CHAPTER I

INTRODUCTION

1.1 Statement and significance of the problem

Nowadays, molecular engineering has contributed immensely to modify the structure of proteins. With the help of genetic engineering, it has been possible to modify the binding affinity, specificity, immunogenicity, pharmacokinetics and molecular size of antibody. It was interesting to reduce the size of antibody to use them inside cells, called intracellular antibodies or intrabodies. Generally, intrabodies are in the form of single-chain Fv antibodies (scFv), in which the heavy and light chain variable domains are fused by a linker peptide to create a single polypeptide. ScFv function by utilizing the antibody binding site coupled with a molecular tag to direct the antibody-antigen interaction to a specific cellular compartment, or by directly interfering with or neutralizing the function of a target protein.

The directing of scFv was occurred by use of signal sequence that allows the scFv to enter cellular compartments, result are blocking of intracellular transport or misdirect the localization of proteins, such as retaining proteins that were destined for cell surface expression in the endoplasmic reticulum (ER). ScFv can be targeted to the ER through the use of a leader sequence and an ER retention signal, such as the KDEL peptide, at the carboxyl termini. ER targeted scFv have been used to retain protein in the ER, thereby not allowing them to perform their normal function.

Several gene delivery systems are available; adenoviruses are attractive candidates for many gene delivery applications in medicine and science. Two

particular feature of adenovirus biology could be critical to successful use in gene therapy, vaccination and molecular genetic experiments. First, adenovirus infects both resting and dividing cells of many types. Second, highly purified virus can easily be produced. Additionally, high levels of transgene expression generally can be obtained. The most common vectors available are derived from human adenovirus type 5, because it is of low pathogenicity and oncogenicity.

The gene delivery systems that can express the scFv gene in each cellular compartment were required for its function at the target site. In this study, scFv expression vectors were constructed using adenovirus vector system to generate scFv in cytoplasmic compartment, ER and as a secreting molecule. The generation and binding activity of scFv in each compartment was investigated. These adenovirus expression systems will be used for many applications in the future.

1.2 Literature reviews

1.2.1 Adenovirus

1.2.1.1 Virus structure

Adenoviruses were first discovered half a century ago. These virus have been the object of intense study, mainly as a model system for basic eukaryotic cellular processes such as transcription, RNA processing, DNA replication, and translation (Rowe *et al*, 1953). Adenoviruses (Ad) are non-enveloped (naked) icosahedral viruses 70-90 nm in size with an outer protein shell surrounding an inner nucleoprotein core (Figure 1.1). The adenoviral capsid consists of 252 subunits called capsomeres, 240 hexon proteins and 12 penton bases. Beside these, the capsid also contains proteins pIIIa, pVI, pVIII and pIX (Majhen and Ambriovic-Ristov, 2006) (Table 1.1).

Table 1.1 Structural proteins of adenovirus

Capsid	
Hexon (II)	240 Trimers, major capsid protein
Penton base (III)	12 Pentamers, base components of 12 penton capsomers
Fiber (IV)	12 Trimers, spike components of 12 penton capsomers
IIIa	60 Monomers of hexon-associated protein
VI	Hexon-associated protein, 60 hexamers
VIII	Hexon-associated protein
IX	Hexon-associated protein, 80 hexamers
Core	
V	DNA-associated protein
VII	DNA-associated protein
Mu (X)	19-aa long DNA-associated protein
TP	Terminal protein

On each of 12 capsid vertices there is a penton base surrounded by 5 hexons. Penton base acts as an anchor for the fiber protein which bulges away from the virus. The fiber is a homotrimer consist of three identical polypeptides which are arranged in parallel orientation and contains three functionally and structurally different domains, i.e. an N-terminus that noncovalently binds fiber to penton base, a C-terminus which forms a knob aids in attachment to the host cell *via* the coxsackie-adenovirus receptor (CAR) on the surface of the host cell. Adenoviral particles have no membranes or lipids and are therefore stable in chemical or physical agents and adverse pH condition, allowing for prolonged survival outside of the body and water.

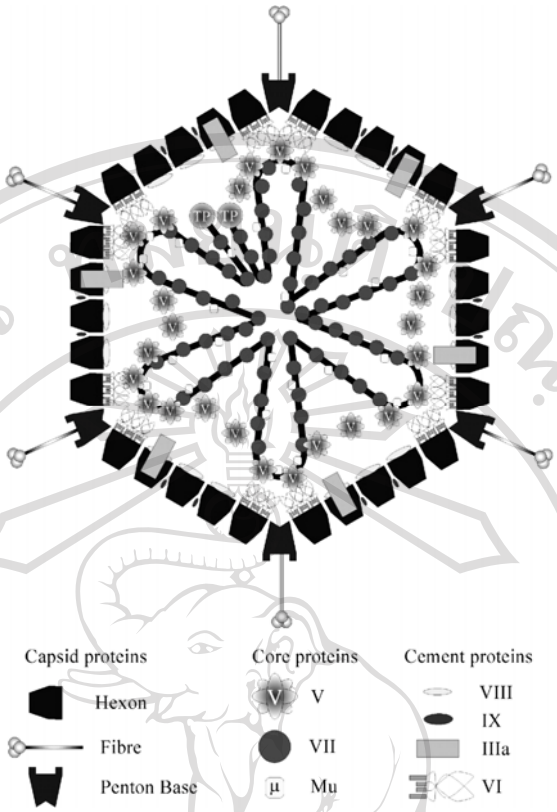

Figure 1.1 Structure of adenovirus. The locations of the capsid and cement components are reasonably well defined. In contrast, the disposition of the core components and the virus DNA is largely conjectural. (http://vir.sgmjournals.org/cgi/content/full/81/11/2573/F1)

There are 51 immunologically distinct human adenovirus serotype, further classified into 6 subgroup (A-F) (Table 1.2) that can cause human infections ranging from respiratory disease, conjunctivitis, pharyngoconjunctival fever and gastroenteritis, but do not show oncogenic potential in humans. Adenoviruses are primarily spread via respiratory droplets, however they can also be spread by fecal routes as well (Russell, 2000; Majhen and Ambriovic, 2006).

