CHAPTER II

LITERATURE REVIEW

1. Human immunodeficiency virus type 1 (HIV-1)

HIV-1 belongs to the *Lentivirus* subfamily of retroviruses and has been shown to be the etiologic agent of acquired immunodeficiency syndrome (AIDS) (Barre-Sinoussi *et al.*, 1983; Gallo *et al.*, 1984). It was genetically divided into two groups: M (major group) and O (outlier group). Group M is pandemic worldwide and contains at least 10 subtypes. The major subtype in Thailand is subtype E while the B is the major subtype in USA and Europe. Recently, new group termed N (non-M-non-O) has now been identified from Cameroon (Simon *et al.*, 1998).

1.1 Genomic organization

HIV-1 proviral DNA is approximately 9.2 kb. Like all retroviruses, HIV-1 has three structural genes: *gag*, *pol* and *env*. The *gag* gene encodes the 55 kDa Gag precursor protein, also called p55, which is cleaved by the viral protease into four proteins that are found in the mature virus: p17, p24, p9 and p6. The *pol* gene encodes three viral enzymes necessary for replication: protease, reverse transcriptase and integrase. The *env* gene encodes a precursor protein (gp160), which is then cleaved by host cell protease into gp120 and gp41. The 5' and 3' long terminal repeats (LTRs) flanking the provirus DNA. The LTRs contain important regulatory regions, especially those for transcription initiation and polyadenylation. The regulatory genes: *tat* and *rev* encode proteins (Tat and Rev, respectively) that modulate transcriptional and post-transcriptional steps of the viral genes. The accessory proteins: Vif, Vpr, Vpu, Vpx, and Nef, which are encoded by the accessory genes: *vif*, *vpr*, *vpu*, *vpx*, *nef*, respectively are not absolutely necessary for viral replication in *in vitro* system, but represent critical virulence factors *in vivo*. The genomic organization of HIV-1 is shown in Figure 1A.

1.2 Virion structure

HIV-1 is spherical in shape and has a diameter of 100 to 120 nm. The outer coat of the virus, known as the viral envelope, is a lipid bilayer that taken from the membrane of the host cell when a nascent virion buds from the infected cell. The envelope contains several cellular membrane proteins, including major histocompatibility antigens (MHC) and viral proteins. The proteins that protrude from the envelope surface consists of gp120 (the surface protein, or SU) and gp41 (the transmembrane protein, or TM). The gp120 contains the binding site for the CD4 receptor, and the seven transmembrane domain chemokine receptors that serve as a co-receptor for HIV-1. The gp41 contains the outer N-terminal hydrophobic domain which initiates the process of virus-cell membrane fusion, the transmembrane and inner cytoplasmic tail that anchor to the lipid bilayer. The inner surface of the envelope lining with matrix proteins (MA, p17) to which it is attached by covalently bound myristic acid. Within the envelope of a mature HIV-1 particle is a cone-shaped capsid, consists of 2,000 copies of major capsid proteins (CA), p24. The capsid surrounds two copies of single strands RNA genome, nucleocapsid protein (NC), p7/p9, three viral enzymes: protease, reverse transcriptase and integrase and the viral accessory proteins: Nef, Vif and Vpr. The structure of HIV-1 is shown in Figure 1B.

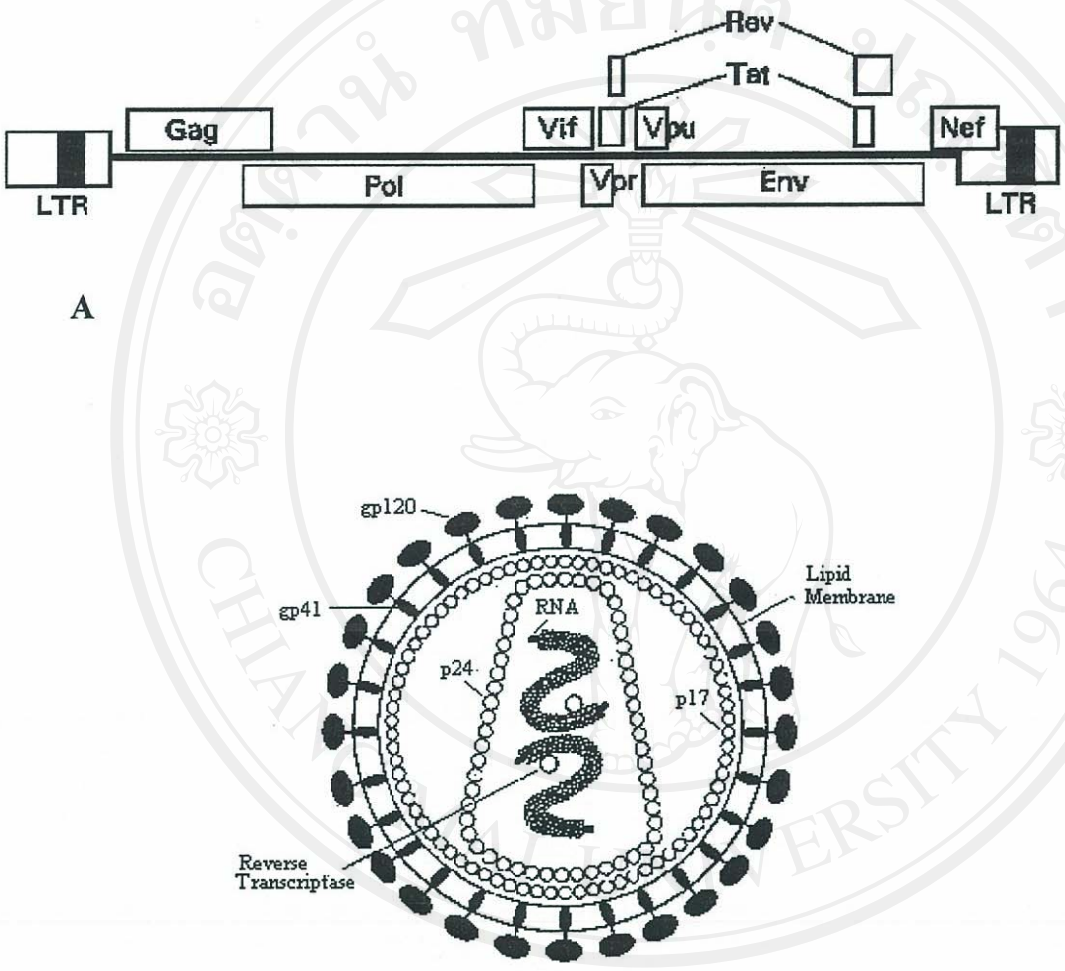

Figure 1 Virion structure and genomic organization of HIV-1; (**A**) structure of HIV-1 DNA, and (**B**) structure of mature HIV-1 virion. (Folks and Hart, 1997)

1.3 Cellular targets

HIV-1 infects the target cells through the specific binding of gp120 to CD4 molecule and co-receptor on the cell surface. CD4+ T cell is the primary target cell for HIV-1 infection. Additionally, the CD4+ monocytes/macrophages including microglia cells, Langerhans cells, and follicular dendritic cells (FDCs) are also infected and may provide an important reservoir of viruses (McElrath *et al.*, 1989; Albright *et al.*, 1999). Not only cells that have the CD4 antigen are infected, some CD4- cells such as those in brain and intestine can be infected via a galactocerebroside (galC) receptor (Harouse *et al.*, 1991). Other cells can be infected in a different way; for example, in macrophages Fc or complement receptor may be used (Homsy *et al.*, 1989; Boyer *et al.*, 1991). In these cases, the HIV-1 must be bound by anti-HIV-1 antibodies that interact with receptors on the cell surface. Thus anything that can up-regulate Fc receptors on macrophages will augment infection.

