HIV Entry:
A Biophysical and Mutational Analysis of gp41-mediated Membrane Fusion and Its Inhibition

By

Tara R. Suntoke

B.S. Biology
Carnegie Mellon University

Submitted to the Department of Biology in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

Signature of Author: _____
Department of Biology
May 20, 2005

Certified by: _____
Peter S. Kim
Former Professor, Department of Biology
President, Merck Research Laboratories
Thesis Co-Supervisor

Certified by: _____
David C. Chan
Professor, Department of Biology
California Institute of Technology
Thesis Co-Supervisor

Accepted by: _____
Stephen P. Bell
Professor, Department of Biology
Chairman, Biology Graduate Committee

Acknowledgments

I thank my parents, Russ and Ava, for their continual support and encouragement. Twenty three years ago, they made the bold decision to leave India and move to the United States in search of better opportunities for their children and themselves. From then until now, they have given us the world. I truly appreciate their hard work, dedication, generosity, and love. I also thank my sister and brother, Amy and Jamshed, for their constant support and companionship.

I am grateful for the encouragement that Peter Kim gave me during my early years of graduate school. I appreciate and admire his big-picture thinking and his rigorous approach to science. I learned a tremendous amount from the graduate students, post-docs, and technicians in the Kim lab, and I thank them for their thoughtful suggestions and help.

I thank David Chan for giving me the opportunity to continue my thesis work in his lab. Throughout my time at Caltech, David has guided my research projects, helped me to analyze and interpret results, and has offered invaluable advice when I was publishing and writing this thesis. He has always had an open-door policy and has been willing to discuss any of my concerns. I appreciate his advice, patience, and efforts in helping me complete my thesis work. I also thank all the Chan lab members, past and present, who have been great labmates, friends, and co-workers.

My committee members, Rick Young, Harvey Lodish, and Bob Sauer, have been instrumental in my progress, and it is through their suggestions that I have been able to carry out successful projects. I am grateful for their input and encouragement.

Finally, Erik, I can't thank you enough for so many things—your unwavering support, encouragement, advice, friendship, and love...just to name a few. You've kept me sane through this whole thing, you make me laugh, you're always up for a trip or a movie, and you are always caring, generous, and considerate. You helped me to believe that I can do this. Thank you so much for everything. I truly value our time together, and can't wait to take the next step, wherever that may be!

HIV Entry:
A Biophysical and Mutational Analysis of gp41-mediated Membrane Fusion and Its Inhibition

By

Tara R. Suntoke

Submitted to the Department of Biology on May 20, 2005
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
in Biology

Abstract

The experiments described in this thesis were designed to elucidate the manner in which the HIV-1 envelope protein (Env) initiates infection of host cells, and to develop inhibitors of viral entry. Env comprises two non-covalently attached subunits, gp120 and gp41, that associate as a trimer on the virion surface. Once gp120 contacts the target cell, gp41 undergoes extensive conformational changes to mediate fusion of viral and cellular membranes. First, a short hydrophobic stretch of residues at the gp41 N-terminus insert into the target membrane, anchoring the protein in both viral and cellular membranes. This 'prehairpin' intermediate structure exposes an N-terminal α-helical coiled coil that is the target of promising antiviral peptides and small molecules. A previously unstudied region of N-terminal trimer was stabilized by fusion to a trimeric scaffold peptide and biophysically characterized (Chapter 2). This hybrid peptide itself potently inhibited HIV fusion, and the basis for this inhibition was assessed by studying mutant molecules. Efforts to use this peptide as an immunogen to elicit anti-gp41 antibodies are also outlined. Similar design strategies may be useful in developing N-terminal peptide inhibitors and HIV vaccine candidates, and in screening for antiviral molecules that bind to this region. As fusion progresses, the prehairpin intermediate resolves into a hairpin structure. This critical transition involves interaction of the N-terminal coiled coil with the gp41 C-terminal region. Evidence suggests that this N-C association provides the energy necessary to promote juxtaposition and merging of viral and cellular membranes. This hypothesis was tested using a biophysical and cell biological approach (Chapter 3), in which residues essential for this transition were mutated and analyzed. These studies confirm the hypothesis and highlight the importance of specific hydrophobic and polar contacts between the N- and C- terminal gp41 regions. This work contributes to a detailed understanding of the gp41 fusion machinery; furthermore, it shows that such knowledge can be used to design effective viral entry inhibitors. Finally, Chapter 4 places this work in the context of a broad overview of current drug and vaccine developments, and addresses some of the significant challenges that confront HIV researchers.

Thesis Supervisors:
Dr. Peter S. Kim, Former Professor of Biology, MIT
(Current Title: President, Merck Research Labs)
Dr. David C. Chan, Professor of Biology, California Institute of Technology

Table of Contents

CHAPTER 1

INTRODUCTION: HIV-1 MEMBRANE FUSION

HIV-1 is the etiological agent of the acquired immunodeficiency syndrome (AIDS) pandemic. The virus is a member of the *Retrovirus* family and the *Lentivirus* genus. As a retrovirus, HIV-1 contains a single-stranded RNA genome that is converted to DNA by the viral enzyme reverse transcriptase and inserted into the host genome. As a lentivirus, HIV-1 is able to infect nondividing cells and causes prolonged, but ultimately fatal, disease progression.

HIV research, from basic biology of the virus to vaccine trials, is fueled by an incentive even greater than scientific knowledge: the desire to prevent, or at least reduce, human disease and suffering caused by this deadly virus. The work in this thesis involves cell biological and biophysical studies of HIV-1 gp41, the viral protein responsible for fusing viral and cellular membranes and thereby initiating infection. This chapter serves as an introduction to the biology of HIV-1 replication, and a summary of the vast body of work detailing the entry of this virus into host cells. Chapter 2 describes an attempt to generate a novel vaccine candidate, in which a portion of the gp41 ectodomain was stabilized as a hybrid protein, biophysically characterized, and injected into guinea pigs. Along the way, this molecule was also examined for its viral inhibitory activity, and subsequent studies led to the generation of a nanomolar inhibitor of HIV-1 entry. Chapter 3 illustrates a biophysical and mutagenic approach to understanding the mechanism by which gp41 mediates fusion. These studies provide further insight into the membrane fusion process and the mechanism of action of entry inhibitors and drug-resistant viruses. Finally, in Chapter 4, current advancements in drug and vaccine research as well as some of the challenges facing HIV researchers today are examined.

Virion structure and genomic organization

The HIV-1 genome (Figure 1*A*) encodes three major polyproteins that are common to all retroviruses: Gag, Pol, and Env. Gag, a polyprotein that is cleaved into matrix (MA), capsid (CA), and nucleocapsid (NC), and Env are structural genes. The viral enzymes protease (PR), reverse transcriptase (RT), and integrase (IN) are encoded by the Pol precursor protein. The virus also contains the regulatory proteins Tat and Rev, as well as the accessory proteins Vif, Vpr, Vpu, and Nef. The function of these proteins in the lifecycle is discussed below.

Approximately 120 nm in diameter, the HIV-1 virion contains at its center two single-stranded RNA molecules coated with viral NC proteins. This core and the trio of viral enzymes,

PR, RT, and IN, are enclosed within a cage of structural CA and MA proteins. The outermost shell of the virus consists of the envelope protein, Env, embedded in a lipid bilayer.

HIV-1 Lifecycle

The HIV-1 lifecycle (Figure 1*B*) begins with attachment of the virion to target cells expressing surface CD4 and one of two chemokine receptors, CCR5 or CXCR4. This stage, as well as the subsequent fusion of viral and cellular membranes, is mediated by gp120 and gp41, the non-covalently–associated subunits of the HIV envelope glycoprotein (Env) (for a review of the HIV lifecycle see (23), (13)).

Upon entering the host cell, the virion is partially uncoated to reveal the genomic RNA and viral RT begins synthesis of double-stranded DNA. The nucleoprotein preintegration complex (PIC), containing some Gag and Pol proteins as well as cellular factors, transports this nascent DNA through the cytoplasm and into the nucleus. The accessory protein Vpr is also involved in this transport. Active nuclear import by the PIC machinery gives viral nucleic acid access to nondividing cells arrested in the G_1 phase of the cell cycle—a feature unique to lentiviruses. The accessory protein Vif, or viral infectivity factor, is highly conserved among lentiviruses. Vif functions to allow viral replication in otherwise non-permissive cells, at a step that follows viral entry and precedes DNA integration (23).

Insertion of DNA into the host genome is mediated by the integrase protein. The site of integration plays a significant role in determining the expression pattern of the provirus and in creating deleterious mutations within the host genome. This step completes the early phase of the viral lifecycle, in which viral proteins perform the majority of the essential functions.

The subsequent late phase is catalyzed primarily by cellular enzymes that perform their normal transcription and translation functions. Viral regulatory proteins and cellular transcription factors contribute to viral RNA synthesis. Tat stimulates expression from the viral long terminal repeat (LTR) by binding to a stem-loop structure in the LTR (called TAR) and enhancing the processivity of RNA Polymerase II. Early transcripts leave the nucleus fully spliced. Later, unspliced and singly spliced mRNA variants are exported from the nucleus by the viral regulatory protein Rev.

Cellular machinery translates viral messages in the cytoplasm. The structural (Gag) and enzymatic (Pol) proteins that are enclosed in the virion, along with full-length genomic RNA, associate in patches of plasma membrane that contain specific lipid content (9, 38). Env is transported through the secretory pathway and is inserted in the plasma membrane, where it associates with newly forming viral particles. The viral accessory protein Vpu prevents premature interaction between Env and its cellular receptor protein CD4 in the endoplasmic reticulum by binding to CD4 and causing its degradation. The importance of Nef, another accessory protein, in *in vitro* studies is debated. However, Nef is necessary to generate high titers of virus in HIV-infected patients. This protein also downregulates cell-surface expression of CD4, the viral receptor, and the major histocompatibility complex I (MHC I) (23).

As the virion assembles at the plasma membrane, it induces membrane curvature that allows it to bud off the plasma membrane. This particle is initially non-infectious and is known as an immature virion. The action of viral protease to cleave the Gag and Pol precursor proteins results in internal structural reorganization to form the mature, infectious virion.

Envelope processing

The envelope glycoprotein (Env) is synthesized from singly spliced viral mRNA. Upon translation, the 160 kilodalton Env precursor, gp160, is translocated into the rough endoplasmic reticulum where it is glycosylated, forms disulfide bonds, and oligomerizes (17, 53). gp160 is then transported into the Golgi network, where cellular furin-like proteases cleave the precursor into two non-covalently attached subunits, gp120 and gp41 (28). This proteolytic cleavage is essential for Env function. The receptor-binding subunit, gp120, is anchored to the viral membrane by its association with the transmembrane glycoprotein gp41. Since this interaction is relatively weak, a substantial portion of gp120 is shed from the virion or infected cell surface. The Env glycoprotein is a trimer on the virion surface (8, 46, 48).

Env is heavily glycosylated (gp120 has 24 potential N-linked glycosylation sites and gp41 has five) and contains several disulfide bonds (nine within gp120 and one within gp41) (Figure 2*A* and Appendix 1). Extensive glycosylation of gp120, the subunit most exposed on the virion surface, permits the virus to evade the host immune response.

gp41 architecture

gp41 contains several important subdomains (Figure 2*A*) and requires proteolytic cleavage of the Env precursor in order to properly function. A stretch of approximately twenty hydrophobic amino acids at the newly created amino terminus is termed the fusion peptide. This region is followed by two amphipathic helical domains, called the N-terminal heptad repeat (N-HR or N-helix) and C-terminal heptad repeat (C-HR or C-helix), that are linked by a short disulfide-bonded loop. The C-HR domain is followed by the transmembrane segment (TM), and a long intraviral portion called the cytoplasmic tail.

Fusion model

gp120 initiates infection by binding to cells expressing CD4, a cell-surface marker for a subset of T cells, macrophage, and other phagocytic cells. Subsequent binding of gp120 to its coreceptor, CXCR4 or CCR5, (both members of a family of chemokine receptors with seven membrane-spanning domains), triggers a series of conformational changes in gp41, which facilitates the merging of viral and cellular membranes.

The current model for gp41-mediated membrane fusion involves the interaction and conformational changes of several of its distinct domains (14, 19, 26, 44). Upon gp120-receptor interaction, the gp41 N-terminal fusion peptide is released from its native state contacts and inserts into the target cell membrane. A prehairpin intermediate then forms, in which the N-terminal helices form a trimeric coiled coil and gp41 is extended between the viral and target cell membranes. This structure collapses into a trimer-of-hairpins, with C-terminal helices packing in an antiparallel manner into the conserved grooves between adjacent N helices, which represents the fusion-active core of gp41 (Figure 2, *B* and *C*). Formation of this six-helix bundle pulls the amino and carboxy termini of the gp41 ectodomain into close proximity. The resulting juxtaposition of viral and cellular membranes leads to membrane fusion.

The six-helix bundle structure is intimately linked to fusion. However, although its formation is thought to provide the energy required to fuse lipid bilayers, the exact mechanism by which this occurs remains unknown. Several models have been proposed. One suggests that trimer-of-hairpins formation could occur concurrently with the merging of lipid bilayers and thus directly provide the energy required for membrane fusion (2). In 1999, Baker and colleagues

solved the structure of the paramyxovirus F (fusion) protein core from simian virus type 5 (SV5). They found that at the N-terminus of the N-peptide, the coiled coil extends into the putative fusion peptide (FP) in a continuous, relatively inflexible, α-helical structure. At the C-terminus of the C-peptide, the seven residues between the crystallized core and the transmembrane domain (TM) were deleted and found to be unnecessary for F glycoprotein function. These results were interpreted to indicate that flexible regions adjacent to the FP and TM domains are not essential for the fusion machinery. Furthermore, they suggested that membranes may merge concurrently with six-helix bundle formation.

In the case of the influenza hemagglutinin 2 (HA2) fusion protein, several residues from the N-terminus of the N trimer form a cap at the end of the coiled coil, ending the helical conformation well before the fusion peptide (44). Thus, in the case of this well-characterized orthomyxovirus, a flexible linker adjacent to the FP and TM domains could theoretically allow opposing membranes to remain apart even after trimer-of-hairpins formation.

The merging of lipid bilayers and mixing of viral and cellular contents are proposed to occur through another gp41 intermediate, termed hemifusion, and formation of a fusion pore. Studies with influenza HA suggest that several six-helix bundles interact with each other at the fusion site to create an initial fusion pore (5, 15, 37). Electrophysiological measurements show that a small, flickering pore eventually expands and allows passage of viral contents into the target cell (12, 24, 45). Hemifusion is a potential intermediate between six-helix bundle formation and full fusion, occurring immediately prior to initial pore formation. The intermediate consists of mixed outer (i.e. opposing) leaflets of the viral and target cell membranes—the inner leaflets remain intact and keep viral and cellular contents separate. As the fusion pore enlarges, the inner leaflets of each bilayer merge, thus completing the fusion process (reviewed in (19, 26)). While hemifusion has not been observed under normal circumstances using currently available techniques, HA containing TM truncations (1), TM replacement by a GPI anchor (30), and a point mutation in the FP domain (39) displays a hemifusion phenotype. The first example of retroviral hemifusion was recently reported for Moloney murine leukemia virus (MoMLV). A point mutation in the receptor-binding subunit caused lipid but not content mixing of dye-labeled virions and target cells, although lipid mixing for the mutant was delayed when compared to wild-type (55). To date, no such observation has been made for HIV.

Structures of viral fusion proteins

At least three distinct conformations of the gp41 envelope glycoprotein have been proposed: the native or pre-fusogenic state, the prehairpin intermediate, and the fusion-active or fusogenic state. Several lines of evidence suggest that the trimer-of-hairpins structure is not present in the native or prehairpin conformations of the envelope glycoprotein. First, a comparison of pre- and post-fusogenic HA2 structures from influenza virus demonstrates that large conformational changes upon exposure to low pH trigger six-helix bundle formation (44). In fact, in the native state structure, only a portion of the N-terminal region of HA2 is helical. Second, peptide inhibitors derived from the HIV gp41 core are likely to inhibit fusion in a dominant negative manner by preventing trimer-of-hairpins formation (7, 18). The extreme thermostability of the six-helix bundle strongly suggests that once formed, it will not dissociate; thus, the mechanism of action of inhibitory peptides supports the hypothesis that the trimer-of-hairpins structure is not present until a late stage in the fusion pathway (25, 31, 32). Third, it is predicted that fusion proteins will adopt the lowest energy conformation when folded in the absence of the receptor binding protein. Indeed, in the absence of HA1, the receptor binding subunit of influenza, HA2 folds into the fusogenic trimer-of-hairpins structure (19).

Although no retroviral native state structure has yet been solved, structures of the fusion-active cores of several viruses provide much insight into the mechanism of viral membrane fusion. Remarkably, the core structures of retroviruses (e.g. HIV, simian immunodeficiency virus (SIV), MoMLV), orthomyxoviruses (influenza), paramyxoviruses (SV5, human respiratory syncytial virus), filoviruses (Ebola), and coronaviruses (SARS-CoV) (14, 19, 54) all reveal the same antiparallel six-helix bundle conformation. This striking similarity implies an evolutionary conservation of the fusion machinery.

In the fusion-active structure of HIV-1 gp41, the N-terminal helices form a trimeric coiled coil (Figure 2*C*). C-terminal helices pack into the grooves between adjacent N-helices in an antiparallel manner. The structure of the gp41 core also reveals a large hydrophobic cavity on the surface of the N-terminal trimer that forms a docking site for three C-terminal residues.

Cavity-containing or cavity-binding inhibitors of infection

X-ray crystal structures of the gp41 core have identified critical contacts between N and C-terminal helices (8, 46, 48). This interface is lined with predominantly hydrophobic residues that are among the most highly conserved within gp41. Peptides derived from the N- or C-terminal helices (termed N or C peptides) are potent inhibitors of HIV-1 infection, acting in a dominant negative manner to prevent formation of the gp41 six-helix bundle. These peptides include enfuvirtide, or T20, (27), the first FDA-approved viral entry inhibitor used to fight HIV-1 infection, several N peptides (4, 18, 33, 52), and a trimer-of-hairpins derivative, called 5-Helix, which lacks one C helix and presumably inhibits infection by sequestering the C-terminal region of gp41 (40). The inhibitory activity of a potent C peptide, C34, depends largely on its ability to form critical contacts within the N-terminal hydrophobic cavity (7). Furthermore, small peptides (20, 43) and synthetic molecules (21) that bind selectively to this coiled-coil cavity also inhibit fusion. A detailed summary of HIV inhibitors, including those that prevent viral entry, is included in Chapter 4.

Mutagenesis of the gp41 trimer-of-hairpins

A vast amount of work describing mutagenesis of gp41 exists in the literature. While some mutants appear to function normally, many affect gp41-gp120 interaction or alter gp41 fusion activity. Since a mutagenic approach to studying gp41 function was taken in this work (Chapters 2 and 3), a review of the relevant mutations is helpful to understanding the basis for mutant selection.

Table 1 is a representative list of several categories of gp41 mutants. Since the gp41 fusion-active core structure was determined in 1997, researchers have biophysically characterized peptide complexes derived from the N- and C-HR regions in order to dissect gp41 function. Table 1 includes a summary of the core structure stability for each set of mutants, if such work has been published. Cell biological and mutagenesis experiments, coupled with these investigations of core stability, have led to an emerging hypothesis correlating the strength of interhelical (i.e. between N- and C-helices) packing interactions with gp41 fusion activity.

The gp41 mutants in Table 1 are divided into six categories. The first represents mutations that have no effect on gp41 function, indicating that, most likely, these residues do not

make critical contacts in the native state or the six-helix bundle structure (reviewed in (41)). The following two categories represent mutants with abnormal precursor processing or Env presentation on the viral surface. It is therefore expected that these mutants are fusion-impaired (10, 11, 35, 47, 51). Mutations that result in a weak trimer-of-hairpins structure often decrease gp41-mediated fusion, as seen in the fourth category (3, 6, 22, 29, 34-36, 49, 51). Since we are interested specifically in mutations that affect gp41 function at the stage of trimer-of-hairpins formation, these mutants are particularly intriguing. Similarly, there is one category—containing the only example of such a mutant in the literature—in which increased trimer-of-hairpins stability is correlated with an increase in fusogenicity (6, 42). Mutations in the sixth and final group were created in the N-HR to abolish interaction between N- and C-helices (16, 49, 50). While trimer-of-hairpins stability was never tested for these mutants, it is a likely explanation of their impaired gp41 function.

Finally, a review of Table 1 reveals that position within gp41 does not dictate the function of a particular residue. Furthermore, the nature of the substitution significantly affects the phenotype of the mutant, as seen with the different substitutions of the isoleucine residue at position 573 (I573).

These mutagenesis experiments strengthen the model for viral membrane fusion, in which trimer-of-hairpins formation drives membrane merger. However, it is difficult to effectively compare the results of these studies. The variety of cell types used for cell-cell fusion experiments, the methods of scoring syncytia, the nature of the substitutions (hydrophobic, polar, conservative, or nonconservative), and the recombinant trimer-of-hairpins peptides used to assess core stability vary considerably. Thus, a systematic and rigorous analysis of hairpin mutants, especially those that affect N-C interaction, would be a useful addition toward understanding how the gp41 fusion machinery functions. Identifying important N-C contacts could lead to the design of improved fusion inhibitors, and may uncover specific subdomains within gp41 that have unique roles in membrane fusion.

References

1. **Armstrong, R. T., A. S. Kushnir, and J. M. White.** 2000. The transmembrane domain of influenza hemagglutinin exhibits a stringent length requirement to support the hemifusion to fusion transition. J Cell Biol **151**:425-37.
2. **Baker, K. A., R. E. Dutch, R. A. Lamb, and T. S. Jardetzky.** 1999. Structural basis for paramyxovirus-mediated membrane fusion. Mol Cell **3**:309-19.
3. **Bar, S., and M. Alizon.** 2004. Role of the ectodomain of the gp41 transmembrane envelope protein of human immunodeficiency virus type 1 in late steps of the membrane fusion process. J Virol **78**:811-20.
4. **Bewley, C. A., J. M. Louis, R. Ghirlando, and G. M. Clore.** 2002. Design of a novel peptide inhibitor of HIV fusion that disrupts the internal trimeric coiled-coil of gp41. J Biol Chem **277**:14238-45.
5. **Blumenthal, R., D. Sarkar, S. Durell, D. Howard, and S. Morris.** 1996. Dilation of the influenza hemagglutinin fusion pore revealed by the kinetics of individual cell-cell fusion events. J. Cell Biol. **135**:63-71.
6. **Cao, J., L. Bergeron, E. Helseth, M. Thali, H. Repke, and 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**:2747-55.
7. **Chan, D. C., C. T. Chutkowski, and P. S. Kim.** 1998. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. PNAS **95**:15613-7.
8. **Chan, D. C., D. Fass, J. M. Berger, and P. S. Kim.** 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell **89**:263-73.
9. **Chazal, N., and D. Gerlier.** 2003. Virus entry, assembly, budding, and membrane rafts. Microbiol. Mol. Biol. Rev. **67**:226-237.
10. **Chen, S. S.** 1994. Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane protein gp41. J Virol **68**:2002-10.
11. **Chen, S. S., C. N. Lee, W. R. Lee, K. McIntosh, and T. H. Lee.** 1993. Mutational analysis of the leucine zipper-like motif of the human immunodeficiency virus type 1 envelope transmembrane glycoprotein. J Virol **67**:3615-9.
12. **Chernomordik, L. V., V. A. Frolov, E. Leikina, P. Bronk, and 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.

13. **Coffin J. M., H. S. H., Varmus H. E.** 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

14. **Colman, P. M., and M. C. Lawrence.** 2003. The structural biology of type I viral membrane fusion. Nat Rev Mol Cell Biol **4**:309-19.

15. **Danieli, T., S. Pelletier, Y. Henis, and J. 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.

16. **Dubay, J. W., S. J. Roberts, B. Brody, and E. Hunter.** 1992. Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. J Virol **66**:4748-56.

17. **Earl, P. L., B. Moss, and R. W. Doms.** 1991. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. J Virol **65**:2047-55.

18. **Eckert, D. M., and P. S. Kim.** 2001. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. PNAS **98**:11187-92.

19. **Eckert, D. M., and P. S. Kim.** 2001. Mechanisms of viral membrane fusion and its inhibition. Annu Rev Biochem **70**:777-810.

20. **Eckert, D. M., V. N. Malashkevich, L. H. Hong, P. A. Carr, and P. S. Kim.** 1999. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell **99**:103-15.

21. **Ferrer, M., T. M. Kapoor, T. Strassmaier, W. Weissenhorn, J. J. Skehel, D. Oprian, S. L. Schreiber, D. C. Wiley, and S. C. Harrison.** 1999. Selection of gp41-mediated HIV-1 cell entry inhibitors from biased combinatorial libraries of non-natural binding elements. Nat Struct Biol **6**:953-60.

22. **Follis, K. E., S. J. Larson, M. Lu, and J. H. Nunberg.** 2002. Genetic evidence that interhelical packing interactions in the gp41 core are critical for transition of the human immunodeficiency virus type 1 envelope glycoprotein to the fusion-active state. J Virol **76**:7356-62.

23. **Freed E. O., M. A. Martin.** 2001. in *Fields Virology*, Fourth Ed, Vol. 2: 1971-2041. Lippincott Williams & Wilkins, Philadelphia.

24. **Frolov, V. A., A. Y. Dunina-Barkovskaya, A. V. Samsonov, and J. Zimmerberg.** 2003. Membrane permeability changes at early stages of influenza hemagglutinin-mediated fusion. Biophys. J. **85**:1725-1733.

25. **Furuta, R. A., C. T. Wild, Y. Weng, and C. D. Weiss.** 1998. Capture of an early fusion-active conformation of HIV-1 gp41. Nat Struct Biol **5**:276-9.

26. **Gallo, S. A., C. M. Finnegan, M. Viard, Y. Raviv, A. Dimitrov, S. S. Rawat, A. Puri, S. Durell, and R. Blumenthal.** 2003. The HIV Env-mediated fusion reaction. Biochim Biophys Acta **1614**:36-50.

27. **Greenberg, M., N. Cammack, M. Salgo, and L. Smiley.** 2004. HIV fusion and its inhibition in antiretroviral therapy. Rev Med Virol **14**:321-37.

28. **Hallenberger, S., V. Bosch, H. Angliker, E. Shaw, H. D. Klenk, and W. Garten.** 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. Nature **360**:358-61.

29. **Ji, H., W. Shu, F. T. Burling, S. Jiang, and M. Lu.** 1999. Inhibition of human immunodeficiency virus type 1 infectivity by the gp41 core: role of a conserved hydrophobic cavity in membrane fusion. J Virol **73**:8578-86.

30. **Kemble, G. W., T. Danieli, and J. M. White.** 1994. Lipid-anchored influenza hemagglutinin promotes hemifusion, not complete fusion. Cell **76**:383-91.

31. **Kilgore, N. R., K. Salzwedel, M. Reddick, G. P. Allaway, and C. T. Wild.** 2003. Direct evidence that C-peptide inhibitors of human immunodeficiency virus type 1 entry bind to the gp41 N-helical domain in receptor-activated viral envelope. J Virol **77**:7669-72.

32. **Kliger, Y., and Y. Shai.** 2000. Inhibition of HIV-1 entry before gp41 folds into its fusion-active conformation. J Mol Biol **295**:163-8.

33. **Louis, J. M., C. A. Bewley, and G. M. Clore.** 2001. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. J Biol Chem **276**:29485-9.

