–508.
[28] F. Sierro, E. Pringault, P.S. Assman, et al., Transient expression of M-cell phenotype by enterocyte-like cells of the follicle-associated epithelium of mouse Peyer's patches, Gastroenterology 119 (2000) 734–743.
[29] K.A. Knoop, N. Kumar, B.R. Butler, et al., RANKL is necessary and sufficient to initiate development of antigen-sampling M cells in the intestinal epithelium, J. Immunol. 183 (2009) 5738–5747.

Vaccine 28S (2010) B60–B67

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

HIV-2 capsids distinguish high and low virus load patients in a West African community cohort

Clayton O. Onyango[a], Aleksandra Leligdowicz[a,b], Masaru Yokoyama[c], Hironori Sato[c], Haihan Song[d], Emi E. Nakayama[d], Tatsuo Shioda[d,*], Thushan de Silva[a], John Townend[a], Assan Jaye[a], Hilton Whittle[a], Sarah Rowland-Jones[a,b], Matthew Cotten[a,*]

[a] *Medical Research Council Laboratories, Fajara, Atlantic Road, P.O. Box 273, The Gambia*
[b] *Weatherall Institute of Molecular Medicine, Medical Research Council Human Immunology Unit, John Radcliffe Hospital, Oxford, OX3 9DS, United Kingdom*
[c] *Laboratory of Viral Genomics, Center for Pathogen Genomics, National Institute of Infectious Diseases, Tokyo 208-0011, Japan*
[d] *Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, Osaka 565-0871, Japan*

ARTICLE INFO

Article history:
Received 15 May 2009
Received in revised form 13 August 2009
Accepted 17 August 2009

Keywords:
HIV-2
Capsid
TRIM5α

ABSTRACT

HIV-2 causes AIDS similar to HIV-1, however a considerable proportion of HIV-2 infected patients show no disease and have low plasma virus load (VL). An analysis of HIV-2 capsid (p26) variation demonstrated that proline at p26 positions 119, 159 and 178 are more frequent in lower VL subjects while non-proline residues at all three sites are more frequent in subjects with high VL. *In vitro* replication levels of viruses bearing changes at the three sites suggested that these three residues influence virus replication by altering susceptibility to TRIM5α. These results provide new insights into HIV-2 pathogenesis.

1. Introduction

HIV-2 was discovered in West Africa patients [1,2] shortly after the discovery of HIV-1, with HIV-2 entering humans via zoonoses distinct from the entry of HIV-1 some time in the early 20th century [3]. The two viruses have 60–80% sequence homology and have similar genomic organization, yet the viruses have distinct transmission rates and disease associations (reviewed in Ref. [4]). Although some HIV-2-infected patients progress to AIDS, the majority control infection [5–7] and patients with low VL survive longer [8]. Early descriptions of HIV-2 observed differences in the virus genetics [9] and noted that the HIV-2 epidemic behaved liked a mixture of pathological and non-pathological viruses [5,10]. Studying and understanding HIV-2 represents a possibility to discover how a potentially lethal lentivirus infection can be controlled by humans.

There are limited data on HIV-2 sequence variation associated with either long-term control or progression to disease [9,11,12]. Although some correlation of Nef variation with disease has been observed [13], most HIV-2 sequences in the public databases are derived from viruses isolated from AIDS patients where high VLs facilitate virus isolation and sequencing which could introduce bias in the data. Sequences derived from patients with low VL are essential for understanding disease non-progression in HIV-2 infection. A community cohort to study HIV was established in Caio, Guinea Bissau in 1988 [50], reviewed in Ref. [4]. HIV screening of the entire adult population in the Caio community cohort have identified HIV-2-positive subjects with a wide spectrum of clinical symptoms and survival [4,10,14]; a sequence analysis of the virus among these patients could identify virus variations that influence outcome.

The capsid (CA) protein (p24 in HIV-1; p26 in HIV-2) has a structure that is conserved among retroviruses [15,16]. p26 accumulates during replication as a Gag polyprotein, assembles into spherical structures that package viral RNA and is subsequently processed by the viral protease and reassembled into the mature virus cores [16–18]. The CA has a distinct amino-terminal domain (residues 1–145), that is exposed on the surface of cores, and a C-terminal domain (residues 146–230) which is required for oligomerization [19]. The 20 amino acid major homology region (MHR) in the C-terminal domain is highly conserved; changes in this motif can interfere with CA assembly, maturation and early stages of infec-

* Corresponding authors.
E-mail addresses: conyango@mrc.gm (C.O. Onyango), aleksandra.leligdowicz@mail.mcgill.ca (A. Leligdowicz), yokoyama@nih.go.jp (M. Yokoyama), hirosato@nih.go.jp (H. Sato), haihansong@hotmail.com (H. Song), emien@biken.osaka-u.ac.jp (E.E. Nakayama), shioda@biken.osaka-u.ac.jp (T. Shioda), tdesilva@mrc.gm (T. de Silva), jtownend@mrc.gm (J. Townend), ajaye@mrc.gm (A. Jaye), hwhittle@mrc.gm (H. Whittle), sarah.rowland-jones@ndm.ox.ac.uk (S. Rowland-Jones), mcotten@mrc.gm (M. Cotten).

