Protein grafting of an HIV-1-inhibiting epitope

Samuel K. Sia*†‡ and Peter S. Kim*§

*Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142; and †Committee on Higher Degrees in Biophysics, Harvard University, Cambridge, MA 02138

Contributed by Peter S. Kim, June 24, 2003

Protein grafting, the transfer of a binding epitope of one ligand onto the surface of another protein, is a potentially powerful technique for presenting peptides in preformed and active three-dimensional conformations. Its utility, however, has been limited by low biological activity of the designed ligands and low tolerance of the protein scaffolds to surface substitutions. Here, we graft the complete binding epitope (19 nonconsecutive amino acids with a solvent-accessible surface area of >2,000 Å^2) of an HIV-1 C-peptide, which is derived from the C-terminal region of HIV-1 gp41 and potently inhibits HIV-1 entry into cells, onto the surface of a GCN4 leucine zipper. The designed peptide, named C34coil, displays a potent antiviral activity approaching that of the native ligand. Moreover, whereas the linear C-peptide is unstructured and sensitive to degradation by proteases, C34coil is well structured, conformationally stable, and exhibits increased resistance to proteolytic degradation compared with the linear peptide. In addition to being a structured antiviral inhibitor, C34coil may also serve as the basis for the development of an alternative class of immunogens. This study demonstrates that ''one-shot'' protein grafting, without subsequent rounds of optimization, can be used to create ligands with structural conformations and improved biomedical properties.

Linear peptides derived from the C-terminal heptad repeat region of HIV-1 gp41 (C-peptides) are promising candidates for inhibiting HIV-1 entry into cells (1, 2), both as a therapeutic (1, 3) and potentially as an immunogen (2, 4). C-peptides bind in a dominant-negative manner to the gp41 N-terminal heptad repeat (N-peptide) region; this binding event prevents formation of the trimer-of-hairpins structure necessary for the fusion of the viral and cellular membranes (ref. 2 and Fig. 1*A*). As therapeutics, C-peptides are potent (nanomolar) inhibitors of HIV-1 infection *in vitro* (5, 6), and clinical trials demonstrate a strong antiviral activity in humans (3, 7). In addition, the gp41 C-peptide region itself is an attractive target for inhibition of viral entry, because several agents [such as 5-Helix (8)] bind to the gp41 C-peptide region and effectively neutralize infection of cells by diverse strains of HIV-1. [Also, the extended epitope of the broadly neutralizing 2F5 antibody overlaps with the C-peptide region (9, 10).] As such, C-peptides can also be potentially used as immunogens to elicit neutralizing antibodies. The feasibility of both uses of C-peptides is underscored by the observation that the amino acid sequences of the region targeted by C-peptides (5, 11, 12), and the C-peptide region itself (8–10, 12), are highly conserved among diverse HIV-1 subtypes.

C-peptides are conformationally unstructured in isolation (13), but are α-helical when bound to their target, the gp41 N-peptide region (11, 14, 15). Also, the gp41 C-peptide region itself is recognized in a helical conformation by a broadly neutralizing agent, 5-Helix (8). Two lines of evidence suggest that the presentation of C-peptides in a preformed, α-helical conformation could increase their biomedical potency. First, linear C-peptides are sensitive to proteolytic degradation and exhibit short half-lives in clinical trials [<2 h using intravenous administration (7)]; in comparison, structured C-peptides are likely to show increased resistance to proteolytic degradation. Second, immunizations of animals with linear peptides corresponding to the C-peptide region (16, 17) have failed to elicit antibodies that neutralize HIV-1 infection; presenting these peptides in a structured helical conformation may be crucial to elicit antibodies with the binding properties of 5-Helix.

We seek to create a fully α-helical and biologically active version of a C-peptide, C34 (6), by protein design. Previous studies (18) of C-peptides using conventional helix-stabilization methods such as chemical cross-linking and amino acid substitutions produced peptides that were not fully helical and substantially less potent than the native C-peptide (19, 20). An alternative strategy is protein grafting, which transfers the critical binding residues of the ligand onto the surface of a stably folded protein (21–28). In previous protein grafting studies of other protein-ligand systems (24, 26), the designed peptides had impressive biological activity given the grafting of only a small number of residues, but their binding affinities were still significantly less than those of the native ligands. The grafting of a large number of key binding residues generated peptides with strong binding activity, but disrupted the native protein scaffold (24, 25, 27). Substantial optimization was needed after the initial grafting to either rescue the scaffold structure (27), or to produce a potent ligand (29). Therefore, an important challenge is to create ligands that exhibit a high biological activity, and also retain the stable, native fold of the scaffold protein.

Materials and Methods

Peptide Design. The sequences of the peptides are given in Fig. 1*D*. The grafted epitope consists of 19 residues of C34 in the **a**′, **d**′, **e**′, and **g**′ positions (W628, W631, D632, E634, I635, Y638, T639, L641, I642, L645, I646, E648, S649, Q652, Q653, K655, N656, E659, and L660). The residues that are replaced in GCN4 are 14 residues in the **b**, **c**, and **f** positions (K3, Q4, D7, E10, E11, S14, Y17, H18, N21, A24, R25, K28, G31, and E32), with 4 residues in the **e** position (E6, L13, E20, and K27) overlapping with the 4 residues in the **g**′ position of C34. The residues that are retained in GCN4 are 15 residues in the **a**, **d**, **e**, and **g** positions (R1, M2, L5, K8, V9, L12, K15, N16, L19, E22, V23, L26, L29, V30, and R33), with the 4 residues in the **e** position.