Table 1.2 Human adenovirus serotype

Group	Serotypes
A	12, 18, 31
B	3, 7, 11, 14, 16, 21, 34, 35, 50
C	1, 2, 5, 6
D	8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, 42-49, 51
E	4
F	40, 41

1.2.1.2 Binding and entry

The primary receptor responsible for attachment of all adenovirus serotypes except those from group B is CAR (Bergelson *et al*, 1997). CAR is a 46-kDa protein and a member of the immunoglobulin (Ig) superfamily that has two extracellular Ig-like domains, a single-pass transmembrane (TM) domain, and a cytoplasmic. The most distal extracellular Ig loop (D1 domain) of CAR facilitates a homophilic interaction that allows aggregation of cells expressing CAR. CAR localizes to the region of tight and adherence functions in polarized epithelia and expresses at variable levels on different tissue at distinct developmental stages (Seidman *et al*, 2001). Although, CAR's tissue distribution in human is not well define but mRNA is present in a number of organs, including the heart, brain, pancreas, and intestine (Tomko *et al,* 1997) as well as the lung, liver and kidney (Fechner *et al*, 1999). Recent research has shown that besides CAR adenovirus can use some other molecules as receptors as well. The best known are major histocompatibility complex I (MHCI) and heparan sulfate glycosaminoglycans.

Adenovirus entry into cells, as defined by experiments with cultured cells, generally involves attachment to a primary receptor, followed by interaction with a

secondary receptor responsible for internalization (Zhang and Bergelson, 2005). The virus enters the cell in a clathrin-coated vesicle and is transported to endosomes, where acidification results in partial disassembly of the capsid, the altered virion escapes into the cytoplasm and is transported to the nucleus, where replication occurs (Figure 1.2) (Meier and Greber, 2004). The route of intercellular trafficking is influenced by the fiber knob and, thus, by interaction with a specific primary receptor. Adenovirus 5, whose fiber binds to CAR escapes rapidly from endosomes into the cytoplasm. However, when it is pseudotyped with fibers that bind to another primary receptor, Adenovirus 5 is retained in the endosomal pathway (Miyazawa *et al*, 1999). After the initial interaction of the virus with the fiber receptor, entry of the virus is preceded. The critical recognition mechanism for this process is an RGD motif is exposed on the penton base (Stewart *et al*, 1997) and interacts with cellular integrins, members of a large family of heterodimeric (αß) adhesion receptors. The engagement of integrins ($\alpha_v\beta_3$, $\alpha_v\beta_5$ and $\alpha_v\beta_1$ integrins) by the penton base induces signals, including activation of PI3 kinase (Amalfitano *et al*, 1998), $p130^{CAS}$ (Li *et al*, 2000), and Rho GTPases (Li *et al*, 1998), that are important for rearrangements in the actin cytoskeleton and initiation of virus internalization. Interaction with integrins is also important for virus escape from the endosome (Wickham *et al*, 1994).

As noted above, progress of the virus through the endosomes and into the cell cytoplasm is normally mediated by clathrin and the coated pit pathway (Wang *et al*, 1998). Thereafter, the virus-encoded protease appears to assist in the further disruption of the virus capsid by the proteolysis of the structural protein VI (Greber *et al*, 1996), which functions as a linker between the capsid and the core components (Russell, 2000). The partially disrupted virus is then transported to the nuclear

membrane and the genome is passaged through the nuclear pore and into the nucleus, where the primary transcription events are initiated. The passage through the cytoplasm to the nucleus has been postulated to be mediated by the association of the virus core with p32 cellular protein. The p32 protein is primarily located in the mitochondria but can also be detected in the nucleus, and it has been suggested that it is a component of a cellular transport system that shuttles between the mitochondria and the nucleus and that the virus can hijack this system to gain access to the nucleus. This passage to the nucleus is relatively rapid and also involves the participation of dynein and microtubules (Leopold *et al*, 2000).