0264-410X/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2009.08.060

Table 1
Comparison of the cohort p26 residues 119, 159 and 178 to HIV-2_{ROD} p26; variation with viral load.

ID	Source	VL	119	159	178	ID	Source	VL	119	159	178
ROD			P	P	P	ROD			P	P	P
CO310	RNA	<100	.	.	.	CO309	RNA	283	Q	S	.
CO315	DNA	<100	.	.	.	CO316	RNA	372	.	.	.
CO318	PRO	<100	.	.	.	CO364	RNA	387	A	.	.
CO319	RNA	<100	.	.	.	CO314	DNA	393	.	S	.
CO324	RNA	<100	.	.	.	CO354	DNA	413	.	.	.
CO326	RNA	<100	.	.	.	CO323	RNA	497	A	.	.
CO327	DNA	<100	.	.	.	CO365	RNA	523	A	S	A
CO339	RNA	<100	.	.	.	CO350	RNA	540	A	S	A
CO349	RNA	<100	.	.	.	CO322	RNA	610	.	S	A
CO359	RNA	<100	.	.	.	CO355	RNA	651	A	.	A
CO360	RNA	<100	.	.	.	CO357	RNA	813	A	S	Q
CO344	DNA	<100	A	.	.	CO332	RNA	1085	A	.	V
CO346	DNA	<100	A	.	.	CO305	RNA	1343	A	.	.
CO351	RNA	<100	A	.	.	CO340	RNA	1587	.	.	.
CO308	RNA	<100	Q	.	.	CO321	RNA	1608	.	.	.
CO337	DNA	<100	.	T	.	CO342	RNA	1907	A	S	.
CO361	RNA	<100	A	S	.	CO325	RNA	1999	.	.	.
CO366	RNA	<100	A	S	.	CO330	RNA	2653	A	S	.
CO367	DNA	<100	A	S	.	CO331	RNA	2949	A	S	A
CO363	DNA	<100	Q	S	.	CO320	RNA	3241	A	.	.
CO335	RNA	<100	A	.	A	CO307	RNA	3764	G	S	.
CO336	RNA	<100	A	.	A	CO368	RNA	6431	A	T	S
CO341	RNA	<100	A	.	A	CO345	RNA	9659	A	.	.
CO362	DNA	<100	A	.	A	CO312	RNA	9979	A	S	A
CO301	DNA	<100	A	S	.	CO334	RNA	10752	.	S	.
CO306	DNA	<100	A	S	A	CO311	RNA	14104	A	S	A
CO333	RNA	<100	A	S	A	CO313	RNA	17067	A	S	A
CO302	RNA	109	.	.	.	CO369	RNA	22446	Q	S	.
CO304	DNA	114	.	.	.	CO343	RNA	25836	A	S	.
CO303	RNA	154	A	S	A	CO338	RNA	28581	Q	S	A
CO356	DNA	182	.	.	.	CO317	RNA	37503	A	S	A
CO328	RNA	198	A	.	.	CO347	RNA	146284	A	.	.
CO353	DNA	234	.	.	.	CO348	RNA	148593	A	.	A
CO329	RNA	267	A	S	S	CO358	RNA	283542	A	S	.
CO352	RNA	275	.	S	.						

Samples from 69 HIV-2 singly-infected patients collected in 2006 were used for sequencing (see Section 2). Patient plasma VL was assayed by RT-PCR with a lower limit of detection of 100 virus copies/ml. For statistical analysis, undetectable viremia was given a value of 50 copies/ml. Viral loads (copies/ml) are indicated in second column. The source of template for the PCR reaction (circulation RNA (RNA) or proviral DNA (DNA)) is indicated. The reference sequence derived from HIV-2_{ROD} p26 is listed in the first row, in the subsequent rows identity with the HIV-2_{ROD} p26 sequence is indicated with a period (.).

tion [19,20]. The multiple functions impose strict constraints on the amino acid changes allowed in CA. Studying the limited variations that do occur in p26 can provide important information on both protein function and the selective forces acting on HIV-2. In this study, we have tested the hypothesis that HIV-2 p26 capsid variants modulate HIV-2 viral load.

2. Methods

2.1. Patient cohort.

Study subjects were recruited from the community-based cohort in Caio, Guinea Bissau (study subjects described in Ref. [14]). Briefly, HIV screening was performed with a Murex ICE HIV-1.2.0 immunoassay (Murex Diagnostics), confirmed and differentially diagnosed (HIV-1/HIV-2) using HEXAGON HIV (Human GmbH). Dually HIV-1+2 positive samples were subjected to peptide-based assays (Pepti-Lav 1-2, Sanofi Diagnostics Pasteur) and HIV-1 and HIV-2-specific polymerase chain reaction (PCR) with primers targeted to the long terminal repeats [7,8]. HIV-1+2 infected patients were excluded from the study. HIV-2 plasma viral load was quantified by PCR based method [7,21] with a lower limit of detection of 100 copies/ml, for analytic purposes, undetectable samples were assigned a value of 50 copies/ml. All participants were antiretroviral naïve at the time of the study and provided informed consent. The study was approved by the Gambian Government/MRC Ethics Committee, The Republic of Guinea Bissau Ministry of Health, and the Oxford Tropical Research Ethics Committee (OXTREC), UK.

