CHAPTER 2

REVIEW OF LITERATURE
Introduction to HIV-1 envelope

HIV has the potential to evolve at an extremely high rate due to the combination of a high mutation rate (Hu et al., 1997), a short generation time (Perelson et al., 1996), and a large population of productively infected cells within the host (Chun et al., 1997). Among genes within HIV genome, the envelope gene evolves particularly at high rate (1-2% per year) at the population (Thomas et al., 1997) and the individual level. After infection, genetic diversity in env begins low, undergoes a drop, and then increases to a peak several years into the infection. Genetic divergence from the infecting strain increases during infection, finally reaching a plateau several years into infection. Consequently, env is highly genetically diverse, posing a significant challenge to vaccine development (Simon et al., 2005).

HIV-1 infection requires two sequential and specific binding steps by the envelope glycoprotein: first is the attachment to cell surface receptor CD4 present in CD4+ T cells, monocytes/macrophages and other cells; and second to a member of the chemokine receptor subfamily, within the G-protein coupled seven transmembrane domain family of receptors, mainly CCR5 and/or CXCR4 (Peters et al., 2008; Clapham et al., 2002; Rossi et al., 2008). Envelope is a heterodimer of a transmembrane glycoprotein (gp41) and a surface glycoprotein (gp120), and forms trimer on the surface of the viral membrane (Liu et al., 2008). The resulting complex protrudes from the virus surface and is often described as a spike, of which there may be as few as 10 on the average HIV-1 virion (Zhu et al., 2003). The precise number of Env trimers on the virion surface probably varies from strain to strain and is affected by a virus passage history in culture (Hartley et al., 2005). Each monomer of gp120 contains a binding site for CD4. Engagement of one CD4 molecule by a single gp120 in the trimeric spike is sufficient to induce conformational changes in all three glycoprotein monomers of the trimer (Salzwedel and Berger, 2000). The envelope glycoprotein of HIV-1 is initially synthesized as type 1 membrane precursor molecule gp160, which is cotranslationally targeted to the ER by its 30 amino acids N-terminal signal sequence (Li et al., 1994). Within the ER, gp160 receives approximately 30 N-linked glycans and is assisted in its maturation by the chaperones Bip, Calnexin and
calreticulin as it undergoes extensive disulfide bond formations. Once gp160 has reached its native state with ten disulfide bonds and its signal sequence has been cleaved post translationally, it assembles into trimers and is exported to the golgi (Jejeic et al., 2009). Within golgi gp160 is cleaved by cellular endoproteases like Furin and LPC/PC7, yielding the transmembrane protein gp41 and the non covalently associated surface protein gp120 (Jejeic et al., 2009; Hallenberger et al., 1997). Thereafter, this complex is transported to the plasma membrane, where it is incorporated into the envelope of assembling HIV-1 particles (Jejeic et al., 2009). simian immunodeficiency virus (SIV)-infected monkeys have shown that immunogens capable of inducing CTL responses can reduce viral set point and slow disease progression (Letvin, 2005). These observations suggested that vaccines capable of eliciting strong T-cell mediated immune responses may be beneficial even if they do not induce sterilizing immunity. However, this type of response does not clear the virus reservoirs and resistance variants can emerge later in infection (Barouch et al., 2003). The failure of Merck's candidate HIV-vaccine which aimed to elicit cell mediated immunity has further questioned this vaccine approach (http://www3.niaid.nih.gov/news/newsreleases/2007/step_statement.htm). There is renewed interest in neutralizing antibodies; prompted mainly by several observations that passive transfer of neutralizing antibodies is able to confer sterilizing immunity in animal studies (Ruprecht, Mascola, 2002). In addition, it has been shown that although effector T cells can limit viral replication, they are not able to assist humoral immunity to prevent the establishment of initial infection (Mascola et al., 2003). As such, an enormous effort is currently being invested in the "intelligent" design of immunogens capable of inducing broadly cross-reactive neutralizing antibodies against HIV-1 primary isolates. Such an endeavor requires not only an in-depth understanding of the viral Envelope glycoprotein structure and function, but also the role of neutralizing antibodies in natural HIV-1 infection.
The main targets of the anti-HIV neutralizing antibodies are the glycoprotein spikes on the virus Envelope membrane. The antibodies directed against this protein block the interaction of this protein with the cellular receptor and therefore these antibodies are referred as neutralizing antibodies. This glycoprotein complex interacts with the receptors present on the surface of target cell initiating the viral entry process. The functional Envelope spike consists of a trimer of heterodimers formed by two glycoproteins, gp120 (the surface protein) and gp41 (the transmembrane glycoprotein/fusion protein). Three gp120 elements interact noncovalently with three gp41 units forming an oligomer, where the trimeric structure is maintained by the interactions between the gp41 domains and the intermolecular interactions among elements of gp120. HIV-1 Envelope glycoprotein (gp160) is a type 1 integral membrane protein and is synthesized as a single polypeptide of approximately 845 to 870 amino acids that plays a critical role in virus infection. gp160 is co-translationally translocated into the endoplasmic reticulum (ER). As soon as the nascent chain enters the ER lumen, it starts to fold, disulfide bonds are formed and the protein becomes glycosylated. gp160 contains approximately 30 potential N-linked glycosylation sites, most of which are used (Gangster et al. 1999).

Amongst the many different strains of the virus, the number and also the location of most sites is fairly conserved. Chaperones and folding enzymes, such as BiP, calnexin, calreticulin, and PDI, assist the folding of gp160. Cleavage of the signal peptide occurs completely at posttranslational level (Song et al., 2005). When gp160 attains its native conformation towards a trimeric form, it is transported to the Golgi complex, where the accessible glycans are trimmed and modified to complex oligosaccharides and in the trans-Golgi, cellular proteases cleave the gp160 molecule into gp120 and gp41 (Decroly et al., 1997; Hallenberger et al., 1997). Finally, the gp160 glycoprotein is transported to the plasma membrane, particularly to the detergent insoluble membrane domains, known as lipid rafts (Rousso et al., 2000).
where the virus assembly takes place and Envelope spikes are incorporated into the virion.

**The gp120 subunit (Surface Protein)**

The amino acid sequence of gp120 consists of five relatively conserved regions (C1-C5) interposed with five variable regions (V1-V5) which, with the exception of V5, are bracketed by cysteines forming disulfide bonds (Leonard et al., 1990). gp120 is a highly glycosylated protein with half of its mass being N-linked glycans (Lasky et al., 1986), far more glycosylated than the surface proteins from other retroviruses of similar size such as HTLV-1 and MuLV (Polonoff, Machida, and Kabat, 1982).

Two types of N-linked glycosylation are found on the surface of gp120, mannose-rich and complex glycans. Structural modeling suggests that the high mannose glycans are clustered on one side of the surface while the complex glycans are localized within a distinct region of the gp120 (Zhu et al., 2000).

**Structural domains of gp120**

Full-length gp120 has eluded structural analysis due to its lack of stability and high flexibility. To obtain crystal structures, HIV-1 and SIV gp120s have been deglycosylated and the N- and C-terminals, V1/V2 and V3 regions truncated, to generate what is commonly referred to as the “gp120 core”. The first gp120 structure was obtained using the gp120 core of an HIV-1 virus stabilized with the D1D2 fragment of CD4 and a CD4 induced epitope-binding antibody (17b) (Kwong et al., 1998).
Fig 2.1 Schematic representation of the functional domains of gp41. (a) Functional motifs are indicated in a linear diagram of gp41. The membrane-spanning domain separates the ectodomain from the cytoplasmic tail. (b) The N-linked glycan moieties are shown all over the Env protein which makes around half of the mass of the total gp160. (Adapted from Frey et al (Frey et al., 2008))

Based on this structure, gp120 is organized into three regions: the inner domain, the outer domain and the bridging sheet (Figure 2.2). The inner domain is formed mainly by the C1 and C5 regions and is largely devoid of glycans. It has long been suggested that this domain constitutes the major contact interface with the gp41 transmembrane unit (Helseth et al., 1991; Moore et al., 1994). The outer domain of gp120 possesses important sites that interact with cellular receptors. This outer domain is heavily glycosylated and modeling of the Envelope oligomer suggests that these glycans cover the solvent-exposed part of the spike, protecting it from antibody recognition. In between the outer and inner domains is the bridging sheet region, formed by four anti parallel β-sheets: β2 and β3, which constitute the stem of the deleted V1/V2 loop; and the β20 and β21 of the C4 region. The bridging sheet plays an important role in maintaining the conformation of the entire gp160 molecule.
FIG 2.2 Structure of HXBc2 and YU2 gp120. In the orientation shown in the figure, the viral membrane is positioned above the molecule; the target cell membrane is below. Ribbon diagram of gp120 core structure. In the left the HXBc2 core is depicted in red (α helices) and salmon (β strands) except for β15 in yellow, which hydrogen bonds to the C" strand of CD4. The inner domain (N terminus-α1, β4-β8, and α5-C terminus), bridging sheet (β2, β3, β20, and β21), and outer domain (β9-β19 and β22-β24) are labeled. In the right panel, the YU2 core is depicted in green (α helices) and light green (β strands) except for β15. The N terminus (N) and C terminus (C) are labeled, as are the sequence-variable loops. Taken from Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing antibody. Nature 393: 648-659.

More recently the gp120 of an SIV virus was resolved in the absence of CD4. This unliganded gp120 structure is presumed to represent the native state conformation(Chen et al., 2005). Comparison between liganded gp120 of HIV and unliganded gp120 of SIV revealed marked difference in structure of inner domain and the bridging sheet, whereas the outer domain is not much differed. This suggests that binding of CD4 induces drastic conformational changes in these regions. The bridging sheet observed in the CD4 bound conformation appears segregated in then ative structure (Figure 2.3a), with the β2 and β3 sheets, laying proximal to the inner-domain, displaced 20-25 Å from the β20 and β21 loop. A third gp120 structure was resolved using an HIV-1 gp120 core that included the V3 loop, complexed to CD4 and the Fab X5 (Huang et al., 2005). In this structure the core resembles the liganded conformation. The V3 loop protrudes as an elongated structure from two antiparallel
β-sheets in the outer domain. A disulphide bond between β12 and β13 stabilizes its base, while a long flexible stem extends away from the core ending in a conserved β-turn tip. The trimeric model of this structure predicts the projection of the tip 30 Å towards the cell membrane, consistent with its positioning towards the coreceptor molecule (Huang et al., 2005) (Figure 2.4).

![Diagram of HIV gp120 core structures](image)

**FIG2.3** Comparison of crystal structures of HIV and SIV gp120 core. A) Structure of CD4 liganded HIV-1 gp120 core (HXB2) viewed from the perspective of CD4. The gp120 inner domain (blue), outer (yellow) bridging sheet orange are shown. The location of various gp120 region are also denoted. B) Structure of unliganded SIV gp120 core (Protein Data Bank ID 2BFI) viewed from perspective of CD4 as in A). C) HIV core in the same orientation as in A) depicting CD4 contact residues (orange) and residues that influence coreceptor binding (green). D) SIV gp120 core in the same orientation as in B) coloured according to scheme of HIV gp120 in C). Taken from Pantoplot R. et al Annu Rev Immunol, 2006.

However, it is not clear if V3 displays an extended structure in the context of the native oligomer. It has been suggested that the V3 is partially occluded by the
V1/V2 loop of the adjacent protomer in the trimeric Envelope glycoprotein, and it is only extended after CD4 interaction (Wyatt et al., 1998). The widely assumed trimeric structure of the Envelope complex has recently been visualized using cryoelectron tomography microscopy by two separate groups (Zanetti et al., 2006; Zhu et al., 2006). Both studies have demonstrated similar dimensions and three-fold symmetry for the Envelope spike, but the actual map images displayed some distinct features. The Zhu et al. (Zhu et al., 2006) structure presents multiple lobes emanating from the core and tripod-like configuration of gp41, while the Zanetti et al. (Zanetti et al., 2006) reconstruction constitutes a smooth structure with only three clear lobes and a single stem linking with the membrane (Roux and Taylor, 2007). These differences have been attributed to the techniques used for data collection and analysis (White et al., 2010). Despite these limitations, this rapid evolving methodology promises to assist in solving the native structure of the Envelope spike in more physiologically relevant environments, i.e. in the context of a viral or cell membrane.

![Diagram](https://via.placeholder.com/150)

**FIG 2.4** Modeled trimer and coreceptor schematic. (A) V3 in the context of a trimer at the target cell surface. gp120 engagement with the CD4 rearranges the trimer gp120 structure around the recessed V3 loop is projected toward the cell membrane for binding with the coreceptor. In this orientation, the target cell membrane and coreceptor are expected to be positioned toward the bottom of the page. (B) Schematic of coreceptor interaction. CCR5 (green) is shown with its tyrosine-sulfated N terminus (at residues 3, 10, 14, and 15) and three extracellular loops (ECLs). V3 (red) is shown with its conserved base interacting with the sulfated CCR5 N terminus and its flexible legs allowing its conserved V3 tip to reach the second ECL of CCR5. (Taken from Huang et al. (Huang et al., 2005).)
The inner domain is formed mainly by the C1 and C5 regions and is largely devoid of glycans (Kwong et al., 1998; Wyatt et al., 1998), which strongly supports the proposition that C1 and C5 function as the major contact interface with the gp41 transmembrane unit. The outer domain, in contrast, is largely covered by glycans (Kwong et al., 1998; Pantophlet and Burton, 2006; Wyatt et al., 1998). Comparison of the liganded gp120 structure of HIV-1 (Figure 2.3A) and the unliganded gp120 structure of SIV (Figure 2.3B) shows that the respective outer domains are highly similar (Chen et al., 2005). The gp120 structure is divided into two major domain, inner domain and outer domain. The inner domain provides conformational flexibility to the gp120 in the absence of CD4. Comparison of the inner domain substructures in the unliganded and liganded core structures suggests that, upon CD4 binding, these substructures are repositioned somewhat independently of each other, rather than a shift of the inner domain as a single unit (Chen et al., 2005). CD4 entails conformational change in gp120 structure which is associated with large energy change (Myszka et al., 2000). These conformational changes influence the formation of the bridging sheet and the interactive surface structure with gp41 and required for forming the coreceptor binding site (Wyatt et al., 1998). The structural changes occurred post attachment of CD4 affect the fusion process. The conformational changes that occur within the inner domain also affect the formation of the bridging sheet, which links the inner and outer domains. In the CD4- liganded gp120 conformation, the bridging sheet is folded into a compact anti parallel, four stranded β-sheet (β2-β3 and β20-β21) (Kwong et al., 1998; Wyatt et al., 1998). However, in the unliganded structure, the β-strands that constitute the bridging sheet lie separated in pairs at a distance of approximately 20 Å (Chen et al., 2005); the two β-strands (β2-β3) that constitute the V1/V2 stem are located in the vicinity of the inner domain, whereas the other two strands (β20-β21) are situated near the outerdomain in approximately the same location as they are on the liganded structure (Pantophlet and Burton, 2006).
Organization of gp120 on the Viral Surface

Knowledge of the oligomeric structure of the gp120-gp41 complex is important for vaccine design strategies. Experimental evidence suggests that the functional unit of the Envelope spike is a heterodimeric trimer complex of gp120 and gp41. For example, a recent electron tomography study revealed structures on the surface of negatively stained virions of SIV and HIV-1 that appear to be tri-lobed Envelope glycoproteins (Zhu et al., 2003; Zhu et al., 2006) (Fig 2.5). Furthermore, the HIV core matrix that interacts with gp41 is organized in a trimeric configuration, and the crystal structures of HIV-1 gp41 cores resemble the transmembrane proteins of other viruses that have been shown to display trimeric Envelope spikes (Cha et al., 1997; Tan et al., 1997; Weissenhorn, 1997). There is, however, also evidence that other Envelope species may be present on the surface of HIV-1. For example, atomic force microscopy analyses have failed to reveal any uniform trimeric Envelope species on the surface of virions (Kuznetsov et al., 2003). Also, in recent studies, it has been shown that viruses can be captured onto ELISA plate wells using antibodies that are unable to neutralize viral particles in solution.

![FIG 2.5 3D tomograms of SIV and HIV pseudovirions. (a) Tomogram section from an SIV particle with a truncated cytoplasmic domain and a high level of Envelope spike expression on the virion surface. (b) Tomogram of an HIV-1 particle expressing the full-length Envelope glycoprotein. Representative tri-lobed structures.](image)
presumably of the Envelope spikes, are indicated (white arrows). Tomograms courtesy of Drs. Ping Zhu and Ken Roux (Florida State University, Tallahassee, Florida).