34. **Lu, M., H. Ji, and S. Shen.** 1999. Subdomain folding and biological activity of the core structure from human immunodeficiency virus type 1 gp41: implications for viral membrane fusion. J Virol **73**:4433-8.

35. **Lu, M., M. O. Stoller, S. Wang, J. Liu, M. B. Fagan, and J. H. Nunberg.** 2001. Structural and functional analysis of interhelical interactions in the human immunodeficiency virus type 1 gp41 envelope glycoprotein by alanine-scanning mutagenesis. J Virol **75**:11146-56.

36. **Markosyan, R. M., X. Ma, M. Lu, F. S. Cohen, and G. B. Melikyan.** 2002. The mechanism of inhibition of HIV-1 Env-mediated cell-cell fusion by recombinant cores of gp41 ectodomain. Virology **302**:174-84.

37. **Markovic, I., E. Leikina, M. Zhukovsky, J. Zimmerberg, and L. V. Chernomordik.** 2001. Synchronized activation and refolding of influenza hemagglutinin in multimeric fusion machines. J. Cell Biol. **155:**833-44.

38. **Nguyen, D. H., and J. E. K. Hildreth.** 2000. Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. J. Virol. **74:**3264-72.

39. **Qiao, H., R. T. Armstrong, G. B. Melikyan, F. S. Cohen, and J. M. White.** 1999. A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype. Mol Biol Cell **10:**2759-69.

40. **Root, M. J., M. S. Kay, and P. S. Kim.** 2001. Protein design of an HIV-1 entry inhibitor. Science **291:**884-8.

41. **Sanders, R. W., B. Korber, M. Lu, B. Berkhout, and J. P. Moore.** 2002. HIV Sequence Compendium, 2002. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

42. **Shu, W., J. Liu, H. Ji, L. Radigen, S. Jiang, and M. Lu.** 2000. Helical interactions in the HIV-1 gp41 core reveal structural basis for the inhibitory activity of gp41 peptides. Biochemistry **39:**1634-42.

43. **Sia, S. K., P. A. Carr, A. G. Cochran, V. N. Malashkevich, and P. S. Kim.** 2002. Short constrained peptides that inhibit HIV-1 entry. PNAS **99:**14664-9.

44. **Skehel, J. J., and D. C. Wiley.** 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem **69:**531-69.

45. **Spruce, A., A. Iwata, and W. Almers.** 1991. The first milliseconds of the pore formed by a fusogenic viral envelope protein during membrane fusion. PNAS **88:**3623-7.

46. **Tan, K., J. Liu, J. Wang, S. Shen, and M. Lu.** 1997. Atomic structure of a thermostable subdomain of HIV-1 gp41. PNAS **94:**12303-8.

47. **Wang, S., J. York, W. Shu, M. O. Stoller, J. H. Nunberg, and M. Lu.** 2002. Interhelical interactions in the gp41 core: implications for activation of HIV-1 membrane fusion. Biochemistry **41:**7283-92.

48. **Weissenhorn, W., A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley.** 1997. Atomic structure of the ectodomain from HIV-1 gp41. Nature **387:**426-30.

49. **Weng, Y., and C. D. Weiss.** 1998. Mutational analysis of residues in the coiled-coil domain of human immunodeficiency virus type 1 transmembrane protein gp41. J Virol **72:**9676-82.

50. **Weng, Y., Z. Yang, and C. D. Weiss.** 2000. Structure-function studies of the self-assembly domain of the human immunodeficiency virus type 1 transmembrane protein gp41. J Virol **74:**5368-72.

51. **Wild, C., J. W. Dubay, T. Greenwell, T. Baird, Jr., T. G. Oas, C. McDanal, E. Hunter, and T. 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. PNAS **91:**12676-80.

52. **Wild, C., T. Oas, C. McDanal, D. Bolognesi, and T. Matthews.** 1992. A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. PNAS **89:**10537-41.

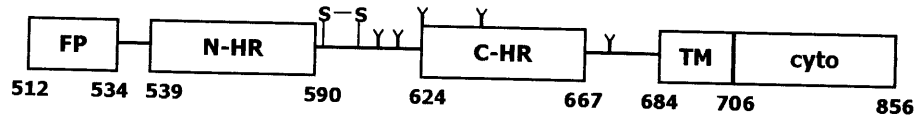
53. **Willey, R. L., J. S. Bonifacino, B. J. Potts, M. A. Martin, and R. D. Klausner.** 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. PNAS **85:**9580-4.

54. **Xu, Y., Z. Lou, Y. Liu, H. Pang, P. Tien, G. F. Gao, and Z. Rao.** 2004. Crystal structure of SARS-CoV spike protein fusion core. J. Biol. Chem.**279:**49414-9.

55. **Zavorotinskaya, T., Z. Qian, J. Franks, and L. M. Albritton.** 2004. A point mutation in the binding subunit of a retroviral envelope protein arrests virus entry at hemifusion. J. Virol. **78:**473-81.

Table 1: Summary of gp41 mutagenesis

gp41 SCHEMATIC [a]:



PHENOTYPE:

Trimer-of-hairpins stability[b]	gp41 fusogenicity	Comments	Representative mutants[c]	References
normal	normal	normal Env biosynthesis	V549A, Q551A, Q563A, Q577A Y638A, L645A, S649A, Q652A	Reviewed in Sanders 2002
normal	decreased	Abnormal Env expression, processing, or gp41-gp120 association	G572A, R579A, I642A	Lu 2001, Wang 2002
decreased	decreased	Abnormal Env expression, processing, or gp41-gp120 association	I573P W628A, W631A, I635A	Wild 1994, Chen 1993, Chen 1994 Wang 2002
decreased	decreased	normal Env biosynthesis	L565A, L568A, V570A I573A,S L556A	Cao 1993, Ji 1999, Lu 2001, Bar 2004 Lu 1999, Markosyan 2002, Wild 1994 Follis 2002, Lu 2001, Weng 1998
increased	increased	normal Env biosynthesis	Q652L	Cao 1993, Shu 2000
N.D.	decreased	normal Env biosynthesis	I573G,E,D and V570R,D,	Dubay 1992, Weng 1998, Weng 2000

[a] gp41 subdomains are labeled according to Figure 2A. Amino acids are numbered according to their position in the HXB2 strain. Also shown are N-linked glycosylation sites (Y) and the disulfide bond between the N- and C-HR domains.

[b] N.D.: not determined

[c] The mutants included here are representative of their phenotypic catagories, but the list is not exhaustive. The nomenclature indicates the original as well as the substituted residue (e.g., L565A indicates that the leucine residue at position 565 was mutated to alanine). Substitutions separated with a comma indicate separate mutations (e.g., V570R,D indicates mutation of Val-570 to Arg or Asp).

Figure 1. Viral genome and lifecycle.

(A) HIV-1 genomic organization. Expression of viral genes is driven from the long terminal repeat (LTR), a repeated motif that is also essential for reverse transcription of viral DNA and for genomic integration. The HIV Gag proteins are synthesized as a single polyprotein precursor, Pr55, that is cleaved to generate the matrix (MA), capsid (CA), and nucleocapsid (NC) proteins. Downstream of Gag is Pol, a highly conserved region of the genome that encodes the viral enzymes protease (PR), reverse transcriptase (RT), and integrase (IN). The envelope protein (Env) is synthesized from a bicistronic Vpu/Env mRNA. Env mediates viral entry into the host, while Vpu assists in viral budding and intracellular CD4 degradation. In contrast to simple retroviruses that only encode Gag, Pol, and Env, HIV contains several regulatory and accessory proteins. Tat is required for efficient transcription of viral genes, and Rev facilitates nuclear export of unspliced or singly-spliced viral RNA. The accessory proteins, Vif, Vpr, Vpu, and Nef, were initially thought to be dispensable for viral replication, but this view is changing following more recent *in vivo* experiments. Vif may either increase viral infectivity or may allow productive infections of otherwise nonpermissive cell types. Vpr functions early in the viral lifecycle, most likely in transport of the viral genome to the nucleus. Surprisingly, Vpu is only present in HIV-1 and not the closely-related HIV-2 or SIV. Vpu enhances virus budding and degrades cellular CD4. Nef also has similar functions, and appears to be required for disease induction *in vivo*.

(B) Viral lifecycle. Interaction of HIV Env with the CD4 receptor and a chemokine coreceptor (CCR5 or CXCR4) on the target cell surface triggers the first step in the viral lifecycle: fusion of viral and target cell membranes. Exposure of the virion core to the cytoplasm initiates reverse transcription of the viral RNA genome; this process occurs four to eight hours after infection. A complex of viral and cellular proteins then transports the viral nucleic acid to the nucleus, where it is stably integrated into the host genome. Expression of the provirus is driven off the viral long terminal repeat (LTR) sequence, with the help of viral accessory proteins. Translation and processing of the Env precursor occurs in the rough ER and Golgi, while other viral proteins (such as Gag and Pol) are synthesized on free ribosomes in the cytoplasm. Virion assembly occurs at the cell surface, and budding of the viral particle initially yields an immature, non-infectious virus. Viral protease then cleaves precursor structural proteins, which causes rearrangement of these proteins within the virion and yields mature, infectious virus.

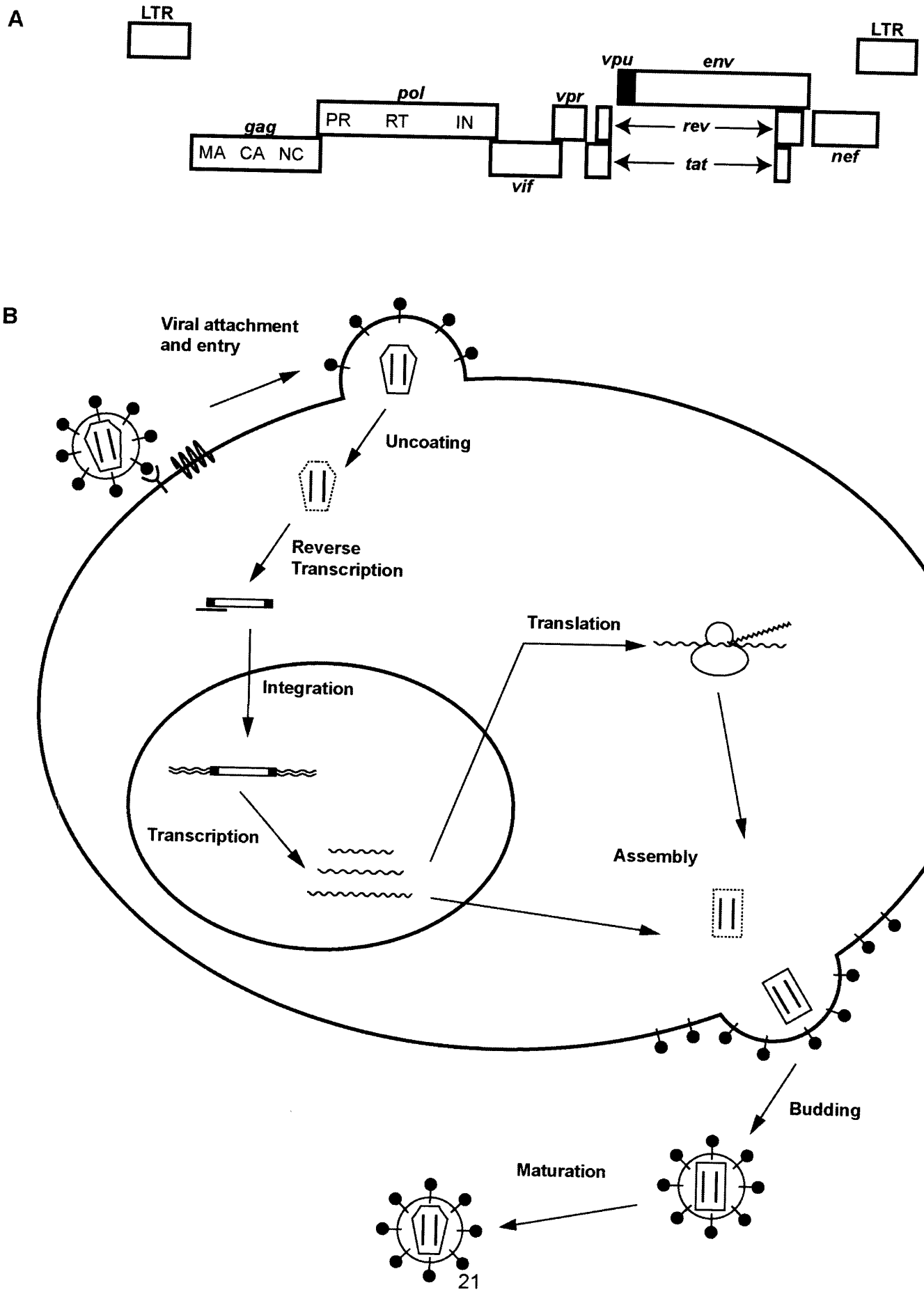

A
LTR
LTR
vpu
env
pol
vpr
gag
PR
RT
IN
rev
MA
CA
NC
tat
nef
vif
B
Viral attachment
and entry
Uncoating
Reverse
Transcription
Translation
Integration
Transcription
Assembly
Budding
Maturation

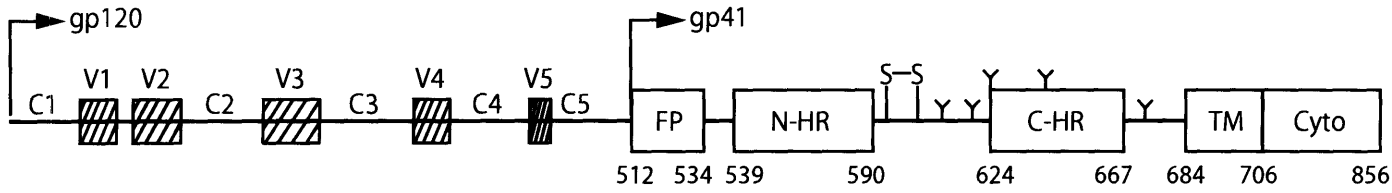
Figure 2. Env architecture and fusion model.

(A) Important subdomains within the Env glycoprotein (not drawn to scale). gp120 contains five variable domains (V1-V5) that are marked with hatched boxes and are interspersed with five constant domains (C1-C5). These domains indicate sequences that readily mutate (variable) and ones that remain relatively unchanged (constant) among diverse HIV-1 isolates. In addition, there are 24 N-linked glycosylation sites and nine disulfide bonds within gp120 (not shown). gp41 starts at residue 512 of Env, the first amino acid of the hydrophobic fusion peptide (FP). Two heptad repeat regions at the N- and C- termini of gp41 (N-HR and C-HR, respectively) are linked by a non-helical disulfide-bonded loop. The transmembrane domain (TM) anchors the Env glycoprotein to the viral membrane, and is followed by a long intraviral portion termed the cytoplasmic tail (Cyto). The five glycosylation sites in gp41 (at positions 611, 616, 624, 637, and 674) are denoted with a "Y". Amino acids are numbered according to the HXB2 strain (a well-studied lab-adapted HIV-1 strain) and are written below the schematic.

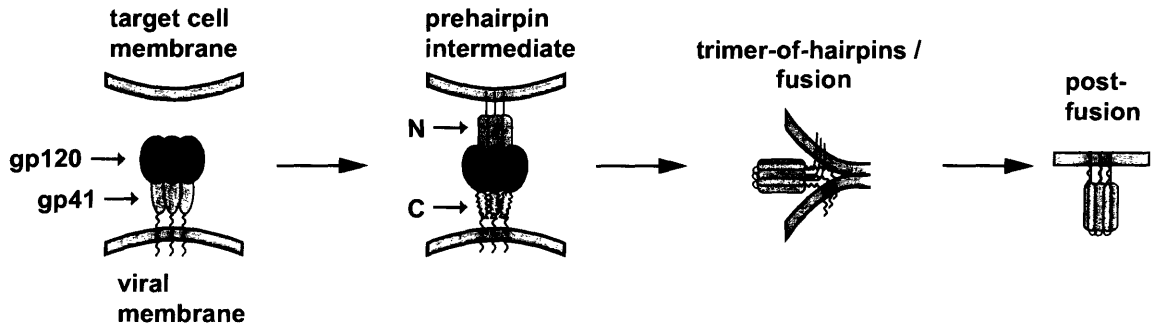
(B) Model of gp41-mediated fusion. Initial contact with cellular receptors and coreceptors is made by gp120. This interaction triggers a series of conformational changes within the Env protein to promote fusion. Specifically, the hydrophobic fusion peptide is released from the native state and inserts into target cell membrane. Thus, gp41 is anchored between the target cell (via the FP region) and the virion (via the TM domain), in a prehairpin intermediate. The trimeric N-HR region is exposed, as are the unstructured C-helices (labeled N and C, respectively). This structure resolves into a trimer-of-hairpins, in which the C-helices bind the grooves in the central N-terminal trimer. This transition juxtaposes the viral and cellular membranes. In a poorly understood process—perhaps involving interaction of multiple trimer-of-hairpin structures, or mixing of the outer leaflets of each membrane—the two membranes fuse and viral contents are released into the host cell.

(C) Structure of the HIV-1 trimer of hairpins. gp41 ectodomain fragments from the N- and C-HR domains assemble in a six-helix bundle structure. At the center are three N-peptides (blue helices). The grooves between adjacent N-peptides form binding sites for the C-peptides (yellow helices). These C-helices pack on the outside of the N-terminal trimer in an antiparallel manner. The location of the FP (blue spheres) and TM (red spheres) regions are indicated. Reprinted with permission from ref. 14.

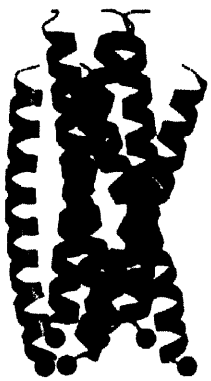
A


gp120
gp41
C1
V1
V2
C2
V3
C3
V4
C4
V5
C5
FP
N-HR
S—S
C-HR
TM
Cyto
512
534
539
590
624
667
684
706
856

B


target cell membrane
prehairpin intermediate
trimer-of-hairpins / fusion
post-fusion
gp120
gp41
N
C
viral membrane

C



CHAPTER 2

GENERATION OF A STABLE gp41 ECTODOMAIN HYBRID PEPTIDE FOR USE AS A FUSION INHIBITOR AND A VACCINE CANDIDATE

With the following exceptions, all work described in this chapter was performed by TRS. Heng Chhay did the viral infectivity assays, and Mike Burgess and Ben Sanford chemically synthesized all peptides. In addition, peptide injections into guinea pigs, as well as neutralization and ELISA assays, were performed by Michael Citron, Xiaoping Liang, and associates, at Merck Research Labs in West Point, PA.

Abstract

We characterized a region of the gp41 N-HR trimer located between the fusion peptide and the hydrophobic cavity. This 26 residue gp41 fragment, which alone forms aggregates, was stabilized when fused to a trimeric scaffold peptide, GCN4-pIqI. The hybrid molecule, IQN26, was helical by circular dichroism, exhibited a cooperative melting curve, and was mostly trimeric with slight aggregation. IQN26 also inhibited fusion at a low nanomolar concentration. Circular dichroism studies showed that IQN26 interacts with a C-HR derived peptide, C29. We hypothesized that IQN26 may function as an inhibitor by sequestering the corresponding C-HR region of gp41. This hypothesis was tested by designing mutations in the N-HR region of IQN26 that abolish its interaction with C29. As expected, the mutant IQN26 molecule no longer interacted with C29 by circular dichroism and was unable to inhibit fusion.

In order to generate a more potent inhibitor of fusion, the same N26 region of gp41 was fused to a more stable 'isoleucine zipper' scaffold protein, IZ. This resulting molecule, IZN26, was indeed more helical, stable, and trimeric than IQN26. Surprisingly however, IZN26 showed little interaction with C29 and possessed no inhibitory activity up to 10μM. The basis for this lack of inhibitory activity remains unknown.

Finally, to test their efficacy as vaccine candidates, IQN26 and IZN26 were injected into guinea pigs. Preliminary high throughput screens did not detect any neutralizing activity in the sera.

In the late 1980s and early 1990s computer modeling of gp41 suggested that it contained an amphipathic α-helical region (13, 20). Synthetic peptides containing these sequences were shown to be helical and capable of inhibiting viral fusion. Secondary structure seemed essential for this property (36). Subsequent gp41 proteolysis experiments in the mid 1990s revealed the presence of stable, interacting N-HR and C-HR regions (26, 27). The minimum stable gp41 ectodomain that was identified in these experiments contained 36 residues from the N-HR region (N36) and 34 residues from the C-HR (C34). Shortly thereafter, in 1997, the first atomic resolution structures of gp41 were solved (8, 33, 34).

In general, peptides derived from the gp41 N- and C-HR regions are thought to inhibit fusion by preventing gp41 trimer-of-hairpins formation. Interestingly, C-peptides are more potent fusion inhibitors than are N-peptides. Several lines of evidence support the hypothesis that these peptides bind the N-HR region of gp41 and thus prevent trimer-of-hairpins formation. First, the antiviral activity of a potent C-peptide can be titrated by addition of exogenous N-peptide (26). Second, viruses that develop resistance to C-peptides often contain mutations that map to the N-HR of gp41 (1, 2, 31). Third, C-peptides that are mutated in the N-HR-binding region are less efficient inhibitors of fusion (7, 35).

Peptides derived from the N-HR region of gp41 are also capable of preventing trimer-of-hairpins formation. Due to their tendency to aggregate, these peptides are often solubilized and stabilized by fusion to stable trimeric peptides. There are at least two possible mechanisms for N-peptide inhibition of fusion. Some reports using stabilized N-peptide cores suggest that they can, in their trimeric conformation, bind and sequester the C-HR region of gp41 and thus prevent trimer-of-hairpins formation (15, 24). Alternatively, other evidence indicates that N-peptides may intercalate into the N-HR region of gp41, again disrupting the native trimer-of-hairpins structure (3).

This study describes a novel N-HR-derived peptide inhibitor of fusion. Unlike other N-peptides mentioned above, we sought to characterize a region of the N-HR that does not contain a deep hydrophobic cavity shown to be an essential binding site for several C-HR amino acids (7). The 26 N-HR residues studied here are flanked by the fusion peptide at the N-terminus and the hydrophobic pocket at the C-terminus. We aimed to determine whether the cavity is necessary to maintain folded structure, whether peptides lacking this region contain important determinants of C-HR binding, and whether these peptides are also potent HIV inhibitors. Using a chimeric

peptide containing this region fused to a scaffold peptide, we first determined that folding of the gp41 core is possible in the absence of the pocket. Our hybrid peptide displays an IC_{50} value in the low nanomolar range. Further biophysical experiments suggest that this molecule inhibits fusion by sequestration of the C-HR. Thus, this region of the N-HR does contain important determinants of C-peptide binding. Finally, in an attempt to generate a novel vaccine candidate, the serum of guinea pigs immunized with these chimeric peptides was analyzed for a neutralizing response. None was found.

Materials and methods

Plasmid construction—A 29 residue fragment from the N-HR region of HIV-1 gp41 (residues 539-567 of HIV-1 strain HXB2) was fused to a GCN4-pIqI trimeric scaffold protein; this molecule, IQN29, was then expressed in *E.coli* for purification. A longer 51 residue N-HR fragment (residues 539-590 of HXB2) was initially PCR amplified and cloned into the p-GCN4pIqIn vector (gift of Mike Milhollen) as a Sph1/BamH1 fragment. A stop codon after residue 567 was introduced by site directed mutagenesis (QuickChange Site-Directed Mutagenesis, Stratagene). The resulting vector, pGCN4-N29, encodes a 60 amino acid chimeric peptide containing an N-terminal GCN4-pIqI fragment fused to 29 residues of the N-HR from the fusion peptide up to and including the first three residues of the hydrophobic cavity. The sequence of this peptide, IQN29, is: NH_2-**MRMKQIEKKIEEIEKKQYKIENEIARIKKLI** VQARQLLSGIVQQQNNLLRAIEAQQHLLQ-COOH where the boldface residues are from the GCN4-pIqI moiety. Additional, tagged versions of IQN29 were also generated. PCR-amplified IQN29 was subcloned into the pET-16b vector (Novagen), introducing a 6X-histidine tag (HIS-tag) followed by a Factor Xa protease cleavage site at the N-terminus of IQN29 (yielding a protein denoted HIS-IQN29). Two maltose-binding protein (MBP) fusion constructs were generated by subcloning IQN29 into the pMAL-c2G and pMAL-p2G vectors (New England BioLabs); the pMAL-c2G vector targets the fusion protein to the bacterial cytoplasm while the pMAL-p2G vector targets it to the periplasmic space. These fusion proteins are called c- or p-MBP-IQN29 and, in both cases, the MBP portion can be cleaved off with the enzyme genenase. All constructs were fully sequenced.

Bacterial protein expression and purification—Protein expression was induced in late log phase cultures of BL21pLysS *E.coli*. One liter cultures were harvested after induction with 1mM IPTG for 2 to 4 hours at 37°C. Numerous attempts were made to purify untagged IQN29 from bacterial

cultures. Cell pellets were resuspended in 50mM Tris pH 8.5, 15% glycerol with or without 1% nonidet P-40 and 1mM EDTA, and sonicated to homogeneity. After two cycles of centrifugation (35,000 x g, 30 minutes) and sonication, IQN29 remained mostly in the soluble fraction. This fraction was then purified over anion (DEAE) or cation (CM sepharose) exchange columns. In both cases, IQN29 was present in the flow-through fraction; we were unable to find conditions in which IQN29 bound the column resin at several different pHs. Noting that solubility of IQN29 changed dramatically with pH, we performed several dialyses into buffers of varying pHs. Ammonium sulfate precipitation as well as gel filtration were also employed as tools for further purification. Using these methods, we were unable to obtain pure fractions of IQN29.

Subsequent purification attempts were made with IQN29 containing either a HIS- or a MBP tag. Bacterial cultures expressing HIS-IQN29 were lysed as described above, and the soluble fraction was purified over a Ni-NTA (Qiagen) column. HIS-IQN29 eluted in 10 mM Tris pH 8 and 50 mM sodium phosphate containing 250 mM imidazole. However, digestion with the factor Xa protease did not remove the HIS-tag from the N-terminus of IQN29. The pMBP-IQN29 protein was expressed at very low amounts in the periplasm, so no further purification attempts were made. Cytosolic cMPB-IQN29, on the other hand, was well expressed in BL21pLysS cells. Upon sonication and centrifugation, as described above, the soluble fraction was separated over an amylose column and cMBP-IQN29 was eluted with 10 mM maltose. Using the genenase protease, we next attempted to cleave off the N-terminal MBP fragment. After a 19 hour digest, no protease cleavage was detected. We suspect that the presence of the strongly helical IQN29 moiety near the factor Xa or genenase cleavage sites may impart some secondary structure to the cleavage site, making it unrecognizable to the protease.

Since none of the above strategies yielded sufficiently pure, untagged IQN29, we decided to obtain chemically synthesized peptides for future experiments.

Peptide Synthesis and Purification—All peptide synthesis work was performed by Mike Burgess and Ben Sanford in the Kim laboratory at MIT. HPLC purification of all peptides was done by TRS.

