1. REVIEW OF LITERATURE

1.1 Immune system

Immunity is the ability of the body to resist microbial infections and allergic toxins which may sickness and harm to normal body functions. Immune responses involve recognition of pathogens or other foreign materials followed by reactions to eliminate them. The immune system is comprised of cells, tissues and molecules that protect the body from pathogenic microbes in the environment. Protective immunity results from the interplay between two key systems: nonspecific innate immunity and antigen-specific adaptive immunity. These two systems work in concert to provide a high degree of protection for vertebrate species.

1.1.1 Innate immune system

The innate immune system is the first line of defence. It is a nonspecific defence mechanism which consists of cells and proteins that are always present and ready to mobilize and fight microbes at the site of infection within minutes of infection. This natural defence system is conserved from insects to mammals. It encompasses physical barriers such as skin, immune system cells such as leukocytes, dendritic cells, Natural killer cells (NK cells) and circulating plasma proteins.

1.1.2 Adaptive immune system

Adaptive immune system refers to the antigen-specific immune response. It is able to recognise the structures present in the foreign material. Adaptive immunity is believed to have appeared approximately 500 Mya at the beginning of vertebrate evolution (Deng et al., 2013; Pancer and Cooper, 2006). This system is more complex than the innate immune system. It comes into play when innate defences are evaded or overcome by the invading pathogen. Once an antigen has been recognized and processed, the adaptive immune system becomes active and creates an army of immune cells specifically designed to attack that antigen. Adaptive responses are characterized by a latent period of clonal expansion lasting several days before the proliferating lymphocytes mature into effector cells capable of eliminating an infection. There are two types of adaptive immune responses: humoral, mediated by antibodies produced by B lymphocytes (B cells) and cell-mediated immunity, mediated by T lymphocytes (T cells). B cells are derived from the stem cells of the bone marrow and produce antibodies. T cells are produced in the bone marrow but are processed in the thymus. In cell-mediated immunity; antigen presenting cells engulf the antigens, process them internally and display part of the antigens on their surface along with their own proteins. This stimulates T cells to recognise the presented antigens. They secrete lymphokines that

stimulate cytotoxic T cells and B cells to grow and divide, attract neutrophils and boost the ability of macrophages to engulf and destroy the microbes. Cells of immune system make memory cells to respond to the next round of infection by same pathogen. In humoral immunity, immature B-lymphocytes are stimulated to mature when an antigen binds to their surface receptors. These cells make antibodies specific to the antigen and facilitate removal of the antigen (Richard et al., 2000)

1.2 Inflammation

Inflammation is a general form of defence which is a protective reaction of cells/ tissues of the body to allergic or chemical irritation, injury and/or infections. It is broadly defined as a nonspecific response to tissue malfunction and is employed by both innate and adaptive immune systems to combat pathogenic intruders. The symptoms of inflammation are characterized by pain, heat, redness, swelling and loss of function that result from dilation of the blood vessels leading to an increased blood supply and increased infiltration of leukocytes, protein and fluids into the intercellular spaces resulting in inflamed regions (Richard et al., 2000). A number of chemical mediators are released by the cells of innate and adaptive immune system. These mediators, depending on the duration of injury determine the severity of inflammation and are termed pro-inflammatory factors. These

substances bind to specific target receptors on the cells and may increase vascular permeability, promote neutrophil chemotaxis, stimulate smooth muscle contraction, increase direct enzymatic activity, induce pain and/or mediate oxidative damage (Coleman, 2002). Process of inflammation is divided into two parts i.e. acute and chronic which could either be beneficial or detrimental. Acute inflammation is characterized by rapid onset and is of short duration. It is characterised by the exudation of fluids and plasma proteins; and the migration of leukocytes, most notably neutrophils into the injured area. Chronic inflammation is for a prolonged duration and manifests histologically by the presence of lymphocytes and macrophages, resulting in fibrosis and tissue necrosis. Persistent chronic inflammation increases the development of degenerative diseases such as autoimmune disease, heart disease, Alzheimers, asthma, etc. (Iwalewa et al., 2007).

1.3 Autoimmune disease

Autoimmune disease is a condition which is triggered by the disregulation of the immune system initiating an attack on self-molecules. It encompasses more than 100 diseases, including rheumatoid arthritis, Crohn's disease, psoriasis, multiple sclerosis, type-1 diabetes and lupus. 5-8 % of population is affected by these diseases out of it more than 75 % are women. It is one of the top 10 leading causes of death of women above age of 65. The initiation of attacks against the body's

self-molecules in autoimmune diseases, in most cases is unknown, but a number of studies suggest that they are strongly associated with factors such as genetics, infections and /or environment (Vyse and Todd, 1996; Raberg et al., 1998; Smith and Germolec, 1999; Graham et al., 2005; Virginia et al., 2010). Genome-wide association studies for 21 autoimmune diseases from genotyping data have identified loci underlying human diseases. The study involved mapping RNA and chromatin in primary immune cells (Farh et al., 2015). Results indicated that 90% of causal variants are non-coding, with 60% mapping to immune-cell enhancers. Gene expression profiling in rheumatoid arthritis (RA) synovial tissue to osteoarthritis revealed up-regulation in RA of transcripts specific to antigen-presenting cells, T cells and B cells (Devauchelle et al., 2004). Similarly a reduction in the diversity of the intestinal microbiota was found in various autoimmune and allergic diseases in humans (Brown et al., 2011; Nylund et al., 2013; de Goffau et al., 2013; Candon et al., 2015) The gut microbiota of patients with Psoriatic arthritis and patients with skin psoriasis was found to be less diverse when compared to that in healthy controls (Scher et al., 2015). These studies indicate that Gut microbiota also plays an important role in onset of the autoimmune disease.

1.4 Tumor Necrosis Factor α

Tumor Necrosis Factor alpha (TNFα) has been reported more than a century ago (1868), when the German physician Dr P. Brunes noted spontaneous regression of tumours in patients following acute bacterial infections (Old, 1985; Aggarwal, 1987). The first use of the name "tumor necrotizing factor" was initiated in 1962 to describe the regression of a sarcoma tumor induced by the serum of mice treated with *Serratia marcescens* polysaccharide. TNFα is mainly secreted by activated macrophages in response to various inflammatory stimuli, but it is also produced by a wide variety of different cell types like T lymphocytes, NK cells, dendritic cells, endothelial cells, mast cells, Kupffer cells, and many more. Large amounts of soluble TNFα are released in response to lipopolysaccharide and other bacterial products (Aggarwal et al., 2001).

TNFα shares sequence homology (30%) with another peptide hormone, lymphotoxin (LT or TNF-β). A comparison of the amino acid sequences of TNFα and TNFβ revealed that TNFα and TNFβ are structurally homologous and subsequently it was proved that they share a common receptor (Aggarwal et al., 1985). In humans, the gene coding for TNFα is located within the major histocompatibility complex region on chromosome six. The TNF superfamily is composed of 19 ligands and 29 receptors.

