CHAPTER I

INTRODUCTION

1.1 Statement and significance of the problem

Antibodies have received considerable attention for many decades, regarding to their indispensable biological properties. The broad use for diagnostics, structural and functional studies, and more recently as protein therapeutics makes antibodies become the valuable and powerful tools in the analyses of proteins. In general, one of the most important steps in antibody production is the preparation of the immunogen. To obtain high quality of specific antibodies, purity of immunogen is crucial. However, the processes to obtain enough amount of purified immunogen are usually cumbersome. The availability of recombinant DNA technologies make it possible for cloning any gene fragments and expressing the target proteins. Prokaryotic expression systems utilizing *Escherichia coli* are still the most extensively used for large-scale production of recombinant proteins. By manipulating the cloning strategies, the target proteins can be expressed as hybrid molecules with affinity tags to allow the affinity purification of the fusion proteins.

The biotin–avidin/streptavidin system is used in numerous biotechnological and diagnostic applications, primarily due to the high affinity of the proteins avidin and streptavidin to a small biotin molecule. Only one natural protein that is post-translationally biotinylated in *E. coli* is the biotin carboxyl carrier protein (BCCP), a

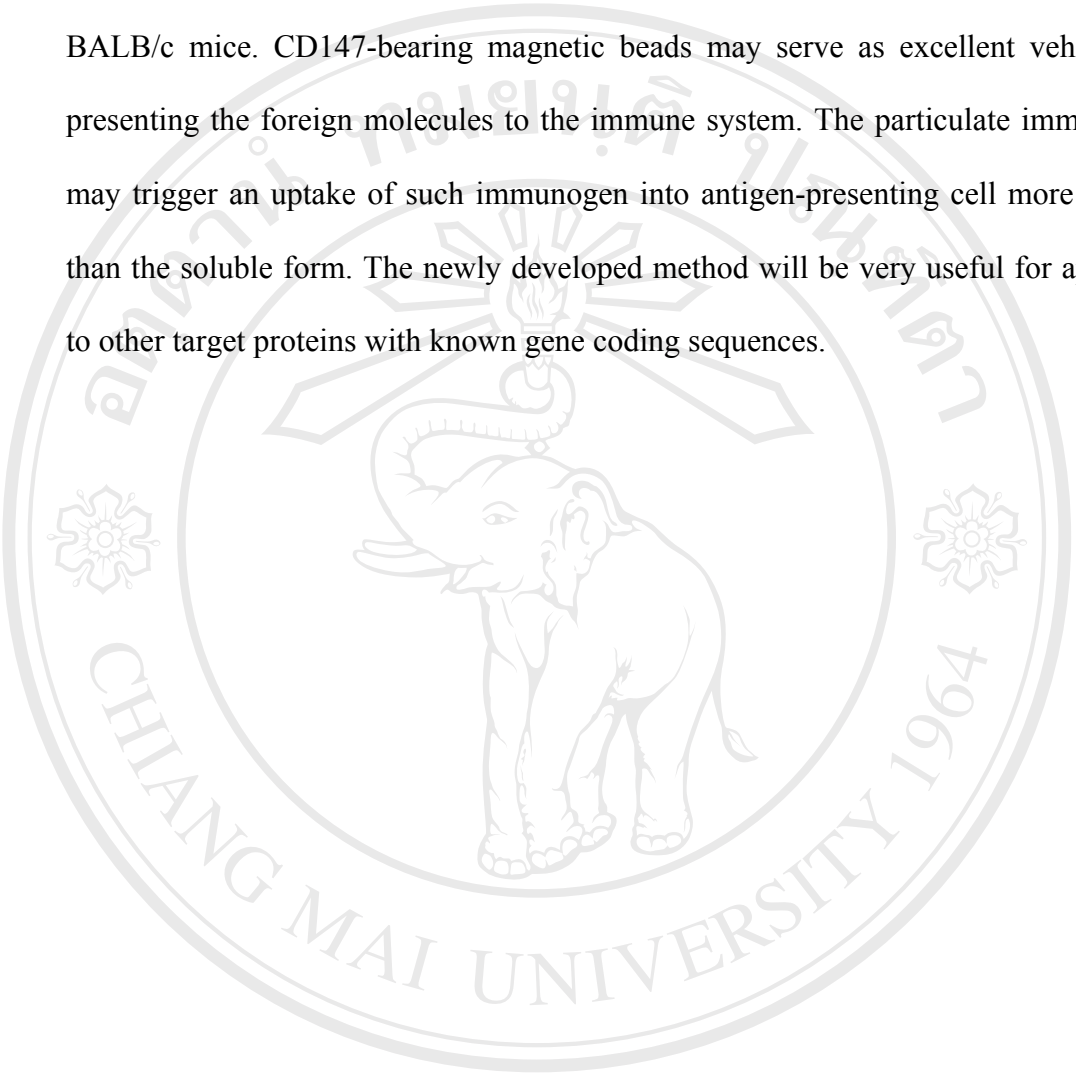
subunit of enzyme acetyl-CoA carboxylase. The reaction is carried out by the endogenous enzyme biotin protein ligase (BirA) which catalyse the formation of a covalent linkage between the carboxyl group of biotin and the ε-amino group of the specific lysine residue located inside the C-terminal half of the BCCP molecule. This domain can be fused to any proteins and therefore serves as target for *in vivo* biotinylation by the endogenous biotin ligase of *E. coli*. In comparison with chemical biotinylation, the number and location of attached biotin are variable, leading to the alteration of biological functions or disruption of the protein structure. In the other hand, the enzymatic biotinylation guarantees the site-specific attachment of only one biotin molecule per protein, a property which can be utilized, for example, for the immobilization of protein molecules in a uniform, bioactive orientation.

The most important reason that the recombinant proteins expressed in *E. coli* should be translocated to the periplasm is the extreme reducing environment of the bacterial cytoplasm. This reducing condition, maintained by two systems (thioredoxin system and glutaredoxin system), keeps proteins in the reduced form and disfavors disulfide bond formation, impairing the production of functional proteins in the cytoplasm. However, the enzymatic biotinylation carried out by biotin protein ligase can occur only inside the cells. Therefore, the special *E. coli* strain with modified redox potential was introduced. This strain, FA113 (commercially available under the name Origami), contains double mutation on thioredoxin reductase (*trxB*) and glutathione reductase (*gor*), rendering the efficient formation of disulfide bond in the cytoplasm. They grow normally and do not need any exogenous reductant even though the cytoplasm is highly oxidized.

In this study, the extracellular domain of CD147 was expressed in cytoplasm of *E. coli* Origami B as fusion protein with biotin acceptor domain of BCCP to generate biotinylated form of CD147. The strain Origami B carries the same mutation as Origami (*trxB gor*) plus the characteristics of both BL21 and Tuner™ strain (deficient in *lon* and *omp-t* proteases, and *lac* permease mutation, respectively). CD147 is a member of human leukocyte surface molecule of the immunoglobulin superfamily and found on the surface of various cell types e.g. cancer cells and activated T-lymphocyte. Although the information of mechanism in signal transduction *via* CD147 is not clearly proposed and its surface receptor has not yet been identified, CD147 is speculated to be an essential molecule in the immune system. Since the extracellular portion of CD147 contains 2 domains that are similar to the domain of human immunoglobulins, the formation of the intradomain disulfide bridges can occur by this approach.