Peptides were synthesized on a PE Biosystems 431A peptide synthesizer using standard Fmoc chemistry (19) and then desalted over a Sephadex G-25 column (Pharmacia). All peptides are acetylated at the N-terminus and amidated at the C-terminus. Final purification of all peptides

was performed by reverse-phase HPLC (Waters) over as C18 preparative column (Vydac) using a linear gradient of acetonitrile and 0.1% trifluoroacetic acid. Peptides were then lyophilized and stored under desiccation. Peptide identity was confirmed by matrix-assisted laser desorption ionization mass spectrometry. The sequences of all synthesized peptides are shown in Figure 1*C*. A three amino acid sequence of Tyr-Gly-Gly was added to the N-terminus of N26, IZ, and IZN26 since these three peptides have no chromophoric residues and therefore no way to accurately determine stock concentrations.

Circular Dichroism—Lyophilized peptides were resuspended in water and concentrations were determined by absorbance at 280 nm in 6 M guanidine hydrochloride as described (18). Samples were then diluted to 1 to 40 μM in PBS (50 mM sodium phosphate pH 7.0, 150 mM sodium chloride).

All circular dichroism (CD) measurements were taken on an Aviv 62 or 62A DS spectrometer (Aviv Associates). All wavelength scans were performed with 10 μM samples, and all thermal denaturation scans with 1 μM samples, unless otherwise indicated. Theoretical 100% helicity was calculated according to the method of Chen et al. (10), and was estimated to be $-35{,}400$ deg cm^2 dmol^{-1} for IQN26, IZN26, and mIQN26(3A). Wavelength scans were performed at 25°C, in a 1 mm cuvette, from 200 to 300 nm with a 5 second averaging time, in triplicate. Thermal denaturation scans were performed in a 1 cm cuvette in 2°C increments starting at 0°C, with a 1.5 minute equilibration time and a 30 second averaging time.

Peptide mixing experiments were performed using a 1 cm dual compartment mixing cuvette that contains a barrier between two chambers. In all experiments, 20 μM of the N-peptide (IQN26, IZN26, mIQN26, or N26) was mixed with 22 μM of C29. The 'before mixing' scans were obtained when the contents of the chambers remained separate. Parafilm was then placed over the cuvette opening and the cuvette was inverted several times in order to thoroughly mix the contents. After several minutes to hours, the mixed sample was then scanned and recorded as the 'after mixing' sample. The parameters for these wavelength scans are the same as those described above.

Sedimentation Equilibrium—All peptide samples were centrifuged at 25°C in a Beckman XL-A analytical ultracentrifuge (Beckman) with a An-60 Ti rotor. Peptides were dialyzed overnight

against PBS, diluted to three concentrations ranging from 20 to 80 μM, and centrifuged at rotor speeds of 15, 20, and 25 krpm. Data were collected at 230, 235, and 280 nm.

HIV Infectivity Assay—TRS provided peptide dilutions, and Heng Chhay (Kim Laboratory, MIT) performed all viral infectivity assays. Replication-incompetent virus was made as described (9). Specifically, 293T cells were cotransfected with a packaging vector containing all viral structural genes except Env and Nef (Nef was replaced with the luciferase gene), and an expression vector encoding HXB2 gp160. This virus was used to infect HOS-CD4/fusin cells (N. Landau, NIH AIDS Reagent Program) in the presence of inhibitory peptides. Cells were lysed two days post-infection and luciferase activity was measured in a Wallac AutoLumat luminometer. The peptide concentration that results in 50% inhibition of viral infection (IC_{50}) was calculated by fitting the data to a Langmuir equation [$y=k/(1+[peptide]/IC_{50})$] where y is the luciferase activity and k is a scaling constant.

Cell-cell Fusion Assay—Similar to the viral infectivity assay, the ability of gp41-derived peptides to inhibit cell-cell fusion was determined in a syncytia assay. Chinese hamster ovary (CHO) cells expressing gp160, Tat, and Rev (23) and HeLa-LTR-β-gal cells expressing CD4 (M. Emerman, NIH AIDS Reagent Program) were used. Fusion of these cell types to form syncytia (multinucleate cells) allows Tat-mediated expression of β-gal from the viral LTR promoter. Approximately $2x10^4$ CHO cells and $4x10^4$ HeLa cells were mixed in 8-well slides. After 18-20 hours of coculture, cells were fixed with 0.5% glutaraldehyde and stained with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). Syncytia containing three or more nuclei were counted under bright field microscopy.

Animal Immunizations—IQN26 and IZN26 were sent to Merck Research Labs for injection into guinea pigs and characterization of the sera. This work was done by Michael Citron, Xiaoping Liang, and colleagues. Female guinea pigs were injected intramuscularly three times (t = 0, 4, 8 weeks), with 100 μg peptide and 40 μg QS21 adjuvant per injection. Serum was sampled at weeks 3, 8, and 12, and tested in ELISA and neutralization assays.

Results and discussion

These studies aim to characterize a 26 amino acid N-HR segment of gp41, which starts five residues after the fusion peptide (12) and spans virtually the entire region between the fusion

peptide and the hydrophobic cavity (the sequence of this region is shown in Figure 1). This fragment is denoted N26, and is composed of gp41 residues 539-564. The N26 region has been crystallized in the context of a longer N-helix segment (34) and is shown to be helical and trimeric. Circular dichroism analysis of the N26 peptide alone, however, shows that it lacks secondary structure (Table 1 and Figure 2*A*) and therefore does not properly represent the N-HR coiled coil observed in gp41 x-ray structures. Furthermore, unlike many N- and C-HR-derived peptides, N26 does not inhibit fusion (Table 1).

Most isolated N-HR peptides (those in the literature contain the hydrophobic cavity) exhibit little helicity and tend to aggregate (15, 26, 27). This property can be overcome by mixing N-HR peptides with their C-peptide binding partners; the complex often adopts a more helical and stable conformation, as seen with both HIV and SIV (4, 26). Alternatively, N-peptides can be fused to heterologous peptides that themselves form stable trimers; these scaffolds impart structural stability to the N-HR region (15, 16, 24, 34). Mixing N26 with its C peptide binding partner failed to increase the helicity of either N26 or the C peptide (Figure 2*A*). This lack of secondary structure, even in the presence of its binding partner, suggests that N26 alone may be too aggregated to form complexes, or that this region of the N-HR requires neighboring residues in order to make the necessary C-peptide contacts. Thus, we decided to utilize the second approach in order to better study this region in the context of its native conformation.

To improve the properties of the N-HR peptide, we fused N26 to GCN4-pIqI, a stable and trimeric peptide (15-17); the resulting chimeric molecule is termed IQN26. GCN4-pIqI is derived from the GCN4 coiled-coil model, in which the oligomerization state of coiled coils has been changed by mutation of residues buried at the helical interface (17, 22, 30). Figure 1 shows the arrangement of gp41 N-HR and C-HR residues on a helical wheel, as well as the sequence of IQN26 and other chimeric peptides used in this study.

In contrast to the N26 peptide, circular dichroism experiments reveal that IQN26 forms a stable, helical structure and has a melting temperature of 70°C (Table 1). Analytical ultracentrifugation experiments suggest that the peptide is a slightly aggregated trimer in solution. The overall stable secondary structure of IQN26 indicates that the cavity-containing portion of gp41 is not required for N-HR folding.

Next, we performed cell-cell fusion assays and viral infectivity assays to determine whether IQN26 is capable of inhibiting fusion. In these assays, N26 alone fails to demonstrate any inhibition potential (Table 1), most likely because of its aggregated or non-native structure. IQN26, on the other hand, is indeed a potent inhibitor, exhibiting IC_{50} values of 70 nM and 40 nM in each assay, respectively (Table 1). Similar chimeric molecules containing the cavity region also inhibit fusion (in the ~200 nM IC_{50} range) while the GCN4-pIqI moiety alone has no inhibitory activity (15).

To distinguish between two possible mechanisms for IQN26 inhibition—sequestration of the C-HR or intercalation into the N-HR—we assessed its ability to bind to a C-peptide, mimicking the intramolecular gp41 interaction. Since most C peptides studied to date were designed to interact with the pocket-containing N-helix region, we used the SIV and HIV solution and crystal structures (5, 6, 8, 26, 28, 34) to align N26 with the C-HR region. The resulting C-peptide, C29, closely overlaps the N26 region when visualized in the context of the six-helix bundle, and alone lacks secondary structure (Figure 1 and Table 1). We used a circular dichroism peptide mixing assay to measure the interaction between the N- and C-HR peptides. CD scans were obtained before and after mixing 22 μM IQN26 and 20 μM C29 solutions in a dual chamber mixing cuvette. IQN26 and C29 appear to interact, as evidenced by a considerable increase in helical signal upon mixing (Figure 2*B*). In contrast, both N26 and C29 remain unstructured during mixing (Figure 2*A*), lending further support to the hypothesis that N26 is primarily aggregated in PBS. These initial results suggest that IQN26 may inhibit by binding and sequestering the C-HR region of gp41; however, these studies alone do not rule out the intercalation model.

To further test the mechanism of IQN26 inhibition, we made mutations in IQN26 that were designed to abolish C-peptide interaction. Residues Leu-544, Gln-551, and Gln-563 at the **e** and **g** positions of N26 (Figure 1, *B* and *C*) were chosen based on their nearest-neighbor contacts (using the Contact and Insight II software programs ((11) and Accelrys, San Diego)) and the reported crystal structures. A mutant molecule, termed mIQN26(3A), contained mutations of the above three residues to alanine. Biophysical experiments confirm that mIQN26(3A) has properties similar to the parent IQN26 molecule: 67% helicity and a 70°C thermal denaturation midpoint. In syncytia assays, mIQN26(3A) shows no fusion inhibition at concentrations up to 10 μM (Table 1). As expected based on its design, mIQN26(3A) does not interact with C29 by CD mixing (Figure 2*C*), as seen by the overlapping spectra before and after mixing. These results

show that abolishing interaction with the C-peptide debilitates N-peptide inhibitory activity and strongly suggest that IQN26 inhibits fusion by sequestering the C-HR region.

In addition to GCN4-pIqI, we used a designed trimeric peptide (isoleucine zipper or IZ) as a scaffold for N26. The IZ scaffold used here is modified from a synthetic, designed coiled coil that contains charged residues on the outer face (at the **e** and **g** positions, see Figure 1 for description), thus allowing electrostatic interactions between adjacent helices (32). It has been previously used to stabilize cavity-containing N-HR peptides (15). IZN26 exhibits an even more stable structure than IQN26, with a melting temperature of 82°C. Furthermore, IZN26 does not display any signs of aggregation up to 60 μM (Table 1). Based on this increased stability, we hypothesized that IZN26 would interact more strongly with C29 and thus be a more potent fusion inhibitor. Surprisingly, this hypothesis proved incorrect. IZN26 shows little interaction with C29; unlike the 41 mdeg increase in helical signal upon IQN26 and C29 mixing, mixing IZN26 and C29 increases the helicity by only 15 mdeg (Figure 2, *B* and *D*). Since the GCN4-pIqI portion does not inhibit fusion (15), it is unlikely to interact with C29; however, this control was not performed in these CD mixing studies. Furthermore, IZN26 did not inhibit cell-cell fusion up to the highest concentration tested (10 μM) (similar studies with the IZ portion indicate that it alone does not inhibit fusion (15)).

The reason for the difference between IQN26 and IZN26 in terms of C-peptide binding and inhibitory properties remains unclear. Perhaps the N26 region in a tightly-associated coiled coil (i.e. IZN26) is unable to efficiently interact with the C-helix. IQN26, in a more dissociated form, may still be capable of this interaction. While this explanation could be valid for the N26 region, other work with cavity-containing N-HR peptides has shown a different result. Simultaneous experiments were performed in the Kim lab with the N-HR hydrophobic cavity residues fused to GCN4-pIqI (IQN17) or IZ (IZN17) (15). These studies show that stability is correlated to inhibitory activity; IZN17 is the more stable of the two molecules, and possesses a tenfold greater inhibitory potency. Furthermore, since IZN17 is a potent inhibitor, it is unlikely that the highly charged IZ moiety in IZN26 interacts unfavorably with gp41. It is unclear why this correlation does not hold true for non-cavity N-HR peptides. Cumulatively, these results suggest that folding and stability are not the only determinants of inhibitory potential.

It is indeed possible that in the pre-hairpin intermediate, the cavity region is a tightly-associated trimer that then splays apart toward the fusion peptide. Synthetic fusion peptides have

been studied in isolation, a difficult task given their extremely low solubility. Evidence indicates that the HA fusion peptide inserts obliquely into the outer leaflet of the target membrane and consists of two α-helices that are interrupted by a kinked region (14). While the structure of a trimeric fusion peptide molecule is unknown, it is possible that each monomeric fusion peptide, and perhaps some neighboring N-HR residues, may splay apart and insert at the same oblique angle into the target membrane. The resulting difference in structural properties could lead to a cavity-containing peptide being a better inhibitor as a stable trimer, and a region closer to the fusion peptide being more potent in a slightly dissociated form.

The CD mixing experiments contain some caveats, leading to the possibility that this method does not accurately measure the level of N- and C-peptide interaction. First, the increase in helicity upon IQN26 and C29 mixing could be a result of both the C-peptide and IQN26 adopting a more helical conformation. Since IZN26 is already more helical and stable than IQN26, perhaps there is no concomitant increase in IZN26 helicity when it is mixed with C29. By extension, the difference in C29 interaction between the IQN26 and mIQN26(3A) (Figure 2, *B* and *C*) may be more accurately reflected in the CD mixing assay since both N-HR chimeric peptides are equally stable and helical at the outset. A second, minor caveat is that we have not ruled out the possibility that the GCN4-pIqI moiety of IQN26, and not the N-HR region or the IZ scaffold, is interacting with C29. This is highly unlikely since the GCN4-pIqI moiety would then presumably demonstrate some inhibitory activity.

Upon closer biophysical examination of the chimeric N-peptides, we determined that all three N-HR peptides in this study display a concentration-dependent helical signal (Figure 3). As they are diluted from 40 or 10 μM to 1 μM, these peptides lose helical signal. Our model of IQN26 inhibition suggests that it interacts with C29 as a helical trimer and that this interaction is correlated with its inhibitory potency. Other studies also suggest that N-peptides inhibit in their helical, trimeric conformation (15, 24, 25) and their potency is often improved with greater helicity. Due to the limits of detection in CD, we cannot perform biophysical studies at low nanomolar (IC_{50}) concentrations. However, the CD dilution experiments indicate that the helical N-trimers dissociate at low micromolar concentrations, leading to the possibility that at low nanomolar (IC_{50} range) concentrations these peptides are in fact monomeric. In addition, these results suggest that dilution of IQN26 from 22 μM to 11 μM in the C29 mixing experiments may slightly decrease IQN26 helicity and may be another potential explanation why the post-mixing spectrum is not as helical as IQN26 before mixing. Future experiments to measure the interaction

between chimeric N-HR peptides and C29 would include thermal denaturation and analytical ultracentrifugation of peptide complexes.

Finally, the trimeric N26 region of gp41 is thought to be exposed as a transient intermediate in the fusion pathway. While it is possible that antibodies directed against this region may prevent gp41 trimer-of-hairpins formation, no such antibody has yet been identified. This could be due to the transient nature of the prehairpin intermediate, or to the possibility that binding of antibodies to this region may be sterically hindered (21).

In order to examine their potential as vaccine candidates, our chimeric N-HR peptides were used as immunogens to elicit an antibody response. Since both peptides were well-characterized biophysically, and the structure of the N-HR region in the prehairpin intermediate is unknown, both the weaker (IQN26) and the stronger (IZN26) N-HR trimers were used for these experiments. The peptides were injected intramuscularly into guinea pigs and the sera were analyzed using ELISA and neutralization assays (Merck Research Labs). Neither peptide had neutralizing activity in these preliminary screens, and therefore no further immunization studies were conducted.

A detailed understanding of gp41 conformational changes, by biophysical, mutagenic, and crystallographic analysis, will aid in the design of peptide immunogens and inhibitors. For example, if future work shows that the N26 region is indeed splayed apart, immunogens can be designed to weakly trimerize these residues. Perhaps a scaffold that is more weakly trimeric than GCN4-pIqI may be tried.

An important future experiment should switch the location of the scaffold in the chimeric peptide. The N-terminus of N26 is close to the fusion peptide and the target membrane, whereas the C-terminus is adjacent to the strongly trimeric hydrophobic cavity. In the present study, the trimeric IQ or IZ moiety is anchored to the N-terminus of N26, adjacent to the target membrane. A molecule with IQ or IZ at the C-terminus, essentially replacing the trimeric hydrophobic cavity, may serve as a better gp41 mimic. Additionally, design of such a molecule would reduce the chance of the scaffold interacting unfavorably with the target phospholipid bilayer.

gp41 is highly immunogenic in HIV infected patients, and sera from these patients react with peptides that mimic all regions of the protein (such as the fusion peptide, N-HR, C-HR, and

the six-helix bundle structure) (29). However, it is unclear when reactive antibodies actually bind the gp41 protein *in vivo*. They could bind gp41 when Env is in its native conformation, before fusion has been triggered. Alternatively, antibody binding could occur during the fusion process when intermediate structures are transiently exposed, or after fusion is complete and the Env protein lies exposed on the surface of the newly infected cell. Finally, antibody binding could also be induced during premature Env triggering when there is no target cell present. The main questions in vaccine design that remain to be answered are (1) Can antibodies access epitopes that are exposed between the virus and target cell (i.e., in the prehairpin intermediate), or are they sterically inhibited? (2) What are the key neutralizing epitopes? (3) How specific does the antibody-epitope interaction have to be (i.e., can an antibody bind anywhere on the N-HR or does it have to bind specifically to the cavity, for example, in order to prevent N-C interaction)? and (4) How high does the neutralizing titer in the sera have to be in order to efficiently prevent infection? These are areas of intense investigation, both in the basic science realm and in clinical trials. Chapter 4 further explores advances in vaccine research and other approaches being taken to prevent viral entry.

References

1. **Armand-Ugon, M., A. Gutierrez, B. Clotet, and J. A. Este.** 2003. HIV-1 resistance to the gp41-dependent fusion inhibitor C-34. Antiviral Research **59:**137-42.
2. **Baldwin, C. E., R. W. Sanders, Y. Deng, S. Jurriaans, J. M. Lange, M. Lu, and B. Berkhout.** 2004. Emergence of a drug-dependent human immunodeficiency virus type 1 variant during therapy with the T20 fusion inhibitor. J. Virol. **78:**12428-37.
3. **Bewley, C. A., J. M. Louis, R. Ghirlando, and G. M. Clore.** 2002. Design of a novel peptide inhibitor of HIV fusion that disrupts the internal trimeric coiled-coil of gp41. J Biol Chem **277:**14238-45.
4. **Blacklow, S. C., M. Lu, and P. S. Kim.** 1995. A trimeric subdomain of the simian immunodeficiency virus envelope glycoprotein. Biochemistry **34:**14955-62.
5. **Caffrey, M., M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, D. G. Covell, A. M. Gronenborn, and G. M. Clore.** 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. Embo J **17:**4572-84.
6. **Caffrey, M., M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, A. M. Gronenborn, and G. M. Clore.** 1997. Determination of the secondary structure and global topology of the 44 kDa ectodomain of gp41 of the simian immunodeficiency virus by multidimensional nuclear magnetic resonance spectroscopy. J Mol Biol **271:**819-26.
7. **Chan, D. C., C. T. Chutkowski, and P. S. Kim.** 1998. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. PNAS **95:**15613-7.
8. **Chan, D. C., D. Fass, J. M. Berger, and P. S. Kim.** 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell **89:**263-73.
9. **Chen, B. K., K. Saksela, R. Andino, and D. Baltimore.** 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:**654-60.
10. **Chen, Y. H., J. T. Yang, and K. H. Chau.** 1974. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. Biochemistry **13:**3350-9.
11. **Collaborative Computational Project.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr **50:**760-3.
12. **Delahunty, M. D., I. Rhee, E. O. Freed, and J. S. Bonifacino.** 1996. Mutational analysis of the fusion peptide of the human immunodeficiency virus type 1: identification of critical glycine residues. Virology **218:**94-102.

13. **Delwart, E. L., G. Mosialos, and T. Gilmore.** 1990. Retroviral envelope glycoproteins contain a "leucine zipper"-like repeat. AIDS Res Hum Retroviruses **6:**703-6.

14. **Earp, L. J., S. E. Delos, H. E. Park, and J. M. White.** 2005. The many mechanisms of viral membrane fusion proteins. Curr Top Microbiol Immunol **285:**25-66.

15. **Eckert, D. M., and P. S. Kim.** 2001. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. PNAS **98:**11187-92.

16. **Eckert, D. M., V. N. Malashkevich, L. H. Hong, P. A. Carr, and P. S. Kim.** 1999. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell **99:**103-15.

17. **Eckert, D. M., V. N. Malashkevich, and P. S. Kim.** 1998. Crystal structure of GCN4-pIQI, a trimeric coiled coil with buried polar residues. J Mol Biol **284:**859-65.

18. **Edelhoch, H.** 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry **6:**1948-54.

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

20. **Gallaher, W. R., J. M. Ball, R. F. Garry, M. C. Griffin, and R. C. Montelaro.** 1989. A general model for the transmembrane proteins of HIV and other retroviruses. AIDS Res Hum Retroviruses **5:**431-40.

21. **Hamburger, A. E., S. Kim, B. D. Welch, and M. S. Kay.** 2005. Steric accessibility of the HIV-1 gp41 N-trimer region. J. Biol. Chem. **280:**12567-72.

22. **Harbury, P. B., P. S. Kim, and T. Alber.** 1994. Crystal structure of an isoleucine-zipper trimer. Nature **371:**80-3.

23. **Kozarsky, K., M. Penman, L. Basiripour, W. Haseltine, J. Sodroski, and M. Krieger.** 1989. Glycosylation and processing of the human immunodeficiency virus type 1 envelope protein. J Acquir Immune Defic Syndr **2:**163-9.

24. **Louis, J. M., C. A. Bewley, and G. M. Clore.** 2001. Design and properties of N(CCG)-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. J Biol Chem **276:**29485-9.

25. **Louis, J. M., I. Nesheiwat, L. Chang, G. M. Clore, and C. A. Bewley.** 2003. Covalent trimers of the internal N-terminal trimeric coiled-coil of gp41 and antibodies directed against them are potent inhibitors of HIV envelope-mediated cell fusion. J Biol Chem **278:**20278-85.

26. **Lu, M., S. C. Blacklow, and P. S. Kim.** 1995. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nat Struct Biol **2:**1075-82.

27. **Lu, M., and P. S. Kim.** 1997. A trimeric structural subdomain of the HIV-1 transmembrane glycoprotein. J Biomol Struct Dyn **15:**465-71.

28. **Malashkevich, V. N., D. C. Chan, C. T. Chutkowski, and 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. PNAS **95:**9134-9.

29. **Opalka, D., A. Pessi, E. Bianchi, G. Ciliberto, W. Schleif, M. McElhaugh, R. Danzeisen, R. Geleziunas, M. Miller, and D. M. Eckert.** 2004. Analysis of the HIV-1 gp41 specific immune response using a multiplexed antibody detection assay. J Immunol Methods **287:**49-65.

30. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science **243:**538-42.

31. **Rimsky, L. T., D. C. Shugars, and T. J. Matthews.** 1998. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. J Virol **72:**986-93.

32. **Suzuki, K., H. Hiroaki, D. Kohda, and T. Tanaka.** 1998. An isoleucine zipper peptide forms a native-like triple stranded coiled coil in solution. Protein Eng 11:1051-5.

33. **Tan, K., J. Liu, J. Wang, S. Shen, and M. Lu.** 1997. Atomic structure of a thermostable subdomain of HIV-1 gp41. PNAS **94:**12303-8.

34. **Weissenhorn, W., A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley.** 1997. Atomic structure of the ectodomain from HIV-1 gp41. Nature **387:**426-30.

35. **Wild, C., T. Greenwell, D. Shugars, L. Rimsky-Clarke, and 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:**323-5.

36. **Wild, C., T. Oas, C. McDanal, D. Bolognesi, and T. Matthews.** 1992. A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. PNAS **89:**10537-41.

Table 1: Biophysical properties and inhibitory activity of N26 chimeric peptides

Peptide	CD Wavelength Scan [a]		Thermal Denaturation [b]		Analytical Ultracentrifugation [b]			Fusion Inhibition [b,e]	
	θ_{222} (deg cm^2 dmol^{-1})	% helicity	T_m (°C)	Properties	Concentration	$MW_{obs/calc}$ [c]	Residuals [d]	Cell-cell IC_{50}	Virus-cell IC_{50}
N26	unstructured	--	ND	--	ND	--	--	> 10 μM	> 10 μM
IQN26	-24,500	69%	70°C	cooperative, reversible	20 μM 50 μM 70 μM	3.25 3.54 3.65	slight aggregation slight aggregation slight aggregation	70 nM	40 nM
mIQN26(3A)	-23,700	67%	70°C	cooperative, reversible	ND	--	--	> 10 μM	ND
C29	unstructured	--	ND	--	ND	--	--	>0.63 μM	ND
IZN26	-32,200	91%	82°C	cooperative, reversible	20 μM 40 μM 60 μM	2.87 3.10 3.02	random random random	> 10 μM	ND

[a] The CD signal at 222 nm (θ_{222}) and the percent helicity are reported for 10 μM peptide solutions in PBS. The theoretical 100% helicity value for IQN26, mIQN26(3A), and IZN26 is -35,400 deg cm^2 dmol^{-1}.

[b] ND: experiment not done.

[c] Sedimentation equilibrium results are shown as the ratio of observed molecular weight (M_{obs}) to calculated molecular weight for a monomer (M_{calc}). An ideal trimer has an M_{obs}/M_{calc} ratio of 3.00.

[d] For IQN26, some non-random residuals were observed for all three concentrations of peptide. In those cases, the upward curvature of the residuals indicated slight aggregation of the peptide.

[e] Syncytia assays and viral infectivity assays were used to assess the ability of N-HR and C-HR peptides to inhibit cell-cell and virus-cell fusion. The IC_{50} values reported are from a single representative assay done in duplicate.

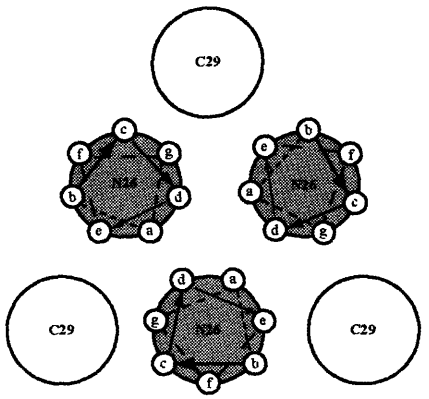
Figure 1. Helical wheel alignment of gp41 and peptide sequences.

(A) In the fusion-active structure of HIV-1 gp41, the N-terminal helices form a trimeric coiled coil. They contain a characteristic heptad repeat motif, in which the first (denoted the **a** position) and fourth (**d** position) residues of each seven-residue repeat interact in the interior of the coiled coil. Residues at these positions are generally, although not always, hydrophobic. In contrast, residues on the outer face of the coiled coil are often polar. In the gp41 core, C-terminal helices pack into the grooves between adjacent N-helices in an antiparallel manner. N-HR **e** and **g** position amino acids form contacts with these C-helices. This pattern is more clearly viewed when arranged on a helical wheel, as shown here.

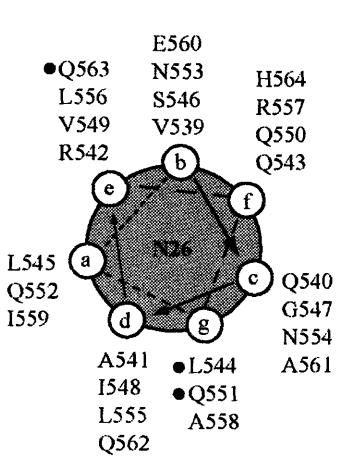
(B) The sequence of the N26 and C29 residues are shown on a helical wheel diagram. Shaded dots indicate the three residues mutated to alanine in mIQN26(3A).