Peptide Synthesis and Purification. Peptides were synthesized by standard fluorenylmethoxycarbonyl chemistry with acetylated N termini and either free or amidated C termini (sequences shown in Fig. 1*D*). The identities of all peptides were confirmed by (Perceptive Biosystems Voyager Elite) matrix-assisted laser desorption ionization MS (MALDI-MS) to be within 0.1% of the predicted masses. Reverse-phase HPLC (C18 column; Vydac, Hesperia, CA) was used to purify the peptides. To synthesize the heterodimer, approximately equimolar amounts of the (Cys)C34-GCN4 and (Cys)GCN4 peptides were mixed in 6 M guanidine hydrochloride and 100 mM Tris, pH 8.0. After 24 h of air oxidation, the heterodimer was purified by reverse-phase HPLC. The purity and molecular weight of the heterodimer were verified by liquid chromatography-MS (LC-MS) (Finnigan-

Abbreviations: C-peptides, peptides corresponding to the C-terminal heptad repeat of HIV-1 gp41; N-peptides, peptides corresponding to the N-terminal heptad repeat of HIV-1 gp41.

‡Present address: Department of Chemistry and Chemical Biology, 12 Oxford Street, Cambridge, MA 02138.

§To whom correspondence should be sent at the present address: Merck Research Laboratories, 770 Sumneytown Pike, West Point, PA 19486. E-mail: peter_kim@merck.com.