Regarding the fact that affinity purification using column chromatography requires sophisticated steps, time-consuming and cost-effective, the alternative purification method of biotin-containing molecules using streptavidin-coated magnetic beads was explored. The biotinylated CD147 could be easily immobilized on the streptavidin magnetic beads and separated from the crude bacterial extracts by magnetic field sorting. In order to produce antibody responses to CD147, the CD147 fusion protein coated beads were used as immunogen. The method of using biotinylated fusion protein captured on streptavidin magnetic beads as an immunogen was proven to be rapid and did not require the cumbersome purification of recombinant antigen.

Herein, we introduced a novel strategy to prepare *in vivo* biotinylated CD147-BCCP fusion protein for the production of polyclonal antibodies against CD147 in BALB/c mice. CD147-bearing magnetic beads may serve as excellent vehicle for presenting the foreign molecules to the immune system. The particulate immunogen may trigger an uptake of such immunogen into antigen-presenting cell more readily than the soluble form. The newly developed method will be very useful for applying to other target proteins with known gene coding sequences.



1.2 Literature review

1.2.1 Recombinant protein expression in *Escherichia coli*

High-level production of recombinant proteins as a prerequisite for subsequent purification has become a standard technique. Important applications of recombinant proteins are; a) immunization, b) biochemical studies, c) three-dimensional analysis of the protein, and d) biotechnological and therapeutic use. Production of recombinant proteins involves cloning of the appropriate gene into an expression vector under the control of an inducible promoter. Efficient expression of the recombinant gene depends on a variety of factors such as optimal expression signals (both at the level of transcription and translation), correct protein folding and cell growth characteristics. In addition, the selection of a particular expression system requires a cost breakdown in terms of design, process and other economic considerations (Schumann and Ferreira, 2004).

Many advantages of *Escherichia coli* have ensured that it remains a valuable organism for the high-level production of recombinant proteins. Although the *E. coli* expression systems still have a major drawback that it is unable to perform many of the posttranslational modifications found in eukaryotic proteins, many eukaryotic proteins retain their full biological activity in a non-glycosylated form and therefore can be produced in *E. coli* (Fuh et al., 1990; Tucker and Grisshammer, 1996).

1.2.1.1 Configuration of expression vectors

The construction of an expression plasmid requires several elements whose configuration must be carefully considered to ensure the highest levels of protein synthesis. The promoter is positioned upstream of the ribosome-binding site (RBS) and is under the control of a regulatory gene, which may be present on the vector itself or integrated in the host chromosome. A promoter should have certain characteristics to render it suitable for high-level protein synthesis (Goldstein and Doi, 1995). First, it must be strong, resulting in the accumulation of protein making up 10 to 30% or more of the total cellular protein. Second, it should exhibit a minimal level of basal transcriptional activity. Tight expression of transcription of recombinant genes is often desirable or necessary since leaky expression can be detrimental or even lethal to cell growth. And third, its induction is simple and cost-effective. Promoter induction is either thermal or chemical and the common used inducer is, for example, sugar analog isopropyl-b-D-thiogalactopyranoside (IPTG).

Downstream of the promoter is the RBS which include the Shine-Dalgarno (SD) sequence and a translation initiation codon. SD site interacts with the complementary 3' end of the 16S rRNA during translation initiation. The transcription terminator is located downstream of the coding sequence and serves both as a signal to terminate transcription and as a protective element protecting the mRNA from exonucleolytic degradation and extending the mRNA half-life (Sorensen and Mortensen, 2005).

In addition to the above elements that have a direct impact on the efficiency of gene expression, vectors contain a gene that confers antibiotic resistance on the host to aid in plasmid selection and propagation. The most common drug

resistance markers confer resistance to, for example, ampicillin, kanamycin, chloramphenicol or tetracycline. The last essential component in expression vector is the origin of replication which determines the plasmid copy number.

1.2.1.2 Protein folding in *E. coli*

Newly synthesized polypeptide chains must fold and assemble into unique three-dimensional structures in order to attain their biological function. In general, overexpressed recombinant proteins often require the assistance of folding modulators; molecular chaperones and folding catalysts, for efficiently reach their native conformation. Molecular chaperones, a ubiquitous class of folding modulators, play a central role in the conformational quality control of the proteome by interacting with, stabilizing and remodeling a wide range of nonnative polypeptides. Although constitutively expressed under balanced growth conditions, many chaperones are upregulated upon heat shock or other insults that increase protein misfolding and are therefore classified as stress or heat shock proteins (Hsps). Mechanistically, molecular chaperones rely on the differential exposure of structured hydrophobic domains to the solvent to bind nonpolar segments that would normally be buried within the core of their substrates (Baneyx and Mujacic, 2004).

Molecular chaperones can be divided into three functional subclasses based on their mechanism of action. 'Folding' chaperones (e.g., DnaK and GroEL) rely on ATP-driven conformational changes to mediate the net refolding/unfolding of their substrates. 'Holding' chaperones (e.g., IbpB) maintain partially folded proteins on their surface to await availability of folding chaperones upon stress abatement.

Finally, the 'disaggregating' chaperone ClpB promotes the solubilization of proteins that have become aggregated as a result of stress (Baneyx and Mujacic, 2004).

1.2.2 Disulfide bond in *E. coli*

Disulfide bonds play a number of different roles in protein structure and activity. For many proteins, disulfide bonds are permanent features of the final folded product. The formation of these bonds may be essential steps in the folding pathway or they may simply enhance the stability of the protein (Creighton, 1997; Vanhove et al., 1997). Certain proteins that act as reductases, oxidation and reduction of cysteines residues at active sites confer the biological activity (Rietsch and Beckwith, 1998). Recently, a finding shows that disulfide bonds can function to promote a folding pathway even when the final protein product does not contain any such bonds (Robinson and King, 1997).

1.2.2.1 Disulfide bond reducing pathway

The most important reason that the recombinant proteins expressed in *E. coli* should be translocated to the periplasm is the extreme reducing environment of the bacterial cytoplasm. There are three reasons to keep proteins in the reduced form: a) a number of enzymes rely on a reduced cysteine residue in their active sites, b) most proteins present in the periplasm are translocated in an unfolded conformation, and c) a number of virulence factors and toxins contain multiple disulfide bonds (Schumann and Ferreira, 2004). This reducing condition, maintained by two systems (thioredoxin system and glutaredoxin system), keeps proteins in the reduced form and disfavors disulfide bond formation (Rietsch and Beckwith, 1998). In general, the *E.*

coli cytoplasm contains two thioredoxins, TrxA and TrxC, and three glutaredoxins (Aslund and Beckwith, 1999). Thioredoxins are small 12-kDa proteins with the Trp-Cys-Gly-Pro-Cys-Lys active site motif (Watson et al., 2004). The glutaredoxins also contain this motif and a thioredoxin-like fold (Holmgren and Aslund, 1995; Holmgren and Bjornstedt, 1995). Catalyzing of disulfide bond reduction in protein substrates makes both thioredoxins and glutaredoxins become oxidized. The oxidized form of these proteins can catalyze the formation of disulfide bonds in peptides. However, both the thioredoxins and the glutaredoxins are maintained in a reduced state by the action of thioredoxin reductase (TrxB) and glutathione, respectively. In *E. coli*, glutathione is synthesized by the *gshA* and *gshB* gene products. The enzyme glutathione oxidoreductase, the product of the *gor* gene, is required to reduce oxidized glutathione and complete the catalytic cycle of the glutathione-glutaredoxin system (Figure 1.1).

