

Matthew Engel
 Biology of HIV-1 Retrovirus
 BME 501

Human immunodeficiency virus (HIV) is a growing epidemic of global proportions. In developing countries, the population most at risk includes young people ages 10 to 24 which specifically includes sex workers, injection drug users and men who have sex with men . Public health interventions such as HIV community education centers, school-based sex education programs, and peer advisory groups have proven most effective in reducing the risk of transmission among target populations and should be implemented more frequently . Mass media intervention has shown efficacy in prevention demonstrated by published studies. The most common outcomes included condom use, demonstrated knowledge of transmission routes, reduction of high-sexual behaviors, increased interpersonal communication about HIV, and abstinence .

In 2005, between 1-3 million people were taking anti-retroviral medication for HIV . Globally, 38.6 million people live with HIV-1 while 25 million lives have already been claimed . There has been nine separate HIV-1 serotypes classified as the result of genetic recombination and selective pressures while it's continuing diversity make an all-encompassing treatment unlikely. However, on a molecular scale, knowledge of HIV structure, fusion, reverse transcription, integration, expression and post-translational processing are increasing. Each of these sectors represent potentials site for antibody, peptide or small-molecule inhibition.

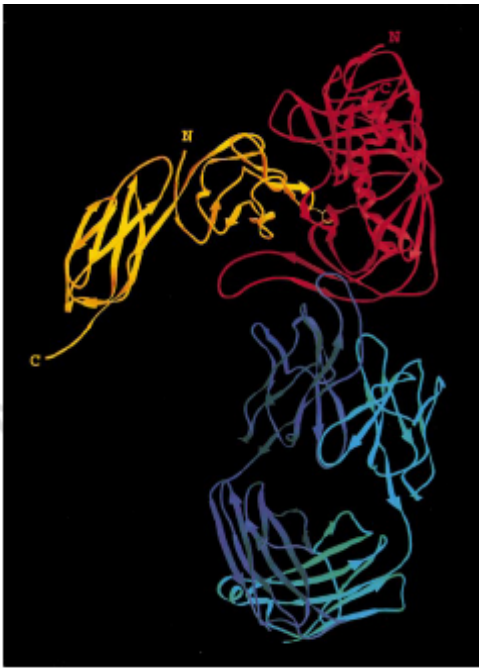


Figure 1. Structure of HIV gp120 (red), N-terminal CD4 domain (yellow) and Fab17 human antibody (blue). Notice how Fab17 interaction with gp120 do not inhibit CD4 binding.

Kwong, P. D., et al. (1998), *Nature* 393, 48-659.

Retroviruses, such as HIV, contain genomic RNA but are replicated by an integrated DNA complex. The retrovirus ~9kb genome contains a promoter at the 3' end which drives the expression of reverse transcriptase directly from this material. During reverse transcription the promoter becomes juxtaposed upstream of the coding sequences, only then resembling the classical gene structure: promoter→ CDS arrangement . The viral genes *gag*, *pol*, *pro*, and *env* encode the major retroviral proteins. *Gag* encodes a full-length protein which is proteolytically cleaved into mature protein (MA), capsid (CA), and the nucleocapsid (NC). *Pol* encodes for reverse transcriptase and integrase enzymes.

HIV-1 infection is initiated by high-affinity binding of envelope proteins (ENV) to the human CD4 receptor. A pivotal coreceptor in the immune-pathway, CD4 stimulates association between T-cell receptors and major histocompatibility complex II which presents antigens on the surface of infected cells. The CD4 extracellular domain exhibits dimeric properties

which may behave in a hinge-like manner important for immune recognition and HIV binding. Deletion of the flexible portion linking this soluble domain to the integral membrane prevents HIV mediated fusion. CD4⁺ lymphocytes are efficiently destroyed by HIV-1 and HIV-2 leading to acquired immunodeficiency syndrome (AIDS). HIV entry is dependent on interactions between the CD4 glycoprotein and gp120 – an exterior viral envelope glycoprotein and immunological target during infection. High resolution crystal structures of this antigen-antibody-receptor complex (figure 1, see Kwong et al. [1998]) reveals seven disulphide bridges in gp120 conserved among HIV strains. Gp120 is folded into two parallel domains (inner/outer) and composed of five variable regions (V1-V5) secured to the viral membrane by gp41, a transmembrane envelope glycoprotein. The peripheral region of this complex forms a spike, composed primarily of gp120 affixed by non-covalent interactions to gp41. This complex is utilized in direct membrane penetration and elicits high titers of ineffectual antibodies throughout infection. Upon binding to CD4, a conformational change arises in gp120 exposing sites for further coreceptor attachment. This rearrangement was discovered during experiments in which exposure of gp120 to CD4 improved antibody adherence and prevented gp120-CD4 aggregates from attaching to their coreceptors.

