

**The Protein Structures Underlying Receptor Binding and Membrane Fusion
of Ecotropic Murine Leukemia Viruses**

by
Deborah Fass

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

DOCTOR OF PHILOSOPHY
in Biology
at the

Massachusetts Institute of Technology
May 1997

© 1997 Massachusetts Institute of Technology
All rights reserved

Signature of Author

Department of Biology
May 19, 1997

Certified by

Peter S. Kim
Professor, Department of Biology
Thesis Supervisor

Accepted by

Richard A. Young
Chairman, Biology Graduate Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

MAY 28 1997

ARCHIVES

LIBRARIES

Dedication

To Mom and Dad,

Thank you for supporting my education and for encouraging
my love of biological sciences

Acknowledgments

I would first like to acknowledge my teachers, beginning with my undergraduate tutor Carolyn Doyle. I thank Carolyn for teaching me to read manuscripts critically and for being an inspiration, both scientifically and personally. I thank Steve Harrison and Tom Ellenberger for introducing me to crystallography, and Steve Gamblin and James Berger for being continuous resources for instruction and advice.

The Kim lab past post-docs Martha Oakley, Pete Petillo, Jamie McKnight, and Steve Blacklow were always generous with practical and emotional help. I also acknowledge the Kim lab past graduate students. Chave Carr inspired my thesis project and taught me a number of the techniques I needed to carry it out. I thank Dan Minor for putting up with me when I was still getting my feet (and his) wet in the lab. I thank Pehr Harbury for showing me that even if only 1% of your ideas work, as long as you have enough ideas, no one notices. Mike Milhollen's bright outlook was contagious; unfortunately his organizational skills were not. Lawren Wu has been my graduate school companion from day one, and I value his generosity, patience, and calm demeanor, as well as his forthrightness and honesty. I thank Brenda Schulman, my dear, dear friend, for the many, many chats about lab, life, and love. Brenda is a model of energy, intelligence, and perseverance. I thank her for her trust and her encouragement.

I thank Rob Davey, Chris Hamson, and Jim Cunningham for a pleasant and rewarding collaboration.

The members of my thesis committee, Carl Pabo and Bob Sauer, have generously taken the time to consider my work and offer direction. I also thank Don Wiley for serving as my outside thesis reader.

The person to whom I owe the deepest gratitude for the completion of this thesis is my adviser, Peter Kim. I thank Peter for providing me the opportunity, or in fact the honor, of working and studying with him. I do not take for granted that, during my four years in the lab, Peter never once asked after the fact, "Why didn't you do it another way?" but instead suggested, "Maybe if you try it this way next time it will work better." I appreciate Peter's foresight and insight, his cool-headedness and unshakable dignity. Among supervisors, Peter is the rarest and most precious of breeds.

Finally, I thank my husband, Russell Rothstein, for not begrudging me the time I spend in the lab, for listening to me talk endlessly about my projects, and for learning the difference between proteins and DNA. Between meeting, courting, marrying, and carrying our child, I am amazed that I was able to get any work done at all. If I was, it was only through his sacrifices and support.

The Protein Structures Underlying Receptor Binding and Membrane Fusion of Ecotropic Murine Leukemia Viruses

by

Deborah Fass

Submitted to the Department of Biology on May 19, 1997
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Biology

Abstract

A combination of protein dissection methods and X-ray crystallographic analyses was used to identify and characterize the fundamental substructures that support cell penetration by mammalian C-type leukemia viruses. A stable, protease-resistant domain was located within the transmembrane (TM) subunit of these retroviruses (Chapter 2), and the high-resolution structure of this domain was determined (Chapter 3). The structure contains a three-stranded coiled coil, stabilized by a hydrophobic cluster at its base, similar to the low pH converted structure of influenza hemagglutinin. The stability of the TM domain in isolation prompted an examination of the relationship of the TM subunit to the receptor-binding (SU) subunit on the surface of intact virions (Chapter 4). The covalent association of SU and TM via a labile disulfide bond was discovered to be significantly less stable than the structure of free TM, consistent with models for retrovirus-mediated membrane fusion that involve loss of the SU subunit. Finally, the high-resolution structure of a minimal receptor-binding domain from an ecotropic murine leukemia virus SU subunit was determined (Chapter 5). The pattern of conserved and variable sequences in the domain reveals the structural principle for generating the range of receptor-specificities found in the mammalian C-type leukemia viruses. The appendix describes a molecular modeling study in which the receptor-binding domain monomer structure is assembled into a putative trimer, as it may exist in the intact glycoprotein complex.

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

Table of Contents

Dedication	2
Acknowledgments	3
Abstract	4
Table of contents	5
Chapter 1	7
Introduction: The Protein Structural Basis of Enveloped Virus Entry	
Chapter 2	26
Chapter 2 has been published as D. Fass and P. S. Kim, "Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin." <i>Current Biology</i> 5, 1377-1383 (1995). © Current Biology Ltd.	
Chapter 3	56
Chapter 3 has been published as D. Fass, S. C. Harrison, and P. S. Kim, "Retrovirus envelope domain at 1.7 Å resolution." <i>Nature Structural Biology</i> 3, 465-469 (1996). © Macmillan Magazines Ltd.	
Chapter 4	79
Characterization of the Labile Covalent Association between the Murine Leukemia Virus Envelope Subunits	

Chapter 5	102
-----------	-----

Chapter 5 has been submitted as D. Fass, R. A. Davey, C. A. Hamson, P. S. Kim, J. M. Cunningham, and J. M. Berger, "Structure of a retrovirus receptor-binding glycoprotein: murine leukemia virus SU domain at 2.0 Å resolution."

Appendix I	125
------------	-----

Model for Trimer Packing of the Receptor-Binding Domain from Ecotropic Murine Leukemia Viruses

Biographical Note	141
-------------------	-----

CHAPTER 1

INTRODUCTION: THE PROTEIN STRUCTURAL BASIS OF ENVELOPED VIRUS ENTRY

Viruses are intracellular parasites that have circumvented their inability to replicate autonomously by developing the capacity to control the complex cellular environment for their own propagation. Throughout history, viruses have been major causes of human disease, from the common cold to diverse and often deadly infections such as smallpox, poliomyelitis, AIDS, and hemorrhagic fever. In this century, however, viruses have also emerged as a powerful tool with which to study fundamental questions in molecular biology. For example, viral infection of bacteria was used to determine that nucleic acids, and not proteins, carry genetic information (Hershey and Chase, 1952). In addition, viruses have served as model systems for uncovering principles of DNA transcriptional regulation (Ptashne, 1987), with direct applications to cancer and developmental biology. More recently, viruses are being considered as potential gene therapy agents, vectors with which to shuttle functional versions of defective genes into patients (Smith, 1995). In fact, an understanding of viral processes forms the axle from which the spokes of modern biology radiate.

The viral infection cycle

Viral spread is a cyclic process in which the genetic material of the virus is shuttled into and out of host cells to be alternately replicated and dispersed (Figure 1). This process can be divided into three main steps: entry of the virus into a host cell, replication and translation of the viral genetic material, and release of new viral particles from the cell.

During viral entry, the virus must first bind to an appropriate host cell. Binding occurs between viral surface proteins and cellular transmembrane proteins or other cell-surface components. These cellular binding partners have diverse primary functions, but function secondarily as viral receptors. Subsequent to binding, the virus penetrates the host cell membrane and sheds the protective protein layers surrounding its genome.

After the viral genome is deposited in the host cell cytosol, it is then both copied to make more genomes and used to synthesize structural and regulatory proteins that will make up the new virus particles. The precise manner and time frame in which these steps occur depend on the specific virus and the

nature of the viral genetic material. Some viral genomes integrate into the host cell genomes and may delay expression of new viral components; genomes of other viruses are utilized directly.

To complete the infection process, newly-synthesized viral proteins assemble around a copy of the genome, and the fresh virion is released from the cell in one of two distinct manners. Release from the cell for some viruses occurs when the cell bursts and frees its contents. Other viruses, called "enveloped" viruses, exit the cell by the more subtle process of "budding." During budding, the virion core in the cell cytoplasm docks against the lipid bilayer membrane of the cell, the bilayer folds around the viral protein assembly, and the membrane pinches off to form a viral membrane distinct from the cell membrane. The virus is thereby liberated from the cell. Unlike the first method of release, bursting of the cell, budding does not require that the cell be destroyed.

During the budding process, the virion acquires not only a lipid bilayer membrane, but also the virally-encoded proteins, often called "envelope" glycoproteins, that are anchored in the membrane and form the outer layer of the complete virus particle. These viral envelope glycoproteins, being the only exposed protein component of the virion, are major targets for the anti-viral immune response in infected hosts. In addition, these surface glycoproteins are the viruses' tickets to entry into the next cycle of viral propagation.

Viral envelope glycoproteins

Viral envelope glycoproteins mediate the two key steps of viral entry into host cells: binding of the virus to the target cell, and fusion of the viral and cellular membranes to reintroduce the viral core into the cell cytoplasm. These two events are dramatically different in nature. The first is somewhat familiar; the receptor-binding event is likely to involve principles common to many protein-protein interactions, such as complementary Van der Waals surfaces, a lack of charge repulsion, and the liberation of water at the intermolecular interface. The second function, membrane fusion, is relatively uncharted. This event involves an interplay of protein components, lipids, and counterions in a dynamic process that has yet to be detailed.

Two goals guide an investigation, from a structural biological perspective, into the functions of viral envelope proteins in the process of entry into cells. The first goal is to determine the protein structures that carry out the distinct events of receptor binding and membrane fusion. The second is to determine how the viral glycoprotein complex structurally converts between these two dramatically different functions. This second goal is multi-faceted and can be answered in any of a variety of terms: a dynamic view of the complex as it sheds one function and assumes the next, a thermodynamic study of the various functional states and their inter-relationship, or identification of the agent or external condition that initiates conversion from one functional state to the next. A variety of approaches, on a number of viral systems, are currently being taken to develop an understanding of the protein structural basis of viral entry.

The prototype: influenza hemagglutinin

The most thoroughly studied surface protein of an enveloped virus is the hemagglutinin (HA) from the orthomyxovirus influenza (for reviews see Wiley and Skehel, 1987; Stegmann and Helenius, 1993). Several features of influenza make its HA protein ideal for structural and mechanistic studies. First, virus quantities large enough for biochemical and biophysical analyses can be produced in the laboratory from chicken embryos. Second, conversion of HA from a receptor-binding state to a membrane-fusion competent conformation can be readily controlled. Penetration of the cell membrane can be artificially induced in the laboratory simply by lowering the pH of the ambient solution. This procedure mimics the natural infection process in which influenza virus binds sialic acid on the surface of cells, is internalized into vesicles, and enters the cell cytoplasm when the pH of these vesicles drops. Third, influenza HA is very stable when maintained at neutral pH, even when cleaved from the viral membrane. Finally, sequence data from natural isolates, gathered from repeated outbreaks caused by antigenic variation, provides a valuable database for identifying regions of HA that are structurally and functionally essential, as opposed to those regions that can vary to allow the virus to escape the host immune response.

A large body of work has shed light on the structures and functions of HA. The HA protein complex is synthesized as a precursor, HA0, which is then

proteolytically cleaved into two fragments, HA1 and HA2 (Lazarowitz et al., 1971; Skehel and Waterfield, 1975). HA2 is membrane-anchored, the receptor-binding HA1 is attached to HA2 via a disulfide bond, and the HA1/HA2 covalent complex associates non-covalently into trimers (Wiley et al., 1977). The ectodomain of the influenza HA glycoprotein “spike” in its native conformation was the first enveloped viral glycoprotein structure to be determined crystallographically (Figure 2) (Wilson et al., 1981). The top of the spike is formed by a globular β -strand rich domain of HA1, while HA2 constitutes much of the stalk of the spike and forms a three-stranded coiled coil at the trimer core.

The native HA structure served as a backdrop for studies on the apparent large-scale and irreversible conformational change that occurs upon exposure of HA to low pH. For example, electron microscopy studies revealed that HA appears to lengthen and dissociate when treated with low pH (Doms et al., 1985; Ruigrok et al., 1986). Furthermore, exposure of hydrophobic regions of HA upon low pH treatment hinted at a mechanism for the activation of membrane fusion (Skehel et al., 1982). Circular dichroism and fluorescence studies, however, suggested that relatively small changes in total secondary structure content and core packing occur (Skehel et al., 1982; Sato et al., 1983).

Recently, two studies described the nature of the end-state of the large-scale conformational change of influenza HA. The first study involved HA2 peptides and demonstrated that a region existing as a loop in the native HA structure forms a trimeric helical assembly in solution (Carr and Kim, 1993). The second study supported and extended the peptide studies, presenting a high-resolution view of the core of low-pH converted HA and confirming that the native-state loop is in fact recruited in the low-pH converted structure to extend the trimeric coiled coil at the HA2 core (Figure 3) (Bullough, et al. 1994).

Although these detailed snapshots of influenza HA before and after exposure to low pH are available, many questions remain regarding the role of HA in the mechanism of membrane fusion. For example, in the low-pH converted structure, the “fusion peptide” region of HA2, which inserts into the target cell membrane (Stegmann et al., 1991; Tsurudome et al., 1992), is at least 100 Å away from the transmembrane region. How the viral and cell membrane are brought together for fusion has yet to be revealed. In addition, the end-state

of fusion must involve only a single membrane, but the orientation of the low-pH converted HA trimer with respect to this membrane is unknown.

An alternative: tick-borne encephalitis virus E protein

The second enveloped virus to have a surface protein revealed crystallographically is the flavivirus tick-borne encephalitis virus (TBE) (Rey et al., 1995). The structure of a soluble fragment of the TBE E protein reveals a dramatically different assembly from that of influenza HA (Figure 4). First, TBE E is dimeric, not trimeric. In addition, E protein is likely to lie flat against the viral membrane, perhaps forming a lattice over the surface of the TBE virion, rather than extending from the viral surface as a knob or spike. In support of this proposed orientation, the elongated E protein dimer forms an arc with a radius comparable to that of the viral particle, and carbohydrate moieties are found on the convex surface. The domain structure of TBE E protein is also very different from that of influenza HA. TBE E protein is composed of three domains: a β -barrel, an immunoglobulin-like domain, and an elongated dimerization domain. The complex has very little helix content, in contrast to influenza HA which is maintained as a trimer by contacts within its largely helical HA2 subunit.

Flavivirus E proteins, like influenza HA, undergo a low-pH induced conformational change to initiate membrane fusion (Kimura and Ohyama, 1988). This conformational change is accompanied by an oligomerization switch, and the TBE E protein becomes trimeric after exposure to low pH (Allison et al., 1995). It has been suggested that this oligomerization switch can theoretically be accomplished with only minor rearrangements to the E protein lattice (Allison et al., 1995). Disulfide bonds, particularly in the dimerization domain, may restrict intradomain conformational changes. Nevertheless, significant changes in monoclonal antibody binding are seen upon exposure to low pH (Roehrig et al., 1990; Heinz et al., 1994), suggesting that the relative orientation of the three domains may change, with regions that were previously buried being exposed. The molecular details of this trimeric fusogenic conformation have yet to be described.

The emerging retrovirus envelope

The subunit organization of the retrovirus envelope is similar to that of influenza HA (for review see Hunter and Swanstrom, 1990). The surface glycoproteins of retroviruses are also synthesized as single-chain precursors, which are proteolytically cleaved into two fragments, called generically for retroviruses the surface (SU) and transmembrane (TM) subunits. SU is the functional analog of influenza HA1, as it contains the receptor-binding regions. TM, like influenza HA2, contains the hydrophobic segments that are likely to anchor the protein bridge between the viral and cellular membranes during membrane fusion. Some studies report trimers as the preferred oligomerization state for retroviral envelope protein complexes (Einfeld and Hunter, 1988; Weiss et al., 1990; Kamps et al., 1991), but other states have also been proposed for lentiviruses (Rey et al., 1990; Doms et al., 1991).

There is preliminary evidence that retroviral envelope proteins can undergo a conformational change during viral entry, although this reorganization appears to be provoked by receptor-binding rather than by low pH as in influenza or TBE. Evidence for a conformational rearrangement comes from changes in reactivity to monoclonal antibodies (Sattentau et al., 1993) or in protease sensitivity (Gilbert et al., 1995) after exposure to soluble versions of the viral receptors. In addition, mutations have been made in retroviral envelope glycoproteins that appear to have no effect on events required for receptor binding, such as glycoprotein cell-surface expression and stability of the SU/TM interaction, and yet have dramatic effects on the membrane fusion function (Cao et al., 1993).

Further examination of the structure and structural changes of the retroviral envelope are complicated by a number of features. First, intact retroviral glycoprotein complexes are difficult to produce in yields high enough for many biophysical studies. Second, retroviral envelope protein complexes are unstable and can dissociate during purification, or even during expression in cells (Saha et al., 1994; Earl et al., 1991). Finally, the use of purified, concentrated receptor to induce conformational changes in the retroviral envelope is more difficult than altering the pH of solution to study influenza HA. Since many retroviral receptors are multipass transmembrane proteins, the technical

difficulties become prohibitive. Moreover, while exposure to low pH results in quantitative conversion of all HA molecules on the influenza viral surface, it is likely that only a small fraction of retroviral envelope proteins actually contact receptor during viral entry, and perhaps not simultaneously. Taken together, the limits on material, the instability of the SU/TM complex, and the inability to achieve quantitative conversion of the retroviral envelope, mean that heterogeneous populations or events are often studied. Therefore, firm structural and mechanistic insights regarding the retroviral entry process have lagged behind progress in other viral systems.

Purpose

In spite of the medical and therapeutic significance of retroviruses, structural biological conclusions have been limited for retroviral envelope proteins, and no high-resolution structural frameworks have been available in which to ground mutagenesis and other data. This thesis describes a protein dissection approach (Chapter 2) toward generating detailed structural models for TM (Chapter 3) and SU (Chapter 5) components of the retroviral surface, using as an experimental system the ecotropic murine leukemia viruses. The dissection approach is pursued together with a broader holistic approach, and the relationship between the protein components of the retrovirus envelope is demonstrated in context of the intact virus (Chapter 4).

References

- Allison, S.L., Schalich, J., Stiasny, K., Mandl, C.W., Kunz, C., and Heinz, F.X. Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J. Virol.* **69**, 695-700 (1995).
- Bullough, P.A., Hughson, F.M., Skehel, J.J., and Wiley, D.C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43 (1994).
- Cao, J., Bergeron, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 envelope glycoprotein. *J. Virol.* **67**, 2747-2755 (1993).
- Carr, C.M., and Kim, P.S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**, 823-832 (1993).
- Doms, R.W., Earl, P.L., and Moss, B. The assembly of the HIV-1 *env* glycoprotein into dimers and tetramers. *Adv. Exp. Med. Biol.* **300**, 203-219 (1991).
- Doms, R.W., Helenius, A., White, J. Membrane fusion activity of the influenza virus hemagglutinin. The low pH-induced conformational change. *J. Biol. Chem.* **260**, 2973-2981 (1985).
- Earl, P.L., Moss, B. and Doms, R.W. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *J. Virol.* **65**, 2047-2055 (1991).
- Einfeld, D., and Hunter, E. Oligomeric structure of a prototype retrovirus glycoprotein. *Proc. Natl. Acad. Sci. USA* **85**, 8688-8692 (1988).
- Gilbert, J.M., Hernandez, L.D., Balliet, J.W., Bates, P., and White, J.M. Receptor-induced conformational changes in the subgroup A avian leukosis and sarcoma virus envelope glycoprotein. *J. Virol.* **69**, 7410-7415 (1995).
- Heinz, F.X., Stiasny, K., Puschner-Auer, G., Holzmann, H., Allison, S.L., Mandl, C.W., and Kunz, C. Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. *Virology* **198**, 109-117 (1994).
- Hershey, A.D., and Chase, M. Independent Function of Viral Protein and Nucleic Acid on Growth in Bacteriophage. *J. Gen. Physiol.* **36**, 39-56 (1952).

- Hunter E., and Swanstrom. R. Retrovirus envelope glycoproteins. *Cur. Top. in Microbiol. and Immun.* **157**, 187-253 (1990).
- Kamps, C.A., Lin, Y.-C., and Wong, P.K.Y. Oligomerization and transport of the envelope protein of Moloney murine leukemia virus-TB and of ts1, a neurovirulent temperature-sensitive mutant of MoMuLV-TB. *Virology* **184**, 687-694 (1991).
- Kimura, T., and Ohyama, A. Association between the pH-dependent conformational change of West Nile flavivirus E protein and virus-mediated membrane fusion. *J. Gen. Virol.* **69**, 1247-1254 (1988).
- Lazarowitz, S.G., Compans, R.W., Choppin, P.W. Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. *Virology* **46**, 830-843 (1971).
- McClure, M.O., Sommerfelt, M.A., Marsh, M., and Weiss, R.A. The pH independence of mammalian retrovirus infection. *J. Gen. Virol.* **71**, 767-773 (1990).
- Ptashne, M. A Genetic Switch: *Gene Control and Phage λ* , Cell Press and Blackwell Scientific Publications, Cambridge, Massachusetts, 1987.
- Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., and Harrison, S.C. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **375**, 291-298 (1995).
- Rey, M.A., Laurent, A.G., McClure, J., Krust, B., Montagnier, L., and Hovanessian, A.G. Transmembrane envelope glycoproteins of human immunodeficiency virus type 2 and simian immunodeficiency virus SIV-mac exist as homodimers. *J. Virol.* **64**, 922-926 (1990).
- Roehrig, J.T., Johnson, A.J., Hunt, A.R., Bolin, R.A., and Chu, M.C. Antibodies to dengue 2 virus E-glycoprotein synthetic peptides identify antigenic conformation. *Virology* **177**, 668-675 (1990).
- Ruigrok, R.W.H., Wrigley, N.G., Calder, L.J., Cusack, S., Wharton, S.A., Brown, E.B., and Skehel, J.J. Electron microscopy of the low pH structure of influenza virus hemagglutinin. *EMBO J.* **5**, 41-49 (1986).
- Saha, K., Lin, Y.-C., Wong, P.K.Y. A simple method for obtaining highly viable virus from culture supernatant. *J. Virol. Meth.* **46**, 349-352 (1994).
- Sato, S.B., Kawasaki, K., Ohnishi, S.I. Hemolytic activity of influenza virus hemagglutinin glycoproteins activated by mildly acidic environments. *Proc. Natl. Acad. Sci. USA* **80**, 3153-3157 (1983).

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

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

Skehel, J.J., and Waterfield, M.D. Studies on the primary structure of the influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. USA* **72**, 93-97 (1975).

Smith, A.E. Viral vectors in gene therapy. *Ann. Rev. Microbiol.* **49**, 807-838 (1995).

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

Tsurudome, M., Glück, R., Graf, R., Falchetto, R., Schaller, U., and Brunner, J. Lipid interactions of the hemagglutinin HA2 NH₂-terminal segment during influenza virus-induced membrane fusion. *J. Biol. Chem.* **267**, 20225-20232 (1992).

Weiss, C.D., Levy, J.A., and White, J.M. Oligomeric organization of gp120 on infectious human immunodeficiency virus type I particles. *J. Virol.* **64**, 5674-5677 (1990).

Wiley, D.C., and Skehel, J.J. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Ann. Rev. Biochem.* **56**, 365-394 (1987).

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

Figure 1 Schematic diagram of the infection cycle of an enveloped virus

Using its envelope glycoproteins, shown as dark gray triangles, the virus binds to a receptor on the target cell (step 1). The subsequent fusion of the viral and cellular membranes, which are shown as thin black lines, results in deposition of the viral capsid into the cell cytoplasm (step 2), represented by the thick light gray circle. The proteins that surround the viral genome then dissociate (step 3), freeing the genome to be used as a template for synthesizing viral proteins (step 4a), and for generating additional copies of the genome itself (step 4b). The viral proteins assemble around these genomes to form new viral core particles (5). A piece of the cell membrane, containing the viral envelope proteins, folds around the viral cores, and eventually separates from the rest of the cell membrane to complete the new viral particle (step 6).

Figure 1

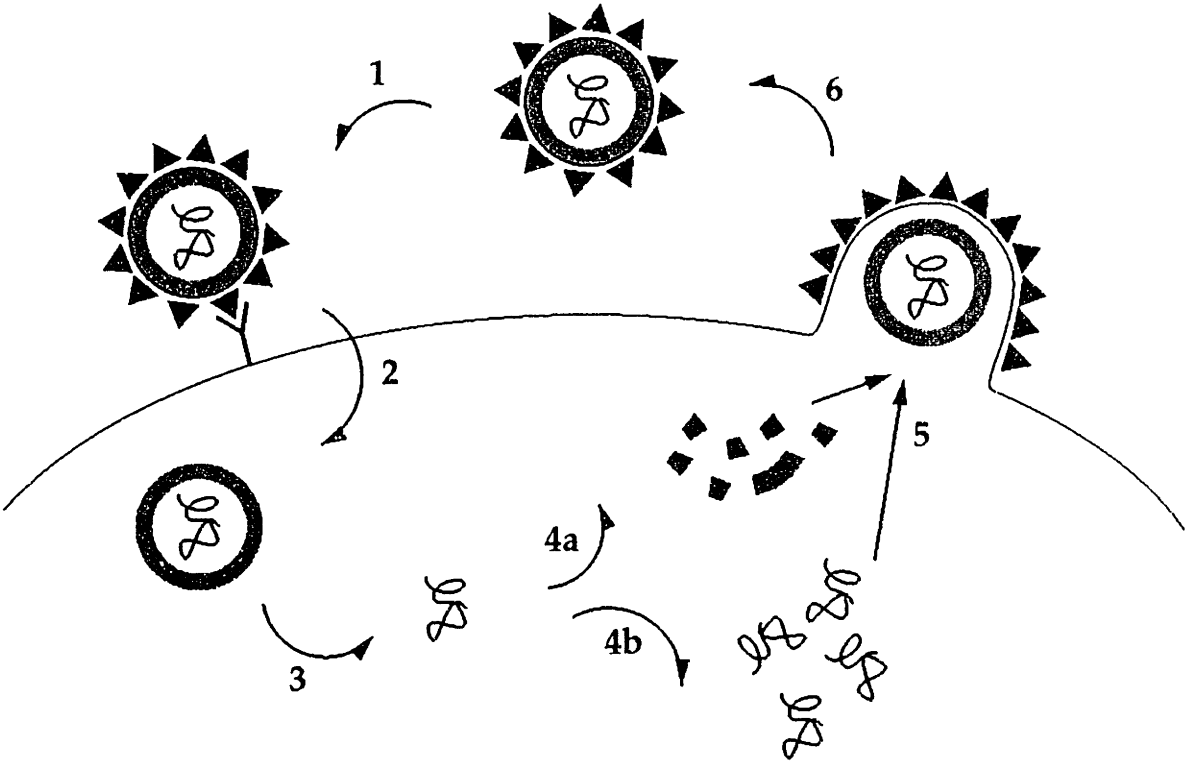


Figure 2 X-ray crystal structure of the native influenza HA trimer

The three components of the native influenza HA homotrimer (Wilson et al., 1981) are shown in shades of purple. Residues (38-53) and (54-77) in one HA2 subunit are highlighted in red and yellow, respectively. These regions form a three-stranded coiled coil in isolated HA2 peptides.