1.2.2.2 Disulfide bond forming pathway

In cytoplasm, not only those two major disulfide bond-reducing pathways, the absence of enzymes catalyzing disulfide oxidation also contributes to disulfide bond prevention. The periplasm contains several enzymes involved in the formation of disulfide bonds which are grouped into two pathways, the oxidation and the isomerization pathway (Bardwell et al., 1991; Rietsch et al., 1996). In the oxidation pathway, DsbA with two oxidized thiol groups transfers its disulfide to pairs of cysteines in substrate proteins by a thiol-disulfide exchange reaction and becomes reduced. To get oxidized again, it interacts with DsbB, an integral membrane

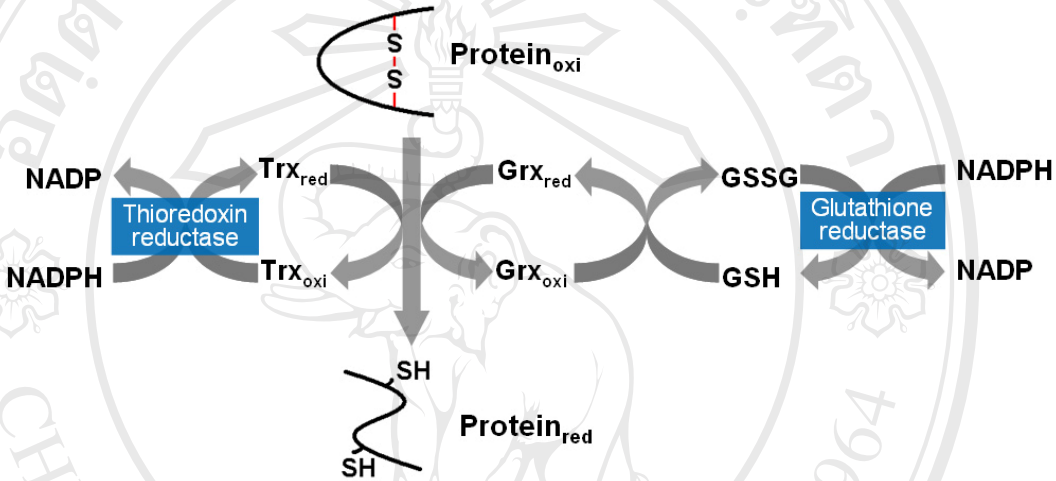


Figure 1.1 Cytoplasmic disulfide bond reducing system in *E. coli*. The two separate pathways, thioredoxin (Trx) and glutaredoxin (Grx) pathways, transferring electron from NADPH via the activity of two flavoenzymes, thioredoxin reductase and glutathione reductase, to reduce the protein disulfide bond in the cytoplasm.

protein which contains two disulfide bonds. If the target protein contains more than two thiol groups, DsbA may form a wrong disulfide bond. This is recognized by the isomerization system which consists of three proteins. The reduced forms of DsbC and DsbG can recognize wrongly formed disulfide bonds on target proteins and catalyze the formation of the correct bonds thereby becoming oxidized. Reduction of the disulfide bonds occurs through interacting with the integral membrane protein DsbD which in turn becomes reduced again through interaction with thioredoxin (Hiniker and Bardwell, 2003) (Figure 1.2).

1.2.3 Inclusion body and how to prevent its formation

Stress situations including heat shock, starvation, exposure to toxic compounds, recombinant protein over-expression, and oxidative stress, impair protein folding and cause the formation of folding intermediates and protein misfolding. In recombinant bacteria, the rapid production of plasmid-encoded gene products triggers the accumulation of high concentrations of folding intermediates which further form insoluble aggregates designated as inclusion body (Kiefhaber et al., 1991; Villaverde and Carrio, 2003). These aggregates are large, spherical particles that do not consist of pure recombinant polypeptide chains, but contain several impurities such as host proteins (RNA polymerase, outer membrane proteins), ribosomal components and circular and nicked forms of plasmid DNA (Schumann and Ferreira, 2004). The factors contributing to the formation of inclusion body are the use of high inducer concentration, strong promoter and the failure of the quality control system to repair or remove misfolded or unfolded protein (Baneyx and Mujacic, 2004).

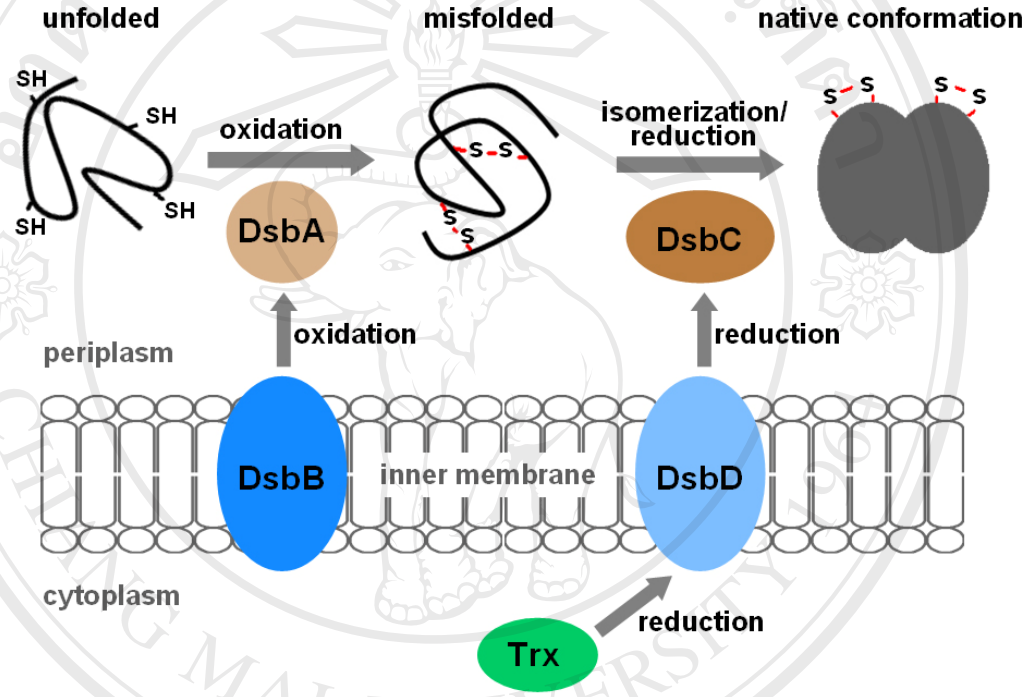


Figure 1.2 Oxidative pathways for protein disulfide bond formation in periplasm. Cysteine pairs in proteins containing disulfide bonds are oxidized by DsbA whereas incorrect disulfides are isomerized by DsbC. These oxidoreductasees are reactivated by DsbB and DsbD, respectively.