1.4 Viral tropism

Primary isolates of HIV-1 exhibit distinct differences in cellular tropisms and syncytium-inducing capability that have important implications for viral pathogenesis and disease progression. Generally, the HIV-1 strains that are most commonly transmitted between individuals and predominate in the early stages of infection have been referred to as "macrophage tropic (M-tropic)" due to their ability to infect and replicate in macrophages and primary CD4+ T cells but fail to infect T-cell lines, "non-syncytium inducting (NSI)" due to their inability to induce syncytia in MT-2 cultures, or "slow-low (SL)" due to their replication kinetics in culture. While the strains that are able to infect T cell lines but fail to infect macrophages and are correlated with accelerated disease progression have been variously referred to as "T cell tropic (T-tropic)", "syncytium-inducing (SI)", or "rapid-high (RH)" according to the nomenclature schemes mentioned as above. Finally, the strains which can infect T cell lines in addition to macrophages and primary CD4+ T cells and may represent an intermediate form during the evolution from M to T tropisms (typically about 4 to 5 years after infection) are referred to as "dual-tropic".

1.5 Replication cycle

1.5.1 Binding and entry

The infection of cells by HIV-1 begins with the interaction of the viral envelope proteins with two cell surface proteins, CD4 molecule and co-receptor. The CD4 is an extracellular immunoglobulin (Ig)-like structure containing 4 domains and is expressed on the surface of a subset of T-lymphocytes and some macrophages. The gp120 of Env bound to the CDR2-like region in the first extracellular Ig-like domain of CD4. Binding of the gp120 to CD4 induces conformational changes in the gp120 that create or expose a binding site for the co-receptor. The association of gp120 with co-receptor drives additional conformational changes within the entire trimeric gp120/gp41 complex and eventually led to the fusion between the viral envelope and the host cell membrane.

1.5.2 Reverse transcription and DNA synthesis

Subsequent to internalization and uncoating, the HIV-1 genomic RNA was reverse transcribed to cDNA and then double-stranded proviral DNA using viral Pol proteins containing the RT and ribonuclease H enzymatic activities (RT/RNase H). The reverse transcription process also generates the LTR regions on the both 5' and 3' ends of the proviral DNA that are characteristic of retroviruses and necessary for integration into the cellular chromosomal DNA.

1.5.3 Nuclear transport of the pre-integration complex

The double-stranded viral DNA pre-integration complex contains the Gag matrix, Pol integrase-RT proteins and the nuclear localization signals that may pivotal in targeting the nucleoprotein complex to the nucleus. In this step, the pre-integration complex has to transverse an intact nuclear membrane for access to the host chromosomal DNA.

1.5.4 Proviral DNA integration

The linear double-stranded viral DNA is capable of integrating into the host chromosomal DNA with the help of HIV-1 integrase that contained in the pre-integration complex. Once incorporated into the host chromosomal DNA, HIV-1 DNA is called a provirus.

The proviral DNA can persist for many years in a latent state and secretly carry the genetic instructions for making new virions.

1.5.5 Viral transcription and protein synthesis

When the host cell is activated, the proviral DNA can be transcribed by the host cell polymerases into viral mRNA, and consequently translated into the viral proteins. Together, the viral genomic RNA and proteins then migrate to the host cell membrane, where they assembled to new virions.

1.5.6 Assembly and budding of virus

Before the virion can be released from the cell, the viral proteins must coordinately assemble. The viral enzyme and genomic RNA gather just inside the cell membrane, while the viral envelope proteins aggregate within the membrane. An immature viral particle forms and pinches off from the cell, acquiring an envelope that includes both cellular and HIV-1 proteins from the cell membrane. During this step, the core of the virus is still immature and the virus is not yet infectious. The long chains of proteins and enzymes that make up the immature viral core are now cleaved into smaller pieces by a viral enzyme called protease. This step results in infectious viral particles.

1.6 Mode of transmission

HIV-1 can be transmitted in three major routes including sexual intercourse, from mother to child and blood injection.

1.6.1 Sexual transmission

Sexual intercourse is the main transmission mode of HIV-1 and is estimated that 75% of HIV-1 infected cases worldwide are acquired by sexual contact (Laga and Schwartländer, 1999). An unprotected sexual intercourse by homosexual, bisexual men and heterosexual couples is the effective mode of sexual transmission of HIV-1. However, variation factors among sexual intercourse may reflect differences in the risk of transmission, including the inherently heterogeneous nature of sex (e.g., frequency and types of sexual partnerships, patterns of sexual

mixing, types of sexual practices), variation in viral strains, concomitant risk factors (e.g. other STDs), the use of vaginal desiccants, and having sex during menses, or to variation in average infectiousness at the community level in different phases of the epidemic (Mertens and Piot, 1997).

1.6.2 Mother to child transmission

On a worldwide scale, transmission of HIV-1 from mother to child during or after pregnancy is the second most common mode of spread of HIV-1 (Laga and Schwartländer, 1999). It is the major source of HIV-1 infection in children. HIV-1 is transmitted to fetus or infant by 13% to 48% of infected mothers (Mertens and Piot, 1997). The mother-to-child transmission of HIV-1 includes transmission during pregnancy, during delivery and through breast-feeding. The risk factors of transmission are differences in the distribution of determinants of transmission, such as the degree of maternal immune deficiency, the presence of chorioamnionitis, maternal vitamin A deficiency, or exposure to infected breast milk (Mertens and Piot, 1997).

1.6.3 Transmission by blood products and contaminated equipment

Sharing of injection equipment among injecting drug users (IDUs) is a major route of HIV-1 transmission by blood, particularly in industrialized countries (Mertens and Piot, 1997), while blood transfusion is the most efficient mode of HIV-1 transmission (Laga and Schwartländer, 1999). Being unable to implement screening of blood donors and other blood supplies led to transmission of HIV-1 through contaminated blood products. However, the role of injection associated nosocomial HIV-1 transmission is rarely possible.

2. Clinical course of HIV-1 infection

HIV-1 infected persons have been classified into three major groups: the typical progressors, rapid progressors and long-term nonprogressors. The classifying criteria are based on the duration of disease progressing period, the immune responses and the kinetics of viremia.

2.1 Typical progressors (TPs)

Normally, 80% to 90% of HIV-infected persons are in the typical progressor group

that experiences a course of HIV-1 disease with a median survival time of approximately 10 years (Figure 2A). The typical course of HIV-1 infection consists of three stages including primary infection, clinical latency and stage of advanced disease.

Primary infection stage generally refers to the period, usually a few weeks to months, starting from initial infection to the immune response to HIV-1. At this stage, a minority of newly infected individuals experience signs and symptoms which may include fever, malaise, rash, lymphadenopathy, pharyngitis, headache, diarrhea and occasionally neurologic manifestations. The primary HIV-1 infection is characterized by high levels of plasma viremia and precipitous decline in CD4+ T cell counts. During this period, the virus disseminates throughout the body and seeds lymphoid organs, where its replication is incompletely suppressed. The initially high levels of HIV-1 replication and plasma viremia are generally decreased with the appearance of an HIV-1 specific immune response; these levels stabilize within 6 months to 1 year at a virologic set-point.

The level of plasma viremia at a set-point is a significance of the beginning of a clinically latent period that may last for years. This phase is characterized by chronic immune activation and persistent viral replication despite a lack of consistent signs or symptoms of disease. Typically, this is the longest lasting of the three stages of HIV-1 infection. During this stage, the number of circulating CD4+ T cells slowly declines by 50 to 70 cells/μl per year by cytotoxicity immune mediated elimination of infected cells and failure to replace adequately the dying CD4+ T cells (D'Souza and Fauci, 1999). Despite the number of infected cells in the peripheral blood, low level of the plasma viral RNA levels and the number of cells expressing virus in lymphoid tissue are detected. In contrast, the large quantity of virus trapped in the follicular DCs (FDCs) in the germinal centers of the lymph nodes was found (Smith *et al.*, 2001).

