CHAPTER 1

INTRODUCTION

1.1 Statement and significance of the problem

Highly active antiretroviral therapy (HAART) has been offered as a discipline to HIV-infected patients. The combined use of powerful antiretroviral drugs can prolong life of patients since it significantly control HIV replication. However, undesirable side effects and the prevalence of drug-resistant strains are the major problematic effects in patients receiving long-term treatment (Booth and Geretti, 2007; D'Aquila and Walker, 1999; Hammer et al., 2006). For these reasons, new therapeutic approaches are prerequisite for the purpose of limiting HIV replication. Antibodies are considered as natural therapeutic agents and have become increasingly accepted to slow down HIV propagation (Huber, Olson, and Trkola, 2008; Parren and Burton, 2001).

A key feature of antibodies for HIV treatment is the ability to bind to targets with high specificity and affinity, leading to the interference of HIV replication (David, 2007; Maynard and Georgiou, 2000; Ugolini et al., 1997). Extracellular antibodies can prevent HIV-1 entry by neutralization of virus attachment at the early phase (Parren and Burton, 2001; Stiegler and Katinger, 2003; Ugolini et al., 1997). On the other hand, intracellular antibodies or intrabodies can block the propagation of viral progeny by interfering different step i.e. HIV assembly. In order to propose for

therapeutic interventions, antibodies have been used extensively as diagnostic tools in HIV infection (Sharma A., 2008; Zhang, 2002).

Focusing on the HIV matrix (MA) protein, MA has been investigated as a target molecule for combating HIV replication due to its multiple structural and physiological functions during viral infection and replication, both in early stages and late stages (Bukrinsky et al., 1993; Fäcke et al., 1993; Spearman et al., 1994) using MA-specific antibodies (Naylor et al., 1987; Papsidero, Sheu, and Ruscetti, 1989). Tewari et al. (1998; 2003) reported that the intracellular expression of a single chain variable fragment (scFv-MA HB-8975) derived from a monoclonal anti-matrix antibody (MH-SVM33C9/ATCC HB-8975) (anti-MA mAb HB-8975) was able to inhibit HIV-1 replication. This anti-MA mAb specifically recognizes the C-terminal epitope (DTGHSSQVSQNY) of HIV-1 MA domain located in the HIV protease (HIV-PR) cleavage site (Tewari et al., 1998; Tewari, Notkins, and Zhou, 2003). Consequently, it would be valuable to evolve and apply scFv-MA HB-8975 as biotechnological and therapeutic agents for HIV therapy.

The antibody-based therapeutic was established to have a high antiviral potential against HIV-1, however viral-specific antibodies still necessitated to be improved for their binding affinity and specificity to ensure in treatment efficacy (Abela, Reynell, and Trkola, 2010; Stiegler and Katinger, 2003). As a consequence, many antibody engineering strategies have been adopted to overcome this. Recent advances in computation-based protein analysis have provided the fundamental understanding of the antibody-antigen interaction that is important for improving the

ability and affinity of antibody (Correia et al., 2011; Malik et al., 2010; Marvin, 2005). In deed, the binding activity and the key amino acids involved in antibody-antigen recognition can be accurately predicted and this predicted information can guide the design of antibody with improved binding affinity and specificity by performing site-specific mutagenesis at the key residues (Honegger et al., 2005; Kamei, Shimazaki, and Nishi, 2001; Kipriyanov, 1997). Accordingly, we aimed to characterize the key residues of scFv-HB-8975 interacting with HIV MA protein using the computational assisted modeling and structural-based protein analysis. The relevant data from computational analysis would be applicable to enhance capabilities of scFv-MA HB-8975 by performing site-direct mutation to develop a potential therapeutic antibody for HIV treatment.

Beside the improvement of scFv by performing site-specific randomized mutagenesis, the use of display vectors and *in vitro* selection strategy have been proposed for isolating efficient scFvs against defined antigen (Bradbury et al., 2011; Hoogenboom et al., 1998). One of the most successful display technologies is the phage display that can isolate antibodies from large combinatorial libraries (Hoogenboom et al., 1998; Marks et al., 1991; Sblattero and Bradbury, 2000). Nevertheless, the folding and post-translation modification of complex proteins in bacteria host are a severe constraint on phage display strategy, eukaryotic virus display has been resolved these restrictions. The baculovirus display-technology has been successfully used for several different applications and has lately been diversified into presentation of targeting moieties on the surface of baculovirus

(Grabherr et al., 2001; Mottershead et al., 1997; Mottershead et al., 2000; Ojala et al., 2001). Moreover, gene delivery to numerous mammalian cell lines of different origin has been successfully demonstrated. Therefore baculovirus can be regarded as an attractive candidate for targeted gene delivery (Ghosh et al., 2002; Kenoutis et al., 2006; Kost, Condreay, and Jarvis, 2005). Thus, our goal was to engineer the scFv-MA HB-8975 and generate a baculovirus-displayed scFv-MA HB-8975 to be applied as gene transfer vehicle in HIV gene therapy. By combining computational approaches and baculovirus display technology would be capable to produce and appropriately deliver improved scFv intrabodies against HIV MA protein whicn will block the production of virus.

Considering the particular recognition of anti-MA mAb HB-8975 specific to the free C-terminal MA located on HIV-PR cleavage site, we proposed to apply this anti-MA mAb HB-8975 for discriminating HIV-PR activity based on an enzyme-linked immunosorbent assay. We successfully established the simplified ELIS-based HIV-PR activity assay or ELIB-PA using the H_6MACA protein as a substrate and two specific mAbs, anti-MA mAb HB-8975 and anti-CA mAb, for detection purpose. In addition, the ELIB-PA could be used to evaluate the efficacy of HIV protease inhibitors i.e. lopnavir (LPV), ritonavir (RTV), and nelfinavir (NFV), against HIV-PR by determining the median 50% inhibitory concentration (IC_{50}) and inhibition rate. Accordingly, the ELIB-PA would be capable for investigation of HIV-PR susceptibility from drug-resistant individual with different types of PI prior to drug administration. Additionally, the ELIB-PA can be applied as a high-throughput

screening platform to identify novel HIV-PR inhibitors from a large computerized drug design which thus will accelerate the drug discovery in the pharmaceutical-industry.

1.2 Literature review

1.2.1 Antibody-based therapeutic for HIV/AIDS therapy

1.2.1.1 HIV-1 and AIDS

Human Immunodeficiency Virus 1 (HIV-1) has been identified as the causative agent of Acquired Immuno Deficiency Syndrome (AIDS) in 1983. In 2009, there were over 33 million people were living with HIV and the current number of new infected patients is increasing (UNAIDS/WHO Report 2009). HIV-1 is one of the five common causes of death worldwide since more than 25 million people have died from AIDS (Barre-Sinoussi, Chermann et al. 1983; Piot, Bartos et al. 2001).

1.2.1.2 A standard regimen of HIV/AIDS treatment

The recommended treatment of HIV is a combination therapy with multiple-class of antiretroviral drugs. Highly Active Antiretroviral Therapy (HAART) is the mainstay of treatment for HIV-infected patients since HAART introduced the decline of mortality and mobility rates in HIV-infected individuals. Currently, there are six classes of HAART consisting of non-nucleoside reverse transcriptase inhibitors (NNRTIs), nucleoside reverse transcriptase inhibitors (NRTIs), protease inhibitors (PIs), fusion inhibitors, CCR5 antagonists, and integrase inhibitors. The

recommend HIV treatment regimens typically include three or four inhibitors from at least two different drug classes (Thompson et al., 2010). Although HAART regimens are able to profoundly suppress HIV replication resulting in rapid increase number of the survival of HIV/AIDS patients, the emergence of HIV-drug resistance and side effects have limited the usefulness of all classes of HIV inhibitors (Waninger et al., 2004). Therefore, additional approaches have been investigated for controlling HIV replication. Recently, an antibody-based therapeutics regarding to passive immunization therapy and gene therapy based on intrabodies has been represented as alternative approaches (Tewari, Goldstein et al. 1998; Zwick, Labrijn et al. 2001).

1.2.1.3 Antibodies and antibody fragment

Immunoglobulins (Ig) are highly specific and naturally evolved molecules that recognize and eliminate foreign antigens. There are five classes of immunoglobulins; IgM, IgG, IgE, IgA, and IgD (IgG is the most common use of antibodies in a biological perspective). The basic structure of IgG consists of two identical heavy chains (50 kDa) and two identical light chains (25 kDa). The heavy chains contain a variable (V) domain (V_H) characterized by the variability in amino acid sequence, and three constants (C) domains (C_H), while light chains consist of a V domain (V_L) and a single C domain (C_L). Importantly, each V_H and V_L contains three hypervariable regions called complementarity determining regions (CDRs) which form an antigen binding site or a pocket to receive the antigen. The CDRs regions are mainly responsible for the specificity and the affinity of antibodies to the antigen. To

form antibody structure, heavy chains and light chains are held together by disulfide bonds leading to a 'Y' shaped protein molecule **(Figure 1.1A and 1.1C).**

Protein engineering technology has been exploited to create a wide variety of functional antigen-binding fragments to increase number of applications in biotechnology and therapeutic (Wörn and Plückthun, 1998). Recently, the most frequently used antibody format is a single chain variable fragment or scFv which is composed of the variable regions of heavy chain (V_H) and light chain (V_L) joined together by a peptide linker such as a glycine-serine linker, generally $(Gly_4Ser)_3$ (Williams and Zhu, 2006). This scFv format is the smallest antibody fragment which retains the binding specificity as the parental molecule as shown in **Figure 1.1B**.

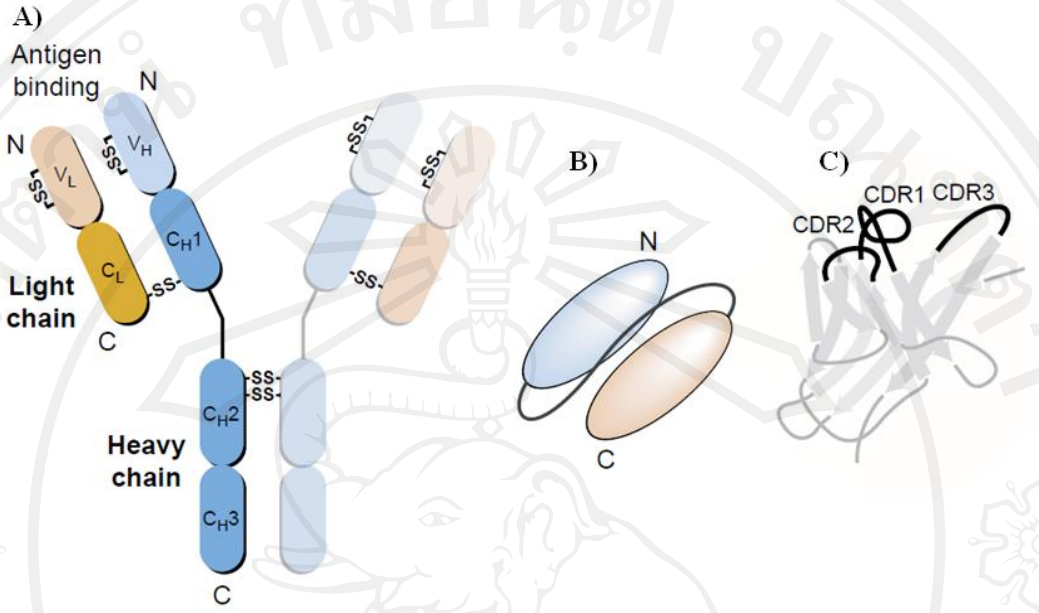

Figure 1.1 Antibodies and antibody fragments. **(A)** Antibodies comprise two identical heavy (H) chains (blue) and two identical light (L) chain (orange), linked by interchain and intrachain disulphide bonds (-SS-). **(B)** The single-chain Fv (scFv) antibody fragment consists of a V_H (blue) and a V_L (orange) domain, connected by a flexible linker peptide **(C)** Tertiary structure of a single V_H or V_L domain of scFv, showing the three complementarity-determining regions (CDRs) that form the antibody binding site (Lobato and Rabbitts, 2003).

1.2.1.4 Therapeutic antibodies of HIV treatment

Extracellular antibodies

The protective effects of antibodies have been known for almost century. Several human monoclonal antibodies (hmAbs) were established to exhibit a high potential antiviral against HIV-1. Neutralizing antibodies employed multiple mechanism to interfere HIV replication. A primary target of neutralizing antibodies (NAb) is the envelope proteins, gp120 (Buchacher 1994; Trkola, Purtscher et al. 1996; Parren, Mondor et al. 1998) and gp41 protein, which are essential in viral infection. The main function of NAb is to neutralize free virions by protecting receptor engagement or interfering with the fusion process resulting in the inhibition of HIV entry **(Figure 1.2A).** On the other hand, even though non-neutralizing antibodies lack of direct inhibition of HIV infection, these antibodies could potentially activate several effector functions. Non-neutralyzing antibodies are capable of viral clearance through phagocytosis of immune complexed virus **(Figure 1.2D)**, lysis virus particles by complement activation **(Figure 1.2C)** and lysis of infected cells by activation of antibody-dependent cellular cytotoxicity (ADCC) (Sinclair 1988; Blue, Spiller et al. 2004; Hangartner 2006) **(Figure 1.2B).** Therefore, extracellular antibodies could be useful in HIV-1 treatment.

