

## DESIGNING A SOLUBLE NEAR FULL-LENGTH HIV-1 GP41 TRIMER \*

Guofen Gao<sup>1</sup>, Lindsay Wieczorek<sup>2</sup>, Kristina K. Peachman<sup>2</sup>, Victoria R. Polonis<sup>3</sup>, Carl R. Alving<sup>3</sup>,  
Mangala Rao<sup>3</sup> and Venigalla B. Rao<sup>1</sup>

From <sup>1</sup>Department of Biology, The Catholic University of America, Washington, DC.

<sup>2</sup>US Military HIV Research Program, Henry M. Jackson Foundation for the Advancement of Military  
Medicine, Silver Spring, MD.

<sup>3</sup>US Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD.

\*Running Title: *Soluble HIV-1 gp41 trimer*

**Keywords:** HIV-1 envelope; gp41 trimer; bacteriophage T4 display; prehairpin fusion intermediate.

**Background:** The envelope glycoprotein gp41 is a key component of HIV-1 virus entry into host cells.

**Results:** Structure-based mutagenesis and biochemical approaches lead to the design of soluble gp41 trimers.

**Conclusion:** Trimers are in a prehairpin structure, similar to that observed during virus entry.

**Significance:** The new gp41 recombinants could lead to the development of novel diagnostics, therapeutics, and vaccines.

### SUMMARY

The HIV-1 envelope spike is a trimer of heterodimers composed of an external glycoprotein gp120 and a transmembrane glycoprotein gp41. gp120 initiates virus entry by binding to host receptors whereas gp41 mediates fusion between viral and host membranes. Although the basic pathway of HIV-1 entry has been extensively studied, the detailed mechanism is still poorly understood. Design of gp41 recombinants that mimic key intermediates is essential to elucidate the mechanism as well as to develop potent therapeutics and vaccines. Here, using molecular genetics and biochemical approaches, a series of hypotheses were tested to overcome the extreme hydrophobicity of HIV-1 gp41 and design a soluble near full-length gp41 trimer. The two long heptad repeat helices HR1 and HR2 of gp41 ectodomain were mutated to disrupt intra-molecular HR1-HR2 interactions but not inter-molecular HR1-HR1 interactions. This resulted in reduced

aggregation and improved solubility. Attachment of a 27-aa foldon at the C-terminus and slow refolding channeled gp41 into trimers. The trimers appear to be stabilized in a prehairpin-like structure, as evident from binding of a HR2 peptide to exposed HR1 grooves, lack of binding to hexa-helical bundle-specific NC-1 mAb, and inhibition of virus neutralization by broadly neutralizing antibodies 2F5 and 4E10. Fusion to T4 small outer capsid protein, Soc, allowed display of gp41 trimers on the phage nanoparticle. These approaches for the first time led to the design of a soluble gp41 trimer containing both the fusion peptide and the cytoplasmic domain, providing insights into mechanism of entry and development of gp41-based HIV-1 vaccines.

Acquired immunodeficiency syndrome (AIDS) caused by the human immunodeficiency virus type 1 (HIV-1) is a major global health epidemic. Although effective chemotherapeutics have been discovered, these inhibit virus replication after infection has already occurred (1,2). A preventative vaccine that can block HIV-1 entry at the site of infection is probably the best strategy to control the epidemic (3-5). Of the four large vaccine efficacy trials conducted in humans so far, only the RV144 trial showed a modest but significant protection (31.2%) from HIV-1 infection (6). Development of an effective HIV-1 vaccine remains as one of the biggest challenges, mainly because of the extreme genetic diversity of HIV-1 (7). Coupled with this diversity are the masking of essential epitopes by glycosylation and the extraordinary evolution of viral envelope to

| Report Documentation Page                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                    |                                     |                                                           | Form Approved<br>OMB No. 0704-0188                  |                                 |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-------------------------------------|-----------------------------------------------------------|-----------------------------------------------------|---------------------------------|
| Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. |                                    |                                     |                                                           |                                                     |                                 |
| 1. REPORT DATE<br><b>26 NOV 2012</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                    | 2. REPORT TYPE                      |                                                           | 3. DATES COVERED<br><b>00-00-2012 to 00-00-2012</b> |                                 |
| 4. TITLE AND SUBTITLE<br><b>Designing a Soluble Near Full-Length HIV-1 GP41 Trimer</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                    |                                     |                                                           | 5a. CONTRACT NUMBER                                 |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5b. GRANT NUMBER                                    |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5c. PROGRAM ELEMENT NUMBER                          |                                 |
| 6. AUTHOR(S)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                    |                                     |                                                           | 5d. PROJECT NUMBER                                  |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5e. TASK NUMBER                                     |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5f. WORK UNIT NUMBER                                |                                 |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><b>The Catholic University of America, Department of Biology, 620 Michigan Ave. NE, Washington, DC, 20064</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                    |                                     |                                                           | 8. PERFORMING ORGANIZATION REPORT NUMBER            |                                 |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                    |                                     |                                                           | 10. SPONSOR/MONITOR'S ACRONYM(S)                    |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)              |                                 |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT<br><b>Approved for public release; distribution unlimited</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                    |                                     |                                                           |                                                     |                                 |
| 13. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                    |                                     |                                                           |                                                     |                                 |
| 14. ABSTRACT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                    |                                     |                                                           |                                                     |                                 |
| 15. SUBJECT TERMS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                    |                                     |                                                           |                                                     |                                 |
| 16. SECURITY CLASSIFICATION OF:                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     | 17. LIMITATION OF ABSTRACT<br><b>Same as Report (SAR)</b> | 18. NUMBER OF PAGES<br><b>25</b>                    | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT<br><b>unclassified</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | b. ABSTRACT<br><b>unclassified</b> | c. THIS PAGE<br><b>unclassified</b> |                                                           |                                                     |                                 |

evade host immune responses (8). A major goal of HIV-1 vaccine development, therefore, is to understand the entry mechanism in detail and identify conserved intermediates that could serve as immunogens as well as targets for therapeutics and antibodies (Abs) that can block virus entry (4,9).

HIV-1, a “spherical” enveloped retrovirus, fuses with the plasma membrane of a host cell and delivers the mature core into the cytosol. A key component of entry is the trimeric spike embedded in the lipid bilayer of the viral envelope. It is a trimer of heterodimers, each dimer consisting of an extracellular glycoprotein gp120 and a transmembrane glycoprotein gp41 that are derived from proteolytic cleavage of the precursor protein gp160 (10). HIV-1 entry involves a series of initial interactions between the virus and host cell receptors. The virus is first captured through relatively weak interactions between gp120 and surface molecules, such as  $\alpha_4\beta_7$  integrin and DC-SIGN (11-13), which then leads to high affinity interactions with CD4, the primary receptor on CD4<sup>+</sup> T cell (14). A conformational change in gp120 exposes the binding site for the chemokine co-receptor, CCR5 or CXCR4 (15). Further conformational changes lead to the opening up of gp41’s two long helices containing heptad repeat (HR) sequences HR1 and HR2 and insertion of the N-terminal fusion peptide into the host cell membrane (16,17). A prehairpin intermediate, a three-stranded coiled coil stabilized by inter-molecular interactions between HR1 helices, is formed (Fig. 1A and B). gp120 subunits dissociate allowing the HR2 helices at the base of the spike to fold back and interact with the HR1 helices. The hexa-helical bundle thus formed brings the host and viral membranes in close proximity facilitating membrane fusion and release of the mature core into the cytosol (18-20).

Understanding the structure and function of the intermediates is essential to design immunogen mimics that induce broadly neutralizing antibodies (bnAbs) against genetically diverse HIV-1 viruses (4,21,22). In fact, the conserved membrane proximal external region (MPER), which is present at the base of the spike between the HR2 helices and the transmembrane domain (Fig. 1B), consists of epitopes that are recognized by a series of bnAbs, 2F5 and 4E10 being the most well-characterized among them (23-26). Passive

immunization with these bnAbs reduced viremia in HIV-1 infected individuals and nonhuman primates (27-29). The MPER epitopes are well exposed in the prehairpin intermediate (Fig. 1B), the most extended conformation of gp41 ectodomain, making it as a prime target for immunogen design (30-32).

Although the crystal structure of the hexa-helical bundle intermediate (see Fig. 1B and Fig. 3A), the core of fusion-active gp41, has been determined (33), very little is known about the structure and function of the prehairpin intermediate (31,34,35). Attempts to produce any form of full-length gp41 in a soluble, trimeric state have not been successful because of the unusually high hydrophobicity of gp41 and its extreme propensity to precipitate (36). Only certain truncated or structurally constrained versions of gp41 ectodomain, containing only HR1, HR2, and MPER motifs, have been produced but most of these collapse into a hexa-helical bundle conformation and induce either weak or no bnAbs (36-40). Other components of the gp41 molecule, such as the fusion peptide and the cytoplasmic domain might be necessary to generate a structure that mimics the native prehairpin intermediate, displaying the MPER and other functional motifs in the right conformation (41,42). However, there have been no reports of soluble, structurally-defined, gp41 oligomers containing the fusion peptide and/or cytoplasmic domain.

Here, we report the design of near full-length soluble gp41 recombinants containing the fusion peptide, the ectodomain, and the cytoplasmic domain. Our design includes introduction of mutations that weaken intra-molecular interactions between HR1 and HR2 helices while retaining inter-molecular interactions between HR1 helices. Such mutations minimized nonspecific interactions and improved the solubility of gp41. Attachment of foldon, a phage T4 trimerization tag along with slow refolding led to folding of gp41 protein into trimers and defined oligomers. These gp41 trimers were displayed on bacteriophage T4 capsid nanoparticles by attaching to the small outer capsid protein (Soc), which also forms trimers by binding to the quasi-3-fold axes of the virus capsid (43). These gp41 recombinants potently inhibited HIV-1 virus neutralization by 2F5 and 4E10 mAbs, presumably by competing with the prehairpin structure formed during virus

entry. These approaches have led to the design of soluble near full-length, gp41 trimers in a prehairpin-like structure that could be utilized to understand the mechanism of viral entry and to develop HIV-1 vaccines, diagnostics, and therapeutics.

## EXPERIMENTAL PROCEDURES

**Construction of the Expression Vectors**—All the gp41 constructs were generated by splicing-by-overlap extension PCR using wild-type HXB2 gp41 DNA as a template (44). Mutations were introduced using primers containing the desired mutations in the nucleotide sequence. For construction of gp41 fusion recombinants, the DNA fragments corresponding to gp41, Soc, and foldon were amplified by PCR using the respective DNA templates and overlapping primers containing additional amino acids (aa) SASA as a linker between each fragment. The fragments were then stitched together and the stitched DNA was amplified using the end primers containing unique restriction sites, Xho I or Nco I. The final PCR product was digested with Xho I and Nco I and ligated with the linearized and dephosphorylated pTriEx-4 Neo plasmid vector (Merck KGaA, Darmstadt, Germany). The recombinant DNA was transformed into *E. coli* XL-10 Gold competent cells (Stratagene, La Jolla, CA), and miniprep plasmid DNA was prepared from individual colonies. The presence of DNA insert was identified by restriction digestion and/or amplification with insert-specific primers. The accuracy of the cloned DNA was confirmed by DNA sequencing (Davis Sequencing, Inc., Davis, CA). The plasmids were then transformed into *E. coli* BL21 (DE3) RIPL competent cells (Stratagene) for protein expression.

**Expression and Solubility Testing of gp41 Recombinant Proteins**—*E. coli* BL21 (DE3) RIPL cells containing gp41 clones were induced with 1 mM IPTG at 30 °C for 3 h. The cells were lysed using bacterial protein extraction reagent (B-PER; Thermo Fisher Scientific Inc., Rockford, IL) and centrifuged at 12,000 g for 10 min. The soluble supernatant and insoluble pellet fractions were analyzed by SDS-PAGE. The pellets containing the insoluble inclusion bodies were treated with different denaturing reagents, SDS, urea, or guanidine hydrochloride (GnHCl). After centrifugation at 12,000 g for 10 min, the

supernatants and pellets were analyzed by SDS-PAGE.