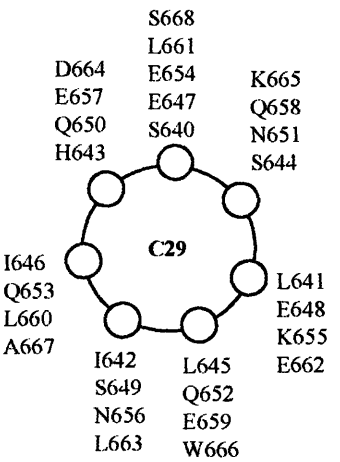
(C) The sequences of all chemically synthesized peptides are shown. All peptides are acetylated at the amino terminus and amidated at the carboxy terminus.

A



B



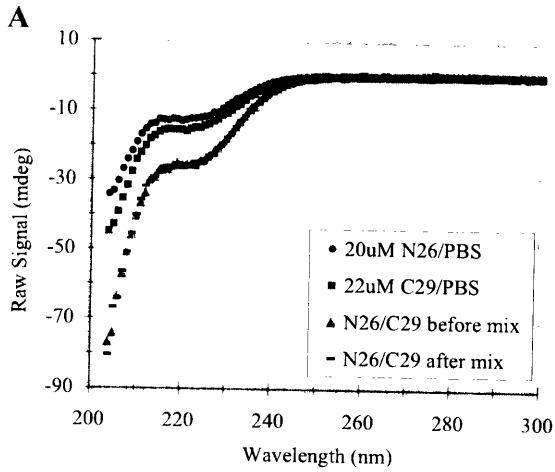


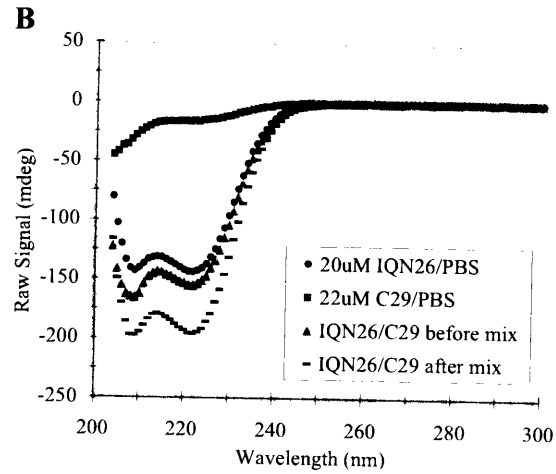
C

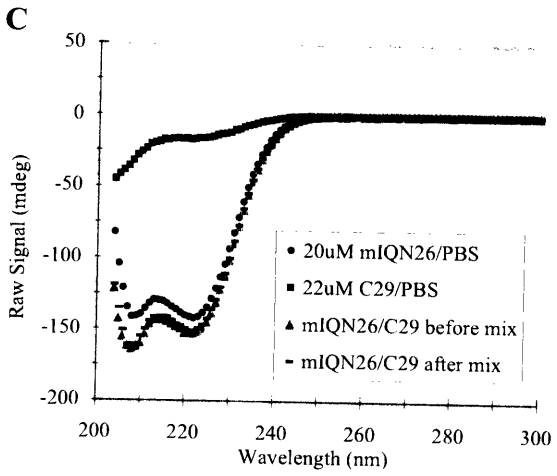
Peptide	Sequence
N26	Ac-YGGVQARQLLSGIVQQQNNLLRAIEAQQH-NH_2
IQ	Ac-RMKQIEDKIEEIESKQYKIENEIARIKKLI-NH_2
IQN26	Ac-RMKQIEDKIEEIESKQYKIENEIARIKKLIVQARQLLSGIVQQQNNLLRAIEAQQH-NH_2
mIQN26(3A)	Ac-RMKQIEDKIEEIESKQYKIENEIARIKKLIVQARQALSGIVQAQNNLLRAIEAQAH-NH_2
IZ	Ac-YGGIKKEIEAIKKEQEAIKKKIEAIEKEI-NH_2
IZN26	Ac-YGGIKKEIEAIKKEQEAIKKKIEAIEKEIVQARQLLSGIVQQQNNLLRAIEAQQH-NH_2
C29	Ac-SLIHSLIEESQNQQEKNEQELLELDKWAS-NH_2

Figure 2. Interaction of chimeric N-HR peptides with C29.

Circular dichroism wavelength scans of peptide solutions in PBS were obtained using a dual compartment mixing chamber. Initial scans of the N-helix peptides (IQN26, IZN26, mIQN26(3A), or N26) were performed with a 20 μM peptide sample in one chamber and PBS buffer in the other. Similar initial scans were performed with 22 μM C29 in one chamber and buffer in the other. A 'before mix' spectra was obtained with 20 μM N-peptide in one chamber and 22 μM C29 in the other. The cuvette was inverted to mix the samples, and an 'after mix' scan was then performed. The wavelength scans for N26 **(A)**, IQN26 **(B)**, mIQN26(3A) **(C)**, and IZN26 **(D)** mixed with C29 are shown.


A
10
-10
-30
-50
-70
-90
Raw Signal (mdeg)
200
220
240
260
280
300
Wavelength (nm)
• 20uM N26/PBS
▪ 22uM C29/PBS
▲ N26/C29 before mix
- N26/C29 after mix


B
50
0
-50
-100
-150
-200
-250
Raw Signal (mdeg)
200
220
240
260
280
300
Wavelength (nm)
• 20uM IQN26/PBS
▪ 22uM C29/PBS
▲ IQN26/C29 before mix
- IQN26/C29 after mix


C
50
0
-50
-100
-150
-200
Raw Signal (mdeg)
200
220
240
260
280
300
Wavelength (nm)
• 20uM mIQN26/PBS
▪ 22uM C29/PBS
▲ mIQN26/C29 before mix
- mIQN26/C29 after mix

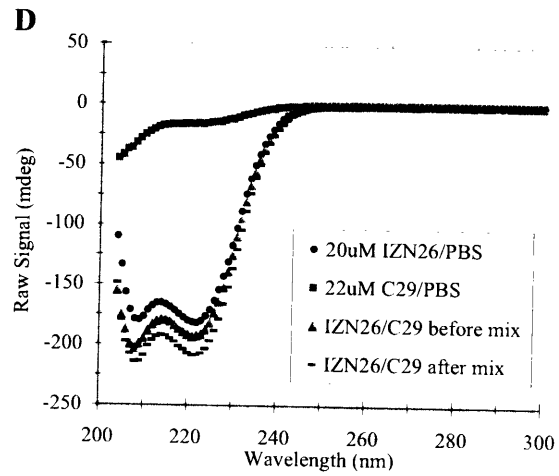
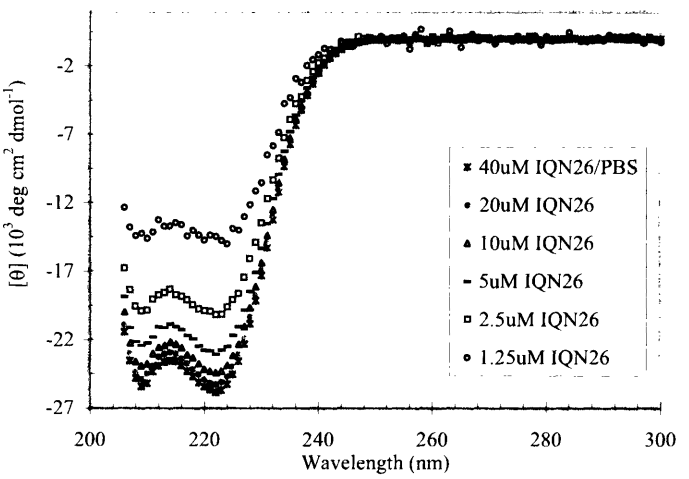

D
50
0
-50
-100
-150
-200
-250
Raw Signal (mdeg)
200
220
240
260
280
300
Wavelength (nm)
• 20uM IZN26/PBS
▪ 22uM C29/PBS
▲ IZN26/C29 before mix
- IZN26/C29 after mix

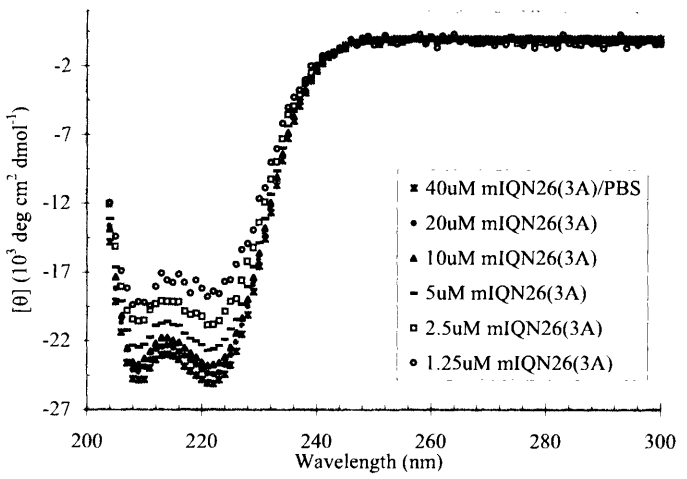
Figure 3. Concentration-dependent secondary structure of chimeric N-HR peptides.

N-HR chimeric peptides were serially diluted and assayed for helicity by circular dichroism. For IQN26 **(A)** and mIQN26(3A) **(B)**, samples were diluted in a mixing cell. For IZN26 **(C)**, serial dilutions were scanned in a 1 cm cuvette.

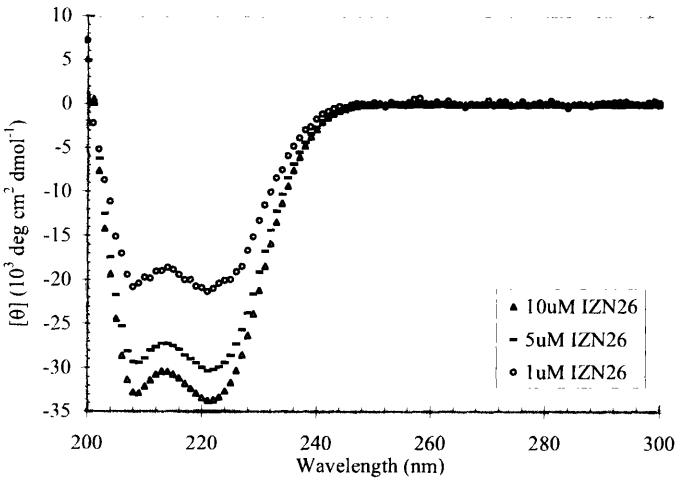
A


-2
-7
-12
-17
-22
-27
200
220
240
260
280
300
Wavelength (nm)
[θ] (10³ deg cm² dmol⁻¹)
40uM IQN26/PBS
20uM IQN26
10uM IQN26
5uM IQN26
2.5uM IQN26
1.25uM IQN26

B


-2
-7
-12
-17
-22
-27
200
220
240
260
280
300
Wavelength (nm)
[θ] (10³ deg cm² dmol⁻¹)
40uM mIQN26(3A)/PBS
20uM mIQN26(3A)
10uM mIQN26(3A)
5uM mIQN26(3A)
2.5uM mIQN26(3A)
1.25uM mIQN26(3A)

C


10
5
0
-5
-10
-15
-20
-25
-30
-35
200
220
240
260
280
300
Wavelength (nm)
[θ] (10³ deg cm² dmol⁻¹)
10uM IZN26
5uM IZN26
1uM IZN26

CHAPTER 3

THE FUSION ACTIVITY OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) GP41 DEPENDS ON INTERHELICAL INTERACTIONS

Chapter 3 has been published as:

Tara R. Suntoke and David C. Chan,

"The Fusion Activity of HIV-1 gp41 Depends on Interhelical Interactions."

Journal of Biological Chemistry **280**, 19852-19857 (2005).

Abstract

Infection by human immunodeficiency virus type I (HIV-1) requires the fusogenic activity of gp41, the transmembrane subunit of the viral envelope protein. Crystallographic studies have revealed that fusion-active gp41 is a "trimer-of-hairpins," in which three central N-terminal helices form a trimeric coiled coil surrounded by three antiparallel C-terminal helices. This structure is stabilized primarily by hydrophobic, interhelical interactions, and several critical contacts are made between residues that form a deep cavity in the N-terminal trimer and the C-helix residues that pack into this cavity. In addition, the trimer-of-hairpins structure has an extensive network of hydrogen bonds within a conserved glutamine-rich layer of poorly understood function. Formation of the trimer-of-hairpins structure is thought to directly force the viral and target membranes together, resulting in membrane fusion and viral entry. We test this hypothesis by constructing four series of gp41 mutants with disrupted interactions between the N- and C-helices. Notably, in the three series containing mutations within the cavity, gp41 activity correlates well with the stability of the N-C interhelical interaction. In contrast, a fourth series of mutants involving the glutamine layer residue Gln-653 show fusion defects even though the stability of the hairpin is close to wild-type. These results provide evidence that gp41 hairpin stability is critical for mediating fusion and suggest a novel role for the glutamine layer in gp41 function.

Structural analyses of the transmembrane subunits of several viral envelope proteins (Env)[1] have provided much insight into the mechanism of Env-mediated membrane fusion. Remarkably, the putative fusion-active structures of these proteins from retroviruses (e.g. HIV, SIV, MoMLV), orthomyxoviruses (influenza), paramyxoviruses (SV5, HRSV), filoviruses (Ebola), and coronaviruses (SARS-CoV) (1,2) all reveal the same six-helix bundle conformation. This six-helix structure can also be described as a trimer-of-hairpins, in which each hairpin consists of an N-terminal helix (N-helix) packed against an antiparallel C-terminal helix (C-helix). Three central N-helices form a coiled-coil surface with three prominent hydrophobic grooves that each allow C-helix binding (3-6). Thus, the trimer-of-hairpins structure is largely stabilized by hydrophobic packing between the three internal N-helices and between the N- and C-helices (Figure 1, *A* and *B*). This fusion-active configuration juxtaposes the amino terminus of gp41, which contains a fusion peptide inserted into the target cell membrane, with the carboxyl terminus, which contains the transmembrane anchor in the viral membrane. The current model for membrane fusion proposes that formation of the hairpin structure provides the energy necessary for viral and cellular membrane apposition, leading to membrane fusion (1,3,6).

There are two prominent features of the trimer-of-hairpins structure in HIV and SIV gp41. First, three symmetric hydrophobic cavities on the surface of the N-helix coiled coil each forms a docking site for three residues from a C-helix. Notably, small peptides or synthetic molecules targeted to these cavities prevent trimer-of-hairpins formation and inhibit fusion (7,8); the efficacy of these molecules underscores the critical role of interhelical interactions within the cavity in gp41 function. Second, a planar cluster of conserved glutamine and asparagine residues, from both the N- and C-helices, form a hydrogen-bonded layer within the six-helix bundle (5,6,9). Though a novel structural feature, the function of this 'glutamine layer' remains unclear (6,9); indeed, none of the reported mutations in this region reduce gp41 activity. In fact, such substitutions can increase gp41 core stability (10-12) and, in the unusual case of a Gln-652 to leucine mutation (Q652L), appear to increase fusogenicity (13).

Several mutational and biophysical analyses of gp41 support the hypothesis that hairpin formation provides the driving force for membrane fusion. For example, the N-helix residues Leu-565, Leu-568, and Val-570, located on the cavity surface, form hydrophobic contacts with C-helices as determined by x-ray crystallography. Alanine mutations at these positions decrease six-helix bundle stability and reduce envelope-mediated fusion (12-16). These mutations are notable because they likely represent a specific defect in the fusion pathway, unlike some N- and

C-helix mutants that fail to correctly process the gp160 precursor into gp120 (the receptor binding subunit) and gp41 (11,12,16). Although these studies support a role for hairpin formation in membrane fusion, a systematic analysis of the correlation between hairpin stability and gp41 activity would provide a more rigorous test of the model.

In this study we seek to identify the energetic contribution of the N-C helical interaction to the fusion reaction. We made a series of substitutions at each of four positions in the N-C helical interface. We report that for mutations within the N-helix cavity, the fusion activity of gp41 depends on the stability of the N-C interface. In contrast, for the series of substitutions in the glutamine layer, we find that gp41 activity is abolished even in the presence of stable N-C interactions, thereby providing evidence for an important role for the glutamine layer in gp41-mediated fusion.

Materials and methods

Plasmid Construction and Mutagenesis—Env mutations were made using QuickChange site-directed mutagenesis (Stratagene) of pBS-SKII containing HXB2 gp160 (pBS-SKII-gp160), and verified by DNA sequencing. For expression in 293T cells, all mutations were subcloned into pcDNA1.1/Amp (Invitrogen). For generation of stable cell lines, the mutations were subcloned into the pB-IRES-CD2 retroviral vector (gift from Jonathan Bogan, Whitehead Institute, Cambridge MA).

For expression of recombinant proteins, mutations were cloned into the pAED4 vector (4). Mutant N34 and wild-type C28 fragments were PCR amplified from pBS-SKII-gp160 using the following primers: N34for (5' GCAATTCCATATGTCTGGTATAGTGCAGCAG) and L6N34rev (5' ACCGCGACCACCGCTTCTTGCTTGGAGCTGCTTG) for the N34 fragment, and L6C28for (5' GGTGGTCGCGGTGGTTGGATGGAGTGGGACAGAG) and C28-HISrev (5' GGAATTCCTAGTGATGATGATGATGATGCTTTTCTTGCTGGTTTTGC) for the C28 fragment. These primers introduce a six residue linker (Ser Gly Gly Arg Gly Gly) between N34 and C28, and a 6X-histidine tag (HIS-tag) 3' of C28. For mutant C28 fragments, C28-HISrev primers encoding the mutations were used. The N and C fragments were joined by overlap PCR with primers N34for and C28-HISrev, and subcloned into pAED4.

Protein Expression and Purification—Bacterial cultures expressing N34(L6)C28-HIS variants were induced with 1 mM IPTG for 3 to 4 hours. Cell pellets were resuspended in lysis buffer (50 mM Tris (pH 8), 15% glycerol, and 300 mM NaCl) containing 10 μM $MgCl_2$, 1 μM $MnCl_2$, and 100 μg/mL DNase I, lysed by sonication, and centrifuged (27,000 x *g* for 30 minutes) to separate the soluble and inclusion body fractions. The pellet was washed again in lysis buffer and supernatants from both washes were pooled. With the exception of the glycine mutants, all peptides were present in the soluble fraction and were affinity-purified using nickel-nitrilotriacetate (Ni-NTA) (Qiagen, Inc.). For the glycine mutants, inclusion body pellets were solubilized in lysis buffer containing 6 M guanidine-HCl before purification using Ni-NTA beads. Peptide-containing fractions were dialyzed into 5% acetic acid, purified to >95% purity over a preparative C5 column by reverse-phase HPLC, and lyophilized.

Circular Dichroism Spectroscopy—Lyophilized peptide was dissolved in water, and concentrations were determined by absorbance at 280 nm in 6 M guanidine-HCl (17). Peptides were then diluted to 10 μM in PBS (50 mM sodium phosphate, 150 mM sodium chloride (pH 7)). Circular dichroism (CD) spectra were obtained on an AVIV 62DS CD spectrometer. To determine helical content, a value of -33,000 deg cm^2 $dmol^{-1}$ was used for theoretical 100% helicity (12). Thermal denaturation scans were performed on 10 μM peptide samples in PBS. The midpoint of the thermal unfolding transition (T_m) was calculated as the maximum value of the first derivative of the molar ellipticity [θ] at 222 nm with respect to temperature. Consistent with previous reports for N34(L6)C28 (10,11), thermal denaturation of these recombinant peptides was cooperative but not reversible.

Sedimentation Equilibrium—All measurements were recorded on a Beckman XL-I analytical ultracentrifuge with a Ti-60a rotor. Wild-type N34(L6)C28-HIS was dialyzed overnight in PBS or sodium acetate buffer (50 mM sodium acetate, 150 mM NaCl (pH 5.5)), and diluted to 5, 10, and 30 μM. Samples dialyzed in PBS formed some precipitate, while samples in sodium acetate remained fully soluble. Data were collected at 20°C, with rotor speeds of 20, 30, and 42 krpm and wavelengths of 229 and 280 nm. Data sets were fitted simultaneously with WinNONLIN (version 1.06) (18). No systematic residuals were observed, and both samples fit well to a single species model.

Because wild-type N34(L6)C28-HIS formed stable trimers in both buffers and remained fully soluble after overnight dialysis at low pH, mutant peptides were analyzed in the sodium

acetate buffer. All N34(L6)C28-HIS variants except L565G fit best to a single species model, showed no systematic residuals, and had molecular weights within 10% of those calculated for an ideal trimer. The L565G peptide was predominantly trimeric in solution, but the data fit slightly better as a monomer-trimer equilibrium containing 80% trimers.

Envelope Glycoprotein Expression and Western Blot Analysis—Mutant envelope glycoproteins were expressed in 293T cells by transient transfection for expression analysis and cell-cell fusion assays. 293T cells were plated in gelatin-coated 12-well plates. At ~30 to 50% confluency, cells were transfected with 800 ng pcDNA-gp160 expressing mutant gp160 and 400 ng pnGFP (a plasmid expressing nuclear GFP) using Lipofectamine 2000 (Invitrogen). Transfection efficiency was assessed by GFP fluorescence, and was similar for all mutants (~30-50%).

Envelope expression was measured by Western blot analysis. gp41 and gp160 from cell lysates were analyzed with an anti-gp41 monoclonal antibody, Chessie 8 (National Institutes of Health AIDS Research and Reference Reagent Program). Shed gp120 was immunoprecipitated from the culture supernatant using an anti-gp120 polyclonal antibody (gift of Peter Kim, Merck Research Laboratories, West Point PA), and protein-A Sepharose beads (Sigma). To determine cell surface expression levels of Env, transfected cells were washed with cold PBS (pH 8.0) and biotinylated using 5 mM Sulfo-NHS-LC-Biotin (Pierce) in PBS for 30 minutes at 4°C. After washing with PBS containing 100 mM glycine to remove excess biotin, cells were lysed in PBS containing 1% NP-40 (Sigma). Biotinylated surface proteins were precipitated using immobilized NeutrAvidin (Pierce) and separated by SDS-PAGE.

Stable cell lines expressing the alanine mutants were made using retroviral transduction of NIH3T3 cells. To generate retrovirus, 293T cells were plated on gelatin-coated 60 mm plates, and transfected with pBI-CD2-gp160 encoding mutant gp160 and a MoMLV packaging vector, pCLeco (gift from C. Lois, Massachusetts Institute of Technology). A 1:2 dilution of viral supernatant was used to infect NIH3T3 cells in the presence of 4 μg/ml polybrene. After several passages, the infected NIH3T3 cell lines were used for dye transfer assays.

Cell-Cell Fusion Assay—Forty-eight hours after transient transfection, 293T cell monolayers were resuspended and counted. A mixture of 4×10^4 envelope-expressing 293T cells and 4×10^4 target cells, 3T3-T4-CXCR4 (National Institutes of Health AIDS Research and Reference Reagent Program), was plated in duplicate into gelatin-coated 8-well slides and incubated for 21

hours at 37°C. After fixation and hematoxylin staining, the number of syncytia per well was counted under phase microscopy. Cells containing three or more nuclei were scored as syncytia.

Dye Transfer Fusion Assay—3T3 cells stably expressing gp41 mutations were plated at a density of 1 x 10^6 cells per 60 mm dish. Eight hours later, cell monolayers were incubated with 420 nM calcein AM (Molecular Probes) in PBS for 10 minutes at 37°C and gently washed twice with PBS before addition of labeled target cells. Target cells, 3T3-T4-CXCR4, were labeled with the lipid dye octadecyl rhodamine B (R18) (Molecular Probes). 2 x 10^6 target cells were washed in PBS, incubated in PBS containing 2 μM R18 for 15 minutes at room temperature in the dark, and washed twice in PBS before resuspending in growth media. Approximately 4 x 10^5 target cells were added to envelope-expressing monolayers, and dishes were immediately incubated at 37°C. Cells were examined for dye transfer every 20 to 30 minutes by counting the number of co-labeled cells in five separate fields at 200X magnification.

Results

Design of four allelic series—Residues Leu-565, Leu-568, and Val-570 are located on the surface of the N-terminal coiled coil within a large cavity and interact with the internal face of the C-helix (Figure 1, *A* and *B*). In previous studies, mutation of these residues to alanine reduced fusion but did not affect gp160 processing or gp120-gp41 association (12-15), suggesting a specific defect in the fusion pathway. In order to obtain a range of gp41 core stabilities, each position was mutated to glycine, alanine, and phenylalanine. We chose glycine for its helix destabilizing properties (19), alanine for its reduced hydrophobic bulk, and phenylalanine for its increased bulk over the native leucine or valine residue.

The polar contacts within the glutamine layer were disrupted by mutagenesis of Gln-653 (Figure 1*C*). This residue is buried in the interhelical interface and forms hydrogen bonds with neighboring helices in the core crystal structure (6,20). Furthermore, it is highly conserved; only two out of 630 viruses in the Los Alamos HIV Sequence Database show mutations at this position (http://www.hiv.lanl.gov). The Q653A and Q653L substitutions were designed to abolish polar contacts with neighboring glutamine residues, whereas the Q653E substitution retains bulk and polarity.

Biophysical analysis of recombinant gp41 core structures—We used a recombinant peptide model to measure the biophysical properties of mutant gp41 cores. This gp41 core peptide, termed N34(L6)C28-HIS, contains N- and C- terminal helices that are joined by a six residue linker (Figure 1*A*). It is similar to the previously described N34(L6)C28 peptide (21), but contains a C-terminal HIS-tag to facilitate purification. Circular dichroism indicates that the molecule is >90% helical and thermostable, with a thermal denaturation midpoint (T_m) of 74°C. In addition, it forms a discretely trimeric species as determined by sedimentation equilibrium (Figure 2 and Table 1). When compared with published data for N34(L6)C28, N34(L6)C28-HIS has a 4°C higher melting temperature. We speculate that addition of the C-terminal tag promotes stability by attenuating the helix dipole moment or by reducing fraying of the helical ends (19,22).

For the three mutant series involving the N-helix, biophysical analysis indicated that we successfully generated a wide range of stabilities in the gp41 core (Table 1). In each position, all mutants are less stable than wild-type, with the relative order Gly < Ala < Phe < WT. At position 568, the phenylalanine mutant (T_m of 61°C) is only slightly more stable than the alanine mutant (T_m of 60°C). The low stability of the L568F variant may be caused by local distortion to accommodate the large aromatic ring. All mutants are highly helical and, with the exception of L565G (see Materials), the least thermostable variant in this study, all form clean trimers as evidenced by sedimentation equilibrium analysis (Table 1).

Interestingly, mutations at Gln-653 do not appreciably destabilize the six-helix bundle structure. The Q653A, Q653L, and Q653E mutants have T_m values close to wild-type, and form stable trimers (Table 1). Thus, in contrast to the effects observed with the cavity mutations, altering the native interhelical interaction within the glutamine layer does not substantially affect gp41 core stability.