(Herrera et al., 2005; Poignard et al., 2003). Taken together, these observations suggest that, although trimers may likely represent the functional Envelope spike, both functional and nonfunctional forms of the Envelope may be present on the virion surface. These nonfunctional Envelope entities may be monomers, dimers, or tetramers and could possibly arise as the result of (a) the dissociation of functional gp120-gp41 complexes, which could perhaps cause gp120 to be shed from the viral surface, or (b) inefficient trimerization of the spike in the Golgi (Burton, 1997; Parren et al., 1997; Wyatt et al., 1998; Wyatt and Sodroski, 1998).

Functional sites of gp120

CD4 binding site (CD4bs)

The CD4bs of gp120 constitutes a conformational region suggested to be only apparent in the context of the liganded structure of gp120. It is characterized by a hydrophobic pocket at the interface of the inner domain, the bridging sheet and outer domain. The residues in the cavity lining are mostly conserved and strictly hydrophobic. In the CD4 bound structure the side chain of the Phe43 of CD4 is buried in this pocket, otherwise known as the "Phe43 cavity. In the unliganded structure many of the residues involved in the interaction with CD4 are distributed around the interface of the inner and outer domain (Chen et al., 2005). These residues are from stem of V1/V2; loop L2, the B15-a3 excursion. Direct inter atomic contacts are made between 22 CD4 residues and 26 gp120 amino acid residues (Kwong et al., 1998). It has been proposed that the CD4 molecule first interacts with the internal face of the conformationally stable outer domains.

However, this interaction is not energetically favorable, incurring a substantial drop in entropy (Kwong et al., 2002), and is only stable at the cell surface where the presence of multiple CD4 molecules, binding the trimer simultaneously, increases the avidity of this interaction (Zhou et al., 2007). The binding of CD4 induces large conformational changes in the inner domain, which leads to the formation of the
bridging sheet and coreceptor binding site. The CD4 binding with the gp120 also changes the conformation and orientation of the V1V2 and V3 loops.

**Coreceptor binding site**

In addition to CD4 receptors, HIV requires the presence of coreceptor molecules on the surface of the target cells. Most primary isolates use the β-chemokine receptor CCR5 as an entry coreceptor, although some viruses undergo a coreceptor switch mainly to CXCR4 usage. Other minor coreceptors such as CCR1, CCR2b, CCR3, CCR8, CXCR6 (Bonzo/STRL33), Bob/GPR15 and GPR1, have also been shown to mediate virus entry *in vitro*, although their use *in vivo* is less certain (Moore *et al.*, 2004). The V3 loop has been mapped as the major determinant of coreceptor switching, which demonstrates its involvement in the coreceptor-binding site. Other conserved structures of gp120 also form part of this functional region, such as the bridging sheet and the stem of the V3 loop (Rizzuto and Sodroski, 2000; Rizzuto *et al.*, 1998). These residues are segregated in the unliganded gp120 structure and only converge after CD4 binding, to form a conserved pocket that harbors the sulfotyrosines on the N terminus of CCR5 (Figure 2.4). This interaction zips the flexible V3 stem into a rigid β-hairpin (Huang *et al.*, 2007). However, it is not clear if these changes occur before or after the tip of the V3 interacts with the second extracellular loop of CCR5 (Huang *et al.*, 2007).

The gp41 Subunit (Fusion Protein)Env protein gp41 anchors the infectious spike to the viral membrane and plays an important role in cell entry. It consists of 345 amino acids (aa) with a molecular mass of 41 kDa, does not contain clearly defined variable regions, and is more conserved than gp120. As shown in Figure 2.1, it is divided into three major domains (Gabuzda *et al.*, 1992): the extracellular region, also called the ectodomain (aa 512 to 683; numbering is based on HIV-1 HXB2 (Kuiken *et al.* HIV sequence compendium. Los Alamos National Laboratory, Los Alamos, New Mexico, USA) unless otherwise specified), the transmembrane (TM) domain (aa 684 to 705), and the cytoplasmic tail (CT) (aa 705 to 856). The ectodomain contains several distinct functional determinants involved in the fusion of viral and host cell membranes: (i) an N-terminal hydrophobic region that functions as a fusion peptide (FP) (aa 512 to 527); (ii) a polar region (PR) (aa 525 to 543); (iii) two α-helix repeat
regions referred as the N-terminal heptad repeat (NHR) (aa 546 to 581) and the C-terminal heptad repeat (CHR) (aa 628 to 661) (Chan et al., 1997); (iv) a disulfide-bridged hydrophilic loop that connects the two heptad repeats, also known as the connecting loop, the cluster I epitope, or the immunodominant (ID) loop (aa 598 to 604); (v) a Trp-rich region known as the MPER (aa 660 to 683); and (vi) a membrane-spanning domain, also called TM, and the CT. Thus, gp41 is comprised of several distinct regions that each contributes unique functions. The structure of whole gp41 is not clearly defined, since the available crystal structures of the HIV-1 gp41 ectodomain core do not contain the FP and the ID loop and have deletions of the MPER (Chan et al 1997, Tan et al 1997, Weissenhorn et al 1997). These studies revealed a six-helix bundle (6HB), which is considered to be the post fusion structure of the ectodomain and different from the native metastable structure of gp41 in the viral spike.

Recently, two different structural models for the viral spike of the whole virion have been solved using cryoelectron tomography (Zanetti et al., 2006; Zhu et al., 2006).
Roux and colleagues proposed that the MPER and TM "stalk" of each trimer are composed of three separate legs that project obliquely from the head of the trimers a tripod-like structure. In contrast, the structural model proposed by Zanetti et al. (Zanetti et al., 2006) shows the TM domain of simian immunodeficiency virus (SIV)gp41 as a stem in the viral surface; the tripod-like structure model correlates with the current view of the possible membrane involvement of the gp41 MPER (Roux and Taylor, 2007). This technique is still being developed, and these conflicting structures could be attributed to the methods used to collect the data and/or the computational approaches used to determine the structures. Thus, the structure of the native gp41 protein in the viral spike remains an open question to be revealed by future studies. The lack of agreement between those two studies demonstrates one of the main challenges involved in understanding the complexities of this virus.

How many Env proteins are required for infection? This question has intrigued HIV-1 researchers for years. A study reported by Yang et al. supports the notion that only a single viral spike could be required for fusion (Yang et al. 2005); however, newer structural studies indicate a different picture. Recently, the structure of the HIV-1–
Tcell interaction at the time of infection has been solved using microscopictomography. Sougrat et al. (Sougrat et al., 2007) observed the arrangement of about six to seven dense rod-like structures at the contact site between the cell and viral membrane. These rods were hypothesized to comprise a single Env spike. This structurally unique contact zone comprises an arrangement of closely packed rods (presumed to be Env) extending from the virus into the membrane of the target cell; this structure was named the "entry claw." However, at this atomic level of resolution (~20 Å), structural information is not refined enough to reveal how the fusion process occurs. Those authors proposed two different models describing the last step of fusion between virus and the host cell membranes. In the first model, the entire width of the "entry claw" is required for the fusion of the viral and the host membranes. This results in the merging of the contents of the two membranes. In the second model, a rod at the center of the "entry claw" creates a pore in the center of the contact zone, which permits the release of the viral core into the cell (Sougrat et al., 2007). Thus, gp41 appears to change its structure repeatedly and to change its location...
on the surface of the virion during infection. This structural flexibility maybe yet another reason behind the difficulties in designing immunogens that elicit bNtAbs.

HIV Receptors

CD4 Receptor

CD4 is a transmembrane glycoprotein expressed on the surface of T-helper cell, regulatory T-cells, monocytes, macrophages and dendritic cells (Bourgeois et al., 2006). The CD4 antigen is a membrane glycoprotein of T lymphocytes that interacts with major histocompatibility complex class II antigens and is also a receptor for the human immunodeficiency virus (Raymond et al., 1991; Lange et al., 1994). HIV infection of CD4+ cells is initiated by binding of the virus to the cell surface, via a high affinity interaction between CD4 and the HIV outer envelope glycoprotein, gp120 (Sattentau 1992). The development of model systems using soluble recombinant forms of CD4 (sCD4) has allowed kinetic and thermodynamic analysis of CD4 binding to gp120, and study of the post-binding events leading to virus-cell membrane fusion. It has thus been demonstrated that the affinity of sCD4 for gp120 on virions or HIV-infected cells depends on both the primary sequence and the tertiary structure of gp120 in the membrane (Sattentau 1992).

The term CD stands for “clusters of differentiation” and was introduced to classify monoclonal antibodies that recognised identical subgroups from a panel of leucocyte cell lines (Bernard et al., 1984). The T-cell receptor for antigen (TCR) is a membrane bound member of the immunoglobulin superfamily (IgSF) that recognizes a complex of peptide antigen bound to self major histocompatibility complex (MHC) molecules. Antigen associated with class II MHC molecules generally elicits a helper response. CD4 participates in antigen recognition through interactions with nonpolymorphic regions of class II molecules by its extracellular domains (Leahy 1995).

CD4 is predominantly expressed as a 55 kDa monomer, but CD4 dimers and trimers (110 and 220 kDa) have also been found to be expressed on the cell surface.
of T-lymphocytes and monocytes and macrophages (Bourgeois et al., 2006). Bourgeois et al. proposed that monomeric forms of CD4 are preferentially used by HIV-1 to gain entry into the target cells. This implies that dimer/monomer ratio at the surface of the target cell may modulate the efficiency of HIV entry. CD4 has 428 amino acids in total (Wu et al., 1996) which are distributed as segments in immunoglobulin like (Ig like) extracellular domains of 372 amino acids (Barclay et al., 1993; Bour et al., 1995), a transmembrane segment of 23 amino acids and, a cytoplasmic tail of 33 amino acids on C terminal end that interacts with p56 Lck - a src like tyrosine kinase (Bour et al., 1995; Leach, 1995, Wu et al., 1996). It contains two cysteines which are essential for the interaction with Lck. This enzyme is essential for activating many molecules involved in the signalling cascade of an activated T-cell. The extracellular region of the CD4 protein is a single chain molecule that is made up of four tandemly arranged immunoglobulin-like domains. A major advance in our knowledge of CD4 structure came with the solutions of 3-D structures of CD4 fragments consisting either of the first two domains of human CD4 (D1D2) (Wang et al., 1990; Ryu et al., 1990) or the membrane proximal domains of rat CD4 (D3D4) (Brady et al., 1993; Lange et al., 1994). These domains are categorized by their relation to the membrane in that there is a two domain segment called the membrane distal fragment (Which consists of Domain 1 or D1, and Domain 2 or D2), as well as a two domain segment called the membrane proximal fragment (D3 and D4) (Wu et al., 1996) (Figure 2.7). The domain that is of interest, when it comes to HIV and the gp120 protein, is domain D1. This domain is homologous to the variable domain of antibodies that exhibit the structure of two β sheets four or five strands long (Wu et al., 1996). The stucture of D1 (as well as D3) is consistent with that of V-Domain, while D2 and D4 are categorized as C-Domains. The V or variable domain is often accompanied by a joining region (Wuet al., 1996). The two combined regions are classified as the V-J region and it is at this region (known as D1) that CD4 binds to β2 domain of MHC class II molecules.

The first domain which contains high affinity binding site for gp120 is composed of nine β strands. By analogy with the antibody V domains the nine strands are termed A, B, C, C’, C”, D, E, F, G; four of them (ABED) form an antiparallel β
A sheet which is packed against another antiparallel β sheet formed by CC'C'EFG.

Also by analogy with the hypervariable complementarity determining regions (CDR) of Ig V domains, the loop between the loops between the strands of B and C is termed as CDR1, that between C' and C'' – CDR2, and that between F and G – CDR3.

The second CD4 domain contains seven β strands (A, B, C, C', E, F, G). The A strand is a continuation of the D1 G strand and connects D1 and D2. The size and overall structure of D2 resembles the constant (C) domains of Ig but has patches of sequences similar to V domains. The strand between C and E strands termed C' strand forms hydrogen bonds to strand C rather than to E as observed for strand D in C domains of Ig. This leads to formation of two antiparallel β sheet packed against each other and composed of ABE and CC'EFG instead of ABDE and CFG as in Ig C domain. The CD4 D2 doesnot have a disulfide bond between the two sheets but an intrasheet disulfide bond which connects F and C strands.

The structure of the fragment from the third and fourth domain of rat CD4 somewhat resembles that of human D1D2 fragment. The two domains D3 and D4 are connected by a continuous β strands and D3 are folded like an IgV domain but doesnot have a disulfide bond. It has been suggested that D2 and D3 are likely to be connected by a β strand and the interactions between D3 and D4 are similar to those between D1 and D2The structure of an N-terminal fragment of CD4 has been determined to 2.4 Å resolutions. It has two tightly abutting domains connected by a continuous β strand. Both have the immunoglobulin fold, but domain 2 has a truncated β barrel and a non-standard disulphide bond. The binding sites for monoclonal antibodies, class II major histocompatibility complex molecules, and human immunodeficiency virus gp120 can be mapped on the molecular surface (Wang et al., 1990). The crystal structure of the third and fourth domains of rat CD4 was solved at 2.8 angstrom resolution and shows that both domains have immunoglobulin folds (Brady et al., 1993). Domain 3, however, lacks the disulfide between the β sheets; this results in an expansion of the domain and is associated with an alteration in packing of β sheets, which may be important for interactions with domain 2 in the overall receptor structure (Brady et al., 1993; Lange et al., 1994). Domains 3 and 4 of CD4 show considerable similarity to of N-linked glycosylation sites on one face of domain
3 appears to preclude the dimerization i.e. observed in antibodies \( (\text{Lange et al., 1994}) \). D1 and D2 are not glycosylated, but D3 and D4 have two glyosylation sites; the N-linked carbohydrates can impose significant limitations on the accessibility to the glycosylated face because domains 1 and 2, although there is a 25° rotation in the relative positions of the domains with respect to one another.

![Structure of CD4](image.png)

FIG: 2.7 Structure of CD4
Biological function of CD4

Now that we have determined the structure of the T-cell surface glycoprotein CD4 we can now begin to understand the function of the protein. As mentioned above, the CD4 has tandem extracellular domains (D1, D2, D3, and D4) (Figure 2.7). In a normal, virus free organism, these termini are used by CD4 for antigen recognition when associated with extracellular major histocompatibility complex class II (MHC class II) (Wu et al., 1996). In past studies, it has been shown that CD4 is the primary binding site for the antigen T4/leu3 (Wu et al., 1996). Although this reaction occurs extracellularly, CD4 can also undergo intracellular reactions with the src-like lymphocyte tyrosine kinase (Lck). CD4 is a transmembrane glycoprotein of the immunoglobulin superfamily, expressed on developing thymocytes, major histocompatibility class II (class II MHC) restricted mature T lymphocytes and, in humans, on cells of the macrophage/monocyte lineage. On lymphoid cells, CD4 plays a critical role during thymocyte ontogeny and in the function of mature T cells. CD4 binds to non-polymorphic regions of class II MHC acting as a co-receptor for the T-cell antigen receptor (TCR). It increases avidity between thymocytes and antigen-presenting cells and contributes directly to signal transduction through association with the Src-like protein tyrosine kinase p56Lck. CD4 is also a co-receptor for the human and simian immunodeficiency viruses (HIV-1, HIV-2 and SIV). Clinically, CD4 antibodies may be used to achieve immunological tolerance to grafts and transplants (Bowers et al., 1997).

The interaction of CD4 with MHC molecules during both T-cell development and response to antigen are critical for T-cell activation with CD4 serving as a coreceptor for T-cell receptor-antigen engagement (Satoh et al., 1997; Leahy et al., 1995). Numerous mutation studies have been performed to determine the regions of CD4 involved in MHC-II binding and like many protein-protein complexes, the interface is generally believed to involve a large surface area of both D1 and D2, with many contact sites (Satoh et al., 1997). CD41 T cells participate in the pathogenesis of a number of immune-based human conditions, including autoimmune diseases, allogenic organ transplant rejection, and graft versus host disease (GVHD) following...
allogenic bone marrow transplantation. Small molecular inhibitors of the CD4-MHC class II interaction could potentially block the undesirable activation of CD4 T cells and could thus serve as effective immunosuppressive. The molecular nature of these interactions and signals arising from them are of considerable interest. CD4 appear capable the length of a TCR and a portion of an MHC molecule to bring their NH2 terminal domains into contact with the membrane proximal domains of an MHC molecules on an opposing cell.

