CHAPTER 1

Introduction

1.1 Statement and significance of the problems

Alpha-thalassemia syndromes are the major health problem in the Mediterranean countries, the Middle East, Southern China, India and Southeast Asia including Thailand [1]. Population of northern Thailand is one of the highest frequencies of α-thalassemias in the world leading to the spread of the severe α-thalassemia diseases including HbH disease and Hb Bart's hydrops fetalis [2]. The more severe form of α-thalassemias, HbH disease, suffers from anemia of varying severity marked by hypochromia and microcytosis while the most severe form of α-thalassemias, Hb Bart's hydrops fetalis, dies either *in utero* or soon after birth [3]. One way to prevent the spread of the severe α-thalassemia diseases is to identify genotypes of α-thalassemia and subsequent genetic counseling. In this study, therefore, the common genotypes of α-thalassemia in northern Thai volunteers were identified by the standard methods including DNA mapping, gap-PCR and PCR-RFLP [4-7].

However, when gap-PCR and PCR-RFLP were implemented for the determination of α-thalassemia genotypes, some limitation of the techniques had been encountered. Gap-PCR and PCR-RFLP required sophisticated equipments, high cost of reagents, well trained technician and tedious laboratory intervention, therefore, these standard methods were not suitable for screening of large population. A simple and practical method such as antigen-antibody based assay need to be developed as an alternative method for the determination of α-thalassemia genotypes.

The antibody involved in antigen-antibody based assay was developed for the detection of a protein marker in blood, namely Hb Bart's. There is correlation between the quantity of Hb Bart's and α-thalassemia genotypes [8-11]. Methods for the detection of Hb Bart's using rabbit polyclonal antibody were previously reported [12-14]. However,

the low specificity and batch to batch variation of polyclonal antibody limited their applications. To increase the specificity and sensitivity, monoclonal antibody highly specific to Hb Bart's was firstly established by our research group [15]. Methods for detection of Hb Bart's in blood using monoclonal antibody were also reported [16-19].

Although, monoclonal antibodies show high binding specificity to Hb Bart's, there is disadvantage of hybridoma secreting specific monoclonal antibody that might be gradually lost of the synthesis and secretion of monoclonal antibody during long-term cultivation [20]. This disadvantage could be solved by the recombinant antibody which produced in bacterial cells. The most common recombinant antibody is single-chain variable fragment (scFv) which consists of the variable domains of immunoglobulin heavy (V_H) and light (V_L) chains connected via a short peptide linker [21]. Recombinant scFv anti-Hb Bart's might be used as an alternative to monoclonal antibody for α-thalassemia diagnosis.

In this study, therefore, I tried to produce a novel recombinant scFv antibody against Hb Bart's and characterized the specificity, sensitivity and affinity constant of the produced scFv antibody.

1.2 Objectives

1.2.1 To determine the α-thalassemia genotypes in volunteers from Maharaj Nakorn Chiang Mai Hospital.

1.2.2 To produce the novel recombinant scFv antibody derived from a mouse hybridoma producing monoclonal antibody highly specific to Hb Bart's.

1.2.3 To characterize the specificity, sensitivity and affinity constant of this recombinant scFv antibody.

1.3 Literature review

1.3.1 Alpha-thalassemia

Alpha-thalassemia is genetic disorders caused by deletion or inactivation of α-globin genes in chromosome 16. A normal α-globin genotype can be represented as αα/αα. When both α-globin genes on the same chromosome are defected, the condition is called α^0-thalassemias. While, the term of

α^+-thalassemias result from one of α-globin genes on the same chromosome is defected. The two most common α^0-thalassemias are $--^{SEA}$ and $--^{MED}$, whereas, the two most common α^+-thalassemias are $-\alpha^{3.7}$ and $-\alpha^{4.2}$ types [3].

The most severe form of α-thalassemias is homozygous α^0-thalassemia (--/--) or Hb Bart's hydrops fetalis who die *in utero* or soon after birth due to the highest affinity O_2 binding of Hb Bart's results in hypoxia. The moderate severe form is HbH disease (--/-α or $--/\alpha^T\alpha$) who suffer from anemia of varying severity marked by hypochromia and microcytosis [3].