**Purification of Recombinant Proteins**—The cells after IPTG induction were harvested by centrifugation at 8,200 g for 15 min at 4 °C and lysed using an Aminco French press (Thermo Fisher Scientific Inc.). The inclusion bodies containing the gp41 recombinant protein were separated from the soluble fraction by centrifugation at 34,000 g for 20 min. The inclusion bodies pellet from 1 L culture was dissolved in 50 ml of 50 mM Tris-HCl (pH 8), 300 mM NaCl, and 20 mM imidazole buffer containing 8 M urea. After incubation at room temperature for 30 min, the sample was centrifuged at 34,000 g for 20 min to remove cell debris. The supernatant was loaded onto a HisTrap HP column (GE Healthcare) pre-equilibrated with the same buffer. The bound protein was eluted with 20-500 mM linear imidazole gradient in the same buffer at 4 °C. A slow refolding procedure (described below) was performed to refold the purified protein. The protein was further purified by Superdex 200 gel filtration chromatography (Hiload prep grade; GE Healthcare) in 20 mM Tris-HCl (pH 8) and 100 mM NaCl buffer at 4 °C. For the gp41 recombinants expressed as soluble proteins, the supernatant of cell lysate was purified by HisTrap and Superdex 200 gel filtration columns. The purified proteins were stored frozen at -80 °C.

**Refolding of gp41 Recombinants**—Following purification by HisTrap chromatography in 8 M urea, the protein was refolded by slow dialysis with incrementally decreasing the urea concentration (6 M, 4 M, 2 M, 1 M, 0.5 M, or no urea). The dialysis buffer in addition contained 20 mM Tris-HCl (pH 8), 100 mM NaCl, 200 mM L-Arg, and 5 mM DTT. Protein was dialyzed for at least 8 h at 4 °C before changing to another buffer with decreasing concentration of urea. At the last step, the protein was dialyzed against either 20 mM Tris-HCl (pH 8) and 100 mM NaCl buffer, or PBS (pH 7.4), for 6 h and the buffer was changed every 2 h.

**SDS-Polyacrylamide Gel Electrophoresis (PAGE), and Native PAGE**—Twelve percent SDS-polyacrylamide gel (PAG) was used to determine the expression, solubility, and purification quality of gp41 recombinant proteins. The proteins were stained with Coomassie Blue R-

250 (Bio-Rad, Hercules, CA). Native-PAGE (4-20% gradient gels; Invitrogen) was used to determine the folding and oligomeric states of the recombinant proteins. The proteins were stained with Bio-safe Coomassie Stain (Bio-Rad).

**Enzyme-linked Immunosorbent Assay (ELISA)**—The gp41 recombinant proteins (or the control Soc protein) were coated on the microtiter plates at 0.25 µg/well (coating buffer: 100 mM NaHCO<sub>3</sub>, pH 9.6). Plates were incubated at 4 °C overnight and blocked with 5% BSA in PBS at 37 °C for 1 h. One hundred microliters of the monoclonal antibodies (mAbs) 2F5, 4E10 (Polymun), or NC-1 (AIDS Reagent Program) were added at 5-fold serial dilutions starting from 10 µg/ml and the plates were incubated at 37 °C for 1 h. The HRP-conjugated anti-human IgG (for 2F5 and 4E10) or anti-mouse IgG (for NC-1) were added and the samples were incubated further for 1 h at 37 °C. The TMB Microwell Peroxidase Substrate (KPL, Inc.) was added and the reaction was terminated by adding the TMB BlueSTOP Solution (KPL, Inc.). The absorbance at 650 nm was measured using the VersaMax microplate reader.

**Peptide Binding Assay**—The HR1 (N36) or HR2 (C34) peptides were added to the Soc-gp41M-Fd protein purified by HisTrap chromatography as described above. The protein was refolded by slow dialysis using a 2 kDa cut-off membrane (Amicon). An aliquot of the sample was then analyzed by native PAGE to determine the folding and oligomeric state of the protein. The rest of the sample was further dialyzed overnight against PBS using a 10 kDa cut-off membrane to remove the unbound peptide. The final sample was electrophoresed on a SDS-PAGE, which separates the Soc-gp41M-Fd-peptide complex into two bands. The amount of bound peptide was quantified by laser densitometry.

**Pseudovirus Neutralization Competition Assay**—TZM/bl cells were used to determine HIV-1 neutralization by 2F5 and 4E10 mAbs (45,46). The mAb was titrated in 3-fold serial dilutions starting at 50 µg/ml in the growth medium [DMEM with 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine (Quality Biologics Inc.), and 15% fetal calf serum (Gemini Bio-Products)]. On a 96-well flat-bottom black plate, 12.5 µl of the mAb at different dilutions was mixed with 12.5 µl of gp41 recombinant proteins

or other control competitors at a concentration of 120 nM for 2F5 neutralization, and 200 nM for 4E10 neutralization. The samples were incubated for 30 min at 37 °C and 25 µl of pseudovirus SF162 at a dilution optimized to yield ~150,000 relative luminescence units (RLUs) was added. The pseudovirus SF162 is a neutralization sensitive B clade virus. It was prepared by transfecting 5 x 10<sup>6</sup> exponentially dividing HEK 293T cells in 20 ml growth medium (DMEM) with 8 µg of env expression plasmid and 24 µg of an env-deficient HIV-1 backbone vector (pSG3ΔEnv) using FuGene as the transfection reagent (Roche, Indianapolis, IN). The SF162 env plasmid was obtained from the AIDS Reagents Program. Pseudovirus-containing culture supernatants were harvested 3 days after transfection, centrifuged, titered, and stored at -80 °C in 1 ml aliquots. The SF162 pseudovirus thus obtained is entry-competent but not replication-competent. Upon entry, it activates the HIV-1 Tat controlled Luciferase reporter gene.

Following addition of SF162 pseudovirus as above, the samples were incubated for an additional 30 min. TZM/bl cells (50 µl; 2 x 10<sup>5</sup> cells/ml in growth medium containing 60 µg/ml DEAE-dextran) was added to each well. Each plate included wells with cells and pseudovirus (virus control) or cells alone (background control). The assay was also performed by omitting the first incubation of gp41 with 2F5 or 4E10. The plates were incubated for 48 h, and then 100 µl/well of reconstituted Brite Lite Plus (Perkin-Elmer) was added. The RLUs were measured using a Victor 2 luminometer (Perkin-Elmer). The percent inhibition due to the presence of the mAb was calculated by comparing RLU values from wells containing mAb to well with virus control. IC50 was calculated for each mAb alone and mAb pre-mixed with gp41 recombinant proteins or other control competitors. Two independent assays were performed and the results were averaged (45,46).

**In vitro Display of Soc-gp41 Trimers on Phage T4 Capsid**—*hoc*<sup>-</sup>*soc*<sup>-</sup> phage was purified by velocity sucrose gradient centrifugation. About 2 x 10<sup>10</sup> PFU of purified *hoc*<sup>-</sup>*soc*<sup>-</sup> phage were centrifuged in 1.5 ml LoBind Eppendorf tubes at 18,000 g, 4 °C for 45 min. The pellets were resuspended in 10 µl PBS buffer. Purified Soc-gp41 fusion proteins were added at the desired concentration and the reaction mixture (100 µl)

was incubated at 4 °C for 45 min. Phage was sedimented by centrifugation as described above, and the pellets were washed twice with 1 ml PBS and resuspended in 10 to 20 µl of the same buffer. The sample was transferred to a fresh Eppendorf tube and analyzed by SDS-PAGE. The density volumes of bound and unbound proteins were determined by laser densitometry. The copy number of displayed gp41 was calculated in reference to the known copy number of the major capsid protein gp23\* (930 copies per phage) (“\*” represents the cleaved form of the major capsid protein gp23) or the tail sheath protein gp18 (138 copies per phage) in the respective lane. The data were plotted as one site saturation ligand binding curve and fitted by non-linear regression using the SigmaPlot10.0 software.

## RESULTS

**gp41 Recombinant Design**—The design of gp41 recombinant proteins has been extremely difficult for several reasons. First, gp41 structure is stabilized by interactions with gp120 in the native envelope trimer (47). Separation from gp120 leads to exposure of highly hydrophobic regions such as fusion peptide, HR1 and HR2 helices, and MPER (Fig. 1A). Nonspecific high-avidity interactions between these regions during heterologous protein expression lead to aggregation of nascent polypeptide chains. Second, a series of interacting residues (hydrophobic and charged) of HR1 and HR2 helices (see Fig. 3A) favor combinatorial, rather than unique, interactions among the polypeptide chains (33). Third, gp41 contains four cysteines (Fig. 1A), which can form nonspecific crosslinks, especially in an aggregated state where the tightly packed polypeptide chains exclude water. We hypothesized that these problems can be addressed by rational modification of gp41 sequence and structure by: i) introduction of mutations, ii) attachment of tags, and iii) controlling folding kinetics (Fig. 1C).

**Mutations.** Introduction of mutations that disrupt intra-molecular HR1-HR2 interactions should disfavor the formation of hexa-helical bundle and stabilize gp41 in a prehairpin intermediate structure, where the chains would be held by inter-molecular HR1-HR1 interactions and the MPER epitopes would be better exposed (30,31).

**Tags.** Attachment of a trimerization tag such as the phage T4 foldon might help nucleate gp41 folding into a trimer (48). Fusion to Soc, which forms a trimer on T4 capsid, would display gp41 trimers on the phage nanoparticle (43,49).

**Folding.** Controlling the kinetics of gp41 folding by varying the protein concentration, denaturants, and reducing conditions could channel the folding pathway towards trimer assembly.

**Deletion of Immunodominant (ID) Region**—We hypothesized that deletion of part of the apical loop between HR1 and HR2 helices (Fig. 2A; aa Q577-T605), will have important consequences for gp41 recombinant design: i) this sequence was reported to consist of ID epitopes (50-52). Although strong Ab responses are directed towards this region, these Abs do not neutralize the virus. On the other hand, they might enhance HIV-1 infection through a complement-mediated mechanism (53,54). Deletion of this region therefore could improve the immunogenicity of gp41 by diverting the Ab responses to the relatively poorly immunogenic MPER epitopes (32). ii) since this sequence consists of two cysteine residues (C598 and C604), their deletion would minimize disulphide crosslinking and insolubilization. iii) deletion of 24 of the 46 aa of the loop would favor the tri-helical prehairpin structure rather than the hexa-helical bundle that requires kinking of the intervening loop (Fig. 1B).

Two near full-length recombinant gp41 proteins were constructed, one with the ID sequence (Soc-gp41) and another without it (Soc-gp41ΔID), containing the fusion peptide, the ectodomain and the cytoplasmic domain, but not the 22-aa transmembrane domain (L684 to V705) (Fig. 2A) (transmembrane domain was found to be toxic to *E. coli*; data not shown). Soc-fusions with a 4-aa flexible linker (SASA) in between Soc and gp41 were used in these experiments because the constructs are eventually displayed on T4 phage (see below). Both Soc-gp41 and Soc-gp41ΔID recombinant proteins were over-expressed in *E. coli* (~20% of total cell protein) (Fig. 2B and 2C, lanes 3) and as predicted, partitioned into insoluble inclusion bodies (Fig. 2B and 2C; compare lanes 4 of soluble fraction with lanes 5 of insoluble fraction). However, they exhibited distinct solubilization behavior (Fig. 2D); Soc-gp41 could not be solubilized either with 8 M urea

or 6 M GnHCl (Fig. 2D, upper panel, lanes 4 and 6), whereas Soc-gp41 $\Delta$ ID was nearly completely solubilized under the same conditions (Fig. 2D, lower panel, lanes 4 and 6) and could be purified to near homogeneity by HisTrap affinity chromatography (Fig. 2E). Furthermore, the concentration of urea could be reduced to 2 M and the protein remained in solution. However, precipitation occurred when the urea concentration was further reduced. On the other hand, the Soc-gp41 protein required SDS, a strong ionic detergent, for solubilization. Even with SDS, only partial solubilization was achieved (Fig. 2D, upper panel, lane 2), and SDS was required throughout purification to maintain solubility.

**Mutations in HR1 and HR2 Helices**—A series of interactions between HR1 and HR2 helices are central to the assembly of a trimeric envelope structure and these interactions dynamically change during membrane fusion and virus entry (18,33,55) (Fig. 1B and Fig. 3A). These include inter-molecular interactions between the HR1 helices leading to trimerization, and intra-molecular interactions by the looping back of HR2 helices into the hydrophobic grooves between two HR1 helices (Fig. 1B and Fig. 3A) (33). We hypothesized that destabilization of the intra-molecular interactions would reduce nonspecific aggregation, but importantly, it would favor the tri-helical prehairpin structure, not the hexa-helical bundle because a combination of shortened apical loop and mutations would make it energetically unfavorable.