Gp120 displays only 35% amino acid sequence homology between HIV-2 and HIV-1 serotypes. Within the HIV-1 clade itself, different subtypes retain marked resemblance displaying up to 77% homology between HIV-1 C and HIV-1 HXBc2 strains. In particular, the CD4 binding site displays excellent conservation between species and has become an attractive anti-viral target. These drugs prevent gp120 from assuming conformations that favor CD4 interaction. Preclinical trials suggest gp120 microbicide would act positively in a cocktail with inhibitors of concomitant mechanisms.

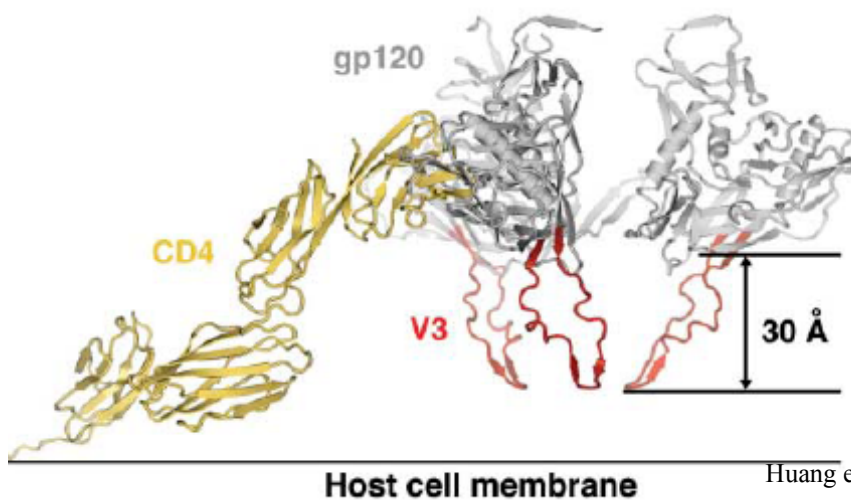


Figure 2. The structure of CD4 complexed with gp120 and the third variable region, V3. V3 is shown in a trimeric state positioned toward the coreceptor on the host cell membrane. This domain is the primary location of antigenic activity and conservation of the V3 tip plays a pivotal role of successful infection.

Huang et al. *Science* 310, 1025-1028. 2005.

The trimeric envelope spike responsible for HIV entry consists of three external gp120 sub-units and three gp41 transmembrane domains. Further coreceptors CCR5 or CXCR4 are essential for infection and interact with the third variable gp120 region, V3. This 35 amino acid region will determine which coreceptor is allocated for entry and is targeted immunodominantly by the humoral response during vaccination. Crystallization

experiments of V3 and gp120 illustrate small differences in the N-terminus and flexible loops of gp120 which are the subject of profuse variation.

V3 emanates from the outer domain of gp120 protruding 30 Å from its root, and is only 15 Å wide (figure 2 [Huang et al. 2005]). The base is highly conserved among species and is linked to a β -hairpin tip by a flexible stem. Interaction of V3 with the X5 antibody caused a 17 Å shift of the X5 α -carbon backbone, one of the most dramatic conformational changes found in antibody binding. Human immunization with gp120 or gp41/gp120 evokes a humoral response directed almost completely at V3. Further studies of V3 peptide-antibody complexes show varying conformations, however the Pro-Gly tip of V3 is conserved. Superposition of the native V3 tip and the peptide-antibody-gp120 complexes indicate that V3 becomes completely surrounded by these immunoglobulins .

A major roadblock to the development of an effective HIV-1 vaccine is that systemic antibodies secreted by humans are nonneutralizing or easily countered by mutations. Only five neutralizing antibodies have been isolated: two are directed against the outer domain of gp120 and the other three target the gp41 transmembrane domain. 2G12 recognizes the gp120 epitope and owes its success to domain swapping, bypassing the carbohydrate cloak .

Once receptor recognition and envelope fusion has occurred, HIV-1 uncoats during entry and takes advantage of the host cell machinery translating reverse transcriptase (RT), a major enzyme in the viral pathway. Highly active antiretroviral therapy contains inhibitors which always includes a small molecule to prevent the function of this enzyme. Composed of two subunits, p66 and p51, HIV-RT exhibits RNase, DNA-dependent and RNA-dependent DNA polymerase activities. P66 contains the active sites for polymerase and RNase governing the conversion of viral ssRNA to dsDNA. The preference of HIV-1 for terminally differentiated CD45RO⁺ memory T lymphocytes over immature CD45RA⁺ lymphocytes is due to HIV's inability to complete reverse transcription in the latter cell type, despite its ability to penetrate the host membrane .

