Trimerization Specificity in HIV-1 gp41: Analysis with a GCN4 Leucine Zipper Model[†,‡]

Wei Shu, Hong Ji, and Min Lu*

Department of Biochemistry, Weill Medical College of Cornell University, New York, New York 10021

Received January 26, 1999; Revised Manuscript Received March 4, 1999

ABSTRACT: The envelope glycoprotein of human immunodeficiency virus type 1 (HIV-1) consists of a complex of two noncovalently associated subunits, gp120 and gp41. Formation of gp120/gp41 oligomers is thought to be dependent on a 4-3 hydrophobic (heptad) repeat located in the amino-terminal region of the gp41 molecule. We have investigated the role of this heptad repeat in determining the oligomeric structure of gp41 by introducing its buried core residues into the first (***a***) and fourth (***d***) positions of the GCN4 leucine-zipper dimerization domain. The mutant peptides fold into trimeric, helical structures, as shown by circular dichroism and equilibrium sedimentation centrifugation. The 2.4 Å resolution crystal structure of one such trimer reveals a parallel three-stranded, α-helical coiled coil. Thus, the buried core residues from the gp41 heptad repeat direct trimer formation. We suggest that the conserved amino-terminal heptad repeat within the gp41 ectodomain possesses trimerization specificity.

The human immunodeficiency virus type 1 (HIV-1)[1] enters cells by the envelope protein-mediated membrane-fusion process that allows invasion of the infectious viral genome. The envelope protein consists of the surface subunit gp120 and the transmembrane subunit gp41 (*1*). These two subunits are noncovalently associated and oligomerize on the surface of the virion. gp120 determines viral tropism through a sequential interaction with the cellular receptors, CD4, and members of the chemokine receptor family at the T cell surface (*2, 3*). These interactions trigger the membrane-fusion activity of gp41 and thus initiate membrane fusion. A number of studies support the notion that, upon binding of gp120 to target cell receptors, gp41 undergoes conformational changes from a native (nonfusogenic) to a fusion-active (fusogenic) state (see refs *4* and *5*). This conformational transition likely facilitates exposure of the hydrophobic, glycine-rich sequence referred to as the fusion peptide at the amino terminus of gp41, leading to insertion into the lipid bilayer of the target membrane (*6*−*9*).

The extracellular region (ectodomain) of the gp41 molecule mediates association with gp120 and is directly involved in the membrane-fusion reaction (*10*). The amino-terminal fusion peptide of gp41 is followed by two regions with a 4-3 hydrophobic (heptad) repeat predicted to form coiled-coil structures (*11*−*13*) (Figure 1). Between these heptad-repeat regions is a disulfide-bonded loop region. The N-terminal heptad repeat is located adjacent to the fusion peptide, whereas the C-terminal heptad repeat precedes the transmembrane segment (Figure 1). Protein-dissection studies showed that these two heptad-repeat regions form a stable helical trimer of antiparallel dimers (*14*−*16*). X-ray crystallographic analyses confirmed that this gp41 core is a six-helix bundle (*17*−*19*). Three N-terminal helices form an interior, parallel trimeric coiled coil, while three C-terminal helices pack in an antiparallel orientation into highly conserved hydrophobic grooves on the surface of this trimeric coiled coil. The structural similarities between the gp41 core and the fusion-active hemagglutinin protein (HA) of influenza virus suggest that the six-helix bundle represents the fusion-active conformation of gp41 (*17*−*19*). This conclusion has also been drawn from antibody binding studies which show that a monoclonal antibody specifically recognizing the six-helix bundle binds to the surface of HIV-1-infected cells only after interaction of the envelope protein complex with soluble CD4 (*20*).

The N-terminal heptad-repeat sequence of gp41 is one of the most highly conserved regions in the HIV-1 envelope protein. It appears to determine the trimerization state of gp41 in its fusion-active conformation (*14*−*19*). On the other hand, an isolated peptide corresponding to this coiled-coil sequence motif has been reported to form an α-helical tetrameric structure (*21*−*24*). In addition, genetic studies indicate that the putative N-terminal coiled-coil domain of gp41 is required for membrane-fusion activity but is not necessary for the oligomerization of the native gp120/gp41 complex, as found on the surface of the virion (*25*−*28*). It has therefore been postulated that the receptor-activated conformational change in the HIV-1 envelope protein complex, as in the influenza hemagglutinin (*29, 30*), is driven by formation of an α-helical coiled coil (*28, 31*). Therefore, it is of interest

[†] This work was supported by National Institutes of Health Grant AI42382.

[‡] The atomic coordinates have been deposited in the Protein Data Bank, Biology Department, Brookhaven National Laboratory, Upton, NY 11973 (PDB file name 1ce0).

* To whom correspondence should be addressed: Department of Biochemistry, Weill Medical College of Cornell University, 1300 York Ave., New York, NY 10021. Phone: (212) 746-6562. Fax: (212) 746-8875. E-mail: mlu@mail.med.cornell.edu.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; SIV, simian immunodeficiency virus; $[\theta]_{222}$, molar ellipticity at 222 nm; CD, circular dichroism; HPLC, high-performance liquid chromatography; LB, Luria-Bertani; T_m, temperature at the midpoint of thermal denaturation; CNBr, cyanogen bromide; GdmCl, guanidinium hydrochloride; IPTG, isopropyl thio-β-galactoside; PBS, neutral pH phosphate-buffered saline; rms, root-mean-square.