From the crystal structure of gp41 hexa-helical bundle (Fig. 3A) (33), we identified the interactions that if mutated would weaken the HR1-HR2 interactions but not the HR1-HR1 interactions. For instance, mutation of Arg557 to Glu would change the electrostatic attraction between Arg557 and Glu648 to electrostatic repulsion (Fig. 3B), and introduction of Glu at Leu568 would disrupt the hydrophobic interactions between Leu568 and Ile635 and at the same time create electrostatic repulsion with Glu634 (Fig. 3C). Using these principles, six mutant clones were constructed in the background of Soc-gp41 $\Delta$ ID and their solubility was compared (Fig. 3D). All the mutants over-expressed gp41 but the Mutant 5 - R557E, L565R, L568E, I635E, L645E – gave the best results, expressing the protein in soluble form (about 40% of the

expressed protein was in the soluble fraction; lane 3, marked with an arrow). Hence this construct, namely Soc-gp41 mutant (Soc-gp41M), was selected for further design.

Attempts to purify Soc-gp41M protein from cell lysate, however, were not successful as it did not bind to HisTrap column probably because the protein was misfolded and the histidine tag was buried in the structure. On the other hand, the 8 M urea solubilized protein bound to the column efficiently and could be purified to >95% purity (Fig. 3E). The protein remained soluble upon “fast” dialysis against PBS (one-step transition from 8 M urea to PBS), but the resultant protein behaved as a very high mol. wt. species by size exclusion gel filtration chromatography (Fig. 3F, blue curve). Also, it migrated as a smear by native PAGE (Fig. 3G, lane 1) suggesting that the mutant protein, even though soluble, formed hetero-disperse aggregates but not defined oligomers.

**Slow Refolding**—We then hypothesized that the folding kinetics of the extremely hydrophobic gp41 must be controlled in order to channel the process towards the correct folding and oligomerization pathway. A number of variables including protein concentration, pH, reducing agents, L-arginine, and “slow” dialysis (see Experimental Procedures) were optimized to control folding kinetics, using native PAGE as an assay [L-arginine suppresses protein aggregation and enhances refolding (56)]. Misfolded and aggregated protein would not enter the native gel or migrate as a smear, whereas the folded species would show distinct bands.

Data from a large series of experiments showed that slow dialysis against Tris-HCl buffer, pH 8.0-9.0, protein concentration between 0.25 to 1 mg/ml, 5 mM DTT, and 200 mM L-arginine gave the best results. The gel filtration elution profile of the refolded gp41 under these conditions showed a shift from large aggregates (void volume; Fig. 3F, blue curve) to oligomers (Fig. 3F, pink curve). Native PAGE showed that a portion of gp41 folded into defined oligomers as evident by the appearance of a ladder of bands (Fig. 3G, lane 2, indicated by arrows). However, most of Soc-gp41M still remained as soluble aggregates and stayed near the well (see Fig. 3G, lane 2).

**Trimerization Using Foldon Tag**—Foldon, a 27-aa trimerization domain of T4 fibrin, has been extensively used to trimerize foreign domains and

proteins (31,48). We hypothesized that attaching the foldon sequence to gp41 might nucleate trimerization of gp41 at the initial step of the folding pathway. We constructed Soc-gp41M-Fd as well as Soc-gp41ectoM-Fd in which the cytoplasmic domain was deleted (Fig. 4A). Both the proteins were over-expressed and purified. The results showed that foldon, as predicted, dramatically altered the folding and oligomeric states of gp41, producing trimers and higher order oligomers, and the solubility was also further improved. The Soc-gp41M-Fd protein purified from either the soluble fraction (~500 µg/L culture, Fig. 4B, lane 2), or the insoluble fraction (~20 mg/L culture, Fig. 4B, lane 3) behaved similarly, producing trimers and defined oligomers (Fig. 4C, lanes 1 and 2). That the lowermost band in the ladder is a trimer was determined by the elution volume (Fig. 4D) of this species in comparison with the elution volumes of a series of known standard proteins used to calibrate the gel filtration column. The next higher oligomer band in the ladder was determined to be a hexamer. Indeed, unlike the Soc-gp41M which produced mostly aggregates, essentially all the foldon-attached Soc-gp41M-Fd and Soc-gp41ectoM-Fd proteins were recovered as trimers and oligomers (Fig. 4C, lanes 1-3). The gp41 trimers and oligomers could be separated on a size exclusion column (Fig. 4D and E). Indeed, fractions containing mostly trimers could be purified by this method. The distribution of the oligomers did not, however, change by a second round gel filtration of trimer fractions suggesting that the gp41 subunit interactions are of high avidity and not in a dynamic equilibrium. We speculate that the basic gp41 oligomer unit is a trimer. Hexamers (and higher order oligomers) are most likely dimers (or multimers) of trimers formed by (nonspecific) interactions between gp41 trimers. Although both Soc-gp41M-Fd and Soc-gp41ectoM-Fd gave similar oligomerization patterns (Fig. 4D and 4E), we found that a greater fraction of the near full-length gp41 oligomerized into trimers than that of the ectodomain construct [Fig. 4C, compare lane 2 (Soc-gp41M-Fd) and 3 (Soc-gp41ectoM-Fd); compare Fig. 4D (Soc-gp41M-Fd) and E (Soc-gp41ectoM-Fd)], suggesting that the bulky cytodomain might have stabilized trimers, probably by restricting trimer-trimer interactions.

*gp41 Trimers Have a Prehairpin-like Structure*—For the reasons described above, the gp41M-Fd mutants are predicted to be stabilized in a prehairpin structure. We tested this prediction by two approaches. First, if the gp41M-Fd construct has a prehairpin-like structure, it should not be recognized by the NC-1 mAb which is raised against, and specific to, the hexa-helical bundle structure (40,57). However, NC-1 mAb also binds to the trimer of HR1 peptide probably because its structure is similar to that of the HR1 trimer within the hexa-helical bundle (58). Our ELISA data showed that both the Soc-gp41M-Fd and Soc-gp41ectoM-Fd proteins do not bind to NC-1 mAb, whereas the Soc-gp41 and Soc-gp41ΔID proteins that lacked the mutations bound strongly (Fig. 5A). These data indicate that the structure of Soc-gp41M-Fd trimer is distinct from that of the classic hexa-helical bundle, probably prehairpin-like, consistent with the recent evidence that the HR1 helices are less tightly packed in the pre-fusion state (55). However, since the specific epitope sequence recognized by the NC-1 is unknown, the possibility that the mutations in Soc-gp41M-Fd also affected NC-1 mAb binding cannot be ruled out.

Second, in a prehairpin structure, the groove between HR1 helices would be well exposed. Hence, an externally added HR2 peptide should be able to interact with the groove (16,33). To test this hypothesis, a 34-aa HR2 peptide (C34, 4 kDa) was added to Soc-gp41M-Fd and the unbound peptide was removed by extensive dialysis using a 10 kDa cut-off membrane. If gp41 trimer is in prehairpin state, it would capture the C34 peptide and form a gp41-C34 complex. The results demonstrated that the C34 peptide was retained with gp41 (Fig. 5B, lane 1). In fact, the ratio of gp41 to C34 in the complex remained the same whether the molar amount of C34 used was 2-times that of gp41 (Fig. 5B, lane 1) or 20-times that of gp41 (Fig. 5B, lane 2). On the other hand, addition of a 36-aa HR1 (N36) peptide resulted in the precipitation of gp41 probably due to uncontrolled HR1-HR1 interactions. Secondly, the folding pattern of gp41 was unaffected by C34 (Fig. 5C, compare lane 1 without C34 to lane 2 with C34), which means that the conformation of gp41 with and without C34 binding was the same. Since C34 binding to HR1 is expected to occur only in the prehairpin conformation, it can be



inferred that gp41 folded into the same conformation even in the absence of C34.

**Neutralizing MPER Epitopes are Well-exposed in gp41 Trimers**—The bnAbs 2F5 and 4E10 bind to the conserved MPER epitopes of gp41 and block HIV-1 entry, presumably by arresting fusion at the prehairpin stage where the epitopes would be well-exposed and the ectodomain is most extended (31,34,35) (see Fig. 1B). Consistent with this hypothesis, these mAbs have the highest affinity to the prehairpin gp41 intermediate (30,31). If the trimeric Soc-gp41M-Fd and Soc-gp41ectoM-Fd have a prehairpin structure, they should bind to 2F5 and 4E10 mAbs at high affinity and inhibit their ability to neutralize HIV-1 infection. ELISA data showed that both of the constructs bound to 2F5 and 4E10 mAbs strongly (Fig. 6A and B). To test if this interaction can titrate out the mAbs and block their ability to neutralize HIV-1, a virus neutralization competition assay was performed. Soc-gp41M-Fd and Soc-gp41ectoM-Fd were added to the TZM/bl pseudovirus neutralization reaction mixture at varying molar ratios of gp41 to mAb, and the amounts of Abs for 50% virus neutralization inhibition (IC<sub>50</sub>) were determined. The data demonstrated that both the constructs potently inhibited virus neutralization (Fig. 6C and D). gp41 concentration as low as 120 nM was sufficient to compete with the virus for binding to 2F5 and 4E10, causing a 7 to 10-fold raise in the IC<sub>50</sub> values (Fig. 6C and D). At a 1:1 molar ratio of gp41 to mAb, 45-76% inhibition of virus neutralization was observed. The near full-length gp41 showed slightly higher inhibition than the ectodomain construct. No significant difference was observed whether gp41 was preincubated with the mAb or added directly to the neutralization mixture. Validating these results, the 23-aa MPER linear peptide, but not the scrambled MPER peptide, inhibited 2F5 neutralization (Fig. 6C and D). Also, the MPER peptide did not affect 4E10 neutralization, consistent with the fact that the 4E10 mAb recognizes a conformational epitope. Neither the gp41 cytoplasmic domain (Soc-cyto) nor Soc controls showed significant inhibition, attesting to the specificity of gp41-2F5/4E10 interactions. These results further support that the trimeric gp41M-Fd constructs are stabilized in a prehairpin structure exposing the MPER

neutralization epitopes in a functionally relevant conformation.

**Display of gp41 Trimers on the Bacteriophage T4 Nanoparticle**—Eight hundred and seventy copies of a small outer capsid protein, Soc (9 kDa), decorate the surface of T4 capsid (43,59). Soc is a monomer in solution but trimerizes upon binding to capsid at the quasi-3-fold axes (Fig. 7A). Each Soc molecule binds to two gp23\* major capsid protein subunits clamping adjacent capsomers and reinforcing the capsid structure. Both the C- and N-termini are exposed on the capsid surface, with the C-termini at the quasi-3-fold axes and the N-termini at the quasi-2-fold axes (Fig. 7A). We hypothesized that by fusing gp41 to the C-terminus of Soc and displaying it on T4, the trimeric gp41 would be stably displayed at the 3-fold axes of the phage capsid. Such particles with arrays of gp41 trimers would allow structure-function studies as well as enhance immunogenicity (64). The gp41 trimers assembled on *hoc<sup>-</sup>soc<sup>-</sup>* capsids nearly as efficiently as native Soc (Fig. 7B-F). Soc-gp41M-Fd binding increased with increasing ratios of Soc-gp41 molecules to capsid binding sites, reaching saturation at a ratio of ~20:1 (Fig. 7B). The apparent association constant ( $K_d$ ) calculated from the saturation binding curve (Fig. 7E) was 121 nM and the maximum copy number of bound gp41 ( $B_{max}$ ) was about 859 per capsid, which is close to the copy number of 870 when all the Soc binding sites are occupied. Similar binding behavior as well as  $K_d$  and  $B_{max}$  values was observed for Soc-gp41ectoM-Fd (Fig. 7C and F).

To further improve the gp41 nanoparticle design, a 13-aa cell penetration peptide (CPP), CPP-Tat (PGRKKRRQRRPPQ), was attached to the N-terminus of Soc-gp41. CPPs are 10-30 aa peptides rich in basic aa that facilitate passage of attached cargo molecules across the cell membrane (60). The CPP-Tat derived from HIV-1 trans-activator protein, TAT, is one of the most efficient CPPs (60). Our recent experiments show that T4 particles displaying targeting molecules attached to Soc are taken up by cells at high efficiency (unpublished results). CPP-Soc-gp41M-Fd could be over-expressed, purified, and bound to T4 capsid efficiently, and the binding parameters are also similar (Fig. 7D and F). Thus, CPP or another molecule such as the CD40 Ligand (61) can be oriented at the quasi-2-fold axes for

targeting of the nanoparticle to antigen presenting cells such as the dendritic cells.