Nucleoside reverse transcriptase inhibitors (NRTIs), such as AZT – the first anti-HIV drug, are incorporated into growing proviral DNA strands terminating reverse transcription. NRTI's are similar in structure to deoxynucleotides (dNTPs) but lack the 3' OH group necessary for chain elongation. Resistant strands of HIV exhibiting RT mutations demonstrate a selective advantage over the original genotype. One method of resistance is conferred by the branched β -chains of amino acid substitutions (Val, Ile, Thr) at position 184, replacing methionine. The functional groups of these residues sterically collide with the β -L-oxathialone ring of 3TCTP, a NRTI, preventing the drug's incorporation into the polymer chain. There is no steric conflict with naturally occurring dNTP's .

Nonnucleoside reverse transcriptase inhibitors (NNRTIs) bind to a hydrophobic pocket approximately 15 Å from the NRTI site. The inhibition is caused by a global conformational change in the RT rendering the enzyme inactive. Because different mutations in RT circumvent positive drug interactions differently, there is a constant need for novel medicines. Often, computational simulations are successful in identifying favorable new models for inhibition based on a general drug scaffold. Binding affinity is calculated by a change in free energy so that many new functional groups can be screened without actually synthesizing them. Free energy perturbation and

thermodynamic integration methods rely on molecular mechanics to generate data describing efficacy of new analogues. These simulations can illustrate how particular hydrogen bonds may react to the ligand in an aqueous solution. The water bridges between NNRTIs and HIV protease are particularly important .

Following integration of the HIV provirus into the host DNA, expression of viral mRNA transcripts proceeds. However, the provirus can remain dormant for extended periods before transcription occurs. After the host DNA polymerase transcribes these genes, the transcripts are transported out into the cytosol. Here, the ribosomes translate viral polypeptides which must undergo cleave to become activated. This non-trivial step is accomplished by the homodimeric HIV protease . Its method of substrate/sequence recognition is still relatively unknown. Processed, mature viral proteins accumulate near the cell membrane in preparation for fusion. Finally, viruses are released from the cell by budding or lysis.

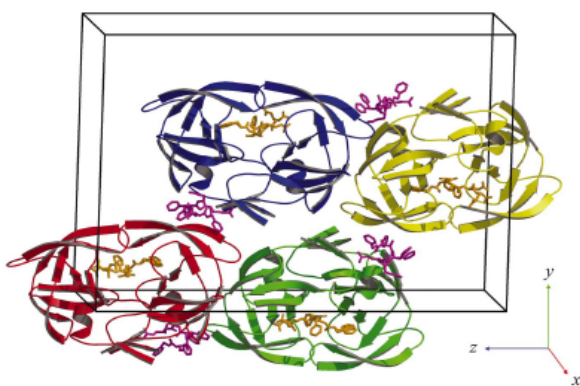


Figure 3
Crystal packing. One unit cell is shown; the orientation of its axes is marked at the bottom right. Protein molecules are represented by ribbon models coloured red for the molecule in the asymmetric part of the unit cell and coloured green, blue, and yellow for the first, second and third symmetry-related molecules, respectively. Stick models, coloured orange for the molecules bound in the active site and magenta for the molecules contacting the outer protein surface, represent the inhibitor molecules.

Brynda, P. R et al. (2004) *Acta Crystallographica* Section D: Biological Crystallography D60, 1943-1948.

Retroviral protease has been extensively studied by X-ray crystallography. For example, after an inhibitor passes a pharmaceutical assay it is often cocrystallized with the protein target in an attempt to understand its mechanism. This can be accomplished by soaking the crystals in a ligand solution or adding the inhibitors directly to the crystallization reagents. Figure 3 (Brynda et al., 2004) demonstrates the arrangement of HIV protease in a crystal lattice bound to two different inhibitors . One peptide can be found within the active site and another as an outer ligand.

The forthcoming patent will address screening methods of molecular libraries for viral protein inhibitors. The assays under development will utilize

existing high energy synchrotron radiation at the National Synchrotron Light Source (NSLS) to search for small inorganic molecules or proteins which initiate a global conformation change in their substrate. The said methods will include synchrotron radiation circular dichroism and small-angle X-ray scattering.

1. Hoffmann, O., Boler, T., and Dick, B. (2006) Achieving the global goals on HIV among young people most at risk in developing countries: young sex workers, injecting drug users and men who have sex with men, *World Health Organization technical report series 938*, 287-315; discussion 317-241.