10.1021/bi990199w CCC: $18.00
Published on Web 04/09/1999

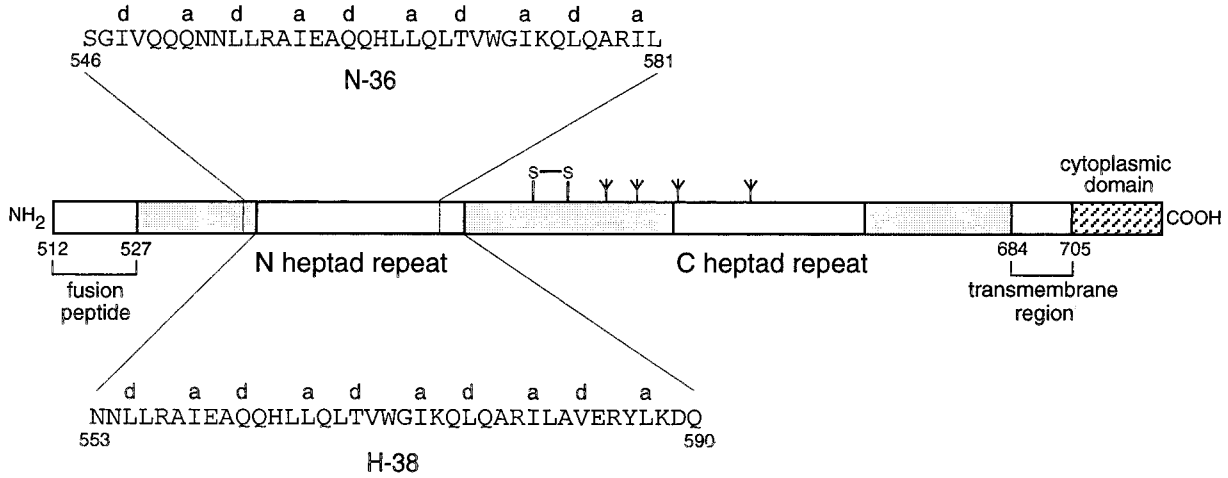

FIGURE 1: α-Helical coiled-coil motif within the N-terminal portion of the gp41 ectodomain. Schematic diagram of gp41 showing important functional features. The locations of the N- and C-terminal hydrophobic heptad-repeat regions are denoted. The amino acid sequences of the H-38 and N-36 segments are shown, along with the ***a*** and ***d*** positions above the sequences. The disulfide bond and four potential N-glycosylation sites are depicted. The residues are numbered according to their position in gp160.

and importance to understand the oligomeric specificity of the N-terminal coiled-coil motif unequivocally.

Coiled-coil motifs have been shown to be important for mediating protein oligomerization (*32*). These motifs share a characteristic seven-amino acid (heptad) repeat, $(\boldsymbol{abcdefg})_n$, with hydrophobic residues at the first (***a***) and fourth positions (***d***) and polar residues generally elsewhere (*33*−*35*). The coiled-coil sequences with this shared heptad-repeat pattern can adopt parallel or antiparallel dimeric as well as trimeric and tetrameric conformations (*32*, *35*−*47*). A 33-amino acid peptide corresponding to the leucine-zipper dimerization domain from the yeast transcription factor, GCN4, forms a parallel, dimeric coiled coil (*36*). Previous studies demonstrated that core mutants of this peptide (that is, substituted only at positions ***a*** and ***d***) can form dimers, trimers, and tetramers (*37*, *38*). Thus, the buried "core" residues in the GCN4 leucine zipper are sufficient to define the oligomerization state of the molecule. In addition, some combinations of core residues lack structural specificity, and can populate more than one oligomeric state (*37*, *39*−*47*). Thus, the stringently defined oligomeric properties of the GCN4 leucine zipper provide a sensitive test-bed for defining the determinants of structural specificity in this class of molecules.

Here we use the GCN4 leucine zipper as a model system to investigate the role of the buried core residues within the N-terminal heptad-repeat region of gp41 in determining oligomeric specificity. We have introduced these core residues into their equivalent ***a*** and ***d*** positions of the GCN4-p1 peptide. Physicochemical characterization shows that the leucine zipper-gp41 hybrid peptides form trimeric helical bundles and that they exhibit a cooperative thermal unfolding transition. The X-ray crystal structure of one such helical bundle at 2.4 Å resolution is a parallel three-stranded, α-helical coiled coil. Thus, the conserved core residues within the N-terminal heptad-repeat region of gp41 direct trimer formation in isolation. We suggest that the N-terminal coiled-coil sequence of gp41 possesses exclusively trimerization specificity.

MATERIALS AND METHODS

Plasmid Construction. Plasmids for the expression of chimeric proteins were used to generate peptides H38-p1 and N36-p1 upon cleavage by cyanogen bromide (CNBr). Synthetic genes encoding H38-p1 and N36-p1 were prepared with optimal codon usage for *Escherichia coli* and verified by DNA sequencing. The genes were subsequently cloned into the *Hin*dIII−*Bam*HI sites of a phagemid-T7 expression vector, pTMHa (*48*), to produce plasmids pN36-p1 and pH38-p1. In the pTMHa vector, the desired sequences are expressed as chimeric proteins containing a modified form of the TrpLE leader sequence, in which an NH_2-terminal $(His)_9$ tag has been added. Standard recombinant DNA techniques were used (*49*).

Protein Production and Purification. Chimeric proteins H38-p1 and N36-p1 were expressed in *E. coli* BL21(DE3) pLysS using the T7 expression system as previously described (*14*). Cells, freshly transformed with an appropriate plasmid, were grown to late log phase. Protein expression was induced by addition of 0.5 mM isopropyl thio-β-D-galactoside (IPTG). After growth for another 3 h at 37 °C, the bacteria were harvested by centrifugation and the pellets were stored frozen. Inclusion bodies of the cell lysate were isolated as described previously (*14*). The inclusion bodies were solubilized in 6 M guanidinium hydrochloride (GdmCl), 0.1 M sodium phosphate, and 10 mM Tris (pH 8.0). The solution was passed over a nickel-chelating column (Ni^{2+}−NTA−agarose, Qiagen). Bound fractions containing the His tag were eluted with a 6 M GdmCl buffer containing 0.2 M acetic acid. After dialysis and lyophilization, proteins were treated with CNBr (0.03 g/mL, 70% formic acid, 2 h). After removal of CNBr, the proteins were dialyzed against 5% acetic acid and lyophilized. The resulting pellet was solubilized in 6 M GdmCl, 0.1 M sodium phosphate, and 10 mM Tris (pH 8.0). The solution was again passed over a nickel-chelating column to remove the leader sequence containing the His tag and any uncleaved chimeric protein. Proteins were purified to homogeneity by reverse-phase HPLC, using a Vydac C-18 preparative column and a linear gradient of acetonitrile containing 0.1% trifluoroacetic acid. The protein identity was confirmed by mass spectrometry. Protein concentrations were determined by tyrosine absorbance at 280 nm in 6 M GdmCl (*50*).