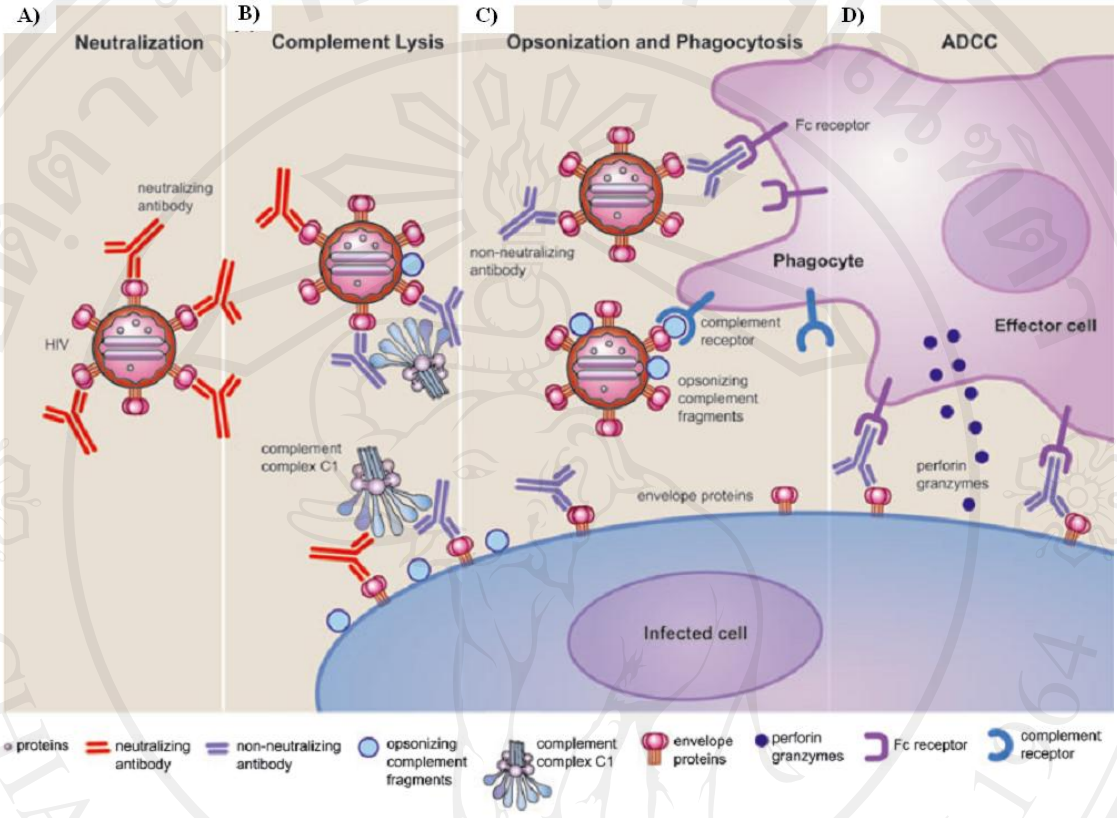

Figure 1.2 How antibodies combat HIV-1. (A) Neutralization of free virus by antibodies, **(B)** complement-mediated lysis of free virus and infected cells triggered by antibodies, **(C)** opsonization of virus particles by antibodies and phagocytosis of virus particles via Fc- or complement-receptors, **(D)** antibody-dependent cellular cytotoxicity (ADCC) against infected cells. Neutralizing antibodies (red), non-neutralizing (blue), Fc-receptors (violet), complement components (light-blue), complement-receptors (black) (Huber and Trkola, 2007).

Intracellular antibodies

Intrabodies are defined as antibody molecules which are expressed intracellularly and directed to defined subcellular compartments and the common target cells are mammalian cells (Williams and Zhu, 2006). A scFv is the most frequently used recombinant antibody format of intrabody since they are small enough to be easily introduced and functioned inside cells. The rational of using intrabodies is to induce a phenotypic knockout of a relevant target molecule either by directly inhibiting the function of the antigen or by diverting it from its normal intracellular location. In some cases, intrabodies have also been used to restore the function of a target antigen and thus rescuing a phenotype. Thus, intrabody therapy combines the specificity of antibodies with a gene-therapeutic strategy to selectively affect an intracellular target protein. As the intrabodies are produced only inside the cells, this strategy has advantages regarding safety and efficacy (Kontermann, 2004).

Indeed, intracellular HIV-1-specific scFv or intrabodies can target and redirect essential HIV-1 proteins away from required subcellular compartments and block the function or processing of such essential proteins as HIV-1 gp120 (Marasco, Haseltine, and Chen, 1993), Rev (Vercruysse et al., 2010), Gag (Levin et al., 1997), reverse transcriptase (RT) (Duan et al., 1994), and integrase (IN) (Levy-Mintz et al., 1996) **(Figure 1.3).**

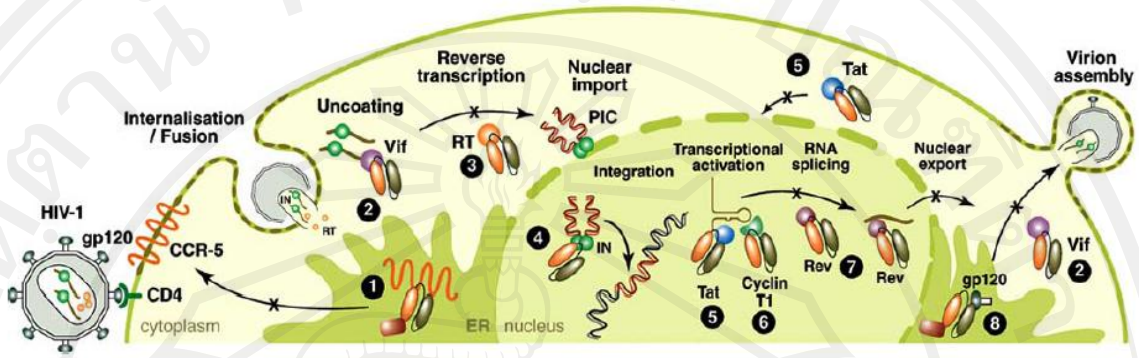

Figure 1.3 Schematic representation of modes of intrabody action against potential therapeutic targets in HIV-1 infected. The figure shows a HIV-1 virus and all the HIV and cellular proteins targeted by specific scFv intrabodies in infected CD4+ T cell (Lobato and Rabbitts, 2003)

1.2.2 Role of HIV matrix (MA) protein in HIV life cycle

The HIV-1 matrix (MA) protein, 17 kDa, is generated after cleavage of the Gag polyprotein by the HIV protease (HIV-PR) and generally lines the inner surface of the viral membrane to form a matrix layer within the HIV virion (Eric O, 1998). MA has long been known to play a central role during HIV life cycle as shown in **Figure 1.4.** In the early stage of a new infection MA dissociates from the viral membrane and targets the core-derived pre-integration complex (PIC) to the nucleus of host cells *via* its nuclear localization signal (Bukrinsky, Haggerty et al. 1993; Gallay, Swingler et al. 1995). Furthermore, newly synthesized MA migrates and accumulates into the cytoplasm and binds to viral RNA promoting its transfer to the plasma membrane (Yuan et al., 1993). Finally, MA plays a role in the assembly and

release of mature viral particle (Fäcke et al., 1993; Spearman et al., 1994). For these reasons, MA is the target of neutralizing antibodies against HIV-1 (Naylor, Naylor et al. 1987; Buratti, Tisminetzky et al. 1997).

Focusing on maturation step, HIV assembly initially occurred by polymerization of the Gag polyprotein, which forms a spherical shell associated with the inner membrane of freshly budding particle. The MA domain is tightly associated with inner surface to stabilize the viral membrane **(Figure 15.B; left part)** (Morikawa, Goto, and Sano, 1999). After budding, HIV-PR is prerequisite to cleave the Gag polyprotein at the protease cleavage sites into distinct smaller proteins: matrix (MA), capsid (CA), p2, nucleocapsid (NC), p1, and p6 as shown in **Figure 1.5A** (Eric O, 1998). The initial cleavage appears between p2 and NC, then between MA and CA, and the last digestion between CA and p2 as shown in **Figure 1.5A; red mark and Figure 1.5C** (Wiegers et al., 1998). The processed proteins rearrange to form mature virion or infectious virus. Upon maturation, the MA remains associated with the inner viral membrane, while CA, NC, and the viral RNA condense into the center of virus **(Figure 15.B).** (Kräusslich et al., 1995)

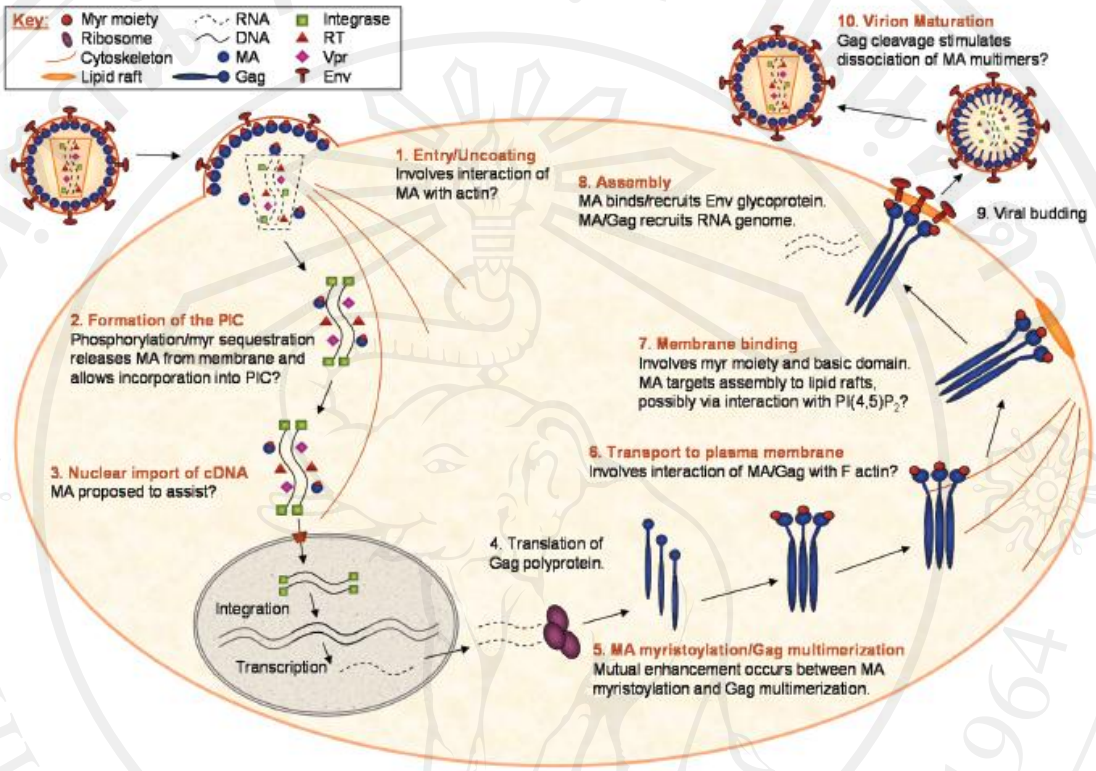

Figure 1.4 Role of MA protein in HIV life cycle. Schematic outlining the steps of HIV-1 infection where MA is thought to play a fundamental role including possible regulatory mechanisms involved in mediating each of these processes as listed. Mechanisms with a question mark are those for which there is some evidence, but further characterization is required. N.B. Detail is not given for stages of HIV-1 infection of which there is involvement of MA (Hearps and Jans, 2007).

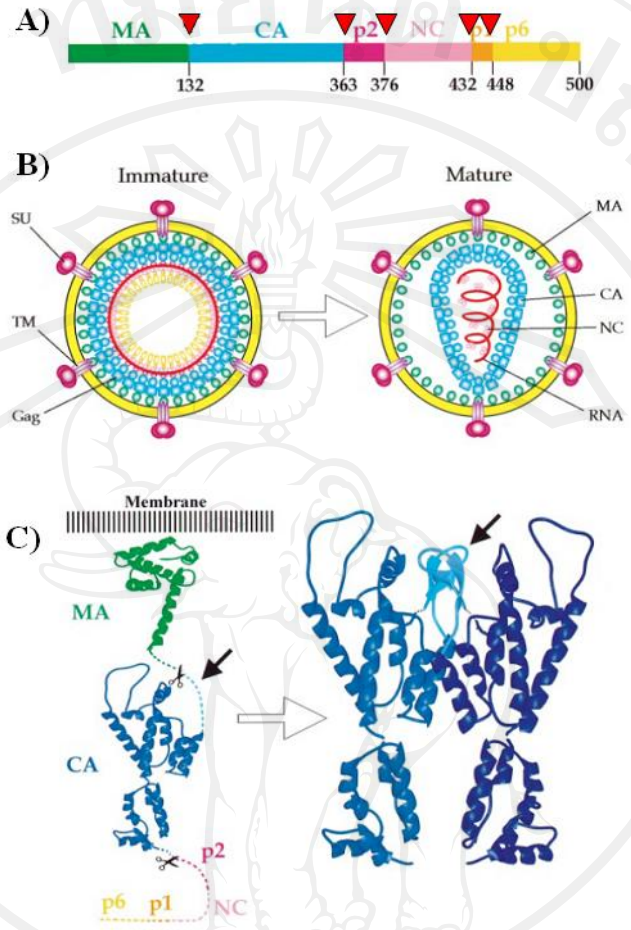

Figure 1.5 Structure and maturation of the HIV-1 virion. **(A)** Domain structure of the HIV-1 Gag polyprotein. Color coding for the different domains of Gag is the same throughout this figure. The HIV protease cleavage sites are indicated in red mark. **(B)** Schematic structures of immature and mature HIV-1 particles. The figure summarizes current models for the locations of the major virion components and emphasizes the dramatic structural rearrangements that accompany viral maturation. **(C)** Model for structure of Gag polyprotein with proteolytic cleavage sites of HIV-PR. Ribbon diagrams of the MA and CA domains of the unprocessed Gag protein are based upon the crystal structures of MA (Hill, 1996). Modified from (von Schwedler et al., 1998)

1.2.3 The effective antibody against MA (anti-MA mAb HB-8975)

Tewari et al. (1998) generated a scFv derived from the anti-MA HB-8975 (scFv-MA HB-8975) murine hybridoma cell line, MH-SVM33C9/ATCC HB-8975, which was obtained from the American Type Culture Collection (ATCC, Manassas, VA). This hybridoma cell line was generated by immunizing animals with lysates of purified HIV-1 then spleen cells were fused with NS-1 myeloma cells. The reactivity of anti-MA HB-8975 was investigated with Gag MA on HIV-1 Western blot strips (Dupont New England Nuclear, Boston, MA) which bound to a C-terminal epitope (DTGHSSQVSQNY) on MA domain located in the HIV-PR cleavage site. (Robert-Hebmann et al., 1992). Interestingly, Tewari et al. (1998; 2003) demonstrated that intracellular expression of this scFv-MA HB-8975 markedly reduced HIV-1 replication when scFv-MA HB-8975 was expressed in cytoplasm rather than nuclease (Tewari et al., 1998). Furthermore, the scFv-MA HB-8975 fused with C_κ (scFv-MA/C_κ) was constructed and expressed in Jurkat cells. As a result, HIV-1 replication was inhibited nearly 90% in scFv-MA/C_κ-tranduced Jurkat cells. More importantly, HIV-1 replication in primary human T cells was abolished as much as 99% when the scFv-MA/C_κ was expressed. Therefore, the expression of scFv-MA/C_κ in primary human T cells renders these T cells resistant to HIV-1 (Tewari, Notkins, and Zhou, 2003). The potential use of scFv-MA HB-8975 may be useful in anti-HIV-1 therapy.

