Cell, Vol. 99, 103–115, October 1, 1999, Copyright ©1999 by Cell Press

Inhibiting HIV-1 Entry: Discovery of D-Peptide Inhibitors that Target the gp41 Coiled-Coil Pocket

Debra M. Eckert, Vladimir N. Malashkevich, Lily H. Hong, Peter A. Carr, and Peter S. Kim*
Howard Hughes Medical Institute
Whitehead Institute for Biomedical Research
Department of Biology
Massachusetts Institute of Technology
Nine Cambridge Center
Cambridge, Massachusetts 02142

Summary

The HIV-1 gp41 protein promotes viral entry by mediating the fusion of viral and cellular membranes. A prominent pocket on the surface of a central trimeric coiled coil within gp41 was previously identified as a potential target for drugs that inhibit HIV-1 entry. We designed a peptide, IQN17, which properly presents this pocket. Utilizing IQN17 and mirror-image phage display, we identified cyclic, D-peptide inhibitors of HIV-1 infection that share a sequence motif. A 1.5 Å cocrystal structure of IQN17 in complex with a D-peptide, and NMR studies, show that conserved residues of these inhibitors make intimate contact with the gp41 pocket. Our studies validate the pocket per se as a target for drug development. IQN17 and these D-peptide inhibitors are likely to be useful for development and identification of a new class of orally bioavailable anti-HIV drugs.

Introduction

Currently, anti-HIV-1 combination drug therapy only targets HIV-1 protease and reverse transcriptase. Recently, there has been an influx of structural and functional information on the HIV-1 membrane fusion process (Chan and Kim, 1998; Kwong et al., 1998; Berger et al., 1999; and references therein). These studies have raised hopes that viral entry is another part of the HIV-1 life cycle that can be successfully inhibited.

Infection of cells by HIV-1 requires fusion of the cellular and viral membranes, a process mediated by the viral envelope glycoprotein complex (gp120/gp41) and cell surface receptors on the target cell. A working model for this entry process (Figure 1) involves multiple steps that have been delineated through numerous studies (for review, Chan and Kim, 1998). Binding of gp120/gp41 to cellular receptors (CD4 and a chemokine coreceptor such as CCR-5 or CXCR-4) induces a conformational change in the envelope glycoprotein. A transient species results, termed the prehairpin intermediate, in which gp41 exists as a membrane protein simultaneously in both the viral and cellular membranes (Chan and Kim, 1998; Furuta et al., 1998; Jones et al., 1998; Munoz-Barroso et al., 1998). The prehairpin intermediate resolves to a trimer-of-hairpins structure that likely represents the fusion-active state of gp41 (Blacklow et al., 1995; Lu et al., 1995) as seen in the X-ray crystal structure of a protease-resistant core of gp41 (Chan et al., 1997; Tan et al., 1997; Weissenhorn et al., 1997). It is unclear whether hairpin formation occurs before, or simultaneously with, the actual fusion of the two bilayers.

The trimer-of-hairpins structure is a common feature of diverse viral membrane fusion proteins (Singh et al., 1999, and references therein). In gp41, a central three-stranded coiled coil (formed by the N-terminal regions of gp41) is surrounded by helices derived from the C-terminal end of the gp41 ectodomains, packed in an antiparallel manner around the outside of the coiled coil (Figure 1, inset). Peptides corresponding to these regions of gp41 are referred to as N-peptides and C-peptides, respectively.

Synthetic C-peptides are potent inhibitors of HIV-1 infection. These inhibitors, such as C34 or DP178 (see Figure 1 legend), inhibit HIV-1 infection and syncytia formation at nanomolar concentrations in cell culture experiments (Jiang et al., 1993; Wild et al., 1994; Lu et al., 1995; Chan et al., 1998; Rimsky et al., 1998). There is substantial evidence to indicate (see Discussion) that C-peptides act in a dominant-negative manner (Herskowitz, 1987) by binding to the transiently exposed coiled-coil N-peptide region in the prehairpin intermediate (Figure 1).

A phase I clinical trial showed that DP178 (recently renamed T-20) has an antiviral effect when injected into infected individuals (Kilby et al., 1998). This result provides compelling evidence that the prehairpin intermediate is a useful target for anti-HIV therapy in humans. However, very large amounts of T-20 (~200 mg/day) are required to observe an antiviral effect, whereas nanomolar concentrations are efficacious in cell culture. Thus, bioavailability issues, presumably including rapid proteolytic degradation, need to be addressed before C-peptides can be widely used as therapeutic agents.

It would be highly desirable if an orally bioavailable drug could be identified that prevents HIV-1 entry by binding to the prehairpin intermediate of gp41. There are many hurdles to the development of orally bioavailable drugs. A major issue is that oral bioavailability generally requires a molecular weight of less than 1000 Da. (In comparison, T-20 and C34 have molecular weights greater than 4000 Da). Based on the crystal structure of the gp41 core, the prehairpin intermediate is expected to contain three prominent, symmetry-related pockets on the surface of the central trimeric coiled coil (Figure 1). Each pocket has an internal volume of roughly 400 $Å^3$ and could be filled by a molecule with a molecular weight of approximately 500 Da, raising the possibility that it could be targeted by small-molecule drugs.

For several additional reasons, these gp41 pockets are expected to be attractive drug targets (Chan et al., 1997, 1998). First, mutagenesis studies indicate that the N-peptide residues forming the pocket are critical for membrane fusion (Dubay et al., 1992; Cao et al., 1993; Chen et al., 1993; Wild et al., 1994; Weng and Weiss, 1998; D. C. Chan and P. S. K., unpublished results). Second, studies of C34 variants show that C34 inhibitory

*To whom correspondence should be addressed (e-mail: tocio@wi.mit.edu).

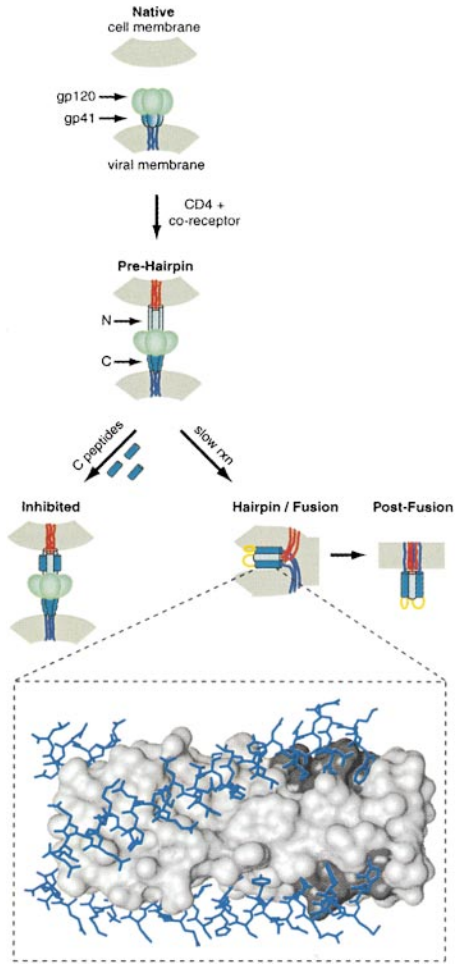


Figure 1. Model of HIV Membrane Fusion and Structure of the gp41 Core

Schematic representation of a working model for HIV membrane fusion (for review, Chan and Kim, 1998). In the native state of the trimeric gp120/gp41 complex ("Native"), the fusion peptide and N-peptide regions of gp41 are not exposed. Following interaction with cellular receptors (CD4 and coreceptor), a conformational change results in formation of the transient prehairpin intermediate ("Pre-Hairpin"), in which the fusion peptide regions (red lines) are inserted into the cell membrane and the coiled coil of the N-peptide region of gp41 (indicated as "N") is exposed. However, the C-peptide region of gp41 (indicated as "C") is constrained and unavailable for interaction with the coiled coil. Thus, exogenous C-peptides can bind to the prehairpin intermediate and inhibit fusion in a dominant-negative manner ("Inhibited"). In the absence of inhibitors, the prehairpin intermediate resolves to the hairpin structure and membrane fusion results ("Hairpin/Fusion"), although it is not known whether hairpin formation precedes membrane fusion per se. The C-peptides discussed in this paper (and corresponding residues in gp41, numbered according to their position in gp160 of the HXB2 HIV-1 strain) are as follows: C34 [628–661]; DP178, also called T-20 [638–673]; and T649 [628–663]. Adapted from Chan and Kim (1998).

The inset depicts the 2.0 Å X-ray crystal structure of N36/C34, a peptide version of the HIV-1 gp41 core (Chan et al., 1997). Three central N-peptides form a coiled coil, shown here as a surface representation, and three helical C-peptides pack along conserved grooves on the surface of the coiled-coil trimer. There are three symmetry-related hydrophobic pockets on the surface of the N-peptide coiled coil (shaded). The pocket region is highly conserved among HIV-1 isolates. There are 11 residues that comprise the lining of the hydrophobic pocket (see Figure 7 of Chan et al., 1997): Leu-565, Leu-566, Leu-568, Thr-569, Val-570, Trp-571, Gly-572, Ile-573, Lys-574, Leu-576, and Gln-577 of HXB2. These 11 residues are completely conserved in 158 of 202 fully sequenced M group HIV-1 strains (HIV Sequence Database [1998/1999 alignments], Los Alamos National Laboratory, http://hiv-web.lanl.gov). Of the remaining 44 isolates, 33 possess only a single conservative methionine substitution for Leu-565.

activity depends on its ability to bind to the pocket (Chan et al., 1998). Third, therapeutic agents that target this pocket would likely be relatively elusive to the emergence of resistant viral strains because (1) residues comprising this region are highly conserved among HIV-1 isolates (see Figure 1 legend), and (2) the segment of mRNA encoding these residues is part of the structured RNA region of the Rev-responsive element (RRE) that is important for Rev response (Malim et al., 1989; Zapp and Green, 1989). Indeed, C-peptides that do not contain pocket-binding residues, such as T-20, are more vulnerable to the emergence of resistant virus than are C-peptides that contain pocket-binding residues, such as T649 (Rimsky et al., 1998).