Human TNFα is a 25 kDa type II transmembrane proteins of 233 amino acids. This membrane-integrated form is released via proteolytic cleavage by the metalloprotease, TNF alpha converting enzyme (TACE) (Black et al., 1997). Mature soluble human TNF (sTNFα) is a 17kDa non-glycoprotein containing a single disulfide bridge and exists as a homotrimer in aqueous solution (Aggarwal et al., 1985). The crystal structure of TNF revealed that each monomer consists of two antiparallel β -pleated sheets with a jelly roll topology that interact with each other in a head-to-tail fashion to form a heterotrimeric structure (Eck and Sprang, 1989; Jones et al., 1989).

TNFα mediates its action through two distinct receptors, TNF-R1 (TNF receptor type 1; CD120a; p55/60) and TNF-R2 (TNF receptor type 2; CD120b; p75/80) each encoded by a separate gene. Both of them belong to the TNF α receptor superfamily based on the presence of homology in cysteine rich pseudo repeats of approximately 40 amino acids. Any one of the receptors is present on almost all nucleated cells. Extracellular domains of these receptors are homologous but they are distinct in their intracellular domains. Most cells express 1000 - 5000 receptor sites on their surface with an affinity ranging from 0.1 to 1 nM. TNF-R1 is expressed in most tissues and is a key mediator of TNF signaling whereas TNF-R2 is typically found in cells of the immune system. TNF-R2 is activated by

membrane bound TNF, is highly regulated and plays a major role in the lymphoid system (Grell et al., 1995). The extracellular domains of both receptors can be proteolytically cleaved by metalloprotease, yielding soluble receptor fragments which have the capacity of neutralizing sTNFα (Wallach et al., 1991; Ostade et al., 1994).

TNFα is one of the most pleotropic cytokines capable of signalling a large number of physiological and biological responses. It induces at least 5 different types of signals that include activation of NF-kB, apoptosis pathways, extracellular signal regulated kinase (ERK), p38 mitogen-activated protein kinase (p38MAPK) and c-Jun N-terminal kinase (Morgan and Liu, 2010; Aggarwal et al., 2012). Through these signals it mediates a wide spectrum of systemic and cellular responses, including induction of other cytokines and immunoregulatory molecules, fever, shock, tissue injury, tumor necrosis, cell proliferation, differentiation and apoptosis.

TNF is chemotactic to monocytes and neutrophils. Stimulation of these cells with TNF induces phagocytosis, adherence of these cells to endothelial cells and generation of free radicals of oxygen-superoxide anion and hydrogen peroxide. TNF-induced endothelial cell activation leads to the structural reorganization of the endothelium, resulting in vascular leakiness (Giraudo et al., 1998). It

stimulates mononuclear phagocytes and other cells to secrete IL-6 and IL-1 that amplify the inflammatory response, and to secrete other inflammatory cytokines such as IL-8, which leads to migration, degranulation, and respiratory burst response of neutrophils. IL-8 also stimulates the interaction of ICAM-1 and neutrophils, a very important interaction during transvenule migration of the cells. TNF induces the synthesis of other chemokines, such as MCP-1, that promotes accumulation of monocytes at the site of inflammation. In most inflammatory diseases like rheumatoid arthritis, TNF is produced at the inflammatory site (Feldmann et al., 1995), suggesting its role in tissue degradation as well as in increase of inflammatory response (**Figure 1A**) (Aggrawal, 2001).

The physiological action of TNFα depend on the quantity of TNFα secreted. At a low concentration it induces local inflammatory responses such as leukocyte extravasation. At moderate concentration it induces systemic effects such as fever, while at high concentration it affects cardiovascular systems causes hypotension, thrombosis and induces hypoglycaemia etc. These physiological responses result from two major biological activities of TNFα: cytotoxicity and transcriptional activation. Cytotoxicity is responsible for the killing activity of TNFα on many tumor cell lines, while transcriptional activation is responsible for cell

Figure 1 : **Cellular effect and signalling of TNFα in immune/ normal cells**

A: Pro-inflammatory actions of tumour-necrosis factor (TNF) on haematopoietic and non-haematopoietic cells

TNF α can trigger immune responses by inhibiting T-cell receptor signalling, dendritic cell (DC), co-stimulation and promoting lymphoid T-cell apoptosis, inducing other cytokines that inhibit cell-mediated immunity. iNOS, inducible nitric oxide synthase; APCs, antigen-presenting cells; MHC, major histocompatibility complex, $O_2^{\cdot -}$, superoxide; IL, interleukin; TGF-β, transforming growth factor-β. (Adapted from Nature Reviews Immunology, 2002).

B. Autocrine, paracrine, reverse and cell to cell signaling pathways for TNF-α

TNF-α expressed both transmembrane bound and in soluble forms. The transmembrane form appears to mediate therapeutic effects, soluble ligand is linked to pathologic effects of TNF-α.

proliferation and survival, immuno-regulation and inflammatory responses (**Figure 1B**)(May and Ghosh, 1998). TNFα is also involved in remodelling of connective tissue. TNFα induces the release of collagenase and other matrix metalloproteases from fibroblasts and synovial cells. It inhibits collagen synthesis in bone. In explanted cartilage it reduces the content of alkaline phosphatase and induces the resorption of proteoglycan.

Administration of TNFα to various animals leads to haemorrhage, necrosis, local inflammation, shock and death (Broukaert et al., 1992). Subcutaneous and intradermal injection of TNFα in mice or rabbits rapidly leads to neutrophil margination, edema formation, skin necrosis and intense neutrophil infiltration (Dunn et al., 1989; Rampart et al., 1989, Sheehan et al., 1995). LPS is one of the major inducers of TNFα production both *in vitro* and *in vivo*. Administration of high doses of LPS or infection with gram-negative bacteria leads to septic shock that causes shock and tissue injury. Thus, TNFα is required for protection against bacterial, fungal, parasitic and viral infections and other stressful stimuli. There is also evidence that TNFα is involved in tumorogenesis and metastasis (Orosz et al., 1993).

1.5 Antibodies

Antibodies, also called immunoglobulin are immunologically active glycoproteins produced and secreted by differentiated B-lymphocytes upon exposure to potentially hazardous pathogens which in turn leads to elimination and destruction of pathogens. The basic structure of an antibody is shown in **Figure 2**. The general shape of the molecule is believed to be that of a distorted Y or T (supported by the structure of intact murine antibody), display extreme asymmetry, suggesting that antibody molecules are extraordinary flexible, (Harris et al., 1992; Harris et al., 1998; Saphire et al., 2001). Antibodies consist of four polypeptide chain, two identical heavy and two identical light chains which are linked together via disulfide bonds of their cysteine residues.

