of colipase and angiogenin-3 mRNAs in these ex–germ-free mice were comparable to those of age-matched mice conventionally raised since birth (Fig. 2). In contrast, the response of sprr2a to colonization depended on the colonizing species: *B. infantis* and *E. coli* produced only small increases (Fig. 2). Mdr1a and glutathione *S*-transferase also exhibited species-specific responses. *Bacteroides thetaiotaomicron* suppressed and *E. coli* and *B. infantis* stimulated expression of both genes, whereas the multicomponent ileal/cecal flora produced no significant (i.e., ≥twofold) change in levels of either mRNA compared with germ-free controls. The differing Mdr1a/GST responses suggest that variations in xenobiotic metabolism between individuals may arise, in part, from differences in their resident gut flora.

The only *B. thetaiotaomicron* genes currently known to link changes in bacterial metabolism with host responses are those involved in fucose utilization (*3*). Transposon-mediated mutagenesis of *fucI* (encoding fucose isomerase) blocks the organism's ability to use fucose as a carbon source and to signal fucosylated glycan production in the ileal epithelium (*3*). Microarray analysis of the host response to colonization revealed no appreciable differences between isogenic mutant and wild-type strains. This similarity extends to all genes in Table 1. Future identification of microbial factors that interlink microbial and host physiology will require characterization of changes in *B. thetaiotaomicron* gene expression as a function of colonization.

In summary, the studies described above provide a broad-based in vivo characterization of transcriptional responses to colonization with a prototypic gut commensal. Our results reveal that commensals are able to modulate expression of host genes that participate in diverse and fundamental physiological functions. The species selectivity of some of the colonization-associated changes in gene expression emphasizes how our physiology can be affected by changes in the composition of our indigenous microflora. The fusion of germ-free technology, functional genomics, and LCM/qRT-PCR makes it possible to use in vivo systems to quantify the impact of a microbial population on host cell gene expression.

References and Notes

1. W. E. C. Moore, L. V. Holdeman, *Appl. Microbiol.* **27**, 961 (1974).
2. T. Ushijima, M. Takahashi, K. Tatewaki, Y. Ozaki, *Microbiol. Immunol.* **27**, 985 (1983).
3. L. V. Hooper, J. Xu, P. G. Falk, T. Midtvedt, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9833 (1999).
4. Age-matched groups of 7- to 15-week-old germ-free NMRI/KI mice were maintained in plastic gnotobiotic isolators on a 12-hour light cycle and given free access to an autoclaved chow diet (B&K Universal). Males were inoculated with *B. thetaiotaomicron* strain VPI-5482 (*3*), the isogenic Fu-4 strain lacking a functional *fucI* gene (*3*), *E. coli* K12, or *Bifidobacterium infantis* (ATCC15697). Mice were killed 10 days later, 2 hours after lights were turned on. The distal 1 cm of the small intestine was used to define CFU/ml ileal contents. The 3 cm of intestine just proximal to this segment was used to isolate total ileal RNA (Qiagen RNeasy kit).
5. L. Bry, P. G. Falk, T. Midtvedt, J. I. Gordon, *Science* **273**, 1380 (1996).
6. Total ileal RNA samples, prepared from four mice from three independent colonizations and from age-matched male germ-free controls ($n = 8$), were each pooled in equal amounts for generation of biotinylated cRNA targets. Two targets were prepared, independently, from 30 μg of each total cellular RNA pool (*30*). Each cRNA was hybridized to Mu11K and Mu19K chip sets according to Affymetrix protocols. The overall fluorescence intensity across each chip was scaled to a target intensity of 150. Pairwise comparisons of "germ-free" versus "colonized" expression levels were performed. A ≥twofold difference was considered significant if three criteria were met: The GeneChip software returned a difference call of Increased or Decreased, the mRNA was called Present by GeneChip software in either germ-free or colonized RNA, and the difference was observed in duplicate microarray hybridizations.
7. See supplemental material at http://gordonlab.wustl.edu.
8. B. S. Wostmann, C. Larkin, A. Moriarty, E. Bruckner-Kardoss, *Lab. Anim. Sci.* **33**, 46 (1983).
9. SYBR Green-based qRT-PCR (*31*) used gene-specific primers (*7*) and deoxyribonuclease-treated RNAs. Control experiments established that amplicons were derived from cDNA and not from primer-dimers or genomic DNA. Signals were normalized to glyceraldehyde 3-phosphate dehydrogenase mRNA. Normalized data were used to quantitate relative levels of a given mRNA in germ-free and colonized ileums ($\Delta\Delta C_T$ analysis; User Bulletin #2, ABI Prism 7700 Sequence Detection System).
10. S. Kersten *et al.*, *J. Biol. Chem.* **275**, 28488 (2000).
11. Seven-micrometer-thick sections were cut from ileums harvested immediately after the death of three sets of germ-free and colonized mice that were independent of those used for microarray profiling. LCM was conducted with the PixCell II system (Arcturus, 7.5-μm-diameter laser spot) and RNA was prepared (Stratagene RNA Micro-Isolation Kit; protocols in http://dir.nichd.nih.gov/lcm/protocol.htm).
12. M. E. Lowe, M. H. Kaplan, L. Jackson-Grusby, D. D'Agostino, M. J. Grusby, *J. Biol. Chem.* **273**, 31215 (1998).
13. P. G. Falk, L. V. Hooper, T. Midtvedt, J. I. Gordon, *Microbiol. Mol. Biol. Rev.* **62**, 1157 (1998).
14. A. J. Macpherson *et al.*, *Science* **288**, 2222 (2000).
15. R. C. DeLisle, M. Petitt, K. S. Isom, D. Ziemer, *Am. J. Physiol.* **275**, G219 (1998).
16. L. Thim, E. Mørtz, *Regul. Pept.* **90**, 61 (2000).
17. P. M. Steinert, L. N. Marekov, *Mol. Biol. Cell* **10**, 4247 (1999).
18. R. W. Johnstone, A. A. Ruefli, M. J. Smyth, *Trends Biochem. Sci.* **25**, 1 (2000).
19. M. Ingelman-Sundberg, M. Oscarson, R. A. McLellan, *Trends Pharmacol. Sci.* **20**, 342 (1999).
20. E. Husebye, P. M. Hellstrom, T. Midtvedt, *Dig. Dis. Sci.* **39**, 946 (1994).
21. P. A. Skehel, R. Fabian-Fine, E. R. Kandel, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1101 (2000).
22. M. T. Liu, J. D. Rothstein, M. D. Gershon, A. L. Kirchgessner, *J. Neurosci.* **17**, 4764 (1997).
23. S. D. Krasinski *et al.*, *Am. J. Physiol.* **267**, G584 (1994).
24. G. D. Lux, L. J. Marton, S. B. Baylin, *Science* **210**, 195 (1980).
25. J. M. Chinsky *et al.*, *Differentiation* **42**, 172 (1990).
26. J. Nilsson, S. Koskiniemi, K. Persson, B. Grahn, I. Holm, *Eur. J. Biochem.* **250**, 223 (1997).
27. X. Fu, M. P. Kamps, *Mol. Cell Biol.* **17**, 1503 (1997).
28. X. Fu, W. G. Roberts, V. Nobile, R. Shapiro, M. P. Kamps, *Growth Factors* **17**, 125 (1999).
29. V. Nobile, B. L. Vallee, R. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4331 (1996).
30. C. K. Lee, R. G. Klopp, R. Weindruch, T. A. Prolla, *Science* **285**, 1390 (1999).
31. N. Steuerwald, J. Cohen, R. J. Herrera, C. A. Brenner, *Mol. Hum. Reprod.* **5**, 1034 (1999).