1.2.4 Recombinant protein expression

1.2.4.1 Recombinant protein expression in *E. coli*

E. coli is one of the most widely used hosts for the production of heterologous protein because of its ability to grow rapidly and at high density on inexpensive substrates, its well-characterized genetics and the availability of an increasingly large number of cloning vectors and mutant host strains. The genetics of *E. coli* are far better characterized than those of any other microorganism. Recent progress in the understanding of transcription, translation, and protein folding in *E.* coli, together with serendipitous discoveries and the availability of improved genetic tools are making this bacterium more valuable than ever for the expression of complex eukaryotic proteins. Although there is no guarantee that a recombinant gene product will accumulate in *E. coli* at high levels in a full-length and biologically active form, a considerable amount of effort has been directed at improving the performance and versatility of this workhorse microorganism (Makrides 1996; Baneyx and Mujacic 2004).

Protein translocation across bacterial cytoplasmic membrane

The targeting and transport of proteins across biological membranes into the periplasm is one of the fundamental features of cellular life. Proteins located within the periplasmic space perform many crucial roles. For example, the detoxifying enzymes play a role in the inhibition of the activity of molecules which are toxic to cell. Nucleases, peptidases and other scavenging enzymes metabolize

large complex molecules into simpler ones that can be utilized by the cell. Proteins that are exported to the bacterial periplasm are usually synthesized with cleavable N-terminal signal sequences, termed signal peptides, which direct the protein to a specific transporter complex in the cytoplasmic membrane. The signal sequences have generally a tripartite structure where a short, basic N-region precedes a longer hydrophobic stretch of amino acid (h-region), followed by the c-region, which normally contains a recognition sequence for the enzyme signal peptidase. In bacteria, three major routes are used to achieve protein translocation across the cytoplasmic membrane *e.g.* the secretory (Sec), signal recognition particle (SRP) and the twin-arginine translocation (Tat) pathway as shown in **Figure 1.**6.

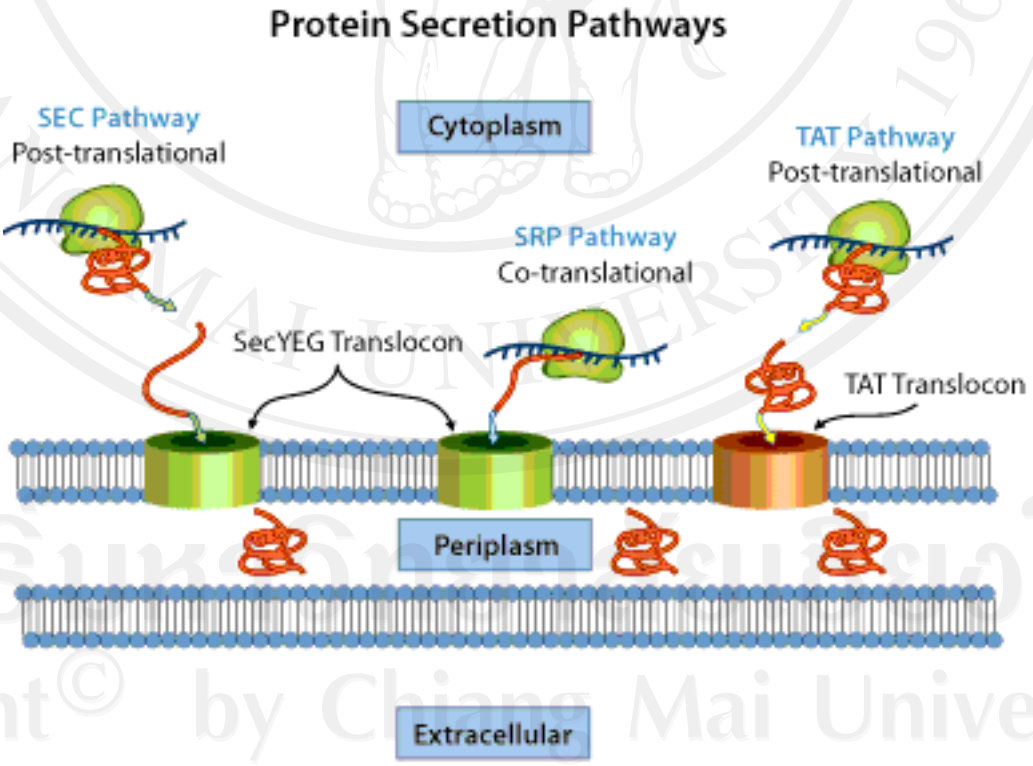

Figure 1.6 Diagram of the secretion pathways of *E. coli*.

(http://www.athenaes.com/tech_brief_ACESyebf.php)

The general secretory (Sec) pathway

The major route of protein translocation in bacteria is the so-called general secretion pathway (Sec pathway) (Hammer et al., 2006). This transporter complex is build by multimeric proteins which are spanning through the inner membrane of *E. coli*. Secretion *via* this pathway involves 9 different components: Trigger factor (TF), SecA, SecB, SecY, SecE, SecG, SecD, SecF and YajC. The core of the translocase consists of a proteinaceous channel formed by the protein complex of SecYEG and the peripheral adenosine triphosphatase (ATPase) SecA as molecular motor. Ribosome-associated nascent chains of secreted proteins bind TF (**Figure 1.7**), which is bound to the ribosomes (Maier et al., 2003). This association is maintained until the preprotein leaves the ribosome, thus preventing cotranslational binding of the nascent chain to SRP components (Mergulhao, Summers, and Monteiro, 2005).

Secreted proteins targeted to the SecB-dependent pathway contain an amino-terminal signal peptide that functions as a targeting and recognition signal. These signal peptides are usually 18–30 amino acid residues long and are composed of a positively charged amino terminus (n-region), a central hydrophobic core (h-region), and a polar cleavage region (c-region) (Fekkes and Driessen 1999; Choi and Lee 2004). The n-region is believed to be involved in targeting the preprotein to the translocase and binding to the negatively charged surface of the membrane lipid bilayer. Increasing the positive charge in this region has been shown to enhance translocation rates, probably by increasing the interaction of the preprotein with SecA (Fekkes and Driessen, 1999). The h-region varies in length from 7 to 15 amino acids.

Translocation efficiency increases with the length and hydrophobicity of the h-region, and a minimum hydrophobicity is required for function (Kollman et al., 2000).

Secreted proteins are kept in a translocation-competent state by the chaperone SecB (de Gier and Luirink, 2001), which interacts with the mature region of the preprotein to prevent premature folding (Khokhlova and Nesmeianova 2003) and targets it to SecA **(Figure 1.7).** In the presence of preprotein, SecB binds SecA (Fekkes, de Wit et al. 1998; Woodbury, Topping et al. 2000), thus releasing the precursor protein that is transferred to SecA (Fekkes and Driessen, 1999). SecA binding to the preprotein is facilitated by the signal peptide, which it recognizes specifically. At this point SecA is bound to the SecY subunit of the SecYEG complex. SecYEG constitutes a pathway (‘channel’) for polypeptide movement. Binding of ATP at one of the two ATP-binding sites on SecA causes the release of SecB from the membrane (Mergulhão, Summers, and Monteiro, 2005). There is no consensus on how the Sec components form a functional translocon, and monomeric, dimeric and oligomeric translocons have been proposed (Mergulhao, Summers, and Monteiro, 2005). Binding of the preprotein to membrane-bound SecA results in the translocation of approximately 20 amino acids, and subsequent binding of ATP to SecA promotes SecA membrane insertion and translocation of additional 15-20 amino acids. ATP hydrolysis releases the preprotein from SecA into the translocation channel (Driessen, Fekkes, and van der Wolk, 1998). ADP is then released and SecA deinserts from the membrane where it can be exchanged with cytosolic SecA. Multiple rounds of SecA insertion and deinsertion promote protein translocation through the channel (de Keyzer, van der Does, and Driessen, 2003). Proton-motive force (PMF) can complete

translocation when the preprotein is halfway through the translocase, even in the absence of SecA (Nishiyama, Fukuda et al. 1999). The mechanism by which PMF drives translocation is unknown but it has been suggested that PMF assists in the initiation phase of protein translocation (Kang et al., 2003) and that it accelerates SecA membrane deinsertion (Nishiyama, Fukuda et al. 1999). Finally, the folded substrate protein is released into the periplasmic space. SecD, F and YajC are accessory proteins that aid in translocation into the periplasm, where proteins are folded into their final conformation.

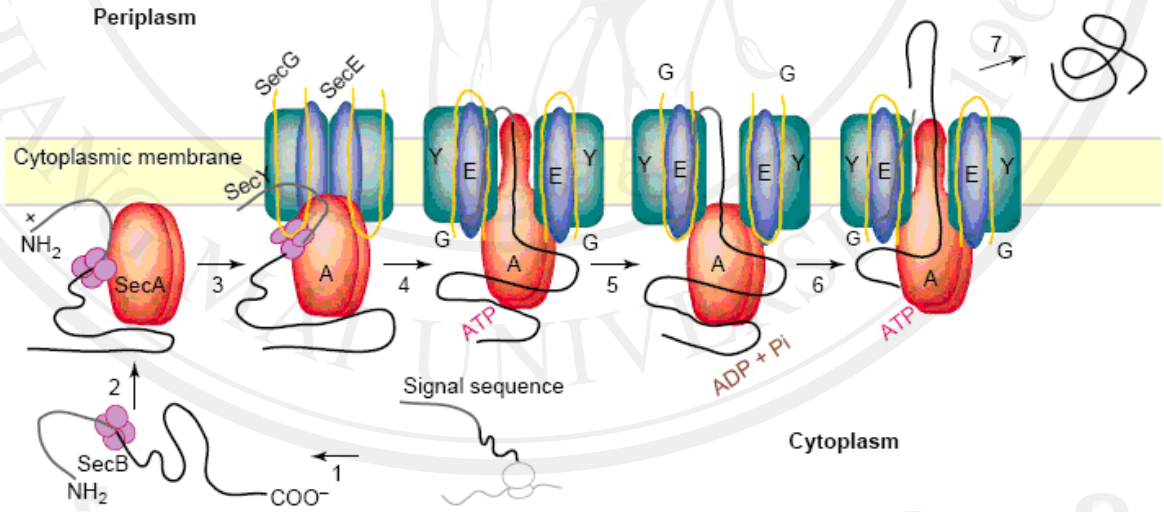

Figure 1.7 The general Sec pathway in bacteria. (Mori and Ito, 2001)

1.2.4.2 Recombinant protein expression in insect cell *via* baculovirus expression system

The Baculovirus Expression Vector System (BEVS) has been widely used in research and scientific industrial communities for the production of high levels (up to 1000 mg/mL) of properly post-translationally modified (folding, disulfide bond formation, oligomerization, glycosylation, acylation, and proteolytic cleavage), biologically active and functional recombinant proteins. The concept of this system is based on the introduction of a foreign gene into a plasmid-based transfer vector, which is then incorporated into the genome of the baculovirus *via* homologous recombination resulting in recombinant baculovirus encoding heterologous protein which can be expressed in cultured insect cells and insect larvae (Huynh and Zieler, 1999). Currently, the viral vectors used in this system have mainly been isolated and constructed form *Autographa californica* nuclear polyhedrosis virus (*Ac*NPV) or *Bombyx mori* nuclear polyhedrosis (*Bm*NPV). *Ac*NPV can infect dozens of moth species and all *Spodoptera frugiperda* ovarian cell lines. In contrast, only cells of silkwarm origin are susceptible to *Bm*NPV. Nearly a thousand of high-value foreign proteins have been successfully produced using the baculovirus system. More recently, the insect baculovirus was applied to production of vaccines (Kang et al., 1997), gene therapy (Ghosh et al., 2002) and recombinant baculovirus insecticides (Bonning, Possee, and Hammock, 1999).