Both heavy and light chains are highly variable among different antibodies, but the variability is restricted mainly to the functional domains of the amino-terminal ends of the chains, while the remaining domains are constant. The amino terminal variable domains of heavy (V_H) and light chain (V_L) together form the structure responsible for the antigen binding ability of the antibody, while the constant parts (C_H and C_L) mediate the effector functions. The residues of the variable domains can be divided into hypervariable or complementarity determining regions (CDR) and framework regions (FR) based on the observed

sequence variation (Wu & Kabat, 1970). Conserved framework regions provide backbone structure to the variable domains of antibody and can also have an influence on antigenic specificity (Maher, 2002). Three CDRs, in each chain are located between the framework regions, and together these six loops of the heavy and light variable domains form the antigen-binding site of the antibody (Richard, 2001; Susumu, 1983). The antibody diversity is generated during B-lymphocyte differentiation in the bone marrow, where antibody encoding gene segments are randomly shuffled by a regulated genetic system. In combinatorial joining of germline gene segments, two additional phenomena are also active namely, junctional flexibility and N-region nucleotide addition. The result is a mature immunocompetent B-lymphocyte, which contains a single functional DNA sequence for the heavy-chain variable-region (V_H) and one for the light-chain variable-region (V_L). The immune system is capable of generating more than 10^8 different antibody molecules. Upon antigenic stimulation this number is further increased by somatic hypermutation of the complementary determining regions in the V_L and V_H domains (Richard, 2000, Storb et al., 2001). The complete repertoire of germline genes that encode antibodies in both humans as well as other vertebrates have recently also been elucidated and are available on the worldwide web (IMGT, the international ImMunoGene Tics database, http://imgt.cines.fr, Lefranc, 2001).

Figure 2: Structure of antibody molecule

1.6 Platforms for generating monoclonal antibodies

Different platforms have been designed to generate a monoclonal antibody against an antigen. An overview is presented in **Figure 3.** Some of these are described below.

1.6.1 Hybridoma technology

Georges Kohler and Cesar Milstein invented Hybridoma technology for cloning individual antibodies. Monoclonal antibodies (mAbs) are those which originate from clonal population on a single B-lymphocyte (**Figure 3a**). They specifically recognize one epitope of the cognate antigen. Such antibodies are obtained by *in vitro* fusion of activated B-lymphocytes with a myeloma cell line. The fused hybridoma cells are selected using HAT (Hypoxanthine, Aminopterin and Thymidine)-containing medium. Secreted antibodies in conditioned media from each hybridoma are tested to select for a single monoclonal antibody-producing hybridoma cell. The resulting hybrid cells (hybridomas) behave like tumor cells and undergo continuous proliferation (Bartal and Hirshaut, 1987; Kohler, 1975). The antibodies secreted by a clone derived from a single hybridoma cell are monoclonal. Monoclonal antibodies (mAbs) are powerful biological reagents for research, diagnostics and therapy that allow uniformity in assays that would be impossible with polyclonal serum (Good & Lorenz, 1991).

Figure 3: Human antibody techniques

a. Phage display exemplifies human antibody library display techniques (phage, bacteria, yeast, mammalian cell and ribosome). Three steps are included in this technique: antibody library construction and display onto the phage surface, selection by panning the library against antigen (Ag) targets, and screening for desired specificity.

b. Transgenic mouse. The mouse immunoglobulin genes have been genetically knocked out and replaced with human counterparts. The transgenic mouse will make human antibodies after foreign antigen immunization. The B cells harvested from immunized mice are immortalized by fusion with a myeloma cell line, as in traditional hybridoma technology. The hybridomas are then screened for desired specificity.

c. Memory B-cell immortalization. Memory B cells are isolated from peripheral blood mononuclear cells (PBMCs). They are immortalized by EBV in the presence of a CpG oligodeoxynucleotide and irradiated allogeneic PBMCs. The culture supernatants are then screened directly for specific antibodies. Positive cultures are further cloned by limiting dilution and fully human mAbs can then be produced from the cloned B cells.

d. CDR grafting exemplifies humanization. CDR residues from variable region of a mouse mAb are transferred to human antibody frameworks that have high sequence homology with the mouse counterparts. (Figure adopted from Nature Biotechnology, 2007)

1.6.2 Transgenic humanized mice

The technique involves introduction of human immunoglobulin heavy and light chain genes in the germline of mice (**Figure 3b**). Transloci were obtained by plasmid assembly, cloned in yeast artificial chromosomes. Translocus integration and maintenance in transgenic mouse strains is achieved by pronuclear DNA injection into oocytes and various transfection methods using embryonic stem cells. The human DNA segments rearrange faithfully in the mouse and produce extensive V (D) J combinations. Specific human monoclonal antibodies of high affinity for use in therapeutic applications have been produced from these translocus mice (Neuberger, 1997). Human mAbs against severe acute respiratory syndrome coronavirus (SARS-CoV) and rabies with high affinity were generated successfully using this method (Greenough, 2005; Sloan, 2007; Coughlin, 2007). Major limitation in this approach is that immune response in transgenic mice is sometimes less robust than in strains that are used to generate mouse mAbs and hence more immunizations or antibody screens are often required (Marasco & Sui, 2007).

1.6.3 Recombinant antibody technology

Structures and properties of antibodies can be genetically engineered by recombinant DNA techniques (**Figure 3c**). The modular nature and conserved

domain structure of antibodies makes them attractive candidates for protein engineering. The initial recombinant antibodies were made by grafting the variable domains (VL and VH) of murine hybridomas on human constant domains (Boulianne et al., 1984). Later, mouse complementary determining regions were put in a human framework (Jones et al., 1986; Verhoeyen et al., 1988; Queen et al., 1989) and expression of the chimeric antibody genes in mammalian cells yielded humanized antibodies.

Due to the domain structure of antibodies, it is particularly suitable to modify or to produce only the domains of the antibodies. A vast number of structural variations of the antibody fragments have been developed (**Figure 4**). The single-chain variable fragment (scFv) consists of heavy and light chain variable domains of the full immunoglobulin antibody joined using a flexible glycine-serine linker. This small ~30 kDa fragment retains the specificity and affinity of the 150 kDa parental antibody despite the exclusion of the constant domains. The antigen-binding Fab fragment consists of the variable heavy and light chain joined the first fragment of with the constant heavy and light chain genes developing a more stable antibody construct. Sub-fragments such as Fabs (V_HC_{H1} + V_LC_L) or scFvs (V_H + V_L) can be expressed well in *E. coli*. Assembly of functional Fab fragments was observed when V_HC_{H1} and V_LC_L domains are expressed as separate molecules. The key to

expression of functional antibody fragments in *E. coli* appears to be export from the reducing environment of the cytoplasm into the oxidizing environment of the periplasm. In contrast to Fab fragments, Fv fragments showed a tendency to dissociate, particularly at low concentrations. This was overcome by linking the C-terminus of one domain to the N-domain of the other with a flexible linker peptide and expressing V_H and V_L domains as single-chain Fv (scFv) fusion proteins (Pliickthun, 1991; Pliickthun and skerra, 1989; Hoogenboom, 2005).

Recombinant antibody fragments are becoming essential tools in research due to their intrinsic properties such as greater penetrability, ability to maintain antigen recognition, small size and ease of production as compared to complete antibody. In medicine, biotechnology and therapeutics, use of recombinant antibody fragments expressed in *E. coli* cells is becoming increasingly popular (Hoogenboom, 2005).