www.pnas.org/cgi/doi/10.1073/pnas.1733910100

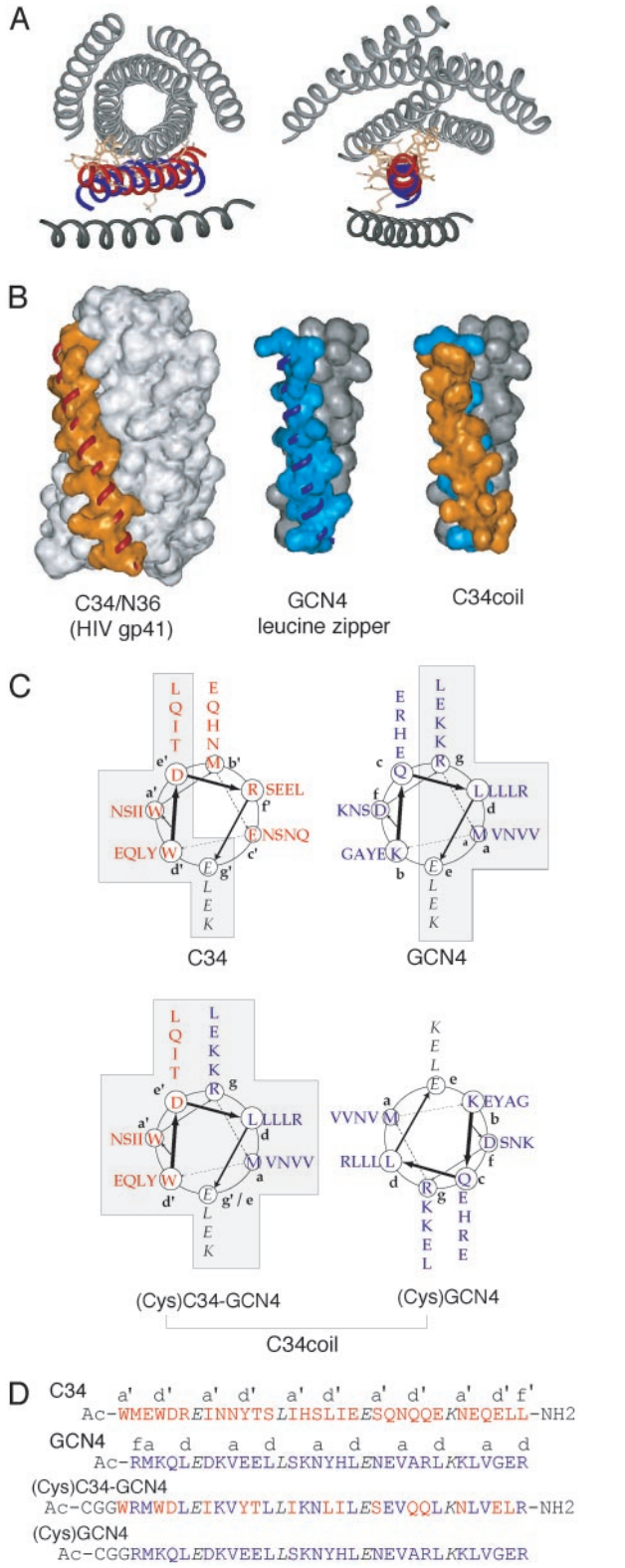

Fig. 1. Design of C34coil. (*A*) Superpositions of one C34 helix from the HIV-1 gp41 core (red) and one helix from the GCN4 leucine zipper (blue). The rest of gp41 is light gray, and the rest of GCN4 is dark gray. Shown are views down the 6-helix bundle of gp41 (*Left*) and down the superimposed helices (*Right*). The main-chain atoms of C34 (residues 632–657) and GCN4 (residues 4–29) superimpose to 1.2-Å rms deviation. (For the whole helix, from residues 629–659 of C34 and residues 1–31 of GCN4, the rms deviation is 1.7 Å due to helix fraying at the ends.) (*B*) Surface diagrams of the 6-helix bundle of HIV-1 gp41 with one C34 helix shown in ribbon format (backbone of C34 helix in red; *Left*) the dimeric GCN4 leucine zipper with one GCN4 helix shown in ribbon format (backbone of GCN4 helix in dark blue; *Center*), and a model of C34coil (*Right*). Both the C34 binding epitope (orange) and the GCN4 hydrophobic core (light blue) are incorporated into C34coil. The solvent-accessible surface area of the C34 binding epitope is 2,061 Å^2 (see *Materials and Methods* for calculations). (*C*) Helical wheel representations of C34 (red), GCN4 (blue), and C34coil (red and blue). The four residues at the **g**′ position of C34 and the four residues at the e position of GCN4 are identical and are shown in italics. C34coil is a covalent heterodimer of two peptides, (Cys)C34-GCN4 and (Cys)GCN4, connected by means of a disulfide bond [(Cys) denotes the addition of a Cys-Gly-Gly sequence to the N terminus of the peptide]. The boxed residues of C34 (at **a**′, **d**′, **e**′, and **g**′ positions) and GCN4 (at **a**, **d**, **e**, and **g** positions) are incorporated into the C34-GCN4 peptide and are equivalent to the binding epitope highlighted in orange and blue, respectively, in *B*. (*D*) Sequences of the peptides C34 (red), GCN4 (blue), (Cys)C34-GCN4 (red and blue), and (Cys)GCN4 (blue). The four residues in common in all peptides are shown in italics. The C34-GCN4 peptide, which lacks the N-terminal Cys-Gly-Gly, was also synthesized. Ac, an acetylated N terminus; NH2, an amidated C terminus. Note that the GCN4 sequence used in this study corresponds to GCN4-p1 in previous studies (32). See *Materials and Methods* for more details on the peptide design.

MAT, San Jose, CA; LCQ). Peptide concentrations were determined by absorbance at 280 nm in 6 M guanidine hydrochloride, using extinction coefficients of 5,690 $M^{-1}\cdot cm^{-1}$ for tryptophan, 1,280 $M^{-1}\cdot cm^{-1}$ for tyrosine, and 120 $M^{-1}\cdot cm^{-1}$ for cystine (30).

CD Spectroscopy and Guanidine Denaturation. CD spectra were measured on an Aviv 60DS spectrapolarimeter at 25°C in PBS buffer (10 mM sodium phosphate/150 mM NaCl, pH 7.0). The peptide concentrations used were: 20 μM C34, 16 μM C34-GCN4, and 1 μM C34coil. For chemical denaturation experiments, the $[\theta]_{222}$ values of C34coil (1 μM) were measured as a function of guanidine hydrochloride concentration in PBS. Guanidine concentrations were measured by refractometry. The titrations were fitted to the standard six-parameter two-state transition equation by using weighted averages to yield standard folding free energy (31).

Gel Filtration Chromatography. Gel filtration chromatography was performed on a TosoHaas G4000SWXL column at room temperature and monitored at 280 nm, using PBS as the running buffer. C34coil (25 μl and ≈100 μM) was injected into the column running at 1 ml/min. A Bio-Rad gel filtration standard was run as a molecular weight standard.

Proteolysis. Proteinase K (Sigma) was added to the peptides (at concentrations of either 10 or 20 μM) in PBS and incubated at 37°C. For the indicated relative ratios of proteinase K digestion, the amounts of proteinase K and incubation times were: 1× for 0.002 μg/ml for 1 min, 10× for 0.02 μg/ml for 1 min, 600× for 0.2 μg/ml for 60 min, and 10,000× for 0.2 μg/ml for 17 h. After incubation, the reactions were quenched by bringing the solutions to a final concentration of 1 mM PMSF and 5% acetic acid. The reaction products were run on a Microsorb-MV C18 column at a speed of 1 ml/min and a gradient of 1% acetonitrile per min.

Cell–Cell Fusion and Viral Infectivity Assays. In both inhibition assays, peptide stocks were dissolved in dimethyl sulfoxide, and their concentrations were determined by absorbance in guanidine at 280 nm. The final concentration of dimethyl sulfoxide in tissue culture media was 1% in all experiments. Inhibition of cell–cell fusion (syncytia formation) was performed as described, with a 20-h incubation (6). Inhibition of viral infectivity was performed as described (6). In both assays, the IC_{50}, the concentration needed to inhibit 50% of cell–cell fusion events or luciferase activity, was calculated from fitting the data to the equation, $y = k/(1 + [\text{peptide}]/IC_{50})$, where y = number of syncytia or luciferase activity and k is a scaling constant.

Calculations of Solvent-Accessible Surface Areas. Solvent-accessible surface areas were calculated by using the program INSIGHT II (Accelrys), with a probe radius of 1.4 Å. The coordinates of C34 were taken from ref. 11, and those of GCN4 were taken from ref. 32. For C34, the coordinates of only one of the C-peptides were used (without the N-peptides). For an equivalent comparison between C34 and GCN4, the 17 residues present in both structures were used for the calculations (**a**′, **d**′, **e**′, and **g**′ positions of C34 from W631 to E659, and **b**, **c**, **e**, and **f** positions of GCN4 from K3 to G31).