1.2.3.1 Decrease the recombinant protein production rate

Strategies to prevent the formation of inclusion bodies are aimed to slow down the production of recombinant proteins by regulating the parameters such as temperature, expression rate, and host metabolism. A lower level of protein synthesis from low-copy number vectors, weaker promoter or from a strong promoter under conditions of partial induction is found to result in a higher amount of soluble proteins (Hockney, 1994). Growth at lower temperatures is a well known technique for facilitating correct folding and has proven effective in improving the solubility of a number of difficult proteins including human interferon α-2, bacterial luciferase, Fab fragments and β-lactamase (Vasina and Baneyx, 1997). The reason why a lower temperature favors the native state is related to a number of factors, including a decrease in the driving force for protein self-association, a slower rate of protein synthesis, changes in the folding kinetics of the polypeptide chain, etc (Vasina and Baneyx, 1997; Schumann and Ferreira, 2004).

1.2.3.2 Co-expression of molecular chaperones

A possible strategy for the prevention of inclusion body formation is the co-expression of molecular chaperones. DnaK (Hsp70 chaperone family) prevents the formation of inclusion bodies by reducing aggregation and promoting proteolysis of misfolded proteins. GroEL (Hsp60 chaperone family) operates the protein transit between soluble and insoluble protein fractions and participates positively in disaggregation and inclusion body formation (Mogk et al., 2002).

1.2.3.3 *E. coli* strains for improving soluble expression

Many proteins contain disulfide bridge that serves an important function in stabilizing the protein structure. The cytoplasmic reducing environment prevents the efficient formation of disulfide bonds and this usually impairs the production of functional proteins inside the cells. However, the redox-modified *E. coli* strain Origami (Novagen), which contains double mutation on both thioredoxin reductase (*trxB*) and glutathione reductase (*gor*), allowed the formation of disulfide bond in the cytoplasm. In this strain, the absence of both reductase enzymes cause the accumulation of thioredoxins and glutaredoxins in the oxidized form, which actively promote disulfide bond formation in certain substrate proteins (Rietsch and Beckwith, 1998; Bessette et al., 1999). This *E. coli* strain was shown to enable efficient cytoplasmic expression of various recombinant proteins (Jurado et al., 2002; Venturi et al., 2002; Kersteen et al., 2004; Santala and Lamminmaki, 2004). Overexpression of the periplasmic foldase DsbC in the cytoplasm stimulates disulfide bond formation further (Bessette et al., 1999).

1.2.3.4 Use of protein fusion

The use of fusion tags in recombinant protein purification has a long tradition. In addition to facilitate purification strategies, fusion tags have been observed to improve protein yield, to prevent proteolysis and to increase solubility *in vivo* (Sorensen and Mortensen, 2005). Although many proteins are highly soluble, they are not all effective as solubility enhancers. Among the most potent solubility enhancing proteins characterized to date is the *E. coli* maltose binding protein (MBP). MBP acts as solubility enhancing partner and is especially suited for the expression of

proteins prone to form inclusion bodies. A precise mechanism for the solubility enhancement of MBP has not been found. However, MBP might act as a chaperone by interactions through a solvent exposed "hot spot" on its surface which stabilizes the insoluble passenger protein (Bach et al., 2001; Fox et al., 2001).

1.2.4 Fusion technology and its role in recombinant protein expression

Escherichia coli, the preferred host for recombinant protein expression, present many challenges such as proteolytic degradation of target proteins, protein misfolding, poor solubility, and the necessity for good purification methodologies. Much advancement has been made toward overcoming those challenges including the development of strong promoters, co-expression with chaperones, and through the use of protein fusions (Butt et al., 2005). A wide range of fusion systems has been developed in order to improve heterologous expression and simplify purification methodologies (Table 1.1). These fusion partners, or affinity tags, share the following features; a) one-step affinity purification, b) a minimal effect on tertiary structure and biological activity, c) easy and specific removal to produce the native protein, d) simple and accurate detection of the recombinant protein, and e) applicability to a number of different proteins (Terpe, 2003). However, there is no single affinity tag that is ideally suited for all of these purposes, it also depends on expression conditions and the target proteins itself. Thus, the strength and weakness of certain affinity tags must be considered in order to obtain the highest outcome on each application.

1.2.4.1 Increasing the yield of recombinant proteins

Recombinant proteins are often viewed as unwanted by bacterial cells and are subjected to proteolytic degradation, decreasing the level of recombinant protein expression (Rozkov and Enfors, 2004). Fusions between the N-terminus, C-terminus, or dual C- and N-terminus of target proteins and some fusion tags, e.g. N-utilization substance A (NusA) and glutathione S-transferase (GST), have been shown to protect the target protein from degradation (Terpe, 2003; De Marco et al., 2004). Moreover, fusions can promote the translocation of their partner proteins to different cellular compartments, decreasing the concentration of the recombinant protein in the protease-rich cytosol. For example, MBP can be translocated to the membrane compartment of the cell (Nikaido, 1994). Addition advantage of fusion tags in improving the yield of recombinant proteins is that they provide a reliable context for efficient translation initiation. Many different tags can enhance the efficiency of translation initiation when fused at the N-terminus of the target proteins (Waugh, 2005).

1.2.4.2 Enhancing the solubility of recombinant protein

One disadvantage when heterologous proteins are over-produced in *E. coli* is that proteins frequently aggregate and form inclusion bodies. Certain affinity tags have been shown to promote the solubility of their partners, although the exact mechanisms by which they improve solubility have not been described (Waugh, 2005). Solubility-enhancing tags tend to be large proteins rather than peptide, for example MBP, NusA, DsbA and TrxA (Terpe, 2003). It has been hypothesized that these tags may act as molecular chaperones by binding to aggregation-prone folding

Table 1.1 Characteristics, advantages and disadvantages of some commonly used fusion tags

Tag	Residues	Size (kDa)	Matrix	Advantages	Disadvantages
His_6-tag	6	0.84	Ni^{2+}-NTA	Low metabolic burden Inexpensive affinity resin	Specificity is rather moderate Does not enhance solubility
FLAG-tag	8	1.01	Anti-FLAG mAb	Low metabolic burden High specificity	Expensive affinity resin Harsh elution condition
Strep-tag II	8	1.06	Strep-Tactin	Low metabolic burden High specificity	Expensive affinity resin Does not enhance solubility
S-tag	15	1.75	S-protein	Low metabolic burden High specificity	Expensive affinity resin Does not enhance solubility Harsh elution condition
Calmodulin-binding peptide	26	2.96	Calmodulin	Low metabolic burden High specificity	Expensive affinity resin Does not enhance solubility
Glutathione S-transferase	211	26.00	Glutathione	Efficient translation initiation Mild elution condition	High metabolic burden Homodimeric protein
Maltose-binding protein	396	40.00	Cross-linked amylose	Efficient translation initiation Enhance solubility	High metabolic burden

intermediates of passenger proteins and preventing their self-association (Fox et al., 2001). However, some disadvantages are that these large tags may interfere with the passengers, cause high metabolic burden and, in some case, can not be purified with specific affinity matrices (e.g. TrxA, DsbA). Therefore, many fusion constructs use the combination of solubility-enhancing tags with a small affinity tags for purification purpose (Tucker and Grisshammer, 1996).