However, this system is inferior to prokaryotic and yeast systems in terms of its capacity for continuous fermentation and have some limitations, for

examples, the host cells which are infected by nuclear polyhedrosis virus will eventually die, so the heterologous gene cannot be expressed continuously. Every round of protein production requires the infection of new insect cells. Moreover, insect cells and mammalian cells differ in their glycosylation patterns, such as in the lengths of oliogsaccharides and in mannose content, so bioactivity and immunogenicity of insect expression products are somewhat different from those of the natural product (Kulakosky, Shuler, and Wood, 1998). However, researchers have successfully addressed this limitation by transforming established insect cell lines with constitutively expressible mammalian genes. This approach has yielded the transgenic insect cell lines with normal growth properties that can support viral infection, have new *N*-glycan processing enzyme activities, and can produce humanized recombinant glycoproteins (Donald L, 2003)

1.2.5 Phage display technology

1.2.5.1 Introduction

Phage display was first developed with the *E. coli* specific bacteriophage M13 (Volkman and Goldsmith, 1985), and the success of M13 phage display has prompted the development of numerous alternative display systems (Ren and Black 1998; Santini, Brennan et al. 1998). Phage display is a powerful method for selecting and engineering polypeptides with desired binding specificities. This technology was applied to the field of immunology, cell biology and pharmaceutical biotechnology (Sidhu, 2000). Displaying peptides and gene fragments enables the analysis of protein-protein interactions such as structural mapping of epitopes (Tayapiwatana,

Arooncharus et al. 2003; Abbasova, Shcheprova Zh et al. 2007), characterization of receptor and ligand interaction (Jager, Jahnke et al. 2007; Mohrluder, Stangler et al. 2007; Casey, Coley et al. 2008), functional analysis (Intasai, Mai et al. 2006; Yang, Meng et al. 2006) and immunodiagnosis (Robles, Gonzalez et al. 2005; Hell, Amim et al. 2009). This method is accomplished by inserting the gene fragments encoding the protein of interest into a phagemid genome as a fusion with M13 coat protein genes. These fusion genes can be incorporated in bacteriophage particles that also display the heterologous proteins on their surfaces. In this way, a physical linkage is established between phenotype and genotype of the expressed protein.

1.2.5.2 Biology and structure of M13 filamentous bacteriophage

The filamentous phage constitutes a large family of bacterial viruses that infect a variety of Gram-negative bacteria, using pili as receptors. The best characterized are the very similar phages M13, fd and f1, that infect *E. coli via* F pili. The filamentous bacteriophages contain a circular single-stranded DNA genome encased in a long protein capsid cylinder. The Ff phage particles are flexible rods about 1 μm long, less than 10 nm in diameter, and mass of 16.3 MD, of which 87% is contributed by protein. The particle consists of a single-stranded DNA core surrounded by a proteinaceous coat. The length of the cylinder consists of approximately 2,700 molecules of 50 amino acid major coat proteins, called gene VIII protein (gpVIII). The four minor coat proteins are present at about 5 copies per particle; protein-VII and protein-IX (gpVII and gpIX) cap one end of the particle while protein-III and protein-VI (gpIII and gpVI) cap the other end (**Figure 1.**8 and

Table 1.1). Each of the five different coat proteins has been successfully used as a platform for the functional display of heterologous polypeptides as either N- or C-terminal fusions. Protein VI and gpIII are crucial for host recognition and phage infectivity, whereas gpVII and gpIX are required for phage assembly (Gailus et al., 1994). The gpIII is the most commonly used coat protein for display. This protein is made up of three domains separated by glycine-rich regions. These three domains have been designated N1 or D1, N2 or D2 and CT or D3 by different groups. The first domain, N1, is required during infection for the translocation of the DNA into the cytoplasm and insertion of the coat proteins into membrane. N2 is responsible for binding to F pilus (Deng, Malik, and Perham, 1999). The carboxy-terminal end makes up the CT domain and is essential for forming a stable phage particle. The DNA is oriented within the virion such that a 78 nucleotide hairpin region called the packaging signal (PS) is always located at the end of the particle containing the gpVII and gpIX proteins.

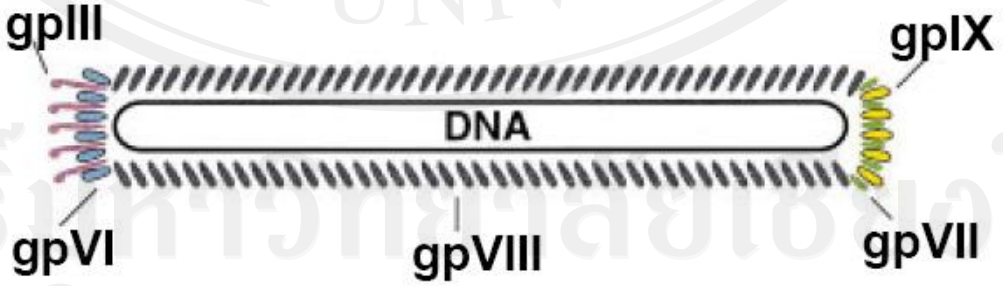

Figure 1.8 Structure of a filamentous bacteriophage. A diagram of the bacteriophage particle represents the single-stranded DNA core surrounded by a proteinaceous coat (Sidhu, 2001).

Table 1.1 Phage coat proteins.

Protein	Number of amino acids	Molecular weight	Copies per phage	Function
gpIII	**406**	**42,500**	**~5**	**Minor capsid protein**
gpVI	**112**	**12,300**	**~5**	**Minor capsid protein**
gpVII	**33**	**3,600**	**~5**	**Minor capsid protein**
gpVIII	**50**	**5,200**	**~2,700**	**Major capsid protein**
gpIX	**32**	**3,600**	**~5**	**Minor capsid protein**

1.2.5.3 M13 filamentous bacteriophage life cycle

Infection of *E. coli* by a filamentous phage is a multistep process requiring interactions of the phage gpIII protein with the F conjugative pilus and the bacterial TolQ, R and A cytoplasmic membrane proteins. These three Tol proteins are bacterial proteins that appear necessary for maintaining the integrity of the bacterial outer membrane, and avoid a leak of periplasmic proteins into the culture medium (Lazzaroni et al., 1999). The TolQRA proteins are required during phage infection for translocation of the Ff phage DNA into the cytoplasm and translocation of the phage coat proteins into the cytoplasmic membrane (Russel, Whirlow et al. 1988; Russel, Linderoth et al. 1997). Infection is initiated by attachment of the N2 domain of gpIII

to the tip of the pilus; this is the end of the particle that enters the cell first (Holliger and Riechmann 1997; Witty, Sanz et al. 2002; Karlsson, Borrebaeck et al. 2003). This binding releases N1 from N2, and allows N1 to interact with TolA. As the process continues, the coat proteins dissolve into the periplasm. The major coat protein, gpVIII, gpVII and gpIX minor capsid proteins, and the uncoated ssDNA concomitantly enters the cytoplasm (**Figure 1.9).** Once the viral (+) strand DNA enters the cytoplasm, the complementary DNA (-) strand is synthesized by bacterial enzymes, resulting in a covalently closed, supercoiled and double-stranded replicative form (RF). The (-) strand of this RF is a template for transcription and resulting mRNAs are translated into all of phage proteins. Of the 11 phage-coated proteins, three (gpII, gpX, gpV) are required to generate ssDNA, three (gpI, gpXI, gpIV) are required for phage assembly and five (gpIII, gpVI, gpVII, gpVIII, gpIX) are components of the phage particle. After one round, gpII circularizes the displace viral (+) strand DNA, which then is converted to a covalently closed, supercoiled and double-stranded RF molecule by bacterial enzyme. The gpV dimers bind cooperatively to newly generated (+) strand RF and prevent its conversion to RF DNA. The RF DNA synthesis continues until the amount of gpV reaches a critical concentration. The gpX is crucial for the proper replication of the phage DNA and functions as an inhibitor of gpII (Fulford and Model, 1984). The RF replicates to make progeny RFs and is also the template for transcription of phage genes and synthesis of progeny ssDNAs. Assembly occurs at specific sites in the bacterial envelope where the cytoplasmic and outer membranes are in close contact by the interaction of gpI, gpVI and gpXI and form a gated pore complex that spans the inner

and outer membranes. Phage assembly is initiated by the incorporation of gpVII and gpIX at one end of the particle. This process continues until the end of the DNA is reached and the assembly is terminated by the incorporation of gpVI and gpIII (**Figure 1.10**). Progeny virions are secreted continuously without lysis of the *E.coli* host; chronically infected cells continue to divide, though at a slower rate than uninfected cells.

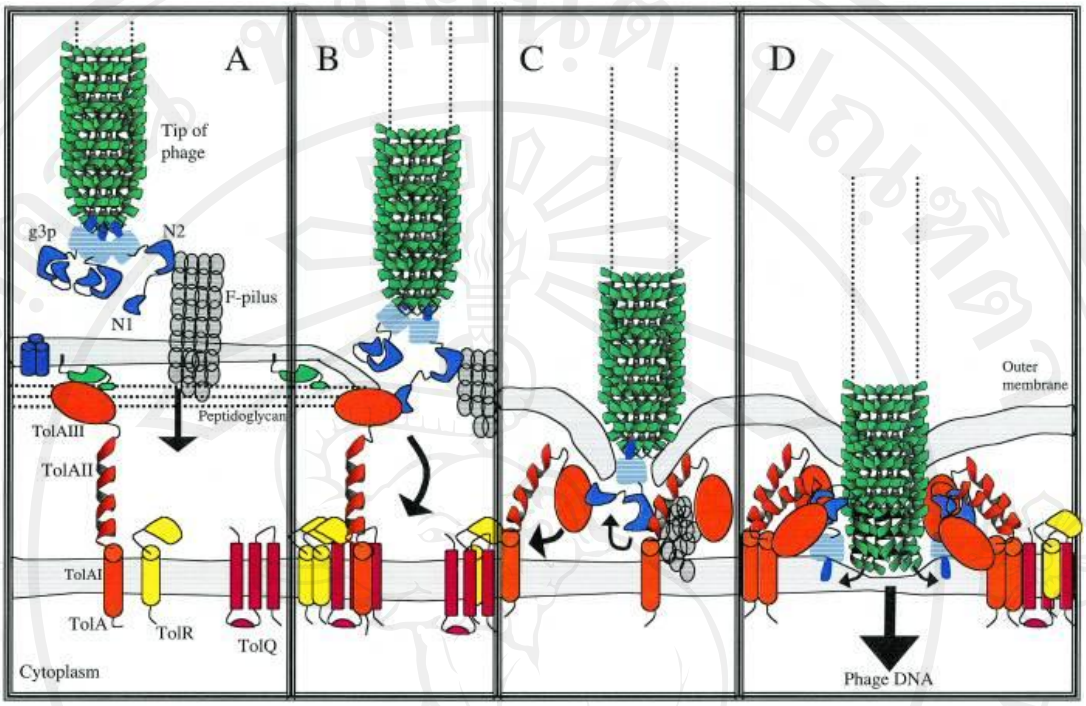

Figure 1.9 Model of filamentous phage infection. A) The phage N2 domain of pIII protein interacts with the F-pilus on the outside of the bacteria. The outer membrane proteins OmpF (blue cylinders) and Pal lipoprotein (greenish) are also included, as there have been reports of TolA interacting with these proteins prior to infection (Lazdunski, Bouveret et al. 1998; Cascales, Gavioli et al. 2000). **B)** After F-pilus retraction, N1 domain of gpIII binds to the C-terminal domain of bacterial TolA domain III (TolAIII). **C)** The retracting pilus brings phage gpIII domains in closer contact with TolA domains. As a consequence, TolA can assume a more compact state of assembly, thus bringing the outer and inner membranes of the bacteria closer together. At this stage, the central domain of TolA (TolAII) has the possibility to interact with the N2 domain of gpIII. **D)** The phage gpIII is inserted into the inner membrane, and the cap of the phage head is opened to allow phage DNA to enter the bacteria (Karlsson et al., 2003).