The carrier states including heterozygous α^0-thalassemia (--/αα) and homozygous α^+-thalassemia (-α/-α) show significant anemia, while heterozygous α^+-thalassemia (-α/αα) represent normal blood picture [3]. The frequency of α^0-thalassemia carrier have been reported sporadically in the Middle East and India while found highest frequency in Southeast Asia as more than 10%. The α^+-thalassemia carrier occurs at a much higher frequency right across the tropical belt. The carrier's frequency of α^+-thalassemia ranged from 2-70% [22], for example, in India varied from 3.84-18% depended on the regions [23-25], south China was 4% [26], in certain parts of Africa and Polynesia was up to 30% [3]. Interestingly, the carrier's frequency of southern Nepal and northern coast of Papua New Guinea were 80% or more [22].

In addition to these deletional α-thalassemias, the less frequency nondeletional α-thalassemias are caused by point mutations ($\alpha^T\alpha$) or other small alterations in the structural genes such as HbCS, Hb Pakse′, Hb Quong Sze and Hb Adanna. The most common of all these nondeletional types is HbCS ($\alpha^{CS}\alpha$) which is also associated with the more severe form of HbH disease ($--/\alpha^{CS}\alpha$) [3]. The main area of HbCS distribution is Southeast Asia [27], especially in northeast Thailand as up to 10.6% [28].

1.3.2 Molecular defects of α-thalassemias

1) The α^0-thalassemia due to deletions of α-globin genes

There are many different length deletions have been found in patients with α^0-thalassemia as shown in Figure 1.1. All of deletions either completely or partially ($-\alpha^{5.2}$ and $-\alpha^{20.5}$) delete both α-globin genes, and therefore no α-chain synthesis is directed by these chromosome *in vivo*. The high frequencies deletions, $--^{SEA}$ and $--^{MED}$, occur in Southeast Asia and the Mediterranean Basin, respectively. Both deletions are removed both α-globin genes but spare the function ζ gene [29]. The α^0-thalassemia $--^{SEA}$ type was deleted from the region close to the third exon of the ψζ-gene to the 3′ hypervariable region whereas the α^0-thalassemia $--^{MED}$ type was deleted in approximately the same 3′ region as in $--^{SEA}$ type, but extends upstream to 5′ to *Bgl* II site in the ψζ-gene. The deletions range in size from rather small 5.2 kb to those which remove the entire cluster, $--^{FIL}$ and $--^{THAI}$ [30], and a deletion of ≤ 47 kb in a northern European family [31].

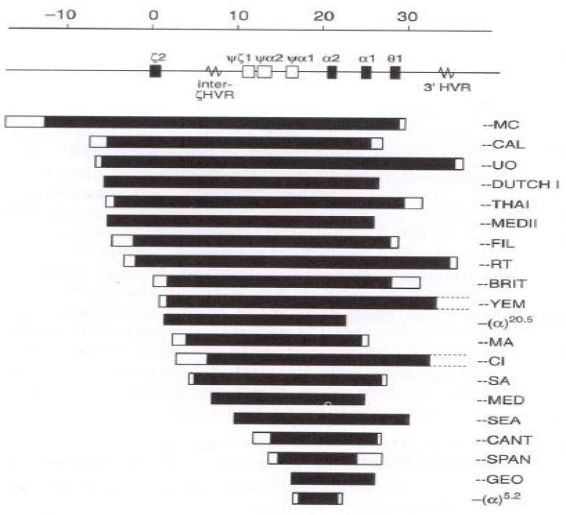

Figure 1.1 Deletions that cause α^0-thalassemia. Above: the α-gene complex is shown (scale in kb, 0 indicates the ζ2-globin mRNA CAP site). Below: the extent of each deletion is known by a black bar. Regions of uncertainly for each breakpoint are shown by white boxes [3].

