AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

STRUCTURE AND ASSEMBLY

Conformational States of a Soluble, Uncleaved HIV-1 Envelope Trimer

Yuhang Liu,[a,b]* Junhua Pan,[a,c] Yongfei Cai,[a] Nikolaus Grigorieff,[b] Stephen C. Harrison,[a,c,d] Bing Chen[a]

Division of Molecular Medicine, Boston Children's Hospital, and Harvard Medical School, Boston, Massachusetts, USA[a]; Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, Virginia, USA[b]; Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA[c]; Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts, USA[d]

ABSTRACT The HIV-1 envelope spike [Env; trimeric $(gp160)_3$ cleaved to $(gp120/gp41)_3$] induces membrane fusion, leading to viral entry. It is also the viral component targeted by neutralizing antibodies. Vaccine development requires production, in quantities suitable for clinical studies, of a recombinant form that resembles functional Env. HIV-1 gp140 trimers—the uncleaved ectodomains of $(gp160)_3$—from a few selected viral isolates adopt a compact conformation with many antigenic properties of native Env spikes. One is currently being evaluated in a clinical trial. We report here low-resolution (20 Å) electron cryomicroscopy (cryoEM) structures of this gp140 trimer, which adopts two principal conformations, one closed and the other slightly open. The former is indistinguishable at this resolution from those adopted by a stabilized, cleaved trimer (SOSIP) or by a membrane-bound Env trimer with a truncated cytoplasmic tail (EnvΔCT). The latter conformation is closer to a partially open Env trimer than to the fully open conformation induced by CD4. These results show that a stable, uncleaved HIV-1 gp140 trimer has a compact structure close to that of native Env.

IMPORTANCE Development of any HIV vaccine with a protein component (for either priming or boosting) requires production of a recombinant form to mimic the trimeric, functional HIV-1 envelope spike in quantities suitable for clinical studies. Our understanding of the envelope structure has depended in part on a cleaved, soluble trimer, known as SOSIP.664, stabilized by several modifications, including an engineered disulfide. This construct, which is difficult to produce in large quantities, has yet to induce better antibody responses than those to other envelope-based immunogens, even in animal models. The uncleaved ectodomain of the envelope protein, called gp140, has also been made as a soluble form to mimic the native Env present on the virion surface. Most HIV-1 gp140 preparations are not stable, however, and have an inhomogeneous conformation. The results presented here show that gp140 preparations from suitable isolates can adopt a compact, native-like structure, supporting its use as a vaccine candidate.

KEYWORDS cryoEM, envelope, human immunodeficiency virus, immunogen

The HIV-1 envelope glycoprotein (Env) initiates infection by inducing fusion of viral and target cell membranes (1). The mature, functional Env spikes $[(gp120/gp41)_3]$ are derived from a trimeric precursor, $(gp160)_3$, after cleavage by a furin-like protease into a receptor-binding fragment, gp120, and a fusionogenic fragment, gp41 (1). Env is the major viral B-cell antigen. To induce fusion, HIV-1 gp120 must bind first to the host primary receptor, CD4, and then to a coreceptor (e.g., CCR5 or CXCR4). These binding events promote large structural rearrangements, including an irreversible refolding of gp41 that effectively brings the two membranes together, leading to membrane fusion and viral entry (2, 3). During this process, the conformational plasticity of Env allows it

Received 31 January 2017 **Accepted** 21 February 2017

Accepted manuscript posted online 1 March 2017

Citation Liu Y, Pan J, Cai Y, Grigorieff N, Harrison SC, Chen B. 2017. Conformational states of a soluble, uncleaved HIV-1 envelope trimer. J Virol 91:e00175-17. https://doi.org/10.1128/JVI.00175-17.

Editor Wesley I. Sundquist, University of Utah

Address correspondence to Bing Chen, bchen@crystal.harvard.edu.

* Present address: Yuhang Liu, Department of Structure Biology and Biophysics, Pfizer Inc., Groton, CT, USA.

Y.L. and J.P. contributed equally to this article.

to adopt multiple conformational states with large-scale differences among them: an untriggered prefusion state, a series of fusion intermediate states, and a six-helix-bundle, postfusion conformation of the gp41 fragment.

The prefusion conformation of Env is an appropriate candidate for vaccine development because it is recognized by most broadly neutralizing antibodies (bnAbs) (4–6). Thus, production of a recombinant form that closely resembles the conformation of Env spikes on the surfaces of virions is one important objective in Env-based B-cell immunogen design. Progress during the past few years in our understanding of Env trimer structure has depended in part on SOSIP.664, a soluble form that is stabilized by several modifications, including an added disulfide bond between gp120 and gp41, an Ile-to-Pro mutation in gp41, deletion of the membrane-proximal external region (MPER), and replacement of the furin cleavage site with a string of six arginines (6, 7). Its structure, determined by both electron cryomicroscopy (cryoEM) and X-ray crystallography (8–10), and a structure determined by cryoEM of a detergent-solubilized clade B Env with the cytoplasmic tail (CT) deleted and with stabilizing Fab bound (JR-FL EnvΔCT) (11), are very similar to each other. Both of these Env constructs are fully cleaved, with the former stabilized by engineered modifications and the latter by bound antibodies. Although they both have compact structures and certain desired antigenic properties, large-scale production of these cleaved Env trimers with the quantity, stability, and homogeneity required for a protein vaccine remains technically very challenging (12, 13).

gp140, the uncleaved ectodomain of $(gp160)_3$, is generally regarded as a likely surrogate for the native Env spike. Most HIV-1 gp140 preparations without stabilizing modifications are biochemically unstable and heterogeneous, with a tendency to aggregate or dissociate (14), leading to claims that all uncleaved gp140 proteins adopt an irrelevant, nonnative structure with three gp120s flexibly associated with a rearranged, postfusion, trimeric gp41 (15). We previously reported that certain HIV-1 gp140 preparations, derived from selected, difficult-to-neutralize strains, such as 92UG037.8 and C97ZA012, are biochemically stable and homogeneous, with many of the antigenic properties of Envs expressed on cell surfaces (5). These trimers are also more immunogenic than sequence-matched monomeric gp120s (4). One of them has been manufactured to good manufacturing practice (GMP) standards by Crucell Holland BV, for a clinical trial now in progress (ClinicalTrials.gov registration no. NCT02315703).

We report in this paper a low-resolution (20 Å) cryoEM structure of the soluble C97ZA012 gp140 trimer, reconstructed by selecting single particles from images of *in vitro*-generated rosettes. We used this strategy to introduce additional constraints for particle orientation to facilitate alignment and to reduce potential problems associated with air-water interface denaturation effects. Using contemporary three-dimensional (3D) classification methods, we show that this soluble trimer has two principal conformations, one slightly more open (i.e., with its gp120 domains spread away from the 3-fold axis) than the other. The closed conformation is fully compact and indistinguishable at this resolution from the conformations of cleaved SOSIP trimers and the JR-FL EnvΔCT trimer. The slightly open conformation more closely resembles a partially open Env trimer (16) than the fully open conformation induced by CD4. Our findings indicate that uncleaved C97ZA012 gp140 trimers have a biochemically stable and compact native structure.

