STRUCTURAL BIOLOGY

An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation

Florian K. M. Schur,[1,2] Martin Obr,[2,3] Wim J. H. Hagen,[1] William Wan,[1] Arjen J. Jakobi,[1,4] Joanna M. Kirkpatrick,[5*] Carsten Sachse,[1] Hans-Georg Kräusslich,[2,3] John A. G. Briggs[1,2†]

Immature HIV-1 assembles at and buds from the plasma membrane before proteolytic cleavage of the viral Gag polyprotein induces structural maturation. Maturation can be blocked by maturation inhibitors (MIs), thereby abolishing infectivity. The CA (capsid) and SP1 (spacer peptide 1) region of Gag is the key regulator of assembly and maturation and is the target of MIs. We applied optimized cryo-electron tomography and subtomogram averaging to resolve this region within assembled immature HIV-1 particles at 3.9 angstrom resolution and built an atomic model. The structure reveals a network of intra- and intermolecular interactions mediating immature HIV-1 assembly. The proteolytic cleavage site between CA and SP1 is inaccessible to protease. We suggest that MIs prevent CA-SP1 cleavage by stabilizing the structure, and MI resistance develops by destabilizing CA-SP1.

The major structural protein of HIV-1, Gag, oligomerizes at the plasma membrane of infected cells, leading to membrane bending and release of immature virus particles. Ordered cleavage of the Gag polyprotein at five sites by the viral protease (PR) then causes a dramatic structural rearrangement of the virus to produce the mature, infectious virion (fig. S1, A to C). Gag-Gag interactions in the immature virus are mediated by the adjacent CA (capsid) domain and SP1 (spacer peptide 1). After maturation, CA forms a conical core encapsulating the viral RNA-nucleoprotein complex (*1*). The final proteolytic cleavage in Gag occurs between CA and SP1. Even small remnants of uncleaved CA-SP1 have a dominant-negative effect on infectivity (*2*, *3*). Cleavage at this site is inhibited by maturation inhibitors (MIs) (fig. S1B) (*4*), and polymorphisms in this region cause resistance against the first-in-class MI bevirimat (BVM) (*5*). Other MIs inhibiting CA-SP1 cleavage are currently undergoing clinical trials, but the precise mechanisms of their inhibitory action and of resistance against them are unclear.

Unprocessed Gag is assembled into irregular curved lattices whose structures cannot be determined using conventional structural biology techniques. The best structural model for CA in the immature virus has been derived by fitting nuclear magnetic resonance and crystal structures of the folded N-terminal and C-terminal domains (CA-NTD and CA-CTD, respectively) into a structure of the immature CA lattice acquired through cryo-electron tomography and subtomogram averaging (*6*). Subtomogram averaging can resolve protein structures within complex environments ranging from cells to enveloped viruses (*7*), but it has been limited to resolutions of ~8 Å. The positions of α-helices are visible at this resolution, but those of amino acids are not. Furthermore, 8 Å resolution is not sufficient to generate ab initio structural models for unknown structures such as the Gag regions downstream of the crystallized CA-CTD, which ends at residue 220 of the 231–amino acid CA domain (fig. S1D). This downstream region consists of a sequence of amino acids that is thought to assemble into a flexible hinge, followed by a sequence that is predicted to form a six-helix bundle spanning the C-terminal residues of CA, the critical CA-SP1 cleavage site, and the N-terminal residues of SP1 (*8*). This region includes the majority of the amino acids that are essential for virus assembly (*9*–*11*). Obtaining a high-resolution structure of CA-SP1 in the immature arrangement is vital for a mechanistic understanding of HIV-1 assembly, maturation, and inhibition.

The HIV-1 Gag construct ΔMACANCSP2 (MA, matrix; NC, nucleocapsid) (*11*) assembles into immature virus-like particles (VLPs) in vitro (Fig. 1A and fig. S1C). We assembled ΔMACANCSP2 particles in the absence or presence of 100 μg/ml BVM. We also purified intact, immature HIV-1 virus particles carrying an inactivating PR mutation, D25A. For all three cases, we used an optimized data collection scheme to acquire cryo-electron tomography tilt series (table S1)

