In Bacteria, Yeast and Plants

Thesis submitted for the degree of
Doctor of Philosophy
at the
University of Leicester

by

Tan Chon Seng

Department of Botany, Leicester University,
Leicester LE1 7RH, U.K.

1997

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I am deeply grateful to my supervisors Dr. G.C. Whitelam and Dr. B. Cockburn for their excellent supervision, encouragement, and helpful advice throughout the period of this study.

I also wish to express my sincere gratitute to Dr. M.R.L.Owen, Dr. Z.C. Chen, and Dr. Y.Y. Liu for their technical assistance. My thanks extend to other members of the Botany Department for their guidance.

My special thanks are due to my wife Suat Ean for her moral support and patience, and my three children Xiang-Yi, Xiang-Wei and Xiang-Zhi for accompanying me through this difficult period.

Finally, I would like to acknowledge the Malaysian Agricultural Research and Development Institute for the financial support of my study.

Constructions and Expressions of anti-ABA scFv genes in Bacteria, Yeast and Plants

Chon-Seng Tan
Botany Department, Leicester University.

ABSTRACTS

Three anti-ABA single-chain Fv (scFv) antibody genes, namely ABA26, MAC61, MAC252 scFvs, had been constructed from mouse and rat hybridomas and expressed in bacteria, yeast and plants. All of these scFv genes could be expressed in *E. coli* using the T7 promoter, either targeted to the *E. coli* periplasm or cytoplasm, albeit at comparatively low levels. The cytoplasmically located scFv proteins were in the form of insoluble fraction and did not therefore exhibit any binding activities to the ABA. The majority of the periplasmically located scFv proteins were retained in the bacterial cytoplasm as insoluble bodies with the *ompA* signal peptide attached to them. Nevertheless, the positive signal in ELISA test indicated that a small portion of scFv proteins were in the form of soluble functional scFv proteins. All the three periplasmically expressed scFv proteins had specific ABA binding activities. However, the affinity constants of the scFv proteins were found to be 5 to 10 fold lower than those of their parental MAbs. In addition, the MAC61 and MAC252 scFv proteins possessed less specific binding sites compared to their parental monoclonal antibodies.

In the yeast expression, although no scFv proteins were detected in the case of ABA26 and MAC252 scFvs with the *Pichia pastoris* expression system, MAC61.scFv proteins were expressed and some of the product was secreted into the culture media. The expression of ABA26, MAC61, and MAC252 scFv genes in plants, either by transient expression using the PVX viral vector or by stable *Agrobacterium*-mediated transformation, did not result in the accumulation of scFv protein to a detectable level either targeted in cells cytosol, secreted to the apoplast or retented in the endoplasmic reticulum.

The low expression levels in *E. coli*, the difficulties encountered in the *P. pastoris* system, and the inability plants to express ABA26, MAC61, and MAC252 scFv proteins appeared to be due to the nature of those particular scFv genes. Investigation indicated that the single V_L domain of MAC61 limited the expression level in *E. coli*. The single V_L domain of MAC61 was expressed at least 10 fold lower than the single V_H domain. While a hybrid scFv gene which contained the V_L domain of MAC61 and the V_H domain of a readily expressed scFv was found to express at a much lower level than the hybrid scFv gene comprised of the V_L domain of a readily expressed scFv and the V_H domain of MAC61. Future experiments involving mutations of the amino acid sequence of the V_L domain could allow precise identification of the characteristics of the V_L domain of MAC61 which are responsible for the adverse effects on the expression in bacteria, yeast and plants.

Contents

List of Figures

List of Tables

List of Abbreviations

The standard scientific metric units and chemistry formula were used throughout the thesis. Other abbreviations were list as follows:

ABA	abscisic acid
bp	base pairs
BSA	bovine serum albumin
CaMV	Cauliflower mosaic virus
cDNA	copy deoxyribonucleic acid
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
ELISA	enzyme linked immunosorbent assay
Fig.	Figure
gfp	green fluorescent protein
hr	hour
kb	kilo base pairs
kD	kilo daltons
LB	Luria broth
lb	pound
MAb	monoclonal antibody
MBP	maltose binding protein
PAGE	polyacrylamide gel electrophoresis
pfu	plaque forming units
RNA	ribonucleic acid
sDNA	single stranded deoxyribonucleic acid
SDS	sodium dedecyl sulphate
tRNA	transfer ribonucleic acid
UV	ultraviolet
xg	times gravity

CHAPTER 1

INTRODUCTION

The antibody has been recognised as a crucial component in the immune response since the late nineteenth century. After more than a century of intensive research, from polyclonal to monoclonal, the antibody technology has proved to possess many important applications. Recent development of antibody technology based upon molecular biology and genetic engineering will extend even further the usefulness of antibody molecules. In this introduction, an overview of the antibody technology and developments in antibody engineering will be presented.

1.1 Immunology Concepts.

1.1.1 The Vertebrate Immune System.

Living organisms have developed several systems which will respond to the outside world. One of them is the **immune system** in the vertebrate. The immune system can detect foreign molecular entities which are immunogenic, transfer information between different structural elements of the system and respond by producing cells and molecules which will finally lead to the removal of the foreign matter. The function of the immune system is therefore to defend organisms against any foreign invaders.

The immune system is composed of lymphatic organs, such as the thymus, spleen, bone marrow, lymph nodes, and lymphocytes which travel as single cells throughout the body except the brain (Klein, 1982; Paul, 1984). The lymphocytes, which can be divided into B lymphocytes and T lymphocytes, are involved in the distinction between self and non-self components. To carry out their function, B and T lymphocytes synthesise certain receptor molecules which can bind to foreign molecules, correctively called **antigens**. In the case of B lymphocytes, the antigen

receptor is the **immunoglobulin** or **antibody** molecule, either as a membrane-bound molecule or as a molecule secreted by the B cell.

The immunoglobulin molecules, synthesised by B cells, can bind soluble antigen directly. Studies of different immunoglobulin molecules have shown that they are all related and are probably derived from a common ancestral gene (Hood *et al.*, 1985). It appears that the primordial gene encoded the antibody domains which were capable of interacting with itself and with other protein domains. However, the most important feature of the antibody domains is the regions which can accommodate, within its structure, a large variety of different amino acid sequences (Hood *et al.*, 1985). These variable parts of antibody domains form part of the antigen-binding site and can be generated to exhibit exquisite binding specificity. However, in order to recognise a large number of foreign antigens, the immune system has a complete mechanism to generate, within the lifetime of an individual, a large repertoire of distinct antibody molecules to capture different types of foreign molecules (Hood *et al.*, 1985).

1.1.2 The Molecular Structure of Immunoglobulin.

Immunoglobulins are a large family of glycoproteins that share key structural features. Structurally, immunoglobulins are composed of one or more copies of characteristic unit that can be visualised as forming a Y shape. Based on the number of Y-like units and the type of heavy-chain polypeptide they contain, immunoglobulins are divided into five classes, IgG, IgM, IgA, IgE, and IgD (Table 1.1). All of them share the similar structural features and key properties. For example, IgG molecules (Fig. 1.1) which contain only one structural Y unit and are the most abundant in serum, are composed of four polypeptide chains: two identical light chains (25 kD) and two identical heavy chains (55 kD) bound to each other by disulphide bridges (Kabat *et al.*, 1977). One complete light chain and the amino-terminal of a heavy chain are required for forming the arm of Y unit while the carboxyl-terminal ends of both heavy chains comprised the base of Y unit. The arms of an IgG molecule are known as **Fab** fragments (the fragments that are still able to bind to antigen after partial proteolytic

Figure 1.1 : The molecular structure of an IgG molecule.

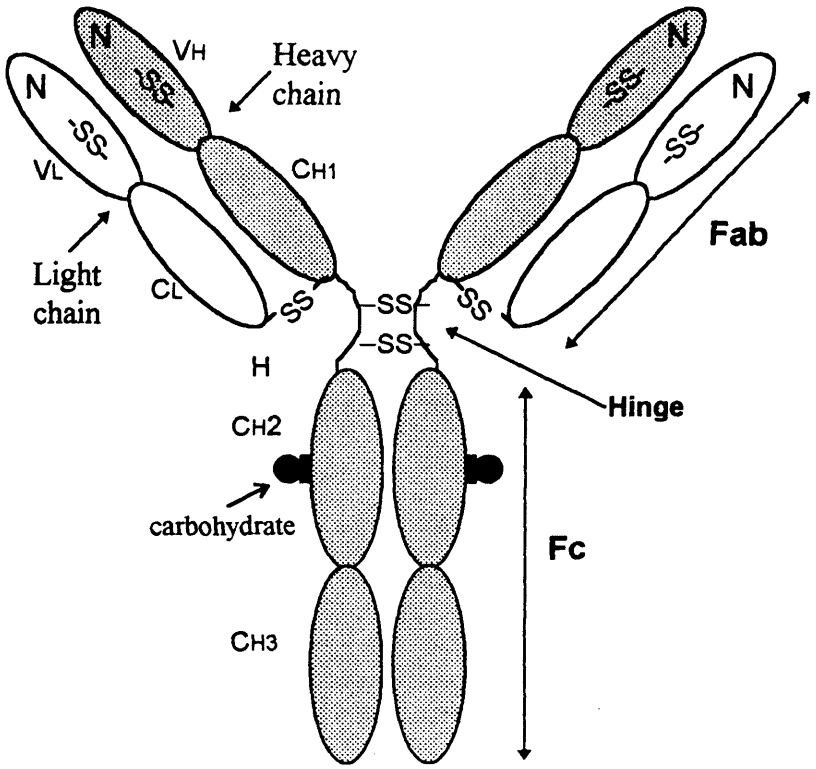
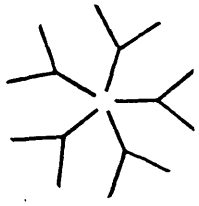
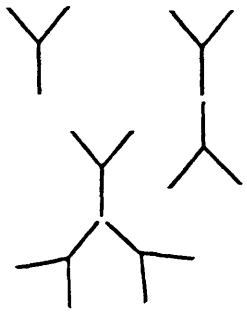


Table 1.1 : Classes of immunoglobulins.

Characteristics	IgG	IgM	IgA	IgE	IgD
Heavy chain	γ	μ	α	ε	δ
Light chain	κ or λ	κ or λ	κ or λ	κ or λ	κ or λ
Y units	1	5	1, 2, or 3	1	1
Molecular structure					

digestion). Whereas the remaining fragment, the protein domains that are involved in immune regulation, is called the **Fc** fragment (the fragment that can be crystallised).

The systematic homologies in internal amino acid sequences suggested that all IgG molecules consist of 14 domains, five domains for each heavy chain and two domains for each light chain (Kabat *et al.*, 1977). The heavy chain domains are denoted V_H, C_{H1}, H, C_{H2}, and C_{H3} where the V_H is a variable-region domain, the C_{H1}, C_{H2}, and C_{H3} are three constant-region domains, and the H is the hinge region. The light chain is composed of a variable-region domain V_L and a constant-region domain C_L. The distinction of variable and constant regions in both light and heavy chains is based on the fact that different chains of the same light chain type or the same heavy chain class show amino acid sequence variation in their amino-terminal portions whereas the carboxyl-terminal parts are constant. The structural division also reflects a functional division of the molecule: the variable regions of IgG molecules bind foreign antigens while the constant portions exert certain effector functions.

The IgG domains have a characteristic tertiary structure. The IgG molecule consists of two twisted anti-parallel ß-sheets which form a ß-sheet sandwich structure, and a minor α-helix component. The two ß-sheets of the variable domains are inclined at 30° to each other, and pinned together by a disulphide (S-S) bond between cysteine residues at conserved positions in the sequence. This characteristic folding is termed the immunoglobulin fold (Chothia *et al.*, 1985; Padlan, 1993). X-ray crystallography of IgG molecules has established that the binding site for antigen is formed by both heavy and light chain variable domains and in particular by so-called **hypervariable regions**, of which there are three in each variable domain (Wu and Kabat, 1970; Kabat and Wu, 1991). These hypervariable regions may vary in both length and sequence of amino acids and are also known as **complementarity-determining regions (CDR)**. However, the hypervariable regions are not completely random in sequence; most hypervariable regions (with the exception of heavy chain CDR3) can fit into one of a few main chain canonical structures (Chothia and Lesk, 1987, Chothia, *et al.* 1989). Studies have also shown that the hypervariable regions are flanked by **framework region (FR)** which are more conserved. The framework residues are presumably

important for the proper three-dimensional structure of variable domains (Wu and Kabat, 1970).

1.1.3 The Molecular Genetics of Immunoglobulins.

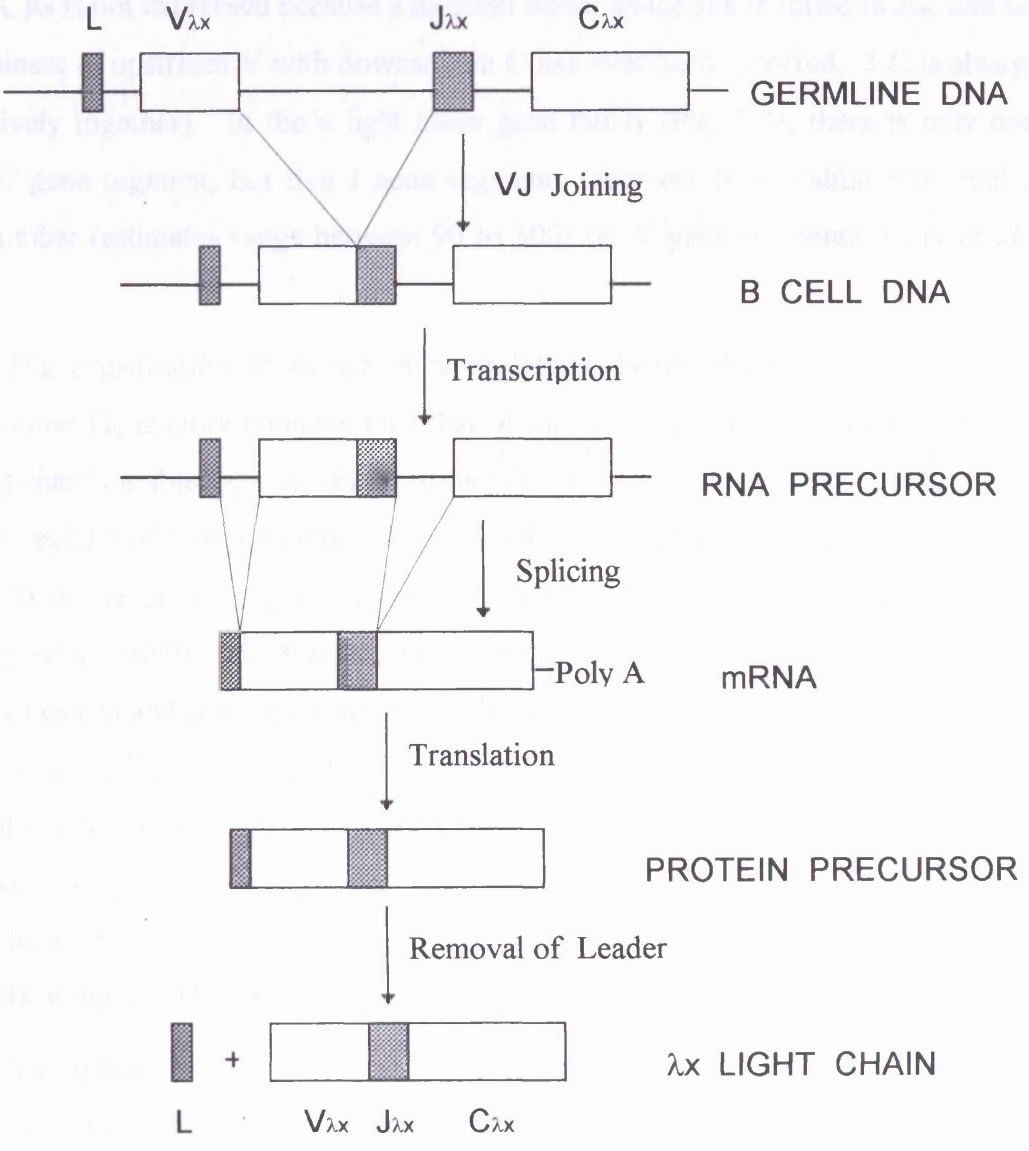
It has long been known that the immune-system is capable of generating a huge diversity of immunoglobulin molecules thereby recognising a wide range of foreign molecules. If each individual antibody molecule were to be encoded by an individual gene, the genes required to encode the entire immunoglobulin repertoire would constitute a large portion of the eukaryotic genome. In 1965, Dreyer and Bennett proposed their revolutionary hypothesis that variable and constant regions of immunoglobulin molecules were encoded by separate genes in the germline DNA. Each heavy and light chain has only a single constant-region gene whereas multiple genes encoding the variable regions are present in the germline DNA. For the production of a complete heavy and light chain polypeptide, one of the many variable-region genes would be fused to the single constant-region gene. This model, although revolutionary, turned out to be essentially correct. The fusion of the coding information for variable and constant regions occurs at the DNA level, and not at the RNA or protein level. This knowledge forms the basis for understanding the genetics of immunoglobulins.

Mouse and human immunoglobulin genes have been the most extensively studied by immunologist. Both group of genes were found to possess a similar complexity and organisation. Both mouse and human have three immunoglobulin gene families encoding three distinct families of polypeptide chains: two different families of light chains (λ and κ, a given immunoglobulin molecule has either two identical λ or two identical κ light chains) and a family of heavy chains (Leder, 1982; Honjo, 1983; Max, 1984).

The DNA rearrangement of an immunoglobulin gene occurs when the B-precursors differentiate into mature antibody-production cells. For instance, in the formation of a mouse immunoglobulin light chain (Fig. 1.2), the gene in germline DNA is split into parts called variable segments (V), J-segments (J) and constant segments

Figure 1.2 : Rearrangement of a λ light chain gene in mouse.

The V gene segment encodes the first 98 amino acids of the variable region, the J gene segment encodes the remaining 12 amino acids of the variable region and the C exon encodes the constant region of the light chain. The L exon encodes the leader or signal peptide which is required for proper intracellular transport of light chains and is cleaved off during production of the mature molecule.



(C) (Tonegawa *et al.*, 1977; Bernard *et al.*, 1978). In B-cell differentiation, a DNA rearrangement results in the fusion (joining) of V and J gene segments (Bernard *et al.*, 1978) to a functional light chain gene. After transcription and splicing, the mRNA is translated into a light chain polypeptide.

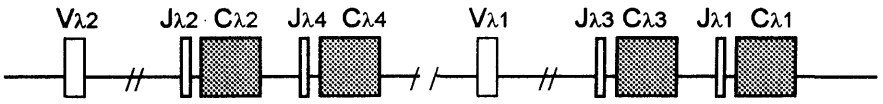
Molecular cloning has shown that mouse λ light chain contains a total of four different J-C linked exons in two gene clusters (Fig. 1.3), each with an upstream V gene segment, located on chromosome 16 (Blomberg *et al.*, 1981; Miler *et al.*, 1981). This arrangement gives the synthesis of light chain polypeptides with the arrangement of $V\lambda_1$-(J-C) λ_1, $V\lambda_1$-(J-C)λ_3 and $V\lambda_2$-(J-C)λ_2 (the arrangement of $V\lambda_2$-(J-C)λ_4 is not expressed because a mutated donor splice site is found in $J\lambda_4$, and no recombinant of upstream V with downstream C has ever been reported, J-C is always distinctively together). In the κ light chain gene family (Fig. 1.3), there is only one single C gene segment, but five J gene segments (segment J_3 is malfunction) and a large number (estimates range between 90 to 300) of V gene segments (Cory *et al.*, 1981).

The organisation of mouse immunoglobulin heavy chain genes, located on chromosome 12, is more complex than that of the light chain genes. In contrast with the light chain variable regions which are encoded by two gene segments (V and J), the variable regions of heavy chains are encoded by three gene segments, V, D, and J (where D is the diversity gene segment) (Early *et al.*, 1980; Sakano *et al.*, 1980; Schilling *et al.*, 1980). The formation of a complete heavy chain gene requires firstly joining of one D and one J gene segment followed by second fusion of a V to a (D-J) gene segment (Alt *et al.*, 1984). The V-D-J joining which involves a third gene segment would obviously increase the diversity in an important region of the antibody molecules. It has been shown that heavy chain genes contain 4 different J gene segments, at least 12 different D gene segments, and an estimated number of V gene segments of more than 100 (Shimizu *et al.*, 1982; Wood and Tonegawa, 1983).

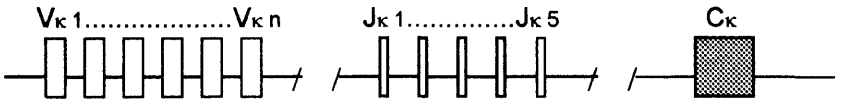
The mouse heavy chain gene family has been reported to be composed of 8 different constant (C) gene segments ($C\gamma_1$, $C\gamma_{2a}$, $C\gamma_{2b}$, $C\gamma_3$, $C\mu$, $C\alpha$, $C\varepsilon$, and $C\delta$), each corresponding to the 8 different classes of immunoglobulin, IgG_1, IgG_{2a}, IgG_{2b}, IgG_3, IgM, IgA, IgE, and IgD respectively (Shimizu *et al.*, 1982). The heavy chain genes

Figure 1.3 : Organisation of immunoglobulin gene families in mouse.

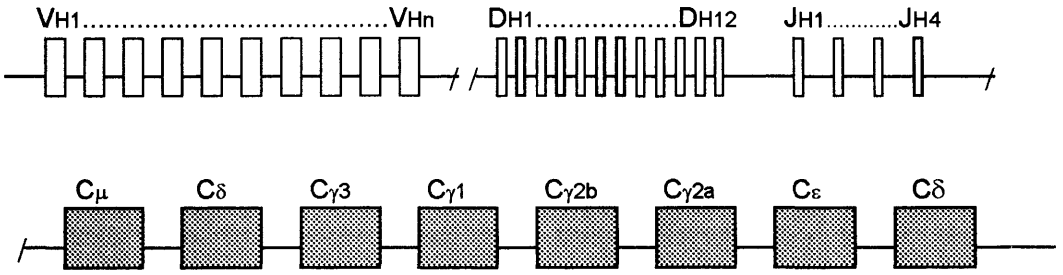
λ Light Chain : (Chromosome No. 16)



κ Light Chain : (Chromosome No. 6)



Heavy Chain : (Chromosome No. 12)



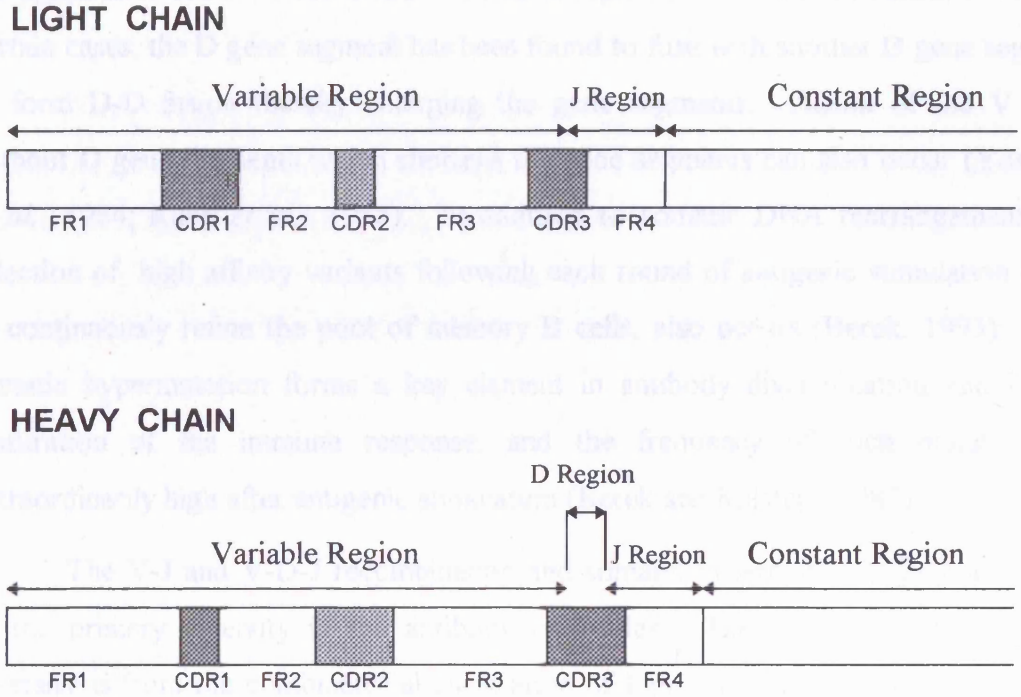
can undergo class switching in the process of B cell differentiation. At the stage of a pre-B cell, the mRNA consists of V region sequence linked to the C_μ sequence. However, heavy chain switching can occur by replacing the C_μ gene with any one of the other constant gene segments resulting in the formation of the different classes of antibody molecules.

The human immunoglobulin genes, on the other hand, have a similar organisation but contain different numbers of gene segments compared to mouse. Human immunoglobulin genes have more than six $C\lambda$ genes compared to four in mouse. Human also contains more heavy chain constant-region gene segments. However, there are only about 20-50 V_κ genes in human compared to 90-300 V_κ genes in mouse (Honjo, 1983).

1.1.4 Diversity of Antibody Molecules.

Recombination between mouse V and J, or V, D and J gene segments can generate a large number of different variable domains of light and heavy chain polypeptide sequences. In immunoglobulin light chains, the first two hypervariable loops, CDR1 and CDR2 are derived from the V gene segments. The light chain CDR3 loop is randomly drawn from the combination of V and J gene segments (Fig 1.4). This combination has limited variability, only 3 copies for λ light chains and approximately 800 copies (estimate to be 200V x 4J = 800 VJ) for κ light chains. However, in immunoglobulin heavy chains, the first two hypervariable loops CDR1 and CDR2 are still drawn from a limited number of V gene segments, but the third hypervariable loop, CDR3, is created from the combination of V, D, and J gene segments (Fig. 1.4). This loop is exceptionally variable in sequence due to the fact that insertion of D gene segments can produce functional proteins in all three reading frames (Sanz, 1991), as a result 12D gene segments can code for 36 different protein sequences. Because of these multiple genes, it is believed that there are at least 7200 (estimated half of 100V x (12D x 3) x 4J = 14400 VDJ) different possibilities of variable domains in heavy chains.

Figure 1.4 : Construction of antibody variable domains by DNA rearrangement.



Key : CDR = Complementarity-determining regions.
FR = Framework region.

Furthermore, antibody repertoires are modified through several rounds of somatic hypermutation. Besides the joining combinations of VJ or VDJ, somatic DNA rearrangement also involves other mechanisms to amplify the diversity of antibody molecules. Studies have shown that the joining at VJ or VDJ is not precise. Nucleotides can be randomly inserted and deleted at the recombination site leading to the appearance of new amino acids in a critical region of the antibody (Snaz, 1991). In certain cases, the D gene segment has been found to fuse with another D gene segment to form D-D fusion thereby enlarging the gene segments. Fusion of the V gene without D gene segments which shortens the gene segments can also occur (Yoshikai *et al.*, 1984; Rubb *et al.*, 1985). In addition to somatic DNA rearrangement, the selection of high affinity variants following each round of antigenic stimulation so as to continuously refine the pool of memory B cells, also occurs (Berek, 1993). This somatic hypermutation forms a key element in antibody diversification and in the maturation of the immune response, and the frequency of such mutations is extraordinarily high after antigenic stimulation (Berek and Milstein, 1987).

The V-J and V-D-J recombination and somatic hypermutation generate much of the primary diversity in the antibody molecules. The major way to antibody diversity is from the combinatorial association of light and heavy chains because the antigen-combining site is formed by hypervariable regions of heavy and light chain variable domains. The random or almost random association of several hundreds of distinct light chains and several thousands distinct heavy chains in a combinatorial fashion can generate several million different types of antibody molecules (Steinmetz, 1986).

1.1.5 Antibody-antigen Interactions.

The region of an antigen that interacts with an antibody is defined as an **epitope**. X-ray crystallographic studies revealed that the size of an epitope is relatively small. Only relatively few of the amino acid side chains of the six CDRs are in close contact with the antigen epitope. Under the X-ray studies, the antibody-antigen binding site can be visualised as a specific cleft or pocket into which the epitope is

docked (Berzofsky and Berkower, 1984). This make the binding between an antibody and an antigen very precise or **specific**. The remarkable specificity of antibodies has been demonstrated by a large number of experiments. For example small changes in the structure of an epitope can prevent antigen recognition and enable the discrimination between two very closely related molecules. Indeed, antibodies have been isolated that will differentiate between protein conformations and detect a single amino acid substitution in a protein antigen (Padlan *et al.*, 1995).

The binding of an antibody to an antigen is entirely dependent on noncovalent interactions, and is stabilised by the combination of weak interactions that depend on the precise alignment of the antigen and antibody (Berzofsky and Berkower, 1984). These forces can occur between side chains or the polypeptide backbones through hydrogen bonds, van der Waals forces, coulombic interactions, and hydrophobic bonds. The binding forces at antibody binding site create a microenvironment of the combining site that can accommodate highly charged, as well as highly hydrophobic molecules. Antigens can consist of proteins, carbohydrates, lipids, nucleic acids, amino acids, as well as a wide range of synthetic organic chemicals. The repertoire of possible binding sites is enormous, and antibodies that are specific to novel compounds have been derived readily.

X-ray studies have also shown that the epitopes are surface located. Studies on big antigen molecules such as lysozyme revealed that the amino acids that form the epitope actually come from different distant stretches of the primary sequence (Benjamin *et al.*, 1984). Thus, the epitopes of a protein antigen are local surface structures that can be formed by continuous or non-continuous amino acid sequences. As antibodies can recognise relatively small regions of antigens, occasionally they can find similar epitopes on the other molecules. This forms the molecular basis for **cross-reaction**.

The binding of antibody and antigen is reversible. The strength or degree of binding between an antibody and an antigen can be described thermodynamically as **affinity** and **avidity**. The affinity is the interactive force between antibody and antigen whereas the avidity is the overall stability of antibody-antigen complexes. The affinity of antibodies can be defined as the amount of antibody-antigen complex formed under

equilibrium conditions. The affinity constant can be measured as equilibrium association constant (K_{eq}) or equilibrium dissociation constant (K_D) where the K_{eq} = [Ag-Ab]/[Ag]•[Ab] and K_D = [Ag]•[Ab]/[Ag-Ab], one is the inverse of the other. The affinity constants for antibody-antigen binding vary widely and the K_D can range from 10^5 M^{-1} to above 10^{12} M^{-1} (Harlow and Lane, 1988). In practice, antibodies with high affinity, will bind larger amounts of antigen in a shorter period than low-affinity antibodies.

The avidity is the intrinsic affinity of the antibody for the epitope and is closely related with the valency of the antibody and antigen. Antibodies are multivalent, IgGs and most IgAs are bivalent, whereas IgMs are decavalent. Antigens can possess multiple epitopes either because they contain multiple copies of the same epitope or because they contain several different epitopes recognised by different antibodies. Multimeric interactions allow antibodies and antigens to bind tightly (Harlow and Lane, 1988).

The high affinity and avidity of antibodies, combined with their specificity, make the antibody molecule be a powerful tool in science, medicine and biotechnology.

1.2 Antibody Technology.

1.2.1 Polyclonal Antibody.

Vertebrate animals produce antibodies when exposed to antigens. The serum so produced will contain different types of antibody with different specificities and affinities. These mixed populations of different types of antibody are termed **polyclonal antibodies**. Polyclonal antibodies are not specific to one antigen. Even in hyperimmune animals, there is seldom more than one-tenth of the circulating antibodies that are specific for one antigen (Harlow and Lane, 1988). This produces a variety of unspecific antigen binding problems in most immunochemical applications such as affinity purification when specific antigen-binding is needed. Therefore, the production of homogenous antibodies is important and this has been achieved with the development of monoclonal antibody technology. Monoclonal antibodies only contain

a single antibody molecule, avoid unspecific antigen-binding problems and in turn greatly increase the scope for the exploitation of antibody-antigen interactions.

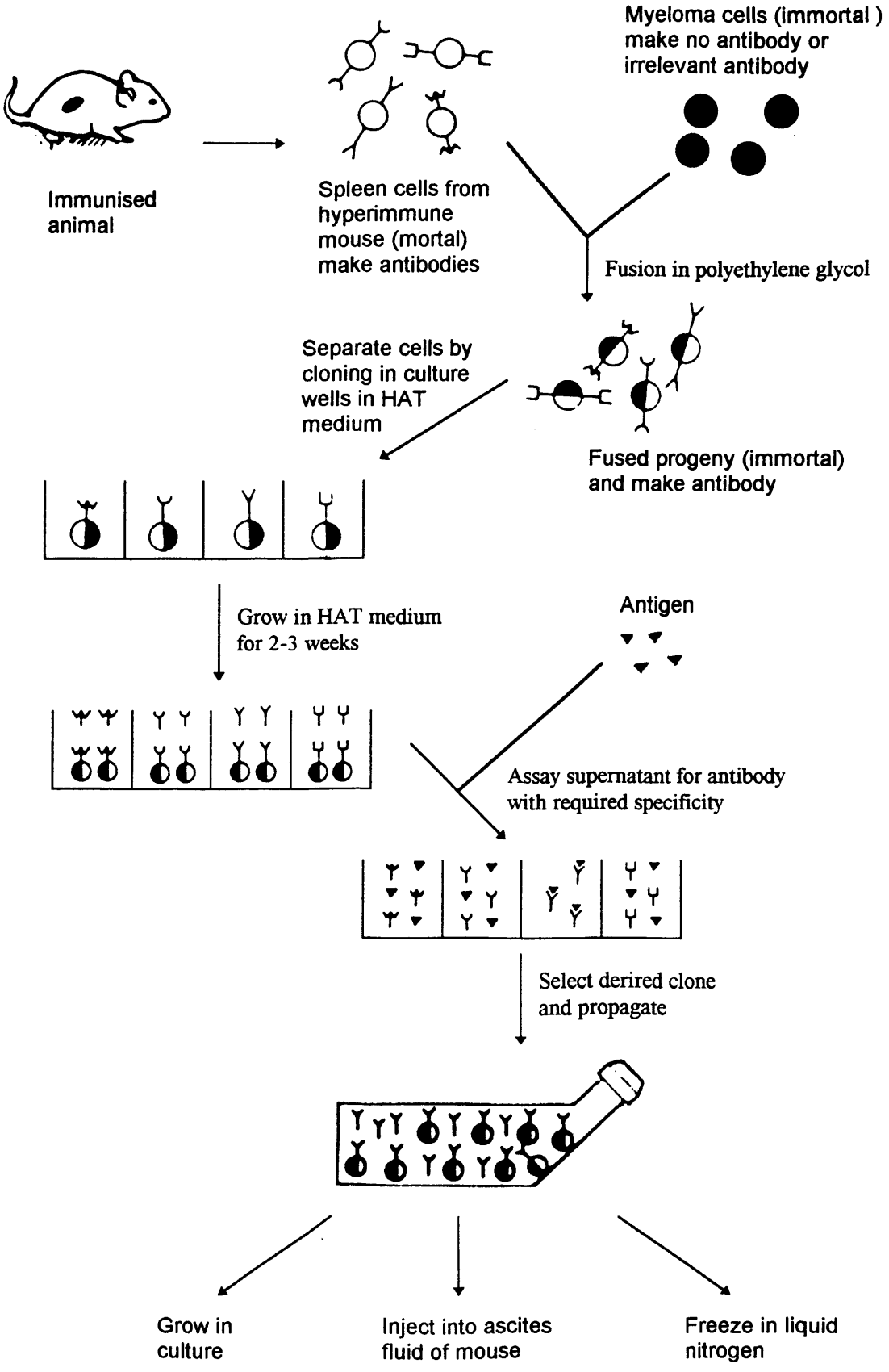
1.2.2 Monoclonal Antibody (MAb) Technology.

In animals, antibodies are synthesised primarily by plasma cells which cannot be grown in tissue culture as an *in vitro* source of antibodies. Therefore, the death of a single antibody producing animal could cause major problems in a diagnostic laboratory. In 1975, Kohler and Milstein developed a technique that allows the growth of clonal populations of cells secreting antibodies with a defined specificity. In this technique, an antibody secreting cell (such as the spleen cell), isolated from an immunised animal, is fused with a tumour B cell or known as **myeloma** cell which can be grown in tissue culture. The hybrid cell was termed as **hybridoma** and can be maintained *in vitro* and will continue to secrete antibodies. Homogenous hybridomas can be then selected for producing homogenous antibodies, named **monoclonal antibodies**.

Fig. 1.5 summaries the general procedures for producing monoclonal antibodies. The usefulness of monoclonal antibodies stems from their specificity of binding, and homogeneity. The production of monoclonal antibodies allows the isolation of reagents for testing and quantitating the presence of a unique, chosen specificity epitope since all the antibodies produced by descendants of one hybridoma cells are identical. In addition, the hybridoma cell lines are able to provide an unlimited supply of antibodies. However, the production of monoclonal antibodies is hard-work, time-consuming and costly. These are problems which can be overcome by genetic engineering.

After the report by Kohler and Milstein (1975, 1976), the potential use of monoclonal antibodies as reagents was soon widely appreciated (Galfre *et al.*, 1977; Williams *et al.*, 1977; Barnstable *et al.*, 1978; White *et al.*, 1978), leading to a rapid growth in the field. There is now an entire field of science research and medical diagnosis which would be impossible without monoclonal antibodies. However, the therapeutic application of monoclonal antibodies, whilst still limited in scope, promises

Figure 1.5 : The general procedure for mouse monoclonal antibody production.



to break its substantial shackles and realise the potential forecast for it by its proponents.

1.2.3 Application of Antibodies.

At present, antibodies have become very much a part of science, medicine, as well as industry. The applications of antibodies are all based on the same fundamental property of the antibody molecule mainly its ability to bind specifically, or selectively, to a particular molecular structure. In some of the applications, other properties of the antibody molecule, such as the ability to bind complement and lyse target cells or hydrolyse molecules, are also involved. However, in most cases, the only property of antibody molecule that is useful is its antigen-binding capacity.

1.2.3.1 Immuno-Diagnosis.

Monoclonal antibodies have displayed a phenomenal success as *in vitro* and diagnostic reagents. The ability of antibodies to bind specifically and selectively to particular molecules has made immuno-diagnosis a powerful tool to detect, identify and quantify the presence of antigen of interest. This success is not only limited to immunology, it is also widely used in other branches of medicine and biology to measure molecules, hormones, cytokines, and enzymes; to identify cells type and pathogens in human, animal or plant; and to detect trace elements and contaminants in the food industry and in environmental monitoring.

The methodology of immuno-diagnosis is based on the use of antibodies to carry out an immunochemical assay. There are many variations on the ways in which immunoassays can be performed (Campbell, 1991). The basic procedures for an immunoassay involves first attaching the antibody (or antigen) to a solid support and then allowing labelled antigen (or labelled antibody) to bind. After washing, the assay is quantified by measuring the amount of labelled antigen (or labelled antibody) retained on the solid support.

A more sophisticated type of immunoassay, known as the "sandwich" immunoassay, is the most commonly used in immuno-diagnosis. The first and second steps are similar to basic immunoassay but using non-labelled antigen or antibody, however, a labelled second antibody is used to detect the bound first antibody (or antigen) at a different epitope. The bound labelled second antibody is then quantified after the unbound second antibody is washed away.

Competition immunoassay is normally used for antigen quantitation. In this assay, a constant amount of labelled antigen (or labelled antibody in another related type of assay) is mixed with the test solution which contains an unknown amount of antigen. The solution is then allowed to bind to a subsaturating amount of antibody (or antigen) bound to a solid support. The presence of antigen in the test solution will reduce the amount of labelled antigen (or labelled antibody) that can bind. The measurement of the labelled antigen (or labelled antibody) retained on the solid support will correlate with the quantity of antigen in the test solution and the exact quantity can then determined from a standard curve of similar assays using known amount of antigen.

Immuno-diagnosis is not only useful in identifying and quantifying antigens. Its applications have been further extended by conjunction with other techniques. For example, radiolabelled or gold-labelled antibodies together with electron-microscopy have been widely used for localise the antigen inside the cells, tissues, even living organisms. Immuno-diagnosis is also a powerful tool in studying types of antigen. An example of a sophisticated study using monoclonal antibodies is in the analysis of human leukocyte differentiation antigens (Knapp *et al.*, 1989; Schlossman, *et al.*, 1994). A group of monoclonal antibodies have been used to identify and characterise more than 100 leukocyte surface antigens, which were not known before monoclonal antibodies against them were prepared. This type of work is important to chart cell phenotype and differentiation status. This system is now known as "Cluster Designation" (CD). In principle, there is no reason why it should not also be possible to type other cell lineages in this manner, such as the differentiation of HIV virus-infected helper T cells.

1.2.3.2 Purification Applications.

Monoclonal antibodies can be utilised in immunoaffinity chromatography to purify antigens from complex mixtures. This method is now a standard in purification. The basic principle is to immobilise the antibody on a resin surface in a column. The mixture containing antigen is then passed through the column to allow the antigen to bind to the antibody. After washing off all unbound contaminant components, the antigen can be eluted by high ionic strength solution. However, the uses of monoclonal antibodies for purifying minor antigens in a complex mixture is effective only when the monoclonal antibodies have high affinity. The use of high affinity monoclonal antibodies may prevent elution of the bound antigen by conditions which do not denature the antigen or antibody.

1.2.3.3 Therapeutic Applications.

The majority of these applications are still at an early stage. The murine OKT3 antibody used in preventing or reversing transplant rejection is the only monoclonal antibody used extensively so far (Jaffers *et al.*, 1986). The mouse monoclonal antibody IgM E5, which recognises bacterial endotoxin, was also evaluated and proven effective in treating human diseases (Foon, 1989).

Antibody therapies are different from drug treatment, choosing a proper antibody is extremely important. Mouse monoclonal antibodies, by nature, are foreign proteins for man and can have the effect of eliciting a human anti-mouse antibody (HAMA) response to shorten the *in vivo* half-life (LoBuglio *et al.*, 1989). In some sensitive patients there may also be a risk of anaphylaxis, and the ability to re-treat a patient could thereby be compromised. Therefore engineered antibodies or engineered antibody fragments which are not recognised as non-self may be useful in this application in the near future.

1.2.3.4 Antibody as Catalysts (abzymes).

This application of monoclonal antibodies was first described in 1986 (Tramontano *et al.*, 1986; Pollack *et al.*, 1986). The basic concept is that a

monoclonal antibody generated to a molecule or molecules which mimic the transition state between substrate and product may accelerate the reaction by forcing the substrate into a suitable conformation for conversion into product. This technology is still in its early stages. The major problem of catalytic monoclonal antibodies is their low K_{cat} values compared to those of enzymes (Benkovic, 1990). However, the catalytic efficiency may be improved by appropriately altering the binding site structure and this might be possible to achieve through antibody gene engineering.

1.2.3.5 Antibodies as Vaccines.

Monoclonal antibodies are not commonly used as parasitic diseases vaccines or passive immunisations as they are too expensive to contemplate on the vast scale in which such therapy would be required. Active vaccination rather than therapy is more appropriate.

The concept of monoclonal antibodies alone as vaccines was largely based on the "anti-iodiotype" theory. This theory suggested that an antibody combining site could be used in immunisation to create a second antibody which was a mirror image of the antigen. This second antibody, being non-pathogenic, could, in theory, therefore safely used in large amounts to elicit an immune response to pathogens such as viruses or tumour cells. However, studies have found that the anti-iodiotype antibodies always have low-affinity binding. Furthermore, an antigen-MAb reaction and anti-iodiotype-MAb reaction for the same monoclonal antibody and antigen were shown to have unrelated recognition patterns (Bentley *et al.*, 1990). Many problems remain to be solved, although antibody engineering or peptide display library technology may ultimately provide solutions.

1.3 Antibody Engineering.

1.3.1 Aims of Antibody Engineering.

Recombinant DNA technology has contributed in several ways to the improvement of antibody technology. Current recombinant antibody techniques are

based on two main concepts. The first is that conventionally-generated MAbs may be improved by the application of DNA technology to alter the quantity or quality of MAbs which have been initially produced by the basic rodent fusion system. The second, is that the constraints of antibody evolution is circumvented by bypassing the immunisation entirely and generating random genetic libraries of heavy and light chains from which a wider range of monoclonal antibodies may be produced.

The aims of antibody engineering include the following :

a) to ensure safe storage of the antibody coding genes of rodent hybridomas as a contingency against failure of hybridoma maintenance.

b) to give more stable expression or safe storage of the relevant antibody coding genes for human hybridomas which are frequently unstable.

c) to alter the constant regions of the MAbs so that they can be used in other species (e.g. man) with minimal rejection.

d) to alter the effector functions (constant domains) of human MAb in order to substitute a suitable class of effector for therapy.

e) to alter the framework sites and constant domains of a MAb, so that it can be used in other species with even more minimal rejection.

f) to alter the CDR in variable domain of a MAb in order to improve its affinity or specificity.

g) to produce larger amounts of MAb in novel expression systems.

h) to amplify the antibody coding genes of spleen, blood or lymph node cells to construct combinatorial antibody gene libraries.

i) to create functional antibody activities in species which do not produce antibodies.

1.3.2 Cloning Strategies.

To carry out antibody engineering, the very first step is to isolate the required antibody genes. The source of antibody genes can be either from germline DNA or directly expressed cells such as hybridomas, peripheral blood lymphocytes, spleen cells

or lymph node cells. The earliest procedure developed for antibody engineering was using hybridoma DNA (Morrison *et al.*, 1984). Later procedures involves construction of cDNA libraries from hybridoma mRNA using either oligo-dT or more specific primers which are complementary to sequences in the region of the constant domain (Liu *et al.*, 1987, White *et al.*, 1987).

At present, polymerase chain reaction (PCR) technology is the most commonly used procedure for extracting antibody genes from their sources. However, in PCR, the size of DNA which may be amplified efficiently and accurately is currently small, e.g. less than 5 kb. Consequently, for antibody genes, this technique has been restricted to cDNA amplification of expressed antibody genes or to amplify variable regions from germlines. Furthermore, PCR techniques require knowledge of the 18-25 bp complementary sequences at either end of the gene sequences which is to be amplified. However, primers for amplification of most antibody genes can be identified as nucleotide sequences at the 5' and 3' ends of heavy and light chain genes are relatively conserved (Orlandi, *et al.*, 1989).

The 3'-end primer can be designed easily as it can be located in the constant domain which is highly conserved, or even in the framework region of J sequences in order to clone the immunoglobulin variable domain only. The 5'-end primer is less straightforward, as it is more variable than the constant regions. The 5'-end primer can only be located at the leader sequence or at the first framework region of variable domains (Padlan, 1993, Larrick *et al.*, 1989).

Identifying primers for PCR has now been resolved as the coding sequences for constant regions flanking the variable domains have been well characterised (Marks *et al.*, 1991) and degenerate primers have been designed (Larrick *et al.*, 1989) for amplification of antibody gene cDNA from a donor tissue mRNA. The definitive sets of PCR primers that are needed for amplifying mouse antibody genes are described by Orlandi *et al.* (1989) and Kettleborough *et al.* (1993), rabbit antibody genes by Ridder *et al.* (1995), and for human antibody genes by Marks *et al.* (1991).

1.3.3 Engineering Antibodies and Antibody Derivatives.

Recombinant DNA technology can be used to engineer a wide range of antibodies and antibody fragments which possess novel structures and multiple functions. Recombinant antibodies ranging from the whole molecule to its fragments, and even single CDR loops, have been cloned and expressed (Winter and Milstein, 1991; Wright *et al.*, 1992; Winter *et al.*, 1994). Fig. 1.6 illustrates a number of the many possible variants on this theme.

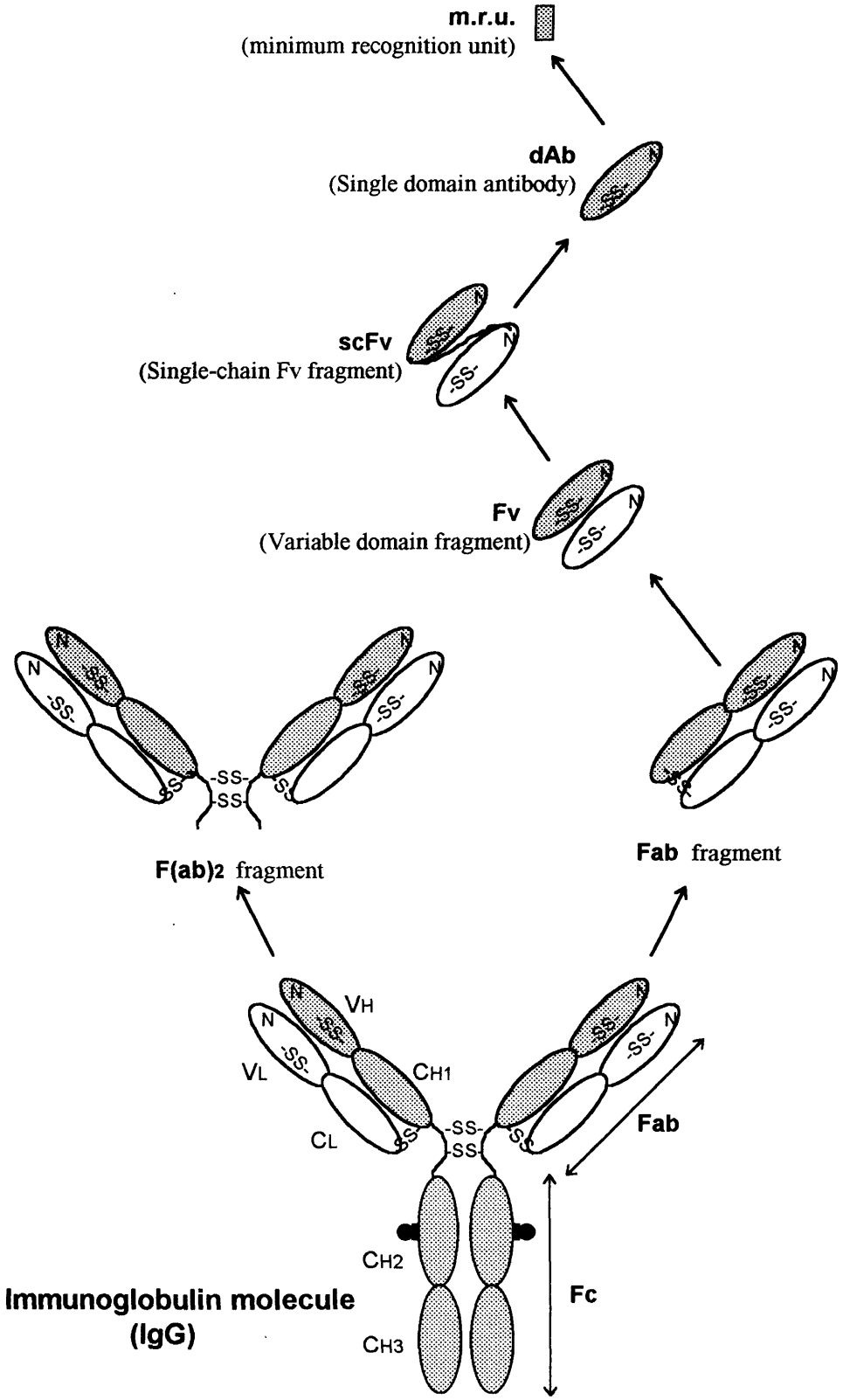
All of these immunoglobulins and derivatives can be successfully assembled in both prokaryotic and eukaryotic (non-lymphoid) cells (Oi *et al.*, 1983; Cabilly *et al.*, 1984). It has also been shown that recombinant antibodies can be retained inside the cytoplasm, targeted to the nucleus (Biocca *et al.*, 1990), or secreted out from the cell cytoplasm (Cattaneo and Neuberger, 1987; Skerra and Plückthun, 1988; Better *et al.*, 1988).

1.3.3.1 Engineering the Whole Antibody.

The whole antibody molecule is required in applications which involve accessory cells of the immune system in ADCC (antibody dependent cell mediated cytotoxicity) which is generally considered to operate through Fc receptors (Campbell, 1991). Engineered antibodies to be used for this purpose require whole antibody molecules and should preferably have the native glycosylation pattern. The following are several applications of engineered whole antibodies.

Chimeric mouse-human antibody: The human anti-mouse antibody (HAMA) response has been found to restrict the therapeutic applications of mouse monoclonal antibodies. Chimeric antibodies comprising murine variable domains and human constant domains may greatly reduce the HAMA response (Morrison *et al.*, 1984). Several chimeric antibodies have been produced and some clinical studies have shown that chimeric antibodies can be as immunogenic as their murine counterparts, while others may be only weakly immunogenic (LoBuglio *et al.*, 1989; LoBuglio and

Figure 1.6 : Structure of various antibody fragments.



Khazaeli, 1993; Meredith *et al.*, 1992). As it stands, this technology has probably been replaced by later versions of humanised antibody.

Humanised antibody: To counteract the HAMA response, the use of a completely humanised antibody has been proposed. The first humanised antibody, the rat antibody YTH34.5HL, which recognises the CAMPATH-H1 (Cdw52) antigen has been engineered by a CDR grafting technique where the six CDRs from the rat MAb were grafted on to a human antibody framework (Riechmann *et al.*, 1988a). However, the affinity of this humanised antibody dropped nearly 40 fold compared to the native rat MAb. Subsequent attempts to humanise murine antibodies in which some non-CDR residues were preserved have been more successful in terms of binding affinity and specificity (Gorman *et al.*, 1991; Tempest, *et al.*, 1991; Studnicka *et al.*, 1994). Clinical studies of those humanised antibodies have demonstrated that the native humanised antibodies are not immunogenic and that their *in vivo* half life is longer than the conventional mouse antibodies (Stephens *et al.*, 1993; Roguska *et al.*, 1994).

***Catalytic antibodies*:** Up to date, about 60 catalytic antibodies have been produced (Stewart *et al.*, 1993; Lerner *et al.*, 1991). Nevertheless, so far, none of the catalytic antibodies compare well with natural enzymes. Several genetic strategies have been used to improve the catalytic activity of catalytic antibodies (Wright *et al.*, 1992; Posner *et al.*, 1994). As yet it remains true that no catalytic antibody can be considered on a par with natural enzymes.

1.3.3.2 Engineering Antibody Derivatives.

Recombinant technology has made it possible to construct many types of antibody derivatives which are not produced naturally. Such molecules may have the potential to revolutionise some aspects of biomedicine, biotechnology and basic biology. Engineered antibody derivatives, can be as small as a single CDR loop, or as

large as an immunoglobulin, or multi-immunoglobulin fusions (Winter and Milstein, 1991; Shu *et al.*, 1993).

Fab fragments: Although the Fab has the same binding specificity as the whole antibody from which it is derived, it lacks the ability of whole antibody to kill cells by ADCC or complement. Additionally, the Fab lacks avidity because of monovalence and it has a very short half life *in vivo*. Some positive reasons for engineering Fab fragments include the fact that Fab proteins can pass more readily into extracellular spaces and can be generated more readily than whole antibodies in bacteria. The fact that Fabs can be generated by bacterial recombinant systems allows their use in a wide range of applications such as research reagents, catalysis, high affinity binding Fab proteins for biosensor construction etc.

Technically, Fabs have been found to be more stable than other small fragments (a factor which might be important for screening purposes). However, the yield of functional Fab in bacteria is low and there is a difficulty to secrete it out from the inner membrane. Studies have shown that by modifying the protein sequences it is possible to improve the yield (Stemmer *et al.*, 1993) and increase the secretion level. For instance, the amount of a Fab antibody secreted from *E. coli* increased after the C_κ domain was replaced by a C_λ domain (MacKenzie *et al.*, 1994a).

Variable domain fragment (Fv) and single chain Fv (scFv): Due to its small size and relative stability, the variable domain fragment (Fv) which consists of noncovalently associated V_H and V_L domains, has been commonly engineered for immunotargeting applications. The idea of constructing the Fv is based on the fact that in certain *in vivo* conditions, IgA and IgG antibodies have been commonly found to be refractory to simple proteolytic Fv release (Huston *et al.*, 1991). Analysis has shown that Fv fragments have the same antigen-binding properties as the parent antibody. Upon dilution however, Fv fragments have a tendency to dissociate into non-functional V_H and V_L domains (Riechmann *et al.*, 1988b).

A number of attempts had been made to stabilise the Fv structures by covalent cross-linking. Initially, the cross-linking was done chemically using glutaraldehyde, and then by introducing an intermolecular disulphide bond (Briggs and Giorasch, 1986). A far simpler solution was devised in 1988 by gene expression of a single polypeptide in which the V_H and V_L domains were connected to each other by a polypeptide linker (Bird *et al.*, 1988; Huston *et al.*, 1988). Some scFv proteins have been found to have equivalent affinity to Fv, Fab and even the parent antibody (Milenic *et al.*, 1991; Denzin and Voss, 1992; Malby *et al.*, 1993).

The primary function of the linker in the scFv molecule is to maintain the close proximity between the V_H and V_L domains. The linker is not required to orientate the variable domains into a specific conformation in order to form the active binding site since this occurs spontaneously (Huston *et al.*, 1991). However, the polypeptide linker must span at least 35 Å (estimated from monomeric Fv models) in distance between the carboxyl terminus of one variable domain and amino terminus of the other for proper folding (Bird *et al.*, 1988; Huston *et al.*, 1988). Either V_L or V_H may be the amino terminal domain, as 5'-V_L-linker-V_H-3' or 5'-V_H-linker-V_L-3' configurations usually exhibit identical affinity and specificity (Whitlow and Filpula, 1991).

Peptide sequences whose lengths and conformations are compatible with bridging the variable domains without serious interference can be used as linkers. The most commonly used linker comprises 15 amino acid residues of three segments of GGGGS (Bird *et al.*, 1988; Takkinen *et al.*, 1991; Glockshuber, *et al.*, 1994). This linker contains stretches of glycine (G) for flexibility and serine (S) residues to provide hydrophilicity to prevent association with hydrophobic region of variable domains. Investigations have also revealed that the linker does not disrupt antigen-binding activity (Huston *et al.*, 1988; Glockshuber *et al.*, 1990) and both of the possible orientations of the linker connection, namely 5'-V_H-linker-V_L-3' or 5'-V_L-linker-V_H-3', produced scFv proteins that give similar antigen binding activity (Huston *et al.*, 1988).

Beside the $(GGGGS)_3$ linker, evidence has been presented which indicates that certain longer linkers can increase an antibody fragment's affinity and decrease the formation of aggregates (Whitlow *et al.*, 1993). Such linkers lacking a well-ordered

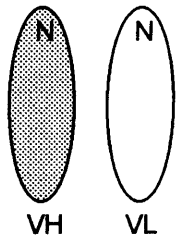
3-dimensional structure should prove beneficial in scFv with therapeutic applications for which reduced immunogenicity is desired. Other more stable linkers with helical structures have also been constructed to suit non-physiological or industrial applications (Pantoliano *et al.*, 1991).

Diabodies: Engineered antibody fragments such as Fab, Fv, and scFv are monovalent while native antibody molecules are divalent. However, by constructing the antibody fragment in the form of a diabody or other dimeric form, it becomes bivalent and even possibly as bispecific (Holliger *et al.*, 1993). Bivalency is important as it can allow antibodies to bind to multimeric antigens with great avidity while bispecificity can allow the cross-linking of two different types of antigens (Staerz *et al.*, 1985).

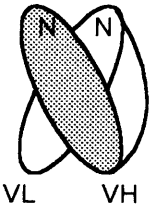
The bivalent diabodies can be made by linking the V_H and V_L in tandem with a short linker fragment, or even without a linker fragment, so that normal folding between V_H and V_L domains cannot take place, but pairing with V_L and V_H domains of other scFv molecule to form dimeric scFv polypeptides with two antigen-binding sites can occur (Holliger *et al.*, 1993). Bispecific diabodies are similar to bivalent diabodies except that the scFv has V_H and V_L domains from two different cross-over antibodies, i.e. V_{HA}-V_{LB} and V_{HB}-V_{LA}. Therefore, bispecific diabodies are formed when two different scFv polypeptides pair together (Holliger *et al.*, 1993).

Other attempts have also been made to dimerise two antibodies fragments (Fig. 1.7). These included chemical cross-linking of the hinge cysteine residues (Shalaby *et al.*, 1992), or by the inclusion a carboxyl terminal peptide that promotes dimerisation such as the leucine zipper regions of the *fos* and *jun* oncogene products or two amphipathic helics at both V_H and V_L domains of scFv polypeptides (Pack and Plückthun, 1992). Bivalent scFv proteins can also be produced in eukaryotic systems using human Fc sequences fused to the carboxyl terminal of an scFv; association between two of these molecules will create a disulphide bridge resulting in formation of a functional Fc (Shu *et al.*, 1993).