CD4 is also involved in signalling. One implication for signalling that does seem clear from the CD4 structures is that any signalling through CD4 is unlikely to occur via a conformational change transmitted across the cell membrane. Any transmembrane signalling through CD4 seems likely to result in some fashion through their associations with other molecules. One of the most important observations suggesting how CD4 may signal across the membranes is that cytoplasmic regions of both these molecules are associated with same src like tyrosine kinases p561ck (Figure 2.7). Signalling through CD4 and TCR can result in phosphorylation of the zeta chain of TCR / CD3 complex and it has been suggested that MHC mediated cross-linking of TCR with CD4 initiates by bringing p561ck into proximity with the cytoplasmic regions of TCR and CD3. P561ck has been shown to be required for cytotoxicity and T-cell development (Leahy et al., 1995).

In addition to its central role in activation of T-helper cells, CD4 may have other physiological functions. Its interaction with IL-16 leads to an increase in intracytoplasmic calcium and inositol triphosphate and migratory responses (Parada et al., 1996). Interleukin -16 or IL-16, is a cytokine that has chemoattractant activity on CD4+ T lymphocytes. It has long been known that eosinophils and CD4+ T lymphocytes are recruited to sites of allergic inflammations, but the molecular mechanism are poorly understood. IL-16 also known as lymphocyte chemoattractant factor (Parada et al., 1996) is secreted by activated eosinophils as part of allergic response along with RANTES, an additional cytokine. Once bound to its cognate receptor, CD4 IL-16 initiates a signal cascade that results in the activation of the PKC family (Santa Crutz Biotechnology Inc., IL-16(323): sc- 52868).
Parada et al., 1996 investigated the role of protein kinase C (PKC) in CD4 mediated migratory response by IL-16. They found recombinant IL-16 induced rapid translocation of PKC from the cytosol to the membrane. PKC inhibitors H7, calphostin C, chelerythrine, and bisindolylmaleimide completely block IL-16 induced lymphocyte migration as well as the motile response induced by HIV-1 gp120 and anti-CD4 antibodies. Taken together the data suggested a role for PKC in CD4-mediated migratory responses. IL-16's activity on CD4 T-cells defines a role for CD4 on eosinophil and monocyte and broadens the scope of functions of CD4 on T-cells.

Alternative receptors

It is now widely accepted that CD4 is the main receptor for HIV, determining viral tropism for CD4 lymphocytes and possibly, monocytes and macrophages as well. However, HIV is also known to infect HIV negative human cell lines including those derived from colorectal carcinoma, rhabdomyosarcoma, bone marrow precursor cells, chondrocytes, synovial cells, hepatoma, glioma and neuroblastoma (Norkin, 1995). While in some cases the infection can still be mediated by low but undetectable amounts of CD4 in many cell systems, clearly demonstrating that the infection is mediated by molecules other than CD4 Some cell types targeted by HIV in vivo express high levels of CD4 e.g. T-cells, others including macrophages and dendritic cells (DCs), express barely detectable amounts. In these situations HIV may attach to cells by CD4-independent interactions involving sugar groups on the envelope glycoproteins with sugar and lectin like domains on cell surface receptors, such as mannose specific macrophage endocytosis receptor (Clapham et al., 2002). Similarly, Dendritic cells (DCs) also play important role in trafficking of whole virus particles to the T-cells. DC-SIGN (DC specific intercellular adhesion molecule 3-grabbing non-integrin), a type II membrane protein with a C-terminal C-type (i.e. Calcium dependent) lectin domain, is responsible for the ability of dendritic cells to efficiently capture and present HIV-1 to receptor positive cells. DC-SIGN appears to be a universal attachment factor for primate lentiviruses that can bind and transmit HIV-1, HIV-2 and Simian Immunodeficiency virus (SIV) (Coleman et al., 2009: Pohlmann...
Besides direct interactions of the envelope glycoprotein with cell surface receptors, interactions also occur between cell-derived molecules incorporated onto virions and their ligands. For instance Coleman et al., in 2009 reported that the interaction between ICAM-1 (Intracellular Adhesion Molecule-1) on DCs and LFA-1 (Leukocyte Function Associated Molecule) on T-cells plays an important role in DC mediated HIV-1 transmission. Pohlmann et al., in 2000 described a homologue of DC-SIGN, termed DC-SIGNR (DC-SIGN related) that exhibits 77% amino acid identity with DC-SIGN and also functions as a universal attachment factor for primate lentiviruses that can transmit multiple HIV-1, HIV-2 and SIV strains to receptor positive cell lines and to human PBMCs. DC-SIGNR is expressed on sinusoidal endothelial cells in the liver, lymph node sinuses and a significant portion of capillary endothelial cells in placenta but was not expressed at appreciable levels in peripheral blood derived DCs. The presence of an efficient virus attachment and presentation factor in these cell types indicates that DC-SIGNR could influence vertical transmission and results in enhanced infection of receptor positive cell types in lymph nodes. The expression pattern of DC-SIGN and DC-SIGNR suggests that they might play roles in both horizontal and vertical transmissions as well as in the dissemination of virus in the host cell (Pohlmann et al., 2001). It is now evident that cell membrane proteins, including the adhesion molecule LFA-1 (CD11a/CD18) and its ligands ICAM-1, ICAM-2 and ICAM-3 are involved in HIV infection. These molecules are expressed on cells that serve as hosts for the virus, as well as on the envelopes of HIV virions. ICAM-1 molecules incorporated into the envelope of HIV virions are functional and capable of interacting with LFA-1 receptor on the target cell surface and this interaction facilitates virus binding and increased infectivity of the virus by 2 to 10 fold. LFA-1 and its ICAM ligands have also been shown to be necessary for syncytium formation in HIV infected cultures and for efficient cell to cell transmission of the virus (Hioe et al., 2001). A sphingolipid, galactosylceramide (Gal-C) and its derivative galactosyl sulfatide were also identified as putative receptors for HIV in neuronal and glia cells in brain and epithelial cells lining the gastrointestinal tract and importantly on macrophages rather then CD4 (Norkin et al., 1995 : Yahi et al., 1992 : Harrouse et al., 1991). Similarly, Galactosyl-alkylacyl-
glycerolipid (Gal AAG) is structurally and antigenically related to Gal-C which when expressed on epithelial and neuronal cells, is a CD4 independent infectivity receptor for HIV-1, HIV-2 and SIV. HIV virus binds to the sperm through the interaction of envelope glycoprotein gp120 with Gal-AAG (Yeamann et al., 2003). Heparan sulphate, a highly sulphated polysaccharide, is present on the surface of mammalian cells and in the extracellular matrix in large quantities. The sulphated monosaccharide sequences within heparan sulfate determine the protein binding specificity and regulate biological functions. Numerous viruses utilize cell surface heparan sulfate as receptors to infect the target cells (Liu and Thorp, 2002). Argyris et al., 2003 also demonstrated the role of cell surface heparan sulphate proteoglycans (HSPGs), particularly syndecans in HIV entry. Syndecans serve as an attachment receptors for HIV-1 on macrophages. They also showed that syndecans concentrate into membrane rafts when they bind large, multivalent ligands. Cannon et al., in 2008 have shown that the scavenger receptor cysteine rich protein gp340 expressed on tissue macrophages and epithelial cells may function to enhance infection and suggests important new opportunities for HIV-1 pathogenesis investigation and therapy. In the genital tract, its expression by cervical and vaginal epithelial cells promotes HIV trans-infection and may play role in sexual transmission. gp340 is alternatively spliced product of the deleted in malignant brain tumors 1 (DMBT1) (Cannon et al., 2008).
Table 2.1: Alternate receptors used by HIV for entry (Taken from Clapham P.R. and Macknight A., 2002).

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Expression</th>
<th>Role in attachment and infection</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal-C</td>
<td>Neuronal and glia cells</td>
<td>Confers inefficient infection presumably by aiding attachment</td>
<td>Harouse <em>et al.</em> (1991)</td>
</tr>
<tr>
<td>Sulphatide (sulphate derivative of Gal-C)</td>
<td>Colorectal epithelial cells and primary macrophages</td>
<td>Confers efficient CD4-independent infection by NDK, a TCLA HIV-1 strain requires CXCR4 coreceptor</td>
<td>Fantini <em>et al.</em> (1993); Seddiki <em>et al.</em> (1994); Delezay <em>et al.</em> (1997)</td>
</tr>
<tr>
<td>Placental membrane binding protein</td>
<td>Cloned from a placental cDNA library</td>
<td>Binds virus particles to the cell surface and thus enhances infectivity via CD4 and coreceptors. May trap HIV in the periphery and carry to T-cells in lymph nodes</td>
<td>Curtis <em>et al.</em> (1992); Geijtenbeek <em>et al.</em> (2000)</td>
</tr>
<tr>
<td>DC-SIGN</td>
<td>On dendritic cells</td>
<td>⏰</td>
<td></td>
</tr>
<tr>
<td>DC-SIGNR</td>
<td>Endothelial cells, such as liver, sinusoidal and lymph node sinus endothelial cells</td>
<td>Acts in the same way as DC-SIGN</td>
<td>Pohlmann <em>et al.</em> (2001)</td>
</tr>
<tr>
<td>Heparans</td>
<td>Many cell types</td>
<td>Attach virus particles to cell surfaces via an interaction with the V3 loop thus enhancing infectivity via CD4 and coreceptors. Acts predominantly for CXCR4 using viruses.</td>
<td>Mondor <em>et al.</em> (1998)</td>
</tr>
<tr>
<td>LFA-1 / ICAM-1</td>
<td>LFA-1 is expressed on hematopoietic cells, ICAM-1 is on different cell types</td>
<td>ICAM-1 incorporated onto virions enhances attachment and infection of LFA-1 + cells.</td>
<td>Fortin <em>et al.</em> (1999); Paquette <em>et al.</em> (1998)</td>
</tr>
</tbody>
</table>
HIV Coreceptors

The Discovery of coreceptors

Soon after CD4 was shown to be the main receptor for HIV and SIV, evidence started to accumulate indicating that CD4 while capable of supporting virus binding was not sufficient to trigger the conformational changes in the viral envelope protein required for the membrane fusion event necessary for HIV-1 to enter a cell. Rather, a specific cellular cofactor or coreceptor was required in conjunction with CD4 to support virus entry (Clapham et al., 2002; Lee et al., 1998). In 1995, the chemokines RANTES, MIP-1α and MIP-1β were identified as the major HIV suppressive factors (Cocchi et al., 1995). Soon after in 1996, a cofactor for HIV-1 fusion and entry was identified, designated as “fusin” which was a putative G-protein coupled receptor with seven transmembrane segments (Feng et al., 1996). Fusin is the chemokine receptor that is now known as CXCR4, with its ligand, SDF-1 / CXCL12 endowed with anti-coreceptor activity. Deng et al., in 1996 further reported that CCR5 acts as a potent coreceptor, a receptor for β chemokines RANTES, MIP-1 α and MIP-1 β together with CD4 allows entry of macrophage tropic HIV-1 into cells, in similar way as CD4 and fusin is required for entry of HIV-1 into T-cell line adapted virus (Deng et al., 1996).

Classification of HIV on basis of coreceptor usage

Primary HIV isolates can be subdivided into two distinct groups according to their biological phenotype. Fast replication to high titres and syncytium formation in PBMC and the capacity to infect and replicate in a broad range of T-lymphoid and monocytoid cell lines often characterize viruses isolated from immunodeficient patients. In contrast, slow replication to low titres and induction of small, if any, syncytia in PBMC are characteristic of viruses isolated from individuals with no or mild symptoms of HIV infection. These viruses lack the capacity to infect and replicate in established cell lines and thus, donot induce syncytia in these cells. These two distinct groups of primary HIV isolates were classified as rapid / high or syncytium – inducing (SI) / T-tropic and slow / low or non-syncytium inducing (NSI) / M- tropic isolates, respectively (Bjorndal et al., 1997; Kuhmann et al., 2008).
However, even before the coreceptors were discovered, the phenotypic variants were differentiated by their ability to replicate, and form syncytia in the MT-2 cell line. MT-2 cell line is a CD4 T-cell line which expresses CXCR4 but not CCR5 coreceptor on its cell surface. The viruses that were able to form syncytia in MT-2 cells were named as “syncytium inducing” (SI) viruses while those that could not were designated as “Non-syncytium inducing” (NSI) viruses. However, with identification of coreceptors, the SI / NSI nomenclature became archaic, once it was realized that SI viruses were able to use CXCR4 whereas NSI used CCR5 only (Moore et al., 2004). Moreover, the viruses with different biological phenotypes differ in the choice of chemokine receptors. CCR5 usage corresponds to a slow / low, non-syncytium inducing (NSI) phenotype whereas isolates with a rapid / high, syncytium inducing capacity (SI) phenotype use CXCR4. Some, isolates are dualtropic capable of using both CCR5 and CXCR4 (Morner et al., 1999). A new alternative HIV-1 phenotype classification system based on coreceptor usage was proposed by Morner, in this system, isolates using CCR5 are called as R5 tropic, while those using CXCR4 are called as X4 tropic. Isolates using both coreceptors were known as R5X4 viruses or dual / mixed (R5X4) tropic viruses. Laboratory derived T-cell lines used to define the T-tropic phenotype express CXCR4 only and hence supported growth of viruses capable of using only X4 as a coreceptor. The macrophage used to define the M-tropic phenotype expressing both the coreceptors but the levels of CXCR4 are low relative to CCR5, hence favoured the replication of viruses using CCR5 as the coreceptor (Kuhmann et al., 2008). Coreceptor choice or coreceptor tropism is determined exclusively by structures present on gp120 subunits of HIV envelope glycoprotein, however it is the sequence of the V3 loop considered to be most important as it determines which coreceptor (CCR5 or CXCR4) is used to trigger the fusion potential of the envelope complex and hence which cells the virus will infect (Hartley et al., 2005).
Ph.D. Thesis
Sampurna Mukhopadhyay

Structure of chemokine receptors and residues involved in CCR5 coreceptor that modulates interaction with CD4-gp120 complex

Periera et al., reported that till now more than 40 chemokines and 19 chemokine receptors have been identified that serve as HIV coreceptors for entry, however despite this extensive range of potential coreceptors, only CCR5 and CXCR4 seem to be relevant in transmission and in the pathogenesis of HIV infection (Periera et al., 2009). CCR5, CXCR4 and other chemokine receptors belong to the rhodopsin superfamily of seven transmembrane G-protein coupled receptors (GPCRs), characterized by a heptapeptidic helical fold (7-TM) spanning the plasma membrane (Alkhatib et al., 2009; Paterlini et al., 2002). Chemokines are chemoattractant proteins with roles in immune development, inflammation, immunity, embryogenesis, and development and also mediates leucocyte chemotaxis. Chemokines generally consists of 70 to 120 residue polypeptides with a common folding pattern that form an α helix underlying three anti-parallel β strands and a less structured N-terminus (Clapham et al., 2002). These are characterised by a series of shared structural determinants including conserved cysteine residues that form disulfide bonds in the chemokine tertiary structure. There are four types of chemokines two major and minor superfamilies (Fernandez et al., 2002), characterised on the position of these cysteine residues. CXC or α chemokines are the ones, where first two cysteine residues in the primary amino acid sequence are separated by a single amino acid (X represents any intervening amino acid) and CC or β chemokines, are those in which cysteine residues lies next to each other. Most chemokines belong to one of these two classes, two additional branches of the chemokine superfamily each containing a single member have been described recently. The C chemokine lymphotactin lacks both the first and third cysteine in the “4 cystein motif”, but shares homology at its carboxyl end with the CC chemokines. Fractalkin / CX3C chemokine has three intervening amino acids between first two cysteine residues (Segerer et al., 2000). CCR5 and CXCR4 which are the two main coreceptor involved in HIV entry belongs to CC and CXC chemokine groups, respectively (Wang et al., 1998). This chemokine receptor consists of four domains exposed on the cell surface: the N-terminus in the
extracellular region, three extracellular loops (E1, E2 and E3), seven transmembrane helices spanning the membrane and carboxy terminal end lying within the intracellular region (Clapham et al., 2002; Paterlini et al., 2002). The schematic representation of CCR5 sequence is shown in Figure 2.8.