2) The α^+-thalassemia resulting from deletions

The α-globin genes are embedded within two highly homologous 4 kb duplication unit, the sequence identity of which has been maintained throughout evolution by gene conversion and unequal crossover events [32-35]. These regions are divided into homologous subsegments (X, Y and Z) by non-homologous elements (I, II and III) as shown in Figure 1.2 and 1.3. Unequal crossing over between Z segments which are 3.7 kb apart, produces a chromosome with only one α-globin gene ($-\alpha^{3.7}$, rightward deletion) [36] and one with three α-globin gene ($\alpha\alpha\alpha^{anti3.7}$) [37]. There are three varieties of this type of α^+-thalassemia, depending on exactly where within the Z box crossover took place, $-\alpha^{3.7I}$, $-\alpha^{3.7II}$ and $-\alpha^{3.7III}$ [38] as shown in Figure 1.3. Non-reciprocal crossovers between homologous X boxes, which are 4.2 kb apart, also gives rise to an α^+-thalassemia ($-\alpha^{4.2}$) [36]. Three additional rare deletion that produce α^+-thalassemia have been described in Figure 1.3. The 3.5 kb deletion has been observe in two Asian Indians [39]. Another deletion, referred to as $(\alpha)-\alpha^{5.3}$, was observed in a family from Italy [40]. The third rare allele was described in a Chinese patient with HbH disease, 2.7 kb deletion [41].

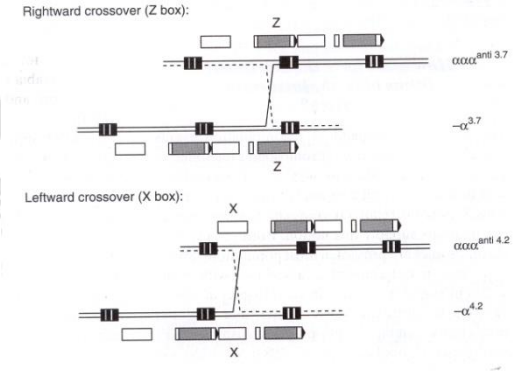

Figure 1.2 The mechanism by which the common deletions underlying α^+-thalassemia occur. Crossover between Z boxes give rise to the $-\alpha^{3.7}$ and $\alpha\alpha\alpha^{anti3.7}$ chromosomes. Crossover between misaligned X boxes give rise to the $-\alpha^{4.2}$ and $\alpha\alpha\alpha^{anti4.2}$ chromosomes [3].

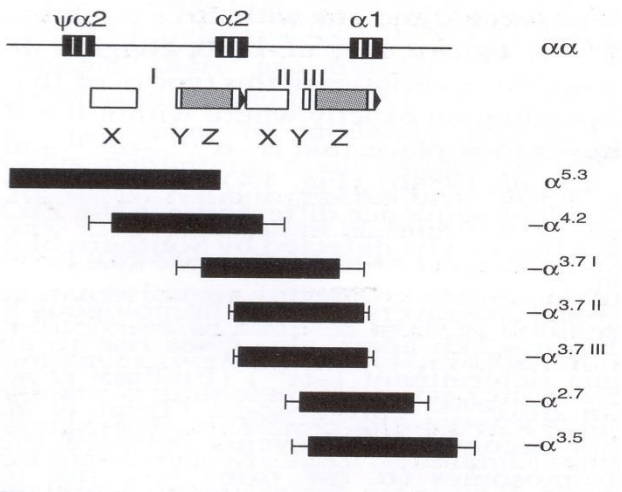

Figure 1.3 Deletions that cause α^{+}-thalassemia. Above: the α-globin genes are shown with the duplication units divided into X, Y and Z boxes with regions of non-homology (I, II, III). Below: the extent of each deletion, represent by a black bar [3].

3) α^{+}-Thalassemia resulting from nondeletions

Nondeletional types of α^{+}-thalassemia frequently occur less than deletional types. Nondeletional α^{+}-thalassemia can be classified according to mechanism; mutations that affect RNA splicing, mutations affecting the poly(A) addition signal, mutations affecting initiation of mRNA translation, chain-termination mutation, and unstable α-chain variants associated with α-thalassemia. Example of the mutations affecting RNA splicing, ($\alpha^{IVS1,5bp\ del}\alpha$), results from a pentanucleotide deletion at the 5′ donor site of IVS1 of the α2-globin genes [42, 43]. Example of the mutations affecting the poly(A) addition signal, the poly(A) sequence is required for transcriptional termination; when mutated, transcription may proceed into neighboring genes and interfere with their expression. The mutation of the α-2 gene polyadenylation signal have been described; AATAAA→AATAAG or $\alpha^{PA6:A\rightarrow G}$, was found in the Saudi Arabian population [44]. The nondeletion mutation affect mRNA translation which disrupt the initiation consensus sequence, CCRCCATG. The mutation