Results

Design Principles. We constructed a helical, stable, and biologically potent C-peptide by grafting all 19 amino acids of the C34 binding epitope onto the surface of a GCN4 leucine zipper, a stable, homodimeric coiled coil (ref. 32 and Fig. 1*A* and *B*). The coiled coil is a good candidate to act as a scaffold protein, because it is a stable and protease-resistant structure (13, 33, 34), and helix bundles can effectively present binding epitopes on their surfaces (24, 25, 35). To design the sequence of the helical C-peptide, we observe that in coiled coils, the residues forming the hydrophobic core are most critical for formation of the helical structure, whereas the solvent-exposed residues are less critical for stability or dimerization (33, 36). Therefore, the 19 amino acids in the hydrophobic core of GCN4 (residues shown in blue in Fig. 1*B* and boxed in Fig. 1*C*) were retained in the designed peptide. Next, all 19 residues in the epitope of C34 for binding to the N-peptide region of gp41 (residues shown in orange in Fig. 1*B* and boxed in Fig. 1*C*) were transferred onto the solvent-exposed face of the GCN4 helix. Fortuitously, the 4 amino acids at positions **g**′ of C34 and **e** of GCN4 [for an amino acid sequence with a heptad repeat of the form $(\mathbf{abcdefg})_n$] (residues shown in italics in Fig. 1 *C* and *D*) are identical. The designed peptide, denoted C34-GCN4, is 34 residues in length: 19 from C34, 19 from GCN4, with 4 overlapping residues.

The GCN4 leucine zipper was chosen as a particularly suitable scaffold protein in this study for several reasons. First, the determinants of the structure and stability of the GCN4 leucine zipper are well characterized. In particular, like other coiled coils (33, 36), the overall helical structure of GCN4 is tolerant to large-scale substitutions at the solvent-exposed residues (at positions **b**, **c**, and **f**), and the effects of many specific mutations of GCN4 are known. This knowledge allows one to design the grafting in a way as to minimize disturbance to the coiled-coil structure while maximizing the amount of replaced residues. Second, compared with other coiled coils, the GCN4 leucine zipper is only moderately stable, with a midpoint transition temperature of 57°C (37). Therefore, on binding to the gp41 target, the GCN4 coiled coil can potentially adjust its helical structure to approach the optimal conformation for the grafted C34 residues. Third, GCN4 is uniquely suited for the grafting of C34 binding residues because the four amino acids at position **e** of GCN4 and **g**′ of C34 are identical (residues shown in italics in Fig. 1 *C* and *D*). The electrostatic interactions in the **e**/**g** positions are important for maintaining the structure (32, 38) and stability (39) of the GCN4 coiled coil, and they may have been important in a previous grafting study (24), where the helical structure of the GCN4 scaffold was disrupted on substitutions at the **e**/**g** positions. This fortuitous match allows both the complete core of GCN4 and the complete binding epitope of C34 to be recapitulated in the designed peptide. Most other helix bundles and coiled coils would not meet this important criterion due to the low probability that all residues at an **e**/**g** position would match to those of C34. Fourth, the solvent-accessible surface area of the C34 epitope closely matches that of the region on GCN4 available for grafting (in an equivalent comparison, ≈1,700 Å^2 for C34 and 1,800 Å^2 for GCN4; see *Materials and Methods*), thereby minimizing disruption to both the scaffold and the binding epitope in the grafted protein. Fifth, both GCN4 and C34 contain four full heptad repeats. The matching length ensures that the full set of C34 binding residues can be grafted, but that the grafted protein does not contain other regions that interfere with binding to the target. Sixth, the backbone rms deviation between GCN4, a canonical coiled coil (32), and C34, which does not exhibit knobs-into-holes packing typical of coiled coils (11), is reasonable but not extremely close (1.2 Å for most of the helix; see the Fig. 1*A* legend). Superposition of crystal structures of the two helices (Fig. 1*A*) shows that the deviation results mostly from a difference in the supercoil. Also, the pitch of α-helices in coiled coils such as GCN4 is generally larger than those of noncoiled coils (40).

Characterization of the C34-GCN4 Peptide and C34coil. The C34-GCN4 peptide is very hydrophobic, and as a result, exhibits low solubility (<1 μM in PBS). The CD spectrum of the C34-GCN4 peptide (in a low amount of guanidine hydrochloride to increase peptide solubility) indicates a mostly random coil structure (Fig. 2*A*).

To overcome the poor helical content and solubility of the C34-GCN4 peptide, we made a disulfide-linked heterodimer, denoted C34coil, composed of two peptides: (Cys)C34-GCN4 and (Cys)GCN4 (Fig. 1*C*). Compared with the C34-GCN4 peptide, C34coil is expected to be more soluble because the second helix of GCN4 contains many hydrophilic amino acids, and to have a more stable helical structure because the disulfide linkage stabilizes the helical conformation of coiled-coil dimers (37). Indeed, C34coil exhibits enhanced solubility (up to 100 μM in PBS) and a helical content of >90% (Fig. 2*A*). Moreover, the helical structure of C34coil is stable, as shown by chemical denaturation (Fig. 2*B*). Importantly, C34coil is monomeric and exhibits no aggregation under physiological conditions, as determined by gel filtration chromatography (Fig. 2*C*). Thus, C34coil exhibits a stable, helical structure, despite substantial mutations to the surface of the GCN4 leucine zipper, and the preformed helical conformation of the C34 binding epitope does not induce aggregation.

To test how well the HIV-1-neutralizing epitope was recapitulated in the preformed α-helical structure of C34coil, two assays were used to measure the potency of the designed peptides in inhibiting HIV-1 envelope-mediated membrane fusion. In both the cell–cell fusion (Fig. 3*A*) and viral infectivity (Fig. 3*B*) assays, C34coil exhibits potent antiviral activities, with IC_{50} values of 3 nM and 16 nM, respectively, which are within an order of magnitude of those of C34 (6). In addition, the GCN4 homodimer, a covalent homodimer of (Cys)GCN4 (Fig. 1*D*), shows no activity in both assays up to 50 μM, confirming that the C34-binding epitope, and not the GCN4 portion, is responsible for the inhibitory activity of C34coil. The potent inhibitory activity of C34coil demonstrates that the helical C34 epitope is presented in a biologically active conformation. Interestingly, C34-GCN4 also potently inhibits HIV-1 entry (Fig. 3).