Figure 2.8 Schematic representation of CCR5 sequence. Grey rectangles outline residues in the 7-transmembrane (TM) region. TM1 through TM7. EI and IL denotes extracellular and intracellular loop regions, respectively. Disulfide bridges between C20 and C269, and C101 and C178 are shown as lines connecting these cystein residues. Gray circles denote conservation of strong groups in CCR1 through CCR5. Gray circles with heavy outline denote identical residues in CCR1 through CCR5. Black circle denotes highly conserved residues in the rhodopsin family of GPCR (Taken from Paterlini et al., 2002).

Many studies have focused on CCR5 interaction with HIV-1 and have tried to identify CCR5 residues that are involved in chemokine binding and subsequent signal transduction. Melikyan et al., in their genetic analysis findings revealed that it is mainly the N-terminal segment and the extracellular loop (ECL2) of CCR5 that is involved in interaction of CCR5 with gp120 / CD4 complex (Melikyan, et al., 2007). Liu et al., 2003 build structural models of CCR5 in complex with gp120 and CD4 using protein structural modelling, docking and molecular dynamics tools and these models revealed that initial interactions of CCR5 with gp120 are involved in negatively charged N-terminus of CCR5 and positively charged bridging sheet region of gp120. Further interactions occurred between ECL2 of CCR5 and base of V3 loop regions of gp120. These interactions results in conformational changes in gp120 and lead to the final entry of HIV into the cell (Liu et al., 2003). Clapham et al., reported...
that N-terminus of CCR5 is important coreceptor activity for HIV-1 R5 viruses, while ECL-2 is critical for X4 strains.

Electrostatic charge interactions are also likely to enhance gp120 and coreceptor interactions. The N-terminus of CCR5 is highly negatively charged / acidic because of the presence of two sulphated tyrosines critical for gp120 binding and plays an important role in the entry of R5 isolates. Sulfation of tyrosine residues in the N-terminal region of CCR5 is important not only in binding affinity to chemokine ligands CCL3 / MIP-1 $\alpha$, CCL4 / MIP-1 $\beta$, CCL5 / RANTES and CCL8 / MCP-2 but is also crucial for interactions with HIV-1 gp120. The region spanning amino acids 2-18 of CCR5 (CCR5 2-18, DYQVSSPIYDINYYTSE) contributes to interaction with HIV-1 gp120 / CD4 complex and it has four tyrosines that can be sulphated (Tyr 3, 10, 14 and 15) (Jen et al., 2009). Residues in the third variable loop (V3 loop) of gp120 also interact with the second extracellular loop of CCR5 and determine coreceptor specificity (Brower et al., 2009). These negative residues interact with positive amino acids in and around the bridging sheet on gp120. Zhou et al in 2000, also identified residues Tyr 10 and Lys26, in the N-terminus of CCR5 which are crucial for ligand binding and signalling and mutational changes in these residues hampers their functions. There are four cysteine residues in the extracellular region that forms two disulfide bonds. These extracellular loops are generally 10-30 residues in length. Extracellular loop 2 (ECL2) is attached to the end of extracellular loop 1 (ECL1) by a conserved disulfide bond between Cys101 and Cys178 (Qin et al., 2000) similarly another disulfide bond is formed between the N-terminus and third extracellular loop (ECL3) (Paterlini et al., 2002). Paterlini et al., also reported that there are specific monoclonal antibodies which recognizes epitope residues within these extracellular regions and thus can help to map the 3- dimensional structure. The antibody mAb-2D7 recognizes residues in both ECL1 and ECL2, while mAbs PA9 and PA14 binds to amino acids in both the N-terminus and ECL2. The ECL2 of GPCRs plays important role in ligand binding, signal transduction and CCR5 interactions with HIV-1 envelope. Maeda et al., by their structural analysis study provided an insight on CCR5 transmembrane residues such as Y108, Y251 and E283 that are important for gp120 fusion, HIV infectivity and inhibitor binding. The loss of hydrogen bond interactions
among these key transmembrane residues and interactions between E283 and S180, which forms the binding pocket for CCR5 inhibitors, may alter ECL2 conformation and thus provides insight to the mechanism of gp120 inhibition (Maeda et al., 2008). Identification of the residues involved in interaction of CCR5 with gp120 / CD4 complex thus provides new insights into the structural basis for CCR5 receptor ligand interaction and may thus guide the design of novel inhibitors (Zhou et al., 2000).

Alternate coreceptors

As already known CCR5 and CXCR4 are the major coreceptors used in vivo, there are at least 12 other members of the chemokine receptor family, and related “orphan receptors” that can support infection of indicator cell lines in vitro. These include CCR1, CCR2b, CCR3, CCR4, CXCR4, CCR8, GPR1, GPR15 / Bob, CXCR6 / STRL33 / Bonzo / TYMSTR, V28, CX3CR1, APJ, CMKCR1 / ChemR23, FPLR1, D6 and RDC1 (Cellier et al., 2005; Wiley et al., 2003; Edinger et al., 1998; Xiao et al., 1998; Sharron et al., 2000; Nedellec et al., 2009). Use of alternate coreceptors could expand HIV tropism to a broader variety of cell types and may be important for infection of certain tissues or compartments. Therefore, entry via coreceptors other than CCR5 and CXCR4 could play a role in progression of the disease and the efficiency of virus transmission (Pohlmann et al., 1999). The emergence of viral variants capable of utilizing a broad range of coreceptors correlates with HIV-1 disease progression and preceded with CD4 cell decline to <200 X 10^6 / l and finally correlated with the onset of AIDS (Xiao et al., 1998; Connor et al., 1997). HIV-2 and SIV strains use a wider range of these alternative coreceptors than HIV-1, as efficiently as they use CCR5 and / CXCR4 (Nedellec et al., 2009). SIV strains can be mediated by CCR5, but not CXCR4. SIV strains can also infect CD4 °T-cells that lack CCR5, which suggest that these strains can use alternate receptors as well (Deng et al., 1997). Certain groups have also investigated whether these orphan receptors could function as efficient coreceptors for both HIV and SIV envelopes in comparison with principal coreceptors CCR5 and CXCR4. For instance, Edinger et al., 1998 reported that GPR15 and STRL33 are rarely used by HIV-1 but more frequently used by SIV strains but in a manner that does not correlate with SIV tropism (Edinger et al.,...
However, Zhang et al., in 2003 reported that BONZO / STRL33 coreceptor can also be used by HIV-1 isolates and it is the four amino acid residues deletion in its V3 loop which differentiated this virus from other strictly HIV-1 R5 using isolates (Zhang et al., 2001; Sharron et al., 2000). The most efficient among these SIV coreceptors are the ones variously designated BOB / GPR15 and BONZO / STRL33/TYMSTR (Zhang et al., 2000).

However, the ability of HIV-1 strains to exploit alternate coreceptor on the surfaces of cell lines does not provide a true indication of coreceptor usage properties in vivo. The capacity of naturally expressed coreceptors (other than CCR5 and CXCR4) to support HIV infection of primary cell cultures may thus provide a stronger indication for their use in vivo (Willey et al., 2003; Yang et al., 2000). Celliers et al., in their study provided the evidence that some R5 isolates grew in Δ32/Δ32 (32 base pair deletion in CCR5) CCR5 PBMC in the presence or absence of AMD3100, a CXCR4 specific inhibitor, indicating that it uses a receptor other than CCR5 or CXCR4 on primary cells (Celliers et al., 2005). Similarly, Yang et al. characterized the expression pattern and coreceptor activity of STRL33 which indicated that it may be a relevant coreceptor in vivo (Yang et al., 2000). It is also reported that a maternal isolate used CXCR6 in addition to CCR5 and CXCR4 on indicator cell lines and replicated in CCR5 deficient peripheral blood mononuclear cells (PBMCs) in the presence of a CXCR4 inhibitor (Sharron et al., 2000; Zhang et al., 1998). Similarly, Lee et al., reported that CCR8 supported infection of primary thymocytes by particular HIV-1 isolates (Lee et al., 2000).

Identification of alternative HIV-1 coreceptor is hampered by the lack of suitable ligands although in some instances chemokines have been used to provide suggestive evidence (Celliers et al., 2005). Willey et al., 2003 identified a subset of HIV and SIV strains that are able to exploit an unknown coreceptor naturally expressed on brain microvascular endothelial cells (BMVECs) and astrocyte cultures and was inhibited by the chemokine vMIP-1, which implicated an unknown vMIP-1 receptor as the coreceptor involved. Similarly, the orphan receptor Chem R1 and CCR8 have been shown to be a receptor for CC chemokine I309 (Horuk et al, 1998; Endres et al., 1999). Given the importance of chemokines and their receptors in
inflammation, autoimmunity and the pathogenesis of AIDS, the identification and characterization of these proteins will be important to initiate approaches for therapeutic interventions.

Table 2.2 Alternate co-receptors used by HIV for entry (Taken from Clapham et al., 2002).

<table>
<thead>
<tr>
<th>Coreceptor</th>
<th>Ligand</th>
<th>HI V-1</th>
<th>HI V-2</th>
<th>SI V</th>
<th>References for coreceptor use</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCR1</td>
<td>MIP-1α, MIPF-1, MCP-3, RANTES</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>Bron et al. (1997); McKnight et al. (1998)</td>
</tr>
<tr>
<td>CCR2b</td>
<td>MCP-1, MCP-2, MCP-3</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>Doranz et al. (1996)</td>
</tr>
<tr>
<td>CCR3</td>
<td>Eotaxin, eotaxin-2, MCP-3, MCP-4, RANTES</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>Choe et al. (1996); Doranz et al. (1996)</td>
</tr>
<tr>
<td>CCR4</td>
<td>MDC, TARC, RANTES, MIP-1α</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>McKnight et al. (1998)</td>
</tr>
<tr>
<td>CCR5</td>
<td>MIP-1α, MIP-1β, RANTES, MCP-2</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>Alkhatib et al. (1996); Deng et al. (1996); Dragic et al. (1996)</td>
</tr>
<tr>
<td>CCR8</td>
<td>IL-8, NAP-2, ELR+ CXC</td>
<td>-</td>
<td>+</td>
<td>N</td>
<td>Ruck et al. (1997)</td>
</tr>
<tr>
<td>CCR9</td>
<td>TECK</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Choe et al. (1998)</td>
</tr>
<tr>
<td>CXCR2</td>
<td>IL-8, NAP-2, ELR+ CXC</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>Bron et al. (1997)</td>
</tr>
<tr>
<td>CXCR4</td>
<td>SDF-1</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>Feng et al. (1996)</td>
</tr>
</tbody>
</table>
As it is widely known that entry of HIV into target cells requires CD4 receptor and one of the coreceptors either CCR5 or CXCR4. CCR5 using (R5) virus predominates early in infection, but as the disease progresses, in about 50% of subtype B infected individuals mainly caucassian population, emergence of CXCR4 using (X4) virus appears and either coexists with R5 viruses. Emergence of CXCR4 variants is associated with more rapid decline of CD4^+ T-cells leading to faster progression of the disease and ultimately AIDS (Kamp., 2009; Ho et al., 2007; Xu et al., 2007; Coetzer et al., 2008; Tasca et al., 2008; Dash et al., 2008). The basis for X4 emergence late in infection or expansion of coreceptor or coreceptor usage use is yet not fully understood but different factors that can contribute to this phenomenon may include envelope gene mutations by chance, CCR5 bearing target cell limitation, and differential immune recognition of R5 and X4 viruses (Tasca et al., 2008; Ho et al., 2007). Furthermore, it is unclear whether X4 viruses evolve during the course of

---

**HIV Switch**

- Rarely or never used as a coreceptor (-); occasional use by few isolates (+); used by 5-20% of isolates (++); frequent use by many isolates or by major subgroup (for example, CXCR4 used by R5X4 and X4 isolates, indicated in bold (+++); major coreceptor used by predominant virus in vivo (for example, CCR5 used by HIV and SIV, indicated in bold) (+++). NT not tested. (Adapted from Clapham and Mac Knight, 2002).

* Table showing coreceptor usage patterns.*
infection or are transmitted or selected during the course of infection (Ho et al., 2007). Studies on coreceptor switching and genetic analysis of viral envelope sequences, suggests that viruses at intermediate stage between CCR5 and CXCR4 use are less fit then parenteral R5 populations. Intermediate stages in coreceptor switching may also differ in preference for CCR5 versus CXCR4 and these stages are designated as dual R (CCR5 preference) and dual X (CXCR4 preference) (Coetzer et al., 2008). Though these R5X4 intermediates are less fit cousins but still the switch from R5 to X4 requires these intermediates which arises by chance as a result of high virus replication and mutation rates (Tasca et al., 2008). Therefore, predicting the emergence of X4 has potential value for understanding pathogenesis, monitoring disease progression and making treatment decisions. Coetzer et al., 2008) provided evidence that envelope gene diverged at the rate of 5.4 % from 4.10 to 5.76 years of infection, with the V3 and V4 / V5 regions showing the greatest divergence and evidence of positive selection. These observations suggested that a decline in the fitness of R5 virus population may be one driving force that permits the emergence of R5X4 variants (Coetzer et al, 2008). Similarly, Kamp proposed a model which suggests that a dynamical change in the underlying environmental conditions in the host may be responsible for coreceptor switch (Kamp 2009). V3 loop is considered to be most important determinant of cellular tropism and presence of positive amino acid residues at position 11/ 24/ 25 and high net charge on V3 loop determines X4 tropism (Xu et al., 2007; Ho et al., 2007). Xu et al., 2007 suggested that the net charge, polarity of V3 loop and five V3 sites (especially sites 22, 12, 18, 13) are seven most important features for predicting HIV-1 coreceptor usage or phenotype (Xu et al., 2007). Such bioinformatic approaches based on V3 loop sequences have been developed for predicting HIV-1 biological phenotype. These approaches employed a multiple linear regression (Briggs et al., 2000), a neural network strategy (Resch et al., 2001), a machine learning method (Pillai et al., 2003), or a position specific scoring matrix (PSSM) (Jensen et al., 2003; Jensen et al., 2006), to predict HIV-1 coreceptor usage and phenotype. However, if we look in case of subtype C infection, coreceptor switch from CCR5 to CXCR4 is less common, the pathological significance of which is not
understood (Coetzer et al., 2007). Dash et al., 2008 identified two such HIV-1 subtype C envelope clones which in addition to CCR5 and CXCR4 used many other coreceptors as well. Genetic analysis revealed that perhaps it is not the V3 loop but longer V1, V2 and V4 loops in both of the molecular clones are longer due to insertion of several amino acid residues which generated potential N-linked glycosylation sites, apart from this both the clones also had NF-xB site variation in the LTR and non-truncated Rev and which perhaps contributed for its expanded coreceptor usage. Thus, it is not only the variable loops but also the length polymorphism and the degree of N-linked glycosylation sites which can influence coreceptor requirement of the viral strain. (Dash et al., 2008). The V3 loop of a vast majority of HIV-1 subtypes contains a single glycan, removal of which is strongly associated with loss of CCR5 but gain of CXCR4 (Ho et al., 2007; Dash et al., 2008).

Thus, the emergence of the X4 virus is essential for disease progression, as seen in subtype B infection and in the case of other viral subtypes, the absence of coreceptor switch in the subtype C infection could be suggestive of the relatively less pathogenic nature of subtype C. In contrast to this assumption, subtype C isolates are responsible for the largest number of infections in the world. Therefore, a better understanding of the conditions leading to coreceptor switch is especially of interest as CCR5 blockers have recently been licensed as drugs which suppress R5 viruses but do not seem to necessarily induce a coreceptor switch (Kamp 2009).