abolishes translation of mRNA such as, -$\alpha^{IN:A\rightarrow G}$/-α, affected person had the typical hematological features of HbH disease [45]. In chain-termination mutation, there are potentially single nucleotide variants of the natural termination codon (TAA) of the α2-globin gene. When mutations change the stop codon to one that encodes an amino acid it allows mRNA translation to continue to the next in-phase termination codon (UAA) located within the polyadenylation signal (AAUAAA), in each case extending the α-chain by 31 amino acid from the natural C-terminal arginine (codon 141). Of the six predicted α2 variants, five have been described; Hb Constant Spring (α142 Gln), Hb Icaria (α142 Lys), Hb Koya Dora (α142 Ser), Hb Seal Rock (α142 Glu) and Hb Pakse′ (α142 Tyr). Nondeletional α^{+}-thalassemia caused by unstable α-chain variants associated with α-thalassemia. Some globin variants alter the tertiary structure of the hemoglobin molecule, making the dimer (αβ) or tetramer ($\alpha_2\beta_2$) unstable. Such molecules may precipitate within the RBC, forming insoluble inclusions (Heinz bodies) which damage RBC membrane. The unstable α-globin chain variants have been shown to produce this phenotype to a greater or lesser extent as shown in Table 1.1

Table 1.1 Phenotype of unstable α-chain variants [3].

Variant	Base change	Residue/substitution	Phenotype
Hb Agrino	CD29 CTG→CCG	α29 (B10) Leu→Pro	α^{+}-thalassaemia trait
Hb Taybee	CD38/39 –ACC	α38/39 (C3/4) Thr→O	α^{+}-thalassaemia trait
Hb Torino	—	α43 (CE1) Phe→Val	Heinz-body haemolytic anaemia
Hb Hirosaki	CD43 TTC→TTG	α43 (CE1) Phe→Leu	Haemolytic anaemia
Hb Adana	CD59 GGC→GAC	α59 (E8) Gly→Asp	α^{+}-thalassaemia trait
Hb Sallanches	CD104 TGC→TAC	α104 (G11) Cys→Tyr	α^{+}-thalassaemia trait
Hb Suan Dok	CD109 CTG→CGG	α109 (G16) Leu→Arg	α^{+}-thalassaemia trait
Hb Pehah Tikva	—	α110 (G17) Ala→Asp	α^{+}-thalassaemia trait
Hb Quong Sze	CD125 CTG→CCG	α125 (H8) Leu→Pro	α^{+}-thalassaemia trait
Hb Tunis-Bizerte (Hb Utrecht)	CD129 CTG→CCG	α129 (H12) Leu→Pro	α^{+}-thalassaemia trait
Hb Sun Prairie	CD130 GCT→CCT	α130 (H13) Ala→Pro	α^{+}-thalassaemia trait
Hb Questembert	—	α131 (H14) Ser→Pro	α^{+}-thalassaemia trait
Hb Bibba	—	α131 (H19) Leu→Pro	Heinz-body haemolytic anaemia
Hb Toyama	—	α136 (H19) Leu→Arg	Heinz-body haemolytic anaemia

1.3.3 Laboratory diagnosis of α-thalassemias

1) Hb Bart's hydrops fetalis

Due to the absence of α-globin chain synthesis, the death fetus does not have either HbF or HbA. The excess γ-chain is able to form soluble homotetramer of γ-chain (γ_4 or Hb Bart's), therefore, hemoglobin electrophoresis shows large amounts of Hb Bart's (85-90%) and about 10-15% of Hb Portland [2].

