Crystal Structure of a Mutation in the HIV gp41 Core: Effect on a Deep Cavity and Implication for Drug Development

by

Brian Cameros

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

Master of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 1999

September 1999

Author .....
Department of Physics
August 31, 1999

Certified by .....
Peter S. Kim
Professor of Biology
Thesis Supervisor

Certified by .....
Simon Mochrie
Associate Professor of Physics
Thesis Supervisor

Accepted by .....
Thomas J. Greytak
Professor, Associate Department Head for Education
Department of Physics

Science

Crystal Structure of a Mutation in the Human Immunodeficiency Virus gp41 core: Effect on a deep cavity and implication for drug development

by

Brian Cameros

Submitted to the Department of Physics on August 31, 1999, in partial fulfillment of the requirements for the degree of Master of Science in Physics

Abstract

Gp41, part of the envelope protein (Env) of Human Immunodeficiency Virus, Type-1 (HIV-1), mediates cell-virion fusion during HIV infection. The mechanism of fusion remains largely unclear, but the crystallographic structure of a partial model of gp41 is known. The structure is believed to correspond to the fusion-active core of the protein and contains a deep cavity important to the stability and infectivity of HIV-1. We solve the x-ray structure of this core with an important mutation (W631A) and investigate the effect on the cavity and its potential as a drug target.

Thesis Supervisor: Peter Kim
Title: Professor of Biology

Thesis Co-Supervisor: Simon Mochrie
Title: Professor of Physics

Introduction

The surface glycoproteins of enveloped viruses mediate the initial events in viral infection: identifying and fusing with the correct cell target. The envelope glycoprotein of human immunodeficiency virus type 1 is a complex of two noncovalently associated proteins, gp120 and gp41, which are derived from proteolytic cleavage of the precursor protein gp160 (review, Luciw, 1996). gp120 recognizes target cells by binding with the cell-surface receptor CD4 and one of several co-receptors that are members of the chemokine receptor family (Broder and Dimitrov, 1996; D'Souza and Harden, 1996; Wilkinson, 1996). Binding of gp120 is thought to expose gp41, which initiates fusion of the cellular and viral membranes.

The HIV envelope glycoprotein has similarities to other viral fusion proteins, for instance the hemagglutinin (HA) protein of influenza virus. HA is also proteolytically processed to generate a receptor-binding subunit (HA_1) and a membrane-spanning subunit (HA_2). Both gp41 (Fig. 1) and HA_2 contain a hydrophobic region which begins at the new amino terminus that results from proteolytic processing. This "fusion peptide" region is known to insert into target membranes as an early step in the fusion process (Stegmann et al., 1991; Tsurudome et al., 1992). Following the fusion peptide region (in gp41, HA_2 and other viruses) is a region with high alpha-helical propensity and a 4-3 hydrophobic repeat characteristic of coiled coils (Chambers et al., 1990; Delwart et al., 1990; Gallaher et al., 1989). X-ray structures of fragments of gp41, HA_2, and the transmembrane (TM) subunit of Moloney murine leukemia virus (Mo-MLV) show that those regions do form coiled coils (Chan et al., 1997; Bullough et al., 1994; Fass et al., 1996).

The mechanism by which gp41 directs fusion is still largely unclear. Numerous studies have indicated the existence of both native (nonfusogenic) and fusion-active conforma-

tions of fusion proteins. In the native influenza complex, for instance, part of the heptad-repeat region of HA_2 folds as a non-helical loop but converts to a coiled coil at low pH (Wilson et al., 1981; Carr and Kim, 1993; Bullough et al., 1994). Since low pH also activates influenza fusion, the low pH conformation is considered fusion-active. The HIV envelope complex is also thought to undergo conformational changes in the transition from the native to the fusion-active state (review, Chan and Kim, 1998). CD4 binding induces structural changes in both gp41 and gp120 which are thought to expose the hydrophobic fusion peptide region essential to membrane fusion (Sattentau and Moore, 1991, 1993; Allan et al., 1990; Sullivan et al., 1995).

It has not yet been possible to solve the crystal structure of gp41, but through the use of protein digests it has been possible to identify and crystallize what is thought to be an important substructure of the fusion-active protein. Limited proteolysis of a fragment corresponding to the ectodomain of gp41 led to the identification of a stable, soluble complex composed of two peptide fragments, N36 and C34 (Fig. 1), which are derived from the N- and C-terminal domains of the ectodomain, respectively (Lu et al., 1995; Chan et al., 1997). The N36 peptide corresponds to the heptad-repeat region, while the C34 peptide is from the region prior to the transmembrane section. The crystal structure of the complex (Fig. 1b) indicates that the peptides form a six-helix bundle, with three C34 peptides packed against a coiled coil trimer of N36 peptides (Chan et al., 1997; Weissenhorn et al., 1997; Tan et al., 1997).