RESULTS

Protein production. In our initial attempt to obtain a 3D cryoEM reconstruction of the C97ZA012 gp140 trimer, refinement failed to converge on a reproducible and interpretable density map, suggesting conformational heterogeneity when the sample was vitrified on holey carbon grids. This result was unexpected, because it was inconsistent with the biochemical and antigenic properties of the trimer in solution (4, 17). A likely cause is the intrinsic conformational flexibility of C97ZA012 gp140, further compounded by the instability induced by the air-water interface created during the standard vitrification procedure (18). Moreover, the "tetrahedral" shape (three gp120

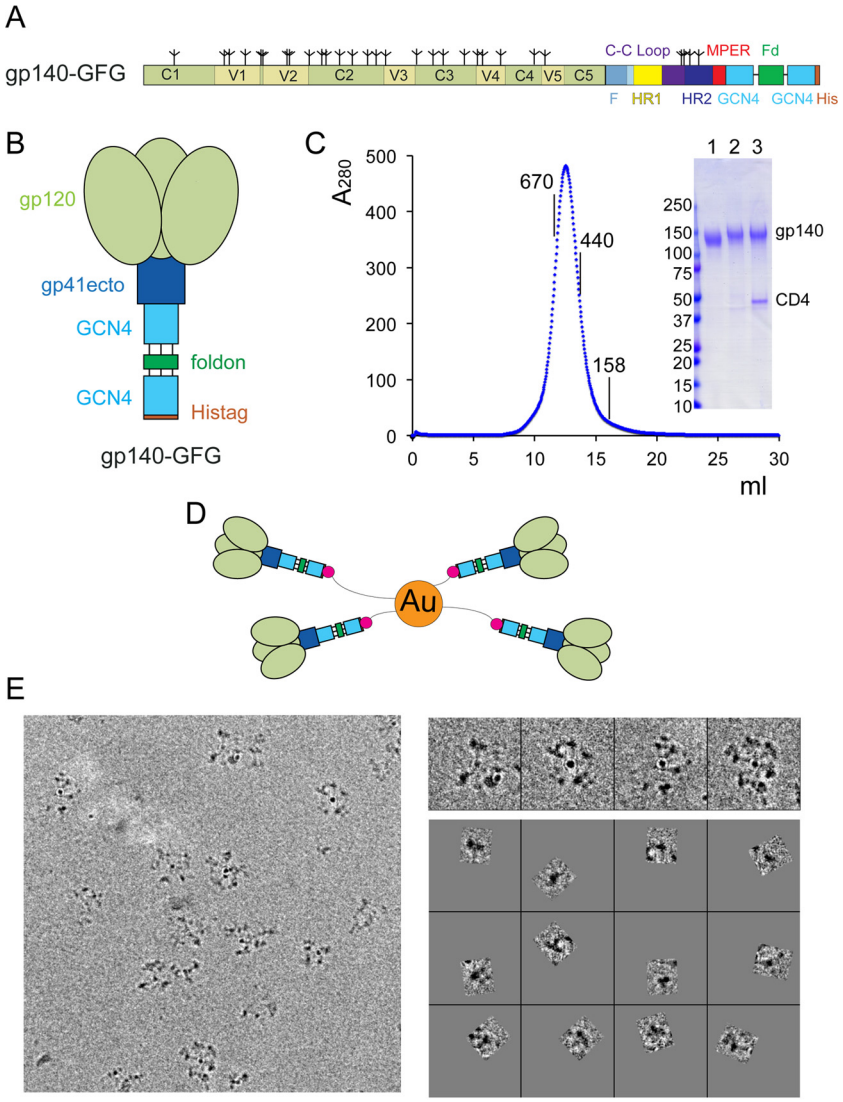

FIG 1 Production and characterization of the C97ZA012 gp140 trimer for cryoEM analysis. (A) Schematic representation of the gp140 construct used in the EM study. The gp140 region includes the following segments: C1 to C5, conserved regions 1 to 5; V1 to V5, variable regions 1 to 5; F, fusion peptide; HR1, heptad repeat 1; C-C loop, immunodominant loop with a conserved disulfide bond; HR2, heptad repeat 2; and MPER, membrane-proximal external region. The C-terminal end of gp140 was fused directly to a GCN4 trimer sequence, followed by a GSGG linker, a foldon sequence, a GSG linker, and another GCN4 sequence ending with a histidine tag. (B) 3D organization of the gp140 trimer. The length of the added stem region can be ~100 Å in an extended conformation. (C) The His_6-tagged gp140-GFG protein was purified from supernatants of 293T cells stably transfected with the expression construct. Purified gp140 was resolved by gel filtration chromatography on a Superose 6 column. The molecular mass standards included thyroglobulin (670 kDa), ferritin (440 kDa), γ-globulin (158 kDa), and ovalbumin (44 kDa). Peak fractions were pooled and analyzed by Coomassie-stained SDS-PAGE (inset; lane 1, gp140 described in reference 4; lane 2, gp140-GFG; and lane 3, purified complex of gp140-GFG and 4-domain CD4). (D) Diagram of rosette formation when the gp140-GFG protein was incubated with Ni-NTA (in red)-conjugated gold beads (in orange). (E) Representative area of a micrograph of gp140 rosettes under cryo conditions. Selected rosettes from this area are shown in the upper right panels, and the extracted gp140 particles from these rosettes are shown in the lower right panels.

monomers and one gp41 trimer) of the gp140 trimer brought additional challenges for classification, alignment, and refinement, even for images recorded on a K2 direct detector camera (Gatan) with excellent contrast and processed with a maximum likelihood algorithm (19).

To overcome these challenges, we added a long stem to the C-terminal end of the gp140 construct (Fig. 1) to make it suitable for a strategy involving formation of rosettes

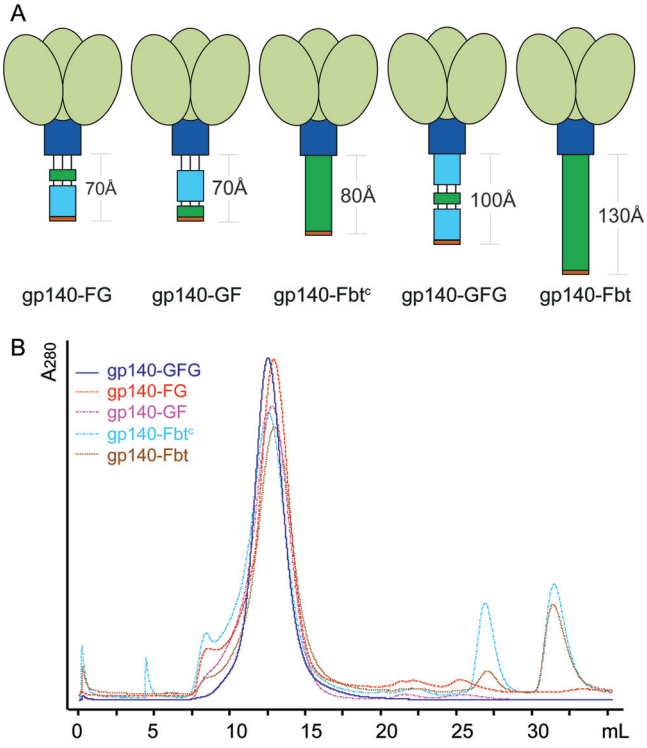

FIG 2 Various gp140-stem-His_6 constructs. (A) Schematic diagrams demonstrating the organization of various gp140 constructs. gp140-FG, the C-terminal end of C97 gp140 is connected to a T4 fibritin foldon through an IEGRGSGG linker, followed by a GSG linker, GCN4, another GSG linker, and a His_6 tag; gp140-GF, the C-terminal end of 92UG gp140 is connected to GCN4 through a GSG linker, followed by an IEGRGSG linker, a foldon, a glycine, and a His_6 tag; gp140-Fbt^c, the C-terminal end of 92UG gp140 is fused in phase to the C-terminal portion of a bacteriophage T4 fibritin unit (34), followed by a glycine and a His_6 tag; gp140-Fbt, the C-terminal end of 92UG gp140 is fused in phase to the full-length fibritin unit, followed by a glycine and a His_6 tag. The coloring scheme follows that in Fig. 1. (B) The constructs were expressed in 293T cells transiently transfected with the expression constructs. The supernatants were harvested at 5 days posttransfection and affinity purified by elution from Ni-NTA resin. The purified gp140 constructs were then subjected to gel filtration chromatography on a Superose 6 column (GE). All of these proteins eluted as a single sharp and symmetrical peak at about the correct elution volume. The variation in the peak is due to the different hydrodynamic radii of these proteins and to different loading volumes.

with nanogold beads, which we also used to study influenza virus hemagglutinin (HA) (Y. Liu, J. Pan, S. Jenni, D. D. Raymond, T. Caradonna, K. T. Do, A. G. Schmidt, S. C. Harrison, and N. Grigorieff, submitted for publication). Within the rosette arrangement, the gp140 molecules that we could recognize and subject to single-particle averaging tended to lie flat in the ice layer, with the trimer axis aligned parallel to the air-water interface. This orientation limits the potential interactions between the molecules and the air-water interface. It also allows us to impose constraints in the image processing step as described below. We tested several versions of the added stem with combinations of trimeric GCN4 (20) and a foldon trimerization domain (21) connected by glycine- and serine-based flexible linkers (Fig. 2). All these proteins were stable and homogeneous, with nearly identical biochemical properties (Fig. 2), suggesting that our manipulations at the C terminus of gp140 had not detectably altered the trimer structure. For further analysis, we chose a construct that we designated gp140-GFG (gp140-GCN4-GSGG-foldon-GSG-GCN4-His_6; the sequences of two Gly-Ser linkers are shown in single-letter code). The C97ZA012 gp140-GFG construct that we prepared had a stem designed to be ~100 Å long in an extended conformation, but with considerable flexibility in the linkers. The purified gp140-GFG protein was monodisperse as judged by gel filtration chromatography; it eluted with an apparent molecular mass of ~600 kDa from a Superose 6 column (Fig. 1C) and gave a single band with a molecular mass of ~150 kDa on SDS-PAGE (inset of Fig. 1C), as expected. Moreover, it showed the same antigenic properties as the gp140 trimer that we reported previously (Fig. 3 and Table 1) (4).