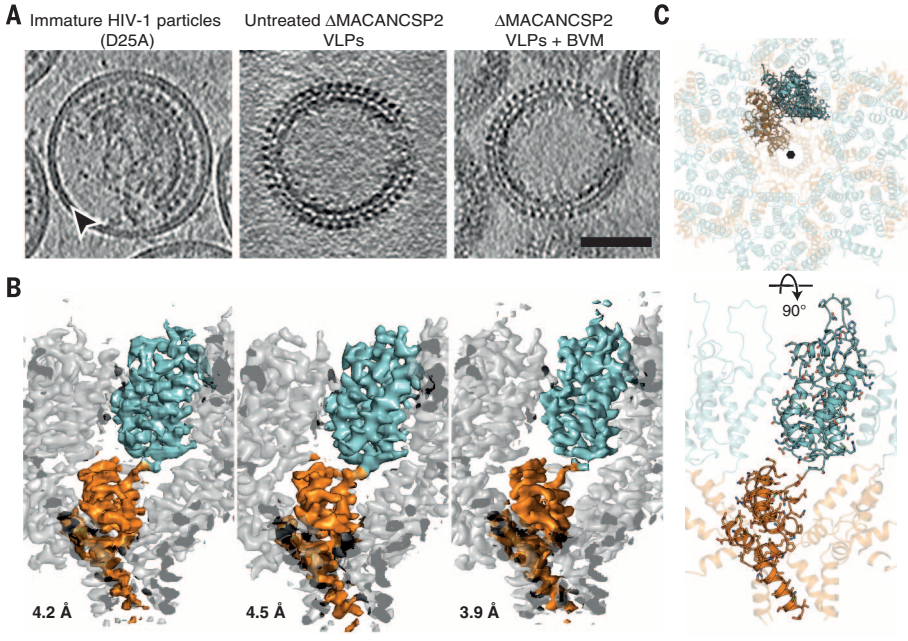

Fig. 1. Structure of the immature HIV-1 CA-SP1 lattice at 3.9 Å. (**A**) Computational slices through tomograms of immature HIV-1 (D25A mutant), untreated ΔMACANCSP2 VLPs, and BVM-treated ΔMACANCSP2 VLPs. The arrowhead marks the membrane bilayer in the left panel. Scale bar, 50 nm. (**B**) Electron densities of CA-SP1 from the samples shown in (A), viewed perpendicular to the lattice, generated by subtomogram averaging. One CA-SP1 monomer is highlighted in color, with the CA-NTD in cyan and the CA-CTD and SP1 in orange. The resolutions of the determined structures are noted. (**C**) The refined atomic model, viewed from outside of the virus (top) and rotated by 90°, shown in the same view as in (B) (bottom). The sixfold symmetry axis is indicated with a hexagon.

[1]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany. [2]Molecular Medicine Partnership Unit, European Molecular Biology Laboratory–Universitätsklinikum Heidelberg, Heidelberg, Germany. [3]Department of Infectious Diseases, Virology, Universitätsklinikum Heidelberg, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany. [4]Hamburg Unit c/o DESY (Deutsches Elektronen-Synchrotron), European Molecular Biology Laboratory, Notkestraße 85, 22607 Hamburg, Germany. [5]Proteomics Core Facility, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany.
*Present address: Leibniz Institute on Aging–Fritz Lipmann Institute, Beutenbergstraße 11, 07745 Jena, Germany.
†Corresponding author. Email: john.briggs@embl.de

(*12*, *13*). VLPs with or without BVM were indistinguishable, showing densities for the Gag lattice that were similar to those in the immature virus particles (Fig. 1A). Subtomogram averaging was performed independently for each data set, using an optimized workflow with features including frame-based motion correction (*14*) and exposure filtering (*15*). The resolutions of the CA-SP1 layer in the final averages for untreated VLPs, BVM-treated VLPs, and immature virus particles were 4.5, 3.9, and 4.2 Å, respectively (Fig. 1B and fig. S2, A and B).

We compared the three structures and found no clear differences in the protein densities of CA or SP1; they varied only in the presence or absence of densities at the center of the hexamer near SP1 (fig. S2, C to I, and fig. S3C). Given the high degree of structural similarity between the maps, we used the 3.9 Å structure to build and refine a complete atomic model for the CA-SP1 region from Gag residues 148 to 371 (Figs. 1C and 2, figs. S3 and S4, table S2, and movies S1 and S2). Residues 356 to 371, covering the C terminus of CA and the first eight residues of SP1, assemble into a six-helix bundle.