Expression and processing of gp41 variants—We examined the expression and processing of envelope mutants after transient transfection of mutant constructs into 293T cells. Transfection efficiency, assessed by co-transfection of a GFP expression vector, was similar for all Env proteins (T.S. and D.C., unpublished data). Western blot analysis of the cell lysates with an anti-gp41 monoclonal antibody reveals that all gp41 mutants, except L565A, are expressed and processed at levels similar to wild-type (Figure 3, *A* and *B*). These results are quantitated by gel densitometry in Table 2 and are consistently similar to wild-type levels (80-120%). L565A gp41

is expressed at 60% of wild-type. The levels of cell surface Env were quantitated by surface biotinylation and precipitation using avidin-conjugated beads. Expression levels of mutant gp160 and gp41 on the cell surface were similar to wild-type (data not shown) and the ratio of surface gp41 to gp160 confirmed that proteolytic processing of the Env precursor was normal for all mutants (Table 2).

Under normal conditions, the noncovalent interaction between gp120 and gp41 is labile, and a portion of gp120 is shed from the surface of virions or Env-expressing cells. Some mutations in gp41 affect its association with gp120, leading to increased release of gp120 into the culture supernatant. Immunoprecipitation of gp120 from the culture medium followed by Western blot analysis shows that steady-state levels of shed gp120 are similar for wild-type and all mutants (Figure 3 and Table 2). These results further indicate that mutations at positions 565, 568, 570, and 653 in gp41 do not affect the native conformation of the envelope glycoprotein.

Fusogenic potential of mutant envelope glycoproteins—The gp41 mutants were tested for membrane fusion activity with syncytia assays. For the N-helix mutants, there is a good correlation between thermostability of the gp41 core and fusion activity (Table 2). The glycine mutants, which are the least stable, have barely detectable levels of fusion. As the stabilities of the mutants increase, there is a simultaneous increase in fusion activity. This trend becomes clear when the fusion activity of the mutants is plotted as a function of gp41 stability (Figure 4). This plot also reveals that the V570A mutant has unusually low fusion activity. V570A has intermediate stability (T_m of 62°C) that is comparable to L565F, L568A, and L568F; however, unlike these three mutants, it is severely deficient in cell fusion activity. The reason for this anomaly remains to be determined. Our results agree with and extend previously published data for alanine mutants at positions 565, 568, and 570 (12,14-16), which suggest a similar correlation between fusion and core stability.

The Gln-653 mutants provide a striking exception to this correlation. Despite its highly stable core structure, the Q653E substitution caused a 32% decrease in gp41 fusion activity compared to wild-type. A more severe mutation, Q653L, which abolishes polarity, further decreased gp41 function (55% of wild-type), while removal of both polarity and bulk with the Q653A mutation almost completely abrogated fusion. These results provide the first evidence that disruption of polar contacts within the gp41 glutamine layer decreases envelope-mediated

fusion. The low fusogenicity of these 653 mutants despite their extremely thermostable core structures implicates the glutamine layer in a novel role in gp41 function.

Fusion kinetics of alanine mutants—In gp41-mediated fusion, there is a significant lag time after mixing cells before the first fusion events are observed (23). This delay is thought to reflect the slow conformational changes that occur in the envelope protein during the fusion process. Our N-helix mutants likely show fusion defects because weakened N-C helical interactions reduce the efficiency of membrane apposition. Alternatively, it is possible that attenuated N-C helical interactions also cause a delay in membrane apposition, resulting in a kinetic defect. In this scenario, Env molecules would display slower-than-wild-type kinetics of fusion, leading to a delay in the onset of initial fusion events.

To test this possibility, the L565A, L568A, and V570A mutants were analyzed further. The alanine mutants were chosen because they display low but detectable levels of fusion. The kinetics of fusion at early timepoints was analyzed using 3T3 cell lines stably expressing L565A, L568A, and V570A mutant envelope glycoproteins. The alanine mutants displayed levels of expression and processing similar to wild-type and fusion phenotypes analogous to those seen with transient transfections (4% for L565A, 17% for L568A, and 2% for V570A) (T.S. and D.C., unpublished data). Dye labeling of envelope-expressing cells with a cytosolic marker, calcein-AM, and target cells with a lipid marker, R18, allowed initial fusion events to be detected as dual-labeled cells. For each cell line, fusion was monitored approximately every 25 minutes. Figure 5 shows the fusion progression of the alanine mutants at early timepoints after co-incubation of effector and target cells. A small number of dye transfer events was initially detected for wild-type gp41 between 15 and 30 minutes post-mixing, followed by a dramatic increase starting at 80 minutes. These results correspond well with reported fusion kinetics experiments, in which initial fusion events are detected after time lag, followed by rapid acceleration ((23) and references therein). With the alanine mutants, it is difficult to reliably measure the onset of initial fusion events due to the low efficiency of fusion. Nevertheless, the acceleration phase for all the alanine mutants occurs between 80 to 100 minutes, indicating that the early kinetics for wild-type and mutant cell lines are similar.

Discussion

Structural studies of the HIV-1 envelope protein suggest an elegant mechanism through which gp41 mediates apposition of viral and target cell membranes: the interaction of N- and C-helices to form the six-helix bundle forces the target cell membrane (containing the inserted fusion peptide) close to the viral membrane (containing the gp41 transmembrane segment). If this model is correct, the efficiency of gp41-mediated membrane fusion should depend on the strength of the N-C interhelical interaction. Indeed, with three sets of N-helix mutants, we observe a strong correlation between trimer-of-hairpins stability and membrane fusion (Figure 4). Fusion activities decrease rapidly with decreased stability, and all mutants with a T_m of 57°C or lower lack significant activity, suggesting that there may be a threshold stability necessary to mediate membrane apposition. To further test this model, it will be important to develop experimental systems to directly measure the free energy change during the transition of gp41 from its native conformation to its fusogenic conformation.

Our fourth mutant series at C-helix position 653 shows markedly different results and suggests a significant role of the glutamine layer in gp41 function. Substitutions at position 653 do not significantly destabilize the hairpin but nevertheless reduce gp41-mediated membrane fusion (Tables 1 and 2). Although the unusual structural features of the glutamine layer were noted in earlier structural studies (5,6,9), this is the first evidence of a functional role for this domain in membrane fusion.

The basis for the glutamine layer requirement remains to be elucidated. Previous mutations in this region have not supported an essential role for this region in membrane fusion (10-13). Interestingly, in the N-helix, residues contributing to the glutamine layer are part of a small region implicated in resistance to the potent viral entry inhibitor enfuvirtide (reviewed in (24), 25). Viruses containing such mutations have reduced replicative fitness *in vitro*, and revert back to wild-type in the absence of inhibitor (25,26). In addition, SNARE helical bundles involved in intracellular vesicle fusion appear to have a similar ionic layer, which may be important for efficient recycling of SNARE complexes (27).

Numerous studies using designed coiled coils suggest that buried polar residues can play important roles in interhelical interactions. For instance, a buried glutamine can decrease the strength of interaction between helices (28). In addition, a buried polar residue at the interface of

a short dimeric coiled coil can influence the orientation of the dimer, facilitating reversibility of assembly (29); this destabilization could permit exchange of binding partners and may be biologically relevant. While our Gln-653 mutations do not significantly change the stability or the oligomeric state of the gp41 core, it remains to be determined whether they have more subtle effects on gp41 protein structure.

Acknowledgments

We thank Drs. R. Olsen and A. Herr for assistance with sedimentation equilibrium experiments; Dr. J. Kaiser for help with Figure 1*C*; Drs. D. Eckert and M. Kay for critical review of the manuscript; and E. Griffin and members of the Chan lab for helpful discussions.

[1] The abbreviations used are: Env, envelope glycoprotein; HIV-1, human immunodeficiency virus type 1; N-helix, N-terminal heptad repeat region of gp41; C-helix, C-terminal heptad repeat region of gp41; T_m, midpoint of thermal unfolding transition.

References

1. Eckert, D. M., and Kim, P. S. (2001) *Annu Rev Biochem* **70**, 777-810
2. Xu, Y., Lou, Z., Liu, Y., Pang, H., Tien, P., Gao, G. F., and Rao, Z. (2004) *J Biol Chem* **279**, 49414-9
3. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell* **89**, 263-273
4. Lu, M., Blacklow, S. C., and Kim, P. S. (1995) *Nat Struct Biol* **2**, 1075-1082
5. Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997) *PNAS* **94**, 12303-12308
6. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature* **387**, 426-430
7. Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A., and Kim, P. S. (1999) *Cell* **99**, 103-115
8. Ferrer, M., Kapoor, T. M., Strassmaier, T., Weissenhorn, W., Skehel, J. J., Oprian, D., Schreiber, S. L., Wiley, D. C., and Harrison, S. C. (1999) *Nat Struct Biol* **6**, 953-960
9. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *PNAS* **95**, 9134-9139
10. Shu, W., Liu, J., Ji, H., Radigen, L., Jiang, S., and Lu, M. (2000) *Biochemistry* **39**, 1634-1642
11. Wang, S., York, J., Shu, W., Stoller, M. O., Nunberg, J. H., and Lu, M. (2002) *Biochemistry* **41**, 7283-7292
12. Lu, M., Stoller, M. O., Wang, S., Liu, J., Fagan, M. B., and Nunberg, J. H. (2001) *J Virol* **75**, 11146-11156
13. Cao, J., Bergeron, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. (1993) *J Virol* **67**, 2747-2755
14. Ji, H., Shu, W., Burling, F. T., Jiang, S., and Lu, M. (1999) *J Virol* **73**, 8578-8586
15. Bar, S., and Alizon, M. (2004) *J Virol* **78**, 811-820
16. Mo, H., Konstantinidis, A. K., Stewart, K. D., Dekhtyar, T., Ng, T., Swift, K., Matayoshi, E. D., Kati, W., Kohlbrenner, W., and Molla, A. (2004) *Virology* **329**, 319-327
17. Edelhoch, H. (1967) *Biochemistry* **6**, 1948-1954
18. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys J* **36**, 575-588
19. Fersht, A. (1999) *Structure and Mechanism in Protein Science* (Julet, M., Ed.), W. H. Freeman and Company, New York. p. 508-539.

20. Yang, Z. N., Mueser, T. C., Kaufman, J., Stahl, S. J., Wingfield, P. T., and Hyde, C. C. (1999) *J Struct Biol* **126**, 131-144
21. Lu, M., Ji, H., and Shen, S. (1999) *J Virol* **73**, 4433-4438
22. Thomas, S. T., Loladze, V. V., and Makhatadze, G. I. (2001) *PNAS* **98**, 10670-10675
23. Gallo, S. A., Finnegan, C. M., Viard, M., Raviv, Y., Dimitrov, A., Rawat, S. S., Puri, A., Durell, S., and Blumenthal, R. (2003) *Biochim Biophys Acta* **1614**, 36-50
24. Greenberg, M., Cammack, N., Salgo, M., and Smiley, L. (2004) *Rev Med Virol* **14**, 321-337
25. Greenberg, M. L., and Cammack, N. (2004) *J Antimicrob Chemother* **54**, 333-340
26. Lu, J., Sista, P., Giguel, F., Greenberg, M., and Kuritzkes, D. R. (2004) *J Virol* **78**, 4628-4637
27. Scales, S. J., Yoo, B. Y., and Scheller, R. H. (2001) *PNAS* **98**, 14262-14267
28. Eckert, D. M., Malashkevich, V. N., and Kim, P. S. (1998) *J Mol Biol* **284**, 859-865
29. Knappenberger, J. A., Smith, J. E., Thorpe, S. H., Zitzewitz, J. A., and Matthews, C. R. (2002) *J Mol Biol* **321**, 1-6

Table 1. Summary of Biophysical Data for N34(L6)C28-HIS Mutants.

Mutant	$-[\theta]_{222}$ (10^3 deg cm^2 $dmol^{-1}$) [a]	T_m (°C) [a]	M_{obs}/M_{calc} [b]
N34(L6)C28-HIS	-30.9	74	2.8
L565G	-23.1	38	2.5
L565A	-24.2	57	2.7
L565F	-24.8	63	3.0
L568G	-25.8	43	2.9
L568A	-27.4	60	3.0
L568F	-23.0	61	2.8
V570G	-27.1	49	2.9
V570A	-27.0	62	3.1
V570F	-28.0	68	2.9
Q653A	-30.5	74	2.9
Q653L	-28.2	76	3.1
Q653E	-28.9	73	3.0

[a] All circular dichroism wavelength scans and thermal melts were performed on 10μM peptide samples in PBS (pH 7.0).

[b] Sedimentation equilibrium results are shown as the ratio of observed molecular weight (M_{obs}) to calculated molecular weight for a monomer (M_{calc}).

Table 2. Summary of Env Expression, Processing, and Fusion

Env	gp160 expression [a]	gp41 expression [a]	Cell surface gp41/gp160 [b]	gp120 secretion [c]	Cell-cell fusion (%) [d]
WT	1.0	1.0	1.0	1.0	100 (±7)
L565G	1.2	1.0	1.0	1.0	1 (±0.1)
L565A	0.8	0.6	1.0	1.2	5 (±0.6)
L565F	1.2	1.1	1.1	1.3	38 (±3.1)
L568G	0.8	0.8	1.1	0.6	5 (±0.6)
L568A	1.4	1.1	1.2	1.2	35 (±2.3)
L568F	1.1	1.0	1.2	1.1	32 (±1.4)
V570G	1.1	0.9	1.2	1.3	2 (±0.2)
V570A	0.8	1.0	1.2	1.2	4 (±0.1)
V570F	0.6	0.9	1.1	1.1	48 (±2.7)
Q653A	1.0	1.1	0.8	1.3	7 (±1.7)
Q653L	1.1	1.2	1.0	1.0	55 (±6.1)
Q653E	1.1	1.1	0.9	1.1	68 (±0.5)

[a] Relative Env expression was quantitated by densitometry. Cellular expression levels of gp160 or gp41 normalized to the amount of total protein in the lysate (β-actin) are shown.

[b] Env precursor processing is measured from cell surface proteins isolated by surface biotinylation and avidin precipitation. The reported number is the ratio (mutant gp41/mutant gp160) / (wt gp41/wt gp160), where wt stands for wild-type.

[c] The amount of shed gp120 (sgp120) is measured as the ratio (mutant sgp120/mutant cellular gp41) / (wt sgp120/wt cellular gp41), where wt stands for wild-type.

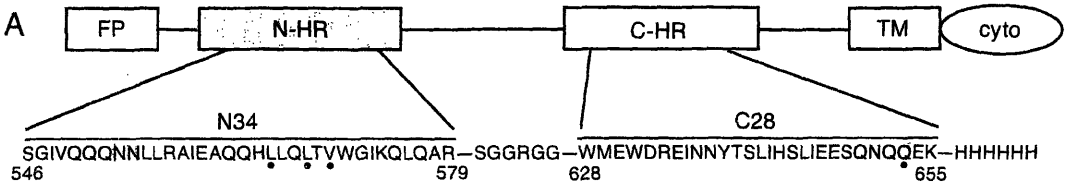
[d] To monitor cell-cell fusion activity, Env-expressing 293T cells were co-cultured with 3T3-T4-CXCR4 target cells for 21 hours at 37°C. Following fixation and staining, fusion events were determined microscopically. Experiments were repeated in duplicate at least twice. The percent of wild-type fusion in a representative fusion assay is shown. Numbers in parenthesis indicate standard deviation. The percent of fusion for cells transfected with a control vector lacking Env is 1.4 (±0.7).

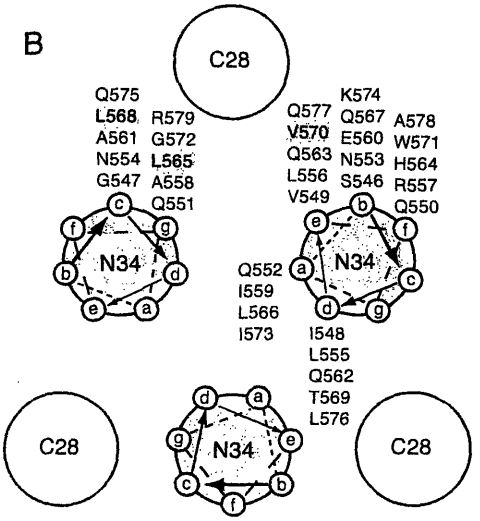
Figure 1. Structure of the HIV-1 gp41 core.

(A) A schematic diagram of HIV-1 HXB2 gp41. Important functional domains include the fusion peptide (FP), N- and C-terminal heptad repeats (N-HR and C-HR), transmembrane domain (TM), and the cytoplasmic tail (cyto). The sequence of the recombinant peptide N34(L6)C28-HIS, containing N-HR residues 546-579, a six residue linker, C-HR residues 628-655, and a C-terminal hexahistidine tag, is shown below. Residues are numbered according to their positions in HXB2 gp160. Dots below specific amino acids indicate residues that are mutated in this study, and highlighted glutamine and asparagine residues denote amino acids within the glutamine layer.

(B) Helical wheel representation of N34 and C28. Residues at the **a** and **d** positions of the N-helices interact to form the interior of the trimeric coiled-coil. C-terminal helices pack against the grooves between adjacent N-terminal helices. N-HR residues that are mutated in this study are highlighted.

(C) The glutamine layer of HIV-1 gp41. A cross-section of the gp41 core reveals glutamine and asparagine residues from both the N- and C-HR domains that form an interconnected polar layer. The C-helix residue Gln-653 protrudes toward the center of the six-helix bundle and is intimately involved in the polar network. Crystallographic studies suggest that this side chain in both HIV and SIV forms hydrogen bonds with the backbone and a glutamine residue on adjacent N helices (6,20). These putative hydrogen bonds are shown as dashed lines.


A
FP
N-HR
C-HR
TM
cyto
N34
C28
SGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQAR—SGGRGG—WMEWDREINNYTSLIHSLIEESQNQQEK—HHHHHH
546
579
628
655


B
C28
Q575
L568 R579
A561 G572
N554 L565
G547 A558
Q551
K574
Q577 Q567 A578
V570 E560 W571
Q563 N553 H564
L556 S546 R557
V549 Q550
N34
Q552
I559
L566
I573
I548
L555
Q562
T569
L576
C28
N34
C28

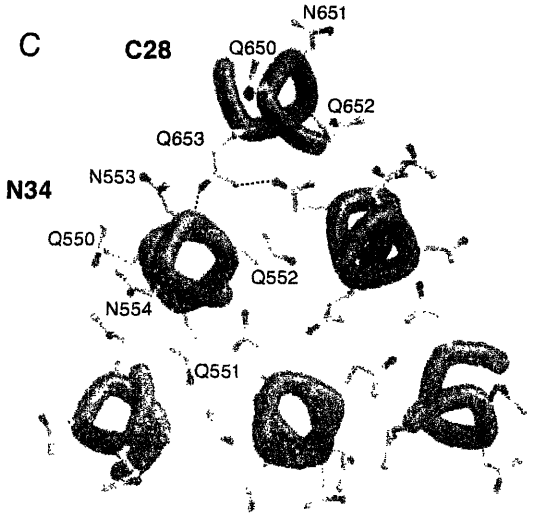

C
C28
N34
N651
Q650
Q652
Q653
N553
Q550
Q552
N554
Q551

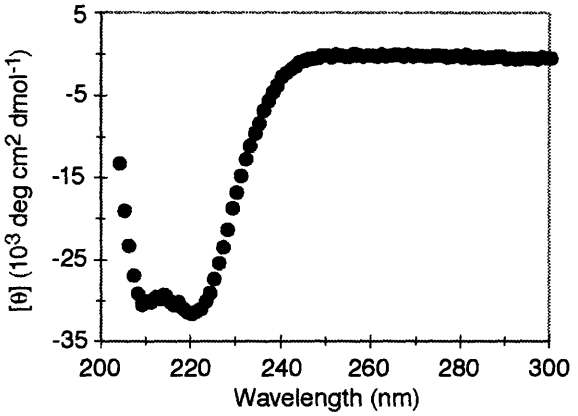
Figure 2. Biophysical properties of N34(L6)C28-HIS.

(A) The CD spectra of N34(L6)C28-HIS (10 μM) in PBS at 0°C.

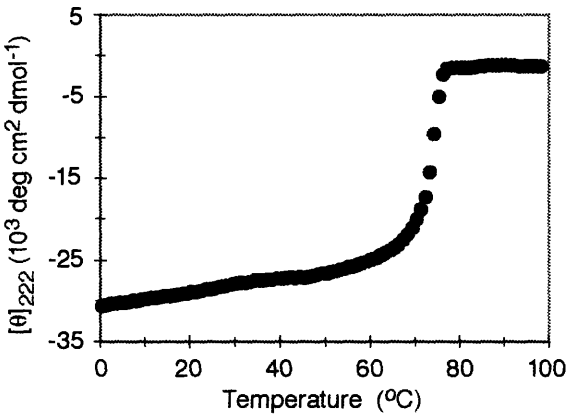
(B) Thermal dependence of the CD signal at 222 nm for a 10 μM peptide sample in PBS.

(C) Representative sedimentation equilibrium data for N34(L6)C28-HIS (10 μM) in sodium acetate buffer collected at 20 krpm and 20°C. The deviation of the data from the trimer model is plotted in the top panel. The bottom panel shows the data and a line representing the fit using an ideal single-species model.

A


5
-5
-15
-25
-35
200
220
240
260
280
300
[θ] (10³ deg cm² dmol⁻¹)
Wavelength (nm)

B


5
-5
-15
-25
-35
0
20
40
60
80
100
[θ]222 (10³ deg cm² dmol⁻¹)
Temperature (°C)

C

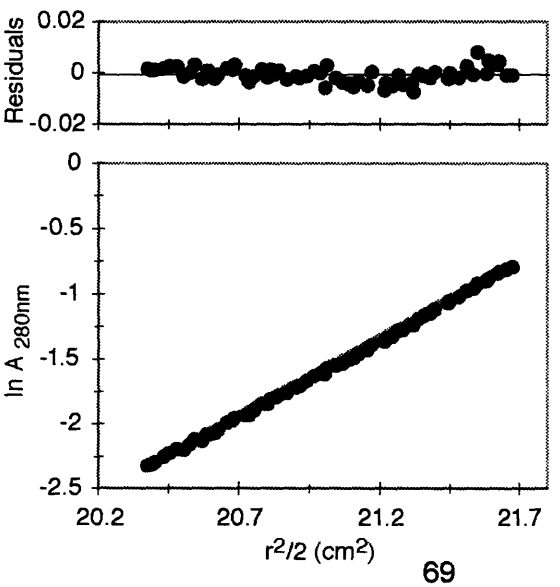

Residuals
0.02
0
-0.02
0
-0.5
-1
-1.5
-2
-2.5
ln A 280nm
20.2
20.7
21.2
21.7
r²/2 (cm²)

Figure 3. Expression and processing of mutant envelope glycoproteins.

(A) Expression analysis of N-helix mutant series. The top panel shows cellular gp160 and gp41 levels detected by Western blot analysis; the middle panel indicates the β-actin loading control; the lower panel depicts the amount of gp120 secreted into the culture medium.

(B) Expression analysis of C-helix mutants at position 653. Panels are the same as in (A).

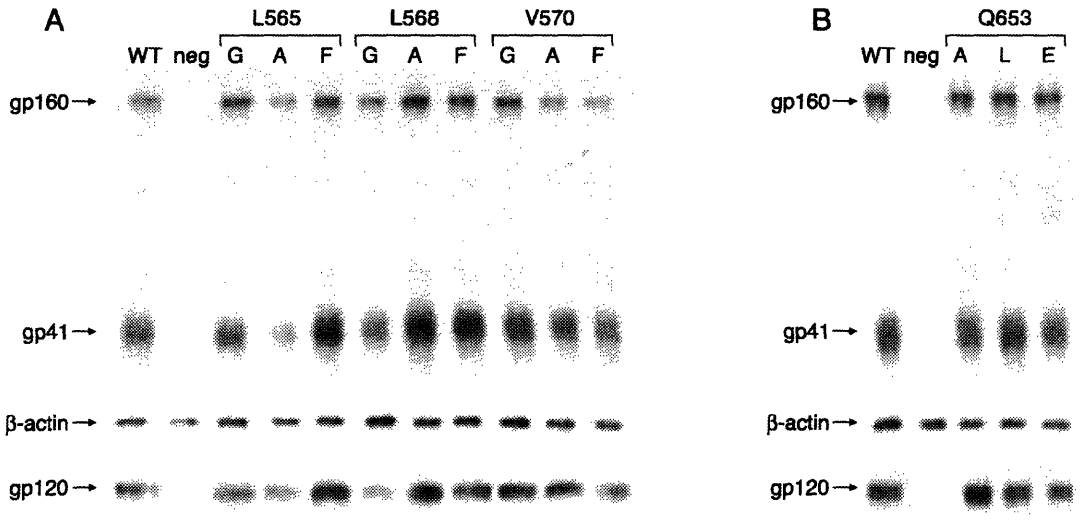

A
L565
L568
V570
WT neg G A F G A F G A F
gp160
gp41
β-actin
gp120
B
Q653
WT neg A L E
gp160
gp41
β-actin
gp120

Figure 4. Correlation between gp41 core melting temperature and mutant fusion activity.

The N-helix mutants are shown as circles (L565 series), triangles (L568 series), and squares (V570 series). For these mutants, the symbols are black for glycine mutants, shaded for alanine mutants, and empty for phenylalanine mutants. The Q653 series is shown as empty diamonds, and wild-type is marked with an X. The N- and C-helix mutant series are each connected by a line showing the correlation between T_m and fusion.

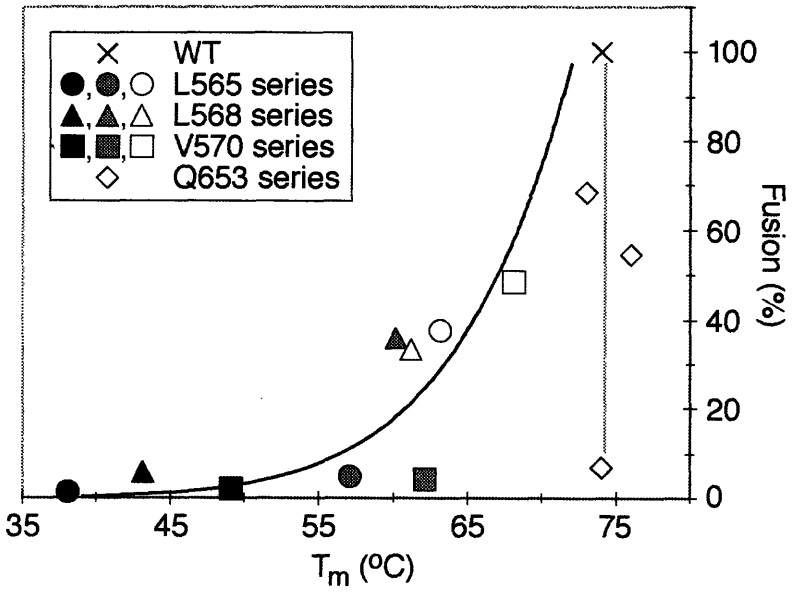

WT
L565 series
L568 series
V570 series
Q653 series
100
80
60
40
20
0
Fusion (%)
35
45
55
65
75
T_m (°C)

Figure 5. Fusion kinetics of N-helix alanine mutants.

3T3 cell lines stably expressing envelope glycoprotein as well as an empty vector control were labeled with the cytosolic dye calcein-AM. Target cells, 3T3-T4-CXCR4, were labeled with the lipid dye R18. Following co-culture at 37°C, fusion events were detected as dual-labeled cells, and were counted over a 3 hour time course. The fusion kinetics of wild-type (filled diamonds), L565A (open circles), L568A (open triangles), V570A (open squares), and empty vector-transfected (neg, filled squares) cells are shown.