2) HbH disease

HbH disease is characterized by anemia accompanied by typical thalassemic changes of the RBC in which HbH inclusions can be generated. The diagnosis should be confirmed by hemoglobin electrophoresis; HbH can be demonstrated by electrophoresis on cellulose acetate or starch gel either at alkaline pH or at neutral or acid pH. Hemoglobin electrophoresis on cellulose acetate at alkaline pH shows an abnormal band anodal to HbA comprising 1-40% of the total Hb. If one of the α-chain termination mutants such as HbCS is involved it can be demonstrated by starch gel or cellulose acetate electrophoresis, or isoelectric focusing (IEF) [3].

3) The α-thalassemia carrier states

The diagnostic problems involved in identifying the different α-thalassemia carrier states were difficult. The heterozygous states for the deletion (-α/αα) and nondeletion (α^Tα/αα) forms of α-thalassemia usually show minimal hematological changes and no abnormalities of the hemoglobin pattern. The heterozygous α^0-thalassemia (--/αα) and homozygous α^+-thalassemia (-α/-α) are characterized by significantly reduced levels of the mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH) together with hypochromic red blood cell (RBC). There are no changes in the hemoglobin pattern in adult life but in the neonatal period from 1 to 6% of Hb Bart's is found. The standard methods for identification

of α-thalassemia carriers are globin chain synthesis analysis, Southern blot analysis [5, 46] and PCR [4, 6, 7].

1.3.4 Hb Bart's

Hb Bart's is the reflect accumulation of excess γ-chains due to defective α-chains synthesis in α-thalassemia. Chemical study showed that Hb Bart's is a γ-chain tetramer (γ_4) [47, 48]. Hb Bart's was firstly described as abnormal hemoglobin in the umbilical cord blood of an infant whose parents showed evidence of thalassemia [49] and reported in a nine-month-old infant with the blood picture of thalassemia who was a patient in St. Bartholomew's Hospital, London [50].

Previous study reported the correlation between the level of Hb Bart's in blood and different genotypes of α-thalassemia as shown in Table 1.2 [51]. Hb Bart's was used as a marker for identification of the α-thalassemia states in neonatal by cellulose acetate electrophoretic method [11]. The amount of Hb Bart's in cord bloods were also quantified by elution from cellulose acetate after electrophoresis and starch-gel electrophoresis [10]. Recently, the level of Hb Bart's was quantified by capillary electrophoresis system [8, 9].

By immunoassay, the capillary tube precipitin test using rabbit polyclonal antibody was developed for identifying the α-thalassemia states in parents and offspring of α-thalassemia disease [14]. The more sensitive method using rabbit polyclonal antibody coupled with radioimmunoassay were developed for the detection of low amount of Hb Bart's in adult bloods [12, 13]. The more specific methods using monoclonal antibody for capturing trace amount of Hb Bart's in adult bloods were reported such as enzyme linked immunosorbent assay (ELISA) [16, 18], an ELISA strip [17] and sandwich-type immunochromatographic strip [19].

Table 1.2 Hb Bart's levels in cord blood of difference genotypes of α-thalassemia [51].

Genotypes of α-thalassemia	Hb Bart's (%)
Hb Bart's hydrops fetalis (--/--)	80-90
HbH disease (--/-α or --/$\alpha^T\alpha$)	24-26
α^0-Thalassemia carrier (--/αα)	5-6
α^+-Thalassemia carrier (-α/αα or $\alpha^T\alpha$/αα)	1-2

1.3.5 Recombinant antibody [52]

Antibodies are our bodies modular defense system, used to identify and attack foreign intruders. The antigen-binding sites of antibodies are located at the upper tips of the Y- or T-shaped immunoglobulin molecules. Each tip includes 6 hypervariable loops, which constitute the surface of the antigen-binding site (Figure 1.4). During the past decade, advances in molecular biology have greatly facilitated the genetic manipulation, recombinant production, identification and conjunction of antibody fragments. Furthermore, genetic fusion and recombinant expression has led to the development of a large variety of engineered antibody molecules for research, diagnosis and therapy. Antibody molecules consist of light and heavy chains, each chain composed of one variable domain and between one and four constant domains, which assemble into molecules exhibiting two or more antigen-binding sites. The antigen-binding sites of immunoglobulins are embedded into the variable heavy and light domains (V_H, V_L) and are specially separated from the effector function mediating regions located in the Fc fragment. The Fv fragment is the smallest fragment of an antibody molecule which is able to provide antigen specificity.

