Inhibiting HIV-1 Entry: Utilizing a Transient Intermediate of Viral Membrane Fusion as a Target for Drug Discovery

by

Debra M. Eckert

B.S. Biology
Texas Christian University

Submitted to the Department of Biology in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN BIOLOGY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2000

Signature of Author: _____
Department of Biology
January 31, 2000

Certified by: _____
Peter S. Kim
Professor, Department of Biology
Thesis Supervisor

Accepted by: _____
Alan D. Grossman
Chairman, Biology Graduate Committee

Dedication

To Jeff,
for moving to Boston

Acknowledgments

First I would like to thank my parents who have always fostered my education and encouraged my desire to learn. I appreciate their support and love. I also cherish the companionship I have shared with my brother, Drew, throughout the years.

Second, I thank my undergraduate advisors, Pat Paulus, Phil Hartman and Roger Hewitt. Each one of them taught me science is fun and greatly impacted my decision to pursue a graduate education.

Third, I would like to thank some of the people I have met here in the Kim Lab. Debbie Fass was an excellent mentor to me. I had never even worked with proteins before joining the lab, so she had her hands full teaching me the basics. Her tireless and encouraging assistance gave me the ability to become an independent scientist. Lawren Wu, my baymate, was a confidant and also provided emotional and technical support. I hope to call him my life-long friend. I would also like to thank my co-workers who have contributed significantly to the work in this thesis: Lily Hong, Peter Carr and Vladimir Malashkevich. Lily has been an incredible worker and a good friend. I greatly appreciate her patience and diligence in performing important, yet often mundane, tasks. Peter and Vladimir exerted enormous energy that made my final graduate project far better than it would have been without them. Mike Burgess, James Pang, Roberta Moro and Ben Sanford provided extensive peptide synthesis support for which I thank them.

Fourth, I thank my advisor, Peter S. Kim. His unfaltering guidance has dramatically effected the course of my graduate research. He always expected more than I thought I could give, and I attribute most of my success to his talent and encouragement. Also, I hope I have learned at least a little from his ability to extract the important story from scientific data and to communicate it effectively.

Finally, I thank Jeff Eckert, for his endless support. He has been my best friend both throughout and before graduate school, and I am thrilled to now call him my husband.

Inhibiting HIV-1 Entry:
Utilizing a Transient Intermediate of Viral Membrane Fusion
as a Target for Drug Discovery

by

Debra M. Eckert

Submitted to the Department of Biology on January 31, 2000
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Biology

Abstract

In order to reproduce, enveloped viruses must deposit their genomes into host cells. The first step of this process is the fusion of viral and cellular membranes, which allows the release of the viral contents into the cell. An envelope glycoprotein on the surface of the virus is responsible for fusion. It is usually composed of two subunits, a surface subunit that attaches the virus to the host cell, and a transmembrane subunit that mediates the fusion process. Recent biochemical and structural studies on the transmembrane subunit of the HIV-1 virus, gp41, have revealed a transient intermediate of the fusion process that is a potential target for anti-viral therapy. After virion attachment to the cell, gp41 undergoes a conformational change and inserts into the target cell membrane. Concomitantly, a conserved region of gp41 that is hidden both before and after this stage is accessible. The exposed core contains a trimeric coiled coil. A hydrophobic pocket on the surface of this coiled coil has been previously identified as a promising drug target. However, synthetic peptides corresponding to this region aggregate, and are therefore not useful for drug screens. This thesis describes the development of a hybrid molecule that accurately presents the gp41 transient pocket and the use of this hybrid in a screen for potential anti-HIV molecules (Chapter 2 and Chapter 3). Extensive biochemical and structural studies on other viral envelope glycoproteins imply that many diverse viruses utilize similar mechanisms of viral entry (reviewed in Chapter 4). Thus, the methods described for targeting the transient intermediate of gp41-mediated fusion should be useful for combating many diverse viruses.

Thesis Supervisor: Dr. Peter S. Kim, Professor of Biology

Table of Contents

CHAPTER 1

INTRODUCTION: INHIBITING HIV-1 INFECTION

In the early 1980's, modern science was confronted by a serious health threat, the human immunodeficiency virus (HIV). In 1983 HIV was determined to cause acquired immune deficiency syndrome (AIDS), a newly identified condition in which patients with compromised immune systems developed rare illnesses such as Pneumocystis carinii pneumonia and Kaposi's sarcoma. Quickly, the HIV virus began to endanger world health, even in the face of rapid drug discovery fueled by powerful scientific tools. Now, less than twenty years after the discovery of the virus, there are approximately 33.6 million people infected with HIV, corresponding to one HIV-positive person for every 200 people in the world. In 1999, a new person was infected with the HIV virus every six seconds. Safe, effective HIV prevention and treatment is drastically needed.

HIV-1 Virion

HIV belongs to the retrovirus family of enveloped viruses. The virus particle contains two copies of a single-stranded RNA genome enclosed in a nucleocapsid that is surrounded by a lipid bilayer (Figure 1). The genome contains five essential genes (gag, pol, env, tat and rev) and four accessory genes not required for viral replication (vif, vpr, vpu and nef). The gag gene encodes the core structural proteins that comprise the nucleocapsid. The pol gene encodes the essential viral enzymes, reverse transcriptase, (RT), protease (PR) and integrase (IN). The env gene encodes the envelope protein that is required for viral entry into the host cell. The other two essential genes (rev and tat) encode for small proteins that do not associate with the virus particle but are required for viral replication.

Life Cycle

Like all viruses, the HIV virus is analogous to an obligate parasite, requiring the host cell to replicate. A cycle of viral replication is initiated when HIV attaches to the host cell via its envelope protein and cell-surface receptors (Figure 2). The lipid bilayers of the virus and the host cell fuse, releasing the viral nucleocapsid into the cell. Following partial uncoating of the particle, the viral RNA genome is reverse transcribed to DNA. Then the DNA, along with attached proteins, is shuttled to the nucleus of the cell to integrate into the host's DNA. The host cell machinery transcribes the viral genome along with its own genome. The resultant viral mRNAs (both full-length and processed) are moved to the cytoplasm. The processed transcripts serve as templates for

the translation of viral proteins. The viral proteins and the full-length RNA assemble into intact virions and bud from the cell.

Drug Discovery

Luckily, at the time of the discovery of the HIV virus, there was already extensive knowledge in the retrovirology field, allowing a rational drug development approach. The most obvious way to combat the virus was to attack its life cycle. Almost immediately, essential HIV enzymes were utilized as targets for anti-viral therapy.

The first line of attack was inhibiting the HIV reverse transcriptase (RT). RT had been identified as an essential component of the retrovirus particle in 1970. It uses the RNA genome as a template for DNA synthesis, hydrolyzes the RNA portion of the resultant RNA/DNA hybrid molecule, and synthesizes the remaining DNA strand. Since reverse transcriptase is not found in humans, it is a promising target unique to the invading virus. Nucleoside analogs were immediately pinpointed as potential reverse transcriptase inhibitors. These analogs are competitive inhibitors for the nucleoside triphosphate binding site and terminate chain elongation. Prior to the discovery of HIV, nucleoside analogs had been studied for cancer and anti-viral therapy. Therefore, several companies already possessed reasonably extensive libraries of nucleoside analogs, and these libraries were then screened for the ability to inhibit HIV replication. The first anti-HIV drug to be discovered was the nucleoside analog 3'-azidothymidine (AZT), and it was approved by the Food and Drug Administration (FDA) for anti-HIV therapy just four short years after the discovery of the virus (Figure 3). Currently there are six nucleoside analog RT inhibitors approved by the FDA for HIV treatment.

Unfortunately, because of their ability to serve as DNA synthesis inhibitors in general, nucleoside analogs are often riddled with toxic side effects. Therefore, efforts have increasingly focused on targeting RT with non-nucleoside analogs. By screening chemical libraries in viral replication assays or in *in vitro* reverse transcriptase assays, structurally distinct molecules were identified which are noncompetitive inhibitors of RT. Although over thirty different structural classes of non-nucleoside reverse transcriptase inhibitors (NNRTIs) have been identified, they each appear to bind the same region on RT, a pocket distinct from, but near to, the nucleoside-binding active site. They also each conform to a similar geometry when bound to the pocket (Figure 4). The availability of many high-resolution structures of NNRTIs bound to RT opens the door to the design of

more potent inhibitors. There are currently three NNRTIs approved by the FDA for HIV therapy.

The second line of attack in anti-HIV drug development was HIV-1 protease (PR). PR is an aspartic protease which autocatalyzes the cleavage of the Gag/Pol polyprotein and then clips the Gag protein, creating the mature viral enzymes and structural proteins. In 1988 PR was identified as a potential therapeutic target: PR with mutations at the active-site aspartate residues caused the production of non-infectious immature viral particles. The discovery that HIV-1 PR, in contrast to other proteases, cleaves the amide bond of proline when preceded and followed by a hydrophobic residue such as tyrosine, led to the design of peptide-derived transition state mimetics that would preferentially bind to HIV-1 protease over other cellular proteases. At this point, the story of the design of effective inhibitors took a more modern turn. As the design of transition-state mimetic inhibitors progressed, many high-resolution x-ray crystal structures of HIV-1 PR both complexed with inhibitors and unbound were solved. The atomic resolution structures facilitated the use of structure-assisted design, such as molecular modeling. These techniques improved the potency of peptido-mimetic compounds, and also allowed structural design of non-peptide-derived compounds. Ultimately the process of PR inhibitor discovery involved multiple steps: initial design, structural determination, design modification and appropriate clinical assays. Often there were numerous intermediate potential inhibitors before the final molecule was identified. For example, at least 150 compounds preceded the final design of Indinavir, an FDA-approved PR inhibitor. Ultimately, this multistep, rational process led to the discovery of many high potency inhibitors, and there are now six FDA-approved PR inhibitors.

Despite the rapid discovery of drugs that target the essential HIV enzymes, these drugs are unable to control or eradicate HIV infection. Viral turnover is very fast, and the reverse transcriptase lacks proofreading ability. These two circumstances allow the virus to rapidly mutate and easily avoid the drugs. Indeed, the HIV virus has a 65-fold faster mutation rate than the influenza virus, and emergent viruses often dominate several weeks after exposure to a single anti-RT or anti-PR molecule. To combat the viral evasion, the available drugs have been administered in combination. This type of treatment, called "combination drug therapy", has been much more effective than treatment with a single agent. Unfortunately, some patients are intolerant of such therapies, and some of those that are tolerant are starting to develop long-term adverse effects. Therefore, there is still a strong need for new avenues of combating HIV.

Additional drugs, especially those that are less susceptible to viral emergence and are less toxic, are highly desirable. In order to identify more compounds, scientists will have to be increasingly inventive in their drug targeting strategies.

gp41

The therapies described above target the virus after it has gained admittance to the cell. It would be useful if an inhibitor could be designed that stops the virus from getting into the host cell in the first place. Recently, evidence has been mounting that this will be possible.

The envelope protein on the surface of the virus mediates viral entry into the cell. It is composed of two subunits, a surface subunit called gp120, and a transmembrane subunit called gp41. The two subunits are produced from a single precursor that is proteolytically cleaved by a cellular convertase. After cleavage, the two subunits remain noncovalently associated. gp120 facilitates attachment to the host cell by binding to specific receptors, and gp41 mediates the fusion of the viral and cellular membranes. Recently, structural and biochemical studies on gp41 have identified it as a viable therapeutic target. If gp41 could be successfully targeted, the fusion of the viral and cellular membranes, and therefore the viral life cycle, could be halted.

HIV-1 gp41 is anchored in the viral membrane via a hydrophobic transmembrane domain. The ectodomain contains an N-terminal hydrophobic region termed the fusion peptide, which is required for membrane fusion, and two predicted helical regions. As in the thoroughly studied influenza virus fusion protein, hemagglutinin (HA), HIV gp41 appears to proceed through conformational transitions leading to membrane fusion. These conformational changes are triggered by the binding of gp120 to the host cell receptor (CD4) and co-receptor (one of a family of seven transmembrane domain chemokine receptors).

After gp120 binds to the host cell, but before membrane fusion occurs, there appears to be an intermediate structure of gp41, in which the transmembrane domain is anchored in the viral membrane and the fusion peptide domain is anchored in the host cell membrane. Particular regions of gp41 that are hidden both before and after this stage are accessible. gp41-derived peptides with inhibitory activity seem to target this intermediate, verifying its potential as a possible drug target.

Vaccine Development

This chapter has described efforts of combating HIV with drug therapy. However, the majority of HIV-infected individuals are located in developing countries, where the expense of drug therapy makes it an impossible option for the control of the virus. The only way to truly sequester the growing numbers of HIV infections would be with an inexpensive preventive therapy, such as a vaccine. The traditional method of vaccination, using killed or attenuated viruses, has been unsuccessful with HIV. These vaccines either are incapable of offering protection against a viral challenge or, in extreme cases, eventually lead to disease. The HIV virus, a master of disguise, can avoid the host's immune system like it avoids anti-viral drugs.

Just as for drug development, inventive strategies are required to design an effective HIV vaccine. Many of these strategies have focused on gp120, since this is the first viral protein the host immune system encounters. gp120, similar to RT and PR, is a difficult target. It is highly glycosylated, and the exposed epitopes are extremely variable between viral strains. It is feasible that the gp41 intermediate could be a useful component of a unique vaccine strategy, if it could be properly displayed to the host's immune system.

Purpcse

The HIV-1 virus has eluded eradication by drug therapy approaches and vaccine development. Recent evidence suggests that a conformation of gp41 that predominates only transiently during the viral infection process is a viable target for drug development and possible vaccine efforts. This thesis describes the design of a hybrid peptide molecule that accurately represents a portion of the transient gp41 structure and efforts to target this molecule for anti-HIV therapy (Chapter 3). Chapter 2, while seemingly divergent, discusses an essential component of the designed hybrid molecule. The concluding chapter of the thesis reviews the current knowledge of enveloped virus membrane fusion and inhibition, and therefore places the significance of the work from the previous chapters in proper context.

Selected References

Carpenter, C.C.J., Fischl, M.A., Hammer, S.M., Hirsch, M.S., Jacobsen, D.M., Katzenstein, D.A., Montaner, S.G., Richman, D.D., Saag, M.S., Schooley, R.T., Thompson, M.A., Vella, S., Yeni, P.G., Volberding, P.A. Antiretroviral therapy for HIV infection in 1998: updated recommendations of the International AIDS Society – USA panel. *J. Am. Med. Assoc.* **280**, 78-86 (1998).

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

Coffin, J.M., Hughes, S.H., Varmus, H.E., ed. *Retroviruses.* Plainview, N.Y., CSHL Press (1997).

De Clerq, E. Perspectives of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Farmaco* **54**, 26-45 (1999).

Esnouf, R.M., Ren, J. Hopkins, A.L., Ross, C.K., Jones, E.Y., Stammers, D.K., Stuart, D.I. Unique features in the structure of the complex between HIV-1 reverse transcriptase and the bis(heteroaryl)piperazine (BHAP) U-90152 explain resistance mutations for this nonnucleoside inhibitor. *Proc. Natl. Acad. Sci. USA* **94**, 3984-3989 (1997).

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M. Rapid turnover of plasma virions and CD4+ lymphocytes in HIV-1 infection. *Nature.* **373**, 123-126 (1995).

Johnston, M.I. and Hoth, D.F. Present status and future prospects for HIV therapies. *Science* **260**, 1286-1293 (1993).

Kaldor, S., Kalish, V.J., Davies II, J.F., Shetty, B.V., Fritz, J.E., Appelt, K., Burgess, J.A., Campanale, K.M., Chirgadze, N.Y., Clawson, D.K., Dressman, B.A., Hatch, S.D., Khalil, D.A., Kosa, M.B., Lubbehusen, P.P., Muesing, M.A., Patrick, A.K., Reich, S.H., Su, K.S., Tatlock, J.H. Viracept (nelfinavir mesylate, Ag1343): a potent orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* **40**, 3979-3985 (1997).

Katz, R.A. and Skalka, A.M. The retroviral enzymes. *Annu. Rev. Biochem.* **63**, 133-173 (1994).

Levy, J. *HIV and the Pathogenesis of AIDS.* Washington, D.C., ASM Press (1998).

Mitsuya, H., Weinhold, K.J., Furman, P.A., St Clair, M.H., Lehrman, S.N., Gallo, R.C., Bolognesi, D., Barry, D.W., Broder, S. 3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. *Proc. Natl, Acad. Sci. USA* **82**, 7096-7100, (1985).

Stryer, L. ed. *Biochemistry.* New York, N.Y., W.H. Freeman and Co. (1995).

Wang, J., Smerdon, S.J., Jaeger, J., Kohlstaedt, L.A., Friedman, J., Rice, P.A., Steitz, T.A. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse trancriptase heterodimer. *Proc. Natl. Acad. Sci. USA* **91**, 7242-7246 (1994).

Wlodawer, A. and Vondrasek, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 249-284 (1998).

Figure 1. HIV-1 virion.

Schematic representation of the HIV-1 virion. The locations of the labeled molecules are approximate. The labeled proteins are the products of the gag, pol, and env genes. Products of the gag gene are capsid, matrix, and nucleocapsid. The products of the pol gene are reverse transcriptase, protease and integrase. The env products are gp41 and gp120.

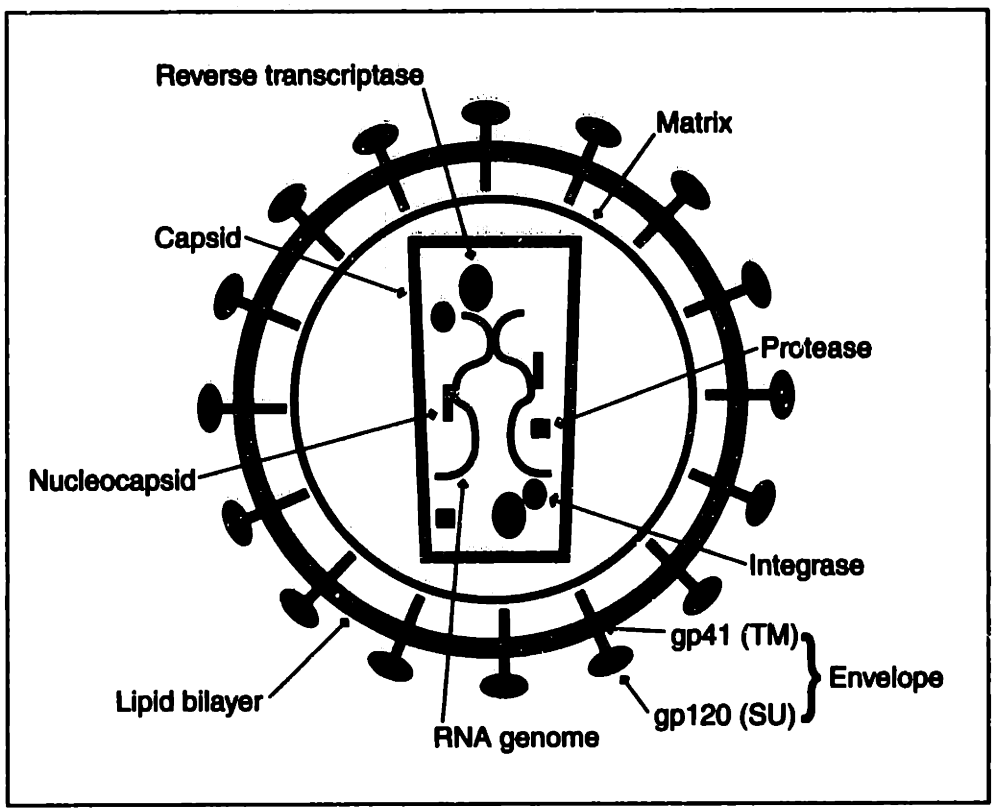

Reverse transcriptase
Matrix
Capsid
Protease
Nucleocapsid
Integrase
gp41 (TM)
Envelope
Lipid bilayer
gp120 (SU)
RNA genome

Figure 2. The HIV-1 viral life cycle.

Representation of one round of HIV-1 infection. To begin a round of infection, the HIV env protein interacts with host cell receptors, CD4 and a seven transmembrane domain chemokine receptor. Subsequently the viral and cellular membranes fuse, releasing the capsid containing the HIV genome into the cell. Then, the capsid particle dissociates and reverse transcriptase uses the single-stranded RNA genome as a template to produce double stranded DNA. The DNA is then shuttled to the nucleus and integrated into the host's genome. The HIV genome is transcribed along with the host's DNA. The full-length and spliced transcripts are transported back to the cytoplasm. Viral proteins are synthesized by the host's translation machinery. Viral particles assemble and bud from the host cell membrane. Following budding, the viral protease activity leads to a mature virus particle.

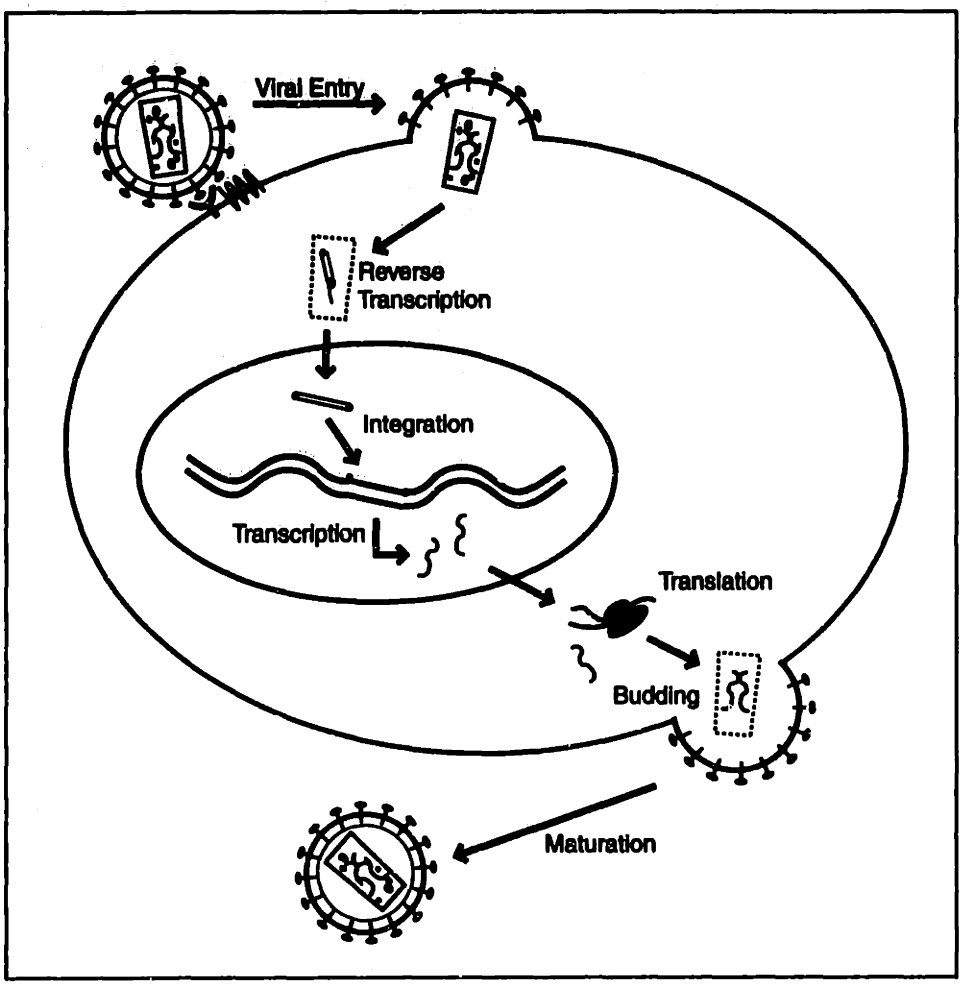

Viral Entry
Reverse
Transcription
Integration
Transcription
Translation
Budding
Maturation

Figure 3. AZT, the first HIV-1 reverse transcriptase inhibitor.

A representation of the structure of AZT, an FDA-approved anti-HIV drug. AZT is a nucleoside analog, and competes with thymidine (shown on the right) as a substrate for HIV-1 reverse transcriptase. AZT lacks a 3' hydroxyl, and therefore serves as a chain terminator during DNA synthesis.

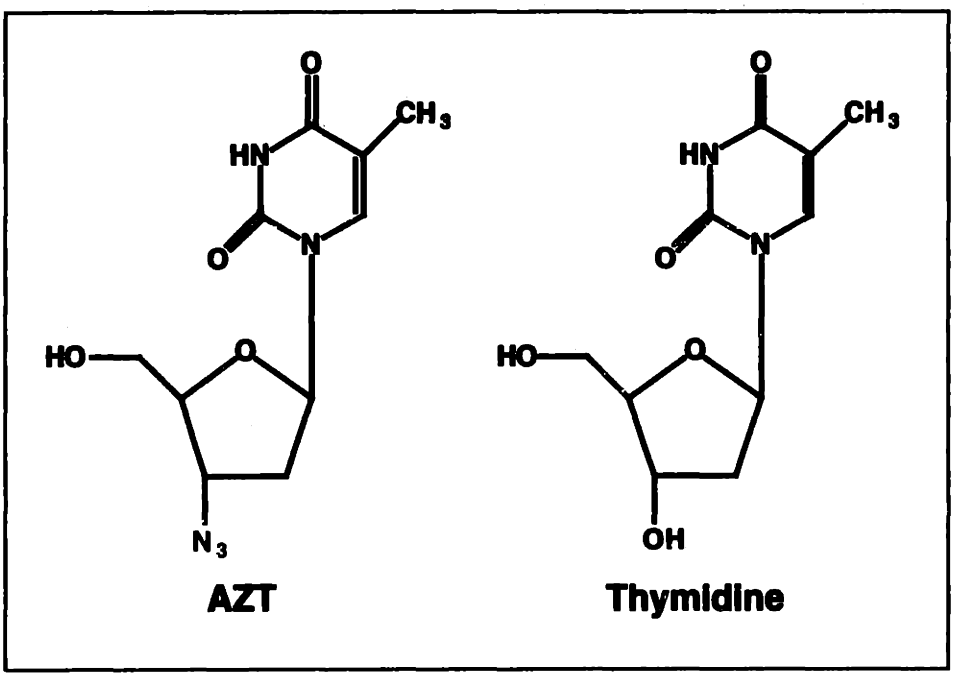

O
CH3
HN
O
N
HO
O
N3
AZT
O
CH3
HN
O
N
HO
O
OH
Thymidine

Figure 4. Non-nucleoside reverse transcriptase inhibitors.

A. Representation of the structures of two-FDA approved non-nucleoside analog reverse transcriptase inhibitors (NNRTIs): nevirapine and delavirdine.

B. The x-ray crystal structures of nevirapine and delavirdine bound to HIV-1 reverse transcriptase (RT). The two structures were overlaid by the alpha carbons of the reverse transcriptase. The secondary structure of RT is depicted in green. Nevirapine is depicted in blue and delavirdine is in yellow. Both inhibitors, although structurally distinct, occupy the same binding site and have the same general "butterfly" conformation.