Circular Dichroism (CD) Spectroscopy. CD spectra were acquired at a peptide concentration of 200 μM in neutral

pH phosphate-buffered saline (PBS) [50 mM sodium phosphate (pH 7.0) and 150 mM NaCl] with an Aviv 62 DS spectrometer as described previously (*14*). The wavelength dependence of molar ellipticity, $[\theta]$, was monitored at 0 or 20 °C as the average of five scans, using a 5 s integration time at 1.0 nm wavelength increments. Spectra were baseline-corrected against the cuvette with buffer alone. The helix content was estimated from the CD signal by dividing the mean residue ellipticity at 222 nm by the value expected for 100% helix formation by helices of comparable size [−33000 deg cm^2 dmol^{-1} (*52*)]. Thermal stability was determined by monitoring the change in CD signal at 222 nm as a function of temperature, and thermal melts were performed in 2 °C intervals with a 2 min equilibration at the desired temperature, and an integration time of 30 s. Reversibility was checked by repeated scans. All melts were reversible, with superimposable folding and unfolding curves and >90% of the signal regained upon cooling. The midpoint of the thermal unfolding transition (T_m) was determined from the maximum of the first derivative, with respect to the reciprocal of the temperature, of the $[\theta]_{222}$ values (*51*). The error in the estimation of T_m is ±1 °C.

Sedimentation Equilibrium. Sedimentation equilibrium studies were performed on a Beckman XL-A analytical ultracentrifuge as described previously (*14*). Protein solutions were dialyzed overnight against PBS, loaded at initial concentrations of 200, 400, and 1000 μM, and analyzed at rotor speeds of 30 000 and 33 000 rpm. Data sets (six per protein) were fitted simultaneously to a single-species model with the program NONLIN (*53*). The protein partial specific volume and solvent density were calculated with constants from Laue et al. (*54*). Molecular masses were all within 10% of those calculated for an ideal trimer, with no systematic deviation of the residuals.

Crystallization, Data Collection, and Structural Determination. The H38-p1 peptide was crystallized by hanging-drop vapor diffusion at room temperature. Initial crystallization conditions were screened by using sparse matrix crystallization kits (Crystal Screen I and II, Hampton Research, Riverside, CA) and then optimized. To grow crystals, a 7 mg/mL stock of the H38-p1 peptide was diluted 1:1 with 0.1 M Hepes (pH 7.5), 1.5 M Li_2SO_4, and 20% (v/v) glycerol (reservoir) and allowed to equilibrate against the reservoir solution. Primitive hexagonal crystals were transferred to a cryoprotectant solution containing 25% (v/v) glycerol in the corresponding mother liquor. Cryoprotected crystals were frozen in propane before data collection. Data to 2.4 Å resolution were collected at 95 K using a Mar research 300 image plate scanner at the X12B beamline of the National Synchrotron Light Source. Data were integrated and scaled with the HKL suite (*55*).

The structure of the H38-p1 trimer was determined by molecular replacement using AMoRe (*56*). The 1.8 Å structure of the isoleucine-zipper trimer GCN4-pII (PDB code 1GCM, all side chains truncated to alanine) was used in a combined rotation−translation search (with 8.0−3.5 Å data) to yield a solution for H38-p1 (correlation coefficiency = 66.5%; R factor = 43.0%). The model was refined with simulated annealing and atomic displacement parameter (ADP) refinements using the program X-PLOR (*57*). The model was rebuilt to reflect the sequence of H38-p1 by using conventional $(2F_o - F_c)\Phi_{calc}$ and $(F_o - F_c)\Phi_{calc}$ maps with the program O (*59*) and improved with the program ARP (*58*). The final model contains residues 1−35 (monomers A and C), 1−36 (monomer B), and 79 water molecules. Gln4 (monomers B and C) is modeled as alanine. The last two residues of monomers A and C and the last residue of monomer B were left out of the model because of the absence of interpretable electron density for these atoms.

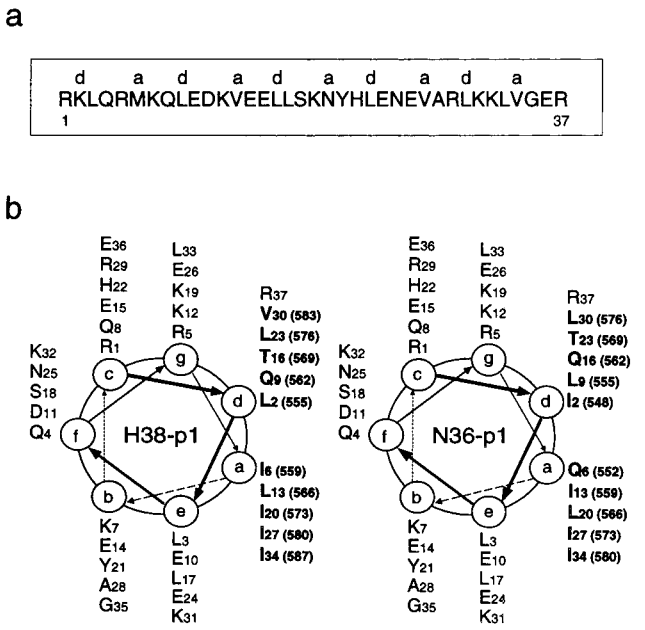

FIGURE 2: Chimeras of gp41 and the GCN4 dimerization domain. (a) Sequence of the GCN4 leucine zipper (amino acids 245−281). The ***a*** and ***d*** positions that form the core of the dimer interface are shown. (b) Helical wheel representation of the H38-p1 (left) and N36-p1 (right) peptides. Residues in positions ***a*** and ***d*** of the GCN4 leucine zipper are replaced with their equivalents from the sequences of H-38 and N-36, respectively. Residues from the gp41 heptad repeat are denoted in bold type; their positions in gp160 are denoted in parentheses. The view is from the NH_2 terminus looking toward the COOH terminus.