32. We thank C. Hong, E. Skogman, and B. Basta for technical assistance. This work was funded by NIH grant DK30292 and AstraZeneca.

16 October 2000; accepted 4 January 2001

Protein Design of an HIV-1 Entry Inhibitor

Michael J. Root, Michael S. Kay, Peter S. Kim*†

Human immunodeficiency virus type–1 (HIV-1) membrane fusion is promoted by the formation of a trimer-of-hairpins structure that brings the amino- and carboxyl-terminal regions of the gp41 envelope glycoprotein ectodomain into close proximity. Peptides derived from the carboxyl-terminal region (called C-peptides) potently inhibit HIV-1 entry by binding to the gp41 amino-terminal region. To test the converse of this inhibitory strategy, we designed a small protein, denoted 5-Helix, that binds the C-peptide region of gp41. The 5-Helix protein displays potent (nanomolar) inhibitory activity against diverse HIV-1 variants and may serve as the basis for a new class of antiviral agents. The inhibitory activity of 5-Helix also suggests a strategy for generating an HIV-1 neutralizing antibody response that targets the carboxyl-terminal region of the gp41 ectodomain.

Infection by HIV-1, the virus that causes AIDS, requires fusion of the viral and cellular membranes (*1–3*). This membrane-fusion process is mediated by the viral envelope glycoprotein complex (gp120/gp41) and receptors on the target cell. Binding of gp120/gp41 to cell-surface receptors (CD4 and a coreceptor, such as CCR5 or CXCR4) triggers a series of conformational changes in the gp120/gp41 oligomer that ultimately lead to formation of a trimer-of-hairpins structure in gp41 (Fig. 1A).

The trimer-of-hairpins is a common structural element involved in the fusion process of many enveloped viruses, suggesting a critical role for this motif in promoting mem-

brane fusion (*3–5*). In HIV-1 gp41, the core of the trimer-of-hairpins is a bundle of six α-helices (Fig. 1B): three α-helices (formed by the COOH-terminal regions of three gp41 ectodomains) pack in an antiparallel manner against a central, three-stranded coiled coil (formed by the NH_2-terminal regions of the gp41 molecules) (*6–9*). The fusion peptide region, which inserts into the cellular membrane, is located at the extreme NH_2-terminus of gp41, and the COOH-terminal region is adjacent to the transmembrane helix anchored in the viral membrane. Thus, the trimer-of-hairpins motif brings the two membranes together (Fig. 1A).

Agents that interfere with formation of the gp41 trimer-of-hairpins structure can inhibit HIV-1 infection. Peptides derived from the COOH-terminal region of the gp41 ectodomain, referred to as C-peptides (corresponding to the outer helices of the six-helix bundle), are effective inhibitors of HIV-1 infection (*10–12*). Studies from several groups support a mechanism of dominant-negative inhibition in which C-peptides bind to a transient gp41 species known as the prehairpin intermediate (Fig. 1A) (*1, 6, 12–15*). In this prehairpin intermediate, the gp41 fusion peptide is embedded in the target-cell membrane, exposing the NH_2-terminal three-stranded coiled coil [compare (*16*)]. Binding of C-peptides to the NH_2-terminal region of the prehairpin structure prevents formation of the gp41 trimer-of-hairpins, ultimately leading to irreversible loss of membrane-fusion activity. C-peptides potently inhibit HIV-1 entry, with a mean inhibitory concentration (IC_{50}) as low as 1 nM in vitro (*11, 12*). One such C-peptide is in clinical trials and shows antiviral activity in humans (*2, 17*). More recently, efforts to target a prominent pocket on the surface of the NH_2-terminal coiled coil of the prehairpin intermediate have led to the discovery of small, cyclic D-peptides that inhibit HIV-1 infection, thereby validating the pocket as a potential target for development of small, orally bioavailable HIV-1 entry inhibitors (*18*).