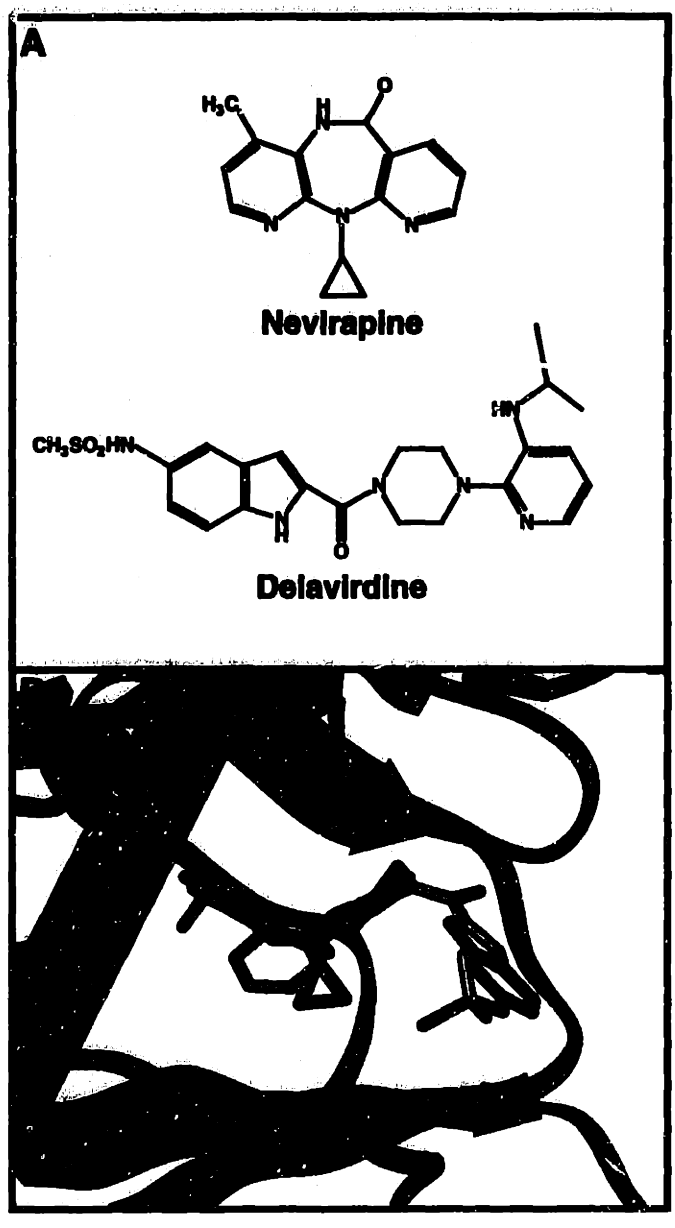
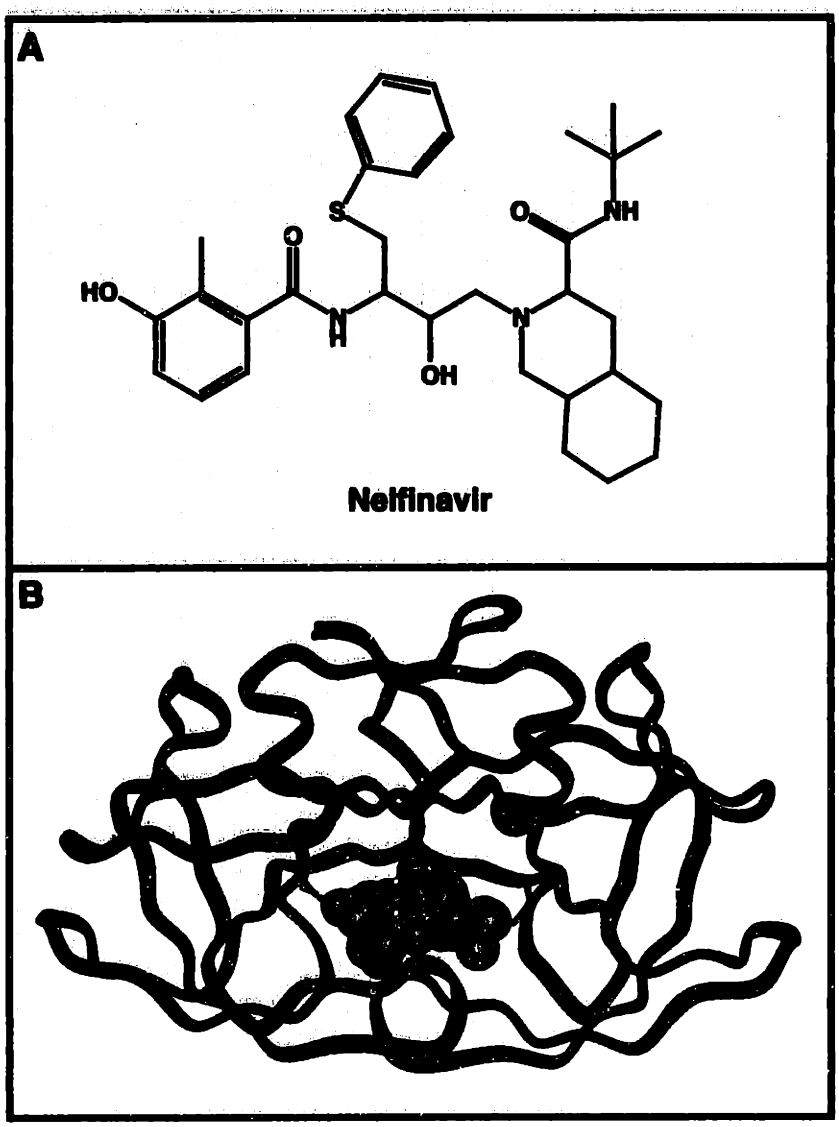

A
Nevirapine
Delavirdine

Figure 5. An FDA-approved protease inhibitor.

A. Representation of the structure of nelfinavir, an FDA-approved, non-peptidic protease inhibitor.

B. The high-resolution x-ray crystal structure of nelfinavir (yellow) bound to the active site of HIV-1 protease. One subunit of the protease dimer is colored purple, the other turquoise. The side chains of the active-site apartic acids are shown.


A
S
O
NH
HO
O
N
H
N
OH
Nelfinavir
B

CHAPTER 2

CRYSTAL STRUCTURE OF GCN4-pI_QI, A TRIMERIC COILED COIL WITH BURIED POLAR RESIDUES

Several people contributed to the work described in Chapter 2. Peptide synthesis was performed by Mike Burgess and James Pang. The circular dichroism and sedimentation studies were performed by me. Both Vladimir Malashkevich and I contributed equally to the x-ray crystallographic analysis.

SUMMARY

Coiled coils consist of two or more α-helices wrapped around each other with a superhelical twist. Hydrophobic amino-acid residues packed in a "knobs-into-holes" arrangement form the interfaces between helices of a coiled coil. Most naturally occurring coiled coils, however, also contain buried polar residues, as do the cores of the majority of naturally occurring globular proteins. Two common buried polar residues in both dimeric and trimeric coiled coils are asparagine and glutamine. In dimeric coiled coils, buried asparagines, but not glutamines, have been shown to confer specificity of oligomerization. We have placed a glutamine in the otherwise hydrophobic interior of a stable trimeric coiled coil, GCN4-pII, to study the effect of this buried polar residue in a trimeric coiled-coil environment. The resulting peptide, GCN4-pI_QI, is a discrete, trimeric coiled coil with a lower stability than GCN4-pII. The crystal structure determined to 1.8 Å shows that GCN4-pI_QI is a trimeric coiled coil with a chloride ion coordinated by one buried glutamine from each monomer.

Keywords : coiled coil; GCN4; buried polar residues; crystallography; ion binding

The coiled coil is a protein structural motif consisting of two to five α-helices wrapped around each other with a superhelical twist (Cohen & Parry, 1990; Malashkevich *et al.*, 1996). The fold of a coiled coil is determined by a simple pattern of amino acids, in which there is a characteristic heptad repeat designated by the letters **a** through **g** (McLachlan & Stewart, 1975). The first and fourth positions of the repeat, the **a** and **d** positions respectively, form the interior of the interacting strands of the coiled coil and are generally hydrophobic (Berger *et al.*, 1995; Hodges *et al.*, 1972; Lupas *et al.*, 1991; McLachlan & Stewart, 1975; Parry, 1982). Nevertheless, most coiled coils also contain polar residues at some of the **a** and **d** positions (Parry, 1982; Woolfson & Alber, 1995). Interestingly, buried polar residues are also a common feature of naturally occurring globular proteins (Barlow & Thornton, 1983; Rashin & Honig, 1984), even though it is energetically unfavorable to sequester a polar residue in a hydrophobic protein core (Harbury *et al.*, 1993; Hendsch *et al.*, 1996; Lumb & Kim, 1995; Sauer & Lim, 1992).

Buried polar residues have been shown to contribute to the specificity of the protein fold at the expense of stability in the GCN4 coiled-coil model system (Harbury et al., 1993). GCN4-p1, a dimeric peptide model of the coiled-coil domain of the yeast transcription activator, GCN4 (O'Shea *et al.*, 1989), contains an asparagine in an **a** position that forms hydrogen bonds with the corresponding asparagine in the other subunit of the homodimer (O'Shea *et al.*, 1991). When this asparagine is replaced with valine, the coiled coil becomes more stable, but oligomerization specificity is lost, as both dimers and trimers are formed (Harbury et al., 1993).

In a designed heterodimeric coiled coil, "Peptide Velcro," a buried asparagine promotes specificity of both helix orientation and oligomerization state. "Peptide Velcro" has a buried asparagine in an **a** position and forms a parallel heterodimer (O'Shea *et al.*, 1993). When the asparagine is mutated to leucine, heterotetramers with both parallel and anti-parallel helix orientation are formed (Lumb & Kim, 1995). Presumably, an anti-parallel orientation is disfavored in "Peptide Velcro" because the polar asparagine sidechains would be buried in an apolar environment, and their hydrogen bonding potential would be unfulfilled. Indeed, it has been found recently that an anti-parallel orientation can be promoted by repositioning the asparagine residues in "Peptide Velcro" (Oakley & Kim, 1998).

Another common polar residue found in the interior of coiled coils is glutamine. Nonetheless, when the asparagine of the dimeric GCN4-p1 is changed to glutamine, oligomeric specificity is lost, and a mixture of dimers and trimers is formed (Gonzalez Jr. *et al.*, 1996). Although glutamine, in the GCN4-p1 backbone, does not specify a discrete oligomeric state, there is a strong preference for glutamine in coiled-coil trimers; glutamine is found approximately three times more frequently than asparagine in the **a** position (Wolf *et al.*, 1997; Woolfson & Alber, 1995). Because of its predominance in naturally occurring coiled-coil trimers, we investigated the effect of burying a glutamine in the hydrophobic core of a well-characterized, trimeric version of the GCN4 coiled coil.

The oligomeric state of a coiled coil is determined by both the hydrophobic and polar residues in the **a** and **d** positions. In GCN4-p1, the oligomeric state can be altered by substituting other hydrophobic residues at the **a** and **d** positions (Harbury et al., 1993). For example, when all the **a** and **d** positions are changed to isoleucine, the peptide folds into an extremely stable, discretely trimeric state (Harbury *et al.*, 1994; Harbury et al., 1993). This peptide, GCN4-pII, provides the backbone for the current studies. We inserted a glutamine into the interior of GCN4-pII at position 16, an **a** position. The resulting peptide, GCN4-pI_QI, has seven isoleucines and one glutamine in the core. A helical wheel projection of GCN4-pI_QI is shown in Figure 1.

As determined by circular dichroism, GCN4-pI_QI has greater than 90 percent helix content in PBS pH 7.0, 4°C. At a concentration of 10 μM, GCN4-pI_QI undergoes a reversible, cooperative thermal unfolding transition with a midpoint of 68°C, as compared to a midpoint exceeding 100°C for GCN4-pII at the same concentration (Figure 2). Sedimentation equilibrium experiments indicate that GCN4-pI_QI exists in a discretely trimeric state across a concentration range from 10 μM to 100 μM (Figure 3). Therefore, the introduction of buried glutamine residues into the core of the trimeric coiled coil does not alter the oligomeric state of the coiled coil but does cause a decrease in stability. In order to determine if there were any structural perturbations not detectable by the above experiments, GCN4-pI_QI was crystallized, and its structure was determined to 1.8 Å resolution.

The crystals of GCN4-pI_QI belong to the space group P321, and contain a monomer in the asymmetric unit, with the trimer formed around the crystallographic three-fold axis. An initial model was determined by molecular replacement methods (Navaza, 1994) using GCN4-pII (Ile16Ala) as the search model. The identical solution

was revealed when a poly-alanine version of GCN4-pII was used as the search model. The glutamine side chain was added after the first round of refinement, as its electron density was apparent in a $2F_O$-F_C map. The structure was refined to a crystallographic R-factor of 21.1% with an R_{free} of 23.6% (Brünger, 1996) over a resolution range of 20.0 to 1.8 Å. The final model consists of 31 amino acid residues and incorporates 47 waters molecules and a chloride ion, which is located on the crystallographic three-fold axis. The model exhibits excellent geometry as determined by PROCHECK (Laskowski *et al.*, 1993). Since molecular replacement methods can cause the final model to be biased by the search model, the final model was verified by simulated-annealing omit maps. The details of crystal growth, data collection, molecular replacement, and refinement are in the legend to Table 1.

GCN4-pI_QI, like GCN4-pII, forms a trimeric, parallel coiled coil with a left-handed superhelical twist (Figure 4). The width of the trimer is ~24 Å and the length is ~47 Å. The isoleucine residues at the core positions show the characteristic "knobs-into-holes" packing (Crick, 1953), in the "acute" mode (Harbury et al., 1993), as seen in the crystal structure of GCN4-pII (Harbury et al., 1994). Interhelical electrostatic interactions are formed between the residues in the **g** position of one heptad ($\mathbf{g}_n$) and the residues in the **e** position of the following heptad in the neighboring strand of the coiled coil ($\mathbf{e}'_{n+1}$), a trait common to coiled-coil structures. All three of the possible interhelical ionic interactions per monomer (between residues Arg1 and Glu6', Lys15 and Glu20', and Glu22 and Lys27') are visible in the final model. The high B-factors, however, of the Lys15 $N\zeta$ atom and Glu20 $O\varepsilon$ atoms (~37 $Å^2$ as compared to an average B factor for the model of 29 $Å^2$) imply that these residues are flexible. Structural superposition of the final model of GCN4-pI_QI and the crystal structure of GCN4-pII results in a root mean square deviation (r.m.s.d.) of 0.82 Å for $C\alpha$ atoms, with the largest deviations occurring at the ends of the helices. The r.m.s.d. using only the $C\alpha$ atoms of the two models without three residues on each end of each helix is 0.35 Å, indicating that the I16Q mutation has a very minor effect on the overall structure of the trimeric coiled coil.

During refinement, a strong spherical peak of electron density appeared on the coiled-coil trimer axis near Gln16 (Figure 4b). This peak is modeled as a chloride ion, present in the crystallization buffer, for two reasons. First, modeling this peak as a water molecule resulted in an anomalously low individual B factor value of 8 $Å^2$ as compared with a value of 31 $Å^2$ for the interacting $O\varepsilon 1$ atom of Gln 16, when the $O\varepsilon 1$ atom of the

glutamine was facing into the center of the coiled coil, and an individual B factor value of 2 $Å^2$ as compared to 19 $Å^2$ for the $N_{\varepsilon 2}$ atom, when the $N_{\varepsilon 2}$ atom is facing in. In contrast, the B-factor of the modeled chloride ion, 24.4 $Å^2$, is quite close to the 24.9 $Å^2$ B-factor of the interacting $N_{\varepsilon 2}$ atom of Gln16. Second, the distance between the initially modeled water molecule and the interacting atom of Gln 16 ($O_{\varepsilon 1}$ or $N_{\varepsilon 2}$) was too long (> 3.2 Å) for optimal hydrogen bonding. The refined distance between the chloride ion and the $N_{\varepsilon 2}$ atom of the glutamine is 3.35 Å, compared to the sum of the van der Waals radii for NH_4^+ and Cl^- of 3.24 Å (Weast, 1980). This distance is also consistent with NH_2 to Cl^- distances previously observed crystallographically in coiled coils (Fass *et al.*, 1996; Malashkevich et al., 1996). In the trimeric coiled coil from the ectodomain of Moloney murine leukemia virus, a chloride ion is coordinated by three asparagines (Fass et al., 1996). The structure of the designed ABC heterotrimer (Nautiyal *et al.*, 1995), which has a core similar to that in GCN4-pI_QI, has recently been solved to 1.8 Å and also shows a chloride ion coordinated by the buried glutamines (S. Nautiyal and T. Alber, in press). Based on these results, it seems likely that anion binding is a common feature of trimeric coiled coils with buried polar residues. Nonetheless, several other crystal structures of trimeric coiled coil domains containing buried glutamine residues, including the HIV gp41 core (Chan *et al.*, 1997; Tan *et al.*, 1997; Weissenhorn *et al.*, 1997b) , the SIV gp41 core (Malashkevich *et al.*, 1998), and a GCN4 variant (Gonzalez Jr. et al., 1996), do not indicate chloride binding.

GCN4-pI_QI should serve as a valuable addition to the repertoire of coiled coils being used to assemble soluble domains of membrane proteins for structural and functional studies. There are many examples where dimeric coiled coils have been used for this purpose (Chang *et al.*, 1994; Cochran & Kim, 1996; Kalandadze *et al.*, 1996; Muir *et al.*, 1994). For example, "Peptide Velcro" was used to replace the transmembrane domain of the T-cell receptor (TCR) in order to express soluble $\alpha\beta$ TCR extracellular domains (Chang et al., 1994). Because many of the membrane-fusion proteins of viruses are trimeric (Allison *et al.*, 1995; Blacklow *et al.*, 1995; Fass & Kim, 1995; Lu *et al.*, 1995; Weissenhorn *et al.*, 1998; Wilson *et al.*, 1981), GCN4-pI_QI will be a useful tool for constructing soluble domains aimed at understanding viral membrane fusion. A potential concern in fusing stable coiled coils to other protein domains is that the native structure of the soluble domain could be distorted by the attached coiled coil. Indeed, the conformation of the cytoplasmic signaling domain of the *Eschericia coli* aspartate receptor, and therefore the extent of activation of the soluble model of the

receptor, can be altered by the length of the linker between that domain and the attached dimeric coiled coil (Cochran & Kim, 1996). GCN4-pII has already been utilized for constructing peptide models of trimeric membrane proteins (Weissenhorn *et al.*, 1997a; Weissenhorn et al., 1998), but the lower stability of GCN4-pI$_Q$I makes it less likely to impose structure on the attached domain.

The coordinates for the GCN4-pI$_Q$I structure have been deposited in the Brookhaven Protein Data Bank (accession code #1piq).

ACKNOWLEDGMENTS

We thank Mike Burgess and James Pang for peptide synthesis and James Berger and Lawren Wu for technical advice. We also thank Deborah Fass for generosity of time and encouragement. Thanks to members of the Kim laboratory for critically reviewing the manuscript. Debra M. Eckert is a National Science Foundation Predoctoral Fellow. This research was supported by the National Institutes of Health (GM44162).

Abbreviations used: r.m.s.d., root mean square deviation; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus; TCR, T-cell receptor; FMOC, fluorenylmethoxycarbonyl; MALDI-TOF, matrix-assisted laser desorption/ionization - time of flight; GuHCl, guanidine hydrochloride; PBS, phosphate buffered saline; MPD, methyl pentanediol.

REFERENCES

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995). Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J. Virol.* **69**, 695-700.

Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867-885.

Berger, B., Wilson, D. B., Tonchev, T., Milla, M. & Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl Acad. Sci. USA* **92**, 8259-8263.

Blacklow, S. C., Lu, M. & Kim, P. S. (1995). A trimeric subdomain of the Simian Immunodeficiency Virus envelope glycoprotein. *Biochemistry* **34**, 14955-14962.

Brünger, A. T. (1996). X-PLOR: A system for X-ray crystallography 3.851 edit. Yale University Press, New Haven.

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., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallog.* **D54**, 905-921.

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

Chang, H.-C., Bao, Z.-Z., Yao, Y., Tse, A. G. D., Goyarts, E. C., Madsen, M., Kawasaki, E., Brauer, P. P., Sacchettini, J. C., Nathenson, S. G. & Reinherz, E. L. (1994). A general method for facilitating heterodimeric pairing between two proteins: application to expression of α and β T-cell receptor extracellular segments. *Proc. Natl Acad. Sci. USA* **91**, 11408-11412.

Cochran, A. G. & Kim, P. S. (1996). Imitation of *Escherichia coli* aspartate receptor signaling in engineered dimers of the cytoplasmic domain. *Science* **271**, 1113-1116.

Cohen, C. & Parry, D. A. D. (1990). α-helical coiled coils and bundles: how to design an α-helical protein. *Proteins Struct. Funct. Genet.* **7**, 1-15.

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

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

Fass, D., Harrison, S. C. & Kim, P. S. (1996). Retrovirus envelope domain at 1.7Å resolution. *Nature Struct. Biol.* **3**, 465-469.

Fass, D. & Kim, P. S. (1995). Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin. *Curr. Biol.* **5**, 1377-1383.

Gonzalez Jr., L., Woolfson, D. N. & Alber, T. (1996). Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nature Struct. Biol.* **3**, 1011-1018.

Harbury, P. B., Kim, P. S. & Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature* **371**, 80-83.

Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401-1407.

Hendsch, Z. S., Jonsson, T., Sauer, R. T. & Tidor, B. (1996). Protein stabilization by removal of unsatisfied polar groups: computational approaches and experimental tests. *Biochemistry* **35**, 7621-7625.

Hodges, R. S., Sodek, J., Smillie, L. B. & Jurasek, L. (1972). Tropomyosin: amino acid sequence and coiled-coil structure. *Cold Spring Harbor Symp. Quant. Biol.* **37**, 299-310.

Jones, T. A., Zou, J. W., Cowan, S. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these methods. *Acta Crystallogr.* **D47**, 110-119.

Kalandadze, A., Galleno, M., Foncerrada, L., Strominger, J. L. & Wucherpfennig, K. W. (1996). Expression of recombinant HLA-DR2 molecules. *J. Biol. Chem.* **271**, 20156-20162.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291.

Lockhart, D. J. & Kim, P. S. (1992). Internal stark effect measurement of the electric field at the amino terminus of an α–helix. *Science* **257**, 947-951.

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

Lumb, K. J. & Kim, P. S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry* **34**, 8642-8648.

Lupas, A., Van Dyke, M. & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* **252**, 1162-1164.

Malashkevich, V., Kammerer, R. A., Efimov, V. P., Schulthess, T. & Engel, J. (1996). The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel? *Science* **274**, 761-765.

Malashkevich, V. N., Chan, D. C., Chutkowski, C. T. & Kim, P. S. (1998). Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: Conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc. Natl Acad. Sci. USA* **95**, 9134-9139.

McLachlan, A. D. & Stewart, M. (1975). Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol.* **98**, 293-304.

Muir, T. W., Williams, M. J., Ginsberg, M. H. & Kent, S. B. H. (1994). Design and chemical synthesis of a neoprotein structural model for the cytoplasmic domain of a multisubunit cell-surface receptor: integrin $\alpha_{IIb}\beta_3$ (platelet GPIIb-IIIa). *Biochemistry* **33**, 7701-7708.

Nautiyal, S. & Alber, T. (1998). Crystal structure of a designed, thermostable, heterotrimeric coiled coil. *Protein Sci.* In the press.

Nautiyal, S., Woolfson, D. N., King, D. S. & Alber, T. (1995). A designed heterotrimeric coiled coil. *Biochemistry* **34**, 11645-11651.

Navaza, J. (1994). AMORE: an automated package for molecular replacement. *Acta. Crystallogr.* **D50**, 760-763.

O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**, 539-544.

O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide 'Velcro*': design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658-667.

O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1989). Evidence that the leucine zipper is a coiled coil. *Science* **243**, 538-542.

Oakley, M. G. & Kim, P. S. (1998). A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry* **37**, 12603-12610.

Otwinoski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp.55-62, SERCDaresbury Laboratory, Warrington, England.

Parry, D. A. D. (1982). Coiled-coils in α-helix containing proteins: analysis of residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Biosci. Rep.* **2**, 1017-1024.

Rashin, A. A. & Honig, B. (1984). On the environment of ionizable groups in globular proteins. *J. Mol. Biol.* **173**, 515-521.

Sauer, R. T. & Lim, W. A. (1992). Mutational analysis of protein stability. *Curr. Opin. Struct. Biol.* **2**, 46-51.

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

Thaller, C., Eichele, G., Weaver, L. H., Wilson, E., Karlsson, R. & Jansonius, J. N. (1985). Diffraction methods for biological macromolecules. Seed enlargement and repeated seeding. *Methods Enzymol.* **114**, 132-135.

Weast, R. C. (1980). *CRC handbook of chemistry and physics*, CRC Press, Inc., Boca Raton, Florida.

Weissenhorn, W., Calder, L. J., Dessen, A., Laue, T., Skehel, J. J. & Wiley, D. C. (1997a). Assembly of a rod-shaped chimera of a trimeric GNC4 zipper and the HIV-1 gp41 ectodomain expressed in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **94**, 6065-6069.

Weissenhorn, W., Calder, L. J., Wharton, S. A., Skehel, J. J. & Wiley, D. C. (1998). The central structural feature of the membrane fusion protein subunit from the ebola virus glycoprotein is a long triple-stranded coiled coil. *PNAS, USA* **95**, 6032-6036.

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

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366-373.

Wolf, E., Kim, P. S. & Berger, B. (1997). MultiCoil: a program for predicting two and three-stranded coiled coils. *Protein Sci.* **6**, 1179-1189.

Woolfson, D. N. & Alber, T. (1995). Predicting oligomerization states of coiled coils. *Protein Sci.* **4**, 1596-1607.

FIGURE 1: Helical wheel representation of the GCN4-pI$_Q$I trimer. Residues in the first heptad, corresponding to the first two helical turns, are circled. The positions in the heptad repeat are labeled in lower case letters. The boxed residue indicates the Ile16Gln mutation from GCN4-pII to GCN4-pI$_Q$I. Dashed lines from one subunit to another represent the potential salt bridges that characteristically form between the **g** residue of one heptad ($\mathbf{g}_n$) and the **e** residue of the next heptad in the neighboring strand ($\mathbf{e}'_{n+1}$). The sequence of GCN4-pI$_Q$I is Ac-RMKQIEDKIEEILSKQYHIENEIARIKKLIGER, with **a** positions singly underlined and **d** positions doubly underlined. GCN4-pI$_Q$I was synthesized by solid phase FMOC peptide synthesis as described previously (Lockhart & Kim, 1992). After cleavage from the resin, the peptide was desalted over a Sephadex G-25 column (Pharmacia) in 5% acetic acid. Following lyophilization, the peptide was resuspended in 5% acetic acid and purified by reverse-phase high performance liquid chromatography (Waters, Inc.) using a Vydac C18 preparative column and a water/acetonitrile gradient with 0.1% trifluoroacetic acid. The peptide elutes at approximately 36% acetonitrile. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry on a Voyager Elite mass spectrometer (PerSeptive Biosystems).

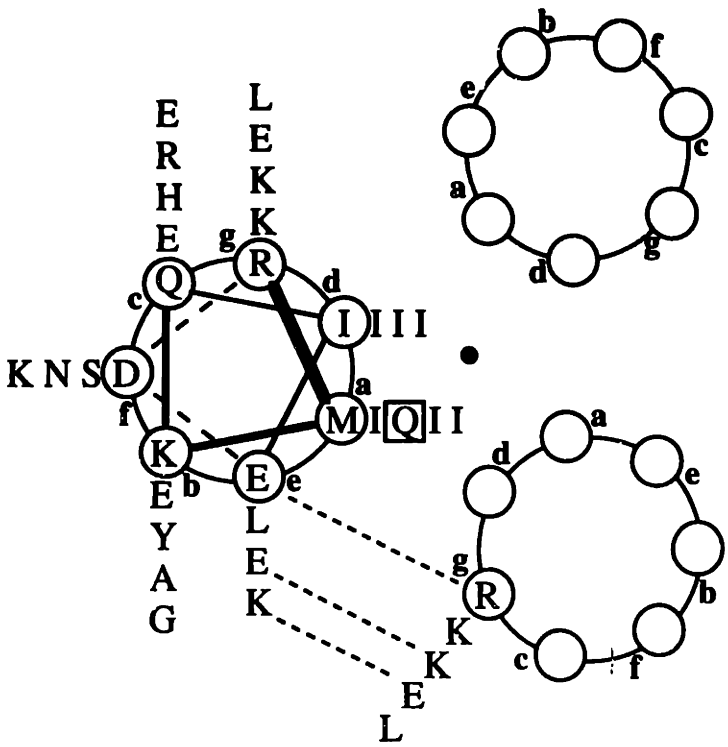

E
R
H
E
L
E
K
K
g
R
d
Q
c
I
I I I
K N S D
a
f
M
I Q I I
K
b
E
e
E
Y
A
G
L
E
K
g
R
K
K
E
L
b
c
f
a
d
e
b
f
e
c
a
g
d

FIGURE 2: Thermal denaturation experiments of GCN4-pI$_Q$I at 10 μM, monitored by circular dichroism, show that it undergoes a reversible cooperative unfolding transition upon heating, with a midpoint of 68°C. The midpoint for GCN4-pII at the same concentration exceeds 100°C. Experiments were performed on an Aviv 62A DS circular dichroism spectrometer. Measurements at 222 nm were performed on 10 μM solutions of peptides in PBS, pH 7.0, in a 10 mm path length cuvette. Samples were heated from 4°C to 84°C, with equilibration times of 1.5 minutes and averaging times of sixty seconds. To obtain a 10 μM sample of the peptide, lyophilized peptide was resuspended first to approximately 500 μM in water. The concentration of the peptide stock was determined by tyrosine absorbance at 280 nm in 6 M GuHCl (Edelhoch, 1967), and the stock was diluted to 10 μM in PBS.

