THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 31, Issue of August 3, pp. 29485–29489, 2001

Printed in U.S.A.

Design and Properties of N_{CCG}-gp41, a Chimeric gp41 Molecule with Nanomolar HIV Fusion Inhibitory Activity*

Received for publication, June 11, 2001
Published, JBC Papers in Press, June 19, 2001, DOI 10.1074/jbc.C100317200

John M. Louis, Carole A. Bewley‡, and G. Marius Clore§

From the Laboratories of Chemical Physics and Bioorganic Chemistry, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

The design and characterization of a chimeric protein, termed N_{CCG}-gp41, derived from the ectodomain of human immunodeficiency virus (HIV), type I gp41 is described. N_{CCG}-gp41 features an exposed trimeric coiled-coil comprising the N-terminal helices of the gp41 ectodomain. The trimeric coiled-coil is stabilized both by fusion to a minimal thermostable ectodomain of gp41 and by engineered intersubunit disulfide bonds. N_{CCG}-gp41 is shown to inhibit HIV envelope-mediated cell fusion at nanomolar concentrations with an IC_{50} of 16.1 ± 2.8 nM. It is proposed that N_{CCG}-gp41 targets the exposed C-terminal region of the gp41 ectodomain in its pre-hairpin intermediate state, thereby preventing the formation of the fusogenic form of the gp41 ectodomain, which comprises a highly stable trimer of hairpins arranged in a six-helix bundle. N_{CCG}-gp41 has potential as a therapeutic agent for the direct inhibition of HIV cell entry, as an anti-HIV vaccine, and as a component of a rapid throughput assay for screening for small molecule inhibitors of HIV envelope-mediated cell fusion. It is anticipated that antibodies raised against N_{CCG}-gp41 may target the trimeric coiled-coil of N-terminal helices of the gp41 ectodomain that is exposed in the pre-hairpin intermediate state in a manner analogous to peptides derived from the C-terminal helix of gp41 that are currently in clinical trials.

Infection by and dissemination of the human immunodeficiency virus (HIV)[1] necessitates virus-cell or cell-cell fusion mediated by envelope (Env) glycoproteins (1). HIV-I Env consists of two non-covalently attached proteins, gp120 and gp41, derived by proteolytic cleavage of gp160 (1). The molecular events leading to fusion include initial binding of gp120 to CD4, triggering conformational changes in gp120 that permit subsequent interactions with the chemokine receptors, CXCR4 or CCR5 (2). This results in a further series of conformational changes in the gp120/gp41 oligomer that lead to insertion of the fusion peptide of gp41 into the target membrane and ultimately membrane fusion. Both gp120 and gp41 offer potential targets for the inhibition of viral entry through either drugs or neutralizing antibodies. Neutralizing antibodies directed at gp120 have been difficult to elicit (3), because the surface of gp120 is heavily glycosylated (4), thereby preventing access to the conserved regions of the molecule that bind the requisite receptors. One very potent inhibitor of fusion directed against gp120 has been discovered, namely, the small protein cyanovirin-N (5) whose selectivity has been shown to arise as a consequence of specific nanomolar binding to $Man_9GlcNac_2$ and the D1D3 isomer of $Man_8GlcNac_2$ present in abundance on the surface of gp120 (6). A potentially more amenable target is afforded by gp41, particularly in the so-called pre-hairpin intermediate state (7).

The solution structure of the complete ectodomain of simian immunodeficiency virus gp41 (8) and crystal structures of various fragments of the ectodomain cores of HIV (9–11) and simian immunodeficiency virus (12) gp41 have been determined. Each gp41 monomer consists of two long helices at the N- and C termini connected by a long linker. The core of gp41 is a trimer of hairpins making up a six-helix bundle; three N-terminal helices form a central parallel coiled-coil around which are packed the C-terminal helices in an antiparallel manner. This structure is thought to represent the fusogenic state of gp41, which serves to bring the viral and cell membranes into close proximity, thereby promoting membrane fusion (7). It has long been known that peptides derived from the C- and N-helices inhibit fusion (13, 14). The C-peptides, which are monomeric in solution (15), have nanomolar IC_{50} values (14, 16), and some of these are currently in clinical trials (17). The activity of the N-peptides, on the other hand, is about three orders of magnitude lower (13), presumably because of aggregation and their inability to form a trimeric coiled-coil in the absence of C-peptide (15). Both the N- and C-peptides are thought to target the pre-hairpin fusion intermediate, which persists for many minutes (7, 18). Formation of the pre-hairpin intermediate in which the N-terminal fusion peptide of gp41 is inserted into the target membrane is postulated to expose the trimeric coiled-coil of N-helices to which the C-peptides bind with high affinity, thereby preventing formation of the fusogenic trimer of hairpins (7, 18). The N-peptides are thought to either hinder the formation of the trimeric coiled-coil of N-helices (19) or bind to the C-terminal region of the gp41 ectodomain corresponding to the C-helix in the fusogenic state of gp41 (7). The N-helix of gp41 has also been targeted by cyclic D-peptide inhibitors derived from phage display (20) and a vari-

* This work was supported in part by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G. M. C. and C. A. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Laboratory of Bioorganic Chemistry, Bldg. 8, NIDDK, National Institutes of Health, Bethesda, MD 20892-0820. Tel.: 301-594-5187; E-mail: caroleb@intra.niddk.nih.gov.