The importance of trimer-of-hairpins formation for HIV-1 entry leads to the hypothesis that the COOH-terminal region on gp41 might also serve as a target for potential membrane-fusion inhibitors (Fig. 1A). If the COOH-terminal region is accessible (at least transiently) before formation of the trimer-of-hairpins, then agents that bind to this region of gp41 may prevent fusion. Consistent with this notion (*6*), peptides derived from the gp41 NH_2-terminal region (referred to as N-peptides) are modest inhibitors of HIV-1 membrane fusion (micromolar IC_{50}) (*6, 19*). The inhibitory mechanism of N-peptides, however, has not been ascertained, in part because these peptides have a strong tendency to aggregate. Indeed, a plausible alternative mechanism of action for the N-peptides is that they intercalate into the gp41 NH_2-terminal coiled coil, thereby disrupting the trimeric interface (*19, 20*).

To directly test the hypothesis that the C-peptide region of gp41 is a potential target for the inhibition of HIV-1 entry, we designed a protein that binds tightly and specifically to this site. The design takes advantage of the binding properties of the N-peptide coiled coil while minimizing the tendency of the N-peptides to aggregate. In this designed protein, denoted 5-Helix, five of the six helices that make up the core of the gp41 trimer-of-hairpins structure are connected with short peptide linkers (Fig. 1C) (*21*). The 5-Helix protein lacks a third C-peptide helix, and this vacancy is expected to create a high-affinity binding site for the COOH-terminal region of gp41.

Under physiological conditions, 5-Helix is well folded, soluble, and extremely stable, with an α-helical content in close agreement with the value predicted from the design (Fig. 2, A and B). In affinity-interaction experiments, 5-Helix interacts strongly and specifically with epitope-tagged C-peptides (Fig. 2C). Moreover, this interaction induces a helical conformation in the bound C-peptide, as judged by the difference in circular dichroism (CD) before and after mixing (Fig. 2D). These properties are consistent with the intended design of 5-Helix.

The 5-Helix protein potently inhibits HIV-1 membrane fusion (nanomolar IC_{50}), as measured by viral infectivity and cell-cell fusion assays (Fig. 3, A and B). In contrast, a control protein, denoted 6-Helix, in which the C-peptide binding site is occupied by an attached C-peptide (i.e., all six helices that constitute the gp41 trimer-of-hairpins have been linked into a single polypeptide) (*21, 22*), does not have appreciable inhibitory activity (Fig. 3A). Likewise, a 5-Helix variant,

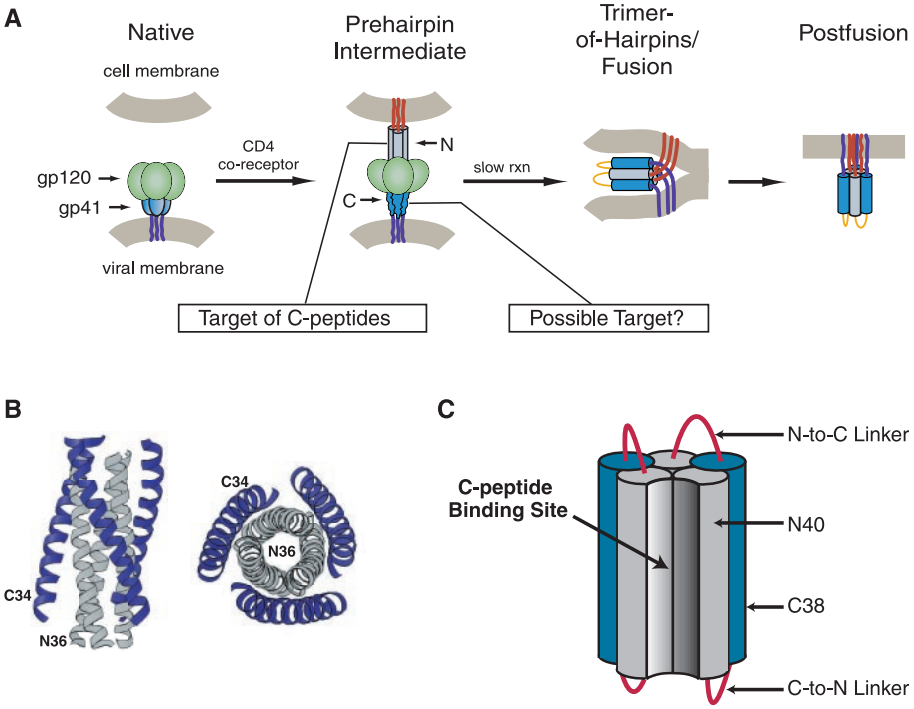


Fig. 1. Targeting HIV-1 membrane fusion. (**A**) A schematic of HIV-1 membrane fusion depicting events that promote formation of the gp41 trimer-of-hairpins [adapted from (*1*)]. The NH_2-terminal fusion peptide of gp41 (red), inaccessible in the native state, inserts into target cell membranes following gp120 interaction with CD4 and coreceptors. Formation of the prehairpin intermediate exposes the NH_2-terminal coiled coil (gray), the target of C-peptide inhibition. This transient structure collapses into the trimer-of-hairpins state that brings the membranes into close apposition for fusion. (**B**) Lateral (left) and axial (right) views of a ribbon diagram representing the core of the gp41 trimer-of-hairpins. The ribbon diagram is derived from the crystal structure of a six-helix bundle formed by N36 (N-peptide, gray) and C34 (C-peptide, blue) (*7*). (**C**) A schematic model of the designed protein 5-Helix. Three N-peptide segments (N40, gray) and two C-peptide segments (C38, blue) are alternately linked (N-C-N-C-N) using short Gly/Ser peptide sequences (red loops) (*21*). The sequences of each segment in single-letter amino acid code are: N40, QLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILA; C38, HTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLE; N-to-C linker, GGSGG; and C-to-N linker, GSSGG.

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142, USA.