Progression to AIDS and clinically apparent disease occurs within 8 to 10 years in typical progressors. The advanced disease is characterized by severe AIDS defining illness, or by opportunistic infections or neoplasms. In this stage, the plasma viremia increases and is correlated to a decline of the levels of circulation CD4+ T cells.

2.2 Rapid progressors (RPs)

Approximately 5% to 10% of HIV-1 infected persons experience a rapid decline in CD4+ T cell levels within 2 to 3 years progress to AIDS within 3 years after primary infection (Figure 2B). Those individuals were defined as rapid progressors (RPs). The RPs uniformly exhibit higher HIV-1 RNA levels in the plasma as well as higher HIV-1 DNA load in PBMCs when compared with nonprogressors. This high viral load is usually detected soon after seroconversion and persists throughout the course of disease. The RPs exhibit more homogeneity among HIV-1 isolates compared with the nonprogressors. This finding implies that the immune response to HIV-1 is less effective. The HIV-1 isolated from RPs has a virulent SI phenotype and uses the CXCR4 as a co-receptor. The levels of antibodies against HIV-1 proteins and neutralizing antibody range from low to absent. The presence of circulating anti-HIV-1 CTL activity against multiple viral proteins has been reported but not effective (Pantaleo *et al.*, 1997). The CD8+ T cells that have been shown to mediated suppression of HIV-1 replication (i.e., mediated by soluble factor) are severely impaired. A series of immunological abnormalities including high percentages of activated CD8+ T cells expressing CD38 and HLA-DR and elevated serum levels of β_2-microglobulin, neopterin, soluble CD8 and interleukin-2 (IL-2) receptors are usually observed (Pantaleo *et al.*, 1997).

2.3 Long-term nonprogressors (LTNPs)

A small percentage (about 5%) of HIV-1 infected persons that have stable CD4+ T cell counts for many years (8 to 10 years at least) and do not have clinical symptoms due to HIV-1 infection and no anti-retroviral therapy are termed as long-term nonprogressors (LTNPs) (Figure 2C). The LTNPs from different cohorts have significantly lower levels of viremia when compared to the progressors. The HIV-1 isolated from LTNPs is heterogeneity and exhibits the NSI phenotype. The HIV-1 isolates from those individuals use CCR5 as a co-receptor. A high frequency of HIV-1-specific memory CTLs against Env, Gag and Pol proteins were presence at high levels when compared to the progressors. To date, it has been suggested that viral strain and phenotype such as virus with defective in the *nef* gene (Kirchhoff *et al.*, 1995) as well as host immune response and host genetic factors may have an effect to clinical outcome of LTNPs.

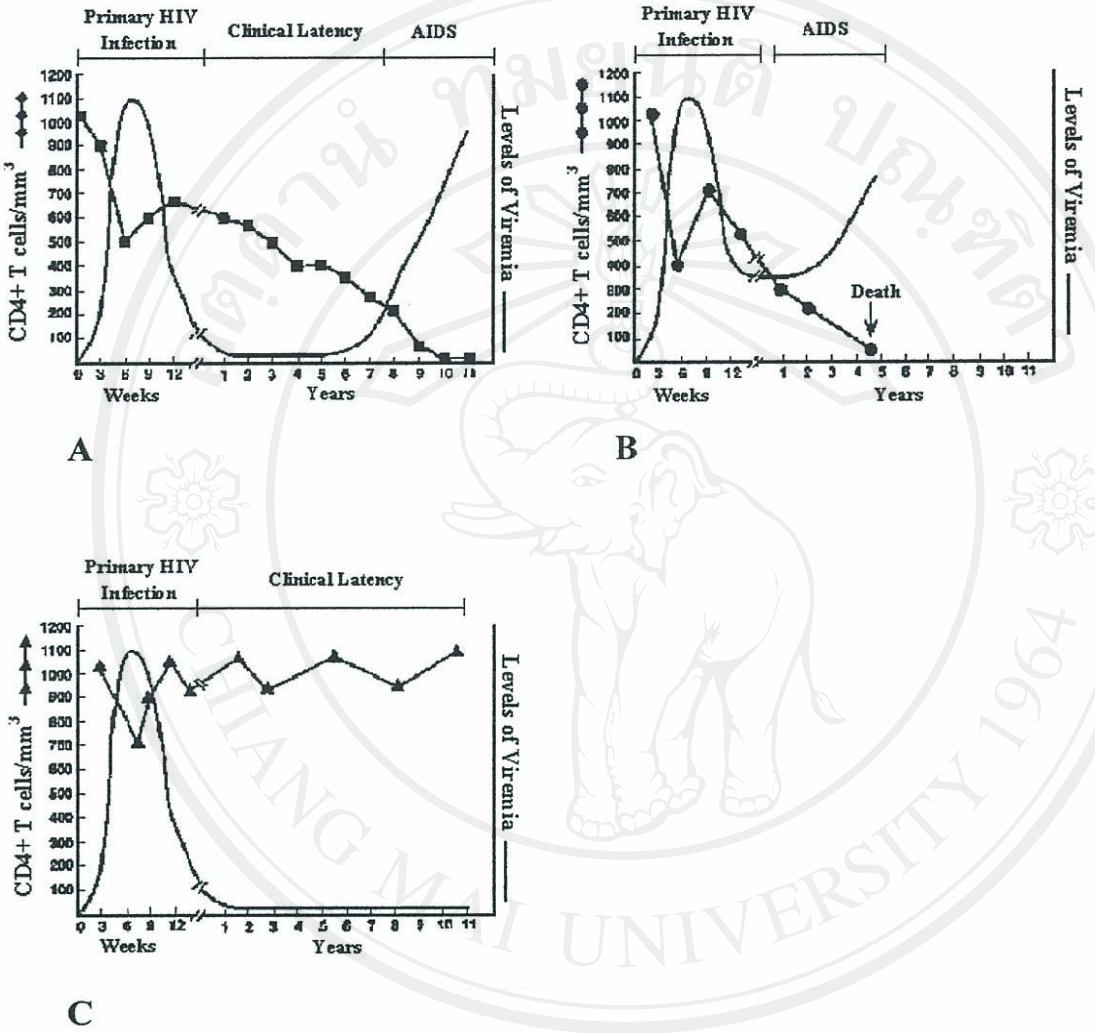

Figure 2 Schematic representation of the decline in CD4+ T cells and the kinetics of viremia during the various courses of HIV-1 infection observed in; **(A)** typical progressors (TPs), **(B)** rapid progressors (RPs), and **(C)** long-term nonprogressors (LTNPs). (Pantaleo *et al.*, 1997)

3. Highly exposed persistently seronegative (HEPS) persons

HIV-1 high risk persons, including commercial sex workers (CSWs), sex partners of HIV-1 infected persons (heterosexual and male homosexual), infants born to HIV-1 infected mothers, recipients of HIV-1 infected blood products, HIV-1 exposed health care workers, and intravenous drug users who have shared needles with HIV-1 infected persons, have particularly high incidence rates to infect by HIV-1. However, the preference of HIV-1 infection depends on the route of exposure, the level of infectious virus and the frequency of HIV-1 exposure. During the past decade, it has become clear that not all people are equally susceptible to infection by HIV-1 and that some individuals in high risk groups may remain HIV-1 seronegative despite multiple repeated exposures to the virus. These individuals are referred to as a highly exposed persistently seronegative (HEPS), exposed uninfected (EU) or exposed seronegative (ES). The term HEPS is used in this study. The characteristic of HEPS individuals are HIV-1 specific IgG negative by all standard serological tests, lacking detectable of HIV-1 proviral DNA and HIV-1 RNA by polymerase chain reaction (PCR) and reverse transcriptase (RT)-PCR respectively. And they have no clinical or laboratory signs of immunodeficiency.