The stable, helical structure of C34coil is expected to give rise to an increased resistance to proteolytic degradation compared with linear peptides, because proteases are most active on unfolded substrates. We compared the sensitivities to proteolytic degradation of C34coil and the linear peptide C34 by digestion with proteinase K, a protease of broad substrate specificity, and analyzed the reaction products by reverse-phase HPLC. C34coil is ≈1,000-fold more resistant to degradation by proteinase K than by C34 (Fig. 2*D*). Because of its increased resistance to proteolytic degradation and potent inhibitory activity, C34coil or its variants may be useful therapeutics for HIV-1 infection in addition to the C-peptides being currently tested (7).

Finally, we characterized a variant of C34coil, named

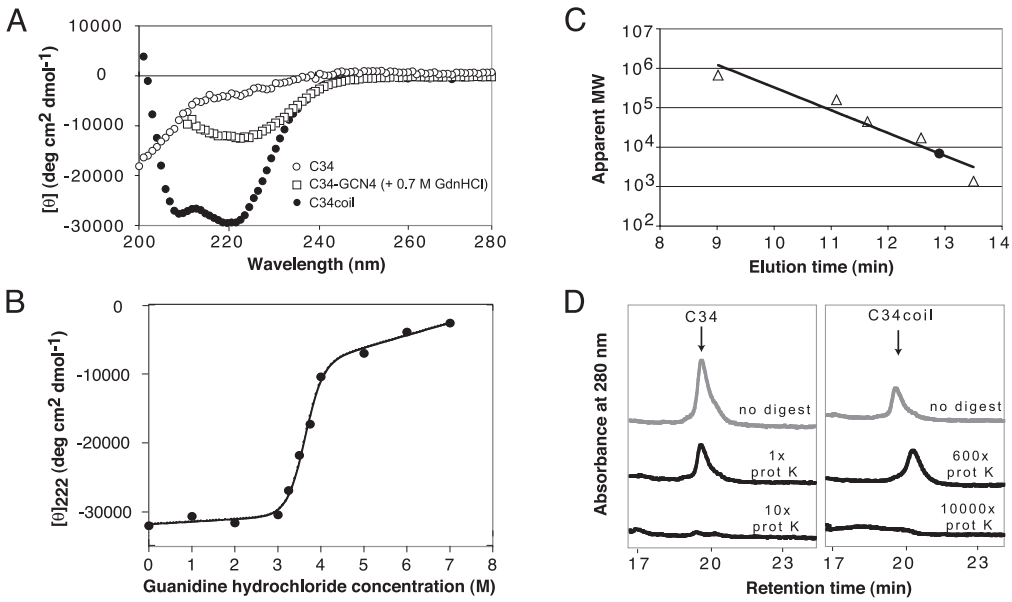

Fig. 2. Biophysical characterization of C34coil. (*A*) CD spectrum of C34, C34-GCN4, and C34coil. The guanidine hydrochloride used to solubilize C34-GCN4 precludes measurements at wavelengths <210 nm. Experimental conditions were: PBS at pH 7.0 at 25°C. C34coil exhibits a helical structure ($[\theta]_{222}$ value of −29,400 deg cm^2·dmol^{-1}), whereas C34 (−3,800 deg cm^2·dmol^{-1}) and the C34-GCN4 peptide (−12,400 deg cm^2·dmol^{-1}) exhibit relatively unstructured conformations. (*B*) Guanidine hydrochloride denaturation of C34coil, as monitored by CD spectroscopy at 222 nm. The fitted curve to a two-state unfolding transition is shown as a black line. C34coil unfolds at a midpoint of 3.6 M guanidine hydrochloride, with a free energy of unfolding of 9.3 kcal/mol. (*C*) Apparent molecular weight (MW) of C34coil, as determined by gel filtration chromatography. Shown are the elution times of molecular weight standards (△) and C34coil (●). Also shown is the best-fit line of apparent MW versus elution time from the molecular weight standards (black line). C34coil exhibits an apparent molecular mass of 7,000 ± 2,000 Da (expected molecular mass of 8,740 Da for a monomer), with no detectable aggregation. (*D*) Sensitivity to proteolytic degradation of C34coil and C34. Peptides are incubated with proteinase K at 37°C, and the reaction products are monitored by reverse-phase HPLC. The relative ratios of proteinase K account for differences in both the protease concentrations and incubation times. Shown are the HPLC chromatograms. C34coil is ≈1,000-fold more resistant to degradation by proteinase K than by C34.

C34coil-N16K, with a mutation of Asn-16 to Lys-16. Compared with C34coil, C34coil-N16K exhibits a less stable helical structure, and inhibits HIV-1 entry with a slightly greater potency (Fig. 4).