*To whom correspondence should be addressed. E-mail: kimadmin@wi.mit.edu

†Present address: Merck Research Laboratories, 770 Sumneytown Pike, West Point, PA 19486, USA.

denoted 5-Helix(D4), in which the C-peptide binding site is disrupted by mutation of four interface residues (V549, L556, Q563, and V570) to Asp (*23*), does not block membrane fusion even at 1 μM (Fig. 3A). We conclude that C-peptide binding is the key determinant of antiviral activity in 5-Helix.

The inhibitory activities of 5-Helix and C-peptides are expected to be antagonistic: when 5-Helix binds C-peptide, the amino acid residues thought to be responsible for the antiviral activities of each inhibitor are buried in the binding interface. Indeed, mixtures of 5-Helix and C34 [a well-characterized and potent C-peptide inhibitor with an $IC_{50} \approx 1$ nM (*12*)] display a dose-dependent antagonistic effect (Fig. 3B). In the presence of 5-Helix, high-potency inhibition by C34 is only observed when the peptide is in stoichiometric excess (Fig. 3B).

The 5-Helix protein inhibits infection by viruses pseudotyped with a variety of HIV-1 envelope proteins (from clades A, B, and D) with similar potency (Fig. 3C). This broad-spectrum inhibition likely reflects the highly conserved interface between the NH_2- and COOH-terminal regions within the gp41 trimer-of-hairpins structure (Fig. 4). The residues in the C-peptide region of gp41 that are expected to make contact with 5-Helix are highly conserved in HIV-1, HIV-2, and simian immunodeficiency virus (SIV) (Fig. 4).

As a potent, broad-spectrum inhibitor of viral entry, 5-Helix may serve as the basis for development of a new class of therapeutic agents against HIV-1 (*24*). Moreover, 5-Helix offers flexibility in the design of variants with better therapeutic characteristics. In principle, 5-Helix can be modified extensively to alter its immunogenic, antigenic, bioavailability, or inhibitory properties (*25*). For example, the C-peptide binding site might be lengthened, shortened, or shifted in the gp41 sequence in order to optimize inhibitory potency by targeting different regions of the gp41 ectodomain.

It would be desirable to generate neutralizing antibodies that mimic the binding properties of 5-Helix. Unstructured C-peptide immunogens may not elicit broadly neutralizing antibodies, because the linear sequence of the gp41 C-peptide region is variable among different HIV-1 strains. The potent and broad-spectrum inhibitory properties of 5-Helix suggest that an HIV-1 neutralizing antibody response might be generated using C-peptide analogs constrained in a helical conformation (i.e., as in the C-peptide region when it binds to 5-Helix).

Interestingly, the epitope for 2F5, a human monoclonal antibody directed against gp41 with broad neutralizing activity, is located immediately COOH-terminal to the C-peptide region targeted by 5-Helix (*26–28*). It is unknown if 2F5 inhibits infection by interfering with trimer-of-hairpins formation. The conformation of the 2F5-

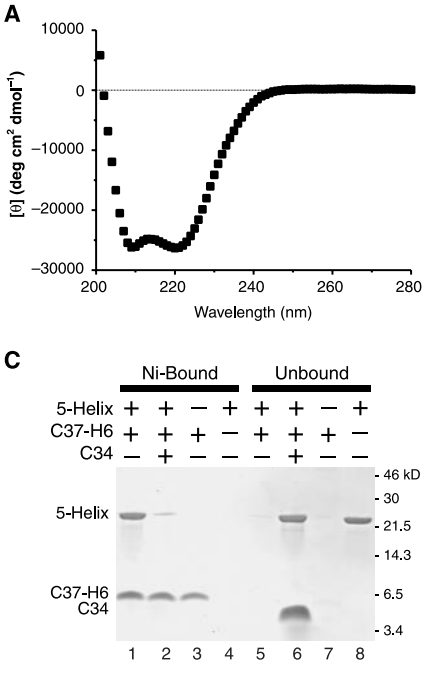
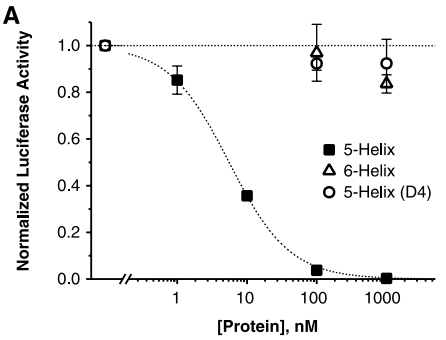


Fig. 2. Properties of 5-Helix (*41*). (**A**) CD spectrum of 5-Helix (10 μM) at 25°C. The spectrum indicates that the 5-Helix protein adopts >95% of the helical content expected from the design. (**B**) Thermal denaturation of 5-Helix monitored by ellipticity at 222 nm in TBS (filled squares) and in 3.7 M GuHCl/TBS (open circles). The denaturation observed in the GuHCl solution is >90% reversible. (**C**) Nickel-NTA precipitation of 5-Helix with a His-tagged C-peptide. Untagged 5-Helix and His-tagged C-peptide [denoted C37-H6 (*39*)] were mixed before Ni-NTA agarose was added in order to precipitate complexes containing C37-H6 (lanes 1 and 5; lanes numbered from left to right). Addition of excess untagged C-peptide (C34) shifts the 5-Helix molecules from the bound to the unbound fraction (lanes 2 and 6). (**D**) CD spectra of 5-Helix and C37-H6 before (filled squares) and after (open circles) mixing in a mixing cuvette (*41*). The increase in ellipticity at 222 nm upon mixing indicates an interaction between the two species that increases the total helical content (corresponding to an additional 28 helical residues per associated C-peptide).