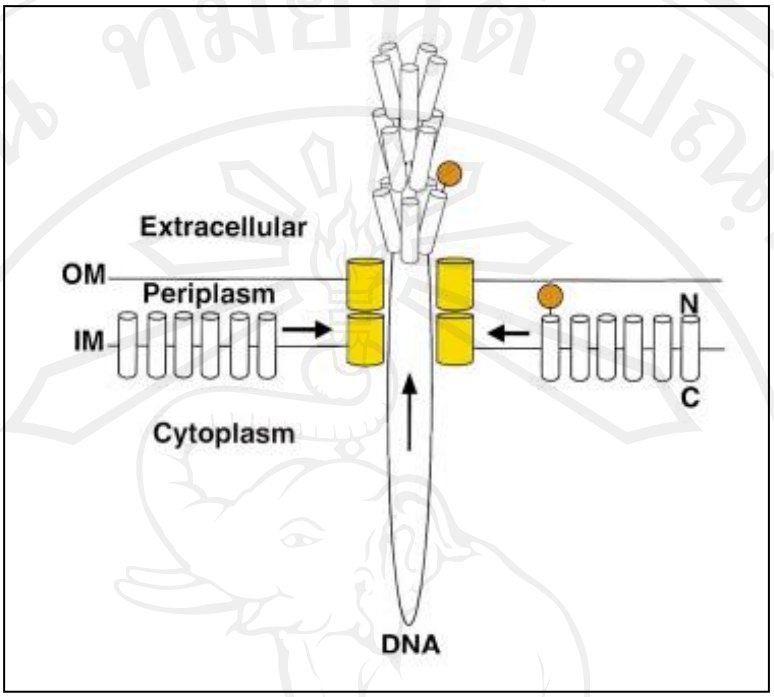

Figure 1.10 M13 bacteriophage assembly. Newly-synthesized coat proteins (white cylinders) are embedded into the inner membrane (IM) with their N termini in the periplasm and their C termini in the cytoplasm. Single-stranded viral DNA is extruded through a pore complex (yellow cylinders) that spans the inner membrane and the outer membrane (OM). Coat proteins also interact with the pore complex, where they surround the DNA and are thus transferred from the bacterial membrane into the assembling phage coat. The assembled phage particle is extruded to the extracellular environment. A heterologous protein (red circle) will be displayed on the phage surface if it is fused to a coat protein that can successfully incorporate into the phage coat (Sidhu, 2001)

1.2.5.4 Filamentous phage-based vector system

Protein can be displayed using vectors based on the natural filamentous phage sequence (phage vector) or using plasmids that contain only the fusion phage gene and no other phage genes (phagemid vectors) (Lowman, 1997). Phagemids, a more popular vector for display, are designed to contain the origins of replications for both the M13 phage and *E. coli*, appropriate multiple cloning sites for insertion of the gene of interest, and an antibiotic-resistance gene for selection and propagation as with typical phagemids (Barbas, 2001). Phagemids can be grown as plasmids or alternatively packaged as recombinant M13 phage with the aid of a helper phage that contains a slightly defective origin of replication (such as M13KO7 or VCSM13) and supplies, *in trans*, all the structural proteins and enzymes required for generating a complete phage. They also carry a kanamycin resistance gene to allow antibiotic selection of helper-infected cells. Thus the phage particles may incorporate either gpIII derived from the helper phage or the polypeptide-gpIII fusion protein, encoded by the phagemid. A major advantage of phagemid vectors is their small size and ease of cloning, as compared to the difficulties of cloning in phage vector without disruption of the structure of gene and promoter. In addition, the large DNA inserts are more readily maintained by phagemid genomes than phage genomes. Two-gene display systems (Type 3+3 and 8+8 phagemid and Type 33 and 88 phage systems) allow modulation of the valency (i.e., the number of copies per phage particle, **Table 1.2**) of the displayed fusion protein. However, the ratios of polypeptide-gpIII fusion protein: wild type gpIII may range between 1:9 and 1:1000 depending on the type of

phagemid, growth conditions, the nature of the polypeptide fused to gpIII, and proteolytic cleavage of antibody-gpIII fusions (Azzazy and Highsmith, 2002).

Table 1.2 Classification of phage-display vectors.

Vector type	Coat protein used for display	Display on all or some copies of coat protein	Number of coat protein genes	Fusion encoded on genome of
3	**pIII**	**All**	**1**	**phage**
8	**pVIII**	**All**	**1**	**phage**
33	**pIII**	**Some**	**2**	**phage**
88	**pVIII**	**Some**	**2**	**phage**
3+3	**pIII**	**Some**	**2**	**phagemid**
8+8	**pVIII**	**Some**	**2**	**phagemid**

1.2.5.5 Applications of phage display

Since the phage display just report of 1985 (Volkman and Goldsmith, 1985), this technology has rapidly evolved into an efficient tool used by structural biologists for the discovery and characterization of diverse ligand-receptor-binding interactions. By this technique, the gene of interest is inserted between the C-terminal of the signal peptide and N-terminal coding region of phage coat protein. The recombinant proteins are synthesized in *E. coli* host together with other coat and accessory proteins of filamentous phage, and incorporated into the phage during the assembly process. The released phage particles expose the recombinant protein as a fusion product to one of the phage coat protein. By inserting different DNA fragments, a library of phage particles bearing different recombinant coat protein can be generated. Each phage particle contains only one type of recombinant coat proteins encoded by the corresponding gene fusion present inside the same phage particle. Individual phages can be rescued from libraries by an interaction of the displayed protein with the cognate ligand by a panning step, which allows the selection hundreds of millions of clones, those few phages that display a peptide that binds the target molecule. These phages can be amplified by infection of bacteria. The recombinant polypeptides or proteins displayed on the phage surface can be used to identify and characterize the interaction with their binding targets (Cesareni, 1992).

Phage display of a functional protein has now become a standard first step of proof of concept prior to the application of combinatorial strategies using the cloned DNA template to evaluate or remodel functional activity (Bass, Greene et al. 1990; Roberts, Markland et al. 1992). Even though a natural functional domain can

represent the end product of a highly directed evolutionary process, phage-display approaches can create variations of the domain with altered binding affinity or fine specificity, or with structural refinements that greatly enhance stability (Lowman, Bass et al. 1991; Lowman and Wells 1993; Hao, Serohijos et al. 2008; Hoyer, Gronwall et al. 2008; Kwong, Baskar et al. 2008; Berntzen, Andersen et al. 2009; Hertveldt, Belien et al. 2009).

Advances in phage display and antibody engineering have led to the development of phage-displayed antibody technology (McCafferty, Griffiths et al. 1990; Sidhu 2000). This technology allows one to isolate antibodies directly from diverse repertoires of antibody genes, generating high affinity binding sites and unique specificity (McCafferty, Griffiths et al. 1990; Winter, Griffiths et al. 1994; Neri, Petrul et al. 1995; Hoogenboom, de Bruine et al. 1998). Phage antibody genes can be easily sequenced, mutated to improve antigen binding, and screened. The advantage of this technology can also be amplified immunoglobulin variable (V) genes from hybridomas or B lymphocytes using polymerase chain reaction (PCR) technology, cloned into phagemid vectors and rescued the monoclonal antibodies from genetically unstable hybridomas. Finally, soluble recombinant antibodies (not displayed on phage) can be produced rapidly and economically and can be used as *in vitro* diagnostic reagents. Various formats of antigen binding fragments, including Fab and scFv have been cloned and displayed on the surface of M13 viral particles with no apparent loss of the antibody's specificity and affinity as shown in **Figure 1.11** (Nissim, Hoogenboom et al. 1994; Azzazy and Highsmith 2002). These

antibodies have become important tools in several fields, including molecular biology (Kwong, Baskar et al. 2008; Kato-Takagaki, Mizukoshi et al. 2009), pharmaceutical and medical research (Sidhu, 2000) as well as in the treatment of diseases such as cancer (Schrama, Reisfeld, and Becker, 2006) and infectious diseases (Mullen et al., 2006).

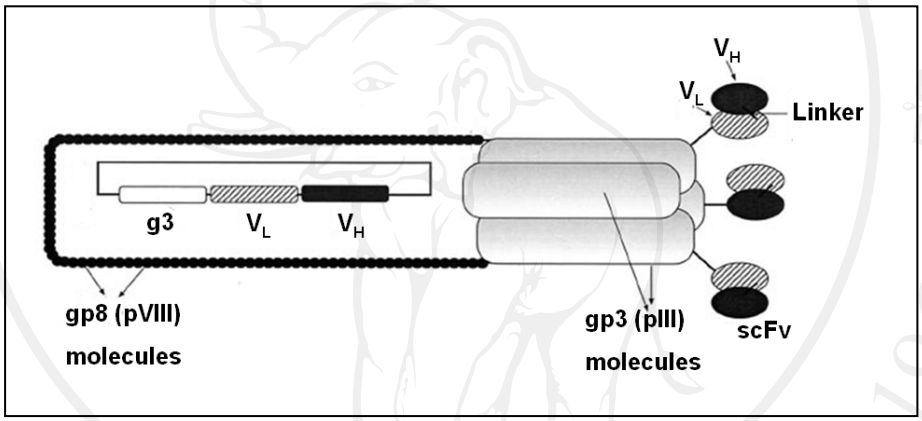

Figure 1.11 Schematic diagram of a filamentous phage displaying single chain variable fragment (scFv) molecules. The phage consists of circular ssDNA surrounded by a coat protein. The genes encoding the variable domains of the scFv and a linker are fused to gene III (g3) in the genome of the filamentous phage. Consequently, the scFv is displayed as a fusion to gp3 (gpIII) protein at the tip of the phage (Azzazy and Highsmith, 2002).

1.2.6 Computational approaches and structural-based protein analysis

Over the last several decades, X-ray crystallography and molecular modeling have highlighted many significant details of the mechanisms of antibody-antigen interactions (Lin, 2008). The antibody-antigen complex structure derived from computational study can aid in understanding the relative potency or binding affinity of antibodies (Kaufmann et al., 2002; Malik et al., 2010).

1.2.6.1.1 Homology modeling

Since the functional specificity and biological function of a protein are linked to its structure, analysis of the three-dimensional (3D) structure of the protein will reveal more information than its sequence. Homology modeling has used computational approaches to generate, model, and predict the 3D structure. The protein structure is determined using a protein that shares a high sequence similarity more than 40% identical residue with the sequence of another (homologous) protein of known structure (Gibney and Baxevanis, 2001). Then the potential structure of the unknown protein can be generated by using homology modeling techniques with high a degree of confidence. As shown in **Figure 1.12**, a protein sequence of unknown structure is initially aligned with sequences of known protein template structures. Then, the alignment is used to find equivalent residues in the sequences. By using relevant information from known structures, the new structure can be predicted and modeled (Schwede et al., 2003; Xu, 2000).

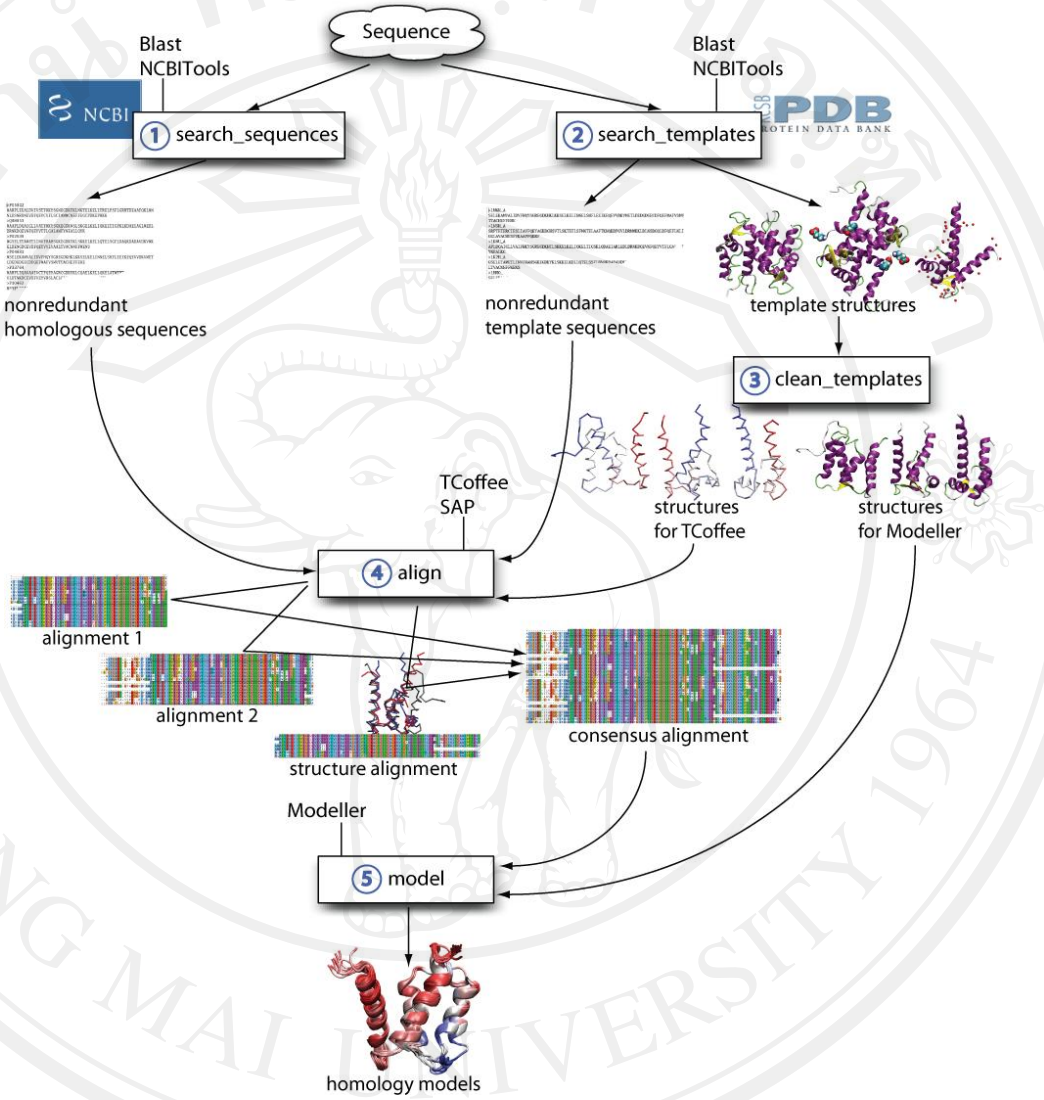

Figure 1.12 Schematic procedure of homology modeling. Initially, a protein sequence is searched for homologous sequences and 3D structures with NCBI tool, Blast. The match sequences from multiple alignments are selected and the structure of unknown protein is modeled. (http://biskit.pasteur.fr/use/workflows/homology-modelling)

1.2.6.2 Molecular docking

Computational generation of protein structure modeling *via* homology modeling and docking of a protein structure with potential interacting partners are two related steps in computational proteomics. Molecular docking is a term used for computational schemes that attempt to find the “best” matching between two model molecules which are a protein and a ligand (Lengauer and Rarey, 1996). The docking strategies are divided into two methods, rigid docking (both protein and ligand are rigid) and flexible docking (flexible ligand and rigid receptor). The bound conformation and interaction energy between a protein and a ligand can be predicted accurately **(Figure 1.13A).** The potential of mean force score or PMF score derive statistical preferences as potentials for protein-ligand atom pair interaction indicating binding affinity of the complex (Muegge, 2005). The high negative PMF score represents the high binding interaction **(Figure 1.13B).** Recently, molecular docking has been widely used for determining the binding affinity of antibodies, scFv, with their target antigens and providing as an effective strategy to create and select the potential antibodies (Barderas et al., 2008; Gu et al., 2010).