The development of the hybridoma technology in 1975 allowed the generation and production of monoclonal antibodies with predefined specificities [53].

Consequently, hybridoma cells provided the starting material for the first recombinant antibody molecules expressed as whole immunoglobulins, Fab, or Fv fragments in lymphoid or non-lymphoid cells [54-57]. However, initial experiments expressing antibodies in bacteria were hampered due to improper folding and aggregation of the polypeptides in the bacterial cytoplasm. These problems were solved by expressing only parts of the immunoglobulin molecule, i.e. Fv or Fab fragments. Further, it soon became evident that due to the non-covalent association of soluble Fv fragments expressed in *Escherichia coli (E. coli)*, these fragments are quite unstable [58]. This drawback has been overcome by engineering single-chain variable fragment.

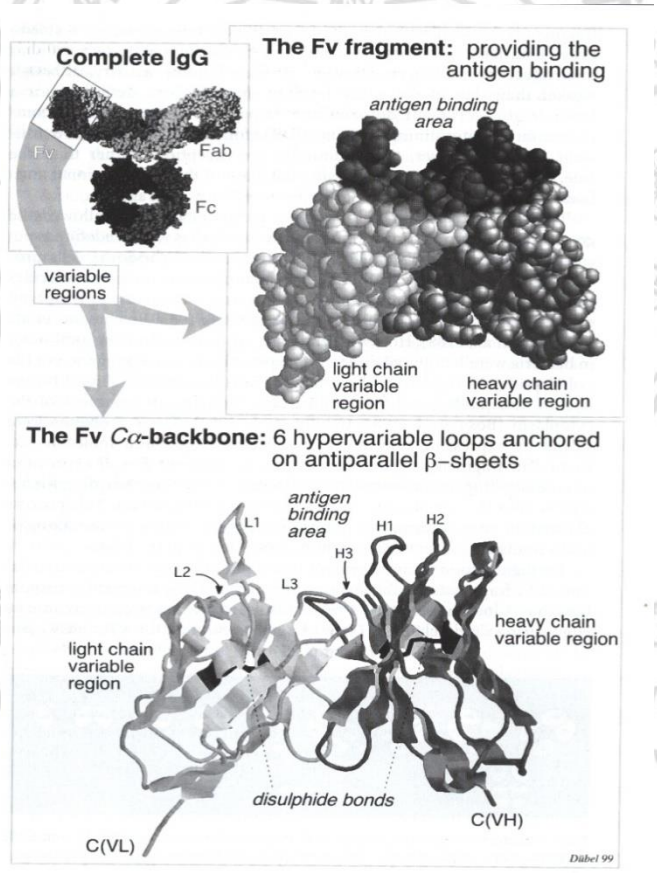

Figure 1.4 The Fv fragment contains the antigen-binding regions of antibodies. An IgG molecule (box upper left) consist of two identical Fv units held together by constant regions. The Cα-cartoon (lower box) visualizes the antiparallel beta-sheet structure of the immunoglobulin fold and the intramolecular disulphide bond in each of variable regions [52].

1.3.6 A single-chain variable fragment

A single-chain variable fragment (scFv) is a fusion protein of the variable regions of the heavy and light chains of immunoglobulins, connected with a short linker peptide of ten to about 15-20 amino acids (Figure 1.5). The connection via linker peptide increases the stability of variable domain under physiological condition [21, 59]. One of the most common used linkers is a stretch of glycine and serine residues of the format $(Gly_4Ser)_3$, but many other linker designs have been successfully employed. The linker is usually rich in glycine for flexibility, as well as serine or threonine for solubility, and can either connect the N-terminus of the V_H with the C-terminus of the V_L, or *vice versa*. This protein retains the specificity of the original immunoglobulin, despite removal of the constant regions and the introduction of the linker [60].