Current evidence suggests that the bundle structure corresponds to the fusion-active conformation of gp41. It shows strong similarity to the fusion-active cores of HA_2 (Bullough et al., 1994) and Mo-MLV (Fass et al., 1996). It folds in the absence of gp120, as expected by the observation that, in some isolates, shedding of gp120 accompanies the conversion from native to fusion-active state (Cohen 1996; Wilkinson 1996). Also, muta-

tions in gp41 which abolish infectivity and membrane fusion often map to residues which appear to be involved in the stabilization of the N36/C34 bundle (Dubay et al., 1992; Chen et al., 1993; Chen, 1994; Wild et al., 1994a; Poumbourios et al., 1997). Particularly interesting are the mutations Leu-568 -> Ala, Trp-571->Arg, and Asn-656->Leu because cells expressing mutant envelope glycoproteins with one of these point mutations are completely defective in membrane fusion, as judged by an inability to form syncytia with CD4-positive human lymphocyte lines, even though the mutant proteins exhibit substantial cell-surface expression, CD4 binding, gp120/gp41 association, gp160 precursor processing, and soluble CD4-induced shedding (Cao et al., 1993). Leu-568 and Trp-571 are N36 residues that line the edge of a deep cavity (Fig. 2a) important to the stability of the N36/C34 bundle (Chan et al., 1998). Asn-656 is in an **a** position of the C34 peptide and packs against the central N36 trimer. The locations of these mutations suggest that interactions between the N36 and C34 helices are critical for membrane fusion.

Another body of evidence, and one of particular relevance to this paper, supports the fusion-active hypothesis. Synthetic peptides containing approximately 40 residues from gp41 which overlap, or include all of, the residues in N36 and C34 can be effective inhibitors (at micromolar to nanomolar concentrations) of HIV infection and syncytia formation (Wild et al., 1992, 1994b; Jiang et al., 1993; Lu et al., 1995). Investigations into the mechanism of these inhibitory peptides implied that they work in a dominant-negative manner (Herskowitz, 1987) by binding to viral gp41 (Lu et al., 1995; Chen et al., 1995a). C-peptides with mutations that disrupt their interactions with the N-trimer also correlate with diminished potency as inhibitors (Wild et al., 1995).

The structure of the gp41 core provides a likely explanation of the dominant-negative inhibition. C-peptides may inhibit fusion by binding to the endogenous N-trimer within viral gp41 in place of the endogenous C-region. N-peptides might disrupt formation of the

endogenous N-trimer within viral gp41 or might compete against the N-trimer to bind endogenous viral C-peptides (Chan et al., 1997).

Both the N- and C-peptide classes of inhibitors are effective against a wide range of HIV strains (Wild et al., 1992; Jiang et al., 1993; Wild et al., 1994b). In contrast, soluble CD4 and many neutralizing antibodies are effective only on a limited subset of HIV strains (e.g. Nara et al., 1988; Palker et al., 1988; Daar et al., 1990; Moore et al., 1995). There is a strong conservation of residues involved in interactions between the N- and C-peptides among isolates and even between HIV and simian immunodeficiency virus (SIV). The broad inhibitory effects of the N- and C- peptides may derive from the sequence conservation of these residues.

The inhibitory potency of the C-peptide suggests the possibility of designing an anti-fusion drug against the endogenous N-trimer. One peptide inhibitor has already shown activity in clinical trials (Kilby et al., 1998). However, it is often preferable to develop small-molecule inhibitors, which can be orally administered, synthesized more easily, and are not as quickly degraded by enzymes in the blood. The structure of the N36/C34 complex displays at least one attractive drug target for small-molecule inhibitors-- a deep cavity at the C-terminus of the N36 trimer (Figs. 1,2). The residues lining the cavity are highly conserved among HIV-1 isolates, and, as mentioned earlier, two of these residues, Leu-568 and Trp-571, are necessary for membrane fusion. It has also been noted that certain C-peptide inhibitors which include the cavity region are less susceptible to escape mutants than inhibitors without this region (Rimsky et al., 1998). Finally, it has been shown that the potency of at least one C-peptide depends upon its cavity interactions. Previous research in our lab (Chan et al., 1998) has demonstrated the importance of the Trp-631 (Figs. 1,2) residue to the potency of the C34 peptide inhibitor. Through a series of amino acid substitutions for Trp-631, it was shown that the ability of the C-peptide to

inhibit membrane fusion (by IC50 assays for viral entry and syncytia formation) is directly related to the hydrophobic bulk of the substitution. As the size of the sidechain was reduced, the potency of the peptide dropped; W631A, for instance, was approximately 30-fold less active. The melting temperature of the six-helix bundle was also found to depend on the size of the substituted sidechain. The relationship between potency and stability suggested that the activity of the C-peptides depends on their ability to bind tightly to the cavity.