Several observations have been postulated to explain resistibility to HIV-1 infection in HEPS individuals. These findings have led to suggestion that multiple factors including host genetic variation in chemokine and chemokine receptor genes, MHC class I restricted cytotoxic T lymphocytes (CTLs), soluble factor of CD8+ mediated suppression of HIV-1 replication, CD16+ natural killer (NK) cells, and mucosal HIV-1 specific IgA may have major role in protection against HIV-1 infection (Simonsen *et al.*, 1998; Hogan and Hammer, 2001; Mascola *et al.*, 2000), while co-infections with sexually transmitted diseases (STDs) including *gonococcal cervicitis*, may involve in increasing the risk of HIV-1 infection (Kaul *et al.*, 2002).

The CSWs and serodiscordant couples (spouses and male homosexuals) are the two particularly groups of HEPS for most studies in resistibility to HIV-1 infection. Unlike other groups of HEPS, the CSWs and serodiscordant couples are multiple repeated exposures to HIV-1 through sexual intercourse in prolonged period of time (longer than 1 year) and their high risk of STDs. Consequently, the CSWs and serodiscordant couples are well-described in resistibility to HIV-1 infection. Despite the frequency of HIV-1 exposure, the CSWs are likely to experience a variety of HIV-1 variants, whereas the serodiscordant couples are a single event.

4. HIV-1 co-receptors

4.1 Chemokines and chemokine receptors

The principle cells targeted by HIV-1 *in vivo* are CD4+ T lymphocyte and cells of the monocyte/macrophage lineage. HIV-1 entry and infection are typically initiated by the binding of gp120 to the CD4 molecule and the co-receptor. The HIV-1 co-receptors are members of the chemokine receptor which binds to the specific chemokine ligands resulting in subtle network of cell-cell signaling and cell trafficking at different sites throughout the body.

Chemokine is a chemo-attractant cytokine. This term is used to describe molecules which combined a chemo-attractant and cytokine properties. The chemokines are a large family of small cytokines, 8 to 10 kDa, and function as an inflammatory mediator of leukocyte activation and chemotaxis to the site of inflammation. About 40 chemokines have now been identified in human. They are divided into four groups according to the number and spacing of conserved cysteines. Three of them named CC, CXC and CX_3C have four conserved cysteines each, whereas the fourth one named C, which has only one known member named lymphotactin, contains only two corresponding to the first and third cysteines in the other groups. The CC-chemokines, also known as the beta-chemokines, have adjacent cysteine residues corresponding to the first and second cysteines. They have a major role in attracting monocytes, lymphocytes, basophils and eosinophils. The CXC-chemokines, known as the alpha-chemokines, contain a single amino acid between the first and second cysteine residues. They are neutrophil and T lymphocyte attractants. Finally, the CX_3C-chemokines, which have only one known member named fractalkine, contain three amino acids between the first and second cysteine residues.

Chemokine receptors are cell surface proteins and belong to the superfamily of seven transmembrane G protein-coupled receptors. The chemokine receptors are grouped into families with the chemokine ligands that they bind. Some receptors are promiscuous, while others are selective in terms of ligand binding. The chemokine receptors are widely distributed on hematopoietic and other cell types such as CXCR4 is expressed on naive T lymphocytes, while Duffy antigen is expressed on surface of erythroid lineage. The effect of chemokine receptor-ligand interactions is usually mediated through G protein-coupled interactions. This induces

alterations in cell function including activation and migration which are usually along a chemokine concentration gradient.

4.2 Chemokine receptors and HIV-1 infection

The soluble factors released by CD8+ T cells, RANTES, MIP-1α, MIP-1β, were found to suppression of HIV-1 replication (Cocchi *et al.*, 1995). This data led to the identification of numerous chemokine receptors that functioned as co-receptors for HIV-1 entry into the target cells. The CC-chemokine receptor named CCR5 was identified as a co-receptor for entry of M-tropic HIV-1 strains (Dragic *et al.*, 1996; Deng *et al.*, 1996; Choe *et al.*, 1996; Doranz *et al.*, 1996; Alkhatib *et al.*, 1996). In addition, RANTES, MIP-1α, MIP-1β, which are ligands for CCR5 were shown to inhibit CCR5-mediated infection by M-tropic HIV-1 isolates but not the T-tropic strains (Cocchi *et al.*, 1995; Samson *et al.*, 1996a). The CXC-chemokine receptor named CXCR4 or fusin was shown to be essential for entry of T-tropic HIV-1 strains (Feng *et al.*, 1996). Thus, a ligand for CXCR4, SDF-1, was found to block infection with T-tropic HIV-1 isolates but not with M-tropic HIV-1 (Oberlin *et al.*, 1996; Bleul *et al.*, 1996).

The discovery of the chemokine receptors as the HIV-1 entry co-receptors suggested a new classification of HIV-1 strains based on co-receptor usage (Berger *et al.*, 1998). M-tropic or NSI HIV-1 isolates that usually bind CCR5 are termed R5 strains, T-tropic or SI isolates that bind CXCR4 are termed X4 strains, and dual-tropic strains that can be used as co-receptors are termed R5X4 strains.

In addition to CCR5 and CXCR4, at least nine other chemokine or orphan receptors have been shown to have the HIV-1 co-receptor activity. These include CCR2B, CCR3, CCR8, CCR9, CX_3CR1, GPR1, GPR15/BOB, STRL33/Bonzo, US28, and V28 (Berger *et al.*, 1999; Michael, 1999).

4.3 Factors influence on susceptibility to HIV-1 infection and disease progression

4.3.1 Host genetic variations

Investigations of host genetic variation potentially associated with susceptibility to HIV-1 infection and disease progression have focused on two major parts including genetic

variations of chemokine receptors and their ligands which mediate the entry of HIV-1 and the HLA types which regulate the host immune response to HIV-1 infection.

The finding that chemokine receptors are used by HIV-1 as co-receptors for cellular entry led to the discovery of genetic mutations in chemokine receptors that associated with resistance to HIV-1 infection or the rate of disease progression. The first mutation type reported was a 32 bp deletion in the CCR5 gene (CCR5Δ32) which resulted in a truncated protein that do not presence on the cell surface (Samson *et al.*, 1996b; Liu *et al.*, 1996). Individuals homozygous for this deletion are highly resistant to HIV-1 infection, whereas the heterozygotes progress more slowly (by approximately 2-4 years) to AIDS after infection than individuals without this deletion (Liu *et al.*, 1996; Huang *et al.*, 1996; Dean *et al.*, 1996; Samson *et al.*, 1996b). The second type of mutation is a single nucleotide polymorphism (T to A substitution at position 303) referred to a CCR5-m303. The CCR5-m303 is a rare genetic variation located in the CCR5 coding region that leads to lacking of CCR5 protein on the cell surface and protects against HIV-1 infection (Quillent *et al.*, 1998). The resent studies found that a single nucleotide deletion at nucleotide 893 (CCR5-893(-)), which is an Asian-specific gene polymorphism, results in truncation of CCR5 and reduced level of CCR5 on the cell surface (Ansari-Lari *et al.*, 1997; Shioda *et al.*, 2001).