Discussion

Protein Grafting. We demonstrated a simple and powerful protein grafting strategy to stabilize an anti-HIV-1 epitope into a structured and functional conformation. Previous successful protein grafting studies (25, 27), where potent binders were generated and the structural integrity of the scaffold was intact, featured substantial optimization after the initial grafting [either phage display or multiple iterations of point mutations (26, 29)]. By comparison, our study uses a simple one-shot protein grafting strategy, which results in a nanomolar binder. Several factors contributed to a successful one-shot protein grafting. First, we minimize disruption of the scaffold structure by retaining all of the key determinants of its stability (**a**, **d**, **e**, and **g** positions for GCN4). Second, we maximize the effectiveness of the grafted epitope by transferring all of the key binding residues of C34. For gp41, transferring such a large number of residues may be particularly important because interactions along the entire interface between the C-peptides and N-peptides are important for the stability of the 6-helix bundle (41). Third, the correspondence between the scaffold and ligand was sufficiently close in structural parameters (accessible surface area and backbone rms deviation). Finally, the conformational flexibility of the scaffold (because the midpoint unfolding temperature of GCN4 is not very high) allows for adjustment of the C34coil helical structure on binding to the gp41 target.

In terms of scale, this study demonstrates that a large number of residues can be successfully grafted onto a scaffold. In particular, 19 residues were successfully grafted in this study [in previous studies, 10 (26) and 8 (24) residues were grafted; in one study, 17 residues were grafted, but the scaffold was unfolded (25, 27)]. The accessible surface area of the grafted anti-HIV-1 epitope is 2,061 Å^2, or 55% of the total surface area of C34. Compared with an average protein–protein interface, where each subunit buries ≈600–1,000 Å^2, comprising 5–20% of its total surface area (42), this study suggests that size alone is likely not a limitation in grafting most protein–protein interfaces.

The binding of C-peptides to their target involves a delicate balance between the enthalpy and entropy of binding (43). Despite presenting the full binding epitope in a helical conformation, C34coil is a slightly less potent inhibitor than the linear peptide C34. Thus, the entropic gain in binding energy from a preformed helical scaffold may be offset by the energy required to change the C34-binding epitope from a coiled-coil conformation (32, 33) to the active helical conformation (11, 14, 15) (the backbone rms deviation between the two conformations, 1.2 Å in the main part of the helix, is reasonably but not extremely close; Fig. 1*A*). To test this hypothesis, we synthesized C34coil-N16K with a mutation of Asn-16 to Lys-16, which introduces greater conformational flexibility to the helical structure. In the GCN4 leucine zipper, the N16K mutation lowered the stability of the coiled coil while preserving a dimeric oligomerization state (44). Indeed, compared with C34coil, C34coil-N16K exhibits cooperative unfolding, but with lower stability, and inhibits HIV-1 entry with slightly greater potency (Fig. 4).

Overall, the constraining method of choice depends on the end goal. Loose covalent cross-linkers may be used to improve the potency of the ligand but not enforce a rigid structure (43), whereas protein grafting may be used to enforce a rigid structure but at the cost of introducing an entropic barrier to

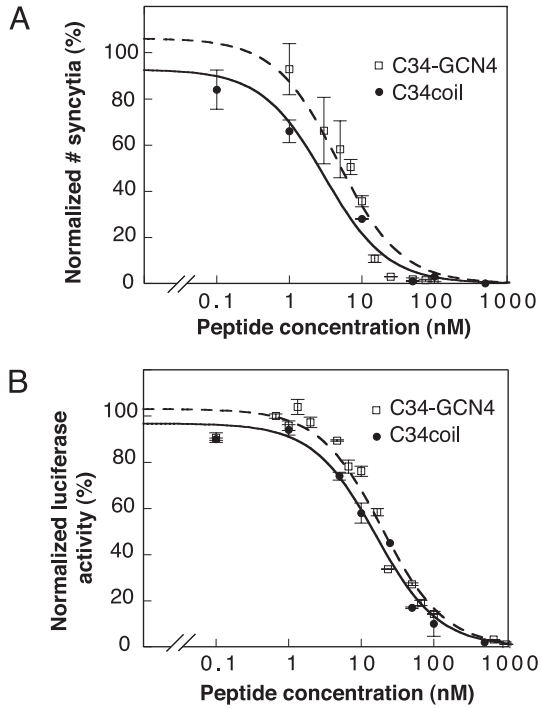

Fig. 3. Inhibition of HIV-1 envelope-mediated membrane fusion. The data represent the mean ± SE of at least two separate experiments. (*A*) Inhibition of cell–cell fusion. The peptides were tested for inhibiting the fusion of HIV-1 Env-expressing cells (CHO gp160) with CD4-expressing cells (CD4 HeLa). The IC_{50} values for C34coil and C34-GCN4 are 3.1 ± 0.8 and 4.6 ± 0.9 nM, respectively. The (Cys)GCN4-homodimer shows no inhibitory activity up to 50 μM. The IC_{50} value of C34 in the cell–cell fusion assay, as reported in ref. 6, was 0.6 nM. (*B*) Inhibition of viral infectivity. The peptides were tested for inhibition of CD4-positive target cells (HOST4) by recombinant, luciferase-expressing HIV-1. The IC_{50} values for C34 coil and C34-GCN4 are 16 ± 2 and 19 ± 3 nM, respectively. The (Cys)GCN4-homodimer shows no inhibitory activity up to 50 μM. The IC_{50} value of C34 in the viral infectivity assay, as reported in ref. 6, was 2 nM.

attaining the optimal conformation of the grafted epitope. For protein grafting, the GCN4 coiled coil is a suitable protein scaffold, because it is sufficiently stable to retain its helical conformation, but not overly stable as to lock the binding epitope into an inactive conformation. GCN4 is also useful in that its helical structure is tolerant of large-scale surface substitutions as long as the key determinants of coiled-coil stability are retained; this property is consistent with previous studies that showed the native fold of a protein can tolerate large-scale substitutions if the substitutions are appropriately chosen (45). Finally, the GCN4 coiled coil can be mutated further to optimize conformational or other properties without losing its ability to serve as a grafting scaffold, as shown by the C34coil-N16K peptide.