Figure 1.7 : Structures of dimeric antibodies.

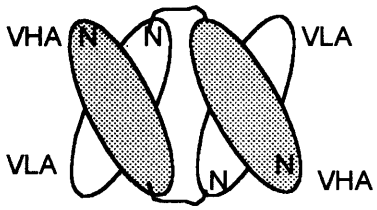


Signle Variable Domains

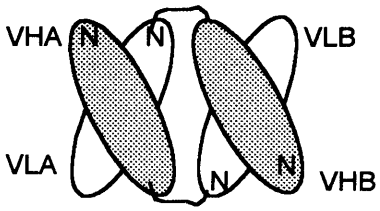


Single-chain Fv (scFv)

Diabodies

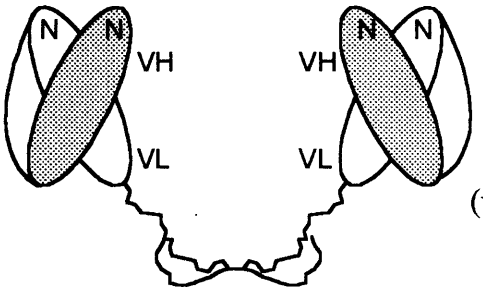


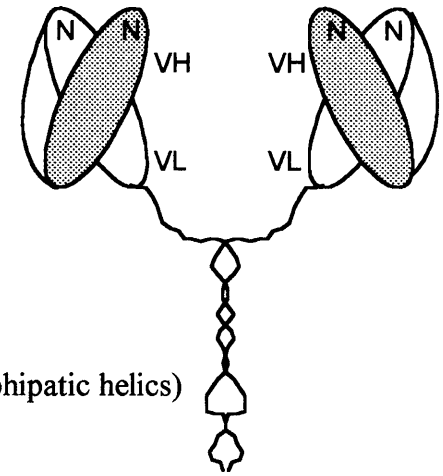
(monospecific)



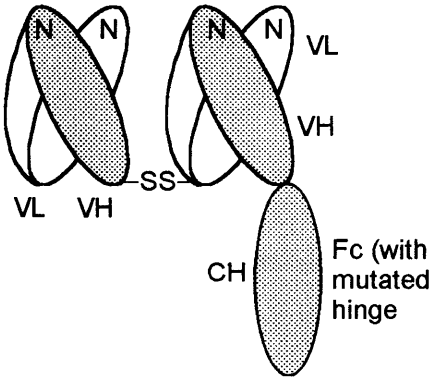
(bispecific)

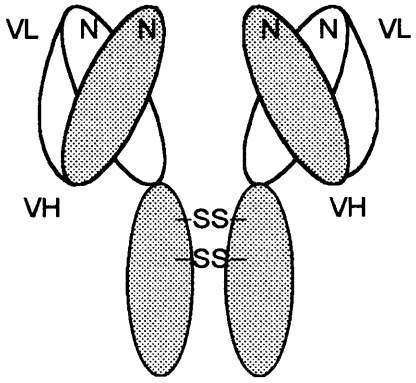
Other Dimeric Antibodies





(with amphipatic helics)





Single Domain antibody (dAb): This term is used to describe binding molecules constructed with only the heavy chain variable domain (Ward *et al.*, 1989). These fragments are reported to be able to bind to soluble protein antigens with moderate affinity but tend to non-specifically cross-react more readily than whole antibodies.

Minimum Recognition unit (m.r.u.): Individual CDR regions can form high-affinity interactions with antigens and individual CDRs have been reported to show similar binding characteristics to the whole antibody from which they have been derived (Williams *et al.*, 1989; Traub *et al.*, 1989). Thus CDR regions are considered to have potential in the field of rational drug design.

1.3.3.3 Engineering Multifunctional Antibodies.

Recombinant fusion proteins are widely used to produce artificial multidomain-multifunctional proteins. In antibody engineering, antibodies, mainly scFv, have been engineered into fusion proteins, combining the antigen specificity of the scFv with the function of its fusion partner. ScFvs are commonly used for this purpose because scFv genes are small in size, easily manipulated, and have binding properties similar to the parental antibody. Several applications of multifunctional antibodies are list as follow :

Immunoreagents: Enzymes conjugated to antibodies, such as alkaline-phosphatase or horse radish peroxidase, have been widely used in immuno-diagnosis. They are normally produced through chemical reaction. However, in antibody engineering, scFvs have been fused with *E. coli.* alkaline phosphatase (Gandecha *et al.*, 1994), β-lactamase (Kolmar *et al.*, 1992), the Fc binding domain of staphylococcal protein A (Tai *et al.*, 1990; Gandecha *et al.*, 1992), and *E. coli* maltose-binding protein (Bregegere *et al.*, 1994) to permit one step immunodetection or provide an affinity handle for purification.

Immunotoxins: Extensive *in vitro* and *in vivo* studies in cancer diagnosis and therapy have been performed using toxins from plants, animals, bacteria, and fungi conjugated chemically to MAbs. ScFvs fused with several of these toxins have now been cloned. These included several antitumor scFv genes fused with *Pseudomonas* exotoxin (PE) (Batra *et al.*, 1990, 1992; Brinkmann *et al.*, 1991, 1993) and diphtheria toxin (DT) (Batra *et al.*, 1991; Chaudhary *et al.*, 1993). The *in vitro* and *in vivo* activities of several scFv immunotoxins have shown that these recombinant immunotoxins are specifically cytotoxic to tumour cells expressing the corresponding antigen (Batra *et al.*, 1991; Brinkmann *et al.*, 1991, 1993). In other cancer research, scFv genes have been fused with cytokines or enzymes to target the tumour cells (Wels *et al.*, 1992; Goshorn *et al.*, 1994; Savage *et al.*, 1993).

Metal-binding antibodies: The ability to engineer metal-binding peptides on to an scFv could greatly improve radioimmunoscintigraphy. The Ca^{2+}-binding protein, calmodulin, has been fused to an antibody against a *Salmonella* polysaccharide as a heavy metal chelating agent for radioimmunodetection (Mackenzie *et al.*, 1994b). The metal-binding fusion can also be used as a versatile tag to facilitate immunodetection of antibody fragments (Neri *et al.*, 1995a).

The design of multipurpose multifunctional antibodies has varied according to the purpose for which the fusion protein was intended. One of the considerations includes the selection of the amino terminal or carboxyl terminal as the site of fusion, a choice which may be influenced by the relationship between the scFv and the fusion partner in the context of the function of both fusion partners. It is suggested that a construct in which the carboxyl terminal of a scFv gene is attached to the amino terminus of the effector-fusion is the most appropriate since in nature, the carboxyl terminus of the variable domain is attached to the constant domain. This configuration has been successfully used to construct immunotoxins (PE40). Such immunotoxins have been shown to possess enhanced cytotoxicity compared to chemically synthesed conjugates (Batra *et al.*, 1990; Kreitman *et al.*, 1990). However, fusion to the amino terminal has also been shown to allow production of functional antibody (Neri *et al.*, 1995b). Indeed, for bivalent or bispecific scFv fusions such as chelating recombinant

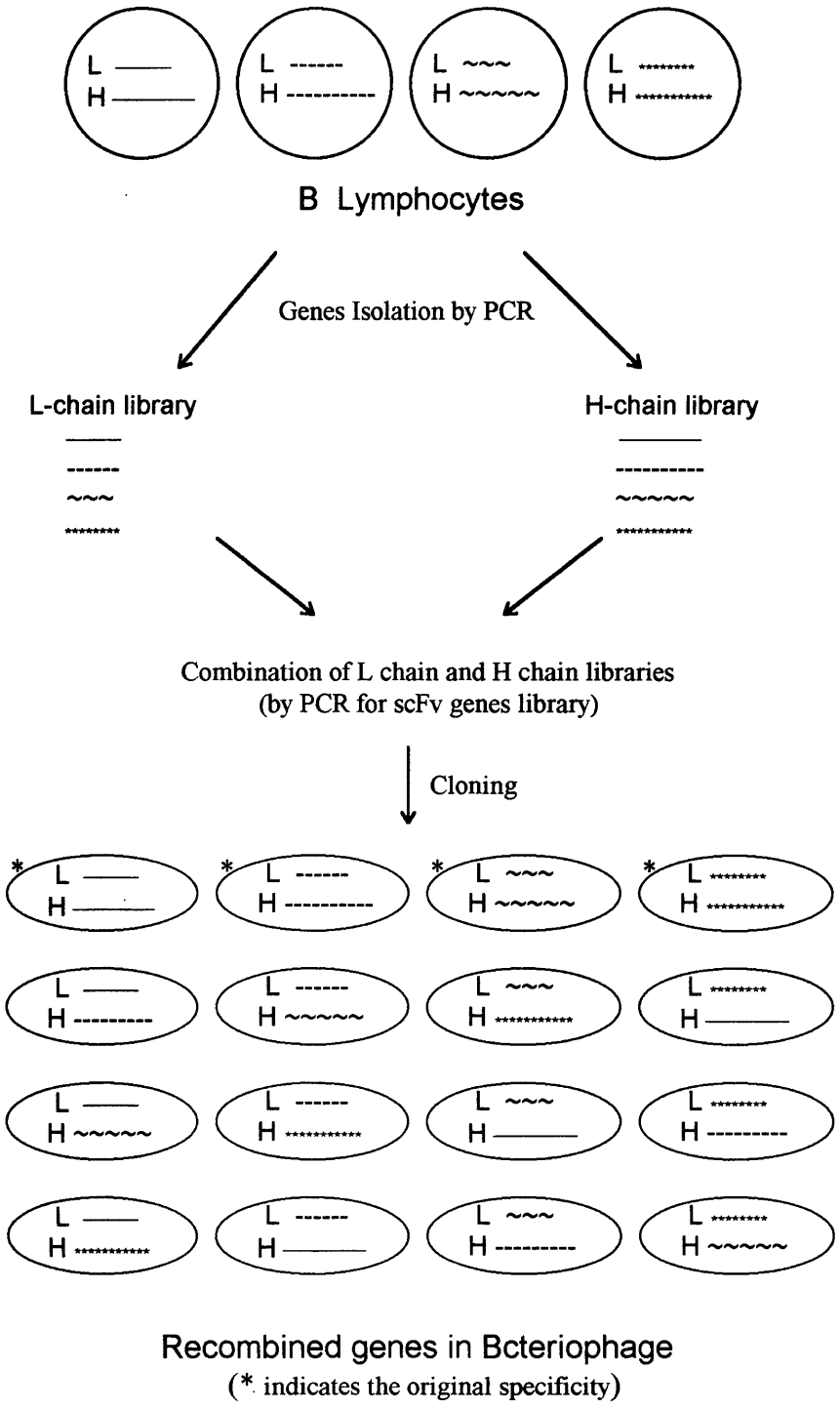
antibodies, or complex multidomain fusion proteins, it may be essential for the fusion partner to be attached to the amino terminal of the scFv. In such cases, a hinge region may be required in between the fusion proteins (Pack and Plückthun, 1992; Shu *et al.*, 1993).

The binding activity of scFv proteins in multifunctional antibodies may influenced by its fusion partner. For example, the use of two different scFv proteins, anti-Tac and anti-TFR, fused separately to the carboxyl terminus of diphtheria toxin, has different effects on scFv specificity and affinity. The specificity and affinity of the anti-Tac fusion is unaltered whereas anti-TFR fusion changed (Chaudhary *et al.*, 1990; Batra, *et al.*, 1991). This dependency must be considered when designing multifunctional antibodies fusion constructs.

1.3.3.4 Antibody Combinatorial Libraries.

The construction of antibody combinatorial libraries involves the amplification of light chain and heavy chain genes to create genes libraries and combination of the two gene libraries to create a huge antibody repertoire (Fig. 1.8). In this technology, mRNA is isolated from antibody producing cells such as spleen or lymph nodes of an immunised rodent, human, or naive subject. cDNA is then made and the variable domain (or variable domain and first constant domain for Fab antibodies library) is amplified by PCR using specific consensus oligonucleotide primers to create a V_L gene library and a V_H gene library. The two libraries are then randomly combined, linked together by PCR to form scFv gene libraries. The PCR amplified scFv gene libraries were then ligated to the phagemid vector which allowed the use of living bacteria to display the scFv fusion proteins on the surface of filamentous fd bacteriophages (Winter and Milstein, 1991). By taking advantage of phagemids to display antibodies on the surface of bacteriophage, scFv antibodies with desired properties can be readily and economically isolated by affinity matrix (panning) against immobilised antigens (Ward *et al.*, 1989; Hoogenboom *et al.*, 1991; Kang *et al.*, 199; Lerner *et al.*, 1992; Griffiths *et al.*, 1993). After the selection through affinity, viable phage was grown in bacterial culture and the cDNA was extracted,

Figure 1.8 : Construction of combinatorial libraries.



analysed, and transferred to other expression vectors for large-scale production of scFv antibodies or expressions in other heterogenous hosts.

This technique dispenses with hybridoma cell fusion procedures entirely and the potential repertoire size is very large. A mouse spleen has 10^8 cells of which one third are B cells. Hence, if all of these B cells offer a different specificity, this approach could offer an immense range of potential binding specification. In theory, it is possible to produce a combinatorial library with a minimum of 10^{10} antibodies which is the primary immune repertoire of a mouse or human (Winter and Milstein, 1991). In contrast, a conventional mouse fusion will only yield 10^3 clones.

1.3.3.5 Total Human Antibody.

Total human antibody is a term coined to mean "an antibody which is totally created by human". The concept of the total human antibody is based upon the use of combinatorial libraries and phage display technology. It is now clear that phage display technology is workable and has many applications. Some fascinating work has been done by the major laboratories in this field, these include

- generating huge naive libraries by extracting the germline V genes (Hoogenboom and Winter, 1992; Nissim *et al.*, 1994; Fisch *et al.*, 1994),
- creating dual-combinatorial libraries by cross-transfecting two different combinatorial libraries from different species to obtain even bigger libraries (Barbas *et al.*, 1991; Hoogenboom *et al.*, 1991, Kang *et al.*, 1991).
- affinity maturation of recombinant antibodies through site directed mutation (Sharon 1990, Hawkins *et al.*, 1992),
- *in vitro* mutagenesis and rational selection by random grafting the CDR loops with the aim to change one antibody to another (Hermes *et al.*, 1989; Gram *et al.*, 1992; Barbas *et al.*, 1992);

- chain shuffling through 'epitope imprinting' techniques to convert an antibody of one species to another (Clackson *et al.*, 1991; Hoogenboom *et al.*, 1994; Figini *et al.*, 1994).

The proponents of these techniques are firm in the view that *in vivo* production of antibodies is now out of date and will soon be replaced by these new methods.

1.3.4 Expression of Recombinant Antibodies.

In normal B cell or myeloma cell lines, antibody molecules undergo a complex series of steps in their final assembly. The leader sequence on heavy and light chains have to direct the chains into the endoplasmic reticulum where the leader sequence splits off and the heavy and light chains are assembled together (Munro and Pelham, 1986; Pelham 1989). The antibody also undergoes glycosylation within the endoplasmic reticulum (Farquhar, 1985), providing the crucial recognition site for recruiting effector functions such as complement-dependent cytotoxicity (CDC) and antibody dependent cell-mediated cytotoxicity (ADCC) for cancer therapy.

Thus, the requirements for a hetero-expression system are that it assembles and folds the protein structures correctly, removes the leader sequences if they are present, glycosylates correctly, and secretes the protein in substantial amounts. Both eukaryote and prokaryote expression systems including bacteria, yeast, baculovirus, plant cells, and mammalian cells have been shown to be able to successfully express recombinant antibody genes. Each system has its own characteristics, properties, advantages, disadvantages, and potential applications.

1.3.4.1 Mammalian Cell Expression Systems.

The early attempts on antibody engineering were made with mammalian cell expression systems. In fact, successful expression of clinically-used recombinant antibodies to date has largely utilised myeloma cell lines which are themselves non-secretors but possess the apparatus to secrete (Neuberger *et al.*, 1983; Sharon

et al., 1984). This is very important in expressing whole antibodies for therapy where correct folding and proper glycosylation may be essential.

Myeloma cells which produce recombinant antibodies (or "transfectomas") in particular have the ability to correctly process and post-translationally modify immunoglobulins. Recombinant antibodies produced by myeloma cells are low in immunogenicity and can serve as a suitable isotype to recruit the desired effector functions. Although the yield in culture media, up to 40 μg/ml (Crowe *et al.*, 1992), is usually low compared to mouse hybridoma, but it can be improved by growing the transfectomas in mouse ascites with preinjection.

To overcome the problems of low expression in myeloma cells, nonlymphoid cell systems have been developed as a practical alternative for recombinant antibody expression (Bebbington, 1991). Several nonlymphoid cells, such as Chinese hamster ovary cells (CHO), C6 glioma, PC12 pheochromocytoma, and HeLa cells, have been tested for the ability to assemble and secrete recombinant antibody. The expression levels were very low using HeLa and C6 glioma cells (Cattaneo and Neuberger, 1987) but rather high in CHO cells (Crowe *et al.*, 1992). Furthermore, the CHO transfectomas can be grown in large scale suspension cultures using serum-free media, and the extensive experience of manufacturing of recombinant proteins using CHO cells also makes them attractive candidates for the production of recombinant antibody.

Research showed that the expression level of recombinant antibodies can be further increased by using a mammalian expression vector rather than integration into genome DNA (Wood *et al.*, 1990; Fouser *et al.*, 1992). Recent research also revealed that using glutamine synthetase (GS) as an selective marker and human CMV MIE as promoter-enhancer to express heavy and light chain genes, the CHO transfectomas were able to secrete recombinant antibody at a level of 200 μg/ml cB72.3 antibody (Bebbington, 1991; Brown *et al.*, 1992) and even up to 560 μg/ml has reported by Brown *et al.* (1992).

In addition to mammalian cell systems, whole mammals such as mice, namely transgenic mice, have been successfully used to produce human antibody light chains

(Davies *et al.*, 1993), chimeric mouse/human anti-human interleukin-2 receptor antibodies (Cahill *et al.*, 1993) and human antibody repertoires (Neuberger, 1994).

1.3.4.2 Bacterial Expression Systems.

Since bacteria have a high growth rate, bacterial expression systems have become the most commonly used tool for producing recombinant proteins in large quantities at low cost. The bacterial expression of recombinant antibodies has thus become a familiar task in molecular biology. However, the reducing environment of the bacterial cytoplasm does not allow the production of functional whole antibodies. Furthermore, the bacterial system's failure to glycosylate proteins makes it unsuitable whenever glycosylation is essential for antibody function (Plückthun and Skerra, 1989). Bacterial systems however, have been extensively used to produce Fab, Fv or scFv fragments. There are many factors that influence antibody expression levels in bacteria. Among these are the primary structure of the antibody genes and amino acids, the structure of the expression vector, the bacterial strains used, and expression conditions (Plückthun, 1991a). The yield is usually low (at a few mg/l) compared to a myeloma expression system, especially in ascites, but the cost is comparatively low.

Several different species of Gram-positive and Gram-negative bacteria have been used for the expression of antibody fragments. Two commonly used Gram-positive bacteria are *Bacillus subtilis* (Wu *et al.*, 1993) and *Streptomyces lividans* (Ueda *et al.*, 1993) with yields of 5 mg/l and 1 mg/l respectively. In the case of Gram-negative bacteria, *Escherichia coli* is the most commonly used (Skerra, 1993). Different *E. coli* strains have been found to exhibit markedly different expression levels. Among several commonly used strains, TG1, BMH71-18, W3110wt, *E. coli* B, Y1089, LE392, and HB101, stains MM294, *E. coli* B, and W3110wt exhibited relatively high expression levels (Duenas *et al.*, 1994, Ayala *et al.*, 1995).

Cytoplasmic expression of antibody fragments in bacteria usually results in the formation of insoluble and inactive protein aggregates or inclusion bodies (Cabilly *et at.*, 1984). Yields of intracellularly produced antibody fragments may be high, but only a small fraction of the yield can be reconstituted *in vitro* to produce active

fragments (Sandhu, 1992). However, secretion to the bacterial periplasm results in the production of active antibody fragments (Plückthun, 1991a). Targeting to the periplasm has been achieved by fusing the antibody gene to a signal peptide, such as *ompA*, *pelB*, *PhoA*, *Bla*, *ompF*, *StII* (Skerra, 1993), *Spssi* (Ueda *et al.*, 1993), or *P43* (Wu *et al.*, 1993). All these signal peptides seem equally successful.

Nevertheless, the expression levels are strongly dependent on the nature of individual recombinant antibody fragments. Expression levels of different antibody fragment in the same expression system, may vary by more than 100 fold (Kellely *et al.*, 1992). However, the reasons for these differences are still unclear. Some studies suggested that the amino acids sequence next to the signal peptide can inhibit protein translocation to the periplasm (Li *et at.*, 1988, Ayala *et al.*, 1995). Several strategies have been used to increase the expression levels of functional antibody fragments in *E. coli*. Growth at low temperature and mild induction (Shibui *et al.*, 1992), fusion with *E. coli* proteins such as the maltose-binding protein (Bregegere *et al.*, 1994) have been used to improve the expression of functional antibody fragments in *E. coli*.

1.3.4.3 Yeast Expression Systems.

Yeast cells have been reported to secrete whole glycosylated (but with inappropriate glycosylation) antibody molecules at very low levels (μg/l) (Horwitz *et al.*, 1988). However, Fab, Fv and scFv fragments can all be produced by yeast cells at much higher levels. In fact, yeast expression systems offer a way to avoid the problems of inclusion bodies and poor *in vitro* assembly of antibody fragments frequently experienced with *E. coli*.

The yeast, *Sacharomyces cerevisiae* was found to be able to recognise mammalian secretion signal sequences to some extent but only about 25% of total immunoglobulin was secreted from the yeast (Wood *et al.*, 1985). Yeast signal sequences are more effective at secreting mammalian proteins from yeast. It has also been shown that expression of the chimeric immunoglobulin L6 in *S. cerevisiae* using yeast invertase signal sequence and driving the expression by yeast 3-phosphoglycerate kinase (PGK) promoter can produce 100 mg/l of light chain and up to 80 mg/l of

heavy chain in culture supernatants. 50% to 70% of the products were associated as a Fab which was functional and identical to the Fab produced by *E. coli* or prepared by papain digested antibody.

The methylotrophic yeast *Pichia pastoris* has been recently used for expressing antibodies fragments (Ridder *et al.*, 1995) due to it capability to produce excellent high-level expression of various heterologous proteins, either intracellularly or secreted into the culture supernatant (Cregg *et al.*, 1993). *P. pastoris* combines the features of an eukaryotic secretion machinery and the features of bacteria like fast growth and a requirement for non-complex growth media. Furthermore, induction can be efficiently regulated by the supplementation of methanol as the carbon source. All these factors make the *P. pastoris* an attractive host for recombinant antibodies expression. Recent research has shown that *P. pastoris* can produce as high as 100 mg/l scFv in media supernatants in small scale cultures (Ridder *et al.*, 1995).

1.3.4.4 Baculovirus Expression Systems.

The baculovirus system is based on the infection of insect cells such as *Spodoptera frugiperda* with the multiple nucleocapsid polyhedrosis virus (MNPV) (Miller, 1988). The polyhedrin gene, which is non-essential for the infection of the cells in culture, is driven by a very strong late promoter and can be replaced by a gene of interest. The baculovirus/insect cell system provides some advantages over conventional prokaryotic expression systems since it is capable in signal peptide cleavage, N-linked glycosylation, correct cellular compartment localisation, and extracellular secretion (Miller, 1988).

This system has also proven to yield substantial amounts of heavy and light chains and a yield of 5 mg/l has been correctly assembled (Haseman and Capra, 1990). This is not a high yield compared to hybridoma but substantially above the *E. coli* system. In addition, correct Fc domain function was demonstrated, indicating that the antibodies were correctly glycosylated (zu Putlitz *et al.*, 1990). However, large-scale production is difficult in insect cultures as viral infection would ultimately cause the death of the cells.

1.3.4.5 Plant Expression Systems.

Similar techniques to those used for expression in plants of a whole range of heterologous proteins have also been employed to produce recombinant antibodies or antibody fragments in transgenic plants. Whole antibody (Hiatt *et al.*, 1989; Hein *et al.*, 1991; Ma *et al.*, 1994; Voss *et al.*, 1995), Fab fragments (De Neve *et al.*, 1993), and scFv fragments (Owen *et al.*, 1992; Tavladoraki *et al.*, 1993) have been successfully expressed in plants. The expression of antibody genes may be driven by different promoters (Hiatt *et al.*, 1989; Düring *et al.*, 1990; Owen *et al.*, 1992) and the antibodies can be fused to targeting signals at amino terminal to direct antibody proteins into the secretory pathway (Firek *et al.*, 1993; Ma *et al.*, 1994), or targeted to different cell compartments (Schouten *et al.*, 1996).

The expression level of recombinant antibodies in plants varies very significantly. In transgenic tobacco, the expression of whole antibody ranges from 0.1 to 1.3% of total soluble protein (Hiatt *et al.*, 1989; Hein *et al.*, 1991; Düring *et al.*, 1990; De Neve *et al.*, 1993; Ma *et al.*, 1994; and Voss *et al.*, 1995), Fab fragments from 0.25 to 0.8% (De Neve *et al.*, 1993), and scFv from 0.06 to 4.8 % (Owen *et al.*, 1992; Firek *et al.*, 1993; Tavladoraki *et al.*, 1993; Artsaenko *et al.*, 1995). It is not possible to deduce a general pattern of yield, since expression levels of recombinant antibodies in plants are controlled by several molecular and cellular factors. These include promoter activity, signal sequences, transcript processing and protein stability (Whitelam *et al.*, 1993).

Antibodies biosynthesised by plants have been shown to exhibit binding activities equivalent to the parental monoclonal antibody (Hiatt *et al.*, 1989; Düring *et al.*, 1990; Ma *et al.*, 1994), and therefore could potentially be used in immunotherapy (Ma *et al.*, 1994; Ma *et al.*, 1995). However, plant synthesised antibodies do not display native glycosylation patterns and may differ in their biological properties such as biodistribution, *in vivo* half-life, and effector functions (Hiatt, 1990).

Plant synthesised antibodies, also have *in vivo* (or more accurately *in planta*) activity. It has been shown that recombinant antibodies in plants can functionally bind

to antigens such as phytochrome or plant hormones to modulate the plant phenotype (Owen *et al.*, 1992; Artsaenko *et al.*, 1995). Recombinant antibodies have also been found to act directly against plant viruses and thus reduce infection (Tavladoraki *et al.*, 1993).

1.4 Aims of the Project.

The work reported in this thesis involves studies on several aspects of recombinant antibodies. The studies began with monoclonal antibody technology, continued with the construction of scFv genes, and followed by their expression in bacteria, yeast, and plants under a range of conditions in order to further understand the factors which may affect recombinant antibodies production. Other studies included the construction and expression of antibody variants such as hybrid scFvs, bivalent antibody (diabodies), and bifunctional antibody molecules by fusing scFv with maltose-binding protein and green fluorescent protein (GFP).

In this study, hybridomas secreting monoclonal antibodies against abscisic acid (ABA), an endogenous plant hormones, were produced and used for construction antibody fragment, scFv.

Abscisic acid was chosen as it is an important compound in plants. ABA has been shown to mediate many physiological and developmental processes throughout the life cycle of plants. Phenomena which involve ABA include responses to water stress, dormancy, bud inhibition, abscission, seed development and germination, root geotropism and wounding (Walton, 1980; Hetherington and Quatrano, 1991).

The ability to specifically modulate the titre of ABA in particular cells or tissues at the whole plant level can provide an experimental tool to assist in the investigation of ABA physiology and transportation. Several approaches have been adopted to study the molecular mechanism of ABA action. This includes the assessment of total ABA levels through the use of ABA-deficient mutant (Finkelstein and Somervilla, 1990), however, it is difficult to see how these methods could be used to study ABA compartmentation and transportation. Recent research which involved targeting ABA

regulated genes and receptors, however, is only an indirect indication of ABA response.

In these studies, it is not only aimed to produce antibody against ABA as a immunoreagents for plant studies, but also for use in immunomodulation studies. It is expected that recombinant antibody bound to ABA in the cytoplasm or apoplast or to particular types of tissue would lead to a sequestration of ABA in that compartment and a reduction in free ABA titre. This in turn would be expected to perturb the physiology of transgenic plant and yield information relevant to the mechanisms involved in ABA-mediated phenomena.

As a low molecular weight, non-protein, mobile antigen, plant component, ABA can be regarded as being a representative of a class of molecules that are also likely to be worthwhile targets for immunomodulation in plants. These include other regulators and signalling molecules, such as hormones and jasmonic acid, pathogen-derived phytotoxins, herbicides, pollutants, volatiles etc.. It is hoped that the knowledge gained from these studies will be relevant to the further exploitation of the immunomodulation approach.

CHAPTER 2

Materials and Methods

2.1 Materials, Chemicals and Reagents.

2.1.1 General Materials.

Most of the common chemicals and reagents were obtained from Sigma, BDH, and GIBCO-BRL unless otherwise stated.

Animal cell culture media, supplements, and antibiotics, were purchased from GIBCO-BRL. MS media for plant tissue cultures, vitamins and hormones, alkaline phosphatase and peroxidase labelled anti-mouse IgG antibodies were obtained from the Sigma Chemical Company. Anti HSV-Tag antibody was supplied by Novagen, anti-MBP antibody was obtained from New England Biolabs, while anti-GFP antibody was supplied by Clontech.

Nucleotides for PCR were purchased from Promega. The cap analogue $m^7G(5')ppp(5')G$ for *in vitro* transcription was supplied by Pharmacia. Oligonucleotides used as primers in PCR reactions were obtained from the DNA synthesising facility in the Protein and Nucleic Acid Chemistry Laboratory of the Biochemistry Department, Leicester University.

Water used for preparing the reagents was deionised distilled water (Q-H_2O) prepared using the Milli-Q reagent water system (Millipore).

2.1.2 Reagent Kits.

cDNA synthesis kits were purchased from Pharmacia, i*n vitro* transcription kits from Promega, chemiluminescent kits from Boehringer Mannheim, agarose gel - DNA extraction kits from Bio-Rad, and Plant RNA miniprep kits from Qiagen.

2.1.4 DNA Modification Enzymes.

Most of the DNA modification enzymes were supplied by GIBCO-BRL unless otherwise stated. *Psp*A I restriction enzyme was obtained from Stratagene and *Taq* polymerase was purchased from Promega.

2.1.5 Molecular Cloning Vectors.

pGEM11Zf(-) was obtained from Promega. pCR, pTrcHis, pRSET, pPIC9K, and pPICZα were purchased from Invitrogen. pMal-c2 and pMal-p2 were supplied by New England Biolabs. pJIT117 and pBin 19 were obtained from the John Innes Institute and pP3C2S 402 was obtained from the Scottish Crop Research Institute.

2.1.6 *Escherichia coli* Strains.

The following bacterial strains were used in gene cloning and bacterial expression of scFv genes and derivatives.

- **λXL-1 Blue** : recA1, endA1, gyrA96, thi$^-$, hsdr17, supE44, relA1, lac [F', proAB lacI ^{q}ZΔM15, Tn10(tetr)].
- **W3110wt** : F$^-$, hsdR$^-$, hsdM$^+$, [p3(kmI)].
- **JM109(DE3)** : recA, endA1, gyrA96, thi, hsdr17 (r_K^-, m_K^-), supE44, relA1, Δ(lac-proAB), [F', proAB, tra D36, proAB, lacI ^{q}ZΔM15].
- **BL21(DE3) trxB$^-$** : F$^-$, ompT$^-$, (r_B^-, m_B^-), (λimm21, lacI, lacUV5, T7 pol, int$^-$), trxB $^-$.

Most of the cloning and plasmid amplification were carried out in *E. coli* XL-1 Blue due to the fact that this strain is a *recA*1 mutant and it lacks the endogenous restriction system, therefore, any undesirable restriction of the cloned DNA and the recombination with host chromosomal DNA can be prevented.

W3110wt was the strain used for expressing those scFv genes driven by the *lac* and *trc* promoters since it has been reported to be the best *E. coli* strain for expressing scFv genes (Dueñas *et al.*, 1994).

JM109(DE3) was used to express those scFv genes driven by the T7 promoter as this strain's genome carried a copy of bacteriophage λ gene containing the T7 RNA polymerase. The XL-1 Blue *E. coli* was also used to express the scFv genes driven by the T7 promoter through M13/T7 phage infection. The F' gene in XL-1 Blue *E. coli* allowed the M13/T7 phage infection thus supplying T7 RNA polymerase and initiating expression. This system was used to block any background expressions prior to phage infection.

The expression of functional scFv proteins requires an oxidising environment. However, the *E. coli* cytoplasm is generally in a reducing state. Hence, BL21(DE3) *trxB* $^{-}$, a thioredoxin minus strain, in which the *TrxB* gene was mutated by the insertion of the Tn*kan*^{R} gene (Holmgren, 1984) was used in the present work as an attempt to create an artificial "oxidising environment" in the E. coli cytoplasm to improve scFv expressions.

2.1.7 Media for Hybridoma Cells Cultures.

Pre-prepared RPMI-1640 (with glutamine and bicarbonate) and DMEM (Modified Eagles Medium) liquid media were supplied directly by GIBCO-BRL. Other related media were prepared according to the recipes described in the laboratory manual published by Campbell (1991).

2.1.8 Media for Bacterial Cloning and Expression.

The LB liquid medium was prepared by dissolving 10 g tryptone, 5 g yeast extract, and 10 g NaCl in 900 ml water and the medium pH was adjusted to 7.5 with NaOH. The solution was made up to 1 litre with water and then sterilised by autoclaving at 121°C and 15 lb/in for 15 min. The LA solid medium was prepared like the LB liquid medium except that an additional 15 g of agar was added after the pH has been adjusted.

The SOB liquid medium was prepared by dissolving 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, and 186 mg KCl in 900 ml water, and then the medium pH was

adjusted to 7.0 with NaOH. The solution was then topped up with water to 1 litre followed by sterilisation by autoclaving. After autoclaving, the medium was allowed to cool to below 60 °C and then 10 ml of sterilised 2M $MgSO_4$ was added. The SOC liquid medium was prepared likewise for the SOB but an additional 10 ml of sterilised 50% glucose was added.

The NB liquid medium was supplied in a powdered form from Difco and was prepared according to the manufacturer's instructions.

2.1.9 Media for Plant Tissue Cultures.

MS (Murashige and Skoog, 1962) salts and vitamins were supplied as pre-prepared powder by Sigma. The MSO liquid medium was prepared by dissolving the MS salts in Q-H_2O with 30g/l of sucrose. The medium pH was then adjusted to pH 5.8 with 0.1 M HCl before it was sterilised by autoclaving. The MSD4x2 liquid medium was prepared like the MSO except with an additional of 0.1 mg/l of naphthalene acetic acid (NAA) and 1 mg/l of 6-benzylamino purine (BAP). The MSD4x2 solid medium was prepared by adding 15 g/l of media agar to the MSD4x2 liquid medium prior to autoclaving.

2.2 Protein Chemistry and Immunochemistry.

2.2.1 Quantification of Proteins.

Proteins were quantified using Bradford reagents as described by Bradford (1976) using the bovine serum albumin (BSA) as a standard

2.2.2 Polyacrylamide Gel Electrophoresis of Proteins.

The discontinuous, denaturing polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli (1970) was routinely used for the separation of denatured proteins. The acrylamide/bis-acrylamide gel was assembled using a Bio-Rad protein mini-gel kit.

Prior to electrophoresis, protein samples were mixed with 1/5 volume of 5X sample dye containing 500 mM Tris (pH 6.8), 10% SDS, 15% glycerol, 5% β-mercaptoethanol, and 0.05% bromophenol blue. The samples were then immersed in boiling water bath for 5 min to denature the proteins. After boiling, the fractions were spun at 10,000 xg for 5 min and the supernatants were used to load the gel. Electrophoresis was carried out at 120V until the bromophenol blue dye reached the other end of the gel. After electrophoresis, the gel was stained with Coomassie blue stain or proteins were transferred to a membrane by electroblotting (The Western Blot).

2.2.2.1 Coomassie Stain.

Coomassie staining of SDS-PAGE gel was carried out by immersing the gel in a staining solution containing 0.1% Coomassie blue R250, 50% methanol and 10% acetic acid for 30 min. The gel was then washed with distilled water and transferred into destaining solution containing 10% methanol and 10% acetic acid on an agitating platform at room temperature for overnight.

2.2.2.2 The Western blot.

Proteins were transferred from SDS-PAGE gels to nitrocellulose membranes which could then be probed with immunodetection systems. The Bio-Rad semidry electroblotting system was used to prepare the Western Blots and was applied according to the protocols provided by the manufacturer.

After blotting, the nitrocellulose was stained with Ponceau staining by immersing and agitating the membrane in a solution containing 0.2% Ponceau and 3% trichloroacetic acid for 2 min and then rinsing with distilled water. Protein bands were stained pink which in addition to staining major sample proteins also allowed the molecular weight markers to be located. Ponceau staining can be washed off with TBS (Tris buffered saline containing 10 mM Tris and 150 mM NaCl at pH 7.4) containing 0.05% Tween-20.

2.2.2.3 Immunodetection of Membrane Bound Protein.

Two immunodetection methods were used, namely the enzymic colorimetric and chemiluminescent systems. Both are basically the same except in the final step where the colorimetric method uses chromogenic enzymes and substrates to develop the stain whereas the chemiluminescent method uses chemiluminescent enzymes and substrates to produce a light signal that can be captured on film.

To carry out immunodetection, the membrane was first placed in a blocking buffer, the commonly used 4% milk powder in TBST (TBS with 0.05% (v/v) Tween-20), and agitated at room temperature for 30 min. The membrane was then exposed to the first antibody solution containing an appropriate dilution of a specific antibody targeting the protein of interest in TBST with 0.5 % of milk powder and agitating at room temperature for 2 hr or overnight at 4°C. The membrane was then washed three times for 10 min with TBST. The membrane was then agitated for 1 hr at room temperature in a secondary antibody solution containing an appropriate dilution in TBST with 0.5% milk powder of enzyme-conjugated antibody which acts against the first antibody. The membrane was then washed four times for 10 min with TBST and finally rinsed with distilled water. The membrane is then ready for the detection procedure.

In the colorimetric method, if the conjugated enzyme was alkaline phosphatase, then the substrate solution containing BCIP and NBT was used to produce a purple reaction product. The enzymic reaction was stopped by rinsing the membrane with running water for about 10 min. In the chemiluminescent method, the membrane was placed in a detection solution supplied by Boehringer according to the instructions given. The film was then processed using Kodak film developing solution.

2.2.3 Molecular Weight Determination of Protein.

A Sephadex G-75 gel filtration column (48 cm x 0.7 cm) was prepared to determine the molecular weight of native scFv proteins. Sample protein in 0.5 ml TBS

buffer was applied to the column and eluted using TBS buffer (pH 8.0) containing 0.1 % phenylalanine. The molecular weight of native scFv protein was determined by comparing the migration values (R_f) of standard proteins.

Blue dextran was used to determine the column void volume while molecular weight calibration was carried out using BSA (67 kD), chicken egg ovalbumin (43 kD), and denatured bacterial expressed scFv (30 kD). The migration value (R_f) was calculated using the following formula.

$$R_f = \frac{Vs - Vo}{Vb - Vo}$$

Where by Vs = Sample volume
Vb = Bed volume
Vo = Void volume

2.2.4 Enzyme Linked Immunosorbent Assay (ELISA).

The sandwich ELISA system was the most frequently used. Each well of the microtitre plate was coated with 100 µl of diluted ABA-BSA conjugate in PBS (Phosphate Buffered Saline, containing 0.2 g/l KCl, 1.44 g/l Na_2HPO_4, 0.24 g/l KH_2PO_4, and 8 g/l NaCl) at 4°C for overnight and the solution was then poured off. The ABA-BSA conjugate at a concentration of 250 µg/ml was used in the detection assays and 2.5 µg/ml was used in the competition assays. After antigen coating, the well surfaces were blocked with a solution containing 4% milk powder in PBS-Tween (PBS with 0.05% Tween-20) at room temperature for 30 min. After that, the blocking solution was poured off and the wells were washed twice with PBS-Tween. Each well was then subjected to 100 µl of the first antibody or test sample and incubated at room temperature for a minimum period of 2 hr. The solution inside the wells was then discarded and the wells were rinsed three times with PBS-Tween. An appropriate amount of enzyme labelled second antibody diluted in 100 µl of PBS-Tween was then added to the wells which were then incubated for a further 1 hr at room temperature. The solution was then drained and the wells were washed for five times using PBS-Tween.

The horse radish peroxidase-labelled secondary antibody conjugate was the most commonly used. The enzymatic reaction was carried out by adding 100 µl of the

substrate solution containing 0.04 mg/ml o-phenylenediamine, 0.01% H_2O_2, 50 mM Na-citrate, and 150 mM Na-phosphate (pH 6.0) to each well. During incubation, the microtitre plate was kept in the dark. The enzymatic reaction was terminated by adding 100 µl of 1M H_2SO_4 to each well and the optical density was measured at 490 nm.

The competition immunoassay was carried out similarly to the normal ELISA assay but with an additional step. For the antigen competition, the first antibody or test sample was mixed together with free abscisic acid ranges from 0 - 4.0 µM in the coated wells and incubated at 4°C overnight to reach the equilibrium. Whereas for the antibody competition, instead of using free abscisic acid, the tested antibody was used.

2.2.5 Antibody Affinity Determination.

The affinity constant of an antibody was measured according to the method developed by Friguet *et al.* (1985). The standard competition ELISA assays were set up over a range of free antigen using a limiting antibody. The competition ELISA assays were then allowed to reach the equilibrium overnight and bound antibody was measured as in the usual ELISA system.

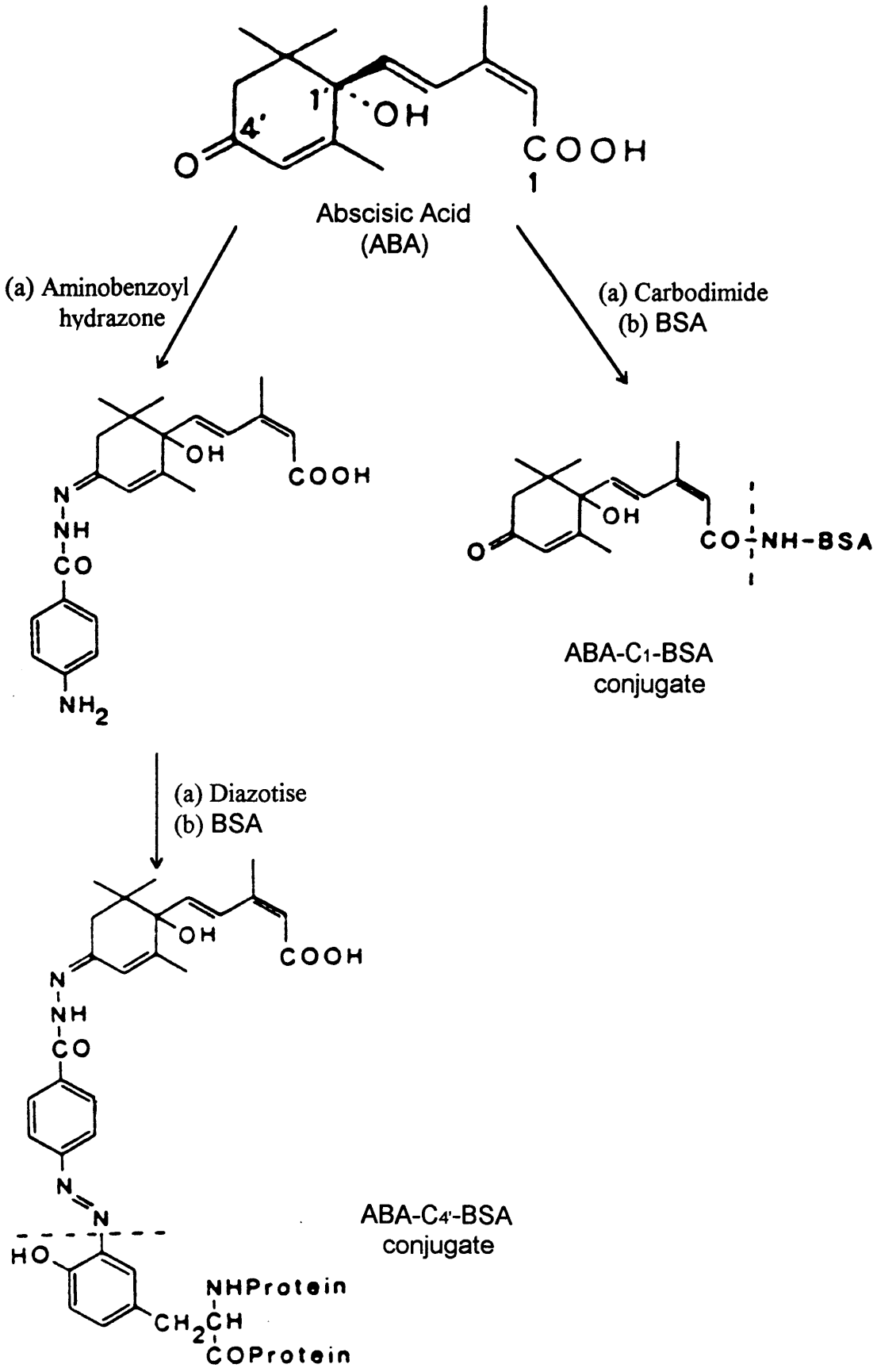
If the absorbance of the well with antibody alone is A_o (no competing antigen) and the absorbance at each concentration was computed $A_o/(A_o-A)$, the plot of the absorbance against 1/(substrate) should give a linear graph which intercepts at 1. The affinity constant of the antibody is given by the slope.

2.3 Monoclonal Antibody Preparation.

2.3.1 Synthesis of Abscisic Acid - Protein Conjugates.

Abscisic Acid (ABA) was conjugated to protein in two ways, through C_1 (-COOH) and $C_{4'}$ (>C=O). Commercially synthesised cis-(±)abscisic acid was used in the coupling reactions (Fig. 2.1).

Figure 2.1 : Preparation of abscisic acid - protein conjugates.



C_1 is the carboxyl functional group and was conjugated to protein amino groups according to the method described by Weiler (1980) using 1-ethyl-3-(3-dimethylaminopropyl) carbodimide. Two carrier proteins, bovine serum albumin (BSA) and chicken egg ovalbumin (EggAlb) were used for conjugating ABA at the C_1 position. The ABA-C_1-EggAlb was used for mouse immunisation while ABA-C_1-BSA was used in hybridoma selection and as antigen conjugate in immunoassays.

$C_{4'}$ is the carbonyl group and was conjugated to protein amino groups according to the method suggested by Quarrie and Galfre (1985) using p-aminobenzoylhydrazone. Only one type of $C_{4'}$ ABA conjugate, ABA-$C_{4'}$-BSA was prepared and used in immunoassays.

2.3.2 Immunisation and Hybridoma Preparation.

Immunisation was carried out using 100 μg of ABA-C_1-EggAlb in 100 μl PBS, emulsified with 100 μl of Freund's complete adjuvant (Sigma), and injected subcutaneously into Balb/c mice. Two booster injections were followed after 2 weeks interval but using Freund's incomplete adjuvant. The antibody responses were monitored by ELISA and a final injection was carried out intravenously using 100 μg conjugate in 100 μl PBS two weeks after the second booster injection.

Four days after the final injection, the spleen from the immunised mice was collected, fused immediately with SP2/0 myeloma cells according to the method described by Campbell (1990), and plated out to undergo media selection. The fusion culture was kept in a CO_2 incubator set at 37°C and 4% CO_2 and fresh media were fed accordingly. After 14 days, the culture was screened for positive hybridoma clones by ELISA using ABA-C_1-BSA conjugate as coating antigen. Two positive hybridoma clones, namely ABA15 and ABA26, were selected for further selection by cells dilution.

2.3.3 Hybridoma Cell Maintenance.

Sterile techniques and sterile environment are important for maintaining hybridoma cells. The mouse hybridomas, ABA15 and ABA26, were maintained in the RPMI complete media (RPMI-1640 with 10% FCS (foetal calf serum) and 1x P/S (Penicillin and Streptomycin)) at a population density of about 10^5-10^6 cells/ml in culture flasks and incubated at 37°C with 4% CO_2.

In addition to the mouse hybridomas, two rat hybridoma cell lines, MAC61 and MAC252, which produced antibodies against ABA-$C_{4'}$-conjugate were obtained from the European Collection of Animal Cells Cultures (ECACC) and included in this study. The rat hybridomas were maintained in the DMEM complete media (DMEM with 10% FCS and 1x P/S) under the same conditions as the mouse hybridomas.

For hybridoma cell storage, cells growing at logarithmic stage were collected by centrifugation at 400 xg for 5 min. The pellet was then resuspended in 90% FCS and 10% DMSO at a cell density of 5 x 10^6 cells/ml. Aliquots of 1.0 ml of cell suspension were then placed into sterile freezing ampoules and kept at -80°C for 1 day before transferring into liquid nitrogen for long term storage.

Frozen hybridoma cells were thawed by incubating the ampoule in a 37°C water bath until it almost turned into liquid. The contents were then transferred into a small culture flask containing 5 ml complete media pre-warmed to 37°C and the flask was then incubated in the CO_2 incubator for 24 hr before subculturing.

2.3.4 Monoclonal Antibody Production.

Monoclonal antibodies were produced by transferring logarithmic growth-stage cells to the RPMI media or DMEM media without FCS and allowing overgrowth for 2 days before the culture supernatant was collected by centrifugation. The collected culture supernatant contains monoclonal antibody and can be used immediately. Proteinase inhibitor such as sodium azide (0.01%) was normally added to prevent bacterial or fungal contamination. For routine usage, culture supernatants containing

monoclonal antibodies were kept at 4°C whereas for long term storage they were kept at -80°C.

2.4 Molecular Cloning.

2.4.1 General Molecular Cloning Techniques.

Most of the methods used for cloning and manipulating DNA fragments such as the extraction of DNA, quantification of DNA and RNA, small-scale or large-scale of plasmid DNA isolation, restriction enzyme digestion, dephosphorylation of DNA fragments, DNA agarose gel electrophoresis, preparation of blunt-ended DNA fragments, and ligation of DNA fragments were carried out according to the laboratory manual by Sambrook *et al.*, (1989).

2.4.2 Partial Restriction Enzyme Digestion.

Partial restriction enzyme digestion was sometimes necessary when a fragment has an internal restriction enzyme site which can be cut by the same enzyme. Partial restriction enzyme digestion was carried out using 3 times more DNA than the usual digestion in a standard 20 µl reaction. The normal amount of restriction enzyme was added and the mixture was incubated. At 5 min intervals, 5 µl of the reaction mixture was transferred into a tube containing 1 µl of 0.5 mM EDTA on ice. After all the reaction mixture was transferred, the mixture was analysed by agarose gel electrophoresis and the DNA fragments of the correct size were collected.

2.4.3 Polymerase Chain Reaction.

The polymerase Chain reaction (PCR) was carried out to amplify DNA fragments from DNA templates using downstream and upstream complementary sequence primers. In a standard 50 µl PCR reaction, 50 - 100 ng of DNA template was mixed with 50 pmol of upstream primer, 50 pmol of downstream primer, 25 nmol each of dNTPs, 1 unit of *Taq* polymerase and 1x PCR buffer containing 50 mM Tris

(pH 8.8), 10 mM ammonium sulphate, 4 mM $MgCl_2$, 10 mM β-mercaptoethanol, and 0.1 mg/ml BSA. The PCR reaction mixture was made up in a 0.5 ml PCR tube and overlaid with one drop of mineral oil to prevent any evaporation. After a brief centrifugation, the tube was put on a Perkin-Elmer Cetus Thermocycler machine to carry out the PCR reaction by denaturing the template DNA at 94°C for 3 min, followed by 25 reaction cycles of denaturation at 94°C for 1 min, annealing at 55°C for 45 sec and extension at 72°C for 1 min, and finally an extended reaction at 72°C for 5 min. When the reactions were completed, 5 μl of the product was analysed by electrophoresis and visualised with ethidum bromide stain under UV illumination.

Some adjustments to the annealing temperature and the number of reaction cycles were sometimes necessary with different types of DNA templates or primers.

2.4.4 DNA Purification by DNA Matrix.

DNA fragments of enzymes digested mixture or PCR products were purified using a DNA matrix. The purification was performed using DNA matrix (glass milk) purchased from Bio-Rad according to the procedures suggested but with some modification. Samples containing DNA fragments were normally applied to agarose gel electrophoresis to separate the desired DNA fragments. The DNA fragments were cut out from the gel and redissolved using gel resolving solution at 60°C. The mixture was then cooled in ice for a few minutes before it was mixed with an appropriate amount of DNA matrix and left for 5 min. The DNA matrix was then separated by centrifugation and the pellet washed once with gel dissolving solution, and twice with 80% ethanol. After that, the DNA matrix was dried completely under vacuum and the DNA fragments were eluted from the matrix using a small amount of Q-H_2O.

2.4.5 DNA Transformation of *E. coli.*

Routine *E. coli* transformations for manipulation of plasmid DNA were carried out according to the method described by Nishimura *et al.* (1990) with minor modifications. Basically, bacteria were allowed to grow in 50 ml culture medium

(LB media supplemented with 10 mM $MgSO_4$ and 0.2% glucose) to mid logarithmic phase (OD_{600} = 0.4-0.7) in a shaking incubator at 37°C before the cells were harvested by centrifugation at 2500 xg for 5 min at 4°C. The cell pellet was then resuspended gently in 0.5 ml culture medium precooled on ice, then mixed without vortexing with 2.5 ml of storage medium (36% glycerol, 12% PEG (M.W. 6000), 12 mM MgSO4 in LB and sterilised by filtration). The competent cells were divided in aliquots of 100 µl each in microcentrifuge tubes and stored at -80°C until needed.

For transformation purposes, the frozen cells were thawed, mixed immediately with up to a maximum of 10 µl of ligation mixture, and incubated on ice for 15 - 30 min. The mixture was then subjected to a heat pulse at 42°C for 60 sec, then chilled on ice for 1 to 2 min, diluted with 1.0 ml of culture medium, then incubated for 1 hr at 37 °C in a shaking bath. The mixture (10 to 200 µl, depending on the DNA amounts used in transformation) was then plated on an agar plate containing antibiotic to allow selection of transformants. The efficiency of this transformation method was found to be higher than the conventional Ca^{2+} method.

2.4.6 Bacterial Colony Screening.

The screening process used to confirm a correct insertion of the required DNA plasmid in bacterial colonies was normally carried out by PCR. Bacterial colonies, normally after bacterial transformation with recombinant DNA, were picked from overnight plates with pipette tips and transferred separately into 0.5 ml PCR tubes and were also inoculated onto media plates for future recovery. The bacterial samples were then mixed with 10 µl of standard PCR reaction mixture, overlaid with mineral oil, and centrifuged before PCR were carried out in a Perkin-Elmer Cetus Thermocycler machine. The PCR procedure involved pre-soaking at 94°C for 5 min, followed by 20 reaction cycles (94°C for 1 min, 55°C for 45 sec, and 72°C for 1 min), and finally an extended reaction at 72°C for 5 min. After the PCR reaction, 5 µl of the reaction mixtures were analysed by agarose gel electrophoresis.

In order to increase the PCR specificity and allow the inserted orientation to be checked, PCR screening normally utilised one primer complementary to the vector sequence and the other complementary to the inserted DNA.

2.5 ScFv Construction.

2.5.1 RNA Isolation from Hybridoma Cells.

A hot phenol method was used to extract total RNA from hybridoma cell lines ABA15, ABA26, MAC61, and MAC252. The hybridoma cells from 30 ml cultures were collected by centrifugation at 400 xg for 5 min. The cell pellets were transferred to clean 1.5 ml microcentrifuge tubes and vortexed for 30 sec with 0.8 ml of hot (preheated at 80°C) extraction buffer containing 50% phenol and 50% of 100 mM LiCl, 100 mM Tris (pH 8.0), 10 mM EDTA, and 1% SDS. After that, 0.4 ml of chloroform : isoamyl alcohol (24:1) was added and vortexed for another 30 sec. The mixture was then centrifuged at 10,000 xg for 5 min. Following centrifugation, the aqueous phase was transferred to a new tube and then 1 volume of 4M LiCl was added. The mixture was then well-mixed and stored at 4°C for overnight.

On the second day, the tube was centrifuged at 10,000 xg for 10 min and the aqueous phase discarded. The pellet was then resuspended in 250 µl of suspension buffer containing 100 mM Tris (pH 8) and 0.2% SDS. The suspension solution was then transferred quickly to a new tube and immediately mixed well with 25 µl of 3M sodium acetate followed by 275 µl of isopropanol. The mixture was then spun at 10,000 xg for 10 min to collect the RNA pellet. The pellet was then washed once with 70% ethanol, dried, and resuspended in 50 µl of RNase free water. The absorbance at 260 nm was measured and used to calculate RNA concentration. 5 µl of the RNA solution was then analysed by agarose gel electrophoresis and the remaining RNA solution was stored at -80°C until needed.

2.5.2 cDNA Synthesis.

First-strand cDNAs of ABA15, ABA26, MAC61, and MAC252 RNA were synthesised using Pharmacia's first-strand cDNA synthesis kit according to the protocol given. Each synthesis reaction involved 5 μg of total RNA obtained from hybridoma cells and oligo-dT was used to prime first-strand synthesis.

2.5.3 Amplification of V_L and V_H domains.

The V_H and V_L domain genes were obtained by PCR amplification using the first-strand cDNA as templates. The following degenerate primers were used to amplify the gene sequences and introduce restriction enzyme sites (underline letters) at the both ends.

Light Chain primers : (L32 was 3'-end primer and others were 5'-end primers.)

L51: 5'-ATACCATGGCAGAC(G/A)TC(C/A)AGAT(G/A)A(T/C)CCAG(T/A)CT(C/A)CA-3'

L52: 5'-ATACCATGGCA(G/A)A(C/A)ATT(G/T)TGCTGAC(T/C)CA(G/A)T(C/T)TCC-3'

L32: 5'-CTACCCGGGTCTGATTTCCAGCTTGGTGCC-3'

Heavy Chain primers : (H33 was 3'-end primer and others were 5'-end primers.)

H51: 5'-AATTGTCGACCAG(C/T)TG(C/G)(A/T)GCA(G/A)TCTGGA-3'

H52: 5'-AATTGTCGACCA(G/A)(C/G)(T/C)GCAGCA(G/A)(C/T)CTGGG-3'

H53: 5'-AATTGTCGAC(C/A)AGCTTCAGGAGTC(A/G)GGACC-3'

H54: 5'-AATTGTCGACA(A/G)G(C/G)TG(G/A)T(G/C)GAGTCTGGAGG-3'

H55: 5'-AATTGTCGACA(A/G)(G/C)(C/G)TGG(T/A)GGAATCTGGAGGA-3'

H56: 5'-AATTGTCGACCAGCTGAAGG/CA(G/C)TCAGGACC-3'

H33: 5'-ACTACTCGAGGGG(A/G)CCA(A/G)(G/T)GGATA(A/G)AC-3'

The light chain primers would introduce the *Nco* I restriction enzyme site at the 5'-end while the *Sma* I site at the 3'-end of the V_L domain genes. The heavy chain primers would introduce the *Sal* I site at the 5'-end and the *Xho* I site at the 3'-end of the V_H domain genes.

In order to obtain a suitable set of primers for the V_L and V_H domain genes amplification, a set of PCR reactions were prepared using different combinations of 5'-end and 3'-end primers. These included L51/L32 and L52/L32 for the V_L domain genes amplification and H51/H33. H52/H33, H53/H33, H54/H33, H55/H33, and H56/H33 for the V_H domain genes amplification. The PCR amplifications were carried out using 0.5 μl of first-strand cDNA synthesis product as DNA templates in a

standard 100 µl of PCR reaction mixture. The mixture was then channelled to undergo PCR procedures by pre-soaking at 94°C for 3 min, followed by 20 reaction cycles (94°C for 1 min, 55°C for 45 sec, and 72°C for 1 min), and finally an extended reaction at 72°C for 5 min. Upon completion of the reaction, the PCR products were analysed by agarose gel electrophoresis.