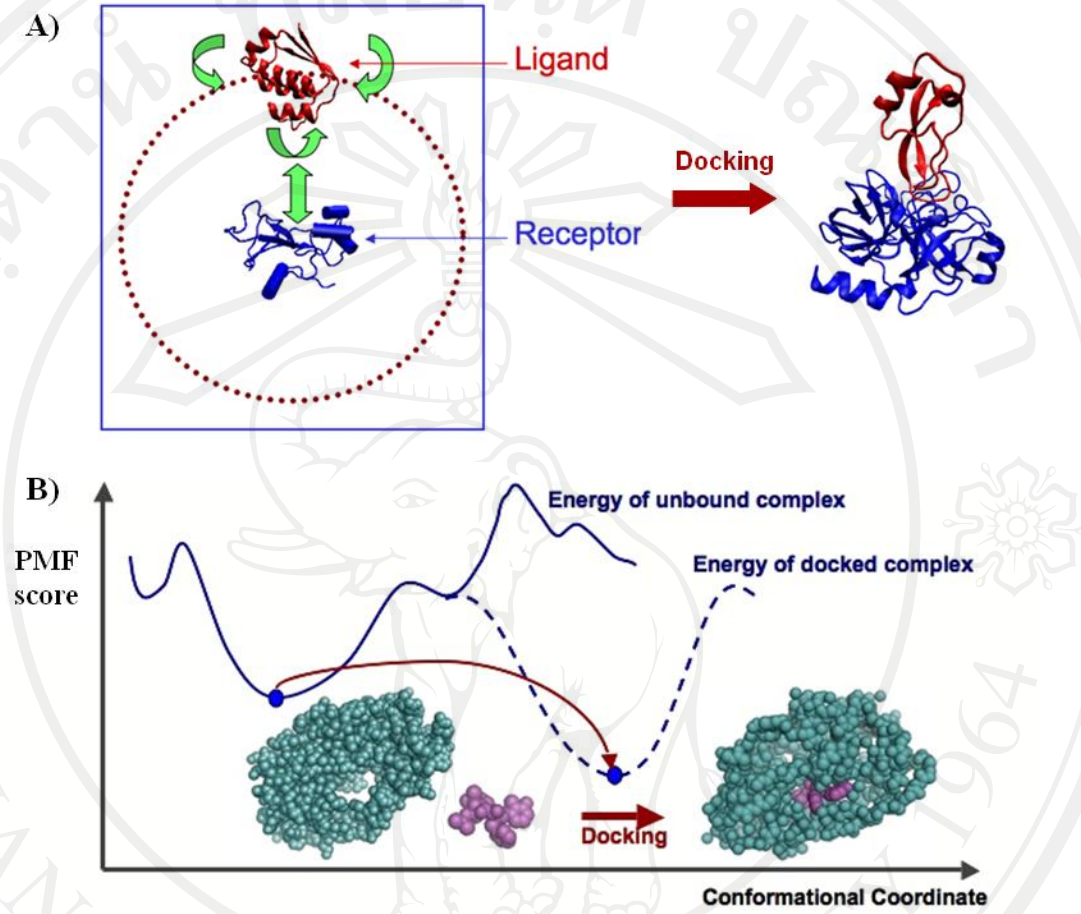

Figure 1.13 Schematic procedures of Molecular docking. **(A)** The flexible docking strategy. The structure of receptor or protein is fixed at the center of the sphere while ligand moves all around the ideal sphere indicated by red dots. The best conformational interaction is shown. Modified form http://www.ihes.fr/~carbone/HCMDproject.html. **(B)** PMF score of protein-ligand complex. The PMF score presents reverse correlation with the interaction. Modified from http://me.engin.umich.edu/research/areas/design.

1.2.6.3 Molecular dynamic simulation (MDs) and binding free energy calculation

Computer simulations have carried the hope of understanding the structural properties of protein-ligand complex molecule interm of their structure and the microscopic interactions. The molecular dynamics simulation (MDs) is the main simulation technique used in the present. MDs and binding free energy calculation have been performed to investigate the binding interaction of protein and ligand. These computational approaches have been applied to identify the binding activity of antibodies or scFvs with their target antigens. Moreover, MDs and binding free energy allow understanding insight the key interacting amino acid of antibodies. These identified residues can be modified for improving the binding affinity (Hu, Zhang, and Chen, 2011; Wang and Duan, 2011).

1.2.7 *In vitro* display technology

Antibody display technology has become the major technology for creating mAbs for human gene therapy. The development of antibody display technology focus on selection platforms of recombinant antibody libraries, methods for selection and screening different types of libraries, antibody affinity and stability optimization strategies, the impact of library-based approaches on antibody humanization. Recently, phage, and ribosome- and yeast-display technologies have turned into mainstream antibody and protein engineering platforms. These technologies include phage display (Hoogenboom, de Bruïne et al. 1998; Knappik, Ge

et al. 2000; Hoet, Cohen et al. 2005), ribosome display and mRNA display (Boder and Wittrup 1997; Hanes, Jermutus et al. 1998) microbial cell display (Boder E T, 2000; Daugherty et al., 1999), microbead display (Sepp, Tawfik, and Griffiths, 2002), and protein-DNA linkage display (Odegrip, Coomber et al. 2004; Reiersen, Løbersli et al. 2005) being the most utilized as shown in **Figure 1.14.**

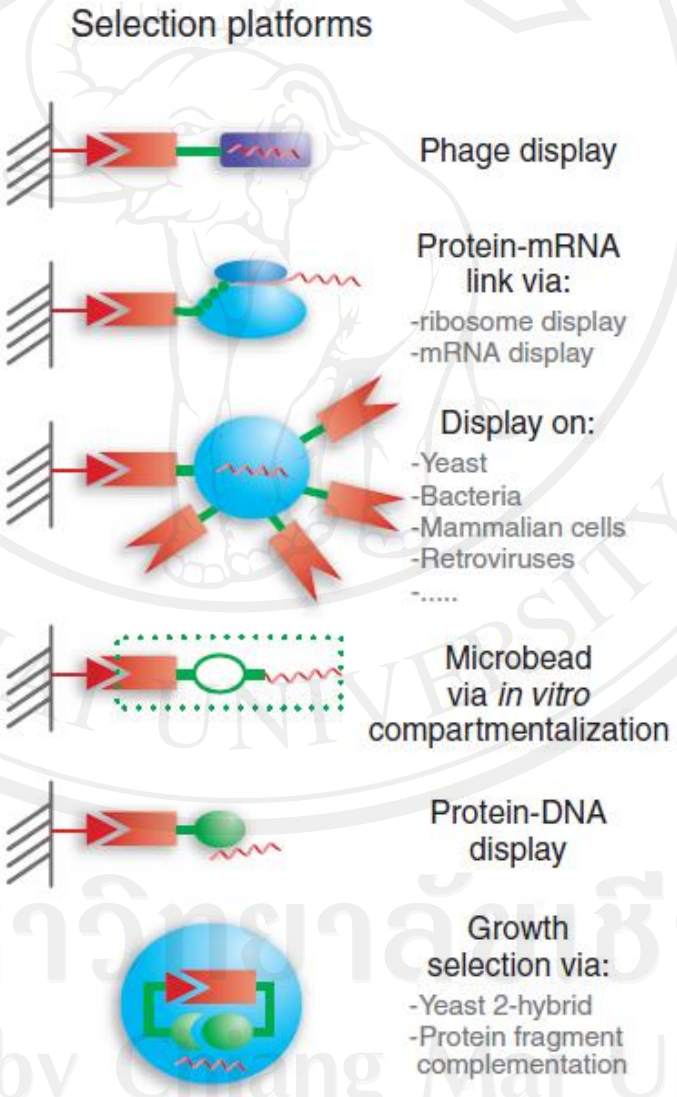

Figure 1.14 Different display technologies for selecting antibodies. (Hoet et al., 2005)

Hence, the applications of engineered antibodies have been used in human, the immunogenicity of antibodies have to be considered (Odegrip et al., 2004). It is accepted that the risk of immunogenicity may be reduced when using antibodies that are as "human" as possible. An efficient viral display technology, baculovirus, has become extremely useful for producing and isolating protein that require post-translation modifications as present only in eukaryotic cells (Boublik, Bonito et al. 1995; Grabherr R 1997; Mottershead, Alfthan et al. 2000; Grabherr, Ernst et al. 2001; Ojala, Mottershead et al. 2001). This eukaryotic display system enables presentation of large complex proteins on the surface of baculovirus particles becoming a versatile system in molecular biology (Smith, Summers et al. 1983; Jarvis, Kawar et al. 1998). Baculovirus surface display vectors have been engineered to contain mammalian promoter elements designed for gene delivery both *in vitro* and *in vivo*. Moreover, baculovirus capsid display has recently been developed; this holds promise for intracellular targeting of the viral capsid and subsequent cytosolic delivery of desired protein moieties such as intrabodies. (Shoji, Aizaki et al. 1997; Kost and Condreay 2002; Kenoutis, Efrose et al. 2006). Finally, the viruses can accommodate large insertions of foreign DNA and replicate only in insect cells (Ayres et al., 1994). Together, these are attributes making baculovirus display technology as important tools for expression and selection of functional antibodies for HIV gene therapy.

1.2.8 Baculovirus display technology

1.2.8.1 Baculovirus (BV)

Baculovirus is a family of insect virus, which is classified mainly into nuclearpolyhedrosisviruses (NPV) and granulosis viruses (GVs). NPVs have many virions within intranuclear crystal called polyhedra while GVs have only one virion inside each crystal or granule (Loy E, 1997). During infection, two forms of the virion are produced. Occluded virions (OVD) are produced and packed into the polyhedras inside the nucleus and are eventually released back to the environment as a consequence of cell lysis (Blissard 1996). In addition, some capsids are transported to cytoplasm and bud from the basal side of the midgut epithelial cells to the hemocoel. These budded viruses (BVs) are responsible for the secondary infection in the host and also infection of cell cultures (Blissard, 1996; Blissard, 1990) **(Figure 1.15).**

1.2.8.2 Biology and structure of budded baculovirus

Autographa californica multi-nucleopolyhedrovirus, (*Ac*MNPV), isolated from the alfalfa looper *Autographa californica*, is the baculovirus most commonly used in the baculovirus expression system (BVES) (Blissard, 1990). The budded form of the virus (BV) has a single rod-shaped nucleocapsid surrounded by a loose-fitting envelope **(Figure 1.15)** The viral DNA (134 kbp) is associated with the DNA binding protein in the rodshaped nucleocapsid. The structure of capsid and viral DNA of BV is similar to ODV, but the lipid and protein composition of the membrane differs, reflecting the distinct roles of the viral forms in the infection process

(Blissard, 1996). The membrane of BV, composed largely of phosphatidylserine, is derived from the plasma membrane of the host cell as the virus buds (Blissard, 1996; Braunagel and Summers, 1994).

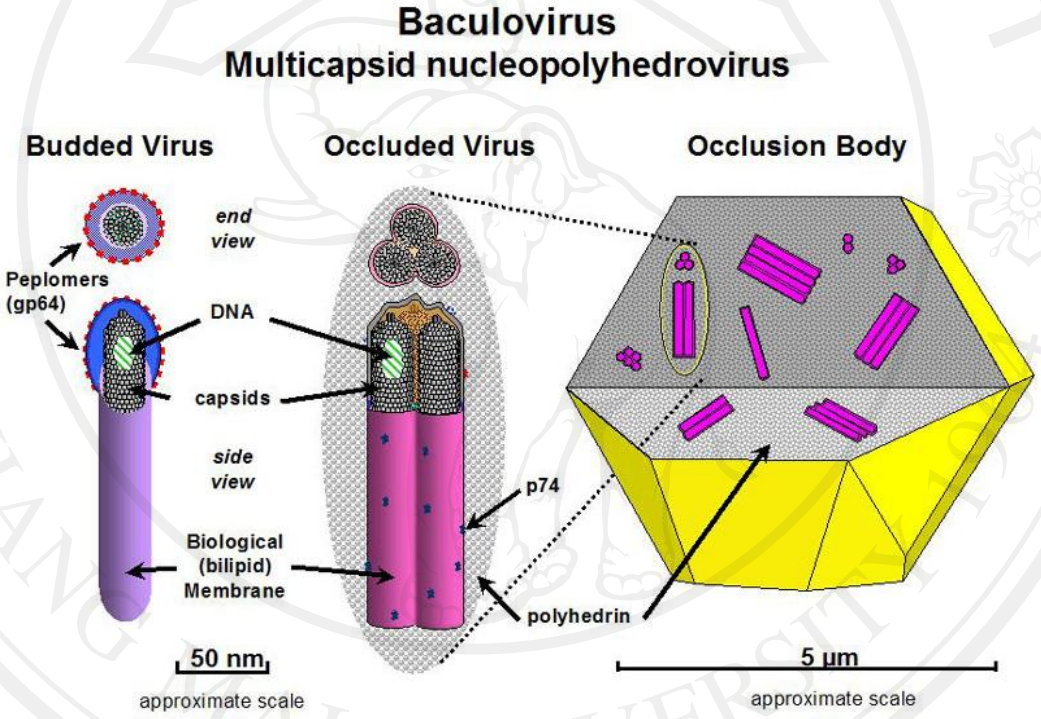

Figure 1.15 Diagram of nucleopolyhedrovirus virions (drawn by Dr. D. Lynn, USDA, Agricutural Research Service, US). Reproduced from *Wikipeidia, The Free Encyclopedia*, 2008.