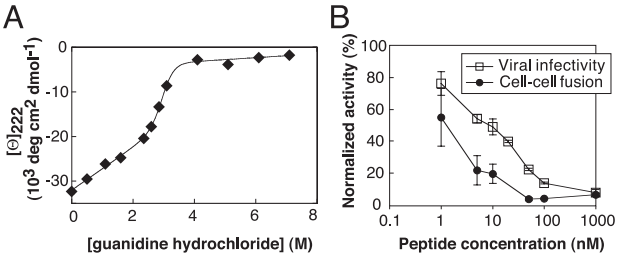

Fig. 4. Characterization of C34coil-N16K. (*A*) Unfolding of C34coil-N16K by guanidine hydrochloride as measured by CD spectroscopy. The midpoint of unfolding occurs at 3.0 M guanidine, with a free energy of unfolding of 9.0 kcal/mol. (*B*) Inhibition of HIV-1 entry by C34coil-N16K, as measured by the cell–cell fusion and viral infectivity assays. Standard deviations are shown as error bars. The IC_{50} values are 1.5 nM in the cell–cell fusion assay and 10 nM in the viral infectivity assay.

Mutability of C34 Region. The potent antiviral activity of C34coil suggests that residues at the solvent-exposed face of the C-peptide helix (opposite from the face that binds to gp41, at **b**, **c**, and **f** positions of the heptad repeat) are not crucial for activity. Consistent with this observation, the unstructured C34-GCN4 peptide, which contains all of the key binding residues of C34, but features completely different residues at the solvent-exposed positions, also shows strong inhibitory potency (Fig. 3). The remarkable tolerance of C34 to amino acid substitutions suggests that significant mutations (for helix stabilization, for example) may be used in the future to create an even more potent linear C-peptide than C34 or T-20.

Future Use and Design. C34coil may be an effective immunogen for eliciting antibodies that recognize the C-peptide region of gp41 in a helical conformation and block membrane fusion (8). Moreover, anti-C34coil antibodies would be expected to recognize only the highly conserved binding epitope of C34; in comparison, antibodies against linear epitopes would recognize both conserved and nonconserved residues (8), which may increase the probability of the emergence of escape mutants. C34coil could also be fused to a 2F5 epitope, which is located immediately C-terminal to and may even overlap with the C34 sequence (10), constrained into a β-turn to form an extended conformationally stabilized epitope [the conformation of the epitope has been shown to be important for recognition by the 2F5 antibody (46, 47)].

Future designs of C34coil can be modified in several ways. A scaffold with a lower backbone rms deviation from C34 may result in a more potent binder; for this purpose, a GCN4-derived tetramer may be appropriate because it exhibits less supercoil than the GCN4 dimer, and it retains the advantages of a GCN4-based system (such as the matching **e**/**g** residues). Also, the current design requires an oxidation and repurification step of two separate peptides; the use of antiparallel coiled coils (48) would circumvent this requirement. Finally, this protein-grafting strategy can be extended to other viruses, such as paramyxovirus (49), parainfluenza (50), and Sendai virus (51), for which analogous C-peptides have been shown to inhibit membrane fusion.

The presentation of protein sequences in preformed, active three-dimensional conformations is a potentially powerful strategy for improving their biomedical properties. In particular, linear peptides and recombinant proteins that are attractive candidates for therapeutics and immunogens often exhibit nonproductive three-dimensional conformations, which can limit their immunogenicity, antigenicity, or binding affinity toward the target (see ref. 52 for HIV-1 immunogens). We demonstrate that protein grafting of a large binding epitope can be performed in a way that exhibits both a potent biological activity and a favorable three-dimensional structure, without subsequent optimization of the grafted protein sequence. The utility of the protein-grafting strategy can be augmented by phage display (27) or structure-based analysis (29) to improve the potency or stability of the designed peptide.

We thank Ben Sanford and Michael Burgess for peptide synthesis; Heng Chhay for help with the viral infectivity assays; Leslie Gaffney for editorial assistance; and Drs. Benjamin Chen, Michael Root, and Amy Keating and the Kim laboratory for helpful suggestions on the manuscript. S.K.S. acknowledges financial support from a Howard Hughes Medical Institute Predoctoral Fellowship and National Science and Engineering Council of Canada Postgraduate Scholarships. This work was supported in part by National Institutes of Health Grant P01 GM56552.