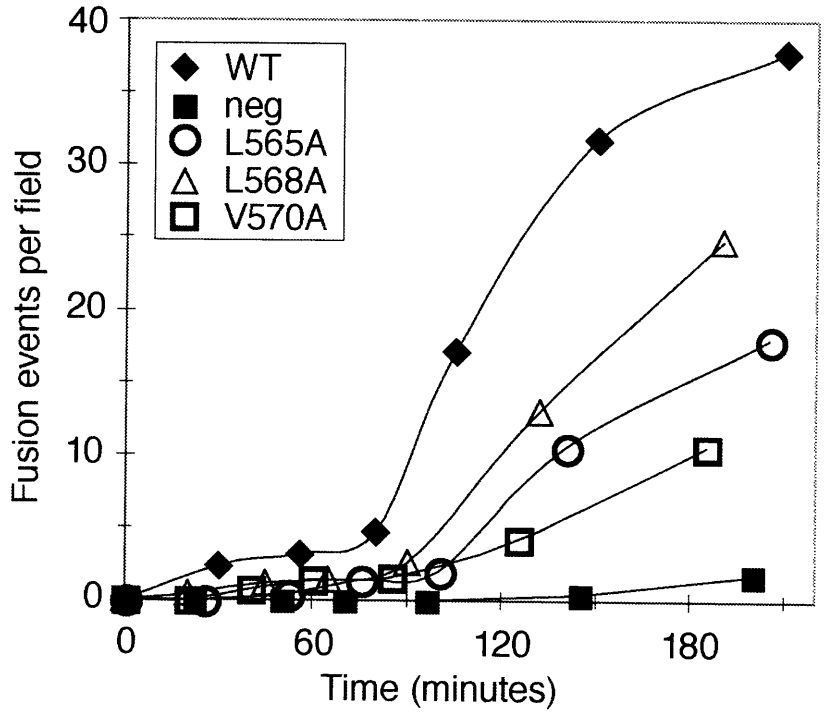
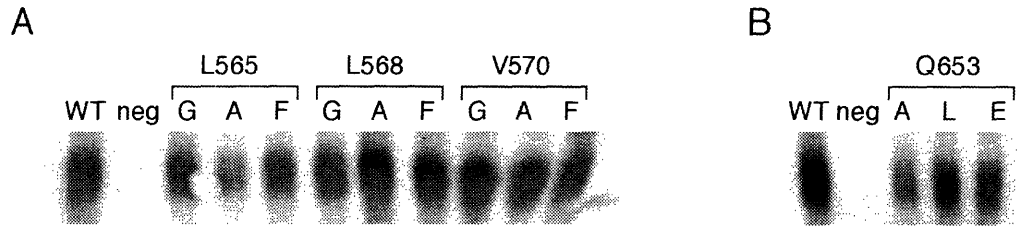

40
30
20
10
0
Fusion events per field
WT
neg
L565A
L568A
V570A
0
60
120
180
Time (minutes)

Figure 6. Cell surface expression of mutant envelope glycoproteins.

(A) Expression analysis of N-helix mutant series. Cell surface proteins were biotinylated and precipitated with avidin. Western blots depicting cell surface gp41 is shown.

(B) Cell surface gp41 expression of C-helix mutants at position 653.

Note: This figure is not a part of the published *Journal of Biological Chemistry* article, but was added as a supplement to this chapter in this thesis.


A
L565
L568
V570
WT neg G A F G A F G A F
B
Q653
WT neg A L E

CHAPTER 4

A CURE FOR AIDS: AN OVERVIEW

Portions of Chapter 4 will be submitted for publication (manuscript in preparation).

This chapter addresses scientific discoveries in HIV drug and vaccine development, novel viral entry inhibitors, and global challenges in combating the pandemic. As a broad overview, Figure 1 shows the current state of the pandemic. There are over 39 million people living with AIDS today. In 2004 alone, 3.1 million people died from AIDS-related illnesses. The effects of this pandemic extend beyond public health, encompassing economic, political, and social considerations as well. Worse, there is no end in sight.

There are eight clades, or subtypes, of the HIV-1 virus worldwide. These divisions are based on complete genome sequence similarity. Sub-Saharan Africa, with the largest infected population in the world, has a high prevalence of clade C virus. Clade E circulates in much of Southeast Asia, and is also found in India and South America. The most well characterized subtype, however, is Clade B, the HIV-1 subtype that is most prevalent in North America, the Caribbean, and Europe.

Anti-HIV drugs

There are several classes of HIV-specific antiretrovirals (ARVs) currently available, and each targets one of three viral proteins: reverse transcriptase (RT), protease (PR), and envelope (Env). This section takes a chronological look at the evolution of these drug therapies and addresses some of the challenges yet to be overcome.

In the late 1980s, within four years of the identification of HIV as the causative agent of AIDS, the first anti-HIV drug was introduced to the market. Still in use today, azidothymidine (AZT), a pyrimidine nucleoside analog, targets the HIV-1 reverse transcriptase (RT) enzyme, preventing chain elongation of nascent DNA from the viral RNA template (reviewed in (14, 24, 45, 56)). AZT belongs to a class of drugs called nucleoside reverse transcriptase inhibitors (NRTIs), in which there are currently a total of seven FDA-approved drugs. NRTIs are 2', 3'-dideoxynucleoside analogs that require three rounds of phosphorylation by cellular kinases to generate their active triphosphate form. A similar class of RT inhibitor is the NtRTI, or nucleotide RT inhibitor. To date, it contains just one member, tenofovir. This drug has the benefit of requiring only two phosphorylation steps in order to be activated. Furthermore, the negative charge of tenofovir allows it to be more easily retained within cells. Both NRTIs and NtRTIs compete with cellular dNTPs to bind the RT active site. A third class of RT inhibitors contains structurally diverse compounds. They are non-nucleoside RT inhibitors (NNRTIs), and

most bind a pocket near the RT active site, causing an allosteric change in conformation that inactivates the protein. Three NNRTI drugs are available for use today. Table 1 summarizes all currently available RT inhibitors, as well as those drugs targeting other viral proteins.

The next group of ARVs to be discovered targets the HIV protease (PR). These protease inhibitors (PIs), of which there are seven on the market, are even more effective than RTIs at reducing viral load. Advances in structure-based drug design in the early and mid-1990s enabled researchers to create peptidomimetic compounds that fit into the active site of the PR dimer (13). Small peptides that mimicked PR substrates soon evolved into a diverse set of chemically synthesized compounds (45). Like NRTIs, all presently available PIs competitively inhibit enzymatic activity by binding the active site.

An important new class of inhibitors has recently gained much attention in the HIV field and in the media. Prior to its introduction, available ARVs only targeted the virus after it had already infected cells. Thus, initial infection events were permitted, allowing a pool of latently infected cells to accumulate. A drug that prevented the infection process would clearly be a useful complement to the arsenal of compounds that work after infection. In 2003, the first entry inhibitor, T20 or enfuvirtide, gained FDA approval (58). It is a 36 amino acid peptide derived from the C-terminal heptad repeat of gp41, and acts in a dominant negative manner to prevent the virus from achieving membrane fusion (60). A detailed summary of other entry inhibitors currently in development is presented below (see the Inhibiting viral entry section).

The availability and use of ARVs has made a substantial impact on reducing the mortality of AIDS patients. For example, pre- and post-natal administration of these drugs to HIV-positive mothers has greatly reduced the incidence of mother-to-child transmission of HIV (22). In the North America and Europe, AIDS-related hospitalizations and deaths have markedly declined (14).

Drug therapy, however, is not a sustainable solution to the HIV and AIDS pandemic. The most formidable obstacle presented by drug therapy is the development of drug-resistant HIV strains. Initial attempts to reduce viral load with monotherapy quickly led to the emergence of drug-resistant strains, often within weeks of administering the drug (4, 14). As a result, HAART, or highly active antiretroviral therapy, has become the norm for drug therapy regimens (13, 14, 45). A combination of three or more ARVs (from the RTI and PI classes) is co-administered, and

the result has generally been successful. When resistance does appear, as it almost always does, patients simply switch to another drug within the same class. Unfortunately, some patients have run out of options. Potential complications with HAART include antagonistic drug interactions, the emergence of multidrug resistant strains, and numerous noxious side effects. Furthermore, while patients with access to HAART are certainly fortunate, they are subjected to an extremely rigorous and demanding dosage schedule that has undoubtedly contributed to patient noncompliance. Some researchers stress the danger of drug-resistant strains that are emerging as patients abandon their drug regimen. Novel HIV inhibitors, such as enfuvirtide, and experimental drugs targeting other viral enzymes (such as integrase) and processes (such as Rev-mediated nuclear export and viral assembly), will provide much needed hope for attenuating disease progression (45).

Finally, this section has only addressed HIV-specific drugs. Since the ultimate cause of disease and mortality in AIDS patients is opportunistic infection, there are many attempts to target those diseases (such as tuberculosis, Kaposi's sarcoma, and Pneumocystis carinii pneumonia) and the immune system itself. Immune system stimulants such as interferons and cytokines boost host defense responses and inhibit viral replication (46). Two promising candidates, recombinant interleukin 2 (IL2) and interleukin 7 (IL7), have been shown to raise CD4 cell counts and stimulate latent proviruses in AIDS patients (14, 59).

The search for a vaccine

To truly reduce HIV/AIDS-related morbidity and mortality, a preventative vaccine is urgently required. The best-case scenario would be a vaccine that prevents infection in challenged patients, i.e. one that elicits sterilizing immunity. However, there is also significant benefit to a vaccine that boosts the host immune defenses, perhaps to the point where plasma viremia are cleared and persistently infected cells are destroyed. In fact, even a vaccine that enables the body to fully function with a basal level of infection would dramatically reduce the effects of the AIDS pandemic.

There are several approaches being taken to develop an HIV vaccine (these are reviewed in (3, 11, 13, 51, 53)). Subunit vaccines that are composed of recombinant viral proteins, usually Env, are intended to stimulate a humoral (antibody) response. B cells that recognize these viral antigens produce antibodies that neutralize free virions in the bloodstream. The most notable

example of a subunit vaccine—as well as the most advanced candidate in clinical trials to date—is a gp120 vaccine developed by the Genentech spin-off company VaxGen (51). Initial immunization with a single laboratory-grown (or lab-adapted) strain of HIV gp120 did not protect people from HIV infection in Phase II trials and, as a result, gp120 from a second, more virulent, strain was added. This 'bivalent' vaccine, termed AIDSVAX, was tested in clinical trails in the United States and Thailand in the late 1990s and early 2000s. Unfortunately, neither AIDSVAX alone or in combination with other candidate vaccines (termed the *prime-boost* approach, see below) has yielded promising results for the gp120 subunit vaccine. In fact, continuation of AIDSVAX drug trials in the face of negative results has caused some disharmony in the scientific community; the best way to integrate basic science, clinical trials, and procedural issues related to vaccine development is still being debated (12). Currently there are no subunit-only vaccines in development. Attempts to generate an antibody response to the virus are being made through other approaches, such as DNA vaccines (see below).

Vector-based vaccines, on the other hand, stimulate the cellular immune response. They consist of recombinant viruses or bacteria (in both replication-competent and replication-incompetent forms) that express HIV proteins. Host cells infected by these vectors stimulate cytotoxic T lymphocytes by presenting HIV antigens on the cell surface. In addition, vectors circulating in the bloodstream are also capable of eliciting a humoral response. A promising approach combining a vector-based and a subunit vaccine that present the same antigen is currently being investigated in clinical trails. In this 'prime-boost' approach, patients are typically initially immunized with a subunit protein, followed by an immune system 'boost' with the live vector (38, 51, 53). The most popular vectors in use are canarypox- (ex. ALVAC), vaccinia- (ex. MVA), and adenovirus- (ex. Ad5) based.

Other approaches to vaccine development include whole-killed viruses, or live-attenuated viruses. While these approaches have been effective in combating diseases such as polio, rabies, and measles, they are met with reasonable and justified skepticism in the HIV field. Preparations of whole-killed virus usually involve inactivation of live virus by chemicals or irradiation (51). If any infectious virus remains after this treatment, the inoculum may indeed cause HIV infection in naïve patients. One controversial vaccine, Remune, which consists of inactivated HIV, is currently in clinical trials and is being used to stimulate the immune response of HIV-positive individuals (52, 56).

A novel and intriguing type of vaccination is the DNA vaccine. Circular DNA encoding one or more viral proteins under the control of a strong promoter is injected intramuscularly or subcutaneously (21). In a process that is not fully understood, the naked DNA is taken up by cells and integrated into the host genome. Antigenic proteins are then expressed on the cell surface. While the method and repercussions of viral DNA uptake and integration into the host genome are unknown (and may potentially be harmful), many DNA vaccines are currently in various phases of clinical trials (53). Although DNA vaccines alone yield poor cytotoxic and humoral immune responses, when coupled with adjuvants or used in a prime-boost combination they may generate a substantial immune response (2). The major advantage to such an approach is the ease and low cost of preparation; if successful, this method would be extremely useful in developing countries.

Macaques (typically rhesus monkeys) are the most commonly used animal in vaccine studies (24, 51). Simian immunodeficiency virus (SIV) is closely related to HIV, but does not cause AIDS-like disease in the natural host. Macaques infected with certain heterologous, pathogenic SIV strains however do show AIDS-like syndrome. Furthermore, researchers have determined that a chimeric virus, termed SHIV, consisting entirely of SIV proteins with the exception of an HIV envelope, can reliably cause disease in Asian macaques (24). In fact, some extremely virulent SHIV strains, obtained by repeated passaging in animals, cause CD4 T cell depletion within one month of infection and death within three to four months. The rapidity of disease progression, and the fact that these chimeric viruses present the HIV envelope protein, makes this animal model particularly well-suited to vaccine studies.

The experimental parameters and protocols used in animal studies are still being developed. For instance, two factors appear to reproducibly influence the outcome: the time and the mode of challenge following vaccination (16). The immune response is strongest shortly after the final vaccine boost, and therefore most studies are initiated at this time. In addition, mucosal challenge typically elicits a stronger immune response than intravenous challenge (51). Unfortunately, even when both parameters are optimized, few vaccine candidates elicit a response strong enough to combat pathogenic viral strains.

The drawback to using non-human primates as a model system is that they do not accurately reproduce the human immune response. The dose of challenge virus used in monkey studies far exceeds the dose of a typical HIV infection in humans. Thus, the danger is that

potentially useful human vaccines may be undetected if they fail to prevent SHIV infection in monkeys. Furthermore, as a necessity to studying vaccines in monkeys, animals that rapidly develop AIDS-like symptoms (within months) are used. However, the human immune system is able to prolong the onset of AIDS for decades. Thus, again, the requirements for an effective animal vaccine may be more stringent than what is needed for a successful human vaccine.

In addition to these hurdles, HIV vaccine researchers face other significant challenges. Despite considerable efforts, the level and type of immune response necessary to reduce disease progression, i.e. the correlates of immunity, are not known. There is much debate about which arm of the immune system is necessary to achieve protective immunity: the T-cell mediated cytotoxic response or the humoral arm with antibody-mediated immunity (reviewed in (44)). To date, each type of response has demonstrated an ability to control infection as well as vulnerability to viral trickery.

The cellular immune response is of paramount importance in clearing infected cells that are potential factories for virus production. The discouraging AIDSVAX trials revived an interest in cellular immunity. Vaccination of macaques with a modified vaccinia Ankara (MVA) vector has reduced viremia and delayed disease progression upon challenge with SIV (43). Similarly, canarypox vaccines elicit a modest cytotoxic cell response, replication-defective alphavirus vectors target and stimulate antigen-presenting cells, and intranasal delivery of replication-competent adenovirus elicits some systemic and mucosal immunity in primates (38). Interestingly, studies in macaques show that immunization with an attenuated pox vector (expressing SIV Env, Pol, and Gag) in combination with antiretroviral therapy stimulates host immunity (38) more than immunization alone. This opens a new avenue of exploration in which drug suppression of viral replication can be used to sustain and strengthen immune defenses. In addition, viral vectors in a prime-boost combination with DNA or subunit vaccines are being actively studied (44). The most compelling reason to pursue cell-based immune strategies comes from broad-based protection of monkeys immunized with live attenuated SIV (16). Similar experiments with a variety of attenuated SIV strains are showing, however, that the immune system may only be capable of controlling challenge from a homologous strain (17). While it is doubtful that live attenuated vectors will be used as an HIV vaccine, due to safety concerns, the correlates of immunity in these animals are being intensively investigated.

Most existing vaccines (such as those for rabies, influenza, hepatitis A and B, and polio) invoke antibody-based protective immunity (44). The most convincing evidence that this approach will work for HIV is that passive transfer of neutralizing antibodies in macaque monkeys confers protection against subsequent viral challenge (37, 39). Although the level of serum antibody in these experiments exceeds the titer produced by vaccination, these studies have generated much hope and interest in humoral immunity. However, generating an antibody response with broadly neutralizing capability remains a significant challenge. Most antibody responses are specific to the immunization strain, and do not neutralize heterologous, challenge virus (17). Notably, a few gp41 monoclonals provide exception to this general rule, and are being extensively studied (see Inhibiting Viral Entry, below).

The route of infection is yet another complicating factor in developing a vaccine. An ideal vaccine would target and stimulate an immune response where it is most needed. For example, since the majority of HIV transmission is through heterosexual intercourse, mucosal cells lining the vaginal wall are important first defenses in an anti-HIV response. In this case, stimulation of immunoglobulin A (IgA) antibodies in mucosa is essential. HIV can also be transmitted from mother to newborn child during childbirth or breastfeeding. While antiretrovirals have virtually eradicated mother-to-child transmission in developed nations, approximately 30% of HIV-positive African women transmit the virus to their children (32, 55). A combination of drug therapy and a preventative vaccine may reduce this transmission worldwide.

The last two years have seen an unprecedented rise in the number of candidate vaccines entering human clinical trials (28). As of January 2005, there were twelve ongoing HIV vaccine trials and eleven trials in various planning stages (53). Even more encouraging, in 2003, the first clade C-derived candidate vaccine moved into clinical trials (see Figure 1*B* for the worldwide distribution of HIV clades or subtypes). The clade C subtype of HIV predominates in South Africa which has the largest infected population in the world. Until recently, most basic science and clinical trials have been performed with clade B virus, which is more prevalent in developed countries. In order to mount an effective response to HIV and AIDS, scientists will have to study and target the globally predominant viral subtypes.

Viral immune evasion

Researchers face a formidable opponent in the HV virus and, despite over two decades of intensified research efforts, AIDS remains a deadly disease. HIV employs several unique strategies to evade the host immune system, four of which are outlined here.

A staggering mutation rate is one of the virus's primary defense mechanisms. Approximately 10^9 to 10^{10} virions are generated every day in an infected patient, and the mutation rate, due to sloppy proofreading capability of RT, is about 3×10^{-5} per nucleotide per replication cycle (30). This allows the virus to continually experiment with different combinations of mutations and, not surprisingly, is the major reason for the rapid evolution of drug resistant strains. In addition, there is another twist in the virus's ability to mutate: recombination between variant strains in a single infected host. This *in vivo* recombination occurs in doubly infected cells and can lead to great genetic variability. The global emergence of these circulating recombinant forms is a significant and threatening source of multi-drug resistant viruses.

Furthermore, HIV primarily infects CD4 T cells, macrophage, and dendritic cells, the very cells that normally effect the immune system's response to a pathogen (13, 30). Early in infection, viral strains are typically macrophage- or M- tropic. They utilize the CCR5 coreceptor, and are nonsyncytium inducing. Later, just before the patient develops full-blown AIDS, the virus often evolves into a more virulent T-tropic strain that infects CD4 T cells, uses the CXCR4 coreceptor, and induces syncytia formation (18). Depletion of these immune cells, directly and indirectly caused by viral infection as well as by cytotoxic responses, weakens the immune response and leaves the host vulnerable to opportunistic infections.

HIV is capable of infecting nondividing cells and the provirus can remain quiescent for long periods of time. Antiretrovirals are able to suppress active viral replication, and a vaccine may potentially destroy dividing infected cells. But, the reservoir of persistently infected cells is invisible to the immune system and will always be able to produce progeny virus (13). Ideally, a long-term solution to fighting HIV infection will involve destruction of this deadly pool of cells.

Finally, the most visible portion of the virus to the humoral immune system is the envelope glycoprotein and the majority of anti-HIV antibodies are directed against Env. Extensive glycosylation of gp120, accounting for about half of its molecular weight, shields

important functional domains, such as receptor and coreceptor binding sites, from antibody binding. As a result, the ability of the humoral immune system to generate neutralizing antibodies is greatly hindered.

Inhibiting viral entry

One of the most powerful ways to prevent the spread of the HIV pandemic is to inhibit early infection events. As stated above, once infection is established, it is extremely difficult to remove the pool of latently infected cells from the body. This section discusses the different avenues being explored to impair the function of the envelope protein, as summarized in Figure 2.

Since gp120 facilitates initial contact between virus and host cell, it is a logical target for anti-HIV drugs. gp120 contains five regions of variable sequence (V1-V5) that are interspersed with five constant domains (C1-C5) (Figure 3). Initial adhesion of the virion to host cells occurs through an ionic interaction between the positively charged gp120 V3 loop and the negatively charged cellular heparin sulfate proteoglycans, lectins, and adhesion factors (41). Several polyanionic and cationic molecules block viral attachment *in vitro*, but only one, AR177 or zintevir, has progressed to clinical trials so far (45). Semi-specific inhibitors of attachment may be enormously useful as topical microbicides.

The binding site for CD4, the major cell surface HIV receptor, is a recessed pocket that is shielded by heavily glycosylated variable regions. The gp120 C3 and C4 domains make several contacts with CD4, but the entire binding surface is discontinuous and contains other gp120 regions. Initial attempts were made to block the CD4-gp120 interaction using soluble CD4 (sCD4); however, this monomeric molecule only exhibited antiviral activity against laboratory-adapted HIV strains, and not the more virulent clinical isolates (41). A multimer of soluble CD4 domains tethered together via a tetrameric immunoglobulin domain led to the development of CD4-IgG2 or PRO 542, an antagonist of the CD4-gp120 interaction (29). This molecule is currently in drug trials. In addition, monoclonal antibodies against CD4 also inhibit fusion. One example is TNX-355, which is administered intravenously and inhibits a broad range of clinical isolates (45). In 2003, the compound BMS-378806 was reported to bind specifically to the CD4 binding pocket of gp120, based on competitive inhibition experiments (26). Subsequent studies reported resistant mutations and indicated that BMS-378806 may bind near, but not directly in, the CD4-binding pocket and may perhaps influence gp120-gp41 interactions (48). A recent

structure of SIV gp120 in its native, unliganded form (9, 10) suggests that this compound likely binds a groove in gp120 that eventually forms the CD4-binding pocket. BMS-378806 is currently being evaluated in clinical trials. Finally, cyanovirin-N (CV-N), a protein from the cyanobacterium *N. ellipsosporum*, is another potent viral inhibitor. It targets specific N-linked oligosaccharides on gp120 and an encouraging result shows little development of resistance to this protein (45).

Inhibiting gp120-coreceptor interaction is an attractive option for preventing viral entry, since a small group of people with a CCR5 mutant allele are extremely resistant to HIV infection yet are otherwise healthy (see Rare cases of resistance, below). There are about twelve coreceptors for HIV *in vitro*, although two (CCR5 and CXCR4) are most commonly used *in vivo* (41). These molecules belong to the seven-transmembrane-spanning G-protein-coupled receptor family and function by binding chemokines and sending downstream signals that dictate cellular migration patterns. While the natural chemokines (such as RANTES, MIP-1α and -1β for CCR5 and SDF1 for CXCR4) can potentially inhibit fusion, their function is to downregulate their receptors and initiate cellular signaling (41). These undesired effects make such ligands unpromising anti-HIV drug candidates; nonantagonistic chemokine derivatives and antibodies are better options. Two promising antagonists of the gp120-CCR5 interaction are PRO 140, an anti-CCR5 antibody, and TAK-779, a small molecule that binds a CCR5 cavity and induces conformational changes that preclude gp120 binding (41, 45). Neither compound initiates downstream cellular signaling. Similarly, a CXCR4 antagonist, the bicyclam AMD3100, is showing promising results in early clinical trials. This molecule may prevent the virus from switching from the less pathogenic CCR5 coreceptor to CXCR4, as is often seen immediately prior to the onset of AIDS (18). An AMD3100 derivative, AMD3465, is even more potent at inhibiting CXCR4 virions (45). Blocking gp120-coreceptor interactions is indeed a promising avenue to explore, but some caution will be necessary so that directed viral evolution toward a more virulent CXCR4-tropic strain is not an unwelcome byproduct.

The process of gp41-mediated viral fusion is well studied. Mutagenic studies, such as the one described in Chapter 3, have highlighted critical residues and regions necessary for gp41 function. Using the available gp41 x-ray structures, potent peptide inhibitors, such as IQN26 (Chapter 2) have been designed to specifically target fusion pathway intermediates. gp41 epitope-mapping and *in vitro* biochemical analyses are also used to further dissect the fusion

pathway. A detailed understanding of gp41 conformational changes during fusion is emerging, and is being used to design gp41 inhibitors.

Enfuvirtide, or T-20, has been a promising and welcome addition to the antiretroviral lineup (see the Anti-HIV Drugs section, above). While this entry inhibitor, corresponding to a 36 amino acid region of the gp41 C-HR, is the first of its kind on the market, there are numerous candidates in the pipeline. A second-generation peptide inhibitor that partially overlaps T-20, called T-1249, is ten times more potent *in vitro* (45). T-1249 is a 39 amino acid peptide composed of C-HR sequences derived from HIV-1, HIV-2, and SIV (20). In January 2004, ongoing early clinical trials of this drug were halted in order to improve the formulation of the product; they are expected to resume once production concerns are resolved (57). In addition, numerous small molecule entry inhibitors are being designed and screened for the potential to bind the gp41 cavity (15, 19, 23, 31). Although none are currently in clinical trials, these non-peptidic small molecules may be better suited for antiretroviral therapy.

Systematic analyses of gp41 have been used to generate chimeric, helical N-peptides, such as IQN26, that have inhibitory activity *in vitro*. One interesting molecule, consisting of three N-helices surrounded by two C-helices, forms a stable six-helix-bundle-like structure that exposes a single C-helix binding site. This recombinant protein, 5-helix, is another potent *in vitro* inhibitor (47). Limited bioavailability and the need to inject peptide inhibitors once or twice daily may hinder the potential of 5-helix and other recombinant peptides as long-term, mainstream fusion inhibitors. However, they may be used as leads to develop nonpeptidic derivatives or as mimics of gp41 intermediates that could serve as screening agents for more effective compounds. In addition, modified versions of such compounds may be useful as immunogens (47).

Env is the most visible viral protein to the immune system; it is therefore the subject of many vaccine trials and studies. Antibodies can neutralize the virus by binding to important functional domains and thus sterically blocking virus-cell interactions. Most HIV antibodies generated by natural infection or immunization have limited or no neutralizing capability. Antibodies that do neutralize a broad spectrum of clades are rare, and once found, their epitopes are difficult to reproduce. There are currently five known broadly neutralizing antibodies against the Env glycoprotein, all derived from individual HIV patients (reviewed in (6)). The first one identified was b12, a gp120-binding antibody that occludes the CD4 binding site via a long protruding loop on the antigen-binding fragment Antibody 2G12 also binds gp120. It is unique,

however, in that it recognizes not only peptidic but also glycan moieties on the gp120 surface. Three other monoclonal antibodies, 2F5, 4E10, and the very similar Z13, recognize adjacent epitopes of the gp41 C-HR region. These antibodies appear to bind more tightly to the native conformation of gp41 rather than the CD4-triggered conformation (40). A recent crystal structure of 4E10 with a 13 residue gp41 peptide shows this tryptophan-rich C-HR region to be helical (7). To date, 4E10 is the most broadly neutralizing antibody known. Since designed immunogens mimicking the C-HR have repeatedly failed to elicit a neutralizing response, identification of the 4E10 epitope structure will certainly aid vaccine design efforts.