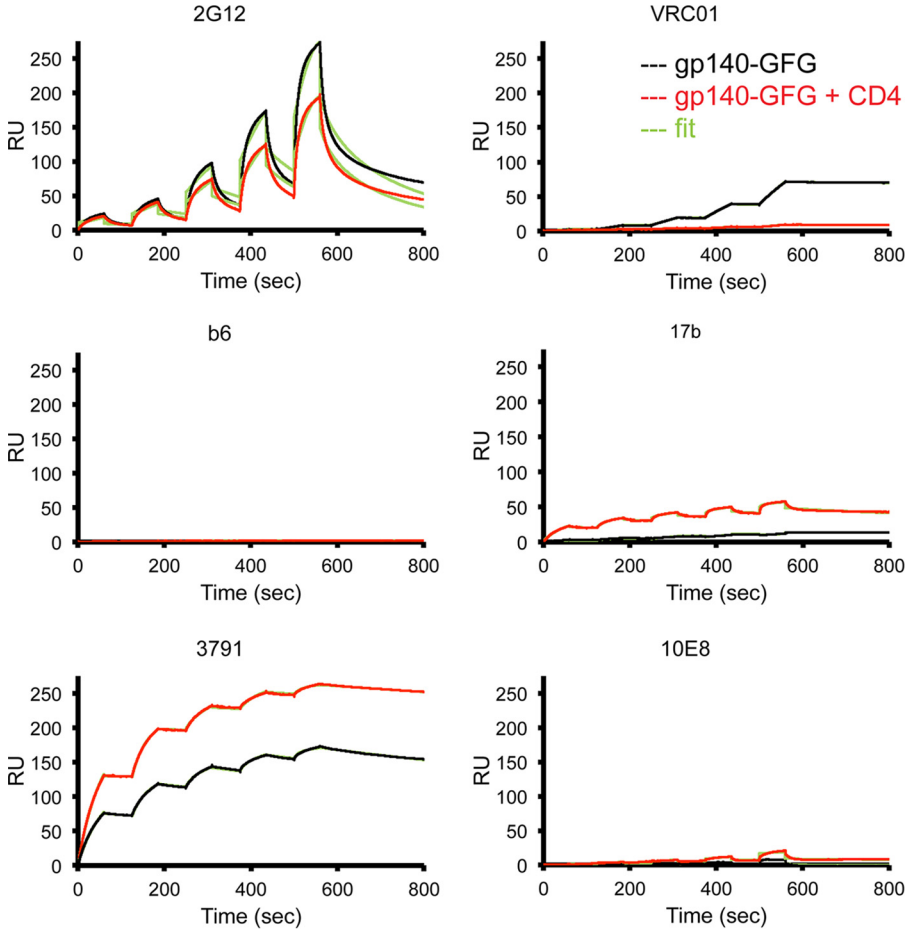

FIG 3 Antigenic properties of the C97ZA012 gp140-GFG trimer. The gp140-GFG protein was analyzed by a surface plasmon resonance assay for binding to the monoclonal antibodies 2G12 (specific for glycan), VRC01 (specific for the CD4 binding site; broadly neutralizing), b6 (specific for the CD4 binding site; nonneutralizing), 17b (CD4i), 3971 (specific for V3; nonneutralizing), and 10E8 (specific for the MPER; broadly neutralizing). gp140-GFG or the purified complex of gp140-GFG and CD4 was captured on the surface of a sensor chip coated with an anti-His_6 antibody to avoid potential artifacts introduced by protein immobilization. Each antibody at various concentrations was passed over the gp140 surface individually without regeneration for single-cycle kinetic analysis. The recorded sensorgrams for gp140-GFG are shown in black, those for the gp140-GFG–CD4 complex in red, and the computational fit in green. Binding constants are summarized in Table 1. RU, response units.

Structure determination. We prepared clusters ("rosettes") of gp140-GFG by mixing the His-tagged protein with Ni-nitrilotriacetic acid (Ni-NTA) nanogold beads (~5 nm in diameter) and found that a ratio of 4 to 6 trimers per gold bead was optimal for imaging (Fig. 1E). We collected 3,138 movies of these rosettes and extracted 121,788

TABLE 1 Equilibrium dissociation constants derived from surface plasmon resonance analysis

	Antibody		gp140 K_D (nM)[a]	
Epitope	Name	Neutralization	Without CD4	With CD4
CD4 binding site	VRC01	+	21.9 ± 0.2	NB
CD4i	b6	−	NB	NB
	17b	−	NB	2.3 ± 0.3
V3	3791	−	1.5 ± 0.2	8.4 ± 0.1
Glycan	2G12	+	239 ± 3	269 ± 4
MPER	10E8	+	NB	NB

[a]NB, no binding.

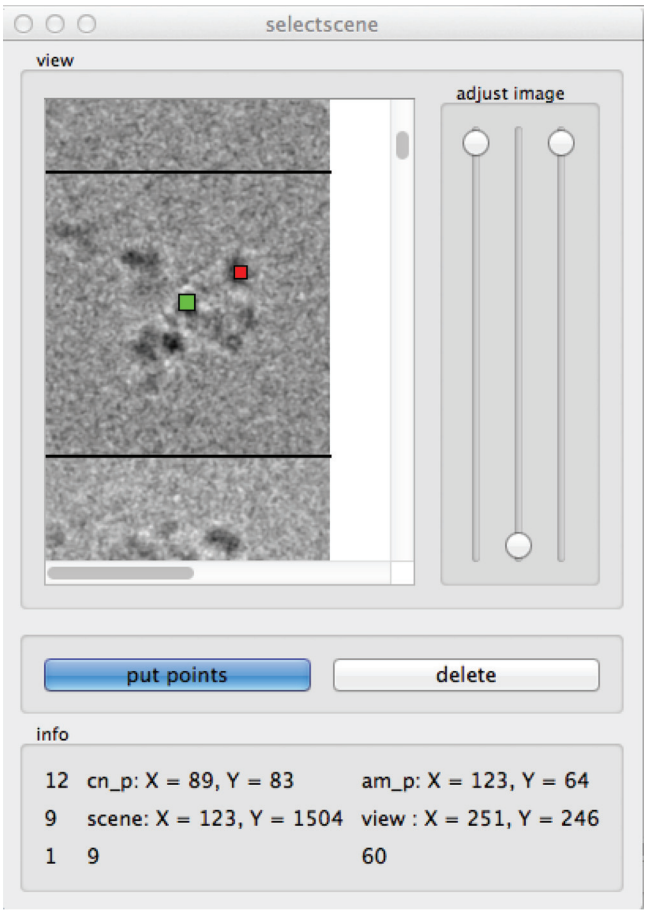

FIG 4 Graphical user interface of a program for picking a trimer from a rosette image. The rosette images are loaded in the upper left panel. The center of the rosette is selected and marked with a green dot; selected positions of individual trimers in the rosette are marked with red dots. The coordinate information from the image is displayed in the lower panel.

individual gp140 trimer particles from the images (Fig. 4). The flexibility of the C-terminal extension ensured that even molecules bound far from the "equator" of the bead would bend into the aqueous layer as it thinned rather than projecting into the air-water interface. Selected particles were aligned to projections of a simplified model made of cylinders to preserve trimer polarity (Fig. 5) (see Materials and Methods). We used cross-correlation and angle-matching criteria to select 95,631 well-aligned particles, which we then subjected to multireference 3D classification with six references and iterative refinement by using Frealign (22). Among the resulting six classes, three had rod-like density at one end, presumably corresponding to the GCN4-foldon tail, and reasonable statistics. These three classes represented ~64% of all selected particles. The remaining three classes (~36% of particles) gave noisy density maps that lacked the tail density. We therefore eliminated the particles in these three classes from the stack. The particles from each of the three "good" classes were subjected to further classification into three subclasses. For classes 1 and 2, the resulting three structures (or subclasses) were all very similar to the structure before subclassification (Fig. 6), suggesting that these reconstructions each represented a relatively homogeneous population. In contrast, the three structures derived from the third class had quite different density distributions (Fig. 6), suggesting that the class 3 structure was an average of multiple conformational states. We applied 3-fold symmetry during image processing until this point. The resolution of the reconstructions for classes 1 and 2 was about 20 Å, based on a Fourier shell correlation (FSC) criterion of 0.143 (Fig. 7).