§ To whom correspondence should be addressed: Laboratory of Chemical Physics, Bldg. 5, Rm. B1-30I, NIDDK, National Institutes of Health, Bethesda, MD 20892-0510. Tel.: 301-496-0782; Fax: 301-496-0825; E-mail: clore@speck.niddk.nih.gov.

[1] The abbreviations used are: HIV, human immunodeficiency virus; Env, viral envelope glycoproteins; HIV-I, HIV, type I; HPLC, high pressure liquid chromatography; N34, N35, and N36 are segments derived from the N-terminal helix of gp41 and comprise residues 546–579, 546–580, and 546–581, respectively, of HIV-I Env; C28 and C34 are segments derived from the C-terminal helix of gp41 and comprise residues 628–655 and 628–661, respectively, of HIV-I Env; N34-(L6)-C28 is a construct comprising peptides N34 and C28 connecting by a six-residue linker (L6).

ety of non-natural binding elements generated by combinatorial chemistry and linked to truncated C-peptides (21); to date the reported inhibitory activities of these compounds are only in the micromolar range.

In this paper we present the design of a chimeric gp41 molecule, which we term N_{CCG}-gp41 (see Fig. 1), in which the N-helix of HIV gp41 is grafted in helical phase onto the N terminus of a minimal thermostable trimeric core (six-helix bundle) of gp41 and stabilized by intermolecular disulfide bridges. We show that N_{CCG}-gp41 presents a stable and exposed trimeric coiled-coil of N-helices that inhibits fusion at nanomolar concentrations. In addition to its fusion inhibitory properties, N_{CCG}-gp41 has been designed with the aim of presenting an epitope suitable for the generation of fusion inhibitory antibodies directed against the exposed N-helices of gp41 in the pre-hairpin intermediate state.

EXPERIMENTAL PROCEDURES

Chemical Synthesis and Cloning of N_{CCG}-gp41—The DNA encoding the N-gp41 chimeric protein was synthesized (Life Technologies, Inc.) in approximately four equal fragments as shown in Fig. 2. The nucleotide sequence was biased for optimal codon usage in *Escherichia coli* (www.kazusa.or.jp/codon/). These fragments were assembled in a manner similar to that described previously (22), and the full-length fragment was cloned into the pET11a vector (Novagen, Madison, WI) for expression in *E. coli* BL21(DE3). The L31C, Q32C, and A33G mutations were introduced into the N-gp41 construct using the Quick-Change mutagenesis protocol (Stratagene, La Jolla, CA) to create the N_{CCG}-gp41 construct.

Expression, Purification, and Folding of N_{CCG}-gp41—Cells were grown at 37 °C either in Luria-Bertani medium or in a modified minimal medium for uniform (>99%) ^{15}N labeling with $^{15}NH_4Cl$ as the sole nitrogen source and induced with 2 mM isopropyl β-D-thiogalactoside for 4 h. Cells derived from 1 liter of culture expressing N_{CCG}-gp41 were suspended in 20 volumes of Buffer A (50 mM Tris-HCl, pH 8.2, 10 mM EDTA, and 10 mM dithiothreitol) and lysed by sonication at 4 °C in the presence of 100 μg/ml lysozyme. The insoluble fraction was washed by resuspension in Buffer A containing 1 M urea and 0.5% Triton X-100 and then in Buffer A. In both cases, the insoluble fraction was pelleted by centrifugation at 20,000 × *g* for 30 min at 4 °C. The final pellet was solubilized in 50 mM Tris-HCl, pH 8.0, 7.5 M guanidine-HCl, 5 mM EDTA, and 20 mM dithiothreitol to yield a concentration not exceeding 20 mg/ml. 30 mg of protein was applied at ambient temperature and a flow-rate of 3 ml/min to a Superdex-75 column (HiLoad, 2.6 × 60-cm; Amersham Pharmacia Biotech) equilibrated in 50 mM Tris-HCl, pH 8, 4 M guanidine-HCl, 5 mM EDTA, and 5 mM dithiothreitol. Peak fractions were then subjected to reverse-phase HPLC on POROS RII resin (Perceptive Biosystems) using a linear gradient of 0 to 60% acetonitrile/0.05% trifluoroacetic acid. Peak fractions were combined, and 7 mg of protein was diluted to a concentration of ~0.2 mg/ml in 35% acetonitrile/water/0.05% trifluoroacetic acid and dialyzed against 2 liters of 50 mM sodium formate buffer, pH 3.0, for 3 h and then against 2 liters of the same buffer overnight at 4 °C. The protein was concentrated to ~3 mg/ml and stored at 4 °C.

A CD spectrum of N_{CCG}-gp41 was recorded at 25 °C on a JASCO J-720 spectropolarimeter using a 0.05-cm path length cell. Quantitative evaluation of secondary structure from the CD spectrum was carried out using the program *k2d* (available at www.bioinformatik.biochemtech.uni-halle.de/), which employs a neural network approach for CD spectra deconvolution (23). A ^1H–^{15}N heteronuclear single quantum coherence correlation spectrum of uniformly ^{15}N-labeled N_{CCG}-gp41 was recorded at 40 °C at 600 MHz on a Bruker DRX600 NMR spectrometer.