2.2. p26 gene amplification strategy, optimization of primers.

p26 gene amplification primers were designed by selecting conserved sequences flanking the p26 coding region from all HIV-2 isolates in the Los Alamos HIV Database to be used as targets for PCR primers. Coupled reverse transcriptase and nested PCR reactions were performed with plasma-isolated virus RNA as template. The following primer sequences in various combinations were used for amplification:

MO017(OF, outer forward) GTCTGCGTCATTTGGTGCAT
MO018(IF, inner forward) CTGCAGAGAAAATGCCAAGCA
MO019(OR outer reverse) GGGCAGTTTGTCATGATGTGTCC
MO020(IR, inner reverse) GCCCTTCCTTTCCACAGTTCCA
MO021(IR) GCCCTTCCTTTCCACAATTCCA
MO030(OF)CACGCAGAAGAGAAAGTGAAAG
MO031(OR) CGGGGAAGTTGCGRGGCTT
MO032(IF) AGTAGACCAACAGCACCACC
MO036(OF) GTGGGCAGCGAATGAATTGG
MO037(OF) GTGGGCAGCGAACGAATTGG
MO038(OR) AAAGAGAGAATTGAGGTGCAGCA.

In case none of the primer combinations amplified a product from RNA, amplification of provirus was performed with the same

series of primers using genomic DNA from the patients' peripheral blood mononuclear cells (PBMCs) as template. The origin of the sequence (viral RNA or proviral DNA) is indicated in Table 1. PCR products were excised from preparative gels, purified by QiaQuick columns, sequenced using the inner PCR primers from both directions and aligned using ClustalW; ambiguities were resolved by direct analysis of sequencing scans using a sequence alignment editor BioEdit [22].

The p26 sequences were monitored to rule out cross-contamination during sample handling or PCR. Contamination among different samples would generate multiple sequences with 100% identity. A phylogenetic analysis showed no identical sequences, with the closest pairs (CO336 and CO335) showing 0.58% difference (4 differences in 690 nucleotides). BLAST analyses revealed no homology greater than 94% with any sequence in the GenBank database (results not shown). All the p26 sequences were consistent with HIV-2 subtype A infections.

2.3. HIV-2 envelope gene amplification

Envelope sequences were obtained from 34 HIV-2 isolates selected randomly from the set of 69 Caio viruses (see below). The following PCR primers targeted to conserved envelope flanking sequences were used:

MO080 (outer forward) CAGTCATCACAGAGTCATGTG
MO076 (outer reverse) TCCTTGTGGATAYGAYCTGT
MO072 (inner forward) TCATGTGAYAAGCACTATTGGGA
MO077 (inner reverse) GGAAGAGAAAACAGGCCTATAGCC.

An approximately 1500 bp fragment (HIV-2_{ROD} C2 to gp41, positions 6783–8285) was amplified in a nested RT-PCR from plasma virus. PCR products were purified as described above, sequenced using inner PCR primers from both directions. If the initial four primers failed, these additional primers were used:

MO122a GTGGACTAACTGCAGAGGAGAATT
MO125 AGTTCTGCCACCTCTGCACT
MO125a AGAAAACCCAAGAACCCTAGCAC.

A fragment of approximately 1300 bp corresponding to HIV-2_{ROD} positions 6849–8210 was used for the phylogenetic analysis. The p26 and Env sequences described in this manuscript were submitted to GenBank (accession numbers GQ485448-GQ485550).

2.4. Molecular modelling

Three-dimensional (3-D) models of HIV-2 CAs were constructed by the homology-modelling technique using the Molecular Operating Environment (MOE) (Chemical Computing Group Inc., Quebec, Canada) as described in Refs. [23–26]. The two crystal structures of the HIV-1 CA proteins were used as templates for the modelling; a CA monomer at a resolution of 3.00 Å (Protein Data Bank (PDB) code: 1E6J [27]) and the dimer of CA C-terminal domain at a resolution of 1.70 Å (PDB code: 1A8O [19]). The amino acid sequence identity of HIV-1 (1E6J) and HIV-2 CA (CO310 in this study) is about 70.5%. The sequence similarity is sufficient to construct a structural model with an r.m.s. deviation of approximately 1.5 Å for the main chain between the predicted and actual structures [28]. The 3-D structures were optimized thermodynamically by energy minimization using MOE and an AMBER99 force field [29] and further refined the physically unacceptable local structures on the basis of evaluation of unusual dihedral angles, ψ and ϕ, by the Ramachandran plot using MOE.