Despite the expectation that the pocket is a promising drug target, there is no direct evidence that molecules binding solely to the pocket (i.e., and not simultaneously to other regions of the gp41 coiled coil) can inhibit HIV-1 infection. It would therefore be desirable to screen a library of compounds for their pocket binding ability. The obvious target for such screens is the N-peptide region that comprises the central trimeric coiled coil of the gp41 crystal structure. Unfortunately, in the absence of C-peptides, N-peptides aggregate (Lu et al., 1995, and references therein; Lu and Kim, 1997), presumably due to the hydrophobic nature of the grooves into which the C-helices pack. A more soluble but accurate representation of the pocket is required for drug screening. We have overcome the aggregation problem by fusing the corresponding gp41 residues to a soluble trimeric coiled coil. The resulting peptide, IQN17, is expected to properly display the pocket of gp41 and thus should be useful for the discovery of compounds that inhibit HIV-1 entry.

Indeed, we have used IQN17 and a previously developed mirror-image phage display method (Schumacher et al., 1996) to identify peptides composed of D-amino acids (which are therefore insensitive to proteolytic degradation) that bind to the hydrophobic pocket of gp41, as shown in NMR studies and a 1.5 Å X-ray cocrystal structure. These D-peptides inhibit gp41-mediated cell–cell fusion and HIV-1 infection, validating the pocket as a target for drug development. The D-peptides may also serve as therapeutic or prophylactic agents, or as leads for the development of such agents, for combating HIV infection. Finally, as pocket-binding molecules, the D-peptides can be used in competitive screens to identify small-molecule drug candidates with similar inhibitory capabilities.

Results

IQN17

Our focus is on the hydrophobic pocket of the N-peptide coiled-coil region of gp41 (Figure 1, inset). A small molecule that binds to this pocket might be expected to

act in a dominant-negative manner, analogous to the synthetic C-peptides, inhibiting formation of the fusion-active structure (Chan et al., 1997). In the absence of C-peptides, N-peptides aggregate (Lu et al., 1995, and references therein; Lu and Kim, 1997) and therefore are not effective for screening small molecules.

We have designed a molecule, denoted IQN17, in which a soluble trimeric coiled coil (GCN4-pI$_Q$I; Eckert et al., 1998) is fused to the portion of the N-peptide that comprises the gp41 hydrophobic pocket (Figure 2A). Analyses of the X-ray crystal structures of GCN4-pI$_Q$I (Eckert et al., 1998) and the N36/C34 core of gp41 (Chan et al., 1997) indicate that the superhelical parameters (Crick, 1953; Harbury et al., 1995) of the two coiled coils are similar. Thus, structural perturbations resulting from creation of the chimera are expected to be minimal.

We initially constructed a peptide in which the first 29 residues of GCN4-pI$_Q$I were fused to the last 17 residues of N36 (there is a 1 residue overlap between the two regions, making the peptide 45 residues long). This peptide did not form a specific trimeric species as determined by sedimentation equilibrium experiments, but instead formed higher-order aggregates (D. M. E. and P. S. K., unpublished results). We therefore made three surface residue mutations, which were expected to increase solubility, in the GCN4-pI$_Q$I region of the molecule (Figure 2A). The resulting peptide, IQN17, is fully helical, as demonstrated by circular dichroism spectroscopy (Figure 2B), and is a soluble trimeric species as determined by sedimentation equilibrium experiments (Figure 2C).

Mirror-Image Phage Display

D-peptides that bind to a target protein of interest can be identified through mirror-image phage display (Figure 3A; Schumacher et al., 1996). The desired target is synthesized chemically with D-amino acids, resulting in a product that is the mirror image of the naturally occurring (L-amino acid) form. This D-target is then used to screen phage expressing a peptide library on one of the phage coat proteins, yielding specific phage clones with L-peptide sequences that bind to the D-target. The mirror images of the phage-expressed L-peptide sequences are chemically synthesized with D-amino acids. By symmetry, these D-peptides should bind to the natural (L-amino acid) form of the target. This technique has been used to identify D-peptides that bind to the SH3 domain of c-Src (Schumacher et al., 1996). By using IQN17 in mirror-image phage display, D-peptides that bind to the pocket of gp41, and are therefore potential inhibitors of HIV-1 infection, can be identified.

The mirror image of IQN17 (denoted D-IQN17) was chemically synthesized using D-amino acids (see Experimental Procedures). As expected, D-IQN17 has a circular dichroism spectrum that is of opposite sign to that observed for the L-amino acid version (D. M. E. and P. S. K., unpublished results). The N terminus of IQN17 was biotinylated to allow immobilization on streptavidin-coated plates for phage panning. A three–amino acid linker (Gly-Lys-Gly) containing an L-lysine was placed between the biotin and IQN17. The L-lysine residue provided a trypsin recognition site, so that bound phage could be eluted by addition of trypsin, rather than by a

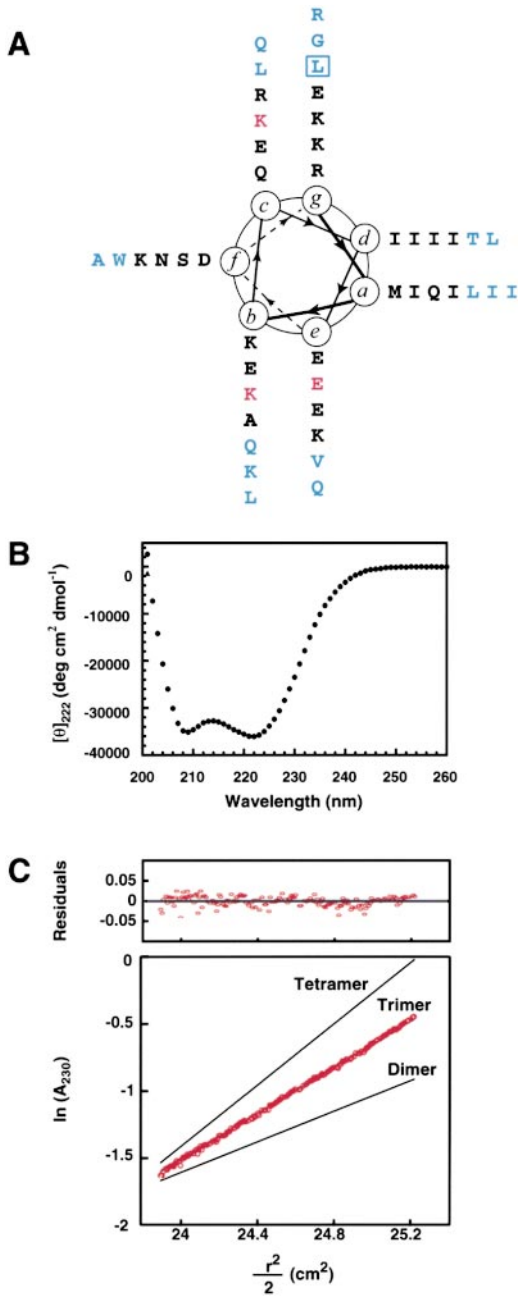


Figure 2. IQN17: Soluble, Trimeric Presentation of the gp41 Pocket

(A) Helical wheel diagram of IQN17. The 17 residues indicated in blue are from gp41 (residues 565–581 of HXB2 gp160). The residues in black are from GCN4-pI$_Q$I, and the residues in red correspond to three surface substitutions in the GCN4-pI$_Q$I region, made to increase solubility. The boxed residue represents the junction, at a leucine (in a "g" position of the coiled coil) that is shared between GCN4-pI$_Q$I and gp41.

(B) Circular dichroism spectrum of a 10 μM solution of IQN17 in PBS (pH 7.4), 4°C. IQN17 is fully helical, with an approximate ellipticity of $-36{,}000$ deg cm^2 dmol^{-1} at 222 nm.

(C) Sedimentation equilibrium data of IQN17 at 20 μM in PBS (pH 7.0). Data are plotted as ln(absorbance) versus half the square of the radius from the axis of rotation. The slope of the data is proportional to the molecular mass of the peptide oligomer. Continuous lines indicate the expected slopes for the different oligomeric states. IQN17 is a discrete trimer under these conditions.

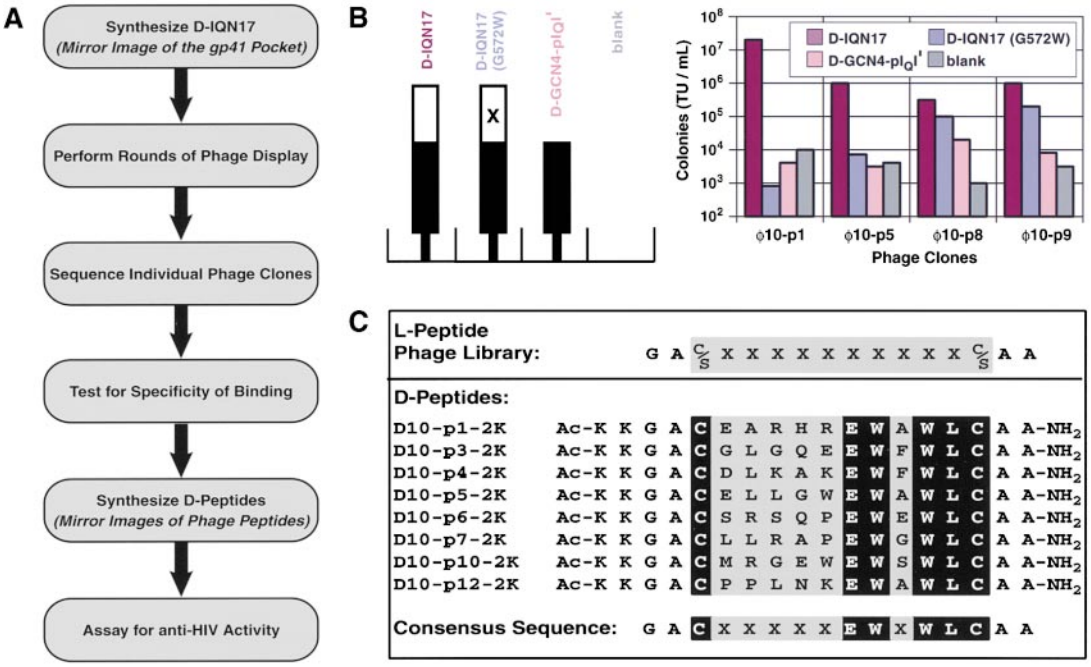


Figure 3. Mirror-Image Phage Display Using D-IQN17 as a Target

(A) Schematic of the mirror-image phage display method (Schumacher et al., 1996) applied to the gp41 pocket. IQN17 is synthesized with D-amino acids to generate a mirror image of the gp41 pocket. A phage-displayed peptide library is subjected to several rounds of selection, and phage that bind to D-IQN17 are sequenced. After testing for specific binding, the mirror images of the phage-expressed peptides are chemically synthesized with D-amino acids and tested for anti-HIV-1 activity.