1.2.4.3 Facilitating the purification and detection of recombinant proteins

The use of protein fusion technology offers the opportunity to simplify and facilitate purification and detection of recombinant proteins. The most important advantage of affinity-based purification procedures carried out by affinity tags is that the target protein never has to interact directly with a chromatographic matrix. The reason is that if the tightest binding conformation of a protein is not the same as its native conformation, contact denaturation can occur (Gerstner et al., 1994). Hence, affinity methods might make it possible to purify proteins that would be difficult or even impossible to obtain by traditional techniques.

Short-peptide epitope, like Flag-tag or c-myc-tag, can be purified on immobilized-antibody affinity chromatography, and these monoclonal antibodies are also utilized in identification of fusion proteins during expression and purification. The octapeptide Strep-tag II, as well as biotin-acceptor peptide which allow *in vivo* biotinylation of fusion proteins can be separated in affinity column immobilizing avidin and Strep-Tactin, respectively (Waugh, 2005). All of these tags exhibit high degree of specificity for their binding partners. However, the resins (immobilized

proteins) that they interact with tend to be expensive, are easily fouled and have relatively low binding capacities. By contrast, large protein tags (e.g. GST, MBP) usually recognize small ligands that make for less expensive and more robust chromatography matrices. Disadvantage of large affinity tags, as mentioned earlier, is that they devour more metabolic energy during overproduction than small tags (Waugh, 2005).

The hexa-histidine tag (His_6-tag), short peptide consisting of six histidine residues, is one of the most commonly used affinity tag. Immobilized metal-affinity chromatography or IMAC (Porath et al., 1975), the technique used in purifying His_6-tag fusion proteins, is based on the interaction between transition metal ion (Ni2+) immobilized on a matrix and a specific amino acid side chain (histidine's immidazole ring). The candidate matrix, nitrilotriacetic acid or NTA (Hochuli et al., 1987), form a quadridentate chelation to Ni2+ and the two remaining valencies of Ni2+ further form coordinate bond with hexa-histidine. His_6-tag combines the advantages of small size with the added benefit of interacting with a chromatographic matrix (e.g. Ni-NTA resin) that is relatively inexpensive and exhibits a high binding capacity. However, like other small affinity tags, His-tag may not enhance the solubility and stability of its passengers (Waugh, 2005).

1.2.5 Biotin carboxyl carrier protein (BCCP)

The biotin–avidin/streptavidin system is used in numerous biotechnological and diagnostic applications, primarily due to the high affinity of the biotin molecule to the proteins avidin and streptavidin (i.e. $K_d = 10^{-15}$ M) (Green, 1975). Although the bacterial-originated streptavidin is distinguished from its

counterpart egg-white avidin by the lack of carbohydrate side chains that present in avidin and the much lower isoelectric point (5-6 and 10 for avidin), both of them are found tetramer in nature and bind 4 moles of biotin per mole of protein (Chaiet and Wolf, 1964). The biotin can be attached to the surface of a protein of choice and the biotinylated protein can then be bound to the avidin/streptavidin, for example, for labeling, immobilization, detection or purification purposes (Wilchek and Bayer, 1990). The biotinylation of proteins is conventionally done by chemical methods in which an activated biotin derivative is conjugated to protein surface residues or carbohydrate moieties (Diamandis and Christopoulos, 1991). Typically, a protein contains several potential target residues for the biotinylation reagent. Thus, the normal outcome of the biotinylation process is a set of protein variants that vary in the number and the location of the biotins attached. Some positions in proteins can be sensitive to this type of modification, for example, biotinylation of the residues in the binding site of antibodies can alter their binding properties, or the labeling of other residues can disrupt the protein structure.

Some natural proteins are post-translationally modified by a formation of a covalent linkage between the biotin moiety and the specific amino acid of the proteins that act as a biotin transporter. Only one such protein that biotinylated in *E. coli* is the biotin carboxyl carrier protein (BCCP), subunit of enzyme acetyl-CoA carboxylase (ACC). *E. coli* ACC catalyzes the first committed, rate-limiting step of fatty acid biosynthesis, the synthesis of malonyl-CoA (Fall and Vagelos, 1972). In this reaction, the biotinylated BCCP serves as a carrier of the carboxyl group which is delivered to the acetyl-CoA substrate. BCCP is specifically biotinylated by BirA biotin ligase (Shenoy and Wood, 1988), the product of the chromosomal *birA* gene, in

a two-step process. The first step is the adenylation of biotin, resulting in the formation of the reactive intermediate biotinyl-AMP, followed by the formation of an amide linkage between the carboxyl group of biotin and the ε-NH_2 group of a specific lysine residue in the biotinylation motif "Met-Lys-Met" (Chapman-Smith and Cronan, 1999) (Figure 1.3). This lysine residue is located at ~35 residues from the C-terminus of the BCCP protein, in an independent domain of about 80 amino acids (residues 77-156).

Full-length BCCP has an additional N-terminal region of 70 to 80 residues, presumed to be the dimerization and intersubunit interaction domain for assembly of the functional acetyl-CoA carboxylase (Li and Cronan, 1992). Structural analysis of which truncated forms of BCCP biotinyl domain are able to biotinylated indicates that a minimum of 35–40 residues on either side of the biotin attachment site is necessary to specify biotinylation (Stolz et al., 1998). Further truncation, which removes residues that contribute to the formation of the hydrophobic core of the folded structure, abolishes biotinylation of the protein (Athappilly and Hendrickson, 1995).

The three-dimensional structures of the apo and biotinylated (holo) forms of the biotin domain of BCCP from *E. coli*, which comprises the C-terminal half of the protein, have been determined by NMR (Yao et al., 1997) and X-ray crystallography (Athappilly and Hendrickson, 1995). The domain is a barrel structure consisting of two antiparallel β-sheets each containing four strands, with the N- and C-termini close together at one end (Figure 1.4). This domain can be fused to recombinant proteins and it can be biotinylated *in vivo* by the endogenous biotin ligase of *E. coli* (Chapman-Smith and Cronan, 1999).