C

Isolate	Clade	Coreceptor	IC_{50} (nM)
HXB2	B	CXCR4	1.9 ± 0.7
UG024.2	D	CXCR4	1.3 ± 0.2
JRFL	B	CCR5	5.6 ± 0.7
RW020.5	A	CCR5	5.9 ± 2.7

Fig. 3. Inhibition of HIV-1 envelope-mediated membrane fusion by 5-Helix (*42*). (**A**) Titration of viral infectivity by 5-Helix (filled squares), 6-Helix (open triangles), and 5-Helix(D4) [open circles (*23*)]. The data represent the mean ± SEM of two or more separate experiments. (**B**) Antagonistic inhibitory activities of 5-Helix and C34. The number of syncytia were measured in a cell-cell fusion assay performed in the absence or presence of 5-Helix, C34, or mixtures of 5-Helix and C34 at the indicated concentrations. The IC_{50} values for 5-Helix and C34 in this assay are 13 ± 3 nM and 0.55 ± 0.03 nM (*12*), respectively. Data represent the mean and range of mean of duplicate measurements, except for the control (mean ± SEM of five measurements). (**C**) Shown is 5-Helix inhibition of pseudotyped virus containing different HIV-1 envelope glycoproteins. The reported IC_{50} values represent the mean ± SEM of three independent experiments.

bound epitope has recently been shown to exist in a hairpin turn (*29*). Although antibodies elicited with fragments of gp41 containing this sequence do not possess significant virus-neutralizing activity (*30, 31*), it is possible that constrained analogs (perhaps with an adjacent helical C-peptide region) will lead to useful immunogens in efforts to develop an AIDS vaccine.

Alternatively, 5-Helix itself is a potential vaccine candidate. The possibility of eliciting an antibody response against transiently exposed conformations of proteins involved in HIV-1 fusion has been suggested (*32*). One possible well-defined target is the NH_2-terminal coiled coil that is exposed in the prehairpin intermediate (*18*). Speculatively, a 5-Helix–like intermediate may be exposed during the fusion process (*33*) and, in this case, antibodies directed against 5-Helix may inhibit viral entry.

Finally, structural (*4*) and computational (*5*) methods predict similar trimer-of-hairpins motifs for viruses in diverse families, including orthomyxoviridae, paramyxoviridae, filoviridae, and retroviridae. In some of these cases, inhibition of viral entry by peptides analogous to the C-peptides of gp41 has been demonstrated (*34–36*). Thus, the 5-Helix design approach may offer a widely applicable strategy for inhibiting viral infections (*37*).

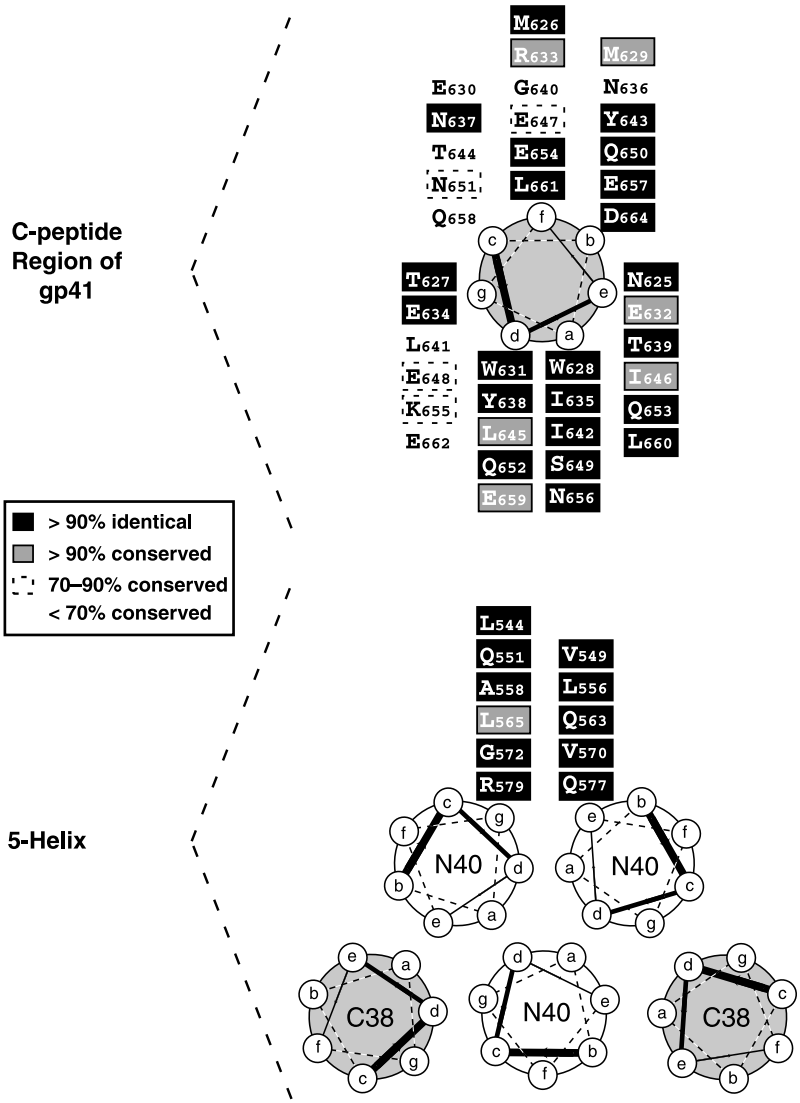


Fig. 4. A helical wheel diagram depicting the interaction of 5-Helix with the C-peptide region of gp41. The (a) through (g) positions in each helix represent sequential positions in the 4,3 hydrophobic heptad repeat in each sequence. The (a) and (d) positions in the gp41 C-peptide region interact with the exposed (e) and (g) positions on the N40 coiled coil of 5-Helix. Residues are boxed according to their degree of conservation as determined from the alignment of 247 sequences from HIV-1, HIV-2, and SIV isolates (HIV-1 sequence database, August 2000, Los Alamos National Laboratory): black rectangle, >90% identical; gray rectangle, >90% conservative substitution (*43*)]; dotted rectangle, 70 to 90% conserved; no box, <70% conserved. Note the high degree of conservation in the (a) and (d) positions of the C-peptide region of gp41, a property markedly lacking in other positions [particularly (c) and (g)] of the C-peptide region not directly involved in binding 5-Helix.

