PROCESS FOR PURIFYING PROTEINS

This invention relates to a process for purifying a recombinant protein from a gram-negative host cell sample or extract thereof. More particularly, the gram-negative host cell sample or extract thereof is adjusted to a low pH to effect the precipitation of protein disulphide isomerase.

Recombinant DNA techniques have rapidly developed and are useful in the production of antibodies, in particular therapeutic antibodies. Systems for the expression of recombinant genes are well known to the person skilled in the field in question. These include expression in mammalian cells, insect cells, fungal cells, bacterial cells and transgenic animals and plants. The choice of expression system is dependent on the features of the encoded protein, for example post-translational modifications. Other considerations include the time and, in particular, the cost involved in the production of the desired quantity of material of the required quality. These latter considerations are particularly important in the production of therapeutic antibodies of the quality required for regulatory approval and in the quantities needed for treatment of large numbers of patients.

A widely used system for the production of recombinant proteins is based on expression in *Escherichia coli* (*E. coli*). A specific problem encountered with the use of *E. coli* is the difficulty in producing material of the required quality in quantities needed for therapy. In particular, the time and costs involved can be prohibitive. One specific problem of note is the loss incurred in the yield of antibodies during purification of the antibodies from *E. coli*.

Although, proportionally, the purification costs are a fraction of the total cost of a therapeutic antibody product, the purification cost proportion will increase further as upstream production costs become cheaper. Thus, improvements in recovery and purification of antibodies will drive production costs down further irrespective of the means of production (Humphreys & Glover, Curr. Opin. Drug Discovery & Development, 2001, 4:172-185). Hence, there is a need for methods that introduce time and/or cost savings into therapeutic antibody production and, in particular, in purification, for example by increasing product recovery and/or improving the quality of the product stream.

Low product yield is often a particular problem noted during the downstream purification steps. A specific problem in downstream purification concerns the separation of host cell proteins (HCP) which are released upon extraction of the therapeutic antibody from the host cell. A number of process modifications have been tried in order to address this problem. Examples of such processes include the following:

The protein of interest may be precipitated and then separated from other HCP contaminants. However, the precipitation of a therapeutic antibody may cause irreversible damage to the antibody. Techniques that specifically precipitate the protein of interest often result in the entrapment of the non-proteinaceous contaminants in the precipitate, rendering the separation ineffective.

Alternatively, the HCP may be precipitated away from the therapeutic antibody. US7169908 describes the addition of a solution of ethacridine lactate to precipitate host cell impurities. However, the use of an agent such as ethacridine lactate is not suitable for the production of therapeutic proteins because it has been used as an abortion agent and may be harmful to the patient if it is not completely removed. The HCPs from a CHO cell culture have been precipitated using a small molecule and pH adjustment in order to purify a protein (Arunakumari A. et al. Advances in Non-Protein A Purification Processes for Human Monoclonal Antibodies Supplement to BioPharm International March 2009 p 22-26).

Whilst such methods have contributed to helping to purifying therapeutic antibodies from host cell contaminants there is still a need to provide improved methods to reduce the quantity of HCPs from microbial host cell systems without negatively affecting the properties of the therapeutic antibody in order to facilitate further downstream processing.

What is more the inventors have found that the expression of certain chaperone proteins, for example a disulfide isomerase, such as DsbC, by the host can be useful in increasing the levels of expression of antibodies or fragments thereof. However, the chaperone protein is expressed at significant levels and becomes a large by-product of the process that requires removal..

Removing large amounts of contaminant can be challenging because methods such as column chromatography find it difficult to remove large amounts of a protein

contaminant due to column fowling and the need to use large amounts of reagents or solvents.

When isolated DsbC is tested it is soluble at low pHs such as pH 3. However, the present inventors have surprisingly found that pH adjustment of 5 or less may be used to precipitate recombinant disulphide isomerase from a gram-negative bacterial host cell and thereby facilitate further processing of a recombinant protein of interest, such as a therapeutic antibody.

Summary of the Invention

In one aspect there is provided a method for purifying a recombinant protein from a host cell, for example a gram-negative bacterial host cell sample expressing a recombinant protein and a recombinant disulphide isomerase or extract thereof; comprising:

a. adjusting the pH of the host cell sample or extract thereof to a pH of 5 or less to precipitate the recombinant disulphide isomerase; and
b. separating precipitated recombinant disulphide isomerase from the recombinant protein to provide a recombinant protein sample.

The present invention also provides the use of a step of adjusting the pH of a gram-negative bacterial host cell sample transformed with an expression vector encoding a recombinant protein or extract thereof to a pH of 5 or less to precipitate host cell recombinant disulphide isomerase, followed by separation of the precipitated host cell recombinant disulphide isomerase from the recombinant protein and provide a purified recombinant protein sample.

In one embodiment there is provided a method for purifying a recombinant antibody or a binding fragment thereof from a gram-negative bacterial host cell sample or an extract of said host cell wherein said host cells expresses the recombinant antibody or binding fragment thereof and a recombinant disulphide isomerase DsbC; wherein the method comprises:

a. adjusting the pH of the host cell sample or extract thereof to a pH in the range 3.5 to 5 to precipitate the recombinant DsbC; and
b. separating precipitated recombinant DsbC to provide a sample containing a purified recombinant antibody or fragment thereof.