Several genetic polymorphisms have been identified within the CCR5 promoter region that may effect HIV-1 transmission or disease progression. HIV-1 infected individuals who are homozygous for CCR5P1 haplotype and without the protective alleles CCR5Δ32, CCR5-64I or SDF-1-3'A showed an accelerated rate of progression (by approximately 4 years) to AIDS compared with those individuals with any of the three protective alleles with or without CCR5P1 haplotype (Martin *et al.*, 1998). Similar results were observed a polymorphism (A/G) at position 59029 (numbering according to Genbank, accession number U95626), where HIV-1 infected individual who are homozygous for 59029A allele was association with rapid progression to AIDS (McDermott *et al.*, 1998). Although several promoter polymorphisms have been reported in relation to the possibility that may effect on the levels of CCR5 expression, but the mechanisms of action in HIV-1 transmission or disease progression is remain unclear.

Other chemokine receptor polymorphisms that have also been reported to have an effect on disease progression are CCR2-64I and CX_3CR1-I249/M280. CCR2 is a minor HIV-1

entry co-receptor which comprises 2 isoforms termed CCR2A and CCR2B. The CCR2B is expressed both on the cell surface and in the cytoplasm, while CCR2A is found almost exclusively in the cytoplasm (Wong *et al.*, 1997). The CCR2-64I is a G to A substitution at position 190 in the CCR2 coding region, causing a V to I change at amino acid position 64 in the first transmembrane domain of CCR2B (Smith *et al.*, 1997) and was associated with delayed disease progression. This type of mutation is relatively common among different ethnic groups, with frequencies of 10% in Caucasians, 15% in African-Americans, 17% in Hispanics, and 25% in Asians. CX_3CR1 is a HIV-1 co-receptor as well as a leukocyte chemotactic/adhesion receptor for fractalkine. Faure *et al.* (2000) identified 2 single nucleotide polymorphisms in the CX_3CR1 gene in Caucasians and demonstrated that HIV-1 infected patients homozygous for I249/M280 progressed to AIDS more rapidly than those with other haplotypes. Functional CX_3CR1 analysis showed that fractalkine binding was reduced among patients homozygous with this particular haplotype. Thus, Faure *et al.* (2000) concluded that CX_3CR1-I249/M280 haplotype was a recessive genetic risk factor for HIV/AIDS.

The genetic variations in chemokine ligands that might influence HIV-1 susceptibility or pathogenesis have been identified. SDF-1 is a ligand for CXCR4, a co-receptor with CD4 for T cell-line tropic HIV-1. A common polymorphism, designated SDF1-3'A, was identified in a conserved segment of the 3'UTR of the SDF-1 structural gene transcript (Winkler *et al.*, 1998). The HIV-1 infected individuals homozygous for SDF1-3'A were shown to exhibit delayed progression to AIDS. The SDF1-3'A has been identified with allele frequencies of 21% in Caucasians, 16% in Hispanics, 5.7% in African-Americans, and 25.7% in Asians (Winkler *et al.*, 1998). RANTES is one of the natural ligands for the CCR5 and potently suppress *in vitro* replication R5 HIV-1 strains. Regarding the polymorphism in the promoter region, RANTES-28G, has been associated with slower rates of CD4+ T cells depletion in a group of Japanese HIV-1 infected individuals (Liu *et al.*, 1999). This mutation increases RANTES expression in HIV-1 infected individuals and thus delayed the progression to AIDS, whereas RANTES-403A haplotype was shown to be a risk factor in HIV-1 transmission (McDermott *et al.*, 2000).

The polymorphisms in HLA gene are believed to associate with different rates of disease progression and varying susceptibility to HIV-1 infection. Individuals heterozygous at HLA loci are able to present a greater variety of antigenic peptides than the homozygote, resulting

in a more effective immune response to a large array of pathogens (Carrington *et al.*, 2001). The association of HLA class I genotype with HIV-1 disease progression was studied by Carrington *et al.* (1999). They reported that an individual who was homozygous at HLA-A, HLA-B and HLA-C presents a limited repertoire of antigenic epitopes relative to an individual who is heterozygous at these loci. HLA-Cw*4 and B*35 alleles, which reduced NK cell activity against HIV-1, were found to associated with accelerate progression, while HLA-B*27 has a protective effect on progression to AIDS (Carrington *et al.*, 2001).

Other genetic associations have been implicated in altering the rate of progression to AIDS. IL-10 is a Th-2 cytokine that inhibits macrophage growth and T cell cytokine secretion, and has been reported to inhibit HIV-1 replication in macrophages. An IL10 promoter polymorphism C to A transversion at position -592 (IL10-5'A) has been associated with decreased IL-10 production and has been linked to accelerated progression to AIDS (Carrington *et al.*, 2001). IL-4 is a pleiotropic cytokine produced primarily by activated CD4+ T lymphocytes, mast cells, and basophils, down-regulated the CCR5 but in contrast up-regulated the CXCR4 production. A polymorphism in the IL-4 promoter region, IL-4-589T, increases IL-4 production in human and thus accelerates the phenotypic switch of HIV-1 from NSI to SI and possibly disease progression of AIDS (Nakayama *et al.*, 2000), but however, the individuals with the IL-4-589T allele have lower serum viral load than in those individuals without this allele (Nakayama *et al.*, 2002). Moreover, three polymorphisms in the promoter region of mannose-binding lectin (MBL) gene; G to D at codon 54, G to E at codon 57, and R to C at codon 52, were found to associated with lower MBL serum concentrations but were associated with opsonization defects and impaired phagocytosis in HIV-1 infected individuals (Carrington *et al.*, 2001).

4.3.2 Immune mechanism

The immune response to an infectious agent tends to be predominantly either the generation of cell mediated immunity (CMI) or the production of antibody through humoral mediated immunity (HMI). The detection of HIV-1 specific IgG antibodies implies HIV-1 infection, but there is undetectable in HEPS individuals despite multiple repeated exposure to the virus. These are led to the concept that cellular immune mechanisms may mediate protective immunity to HIV-1 and are correlated with HIV-1 resistance in HEPS persons in the absence of

systemic humoral immunity. Study in the HEPS individuals show that the CMI may play an important role against HIV-1 infection and the predominant T cell response has been shown to be mediated by MHC class I restricted CTLs. MHC class I restricted CTLs to HIV-1 have been reported in the uninfected infants born to HIV-infected mothers (Rowland-Jones *et al.*, 1993), uninfected heterosexual partners of HIV-1 infected individuals and CSWs in Gambia (Rowland-Jones *et al.*, 1995). In the recent study, a group of resistant female sex workers in Kenya and low-risk controls were determined for CTLs by lysis of autologous EBV-transformed B cell lines infected with recombinant vaccinia viruses containing the HIV-1 structural genes *env*, *gag* and *pol*. Fifteen of 22 HIV-1 resistant women compared with 0 of 12 controls showed CTLs response to HIV-1 *env*, *gag* or *pol* (Fowke *et al.*, 2000).

CD8+ T cells also mediate non-cytolytic suppression of HIV-1 replication through secretion of soluble factors. The β-chemokines MIP1-α, MIP1-β and RANTES which are the natural ligands for CCR5 have been shown to inhibit the replication of M-tropic HIV-1 isolates by blocking the co-receptor (Cocchi *et al.*, 1995). The α-chemokine SDF-1 which is the natural ligands for CXCR4, inhibited fusion of cell membrane with the T-tropic isolates of HIV-1 by binding the CXCR4 (Oberlin *et al.*, 1996). In an acute HIV-1 infection, the high levels of CTLs correlate with reductions in circulating viral loads and elevations of CD4+ T cell numbers. Both non-cytolytic CD8+ T cells mediated suppression of HIV-1 replication and MHC class I restricted HIV-specific CTLs are detected in LTNPs and decline with progression to AIDS.