RESULTS AND DISCUSSION

Core Mutants of the GCN4 Leucine Zipper. The N-terminal coiled-coil segment of HIV-1 gp41 corresponds to residues 553−590 (Figure 1), named H-38 for its length in residues. The N-36 peptide (residues 546−581) identified by protein dissection differs slightly from the predicted H-38 motif (Figure 1). Unlike traditional coiled-coil motifs, H-38 and N-36 also contain a preponderance of hydrophobic residues at positions ***e*** and ***g*** (Figure 1). Having found that the isolated N-36 peptide has a tendency to aggregate (*15*), we surmised that the residues at positions ***a*** and ***d*** of the peptide are responsible for oligomer formation, whereas the hydrophobic residues at positions ***e*** and ***g*** cause aggregation.

We assume that if the buried core residues in the N-terminal heptad-repeat region of gp41 can specify a unique coiled-coil structure (e.g., dimer, trimer, or tetramer), formation of favorable packing interactions at the helix interface involving these residues should contribute to the global stability and oligomeric specificity of the GCN4 leucine zipper (see the introductory section). To test this hypothesis, we generated two core mutants of the GCN4 leucine zipper by replacing the residues at positions ***a*** and ***d*** of the molecule with their equivalents from H-38 and N-36, respectively (Figure 2b). Because H-38 and N-36 have five heptad repeats

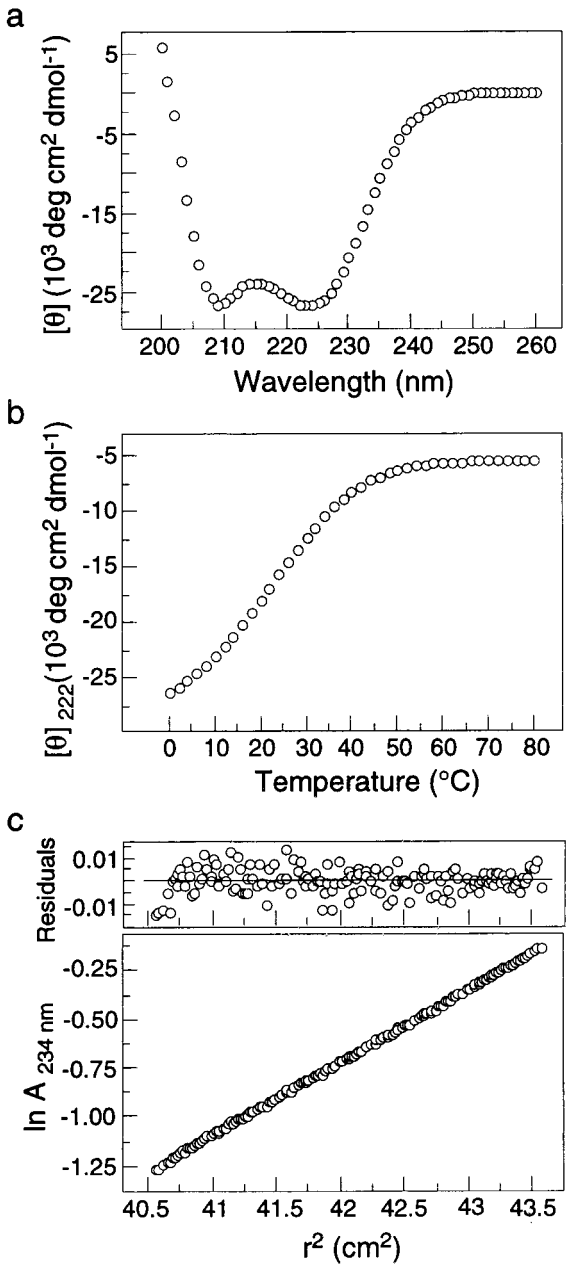

FIGURE 3: Folding of N36-p1 as an α-helical trimer. (a) Circular dichroism (CD) spectrum of the N36-p1 peptide in PBS (pH 7.0) and at a peptide concentration of 200 μM at 0 °C. (b) Thermal melt of N36-p1 monitored by CD at 222 nm. (c) Analytical ultracentrifugation data (27 000 rpm) collected at 4 °C in PBS (pH 7.0) at at a peptide concentration of ~400 μM. The natural logarithm of the absorbance at 234 nm is plotted against the square of the radial position. Deviations from the calculated values are plotted as residuals (upper part).

each, the desired length was achieved by extending four residues at the N terminus of the GCN4-p1 peptide (Figure 2a) (*60*, *61*). These hybrid peptides were named H38-p1 and N36-p1, respectively (Figure 2b). They were produced by bacterial expression and purified by reverse-phase HPLC (see Materials and Methods).

Core Mutants Form Trimeric Helical Structures. The circular dichroism (CD) spectrum of the N36-p1 peptide is typical of an α-helix, displaying the characteristic minima at 208 and 222 nm (Figure 3a). On the basis of the magnitude of the CD signal at 222 nm in PBS (pH 7.0) and the 200 μM peptide concentration at 0 °C, N36-p1 appears to be ~80% helical. Incomplete helix formation is possibly due to fraying at the ends of the helices. Under these conditions, N36-p1 exhibits a cooperative, thermally induced unfolding transition, with a melting temperature (T_m) of 24 °C (Figure 3b). Sedimentation equilibrium measurements were used to determine the oligomerization states of the N36-p1 peptide. Over a 5-fold range of peptide concentrations (200 μM to 1 mM), the average experimental molecular mass of N36-p1 is 12.6 kDa (Figure 3c), indicating that the peptide is a trimer. Taken together, these results indicate that the N36-p1 peptide forms a three-stranded helical bundle in solution.