References and Notes

1. D. C. Chan, P. S. Kim, *Cell* **93**, 681 (1998).
2. M. P. D'Souza, J. S. Cairns, S. F. Plaeger, *JAMA* **284**, 215 (2000).
3. F. Hughson, *Curr. Biol.* **7**, R565 (1997).
4. X. Zhao, M. Singh, V. N. Malashkevich, P. S. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 14172 (2000).
5. M. Singh, B. Berger, P. S. Kim, *J. Mol. Biol.* **290**, 1031 (1999).
6. M. Lu, S. C. Blacklow, P. S. Kim, *Nature Struct. Biol.* **2**, 1075 (1995).
7. D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell* **89**, 263 (1997).
8. W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel, D. C. Wiley, *Nature* **387**, 426 (1997).
9. K. Tan, J.-H. Liu, J.-H. Wang, S. Shen, M. Lu, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12303 (1997).
10. S. Jiang, K. Lin, N. Strick, A. R. Neurath, *Nature* **365**, 113 (1993).
11. C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9770 (1994).
12. D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15613 (1998).
13. R. A. Furuta, C. T. Wild, Y. Weng, C. D. Weiss, *Nature Struct. Biol.* **5**, 276 (1998).
14. L. T. Rimsky, D. C. Shugars, T. J. Matthews, *J. Virol.* **72**, 986 (1998).
15. I. Munoz-Barroso, S. Durell, K. Sakaguichi, E. Apella, R. Blumenthal, *J. Cell Biol.* **140**, 315 (1998).
16. C. M. Carr, P. S. Kim, *Cell* **73**, 823 (1993).
17. J. M. Kilby *et al.*, *Nature Med.* **4**, 1302 (1998).
18. D. M. Eckert, V. N. Malashkevich, L. H. Hong, P. A. Carr, P. S. Kim, *Cell* **99**, 103 (1999).
19. C. Wild, T. Oas, C. McDanal, D. Bolognesi, T. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10537 (1992).
20. Y. Weng, C. D. Weiss, *J. Virol.* **72**, 9676 (1998).
21. The design of 5-Helix was based on the N36/C34 six-helix bundle crystal structure (*7*). For the 5-Helix protein, each peptide region was extended (compared with N36 and C34) by three residues on its NH_2-terminus and one residue on its COOH-terminus, generating the final N40 and C38 segments (representing residues 543 to 582 and 625 to 662, respectively, of HIV-1 HXB2 gp160). Three N40 and two C38 segments were joined using a -GGSGG- linker (*38*) after N40 and a -GSSGG- linker after C38. All constructs include an NH_2-terminal Met for translation initiation. Two distinct 5-Helix proteins that differ only at their COOH-termini were generated for this study: (i) His-tagged 5-Helix, which ends in -GG$(H)_6$, and (ii) untagged 5-Helix, which ends in -GGR. In addition, a third construct, denoted 6-Helix, was generated in which the 5-Helix backbone was connected to the His-tagged C-peptide, C37-H6 [sequence given in (*39*)], through a trypsin-cleavable linker (-GGR-). All DNA constructs were assembled from polymerase chain reaction cassettes sequentially cloned into the pAED4 vector [D. S. Doering, P. Matsudaira, *Biochemistry* **35**, 12677 (1996)] using *Escherichia coli* XL1-Blue (recA$^-$ strain, Stratagene). All proteins were recombinantly expressed in *E. coli* strain RP3098 grown in 2× YT (yeast-tryptone medium) to an optical density (at 590 nm) between 0.5 and 0.7 before induction with isopropyl-β-D-thiogalactopyranoside (0.4 mM) for 3 hours. Bacterial pellets were resuspended in tris/NaCl buffers (Qiaexpressionist booklet, March 1999, Qiagen) supplemented with Complete EDTA-free protease inhibitor tablets (Roche), and subsequently frozen at –20°C until the day of purification. Thawed resuspensions were lysed (sonication or French press) and centrifuged (35,000*g* for 30 min) to separate the soluble fraction from inclusion bodies. His-tagged 5-Helix (generated from plasmid p-5HelixH6) was purified directly from the inclusion bodies resuspended in 8 M urea in tris-buffered saline (TBS) [50 mM tris (pH 8.0), 100 mM

NaCl] and 10 mM imidazole. The mixture was clarified by centrifugation (35,000*g* for 30 min) before binding to a Ni-nitrilotriacetate (Ni-NTA) agarose (Qiagen) column at room temperature. Protein was eluted in 6 M urea/TBS/100 mM imidazole in 40 ml (~5 column volumes). The protein was refolded by slow dripping into a 1-liter, stirred solution of 20 mM tris (pH 8.0) at room temperature. Refolded protein was then reconcentrated by passage over a Ni-NTA agarose column and eluted with 20 ml (~2 column volumes) of 100 mM imidazole in TBS. Untagged 5-Helix was produced via proteolysis of 6-Helix (see below) to generate a 5-Helix/C37-H6 complex. Following digestion with trypsin (1:200 weight ratio in TBS at room temperature for 1 hour, Sigma), the 5-Helix/C37-H6 complex was bound to Ni-NTA agarose and washed extensively to remove excess trypsin. The beads were resuspended in 8 M GuHCl/TBS and heated (70°C) in order to denature the complex. The nonbinding fraction, containing denatured 5-Helix, was sequentially dialyzed into 8 M urea/20 mM tris (pH 8.0) (4 hours at room temperature) and 4 M urea/20 mM tris (pH 8.0) (overnight at 4°C). The protein was loaded onto a DEAE column (Fastflow, Pharmacia) and a reverse urea gradient [4 M to 0 M urea in 20 mM tris (pH 8.0)] was run over 20 column volumes in 4 hours at room temperature. The protein was eluted from the DEAE resin using a NaCl gradient (0 to 300 mM) in 20 mM tris (pH 8.0) (10 column volumes). The 6-Helix protein (generated from plasmid p-6Helix) was purified directly from the soluble fraction of the bacterial lysate. The solution was passed over Ni-NTA agarose column and eluted with an imidazole gradient (10 to 250 mM) in TBS over 10 column volumes. For all proteins, monomers were separated from aggregates by gel filtration (Sephacryl S200 HR or Superdex 75) in TBS. The proteins were >95% pure as judged by SDS–polyacrylamide gel electrophoresis and can be concentrated to at least 3 mg/ml. The concentrations of all peptides and proteins were determined by absorbance at 280 nm in 6 M GuHCl [H. Edelhoch, *Biochemistry* **6**, 1948 (1967)].