In addition to the soluble factors secreted by CD8+ T cells that suppress replication of HIV-1, the ability of endogenous cytokines was also described. The cytokines including IL-1β, IL-2, IL-3, IL-6, IL-12, tumor necrosis factor (TNF)-α, TNF-β, M-CSF, and GM-CSF could induce HIV-1 expression. TNF-α and TNF-β suppressed HIV-1 replication whereas the transforming growth factor (TGF)-β, IL-4, IL-10, IL-13 and IFN-γ either induced or suppressed HIV-1 expression depending on the culture system (Fauci, 1996).

While the HIV-1-specific cellular immune responses are found in a number of HEPS cohorts, low level of HIV-1-specific IgA antibodies have been found in the genital tract and plasma of several HEPS cohorts and can neutralize HIV-1 infection (Beyrer *et al.*, 1999; Devito *et al.*, 2000a). The studies by Devito *et al.*, 2000b and Broliden *et al.* 2001, showed that IgA purified from genital tract, saliva and plasma from Nairobi HEPS sex workers were able to

neutralize infection of PBMC by NSI isolates and were able to inhibit HIV-1 transcytosis across human epithelial cells; this could be important protecting mechanism applicable for the sexual transmission of HIV-1 infection in HEPS individuals.

4.3.3 Sexually transmitted diseases (STDs)

The pre-existing infection with conventional STDs increases susceptibility to HIV-1 infection (Simonsen *et al.*, 1998). Studies in female sex workers in Kenya and a large number of observational studies from different countries have confirmed the association between STDs and HIV-1 infection (Simonsen *et al.*, 1990; Wasserheit, 1992). The observational study from Tanzania has reported reduction in HIV-1 incidence among rural populations who receiving effective treatment for bacterial STDs compared to those without any treatment (Grosskurth *et al.*, 1995). Although the associations of STDs and HIV-1 infection have been reported, the mechanisms by which STDs enhance HIV-1 transmission are not well understood. Genital inflammation with recruitment of CD4+ T lymphocytes, Langerhans' cells and other macrophages to the genital tract might increase susceptibility to HIV-1 infection by presenting a greater population of cells capable of supporting HIV-1 invasion and replication. In addition, breaches of mucosal barrier may provide ready portals for viral entry. Not all conditions causing genital inflammation are necessarily associated with increased susceptibility to HIV infection. Several genital infections such as bacterial vaginosis, trichomoniasis and *candida vaginitis* have not been consistently linked to an increased risk of HIV-1 acquisition (Simonsen *et al.*, 1998). The recent study in CSWs found that *gonococcal cervicitis* was associated with increased viral shedding and plasma viremia. The increasing susceptibility to HIV-1 in the HEPS-CSWs was associated with the reduction of CD8+ T cell responses (Kaul *et al.*, 2002).

5. CC-Chemokine receptor 5 (CCR5)

5.1 Structure and function of CCR5

CCR5 (also know as CC-CKR-5, CMKBR5, or CD195 antigen) is a 40.6 kDa protein containing 352 amino acid residues (Samson *et al.*, 1996a). It is a member of the chemokine receptors that bind β-chemokines, including macrophage inflammatory protein (MIP)-

1α, MIP-1β, RANTES (regulated upon activation, normal T cell expressed and secreted) (Samson *et al.*, 1996a; Raport *et al.*, 1996), monocyte chemoattractant protein (MCP)-2 (Yang *et al.*, 2002). It also serves as an entry co-receptor for M-tropic strains of HIV-1. Like all other chemokine receptors, the CCR5 belongs to a family of seven-transmembrane domain (7TM) G protein-coupled receptors (GPCRs) which are serpentine seven-transmembrane spanning domains structure. The whole structure consists of 3 major parts; an extracellular amino-terminal (N-terminal) end, three extracellular domains connecting with three intracellular loops and a C-terminal cytoplasmic end. The CCR5 contains four cysteines in the extracellular domain; two of them locate at the first and second extracellular domains and are paired to from disulfide bonds while the other two locate at the N-terminal domain and the third extracellular loop. The extracellular domains of CCR5 are involved in binding of chemokines and HIV-1, whereas intracellular domains are involved in cellular signaling. Several studies shown that, a small tyrosine-rich region of N-terminal domain and several charged residues in the second and third extracellular loops are critical for entry of the M-tropic and dual-tropic strains of HIV-1 (Farzan *et al.*, 1998; Berson and Doms, 1998). The C-terminal domain which is rich in serine and theronine residues is the potential sites for phosphorylation by G protein-coupled receptor kinases (Samson *et al.*, 1996a). The CCR5 is expressed on several distinct cell types including a subpopulation of lymphocytes, activated and memory T cells and monocytes/macrophages, primary and secondary lymphoid organs, astrocytes and microglia in the central nervous system, epithelium, endothelium, vascular smooth muscle, and fibroblasts (Rottman *et al.*, 1997). The predicted structure and amino acid sequence of CCR5 is shown in Figure 3A.

5.2 Characterization of the CCR5 gene

The CCR5 gene has been mapped to the short arm of chromosome 3, which is located in region p21.3 (Liu *et al.*, 1996). Mummidi *et al.* (1997) have shown that the CCR5 gene is organized into four exons and two introns which spanning approximately 6 kb. The introns interrupt between exons 1 and 2 and the exons 3 and 4, respectively. The exon 4 contains an open reading frame (ORF), 11 nucleotides of 5'-untranslated regions (UTR) and complete 3'-UTR. The transcripts are initiated from two distinct promoters termed an upstream promoter (Pu) and a downstream promoter (Pd). Pu is located upstream of exon 1 and Pd is located downstream of

exon 1 in the region between intron 1 and exon 2. The downstream promoter appears to be much stronger than the upstream promoter in monocytic and lymphocytic cell lines, as well as CD4+ T cells (Moriuchi *et al.*, 1997; Mummidi *et al.*, 1997; Guignard *et al.*, 1998; Liu *et al.*, 1998). Sequence motifs similar to consensus sequences for a variety of transcription factors have been identified in the Pd promoter region and mutational analysis of several such sequences indicated their importance in transcription of CCR5 gene (Liu *et al.*, 1998). In addition, the complex alternative splicing of four exons of CCR5 gene give rise to multiple CCR5 transcripts that differ in their 5'UTRs (mRNA with or without exon 2). However, the generation of multiple CCR5 transcripts has no effect on the protein sequence of CCR5 (Mummidi *et al.* 1997). The genomic organization and map of CCR5 gene is shown in Figure 3B.