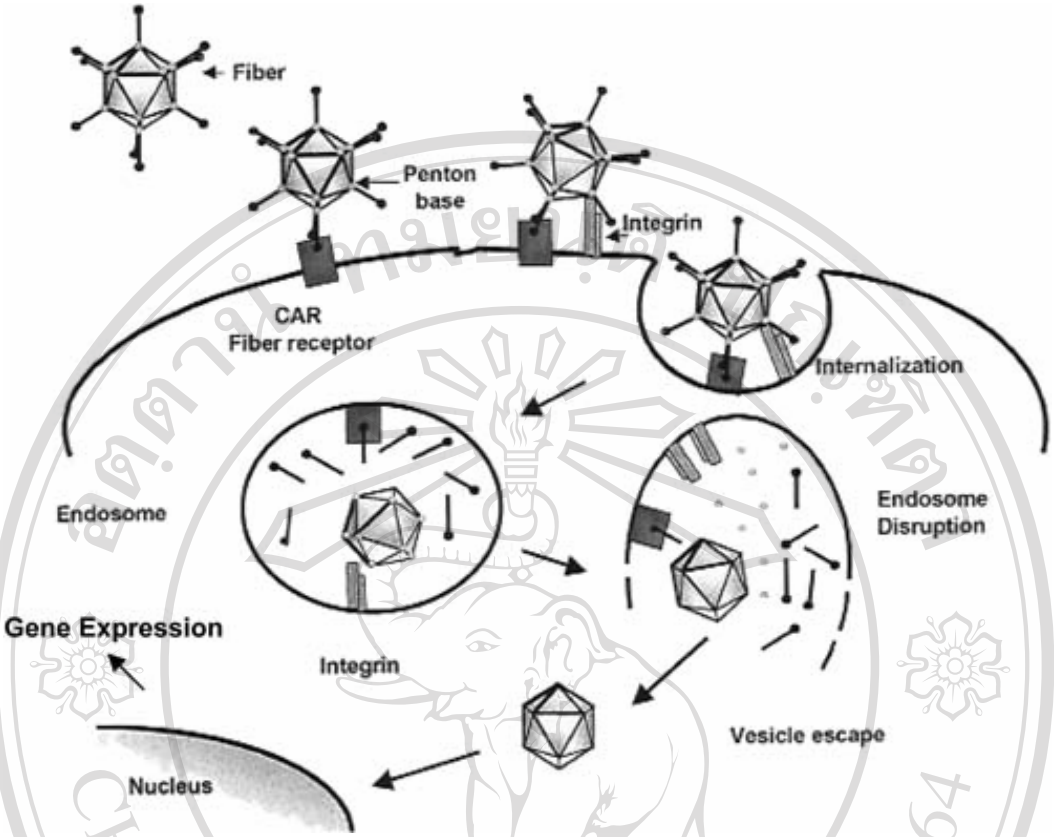

Figure 1.2 Cell-entry pathway of the adenovirus. The adenovirus initially binds to the cell via the specific cellular receptor, coxsackievirus and adenovirus receptor (CAR). After binding, the virion achieves internalization via receptor-mediated endocytosis pathway. This initial entry involves interaction with cellular integrin receptors via an arginine-glycine-aspartame motif in the adenoviral penton capsid protein. After internalization, the virus is localized within the cellular endosomes. Acidification of the endosomes allows the virions to be released within the cytosol and consequently the virion will be translocated into the nucleus to begin gene expression. (http://dels.nas.edu/ilar_n/ilarjournal/45_3/html/v4503contreras.shtml)

1.2.1.3 Transcription and replication

The adenoviral genome consist of a 36 kb-long linear double-stranded DNA with inverted terminal repeats (ITRs) of 100-400 bp in length at both ends acting as a *cis* elements during replication of viral genome. The terminal protein (TP), covalently attached to the 5' termini, serves as initiation primer for viral DNA replication. Besides the DNA molecule, the nucleoprotein core contains pV, pVI, and mu proteins, as well as the histone-like protein pVII. Altogether, the Ad particle has a molecular weight of about 150 MD (Majhen and Ambriovic-Ristov, 2006).

The adenovirus genome is organized in five early (E1A, E1B, E2, E3, and E4), four intermediate (IVa2, IX, VAI, and VAII) and one late transcriptional unit (Figure 1.3). Therefore, the adenovirus infectious cycle can be defined into two phases. The first (early) phase covers the entry of the virus into the host cell and the passage of the virus genome to the nucleus, followed by the selective transcription and translation of the early genes. The early transcription unit to be expressed after arrival in the cell nucleus is E1A. These E1A proteins are primarily concerned with modulating cellular metabolism to make the cell more susceptible to virus replication, E1A proteins activate early transcription units whose products induce the infected cell to enter S phase so as to create an optimal environment for virus replication. E1B proteins bind p53, Bak and Bax proteins and inhibit p53-dependent apoptosis, allowing survival of infected cells. The E2A unit encodes proteins included in replication of the viral genome such as DNA polymerase, preterminal protein and single-stranded DNA binding protein. E3 proteins subvert the host immune response and allow persistence of infected cells, while the E4 transcription unit encodes proteins influencing cell cycle control and transformation. Hence, these early event modulate the functions of

the cell so as to facilitate the replication of the virus DNA and the resultant transcription and translation of the late genes which are transcribed in the form of long precursor transcripts form the major late promoter (MLP) and encode structural capsid proteins. The virus also contains a virus-encode protease (Pr) which is necessary for processing of some of the structural proteins to produce mature infectious virus (Weber, 1976). This leads to the assembly in the nucleus of the structural proteins and the maturation of infectious virus. The early phase in a permissive cell can take about 6–8 h (depending on a number of extraneous factors), while the late phase is normally much more rapid, yielding virus in another 4–6 h (Russell, 2000).