The potential to target antibodies to gp41 intermediates, such as the prehairpin structure, has been tantalizing yet controversial. This approach could face one of two major potential problems. The first is the question of steric hindrance—i.e., can an antibody fit in the space between virus and cell, and can it access exposed N-HR or C-HR regions in the presence of gp120 and nearby cell-surface proteins? Some evidence, using a series of cargo proteins attached to a potent inhibitor that binds the N-HR region, indicates that a ~150 kDa antibody may be sterically occluded from binding the N-HR region (27) . Second, the transient exposure of the prehairpin intermediate could kinetically prevent antibody binding. Some groups report *in vitro* binding and neutralizing activity of anti-N-HR antibodies only at suboptimal temperatures (25). However, antibodies from HIV-positive individuals react with peptide mimics of the N-HR region (very similar to IQN26, described in Chapter 2), some of which may possess multi-strain neutralizing capability (27, 42). gp41 is indeed a difficult protein to biochemically characterize, given its labile native structure (that is largely shielded by gp120) and its dramatic conformational transitions. While the N-HR region is an attractive target due to its sequence conservation and the presence of the hydrophobic cavity, a well-established drug target, scientists will need to become increasingly creative to generate a broadly neutralizing response to these transiently exposed structures.

The biophysical work presented in Chapter 2 suggests that the N-HR region may not be a uniform, stable trimer throughout its length. Whereas hybrid peptides containing the hydrophobic cavity are better inhibitors when they are extremely thermostable and trimeric, a chimeric peptide containing the N26 region only interacted with C-HR and inhibited fusion in a less stable form (see Chapter 2). If further biochemical studies confirm that the N26 region is a weak trimer, perhaps inhibitors that intercalate into this region of the N-HR could be designed. Even if they

are not effective immunogens, hybrid peptides that accurately mimic gp41 intermediates could greatly facilitate drug screening.

Furthermore, this chimeric peptide strategy need not be limited to N-HR-derived peptides. The few broadly neutralizing antibodies available today all bind the C-HR region in a native conformation (see above). Since structural analysis of antibody 4E10 shows its C-peptide epitope to be helical, hybrid molecules similar to IQN26 that stabilize the C-peptide may be even more useful immunogens.

The mutagenesis experiments described in Chapter 3 show that six-helix bundle formation drives membrane fusion. Further confirmation of this model will come from a mutant that is more fusogenic than wild-type Env and also has a more stable six-helix bundle structure. In addition, the effect of these substitutions on gp41 structure can be better studied through the use of longer N- and C- peptides that are not covalently attached to each other (unlinking the peptides may potentially reveal improper helix alignment). The N-HR glycine, alanine, and phenylalanine mutants described in Chapter 3 are likely to be arrested at the prehairpin intermediate state. This hypothesis could be confirmed by assessing fusion kinetics in the presence of C-peptide inhibitors or by immunoprecipitation with antibodies directed against non-cavity N-HR residues (the mutations are all within the cavity). If the prehairpin intermediate is indeed sterically accessible to antibodies (see above), then a vector-based vaccine containing an Env mutation (similar to those in Chapter 3) may allow prolonged exposure of this intermediate to the immune system. Given the difficulties inherent in generating a prehairpin-like immunogen, viral vectors that present prehairpin-arrested Env may be worth exploring. The glutamine layer is an intriguing target, and further mutagenesis and biochemical studies will be useful in elucidating its role in gp41-mediated fusion. This highly polar segment of the mainly hydrophobic N-C interface may be an attractive drug target.

Several important questions in the membrane fusion field remain unresolved. For example, how does the prehairpin intermediate actually resolve into the six-helix bundle structure? There is an inherent problem of topology. If all three C peptides extend toward the target membrane symmetrically, 120° apart from each other, and simultaneously interact with the trimeric N-HR, the viral membrane would surround the N-C linker and the newly-formed six-helix bundle. However, studies with influenza HA and HIV Env suggest that the six-helix bundle remains anchored on the outside of an infected cell (50, 61). It is therefore likely that N-HR

interacts with the three C-helices sequentially. If this occurs, it is unknown whether all three C helices, or just one or two, have to interact with the N-HR in order for the attached membranes to merge. Some studies suggest that six-helix bundle formation may even occur after formation of a small pore between virus and cell (36). An understanding of this process may lead to better design of inhibitors; for instance, tethering two or more C peptides together may increase their potency.

A monumental advance in the membrane fusion field will come from the determination of the gp41 native state structure, and if at all possible, the prehairpin intermediate. Undoubtedly, much of native gp41 is shielded by gp120 and a few C-terminal glycan moieties. However, the most potently neutralizing HIV antibodies to date target the C-HR of gp41, most likely in a native conformation. A native state structure would provide critical hints about the conformational transitions that gp41 undergoes to reach the six-helix bundle. The prehairpin intermediate is perhaps the most promising gp41 target, and it will be important to definitively answer the question of steric accessibility. Molecules that mimic its transiently exposed trimeric structure will serve as immunogens and useful screening agents. Hybrid peptide strategies such as the one discussed in Chapter 2, are currently being used and can be improved upon. For example, measures can be taken to reduce the immune reactivity of the scaffold peptide moiety of these immunogens, such as by surface glycosylation. Alternatively, the outer surface of the scaffold peptide can be engineered to more closely resemble the gp41 trimer (such as in (49)). Atomic resolution structures of Env will certainly advance the rational design of anti-HIV drugs with improved binding to the native or intermediate gp41 structures.

Rare cases of resistance

During the 8—10-year period of clinical latency, there is usually a steady decline in the level of circulating CD4 T cells. However, in some rare cases, this decline is not observed, and patients can live up to twenty years or more after infection. The basis for these long-term nonprogressors is unknown. Some sex workers in Kenya and Gambia, for instance, have remained persistently seronegative despite continuous exposure to the virus (30, 34). They have a vigorous cytotoxic immune response that often recognizes unmapped epitopes. Although some 'resistant' workers eventually do succumb to the disease, they are being carefully studied and provide hope that a vaccine may confer substantial immunity to HIV and AIDS.

Additionally, individuals homozygous for a 32 base pair deletion in CCR5 (Δ32 allele) are highly HIV-resistant, while heterozygotes demonstrate slower progression to AIDS than normal (1). The 32 base pair deletion occurs in an extracellular loop between transmembrane regions four and five of the seven-transmembrane chemokine receptor, and prevents cell surface expression of the mutant protein. CCR5Δ32 may also interact with and downregulate surface expression of other HIV coreceptors (such as CXCR4), further enhancing cellular resistance to viral infection.

Finally, infection with HIV-2, a less virulent and transmissible virus that causes slower disease progression, has been shown to reduce the progression and severity of HIV-1 associated disease (30). Some studies indicate that the basis for this protection is that HIV-2 stimulates production of β-chemokines that are CCR5 ligands (35). Thus, establishment of infection with M-tropic HIV-1 strains that utilize the CCR5 coreceptor is reduced. The humoral and cellular responses to HIV-2, and their cross-reactivity with HIV-1, are under investigation.

Successful strategies

Rising mortality rates and an awareness of the serious health and economic impact caused by AIDS are fueling many significant global efforts to combat the epidemic. A 2003 AIDS Pipeline Project group meeting cites cooperation between many international health agencies. Among them are The Global Fund to Fight AIDS, TB, and Malaria, the Millennium Challenge Account, the World Health Organization 3 by 5 Initiative, the Global Alliance for Vaccines and Immunization, and the International AIDS Vaccine Initiative.

Uganda has been heralded in the media for turning around its AIDS crisis and dramatically lowering infection rates despite its limited resources (33, 54). The ABC approach of 'Abstinence, Be Faithful, and Condomize' has successfully spread the message of prevention to younger generations of Ugandans. The president, Yoweri Museveni, and other high-level government officials have made HIV/AIDS prevention, care, and treatment programs a top priority. Indeed, individuals from all sectors of society—religious leaders, school teachers, community groups—have committed to action. The Ugandan fight against AIDS involves sex education programs in schools, aggressive marketing of condoms (at prices subsidized by the government and other health organizations), and promotion of an STI (sexually transmitted infections) self-treatment kit. Same-day testing and counseling services, launched in 1997, have

not only become widely used, but are also a way to recruit and train peer educators. Importantly, these efforts have also succeeded in lowering the barriers of social stigma that so often hinder care and treatment programs.

The results have been remarkable. Since the early 1990s, infection rates among pregnant women have declined by more than 50% in some areas of the country. According to the World Health Organization, HIV infection rates among men attending STI clinics fell 16% between 1992 and 1998. Condom use is higher than 50% in rural areas, and over 85% in urban areas. The prevalence of HIV in teenage girls has dropped from approximately 4.5% in 1990 to less than 1% in 1997 (54).

The encouraging downward trend in HIV infection is not limited to Uganda. Thailand is also controlling the spread of HIV by social awareness programs and condom marketing. In both Thailand and Senegal, sex workers are reporting increased condom usage.

Even in these countries, of course, much work remains to be done—poverty, gender inequality, and the growing number of AIDS orphans are serious problems. Mother-to-child HIV transmission can be dramatically reduced in developing countries with ARVs, as it is in the North America and Western Europe. Drug and vaccine distribution also poses some obstacles. Many developing nations lack the infrastructure to provide HIV testing and counseling, treatment, and in the future, vaccines, to infected people. Currently, global health and funding organizations are making progress toward addressing this issue. In addition, despite the low profitability of developing HIV vaccines, some drug companies are nevertheless conducting clinical trials and establishing partnerships in developing nations to eventually distribute ARVs and vaccines.

The dedication and leadership of political, community, and health groups in countries like Uganda powerfully demonstrates that concerted effort can alleviate the AIDS crisis. The success of these approaches, combined with continual scientific progress in HIV drug and vaccine development, provide much hope that the AIDS pandemic will someday fade into the pages of human history.

Future prospects

After several years of conducting research on gp41, I have come to appreciate the complexity of the virus and, especially, the entry process. In terms of targeting the virus in order to curtail the spread of this pandemic, I believe that viral entry has the most promising potential. This is because: (1) No other HIV protein is as visible to the immune system, or as immunogenic. Thus, Env can be effectively attacked on two fronts, as a drug and as a vaccine target, to mitigate the effects of resistant mutations. (2) Reducing initial fusion events in a newly infected person will greatly reduce the number of latently infected cells. I see these cells as one of the major obstacles to truly 'curing' AIDS. Unless they are removed, there is always the potential to destroy CD4 cells and develop immune deficiency. (3) Portions of Env, mostly the gp41 rev response element-encoding sequence and the N- and C- heptad repeats (including the glutamine layer residues) are highly conserved among the diverse clades. Therefore, we have the best chance to use the same antiviral strategy for diverse HIV strains if this protein is targeted.

Every step taken toward understanding the mechanistic details of the fusion machinery is vital, applicable, and necessary. Mutagenic studies (Chapter 3) indicate that even conservative mutations in this fusion machinery can greatly impair function. Furthermore, this study suggests a novel gp41 subdomain, the glutamine layer, may be critical to gp41 activity. It is intriguing that a potent peptide inhibitor, T20, can be made tenfold more effective by reengineering its sequence to include conserved HIV-1, HIV-2, and SIV residues (and by lengthening it by three amino acids). Perhaps by combining these two approaches—mutagenesis to identify critical residues and sequence alignment to identify conservation—HIV antigens and peptide inhibitors can become more potent.

An HIV vaccine that confers sterilizing immunity is highly unlikely. Encouraging results with live attenuated SIV and passive antibody transfer in monkeys suggest that a vaccine that reduces disease progression is attainable. A truly effective vaccine will have to stimulate both the cytotoxic and humoral arms of the immune system. The most promising avenue for vaccination is the prime-boost approach, although this method will require administering two separate immunizations several months apart. In addition, it will not be possible to generate a single-strain vaccine that effectively neutralizes a broad spectrum of clades. A broadly-neutralizing vaccine will likely contain components of at least two HIV strains or clades, an approach taken each year with the flu vaccine (8).

On a global scale, increased efforts and resources are being devoted to HIV/AIDS prevention and education programs, distribution of ARVs, and the establishment of care and treatment facilities. The substantial scientific research efforts outlined here, aimed at understanding the viral fusion machinery, will also certainly contribute to the fight against HIV and AIDS. Progress in this direction has already led to the development of antiviral drugs that ease suffering and prolong lives. Now, a vaccine that reduces disease progression and transmission is the next challenge, and perhaps the most powerful way to mitigate the impact of HIV and AIDS worldwide.

References

1. **Agrawal, L., X. Lu, J. Qingwen, Z. VanHorn-Ali, I. V. Nicolescu, D. H. McDermott, P. M. Murphy, and G. Alkhatib.** 2004. Role for CCR5{Delta}32 protein in resistance to R5, R5X4, and X4 human immunodeficiency virus type 1 in primary CD4+ cells. J. Virol. **78:**2277-2287.
2. **Amara, R. R., F. Villinger, J. D. Altman, S. L. Lydy, S. P. O'Neil, S. I. Staprans, D. C. Montefiori, Y. Xu, J. G. Herndon, L. S. Wyatt, M. A. Candido, N. L. Kozyr, P. L. Earl, J. M. Smith, H. L. Ma, B. D. Grimm, M. L. Hulsey, J. Miller, H. M. McClure, J. M. McNicholl, B. Moss, and H. L. Robinson.** 2001. Control of a mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. Science **292:**69-74.
3. **Baltimore, D., and C. Heilman.** 1998. HIV vaccines: prospects and challenges. Scientific American:98-103.
4. **Bartlett, J. G., and R. D. Moore.** 1998. Improving HIV therapy. Sci Am **279:**84-7, 89.
5. **Binley, J. M., R. W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D. J. Anselma, P. J. Maddon, W. C. Olson, and J. P. Moore.** 2000. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. J Virol **74:**627-43.
6. **Burton, D. R., R. C. Desrosiers, R. W. Doms, W. C. Koff, P. D. Kwong, J. P. Moore, G. J. Nabel, J. Sodroski, I. A. Wilson, and R. T. Wyatt.** 2004. HIV vaccine design and the neutralizing antibody problem. Nat Immunol **5:**233-6.
7. **Cardoso, R. M., M. B. Zwick, R. L. Stanfield, R. Kunert, J. M. Binley, H. Katinger, D. R. Burton, and I. A. Wilson.** 2005. Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. Immunity **22:**163-73.
8. **Centers for Disease Control and Prevention.** 2005. Flu vaccine: questions & answers. http://www.cdc.gov/flu/about/qa/fluvaccine.htm.
9. **Chen, B., E. M. Vogan, H. Gong, J. J. Skehel, D. C. Wiley, and S. C. Harrison.** 2005. Determining the structure of an unliganded and fully glycosylated SIV gp120 envelope glycoprotein. Structure **13:**197-211.
10. **Chen, B., E. M. Vogan, H. Gong, J. J. Skehel, D. C. Wiley, and S. C. Harrison.** 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. Nature **433:**834-841.

11. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

12. **Cohen, J.** 2005. Military vaccine studies: on trial, p. 32-33, MIT Technology Review, vol. 108.

13. **Cohen, O. J., and A. S. Fauci.** 2001. *Fields Virology*, Fourth ed, vol. 2. Lippincott Williams & Wilkins, Philadelphia.

14. **Darbyshire, J.** 2000. Therapeutic interventions in HIV infection - a critical view. Trop Med Int Health **5:**A26-31.

15. **Debnath, A. K., L. Radigan, and S. Jiang.** 1999. Structure-based identification of small molecule antiviral compounds targeted to the gp41 core structure of the human immunodeficiency virus type 1. J Med Chem **42:**3203-9.

16. **Desrosiers, R. C.** 2001. *Fields Virology*, Fourth ed, vol. 2. Lippincott Williams & Wilkins, Philadelphia.

17. **Desrosiers, R. C.** 2004. Prospects for an AIDS vaccine. **10:**221-223.

18. **Dragic, T.** 2001. An overview of the determinants of CCR5 and CXCR4 co-receptor function. J Gen Virol **82:**1807-14.

19. **Eckert, D. M., V. N. Malashkevich, L. H. Hong, P. A. Carr, and P. S. Kim.** 1999. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. Cell **99:**103-15.

20. **Eron, J. J., R. M. Gulick, J. A. Bartlett, T. Merigan, R. Arduino, J. M. Kilby, B. Yangco, A. Diers, C. Drobnes, R. DeMasi, M. Greenberg, T. Melby, C. Raskino, P. Rusnak, Y. Zhang, R. Spence, and G. D. Miralles.** 2004. Short-term safety and antiretroviral activity of T-1249, a second-generation fusion inhibitor of HIV. J Infect Dis **189:**1075-83.

21. **Estcourt, M. J., A. J. McMichael, and T. Hanke.** 2004. DNA vaccines against human immunodeficiency virus type 1. Immunol Rev **199:**144-55.

22. **Family Health International.** 2004. Reducing mother-to-child transmission (MTCT) of HIV. http://www.fhi.org/en/HIVAIDS/pub/fact/reducingmtct.htm.

23. **Ferrer, M., T. M. Kapoor, T. Strassmaier, W. Weissenhorn, J. J. Skehel, D. Oprian, S. L. Schreiber, D. C. Wiley, and S. C. Harrison.** 1999. Selection of gp41-mediated HIV-1 cell entry inhibitors from biased combinatorial libraries of non-natural binding elements. Nat Struct Biol **6:**953-60.

24. **Freed, E. O., and M. A. Martin.** 2001. *Fields Virology*, Fourth ed, vol. 2. Lippincott Williams & Wilkins, Philadelphia.

25. **Golding, H., M. Zaitseva, E. de Rosny, L. R. King, J. Manischewitz, I. Sidorov, M. K. Gorny, S. Zolla-Pazner, D. S. Dimitrov, and C. D. Weiss.** 2002. Dissection of human immunodeficiency virus type 1 entry with neutralizing antibodies to gp41 fusion intermediates. J. Virol. **76:**6780-6790.

26. **Guo, Q., H. T. Ho, I. Dicker, L. Fan, N. Zhou, J. Friborg, T. Wang, B. V. McAuliffe, H. G. Wang, R. E. Rose, H. Fang, H. T. Scarnati, D. R. Langley, N. A. Meanwell, R. Abraham, R. J. Colonno, and P. F. Lin.** 2003. Biochemical and genetic characterizations of a novel human immunodeficiency virus type 1 inhibitor that blocks gp120-CD4 interactions. J Virol **77:**10528-36.

27. **Hamburger, A. E., S. Kim, B. D. Welch, and M. S. Kay.** 2005. Steric accessibility of the HIV-1 gp41 N-trimer region. J. Biol. Chem. **280:**12567-12572.

28. **International AIDS Vaccine Initiative.** 2005. http://www.iavi.org.

29. **Jacobson, J. M., R. J. Israel, I. Lowy, N. A. Ostrow, L. S. Vassilatos, M. Barish, D. N. Tran, B. M. Sullivan, T. J. Ketas, T. J. O'Neill, K. A. Nagashima, W. Huang, C. J. Petropoulos, J. P. Moore, P. J. Maddon, and W. C. Olson.** 2004. Treatment of advanced human immunodeficiency virus type 1 disease with the viral entry inhibitor PRO 542. Antimicrob Agents Chemother **48:**423-9.

30. **Janeway, C. A., P. Travers, M. Walport, and J. D. Capra.** 1999. *Immunobiology.* Current Biology Publications and Garland Publishing, London and New York.

31. **Jiang, S., and A. K. Debnath.** 2000. A salt bridge between an N-terminal coiled coil of gp41 and an antiviral agent targeted to the gp41 core is important for anti-HIV-1 activity. Biochem Biophys Res Comm **270:**153-157.

32. **Joint United Nations Programme on HIV/AIDS (UNAIDS).** 2005. Resources: FAQs about HIV and AIDS. http://www.unaids.org/en/Resources/FAQ.asp.

33. **Joint United Nations Programme on HIV/AIDS (UNAIDS).** 2004. UNAIDS at Country Level: Progress Report. http://www.unaids.org.

34. **Kaul, R., S. L. Rowland-Jones, J. Kimani, K. Fowke, T. Dong, P. Kiama, J. Rutherford, E. Njagi, F. Mwangi, T. Rostron, J. Onyango, J. Oyugi, K. S. MacDonald, J. J. Bwayo, and F. A. Plummer.** 2001. New insights into HIV-1 specific cytotoxic T-lymphocyte responses in exposed, persistently seronegative Kenyan sex workers. Immunol Lett **79:**3-13.

35. **Kokkotou, E. G., J.-L. Sankale, I. Mani, A. Gueye-Ndiaye, D. Schwartz, M. E. Essex, S. Mboup, and P. J. Kanki.** 2000. In vitro correlates of HIV-2-mediated HIV-1 protection. PNAS **97:**6797-6802.

36. **Markosyan, R. M., F. S. Cohen, and G. B. Melikyan.** 2003. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. Mol. Biol. Cell **14**:926-938.

37. **Mascola, J. R.** 2002. Passive transfer studies to elucidate the role of antibody-mediated protection against HIV-1. Vaccine **20**:1922-5.

38. **Mascola, J. R., and G. J. Nabel.** 2001. Vaccines for the prevention of HIV-1 disease. Curr Opin in Immunol **13**:489-494.

39. **Mascola, J. R., G. Stiegler, T. C. VanCott, H. Katinger, C. B. Carpenter, C. E. Hanson, H. Beary, D. Hayes, S. S. Frankel, D. L. Birx, and M. G. Lewis.** 2000. Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. Nat Med **6**:207-10.

40. **McGaughey, G. B., G. Barbato, E. Bianchi, R. M. Freidinger, V. M. Garsky, W. M. Hurni, J. G. Joyce, X. Liang, M. D. Miller, A. Pessi, J. W. Shiver, and M. J. Bogusky.** 2004. Progress towards the development of a HIV-1 gp41-directed vaccine. Curr HIV Res **2**:193-204.

41. **Moore, J. P., and M. Stevenson.** 2000. New targets for inhibitors of HIV-1 replication. Nat Rev Mol Cell Biol **1**:40-9.

42. **Opalka, D., A. Pessi, E. Bianchi, G. Ciliberto, W. Schleif, M. McElhaugh, R. Danzeisen, R. Geleziunas, M. Miller, and D. M. Eckert.** 2004. Analysis of the HIV-1 gp41 specific immune response using a multiplexed antibody detection assay. J Immunol Methods **287**:49-65.

43. **Ourmanov, I., C. R. Brown, B. Moss, M. Carroll, L. Wyatt, L. Pletneva, S. Goldstein, D. Venzon, and V. M. Hirsch.** 2000. Comparative efficacy of recombinant modified vaccinia virus Ankara expressing simian immunodeficiency virus (SIV) Gag-Pol and/or Env in macaques challenged with pathogenic SIV. J. Virol. **74**:2740-2751.

44. **Pantaleo, G., and R. A. Koup.** 2004. Correlates of immune protection in HIV-1 infection: what we know, what we don't know, what we should know. **10**:806-810.

45. **Pereira, C. F., and J. T. Paridaen.** 2004. Anti-HIV drug development--an overview. Curr Pharm Des **10**:4005-37.

46. **Pett, S. L., and A. D. Kelleher.** 2003. Cytokine therapies in HIV-1 infection: present and future. Expert Rev Anti Infect Ther **1**:83-96.

47. **Root, M. J., M. S. Kay, and P. S. Kim.** 2001. Protein design of an HIV-1 entry inhibitor. Science **291**:884-8.

48. **Si, Z., N. Madani, J. M. Cox, J. J. Chruma, J. C. Klein, A. Schon, N. Phan, L. Wang, A. C. Biorn, S. Cocklin, I. Chaiken, E. Freire, A. B. Smith, 3rd, and J. G. Sodroski.** 2004. Small-molecule inhibitors of HIV-1 entry block receptor-induced conformational changes in the viral envelope glycoproteins. PNAS **101:**5036-41.

49. **Sia, S. K., and P. S. Kim.** 2003. Protein grafting of an HIV-1-inhibiting epitope. PNAS **100:**9756-61.

50. **Skehel, J. J., and D. C. Wiley.** 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem **69:**531-69.

51. **Snow, B.** ed. 1999. HIV vaccine handbook. AIDS Vaccine Advocacy Coalition (AVAC), Washington DC.

52. **The Immune Response Corporation.** 2005. Remune. http://www.imnr.com/.

53. **The Pipeline Project.** 2005. HIV vaccines in development. http://chi.ucsf.edu/vaccines/.

54. **The World Health Organization.** 2005. Health: a key to prosperity. Success stories in developing countries. http://www.who.int/en/.

55. **The World Health Organization.** 2005. Mother-to-child transmission of HIV. http://www.who.int/reproductive-health/rtis/MTCT/index.htm.

56. **Treatments for HIV & AIDS.** 2004. AIDSmeds.com. http://www.aidsmeds.com/List.htm.

57. **Trimeris Inc.** 2005. T-1249: Expanding HIV treatment options. http://www.trimeris.com/pipeline/t-1249.html.

58. **U.S. Food and Drug Administration.** 2005. Drugs@FDA. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/.

59. **Wang, F. X., Y. Xu, J. Sullivan, E. Souder, E. G. Argyris, E. A. Acheampong, J. Fisher, M. Sierra, M. M. Thomson, R. Najera, I. Frank, J. Kulkosky, R. J. Pomerantz, and G. Nunnari.** 2005. IL-7 is a potent and proviral strain-specific inducer of latent HIV-1 cellular reservoirs of infected individuals on virally suppressive HAART. J Clin Invest **115:**128-37.

60. **Wild, C., D. Shugars, T. Greenwell, C. McDanal, and T. Matthews.** 1994. Peptides corresponding to a predictive {alpha}-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. PNAS **91:**9770-9774.

61. **Zwick, M. B., E. O. Saphire, and D. R. Burton.** 2004. gp41: HIV's shy protein. **10:**133-134.

Table 1: Summary of anti-HIV drugs [a]

Drug Name	Trade Name	Target / Class [c]	Company	Year approved
Zidovudine (AZT, ZDV)	Retrovir	Reverse Transcriptase / NRTI (nucleoside inhibitors)	GlaxoSmithKline	1987
Didanosine (ddI)	Videx		Bristol-Myers Squibb	1991
Zalcitabine (ddC)	Hivid		Hoffmann - La Roche	1992
Stavudine (d4T)	Zerit		Bristol-Myers Squibb	1994
Lamivudine (3TC)	Epivir		GlaxoSmithKline	1995
Abacavir (ABC)	Ziagen		GlaxoSmithKline	1998
Emtricitabine (FTC)	Emtriva		Gilead Sciences	2003
Tenofovir	Viread	NtRTI (nucleotide inhibitor)	Gilead Sciences	2001
Nevirapine (NVP)	Viramune	Reverse Transcriptase / NNRTI (non-nucleoside inhibitors)	Boehringer Ingelheim	1996
Delavirdine (DEL) [b]	Rescriptor		Agouron Pharmaceuticals	1997
Efavirenz (EFV)	Sustiva		Bristol-Myers Squibb	1998
Saquinavir (SQV)	Invirase, Fortovase	Protease	Hoffmann - La Roche	1995
Indinavir (IND)	Crixivan		Merck Pharmaceuticals	1996
Ritonavir (RIT)	Norvir		Abbott Laboratories	1996
Nelfinavir (NEL)	Viracept		Agouron Pharmaceuticals	1997
Amprenavir (AMP)	Agenerase		GlaxoSmithKline	1999
Lopinavir, Ritonavir	Kaletra		Abbott Laboratories	2000
Atazanavir	Reyataz		Bristol-Myers Squibb	2003
Enfuvirtide (T-20)	Fuzeon	Env	Hoffmann - LaRoche	2003

[a] This list is a compilation of data from the U.S. Food and Drug Administration (www.accessdata.fda.gov, reference 58), the AIDS Education Global Information System (www.aegis.com), and from reference 14.