Figure 2

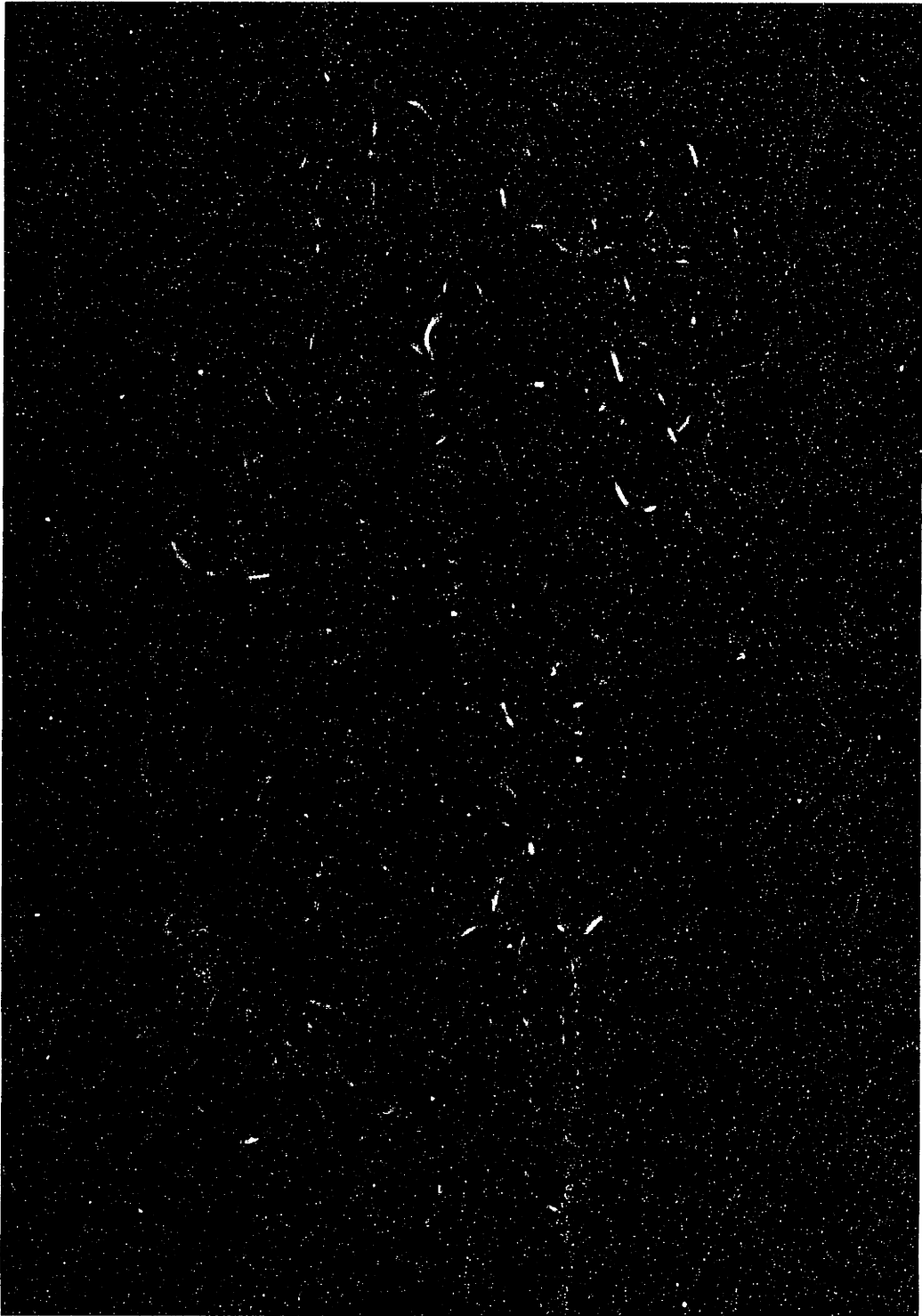


Figure 3 X-ray crystal structure of the low-pH converted HA trimer

The three components of the low-pH converted HA homotrimer (Bullough et al., 1994) are shown in shades of purple. Residues (38-53) and (54-77) in one HA2 subunit are shown in red and yellow, respectively, as in Figure 2. These regions illustrate the essence of the HA conformational change that occurs upon exposure to low pH, although other changes are evident in the structure. For example, the long coiled coil folds back at the bottom, so that a short helix packs against the new base of the coiled coil.

Figure 3

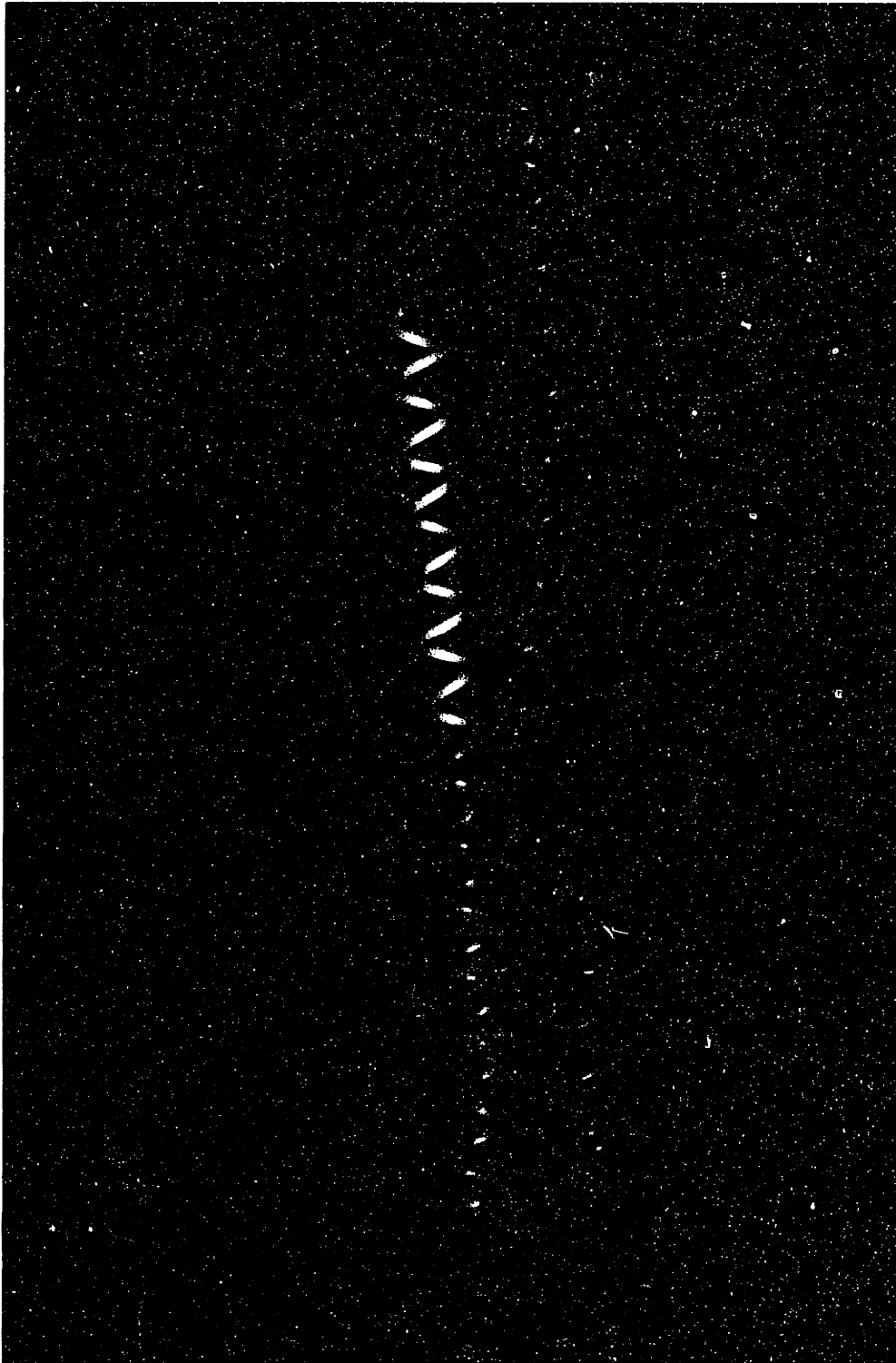
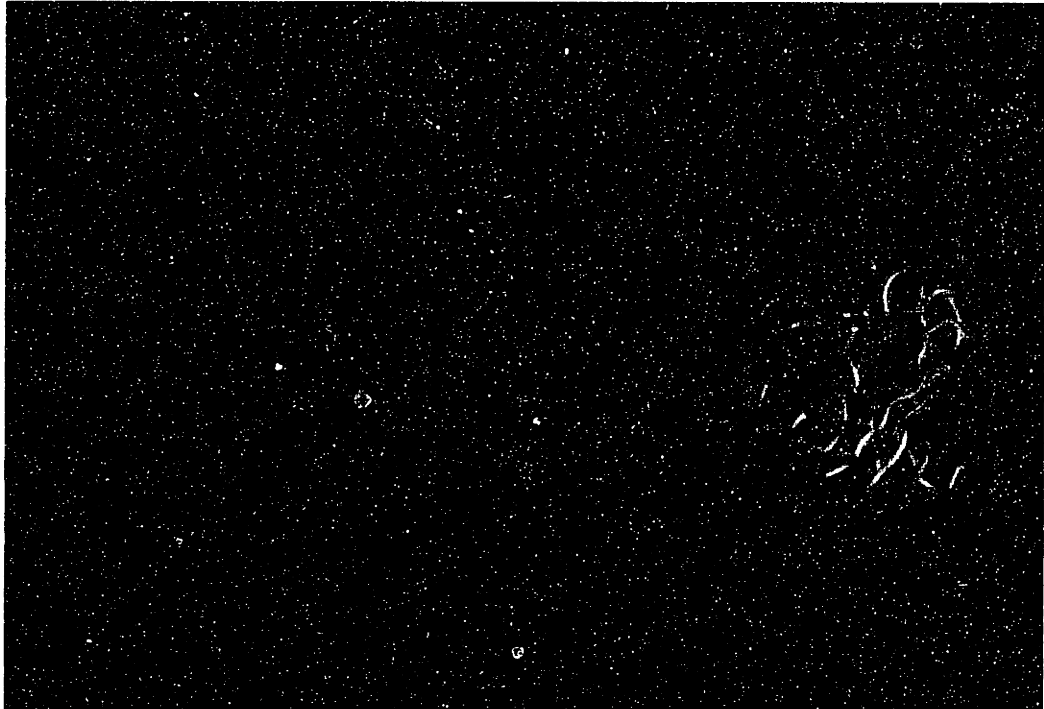


Figure 4 X-ray crystal structure of the tick-borne encephalitis virus E protein

The two subunits of the TBE E protein homodimer (Rey et al., 1995) are shown in red and yellow. The view is from the side, as if the viral membrane were toward the bottom of the page. This view emphasizes the convex (upper) and concave (lower) surfaces of the dimer, with the latter surface mirroring the curvature of the viral membrane.

Figure 4



CHAPTER 2

DISSECTION OF A RETROVIRUS ENVELOPE PROTEIN REVEALS
STRUCTURAL SIMILARITY TO INFLUENZA HEMAGGLUTININ

Background: Retrovirus envelope proteins contain a 4-3 repeat of hydrophobic residues, characteristic of coiled coils. This repeat is located in the transmembrane (TM) subunit adjacent to the fusion peptide, a region that inserts into the host bilayer during the membrane-fusion process. A 4-3 hydrophobic repeat region in an analogous position of the influenza hemagglutinin protein is recruited to extend a three-stranded coiled coil during the conformational change to the fusion-competent state. To determine the conformation of the retrovirus TM subunit and the role of the 4-3 hydrophobic repeat, we constructed soluble peptide models of the envelope protein from Moloney murine leukemia virus (MMLV).

Results: The region of the MMLV TM protein external to the lipid envelope (the ectodomain) contains a stably folded, trimeric, protease-resistant core. As predicted, an α -helical segment spans the 4-3 repeat. A cysteine-rich region carboxy-terminal to the 4-3 repeat confers a dramatic increase in stability and displays a unique disulfide bonding pattern.

Conclusions: Our results demonstrate that MMLV TM can fold into a stable and distinct species in the absence of the receptor-binding "surface" co-subunit (SU) of the envelope complex. As the SU subunit is readily shed from the surface of the virus, we conclude that the TM structure forms the core of the MMLV membrane fusion machinery, and that this structure, like the fusion-active conformation of influenza, contains a three-stranded coiled coil adjacent to the fusion peptide.

Background

The envelope (Env) proteins of retroviruses bind to cell surface receptors and promote fusion with the host membrane [1]. The mechanism of retrovirus-mediated membrane fusion has remained elusive due to the lack of detailed structural information about Env proteins. Structural studies have been hindered by the low virus titer in the natural systems studied and by the large size of the envelope complexes. Furthermore, the association between the Env subunits is labile, and the protein complexes are membrane-bound.

By contrast, the hemagglutinin (HA) protein of the orthomyxovirus influenza has been characterized by X-ray crystallography in both the native [2] and the low-pH activated forms [3]. These high-resolution HA structures, together with other studies, have led to a model for HA-mediated membrane fusion in which dissociation of the receptor-binding domains from one another [4,5] accompanies extension of a coiled coil to bring the fusion peptides into position to interact with the target cell membrane [6]. The fusion-active subunit would then bridge the viral and cellular membranes to form a critical intermediate in the membrane fusion process.

Although the orthomyxovirus and retrovirus families have no obvious evolutionary relationship, features of influenza HA are similar to those of retroviral Env proteins. First, both HA and Env are composed of two fragments cleaved from a common precursor. Like the HA1 domain of influenza, the retrovirus surface glycoprotein (SU) binds the host-cell receptor [1]. The retroviral transmembrane subunit (TM) contains a hydrophobic sequence at its amino terminus analogous to the fusion peptide of the influenza HA2 subunit [7], which inserts into the target membrane [8,9]. Retrovirus TM, like HA2, contains a 4-3 hydrophobic repeat, predicted to form a coiled coil [10,11]. Finally, shedding of the SU subunit of retroviruses such as HIV [12] and murine leukemia virus [13] might parallel the displacement of influenza HA1 during the conformational change to the fusogenic state.

The central features of the influenza fusogenic state were preserved in the crystallized fragment of low-pH converted HA, despite the absence of the fusion peptide, the transmembrane helix, and most of the receptor-binding domain [3].

Furthermore, the “spring” to the fusion-competent conformation was modeled using peptides corresponding to only the coiled-coil regions of HA [6]. We consequently undertook a protein dissection study of a retrovirus Env protein in order to characterize the structure of the retrovirus membrane-fusion apparatus.

Moloney murine leukemia virus (MMLV) was chosen for study because the TM subunit is small (15 kD) and lacks glycosylation sites. Moreover, the intersubunit disulfide bond between MMLV SU and TM can be eliminated by an apparent disulfide rearrangement [13]. The requirement for disruption of a covalent bond during shedding of SU implies a specific mechanism to dispose of SU. Therefore, the structure of isolated TM from MMLV may be important in the membrane-fusion mechanism of this retrovirus.

Peptides were designed to determine if the MMLV TM protein can fold independently. To avoid the difficulty of working with insoluble peptides, we studied the hydrophilic segment of TM by removing the fusion peptide and the transmembrane region from our model constructs.

Results

The TM Model Folds into a Stable Structure

The primary structure of the MMLV TM protein is shown in Figure 1. We produced a peptide, Mo-92, that extends from Asp 45 in the TM sequence to the beginning of the transmembrane region. Asp 45 was chosen because it is at an f position near the beginning of the 4-3 hydrophobic repeat region, placing the positively charged amino terminus away from the hydrophobic core (Fig. 2), and was predicted by the Paircoil algorithm [14] to be near the start of the coiled coil. To verify this choice, a peptide was constructed starting at Met 31 and including the following additional sequence from TM: MATQQFQQLQAAVQ. When this peptide was subjected to proteolysis with thermolysin, the amino terminus was trimmed to Val 43. Although thermolysin readily cleaves before leucine residues, prolonged incubation did not result in cleavage before Leu 47. Therefore, the amino terminus of Mo-92 is within 2 residues of the start of the protease-resistant, and presumably well-folded, region.

Circular dichroism (CD) spectroscopy (Fig. 3) indicates that our peptide model for TM has a helix content of approximately 59% (Table 1). The Mo-92 structure is remarkably stable, with a temperature midpoint for thermal unfolding (T_m) of 86°C under physiological conditions. Carboxy-terminal truncations, denoted Mo-55 and Mo-36 (Fig. 1), were constructed to locate the secondary structure within Mo-92. Mo-36 spans the 4-3 hydrophobic repeat region, and the CD spectrum indicates that approximately 28 residues of this peptide are in a helical conformation. We conclude that the 4-3 repeat region is helical. Since the Mo-55 peptide contains no additional helical residues as judged by CD, but is substantially more stable than Mo-36 (T_m values of 80°C and 30°C respectively), we conclude that the immunosuppressive and cysteine-rich regions, while folded, are in a non-helical conformation.

Mo-92 has a higher helix content than Mo-55, indicating that helical segment(s) are present carboxy-terminal to the cysteine-rich region. However, these segments do not confer a substantial increase in stability. A thermal melt of Mo-92 monitored by CD signal at 222 nm reveals a folded baseline with a slope greater than that of Mo-55. When the estimated number of helical residues in each peptide is plotted vs. temperature (Fig. 4), it appears that additional helical region(s) within Mo-92 melt non-cooperatively and independently, prior to the transition for the amino-terminal helical region.

TM is Trimeric

The oligomerization state of the TM peptide Mo-92 was determined by analytical ultracentrifugation. The data were consistent with a trimer model over a 9-fold concentration range (Fig. 5A). However, a systematic trend is observed in the residuals between the data and the linear fit, becoming more severe with increasing concentration. This behavior is diagnostic of a tendency toward aggregation.

To pare away regions contributing to non-specific aggregation, Mo-92 was subjected to proteolysis, and the products were analyzed by mass spectrometry. Digestion with trypsin (Fig. 1, pink arrow) or chymotrypsin (Fig. 1, blue arrows) removed carboxy-terminal sequences, but left the amino terminus intact, yielding fragments similar to Mo-55. Mo-55 thus corresponds to a protease-

resistant subdomain within the TM subunit. When Mo-55 was subjected to analytical ultracentrifugation, data fit closely to a trimer model, and no systematic residuals were observed (Fig. 5B).

TM Contains a Specific Disulfide Bond and a Free Thiol

The protease-resistant subdomain of MMLV TM contains three cysteines, the presence and spacing of which are preserved in C- and D-type retroviruses [15,16]. Disulfide bond formation between these cysteines is therefore likely to be structurally and/or functionally significant. Mapping of disulfide bonds in Mo-92 was challenging because there are adjacent cysteines that cannot be separated by proteolysis. Instead, oxidized Mo-92 was incubated with 2-nitro-5-thiocyanobenzoic acid (NTCB), which cleaves proteins at free cysteine residues [17].

NTCB cleavage of Mo-92 yielded fragments with masses of 5597 and 5081, corresponding to the amino- and carboxy-terminal products, respectively, of cleavage at the third cysteine (C3) in the TM sequence (Fig. 1). Control experiments with reduced cysteine-to-alanine mutants (used in the disulfide permutation experiments described below) demonstrated that fragments from cleavage at either of the adjacent cysteines could be obtained and readily distinguished from one another by mass spectrometry. Thus, the NTCB cleavage results indicate that only C3 in oxidized Mo-92 is reactive with NTCB, and that C1 and C2 are likely to be protected in an intramolecular disulfide bond. Cleavage of oxidized Mo-55 produced a 5600 dalton amino-terminal fragment, demonstrating that the protease-resistant core is sufficient to specify the unique disulfide bonding pattern that leaves only C3 reactive with NTCB.

To verify that C3 remains free due to the preferential formation of the C1-C2 disulfide bond, we performed a disulfide permutation experiment. Three mutants of Mo-92 were constructed by replacing each cysteine individually with alanine. The peptides were oxidized in the presence of 6 M GuHCl to remove structural barriers to formation of non-native disulfide bonds. After refolding by removal of GuHCl, the stability of each oxidized peptide was assayed by thermal denaturation (Table 2). The mutant with a disulfide bond between C1 and C2, denoted Mo-92.3, is most stable, with a T_m of 86°C. In contrast, the other two

mutants (Mo-92.1 and Mo-92.2) are substantially destabilized, with T_m 's of 76°C and 70°C. The permutation repeated in the background of Mo-55 gave qualitatively similar results (Table 2). Interestingly, for both Mo-92 and Mo-55 backgrounds, the folded baselines of the three disulfide variants are identical, indicating that alternate disulfide pairings accommodate the helical region completely, albeit with impaired stability. These disulfide permutation experiments are consistent with the NTCB cleavage data and confirm that the most stable form of isolated TM contains a disulfide bond between C1 and C2.

Discussion

The structure of the retrovirus TM protein is central to viral entry. First, TM contains the fusion peptide region. In addition, the most highly conserved regions of retrovirus Env lie within TM. Carboxy-terminal to the 4-3 hydrophobic repeat in C- and D-type retroviruses, TM invariably contains an "immunosuppressive sequence" [15]. Peptides with the consensus sequence for this region have been shown to inhibit lymphocyte proliferation [18] and INF- γ production [19]. Adjacent to the immunosuppressive region are the conserved cysteine residues, at least one of which must be involved in the initial covalent association with SU in MMLV [20]. Furthermore, mutagenesis studies have shown that infectivity of MMLV is particularly sensitive to amino acid substitutions [21] and insertions [22] in TM. Finally, a construct expressing the MMLV TM subunit alone partially complements a non-fusogenic Env mutant anchored to the virus membrane by a glycosyl-phosphatidylinositol linkage [23]. These observations strongly suggest that efforts to elucidate the mechanism of retrovirus-mediated membrane fusion should target the TM subunit.

Conflicting reports have emerged on the oligomerization order of retrovirus envelope proteins such as murine and feline leukemia viruses [24,25] and HIV [26-28]. Peptide studies of the HIV TM protein located a helical region in an analogous position to the MMLV 4-3 repeats, but the oligomerization state of these peptides was ambiguous [29]. Our results show that the region of the MMLV TM protein adjacent to the fusion peptide is folded, helical, and trimeric, most likely as a three-stranded α -helical coiled coil.

The non-helical region of TM, which includes the immunosuppressive region and the cysteines, is also folded in our model of isolated TM, as shown by the following observations: i) these regions are present in the protease-resistant core; ii) the thermal stability of Mo-55 is considerably greater than that of Mo-36; and iii) only one of the three potential intramolecular disulfide bonds is observed. The observation that a substantial portion of the TM protein is stably folded in the absence of the SU subunit supports our hypothesis that this conformation of TM corresponds to a functional state of the retrovirus envelope.

The TM protein can be isolated on the surface of the virus by a disulfide rearrangement that results in loss of the SU subunit [13]. If shedding of SU proves to be a required step in membrane fusion by retroviruses, then studying the connectivity of the disulfide bonds in MMLV will provide a novel means of monitoring the progress of a virus envelope protein through the receptor-binding and membrane-fusion events. A disulfide bonding pattern for SU released from a murine leukemia virus has been proposed [30]. Our studies show a preference for disulfide bond formation between the first and second cysteines in the isolated TM subunit, leaving the third cysteine as a free thiol. Identification of the disulfide bonding pattern in the SU/TM complex will reveal the nature of the MMLV disulfide rearrangement required for shedding of SU.

Conclusions

We propose that shedding of SU leaves behind a folded, trimeric fusogenic complex (Fig. 6), similar to the low-pH induced structure of influenza HA [3]. In the native state of influenza, the receptor-binding domains inhibit formation of the extended coiled coil, thereby serving as the basis of the “spring-loaded mechanism” [6] for the HA conformational change. It remains to be seen whether the residues corresponding to the coiled coil identified here are in an alternative conformation in the native SU/TM protein complex. In either case, since the trimeric coiled-coil motif is found in analogous positions in the proposed fusion-competent states of the influenza and Moloney murine leukemia viruses, it is likely that coiled coils will also prove to be central to the membrane fusion mechanism of retroviruses.

Materials and Methods

Protein Production and Purification:

Synthetic Peptides

The Mo-36 peptide and a peptide 6 residues shorter at the carboxy terminus, Mo-30, were synthesized using Fmoc chemistry and purified using a method described previously [31]. These peptides contain non-native carboxy-terminal tyrosines to facilitate concentration determination [32]. Mo-30 was essentially unfolded and was not studied further.

Recombinant Peptides

Mo-92 was generated by PCR from a plasmid containing the gene for the MMLV envelope (kindly provided by D. Sanders and R.C. Mulligan, Whitehead Institute), using primers encoding 5' NdeI and 3' BamHI sites. The amplified fragment was subcloned into the pAED4 vector [33]. Mo-55 was constructed from Mo-92 using a new C-terminal PCR primer. Both recombinant peptides include amino-terminal methionines for expression. The sequences of the Mo-92 and Mo-55 inserts, as well as all mutants, were confirmed by dideoxy sequencing (USB). Mo-92 and Mo-55 were expressed in *E. coli* BL21(DE3) pLysS. Colonies were picked directly from the plate, grown in LB to an A_{600} of 0.6 and induced with 1 mM IPTG. Cells were harvested by centrifugation 3 hr later. Mo-92 was purified by lysing cell pellets in glacial acetic acid and centrifuging 20 min at 12,000 rpm in a Sorvall RC-5B equipped with an SS-34 rotor. The supernatant was collected and diluted to 10% acetic acid with H₂O. The solution was filtered and purified by reversed-phase high performance liquid chromatography (HPLC) using a Vydac preparative C18 column. A linear H₂O/acetonitrile gradient containing 0.1% TFA was used at a flow rate of 10 ml/min. Mo-55 was purified by lysing cell pellets in 100 mM NaCl, 25 mM Tris, pH 8.0, 0.1% (v/v) β -mercaptoethanol. The lysate was centrifuged as described above. The soluble fraction was applied under gravity to a 10 ml bed volume DEAE anion-exchange column. The column was washed with 20 ml 100 mM NaCl, 25 mM Tris, pH 8.0, 0.1% (v/v) β -mercaptoethanol. Elution was performed with 25 mM Tris, pH 8.0, 0.1% (v/v) β -mercaptoethanol and a salt gradient from 100 mM to 500 mM NaCl. The eluate fractions were analyzed on 18% SDS-PAGE gels. Fractions containing Mo-55 were pooled and brought to 10% acetic acid, filtered, and purified by reversed-phase HPLC.

Mutagenesis

Cysteine-to-alanine mutations were accomplished by oligonucleotide-directed mutagenesis of the parent constructs [34]. Mutants were purified by the identical method used for corresponding wild-type peptides. Oxidations for the disulfide permutation experiments were performed by incubating 5 mg/ml mutant in 6 M GuHCl, 0.5 mM oxidized glutathione, 0.5 mM reduced glutathione, 20 mM Tris, pH 8.8, for 24 hr at room temperature. Samples were brought to 10% acetic acid, dialyzed at room temperature against 5% acetic acid for 48 hr with two buffer changes, and lyophilized. The absence of free thiols was confirmed by the lack of reaction with Ellman's reagent in 6 M GuHCl [35]. In each case, disulfide bond formation was intramolecular as judged by the absence of higher-order species on non-reducing SDS-PAGE gels.

Protease Digestion

5 mg/ml Mo-92 and 0.8 mg/ml TPCCK-treated trypsin (Sigma) in 25 mM Tris, pH 8.0, were left overnight at room temperature. The digestion mixture was then diluted 5-fold into 5% acetic acid, and the trypsin-resistant fragment was purified by reversed-phase HPLC. The fragment had an observed mass of 6852 daltons. The expected mass for cleavage at Arg 104 is 6856 daltons. 5 mg/ml Mo-92 and 0.5 mg/ml chymotrypsin (Boehringer Mannheim) were incubated at room temperature in 50 mM KPO₄, pH 7.0. Aliquots were diluted 1:10 in 5% acetic acid for mass spectrometry. Cleavage occurred after either Phe 95 or Tyr 96, with observed masses of 5851 and 6007 daltons respectively. Calculated masses are 5843 and 6007 daltons. Longer digestion times favored the shorter fragment. Cleavage of a peptide containing residues from Met 31 to Asp 98 was carried out with 5 mg/ml peptide and 0.5 mg/ml thermolysin in 25 mM Tris, pH 8.0, at 37°C for 6 hours. A mass of 6290 was observed, as compared to the calculated mass of 6289 for a peptide from Val 43 to Asp 98.

Circular Dichroism

CD spectroscopy was performed with an Aviv 62DS spectrometer equipped with a thermoelectric temperature controller. Peptide concentration was determined by tyrosine absorbance at 275 nm using $\epsilon=1500$ (Mo-30, -36, -55), or by tryptophan and tyrosine absorbance at 280 nm using $\epsilon=6880$ (Mo-92) [31]. The Mo-36, Mo-55,

and Mo-92 peptide concentrations were 23 μM , 10 μM , and 10 μM respectively in 50 mM NaPO_4 /150 mM NaCl , pH 7.0, for wavelength scans. Scans were taken at 2°C, and signal was averaged for 5 sec. All thermal melts were performed at 10 μM in the same buffer. The CD signal was measured in 2-deg steps, with a 1.5-min equilibration at each temperature. Signal was averaged for 10 sec at each temperature. The T_m for each peptide was determined from the peak in the first derivative of the ellipticity vs. $1/T$ curve [36]. All T_m values were determined from thermal unfolding experiments that were >90% reversible. The number of helical residues in each TM peptide was estimated from $[\theta]_{222}$ by assuming that a value of -33,000 deg $\text{cm}^2 \text{dmol}^{-1}$ corresponds to a helix content of 100% [37]. The unfolded baseline of -4000 deg $\text{cm}^2 \text{dmol}^{-1}$ was taken as 0% helical.