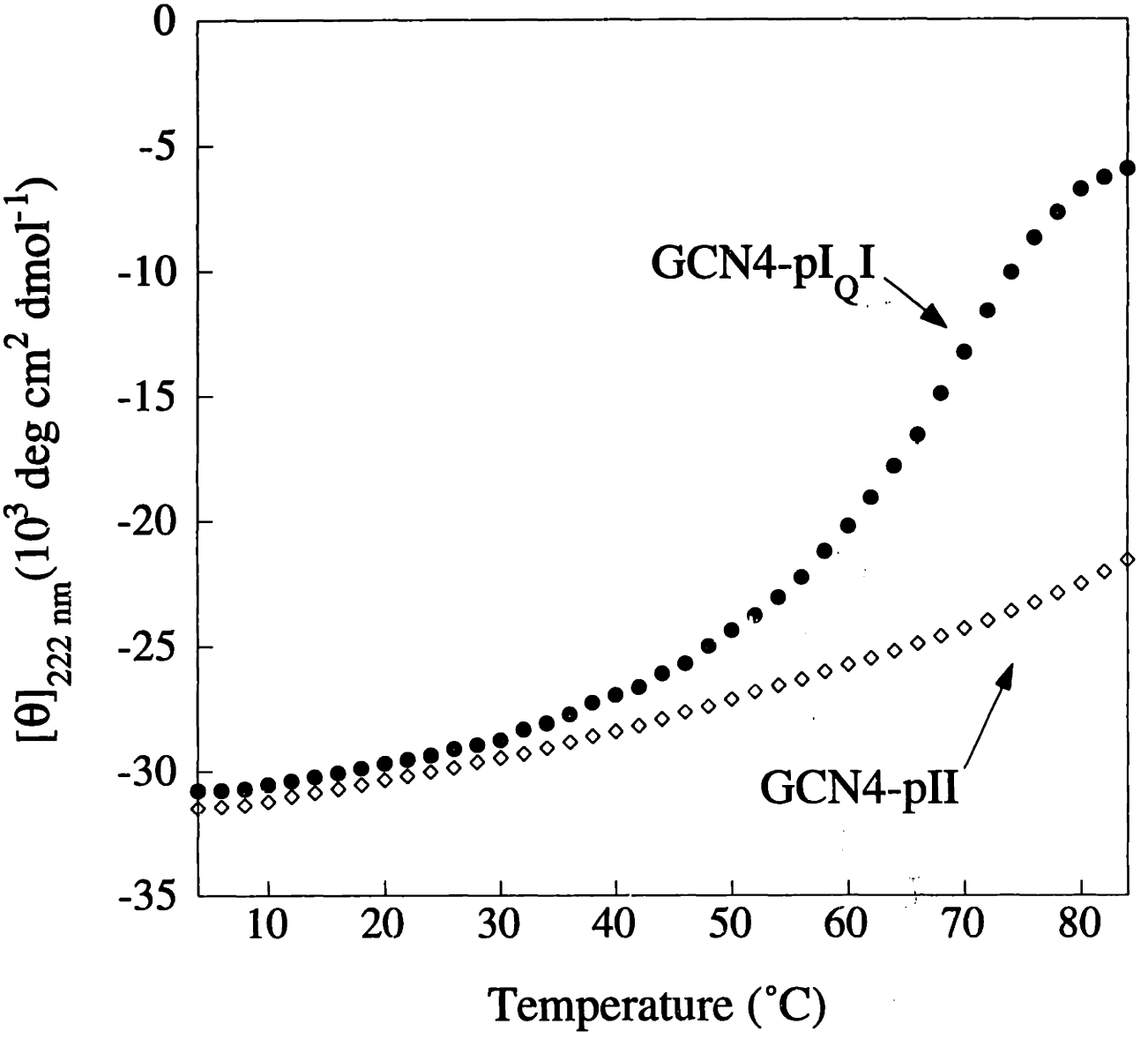

0
-5
-10
-15
-20
-25
-30
-35
10
20
30
40
50
60
70
80
$[\theta]_{222\ nm}(10^3\ deg\ cm^2\ dmol^{-1})$
Temperature (°C)
GCN4-pI$_Q$I
GCN4-pII

FIGURE 3: Sedimentation equilibrium studies of GCN4-pI$_Q$I indicate that it is trimeric across a concentration range from 10 μM to 100 μM. Representative data for 10 μM GCN4-pI$_Q$I are plotted as ln(absorbance) at 231 nm against the square of the radius from the axis of rotation divided by 2. The slope is proportional to molecular weight. Solid lines indicate the calculated lines for the indicated oligomeric states. The deviation in the data from the line calculated for the trimeric species is plotted in the upper panel. Sedimentation equilibrium studies were performed at 4°C on a Beckman XL-A analytical ultracentrifuge using an An-60 Ti rotor at speeds of 28 and 32 krpm and protein concentrations of 10 μM, 20 μM, 50 μM, and 100 μM. The protein solutions were dialyzed overnight against PBS pH 7.0 before spinning in Beckman 6-sector cells.

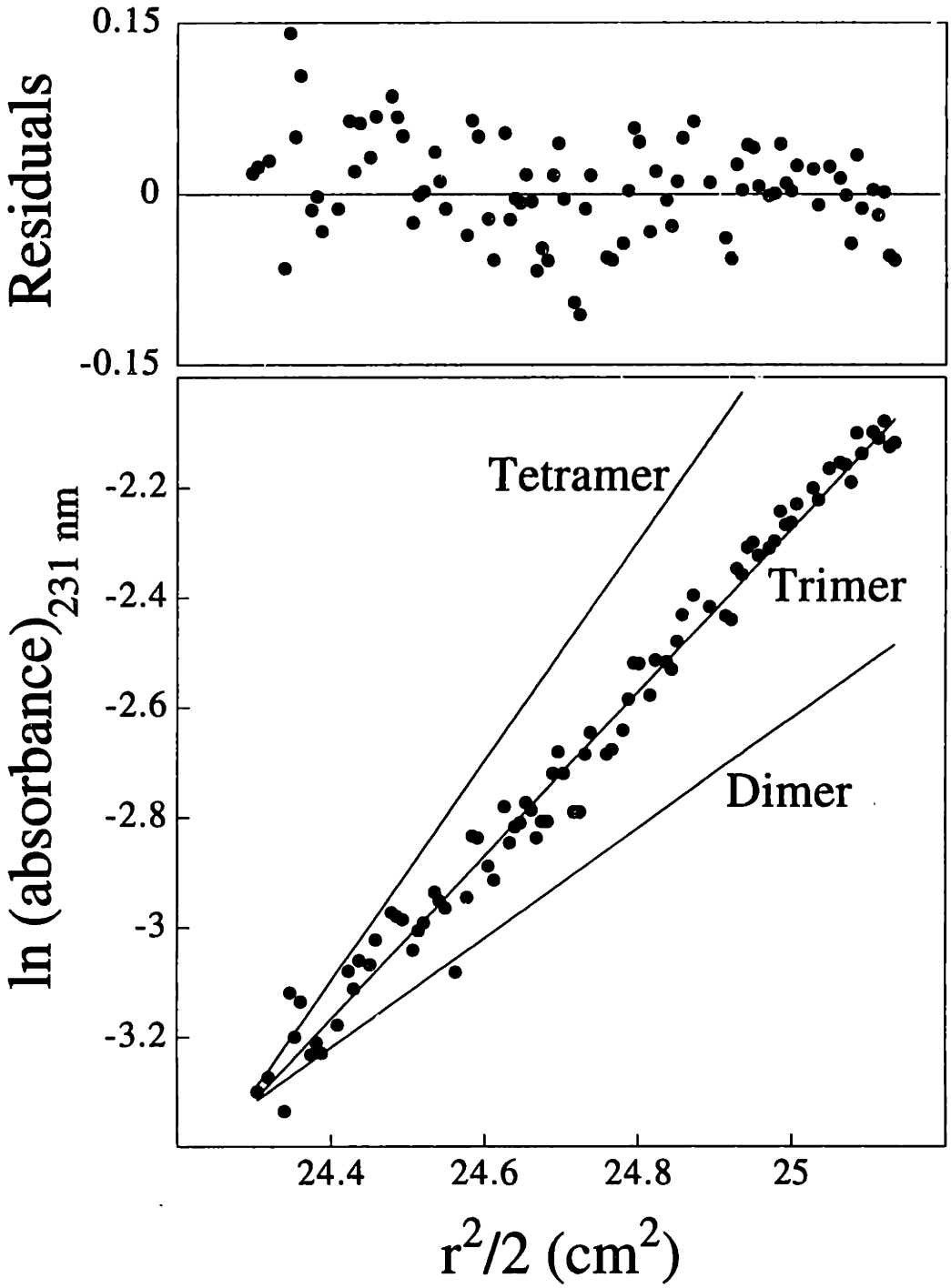

Residuals
0.15
0
-0.15
ln (absorbance)231 nm
-2.2
-2.4
-2.6
-2.8
-3
-3.2
Tetramer
Trimer
Dimer
24.4
24.6
24.8
25
r2/2 (cm2)

FIGURE 4: Crystal structure of GCN4-pI$_Q$I. a. A portion of the 20.0 to 1.8 Å $2F_o$-F_c electron density map contoured at 1.5σ is superposed on the final model. The side view of the helix covers residues 16 to 23. b. Another region of the same map, superposed on the final model, showing the cross section of the trimer at residue 16. A chloride ion (green sphere) is located in the center of the trimer, on the crystallographic three-fold axis, 3.35 Å from the $N_{\varepsilon 2}$ atom of Gln16. The $O_{\varepsilon 3}$ atom of Gln16 forms a hydrogen bond with a water molecule (red sphere) c. A view of the GCN4-pI$_Q$I structure, looking down the helices from the N to the C-terminus. All side chains are displayed and the chloride is depicted as a yellow ball in the center of the trimer. GCN4-pI$_Q$I is a trimeric coiled coil with three right-handed α-helices wrapped around each other in a left-handed superhelical twist. d. A side, stereo view of GCN4-pI$_Q$I. Side chains of the residues in the **a** and **d** positions are displayed. The chloride ion is represented as a yellow ball. Figures a and b were generated with O (Jones *et al.*, 1991), and figures c and d were generated with InsightII (Biosym).

a
b
c
d

TABLE 1: GCN4-pI_QI was crystallized at 20°C using vapor diffusion. 10 mg/ml protein solution was mixed 1:1 with a reservoir solution containing 75 mM NaAc pH 4.5, 20 mM $MgCl_2$, and 23% methyl pentanediol (MPD) and equilibrated against the reservoir solution for several days. Most of the crystals appeared as long, thin needles and diffracted to only 3.0 Å. However, crystals grown using a macroseeding protocol (Thaller *et al.*, 1985) diffracted to 1.8 Å. Crystals belong to the space group P321 with one monomer per asymmetric unit and solvent content of 54%. A crystal with dimensions 0.3 x 0.2 x 0.2 mm^3 was flash-frozen using an X-stream cryogenic crystal cooler (Molecular Structure Corporation, MSC, The Woodlands, Texas). Diffraction data were collected using an Raxis IV detector mounted on a Rigaku RU300 rotating anode generator (MSC). The diffraction data were indexed using DENZO and scaled using SCALEPACK (Otwinowski, 1993). Initial phases were determined by molecular replacement using GCN4-pII (Ile16Ala) as the search model in the program AMORE (Navaza, 1994). Interestingly, using a trimer as a search model produced an unambiguous rotation/translation solution, whereas a search with a monomer did not. This observation is probably due to the non-globular shape of each protomer of the coiled coil. Given an integration radius limited to 15 Å by a small unit cell volume, the trimer provides more specific intramolecular vectors than the monomer. One monomer of the trimeric solution was chosen as a model with which to begin refinement, since the GCN4-pI_QI crystal contains one monomer per asymmetric unit. Refinement was first performed using XPLOR (Brünger, 1996) and subsequently with CNS (Brünger *et al.*, 1998). This involved an initial round of rigid body refinement and group B factor refinement. The glutamine side chain was then built into electron density using O (Jones *et al.*, 1991). Additional refinement steps included rounds of simulated annealing, positional refinement, and individual B factor refinement, as well as model rebuilding using O. When the R_{free} reached approximately 35%, water molecules were added manually. Bulk solvent and anisotropic B-factor corrections were applied, followed by additional positional and individual B factor refinement. The final structure was verified using simulated annealing omit maps, omitting one residue at a time and heating the structure to 1000 K before cooling. The R-factor for the final model is 21.1% with an R_{free} of 23.6%. According to PROCHECK (Laskowski et al., 1993), all bond lengths and angles are good, and all residues lie within the most preferred region of the Ramachandran plot for α-helices.

Table 1. Data Processing and refinement statistics

Space group	P321
Cell constants (Å)	a=b=39.66
	c=47.47
	α=β=90
	γ=120
R_{merge}[a] (%)	3.6
Multiplicity	6.7
Completeness (%)	94.9
No. of reflections (working/free)	3652/376
R_{cryst}[b] (20.0-1.8 Å) (%)	21.1
R_{free}[b] (20.0-1.8 Å) (%)	23.6
Average B-factor (Å^2)	29.0
r.m.s. standard deviations from target values	
Bond length (Å)	0.011
Bond angles (°)	1.2
Dihedrals (°)	14.4

[a] $R_{merge} = \Sigma\Sigma_j|I_j - \langle I\rangle|/\Sigma|\langle I\rangle|$, where I_j is the intensity measurement for reflection j and $\langle I\rangle$ is the mean intensity for multiply recorded reflections.

[b] $R_{cryst,\ free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where the crystallographic and free R factors are calculated using the working and test reflection sets, respectively. The test reflections included 10% of the total reflections, which were chosen before refinement of the initial model and were not used during refinement.

CHAPTER 3

INHIBITING HIV-1 ENTRY: DISCOVERY OF D-PEPTIDE INHIBITORS THAT TARGET THE GP41 COILED-COIL POCKET

The work described in Chapter 3 was a collaborative effort. Michael Burgess, James Pang and Ben Sanford provided peptide synthesis services. I designed and characterized IQN17 biophysically as well as performed the mirror-image phage display experiments. The syncytia assays were performed by me, and the infectivity assays were performed by Heng Chhay under my supervision. IQN17 and D10-p1 were crystallized by Lily Hong under my supervision. Lily and I soaked the crystals to obtain a heavy metal derivative. Vladimir Malashkevich and I contributed equally to early molecular replacement efforts to solve the x-ray crystal structure. I collected the MAD diffraction data that Vladimir used to obtain the final electron density map. Peter Carr and I collaboratively obtained the 1D NMR data, and Peter performed all of the 2D NMR data himself.

ABSTRACT

The HIV-1 gp41 protein promotes viral entry by mediating the fusion of viral and cellular membranes. A prominent hydrophobic pocket on the surface of a central trimeric coiled coil within gp41 was previously identified as an attractive potential target for drugs that inhibit HIV-1 entry into cells. However, identification of pocket-binding molecules has been problematic because this region is only transiently exposed during the viral infection process, and the corresponding peptide fragments of gp41 aggregate. We have designed and constructed a non-aggregating peptide, IQN17, which properly presents this pocket. Utilizing IQN17 and the mirror-image phage-display method, we have identified cyclic, D-peptide inhibitors of HIV-1 infection which share a sequence motif. A 1.5Å co-crystal structure of IQN17 in complex with a D-peptide, and subsequent NMR studies, show that conserved residues of these inhibitors make intimate contact with the gp41 pocket. Our studies provide validation of the pocket per se as a target for drug development. IQN17 and these D-peptide inhibitors are likely to be useful for the 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 (Berger et al., 1999; Chan and Kim, 1998; Kwong et al., 1998; 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 (review, Chan and Kim, 1998). Binding of gp120/gp41 to cellular receptors (CD4 and a chemokine co-receptor such as CCR-5 or CXCR-4) induces a conformational change in the envelope glycoprotein. A transient species results, termed the pre-hairpin 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 pre-hairpin intermediate resolves to a trimer-of-hairpins structure that likely represents the fusion-active state of gp41 (Lu et al., 1995; Blacklow et al., 1995) as seen in the X-ray crystal structure of a protease-resistant core of gp41 (Chan et al., 1997; Weissenhorn et al., 1997; Tan 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 pre-hairpin intermediate (Figure 1).

A phase I clinical trial showed that DP178 (recently renamed T-20) has an anti-viral effect when injected into infected individuals (Kilby et al., 1998). This result provides compelling evidence that the pre-hairpin intermediate is a useful target for anti-HIV therapy in humans. However, very large amounts of T-20 (~200mg/day) are required to observe an anti-viral 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 pre-hairpin 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 daltons. (In comparison, T-20 and C34 have molecular weights greater than 4000 daltons). Based on the crystal structure of the gp41 core, the pre-hairpin 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 daltons, 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; Chan et al., 1998). (i) Mutagenesis studies indicate that the N-peptide residues forming the pocket are critical for membrane fusion (Cao et al., 1993; Chen et al., 1993; Dubay et al., 1992; Weng and Weiss, 1998; Wild et al., 1994; D.C. Chan and P.S.K., unpublished results). (ii) Studies of C34 variants show that C34 inhibitory activity depends on its ability to bind to the pocket (Chan et al., 1998). (iii) Therapeutic agents that target this pocket would likely be relatively elusive to the emergence of resistant viral strains because: (a) residues comprising this region are highly conserved among HIV-1 isolates (see Figure 1 legend) and (b) 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 and Kim, 1997; Lu et al., 1995, and references therein), 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 1.5Å X-ray co-crystal structure. These D-peptides inhibit gp41-mediated cell-cell fusion and HIV-1 infection, providing validation of 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. In the absence of C-peptides, N-peptides aggregate (Lu and Kim, 1997; Lu et al., 1995, and references therein) and therefore are not effective for screening small molecules.

We have designed a molecule, denoted IQN17, in which a soluble trimeric coiled coil (GCN4-pIQI; 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-pIQI (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-pIQI were fused to the last 17 residues of N36 (there is a one 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-pIQI 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 non-specific 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-pIQI' or D-IQN17(G572W). D-GCN4-pIQI' is a D-amino acid version of GCN4-pIQI 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, twelve IQN17-specific phage were identified and labeled ϕ10-p1 through ϕ10-p12 (see Experimental Procedures). Nine of the twelve phage clones are pocket-specific binders, with some showing over a thousand-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 co-receptor, 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 50μM, 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 3.6μM (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 (review, Hendrickson, 1991) of an osmium derivative. The final structure was refined against data to 1.5Å from a native 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 gp41 hydrophobic pocket in the IQN17/D10-p1 complex is almost identical to that in the wild type HIV gp41 structure (Chan et al., 1997), with a C_α RMSD of 0.65Å. The conformation of the GCN4-pIQI fraction of the molecule is also well preserved, with a C_α RMSD from the original GCN4-pIQI (Eckert et al., 1998) of 0.40Å.

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 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 sidechains are significantly different, owing to the opposite handedness of the inhibitors (Figure 6).

Of the sixteen residues in D10-p1, only six 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 $O_{\varepsilon 1}$ – Trp-12 $N_{\varepsilon 1}$ (2.8Å). The potential exists for a second intermolecular hydrogen bond (Gln-577 $N_{\varepsilon 2}$ – Glu-9 O; 3.1Å) with less ideal geometry. It is not apparent why Glu-9 is absolutely conserved (Figure 3C), although its sidechain carboxylate does bend back to accept a hydrogen bond from its own backbone amide (2.7Å), 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 sidechains (relative to Trp-631 and Ile-635 in C34) are accompanied by slight changes in the shape of the pocket: (i) 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; (ii) Gln-577 lining the bottom, pushes further into the pocket, raising the bottom wall (mainly due to a rotation about χ_2); and (iii) a rotamer change of Leu-565 slightly lowers the top wall of the pocket. 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 co-crystal 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 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 $H_{\varepsilon 3}$) 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 sidechain, almost certainly Trp-571. For several of the complexes studied, 2D NOESY experiments also indicate contact between this sidechain 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 pre-hairpin 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 pre-hairpin 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 (Chan et al., 1998; Wild et al., 1995). 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 anti-viral 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 400Å^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. Such drugs would also demonstrate the feasibility of therapeutics which

interfere with protein-protein interactions (as opposed to enzyme activity) and would be a useful addition to the current regimen used to treat HIV-1 infection.

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 (~35 residues). The D-peptides are also larger than desired for oral administration (≥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 sidechains 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 co-crystal structure enables rational drug development approaches.

Towards 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 pre-hairpin intermediate is a vulnerable target for neutralizing antibodies. The most obvious approach towards 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/ml, 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 one hour 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 two hours, 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 hours, and in the seventh round, they were incubated for twelve hours. 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: three minutes, Round 4: five minutes, Round 5: seven minutes, Round 6: 10 minutes, Round 7: 12 minutes). 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 fifteen minutes at 4°C). In all subsequent rounds, the phage were eluted by the addition of two micrograms of trypsin (Sigma, T-8658) in 100 μL of phage-binding buffer and 2.5 mM $CaCl_2$ with one hour incubation at 37 °C. To determine recovery, a dilution of the eluted phage was used to infect K91-kan cells. After a one hour 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). Non-specific 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 six hours for Round 3, two hours for Round 4 and one hour 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-pIQI', 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-Leu). The peptides were cleaved from the resin using Reagent K (King et al., 1990).

IQN17 contains 29 residues derived from GCN4-pIQI on the N-terminus and 17 residues from the C-terminus of N36 on the C-terminus. There is a one residue overlap between GCN4-pIQI and the N36 region, making the peptide 45 residues long. To improve solubility, three amino-acid substitutions were made in the GCN4-pIQI region of IQN17, as compared to the original GCN4-pIQI sequence (Eckert et al., 1998). These substitutions are L13E, Y17K and H18K. Thus, the sequence of IQN17 is:

Ac–RMKQIEDKIEEIESKQKKIENEIARIKKLLQLTVWGIKQLQARIL–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-pIQI' 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 Na 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 re-determined 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 eighteen hours after starting, or changing centrifugation speed, the absorbance was recorded at 229 and 230 nm.

Cell/Cell Fusion Assay

Inhibition of cell/cell fusion (i.e., syncytia formation) was assayed by co-culturing 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 & 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 x 10^4 CHO cells and 4 x 10^4 HeLa cells were then added. Approximately twenty hours after co-culturing 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 [$y = k/(1 + [peptide]/ IC_{50})$], 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 co-transfecting 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 one microliter of the stock added to one microliter 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 five microliters 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, 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 ($\lambda1$, $\lambda2$, $\lambda3$) were scaled with SCALEIT to the remote wavelength ($\lambda4$) 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-pIQI and HIV gp41 N36/C34 structures (Chan et al., 1997; Eckert et al., 1998), with sidechains 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-pIQI 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 was processed in 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 0.3-1.0 mM for individual peptides, and 0.8-1.0 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 55ms (NOESY) and 42 ms (TOCSY). Spectral widths of 11,111 Hz and 5555 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 anti-viral assays, Michael Burgess, James Pang and Ben Sanford for peptide synthesis, Alexandra Evindar for editorial and graphics assistance, and Drs. Craig Ogata and 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, Drs. David Chan and 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).

Protein Data Bank ID Code

The coordinates for the D10-p1/IQN17 complex have been deposited in the Protein Data Bank with the ID code 1CZQ, and are available immediately at the website (http://web.wi.mit.edu/kim).

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. 1999 *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., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta. Crystallogr. *D54*, 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., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996). Protein NMR Spectroscopy: Principles and Practice (San Diego, CA: Academic Press).

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.

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

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.

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., and Kim, P. S. (1998). HIV entry and its inhibition. Cell *93*, 681-684.

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, 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.

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-pIQI, 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. Nature 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, 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.

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. *D47*, 110-119.

Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M., and Saag, M. S. (1998). Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. Nature 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., Blacklow, S. C., and Kim, P. S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nature Struct. Biol. 2, 1075-1082.

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

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, Isaacs, N. & Bailey, S., eds. (SERC, Daresbury Laboratory, Warrington, England), 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., 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.

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.

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.

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

Schematic representation of a working model for HIV membrane fusion (review, Chan & 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 co-receptor), a conformational change results in formation of the transient pre-hairpin 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 pre-hairpin intermediate and inhibit fusion in a dominant-negative manner ("Inhibited"). In the absence of inhibitors, the pre-hairpin intermediate resolves to the hairpin structure and membrane fusion results ("Hairpin/Fusion"), although it is not known if 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: C34 [628-661]; DP178, also called T-20 [638-673]; and T649 [628-663]. Adapted from Chan & Kim (1998).

The inset depicts the 2.0Å 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 eleven 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 eleven 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.

Figure 1

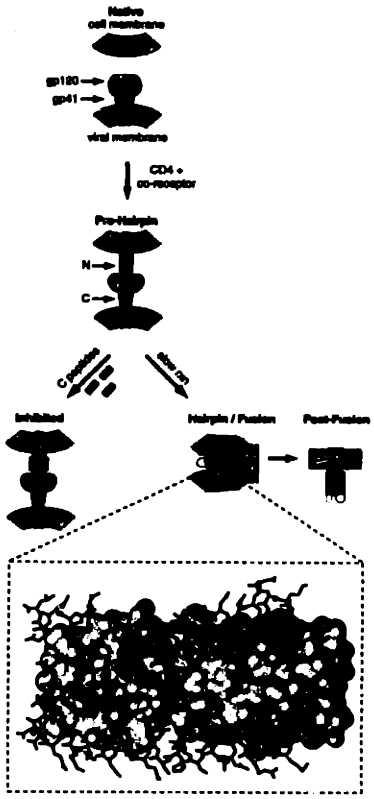

Native cell membrane
gp120
gp41
viral membrane
CD4 + co-receptor
Pre-Hairpin
N
C
C peptide
short DP
Inhibited
Hairpin / Fusion
Post-Fusion

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

(A) Helical wheel diagram of IQN17. The seventeen residues indicated in blue are from gp41 (residues 565-581 of HXB2 gp160). The residues in black are from GCN4-pIQI, and the residues in red correspond to three surface substitutions in the GCN4-pIQI 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-pIQI 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 is 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 2

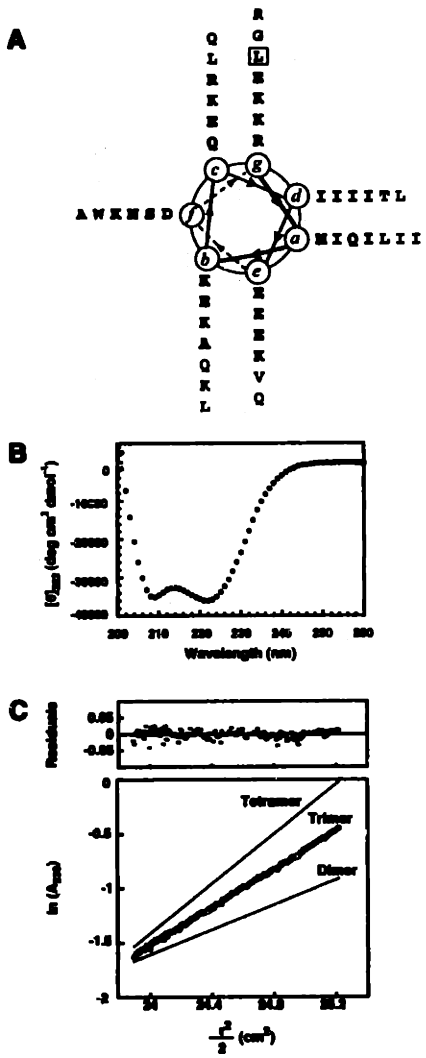


Figure 3. Mirror-Image Phage Display Using D-IQN17 as a Target

(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 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-pIQI', and (iv) no target. In this schematic, the GCN4-pIQI' 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-pIQI' is a D-amino acid version of GCN4-pIQI (Eckert et al., 1998) that contains the same surface mutations as IQN17. The G572W substitution in 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 non-specific 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).