## DISCUSSION

Although the key interactions between HIV-1 and host cell have been well established, the extraordinary genetic diversity of viral envelope and masking of essential epitopes by glycosylation made it difficult to design recombinants that can induce protective immune responses (62,63). However, the HIV-1 virus, like many type-1 fusion viruses, undergoes dynamic transitions during entry, exposing some of the vulnerable sites on the cell surface making them accessible to therapeutics and neutralizing Abs. The prehairpin intermediate is one such target because it is relatively stable with a half-life on the order of several minutes (19), and its ectodomain most extended and the conserved neutralization epitopes most exposed (Fig. 1B) (30,31,62,63). Indeed, Enfuvirtide, a potent 20-aa entry inhibitor approved for clinical use (64), and a series of bnAbs, such as 2F5 and 4E10, arrest virus entry by binding to this intermediate. Design of gp41 recombinants stabilized in a prehairpin structure, therefore, will have important implications for understanding the mechanism as well as for development of effective therapeutics and vaccines.

The extremely hydrophobic gp41 is notoriously prone to aggregation and attempts to produce soluble gp41 have not been successful (36). Previous studies could only produce short truncated parts of the gp41 ectodomain, most containing only the HR1 and HR2 helices (31,38,40,65). These and other synthetic peptide mimics could not elicit potent bnAbs, leading to the hypothesis that other gp41 structural and functional motifs might be essential to mimic the true prehairpin conformation (see Fig. 1B) (41,42). These might include, in addition to HR1/HR2 helices and MPER, the fusion peptide at the N-terminus and the cytoplasmic domain at the C-terminus, but none of the gp41 recombinants produced so far included these highly hydrophobic regions.

We hypothesized that three key problems should be addressed in order to generate a soluble trimeric gp41 stabilized in a prehairpin structure (Fig. 1C). First, the inter-molecular interactions between HR1 and HR2 helices that lead to hexa-

helical bundle formation as well as nonspecific aggregation should be disrupted to stabilize the molecule in a three-stranded coil. This we achieved by deleting part of the apical loop and the five C-terminal aa of HR1 helix, as well as converting some of the complementary charge-charge and hydrophobic interactions into electrostatic repulsion, leaving intact the MPER epitope residues. These modifications greatly enhanced the solubility of gp41, however only a small fraction of the protein oligomerized into trimers (Fig. 8). Attachment of a foldon tag that has strong propensity to trimerize was necessary to trimerize gp41. Presumably, the foldon helped nucleate gp41 folding and assembly into a trimer. Since the tag is present at the C-terminal end, trimerization was probably initiated at this end and propagated through the rest of the molecule leading to folding of the protein into a three-stranded coiled coil through the strong HR1-HR1 interactions. Kinetically slowing down this process at relatively low protein concentration was also necessary, otherwise nonspecific inter-chain interactions presumably channeled the protein into abortive folding pathways leading to rapid and uncontrolled aggregation.

Although our approaches yielded predicted outcomes (Fig. 8), each approach by itself was insufficient to produce gp41 trimers. For instance, introduction of mutations greatly improved solubility but the protein chains still coalesced into aggregates because folding was not trimer-directed. Both trimerization tag attachment and slow refolding were necessary to correct this problem. Although hexamers and higher order oligomers were produced in addition to trimers, the core structure of all the oligomers appears to be a trimer and the higher order oligomers are probably multimers of trimers formed by nonspecific interactions between trimers. This is not unexpected because several hydrophobic patches would be exposed in the gp41 ectodomain, which would otherwise be stabilized by interactions with the gp120 domains in the native spike. These would lead to multimerization of trimers, a commonly observed phenomenon even with the gp140 trimers produced by heterologous expression systems where only short regions of the gp41 ectodomain are exposed.

Evidence indicates that the gp41 trimers have a structure mimicking the prehairpin intermediate

in which the external grooves of the three-stranded HR1 helices were not occupied by HR2 helices (16,33). Consistent with this prediction, a 34-aa HR2 peptide bound to the gp41 trimers and the oligomerization pattern was identical with or without the peptide. The gp41 trimers did not bind to NC-1 mAb that is specific to hexa-helical bundle structure but bound strongly to bnAbs 2F5 and 4E10 that have the highest affinity to the prehairpin structure (30,31). The trimers potently inhibited 2F5 and 4E10 virus neutralization even at an equimolar ratio of gp41 to mAb, and in the presence of excess virus, suggesting that the MPER epitopes are well exposed, as would be expected in a prehairpin intermediate (30,31).

The potential use of gp41 trimer as an immunogen can be further enhanced by linking the recombinants to a robust platform that can induce strong immune responses. The bacteriophage T4 display provides a simple yet powerful strategy to convert soluble antigens into nano-particulate antigens by attaching Soc to one end of the antigen (66,67). We have previously shown that such nanoparticles displaying HIV-1 Gag p24 and other antigens induced strong Ab as well as cellular responses (67,68). Attachment of Soc to the N-terminus did not interfere with the folding or trimerization of gp41, neither did it affect binding to T4 capsid. Indeed, the Soc-binding sites on the capsid were nearly saturated, resulting in the decoration of T4 phage with ~290 trimers of gp41. Since Soc C-termini are projected outward at the quasi-3-fold axes (Fig. 7A) (43), the C-terminally

attached gp41 trimers would be extending away from the capsid surface (Fig. 8), thereby exposing the MPER epitopes for capture by antigen presenting cells. Furthermore, we show that additional targeting ligands, such as CPP, can be incorporated into the displayed gp41 to enhance the uptake of the T4-gp41 particles and potentially induce robust immune responses.

In conclusion, using molecular genetics and biochemical approaches a series of hypotheses were tested (Fig. 8), leading to the generation of soluble near full-length gp41 trimers containing the fusion peptide, the ectodomain, and the cytoplasmic domain, as well as the same arrayed on phage nanoparticles. These, for the first time, allow structure determination of this critical intermediate, screening for novel therapeutics, development of new diagnostics, and design of gp41-based HIV-1 vaccines. For instance, the gp41 trimers could be used for screening peptides that exhibit high affinity to prehairpin structure and effectively block HIV-1 infection. The soluble near full-length gp41 might be an attractive candidate for detection of gp41 Abs in HIV infected individuals. The recent RV144 trial showed a correlation between protection against HIV-1 infection and generation of Abs to the gp120 variable loop V2 (6,69). The near full-length gp41 trimers can be used in conjunction with gp120 to further improve the immunogenicity of the vaccine to induce binding and neutralizing Abs as well as cellular responses.

## REFERENCES

1. Maltez, F., Doroana, M., Branco, T., and Valente, C. (2011) Recent advances in antiretroviral treatment and prevention in HIV-infected patients. *Curr Opin HIV AIDS* **6 Suppl 1**, S21-30
2. Meadows, D. C., and Gervay-Hague, J. (2006) Current developments in HIV chemotherapy. *ChemMedChem* **1**, 16-29
3. Walker, L. M., Huber, M., Doores, K. J., Falkowska, E., Pejchal, R., Julien, J. P., Wang, S. K., Ramos, A., Chan-Hui, P. Y., Moyle, M., Mitcham, J. L., Hammond, P. W., Olsen, O. A., Phung, P., Fling, S., Wong, C. H., Phogat, S., Wrin, T., Simek, M. D., Koff, W. C., Wilson, I. A., Burton, D. R., and Poignard, P. (2011) Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466-470
4. Kwong, P. D., Mascola, J. R., and Nabel, G. J. (2012) Rational Design of Vaccines to Elicit Broadly Neutralizing Antibodies to HIV-1. *Cold Spring Harbor Perspectives in Biology* **4**
5. McElrath, M. J., and Haynes, B. F. (2010) Induction of immunity to human immunodeficiency virus type-1 by vaccination. *Immunity* **33**, 542-554
6. Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premisri, N., Namwat, C., de Souza, M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J. G.,

- Francis, D. P., Stablein, D., Birx, D. L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., Robb, M. L., Michael, N. L., Kunasol, P., and Kim, J. H. (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* **361**, 2209-2220
7. Burton, D. R., Poignard, P., Stanfield, R. L., and Wilson, I. A. (2012) Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. *Science* **337**, 183-186
8. Boutwell, C. L., Rolland, M. M., Herbeck, J. T., Mullins, J. I., and Allen, T. M. (2010) Viral evolution and escape during acute HIV-1 infection. *J Infect Dis* **202 Suppl 2**, S309-314
9. Pejchal, R., and Wilson, I. A. (2010) Structure-based vaccine design in HIV: blind men and the elephant? *Curr Pharm Des* **16**, 3744-3753
10. Wyatt, R., and Sodroski, J. (1998) The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**, 1884-1888
11. Arthos, J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D., Xiao, Z., Veenstra, T. D., Conrad, T. P., Lempicki, R. A., McLaughlin, S., Pascuccio, M., Gopaul, R., McNally, J., Cruz, C. C., Censoplano, N., Chung, E., Reitano, K. N., Kottlil, S., Goode, D. J., and Fauci, A. S. (2008) HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. *Nat Immunol* **9**, 301-309
12. Cicala, C., Martinelli, E., McNally, J. P., Goode, D. J., Gopaul, R., Hiatt, J., Jelacic, K., Kottlil, S., Macleod, K., O'Shea, A., Patel, N., Van Ryk, D., Wei, D., Pascuccio, M., Yi, L., McKinnon, L., Izulla, P., Kimani, J., Kaul, R., Fauci, A. S., and Arthos, J. (2009) The integrin alpha4beta7 forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. *Proc Natl Acad Sci U S A* **106**, 20877-20882
13. Geijtenbeek, T. B., Krooshoop, D. J., Bleijs, D. A., van Vliet, S. J., van Duijnhoven, G. C., Grabovsky, V., Alon, R., Figdor, C. G., and van Kooyk, Y. (2000) DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. *Nat Immunol* **1**, 353-357
14. Dalglish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763-767
15. Berger, E. A., Murphy, P. M., and Farber, J. M. (1999) Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* **17**, 657-700
16. Furuta, R. A., Wild, C. T., Weng, Y., and Weiss, C. D. (1998) Capture of an early fusion-active conformation of HIV-1 gp41. *Nat Struct Biol* **5**, 276-279
17. Colman, P. M., and Lawrence, M. C. (2003) The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol* **4**, 309-319
18. Melikyan, G. B., Markosyan, R. M., Hemmati, H., Delmedico, M. K., Lambert, D. M., and Cohen, F. S. (2000) Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J Cell Biol* **151**, 413-423
19. Chan, D. C., and Kim, P. S. (1998) HIV entry and its inhibition. *Cell* **93**, 681-684
20. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426-430
21. Walker, L. M., and Burton, D. R. (2010) Rational antibody-based HIV-1 vaccine design: current approaches and future directions. *Curr Opin Immunol* **22**, 358-366
22. Phogat, S., and Wyatt, R. (2007) Rational modifications of HIV-1 envelope glycoproteins for immunogen design. *Curr Pharm Des* **13**, 213-227
23. Muster, T., Steindl, F., Purtscher, M., Trkola, A., Klima, A., Himmler, G., Ruker, F., and Katinger, H. (1993) A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* **67**, 6642-6647