Significance of Coreceptor for transmission, replication and pathogenesis

Several genetic polymorphisms in HIV-1 chemokine coreceptors are important host genetic factors that are capable of influencing susceptibility to HIV-1 infection or affecting the rate of disease progression (Ometto et al., 2001). Since, CCR5 is the most important coreceptor involved in HIV entry therefore many studies have been done regarding the polymorphism associated within the coding and 5' regulatory region of the CCR5 gene (Carrington et al., 1999). The CCR5 gene has been mapped to the short arm of chromosome 3 amongst a group of genes that encodes multiple
The mutant allele CCR5 A32, characterized by 32 base pair deletion at position 554 to 585 in single coding exon of the CCR5 gene (Cohen et al., 1998), that induces a frameshift, a premature stop codon, does not produce a functional protein and hence is responsible for providing complete protection to individuals who are homozygous for this particular allele. While individuals heterozygous for this allele, in them slower progression to AIDS occurs (by 2-4 years of age) after HIV-1 seroconversion (Carrington et al., 1999; Deng et al., 2004; Roman et al., 2002; Struyf et al., 2000; Singh et al., 2008; Cohen et al., 1998; Ometto et al., 2001; Picton et al., 2010; Kulkarni et al., 2008). Picton et al., in 2010 reported that the CCR5 A32 allele occurs at a variable frequency of 4-15% in Caucasian populations, with an average of 10% in Europe and yet is rarely found in Asian or African populations (Picton et al., 2010). The CCR5 gene is composed of four exons, two introns and two CCR5 promoters: a weak upstream promoter (PU) and a stronger downstream promoter (PD). This difference in two promoters results in differential cell surface CCR5 expression levels (Picton et al., 2010). Thus, because of this phenomenon the individuals who are heterozygous for Δ32 mutations have diminished levels of CCR5 on the cell surface and low expression of CCR5 correlates with reduced infection of T-cells by R5 isolates of HIV-1 in vitro (Carrington et al., 1999).

Another mutation in the CCR5 open reading frame (ORF), a single nucleotide polymorphism (SNP) substituting a T to A at position 303 (CCR5m303), introduces a premature stop codon that prevents the expression of CCR5 coreceptor and hence conferring resistance to HIV-1 infection when associated with CCR5 A 32 (Ometto et al., 2001; Clapham et al., 2002; Roman et al., 2002). Thus, for both CCR5 mutation in the coding region (CCR5 Δ32 and CCR5m303), the resistance to HIV-1 is attributed to the expression of a truncated, non-functional CCR5 receptor (Roman et al., 2002).

Other mutations in the CCR2, CX(3)CR1, CXCL12 (SDF1) and CCL5 (RANTES) genes have been identified which are associated with resistance or susceptibility to HIV-1 infection and disease progression (Reiche et al., 2007). Since, it is known that CCR5 regulatory regions have several genetic variations which affects rate of disease progression, Ometto et al., reported that there are 10 polymorphic nucleotide positions in CCR5 promoter region 58934-59537 that have been identified. Combinations of the...
10 polymorphic sites specify 10 CCR5 promoter haplotypes: 4 are common (CCR5-P1 to CCR5-P4) and 6 are rare (CCR5-P5 to CCR5-P10). Polymorphism in the CCR5 regulatory region (promoter), CCR5-59653T, is genetically associated with a mutation in the coding region of the CCR2 gene leading to substitution of Valine to Isoleucine (CCR2-64I) (Ometto et al., 2001; Roman et al., 2002). HIV-1 infected individuals who possess the CCR2-64I allele progress less rapidly to disease (Ometto et al., 2001; Roger et al., 1998; Roman et al., 2002; Carrington et al., 1999). However, as compared to CCR5 A32 mutations, which are unique to Caucasian ethnic groups and results in non-functional chemokine receptor, the CCR2b-64I polymorphism is found in many different ethnic groups and encodes a conservative amino acid substitution which does not affect CCR2 protein function or its structure (Ometto et al., 2001). Moreover, Carrington et al., in 1999 reported that genes encoding CCR1-CCR5 are clustered on the same chromosome 3, but CCR2 and CCR5 are separated by approximately 14 kb explaining the near complete linkage disequilibrium between the two genes. Thus, suggesting that CCR2b-64I protection is as strong as and genetically independent of CCR5 A32 influence (Carrington et al., 1999; Roman et al., 2002).

Several SNPs located in the CCR5 promoter are reported to affect the expression of CCR5. One such polymorphism is the -2459 G/A located within the downstream promoter (PD or P1), and individuals who are homozygous for -2459 G (found on CCR5 P2-P4) allele exhibits lower CCR5 receptor density in CD14+ monocytes and hence slower progression to disease by 3-8 years than those who were homozygous for -2459 A (found on CCR5 P1) (Picton et al., 2010; Carrington et al., 1999). Similarly, Ometto et al., reported that persons homozygous for the 59029 A allele progressed to AIDS more rapidly than did persons with the 59020 G/G genotype (Ometto et al., 2001). Other mutations are also reported that are known to accelerate disease progression in HIV+ individuals. Roman et al., reported that G to A substitution at position 801 in a non-coding region of the CXCR4 ligand, SDF 1-β (SDF 1-3′A) has been associated with increased perinatal HIV-1 transmission. Similarly, a variant haplotype of the receptor to fractalkine, CX3CR1, highly expressed in brain tissues and in activated peripheral blood lymphocytes is described to accelerate disease progression. Two SNPs results in valine to isoleucine and
threonine to methionine changes at codons 249 (CX3CR1-V249I) and 280 (CX3CR1-T280M), respectively and thus patients homozygous for CX3CR1-I249M280 progress more rapidly (Roman et al., 2002). Singh et al., reported that the polymorphisms in the promoter region of CCR5 alters the risk of mother to child transmission (MTCT) through the role of CCR5 as dominant coreceptor used by HIV-1 for primary infection of infants. In contrast polymorphism that alters the expression of CX3CR1 likely affects MTCT by modulating the early immunologic response to viral exposure (Singh et al., 2008).

However, there are certain mutations reported that have no effect on the rate of progression to AIDS for instance, the codon 335 mutation resulting in an alanine to valine (A335V) substitution present at a higher frequency in African American population in comparison to Caucasians has no effect on disease progression (Ansari-Lali et al., 1997; Carrington et al., 1997). Thus, a better understanding of the role played by genetic defects in host genes in response to HIV-1 exposure will contribute towards better understanding of protective immunity to HIV-1 and of the disease process in HIV-1 infected individuals (Picton et al., 2010). This in turn helps for the development of therapeutic strategies that targets the interaction of HIV-1 envelope with different coreceptors (Carrington et al., 1999).

The Entry Inhibitors

The main objective in the field of HIV is the discovery of drugs that will combat the disease. Existing antiretroviral drugs (ARV) are categorized according to their mode of action into three main groups: (1). the nucleoside reverse transcriptase inhibitors (NRTI), (2). the non-nucleoside reverse transcriptase inhibitors (NNRTI) and (3). the protease inhibitors (PI) (De Clercq et al., 2002). The combination of above ARV drugs is now administered to produce more efficient treatment / therapy known as HAART (Highly active antiretroviral therapy), which has significantly decreased mortality and morbidity in the developed world (Krambovitis et al., 2005). Despite the fact that the current antiviral treatments have improved prognosis but still these treatment strategies can give rise to significant side effects such as drug resistance,
toxicity and the frequency of treatment change in drug experienced patients (Kuhmann et al., 2008; Krambovitis et al., 2005; Este et al., 2007). Therefore, there is a need for the development of new classes of anti-HIV drugs that could either replace or augment existing treatment regimens (Kuhmann et al., 2008).

Several steps in HIV replication are potential targets of interventions. One of the important steps is the entry of HIV virus into the target cells which comprises of 3 steps: (1) attachment of the viral gp120 to the CD4 T-cell receptor (2) binding of the gp120 to CCR5 or CXCR4 coreceptors and (3) fusion of the viral and cellular membranes (Briz et al., 2006; Tilton and Doms., 2010). Drugs that block HIV entry are collectively known as entry inhibitors (Tilton and Doms, 2010) and they comprises of multiple mechanism of action depending on the stage of entry process. The drugs currently under development fall under 3 categories: gp120-CD4 binding inhibitors, gp120-coreceptor binding inhibitors and membrane fusion inhibitors, which interfere with gp41 conformational changes (Veiga et al., 2006; Ryser et al., 2005; Briz et al., 2006). Two entry inhibitors, Maraviroc: coreceptor inhibitor (Kuhmann et al., 2008; Muniz Medina et al., 2009) and enfuvirtide: fusion inhibitor (Veiga et al., 2006) has been approved for the treatment of HIV-1 infection, and a number of agents are in development (Tilton and Doms, 2010).
Chemokine analogues: a first generation of coreceptor inhibitors

Certain chemokines act as natural antagonists of human immunodeficiency virus (HIV) by blocking key viral coreceptors, such as CCR5 or CXCR4, on the surface of susceptible cells (Nardese et al., 2001; Simmons et al., 1997; Alkhaidi et al., 1997). Importance of CCR5 coreceptor in viral transmission and as a target for therapeutic intervention was underscored by discovery of 32 base pair deletion in human CCR5 coding region as described earlier. Mutations that block CCR5 expression are well tolerated in both humans and mice, but deletion of CXCR4 expression is lethal in mice (M. Strizki and Donald E. Mosier, Book- inhibitors that target gp120 interactions with coreceptors). Cocchi et al., in 1996 provided evidence which suggested that CD8+ T lymphocytes are involved in the control of Human immunodeficiency virus (HIV) infection invivo, either by cytolytic mechanisms or by the release of HIV suppressive factors (HIV-SF). The natural
chemokines RANTES, MIP-1α and MIP-1β which are the natural ligands for CCR5 coreceptor were identified as the major HIV-SF produced by CD8+ T cells. It has also been shown that increased expression or segmental duplication of gene encoding CCL3L1 (MIP-1α), is associated with marked enhanced HIV susceptibility (Gonzalez et al., 2005). These data opened up an avenue that natural chemokines can also have relevance for the prevention and therapy of AIDS (Cocchi et al., 1995; Baggioolini et al., 1997). Though native chemokine binding to specific coreceptors is sufficient to inhibit HIV infection but at the same time these chemokines can give rise to adverse side effects due to leukocyte activation (Baggioolini et al., 1997).

Chemokines are thus engineered to enhance their natural anti-HIV properties. The first engineered anti-HIV chemokine to be described was N-terminally modified RANTES analogues. These molecules include an N-terminal truncation (RANTES), a variant extended by methionine (Met-RANTES), a rationally designed analogue of Met-RANTES (aminooxypentane - RANTES or AOP - RANTES) and a further optimized molecule based on AOP-RANTES (N-nonanoyl- RANTES or NNY-RANTES). Unlike native chemokines, AOP and NNY-RANTES have the additional capacity to inhibit the recycling of internalized CCR5, resulting in remarkably profound and prolonged sequestration of coreceptors (Pastore et al., 2003; Simmons et al., 1997).

**Unusual Inhibitory Mechanisms:**

There are different inhibitory mechanisms of CCR5 inhibitors as proposed by Kuhamann et al., 2008. Inhibition can either be achieved by:

(a). Steric blockade of the coreceptor i.e. engaging the receptor and thereby blocking the viral access example monoclonal antibodies (Figure 2.7a);

(b). Allosteric inhibition i.e. altering the conformation of the receptor (indicated by yellow / red hatching in Figure 2.7b below) so that it is not usable as a coreceptor example small molecule inhibitors; or

c. Removal of the receptor from the cell surface, either by inducing receptor sequestration or by blocking receptor expression example chemokine analogues (Figure 2.7c).
Steric blockade e.g., MAbs

Allosteric modulation e.g., small molecules

Removal from cell surface e.g., RANTES analogues, RNAI

Figure 2.10 Inhibitory mechanisms CCR5 inhibitors. (Taken from Kuhamann and Hartley (2008). “Targeting Chemokine receptors in HIV: A status report”. 48: 425-61).

Examples:

vMIP-1

Macrophase Inflammatory protein (MIP) belongs to the family of chemo tactic cytokines known as chemokines. Sherry et al. (1998) demonstrated 2 protein components of MIP1, called by them α and β. In humans there are two major forms, MIP-1α and MIP-1β that are now officially named CCL3 and CCL4, respectively.

Both the major factors are produced by macrophages after they are stimulated with bacterial endotoxins (Sherry et al., 1998). They activate human granulocytes (neutrophils, eosinophils and basophils) which can lead to acute neutrophilic inflammation and also induces synthesis and release of other pro-inflammatory cytokines such as interleukin 1 (IL-1), IL-6 and TNF-α from fibroblasts and macrophages. The genes for CCL3 and CCL4 are both located on human chromosome (Irving et al., 1990). Many studies have also shown that in addition to roles in...
regulating leukocyte trafficking, the chemokine systems are linked to the biology of infectious disease. This prompted intensive investigation of viral chemokine analogues that have been identified in herpesvirus and poxvirus genomes. Large DNA virus such as pox virus and herpes virus encode protein that target and exploit the chemokine system of their host (Luttichau et al., 2000). The Kaposi-Sarcoma associated herpes virus (KSHV) encodes three predicted CC type chemokines (vMIP-1, vMIP-2 and vMIP-3) (Dairaghi et al., 1999; Endres et al., 1999). Previous studies have shown that vMIP-1 and vMIP-2 interacts with cellular chemokine receptors and stimulates angiogenesis. These studies demonstrate the ability of vMIP-1 to bind to CCR5 and inhibit HIV infection. Although vMIP-1 and vMIP-2 share significant sequence similarity with human chemokines (43% and 52 %, respectively, with MIP-1α), vMIP-3 is more distantly related (Stine et al., 2000). Celliers reported that vMIP-1 could bind to a range of coreceptors, it is an agonist to CCR8 (Celliers et al., 2005; Dairaghi et al., 1999; Endres et al., 1999), while antagonist to CXCR6, CCR3, GPR15 and GPR1. This chemokine has been shown to prevent HIV-1 infection of some isolates on primary cells when the major coreceptors are absent or blocked, suggesting that vMIP-1 might be useful in identifying alternate coreceptors that could be used by HIV-1 isolates (Willey et al., 2003).

vMIP-2

The viral macrophage inflammatory protein (73 amino acids) is encoded in the genome of KSHV (Lindow et al., 2003); Luttichau et al. in 2000 have reported that viral MIP-2 acts as an antagonist or blocks a number of the chemokine receptors such as CCR1, CCR2, CCR4, CCR5, CXCR3, CXCR4, CX3CR1 and lymphotactin receptor CXCR1. Viral encoded chemokines are important in the study of viral pathogenesis, but they can also be used as tools in the investigation of specific chemokine receptors. vMIP-2 has been shown to reduce the inflammatory response in small animal models (Luttichau et al., 2000). Stine et al., (2000) reported that vMIP-2 binds to CCR3 (Lindow et al., 2003) and CCR8 (Sozzani et al., 1998) as an agonist. vMIP-2 is said to block infection of HIV-1 on a CD4^ cell line expressing CCR3 and to a lesser extent on one expressing CCR5, whereas both vMIP-1 and vMIP-2 partially
inhibited HIV infection in peripheral blood mononuclear cells (PBMCs) (Boshoff et al., 1997). vMIP-2 is a selective chemoattractant for T-helper cells (Th2 cells) and it is this capacity of vMIP-2 by which it is able to subvert the host immune response. However, it was found to be inactive on other leukocytes including Th1 cells, dendritic cells and natural killer cells (NK cells) (Sozzani et al., 1998). Weber et al., (2001) showed that vMIP-2 blocks the RANTES – induced firm arrest of monocytes and Th1 type T cells but promotes arrests of eosinophils and Th2 type T cells. This finding explained the functional role of vMIP-2 in leucocyte recruitment. Zhou et al., 2000 have obtained synthetic peptide derived from the N-terminus of vMIP-2 (amino acids 1-21). In contrast to full length protein, which recognizes CXCR4 and CCR5, this peptide strongly binds to CXCR4, but not CCR5. The first five amino acids have been shown to be important for this specific receptor interaction, and assumption is that vMIP-2 probably interacts with other chemokine receptors with different sequence and conformational determinants. The peptide selectively prevents CXCR4 signal transduction and coreceptor function in mediating the entry of HIV-1 isolates.