The H38-p1 peptide is not soluble in PBS (pH 7.0) but becomes soluble up to ~0.5 mM in PBS in the presence of the denaturant guanidinium hydrochloride (GdmCl) at a concentration of 0.7 M. CD experiments indicate that H38-p1 is ~80% helical at 20 °C in PBS (pH 7.0) containing 0.5 M GdmCl at a peptide concentration of 200 μM (Figure 4a). Under these conditions, H38-p1 melts cooperatively and reversibly with a T_m of 58 °C (Figure 4b). Sedimentation equilibrium measurements indicate that the H38-p1 peptide sediments as a trimer and exhibits no systematic dependence of molecular mass on peptide concentration between 200 μM and 1 mM in PBS in the presence of 1 M GdmCl (Figure 4c).

Although the H38-p1 peptide is not soluble under physiological conditions (i.e., PBS buffer), it does retain a cooperative two-state folding−unfolding transition and form a clean trimer in the presence of low concentrations of the denaturant GdmCl, demonstrating a defined trimeric, helical structure. In addition, the sequences of the H38-p1 and N36-p1 peptides differ in several buried core residues (Figure 1), which may give rise to their large difference in melting temperature. For example, the buried polar residue Gln522 may lead to significantly reduced thermal stability of N36-p1 (see ref *46*). Taken together, these results indicate that the conserved core residues in the N-terminal coiled-coil motif of HIV-1 gp41 confer structural specificity for trimer and not tetramer formation. Conclusions regarding the tetrameric form of gp41 based on studies of the H-38 peptide and the chimeric protein are probably misleading because the hydrophobic residues at positions ***e*** and ***g*** cause aggregation upon exposure to solvent.

Crystal Structure of the H38-p1 Trimer. To better understand how the buried core residues from the N-terminal coiled-coil motif of gp41 determine trimer formation, the X-ray crystal structure of the H38-p1 trimer was determined. The H38-p1 peptide was crystallized by hanging-drop vapor diffusion. Primitive hexagonal crystals were obtained from 0.1 M Hepes (pH 7.5), 1.5 M Li_2SO_4, and 20% (v/v) glycerol. The crystals belong to space group *P*61 ($a = b = 50.79$ Å, $c = 119.25$ Å; Table 1), with three monomers in a crystallographic asymmetric unit. The crystal structure of H38-p1 was determined by molecular replacement using a 1.8 Å structure of GCN4-pII (*37*) and refined at 2.4 Å resolution, with crystallographic and free *R* factors of 22.1% and 29.4, respectively, in the resolution range of 8.0−2.4 Å (Table 1). The current model contains residues 1−35 (monomers A and C) and 1−36 (monomer B) as well as 79 water molecules. The root-mean-square (rms) deviations of bond lengths and bond angles from the ideal values are 0.006 Å and 1.303°, respectively (Table 1). All ϕ and ψ angles are in allowed regions of the Ramachandran plot.

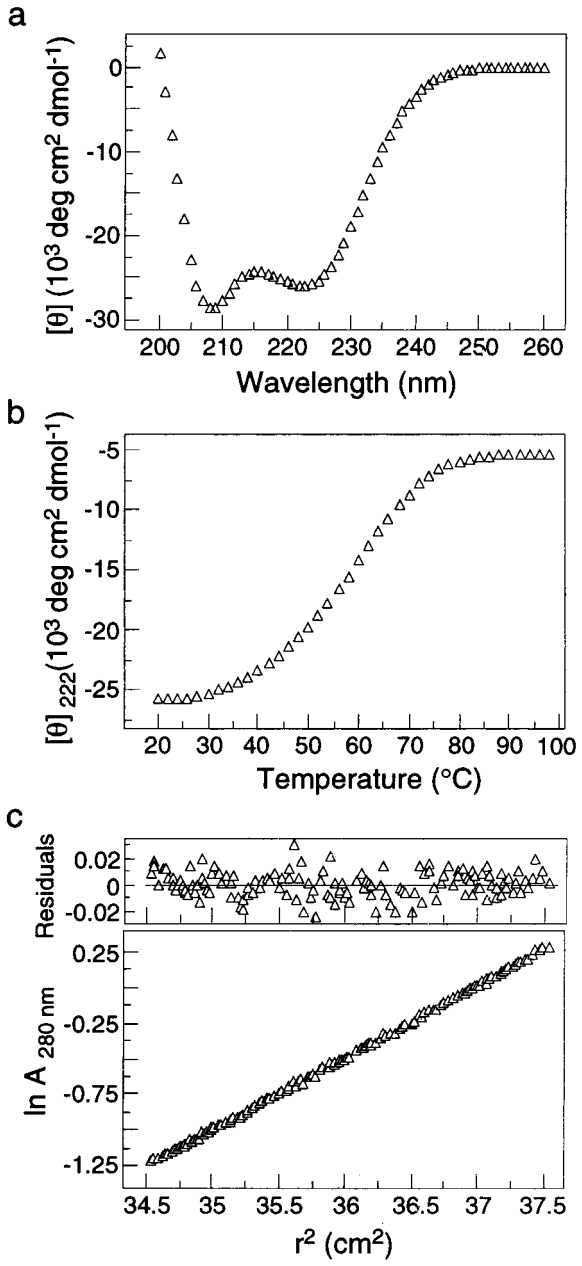

FIGURE 4: H38-p1 forms a helical trimer. (a) CD spectrum of the H38-p1 peptide at 20 °C in PBS (pH 7.0) and at a peptide concentration of 200 μM in the presence of 0.5 M GmdCl, a denaturant. (b) Thermal melt of H38-p1 monitored by CD at 222 nm. (c) Analytical ultracentrifugation data (30 000 rpm) collected at 20 °C in PBS (pH 7.0) at a peptide concentration of ~1 mM in the presence of 1 M GdmCl. The natural logarithm of the absorbance at 280 nm is plotted against the square of the radial position. Deviations from the calculated values are plotted as residuals (upper part).