The binding energies of the p26 dimer models, E_{bind}, were calculated as described elsewhere [25] [30], using the formula $E_{bind} = E_{dimer} - 2E_{monomer}$, where E_{dimer} is the energy of the p26 dimer; $E_{monomer}$ is the energy of the p26 monomer. Spearman's rank correlation coefficient for viral load and capsid binding energy and its statistical evaluation were calculated by SPSS ver 14 (SPSS Inc, Chicago).

2.5. Expression of TRIM5α

Construction of recombinant Sendai viruses (SeVs) carrying human MT4-TRIM5α-tag (Hu-TRIM5α-Sev) and cynomolgus monkey TRIM5α lacking the SPRY domain (CM-SPRY(-)-Sev) was described previously [23].

2.6. Construction of HIV-2 GH123 capsid variants

Mutant DNA constructs of infectious molecular clones of HIV-2 GH123 carrying alanine 119 (GH123/119A) was described previously [23]. Note that because of an insertion in the GH123 p26 protein, positions 120, 160 and 179 from Ref. [23] correspond to positions 119, 159 and 178 in this work. PCR-based mutagenesis with primer pairs P1, P2 and P3, P4 was used to generate mutant GH123 carrying alanine 178 (GH123/178A). Mutant GH123 carrying serine 159 and 178A (GH123/159S-178A) was then constructed using primer pairs P1, P5 and P4, P6 with GH123/178A as a template. Mutant GH123 carrying 119A, 159S and 178A (GH123/119A-159S-178A) was constructed with primer pairs P1, P5 and P4, P6; and both GH123/119A and GH123/178A as templates. Infectious viruses were prepared by transfection of 293T cells with mutant proviral DNA clones.

P1:CTTCCCTGTACAACAGACA, P2:TTTACTGCTGCATCTGTTTGTTCTG,
P3:CAGAACAAACAGATGCAGCAGTAAA,
P4:GTGCAGCAAGTCCTCTGTC
P5:TAGCTCTGGAATGATTCCTTTGGTCC,
P6:GGACCAAAGGAATCATTCCAGAGCTA

3. Results

3.1. p26 CA variation in the Caio cohort.

HIV-2 positive subjects were selected from the Caio community cohort to assess HIV-2 sequence variation among a wide range of VL (≤100 to >280,000 copies/ml). PCR was attempted from a total of 92 subjects and p26 sequences were obtained from 69 subjects; 53 sequences were derived from viral RNA and 16 were from proviral DNA. Additional clinical features of these patients were previously described [14].

Caio p26 sequences were compared to either the Caio 69 p26 consensus sequence (Fig. 1 upper) or to the reference strain HIV-2_{ROD} p26 [31] (Fig. 1 lower) to identify amino acid polymorphisms in p26. HIV-2_{ROD} was chosen as a reference because the virus was isolated early in the epidemic in Cape Verde, islands historically linked to Guinea Bissau. Both comparisons show a similar pattern with variant amino acid sites clustering outside of the alpha helices (marked in black) required for CA folding (Fig. 1). Variations that associate with high or low VL (Mann–Whitney test, $p < 0.05$, except for position 178) are highlighted with a dotted line. The polymorphisms at position 119, 159 and 178 were of special interest because they showed 35–70% variation from HIV-2_{ROD} and involved changes from the reference sequence proline, a change expected to strongly alter protein structure. The non-proline residues were either alanine (38 cases), glutamine (5), or glycine (1) at position 119, serine (29) or threonine (2) at 159, and alanine (18), serine (2), gultamine (1), or valine (1) at 178. Viruses in low VL patients often had proline at these three positions while in higher VL samples these sites were frequently occupied by