Sedimentation Equilibrium—Sedimentation equilibrium experiments were conducted at 20.0 °C and three different rotor speeds (10,000, 12,000, and 14,000) on a Beckman Optima XL-A analytical ultracentrifuge. Protein samples were prepared in 50 mM sodium formate buffer, pH = 3, and loaded into the ultracentrifuge cells at nominal loading concentrations of 0.80 A_{280}. Data were analyzed in terms of a single ideal solute to obtain the buoyant molecular mass, $M(1 - \upsilon\rho)$, using the Optima XL-A data analysis software (Beckman). The value for the experimental molecular mass, M, was determined using calculated values for the density, ρ (determined at 20 °C using standard tables), and partial specific volume, υ (calculated on the basis of amino acid composition; see Ref. 24).

Cells and Reagents—NIH-3T3 and B-SC-1 cells (American Type Culture Collection) grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and gentamycin at 50 μg/ml (all from Life Technologies, Inc.), were used for all assays. Recombinant vaccinia viruses used in this study include vCB41 (25), vCBYF1-fusin (26), vP11T7gene1 (27), and vCB21R-LacZ, which encode HIV-I LAV (Lai) Env (T-cell line tropic), CXCR4, bacteriophage T7 RNA polymerase driven by a vaccinia virus promoter, and the *E. coli lacZ* gene under control of the T7 promoter, respectively. Two-domain-soluble CD4 (1–183) was a gift from E. Berger (National Institute of Allergy and Infectious Diseases, National Institutes of Health) and donated by S. Johnson (Pharmacia Upjohn, Kalamazoo, MI). N36 (residues 546–581 of HIV-I Env) and C34 (residues 628–661 of HIV-I Env) peptides, acetylated at their N termini and amidated at their C termini, were synthesized by solid-phase peptide synthesis (Commonwealth Biotechnologies, Richmond, VA), purified by reverse phase HPLC, and characterized by mass spectrometry.

Cell Fusion Assay—A modification (28) of the vaccinia virus-based reporter gene assay employing soluble CD4 (200 nM final concentration) was used to determine the effect on HIV Env-mediated cell fusion of N_{CCG}-gp41 and C34 and N36 peptides. Assays were conducted essentially as described (28). Briefly, NIH-3T3 and B-SC-1 cells were used for target and effector cell populations, respectively. Target cells were co-infected with vCB21R-LacZ and vCBYF1-fusin (CXCR4) and effector cells with vCB41 (Env) and vP11T7gene1 at a multiplicity of infection of 10. For inhibition studies, N_{CCG}-gp41, C34, or N36 were added to an appropriate volume of Dulbecco's modified Eagle's medium supplemented with 2.5% fetal bovine serum and PBS to yield identical buffer compositions (100 μl), followed by addition of 1×10^5 effector cells (in 50 μl of medium) per well. After incubation for 15 min, 1×10^5 target cells (in 50 μl) and soluble CD4 were added to each well. Following 2.5 h of incubation, β-galactosidase activity of cell lysates was measured (A_{570}; Molecular Devices 96-well spectrophotometer) upon addition of chlorophenol red-β-D-galactopyranoside (Roche Molecular Biochemicals). At least two independent experiments (performed in duplicate) were conducted for each inhibitor. The curves for % fusion *versus* inhibitor concentration, [I], were fit by non-linear optimization to the activity relationship, $\%\text{fusion} = 100/(1 + [I]/IC_{50})$.

Modeling—The structure of N_{CCG}-gp41 was modeled and regularized using the National Institutes of Health version (available by anonymous ftp at portal.niddk.nih.gov in the directory /pub/clore/xplor_nih) of the program XPLOR (29) and visualized using the program VMD-XPLOR (30).

RESULTS AND DISCUSSION

Design of N_{CCG}-gp41—The minimal trimeric core of the ectodomain of HIV gp41, N34-(L6)-C28, which has been solved crystallographically (11), comprises residues 546–579 (N34) and 628–655 (C28) of HIV-I Env covalently linked by a six-residue SGGRGG loop (L6) (Fig. 1*a*). The structure forms a thermostable trimer of hairpins in which N34 and C28 are entirely helical and arranged in a six-helix bundle (*top half* of Fig. 1*b*). We have made use of the N34-(L6)-C28 core as a scaffold to graft a 35-residue sequence (N35) comprising residues 546–580 of HIV-I Env onto its N terminus, taking care that the helical register is continuously maintained from N35 through N34 (Fig. 1*a*). The rationale for this design is that a fully accessible trimeric coiled-coil of N35 helices will be stabilized as a consequence of its being covalently linked to the gp41 ectodomain core (Fig. 1*b*). To further stabilize the trimeric coiled-coil of N35 helices, and to ensure that the entire molecule remains trimeric at very low concentrations, we substituted the sequence Leu-Gln-Ala located at the C-terminal end of N35 (residues 31–33 of the construct) to Cys-Cys-Gly (Fig. 1, *a* and *b*). This approach was used previously in an attempt to stabilize the gp160 envelope protein (31). The location of the two cysteines on the helix face (at positions d and e, respectively) is such that three intermolecular disulfide bonds can readily be formed between the three subunits (A, B, and C), Cys31(A)-Cys32(B), Cys31(B)-Cys32(C), and Cys31(C)-Cys32(A) (Fig. 1, *c* and *d*). The A33G mutation was employed to ensure that minor adjustments of the polypeptide backbone can readily occur to ensure disulfide bond formation.