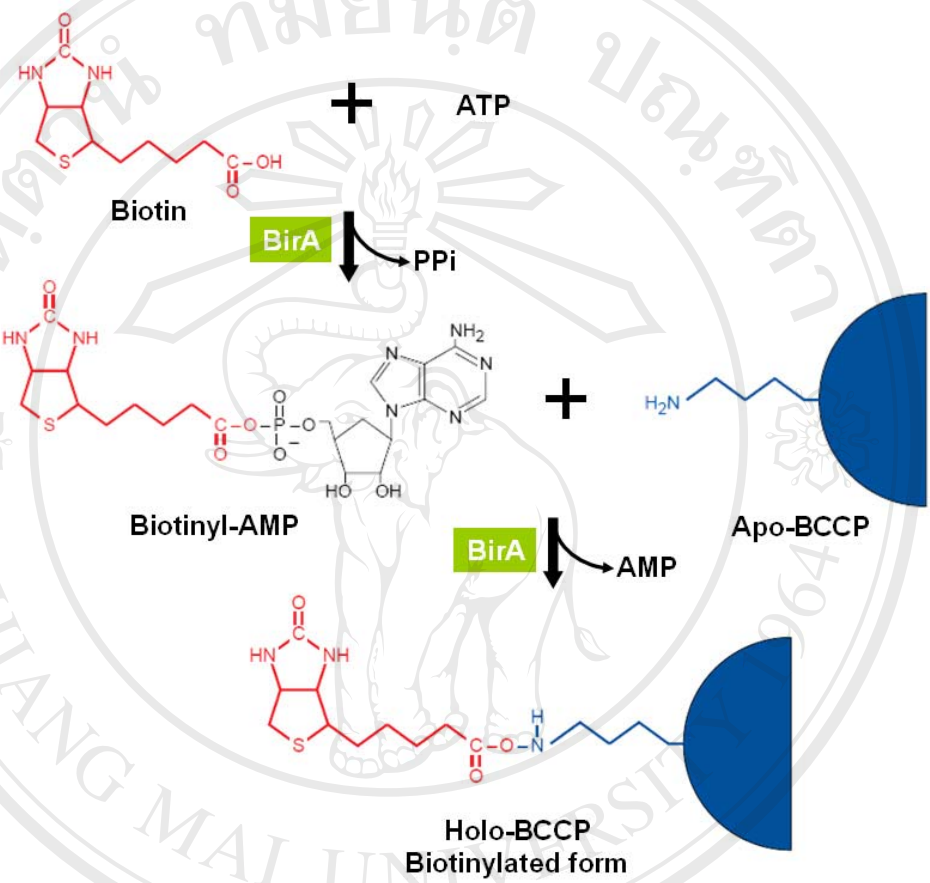


Figure 1.3 Enzymatic biotinylation reaction carried out by *E. coli* biotin protein ligase (BirA) (Chapman-Smith and Cronan, 1999).

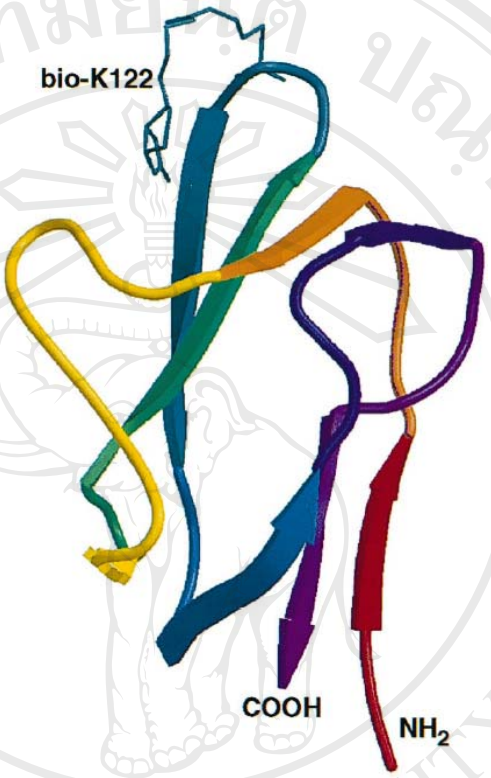


Figure 1.4 Structure of the biotin domain of the *E coli* BCCP. The overall fold of the holoprotein and the biotin prosthetic group attached to lysine (residue 122) are demonstrated (Chapman-Smith and Cronan, 1999).

Using affinity selection of peptide libraries, Schatz generated a short artificial peptide substrate (13 residues), which is sufficient to specify biotinylation by BirA biotin ligase (Schatz, 1993). This peptide has been used for the enzymatic biotinylation of various recombinant proteins both as an N- and a C-terminal fusion (Tsao et al., 1996; Tucker and Grisshammer, 1996; Smith et al., 1998). Kinetic analysis of this peptide indicates that the biotinylation kinetics is very similar to those measured for the natural substrate (Beckett et al., 1999). Although this peptide appears to be unstructured in solution, it might adopt a transient fold that mimics the conformation of a region in the biotin domain that is crucial for recognition by the BirA enzyme.

The BCCP biotinyl domain fusion approach has previously been used for site-specific biotin labeling of various recombinant proteins (Reed and Cronan, 1991; Min et al., 1999; Sibler et al., 1999; Santala and Lamminmaki, 2004). The attached biotin molecule can be used to purify the proteins on avidin resin by affinity chromatography (de Dios Alche and Dickinson, 1998). In addition, the single biotin borne by each protein can be used to tether it to avidin or streptavidin coated surfaces for various purposes, such as the construction of a protein affinity column to capture specific antibody (de Dios Alche and Dickinson, 1998), or immobilization of the proteins for affinity selection with phage-displayed combinatorial libraries (Scholle et al., 2004). The chief advantage of this approach is that, unlike chemical reagents, enzymatic biotinylation assures that all molecules will be immobilized in a uniform, bioactive orientation.

1.2.6 The human CD147

Human leukocyte surface molecule CD147 is a type I transmembrane glycoprotein with a molecular weight of 50-60 kDa. This molecule is also known as human basigin (Miyauchi et al., 1991), leukocytes activation-associated M6 antigen (Kasinrerk et al., 1992) and extracellular matrix metalloproteinase inducer (EMMPRIN) (Biswas et al., 1995). Homologues of CD147 in other species have also been discovered, such as basigin or gp42 in mouse (Miyauchi et al., 1991), OX47 in rat (Watson et al., 1992; Nehme et al., 1995) and 5A11, HT7 or neurothelin in chicken (Seulberger et al., 1992; Fadool and Linser, 1993). It was designated as cluster of differentiation (CD) system in 1997 at the Sixth International Workshop and Conference of Human Leukocyte Differentiation Antigens (HLDA workshop) (Stockinger et al., 1997). CD147 belongs to the immunoglobulin superfamily and contains two Ig-like domains in the N-terminal extracellular region, a single transmembrane domain and a short cytoplasmic domain (Figure 1.5) (Miyauchi et al., 1991; Kasinrerk et al., 1992; Biswas et al., 1995). The extracellular region consists of three Asn-linked oligosaccharides which contribute to almost half of the size of the mature molecule (Kanekura et al., 1991). The transmembrane region shows a very high degree of species conservation, suggesting it may have an important biological function superior to membrane anchorage (Kasinrerk et al., 1992). The presence of the charged amino acid, glutamic acid, in the middle of this region is a common characteristic of proteins with multiple transmembrane domain, but unusual for single transmembrane proteins (Muramatsu and Miyauchi, 2003), implying that CD147 may asscociate with other transmembrane proteins (Kasinrerk et al., 1992).

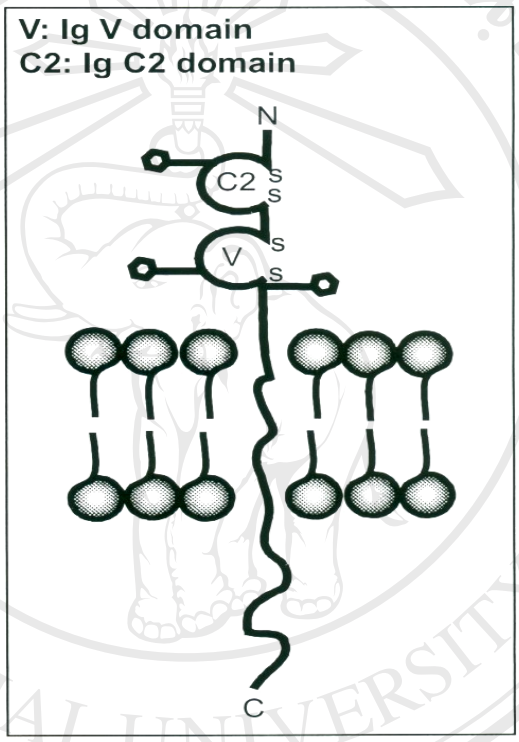


Figure 1.5 Schematic represents the structure of CD147.