PCR products of the appropriate sizes were then cloned directly into the pCR vector purchased from Invitrogen, according to the protocols given. After transformation into *E. coli* XL-1 Blue, at least three colonies were selected to isolate the plasmid DNA and then verified by restriction enzymes. Automatic DNA sequencing was then carried out on the selected plasmids to obtain the sequence for the V_L and V_H domains of ABA15, ABA26, MAC61, and MAC252.

2.5.4 Automatic DNA Sequencing.

The dye-terminator cycle sequencing method was used to determine the DNA sequences of the V_L and V_H domains in pCR vectors and other DNA fragments. Dye-terminator cycle sequencing was carried out according to the manufacturer's instructions (Applied Biosystems). The template DNA used was plasmid DNA or the PCR product of genome DNA. Typically, 1 µg plasmid DNA was used as template in 20 µl reaction mixtures. The PCR reaction involved 25 cycles of 96°C for 30 sec, 50°C for 15 sec, and 60°C for 4 min. After the reaction was completed, the PCR product was purified by phenol/chloroform extraction and precipitated by 2 volumes of absolute ethanol. The DNA pellet was then analysed by using an Automated DNA Sequencer in the Protein and Nucleic Acid Chemistry Laboratory of the Biochemistry Department, Leicester University. Sequence analysis was then performed using SeqEd (Applied Biosystems).

2.5.5 5'-V_L-linker-V_H-tag-3' scFv Assembly.

The scFv genes of ABA26, MAC61, and MAC252 were assembled by cloning. The V_L and V_H domain gene fragments in pCR vectors were obtained by restriction

enzyme digestion and assembled together with a linker between the V_L and V_H domains and a HSV-tag detection signal sequence into pGEM11Zf(-) cloning vector to form 5'-V_L-linker-V_H-tag-3' scFv gene fragment. The linker used was a sequence which encoded 15 amino acids of $(GGGGS)_3$ and the HSV-tag sequence encoded the amino acids OPELAPEDPED. The linker and HSV-tag were previously produced through oligonucleotide synthesis.

In addition, PV4 and PV5, two similarly designed scFv genes derived from monoclonal antibodies against potato virus Y, were obtained from Dr. Chen, Botany Department, Leicester University, and incorporated into certain parts of the study.

2.6 Expression of scFv proteins in *E. coli.*

2.6.1 Bacterial Expression System.

Several expression systems were used for scFv genes expression in *E. coli*. These included the expression systems driven by *lac* or T7 promoters, with the scFv proteins being either localised in the cytoplasm or secreted into the periplasm. The following expression vectors were used.

- pTrcHis expression vector : This vector is a cytoplasmic expression vector driven by the *trc* promoter.
- pRSET expression vector : This vector is a cytoplasmic expression vector driven by the T7 promoter.
- pABE expression vector : This vector is driven by the *tac* promoter and contains *pelE* signal peptide which secretes the recombinant protein into the periplasm.
- pTM-22 expression vector : This vector is driven by the T7 promoter and contains *ompA* signal peptide which secretes the recombinant protein into the periplasm.

The scFv genes were transferred from pGEM11Zf(-) cloning vector into a particular expression vector by molecular cloning. The cloning was first carried out in XL-1 Blue *E. coli* cells but constructs were subsequently transformed into other *E. coli* strains such as JM109, JM109(DE3), W3110wt, BL21(DE3) *TrxB*$^-$ for the

expression studies. Prior to bacterial expressions, the construct were verified by restriction enzyme and PCR analysis.

2.6.2 Bacterial Expression Protocols.

Two types of expression protocols were used depending on the type of expression vector and bacteria strain used.

2.6.2.1 Expression by Direct Induction.

This expression protocol was used for the expression system which involved the use of the *lac* or *trc* promoter in all the *E. coli* strains and the expression with the T7 promoter in *E. coli* strain JM109(DE3).

Bacterial expression was carried out using 100 μl of an overnight culture to inoculate 10 ml of culture media containing 0.2% glucose and 50 μg/ml ampicillin. The LB media were used for expressions by pABE and pTM-22 vectors while the SOB media were used for expressions by pTrcHis and pRSET vectors. The inoculated culture was kept at 37°C with vigorous shaking until it reached an OD_{600} about 0.5, at which time IPTG was added to a final concentration of 3 mM. The culture was allowed to grow for an additional 3 hr under the same conditions after which the cells were collected for analysis.

Variations in these methods included the use of different media recipes, incubation temperatures, IPTG concentrations, and cell harvest times.

2.6.2.2 Expression by M13 Phage Infection.

This expression protocol was used for the expression with the T7 promoter in *E. coli* strains such as the XL-1 Blue which does not carry the T7 RNA polymerase gene but can be infected by the M13/T7 phage.

In this method, 100 μl of an overnight culture was used to inoculate 10 ml of culture media containing 0.2% glucose and 50 μg/ml ampicillin. The LB media were

used for expressions by pTM-22 while the SOB media were used for expressions by pRSET. The culture was kept at 37°C with vigorous shaking until an OD_{600} about 0.3 was reached and IPTG was added to a final concentration of 1 mM. The culture was then grown under the same conditions for an additional 1 hr. The OD_{600} was then measured and the number of cells was calculated based on the fact that 10^6 cell/ml has an OD_{600} of 1.0. An amount of M13/T7 phage equivalent to 5 pfu/cell was then added and the mixture was allowed to grow under the same conditions for an additional 5 hr before the cells were collected for analysis. Variations in these methods included the use of different media recipes, incubation temperatures, IPTG concentrations, and cells harvest times.

The M13/T7 phage solution was prepared by infecting XL-1 Blue *E. coli* cells in the logarithmic growth phase and leaving the culture to grow overnight. The culture supernatant was then collected and the phage titre was estimated. This supernatant was used as the M13/T7 stock solution.

2.6.3 Bacteria Protein Extraction.

The bacterial cells were collected by centrifugation at 4°C at 4000 rpm for 5 min. The cell pellet was then resuspended in 1/25 original volume of extraction buffer containing 20 mM Tris (pH 7.4), 1 mM EDTA, 200 mM NaCl and protease inhibitor at a concentration of 1 μg/ml of leupeptin, 2 μg/ml of aprotinin and 100 μg/ml PMSF. The mixture was then placed in an ice bath and sonicated for three short pulses of 20 sec with 10 sec interval breaks. After sonication, the mixture was centrifuged at 10,000 xg for 10 min. The supernatant which represents the crude soluble extract, was transferred to a new tube. The pellet was then washed once with extraction buffer, spun, resuspended again in the original volume of extraction buffer, and treated as an insoluble fraction. Both protein fractions were used immediately for analysis or kept at -20°C for long term storage.

2.7 Yeast Expression Systems.

2.7.1 Plasmid Construction.

The *Pichia pastoris* yeast expression system was obtained from Invitrogen. Most of the experimental procedures were carried out according to the protocols suggested by the manufacturer. Two expression vectors, pPIC9K and pPICZα, were used in this study. The scFv genes of ABA26, MAC61, MAC252, and PV4 from pGEM11Zf(-) were fused in frame with the signal peptide sequences in pPIC9K and pPICZα by molecular cloning.

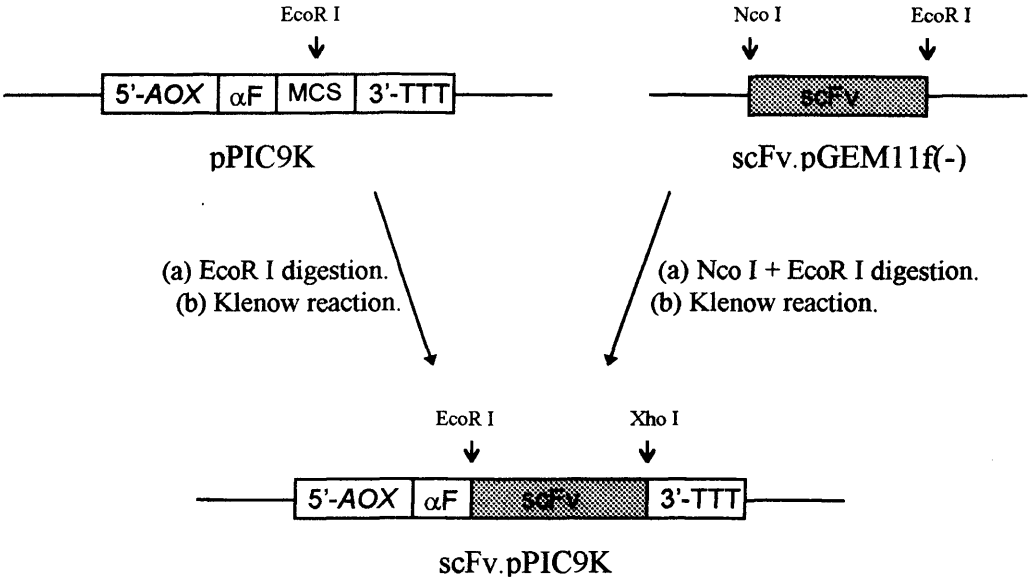
To ensure that the constructs were in the correct frame sequence in pPIC9K, the scFv genes from pGEM11Zf(-) were digested by *Nco* I and *Eco*R I, and the sticky ends were then blunted by Klenow fragment reaction followed by the cloning into pPIC9K vector which was predigested by *Eco*R I and the sticky ends were filled in by Klenow fragment reaction. In this cloning, the correct scFv insertion can be verified by the restriction enzyme digestion as the DNA ligation between blunted *Nco* I digested ends (from scFv) and blunted *Eco*R I digested ends (from pPIC9K vector) would create an *Eco*R I site (Fig. 2.2). Therefore, the correct construct can be verified by *Eco*R I digestion.

While in the pPICZα constructs, the scFv genes (without the HSV-tag tail) from pGEM11Zf(-) were digested by *Nco* I and *Xho* I, the sticky digested ends blunted by Klenow fragment reaction, and cloned into pPICZα which was predigested with *Eco*R I and *Not* I followed by the filling-in of the sticky digested ends by Klenow fragment reaction. In this cloning, the correct scFv insertion can also be verified by restriction enzyme digestion as the DNA ligation between blunted scFv and vector would create an *Eco*R I site at the 5'-end and *Xho* I site at the 3'-end of scFv (Fig. 2.2). Therefore, the correct construct can be verified by both the *Eco*R I and *Xho* I digestion.

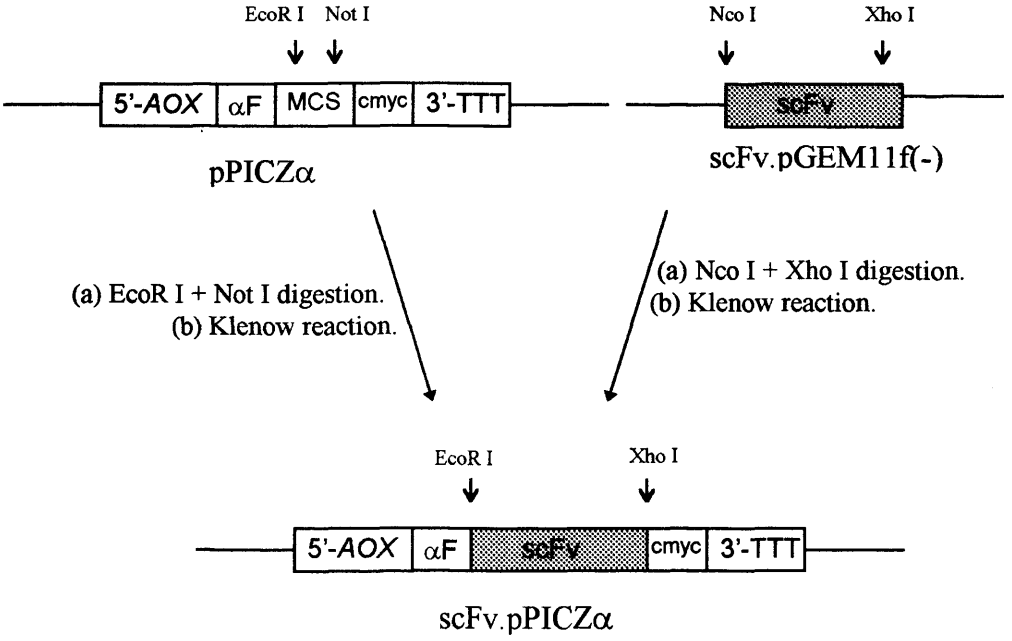
After the constructs had been verified by PCR and restriction enzyme digestion, the pPIC9K and pPICZα derivatives were transformed into *Pichia pastoris* strains, GS115 and KM71 by electroporation. After selection, the integration of the scFv

Figure 2.2 : Constructions of scFv gene in *Pichia* expression vectors.

(a) scFv.pPIC9K genes construction.



(a) scFv.pPICZα genes construction.



Key : *5'-AOX* = 5' AOX1 gene promoter.
α F = α Factor signal sequence.
cmyc = C-myc tag detection signal sequence.
3'-TTT = 3' AOX gene transcription termination sequence.
MCS = Multiple cloning sites.

genes into yeast genome was verified by PCR as described for the screening of bacterial colonies.

2.7.2 Expression and Protein Analysis.

Several positive yeast clones of each construct were selected for expression studies. A small scale expression was carried out by inoculating 10 ml aliquots of yeast media (without glucose but containing glycerol) with several yeast colonies followed by growth in a shaking incubator at 29°C overnight. The yeast cells were collected the following day by centrifugation at 2500 xg for 5 min. The pellet was then resuspended in glucose-free yeast media containing 0.5% methanol at an OD_{600} = 2 - 6 to induce expression. The yeast culture was then incubated in a shaking incubator at 29°C for 24 hr before the samples were collected for analysis.

To carry out total protein analysis, the induced yeast culture was mixed with SDS-polyacrylamide gel buffer and boiled for 5 min. The sample was then centrifuged at 10,000 xg for 10 min. The supernatant (termed the total protein sample) was then transferred to a new tube and ready for loading onto SDS-PAGE.

The induced yeast culture supernatant (termed the soluble protein fraction) was obtained by centrifuging the induced yeast culture at 10,000 xg for 5 min to remove the yeast cells and any other insoluble material.

2.8 Plant Expression Systems.

2.8.1 Plasmid Constructions.

The following derivatives of scFv constructs were used for the expression of scFv genes in plant.

- scFv (scFv only, expressed at cytoplasm.)
- SS.scFv (scFv with PR1-a signal sequence (SS), expressed in apoplast.)
- TP.scFv (scFv with chloroplast targeted peptide (TP).)

- SS.scFv.KDEL (scFv with signal sequence (SS) and KDEL short peptide for ER retention.)

The pJIT117 vector obtained from the John Innes Institute was used to construct the scFv derivatives. The pJIT117 vector contained double 35S promoters which were followed by TP sequence, multiple cloning sites and poly-A sequence (Guerineau *et al.*, 1988). The scFv sequences of ABA26, MAC61, and MAC252 from the pGEM11Zf(-) cloning vectors were transferred by cloning into pJIT117 to form TP.scFv derivatives (Fig. 2.3), known as TP.scFv26, TPscFv61, and TP.scFv252 respectively (a similar naming system was used for other derivatives).

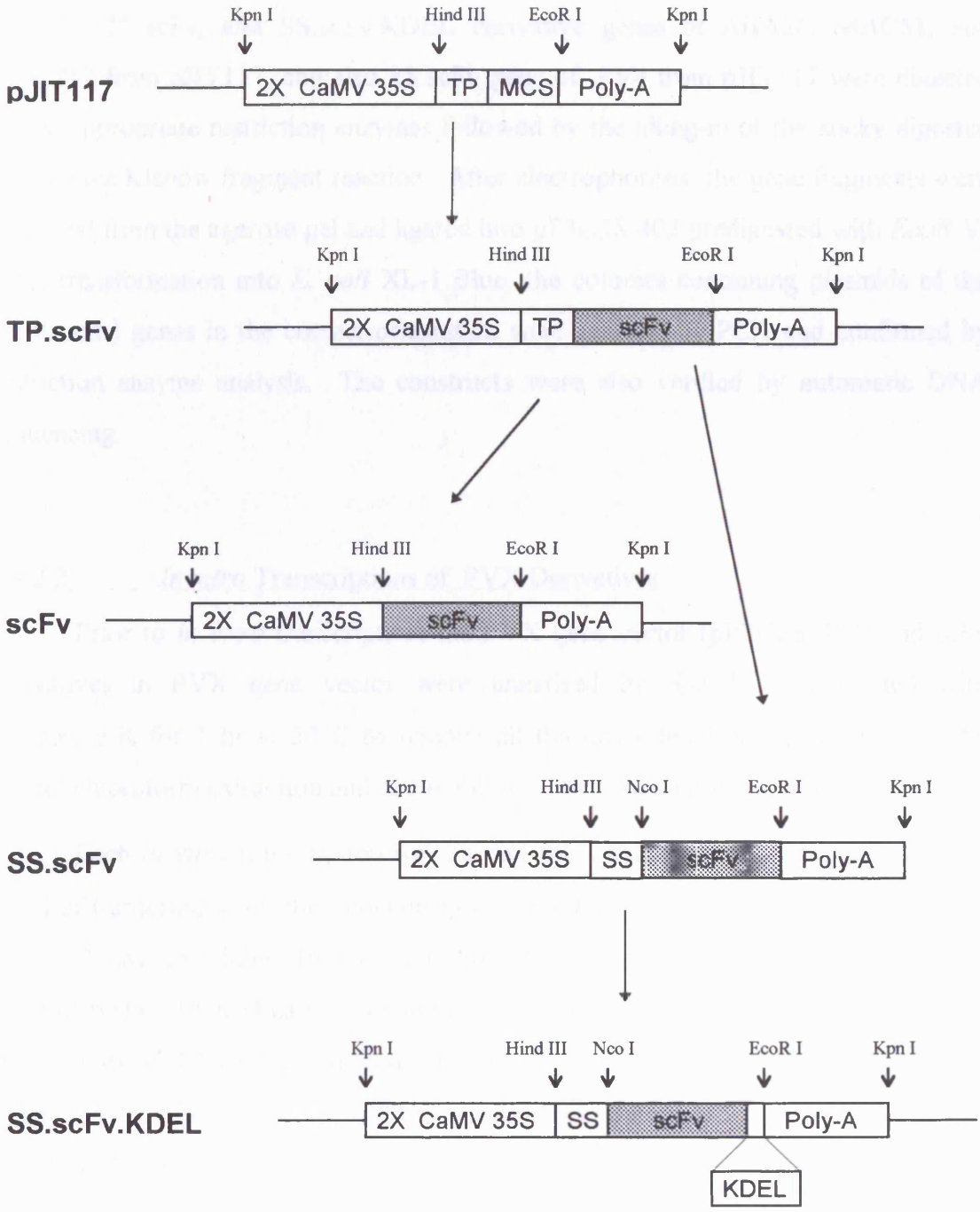
The TP sequence in TP.scFv derivatives was then removed to produce a construct encoding only the scFv (Fig.2.3). The SS.scFv derivative (Fig. 2.3) was constructed by replacing the TP sequence in the TP.scFv construct with a tobacco PR1-a signal sequence obtained from Firek *et al.* (1993). The SS.scFv.KDEL derivative was produced by synthesising the oligonucleotides for the KDEL sequence and ligating it into the 3'-end of SS.scFv construct in the pJIT117 vector (Fig. 2.3). A SS.scFv construct for PV4 antibody was also created by replacing the scFv sequence of MAC61 with the scFv sequence of the PV4 antibody.

The above cloning steps were carried out in *E. coli* XL-1 Blue, and the joining sites of those constructs were verified by restriction enzymes or by automatic DNA sequencing if the restriction enzyme site was destroyed after joining.

2.8.2 Transient Expression System.

The transient expression of various derivatives of scFv genes in plant was carried out using potato virus X (PVX) as a gene vector. The PVX gene vector, pP3C2S 402, containing a full-length cDNA of PVX with an additional initial sequence of coat protein gene for gene insertion (Chapman et al. 1992), was obtained from the Scottish Crop Research Institute.

Figure 2.3 : Construction of various scFv derivatives for plant expression.



Key : 2X CaMV 35S = Double CaMV 35S promoters.
TP = Chloroplast targeted peptide sequence.
SS = Tobacco PR1-a signal peptide sequence.
Poly A = Transcription termination sequence.
KDEL = ER retention peptide sequence.
MCS = Multiple cloning sites.

2.8.2.1 Construction of PVX Derivatives.

The scFv genes of ABA26, MAC61, and MAC252 from pGEM11Zf(-), and SS.scFv, TP.scFv, and SS.scFv.KDEL derivative genes of ABA26, MAC61, and MAC252 from pJIT117, and also SS.scFv gene of PV4 from pJIT117 were digested by the appropriate restriction enzymes followed by the filling-in of the sticky digested ends by the Klenow fragment reaction. After electrophoresis. the gene fragments were extracted from the agarose gel and ligated into pP3C2S 402 predigested with *EcoR* V. After transformation into *E. coli* XL-1 Blue, the colonies containing plasmids of the constructed genes in the correct orientation were selected by PCR and confirmed by restriction enzyme analysis. The constructs were also verified by automatic DNA sequencing.

2.8.2.2 *In vitro* Transcriptions of PVX Derivatives.

Prior to *in vitro* transcription, the PVX gene vector (pP3C2S 402) and scFv derivatives in PVX gene vector were linearised by *Spe* I, then treated with proteinase K for 1 hr at 37°C to remove all the unwanted RNase, and purified by phenol/chloroform extraction and finally followed by ethanol precipitation.

Each *in vitro* transcription was carried out using 5 μg of linearised DNA in 50 μl of transcription reaction containing 40 mM Tris (pH 8.0), 25 mM NaCl, 8 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 2 mM of each rATP, rUTP and rCTP, 0.2 mM rGTP, 0.5 mM cap analogue $m^7G(5')ppp(5')G$, 40 units of RNase inhibitor, and 40 units of T7 RNA polymerase. The reaction was carried out at 37°C for 20 min before the addition of 100 mM GTP to a final concentration of 2 mM. After a further 40 min of incubation at 37°C, another 20 units of T7 RNA polymerase were added and the incubation was carried on at 37°C for another 1 hr. Upon completion of the reaction, 5 μl of the reaction mixture was analysed by gel electrophoresis while the remainder was purified by phenol/chloroform extraction followed by ethanol precipitation. The dried pellet was then stored at -80°C until needed.

2.8.2.3 Plant Inoculations.

The *in vitro* transcription products of PVX and PVX-scFv derivatives were resuspended in 30 µl of RNase-free water containing 50 mM sodium phosphate and 5 mg/ml bentonite. The inoculation involved the scraping of 10 µl of the transcribed solution onto several leaves of the *N. clevelandii* plant, predusted with carborundum. Those inoculated plants were kept in moist conditions overnight by covering them with wet paper towels. The plants were then transferred to a growth cabinet providing 16 hr light period at 25°C and 8 hr dark period at 20°C. The plants were observed daily and samples were taken for protein analysis after virus infection symptoms appeared.

2.8.3 Stable Plant Transformations.

Agrobacterium-mediated transformation based on the methods described by Draper *et al.*(1988) was applied with slight modifications to stabilise the transformation of the scFv derivatives of ABA26, MAC61, MAC252, and PV4 into the tobacco plants.

2.8.3.1 Construction of pBin 19 Derivatives.

The binary vector pBin 19 was used for *Agrobacterium*-mediated stable plant transformation. The scFv derivatives together with the double 35S promoters, and poly-A sequences from pJIT117 were digested by *Kpn* I and cloned into pBin 19 predigested with *Kpn* I. The cloned constructs were also verified by PCR and restriction enzymes.

2.8.3.2 *Agrobacterium* Transformations.

The plasmids of pBin 19 derivatives were transformed directly into the competent cells of *Agrobacterium tumefaciens* strain LB4044. Those competent cells were prepared by inoculating 50 ml of NB media with 2.0 ml of an overnight culture

of *Agrobacterium*. The culture was then left to grow for about 4 hr in a shaking incubator at 29°C to a mid logarithmic phase. The cells were then collected by centrifugation at 2500 xg for 5 min and resuspended in 0.5 ml NB media containing 36% glycerol and 12% PEG (M.W. 6000). After that, the culture was divided into 100 µl aliquots which were used immediately.

For *Agrobacterium* transformation, about 5 µg of DNA in 10 µl water was mixed with 100 µl of *Agrobacterium* competent cells in a microcentrifuge tube, the lid was closed, and the tube was immediately put into liquid nitrogen . The tube was then subjected to heat shock at 37°C for 5 min. After that, 1.0 ml of NB was added and the tube was incubated at 29°C in a shaking incubator for 2 to 4 hr before being plated onto antibiotic plates for selection of transformed colonies which appeared after 2 to 3 days.

2.8.3.3 Plant Transformations.

Stable plant transformation was carried out according to the laboratory manual described by Draper et al. (1988) with minor modifications. Sterile leaf explants derived from 3-6 weeks old *Nicotiana tabacum* (SR1) were inoculated for 15 min with an overnight culture of transformed *Agrobacterium tumefaciens* prediluted with 10 times of MSO media and then subcultivated on MSD4x2 media for 2 days. Inoculated leaf discs were then transferred to the selection media containing MSD4x2 with 100 mg/l kanamycin and 400 mg/l augmentin and kept in a growth room providing 16 hr light period at 25°C and 8 hr dark period at 20°C. After several weeks, regenerated shoots were cut off and put into rooting media consisting of MSO media with 100 mg/l kanamycin, and 400 mg/l augmentin. Plantlets were then transferred to soil after roots developed and finally grown in green house for 16 hr light period at 25°C and 8 hr dark period at 20°C.

2.8.4 Analysis of Plant Protein Extracts.

The extraction of the total plant protein was carried out by homogenising 1 to 2 g of fresh plant tissue at 4°C with an equal volume of SDS-polyacrylamide gel sample buffer supplemented with 1 mM EDTA and protease inhibitor consisting of 1 μg/ml of leupeptin, 2 μg/ml of aprotinin, and 100 μg/ml PMSF. The sample was then boiled for 5 min and the insoluble materials were removed by centrifugation at 10,000 xg for 10 min. The supernatant (termed as total protein extract) was transferred to a new tube, ready for loading onto SDS-PAGE.

The soluble plant protein extract was obtained by homogenising 1 to 2 g of fresh plant tissue with an equal volume of extraction buffer containing 100 mM Tris (pH 7.5), 5 mM EDTA, 20 mM Na_2SO_3, 1 μg/ml of leupeptin, 2 μg/ml of aprotinin, and 100 μg/ml PMSF. The sample was then centrifuged at 10,000 xg for 10 min to remove any insoluble materials. The supernatant (termed as soluble protein extract) was then transferred to a new tube and stored at -20°C if it was not needed immediately.

2.8.5 Plant DNA Isolation for PCR.

The method for isolating plant genomic DNA for PCR analysis was modified from Edwards *et al.* (1991). Plant tissue sample, normally leaf , was collected using the lid of a sterile microcentrifuge tube to pinch out a disc of the leaf into the tube. The sample was then ground with 400 μl of extraction buffer containing 200 mM Tris (pH 7.5), 250 mM NaCl, 25 mM EDTA, and 0.5% SDS using a plastic rod. An equal volume of phenol/chloroform was then added and the mixture was vortexed for 30 sec followed by centrifugation at 10,000 xg for 5 min. 200 μl of the supernatant was then transferred to a new tube, mixed well with 200 μl of isopropanol and left on ice for 5 min. The mixture was then centrifuged at 10,000 xg for 5 min and the pellet was dried under vacuum conditions. The pellet was then redissolved in 50 μl of TE buffer and 1 μl of this sample was normally sufficient for a 20 μl PCR reaction.

2.8.6 RT-PCR of Plant RNA.

Plant RNA miniprep was isolated using "RNeasy Plant" kit (Qiagen) according to the protocol suggested by the manufacturer. The purified plant RNA was then treated with RNase-free DNase I at 37°C for 15 min, extracted once with phenol/chloroform and then precipitated by isopropanol.

The reverse transcription reaction was carried out using Pharmacia's first strand cDNA synthesis kit according to the manufacturer's protocol except that a smaller reaction volume was used. 1 μg of DNase-treated plant RNA and the H33 primer was used in a total of 3 μl of the first strand cDNA synthesis reaction. Upon completion of the reverse transcription reaction, 1 μl of the reaction product was used to carry out a standard 50 μl PCR reaction.

RT-PCR for yeast RNA was also carried out following the same procedures.

2.9 Antibody Variants.

2.9.1 Single V_H or V_L Domain.

The construct of single V_L domain of MAC61 was obtained by removed the V_H domain of MAC61.pRSET construct by *Sal* I and *Xho* I digestion (complementary digested end) and remained the V_L domain, $(GGGGS)_3$ linker and HSV-tag detection peptide after ligation. The construct of single V_H domain of MAC61, on the other hand, was obtained by removed the V_L domain together with $(GGGGS)_3$ linker of MAC61.pRSET construct by *BamH* I and *Sal* I digestion and remained the V_H domain with HSV-tag detection peptide after blunted end ligation.

Prior to bacterial expression, the constructs were verified by PCR, restriction enzymes digestion, and DNA sequencing across the destroyed cloning site.

2.9.2 5'-V_H-linker-V_L-tag-3' scFv.

The scFv genes used for this study were normally in the form of 5'-V_L-linker-V_H-tag-3'. However, the MAC61.scFv was also constructed in the form

of 5'-V_H-linker-V_L-tag-3'. In order to introduced a new restriction sites for facilitated the cloning, PCR amplification was carried out to extract the V_H and V_L domain fragments from pTM-22 plasmid which contained MAC61.scFv. The Oligonucleotides 5'-AAATTTTCCATGGGGCAGCTGAAGGAGTCAGGACCTGG-3' and 5'-ATAT ATCCCGGGTGAGGAGACTGTGACCATGACTCC-3' were used to PCR amplify and introduce *Nco* I restriction enzyme site at the 5'-ends and *Sma* I site at the 3'-end of the V_H domain fragment while oligonucleotides 5'-AAATTTGTCGACATTTTGCT GACCCAATCTCCTACT-3' and 5'-ATATAACTCGAGTCTGATTTCCAGCTTGG TGCCTCC-3' were used to PCR amplify and introduce *Sal* I site at the 5'-end and *Xho* I site at the 3'-end of the V_L domain fragment of MAC61. After PCR amplification, the 5'-V_H-linker-V_L-tag-3' was constructed into the pRSET expressing vector using the similar procedures for the construction of 5'-V_L-linker-V_H-tag-3' scFv gene.

Prior to bacterial expression, the constructs were verified by PCR and restriction enzymes digestion.

2.9.3 Hybrid scFv.

The MAC61.scFv and PV4.scFv in pRSET were used to construct the hybrid scFv genes. Since both the MAC61.scFv and PV4.scFv in pRSET have similar restriction cloning sites, the hybrids can be obtained by switching the V_L domain genes through molecular cloning. The V_L domain gene fragment of MAC61 was obtained by digesting the MAC61.scFv with *Nco* I and *PspA* I before cloning into PV4.scFv predigested with *Nco* I and *PspA* I to form hybrid scFv (VL_{61}-VH_{P4}.scFv) comprising of the V_L domain of MAC61 and the V_H domain of PV4. Following the similar method, the V_L domain gene fragment of PV4 was cloned into the predigested MAC61.scFv to form hybrid scFv (VL_{P4}-VH_{61}.scFv) containing the V_L domain of PV4 and the V_H domain of MAC61.

Prior to bacterial expression, the constructs were verified by restriction enzymes and PCR analysis.

2.9.4 Maltose-binding Protein and scFv Fusions.

Two vectors containing the maltose-binding protein (MBP) gene coding region, pMal-c2 and pMal-p2, were purchased from the New England Biolab. pMal-p2 is basically similar to pMal-c2 except that pMal-p2 contains a *malE* signal sequence.

The scFv genes of ABA26, MAC61, and MAC252 from pGEM11Zf(-) were fused to the C-terminus of the maltose binding protein. The construct was made by digesting pGEM11Zf(-) plasmid containing scFv with *Eco*R I, followed by filling in of the digested ends by the Klenow fragment reaction. The plasmid was then purified by DNA matrix before it was digested by *Xba* I to produce scFv fragments with *Xba* I digested 5'-end and blunted 3'-end. Similarly, the pMal vector was digested by *Hind* III, treated with Klenow fragment reaction, subjected to DNA matrix purification and digested by a second restriction enzyme, *Xba* I. The digested scFv fragments were then ligated into predigested pMal vector. After selection by antibiotic, the colonies were screened by PCR and the constructs verified by restriction enzymes analysis.

Bacterial expression was then carried out on selected positive clones and the proteins produced were analysed.

2.9.5 Diabodies.

The diabodies construct was produced using scFv of PV5 (similar antigen as PV4 but with different epitopes). Two primers were synthesised, one of them was the downstream heavy chain complementary sequences, 5'-AAATTTCCCGGGGGTGGC GGTTCTGTCGACCCGGTGCAG-3', and the other was the scFv short linker followed by upstream heavy chain sequences, 5'-ATAATACTCGAGTGAGGAGA CTGTGAGAGTGGTGCC-3'. Enzyme restriction sites were also included in the primer to facilitate cloning. PCR was carried out using the above primers to synthesise heavy chain with short linker at 5'-end. The PCR fragment was then purified by agarose gel and digested by *Sma* I and *Xho* I before it was cloned into PV5 scFv in pABE where the long linker and heavy chain had been removed by restriction enzymes

digestion. After selection, the construct was verified by PCR and restriction enzymes digestion.

Bacterial expression was then carried out on selected positive clones and the proteins produced were analysed.

2.9.6 Green Fluorescent Protein (GFP) and scFv Fusions.

The green fluorescent protein gene, m-gfp5, was obtained from Dr. Jim Haselhof (MRC, Cambridge). The gene fragment was extracted from the original plasmid by PCR to introduce restriction enzyme sites which facilitate cloning. Two primers were synthesised, namely 5'-ATAGGATCCAATGAGTAAAGGAGAAGA AC-3' and 5'-TAACTCGAGTTGTATAGTTCATCCATGCC-3'. The GFP gene was initially cloned into the pRSET expression vector to verify the sequences by bacterial expression of the fluorescent protein. After the gene was proven to be functional, it was fused separately in frame with amino terminal and carboxyl terminal of MAC61.scFv and PV4.scFv in pRSET expression vector. The constructs were then verified by PCR and restriction enzymes digestion before the bacterial expression of GFP antibody fusions.

CHAPTER 3

Construction of scFv Genes and Expression in Bacteria.

3.1 Introduction.

The expression of heterologous genes in *E. coli* has frequently been used to produce large quantities of recombinant proteins. The exact form of a foreign protein produced in *E. coli*, however, depends on many factors, including both the nature of the expressed proteins and the type of expression system. Proteins expressed in *E. coli* may remain soluble in the cytoplasm, become insoluble inclusion bodies, be transported across the cytoplasmic membrane and accumulate in the periplasmic space, or be secreted into the bacterial culture medium.

In antibody engineering, numerous attempts have been made to produce antibody molecules or antibody fragments in bacteria. All of the techniques that have been used for the expression of heterologous proteins have also been applied to the expression of antibodies or antibody fragments in *E. coli*. However, to date, it has not proved possible to produce functional, inact antibodies in bacterial systems. Furthermore, the failure of bacteria to glycosylate proteins makes them unsuitable hosts whenever glycosylation is essential for the proper functioning of antibodies (Sandhu, 1992). Nevertheless, when whole and glycosylated antibodies are not required, for example, in the expression of antibody fragments for studies of antigen-binding or antibody structure, the *E. coli* expression system is as yet unmatched by any other prokaryotic or eukaryotic system.

At present, there are several ways of expressing antibody fragments in *E. coli*, all of which share some common advantages. These include a well established and convenient gene technology, which allows simple gene manipulation and direct insertion into convenient ready-made expression vectors. The rapid growth of *E. coli* and its comparatively simple fermentation enable relatively straightforward large-scale production of antibody fragments (Plückthun, 1991a). Furthermore, the handling of very large numbers of clones currently necessary in all work with antibody

combinatorial libraries, is only feasible in the *E. coli* systems (Plückthun, 1991b).

The expression of functional antibody fragments in *E. coli* is not a straightforward process. The yields reported for different antibody fragments are rather different and critically dependent on the design of translation initiation region, the proteolytic stability of the antibody protein, the expression system used, and the bacterial host strain (Plückthun, 1991a). In addition, the correct folding of antibody protein is essential for producing functional antibody fragments in *E. coli*. Among the various types of bacterial expression system, those involving the secretion of the expression products appear to offer the greatest promise for the production of functional scFv proteins (Glockshuber *et al.*, 1990; Plückthun, 1992). Antibody fragments produced previously only as inclusion bodies in the *E. coli* cytoplasm have been obtained in functional form following secretion into the bacterial periplasm (Glockshuber *et al.*, 1990).

It is now known that the formation of disulfide bonds serves an important function in determining the correct folding essential for the production of a functional antibody fragment (Hunkapiller and Hood, 1989). Mutagenesis experiments with certain antibody variable domains have indicated that in general it may not be possible to obtain the folded structure in the absence of the disulfide bonds, as the total free energy is then insufficient (Glockshuber *et al.*, 1992). It is also clear that oxidising conditions favour the formation of disulfide bonds (Glockshuber *et al.*, 1992). Thus the relative ease with which functional antibody fragments can be produced in the bacterial periplasm is believed to relate to the presence of the disulfide bond forming machinery of the bacterium and generally oxidising conditions in that compartment.

In contrast, the bacterial cytoplasm is normally in a reducing state, lacks the enzyme protein disulfide isomerase and is therefore unfavourable for the formation of disulfide bonds (Bardwell *et al.*, 1991; Kamitani, 1992). Recently, the use of a thioredoxin reductase (*trxB*$^{-}$) mutant *E. coli* strain in which the cytoplasm is somewhat more oxidising has permitted the formation of disulfide bonds and consequently, the production of functional antibody fragments (Proba *et al.*, 1995).

Secretion of antibody fragments to bacterial periplasm can be achieved by a fusion of the antibody gene to *E. coli* signal sequences such as the outer membrane

protein A signal sequence (*ompA*). There are several advantages to this secretory expression system. First, it directly leads to an assembled functional product with correctly formed disulfide bonds without the need to refold the protein *in vitro*. Secondly, the problem of protease degradation is greatly diminished, as there are fewer proteases in the periplasm than in the cytoplasm. Protection against proteolysis is also achieved by the folding to globular domains which accompanies the oxidation of the disulfide bonds in the periplasm (Plückthun, 1991b).

If the accumulation of antibody fragments as insoluble materials in either the periplasm or cytoplasm is to be avoided, expression rate must be appropriate to the rates of secretion, folding and assembly (Plückthun, 1991a). Therefore, factors such as the promoters used and the expression conditions must also be taken into consideration in expressing functional antibody fragments. Promoters from a range of different sources have been used to drive the expression of antibody genes. Those most commonly used have been *lac* (Skerra and Plückthun, 1988), T7 (Condra *et al.*, 1990), *Para*B (Better *et al.*, 1988), *trp* (Huston *et al.*, 1988), and *tac* (Bregegere *et al.*, 1994). All of these are inducible promoters and the T7 promoter is the strongest among them.

Detection (and sometimes also in purification) of bacterially produced antibody fragments is often facilitated by addition of a fusion "tag" sequence. The tag sequences can be fused either at the carboxyl terminal or amino terminal of an antibody fragment. The commonly used carboxyl terminal "tag" sequences include immunodetectable peptides such as C-myc-tag (Glockshuber *et al.*, 1990), and HSV-tag, and *Strep* tag (Schmidt and Skerra, 1993). In contrast, the short FLAG tag (DYKD) offers the amino terminal "tag tail" which allows sensitive detection and affinity purification (Knappik and Plückthun, 1994). Besides the short peptide tag sequences, fusion proteins have also been employed as a versatile recombinant antibody tag. These include calmodulin which binds to inorganic ligands with high affinity (Neri *et al.*, 1995), *Staphylococcal* protein A (Gandecha *et al.*, 1995), and maltose-binding protein (Bregegere *et al.*, 1994).

Although there has been intensive research into the bacterial expression of recombinant antibodies for the past few years, there is no single expression system or a

particular type of antibody fragment which provides a generic solution for the expression of antibody fragments. In this chapter, the construction and bacterial expression of scFv genes from four different hybridoma cells lines is reported. The bacterially expressed scFv proteins were also characterised.

3.2 Results and Discussion.

The studies described here on the bacterial expression of antibody fragments began with the preparation of hybridoma cells as the source of antibody genes.

3.2.1 Preparation of Hybridomas.

A total of 74 clones of mouse hybridomas from a single fusion between SP2/0 myeloma cells and spleen cells of a mouse immunised with ABA-C_1-EggAlb conjugate exhibited strong positive responses on ELISA screening using ABA-C_1-BSA conjugate as immobilised antigen. Fig. 3.1 shows several developmental stages of the hybridoma cells after cell fusion. Two hybridoma clones, namely ABA15 and ABA26 with different growth rates and apparent high affinities on ELISA tests were subjected for further monoclonal selection through limiting dilution. After three rounds of selection by limiting dilution, the two selected hybridoma clones were propagated to produce monoclonal antibodies (MAbs) and hybridoma cells for RNA extraction.

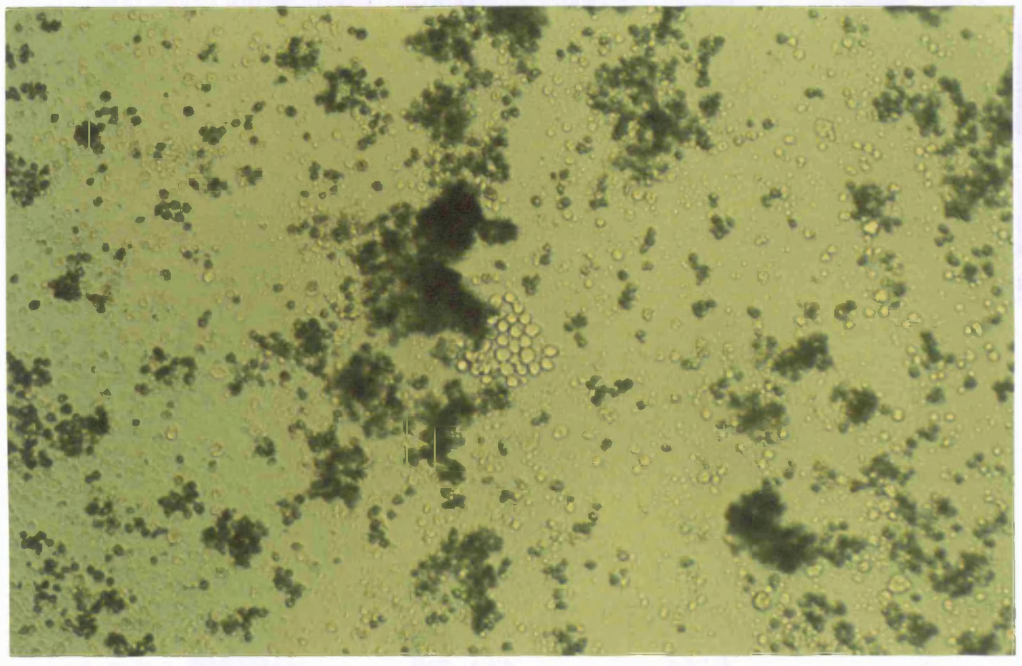
Two rat hybridoma clones, MAC61 and MAC252 which secrete MAbs against ABA-C4'-conjugate were obtained from the European Collection of Animal Cell Cultures (ECACC) and were also propagated for the production of MAbs and hybridoma cells for RNA extraction.

ELISA tests indicated that all four MAbs were able to bind to the types of ABA-conjugates against which they were produced, i.e. ABA15 and ABA26 MAbs bound to ABA-C_1-conjugate while MAC61 and MAC252 MAbs bound to ABA-C4'-conjugate (Fig. 3.2). However, ABA15 and ABA26 MAbs were also found to bind to ABA-C4'-conjugate whereas MAC61 and MAC252 MAbs were unable to bind to ABA-C_1-conjugate. This result indicated that MAC61and MAC252 MAbs were more

Figure 3.1 : Appearance of hybridomas at different stages of development.

The diagrames show the development stage of hybridomas, (a) 1 week after fusion, and (b) formation of hybridomas colony. Note that the surrounding dead myeloma and spleen cells when the hybridoma cells undergo HAT media selection. Cultures were examined using phase contrast optics with an inverted microscope.

(a)



(b)

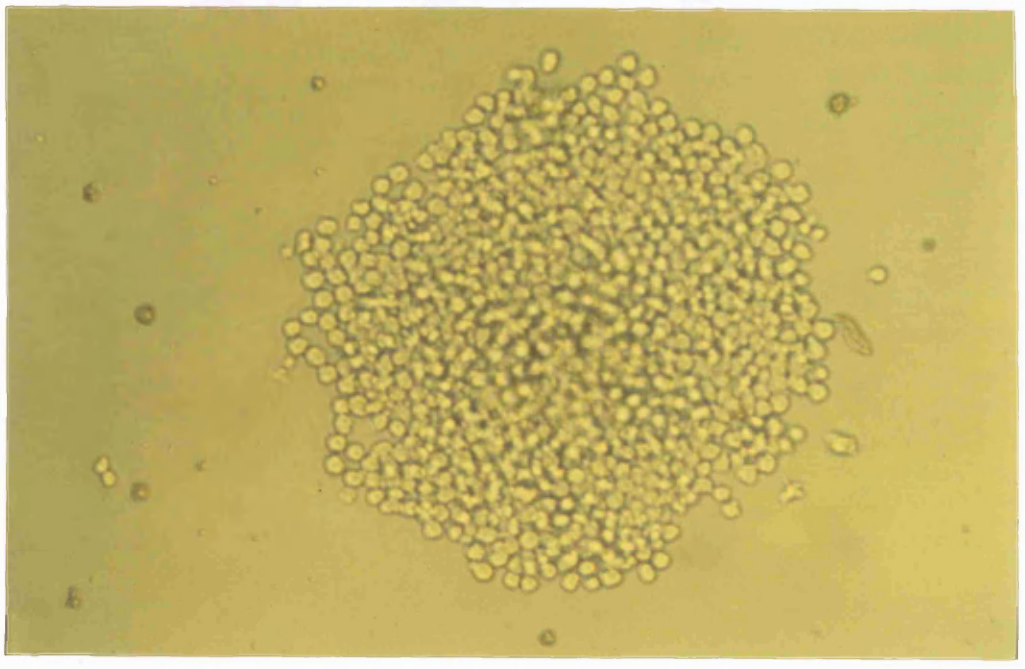
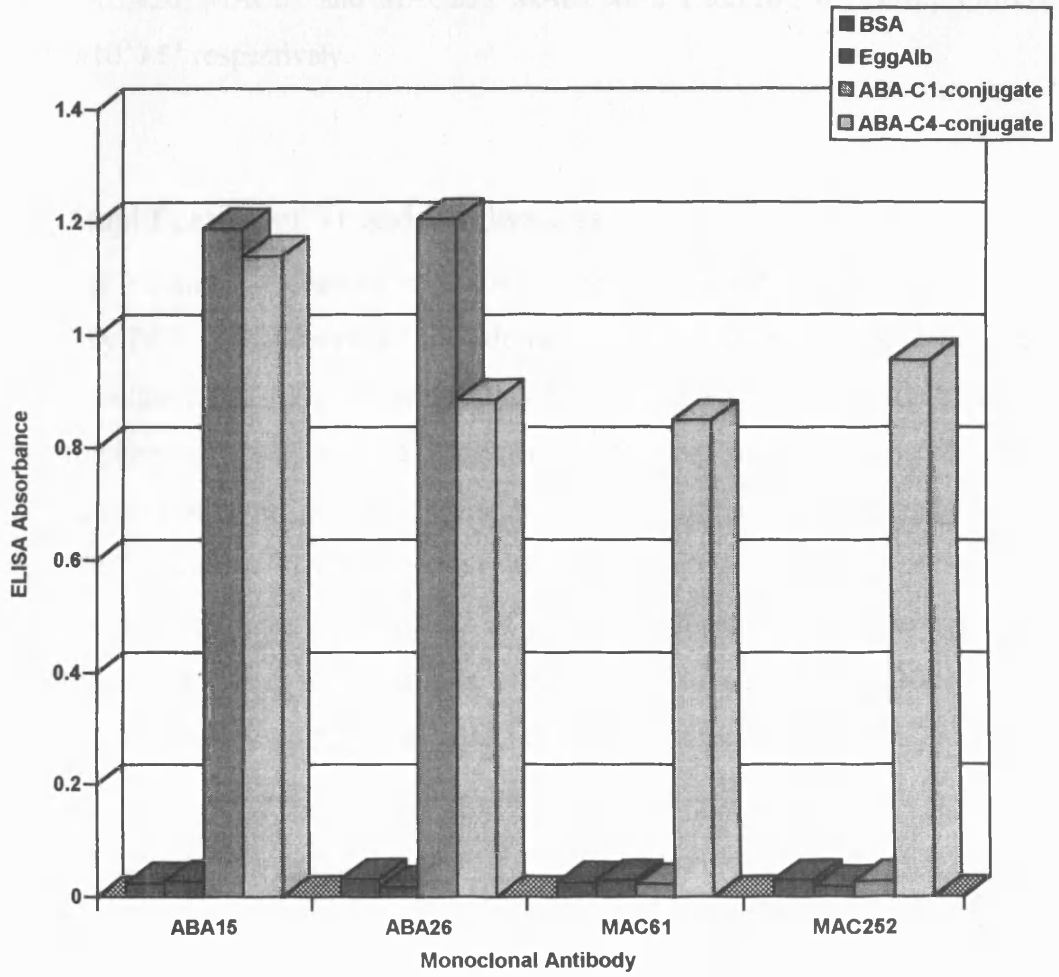


Figure 3.2 : Binding activities of ABA15, ABA26, MAC61, and MAC252 MAbs.

The ELISA assay was carried out using 100 µl of 250µg/ml BAS, EggAlb, ABA-C_1-conjugate, or ABA-$C_{4'}$-conjugate as coated antigen. 100 µl of 500 times dilution of hybridoma culture supernatants of ABA15, ABA26, MAC61, and MAC252 were used as test samples and the horse radish peroxidase-labelled anti-mouse IgG antibody was used as the second antibody.



orientationally specific to ABA-conjugates than ABA15 and ABA26 MAbs.

ELISA tests also demonstrated that the MAbs showed no binding activity to BSA or EggAlb (the carrier proteins used for preparation of the ABA-conjugates), consistent with them being ABA specific (Fig. 3.2). Confirmation of ABA specificity of ABA15, ABA26, MAC61 and MAC252 MAbs was also provided by antigen competition assays in which the binding of all four MAbs to immobilised ABA-conjugates was inhibited by free ABA (Fig. 3.3).

The affinities of the four MAbs for ABA were also measured and found to be moderately high for a small hapten like ABA (Fig. 3.4). The affinity constants (K_D) for ABA15, ABA26, MAC61 and MAC252 MAbs were 1.82×10^6, 6.72×10^5, 1.40×10^6, and 9.60×10^5 M^{-1} respectively.

3.2.2 Amplification of V_L and V_H domains.

The V_L and V_H domains of ABA15, ABA26, MAC61, and MAC252 were obtained by PCR amplification of hybridoma cDNAs using degenerative primers as stated in Section 2.5.3. Fig. 3.5 shows the PCR products after separation by agarose gel electrophoresis. Particular sets of primer combinations resulted in amplification of DNA fragments of about 300 bp, the fragment size equivalent to the variable domain sequence of antibodies. The PCR results also indicated that certain variable domains can be amplified by more than one set of primer combinations. However, only the following sets of primer combinations were chosen to amplify the 300 bp DNA fragments for V_L and V_H domains of ABA15, ABA26, MAC61, and MAC252.

ABA15 :	V_L domain =	L51 / L32 ;	V_H domain =	H51 / H33
ABA26 :	V_L domain =	L51 / L32 ;	V_H domain =	H51 / H33
MAC61 :	V_L domain =	L52 / L32 :	V_H domain =	H56 / H33
MAC252 :	V_L domain =	L51 / L32 ;	V_H domain =	H56 / H33

All of the PCR primers used for amplifying the V_L and V_H domains were complementary sequences targeted at the 5'-end or 3'-end of the variable domain of mouse/rat antibodies except for primer H33 which was targeted at 15 amino acids

Figure 3.3 : Competition immunoassays of ABA15, ABA26, MAC61 and MAC252 MAbs.

The ELISA assay was carried out using 100 µl of 25µg/ml ABA-conjugates as the coating antigen and 100 µl of 1000 times dilution of hybridoma culture supernatants of ABA15, ABA26, MAC61, and MAC252 as the first antibody with horse radish peroxidase labelled anti-mouse-IgG antibody as the second antibody. The antigen competition assay was carried out using 50 µl of 5 µM ABA together with 50µl of 500 times dilution of hybridoma culture supernatants of ABA15, ABA26, MAC61, and MAC252 as the competition mixture.

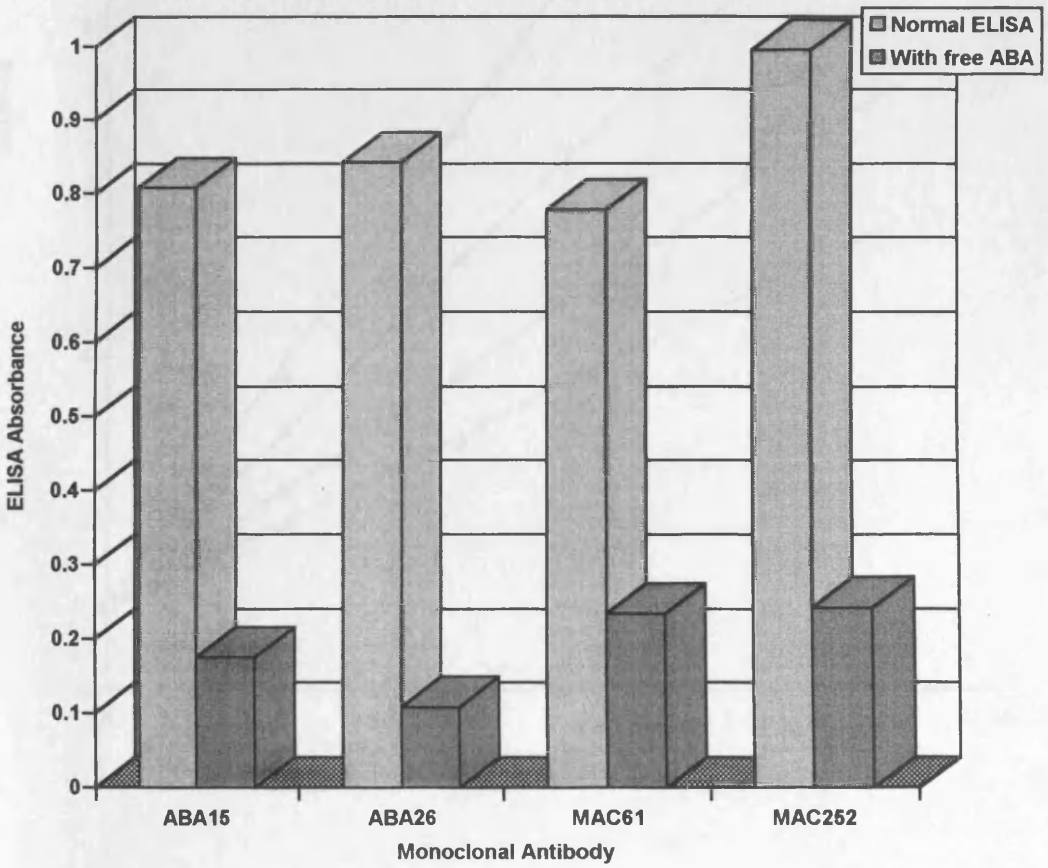


Figure 3.4 : Affinity constants of ABA15, ABA26, MAC61, MAC252 MAbs.

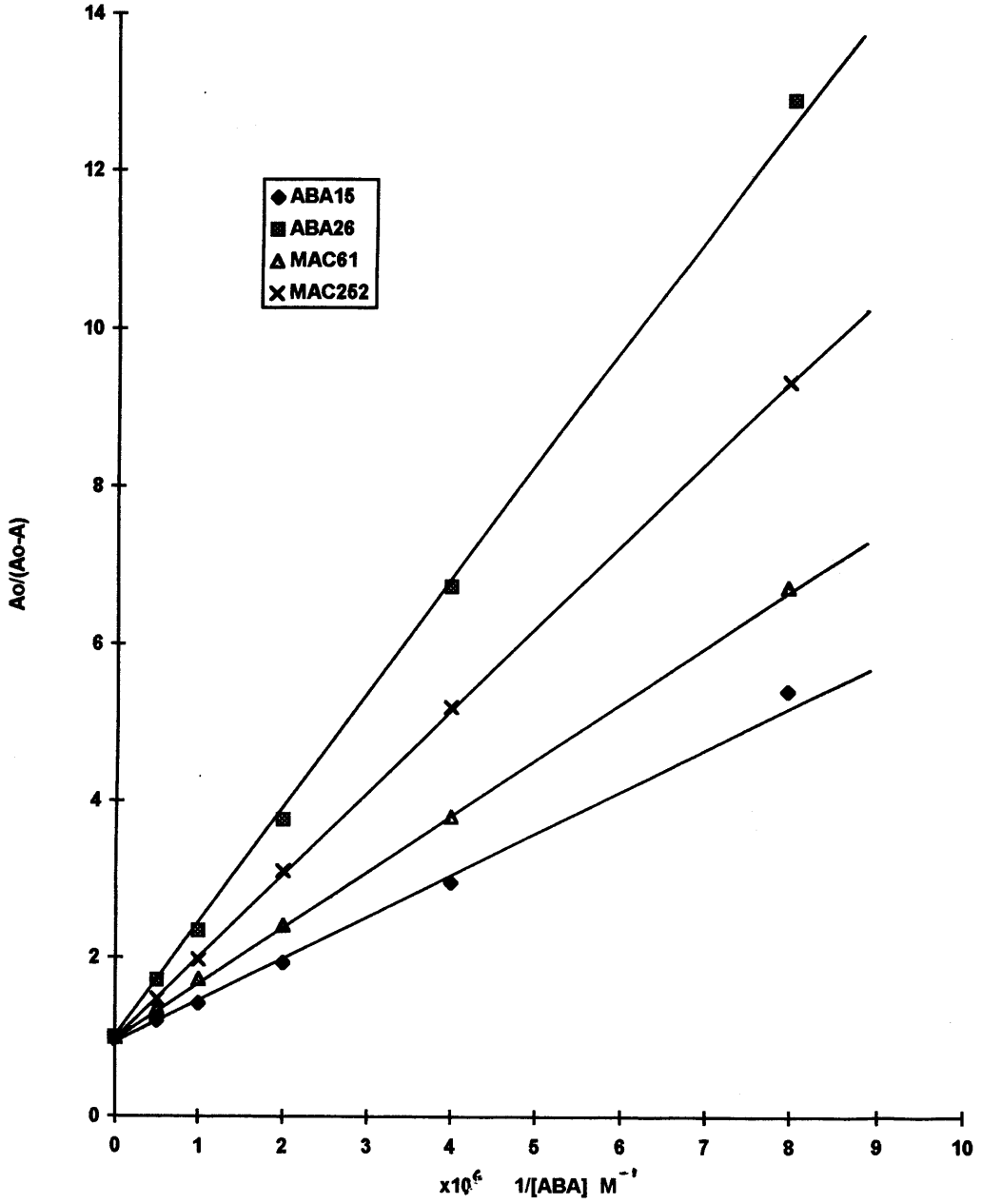
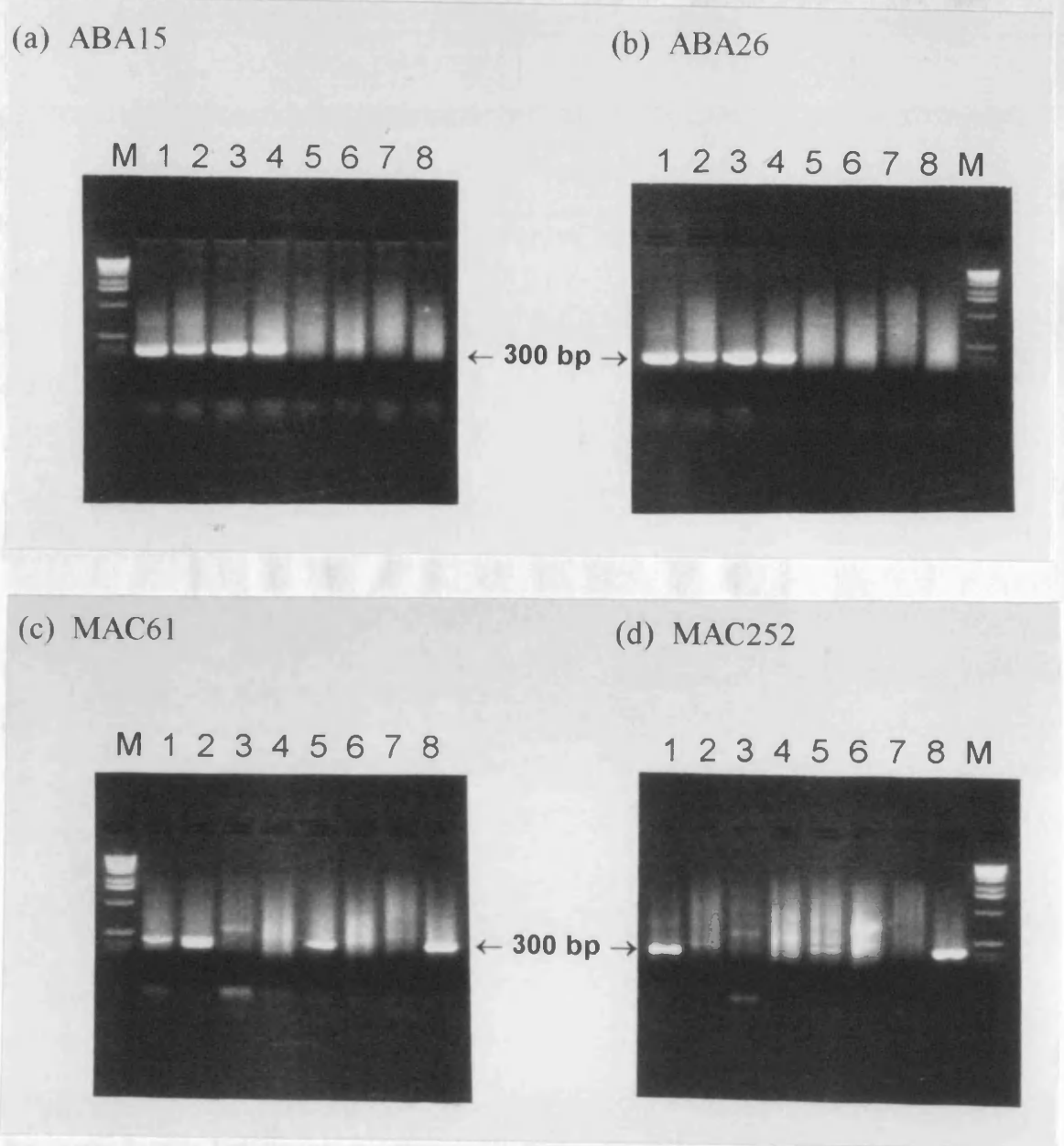


Figure 3.5 : PCR amplification of the V_L and V_H domains from hybridoma cDNAs..