Analytical Ultracentrifugation

Equilibrium ultracentrifugation studies [38] were done with a Beckman XL-A analytical ultracentrifuge, using an An-60 Ti rotor at 20°C. Mo-92 in 1 ml 50 mM Tris, pH 8.8., was dialyzed overnight against 500 ml 50 mM Tris, pH 8.8, 200 mM NaCl , 0.5 mM DTT. Mo-55 in 1 ml 100 mM Tris, pH 8.8, was dialyzed overnight against 500 ml 50 mM Tris, pH 8.8, 100 mM NaCl , 2 mM DTT. Fractions of the sample were then diluted 1:2 and 1:8 with dialysate to generate protein solutions of approximately 100 μM , 33 μM , and 11 μM . Low concentrations of DTT were used as a precaution against intermolecular disulfide bond formation. However, subsequent studies showed reducing agent to be unnecessary. Indeed, similar results were obtained for oxidized Mo-92.3, which does not contain a free thiol, as for Mo-92 in presence or absence of DTT. Mo-92 samples were spun in a 6-sector cell at rotor speeds of 15,000 and 17,000 rpm. Mo-55 was spun at 20,000 rpm, 22,000 rpm, and 24,000 rpm. Data for Mo-92 were collected for each speed and for each protein concentration at 230 nm, 260 nm, 280 nm, and 320 nm. Data for Mo-55 were collected at 229 nm, 247 nm, 280 nm, and 300 nm. Data were analyzed using the equation $M = [2RT / (1 - \bar{v}\rho)\omega^2] [d(\ln c) / dr^2]$ with \bar{v} as 0.732 cm^3/g for Mo-92 and 0.730 cm^3/g for Mo-55. In both cases, r was 1.01 g/cm.

NTCB Cleavage

Mo-92 and Mo-55 were oxidized at 100 μM by stirring overnight in 20 mM Tris, pH 8.8. Samples were brought to 5% acetic acid, dialyzed against 5% acetic acid for 24 hr, and lyophilized. Oxidized Mo-92, Mo-55, or reduced cysteine-to-alanine

mutants were dissolved at 1 mM in 20 mM Tris, pH 8.8. NTCB was added to 10 mM, followed by GuHCl to saturation. The sample was placed in the dark for 15 min at 37°C. 1N NaOH was added until pH ~9.5 was achieved, and the cleavage reaction was allowed to proceed in the dark for 8 hr at 37°C [17]. An aliquot was diluted 1:10 in 70% CH₃CN, 0.1% TFA for mass spectrometry. The observed masses for NTCB cleavage of Mo-92 were 5597 and 5081 daltons. Expected masses for the amino- and carboxy-terminal fragments of cleavage before the third cysteine in Mo-92 are 5593 and 5063 daltons respectively. A methionine sulfoxide in the carboxy-terminal fragment would lead to an expected mass of 5079 daltons. Cleavage of oxidized Mo-55 yielded an amino-terminal fragment of 5600 daltons.

Mass Spectrometry

Laser desorption mass spectrometry was performed on a Voyager Elite BioSpectrometry Research Station (PerSeptive Biosystems), using a matrix of 10 mg/ml α -cyano-4-hydroxycinnamic acid in 70% CH₃CN, 0.1% TFA.

Acknowledgments

We thank Mike Burgess for synthesis of Mo-30 and Mo-36 peptides, Shiufun Cheung for Figure 6, Chave Carr for inspiring peptide studies of MMLV, and Pehr Harbury for insightful suggestions of experimental approaches. We also thank other members of the Kim lab for discussions and comments on the manuscript. D.F. is a National Science Foundation Predoctoral Fellow. This research was supported by the Howard Hughes Medical Institute.

References

1. Hunter E, Swanstrom R: **Retrovirus envelope glycoproteins.** *Cur. Top. in Microbiol. and Immun.* 1990, **157**:187-253.
2. Wilson IA, Skehel JJ, Wiley DC: **Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution.** *Nature* 1981, **289**:366-373.
3. Bullough PA, Hughson FM, Skehel JJ, Wiley DC: **Structure of influenza haemagglutinin at the pH of membrane fusion.** *Nature* 1994, **371**:37-43.
4. White JM, Wilson IA: **Anti-peptide antibodies direct steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin.** *J. Cell. Biol* 1987, **105**:2887-2896.
5. Kemble GW, Bodian DL, Rose J, Wilson IA, White JM: **Intermonomer disulfide bonds impair the fusion activity of influenza virus hemagglutinin.** *J. Virol.* 1992, **66**:4940-4950.
6. Carr CM, Kim PS: **A spring-loaded mechanism for the conformational change of influenza hemagglutinin.** *Cell* 1993, **73**:823-832.
7. White JM: **Viral and cellular membrane fusion proteins.** *Ann. Rev. Physiol.* 1990, **52**:675-697.
8. Stegmann T, Delfino JM, Richards FM, Helenius A: **The HA2 subunit of influenza hemagglutinin inserts into the target membrane prior to fusion.** *J. Biol. Chem.* 1991, **266**:18404-18410.
9. Tsurudome M, Gluck R, Graf R, Falchetto R, Schaller U, Brunner J: **Lipid interactions of the hemagglutinin HA2 NH2-terminal segment during influenza virus-induced membrane fusion.** *J. Biol. Chem.* 1992, **267**:20225-20232.
10. Delwart EL, Mosialos G: **Retroviral envelope glycoproteins contain a "leucine zipper"-like repeat.** *AIDS Res. and Human Retroviruses* 1990, **6**:703-706.
11. Chambers P, Pringle CR, Easton AJ: **Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins.** *J. General Virology* 1990, **71**:3075-3080.
12. Moore JP, McKeating JA, Weiss RA, Sattentau QJ: **Dissociation of gp120 from HIV-1 virions induced by soluble CD4.** *Science* 1990, **250**:1139-1142.

13. Pinter A, Leiman-Hurwitz J, Fleissner E: **The nature of the association between the murine leukemia virus envelope proteins.** *Virology* 1978, 91:345-351.
14. Berger BA, Wilson DB, Wolf E, Tonchev T, Milla M, Kim PS: **Predicting coiled coils using pairwise-residue correlations.** *Proc. Natl. Acad. Sci. USA* 1995, 92:8259-8263.
15. Cianciolo GJ, Kipnis RJ, Snyderman R: **Similarity between p15E of murine and feline leukemia viruses and p21 of HTLV.** *Nature* 1984, 311:515.
16. Schulz TF, Jameson BA, Lopalco L, Siccardi AG, Weiss RA, Moore JP: **Conserved structural features in the interaction between retroviral surface and transmembrane glycoproteins?** *AIDS Res. and Human Retroviruses* 1992, 8:1571-1580.
17. Stark GR: **Cleavage at cysteine after cyanylation.** *Meth. Enzym.* 1977, 47:129-132.
18. Cianciolo GJ, Copeland TD, Oroszlan S, Snyderman R: **Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins.** *Science* 1985, 230:453-455.
19. Ogasawara M, Cianciolo GJ, Snyderman R, Mitani M, Good RA, Day NK: **Human INF- γ production is inhibited by a synthetic peptide homologous to retroviral envelope protein.** *J. Immunol.* 1988, 141:614-619.
20. Pinter A, Honnen WJ: **Topography of Murine Leukemia Virus Envelope Proteins: Characterization of Transmembrane Components.** *J. Virol.* 1983, 46:1056-1060.
21. Berkowitz RD, Goff SP: **Point mutations in Moloney murine leukemia virus envelope protein: effects on infectivity, virion association, and superinfection resistance.** *Virology* 1993, 196:748-757.
22. Gray KD, Roth MJ: **Mutational analysis of the envelope gene of Moloney murine leukemia virus.** *J. Virol.* 1993, 67:3489-3496.
23. Ragheb JA, Anderson WF: **Uncoupled expression of Moloney murine leukemia virus envelope polypeptides SU and TM: a functional analysis of the role of TM domains in viral entry.** *J. Virol.* 1994, 68:3207-3219.
24. Pinter A, Fleissner E: **Structural studies of retroviruses: characterization of oligomeric complexes of murine and feline leukemia virus envelope and core components formed upon crosslinking.** *J. Virol.* 1979, 30:157-165.

25. Kamps CA, Lin Y-C, Wong PKY: **Oligomerization and transport of the envelope protein of Moloney murine leukemia virus-TB and of *ts1*, a neurovirulent temperature-sensitive mutant of MMLV-TB.** *Virology* 1991, **184**:687-694.
26. Einfeld D, Hunter E: **Oligomeric structure of a prototype retrovirus glycoprotein.** *Proc. Natl. Acad. Sci USA* 1988, **85**:8688-86.
27. Schawaller M, Smith, GE, Skehel JJ, Wiley DC: **Studies with crosslinking reagents on the oligomeric structure of the env glycoprotein of HIV.** *Virology* 1989, **172**:367-369.
28. Thomas DJ, Wall JS, Hainfeld JF, Kaczorek M, Body FP, Trus BL, Eiserling FA, Steven AC: **gp160, the envelope glycoprotein of human immunodeficiency virus type 1, is a dimer of 125-kilodalton subunits stabilized through interactions between their gp41 domains.** *J. Virol.* 1991, **65**:3797-3803.
29. Wild C, Oas T, McDanal C, Bolognesi D, Matthews T: **A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and inhibition.** *Proc. Natl. Acad. Sci. USA* 1992, **89**:10537-10541.
30. Linder M, Linder D, Hahnen J, Schott H-H, Stirm S: **Localization of the intrachain disulfide bonds of the envelope glycoprotein 71 from Friend murine leukemia virus.** *Eur. J. Biochem.* 1992, **203**:65-73.
31. O'Shea EK, Lumb KJ, Kim PS: **Peptide "velcro": design of a heterodimeric coiled coil.** *Current Biology* 1993, **3**:658-667.
32. Edelhoch H: **Spectroscopic determination of tryptophan and tyrosine in proteins.** *Biochemistry* 1967, **6**:1948-1954.
33. Doering DS: **Functional and Structural Studies of a Small f-actin Binding Protein.** Ph.D. thesis, Massachusetts Institute of Technology, Cambridge; 1992.
34. Kunkel TA, Roberts JD, Zakour RA: **Rapid and efficient site-selected mutagenesis without phenotypic selection.** *Meth. Enzym.* 1987, **154**:367-382.
35. Riddles PW, Blakeley RL, Zerner B: **Reassessment of Ellman's reagent.** *Meth. Enzym.* 1983, **91**:49-60.
36. Cantor C, Schimmel P: *Biophysical Chemistry*, Part III. New York: W.H. Freeman and Company; 1980.
37. O'Shea EK, Rutkowski R, Kim PS: **Evidence that the leucine zipper is a coiled coil.** *Science* 1989, **243**:538-542.

38. Laue TM, Shah BD, Ridgeway TM, Pelletier SL: **Computer-aided interpretation of analytical sedimentation data for proteins.** In *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Edited by Harding SE, Rowe AJ, JC Horton JC. Cambridge, UK: Royal Society of Chemistry; 1992:90-125.

Table 1. Helix content and stability of MMLV TM peptides monitored by circular dichroism.

Peptide	$-\theta_{222} \times 10^{-3}$	Helical residues	T_m(°C)
Mo-36	26	~28	30
Mo-55	17	~28	80
Mo-92	21	~54	86

The ellipticity at 222 nm and the temperature midpoint for thermal denaturation are indicated for each MMLV TM fragment. Estimates of the number of helical residues in each peptide and determinations of T_m values were made as described in Materials and Methods.

Table 2. Disulfide permutation of MMLV TM peptides.

Peptide	T _m (°C)	-[θ] ₂₂₂ × 10 ⁻³ at 10°C
Mo-92.1	70	20
Mo-92.2	76	20
Mo-92.3	86	20
Mo-55.1	62	17
Mo-55.2	72	17
Mo-55.3	80	17

Disulfide Permutation of Mo-92 and Mo-55. The table displays the T_m values and the folded CD signal at 10°C for the Mo-92 and Mo-55 mutants in the oxidized form. In each set of mutants, the number after the decimal point indicates the position of the alanine. Thus, Mo-92.1 refers to a mutant with a disulfide bond between C2 and C3, and an alanine substitution at C1.

Fig. 1. Primary Structure of the TM protein in Moloney Murine Leukemia Virus. A linear map of TM is shown with the sequence of Mo-92 specified. Regions modeled by the three peptides, named for the number of residues they contain, are indicated above the sequence. Core residues of the 4-3 hydrophobic repeat are enlarged. The three cysteine thiols are numbered. The trypsin cleavage site is marked by a pink arrow, while the two chymotrypsin sites are indicated by blue arrows. A description of the immunosuppressive region is found in the Discussion.

Figure 1

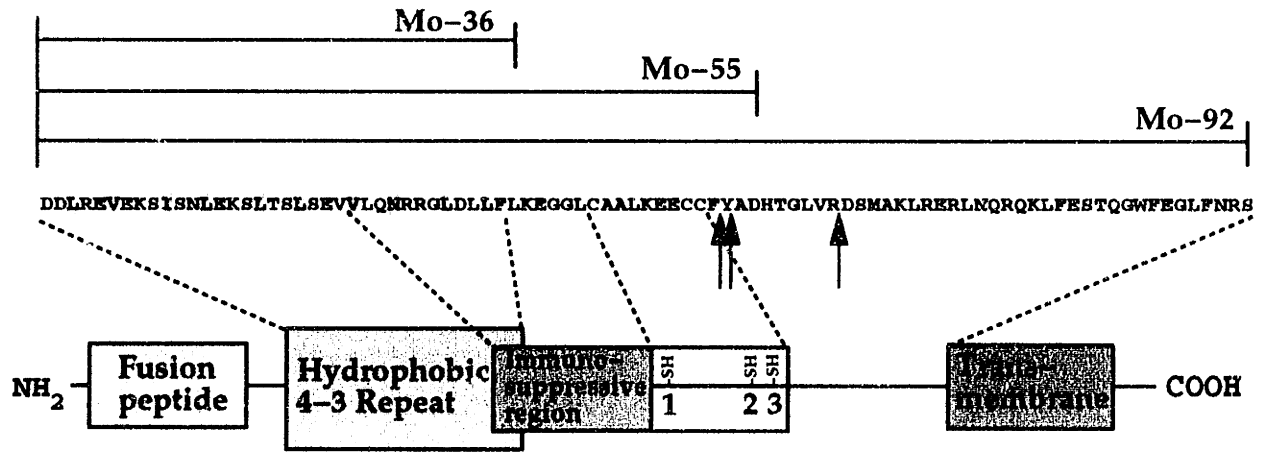


Fig. 2. Helical Wheel Representation of Mo-36. The view is down the helical axis starting at the amino-terminus (position f) of the peptide, which corresponds to residue Asp 45 of TM. Core (a and d) positions are indicated in blue, while other positions are yellow.

Figure 2

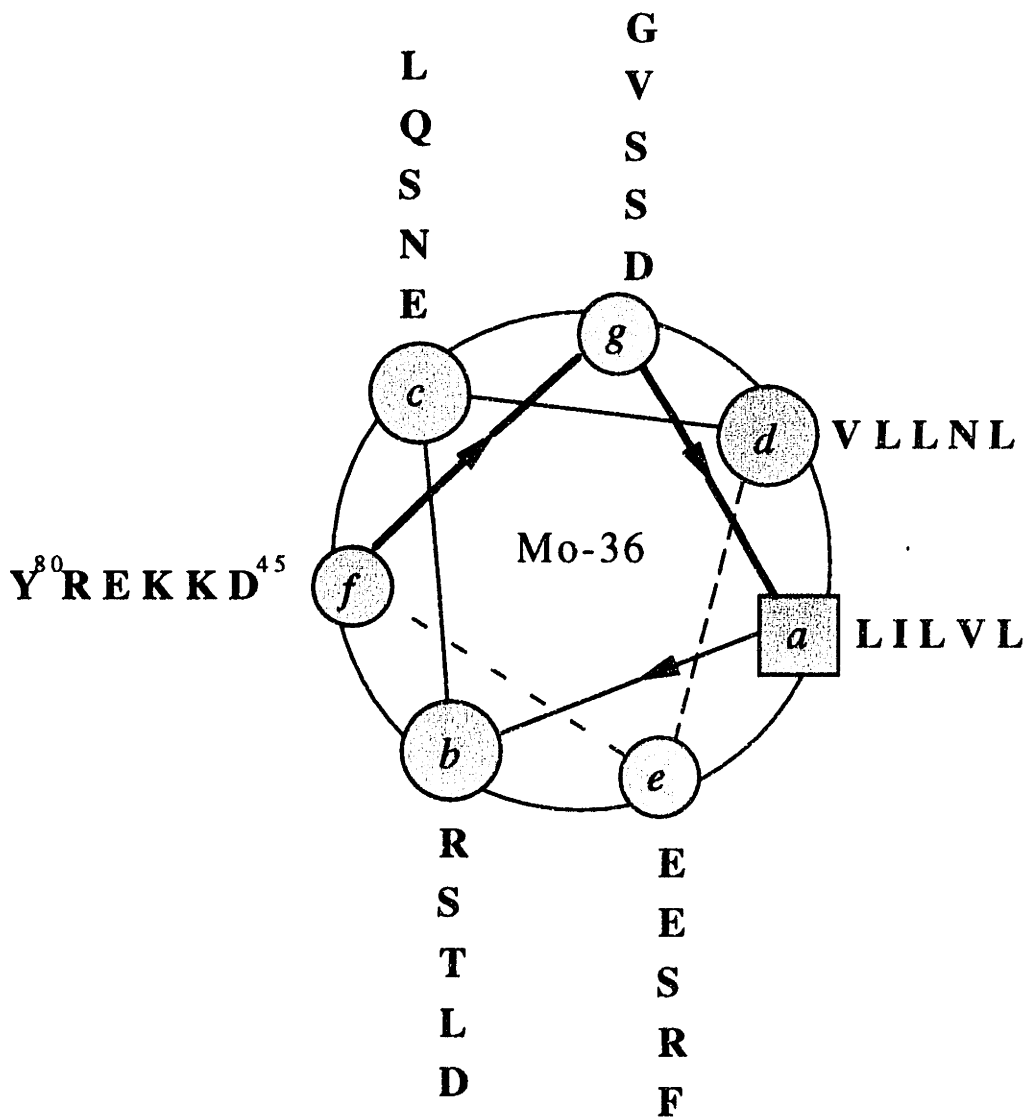


Fig. 3. Circular Dichroism Spectra of TM Fragments. Each fragment displays a characteristic helical spectrum at 2°C with double minima at 222 nm and 208 nm.

Figure 3

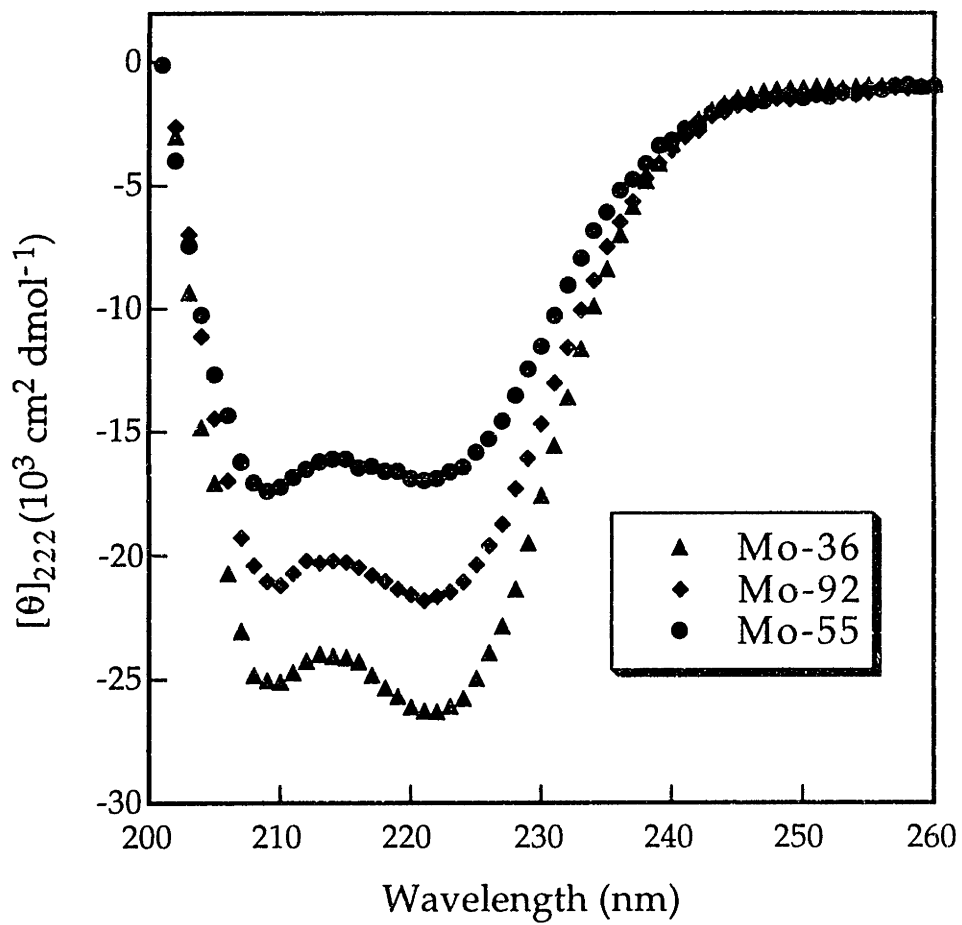


Fig. 4. An estimate (see Materials and Methods *Circular Dichroism*) of the number of helical residues in each fragment is plotted as a function of temperature. Mo-36 and Mo-55 lose most of their helicity during the cooperative unfolding transitions. On the other hand, approximately 20 helical residues, almost 40% of the total helix content, unfold in Mo-92 before the cooperative transition occurs.

Figure 4

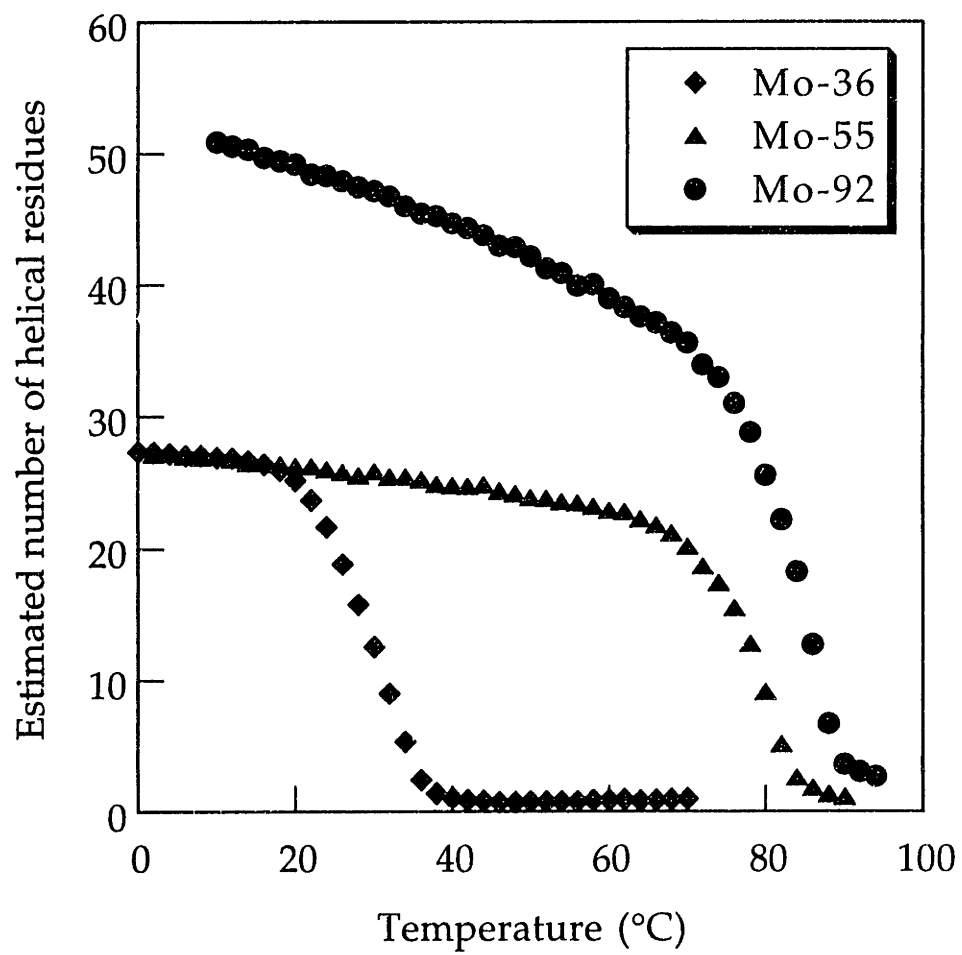
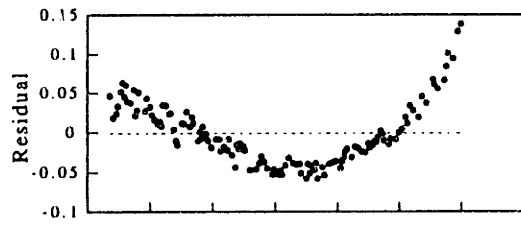


Fig. 5. Equilibrium Sedimentation of TM Peptides. Representative data are plotted as $\ln(\text{absorbance})$ vs the square of the radius from the axis of rotation. The slope is proportional to molecular weight (see Experimental Procedures). Dashed lines with increasing slopes indicate calculated values for monomeric (1), dimeric (2), trimeric (3), and tetrameric (4) peptides. (A) Data for Mo-92 are consistent with a trimeric model, but with systematic residuals. (B) Data for Mo-55 fit closely to a trimeric model, but leave no systematic residuals.

