The Structural Biology of HIV

HIV (human immunodeficiency virus) is composed of two strands of RNA, 15 types of viral proteins, and a few proteins from the last host cell it infected, all surrounded by a lipid bilayer membrane. Together, these molecules allow the virus to infect cells of the immune system and force them to build new copies of the virus. Each molecule in the virus plays a role in this process, from the first steps of viral attachment to the final process of budding.

25 years of research on the structural biology of HIV have revealed the atomic details of these proteins. These structures are all publicly available in the Protein Data Bank (PDB) archive. Using these data, researchers have designed new treatments for HIV infection, including effective drug regimens that halt the growth of the virus. The structures also provide new hope for development of a vaccine.

Viral Enzymes

RT: Reverse transcriptase builds a DNA copy of the viral RNA genome, which is then used to build new viruses. This structure captures the enzyme as it is building a DNA strand (red) from the viral RNA (yellow). It will then destroy the RNA and build a second DNA strand. Many of the drugs currently used to fight HIV infection block the action of reverse transcriptase. PDB entry 1hys.

IN: Integrase takes the DNA copy of the viral genome and inserts it into the infected cellular genome. In this way, HIV can lie dormant in cells for decades, making it incredibly difficult to fight. Anti-HIV drugs that block integrase have been developed. PDB entry 1ex4.

PR: HIV protease is essential for the maturation of HIV particles. The proteins in HIV are built as long polyproteins, which then must be cleaved into the proper functional pieces by HIV protease. Protease inhibitors are widely used as anti-HIV drugs, often in combination with drugs that block reverse transcriptase and integrase. PDB entry 1hpv.

Structural Proteins

MA: Matrix protein forms a coat on the inner surface of the viral membrane. It plays a central role when new viruses bud from the surface of infected cells. This protein assembles into trimers, which then associate side-by-side on the membrane. PDB entry 1hiw.

CA: Capsid protein forms a cone-shaped coat around the viral RNA, delivering it into the cell during infection. It forms stable hexamers, which then assemble like tiles to form geodesic capsids. PDB entry 3h47.

Accessory Proteins

Vpu (viral protein u) helps the virus escape the cell during budding by weakening the interaction of the new envelope proteins with cell receptors. It also forms an ion channel in the viral membrane. PDB entries 1pi7 and 1vpu.

Vif (viral infectivity factor) attacks one of the cell's defense proteins, which forces the cell to destroy it. Only a small portion of Vif (green) is shown in this structure, bound to proteins from the infected cell (purple). PDB entry 3dcg.

Vpr (viral protein r) guides the viral genome into the nucleus following infection. PDB entry 1esx.

P6 is involved in the incorporation of Vpr into new viruses. It is largely unstructured and there is currently no structure for it in the PDB.

SU and **TM**: Envelope proteins gp120 and gp41 bind to receptors on the surface of cells that HIV infects, and then penetrate the surface to infect it with the viral RNA. The spikes formed by these proteins are highly decorated with carbohydrates, making them difficult to recognize by antibodies. The structures shown here include the portion outside the virus, and have all of the carbohydrates removed. PDB entries 1g9m (SU, top) and 2ezo (TM, bottom).

NC: Nucleocapsid protein forms a stable complex with the viral RNA, protecting it. In this structure, a short piece of RNA (yellow) is bound to one copy of nucleocapsid (orange). PDB entry 1a1t.

Nef (negative regulatory factor) forces the infected cell to stop making several proteins that are important in cell defense. Nef is important in the progression of HIV infection to Acquired Immune Deficiency Syndrome (AIDS). PDB entries 1avv and 1qa5.

Rev (regulator of virion) protein binds to a hairpin in the viral RNA and regulates the splicing and transport of viral RNA. The structure shown here includes only the portion of the protein that is bound to the RNA–the whole protein is several times larger. PDB entry 1etf.

Tat (trans-activator of transcription) protein binds to a hairpin in the viral RNA and greatly enhances the amount of protein that is made. PDB entries 1biv and 1jfw.