(B) Test for pocket-specific binding phage. Individual phage clones are screened for binding to wells containing: (i) D-IQN17, (ii) D-IQN17[G572W], (iii) D-GCN4-pI$_Q$I′, and (iv) no target. In this schematic, the GCN4-pI$_Q$I′ region of the molecule is depicted in black, and the HIV-1 gp41 region is depicted in white; the G572W mutation is indicated with an "X." D-GCN4-pI$_Q$I′ is a D-amino acid version of GCN4-pI$_Q$I (Eckert et al., 1998) that contains the same surface mutations as IQN17. The G572W substitution in D-IQN17[G572W] introduces a large protrusion into the gp41 pocket. Phage are considered pocket-specific binders if they bind appreciably to wells containing IQN17, but not to the other wells. Representative data are shown for pocket-specific binding phage (ϕ10-p1 and ϕ10-p5) and nonspecific phage (ϕ10-p8 and ϕ10-p9).

(C) D-peptide sequences. Peptides corresponding to the mirror images of the peptides expressed on pocket-specific phage were synthesized with D-amino acids. Four conserved residues among the ten positions that were randomly encoded in the phage library, as well as two Cys residues that form an intramolecular disulfide bond to generate a cyclic peptide, define a consensus sequence. Flanking residues from the gene III phage protein were added at the N termini (GA) and C termini (AA) to more closely represent the phage-displayed peptides. In addition, lysine residues were added at the N termini of some peptides to increase solubility (peptides with two or four additional lysines are denoted with the suffixes −2K or −4K, respectively; see text). The termini of all peptides were blocked (Ac, acetyl; NH_2, amide).

nonspecific elution procedure such as acid addition (cf. Wrighton et al., 1996).

Phage that bind to D-IQN17 might bind to any region of this target. Our interest was to identify those phage that bind only to the gp41 pocket region. Phage that bound specifically to D-IQN17 (i.e., not to wells that lacked the target) were therefore further tested against a panel of molecules lacking the target pocket structure (Figure 3B). Phage were considered to be pocket specific if they bound to D-IQN17 but not to D-GCN4-pI$_Q$I′ or D-IQN17(G572W). D-GCN4-pI$_Q$I′ is a D-amino acid version of GCN4-pI$_Q$I that contains the same surface mutations as IQN17 and lacks the pocket residues entirely, while D-IQN17(G572W) contains a tryptophan in the position corresponding to Gly-572 of gp41, thereby introducing a large protrusion into the pocket.

Bacteriophage fd expressing a library of L-amino acid peptides fused to the N terminus of the gene III protein were utilized. The phage-expressed peptides (Schumacher et al., 1996) contain ten randomly encoded amino acid residues flanked by either a cysteine or a serine on both sides (Figure 3C). After several rounds of panning, 12 IQN17-specific phage were identified and labeled ϕ10-p1 through ϕ10-p12 (see Experimental Procedures). Nine of the 12 phage clones are pocket-specific binders, with some showing over 1000-fold more binding to D-IQN17 than to the controls (Figure 3B). In addition, eight of these nine contain the consensus sequence CXXXXXEWXWLC (Figure 3C).

Inhibition of Membrane Fusion

D-peptides corresponding to the mirror images of the pocket-specific phage-displayed peptides were synthesized using D-amino acids. The D-peptides are denoted D10-p1 through D10-p12, corresponding to the mirror images of the ϕ10-p1 through ϕ10-p12 sequences (for a total of nine D-peptides; three of the phage peptides were not pocket specific and were therefore not studied further). The peptides (Figure 3C) were synthesized with blocked termini (acetylated on the N terminus and ending in a C-terminal amide) and were oxidized to permit disulfide bond formation. In addition, two flanking residues from the gene III protein were included on each end of each peptide so as to resemble more closely the

peptide on the surface of the phage. To increase water solubility of the D-peptides, many were synthesized with additional N-terminal lysines (see Experimental Procedures). Most experiments described here were performed using D-peptides with two N-terminal lysines (denoted with the suffix "-2K"). However, in several of the experiments other variants were employed including those with no lysines (for example D10-p1 in the X-ray crystallography studies) and those with four lysines (denoted with the suffix "-4K," for example D10-p5-4K in the NMR experiments).

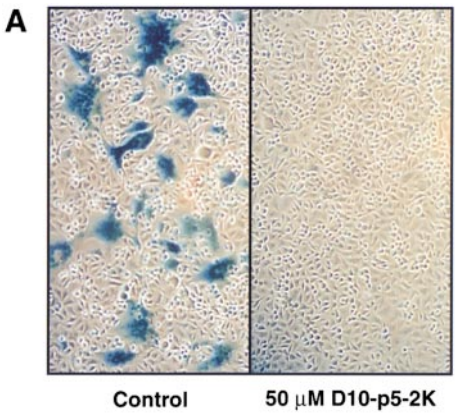
Cell/cell fusion assays, in which cells expressing HIV-1 envelope glycoprotein are mixed with cells expressing CD4 and coreceptor, were performed in the presence of these nine peptides. The eight peptides sharing the consensus sequence inhibit cell/cell fusion with IC_{50} values (50% inhibitory concentration) in the micromolar concentration range, ranging from 3.6 μM for D10-p5-2K to 130 μM for D10-p12-2K (Figure 4). The ninth sequence was toxic to cells at comparable concentrations and did not show inhibitory activity at lower concentrations—it was not studied further (D. M. E. and P. S. K., unpublished results). The D-peptides were also tested for their ability to inhibit HIV-1 infection of cells, using a recombinant luciferase-based HIV-1 infection assay (Chen et al., 1994). All of the peptides tested show inhibition of HIV-1 entry into cells, again in the micromolar concentration range (Figure 4C).

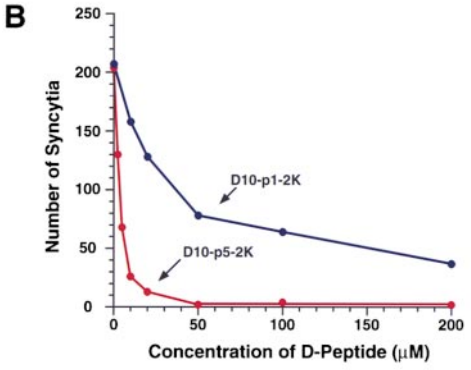
In order to assess the effect of the added lysines, D-peptides with different numbers of lysines were compared in the cell/cell fusion assay (Figure 4C). D10-p1-2K, D10-p4-2K, and D10-p6-2K were found to have IC_{50} values for inhibition of cell/cell fusion (i.e., syncytia formation) approximately 1.5- to 2-fold higher than the respective variants without added lysines. The IC_{50} values for inhibition of syncytia formation of D10-p5-4K was approximately 1.5-fold higher than D10-p5-2K. We conclude that the addition of N-terminal lysine residues to the D-peptides results in only a modest decrease in inhibitory activity.

The pocket-binding phage sequences always contain cysteines in the flanking regions (Figure 3C), even though the library contains either cysteines or serines at these positions, strongly suggesting that an intramolecular disulfide bond is required for pocket binding and viral inhibition by these D-peptides. To more directly address this issue, a derivative of D10-p5-2K in which each cysteine was replaced with alanine was evaluated in the cell/cell fusion assay. Although toxic to cells at concentrations above 50 μM, this control peptide did not noticeably inhibit syncytia formation up to this concentration (D. M. E. and P. S. K., unpublished results). Since D10-p5-2K has an IC_{50} of 3.6 μM (Figure 4C), we conclude that the disulfide bond is crucial for the inhibitory activity of these D-peptides.

Crystal Structure of the IQN17–D10-p1 Complex

We obtained diffraction quality crystals of D10-p1 bound to IQN17 using the hanging-drop vapor diffusion method. The structure of the complex was solved by multiwavelength anomalous dispersion (MAD) analysis (for review, Hendrickson, 1991) of an osmium derivative. The final structure was refined against data to 1.5 Å from a native





C

D-Peptide	IC_{50} (Syncytia) (μM)	IC_{50} (Viral Entry) (μM)
D10-p1	29	
D10-p1-2K	46	85
D10-p3-2K	16	39
D10-p4	64	
D10-p4-2K	83	160
D10-p5-2K	3.6	11
D10-p5-4K	5.6	
D10-p6	53	94
D10-p6-2K	113	210
D10-p7-2K	41	
D10-p10-2K	48	96
D10-p12-2K	130	270

Figure 4. Inhibition of HIV-1 Membrane Fusion by D-Peptides

(A) Photographs of cell culture assays of fusion between CHO cells expressing HIV-1 envelope (Kozarsky et al., 1989) and CD4-HeLa cells with a β-galactosidase reporter gene (Kimpton and Emerman, 1992). Cells were cocultured in the absence or presence of D10-p5-2K, and then incubated with β-galactosidase substrate. Multinucleated cells (syncytia) are stained in blue.