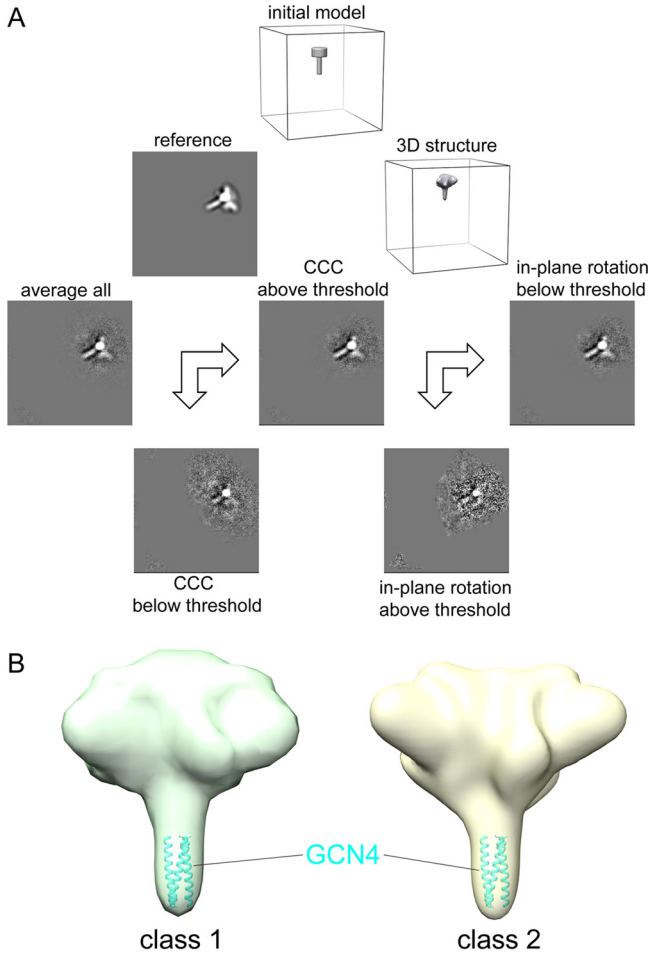

FIG 5 Structure determination for the C97ZA012 gp140 trimer by cryoEM. (A) Strategy to eliminate particles from the reconstruction in the early alignment cycles by using the SPIDER procedure. (First row) An initial model made of a long and thin cylinder connected to a short and thick cylinder was used to perform the first round of 3D reconstruction. The resulting structure served as the reference for the next cycle. (Second row) A representative projection is shown next to one of the refined 3D maps. (Third row) Particles were aligned to their best-matching projections. The "average all" panel represents the average for all particles that matched the projection shown in the second row. The "CCC above threshold" panel shows the average for the particles that have a cross-correlation coefficient (CCC) above the threshold value. For the particles that did not meet the threshold, the average image is shown as "CCC below threshold," in the fourth row. The particles that met the CCC threshold were sorted by the angle-matching criterion. The average for particles below the angular threshold is represented in the "in-plane rotation below threshold" panel, which was included in the final 3D structure reconstruction. (Fourth row) The average for the rejected particles is shown for each criterion. (B) Side views of the two reconstructions corresponding to subclasses 1 and 2, with trimeric GCN4 (in cyan) fit into the tail-like density.

Subclassification of the three classes indicated that there was conformational microheterogeneity (i.e., small variations in density distribution among subclasses) (Fig. 6) within each of them, limiting the resolution of the reconstructions. To identify potential sources of this microheterogeneity, we explored the hypothesis that the gp140 trimer may not have perfect 3-fold symmetry due to conformational dynamics. For example, each of the three protomers might deviate from the 3-fold axis independently, breaking the symmetry. To test this possibility, we performed classification of classes 1 and 2 by using a mask to focus on a single protomer and generated reconstructions without imposing 3-fold symmetry (Fig. 8). A slightly different density distribution within the mask was evident, suggesting that microscale heterogeneity was indeed present within each class, probably due to conformational flexibility of each protomer, residual alignment errors, and heterogeneity in glycosylation.

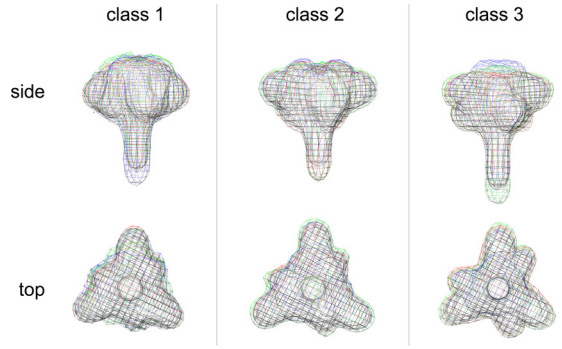

FIG 6 Further subclassification of three retained 3D classes. The particles that belonged to classes 1, 2, and 3 were extracted into separate stacks. Each class was subjected to further classification into three subclasses. The three structures (red, blue, and green) obtained after 60 refinement cycles were overlaid to evaluate the structural homogeneity of each class.

The main differences between the structures for classes 1 and 2 are in the distribution of the density around the trimer apex and at its periphery. We regard these differences as significant because they come from comparison of 3D reconstructions of the same protein on the same grid. The two structures appear to represent "closed" (class 1) and "slightly open" (class 2) conformations. Class 3 may represent intermediates in the transition between these two states, as it appears to have features of both classes 1 and 2 (Fig. 6). We also note that different initial models gave essentially the same 3D reconstructions (Fig. 9), suggesting that the procedure is reproducible and not biased by the input model.

Conformation of the C97ZA012 gp140 trimer. To determine which conformational state(s) the C97ZA012 gp140 trimer might adopt, we superimposed the density maps of classes 1 and 2 onto published EM maps of HIV-1 Env, including the clade B JR-FL EnvΔCT construct (closed conformation), the BG505 SOSIP-sCD4-17b-8ANC195$_{G52K5}$ construct (partially open), and the CD4^{-} and 17b-bound trimer (open) (11, 16, 23). We note that the Envs used in the various studies are different in sequence, construct design, and glycosylation. Some degree of difference in their EM maps is thus expected. Nonetheless, the gp120 part of class 1 well matched the density of the closed trimer (Fig. 10A). The gp120 parts of the partially open and open trimers could also fit into the class 1 map, but these structures do not account for the strong density feature

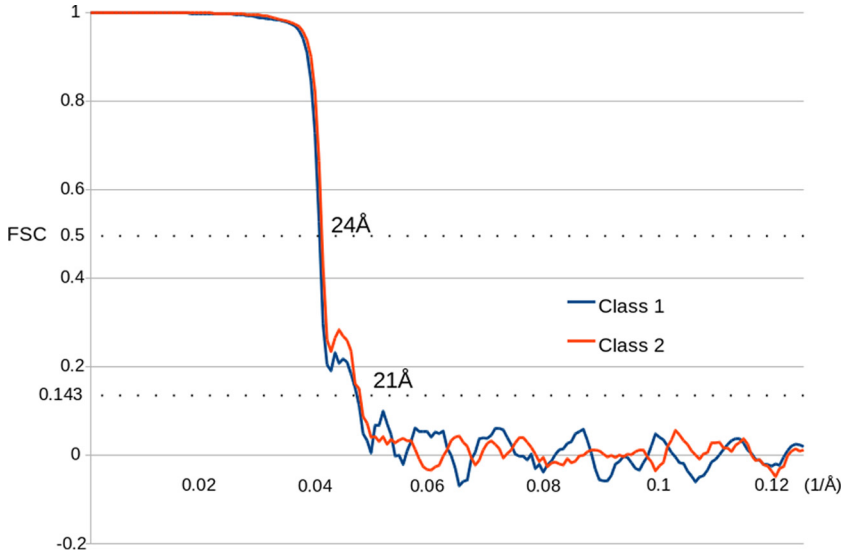

FIG 7 FSC curves for the reconstructions of classes 1 and 2.