24. Zwick, M. B., Labrijn, A. F., Wang, M., Spenlehauer, C., Saphire, E. O., Binley, J. M., Moore, J. P., Stiegler, G., Katinger, H., Burton, D. R., and Parren, P. W. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J Virol* **75**, 10892-10905
25. Barbato, G., Bianchi, E., Ingallinella, P., Hurni, W. H., Miller, M. D., Ciliberto, G., Cortese, R., Bazzo, R., Shiver, J. W., and Pessi, A. (2003) Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. *J Mol Biol* **330**, 1101-1115
26. Brunel, F. M., Zwick, M. B., Cardoso, R. M., Nelson, J. D., Wilson, I. A., Burton, D. R., and Dawson, P. E. (2006) Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. *J Virol* **80**, 1680-1687
27. Hessel, A. J., Rakasz, E. G., Tehrani, D. M., Huber, M., Weisgrau, K. L., Landucci, G., Forthal, D. N., Koff, W. C., Poignard, P., Watkins, D. I., and Burton, D. R. (2009) Broadly neutralizing monoclonal antibodies 2F5 and 4E10 directed against the human immunodeficiency virus type 1 gp41 membrane-proximal external region protect against mucosal challenge by simian-human immunodeficiency virus SHIVBa-L. *J Virol* **84**, 1302-1313
28. Mascola, J. R., Lewis, M. G., Stiegler, G., Harris, D., VanCott, T. C., Hayes, D., Louder, M. K., Brown, C. R., Sapan, C. V., Frankel, S. S., Lu, Y., Robb, M. L., Katinger, H., and Birx, D. L. (1999) Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. *J Virol* **73**, 4009-4018
29. Mascola, J. R., Stiegler, G., VanCott, T. C., Katinger, H., Carpenter, C. B., Hanson, C. E., Beary, H., Hayes, D., Frankel, S. S., Birx, D. L., and Lewis, M. G. (2000) Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat Med* **6**, 207-210
30. Frey, G., Chen, J., Rits-Volloch, S., Freeman, M. M., Zolla-Pazner, S., and Chen, B. (2010) Distinct conformational states of HIV-1 gp41 are recognized by neutralizing and non-neutralizing antibodies. *Nat Struct Mol Biol* **17**, 1486-1491
31. Frey, G., Peng, H., Rits-Volloch, S., Morelli, M., Cheng, Y., and 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
32. Montero, M., van Houten, N. E., Wang, X., and Scott, J. K. (2008) The membrane-proximal external region of the human immunodeficiency virus type 1 envelope: dominant site of antibody neutralization and target for vaccine design. *Microbiol Mol Biol Rev* **72**, 54-84, table of contents
33. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-273
34. Chakrabarti, B. K., Walker, L. M., Guenaga, J. F., Ghobbeh, A., Poignard, P., Burton, D. R., and Wyatt, R. T. (2011) Direct antibody access to the HIV-1 membrane-proximal external region positively correlates with neutralization sensitivity. *J Virol* **85**, 8217-8226
35. de Rosny, E., Vassell, R., Jiang, S., Kunert, R., and Weiss, C. D. (2004) Binding of the 2F5 monoclonal antibody to native and fusion-intermediate forms of human immunodeficiency virus type 1 gp41: implications for fusion-inducing conformational changes. *J Virol* **78**, 2627-2631
36. Scholz, C., Schaarschmidt, P., Engel, A. M., Andres, H., Schmitt, U., Faatz, E., Balbach, J., and Schmid, F. X. (2005) Functional solubilization of aggregation-prone HIV envelope proteins by covalent fusion with chaperone modules. *J Mol Biol* **345**, 1229-1241
37. Buzon, V., Natrajan, G., Schibli, D., Campelo, F., Kozlov, M. M., and Weissenhorn, W. (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog* **6**, e1000880

38. Kim, M., Qiao, Z., Yu, J., Montefiori, D., and Reinherz, E. L. (2007) Immunogenicity of recombinant human immunodeficiency virus type 1-like particles expressing gp41 derivatives in a pre-fusion state. *Vaccine* **25**, 5102-5114
39. Nelson, J. D., Kinkad, H., Brunel, F. M., Leaman, D., Jensen, R., Louis, J. M., Maruyama, T., Bewley, C. A., Bowdish, K., Clore, G. M., Dawson, P. E., Frederickson, S., Mage, R. G., Richman, D. D., Burton, D. R., and Zwick, M. B. (2008) Antibody elicited against the gp41 N-heptad repeat (NHR) coiled-coil can neutralize HIV-1 with modest potency but non-neutralizing antibodies also bind to NHR mimetics. *Virology* **377**, 170-183
40. Wang, J., Tong, P., Lu, L., Zhou, L., Xu, L., Jiang, S., and Chen, Y. H. (2011) HIV-1 gp41 core with exposed membrane-proximal external region inducing broad HIV-1 neutralizing antibodies. *PLoS One* **6**, e18233
41. Hager-Braun, C., Katinger, H., and Tomer, K. B. (2006) The HIV-neutralizing monoclonal antibody 4E10 recognizes N-terminal sequences on the native antigen. *J Immunol* **176**, 7471-7481
42. Lorizate, M., Gomara, M. J., de la Torre, B. G., Andreu, D., and Nieva, J. L. (2006) Membrane-transferring sequences of the HIV-1 Gp41 ectodomain assemble into an immunogenic complex. *J Mol Biol* **360**, 45-55
43. Qin, L., Fokine, A., O'Donnell, E., Rao, V. B., and Rossmann, M. G. (2009) Structure of the small outer capsid protein, Soc: a clamp for stabilizing capsids of T4-like phages. *J Mol Biol* **395**, 728-741
44. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-68
45. Montefiori, D. C. (2009) Measuring HIV neutralization in a luciferase reporter gene assay. *Methods Mol Biol* **485**, 395-405
46. Polonis, V. R., Brown, B. K., Rosa Borges, A., Zolla-Pazner, S., Dimitrov, D. S., Zhang, M. Y., Barnett, S. W., Ruprecht, R. M., Scarlatti, G., Fenyo, E. M., Montefiori, D. C., McCutchan, F. E., and Michael, N. L. (2008) Recent advances in the characterization of HIV-1 neutralization assays for standardized evaluation of the antibody response to infection and vaccination. *Virology* **375**, 315-320
47. Pancera, M., Majeed, S., Ban, Y. E., Chen, L., Huang, C. C., Kong, L., Kwon, Y. D., Stuckey, J., Zhou, T., Robinson, J. E., Schief, W. R., Sodroski, J., Wyatt, R., and Kwong, P. D. (2010) Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A* **107**, 1166-1171
48. Yang, X., Lee, J., Mahony, E. M., Kwong, P. D., Wyatt, R., and Sodroski, J. (2002) Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibritin. *J Virol* **76**, 4634-4642
49. Li, Q., Shivachandra, S. B., Zhang, Z., and Rao, V. B. (2007) Assembly of the small outer capsid protein, Soc, on bacteriophage T4: a novel system for high density display of multiple large anthrax toxins and foreign proteins on phage capsid. *J Mol Biol* **370**, 1006-1019
50. Gnann, J. W., Jr., Nelson, J. A., and Oldstone, M. B. (1987) Fine mapping of an immunodominant domain in the transmembrane glycoprotein of human immunodeficiency virus. *J Virol* **61**, 2639-2641
51. Mathiesen, T., Chiodi, F., Broliden, P. A., Albert, J., Houghten, R. A., Utter, G., Wahren, B., and Norrby, E. (1989) Analysis of a subclass-restricted HIV-1 gp41 epitope by omission peptides. *Immunology* **67**, 1-7
52. Wang, J. J., Steel, S., Wisniewolski, R., and Wang, C. Y. (1986) Detection of antibodies to human T-lymphotropic virus type III by using a synthetic peptide of 21 amino acid residues

- corresponding to a highly antigenic segment of gp41 envelope protein. *Proc Natl Acad Sci U S A* **83**, 6159-6163
53. Robinson, W. E., Jr., Gorny, M. K., Xu, J. Y., Mitchell, W. M., and Zolla-Pazner, S. (1991) Two immunodominant domains of gp41 bind antibodies which enhance human immunodeficiency virus type 1 infection in vitro. *J Virol* **65**, 4169-4176
  54. Robinson, W. E., Jr., Kawamura, T., Lake, D., Masuho, Y., Mitchell, W. M., and Hersh, E. M. (1990) Antibodies to the primary immunodominant domain of human immunodeficiency virus type 1 (HIV-1) glycoprotein gp41 enhance HIV-1 infection in vitro. *J Virol* **64**, 5301-5305
  55. Tran, E. E., Borgnia, M. J., Kuybeda, O., Schauder, D. M., Bartesaghi, A., Frank, G. A., Sapiro, G., Milne, J. L., and Subramaniam, S. (2012) Structural Mechanism of Trimeric HIV-1 Envelope Glycoprotein Activation. *PLoS Pathog* **8**, e1002797
  56. Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., and Arakawa, T. (2004) Role of arginine in protein refolding, solubilization, and purification. *Biotechnol Prog* **20**, 1301-1308
  57. Jiang, S., Lin, K., and Lu, M. (1998) A conformation-specific monoclonal antibody reacting with fusion-active gp41 from the human immunodeficiency virus type 1 envelope glycoprotein. *J Virol* **72**, 10213-10217
  58. Dimitrov, A. S., Louis, J. M., Bewley, C. A., Clore, G. M., and Blumenthal, R. (2005) Conformational changes in HIV-1 gp41 in the course of HIV-1 envelope glycoprotein-mediated fusion and inactivation. *Biochemistry* **44**, 12471-12479
  59. Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B., and Rossmann, M. G. (2004) Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci U S A* **101**, 6003-6008
  60. Heitz, F., Morris, M. C., and Divita, G. (2009) Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *Br J Pharmacol* **157**, 195-206
  61. Melchers, M., Matthews, K., de Vries, R. P., Eggink, D., van Montfort, T., Bontjer, I., van de Sandt, C., David, K., Berkhout, B., Moore, J. P., and Sanders, R. W. (2011) A stabilized HIV-1 envelope glycoprotein trimer fused to CD40 ligand targets and activates dendritic cells. *Retrovirology* **8**, 48
  62. Pantophlet, R., and Burton, D. R. (2006) GP120: target for neutralizing HIV-1 antibodies. *Annu Rev Immunol* **24**, 739-769
  63. Wei, X., Decker, J. M., Wang, S., Hui, H., Kappes, J. C., Wu, X., Salazar-Gonzalez, J. F., Salazar, M. G., Kilby, J. M., Saag, M. S., Komarova, N. L., Nowak, M. A., Hahn, B. H., Kwong, P. D., and Shaw, G. M. (2003) Antibody neutralization and escape by HIV-1. *Nature* **422**, 307-312
  64. Greenberg, M., Cammack, N., Salgo, M., and Smiley, L. (2004) HIV fusion and its inhibition in antiretroviral therapy. *Rev Med Virol* **14**, 321-337
  65. Bianchi, E., Joyce, J. G., Miller, M. D., Finnefrock, A. C., Liang, X., Finotto, M., Ingallinella, P., McKenna, P., Citron, M., Ottinger, E., Hepler, R. W., Hrin, R., Nahas, D., Wu, C., Montefiori, D., Shiver, J. W., Pessi, A., and Kim, P. S. (2010) Vaccination with peptide mimetics of the gp41 prehairpin fusion intermediate yields neutralizing antisera against HIV-1 isolates. *Proc Natl Acad Sci U S A* **107**, 10655-10660
  66. Li, Q., Shivachandra, S. B., Leppla, S. H., and Rao, V. B. (2006) Bacteriophage T4 capsid: a unique platform for efficient surface assembly of macromolecular complexes. *J Mol Biol* **363**, 577-588
  67. Sathaliyawala, T., Rao, M., Maclean, D. M., Birx, D. L., Alving, C. R., and Rao, V. B. (2006) Assembly of human immunodeficiency virus (HIV) antigens on bacteriophage T4: a novel in vitro approach to construct multicomponent HIV vaccines. *J Virol* **80**, 7688-7698
  68. Peachman, K. K., Li, Q., Matyas, G. R., Shivachandra, S. B., Lovchik, J., Lyons, R. C., Alving, C. R., Rao, V. B., and Rao, M. (2011) Anthrax vaccine antigen-adjuvant formulations completely protect New Zealand white rabbits against challenge with *Bacillus anthracis* Ames strain spores. *Clin Vaccine Immunol* **19**, 11-16

69. Barouch, D. H., Liu, J., Li, H., Maxfield, L. F., Abbink, P., Lynch, D. M., Iampietro, M. J., SanMiguel, A., Seaman, M. S., Ferrari, G., Forthal, D. N., Ourmanov, I., Hirsch, V. M., Carville, A., Mansfield, K. G., Stablein, D., Pau, M. G., Schuitemaker, H., Sadoff, J. C., Billings, E. A., Rao, M., Robb, M. L., Kim, J. H., Marovich, M. A., Goudsmit, J., and Michael, N. L. (2012) Vaccine protection against acquisition of neutralization-resistant SIV challenges in rhesus monkeys. *Nature* **482**, 89-93

**Acknowledgements**—We thank Victor Padilla-Sanchez of The Catholic University of America for help with the preparation of Figures 7 and 8, and one of the reviewers for suggesting the testing of gp41M-Fd binding to NC-1 mAb.

## FOOTNOTES

\* This work was partly supported by a grant from The National Institute of Allergy and Infectious Diseases (R01AI02725).

<sup>1</sup>To whom correspondence may be addressed: Venigalla B. Rao, Department of Biology, The Catholic University of America, 620 Michigan Ave. NE, Washington, DC, 20064. Phone: (202) 319-5271; E-mail: rao@cua.edu

<sup>2</sup>The abbreviations used are: AIDS, acquired immunodeficiency syndrome; HIV-1, human immunodeficiency virus type 1; Abs, antibodies; HR, heptad repeat; bnAbs, broadly neutralizing antibodies; mAb, monoclonal antibody; MPER, membrane proximal external region; Soc, small outer capsid protein; ID, immunodominant; PAG, polyacrylamide gel; PAGE, polyacrylamide gel electrophoresis; ELISA, Enzyme-linked immunosorbent assay; RLUs, relative luminescence units; B-PER, bacterial protein extraction reagent; GnHCl, guanidine hydrochloride.