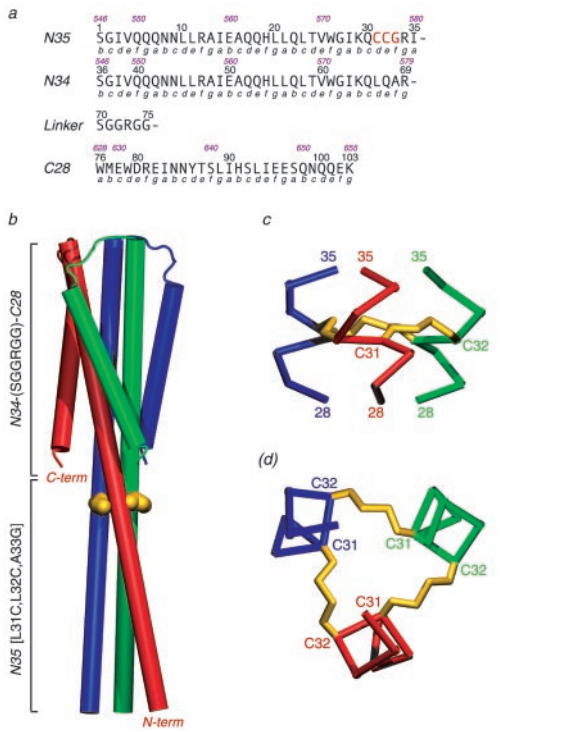

FIG. 1. **Design of the N_{CCG}-gp41 chimera.** *a*, sequence of N_{CCG}-gp41. Residue numbering for N_{CCG}-gp41 is indicated in *black*, and the corresponding residue numbering for HIV-I Env is shown in *purple*. The location of the residues in a helical wheel are indicated by the *small italic letters* (*a–g*) *below* the *sequence*. The sequence comprises residues 546–580 of HIV-I Env (residues 1–35; *N35*) with Leu^{576}, Gln^{577}, and Ala^{578} (residues 31–33 of N_{CCG}-gp41) mutated to Cys, Cys, and Gly, respectively; residues 546–579 of Env (residues 36–69; *N34*); a six-residue linker (residues 70–75); and finally residues 628–655 of Env (residues 76–103; *C28*). *b*, model of N_{CCG}-gp41. The structure of the N34-(L6)-C28 portion of N_{CCG}-gp41 has been solved crystallographically (11). N35 was grafted onto the N-terminal end of the crystal structure to generate a 69-residue continuous α-helix comprising N35 and N34. The three subunits, *A*, *B*, and *C*, are indicated in *blue*, *red*, and *green*, respectively, and the location of the three intersubunit disulfide bonds is shown in *gold*. Detailed side (*c*) and top (*d*) views illustrating that the three intersubunit disulfide bonds can readily be formed with good stereochemistry. The backbone is shown as a Cα trace with the same coloring code as in *b*, and the disulfide bonds are shown in *gold*.

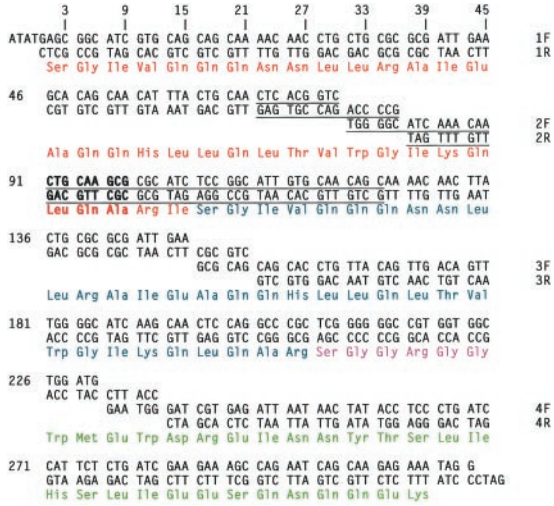

FIG. 2. **Construction of synthetic N-gp41 and N_{CCG}-gp41 genes for expression in *E. coli*.** The N-gp41 coding sequence and its complementary sequence were synthesized in approximately four equal fragments and purified by polyacrylamide gel electrophoresis. *1F*, *2F*, *3F*, and *4F* denote the four fragments, and their complementary fragments are *1R*, *2R*, *3R*, and *4R*, respectively. Each of the fragments was phosphorylated except for 1F and 4R. Fragments 1F and 1R, 2F and 2R, 3F and 3R, and 4F and 4R were annealed and then ligated. The assembled full-length N-gp41 DNA was isolated by gel electrophoresis and subsequently cloned into the *Nde*I and *Bam*HI sites of the pET11a vector. The DNA sequence of two individual clones was confirmed by sequencing. The L31C, Q32C, and A33G mutations were introduced into the N-gp41 construct using the Quick-Change mutagenesis protocol (Stratagene, La Jolla, CA) to generate the N_{CCG}-gp41 construct. The *underlined sequence* denotes the forward and reverse primers used in which the sequence CTGCAAGCG (*bold letters*) was changed to TGT-TGTGGC to encode Cys^{31}, Cys^{32}, and Gly^{33} in the N_{CCG}-gp41 construct. The amino acid sequence is shown *below* the nucleotide sequence with the N35, N34, linker, and C28 portions of the amino acid sequence shown in *red*, *blue*, *purple*, and *green*, respectively. Note that although the amino acid sequences of N34 and the first 34 residues of N35 are identical, their nucleotide sequences are different.