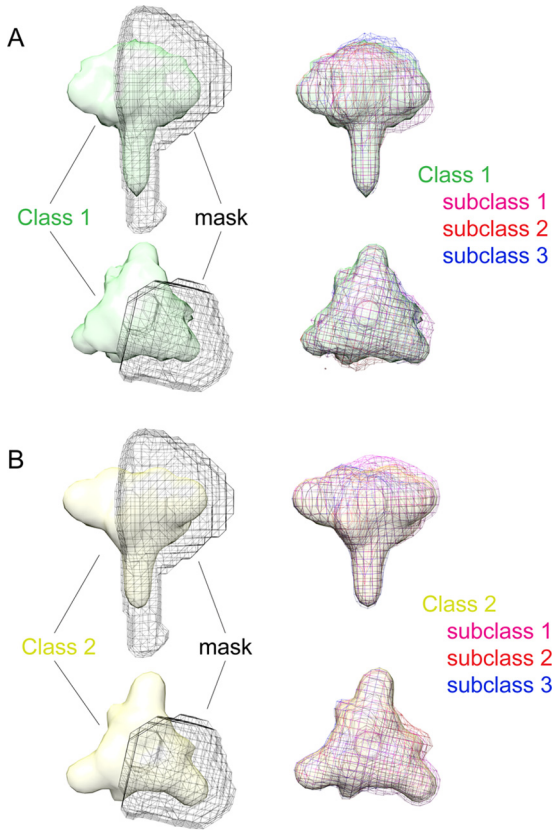

FIG 8 Focused classification of gp140 structures. Focused classification using a mask (shown in black mesh) covering one protomer of the trimer was performed for both class 1 (A; green) and class 2 (B; yellow). Three subclasses for each class are shown, in magenta, red, and blue. Differences within the mask are evident among subclasses, while density outside the mask remains the same.

in the center of the trimer. Class 2 probably represents a partially open conformation. Side views of the class 2 map show that it lacks the density at the trimer apex, contributed in the closed structure by the V1V2 and V3 loops, but that it has more density in that region than would characterize a fully open trimer (Fig. 10B and 11),

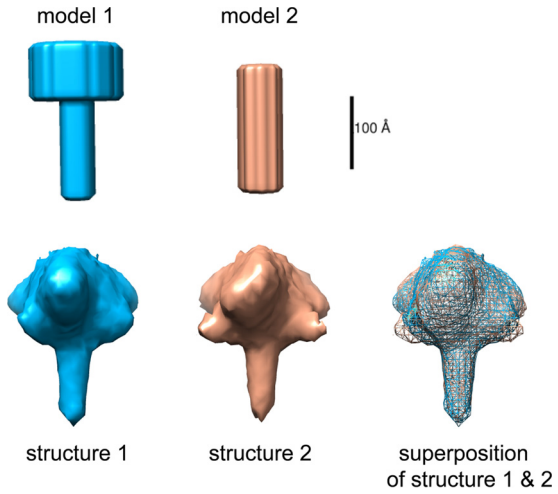

FIG 9 Impact of the initial model on 3D reconstruction. Two different initial models were tested for 3D reconstruction. Model 1, a long and thin cylinder connected to a short and thick cylinder led to structure 1; model 2, a single cylinder gave rise to structure 2. The two structures are almost identical, as shown by the overlay.

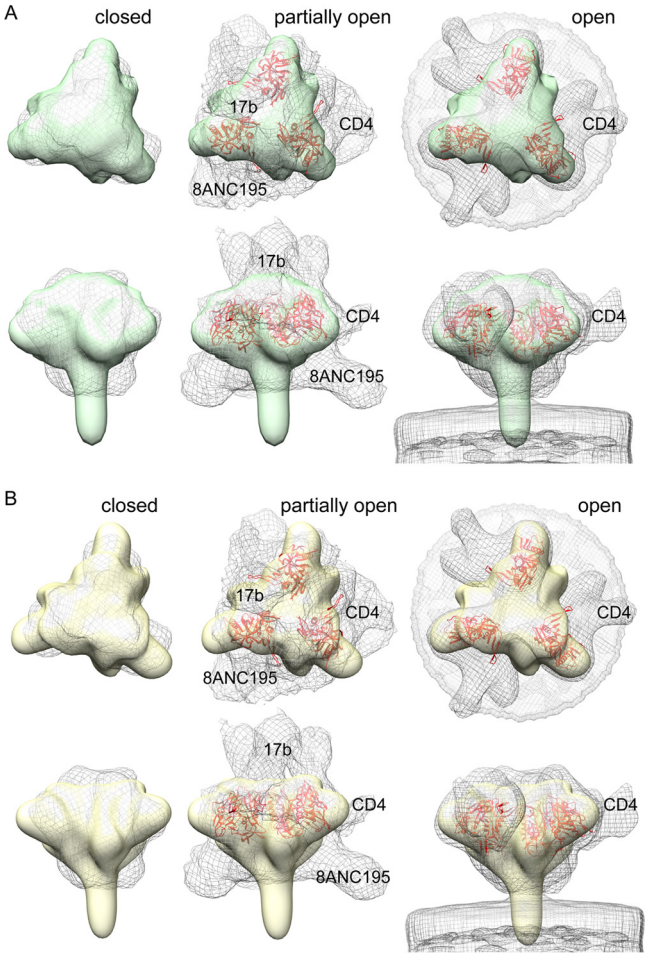

FIG 10 Conformation of the uncleaved C97ZA012 gp140 trimer. (A) CryoEM map of class 1 C97ZA012 gp140 trimer (light green) superimposed on the map of the cleaved, wild-type JR-FL EnvΔCT trimer in complex with PGT151 (structure EMD-3308 filtered to 21-Å resolution [shown as mesh[; the density for PGT151 was removed for clarity) (left), the map of the BG505 SOSIP.664 trimer in complex with soluble 2D CD4, 17b Fab, and 8ANC195 variant G32K5 Fab (structure EMD-3096 [mesh]) (middle), and the map of the BaL Env trimer on the virion surface in complex with soluble 2D CD4 (structure EMD-5455 [mesh]) (right). For the maps representing partially open and open conformations induced by ligand binding, the locations of three gp120s are shown by models, in red, of the gp120 core (PDB entries 5A8H and 3DNO, respectively). Locations of CD4 and antibodies are also indicated. (B) As in panel A, except that the map in light yellow is the cryoEM map of a class 2 C97ZA012 gp140 trimer.

consistent with its lack of binding to the CD4-induced (CD4i) antibody 17b in the absence of CD4 (Fig. 3) (4). These results suggest that the C97ZA012 gp140 sample conformations range between closed and partially open.

A model of the clade B JR-FL EnvΔCT trimer, an antibody-stabilized Env with a structure nearly identical to that of the BG505 SOSIP.664 trimer (11), fit the class 1 map exceptionally well (Fig. 12), especially considering differences in glycosylation between the JR-FL and C97ZA012 Envs. In particular, the density for the variable loops at the trimer apex, which present the quaternary epitopes for trimer-specific antibodies, such as PG9, PG16, and PGT145, was evident in the class 1 map, consistent with our published antigenicity studies showing that C97ZA012 gp140 binds PG9 and PG16 (4). As suggested above, class 2 appears to represent a structure in which the three gp120 molecules have moved apart toward an open conformation, disrupting the trimer-specific epitopes for PG9, PG16, or PGT145 created by 3-fold packing of the V1V2 and V3 loops. If the C97ZA012 gp140 trimer spends a substantial fraction of its time in the class 2 state, then the Env preparation will have, as observed, weaker-than-expected binding to these antibodies in solution (4).