## FIGURE LEGENDS

**FIGURE 1. gp41 recombinant design.** A, Schematic representation of various regions of gp41. *FP*, fusion peptide; *PR*, fusion peptide proximal region; *HR1*, heptad repeat 1; *loop*, apical loop; *HR2*, heptad repeat 2; *MPER*, membrane proximal external region; *TM*, transmembrane helix; *Cyto*, cytoplasmic domain. *Ecto* and *Cyto* correspond to ectodomain and cytoplasmic domains, respectively. Numbers correspond to the first aa for each region using the gp41 sequence from HIV-1 strain HXB2. Positions of the four cysteines (-SH) are shown. B, Schematic diagram of HIV-1 entry with emphasis on gp41 function. Following the interaction of HIV-1 envelope with the host receptors, gp120 and gp41 undergo conformational changes resulting in the exposure of HR1 and HR2 helices and formation of a prehairpin intermediate. The HR2 helix loops back and interacts with the groove between HR1 helices forming a hexa-helical bundle (*6HB*), bringing the viral and cellular membranes to close proximity for fusion. C, Strategies to generate gp41 recombinant trimers in prehairpin structure.

**FIGURE 2. Deletion of ID region improved the solubility of gp41.** A, Schematic representation of Soc-gp41 and Soc-gp41ΔID recombinants. The aa Q577-T605 were deleted in the Soc-gp41ΔID mutant and aa L684-V705 were deleted in both the recombinants. B and C, SDS-PAG (12%) showing the protein patterns of Soc-gp41 (B) and Soc-gp41ΔID (C) without IPTG induction (0 h) or 3 hours after IPTG induction (3 h). The cells were lysed using B-PER reagent and centrifuged at 12,000 g for 10 min. The soluble supernatant (S) and insoluble pellet (P) fractions were analyzed. *Std*, molecular size standards. D, Comparison of the solubility of Soc-gp41 (upper panel) with Soc-gp41ΔID (lower panel). The insoluble inclusion bodies (IB) were treated with various denaturing reagents and centrifuged at 12,000 g for 10 min. The supernatants (S) and pellets (P) were analyzed by SDS-PAGE. E, Elution profile of Soc-gp41ΔID on a HisTrap column. The green curve represents absorbance units and the blue curve, the imidazole gradient. Inset shows SDS-PAG (12%) image of the purified protein. See Experimental Procedures for additional details.



**FIGURE 3. Construction and characterization of HR mutants.** *A*, gp41 6HB structure (PDB: 1AIK). HR1 is shown in *cyan* and HR2 in *yellow*. Side chains are shown as sticks. *B*, Arg557 of HR1 shown in *blue* was mutated to Glu to change the electrostatic interaction to repulsion with Glu648 in HR2, shown in *red*. *C*, Leu568 of HR1 shown in *purple* was mutated to Glu to disrupt hydrophobic interaction with Ile635 in HR2 (shown in *purple*) and create charge repulsion with Glu634 (shown in *red*). *D*, SDS-PAG (12%) showing the protein expression and solubility of various gp41 mutants, without IPTG induction (0 h) and 3 hours after IPTG induction (3 h). *Lanes 3 and 4* represent soluble supernatant (S) and insoluble pellet (P) fractions after the cells were lysed with B-PER reagent followed by centrifugation at 12,000 g. *Red arrow* shows the mutant (Soc-gp41M) that expressed gp41 in soluble form. *E*, SDS-PAG (12%) showing the purified Soc-gp41M protein. The protein was purified from inclusion bodies by 8 M urea denaturation followed by HisTrap column chromatography. The purified protein in 8 M urea was then dialyzed against PBS buffer ("fast" dialysis). *Std*, molecular size standards. *F*, Elution profile of Soc-gp41M by Superdex 200 gel filtration. *Blue curve* represents Soc-gp41M protein renatured by fast dialysis and *pink curve* represents the protein renatured by slow refolding. *G*, Native PAG (4-20% gradient) of purified Soc-gp41M protein renatured by fast dialysis (*lane 1*) or by slow refolding (*lane 2*). See Experimental Procedures for additional details.

**FIGURE 4. The oligomeric state of purified Soc-gp41M-Fd and Soc-gp41ectoM-Fd proteins.** *A*, Schematic representation of Soc-gp41M-Fd and Soc-gp41ectoM-Fd recombinants. *B*, SDS-PAG (12%) of purified Soc-gp41M-Fd protein. *Lane 1*, molecular size standards; *lane 2*, protein purified from the supernatant of cell lysate; *lane 3*, protein purified from the inclusion bodies by urea denaturation and slow refolding. *C*, Native PAG (4-20% gradient) of purified Soc-gp41M-Fd and Soc-gp41ectoM-Fd proteins. *Lane 1*, Soc-gp41M-Fd purified from supernatant; *lane 2*, Soc-gp41M-Fd purified from inclusion bodies; *lane 3*, Soc-gp41ectoM-Fd purified from inclusion bodies. Trimer bands were marked with *red arrows*; Hexamer bands were marked with *blue arrows*. *D and E*, Native gels (4-20% gradient) showing the oligomeric state of Soc-gp41M-Fd (*D*) and Soc-gp41ectoM-Fd (*E*) fractions following Superdex 200 gel filtration. The elution volumes and estimated mol. wt. of the fractions are labeled at the *top* of the *lanes*. The positions of trimer and hexamer are indicated with *red* and *blue arrows* respectively.

**FIGURE 5. Binding of Soc-gp41M-Fd trimers to NC-1 mAb and HR2 peptide.** *A*, The Soc-gp41M-Fd proteins do not bind to NC-1 mAb. The microtiter plates were coated with different Soc-gp41 recombinant proteins or the control Soc protein. Binding to NC-1 mAb was tested as described in Experimental Procedures. (*B and C*) Binding of HR2 peptide to Soc-gp41M-Fd. *B*, SDS-PAG (12%) showing the HR2 peptide C34 bound to Soc-gp41M-Fd. The C34 peptide was added to Soc-gp41M-Fd at a molar ratio of 2 or 20 times C34 to gp41 molecules and gp41 was refolded according to the procedure described in Experimental Procedures. The unbound peptide was removed by extensive dialysis using a 10 kDa cut-off membrane. *Lane 3*, 0.4  $\mu$ g of C34 peptide used as size standard. *C*, Native PAG (4-20% gradient) showing the oligomeric state of Soc-gp41M-Fd with or without the addition of C34 peptide (1:20 molar ratio of Soc-gp41M-Fd to C34). The samples were electrophoresed prior to removing excess C34 by dialysis. *Lane 3*, 3  $\mu$ g of C34 peptide used as size standard. The NC-1 mAb and C34 peptide were provided by the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

**FIGURE 6. Inhibition of virus neutralization by gp41 trimers.** *A and B*, The Soc-gp41M-Fd and Soc-gp41ectoM-Fd bind to 2F5 (*A*) and 4E10 (*B*) mAbs. The microtiter plates were coated with Soc-gp41M-Fd, Soc-gp41ectoM-Fd, or Soc control proteins. Binding to 2F5 and 4E10 mAbs was determined by ELISA as described in Experimental Procedures. *C and D*, Virus neutralization as determined by the TZM/bl assay (45,46). Serial dilutions of purified 2F5 (*C*) or 4E10 (*D*) IgG were added to 96-well plates. gp41 trimers or other control competitors were added to the mAb and incubated for 30 min at 37 °C. SF162 virus was added to the plate and incubated for 30 min at 37 °C, followed by the addition of TZM/bl cells. After incubation for 48 h at 37 °C, the cells were lysed, and concentration of half-maximal inhibition (IC50) was calculated from the luciferase activities determined by luminescence measurements.

The sequence of MPER peptide is LELDKWASLWNWFNITNWLWYIK(amide) and that of MPER scrambled peptide is LSINEAFKWLDWWTLNDLWYIWK(amide). Soc-cyto is the fusion of cytoplasmic domain of gp41 to the C-terminus of Soc. The protein was over-expressed and purified from *E.coli* after 8 M urea denaturation followed by refolding as described in Experimental Procedures.

**FIGURE 7. Display of gp41 trimers on phage T4 nanoparticle.** A, Cryo-EM structure of phage T4 capsid (43,59). The structure of Soc trimer at quasi-3-fold axes is shown. The N and C termini are labeled, and the C-terminus of each subunit is shown as *red* dot. B, Binding of Soc-gp41M-Fd on phage T4. About  $2 \times 10^{10}$  *hoc<sup>-</sup>soc<sup>-</sup>* phage particles were incubated with increasing ratios of Soc-gp41M-Fd molecules to capsid binding sites (1:1 to 40:1, labeled at the *top*) and assembly was carried out as described in Experimental Procedures. *Lanes*: 1, control *hoc<sup>-</sup>soc<sup>-</sup>* phage; 2, 4, 6, 8, 10 and 12, phage displaying the bound fusion protein Soc-gp41M-Fd (*B*); 3, 5, 7, 9, 11 and 13, unbound protein in the supernatant (*U*). The position of the major capsid protein gp23\* is marked with *black arrow*. C, Binding of Soc-gp41ectoM-Fd on phage T4 at a Soc-fusion protein to capsid binding sites ratio of 20:1. The bound Soc-gp41ectoM-Fd protein was indicated with a *red arrow* (*lane 2*). D, Binding of CPP-Soc-gp41M-Fd on phage T4 at a Soc-fusion protein to capsid binding sites ratio of 20:1. Note that the 49 kDa CPP-Soc-gp41M-Fd protein migrates to the same position as the 48.7 kDa gp23\* (indicated with a *red arrow*). E, The saturation binding curve of Soc-gp41M-Fd. The density volumes of bound and unbound proteins from SDS-PAG (12%) were determined by laser densitometry and normalized to that of gp23\* present in the respective lane. The copy numbers were determined in reference to gp23\* (930 copies per capsid). The data were plotted as one site saturation ligand binding curve and fitted by non-linear regression using the SigmaPlot10.0 software and the calculated binding parameters are shown.  $K_d$ , apparent binding constant;  $B_{max}$ , maximum copy number per phage particle. F, The binding parameters of Soc and Soc-gp41 fusion recombinants. Since the CPP-Soc-gp41M-Fd band overlapped with the gp23\* band, gp23\* density was subtracted and the copy number was determined in reference to the tail sheath protein, gp18 (138 copies per phage; marked with a *black arrow* in panel D, *lane 2*).

**FIGURE 8. Recombinant designs leading to soluble and nanoparticle arrayed gp41 trimers.** A flow chart showing a series of approaches to generate soluble as well as phage T4 nanoparticle arrayed gp41 trimers. Schematic diagrams of soluble and displayed trimers are shown at the bottom. The trimers are stabilized in a prehairpin-like structure in which the HR1 helical grooves and MPER epitopes would be well exposed. Shown on the right is an enlarged cut-out of the capsid decorated with gp41 trimers. See Results and Experimental Procedures for additional details.