The proposed structure for the chimeric protein, which we term N_{CCG}-gp41, is shown in Fig. 1*b*. The model was constructed from the x-ray structure of N34-(L6)-C28 (11) and the internal trimeric coiled-coil of N-helices from the x-ray structure of a longer ectodomain fragment of gp41 (comprising residues 541–588 and 628–665 of HIV-I Env) (10). N35 was grafted onto N34 by best-fitting the backbone of residues 581–586 (*i.e.* five residues C-terminal of N35) of the longer HIV gp41 ectodomain onto the backbone of residues 546–551 (*i.e.* the first five residues) of the N34-(L6)-C28 construct, followed by deletion of residues 581–586; substitution of Leu^{576}, Leu^{577}, and Ala^{578} in N35 by Cys, Cys, and Gly, respectively; and regularization to covalently link the N35 and N34 helices of each subunit and form intermolecular disulfide bonds with good stereochemistry. N35 (residues 1–35 of N_{CCG}-gp41) and N34 (residues 36–69 of N_{CCG}-gp41) form a continuous 69 residue α-helix, ~100-Å-long. The internal trimeric coiled-coil of N34 helices is surrounded by three C-terminal helices (C28 corresponding to residues 76–103 of N_{CCG}-gp41), whereas the trimeric coiled-coil of N35 helices, ~50-Å in length, is fully exposed to solvent.

Construction of Synthetic N-gp41 and N_{CCG}-gp41 Genes for Expression in E. coli—The strategy employed for synthesizing the genes for N-gp41 and N_{CCG}-gp41 is shown in Fig. 2. The N-gp41 coding sequence and its complementary sequence were synthesized in approximately four equal fragments, which were ligated to yield full-length N-gp41 DNA and then cloned into the pET11a expression vector. The following two critical features of the nucleotide sequence are noteworthy: first, codon usage was optimized for *E. coli*; second, within these confines, non-identical codon usage was employed for the N35 and N34 portions of the gene. Hence, although the amino acid sequences of residues 1–34 and 36–69 of N-gp41 are identical (Fig. 1*a*), the corresponding nucleotide sequences are different (Fig. 2). This is important, because it allows one to mutagenize at will residues within N35 without affecting N34. The L31C, Q32C, and A33G mutations in N35 were then introduced into the N-gp41 coding sequence to generate the N_{CCG}-gp41 construct.

Biophysical Properties of N_{CCG}-gp41—Folded N_{CCG}-gp41 is fully soluble at low pH values (less than 4) and elutes as a single peak on size-exclusion chromatography (Fig. 3*a*). Quantitative analysis of elution profiles down a Superdex-75 column as described (32) indicates that N_{CCG}-gp41 elutes with an apparent molecular mass of 30,000 Da corresponding to a trimer (which has a calculated molecular mass of 35,442 Da). No evidence of dimer or monomer was apparent at the lowest concentration of about 140 nM tested on the Superdex-75 column. The trimeric nature of N_{CCG}-gp41 was confirmed by sedimentation equilibrium studies, which indicated that N_{CCG}-gp41 behaves as a single monodisperse species with a molecular mass of 35,600 ± 150 Da.

N_{CCG}-gp41 was specifically designed with the aim of generating a protein in which the subunits of the trimer are co-