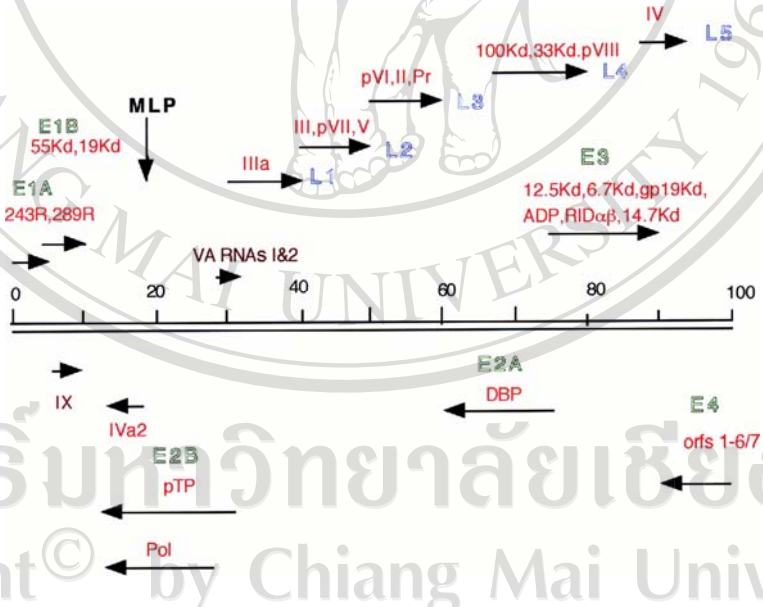

Figure 1.3 Transcription of the adenovirus genome. The early transcripts are outlined in green, the late in blue. Arrows indicate the direction of transcription. The gene locations of the VA RNAs are denoted in brown. MLP, Major Late Promoter. (http://vir.sgmjournals.org/cgi/content/full/81/11/2573/F2)

1.2.2 Adenoviral vector

Adenoviruses can infect a wide variety of cell types and tissue in both dividing and non-dividing cells. This features, together with their relative ease of preparation and purification, led them to be considered potential vectors for gene delivery. The virus can incorporate only 2 kb of foreign DNA without significant affects on its stability or its infectivity (Zhang, 1999). This vector can be utilized for cancer therapy to deliver genes that will lead to tumor suppression and elimination, gene therapy to deliver genes to tissues to augment defective genes, supplementary therapy to deliver genes which will combat disease processes.

In the first generation of vectors, the *E1* and/or *E3* gene cassettes were removed, allowing the introduction of up to 6.5 kb of foreign DNA, under the control of a heterologous promoter. In the case of the *E1* deletion, care was taken to ensure the retention of the ITR and the packaging sequences. Deletion of the *E1* genes had the additional apparent advantage of impairing the transcription of the *E2* genes which are E1 dependent, the replication of virus DNA and the production of the virus capsid proteins. The defective *E1* virus could be propagated by infection of 293 cells (Graham *et al*, 1977), which provide the *E1* gene products. Although many of the initial studies *in vitro* provided much promise, it has been reported that the expression of the *transgene in vivo* was only transient and was depressed because of the overwhelming immune response which against the virus capsid antigens as well as the expressed transgene. Furthermore, removal of the *E1B* products effectively disarmed of the mechanism for combating proapoptotic defenses and provides additional free space.

The next approach was to construct vectors with some or all of the *E2* genes excised and with the capacity to replicate virus DNA and to produce replication-competent adenoviruses (RCA) removed. Generation of RCAs could also be prevented by constructing cell lines that do not contain adenovirus sequences that overlap those in the vector. Nevertheless, the host immune response was still a major impediment to achieving persistent transgene expression and was particularly evident when repeated infections were attempted.

Moreover, the sophisticated vectors have been constructed by deleting other virus genes (Amalfitano *et al*, 1998) and the latest of these have all or nearly all of the virus genes removed. These so-called high 'capacity' or 'gutless' vector originally retained only the ITR and packaging sequences and required helper virus and appropriate complementing cells for propagation.

1.2.3 AdEasy system

Two approaches traditionally have been used to generate recombinant adenoviruses. The first involves direct ligation of DNA fragments of the adenoviral genome to restriction endonuclease fragments containing a transgene (Rosenfeld *et al*, 1991). The low efficiency of large fragment ligation and the scarcity of unique restriction sites have made this approach technically challenging. The second and more widely used method involves homologous recombination in memalian cells capable of complementing defective adenoviruses (Mittal *et al*, 1993). Homologous recombination results in a defective adenovirus that can replicate in the packaging line such as 293 or 911 cell which supplying the missing gene product (Graham *et al*, 1977). The desired recombinants are identified by screening individual plaques

generated in a lawn of packaging cells. Though this approach has proven extremely useful, the low efficiency of homologous recombination, the need for repeated rounds of plaque purification, and the long times required for completion of the viral production process have hampered more widespread use of adenoviral vector technology.

Therefore, the novel methods for generating adenoviral vectors were developed (Figure 1.4). First, the backbone vector, containing most of the adonoviral genome, is used in supercoiled form, obviating the need to enzymatically manipulate it. Second, the recombination is preformed in *Escherichia coli* rather than in mammalian cells. Third, there are no ligation steps involved in generating the adenoviral recombinants, as the process takes advantage of the highly efficient homologous recombination machinery present in bacteria. Fourth, the vectors allow inclusion of up to 10 kb of transgene sequences and allow multiple transgenes to be produced from the same virus. Fifth, some of the new vectors contain a green fluorescent protein (GFP) gene incorporated into the adenoviral backbone, allowing direct observation of the efficiency of transfection and infection, processes that have been difficult to follow with adenoviruses in the past. These modifications resulted in highly efficient viral production systems that often can obviate the need for plaque purification and significantly decrease the time required to generate usable virus (He *et al*, 1998).