Figure 5

A



B

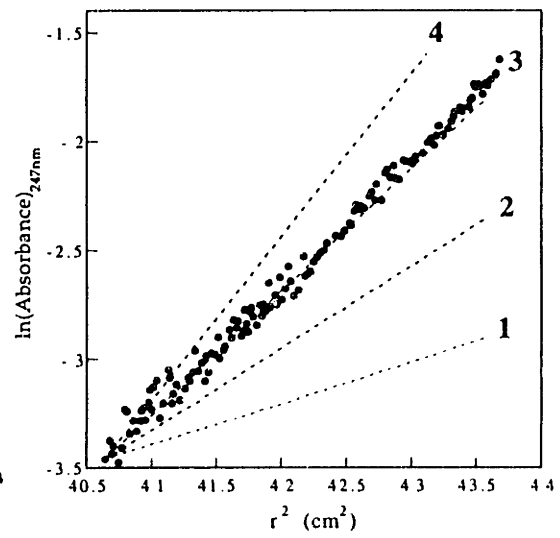
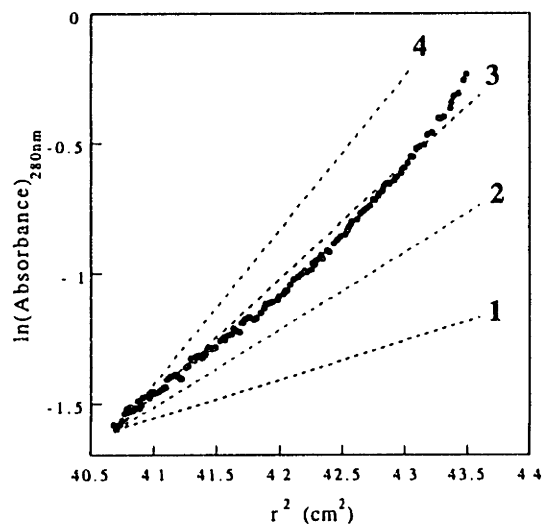
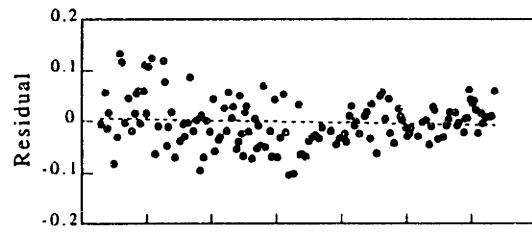
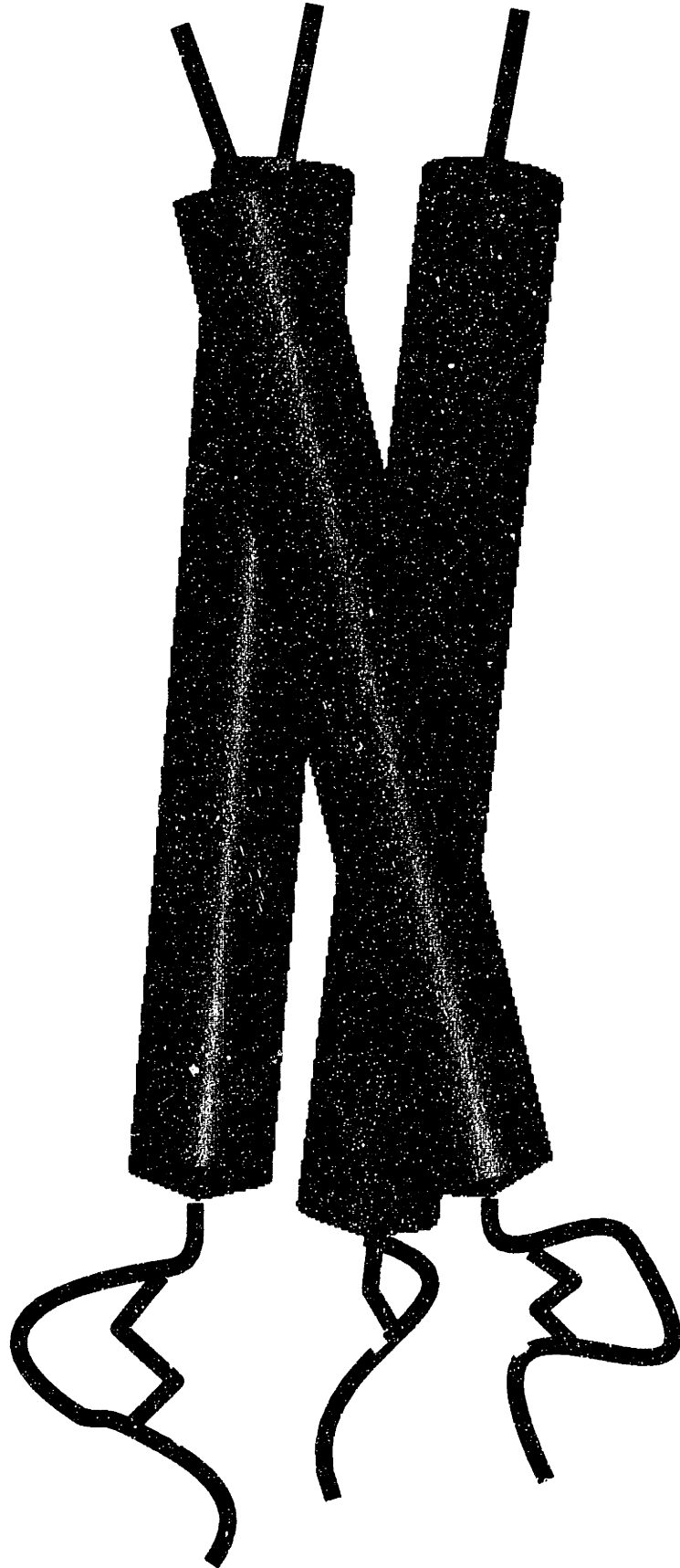


Fig. 6. Proposed Model for the Core of the Fusion-Active Conformation of MMLV TM Protein. Magenta cylinders represent the three-stranded coiled coil adjacent to the fusion peptides (blue). For simplicity, supercoiling of the coiled coil is not depicted. The green loops represent the immunosuppressive and the cysteine-rich regions. Contained in the non-helical region are three cysteines portrayed as half lightning bolts. The first two cysteines readily form a disulfide bond (shown as a full lightning bolt), as judged by NTCB cleavage data and thermal stabilities of cysteine-to-alanine mutants (see text).

Figure 6



CHAPTER 3

RETROVIRUS ENVELOPE DOMAIN AT 1.7Å RESOLUTION

We report the crystal structure of an extraviral segment of a retrovirus envelope protein, the Moloney murine leukemia virus (MoMuLV) transmembrane (TM) subunit. This segment, which comprises a region of the MoMuLV TM protein analogous to that contained within the X-ray crystal structure of low-pH converted influenza hemagglutinin, contains a trimeric coiled coil, with a hydrophobic cluster at its base, and a strand that packs in an antiparallel orientation against the coiled coil. This structure gives the first high-resolution insight into the retrovirus surface and serves as a model for a wide range of viral fusion proteins; key residues in this structure are conserved among C- and D-type retroviruses and the filovirus ebola.

Surface glycoproteins target enveloped viruses to their host cell receptors and mediate fusion of the viral and cellular membranes¹. The envelope proteins of retroviruses are synthesized as a single chain, which is then proteolytically cleaved into an amino-terminal “surface” subunit (SU) and a carboxy-terminal transmembrane subunit (TM). The SU subunit binds the receptor, while the TM subunit contains the hydrophobic “fusion peptide” at its amino terminus¹. To provide a structural framework for studies of retrovirus-mediated membrane fusion, we sought a high-resolution model of a retrovirus TM protein.

Moloney murine leukemia virus (MoMuLV) was chosen for study. MoMuLV SU and TM are linked by a labile disulfide bond that can be stabilized by thiol blocking agents², presumably by preventing a free thiol in the envelope complex from initiating thiol-disulfide rearrangement. This rearrangement eliminates the covalent association between SU and TM². The ectodomain of the TM protein, lacking the fusion peptide, has been shown to fold into a stable structure in the absence of the SU subunit³. Furthermore, because the MoMuLV TM subunit lacks glycosylation sites, it can be represented accurately by synthetic and recombinant peptides.

Conserved patterns in viral TM amino acid sequences

MoMuLV TM (Fig. 1a) contains three amino acid sequence motifs that are highly conserved among C- and D-type retroviruses with hosts ranging from mice to humans. The region from residues 43 to 78 exhibits a 4-3 repeat of hydrophobic amino acids predicted to form a coiled coil^{4,5}. An immunosuppressive peptide sequence (ISP), spanning residues 69 to 85, inhibits cellular immune responses^{6,7}. Finally, three cysteine residues form the unvarying pattern CX₆CC from residues 86 to 94. The ISP and the cysteine motif are also found in envelope proteins the filovirus family⁸, which differs morphologically from the retrovirus family and carries an RNA genome of opposite polarity⁹.

Previous solution studies³ demonstrated that a peptide containing the 4-3 hydrophobic repeat, the ISP, and the conserved cysteine motif displays the significant features of the recombinant ectodomain. The peptide, named Mo-55 for its length in residues, is trimeric. In addition, it assumes the identical disulfide bond connectivity as a larger segment lacking only the fusion peptide. Finally, its thermostability differs from that of the larger peptide by only 6°C. Although the segment of the ectodomain lacking only the fusion peptide does have additional helical residues not present in Mo-55, their structure is lost non-cooperatively upon heating. Furthermore, treatment of the recombinant ectodomain with proteases results in rapid removal of these residues, yielding protease-resistant fragments similar to Mo-55. The entire ectodomain, containing the fusion peptide, aggregates in aqueous solution, suggesting that the fusion peptides are exposed and are not buried in the core of the structure. Thus, the information required to specify the fold of the stable core of the isolated TM subunit is contained within Mo-55.

Structure determination

Mo-55 was crystallized as described in **Methods**. The Mo-55 structure (Fig. 1b) was determined by multiple isomorphous replacement (MIR) using selenomethionine and thulium (Tm) derivatives. An initial model built into a 2.4 Å resolution MIR map was improved by rounds of refinement and rebuilding using O¹⁰. The current model includes residues 46 through 97 of the TM protein, 40 water molecules, and a chloride ion. This model is in good agreement with both the experimental data to 1.7 Å (working R-factor 16.9%, free R-factor 23.4%) and standard geometry (rmsd bonds 0.017 Å, rmsd angles 2.39°). Details of the data collection and refinement are given in Table 1.

MoMuLV TM architecture

The core of Mo-55 is a 33-residue, trimeric coiled coil. Six leucines, two valines, one isoleucine, and one asparagine from each monomer are buried at the center of the coiled coil. The three Asn 71 amide groups of the trimer trap an X-ray scatterer (Fig. 2a, b) identified by solution studies as a chloride ion (Table 2).

Buried polar interactions have been shown to promote specificity in the folding of model coiled coils^{11,12}, and the interaction of chloride with the buried Asn 71 residues may prove to be important in MoMuLV.

At the base of the Mo-55 coiled coil, the polypeptide chain reverses direction and forms a short α -helix (residues 85-89) nearly perpendicular to the coiled-coil axis. Hydrophobic clusters are formed by Phe 79, Leu 85, Ala 88, and Leu 89, together with the coiled-coil core residues Leu 75 and Leu 78, as well as Leu 77 and Leu 80 from an adjacent monomer. The carboxy-terminal residues (90-97) of Mo-55 pack primarily against the counter-clockwise adjacent monomer, as viewed from the top of the coiled coil. Cys 93 in the extended region forms an intramolecular disulfide bond with Cys 86 in the short helix. Cys 94 has a free thiol.

Implications for viral entry

X-ray crystallographic studies of enveloped virus surface proteins illustrate the remarkable diversity of native envelope protein conformations. TBE E protein is a flat, elongated dimer that lies parallel to the surface of the virus¹³, whereas influenza hemagglutinin (HA) is a trimeric "spike" that extends approximately 150 Å from the viral membrane¹⁴. These two structures suggest the possibility that each virus family exhibits a unique solution to the membrane-fusion problem.

It is striking, therefore, that the structure of Mo-55 shares many features in common with low pH-converted influenza HA (TBHA₂)^{15,16} demonstrating that, even in the absence of sequence homology or any obvious evolutionary relationship, diverse viruses may use similar scaffolds for presenting fusion peptides to target cells (Fig. 3). Both TBHA₂ and Mo-55 have trimeric coiled coils in similar positions in their primary sequences. In addition, hydrophobic clusters stabilize both the Mo-55 and TBHA₂ trimers at the bases of their coiled coils. Finally, the carboxy-terminal residues of both fragments extend back toward the amino termini of the coiled coils, apparently away from the viral membranes. The structural similarities between the envelope proteins of the

retrovirus MoMuLV and the well-studied orthomyxovirus influenza open the rich literature on influenza HA (for review see ref. 17) to the question of the retrovirus membrane-fusion mechanism.

Within their common structural framework, however, Mo-55 and TBHA₂ do contain notable differences. First, the TBHA₂ coiled coil spans 65 residues, while that of Mo-55 is only half as long. Furthermore, the Mo-55 coiled coil displays a regular superhelical twist of the order observed in a model trimeric coiled coil¹⁸, and the 4-3 periodicity is preserved throughout the length of the coiled coil. In contrast, the TBHA₂ structure contains two regions with an irregular 3-4-4-3 pattern of core residues. Due to these skips, the TBHA₂ coiled coil has a pitch of 300-400 Å, while that of Mo-55 is approximately 175 Å. The location of disulfide bonds in the two structures is distinct, and the backbone conformations differ between the hydrophobic clusters. Finally, a short helix packs nearly antiparallel to the coiled coil in TBHA₂, a feature absent from the Mo-55 structure.

It is likely that the Mo-55 structure represents the core of the fusion active conformation of MoMuLV TM, at a post-binding stage in which SU is no longer required for envelope function. For influenza, the low pH in the endosome leads to displacement of the receptor-binding HA₁ domain with a consequent dramatic rearrangement of the transmembrane polypeptide HA₂. We expect that the SU subunit of retroviruses similarly aids in burial of the fusion peptide prior to activation, but the structure of TM in the SU/TM complex is unknown. Whether displacement or shedding of SU in retrovirus envelope proteins leads to a conformational change in TM, perhaps in a “spring-loaded” manner¹⁹, remains to be determined.

Methods

Crystallization and data collection

Mo-55 was purified as described previously. Crystals in space group $P2_13$ ($a=b=c=53.1$ Å) were grown by vapor diffusion from 1.6 M NaCl, 50 mM KPO_4 , pH 5.6, 2.5% PEG 8000. A peptide (called Mo-55.3) with a cysteine-to-alanine mutation at residue 94 formed crystals larger than those of the wild-type peptide, although the space group and unit-cell constants for both forms were identical. Crystals were transferred to 1.6 M NaAc, pH 5.6, 10% PEG 8000 at least 24 hours before data collection. Diffraction data were collected on a Siemens-Xentronics multiwire area detector mounted on an Elliott GX13 X-ray generator. Reflected X-ray intensities were integrated with the program Buddha²⁰. Subsequent data analysis was done with CCP4 programs²¹.

Phase determination

The Se derivative was prepared by introducing a leucine-to-methionine mutation at residue 47 of TM and incorporating selenomethionine at this position during expression of Mo-55.3²². The Tm^{3+} derivative was acquired by soaking crystals of Mo-55.3 in 1.6 M NaAc, pH 5.6, 10% PEG 8000, 50 mM $Tm_2(SO_4)_3$ for 24 hours. The Se and Tm^{3+} sites were located by Patterson methods and refined using the program MLPHARE (CCP4 program suite). Anomalous signal from the Tm^{3+} derivative was included. MIR maps calculated to 2.4 Å revealed connected backbone electron density from residue 47 to residue 96. Density for most of the side chains and many of the carbonyl groups was also clearly visible, facilitating the chain trace.

Structure refinement

Positional and B-factor refinements of the initial model were performed using X-PLOR²³, first against data to 2.1 Å. Ten fixed waters were added, and a bulk-solvent correction was applied. Further cycles of rebuilding and refinement were carried out using data to 1.7 Å. Data from wild-type crystals were also

collected, and a difference map phased with the refined mutant model was used to position the Cys 94 side chain. The most common cysteine rotamer fit clearly into the largest peak in the difference map. However, extra density remained near the cysteine sulfur atom, indicating partial derivatization. This density was not modeled. The wild-type model was subjected to B-factor refinement and bulk-solvent correction, yielding an R factor of 0.190 against data (20 to 1.8 Å) for the wild-type protein. Refined coordinates will be deposited in the Protein Data Bank (Brookhaven National Laboratory, Upton, NY).

Thermal denaturation

Proteins with cysteine-to-alanine mutations at position 94 (Mo-55.3) were used to eliminate the possibility of intermolecular disulfide bond formation or thiol rearrangement during denaturation. Thermal denaturation was monitored by circular dichroism at 222 nm using 10 μM peptide. Buffer for all melting experiments was 50 mM NaPO₄, pH 7.0. CD signal was measured and T_m values were calculated as described previously³.

Acknowledgements

We thank J. Berger, S. Gamblin, D. Rodgers, and members of the Harvard structural biology laboratory for their patience and generosity with time, equipment, and advice. We also thank T. Ellenberger for protocols for selenomethionine incorporation. Thanks to P. Harbury, B. Schulman, and other members of the Kim laboratory for support and encouragement. D.F. is a National Science Foundation Predoctoral Fellow. This research was supported by the Howard Hughes Medical Institute.

References

1. Hunter, E. & Swanstrom, R. Retrovirus envelope glycoproteins. *Curr. Top. in Microbiol. and Immun.* **157**, 187-253 (1990).
2. Pinter, A., Lieman-Hurwitz, J. & Fleissner, E. The nature of the association between the murine leukemia virus envelope proteins. *Virology* **91**, 345-351 (1978).
3. Fass, D. & Kim, P.S. Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin. *Curr. Biol.* **5**, 1377-1383 (1995).
4. Chambers, P., Pringle, C.R. & Easton, A.J. Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *J. Gen. Virol.* **71**, 3075-3080 (1990).
5. Delwart, E.L., Mosialos, G. & Gilmore, T. Retroviral envelope glycoproteins contain a "leucine zipper"-like repeat. *AIDS Res. and Human Retroviruses* **6**, 703-706 (1990).
6. Cianciolo, G.J., Copeland, T.D., Oroszlan, S. & Snyderman, R. Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. *Science* **230**, 453-455 (1985).
7. Ogasawara, M., Cianciolo, G.J., Snyderman, R., Mitani, M., Good, R.A. & Day, N.K. Human INF- γ production is inhibited by a synthetic peptide homologous to retroviral envelope protein. *J. Immunol.* **141**, 614-619 (1988).
8. Volchkov, V.E., Blinov, V.M. & Netesov, S.V. The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. *FEBS Letters* **305**, 181-184 (1992).
9. Feldman, H., Will, C., Schikore, M., Slenczka, W. & Klenk, H.-D. Glycosylation and oligomerization of the spike protein of Marburg virus. *Virology* **182**, 353-356 (1991).
10. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110-119 (1991).
11. Harbury, P.B., Zhang, T., Kim, P.S. & Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**, 1401-1407 (1993).

12. Lumb, K.J. & Kim, P.S. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochem.* **34**, 8642-8648 (1995).
13. Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C. & Harrison, S.C. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **375**, 291-298 (1995).
14. Wilson, I.A., Skehel, J.J. & Wiley, D.C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366-373 (1981).
15. Bullough, P.A., Hughson, F.M., Skehel, J.J. & Wiley, D.C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43 (1994).
16. Bullough, P.A., Hughson, F.M., Treharne, A.C., Ruigrok, R.W., Skehel, J.J. & Wiley, D.C. Crystals of a fragment of influenza haemagglutinin in the low pH induced conformation. *J. Mol. Biol.* **236**, 1262-1265 (1994).
17. Hughson, F.M. Structural characterization of viral fusion proteins. *Curr. Biol.* **5**, 265-274 (1995).
18. Harbury, P.B., Kim, P.S. & Alber, T. Crystal structure of an isoleucine-zipper trimer. *Nature* **371**, 80-83 (1994).
19. Carr, C.M. & Kim, P.S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**, 823-832 (1993).
20. Blum, M., Metcalf, P., Harrison, S.C. & Wiley, D.C. A system for collection and on-line integration of X-ray diffraction data from a multiwire area detector. *J. Appl. Crystallogr.* **20**, 235-242 (1987).
21. CCP4, a Suite of Programs for Protein Crystallography (SERC (UK) Collaborative Computing Project No. 4, Daresbury Laboratory, Warrington, 1979).
22. Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105-124 (1993).
23. Brünger, A.T. *X-PLOR (Version 3.1:) A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).

24. Stegmann, T., Delfino, J.M., Richards, F.M. & Helenius, A. The HA₂ subunit of influenza hemagglutinin inserts into the target membrane prior to fusion. *J. Biol. Chem.* **266**, 18404-18410 (1991).
25. Tsurudome, M., Gluck, R., Graf, R., Falchetto, R., Schaller, U. & Brunner, J. Lipid interactions of the hemagglutinin HA₂ NH₂-terminal segment during influenza virus-induced membrane fusion. *J. Biol. Chem.* **267**, 20225-20232 (1992).
26. Pinter, A. & Honnen, W.J. Topography of murine leukemia virus envelope proteins: characterization of transmembrane components *J. Virol.* **46**, 1056-1060 (1983).
27. Weast, R.C., Ed., CRC Handbook of Chemistry and Physics (CRC Press, Inc. Boca Raton, Florida, 1980), p. F-216.
28. Kraulis, P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 924-950 (1991).
29. White, J.M. & Wilson, I.A. Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J. Cell. Biol.* **105**, 2887-2896 (1987).
30. Kemble, G.W., Bodian, D.L., Rose, J., Wilson, I.A. & White, J.M. Intermonomer disulfide bonds impair the fusion activity of influenza virus hemagglutinin. *J. Virol.* **66**, 4940-4950 (1992).

Fig. 1 a, Primary structure of MoMuLV TM protein. TM contains no glycosylation sites. The region between the arrows is Mo-55, the peptide expressed in *E. coli* and crystallized. A hydrophobic and glycine-rich sequence of approximately 30 residues at the amino terminus of TM (underlined) is characteristic of a fusion peptide, which has been shown, for influenza, to insert into the target membrane during viral entry^{24,25}. Near the carboxy terminus is a second hydrophobic stretch (underlined), containing residues 136 to 163, that anchors TM to the virus envelope²⁶. Helical regions in the Mo-55 structure are indicated by striped poles under the sequence. **b**, Ribbon diagram of the Mo-55 trimer. Yellow balls indicate sulfur atoms in cysteine side chains, of which two per monomer are in an intramolecular disulfide bond. Residues in red (shown only in one monomer) are conserved among representative C-type retroviruses (MoMuLV, feline leukemia virus, human T-cell leukemia virus type 1, Gibbon ape leukemia virus, and bovine leukemia virus), as well as the D-type retrovirus Mason-Pfizer monkey virus, and the filovirus ebola.

Figure 1A

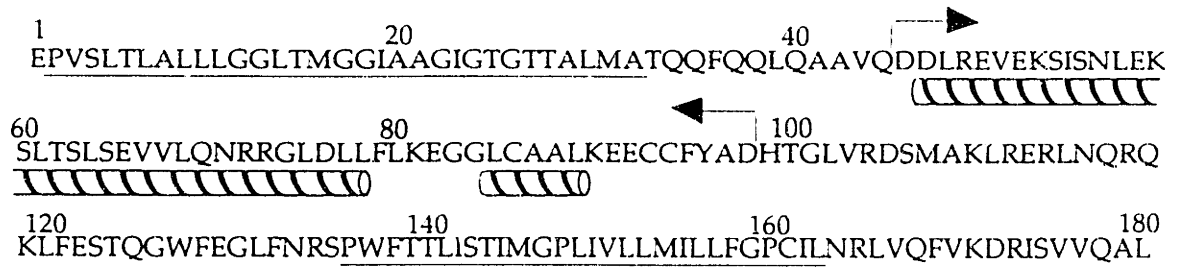


Figure 1B

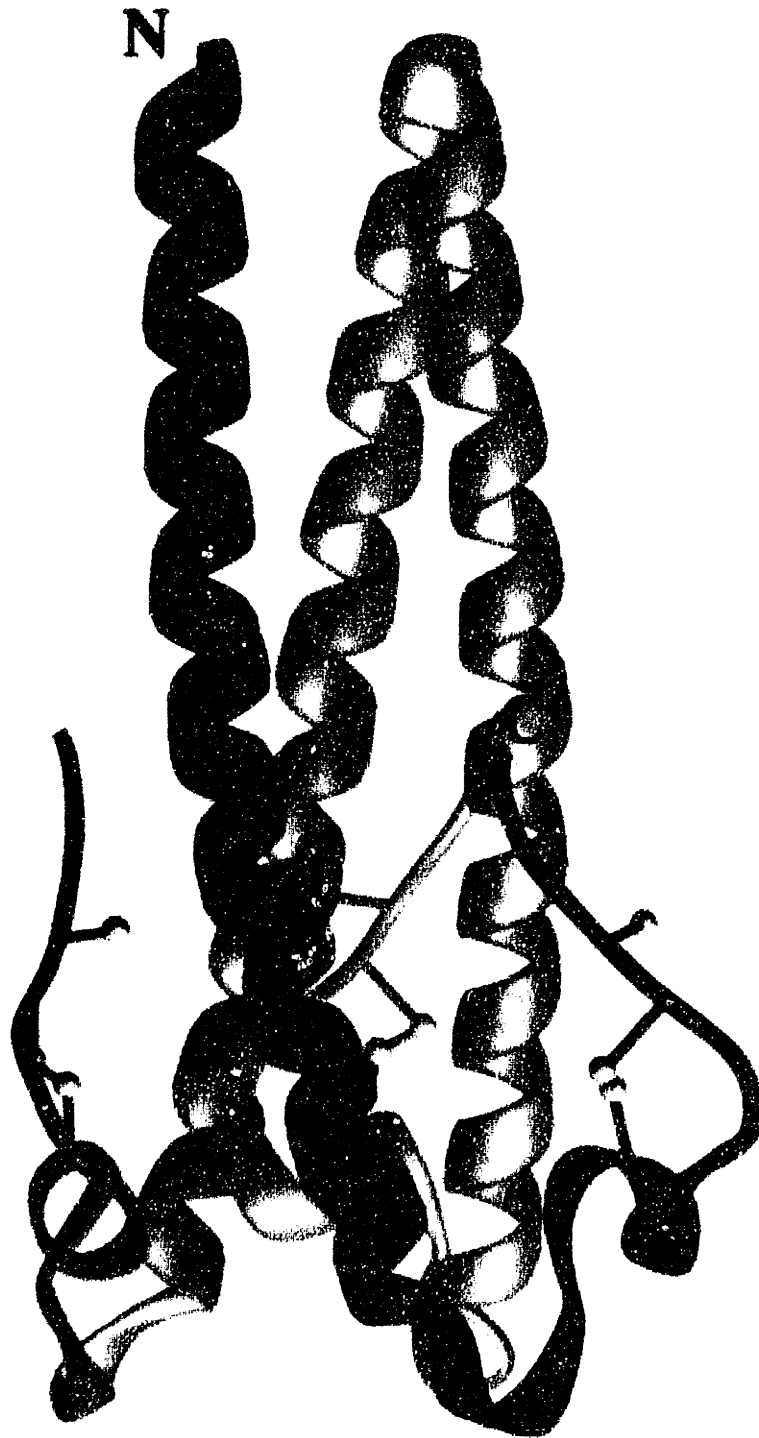


Fig. 2 a, A helical segment of the model is enveloped by a 1.2 s 2.4 Å MIR map. Residues in the core of the coiled coil are numbered. Only one subunit of the coiled coil is shown. b, Asn 71 and symmetry-related residues are viewed along the trimer axis. The distance between the amide nitrogen of Asn 71 and the chloride ion is 3.25 Å. The sum of the van der Waals radii for NH_4^+ (1.43 Å) and Cl^- (1.81 Å) is 3.24 Å²⁷.