(B) Results from individual cell–cell fusion assays. The number of syncytia is plotted as a function of D-peptide concentration.

(C) IC_{50} values for inhibition of cell/cell fusion as determined by the syncytia assay described in (A) and of HIV-1 infection as determined by a recombinant luciferase assay (Chen et al., 1994). For the inhibition of syncytia, average values from four to six determinations are given (standard errors of approximately ±10%). For inhibition of virus infection, average values from two to five determinations are given (standard error of approximately ±20%).

Table 1. Crystallographic and Refinement Statistics

Data Collection

Crystal	λ (Å)	Completeness (%)	R_{sym}[a] (%)	Resolution (Å)
IQN17/D10-p1	1.1197	93.8	4.8	1.5
Os λ1	1.1403	98.6	6.3	2.0
Os λ2	1.1399	96.8	9.7	2.0
Os λ3	1.1393	96.9	7.9	2.0
Os λ4	1.1197	97.0	8.4	2.0

MAD Phasing Statistics (22.0–2.0 Å)

Derivative	R_{iso}[b] (%)	R_{cullis}[c] (Acentric)	R_{cullis}[c] (Centric)	R_{cullis}[c] (Anom.)	Ph. Power[d] (Acentric)	Ph. Power[d] (Centric)	Occ.[e]	Anom. Occ.[e]
Os λ1 vs. λ4	7.3	0.75	0.61	0.47	1.41	1.21	−0.039	0.337
Os λ2 vs. λ4	5.2	0.83	0.71	0.44	1.04	1.15	−0.027	0.533
Os λ3 vs. λ4	3.3	0.97	0.97	0.49	0.35	0.28	−0.005	0.295

Overall figure of merit (before solvent flattening): 0.68

Refinement Statistics

Crystal	Non-hydrogen protein atoms	Waters	Ions	Resolution (Å)	Reflections (Total)	R_{cryst}[f]	R_{free}[f]	Rms Deviations Bonds (Å)	Rms Deviations Angles (°)
IQN17/D10-p1	516	150	1	10.0–1.5	13,549	0.214	0.245	0.012	1.498

[a] $R_{sym} = \Sigma\Sigma_j|I_j - \langle I\rangle|/\Sigma\Sigma_j\langle I\rangle|$, where I_j is the recorded intensity of the reflection j and $\langle I\rangle$ is the mean recorded intensity over multiple recordings.
[b] $R_{iso} = \Sigma||F_{(\lambda i)} \pm F_{(\lambda 4)}| - |F_{(\lambda i)}||/\Sigma|F_{(\lambda 4)}|$, where $F_{(\lambda i)}$ is the structure factor at wavelength λi and $F_{(\lambda 4)}$ is the structure factor at the reference wavelength λ4.
[c] $R_{cullis} = \Sigma||F_{(\lambda i)} \pm F_{(\lambda 4)}| - |F_{h(\lambda i),c}||/\Sigma|F_{(\lambda i)} \pm F_{(\lambda 4)}|$, where $F_{h(\lambda 4),c}$ is the calculated heavy atom structure factor.
[d] Phase power $= \langle F_{h(\lambda i)}\rangle/E$, where $\langle F_{h(\lambda i)}\rangle$ is the root-mean-square heavy atom structure factor and E is the residual lack of closure error.
[e] Occupancies are values output from MLPHARE.
[f] R_{cryst}, $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/|F_{obs}|$, where the crystallographic and free R factors are calculated using the working and test sets, respectively. Test set contained 10% of reflections.

crystal to a crystallographic R factor of 21.4% with an R_{free} of 24.5% (Table 1). The electron density map at this resolution reveals structural details of IQN17 and D10-p1 with clarity (Figure 5C). IQN17 is a continuous, regular three-stranded coiled coil (Figure 5A). Structural superposition indicates that the overall architecture of the HIV-1 gp41 hydrophobic pocket in the IQN17/D10-p1 complex is almost identical to that in the wild-type HIV-1 gp41 structure (Chan et al., 1997), with a C_α rmsd of 0.65 Å. The conformation of the GCN4-pI_QI fraction of the molecule is also well preserved, with a C_α rmsd from the original GCN4-pI_QI (Eckert et al., 1998) of 0.40 Å.

As anticipated from our binding studies (Figure 3B), D10-p1 is bound only to the gp41 region of IQN17 (Figures 5A and 5B). A disulfide bond in D10-p1 between Cys-3 and Cys-14 results in a somewhat circular overall structure. Two segments of this D-peptide, Ala-2 to Ala-5, and Ala-11 to Ala-16, form short left-handed α helices, and the residues between these two segments form a loop. Two of the loop residues (Arg-8 and Glu-9) are also in a left-handed helical conformation. The longer (C-terminal) helical segment of D10-p1 binds the gp41 pocket of IQN17 in an oblique manner, at a similar angle to the packing of the HIV-1 C34 helix against the N36 peptide coiled coil (Chan et al., 1997). In a superposition of the two crystal structures, the overall positions of the D10-p1 and C34 helices closely overlap, but the orientations of most side chains are significantly different, owing to the opposite handedness of the inhibitors (Figure 6).

Of the 16 residues in D10-p1, only 6 interact directly with the gp41 pocket of IQN17 (Figure 5B). These residues are either part of the conserved EWXWL sequence (Trp-10, Trp-12, and Leu-13) or invariant from the original flanking phage sequence (Gly-1, Ala-2, and Ala-16). The side chains of Trp-10, Trp-12, Leu-13, and Ala-16 are deeply buried in the hydrophobic pocket of IQN17. Gly-1 and Ala-2 are not deeply buried but also contact the indole group of Trp-571 in the pocket. (For the gp41-derived residues of IQN17, we use the residue numbers corresponding to those in the HXB2 strain of HIV-1 gp160.) Although the interface between D10-p1 and IQN17 is largely nonpolar, a hydrogen bond is formed between a pocket residue and D10-p1: Gln-577 $O_{\epsilon 1}$–Trp-12 $N_{\epsilon 1}$ (2.8 Å). The potential exists for a second intermolecular hydrogen bond (Gln-577 $N_{\epsilon 2}$–Glu-9 O; 3.1 Å) with less ideal geometry. It is not apparent why Glu-9 is absolutely conserved (Figure 3C), although its side chain carboxylate does bend back to accept a hydrogen bond from its own backbone amide (2.7 Å), and both of its methylene groups contact the δ_1-methyl of Leu-581.

The identities of the conserved D-peptide pocket-binding residues (Trp-10, Trp-12, and Leu-13) are very similar to those of C34 (Trp-628, Trp-631, and Ile-635). However, only the indole rings of Trp-10 (D10-p1) and Trp-628 (C34) occupy the same position when bound to the gp41 pocket (Figure 6). The packing differences of the Trp-12 and Leu-13 side chains (relative to Trp-631 and Ile-635 in C34) are accompanied by slight changes in the shape of the pocket: (1) Lys-574, lining the left side of the pocket (in the orientation shown in Figure 6), is pushed out, widening the pocket as compared to the N36/C34 structure; (2) Gln-577, lining the bottom, pushes further into the pocket, raising the bottom wall (mainly due to a rotation about χ_2); and (3) a rotamer change of Leu-565 slightly lowers the top wall of the pocket.

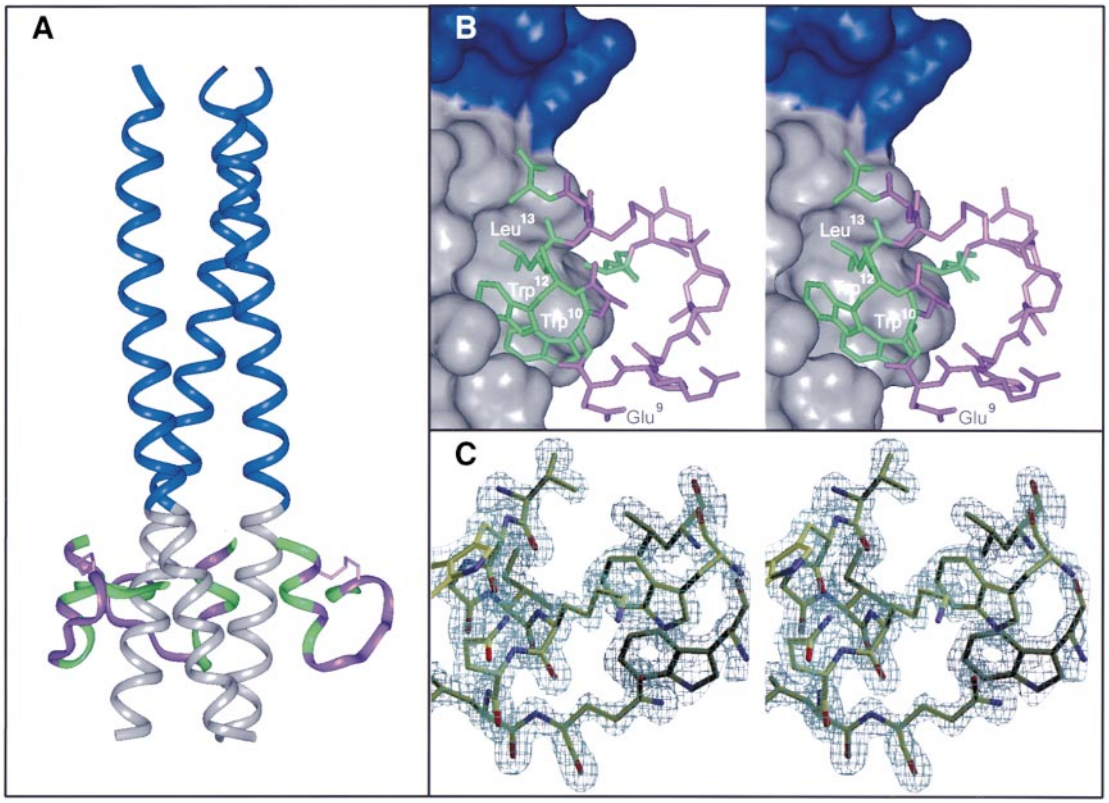


Figure 5. Crystal Structure of a D-Peptide Bound to the gp41 Pocket

(A) Ribbon representation of the overall structure of the IQN17/D10-p1 complex. The GCN4-pI$_O$I′ part of the chimera (dark blue) and the HIV-1 gp41 hydrophobic segment (gray) form a continuous three-stranded coiled coil. Three D10-p1 inhibitors (purple and green) bind solely to the hydrophobic pocket. The six residues of the D-peptide that make direct contact with IQN17 are shown in green (Gly-1, Ala-2, Trp-10, Trp-12, Leu-13, and Ala-16). Figure drawn with Insight II 98.0 (Molecular Simulations Inc.).