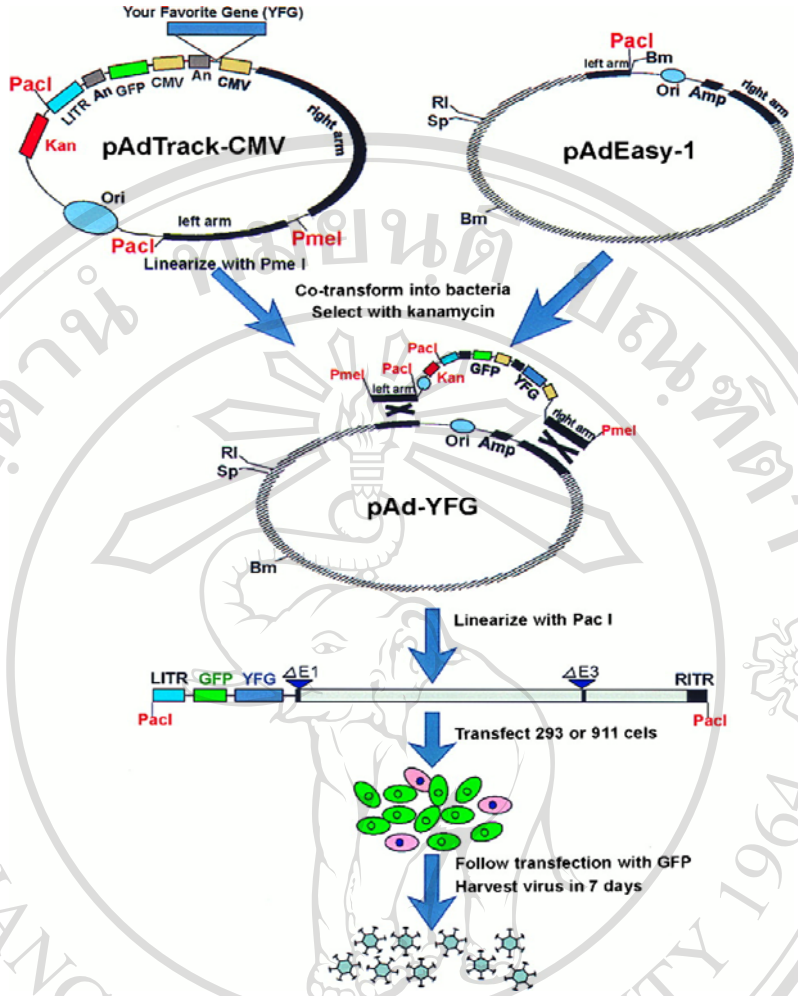

Figure 1.4 Schematic outline of the AdEasy system. The gene of interest is first cloned into a shuttle vector, e.g., pAdTrack-CMV. The resultant plasmid is linearized by digesting with restriction endonuclease *Pme*I, and subsequently cotransformed into *E. coli* BJ5183 cells with an adenoviral backbone plasmid, e.g., pAdEasy-1. Recombinants are selected for kanamycin resistance, and recombination was confirmed by multiple restriction endonuclease analyses. Finally, the linearized recombinant plasmid is transfected into adenovirus packaging cell lines, e.g., 911 or 293 cells. Recombinant adenoviruses typically are generated within 7-10 days. The "left arm" and "right arm" represent the regions mediating homologous recombination between the shuttle vector and the adenoviral backbone vector

1.2.4 Application of adenovirus vectors

The ability of adenovirus vectors to deliver and express genes at high yields has been amply demonstrated over the last 15 years. However, the superiority of the immune response *in vivo* has been a limiting factor in the practical development of vectors. Therefore, it is critical to suppress the immune response to the vector and the transgene. In contrast, cancer therapy may benefit from the induction of a vigorous immune response.

The majority of scientists agree that it is better to destroy cancer than to correct it genetically. Accordingly, the ultimate goal of many genes used in cancer gene therapy is to kill target cells. Cancer gene therapy strategies that developed until now could be divided into three group: (i) tumor suppressors, (ii) molecular chemotherapy, (iii) genetic immunopoteniation and (iv) vaccines.

(i) Tumor suppressors

Development of many cancers is accompanied by changes in expression of either tumor suppressor genes or oncogenes. Mutation in tumor suppressor genes like *p53* or *BRCA1* that lead to loss of function have been implicated in the development of a wide variety of human tumor cells. To induce apoptosis in the tumor cells, a number of vectors incorporating wild-type p53 have been constructed. Initial studies, using appropriate tumor cell lines and then animal model systems, demonstrated the efficacy of these approaches with lung cancer (Swisher *et al*, 1999; Schuler *et al*, 2001), human gliomas (Li *et al*, 1999) and cervical cancer (Hamada *et al*, 1996). Clinical trials testing the efficacies of these vectors in the treatment of lung, head, neck and liver cancers are under way. More recent investigations have illustrated the importance of the ARF-mdm2-p53 interaction in regulating p53 expression, and the

discovery of a range of mutations in ARF and in related transcription factor such as *Twist* (Maestro *et al*, 1999) in variety of tumors have suggested that vectors expressing other components of the p53 pathway might be equally productive.

When cancer is caused by over expression of oncogenes, such as *bcl-2* or *c-myc*, adenovirus vector can be used for introducing antisense oligonucleotides which will block oncogenes expression (Chen *et al*, 2001; Yang *et al*, 2002).

(ii) Molecular chemotherapy

The basis of molecular chemotherapy, also known as suicide gene therapy, is the introduction and expression of so-called suicidal genes, which encode enzymes needed for the metabolic conversion of nontoxic pro-drugs into toxic products capable of killing tumor cells. The many different enzyme/pro-drug combinations such as thymidine kinase from *Herpes simplex virus* (HSV-TK) and ganciclovir are the ones used most often. Additional effect in this approach is lateral diffusion of activated pro-drug into neighboring cells, known also as the bystander effect, which causes death of the cells that were not directly treated (Elshami *et al*, 1996).