CD147 widely expresses on many cell types, such as activated T cells (Kasinrerk et al., 1992), differentiated macrophage (Major et al., 2002), retinal pigment epithelium (Marmorstein et al., 1998), and human keratinocytes (Biswas et al., 1995). Highly expression of CD147 is also observed on human tumor cells and seems to be responsible for stimulating fibroblast matrix metalloproteinases (MMPs) production, leading to extracellular matrix destruction and tumor metastasis (Biswas et al., 1995). The degree of MMP expression by stromal fibroblasts in a wide range of tumors has been shown to be correlated with CD147 expression level (Caudroy et al., 1999; Dalberg et al., 2000; Thorns et al., 2002), suggesting a key role of CD147 in the induction of MMPs at the site of tumor-stroma interaction. CD147 appears to involve in regulation of T cell functions. The first indication for its T cell regulatory roles is that CD147 is strongly up-regulated on T cells upon activation (Kasinrerk et al., 1992; Koch et al., 1999). Triggering of CD147 molecules on T cells by the inhibitory mAbs resulted in modulation of lipid rafts, which associated with impaired signaling for IL-2 receptor α-chain CD25 expression (Staffler et al., 2003). Further more, certain CD147 mAbs induced homotypic cell aggregation of monocytic cell line U937 by LFA-1/ICAM-1 dependent pathway (Kasinrerk et al., 1999). This activation depended upon the activation of protein kinase and reorganization of cytoskeleton (Khunkeawla et al., 2001).

1.2.7 Immune responses

The term *immunity* is derived from the Latin word *immunitas*, which meant protection from disease and, more specifically, infectious disease (Abbas and Lichtman, 2003). The cells and molecules responsible for immunity constitute the

immune system, and their response to the introduction of foreign substances is called the immune response. Historically, the physiologic function of the immune system was defense against infectious microbes, such as bacteria, viruses, fungi and parasites. However, even noninfectious foreign substances can elicit immune response. Therefore, a more inclusive definition of immunity is a reaction to foreign substances, including microbes, as well as to macromolecules such as proteins and polysaccharides (Abbas and Lichtman, 2003).

Two systems contribute to mammalian immunity, the innate and adaptive immune systems. Innate immunity (or native immunity) is the primitive system of host defense that can be activated rapidly to infection. The principal components of innate immunity include a) physical and chemical barrier such as skin and mucous membrane, gastric pH and lysozyme in mucosal secretion; b) phagocytic cells such as neutrophils, macrophages and natural killer (NK) cells; and c) blood proteins including members of complement system, inflammatory mediators and cytokines (Medzhitov and Janeway, 2000; Abbas and Lichtman, 2003). The mechanisms of innate immunity are specific for structures that are common to group of related microbes and may not distinguish fine difference between foreign substances. It provides the early lines of defense against microbes.

In contrast to innate immunity, there are other immune responses that are stimulated by exposure to infectious agents and adapt to the infection, it is called adaptive immunity (or acquired immunity). The defining characteristics of adaptive immunity are exquisite specificity for distinct molecules and an ability to remember and respond more vigorously to repeated exposures to the same microbes (Abbas and Lichtman, 2003). There are two types of adaptive responses, humoral-mediated

immunity (HMI) and cell-mediated immunity (CMI), which are mediated by different components of the immune system. HMI function as defense for extracellular microbes and can prevent the spread of intracellular infection. The mechanism is mediated by secreted molecules called antibodies that are produced by B lymphocytes (or B cells). Antibodies recognize microbial antigens, neutralize the infectivity and assist elimination of those microbes by various effecter mechanisms such as complement activation, opsonization, and antibody-dependent cellular cytotoxicity or ADCC (Abbas and Lichtman, 2003).

CMI is mediated by T lymphocytes, or T cells, which recognize the antigen of intracellular microbes and function to destroy these microbes or the infected cells. T cells consist of functionally distinct populations, helper T cells and cytotoxic T lymphocytes (CTLs). Once activated, helper T cells secrete cytokines that function to stimulate the proliferation and differentiation of other leukocytes, leading to enhanced immunological response. CTLs kill cells that produce foreign antigens, such as cells infected by intracellular pathogens as well as tumor cells.

1.2.8 Mechanisms of adaptive immune responses

The principal feature of adaptive immune response is the use of antigen-specific receptors on T or B cells to drive effector responses in two stages. First, the antigens are presented to and recognized by antigen-specific T or B cells in the lymphoid tissues, leading to cell activation and differentiation (Brodsky and Guagliardi, 1991). Second, effector mechanisms occur either the activated T cells leave the lymphoid tissues and direct to the infected sites, or activated B cells (or plasma cells) secrete antibodies which are then released into blood and tissue fluids.

1.2.8.1 CMI response

T cells respond to peptide fragments of antigens that are displayed by antigen-presenting cells (APCs) such as macrophage, dendritic cells, or B cells. At this phase, T cells recognize self major histocompatibility complex (MHC) molecules, as well as peptide antigens mounted on these MHC molecules, which are presented on the APCs surface. The receptor that recognizes these MHC-peptide complexes is called the T cell receptor (TCR), which is responsible for antigen specificity. Antigen recognition via TCR is necessary for initiation of cell activation. However, additional signal provided by co-stimulatory molecules that are expressed on APCs is required for complete T cell activation. Co-stimulators, such as B7, deliver signal to T cells via accessory molecules CD28 and function in concert with signal from TCR to fully activate the cells (Guerder et al., 1994). Other accessory molecules, such as LFA-1 and CD2, function to increase the avidity of cell-cell interaction, allowing the TCR to be engaged by antigens long enough to transduce the necessary signals (Hogg and Landis, 1993).

Helper T cells can be divided into 2 different types based on the cytokines they produce. The T_H1 cells have primarily function to promote CMI responses by producing cytokines, such as IFN-γ, IL-2 and TNF-β. Conversely, T_H2 cells secreted IL-4, IL-5 and IL-10, which promote antibody production (Abbas and Lichtman, 2003). In contrast, CTLs recognize peptide antigens of microbes residing in the cytoplasm of infected cells. Activation of CTLs results in the release of granule contents that contain perforin and granzymes, which kill the target cells by inducing apoptosis (Henkart, 1994).

1.2.8.2 HMI response

Activation of B cells is initiated by specific recognition of antigens by the surface immunoglobulin (Ig) receptors of the cells. Binding of antigens to B cell receptors deliver initial signal of the activation process. The antigens are then internalized into endosomal vesicles and processed into peptides that are presented in complex with MHC molecules on the B cell surface for recognition by helper T cells. Signal from helper T cells, such as cytokines, further stimulates the proliferation and differentiation of the specific B cell clone. Progeny of the clone may produce IgM antibody or other Ig isotypes, may undergo affinity maturation, or may persist as memory cells (Abbas and Lichtman, 2003).