1.6.4 Phage Display

This technology was first reported to display small peptides on the minor coat protein pIII of bacteriophage Fd (Smith, 1985). It was demonstrated that an antibody scFv gene could be directly incorporated into the phage genome (McCafferty et al., 1990). Phage display involves using a filamentous bacteriophage which has a circular single stranded DNA genome encapsulated within surface

coat protein. Foreign DNA of interest can be inserted into the phage genome and expressed on the surface as fusion coat proteins. These Phages are useful tools to link genotype and phenotype of select recombinant proteins (**Figure 3d**).

The M13 phage replicates in *E. coli,* turning the bacterium into a phage production factory, secreting several hundred phages per cell per division cycle (Ploss, 2010). A filamentous phage does not kill its host bacterium, but rather slows its growth by about half. The M13 phage is made up of 6.4 kb, circular, single-stranded genomic DNA that encodes phage proteins I to XI (**Figure 5**). Five of these proteins are coat proteins. The major coat protein (pVIII) is present in approximately 2800 copies and protects the genome in a cylindrical manner. The minor coat proteins pVII and pIX are necessary for efficient particle assembly, while the minor coat proteins pIII and pVI are necessary for particle stability and infectivity (Hooganhoom, 2002). Infection is initiated by attachment of the phage through the N terminal pIII end to the F pilus of a male *E. coli* cell. This process is necessary for viral uncoating and phage DNA transfer to the cytoplasm of the bacterium. The viral single-stranded circular DNA is replicated by bacterial machinery, and converted into a double-stranded replicative form (Pini and Bracci, 2000). The viral particles are then assembled and extruded through the bacterial envelope (Smith and Petrenko, 1997).

Figure 4: Different format of antibody fragments

(Adapted from Nature Biotechnology, 2005)

Figure 5: Structure of a filamentous phage displaying scFv fragments on its surface

(Adapted from Nature, 1990)

Phage display has proven to be a powerful tool in the development of recombinant antibodies. The technique has two major advantages. First, they can be used to select against an enormous range of antigens with a broad flexibility in the selection conditions that can be adapted to adjust the selection pressure (Ponsel et al., 2011). Second is its, robustness, which gives it a great potential for automation and the high stability of phage is advantageous in terms of applying extreme conditions to the selection procedure (Hoogenboom, 2005;, Ario de Marco 2011; Ahmad et al., 2012].

In the immune system, the B-lymphocytes represent self-replicating packages. They contain the antibody genes that encode the antibody displayed at their surface. Filamentous bacteriophage expressing functional antibody fragments on their surface and containing gene sequence of the same, thus, mimicing the B lymphocyte.

Display of antibody fragments is achieved by fusing the coding sequence of the antibody variable (V) regions as a single-chain Fv (scFv) to the N-terminus of the phage minor coat protein pIII using a phage vector. The scFv sequence is cloned in frame and downstream of gene III signal sequence that directs export of the adsorption protein. In the periplasmic environment, the V_H and V_L domains fold correctly and pair to form a functional scFv (Skerra and Pluckthun, 1988; Better et

al., 1988). Initially, phage vectors that carried all the genetic information required for the phage life cycle were used, but now phagemids have become the most popular vector system for display .These plasmid-based vectors have sequences derived from the intergenic regions of filamentous phage, which enable them to replicate as a single stranded DNA in *E. coli.* The phagemids are packaged into filamentous phage particles (rescued) when cells harbouring them are co-infected with helper phage such as M13K07, that provide all the phage proteins but due to a defective origin they are themselves poorly packaged in competition with the phagemids.

During assembly of M13, the foreign protein is fused to a coat protein and displayed on the surface of the phage. Apart from the minor coat protein III (pIII), major coat protein VIII (pVIII) and the other minor coat proteins of M13 have also been used for recombinant fusions. Phage displays using the pIII and pVIII proteins have different advantages. pVIII protein enables high copy display of the recombinant protein as there are over 2,800 copies of this protein on the surface of the phage. However, a drawback in using the pVIII protein is limitation in the size of the displayed protein. The pVIII protein can only display small peptides less than six amino acids before the function of the coat protein becomes compromised and the number of infectious particles plummets. Recombinant fusions using the pIII protein are not so restricted in the size of the display peptide. The pIII-fusion

protein can display peptides of 100 amino acids or more before the ability of the pIII protein to bind to the F pilus of *E. coli* becomes compromised (Kwasnikowski et al.,2005; Hooganhoom, 2002).

1.7 Combinatorial antibody libraries

Initially, expression of antibody fragments in *E. coli* was achieved by cloning the heavy- and light-chain encoding cDNA sequences from hybridoma cells. Polymerase chain reaction (PCR) provided a means to amplify and clone the VH and VL coding sequences directly from B-lymphocytes (Orlandi et al., 1989; Ward, 1989), thereby omitting the need for the generation of hybridomas. The resulting libraries were combinatorial as the amplified V_H and V_L genes were randomly recombined irrespective of their original pairing. With a large diversity, these libraries are a source of antibodies against a range of different antigens, including self, non-immunogenic and toxic antigens and for this reason these libraries are now extensively used in the industry and in academia (Hoogenboom and Chames, 2000; Hust and Dubel, 2004).

The affinities of the antibody from these libraries are proportional to the size of the library, up to 10 nM for libraries with 10^7 to 10^8 clones and up to 0.1 nM for large repertoire of 10^{10} clones. (Griffiths et al., 1993; Vaughan et al., 1996; Hoet et al., 2005). The most current and successful antibody libraries (natural or synthetic)

display diversity in multiple CDRs and routinely yield single-digit nanomolar and sometimes sub-nanomolar affinity antibodies. These affinities are equal to or better than the affinities of antibodies isolated from immunized mice.

There are no major difference in performance when comparing the best nonimmune and synthetic antibody phage-display libraries, with regard to the frequency of binders and top affinity of selected clones. However, the success of the drug discovery process is influenced by several important differences among libraries, including antibody format and display levels, sequence diversity in selected antibody populations, compatibility with expression screening and with affinity maturation and finally the ease of conversion to other antibody formats, display or selection systems.

1.7.1 Immune library

Repertoires are created using V_H and V_L gene pools that are amplified and cloned from B cells. Source of B cells can be peripheral blood, bone marrow, spleen or tonsils, of immune donors or mice immunized with antigen (Orlandi et al., 1989; Huse et al., 1989). Antibody repertoire will be enriched in antigen-specific antibodies; generally the number of antigen-specific V genes is high, at best < 1/500 (Mullinax et al., 1990) and usually < 1/5000 (Persson et al., 1991). Libraries made with this procedure may form a better reflection of the composition of the

natural immune response compared with random combinatorial libraries with artificially paired chains in immune or non-immune libraries. Such library also facilitates the investigation of humoral immune system at a molecular level (Roovers et al., 2001). These libraries are enriched for rare antibody specificities for the antigen. This method sometimes yields antibodies with higher affinity than obtained from hybridomas, as was reported for an anti-carcinoembryonic antigen (CEA) antibody (Chester et al., 1994). The construction of immune libraries from a variety of species has been reported, including mouse (Kettleborough et al., 1994; Clackson, 1991), human (Barbas et al., 1993, Cai and Garen, 1995), chicken (Yamanaka et al., 1996; Davies et al., 1995), rabbit (Lang et al., 1996) and camel (Ghahroudi et al., 1997). Immune-phage libraries are useful in analyzing natural humoral responses, for example, in patients with autoimmune disease, viral infection, neoplastic diseases, or to study in vitro immunization procedures (Barbas and Burton, 1996; Asra et al., 2012).