1.2.8.3 Major membrane glycoprotein

A major envelope glycoprotein of *Ac*MNPV is generally known as gp64. Gp64 is a type I integral membrane glycoprotein containing an N-terminal signal sequence and a carboxy-proximal transmembrane (TM) domaincc(Blissard and Wenz, 1992). Moreover, gp64 contains four glycosylation sites at asparagine residues, oligomerization and fusion domain. This highly conserved glycoprotein forms spike-like peplomers that are mainly located at the end of the virion and are incorporated into the viral particles during the budding process (Volkman and Goldsmith, 1984). Gp64 is responsible for the host cell receptor binding activity (Hefferon et al., 1999) and is both essential and adequate for mediating pH-dependent membrane fusion during viral entry (Blissard and Wenz, 1992) through the membrane fusion domain (Monsma and Blissard, 1995). The gp64 protein is necessary for cell-to-cell transmission of the BV (Monsma, Oomens, and Blissard, 1996) and is important for efficient virion maturation and budding (Oomens and Blissard, 1999). In addition, structural studies of gp64 have provided the identification of distinct domains which are responsible for oligomer formation and membrane fusion (Duisit et al., 1999). These particular structural characteristics of gp64 are suitable to be a good candidate as a platform of a eukaryotic-based viral surface display system.

1.2.8.4 Baculovirus entry

Entry to insect cells

The *Spodoptera Frugiperda* or Sf9 is the commonly used insect cell in BVES. Previous studies showed that the gp64, a host cell receptor-binding protein, is necessary and sufficient for low pH triggered membrane fusion activity, the process that occur during virion entry by endocytosis pathway (Blissard and Wenz, 1992). Normally, envelope proteins from animal viruses mediate entry into host cells by facilitating two critical functions: binding to host cell receptors and fusion of the viral envelope with the host cell plasma membrane. Therefore, viral binding to a host cell receptor may serve two major functions. First, adhesion of the virus particle to the cell may be a necessary prerequisite for subsequent endocytosis. Second, the viral envelope and host plasma membrane must be closely apposed to fuse. The evidence supports the use of gp64 as receptor binding molecule of BV was demonstrated by after the elimination of N-glycosylation sites in gp64 diminution of the viral binding in insect cells (Jarvis et al., 1998). In addition, importance in the endocytosis was evaluated the blocking of the endocytosis by using neutralizing antibodies against gp64 (Volkman and Goldsmith, 1985). Taken together, gp64 is necessary of baculovirus entry into insect cell.

Entry to mammalian cells

In 19995, Hoffman et al. demonstrated that a recombinant baculovirus can efficiently infect human hepatocytes and deliver functional genes to the nucleus using an immediate early cytomegalovirus gene promoter (Hofmann et al., 1995).

Thus, uptake of baculoviruses by mammalian cells was considered a new phenomenon. Typically, baculovirus enters insect cells by endocytosis followed by low pH-faciliated fusion of the viral envelope, gp64 with endosomal membrane which allows viral entry into the cytoplasm (Blissard and Wenz, 1992). In the case of mammalian cells, it has been demonstrated that electrostatic interactions between haparan sulfate moieties on the cell membrane and gp64 on baculovirus may be prerequisite for baculovirus binding to cell surface of mammalian cells. Moreover, specific surface molecule of mammalian cells as docking point is involved in the viral attachment and viral entry into mammalian cells (Duisit, Saleun et al. 1999; Tani, Nishijima et al. 2001). Furthermore, Hoffman et al. brought attention to the possible hazards for humans in the unrestricted use of BEVS, but did emphasize that properly designed vectors should not be harmful. The development of the BacMam^{TM} System for gene delivery into mammalian cells has accomplished. There has been extremely use of this system (Hu et al., 2003; Kost, Condreay, and Jarvis, 2005) in efficient deliver genes into many cell types

1.2.8.5 Baculovirus applications

The development of insect pathogens as biological pesticides started in 1949. Registration of the first insect pathogen as a pesticide took place in 1975 and since then baculoviruses has become commonly used as pesticides. As a result of increasing resistance to chemical pesticides, new genetically engineered baculovirus pesticides have been tested in the field. The development of *Ac*MNPV expression system of a eukaryotic protein during the early 1980s marked the start of a wide

utilization of the virus to express heterologous proteins in insect cells (Oker-Blom, Airenne, and Grabherr, 2003). Major advantages of the baculovirus expression system include large insertion capacity, relatively easy construction, strong promoters and post-translational processing abilities (O'Reilly, 1994). In contrast to prokaryotic expression, insect cells are able to drive post-translational modifications e.g. proteolytic processing, phosphorylation, N- and O-glycosylation and acylation required for the biological activity of many eukaryotic proteins (O'Reilly, 1994) The baculovirus expression system has been used for a variety of applications. One of them is the production of virus-like particles (VLPs) that have been used to study viral assembly processes, production of novel vaccines and proteins for diagnostic assays and gene transfer (Kost, Condreay, and Jarvis, 2005).

Moreover, baculoviruses can be used as a gene carrier for gene therapy compared with animal viral vectors. Most important advantages are nonreplicative and nontoxic attributes of baculovirus. High species specificity, lack of replication in mammalian cells, and comparatively small cytotoxic effects provide baculovirus vector system with safety as gene vehicle (Mäkelä et al., 2006). In addition, baculoviruses can be produced easily and have a large cloning capacity compared to animal viral vector, along with the property to avoid the problem of pre-existing immunity (Yu - Chen, 2006). Therefore, baculovirus system is regarded as a fresh and attractive candidate for gene therapy application (Mäkelä et al., 2006).

1.2.8.6 Baculovirus display technology

For the purpose of displaying foreign proteins on the surface of baculovirus particles, as well as on infected insect cells, gp64 can serve as a fusion partner that, together with the chosen target protein, becomes incorporated into budded virions **(Figure 1.16A).** In the first report, foreign genes were cloned into a vector providing N-terminal fusion with the gp64 signal peptide and C-terminal fusion with the complete gp64. It was demonstrated that the glutathione-Stransferase-gp64 and gp120-gp64 fusion proteins were presented on the surface of baculovirus particles. The virally incorporated gp120 of human immunodeficiency virus (HIV)-1 had retained its biological activity as it bound to its ligand, CD4, on the surface of the recombinant virions (Boublik, Bonito, and Jones, 1995). These findings clearly indicated that large, complex proteins could be displayed on the surface of baculovirus particles in a functional form. The mechanism of incorporation into the viral particle was oligomerization of the fusion construct with wild-type gp64 (Boublik, Bonito, and Jones, 1995). Similarly, other proteins such as the HIV-1 envelope protein gp41 (Grabherr R, 1997), as well as the green fluorescent protein (GFP) and the envelope glycoproteins E1 and E2 of rubella virus were successfully displayed on the surface of baculovirus particles (Mottershead et al., 1997). Various other proteins have since been exposed on the surface of both baculovirus virions and infected insect cells using similar or identical strategies. Furthermore, enhanced specificity of baculovirus binding mammalian cells has been achieved through fusion

of synthetic immunoglobulin binding domains to gp64, or by utilising single-chain antibody fragments in a similar manner.

In contrast to expressing a second copy of gp64 fused to a target protein, peptides can be directly engineered into the native gp64 of *Ac*MNPV in order to increase the avidity of the displayed target **(Figure 1.16B)**. So far, only small peptides have been inserted into the native gp64, the maximum size being 23 amino acids. By comparing different positions within the gp64 sequence using specific antibody epitopes, it was found that the surface probability of the inserted peptide strongly depends on the position, structural framework and the adjacent amino acids. The position not only affects the viral growth characteristics, also the actual presentation of the epitope is influenced by the structure flanking the corresponding target peptide. The three-dimensional (3D) structure of gp64 would be of great value for further development of display strategies based on insertions into the native envelope protein of AcMNPV.

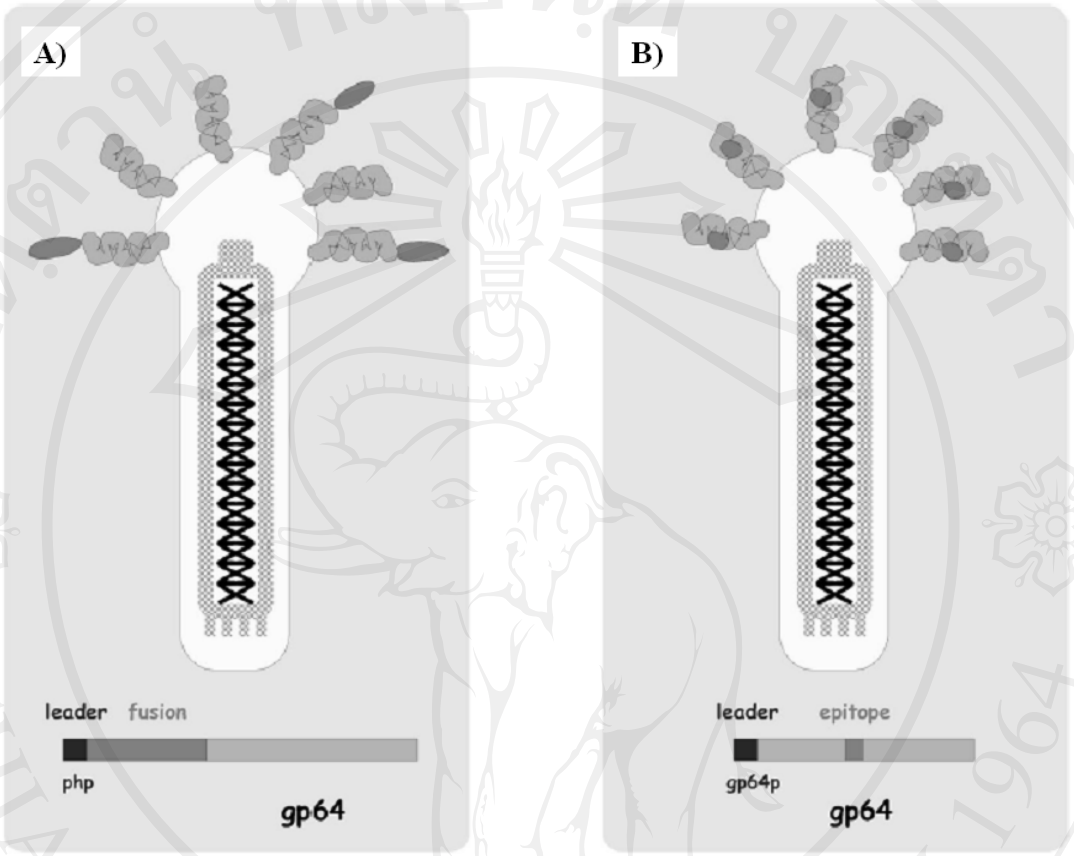

Figure 1.16 Schematic representations of the baculovirus display strategies. (A) gp64 as a fusion partner for baculovirus display. The protein of interest is fused to the N-terminus of the gp64 membrane protein or truncated versions thereof. **(B)** Peptide insertions into the native gp64 protein. The insertions are placed within the gp64 sequence and the modified protein is expressed under the gp64 promoter, where the native gp64 coding sequence has been omitted (Oker-Blom, Airenne, and Grabherr, 2003).

1.2.9 HIV-1 protease (HIV-PR)

The aspartyl protease encoded by HIV-1 play an essential role in viral maturation since newly synthesized Gag polyprotein requires the specific cleaved by HIV protease (HIV-PR) to rearrange the digested proteins and become a mature vision. Therefore, HIV-PR is considered an attractive target for treatment of AIDS (Kohl et al., 1988).

1.2.9.1 HIV-PR structure and mechanism

HIV-PR protein structure has been investigated using X-ray crystallography **(Figure 1.17)**. It exists as a homodimer containing 99 amino acids with only one active site which is C2-symetrix in free form. Each monomer contains an extended β-sheet region known as the flap. The active site of HIV-PR is a triad Asp25-Thr26-Gly27 located between the identical subunits (Kramer et al., 1986). The Asp25 residue from one subunit and Asp25´ residue from the other subunit act as the catalytic residues, thus the enzyme is defined as aspartic proteases (Navia et al., 1989; Pearl and Taylor, 1987). Additionally, HIV-PR has two molecular “flap” which move a distance of up to 7 Å when enzyme becomes associated with substrate (Torbeev et al., 2011).

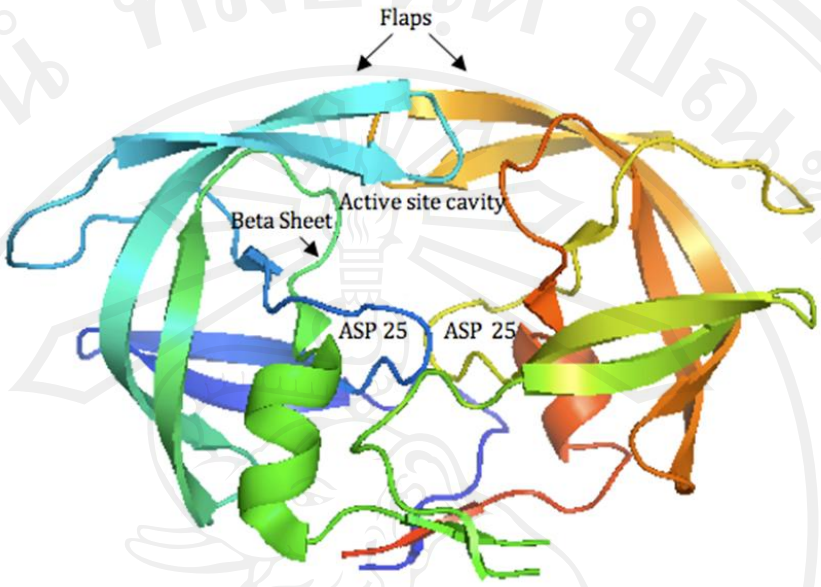

Figure 1.17 The crystal structure of the HIV-1 protease. The two flaps on the top are flexible structures that open to allow the entrance of polyprotein. The two aspartates in the active site present hydrolytic activity. (http://people.mbi.ucla.edu/yeates/153AH_2009_project/sriphanlop.html)

The active-site triad (Asp25-Thr26-Gly27) is located in a loop whose structure is stabilized by a network of hydrogen bonds. The hydrolysis mechanism of the Phe-Pro peptide bond specific to the HIV-1 PR consists of three elementary reactions **(Figure 1.18)**: first, hydration of the amide carbonyl group, with a water molecule accommodated between the two side-chains of the aspartic acid residues 25/25', gives a putative tetrahedral intermediate **(Figure 1.18A-B)**; second, the reaction of the protonation of the proline nitrogen of the substrate **(Figure 1.8C)**; and third, the reaction of the C-N bond cleavage of the substrate **(Figure 1.8D)** (Lee et al., 1996; Okimoto et al., 1999).