I-309

In addition to CCR5 and CXCR4 other chemokine receptor such as CCR8 can act as viral coreceptors (Cota et al., 2000; Schramm et al., 2000). CCR8 chemokine receptor, previously known as TERI, Chem R1 or CKR-L1 is expressed on monocytes and type-2 T-lymphocytes (Haque et al., 2001) in particular on thymus (Zingoni et al., 1998). Th lymphocytes have been functionally separated into type 1 (Th1) and type 2 (Th2) subsets based on their ability to produce discrete set of cytokines. Th1 subsets produce IL-2, IFN-γ, TNF-α and lymphotoksin and participate in cell mediated immunity. The Th2 subsets produce IL-4, IL-5, IL-6, IL-10 and IL-13 and are associated with humoral responses (Zingoni et al., 1998). CCR8 is the sole receptor for the human CC chemokine I-309 and for the viral monocyte inflammatory protein (vMIP-I). Lee et al., in 2000 for the first time provided evidence that CCR8 functions as an HIV-1 coreceptor on primary human cells and suggests that CCR8 may contribute to HIV-1 induced thymic pathogenesis. Several studies have reported that I-309 has been found to inhibit HIV-1 envelope mediated cell-cell fusion and virus
infection (Haque et al., 2001; Horuk et al., 1998; Lynn et al., 2001). One possible mechanism for this inhibition may be the ability of chemokines to activate NK (natural killer cells) cells which express CCR4 and CCR8, as shown by Inngjerdingen et al., 2000 suggesting that these receptors are not exclusive for Th2 cells. 1-309 / CCL1 and its murine homologue, mCCL1 (TCA3) is secreted by activated monocytes and T-lymphocytes and findings of Haque et al. showed that 1-309 is the principal monocyte chemoattractant secreted by human umbilical vein endothelial cells (HUVECs) when incubated with apolipoprotein portion of the atherogenic lipoprotein. Horuk et al., 1998 showed that 1-309 can induce a transient calcium ion flux in cells expressing CCR8. Moreover, Spinetti et al., in 2003 showed that vMIP-1 and CCL1 are able to trigger a CCR8 mediated rescue from dexamethasone induced apoptosis via an extracellular regulated kinase (ERK) dependent mechanism.

Small molecule coreceptor inhibitors

Known small molecule inhibitors:

The first small molecule CCR5 antagonist described in the literature was TAK-779, could not be developed as an anti-HIV-1 agent because of its poor bioavailability (Baba et al., 2005; Kuhmann et al., 2008). Since then, a number of small molecule CCR5 inhibitors (Table 2.3) have been described and some of which have entered clinical trials and one Maraviroc from Pfizer have been approved for clinical use. Maraviroc (UK-427,857) have favorable pharmacological properties and it blocks binding of gp120 to CCR5 to prevent membrane fusion events necessary for viral entry. Maraviroc didnot effect CCR5 cell surface levels or associated intracellular signaling, no detectable invirto cytotoxicities have been observed and is highly selective for CCR5 (Dorr et al., 2005). TAK-220 and TAK-652, novel orally bioavailable CCR5 antagonists, are successors of TAK-779. Both the compounds are highly potent inhibitor of R5 HIV-1 replication and hence are considered to be potential candidates for clinical development (Baba et al., 2007). It was reported by Baba et al., 2005 that TAK-652 have favorable pharmacokinetics profile in humans.
and inhibited the binding of RANTES, MIP-1 and MCP-1 to CCR5 and CCR2b expressing cells. Duong et al., 2007 reported that DCM205 represents a promising new class of HIV entry inhibitor that can also be used as strategy in the prevention of HIV-1/ AIDS. This molecule is found to be effective against CCR5, CXCR4 and dual tropic laboratory adapted and primary strains of HIV-1 (Duong et al., 2007). Garg et al., 2008 synthesized a series of galactosyl ceramide (GalCer) analogues, which is a cofactor in HIV envelope binding to V3 loop and found to inhibit HIV entry in a coreceptor independent manner as they blocked the infection via CCR5, CXCR4 and dual tropic viruses (Garg et al., 2008). These compounds inhibited HIV entry at early steps were found to be ineffective if added at post viral entry steps. However, one such post attachment inhibitor ibalizumab has shown activity in phase 1 and 2 trials and hence further studies are anticipated (Kuritzkes, 2009). SCH-C (SCH 351125) is a small molecule antagonist of the human immunodeficiency virus type 1 (HIV-1) coreceptor CCR5. Tremblay et al., have studied anti-HIV interactions of SCH-C with other anti-retroviral agents in vitro, and found synergistic interactions against all the viruses tested, some of which harbored resistance mutations to reverse transcriptase and protease inhibitors. These findings suggested that SCH-C may be a useful anti-HIV drug in combination regimens and that a combination of chemokine coreceptor inhibitors may be useful in the treatment of multidrug viruses (Tremblay et al., 2002).
TABLE 2.3 Selected small molecule CCR5 inhibitors that have been evaluated in vitro. (Taken from Kuhmann and Hartley (2008). “Targeting Chemokine receptors in HIV: A status report”. 48: 425-61).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Manufacturer Company and clinical development of small molecule inhibitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMPD-167</td>
<td>From Merck Inc.; used in proof-of-principle studies in macaques; licensed to the International Partnership for Microbicides for development as a topical Microbicides</td>
</tr>
<tr>
<td>TAK-779</td>
<td>From Takeda Inc.; first in class; no longer in clinical development</td>
</tr>
<tr>
<td>TAK220</td>
<td>From Takeda; in clinical development</td>
</tr>
<tr>
<td>TAK625</td>
<td>From Takeda; in clinical development</td>
</tr>
<tr>
<td>AD101</td>
<td>From Schering-Plough; no longer in clinical development</td>
</tr>
<tr>
<td>SCH-C</td>
<td>From Schering-Plough; no longer in clinical development</td>
</tr>
<tr>
<td>Vicriviroc</td>
<td>From Schering-Plough; in Phase 2/3 clinical trials</td>
</tr>
<tr>
<td>Aplaviroc</td>
<td>From GlaxoSmithKline; Phase 2/3 trials halted due to hepatotoxicity</td>
</tr>
<tr>
<td>Maraviroc</td>
<td>From Pfizer, recently approved in the United States for treatment-experienced patients.</td>
</tr>
</tbody>
</table>

**Binding site for coreceptor inhibitors:**

Site directed mutagenesis and molecular modelling of CCR5 have identified that all known small molecule CCR5 inhibitors bind in the hydrophobic pocket formed by the seven transmembrane helices (Kuhmann et al., 2008; Ji et al., 2007). These CCR5 antagonists compete for binding to the same pocket, although they may interact with different residues in the helices (Ji et al., 2007). Because CCR5 antagonists sit deep in the pocket, thus CCR5 antagonist inhibit HIV entry via allosteric mechanism i.e stabilization of a conformation of CCR5 that is not compatible with HIV gp120 binding or the binding site for inhibitor does not overlap with the binding site of gp120 thus works in a non-competitive manner (Watson et al., 2005).
Mechanism of HIV-1 escape from small molecule CCR5 inhibitors:

The predominant pathway for the evolution of denovoresistance involves the continued use of CCR5 in an inhibitor insensitive manner, but a more common route to treatment failure involves the expansion of pre-existing CXCR4 using viruses that are naturally insensitive to these inhibitors (Moore et al., 2001). At a molecular level, gp120 binds to CCR5 at two distinct sites: the ‘bridging sheet’ and residues at the V3 base bind to the tyrosine sulphated N-terminus of CCR5 and the V3 crown interacts with ECL2 (Pugach et al., 2008; Pugach et al. 2009; Berro et al., 2009). The envelope sequence changes in the resistant variants might directly or indirectly alter the conformation of V3, such that its crown no longer interacts with CCR5 ECL2 to mediate infection and hence resistant viruses becomes more dependent on the interaction with the N-terminus (Moore, et al., 2001; Nolan et al., 2008). However, there are studies which reports that there are certain resistant mutants which shows no V3 sequence changes, which implies the existence of alternative genetic pathways to the same phenotype (Pugach et al., 2009). For instance Pugach et al., in 2008 showed that two resistant variants (CC101.19 and D1/85.16) derived from the same primary isolate (CC1/85), continued to use CCR5 as the coreceptor for their entry, despite the fact that one resistant variant had sequence changes in V3 loop (CC101.19) while other variant (D1/85.16) has no sequence changes in V3 loop but has three substitutions in the gp41 fusion peptide (Berro et al., 2009) which perhaps create a similar rearrangement of the components of its CCR5 binding site (Pugach et al., 2008).

In order to test their hypothesis, that R5 viruses undergo phenotypic evolution to X4 usage by blocking R5 coreceptor by using specific antagonist in vivo, Trkola et al., conducted some in vitro experiments to characterize the escape pathways used by HIV-1 isolate CC1/85 when replicating in PBMCs in presence of AD101, an R5 inhibitor. Their findings revealed that the AD101 escape mutant did not use CXCR4 but instead gained the ability to use CCR5 in an AD101 insensitive manner (Trkola et al., 2002). This suggests, that main path of resistance to CCR5 antagonists appears to be continued use of CCR5 in the presence of the inhibitor i.e. inhibitor-CCR5 complex while also interacting with free CCR5 (Nolan et al., 2008; Berro et al., 2009).
two resistance mechanisms most relevant to CCR5 inhibitors have been described as competitive and non-competitive. Competitive resistance is defined as resistance that results in a shift in the IC50 (Inhibitory concentration) of an inhibitor to a higher concentration, although complete inhibition may still be achieved at sufficient inhibitor concentrations. It would arise from more efficient use of inhibitor-free CCR5 to gain access to target cells, enabling HIV to scavenge low levels of inhibitor to CCR5. Competitive resistance could arise from the envelope glycoprotein either by developing a higher affinity for CCR5 or from acquiring more rapid fusion kinetics after CCR5 engagement. For non-competitive resistance, inhibition IC50 values are equivalent to those for the fully sensitive virus, but at maximum (plateau) levels inhibition remains incomplete. This implies that the resistant envelope glycoprotein complex has adapted to use the inhibitor bound form of CCR5 as coreceptor. The level of residual inhibition termed the plateau or maximum percentage inhibition (MPI) value once the effect has saturated reflects the efficiency with which the virus can use the inhibitor-CCR5 complex relative to inhibitor – free CCR5 (Kuhmann et al., 2008; Moore et al., 2009). Pugach et al., in 2009 reported that manifestation of resistance to small molecule CCR5 inhibitors varies and that CCR5 density is the most relevant host cell factors that accounts for these differences. They showed that MPI value is an inverse function of the CCR5 cell surface density i.e. high level of resistance (a low plateau or low MPI value) is observed only when the target cells express high levels of CCR5. Moreover, CCR5 cell-surface levels are highly variable in the population and depend on the multiple genetic factors, including polymorphism in the CCR5 promoter and the expression of its chemokine ligands (Pugach et al., 2009). Anastassopoulou et al., in 2007 reported that specific amino acid substitutions conferring AD101 resistance did cause a fitness loss when experimentally introduced into a sensitive clone, but in the naturally selected escape mutant they are probably compensated for by other changes and this work may help understand the development and management of resistance to CCR5 inhibitors now being evaluated clinically to treat HIV-1 infection.
Example of small molecule coreceptor inhibitors

CCR5 Inhibitor: Maraviroc

In 2007, the US Food and Drug Administration (FDA) approved Maraviroc for the treatment of HIV-1 infection, making it first (and to date only) CCR5 antagonist available for clinical use. Many recent studies also indicated that treatment with MVC also resulted in larger increases in CD4+ T-cell counts that can be attributed to its antiretroviral effect alone (Cooper et al., 2010; Asmuth et al., 2010) but the mechanisms responsible has not been identified. However, a study conducted by Funderburg et al., 2010 found that it was decrease in immune activation that actually correlated with increased CD4+ T-cell gains (Funderburg et al., 2010). It has also been shown that topical microbicide gel formulation of CCR5 inhibitor MVC fully protects humanized mice from HIV-1 challenge via vaginal route which is the predominant mode of viral transmission (Neff et al., 2011).

Approval stipulated that maraviroc be used only in treatment-experienced patients who have viremia with a virus using the CCR5 receptor (R5 virus)—the latter being a critical limitation because maraviroc has no antiviral activity against non-CCR5-using viruses (Sax, 2010). Tilton et al., reported the case of a patient who developed high level resistance to Maraviroc (MVC) during the course of treatment, although the viruses obtained from this patient were sensitive to most other CCR5 antagonists. Mutational changes within the V3 and V4 loops contributed for this phenomenon. The finding of this study reported that these MVC resistant envelopes dependent not only on N-terminus of CCR5 (like most other CCR5 antagonist resistant virus) but also on the drug modified form of ECL2 of CCR5 (Tilton et al., 2010).

CCR5 Inhibitor: TAK-779

The first non-peptide CCR5 antagonist found was TAK-779 with limited antiretroviral activity against R5 viruses (Baba et al., 1999). TAK-779 binds to CCR5 coreceptor with high affinity and also binds to CCR2 but with 20 fold lower affinity.
(Takami et al., 2002). This compound (N,N-dimethyl-N-(4-[[[2-(4-methylphenyl)-6,7-dihydro-5H benzocyclohepten8yl]carbonyl]amino]benzyl)-tetrahydro-2H-pyran-4-aminium chloride) (Dragic et al., 2000) has a molecular weight (Mw 531.13) and antagonizes the binding of RANTES to CCR5 expressing cells. In a study conducted by Baba et al., they found that TAK-779 inhibited the replication of CCR5 specific (R5) HIV-1 Ba-L strain in MAGI-CCR5 cells with EC50 and EC90 of 1.2 and 5.7 nM, respectively. However, TAK-779 did not affect CXCR4 specific (X4) HIV-1 HXB2 strain replication at concentrations up to 20 μM. (Este and Telenti, 2007). To further facilitate the development of entry inhibitors as anti-viral drugs, Dragic et al. explored that TAK-779 inhibits HIV-1 replication at the membrane fusion stages by blocking the interaction of the viral surface glycoprotein gp120 with CCR5. They found that extracellular domain of CCR5 did not affect the antiviral action of TAK-779, however alanine scanning mutagenesis of the transmembrane domains revealed that the binding site for TAK-779 on CCR5 is located near the extracellular surface of the receptor, within a cavity formed between transmembrane helices 1, 2, 3 and 7 (Dragic et al., 2000; Dragic et al., in their study also postulated that methylphenylbenzocycloheptenyl group of TAK-779 inserts into the hydrophobic pocket, allowing its charged moiety to contact its polar residues in the extracellular domain of CCR5. Takami et al., demonstrated the protective effects of centrally or systemically administered TAK-779 on ischemic brain using a middle cerebral artery (MCA) occlusion model in mice. However, this compound could not be developed further as an anti-HIV-1 agent because of its poor bioavailability (Baba et al., 2005; Kuhmann et al., 2008).
CXCR4 Inhibitor: AMD3100

The CXCR4 receptor is expressed much more broadly than chemokine receptors in general. In contrast to many chemokine receptors, the CXCR4 receptor is only activated by a single chemokine ligand, stromal cell derived factor (SDF-1 also called as CXCL12). The importance of X4 receptor is emphasized by the fact that targeted deletion of either the gene for CXCR4 or for its ligand in both cases leads to embryologic lethality (Rosenkilde et al., 2004). The development of CXCR4 inhibitors has proceeded more slowly than that of the CCR5 antagonists. One problem unique to CXCR4 inhibitors is that whereas R5 viruses are found on their own in 50% or more of patients, X4 viruses are present as mixtures together with R5 viruses. Inhibition of just the X4 component of the virus population may not lead to measurable declines in overall plasma viremia, thereby complicating assessment of drug activity. Co-administration of CCR5 and CXCR4 antagonists might be uniquely effective, if the safety of such combinations can be established (Krutizikes, 2009).

AMD3100, which is composed of two 1,4,8,11-tetrazacyclotetradecane (cyclam) moieties connected by conformationally constraining heteroatomic phenylenebismethylene linker (Rosenkilde et al., 2004; Rosenkilde et al., 2007). It is a symmetric bicyclam non-peptide antagonist of CXCR4 receptor, that uniquely inhibits the entry of human immunodeficiency virus type 1 (HIV-1) and HIV-2 replication (Rosenkilde et al., 2004; Hendrix et al., 2000). AMD3100 inhibits the binding and function of SDF-1 with high affinity and potency (Rosenkilde et al., 2004; Rosenkilde et al., 2007). AMD3100 does not bind to other physiologically
relevant chemokine receptor, CCR5, which mediates the entry of macrophage tropic (R5) HIV. The specificity of AMD3100 was further confirmed by their ability to block entry of HIV-1 in GHOST CXCR4 transfected cells with no effect on GHOST CCR5 cells (Owen et al., 2002). In vitro, Hendrix et al., reported that AMD3100 inhibits HIV replication with a 90% effective concentration of below 1 to 10 ng/ml and provides complete protection of lymphocytes and monocytes against HIV at concentrations of 10 to 30 ng/ml. Mutational substitutions at 16 positions located in TM-III, 1V, V and VII lining the main ligand binding pocket of the CXCR4 receptor identified 3 acid residues: Asp171 (Asp IV :20), Asp262 (Asp VI :23), and Glu288 (Glu VII :06) as the main interaction point for AMD3100. Molecular modeling suggests that one cyclam ring of AMD3100 interacts with Asp171 in TM-IV, whereas the other ring is sandwiched between the carboxylic acid groups of Asp262 and Glu288 from TM-VI and VII, respectively (Rosenkilde et al., 2004; Rosenkilde et al., 2007).