Figure 3

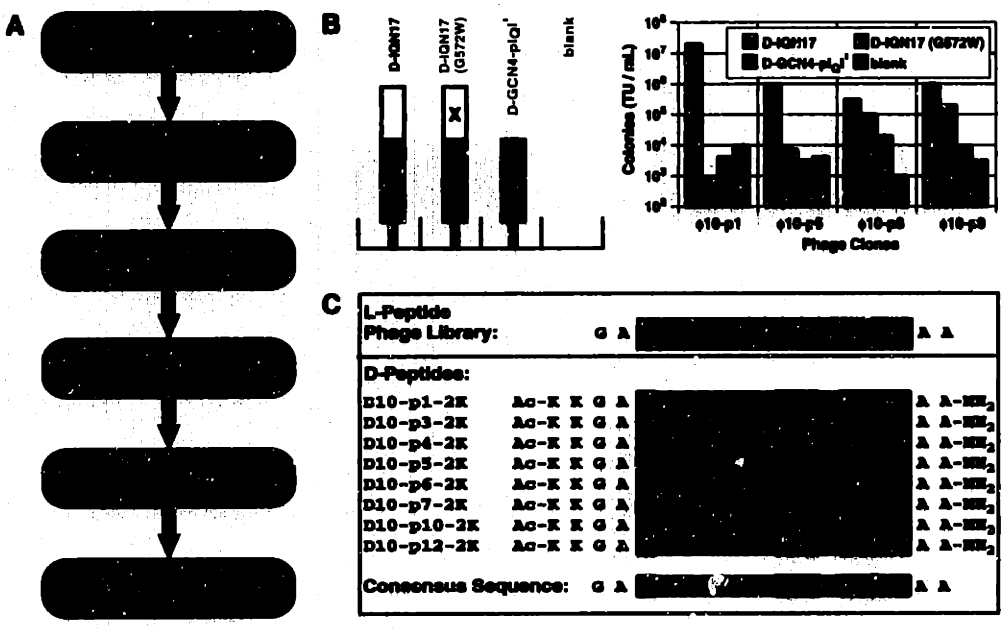

A
B
D-IQN17
D-IQN17 (G572W)
D-GCN4-pIqI'
blank
Colonies (TU / mL)
Phage Clones
ϕ10-p1
ϕ10-p5
ϕ10-p8
ϕ10-p9
C
L-Peptide Phage Library:
D-Peptides:
D10-p1-2K
D10-p3-2K
D10-p4-2K
D10-p5-2K
D10-p6-2K
D10-p7-2K
D10-p10-2K
D10-p12-2K
Ac-K K G A
A A-NH2
Consensus Sequence:

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

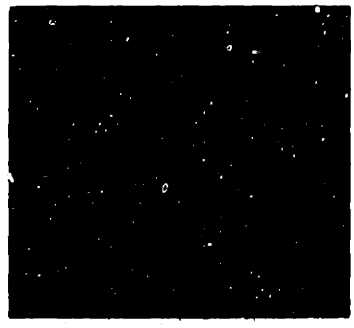
(A) Photographs of cell culture assays of fusion between CHO cells expressing HIV-1 envelope (Kozarsky et al., 1989) and CD4-HeLa cells with a β-galactosidase reporter gene (Kimpton & Emerman, 1992). Cells were co-cultured in the absence or presence of D10-p5-2K, and then incubated with β-galactosidase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

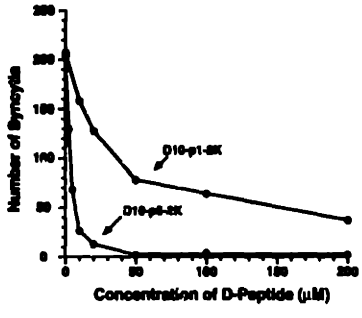
Figure 4

A



Control 50 μM D10-p5-2K

B



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	84	
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	46	96
D10-p12-2K	130	270

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-pIQI' part of the chimera (dark blue) and the HIV gp41 hydrophobic segment (grey) 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) Stereoview 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) Stereoview 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).

Figure 5

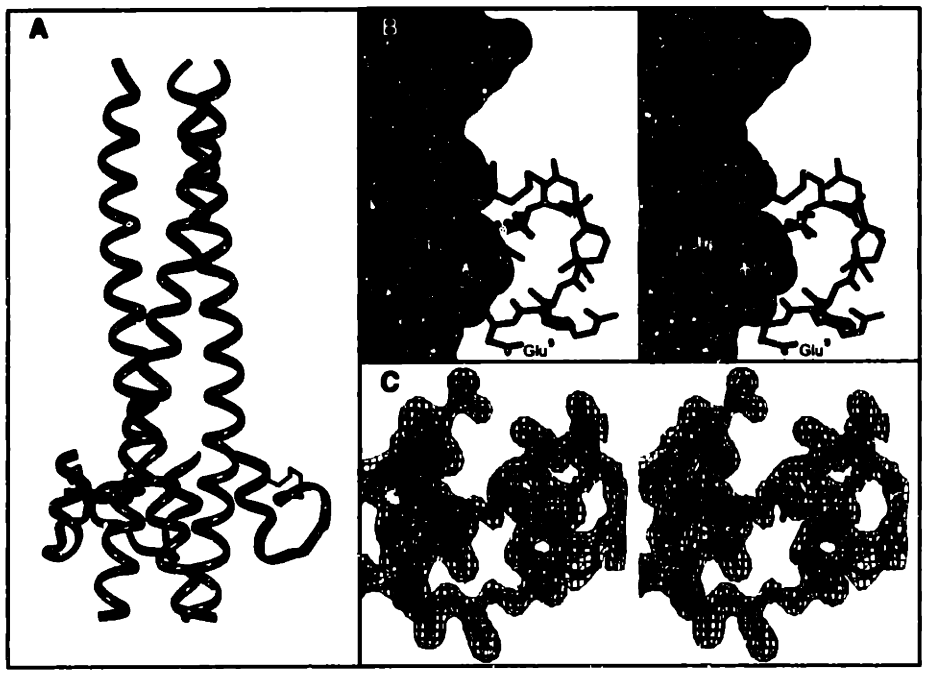


Figure 6. Comparison of the Pocket in the C34 and D10-p1 Bound States

Stereoview of the structural superposition between IQN17/D10-p1 (green) and HIV 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, and Lys-574, as labeled). Figure drawn with Insight II 98.0 (Molecular Simulations Inc.).

Figure 6

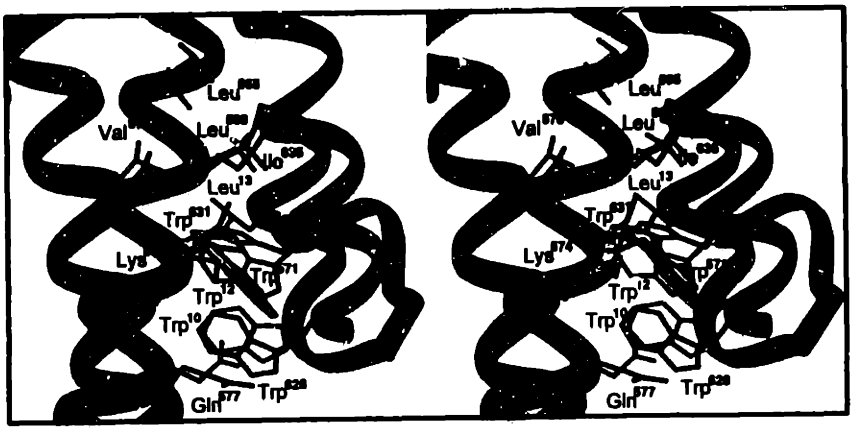

Leu
Val
Leu
Ile
Leu
Trp
Lys
Trp
Trp
Trp
Gln
Trp
Leu
Val
Leu
Leu
Trp
Lys
Trp
Trp
Gln
Trp

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 sidechain 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.

Figure 7

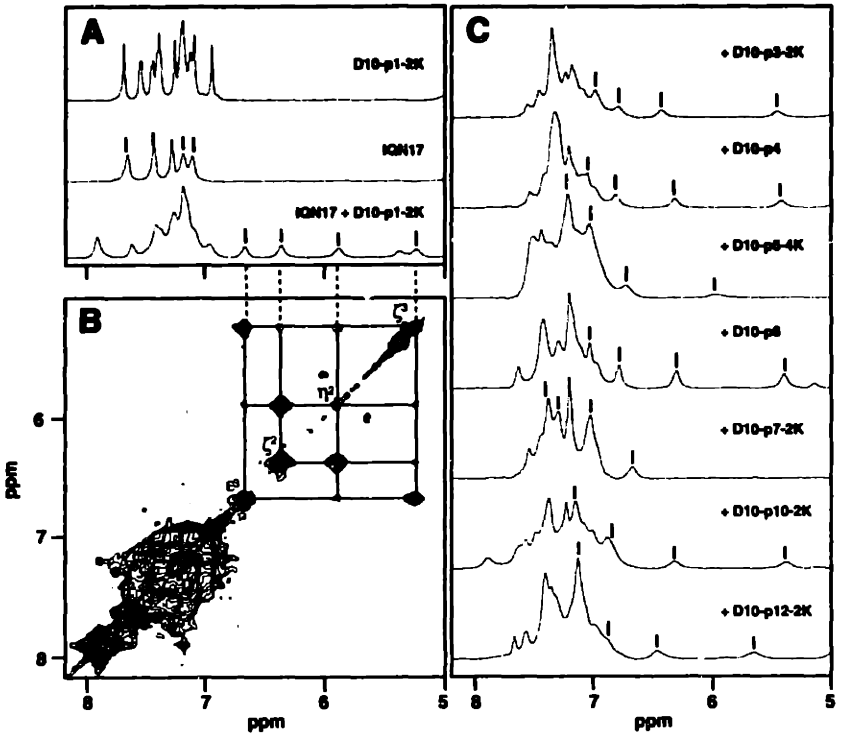

A
D10-p1-2K
IQN17
IQN17 + D10-p1-2K
B
C
+ D10-p3-2K
+ D10-p4
+ D10-p5-4K
+ D10-p6
+ D10-p7-2K
+ D10-p10-2K
+ D10-p12-2K
ppm
ppm
ppm
6
7
8
5

Table 1. Crystallographic and Refinement Statistics

Data collection

Crystal	λ (Å)	Completeness (%)	R_{sym}[1] (%)	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}[2] (%)	R_{cullis}[3] Acentric	R_{cullis}[3] Centric	R_{cullis}[3] Anom.	Ph. Power[4] Acentric	Ph. Power[4] Centric	Occ.[5]	Anom. Occ.[5]
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}[6]	R_{free}[6]	R.m.s. deviations bonds (Å)	angles (°
IQN17/D10-p1	516	150	1	10.0 - 1.5	13549	0.214	0.245	0.012	1.498

[1] $R_{sym} = \Sigma\Sigma_j|I_j - \langle I\rangle| / \Sigma\Sigma_j|\langle I\rangle|$, where I_j is the recorded intensity of the reflection j and $\langle I\rangle$ is the mean recorded intensity over multiple recordings.

[2] $R_{iso} = \Sigma||F_{(\lambda i)} \pm F_{(\lambda 4)}| - |F_{(\lambda i)}|| / \Sigma|F_{(\lambda 4)}|$, where $F_{(\lambda i)}$ is the structure factor at wavelength λi and $F_{(\lambda 4)}$ is the structure factor at the reference wavelength λ4.

[3] $R_{cullis} = \Sigma||F_{(\lambda i)} \pm F_{(\lambda 4)}| - |F_{h(\lambda i),c}|| / \Sigma|F_{(\lambda i)} \pm F_{(\lambda 4)}|$, where $F_{h(\lambda i),c}$ is the calculated heavy atom structure factor.

[4] 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.

[5] Occupancies are values output from MLPHARE.

[6] $R_{cryst,\ free} = \Sigma||F_{obs}| - |F_{calc}|| / |F_{obs}|$, where the crystallographic and free R factors are calculated using the working and test sets, respectively. Test set contained 10% of reflections.

CHAPTER 4

CONCLUSION: VIRAL MEMBRANE FUSION AND ITS INHIBITION

The fusion of lipid bilayers is involved in diverse biological processes such as fertilization, vesicle trafficking, muscle development, and viral infection. The most extensively studied of these events is the membrane fusion of viral infection. When an enveloped virus infects a host cell, it fuses its membrane with the host cell membrane, allowing the virus to empty its contents into the cell. The fusion event is mediated by virally encoded surface glycoproteins.

After binding to the appropriate cellular receptor(s) on the host cell surface, enveloped viruses enter cells in one of two ways: a) they fuse with the host cell membrane at the cellular surface, and therefore at neutral pH, or b) after endocytosis into clathrin-coated pits, they fuse with the endosomal membrane when the pH is lowered. Viruses such as HIV and human respiratory syncytial virus use the former mode of viral entry, whereas viruses like influenza and rabies utilize the latter method.

Despite the point of entry, be it cellular surface or endosome, enveloped virus glycoproteins share a variety of features (1). Most envelope glycoproteins are integral membrane proteins that are synthesized as a precursor and then cleaved into two subunits that remain closely associated with each other. These proteins form higher order oligomers and are glycosylated. Because of their surface exposure, much of the host's immune response targets these glycoproteins. Lastly, these envelope glycoproteins contain a short peptide within their sequence called the "fusion peptide," which is required for mediating membrane fusion (see review in 2, 3).

Potential fusion peptide domains are identified through sequence analysis. They are usually hydrophobic and rich in glycine and alanine residues and extend 20 to 25 amino acids in length. Their importance in membrane fusion is analyzed by point mutations or multiple base mutations. Changes in the fusion peptide sequence either ablate or alter fusion activity (4-9). The fusion peptide region is thought to insert into the target cell membrane, somehow encouraging fusion (10). Indeed, when the influenza, rabies and vesicular stomatitis virus fusion proteins were exposed to target bilayers with photoactivatable lipid probes, only the predicted fusion peptide domains were labeled (11, 12). In addition, synthetic peptides corresponding to this region can bind and destabilize lipid bilayers (13-18). The fusion peptide is almost always found at the N terminus of the membrane anchored glycoprotein subunit following the cleavage event, although in several viruses, such as in Rous sarcoma virus, it is predicted to be located internally in the membrane fusion protein (19). The fusion proteins of two viral families,

togavirus and flavivirus, do not undergo cleavage for fusion activation, and their fusion peptides are also predicted to be located in the interior of the uncleaved fusion protein. The internal localization has been verified by mutagenesis and photolabeling experiments (12, 20, 21).

Generally, viral envelope fusion proteins proceed through a series of conformational changes in order to fuse with the host cell. The initial cleavage of the glycoprotein precursor leaves the complex in a metastable state, primed for fusion, although the fusion peptide is not exposed. Then, either due to binding to host cell receptors or exposure to low pH, conformational changes occur which expose the fusion peptide and allow close juxtaposition of the viral and host cell membranes, leading to membrane fusion.

This chapter will review the mechanisms of enveloped viral membrane fusion and inhibition. First, influenza hemagglutinin will be discussed. It is the only viral envelope glycoprotein for which high-resolution structural information is known for both the neutral-pH (pre-fusogenic) and low-pH conformations (22-25). Because of this and the abundance of biochemical and biophysical data available on hemagglutinin, it has been the paradigm for studying viral membrane fusion for many years (for recent reviews see 1, 26, 27). Next, retroviral membrane fusion will be covered, using HIV-1 as an example. Structural data and inhibition efforts have demonstrated the similarities between HIV viral entry and influenza viral entry. Finally, several other families of enveloped animal viruses will be discussed, such as paramyxoviruses and filoviruses like ebola. These viruses seem to use a similar method of viral membrane fusion. With all of the data from different viruses pooled together, a clear view of membrane fusion emerges, as well as possible strategies for combating the infections of widely different viruses.

Influenza Hemagglutinin

The vast amount of experimental knowledge on influenza hemagglutinin has made it the prototypic enveloped viral membrane fusion glycoprotein (for reviews see 1, 13, 28). Hemagglutinin (HA) is the sole protein on the influenza viral surface required for membrane fusion with the host cell. It is synthesized as a fusion-incompetent precursor, termed HA0. HA0 is then proteolytically cleaved into two subunits, HA1 and HA2, which remain associated via a disulphide linkage (Figure 1A). This cleavage event

activates the fusion potential of hemagglutinin. HA1 is responsible for binding the viral receptor, sialic acid, on the host cell, thereby initiating contact between the virus and the cell. HA2 stays anchored in the viral membrane and contains the characteristic "fusion peptide" at its N terminus and is therefore indispensable for fusion.

Early Evidence for pH-induced Conformational Changes

Several decades of research have illuminated a complex and elegant process by which influenza hemagglutinin mediates viral membrane fusion. Early studies of the influenza virus demonstrated that it enters cells via the endosomes, and that exposure to low pH activates the membrane fusion potential of hemagglutinin (29-31). This observation led to intense scrutiny of the effect of low pH on hemagglutinin structure. Several experiments suggested that hemagglutinin undergoes conformational changes after exposure to low pH. First, the proteolytic vulnerability of HA changes (32). Second, HA becomes more hydrophobic, forming aggregates and acquiring the ability to bind detergents and liposomes (33). Third, electron microscopy analysis identified a change in the shape of hemagglutinin, including a partial dissociation of globular head domains on the top of the molecule (33, 34). Finally, antibody recognition differs (32, 35-39).

Specifically, under an electron microscope, the neutral-pH HA molecule looks like a rod, protruding from the virus (33, 34). Upon exposure to low pH, the protein becomes much thinner and extended, implying significant conformational changes. There are prominent antibody epitopes found on the tip and other regions of the protein. The changes in the antigenicity of HA with a change in pH implied that structural rearrangements occurred from the base of the molecule to the tip, as the recognition of most epitopes was altered upon the induction of low pH (36-38). For example, monoclonal antibodies reactive with the tip of neutral-pH HA are unable to immunoprecipitate HA after low-pH-induced conformational changes (35). Likewise, antibodies specific for epitopes exposed in the low-pH-activated HA, such as antibodies raised against the N-terminal fusion peptide, are unable to immunoprecipitate neutral-pH HA (35, 39).

Fortunately, after much work, atomic structural information of both the neutral-pH (pre-fusogenic) and low-pH conformations of hemagglutinin became available and shed light on the surmised conformational changes in HA (23-25). In addition, there is also a crystallographic structure of HA0, uncleaved hemagglutinin, which shows the non-fusogenic form of HA, before it is primed for fusion (22).

The neutral-pH structure was solved by crystallization of a large portion of the HA ectodomain (BHA) that was cleaved from the viral surface by bromelain (25, 40) (Figure 1B and 2B). BHA lacks the hydrophobic region that anchors hemagglutinin to the viral membrane. It is a long trimer and extends 135 Å from the C terminus (where it would be attached to the viral membrane). At the top of the structure (opposite end from the viral membrane) are three globular head domains consisting entirely of HA1, and responsible for binding to sialic acid. The head domains sit atop a stalk, composed of the remainder of HA1 and of HA2. This stalk contains a central trimeric coiled coil at its core, composed entirely of HA2. Importantly, the N terminus of HA2, which contains the fusion peptide, is buried in the neutral-pH structure, verifying the early biochemical studies. The fusion peptide is located approximately 35 Å from the C terminus of the structure and 100 Å from the distal tip, stabilized by hydrogen bonds.

More recently, the crystal structure of the uncleaved hemagglutinin, HA0, was solved (22) (Figure 1C). For crystallization, hemagglutinin with a mutated cleavage site was expressed recombinantly in vaccinia virus and clipped off the surface of the infected cells with trypsin. The resultant molecule is called R329Q $HA0_S$. The structure is almost identical to the BHA structure, except for eighteen residues surrounding the cleavage site. In the BHA structure, the C terminus of HA1 and the N terminus of HA2 (formed by proteolytic cleavage of HA0) are approximately 22 Å from each other, whereas in the uncleaved molecule, they form an extended, uncleaved, loop. This exposes the fusion peptide region to solvent. Importantly, there is a pocket lined by ionizable residues (two aspartic acids and a histidine) next to the extended loop. In the cleaved, neutral-pH structure, the first ten residues of the fusion peptide are buried in this pocket. Interestingly, the ionizable residues have been shown to be important in the fusion activation caused by exposure to low pH (41).

The biochemical studies predicted that membrane fusion requires the exposure of the fusion peptide that is buried in the neutral-pH structure. Without a high-resolution structure of the low-pH conformation, however, the necessary structural rearrangements required to accomplish this goal were unknown. Twelve years after the neutral-pH HA structure was solved, synthetic peptide experiments led to a valuable structural model for fusion peptide exposure (42). Computational methods identified a region of the HA ectodomain that had high propensity for forming a coiled coil. However, in the neutral-pH structure, this region maintained an extended loop conformation (white residues in Figure 2B). It was proposed that in the fusogenic, low-pH structure, these residues formed a coiled coil. This would extend the central trimeric coiled coil and displace the fusion peptide to the opposite end of the molecule. This proposal was called the "spring-loaded" mechanism (42). One year later, the x-ray crystal structure of the low-pH-activated HA was solved and confirmed the spring-loaded mechanism (Figure 2A) (24).

There are two x-ray crystal structures of the low-pH-converted HA (23, 24). The first was solved using a fragment called TBHA2. TBHA2 was made by exposing BHA to low pH, and then removing the HA1 head domain with trypsin and the fusion peptide with thermolysin. This structure revealed three very significant conformational changes from the neutral-pH structure (Figure 2C). First, the conformational change of a loop to helix at the N terminus of the coiled coil, as described above, propels the short helix to the top of the structure. Although the fusion peptide is not present in the structure, this conformational change would theoretically move the fusion peptide 100 Å from its position in the neutral-pH conformation. Second, at the base of the trimeric coiled coil, a helical region of BHA becomes a loop in TBHA2, reversing the direction of what was the C-terminal end of the coiled coil. Third, the C-terminal residues that were densely packed at the base of the neutral-pH conformation become mostly extended and continue along the side of the central coiled coil in an anti-parallel direction. The final twenty or so residues are disordered and therefore not seen in the structure.

Recently, a new low-pH HA structure provided a more concrete picture of the N and C-terminal regions of the HA ectodomain (23). For this structure, a stable domain was identified through proteolysis of a recombinantly expressed HA2 ectodomain. The identified domain, termed EHA2, was then expressed in *Escherichia coli* . In the EHA2 crystal structure, it is clear that both the N terminus and the C terminus of the ectodomain are located at the same end of the structure. The residues on the C-terminal end of the HA ectodomain are in an extended conformation and continue up the side of the trimeric

coiled coil in an anti-parallel fashion, with each monomer fitting into the groove caused by two of the coiled coil helices. Only seven residues remain between these residues and the transmembrane domain. The N-terminal end of the coiled coil is capped by the final three residues of EHA2, and these residues also form extensive contacts with the C-terminal residues. EHA2 is an extremely stable structure due to this N-terminal cap, and is proposed to be the final low-energy conformation of the cleaved fusion protein. This same orientation, in which both termini of the ectodomain are located at the same end of the folded molecule, is seen for many other enveloped virus fusion proteins and will be discussed later in this chapter.

There are several reasons why the low-pH-converted structures are thought to represent the fusion-active conformation of HA. First, the molecules used for the structural investigations were prepared by exposure to low pH (23,24), and low pH is required to activate the fusion potential of HA (29-31). Second, the resulting conformational changes in the soluble ectodomain fragments of HA are consistent with biochemical and morphological observations of the conformational changes that are required for HA-mediated fusion (32-39). As mentioned above, these observations predicted an exposure of the fusion peptide domain. The low-pH structures display a large displacement of the fusion peptide region from a protected pocket to an exposed tip. The biochemical studies also predicted a partial dissociation of HA1 from HA2. Consistent with this observation, the two structures represent HA2 molecules that are stable in the absence of HA1 (23, 44, 45). Also, the high-resolution structures are consistent with the shape of HA observed previously in electron micrographs of low-pH HA on the viral surface. Third, specific mutations that alter fusion activity map to regions that change conformation between the neutral-pH and low-pH structures (24, 36). Specifically, proline mutations in the region that undergoes the loop to helix transition impaired fusion activity (43). Finally, the low-pH structure is consistent with the biochemical data of inactivated virus. When virus is exposed to low pH in the absence of a target membrane, it becomes inactivated, presumably by inserting its fusion peptide into its own membrane (18, 47, 48). The low-pH structures have the N and C termini at the same end of the molecule as they would be in the inactivated form. Indeed, a monoclonal antibody against residues 105-113 of HA2 (the red loop region in Figure 2C) binds to the tip of HA distal to the viral membrane of inactivated virus (49). Therefore, because of the above observations, the low-pH-converted structures are thought to define a fusion-active conformation of HA, which either occurs concomitantly and/or just after membrane fusion. Hereafter, these structures will be referred to as "fusogenic."

HA-Mediated Fusion of Lipid Bilayers

The structural studies discussed up to this point have not addressed one very important aspect of membrane fusion. How does the merger of two lipid bilayers proceed in HA-mediated fusion? This question has been addressed in a variety of model systems, and seemingly contradictory results have emerged. Commonly, a hemagglutinin-expressing cell is exposed to red blood cells (RBCs) which contain sialic acid. Upon the acidification of the exterior medium, membrane fusion between the two cell types ensues. The fusion can be monitored in one of two ways. First, fusion pore formation can be recorded electrophysiologically with patch-clamp methodologies. Second, the RBCs can be labeled with fluorescent lipid probes and/or water-soluble dyes to monitor the mixing of lipid bilayers and cellular contents respectively.

In early studies, HA-expressing fibroblasts were mixed with RBCs containing a fluorescent dye in their cytoplasm. The pH of the exterior medium was lowered, and subsequently the dye transferred to the fibroblast, revealing content mixing between the cells (50). Patch-clamp studies allowed the same event to be monitored electrophysiologically. Suction is created between a microscopic glass pipette tip and the cellular membrane such that the current of the pipette tip and the lipid bilayer are the same. When a pore opens in the bilayer, the current changes and is detected by a probe attached to the pipette. The patch-clamp studies detected a capacitance increase shortly (less than a minute) after the decrease in pH (50). The increase in capacitance comes and goes in the first few milliseconds. The small capacitance probably corresponds to a modest fusion pore of 1-2 nm that flickers open and close (51). Subsequently, the capacitance increases over tens of seconds as the pore dilates, ultimately to a size that would permit passage of the viral contents into the host cell. In support of the pore dilation, the transfer of small soluble dyes from the RBC to the fibroblast during the fusion process occurs prior to the transfer of larger molecules (52). The same fusion pore dynamics are seen in the experimental system of HA-expressing cells fusing to voltage-clamped planar lipid bilayers (53). Planar lipid bilayers can be formed across holes in thin dividers (about 1mm in width) between two aqueous compartments. A voltage can be applied across the membrane and the current can be detected with electrodes in the separate compartments.

Further fluorescent studies have attempted to delineate the course of events in fusion pore formation. Which occurs first, lipid mixing or content mixing? In two

experiments using fibroblasts and RBCs, fusion pore formation, monitored electrophysiologically, was shown to precede the transfer of lipid dye from one membrane to another (54, 55). In a more recent experiment, however, in which HA-expressing fibroblasts were fused to voltage-clamped planar lipid bilayers, lipid dye transfer was shown to precede fusion pore formation (56).

The results of the earlier experiments, where pore formation occurred before lipid mixing, led to the proposal that the fusion pore is proteinaceous, similar to a gap junction (57), and composed of a pool of HA molecules. Indeed, there is a lag time between acidification and pore formation, and this lag time was shown to depend cooperatively on the density of HA on the cell (58). In addition, quick freeze electron microscopy revealed an ordered array of HA trimers at the fusion site (59).

Alternatively, the results of the latter experiment, where lipid dye transfer preceded pore formation, support the proposal that fusion pore formation occurs through a lipid-dependent intermediate. One such model for membrane fusion, the stalk/pore model (Figure 3), has gained experimental support in the HA-mediated fusion system. In this model, the outer membranes fuse first, forming a stalk between the two enveloped entities, each of which still maintain intact inner membranes. Pore formation occurs when the two inner membranes fuse. Membrane curvature drives stalk/pore fusion, with the outer and inner membranes bending in opposite directions. The outer membrane requires a negative curvature (the curvature opposite to that found in micelles), and the inner membrane requires a positive curvature. Particular lipids can favor either positive or negative curvature. For example, lysophosphatidylcholine (LPC) has an 'inverted cone' shape, and favors positive membrane curvature. Oleic acid (OA), on the other hand, has a cone shape and favors negative membrane curvature. When present in the outer membrane in HA-mediated fusion, LPC inhibits fusion and OA promotes it; when present at the same amounts, no effect on fusion is seen (60). Also, LPC in the inner membrane encourages fusion, supporting the notion that fusion occurs through a lipid stalk intermediate (61). Interestingly, inflated HA-expressing cells that are unable to undergo local bending of the membrane do not undergo fusion with RBCs (62).