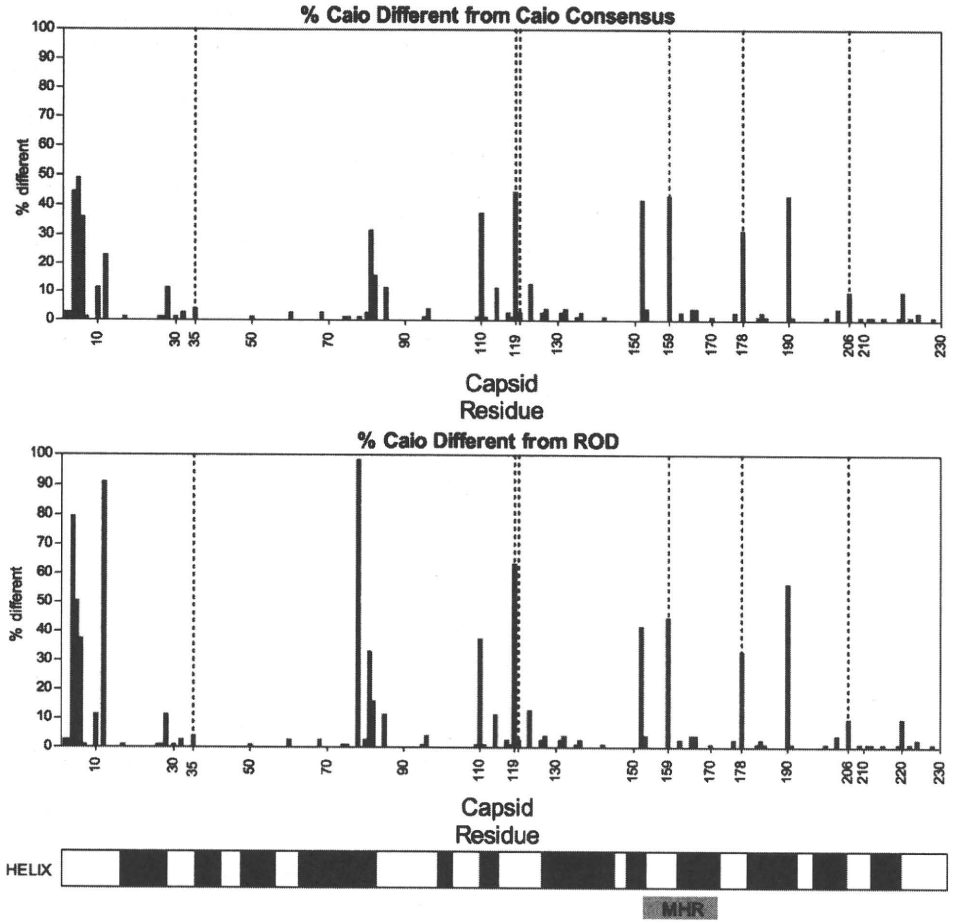


Fig. 1. Amino acid polymorphisms in the Caio HIV-2 p26. Polymorphisms in the p26 coding sequence are displayed as the percentage of the 69 sequences that differ at each position from either the Caio p26 consensus sequence (determined as the majority amino acid at each position, upper panel) or from the HIV-2_{ROD} p26 sequence (middle panel). Conserved alpha helices (in black) and the Major Homology Region (MHR) are indicated in the lower panel. Sites of variation that were associated with VL are indicated with dotted lines (Mann–Whitney test) ($p \leq 0.05$). [NB $p = 0.07$ for P178].

non-proine residues (Table 1). Three of the significantly varying positions (35, 120 and 206) were too infrequent for further study.

3.2. p26 CA variation correlating with VL

The association of proline 119, 159 and 178 with reduced VL becomes apparent when the log-transformed VL for each sample is plotted as a function of the total number of prolines at these three sites Fig. 2A. A Tobit regression analysis showed a clear relationship of increased VL with decreasing prolines in these three positions ($p = 0.003$).

To examine proline variation in more detail, CA types were grouped according the residue at each of the three positions using the code P (proline), N (not proline), or the wild card * (any amino acid) and ordered by increasing VL (Fig. 2B). Among the progression of median viral loads, there was a pattern of decreasing prolines, starting with the PPP group, through the intermediate forms to the complete non-proline NNN group (Fig. 2B). The two exceptions to this trend (the single proline NPN group had a lower median VL than the PNP and NPP groups) indicated that there may be interactions between specific combinations of these prolines in their effects on viral load but the number of examples was too small to demonstrate such effects statistically.

Considering the effect of single proline changes, proline 119 (P119) CAs (the P group) were isolated more frequently from patients with low VL with a 4.9-fold difference in the median VL of P** group compared to the N** group ($p = 0.0205$; Fig. 2B).

Proline 159 (P159) had a stronger effect on VL with at least 6.1-fold difference in the median VL of the *P* group compared to the *N* group ($p = 0.0075$; Fig. 2B). This site is of special interest being within the highly conserved MHR (Fig. 1) [20], essential for virion assembly.

Proline 178 (P178) showed modest variation with VL with only 3.5-fold difference in the median VL of the **P group compared to the **N group ($p = 0.0709$; Fig. 2B,). However, the presence of P178 was linked to the other two prolines: all CAs with both P119 and P159 had P178 (i.e. the PPP group = the PP* group, Fig. 2B). This may be due to a p26 folding requirement and/or genetic linkage.

Stronger associations were observed when the positions were analyzed in combination. The median VL in subjects with PP* viruses differed from NN* viruses by at least 13.6-fold ($p = 0.0028$, Fig. 2B). Importantly, median VL in subjects with viruses having all three prolines (PPP) compared to those lacking a proline at the sites (NNN) differed by at least 18.8-fold ($p = 0.0013$, Fig. 2B).