The following pictures depicted the PCR amplification of the V_L and V_H domains of (a) ABA25, (b) ABA26, (c) MAC61, and (d) MAC252 using hybridoma cDNAs as amplification templates. In all the pictures shown, lanes 1 and 2 represent the amplification of the V_L domain using primers L51/L32 and L52/L32 respectively, while lanes 3 to 8 depict the amplification of the V_H domain using primers H51/H33, H52/H33, H53/H33, H54/H33, H55/H33, and H56/H33 respectively. Lane M represents the DNA molecular weight marker.



downstream of the first constant domain of the heavy chain. Primer H33 was targeted at the constant domain rather than the 3'-end of FR 4 of V_H domain because the FR 4 sequence of MAC61 and MAC252 V_H domains were found to be unsuitable for targeting. Unspecific PCR amplification was observed when cDNAs of MAC65 and MAC252 were amplified using primers targeted at the 3'-end of FR 4 of V_H domain (Fig. 3.6).

3.2.3 DNA Sequence of V_L and V_H domains.

The DNA sequences of the cloned V_L and V_H domains of ABA15, ABA26, MAC61, and MAC252 were determined by the dye-terminal automatic sequencing method using cloned plasmid DNA as sequencing template. Table 3.1 lists the corresponding amino acid sequences of the V_L and V_H domains of ABA15, ABA26, MAC61, and MAC252. Comparison of the amino acid sequences with known antibody variable domain sequences and assignment of framework and CDR regions to the sequences confirmed that all of the cloned V_L and V_H domains possessed the characteristics of variable domain sequences. All of the cloned variable domains also contained two cysteine residues (letter bold **C** in Table 3.1), one in the FR 1 and one in FR 3, characteristic of antibody variable domains and responsible for the formation of internal disulfide bonds.

The DNA sequencing results were quite unexpected as the DNA sequences of the V_L and V_H domains of ABA15 and ABA26 turned out to be identical (Table 3.1) despite the MAbs having different affinity constants. A similar result was obtained when DNA sequencing experiments were repeated using cDNAs of ABA15 and ABA26 as sequencing templates. It is possible that these result from differences in the constant domains which were not investigated in this study. There is evidence from antibody gene rearrangement studies that it is possible for two different classes of antibody with identical variable domains to have different constant regions (Steinmetz, 1986). Furthermore, affinity maturation, post-transcriptional and post-translational modification (Berek, 1993) may also cause ABA15 and ABA26 to possess different affinities despite having identical variable domains. Since both ABA15 and ABA26

Figure 3.6 : PCR of the V_H domain using primer targeted at FR 4.

The PCR amplification of the V_H domain was carried out using a primer targeted at 5'-end of the variable domain (H51 for ABA15 and ABA26; H56 for MAC61 and MAC252) while the other primer, H3N2, was targeted at the 3'-end of FR4 as suggested by Orlandi *et al.* (1989). The oligonucleotide sequence of H3N2 is 5'-ACTCTCGAGGGAGA(G/C)TGTGA (G/C)AGTGGTGCC-3'. Lanes 1 to 4 depict the PCR amplification for ABA15, ABA26, MAC61, and MAC252 respectively while lane M represent the DNA molecular weight marker.

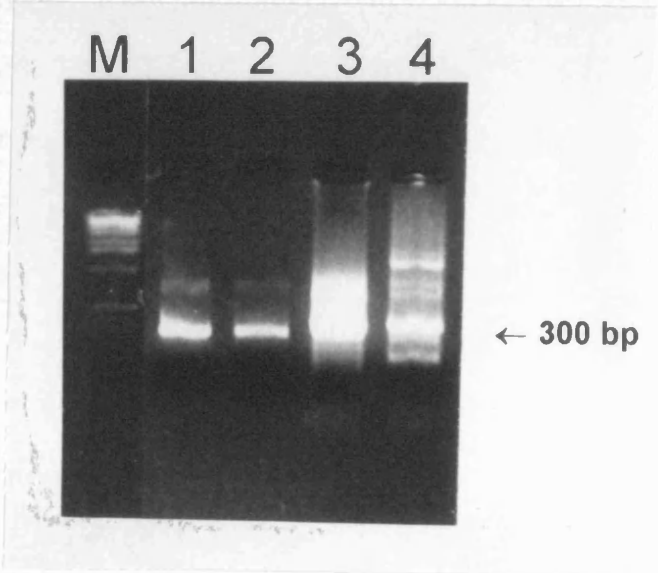


Table 3.1 : The predicted amino acid sequence of V_H and V_L domains of ABA15, ABA26, Mac61, and Mac252.

Amino acid Sequence of V_L domains :

	Framework 1	CDR 1	Framework 2
ABA15	MADVQMIQTPSSLSASLGERVSLT**C**	RASQDIGSSLN	WFQQEPDGTIKRLIY
ABA26	MADVQMIQTPSSLSASLGERVSLT**C**	RASQDIGSSLN	WFQQEPDGTIKRLIY
MAC61	MADILLTQSPT LAVSPGQRATIS**C**	RASQSVTIPSVNLMN	WYQQKPGQQPKLLIY
MAC252	MADVKIIQTTASLSASLGETVSIE**C**	LASEDIYSYLA	WYQQKPGKSPQLLIY

	CDR 2	Framework 3	CDR 3
ABA15	ATSSLDS	GVPKRFSGSRSGSDYSLTISRLESEDFVDYY**C**	LQYASSPWT
ABA26	ATSSLDS	GVPKRFSGSRSGSDYSLTISRLESEDFVDYY**C**	LQYASSPWT
MAC61	HASNLGS	GIPTRFSGSGSGTDFTLTIDPVQADDIATYY**C**	QQSRESPPT
MAC252	ASNRLQD	GVPSRFSGSGSGTQYSLKISGMQPEDEGRYF**C**	LQGSKSPPT

	Framework 4
ABA15	FGGGTKLKLK
ABA26	FGGGTKLKLK
MAC61	FGGGTKLEIR
MAC252	FGGGTKLELR

Amino acid Sequence of V_H domains :

	Framework 1	CDR 1	Framework 2
ABA15	QLVQSGAELVRPGTSVTIS**C**KASGYIFT	NYWLG	WVKQRPGHGLEWIG
ABA26	QLVQSGAELVRPGTSVTIS**C**KASGYIFT	NYWLG	WVKQRPGHGLEWIG
MAC61	QLKESGPGLVQPSETLSLT**C**TVSGFSLT	SYNVH	WVRQPPGGGLEWDG
MAC252	QLKESGPGLVQPSQTLSLT**C**TVSGFSLT	SYNVH	WVRQPTGKGLEWMG

	CDR 2	Framework 3
ABA15	DIYPGVGYTKYNEKFKG	KATLTADTSSSTAHMQLSSLTSEDSAVYF**C**AS
ABA26	DIYPGVGYTKYNEKFKG	KATLTADTSSSTAHMQLSSLTSEDSAVYF**C**AS
MAC61	SYVAWWNYKIIIQVSKS	RLSISRDASKNQLFLKMNSLQSEDTTTYY**C**AR
MAC252	IIWAGGNTAYNSALKS	RLSITRDTSKSQIFLKMNRPQTEDIGTYY**C**VR

	CDR 3	Framework 4	Initial Constant Region
ABA15	TAGNY	WGQGTTLTVSS	AQTTPPSVYPLAPLE...
ABA26	TAGNY	WGQGTTLTVSS	AQTTPPSVYPLAPLE...
MAC61	GYNLYLDY	WGQGVMVTVSS	ARTTAPSVYPLAPLE...
MAC252	GRSIFHYFDY	WGQGVMVTVSS	AETTAPSVYPLAPLE...

possessed identical V_L and V_H domain sequences, a single scFv gene (ABA26.scFv) was constructed containing the variable region domains of these MAbs.

3.2.4 Single-chain Fv (scFv) Genes Construction.

The scFv genes of ABA26, MAC61, and MAC252 were assembled in the pGEM11Zf(-) cloning vector using PCR cloned V_H and V_L domains. Fig. 3.7 shows a schematic diagram for the construction method.

The scFv genes were assembled as 5'-V_L-linker-V_H-3' with a *Nco* I restriction enzyme site located at the 5'-end. The ATG codon at *Nco* I restriction enzyme site encoded the first amino acid of the constructed scFv genes. The linker used was the commonly-used hydrophilic $(GGGGS)_3$ linker and the carboxyl terminal of the scFv genes were attached to a HSV-tag detection peptide. Table 3.2 lists the complete DNA sequences and the corresponding predicted amino acid sequences of the complete ABA26, MAC61, and MAC252 scFv genes.

These scFv genes were constructed along the general lines of scFv genes constructed by other researchers except for the particular restriction enzyme cloning sites and detection peptide used. Although most of the recently published scFv gene constructions were assembled with the arrangement 5'-V_H-linker-V_L-3' (Glockshuber *et al.*, 1990; Anand *et al.*, 1991, ,Ayala *et al.*, 1992, Cheadle *et al.*, 1992, Malby *et al.*, 1993), other work has shown that the 5'-V_L-linker-V_H-3' scFv gene arrangement can also be used in a similar manner (Anand *et al.*, 1991, Pantoliano *et al.*, 1991; Milenic *et al.*, 1991; Denzin and Voss, 1992). However orientation is sometimes important as illustrated by the work of Luo *et al.* (1995) demonstrating that scFv-174 gene can only be expressed in the 5'-V_L-linker-V_H-3' arrangement and not as 5'-V_H-linker-V_L-3'.

Another distinguishing feature of the scFv genes in this study is that the use of the 3'-end primer (H33) in PCR amplification of V_H domain resulted in the inclusion at the carboxyl terminal of the cloned V_H domain of 15 amino acids of the heavy chain constant region. However, it is considered very unlikely that this additional peptide tail has any effect on the antigen-binding properties of the scFv proteins. It is

Figure 3.7 : Schematic diagram of the scFv construction.

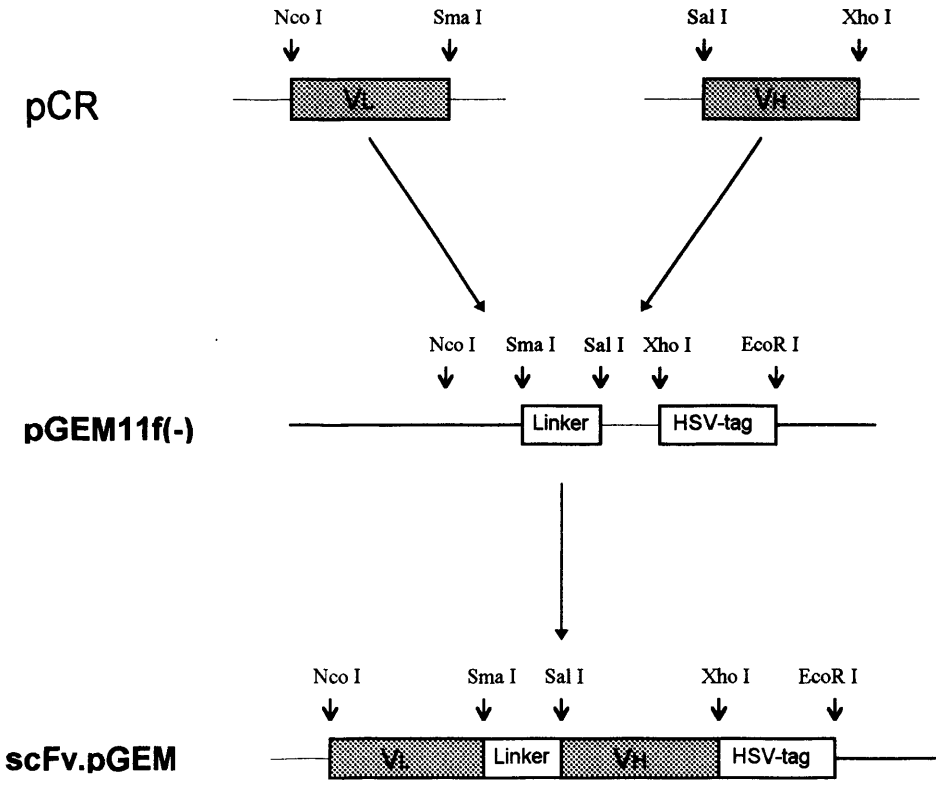


Table 3.2a : DNA and amino acid sequence of ABA26.scFv gene.

```
   Nco I |→
1  ccATGGCAGACGTCCAGATGATCCAGACTCCATCCTCCTTATCTGCCTCTCTGGGAGAAAGA
     M  A  D  V  Q  M  I  Q  T  P  S  S  L  S  A  S  L  G  E  R

61   GTCAGTCTCACTTGTCGGGCAAGTCAGGACATTGGTAGTAGCTTAAACTGGTTTCAGCAG
     V  S  L  T  C  R  A  S  Q  D  I  G  S  S  L  N  W  F  Q  Q

                                VL domain
121  GAACCAGATGGAACTATTAAACGCCTTATCTACGCCACATCCAGTTTAGATTCTGGTGTC
     E  P  D  G  T  I  K  R  L  I  Y  A  T  S  S  L  D  S  G  V

181  CCCAAAAGGTTCAGTGGCAGTAGGTCTGGGTCAGATTATTCTCTCACCATCAGCAGGCTT
     P  K  R  F  S  G  S  R  S  G  S  D  Y  S  L  T  I  S  R  L

241  GAGTCTGAAGATTTTGTAGACTATTACTGTCTACAATATGCTAGTTCTCCGTGGACGTTC
     E  S  E  D  F  V  D  Y  Y  C  L  Q  Y  A  S  S  P  W  T  F

                                ←|Sma I|→                  Linker
301  GGTGGAGGCACCAAGCTGAAACTCAAACCCGGGGGCGGTGGCGGTTCTGGTGGCGGTGGC
     G  G  G  T  K  L  K  L  K  P  G  G  G  G  G  S  G  G  G  G

                     ←|Sal I|→
361  TCTGGCGGTGGCGGTTCTGTCGACCAGCTGGTGCAGTCTGGAGCTGAGCTGGTAAGGCCT
     S  G  G  G  G  S  V  D  Q  L  V  Q  S  G  A  E  L  V  R  P

421  GGGACTTCAGTGACGATATCCTGCAAGGCTTCTGGCTACATCTTCACTAACTACTGGCTA
     G  T  S  V  T  I  S  C  K  A  S  G  Y  I  F  T  N  Y  W  L

481  GGTTGGGTAAAGCAGAGGCCTGGACATGGACTTGAGTGGATTGGAGATATTTACCCTGGA
     G  W  V  K  Q  R  P  G  H  G  L  E  W  I  G  D  I  Y  P  G

                                VH domain
541  GTTGGTTATACTAAGTACAATGAGAAGTTCAAGGGCAAGGCCACACTGACTGCAGACACA
     V  G  Y  T  K  Y  N  E  K  F  K  G  K  A  T  L  T  A  D  T

601  TCCTCCAGCACTGCCCACATGCAGCTCAGTAGCCTGACATCTGAGGACTCTGCTGTCTAT
     S  S  S  T  A  H  M  Q  L  S  S  L  T  S  E  D  S  A  V  Y

                                                              ←|
661  TTCTGTGCCTCTACGGCAGGTAACTACTGGGGCCAAGGCACCACTCTCACAGTCTCCTCA
     F  C  A  S  T  A  G  N  Y  W  G  Q  G  T  T  L  T  V  S  S

                                                Xho I
721  GCCCAAACGACACCCCCATCTGTCTATCCCCTGGCCCCTCTCGAGGGTGCCGGCTCCGGA
     A  Q  T  T  P  P  S  V  Y  P  L  A  P  L  E  G  A  G  S  G

     |→       HSV-tag                ←|        EcoR I
781  CAGCCTGAACTCGCTCCAGAGGATCCGGAAGATTAACTGCAGAATTC
     Q  P  E  L  A  P  E  D  P  E  D  *
```

Table 3.2b : DNA and amino acid sequence of MAC61.scFv gene.

```
   Nco I |→
1  ccATGGCAGACATTTTGCTGACCCAATCTCCTACTTTAGCTGTGTCTCCAGGGCAGAGGGCC
     M  A  D  I  L  L  T  Q  S  P  T  L  A  V  S  P  G  Q  R  A


61   ACCATTTCCTGTAGGGCCAGCCAGAGTGTCACCATTCCTAGCGTTAATTTGATGAACTGG
     T  I  S  C  R  A  S  Q  S  V  T  I  P  S  V  N  L  M  N  W

                                VL domain
121  TACCAACAGAAACCAGGACAACAGCCCAAGCTCCTCATCTATCATGCATCCAACCTAGGT
     Y  Q  Q  K  P  G  Q  Q  P  K  L  L  I  Y  H  A  S  N  L  G


181  TCTGGGATTCCTACCAGGTTCAGTGGCAGTGGGTCTGGGACAGACTTCACCCTCACCATC
     S  G  I  P  T  R  F  S  G  S  G  S  G  T  D  F  T  L  T  I


241  GATCCTGTGCAGGCTGATGATATTGCAACCTATTACTGTCAACAGAGTAGGGAGTCTCCT
     D  P  V  Q  A  D  D  I  A  T  Y  Y  C  Q  Q  S  R  E  S  P

                                          ←|Sma I|→
301  CCGACGTTCGGTGGAGGCACCAAGCTGGAAATCAGACCCGGGGGCGGTGGCGGTTCTGGT
     P  T  F  G  G  G  T  K  L  E  I  R  P  G  G  G  G  G  S  G

        Linker                  ←|Sal I|→
361  GGCGGTGGCTCTGGCGGTGGCGGTTCTGTCGACCAGCTGAAGGAGTCAGGACCTGGTCTG
     G  G  G  S  G  G  G  G  S  V  D  Q  L  K  E  S  G  P  G  L


421  GTGCAGCCCTCAGAGACCCTGTCCCTCACCTGCACTGTCTCTGGGTTCTCATTAACCAGC
     V  Q  P  S  E  T  L  S  L  T  C  T  V  S  G  F  S  L  T  S


481  TATAACGTGCACTGGGTTCGACAGCCTCCAGGAGGAGGTCTGGAGTGGGATGGGAGTTAT
     Y  N  V  H  W  V  R  Q  P  P  G  G  G  L  E  W  D  G  S  Y

                                VH domain
541  GTTGCGTGGTGGAACTACAAGATTATAATTCAGGTCTCCAAATCCCGACTGAGCATCAGC
     V  A  W  W  N  Y  K  I  I  I  Q  V  S  K  S  R  L  S  I  S


601  AGGGACGCCTCCAAGAACCAACTTTTCCTAAAAATGAACAGTCTGCAAAGTGAAGACACA
     R  D  A  S  K  N  Q  L  F  L  K  M  N  S  L  Q  S  E  D  T


661  ACCACTTACTACTGTGCCAGAGGCTATAACTTATACCTTGATTACTGGGGCCAAGGAGTC
     T  T  Y  Y  C  A  R  G  Y  N  L  Y  L  D  Y  W  G  Q  G  V

                     ←|                                          Xho I
721  ATGGTCACAGTCTCCTCAGCCAGAACAACAGCCCCATCTGTTTATCCCCTGGCCCCTCTC
     M  V  T  V  S  S  A  R  T  T  A  P  S  V  Y  P  L  A  P  L

                          |→      HSV-tag              ←|
781  GAGGGTGCCGGCTCCGGACAGCCTGAACTCGCTCCAGAGGATCCGGAAGATTAACTGCAG
     E  G  A  G  S  G  Q  P  E  L  A  P  E  D  P  E  D  *

     EcoR I
841  AATTC
```

Table 3.2c : DNA and amino acid sequence of MAC252..scFv gene.

```
   Nco I |→
1  ccATGGCAGACGTCAAGATAATCCAGACTACAGCTTCCCTGTCTGCATCTCTGGGAGAAACT
     M  A  D  V  K  I  I  Q  T  T  A  S  L  S  A  S  L  G  E  T

61   GTCTCCATCGAATGTCTAGCAAGTGAGGACATTTACAGTTATTTAGCATGGTATCAGCAG
     V  S  I  E  C  L  A  S  E  D  I  Y  S  Y  L  A  W  Y  Q  Q

                              VL domain
121  AAGCCAGGGAAATCTCCTCAGCTCCTGATCTATGCTTCAAATAGGTTGCAAGATGGGGTC
     K  P  G  K  S  P  Q  L  L  I  Y  A  S  N  R  L  Q  D  G  V

181  CCATCACGGTTCAGTGGCAGTGGATCTGGCACACAGTATTCTCTCAAGATCAGCGGCATG
     P  S  R  F  S  G  S  G  S  G  T  Q  Y  S  L  K  I  S  G  M

241  CAACCTGAAGATGAAGGACGTTATTTCTGTCTGCAGGGTTCCAAGTCTCCTCCGACGTTC
     Q  P  E  D  E  G  R  Y  F  C  L  Q  G  S  K  S  P  P  T  F

                                  ←|Sma I|→                Linker
301  GGTGGAGGCACCAAGCTGGAACTCAGACCCGGGGGCGGTGGCGGTTCTGGTGGCGGTGGC
     G  G  G  T  K  L  E  L  R  P  G  G  G  G  G  S  G  G  G  G

                       ←|Sal I|→
361  TCTGGCGGTGGCGGTTCTGTCGACCAGCTGAAGGAGTCAGGACCTGGTCTGGTGCAGCCC
     S  G  G  G  G  S  V  D  Q  L  K  E  S  G  P  G  L  V  Q  P

421  TCACAGACCCTGTCCCTCACCTGCACTGTCTCTGGGTTCTCACTAACCAGCTACAATGTG
     S  Q  T  L  S  L  T  C  T  V  S  G  F  S  L  T  S  Y  N  V

481  CACTGGGTTCGACAGCCTACAGGAAAAGGTCTGGAGTGGATGGGAATAATATGGGCTGGT
     H  W  V  R  Q  P  T  G  K  G  L  E  W  M  G  I  I  W  A  G

                              VH domain
541  GGAAACACAGCTTACAATTCAGCTCTCAAATCCCGACTGAGCATCACCAGGGACACCTCC
     G  N  T  A  Y  N  S  A  L  K  S  R  L  S  I  T  R  D  T  S

601  AAGAGCCAAATCTTCTTAAAAATGAACAGGCCACAAACTGAAGACATAGGCACTTACTAC
     K  S  Q  I  F  L  K  M  N  R  P  Q  T  E  D  I  G  T  Y  Y

661  TGTGTCAGAGGACGTAGCATCTTCCACTACTTTGATTACTGGGGCCAAGGAGTCATGGTC
     C  V  R  G  R  S  I  F  H  Y  F  D  Y  W  G  Q  G  V  M  V

               ←|                                          Xho I
721  ACAGTCTCATCAGCTGAAACAACAGCCCCATCTGTCTATCCCCTGGCCCCTCTCGAGGGT
     T  V  S  S  A  E  T  T  A  P  S  V  Y  P  L  A  P  L  E  G

                  |→      HSV-tag                ←|        EcoR I
781  GCCGGCTCCGGACAGCCTGAACTCGCTCCAGAGGATCCGGAAGATTAACTGCAGAATTC
     A  G  S  G  Q  P  E  L  A  P  E  D  P  E  D  *
```

comprised of antibody-derived amino acid sequence and is located at the carboxyl terminal of the scFv protein. It has been established that antibody fragments in the form of Fabs (containing the whole first constant domain) can be expressed equally as well as Fv or scFv fragments (Plückthun, 1991a). Furthermore scFv proteins do not exhibit any differences in binding activity and affinity when fused with many kinds of fusion proteins at the carboxyl terminal (Skerra, 1993; Luo *et al.*, 1996).

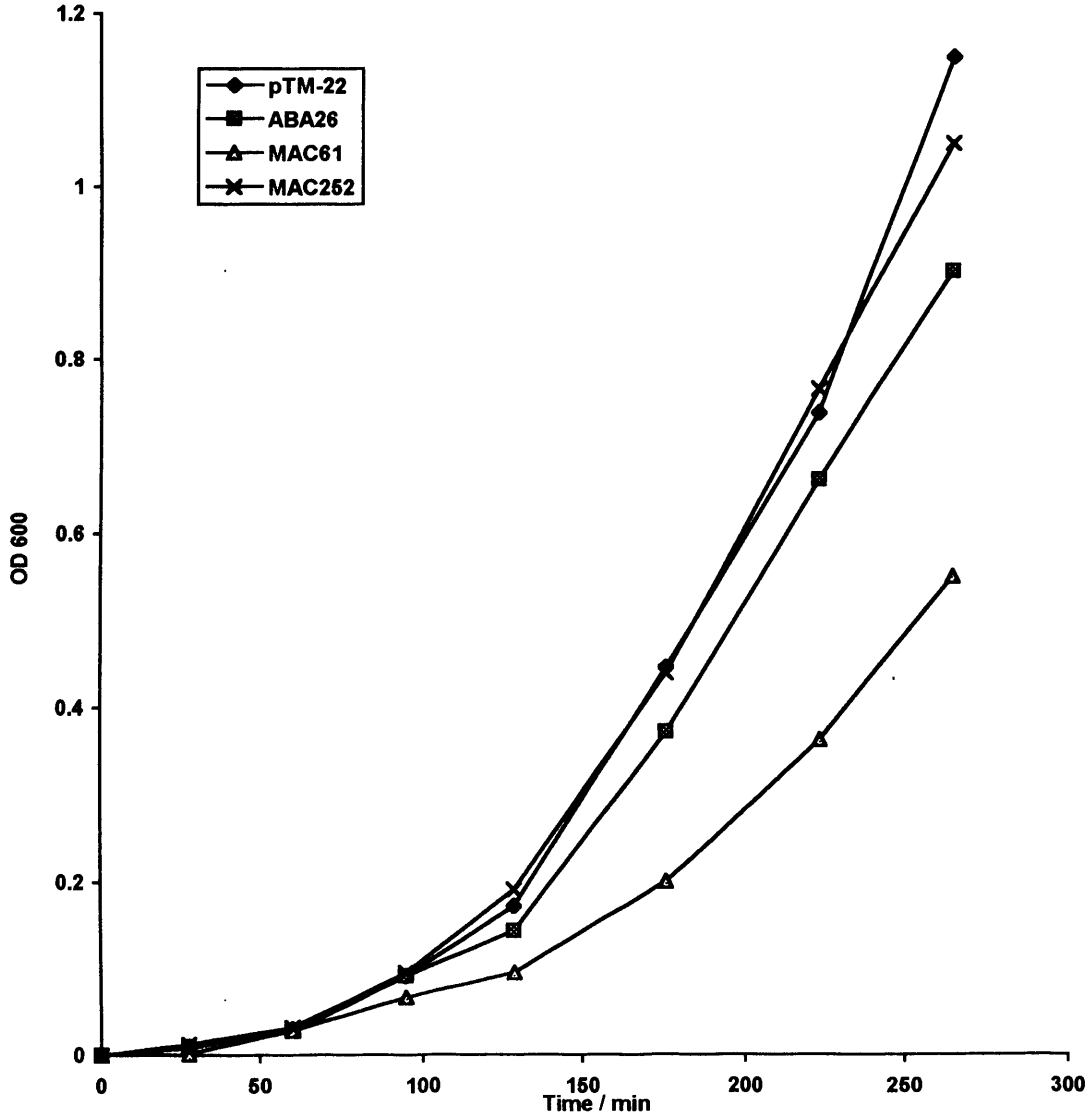
3.2.5 Bacterial Expression of scFv Genes.

Bacterial expression experiments of the scFv genes of ABA26, MAC61, and MAC252 were carried out in four different bacterial expression vectors, namely pTrcHis, pABE, pRSET, and pTM-22. Plasmids were initially constructed in *E. coli* XL-1 Blue and then transferred to other *E. coli* strains appropriate for expression. The *E. coli* W311wt strain was used for expressing scFv genes in pTrcHis and pABE expression vectors while strain JM109(DE3) was used to express scFv genes in pRSET and pTM-22 expression vectors.

Bacterial colonies containing all three scFv genes in the pTrcHis, pABE and pRSET expression vectors grew very well compared to control clones transformed with empty expression vectors. The bacterial colonies transformed with scFv.pTM-22 constructs, however, exhibited a different growth rate from the control clone. The bacterial colonies containing MAC252.pTM-22 had similar growth rate to the control. Colonies containing ABA26.pTM-22 grew more slowly than those of MAC252.pTM-22 whereas colonies containing MAC61.pTM-22 grew the slowest of all. Bacterial colonies containing MAC61.pTM-22 required two days to grow to the size of an overnight colony of cells containing MAC252.pTM-22. A similar result was obtained when growth rates were measured as the increment of OD_{600} at different time intervals (Fig. 3.8). The OD_{600} versus time curve for cells containing MAC252.pTM-22 was similar to the control bacterial culture while cells containing ABA26.pTM-22 had a lower growth rate, and those containing MAC61.pTM-22 displayed the lowest rate of all.

Figure 3.8 : Growth rate of *E. coli* JM109(DE3) containing scFv.pTM-22 genes.

This experiment was carried out using diluted cultures of a saturated culture of *E. coli* JM109(DE3) transformed with ABA26, MAC61 and MAC252 scFv.pTM-22 to OD_{600} at 0.1 using fresh LB media. Cultures were incubated at 37°C and change of OD_{600} was measured.



3.2.5.1 Direct Expression System.

Neither *E. coli* W311wt transformed with scFv.pTrcHis or scFv.pABE, nor *E. coli* JM109(DE3) transformed with scFv.pRSET produced any detectable antibody protein. Only *E. coli* JM109(DE3) transformed with scFv.pTM-22 produced detectable scFv protein as assessed by Western Blots of total protein extract of induced cultures probed with anti-HSV-tag antibody (Fig. 3.9). However, the expression level of the scFv proteins was too low to be detected using colorimetric detection system and only a moderate signal was detected with chemiluminescent detection. Among the three scFv.pTM-22 genes, MAC61.pTM-22 was frequently found to be expressed at a slightly higher level than the others.

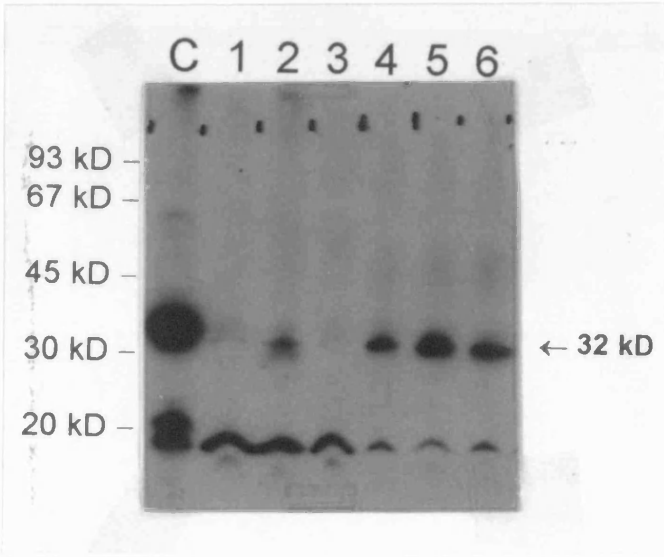
The molecular weight of each expressed scFv protein was about 32 kD as determined by SDS-PAGE. This value was higher than the actual molecular weights calculated on the basis of amino acid sequence. The actual molecular weights of the ABA26, MAC61, and MAC252 scFv proteins were 28.7 kD, 29.6 kD, and 29.4 kD respectively. However, if the *ompA* signal peptide sequences are taken into consideration, than the molecular weight of the expressed scFv proteins would be close to 32 kD.

When soluble and insoluble fractions from positively expressed clones were analysed, the majority of the scFv protein was found to be in the insoluble fraction and only a very small amount (barely detectable) was found in the soluble fraction (Fig. 3.9). This observation conflicts with the common finding that secretion of bacterially expressed scFv proteins to the periplasm results in the production of substantial amounts of scFv proteins in soluble form (Glockshuber *et al.*, 1990; Plückthun, 1990; Skerra, 1993). Previous work has also shown that if expressed scFv proteins are targeted to the cytoplasm they commonly accumulate as insoluble cytoplasmic inclusion bodies (Ayala *et al.*, 1995).

Judging from their molecular weights, it appears that the bacterial expressed scFv proteins in this study possessed signal peptides. This may indicate that they had not been secreted into the periplasm but had remained in the cytoplasm and formed insoluble inclusion bodies. Previous work has shown that, particular sequences within

Figure 3.9 : Expression of scFv.pTM-22 constructed in *E. coli* JM109(DE3).

The protein samples were separated on 12% SDS-acrylamide gel. The Western Blot membrane was probed with the anti-HSV-tag antibody and stained by chemiluminescent detection method. Lanes 1, 2, and 3 represent the soluble fraction while lanes 4, 5, and 6 depict the insoluble fraction of the induced cultures of ABA26, MAC61, and MAC252 scFv.pTM-22 transformants respectively. Lane C represents positive control of a bacterial expressed scFv which contained HSV-tag detection signal peptide.



30 amino acids downstream of the signal peptide can inhibit the translocation of protein into the periplasm (Anderson and Heijne, 1991; Johansson *et al.*, 1993). In fact, Ayala *et al.* (1995) has found that the scFv protein of CB-Hep.1 when fused with the *ompA* signal peptide was not exported to the periplasm but remained in the cytoplasm in an aggregated form with the bacterial insoluble material. In this case the presence of a positively charged arginine close to the signal peptide was shown to be responsible for the association of the scFv protein with the bacterial inner membrane. The scFv genes expressed in this study were found to contain several positively charged amino acids next to the signal peptide sequence. ABA26 and MAC61 scFv proteins contained arginine at residues 20 and 26, whereas MAC252 scFv protein held lysine at residue 5. Mutagenesis of those residues would be required to determine if these positively charged amino acids were responsible for blocking the export of these scFv proteins to the periplasm.

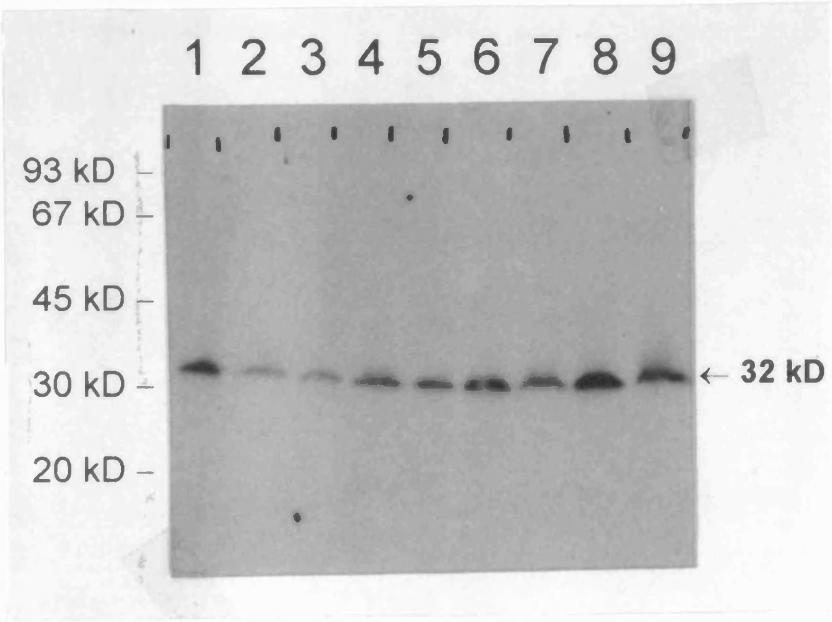
Although the scFv proteins produced following induction of *E. coli* transformed with scFv.pTM-22 were barely detectable in the soluble fraction by Western Blot analysis, ELISA tests using overnight incubation at 4°C of total bacterial protein extracts produced unequivocal positive signals with ABA-conjugate as immobilised antigen. It appears therefore, that a small portion of the bacterially produced scFv proteins were in soluble and functional form. The ability to detect functional scFv proteins in the ELISA tests when soluble scFv protein was barely detectable by Western Blotting analysis may be accounted for by the inherent high sensitivity of ELISA and the fact that 10 times the quantity of the sample was used in the ELISA tests.

3.2.5.2 Indirect Expression System.

Bacterial expression of the scFv genes using pRSET and pTM-22 (T7 promoter) were also carried out with M13/T7 phage infection of *E. coli* XL-1 Blue transformants in order to impose tight regulation of expression. Again the use of all of the scFv.pTM-22 constructs resulted in expression of scFv proteins (Fig. 3.10). However, expression levels were approximately equal to those obtained with *E. coli*

Figure 3.10 : Expression of scFv.pTM-22 constructs in *E. coli* XL-1 Blue through M13/T7 phage infection.

The protein samples were separated on 12% SDS-acrylamide gel. The Western Blot membrane was probed with the anti-HSV-tag antibody and stained by the chemiluminescent detection method. Lanes 1& 2, 3 & 4, and 5 & 6 were crude extract of the induced cultures of scFv.pTM-22 constructs of ABA26, MAC61, and MAC252 in *E. coli* XL-1 Blue. Lanes 7, 8, and 9 show control samples prepared by using crude extract of the induced cultures of scFv.pTM-22 constructs of ABA26, MAC61, and MAC252 in *E. coli* JM109(DE3). Lane N depicts negative control samples of induced *E. coli* transformed with empty pTM-22 vector.



JM109(DE3) and again most of the scFv protein was found in the insoluble fraction with a molecular weight of about 32 kD. Thus there were no substantial differences between the expression of the scFv genes using pTM-22 either by direct expression in transformed *E. coli* JM109(DE3) or indirect expression by M13/T7 phage infection of transformed *E. coli* XL-1 Blue.

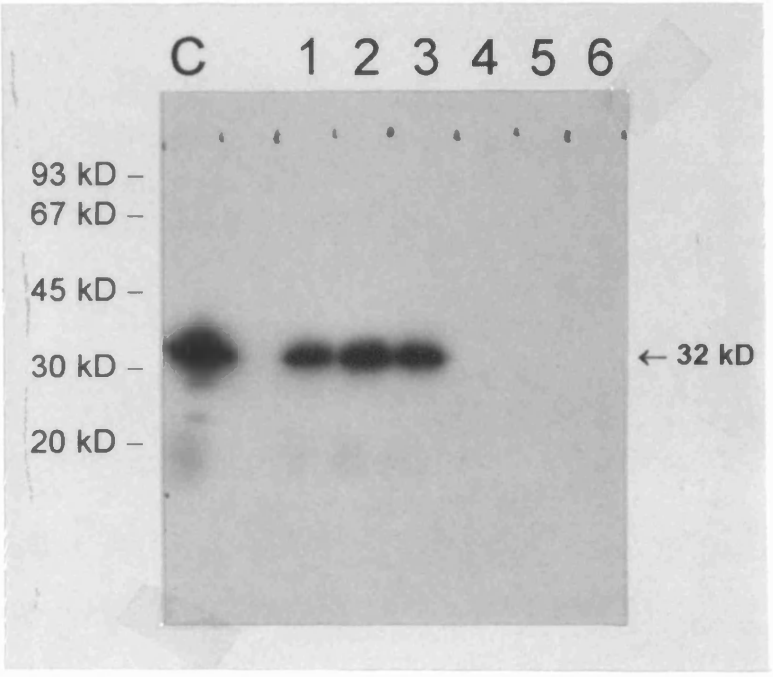
Using the pRSET cytoplasmic expression vector, scFv proteins which were undetectable after attempts to produce them using direct expression in transformed *E. coli* JM109(DE3), were detectable following M13/T7 phage infection of transformed *E. coli* XL-1 Blue (Fig. 3.11). The scFv expression levels obtained were low, similar to those found with the pTM-22 expression vector, and the scFv protein was located in the insoluble fraction. As estimated by Western Blotting analysis, the scFv proteins possessed the predicted molecular weight of about 34 kD (29.5 kD scFv + 4.6 kD of fusion His-tag). However, no positive signals were obtained in ELISA tests for antibody function using the total bacterial protein extract. These findings are consistent with previous work in which targeting of scFv proteins to the bacterial cytoplasm resulted in the formation of insoluble non-functional inclusion bodies (Glockshuber *et al.*, 1990; Plückthun, 1990).

3.2.5.3 Summary of Bacterial Expression.

In summary, it is clear that the ABA26, MAC61, and MAC252 scFv proteins could be expressed at low levels in *E. coli* when expression was driven by the T7 promoter but not when the *tac* or *trc* promoter was used. Among the three scFv.pTM-22 genes, MAC61.scFv was often found to be expressed at a slightly higher level than the others. Although the ABA26, MAC61, and MAC252 scFv genes can be successfully expressed by T7 promoter in both the secretory and cytoplasmic expressions, the secretory expressions by pTM-22 expression vector were at low levels and the majority of the expressed scFv proteins were in the form of insoluble fraction and only a very small amount of the expressed scFv proteins were in soluble functional scFv proteins. While the cytoplasmic expressions by pRSET were at similar levels as expressed by pTM-22 and the expressed scFv proteins were in the form of

Figure 3.11 : Expression of scFv.pRSET in *E. coli* XL-1 Blue through M13/T7 phage infection.

The protein samples were separated on 12% SDS-acrylamide gel. The Western Blot membrane was probed with the anti-HSV-tag antibody and stained by the chemiluminescent detection method. Lane C represents a positive control of bacterially expressed scFv containing HSV-tag detection signal peptide. Lanes 1, 2, and 3 were crude extract of induced cultures of *E. coli* XL-1 Blue while lanes 4, 5, and 6 were crude extract of induced cultured of *E. coli* JM109(DE3) transformed with scFv.pRSET constructs of ABA26, MAC61, and MAC252 respectively.



non-functional insoluble scFv proteins.

The difficulty of expressing the scFv genes with the *tac* or *trc* promoter was unlikely to be caused by errors in the scFv gene constructs or gene manipulation as these factors were well monitored throughout the studies. All of the cloning steps were frequently verified by restriction enzyme analysis of the cloning site or DNA sequencing across the cloning site if the restriction enzyme site was destroyed in the process of cloning. Furthermore, expression experiments always included both negative and positive controls. In addition, since scFv genes constructed with *tac* and *trc* promoters were used without success using the scFv genes from the successfully expressed scFv.pTM-22 constructs, errors in the scFv sequence can be virtually eliminated.

3.2.6 Characterisation of Bacterially-produced scFv Proteins.

ScFv proteins produced by the use of the pTM-22 expression vector were chosen for scFv protein characterisation studies. ELISA tests indicated that scFv proteins of ABA26, MAC61, and MAC252 bind to ABA-conjugate (Fig. 3.12). The specificity of binding to ABA was confirmed by antigen competition assays. As shown in Fig. 3.12, introduction of free abscisic acids into the assay mixture inhibited the binding of the scFv proteins to immobilise ABA-conjugate.

The ELISA results also revealed that the scFv proteins of MAC61 and MAC252 exhibited slightly different binding activities from their parental MAbs. Unlike the parental MAbs which bind only to ABA-C4'-conjugate MAC61.scFv and MAC252.scFv were found also to bind to ABA-C_1-conjugate (Fig. 3.13). This indicates that MAC61 and MAC252.scFv proteins have more flexible binding sites than their parental MAbs, or in other words, they were less orientationally specific than the parental MAbs in terms of their binding activities.

ABA26.scFv protein was also found to bind to ABA-C1-conjugate and ABA-C4'-conjugate (Fig. 3.13). However, this ability was similar to its parental MAb (ABA15 and ABA26) which was also capable of binding to both the

Figure 3.12 : ELISA competition assay of ABA26, MAC61, and MAC252 scFv proteins.

The ELISA wells were coated with 100 µl of 25µg/ml ABA-conjugate. 100 µl of the total protein extract from the induced *E. coli* JM109(DE3) transformed with scFv.pTM-22 constructs were used in normal ELISA assays. The antigen competition assay was carried out using 50 µl of 5 µM ABA together with 50 µl of total protein extract of the induced JM109(DE3) transformed with scFv.pTM-22 constructs. The anti-HSV-tag antibody was used as the second antibody.

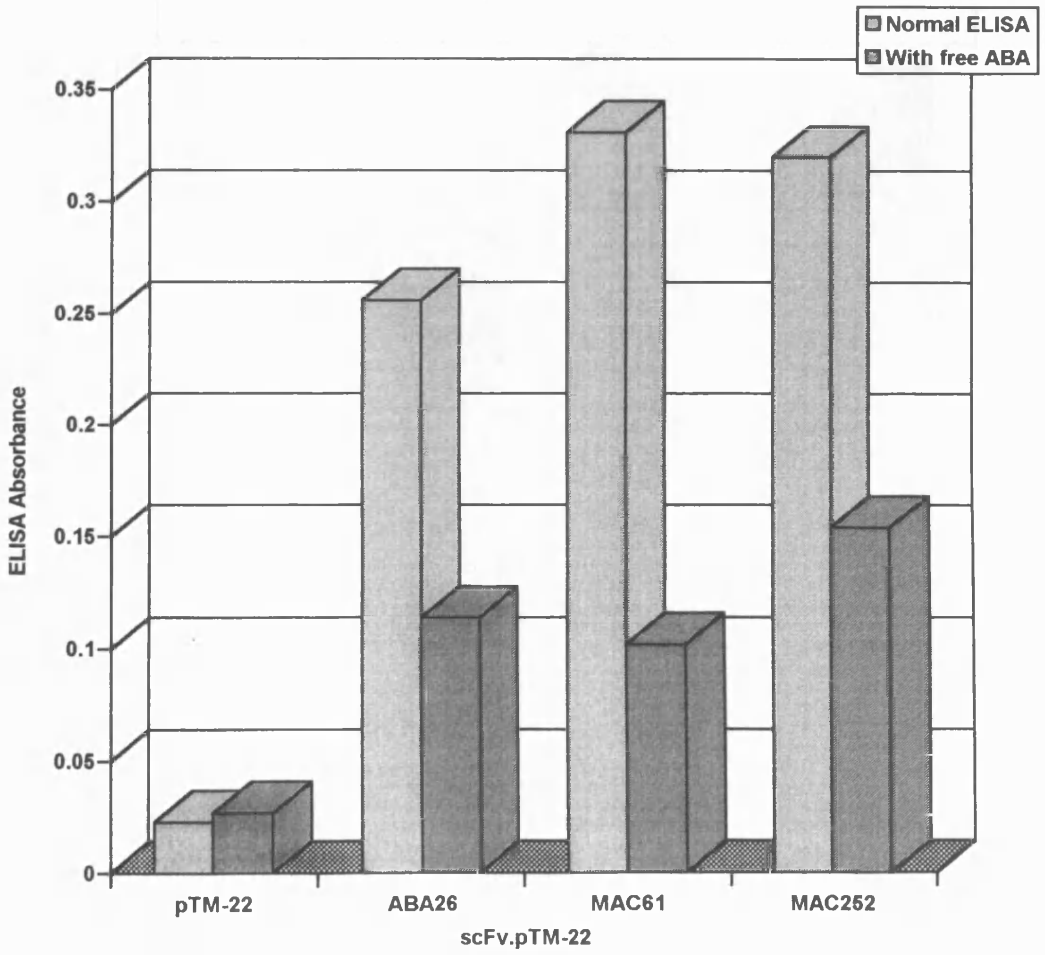
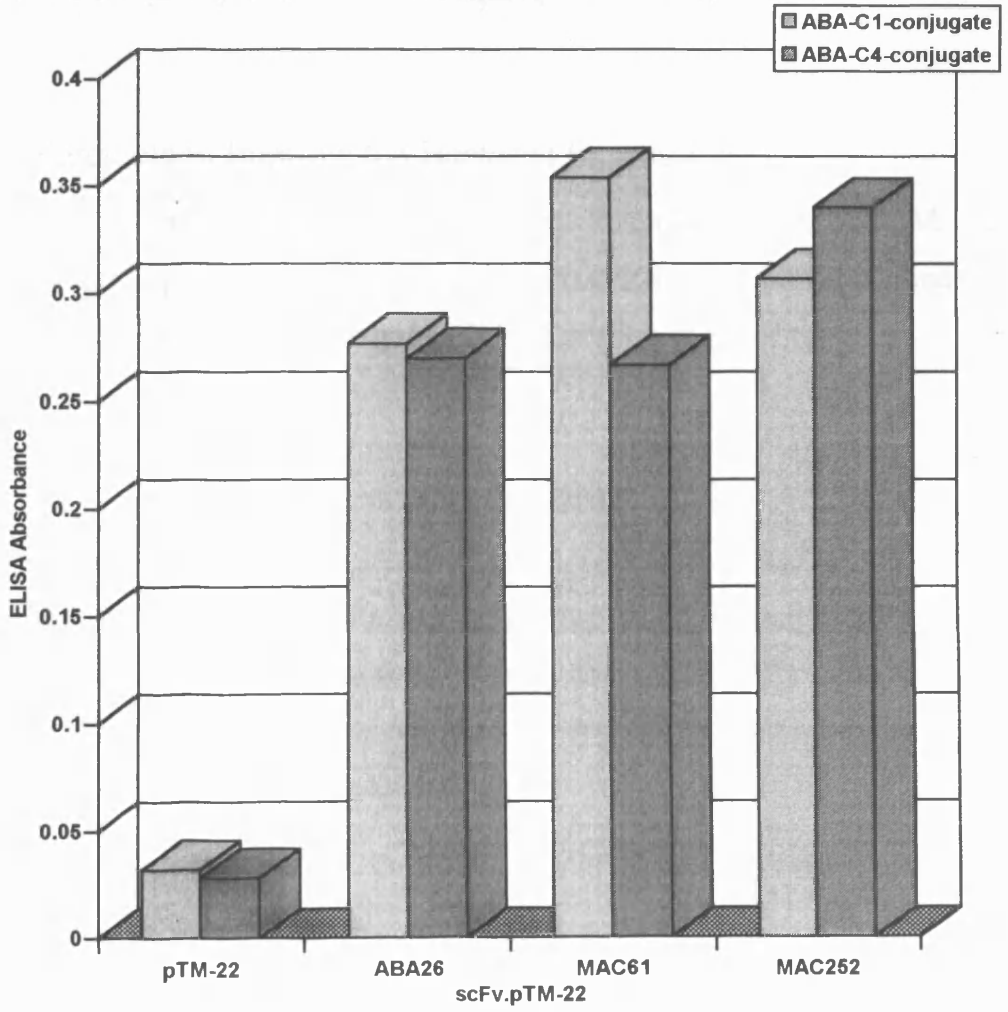


Figure 3.13 : ELISA tests of ABA26, MAC61 and MAC252 scFv proteins using different ABA-conjugates.

In these ELISA tests, 100 µl of 250µg/ml ABA-C_1-conjugate or ABA-$C_{4'}$-conjugate were used as coated antigens. 100 µl of the total protein extract from induced *E. coli* JM109(DE3) transformed with scFv.pTM-22 constructs was used as test samples and anti-HSV-tag antibody was used as the second antibody.



ABA-C1-conjugate and ABA-C4'-conjugate.

The affinity constants of the scFv proteins were also measured and found to be about 5-10 fold lower than those of the parental MAbs. The affinity constants (K_D) of ABA26, MAC61, and MAC252 scFv proteins were measured as $1.84x10^5$, $1.48x10^5$, and $1.42x10^5$ M^{-1} respectively (Fig. 3.14) while the affinity constants of their parental MAbs were $6.72x10^5$, $1.40x10^6$, and $9.60x10^5$ M^{-1} respectively. Although some reports describe scFv proteins with similar affinity constants to their parental MAbs (Skerra and Plückthun, 1988; Glockshuber *et al.*, 1990; Webber *et al.*, 1995), others, like those studied here, have been found to have lower affinity constants than their parental MAbs (Leahy *et al.*, 1988; Desplancq *et al.*, 1994).

3.2.7 Attempts to Improve the Bacterial Expression.

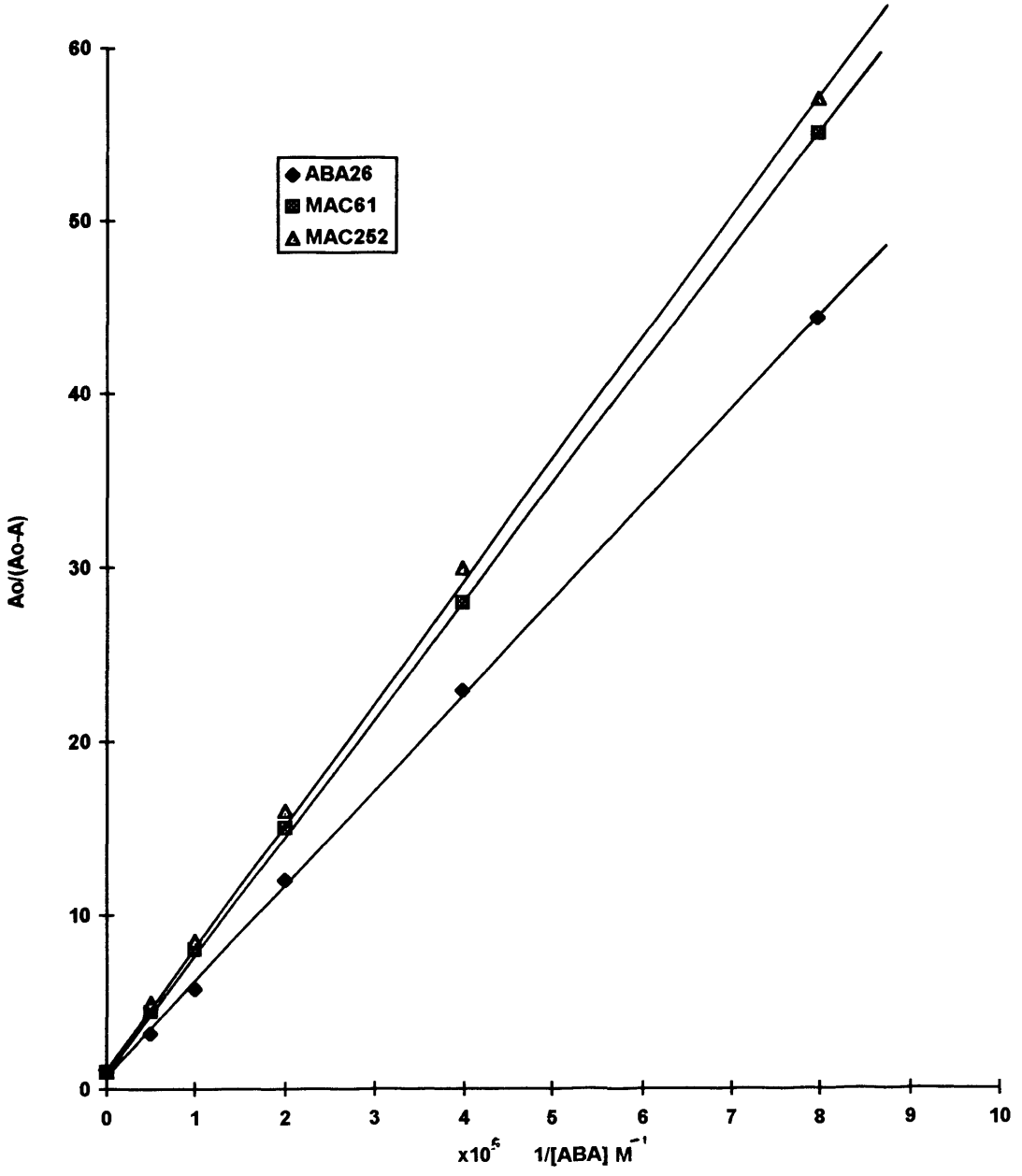
Attempts were made to improve the expression levels of ABA26, MAC61, and MAC252 scFv proteins in *E. coli* by optimisation of expression conditions, by expression in a range of host strains, and by expression as fusion proteins.

3.2.7.1 Modifying the Expression conditions.

The following factors were investigated with respect to effects on scFv protein expression levels: different media recipes, glucose concentration, culture temperature, IPTG concentration, and harvest time. Despite considerable effort only relatively small increases in expression levels were observed. However, the following conditions were found to produce consistent results.

1. LB medium was found to be more suitable than SOB medium for expressing scFv genes using pTM-22.
2. LB medium containing 0.1% glucose gave about two fold higher scFv protein expression levels than medium without glucose. Media containing 0.1%, 0.2% glucose had similar effects on scFv protein production, medium containing 0.3% glucose was found to lower the expression levels whereas a glucose concentration

Figure 3.14 : Affinity constants of ABA26, MAC61, MAC252 scFv proteins.



of 0.5% completely suppressed scFv protein expression.

3. Similar expression levels were observed at incubation temperatures of 37°C and 30°C.
4. Similar scFv protein expression levels were observed following induction with 0.3 or 0.5 mM IPTG at OD_{600} values of 0.3, 0.5 or 0.8.
5. ScFv protein could only be detected after overnight induction. Cultures harvested after 3 hr and 8 hr induction periods did not contain detectable levels of scFv protein.
6. It was found that the use of freshly prepared *E. coli* JM109(DE3) transformants was very important for the expression of scFv.pTM-22 constructs. Especially in the case of MAC61.scFv.pTM-22, transformants tended to lose their ability to express scFv genes after several subcultures.

The somewhat different conditions found to give consistent results in the expression of scFv.pRSET constructs through M13/T7 phage infection of *E. coli* XL-1 Blue transformants, took account of the following observations.

1. Expression levels of scFv proteins were slightly higher in SOB medium than LB medium.
2. Optimal phage infection required growth of the *E. coli* XL-1 Blue transformants in medium containing 25 μg/ml tetracycline.
3. SOB medium containing 0.2% glucose was more effective for scFv gene expression than medium without glucose or containing more than 0.3% glucose.
4. Cultures grown at 37°C and 30°C did not show any significant differences in scFv protein expression levels.
5. Induction of expression by 0.3 mM or 0.5 mM IPTG at culture OD_{600} = 0.3 or 0.5 gave similar expression levels.
6. The amount of M13/T7 phage used was also not critical, a phage titre of around 5pfu/cell ratio was sufficient for infection.

7. ScFv proteins were detected in cultures collected after overnight induction but not in cultures collected after 3 and 5 hr of induction.