(iii) Genetic immunopoteniation

Development of cancer indicates that the immune system was inefficient in recognizing and responding appropriately to tumor antigens. The reason for this is the weak immunogenicity of tumor cells and the fact that during its development, cancer causes neither inflammation nor tissue damage that could be recognized as a danger for organism. Genetic immunopotentiation strategies attempt to achieve active immunization against tumor-associated antigens by gene transfer either to tumor cells to enhance their immunogenicity, or to cellular components of the immune system to enhance their anti-tumor process. The anti-tumor immune response could be

augmented by introducing genes encoding cytokines like interleukin-12 (IL-12) into the tumor cells. Intratumoral injection of adenovirus vector bearing the IL-12 gene under the control of the CMV promoter in animals with a single large tumor nodule implanted in the liver caused significant inhibition of tumor growth (Barajas *et al*, 2001).

(iv) Vaccines

Cancer gene vaccination is an approach based on active immunization or modulation of cellular components of the immune system in order to enhance their capacity for recognizing and rejecting tumor antigens, and is possible only for the treatment of tumors that express specific tumor-associated antigens (Tuting *et al*, 1997). After presenting tumor antigens to the host immune system, effectors of the immune system are expected to be able to destroy tumor cells expressing the same antigens (Henderson *et al*, 2005). Most often the modified cells are antigen presenting cells such as dendritic cells or tumor infiltrating lymphocytes such as NK cells and cytotoxic T-lymphocytes specific for tumor antigens.

1.2.5 The human CD147

CD147, a member of the immunoglobulin superfamily, was discovered twenty five years ago, however, very little is known about its regulation and function *in vivo*. CD147 was identified as a tumor surface protein capable of inducing matrix metalloproteinase (MMP) that expressed in fibroblast, the protein was initially termed tumor cell-derived collagenase stimulatory factor (TCSF) (Ellis *et al*, 1989). Moreover, it was determined to be a transmembrane glycoprotein that was highly homologous with proteins of the immunoglobulin (Ig) superfamily (Kasinrerk *et al*, 1992) and was able to induce the production of several MMPs (Kataoka *et al*, 1993). Thus, it was renamed as the CD147/EMMPRIM (extracellular matrix metalloproteinase inducer) (Biswas *et al*, 1995; Gabison *et al*, 2005). CD147 is a pleiotropic molecule that is critical in fetal development and retinal function and has been shown to play a role in thymic T cell development, as well as many neurological processes ranging from the development of the nervous system to involvement in spatial learning and plaque formation within Alzheimer's stricken brains (Naruhashi *et al*, 1997; Yurchenko *et al*, 2002). Because of ubiquitous expression and pleiotropic nature of this molecule, it was used as a therapeutic targeting (Iacono *et al*, 2007).

CD147 is a transmembrane glycoprotein that was categorized as a member of the immunoglobulin superfamily (IgSF) of receptor (Kasinrerk *et al*, 1992; Biswas *et al*, 1995). CD147 was identified independently in various species and are referenced throughout the literature as EMMPRIN (Extracellular Matrix Metalloproteinase-Inducer), M6 (Human), Neurothelin, 5A11and HT7 (chicken), OX47 and CE9 (rat), and Basigin, gp42 (human and mouse) (Kirsch *et al*, 1997; Iacono *et al*, 2007). The CD147 gene is located of chromosome 19p13.3 and encodes a 29 kD protein, though

migration on SDS-PAGE usually occurs between 35-65 kD, depending on the degree of glycosylation (Kasinrerk *et al*,. 1992; Kirsch *et al*, 1997; Kasinrerk *et al*, 1999). CD147 is composed of an extracellular domain of 187 residues, a 24 residue transmembrane domain and a 40 amino acid cytoplasmic region (Biswas *et al*, 1995). Three glycosylation sites have been identified within the CD147 Ig like domains, two within the membrane proximal Ig domain and one within the distal Ig domain (Hakomori, 1996; Sun and Hemler, 2001). CD147 is broadly expressed on human peripheral blood cells, endothelial cell, hemopoietic and non-hemopoietic cell lines. Moreover, it is expressed significantly in neoplasm of bladder, liver and lung. The molecular function of neither CD147 nor any of its species homologues is fully understood. However, one suggested that CD147 is involved in signal transduction and cell adhesion functioning either directly as a signal transmitting adhesion molecule or as a regulator of adhesion.

1.2.6 Antibody

Plasma or serum glycoproteins are traditionally separated by solubility characteristics into albumins and globulins and may be further separated by migration in an electric field. Most antibodies are found in the third-fastest migrating group of globulins, named gamma globulins. Another common name for antibody is immunoglobulin (Ig), referring to the immunity-confering portion of the gamma globulin fraction. All antibody molecules share the same basic structural characteristics but display remarkable variability in the regions that bind antigens. This variability of the antigen-binding regions accounts for the capacity of different antibodies to bind a tremendous number of structurally diverse antigens. There are

more than 10^7, and perhaps as many as 10^9, different antibody molecules in every individual, each with unique amino acid sequences in their antigen-combining sites. The effector functions and common physicochemical properties of antibodies are associated with the non-antigen-binding portions, which exhibit relatively few variations among different antibodies.

An antibody molecule has a symmetric core structure composed of two identical light chains and two identical heavy chains. Each light chain is about 24 kD, and each heavy chain is 55 to 70 kD. One light chain is covalently attached to one heavy by a disulfide bond, and the two heavy chains are attached to each other by disulfide bonds. Both the light chains and the heavy chains contain a series of repeating homologous units, each about 110 amino acids residues in length, that fold independently in a globular motif that is called and Ig domain. An Ig domain contains two layers of β-pleated sheet, each layer composed of three to five strands of antiparallel polypeptide chain.