Any attempt at structure-based drug design requires atomic-level knowledge about the intended target. Information about the flexibility of target site residues, binding of alternate ligands, and the conformation of the unbound state can be useful in building a tight-binding ligand. In this paper, we investigate these properties of the cavity through the use of a point mutation (W631A) in the C34 ligand. Trp-631 is a bulky and central contact between C34 and the N-trimer cavity; the removal of its sidechain should allow for significant rearrangement of the cavity.

Results and Discussion

Crystals of N36/W631A were grown by hanging-drop vapor diffusion (see Experimental Procedures). Phases for the complex were estimated from a structure of the wild-type core of gp41 previously solved in our lab (Chan et al., 1997). Details of data collection and statistics are listed in Table 1. A representative portion of the solvent-flattened electron density map is shown in figure 3. The structure was refined to 1.9 Å against data from the crystal to give an R_{free} of 0.268 and an R_{cryst} of 0.213 (Table 1).

0.1 Structure of the N36/W631A mutant

The N36/W631A mutant overlaps well (rmsd 0.359 Å for alpha carbons of the N-pep-

tides) with the native N36/C34 complex, except near the cavity. Structural differences in the cavity (Fig. 2b,c) are significant. First, both Gln-577 and Gln-575 change rotamers and shift into the cavity, creating a ridge which splits the cavity in two. Secondly, Trp-571 changes its rotamer so as to lie outside the cavity. This shift removes hydrophobic bulk from the cavity.

The availability of space and the movement of polar sidechains into the cavity allows the binding of five buried, coordinated water molecules (Fig. 4). Gln-577 on the left wall and both Gly-572 and Gln-575 on the right wall bind a cluster of three of the waters-- W6, W3, and W43-- between them. Water W6 binds between the sidechain nitrogen of Gln-577 and the oxygen of Gly-572. W43 binds to the sidechain nitrogen of Gln-577 and to W3. W3 binds to W43 and to the sidechain nitrogen of Gln-575. The sidechain oxygen of Gln-575 also hydrogen bonds to the sidechain nitrogen of Trp-628 (on the W631A ligand).

A B-factor analysis (Fig. 4) implies only partial occupancy for W43 (B-factor of 53.89 compared to 25.98 for Gln-577 NE2 and 29.60 for W3). However, B-factors for W3 and W6 are comparable to their bonding partners, indicating full occupancy. All bonding distances are within the allowable range, and the only bonding angle of interest, W6 -> Gln-577 -> W43, measures 120.10 as expected.

Two other waters bind slightly higher in the cavity. W44 binds to the backbone oxygen of Leu-568 on the right wall. W81 binds to the backbone oxygen of Ala-631 in this region. B-factors indicate that both waters show only partial occupancy. Again, binding distances are within the allowable range.

There are also slight movements to several residues (Leu-565, Arg-579, and Leu-581) which form the top and bottom edges of the cavity (Fig. 2). Arg-579 clearly changes its rotamer. Leu-581 retains its original rotamer but is displaced approximately 1.5 Å from

the HIV structure due to gradual deviations in the backbone. These deviations are most likely associated with the downward displacement of Trp-628 in the mutant. Leu-565 also shifts slightly, with a deviation in its gamma carbon of about 1.7 Å versus the native structure. In addition, the dihedral angle of Leu-565 is reduced nearly 60^0. It should be noted, however, that Arg-579 and Leu-581 are located at the very base of the structure and pack against neighboring complexes in the crystal. Dissimilar crystal contacts could explain the changes.

0.2 Implication for Drug Discovery

The deep cavity at the base of the N36/C34 bundle alters substantially when part of its natural ligand, the sidechain of Trp-631, is removed. In the absence of that hydrophobic bulk, the cavity fills with water, and several flexible sidechains (Trp-571, Gln-575, Gln-577) move to create a network of hydrogen bonds. This rearrangement suggests that the conformations of these sidechains should be allowed to vary in any drug search directed at the N-trimer cavity.

Experimental Procedures

0.3 Peptide Purfication and Crystallization

Peptides N36 and W631A were synthesized by standard FMOC peptide chemistry and have an acetylated N-terminus and a C-terminal amide. N36 corresponds to residues 546-581 of gp160 (HXB2 strain), while W631A corresponds to residues 628-661, with the substitution of alanine for tryptophan at position 631. Peptides were desalted on a Sephadex G-25 column (Pharmacia), lyophilized, and then purified by reverse-phase liquid chromatography (Waters, Inc.) on a Vydac C18 preparative column. The identity of the peptides was verified by mass spectrometry. Peptide concentrations were determined by tyrosine and tryptophan absorption in 6M GuHCl (Edelhoch, 1967).

To grow crystals, a 10.8 mg/ml stock of the N36/W631A complex was diluted 1:1 in a hanging drop and allowed to equilibrate against 30% PEG4000, 0.1M $NaCH_3COO$ pH 4.6 and 0.2M NH_4CH_3COO. Crystals grew as rhombohedral prisms, space group R3 (a=b=c=34.5 Å, alpha=beta=gamma=78.6^0). Crystals were flash frozen in a MSC cryogenic crystal cooler (X-stream), and data was collected at the Howard Hughes Medical Institute beamline X4A of the National Synchroton Light Source at Brookhaven National Laboratory, at a wavelength of 1.1197 Å. Reflections were integrated and scaled with DENZO and SCALEPACK (Otwinowski, 1993).