The H38-p1 structure is a regular trimeric coiled coil, consisting of three parallel α-helices wrapped in a gradual left-handed superhelix (Figure 5). The superhelix creates a cylinder that is ~51 Å long and ~25 Å wide. The α-helices are packed together in the classical acute "knobs-into-holes" arrangement characteristic of trimeric coiled coils, in which the $C_\alpha-C_\beta$ bonds in the side chains (knobs) of the ***a*** and ***d*** residues make an acute angle with respect to the recipient holes formed by spaces between four residues on a neighboring helix (*37*, *62*). All the core leucine and valine residues point into the center of the trimer and adopt preferred rotamers (*63*). Thus, the buried core residues in the N-terminal heptad-repeat region of HIV-1 gp41 are capable of imparting orientational and trimeric specificity to the leucine-zipper coiled coil.

Table 1: X-ray Data Collection and Refinement Statistics

space group	*P*61
a, *b*, *c* (Å)	50.79, 50.79, 119.25
resolution (Å)	30.0−2.4
no. of measured reflections	22736
no. of unique reflections	6765
completeness (%)	98.8
R_{merge}[a] (%)	3.9
Refinement	
resolution (Å)	8.0−2.4
no. of reflections in working set (2σ)	6059
no. of reflections in test set	500
no. of protein non-hydrogen atoms	895
no. of water molecules	79
R_{free}[b] (%)	29.4
R_{cryst}[b] (%)	22.1
average *B*-factor (Å²)	37.5
rms deviations from ideality	
bond lengths (Å)	0.006
bond angles (deg)	1.303
torsion angles (deg)	21.9

[a] $R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the intensity of an individual measurement and $\langle I\rangle$ is the mean recorded intensity over multiple recordings. [b] $R = \Sigma|F_o - F_c|/\Sigma F_{obs}$, where the crystallographic and free R factors are calculated using the working and free reflection sets, respectively.

The coiled-coil packing geometry in the H38-p1 trimer is very similar to that seen in the six-helix bundle of the gp41 core [N34(L6)C28; *19*] (Figure 6). The rms deviation between all C_α atoms of the homotrimeric packing in H38-p1 and N34(L6)C28 is 0.44 Å. In H38-p1, there are 10 layers of homotrimeric interactions in which the core residues at the successive ***a*** and ***d*** positions pack together on the molecular 3-fold symmetry axis. Although these homotrimeric contacts are largely hydrophobic, two polar residues are buried in the interior core packing. The Gln9 side chains, at the ***d*** position, face each other across the 3-fold axis and form a network of hydrogen bonds between the side chain amide group and the backbone carbonyl on the opposite helix (3.15 Å). As a consequence, the carbonyl groups of the Gln9 side chain are exposed to the solvent. In the layer around Thr16 at the ***d*** position, the residue has its hydroxyl group hydrogen-bonded to the backbone carbonyl of Lys12 (2.9 Å) and still uses its hydrophobic methyl group pointing toward the center of the trimer (Figure 5a). These polar residues are conserved in HIV-1 and simian immunodeficiency virus (SIV) and appear to be critical in maintaining the trimeric coiled-coil structure.

Oligomeric Structure of gp41. Attempts to define the oligomerization state of the HIV-1 envelope protein complex have yielded conflicting results. Chemical cross-linking, velocity sedimentation, and electron microscopy studies suggest variously dimeric, trimeric, and tetrameric models for the native envelope protein complex (for a review, see ref *10*; see, however, ref *64*). The highly conserved N-terminal coiled-coil motif of gp41 has been predicted to play a role in oligomer formation and stability (*11*−*13*). Mutagenesis studies have demonstrated that substituting proline for isoleucine or leucine within the N-terminal heptad-repeat region abolishes membrane-fusion activity, but these sub-