Interactions in the CA-NTD layer are described in fig. S4A. The CA-NTD and CA-CTD contact one another in the region of E160 and E161 in the CA-NTD and Q308 in the CA-CTD. The relative positions of the two domains may also be restricted by the extended rigid linker between them, which appears to be stabilized by an interaction with R305 in helix 8. Y277, the last residue of helix 7 in a solution structure (*16*) and in the mature-like CA hexamer (*17*), is rotated out of the helix and packs against the linker at P279 (Fig. 2B). This network of interactions may allow the CA-CTD to be structurally modulated by cleavage upstream of the CA-NTD, and vice versa.

Highly conserved residues of the major homology region (MHR) in the CA-CTD abut the extended chain between the 3_{10}-helix and helix 8 (Fig. 2C). Charged residues in the 3_{10}-helix and extended chain interact with the neighboring CA-CTD molecule within the hexamer (Fig. 2, B and C). Hexamers are linked by a CA-CTD dimer interface (fig. S4B) formed by residues W316 and M317 (*6*, *18*, *19*).

Downstream of helix 11, beyond the residues that are resolved in crystal structures (*19*, *20*), the flexible hinge region (referred to here as the VGG hinge; residues 353 to 355) unexpectedly adopts a rigid structure within the lattice. P356 then marks the start of a helix (referred to here as the CA-SP1 helix) that extends down to residue 371 in SP1 (Fig. 2, A and D) before abruptly ending, indicating that residues C-terminal of 371 are disordered (movie S1). The CA-SP1 helix protrudes up into the CA-CTD layer, where the top of the helix and the VGG hinge from one Gag molecule pack tightly against the CA-CTD from the neighboring Gag molecule. The CA-CTD, VGG hinge, and CA-SP1 helix thus form a single, integrated assembly unit that defines the structure of the hexamer (Fig. 2, D and G).

This assembly unit appears to be stabilized by a three-way interaction between H358 in the CA-SP1 helix, D329 in the base of helix 10, and P356 in the CA-SP1 helix of the neighboring CA molecule (Fig. 2E). K290 in the loop upstream of helix 8 and K359 in the CA-SP1 helix protrude from above and below these residues toward the center of the six-helix bundle, where they coordinate a density, presumably a negatively