1.7.2 Naïve antibody library

The major disadvantage of biased or immune combinatorial antibody libraries, assembled from B lymphocytes of immunized sources, is that for each antigen a new library has to be constructed. Moreover, immunization of humans for this purpose is not ethical and self-antigens are not suited to elicit an immune

response. However, it is possible to bypass immunization and to construct naive combinatorial phage display antibody libraries using V genes of non-immunized sources.

Naïve or non-immune libraries are derived from natural unimmunized, rearranged V genes from the IgM mRNA of B cells. These libraries are useful for selecting antibodies against a wide variety of antigens. These were the first libraries used to isolate anti-self-antibodies; otherwise difficult to obtain by immunization (Griffith et al., 1994).

V genes are amplified from B-cell cDNA using V-gene-family based oligonucleotides and heavy and light chains are randomly combined and cloned to generate a combinatorial library of scFv or Fab antibody fragments. A single naïve library can be used to generate antibodies to a large number of different antigens that include self, non-immunogenic, and toxic (Gram et al., 1992; de Haard et al., 1999, Vaughan et al., 1996,).

1.7.3 Synthetic library

Alternatives for the use of rearranged V repertoires from a non-immunized source are cloned human germ line V_H and V_L sequences. Synthetic antibody libraries are constructed entirely *in vitro* by assembly of V-gene segments and D/J segments

using oligonucleotides that introduce areas of complete or tailored degeneracy into the CDRs of one or more V genes. Synthetic diversity bypasses the natural biases and redundancies of antibody repertoires created *in vivo* and allows control over the genetic makeup of V genes and the introduction of diversity (Winter and Miltein, 1991; Barbas et al., 1992; Hooganhoom et al., 1992; Griffiths et al., 1994). CDR positions that are known to be involved in antigen binding can be identified on structural grounds and randomly shuffled in the context of an antibody with known structure (Sidhu et al., 2004, Silacci et al., 2005). Mostly structural and sequence diversity is found in the CDR3 of the heavy chain so CDR3 has been the target for introduction of diversity in the first synthetic libraries (Nissim et al., 1994). In another design consensus, V-gene segments optimized for expression in *E. coli* and setup for downstream engineering are used in combination with trinucleotide-mediated diversity in the CDR3 regions (Knappik et al., 2000).

Semi-synthetic libraries are made by combinations of natural and synthetic diversity. Such libraries have been created by shuffling natural CDR regions from human B-cells with synthetic CDR1 and CDR2 diversity (Soderlind et al., 2000; Hoet et al., 2005).

1.8 Selection of specific antibodies from combinatorial antibody phage display libraries

Combinatorial antibody phage display libraries encode antibodies of considerable diversity and can theoretically be used to select antibodies against many different antigens. In addition, the phage display system may also prove to be a more efficient method to obtain specific mAbs, as antibody specificity is not biased towards immune-dominant epitopes of antigen.

In the immune system, a large collection of different antibodies is created during maturation of B-cells by the combinatorial assembly of germline-encoded segments. This produces a repertoire of naive B-cell, each expressing a unique antibody on their surface. On exposure to antigen, lymphocytes producing antigen-reactive antibodies are selected and triggered for incorporation of somatic mutations in the V genes. Recombinant antibodies can also be isolated from recombinant antibody libraries in the laboratory, using a selection procedure which mimics this *in vivo* process (**Figure 6**). Phages, encoding displayed antibody fragments, can be affinity selected from combinatorial antibody phage display libraries for binding to particular antigens by passing the phage-antibodies over immobilized antigen (panning). Several formats have been used, such as

Figure 6 : Selection of specific antibodies from combinatorial antibody phage display libraries using biopanning

(Figure adapted from http://utminers.utep.edu)

immobilized antigen on a column matrix, antigen-coated plastic tubes or dishes, biotinylated antigen in solution followed by capture on streptavidin coated beads and whole cells. Phages that are bound to antigen are retained on the surface (McCafferty et al., 1990; Marks et al., 1992; Holt et al., 2000; Hoogenboom, 2005). Non-binding phage antibodies are removed by washing. Bound phage-antibodies can be eluted by a specific elution using acid or alkaline buffers or more specifically using competition with soluble antigen, proteolysis of spacers located between phage and antibody or reduction of disulfide bonds in biotinylated antigens. The selected (eluted) phages are used to infect *E. coli* and can be applied in several sequential rounds of phage growth and selection, allowing even rare phage-antibodies to be isolated (**Figure 6**) (Marks et al., 1991, Clackson et al., 1991; Barbas et al., 1991).This process leads to a diverse collection of recombinant antibody genes, such as those from the B cell of immunized animals. Antibody libraries are enriched by rounds of selection with target antigens and amplification and after a few rounds; individual clones are screened for antigen reactivity (McCafferty and Johnson, 1996; Nissim et al., 1994).

1.8 Screening of selected antibodies

The success of the panning against an antigen can be assessed, using a phage-ELISA for screening the specificity of the selected phage-antibodies. Alternatively,

one of the most useful features in phage display vectors is the presence of an amber stop-codon at the junction between the antibody and the coat protein, which can be used for switching between fused and soluble antibodies. Soluble expression of recombinant antibody fragments can be achieved by changing the bacterial strain. For the selection process the phagemid is expressed in a suppressor *E. coli* strain (XL1-Blue, TG1) which allows for transcription of the pIII fusion protein along with the antibody gene. By incorporating the phage into a non-suppressor *E. coli* strain (HB2151) the amber stop codon is read with soluble expression of the recombinant antibody fragment in the periplasmic space without the pIII gene product. Recombinant antibodies can be expressed with fusion tags (c-myc and His) for downstream detection and purification. Because most recombinant antibodies have a peptide tag fused to the C-terminus, the reaction towards antigen can also be determined in ELISA (Lee et al, 2007; Christoph et al., 2014).