Many other proteins of importance in the immune system contain domains that use the same folding motif and have amino acid sequences that are similar to Ig amino acid sequences all molecules that contain this motif are said to belong to the Ig superfamily, and all the gene segments encoding the Ig domains of these molecules are believed to have evolved from one ancestral gene. Both heavy chains and light chains consist of amino terminal variable (V) regions that participate in antigen recognition and carboxy terminal constant (C) regions. In the heavy chain, the V region is composed of one Ig domain and the C region is composed of three or four Ig domains. Each light chain is made up of one V region Ig domain and one C region Ig domain (Figure 1.5). Variable regions are so named because they contain regions of

variability in amino acid sequence that distinguish the antibodies made by one clone of B cells from the antibodies made by other clones. The V region fo one heavy chain (V_H) is juxtaposed with the V region of one light chain (V_L) to form an antigen-binding site. Because the core structural unit of each antibody molecule contains two heavy chains and two light chains, it has two antigen-binding sites. The C region domains are separate form the antigen-binding site and do not participate in antigen recognition. The heavy chain C regions interact with other effector molecules and cells of the immune system and therefore mediated most of the biologic function of antibodies. In addition, the carboxy terminal ends of the heavy chains anchor membrane-bound antibodies in the plasma membranes of B lymphocytes. The C regions of light chains do not participate in effector functions and are not attached to cell membranes.

Most of the sequence differences among different antibodies are confined to three short stretches in the V regions of heavy and light chains called the hypervariable segments. These hypervariable regions are each about 10 amino acid residues long, and they are held in place by more conserved framework sequences that make up the Ig domain of the V region. In an antibody molecule, the three hypervariable regions of a V_L domain and the three hypervariable regions of a V_H domain are brought together in three-dimentional space to form an antigen-binding surface. Because these sequences form a surface that is complementary to the three-dimentional structure of the bound antigen, the hypervariable regions are also called complementary-determining regions (CDRs). Proceeding from either the V_L or the V_H amino terminus, these regions are called CDR1, CDR2, and CDR3, respectively. The

CDR3s of both the V_H segment and the V_L segment are the most variable of the CDRs.

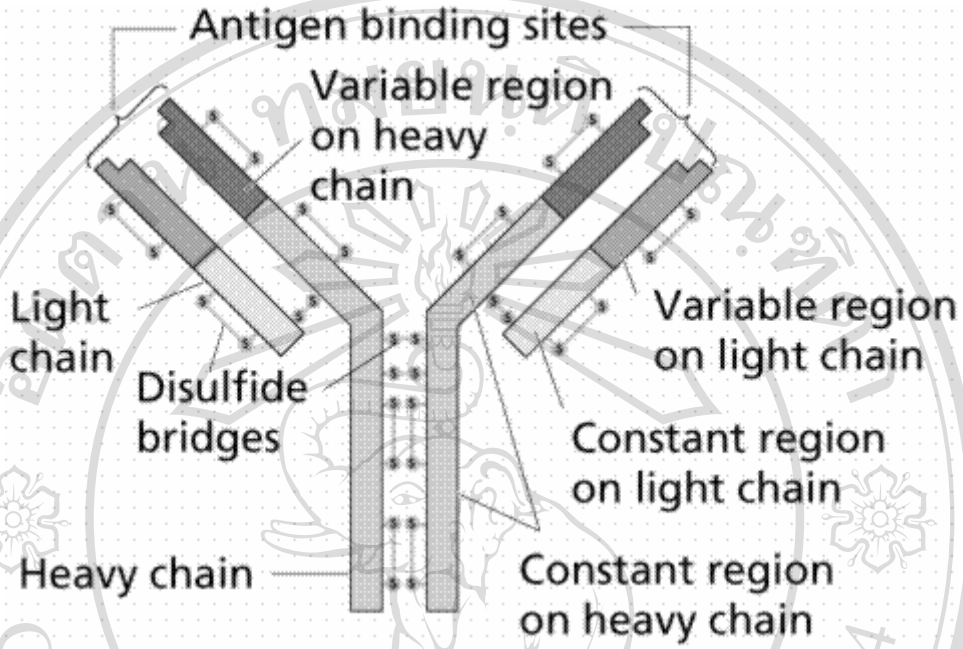

Figure 1.5 Structural regions of an antibody molecule. (http://www.cartage.org.lb/en/themes/sciences/lifescience/generalbiology/Physiology/LymphaticSystem/Antibodymediated/Antibodymediated.htm)

Crystallographic analyses of antibodies reveal that the CDRs form extended loops that are exposed on the surface of the antibody and are thus available to interact with antigen. Sequence differences among the CDRs of different antibody molecules result in unique chemical structures being displayed at the surfaces of the projecting loops and therefore in different specificities for antigens. The ability of a V region to fold into an Ig domain is mostly determined by the conserved sequences of the framework regions adjacent to the CDRs. Confining the sequence variability to three short stretches allows the basic structure of all antibodies to be maintained despite the

variability among different antibodies. Antigen binding by antibody molecules is primarily a function of the hypervariable regions of V_H and V_L. Crystallographic analyses of antigen-antibody complexes show that the amino acid residues of the hypervariable regions form multiple contacts with bound antigen. The most extensive contact is with the third hypervariable region (CDR3), which is also the most variable of the three CDRs. However, antigen binding is not solely a function of the CDRs, and framework residues may also contact the antigen. Moreover, in the binding of some antigens, one or more of the CDRs may be outside the region of contact with antigen, thus not participation in antigen binding (Abbas and Lichtman, 2005).