Data scaling, phase determination, and map generation were all performed using the CCP4 suite of programs (CCP4, 1994). Phases were calculated with AMORE using the N36/C34 structure (Chan, 1997) as a model.

0.4 Model refinement

The polypeptide chain was traced and the side chains positioned into a 1.9Å density-modified map using the program O (Jones and Kjeldgaard, 1992). The initial model of N36/W631A was refined with the program XPLOR (Brunger, 1992a) against data to 1.9Å. A free R set (Brunger, 1992b) was taken from the data prior to refinement. The model was refined by iterative cycles of grouped B-factor and positional refinement. As the refinement preceded, 245 waters were added and a bulk solvent correction was applied. Main chain and side chain geometries were optimized in O between cycles of refinement. The quality of the structure was checked by WHATIF.

Acknowledgments

I wish to thank Peter S. Kim for reviewing this thesis and for welcoming me into his lab despite my meager background in biology. I would also like to thank Vladimir Malashkevich for his instruction in data collection and analysis, Michael Burgess and James Pang

for peptide synthesis, and David Lee for constant help and useful discussions.

12

References

Allan, J.S., Strauss, J., and Buck, D.W. (1990). Enhancement of SIV infection with soluble receptor molecules. Science *247*, 1084-1088.

Broder, C.C., and Dimitrov, D.S. (1996). HIV and the 7-transmembrane domain receptors. Pathobiology *64*, 171-179.

Brunger, A.T. (1992a). A system for X-ray crystallography and NMR. X-PLOR Version 3.1. (New Haven, Connecticut: Yale University Press).

Brunger, A.T. (1992b). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

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

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

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

CCP4 (1994). Collaborative computational project, Number 4. Acta Cryst. *D50*, 760-763.

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

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

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

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

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

Chen, S.S. (1994). Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane glycoprotein gp41. J. Virol. *68*, 2002-2010.

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

Cohen, J. (1996). Investigators detail HIV's fatal handshake. Science *274*, 502.

D'Souza, M.P., and Harden, V.A. (1996). Chemokines and HIV-1 second receptors. Nature Med. *2*, 1293-1300.

Daar, E.S., Li, X.L., Moudgil, T., Ho, D.D. (1990). High concentrations of recombinant soluble CD4 are required to neutralize primary human immunodeficiency virus type 1 isolates. Proc. Natl. Acad. Sci. U.S.A. *87*, 6574-6578.

Delwart, E.L., Moialos, G., and Gilmore, T. (1990). Retroviral envelope glycoproteins contain a leucine zipper-like repeat. AIDS Res. Hum. Retroviruses *6*, 703-706.

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

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

Fass, D., Harrison, S.C., and Kim, P.S. (1996). Structure of Moloney murine virus envelope domain at 1.7 A resolution. Nature Struct. Biol. *3*, 465-469.

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

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

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

Jones, T.A., and Kjeldgaard, M. (1992). O- The Manual (Uppsala, Sweden: http://kaktus.kemi.aau.dk).

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

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

Moore, J.P., Cao, Y., Qing, L., Sattentau, Q.J., Pyati, J., Koduri, R., Robinson, J., Barbas, C.F., III, Burton, D.R., and Ho, D.D. (1995). Primary isolates of human immunodeficiency virus type 1 are relatively resistant to neutralization by monoclonal antibodies to gp120, and their neutralization is not predicted by studies with monomeric gp120. J. Virol. *69*, 101-109.

Nara, P.L., Robey, W.G., Pyle, S.W., Hatch, W.C., Dunlop, N.M., Bess, J.W., Jr., Kelliher, J.C., Arthur, L.O., and Fischinger, P.J. (1988). Purified envelope glycoproteins from human immunodeficiency virus type 1 variants induce individual, type-specific neutralizing antibodies. J. Virol. *62*, 2622-2628.

Otwinowski, Z. (1993). Oscillation data reduction program. In Data Collection and Processing. L. Sawyer, N. Isaacs, S. Bailey, eds. (Warrington, England; SERC, Daresbury Laboratory) p. 56-62.

Palker, T.J., Clark, M.E., Langlois, A.J., Matthews, T.J., Weinhold, K.J., Randall, R.R., Bolognesi, D.P., and Haynes, B.F. (1988). Type-specific neutralization of the human immunodeficiency virus with antibodies to env-encoded synthetic peptides. Proc. Natl. Acad. Sci. U.S.A. *85*, 1932-1936.

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

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

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

Sattentau, Q.J., and Moore, J.P. (1993). The role of CD4 in HIV binding and entry. Phil. Trans. Royal Soc. B *342*, 59-66.