During the development of AMD3100, it was discovered that the compound increases white blood cell counts in the blood and importantly it mobilizes stem cells from the bone marrow. Thus, AMD3100 in combination with G-CSF, can be used for stem cell mobilization for auto transplantation (Rosenkilde et al., 2004). However, AMD3100 does not exhibit oral bioavailability presumably due to the positive charge of each ring at physiological pH.

A number of compounds are being developed to specifically target each of these steps leading to virus entry and some compounds have reached early clinical development. Conversely, agents such as CCR5 antagonist TAK-779 and the CXCR4 antagonist AMD3100 are no longer being thought as relevant anti-HIV agents but have given way to new analogues with improved properties (Este and Telenti, 2007).
An increasing large number of antiviral agents that prevent entry of Human Immunodeficiency virus (HIV) into cells are in preclinical and clinical development. T-20 or Enfuvirtide (formerly DP-178) is the first HIV-1 entry fusion inhibitor approved by U.S. Food and Drugs Administration (FDA) for the treatment of HIV-1 infection (Veiga et al., 2006; Chinnadurai et al., 2005). T-20 (YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF) corresponds to a linear 36-amino-acid sequence corresponding to residues 643-678 (Aquaro et al., 2006) within HR2 from HIV-1 HXB2 envelope glycoprotein (Chinnadurai et al., 2005; Reeves, et al., 2002). T-20 is a synthetic peptide that potently inhibits replication of HIV-1 by interfering with the transition of the transmembrane protein, gp41, to a fusion active state following interactions of surface glycoproteins gp120, with CD4 and coreceptor molecules displayed on target cell (Aquaro et al., 2006). T-20 peptides are proposed to interact with a target sequence within HR1, inhibiting association with native HR2 and preventing apposition of the viral and cellular membranes (Chinnadurai et al., 2005; Melikyan et al., 2000; Wadhwa et al., 2007).

A few studies have explored the enfuvirtide resistance profile and defined mutations at 36-45 in the HR1 domain of HIV-1 gp41 for enfuvirtide (Aquaro et al., 2006; Reeves et al., 2004). Several studies have investigated the variability of HR1
region in HIV-1 untreated infected patients to evaluate the frequency of resistance to T-20. The results demonstrated that HR1 region is usually highly conserved. However, some mutations in the GIV motif and in the hydrophobic pocket were identified, which affected HIV affinity to inhibition to T-20 (Chinnadura et al., 2005). The ability of T-20 to block entry by primary HIV-1 strains varies markedly among isolates. Importantly the efficiency with which T-20 blocks entry is affected by the affinity of CD4-triggered gp120 for the coreceptor, fusion kinetics, and receptor and/or coreceptor density as well as other strain and cell dependent factors (Yi et al., 2008). Early reports suggested that T-20 sensitivity might be greater for strains that use CXCR4 than those that use CCR5, although later reports did not support a clear dichotomy (He et al., 2006; Reeves et al., 2002; Celliers et al., 2004). The fusion process may occur at a range of speeds (fast to slow) depending on the concentration of CD4 and coreceptors on cell surfaces. The lower the receptor/coreceptor concentration on the cell surface, the slower the fusion process and the more time for the inhibitor to bind and hence more sensitivity to T-20 (Moyle, 2003). Reeves et al., in 2002 determined another important factor i.e. envelope coreceptor affinity that too can modulate T-20 sensitivity. This group found that envelopes that bound to coreceptor with high affinity were more resistant to T-20 than those that bound to coreceptor with reduced affinities, it is because increased coreceptor affinity resulted in faster fusion kinetics and hence reducing the kinetic window during which envelope is sensitive to T-20 (Reeves et al., 2002). Thus, receptor expression levels and envelope receptor affinity are cellular and viral determinants, respectively that impact viral sensitivity to T-20 (Reeves et al., 2002). The effectiveness of T-20 for HIV-1 subtype B isolates is well-defined, whereas less is known regarding its activity against HIV-1 subtype C isolates. HIV-1 subtype C now accounts for more than half the new infections globally and, as such, these isolates should be included in an evaluation of any new antiviral compounds (Celliers et al., 2004).
gp120 is a heavily glycosylated protein, with approximately half of its mass being N-linked carbohydrates. Three types of glycans are found in gp120: high-mannose glycans composed of 7 to 9 terminal mannose residues, complex glycans containing terminal sialic acid residues, and hybrid glycans, which are a mixture of both. High-mannose glycans on gp120 are also targets of glycan specific agents, such as lectins. Lectins are proteins of non-immunoglobulin nature, capable of recognition of and reversible (Kabamba B. A et al. 2011, 2010) binding to carbohydrate moieties of complex glycoconjugates without altering the covalent structure of any of the recognized glycosyl ligands. Lectins can be found across a wide variety of different species in nature including prokaryotes, sea corals, algae, fungi, higher plants, invertebrates and vertebrates and are involved in many biological processes, among them host-pathogen interactions, cell-cell communication, induction of apoptosis, cancer metastasis and differentiation, targeting of cells, as well as recognizing and binding carbohydrates. Several lectins have been identified in recent years that potently block the infectivity of viruses, such as HIV and many other classes of enveloped viruses by interacting with the carbohydrates present on the viral envelope. Three lectins namely Griffithsin (GRFT), cyanovirin (CV-N) and scytovirin (SVN) are most widely studied.

Griffithsin is a lectin isolated from the red algae Griffithsia sp. found in the coastal waters of New Zealand. GRFT is a 121-amino-acid dimeric protein and 13-kDa molecular mass lectin. GRFT exists exclusively as a 25 kDa dimer and has a domain-swapped structure in which two β-strands of one monomer combine with 10 β-strands of the other monomer to form a β prism of three four-stranded sheets (Kabamba B. A et al. 2011, 2010). The homodimer has six carbohydrate binding pockets, 3 located at each of the opposite ends of the double prism homodimer. GRFT binds oligomannose glycans, targeting terminal mannose residues found on Man5-9-GlcNAc2 (Dana H et al. 2012).
Cyanovirin isolated from blue green algae *Nostoc sp.* is a 101 amino acid protein that has shown potent anti-HIV activity in vitro. Scytovirin isolated from a cyanobacteria *Scytonema sp.* consists of a single amino acid chain with 95 amino acids.

Many research groups have investigated the antiviral activity of CV-N and GRFT against different classes of enveloped viruses. CV-N is effective against Ebola virus, influenza A and B, hepatitis C virus, measles virus, herpes simplex virus type-1 (HSV-1) and human herpes virus 6 (HHV-6).

Apart from HFV-1, GRFT can prevent hepatitis C virus infection *in vitro* and mitigate hepatitis C virus infection *in vivo*. Also, GRFT displayed low nanomolar activity against SARS-related coronavirus (SARS-CoV) (Dubuisson, *J et al* 2011). In fact, GRFT was active against coronavirus strains that utilize protein-protein interactions for viral targeting (e.g., ACE2 as a cellular receptor, SARS-CoV, and HCoV-NL63) and those that utilize protein-carbohydrate interactions for viral attachment (i.e., α-2,3-linked sialic acid moieties, IBV-CoV, and HCoV-OC43). Mao *X et al* 2013 have shown GRFT is against Japanese encephalitis virus (JEV) infection. (Dana Het *al* 2012).
As a result of their ability to block HIV-1 entry in vitro, GRFT, CV-N and SVN have been proposed as potential microbicides to prevent the sexual transmission of HIV-1. The search for HIV-1 microbicide is driven by the challenges encountered in making a vaccine. The lectins selectively target viruses and prevent its entry into the susceptible cells hence preventing integration. These compounds can be readily and cheaply produced. The commensal lactobacilli can be engineered to produce these compounds and can be used to colonize vaginal mucosa and create environment hostile for HIV-1 transmission. (Puesch et al 2005)

Although these compounds are not HIV-1 specific, they target high-mannose arrays that are present on envelope glycoproteins. Since such arrays are uncommon in mammalian cells, these compounds are not likely to be toxic to human cells in vivo even at relatively high concentrations. (Dana H et al 2012). Furthermore, a recombinant GRFT produced in the tobacco-like plant Nictiona benthamiana was shown to be nontoxic in a rabbit vaginal irritancy model and in human cervical explants (O’Keeffe et al., 2009). Although GRFT, CV-N and have not yet been tested in human clinical trials it is noteworthy that CV-N was shown to be effective in protecting pigtailed macaques after vaginal and rectal challenges with high dose of 89.6 (Tsai et al 2004, 2003). Since GRFT binds high-mannose oligosaccharides, including the one at position 386 that conceals the b12 epitope, we wished to explore whether this lectin affected exposure of the CD4bs. The authors examined binding using both a virus capture assay and neutralization. Studies found that GRFT enhanced HIV-1 binding of b12 and the nonneutralizing CD4bs monoclonal antibody (MAb) b6, as well as CD4-IgG2, which was used here as a surrogate for the CD4 receptor molecule. Importantly, GRFT and b12 synergized to render some HIV-1 isolates more sensitive to neutralization. The glycan at position 386 on gp120 was found to play a role in both enhancement and synergy, suggesting that GRFT could be used to increase exposure of the CD4bs of HIV-1. Despite differences in the mannose rich glycosylation patterns, the HIV subtypes C, B and A GRFT showed similar sensitivity to GRFT (Zhang et al 2004). The study showed that absence of glycans at positions 234 and 295 were associated with natural resistance to GRFT.
BROADLY NEUTRALIZING ANTIBODIES

Neutralizing antibody responses in HIV-1 infected patients

In HIV-1 infection, antibodies capable of blocking virus infection \textit{in vitro} develop in almost all individuals, although whether they are able to perform this function \textit{in vivo} is less clear. Their absence during the acute phase of infection, when viral levels are brought under control, suggests that cellular immune responses may be more critical during this period (Koup, 1994; Moog \textit{et al.}, 1997). The earliest neutralizing antibodies can be detected after 3-12 months of infection; however, there is considerable variation in the kinetic, magnitude and breadth of this response (Kelly\textit{et al.}, 2005; Moog \textit{et al.}, 1997; Pellegrin \textit{et al.}, 1996; Richman \textit{et al.}, 2003; Wei \textit{et al.}, 2003). In general, the initial neutralization response is narrow, only effective against early autologous viruses (Li \textit{et al.}, 2006a; Moog \textit{et al.}, 1997; Richman \textit{et al.}, 2003) and some T-cell line adapted strains (Pilgrim \textit{et al.}, 1997).

Nevertheless, the appearance of neutralization escape variants soon after the autologous response has developed, supports the notion that these antibodies exert immunological pressure on the virus (Richman \textit{et al.}, 2003; Wei \textit{et al.}, 2003). Antibodies capable of neutralizing heterologous viruses develop later in infection, with only a small percentage of chronically infected patients having broadly cross-reactive antibodies against multiple HIV-1 viruses (Braibant \textit{et al.}, 2006; Donners \textit{et al.}, 2002; Pilgrim \textit{et al.}, 1997)). The nature of the antibodies in broadly cross-reactive sera, as well as why breadth develops so rarely, is not well understood. It is clear, however, that a threshold of viremia is necessary to induce neutralizing antibodies, demonstrated by their absence in individuals on highly active antiretroviral therapy (HAART) and in "elite controllers" (Bailey \textit{et al.}, 2006; Binley \textit{et al.}, 2000; Montefiori \textit{et al.}, 2001). In chronic infection, high viremia has been correlated with neutralization breadth (Deeks \textit{et al.}, 2006). On the other hand, rapidly progressing individuals, who lack control over viremia, usually display low neutralizing antibody titers, but this may be attributed to a general immune suppression (Cecilia \textit{et al.}, 1999; Pilgrim \textit{et al.}, 1997). A recent study has suggested that neutralizing antibodies might
protect from HIV-1 super infection, as this is more likely to occur during the early phase of infection when these antibodies are absent (Smith et al., 2006). Other studies have shown that maternal neutralizing antibodies can exert powerful protective and selective effects during perinatal HIV-1 transmission with resistant strains establishing infection in infants (Dickover et al., 2006).

**Mechanisms of evasion from neutralizing antibodies**

HIV-1 has developed multiple escape mechanisms to avoid neutralization. The shedding of gp120 monomers diverts the immune system towards structures otherwise not found on the native trimer (Wyatt and Sodroski, 1998). The Envelope spike is heavily glycosylated, with the poorly immunogenic glycans shielding antibody access to the peptidic structure (Johnston, Sauvignon, and Desrosiers, 2001). Furthermore, changes in glycan packing yield viruses resistant to the autologous neutralizing antibody response. This neutralization escape mechanism is referred to as an “evolving glycan shield” (Wei et al., 2003). The trimeric nature of the Envelope glycoprotein shields conserved regions, while exposing relative amorphous highly glycosylated loop structures. These regions tolerate high levels of variation and therefore can easily escape from neutralizing antibodies (Wyatt et al., 1998).

Multiple studies have suggested that the V1/V2 loops cover conserved epitopes involved in the coreceptor binding site of the neighboring protomer (Kwong et al., 2000) as deletion of these variable loops confers sensitivity to antibodies targeting this region (Sullivan et al., 1998b; Wyatt et al., 1995). The coreceptor binding site is only transiently exposed after receptor engagement and thus out of antibody reach (Labrijn et al., 2003; Wu et al., 1996). The CD4bs on the other hand is exposed for functional reasons; however, a distinct type of camouflage, called “entropic masking”, protects it. Binding to this epitope requires the fixation of the otherwise flexible gp120, imposing an entropic barrier for the high affinity antibody binding required for neutralization (Kwong et al., 2002).
Broadly neutralizing antibodies

Despite all these defense mechanisms, a few rare broadly neutralizing monoclonal antibodies (bnMAbs) have been isolated from HIV-1 subtype B infected individuals. These MAbs have been shown to neutralize several primary isolates from different genetic subtypes, indicating some conserved structures on the Envelope glycoproteins. Their epitopes include regions in gp41 (2F5 and 4E10), the CD4bs(b12), and part of the carbohydrate-masked “silent face” of gp120 (2G12). Crystallographic analyses of these antibodies have revealed that they underwent remarkable structural adaptations to attain virus recognition (Burton, Stanfield, and Wilson, 2005).

Passive immunization of primates challenged with chimeric simian–human immunodeficiency virus (SHIV) strains has shown that human bnMAbs can protect against infection and are effective against intravenous (Baba et al., 2000; Mascola, 1999), oral (Hofmann-Lehmann et al., 2001) or intra vaginal challenges (Mascola et al., 2000; Parren and Burton, 2001; Vezey et al., 2003) in cases where transmission occurs disease by blunting the peak of viremia and lowering the viral set point (Ferrantelli et al., 2007). A recent study in humans showed that in some HIV-infected individuals these bnMAbs can reduce the rate of viral rebound following a structured treatment interruption (Trkola et al., 2005). Furthermore, the existence of 2G12 escape variants in some of the treated patients demonstrated that this bnMAb was indeed functional in vivo (Manrique et al., 2007; Nakowitsch et al., 2005).
Neutralizing antibodies

CD4 binding site (CD4bs):

IgG1b12

The neutralizing antibody b12 was obtained as a Fab through a phage display library strategy (Moulard et al., 2002). The Fab b12 as well as the IgG1 recombinant MAb derived from it, IgG1b12, occlude the CD4bs on gp120 and prevents CD4 attachment (Burton et al., 1994; Roben et al., 1994). A key element of the CD4bs is a recess that forms a contact site for the Phe43 protruding from a loop of CD4 (Wyatt and Sodroski, 1998). The first crystal structure of IgG1b12 revealed that the protruding heavy chain complementarity determining region 3 (CDRH3) was unusually long, allowing it to access the CD4 binding pocket (Saphire et al., 2001). However, it was only the recent crystallized structure of the Fab b12 in complex with gp120 that clarified the mechanism behind this antibody neutralization. The b12 interactions with gp120 are mainly with residues in the structurally invariant outer domain (Figure 2.13). As a result, the b12 binding site, in contrast to the CD4bs, does not differ considerably from the pre- to post-attachment forms of gp120 (Zhou et
Ph D Thesis
Sampurna Mukhopadhyay

explaining the previously shown low entropic cost of this interaction (Kwong et al., 2002).