(B) Stereo view of the IQN17/D10-p1 complex in which IQN17 is represented as a molecular surface and D10-p1 is represented with sticks. The color scheme is as in (A). The four conserved residues of the EWXWL motif (Glu-9, Trp-10, Trp-12, and Leu-13) are labeled. Figure drawn with Insight II 98.0 (Molecular Simulations Inc.).

(C) Stereo view of a region of the final 1.5 Å 2Fo-Fc map, contoured at 2.1σ, superimposed on the final model. The view is approximately the same orientation as in (B). Figure drawn with O (Jones et al., 1991).

Overall, however, the hydrophobic pocket maintains its integrity between the N36/C34 and IQN17/D10-p1 structures (Figure 6).

NMR Studies of the IQN17/D–Peptide Complexes

The finding that Trp-10, Trp-12, and Leu-13 of D10-p1 are buried in the IQN17 pocket, and the conservation of these residues among the selected phage sequences, strongly suggest that the D-peptides all bind in a similar manner. To more directly address this issue, NMR experiments were used to assess the binding of each D-peptide to IQN17. In the cocrystal structure, Trp-10 of D10-p1 and Trp-571 of IQN17 are oriented in a manner that should alter the chemical shifts of several protons

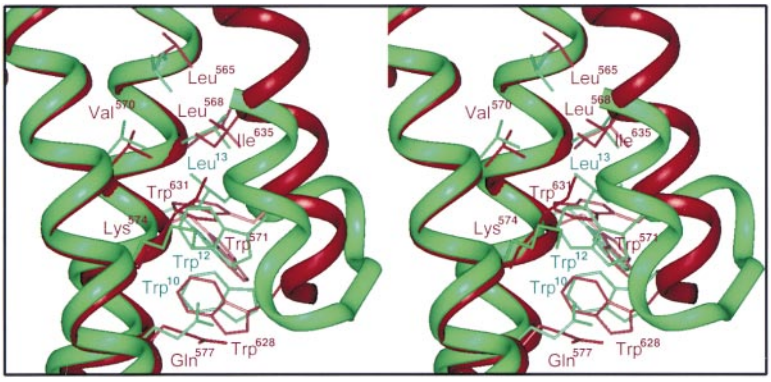


Figure 6. Comparison of the Pocket in the C34 and D10-p1 Bound States

Stereo view of the structural superposition between IQN17/D10-p1 (green) and HIV-1 N36/C34 (red) structures in the region of the hydrophobic pocket. Two of the pocket-binding residues of D10-p1 (Trp-12 and Leu-13) differ in position from two of the pocket-binding residues of C34 (Trp-631 and Ile-635), whereas Trp-10 of D10-p1 and Trp-628 of C34 are closely overlaid. In addition, several residues comprising the hydrophobic pocket occupy slightly different positions in the two structures (Leu-565, Gln-577, Lys-574, as labeled). Figure drawn with Insight II 98.0 (Molecular Simulations Inc.).

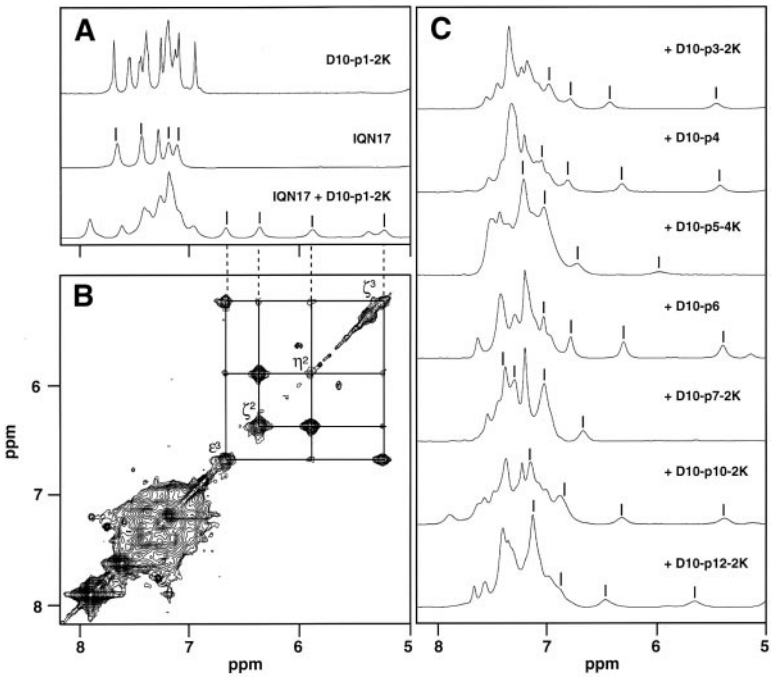


Figure 7. NMR Indicates a Similar Mode of Binding for Many D-Peptides

(A) One-dimensional ^{1}H NMR spectra of D10-p1-2K (top), IQN17 (middle), and a 1:1 complex of D10-p1-2K and IQN17 (bottom). The x axis is the same as for (B) below. The upfield peaks assigned to four aromatic ring protons of Trp-571 are indicated. The unmarked upfield peak of the bottom trace corresponds to an unassigned H_α resonance.

(B) Two-dimensional TOCSY NMR spectrum of the IQN17/D10-p1-2K complex. Cross-peaks linking the same Trp-571 protons are indicated, along with specific assignments. No other type of side chain is expected to display this cross-peak pattern in this chemical shift range.

(C) One-dimensional ^{1}H NMR spectra of 1:1 complexes between IQN17 and each D-peptide (as labeled). The same four protons from Trp-571 (assigned in separate TOCSY and NOESY experiments) are indicated in each spectrum. Some upfield-shifted peaks are selectively broadened compared to others in the same spectra, especially for the complex with D10-p5-4K, suggestive of a chemical exchange process.

due to aromatic ring current effects (Bovey, 1988). Specifically, use of the structure-based chemical shift prediction program SHIFTS (Osapay and Case, 1991) predicted that four protons of the Trp-571 indole ($H_{\zeta 2}$, $H_{\eta 2}$, $H_{\zeta 3}$, and $H_{\epsilon 3}$) would experience a significant upfield shift, especially the $H_{\zeta 3}$ proton (P. A. C. and P. S. K., unpublished results). Figure 7A demonstrates these effects: the aromatic region of the IQN17/D10-p1-2K NMR spectrum displays chemical shifts dramatically different from either of the two separate components, including a set of upfield peaks. A two-dimensional TOCSY experiment was used to confirm that the upfield peaks all correspond to a single Trp indole (Figure 7B).

If each D-peptide binds the IQN17 pocket in the same fashion as D10-p1, a similar juxtaposition of Trp-571 and Trp-10 should occur, also resulting in upfield-shifted peaks. All of the IQN17/D-peptide complexes studied displayed such peaks (Figure 7C). Although there is substantial variation in the magnitudes of these changes, ranging from roughly 0.5 to 2 ppm for the most upfield-shifted proton ($H_{\zeta 3}$, in all cases where it could be assigned), ring current effects can be highly sensitive to distance and orientation. Moreover, the magnitudes of these changes are very large. By comparison, chemical shift differences used to detect binding in studies of SAR (structure–activity relationships) by NMR (Shuker et al., 1996) are frequently much smaller, in the range of 0.05 to 0.2 ppm. Two-dimensional (2D) TOCSY experiments confirm that the upfield-shifted aromatic peaks of each complex belong to a single indole side chain, almost certainly Trp-571. For several of the complexes studied, 2D NOESY experiments also indicate contact between this side chain and another aromatic group, presumably Trp-10 (P. A. C., D. M. E., and P. S. K., unpublished results).

We conclude that in the majority of these IQN17 complexes (i.e., D10-p1-2K, D10-p3-2K, D10-p4, D10-p6, D10-p10-2K, and D10-p12-2K) the D-peptides contact the hydrophobic pocket with very similar binding interfaces, bringing Trp-571 in close contact with the aromatic ring of Trp-10 in roughly the same orientation. For the complexes with D10-p5-4K and D10-p7-2K, this conclusion also seems very likely, although the more limited chemical shift dispersion and broader peaks raise the possibility of some other mode of binding.

Discussion

Validation of the gp41 Pocket as a Target for Drug Development

Inhibition of HIV-1 infection by synthetic C-peptides, and structural studies of gp41, have shed light on the vulnerable but transient prehairpin intermediate of the envelope glycoprotein during the viral entry process. Several lines of evidence suggest that the C-peptides act in a dominant-negative manner by binding to this prehairpin intermediate (Figure 1). First, the C-peptides must be present during the infection process to act as inhibitors. Preincubation of virus with a C-peptide, followed by its removal, does not inactivate the virus (Furuta et al., 1998). Second, the inhibitory potency of C-peptides is greatly reduced in the presence of an equimolar amount of complementary N-peptides (Lu et al., 1995; see also Chen et al., 1995). Third, C-peptide derivatives that have weakened interactions with N-peptides show decreased potency as inhibitors (Wild et al., 1995; Chan et al., 1998). Fourth, interaction between gp41 and a C-peptide has been demonstrated in immunoprecipitation experiments. Importantly, this interaction is detected only after exposure of the envelope complex to cellular receptors (Furuta et al., 1998).