Figure 2A

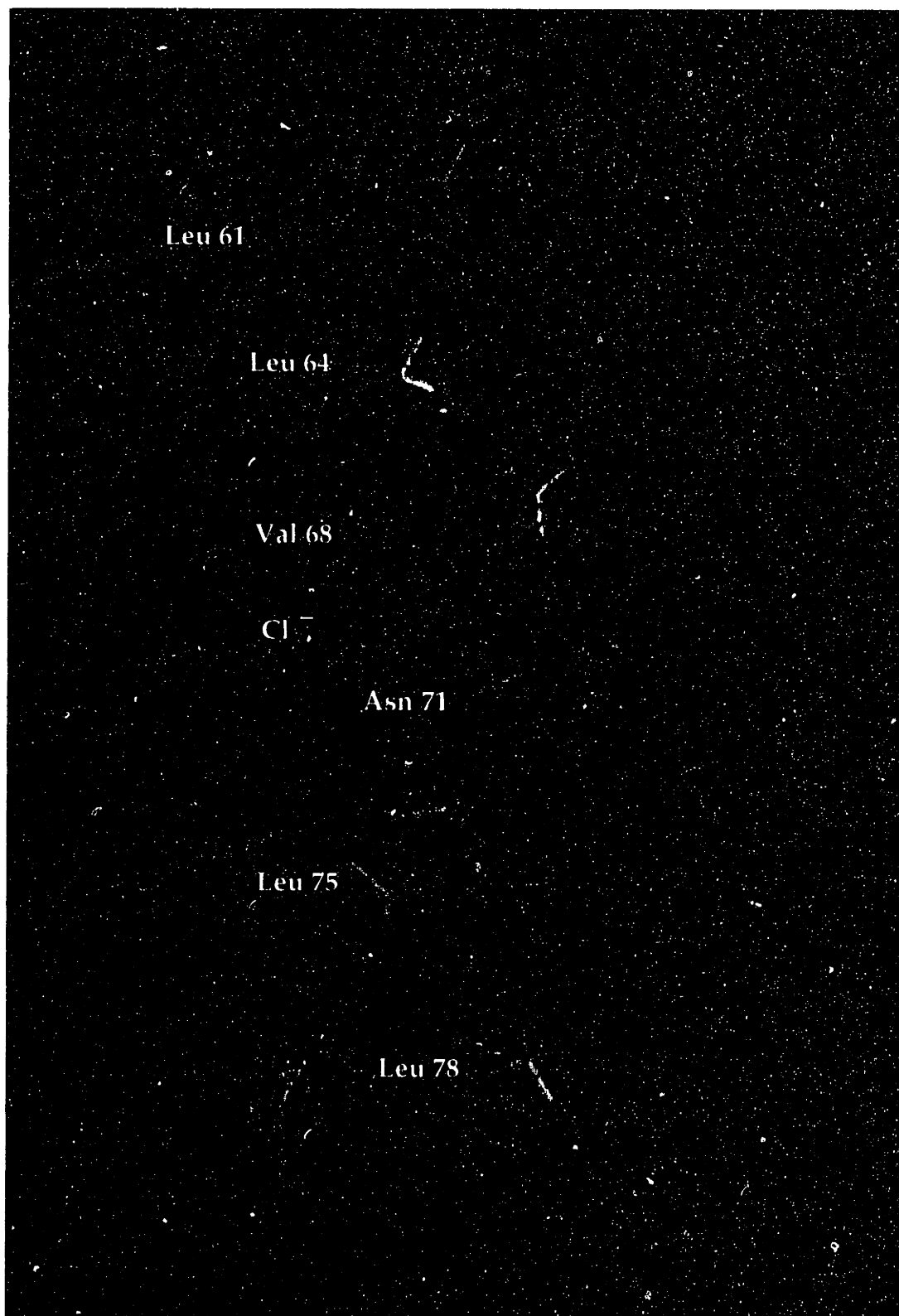


Figure 2B

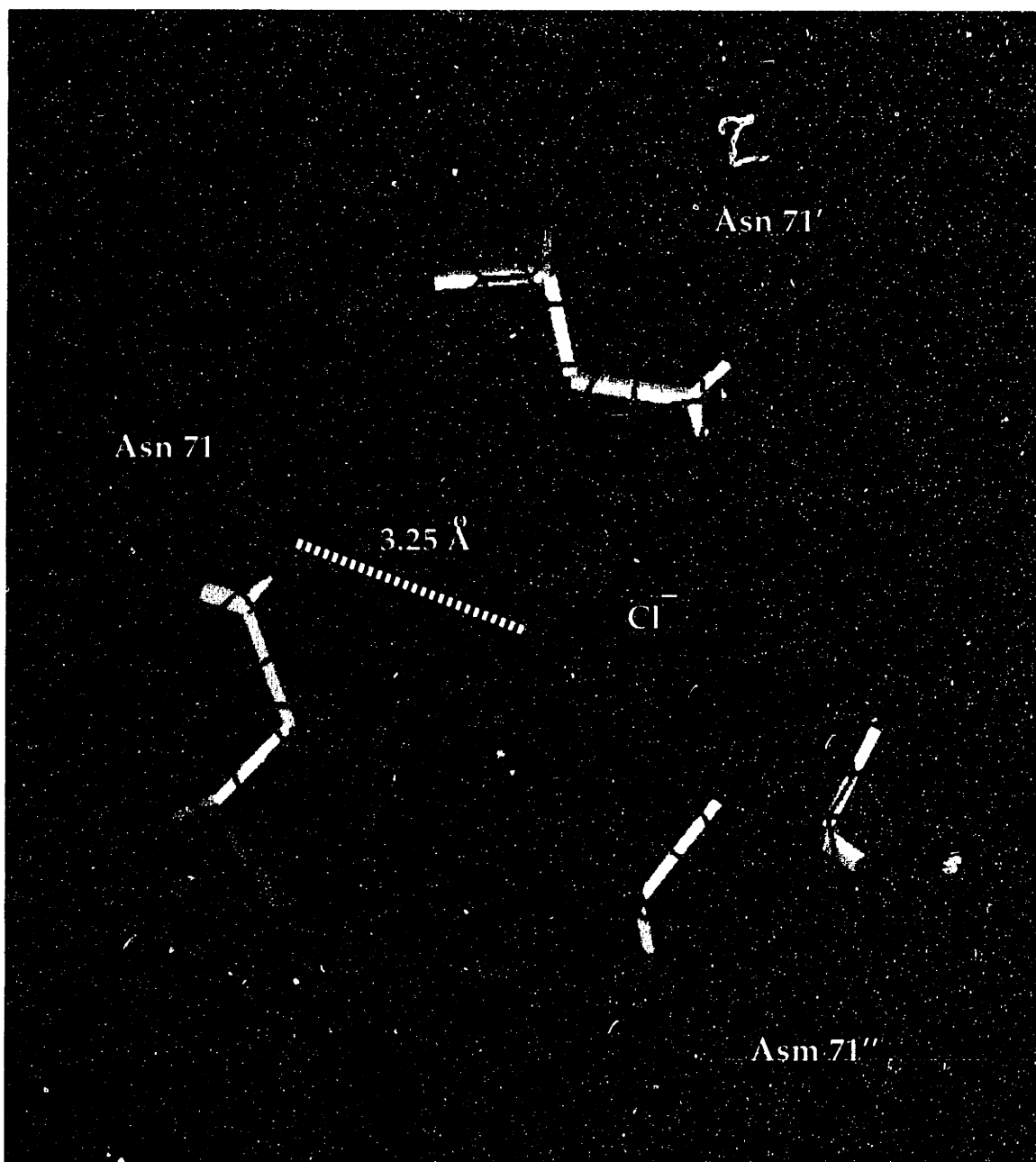


Fig. 3 Comparison of the Mo-55 and influenza TBHA₂¹⁵ structures. Linear maps drawn to approximate scale show the positions of Mo-55 and TBHA₂ (shaded) within TM and HA₂, respectively. The fusion-peptide (fp) and transmembrane (tm) segments are labeled. Both crystallographic models are lacking amino-terminal residues corresponding to the fusion-peptide region, as well as carboxy-terminal segments that form protease-sensitive links^{3,16} to the transmembrane regions. Below, the structures of Mo-55 and TBHA₂ are compared (see text and ref. 15) using MOLSCRIPT²⁸. TBHA₂ is the fragment generated¹⁶ by first cleaving native influenza HA from the viral membrane with the protease bromelain, and then treating this soluble HA with low pH. Low pH causes the receptor binding domains to dissociate from one another^{29,30} and the transmembrane polypeptide HA₂ to undergo a dramatic conformational change involving extension of a three-stranded coiled coil¹⁹. Finally, further proteolysis generates a crystallizable fragment by removal of the dissociated HA₁ domains and the fusion peptides. The TBHA₂ structure is thought to be the HA conformation that mediates membrane fusion¹⁵. White balls in the figure indicate sulfur atoms in cysteine side chains. A short strand from HA₁ is disulfide bonded to HA₂ in the TBHA₂ structure and is shown in dark gray. The apparent bend in the Mo-55 helix is an artifact of representing one subunit of a coiled coil in a two-dimensional diagram.

Figure 3

A

Moloney murine leukemia virus TM subunit



Influenza hemagglutinin HA2 subunit



B

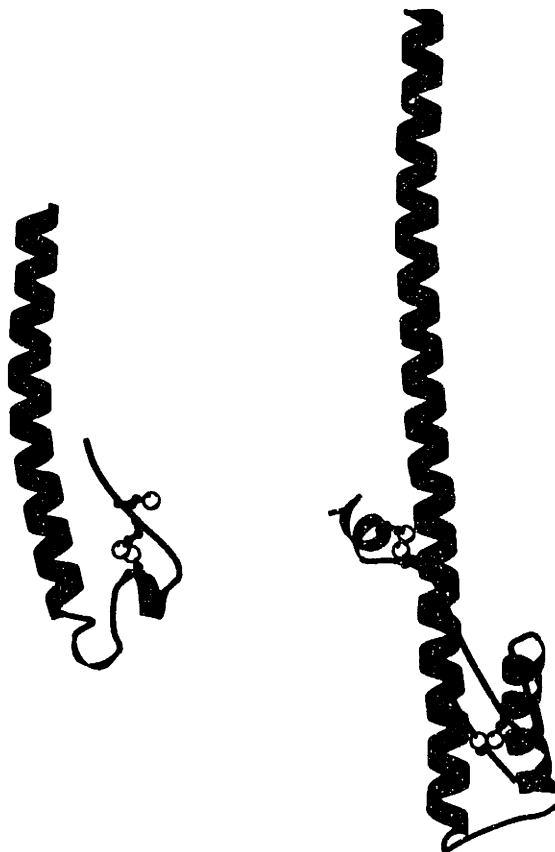


Table 1 Crystallographic and refinement statistics

Statistics are reported for crystals of a cysteine-to-alanine mutant at residue 94 in the TM sequence³.

^a $R_{\text{merge}} = \sum \sum_j |I_j - \langle I \rangle| / \sum I_j$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

^b $R_{\text{iso}} = \sum | |F_{\text{ph}}| - |F_{\text{p}}| | / \sum |F_{\text{p}}|$, where F_{ph} and F_{p} are the derivative and native structure factors, respectively.

^c $R_{\text{c}} = \sum | |F_{\text{ph}} \pm F_{\text{p}}| - |F_{\text{h,c}}| | / \sum |F_{\text{ph}} \pm F_{\text{p}}|$, where $F_{\text{h,c}}$ is the calculated heavy-atom structure factor.

^dPhasing power = $\langle F_{\text{h}} \rangle / E$, where $\langle F_{\text{h}} \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error.

^e $R_{\text{cryst, free}} = \sum | |F_{\text{obs}}| - |F_{\text{calc}}| | / \sum |F_{\text{obs}}|$, where the crystallographic and free R factors are calculated using the working and free reflection sets, respectively. The free reflections (about 7% of the total) were chosen before refinement of the initial model and were not used during refinement.

Table 1 Crystallographic and refinement statistics

<u>Parameter</u>	<u>Native</u>	<u>Se</u>	<u>Tm</u>
Resolution (Å)	20.0-1.7	20.0-2.4	20.0-2.4
Completeness (%)	96.0	97.0	97.3
R _{merge} ^a	0.061	0.064	0.054
R _{iso} ^b	--	0.109	0.245
Number of sites	--	1	1
R _c ^c	--	0.74	0.57
Phasing power ^d	--	1.23	2.13

Refinement statistics

Number of atoms (non-hydrogen)	413
Number of H ₂ O molecules	40
Number of reflections (working)	5030
R _{cryst} ^e	0.169
Number of reflections (free)	414
R _{free} ^e	0.234
Deviations from ideal geometry	0.017 Å (bonds) 2.39° (angles)

Table 2 Identification of buried polar interaction

	<u>Midpoint of Thermal Denaturation</u>	
	<u>Mo-55.3</u>	<u>Asn 71 → Ile mutant</u>
none	64 °C	76 °C
NaF (5 mM)	64 °C	--
NaCl (5 mM)	73 °C	76 °C
NaBr (5 mM)	73 °C	--
NaCl (20 mM)	75 °C	--

The temperature midpoint for thermal unfolding (T_m) of Mo-55, but not an asparagine-to-isoleucine mutant at position 71, is sensitive to the presence of chloride. Chloride and bromide stabilized the Mo-55 structure, but fluoride did not.

CHAPTER 4

CHARACTERIZATION OF THE LABILE COVALENT ASSOCIATION BETWEEN THE MURINE LEUKEMIA VIRUS ENVELOPE SUBUNITS

ABSTRACT

Earlier observations suggested that a labile intersubunit disulfide bond exists between the receptor-binding and membrane-fusion subunits of some retroviral envelope glycoprotein complexes (19). The conservation in oncoretroviruses of a thioredoxin-like CXXC amino acid motif, which is likely to be involved in this labile disulfide, implies a selective pressure to maintain this feature of the envelope glycoproteins. To further our understanding of the structural basis of retroviral envelope function, we sought to determine the conditions under which the disulfide bond is present in the Moloney murine leukemia virus (Mo-MLV) envelope. Our results indicate that the disulfide exists in fresh virions and is favored by the native glycoprotein structure. Conversely, disruption of the native glycoprotein structure by mild treatment with heat or denaturants allows a buried free thiol to initiate an irreversible thiol-disulfide rearrangement that eliminates the intersubunit disulfide bond. This rearrangement exhibits characteristics consistent with its involvement in the cell entry pathway of retroviruses.

INTRODUCTION

Enveloped viruses penetrate the cell surface in two successive steps, receptor binding and membrane fusion. Receptor binding is accomplished in retroviruses by the surface (SU) subunit of the viral envelope glycoprotein complex, while the transmembrane (TM) subunit mediates membrane fusion (for review see ref. 14). The SU and TM subunits are synthesized as a single-chain precursor, which is proteolytically cleaved to generate the two fragments.

The association of the receptor-binding and membrane-fusion subunits may be important in regulating cell entry by retroviruses. Crosslinking of the Rous sarcoma virus (RSV) envelope inhibits fusion but not receptor binding (7), consistent with a requirement for the separation of subunits upon initiation of membrane fusion. For some laboratory-adapted strains of HIV, binding to the receptor CD4 causes the HIV SU subunit, gp120, to dissociate or "shed" from the viral surface (12,18), and can increase viral infectivity (22), suggesting that

receptor-induced changes in subunit association are involved in retroviral entry. The shedding of SU from retroviral virions may parallel the dissociation of the influenza receptor-binding subunits during the conformational change to the fusion-active state of this well-characterized virus (for review see ref 24).

Interestingly, a labile intersubunit disulfide bond has been detected in the SU/TM complex of some retroviruses, implying the existence of a mechanism to modulate the nature of the SU/TM association. The SU and TM envelope proteins of murine leukemia viruses (MLVs) migrate independently by SDS-PAGE, even in the absence of exogenous reducing agent, whereas in virus treated with N-ethylmaleimide (NEM), SU and TM migrate as a covalent complex (19). Some confusion has resulted from a misinterpretation of the original observation, in which NEM was proposed to activate formation of a disulfide bond that was absent in the native virion (19). The phenomenon could be caused instead by the presence of an endogenous reducing agent, such as a free thiol in the viral envelope, which is blocked by the addition of the thio-alkylating agent NEM.

Nearly all oncoretroviruses, including the human T-cell lymphotropic and mammalian C-type retroviruses, share an intriguing sequence motif that suggests that the labile disulfide bond may be a common feature and that the nature of the SU/TM association may indeed be important for retroviral envelope function. The SU subunits of these viruses contain the sequence CXXC, which is reminiscent of the active site of thiol-redox enzymes such as thioredoxin (13) or protein-disulfide isomerase (6). Since the CXXC pattern of cysteine residues is so highly conserved among oncoretroviruses, and since mutagenesis of either of these cysteines yields non-infectious virus (11), these residues are likely to play an important role in the function of the viral envelope. It is possible that this motif is involved in the thiol-disulfide rearrangement that eliminates the intersubunit disulfide bond from these viruses. Although the CXXC cysteines are disulfide bonded to one another in SU that has been released from virions (16), it has recently been proposed that one of the CXXC cysteines is involved in the intersubunit disulfide bond of the SU/TM complex (21).

Since the association of subunits may be linked to the mechanism of cell entry, studies on the mechanism of membrane fusion by oncoretroviruses would benefit from a rigorous analysis of the properties of the labile disulfide bond and the free thiol in the SU/TM complex. We sought to determine the conditions that favor preservation of the intersubunit disulfide, and to verify that these conditions are consistent with the involvement of thiol-disulfide rearrangement in retroviral entry. Our studies led not only to a measure of the stability of the SU/TM covalent association in Moloney murine leukemia virus (Mo-MLV) and to identification of the factors affecting this association, but also to unexpected discoveries regarding the nature of the free thiol in the SU/TM complex.

MATERIALS AND METHODS

Virus production. CL-1 cells (gift of J. Cunningham) were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco BRL) supplemented with 10% calf serum. For Mo-MLV virus production, cells were plated at 50% confluence in DMEM with 1% calf serum and incubated at 32°C for 36 to 48 hrs. Culture supernatant was collected and passed through a 0.8 μ M filter.

Virus purification. Virus was purified using either potassium citrate or sucrose gradients. Sixty percent w/v potassium citrate, pH 7.2, or 60% sucrose in PBS, was diluted with PBS to obtain citrate or sucrose solutions in 10% increments. Density gradients were created by layering 0.8 mL each of 60% to 10% citrate or sucrose. Approximately 7 mL filtered supernatant was layered on top of each gradient. Gradients were centrifuged for 1 hr at 20,000 rpm and 4°C in a Beckmann L8-M ultracentrifuge using a SW41 rotor. The opaque virus band was removed in approximately 0.5 mL with a needle and syringe through the side of each polyallomer centrifuge tube (Beckman), and desalted into PBS or other buffer on a NAP 10 column (Pharmacia Biotech). Approximately 4.5 mL pooled desalted virus was then concentrated at 4°C to ~0.3 mL in a centricon-100 (Amicon).

Thermally-induced disulfide rearrangement. 12 μ L aliquots of concentrated virus were incubated at the indicated temperatures for 15 minutes. Aliquots were then placed at room temperature, and 1.3 μ L of 3% w/v N-ethylmaleimide (NEM) in 10% v/v acetonitrile was added. After 30 min, 6 μ L 50 mM Tris, pH 6.8, 2% w/v SDS, 0.1% w/v bromphenol blue, 20% v/v glycerol (gel-load buffer) was added and proteins were separated by gradient sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels (5 to 20% acrylamide). Protein bands were transferred to Hybond-C nitrocellulose (Amersham). Western blots were performed by shaking gently for one hr each in the following solutions: 1) PBS containing 5% w/v nonfat milk powder and 0.1% w/v Tween-20; 2) 1:2000 dilution of polyclonal antisera (HRP, Inc.) to the TM ectodomain peptide Mo-92 (4) in PBS containing 0.1% Tween-20; 3) 1:5000 dilution horseradish peroxidase-linked goat anti-rabbit secondary antibody (Amersham) in PBS containing 0.1% Tween-20. One 15 min and two 5 min

washes were performed between each step with PBS containing 0.1% Tween-20. ECL (Amersham) western blotting detection reagents were used to visualize immunoreactive bands.

The SU/TM band was identified by size and by reactivity with both α -SU (data not shown) and α -TM peptide antisera. The capsid band was identified by size. On Coomassie-stained gels, four bands were visible with molecular weights under 20 kD (data not shown). These four bands were distinguished by amino-acid sequencing of their N-termini. The top band gave the sequence Glu-Pro-Val, and was therefore identified as the p15E TM protein. The second band, in spite of being the strongest on Coomassie-stained gels, did not yield amino acid sequence, and was perhaps blocked at its amino-terminus. This observation is consistent with the band being the p15 matrix protein arising from the amino-terminus of the gag polyprotein. The third and fourth bands (not detected in the Western blot), gave the sequences Pro-Ala-X-Thr-Pro and Ala-Thr-Val-Val-Ser, consistent with the amino-terminal sequences of the p12 protein and p10 nucleocapsid, respectively. The capsid protein consistently shows a low level of cross-reactivity with the α -TM peptide antisera, while the matrix protein is severely cross-reactive. Affinity purification of the α -TM peptide antibodies does not eliminate this problem (data not shown).

pH and urea dependence. For pH-dependence measurements, sucrose gradients were run in the presence of 200 mM NaCl/25 mM NaPO₄ pH 5.0, 6.0, or 7.0 instead of PBS and desalted into the same buffers. Incubations, NEM blocking, SDS-PAGE, and western blotting proceeded as above.

For urea-dependence measurements, virus was purified in PBS. 6 μ L aliquots of concentrated virus was mixed on ice with 6 μ L urea dissolved in PBS at twice the final desired urea concentration. Incubations, NEM blocking, SDS-PAGE and western blotting proceeded as above.

NEM time and concentration courses. NEM from a 2% w/v stock in 10% acetonitrile was added to 10 μ L virus aliquots to a final concentration of 0.2%, and incubations allowed to proceed at room temperature for the indicated lengths of time before addition of gel-load buffer.

For NEM concentration-dependence measurements, a 10% NEM stock was made in acetonitrile. This stock was diluted to various concentrations with water, and 4 μL of the dilutions were added to 16 μL of gel-load buffer. 10 μL of these mixtures was added to 10 μL aliquots of virus, to achieve the final NEM concentrations indicated. Note that the final concentrations of NEM are and should be determined relative to the total volume, including gel-load buffer, and not to the volume prior to addition of denaturing detergent.

The rate of loss of the SU/TM covalent complex after addition of denaturant was determined by rapidly stirring 20 μL virus aliquots in Reacti-Vials (Pierce), adding 16 μL gel-load buffer, and, after the indicated delay, adding 4 μL of a 3% NEM stock in 10% acetonitrile to achieve a final NEM concentration of 0.3%.

RESULTS

Native protein structure inhibits thiol-disulfide rearrangement. To measure the thermal stability of the covalent SU/TM complex, purified Mo-MLV virus was aliquoted and incubated for 15 minutes at various temperatures. We presumed that thiol-disulfide rearrangement would be an irreversible event, preventing the acquisition of meaningful equilibrium thermodynamic data, so the fifteen minute incubation was chosen to facilitate comparison to the thermally induced irreversible conformational change of influenza hemagglutinin (HA) (3). After the 15 minute incubation, the virus samples were returned to room temperature and NEM was added to preserve the remaining SU/TM covalent complexes. Gel load buffer containing SDS was added to disrupt virions and denature viral proteins, which were then separated by gradient SDS-PAGE. Western blots were probed with polyclonal antisera against a 92-amino acid fragment of the Mo-MLV TM protein (4). The only band detected in the higher molecular weight region of the gel (~90 kD) represents a covalent complex between SU and TM (Fig. 1).

As judged by the loss of the ~90 kD band after incubation at higher temperatures, the SU/TM complex cooperatively undergoes thiol-disulfide rearrangement with a midpoint of $48 \pm 1^\circ\text{C}$ in PBS. To determine the role of glycoprotein structure in the thiol-disulfide rearrangement, we examined the effect of protein denaturants on the temperature midpoint of thiol-disulfide rearrangement. The temperature midpoint for the transition decreased with increasing concentrations of urea present during the incubation (Fig. 2A). Since urea has no effect on inherent thiol reactivity (15), its dramatic effect on the Mo-MLV glycoprotein thiol-disulfide rearrangement must be due to its role as a protein denaturant.

Loss of covalent complex is pH-independent. Heat, denaturants, and low pH can all induce a conformational change in influenza HA (3), which naturally infects via a low pH mediated pathway (24). In all cases, the influenza HA conformational change correlates with conditions that promote membrane fusion (3). Therefore, we inquired whether loss of the SU/TM covalent association in Mo-MLV is consistent with the known pH profile of retroviral infectivity. We examined thiol-disulfide exchange in Mo-MLV as a function of

pH; virus at pHs relevant to cellular endosomal compartments all exhibited the same temperature midpoint for thiol-disulfide rearrangement, $48 \pm 1^\circ\text{C}$ (Fig. 2B), consistent with the pH independence of retroviral cell entry.

All SU and TM may initially be in covalent complex. The above experiments do not depend on an accurate measurement of the amount of SU/TM initially present in covalent complex, or, within limits, on the effectiveness of the thiol-blocking agent. Assuming that a constant fraction of all complexes present are preserved by blocking with NEM, the ratio of ~90 kD band in test and control samples will always provide a measure of the resistance of the covalent complex to the test conditions. However, there lingers an independent question regarding the absolute fraction of SU and TM subunits in covalent complex on the surface of native virions, a measurement that is highly dependent on complete blocking by NEM of all SU/TM covalent complexes. Therefore we performed a titration of NEM to determine a maximally-effective concentration of alkylating agent (Fig. 3).

A typical concentration of NEM used in previous experiments is 0.2% prior to the addition of denaturing gel-load buffers (19). However, we observe that complete blocking does not occur until near or above a final concentration of 0.5% NEM (Fig. 3). Concentrations of NEM above 0.5% become problematic when working in aqueous solution. However, we observe that, to the limits of our detection, the amount of free TM protein approaches zero when high concentrations of NEM are used (Fig. 3). It is appropriate to monitor the free TM band (~20 kD) to determine the extent of covalent complex, since TM is fixed in the viral membrane and cannot be lost during purification. The ~90 kD and ~20 kD bands therefore represent all the TM protein initially present in the native virion.

Free thiol inaccessible in native SU/TM. Since NEM is a very reactive thiol-modifying agent (9), the extremely high concentrations required to block the Mo-MLV envelope are surprising. To investigate this anomaly, we performed a time-course of incubation with NEM prior to denaturation (Fig. 4). The amount of blocked SU/TM complex was found to be independent of the length of incubation. Even if blocking was initiated simultaneously with denaturation, by including the NEM in the gel-load buffer, blocking was as

effective as when NEM was added prior to denaturant (data not shown). Since the length of incubation in the presence of NEM cannot compensate for lower NEM concentrations (data not shown), these results indicate that NEM modifies the free thiol in SU/TM only upon or during denaturation of the complex.

We also studied the effect of NEM addition subsequent to denaturation. The time required for the SU/TM complex to denature and for thiol-disulfide rearrangement to occur was examined by adding gel-load buffer into rapidly-mixed virus aliquots and adding NEM at short time points later. In approximately 5 seconds, half the covalent complexes had been disrupted (Fig. 5). The rate of loss of the SU/TM covalent complex after addition of SDS could reflect the rate of denaturation, with thiol-disulfide rearrangement proceeding rapidly after denaturation. Alternatively, denaturation could occur rapidly, with thiol-disulfide exchange then proceeding on the order of 5 seconds. The first possibility is the more likely, as rate constants for the reaction of free cysteine with N-ethylmaleimide at neutral pH have been estimated to be greater than $1000 \text{ M}^{-1} \text{ s}^{-1}$ (9).

DISCUSSION

In the most thoroughly-studied enveloped virus entry event, that of the orthomyxovirus influenza, the association of the receptor-binding (HA1) subunits with the membrane-fusion (HA2) subunits in the glycoprotein trimer plays an important role in the timing of the two major steps in viral entry. Upon exposure to low pH, which initiates the membrane-fusion event for influenza, the HA1 subunits dissociate from one another. HA1 remains tethered to HA2 by a disulfide bond, but the majority of the contacts between HA1 and HA2 appear to be broken at this stage (10). The crystal structure of low-pH converted HA (1) shows that regions of HA2 held in a loop conformation under the HA1 "clamp" in the native HA structure (25) spring up to form a coiled coil once HA1 is released. This conformational change is thought to catapult the hydrophobic and glycine-rich "fusion-peptide" region at the amino terminus of HA2 toward the target cell membrane (2).

Although it remains to be determined whether other virus envelope proteins, including those of retroviruses, are also "spring-loaded," the observations that the SU/TM complex of retroviruses readily disassembles, and that this disassembly may be involved in viral entry, are provocative. It was previously determined that SDS treatment can result in loss of the covalent association between SU and TM in murine leukemia viruses (19), but it was not clear if milder treatments could promote this phenomenon. This question has direct bearing on the formulation of hypotheses for the mechanism of conversion from the receptor-binding to the membrane-fusion competent state of the retroviral envelope.

Our results show that short exposure to mild heat treatment, approximately 10°C above physiological temperature, causes the cooperative loss of the SU/TM intersubunit disulfide bond. In addition, low levels of urea allow thiol-disulfide exchange to occur readily at physiological temperatures. Although neither heat nor urea treatment are directly relevant to the mechanism of viral entry, these studies indicate that only a low energy barrier exists for conversion from the native state of the viral envelope in which SU and TM are covalently linked. Overcoming this energy barrier during viral entry may be accomplished instead by the binding energy from contact with receptor and may result in formation of the fusion-competent conformation of the viral envelope.

In support of the hypothesis that a low energy barrier must be overcome for viral-mediated membrane fusion to begin, it has been shown for influenza HA that any treatment that destabilizes the native state, such as heat, denaturants, or the physiologically relevant low pH, results in conversion to a common HA2 structure (3); in each case the structural change coincides with acquisition of the capacity to fuse membranes (3). By analogy, destabilization of the native retroviral SU/TM covalent complex, by whatever means, may promote the fusion-active conformation of the retrovirus envelope. In fact, a post-binding temperature-dependent event has been detected for the avian oncoretrovirus RSV; temperatures above 22°C are required for RSV fusion, but not binding (8). This observation can be interpreted as indicating that receptor binding cannot overcome the increased activation energy for conversion to the fusogenic state of RSV when the thermal energy of the native state is decreased.

It is noteworthy that, while the SU/TM covalent complex of Mo-MLV is disrupted at relatively low temperatures, the TM subunit core is highly thermostable (4) and structured (5) in isolation. Studies on soluble TM ectodomain peptides demonstrated that TM denatures only at temperatures greater than 86°C (4). Therefore, the structure observed in the TM peptides would persist through, or perhaps even form during, the structural transition that dissociates SU and TM. This situation is similar to that of influenza in which, at neutral pH, the conformational change to the fusion-active state occurs at 63°C, but a second transition, presumably the denaturation of the HA2 subunit, occurs at higher temperatures (2, 20).

Consistent with the Mo-MLV thiol-disulfide exchange being relevant to the retroviral entry pathway is the observation that the thermal stability of the SU/TM covalent complex is independent of pH in the pH range of endosomal compartments. In contrast, the thermal stability of the native conformation of influenza HA varies dramatically with pH in this range (3), as would be expected for a virus with a low-pH dependent membrane-fusion event. Pre-treatment of influenza with low pH in the absence of target membranes inactivates the fusion apparatus (23). Retroviruses are not inactivated by low pH pre-treatment (17), in keeping with our observation that structural changes in the Mo-MLV envelope resulting in thiol-disulfide rearrangement are not promoted by low pH.