1.9 Yeast Display and affinity maturation

For the isolation of antibodies, phage display is a very powerful technology but the isolated antibodies often require affinity optimisation prior to clinical development, as improved binding efficiency correlates with improved clinical efficacy. Recombinant antibodies can be manipulated by techniques such as

affinity maturation (Schier et al., 1996), chain shuffling, error-prone PCR and site-directed mutagenesis in order to increase affinity and sensitivity (Hoogenboom, 2005). Yeast surface display is a powerful platform for engineering proteins by directed evolution. Selection process of yeast display system utilizes flow cytometry sorting which allowed visualizing binding to the target antigen to its antibody during each selection round. Stringency steps used for the selection in phage display system can also be applied in yeast display system. With FACS sorting changes in stringency can also be made "on the fly" by setting the cell sort gate based on the antigen binding fluorescence. One more advantage compared to the phage system, is that it allowed isolation of loss-of-binding populations as selection of improved clones was less predictable using phage with immobilised antigen. In addition, yeast display also allows analysis of individual scFv clones without having pure proteins (scFv). The yeast display system uses the *S. cerevisiae* α-agglutinin receptor to display heterologous proteins on the cell

Figure 7: Yeast surface display

A single-chain antibody is expressed in yeast as a fusion to the yeast agglutinin Aga2p. Self assembly with mating agglutinin Aga1p yields about 50-100k copies of surface displayed scFv per yeast cell, with each cell expressing a single antibody clone.

surface (Boder and Wittrup, 1997). Antibodies are expressed in a scFv format, which is fused to the adhesion subunit of the yeast agglutinin protein Aga2p, which can bind to the Aga1p by disulfide bonds to attach to the yeast cell walls (**Figure 7**) (Chao et al., 2006). A library of antibody fragments mutagenized using any of the standard technique can easily be expressed on the surface of the yeast. The mutant library generated is screened using the labelled antigen on FACS. The dual color FACS yeast antibody display is generally used for both expression and binding studies. The expressed cells are labelled with anti-c-Myc IgY/ Alexa Fluor-488-conjugated goat anti-chicken IgG for expression detection and biotinylated antigen followed by streptavidin-phycoerythrin (PE) for antigen binding. Cells positive for both scFv expression and antigen binding are analysed (Boder and Wittrup, 1997; Chao et al., 2006).

Affinity maturation of antibody fragments and other proteins has been carried out using yeast surface display technology. Example of protein include T cell receptors (Kieke et al., 1999) and carcino-embryonic antigen (Graff et al., 2004) etc. A small library of scFv mutants was screened for fluorescein and a 3-fold reduced off-rate was obtained (Boder and Wittrup, 1997). Similarly scFv specific for T-cell receptor was affinity matured to have a 3-fold higher affinity (Kieke et al., 1997). One of the strongest affinity antibody fragment (scFv) reported against fluorescein with

an equilibrium dissociation constant (K_D) of 48 fM, was affinity matured by using this technique (Boder et al., 2000).

1.10 Anti TNFα Therapies

The evaluation of cytokine expression in human rheumatoid synovial fluid was a key experiment in defining that cytokines were overexpressed at the disease site (Feldmann and Maini, 2001) and they contribute significantly to inflammation and articular destruction. TNFα blockade reduced the level of other pro inflammatory cytokines which validated TNFα as a therapeutic target for rheumatoid arthritis (RA). Later on study of the pathophysiology of other autoimmune diseases confirmed that TNFα is closely related to the pathologies of many autoimmune diseases, such as rheumatoid arthritis, ulcerative colitis, psoriasis and Crohn's disease (Aggarwal et al., 2012).

Biological therapy with TNFα inhibitors has radically changed the treatment of many inflammatory diseases, and is an integral part of disease management. Since clinical trials results have been very consistent, with all the TNFα inhibitors tested being efficacious (Feldmann and Maini, 2001). Clinical efficacies of anti TNFα drugs in majority of patients have demonstrated that the therapy prevented the irreversible structural damage characteristic of rheumatoid arthritis (RA) and other rheumatic diseases.

There are five FDA approved anti-TNFα drugs. These drugs have been licenced for a variety of diseased conditions like rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis, and psoriasis.

Infliximab

Infliximab is a chimeric (human/ murine) IgG1 monoclonal antibody directed against TNF α. This is the first anti TNFα drug that reached to routine clinical use. It is administered through the intravenous route. Licenced dose of Infliximab is 3 mg /kg at week 0, 2 and 6 followed by 3 mg/kg-1 every 8 weeks. MTX (methotrexate) co-therapy is mandatory and doses are increased up to 7.5 mg/kg for patient with inadequate response (Merk, 2011).

Adalimumab

It is a fully humanized IgG1 monoclonal antibody that inhibits TNF α. It can be administrated with or without MTX. The PREMIER study showed that Adalimumab is most effective when used in combination with MTX, provide a better clinical response and reduced radiographic progression (Breedveld, 2006). Licenced dose of Adalimumab is 40 mg, administered subcutaneously every 2 weeks, the mean terminal half-life of this drug is 2 weeks. (Abbott Laboratories Ltd, 2011)

Etanercept

Etanercept is a recombinant fusion protein that consists of TNF receptor (p75) linked to the Fc portion of human IgG1. It can be prescribed as monotherapy or with MTX and is administered subcutaneously every weak. Its mean elimination half-life is around 70h. (Wheyth pharmaceuticals, 2011)

Golimumab

It is a human IgG1 kappa monoclonal antibody against TNFα. It is also licenced for used with MTX. Golimumab is administered at a dose of 50 mg monthly by subcutaneous injection. Its half-life is approximately 2 weeks (Taylor, 2010)

Certolizumab

It is a recombinant antibody Fab fragment conjugated to polyethylene glycol. PEGylation delays elimination by reducing renal clearance, decreasing proteolysis. It can be administered with or without MTX (Mease, 2011). The licenced dose of Certolizomab is 400 mg at week 0, 2 and 4 followed by 200 mg fortnightly. It has a half-life of 14 days (UCB Pharma Ltd, 2011). It is the only anti TNFα biologic produced in bacteria which improves yield and could reduce production cost.

These drugs are generally well tolerated and serious adverse events are rare. However, their increasing use has revealed a number of immune-mediated adverse effects such as opportunistic infections, tuberculosis, autoimmunity and many more.

Formation of anti-drug antibodies (ADA) has been associated with all five agents. In case of Adalimumab and Infliximab, immunogenicity is linked to the serum drug levels. Clinically data demonstrated that Adalimumab provided the highest rate of treatment response and disease remission, and Etanercept was the best tolerated drug overall with longest drug survival rate (Hetland et al., 2010).

In vitro and animal studies indicate that Adalimumab, Infliximab, and Etanercept have similar binding properties for both soluble TNF and membrane bound TNF (Kaymakcalan et al., 2009). But clinical data from the patients showed that these anti-TNFα drugs behave differently in patients and in different diseases. As Etanercept therapy is not effective in Crohn's disease and sarcoidosis compared to the monoclonal anti TNFα antibodies (Sandborn, 2001). Bifluoresence imaging in mice with collagen induced arthritis to assess the distribution of labelled Certolizumab, Infliximab and Adalimumab in healthy and inflamed murine tissue at multiple time points up to 26h post administration (Palframan, 2009). All three agents preferentially distributed to the inflamed tissue but Certolizumab

penetrated more effectively into arthritic paws and remained for a longer period of time (Palframan, 2009).