Despite the antiviral effect of synthetic C-peptides, orally bioavailable small molecules are more desirable. Small molecules that bind to the hydrophobic pocket of

the gp41 core might be expected to function as inhibitors in the same dominant-negative manner as the C-peptides (Chan et al., 1997). Recent mutagenesis studies on C34 demonstrate that the hydrophobic residues that insert into this pocket are important for inhibitory potency, reinforcing the suggestion that the gp41 pocket is an attractive drug target (Chan et al., 1998). Nonetheless, this hypothesis had not been tested directly, and another potent C-peptide, DP178 (or T-20), lacks these pocket binding residues entirely (Kilby et al, 1998).

Our finding that D10-p1 inhibits HIV-1 infection and binds exclusively to the gp41 pocket (as shown by X-ray crystallography) demonstrates directly that molecules that bind solely to the pocket can inhibit infection. The size of this pocket (approximately 400 $Å^3$) makes it ideal for binding by a small molecule. Thus, validation of the gp41 hydrophobic pocket as a target for drug discovery sets the stage for the development of a new class of orally bioavailable anti-HIV drugs that inhibit viral entry into cells. These drugs would be a useful addition to the current regimen used to treat HIV-1 infection. Such drugs would also demonstrate the feasibility of small-molecule therapeutics that interfere with protein–protein interactions (as opposed to enzyme activity).

D-Peptides and Derivatives as Inhibitors of HIV Entry

The D-peptides identified here can serve as starting points for the development of therapeutic or prophylactic agents that inhibit HIV entry. However, the potency of these D-peptides is only in the micromolar range (Figure 4C). For comparison, C34 and T-20 each have nanomolar potency (in cell culture), but they are subject to proteolysis and are large in size ($\sim$35 residues). The D-peptides are also larger than desired for oral administration ($\geq$16 residues). In contrast, cyclosporin, an effective oral pharmaceutical, is a cyclic 11-residue peptide composed of D-amino acids and N-methylated amino acids. Thus, chemical optimization of these D-peptide HIV inhibitors may lead to derivatives that are directly useful as therapeutic agents.

Several considerations suggest that it may be possible to improve the potency and reduce the size of these D-peptides. First, mirror-image phage display experiments utilize only the 20 naturally occurring amino acid residues, leaving a tremendous range of chemical diversity unsampled for both the side chains and the backbone. Second, several of the remaining residues are involved in a loop remote from the pocket (Figure 5B). Third, IQN17, alone or in combination with a specific pocket-binding molecule (e.g., D10-p1), greatly facilitates the ease with which more sophisticated drug discovery methods (e.g., utilizing combinatorial libraries based on these D-peptides) can be employed to discover better inhibitors. Fourth, the NMR assay employed here can be used as a fast screen to verify that variants bind the gp41 hydrophobic pocket in a similar manner as the lead D-peptides (cf. Shuker et al., 1996). Finally, the availability of a high-resolution cocrystal structure enables rational drug development approaches.

Toward Small-Molecule Inhibitors

Our results also make possible the rapid discovery of other drug candidates with the same mechanism of action as the D-peptides identified here, but with completely unrelated chemical structures. In particular, a molecule that accurately presents the gp41 hydrophobic pocket (e.g., IQN17) either alone or in combination with a pocket-binding molecule (e.g., D10-p1) constitutes the basis for high-throughput drug discovery screens. Chemical libraries can be screened with IQN17 to identify pocket-binding molecules. Alternatively, drug candidates can be identified by their ability to disrupt the interaction between IQN17 and a D-peptide.

Implications for Vaccine Candidates

Finally, the strategy used to create a soluble, trimeric version of a part of the gp41 N-terminal coiled-coil region may also help efforts to develop HIV vaccine candidates. Recent studies have raised the possibility that transiently exposed conformations of proteins required for HIV-1 entry can be useful for eliciting a neutralizing antibody immune response (LaCasse et al., 1999). It is possible that the transiently exposed prehairpin intermediate is a vulnerable target for neutralizing antibodies. The most obvious approach toward eliciting an antibody response against the N-helix region of gp41 would be to use an N-peptide as an immunogen. However, since the isolated N-peptides aggregate, they do not properly present the gp41 N-helix coiled-coil trimer. Accordingly, the same strategy used here to solve the aggregation problem for the gp41 hydrophobic pocket can be employed to develop soluble, trimeric versions of the gp41 N-helix coiled-coil region that might have utility as immunogens.

Experimental Procedures

Phage Display

Bacteriophage fd expressing a library of ten randomly encoded amino acids flanked by Ser or Cys residues on the N terminus of the pIII protein (Schumacher et al., 1996) were utilized here (Figure 3C). The expressed fusion protein contains an N-terminal Ala-Asp-Gly-Ala sequence preceding the peptide library sequence. The amplified phage stock had a titer of approximately 10^{12} transducing units per milliliter, consisting of 3.6×10^8 primary clones.

Neutravidin (Pierce, 10 μg in 100 μl of 100 mM $NaHCO_3$) was added to individual wells of a 96-well high-binding styrene plate (Costar) and incubated overnight on a rocking platform at 4°C. The neutravidin was removed, and the wells were washed four times with a TBS/Tween solution (50 mM Tris [pH 7.5], 150 mM NaCl, 0.5% Tween-20). Biotinylated D-IQN17 (100 μl of a 10 μM peptide solution in 100 mM $NaHCO_3$) was added to the wells and incubated for 1 hr at 25°C. The biotinylated target was removed, and a blocking solution (30 mg/ml nonfat dried milk in 100 mM $NaHCO_3$, freshly made) was added to the wells and incubated for 2 hr, with rocking, at 4°C. The blocking solution was removed, and the wells were coated again with the biotinylated target as above. The target was removed, and the unliganded neutravidin was blocked by the addition of the blocking solution with 5 mM biotin. After removing the biotin, the wells were washed six times with the TBS/Tween solution. The phage stock (Schumacher et al., 1996) was then added to the wells (50 μl of phage stock plus 50 μl of phage-binding buffer: TBS [50 mM Tris pH 7.5, 150 mM NaCl], 0.1% Tween-20, 1 mg/ml nonfat dried milk, 0.05% sodium azide). The incubation time of the phage stock in the wells decreased with increasing rounds of selection. For the first round, the phage were incubated in the wells for 24 hr, and in the seventh round, they were incubated for 12 hr. After

incubation, the phage solution was removed, and the wells were washed with TBS/Tween to remove the unbound phage. In the first round, the wells were washed six times with no incubation. In following rounds, the wells were washed twelve times; the odd-numbered washes were performed quickly, with no incubation time; even-numbered washes were incubated for increasing amounts of time each round of phage selection (round 2, no incubation; round 3, 3 min; round 4, 5 min; round 5, 7 min; round 6, 10 min, round 7, 12 min). In the first round, phage were eluted by acid elution (0.1 N HCl [pH 2.2], and 1 mg/ml BSA were added to the wells and incubated for 15 min at 4°C). In all subsequent rounds, the phage were eluted by the addition of 2 μg of trypsin (Sigma, T-8658) in 100 μl of phage-binding buffer and 2.5 mM $CaCl_2$ with 1 hr incubation at 37°C. To determine recovery, a dilution of the eluted phage was used to infect K91-kan cells. After a 1 hr incubation, 100 μl of cells was removed and 1:10, 1:100, and 1:1000 dilutions in LB were plated on LB/tetracycline plates. Phage recovery was determined as a ratio of transducing units recovered (the titer of the eluted phage) to the input number of transducing units (the titer of the phage stock used that round). Nonspecific phage recovery generally has a ratio in the order of magnitude of 10^{-8} to 10^{-9}, whereas specifically amplified phage have a ratio of 10^{-7} or greater. Individual clones were amplified and sequenced.

ϕ10-p7 was identified after five rounds, and ϕ10-p1, ϕ10-p3, ϕ10-p4, ϕ10-p5, and ϕ10-p6 after seven rounds of phage selection. The phage selection was performed a second time, with shorter incubation times, using the phage eluted from round 2. The incubation times were shortened to 6 hr for round 3, 2 hr for round 4, and 1 hr for round 5. ϕ10-p10 and ϕ10-p12 were identified after these three additional rounds of selection.

To test the specificity of binding of identified phage clones to the pocket of D-IQN17, the phage clones were added to wells of 96-well plates coated as above with either D-IQN17, D-GCN4-pI_QI′, D-IQN17(G572W), or wells with no target (Figure 3B). The phage were incubated on the plates and washed for the same lengths of time as in the round from which they were identified. Eluted phage were used to infect K91-kan cells, and the recovered transducing units were determined as above. Seven of the ten phage clones were shown to bind specifically to D-IQN17. Two additional phage clones, ϕ10-p10 and ϕ10-p12, were not tested in this binding assay but were expected to bind specifically to D-IQN17 because they shared the same consensus sequence as six of the D-IQN17 specific phage clones.

Peptide Purification

IQN17 (and all variants) and the D-peptides were synthesized on a Perkin Elmer Model 431A peptide synthesizer upgraded with conductivity feedback monitoring. All of the peptides have an acetylated N terminus and a C-terminal amide and were synthesized using Fmoc/HBTU chemistry (Fields et al., 1991) modified with DMSO/NMP resin swelling and acetic anhydride capping. In all syntheses involving D-Ile or D-Thr, which have second chiral centers, the exact mirror images of the naturally occurring L-Ile or L-Thr were used (D-allo-Ile and D-allo-Thr). The peptides were cleaved from the resin using Reagent K (King et al., 1990).

IQN17 contains 29 residues derived from GCN4-pI_QI on the N terminus and 17 residues from the C terminus of N36 on the C terminus. There is a 1 residue overlap between GCN4-pI_QI and the N36 region, making the peptide 45 residues long. To improve solubility, three amino acid substitutions were made in the GCN4-pI_QI region of IQN17, as compared to the original GCN4-pI_QI sequence (Eckert et al., 1998). These substitutions are L13E, Y17K, and H18K. Thus, the sequence of IQN17 is Ac–RMKQIEDKIEEIESKQKKIENEIARIKK LLQLTVWGIKQLQARIL–NH_2 (Ac– represents an N-terminal acetyl group, and –NH_2 represents a C-terminal amide; the HIV portion is underlined.)