Although the pH profile of thiol-disulfide rearrangement is consistent with the known profile of retroviral entry, it remains to be determined whether thiol-disulfide rearrangement is essential for, or occurs during viral entry. Mutants of the CXXC cysteine residues produce non-infectious virus, but these mutants also have defects in folding and transport of their envelope glycoproteins (11). To determine whether thiol-disulfide rearrangement is truly important for viral entry, one would ideally destroy the free thiol after it has contributed to folding and transport, and its only remaining role could be in conversion from the receptor-binding to the membrane-fusion competent state of the viral envelope.

The obvious experiment to address the importance of thiol-disulfide rearrangement in oncoretroviral entry is to block the free thiol on mature

virions, remove the excess blocking reagent, and record the infectivity of the modified virions. Our results show that this experiment is not feasible, however, since the free thiol cannot be modified by NEM, a potent thiol-specific agent, while the envelope is in the native folded state. When prolonged incubations in high concentrations of NEM were performed and NEM was then removed by gel filtration, no SU/TM covalent complex was detected after SDS-PAGE (data not shown). Similarly, half-hour incubations with 0.2% NEM were less effective than simultaneous addition of 0.5% NEM with denaturant. While the apparent lack of reactivity of the free thiol in the folded state is discouraging toward a simple and direct determination of the significance of thiol-disulfide rearrangement in viral entry, it serves as an important warning for the interpretation of future data. Any treatments that appear to block the free thiol should be thoroughly examined for deleterious effects on the native protein structure that could have led to exposure of the buried free thiol.

In addition, we conclude from our experiments that previous measurements of the ratio of disulfide bonded to non-bonded SU and TM on the surface of virions depended more on the method and NEM concentration used for blocking the covalent complexes than on the inherent ratio of these two species. We find no evidence that the native glycoprotein structure involves other than disulfide-bonded complexes. However, the buried free thiol, the relative instability of the SU/TM covalent complex, and the irreversibility of thiol-disulfide rearrangement could easily contribute to detection of non-disulfide bonded species.

We have demonstrated that the native structure of the SU/TM complex on the surface of Mo-MLV contains an intersubunit disulfide bond, and that mild perturbations of the environment can cause irreversible loss of this intersubunit disulfide. The methods presented here provide a simple and quantitative means of analyzing the effects of various agents on the native structure of the Mo-MLV envelope, and can theoretically be extended to examine the effect of receptor binding on the SU/TM covalent association. Although the precise details may differ, it is likely that similar principles govern cell entry by all retroviruses, and that conclusions from oncoretroviruses regarding the role of the SU/TM association in regulating the membrane-fusion active state of the envelope may be more generally applicable. However, the labile disulfide bond

that exists in Mo-MLV provides a convenient probe for the state of the viral envelope that other retroviruses, such as HIV, lack. The facility with which thiol-disulfide rearrangement is induced and the pH independence of rearrangement provide support that the extensive conservation of the CXXC motif implies a physiological role for disruption of the SU/TM covalent association.

Acknowledgements

We are grateful to Robert Davey and James Cunningham for virus-producing cells and helpful advice on virus production and purification.

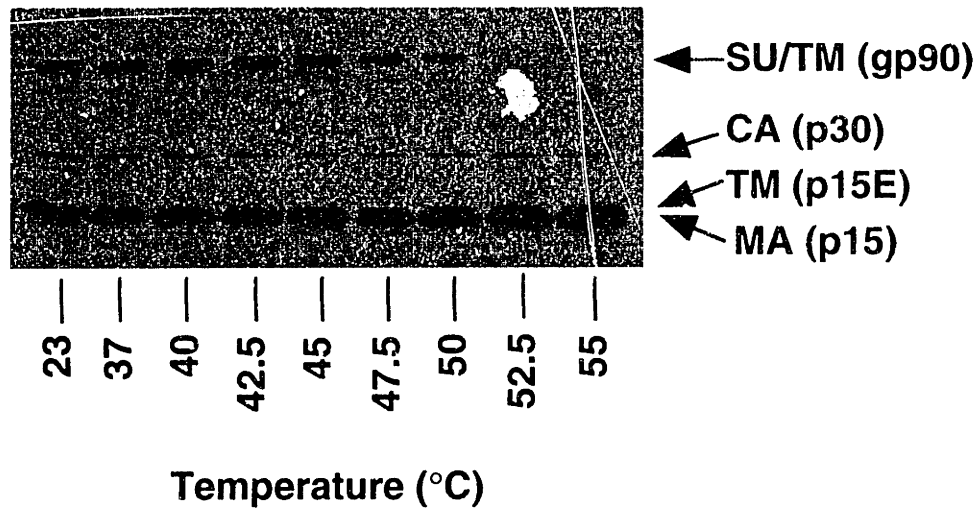
References

1. **Bullough, P.A., F.M. Hughson, J.J. Skehel, and D.C. Wiley.** 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**:37-43.
2. **Carr, C.M., and P.S. Kim.** 1993. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**:823-832.
3. **Carr, C.M., C. Chaudhry, M. Milhollen, and P.S. Kim.** in preparation.
4. **Fass, D., and P.S. Kim.** 1995. Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin. *Curr. Biol.* **5**:1377-1383.
5. **Fass, D., S.C. Harrison, and P.S. Kim.** 1996. Retrovirus envelope domain at 1.7 Å resolution. *Nature Struc. Biol.* **3**:465-469.
6. **Freedman, R.B.** 1984. Native disulfide bond formation in protein biosynthesis: evidence for the role of protein disulphide isomerase. *Trends Biochem. Sci.* **9**:438-441.
7. **Gilbert, J.M., D. Mason, and J.M. White.** 1990. Fusion of Rous sarcoma virus with host cells does not require exposure to low pH. *J. Virol.* **64**:5105-5113.
8. **Gilbert, J.M., L.D. Hernandez, J.W. Balliet, P. Bates, and J.M. White.** 1995. Receptor-induced conformational changes in subgroup A avian leukosis and sarcoma virus envelope glycoprotein. *J. Virol.* **69**:7410-7415.
9. **Gorin, G., P.A. Martic, and G. Doughty.** 1966. Kinetics of the reaction of N-ethylmaleimide with cysteine and some congeners. *Arch. Biochem. Biophys.* **115**:593-597.
10. **Graves, P.N., J.L. Schulman, J.F. Young, and P. Palese.** 1983. Preparation of influenza virus subviral particles lacking the HA1 subunit of hemagglutinin: unmasking of cross-reactive HA2 determinants. *Virology* **126**:106-116.
11. **Gu, J., S. Parthasarathi, A. Varela-Echavarria, Y. Ron, and J. P. Dougherty.** 1995. Mutations of conserved cysteine residues in the CWLC motif of the oncoretrovirus SU protein affect maturation and translocation. *Virology* **206**:885-893.
12. **Hart, T.K., R. Kirsch, H. Ellens, R.W. Sweet, D.M. Lambert, S.R. Petteway, Jr., J. Learly, and P.J. Bugelski.** 1991. Binding of soluble CD4 proteins to human immunodeficiency virus type 1 and infected cells induces release of envelope glycoprotein gp120. *Proc. Natl. Acad. Sci. USA* **88**:2189-2193.

13. **Holmgren, A.** 1989. Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* **264**:13963-13966.
14. **Hunter E., and R. Swanstrom.** 1990. Retrovirus envelope glycoproteins. *Cur. Top. in Microbiol. and Immun.* **157**:187-253.
15. **Lin, T.Y. and Kim, P.S.** 1989. Urea dependence of thiol-disulfide equilibrium in thioredoxin: conformation of the linkage relationship and a sensitive assay for structure. *Biochemistry* **28**:5282-5287.
16. **Linder, M., D. Linder, J. Hahnen, H.-H. Schott, and S. Stirm.** 1992. Localization of the intrachain disulfide bonds of the envelope glycoprotein 71 from Friend murine leukemia virus. *Eur. J. Biochem.* **203**:65-73.
17. **McClure, M.O., M.A. Sommerfelt, M. Marsh, and R.A. Weiss.** 1990. The pH independence of mammalian retrovirus infection. *J. Gen. Virol.* **71**:767-773.
18. **Moore, J.P., J.A. McKeating, R.A. Weiss, and Q.J. Sattentau.** 1990. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. *Science* **250**:1139-1142.
19. **Pinter, A., J. Leiman-Hurwitz, and E. Fleissner.** 1978. The nature of the association between the murine leukemia virus envelope proteins. *Virology* **91**:345-351.
20. **Ruigrok, R.W.H., S.R. Martin, S.A. Wharton, J.J. Skehel, P.M. Bayley, and D.C. Wiley.** 1986. Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. *Virology* **155**:484-497.
21. **Sanders, D., Z. Li, S. Kayman, R. Kopelman, and A. Pinter.** 1995. Characterization of a labile SU-TM disulfide bond that is involved in processing of MuLV envelop protein. *Retroviruses Meeting, Cold Spring Harbor, N.Y.*
22. **Sattentau, Q.J., J.P. Moore, F. Vignaux, F. Traincard, and P. Poignard.** 1993. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *J. Virol.* **67**:7383-7393.
23. **Skehel, J.J., P. Bayley, E. Brown, S. Martin, M. Waterfield, J. White, I. Wilson, and D. Wiley.** 1982. Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc. Natl. Acad. Sci. USA* **79**:968-972.
24. **Wiley, D.C., and J.J. Skehel.** 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Ann. Rev. Biochem.* **56**:365-394.

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

Figure 1



Thermal denaturation of the Mo-MLV SU/TM covalent complex. Each lane contains virus incubated for 15 minutes at the indicated temperatures, and then treated with 0.3%NEM prior to the addition of gel-load buffer. The following virus bands are visible in the blot and are labeled: envelope glycoprotein covalent complex (SU/TM), capsid (CA), free TM protein (TM), and matrix (MA). The midpoint for the thermal transition involving loss of the SU/TM covalent association is $48 \pm 1^\circ\text{C}$.

FIG. 2. Denaturant- and pH-dependence of the thermal transition. (A) The SU/TM covalent complex (gp90) band is shown. Virus was incubated at various temperatures in the presence of increasing concentrations of urea. The midpoint of the thermal transition is highly sensitive to the concentration of urea present. Approximately 1 M urea causes loss of the intersubunit disulfide bond to occur at physiological temperatures over the course of the 15 minute incubation. (B) Virus was incubated at various temperatures in buffers of decreasing pH. The midpoint of the thermal transition shows no dependence on pH in this range.

Figure 2

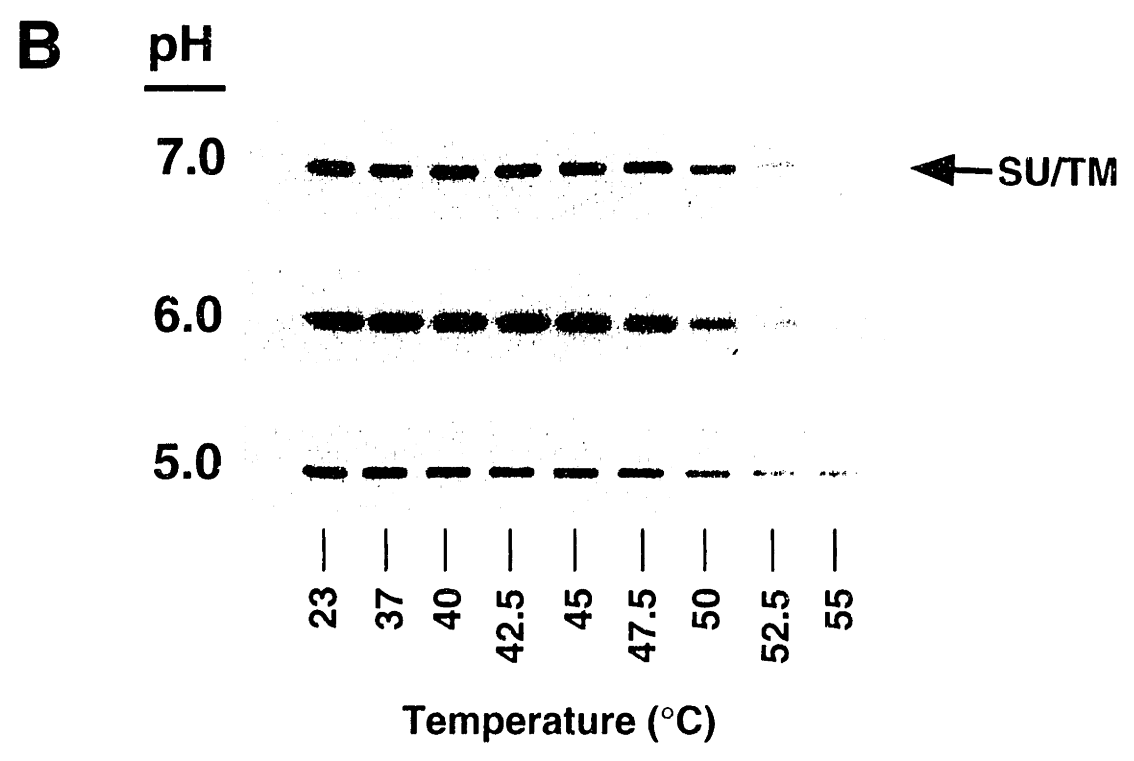
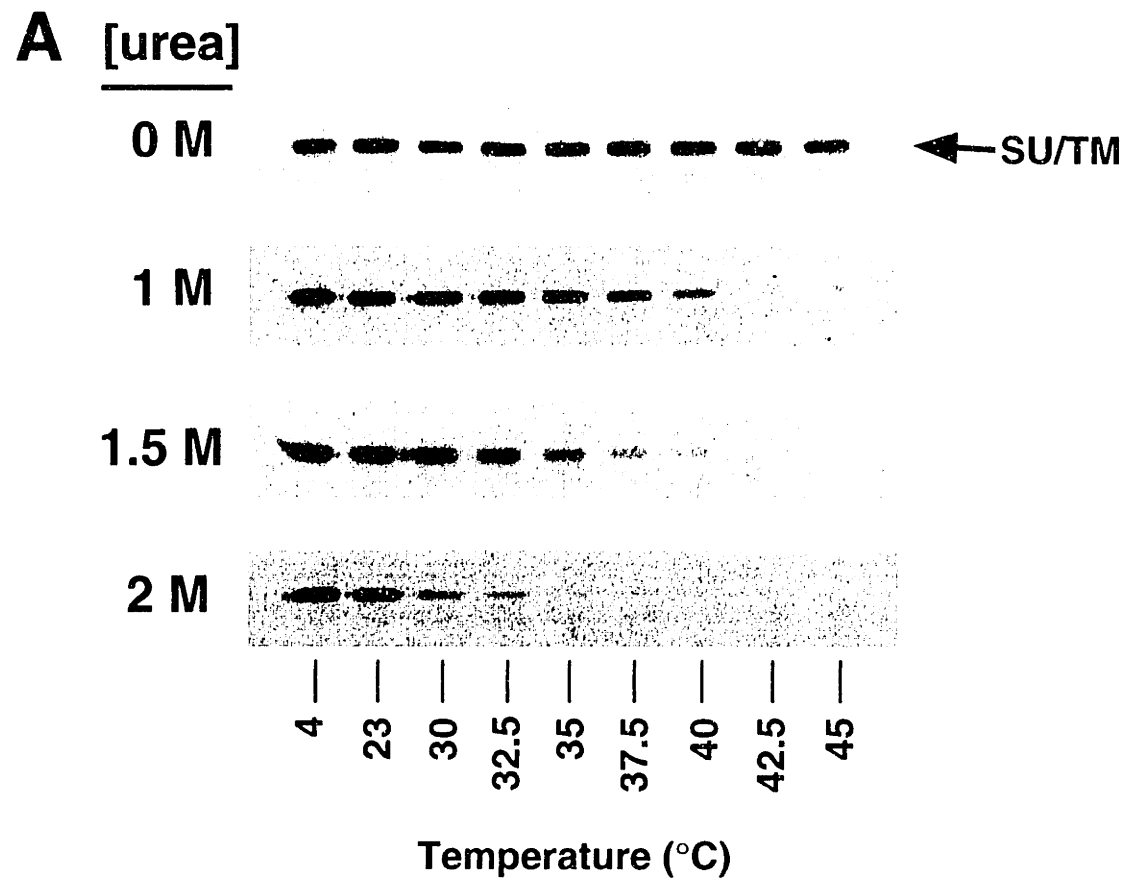
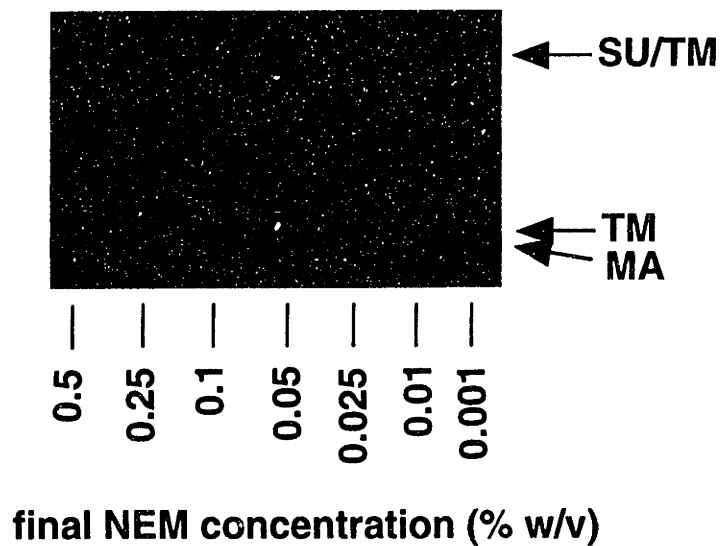
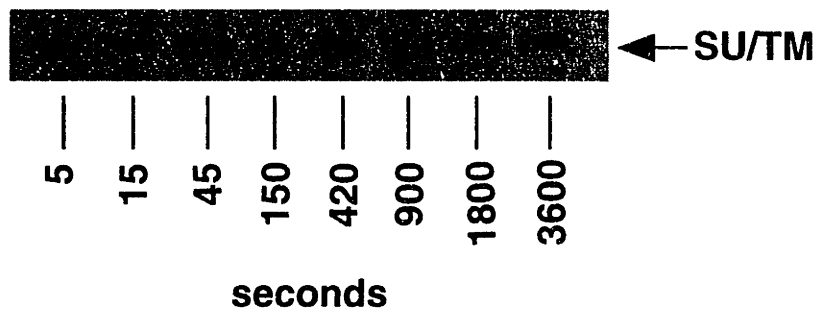


Figure 3



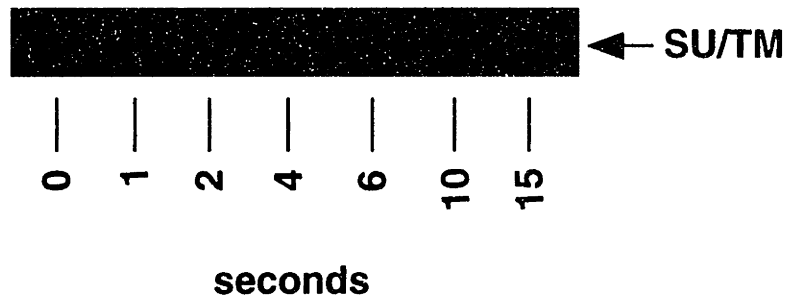
Concentration dependence of NEM in the thiol-blocking reaction. NEM was added to the final concentrations indicated. The amount of SU/TM covalent complex preserved by NEM treatment is highly sensitive to the concentration of NEM used. The intensity of the free TM band approaches zero with the highest concentration of NEM.

Figure 4



Time course of NEM addition prior to denaturation. NEM was added to 0.2% w/v at various times prior to addition of SDS-PAGE load buffer. The amount of SU/TM complex detected is independent of the time of incubation with NEM.

Figure 5



Rate of thiol-disulfide exchange. SDS-PAGE load buffer was added at the indicated times prior to addition of NEM. Approximately 5 seconds are required for half the SU/TM covalent complexes to rearrange.

CHAPTER 5

STRUCTURE OF A RETROVIRUS RECEPTOR-BINDING GLYCOPROTEIN: MURINE LEUKEMIA VIRUS SU DOMAIN AT 2.0 Å RESOLUTION

An essential step in retrovirus infection is the binding of the virus to its receptor on a target cell. The structure of the receptor-binding domain of the envelope glycoprotein from Friend murine leukemia virus, which is functionally analogous to HIV gp120, was determined to 2.0 Å resolution by X-ray crystallography. The core of the domain is an anti-parallel β -sandwich, with two interstrand loops forming a helical subdomain atop the sandwich. The residues in the helical region, but not in the β -sandwich, are highly variable among mammalian type-C retroviruses with distinct tropisms, indicating that the helical subdomain determines the receptor specificity of the virus. This structure provides the first high-resolution insight into a retrovirus receptor-binding glycoprotein, and will serve as a valuable tool for the development of targeted gene therapy vectors.

Retroviruses are simultaneously a profound human medical problem and a potential medical solution. They can be pathogenic, causing immunodeficiency, leukemia, and neurological disease, but they are also actively studied for their proposed utility as gene-therapy vectors. Essential to both roles is the targeting of virus to host cell through interactions between viral-envelope and cell-surface proteins.

Retrovirus envelope glycoproteins (reviewed in 1) are synthesized as single-chain precursors that are subsequently cleaved into two subunits, SU (surface) and TM (transmembrane) (Fig. 1). The SU glycoprotein binds receptor. The TM subunit contains the hydrophobic fusion-peptide and transmembrane segments, and is likely to participate directly in fusion of the viral and cellular membranes following receptor binding.

Efforts to understand the structural basis of retroviral binding and entry and to develop retrovirus-based gene therapy vectors (reviewed in 2) with engineered tropisms have been hindered by the lack of high-resolution models for the receptor-binding regions of retrovirus envelope glycoproteins. To address this problem, we initiated structural studies on the envelope glycoprotein of the mammalian C-type retrovirus, Friend murine leukemia virus (Fr-MLV). The cellular receptor for Fr-MLV and other ecotropic MLVs (3) is the 14-transmembrane-pass cationic amino acid transporter MCAT-1 (4). The receptor-binding function of Fr-MLV has been localized to the amino-terminal region of the SU glycoprotein (5), which has been defined as a domain by proteolysis (6); thus we refer to this region as the Friend receptor-binding domain (Fr-RBD) (Fig. 1). In this report, we present the X-ray crystal structure (7) of Fr-RBD at 2.0 Å resolution (Table 1).

Fr-RBD is a roughly L-shaped molecule consisting of a lopsided anti-parallel β -sandwich and a helical subdomain formed from two extended loops of the sandwich (Fig. 2A, B). Six strands (strands 1, 2, 4, 5, 8, and 9) form the large sheet of the Fr-RBD sandwich, while three strands (strands 3, 6, and 7) make up the second face (8). The large sheet is curved to form nearly half a barrel, and the small sheet packs against the half-barrel at an angle of approximately 40°. The first loop of the helical subdomain (between strands 3 and 4), contains two α -

helices (helices A and D). Also within this loop is an extended coil from Leu 50 to Thr 100 that passes almost 360° around the end of the helical lobe. The second loop (between strands 6 and 7) contains an α -helix (helix G) that packs in an antiparallel orientation against helix D. Fr-RBD contains six disulfide bonds, four of which Linder and coworkers previously identified biochemically, and two of which they predicted correctly by computer modeling (9). Although the overall tertiary structure of Fr-RBD structure is unique, a comparison of the C α coordinates of Fr-RBD against all known folds, performed by the Dali server, Heidelberg (10), reveals modest similarity between the Fr-RBD β -sandwich and proteins with immunoglobulin (Ig) folds (11) (Fig. 3).

Amino acid sequence alignments of receptor-binding domains from C-type MLVs that use distinct receptors have highlighted regions that vary with tropism (Fig. 1). These variable regions, called VRA, VRB, and VRC (12), are mapped onto the Fr-RBD structure in Figure 2A. The sequence throughout most of the β -sandwich is conserved, whereas the sequence corresponding to the helical lobe is highly divergent. VRA, VRB, and, to a lesser extent VRC, vary in length as well as in sequence between MLVs with different tropisms. For example, VRA in amphotropic MLVs is considerably shorter than VRA of ecotropic MLVs (61 residues and 1 putative disulfide bond vs. 98 residues and 3 disulfides), while amphotropic VRB is longer by 19 residues, including two additional cysteines. The pattern of sequence conservation in the receptor-binding domains of MLVs, when viewed in light of the Fr-RBD structure, indicates that the Ig-like core is used as a scaffold for displaying the receptor-binding regions of the envelope glycoprotein.

Chimeric envelope proteins have been constructed by others in attempts to localize the receptor-choice determinants in MLV SU glycoproteins. Although functional chimeras can be constructed readily between various non-ecotropic viruses (13), it has not been possible to construct functional chimeras containing both ecotropic and non-ecotropic variable regions (14). Upon examination of the Fr-RBD structure, this observation is not surprising. VRA and VRB are intimately associated with one another in the helical lobe, and a significant alteration of the length and sequence of either VRA or VRB alone would not easily be accommodated by the rest of the structure (15).

Despite the absence of useful chimeras, mutagenesis experiments identified amino acids in the VRA region of Moloney murine leukemia virus (Mo-MLV) SU that, when altered, inhibit MCAT-1 binding without affecting envelope processing and assembly into virions (16). Substitution of Lys for Asp 84 in Mo-MLV SU (corresponding to Asp 86 in closely related Fr-MLV), blocks binding and infection. When residues Arg 83, Glu 86, and Glu 87 in Mo-MLV (homologous to Arg 85, Asp 88, and Glu 89 in Fr-MLV) are changed individually to Glu, Lys, and Lys respectively, binding is abrogated, although virus titer is not significantly affected (17). The binding surface defined by the positions of these mutants is a textured landscape that includes a hydrophobic pocket and a charged ridge (Fig. 4).

The Fr-RBD structure demonstrates that, in the SU glycoproteins of MLVs, the polypeptide chain is partitioned structurally into a conserved Ig-like anti-parallel β -sheet framework and a variable subdomain. The Fr-RBD architecture will therefore facilitate the development of viruses with novel receptor specificities for the purpose of gene therapy. Current approaches have included tethering large binding domains such as single-chain antibodies to the amino terminus of the envelope protein (18) or substituting erythropoietin in place of SU sequences (19). The resultant recombinant viruses have very low transduction efficiencies, and in some cases require the wild-type glycoprotein to be co-incorporated into virions (19). Using the Fr-RBD structure as a model, modifications can now be directed to the receptor-binding lobe alone, with a detailed understanding of how the components of this subdomain interact with one another and with the core of the protein. The feasibility of replacing the variable subdomain sequences to generate new receptor-specificities and tropisms has already been demonstrated by the success of viral evolution in generating the receptor-diversity of this subgenus.

In spite of a complete lack of sequence homology, the TM proteins of distinct retroviruses have recently been shown to share a similar core structure (20). Coupled to the growing recognition of the importance of multi-pass transmembrane proteins in retroviral entry (4, 21), this observation suggests the intriguing possibility that SU glycoproteins from even distantly-related retroviruses may also share some structural similarities. Although the structure of the HIV SU subunit gp120 is unknown, the Fr-RBD fold provides a conceptual

framework in which to consider models for the structural basis of cell entry by HIV and other retroviruses.