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

Sullivan, N., Sun, Y., Li, J., Hofmann, W., and Sodroski, J. (1995). Replicative function and neutralization sensitivity of envelope glycoproteins from primary and T-cell line-passaged human-immunodeficiency virus type 1 isolates. J. Virol. *69*, 4413-4422.

Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997). Atomic structure of a thermostable

subdomain of HIV-1 gp41. Proc. Natl. Acad. Sci. U.S.A. *94*, 12303-12308.

Tsurudome, M., Gluck, R., Graf, R., Falchetto, R., Schaller, U., and Brunner, J. (1992). Lipid interactions of the hemagglutinin HA_2 NH_2-terminal segment during influenza virus-induced membrane fusion. J. Biol. Chem. *267*, 20225-20232.

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

Wild, C.T., Oas, T., McDanal, C.B., Bolognesi, D., and Matthews, T.J. (1992). A synthetic peptide inhibitor of human immunodeficiency virus replication: correlation between solution structure and viral inhibition. Proc. Natl. Acad. Sci. U.S.A. *89*, 10537-10541.

Wild, C., Dubay, J.W., Greenwell, T., Baird, T., Jr., Oas, T.G., McDanal, C., Hunter, E., and Matthews, T. (1994a). Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. Proc. Natl. Acad. Sci. U.S.A. *91*, 12676-12680.

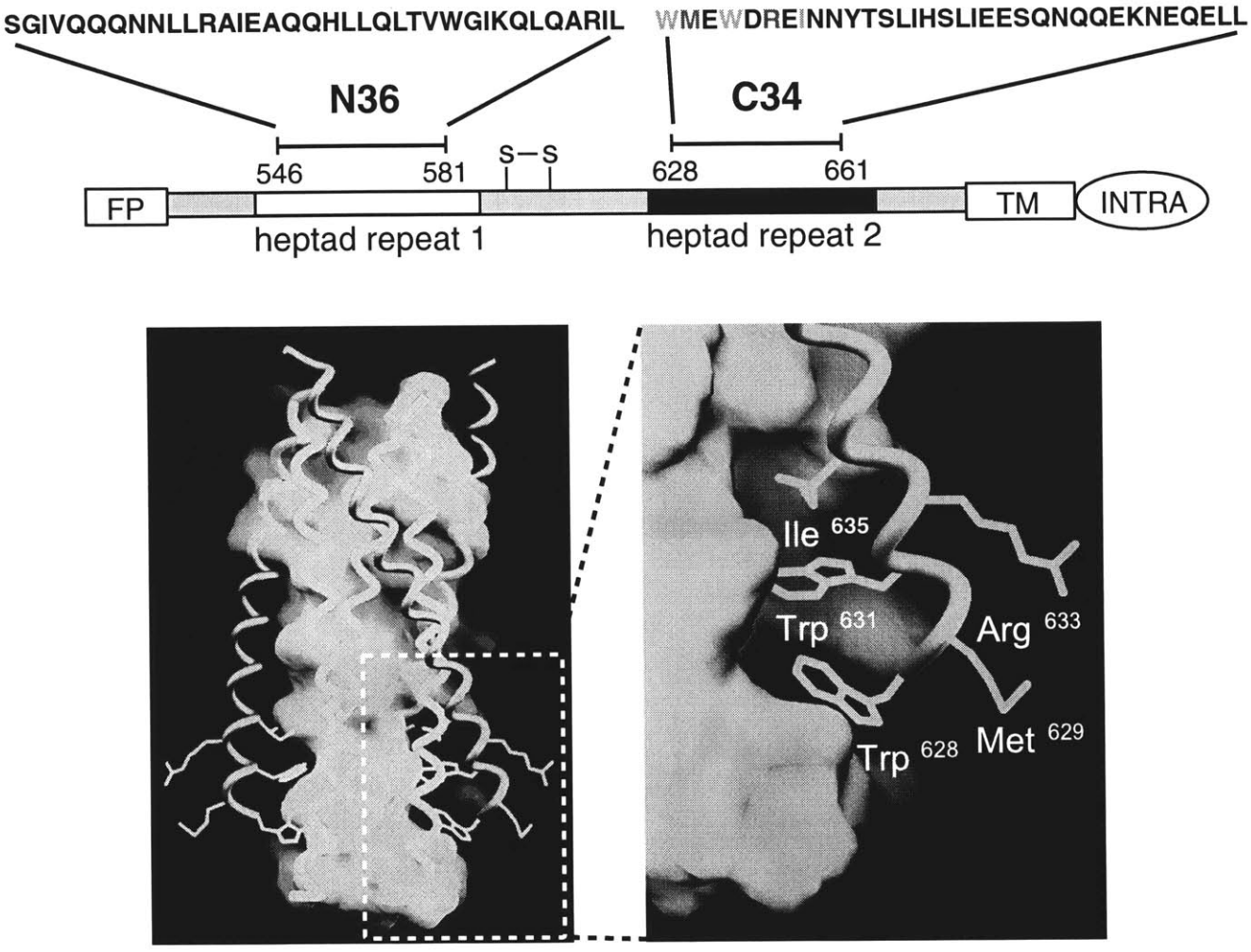
Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B., and Matthews, T.J. (1994b). Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc. Natl. Acad. Sci. U.S.A. *91*, 9770-9774.