22. The 6-Helix construct has an extremely stable helical fold ($[\theta]_{222}$ of −26,000 deg cm^2 dmol^{-1} at 25°C; >95% of the amount predicted by the design) that does not display thermal denaturation (up to 96°C) in phosphate-buffered saline (PBS) [50 mM phosphate (pH 7.5), 150 mM NaCl] or chemical denaturation (up to 8 M GuHCl in PBS) at 25°C.
23. In 5-Helix(D4), four highly conserved residues in the C-peptide binding site of His-tagged 5-Helix (Val^{549}, Leu^{556}, Gln^{563}, and Val^{570}) were mutated to Asp in the final (third) N40 segment. The construct [p-5Helix(D4)] was recombinantly expressed and purified in the same manner as the His-tagged 5-Helix. The His-tagged 5-Helix and 5-Helix(D4) proteins have the same ellipticity: for both, $[\theta]_{222}$ = −28,100 ± 1500 deg cm^2 dmol^{-1} (~100% of the predicted helical content) at 4°C in TBS, and both proteins are extremely stable to thermal denaturation (T_m > 98°C) in TBS, as well as to GuHCl chemical denaturation [C_m values of ~6 M for 5-Helix(D4); ~7.2 M for the His-tagged 5-Helix] at 25°C. The slightly decreased stability of 5-Helix(D4) likely reflects the low helical propensity and charge of the Asp residues which, in this context, are placed within a predominantly hydrophobic groove on the surface of 5-Helix.
24. Although they typically require parenteral administration, protein-based therapeutics can be practical, as exemplified by insulin, growth hormone, tissue plasminogen activator, granulocyte–colony stimulating factor, and erythropoietin. Alternatively, 5-Helix could be expressed endogenously (e.g., via gene therapy) with secretion into the bloodstream. If 5-Helix were expressed endogenously in HIV-1–infected cells, it could inhibit intracellular folding and transport of gp160. The 5-Helix, 5-Helix(D4), and 6-Helix proteins are also potential reagents for small-molecule drug-screening purposes [compare (*18*)].
25. Three potential obstacles to using 5-Helix as a therapeutic include proteolysis, renal filtration, and immune clearance. The stably folded core of 5-Helix is likely to be resistant to serum proteolysis, but it may be desirable to replace the flexible Gly/Ser peptide linker regions with sequences that are more protease-resistant. It should also be possible to generate 5-Helix variants with an increased molecular weight (by oligomerization or tethering to a large protein) to reduce the rate of kidney clearance. Immune reactivity is thought to be reduced by glycosylation [S. Alexander, J. H. Elder, *Science* **226**, 1328 (1984)]. The exterior of the C-peptide helices in 5-Helix each contain an endogenous N-linked glycosylation site (*40*), and additional sites could potentially be engineered into a 5-Helix variant.
26. T. Muster *et al.*, *J. Virol.* **67**, 6642 (1993).
27. M. Purtscher *et al.*, *AIDS* **10**, 587 (1996).
28. The core of the 2F5 epitope (Leu-Asp-Lys-Trp; residues 663 to 666 in the HIV HXB2 gp160 sequence) is highly conserved (81% identity) across the same set of HIV-1, HIV-2, and SIV isolates used to generate Fig. 4. However, some HIV-1 escape variants to 2F5 neutralization do not contain mutations in the epitope sequence (*26, 27*), suggesting that inhibition by 2F5 may involve recognition of additional determinants.
29. E. F. Pai, M. H. Klein, P. Chong, A. Pedyczak, World Intellectual Property Organization (www.wipo.org) patent WO-00/61618 (2000).
30. T. Muster *et al.*, *J. Virol.* **68**, 4031 (1994).
31. L. Eckhart *et al.*, *J. Gen. Virol.* **77**, 2001 (1996).
32. R. A. LaCasse *et al.*, *Science* **283**, 357 (1999).
33. Topological considerations restrict the nature of the transition between the prehairpin intermediate and the trimer-of-hairpins state. In the prehairpin intermediate, there is a threefold axis of symmetry perpendicular to the two membranes. In contrast, the threefold axis of symmetry in the trimer-of-hairpins is thought to be parallel to the membranes (Fig. 1A) (*1, 3, 8, 40*). Thus, any concerted collapse into the trimer-of-hairpins state would proceed through an asymmetric transition state. Alternatively, the reaction might proceed sequentially, one hairpin at a time. Intermediates with one or two hairpins might twist to avoid entanglement of the fusion peptides and transmembrane helices.
34. D. M. Lambert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2186 (1996).
35. D. Rapaport, M. Ovadia, Y. Shai, *EMBO J.* **14**, 5524 (1995).
36. Q. Yao, R. W. Compans, *Virology* **223**, 103 (1996).
37. Cellular fusion events mediated by SNARE (soluble N-ethylmaleimide–sensitive factor attachment protein receptor) proteins involve formation of a 4-helix coiled-coil bundle [R. B. Sutton, D. Fasshauer, R. Jahn, A. T. Brünger, *Nature* **395**, 347 (1998)]. An approach analogous to that used here might also lead to specific inhibitors of such cellular fusion events.
38. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