A fusion intermediate was identified in which lipid mixing occurred in the absence of content mixing. This intermediate, termed hemifusion, was identified when hemagglutinin with a lipid anchor (in place of the transmembrane domain and cytoplasmic tail) was expressed on fibroblasts and mixed with RBCs (63). The lipid-

anchored HA was able to bind RBCs and change conformation in response to acidification just as wild type HA. However, the fusion reaction stopped at the hemifusion state when lipid dye could be transferred between the bilayers of the two cells, but aqueous dye could not be transferred between cells. The lipid transfer is limited to the outer leaflet of the membrane and also occurs in the experimental system that utilizes planar bilayers instead of RBCs (64).

Chernomordik et al. has argued that hemifusion, as described above, is a dead-end route off the physiologically relevant fusion pathway (61). These workers distinguish between two types of hemifusion, "unrestricted" and "restricted." Unrestricted hemifusion is the state described above, in which lipids freely transfer between membranes, but no fusion pore forms. Under certain conditions (non-optimal pH, temperature or HA density) wild type HA can also cause unrestricted hemifusion (61, 65), but even after extended periods of time, according to Chernomordik et al., those hemifusion diaphragms do not to resolve to fusion pores (61). In restricted hemifusion, a lipidic intermediate is formed in the absence of pore formation. However, the membrane lipids are not free to move through the hemifusion diaphragm, and therefore no lipid mixing is observed. This type of hemifusion is proposed to be physiologically relevant. The results of the early membrane fusion experiments, where pore formation preceded lipid dye transfer is consistent with this type of hemifusion diaphragm (54,55). At this point, it is not entirely clear if unrestricted hemifusion is a dead-end pathway. Indeed, under certain conditions, unrestricted hemifusion, caused by lipid-anchored HA or wild type HA under suboptimal conditions, can resolve to pore formation and complete fusion. For example, if the HA-expressing cells are osmotically swollen or if lipids that promote positive membrane curvature (the same curvature found in micelles) are added to the inner membrane, fusion pore formation can proceed from unrestricted hemifusion (64, 65). These latter results imply that unrestricted hemifusion is not a dead-end reaction. Perhaps, though, in extreme cases such as lipid-anchored HA or fusion under sub-optimal conditions, the hemifusion diaphragm extends further than in wild type HA mediated fusion at optimal conditions.

Vast experimental evidence suggests at least several HA molecules conglomerate at the fusion site. There is also a lot of evidence that HA-mediated fusion proceeds through a hemifusion intermediate (be it restricted or unrestricted). One idea does not necessarily preclude the other. Likely, HA molecules as a group encourage the lipid bending required to form the stalk structure. Ideas for how this occurs will be discussed

below. One mystery is left, however. Why does the chronology of pore formation and lipid dye transfer differ when the target membranes differ (RBC vs. planar lipid bilayer)? This is a difficult question, although there are at least a couple of possibilities. First, perhaps in the experiments with the RBCs, the fusion pore detection was much more sensitive than the lipid dye mixing detection. It is likely that even if a few lipid molecules diffused, they would be difficult to see over background. This would imply that a small amount of lipid mixing did occur first, but could not be visualized. Second, the lipid diffusion through a fusion pore in RBCs is known to be inherently slower than the fusion of lipids through a fusion pore with planar lipid bilayers (66-68). Therefore, perhaps a particular component in RBCs, such as RBC-specific transmembrane proteins, impedes lipid fusion in a manner that is unrelated to fusion, and even though there is contact between lipid bilayers in the RBC system, diffusion between them is slow. Despite the difference in the results, in HA-mediated membrane fusion it is possible that the HA transmembrane domains, which are present at high concentrations at the fusion site, impede lipid movement even when there are continuous lipid contacts. As the pore expands, the HA molecules expand and allow more room for lipids to pass. Consistent with that proposal, in the experiment with planar lipid bilayers, one dye (DiI) mixed prior to the detection of a fusion pore, while another dye did not (Rho-PE) (56). DiI has the ability to flip-flop in a lipid bilayer, freely moving from the outer to inner membrane. This would give it the ability to transfer sides of the hemifusion stalk, and therefore transfer quickly from the planar bilayer to the cell.

Mechanism of HA-mediated Membrane Fusion

The high-resolution structures and the membrane fusion experiments, in combination with other data, have yielded significant insight into the membrane fusion mechanism of influenza hemagglutinin. The model described below is derived from several current hypotheses (1, 26, 69, 70) (see Figure 6 for related model for HIV-1) and considers all current experimental data.

In the virus, HA is first expressed as an uncleaved precursor, unable to fuse. Upon cleavage, the N terminus of HA2 containing the fusion peptide is created. The fusion peptide is then tucked into a pocket at the base of HA, primed for low-pH structural rearrangement. Perhaps the protonation of the ionizable residues in the pocket at low pH encourages the movement of the fusion peptide domain. As briefly mentioned above, several pieces of data suggest the neutral-pH, cleaved form of HA is folded in a

metastable state, blocked from its native fold by a kinetic barrier. First, the uncleaved HA0 cannot undergo fusion-activating structural rearrangements at low pH (28). Second, the same conformational changes that are caused by exposure to low pH can also be activated by high temperature or urea denaturation (46). Third, when HA2 is expressed in the absence of HA1, it adopts the fusogenic structure (42, 45)). Most likely when HA folds as a single chain precursor, it folds to its energetically most stable state. Upon cleavage, that state becomes the metastable state for the cleaved molecule, essentially priming the virus for fusion.

Upon exposure to low pH, the fusion peptide is liberated and propelled in a 'spring-loaded' fashion (42) to the N terminus of the central trimeric coiled coil so it can interact with the host cell membrane. Although the HA1 domains are not present in the high-resolution crystal structure, biochemical and electron microscopy studies show that these globular head domains lose their trimeric contacts and slightly dissociate from the compact glycoprotein structure (34, 36-38, 49). It seems likely that this conformational event occurs before the final fusogenic structure, and therefore before the two termini of the HA2 ectodomain are located near the same end of the folded molecule (1). Recently, the proposal of a fusion intermediate has garnered much more extensive experimental support in other enveloped viruses, which will be discussed later in the chapter. The existence of this intermediate may depend on the presence of the host-cell membrane. When radioactive target membranes are present upon exposure to low pH, the HA fusion peptide becomes labeled (18). However, in the absence of a target membrane, HA quickly becomes inactivated after acidification (47, 48). In the inactive state, the fusion peptide interacts with the viral membrane (18), probably forming the fusogenic structure anchored solely in the viral membrane.

In the fusogenic structure, both termini of the HA2 ectodomain are located at the same end of the folded molecule, and the trimeric coiled coil is capped (23). This structure is extremely stable and presumably represents the most thermodynamically stable form of the molecule. After the fusion peptide inserts into the target membrane, HA2 is now spanning two membranes, the viral and host membranes. The subsequent formation of the energetically favorable fusogenic conformation will bring the two membrane-proximal regions close together. Most likely, the energy that is gained by forming this highly stable structure overcomes the unfavorable process of pulling two phosopholipid membranes into close proximity. In order to accomplish bringing the two membranes together, and keeping membrane contact on both ends, the HA molecule

would likely lie on its side in relation to the membranes. In the first fusogenic structure, the C-terminal residues in the HA ectodomain were disordered (24). Flexibility in this region would assist in the tilting of the molecule.

It is not clear why several HA trimers are required at each fusion site. Presumably, the concerted action of HA trimers facilitates stalk formation through local bending of one or both of the membranes involved. It is possible that several fusion peptides inserted in the target membrane can appropriately destabilize that site and contribute to membrane bending. Alternatively, perhaps the concerted effort of these trimers to form their fusogenic structure as described above is needed in order to cause close juxtaposition and fusion of the two lipid bilayers. Finally, in a divergent hypothesis, it has been suggested that the "inactivated" HA2 at the fusion site (those that have inserted the fusion peptide into their own viral membrane) form local dimples in the viral membrane, allowing stalk formation with the closely juxtaposed cellular membrane (71). Regardless of the mechanism of membrane bending, it is clear that the fusogenic structure of HA allows both transmembrane regions of HA2 to occupy the same post-fusogenic membrane, as the viral and cellular membranes become one.

HIV-1

Recently, there has been an increasing emphasis on biochemical and structural studies of the envelope glycoprotein of HIV-1. HIV belongs to the retrovirus family of viruses and as a cause of the fatal disease, AIDS, is a major world health threat. There are currently over 35 million people worldwide infected with the HIV virus. In the past two decades there have been over 50 million deaths attributed to AIDS (Joint United Nations Program on HIV/AIDS, http://www.unaids.org). By gaining knowledge on the different structural conformations of the HIV envelope glycoprotein during fusion, it should be easier to intelligently devise strategies to combat this deadly virus.

Like influenza virus, HIV contains one viral envelope protein responsible for binding to host cells and mediating membrane fusion (for recent reviews see 69, 72, 73). This protein is synthesized as a precursor, termed gp160, and post-translationally cleaved by a cellular convertase into two subunits, gp120 and gp41 (74-77). In the retrovirus family, these subunits are more generally termed the surface subunit (SU) and the transmembrane subunit (TM), respectively. Unlike in influenza HA, these subunits are not covalently linked by a disulphide bond, but rather remain attached noncovalently

(78). gp120 is analogous to HA1 of influenza. It is responsible for attaching the virus to the host cell by binding a receptor, CD4, and then subsequently a coreceptor, one of a family of seven transmembrane domain chemokine receptors (79-87). gp41 is analogous to HA2 of influenza, and remains anchored to the viral membrane by a hydrophobic membrane-spanning domain. The HIV virus initiates membrane fusion at the cellular surface at neutral pH, rather than in the endosomes.

The N-terminal Fusion Peptide

gp41, like HA2, contains a fusion peptide domain at its N terminus (3, 4, 78). This peptide domain is conserved within the retrovirus family (1), and mutagenesis studies have indicated that the hydrophobic nature of particular amino acid positions is required for its membrane fusion potential (4, 5, 9, 88, 89). Also, synthetic peptides corresponding to the fusion peptide have been shown to bind lipids and cause fusion of membranes (90-94). Although the behavior of synthetic peptides may not accurately represent the role of the fusion peptide in the context of the entire envelope glycoprotein, there is at least one correlation between the full protein and the synthetic peptide studies that lend credibility to the peptide studies. Specifically, a valine to glutamate mutation in the fusion peptide region was shown to abolish the fusion activity of gp41 (5). Likewise, the corresponding synthetic fusion peptide was unable to cause vesicle membrane fusion, while the wild type peptide could (95). Therefore, from the studies described above and by analogy to influenza HA, it is assumed the HIV fusion peptide inserts into the target membrane and that this insertion is necessary for the membrane fusion activity of gp41. No direct studies, such as photolabeling experiments in the context of whole envelope protein, have yet been reported to verify this assumption.

Receptor-Induced Conformational Changes

Although there is no pH change to activate membrane fusion, there was early evidence that at least two conformational forms of HIV env exist, and that the conformational change is activated by interaction with CD4 (reviewed in 72). Several conformational probes detected differences in HIV env after exposure to receptor. First, there are differences in epitope accessibility. For example, the V3 loop on gp120 can be recognized after receptor binding (96), and epitopes on gp41 become more accessible (97). Second, gp120 was shown to "shed" from gp41 in laboratory-adapted strains,

perhaps activating gp41 for fusion such as in "spring-loaded" HA (98-101). Third, even in the absence of gp120 shedding in primary isolates, the proteolytic susceptibility of the envelope proteins is altered (102). Fourth, there is increased binding of a fluorescent probe, bis-ANS, that is proposed to contact hydrophobic groups (103). Interestingly, this was also seen in influenza HA after exposure to low pH (104). Also, mutagenesis studies implied the existence of different conformations. Specific point mutations in the coiled-coil domain of gp41 do not alter the proteolytic processing or surface expression of the envelope, yet they jeopardize infectivity and gp41-mediated membrane fusion. Most likely, these mutations disrupt the fusion-active state of the protein, but not the early, non-fusogenic form (105-111).

Biochemical studies have revealed some of the details of the early gp120 conformational changes. Following binding of CD4, novel epitopes are revealed on the surface of gp120 which allow recognition by neutralizing antibodies (112). These neutralizing antibodies, such as 17b, can block binding of gp120/CD4 complexes to coreceptor (113, 114), implying that CD4 binding reveals the previously hidden coreceptor binding site (115). Consistent with this chain of events, after the HIV envelope is exposed to soluble CD4 (116), it can fuse with CD4- negative cells expressing coreceptor (117).

High-Resolution View of gp120

The first high-resolution structural glance of gp120 bound to CD4 and a molecule mimicking the coreceptor was recently provided (118 and reviewed in 120, 121) (Figure 4). The primary sequence of gp120 is composed of five regions of high sequence variability among known HIV isolates (termed V1-V5) interspersed with five conserved regions. The x-ray crystal structure reports a mostly de-glycosylated gp120 conserved core (missing V1, V2, V3 and the N and C termini) bound to two domains of CD4 and the Fab fragment of the monoclonal antibody 17b which represents the coreceptor binding sight. gp120 contains two structural domains, an 'inner' domain that is presumed to contain the gp41 contacts and an 'outer' domain. These regions are joined by a small bridging domain. The CD4 binding site is large and is located at an interface of all three domains. The binding surface contains both conserved residues required for CD4 binding as well as variable residues to which CD4 binds the main chains. Also, there are cavities within the binding surface which are not contacted by CD4. Immune recognition

of this CD4 binding site is probably difficult because: a) there is a mix of conserved and variable regions in the binding site, b) the V1 and V2 regions likely mask the CD4 binding site prior to binding (121) and c) the CD4-bound state probably represents an otherwise energetically unfavorable conformation of gp120 (118, 120). The Fab binds to the opposite side of the bridging domain composed of part of the fourth conserved domain and the base of the V1 and V2 stem. This region is presumably the coreceptor binding site. Although the V3 loop is absent in the x-ray crystal structure, there is a large gap between the core gp120 and the 17b antibody light chains that would most likely be filled by the V3 loop (121). For several reasons it seems likely that this domain is shielded by the V3 loop before CD4 binding. First, the base of the V3 loop is in the proper location in the x-ray crystal structure of core gp120. Second, antibodies that recognize the V3 loop can inhibit coreceptor binding (113). Third, V3-recoginizing antibodies compete for gp120 binding with the neutralizing antibodies such as 17b that recognize the coreceptor binding site (122).

The Transmembrane Subunit, gp41

gp41, the transmembrane subunit of the HIV envelope protein, is predicted to contain a helical region just downstream of the fusion peptide (123-125). A similar heptad repeat region is seen in influenza hemagglutinin and Moloney murine leukemia virus, both of which form coiled-coil structures as determined by x-ray crystallography (24, 126). Protein dissection was performed on a recombinantly expressed gp41 ectodomain that lacked the N-terminal fusion peptide in order to obtain structural information on this subunit (127). After exposure to proteolysis, a stable, trimeric, helical sub-domain of gp41 remains. It is composed of two discontinuous peptides, N51 and C43. More generally, peptides from these two regions of gp41 are called N peptides and C peptides because they originate from the N-terminal and the C-terminal region of the gp41 ectodomain, respectively. Biophysical analysis led to the proposal that the N peptides form a central, trimeric coiled coil, with helical C peptides bound to the outside of the coiled coil. The same protein dissection and biophysical analyses were performed on the ectodomain from simian immunodeficiency virus (SIV), a very closely related retrovirus, and yielded the same results (128). Shortly after, the x-ray crystal structure of N and C peptides was determined, and verified the proposed model in both HIV (Figure 5) and SIV (129-132). The structure can be described as a trimer of helical hairpins. Three helical C peptides bind to the outside of the coiled coil core of N peptides in an

antiparallel fashion. Therefore, the intervening sequence of gp41 (not present in the structure) would be required to loop around from the base of the coiled coil to return the C peptide back to the same end of the molecule. In fact, a solution structure of the SIV gp41 shows the loop between the two helical regions (133).

As discussed above, the influenza HA2 low-pH-converted structure presumably represents the fusogenic conformation of HA. The HA2 ectodomain was prepared for crystallization by exposure to low pH, mimicking the fusion-activating conformational changes of HA seen on the surface of the virus upon reduction of pH (24). HIV, however, is not activated by a change in pH, so it is not as easy to verify the conformational state of observed envelope fragments. However, for several reasons the trimer of helical hairpins is believed to represent the fusogenic form of gp41 (129). First, in the absence of gp120, gp41 would be expected to fold to the fusogenic state. Indeed, conformational changes induced by receptor binding and required for fusion activation include the shedding of gp120 (101). In comparison, HA2 of influenza was shown to acquire its low-pH form when expressed in the absence of HA1 (42, 45). Second, the N and C peptide complex is extremely stable and is unlikely to dissociate after it is formed (127). In an analogy to influenza, the final fusogenic state of hemagglutinin is an extremely stable species (23), whereas the native, metastable state is easily altered by low pH, heat or chemical denaturants (46). Third, also as seen in the influenza fusogenic structure, the N and C termini of the envelope ectodomain are located at the same end of the folded molecule. This type of organization is present in the structure of the ectodomain of Moloney murine leukemia virus, another retrovirus, as well, which was proposed to be in its fusion-active state (126) and was earlier shown to posses similar structural features to influenza HA (134). Fourth, several of the mutations mentioned above, those that do not affect envelope protein processing or cell-surface expression yet ablate infectivity and membrane fusion are mapped to the core trimeric coiled coil (105-111). Finally, if the x-ray crystal structure represents the fusogenic conformation, it helps to explain the inhibitory activity of gp41-derived peptides that will be discussed below.

gp41-Derived Inhibitory Peptides

Several years before the structural studies on gp41, synthetic N and C peptides were shown to inhibit HIV infection (127, 135-137). The C peptides, which are effective at nanomolar concentrations, are much more potent than N peptides, which require

micromolar concentrations for effectiveness. It was proposed that the C peptides act by binding to or near to the predicted helical region downstream from the fusion peptide, which corresponds to the N peptide region (137, 138) and therefore inhibit infection in a dominant negative manner (127). Several experimental observations support this hypothesis. First, if you mix equimolar amounts of N and C peptides, the inhibitory activity of the C peptide is abolished (127). Second, mutant C peptides which are shown to destabilize the formation of the trimer of helical hairpins structure also demonstrate weakened anti-viral potency (139, 140). Third, the hydrophobic binding surface on the N peptide coiled coil core to which the C peptides bind is highly conserved between HIV-1 and SIV. Indeed, the comparable SIV C peptide can inhibit HIV-1 mediated membrane fusion (130). And, fourth, viruses with prolonged exposure to C peptides develop resistance by mutation of several residues in the N peptide region (141). Therefore, it seems likely that the C peptides do inhibit formation of the trimer of helical hairpins in a dominant negative manner (142) by binding to the N peptide region of gp41. This leaves a conundrum, however (69). How can the C peptides, in an intermolecular interaction which occurs at nanomolar concentrations, successfully compete against an intramolecular interaction (the C peptide region of gp41 folding up to interact with the N peptide region of gp41 to form the fusogenic structure) with a high effective concentration?

A Transient Fusion Intermediate

This quandary is partially solved by the proposal of a transient intermediate in the fusion process, an intermediate formed after receptor binding but before the formation of the helical hairpin (and therefore termed "pre-hairpin") (69, 132, 143). In the pre-hairpin intermediate the N peptide region is exposed, vulnerable to binding by synthetic C peptide. To finish solving the quandary, when the C peptide binds, it is proposed to irreversibly inactivate the protein for fusion. There are several observations that support the existence of this intermediate. First, C peptides must be present during or after exposure to the host cell in order to be effective. If the virus is preincubated with C peptide, followed by C peptide removal before addition of target membrane, infection is not inhibited (143). The kinetics of this inhibition are quite startling. Conformational changes due to receptor binding begin in under a minute (103), whereas the C peptides maintain potency even if added up to fifteen minutes after receptor exposure (144), implying that the transient intermediate experiences a lengthy lifetime. Second, epitope-

tagged C peptide can immunoprecipitate gp41, but only after exposure to CD4 and coreceptor, demonstrating that a receptor-mediated conformational change is required to expose the N peptide region (143). In addition, such an intermediate has been proposed for influenza HA (1, 70), although it has yet to be verified experimentally. In influenza it is likely that the conformational changes between the neutral-pH structure and the fusogenic structure occur sequentially. First, the "spring-loaded" mechanism occurs, flinging the fusion peptide up to the target membrane; this would be the pre-hairpin intermediate with a vulnerable "N peptide" region. Next, the bottom of the trimeric coiled coil flips up to fold the molecule in half, bringing the two membranes into close contact, just as the trimer of helical hairpins would do for HIV.

It was mentioned above that N peptides also exhibit anti-HIV characteristics. Since the most obvious interpretation of N peptide inhibition is through binding the C peptide region of gp41, it is conceivable that the C peptide region of gp41 is exposed in the pre-hairpin intermediate. However, it is possible that N peptides, which have a potential to oligomerize, interact with the N peptide region of gp41, and that the C peptide region is not exposed in the intermediate.

Little is known about HIV envelope-mediated fusion at the lipid level. It is generally assumed to be a comparable process to HA-mediated fusion. There is, however, experimental evidence that HIV fusion may proceed through the hemifusion state, like influenza. In some cases inhibition by C peptides allows lipid mixing, but not content mixing (144). Further experiments are needed to address this issue.

A Mechanism for HIV-1 Viral Entry

With all of the above experimental observations, the following mechanism is suggested for HIV-mediated viral membrane fusion (Figure 6), a mechanism very similar to that of influenza hemagglutinin-mediated membrane fusion. On the surface of the virus, the envelope protein complex exists in a pre-fusogenic, metastable state. Most of the surface area exposed consists of gp120, with much of gp41, including the fusion peptide domain, sequestered below. Binding of the gp120 to CD4 enacts conformational changes in gp120, allowing the attachment to the coreceptor followed by conformational changes both in gp120 and gp41. These conformational changes greatly weaken the interactions between gp120 and gp41. The transient, pre-hairpin intermediate of gp41 is

formed, freeing the previously hidden fusion peptide to interact with the host cell membrane and exposing the N peptide region (and possibly the C peptide region). It is unknown if there is a "spring-loaded" mechanism involved in forming the pre-hairpin intermediate as in influenza (42) or if the fusion peptide maintains the same position as in the pre-fusogenic state and is simply uncovered by the movement of gp120. The pre-hairpin maintains a durable half-life at which time it is susceptible to inhibitory peptides. However, in the absence of such peptides, the structure resolves to the highly stable trimer of helical hairpins seen in the x-ray crystal structures. This structure brings the N terminus and C terminus of the ectodomain into close proximity and most likely tilts to adapt a parallel orientation to the membrane (as first suggested in 128, also 69, 132), closely juxtaposing the two membranes. Membrane fusion ensues, probably in a manner similar to that of influenza virus.

Inhibiting HIV-1 Entry

Currently, only two HIV proteins are targeted by FDA approved drugs. These are reverse transcriptase, which is responsible for transcribing the HIV RNA genome to DNA, and protease, which autocatalyzes the digestion of the HIV Gag/Pol polyprotein and clips the subsequent Gag protein (145). Because of the high rate of viral turnover (146) and error-prone reverse transcriptase, resistant viruses often emerge. Combination drug therapy in which three or more drugs are administered combats the viral resistance. However, there are still significant problems associated with combination chemotherapy. Some patients develop immediate adverse effects and are therefore intolerant of available drugs. For those patients who are tolerant, they face an expensive, arduous treatment. Among these patients, some still experience viral emergence, and long-term adverse effects are developing (145, 147, 148). Also, because of increasing viral resistance, the threat of an outbreak of virus immune to all available drugs is rising. Therefore, any addition to the repertoire of anti-HIV drugs would be welcome, especially if the drug is less toxic and less susceptible to viral emergence.

Now that the process of HIV entry has been dissected extensively, it is feasible that soon drugs will be discovered to target this additional step of the viral life cycle. The recent identification of the coreceptor (review by 149), as well as a high-resolution structure of gp120 bound to CD4 and an antibody mimicking the coreceptor (118), open the door to discovering molecules that will inhibit the HIV virus from binding to the cell. Also, the inhibitory activity of the C peptides demonstrates the feasibility of targeting a

transient gp41 structure displayed during viral infection. Anti-HIV molecules which inhibit entry will stop the virus before it gets into the cell in the first place, unlike the currently used drugs which act after the virus has already infected the cell. Quite possibly identified drugs which could stop the virus from invading cells would be useful as prophylactic agents, creating a barrier to the initial infection event. Hopefully, the knowledge utilized for drug design for inhibiting HIV entry will provide useful leads for effective HIV preventive vaccines.

Soluble CD4

Early efforts to inhibit entry of HIV virus focused on inhibiting binding of the envelope protein to CD4 on the surface of cells. Soluble CD4 (sCD4) was shown to inhibit laboratory-adapted strains (116). However, when used on primary isolates, sCD4 was much less successful and actually increased the infectivity of some isolates (reviewed in 150). Several more recent experimental observations have explained why sCD4 is effective with laboratory-adapted strains and not primary isolates. First, sCD4 induces shedding of gp120 in laboratory adapted strains (101). Therefore in lab-adapted strains, sCD4 ablates the virus's potential to bind to and fuse with host cells by the complete removal of gp120 from the viral surface. Second, the same extent of shedding is not observed in primary isolates (153), and third, CD4 binding induces changes in gp120 that allow it to bind the coreceptor with higher affinity (117,151,152 and reviewed in 149). These last two observations explain why sDC4 is not effective against primary isolates. In the lack of shedding of gp120, sCD4 likely induces changes on the surface of the virus that encourage the virus to bind and fuse to cells expressing the proper coreceptor. So, in primary isolates, when gp120 shedding does not occur, sCD4 facilitates fusion.

Chemokine Receptors

The chemokine receptors provide an additional therapeutic target. Most HIV-1 variants use one or both of two specific chemokine receptors, CXCR4 and CCR5 (reviewed in 149 and 154). CXCR4 is expressed on T-cells and is therefore utilized by T-cell tropic HIV virus. CCR5 is expressed on macrophages and is consequently the coreceptor of choice for macrophage-tropic HIV. Both receptors are expressed on primary T-cells. These cells are vulnerable to infection by T-cell, macrophage and dual-

tropic viruses. Although those are the two most common coreceptors, at least eleven additional have been reported.

Each of the coreceptors have specific chemokine ligands, and these ligands are effective in blocking entry of HIV isolates which utilize that specific coreceptor. For example, the chemokine ligands of CCR5 (RANTES, MIP-1α, and MIP-1β) are effective inhibitors of macrophage-tropic HIV strains (155). Treatment with chemokines may cause unfavorable side effects, however, such as unwanted activation or interference with normal signaling pathways. Therefore, chemokine derivatives have been investigated, including synthetic peptide analogs. Some of these variants have uncoupled viral inhibition from the natural chemokine receptor function, and therefore serve as promising therapeutic leads (156-159). In addition, in a step toward potent, small molecule drugs, several low molecular weight compounds have been identified which bind to either CXCR4 or CCR5 and inhibit HIV infection (160-164)

As host proteins, the chemokine receptors are relatively static targets not prone to mutation, compared to the HIV virus "moving target" which rapidly mutates. However, despite the advantage to targeting the chemokine receptor directly, there are disadvantages to this approach. First, it is possible that blocking chemokine receptors will have negative side effects. Genetic data imply that this will not be the case when blocking CCR5 function (165). However, the absence of CXCR4 appears to be more detrimental, with CXCR4 knockout mice dying during embryogenesis (166, 167). Second, there is a concern that if specific coreceptors are blocked, there will be a strong selection for viruses with altering tropism. In a SCID mouse model sensitive to HIV infection, a rapid switch in coreceptor usage was seen after short treatment with an anti-CCR5 agent (168). CCR5, usually the coreceptor of choice early in an HIV infection, appears to be the safest target for avoiding toxicity. However, in an HIV infection, a switch to a T-tropic viral population usually coincides with onset of AIDS. Therefore, a CCR5-blocking agent that causes a selective pressure encouraging the emergence of T-tropic viruses could be extremely detrimental. The most useful therapy may be targeting all chemokine coreceptors. Only experimental trials will determine if this can be performed safely.

Targeting the Transient Fusion Intermediate

The transient intermediate of gp41 during the viral entry process is also a promising target for inhibiting entry. C peptides, which bind to this intermediate, have shown reasonable success in human clinical trials (169). The participants who received the largest dose of the C peptide, T20, experienced viral reduction levels similar to that in patients treated with one reverse transcriptase or protease inhibitor. The C peptides may be very useful in combination therapies or as viable short-term treatments, such as preventing the vertical transmission of the virus from mother to baby during the birth process. Also, there is the potential for utilizing C peptide derivatives that may be even more potent than wild type C peptides. For example, a C peptide containing a chemical linker which physically constrains the helical conformation lowers the entropy cost of binding gp41 and increases the potency of the C peptide (170). However, there are disadvantages to the therapeutic use of C peptides. They are large and are not amenable to oral routes of entry. They must be administered intravenously and at high quantity. This process is not only tedious, but also expensive. It would be much more preferable to identify a small molecule that could accomplish the function of the C peptide and be orally bioavailable.