1. D'Souza, M. P., Cairns, J. S. & Plaeger, S. F. (2000) *J. Am. Med. Assoc.* **284,** 215–222.
2. Eckert, D. M. & Kim, P. S. (2001) *Annu. Rev. Biochem.* **70,** 777–810.
3. Moore, J. P. & Stevenson, M. (2000) *Nat. Rev. Mol. Cell Biol.* **1,** 40–49.
4. Nabel, G. J. (2001) *Nature* **410,** 1002–1007.
5. Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B. & Matthews, T. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9770–9774.
6. Chan, D. C., Chutkowski, C. T. & Kim, P. S. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 15613–15617.
7. Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., *et al.* (1998) *Nat. Med.* **4,** 1302–1307.
8. Root, M. J., Kay, M. S. & Kim, P. S. (2001) *Science* **291,** 884–888.
9. Muster, T., Steindl, F., Purtscher, M., Trkola, A., Klima, A., Himmler, G., Rüker, F. & Katinger, H. (1993) *J. Virol.* **67,** 6642–6647.
10. Parker, C. E., Deterding, L. J., Hager-Braun, C., Binley, J. M., Schülke, N., Katinger, H., Moore, J. P. & Tomer, K. B. (2001) *J. Virol.* **75,** 10906–10911.
11. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. (1997) *Cell* **89,** 263–273.
12. Hanna, S. L., Yang, C., Owen, S. M. & Lal, R. B. (2002) *AIDS* **16,** 1603–1608.
13. Lu, M., Blacklow, S. C. & Kim, P. S. (1995) *Nat. Struct. Biol.* **2,** 1075–1082.
14. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1997) *Nature* **387,** 426–430.
15. Tan, K., Liu, J., Wang, J., Shen, S. & Lu, M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12303–12308.
16. Rosny, E., Vassell, R., Wingfield, P. T., Wild, C. T. & Weiss, C. D. (2001) *J. Virol.* **75,** 8859–8863.
17. Muster, T., Guinea, R., Trkola, A., Purtscher, M., Klima, A., Steindl, F., Palese, P. & Katinger, H. (1994) *J. Virol.* **68,** 4031–4034.
18. Andrews, M. J. I. & Tabor, A. B. (1999) *Tetrahedron* **55,** 11711–11743.
19. Judice, J. K., Tom, J. Y. K., Huang, W., Wrin, T., Vennari, J., Petropoulos, C. J. & McDowell, R. S. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13426–13430.
20. Jin, B. S., Ryu, J. R., Ahn, K. & Yu, Y. G. (2000) *AIDS Res. Hum. Retroviruses* **16,** 1797–1804.
21. Nygren, P. & Uhlén, M. (1997) *Curr. Opin. Struct. Biol.* **7,** 463–469.
22. Fernandez-Carneado, J., Grell, D., Durieux, P., Hauert, J., Kovacsovics, T. & Tuchscher, G. (2000) *Biopolymers* **55,** 451–458.
23. Cochran, A. G. (2000) *Chem. Biol.* **7,** R85–R94.
24. Domingues, H., Cregut, D., Sebald, W., Oschkinat, H. & Serrano, L. (1999) *Nat. Struct. Biol.* **6,** 652–656.
25. Zondlo, N. J. & Schepartz, A. (1999) *J. Am. Chem. Soc.* **121,** 6938–6939.
26. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Ménez, A., Roumestand, C., Yang, Y. S., Ylistigui, L., Benjouad, A. & Gluckman, J. C. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13091–13096.
27. Chin, J. W. & Schepartz, A. (2001) *J. Am. Chem. Soc.* **123,** 2929–2930.
28. Montclare, J. K. & Schepartz, A. (2003) *J. Am. Chem. Soc.* **125,** 3416–3417.
29. Martin, L., Stricher, F., Misse, D., Sironi, F., Pugniere, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., *et al.* (2003) *Nat. Biotech.* **21,** 71–76.
30. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954.
31. Pace, C. N. (1986) *Methods Enzymol.* **131,** 266–280.
32. O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991) *Science* **254,** 539–544.
33. Cohen, C. & Parry, D. A. D. (1990) *Proteins* **7,** 1–15.
34. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994) *Nature* **371,** 37–43.
35. Ghirlanda, G., Lear, J. D., Lombardi, A. & DeGrado, W. F. (1998) *J. Mol. Biol.* **281,** 379–391.
36. O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1992) *Cell* **68,** 699–709.
37. O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1989) *Science* **243,** 538–542.
38. Zeng, X., Zhu, H., Lashuel, H. A. & Hu, J. C. (1997) *Protein Sci.* **6,** 2218–2226.
39. Spek, E. J., Bui, A. H., Lu, M. & Kallenbach, N. R. (1998) *Protein Sci.* **7,** 2431–2437.
40. Seo, J. & Cohen, C. (1993) *Proteins* **15,** 223–234.
41. Lu, M., Stoller, M. O., Wang, S., Liu, J., Fagan, M. B. & Nunberg, J. H. (2001) *J. Virol.* **75,** 11146–11156.
42. Stites, W. E. (1997) *Chem. Rev.* **97,** 1233–1250.
43. Sia, S. K., Carr, P. A., Cochran, A. G., Malashkevich, V. N. & Kim, P. S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 14664–14669.
44. Gonzalez, L., Woolfson, D. N. & Alber, T. (1996) *Nat. Struct. Biol.* **3,** 1011–1018.
45. Dalal, S., Balasubramanian, S. & Regan, L. (1997) *Nat. Struct. Biol.* **4,** 548–552.
46. Pai, E. F., Klein, M. H. & Pedyczak, A. (2000) in *World Intellectual Property Organization*, Patent Cooperation Treaty WO 00/61618.
47. Ho, J., MacDonald, K. S. & Barber, B. H. (2002) *Vaccine* **20,** 1169–1180.
48. Oakley, M. G. & Hollenbeck, J. J. (2001) *Curr. Opin. Struct. Biol.* **11,** 450–457.
49. Lambert, D. M., Barney, S., Lambert, A. L., Guthrie, K., Medinas, R., Davis, D. E., Bucy, T., Erickson, J., Merutka, G. & Petteway, S. R., Jr. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2186–2191.
50. Yao, Q. & Compans, R. W. (1996) *Virology* **223,** 103–112.
51. Rapaport, D., Ovadia, M. & Shai, Y. (1995) *EMBO J.* **14,** 5524–5531.
52. Burton, D. & Moore, J. P. (1998) *Nat. Med.* **4,** 495–498.