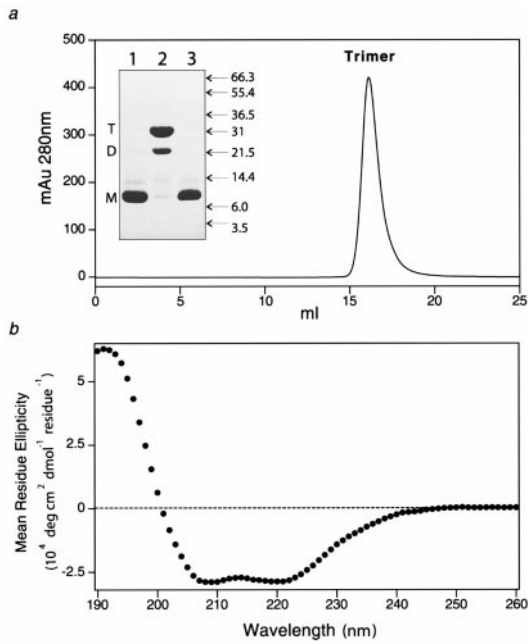

FIG. 3. **Characterization of N_{CCG}-gp41.** *a*, analysis of purified native N_{CCG}-gp41 by size-exclusion column chromatography. N_{CCG}-gp41 was fractionated on a Superdex-200 column in 50 mM sodium formate buffer, pH 3.0, and 0.2 M GnHCl at room temperature (*mAu*, milliabsorbance units). The *inset* shows the SDS-polyacrylamide gel electrophoresis analysis of N_{CCG}-gp41. *Lane 1*, purified N_{CCG}-gp41 under non-reducing conditions after reverse-phase HPLC but before protein folding; *lane 2*, N_{CCG}-gp41 under non-reducing conditions after protein folding by dialysis of the protein from 35% acetonitrile/0.05% trifluoroacetic acid into 50 mM sodium formate buffer, pH 3; *lane 3*, the same fraction of N_{CCG}-gp41 shown in *lane 2* under reducing conditions following treatment with 0.55 M 2-mercaptoethanol. The positions of N_{CCG}-gp41 trimer, dimer, and monomer are indicated by the letters *T*, *D*, and *M*, respectively, on the *left-hand side* of the *inset*; the positions of molecular mass markers (in kDa) are indicated on the *right-hand side* of the *inset*. All samples were heated to 90 °C for 2 min in the presence of 1.5% SDS and 17 μM protein in 50 mM Tris-HCl, pH 8.0, with or without 2-mercaptoethanol, prior to loading on the gel. *b*, CD spectrum of N_{CCG}-gp41 (9.94 μM trimer in 10 mM sodium formate buffer, pH 3, at ambient temperature).

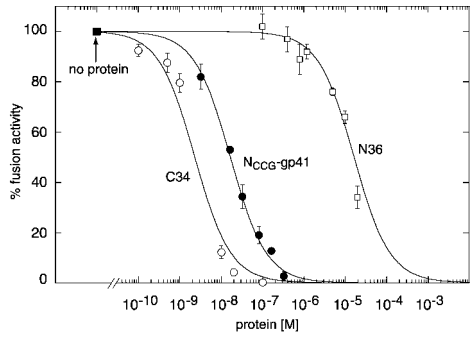

FIG. 4. **Inhibition of HIV-I envelope-mediated cell fusion by the N_{CCG}-gp41 chimera and the gp41-derived peptides C34 and N36.** The C34 peptide corresponds to residues 628–631 on HIV-I Env, which comprise the C-helix of gp41; the N36 peptide corresponds to residues 546–581 of Env, which comprise the N- helix of gp41. *Solid circles*, N_{CCG}-gp41; *open circles*, C34; *open squares*, N36. *Vertical bars* indicate standard. The *solid lines* represent best-fits to the data using the following simple activity relationship: %fusion = $100/(1+[I]/IC_{50})$, where [I] is the inhibitor concentration. The IC_{50} values for N_{CCG}-gp41, C34, and N36 are 16.1 ± 2.8 nM, 2.3 ± 0.5 nM, and 16.4 ± 1.8 μM, respectively.

valently linked by three intermolecular disulfide bonds, located at the C-terminal end of the N35 portion of the construct (Fig. 1, *b–d*), thereby ensuring that dissociation into monomers cannot occur. This is indeed found to be the case experimentally. SDS-polyacrylamide gel electrophoresis of refolded N_{CCG}-gp41 demonstrates that under non-reducing conditions the majority (~90%) of the protein migrates as a trimer with about 10% as a dimer (see Fig. 2*a*, *lane 2*). This result is consistent with mass spectroscopic analyses of non-reduced N_{CCG}-gp41, which shows the presence of trimer and minor dimer forms with experimental masses of *m/z* 35,442 and 23,629, respectively. These values are essentially identical to the expected values of 35441.7 and 23627.8 for the trimer and dimer, respectively. When the disulfide bonds are reduced, however, by the addition of 2-mercaptoethanol, all the N_{CCG}-gp41 migrates as a monomer as expected (see Fig. 2*a*, *lane 3*).

The CD spectrum of N_{CCG}-gp41 (Fig. 3*b*) displays the characteristic signature of an α-helical protein with double minima at 208 and 222 nm. Deconvolution of the CD spectrum with the neural network program *k2d* (23) yields an α-helical content of 96%, in complete agreement with the model in Fig. 1*b* in which the only non-helical residues are located in the six-residue loop connecting the N34 and C28 helices. The CD results are also completely consistent with the $^1H–^{15}N$ correlation (heteronuclear single quantum coherence) NMR spectrum of N_{CCG}-gp41 (data not shown), which is reminiscent of that of the complete ectodomain of simian immunodeficiency virus gp41 (33) and displays rather limited dispersion of the backbone amide proton resonances (9.3–6.5 ppm), as expected for a predominantly helical protein.