There are practical difficulties of sequencing circulating RNA genomes, especially from patients with <100 copies viral RNA per ml of plasma. Sixteen of the 69 sequences were derived from proviral DNA (listed as DNA in Table 1) because multiple attempts to obtain RNA sequence failed. To test if proline/VL associations were biased by the inclusion of these proviral sequences, the analysis was repeated after excluding the data from these 16 samples. The association of higher VL with non-proline residues remained significant at positions 119 ($p = 0.0123$) and 159 ($p = 0.0043$), and for the combined positions 119 + 159 ($p = 0.0012$) and 119 + 159 + 178

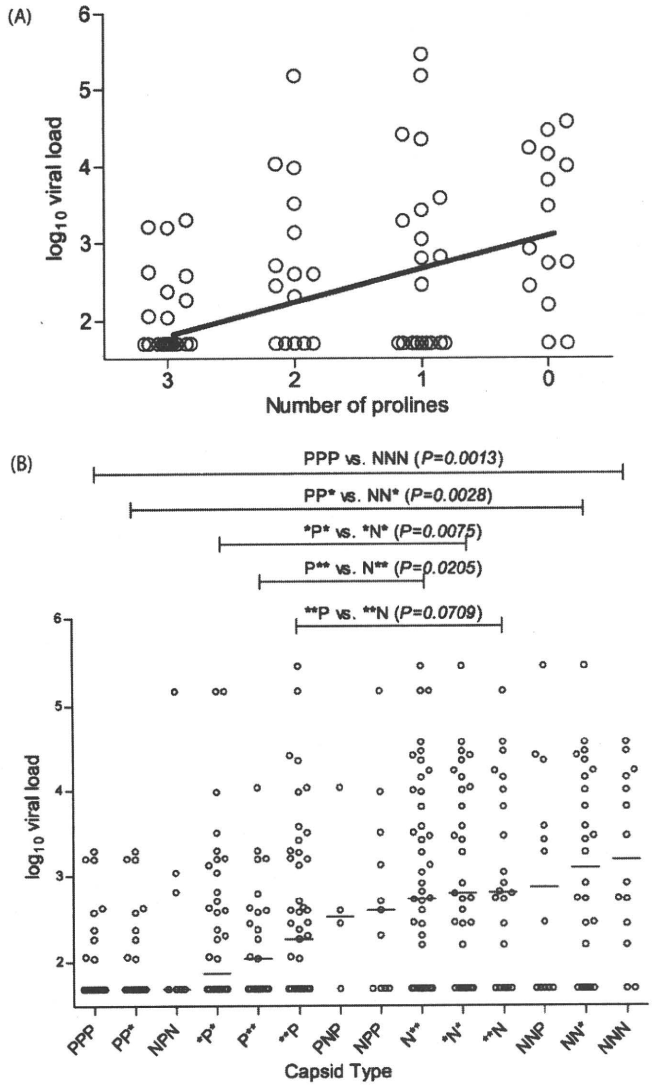


Fig. 2. HIV-2 VL correlates with amino acid variation at three p26 sites. (A) Relationship between VL and the number of prolines. Log-transformed VL for each sample was plotted as a function of the total number of prolines at p26 positions 119, 159 and 178. Tobit regression (Stata10, StatCorp TX, USA) was used to investigate the relationship between VL and the number of prolines at positions 119, 159 and 178. This form of regression is able to allow for censoring of viral loads below 100 in the dataset. The regression line was drawn with the equation $\log_{10}VL = 3.15 - 0.46 \times$ No. of prolines. (B) Relationship between VL and type of PPP motif. Log-transformed VL for each sample is plotted as a function of amino acid variation at p26 position 119, 159 and 178. Median VL are indicated by horizontal bars. The patients were stratified by the presence of proline (P) no proline residue (N) or any amino acid (*) at each of the positions 119, 159 and 178. For example, PPP = proline resides at positions 119, 159 and 178, NNN = no proline at the three positions. Comparisons of plasma VLs for different amino acid polymorphisms of particular interest were made using the non-parametric Mann–Whitney test (GraphPad Prism 5). *p*-values for these are shown above the figure.

($p = 0.0024$) (Table S1, lower panel). We conclude that independent of the origin of the sequences, there exists an association between low VL and proline residues at positions 119 and 159 independently, and with 119, 159 and 178 combined.

3.3. p26 CA variation influences susceptibility to TRIM5α

The TRIM5α was identified as a limit to cross-species retroviral infection [32–35]. This has led to a model of TRIM5α blocking retroviral infection by binding to the CA during entry, accelerating virus uncoating, and limiting subsequent steps in the infection process [32,34,36,37]. P119 was recently identified as a determinant of TRIM5α restriction [23] with HIV-2 CAs derived from TRIM5α sensitive viruses bearing a proline, and resistant strains having an alanine or glutamine at this site. It is possible that the reduced replication of the PPP viruses we observed in patients was part of the same phenomenon. Accordingly, the contribution of all three proline residues to TRIM5α restriction of replication was directly examined *in vitro*. Starting with the HIV-2 strain GH-123 as a PPP virus, P119, P159 plus P178, or all three prolines were altered to alanine or serine using site-directed mutagenesis. The growth of these variant viruses was compared in cells modified to express human TRIM5α; a parallel cell line expressing TRIM5α missing the SPRY domain, essential for p26 interaction, was used as a control to determine if these p26 changes altered virus replication independent of TRIM5α function [23]. Altering P119 (to produce APP) or P159 + P178 (to produce PSA) allowed 3-fold greater virus replication Fig. 3A. Alteration of all three prolines to ASA resulted in a 6-fold increase in virus replication. Thus the GH-123 variants displayed replication *in vitro* that closely mimic the behavior of the HIV-2 variants *in vivo*.