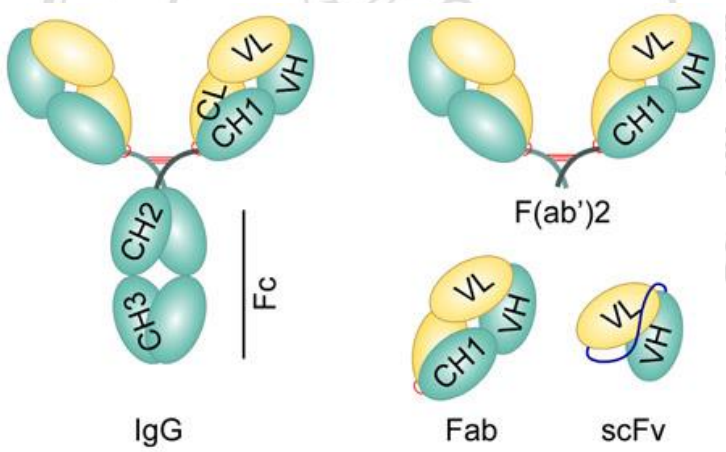

Figure 1.5 Antibody model showing subunit composition. Each oval represents an immunoglobulin folding domain. VL indicates variable domain light chain; VH, variable domain heavy chain; CL, constant domain light chain; CH, constant domain heavy chain; Fc, Fc fusion; IgG, immunoglobulin G; F(ab′)2, dimeric antigen binding fragment; Fab, antigen binding fragment; scFv, single chain antigen binding fragment [60].

The scFv antibodies have many uses such as flow cytometry, immunohistochemistry and as antigen-binding domains of artificial T cell receptors. In medical application, the scFv used as an immunoinhibitor to block the TNF-α has been recently approved by FDA to be used for the treatment of Crohn's disease [61] and rheumatoid arthritis [62]. In addition, scFv anti-ErbB2 act as anti-cancer intrabody which expressed within the cell for the treatment of ErbB2-overexpressing ovarian cancer [63]. Another study proved that anti-cyclin E scFv intrabody could inhibit the growth of breast cancer cell line [64].

Various other hosts have been employed for the expression of recombinant antibodies, including gram-positive bacteria (*Bacillus subtilis*), fungae (*Saccharomyces cerevisiae*), plant cells (*Nicotiana tabacum*), Bacculovirus infected insect cells and various mammalian cell lines (CHO, COS, HEK293) [52]. Unlike monoclonal antibodies, which are often produced in mammalian cell cultures, scFv antibodies are more often produced in bacteria cell cultures such as *E. coli* [60].

Single-chain variable fragments lack the constant Fc region found in complete antibody molecules, and, thus, the common binding sites used to purify antibodies. These fragments can often be purified or immobilized using Protein L, since Protein L interacts with the variable region of kappa light chains. More commonly, scientists incorporate a six histidine tag on the C-terminus of the scFv molecule and purify them using immobilized metal affinity chromatography (IMAC).

1.4 Principle and rationales

Previous reports of the high prevalence of α-thalassemias in northern Thailand led to the question of prevention and control of this genetic disorder in public health. If we were able to determine the common genotypes of α-thalassemia in pregnant women followed by appropriate genetic counseling, it should be an effective mean of prevention and control of this genetic disorder. The most important group among pregnant women was the first time of pregnancy (primigravidarum). When their genotypes of α-thalassemia were identified coupled with genetic counseling to the at

risk parents, it should be an effective prevention and control of α-thalassemias for their family planning.

The methods of gap-PCR and PCR-RFLP for the determination of α-thalassemia genotypes had some disadvantages. They required sophisticated equipments, high cost of reagents, well trained technician and tedious laboratory intervention. Therefore, recombinant DNA technology was introduced into this study to generate a novel recombinant scFv antibody specific to Hb Bart's.

Mouse hybridoma producing monoclonal antibody highly specific to Hb Bart's was the starting material. This novel scFv antibody should retained the same specificity, sensitivity and affinity for Hb Bart's as the parent monoclonal antibody. A high sensitivity, high specificity and simple immunoassay for the diagnosis of α-thalassemias might be developed using the produced scFv antibody. This immunoassay might be used for clinical or mass populations screening for the presence of α-thalassemias. In the future, the antibody gene construct of heavy and light chains of the scFv antibody might be used as the template for peptide synthesis to synthesize highly specific and purified scFv antibody.