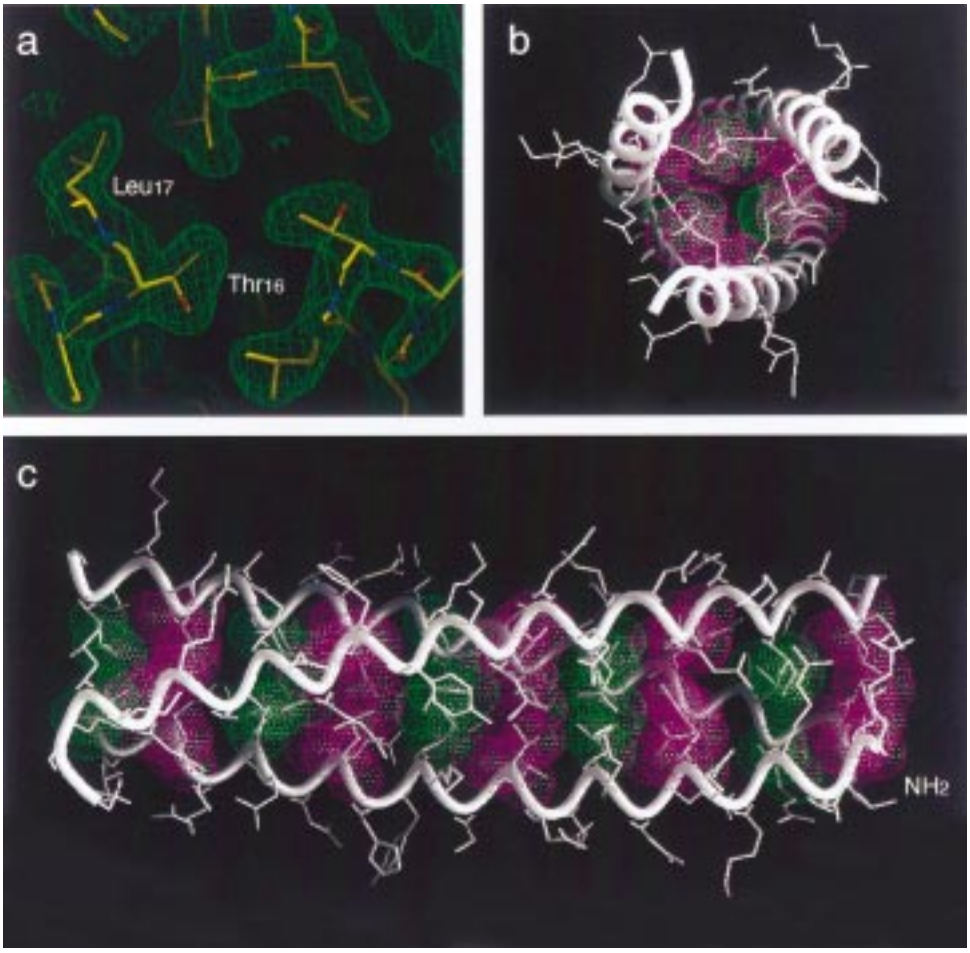

FIGURE 5: Crystal structure of H38-p1. (a) Final $2F_o - F_c$ electron-density map of the H38-p1 trimer with the refined model superimposed. The map is contoured at 1.4σ. This figure was created with the program O (*59*). (b) Axial view of the H38-p1 trimer from the NH_2 terminus looking down the 3-fold axis of the trimer. The van der Waals surfaces are colored purple for residues at the ***a*** positions and green for residues at the ***d*** positions. (c) Side view of the trimer showing the van der Waals surfaces of residues at the ***a*** (purple) and ***d*** (green) positions superimposed on the helix backbone. The NH_2 terminus is labeled. Panels b and c were generated with the program GRASP (*65*).

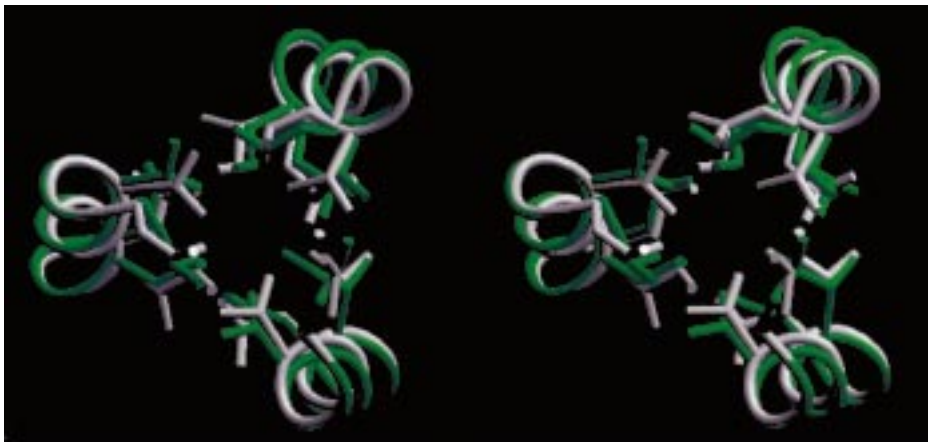

FIGURE 6: Stereoview of the superimposed backbone ribbons of the N-terminal trimeric coiled coil in H38-p1 (green) and in N34(L6)C28 (white), and overlayed using helices from H38-p1 (residues 13−23) and N34(L6)C28 (residues 21−31). This figure was generated with the program SETOR (*66*).

stitutions have little or no effect on the oligomeric structure of the HIV-1 envelope protein complex on the surface of the virion (*25*−*28*). This finding suggests that the N-terminal coiled-coil motif must be "sequestered" in the native structure and that the gp41 conformational change is driven by formation of the N-terminal coiled-coil structure (*28*, *31*). An understanding of the factors that determine oligomerization specificity in gp41 has ramifications that extend into more general considerations of the conformational change in gp41 during activation of HIV-1 fusion.

In the high-resolution structure of the fusion-active gp41 core, the residues at the ***a*** and ***d*** positions of the N-terminal heptad-repeat region form the hydrophobic interface of the central coiled-coil trimer, while three C helices pack along this central coiled-coil surface (*17*−*19*). We have previously shown that formation of the N-terminal coiled-coil structure

is required to specify the overall fold of the six-helix bundle (*16*). Moreover, the folding and stability of the gp41 core correlate well with severity of the in vivo phenotypes observed in cells expressing the full-length HIV-1 envelope protein complex bearing mutations at Ile573, an ***a*** heptad position in the N-terminal coiled coil (*16*, *25*). These results provide direct evidence for the notion that the N-terminal coiled-coil domain is a critical component of the fusion-active gp41 core (*14*).

Studies from several groups have suggested that an isolated peptide corresponding to the N-terminal coiled-coil sequence of HIV-1 gp41 forms an α-helical, tetrameric structure (*21*−*24*). However, the reported analytical sedimentation data fit well to a self-associating monomer−tetramer equilibrium rather than a single-species (i.e., tetramer) model. In this study, we demonstrate that the buried core residues at positions ***a*** and ***d*** of the N-terminal heptad-repeat region direct trimer formation of the leucine-zipper model peptide. The crystallographic analysis reveals that these core residues pack in the characteristic "acute" packing geometry for the parallel three-stranded, α-helical coiled coil. These results strongly suggest that the highly conserved N-terminal coiled-coil motif of gp41 exclusively possesses trimerization specificity. This work also demonstrates the feasibility of using the GCN4 leucine-zipper system to determine the oligomerization state conferred by the core residues of a coiled-coil sequence.

ACKNOWLEDGMENT

We thank Malcolm Cupel of beamline X12B at National Synchrotron Light Source laboratories for support and Jennifer Poitras for secretarial help, Temple Burling for assistance with data collection, Temple Burling, Chris Lima, and Jun Dong for suggestions on structural refinement, and Neville Kallenbach for critical reading of the manuscript.