Inhibition of HIV-I Env-mediated Cell Fusion by N_{CCG}-gp41—A quantitative vaccinia-virus-based reporter gene assay (28) was used to assess the ability of N_{CCG}-gp41 to inhibit HIV-I Env-mediated cell fusion. In this assay, the extent of HIV-I Env-mediated cell fusion upon addition of soluble CD4 between effector cells bearing HIV-I Env (LAV) on their surface and target cells expressing the chemokine receptor CXCR4 is directly monitored via rates of β-galactosidase activity. Fusion activity as a function of N_{CCG}-gp41 concentration is shown in Fig. 4 where it can be clearly seen that N_{CCG}-gp41 inhibits HIV-I Env-mediated cell fusion at nanomolar concentrations with an IC_{50} of 16.1 ± 2.8 nM. For comparison we also carried out a set of parallel experiments with the C34 and N36 peptides (15) derived from the N- and C-terminal helices of gp41. (C34 comprises residues 628–661 of HIV-I Env and corresponds to the C28 portion of N_{CCG}-gp41 plus an additional six residues at its C terminus; N36 comprises residues 546–581 of HIV-I Env and corresponds to the N35 portion of N_{CCG}-gp41 plus an additional residue at its C terminus; *cf.* Fig. 1*a*.) The C34 peptide has an IC_{50} of 2.2 ± 0.5 nM, in agreement with previous studies (16, 21). For comparison, DP-178, the original C-peptide shown to have fusion inhibitory activity (14) and comprising residues 638–673 of HIV-I Env, which overlaps with the C-terminal half of the C34 peptide, has an IC_{50} of ~50 nM (21). The fusion inhibitory activity of the N36 peptide, however, is much lower, in the micromolar range with an IC_{50} of 16.4 ± 1.8 μM, also consistent with previous work (13).

N_{CCG}-gp41 is three orders of magnitude more potent than the N36 peptide in inhibiting HIV-I Env-mediated fusion (Fig. 4). This must be due to the fact that the N36 peptide aggregates and does not form a trimeric coiled-coil in the absence of C34 peptide (15). In contrast, the exposed N35 portion of N_{CCG}-gp41 is maintained in the trimeric helical coiled-coil conformation as a consequence of its being directly linked to the stable trimeric, six-helix core of gp41 and further stabilized by the presence of three intersubunit disulfide bonds (Fig. 1*b*). Note that the trimeric six-helix core of gp41 possesses no inhibitory activity (34, so that the activity of N_{CCG}-gp41 is entirely due to the presence of the exposed trimeric helical coiled-coil formed by the N35 portion of the molecule.

Because the N35 portion of N_{CCG}-gp41 forms a stable, covalently linked, trimeric coiled-coil that is entirely helical, N_{CCG}-gp41 must target the accessible C-terminal region of the gp41 ectodomain in its pre-hairpin intermediate state. Thus, whereas the C-terminal region of the gp41 ectodomain forms an outer helical shell that interacts with the internal trimeric coiled-coil of N-helices to form a six-helix bundle in the fusogenic state solved by NMR (8) and crystallography (9–12), in the pre-hairpin intermediate state, the N- and C-terminal regions of the gp41 ectodomain cannot be interacting with each other, consistent with current models (7, 34).

We also carried out a series of experiments in which fully inhibitory concentrations of N_{CCG}-gp41, C34, or N36, in the presence or absence of CD4, were added to effector cells, followed by repeated washing prior to adding the effector cells to target cells in the presence of soluble CD4. Under these conditions, cell fusion is observed in all three cases, suggesting that all three molecules act on a fusion intermediate of gp41 generated subsequent to the interaction of HIV-I envelope with cellular receptors, in agreement with previous studies on peptides derived from the C-helix of gp41 (18).

Concluding Remarks—In this paper we have presented the design and biophysical characterization of a chimeric protein, N_{CCG}-gp41, in which the solvent-exposed trimeric N-helix coiled-coil of gp41 is stabilized by directly attaching it to a thermostable core fragment of gp41 comprising a trimer of hairpins (six-helix bundle) and covalently linking the subunits by engineering intermolecular disulfide bridges. We also demonstrated that N_{CCG}-gp41 is a highly potent inhibitor of HIV-I Env-mediated cell fusion at nanomolar concentrations with an IC_{50} of ~16 nM. While this work was in progress, Kim and colleagues (34) described the properties of an engineered protein, which they termed 5-helix and which comprises a single 200-residue polypeptide chain (N36-L6-C34-L6-N36-L6-C34-L6-N36, where L6 is a six residue linker) in which the N36 segments form a parallel triple helix coiled-coil at the center of the molecule, surrounded by two antiparallel (C34) helices, thereby leaving one N36 helix exposed to solvent. The inhibitory activity of 5-helix is comparable with that of N_{CCG}-gp41, and it is likely that both function in a very similar manner. Both N_{CCG}-gp41 and 5-helix could potentially be used as therapeutic agents (either parenteral or possibly as a topical microbicide), in a manner analogous to peptides derived from the C-helix of gp41 (14, 16) that are currently undergoing clinical trials (17). N_{CCG}-gp41 offers a number of unique advantages in this regard. N_{CCG}-gp41 is covalently linked by intersubunit disulfide bridges between the exposed N-helices and is about twice the size of 5-helix. In principle, these features should render N_{CCG}-gp41 immune from the effects of potential proteolytic cleavage at the linkers between the N- and C-helices and increase its lifetime by reducing its degradation rate in the liver and clearance by the kidneys. Note that cleavage of the linkers in 5-helix will eventually result in repartitioning of the N- and C-helices to yield a significant proportion of inactive six-helix bundle. Moreover, at very low concentrations dissociation of the N- and C-helices could occur; cleavage of the linkers in N_{CCG}-gp41, on the other hand, may lead to dissociation of the C-helices at very low concentrations, but the trimeric coiled-coil of N-helices will still be preserved because of the presence of the intersubunit disulfide bonds. In addition, N_{CCG}-gp41 exposes the complete internal trimeric N-helix coiled-coil of gp41 in a stable manner. This feature, together with the larger size of N_{CCG}-gp41, should be advantageous with respect to the possible use of N_{CCG}-gp41 as an anti-HIV vaccine. We anticipate that immunization with N_{CCG}-gp41 will generate antibodies directed against the internal trimeric N-helix coiled-coil of gp41, which is exposed in the pre-hairpin intermediate state. We predict that such antibodies will function in a manner analogous to the highly inhibitory peptides derived from the C-helix of gp41. This hypothesis is currently being tested in our laboratory. Finally, one can also anticipate that N_{CCG}-gp41, in combination with the C34 or C28 peptide, could be used for high throughput screening for small molecule inhibitors that bind to the exposed trimeric N-helix coiled-coil of N_{CCG}-gp41, thereby inhibiting HIV Env-mediated cell fusion.