Recent progress has been made towards identifying such molecules. In the gp41 x-ray crystal structure there is a small pocket in the conserved hydrophobic groove of the N peptide trimeric coiled coil (Figure 5). Three hydrophobic residues from the C peptide, two tryptophans and an isoleucine, bind into this pocket. The pocket was proposed to be a good target for drug discovery for many reasons (129, 140). First, the pocket is small (400 $Å^3$) and well defined, ideal for binding by a small molecule of 500-600 Daltons. Second, the residues comprising the pocket are highly conserved among all known HIV-1 isolates as well as among SIV. Third, mutations of several of the pocket-lining residues impair HIV infectivity (105, 106, 108-111). Fourth, the RNA encoding this region is an integral part of the structured Rev-response element (171, 172), implying that there is selective pressure at both a protein and an RNA level. Fifth, the hydrophobic residues of the C peptide that bind to the pocket are important for C peptide inhibitory activity (140). Finally, although there are effective C peptides that do not contain the pocket-binding residues (e.g. T20), they are more vulnerable to the emergence of resistant viruses than those containing the pocket-binding residues (e.g. T649) (141).

Several methods have been successful for identifying compounds that bind to this promising potential drug target. First, mirror-image phage display (173) was employed to identify small peptides composed of D-amino acids that bind to the hydrophobic pocket of gp41 (174). Since gp41 N peptides aggregate in the absence of C peptides, a hybrid coiled-coil molecule, IQN17, was used to properly present the gp41 pocket. In IQN17, a soluble, trimeric coiled coil, GCN4-pIQI (175) was fused to the pocket forming residues of gp41. Peptides were identified which bound to IQN17, but not to a control molecule with a point mutation that occludes the pocket. A 1.5 Å cocrystal structure showed that the D-peptides bind to the gp41 pocket-region of IQN17. These D-peptides are less than half the size of C peptides, but still slightly too large for oral administration. Their potency is in the micromolar range. Although the D-peptides themselves may not prove useful for therapy, their identification has verified the concept that targeting the gp41 coiled coil pocket, and only the pocket, is a viable therapeutic option. Also, the D-peptides in combination with the IQN17 target provide the tools needed for a competitive screen for identifying other potentially useful compounds. Second, combinatorial chemical libraries were used in place of the pocket-binding residues of the C peptide to identify tighter binders to the N peptide (176). The chemical libraries were attached to the pocket-binding site of a shortened, eighteen residue C peptide. This shortened C peptide, called a "biasing element" was meant to target the library to the correct binding site. The library was screened for compounds that bound to the N peptide better than the biasing element alone. One member of the library passed the screen, although it was ineffective at inhibiting HIV infection. However, placed on a 30 residue biasing element, the identified chemical ligand increased the peptide's potency an order of magnitude. Again, the potency is in the micromolar range. Although this particular ligand will not be useful for anti-HIV therapy, the experiment further emphasized the importance of targeting the conserved pocket. Finally, utilizing the x-ray crystal structure of the gp41 core and molecular docking techniques, a database of organic molecules was screened for potential fitting into the hydrophobic pocket (177). Of the sixteen compounds that were determined to have the best fit, two were shown to inhibit HIV infection at micromolar concentrations. The above experiments demonstrate that there are many different methods that can be utilized to attempt to target the gp41 pocket. Specifically, the D-peptide work verifies that molecules that bind to the pocket, and only the pocket, of gp41 are capable of inhibiting HIV-mediated membrane fusion. It's a good beginning, but much more work is necessary before there will be a small, potent, orally -bioavailable drug.

The Viral Fusion Proteins of Other Enveloped Viruses

During the past several years, there has been an onslaught of x-ray crystal structures of segments of enveloped viral fusion proteins (23, 126, 129-132, 178-182). The four viral families with the most extensive structural information are orthomyxovirus, retrovirus, paramyxovirus and filovirus. Most of the new structures depict a very similar theme described above for influenza and HIV and will be discussed later in this section. One other virus envelope structure has been solved at high resolution, TBE of the flavivirus family (181). Its structure is exceedingly different than the common structure previously described.

Tick-borne Encephalitis E Protein, A Structurally Distinct Viral Fusion Protein

Tick-borne encephalitis virus (TBE) belongs to the flavivirus family of enveloped viruses. It contains two membrane-anchored proteins, M and E. Low-pH-induced cleavage of the M protein is required for infectivity (183). The E protein has been proposed to be the fusion protein of TBE, although antibodies directed against either membrane protein are capable of inhibiting viral infectivity (184, 185). E is proposed to contain a fusion peptide domain. The domain is highly conserved within the flavivirus family and is hydrophobic, and monoclonal antibodies that recognize the potential fusion peptide inhibit membrane fusion (181, 184). TBE viral entry is initiated by low pH, which seems to switch the oligomeric state of E from dimer to trimer (186). The oligomeric change leads to an increased hydrophobicity of the E protein and exposure of the potential fusion domain (187-189).

The crystal structure of the neutral-pH, dimeric form of E was solved at 2.0 Å resolution, and it has a very different architecture than the HIV and influenza fusion proteins (181) (Figure 7). In order to produce a crystallizable fragment of the ectodomain, TBE virions were treated with trypsin to create a soluble ectodomain of E (sE). Overall it is a long, flat molecule (about 150 Å by 55 Å by 30 Å) composed of three structural domains and is proposed to lie parallel to the viral surface. Several structural details support this proposal. First, the molecule is curved and could therefore conform to the shape of the virus. Second, the outer face contains the glycosylation sites and third, the outer face also contains the epitopes for most neutralizing antibodies. The proposed fusion peptide domain (residues 98-120) forms a loop that is located in the

interface between the two monomers. The low-pH-activated conformational change would presumably alter the intermolecular interactions, exposing the fusion peptide sequence. So, even though the quaternary structure seems very different from the fusion proteins of other viruses, it seems likely that the low-pH-induced conformational changes in TBE serve the same purpose as the trimer of helical hairpin structure, to bring the two membranes together to initiate fusion.

Fusion Proteins of Retroviruses, Paramyxoviruses and Filoviruses

As mentioned above, there are many new high-resolution structures available for the fusion proteins of viruses in the retrovirus, paramyxovirus and filovirus families, and they all closely resemble the fusogenic structures of influenza HA and HIV gp41. In these structures there is a characteristic central trimeric coiled coil, presumably posed to present the fusion peptide at its tip. At the base of the coiled coil, the chain folds back, and supporting structures bind to the outside of the coiled coil, in most cases at least part of the outer structure contains a helix (Figure 8).

Heptad repeat regions were identified in many viral envelope proteins over a decade ago (123-125, 190). However, improved computational methods (191) based on known crystal structures (126, 129-132, 180, 182) have provided a broader, more accurate prediction for many viruses within the retrovirus, paramyxovirus and filovirus families . These computational methods have been verified by crystallographic analysis not included in the algorithm (178, 179). First, within the lentivirus genus of retroviruses, the envelope protein is predicted to contain two helical regions. This corresponds to the trimer of hairpin dimer structures seen in HIV (129, 131, 132), SIV (130, 133), and Visna virus (Malashkevich and Kim, unpublished observations). Second, the other genera of retroviruses, mammalian C-type, avian C-type, D-type and BLV-HTLV, are predicted to contain one coiled coil region and to be very similar to the filovirus family (126, 191, 192). Indeed, the x-ray crystal structures of the glycoproteins of Moloney murine leukemia virus (a mammalian C-type retrovirus), human T cell leukemia virus (a BLV-HTLV) and the filovirus, ebola, show marked similarity to each other (126, 179, 180, 182) (Figure 8). In comparison to the retroviruses, the fusion peptide of ebola is not predicted to be located at the very N terminus of the transmembrane domain, but rather 22 residues removed from the N terminus (192-194). The similarity in structure, however, implies that this internal fusion peptide can be

exposed to the target membrane in the same manner as an N-terminal fusion peptide. Finally, the paramyxovirus family of retroviruses is strongly predicted to contain both helical N peptides and C peptides across all genera (191). Although these proteins contain many more residues between the predicted helical regions, X-ray crystallographic studies of simian parainfluenza virus 5 (SV5) (178) and human respiratory synctyial virus (Zhao, Malashkevich and Kim, unpublished observations) verify formation of the trimer of helical hairpins (Figure 8). Interestingly, the C peptide regions of many of the paramyxoviruses, such as Sendai virus, measles virus, Newcastle disease virus (NDV), human parainfluenza virus (HPIV), respiratory syncytia virus (RSV) and simian parainfluenzavirus 5 (SV5), inhibit virus infectivity (195-201), implying that they all mediate fusion with a trimer of hairpin dimer conformation. Hopefully the C peptides of these viruses, as well as of others with similar structures, will prove to be of therapeutic use.

Because of the strong structural similarity between viruses of different families, it seems likely that they all share a common mechanism for viral membrane fusion as proposed above for influenza and HIV. In addition, viral inhibition strategies employed for one virus are likely to be relevant for many other viruses. This has already been demonstrated with the envelope-derived inhibitory peptides of paramyxoviruses. Hopefully the elucidation of more structures of the envelope proteins of health-threatening enveloped viruses will lead to small molecule identification as in the case of HIV.

Conclusion

In conclusion, extensive biochemical and structural studies on viruses from different families have provided a general mechanism for viral envelope glycoprotein-mediated membrane fusion. Viruses synthesize their fusion proteins in an inactive form, and subsequent cleavage of the glycoprotein activates the fusion potential of the virus. The protein waits in a metastable state for the appropriate activation signal, be it induction of low pH or receptor binding. After the signal arrives, the glycoprotein unleashes its fusion potential. Through a spring-loaded mechanism, at least in the case of influenza HA, the fusion peptide domain is propelled out of the interior of the protein and inserted into the target membrane. Then, the protein adopts its most stable fold. Most

likely the energy harnessed through acquisition of the stable state encourages the energetically unfavorable fusion of two membranes.

Recent evidence suggests that the helical hairpin may be a global motif for promoting membrane fusion events. Specifically, SNARES, the proteins required for vesicle fusion, adopt a similar conformation. In synaptic vesicle fusion, the vesicle contributes one SNARE (synaptobrevin), and the plasma membrane two (syntaxin and SNAP-25). Between the three of them, they form a highly stable bundle of four helices, and it has been proposed that this bundle closely juxtaposes the two membranes, leading to membrane fusion (202, 203). Therefore, the mechanisms of fusing two distinct biological membranes in these two largely different biological processes (viral membrane fusion and vesicle trafficking) seem extremely similar. Perhaps future scientific efforts will uncover similarities in other membrane fusion events, such as sperm and egg fusion in fertilization.

References

1. Hernandez, L. D., L. R. Hoffman, T. G. Wolfsberg, J. M. White. 1996. Virus-cell and cell-cell fusion. *Annu. Rev. Cell Dev. Biol.* 1996: 627-61.

2. Pecheur, E. I., J. Sainte-Marie, A. Bienvenue, D. Hoekstra. 1999. Peptides and membrane fusion: towards an understanding of the molecular mechanism of protein-induced fusion. *J. Mem. Biol.* 167: 1-17.

3. Durell, S. R., I. Martin, J.-M. Ruysschaert, Y. Shai, R. Blumenthal. 1997. What studies of fusion peptides tell us about viral envelope glycoprotein-mediated membraen fusion (review). *Mol. Mem. Biol.* 14: 97-112.

4. Bosch, M. L., P. L. Earl, K. Fargnoli, S. Picciafuoco, F. Giombini, F. Wong-Staal, G. Franchini. 1989. Identification of the fusion peptide of primate immunodeficiency viruses. *Science.* 244: 694-697.

5. Freed, E. O., D. J. Myers, R. Risser. 1990. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc. Natl. Acad. Sci. USA.* 87: 4650-4654.

6. Gething, M. J., R. W. Doms, D. York, J. White. 1986. Studies on the mechanism of membrane fusion: site-specific mutagenesis of the hemagglutinin of influenza virus. *J. Cell Biol.* 102: 11-23.

7. Horth, M., B. Lambrecht, M. C. Khim, F. Bex, C. Thiriart, J. M. Ruysschaert, A. Burny, R. Brasseur. 1991. Theoretical and functional analysis of the SIV fusion peptide. *EMBO J.* 10: 2747-2755.

8. Steinhauer, D. A., S. A. Wharton, J. J. Skehel, D. C. Wiley. 1995. Studies of the membrane fusion activities of fusion peptide mutants of influenza hemagglutinin. *J. Virol.* 69: 6643-6651.

9. Steffy, K. R., G. Kraus, D. J. Looney, F. Wong-Staal. 1992. Role of the fusogenic peptide sequence in syncytium induction and infectivity of human immunodeficiency virus type 2. *J. Virol.* 66: 4532-4535.

10. Stegmann, T., J. M. Delfino, F. M. Richards, A. Helenius. 1991. The HA2 subunit of influenza hemagglutinin inserts into the target membrane prior to fusion. *J. Biol. Chem.* 266: 18404-18410.

11. Durrer, P. l., C. Galli, S. Hoenke, C. Corti, R. Gluck, T. Vorherr, J. Brunner. 1996. H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. *J. Biol. Chem.* 271: 13417-13421.

12. Durrer, R., Y. Gaudin, R. W. H. Ruigrok, R. Graf, J. Brunner. 1995. Photolabeling identifies a putative fusion domain in the envelope glycoprotein of rabies and vesicular stomatitis viruses. *J. Biol. Chem.* 270: 17575-17581.

13. Gaudin, Y., R. W. H. Ruigrok, J. Brunner. 1995. Low-pH induced conformational changes in viral fusion proteins: implications for the fusion mechanism. *J. Gen. Virol.* 76: 1541-1556.

14. Clague, M. J., J. R. Knutson, R. Blumenthal, A. Herrmann. 1991. Interation of influenza hemagglutinin amino-terminal peptide with phospholipid vesicles: a fluorescence study. *Biochemistry.* 30: 5491-5497.

15. Lear, J. D., W. F. DeGrado. 1987. Membrane binding and conformational properties of peptides representing the NH_2 terminus of influenza HA_2. *J. Biol. Chem.* 262: 6500-6505.

16. Rafalski, M., A. Ortiz, A. Rockwell, L. C. Van Ginkel, J. D. Lear, W. F. DeGrado, J. Wilschut. 1991. Membrane fusion activity of the influenza virus hemagglutinin: interaction of HA_2 N-terminal peptides with phospholipid vesicles. *Biochemistry.* 30: 10211-10220.

17. Wharton, S. A., S. R. Martin, R. W. H. Ruigrok, J. J. Skehel, D. C. Wiley. 1988. Membrane fusion by peptide analogues of influenza virus haemagglutinin. *J. Gen. Virol.* 69: 1847-1857.

18. Harter, C., P. James, T. Bachi, G. Semenza, J. Brunner. 1989. Hydrophobic binding of the ectodomain of influenza hemagglutinin to membranes occurs through the "fusion peptide.". *J. Biol. Chem.* 264: 6459-64.

19. Hunter, E., E. Hill, M. Hardwick, A. Bhown, D. E. Schwarz, R. Tizard. 1983. Complete sequence of the Rous sarcoma virus env gene: identification of structural and functional regions of its product. *J. Virol.* 46: 920-936.

20. Fredericksen, B. L., M. A. Whitt. 1995. Vesicular stomatitis virus glycoprotein mutations that affect membrane fusion activity and abolish virus infectivity. *J. Virol.* 69: 1435-1443.

21. Zhang, L., H. P. Ghosh. 1994. Characterization of the putative fusogenic domain in vesicular stomatitis virus glycoprotein G. *J. Virol.* 68: 2186-2193.

22. Chen, J., K. H. Lee, D. A. Steinhauer, D. J. Stevens, J. J. Skehel, D. C. Wiley. 1998. Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. *Cell.* 95: 409-417.

23. Chen, J., J. J. Skehel, D. C. Wiley. 1999. N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA subunit to form an N cap that terminates the triple-stranded coiled coil. *Proc. Natl. Acad. Sci. USA*. 96: 8967-8972.

24. Bullough, P. A., F. M. Hughson, J. J. Skehel, D. C. Wiley. 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*. 371: 37-43.

25. Wilson, I. A., J. J. Skehel, D. C. Wiley. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*. 289: 366-373.

26. Chernomordik, L. V., E. Leikina, M. M. Kozlov, V. A. Frolov, J. Zimmerberg. 1999. Structural intermediates in influenza haemagglutinin-mediated membrane fusion. *Mol. Mem. Biol.* 16: 33-42.

27. Ramalho-Santos, J., M. C. Pedroso de Lima. 1998. The influenza virus hemagglutinin: a model protein in the study of membrane fusion. *Biochim. Biophys. Acta*. 1376: 147-154.

28. Wiley, D. C., J. J. Skehel. 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Ann. Rev. Biochem.* 56: 365-94.

29. Maeda, T., S. Ohnishi. 1980. Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes. *FEBS Lett.* 122: 283-287.

30. Maeda, T., K. Kawasaki, S. Ohnishi. 1981. Interaction of influenza virus hemagglutinin with target membrane lipids is a key step in virus-induced hemolysis and fusion at pH 5.2. *Proc. Natl. Acad. Sci. USA*. 78: 4133-4137.

31. Simons, K., H. Garoff, A. Helenius. 1982. How an animal virus gets into and out of its host cell. *Sci Am.* 246: 58-66.

32. Doms, R. W., A. Helenius, J. White. 1985. Membrane fusion activity of the influenza virus hemagglutinin. The low pH-induced conformational change. *J. Biol. Chem.* 260: 2973-2981.

33. Skehel, J. J., P. M. Bayley, E. B. Brown, S. R. Martin, M. D. Waterfield, J. M. White, I. A. Wilson, D. C. Wiley. 1982. Changes in the conformation of influenza virus at the pH optimum of virus-mediated membrane fusion. *Proc. Natl Acad. Sci. USA*. 79: 968-972.

34. Ruigrok, R. W., N. G. Wrigley, L. J. Calder, S. Cusack, S. A. Wharton, E. B. Brown, J. J. Skehel. 1986. Electron microscopy of the low pH structure of influenza virus haemagglutinin. *EMBO J.* 5: 41-49.

35. Copeland, C. S., R. W. Doms, E. M. Bolzau, R. G. Webster, A. Helenius. 1986. Assembly of influenza hemagglutinin trimers and its role in intracellular transport. *J. Cell Biol.* 103: 1179-1191.

36. Daniels, R. S., A. R. Douglas, J. J. Skehel, D. C. Wiley. 1983. Analysis of the antigenicity of influenza haemagglutinin at the pH optimum for virus-mediated membrane fusion. *J. Gen. Virol.* 64: 1657-1662.

37. Yewdell, J. W., W. Gerhard, T. Bachi. 1983. Monoclonal anti-hemagglutinin antibodies detect irreversible antigenic alterations that coincide with the acid activation of influenza virus A/PR/834-mediated hemolysis. *J. Virol.* 48: 239-248.

38. Webster, R. G., L. E. Brown, D. C. Jackson. 1983. Changes in the antigenicity of the hemagglutinin molecule of H3 influenza virus at acidic pH. *Virology.* 126: 587-99

39. Takahashi, S. 1990. Conformation of membrane fusion-active 20-residue peptides with or without lipid bilayers. Implication of alpha-helix formation for membrane fusion. *Biochem.* 29: 6257-6264.

40. Weis, W. I., S. C. Cusack, J. H. Brown, R. S. Daniels, J. J. Skehel, D. C. Wiley. 1990. The structure of a membrane fusion mutant of the influenza virus haemagglutinin. *EMBO J.* 9: 17-24.

41. Daniels, R. S., J. C. Downie, A. J. Hay, M. Knossow, J. J. Skehel, M. L. Wang, D. C. Wiley. 1985. Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell.* 40: 431-439.

42. Carr, C. M., P. S. Kim. 1993. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell.* 73: 823-832.

43. Qiao, H., S. L. Pelletier, L. Hoffman, J. Hackert, R. T. Armstrong, J. M. White. 1998. Specific single or double proline substitutions in the "spring-loaded" coiled-coil region of the influenza hemagglutinin impair or abolish membrane fusion activity. *J. Cell Biol.* 141: 1335-1347.

44. Ruigrok, R. W., A. Aitken, L. J. Calder, S. R. Martin, J. J. Skehel, S. A. Wharton, W. Weis, D. C. Wiley. 1988. Studies on the structure of hte influenza virus hemagglutinin at the pH of membrane fusion. *J. Gen. Virol.* 69: 2785-2795.

45. Chen, J., S. A. Wharton, W. Weissenhorn, L. J. Calder, F. M. Hughson, J. J. Skehel, D. C. Wiley. 1995. A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in Escherichia coli into the low-pH-induced conformation. *Proc. Natl. Acad. Sci. USA.* 19: 12205-12209.

46. Carr, C. M., C. Chaudhry, P. S. Kim. 1997. Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc. Natl. Acad. Sci. USA.* 94: 14306-14313.

47. Stegmann, T., F. P. Booy, J. Wilschut. 1987. Effects of low pH on influenza virus. Activation and inactivation of the membrane fuson capacity of hte hemagglutinin. *J. Biol. Chem.* 262: 17744-17749.

48. White, J., J. Kartenbeck, A. Helenius. 1982. Membrane fusion activity of influenza virus. *EMBO J.* 1: 217-222.

49. Wharton, S. A., L. J. Calder, R. W. H. Ruigrok, J. J. Skehel, D. A. Steinhauer, D. C. Wiley. 1995. Electron microscopy of antibody complexes of influenza virus haemagglutinin in the fusion pH conformation. *EMBO J.* 14: 240-246.

50. Spruce, A. E., A. Iwata, J. M. White, W. Almers. 1989. Patch clamp studies of single cell-fusion events mediated by a viral fusion protein. *Nature*. 342: 555-558.

51. Spruce, A. E., A. Iwata, W. Almers. 1991. The first milliseconds of the pore formed by a fusogenic viral envelope protein during membrane fusion. *Proc. Natl. Acad. Sci. USA*. 88: 3623-3627.

52. Sarkar, D. P., S. J. Morris, O. Eidelman, J. Zimmerberg, R. Blumenthal. 1989. Initial stages of influenza hemagglutinin-induced cell fusion monitored simultaneously by two fluorescent events: cytoplasmic continuity and lipid mixing. *J. Cell Biol.* 109: 113-122.

53. Melikyan, G. B., W. D. Niles, M. E. Peeples, F. S. Cohen. 1993. Influenza hemagglutinin-mediated fusion pores connecting cells to planar membranes: flickering to final expansion. *J. Gen. Physiol.* 102: 1131-1149.

54. Tse, F. W., A. Iwata, W. Almers. 1993. Membrane flux though the pore formed by a fusogenic viral envelope protein during cell fusion. *J. Cell Biol.* 121: 543-552.

55. Zimmerberg, J., R. Blumenthal, D. P. Sarkar, M. Curran, S. J. Morris. 1994. Restricted movement of lipid and aqueous dyes through pores formed by influenza hemagglutinin during cell fusion. *J. Cell Biol.* 127: 1885-1894.

56. Razinkov, V. I., G. B. Melikyan, F. S. Cohen. 1999. Hemifusion between cells expressing hemagglutinin of influenza virus and planar membranes can precede the formation of fusion pores that subsequently fully enlarge. *Biophys. J.* 77: 3144-3151.

57. Simon, A. M., D. A. Goodenough. 1998. Diverse functions of vertebrate gap junctions. *Trends Cell Biol.* 8: 477-483.

58. Danieli, T., S. L. Pelletier, Y. I. Henis, J. M. White. 1996. Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. *J. Cell Biol.* 133: 559-569.

59. Kanaseki, T., K. Kawasaki, M. Murata, Y. Ikeuchi, S.-I. Ohnishi. 1997. Structural features of membrane fusion between influenza virus and liposome as revealed by quick-freezing electron microscopy. *J. Cell Biol.* 137: 1041-1052.

60. Chernomordik, L. V., E. Leikina, V. Frolov, P. Bronk, J. Zimmerberg. 1997. An early stage of membrane fusion mediated by the low pH conformation of influenza hemagglutinin depends upon membrane lipids. *J. Cell Biol.* 136: 81-94.

61. Chernomordik, L. V., V. A. Frolov, E. Leikina, P. Bronk, J. Zimmerberg. 1998. The pathway of membrane fusion catalyzed by influenza hemagglutinin: restriction of lipids, hemifusion, and lipidic fusion pore formation. *J. Cell Biol.* 140: 1369-1382.

62. Markosyan, R. M., G. B. Melikyan, F. S. Cohen. 1999. Tension of membranes expressing the hemagglutinin of influenza virus inhibits fusion. *Biophys. J.* 77: 943-952.

63. Kemble, G. W., T. Danieli, J. M. White. 1994. Lipid-Anchored Influenza Hemagglutinin Promotes Hemifusion, Not Complete Fusion. *Cell.* 76: 383-391.

64. Melikyan, G. B., J. M. White, F. S. Cohen. 1995. GPI-anchored influenza hemagglutinin induces hemifusion to both red blood cell and planar bilayer membranes. *J. Cell Biol.* 131: 679-691.

65. Melikyan, G. B., S. A. Brener, D. C. Ok, F. S. Cohen. 1997. Inner but not outer membrane leaflets control the transition from glycosylphosphatidylinositol-anchored influenza hemagglutinin-induced hemifusion to full fusion. *J. Cell Biol.* 136995-1005.

66. Niles, W. D., F. S. Cohen. 1991. Fusion of influenza virions with a planar lipid membrane detected by video fluorescence microscopy. *J. Gen. Physiol.* 97: 1101-1119.

67. Lowy, R. J., D. P. Sarkar, Y. Chen, R. Blumenthal. 1990. Observation of single inluenza virus-cell fusion and measurement by fluorescence video microscopy. *Proc. Natl. Acad. Sci. USA.* 87: 1850-1854.

68. Georgiou, G. N., I. E. Morrison, R. J. Cherry. 1989. Digital fluorescence imaging of fusoin of influenza virus with erythrocytes. *FEBS Lett.* 250: 487-492.

69. Chan, D. C., P. S. Kim. 1998. HIV entry and its inhibition. *Cell.* 93: 681-684.

70. Hughson, F. M. 1997. Enveloped viruses: A common mode of membrane fusion? *Curr. Biol.* 7: R565-R569.

71. Kozlov, M. M., L. Chernomordik. 1998. A mechanism of protein-mediated fusion: coupling between refolding of the influenza hemagglutinin and lipid rearrangements. *Biophys. J.* 75: 1384-1396.

72. Weissenhorn, W., A. Dessen, L. J. Calder, S. C. Harrison, J. J. Skehel, D. C. Wiley. 1999. Structural basis for membrane fusion by enveloped viruses. *Mol. Membr. Biology*. 16: 3-9.

73. Wyatt, R., J. Sodroski. 1998. The HIV-1 envelope glycoproteins: fusogens, antigens and immunogens. *Science*. 280: 1884-1888.

74. Hallenberger, S., V. A. Bosch, H., E. Shaw, H. D. Klenk, W. Garten. 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature*. 360: 358-361.

75. Allan, J. S., J. E. Coligan, F. Barin, M. F. McLane, J. G. Sodroski, C. A. Rosen, W. A. Haseltine, T. H. Lee, M. Essex. 1985. Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science*. 228: 1091-1094.

76. Robey, W. G., B. Safai, S. Oroszlan, L. O. Arthur, M. A. Gonda, R. C. Gallo, P. J. Fischinger. 1985. Characterizatoin of envelope and core structural gene products of HTLV-III with sera from AIDS patients. *Science*. 228: 593-595.

77. Veronese, F. D., A. L. DeVico, T. D. Copeland, S. Oroszlan, R. C. Gallo, M. G. Sarngadharan. 1985. Characterizatio of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene. *Science*. 229: 1402-1405.

78. Kowalski, M., J. Potz, L. Basiribour, T. Dorfman, W. C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. Haseltine, J. Sodroski. 1987. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science*. 237: 1351-1355.

79. Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, E. A. Berger. 1996. CC CKR5: a RANTES, MIT-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*. 272: 1955-1958.

80. Doranz, B. J., J. Rucker, Y. Yi, R. J. Smyth, M. Samson, S. C. Peiper, M. Parmentier, R. G. Collman, R. W. Doms. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b a fusion cofactors. *Cell*. 85: 1149-1158.

81. Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Doup, J. P. Moore, W. A. Paxton. 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*. 381: 667-673.

82. Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 381: 661-666.

83. Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, J. Sodroski. 1996. The beta-chemokine receptors CCR3 adn CCR5 facilitate infection by primary HIV-1 isolates. *Cell.* 85: 1135-1148.

84. Feng, Y., C. C. Broder, P. E. Kennedy, E. A. Berger. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science.* 272: 872-877.

85. McDougal, J. S., M. S. Kennedy, J. M. Sligh, S. P. Cort, A. Mawle, J. K. Nicholson. 1986. Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science.* 231: 382-385.

86. Dalgleish, A. G., P. C. Beverley, P. R. Chapham, D. H. Crawford, M. F. Greaves, R. A. Weiss. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature.* 312: 763-767.

87. Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J. C. Gluckman, L. Montagnier. 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature.* 312: 767-768.

88. Delwart, E. L., G. L. Buchschacher Jr., E. O. Freed, A. T. Panganiban. 1992. Analysis of HIV-1 envelope mutants and pseudotyping of replication-defective HIV-1 vectors by genetic complementation. *AIDS Res. Hum. Retro.* 8: 1669-1677.

89. Schaal, H., M. Klein, P. Gehrmann, O. Adams, A. Scheid. 1995. Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. *J. Virol.* 69: 3308-3314.

90. Mobley, P. W., H. F. Lee, C. C. Curtain, A. Kirkpatrick, A. J. Waring, L. M. Gordon. 1995. The amino-terminal peptide of HIV-1 glycoprotein 41 fuses human erythrocytes. *Biochim. Biophys. Acta.* 1271: 304-314.

91. Rafalski, M., J. D. Lear, W. F. DeGrado. 1990. Phospholipid interactoins of synthetic peptides representing the N-terminus of HIV gp41. *Biochemistry.* 29: 7917-7922.

92. Slepushkin, V. A., S. M. Andreev, M. V. Sidorova, G. B. Melikyan, V. B. Grigoriev, V. M. Chumakov, A. E. Grinfeldt, R. A. Manukyan, E. V. Karamov. 1992. Investigation of human immunodeficiency virus fusion peptides. Analysis of interrelations between their structure and function. *AIDS Res. Hum. Retroviruses.* 8: 9-18.

93. Martin, I., F. Defrise-Quertain, E. Decroly, M. Vandenbranden, R. Brasseur, J. M. Ruysschaert. 1993. Orientation and structure of the NH2-terminal HIV-1 gp41 peptide in fused and aggregated liposomes. *Biochim Biophys Acta.* 1145: 124-133.

94. Nieva, J. L., S. Nir, A. Muga, F. M. Goni, J. Wilschut. 1994. Interaction of the HIV-1 fusion peptide with phospholipid vesicles: different structural requirements for fusion and leakage. *Biochemistry*. 33: 3201-3209.

95. Kliger, Y., A. Aharoni, D. Rapaport, P. Jones, R. Blumenthal, Y. Shai. 1997. Fusion peptides derived from the HIV type 1 glycoprotein 41 associate within phospholipid membranes and inhibit cell-cell fusion. *J. Biol. Chem.* 272: 13496-13505.

96. Sattentau, Q. J., J. P. Moore. 1991. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J. Exp. Med.* 174: 407-415.

97. Sattentau, Q. J., J. P. Moore, F. Vignaux, F. Traincard, P. Poignard. 1993. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *J. Virol.* 67: 7383-7393.

98. Kirsh, R., T. K. Hart, H. Ellens, J. Miller, S. A. Petteway Jr., D. M. Lambert, J. Leary, P. J. Bugelski. 1990. Morphometic analysis of recombinant soluble CD4-mediated release of the envelope glycoprotein gp120 from HIV-1. *AIDS Res. Hum. Retroviruses*. 6: 1209-1212.

99. Hart, T. K., R. Kirsh, H. Ellens, R. W. Sweet, D. M. Lambert, S. R. Petteway Jr., J. Leary, P. J. Bugelski. 1991. Binding of soluble CD4 proteins to human immunodeficiency virus type 1 and infected cells induces release of envelope glycoprotein gp120. *Proc. Natl. Acad. Sci. USA*. 88: 2189-2193.

100. McKeating, J. A., A. McKnight, J. P. Moore. 1991. Differential loss of envelope glycoproteing gp120 from virions of human immunodeficiency virus type 1 isolates: effects on infectivity and neutralization. *J. Virol.* 65: 852-860.

101. Moore, J. P., J. A. McKeating, R. A. Weiss, Q. J. Sattentau. 1990. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. *Science*. 250: 1139-1142.

102. Clements, G. J., M. J. Price-Jones, P. E. Stephens, C. Sutton, T. F. Schulz, P. R. Clapham, J. A. McKeating, M. O. McClure, S. Thomson, M. Marsh, e. al. 1991. The V3 loops of the HIV-1 and HIV-2 surface glycoproteins contain proteolytic cleavage sites: a possible function in viral fusion? *AIDS Res. Hum. Retroviruses*. 7: 3-16.

103. Jones, P. L., T. Korte, R. Blumenthal. 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.

104. Korte, T., A. Herrmann. 1994. pH-dependent binding of the fluorophore bis-ANS to influenza virus reflects the conformational change of hemagglutinin. *Eur. Biophys. J.* 23: 105-113.

105. Cao, J., L. Bergeron, E. Helseth, M. Thali, H. Repke, J. Sodroski. 1993. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. *J. Virol.* 67: 2747-2755.

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

107. Chen, S. S. 1994. Functional role of the zipper motif region of human immunodeficiency virus type 1 envelope transmembrane glycoprotein. *J. Virol.* 68: 2002-2010.

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

109. Poumbourios, P., K. A. Wilson, R. J. Center, W. El Ahmar, B. E. Kemp. 1997. Human immunodeficiency virus type 1 envelope glycoprotein oligomerization requires the gp41 amphipathic alpha-helical/leucine zipper-like sequence. *J. Virol.* 71: 2041-2049.

110. Wild, C. T., D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews. 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.

111. Weng, Y., C. D. Weiss. 1998. Mutational analysis of residues in the coiled-coil domain of human immunodeficiancy virus type 1 transmembrane protein gp41. *J. Virol.* 72: 9676-9682.

112. Thali, M., J. P. Moore, C. Furman, M. Charles, D. D. Ho, J. Robinson, J. Sodroski. 1993. Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* 67: 3978-3988.

113. Trkola, A., T. Dragic, J. Arthos, J. M. Binley, W. C. Olson, G. P. Allaway, C. Cheng-Mayer, J. Robinson, P. J. Maddon, J. P. Moore. 1996. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature.* 384: 184-187.

114. Wu, L., N. P. Gerard, R. Wyatt, H. Choe, C. Parolin, N. Ruffing, A. Borsetti, A. A. Cardoso, E. Desjardin, W. Newman, C. Gerard, J. Sodroski. 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature.* 384: 179-183.

115. Doranz, B. J., R. W. Doms. 1999. Use of a gp120 binding assay to dissect the requirements and kinetics of human immunodeficiency virus fusion events. *J. Virol.* 73: 10346-10358.

116. Deen, K. C., J. S. McDougal, R. Inacker, G. Folena-Wasserman, J. Arthos, J. Rosenberg, P. J. Maddon, R. Axel, R. W. Sweet. 1988. A soluble form of CD4 (T4) protein inhibits AIDS virus infection. *Nature*. 331: 82-84.

117. Salzwedel, K., E. D. Smith, B. Dey, E. A. Berger. 2000. Sequential CD4-Coreceptor Interactions in Human Immunodeficiency Virus Type 1 Env Function: Soluble CD4 Activates Env for Coreceptor-Dependent Fusion and Reveals Blocking Activities of Antibodies against Cryptic Conserved Epitopes on gp120. *J. Virol.* 74: 326-333.

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

119. Turner, B. G., M. F. Summers. 1999. Structural Biology of HIV. *J. Mol. Biol.* 285: 1-32.

120. Sattentau, Q. J. 1998. HIV gp120: double lock strategy foils host defenses. *Structure*. 6: 945-949.

121. Wyatt, R., P. D. Kwong, E. Desjardins, R. W. Sweet, J. Robinson, W. A. Hendrickson, J. G. Sodroski. 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature*. 393: 705-711.

122. Moore, J. P., J. Sodroski. 1996. Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelope glycoprotien. *J. Virol.* 70: 1863-1872.

123. Gallaher, W. R., J. M. Ball, R. F. Garry, M. C. Griffin, R. C. Montelaro. 1989. A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS Res. Hum. Retroviruses*. 5: 431-440.

124. Delwart, E. L., G. Moealos, T. Gilmore. 1990. Retroviral envelope glycoproteins contain a leucine zipper-like repeat. *AIDS Res. Hum. Retroviruses*. 6: 703-706.

125. Chambers, P., C. R. Pringle, A. J. Easton. 1990. Heptad repeat regions are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *J. Gen. Virol.* 71: 3075-3080.

126. Fass, D., S. C. Harrison, P. S. Kim. 1996. Retrovirus envelope domain at 1.7Å resolution. *Nature Struct. Biol.* 3: 465-469.

127. Lu, M., S. C. Blacklow, P. S. Kim. 1995. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature Struct. Biol.* 2: 1075-1082.

128. Blacklow, S. C., M. Lu, P. S. Kim. 1995. A trimeric subdomain of the simian immunodeficiency virus envelope glycoprotein. *Biochemistry*. 34: 14955-14962.

129. Chan, D. C., D. Fass, J. M. Berger, P. S. Kim. 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*. 89: 263-273.

130. Malashkevich, V. N., D. C. Chan, C. T. Chutkowski, P. S. Kim. 1998. Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: Conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc. Natl Acad. Sci. USA*. 95: 9134-9139.

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

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

133. Caffrey, M., M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, D. G. Covell, A. M. Gronenborn, G. M. Clore. 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J.* 17: 4572-4584.

134. Fass, D., P. S. Kim. 1995. Dissection of a Retrovirus Envelope Protein Reveals Structural Similarity to Influenza Hemagglutinin. *Curr. Biol.* 5: 1377-1383.

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

136. Wild, C. T., T. Greenwell, T. Matthews. 1993. A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS Res. Hum. Retroviruses*. 9: 1051-1053.

137. Wild, C. T., D. C. Shugars, T. K. Greenwell, C. B. McDanal, T. J. Matthews. 1994. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc. Natl. Acad. Sci. USA*. 91: 9770-9774.

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

139. Wild, C. T., T. Greenwell, D. Shugars, L. Rimsky-Clarke, T. Matthews. 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.

140. Chan, D. C., C. T. Chutkowski, P. S. Kim. 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.

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

142. Herskowitz, I. 1987. Functional inactivation of genes by dominant negative mutations. *Nature*. 329: 219-222.

143. Furuta, R. A., C. T. Wild, Y. Weng, C. D. Weiss. 1998. Capture of an early fusion-active conformation of HIV-1 gp41. *Nature Struct. Biol.* 5: 276-279.

144. Munoz-Barroso, I., S. Durell, K. Sakaguchi, E. Appella, R. Blumenthal. 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.

145. Carpenter, C. C. J., M. A. Fischl, S. M. Hammer, M. S. Hirsch, D. M. Jacobsen, D. A. Katzenstein, J. S. G. Montaner, D. D. Richman, M. S. Saag, R. T. Schooley, M. A. Thompson, S. Vella, P. G. Yeni, P. A. Volberding. 1998. Antiretroviral therapy for HIV infection in 1998: updated recommendations of the International AIDS Society-USA panel. *J. Am. Med. Assoc.* 280: 78-86.

146. Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*. 373: 123-126.

147. McDonald, C. K., D. R. Kuritzkes. 1997. Human immunodeficiency virus type 1 protease inhibitors. *Arch. Intern. Med.* 157: 951-959.

148. Saag, M. S. 1992. Treatment of HIV infection: the antiretroviral nucleoside analogues. Nucleoside analogues: adverse effects. *Hosp. Pract. (Off. Ed.)*. 27 Suppl 2: 26-36.

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

150. Daar, E. S., D. D. Ho. 1991. Relative resistance of primary HIV-1 isloates to neutralization by soluble CD4. *Am J. Med.* 90: 22S-26S.

151. Sullivan, N., Y. Sun, J. Binley, J. Lee, C. F. Barbas III, P. W. Parren, D. R. Burton, J. Sodroski. 1998. Determinants of human immunodeficiency viurs type 1

envelope glycoprotein activation by soluble CD4 adn monoclonal antibodies. *J. Virol.* 72: 6332-6338.

152. Sullivan, N., Y. Sun, Q. Sattentau, M. Thali, D. Wu, G. Denisova, J. Gershoni, J. Robinson, J. Moore, J. Sodroski. 1998. CD4-induced conformational changes in the human immundeficiency virus type 1 gp120 glycoprotein: consequences for virus entry and neutralization. *J. Virol.* 72: 4694-4703.

153. Moore, J. P., J. A. McKeating, Y. X. Huang, A. Ashkenazi, D. D. Ho. 1992. Virions of primary human immunodeficiency virus type 1 isolates resistant to soluble CD4 (sCD4) neutralization differ in sCD4 binding and glycoprotein gp120 retention from sCD4-sensitive isolates. *J. Virol.* 66: 235-243.

154. Littman, D. R. 1998. Chemokine receptors: keys to AIDS pathogenesis? *Cell.* 93: 677-680.

155. Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, P. Lusso. 1995. Identification of RANTES, MIP-1α, and MIP-β as the major HIV-suppressive factors produced by CD8+T cells. *Science.* 270: 1811-1815.

156. Proost, P., S. Struyf, D. Schols, G. Opdenakker, S. Sozzani, P. Allavena, A. Mantovani, K. Augustyns, G. Bal, A. Haemers, A. Lambeir, S. Scharpe, J. Van Damme, I. De Meester. 1999. Truncation of macrophage-tropic chemokine by CD26/depeptidyl-peptidase IV beyond its predicted cleavage site affects chemotactic activity and CC chemokine receptor 4 interaction. *J. Biol. Chem.* 274: 3988-3993.

157. Heveker, N., M. Montes, L. Germeroth, A. Amara, A. Trautmann, M. Alizon, J. Schneider-Mergener. 1998. Dissociation of the signalling and antiviral properties of SDF-1-dervided small peptides. *Curr. Biol.* 8: 369-376.

158. Crump, M. P., J. H. Gong, P. Loetscher, K. Rajarathnam, A. Amara, F. Arenzana-Seisdedos, J. L. Virelizier, M. Baggiolini, B. D. Sykes, I. Clark-Lewis. 1997. Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J.* 16: 6996-7007.

159. Arenzana-Seisdedos, F., J.-L. Virelizier, D. Rousset, I. Clark-Lewis, P. Loetscher, B. Moser, M. Baggiolini. 1996. HIV blocked by chemokine antagonist. *Nature.* 383: 400.

160. Donzella, G. A., D. Schols, S. W. Lin, J. A. Este, K. A. Nagashima, P. J. Maddon, G. P. Allaway, T. P. Sakmar, G. Henson, E. De Clercq, J. P. Moore. 1998. AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor. *Nat. Med.* 4: 72-77.

161. Howard, O. M., T. Korte, N. I. Tarasova, M. Grimm, J. A. Turpin, W. G. Rice, C. J. Michejda, R. Blumenthal, J. J. Oppenheim. 1998. Small molecule inhibitor of HIV-1 cell fusion blocks chemokine receptor-mediated function. *J. Leukoc. Biol.* 64: 6-13.

162. Doranz, B. J., K. Grovit-Ferbas, M. P. Sharron, S. H. Mao, M. B. Goetz, E. S. Daar, R. W. Doms, W. A. O'Brien. 1997. A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 coreceptor. *J. Exp. Med.* 186: 1395-1400.

163. Baba, M., O. Nishimura, N. Kanzaki, M. Okamoto, H. Sawada, Y. Iizawa, M. Shiraishi, Y. Aramaki, K. Okonogi, Y. Ogawa, K. Meguro, M. Fujino. 1999. A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proc. Natl. Acad. Sci. USA.* 96: 5698-5703.

164. Schols, D., S. Struyf, J. Van Damme, J. A. Este, G. Henson, E. De Clercq. 1997. Inhibition of T-tropic HIV strains by selective antagonization of the chemokine receptor CXCR4. *J. Exp. Med.* 186: 1383-1388.

165. Nguyen, G. T., M. Carrington, J. A. Beeler, M. Dean, L. M. Aledort, P. M. Blatt, A. R. Cohen, D. DiMichele, M. E. Eyster, C. M. Kessler, B. Konkle, C. Leissinger, N. Luban, S. J. O'Brien, J. J. Goedert, T. R. O'Brien. 1999. Phenotypic expressions of CCR5-delta32/delta32 homozygosity. *J. Acquir. Immune Defic. Syndr.* 22: 75-82.

166. Zou, Y. R., A. H. Kottmann, M. Kuroda, I. Taniuchi, D. R. Littman. 1998. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature.* 393: 595-599.

167. Tachibana, K., S. Hirota, H. Iizasa, H. Yoshida, K. Kawabata, Y. Kataoka, Y. Kitamura, K. Matsushima, S. Nishikawa, T. Kishimoto, T. Nagasawa. 1998. The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract. *Nature.* 393: 591-594.

168. Mosier, D. E., G. R. Picchio, R. J. Gulizia, R. Sabbe, P. Poignard, L. Picard, R. E. Offord, D. A. Thompson, J. Wilken. 1999. Highly potent RANTES analogues either prevent CCR5-using human immunodeficiency virus type 1 infection in vivo or rapidly select for CXCR4-using variants. *J. Virol.* 73: 3544-3550.

169. Kilby, J. M., S. Hopkins, T. M. Venetta, B. DiMassimo, G. A. Cloud, J. Y. Lee, L. Alldredge, E. Hunter, D. Lambert, D. Bolognesi, T. Matthews, M. R. Johnson, M. A. Nowak, G. M. Shaw, M. S. Saag. 1998. Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. *Nature Med.* 4: 1302-1307.

170. Judice, J. K., J. Y. K. Tom, W. Huang, T. Wrin, J. Vennari, C. J. Petropoulos, R. S. McDowell. 1997. Inhibition of HIV type 1 infectivity by constrained alpha-helical peptides: implications for the viral fusion mechanism. *Proc. Natl. Acad. Sci. USA.* 94: 13426-13430.

171. Zapp, M. L., M. R. Green. 1989. Sequence-specific RNA binding by the HIV-1 rev protein. *Nature.* 342: 714-716.

172. Malim, M. H., J. Hauber, S.-Y. Le, J. V. Maizel, B. R. Cullen. 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.

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

174. Eckert, D. M., V. M. Malashkevich, L. H. Hong, P. A. Carr, P. S. Kim. 1999. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target te gp41 coiled-coil pocket. *Cell*. 99: 103-115.

175. Eckert, D. M., V. N. Malashkevich, P. S. Kim. 1998. Crystal structure of GCN4-pIQI, a trimeric coiled coil with buried polar residues. *J. Mol. Biol.* 284: 859-865.

176. Ferrer, M., T. M. Kapoor, T. Strassmaier, W. Weissenhorn, J. J. Skehel, D. Oprian, S. Schreiber, D. C. Wiley, S. C. Harrison. 1999. Selection of gp41-mediated HIV-1 cell entry inhibitors from biased combinatorial libraries of non-natural binding elements. *Nature*. 6: 953-960.

177. Debnath, A. K., L. Radigan, S. Jiang. 1999. Structure-based identification of small molecule antiviral compounds targeted to the gp41 core structure ofthe human immunodeficiency virus type 1. *J. Med. Chem.* 42: 3202-3209.

178. Baker, K. A., R. E. Dutch, R. A. Lamb, T. S. Jardetzky. 1999. Structural basis for paramyxovirus-mediated membrane fusion. *Mol. Cell*. 3: 309-319.

179. Kobe, B., R. J. Center, B. E. Kemp, P. Poumbourios. 1999. Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *Proc. Natl. Acad. Sci. USA*. 96: 4319-4324.

180. Malashkevich, V. N., B. J. Schneider, M. L. McNally, M. A. Milhollen, J. X. Pang, P. S. Kim. 1999. Core structure of the envelope glycoprotein GP2 from ebola virus at 1.9 Å resolution. *Proc. Natl. Acad. Sci. USA*. 96: 2662-2667.

181. Rey, F. A., F. X. Heinz, C. Mandl, C. Kunz, S. C. Harrison. 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature*. 375: 291-298.

182. Weissenhorn, W., A. Carfi, K. H. Lee, J. J. Skehel, D. C. Wiley. 1998. Crystal structure of the Ebola viurs membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol. Cell*. 2: 605-616.

183. Stadler, K., S. L. Allison, J. Schalich, F. X. Heinz. 1997. Proteolytic activation of tick-borne encephalitis virus by furin. *J. Virol.* 71: 8475-8481.

184. Volkova, T. D., M. F. Vorovitch, V. T. Ivanov, A. V. Timofeev, O. M. Volpina. 1999. A monoclonal antibody that recognizes the predicted tick-borne encephalits virus E protein fusion sequence blocks fusion. *Arch. Virol.*: 1035-1039.

185. Iacono-Connors, L. C., J. F. Smith, T. G. Ksiazek, C. L. Kelley, C. S. Schmaljohn. 1996. Characterization of Langat virus antigenic determinants defined by monoclonal antibodies to E, NS1 and preM and identification of a protective, non-neutralizing preM-specific monoclonal antibody. *Virus Res.* 43: 125-136.

186. Allison, S. L., J. Schalich, K. Stiasny, C. W. Mandl, C. Kunz, F. X. Heinz. 1995. Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J. Virol.* 69: 695-700.

187. Roehrig, J. T., A. T. Johnson, A. R. Hunt, R. A. Bolin, M. C. Chu. 1990. Antibodies to dengue 2 virus E-glycoprotein synthetic peptides identify antigenic conformation. *Virol.* 177: 668-675.

188. Roehrig, J. T., A. R. Hunt, A. J. Johnson, R. A. Hawkes. 1989. Synthetic peptides derived from the deduced amino acid sequence of the E glycoprotein of Murray valley encephalitis virus elicit antiviral antibody. *Virol.* 171: 49-60.

189. Heinz, R. X., K. Stiasny, G. Puschner-Auer, H. Holzmann, S. L. Allison, C. W. Mandl, C. Kunz. 1994. Structural changes and functional control of the tick-borne encephalitis virus glycoportein E by the heterodimeric association with protein prM. *Virol.* 198: 109-117.

190. de Groot, R. J., W. Luytjes, M. C. Horzinek, B. A. van der Zeijst, W. J. Spaan, J. A. Lenstra. 1987. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *J. Mol. Biol.* 196: 963-966.

191. Singh, M., B. Berger, P. S. Kim. 1999. LearnCoil-VMF: Computational Evidence for Coiled-coil-like Motifs in Many Viral Membrane-fusion Proteins. *J. Mol. Biol.* 290: 1031-1041.

192. Gallaher, W. R. 1996. Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. *Cell.* 85: 477-478.

193. Hiroshi, I., S. Watanabe, A. Sanchez, M. A. Whitt, Y. Kawaoka. 1999. Mutational analysis of the putative fusion domain of Ebola virus glycoprotein. *J. Virol.* 73: 8907-8912.

194. Ruiz-Arguello, M. B., F. M. Goni, F. B. Pereira, J. L. Nieva. 1998. Phosphatidylinositol-dependent membrane fusion induced by a putative fusogenic sequence of Ebola virus. *J. Virol.* 72: 1775-1781.

195. Young, J. K., D. Li, M. C. Abramowitz, T. G. Morrison. 1999. Interaction of peptides with sequences from the Newcastle disease virus fusion protein heptad repeat regions. *J. Virol.* 73: 5945-5956.

196. Young, J. K., R. P. Hicks, G. E. Wright, T. G. Morrison. 1997. Analysis of a peptide inhibitor of paramyxovirus (NDV) fusion using biological assays, NMR and molecular modeling. *Virology.* 238: 291-304.

197. Yao, Q., R. W. Compans. 1996. Peptides corresponding to the heptad repeat sequence of human parainfluenza virus fusion protein are potent inhibitors of virus infection. *Virology.* 223: 103-112.

198. Wild, T. F., R. Buckland. 1997. Inhibition of measles virus infection and fusion with peptides corresponding to the leucine zipper region of the fusion protein. *J. Gen. Virol.* 78: 107-111.

199. Rapaport, D., M. Ovadia, Y. Shai. 1995. A synthetic pepitde corresponding to a conserved heptad repeat domain is a potent inhibitor of Sendai virus-cell fusion: an emerging similarity with functional domains of other viruses. *EMBO J.* 14: 5524-5531.

200. Joshi, S. B., R. E. Dutch, R. A. Lamb. 1998. A core trimer of the paramyxovirus fusion protein: parallels to influenza virus hemagglutinin and HIV-1 gp41. *Virology.* 248: 20-34.

201. Lambert, D. M., S. Barney, A. L. Lambert, K. Guthrie, R. Medinas, D. E. Davis, T. Bucy, J. Erickson, G. Merutka, S. R. Petteway. 1996. Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. *Proc. Natl. Acad. Sci. USA.* 93: 2186-2191.

202. Sutton, R. B., D. Fasshauer, R. Jahn, A. T. Brunger. 1998. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature.* 395: 347-353.

203. Poirier, M. A., W. Xiao, J. C. Macosko, C. Chan, Y.-K. Shin, M. K. Bennett. 1998. The synaptic SNARE complex is a parallel four-stranded helical bundle. *Nat. Struct. Biol.* 5: 765-769.

204. Weber, T., B. V. Zemelman, J. A. McNew, B. Westermann, M. Gmachl, F. Parlati, T. H. Sollner, J. E. Rothman. 1998. SNAREpins: minimal machinery for membrane fusion. *Cell.* 92: 759-772.

Figure 1. A comparison of cleaved and uncleaved hemagglutinin at neutral pH.

A. The primary structure of cleaved hemagglutinin, including a disulphide bond between residue 14 of HA1 and residue 137 of HA2. The sequences present in the x-ray crystal structures are colored in green, red and yellow. The remaining region of HA2 not present in the structures include approximately ten residues of the C-terminal portion of the ectodomain, the membrane anchoring sequence (TM) and the short intraviral domain.

B. The x-ray crystal structure of the bromelain-released form of cleaved HA (BHA), including all of HA1 and the first 175 residues of HA2 (25). Two HA1 subunits of the trimer are colored light blue, while two of the HA2 subunits are dark blue. One monomer is displayed with green HA1 and yellow HA2, as indicated in A. The eighteen residues that differ between this structure and the uncleaved HA0 structure are colored red. These include the six C-terminal residues of HA1 and the twelve N-terminal residues of HA2. The globular head domains composed of HA1 sit atop a stalk structure that is supported by an HA1 trimeric coiled coil. The fusion peptide domain (the N-terminus of HA2) is buried in the interior of the trimeric structure. The two termini that are produced from the cleavage event are labeled; they are approximately 22 Å from each other.

C. One monomer of the x-ray crystal structure of uncleaved HA (R329Q $HA0_S$) (22). The molecule is colored as in B. The residues forming the cleavage site form a loop that is exposed to solvent.