Figure 1

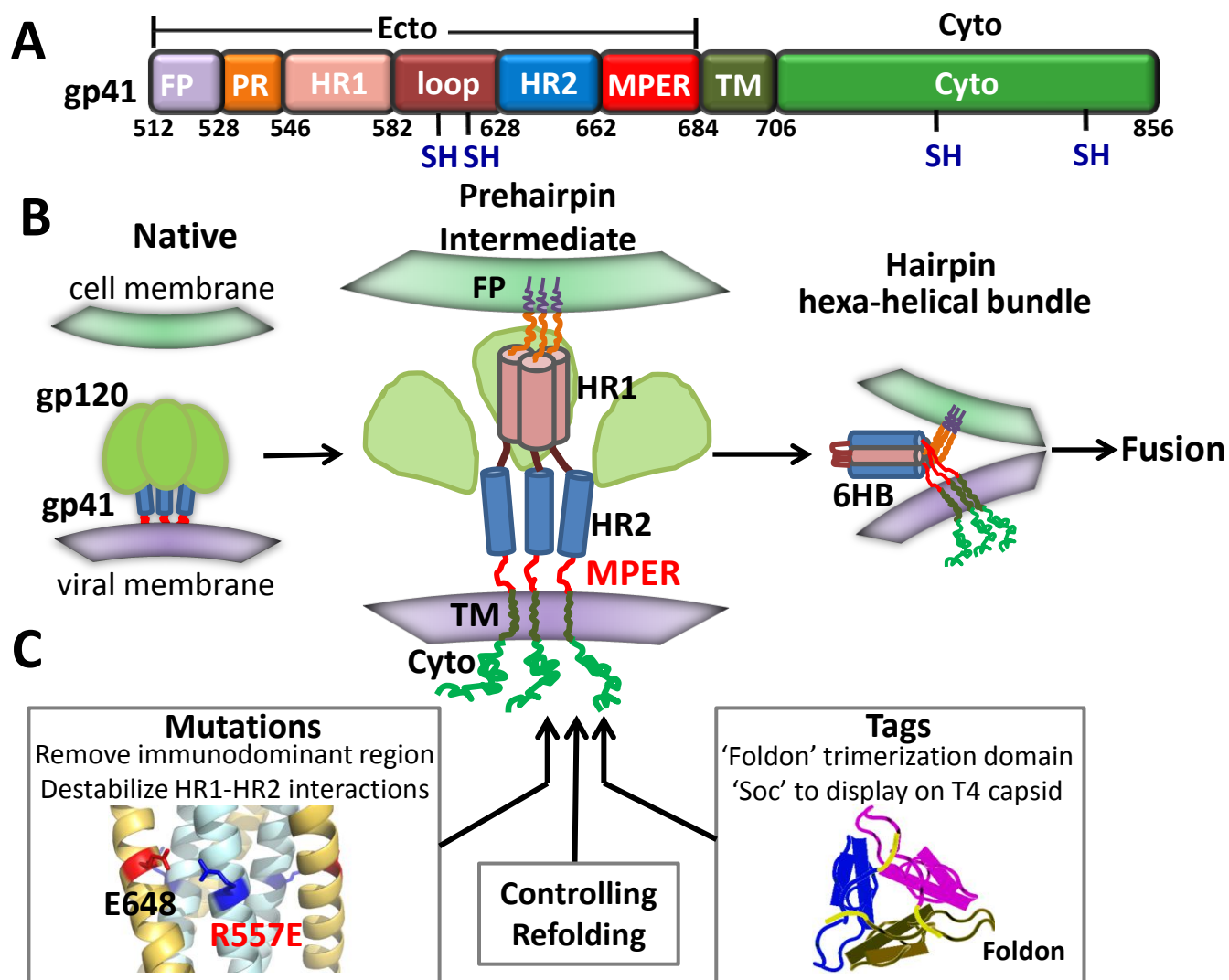


Figure 2

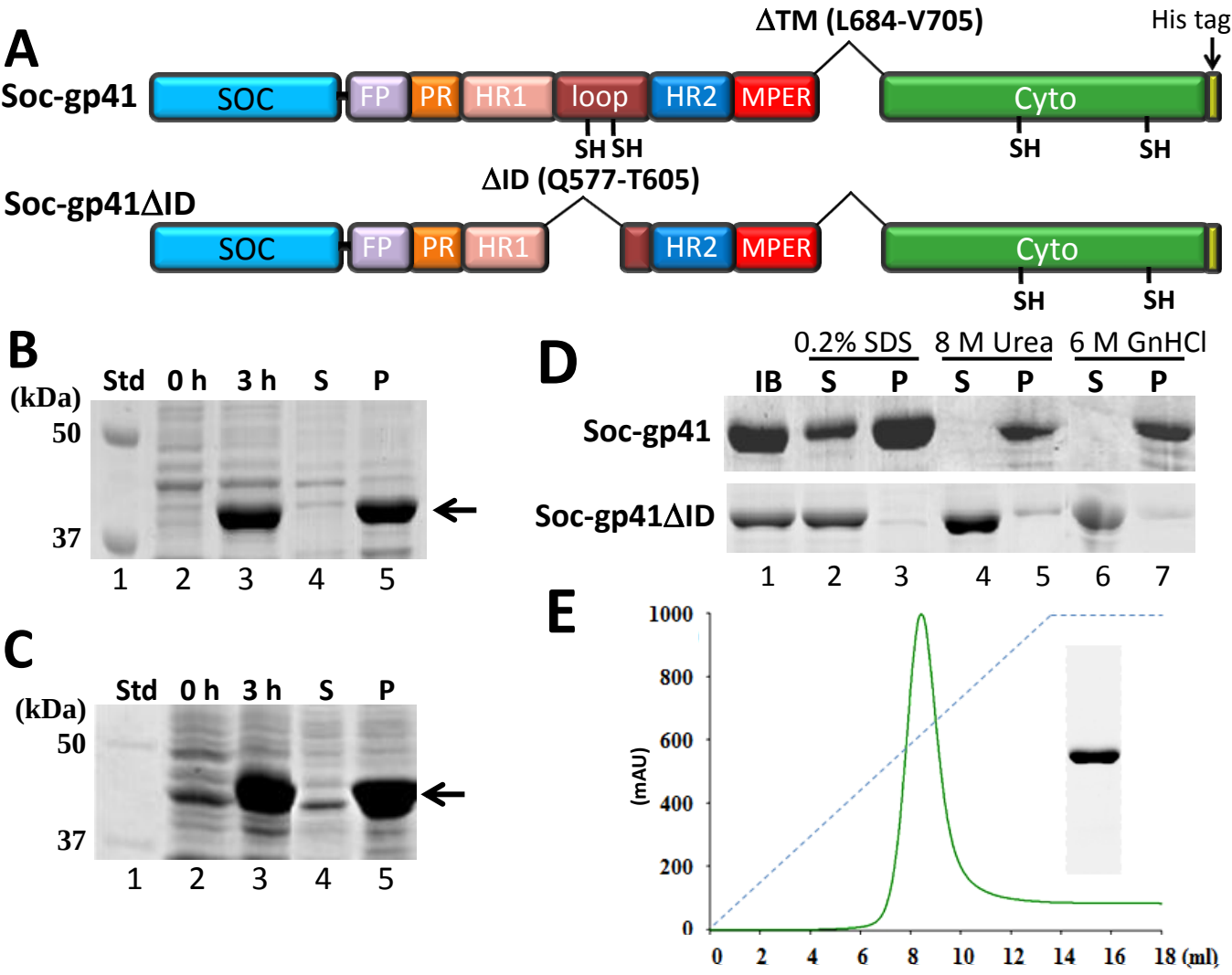


Figure 3

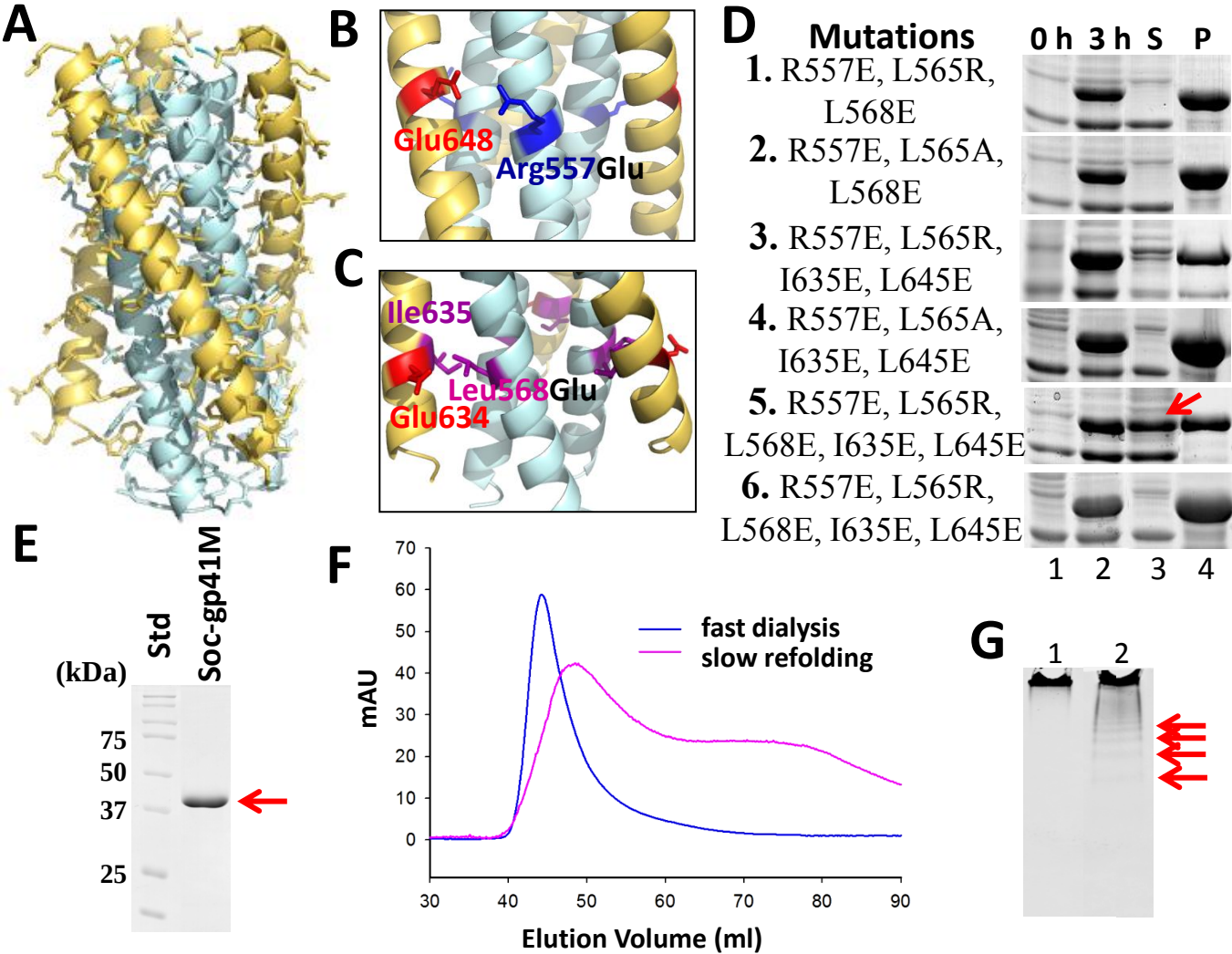


Figure 4

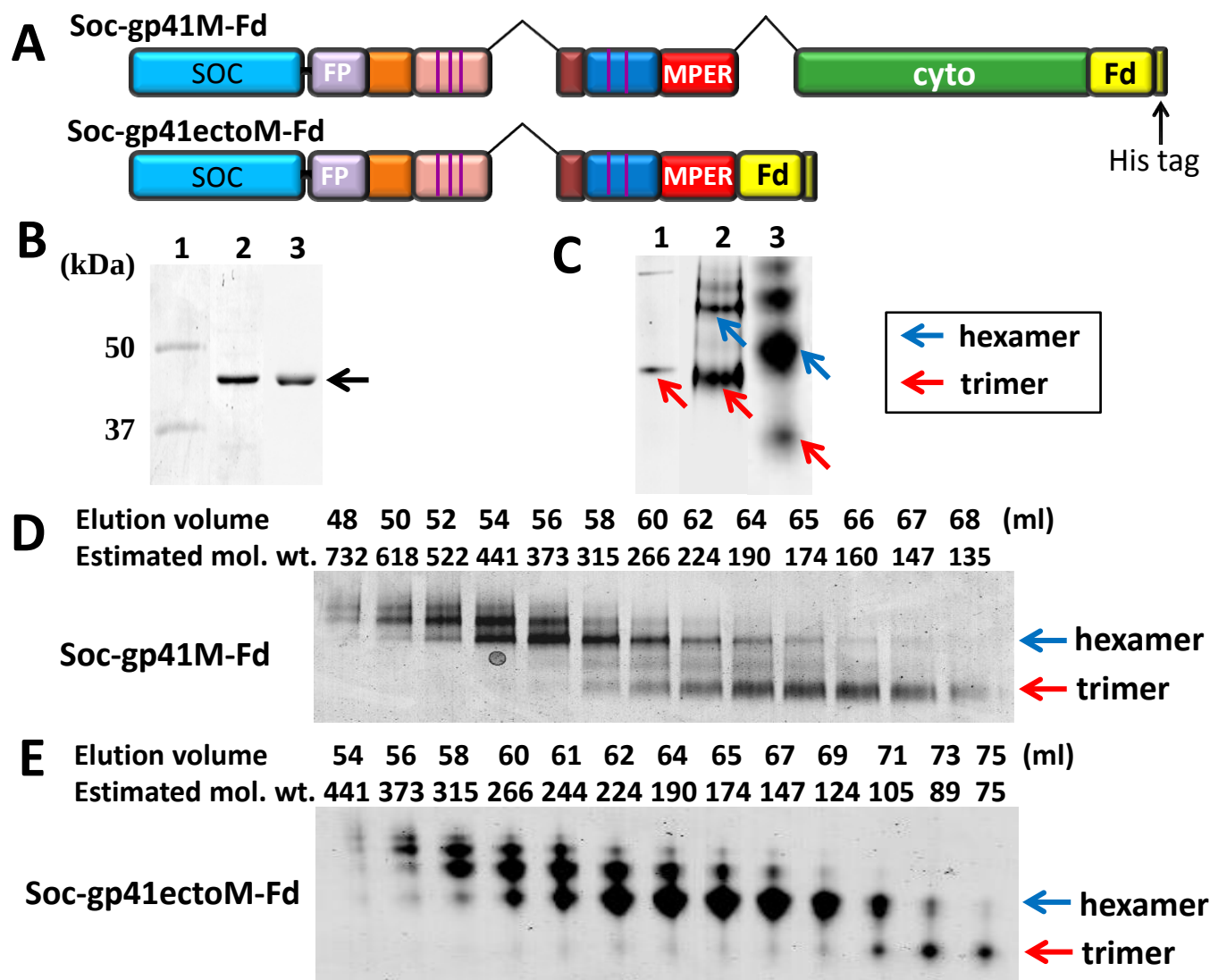