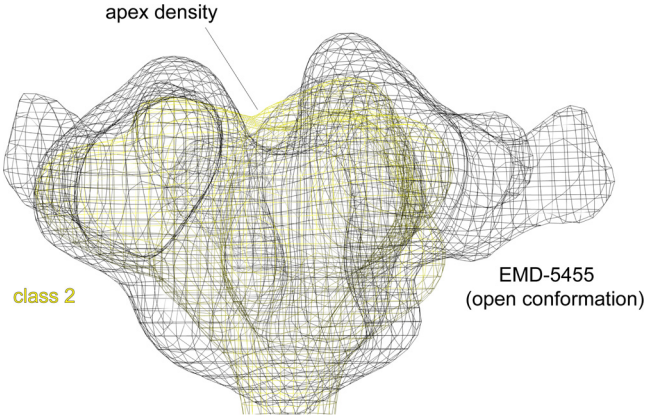

FIG 11 Main difference between the maps for class 2 and the open conformation. The cryoEM map of a class 2 C97ZA012 gp140 trimer (yellow) is superimposed on the map of the BaL Env trimer on the virion surface in complex with soluble 2D CD4 (structure EMD-5455 [black mesh]). The main difference between the two maps, next to the trimer apex, is indicated.

The fit of the gp41 part of the clade B JR-FL EnvΔCT model to the class 1 map is imperfect, particularly near the C-terminal end, suggesting that there may be some differences in the gp41 structure between the two trimers. This difference is not surprising, as the JR-FL EnvΔCT trimer is solubilized by detergent and its MPER and transmembrane domain (TMD) are disordered, while the MPER of C97ZA012 gp140 is directly constrained by a GCN4 or foldon trimerization domain. We note that our gp140 trimer does not bind MPER-directed bnAbs, fully consistent with the antigenic properties of the corresponding full-length, functional Env spikes on the cell surfaces (5).

DISCUSSION

The structural data presented here confirm that the uncleaved, soluble C97ZA012 gp140 trimer, currently being tested in a clinical trial, has a compact structure very similar to those of disulfide-stabilized, cleaved SOSIP trimers as well as that of a

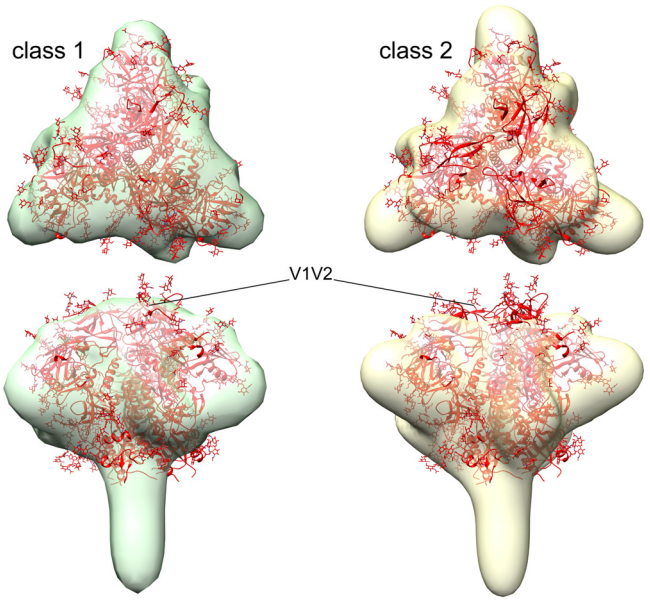

FIG 12 Fit of the JR-FL EnvΔCT trimer into cryoEM maps of C97ZA012 gp140. The model of the JR-FL EnvΔCT trimer (PDB entry 5FUU), shown as a ribbon diagram in red, was fit by manual adjustment in UCSF Chimera into the density maps of the 3D reconstructions of class 1 (light green) and class 2 (light yellow). The location of V1V2, which presents the trimer-specific, broadly neutralizing epitopes, is indicated.

detergent-solubilized, cleaved EnvΔCT trimer (8–11). The inherent conformational flexibility of the trimer without constraints from bound Fabs or engineered disulfides and the alignment errors resulting from this microheterogeneity have limited the resolution of the cryoEM reconstructions to about 20 Å. However, most C97ZA012 gp140 trimers appear to sample a narrow range of conformations, as 3D classification shows two principal states: one closed and the other partly open. This range may simply reflect the intrinsic dynamics expected for any functional protein. Both conformations are well within the range of those observed for functional Envs (Fig. 10). Whatever the detailed sources of microheterogeneity, our observations are inconsistent with the assertion that all the uncleaved gp140 trimers have nonnative structures (15). As has been well documented for influenza virus HA, we can expect to find local structural differences, near the cleavage site, between the uncleaved precursor and the fully cleaved and fusion-activated trimer (24, 25), but if the HA precedent is relevant, we do not expect global changes following cleavage. The cryoEM structures described here are fully consistent with these expectations.

Our recent studies of two conformationally homogeneous, cell surface-expressed gp160s also suggest that the prefusion conformation of HIV-1 Env is independent of the cleavage between gp120 and gp41. In those studies, a substantial amount of uncleaved gp160 trimer was also present, but the cells showed no binding to any of the nonneutralizing antibodies that recognize various epitopes over the entire Env ectodomain (5). That is, the antigenic properties of the uncleaved gp160 protein and the cleaved trimer are virtually identical if both are sufficiently stable. The CT (cytoplasmic tail)-deleted Env, while fully functional in inducing membrane fusion, exposes certain nonneutralizing epitopes which are accessible on intact Env only upon CD4 binding, and it binds less tightly to the trimer-specific antibodies, which recognize the closed state (5). Thus, this truncated, functional Env may adopt the partially open conformation observed in the present study.

Why are most soluble gp140 preparations not stable and homogeneous? Our recently reported nuclear magnetic resonance (NMR) structure of the TMD of HIV-1 Env shows that it forms a tightly packed trimer in a lipid bilayer environment (26). The TMD trimer resists denaturation by SDS *in vitro* and cannot be disrupted by simple mutations at the trimer interface. The extraordinary stability of the TMD results from the contributions of two structural elements: a "standard," N-terminal coiled-coil segment with a hydrophobic core around the 3-fold axis and a C-terminal segment with a hydrophilic core (26). Either one is sufficient to maintain the TMD trimer structure. Destabilizing the TMD trimers leads to major changes in the antigenic structure of the ectodomain of Env, suggesting that the TMD contributes to the stability of the entire Env spike. These data can explain why most recombinant soluble Env preparations with the TMD deleted are unstable and conformationally heterogeneous, except for those of a few selected strains (14, 17), such as the C97ZA012 gp140 studied here.

Stabilized SOSIP trimers have failed to induce any heterologous, tier 2 responses in animal models, despite expression of most known bnAb epitopes (27). We do not know whether the modifications introduced into these trimers have prevented them from recapitulating all the properties, including conformational dynamics, of the native, functional Env spikes on the virion surface or whether there are unmet immunological requirements, such as reactivity with unmutated ancestors of bnAbs. We also do not know whether a rigidified protein, such as a SOSIP trimer, is a better immunogen than the C97ZA012 gp140 trimer described here. For example, the intrinsic dynamics expected for a functional protein may influence its antigenic or immunogenic properties. Indeed, the cryoEM structure of JR-FL EnvΔCT in complex with the Fab of the bnAb PGT151 showed that the antibody recognizes the fusion peptide in a conformation that cannot be stable in the absence of the antibody (11), suggesting that protein flexibility may be necessary to induce an antibody like PGT151. Thus, there are potential benefits of using an immunogen, such as C97ZA012 gp140, without the engineered disulfides and with native conformational flexibility. A potential immunization strategy is to prime with a construct such as C97ZA012, which samples a range of conformations that a

functional Env can adopt, and then boost with a more rigid, disulfide-cross-linked SOSIP trimer to focus the immune response on a conformational state that selectively presents trimer-specific bnAb epitopes.

In summary, uncleaved C97ZA012 gp140 trimers sample two principal conformational states, both of which are relatively compact. They may represent two functional states of the unliganded Env or simply two components of the dynamic ensemble that characterizes it. One of them is indistinguishable, at the resolution we examined, from the closed conformation seen for cleaved Env trimers. These data support the use of stable, soluble HIV-1 Env trimers in HIV vaccine research, including human efficacy trials.