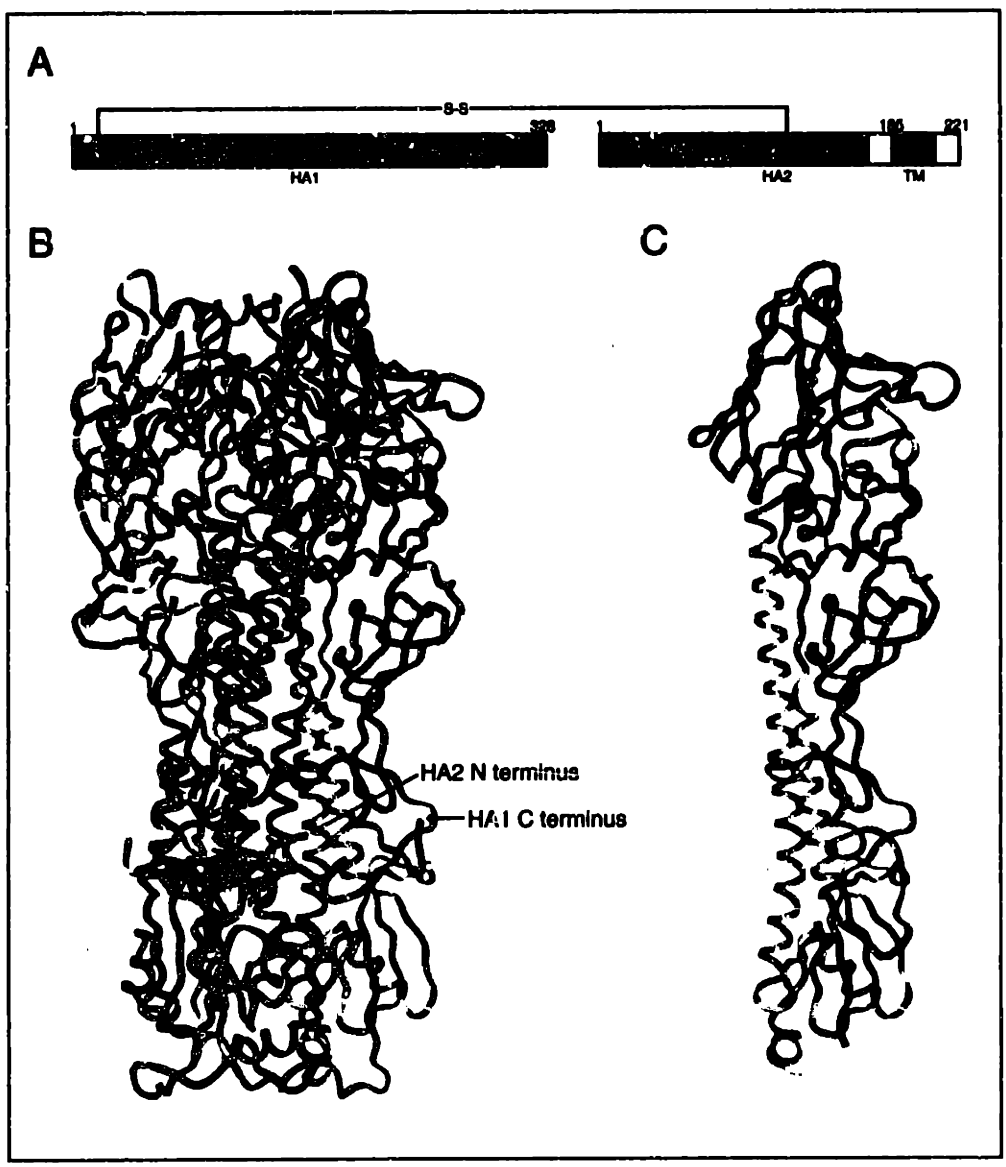

A
S-S
HA1
HA2
TM
B
C
HA2 N terminus
HA1 C terminus

Figure 2. A comparison of the structures of HA at neutral (non-fusogenic) pH and at low (fusogenic) pH.

A. A representation of the primary structure of cleaved hemagglutinin, depicting the color scheme for the remainder of the figure. The HA1 region contained in the neutral pH structure is indicated in light blue, and the HA2 region of that structure includes the dark blue, orange, grey, yellow, red and green colored areas. The low pH structure contains the HA1 residues indicated by the small light-blue box, and most of the HA2 residues colored orange, grey, yellow, red and green (residues 40-162).

B. The neutral pH structure of the bromelain-released form of HA (BHA), including all of HA1 and the first 175 residues of HA2 (25). All of the HA1 subunits are colored in light blue, and two of the three HA1 subunits are colored in dark blue. The remaining HA1 monomer is colored according to the color scheme in A with residues 38-54 (orange), 55-75 (grey), 76-104 (yellow), 1105-128 (red), and 129-175 (green). The N and C termini of the HA2 subunit are indicated. The HA2 subunit contains a central trimeric coiled coil composed of residues from the yellow and red regions. The green residues form a supportive hydrophobic base at the bottom of the coiled coil. At the top of the coiled coil, the grey residues loop around, such that another helical region (orange) lies anti-parallel to the coiled coil. The remaining N-terminal residues form a β-sheet and then tuck the fusion peptide into a hydrophilic pocket.

C. Two x-ray crystal structures of low-pH converted HA. On the left is TBHA2, containing residues 10-17 of HA1 and 40-162 of HA2 (24). These two fragments are disulphide bonded to each other. On the right is EHA2, containing residues 33-185 of HA2 (23). As in B, all of the HA2 subunits are colored in light blue (when present), and two of the HA1 subunits in each structure are in dark blue. The remaining HA2 monomers are colored as indicated in A. Significant differences between this structure and the neutral, non-fusogenic structure are very apparent. First, the grey region has formed a continuation of the central trimeric coiled coil, bringing the orange helical region to the N-terminus of the coiled coil. Although the fusion peptide region is not present in the structure, this significant structural change would relocate the fusion peptide up to 100 Å away from its position in the neutral pH structure and give it access to the target cell membrane. Second, part of the red region forms a loop at the bottom of the coiled coil, locating the remainder of the red helix anti-parallel to the coiled coil. Third, the green region continues in the same anti-parallel direction, forming a mostly

extended structure along the outside of the coiled coil. The EHA2 structure contains more ordered residues, and it is clear that the anti-parallel extension of the C-terminal residues continues. The top of this structure is capped. Several of the capping residues, located on the N-terminus of the HA2 subunit, are indicated in pink.

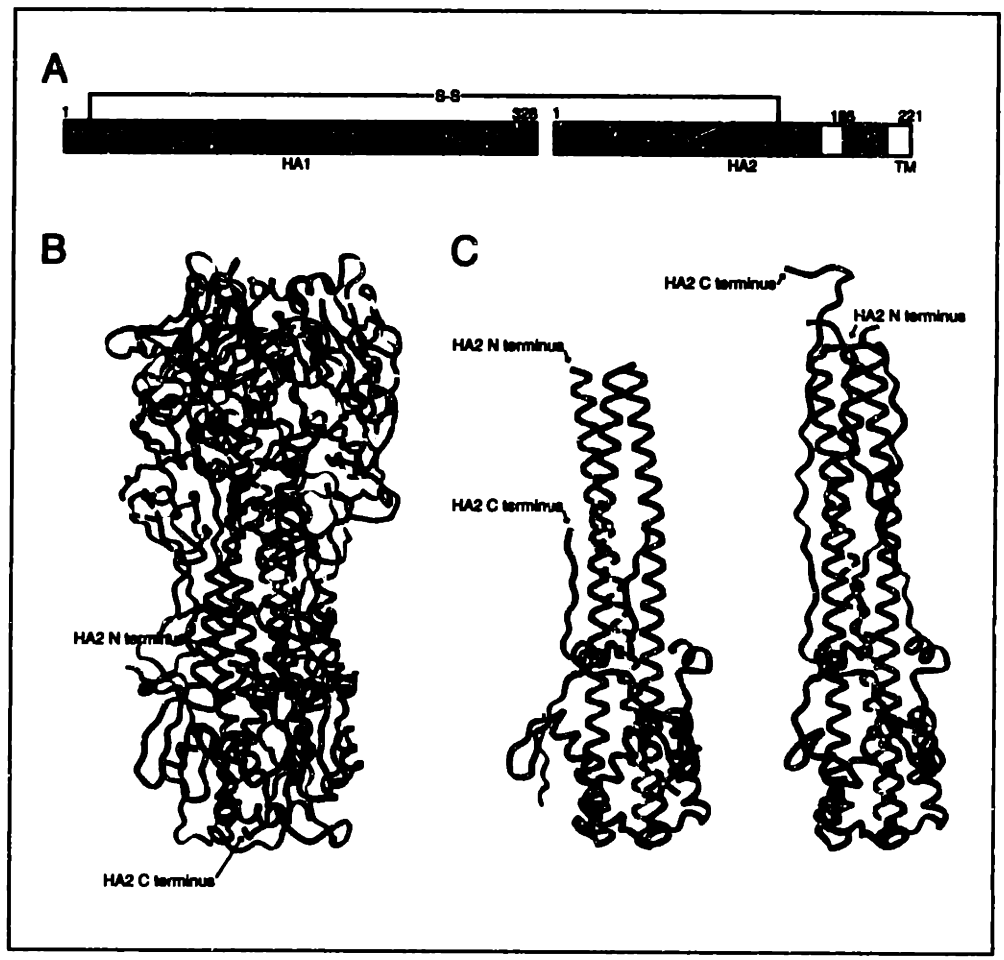

A
S-S
1
328
1
185
221
HA1
HA2
TM
B
HA2 N terminus
HA2 C terminus
C
HA2 C terminus
HA2 N terminus
HA2 N terminus
HA2 C terminus

Figure 3. The stalk/pore model of membrane fusion.

A schematic of the stalk/pore model of membrane fusion. The lipids of the outer membranes are colored blue, and the lipids in the inner membranes are colored black. The image on the left represents two distinct lipid bilayers, such as those surrounding a virus and host cell, respectively. The first step of membrane fusion involves formation of a stalk structure (represented in the middle) in which the two outer membranes have fused, but the two inner membranes remain intact. Stalk formation requires negative curvature and can be encouraged by cone-shaped lipids, as signified. The final stage of membrane fusion is pore formation, when the inner membranes fuse, allowing mixing of the contents of the virus and the cell. Pore formation requires positive curvature in the inner membrane and is facilitated by inverted cone-shaped lipids, as indicated.

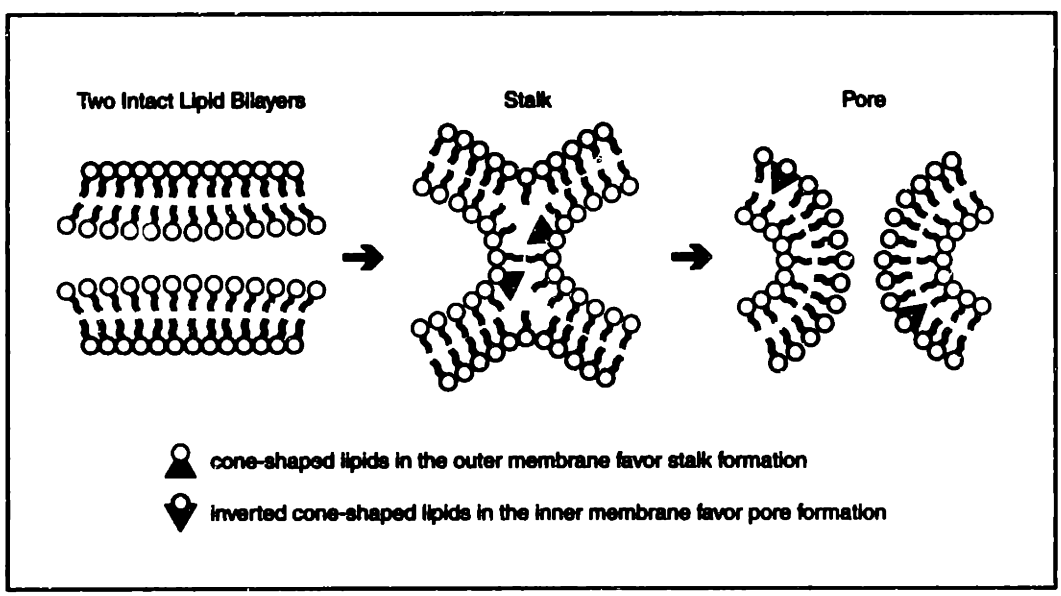

Two Intact Lipid Bilayers
Stalk
Pore
cone-shaped lipids in the outer membrane favor stalk formation
inverted cone-shaped lipids in the inner membrane favor pore formation

Figure 4. gp120 complexed with CD4 and a Fab representing the chemokine receptor binding site.

A representation of the 2.5 Å x-ray crystal structure of the gp120 core bound to CD4 and Fab 17b (118). CD4 and 17b are depicted with surface representations. CD4 is in light grey, and located behind the gp120 molecule. 17b is in darker grey and located beneath gp120. The secondary structure of gp120 is depicted, with β-sheets in maroon, α-helices in yellow, and loops and random coils in dark grey. The three domains of gp120 are indicated. The outer domain is located on the left in this representation. It is composed of two β-barrels. The barrel on top contains six β-sheets and enfolds an α-helix. The one on bottom contains seven β-sheets. The inner domain is located on the upper right, and contains a five-stranded β-sandwich on top, with a helix, two β-strand, helix bundle beneath. It contains both the N and C termini of the gp120 core, which are indicated. The bridging sheet is composed of four β-sheets on the lower right. Although the V1, V2, and V3 loops are absent in the structure, their bases are indicated. CD4 interacts with all three domains, while 17b mainly interacts with the bridging sheet.

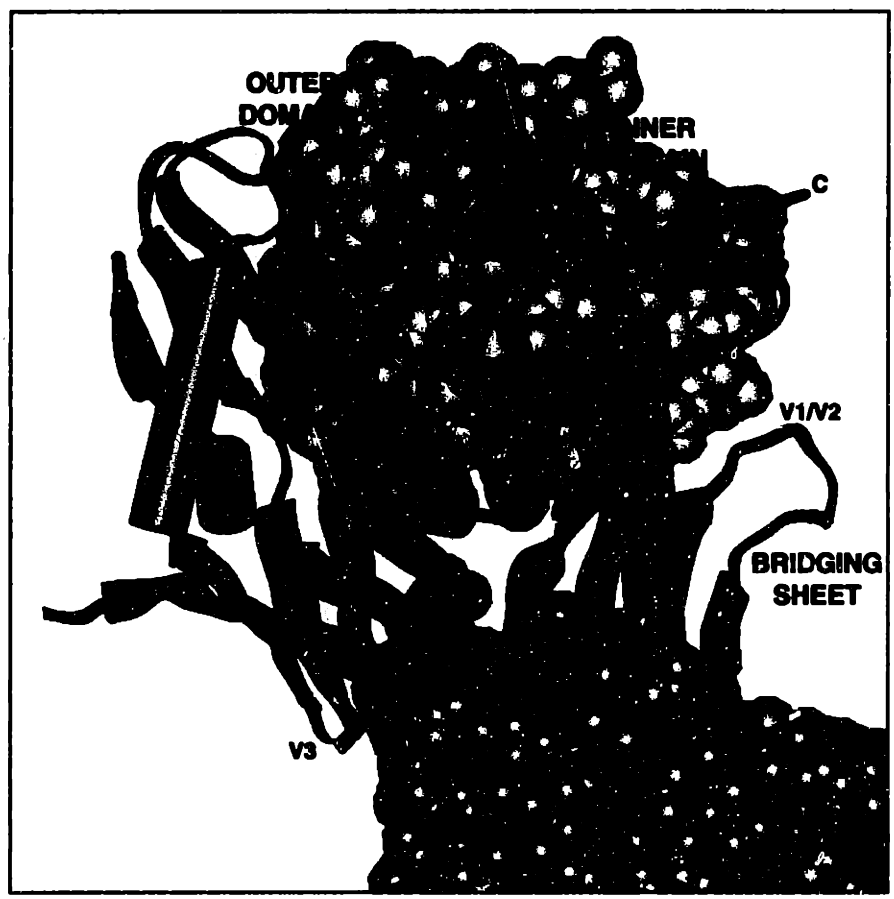

C
V1/V2
BRIDGING
SHEET
V3

Figure 5. The core structure of gp41.

A. A representation of the primary structure of HIV-1 gp41. The two hydrophobic regions, the fusion peptide (FP) and membrane-spanning domain (TM) are indicated in black. The two discontinuous peptides that were isolated through protein dissection techniques are indicated: the N peptide is grey, and the C peptide is red.

B. The x-ray crystal structure of a stable subdomain of HIV-1 gp41 (129), looking down the axis of symmetry from the N to the C terminus of the N peptides. The N peptides form a central trimeric coiled coil, shown in grey, and the C peptides, shown in red, wrap around the outside of the coiled coil in an anti-parallel fashion. Three residues of each C peptide are depicted in yellow, Trp 628, Trp 631, and Ile 635.

C. Another representation of the same structure. The central N peptides are depicted as a surface representation, and the C peptides as sticks. The yellow residues are the same as in B. They are shown to interact with a hydrophobic pocket on the surface of the N peptide coiled coil, which is highlighted in dark grey.

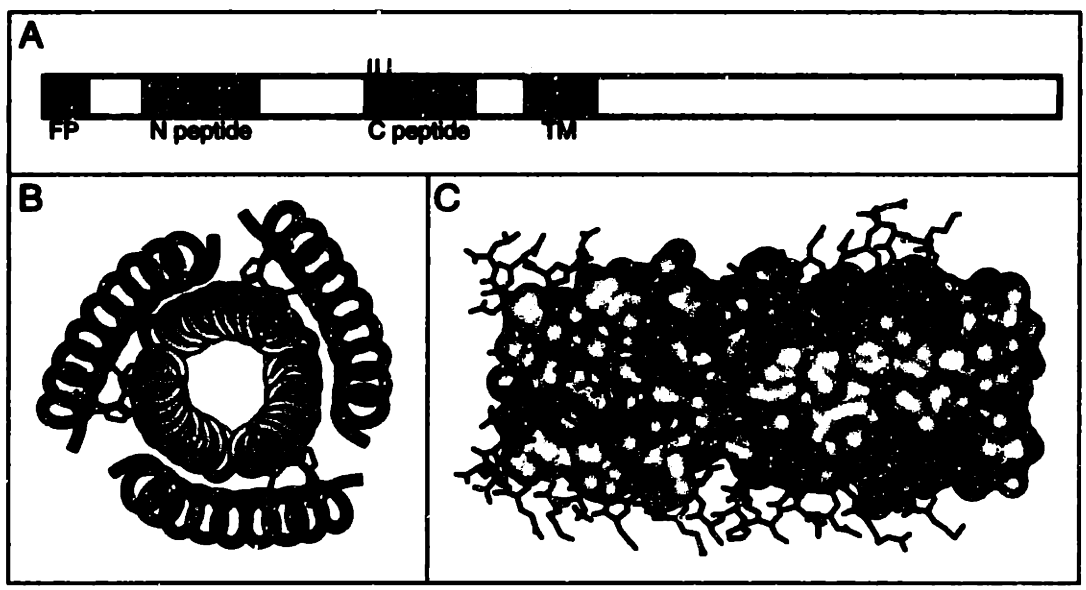

A
FP
N peptide
C peptide
TM
B
C

Figure 6. Model of HIV-1 membrane fusion.

A working model for HIV envelope protein-mediated membrane fusion. First gp120 binds CD4 and coreceptor, initiating conformational changes in gp41. gp41 inserts its fusion peptide domain into the host cell membrane. Transiently, a pre-hairpin intermediate exists, and is vulnerable to inhibition by C peptides or by D peptides (see 'Inhibiting HIV-1 Entry' section). When inhibitory peptides are not present, gp41 slowly resolves to the fusogenic trimer of hairpins structure.

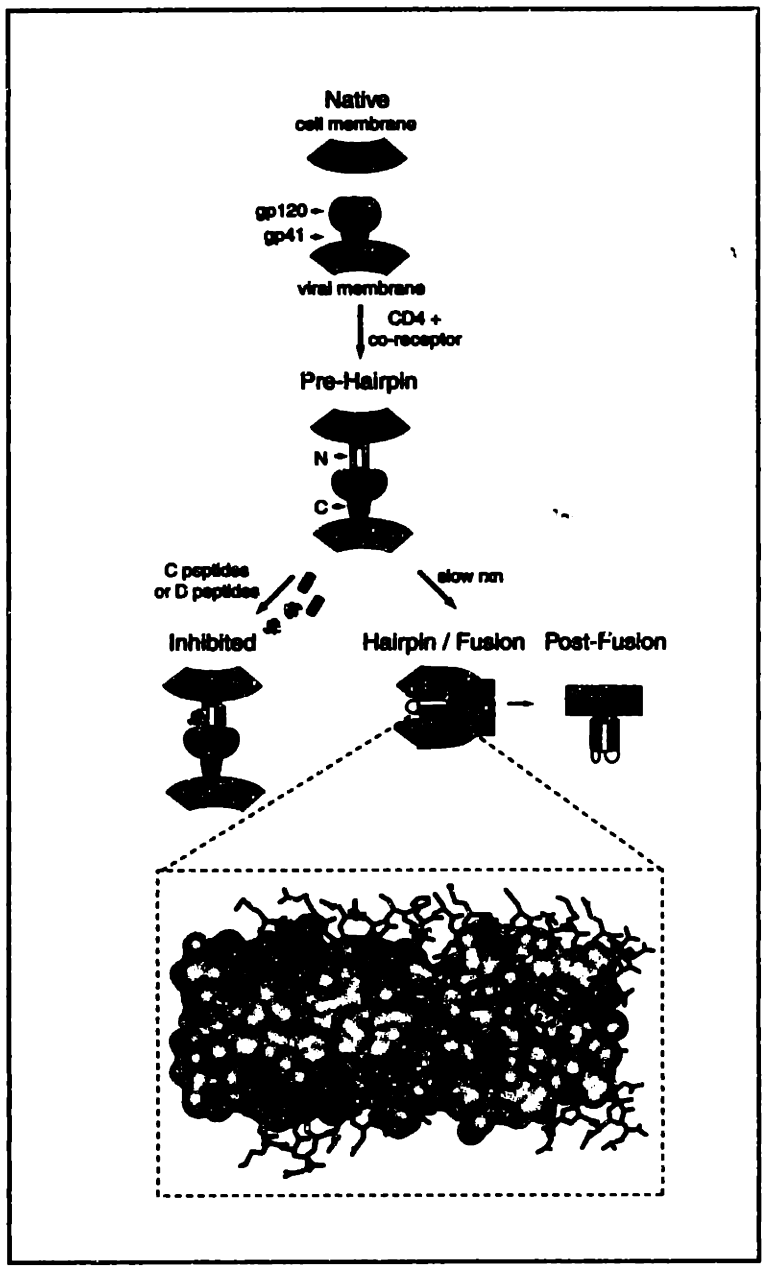

Native
cell membrane
gp120
gp41
viral membrane
CD4 +
co-receptor
Pre-Hairpin
N
C
C peptides
or D peptides
slow non
Inhibited
Hairpin / Fusion
Post-Fusion

Figure 7. The tick-borne encephalitis fusion protein.

Representation of the 2.0 Å x-ray crystal structure of dimeric E protein from tick-borne encephalitis virus (TBE) (181). The structure contains residues 1-395 of the E ectodomain. The view is looking down at the external face of the dimer. This would place the viral surface underneath the structure. One of the monomers depicts the secondary structure, with β-sheets in pink, α-helices in blue, and turns and random coils in white. The second monomer is depicted with ribbons and distinguishes the three structural domains: domains I, II and III, colored in light blue, blue and dark blue, respectively. Domain I, the central domain, is composed of an eight-stranded β-barrel. Domain II is the dimerization domain. The region of Domain II nearest to Domain I contains a five-stranded β-barrel and two α-helices. The domain extends toward Domain III of the other subunit with a three-stranded β-sandwich. The proposed fusion peptide (yellow) is located on a loop that interacts with Domain III of the other monomer. The third domain is folded into an IgC-like fold.

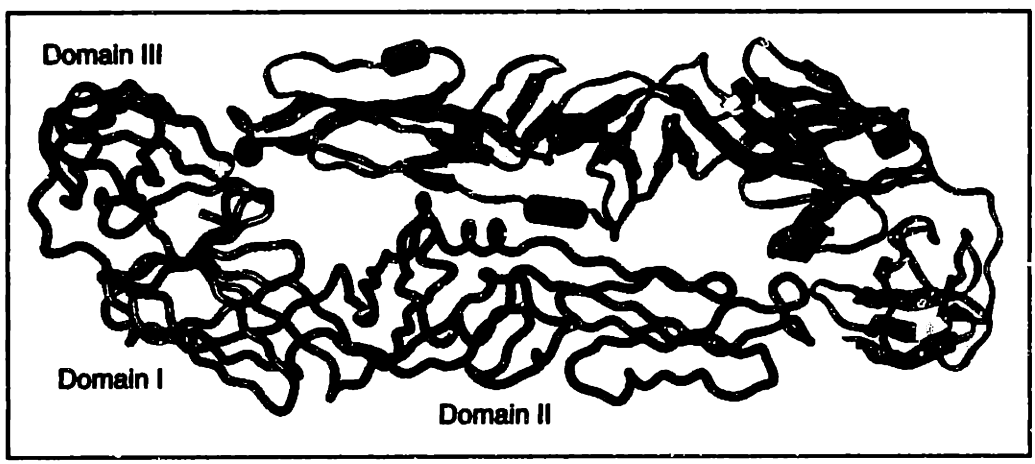
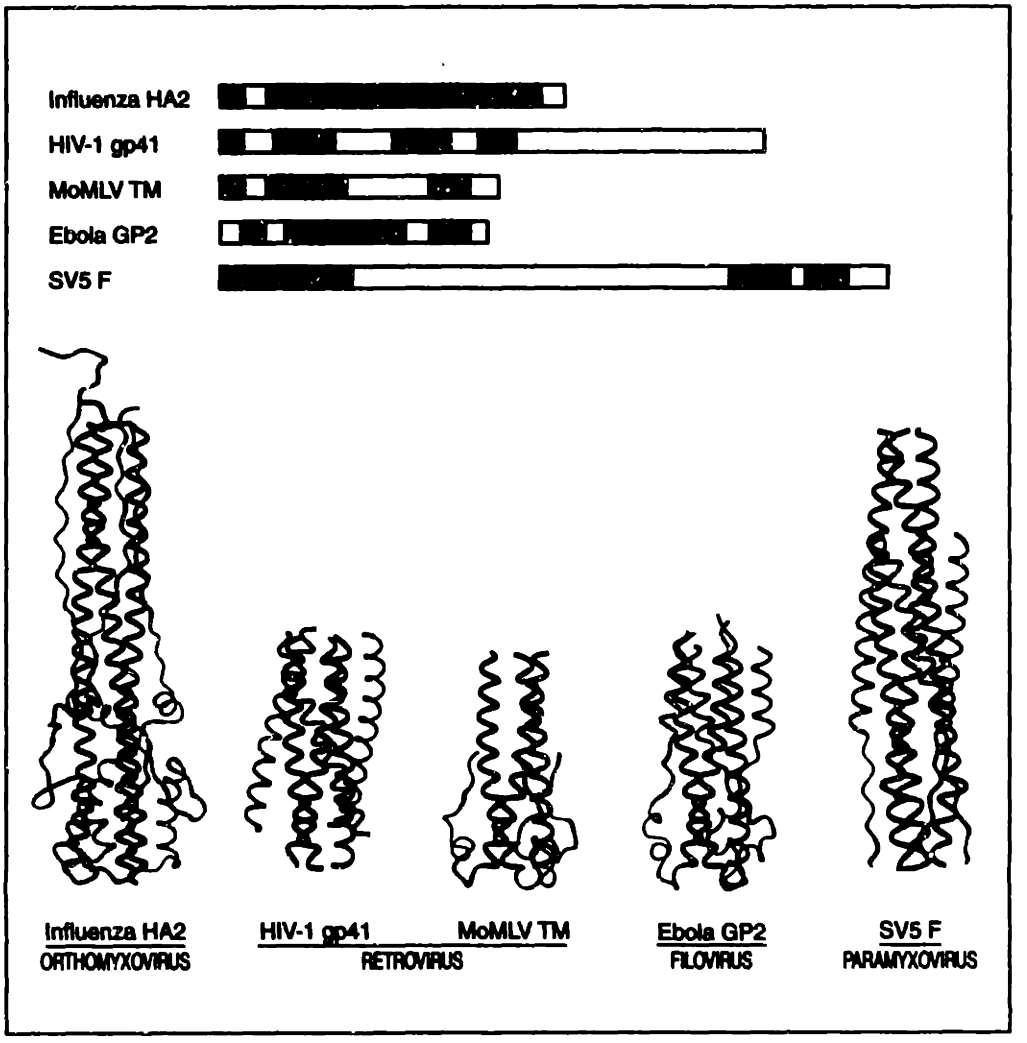

Domain III
Domain I
Domain II

Figure 8. Structures of fusion proteins from diverse enveloped viruses.

Five structures of the fusion proteins from five different viruses and four different viral families are depicted: influenza HA2 (23), HIV-1 gp41 (129), MoMLV TM (126), Ebola GP2 (180), and SV5 F (178). The primary representation of each fusion protein is depicted at the top of the diagram. They are approximately to scale to provide a context for each structural representation. The fusion peptide domains are indicated in red, and the transmembrane domains are indicated in black. The central region of each structure is indicated in blue, with the anti-parallel, outer regions in yellow. The bottom of the diagram has a ribbon depiction of each of the five structures with the name of the virus and the viral family underneath each structure.


Influenza HA2
HIV-1 gp41
MoMLV TM
Ebola GP2
SV5 F
Influenza HA2
ORTHOMYXOVIRUS
HIV-1 gp41
MoMLV TM
RETROVIRUS
Ebola GP2
FILOVIRUS
SV5 F
PARAMYXOVIRUS