It is noteworthy that irrespective of the expression conditions employed no scFv protein was detected in constructs which used the pTrcHis and pABE expression vectors.

3.2.7.2 Expression in Different *E. coli* Strains.

Several *E. coli* strains were used as hosts for expressing scFv genes. These included DH5α, XL-1 Blue, TG1, JM109, W311wt, JM109(DE3), and BL21(DE3) *trxB*$^-$.

Despite the finding by Dueñas *et al.* (1994) that *E. coli* W311wt was particularly suitable for the expression of their scFv genes, none of the scFv genes in the present work which utilised either pTrcHis or pABE expression vectors produced detectable scFv protein in this strain.

Comparison of the expression of scFv.pRSET constructs in *E. coli* BL21(DE3) *trxB*$^-$ with that of the same scFv.pRSET constructs in *E. coli* XL-1 Blue (through M13/T7 phage infection) provided some interesting results. Expression experiments were carried out using SOB medium, induction with 0.3 mM IPTG when culture OD_{600} reached 0.5, and collection of the induced cultures after overnight induction.

All three scFv.pRSET constructs were found to produce at least 5 fold (estimated from Western analysis) more scFv proteins in *E. coli* BL21(DE3) *trxB*$^-$ compared to those expressed using *E. coli* XL-1 Blue (Fig. 3.15). As estimated by Western blotting, the scFv proteins possessed the predicted molecular weight of about 34 kD (29.5 kD scFv + 4.6 kD fusion His-tag). Although the *E. coli* BL21(DE3) *trxB*$^-$ transformants were found to produce a similar amount of total scFv proteins when expressed at either 37°C or 22°C, more soluble scFv proteins were detected at 22°C (Fig. 3.16). From Western analysis, expression at 22°C was found to produce about half of the total scFv proteins in the soluble form while expression at 37°C

Figure 3.15 : Expression of scFv.pRSET in *E. coli* BL21(DE3) *trxB*⁻ strain.

The 12 % SDS-acrylamide gel was used to analyse the total protein extract of induced bacterial cultures. The Western Blot membrane was probed using the anti-HSV-tag antibody and stained by the colour development reaction. Lanes 1, 2, and 3 represent samples prepared by using the crude extract from expressing cultures of *E. coli* XL-1 Blue transformed with scFv.pRSET constructs of ABA26, MAC61, and MAC252 respectively. Lanes 4, 5, and 6 were crude extract of the induced cultures of scFv.pRSET constructs of ABA26, MAC61, and MAC252 in *E. coli* BL21(DE3) *trxB*⁻ strain. While lane N depicts the negative control sample using *E. coli* XL-1 Blue containing empty pRSET vector.

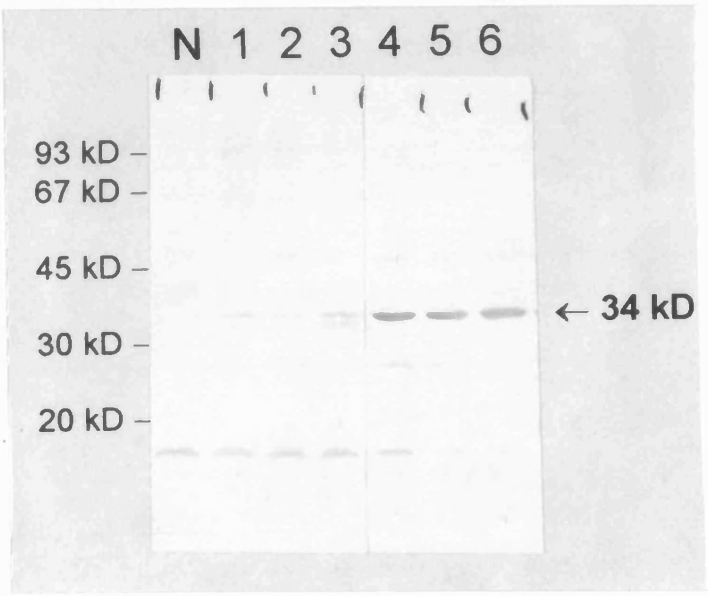
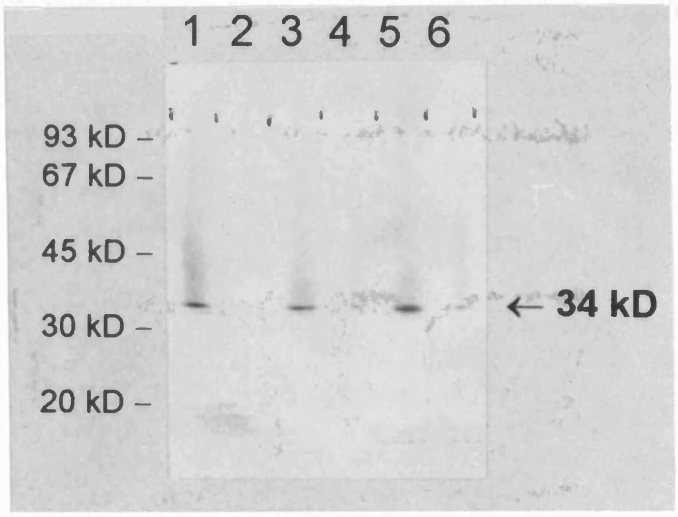
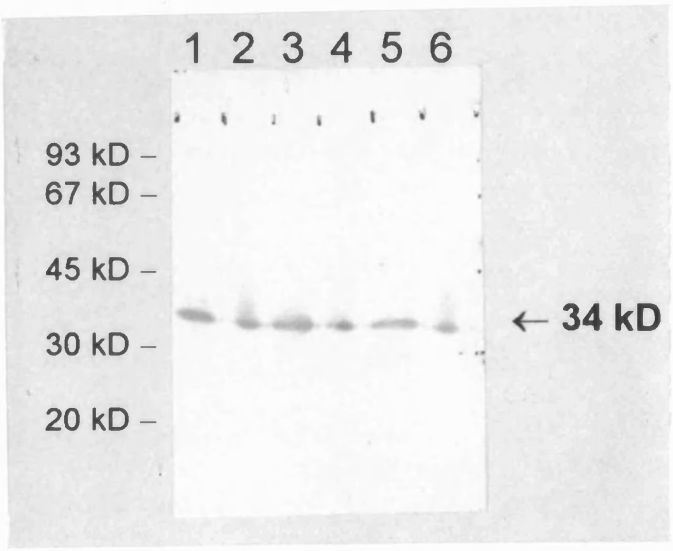


Figure 3.16 : Expression of scFv.pRSET in *E. coli* BL21(DE3) *trxB*$^-$ at 22 °C and 37°C.

The 12 % SDS-acrylamide gel was used to analyse the total protein extract of induced bacterial cultures. The Western Blot membrane was probed using the anti-HSV-tag antibody and stained by the colour development reaction. Panel (a) shows culture samples expressed at 37°C while panel (b) shows culture samples expressed at 22°C. In both panels, lanes 1, 3, and 5 depict the insoluble fraction while lanes 2, 4, and 6 represent the soluble fraction of the bacterial total protein extract of ABA26, MAC61, and MAC252 scFv.pRSET transformants respectively.





produced about one tenth.

When the total protein extracts from induced *E. coli* BL21(DE3) *trxB*⁻ transformants were tested by ELISA using ABA-BSA conjugate as immobilised antigen, all the expressed scFv proteins displayed significant ELISA signals (Fig. 3.17). Based on this result, it was clear that functional scFv proteins can be produced cytoplasmically using *E. coli* BL21(DE3) *trxB*⁻ as the host. In addition, total protein extracts produced at 22°C in *E. coli* BL21(DE3) *trxB*⁻ transformants exhibited stronger ELISA signals than protein produced at 37°C (Fig. 3.17). These differences might be due to the fact that more soluble scFv proteins were produced at 22°C than at 37°C.

The *trxB*⁻ mutant has a more oxidising cytoplasm than wild type *E. coli* and it is possible that this environment may aid the production of functional scFv proteins. This finding agrees with the work of Proba *et al.* (1995) which indicated that the oxidising cytoplasmic environment of this *E. coli* strain assists in the formation of disulfide bonds which in turn enable the correct folding required for the production of functional scFv proteins.

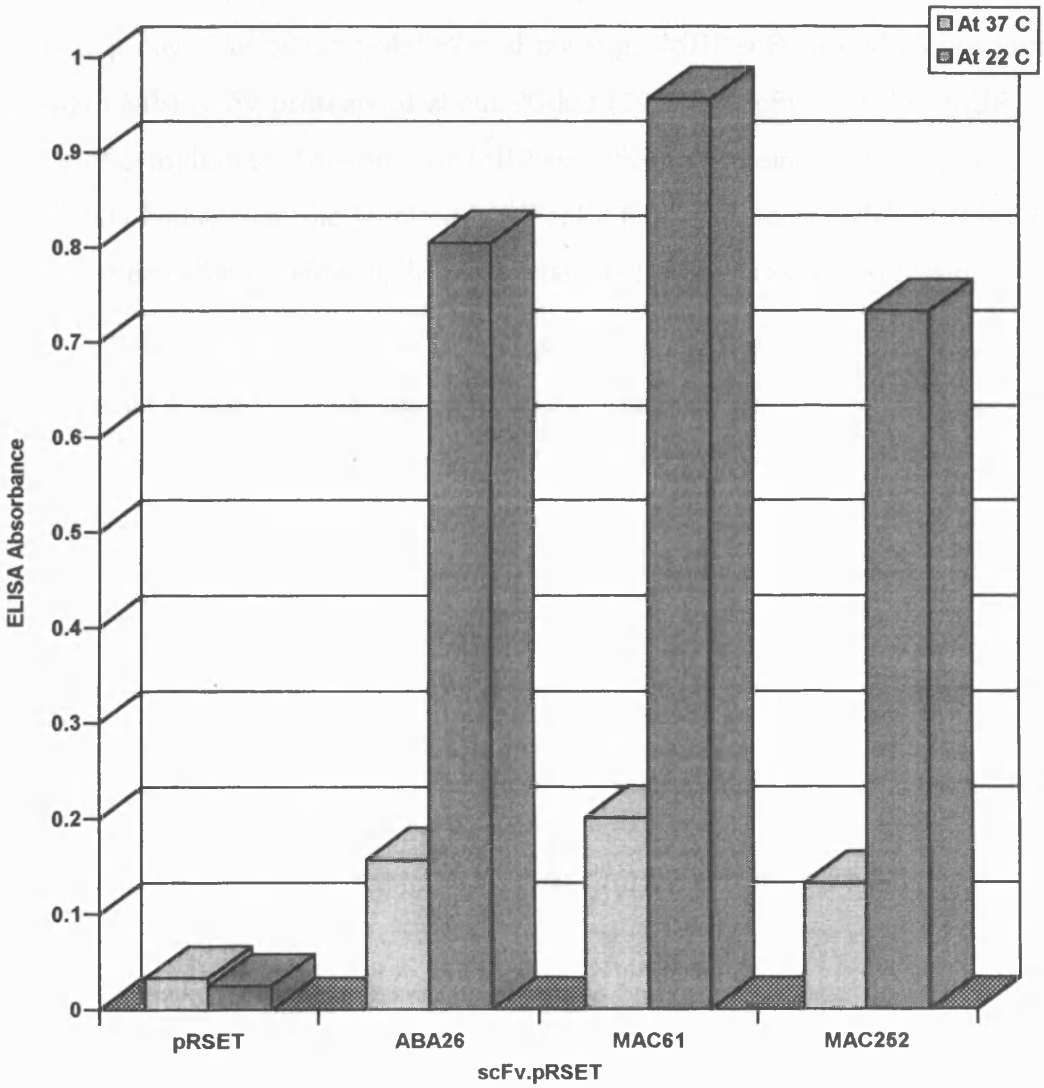
3.2.7.3 Maltose-binding Protein and scFv Gene Fusion.

It is frequently observed that the production of recombinant protein is increased when it is fused at the carboxyl terminal of a highly expressed protein. The *E. coli* maltose-binding protein (MBP) is one of the commonly used fusion protein partners in such strategies (Duplay *et al.*, 1984; Guan *et al.*, 1987; Takagi *et al.*, 1988; Maina *et al.*, 1988). Indeed, Bregegere *et al.* (1994) have fused the variable domains of an immunoglobulin to MBP to increase the expression levels and to facilitate the purification of fusion proteins.

In this study, the effects of fusing the ABA26, MAC61, and MAC252 scFv genes to the carboxyl terminal of MBP were investigated. The fusion was constructed using both the pMal-c2 and pMal-p2 bacterial expression vectors, both of which contain the MBP genes. These expression vectors both utilise the *tac* promoter but

Figure 3.17 : ELISA tests of *E. coli* BL21(DE3) *trxB*⁻ expressed scFv proteins.

In these ELISA tests, 100 µl of 250µg/ml ABA-conjugate was used as coated antigen. 100 µl of the total protein extract from the induced *E. coli* BL21 (DE3) *trxB*⁻ transformed with scFv.pRSET constructs were used as test samples and the anti-HSV-tag antibody was used as the second antibody.



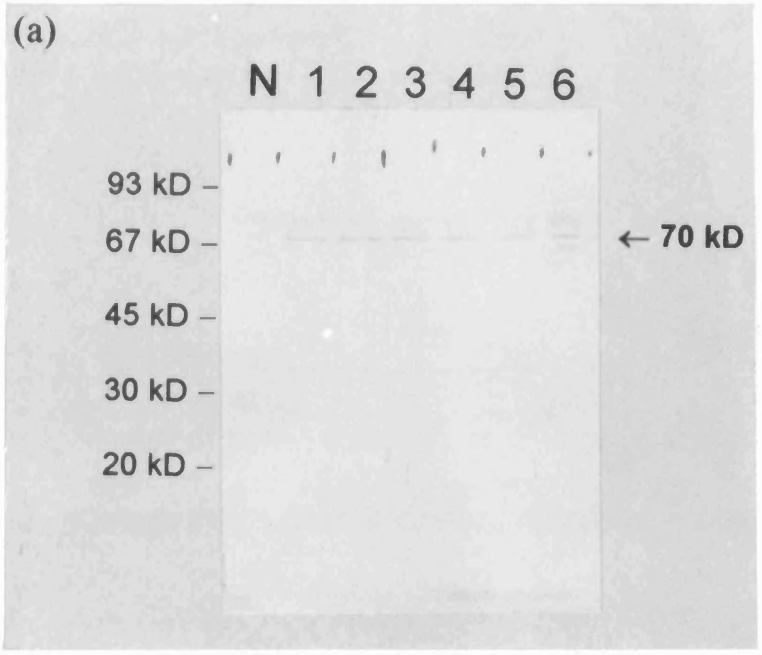
only pMal-p2 includes the *malE* signal peptide sequences to direct secretion of product to *E. coli* periplasm. In order to investigate the effects of fusion of scFv genes to MBP, experiments were also conducted with only the scFv genes which utilised the pABE expression vector. The pABE expression vector is derived from the pMal-p2 expression vector by removal of the MBP gene. The *malE* signal peptide sequence remains intact.

Fig. 3.18 displays bacterial expression data for all of the MBP.scFv gene fusion constructs in both pMal-c2 and pMal-p2 using anti-HSV-tag antibody as scFv protein detection probe. All of the pMal-c2 and pMal-p2 MBP.scFv tested clones produced full-length MBP.scFv proteins of about 70 kD (29.5 kD scFv + 40 kD MBP). The amount of cytoplasmically expressed MBP.scFv fusion proteins (pMal-c2) in this study was slightly higher than the secreted MBP.scFv fusion proteins pMal-p2 (Fig. 3.18b). In contrast no scFv protein was detected when pABE was used as expression vector.

Although these data show an increase in expression level when the scFv genes are expressed as MBP fusions compared to expression of the scFv gene alone using the same promoter, expression levels were still comparatively low. An indication that it is the nature of these particular anti-ABA scFv genes that is responsible for the low expression levels is indicated by experiments using another scFv, G9 (Li, 1996), in the same expression system. Cytoplasmic and secreted MBP.G9 fusion proteins were produced at levels of at least 50 fold and 20 fold greater respectively than the MBP.scFv fusion proteins expressed in this study.

It has been reported that fusion proteins expressed in *E. coli* from pMal-c2 can constitute up to 20-40% of the total cellular protein, while fusion proteins expressed by pMal-p2 can constitute about 5-10% of the total cellular protein (Takagi *et al.*, 1988). However, estimates based on Western blotting indicated that the MBP.scFv fusion proteins expressed in this study represented not more than 0.05 % of the total cellular protein.

When expression was followed over time, it was found that the MBP.scFv fusion proteins were detectable 1 hr after IPTG induction and reached their optimum levels after 2 hr induction (Fig. 3.19). These maximum levels were maintained for at


(a)
N 1 2 3 4 5 6
93 kD –
67 kD –
45 kD –
30 kD –
20 kD –
← 70 kD

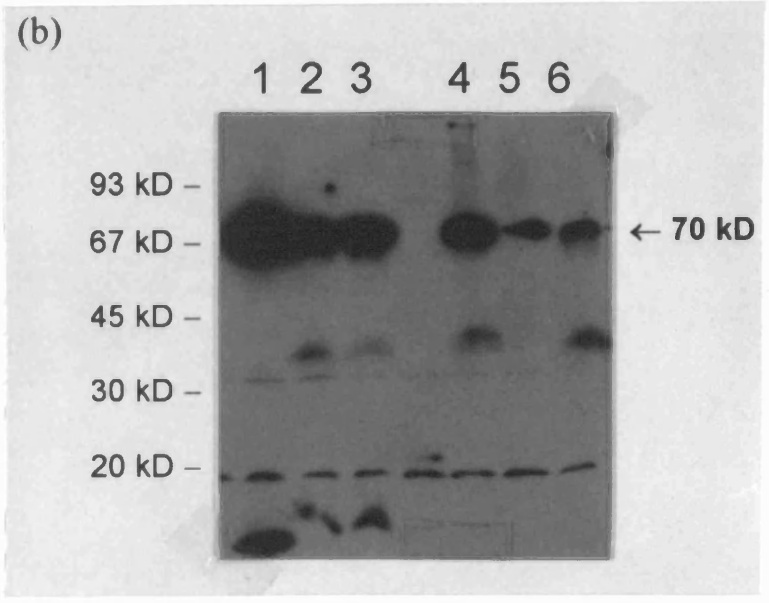
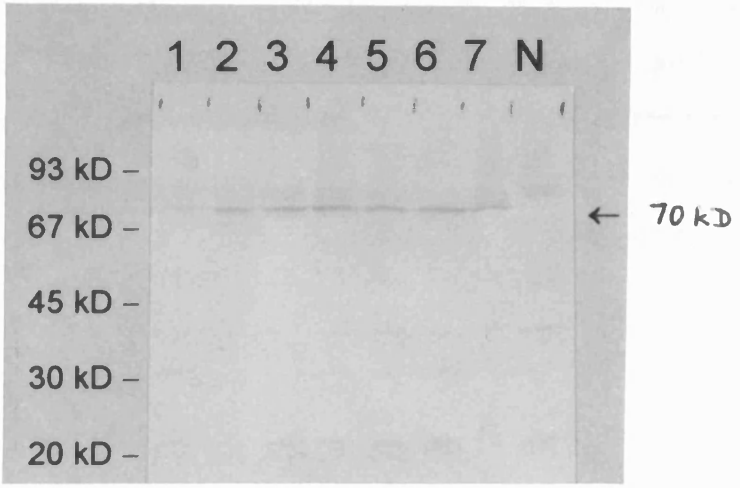

(b)
1 2 3 4 5 6
93 kD –
67 kD –
45 kD –
30 kD –
20 kD –
← 70 kD

Figure 3.19 : Amount of MBP.scFv fusion protein after induction.

The 10 % SDS-acrylamide gel was used to analyse total protein extracts of induced bacterial culture containing MAC61.pMal-c2. The Western Blot membrane was probed using the anti-HSV-tag polyclonal antibody and stained by the colour development reaction. Lanes 1 to 7 represent samples collected after 1, 2, 3, 4, 6, 8, and 10 hr of IPTG induction respectively. Lane N depicts negative control sample of induced *E. coli* transformed with empty vector pMal-c2 after 3 hr of IPTG induction. (Similar results were obtained for the other scFv.pMal-c2 and scFv.pMal-p2 constructs.)

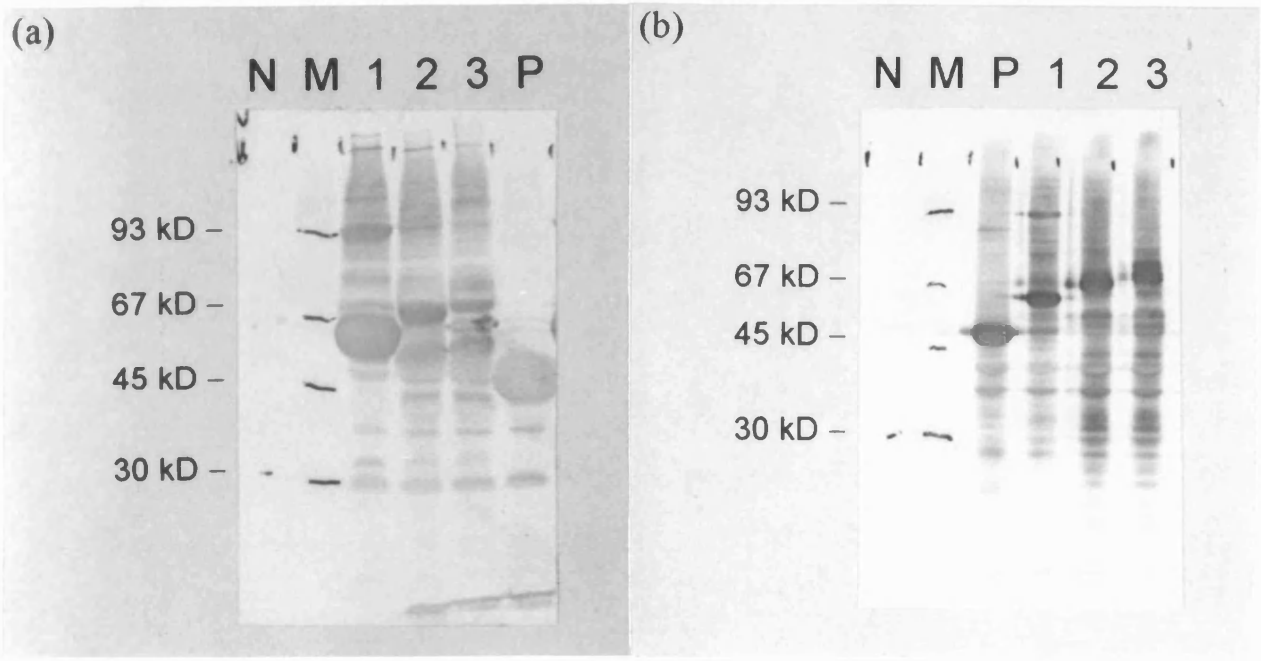


least 10 hr. However, it was difficult to prove whether the expressions stopped after reaching the maximum MBP.scFv fusion proteins level or a balance between production and consumption or degradation was maintained.

Further analysis of the total protein extracts of induced *E. coli* were carried out using anti-MBP polyclonal antibody as the detection probe (Fig. 3.20). Many stained bands were observed. However the polyclonal anti-MBP antibody used was raised using recombinant bacterially-produced MBP as antigen and it is possible that some of the observed bands represented the effects of contaminating anti-bacterial antibodies. However, based on comparisons with the negative control sample (lane N) which did not show any stain, the stained bands are most likely to represent induced MBP fusion proteins. As seen from Fig. 3.20, a major band was observed at all the MBP.scFv transformants. However, the molecular weight of these major stained bands was smaller than the actual molecular weight of MBP.scFv fusion proteins. If these major bands were assumed to be the MBP.scFv fusion protein, then the MBP.scFv fusion proteins were not produced in a complete length. However, it was difficult to know whether those incomplete MBP.scFv fusion proteins were incompletely translated or degraded.

Figure 3.20 : Bacterial expression of MBP.scFv fusion protein probed with anti-MBP antibody.

The 10 % SDS-acrylamide gel was used to analyse total protein extracts of induced bacterial cultures. The Western Blot membrane was probed using the anti-MBP polyclonal antibody and stained by the colour development reaction. Panel (a) shows the results for pMal-c2 constructs while panel (b) depicts the results for pMal-p2 constructs. In both panels, lanes 1, 2, and 3 show the constructs of ABA26, MAC61, and MAC252 MBP.scFv fusions. Lane P and lane N represent positive and negative control samples of induced and uninduced *E. coli* containing pMal vector which expressing 45 kD MBP-fusion protein. Lane M shows the molecular weight markers.



3.2.8 Efforts Towards a Further Understanding of Bacterial Expression of scFv Genes.

PV4.scFv, derived from a monoclonal antibody against potato virus Y, was also included in this study in order to assess the effectiveness of methods used in attempts to express ABA26, MAC61, and MAC252 scFv genes and to provide further general information on the expression of scFv genes in *E. coli* systems. PV4scFv was constructed by Dr. Z.C. Chen, Botany Department, Leicester University; using methods similar to those used in the construction of the other scFv genes in this study.

3.2.8.1 Bacterial Expression of PV4.scFv.

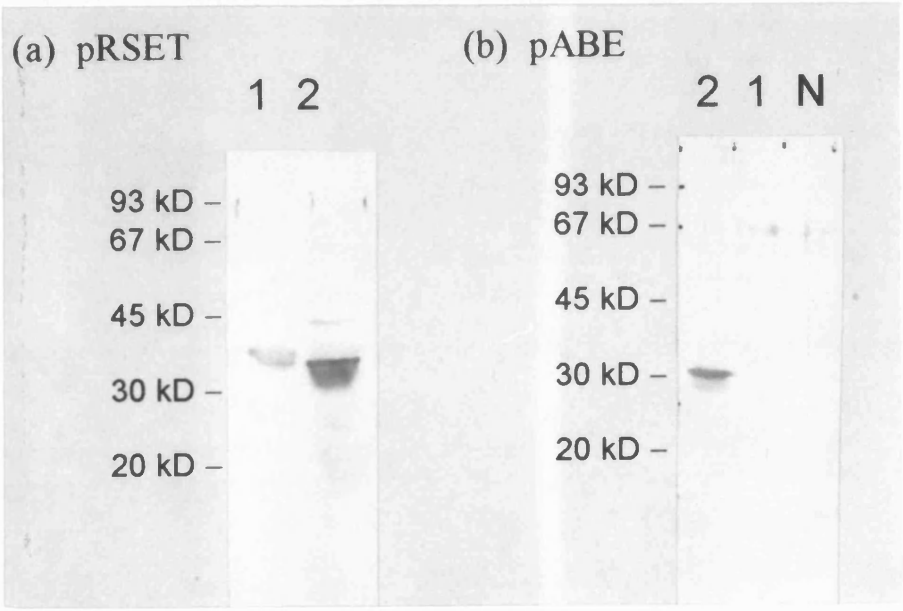
In order to compare its expression level with those of the ABA26, MAC61, and MAC252 scFv proteins, PV4.scFv gene was cloned into the pABE and pRSET expression vectors. *E. coli* W311wt and *E. coli* BL21(DE3) *trxB*$^-$ were used as host strains and expression experiments were carried out side by side with experiments using MAC61.scFv gene.

When the scFv proteins were secreted to the periplasm of *E. coli* W311wt by the pABE expression vector PV4.scFv, in marked contrast to MAC61.scFv, was found to express scFv protein which was easily detected on Western Blot using colorimetric detection method (Fig. 3.21a). In the case of cytoplasmic expression with the pRSET vector using *E. coli* BL21(DE3) *trxB*$^-$ as host, both the PV4.scFv and the MAC61.scFv proteins were detectable although the level of PV4.scFv protein was at least 10 times (based on the Western Blot) more than that of the MAC61.scFv protein (Fig. 3.21b).

Despite the close similarity of the gene constructs encoding the four scFv genes in this study they vary greatly in their 'expressibility' in *E. coli*. PV4.scFv gene is readily expressed, MAC61.scFv gene is difficult to express and ABA26.scFv and MAC252 genes proved to be very difficult.

Figure 3.21 : Comparison of bacterial expression levels of PV4 and MAC61 scFv proteins.

The 12 % SDS-acrylamide gel was used to analyse total protein extract of induced bacterial cultures containing (a) pRSET constructs and (b) pABE constructs. The Western Blot membrane was probed using the anti-HSV-tag antibody and stained by the colour development reaction. In both panels, lane 1 shows the MAC61.scFv protein while lanes 2 depicts the PV4.scFv protein. Lane N represents the negative control sample of induced *E. coli* transformed with empty vector.



3.2.8.2 Expression of Single V_L or V_H Domains in *E. coli*.

To investigate characteristics of the scFv genes which might influence expression of scFv genes in *E. coli*, single V_L and V_H domains of MAC61.scFv gene were expressed separately. The separate, single variable domains of MAC61 gene were cloned into the pRSET expression vector and expressed using *E. coli* BL21(DE3) *trxB*$^-$ as host strain. Fig. 3.22 shows that the V_H domain of MAC61 accumulated to high levels whereas the V_L domain accumulated to much lower levels. Estimated from Western Blots the level of accumulation of the single V_H domain of MAC61 was at least 10 fold more than that of the complete MAC61.scFv protein. In contrast, the single V_L domain of MAC61 accumulated to around the same level as the complete MAC61.scFv protein. These data suggest that the V_L domain of MAC61 might be responsible for the difficulties encountered in expressing the complete MAC61scFv gene.

3.2.8.3 Hybrid scFvs.

In order to further investigate the effects of individual variable domains on the expression levels of scFv proteins in *E. coli*, hybrid scFv genes were constructed. The V_L and V_H domains of MAC61 were fused to the amino terminals V_H and V_L domains of PV4.scFv respectively to create the hybrid scFv genes of VL_{61}-VH_{P4}.scFv and VL_{P4}-VH_{61}.scFv. The hybrid scFv genes were then expressed using the pRSET expression vector and *E. coli* BL21(DE3) *trxB*$^-$ as host.

Hybrid VL_{P4}-VH_{61}.scFv protein was found to accumulate to very high levels whereas hybrid VL_{61}-VH_{P4}.scFv protein accumulated to a much lower level, similar in fact to the level of the intact MAC61.scFv protein expressed by similar vector and *E. coli* host (Fig. 3.23). Thus scFv genes which carried the V_L domain of MAC61 were expressed at low levels, an effect not observed in scFv genes which carried the V_H domain of MAC61. This finding is again consistent with the possibility that some characteristic of the V_L domain of MAC61 adversely affects the expression in *E. coli* of MAC61.scFv gene.

Figure 3.22 : Bacterial expression of the single V_L and V_H domains of MAC61.

The 12 % SDS-acrylamide gel was used to analyse total protein extract of induced bacterial cultures. The Western Blot membrane was probed using the anti-HSV-tag antibody and stained by the colour development reaction. Lane N represents the negative control sample transformed with empty pRSET vector, lanes 1 shows the MAC61.scFv protein while lane 2 and 3 depict the single V_L and V_H domains of MAC61 respectively.

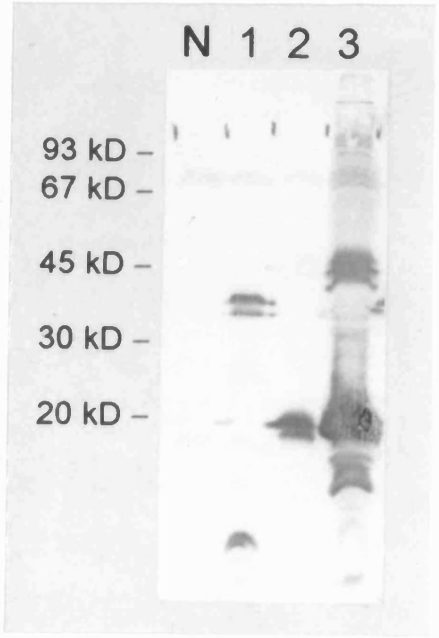
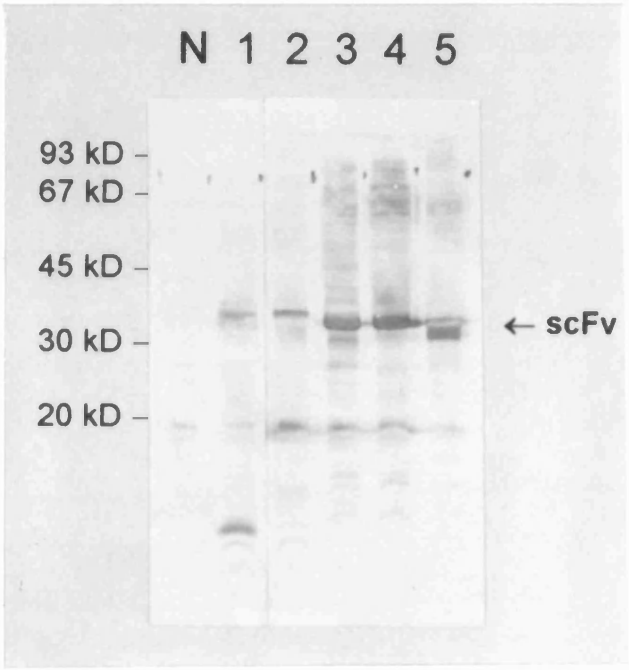


Figure 3.23 : Bacterial expression of PV4 and MAC61 hybrid scFvs.

The 12 % SDS-acrylamide gel was used to analyse the total protein extracts of induced bacterial cultures. The Western Blot membrane was first probed using the anti-HSV-tag antibody and later stained by the colour development reaction. Lane N represents the negative control sample of induced *E. coli* transformed with an empty pRSET vector. Lane 2 and lanes 3 & 4 represent the hybrid scFv proteins of VL_{61}-VH_{P4}.scFv and VL_{P4}-VH_{61}.scFv respectively. Lane 1 shows the MAC61.scFv protein while lane 5 shows the PV4.scFv protein.



3.2.8.4 Domain Orientation in scFv Gene.

All of the anti-ABA scFv genes in this study were initially constructed with the orientation, 5'-V_L-linker-V_H-tag-3'. In order to investigate possible effects of domain orientation on expression of the scFv genes in *E. coli*, another version of the MAC61.scFv gene with the orientation 5'-V_H-linker-V_L-tag-3' was constructed. An additional possible effect of reversing the variable domains relates to the fact that the most highly expressed of the two domains (V_H domain) would then be translated first. Translation of a highly expressed fusion partner in advance of the gene of interest is believed to be the basis for the expression-enhancing effects of fusion protein strategies such as that employing MBP described above. Thus translation of the V_H domain in advance of the V_L domain might increase the level of expression of the MAC61.scFv gene.

Fig. 3.24 illustrates the results of bacterial expression experiments for both the 5'-V_H-linker-V_L-tag-3' and 5'-V_L-linker-V_H-tag-3' versions of MAC61.scFv gene. Regardless of orientation of the MAC61.scFv variable domains, the levels of expression were similar. Thus it can be concluded that in the case of MAC61.scFv gene the orientation of the V_H and V_L domains does not significantly affect expression levels.

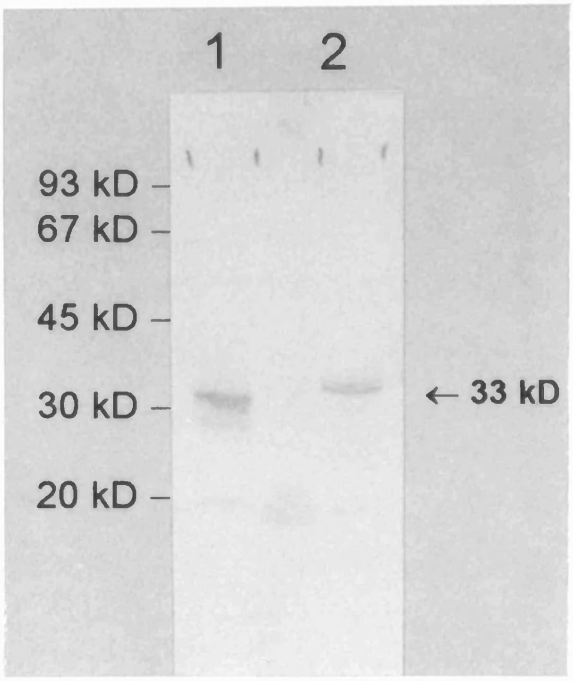
3.2.8.5 DNA Sequences.

The results of the experiments reported here indicated that the difficulties encountered in expressing ABA26, MAC61, and MAC252 scFv genes in *E. coli* is accounted for by characteristics of their gene sequences which are specific to these particular scFv genes. In the case of MAC61.scFv gene, it appears that it is the V_L domain which limits expression of scFv genes containing it in *E. coli*.

In an attempt to determine the characteristics of scFv genes which influence the expression level in *E. coli*, comparisons were made between the DNA (and the corresponding amino acids) sequences of ABA26, MAC61, and MAC252 with other expressible scFv genes for which there is information on expression level in *E. coli*.

Figure 3.24 : Bacterial expression of MAC61.scFv with different domain arrangements.

The 12 % SDS-acrylamide gel was used to analyse total protein extract of induced bacterial cultures. The Western Blot membrane was probed using the anti-HSV-tag antibody and stained by the colour development reaction. Lane 1 represents the 5'-V_H-linker-V_L-tag-3' and lane 2 depicts the 5'-V_L-linker-V_H-tag-3' of MAC61.scFv proteins. (Note that the size of the 5'-V_H-linker-V_L-tag-3' protein was smaller than the 5'-V_L-linker-V_H-tag-3' protein. This result from removed of the 15 amino acids constant region tail at the carboxyl terminal of the V_H domain of 5'-V_H-linker-V_L-tag-3' construct to suit the linker used.)



However, due to the huge variation in DNA and amino acid sequences and only a small number (23 in total) of expressible scFv genes for which information was obtainable, no particular genetic patterns or trends could be identified for either the DNA or amino acid sequences.

Some interesting features were noted in relation to the VL domains of MAC61 and PV4 scFv genes. Both of the VL domains showed strong similarities, especially in the FR regions. Only 18 out of 85 amino acid residues in the FR regions and 16 out of 31 amino acid residues of the CDR regions were different (Table 3.3). Despite these similarities, the two scFv genes were expressed very differently in *E. coli*. It appears that some aspect of these relatively small differences in the VL amino acid sequence are responsible for the low expression of the MAC61.scFv gene in *E. coli*.

Other characteristics of the VL domain of MAC61.scFv gene were also taken into consideration. These were the translation initiation region, codon usage favoured by *E. coli*, and possible proteinase digestion sites.

It appears unlikely that the translation initiation region of the VL domain of MAC61.scFv gene limits the expression level in *E. coli* for the following reasons. Among the first 24 amino acid residues of the VL domain of the MAC61.scFv gene, only 5 amino acid residues, residues 3, 9, 11(12), 16, and 17 were different from the VL domain of the PV4.scFv gene (Table 3.4). Furthermore most of the differences involve amino acids with similar characteristics, i.e. acidic aspartic acid at residue 3 of the VL domain of the MAC61 is substituted by an acidic glutamic acid at residue 3 of the VL domain of PV4 (Table 3.4).

In terms of codon usage (Table 3.5), most of the codons in the VL domain of MAC61.scFv were of similar patterns to those used by bacteria except the codons for arginine. Previous work by Spanjaard *et al.* (1990) demonstrated that gene expression in *E. coli* was suppressed by either AGA and AGG arginine codons since in *E. coli* there is limiting amounts of tRNA for AGA and AGG. These codons are rarely found in bacterial genes whereas the VL domain of the MAC61 gene was found to contain 12 arginine residues of which 3 were encoded by AGA and 5 by AGG. The presence of these codons does not however necessarily preclude the expression of scFv genes

Table 3.3: Alignment of ABA26, MAC61, and MAC252 scFv amino acid sequences with PV4 V_L domain.

V_L domain	FR1	CDR1
PV4	MA EILL TQF PAS LAVPL GQRATISY	RASKSV STSGYSYMH
MAC61	-- D--- --S -T ---S- -------C	---Q-- TIPSVNL-N
ABA26	-- DVQM I-T PS- -SAS- -ERVSLTC	---Q DIGSSLN
MAC252	-- DVKI I-T T-- -SAS- -ETVSIEC	LASE DIYSYLA
	FR2	CDR2
PV4	WNQQ KPGQ PPR LLIY	LVSN LES
MAC61	-Y-- ---- Q-K ----	HA-- -G-
ABA26	-F-- E-DG TIK R---	AT-S -DS
MAC252	-Y-- K--K S-Q ----	ASNR -QD
	FR3	CDR3
PV4	GVPA RFSGSGSG TDFTL NIHPVEE EDPAT YYC	QHIRE SPLT
MAC61	-I-T -------- ----- T-D--QA D-I-- ---	-QS-- --P-
ABA26	---K -----R-- S-YS- T-SRL-S --FVD ---	LQYAS --W-
MAC252	---S -------- -QYS- K-SGMQP --EGR ---	LQGSK --P-
	FR4	
PV4	FGGG TKLESK PG	
MAC61	---- ----IR --	
ABA26	---- ---KL- --	
MAC252	---- ----LR --	

Note : Symbol "-" indicated the similar amino acid as compared to PV4.

Table 3.4 : Translation initiation region of MAC61.scFv and PV4.scFv genes.

The table shows the first 24 amino acids sequence in the translation initiation region of MAC61.scFv and PV4.scFv genes. The differences between the two sequences are highlighted.

MAC61.scFv : 5'-Met-Ala-**Asp**-Ile-Leu-Leu-Thr-Gln-**Ser**-Pro-**Thr**-==-Leu-Ala-Val- **Ser-Pro**-Gly-Gln-Arg-Ala-Thr-Ile-Ser----------3'
PV4.scFv : 5'-Met-Ala-**Glu**-Ile-Leu-Leu-Thr-Gln-**Phe**-Pro-**Ala-Ser**-Leu-Ala-Val- **Pro-Leu**-Gly-Gln-Arg-AlaThr-Ile-Ser----------3'

in *E. coli* since the AGA and AGG codons are commonly found in the V_L domains of other highly expressed scFv genes including PV4.

The presence of possible proteinase digestion sites in the V_L domain of MAC61 were investigated by the Genetics Computer Group Inc. (GCG) computer programme. Both the V_L domains of MAC61 and PV4 were discovered to be potentially digestible by Tryp, Chym, Clos, CnBr, IBzO, Myxo, NTCB, ProEn, Staph, Tryp K, Tryp R, and AspN.

On the basis of these DNA and amino acid sequence analysis, it is difficult to draw any conclusion as to possible ways of improving the expression of the MAC61.scFv gene in *E. coli*. However, further progress in characterising differences which are significant in terms of affecting expression levels in *E. coli* could be achieved by mutation experiments focusing on the differences in amino acid sequences between the V_L domains of MAC61, and PV4.

Table 3.5 : Codons usage in MAC61.scFv gene.

The table shows the number of each codon used by MAC61.scFv gene. While the number in the brackets indicates the frequency of codons usage by *E. coli* genes (Gratham *et al.*, 1981).

Ala	GCA = 3,(31)	Gly	GGA = 7,(4)	Pro	CCA = 5,(7)
Ala	GCC = 8,(20)	Gly	GGC = 11,(29)	Pro	CCC = 4,(4)
Ala	GCG = 1,(23)	Gly	GGG = 7,(6)	Pro	CCG = 2,(19)
Ala	GCT = 3,(37)	Gly	GGT = 11,(33)	Pro	CCT = 9,(5)
Arg	**AGA = 3,(5)**	His	CAC = 1,(9)	Ser	AGC = 5,(9)
Arg	**AGG = 5,(2)**	His	CAT = 1,(16)	Ser	AGT = 7,(7)
Arg	**CGA = 4,(3)**			Ser	TCA = 4,(8)
Arg	**CGC = 0,(21)**	Ile	ATA = 1,(5)	Ser	TCC = 8,(13)
Arg	**CGG = 0,(4)**	Ile	ATC = 4,(32)	Ser	TCG = 0,(9)
Arg	**CGT = 0,(30)**	Ile	ATT = 7,(24)	Ser	TCT = 9,(17)
Asn	AAC = 7,(25)	Leu	CTA = 0,(2)	Thr	ACA = 5,(5)
Asn	AAT = 1,(10)	Leu	CTC = 6,(8)	Thr	ACC = 12,(22)
		Leu	CTG = 9,(47)	Thr	ACG = 1,(10)
Asp	GAC = 5,(27)	Leu	CTT = 2,(8)	Thr	ACT = 3,(21)
Asp	GAT = 7,(25)	Leu	TTA = 3,(7)		
		Leu	TTG = 2,(7)	Trp	TGG = 6,(12)
Cys	TGC = 1,(6)				
Cys	TGT = 3,(5)	Lys	AAA = 3,(46)	Tyr	TAC = 7,(12)
		Lys	AAG = 5,(18)	Tyr	TAT = 6,(14)
Gln	CAA = 7,(13)				
Gln	CAG = 11,(29)	Met	ATG = 4,(22)	Val	GTC = 7,(9)
				Val	GTG = 4,(17)
Glu	GAA = 4,(37)	Phe	TTC = 5,(18)	Val	GTT = 4,(28)
Glu	GAG = 6,(18)	Phe	TTT = 0,(18)		

3.3 Summary and Conclusions.

Three anti-ABA scFv genes, namely ABA26, MAC61, and MAC252, have been constructed using PCR amplification of antibody variable domains from mouse and rat hybridoma cDNAs. All of these scFv genes could be expressed in *E. coli* using the T7 promoter, albeit at comparatively low levels. When expression was driven by the *tac* or *trc* promoter, scFv protein was not detectable for any of the constructs.

No significant differences in expression levels were observed between production of the scFv proteins in the bacterial cytoplasm using the pRSET expression vector (through M13/T7 phage infection) or when they were targeted to the periplasm with the pTM-22 expression vector. The cytoplasmically located scFv proteins were in the form of insoluble fraction and did not therefore exhibit any binding activities to the ABA.. The molecular weight of each pRSET expressed scFv protein (with the His-tag fusion partner) was 34 kD, consistent with calculations based on the DNA sequences. Expression of the scFv genes, using pTM-22 also resulted in the major proportion of the scFv protein being accumulated in insoluble form but in this case the observed molecular weight was 32 kD. This is equivalent to the molecular weight of the scFv protein plus the molecular weight of the *ompA* signal peptide. It was concluded that the majority of the pTM-22 expressed scFv proteins retained the *ompA* signal peptide and were probably therefore also located in the bacterial cytoplasm as insoluble inclusion bodies. However, based on the positive signal in ELISA test, it was believed that a small portion of the pTM-22 expressed scFv proteins were in the form of soluble functional scFv proteins.

Competition ELISA using free ABA demonstrated that all three pTM-22 expressed scFv proteins, ABA26, MAC61, and MAC252, had specific ABA binding activity. However, the affinity constants of the scFv proteins were found to be 5 to 10 fold lower than those of their parental MAbs. The pTM-22 expressed MAC61 and MAC252 scFv proteins were also found to possess less specific binding sites compared to their parental MAbs. Thus MAC61.scFv and MAC252.scFv proteins were able to bind to both ABA-C1-conjugate and ABA-C4'-conjugate although their parental MAbs were specific for the ABA-C4'-conjugate against which they were raised. In common

3.3 Summary and Conclusions.

Three anti-ABA scFv genes, namely ABA26, MAC61, and MAC252, have been constructed using PCR amplification of antibody variable domains from mouse and rat hybridoma cDNAs. All of these scFv genes could be expressed in *E. coli* using the T7 promoter, albeit at comparatively low levels. When expression was driven by the *tac* or *trc* promoter, scFv protein was not detectable for any of the constructs.

No significant differences in expression levels were observed between production of the scFv proteins in the bacterial cytoplasm using the pRSET expression vector (through M13/T7 phage infection) or when they were targeted to the periplasm with the pTM-22 expression vector. The cytoplasmically located scFv proteins were in the form of insoluble fraction and did not therefore exhibit any binding activities to the ABA.. The molecular weight of each pRSET expressed scFv protein (with the His-tag fusion partner) was 34 kD, consistent with calculations based on the DNA sequences. Expression of the scFv genes, using pTM-22 also resulted in the major proportion of the scFv protein being accumulated in insoluble form but in this case the observed molecular weight was 32 kD. This is equivalent to the molecular weight of the scFv protein plus the molecular weight of the *ompA* signal peptide. It was concluded that the majority of the pTM-22 expressed scFv proteins retained the *ompA* signal peptide and were probably therefore also located in the bacterial cytoplasm as insoluble inclusion bodies. However, based on the positive signal in ELISA test, it was believed that a small portion of the pTM-22 expressed scFv proteins were in the form of soluble functional scFv proteins.

Competition ELISA using free ABA demonstrated that all three pTM-22 expressed scFv proteins, ABA26, MAC61, and MAC252, had specific ABA binding activity. However, the affinity constants of the scFv proteins were found to be 5 to 10 fold lower than those of their parental MAbs. The pTM-22 expressed MAC61 and MAC252 scFv proteins were also found to possess less specific binding sites compared to their parental MAbs. Thus MAC61.scFv and MAC252.scFv proteins were able to bind to both ABA-C_1-conjugate and ABA-C4'-conjugate although their parental MAbs were specific for the ABA-C4'-conjugate against which they were raised. In common

with the monoclonal antibody from which it was derived, ABA26.scFv protein displayed binding activities to both the ABA-C1-conjugate and ABA-C4'-conjugate.

To investigate factors determining expression levels of scFv proteins in bacterial systems, several different expression strategies were employed. With pABE and pTrcHis expression vectors, neither modification of expression conditions nor the bacterial host strain used resulted in any detectable scFv protein with any of the anti-ABA scFv genes investigated.

Expression to detectable levels was observed however when the scFv proteins were produced as fusion proteins downstream of maltose-binding protein (MBP) using pMal-c2 and pMal-p2 expression vectors. Although cytoplasmically expressed MBP.scFv proteins showed slightly higher expression level compared to the periplasmically targeted MBP.scFvs, expression levels were generally low.

Expression of scFv genes using pRSET was found to result in the accumulation of higher levels of scFv protein when *E. coli* BL21(DE3) *trxB*$^-$ was used as host strain compared to *E. coli* XL-1 Blue with M13/T7 phage infection. A proportion of the *E. coli* BL21(DE3) *trxB*$^-$ produced scFv proteins were found to be soluble and functional in terms of binding activity to ABA.. Although similar amounts of scFv proteins accumulated at both 22°C and 37°C, using *E. coli* BL21(DE3) *trxB*$^-$ expression at 22°C produced a greater proportion of soluble , functional scFv protein.

• An investigation was also conducted to investigate possible causes of the low expression levels in *E. coli* of one of the anti-ABA scFv protein, namely MAC61. Separate expression of single V_L and V_H domains of MAC61 in *E. coli* indicated that the single V_H domain accumulates to levels at least 10 fold higher than the single V_L domain (which accumulates to approximately to the same level as the complete MAC61scFv).

A hybrid scFv gene which contained the V_L domain of MAC61 and the V_H domain of a readily expressed scFv (PV4.scFv) was found to express at a much lower level than the hybrid scFv gene comprised of the V_L domain of PV4 and the V_H domain of MAC61. In fact the hybrid scFv which contained the V_H domain of MAC61 was expressed at a similar level to the complete PV4.scFv. Thus it appears

that it is the V_L domain of MAC61.scFv gene which has adverse effects on the expression of this scFv gene in *E. coli*.

The amino acid sequence of the V_L domain of MAC61 gene was found to be very similar to the V_L domain of PV4 gene, especially in the FR regions. Future experiments involving mutations of the relatively few different amino acids could allow precise identification of the characteristics of the V_L domain of MAC61 which are responsible for the adverse effects on expression in *E. coli*.

CHAPTER 4

Expression of scFv genes in Yeast and Plants.

4.1 Introduction.

Although the original source of antibody genes is usually mammalian cells, recombinant antibodies or antibody fragments can be expressed in functional form in a range of non-mammalian eukaryotic cells such as yeast (Horwitz *et al.*, 1988, Wood *et al.*, 1985), insects (Potter *et al.*, 1993), and plants (Hiatt and Ma, 1993).

In this study, yeast and plants were investigated as hosts for the expression of the anti-ABA scFv genes which had proved moderately difficult (MAC61.scFv) or very difficult (ABA26.scFv and MAC252.scFv) to express using *E. coli* as host. The PV4.scFv which was readily expressed in bacterial systems was also included in the study.

4.1.1 Yeast Expression Systems.

As a host for the production of heterologous eukaryotic proteins, yeast combines the advantages of the ease of molecular genetic manipulation and the growth characteristics of prokaryotic organisms together with the eukaryotic subcellular machinery for post-translational protein modification. Although the species of yeast, *Saccharomyces cerevisiae* was the first to be used for this purpose, recent developments have shown that the methylotrophic yeast, *Pichia pastoris* has many advantages as an expression host (Buckholz and Gleeson, 1991).

A significant feature of the *P. pastoris* expression system is a direct consequence of the inherent transcriptional properties of the promoter most commonly used to control foreign gene expression. This promoter is derived from the alcohol oxidase (*AOX*) genes of *P. pastoris* that are expressed in methanol-grown cells. The *AOX* enzyme is undetectable in normal *P. pastoris* cells cultured on carbon sources such as glucose, glycerol, or ethanol, but constitutes up to 30% of the total soluble protein in methanol-grown cells (Coudere, and Baratti, 1988). It was shown that *AOX*

synthesis was regulated at the transcriptional level and that the promoter from this gene would be most useful for controlling the expression of foreign genes. The *P. pastoris* expression system has now been successfully utilised to produce a number of heterologous proteins at high concentrations, some at levels above 10 g/l (Sreekrishna, *et al.*, 1989; Clare, *et al.*, 1991).

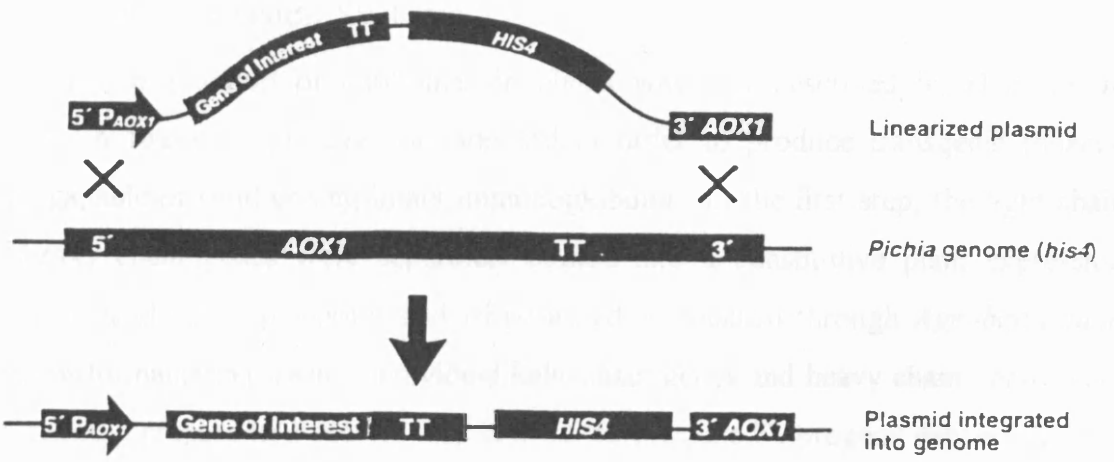
Three different *Pichia pastoris* expression systems are currently available commercially, namely the "original" system, the "multi-copy" system, and the "easy-selection" system (Fig. 4.1). The "original" system makes use of *Pichia* expression vectors which carry the *AOX1* promoter for high level expression, the *HIS4* gene for selection in the *his4* mutant strain, and the *3'AOX1* gene for targeting the integration into the *Pichia* genome. The "multi-copy" system is basically similar to the original system but contains an additional kanamycin resistance gene that confers resistance to G418 for selection of multiple gene insertions in *P. pastoris*. The "easy-selection" system uses smaller *Pichia* expression vectors which carry the *AOX1* promoter for high level expression and the Zeocin resistance gene with *EM7* promoter and *TEF1* promoter for clone selection in *E. coli* and also in *P. pastoris*.

Homologous recombination between the linearised transforming vector and the regions of homology within the *Pichia* genome is the method most commonly used to transform genes of interest into the *Pichia* genome (Cregg *et al.*, 1989). Transgene integration can be achieved either by gene replacement or gene insertion (Fig. 4.1). In the case of gene insertion, a single crossover occurs between the loci and the similar DNA regions (i.e. 3'-*AOX1* region) on the vector. In gene replacement, a double crossover occurs between the vector and the genome (i.e. the *AOX1* promoter and 3'-*OAX1* region). It has been found that the single crossover events (gene insertions) are much more likely to occur than the double crossover events (gene replacements), and that multiple copy genes insertion events occurred spontaneously at about 1-10% of the gene insertion events (Cregg *et al.*, 1989). Among the three *Pichia* expression systems, only the "original" system is capable of performing both gene replacement and gene insertion. Although the "multi-copy" system can perform both gene replacement and gene insertion, the integration of multiple copy genes is only possible through gene insertion. The "easy-selection" system, on the other hand, can only carry out gene

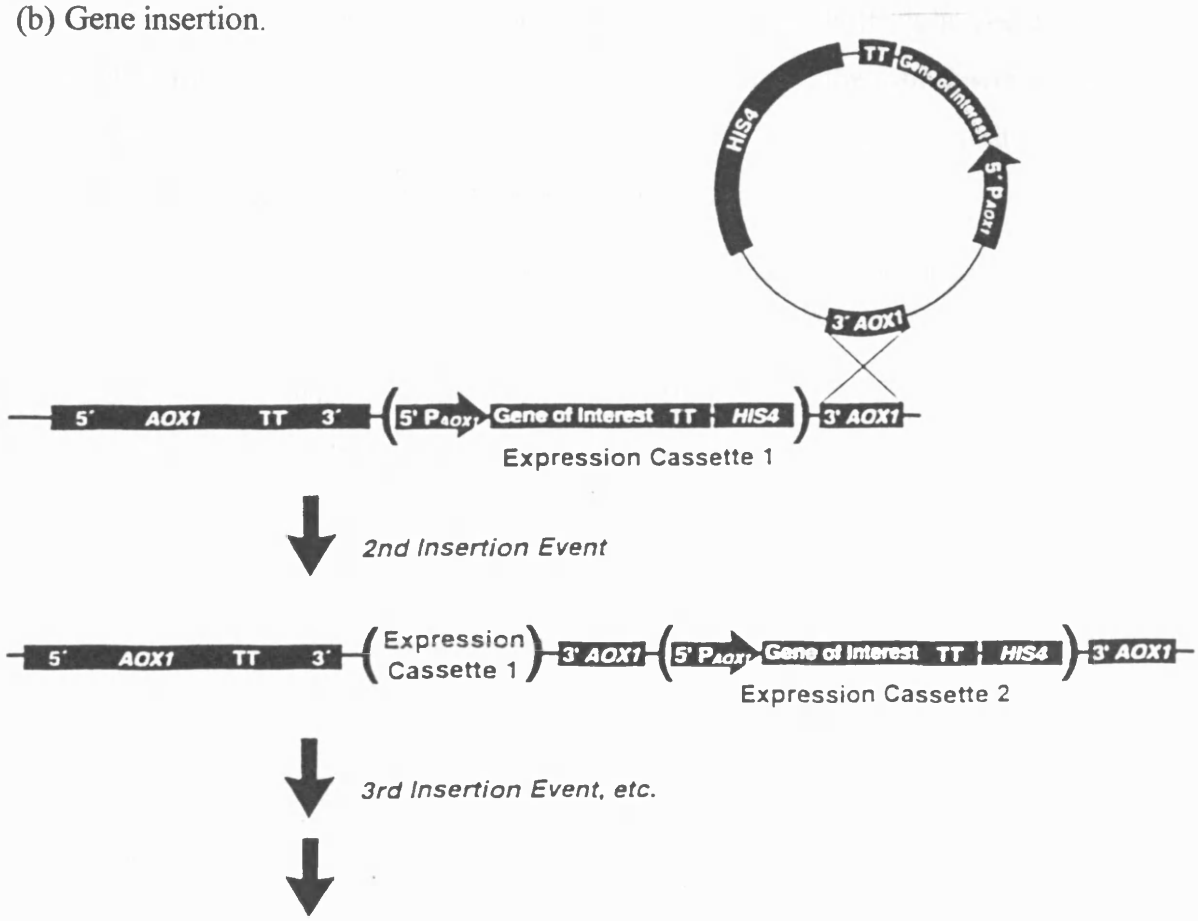
Figure 4.1 : Gene integration in *P. pastoris* expression systems.