References

1. E. Hunter and R. Swanstrom, *Curr. Top. in Microbiol. and Immun.* **157**, 187 (1990).
2. A.E. Smith, *Ann. Rev. Microbiol.* **49**, 807 (1995).
3. MLVs have been grouped on the basis of superinfectivity resistance, the phenomenon whereby cells preinfected with an MLV cannot be infected by an exogenous virus that presumably utilizes the same receptor. These groups have been named according to their host-ranges: ecotropic, amphotropic, polytropic, and xenotropic. Fr-MLV is an ecotropic MLV and infects only murine and certain other rodent cells. Amphotropic and polytropic viruses infect both murine and non-murine cells, while xenotropic viruses infect only non-murine cells.
4. L.M. Albritton, L. Tseng, D. Scadden, J.M. Cunningham, *ibid.* **57**, 659 (1989); J.W. Kim, E.I. Closs, L.M. Albritton, J.M. Cunningham, *Nature* **352**, 725 (1991); H. Wang, M.P. Kavanaugh, R.A. North, D. Kabat *ibid.* **352**, 729 (1991).
5. The amino-terminal domain of SU expressed in isolation is sufficient to mediate superinfectivity resistance in both ecotropic and amphotropic MLVs (26). Furthermore, this domain binds MCAT-1 *in vitro* (R.A.D., C.A.H., and J.M.C., in preparation).
6. A. Pinter, W.J. Honnen, J.-S. Tung, P.V. O'Donnell, U. Hammerling, *Virology* **116**, 499 (1982).
7. We expressed in insect cells, from a baculovirus vector, a truncated form of the ecotropic envelope from the Friend isolate (GenBank accession number: J02192) (27). The coding sequence was fused in-frame to a Factor Xa protease site (IEGR), three copies of the HA epitope (YPYDVPYA) (28), and a 6-histidine motif. Details of this vector and of protein production will be provided elsewhere (R.A.D., C.A.H., and J.M.C., in preparation). Briefly, insect cells in log growth phase were infected with recombinant virus. Culture supernatants were collected three days post-infection, protein was purified over a nickel-primed chelating resin (Pharmacia), and the tags were removed with Factor Xa protease (NEB). Cleaved protein, containing residues 1 through 236 of mature Fr-MLV, as well as 6 additional residues remaining on the carboxy-terminus from the factor Xa cleavage site, was repurified using a Mono S ion exchange column (Pharmacia). The eluted protein was dialyzed against 20 mM sodium acetate, pH 5.5, and concentrated to 10 mg/ml using a centricon-3 microconcentrator (Amicon).
Crystals of Fr-RBD were initially grown by the hanging drop method at 20°C by equilibrating over a reservoir containing 150 mM calcium acetate, 100 mM sodium cacodylate, pH 6.5, 19% PEG 8000. In 3-4 days, crystals with a long rod- or needle-like morphology appeared. When zinc acetate was substituted for

calcium acetate in the well solution, crystals did not grow readily, and the crystals that were observed were clustered. However, streak-seeding with a human eyelash from the crystals grown in the presence of calcium to protein pre-equilibrated for 12 hours in zinc solution yielded single, block-like crystals in the space group $P2_12_12_1$ ($a = 55.3 \text{ \AA}$, $b = 68.4 \text{ \AA}$, $c = 80.4 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$), with a single Fr-RBD monomer in the asymmetric unit. Subsequent crystals were grown by streak-seeding either from crystals grown in calcium or from zinc crystals that were themselves grown by streak-seeding. 7% glycerol was later included in the crystallization conditions as a cryo-protectant. Solvent content of the crystals is approximately 52%.

Diffraction data were collected using a rotating anode source (Rigaku RU200) with an RAXIS IIC detector. All data were collected at -180°C using an MSC cryogenic crystal cooler (X-stream). Crystals were mounted in a $20 \mu\text{m}$ rayon fiber loop and flash-frozen in the gaseous nitrogen stream. Reflections were indexed and integrated with the program DENZO (29) and were scaled using SCALEPACK (29). Subsequent data manipulations were carried out using the CCP4 package (30). Heavy atom derivatization was accomplished by harvesting crystals into reservoir solution, adding solid heavy atom to 1 mM, and soaking overnight. Heavy atom sites were located for each derivative by Patterson methods. Heavy atom refinement and phasing were carried out with MLPHARE (31), and the sites for each derivative were verified by calculating difference fourier maps using phases from all other derivatives.

Density modification was performed on the MIR map using the program DM (32). The 3 \AA density-modified map was displayed with the program O (33) for tracing of the polypeptide chain. The Fr-RBD model contains all residues in the fragment except the first 8 at the amino terminus and the carboxy-terminal residues from the Factor Xa cleavage site. The model was refined against data to 2.0 \AA with the program XPLOR (34). Cycles of positional refinement, B-factor refinement, and rebuilding were continued until the free R factor (35) fell below 35%, at which point ordered waters were added and a bulk solvent correction was applied. The quality of the structure was verified by Procheck (30), with no residues falling in disallowed regions of Ramachandran space, and by 1D-3D (36), which gave a score consistent with all sidechains being located in acceptable environments.

The Fr-RBD model contains one N-acetylglucosamine (GlcNAc) at each of the two N-linked glycosylation sites, residues 12 and 168. There is clear electron density in the MIR map for the first GlcNAc at position 168, but density is not visible for additional carbohydrate units. One GlcNAc was placed at position 12, but the precise orientation of this carbohydrate unit should be considered ambiguous because it is difficult to interpret its electron density.

Three zinc ions are bound in the Fr-RBD crystals: one at His 55, and two sites spaced 3.58 \AA apart from one another at a crystal contact, bridging Asp 21 and Ser 51 of one molecule and Asp 86 of another. Data to 2.8 \AA were subsequently collected on a crystal grown in calcium acetate. The lengths of unit cell edges b and c were decreased in these crystals by 2.8% and 3.9%, respectively

($a = 55.3 \text{ \AA}$, $b = 66.5 \text{ \AA}$, $c = 77.3 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$). Molecular replacement using the program Amore (37) and refinement to a free R factor of 29.4% demonstrated that the packing is shifted slightly but significantly in these crystals, diminishing the distance between symmetry-related molecules and eliminating the intersubunit zinc ion binding sites. Therefore, the zinc ion binding sites are not integral to the Fr-RBD structure. There are no significant differences between the protein structures refined against data from the zinc- or calcium-containing crystals.

8. Although strand 3 contains two prolines that prohibit it from being considered a β -strand according to standard secondary structure classification schemes such as that of Kabsch and Sander (38), we consider it a β -strand here because it is in a position to make four backbone hydrogen bonds to strand 7.

9. M. Linder, D. Linder, J. Hahnen, H.-H. Schott, S. Stirm, *Eur. J. Biochem.* **203**, 65 (1992). Four disulfide bonds (cysteines 46-98, 72-87, 73-83, and 178-184) are in the helical lobe, while 2 disulfide bonds (cysteines 121-141, and 133-146) are in a loop between strands 4 and 5. Only one cysteine, at position 121, lies in a β -strand.

10. L. Holm and C. Sander, *J. Mol. Biol.* **233**, 123 (1993); L. Holm and C. Sander, *Nuc. Acids Res.* **24**, 206 (1996).

11. The following proteins gave Z scores ranging from 3.4 to 2.0 and α -carbon RMSDs between 3.1 and 4.9 \AA over approximately 90 residues each (an alignment was found for strands 3 through 9, but the helical lobe and strands 1 and 2 could not be aligned): Single chain FV fragment, light chain (pdb1mfa.ent); human CD2 (pdb1hnf.ent); OPG2 Fab fragment, heavy chain (pdb1opg.ent); rat CD4, domain 3 (pdb1cid.ent); human CD8, domain 1 (pdb1cd8.ent); Fab fragment, light chain (pdb1mlb.ent); Beta chain of murine T cell antigen receptor, variable domain (pdb1bec.ent); human CD4, domain 1 (pdb3cd4.ent); human Coagulation Factor XIII (pdb1ggt.ent); E. Coli PAP*D chaperone (pdb3dpa.ent); human p53 tumor suppressor protein (pdb1tup.ent); human VCAM-1, domain 1 (pdb1vca.ent); murine N-cadherin, domain 1 (pdb1nci.ent).

12. These variable regions have previously been referred to as VRA and VRB (39). Biased by the Fr-RBD structure, and given the existence of a conserved region near the end of VRA, we have divided VRA into two distinct variable regions, one which we still refer to as VRA and the other which we now call VRC.

13. D. Ott, and A. Rein, *J. Virol.* **66**, 4632 (1992).

14. C. Peredo, L. O'Reilly, K. Gray, M.J. Roth, *ibid.* **70**, 3142 (1996); R. A. Morgan, O. Nussbaum, D.D. Muenchau, L. Shu, L. Couture, and W.F. Anderson, *ibid.* **67**, 4712 (1993).

15. We have generated a packing model for Fr-RBD as it would exist in the trimeric envelope glycoprotein complex. This trimer packing model was constructed manually by placing the carbohydrates toward solvent, keeping variable sequences surface-exposed rather than at trimer contacts, and maximizing contact area while avoiding steric clashes. Although VRC is distant from VRA and VRB within an Fr-RBD monomer, VRC from one subunit is close to VRB of the adjacent subunit of the trimer model, such that there are only three variable lobes per trimer, with each being composed of sequences from two Fr-RBD subunits (D. F., Ph.D. thesis, Massachusetts Institute of Technology, 1997).
16. A.J. MacKrell, N.W. Soong, C.M. Curtis, W.F. Anderson, *ibid.* **70**, 1768 (1996).
17. MacKrell and coworkers (16) also identified two mutants, His 123 and Arg 124 (His 125 and Arg 126 in Fr-MLV SU), that severely reduce incorporation of envelope glycoprotein into virions, consistent with the mutant glycoproteins having defects in folding or assembly. The side chain of His 125 in Fr-RBD is located near a buried aspartic acid, Asp 161, that is conserved among all MLVs. Replacing His 125 with a different hydrophobic residue would isolate the buried charge of Asp 161, potentially destabilizing either VRC, the core of the domain, or both regions. The other structurally significant residue, Arg 126, is partially buried and in position to form a hydrogen bond to the backbone carbonyl of residue 134. This interaction could anchor the VRC loop or aid in proper disulfide bond formation of nearby cysteines and is apparently also critical for proper folding and/or processing.
18. S. Valsesia-Whittman *et al.*, *J. Virol.* **70**, 2059 (1996).
19. N. Kasahara, A.M. Dozy, Y.W. Kan, *Science* **266**, 1373 (1994).
20. D. Fass, S.C. Harrison, P.S. Kim, *Nature Struc. Biol.* **3**, 465 (1996); D. C. Chan, D. Fass, J.M. Berger, and P.S. Kim, *Cell* **89**, 263 (1997).
21. R.A. Weiss and C.S. Taylor, *Cell* **82**, 531 (1995); M. Samson *et al.*, *Nature* **382**, 722 (1996); R. Liu *et al.*, *Cell* **86**, 367 (1996); M. Dean *et al.*, *Science* **273**, 1856 (1996).
22. M. Carson, *J. Appl. Crystallogr.* **24**, 958 (1991).
23. P. Kraulis, *ibid.* **24**, 924 (1991).
24. P. Bork, L. Holm, L., C. Sander, *J. Mol. Biol.* **242**, 309 (1994).
25. A. Nicholls, K.A., Sharp, B. Honig, *Proteins* **11**, 281, 1991.

26. J.M. Heard and O. Danos, *J. Virol.* **65**, 4026 (1991); J.-L. Battini, O. Danos, J.M. Heard, *ibid.* **69**, 713 (1995).
27. C. Friend, *J. Exp. Med.* **105**, 307 (1957); W. Koch, G. Hunsmann, R. Friedrich, *J. Virol.* **45**, 1 (1983).
28. D.J. Goldstein and R. Schlegel, *EMBO J.* **9**, 137 (1990).
29. Z. Otwinowski, In *Data Collection and Processing*, L. Sawyer, N. Isaacs, S. Bailey, eds. (Warrington, United Kingdom: SERC Daresbury Laboratory, 1993), pp. 56-62.
30. Collaborative Computational Project, Number 4, *Acta Cryst.* **D50**, 760 (1994).
31. Z. Otwinowski, In *Isomorphous Replacement and Anomalous Scattering*, W. Wolf, P.R. Evans, and A.G.W. Leslie, eds. (Warrington, United Kingdom: SERC Daresbury Laboratory, 1991) pp. 80-86.
32. K. Cowtan, *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34 (1994).
33. T.A. Jones and M. Kjeldgaard (1992). O -- The Manual (Uppsala, Sweden: <http://kaktus.kemi.aau.dk>).
34. A.T. Brünger, *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
35. A.T. Brünger, *Nature* **355**, 472 (1992).
36. K.Y. Zhang and D. Eisenberg, *Protein Sci.* **3**, 687 (1994).
37. J. Navaza, *Acta Cryst.* **A50**, 157 (1994).
38. W. Kabsch, and C. Sander, *Biopolymers* **22**, 2577 (1983).
39. J.-L. Battini, J.M. Heard, O. Danos, *J. Virol.* **66**, 1468 (1992).
40. We thank Stephen C. Harrison and Steven J. Gamblin for helpful discussions and Ndubuisi Azubuine for media preparation. J.M.B. is a Whitehead Fellow and acknowledges support from the W. M. Keck Foundation. This work was funded by the Howard Hughes Medical Institute and utilized the W.M. Keck Foundation X-ray Crystallography Facility at the Whitehead Institute. The coordinates will be deposited in the Brookhaven Protein Data Bank.

Fig. 1. Domain organization of the Fr-MLV envelope glycoprotein and sequence of Fr-RBD. **(A)** The Fr-MLV envelope glycoprotein is synthesized as a single polypeptide chain, shown schematically (SS = signal sequence, FP = fusion peptide region, TM = transmembrane segment). Maturation involves removal of the signal sequence, cleavage to generate distinct SU and TM subunits, and removal of a short peptide at the carboxy terminus. Cleavage sites are indicated by dotted lines. The crystallized Fr-RBD region is shaded. Amino acid residues at the junctions of each indicated segment are numbered. **(B)** Secondary structural elements are indicated above the amino acid sequence of Fr-RBD. β -strands are numbered and shown as black bars, with strand 3 dotted because it does not classify as a true strand (8). α -helices are lettered and shown as gray bars, and 3_{10} -helices are dotted in gray. N-linked glycosylation sites are underlined. Regions conserved among MLVs are boxed in gray, while regions whose lengths and sequences vary with the tropism of the virus are boxed in blue and labeled VRA, VRB, and VRC (12).

Figure 1A

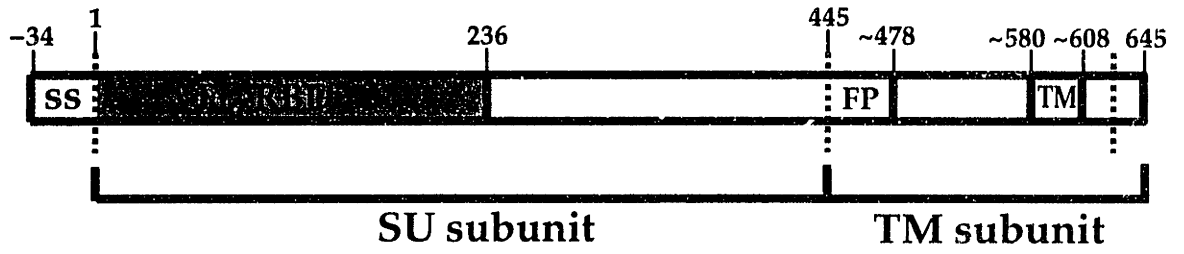


Figure 1B

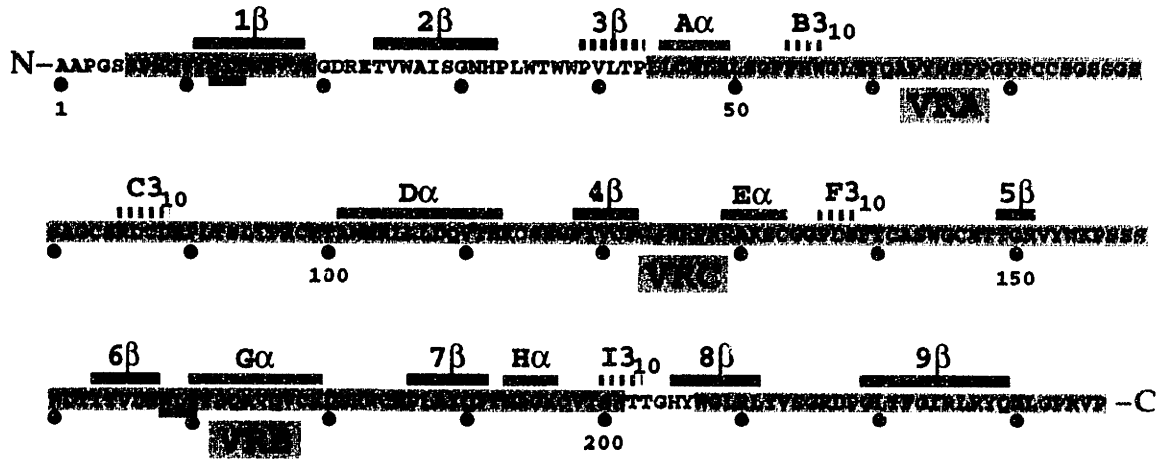


Fig. 2. Structure of Fr-RBD. **(A)** A ribbon representation is shown with disulfide bonds and carbohydrates indicated as ball-and-stick representations. Strands are numbered and helices are lettered to correspond to Fig. 1. Variable regions are indicated (12). VRA includes the extended coil in the helical subdomain (residues 50 to 100) and helix D, while VRB corresponds to helix G. VRC is located in the loop between strands 4 and 5 and is separated from VRA and VRB in the Fr-RBD structure by conserved elements such as strands 4, 5, 8, and 9 (15). The figure was generated using Ribbons (22). **(B)** A stereo diagram of Fr-RBD is numbered every 10 residues. The figure was generated using Molscript (23).

Figure 2A

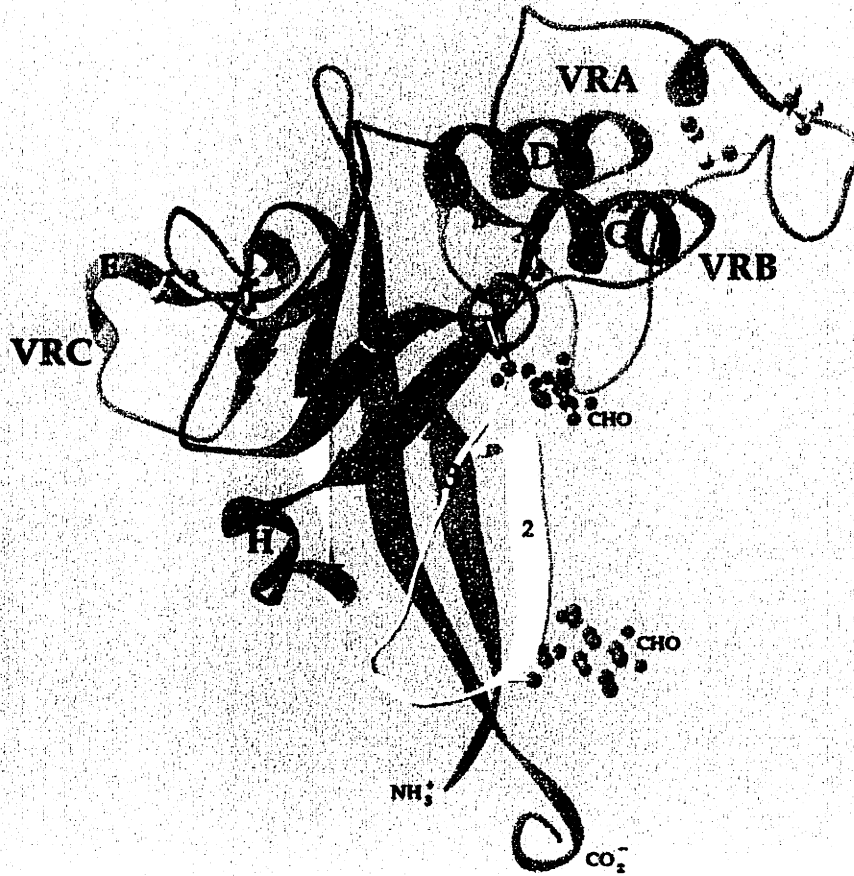


Figure 2B

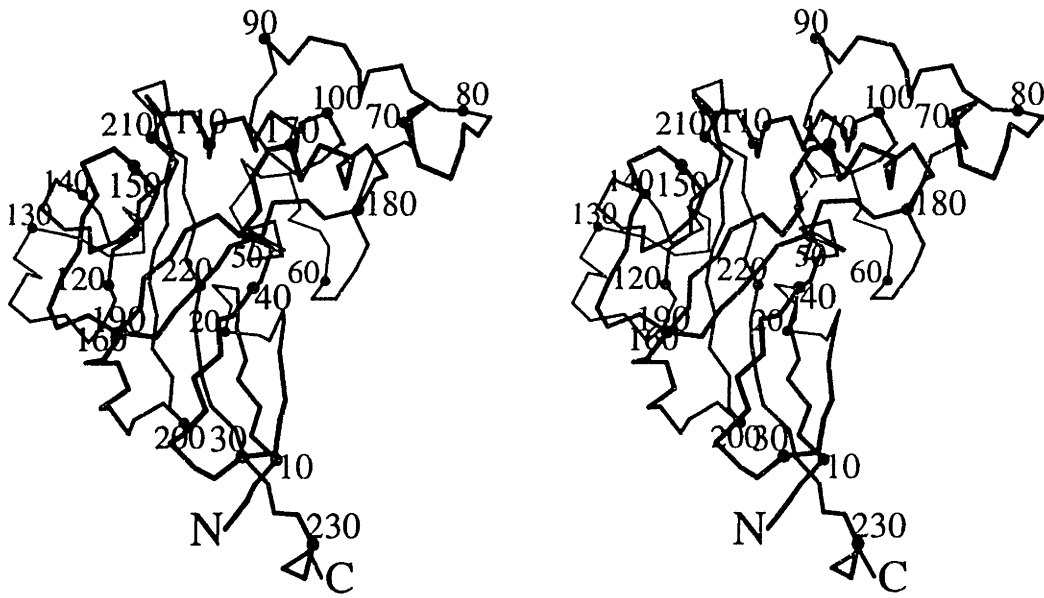
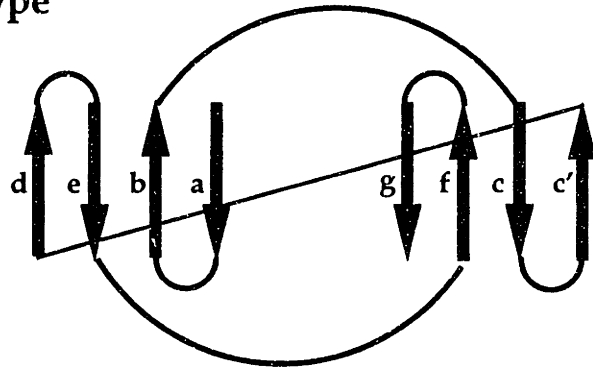


Fig. 3. Topological relationship between Fr-RBD and the Ig fold. A side view of the topologies of an h-type Ig fold and of Fr-RBD are shown with the β -sandwiches splayed open between strands c' and d in the Ig fold and between strands 5 and 6 in Fr-RBD. Strands 1 and 2 of Fr-RBD are shown in light gray to indicate that they are a modification of the classical h-type Ig fold topology (24).

Figure 3

Ig fold
h-type



Fr-RBD

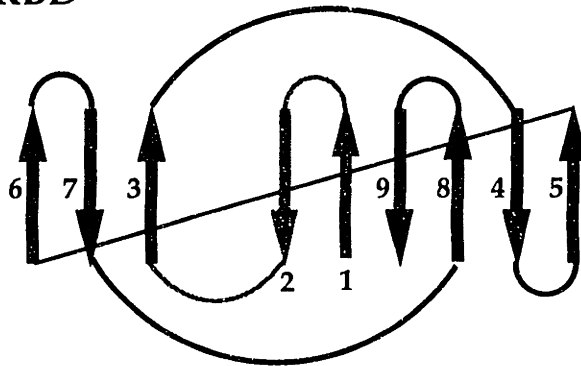


Fig. 4. The putative receptor-binding surface. A molecular surface is superimposed on the structure of the region containing likely receptor-contacting residues. Residues Arg 85, Glu 89, Pro 90, and Thr 92 form a ridge across the top of this region. The face just below the ridge contains the critical Asp 86 (see text), which lies adjacent to a hydrophobic pocket. This pocket is lined with the residues Pro 90, Leu 94, and Trp 102, as well as Asp 86, and its base is formed by Leu 91. The relationship of the putative receptor-binding surface to the intact domain can be derived from the inset, in which the expanded region is boxed in red, and the helices are lettered as in Figs. 1 and 2A. The orientation is roughly a 180° rotation about the y-axis from the views in Fig. 2. The figure was generated using Grasp (25).

Figure 4

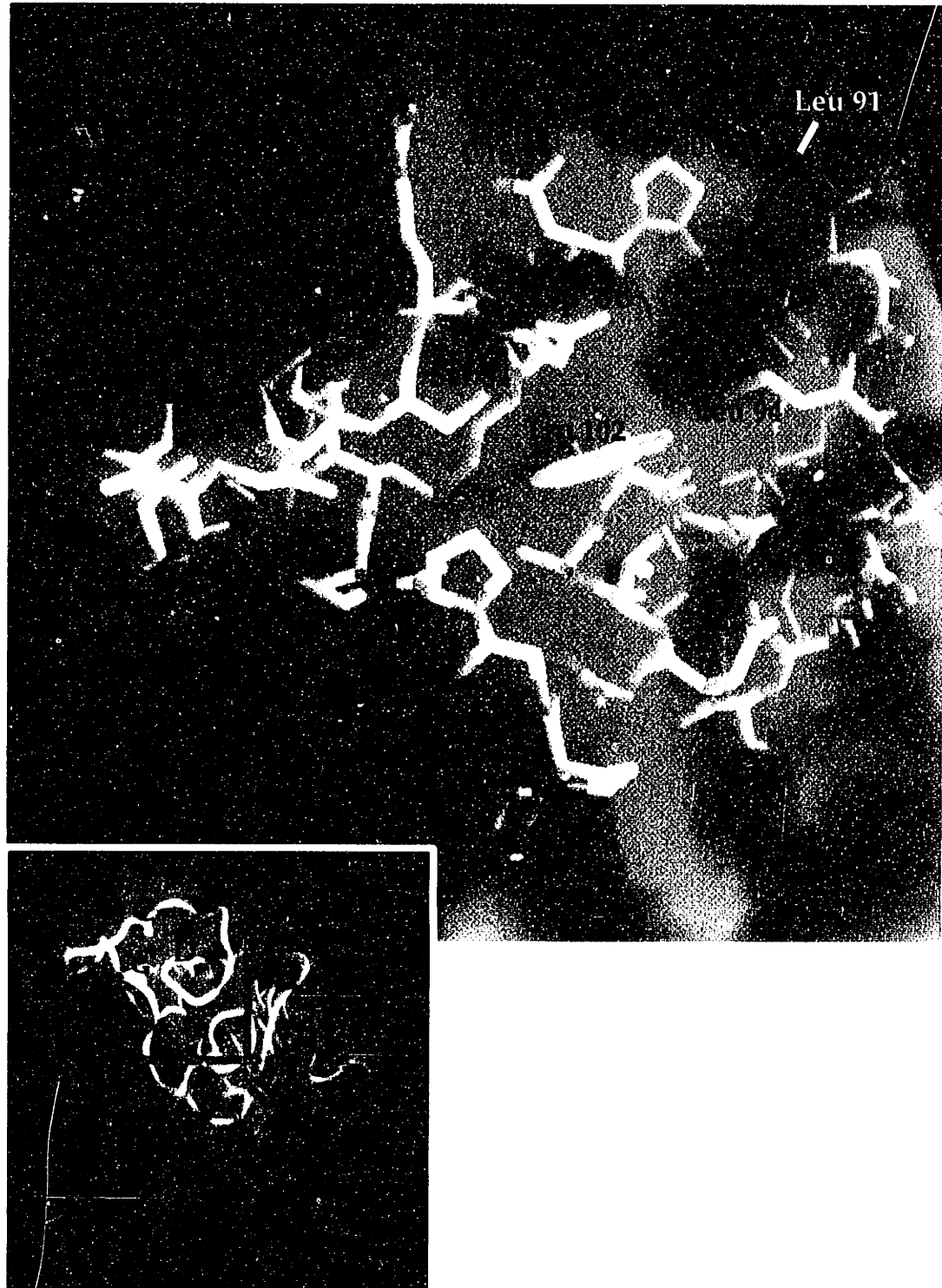


Table 1 Multiple isomorphous replacement and refinement statistics

* $R_{\text{sym}} = \sum \sum_j |I_j - \langle I \rangle| / \sum I_j$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

† $R_{\text{iso}} = \sum | |F_{\text{ph}}| - |F_{\text{p}}| | / \sum |F_{\text{p}}|$, where F_{ph} and F_{p} are the derivative and native structure factors, respectively.

‡Phasing power = $\langle F_{\text{h}} \rangle / E$, where $\langle F_{\text{h}} \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error.