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

Wilkinson, D. (1996). HIV-1: cofactors provide the entry keys. Curr. Biol. *6*, 1051-1053.

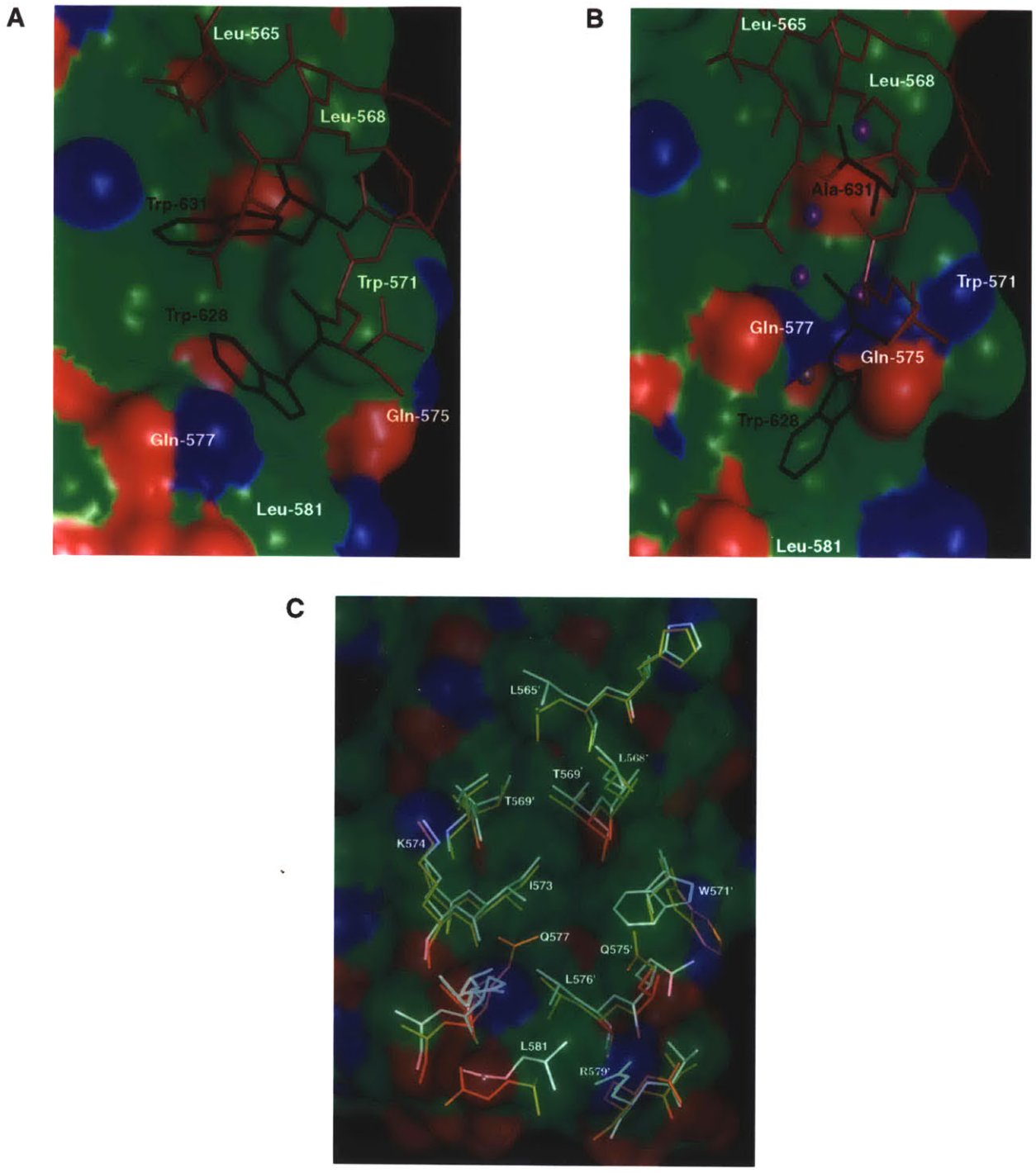
Wilson, I.A., Skehel, J.J., and Wiley, D.C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature *289*, 366-373.