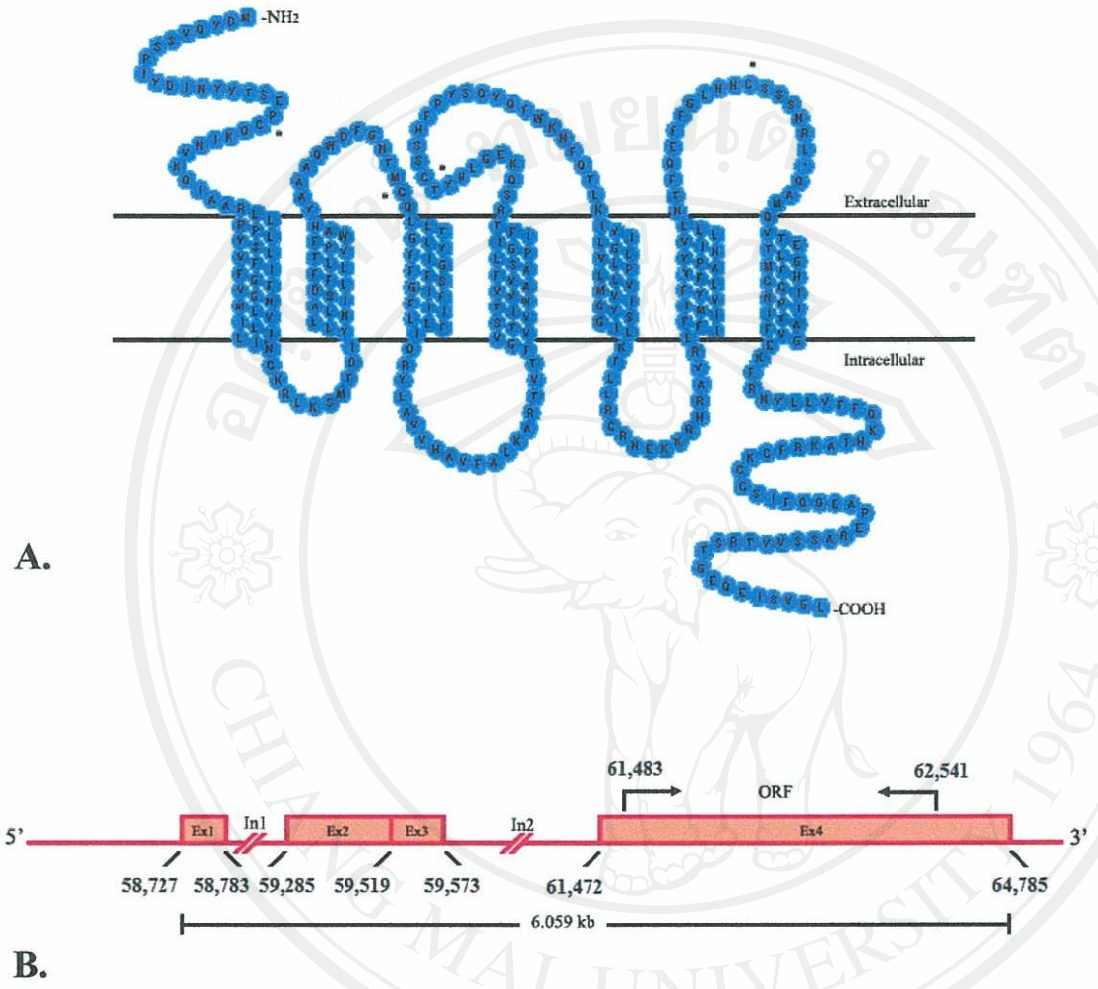

Figure 3 The schematic structure and genomic organization of CCR5; (**A**) the predicted structure and amino acid sequence of CCR5 (McNicholl *et al.*, 1997). Amino acids are listed with a single letter code. Extracellular cysteine residuals are indicated by asterisks, (**B**) the genomic organization of CCR5 gene (not to scale). The four CCR5 exons (open boxes) and two introns are numbered and the ORF is in exon 4. CCR5 gene numbering is based on the reported sequence according to GenBank accession number U95626.

5.3 Variation in the coding region of CCR5 gene

5.3.1 CCR5Δ32

In 1996, Samson *et al* and Liu *et al* identified a 32 bp deletion mutation in the coding region of CCR5 gene (CCR5Δ32). The deletion beginning at nucleotides 794 to 825 of the coding region, which encoded the third extracellular domain of CCR5 protein, results in a frame shift and premature stop codon in the fifth transmembrane domain. The mutant protein with 215 amino acids long is not presence on the cell surface (Liu *et al.*, 1996).

The CCR5Δ32 allele has been found to associate with resistance to HIV-1 infection and the rate of disease progression. The cohorts of HEPS, LTNP and HIV-1 infected individuals have also screened for the presence of the CCR5Δ32 allele. Several homozygous CCR5Δ32 genotype were found in many HEPS cohorts but not in HIV-1 infected individuals (Liu *et al.*, 1996; Samson *et al.*, 1996b; Dean *et al.*, 1996). The primary mononuclear cells from homozygous CCR5Δ32 individuals were resistant to *in vitro* infection with R5 but not X4 HIV-1 isolates (Connor *et al.*, 1996; Paxton *et al.*, 1996; Michael, 1999). This result demonstrated that the homozygous CCR5Δ32 genotype was protective against HIV-1 infection. However, rare HIV-1 infected individuals with homozygous CCR5Δ32 genotype have also been reported (Biti *et al.*, 1997; Theodorou *et al.*, 1997), this indicates that the protective effect of this allele is not absolute. Presumably these individuals were infected by a viral strain that used non-CCR5 dependent mechanisms. The biological study of virus isolated from HIV-1 infected individuals with homozygous CCR5Δ32 genotype demonstrated that the virus isolated was the T-tropic that used CXCR4 for the entry (Michael *et al.*, 1998). The heterozygous CCR5Δ32 genotype has been found in several studies of LTNPs (Michael *et al.*, 1997; Eugen-Olsen *et al.*, 1997). HIV-1 infected individuals who are heterozygous CCR5Δ32 genotype had 2 to 4 years delay progression to AIDS. In addition, the slower rates of CD4+ T cells depletion and lower viral loads were observed when compared to those who were homozygous wildtype genotype (Berger *et al.*, 1999). However, it does not necessary mean that all heterozygous CCR5Δ32 individuals will be the LTNPs.

Distributions of CCR5Δ32 mutation have been determined in populations of different geographical area. The results revealed that CCR5Δ32 allele is highly present in Caucasian populations of European descent. The homozygous individuals were detected in approximately 1% in northern European populations while the heterozygous and wild-type individuals comprise approximately 16% and 83% respectively (Samson *et al.*, 1996b; Martinson *et al.*, 1997). The heterozygosity is found in small numbers of non-Caucasians approximately 6% of African-Americans, 7% of Hispanics, 13% of Native Americans and less than 1% of Asians (McNicholl *et al.*, 1997). However, the individuals who carried one or two CCR5Δ32 alleles have no apparent health problems, suggesting that the normal function of CCR5 is dispensable, perhaps because of compensating function by other chemokine receptors with a similar leukocyte distribution.

5.3.2 CCR5-m303

CCR5-m303 (also know as C101X) is a single point mutation (T to A substitution) at the position 303 of the CCR5 coding region (Quillent *et al.*, 1998), it results in a cys101-to-ter substitution. The m303 mutation generates a premature stop codon in a coding sequence of the first extracellular loop of CCR5. This leads to lacking of the functional CCR5 on the cell surface. Similarly with CCR5Δ32 allele, the resistibility against HIV-1 infection in individual with m303 allele was described. This mutation was first observed in 1 of 18 men who had frequent unprotected sexual intercourses with a HIV-1 seropositive partner and whose other allele carried CCR5Δ32. The PBMCs from him were resistant to *in vitro* infection by R5 HIV-1 viruses (Quillent *et al.*, 1998).

5.3.3 CCR5-893(-)

CCR5-893(-) (also know as 893delC) is a single nucleotide deletion at nucleotide 893 which is observed exclusively in Asian populations (Ansari-Lari *et al.*, 1997). This deletion causes a frame shift at codon 299 and results in premature termination of translation. The CCR5-893(-) gene product lacks 51 amino acid residuals in the C-terminal cytoplasmic tail and 3 residuals in the last transmembrane segment, and gains 10 amino acid residuals encoded in the -1 ORF in its C terminus (Shioda *et al.*, 2001). The mutant protein is composed of 308 amino acids,

whereas the wild-type protein is composed of 352 amino acids. Peripheral blood CD4+ T cells obtained from individuals heterozygous for this allele expressed very low levels of CCR5 on the cell surface, suggesting that the CCR5-893(-) mutation affects intracellular transport of CCR5 and raise the possibility that this mutation affects the HIV-1 transmission and disease progression (Shioda *et al.*, 2001). The CCR5-893(-) had an allelic frequency of approximately 4% in Japanese and Chinese populations (Ansari-Lari *et al.*, 1997). Although, this mutation was absent from 22 of HIV-1 discordant Thai couples (Louisirirotchanakul *et al.*, 2002).