For mirror-image phage display, D-IQN17 was synthesized using D-amino acids. Protected D-amino acids were obtained from Peptides International and Bachem Bioscience. The N terminus of the peptide was biotinylated prior to cleavage from the resin, using NHS-LC-biotin II (Pierce, catalog #21336). Between the biotin and the IQN17 sequence was a three–amino acid linker (Gly-Lys-Gly), with the lysine in the naturally occurring L-form for use as a trypsin recognition site. Similarly, D-GCN4-pI_QI′ and D-IQN17(G572W) were synthesized with biotinylated N termini and Gly-Lys-Gly linkers.

D-peptides corresponding to the mirror images of the phage-displayed sequences that bound D-IQN17 were synthesized with two flanking residues on each end. These residues (GA at the N terminus and AA at the C terminus) correspond to the mirror images of phage-encoded sequences that flank the expressed peptide sequences. In addition, in order to improve water solubility for the membrane fusion inhibition studies and NMR studies, multiple D-Lys residues were added at the N termini of the D-peptides. Peptide variants with two N-terminal lysines are denoted with the suffix "-2K," and peptides with four N-terminal lysines are denoted with "-4K."

After cleavage from the resin, peptides were desalted on a Sephadex G-25 column (Pharmacia) in 5% acetic acid, purified by reverse-phase high-performance liquid chromatography (Waters, Inc.) on a Vydac C18 preparative column, using a water-acetonitrile gradient in the presence of 0.1% trifluoroacetic acid, and lyophilized. D-peptides were also air oxidized by dissolving the lyophilized powder in 20 mM Tris (pH 8.2) and stirring at room temperature for several days. Oxidized peptides were then HPLC purified again. The expected molecular weights of all peptides were verified using MALDI-TOF mass spectrometry (PerSeptive Biosystems).

Circular Dichroism

CD experiments were performed on an Aviv 62A DS circular dichroism spectrometer. Measurements from 200 to 260 nm were performed on a 10 μM solution of IQN17 in PBS (50 mM sodium phosphate, 150 mM NaCl [pH 7.4]) in a 10 mm path length cuvette.

Sedimentation Equilibrium

Sedimentation equilibrium experiments were performed on a Beckman XL-A analytical ultracentrifuge using an An-60 Ti rotor. A stock of IQN17 at approximately 200 μM was dialyzed overnight against PBS (pH 7.0), and the concentration was redetermined following dialysis (Edelhoch, 1967). The sample was diluted to 20 μM using the dialysis buffer. The sample was spun at 22,000 and 24,000 rpm; approximately 18 hr after starting, or changing centrifugation speed, the absorbance was recorded at 230 nm.

Cell/Cell Fusion Assay

Inhibition of cell/cell fusion (i.e., syncytia formation) was assayed by coculturing Chinese hamster ovary (CHO) cells expressing HXB2 envelope, tat and rev (Kozarsky et al., 1989), and HeLa-CD4-LTR-Beta-gal cells (M. Emerman, National Institutes of Health AIDS Reagent Program) in the presence of varying concentrations of D-peptide. When these cells are mixed in the absence of an inhibitor, they form syncytia, or multinucleated cells, which express β-galactosidase (Kimpton and Emerman, 1992). Peptide stocks for these assays were prepared at approximately 20 mM in DMSO, and the precise concentrations were determined using tyrosine, tryptophan, and cysteine absorbance at 280 nm in 6 M GuHCl (Edelhoch, 1967). These stocks were diluted in media to the concentrations required for the experiment (ranging from 2.5 μM to 200 μM); the final concentration of DMSO was kept constant at 1%. The peptide sample was added to wells of an 8-well permanox chamber slide (Lab-Tek), and an estimated 2×10^4 CHO cells and 4×10^4 HeLa cells were then added. Approximately 20 hr after coculturing the cells, the monolayers were stained with 5-bromo-4-chloro-3-indolyl-β-D-galactoside to detect the syncytia, which were visualized by microscope and counted manually. (The entire well was counted; a syncytium was scored as a fused cell containing three or more nuclei.) The IC_{50} was calculated from fitting the data to a Langmuir equation [$y = k/(1 + [\text{peptide}]/IC_{50})$], where y = number of syncytia, and k is a scaling constant.

HIV Infectivity Assay

The ability of the D-peptides to inhibit HIV-1 infection was assayed using a recombinant luciferase-encoding HIV-1 (Chen et al., 1994). As in the cell/cell fusion assays, peptide stocks were prepared in DMSO. The final concentration of DMSO was kept constant at 1%. The virus was produced by cotransfecting an envelope-deficient HIV-1 genome NL43LucR^-E^- (Chen et al., 1994) and the HXB2 gp160

expression vector pCMVHXB2 gp160 (D. C. Chan and P. S. K., unpublished data; see Chan et al., 1998) into 293T cells. Low-speed centrifugation was used to clear the viral supernatants of cellular debris. The supernatant was used to infect HOS-CD4/Fusin cells (N. Landau, National Institutes of Health AIDS Reagent Program) in the presence of the D-peptides with concentrations ranging from 0 to 500 μM. Forty-eight hours postinfection, the cells were harvested, and luciferase activity was monitored using a Wallac AutoLumat LB953 luminometer (Gaithersburg, MD). The IC_{50} was calculated from fitting the data to a Langmuir equation [$y = k/(1 + [peptide]/IC_{50})$], where y = luciferase activity, and k is a scaling constant.

Crystallization

A 10 mg/ml stock (total peptide concentration) of a mixture of IQN17 and D10-p1 was prepared in water. The final concentration of IQN17 was about 1.4 mM, and the final concentration of D10-p1 was about 1.5 mM. Initial crystallization conditions were found using Crystal Kits I and II (Hampton Research) and then optimized. The best diffracting crystals grew from 1 μl of the stock added to 1 μl of the reservoir buffer (10% PEG 4000, 0.1 M sodium citrate [pH 5.6], 20% 2-propanol) and allowed to equilibrate against the reservoir buffer. Crystals belong to the space group P321 ($a = b = 41.83$ Å; $c = 84.82$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$) and contain one IQN17/D10-p1 monomer in the asymmetric unit. A useful heavy atom derivative was produced by increasing the concentration of PEG 4000 in the reservoir solution by 4%, adding $(NH_4)_2OsCl_6$ to the reservoir solution to a final concentration of 5 mM and adding 5 μl of the resulting solution to the drop containing the protein crystal. Prior to data collection, native and heavy atom derivative crystals were transferred into cryosolution containing 20% PEG 4000, 0.1 M sodium citrate (pH 5.6), 20% 2-propanol and flash-frozen using an X-stream cryogenic crystal cooler (Molecular Structure Corporation).

X-Ray Data Collection and Processing

Initial data were collected on a Rigaku RU300 rotating anode X-ray generator mounted to an R-AXIS IV area detector (Molecular Structure Corporation). Final native and multiwavelength anomalous dispersion (MAD) data sets for the IQN17/D10-p1 complex were collected at the Howard Hughes Medical Institute Beamline X4A at Brookhaven National Laboratory using an R-AXIS IV detector. For MAD data, four wavelengths near the osmium KIII absorption edge were selected based on the fluorescence spectrum of the Os derivative crystal (Table 1). The four wavelengths were 1.1398 Å, 1.1403 Å, 1.1393 Å, and 1.1197 Å. Data sets were collected in 20° batches, allowing each batch to be collected at each wavelength before moving to the next, in order to minimize crystal decay between data sets. Reflections were integrated and scaled with the programs DENZO and SCALEPACK (Otwinowski, 1993).

Further diffraction data processing, phase determination, and map calculations were performed using the CCP4 suite of programs (CCP4, 1994). Intensities were reduced to amplitudes with the program TRUNCATE, and the data sets for the wavelengths closest to the Os KIII absorption edge (λ1, λ2, λ3) were scaled with SCALEIT to the remote wavelength (λ4) data set (Table 1).

Phase Determination and Crystallographic Refinement

Phase determination for IQN17/D10-p1 was initially attempted with the molecular replacement technique, using a replacement model built from the published GCN4-pI_QI and HIV-1 gp41 N36/C34 structures (Chan et al., 1997; Eckert et al., 1998), with side chains truncated to generate a polyserine chain. The resultant molecular replacement solutions were ambiguous, and the electron density map did not reveal the conformation of the D10-p1 peptide. However, the molecular replacement phases were sufficient to locate a single Os atom in the corresponding derivative using difference and anomalous Fourier maps. The heavy atom binds on the crystallographic three-fold axis (0.333, 0.667, 0.047). MAD phases were then generated with the program MLPHARE (Table 1) and extended to higher resolution with the program DM (CCP4, 1994). Electron density map interpretation and model building were done with the program O (Jones et al., 1991). The structure of the IQN17/D10-p1 complex was refined at 1.5 Å resolution using the program CNS (Brünger et al., 1998). The correctness of the structure was checked with simulated annealing omit maps and with the program WHAT-CHECK (Hooft et al., 1996).

All residues of IQN17 and of D10-p1 (when converted into its mirror image) occupy most preferred areas of the Ramachandran plot. The conformations of the majority of the residues are well defined except for the two most N-terminal residues of IQN17 and the side chains of Arg-6 and Arg-8 of D10-p1. As in the original GCN4–pI_QI structure (Eckert et al., 1998), a chloride ion is bound in the coiled-coil core near the side chains of Gln-16. Its temperature factor ($B = 45.3$ Å^2) is higher than those of the interacting nitrogen atoms of Gln-16 ($B = 32.7$ Å^2) and indicates only partial occupancy.