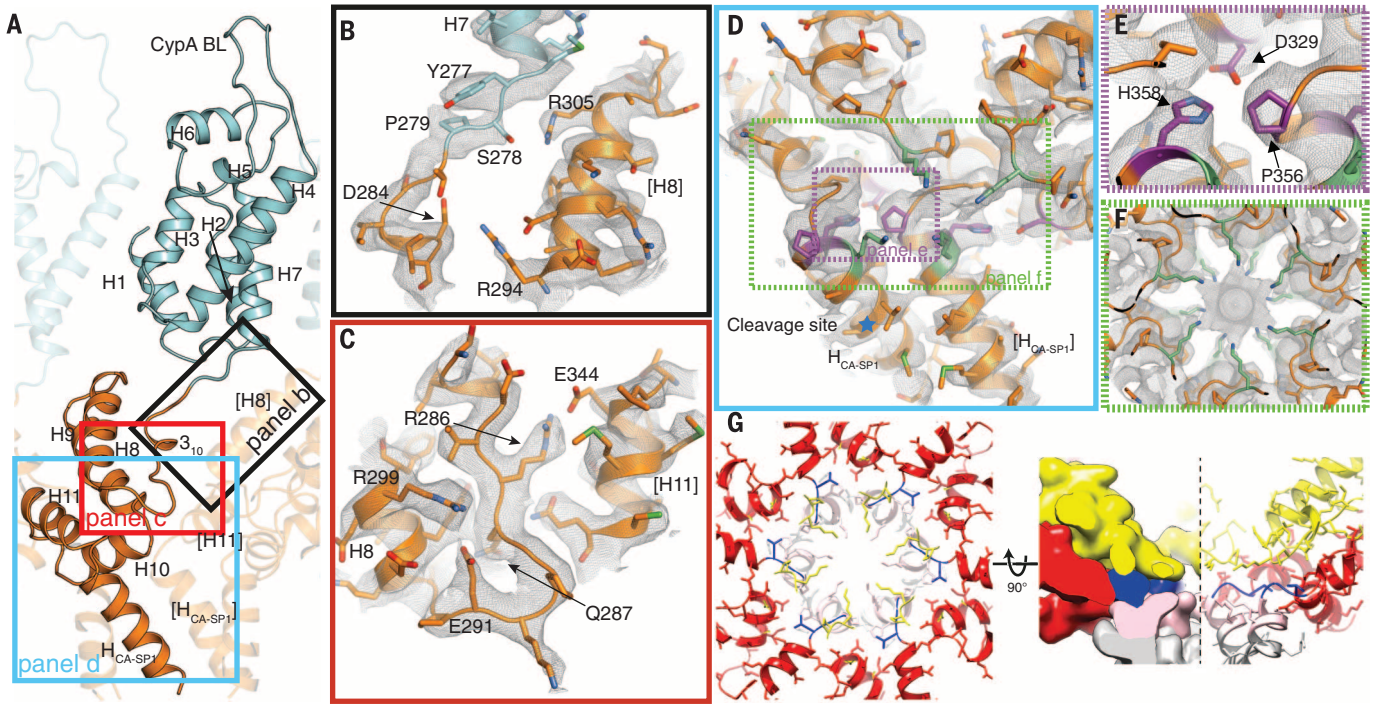

Fig. 2. Structural features in the immature CA-SP1 lattice. (**A**) A single CA-SP1 monomer, as in Fig. 1C. α-helices (H1 to H11 and H_{CA-SP1}) and the cyclophilin A binding loop (CypA BL) are labeled (α-helices from neighboring CA monomers are shown in brackets). Colored rectangles indicate regions enlarged in (B) to (D). (**B**) The CA-NTD–CA-CTD linker is in an extended conformation, with Y277 binding to the linker and S278 approaching R305. (**C**) The highly conserved residues in the MHR (Q287, E291, and R299) stabilize the linker connecting the 3_{10}-helix and helix 8. Residues in this linker can interact with an adjacent CA monomer around the hexameric ring (e.g., R286 with E344 and D284 with R294); point mutations of these residues do not abolish assembly (*22*), suggesting some redundancy in these interactions. (**D**) The CA-CTD, the VGG hinge, and the top of the CA-SP1 helix form an integrated structural assembly unit. The CA-SP1 cleavage site is marked by a blue star. Dashed rectangles indicate the approximate positions of the regions enlarged in (E) and (F). Residues are colored as in (E) and (F). (**E**) Residues D329, P356, and H358 (in purple) form a three-way linkage between two neighboring CA-SP1 helices and the base of the CA-CTD. (**F**) K290 and K359 (in green) protrude from above and below the region shown in (E) to the center of the hexamer, where they coordinate a strong density. (**G**) Horizontal (left) and vertical (right) slabs through the structure illustrate that the MHR (yellow), other residues in the CA-CTD base (red), the VGG hinge (blue), and the top of the CA-SP1 helix (pink) come together to form the hexameric assembly unit. In the vertical slab, one-half of the hexamer is represented in a surface view. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; G, Gly; H, His [except when used to indicate helices, as in (A) to (D)]; K, Lys; L, Leu; M, Met; P, Pro; Q, Gln; R, Arg; S, Ser; V, Val; W, Trp; and Y, Tyr.

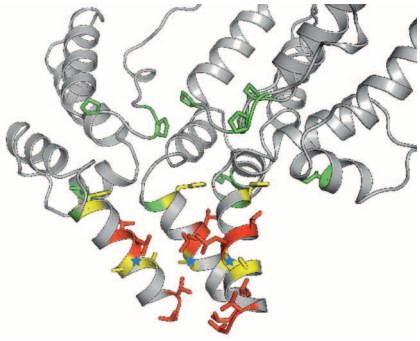

Fig. 3. Mutations that confer resistance to MIs destabilize the immature lattice. Resistance mutations or naturally occurring polymorphisms that render HIV-1 resistant to BVM (red) or PF-46396 (green) have been mapped onto the atomic model. Mutations that confer resistance for both compounds are colored yellow. The CA-SP1 cleavage site is marked by blue stars. For clarity, only three helices of the six-helix bundle are displayed.

charged ion cluster (Fig. 2F and movie S2). This arrangement is reminiscent of the six arginine residues that protrude into the center of the NTD hexamer in mature HIV CA (*17*, *21*). Twelve essential amino residues have been identified in the CA-CTD, where mutation to alanine abolishes virus assembly (*22*, *23*). These include W316 and M317 in the hydrophobic dimeric interface; V353, G354, and G355 in the VGG hinge; K290, D329, P356, H358, and K359, which together form the intricate network of interactions that defines the assembly unit; and A360 and L363, which appear to have hydrophobic interactions within the six-helix bundle. There is therefore a close correlation between the sensitivity of a residue to mutation and whether it has a role in mediating interactions in the CA-SP1 assembly unit, confirming the importance of these interactions in virus assembly.