Figure 5

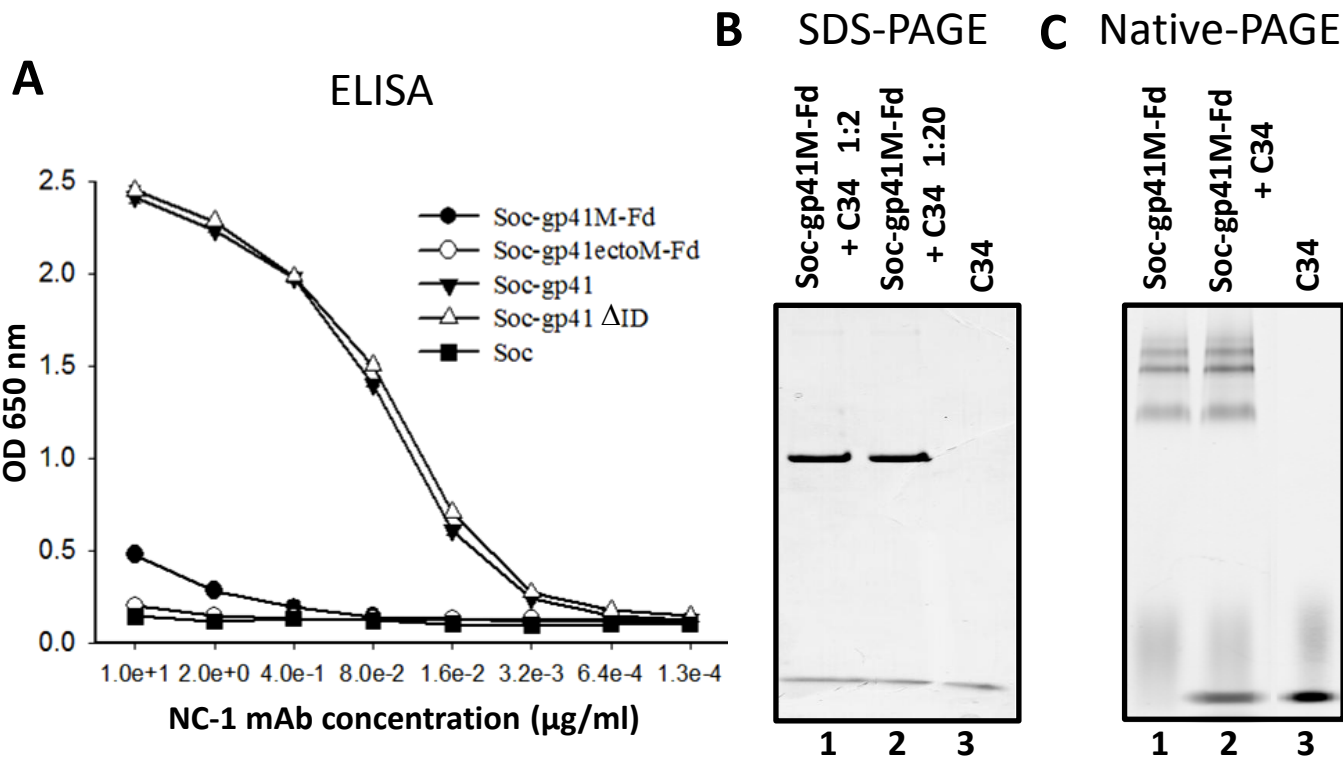


Figure 6

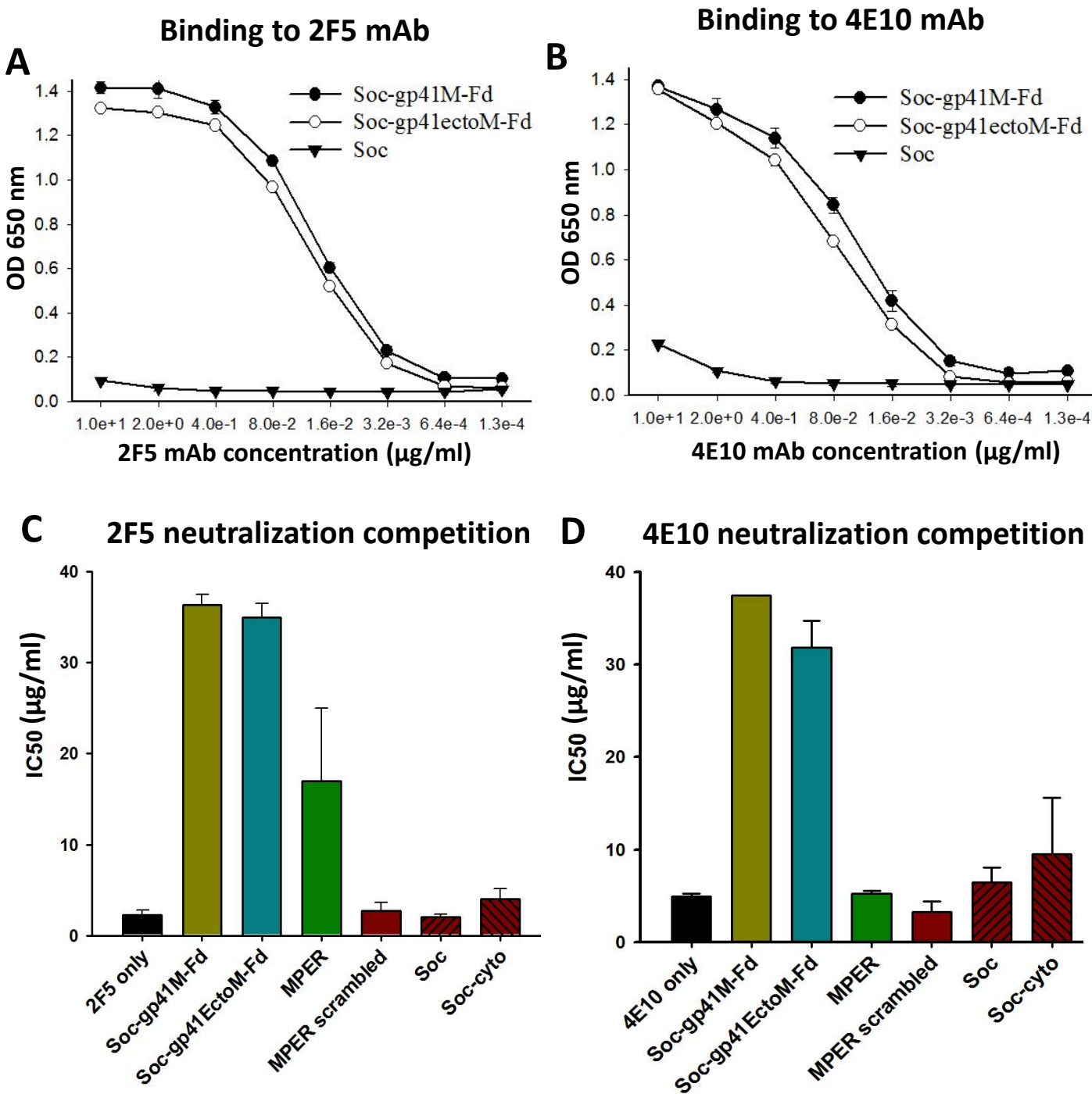




Figure 7

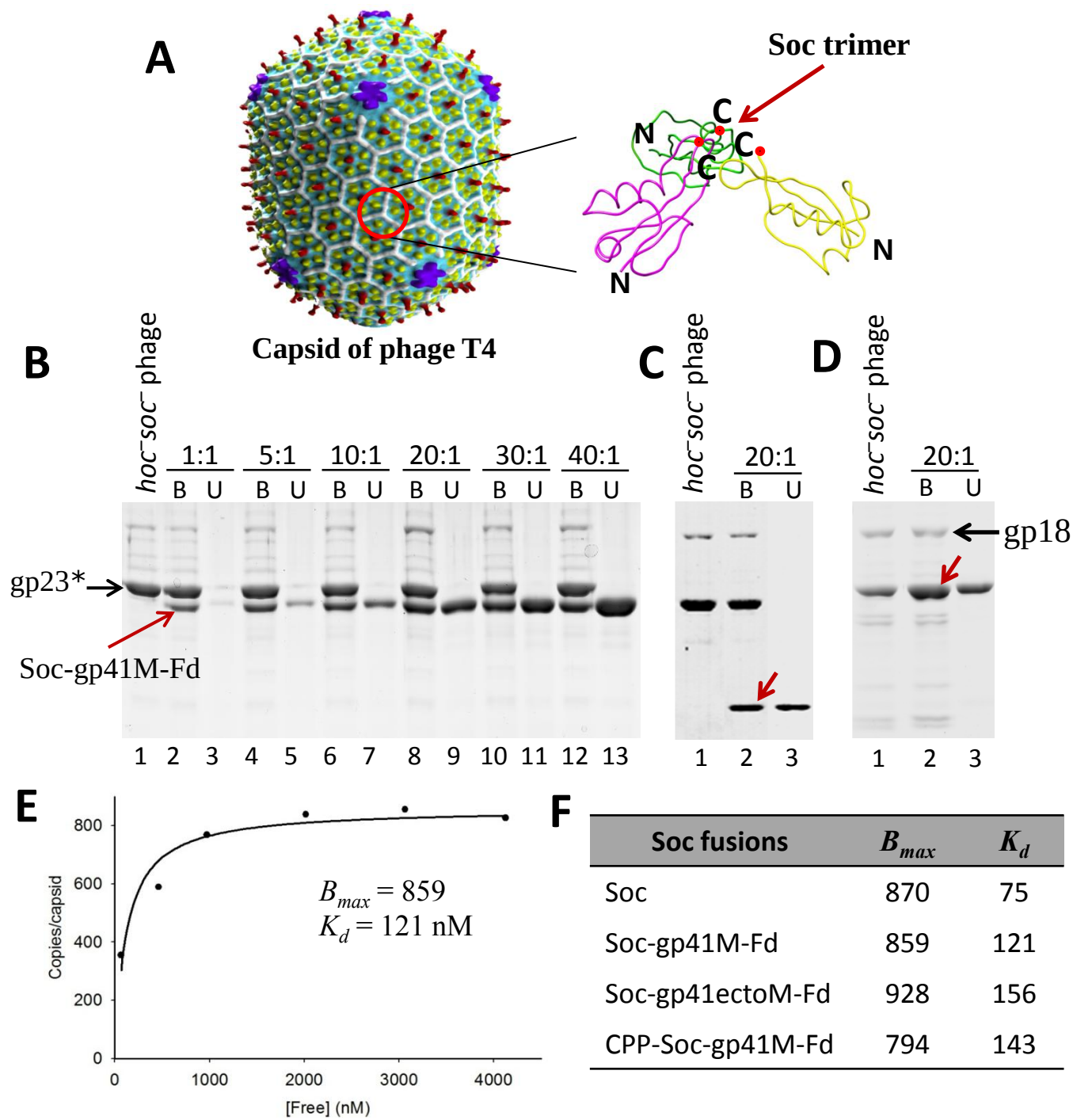


Figure 8