5.3.4 Additional coding region variants of the CCR5 gene

Twenty additional alleles of the CCR5 coding region have been described (Ansari-Lari *et al.*, 1997; Carrington *et al.*, 1997; Carrington *et al.*, 1999; Magierowska *et al.*, 1999) (Table 1). Six of twenty alleles (I12T, C20S, A29S, I42F, L55Q and A73V) were studied for the ligand signaling and influence on HIV-1 infection. The results demonstrated that single amino acid changes in the extracellular domains (I12T, C20S and A29S) had affected both HIV-1 co-receptor and specific ligand-induced functions, whereas the mutations in the transmembrane domain (I42F, L55Q and A73V) only affect the response to chemokine ligands (Howard *et al.*, 1999).

Table 1 Genetic variations of the coding region of CCR5 gene

Position	Amino acid variation	Nucleic acid variation[1]
Extracellular domain 1	I12T	A25C
Extracellular domain 1	C20S	T58A
Extracellular domain 1	A29S	G85T
Transmembrane domain 1	I42F	A124T
Transmembrane domain 1	L55Q	T164A
Intracellular domain 1	R60S	G180T
Intracellular domain 1	S63C	A187T
Transmembrane domain 2	A73V	C218T
Transmembrane domain 2	S75S	T215C
Extracellular domain 2	C101X	T303A[2]
Transmembrane domain 4	I164I	C492A
Extracellular domain 3	C178R	T532C
Extracellular domain 3	Δ32 (185)	del32[3]
Transmembrane domain 5	L215S	C664T
Intracellular domain 3	R223Q	G668A
Intracellular domain 3	228delK	680del3
Transmembrane domain 7	299 (FS)	893delC[4]
Transmembrane domain 7	V300V	C900A
Transmembrane domain 7	G301V	G902T
Intracellular domain 4	R319H	G956A
Intracellular domain 4	P332P	C996T
Intracellular domain 4	A335V	C1004T
Intracellular domain 4	Y339F	A1016T

[1] The numbering system used designates the first nucleotide of the translation start site of CCR5 gene as position 1 (the CCR5 AIDS symposium held at the NCI-FCRDC, Frederick, MD, on April 30, 1999)

[2] CCR5-m303

[3] CCR5Δ32

[4] CCR5-893(-)

5.4 Variation in the promoter region of CCR5 gene

The genetic polymorphisms within the CCR5 promoter region have been identified, some of which could potentially disrupt transcription factor binding motifs accounting for some of the heterogeneity in CCR5 expression among individuals (Carrington *et al.*, 1999). Several studies have demonstrated that the polymorphisms in the CCR5 promoter region may affect the HIV-1 transmission or disease progression and possibly effects on levels of CCR5 expression. Ten nucleotide polymorphisms in the CCR5 promoter region which compose of 4 common multisite alleles (CCR5P1-P4) and 6 rare multisite alleles (CCR5P5-P10) (Martin *et al.*, 1998) have been identified (Table 2). HIV-1 infected individuals homozygous for the P1 haplotype and without the protective alleles (CCR5Δ32, CCR2-64I or SDF1-3'A) showed an accelerated rate of progression to AIDS when compared to those individuals with any of the protective alleles with or without CCR5P1, particularly in the early years after infection. Similar results were observed in another study for an A/G polymorphism at position 59029 (according to GenBank accession number U95626), where HIV-1 infected individuals with CCR5-59029A/A genotype progressed to AIDS (by approximately 4 years) more rapidly than those with CCR5-59029G/G genotype (McDermott *et al.*, 1998). Other promoter polymorphisms have been identified and also associated with disease progression. The mutant CCR5-59356T allele is relatively common in African-Americans and rare in Caucasians and Hispanics. Homozygosity for CCR5-59356T has been strongly associated with an increase rate of perinatal HIV-1 transmission (Kostrikis *et al.*, 1999).

The nucleotide variations found in the coding and promoter regions of CCR5 gene are summarized in Figure 4.

Table 2 CCR5P alleles

		Position on the CCR5 promoter region									
Numbering systems	New [1]	-2554	-2150	-2136	-2135	-2132	-2115	-2086	-2078	-2048	-1951
	Martin *et al.* [2]	208	612	626	627	630	647	676	684	714	811
	GenBank no. U95626 [3]	58934	59338	59352	59353	59356	59373	59402	59410	59440	59537
CCR5P haplotypes	P1	G	A	C	C	C	C	A	T	C	G
	P2	.	.	.	T	.	.	.	.	.	.
	P3	T	.	.	T	T	.	.	.	.	.
	P4	T	.	.	T	.	.	G	.	.	.
	P5	-	G	.	T	.	.	.	.	.	.
	P6	-	.	.	.	T	.	.	C	.	.
	P7	-	.	A	.	.	.	.	.	.	.
	P8	-	.	.	T	.	A	.	.	.	.
	P9	-	.	.	T	.	.	.	.	.	A
	P10	-	.	.	T	.	.	G	.	G	.

Symbols: (.), consensus to P1; (-), sequence not known

[1] The numbering system used designates the first nucleotide of the translation start site of CCR5 gene as position 1 (the CCR5 AIDS symposium held at the NCI-FCRDC, Frederick, MD, on April 30, 1999).

[2] according to Martin *et al.*, 1998

[3] according to GenBank accession number U95626

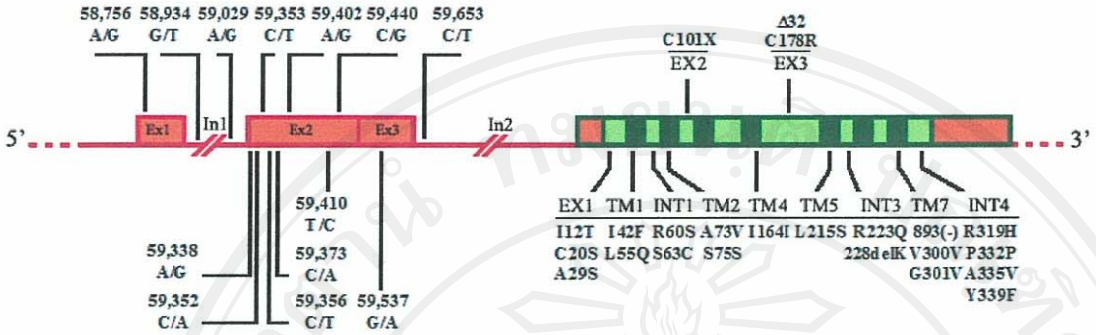

Figure 4 CCR5 polymorphisms; the map of the reported polymorphisms found in the CCR5 promoter and coding region. The numbers of the nucleotide polymorphisms in the promoter region refer to nucleotide positions in GenBank accession numbers U95626, while the numbers of the amino acid variants in the coding region refer to amino acid position of CCR5 protein.

The central role of the CCR5 molecule in HIV-1 infection has led to characterizing its functional and genetic properties. In summary, the mutant CCR5Δ32 allele has been shown to provide almost completes protection against HIV-1 infection in homozygous individuals. Heterozygous individuals also exhibit slower progression to AIDS after seroconversion. The promoter region of CCR5 gene has been characterized by several groups, and it appears that polymorphisms in this region may have effect on AIDS progression, possibly due to their effect on levels of CCR5 expression. These studies should prove useful not only in predicting outcome to HIV-1 infection, but also in developing novel therapeutic strategies.