RCSB PDB
PROTEIN DATA BANK
www.pdb.org • info@rcsb.org

Structure References

VIRAL ENZYMES: 1hys: S. G. Sarafianos, K. Das, C. Tantillo, A. D. Clark Jr., J. Ding, J. Whitcomb, P. L. Boyer, S. H. Hughes, E. Arnold (2001) Crystal structre of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J* **20**: 1449-1461. 1ex4: J. C. Chen, J. Krucinski, L. J. Miercke, J. S. Finer-Moore, A. H. Tang, A. D. Leavitt, R. M. Stroud (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc Natl Acad Sci USA* **97**: 8233-8238. 1hpv: E. E. Kim, C. T. Baker, M. D. Dwyer, M. A. Murcko, B. G. Rao, R. D. Tung, M. A. Navia. (1995) Crystal structure of HIV-1 protease in complex with Vx-478, a potent and orally bioavailable inhibitor of the enzyme. *J Am Chem Soc* **117**: 1181-1182.

STRUCTURAL PROTEINS: 1hiw: C. P. Hill, D. Worthylake, D. P. Bancroft, A. M. Christensen, W. I. Sundquist (1996) Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proc Natl Acad Sci USA* **93**: 3099-3104. 3h47: O. Pornillos, B. K. Ganser-Pornillos, B. N. Kelly, Y. Hua, F. G. Whitby, C. D. Stout, W. I. Sundquist, C. P. Hill, M. Yeager (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137**: 1282-1292. 1g9m: P. D. Kwong, R. Wyatt, S. Majeed, J. Robinson, R. W. Sweet, J. Sodroski, W. A. Hendrickson (2000) Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure* **8**: 1329-1339. 2ezo: M. Caffrey, M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, D. G. Covell, A. M. Gronenborn, G. M. Clore (1998) Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J* **17**: 4572-4584. 1a1t: R. N. De Guzman, Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer, M. F. Summers (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**: 384-388.

ACCESSORY PROTEINS: 1pi7: S. H. Park, A. A. Mrse, A. A. Nevzorov, M. F. Mesleh, M. Oblatt-Montal, M. Montal, S. J. Opella (2003) Three-dimensional structure of the channel-forming trans-membrane domain of virus protein "u" (Vpu) from HIV-1. *J Mol Biol* **333**: 409-424. 1vpu: D. Willbold, S. Hoffmann, P. Rosch (1997) Secondary structure and tertiary fold of the human immunodeficiency virus protein U (Vpu) cytoplasmic domain in solution. *Eur J Biochem* **245**: 581-588. 3dcg: B. J. Stanley, E. S. Ehrlich, L. Short, Y. Yu, Z. Xiao, X. F. Yu, Y. Xiong (2008) Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *J Virol* **82**: 8656-8663. 1esx: K. Wecker, N. Morellet, S. Bouaziz, B. P. Roques (2002) NMR structure of the HIV-1 regulatory protein Vpr in H2O/trifluoroethanol. Comparison with the Vpr N-terminal (1-51) and C-terminal (52-96) domains. *Eur J Biochem* **269**: 3779-3788. 1avv: S. Arold, P. Franken, M. P. Strub, F. Hoh, S. Benichou, R. Benarous, C. Dumas.(1997) The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling. *Structure* **5**: 1361-1372. 1qa5: M. Geyer, C. E. Munte, J. Schorr, R. Kellner, H. R. Kalbitzer (1999) Structure of the anchor-domain of myristoylated and non-myristoylated HIV-1 Nef protein. *J Mol Biol* **289**: 123-138. 1etf: J. L. Battiste, H. Mao, N. S. Rao, R. Tan, D. R. Muhandiram, L. E. Kay, A. D. Frankel, J. R. Williamson (1996) Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science* **273**: 1547-1551. 1biv: X. Ye, R. A. Kumar, D. J. Patel (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem Biol* **2**: 827-840. 1jfw: J. M. Péloponèse Jr., C. Grégoire, S. Opi, D. Esquieu, J. Sturgis, E. Lebrun, E. Meurs, Y. Collette, D. Olive, A. M. Aubertin, M. Witvrow, C. Pannecouque, E. De Clercq, C. Bailly, J. Lebreton, E. P. Loret (2000) 1H-13C nuclear magnetic resonance assignment and structural characterization of HIV-1 Tat protein. *C R Acad Sci III* **323**: 883-894.