During maturation, the final proteolytic cleavage occurs between L363 and A364. In our structure, this site is in the middle of the CA-SP1 helix bundle, where it is inaccessible to PR, which acts on extended protein chains (*24*) (Fig. 2D and movie S2). Disassembly of the immature lattice and full cleavage between CA and SP1 only take place once cleavage has occurred both between MA and CA and between SP1 and NC (*10*, *25*). Together, these observations support a model in which the final step in maturation is regulated by limiting the access of PR to its substrate: Cleavage upstream of CA and downstream of SP1 together destabilize the immature lattice and the CA-SP1 helix, thereby allowing PR to cleave between CA and SP1.

MIs block HIV infection by preventing cleavage at the CA-SP1 site (*4*, *13*, *26*) (fig. S5) and may also stabilize the immature lattice (*27*, *28*). We mapped mutations, deletions, and polymorphisms that confer resistance to BVM (*5*, *29*, *30*) and PF-46396 [another MI (*31*, *32*)] onto our structural model (Fig. 3). This revealed that the positions of resistance mutations do not map out potential drug-binding pockets. Instead, they are located at protein-protein interfaces within the CA-SP1 lattice. Together with our observation that the CA-SP1 cleavage site is inaccessible in the immature virus, this implies that the mode of action of MIs is not steric inhibition of proteolysis but, instead, stabilization of the immature Gag lattice. Our data suggest that BVM stabilizes the lattice by binding to a site in the center of the six-helix bundle (fig. S2). The reduced cleavage at the CA-SP1 boundary is a downstream effect of stabilization because cleavage of this site requires unfolding to an extended chain. HIV-1 appears to develop MI resistance by destabilizing its immature form, thus directly counteracting the stabilizing effects of MIs.

Note added in proof: In a concurrent publication, Wagner *et al.* report a crystal structure of a CA-CTD-SP1 construct that adopts a similar conformation to the structure reported here (*33*). They also report that the CA-SP1 cleavage site is within a six-helix bundle and suggest that MIs prevent proteolytic cleavage by stabilizing this structure.

REFERENCES AND NOTES

1. W. I. Sundquist, H. G. Kräusslich, *Cold Spring Harb. Perspect. Med.* **2**, a006924 (2012).
2. B. Müller *et al.*, *J. Biol. Chem.* **284**, 29692–29703 (2009).
3. M. A. Checkley, B. G. Luttge, F. Soheilian, K. Nagashima, E. O. Freed, *Virology* **400**, 137–144 (2010).
4. C. S. Adamson, K. Salzwedel, E. O. Freed, *Expert Opin. Ther. Targets* **13**, 895–908 (2009).
5. C. S. Adamson, M. Sakalian, K. Salzwedel, E. O. Freed, *Retrovirology* **7**, 36 (2010).
6. F. K. Schur *et al.*, *Nature* **517**, 505–508 (2015).
7. J. A. Briggs, *Curr. Opin. Struct. Biol.* **23**, 261–267 (2013).
8. M. A. Accola, S. Höglund, H. G. Göttlinger, *J. Virol.* **72**, 2072–2078 (1998).
9. S. A. Datta *et al.*, *J. Virol.* **85**, 4111–4121 (2011).
10. A. de Marco *et al.*, *PLOS Pathog.* **6**, e1001215 (2010).
11. I. Gross *et al.*, *EMBO J.* **19**, 103–113 (2000).
12. W. J. Hagen, W. Wan, J. A. Briggs, *J. Struct. Biol.* 10.1016/j.jsb.2016.06.007 (2016).
13. Materials and methods are available as supplementary materials on *Science* Online.
14. X. Li *et al.*, *Nat. Methods* **10**, 584–590 (2013).
15. T. Grant, N. Grigorieff, *eLife* **4**, e06980 (2015).
16. C. Tang, Y. Ndassa, M. F. Summers, *Nat. Struct. Biol.* **9**, 537–543 (2002).
17. A. T. Gres *et al.*, *Science* **349**, 99–103 (2015).
18. J. A. Briggs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11090–11095 (2009).
19. V. Bartonova *et al.*, *J. Biol. Chem.* **283**, 32024–32033 (2008).
20. T. R. Gamble *et al.*, *Science* **278**, 849–853 (1997).
21. O. Pornillos *et al.*, *Cell* **137**, 1282–1292 (2009).
22. U. K. von Schwedler, K. M. Stray, J. E. Garrus, W. I. Sundquist, *J. Virol.* **77**, 5439–5450 (2003).
23. D. Melamed *et al.*, *J. Virol.* **78**, 9675–9688 (2004).
24. M. Prabu-Jeyabalan, E. Nalivaika, C. A. Schiffer, *Structure* **10**, 369–381 (2002).
25. K. Wiegers *et al.*, *J. Virol.* **72**, 2846–2854 (1998).
26. M. Sakalian *et al.*, *J. Virol.* **80**, 5716–5722 (2006).
27. P. W. Keller *et al.*, *J. Virol.* **87**, 13655–13664 (2013).
28. P. W. Keller, C. S. Adamson, J. B. Heymann, E. O. Freed, A. C. Steven, *J. Virol.* **85**, 1420–1428 (2011).
29. C. S. Adamson *et al.*, *J. Virol.* **80**, 10957–10971 (2006).
30. N. A. Margot, C. S. Gibbs, M. D. Miller, *Antimicrob. Agents Chemother.* **54**, 2345–2353 (2010).
31. K. Waki *et al.*, *PLOS Pathog.* **8**, e1002997 (2012).
32. W. S. Blair *et al.*, *Antimicrob. Agents Chemother.* **53**, 5080–5087 (2009).
33. J. M. Wagner *et al.*, *eLife* **5**, e17063 (2016).