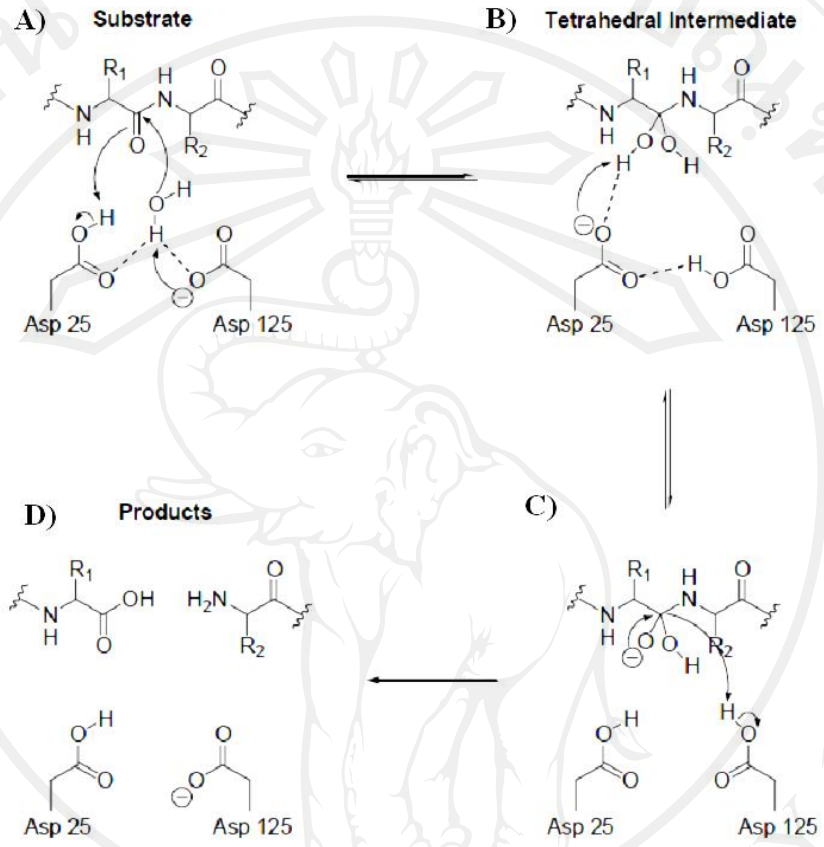

Figure 1.18 Schematic representation of the HIV protease cleavage mechanism. (Pokorná et al., 2009).

1.2.9.2 HIV-PR inhibitors

Since the combine use of HIV-PR inhibitors in HAART regimens have been introduced into the clinical practice. Such drugs inhibit specific steps in viral life cycle resulted in the abolishment of HIV replication. Nine protease inhibitors are currently on the market i.e. ritonavir (RTV), saquinavir (SQV), indinavir (IND), nelfinavir (NFV), lopinavir (LPV), amprenavir (AMV), fosamprenavir (FPV), atazanavir (ATV), tipranavir (TPV), and darunavir (DRV). They are competitive inhibitors of HIV-PR and all but one is peptidomimetics of the polyprotein cleavage sites. The first true nonpeptidic inhibitor is tipranavir, which was approved in 2005. The currently available PIs are also active against the HIV-2 protease, but with some exceptions – e.g., lopinavir is less effective in the therapy of HIV-2 infection. As mentioned above, important characteristics of any drug that are very relevant for its clinical application are toxicity and undesired side effects. The pathogenesis of some side effects has already been satisfactorily explained (e.g., nephrolithiasis occasionally caused by indinavir), but pathogenetic mechanisms of numerous side effects (e.g., lipodystrophy syndrome, insulin resistance, inhibition of glucose uptake or disturbances of bone metabolism) is still not entirely elucidated (Barbaro and Iacobellis, 2009; Duvivier et al., 2009; Mallewa et al., 2008).

The fist generation of HIV protease inhibitors are included several drugs. Saquinavir (SQV or Ro-31-8959), the first a HIV-PR inhibitor, was approved by FDA in 1995. This drug was designed based on unusual specificity of HIV-PR for cleavage of the Phe-Pro and Tyr-Pro bond. The scissile bond was replaced by a hydroxyethylamine isoester as shown in **Figure 1.19A** (Roberts et al., 1990).

Ritonavir (RTV, ABT-538) was then approved by FDA in 1996 and was constructed by concentrating on C-2 symmetric molecules of HIV-PR. The complexes formation of inhibitors and enzymes resulted in asymmetrical structure of HIV-PR which can not function **(Figure 1.19B)** (Kempf et al., 1994). Hence prolong-treatment of HAART using the first generation of HIV-PR inhibitors has led to the development of drug resistant. Therefore, the second generation inhibitors were designed to inhibit HIV-PR species resistant to the first generation. Lopinavir (LPV, ABT-378) is the second generation inhibitor most widely used in PI naïve patients. LPV was created to inhibit the resistant strains that contain the mutation V82A and was designed based on the structure of RTV. The P3 isopropylthiazolyl group of RTV that interacts with the Val82 residue of the wild-type HIV-PR was eliminated **(Figure 1.19C)** (Sham et al., 1998). RTV is currently used at lower dose as a pharmacokinetic boosting agent to increase the concentration of PIs in blood and prolong PIs therapeutic effects (Moyle and Back, 2001).

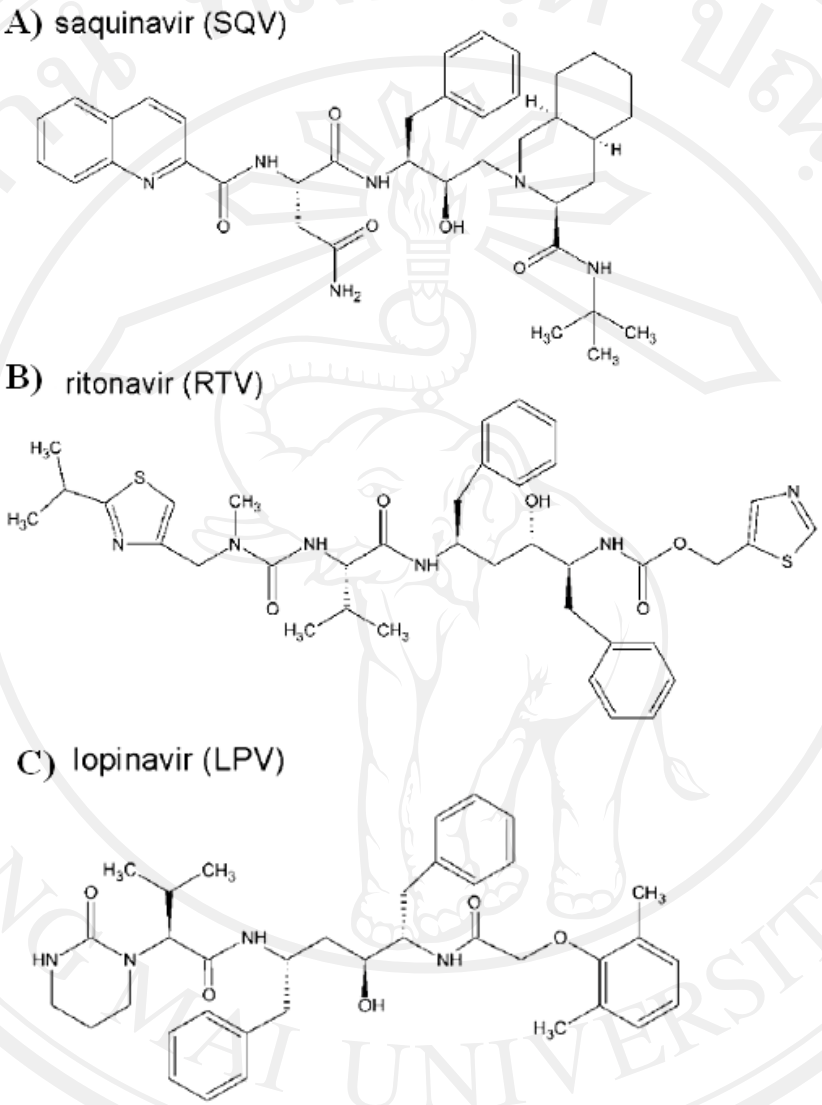

Figure 1.19 Chemical structures of HIVPR inhibitors. (A) saquinavir (SQV), **(B)** ritonavir (RTV), **(C)** lopinavir (LPV) (Pokorná et al., 2009).

1.2.9.3 HIV-PR resistant strains

The HIV-PR inhibitors are efficacious in controlling HIV replication. Problem of antiviral drug resistance can occur in patients who receive long-term treatment. The major cause of resistance is a high mutation rate due to the lack of proofreading activity of the viral reverse transcriptase, the dynamic viral replication, and the potential dual infection and insufficient effect of drugs (Kožíšek et al., 2008; Winters and Merigan, 2005). The pattern of mutations associated with viral resistant is extremely complicated as shown in **Table 1.3** (Johnson, 2008)**.** The mutations are appeared not only in the substrate binding region of protease, but also outside the active site of enzyme.

Table 1.3 Mutations in the protease gene associated with resistance to PIs

HIV-PR inhibitors	Major mutations	Minor mutations
Atazanavir +/- ritonavir	50, 84, 88	10, 16, 20, 24, 32, 33, 34, 36, 46, 48, 53, 54, 60, 62, 64, 71, 73, 82, 85, 90, 93
Darunavir	50, 54, 76, 84	11, 32, 33, 47, 74, 89
Fosamprenavir	50, 84	10, 32, 46, 47, 54, 73, 76, 82, 90
Indinavir	46, 82, 84	10, 20, 24, 32, 36, 54, 71, 73, 76, 77, 90
Lopinavir	32, 47, 82	10, 20, 24, 33, 46, 50, 53, 54, 63, 71, 73, 76, 84, 90
Nelfinavir	30, 90	10, 36, 46, 71, 77, 82, 84, 88
Saquinavir	48, 90	10, 24, 54, 62, 71, 73, 77, 82, 84
Tipranavir	33, 47, 58, 74, 82, 84	10, 13, 20, 35, 36, 43, 46, 54, 69, 83, 90

1.2.9.4 Drug discovery

The virtual screening approach for docking small molecules into a known protein structure is a powerful tool for drug design. Importantly, the use of virtual screening to discover new inhibitors is becoming a common practice in modern drug discovery (Shoichet, 2004). The design and screening of known compounds with ligands from database using computer-aided drug design has provided a high-throughput screening for early steps. Furthermore, the ligand candidates can be investigated *in vitro* and *in vivo*. Several studies have developed the optimized computation strategies for isolating a potential ligand. Moreover, the structure, binding activity, and active site between known protein and ligiand can be identified (Kitchen et al., 2004). Accordingly, virtual screening by computational approaches has been applied for creating and screening HIV-PR inhibitors (Chang et al., 2010; Sangma, 2005). Even though the virtual screening offers the high-throughput screening, isolation and investigation *in vitro* or *in vivo* also require a high-throughput platform to obtain potential HIV-PR inhibitors.

An HIV-PR activity assay has been developed to be used as a screening tool for drug discovery. A colorimetric assay of HIV-PR activity *in vitro* was generated based on the presence of digested product, 2,4,6-trinitrobenzenenesulfonic acid that is no longer available in crystalline form (Stebbins, 1997). On the other hand, selection of HIV-PR inhibitors *in vivo* was established based on a new cell line expressing HIV-PR in a chimeric protein with green fluorescence protein (GFP) and the function of HIV-PR is measured by

fluorescent signal from GFP (Fuse et al., 2006). However, this strategy is cost-ineffective and requires particular instruments that are not suitable as a high-throughput screening method. Therefore alternative simplified HIV-PR assay is prerequisite as high-throughput platform for isolating novel HIV-PR inhibitors from a large combinatorial library.

1.3 Objectives

1.3.1 To construct recombinant scFv anti-MA HB-8975 in bacterial expression system

1.3.3 To investigate the involved amino acids and binding activity of scFv anti-MA HB-8975 with its epitope and natural mutant epitopes using information from molecular modeling

1.3.4 To engineer scFv-MA HB-8975 and create the baculovirus display antibody technology

1.3.5 To produce the fusion protein of HIV-1 matrix and capsid (H_6MA-CA) protein substrate using baculovirus expression system

1.3.6 To establish ELIS-based HIV-PR activity assay using two specific anti-MA mAb and anti-CA mAb to investigate the reactivity of isolated scFvs against HIV protease substrate.