Figure 1: HIV-1 gp41 Core Structure



(A) Schematic of HIV-1 gp41 showing the N36 and C34 peptides, located within two regions containing 4,3 hydrophobic heptad repeats (white and black boxes). Residues in light gray project into the N36 cavity, whereas residues in dark grey do not. FP, fusion peptide; S-S disulfide bond; TM, transmembrane region; INTRA, intraviral region. **(B)** The N36/C34 crystal structure of the HIV-1 gp41 ectodomain core. The trimeric N36 coiled coil is represented by the three central helices overlaid with a semitransparent molecular surface *(Left)* . Three C34 helices (shown with selected side chains) pack against this coiled coil surface. The bottom of the N36 surface contains three symmetry-related cavities (one is outlined by the box), each of which accommodates three hydrophobic residues (Ile, and two Trp residues) from a C34 helix. In contrast, Arg and Met project outwards and do not make contacts with the coiled coil.

Figure 2: Comparison of Native and Mutant Cavities



(A) N36/C34 cavity indicating the location of important cavity sidechains (white) and ligand sidechains (black). Carbon is green, oxygen red, and nitrogen blue. Colors do not indicate electrostatic potential and are meant only as structural cues. **(B)** N36/W631A cavity. Gln-577 and Gln-575 form a ridge across the cavity, and Trp-571 spins ~120 degrees about its dihedral angle. Five buried waters (purple) hydrogen bond in the cavity. **(C)** Superposition of cavity sidechains for wild-type (white) and mutant (yellow) structures.

Table 1: Data Collection and Refinement Statistics

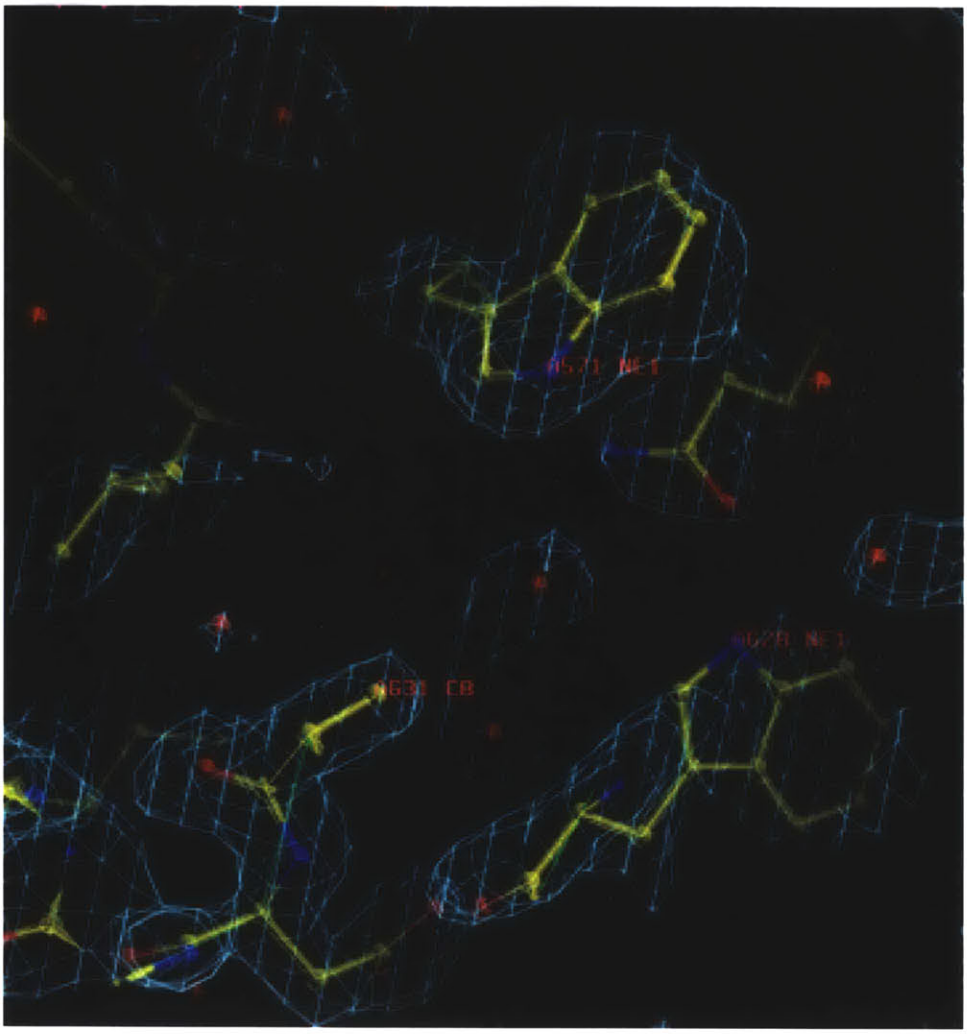
Data Collection

Highest resolution	1.9 Å
Observed reflections	12,666
Unique reflections	3,716
Completeness, %	91.7
R_{merge}	0.040
R_{merge} at 1.9 Å	0.112