Acknowledgments—We thank L. C. Chang and I. Neshiwat for expert technical assistance, E. Berger and P. Kennedy for recombinant vaccinia viruses and soluble CD4, R. Ghirlando for analytical ultracentrifugation, and L. Pannell for mass spectrometry.

REFERENCES

1. Freed, E. O., and Martin, M. A. (1995) *J. Biol. Chem.* **270,** 23883–23886
2. Moore, J. P., Trkola, A., and Dragic, T. (1997) *Curr. Opin. Immunol.* **9,** 551–562
3. McMichael, A. J., and Hanke, T. (1999) *Nat. Med.* **5,** 612–614
4. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature* **393,** 648–659
5. Boyd, M. R., Gustafson, K. R., McMahon, J. B., Shoemaker, R. H., O'Keefe, B. R., Mori, T., Gulakowski, R. J., Wu, L., Rivera, M. I., Laurencot, C. M., Currens, M. J., Cardellina, J. H., Buckheit, R. W., Nara, P. L., Pannell, L. K., Sowder, R. C., and Henderson, L. E. (1997) *Antimicrob. Agents Chemother.* **41,** 1521–1530
6. Bewley, C. A., and Otero-Quintero, S. (2001) *J. Am. Chem. Soc.* **123,** 3892–3902
7. Chan, D. C., and Kim, P. S. (1998) *Cell* **93,** 681–684
8. Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) *EMBO J.* **17,** 4572–4584
9. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell* **89,** 263–273
10. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature* **387,** 426–430
11. Tan, K. J., Liu, J., Wang, S., Shen, S., and Lu, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12303–12308
12. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 9134–9139
13. Wild, C. T., Oas, T., McDanal, C. B., Bolognesi, D., and Matthews, T. J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 10537–10541
14. Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B., and Matthews, T. J. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 9770–9774
15. Lu, M., Blacklow, S. C., and Kim, P. S. (1995) *Nat. Struct. Biol.* **2,** 1075–1082
16. Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 15613–15617
17. Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M., and Saag, M. S. (1998) *Nat. Med.* **4,** 1302–1307
18. Furuta, R. A., Wild, C. T., Weng, Y., and Weiss, C. D. (1998) *Nat. Struct. Biol.* **5,** 276–279
19. Weng, Y., and Weiss, C. D. (1998) *J. Virol.* **72,** 9676–9682
20. Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A., and Kim, P. S. (1999) *Cell* **99,** 103–115
21. Ferrer, M., Kapoor, T. M., Strassmaier, T., Weissenhorn, W., Skehel, J. J., Oprian, D., Schreiber, S. L., Wiley, D. C., and Harrison, S. C. (1999) *Nat. Struct. Biol.* **6,** 953–960
22. Louis, J. M., Wondrak, E. M., Copeland, T. D., Smith, C. A. D., Mora, P. T., and Oroszlan, S. (1989) *Biochem. Biophys. Res. Comm.* **159,** 87–94
23. Andrade., M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) *Protein Eng.* **6,** 383–390
24. Perkins, S. J. (1986) *Eur. J. Biochem.* **157,** 169–180
25. Broder, C. C., and Berger, E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 9004–9008
26. Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996) *Science* **272,** 872–877
27. Alexander, W. A., Moss, B., and Fuerst, T. T. (1992) *J. Virol.* **66,** 2934–2942
28. Salzwedel, K., Smith, E., Dey, B., and Berger, E. J. (2000) *J. Virol.* **74,** 326–333
29. Brünger, A. T. (1993) *X-PLOR: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT
30. Schwieters, C. D., and Clore, G. M. (2001) *J. Magn. Reson.* **149,** 239–244
31. Farzan, M., Choe, H., Desjardins, E., Sun, Y., Kuhn, J., Cao, J., Archambault, D., Kolchinski, P., Koch, M., Wyatt, R., and Sodroski, J. (1998) *J. Virol.* **72,** 7620–7625
32. Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M., and Wlodawer, A. (1999) *J. Mol. Biol.* **288,** 403–412
33. Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Gronenborn, A. M., and Clore, G. M. (1997) *J. Mol. Biol.* **271,** 819–826
34. Root, M. J., Kay, M. S., and Kim, P. S. (2001) *Science* **291,** 884–888