MATERIALS AND METHODS

Production of HIV-1 envelope proteins. A 293T cell line stably transfected with the gp140-GFG gene was generated using previously described methods (4, 5). The stable cell line was grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) to confluence and then changed to Freestyle 293 expression medium (Invitrogen). The cell supernatants were harvested 96 to 108 h after the medium change. The His_6-tagged envelope proteins were purified by use of Ni-NTA (Qiagen) followed by gel filtration chromatography as described previously (4, 28). The purified proteins were concentrated, frozen in liquid nitrogen, and stored at −80°C.

Surface plasmon resonance binding assay. All binding experiments were performed in duplicate with a Biacore T200 instrument (GE Healthcare). Interactions of antibodies with envelope constructs were analyzed using single-cycle kinetics, consisting of four cycles of a 1-min association phase and a 1-min dissociation phase, without regeneration between injections, followed by an additional cycle of a 1-min association phase and a 4-min dissociation phase, with a flow rate of 30 μl/min. All experiments were carried out in HBS-EP (10 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM EDTA, and 0.005% P-20). For each antibody, the same amount (in moles) of gp140 trimer or purified gp140-CD4 complex was captured by use of an anti-His antibody (Millipore) immobilized on a CM5 sensor chip. Fab was then passed over the surface at increasing concentrations from 62.5 nM to 1,000 nM. Immobilization of the anti-His antibody on CM5 chips was performed following the standard amine coupling procedure as recommended by the manufacturer. The surface was regenerated after each experiment by two injections (5 min each) of 10 mM glycine-HCl (pH 2.1) at 10 μl/min followed by 5 min of equilibration in HBS-EP. Injections over blank surfaces were subtracted from the data. Single-cycle binding kinetics sensorgrams were fit to a 1:1 binding model by use of BiaEvaluation software (GE Healthcare).

CryoEM grid preparation and imaging. We screened mixtures with different ratios of gp140-GFG protein and Ni-NTA nanogold (Nanoprobes); a ratio of ~4 to 6 trimers/gold bead in each rosette was optimal for obtaining enough rosettes per micrograph. For data collection, the rosette sample was prepared with gp140-GFG at 0.05 mg/ml, nanogold beads at ~60 nM, and 0.07% β-octylglucoside. To prepare cryogrids by use of a Vitrobot plunger, 3.5 μl sample solution was applied to a glow-discharged Quantifoil grid (SPI Supplies), which was then blotted for 4 s, with the offset set to 1, in a 100% humidity chamber. The grid was immediately plunged into liquid ethane, transferred to a small cryobox, and stored in liquid nitrogen before imaging.

Images were recorded with a Polara electron microscope (FEI) operated at 300 kV by use of a K2 Summit camera (Gatan) in superresolution counting mode, with a pixel size of 0.99 Å. The dose rate of the camera was ~8 electrons/physical pixel/s. We collected 38 frames for each movie, with a total dose of ~40 electrons/Å^2 on the sample, using semiautomated data collection with SerialEM (29). We recorded 3,138 movies with a defocus range of 1 to 4 μm.

Image processing and 3D reconstruction. All movies were aligned and averaged with the program Unblur (30). The drift-corrected images were binned by a factor of 4 and low-pass filtered to 15 Å for particle picking by using proc2d of EMAN2 (31). We manually picked 57,913 rosettes from the images by using the e2boxer program (31), with a box size of 180 pixels on the 4×-binned micrographs. The rosette center and the angle α between a fixed axis and the axis of the gp140-GFG trimers in each individual rosette were determined with an in-house program (Fig. 2). Based on these two parameters, 121,788 gp140 trimers were extracted from the rosettes on unbinned images by use of SPIDER (32). The image stack was then binned 4 times with the resample tool from the Frealign package (22) and used for refinement.

Procedures in SPIDER (32) were used for initial alignment of the particles. We created a simple, elongated model (Fig. 5) to represent a gp140-GFG trimer in a rosette and generated projections of this model as references for the image stack. During the alignment, the maximum search range was set to 15 pixels (~60 Å) to preserve the polarity of the gp140 trimers, which projected away from the center of the gold bead. After initial alignment, we applied the following angle-matching criterion. Because particles were picked from rosettes, their orientations were constrained in two ways. First, the C-terminal His_6 tag attaching the trimer to the bead determined the polarity of each molecule within the rosette. Second, each trimer projected from the bead at an angle (with respect to image coordinates) that could be estimated directly from the display of the picked particle. We therefore required that the orientation determined by alignment with the reference agree, to within a threshold level of uncertainty, with the orientation estimated directly from the image. Particles that did not meet this angle-matching criterion were excluded, leaving 95,631 particles to be used for image reconstruction. For these particles, in-plane rotations and shifts were applied to generate a new particle stack, and the Euler angles from the best-matching reference were used for reconstruction in Frealign. The resulting structure was then used

as a reference for the next iteration. We carried out nine iterative cycles in all. The cross-correlation (percentage) and angle-matching (degrees) cutoffs for the iterations were as follows: 1, 1, 10, 10, 10, 10, 10, 10, and 20 for cross-correlation and 20, 20, 20, 20, 20, 20, 20, 20, and 18 for angle matching. In this initial alignment procedure, the 3D model from each cycle was low-pass filtered to 15 Å before generation of projection images.

Following the projection matching just described, the 95,631 particles were subjected to multireference 3D classification and refinement with Frealign. First, the model was gradually shifted to the center of the volume in steps of 5 pixels by use of proc3d (31). After each shift, several cycles of refinement were performed with the angle ϕ fixed. The centering made it easier to mask the reconstructions to improve refinement. The stack was then refined against six references for classification, with all the angle and shift parameters free to vary. After cycle 60 of multireference refinement, the particle distribution among the six classes was as follows: ~18% for class 1, ~24% for class 2, ~22% for class 3, ~10% for class 4, ~13% for class 5, and ~12% for class 6. The first three classes showed elongated density at one end, corresponding to the GCN4 tail. The last three classes, each with fewer particles, did not have a clear elongation, and we excluded the corresponding particles from further processing. The particles that belonged to the three retained classes were extracted, and each class was subclassified into three classes for 60 cycles. For classes 1 and 2, the three resulting structures from subclassification looked very similar, while the structures from class 3 had rather different density distributions. We applied 3-fold symmetry during all image processing steps except focused classification (see below). During the entire Frealign multireference refinement, the first few cycles were performed with a resolution cutoff of 40 Å, which was gradually changed to the final resolution cutoff of 30 Å. The resolution was calculated from half maps by use of SPIDER, and the final volumes were filtered at an estimated resolution of 21 Å (Fig. 7) and sharpened with a B-factor of −300 Å^2. The final images were prepared with UCSF Chimera (33).

For the focused classification procedure, we generated a mask covering a single protomer by performing the following steps. (i) We created a 120-degree wedge by overlaying the edges of two cube volumes made with SPIDER commands (32). (ii) We then placed the wedge along the 3-fold symmetry axis of the class 1 or class 2 density map to mask out a region corresponding to a single protomer including both gp120 and gp41. (iii) We used the proc3d command (EMAN2) (31) to generate a binary mask from the density map in step 2. With this mask, focused classification into three subclasses was then performed for each class by using Frealign (22). Particle alignment parameters were kept fixed, and no 3-fold symmetry was applied during this subclassification.

EM density maps and Protein Data Bank (PDB) models were fitted into reference density maps by rigid body fitting with the UCSF Chimera program (33); the fits were adjusted manually to ensure good overlaps for common features.

Accession number(s). Data were deposited in the EMDatabank (http://www.emdatabank.org/index.html) under accession codes EMD-8629 and EMD-8631.

ACKNOWLEDGMENTS

We thank Chen Xu for technical assistance.

This work was supported by NIH grants AI084794 (to B.C. and Dan H. Barouch), P01 AI089618 (to S.C.H.), R01 AI106488 (to B.C.), and R56AI112489 (to B.C.), by Collaboration for AIDS Vaccine Discovery (CAVD) grant OPP1040741 (to Dan H. Barouch from the Bill and Melinda Gates Foundation), and by Center for HIV/AIDS Vaccine Immunology-Immunogen Discovery (CHAVI-ID) grant UMI-AI100645 (to Barton F. Haynes). S.C.H. and N.G. are investigators at the Howard Hughes Medical Institute.