1.2.7 Single-chain Fv

The driving forces for the development of small antibody fragments was the make these molecules suitable for the radioimmunotherapy of cancer. The large molecular size of intact IgG resulted in increased circulation times, which caused dose-limiting myelotoxicity in therapeutic applications and a high background in imaging. In addition, effector functions mediated by the Fc region are generally undesirable in some applications. The initial attempts to make smaller fragments were based on the digestion of intact IgGs by the proteolytic enzymes pepsin or papain, to generate $F(ab')_2$ or Fab' fragments. These fragments retained their specific antigen-binding capabilities and exhibited faster blood clearance and homogenous tumor penetration compared with intact IgGs. Advances in recombinant DNA technology facilitated the production of smaller recombinant antibody fragments (Jain *et al*, 2007). Antibody engineering started with the development of antibody fragments encoded by a single gene, wherein only the regions involved in antigen binding, the

V_H and V_L, were linked by a short peptide. The resulting protein was called a single-chain Fv (scFv) and was monovalent in binding (Batra *et al*, 2002). In general, the affinity of a given scFv is comparable with the parental Fab fragment but owing to their monovalent binding such antibodies are of limited therapeutic significance. However, because of their small size (30 kD), most scFvs have a short serum half life compared with the intact IgG, which can make them useful for diagnostic applications such as imaging. Moreover, scFv are therefore the most commonly used format for intracellular expression. Their small size facilitates expression and assembly of functional molecules (Sanz *et al*, 2005).

1.2.8 Protein translocation

The cell faces a number of challenges in nescent proteins (proteins whose synthesis has been newly initiated) into the secretory pathway. First, the proteins must be selectively recognized and brought to their sites of translocation at the ER membrane. Next, they must translocate across the ER membrane either completely (in the case of soluble proteins) or partially (in the case of membrane proteins). This must occur without allowing the general exchange of other molecules between the lumen of the ER and the cytosol. Finally, all translocated proteins must be properly folded and in many cases must also be modified or assembled with other proteins in the ER lumen.

The ER is only one of several membranes enclosed organelles to which targeting can occur. Proteins that are destined for the ER must be distinguished from those that are bound for other organelles or destined to remain in the cytosol. Cells accomplish this segregation through the use of signal sequences. These sequences are

discrete stretches of amino acids within a protein's primary structure that are recognized by machinery associated with the target organelle.

The mechanism by which a signal sequence for the ER is recognized influences how the protein is translocated. The most common form of translocation into the ER is cotranslational translocation, meaning that it occurs while the protein is being synthesized by a ribosome bound to the membrane. This form of translocation is initiated when the signal sequence is recognized in the cytosol by a complex known as the signal recognition particle, or SRP. SRP binds to the signal sequence and brings both the protein and the ribosome that is synthesizing it to the ER by interacting with a receptor on ER membrane. However, some ER signal sequences do not interact with SRP. This causes the proteins to be translocated posttranslationally, after their synthesis in the cytosol is complete. Different organisms use the two form of translocation to different extents. In mammals, almost all translocation is cotranslational, but in simpler eukaryotes, such as the yeast *Saccharomces cerevisiae*, both forms occur.

When a protein has been targeted to the ER, it must cross the lipid bilayer enclosing the organelle. This occurs through a channel that provides an aqueous passageway across the hydrophobic membrane. The proteins that make up and associate with this channel are collectively called the translocon to emphasize their function as an integrated unit. The channel is gated, meaning that it opens only during the translocation of a nascent protein. Gating prevents the passage of other material, such as small molecules, ions, and other proteins, and allows the cytosol and the interior of the ER to be maintained as distinct compartments.

When translocation begins, nascent transmembrane proteins must be distinguished form proteins that need to pass through the channel completely. This distinction, like recognition of the signal sequence, is made by the channel. Movement across the lipid bilayer halts when a transmembrane domain, a hydrophobic segment destined to span the lipid bilayer, is recognized by the translocon and moved laterally out of the channel into the lipid bilayer. This process can occur multiple times in a single polypeptide, helping create the complex topologies of large multispanning membrane protein.

For both secretory and transmembrane proteins, translocation is also coordinated with activities that modify the polypeptide. For instance, most signal sequences are removed early in translocation, many translocating proteins also have complex carbohydrate structures added. Others are cleaved near the end of translocation and covalently attached to phospholipids.

Each translocated in the ER assist nascent proteins in the folding process. Some, called molecular chaperones, bind to nascent proteins and protect them from either misfolding or aggregating; others rearrange disulfide bonds on the nescent proteins or assist in the assembly of multimeric proteins. Together, all of these proteins form a system of quality control that ensures the proper folding and assembly of proteins within the ER. Closely coordinated with quality control is the retrograde translocation system, which identifies misfolded proteins and returns them to the cytosol for degradation. Only when all the steps of quality control have been satisfied the secretory and membrane proteins can leave the ER and proceed to their final destinations *via* the secretory pathway (Lewin *et al*, 2007).

1.3 Objective

1. To construct the adenoviral plasmids and produce recombinant adenoviruses expressing scFv-M61B9 in cytoplasmic compartment and as a secreting molecule.
2. To produce the scFv-M61B9 in mammalian cell compartments.
3. To study the binding activity and specificity of produced scFv-M61B9 in each cell compartments.