Refinement

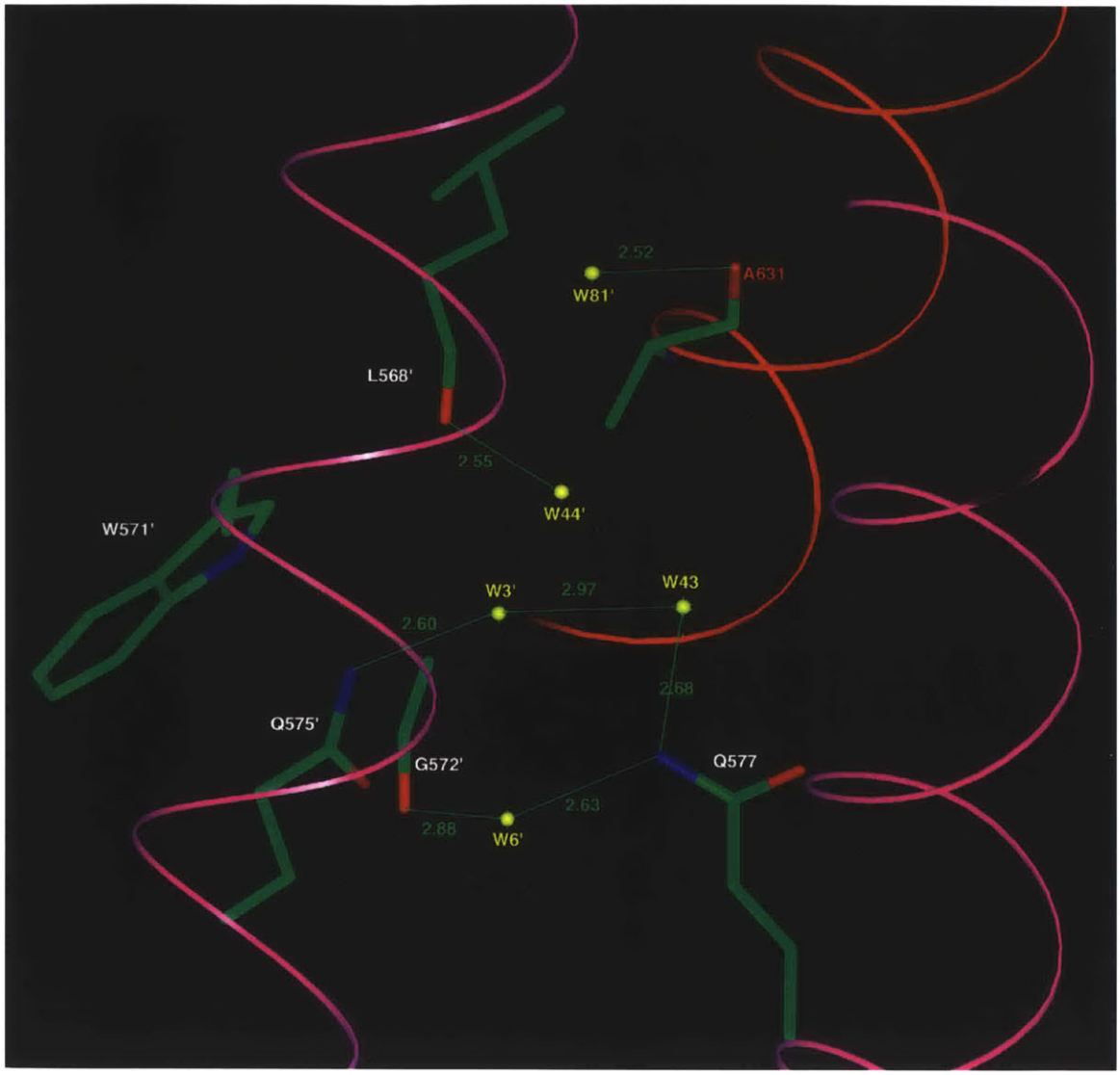
Resolution, Å	20.0-1.9
Protein nonhydrogen atoms	1,761
Water molecules	245
R_{crys} / R_{crys} at 1.9 Å	0.213 / 0.242
R_{free} / R_{free} at 1.9 Å	0.268 / 0.322
Average B-factor, $Å^2$	23.1
Rmsd from ideal geometry	
Bond length, Å	0.011
Bond angles	1.34^0
Torsion angles	16.0^0

Figure 3: Experimental Electron Density Map



A representative portion of the electron density map calculated from experimental structure-factor amplitudes and estimated phases (see Experimental Procedures) is shown with the refined molecular model. The map is contoured to 1.0 standard deviations above the mean density. Waters are shown as small red spheres. The figure was generated with the program O (Jones and Kjeldgaard, 1992).

Figure 4: Hydrogen Bonding of Buried Waters



Atom	B-factor	Atom	B-factor	Atom	B-factor
G572 O	10.85	L568 O	14.58	A631 O	27.60
Q577 NE2	25.98	W44	31.76	W81	41.93
W6	19.91				
W3	29.60				
W43	53.89				

Hydrogen bonding pattern and B-factors for buried waters. View is from the center of the N-trimer. Purple helices correspond to the N36 peptides, and red helix corresponds to the C34 peptide. Waters are shown in yellow.