REFERENCES

1. Harrison SC. 2005. Mechanism of membrane fusion by viral envelope proteins. Adv Virus Res 64:231–259. https://doi.org/10.1016/S0065-3527(05)64007-9.
2. Chan DC, Fass D, Berger JM, Kim PS. 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell 89:263–273. https://doi.org/10.1016/S0092-8674(00)80205-6.
3. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. 1997. Atomic structure of the ectodomain from HIV-1 gp41. Nature 387:426–430. https://doi.org/10.1038/387426a0.
4. Kovacs JM, Nkolola JP, Peng H, Cheung A, Perry J, Miller CA, Seaman MS, Barouch DH, Chen B. 2012. HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. Proc Natl Acad Sci U S A 109:12111–12116. https://doi.org/10.1073/pnas.1204533109.
5. Chen J, Kovacs JM, Peng H, Rits-Volloch S, Lu J, Park D, Zablowsky E, Seaman MS, Chen B. 2015. HIV-1 envelope. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. Science 349:191–195. https://doi.org/10.1126/science.aaa9804.
6. Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, Kim HJ, Blattner C, de la Pena AT, Korzun J, Golabek M, de Los Reyes K, Ketas TJ, van Gils MJ, King CR, Wilson IA, Ward AB, Klasse PJ, Moore JP. 2013. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS Pathog 9:e1003618. https://doi.org/10.1371/journal.ppat.1003618.
7. Pugach P, Ozorowski G, Cupo A, Ringe R, Yasmeen A, de Val N, Derking R, Kim HJ, Korzun J, Golabek M, de Los Reyes K, Ketas TJ, Julien JP, Burton DR, Wilson IA, Sanders RW, Klasse PJ, Ward AB, Moore JP. 2015. A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. J Virol 89:3380–3395. https://doi.org/10.1128/JVI.03473-14.
8. Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, Klasse PJ, Burton DR, Sanders RW, Moore JP, Ward AB, Wilson IA. 2013. Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science 342:1477–1483. https://doi.org/10.1126/science.1245625.
9. Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, Burton DR, Sanders RW, Moore JP, Carragher B, Wilson IA, Ward AB. 2013. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science 342:1484–1490. https://doi.org/10.1126/science.1245627.
10. Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, Huang J, Acharya P, Chuang GY, Ofek G, Stewart-Jones GB, Stuckey J, Bailer RT, Joyce MG, Louder MK, Tumba N, Yang Y, Zhang B, Cohen MS, Haynes

BF, Mascola JR, Morris L, Munro JB, Blanchard SC, Mothes W, Connors M, Kwong PD. 2014. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature 514:455–461. https://doi.org/10.1038/nature13808.
11. Lee JH, Ozorowski G, Ward AB. 2016. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. Science 351:1043–1048. https://doi.org/10.1126/science.aad2450.
12. Chung NP, Matthews K, Kim HJ, Ketas TJ, Golabek M, de Los Reyes K, Korzun J, Yasmeen A, Sanders RW, Klasse PJ, Wilson IA, Ward AB, Marozsan AJ, Moore JP, Cupo A. 2014. Stable 293 T and CHO cell lines expressing cleaved, stable HIV-1 envelope glycoprotein trimers for structural and vaccine studies. Retrovirology 11:33. https://doi.org/10.1186/1742-4690-11-33.
13. Kong L, He L, de Val N, Vora N, Morris CD, Azadnia P, Sok D, Zhou B, Burton DR, Ward AB, Wilson IA, Zhu J. 2016. Uncleaved prefusion-optimized gp140 trimers derived from analysis of HIV-1 envelope metastability. Nat Commun 7:12040. https://doi.org/10.1038/ncomms12040.
14. Jeffs SA, Goriup S, Kebble B, Crane D, Bolgiano B, Sattentau Q, Jones S, Holmes H. 2004. Expression and characterisation of recombinant oligomeric envelope glycoproteins derived from primary isolates of HIV-1. Vaccine 22:1032–1046. https://doi.org/10.1016/j.vaccine.2003.08.042.
15. Ringe RP, Sanders RW, Yasmeen A, Kim HJ, Lee JH, Cupo A, Korzun J, Derking R, van Montfort T, Julien JP, Wilson IA, Klasse PJ, Ward AB, Moore JP. 2013. Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. Proc Natl Acad Sci U S A 110:18256–18261. https://doi.org/10.1073/pnas.1314351110.
16. Scharf L, Wang H, Gao H, Chen S, McDowall AW, Bjorkman PJ. 2015. Broadly neutralizing antibody 8ANC195 recognizes closed and open states of HIV-1 Env. Cell 162:1379–1390. https://doi.org/10.1016/j.cell.2015.08.035.
17. Kovacs JM, Noeldeke E, Ha HJ, Peng H, Rits-Volloch S, Harrison SC, Chen B. 2014. Stable, uncleaved HIV-1 envelope glycoprotein gp140 forms a tightly folded trimer with a native-like structure. Proc Natl Acad Sci U S A 111:18542–18547. https://doi.org/10.1073/pnas.1422269112.
18. Cheng Y, Grigorieff N, Penczek PA, Walz T. 2015. A primer to single-particle cryo-electron microscopy. Cell 161:438–449. https://doi.org/10.1016/j.cell.2015.03.050.
19. Scheres SH. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol 180:519–530. https://doi.org/10.1016/j.jsb.2012.09.006.
20. Harbury PB, Zhang T, Kim PS, Alber T. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science 262:1401–1407. https://doi.org/10.1126/science.8248779.
21. Boudko SP, Londer YY, Letarov AV, Sernova NV, Engel J, Mesyanzhinov VV. 2002. Domain organization, folding and stability of bacteriophage T4 fibritin, a segmented coiled-coil protein. Eur J Biochem 269:833–841. https://doi.org/10.1046/j.1432-1033.2002.02734.x.
22. Grigorieff N. 2016. Frealign: an exploratory tool for single-particle cryo-EM. Methods Enzymol 579:191–226. https://doi.org/10.1016/bs.mie.2016.04.013.
23. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. 2008. Molecular architecture of native HIV-1 gp120 trimers. Nature 455:109–113. https://doi.org/10.1038/nature07159.
24. Wilson IA, Skehel JJ, Wiley DC. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature 289:366–373. https://doi.org/10.1038/289366a0.
25. Chen J, Lee KH, Steinhauer DA, Stevens DJ, Skehel JJ, Wiley DC. 1998. Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. Cell 95:409–417. https://doi.org/10.1016/S0092-8674(00)81771-7.
26. Dev J, Park D, Fu Q, Chen J, Ha HJ, Ghantous F, Herrmann T, Chang W, Liu Z, Frey G, Seaman MS, Chen B, Chou JJ. 2016. Structural basis for membrane anchoring of HIV-1 envelope spike. Science 353:172–175. https://doi.org/10.1126/science.aaf7066.
27. Sanders RW, van Gils MJ, Derking R, Sok D, Ketas TJ, Burger JA, Ozorowski G, Cupo A, Simonich C, Goo L, Arendt H, Kim HJ, Lee JH, Pugach P, Williams M, Debnath G, Moldt B, van Breemen MJ, Isik G, Medina-Ramirez M, Back JW, Koff WC, Julien JP, Rakasz EG, Seaman MS, Guttman M, Lee KK, Klasse PJ, LaBranche C, Schief WR, Wilson IA, Overbaugh J, Burton DR, Ward AB, Montefiori DC, Dean H, Moore JP. 2015. HIV-1 vaccines. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science 349:aac4223. https://doi.org/10.1126/science.aac4223.
28. Frey G, Peng H, Rits-Volloch S, Morelli M, Cheng Y, Chen B. 2008. A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. Proc Natl Acad Sci U S A 105:3739–3744. https://doi.org/10.1073/pnas.0800255105.
29. Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol 152:36–51. https://doi.org/10.1016/j.jsb.2005.07.007.
30. Grant T, Grigorieff N. 2015. Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6. eLife 4:e06980. https://doi.org/10.7554/eLife.06980.
31. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ. 2007. EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol 157:38–46. https://doi.org/10.1016/j.jsb.2006.05.009.
32. Shaikh TR, Gao H, Baxter WT, Asturias FJ, Boisset N, Leith A, Frank J. 2008. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat Protoc 3:1941–1974. https://doi.org/10.1038/nprot.2008.156.
33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.
34. Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG. 1997. Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure 5:789–798. https://doi.org/10.1016/S0969-2126(97)00233-5.