[b] Only available in the United States.

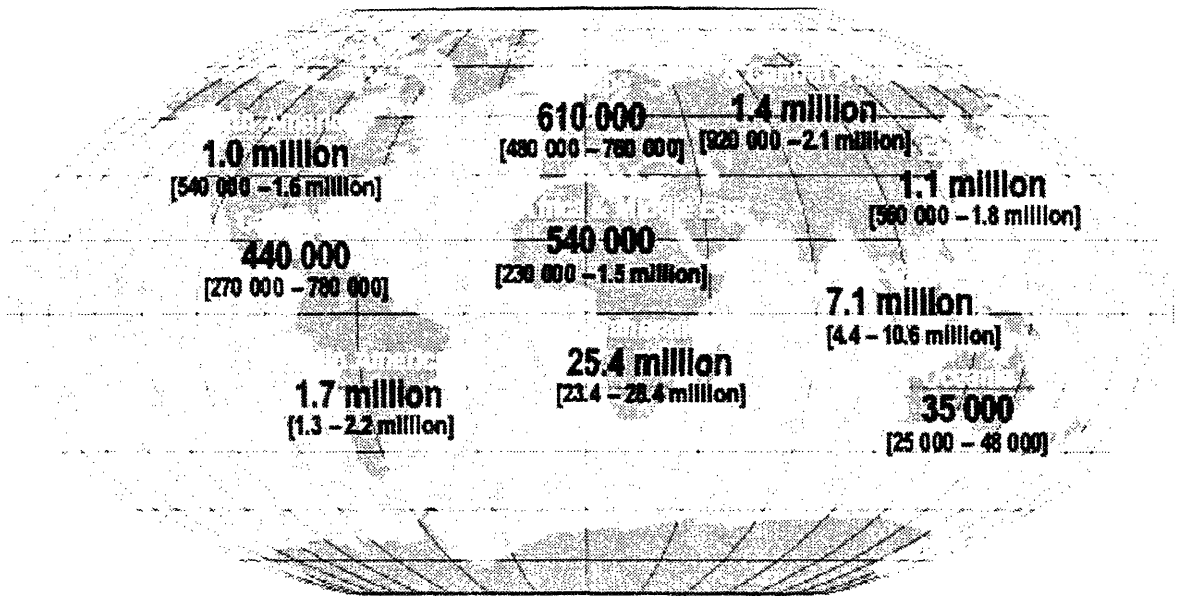
[c] Classes of inhibitors are divided into nucleoside and nucleotide reverse transcriptase inhibitors (NRTI and NtRTI, respectively), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors, and envelope inhibitors.

Figure 1. The global impact of HIV/AIDS.

(A) Adults and children estimated to be living with HIV at the end of 2004. Sub-Saharan Africa is currently at the heart of the pandemic, although Southeast Asia, India, and Latin America have alarming rates of infection. These data are from the Joint United Nations Programme on HIV/AIDS (UNAIDS) / World Health Organization Report on the Global AIDS Epidemic, December 2004 (http://www.unaids.org).

(B) Prevalence of HIV clades worldwide. The three genetic subtypes of HIV are designated M (for major), O (for outlier), and N (for non-M and non-O). The M group includes over 95% of all viral isolates worldwide, and is divided into eight clades (A, B, C, D, F, G, H, and J. Note that Clade E is a circulating recombinant form.) based on the sequence of complete viral genomes. Viruses belonging to the O group have been detected primarily in patients from West Africa, and those in the N group have been discovered only in certain patients from Cameroon. The clade distribution of group M isolates is shown here, with the predominant clade in a geographical region indicated by a large circled letter. The distribution map, based on data collected by UNAIDS, is reproduced with permission from reference (24).

A



Total: 39.4 (35.9 - 44.3) million

B

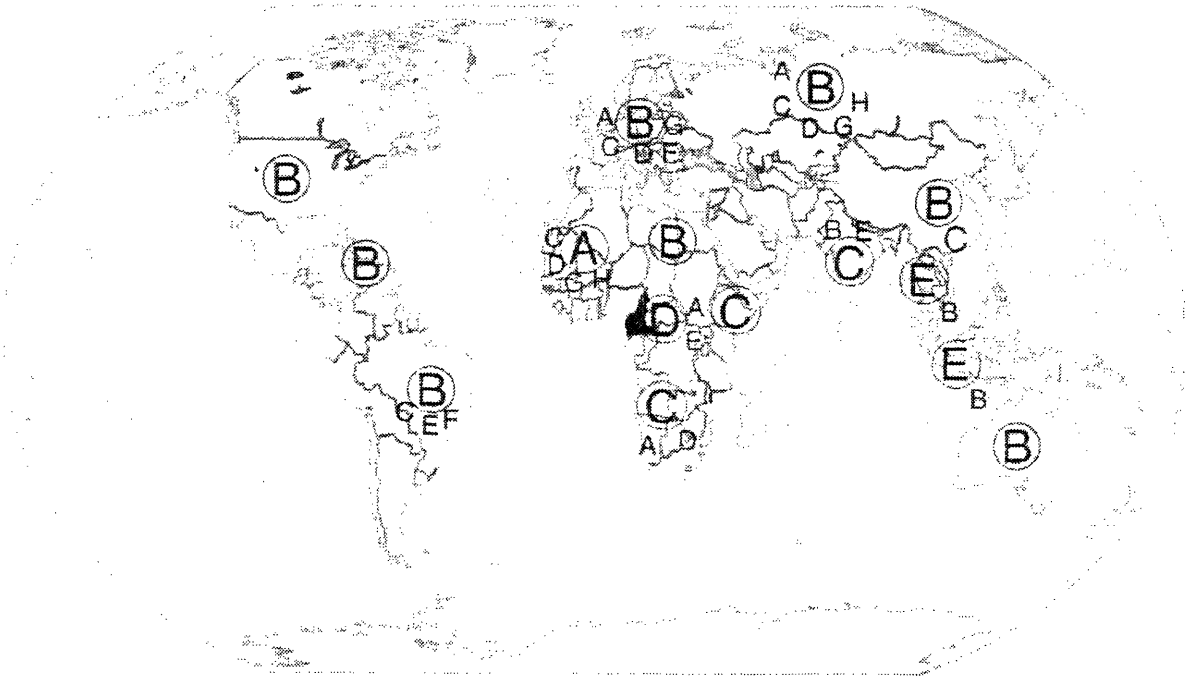


Figure 2. Targets for HIV entry inhibitors.

The four major targets in the Env-mediated HIV fusion pathway and the experimental drugs and monoclonal antibodies (mAb) that block progression through this pathway are depicted. The Env protein is depicted as in Chapter 1, Figure 2. Drugs and antibodies can prevent virus attachment, gp120-binding to receptor or coreceptor, or gp41 conformational transitions. The conformation of gp41 in the prehairpin intermediate (with the N-HR and C-HR regions labeled N and C, respectively) is unknown. The monoclonal antibody 2G12 binds carboyhdrate moieties on gp120 and nonspecifically interferes with viral entry. The gp41-binding mAbs 2F5, 4E10, and Z13 preferentially recognize the native gp41 ectodomain, but do not interfere with receptor attachment. Therefore, they are placed at the 'six-helix bundle transition' step, since they likely inhibit gp41 conformational changes during fusion.

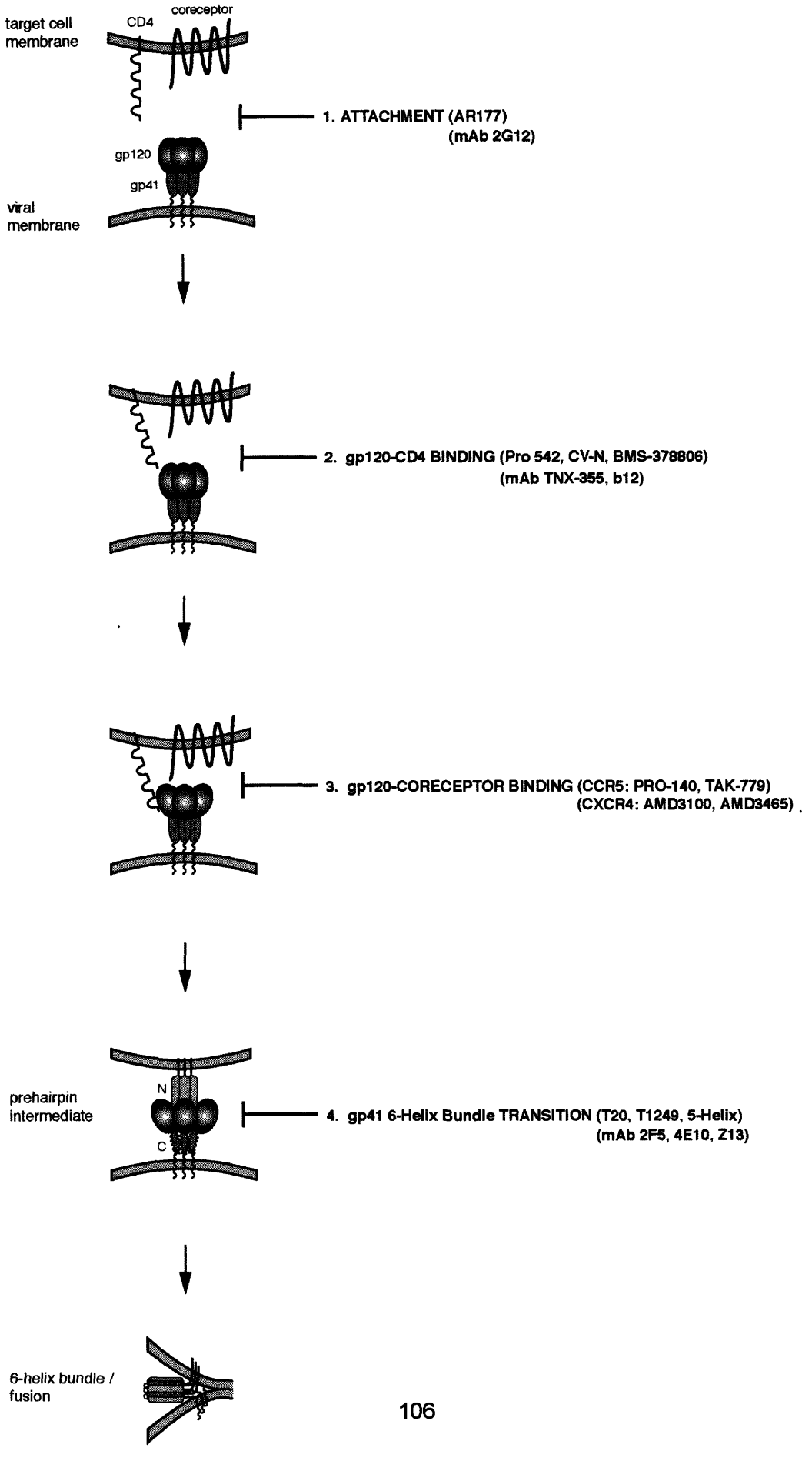
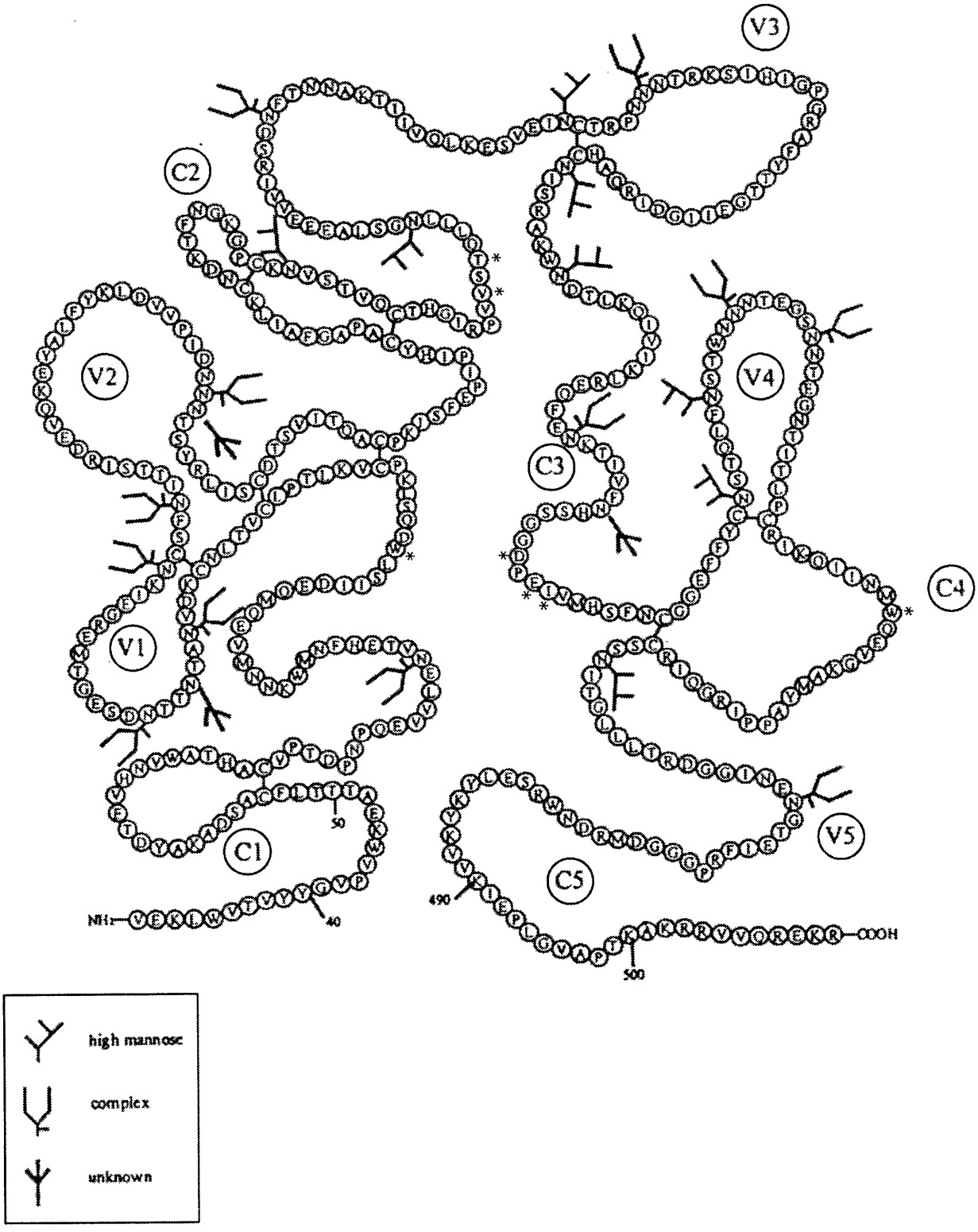

target cell membrane
CD4
coreceptor
1. ATTACHMENT (AR177)
(mAb 2G12)
gp120
gp41
viral membrane
2. gp120-CD4 BINDING (Pro 542, CV-N, BMS-378806)
(mAb TNX-355, b12)
3. gp120-CORECEPTOR BINDING (CCR5: PRO-140, TAK-779)
(CXCR4: AMD3100, AMD3465)
prehairpin intermediate
N
C
4. gp41 6-Helix Bundle TRANSITION (T20, T1249, 5-Helix)
(mAb 2F5, 4E10, Z13)
6-helix bundle / fusion

Figure 3. gp120 subdomains.

A diagram structure of gp120 is shown. The sequence is based on a macrophage-tropic primary isolate, JRFL. Glycosylation sites (high mannose, complex, or unknown) are indicated by branched or U-shaped structures and the key CD4-binding residues are indicated with by a (*). Modified and reprinted with permission from reference (5).


V3
C2
V2
V4
C3
C4
V1
C1
V5
C5
high mannose
complex
unknown

APPENDIX 1

gp41 GLYCOSYLATION

The N-linked glycans on gp41 are located at positions 611, 616, 624, 637, and 674 (numbered according to the HXB2 strain). Of these, the glycosylation pattern at positions 616, 624, and 637 is well conserved among HIV-1, HIV-2, and SIV isolates (2) and all glycosylation sites except the one at position 674 are highly conserved among HIV-1 isolates (4). gp41 is modified with high-mannose carbohydrates (1) (in contrast, about half the oligosaccharides on gp120 are high-mannose oligosaccharides, while the remaining half contains primarily complex, branched structures with a few hybrid oligosaccharides. The type of glycan modification can vary, however, when different cell lines are used for Env biosynthesis (7)). Site-directed mutations that disrupt the canonical glycosylation sequence (Asn-X-Ser/Thr) at all five gp41 sites have been made, and their effects are summarized here.

Johnson and colleagues mutated the four conserved asparagine residues (positions 611, 616, 624, and 637) to glutamine, both singly and as double, triple, or quadruple mutants. Single glutamine substitutions at each of these positions showed only minor differences in expression and processing using 293T cells and no significant effect on viral replication in human and rhesus monkey T cell lines, as well as in rhesus peripheral blood mononuclear cells (3). The quadruple mutant showed increased shedding of gp120 into the culture medium and a defect in viral replication, perhaps due to the altered stability and fold of the native Env glycoprotein. A triple mutant containing glutamine substitutions at positions 611, 616, and 624 exhibited reduced replication in some human lymphocytes (CEMx174 cells) but not in others (C8166-45 cells). Cell-cell fusion activity was assessed for a few combinations of mutants. Env lacking oligosaccharides at positions 616 and 637, or at positions 611, 616, and 624, showed substantial cell-cell fusion activity. Cells lacking all four N-glycans were severely fusion impaired, in agreement with the viral replication studies. However, these studies suggest that removal of the individual oligosaccharides at these four positions does not impair Env expression or gp41 function.

In a similar study, a triple mutant containing glutamine substitutions at positions 616, 624, and 637 exhibited expression levels were similar to wild-type, with precursor cleavage slightly decreased (~90% of wild-type) (2). This triple mutant showed a pattern of immunoreactivity with a panel of gp41 antibodies that was the same as wild-type Env, and was capable of binding to CD4. The cell-cell fusion activity of this mutant was about 30% of wild-type levels however, suggesting that the overall bulk and hydrophilicity of the oligosaccharide moieties may specifically affect gp41 function as fusion progresses.

Another group created mutations both at the Asn and the Ser or Thr of the canonical Asn-X-Ser/Thr glycosylation sequence, to further test how removal of the oligosaccharide (and not the substituted residue itself) affects gp41 function. Mutation of Asn-611 to His or Ser-613 to Ala resulted in normal expression and processing in COS-7 cells, and replication kinetics in SupT cells that were similar to wild-type (4). These results further indicate that glycosylation at position 611 is not essential to Env processing or viral replication. At position 624, Asn was mutated to His, again without effect on processing or growth kinetics. The additional mutation at position 626 was not studied. Substitution of Asn-616 to His showed a slight decrease in growth kinetics, whereas the Ser-618 to Ala mutation caused an increase in gp120 shedding and very little viral replication. Thus, while glycosylation at residue 616 may not be essential for gp41 function, the neighboring residues (Ser-618) may form important contacts with gp120 in the native state. Unlike the later result by Johnson *et al* (in 2001, see above), who mutated Asn-637 to Gln, this 1992 study showed that N637H had delayed viral replication kinetics in SupT host cells. Viral reverse transcriptase activity was not detected in the culture media until day 27 post-infection, versus day 10 for wild-type virus. A similar result was seen with T639V virions. Since reduced viral replication was observed for both the N637H and T639V mutants, the lack of glycosylation may be affecting gp41 function. However, it should be noted that some of the effect of the N637H mutant may be due to the added histadine residue; a more conservative change to glutamine (as in Johnson *et al*) has less potential to disrupt Env structure. Interestingly, mutation of the least well-conserved site, N674H, showed no difference in migration of gp41 by western blot when compared to wild-type (size differences were observed for all other substitutions); therefore, it is likely that this site is not glycosylated in HXB2 Env.

Overall, these studies do not point to a critical role for individual oligosaccharides in gp41 function. In most cases, cell surface expression of cleaved mutant Env is normal (at least by western blot analysis), indicating that the gp41 oligosaccharides must not be essential for stabilizing the native state. Furthermore, removal of individual N-glycans does not alter viral replication kinetics. The effects are made more severe by creating double or triple mutants; however, it is difficult to pinpoint the exact role of N-glycans in gp41 function with such dramatic substitutions. Moreover, many of these studies assess only expression and processing, and viral growth kinetics. The potential role of oligosaccharides in stabilizing gp41 or the six-helix bundle structure has not been investigated until recently.

Since the resolution of gp41 core x-ray structures in the late 1990s, gp41 structure/function studies have been performed with isolated N- and C-peptides. These peptides lack N-linked glycans since their chemical synthesis has been difficult to achieve. Recently, these methods have been improved and have been used to study the effect of glycosylation on the inhibitory activity of a potent C peptide, C34.

C34 contains a single N-linked oligosaccharide, at position 637. Three versions of the C34 glycopeptide were constructed, containing a monosaccharide, a disaccharide, or a native-like high mannose oligosaccharide (6). C34 containing a single glycan moiety (*N*-acetylglucosamine or GlcNAc) at position 637 was synthesized according to standard techniques. To make a native N-linked glycopeptide, a high mannose oligosaccharide was transferred onto GlcNAc-C34 using the *Arthrobacter* endo-β-*N*-acetylglucosaminidase (Endo-A) enzyme and an oligosaccharide donor ($Man_9GlcNAc_2Asn$). Endo-A transfers an intact oligosaccharide moiety (in this case, $Man_9GlcNAc$) to a GlcNAc-containing peptide. The exact linkage between the transferred oligosaccharide and GlcNAc-C34 was confirmed by Endo-F1 treatment and HPLC analysis. Placement of the oligosaccharide at residue 637 was confirmed by treatment with pronase (which hydrolyzes peptide bonds) followed by chromatographic analysis. Using a similar method, construction of a disaccharide-C34 peptide (called LacNAc-C34) was performed using the β-1,4-galactosyltransferase (GalT) enzyme.

The disaccharide and high-mannose versions of C34 were more soluble than the C-peptide alone—presumably due the negative charge of the high-mannose moiety—although they remained unstructured in solution.

The glyco-C34 peptides formed less stable six-helix bundle structures with the N36 peptide when compared to unglycosylated C34. Melting temperatures for the mono-, di-, and high mannose C34/N36 complexes were 54, 56, and 51 °C, compared to 62 °C for the wild-type complex. These six helix structures were also less helical than wild-type (58, 58, and 47%, respectively, compared to 96%). Despite these structural differences however, the modified C peptides still retained potent antiviral activity. In cell-cell fusion assays, mono- and di-saccharide C34 peptides had IC_{50} values comparable to wild-type (~1.3 nM compared to 1.1 nM). The high mannose-C34 had an IC_{50} value of 7.7 nM. These results show that the added oligosaccharide does not substantially affect binding of the C-peptide to the trimeric N core.

These results indicate that glycosylation of the C-terminal region may destabilize the N36/C34 gp41 core more than previously suggested with unglycosylated complexes. However, the glyco-C34 peptides cause a severe decrease in helicity of the bundle, and the authors have not analyzed the oligomeric state or looked at the aggregation of these peptide structures. Furthermore, the *in vivo* data suggests that Env still retains normal function without an oligosaccharide at position 637. Thus, it is possible that glycosylating C34 has a more pronounced effect on the peptide model than it does on native Env. In the context of the N34(L6)C28 peptide model, the stability of the cavity mutants presented in Chapter 3 may be shifted by addition of the oligosaccharide, but there is no evidence that the severity of the cavity mutants with respect to each other (i.e. the Gly < Ala < Phe < WT trend) would be affected.

The effects of other N-glycans on C-peptide or six-helix bundle stability have not been studied. The longest C peptide in the gp41 core that has been obtained by limited proteolysis of the ectodomain and biophysically characterized is C43 (5). This peptide spans HXB2 residues 624 to 666, and contains two N-linked glycosylation sites (624 and 637). Thus, it is likely that the remaining two (or three, if residue 674 is glycosylated) gp41 glycosylation sites fall outside the C-terminal helical region. If the C-helix were plotted on a helical wheel (although it does not conform to a strict heptad repeat motif), Asn-624 would fall at the d position, presumably making close contact with the N-helix. Asn-637 would be placed at the c position, suggesting that this N-linked oligosaccharide would be displayed on the outer surface of the six-helix bundle structure.

The labile, metastable native structure and the low solubility of the Envelope protein make it a difficult protein to study in its entirety. Thus, limited proteolysis experiments have been used to identify and study minimal, stable subdomains. N34/C28 and the slightly longer N36/C34 represent the minimal protease-resistant gp41 fragments. These peptides were derived from proteolysis of longer N51/C43 and N62/C52 fragments (5), which tend to form protein aggregates, especially in isolation. Addition of oligosaccharides to these peptides may destabilize these core peptide structures, and may make them less amenable to biophysical characterization. However, the fact that antibodies directed against C-peptides or the six-helix bundle do not have broadly neutralizing capability may be due to the shielding effect of the oligosaccharides. Thus, glycosylated six-helix bundles, perhaps with longer N-C peptides than those studied here, may serve as better immunogens in vaccine studies.

References

1. **Adams, E. W., D. M. Ratner, H. R. Bokesch, J. B. McMahon, B. R. O'Keefe, and P. H. Seeberger.** 2004. Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions. Chem Biol **11**:875-81.
2. **Fenouillet, E., I. Jones, B. Powell, D. Schmitt, M. P. Kieny, and J. C. Gluckman.** 1993. Functional role of the glycan cluster of the human immunodeficiency virus type 1 transmembrane glycoprotein (gp41) ectodomain. J Virol **67**:150-60.
3. **Johnson, W. E., J. M. Sauvron, and R. C. Desrosiers.** 2001. Conserved, N-linked carbohydrates of human immunodeficiency virus type 1 gp41 are largely dispensable for viral replication. J Virol **75**:11426-36.
4. **Lee, W. R., X. F. Yu, W. J. Syu, M. Essex, and T. H. Lee.** 1992. Mutational analysis of conserved N-linked glycosylation sites of human immunodeficiency virus type 1 gp41. J Virol **66**:1799-803.
5. **Lu, M., S. C. Blacklow, and P. S. Kim.** 1995. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nat Struct Biol **2**:1075-82.
6. **Wang, L. X., H. Song, S. Liu, H. Lu, S. Jiang, J. Ni, and H. Li.** 2005. Chemoenzymatic Synthesis of HIV-1 gp41 Glycopeptides: Effects of Glycosylation on the Anti-HIV Activity and alpha-Helix Bundle-Forming Ability of Peptide C34. Chembiochem.
7. **Zhu, X., C. Borchers, R. J. Bienstock, and K. B. Tomer.** 2000. Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. Biochemistry **39**:11194-204.