NMR

^{1}H NMR experiments were performed at 25°C on a Bruker AMX 500 spectrometer. Data were processed with Felix 98.0 (Molecular Simulations, Inc.) on Silicon Graphics computers, and all spectra were referenced to DSS. All samples were dissolved in 100 mM NaCl, 50 mM sodium phosphate (pH 7.5); the buffers used were >99.7% D_2O, to remove overlapping resonances from solvent-exchangeable backbone and side chain protons. Solute concentrations ranged from 0.3 to 1.0 mM for individual peptides, and from 0.8 to 1.0 mM for 1:1 complexes of IQN17 with each D-peptide. Several samples of IQN17/D-peptide complexes showed slight precipitation over time. 2D NOESY and TOCSY experiments were performed as described (Cavanagh et al., 1996) on samples of IQN17 and of each complex, with mixing times of 55 ms (NOESY) and 42 ms (TOCSY). Spectral widths of 11,111 Hz and 5,555 Hz were used in the acquisition (t_2) and indirect (t_1) dimensions, respectively. TOCSY experiments employed the DIPSI-2rc mixing sequence (Cavanagh and Rance, 1993). Chemical shift predictions were made using the program SHIFTS (version 3.02b, D. Sitkoff, K. Osapay, and D. A. Case; The Scripps Research Institute).

Acknowledgments

We thank Heng Chhay for performing antiviral assays; Michael Burgess, James Pang, and Ben Sanford for peptide synthesis; Alexandra Evindar for editorial and graphics assistance; and Dr. Craig Ogata and Dr. Xun Zhao for assistance at the synchrotron beamline. We also thank Dr. Ton Schumacher for providing the phage library; Michael Milhollen, David Akey, and Brian Schneider for technical advice; Dr. David Chan and Dr. David Lee for helpful suggestions; and members of the Kim lab for stimulating discussions and other contributions. This research was funded by the National Institutes of Health (PO1 GM56552) and utilized the W. M. Keck Foundation X-ray Crystallography Facility (Whitehead Institute) and the Howard Hughes Medical Institute beamline (X4A) at the National Synchrotron Light Source (Brookhaven National Laboratory).

Received August 19, 1999; revised September 7, 1999.

References

Berger, E.A., Murphy, P.M., and Farber, J.M. (1999). Chemokine receptors as HIV-1 coreceptors: role in viral entry, tropism and disease. Annu. Rev. Immunol. *17*, 657–700.

Blacklow, S.C., Lu, M., and Kim, P.S. (1995). A trimeric subdomain of the simian immunodeficiency virus envelope glycoprotein. Biochemistry *34*, 14955–14962.

Bovey, F.A. (1988). Nuclear Magnetic Resonance Spectroscopy (San Diego, CA: Academic Press).

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

Cao, J., Bergeron, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. (1993). Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. J. Virol. *72*, 2747–2755.

Cavanagh, J., and Rance, M. (1993). Suppression of cross-relaxation

effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. J. Magn. Reson. Ser. A *105*, 328.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., and Skelton, N.J. (1996). Protein NMR Spectroscopy: Principles and Practice (San Diego, CA: Academic Press).

CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta. Crystallogr. D *50*, 760–763.

Chan, D.C., and Kim, P.S. (1998). HIV entry and its inhibition. Cell *93*, 681–684.

Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. Cell *89*, 263–273.

Chan, D.C., Chutkowski, C.T., and Kim, P.S. (1998). Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. Proc. Natl. Acad. Sci. USA *95*, 15613–15617.

Chen, S.S.-L., Lee, C.-N., Lee, W.-R., McIntosh, K., and Lee, T.-H. (1993). Mutational analysis of the leucine zipper-like motif of the human immunodeficiency virus type 1 envelope transmembrane glycoprotein. J. Virol. *67*, 3615–3619.

Chen, B.K., Saksela, K., Andino, R., and Baltimore, D. (1994). Distinct modes of human immunodeficiency virus type 1 proviral latency revealed by superinfection of nonproductively infected cell lines with recombinant luciferase-encoding viruses. J. Virol. *68*, 654–660.

Chen, C.-H., Matthew, T.J., McDanal, C.B., Bolgnesi, D.P., and Greenberg, M.L. (1995). A molecular clasp in the human immunodeficiency virus (HIV) type 1 TM protein determines the anti-HIV activity of gp41 derivatives: implication for viral fusion. J. Virol. *69*, 3771–3777.

Crick, F.H. C. (1953). The packing of α-helices: simple coiled coils. Acta. Crystallogr. *6*, 689–697.

Dubay, J.W., Roberts, S.J., Brody, B., and Hunter, E. (1992). Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. J. Virol. *66*, 4748–4756.

Eckert, D.M., Malashkevich, V.N., and Kim, P.S. (1998). Crystal structure of GCN4-pI_{Q}I, a trimeric coiled coil with buried polar residues. J. Mol. Biol. *284*, 859–865.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry *6*, 1948–1954.

Fields, C.G., Lloyd, D.H., Macdonald, R.L., Otteson, K.M., and Noble, R.L. (1991). HBTU activation for automated Fmoc solid-phase peptide synthesis. Pept. Res. *4*, 95–101.

Furuta, R.A., Wild, C.T., Weng, Y., and Weiss, C.D. (1998). Capture of an early fusion-active conformation of HIV-1 gp41. Nat. Struct. Biol. *5*, 276–279.

Harbury, P.B., Tidor, B., and Kim, P.S. (1995). Repacking protein cores with backbone freedom: structure prediction for coiled coils. Proc. Natl. Acad. Sci. USA *92*, 8408–8412.

Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51–58.

Herskowitz, I. (1987). Functional inactivation of genes by dominant negative mutations. Nature *329*, 219–222.

Hooft, R.W., Vriend, G., Sander, C., and Abola, E.E. (1996). Errors in protein structures. Nature *381*, 272.

Jiang, S., Lin, K., Strick, N., Neurath, A.R. (1993). HIV-1 inhibition by a peptide. Nature *365*, 113.

Jones, T.A., Zou, J.W., Cowan, S., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these methods. Acta. Crystallogr. D *47*, 110–119.

Jones, P.L., Korte, T., and Blumenthal, R. (1998). Conformational changes in cell surface HIV-1 envelope glycoproteins are triggered by cooperation between cell surface CD4 and co-receptors. J. Biol. Chem. *273*, 404–409.

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). Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. Nat. Med. *4*, 1302–1307.

Kimpton, J., and Emerman, M. (1992). Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. J. Virol. *66*, 2232–2239.

King, D.S., Fields, C.G., and Fields, G.B. (1990). A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. Int. J. Pep. Prot. Res. *36*, 255–266.

Kozarsky, K., Penman, M., Basiripour, L., Haseltine, W., Sodroski, J., and Krieger, M. (1989). Glycosylation and processing of the human immunodeficiency virus type 1 envelope protein. J. Acquir. Immune Defic. Syndr. *2*, 163–169.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature *393*, 648–659.

LaCasse, R.A., Follis, K.E., Trahey, M., Scarborough, J.D., Littman, D.R., and Nunberg, J.H. (1999). Fusion-competent vaccines: broad neutralization of primary isolates of HIV. Science *283*, 357–362.

Lu, M., and Kim, P.S. (1997). A trimeric structural subdomain of the HIV-1 transmembrane glycoprotein. J. Biol. Struct. Dyn. *15*, 465–471.

Lu, M., Blacklow, S.C., and Kim, P.S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nat. Struct. Biol. *2*, 1075–1082.

Malim, M.H., Hauber, J., Le, S.-Y., Maizel, J.V., and Cullen, B.R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature *338*, 254–257.

Munoz-Barroso, I., Durell, S., Sakaguchi, K., Appella, E., and Blumenthal, R. (1998). Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41. J. Cell Biol. *140*, 315–323.

Osapay, K., and Case, D.A. (1991). A new analysis of proton chemical shifts in proteins. J. Am. Chem. Soc. *113*, 9436–9444.

Otwinowski, Z. (1993). Oscillation data reduction program. In Data Collection and Processing, L. Sawer, N. Isaacs, and S. Bailey, eds. (Warrington, UK: SERC Daresbury Laboratory), pp. 55–62.

Rimsky, L.T., Shugars, D.C., and Matthews, T.J. (1998). Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. J. Virol. *72*, 986–993.

Schumacher, T.N.M., Mayr, L.M., Minor, D.L., Milhollen, M.A., Burgess, M.W., and Kim, P.S. (1996). Identification of D-peptide ligands through mirror-image phage display. Science *271*, 1854–1857.

Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.

Singh, M., Berger, B., and Kim, P.S. (1999). LearnCoil-VMF: Computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins. J. Mol. Biol. *290*, 1031–1041.

Tan, K., Liu, J., Wang, J.-H., Shen, S., and Lu, M. (1997). Atomic structure of a thermostable subdomain of HIV-1 gp41. Proc. Natl. Acad. Sci. USA *94*, 12303–12308.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. Nature *387*, 426–430.

Weng, Y., and Weiss, C.D. (1998). Mutational analysis of residues in the coiled-coil domain of human immunodeficiency virus type 1 transmembrane protein gp41. J. Virol. *72*, 9676–9682.

Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B., and Matthews, T.J. (1994). Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. Proc. Natl. Acad. Sci. USA *91*, 12676–12680.

Wild, C.T., Greenwell, T., Shugars, D., Rimsky-Clarke, L., and Matthews, T. (1995). The inhibitory activity of an HIV type 1 peptide correlates with its ability to interact with a leucine zipper structure. AIDS Res. Hum. Retroviruses *11*, 323–325.

Wrighton, N.C., Farrell, F.X., Chang, R., Kashyap, A.K., Barbone, F.P., Mulcahy, L.S., Johnson, D.L., Barrett, R.W., Jolliffe, L.K., and Dower, W.J. (1996). Small peptides as potent mimetics of the protein hormone erythropoietin. Science *273*, 458–463.

Zapp, M.L., and Green, M.R. (1989). Sequence-specific RNA binding by the HIV-1 Rev protein. Nature *342*, 714–716.

Protein Data Bank ID Code

The coordinates for the D10-p1/IQN17 complex have been deposited with the ID code 1czq; they are available immediately at the website <http://web.wi.mit.edu/kim>.