39. Peptide C37-H6 is a His-tagged C-peptide of the following sequence: GGHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLGHHHHHH. The peptide is derived from HIV-1 HXB2 residues 625 to 661 (underlined) and contains the entire C34 sequence (W628 to L661). Peptide C37-H6 is produced from the tryptic digestion of a recombinantly expressed construct, p4-NC1.1, consisting of one N40 segment joined to C37-H6 through a -GGR- linker. Following expression, NC1.1 is purified from the soluble fraction of bacterial lysates in the same manner as 6-Helix. Trypsin digestion (same conditions as for untagged 5-Helix) generates C37-H6, which is then purified to homogeneity by reverse-phase high-pressure liquid chromatography using a Vydac C-18 column and a linear gradient of acetonitrile in water containing 0.1% trifluoroacetic acid. The identity of C37-H6 was confirmed by mass spectrometry [MALDI-TOF (matrix-assisted laser desorption/ionization–time-of-flight), PerSeptive]. Like C34, C37-H6 is a potent inhibitor of HIV-1 membrane fusion, with an $IC_{50} \approx 1$ nM in the cell-cell fusion assay (M. J. Root, P. S. Kim, unpublished results).
40. S. C. Blacklow, M. Lu, P. S. Kim, *Biochemistry* **34**, 14955 (1995).
41. The data in Fig. 2 were generated using the untagged version of 5-Helix, but similar results were obtained with the His-tagged version (*23*). The CD (Aviv 62 DS) experiments were performed in TBS buffer unless otherwise stated. In Fig. 2B, the protein concentration was 1 μM for the TBS sample and 0.54 μM for the GuHCl/TBS sample. In Fig. 2D, a quartz mixing cell (Helma) with 1-ml chambers (4.375 mm per chamber pathlength) was used. The polypeptides were at a concentration of 5.9 μM (5-Helix) and 6 μM (C37-H6) in 20 mM tris (pH 8.0)/250 mM NaCl before mixing. C37-H6 does not have a significant helical content in isolation (M. J. Root, P. S. Kim, unpublished results). The 5-Helix precipitation experiment (Fig. 2C) was performed in 20 μl of TBS with 16 μM untagged 5-Helix, 30 μM His-tagged C37-H6, and/or 150 μM C34. The solution was added to 10 μl of Ni-NTA agarose and incubated at room temperature for 10 min. After the unbound supernatant was removed, the beads were washed twice with 1 ml of TBS and then eluted with 500 mM imidazole. The Ni-bound and unbound samples were run on a 16.5% tris-Tricine polyacrylamide gel (Bio-Rad) and stained with Gel-code Blue (Pierce).
42. All data in Fig. 3 were generated using His-tagged 5-Helix (*21*). The cell-cell fusion assays (Fig. 3B) were performed as described (*12*). Inhibition of viral infectivity was studied using a recombinant luciferase reporter assay slightly modified from that previously detailed (*12*). Briefly, pseudotyped viruses were generated from 293T cells cotransfected with an envelope-deficient HIV-1 genome NL43LucR$^-$E$^-$ [B. K. Chen, K. Saksela, R. Andino, D. Baltimore, *J. Virol.* **68**, 654 (1994)] and one of four gp160 expression vectors: pCMV-HXB2 (*12*), pEBB-JRFL (kindly provided by B. K. Chen), pSVIII-UG024.2, and pSVIII-RW020.5. The plasmids pSVIII-UG024.2 and pSVIII-RW020.5 were obtained from the NIH AIDS Reagent Program (F. Gao, B. Hahn, and the Division of Acquired Immunodeficiency Syndrome, National Institute of Allergy and Infectious Diseases) and code for envelope protein from primary HIV-1 isolates. Supernatants containing virus were prepared as described (*12*) and used to infect either human osteosarcoma (HOS)–CD4 cells (HXB2 and UG024.2) or HOS-CD4-CCR5 cells (JRFL and RW020.5). Cells were obtained from the NIH AIDS Reagent Program (N. Landau). In Fig. 3A, viral infectivity assays were performed in the standard 24-well format (*12*). The data in Fig. 3C were obtained from assays conducted in 96-well format (B. K. Chen and P. S. Kim, unpublished results): virus-containing supernatant (10 μl) and media (90 μl) were overlaid onto HOS cells at 50% confluency. Following 2 days of incubation at 37°C, the cells were harvested in 100 μl of lysis buffer (Luciferase Assay System, Promega), of which 10 μl was analyzed per manufacturer's protocol. The IC_{50} values were calculated by fitting the 5-Helix titration data to a Langmuir function {normalized luciferase activity = 1/(1 + [5-Helix]/IC_{50})}.
43. To generate Fig. 4, we considered substitutions within the following groups of amino acid residues to be conservative: [Asp, Glu], [Lys, Arg], [Asn, Gln], [Phe, Tyr], [Ser, Thr], and [Val, Ile, Leu, Met].
44. We thank H. Chhay and B. K. Chen for their assistance with the cell-cell fusion and viral infection assays; I. Rousso and D. Chan for helpful discussions; M. Burgess for peptide synthesis; A. Evindar and L. Gaffney for editorial assistance; and members of the Kim laboratory for stimulating discussions and other contributions. M.J.R. and M.S.K. are postdoctoral fellows of the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation. This research was funded by NIH (P01 GM56552).

14 November 2000; accepted 27 December 2000
Published online 5 January 2001;
10.1126/science.1057453
Include this information when citing this paper.