§Cullis $R = \sum | |F_{\text{ph}} \pm F_{\text{p}}| - |F_{\text{h,c}}| | / \sum |F_{\text{ph}} \pm F_{\text{p}}|$, where $F_{\text{h,c}}$ is the calculated heavy-atom structure factor.

|| $R_{\text{work, free}} = \sum | |F_{\text{obs}}| - |F_{\text{calc}}| | / \sum |F_{\text{obs}}|$, where the working and free R factors are calculated using the working and free reflection sets, respectively. The free reflections (6.9 % of the total) were held aside throughout refinement.

Table 1. Multiple isomorphous replacement and refinement statistics

			Pt(II)(2,2':6',2"·	Phenyl		
Data collection	Native	HgCl ₂	terpyridine)Cl	HgAc	AgNO ₃	PtCN ₄
Resolution (Å)	2.0	3.0	3.0	4.0	3.0	3.0
R _{sym} * (%)	6.7	13.5	10.7	7.9	7.9	15.9
Completeness (%)	97.2	99.6	99.2	84.4	98.6	80.2
R _{iso} † (%)		22.0	20.4	20.0	15.8	19.7
Number of sites		4	1	4	4	2
Phasing power‡		1.11	1.33	1.31	1.02	0.74
Cullis R§		0.812	0.745	0.758	0.818	0.904
Overall figure of merit 0.599						
Refinement			Structure and Stereochemistry			
Resolution (Å)	20.0-2.0		Number of atoms			
Number of reflections			Protein			1786
working	18928		Carbohydrate			28
free	1393		Water			191
R _{work} /R _{free} (%)	22.3/26.4		Zinc ions			3
			rmsd bond lengths (Å)			0.014
			rmsd bond angles (degrees)			2.99

APPENDIX I

MODEL FOR TRIMER PACKING OF THE RECEPTOR-BINDING DOMAIN FROM ECOTROPIC MURINE LEUKEMIA VIRUSES

The glycoprotein complexes of many enveloped viruses are trimeric (for review see White, 1990). This generality may reflect one of the major functions of these protein assemblies, to initiate fusion of the viral and cellular membranes during viral entry. If DNA-binding proteins are often dimeric because two points determine, or regulate, the linear DNA, then membrane-fusion proteins may be trimeric because three points determine, or modulate, the plane of the lipid bilayer.

Trimeric assemblies have been detected, using various techniques and to various resolutions, in numerous families of enveloped viruses. The high-resolution structures of both the native (Wilson et al., 1981) and the low-pH converted (Bullough et al., 1994) HA proteins from the orthomyxovirus influenza reveal trimers organized around three-stranded coiled coils. The envelope glycoprotein of the avian oncoretrovirus Rous sarcoma virus has been shown to be trimeric by velocity sedimentation (Einfeld and Hunter, 1988). Some reports claim that the lentivirus envelope glycoprotein is trimeric (Weiss et al., 1990), although other oligomerization states have been proposed (Rey et al., 1990; Doms et al., 1991). Even the flavivirus E protein, for which a high-resolution dimeric native structure is known (Rey et al., 1995), converts to a trimeric membrane-fusion competent structure (Allison et al., 1995). Complementing these studies on intact glycoprotein complexes, peptide studies have identified regions of viral membrane-fusion subunits that form trimers in solution. For example, high-resolution trimeric structures are available for the cores of the TM subunits of MLV (Fass et al., 1996) and HIV (Chan et al., 1997).

Although fragments of the transmembrane subunits of some enveloped viruses reflect the oligomerization of the intact glycoprotein complexes (Carr and Kim, 1993; Fass and Kim, 1995), regions of receptor-binding subunits studied in isolation may not. Fr-RBD, the receptor-binding domain of Friend murine leukemia virus (Fr-MLV) is monomeric in solution (R. Davey, unpublished data) and crystallizes without three-fold symmetry axes that could indicate how the domain packs in a native glycoprotein trimer (Chapter 5). This result is not unexpected, since a comparable situation was previously found for influenza HA. The influenza HA "tops" domain (residues 28-328 of HA1), which contains

the sialic acid binding site, makes trimer contacts in the native HA structure (Wilson et al., 1981). These contacts are broken during the low-pH induced conformational change (Graves et al., 1983), and the domain can be proteolytically cleaved from the remainder of the structure. This isolated tops domain is monomeric in solution (Ruigrok et al., 1988) and crystallizes as such in complex with a Fab fragment (Bizebard et al., 1995). Nevertheless, the tops domain is in the same conformation either as part of the native HA trimer or as bound to the Fab fragment in monomeric form after low pH and protease treatment.

If the structural and functional analogy between Fr-RBD and the influenza tops domain is sufficiently strong, then Fr-RBD may also make direct, if weak, trimer contacts in the native glycoprotein complex of Fr-MLV. Low resolution studies support this likelihood. First, the native MLV envelope glycoprotein assembly is known to be trimeric (Kamps et al., 1991). In addition, electron micrographs of MLV show the surface of the virions to be studded with compact knobs or spikes (Nermut et al., 1972), suggesting that the various domains of the complex are organized closely around the three-fold axis, which extends perpendicular to the membrane. Fr-RBD is the largest domain in the complex and is probably the farthest removed from the membrane, since it contains the receptor-binding regions. Therefore, the close-packed knobs in the electron micrographs primarily reflect the position of Fr-RBD. These features of the Fr-MLV envelope suggested the possibility that the Fr-RBD structure could be reassembled to model the packing of the domain in the intact glycoprotein trimer.

A reliable computational method for optimal docking of protein domains is not available. Nevertheless, a number of features of the Fr-RBD structure were used to indirectly suggest trimer packing models. First, carbohydrates should be exposed to solvent. Also, residues shown to be important for receptor binding should face roughly away from where the viral membrane would be. The carboxy terminus of Fr-RBD should point approximately in the direction of the viral membrane, since the carboxy-terminal domain of SU, which is not part of the crystallized fragment, makes contacts to TM which is in turn fixed in the viral membrane. The packing interface should contain relatively conserved residues, while variable regions should be absent from the core of the trimer.

After these considerations were applied to roughly orient the Fr-RBD molecules with respect to one another, reassembly was completed by manually attempting to bury as much surface area as possible while avoiding steric clashes.

Using the above constraints, which greatly reduced the number of possible subunit orientations, a plausible model was constructed for trimer formation between Fr-RBD monomers (Figures 1, 2). In this model, 834 Å², or 9.4%, of the surface of each monomer becomes inaccessible as calculated for a sphere with radius 1.4 Å using Grasp (Nicholls et al., 1991). For comparison, surface accessibility was calculated for the globular portion of influenza HA1 tops domain (residues 55 to 276). This region of HA1 was selected because it is structurally similar to Fr-RBD in that it is relatively compact, it is the most membrane-distal region of HA1, and it contains the receptor-binding sites. The trimer contacts within these HA1 domains, ignoring contacts to the rest of the glycoprotein spike, bury a comparable 895 Å², or 10.4%, of the surface of each HA1 globular domain monomer. The contacts made by the HA1 globular domain to both the remaining HA1 regions and to HA2 were ignored to formulate an appropriate comparison, since it is not known how Fr-RBD associates with the carboxy-terminal region of SU or with the TM subunit.

The positions of variable regions in the proposed Fr-RBD trimer model are aesthetically satisfying and may provide a means to test the model. In an Fr-RBD monomer (Chapter 5), VRA and VRB pack intimately with one another, but VRC is separated from VRA and VRB by a section of β -sheet. In the Fr-RBD trimer model, VRC from each monomer packs near VRB from the adjacent Fr-RBD subunit (Figure 3). Each helical lobe in the model is thereby composed of variable regions from two separate subunits. The sequence database of MLVs is not large or diverse enough to determine rigorously if the sequences of VRB and VRC co-vary in terms of size, charge, or other character. However, it would be possible to directly test the importance of an interaction between VRB and VRC by swapping VRC regions and measuring the envelope stability and/or infectivity of the resulting hybrid virions.

The overall shape of the trimer model, as viewed down the three-fold axis, is roughly triangular, with the putative receptor-binding residues located near the apices (Figures 1, 3). In the side view, a cavity can be seen underneath

the trimer (Figure 2). Interestingly, a similar cavity exists under the globular domain of influenza tops in the native HA trimer (Figure 4). In the intact HA structure (Wilson et al., 1981), this cavity accommodates the HA2 subunit in its "spring-loaded" conformation (Carr and Kim, 1993). It is possible that the TM subunit of Fr-MLV is similarly accommodated under the Fr-RBD trimer, and that the carboxy-terminal regions of the Fr-MLV SU subunits extend down along the outside of TM. It is not known whether retroviral fusion peptides are buried at the core of the TM subunits in the native glycoprotein complex, analogous to the influenza HA fusion peptides, or whether they are shielded from solvent directly by the SU subunit. In either case, however, Fr-RBD could serve to regulate the fusion-active state of the retroviral envelope, preventing premature exposure of the fusion peptide regions.

The proposed trimer model has several implications for the mechanism of receptor binding and viral entry. First, the monomeric state of Fr-RBD in solution, in spite of a reasonable packing model, supports an analogy between Fr-RBD and the flu tops domain. Once dissociation occurs, the isolated Fr-RBD subunits, like those of flu HA1, may not readily retrimmerize. Thus, dissociation of the Fr-RBD trimer might be necessary for the TM fusion peptides to be exposed, analogous to the dissociation of the HA1 domain in flu HA upon exposure to low pH (White and Wilson, 1987; Kemble et al., 1992). The lack of trimerization of both the Fr-RBD and the influenza tops domains exists in marked contrast to the highly stable trimeric association of fragments of the MLV and influenza transmembrane subunits (Fass and Kim, 1995; Carr and Kim, 1993)

Another implication of the Fr-RBD trimer model addresses the stoichiometry of binding between viral envelope and receptor. The amino acid sequence of the ecotropic MLV receptor, MCAT-1 (Albritton et al., 1989), does not imply any internal 3-fold structural symmetry, and it is therefore unlikely that an SU trimer would simultaneously use the putative receptor-binding region from each subunit to interact with a single receptor monomer. In the Fr-RBD trimer model, the three receptor binding regions are located far ($\sim 60 \text{ \AA}$) from one another, so it is possible that the Fr-RBD trimer can interact with three MCAT-1 molecules. The Fr-RBD domain alone can inhibit superinfectivity of the expressing cell (Heard and Danos, 1991), suggesting that the Fr-RBD monomer is

capable of binding to receptor when not present in the trimer complex. In addition, direct binding of the Fr-RBD domain to MCAT-1 has been observed in vitro (R. Davey, and J. Cunningham, unpublished results). These results support the hypothesis that each SU monomer in an intact trimer complex could independently contact a receptor molecule. Whether simultaneous binding of the glycoprotein trimer to multiple receptor molecules is required for Fr-MLV entry remains to be determined.

The high-resolution structure of an intact retroviral glycoprotein complex may not be readily forthcoming. The ease with which the SU subunit of retroviruses dissociates from the virion and the difficulty of producing large quantities of soluble versions of intact retroviral glycoprotein complexes require that alternative approaches be taken to generate models for the higher-order assembly and overall function of the retroviral envelope. A protein dissection approach to generating detailed structural models of components of the retroviral envelope can successfully provide structural frameworks to guide further studies. To understand the functions and interactions of these retroviral envelope components in the larger process of viral entry, it becomes important to formulate hypothetical structural models that can be tested with functional assays. The Fr-RBD trimer model is one such proposal, and represents an attempt to inspire the experiments and discussions that will lead toward a greater understanding of the structural basis for cell entry by enveloped viruses.

References

- Albritton, L.M., Tseng, L., Scadden, D., and Cunningham, J.M. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* **57**, 659-666 (1989).
- Allison, S.L., Schalich, J., Stiasny, K., Mandl, C.W., Kunz, C., and Heinz, F.X. Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J. Virol.* **69**, 695-700 (1995).
- Bizebard, T., Gigant, B., Rigolet, P., Rasmussen, B., Diat, O., Bosecke, P., Wharton, S.A., Skehel, J.J., and Knossow, M. Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature* **376**, 92-94 (1995).
- Bullough, P.A., Hughson, F.M., Skehel, J.J., and Wiley, D.C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37-43 (1994).
- Carr, C.M., and Kim, P.S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**, 823-832 (1993).
- Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-273 (1997).
- Doms, R.W., Earl, P.L., and Moss, B. The assembly of the HIV-1 *env* glycoprotein into dimers and tetramers. *Adv. Exp. Med. Biol.* **300**, 203-219 (1991).
- Einfeld, D., and Hunter, E. Oligomeric structure of a prototype retrovirus glycoprotein. *Proc. Natl. Acad. Sci. USA* **85**, 8688-8692 (1988).
- Fass, D., and Kim, P.S. Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin. *Curr. Biol.* **5**, 1377-1383 (1995).
- Fass, D., Harrison, S.C. and Kim, P.S. Retrovirus envelope domain at 1.7 Å resolution. *Nature Struc. Biol.* **3**, 465-469 (1996).
- Graves, P.N., Schulman, J.L., Young, J.F. and Palese, P. Preparation of influenza virus subviral particles lacking the HA1 subunit of hemagglutinin: unmasking of cross-reactive HA2 determinants. *Virology* **126**, 106-116 (1983).
- Kamps, C.A., Lin, Y.-C., and Wong, P.K.Y. Oligomerization and transport of the envelope protein of Moloney murine leukemia virus-TB and of ts1, a

neurovirulent temperature-sensitive mutant of MoMuLV-TB. *Viol.* **184**, 687-694 (1991).

Kemble, G.W., Bodian, D.L., Rose, J., Wilson, I.A. and White, J.M. Intermonomer disulfide bonds impair the fusion activity of influenza virus hemagglutinin. *J. Virol.* **66**, 4940-4950 (1992).

Nermut, M.V., Frank, H., and Schäfer, W. Properties of mouse leukemia virus, III. Electron microscopic appearance are revealed after conventional preparation techniques as well as freeze-drying and freeze-etching. *Viol.* **49**, 345-358 (1972).

Nicholls, A., Sharp, K.A., and Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296 (1991).

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

Rey, M.A., Laurent, A.G., McClure, J., Krust, B., Montagnier, L., and Hovanessian, A.G. Transmembrane envelope glycoproteins of human immunodeficiency virus type 2 and simian immunodeficiency virus SIV-mac exist as homodimers. *J. Virol.* **64**, 922-926 (1990).

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

Weiss, C.D., Levy, J.A., and White, J.M. Oligomeric organization of gp120 on infectious human immunodeficiency virus type I particles. *J. Virol.* **64**, 5674-5677 (1990).

White, J.M. and Wilson, I.A. Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J. Cell. Biol.* **105**, 2887-2896 (1987).

White, J.M. Viral and cellular membrane fusion proteins. *Annu. Rev. Physiol.* **52**, 675-697 (1990).

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

Figure 1 **Top view of hypothetical trimer packing model**

A surface representation, generated using Grasp (Nicholls et al., 1991), of the proposed Fr-RBD trimer is viewed down the three-fold axis from above. Individual subunits are colored red, orange, and yellow. In this model, the subunits close-pack and interdigitate, but no steric clashes exist.

Figure 1

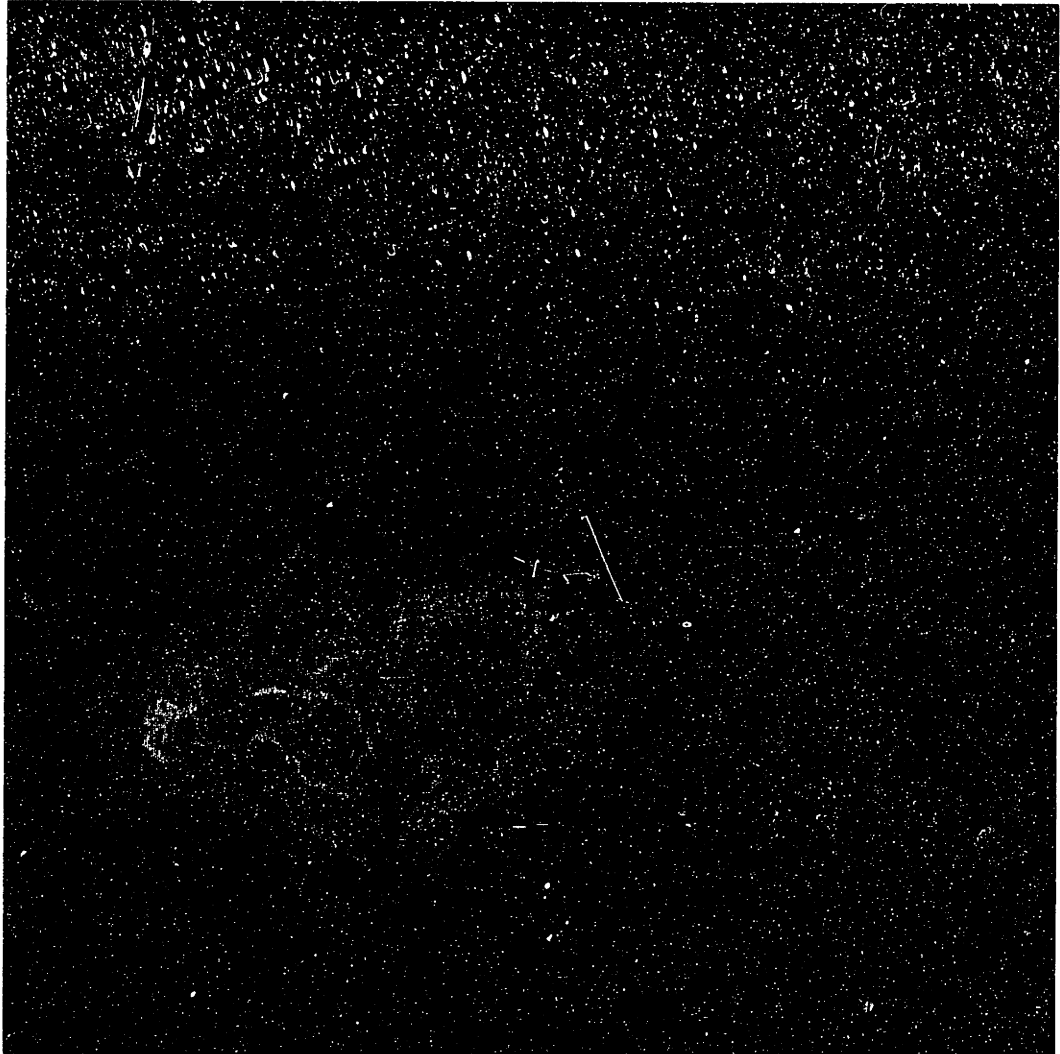


Figure 2 Side view of hypothetical trimer packing model

The surface representation from Figure 1 is viewed from the side. The cavity under the trimer assembly is formed from the amino- and carboxy-terminal regions of each Fr-RBD monomer, which are splayed out from the trimer axis.

Figure 2

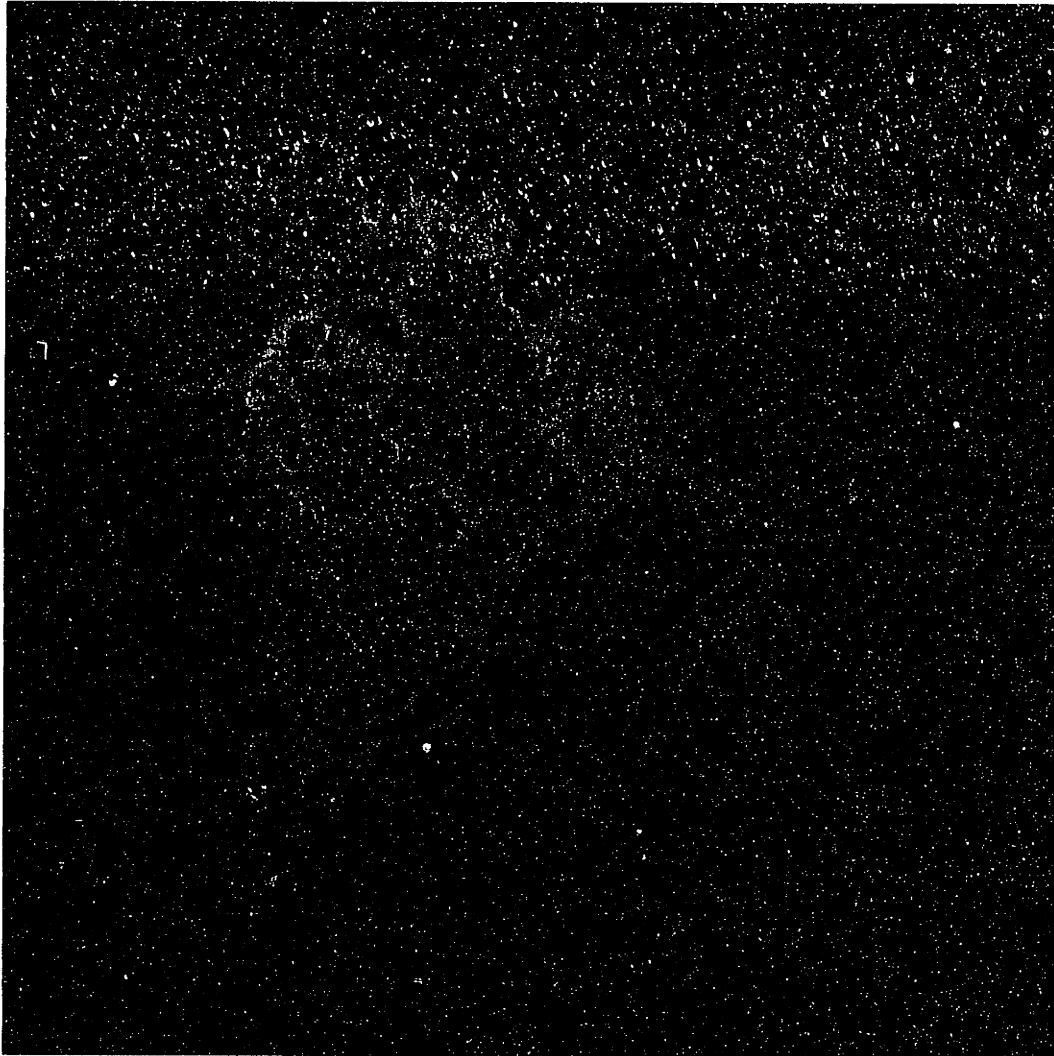


Figure 3 Variable regions within the Fr-RBD trimer model

A ribbons representation of the trimer model, viewed from the top, illustrates the positions of VRA, VRB, and VRC, colored white, yellow, and orange, respectively. The 3_{10} -helix containing putative receptor-contacting residues is circled.

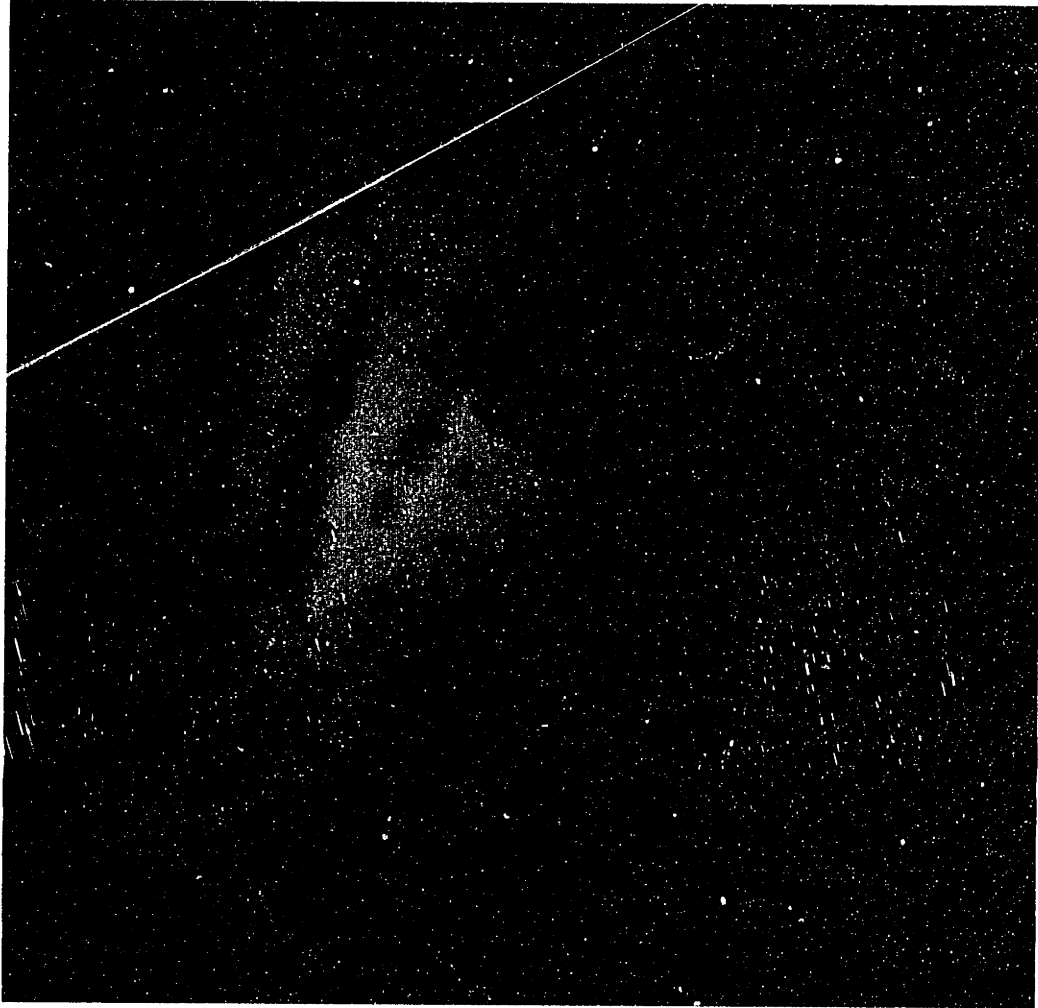
Figure 3



Figure 4 Trimeric assembly of the globular domain from influenza HA1

For comparison to the Fr-RBD trimer model, the globular region (residues 55 to 276) of the influenza HA1 subunit are shown as it packs in the native HA trimer (Wilson et al., 1981). The surface representation, generated using Grasp (Nicholls et al., 1991), is colored according to surface charge (blue is basic, red is acidic). A cavity similar to that under the Fr-RBD trimer can be seen at the base of this HA1 globular domain trimer.

Figure 4



Biographical Note

Deborah Fass

Education

- 1992 - present Doctoral candidate in Biology
Massachusetts Institute of Technology, Cambridge, MA
- 1991 B.A., Biochemistry, Magna cum Laude with Honors
Harvard-Radcliffe University, Cambridge, MA

Honors and Awards

- 1992-1995 National Science Foundation Predoctoral Fellow
- 1991 Phi Beta Kappa
- 1991 Bowdoin Prize for Essay in the Natural Sciences

Publications

Gupta, S., Fass, D., Shimizu, M., and Vayuvegula, B. "Potentiation of Immunosuppressive Effects of Cyclosporin A by 1 α ,25-Dihydroxyvitamin D₃," *Cellular Immunology*, **121**, 290-297 (1989).

Ponath, P.D., Fass, D., Liou, H.C., Glimcher, L.H., and Strominger, J.L. "The Regulatory Gene, hXBP-1, and its target, HLA-DRA, utilize both common and distinct regulatory elements and protein complexes," *J Biol Chem* **268**, 17074-17082 (1993).

Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S.C. "Crystal Structure of Transcription Factor E47 E-box Recognition by a Basic Region Helix-Loop-Helix Dimer," *Genes and Development*, **8**, 970-980 (1994).

Fass, D., and Kim, P.S. "Dissection of a Retrovirus Envelope Protein Reveals Structural Similarity to Influenza Hemagglutinin," *Current Biology* **5**, 1377-1383 (1995).

Fass, D., Harrison, S.C., and Kim, P.S. "Retrovirus Envelope Domain at 1.7Å Resolution," *Nature Structural Biology* **3**, 465-469 (1996).

Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. "Core Structure of gp41 from the HIV Envelope Glycoprotein," *Cell* **89**, 263-273 (1997).

Fass, D., Blacklow, S., Kim, P.S., and Berger, J.M. "Structure of a Calcium-Coordinating LDL Receptor Module," *Nature*, in press.

Fass, D., Davey, R.A., Hamson, C.A., Kim, P.S., Cunningham, J.M., and Berger, J.M. "Structure of a Retrovirus Receptor-Binding Glycoprotein: Murine Leukemia Virus SU domain at 2.0 Å Resolution," submitted.

Personal

Date of birth: February 21, 1970
Place of birth: Seattle, Washington
Citizenship: U.S.