Schematic diagrams showing three different versions of *P. pastoris* expression vectors and the possible integration of scFv genes into the *P. pastoris* genome, (a) gene replacement and (b) gene insertion.

(a) Gene replacement.



(b) Gene insertion.



insertion because the vector only contains a single homologous region (the 3'*AOX* region is not present).

All three *Pichia* expression systems were investigated in this study for expressing scFv genes.

4.1.2 Plant Expression Systems.

The production of antibodies in plants was first described by Hiatt *et al.* (1989). A two-step strategy was adopted in order to produce transgenic tobacco plants capable of synthesising intact immunoglobulin. In the first step, the light chain and heavy chain genes were separately ligated into a constitutive plant expression vector using the 35S promoter and transformed to tobacco through *Agrobacterium*. The transformants expressing individual light chain genes and heavy chain genes were then sexually crossed in a second step in order to produce F1 progeny expressing both heavy and light chain genes. In an alternative strategy, double transformations were performed using a single expression cassette containing both light and heavy chain genes (Düring *et al.*, 1990; Voss *et al.*, 1994). Routinely, the cauliflower mosaic virus (CaMV) 35S promoter has been used to drive the expression of antibody genes in tobacco following *Agrobacterium* mediated transformation.

Plant transformation with scFv is more straightforward, as scFv proteins are encoded by single continuous genes. The correct folding of scFv protein is a prerequisite for antigen binding but scFv protein does not have the assembly requirements of complete antibodies. The first functional scFv in plants, AS32.scFv was reported by Owen *et al.* (1992). The AS32.scFv gene was transformed into tobacco where expression was driven by the CaMV 35S promoter. Despite the reported successes, intracellular expression of scFv proteins in plants is not always straightforward. Intracellular expression levels are usually low and are sometimes undetectable in most of the transformants produced. (Owen *et al.*, 1992; Artsaenko *et al.*, 1995). However, the secretory version of scFv genes always produces more scFv proteins in plants than the cytoplasmically targeted version (Firek *et al.*, 1993, Artsaenko *et al.*, 1995). Furthermore, scFv expression level has been shown to be

enhanced when the C-terminal of the scFv was fused to an endoplasmic reticulum (ER) retention signal peptide, KDEL (Artsaenko *et al.*, 1995; Schouten *et al.*, 1996). Consequently, several derivatives of scFv genes have been constructed in this study to target to a range of different subcellular compartments. These include the cytoplasmic expression of scFv genes (scFv only), secretory scFv genes which contained the signal peptide (SS) coding region derived from the tobacco PR1-a gene, chloroplast targeted scFv using the TP signal peptide, and scFv targeted to and retained in the ER by the PR1-a signal peptide and KDEL signal peptides.

Two types of plant expression systems, the transient plant expression and stable plant transformation, were used to express the various PV4, MAC61, ABA26, and MAC252 scFv derivatives in tobacco plants.

4.1.2.1 Transient Plant Expression.

Episomal plant viral vectors have been developed as an alternative to stable transformation for the expression of recombinant genes in plants. Such vectors may have problems of genetic instability and, of course, the transgene is not transmitted to progeny. They offer the important advantage of speed to assess the expression of transgene. Plant expression by viral vectors usually can be assessed within 7 to 10 days while the stable plant transformation requires several weeks to obtain transgenic plants for assessment. Furthermore, viral expression systems often allow the accumulation of substantial quantities of the transgene product (Joshi *et al.*, 1990; Ahiquist and Pacha, 1990).

Several plant viral vectors have been produced for this purpose. These include the DNA viral vectors constructed from CaMV (Brisson *et al.*, 1984) and the cassava latent virus (Haynes *et al.*, 1988), and RNA viral vectors from TMV (Dawson *et al.*, 1989; Donson *et al.*, 1991) and PVX (Chapman *et al.*, 1992). The PVX vector was used for the transient expression of scFv genes in tobacco plants in this study.

The PVX expression vector was created from potato virus X, which is the type member of the potex virus group (Chapman *et al.*, 1992). PVX is a single-stranded RNA virus which is capped at the 5'-end and polyadenylated at the 3'-end. The

positive-sense genome (Fig. 4.2) contains five large open reading frames (ORFs). The farthest 5'-end of these genes encodes a protein of 166 kD which is involved in viral replication (Huisman *et al.*, 1988), followed by three partially overlapping ORFs encoding proteins of 25, 12 and 8 kD which are believed to be involved in plant infection or cell to cell movement, and finally at the 3'-end, is the gene which encodes the coat protein.

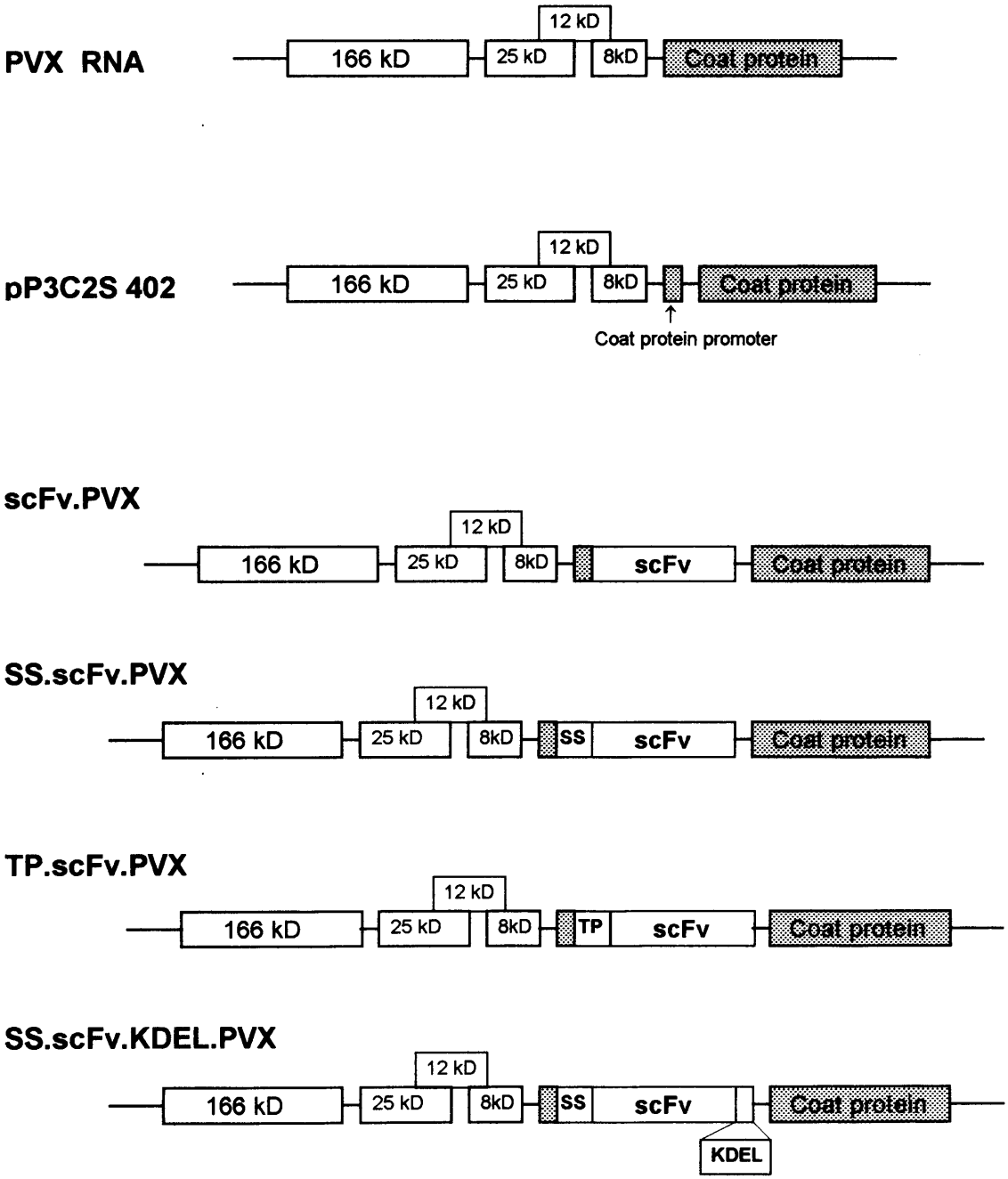
The PVX plant viral expression vector (pP3C2S 402) contains the full-length reverse transcribed PVX viral RNA and a duplicate part of the translation initiation region of coat protein gene into which transgenes are inserted (Fig. 4.2). This arrangement enables the expression of the inserted transgene to be under the control of a very strong promoter sequence that controls production of the coat protein mRNA during virus replication in infected plants.

4.1.2.2 Stable Plant Transformation.

Stably transformed plants can be produced by *Agrobacterium*-mediated gene transfer (Zambryski *et al.*, 1983), particle bombardment, electroporation or microinjection (Cairns *et al.*, 1978; Fromm *et al.*, 1985; Klein *et al*, 1987), or other methods. In this study scFv genes were transformed into tobacco by the *Agrobacterium*-mediated method, the most commonly used procedure for this species.

Figure 4.2 : The PVX plant viral expression vector.

The schematic diagrams displayed the PVX expression vector and the scFv.PVX constructs assembled in this study.



4.2 Results and Discussion.

4.2.1 Expression of scFv Genes in *P. pastoris*.

Since the *Pichia pastoris* expression system has been reported to produce massive quantities of recombinant proteins, it was investigated in this study as an expression system for the refractory anti-ABA scFv genes of ABA26, MAC61, and MAC252. PV4.scFv, which was readily expressed *in E. coli*, was also included in this study.

The "multi-copy" *P. pastoris* expression system was used since it has been shown that this system generally increased expression levels of recombinant proteins. (Clare *et al.*, 1991; Scorer *et al.*, 1993).

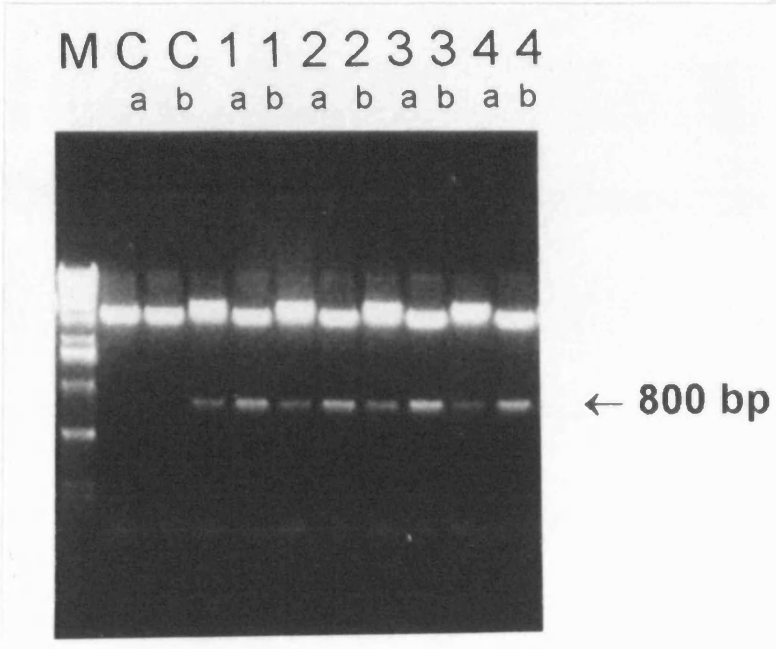
4.2.1.1 Construction and Expression of scFv.pPIC9K in *P. pastoris*.

Genes encoding PV4, MAC61, ABA26, and MAC252 scFvs were inserted into the secretory "multi-copy" *P. pastoris* expression vector, pPIC9K. All scFv.pPIC9K constructs were verified by restriction enzyme analysis before *P. pastoris* transformation (Fig. 4.3). Prior to the *P. pastoris* transformation, the plasmids were linearised by *Sst* I restriction enzyme to allow multiple-copy gene insertion at *5'AOX1*. The linearised plasmids were then transformed into *P. pastoris* GS115 (Mut^+) and KM71 (Mut^S) strains by electroporation. Transformed *P. pastoris* cells were first selected on minimal media (without histidine) for His^+ clones followed by G418 at different concentrations for the selection of transformants containing multiple gene copies.

More than $5x10^{10}$ colonies of GS115 transformants and $2x10^{10}$ colonies of KM71 transformants were obtained from a single electroporation using 5 μg of linearised plasmid DNA of each of the scFv.pPIC9K constructs. Between 5 to 20 colonies of each scFv.pPIC9K constructs were found to be resistant to the antibiotic G418 up to a concentration of 2 mg/ml. This resistance to high concentrations of G418 indicated that the transformation has produced *P. pastoris* transformants which contained multiple-copies of the inserted genes - so-called

Figure 4.3 : Verification of scFv.pPIC9K constructs by restriction enzyme analysis.

The assembled scFv.pPIC9K plasmids were verified by restriction enzymes (a) *Xho* I only, and (b) *EcoR* I and *Xba* I. The correct scFv.pPIC9K constructs would give a band corresponding to 800 bp when digested by *Xho* I, and 1.6 kb when digested by *EcoR* I and *Xba* I. Lanes 1, 2, 3, and 4 represent scFv.pPIC9K constructs of ABA26, MAC61, MAC252, and PV4 respectively. Lane M shows the DNA molecular weight marker.



"jackpot" clones (Clare *et al.*, 1991; Scorer *et al.*, 1993).

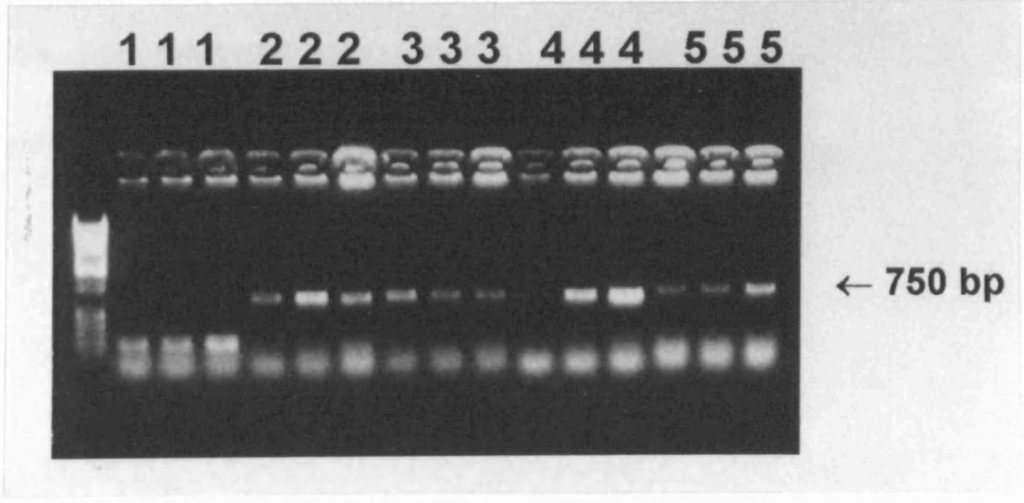
After selection, 8 to 10 transformed clones were chosen and analysed by PCR to confirm integration of the inserted gene into the *P. pastoris* genome (Fig. 4.4). For each of the scFv.pPIC9K constructs, two PCR-verified clones from the minimal media plates (with a single-copy of transgene) and two verified clones from the highest G418 selection plates (with multiple transgene copies) were selected for *P. pastoris* expression. Together with two control clones transformed with empty pPIC9K vector, this gave a total of 18 clones each of GS115 and KM71 transformants. Two expression procedures, the Mut^+ expression procedure for GS115 clones (His^+, Mut^+) and the Mut^S expression procedure for KM71 clones (His^+, Mut^S), were employed to express these selected clones.

Only those clones containing multiple copies of PV4.pPIC9K exhibited detectable expression of scFv protein when the media supernatant of methanol induced cultures were separated by SDS-PAGE followed by Western analysis using anti-HSV-tag antibody as detection probe (Fig. 4.5). No scFv protein was detected with other multi-copy scFv.pPIC9K clones or single-copy scFv.pPIC9K clones (including the single-copy PV4.pPIC9K clones) in either unconcentrated or tenfold concentrated media supernatants. With multi-copy PV4.pPIC9K clones, the scFv protein was first detected after 3 days of methanol induction and reached its maximum level at day 4. Similar results were obtained with both the GS115 (Mut^+) and KM71(Mut^S) transformants, although, the expression level of the PV4.scFv was lower in the KM71 strain compared to the GS115 strain. It is possible that this resulted from the low metabolic rate of KM71 *P. pastoris* in methanol media which in turn caused the scFv proteins to be produced at a much lower level.

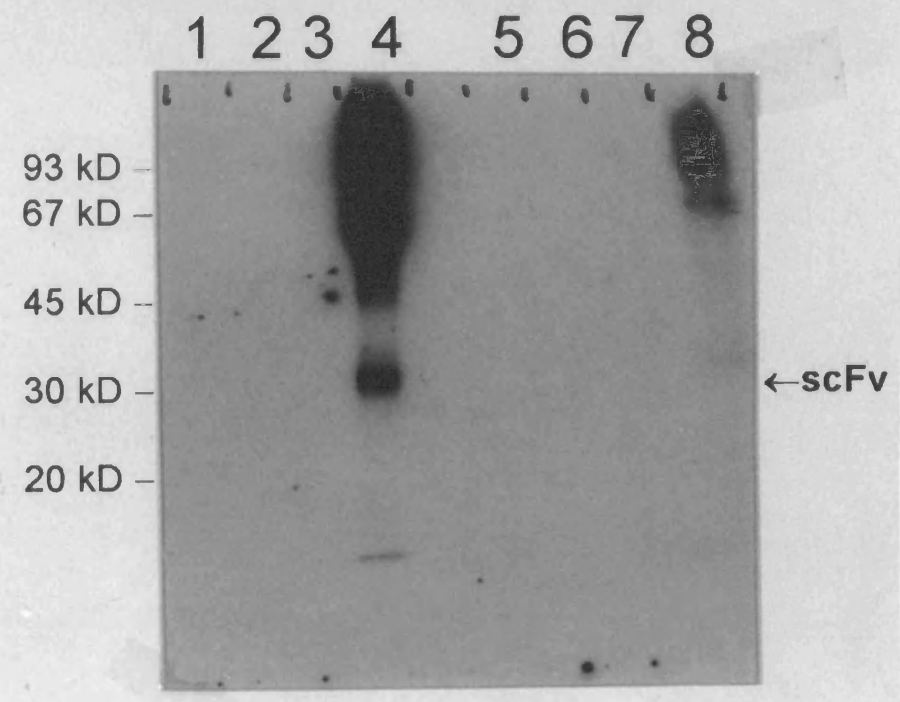
On Western analysis, although a stained scFv band appeared at the predicted molecular weight of 30 kD, strongly stained bands were also visible at molecular weights above 60 kD. These stained bands were unlikely to be the unspecific staining as they were not found in other samples including those from unexpressed transformants. The stained bands at higher molecular weight may resulted from association of scFv polypeptides or with other polypeptides.

Figure 4.4 : Verification by PCR of scFv.pPIC9K integration into the *P. pastoris* genome.

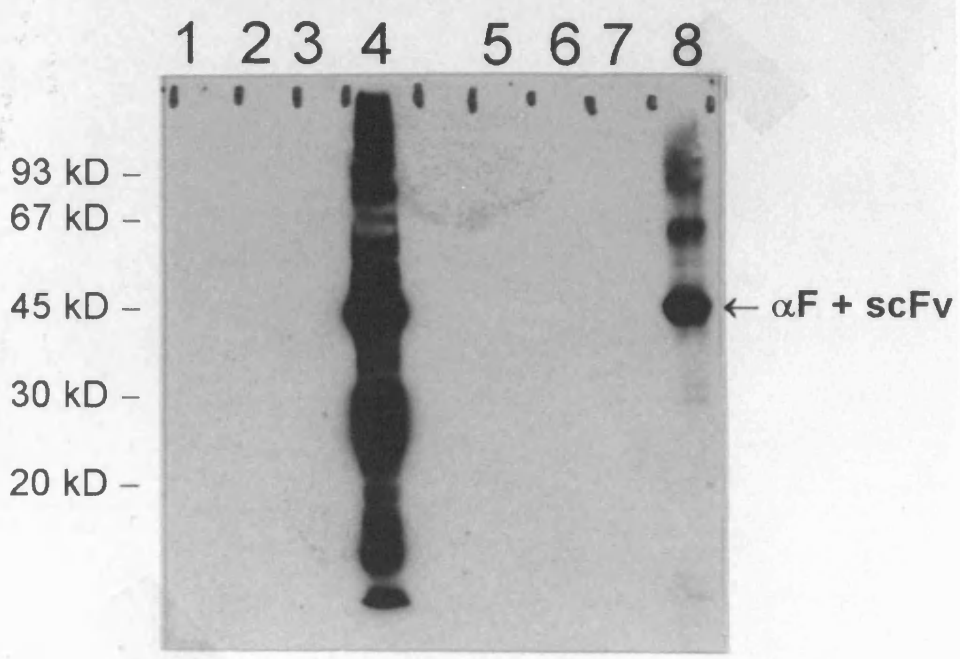
P. pastoris colonies were verified directly by a standard PCR protocol using primers targeted at 5'-P_{AOX1} promoter and 3'-end of scFv. *P. pastoris* transformants containing inserted scFv.pPIC9K would produce an amplified DNA fragment at 750 bp. Lanes 1 show negative control clones transformed with pPIC9K while lanes 2, 3, 4, and 5 represent three different selected clones of scFv.pPIC9K transformants of ABA26, MAC61, MAC252, and PV4 respectively. Lane M shows the DNA molecular weight marker.



(a) Culture supernatants.



(b) Cells pellets.



The cell pellets of the induced *P. pastoris* cultures were also analysed by SDS-PAGE and Western blotting. Only pellets of cells containing multiple copies of PV4.pPIC9K genes produced positive signals when probed with anti-HSV-tag antibody (Fig. 4.5). The major stained band, equivalent to a molecular weight of 40 kD, may represent scFv plus the 10 kD α-factor signal peptide. Similar results were obtained with both GS115 and KM71 transformants, although a stronger scFv signal was observed for the GS115 clones.

However, based on the chemiluminescent intensity detection signal, it was obvious that the level of the expressed PV4.scFv protein inside the cells was much higher than that in the culture medium. The Western analysis of cell pellets of PV4.pPIC9K gave a very strong signal which can be detected within 1 minute (Fig. 4.5b) while more than 30 minutes were required for detecting the signal (Fig. 4.5a) from samples using media supernatant. It appears that PV4.scFv may not have been effectively secreted and that as a consequence a substantial proportion remained in the cytoplasmic compartment. In contrast, when the 4715.scFv was expressed in *P. pastoris*, the expressed 4715.scFv protein was detectable only in the culture medium (4715.pPIC9 clones were obtained from Dr. Paul van der Logt, Unilever, UK., as positive controlled clones).

Based on the results obtained from the multi-copy *P. pastoris* expression system, it was clear that only the *P. pastoris* transformed with multiple-copies of PV4.pPIC9K produced scFv protein. The scFv genes of ABA26, MAC61, and MAC252 were not expressed at a detectable level in this system. This was unlikely to be caused by any experimental errors since all of the scFv.pPIC9K constructs were correctly assembled (verified by restriction enzymes) and the scFv genes were shown to be incorporated into the *P. pastoris* genome (verified by PCR and G418 resistance). Thus some characteristics of these anti-ABA scFv genes were incompatible with expression using the pPIC9K vector in *P. pastoris* and the possibility of expressing the genes with another *Pichia* vector, pPICZα was investigated .

4.2.1.2 Construction and Expression of scFv.pPICZα in *P. pastoris*.

PV4, MAC61, ABA26, and MAC252 scFv genes were constructed using the pPICZα expression vector and transformed into *P. pastoris* GS115 (Mut^+) and KM71 (Mut^S) strains for the so-called "easy-selection" expression system. The scFv.pPICZα construction was easier than the construction using the pPIC9K vector as a result of the comparatively small size of pPICZα vector (3.6 kD compared to 11 kD for pPIC9K). In addition, the inclusion in the vector of the gene for resistance to the antibiotic zeocin greatly simplified selection of both bacterial and *P. pastoris* transformants. However scFv genes could only be integrated into the *P. pastoris* genome by gene insertion and thus, the production of cells containing multiple copies of the transgene was not possible. The scFv.pPICZα constructs were analysed by restriction enzyme digestion (Fig. 4.6) and the integration of scFv.pPICZα into the *P. pastoris* genome was verified by PCR (Fig. 4.7).

A total of 13 verified clones (3 each from PV4, MAC61, ABA26 and MAC252 scFv genes plus 1 control clone transformed with empty pPICZα vector) each of GS115 and KM71 transformants were selected for expression. However, as assessed by Western blot analysis using anti-HSV-tag antibody of the methanol-induced culture supernatants and cell pellets on a daily basis up to seven days, none of the clones produced detectable scFv protein.

Again it was possible to be confident that the problem did not result from an error in the construction of the genes and also that the scFv genes had indeed been incorporated into the *P. pastoris* genome. It is concluded therefore that there are characteristics of these scFv genes which are incompatible with expression by the "easy-selection" *Pichia* expression system.

4.2.1.3 Attempts to Improve scFv Expression in *P. pastoris*.

The difficulties encountered in expressing scFv genes using both the pPIC9K and pPICZα expression vectors instigated a thorough analysis of the *Pichia* expression systems. Based on the published reports on successful scFv expression and discussions with fellow researchers working with *Pichia*, it appeared to be a general rule that if a

Figure 4.6 : Verification of scFv.pPICZα constructs by restriction enzyme analysis.

Assembled scFv.pPICZα plasmid were subjected to restriction enzymes digestion using (a) *Xho* I only, and (b) *EcoR* I and *Xho* I. Separated by agarose gel electrophoresis and visualised with ethidium bromide. Correct construct produce a band of 800 bp. Lanes 1, 2, 3, and 4 depict scFv.pPICZα plasmids of ABA26, MAC61, MAC252 and PV4 respectively. Lane C represents the negative control plasmid of pPICZα vector while lane M shows the DNA molecular weight marker.

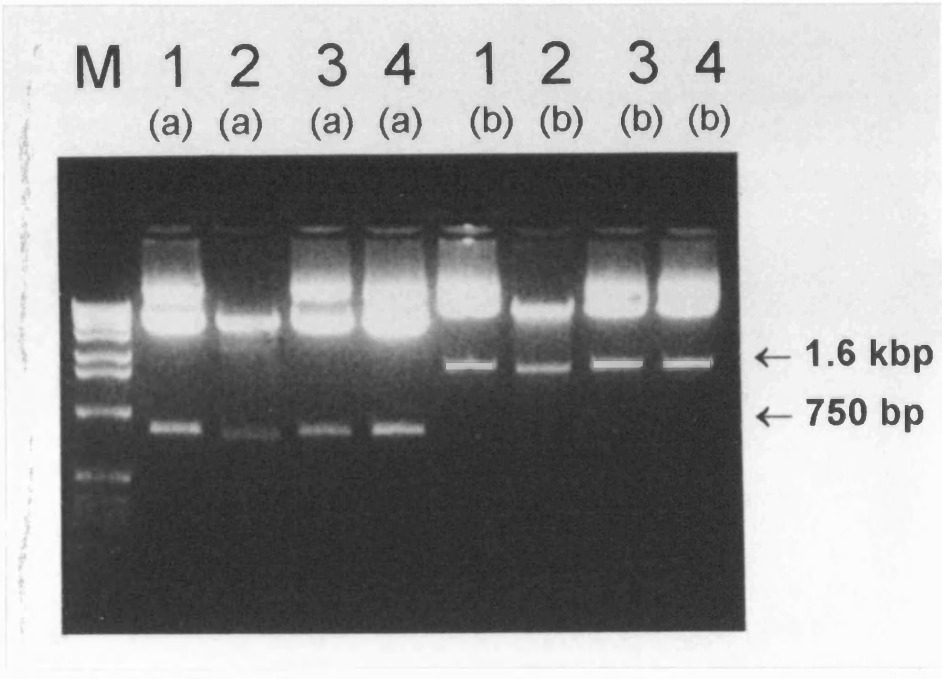
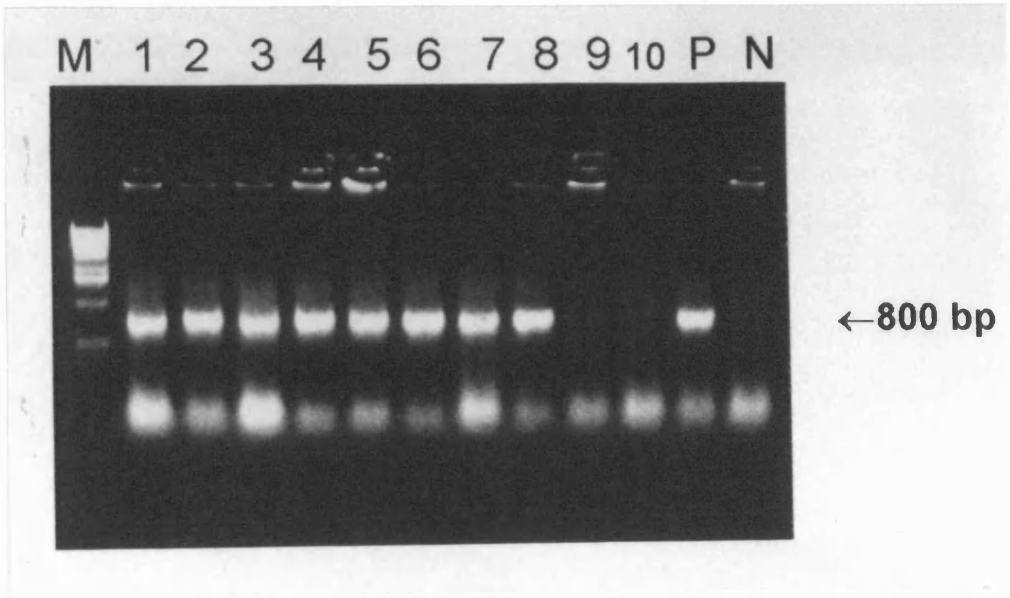


Figure 4.7 : Verification by PCR of scFv.pPICZα integration into *P. pastoris* genome.

The *P. pastoris* colonies were verified directly using the standard PCR protocol with primers targeted at *AOX1* gene promoter and the V_H domain of scFv. *P. pastoris* transformants containing inserted scFv.pPICZα should give an amplified DNA fragment at 800 bp. Lanes 1 & 2, 3 & 4, 5 & 6, and 7 & 8 show the selected clones of scFv.pPICZα transformants of ABA26, MAC61, MAC252, and PV4 respectively. Lanes 9 & 10 represent the negative *P. pastoris* control clones transformed with empty pPICZα vector. Lanes P and N were the positive and negative control samples using plasmid of MAC61.scFv and pPICZα respectively while lane M shows the DNA molecular weight marker.



recombinant protein could be readily expressed in *E. coli* it was very likely that it could also be readily expressed in the *Pichia* systems. Similarly if difficulties were encountered with the bacterial system it was likely that the protein would also be difficult to express in *Pichia*. Since in general, scFv genes do not express readily to high levels in *E. coli* difficulties in expressing them in the *Pichia* systems are perhaps not surprising.

Furthermore, since the *Pichia* expression system has come into use relatively recently there have been relatively few reports on the expression of antibodies (especially scFvs) in this system and consequently not very much information relevant to the expression of scFv genes is available. Although several scFv genes have been successfully expressed in *P. pastoris*, all of them used the Mut^S mutant of GS115 transformants (Luo *et al.*, 1995; Ridder *et al.*, 1995; Dr. Paul van der Logt, communication). At the time of writing this thesis there have been no reports on the expression of scFv genes in Mut^+ transformants.

Mut^S mutant (KM71 transformants) were produced in this study with the aim of expressing scFv genes. However no expression of the scFv genes was detected including PV4.pPICZα which was included as a gene likely to be expressed and which might act as a positive control in the event that the anti-ABA scFv genes were not expressed. Thus, use of the *Pichia* Mut^S mutant did not result in the expression of the scFv genes used in this study.

In all of the reports of scFv expression in *Pichia* , the "original" *P. pastoris* expression system was employed and the genes were transformed into *P. pastoris* GS115 strain. In this system, the GS115 Mut^S mutant would only be formed when the *AOX1* gene was replaced by scFv transgene. In other words, it was only through **gene replacement** and not by gene insertion that the Mut^S mutant was produced in strain GS115. In contrast, in the KM71 *P. pastoris* strain, a mutated Mut^S mutant is created when the *AOX1* gene is substituted with a **functional** *AGR4* gene, and the gene of interest (such as the scFv) is integrated into the KM71 genome by gene insertion.

As a result, the transgenic integrations by gene replacement or gene insertion exhibited the following significant differences: (P_{AOX} = *AOX1* promoter)

a) scFv gene insertion in GS115 creates the Mut^+ transformant which contains the gene encoding the 5'-PAOX-*AOX1* and the 5'-PAOX-scFv (multiple-copy insertion may occur).

b) scFv gene insertion in KM71 creates the Mut^S mutant which contains the gene encoding the 5'-PAOX-*AGR4* and the 5'-PAOX-scFv (multiple-copy insertions may occur).

c) scFv gene replacement in GS115 creates the Mut^S mutant which contains only the gene encoding the 5'-PAOX-scFv.

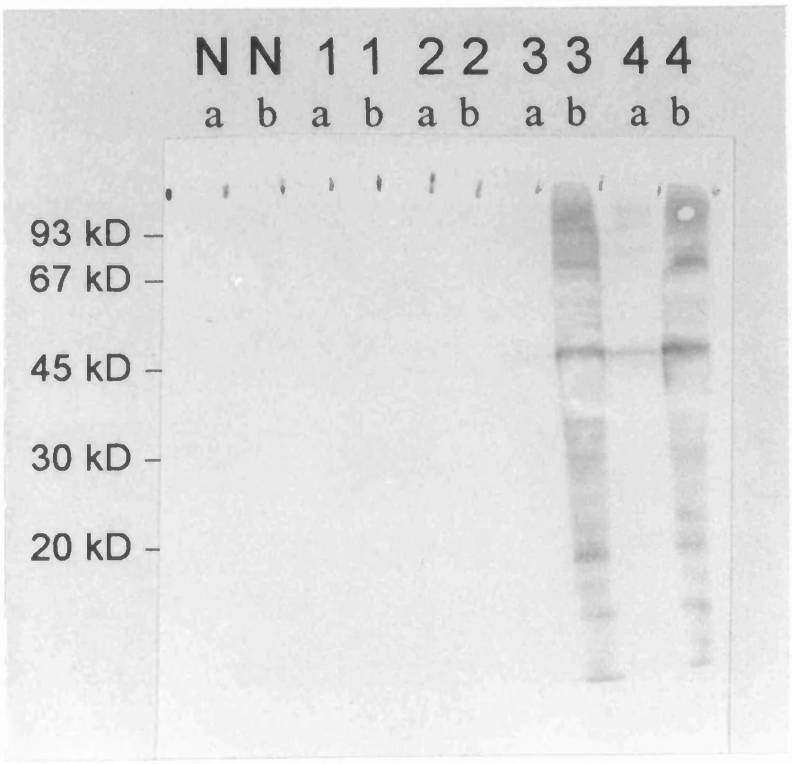
The GS115 Mut^S mutant produced by gene replacement contains a single gene under the control of 5'-PAOX promoter. When gene insertion is employed however, the *P. pastoris* transformants (either Mut^+ GS115 or Mut^S KM71) contain two different genes under the control of the same promoter. This may adversely affect expression of gene driven by the duplicated promoter, in this case expression of the scFv gene. To investigate this possibility, the scFv.pPIC9K constructs were integrated into *P. pastoris* GS115 genome by gene replacement (in this case, formation of multiple-copy gene was nearly impossible). The Mut^S *P. pastoris* transformants were then selected and expression of scFv genes was assessed.

ScFv proteins were detected in cultures supernatant of four-day methanol induced cultures of Mut^S mutants transformed with single-copies PV4.pPIC9K and MAC61.pPIC9K (Fig. 4.8). Thus in addition to PV4.scFv which was detectable with the "multi-copy" clones the single copy clones also allowed the production of detectable MAC61.scFv protein. However, no scFv proteins were detected with ABA26.pPIC9K and MAC252.pPIC9K in the Mut^S mutants. The results mirror the findings of the bacterial expression experiments and supports the general rule that performance in *E. coli* expression systems gives a good indication of likely performance in *Pichia* expression systems.

As a result of this study, it may be concluded that in the *P. pastoris* expression system the presence of another endogenous functional gene driven by the same promoter as the transgene suppressed the expression of the scFv genes.

Figure 4.8 : Expression of *P. pastoris* transformants by gene replacement.

The protein contents from (a) culture supernatants and (b) cell pellets of 3 day methanol induced *P. pastoris* transformants were analysed using 12.5% SDS-PAGE. The proteins blot was probed with anti-HSV-tag antibody using the colorimetric detection system. Lanes 1, 2, 3, and 4 show GS115 Mut^S *P. pastoris* transformants of ABA26, MAC252, MAC61, and PV4 scFv.pPIC9K respectively. Lane N was the negative control GS115 Mut^S clone transformed with pPIC9K vector.



4.2.2 Expression of scFv Genes in Plants.

Two plant expression systems were used with ABA26, MAC61, MAC252 and PV4 scFv genes, a transient expression system using a viral expression vector and the *Agrobacterium*-mediated stable transformation system.

4.2.2.1 Transient Expression System.

The PVX viral expression vector pP3C2S-402 was used to assemble the scFv.PVX constructs to be used for transient expression in tobacco plants. A total of 13 different constructs were assembled into the PVX expression vector, namely : a) ABA26, b) SS.ABA26, c) SS.ABA26.KDEL, d) TP.ABA26, e) MAC61, f) SS.MAC61, g) SS.MAC61.KDEL, h) TP.MAC61, i) MAC252, j) SS.MAC252, k) SS.MAC252.KDEL, 1) TP.MAC252, and m) SS.PV4.

Since PVX is an RNA virus, the constructed scFv.PVX derivatives were transcribed *in vitro* prior to tobacco (*N. clevelandii)* inoculation. The tobacco plants started to show viral infection symptoms 7 to 10 days after inoculation with transcribed RNA (Fig. 4.9). The viral infection symptoms were slightly delayed (about 1 to 2 days) on those plants inoculated with scFv.PVX constructs compared to the plants inoculated with transcribed RNA from the PVX vector only. This may be due to the inserted scFv gene influencing the biological activity of the virus.

However, when the protein extracts from the virus infected leafs were analysed by SDS-PAGE followed by Western analysis using anti-HSV-tag antibody as detection probe, only plants inoculated with SS.PV4.PVX RNA were found to produce scFv protein to a detectable level, and no scFv proteins were detected in other inoculated plants (Fig. 4.10). The molecular weight of the scFv produced using the SS.PV4.PVX construct was estimated to be 30 kD, as predicted from calculations based on its DNA sequence.

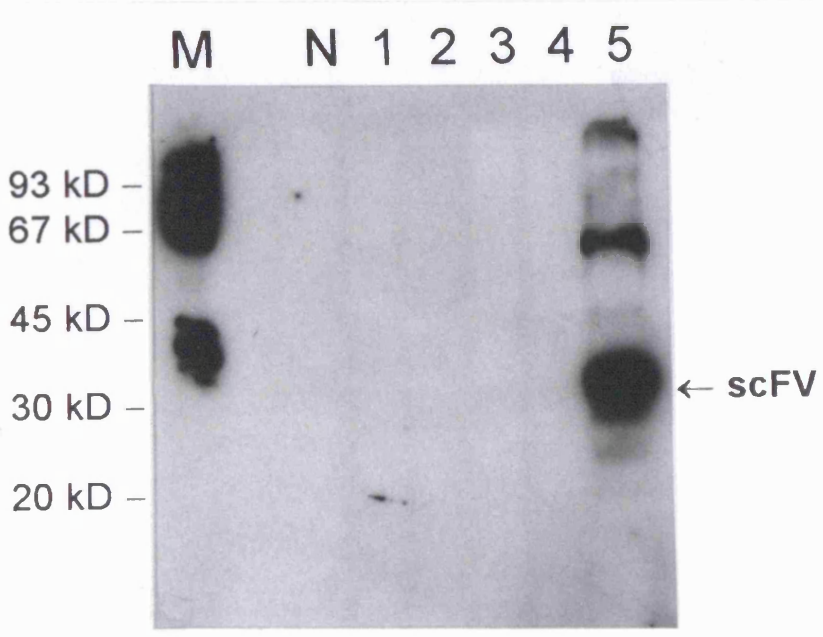
Thus it did not prove possible to express ABA26, MAC61, and MAC252 scFv proteins to a detectable level using the PVX expression system in plants. The factors responsible for this are unknown. However the monitoring of the production of the

Figure 4.9 : The scFv.PVX infection plants.

The diagram showed the *N. clevelandii* plants after 10 days inoculation with in vitro transcribed RNA of PVX plasmids and scFv.PVX constructs. Note that the viral infected symptoms were showed as the chlorotic mosaic and curly young leaves.

Figure 4.10 : Expression of scFv.PVX in plants.

The protein extracts from the infected leaves were analysed using 12.5% SDS-PAGE. The protein blot was then probed with anti-HSV-tag antibody using the chemiluminescent detection system. Lanes 1, 2 , 3, 4, and 5 show plant samples inoculated with transcribed RNA of PVX constructs of MAC61, ..MAC61, .MAC61.KDEL, TP.MAC61, and ..PV4 respectively. Lane N was the negative control sample inoculated with PVX RNA while lane M shows the positive marker using bacterial expressed HSV-tag fusion protein.



constructs and the expression experiments allows a number of possibilities to be eliminated.

1. Since all of the constructs were not assembled at the same time, it was very unlikely that the same experimental errors could occur twice or even three times. Furthermore, the MAC61.PVX constructs, which did not express, were assembled at the same time and with the same procedures as the PV4.PVX constructs, which did express.

2. The scFv.PVX constructs were always verified by PCR and restriction enzyme analysis. PCR screening was carried out using a 5'-end primer targeted at the vector and a 3'-end primer targeted at the scFv to verify correct orientation of the scFv in the scFv.PVX construct (Fig. 4.11). The plasmids from the PCR selected clones were then verified by restriction enzyme digestion. Correct scFv.PVX constructs would show a DNA fragment at 750 bp when the plasmids were digested by *Xho* I restriction enzyme (Fig. 4.12). The PCR and restriction enzyme verified clones were further authenticated by automatic DNA sequencing across the cloning site. This was required because the constructs were assembled by blunt end ligation and the restriction site was destroyed by the cloning procedure.

3. To ensure that full-length scFv.PVX transcripts were used for inoculation, a small portion of the *in vitro* transcribed products were analysed by agarose gel. Full length transcription would produce a band at about 3.3 kbp as predicted from its sequence (Fig. 4.13a). RNase liability was used to confirm that gel bands were composed of RNA (Fig. 4.13b).

4. At least three tobacco plants were used for the inoculation of each construct. Only full length transcribed RNA caused viral infection symptoms after inoculation. Since the virus coat protein gene is located at the end of the transcribed sequence, shorter RNAs from prematurely terminated *in vitro* transcriptions do not cause viral infection symptoms. Plants inoculated with truncated RNA were sometimes included as a negative control to ensure that viral infection symptoms did not result from contamination or cross infection from infected plants. In addition, the positive control plants, inoculated with RNA transcribed from the PVX vector

Figure 4.11 : Selection of scFv.PVX clones by PCR screening.

The scFv.PVX transformants were screened using one primer targeted at the PVX vector (upstream of scFv) and the other at the 3'-end of scFv. The transformants containing the correct scFv orientation would give amplified DNA fragments of 750 bp while the incorrect orientation would yield no amplified fragments.

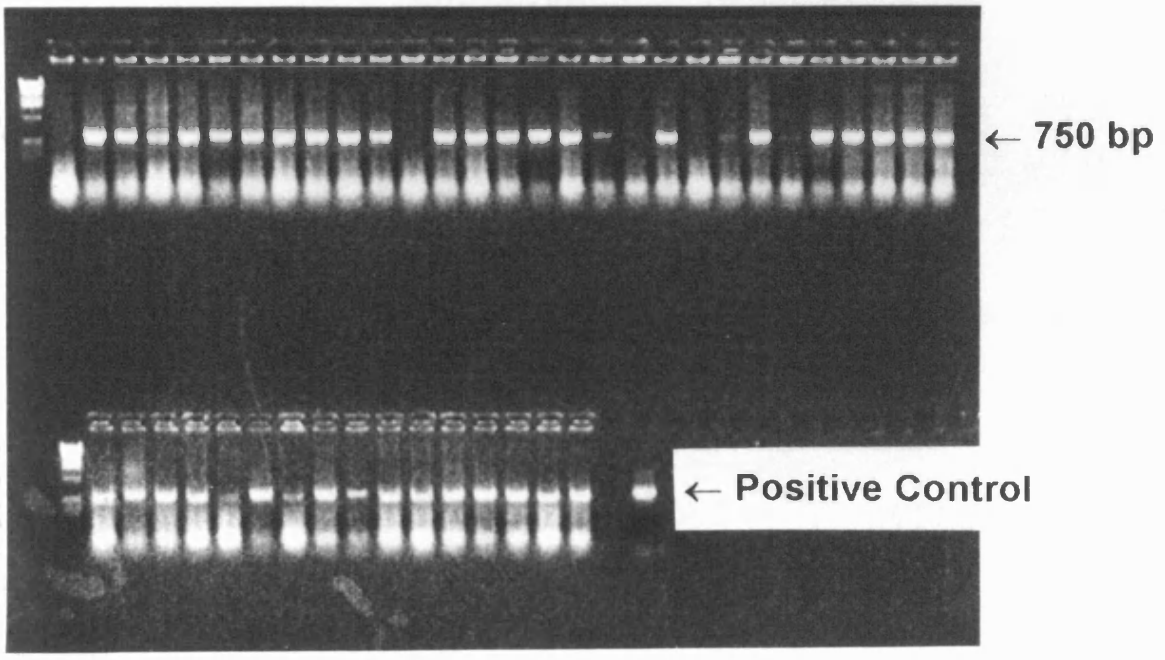
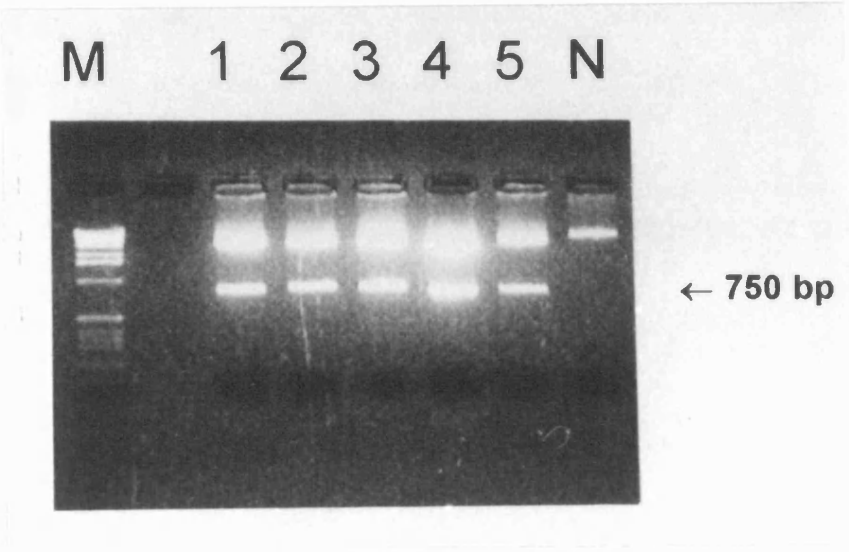
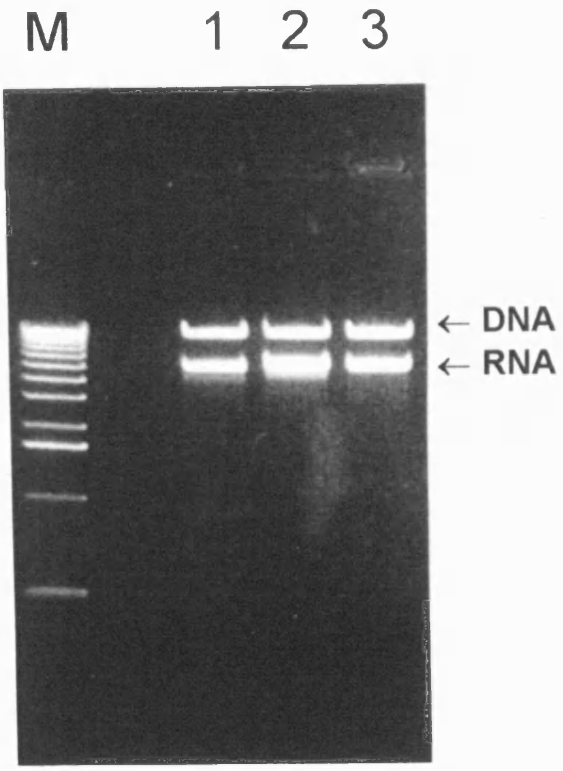


Figure 4.12 : Verification of scFv.PVX constructs by *Xho* I restriction enzyme analysis.

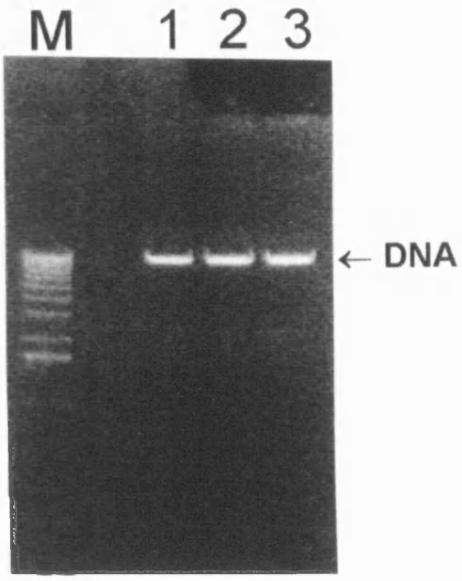
The assembled scFv.PVX plasmids were verified by the restriction enzyme *Xho* I. Correct scFv.PVX constructs would produce a band of 750 bp when digested by *Xho* I. Lanes 1, 2, 3, 4, and 5 show PVX constructs of MAC61, SS.MAC61, SS.MAC61.KDEL, TP.MAC61, and SS.PV4 respectively. Lane N represents the digested PVX plasmid while lane M shows the DNA molecular weight marker.



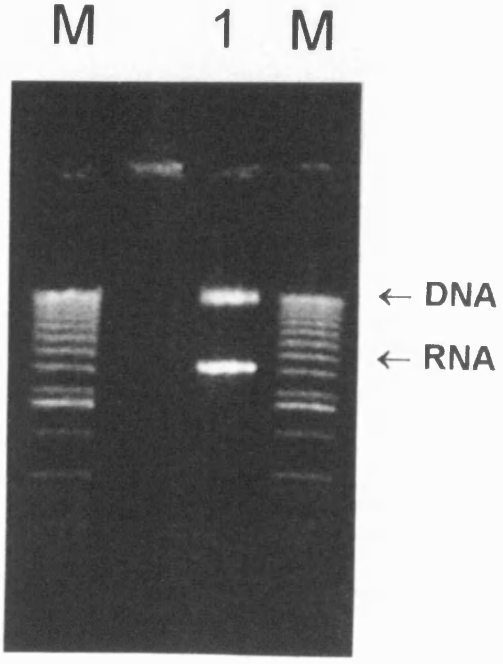
(a)



(b)



(c)



were separated from other inoculated plants.

5. Protein extracted from inoculated plants was also analysed by SDS-PAGE and probed by the anti-PVX antibody. The results showed that all plants showing viral infection symptoms contained similar protein patterns to samples obtained from plants inoculated with transcribed PVX RNA (Fig. 4.14). No virus proteins were detected in control plants which did not show any viral infection symptoms.
6. RNA extracts from leaves of inoculated plants were also analysed by RT-PCR. Leaves used for this purpose showed viral infection symptoms but not the leaves which had been inoculated with transcribed RNA. The results indicated that all plants inoculated with scFv.PVX RNA contained scFv RNA (Fig. 4.15). It was also demonstrated that, the scFv RNA produced spread into all of the virus infected parts of the plants. Nevertheless scFv protein was not detected.

At the time of writing this thesis there were no published reports on the expression of scFv genes in plants using a viral vector and it was therefore not possible to make comparisons with work utilising other scFv genes. The only information available comes from Dr. Z.C. Chen, a fellow researcher in the Botany Department Leicester University, who was expressing PV4.PVX constructs. In his later findings, PV4.PVX constructs which contained a signal peptide (SS.PV4.PVX constructs) or were targeted to the chloroplast stoma by means of a transit peptide (TP.PV4.PVX construct) produced detectable scFv protein. Cytoplasmically targeted scFv (PV4.PVX construct) can produce detectable scFv protein, but at very low level.

Figure 4.14 : Viral proteins analysis of scFv.PVX infected plants.

The protein extracts from the infected leaves were analysed using 12.5% SDS-PAGE. The protein blot was then probed with anti-PVX antibody using the chemiluminescent detection system. Lanes 1, 2 , 3, 4, and 5 depicted plant samples inoculated with transcribed RNA of PVX constructs of MAC61, SS.MAC61, SS MAC61.KDEL, TP.MAC61, and SS.PV4 respectively. Lane N was the negative control sample inoculated with PVX RNA.

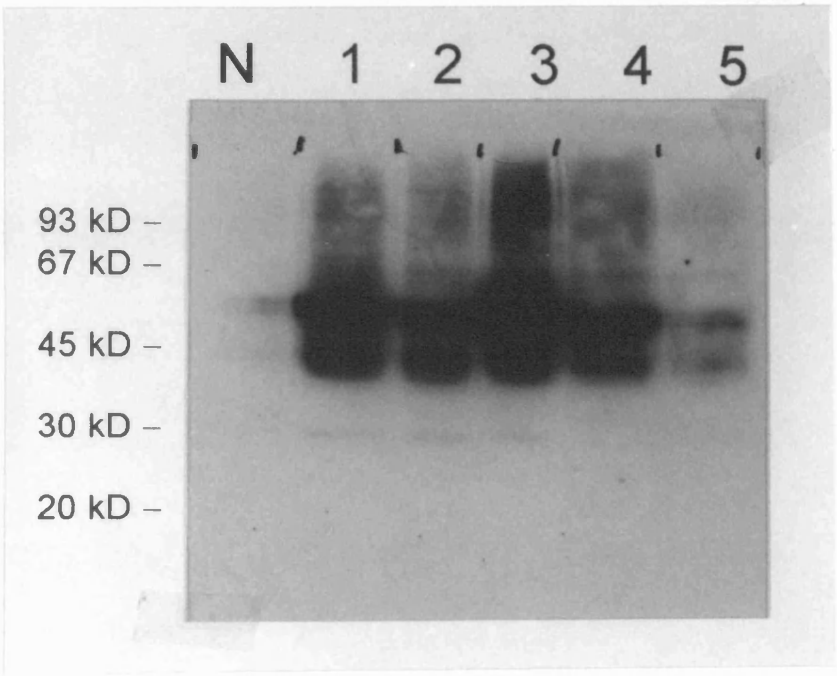
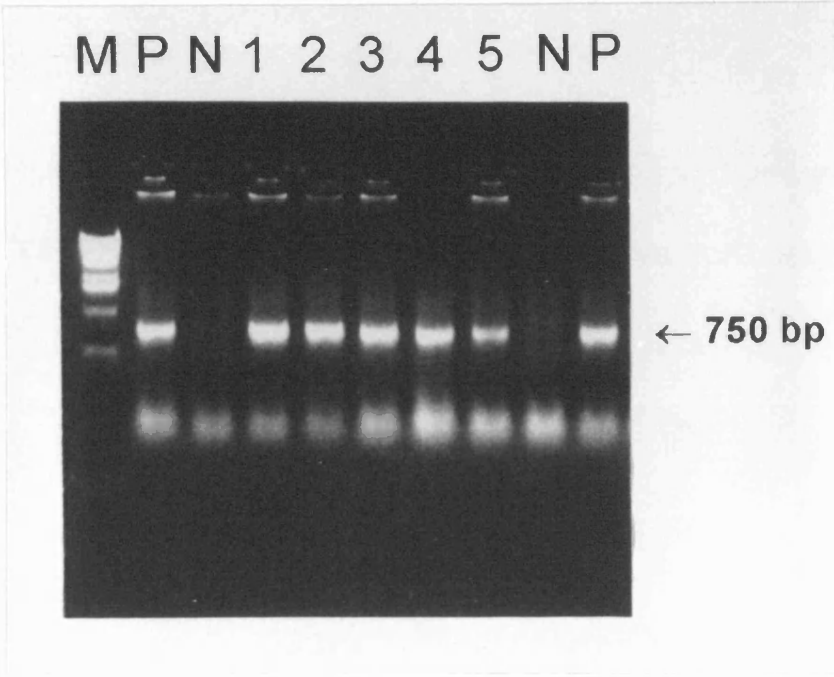


Figure 4.15 : RNA analysis of scFv.PVX infected plants by RT-PCR.

RNA extracts from infected leaves were analysed by RT-PCR. Lanes 1, 2 , 3, 4, and 5 represent plant samples inoculated with transcribed RNA of PVX constructs of MAC61, SS.MAC61, SS.MAC61.KDEL, TP.MAC61, and SS.PV4 respectively. Lane N shows the control sample inoculated with PVX RNA, lane P shows a positive PCR control using MAC61.PVX plasmid while lane M shows the DNA molecular weight marker.



4.2.2.2 Stable Plant Transformation.

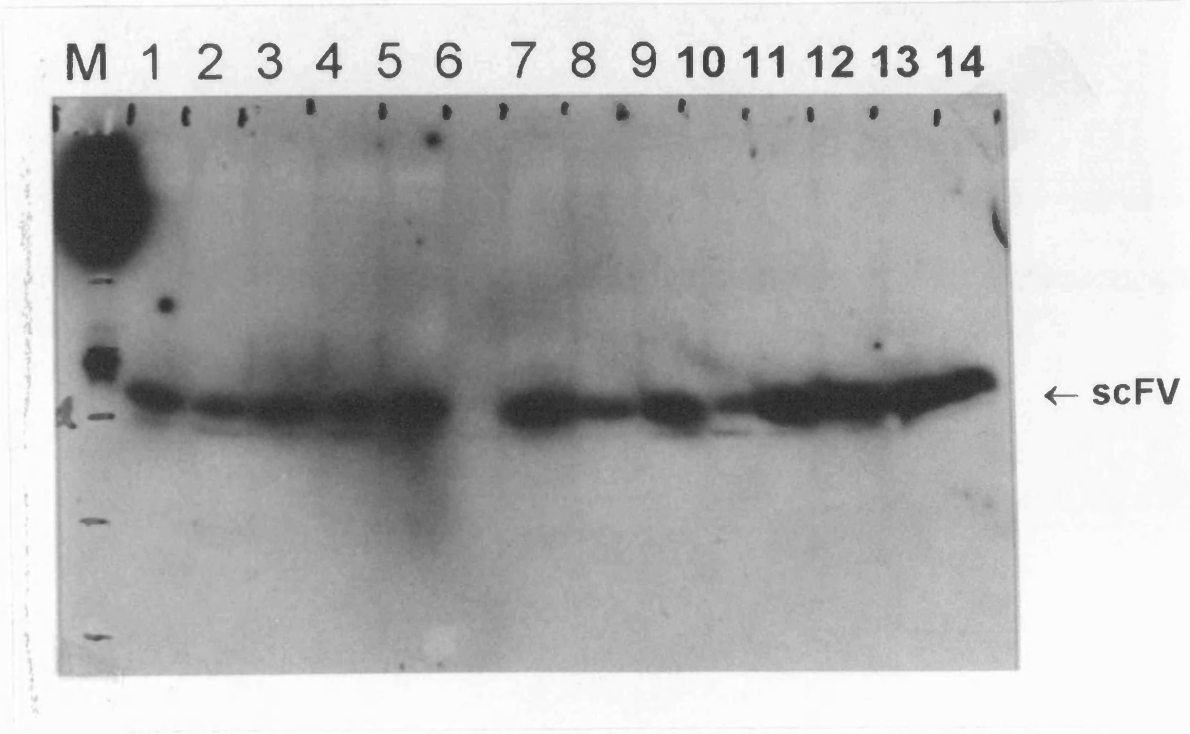
Agrobacterium-mediated stable plant transformation was used to transform various scFv constructs into the tobacco plant genome. A total of 13 different constructs were assembled into pBin 19 binary expression vectors. These were a) ABA26, b) SS.ABA26, c) SS.ABA26.KDEL, d) TP.ABA26, e) MAC61, f) SS.MAC61, g) SS.MAC61.KDEL, h) TP.MAC61, i) MAC252, j) SS.MAC252, k) SS.MAC252.KDEL, 1) TP.MAC252, and m) SS.PV4.