3.4. Phylogenetic analysis of p26 CA variation

One possible origin of the PPP form of p26 is that such a CA was encoded by a founder variant of HIV-2. The reduced growth of PPP viruses could be due to the p26 itself or to other shared and co-evolved features in these viruses. Alternately, changes at these three codons could occur more frequently. A phylogenetic analysis of Caio HIV-2 was performed to distinguish the PPP founder virus model from a multiple occurrence model. A founder effect with the appearance and spread of a PPP virus would appear as phylogenetic clustering of these variants. Ongoing selection for or against prolines in p26 would result in a phylogeny lacking PPP clustering. Because results derived from the 960 bp containing p26 coding could be dominated by variation in codons 119, 159 and 178, phylogeny was also inferred from a larger sequence spanning approximately 1300 bp of the envelope gene and including the highly variable V3 and V4 loops. The inferred phylogenies show that HIV-2 isolates encoding PPP p26 are distributed throughout the p26 and envelope trees (Fig. S1), supporting the conclusion that the occurrence of PPP p26 is not associated with a specific phylogenetic branch of Caio HIV-2 and is unlikely to be associated with a single occurrence of this CA motif. This conclusion is also supported by the high bootstrap values for some of the branches. These results are consistent with selection for and multiple appearance of the PPP p26 in the Caio population.

3.5. Modelling p26 CA sequence variation on structure and dimer formation

The bulky and constrained structure of proline strongly influences protein secondary structure; proline is inimical to alpha-helices and can kink otherwise flexible loops [38]. Thus polymorphisms involving prolines residues could alter the p26 structure. Using homology modelling, three-dimensional structures of six of the Caio HIV-2 p26 molecules (2 PPP, 2 ASA and 2 intermediate forms, APP and ASP) were constructed. The thermodynamically optimized 3-D structure models showed that the HIV-2 p26 consists of two packed core structures of N-terminal and C-terminal domains, a similar conformation to HIV-1 p24 [39]. Superimposition of the six HIV-2 p26 models showed that the overall 3-D structures of the variants were very similar with an exception: the amino acid substitution at position 119 from proline to alanine induced marked changes in the configuration of the loop between helices 6 and 7, as found previously with HIV-2 p26 N-terminal domain model [23]. In contrast, the substitution

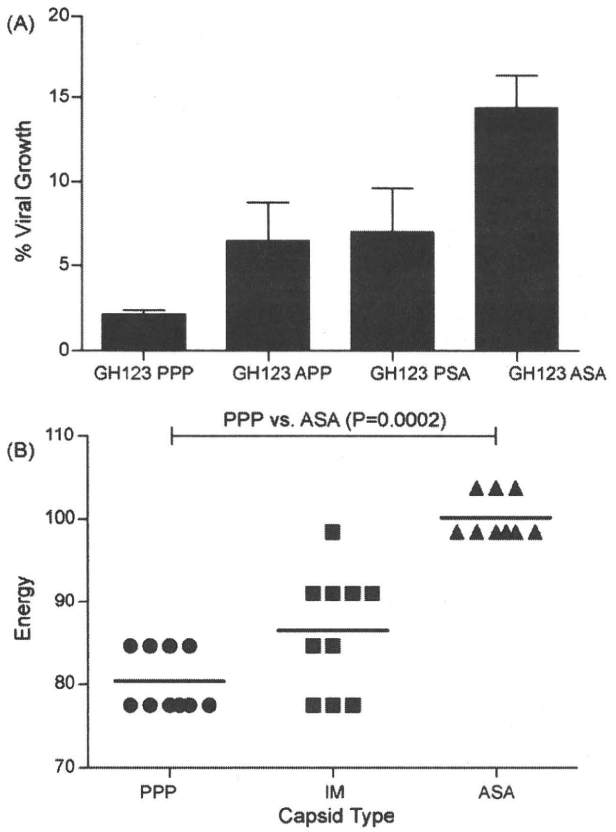


Fig. 3. HIV-2 capsid changes alter TRIM5α susceptibility and capsid stability. (A) Growth of HIV-2 GH123 (PPP) and its mutant viruses HIV-2 GH123/119A (APP), HIV-2 GH123/159S-178A (PSA), and HIV-2 GH123/119A-159S-178A (ASA) in the presence of human TRIM5α. MT4 cells (10^5) were infected with Hu-TRIM5α-SeV or CM-SPRY(-)-SeV at a multiplicity of infection of 10 plaque forming units per cell. 9 h after infection, the cells were superinfected with 20 ng of p26 of HIV-2 GH123, HIV-2 GH123/119A, HIV-2 GH123/159S-178A, or HIV-2 GH123/119A-159S-178A viruses. The culture supernatants were collected 6 days after infection, and the level of p26 was measured by using a RETROtek antigen ELISA kit (ZeptoMetrix Corp., Buffalo, NY). Ratios of HIV-2 CA levels of Hu-TRIM5α-SeV-infected cells to those of CM-SPRY(-)-SeV-infected cells are shown as percent growth. Error bars denote standard deviations in quadruplicate samples. A Kruskal–Wallis test for the entire data set clearly detected the difference of viral growth among those four viruses ($p = 0.006$). Furthermore, comparison of each mutant virus with GH123(PPP) using the Dunnett test also showed statistically significant differences of PPP vs. APP $p < 0.05$, PPP vs. PSA $p < 0.05$, and PPP vs. ASA $p < 0.01$. (B) Dimer binding energies as a function of capsid type. Samples were grouped into PPP, IM (intermediate, with proline at one or two of the three sites) or ASA, based on the amino acid at position 119, 159 and 178. The capsid dimer binding energies (absolute value) for each sequence were determined by homology modelling (see Section 2); the mean values for each group are indicated. Comparison between the PPP and the ASA group using the non-parametric Mann–Whitney test (GraphPad Prism 5) provided the indicated *p*-value.