Similarly RA patients treated with either Adalimumab or Infliximab have a 7-17 fold higher risk of developing a Mycobacterium tuberculosis (TB) infection compared to RA patients treated with etanercept (Sandborn, 2001; Utz et al., 2003; Doty, 2005; Papadakis, 2009; Tubach, 2009). These data suggest that these drugs act differently in RA patients.

1.11 Protein Expression Systems

1.11.1 *Pichia pastoris*

The ability of certain yeast species to utilize methanol was described first by Koichi Ogata (Ogata et al., 1969). Methylotrophic yeasts belonging to the four genera: *Hansenula, Pichia, Candida* and *Torulopsis*, are capable of growth on methanol as sole carbon and energy source (Faber et al., 1995). Due to their ability to grow to high cell densities (>130 g l^{-1} dry cell weight) on methanol in continuous cultures, they were initially considered to be good candidates as a potential source of single-cell protein (SCP). Subsequently, with the increasing availability of the relative genetic elements this limited scope has been extended to develop several of these methylotrophic species into powerful expression systems.

One of the most successful examples is the development of *Pichia* expression systems. Researchers at Salk Institute Biotechnology/Industrial Associates, Inc. (SIBIA, La Jolla, CA) isolated the gene and promoter for alcohol oxidase, and generated series of vectors, host strains, and corresponding protocols for the molecular genetic manipulation of *P. pastoris*. The combination of strong and tightly regulated alcohol oxidase promoter along with the optimized fermentation methods developed for the SCP process resulted in surprisingly high levels of foreign protein expression in *P. pastoris*.

1.11.1.1 Methanol Metabolism in Yeasts

A specific methanol utilization pathway involving several unique enzymes exists in all the methylotrophic yeasts. A key feature that distinguishes methylotrophic yeasts from bacteria is the reaction mechanism used to oxidize methanol to formaldehyde, catalyzed by an oxygen-dependent oxidase as the electron acceptor (Veenhuis et al., 1983; Veenhuis and Harder, 1987). Whereas, in bacteria this reaction is catalyzed by a dehydrogenase linked to the electron transport chain.

The key steps involved in the methanol metabolism are outlined in the Figure 3. The enzymes that catalyze the first three steps of methanol metabolism are localized in the matrix of single membrane bound organelles, called peroxisomes. The peroxisomes are indispensable during growth on methanol, as they harbour

alcohol oxidase, catalase and dihydroxyacetone synthase, the key enzymes involved in methanol metabolism. Besides this, peroxisomes also help in protecting the cell from the toxicity of the by-products (such as hydrogen peroxide) generated during methanol metabolism; by sequestering them away from the rest of the cell. The enzyme alcohol oxidase (*AOX*; EC 1.1.3.13) catalyzes the first step in methanol metabolism, the oxidation of methanol to formaldehyde using oxygen as an electron acceptor.

The hydrogen peroxide generated in this reaction is converted to oxygen and water by the peroxisomal catalase (*CAT*). A portion of the formaldehyde generated by AOX leaves the peroxisome and is oxidized subsequently to formate and carbon dioxide by two cytoplasmic dehydrogenases, the formaldehyde dehydrogenase (*FLD1*) and formate dehydrogenase (*FMD1*), respectively. These reactions represent the source of energy for cells growing on methanol. The remaining formaldehyde is assimilated to form cellular constituents by a cyclic pathway that starts with the condensation of formaldehyde with xylulose 5-monophosphate, a reaction catalyzed by third peroxisomal enzyme dihydroxyacetone synthase (*DHAS*). The products of this reaction, glyceraldehyde 3-phosphate and dihydroxyacetone, leave the peroxisome and enter a cytoplasmic pathway that

regenerates xylulose 5-monophosphate. One molecule of glyceraldehyde 3-phosphate is generated for every three cycles.

In addition to methanol, the *AOX* from methylotrophic yeast also oxidizes other lower primary aliphatic alcohols, so it has been named as alcohol oxidase. The enzyme functions as an octamer of identical subunits of ~74 kDa, each containing one non-covalently bound flavin adenine dinucleotide moiety. Alcohol oxidase has a poor affinity for oxygen, and *P. pastoris* compensates for this by synthesizing large amounts of this enzyme. In methylotrophic yeasts, AOX levels are induced dramatically in presence of methanol and account for >30 % of the total soluble protein (Couderc and Baratti, 1980). *P. pastoris* has two alcohol oxidase genes, namely *AOX1* and *AOX2*. Protein coding regions of the genes are closely are closely homologous, 92% at the nucleotide sequence level and 97% at the amino acid sequence levels, whereas the DNA sequences of the promoter region share no homology. Although both genes encode equally functional alcohol oxidase enzymes, *AOX1* is responsible for majority of alcohol oxidase activity in the cell, due to the relative strength of its promoter (Ellis et al., 1985; Cregg et al., 1989). In methanol-grown cells, ~5% of polyA$^+$ RNA is from *AOX1*; however, in cells grown on most other carbon sources, *AOX1* message is undetectable (Cregg and Madden, 1988). Consequently, the *AOX1* promoter, which regulates 85% of the alcohol

oxidase activity in the cell, is the most commonly utilized promoter for controlling heterologous gene expression in *P. pastoris*.

1.11.1.2 Regulation of Methanol Metabolism in Yeast

Regulation of methanol metabolism in yeast is a very complex process including the control of synthesis and activation of the corresponding enzymes as well as their degradation. The enzymes involved in the methanol metabolism (*AOX*, *CAT* and *DHAS*) are present at high levels in cells grown on methanol, but are not detectable in cells grown on most other carbon sources (e.g., glucose, glycerol, or ethanol). Synthesis of these enzymes is regulated at the level of transcription; it is induced in presence of methanol and repressed by glucose and ethanol. The addition of glucose or ethanol to methanol-grown cells results in not only repression of methanol metabolizing enzymes but also inactivation of these enzymes. The regulation of the *AOX1* gene appears to involve two mechanisms: a repression/depression mechanism; plus an induction mechanism, similar to the regulation of several of the *Saccharomyces cerevisiae GAL* genes. The presence of methanol appears to be essential to induce high levels of *AOX1* transcription (Tschopp et al., 1987). However, unlike *GAL*-pathway regulation, the absence of a repressing carbon source, such as glucose in the medium, does not result in substantial transcription of *AOX1*.

1.11.1.3 Methanol utilization phenotypes in *P. pastoris*

P. pastoris strains exhibit three different phenotypes based on their ability to utilize methanol. The strains with Mut^+, or methanol utilization plus phenotype, like the wild type exhibit a high methanol utilization rate and grow well on methanol. These strains require high feeding rates of methanol in large-scale fermentations (Cereghino and Cregg, 2000). The strains with Mut^s or methanol utilization slow phenotype have a disrupted or a non-functional *AOX1* gene. Since the cells must then rely on the weaker *AOX2* for methanol metabolism, a slower growing and slower methanol utilization strain is produced. One of the advantages of this phenotype is that low growth rates may be desirable for production of certain recombinant products (Cregg et al., 1987; Cregg and Madden, 1988). The Mut^- strains, or methanol utilization minus phenotype, are unable to grow on methanol, due to deletion in both *AOX* genes. Currently, the majority of researchers use the Mut^+ phenotype (Hohenblum et al., 2004; Slibinskas et al., 2004), although some researchers are also using the Mut^s phenotype (Paramsivam et al., 2002; Aoki et al., 2003; McKinney et al., 2004; Yang et al., 2004).