VRC01 and VRC03
Three CD4bs-directed MAbs, VRC01, -02, and -03 (Wu et al., 2010) were recently isolated from a donor whose serum neutralization activity mapped (Wu et al., 2010) predominantly to the CD4bs (Li et al., 2007). VRC01 and VRC02 is a closely related pair of somatic variants that neutralize over 90% of diverse HIV-1 primary isolates. The structure of VRC01 in complex with the gp120 core reveals that the VRC01 heavy chain binds to the gp120 CD4bs in a manner similar to that of the primary receptor, CD4 (Zhou et al., 2010). However, the gp120 core lacks the major variable regions, as well as the N and C termini, and may only partially reflect the VRC01 interaction with full-length gp120, or with gp120 in the context of the native viral spike. It was later shown that VRC01 interacts with the functional spike in a manner distinct from that of CD4. VRC01 achieves potent neutralization by precisely targeting a highly conserved region of the CD4bs without requiring the alterations of the Env functional spike configuration that occur upon CD4 ligation (Li et al., 2011). This helps to explain how VRC01 can access the CD4bs on the large majority of virus isolates and why VRC01 resistance based on the quaternary structure of the HIV-1 Env is uncommon. These studies highlight the unique features of this broadly neutralizing CD4bs-directed antibody and point to further efforts regarding HIV immunogen design based on its gp120 (Li et al., 2011; Zhou et al., 2010).
FIG 2.14 Structure of antibody VRC01 in complex with HIV-1 gp120. Atomic-level details for effective recognition of HIV-1 by a natural human antibody are depicted with polypeptide chains in ribbon representations. The gp120 inner domain is shown in gray, the bridging sheet in blue, and the outer domain in red, except for the CD4-binding loop (purple), the D loop (brown), and the V5 loop (orange). The light chain of the antigen-binding fragment (Fab) of VRC01 is shown in light blue with complementarity determining regions (CDRs) highlighted in dark blue (CDR L1) and marine blue (CDR L3). The heavy chain of Fab VRC01 is shown in light green with CDRs highlighted in cyan (CDR H1), green (CDR H2), and pale yellow (CDR H3). Both light and heavy chains of VRC01 interact with gp120; the primary interactive surface is provided by the CDR H2, with the CDR L1 and L3 and the CDR H1 and H3 providing additional contacts.

2G12: Sugar binding antibody

2G12 recognizes a unique epitope on the surface of gp120 that is not directly associated with the receptor binding sites (Sanders et al., 2002). Antibody mapping studies using monomeric gp120 showed that 2G12 forms a unique competition group.
in that no other MAb is able to prevent its binding to gp120 and vice versa (Moore and Sodroski, 1996). 2G12 binds to high mannose and/or hybrid glycans, with mannose residues as essential components (Figure 2.15). Mutagenesis studies have implicated the glycans at positions 295, 332 and 392 in gp120 as being the most critical for 2G12 binding (Sanders et al., 2002; Scanlan et al., 2002). 2G12 binds to the N-linked glycans at position 332 and 392 in the primary combining sites, with a potential interaction with glycan 339 at the VH/VH' interface of the antibody (Calarese, 2003). They also proposed that the glycan at position 295 plays an indirect role by preventing further processing of the glycan at 332 and maintaining its oligomannose structure as the one recognized by 2G12.

FIG2.15A model of mAb 2G12 Fab2 bound to the HIV-1 Env spike. The heavy chains of 2G12 are shown in dark blue and light red, and the light chains are shown in azure. The domain-swapped structure of the heavy chains (VH) of 2G12 interlock with the light chains (VL) of 2G12 in both the uncomplexed and complexed with Man9GlcNAc2 is described in ref Calarese et al 2003. Docking of the structure of 2G12, complexed in the conventional VH-VL combining sites with Man9GlcNAc2, onto gp120 places the GlcNAc2 groups very close to N332 and N392 (outer dark red moieties). The Man9GlcNAc2 group attached to N339 (middle dark red) can readily modeled to interact with the nonconventional VH-VH interface region. Figure taken from Burton et al (Burton, Stanfield, and Wilson, 2005).
Antibodies to Membrane proximal external region (MPER) of gp41: 2F5 and 4E10 epitopes

2F5 and 4E10 recognize two adjacent highly conserved epitopes in the extreme C terminal of the gp41 ectodomain (Figure 2.16). This region is particularly attractive for vaccine design because it mediates the viral entry process and is highly conserved between viral strains. The 2F5 epitope has been mapped to the motif ELDKWA at the end of the HR2 region of gp41 (Muster et al., 1993), where the core residues D664, K665 and W666 are indispensable for antibody recognition (Zwick et al., 2005).

FIG2.16(A) Models of the mAbs 4E10 and 2F5 Fabs bound to their epitopes. The schematic representation of 4E10 and 2F5 mAbs bound to their epitopes present in gp41 subunit and transiently exposed during the fusion process is shown. (B) 2F5 and 4E10 are lipid reactive antibodies and the structures of Fabs are shown in the vicinity of the viral membrane (Figure adopted from The Scripps Research Institute website (www.scripps.edu)).
Structural data of 2F5 MAb epitope complexes shows that this region adopts an extended conformation with a type I β-turn at the core of the epitope. Interestingly, the hydrophobic apex of the CDR H3 loop of 2F5 does not interact with the epitope directly (Ofek et al., 2004). It has been suggested that this region mediates interactions with the epitope-proximal viral membrane, explaining early evidence that 2F5 binding was enhanced in the presence of lipids (Grundner et al., 2002). The bnMAb 4E10 recognizes a contiguous epitope at the C-terminus of the 2F5 binding region (Stiegler et al., 2001; Zwick et al., 2001). Mutagenesis experiments have demonstrated that the residues W672, F673 and W680 of the Trprich region of gp41 are indispensable for recognition by 4E10 (Zwick et al., 2005). The crystal structure of 4E10 bound to a 13-residue peptide revealed that this epitope assumes an unusual helical conformation. The hydrophobic face of this amphipathic helix is buried in the antibody combining site, where amino acids W672, F673, I675 and T676 are the key residues in this interaction (Cardoso et al., 2005). Further structural analysis of this epitope has extended it to the motif 672-WFx(I/L)(T/S)x(L/I)W-680, where x does not play a major role in 4E10 binding (Cardoso et al., 2007).

**V3 loop**

In addition to those described above, there are other epitopes able to induce neutralizing antibodies, but in a more limited way. This is the case for the V3 loop of gp120, which was previously considered the principal neutralizing determinant (Palker et al., 1988; Rusche et al., 1988). Later research demonstrated that this was only applicable to TCLA strains, where numerous passages in cell culture rendered these viruses highly sensitive to neutralization by anti-gp120 monoclonal antibodies, patients’ sera, and soluble forms of CD4 (Follis et al., 1998; Verrier et al., 2001; Wrin et al., 1995). The mechanism by which sensitivity to neutralizing ligands is acquired is not clear and is manifested only in the context of the functional trimeric Envelope spike. Monomeric gp120s derived from either a TCLA strain or a primary isolate exhibit similar affinities for sCD4 (Moore et al., 1991; Moore, Morikawa, and Jones, 1991). By contrast the trimeric Envelope glycoprotein of TCLA viruses bind the CD4 molecule more efficiently than the primary isolates (Kabat et al., 1991).
The exposure of epitopes on TCLA viruses may reflect an optimization of the virus-cell interactions, particularly the CD4-gp120, in the absence of selective pressure provided by serum-neutralizing antibodies (Moore, 1995). It is believed that the V3 loop is less important for primary isolate neutralization, presumably because this region is occluded in the trimeric structure prior to receptor binding. Furthermore, due to the variable nature of this region, most anti-V3 antibodies are isolate-specific. However, a group of these antibodies recognizes conformation-sensitive epitopes on V3 and they are able to neutralize a range of primary isolates. This is the case for the MAb 447-52D, which recognizes the GPGR motif at the tip of the V3 and main-chain atoms along the N-terminal side of the loop (Gorny et al., 1992; Huang et al., 2005; Stanfield et al., 2004). 447-52D neutralizes laboratory strains (Gorny et al., 1992; Gorny et al., 1993) and clinical isolates from various clades (Conley et al., 1994; Nyambi et al., 1998). Although, its activity is limited to viruses containing the GPGR sequence at the apex of V3 loop, the relatively broad neutralizing activity of 447-52D highlights the existence of conserved structures in the V3 loop and makes this region a potential vaccine target.

Figure 2.17: The structure of a V3-loop peptide in the binding site of the antibody 447 adopted from Stanfield et al (Stanfield et al., 1999). The CDRH3 loop (pink, numbered) forms a mixed β-sheet with the V3 loop (blue). GPGR forms the turn in the peptide structure and interacts with the base of the CDRH3 loop. Main-chain interactions dominate the interaction of the peptide with the CDRH3 loop.
Coreceptor binding site and/or CD4 induced epitope (CD4i)

The binding of CD4 to gp120 induces conformational changes that lead to the formation of the coreceptor binding site and enhanced binding of a group of antibodies, referred to as CD4i antibodies, such as: 17b, 21c, 23e, 48d, 49e (Xiang et al., 2002), X5 (Moulard et al., 2002), E51 (Xiang et al., 2003) and 41d (Xiang et al., 2005). Crystal structure and mutagenesis data have shown that the epitope recognized by these antibodies overlaps significantly with the highly conserved coreceptor binding site (Huang et al., 2005; Kwong et al., 1998; Xiang et al., 2002; Xiang et al., 2003). In many cases these antibodies mimic the coreceptor molecule by presenting sulfated tyrosine in their CDRH3 (Huang et al., 2004; Huang et al., 2007). CD4i antibodies are commonly found in HIV-infected individuals (Decker et al., 2005) suggesting that this epitope is highly immunogenic. Despite the extremely broad recognition of CD4i antibodies, neutralization is usually impaired. Several studies have shown that virus strains that do not require CD4 for entry are highly sensitive to neutralization by CD4i antibodies (Edwards et al., 2001; Kolchinsky, Kiprilov, and Sodroski, 2001). Moreover, CD4 dependence assures that this immunogenic and conserved region is only exposed after receptor engagement, where the close proximity of the viral and host membranes somewhat restricts the access to this region. This is supported by the fact that small forms of these antibodies, such as Fabs or single chains, display better neutralizing activity (Labrijn et al., 2003). Taken together, these observations preclude this epitope as a good target for vaccine design.

Quaternary Site Epitope (QNE) binding antibodies

Recently, broadly neutralizing monoclonal antibodies (MAbs) to HIV-1 have been discovered that were found to be very potent in nature and which primarily targets quaternary epitopes (such as PG9, PG16 and PGT antibodies) (Walker et al., 2011; Walker et al., 2009) in variable loops. Monoclonal antibodies PG9 and PG16 were isolated from a African HIV positive patient infected with clade A HIV-1 and who had developed a potent and broadly neutralizing serum. PG9/16 neutralize 80%
of HIV-1 isolates across all clades with extraordinary potency and target novel epitopes preferentially expressed on Env trimers. As these antibodies are very potent in nature the structural properties of these antibodies have been investigated which showed that these antibodies bear a very long CDR-3 loop. The crystal structure of the antigen-binding fragment (Fab) of PG16 at 2.5 Å resolutions revealed its unusually long, 28-residue, complementarity determining region (CDR) H3 forms a unique, stable subdomain that towers above the antibody surface and this is the peculiar structure of these antibodies that enhanced their neutralization potency. The PG16 electron density maps shows that a CDR H3 tyrosine is sulfated, and it increases the binding energy of these antibodies with the epitope on the Env trimer (McLellan et al., 2011; Pejchal et al., 2010). PG9 and PG16 use unique structural features to mediate potent neutralization of HIV-1 that may be of utility in antibody engineering and for high-affinity recognition of a variety of therapeutic targets. Notably, their potency is about an order of magnitude higher than that of previously described broadly neutralizing mAbs to HIV-1. These two antibodies differ by somatic variation and bind to overlapping, but distinct, gp120 epitopes composed of conserved determinants in V2, V3, the V1/V2 stem, and perhaps elements of the coreceptor binding site (CoRbs) (Walker et al., 2009). The epitopes of these antibodies are preferentially presented on the trimeric form of Env rather than monomeric gp120 indicating that these are quaternary structural epitopes. The sensitivity of PG9/16 depends on the V1, V2 ad V3 loop sequence and structure. PG9 is sensitive to mutations in V1 and V2 while changes in V3 affect PG16 sensitivity more than PG9 (Walker et al., 2011). The V1V2 and V3 loops are situated at the apex of the trimer as determined by fitting crystal structures of coregp120 in complex with the antigen-binding fragment (Fab) into cryoelectron tomography reconstructions of the native trimer (Liu et al., 2008) and are thought to mediate gp120 protomer association within the trimer (Liu et al., 2008; Schief, Ban, and Stamatatos, 2009). However, no gp120 crystal structure with gp41 or the crystal structure with variable loops exists to model the epitopes. Also the atomic details of the gp120 interactive surface are not known and therefore it is difficult to exactly mimic the epitopes of these antibodies. However, it can be inferred that the epitopes are presented only on trimeric Env and on
unliganded Env as the binding of PG antibodies is affected after the Env trimer is engaged with soluble CD4 (Walker et al., 2011). CD4 binding entails the major structural changes in the core structure probably displacing variable loops of Env and PG9/16 competes with sCD4 for the binding to Env. These Abs do not compete with CD4bs antibodies like b12 as its binding involves small structural changes (Liu et al., 2008). The binding and neutralization of PG antibodies is also shown to be dependent on the glycans on the trimeric Env (Doores and Burton, 2010). Walker et al (Walker and Burton, 2010) and McLellan et al (McLellan et al., 2011) have shown that the glycans in V1V2 region at positions 156 in V1 and 160 in V2 are indispensable for the binding of PG antibodies. These glycans make the part of the epitope and the antibody CDR3 loop directly interacts with these glycans.

FIG 2.18 Overall structure of the V1/V2 domain of HIV-1 gp120 in complex with PG9. V1/V2 from the CAP45 strain of HIV-1 is shown, in magenta ribbons, in complex with Fab of antibody PG9. The PG9 heavy and light chains are shown as yellow and blue ribbons, respectively, with CDRs in different shades. Although the rest of HIV-1 gp120 has been replaced by the 1FD6 scaffold (shown in white ribbons), the positions of V1/V2, PG9 and the scaffold are consistent with the proposal that the viral spike, and hence the viral membrane, is positioned towards the top of the page. The extended CDR H3 of PG9 is able to penetrate the glycan shield that covers the V1/V2 cap on the spike and to reach conserved elements of polypeptide, while residues in heavy- and light-chain-combining regions recognize N-linked glycans. The disordered region of the V2 loop is represented by a dashed line. Taken from McLellan et al. (McLellan et al., 2011).
Recently, newer antibodies have been discovered that mediate even broader neutralization coverage than PG9 and PG16. The neutralizing antibody repertoires of our Human Immunodeficiency Virus (HIV)-infected donors with remarkably broad and potent neutralizing responses and rescued 17 new monoclonal antibodies (PGT antibodies) that neutralize broadly across clades. Antibody variable genes were rescued from B-cell cultures that showed cross-clade neutralizing activity and expressed as full-length IgGs. Many of the new monoclonal antibodies are almost tenfold more potent than the recently described PG9, PG16 and VRC01 broadly neutralizing monoclonal antibodies and 100-fold more potent than the original prototype HIV broadly neutralizing monoclonal antibodies (Wu et al. 2010, Walker 2009). These antibodies are found to target new epitopes and they are These PGT antibodies prototyped by PGT 128 manage to bind to two closely spaced glycans and at the same time reach through the rest of the "glycan shield to take hold of a small part of structure on gp120 known as the V3 loop (Walker et al. 2011).