2. Kirby, D., Obasi, A., and Laris, B. A. (2006) The effectiveness of sex education and HIV education interventions in schools in developing countries, *World Health Organization technical report series 938*, 103-150; discussion 317-141.
3. Bertrand, J. T., O'Reilly, K., Denison, J., Anhang, R., and Sweat, M. (2006) Systematic review of the effectiveness of mass communication programs to change HIV/AIDS-related behaviors in developing countries, *Health education research 21*, 567-597.
4. Schwartlander, B., Grubb, I., and Perriens, J. (2006) The 10-year struggle to provide antiretroviral treatment to people with HIV in the developing world, *Lancet 368*, 541-546.
5. Simon, V., Ho, D. D., and Abdool Karim, Q. (2006) HIV/AIDS epidemiology, pathogenesis, prevention, and treatment, *Lancet 368*, 489-504.
6. John Coffin, S. H., Harold Varmus (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press.
7. Wu, H., Kwong, P. D., and Hendrickson, W. A. (1997) Dimeric association and segmental variability in the structure of human CD4, *Nature 387*, 527-530.
8. Kwong, P. D., Doyle, M. L., Casper, D. J., Cicala, C., Leavitt, S. A., Majeed, S., Steenbeke, T. D., Venturi, M., Chaiken, I., Fung, M., Katinger, H., Parren, P. W., Robinson, J., Van Ryk, D., Wang, L., Burton, D. R., Freire, E., Wyatt, R., Sodroski, J., Hendrickson, W. A., and Arthos, J. (2002) HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites, *Nature 420*, 678-682.
9. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody, *Nature 393*, 648-659.
10. Kadow, J., Wang, H. G., and Lin, P. F. (2006) Small-molecule HIV-1 gp120 inhibitors to prevent HIV-1 entry: an emerging opportunity for drug development, *Curr Opin Investig Drugs 7*, 721-726.
11. Wang, T., Zhang, Z., Wallace, O. B., Deshpande, M., Fang, H., Yang, Z., Zadjura, L. M., Tweedie, D. L., Huang, S., Zhao, F., Ranadive, S., Robinson, B. S., Gong, Y. F., Ricarrdi, K., Spicer, T. P., Deminie, C., Rose, R., Wang, H. G., Blair, W. S., Shi, P. Y., Lin, P. F., Colonno, R. J., and Meanwell, N. A. (2003) Discovery of 4-benzoyl-1-[(4-methoxy-1H-pyrrolo[2,3-b]pyridin-3-yl)oxoacetyl]-2-(R)-methylpiperazine (BMS-378806): a novel HIV-1 attachment inhibitor that interferes with CD4-gp120 interactions, *Journal of medicinal chemistry 46*, 4236-4239.
12. Meanwell, N. A., and Kadow, J. F. (2003) Inhibitors of the entry of HIV into host cells, *Current opinion in drug discovery & development 6*, 451-461.
13. Dorfman, T., Moore, M. J., Guth, A. C., Choe, H., and Farzan, M. (2006) A Tyrosine-sulfated Peptide Derived from the Heavy-chain CDR3 Region of an HIV-1-neutralizing Antibody Binds gp120 and Inhibits HIV-1 Infection, *The Journal of biological chemistry 281*, 28529-28535.
14. Huang, C. C., Tang, M., Zhang, M. Y., Majeed, S., Montabana, E., Stanfield, R. L., Dimitrov, D. S., Korber, B., Sodroski, J., Wilson, I. A., Wyatt, R., and Kwong, P. D. (2005) Structure of a V3-containing HIV-1 gp120 core, *Science 310*, 1025-1028.
15. Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J. R., Wyatt, R., and Kwong, P. D. (2004) Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope, *Journal of virology 78*, 10724-10737.
16. Lawn, S. D., Butera, S. T., and Folks, T. M. (2001) Contribution of immune activation to the pathogenesis and transmission of human immunodeficiency virus type 1 infection, *Clinical microbiology reviews 14*, 753-777, table of contents.
17. Tuske, S., Sarafianos, S. G., Clark, A. D., Jr., Ding, J., Naeger, L. K., White, K. L., Miller, M. D., Gibbs, C. S., Boyer, P. L., Clark, P., Wang, G., Gaffney, B. L., Jones, R. A., Jerina, D. M., Hughes, S. H., and Arnold, E. (2004) Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir, *Nature structural & molecular biology 11*, 469-474.

18. Rizzo, R. C., Udier-Blagovic, M., Wang, D. P., Watkins, E. K., Kroeger Smith, M. B., Smith, R. H., Jr., Tirado-Rives, J., and Jorgensen, W. L. (2002) Prediction of activity for nonnucleoside inhibitors with HIV-1 reverse transcriptase based on Monte Carlo simulations, *Journal of medicinal chemistry* 45, 2970-2987.
19. Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., and Schiffer, C. A. (2003) Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy, *Journal of virology* 77, 1306-1315.
20. Jiri Brynda, P. R., Milan Fabry, M., Horejsi, R. S., Milan Soucek, M., Hradilek, J. K. a., and Sedlacek, J. (2004) Inhibitor binding at the protein interface in crystals of a HIV-1 protease complex, *Acta Crystallographica Section D: Biological Crystallography* D60, 1943-1948.