ACKNOWLEDGMENTS

The authors thank B. Glass for preparation of the immature HIV-1 (D25A) sample; J. Plitzko and D. Tegunov for providing the K2Align software; and S. Mattei, N. Hoffman, F. Thommen, A. Sonnen, and S. Dodonova for technical assistance and/or discussion. This study was supported by Deutsche Forschungsgemeinschaft grants BR 3635/2-1 (to J.A.G.B.) and KR 906/7-1 (to H.-G.K.). The Briggs laboratory acknowledges financial support from the European Molecular Biology Laboratory (EMBL) and from the Chica und Heinz Schaller Stiftung. W.W. was supported by a European Molecular Biology Organization Long-Term Fellowship (ALTF 748-2014). A.J.J. acknowledges support by the EMBL Interdisciplinary Postdoc Program under the Marie Curie Action COFUND (PCOFUND-GA-2008-229597) and by the Joachim Herz Stiftung. This study was technically supported by the EMBL information technology services unit and the EMBL Proteomics Core Facility. F.K.M.S., M.O., H.-G.K., and J.A.G.B. designed the experiments, with J.M.K. assisting in the design of those involving mass spectrometry. F.K.M.S. and M.O. prepared samples. W.J.H.H. implemented tomography acquisition schemes. F.K.M.S. and W.J.H.H. acquired the data. F.K.M.S. and W.W. processed images. F.K.M.S., A.J.J., and C.S. refined the model. F.K.M.S., M.O., and J.A.G.B. analyzed the data. F.K.M.S. and J.A.G.B. wrote the manuscript with support from all authors. Representative tomograms and the final electron microscopy structures have been deposited in the Electron Microscopy Data Bank with accession numbers EMD-4015, EMD-4016, EMD-4017, EMD-4018, EMD-4019, and EMD-4020. The refined HIV-1 CA-SP1 model has been deposited in the Protein Data Bank with accession number 5L93.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/353/6298/506/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References (*34*–*59*)
Movies S1 and S2

26 April 2016; accepted 29 June 2016
Published online 14 July 2016
10.1126/science.aaf9620

Editor's Summary

Maturation and inhibition of HIV-1

HIV-1 undergoes a two-step assembly process controlled largely by a single region of its Gag protein. Schur *et al.* determined a complete atomic model for this region within an assembled Gag protein lattice using cryo-electron tomography together with subtomogram averaging. Amino acids from different parts of multiple Gag molecules come together to form an intricate network of interactions that drive HIV-1 assembly. The final step of maturation into the infectious HIV-1 virus is controlled by structural changes in Gag that alter the accessibility of the final cleavage site to the viral protease.

Science, this issue p. 506