The effector functions of antibodies include neutralization of microbial infectivity and their toxins, opsonization of antigens for phagocytosis by phagocytes, activation of complement classical pathway, and elimination of infected cells by ADCC that mediated by NK cells and macrophages.

1.2.8.3 Roles of antigen-presenting cells in immune responses

Whereas the B cell receptors often directly recognize a foreign antigen, the TCR recognizes an antigen that has been processed. The cells that serve this role are called APCs. These cells aid two important functions in activation of adaptive imunity. First, APCs endocytose protein antigens and convert to peptides. An appropriate peptide forms a complex with class II MHC molecule, which is expressed at the cell surface for recognition by $CD4^+$ T cells. But if the protein antigens are endogenously synthesized in the cytosol of APCs, they are degraded into peptides that may associate with class I MHC molecule, which is recognized by $CD8^+$ T cells.

Second, some APCs provide co-stimulation to the T cells beyond those initiated by recognition of peptide-MHC complexes by TCR. Without this second signal, a naïve T cell is not activated and becomes anergy (Abbas and Lichtman, 2003).

Cells that express high levels of MHC and co-stimulatory molecules are often referred to as professional APCs. While macrophages and B cells efficiently present antigens to differentiated helper T cells in the effector phase of CMI and HMI responses, respectively, the APCs that are most potent at activating naïve T cells at initial phase are dendritic cells (Abbas and Lichtman, 2003; Plotkin and Orenstein, 2004). Immature dendritic cells are present in epithelial of the skin, gastrointestinal and respiratory tracts, capturing microbial antigens that enter from the external environment. The protein antigens are endocytosed and processed into peptides capable of binding to MHC molecules. The microbes, in combination with inflammatory cytokines, trigger dendritic cell maturation. The cells begin to express CCR7 chemokine receptor and migrate to T cell zone of draining lymph nodes, where the corresponding chemokines occur (Sato et al., 2001). Antigenic peptide-bearing dendritic cells mature during this migration by up-regulating the class II MHC molecules with bound peptides as well as co-stimulatory molecules. Once in the T cell-rich area of the nodes, naïve T cells that are specific for the display peptide-MHC complexes are activated, and an immune response is initiated.

1.2.9 Immunological adjuvants

To induce an immunological response to a protein antigen in a vaccine or experimentally, the antigen may be administered with substance called adjuvants. Immunological adjuvants are agents usually incorporated into vaccine formulation to

enhance the immunogenicity of the vaccine antigens. The immunogenicity of a vaccine is defined as its ability to evoke an immune response in the vaccinated individual. Classification of adjuvants, based on physical and chemical properties, is shown in table 1.2 (Plotkin and Orenstein, 2004).

Adjuvants promote T cell activation by several mechanisms. The original mechanism of action attributed to adjuvants was so-called depot effect, in which the rapidly clearance of soluble antigens from the injection sites can be diminished using mineral salts or emulsion-based adjuvants. Upon administration, the aluminum-containing adjuvants and the adsorbed antigens remain at the site of injection, induce inflammation, and stimulate the influx of APCs to sites of antigen exposure (Gupta et al., 1995). Adjuvants can also induce the production of various cytokines and chemokines, which then act on helper T cell subsets to modulate immune responses. One of such adjuvants is the immunostimulatory sequences in bacterial DNA (Plotkin and Orenstein, 2004). These are short sequences of DNA containing unmethylated cytosine and guanine dinucleotides, so-called CpG motifs. They trigger B cell activation and induce cytokine secretion that result in T_H1 response (Weeratna et al., 2000). Another adjuvant that affects cytokine production is MPL® adjuvant (Corixa, Inc.). MPL, or monophosphoryl lipid A, is a derivative of lipid A portion of gram-negative bacteria's lipopolysaccharide (LPS) endotoxin. It retains the immunomodulating activity of LPS but exhibits greatly reduced toxicity (Plotkin and Orenstein, 2004). MPL activates APCs, leading to the release of a cascade of cytokines including IL-1, IL-12, TNF-α, and GM-CSF (De Becker et al., 2000). MPL also stimulates the production of IL-2 and IFN-γ, leading to the CMI responses (Salkowski et al., 1997).

Table 1.2 Classification of immunologic adjuvants (Plotkin and Orenstein, 2004)

Types of adjuvants	Examples
Mineral salt	Aluminum hydroxide/phosphate (alum) Calcium phosphate
Microbial	Muramyl dipeptide (MDP) Monophosphoryl lipid A (endotoxin adjuvant)
Particulate	Biodegradable polymer microspheres Immune-stimulating complexes (ISCOMs)
Oil-emulsion	Freund's adjuvant MF59 (microfluidized emulsion)
Synthetic	Murabutide (muramyl dipeptide derivative) Polyphosphazene (PCPP)

The immune stimulating complexes, so-called iscoms, are the 40-nm cage-like, uniform stable particulate complexes consisting of cholesterol, phospholipids, saponin, and antigens (Morein and Bengtsson, 1999). Iscom-based vaccines have been reported to enhance the antigen targeting, uptake, and activity of APCs. The most likely explanation is that the antigens are organized in multiple copies into small particles and, because of their particulate nature, efficiently targeted to and taken up by the APCs simultaneously with the adjuvant-active saponin (Morein and Bengtsson, 1999; Sanders et al., 2005). In animal models and *in vitro* studies, influenza virus protein-containing iscoms have been shown to be taken up with high efficiency by APCs (Villacres et al., 1998), stimulate the secretion of IL-1 and IL-12 (Villacres-Eriksson et al., 1997), and increase the expression of class II MHC molecules by APCs (Watson et al., 1992). In clinical studies, iscom-based vaccines have been tested with a number of antigens including influenza virus proteins, or purified recombinant proteins such as HPV E6E7 fusion protein and HCV core protein (Sambhara et al., 2001; Frazer et al., 2004). In these studies, regardless of the antigens, antibody responses were detected in all of the subjects that received iscom-based vaccines.

Other particulate adjuvants, such as liposomes and microspheres, can also target antigens to APCs. Phagocytosis by macrophages of liposomes containing encapsulated antigens and lipid A has been shown to enhance antigen presentation by APCs and induce more potent humoral responses in mice when compare with lipid A alone (Verma et al., 1992). Biodegradable microspheres, such as poly(lactic-co-glycolic acid) or PLGA, were demonstrated to have adjuvant activity via their depot effect, the ability to target antigens to APCs, and their localization in lymph nodes

(Singh and O'Hagan, 1999; Peyre et al., 2004). The capacity of inducing long-lasting antibody responses was reported in numerous animal models following a single dose of PLGA microparticle-encapsulated antigens (Gupta et al., 1998; Johansen et al., 2000).

1.3 Objectives

1. To produce the recombinant CD147Ex-BCCP fusion protein with biotinylation in *E. coli* strain Origami B.
2. To develop a method for preparing biotinylated CD147Ex-BCCP as an immunogen to produce anti-CD147 polyclonal antibody by utilizing the streptavidin-coated magnetic bead.
3. To evaluate the capability in inducing the specific antibody production of the immunogen preparation strategy.