REFERENCES

1. Luciw, P. A. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melinick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds.) pp 1881−1952, Lippincott-Raven Publishers, Philadelphia.
2. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature 393*, 648−659.
3. Rizzuto, C. D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P. D., Hendrickson, W. A., and Sodroski, J. (1998) *Science 280*, 1949−1953.
4. Moore, J. P., Jameson, B. A., and Dragic, T. (1997) *Curr. Opin. Immunol. 9*, 551−562.
5. Chan, D. C., and Kim, P. S. (1998) *Cell 93*, 681−684.
6. Carr, C. M., and Kim, P. S. (1993) *Cell 73*, 823−832.
7. Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) *Nature 371*, 37−43.
8. Hughson, F. M. (1995) *Curr. Opin. Struct. Biol. 5*, 507−513.
9. Hernandez, L. D., Hoffman, L. R., Wolfsberg, T. G., and White, J. M. (1996) *Annu. Rev. Cell Dev. Biol. 12*, 627−661.
10. Moore, J. P., Jameson, B. A., Weiss, R. A., and Sattentau, Q. J. (1993) in *Viral Fusion Mechanisms* (Bentz, J., Ed.) pp 233−289, CRC Press, Boca Raton, FL.
11. Delwart, E. J., Mosialos, G., and Gilmore, T. (1990) *AIDS Res. Hum. Retroviruses 6*, 703−706.
12. Chambers, P., Pringle, C. R., and Easton, A. J. (1990) *J. Gen. Virol. 71*, 3075−3080.
13. Gallaher, W. R., Ball, J. M., Garry, R. F., Griffin, M. C., and Montelaro, R. C. (1989) *AIDS Res. Hum. Retroviruses 5*, 431−440.
14. Lu, M., Blacklow, S. C., and Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 1075−1082.
15. Lu, M., and Kim, P. S. (1997) *J. Biomol. Struct. Dyn. 15*, 465−471.
16. Lu, M., Ji, H., and Shen, S. (1999) *J. Virol. 73*, 4433−4438.
17. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell 89*, 263−273.
18. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature 387*, 426−430.
19. Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12303−12308.
20. Jiang, S., Lin, K., and Lu, M. (1998) *J. Virol. 72*, 10213−10217.
21. Rabenstein, M., and Shin, Y. K. (1995) *Biochemistry 34*, 13390−13397.
22. Rabenstein, M., and Shin, Y. K. (1996) *Biochemistry 35*, 13922−13928.
23. Lawless, M., Barney, S., Guthrie, K. I., Bucy, T. B., Petteway, S. R., Jr., and Merutka, G. (1996) *Biochemistry 35*, 13697−13708.
24. Shugars, D. C., Wild, C. T., Greenwell, T. K., and Mathews, T. J. (1996) *J. Virol. 70*, 2982−2991.
25. Dubay, J. W., Roberts, S. J., Brody, B., and Hunter, E. (1992) *J. Virol. 66*, 4748−4756.
26. Cao, J., Bergeron, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. (1993) *J. Virol. 67*, 2747−2755.
27. Chen, S. S., Lee, C. N., Lee, W. R., McIntosh, K., and Lee, T. H. (1993) *J. Virol. 67*, 3615−3619.
28. Wild, C. T., Dubay, J. W., Greenwell, T. K., Baird, T., Jr., Oas, T. G., McDanal, C. B., Hunter, E., and Matthews, T. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 12676−12680.
29. Carr, C. M., and Kim, P. S. (1993) *Cell 73*, 823−832.
30. Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) *Nature 371*, 37−43.
31. Chen, C. H., Matthews, T. J., McDanal, C. B., Bolognesi, D. P., and Greenberg, M. L. (1995) *J. Virol. 69*, 3771−3777.
32. Cohen, C., and Parry, A. D. (1990) *Protein 7*, 1−15.
33. Hodges, R. S., Sodek, J., Smillie, L. B., and Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol. 37*, 299−310.
34. McLachlan, A. D., and Stewart, M. (1975) *J. Mol. Biol. 98*, 293−304.
35. Conway, J. F., and Parry, D. A. (1990) *Int. J. Biol. Macromol. 12*, 328−334.
36. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. A. (1991) *Science 254*, 539−544.
37. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature 371*, 80−83.
38. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science 262*, 1401−1407.
39. Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) *Protein Sci. 2*, 383−394.
40. Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., and Eisenberg, D. (1993) *Science 259*, 1288−1293.
41. Betz, S. F., Bryson, J. W., and DeGrado, W. F. (1995) *Curr. Opin. Struct. Biol. 5*, 457−463.
42. Woolfson, D. N., and Alber, T. (1995) *Protein Sci. 4*, 1596−1607.
43. Lumb, K. J., and Kim, P. S. (1995) *Biochemistry 34*, 8642−8648.
44. Gonzalez, L., Jr., Brown, R. A., Richardson, D., and Alber, T. (1996) *Nat. Struct. Biol. 3*, 1002−1010.
45. Gonzalez, L., Jr., Plecs, J. J., and Alber, T. (1996) *Nat. Struct. Biol. 3*, 510−515.
46. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) *Nat. Struct. Biol. 3*, 1011−1018.
47. Potekhin, S. A., Medvedkin, V. N., Kashparov, I. A., and Venyaminov, S. U. (1994) *Protein Eng. 7*, 1097−1101.
48. Staley, J. P., and Kim, P. S. (1994) *Protein Sci. 3*, 1822−1832.
49. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
50. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.

51. Cantor, C., and Schimmel, P. (1980) *Biophysical Chemistry*, Part III, pp 1131−1132, W. H. Freeman and Co., New York.
52. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
53. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.
54. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds) pp 90−125, Royal Society of Chemistry, Cambridge, U.K.
55. Otwinowski, Z., and Minor, W. (1996) *Methods Enzymol. 276*, 307−326.
56. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
57. Brunger, A. T. (1992) *X-PLOR: A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
58. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−147.
59. Jones, T. A., Zou, J.-Y., and Cowan, S. W. (1991) *Acta Crystallogr. A47*, 110−119.
60. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science 243*, 538−542.
61. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6442−6446.
62. Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 689−697.
63. Ponder, J. W., and Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.
64. Farzan, M., Choe, H., Desjardins, E., Sun, Y., Kuhn, J., Cao, J., Archambault, D., Kolchinsky, P., Koch, M., Wyatt, R., and Sodroski, J. (1998) *J. Virol. 72*, 7620−7625.
65. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.
66. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.

BI990199W