After plant transformation and kanamycin selection, the protein content from transformed tobacco plants which showed resistance to kanamycin were analysed by SDS-PAGE and probed with anti-HSV-tag antibody. More than 100 plants from each construct were screened. ScFv protein was only detected in protein extracts from some of the tobacco plants transformed with the SS.PV4 construct. (Fig. 4.16). The molecular weight of scFv proteins obtained from the SS.PV4 transgenic plants was about 30 kD, as predicted from its DNA sequence. Further studies on plant expressed PV4.scFvs were not carried out as these were undertaken by Dr. Chen, Botany Department, Leicester University. In his findings, transgenic plants transformed with SS.PV4 and TP.PV4 constructs were found to express scFv proteins to high levels whereas the cytoplasmically targeted scFv genes were only detected at very low levels. All of these plant-expressed scFv proteins were shown to possess antigen binding activities.

Thus the plant expression systems investigated here did not produce detectable amounts of ABA26, MAC61, and MAC252 scFv proteins. It has previously been found that production of scFv proteins in the plant cytoplasm often results in very low levels of accumulation. Owen *et al.* (1992) reported that screening of 119 transgenic plants transformed with AS32.scFv constructs revealed only one plant which expressed the scFv and furthermore the expression level was low (<0.1 % of the total soluble protein fraction). A similar low expression level of 0.1% of the total soluble protein was obtained by Tavladoraki *et al.* (1993) when the F8.scFv was expressed in tobacco plants. Fusion of scFv genes to a signal peptide which directed the scFv into the secretory pathway has been found to increase substantially the levels of scFv

Figure 4.16 : Analysis of plants expressed scFv proteins.

This diagram showed the protein analysis of several SS.PV4 transgenic plants (lanes 1 to 14). Protein extracts were analysed using 12.5% SDS-PAGE and the protein blot was probed with anti-HSV-tag antibody using the chemiluminescent detection system. Lane M shows the positive control marker using bacterial expressed HSV-tag fusion protein.



accumulating. Firek *et al* (1995) demonstrated that the secreted version of AS32.scFv accumulated to levels ten times those observed when the scFv was targeted to the cytosol. Application of this strategy to the anti-ABA scFvs, ABA26, MAC61, and MAC252 scFvs in this study did not however result in the accumulation of detectable levels of scFv protein.

Schouten *et al.* (1996) reported that cytoplasmic targeting of 21C5.scFv in tobacco did not result in the accumulation of detectable levels of scFv protein. The same scFv accumulated to maximum of 0.01% of the total soluble proteins when targeted to the plant apoplast, 0.2% of the total soluble protein with KDEL sequence, and up to 1% of the total soluble proteins when secreted and retained in the ER by means of both the secretory signal sequence and KDEL ER retention sequence. Somewhat similar observations were made by Artsaenko *et al.* (1995) for the expression of 15-I-C5.scFv. No scFv proteins were found in transgenic plants transformed with cytoplasmic or secreted versions of the scFv genes, although the scFv protein was detected with the KDEL ER retention version. The expression levels in these experiments were very varied. The majority of the transgenic plants expressed scFv proteins at relatively low levels ranging from 0.05 to 0.5% of the total soluble protein, although 5 out of 55 scFv expressed plants showed high expression levels of 2.1 to 4.8% of the total soluble protein.

Based on the information on expression of scFv genes in plants available at the time of writing it is clear that scFv expression in plants is not a straightforward process and accumulation of scFv proteins in the plant cytosol is particularly problematical. The formation of the two important intracellular disulphide bonds (one each in the V_H and V_L domains) is not favoured in the reducing environment of the cytosol. In general, either secretion (Firek *et al.*, 1993) or ER retention (Schouten *et al.*, 1996) of scFv genes tends to increase expression levels while scFv genes with both the secretion and ER retention sequences can further increase the expression levels (Artsaenko *et al.*, 1995; Schouten *et al.*, 1996). In an application of this approach in the present study, secretory and KDEL versions of ABA26, MAC61, and MAC252 scFv genes were produced and transformed into tobacco. None of the plants produced detectable levels of scFv protein.

Perhaps most importantly in the context of the present work, there appears to be significant differences in the levels of accumulation of different individual scFv proteins which presumably relates to the particular characteristics of the individual scFv genes (Firek *et al.*, 1993; Artsaenko *et al.*, 1995; Schouten *et al.*, 1996). It appears that ABA26, MAC61, and MAC252 scFv genes have individual characteristics which are incompatible with accumulation to detectable levels in plants. Thus these particular scFv genes have proved refractory or impossible to express in all of the expression systems utilised in the work described in this thesis.

The following points allow elimination of some of the possible artefactual explanations for the difficulties experienced in expressing ABA26, MAC61, and MAC252 scFv in stably transformed plants.

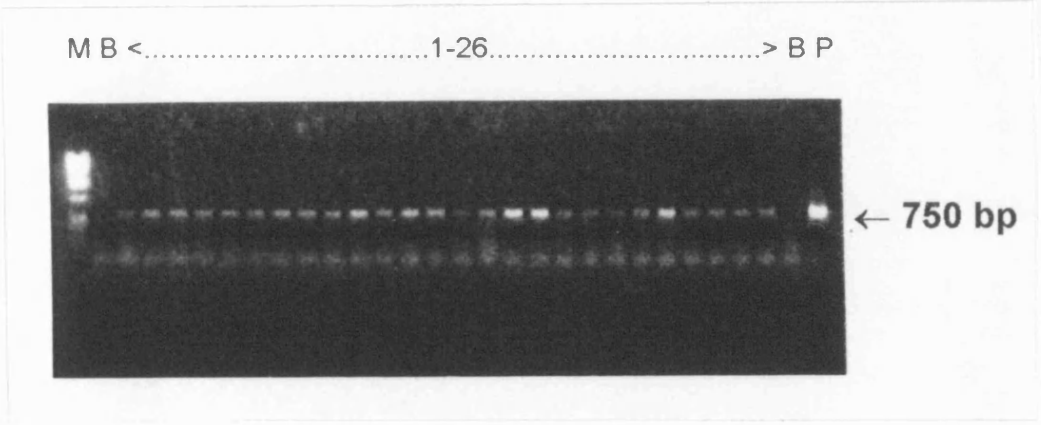
1. Since all the constructs were not assembled at the same time, it was very unlikely that the same experimental errors occurred in all cases. Furthermore the scFv.pBin constructs were not assembled and transformed into tobacco leaf disks at the same time. The MAC252.pBin constructs were the first set of constructs assembled. It was only after all of the MAC252.pBin transformants failed to produce any detectable scFv proteins that the MAC61.pBin constructs were assembled together with SS.PV4.pBin as a positive control. In this batch of constructs, however, only SS.PV4 transformants accumulated detectable levels of scFv protein., The ABA26.pBin constructs were last to be assembled. Accumulation of detectable SS.PV4 scFv protein provided verification of the technical procedures that were also used for ABA26, MAC61, and MAC252 scFvs.

2. The constructed scFv.pJIT and scFv.pBin derivatives were always verified by restriction enzymes and PCR analysis. Automatic DNA sequencing was also carried out across cloning sites if restriction sites were destroyed in the cloning process.

3. The presence of the scFv genes in the plant genome was verified by PCR. The result confirmed the presence of scFv genes in putative transformants while the control plants, healthy tobacco plants and the pBin 19 transformants, did not show any amplification of scFv genes (Fig. 4.17).

Figure 4.17 : DNA analysis of transgenic plants by PCR.

The transgenic plant DNA extracts were amplified by PCR using primers targeting the scFv gene. Transgenic plant with scFv transgene would shown a amplified DNA fragments of 750 bp. Lanes 1-26 show transgenic plants of MAC61, SS.MAC61, SS..MAC61.KDEL, and TP.MAC61. Lane B represents transgenic plant of pBin 19, lane P shows the positive control using MAC61.scFv plasmid while lane M shows the DNA molecular weight marker.



4. RNA extracted from the leaf of transgenic plants was also analysed by RT-PCR. The extracted RNA was first treated with DNase to remove any traces of DNA before RT-PCR was carried out. The results confirmed that putative scFv.pBin transformants contained scFv RNA while controls using RNA from pBin 19 transgenic plants gave negative results (Fig. 4.18). Thus despite the absence of detectable scFv protein the plants contained transcribed scFv messenger.

5. In addition to the above experiments, the scFv genes from the pBin 19 constructs were transferred to pRSET bacterial expression vector and expressed in *E. coli* BL21(DE3) *trxB*$^-$. ScFv protein could be detected following bacterial expression which verified that no mutations had been introduced into the scFv genes during the construction of the genes in pBin 19 for plant expression (Fig. 4.19).

Since the anti-ABA scFv protein of ABA26, MAC61, and MAC252 did not accumulate to detectable levels in plants, experiments planned to investigate the effects *in vivo* of the binding of ABA to the ScFv proteins, such as those subsequently published by Artsaenko *et al.* (1995) had to be abandoned. Artsaenko *et al.* (1995) reported that transgenic tobacco plants were capable of producing functional anti-ABA scFv proteins up to levels as high as 4% of total soluble protein. This finding clearly shows that it is not the ability to bind to ABA which prevents the anti-ABA scFv of ABA26, MAC61 and MAC252 being expressed in plants.

Figure 4.18 : RNA analysis of transgenic plants by RT-PCR.

The transgenic plant RNA extracts were analysed by RT-PCR. Transgenic plant which transcribed the scFv mRNA would shown a amplified DNA fragments of 750 bp. Lanes 1, 2 , 3, 4, and 5 depict RNA extracts from transgenic plants containing pBin 19, MAC61, SS.MAC61, SS.MAC61.KDEL, and TP.MAC61 respectively. Lane C represents RNA extract of SS.MAC61 transgenic plant without cDNA synthesis, lane P shows the positive control using SS.MAC61 plasmid DNA while lane M shows the DNA molecular weight marker.

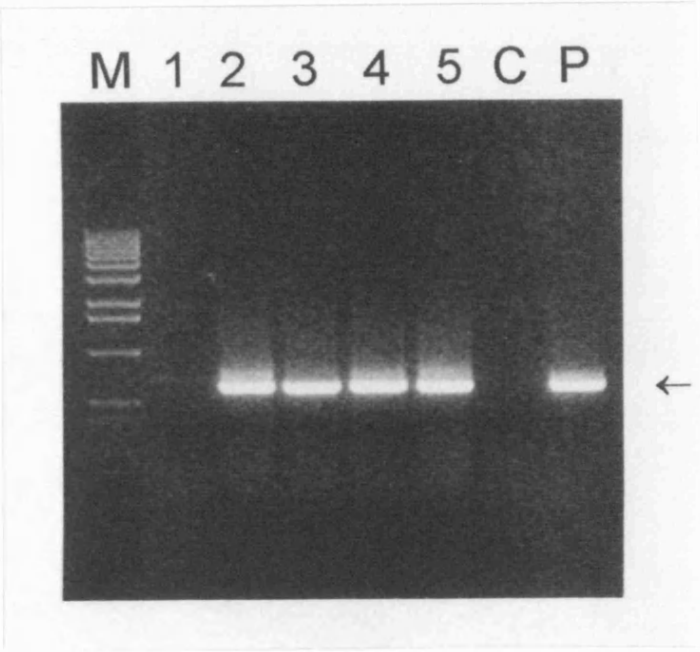
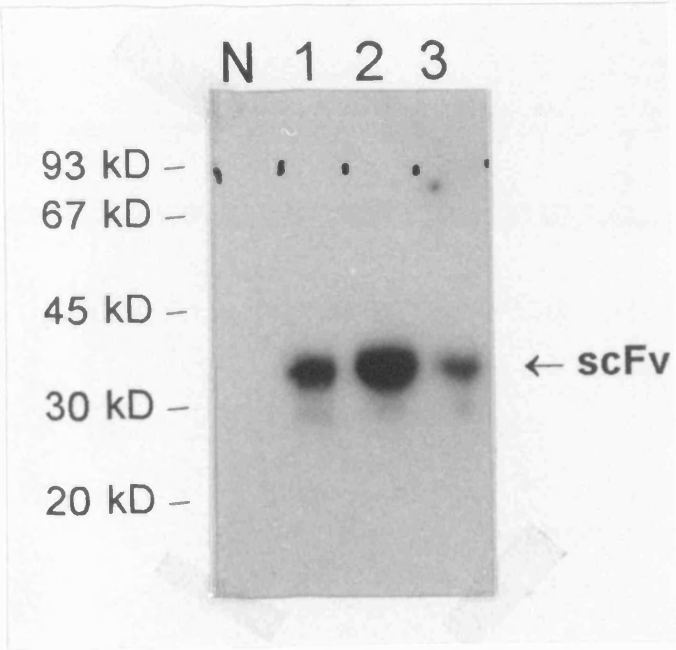


Figure 4.19 : Bacterial expression of scFv genes from pBin-19 constructs.

This panel shows bacterial expression of scFv genes transferred from scFv.pBin 19 constructs. The total bacterial proteins from induced bacterial cultures were analysed using 12.5% SDS-PAGE and a protein blot was probed with anti-HSV-tag antibody using the chemiluminescent detection system. Lanes 1, 2, and 3 depict constructs for ABA26, MAC61, and MAC252 scFvs respectively. Lane N shows the negative control.



4.3 Summary and Conclusions.

The expression of ABA26, MAC61, MAC252, and PV4 scFv genes was investigated in two types of eukaryotes, namely *Pichia pastoris* and tobacco plants .

Although no scFv protein was detected in the case of ABA26 and MAC252 scFvs with the *Pichia pastoris* expression system, PV4 and MAC61 scFv proteins were expressed and some of the product was secreted into the culture media. Secretion was accompanied by the anticipated excision of the signal sequence as evidenced by the observation of the predicted molecular weight of about 30 kD. In addition to the secreted scFv proteins of PV4 and MAC61 the *P. pastoris* cytosol contained a large portion of unsecreted scFv proteins to which the signal peptide remained attached. This inefficient secretion of PV4 and MAC61 scFv proteins contrasted with the behaviour of the positive control 4715.scFv which was almost all secreted into the culture medium.

Although it did prove possible to express PV4 and MAC61 scFv genes in the *Pichia* expression system, there were differences between them in this respect. When driven by the *AOX1* promoter, expression of the MAC61.scFv gene requires it to be integrated into the *P. pastoris* genome by replacement of the entire *AOX1* gene. In contrast expression of the PV4.scFv gene occurs with integration either by replacing the *AOX1* gene or by insertion of multiple copies of scFv genes. It is possible that depression of transgene expression, possibly by direct effects on the *AOX1* promoter or through the presence of an actively expressed gene driven by the same promoter as the transgene has a more significant effect on MAC61scFv gene expression than on the more readily expressible PV4.scFv gene.

It did not prove possible to express ABA26, and Mac 252 scFv proteins to detectable levels by any variant of the *P. pastoris* expression system, including gene replacement. It appears that ABA26 and MAC252 scFv genes have characteristics which adversely affect expression in both bacterial and *P. pastoris* systems.

Expression of ABA26, MAC61, and MAC252 in plants, either by transient expression using the PVX vector or by stable *Agrobacterium*-mediated transformation,

did not result in the accumulation of scFv protein to a detectable level. Secretion to the apoplast and ER retention, both of which normally increase accumulation of scFv protein, did not result in the accumulation of detectable levels of these particular scFv proteins. The secreted form of PV4.scFv was however readily detectable in tobacco plants transformed and analysed by the same procedures as employed for the anti-ABA scFvs.

The inability of the plant expression systems to express ABA26, MAC61, and MAC252 scFvs to a detectable level appears to be due to the nature of these particular scFv genes. The actual characteristics which have this adverse affect on expressibility are unknown although it is clear that they are deleterious to expression in bacterial, yeast and plant-based expression systems. It may be that the general rule, that an scFv which is difficult to produce in *E. coli* is also likely to be difficult to produce in *Pichia*, may be extended to predict that the scFv will also be difficult to produce in tobacco.

CHAPTER 5

Antibody Variants.

5.1 Introduction.

As well as allowing the production of antibody fragments, antibody engineering technology provides ways of altering antibody properties and combining them with other effector activities to produce immunoreagents tailored for specific purposes. Such engineered immunoreagents include: catalytic antibodies; chimeric antibodies; antibody fusion proteins such as therapeutic immunotoxins and antibody/detector function fusions; diabodies and other multimerised fragments.

In the work reported here a study was made of two such variants of particular, though not exclusive relevance to plant science. Firstly, since the monovalent nature of the scFv results in a lower avidity than that of the bivalent parental antibody and also precludes the possibility of agglutination of antigen which may be of value in anti-pathogen strategies. Consequently a bivalent diabody version of an scFv of potential use in anti-plant virus strategies was constructed and studied. Secondly an scFv fusion to the extremely versatile green fluorescent protein was constructed and studied. Possible applications of such a fusion include its use as a single step immunodiagnostic and an *in-vivo* marker for immunoreagent presence, quantification and localisation.

5.1.1 Diabodies - Dimeric scFv Proteins.

Two forms of diabody or dimeric scFv protein have been described namely, bivalent and bispecific antibodies. Bivalent antibodies contain two identical antigen binding sites whereas bispecific diabodies contain two different antigen binding sites. Several different strategies have been used to construct dimeric antibodies, including chemical cross-linking (Shalaby *et al.*, 1992) or by introducing a carboxyl terminal peptide at scFv to promote dimerisation (Pack and Plückthun, 1992; Shu *et al.*, 1993).

Holliger *et al.* (1993) described an elegant method which utilises a short linking peptide between the V_H and V_L domains which prevents association of V_H and V_L domains on the same polypeptide but allows associations between variable domains on separate polypeptides, thus forming dimeric scFv proteins with two antigen-binding sites. In this packing, the antigen binding sites of the dimeric scFv proteins (diabodies) are pointed outward in opposite directions and the two sets of paired domains are related by a dyad axis (Holliger *et al.*, 1993).

In this study, a gene encoding a diabody derived from 'normal' version of PV5.scFv (with 15 amino acid linker between V_H and V_L domains) which recognises an epitope of potato virus Y was constructed, expressed and characterised using *E. coli*. The PV5 diabody was constructed using a 5 residue 'short' linker (GGGGS) between the V_L and V_H domains of PV5.scFv. This short linker prevents association of V_L and V_H domains on the same polypeptide but imposes major constraints on the ways by which two separate scFv chains can associate and tends to cause the formation of dimers rather than larger multimers. Furthermore, X-ray crystallographic structure studies and computer graphics modelling showed that the length of an scFv linker can be as short as 7Å in dimeric scFv polypeptides (Fischmann *et al.*, 1991; Holliger *et al.*, 1993). This length requirement can be satisfied by a 5-residue amino acid linker used in this work.

5.1.2 Green Fluorescent Protein - ScFv Fusion Proteins.

Many of the applications of antibodies require coupling of the antibody to other functional partners to create multifunctional immunoreagents. In conventional antibody technology, antibodies have been coupled chemically to a very wide range of substances including enzymes, toxins, radioactive compounds, fluorescent organic compounds, and even gold particles. The possibilities for the generation of multifunctional immunoreagents have been greatly extended by the application of recombinant antibody technology. Antibodies and fragments, mainly scFvs, have been fused at the level of the gene with a wide range of fusion partners such as enzymes (Kolmar *et al.*, 1992; Gandecha *et al.*, 1994), toxin proteins (Batra *et al.*, 1990;

Brinkmann *et al.*, 1991; Chaudhary *et al.*, 1993), protein A (Tai *et al.*, 1990; Gandecha *et al.*, 1992), maltose-binding proteins (Bregegere *et al.*, 1994; Li, 1996), and metal-binding proteins (MacKenzie *et al.*, 1994b).

Investigations are reported here of the use of jellyfish green fluorescent protein (GFP) as scFv fusion partner. GFP is a bioluminescent protein produced by the jellyfish *Aequorea victoria* which emits a bright green fluorescence when illuminated by blue or UV light (Inouye and Tsuji, 1994). GFP offers the advantage over other bioluminescent systems that it does not require any exogenously added substrate or co-factors or additional gene products from *A. victoria*. The non-invasive nature of the treatment required to elicit fluorescence, namely simple illumination, allows numerous *in-vivo* applications of this remarkable protein (Inouye and Tsuji, 1994).

The jellyfish GFP has been widely used as a fusion partner to many proteins and has been successfully expressed in bacteria (Chalfie *et al.*, 1994; Inouye and Tsuji, 1994), yeast (Flach *et al.*, 1994; Nabeshima *et al.*, 1995), insects (Wang and Hazelrigg, 1994; Chalfie *et al.*, 1994; Marshall *et al.*, 1995), mammalian cells (Kain *et al.*, 1995; Olson *et al.*, 1995), vertebrates (Amsterdam *et al.*, 1995; Wu *et al.*1995) and plants (Haselhoff and Amos, 1995; Baulcombe *et al.*, 1995). GFP can be fused to either the amino terminal or the carboxyl terminal of the fusion partner and in many cases, maintains fluorescent properties similar to the native GFP, as well as retaining the normal biological activity of the fusion partner.

GFP was selected as fusion partner in this study because of the many potential applications of fluorescent immunoreagents, particularly as immunodiagnostics. In addition, the jellyfish green fluorescent protein could be used as a "fusion tag" or *in vivo* marker to detect and localise GFP.scFv fusion proteins.

The genetic variant of GFP, m-gfp5 which was created by mutating several amino acid residues of wild type GFP, was used for this study. Variant m-gfp5 has proven to enhance the fluorescence intensity 5 to 10 fold compared with wtGFP. It is thermo-stable and it can be expressed in the bacterial cytoplasm at temperatures up to 37°C whereas wtGFP would only fluorescence when expressed at temperatures lower than 25°C (Dr. Jim Haselhof, MRC, Cambridge., Communication).

5.2 Results and Discussion.

5.2.1 Diabodies - Dimeric scFvs.

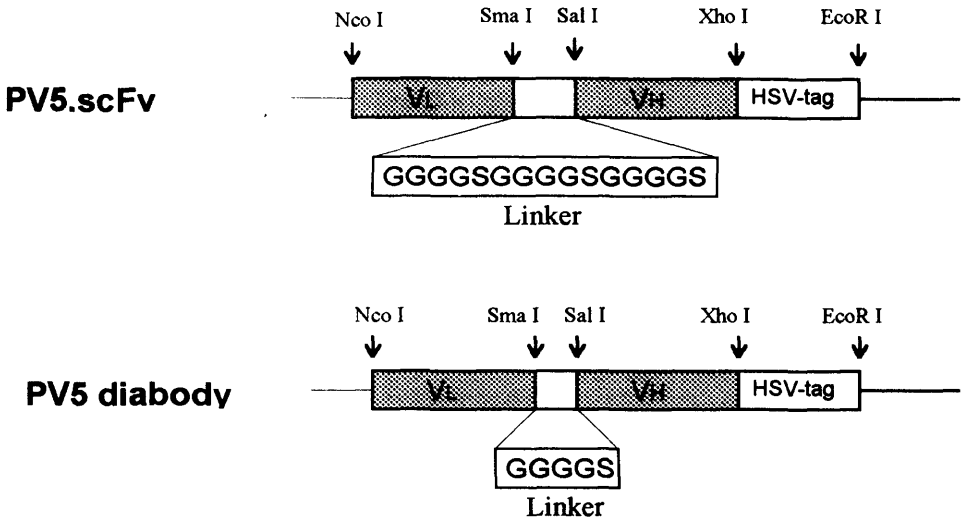
The diabody version of PV5.scFv (PV5 diabody) was assembled in the pABE expression vector. Fig. 5.1 shows the schematic construct and the amino acids sequence of the diabody version of PV5.scFv. This expression vector incorporates the signal peptide coding region of the *malE* gene to direct scFv proteins to the bacterial periplasm. It is well established that export to the *E. coli* periplasm allows formation of disulphide bonds and soluble scFv protein (Plückthun, 1992). This feature was important in this study to facilitate correct folding of both variable domains and to allow dimerisation of scFv proteins in formation of diabodies.

Bacterial expression of PV5 diabody was undertaken using *E. coli* W3110wt after the construct had been verified by restriction enzyme digestion and PCR analysis. Bacterial expression experiments using the normal version of PV5.scFv in the same expression vector and bacterial host were also carried out.

Total protein extracts obtained from induced bacteria transformed with PV5 diabody were analysed using SDS-PAGE followed by Western blotting and probed with anti-HSV-tag antibody. Fig. 5.2a shows that the denatured monomers of PV5 diabody protein possessed a molecular weight of about 29 kD, corresponding to that calculated from the DNA sequence. The expression level of PV5 diabody protein was similar to that for the normal version of PV5.scFv protein. Based on Western analysis, about 50% of the total expressed PV5 diabody protein was found in the form of soluble fraction (Fig. 5.2b), similar proportion in expressing of PV5.scFv protein.

The molecular weight of the putative PV5 diabody was estimated using Sephadex G-75 gel filtration chromatography. Eluted fractions were immobilised on microtitre plates and probed with anti-HSV-tag antibody revealing that the majority of the detected signal was in a fraction corresponding to a molecular weight of about 60 kD (Fig. 5.3). This molecular weight is equivalent to the sum of the molecular weights of two PV5.scFv polypeptides and indicates that the PV5 diabody polypeptides were indeed in the form of dimeric proteins. From this result, it is clear that diabodies can be constructed using the short (GGGGS) linker between the V_H and

Figure 5.1 : Schematic diagram and amino acid sequence of PV5 diabody.



Amino acids sequence of PV5 diabody :

Light chain sequence :

```
  1-50 MADVQIIQTTASLSASVGETVTITCRASEHIYSYLAWYQQKQGKSPQLLV
 51-100 YSAKTLAEGVPSRFSGSGSGTQFSLKINSLQPEDFGSYYCQHHYDTPRTF
101-124 GGGTKLEIRRADAAPTVSIFPPSS
```

Heavy chain sequence :

```
  1-50 QVQQPGAELVRSGASVKMSCKASGYTFTSYNMHWVKQTPGQGLEWIGYIY
 51-100 PGNGGTIYNQKFKGKATLTADTSSSTANMQISSLTSEDSAVYFCARGDYR
101-117 NDPFDFWGQGTTLTVSS
```

HSV-tag dectection signal sequence:

```
LEGAGSGOPELAPEDPED*
```

Figure 5.2 : Bacterial expression of diabody version of PV5.scFv.

The total protein extracts of induced bacterial cultures were analysed using 12.5% SDS-PAGE and the protein blot was probed with anti-HSV-tag antibody using colorimetric detection system. Lane 1 represents the original version of PV5.scFv, lanes 2 & 3 were two selected diabodies versions of PV5.scFv. In this panel, (a) represents total bacterial protein extracts, (b) represents soluble protein fractions, while (c) represents insoluble protein fractions.

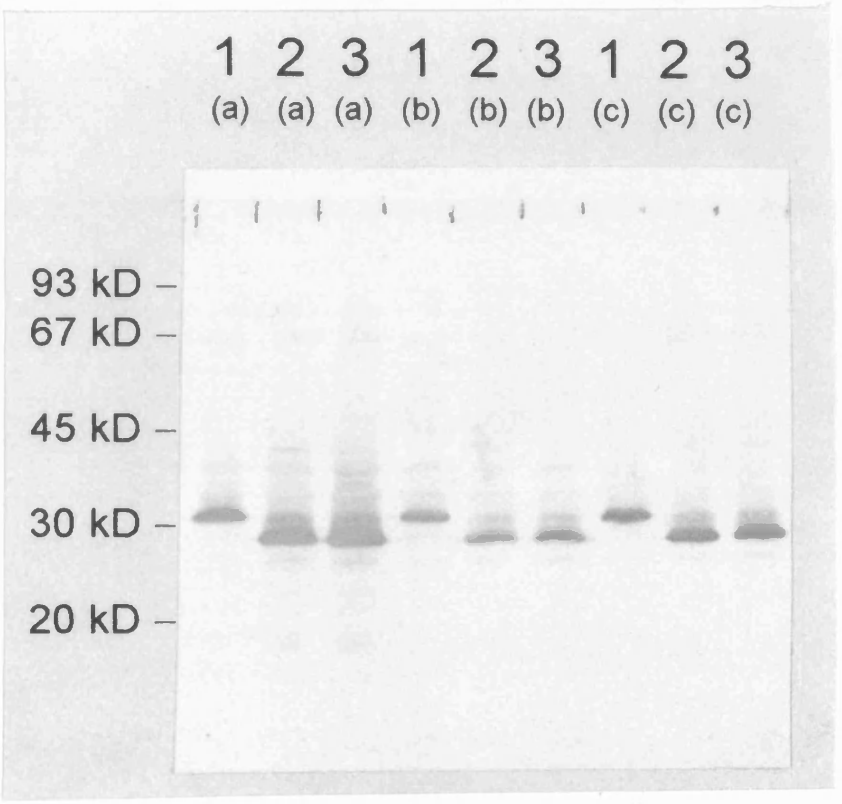
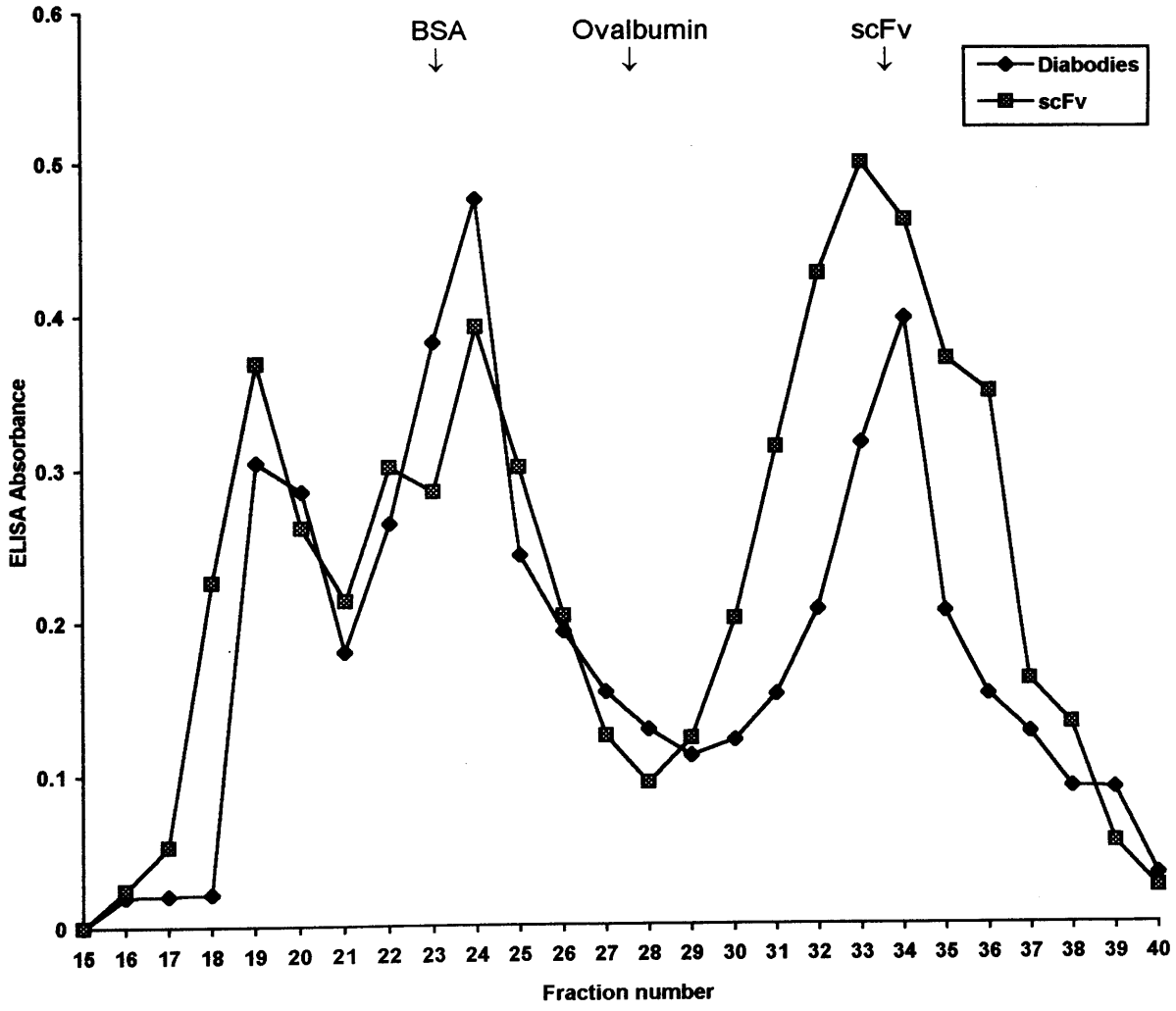


Figure 5.3 : Fractionation of bacterially expressed PV5.scFv proteins by Sephadex G-75.

The eluted fraction was then collected in a volume of 0.375 ml per tube and the scFv protein was detected by immobilised the fraction content on a microtitre plate and probed using anti-HSV-tag antibody. The molecular weight of native scFv protein was determined by comparing the migration values (Rf) of standard proteins.

Blue dextran was used to determine the column void volume while molecular weight calibration was carried out using BSA (67 kD), chicken egg ovalbumin (43 kD), and denatured bacterial expressed scFv (30 kD).



V_L domains of this scFv protein. This finding is consistent with the work of Holliger *et al.* (1993) in which most of the expressed scFv proteins containing the 5 amino acid linker appeared to be in the form of dimeric complexes. As well as the dimeric PV5 diabody fraction, two other fractions, one eluting in the void volume fraction (high molecular weight) and the other corresponding to a molecular weight of about 29 kD were detected. This may indicate that some of the expressed PV5 diabody polypeptides associated to produce complexes of high molecular weight while some remained as monomeric polypeptide.

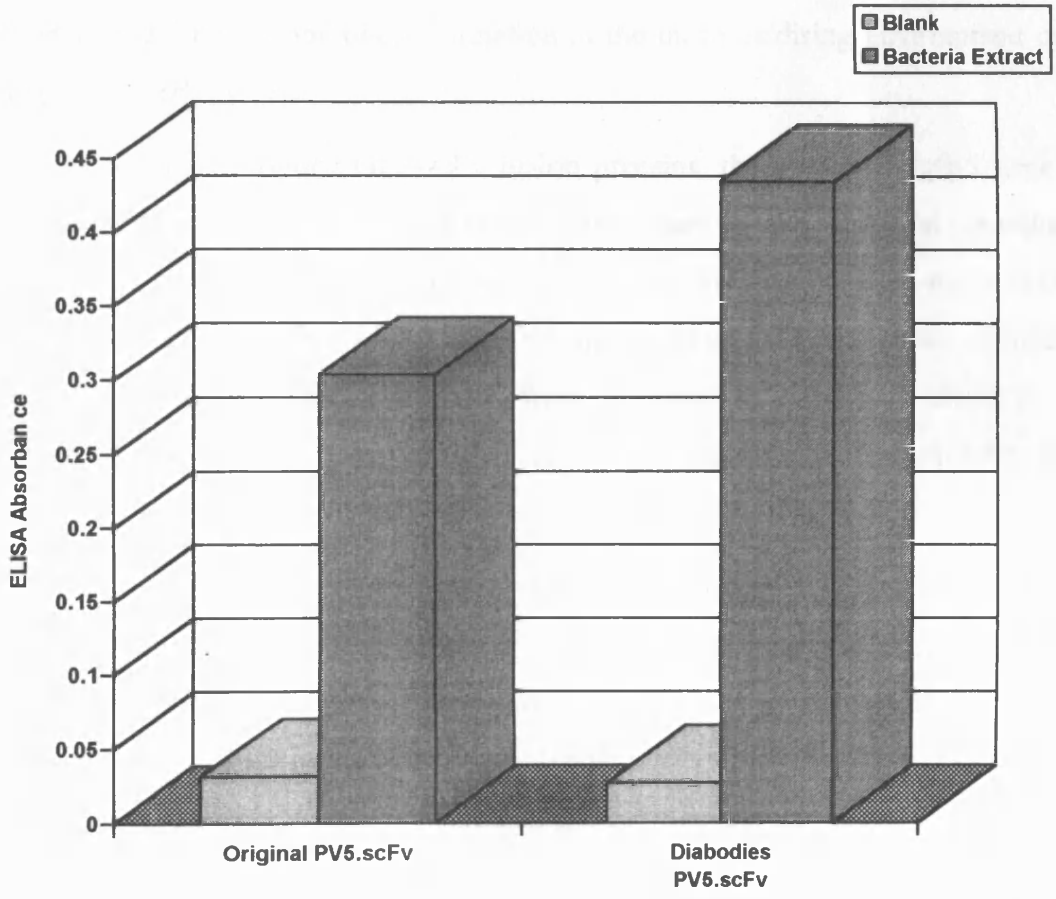
A similar result was obtained when the normal version of PV5.scFv protein was fractionated on the same Sephadex G-75 column (Fig. 5.3). Based on this result, some of the PV5.scFv associated in the form of dimeric complexes. It has been commonly found that, the "conventional" scFv proteins can associate spontaneously to form dimers or higher multimers (Holliger *et al.*, 1993). However, only about one third (as estimated from the sums of the detection signals in Fig. 5.3) of the normal version of PV5.scFv proteins were in dimeric complexes while about half of the diabody version of PV5.scFv was dimeric.

The binding activity of the PV5 diabodies was determined by ELISA using total protein extract of isolated PVY virus as immobilised antigen. As Fig. 5.4 shows the PV5 diabodies were able to bind to PVY. In addition, using equal amounts of antibody protein (based on Western blot analysis), the PV5 diabodies were found to produce about 30% higher ELISA signal than the PV5.scFv protein. Although inconclusive, this is consistent with an increase in avidity predicted for the PV5 diabodies as a consequence of their bivalence.

Further studies on expression of the PV5 diabody in plants and its effects on virus infection in plants are currently being undertaken by another worker.

Figure 5.4 : Antigen binding activity of PV5 diabodies.

The ELISA assay was carried out using 100 µl of 100µg/ml PVV and PVY extract as coated antigen. 100 µl of 2x diluted total protein extracts of induced bacterial cultures were used as test samples. The assays were then probed with the anti-HSV-tag antibody and the horse radish peroxidase-labelled anti-mouse IgG antibody was used as the second antibody.



5.2.2 Green Fluorescent Protein - ScFv Fusion Proteins.

The m-gfp5 gene was first extracted from the plasmid in which it was supplied by PCR amplification and was subsequently cloned into the pRSET cytoplasmic expression vector. After the construct had been verified by restriction enzyme digestion and PCR analysis, the cloned m-gfp5 plasmids were transformed into *E. coli* JM109(DE3) and BL21(DE3) *trxB*$^-$. Transformed colonies of both of these strains were found to be fluorescent when illuminated with UV light even without IPTG induction (Fig. 5.5). The intensity of fluorescence of BL21(DE3) *trxB*$^-$ transformants was markedly stronger than that of the JM109(DE3) transformants. This is consistent with enhanced GFP chromophore formation in the more oxidising environment of the BL21(DE3) *trxB*$^-$ cytosol.

In order to produce GFP-scFv fusion proteins, the verified m-gfp5 gene was fused separately to the amino terminal and also to the carboxyl terminal of MAC61.scFv and PV4.scFv to create 4 different variants, the 5'-mgfp.MAC61, 5'-MAC61.mgfp, 5'-mgfp.PV4, and 5'-PV4.mgfp. As a result of the data obtained on fluorescence intensity using free GFP from experiments described above *E. coli* BL21(DE3) *trxB*$^-$ was chosen as host strain for expression of mgfp.scFv fusion proteins

Among the transformed BL21(DE3) *trxB*$^-$ colonies, the mgfp.MAC61, mgfp.PV4, and PV4.mgfp transformants emitted fluorescence under UV illumination (Fig. 5.6). Although the fluorescence intensity was much lower than that of positive control colonies transformed with m-gfp5 gene alone, the fluorescent colonies could readily be differentiated from non-fluorescent colonies by the naked eye. Thus the GFP moiety of the GFP fusion could be used as an *in situ* marker for direct selection of transformed bacterial colonies. Of the three GFP-scFv fusions which produced fluorescent colonies, the mgfp.PV4 transformants showed the highest fluorescence intensity followed by the PV4.mgfp transformants and then the 5'-mgf.MAC61 transformants. The colonies of 5'-MAC61.mgfp transformants were not fluorescent under UV illumination.

Figure 5.5 : Fluorescent bacterial colonies containing m-gfp5.pRSET.

The diagrams displayed the fluorescent bacterial colonies of mgfp.pRSET transformants under the illumination of UV light. (a) shows JM109(DE3) transformant while (b) shows BL21(DE3) *trxB*$^{-}$ transformant.

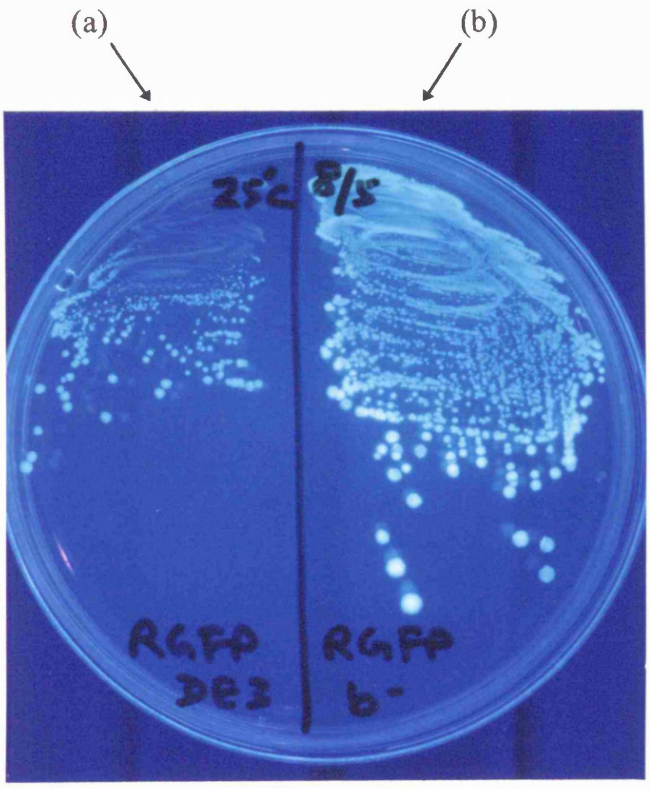
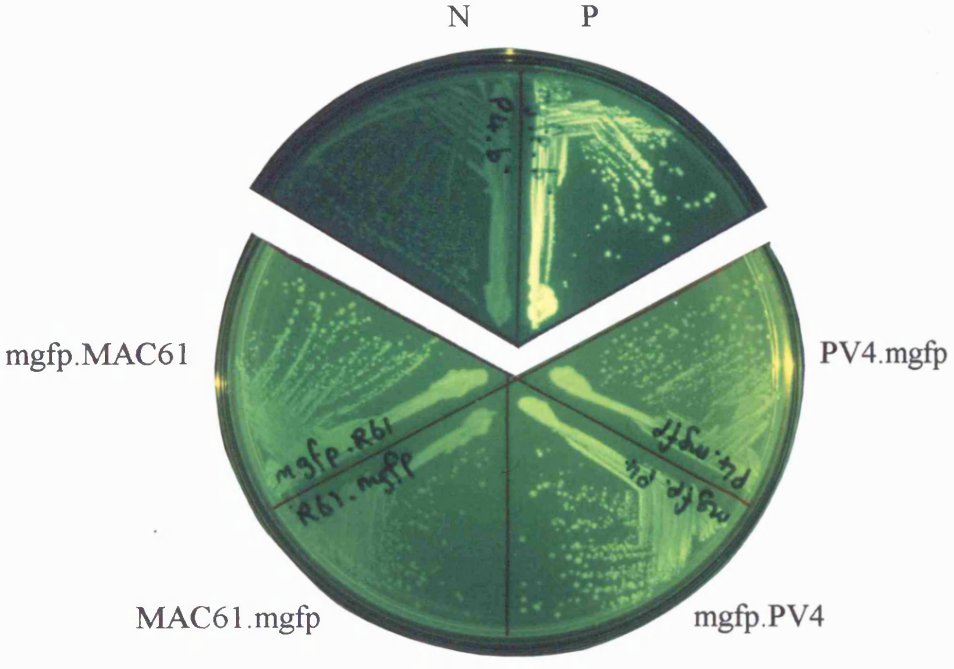


Figure 5.6 : Fluorescent bacterial colonies of GFP-scFv transformants.

The diagram depicts the fluorescent bacterial colonies of GFP-scFv fusion transformants under illumination with UV light. Culture P was a positive controlled sample using m-gfp5.pRSET transformants, while culture N was negative control using PV4.pRSET transformants.



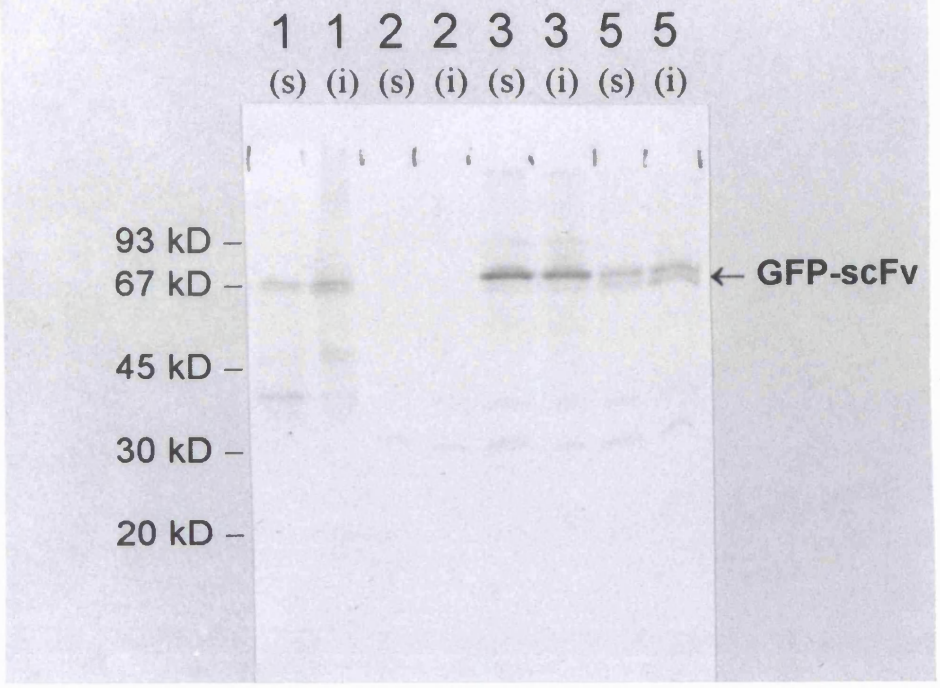
Total protein extracts of induced bacteria transformed with GFP.scFv fusion constructs were analysed by SDS-PAGE followed by Western Blotting and probed with anti-GFP antibody (Fig. 5.7a). It was found that fusion proteins at a molecular weight of about 60 kD (consistent with predictions based on 4.5 kD His-tag + 26 kD mgfp + 29.5 kD scFv), were produced with the mgfp.MAC61, mgfp.PV4 and PV4.mgfp constructs but not the MAC61.mgfp construct. This result agreed with the fluorescent colonies result in which only the fluorescent colonies of GFP-scFv transformants contained GFP-scFv fusion proteins and the fluorescence intensities were co-related with the amount of GFP-scFv fusion proteins showed on Western analysis.

Further Western analysis was also carried out on the mgfp.MAC61 and mgfp.PV4 products using anti-HSV-tag antibody as detection probe. (It should be noted that constructs MAC61.mgfp and PV4.mgfp did not contain HSV-tag as a result of cloning procedures involved in their production). Both the mgfp.MAC61 and mgfp.PV4 transformants produced fusion proteins at the appropriate molecular weight of about 60 kD (Fig. 5.7b). Since the HSV-tag was located at the carboxyl terminal of the fusion protein, this result confirms that complete full-length mgfp.MAC61 and mgfp.PV4 GFP-scFv fusion proteins had been produced.

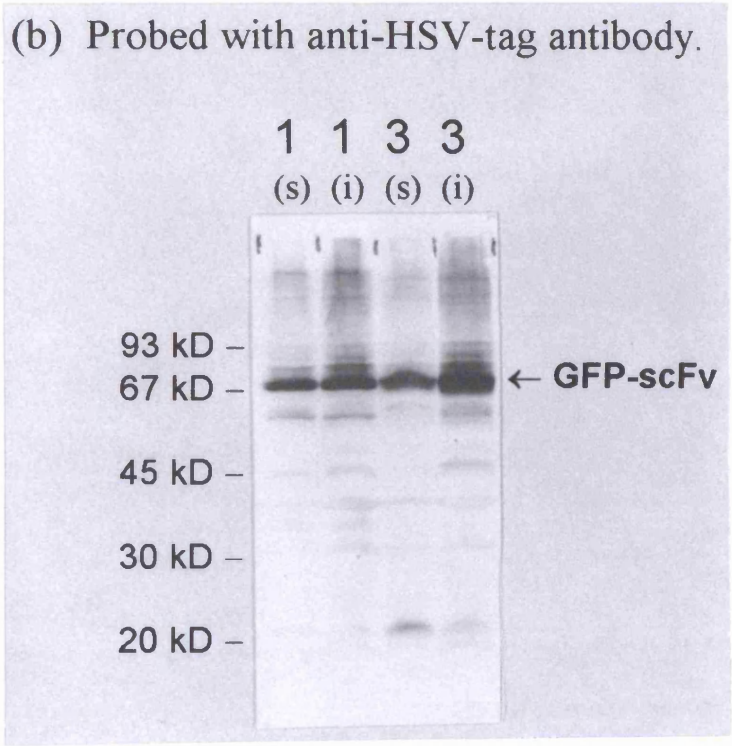
Western analysis showed that the expression level of mgfp.PV4 fusion protein was higher than that of PV4.mgfp while the expression level of mgfp.MAC61 was higher than that of MAC61.mgfp (Fig. 5.7a). This suggested that the expression levels were higher when GFP protein was attached to the amino terminal of scFv rather than the carboxyl terminal. This might be due to the low expression level of scFv genes in bacteria compared to the m-gfp5 gene. The presence of low expressing scFv genes at the amino terminal of the fusion protein could depress the expression of GFP-scFv fusion genes.

Western analysis also demonstrated that nearly half of the expressed mgfp.MAC61, mgfp.PV4 and PV4.mgfp GFP-scFv fusion proteins was in the soluble fraction. However, fluorescence intensity was not related to the solubility of the GFP-scFv fusion protein. When both the soluble and insoluble fractions of the

(a) Probed with anti-GFP antibody.



(b) Probed with anti-HSV-tag antibody.



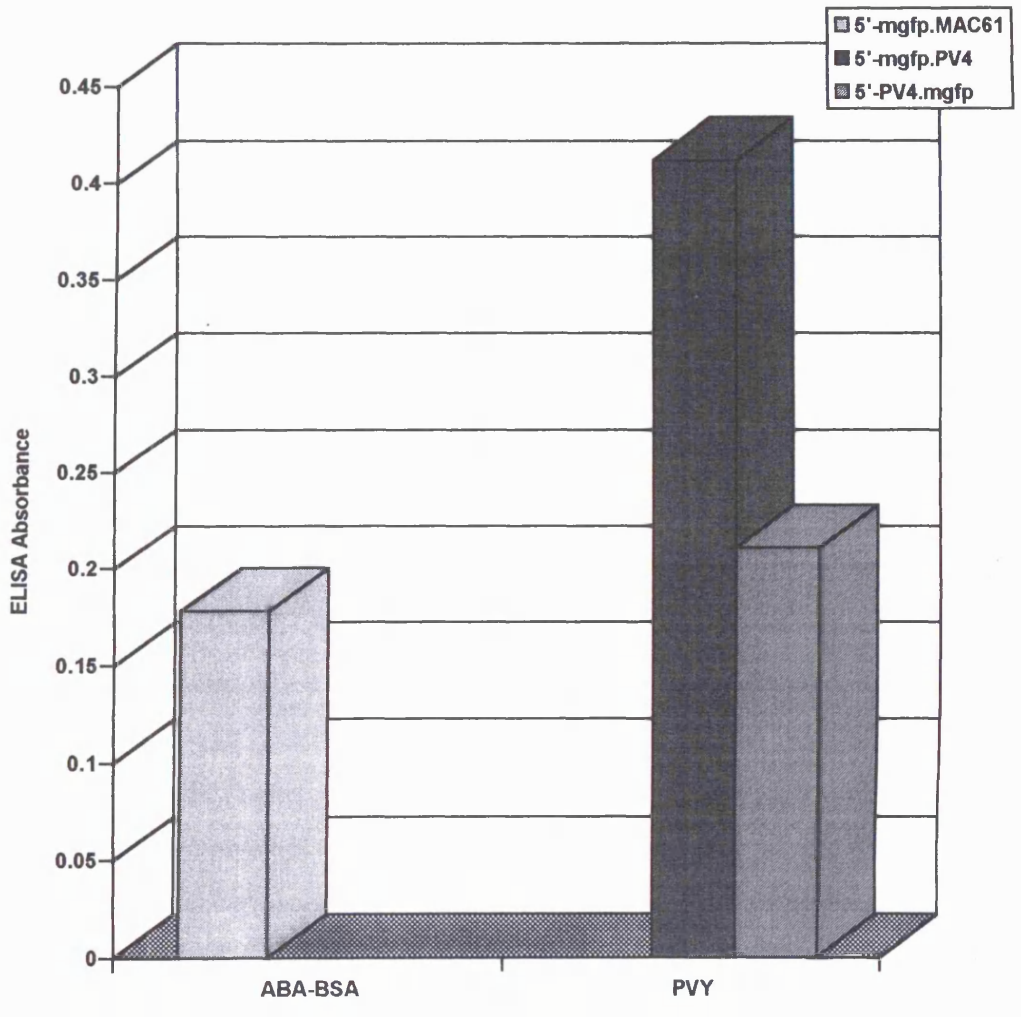
expressed mgfp.MAC61, mgfp.PV4, and PV4.mgfp GFP-scFv fusion proteins were viewed under UV illumination, both fractions exhibited similar fluorescence intensity (result not shown). Thus, both soluble and insoluble forms of the GFP.scFv fusion proteins were in a functional fluorescent conformation. This contrasts with the situation in the earlier reports on the free GFP proteins which demonstrated that formation of the functional chromophore conformation required the protein in soluble form (Cody *et al.*, 1993; Inouye and Tsuji, 1994; Heim *et al.*, 1994). It is conceivable that although the whole GFP-scFv fusion protein in the 'insoluble' fraction was not soluble the relatively hydrophilic GFP moiety might be exposed to the solvent environment in a similar way to the soluble free GFP protein - thus allowing the production of the functional fluorescent chromophore.

ELISA tests of functionality of the scFv moieties of the fusion proteins were carried out using total protein extracts of induced bacterial cultures using the anti-GFP antibody as detection probe. The result shown in Fig. 5.8 indicated that mgfp.MAC61, mgfp.PV4, and PV4.mgfp fusion proteins possessed antigen binding activity. Thus functional fusions between GFP and scFv genes are possible with both of the possible orientations of the fusion partners.

Preliminary attempts were made to detect antigen-bound GFP-scFv fusion proteins by fluorescence in an ELISA test analysed with a spectrofluorometer. No significant signals were obtained. This might be due to inadequate sensitivity of the detection instrument or loss of fluorescence on binding to the antigen. If the reason for the absence of detectable signal is inadequate sensitivity then the possibility of devising practicable immunoassays is seriously compromised. However the alternative possibility that reduction of fluorescence occurs with antigen binding might form the basis of a detection system.

Figure 5.8 : Antigen binding activity of GFP-scFv fusion proteins.

The ELISA assay was carried out using 100 µl of 250µg/ml of ABA-BSA conjugate and 100µg/ml of PVY extracts as coated antigen. 100 µl of 2x diluted total protein extracts of induced bacterial cultures were used as test samples. The assays were then probed with anti-GFP antibody and horse radish peroxidase-labelled anti-rabbit IgG antibody as second antibody.



5.3 Summary and Conclusions.

Diabodies constructed using a short (GGGGS) linker between the V_L and V_H domains of PV5.scFv were found to be expressed in bacteria at similar levels to the 'conventional' normal version of PV5.scFv gene. As estimated by Sephadex G-75 column chromatography about half of the bacterially expressed diabody version of PV5.scFv protein was found to be dimeric with a molecular weight of about 60 kD. However, the diabody version of PV5.scFv polypeptide was also found in fractions with molecular weights corresponding to monomeric or higher multimeric complexes. A similar fractionation pattern was obtained for the bacterially expressed normal version of PV5.scFv protein. In this case, the majority of the PV5.scFv polypeptides were monomeric and one third of the total PV5.scFv polypeptides were corresponding to the molecular weight of dimeric complexes.

About half of the bacterially expressed diabody version of PV5.scFv protein was found in the soluble fraction and exhibited stronger antigen binding activities than the normal version of PV5.scFv protein. This difference in binding activity is consistent with the diabodies possessing a higher avidity as a result of their bivalence.

Besides diabodies, a representative of a second class of engineered antibody fragment, the multifunctional immunoreagent namely green fluorescent protein scFv fusion protein (GFP-scFv) was constructed, expressed in bacteria and characterised. Fusions were constructed in which the GFP mutant, m-gfp5, was fused to either the amino terminal or the carboxyl terminal of both PV4 and MAC61 scFv genes. The constructed mgfp.PV5, PV5.mgfp, and the mgfp.MAC61 fusion proteins were expressed in bacteria and were shown to be fluorescent under UV illumination. However, no fluorescence were detected in bacteria transformed with MAC61.mgfp.

The expression level was consistently higher when the GFP protein was fused at the amino terminal rather than the carboxyl terminal of the scFvs. Maximum observed expression levels of GFP-scFv fusion proteins were significantly lower than those observed with free GFP and were in fact similar to the levels observed when free

scFvs were expressed. The GFP-scFv fusion proteins possessed antigen binding activity irrespective of whether the GFP was fused to the amino or carboxyl terminus of the scFv. Fusion to GFP certainly allowed direct visual identification of transformed bacterial colonies although further work will be required to determine if the fluorescence function of the GFP-scFv fusion protein could have application in diagnostic assays or *in vivo* localisation of antigen.

CHAPTER 6

Overall Discussion and Conclusions.

Antibody molecules have been extensively used in biological science, medicine, and industry. At present, the major proportion of the antibodies utilised in these applications (polyclonal or monoclonal) are produced 'conventionally' in animal cells and used in the form that they are produced with perhaps a chemically conjugated effector function added. The emergence of antibody engineering techniques has opened the door to novel approaches to manipulate and modify antibody molecules for the construction and production of immunoreagents with entirely new properties and effector moieties tailored specifically for particular applications.

In this study, antibody fragments, namely ABA26, MAC61, and MAC252 scFvs, were constructed using hybridoma cDNAs. The ABA26 hybridomas was derived from immunised mouse while MAC61 and MAC252 were derived from immunised rat. The ABA26 monoclonal antibody was capable of binding to ABA-C1-conjugate while MAC61 and MAC252 monoclonal antibodies were specific to ABA-C4'-conjugate.

Although selection and extensive optimisation of particularly suitable systems was found to be necessary, it did eventually prove possible to produce all three scFvs in *E. coli* expression systems - albeit at very low levels. Of the three scFvs, only MAC61.scFv was expressed at detectable, though low, levels in *P. pastoris* expression systems. Neither *Agrobacterium*-mediated stable transformation of tobacco plants with the scFv genes or transient expression of the genes using the PVX viral vector system resulted in the production of detectable scFv protein. The gene constructs were rigorously checked and the use of positive controls to verify techniques and eliminate technical artefactual explanations for the difficulties experienced in expressing ABA26, Mac61, and Mac252 scFv genes.

Thus although some antibodies can be readily expressed in heterologous systems (Plückthun, 1991a), including the positive controls used in this study, it is

clear from the work described in this thesis that there are very substantial differences in the expressibility of different individual antibodies. Other laboratories have also encountered difficulties in working with particular antibodies (Pietersz and MacKenzie, 1995) and informal discussions with other workers in the field suggests that there are unreported examples of antibodies which proved difficult to express.

The experimental results obtained in this study indicate that the difficulties experienced in expressing ABA26, MAC61, and MAC252 scFv genes in bacterial, yeast and plant expression systems relates to characteristics of these particular antibodies. Obviously the inability to produce detectable levels of these anti-ABA scFv genes in plant systems rules them out as candidates for immunomodulation studies of ABA physiology in plants. For such purposes it would be necessary to be able to produce plants which accumulated substantial amounts of anti-ABA antibody in a range of cellular compartments.