at positions 159 or 178 induced no major changes in the main-chain backbone of HIV-2 CA (data not shown), suggesting that the structure of the CA can accommodate alternative residues at these sites.

Positions 159 and 178 are located in the C-terminal portion of p26 required for dimer formation as well as virion shell assembly [18]. Homology modelling of the p26 dimer was used to calculate binding energy for dimer formation. Since the C-terminal sequences used to prepare the initial six structures were common to a larger set of p26 sequences, dimer binding energies for 29 of the Caio CAs could be calculated. These calculations revealed that the PPP p26 dimers had weaker binding energies, the ASA dimers had stronger binding energies and the intermediate forms with only one or two of the prolines altered (IM) had intermediate binding energies (Fig. 3B). In addition, the viral loads of the 29 patients were modestly correlated with the absolute values of the binding energy of the viral CA ($r_s = 0.383$, $p = 0.040$). We conclude that the amino acid changes at positions 159 and 178 influence p26 dimer stability, with ASA CA dimers having a higher stability than PPP or APP CA dimers and the increased stability may be related to elevated VL in these patients.

4. Discussion

This study represents a detailed examination of HIV-2 p26 sequence variation within a set of 69 sequences isolated from subjects with both high and low VL. Disease progression after HIV-2 infection is highly dependent on VL [40] with subjects who control HIV-2 replication continuing to do this over a period of many years, suggesting that virus-host interactions result in a stable set-point of virus replication. The current study identified three sites of p26 variation correlating with VL. This study reveals a previously undescribed pattern of variation in the highly conserved p26 and indicates that the outcome of HIV-2 infection is partially predicted by the form of the p26 carried by the virus. In addition to confirming the importance of P119 as a determinant of TRIM5α restriction, the current study identified two additional amino acid positions (159 and 178) whose identities correlate with virus load. Although these associations were not significant after stringent adjustment for multiple comparisons, HIV-2 encoding p26 specifically modified at these three positions showed *in vitro* replication levels consistent with the *in vivo* VL data and further supported the conclusion that the three residues 119, 159 and 178 are important determinants of virus growth and influence TRIM5α restriction (Fig. 3).

Virus replication *in vivo* is influenced by a large number of host and viral factors and it would be naïve to conclude that these three residues are the sole determinants. That additional factors influence the course of infection is reflected in the VL data and the exceptions from the pattern (e.g. the three PPP viruses with greater than 10e3 VL and the 2 NNN viruses with undetectable VL). However, the power of such a population study is that patterns of HIV-2 behavior appear when large numbers of infection are monitored. As shown in Fig. 2, the substantial and statistically significant change in VL that accompanies the variation from PPP to NNN viruses strongly supports our conclusions that this p26 motif is an important determinant of the course of infection.

How might the PPP motif function? Our data support a destabilization of the CA by the three proline residues. P119 may directly form a recognition signal for TRIM5α binding, and the three proline residues may result in less tightly packed core that is more readily dismantled and processed after TRIM5α recognition. Our preliminary immunological studies show that patients with PPP virus mount stronger T cell responses to p26 and to the entire HIV-2 proteome (A.L., S.R.J. unpublished results) and this increased immune exposure might be a consequence of TRIM5α recognition and more efficient antigen presentation.

These results are consistent with TRIM5α restriction playing a direct role in limiting HIV-2 replication and a more indirect role in enhancing the immune response to the virus. The *in vitro* studies (Fig. 3A), although using manipulated cells and monitoring virus replication only over a short period of replication, demonstrated that variation in these three CA residues influence the susceptibility of HIV-2 replication to TRIM5α. *In vivo*, it is likely that TRIM5α effects are both manifested over multiple rounds of infection and TRIM5α may cooperate with other processes such as the adaptive immune response; *in vitro* cell culture conditions and growth in immortalized cell lines are unlikely to fully recreate these processes. We believe that our *in vivo* virus load data are the strongest support for the hypothesis that the PPP motif modulates VL.

If our hypothesis is correct, incident infections in Caio, infected patients that progress to require anti-retroviral therapy and moth-