AOX1 promoter

Expression of the *AOX1* gene is tightly regulated and induced by methanol to high levels, typically >30% of the total soluble protein in cells grown with methanol as

the carbon source. In methanol-grown cells the *AOX1* represents approximately 5% of the polyA$^+$ RNA. The *AOX1* promoter has been widely used to drive expression of the gene of interest in *P. pastoris* (Ellis et al., 1985; Tschopp et al., 1987; Koutz et al., 1989). Most expression vectors have an expression cassette composed of a 0.9-kb fragment from *AOX1* composed of the 5' promoter sequences and a second short *AOX1*-derived fragment with sequences required for transcription termination (Koutz et al., 1989). Expression of foreign gene under the control of *AOX1* promoter often results in a high level of production. In addition, for secretion of foreign proteins, vectors are available where in-frame fusions of foreign proteins and the secretion signals of *P. pastoris* acid phosphatase (*PHO1*) or *S. cerevisiae* α-mating factor (α-MF) can be generated.

Integration of expression vectors in *P. pastoris* genome

Stability of recombinant *Pichia* strain is high due to integration of expression vectors in to the *P. pastoris* genome (Romanos et al., 1992). Expression vector are directed to integrate into the *Pichia* genome in one of the two ways, depending on where the vector DNA is cut before transformation. Linearization of the vector, by cutting either 5' side of AOX1 promoter or within the HIS4 marker, direct integration of the plasmid occurs at the homologous sites in the genome. In *Pichia pastoris* GS115 it gives rise to Mut^s phenotype with integration at *AOX1* or *HIS4*,

multicopy transformants (up to 10) can arise from repeated recombinant events (Cereghino and Cregg, 2000). For generation of Mut^{s} strain, *P. pastoris* expression vector are digested in such a way that the expression cassette and marker gene are released, flanked by 5' and 3' *AOX1* sequence. Approximately, 10-20% of transformation events are the result of a gene replacement event in which the *AOX1* gene is deleted and replaced by the expression cassette and marker gene. This disruption of *AOX1* gene forces these strain to rely on the transcriptionally weaker *AOX2* gene for growth on methanol (Cregg and Maddan, 1987).

A strain that contain multiple integrated copies of an expression cassette can sometimes yield more recombinant protein than single copy strain (Clare et al., 1991). Optimization of protein expression therefore often includes the isolation of multicopy expression strain. In a detail study of expression of tetanus toxin fragment C in *P. pastoris*, it was shown that expression was directly correlated with copy number (1-14 copies), whereas site of integration and Mut phenotype had, only a minor effect on yield (Clare et al., 1991). In numerous reports, the isolation of multicopy integrants has resulted in dramatically higher yields of protein production (sreekrishna et al., 1989; Romanos et al., 1991).

Secretion signal selection

Since Pichia secretes very low level of endogenous proteins, the secreted recombinant protein constitutes a vast majority of total protein in the medium. Several different signal sequences, including the native secretion signal present on heterologous proteins, have been used successfully for the secretion of foreign protein in *Pichia pastoris* (Leonardo, 2012). The *S. cerevisiae* α factor prepro-peptide has been used with the best results. This signal sequence consist of a 19 amino acid signal (pre) sequence followed by a 66 residue (pro) sequence containing three consensus *N*- linked glycosylation sites and the dibasic Kex2 endopeptidase processing site (Kurjan and Herskowitz, 1982). In some cases, standard α MF or *PHO1* secretion signal have not worked, so synthetic leaders have been created. Martinez-Ruiz and coworkers made mutations in the native leader sequence to reconstruct a more efficient Kex2p recognition motif (Lys-Arg). This aided in secretion of the ribosome-inactivation protein, α-sarcin from *Aspergillus giganteus* (Martinez et al., 1998). Another solution for secretion was to create an entirely synthetic prepro leader. For the expression of human insulin, a synthetic leader and spacer was found to improve secretion and protein yield (Kjeldsen et al., 1999).

1.11.1.4 *Pichia* as a Host for Heterologous Protein Production

The methylotrophic yeast *Pichia pastoris* has proved to be an outstanding host for heterologous gene expression, due to high expression levels of heterologous proteins. Over four hundred heterologous proteins of commercial interest have been successfully produced using the *P. pastoris* host vector system and over 500 reports describing its use have been published (http://faculty.kgi.edu/cregg/index.htm). Most of them rely on the ability of *AOX1* promoter to induce high levels of gene expression in presence of methanol. Like other yeast vectors, all the *Pichia* vectors are *E. coli/P. pastoris* shuttle vectors, containing an origin of replication for plasmid maintenance in *E. coli* and selectable markers. Some of the commercialized proteins produced in *Pichia* are: Phytase (Phytex, Sheridan, IN, USA), Trypsin (Roche Applied Science, Germany), nitrate reductase (The Nitrate Elimination Co., Lake Linden, MI, USA), phospholipase C (Verenium, San Diego, CA, USA/DSM, The Netherlands), Collagen (Fibrogen, San Francisco, CA, USA).

Kalbitor® (Ecallantide), is a plasma kallikrein inhibitor indicated against hereditary angioedema, is produced in *P. pastoris* by Dyax (Cambridge, MA, USA). This product was the first biopharmaceutical to be approved by the FDA for market release in 2009 (Walsh 2010). Indian company Biocon produces recombinant human insulin and analogues thereof (Insulin, Glargine) (Gonçalves

et al., 2013). *P. pastoris* has the ability to produce proteins of therapeutic and commercial interest in concentrations ranging from milligrams to grams per liter (Macauley-Patrick et al., 2005). The first reported expression of fully functional antibody produced in *P. Pastoris* (Ogunijimi et al., 1999). Engineered Pichia was optimised for the glycosylation of the glycoproteins with uniform human N-glycans (Li et al., 2008; Barnard et al., 2010). GlycoSwitch and GlycoFi technologies have successfully produced glycoengineered antibody up to 1.6 g/l (Potgieter, 2009; Ye, 2011). During the last two decades, *P. pastoris* has been proven to be a powerful candidate for high level expression of functional antibody fragments with reports of yields ranging from 10 mg up to 4.88 gram per liter of culture (Fischer et al., 1999; Freyre et al., 2000; Damasceno et al., 2004; Khatri and Hoffmann, 2006).

Considering the merits of scfv for antibody generation and value of therapeutic anti TNFα in a much effective and economic manner, this study has focused on isolation of the efficient clones of TNFα scFv from the Tomlinson scFv phage display libraries for the development of novel hTNFα binding antibodies, and cost-effective expression of anti TNF scFv (Recab) in non conventional yeast *P. pastoris*.