Since the vertebrate immune system is capable of producing a large number of different types of antibody genes when stimulated by immunisation it seems probable that some of the antibody genes generated would yield derivatives which are readily expressible in heterologous systems. However, to identify and isolate such genes from the family of different antibodies produced in response to immunisation would be an impractically time consuming process. Many scFv genes would need to be constructed and characterised in terms of both function and expressibility. Thus the selection of antibodies from which to derive scFv genes in the present work could not include criteria relating to their potential expressibility in heterologous systems and their characteristics in this respect were a matter of chance.

At the time this study commenced the likelihood of obtaining poorly expressible scFv genes was recognised though not considered high. Therefore a relatively small number (in total of 4 different hybridomas cells lines) of individual anti-ABA antibodies were selected to allow unsuitable antibodies to be discarded with work being eventually concentrated on one or two of them. The work reported here shows that by chance, none of the anti-ABA antibodies selected for study was suitable for applications involving scFv expression in heterologous systems.

Recent advances in phage display technology of antibody combinatorial libraries have illustrated the potential of such methods for direct and relatively straightforward selection of antibody genes which in addition to possessing desired antigen binding properties must of necessity be expressible in bacteria because the selection method itself involves their production in this heterologous host (Winter and Milstein, 1991). Combinatorial antibody gene libraries were first proposed by Ward *et al.* (1989) and Huse *et al.* (1989). The initial work involved the use of cells rich in B-lymphocytes which express antibody genes such as mouse spleen cells as the source of antibody genes.

Although it has proved possible to obtain antibody genes which encode antibodies against particular antigens of interest from naive libraries, the phage antibodies selected were generally low in affinity (Marks *et al.*, 1991). Although it is possible to improve the affinity of these antibodies an alternative approach to produce phage displayed libraries containing high affinity antibodies for a particular antigen involves using antibody genes from animals immunised with the antigen of interest. In this case advantage is taken of the ability of the animal immune system to rearrange and select high affinity antibodies (upon affinity maturation) prior to the construction of the phage display libraries.

Obtaining antibody genes from phage display libraries may offer some advantages over obtaining them from hybridomas. Although the final goal of the two strategies is the same, that is to obtain an engineered antibody against a particular antigen, the processes involved are quite different. In phage display libraries, the antigen binding activity and the abilities of individual original antibody genes to be expressed in heterologous hosts were not readily taken into consideration, but the engineered antibody genes (in phagemid) which were capable of expressing antibodies with antigen binding activities were selected based on their affinities. In contrast, with antibody genes from hybridomas, the properties of the antibodies produced by individual hybridoma are already known, thus antibody genes derived from hybridoma cells would be expected to produce recombinant antibodies with the same properties as the parental monoclonal antibody. In broad terms this is true although the engineered antibody fragments often have somewhat lower affinities than the parental monoclonal

(Leahy *et al.*, 1988; Desplancq *et al.*, 1994). Significantly in the context of the present work, the expressibility of the hybridoma-derived antibody genes in heterologous hosts is unknown.

The selection of an antibody genes from phage display libraries, therefore, has the advantage that such antibody genes must be expressible in *E. coli* (Hoogenboom and Winter, 1992; Griffiths *et al.* 1993). This is because bacteria was the host for producing the phage proteins, including the scFv fusion proteins. If the scFv genes are unexpressible in bacteria, it would not be displayed from the beginning at the surface of phage for selection. The work described here and other unpublished observations of other workers in the field supports, also gives a good general indication, that readily expressible scFv genes in bacteria also be expressible in other heterologous hosts such as yeast and plants.

Furthermore, the relative speed and ease of the generation and selection of antibody genes by phage display technology makes it feasible to investigate a greater number of potentially useful antibody genes in parallel than is practicable with the labour intensive and time consuming process of producing suitable monoclonal cell lines, characterising and isolating genes from them. An additional advantage of selection of antibodies from phage displayed antibody gene libraries is that gene combinations which might be rejected by the animal immune system remain available for selection.

Although phage display antibody technology encountered certain technical problems such as library size, phage stability and low affinity antibodies several years ago, those problems have largely been solved (Marks *et al.*, 1991; Nissim *et al.*, 1994). In addition, many techniques and much knowledge is now available to improve the affinity of the selected phage antibodies. These include site directed mutation or affinity maturation (Hawkins *et al.*, 1992), *in vitro* mutagenesis or CDR loop grafting (Gram *et al.*, 1992; Barbas *et al.*, 1992), chain shuffling or epitope imprinting (Hoogenboom *et al.*, 1994; Figini *et al.*, 1994). All of these could be used to improve the affinity of the selected antibody genes.

Both yeast and plants are possible candidates as heterologous hosts for large and very large scale antibody production respectively (Moffat, 1995). Production

costs will obviously be a significant factor in determining the viability of therapeutic recombinant antibodies and will be even more significant for industrial and other non-medical applications. More information on a number of aspects of the production of antibodies in these heterologous hosts will be required before their relative merits can be properly assessed. These include expression, accumulation, protein stability, and purification of recombination antibodies.

The work described in this thesis illustrates some technical problems with the *Pichia pastoris* expression systems. Although some antibodies have been shown to be expressed at high level, comparable to hybridoma cultures (Ridder *et al.*, 1995), however, others did not express well in this system. It is proposed here that the expression of antibody genes under the control of the *AOX1* promoter is depressed by the presence in the yeast cells of other naive functional genes also driven by the same *AOX1* promoter. The development of expression systems using other promoters might allow enhanced expression of certain antibodies in this system.

For plant expression, transient expression using the PVX plant virus expression vector proved to be an effective and rapid method for the production of sufficient quantities of plant-synthesised antibody for evaluation purposes prior to the time-consuming process of stable plant transformation. The results from this study suggest that scFv genes which could be expressed by transient expression were also expressible in stably transformed plants. Similarly those scFv genes which did not produce detectable scFv protein in the transient expression system did not produce detectable product following stable transformation. For those antibodies used in the present study which reached detectable levels, targeting to and retention in the plant ER resulted in the accumulation of higher levels of scFv protein than targeting to the cytoplasm or apoplast. Regardless of the location to which they were targeted none of the anti-ABA scFv proteins accumulated to detectable levels.

ScFv antibody genes that derived from hybridomas or obtained from phage display libraries can both be manipulated and modified in many ways to fit them for particular applications. One such variant of possible use in plant protection strategies is the diabodies or dimeric scFv proteins. A diabody derived from a monoclonal antibody which was able to bind to potato virus Y was constructed using a short

(GGGGS) linker between the scFv V_L and V_H domains. The diabody was expressed in *E. coli* and shown to associate into dimer. In functional tests it was shown to bind antigen more strongly than the corresponding scFv, presumably indicating increased avidity resulting from bivalence. Further work will be required to determine how it compares with the scFv version in plant protection strategies although bivalence of this antibody variant must allow the possibility of interference with pathogenic processes in ways not possible with monomeric scFv proteins.

Of the many possible types of recombinant antibody derivatives, antibody fusion proteins are amongst the most versatile. Recombinant antibody fusions with various effector function fusion partners will probably supersede chemically-bonded monoclonal antibody conjugates carrying detection and quantification probes. Selfector functions include enzymes, fluorophores and radioisotopes which have long been extensively used in immunodiagnosis. Problems arising from the use of those chemically linked effector function such as loss of antibody binding activity during chemical conjugation processes and poor reproducibility of production are eliminated with recombinant fusions. Furthermore, the use of such genetic engineering techniques may provide more economical production methods than chemical conjugation.

In this study, as an example of a recombinant antibody fusion carrying a fluorescent detection and localisation marker, scFv antibody fragments were fused with jellyfish green fluorescent protein at both the carboxyl and amino terminal ends of an scFv gene. The fusion protein was shown to be bifunctional in that it bound to its antigen and was also fluorescent when irradiated with UV light. Although further development would be required to increase detection sensitivity such recombinant fluorescent immunoreagents could find commercial application.

Although it has been proved that antibody engineering is generally workable, the potential of antibody engineering can only be realised if suitable methods can be devised to produce recombinant antibodies in sufficient amounts at an acceptable cost. This will only be achieved if appropriate expression system and in some cases suitable recombinant antibody purification methods can be developed. The investigation of antibody expression systems represents the main body of the work reported in this thesis.

In conclusion, the engineering of antibody genes has the potential to make major impacts in many areas. They provide for example a means to realise the concept of the therapeutic 'magic bullet' by P. Ehrlich 15 years ago (Ehrlich *et al.*, 1982). In the near future, genetic engineering will be used to produce antibodies or antibody fragments that have specific applications. The scope of application of these new human engineered molecules is enormous, it was a mixture of human imagination and inspiration that developed this technology in the first place, and it will be these same human qualities that finally determines its full potential.

REFERENCES

Alt, F.W., Yancopoulos, G.D., Blackwell, T.K., Wood, C., Thomas, E., Boss, M., Coffmann, R., Rosenberg, N., Tonegawa, S., and Baltimore, D. (1984). Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.*, **3**: 1209-1219.

Amsterdam, A., Lin, S., and Hopkins, N. (1995). The Aequorea victoria green protein can be used as a reporter in live zebrafish embryos. *Devel. Biol.*, **171**: 123-129.

Anand, N.N., Mandal, S., MacKenzie, C.R., Sadowska, J., Sigurskold, B., Young, N.M., Bundle, D.R., Narang, S.A. (1991). Bacterial expression and secretion of various single-chain Fv genes encoding proteins specific for a Salmonella serotype B O-antigen. *J. Biol. Chem.*, **266**: 21874-21879.

Anderson, H. and von Heijne, G. (1991). A 30-residues long "export initiation domain" adjacent to the signal sequence is critical for protein translocation across the inner membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.*, **88**: 9751-9754.

Artsaenko, O., Peisker, M., zur Nieden, U., Fiedler, U., Weiler, E. W., Muntz, K., and Conrad, U. (1995). Expression of a single-chain Fv antibody against abscisic acid creates a wilty phenotype in transgenic tobacco. *Plant J.*, **8**: 745-750.

Ayala, M., Balint, R.F., Fernandez-de-Cossio, M.E., Canaan-Haden, L., and Gavilondo, J.V. (1995). Variable region sequence modulates periplasmic export of a single-chain Fv antibody fragment in *Escherichia coli*. *BioTechniques*, **18**: 832-842.

Ayala, M., Due-as, M., Santos, A., Vazquez, J., Menendez, A., Silva, A., Gavilondo, J.V. (1992). Bacterial single-chain antibody fragments, specific for carcinoembryonic antigen. *BioTechniques*, **13**: 790-799.

Barbas III, C.F., Bain, J.D., Hoekstra, D.M., Lerner, R.A. (1992). Semisynthetic combinatorial antibody libraries : a chemical solution to the diversity problem. *Proc. Natl. Acad. Sci.*, **89**: 4457-4461.

Barbas III, C.F., Kang, A.S., Lerner, R.A., and Benkovic, S.J. (1991). Assembly of combinatorial antibody libraries on phage surfaces : The gene III site. *Proc. Natl. Acad. Sci. USA.*, **88**: 7978-7982.

Bardwell, J.C.A., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulphide bond formation in vitro. *Cell*, **67**: 581-589.

Batra, J. K., FitzGerald, D. J., Chaudhary, V. K., & Pastan, I. (1991). Single-chain immunotoxins directed at the human transferring receptor countering *Pseudomonas* exotoxin A or diphtheria-toxin:anti-TFR(Fv)-PE40 and DT388-anti-TFR(Fv). *Mol. Cell. Biol.*, **11**: 2200-2205.

Batra, J. K., FitzGerald, D., Gately, M., Chaudhary, V. K., & Pastan, I. (1990). Anti-Tac(Fv)-PE40, a single chain antibody-pseudomonas fusion protein directed at interleukin-2 receptor bearing cells. *J. Biol. Chem.*, **256**: 15198-15202.

Batra, J.K., Kasprzyk, P.G., Bird, R.E., Pastan, I., and King, C.R. (1992). Recombinant anti-erbB2 immunotoxins containing *Pseudomonas* exotoxin. *Proc. Natl. Acad. Sci.*, **89**: 5867-5871.

Baulcombe, D.C., Champman, S., and Santa Curz, S. (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *Plant. J.*, 7: 1045-1053.

Bebbington, C.D. (1991). Expression of an antibody genes in non-lymphoid mammalian cells. *Methods*, **2**: 136-145.

Benjamin, D.C., Berzofsky, J.A., East, I.J., Gurd, F.R.N., Hannum, C., Leach, S.J., Margolish, E., Michea, J.C., Miler, A., Prager, E.M., Reichlin, M., Sercarz, E.E., Smith-Gill, S.J., Todd, P.E., and Wilson, A.C. (1984). The antigenic structure of proteins : A reappraisal. *Annu. Rev. Immunol.*, **2**: 67-101.

Benkovic, S.J., Adams, J.A., Borders, C.L., Janda, K.D., and Lerner, R.A.(1990). The enzymatic nature of antibody catalysis development of multistep kinetic processing. *Science*, **250**:1135-1143

Bentley, G.A., Baulot, G., Riottot, M.M. and Poljak, R.J. (1990). The dimensional structure of an idiotope anti-idiotape complex. *Nature*, **348**:254-257.

Berek, C. (1993). Somatic mutation and memory. *Curr. Opin. Immunol.*, **5**: 218-222.

Berek, C. and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.*, **96**:23-41.

Bernard, O., Hozumi, N., and Tonegawa, S. (1978). Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell*, **15**: 1133-1144.

Berzofsky, J.A. and Berkower, I.J. (1984). Antibody-antigen interaction. *In* : Fundamental immunology. *Paul W.E. ed.*, Raven Press, New York. 595-644.

Better, M., Chang, C.P., Robinson, R.R., and Horwitz, A.H. (1988). *Escherichia coli* secretion of an active chimerical antibody fragment. *Science*, **240**: 1041-1043.

Biocca, S. Neuberger, M.S., and Cattanea, A. (1990). Expression and targeting of intracellular antibodies in mammalian cells. *EMBO J.*, **9**: 101-108.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Johson, S., Kaufman, B.M., Riordan, G.S., and Whitlow, M. (1988). Single-chain binding proteins. *Science*, **242**: 423-426.

Blomberg, B., Traunecker, A., Eisen, H., and Tonegawa, S. (1981). Organisation of four mouse λ light chain immunoglobulin genes. *Proc. Natl. Acad. Sci.*, **78**:3765-3769.

Bradford M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.*, **72**: 248-254.

Bregegere, F., Schwartz, J., and Bedouelle, H. (1994). Bifunctional hybrids between the variable domains of an immunoglobulin and the maltose-binding proteins of *Escherichia coli* - production purification and antigen-binding. *Protein Eng.*, 7: 271-280.

Briggs, M.S., Cornell, D.G., Dluphy, R.A., and Gierasch, L.M. (1986). Conformation of signal peptides included by lipids suggest initial steps in protein export. *Science*, **233**: 206-208.

Brinkmann, U., Pai, L.H., FitzGerald, D.J., Willingham, M., and Pastan, I. (1991). B3(Fv)-PE38KDEL, a single chain immunotoxin that causes complete regression of a human carcinoma in mice. *Proc. Natl. Acad. Sci.*, **88**: 8616-8620.

Brisson, N., Paszkowski, J., Penswick, J.R., Gronenborn, B., Potrykus, I., and Hohn, T. (1984). Expression of a bacteria gene in plants by using a viral vector. *Nature*, **310**: 511-514.

Brown, M.E., Renner, G., Field, R.P., Hassell, T. (1992). Process development for the production of recombinant antibodies using the glutamine synthetase (GS) system. *Cytotechnology*, **9**: 231-236.

Buckholz, R.G. and Gleeson, M.A.G. (1991). Yeast systems for the commercial production of heterologous proteins. *Bio/Technology*, **9**:1067-1072.

Cabilly, S., Riggs, A.D., Pande, H., Shively, J., Holmes, W.E., Rey, M., Perry, L.J., Wetzel, R. and Heyneker, H.L. (1984). Generation of antibody activity from immunoglobulin polypeptide chains produced in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, **81**: 3273-3277.

Cahill, K.B., Wirak, D.O., Mierz, D.V., Wunderlich, D.A., Rose, B.W., Shedd, D.G., Gillespire, A., and Zerler, B.R. (1993). Transgenic mice and transhybridomas producing mouse/humane anti-human interleukin 2 receptor recombinant antibodies. *Gene*, **137**: 101-107.

Cairns, E., Geschewender, H.H., Primke, M., Yamakawa, M., Traub, P., and Schweiger, H.G. (1978). Translation of animal virus RNA in the cytoplasm of a plant cell. *Proc. Natl. Acad. Sci. USA*, **75**: 5557-5559.

Campbell, A.M. (1991). *In* : Monoclonal antibody and immunosensor technology, Laboratory techniques in Biochemistry and Molecular Biology. *Elsevier, Amsterdam.*

Cattaneo, A. and Neuberger, M.S. (1987). Polymeric immunoglobulin M is secreted by transfectants of nonlymphoid cells in the absence of immunoglobulin J chain. *EMBO J.*, **6**: 2753-2758.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science*, **263**: 802-805.

Chapman, S., Kavanagh, T., and Baulcombe, D. (1992). Potato virus X as vector for gene expression in plants. *Plant. J.*, **2**: 549-557.

Chaudhary, V.K., Gallo, M. G., FitzGerald, D. J., and Pastan, I. (1990). A recombinant single-chain immunotoxin composed of anti-Tac variable regions and a truncated diphtheria toxin. *Proc. Natl. Acad.* Sci., **87**: 9491-9494.

Cheadle, C., Hook, L.E., Givol, D., Ricca, G.A. (1992). Cloning and expression of variable regions of mouse myeloma protein MOPC315 in *E. coli* : recovery of active Fv fragments. *Mol. Immunol.*, **29**: 21-30.

Chothia, C., and Lesk, A.M. (1987). Canonical structure for the hypervariable regions of immunoglobulins. *J. Mol. Biol.*, **196**: 901-917.

Chothia, C., Novotny, J., Bruccoleri, R.E., and Karplus, M. (1985). Domain association in immunoglobulin molecules : The packing of variable domains. *J. Mol. Biol.*, **186**: 651-663.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D., and Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature*, **352**: 624-628.

Clare J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. (1991). High-level expression of tetanus toxin fragments c in *Pichia pastoris* strains containing multiple tanden integrations of the genes. *Bio/Technology*, **9**: 455-460.

Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G., and Ward, W.W. (1993). Chemical structure of the hexapeptide chromophore of Aequorea green-fluorescent protein. *Biochemistry*, **32**: 1212-1218.

Condra, J.H., Sardana, V.V., Tomassini, J.E., Schlabach, A.J., Davis, M.E., Lineberger, D.W., Graham, D.J., Gotlib, L., and Colonno, R.J. (1990). Bacterial expression of antibody fragments that block human rhinovirus infection of cultured cells. *J. Biol. Chem.*, **265**: 2292-2295.

Cory, S., Tyler, B.M., and Adams, J.M. (1981). Sets of immunoglobulin Vκ genes : implications for the number of germline V genes. *J. Mol. Appl. Genet.*, **1**: 103-116.

Coudere, R. and Baratti, J. (1980). Oxidation of methanol by the yeast Pichia pastoris : purification and properties of alcohol oxidase. *Agric. Biol. Chem.*, **44**: 2279-2289.

Cregg, J.M., Madden, K.R., Barringer, K.J., Thill, G., and Stillman, C.A. (1989). Functional characterisation of the two alcohol oxidase genes from the yeast, Pichia pastoris. *Mol. Cell. Biol.*, *9*: 1316-1323.

Cregg, J.M., Vedvick, T.S., and Raschke, W.C. (1993). Recent advances in the expression from the yeast, Pichia pastoris. *Bio/Technology*, 7: 160-164.

Crowe, J.S., Hall, V.S., Smith, M.A., Cooper, H.J., and Tite, J.P. (1992). Humanised monoclonal antibody CAMPATH-1H : myeloma cell expression of genomic constructs, nucleotide sequences of cDNA constructs and comparison of effector mechanisms of meyloma and Chinese hamster ovary cell derived material. *Clin. Exp. Immunol.*, **87**:105-110.

Düring, K., Hippe, S., Kreuzaler, F., and Schell, J. (1990). Synthesis and self-assemble of a functional monoclonal antibody in transgenic *Nicotiana tabacum*. *Plant Mol. Biol.*, **15**: 281-293.

Davies, N.P., Rosewell, I.R., Richardson, J.C., Cook, G.P., Neuberger, M.S., Brownstein, B.H., Norris, M.L., and Bruggemann, M. (1993). Creation of mice expression human antibody light chains by introduction of a yeast artificial chromosome containing the core region of the human immunoglobulin *k* locus. *Bio/Technology*, **11**: 911-914.

Dawson, W.O., Lewandowski, D.J., Hilf, M.E., Bubrick, P., Raffo, A.J., Shaw, J.J., Grantham, G.L., and Dejardins, P.R. (1989). A tobacco mosaic virus-hybrid expression and loses an added gene. *Virology*, **172**: 285-292.

De Neve, M., De Loose, M., Jacobs, A., Van Houdt, H., Kauluza, B., Weildle, U., Van Montagu, M., Depicker, A. (1993). Assembly of an antibody and its derived antibody fragment in *Nicotiana* and *Arabidopsis*. *Transgen Res.*, **2**: 227-237.

Denzin, L.K. and Voss, E.W. Jr. (1992). Construction, characterisation and mutagenesis of an anti-fluorescent sing-chain antibody idiotype family. *J. Biol. Chem.*, **267**: 8925-8931.

Desplancq, D., King, D.J., Lawson, A.D.G., and Mountain, A. (1994). Multimerization behaviour of single chain Fv variants for the tumour-binding antibody B72.3. *Protein Eng.*, 7: 1027-1033.

Donson, J., Kearney, C.M., Hilf, M.E., and Dawson, W.O. (1991). Systemic expression of a bacterial gene by tobacco mosaic virus based vector. *Proc. Natl. Acad. Sci. USA*, **88**: 7204-7208.

Draper, J., Scott, R., and Armitage, P. (1988). *In* : Plant genetic, transformation and gene expression, a laboratory manual. *Backwell Sc. Pub., Oxford.*

Dreyer, W.J. and Bennett, J.C. (1965). The molecular basis of antibody formation : a paradox. *Proc. Natl. Acad. Sci.*, **54** : 864-869.

Dueñas, M., Vazquez, J., Ayala, M., Söderlind, E., Ohlin, M., Pérez, L., Borrebaeck, C.A.K., and Gavilondo, J.V. (1994). Intra- and Extracellular expression of an scFv antibody fragment in *E. coli* : Effect of bacteria strains and pathway engineering using GrpES/L chaperonins. *BioTechniques*, **16**: 476-483.

Early, O., Huang, H., Davis, M., Calame, K., and Hood, L. (1980). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA, V, D and J. *Cell*, **19**:981-992.

Edwards, K., Johstone and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nuc. Acids Res.*, **19**: 1349.

Ehrlich, P., Moyle, W.R., Moustafa, Z.A., and Canfield, R.E. (1982). Mixing 2 monoclonal antibodies yields enhanced affinity for antigen. *J. Immunol.*, **128**: 2709-2713

Farquhar, M.G. (1985). Progress in unravelling pathways of golgi traffic. *Ann. Rev. Cell. Biol.*, **1**: 447-488.

Figini, M., Marks, J.D., Winter, G., and Griffiths, A.G. (1994). *In vitro* assembly of repertories of antibody chains on the surface of phage by renaturation. *J. Mol. Biol.*, **239**: 68-78.

Friguet, B., Chaffotte, A.F., Djavadi-Ohaniance, L., and Goldberg, M. (1985). Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. *J. Immunol. Methods*, 77: 305.

Finkelstein, R.R. and Somervilla, C.R. (1990). Three classes of abscisic acid (ABA)-insensitive mutants of Arabidopsis define genes that control overlapping subsets of ABA responses. *Plant Physiol.*, **94**: 1172-1179.

Firek, S., Draper, J., Owen, M.R.L., Gandecha, A., Cockburn, B., and Whitelam, G.C. (1993). Secretion of a functional single-chain Fv protein in transgenic tobacco plants and cell suspension cultures. *Plant Mol. Biol.*, **23**: 861-870.

Fisch, I., Griffiths, A., and Winter, G. (1994). Generation of large phage antibody repertoires using in vivo recombination. *Proc. UCLA Meet., California.* Abstract T201.

Fischmann, T.O., Bentley, G.A., Bhat, T.N., Boulot, G., Mariuzza, R.A., Phillipps, S.E.V., Tello, D., and Poljak, R.J. (1991). Crystallographic refinement of the three-dimensional structure of the Fab D1.3 lysozyme complex at 2.5-A resolution. *J. Biol. Chem.*, **266**: 12915-12920.

Flach, J., Bossie, M., Vogel, J., Corbett, A., Jinks, T., Willins, D.A., and Silver, P.A. (1994). A yeast RNA-binding protein shuttles between the nucleus and the cytoplasm. *Mol. Cell. Biol.*, **14**: 8399-8407.

Foon, K.A. (1989). Biological response modifiers the new immunotherapy. *Cancer Res.*, **49**: 1621-1639.

Fouser, L.A., Swanberg, S.L., Lin, B.Y., Benedict, M., Kelleher, K., Cumming, D.A., Riedel, G.F. (1992). High level expression of chimeric anti-ganglioside GD2 antibody : genomic kappa sequences improve expression in COS and CHO cells. *Bio/Technology,* **10**:1121-1127.

Fromm, M., Taylor, L.P., Walbot, V. (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc. Natl. Acad. Sci. USA*, **82**: 5824-5828.

Galfre, G., Howe, S,C., Milstein, C., Butcher, G.W., Howard, J.C. (1977). Antibodies to major histocompatibility antigens produced by hybrid cell lines. *Nature*, **266**: 550-552.

Gandecha, A. R., Owen, M.R.L., Cockburn, B., and Whitelam, G.C. (1994). Antigen detection using recombinant, bifunctional single-chain Fv fusion proteins synthesised in *Escherichia coli. Protein Expr. Pruf.*, **5**: 385-365.

Gandecha, A.R., Owen, M.R.L., Cockburn, B., and Whitelam, G.C. (1992). Production and secretion of a bifunctional staphylococcal protein A::antiphytochrome single-chain Fv fusion protein in *Escherichia coli. Gene*, **122**: 361-365.

Glockshuber, R., Malia, M., Pfitzinger, I., and Plückthun, A. (1990). A comparison of strategies to stabilise immunoglobulin Fv-fragments. *Biochemistry*, **29**: 1362-1367.

Glockshuber, R., Schmidt, T., and Plückthun, A. (1992). The disulphide bonds in antibody variable domains : effects on the stability, folding in vitro, and functional expression in *Escherichia coli. Biochemistry*, **31**: 1270-1279.

Gorman, S.D., Clark, M.R., Routledge, E.G., Cobbold, S.P., and Waldmann, H. (1991). Reshaping a therapeutic CD4 antibody. *Proc. Natl. Acad. Sci.*, **88**: 4181-4185.

Goshorn, S.C., Svensson, H.P., Kerr, D.E., Somervilla, J.E., Senter, P.D., and fell H.P. (1993). Genetic construction, expression, and characterisation of a single chain anti-carcinoma antibody fused to beta-lactamase. *Cancer Res.*, **53**: 2123-2127.

Gram, H., Marconi, L.A., Barbas III, C.F., Collet, T.A., Lerner, R.A., and Kang, A.S. (1992). *In vitro* selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library. *Proc. Natl. Acad. Sci. USA*, **89**: 3576-3580.

Grantham, R., Gautier, C., Gouy, M., Jacobzone, M., and Mercier, R. (1981). Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucl. Acids Res.*, **9**: 43-69.

Griffiths, A.D. (1993). Production of human antibodies using bacteriophage. *Curr. Opin. Immunol.*, **5**: 263-267.

Griffiths, A.D., Malmqvist, M., Marks, J.D., Bye, J.M., Embleton, M.J., McCafferty, J., Balier, M., Holliger, K.P., Gorick, B.T., Et, A.L. (1993). Human anti-self antibodies with high specificity from phage display libraries. *EMBO J.*, **12**: 725-734.

Guan, C., Li, P., Riggs, P.D., and Inouye, H. (1987). Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. *Gene*, **67**: 21-30.

Guerineau, F., Woolston, S., Brooks, L., and Mullineaux, P. (1988). An expression cassette for targeting foreign proteins into chloroplasts. *Nucl. Acids Res.*, **16**: 11380.

Harlow, E. and Lane, D. (1988). Antibody-antigen interactions. *In* : Antibodies, a laboratory manual. Cold Spring Harbor Laboratory. 23-36.

Haseloff, J. and Amos, B. (1995). GFP in plants. *Trend Genet.*, **11**: 328-329.

Haseman, C.A., and Capra, J.D. (1990). High-level production of a functional immunoglobulin heterodimer in a baculovirus expression system. *Proc. Natl. Acad. Sci.*, **87**: 3942-3949.

Hawkins, R.E., Russell, S.J., Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J. Mol. Biol.*, **226**: 889-896.

Hein, M.B., Tang, Y., McLeod, D.A., Janda, K.D., and Hiatt, A. (1991). Evaluation of immunoglobulins from plant cells. *Biotechnol. Prog.*, **7**: 455-461.

Hermes, J.D., Parekh, S.M., Blacklow, S.C., Koster, H., and Knowles, J.R. (1989). A reliable method for random mutagenesis : the generation of mutant libraries using spiked oligodeoxyribonucleotide primers. *Gene*, **84**: 143-151.

Hetherington, A.M. and Quatrano, R.S. (1991). Mechanisms of action of abscisic acid at the cellular level. *New Phytol.*, **119** : 9-32.

Hiatt, A. (1990). Antibody produced in plants. *Nature*, **344**: 469-470.

Hiatt, A. and Ma, J.K. (1993). Characterisation and applications of antibodies produced in plants. *Int. Rev. Immunol.*, **10**: 139-152.

Hiatt, A., Cafferkey, R., and Bowdish, K. (1989). Production of antibodies in transgenic plants. *Nature*, **342**: 76-78.

Holliger, P, Prospero, T., and Winter, G. (1993). Diabodies : Small bivalent and bispecific antibody fragments. *Proc. Natl. Acad. Sci.*, **90**: 6444-6448.

Holmgren, A. (1984). Enzymatic reduction-oxidation of protein disulphides by thioredoxin. *Methods Enzymol.*, **107**: 295-300.

Honjo, T. (1983). Immunoglobulin genes. *Ann. Rev. Immunol.*, **1**: 499-528.

Hood, L., Kronenberg, M., and Hunkapiller, T. (1985). T cell antigen receptors and the immunoglobulin supergene family. *Cell*, **40**: 225-229.

Hoogenboom, H.R. and Winter, G. (1992). By-passing immunisation-human antibodies from synthetic repertoires germline VH gene segments rearranged *in vitro*. *J. Mol. Biol.*, **227**: 381-388.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.*, **19**: 4133-4137.

Hoogenboom, H.R., Jespers, L., Mahler, S., Roberts, A.J., Chiswell, D.J., and Winter, G. (1994). Human antibodies derived from rodent antibodies via epitope imprinted selection (EIS). *Proc. UCLA Meet., California.* Abstract T316.

Horwitz, A.H., Chang, C.P., Better, M., Hellstrom, K.E, and Robinson, R.R. (1988). Secretion of functional antibody and Fab fragment from yeast cells. *Proc. Natl. Acad. Sci. USA*, **85**: 8678-8682.

Huisman, M.J., Linthorst, H.J.M., Bol, J.F., and Cornelissen, B.J.C. (1988). The complete nucleotide sequence of potato virus X and its homologies at the amino acid level with various plus-stranded RNA viruses. *J. Gen. Virol.*, **69**: 1789-1798.

Hunkapiller, T. and Hood, L. (1989). Diversity of the immunoglobulin gene superfamily. *Adv. Immunol.*, **44**: 1-63.

Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting, M.M., Burton, D.R., Benkovic, S.J., Lerner, R.A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science*, **246**: 1275-1281.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., and Oppermann, H. (1988). Protein engineering of antibody binding sites : recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, **85**: 5879-5883.

Huston, J.S., Mudgett-Hunter, M., tai, M.S., McCartney, J., Warren, F., Haber, E., and Oppermann, H. (1991). Protein Engineering of single-chain Fv analogues and fusion protein. *Methods Enzym.*, **203**: 47-98.

Inouye, S. and Tsuji, F.I. (1994). Aequorea green fluorescent protein : Expression of the gene and fluorescent characteristics of the recombinant protein. *FEBS Letters*, **341**: 277-280.

Jaffers, G.J., Fuller, T.C., Cosimi, A.B., Russell, P.S., Winn, H.J., and Colvin, R.B. (1986). Monoclonal antibody therapy. Anti-iodiotypic and non-anti-iodiotypic antibodies to OKT3 arising despite intense immunosuppression. *Transplantation*, **41**: 572-578.

Johansson, M., Nilsson, I., and von Heijne, G. (1993). Positively charged amino acids placed next to the signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. *Mol. Gen. Genet.*, **239**: 251-256.

Kabat, E.A. and Wu, T.T. (1991). Identical V region amino acid sequences and segments of sequences in antibodies of different specificity. *J. Immunol.*, **14**: 1709-1719.

Kabat, E.A., Wu, T.T., and Bilfsky, H. (1977). Unusual distribution of amino acid in complementary-determine (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles. *J. Biol. Chem.*, **252**: 6609-6616.

Kain, S.R., Adams, M., Kondepudi, A., yang, T.T., Word, W.W., and Kitts, P. (1995). The green fluorescent protein as a reporter of gene expression and protein localization. *BioTechniques*, **19**: 650-655.

Kamitani, S., Akiyama, Y., and Ito, K. (1992). Identification and characterisation of an Escherichia coli gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. *EMBO J.*, 11: 57-62.

Kang, A.S., Bargbas, C.F., Janda, K.D., Benkovic, S.J., and Lerner, R.A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA*, **88**: 4363-4366.

Kelley, E.F., O'Connell, M.P., Carter, P., Presta, L.C.E., Covarrubias, M., Snedecor, B., Bourell, J.H., and Vetterlein, D. (1992). Antigen binding thermodynamics and antiproliferative effects of chimeric and humanised anti-p185^{HER2} antibody Fab fragments. *Biochemistry*, **31**: 5435-5441.

Kettleborough, C.A., Saldanha, J., Ansell, K.H., and Bendig, M.M. (1993). Optimisation of primers of cloning libraries of mouse immunoglobulin gene using PCR. *Eur. J. Immunol.*, **23**: 206-211.

Klein, J. (1982). *In* : Immunology : the Science of self-nonself discrimination. Wiley, New York.

Klein, T.M., Wolf, E.D., Wu, R., Sanford, J.C. (1987). High velocity microprojectiles for delivering nucleic acids into living cells. *Nature*, **327**: 70-73.

Knapp, W., Dorken, B., Gilks, W.R., Rieber, E.P., Schmidt, R.E., Stein, H., von dem Borne, AEGKr (1989). *In* : Leukocyte typing IV. *Oxford University Press, Oxford.*

Knappik, A. and Plückthun, A. (1994). An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. *BioTechniques*, **17**: 754-761.

Kohler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, **256**: 495-497.

Kohler, G. and Milstein, C. (1976). Derivation of specific antibody-producing tissue cultures and tumour lines by cell fusion. *Eur. J. Immunol.*, **6**: 511-519.

Kolmar, H., Ferando, E., Hennecke, F., Wippler, J., and Fritz, H.J. (1992). General mutagenesis/gene expression produce for the construction of variant immunoglobulin domains in *Escherichia coli*. *J. Mol. Biol.*, **228**: 359-365.

Kreitman, R.J., Chaudhary, V.K., Waldmann, T., Willingham, M.C., FitzGerald, D.J., and Pastan, I. (1990). The combinant immunotoxin anti-TAC(Fv)-pseudomonas exotoxin 40 is cytotoxic toward peripheral-blood malignant-cells from patients with adult T-cell leukaemia. *Proc. Natl. Acad. Sci.*, **87**: 8291-8295

Larrick, J.W., Wallace, E.F., Coloma, M.J., Bruderer, U., Lang, A.B., and Fry, K.E. (1989). Therapeutic human antibodies derived from PCR amplification of B-cell variable regions. *Immunol. Rev.*, **130**:69-85.

Leahy, D.J., Rule, G.S., Whittaker, M.M., and McConnell, H.M. (1988). Sequences of 12 monoclonal anti-dinitrophenyl spin-label antibodies for NMR studies. *Proc. Acad. Sci. USA.*, **85**: 3661-3665.

Leder, P. (1982). The genetic of antibody diversity. *Sci. Am.*, **246**: 72-83.

Lerner, R.A., Benkovic, S.J. and Schultz, P.G. (1991). At the crossroads of chemistry and immunology - catalytic antibodies. *Science*, **252**: 659-667.

Lerner, R.A., Kang, A.S., Bain, J.D., Burton, D.R., and Barbas C.F.III. (1992). Antibodies without immunisation. *Science*, **258**: 1313-1314.

Li, P., Bechwith, J., and Inouye, H. (1988). Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, 88: 7685-7689.

Li, Y. (1996). *In thesis* : Antibody engineering studies using filamentous bacteriophage display technology. *Botany Department, Leicester University, UK.*

Liu, A.Y., Robinson, R.R., Hellstrom, K.E., Murray, E.D. Jr, Chang, C.P., Hellstrom, I. (1987). Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells. *Proc. Natl. Acad. Sci.*, **84**:3439-3443.

LoBuglio, A.F. and Khazaeli, M.B. (1993). Clinical studies with murine and chimeric 17-1A in colorectal cancer patients. *In* : Protein Engineering of Antibody Molecules for Prophylactic and Therapeutic Applications in Man. (Clark, M. ed.) *Academic Titles, Nottingham.* 171-177.

LoBuglio, A.F., Wheeler, R.H., Trang, J. 'em (1989). Mouse/human chimeric monoclonal antibody in man : kinetics and immune response. *Proc. Natl. Acad. Sci.*, **86**: 4220-4224.

LoBuglio, A.F., Wheeler, R.H., Trang, J., Haynes, A., Rogers, K., Harvey, E.B., Sun, I., Ghrayeb, J., Khazaeli, M.B. (1989). Mouse/human chimeric monoclonal antibody in man : Kinetics and immune response. *Proc. Natl. Acad. Sci.*, **86**: 4200-4224.

Luo, D., Mah, N., Wishart, D., Zhang, Y., Jacobs, F., and Martin, L. (1996). Construction and expression of bi-functional proteins of single-chain Fv with effector domains. *J. Biochem.*, **120**: 229-232.

Ma, J.K.C., Hiatt, A., Hein, M., Vine, N.D., Wang, F., Stabila, P., Vandolleweerd, C., Mostov, K., and Lenner, T. (1995). Generation and assemble of secretary antibodies in plants. *Science*, **268**: 716-719.

Ma, J.K.C., Lehner, T., Stabila, P., Fux, C.I., and Hiatt, A. (1994). Assembly of monoclonal antibodies with IgG1 and IgA heavy chain domains in transgenic tobacco plants. *Eur J. Immunol.*, **24**: 131-138.

Ma, J.K.C., Lehner, T., Stabila, P., Fux, C.I., and Hiatt, A. (1994). Assemble of monoclonal antibodies with IgG1 and IgA heavy chain domains in transgenic tobacco plants. *Eur. J. Immunol.*, **24**: 131-138.

MacKenzie, C.R., Sharma, V., Brummell, D., Bilous, D., Dubuc, G., Sadowska, J., Young, N.M., Bundle, D.R., and Narang, S.A. (1994a). Effect of Cλ and Cκ domain switching on Fab activity and yield in *E. coli* : synthesis and expression of gene encoding two anti-carbohydrate Fabs. *Bio/Technology*, **12**: 390-395.

MacKenzie, R. Szabo, A., Clark, I., Evans, S., Hill, I., MacManus, J., Bundle, D., Young, M., and Narang, S. (1994b). Characterisation of bifunctional fusion proteins consists of a single-chain Fv and a metal-binding component. *Proc. IBC Int. Conf.* , Abstract 45.

Maffat, A.S. (1995). Exploring transgenic plants as a new vaccine source. *Science*, **268**: 658-660.

Maina, C.V., Riggs, P.D., Grandea, A.G. III, Slatko, B.E., Moran, L.S., Tagliamont, J.A., McReynolds, L.A., and Guan, C. (1988). A vector to express and purify foreign proteins in *Escherichia coli* by fusion to, and separation from, maltose binding protein. *Gene*, **74**: 365-373.

Malby, R.L., Caaldwell, J.B., Gruen, L.C. 'em (1993). Recombinant single-chain antibody : expression, characterisation and crystallisation in complex with antigen. *Proteins, Struc. Funct. Gen.*, **16**: 57-63.

Marks, J.D., Tristrem, M., Karpas, A., and Winter, G. (1991). Oligonucleotide primers for polymerase chain reaction amplification of human immunoglobulin variable genes and design of family-specific oligonucleotide probe. *Eur. J. Immunol.*, **21**: 985-991.

Marshall, J., Molly, R., Moss, G.W.J., Howe, J.R., and Hughes, T.E. (1995). The jellyfish green fluorescent protein : a new tool for studing ion channel expression and function. *Neuron*, **14**: 211-215.

Max, E.E. (1984). Immunoglobulins : molecular genetic. In : Fundamental Immunology. 167-204. Raven Press, New York.

Meredith, R.F., Khazaeli, M.B., Plott, W.E. 'em (1992). Phase I trial of iodine-131-chimeric B72.3 (human IgG4) in metastatic colorectal cancer. *J. Nucl. Med.*, **33**:23-29.

Merrison, S.L., Johson, M.J., Herzenberg, L.A. and Oi, A.T. (1984). Chimeric human antibody molecules : mouse antigen binding domains with human constant region domains. *Proc. Natl. Acad. Sci.*, **81**: 6851-6855.

Milenic, D.E., Yokota, T., Filpula, D.R., Finkelman, M.A.J., Dood, S.W., Wood, J.F., Whitlow, M., Snoy, P., and Schom, J. (1991). Construction, binding properties, metabolism, and tumour targeting of a single-chain Fv derived from the pancarcinoma monoclonal antibody CC49. *Cancer Res.*, **51**: 6363-6371.

Miler, J., Bothwell, A., and Storb, U. (1981). Physical linkage of the constant region genes for immunoglobulins λ1 and λ3. *Proc. Natl. Acad. Sci.*, **78**: 3829-3833.

Miller, L.K. (1988). Baculovirus as gene expression vectors. *Annu. Rev. Microbial.*, **42**: 177-200.

Munro, S. and Pelham, H.R.B. (1986). An HSP-70-like protein in the endoplasmic reticm identity with the 78-kilodalton glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell*, **46**: 291-300.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.*, **15**: 473-497.

Nabeshima, K., Kurooka, H., Takeuchi, M., Kinoshita, K., Nakaseko, Y., and Yanagida, M. (1995). The p93dis, which is required for sister chromatid separation, is a novel microtubule and spindle pole body-associating protein phosphorylated at the Cdc2 target sites. *Genes Devel.*, **9**: 1572-1585.

Neri, D. de Lalla, C., Petrul, H., Neri, P., and Winter, G. (1995a). Calmodulin as a versatile tag for antibody fragments. *Bio/Technology*, **13**: 373-377.

Neri, D., Momo, M., Prospero, T., and Winter, G. (1995b). High-affinity antigen binding by chelating recombinant antibodies (CRAbs). *J. Mol. Biol.*, **246**: 367-373.

Neuberger, M.S. (1983). Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. *EMBO J.*, **2**: 1373-1378.

Neuberger, M.S. (1994). Antibody repertoires from transgenic mice. *In* : The Shape of Antibody to Come. *London*, 16-41.

Nishimura, A., Morita, M., Nishimura, Y., and Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucl. Acids Res.*, **18**: 20.

Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D., and Winter, G. (1994). Antibody fragments from a "single pot" phage display library as immunochemical reagents. *EMBO J.*, **13**: 692-698.

Oi, V.T., Morrison, S.L., Herzenberg, L.A., and Berg, P. (1983). Immunoglobulin gene expression in transformed lymphoid cells. *Proc. Natl. Acad. Sci.*, **80**: 825-829.

Olson, K.R., McIntosh, J.R., and Olmsted, J.B. (1995). Analysis of MAP4 function in living cells using green fluorescent protein (GFP) chimeras. *J. Cell. Biochem.*, **130**: 639-650.

Orlandi, R., Gussow, D.H., Jones, P.T.,, Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc. Natl., Acad. Sci.*, **86**: 3833-3837.

Owen, M.RL., Gandecha, A., Cockburn, B., and Whitelam, G.C. (1992). Synthesis of a functional anti-phytochrome single-chain Fv protein in transgenic tobacco. *Bio/Technology*, **10**: 790-794.

Pack, P. and Plückthun, A. (1992). Miniantibodies : use of amphipathic helices to produce functional flexibly linked dimeric Fv fragments with high avidity in *Escherichia coli*. *Biochemistry*, **31**: 1579-1584.

Padlan, E.A. (1993). Anatomy of the antibody molecule. *Mol. Immunol.*, **31**: 169-217.

Padlan, E.A., Abergel, C., and Tipper, J.P. (1995). Identification of specificity-determining residues in antibodies. *FASEB J.*, **9**: 133-139.

Pantoliano, M.W., Bird, R.E., Johson, S., Asel, E.D., Dodd, S.W., Wood, J.F., and Hardman, K.D. (1991). Conformational stability folding and ligand-binding affinity of single-chain Fv immunoglobulin fragments expressed in *Escherichia coli*. *Biochemistry*, **31**: 10117-10125.

Paul, W.E. (1984). *In* : Fundamental Immunology. Raven Press, New York.

Pelham, H.R.B. (1989). Control of protein exit from the endoplasmic reticulum. *Ann. Rev. Cell. Biol.*, **5**: 1-24.

Pietersz , G.A. and MacKenzie, F.C. (1995). In : Monoclonal antibody, the second generation. Chap. 5 : The genetic engineering of antibody constructs for diagnosis and therapy. Heddy Zola Ed., BIOS Scientific Publish.

Plückthun, A. (1991a). Antibody engineering : Advances from the use of *Escherichia coli* expression systems. *Biotechnology*, **9**: 545-551.

Plückthun, A. (1991b). Antibody engineering. *Curr. Opin. Biotech.*, **2**: 238-246.

Plückthun, A. and Skerra, A. (1989). Expression of functional antibody Fv and Fab fragments in *Escherichia coli*. *Methods Enzymol*, **178**: 498-515.

Pollack, S.J., Jacobs, J.W., and Schultz, P.G. (1986). Selective chemical catalysis by an antibody. *Science*, **234**: 1570-1573.

Posner, B., Smiley, J., Lee, I., and Benkovic, S. (1994). Catalytic antibodies : perusing combinatorial libraries. *TIBS*, **19**:145-150.

Potter, K.N., Li, Y., and Capra, J.D. (1993). Antibody production in the baculovirus expression system. *Int. Rev. Immunol.*, **10**: 103-112.

Proba, K., Ge, L., and Plückthun, A. (1995). Functional antibody single-chain fragments from the cytoplasm of *Escherichia coli* : influence of thioredoxin reductase (TrxB). *Gene*, **159**: 203-207.

Quarrie, S.A. and Galfre, G. (1985). Use of different hapten-protein conjugates immobilised on nitro-cellulose to screen monoclonal antibodies to abscisic acid. *Anal. Biochem.*, **151**: 389-399.

Ridder, R., Schmitz, R., Legay, F., and Gram, H. (1995). Generation of rabbit monoclonal antibody fragments from a combinatorial phage display library and their production in the yeast Pichia pastoris. Bio/Technology, **13**: 255-260.

Riechmann, L., Clark, M., Waldman, H., Winter, G. (1988a). Reshaping human antibodies for therapy. *Nature*, **328**: 731-734.

Riechmann, L., Foote, J., and Winter, G. (1988b). Expression of an antibody Fv fragment in myeloma cells. *J. Mol. biol.*, **203**: 825-828.

Roguska, M.A. Pedersen, J.T., Keddy, C.A., Henry, A.H., Searle, S.J., Lambert, J.M., Goldmacher, V.S., Blattler, W.A., Rees, A.R., and Guild, B.C. (1994). Humanisation of murine monoclonal antibodies through variable domain resurfacing. *Proc. Natl. Acad. Sci.*, **91**: 969-973.

Rubb, F., Acha-Orbea, H., Hengartner, H., Zinkernagel, R., and Joho, R. (1985). Identical V T cell receptor genes used in alloreactive cytotoxic and antigen plus I-A specific helper T cells. *Nature*, **315**: 425-427.

Sakano, H., Maki, R., Kurosawa, Y., Roeder, W., and Tonegawa, S. (1980). Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. *Nature*, **286**:676-683.

Sandhu, J.S. (1992). Protein engineering of antibodies. *Crit. Rev. Biotech.*, **12**: 437-462.

Sanz, I. (1991). Multiple mechanisms participate in the generation of diversity of Herman H chain CDR3 regions. *J. Immunol.*, **147**: 1720-1729.

Savage, P., So, A., Spooner, R.A., and Epenetos, A.A. (1993). A recombinant single chain antibody interleukin-2 fusion protein. *Br. J. Cancer*, **67**: 304-310.

Schilling, J., Clevinger, B., Davie, J.M., and Hood, L. (1980). Amino acid sequences of homogeneous antibodies to dextran and DNA rearrangements in heavy chain V-region gene segments. *Nature*, **283**:35-40.

Schlossman, S.F., Boumsell, L., Gilks W., 'em (1994). CD antigens 1993. *The Immunologist*, **2**: 28-29.

Schmidt, T.G.M. and Skerra, A. (1993). The random peptide library-assisted engineering of a C-terminal affinity peptide useful for the detection and purification of a functional Ig Fv fragment. *Protein Eng.*, **6**: 109-122.

Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A.M., Borst-Vrenssen, A.W.M., Zilverentant, J.F., Bosch, D., Stiekema, W., Gommers, F.J., Schots, A., and Bakker, J. (1996). The C-terminal KDEL sequence increase the expression level of a single chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol. Biol.*, **30**: 781-793.

Scorer, C.A., Buckholz, R.G., Clare, J.J., and Romanos, M.A. (1993). The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast Pichia pastoris. *Gene*, **136**: 111-119.

Sharon, J. (1992). Structural correlates of high affinity : three engineered amino acid substitutions can increase the affinity of an anti-p-azophenylarsonate antibody 200 fold. *Proc. Natl. Acad. Sci. USA*, **87**: 4814-4817.

Sharon, J., Geffer, M.L., Manser, T., Morrison, S.L., Oi, V.T., and Ptashne, M. (1984). Expression of a VHCκ chimeric protein in mouse myeloma cells. *Nature*, **309**: 364-367.

Shibui, T., and Nagahari, K. (1992). Secretion of a functional Fab fragment in *Escherichia-coli* and the influence of culture conditions. *Appl. Micro. Biotech.*, **37**: 352-357.

Shimizu, A., Takahashi, N., Yaoita, Y., and Honjo, T. (1982). Organisation of the constant region gene family of the mouse immunoglobulin heavy chain. *Cell*, **28**: 499-506.

Shu, L., Qi, C.F., Schlom, J., and Kashimiri, S.V.S. (1993). Secretion of a single-gene-encoded immunoglobulin from myeloma cells. *Proc. Natl. Acad. Sci.*, **90**: 7995-7999.

Skerra, A. (1993). Bacteria expression of immunoglobulin fragments. *Curr. Opin. Immunol.*, **5**: 256-262.

Skerra, A. and Plückthun, A. (1988). Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli. Science.*, **240**: 1038-1040.

Sreekrishna, K., Nelles, L., Potenz, R., Cruse, J., Mazzaferro, P., Fish, W., Fuke, M., Holden, K., Phelps, D., Wood, P., and Parker, K. (1989). High level expression, purification and characterisation of recombinant human tumour necrosis factor synthesised in the methylotrophic yeast Pichia pastoris. *Biochemistry*, **28**: 4117-4125.

Staerz, U.D., Kanaggawa, O., and Bevan, M.J. (1985). Hybrid antibodies can target sites for attack by T-cells. *Nature*, **314**: 628-631.

Steinmetz, M. (1986). Genes of the immune system. *In* : Genetic Engineering. **5**: 117-158. *Academic Press, New York.*

Stemmer, W.P.C., Morris, S.K., Kautzer, C.K., and Wilson, B.S. (1993). Increased antibody expression from *E. coli* through wobble-base library mutagenesis by enzymatic inverse PCR. *Gene*, **123**: 1-7.

Stephens, S., Athwal, D., Chaplin, L., Entage, S., Sopwith, M., Temple, G., Vetterlein, O., Bodmer, M. (1993). Pharmacokinetics of CDP571, an anti-TNFα engineered human antibody in the cynomologous monkey and man. *Circ. Shock., Suppl.*, **1**:60.

Stewart, J.D. and Benkovic, S.J. (1993). Catalytic antibodies - mechanistic and practical considerations. *Chem. Soc. Rev.*, **22**: 213-219.

Studnicka, GM, Soares, S., Better, M., Williams, R.E., Nadell, R., and Harwitz, A.H. (1994). Human engineered monoclonal antibodies retain full specific binding activity by preserving non-CDR complementary-modulating residues. *Protein Eng.*, **7**: 805-814.

Tai, M. S. (1990). A bifunctional fusion protein containing Fc-Binding B of stachlococcal Protein A Amino Terminal to Antidigoxin Single chain Fv. *Biochemistry*, **29**: 35.

Takagi, H., Morinaga, Y., Tsuchiya, M., Ikemura, G., and Inouye, M. (1988). Control of folding of proteins secreted by a high expression secretion vector, pIN-III-*ompA* : 16-fold increase in production of active subtilisin E in *Escherichia coli. Bio/Technology*, **6**: 948-950.

Takkinen, K., Laukanen, M.L., Sizmann, D., Alfthan, K., Immonen, T., Vanne, L., Kaartinen, M., Knowles, J.K.C., and Teeri, T.t. (1991). An active single-chain antibody containing a cellulase linker domain is secreted by *Escherichia coli. Protein Eng.*, **4**: 837-841.

Taub, R., Gould, R.J, Garsky, V.M., Ciccarone, T.M., Huxie, J., Friedman, P.A., and Shattil, S.J. (1989). A monoclonal antibody against the platelet fibrinogen receptor contains a sequence that mimics a receptor recognition domain in fibrinogen. *J. Biol. Chem.*, **264**: 259-265.

Tavladoraki, P., Benvenuto, E., Trinca, S., De Martinis, D., Cattaneo, A., and Galeffi, P. (1993). Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack. *Nature*, **366**: 469-472.

Tempest, P.R., Bremner, P., Lambert, M., Taylor, G., Furze, J.M., Carr, F.J., and Harris, W.J. (1991). Reshaping a human monoclonal antibody to inhibit human respiratory syncytial virus infection in vivo. *Bio/Technology*, 9: 266-271.

Tonegawa, S., Brack., Hozumi, N., and Schuller, R. (1977). Cloning of an immunoglobulin variable region gene from mouse embryo. *Proc. Natl. Acad. Sci.*, **74**: 3518-3522.

Tramontano, A., Janda, K.D., and Lerner, R.A. (1986). Catalytic antibodies. *Science*, **234**: 1566-1570.

Ueda, Y., Tsumoto, K., Watanabe, K., and Kumagai, I. (1993). Synthesis and expression of a DNA encoding the Fv domain of an anti-lysozyme monoclonal antibody, HyHEL10, in *Streptomyces lividans*. *Gene*, **129**: 129-134.

Voss, A., Nierbach, M., Hain, R., Hirsch, H.J., Liao, Y., Kreuzaler, F., and Fischer, R. (1995). Reduced virus infectivity in *N. tabacum* secreting a TMV-specific full size antibody. *Mol. Br.*, **1**: 15-26.

Walton, D.C. (1980). Biochemistry and physiology of abscisic acid. *Ann. Rev. Plant Physiol.*, **31** : 453-89.

Wang, S. and Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. *Nature*, **369**: 400-403.

Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature*, **34**: 544-546.

Webber, K.O., Reiter, Y., Brinkmann, U., Kreitman, R.T., and Pastan, I. (1995). Preparation and characterisation of a disulphide-stabilised Fv fragment of the anti-Tac antibody : comparison with its single-chain analogue. *Mol. Immunol.*, **32**: 249-258.

Weiler, E.W. (1980). Radioimmunoassays for the differential and direct analysis of free and conjugated abscisic acid in plant extracts. *Planta*, **148**: 262-272.

Wels, W., Harwerth, I.M., Zwicki, M., Hardman, N., Groner, B., and Hynes, N.E. (1992). Construction, bacterial expression and characterisation of a bifunctional single-chain antibody-phosphatase fusion protein targeted to the human erB-2 receptor. *Bio/Technology*, **10**: 1128-1132.

White, N., Adair, J., Lloyd, C., Jenkins, L., Devine, J., Schlom, J., Raubitshek, A., Colcher, D., Bodmer, M. (1987). Expression in COS cells of a mouse/human chimeric B72.3 antibody. *Protein Eng.*, **1**: 499-505.

White, R.A., Mason, D.W., Williams, A.F., Galfre, G., Milstein, C. (1978). T-lymphocyte heterogeneity in the rat : separation of functional subpopulations using a monoclonal antibody. *J. Exp. Med.*, **148**: 664-673.

Whitlow, M. and Filpula, D. (1991). Single-chain Fv proteins and their fusion proteins. *Methods*, **2**: 97-105.

Whitlow, M., Bell, B.A., Feng, S.L., Filpula, D., Hardman, K.D., Hubert, S.L., Rollence, M., Wood, J.F., Schott, M.E., Milenic, D.E., Yokota, T., and Schlom, J. (1993). An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability. *Protein Eng.*, **6**: 989-995.

Williams, A.F., Galfre, G. and Milstein, C. (1977). Analysis of cell surfaces by xenogenic myeloma hybrid antibodies : differentiation antigens of rat lymphocytes. *Cell*, **12**: 663-673.

Williams, W.V., Moss, D.A., Kieber-Emmons, T., Cohen, J.A., Myers, J.N., Weiner, D.B. and Greene, M.I. (1989). Development of biologically active peptide based on antibody structure. *Proc. Natl. Acad. Sci*, **86**: 5537-5541.

Winter, G. and Milstein, C. (1991). Man-made antibodies. *Nature*, **349**: 293-299.

Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. *Ann. Rev. Immunol.*, **12**: 433-455.

Wood, C.R. and Tonegawa, S. (1983). Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation : implications for the joining mechanism. *Proc. Natl. Acad. Sci.*, **80**: 3030-3034.

Wood, C.R., Boss, M.A., Kenten, J.H., Calvert, J.E., Roberts, N.A., and Emtage, J.S. (1985). The synthesis and in vitro assembly of functional antibodies in yeast. Nature, **314**: 446.

Wood, C.R., Dorner, A.J., Morris, G.E., Alderman, E.M., Wilson, D., O'Hara, R.Jr, Kaufman, R.J. (1990). High level synthesis of immunoglobulins in Chinese hamster ovary cells. *J. Immunol.*, **145**: 3011-3016.

Wright, A., Shin, S.U., and Morrison, S.L. (1992). Genetically engineered antibodies : progress and prospects. *Curr. Rev. Immunol.*, **12**:123-168.

Wu, G.I., Zou, D.J., Koothan, T., and Cline, H.T. (1995). Infection of frog neurons with vaccinia virus permits in vivo expression of foreign proteins. *Neuron*, **14**: 681-684.

Wu, T.T., and Kabat, E.A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementary. *J. Exp. Med.*, **132**: 211-250.

Wu, X.C., Ng, S.C., Near, R.I., and Wong, S.L. (1993). Efficient production of a functional single-chain antidigoxin antibody via an engineered *Bacillus subtilis* expression-secretion system. *Bio/Technology*, **11**: 71-76.

Yoshikai, Y., Anatoniou, D., Clark, S.P., Yanagi, Y., Sangster, R., van den Elsen, P., Terhorts, C., and Mark, T.W. (1984). Sequences and expression of transcripts of the human T-cell receptor β-chain genes. *Nature*, **312**:521-524.

Zambryski, P., Joos, H., Gentello, J., Leemans, J., van Montagu, M., Schell, J. (1983). Ti-plasmid vector for introduction of DNA into plant cells without altering their normal regeneration capacity. *EMBO J.*, **2**: 2143-2150.

zu Putlitz, J., Kubasek, W.L., Duchene, M., Marget, M., von Spcht, B.U., and Domdey, H. (1990). Antibody production in baculovirus-infected insect cells. *Bio/Technology*, **8**: 651.