Surprisingly this pH adjustment step precipitates a significant quantity of the host cell recombinant disulphide isomerase from the solution and thereby allows separation of the host cell recombinant disulphide isomerase from the recombinant protein of interest. This facilitates downstream processing, particularly any subsequent chromatography steps such as non-affinity chromatography capture, by reducing the host cell disulphide isomerase in the solution competing with the recombinant protein for binding sites on the chromatography column. Therefore, the load of HCPs on the chromatography column is minimized and yield of the recombinant protein of interest is improved when the same size and number of chromatography columns are used. This reduces the time and cost of further downstream purification.

Interestingly pure DsbC does not precipitate at a pH in the range 3.0 to 5 (see Figure 4). However, when present in the complex environment of a gram-negative cell expressing an antibody or a binding fragment thereof or alternatively in the presence of part or all of the contents of the said cells then the DsbC precipitates. Whilst not wishing to be bound by theory it is hypothesized that the complex matrix environment of the host cell or extract therefrom catalyzes the precipitation of the disulphide isomerase, such as DsbC, in the pH range even though normally the protein is soluble at a pH in the range 3.5 to 5. Perhaps once the precipitation is initiated it continues and escalates.

Thus, the method according to the present disclosure is suitable for use on large scales and therefore is very useful in commercial processing of proteins such as antibodies.

Brief Description of the Drawings

Figure 1 shows the observed precipitation of the host cell solution after the pH has been adjusted to pH 5.0, pH 4.5 or pH 3.0 compared to the control pH of 6.9.

Figure 2 shows the chromatogram of a reverse phase HPLC analysis at Time (T) = 0 for host cell solutions after pH adjustment to pH5, pH4.5 or pH3 compared to the control pH of 7.

Figure 3 shows the separate chromatograms of a reverse phase HPLC analysis at T=0 for host cell solutions after pH adjustment to pH5, pH4.5 or pH3 compared to the control pH of 7.

Figure 4 shows the separate chromatograms of a reverse phase HPLC analysis for a solution comprising DsbC after pH adjustment to pH3 after different time intervals compared to the control pH of 7.

Figure 5 shows the total peak area for DsbC from a reverse phase HPLC analysis at various time points for host cell solutions after pH adjustment to pH5, pH4.5 or pH3 compared to the control pH of 7.

Figure 6 shows the total peak area for DBP from a reverse phase HPLC analysis at various time points for host cell solutions after pH adjustment to pH5, pH4.5 or pH3 compared to the control pH of 7.

Figure 7 shows the SDS-PAGE Analysis Gel for host cell solutions after pH adjustment to pH5, pH4.5 or pH3 compared to the control pH of 6.9.

Figure 8 shows sequences SEQ ID NOs: 1 to 16.

Figure 9 shows the structure of a compound called CDP870 comprising a modified anti-TNF Fab' fragment covalently linked via a cysteine residue to a lysyl-maleimide linker wherein each amino group on the lysyl residue has covalently attached to it a methoxy PEG residue wherein n is about 420.

Brief Description of the Sequences

SEQ ID NO: 1 shows the amino acid sequence of CDRH1 of CDP870.

SEQ ID NO: 2 shows the amino acid sequence of CDRH2 of CDP870.

SEQ ID NO: 3 shows the amino acid sequence of CDRH3 of CDP870.

SEQ ID NO: 4 shows the amino acid sequence of CDRL1 of CDP870.

SEQ ID NO: 5 shows the amino acid sequence of CDRL2 of CDP870.

SEQ ID NO: 6 shows the amino acid sequence of CDRL3 of CDP870.

SEQ ID NO: 7 shows the nucleotide and predicted amino acid sequence of the light chain variable region CDP870.

SEQ ID NO: 8 shows the nucleotide and predicted amino acid sequence of the heavy chain variable region CDP870.

SEQ ID NO: 9 shows the amino acid sequence of a grafted anti-TNFα Fab CDP870 light chain.

SEQ ID NO: 10 shows the amino acid sequence of a grafted anti-TNFα Fab CDP870 heavy chain.

SEQ ID NO: 11 is the nucleotide sequence of his-tagged DsbC.

SEQ ID NO: 12 is the amino acid sequence of his-tagged DsbC.

SEQ ID NO: 13 is the sequence of the wild-type spr gene including the signal sequence which is the first 26 amino acid residues.

SEQ ID NO: 14 is the sequence of the wild-type spr gene without the signal sequence.

SEQ ID NO: 15 is the DNA sequence of a mutated knockout Tsp gene including the 6 nucleotides ATGAAT upstream of the start codon.

SEQ ID NO: 16 shows the oligonucleotide cassette encoding intergenic sequence 2 (IGS2) for *E. coli* Fab expression.

Detailed Description of the Preferred Embodiments of the Invention

The present invention will now be described in more detail.

The terms "protein" and "polypeptide" are used interchangeably herein, unless the context indicates otherwise. "Peptide" is intended to refer to 10 or less amino acids.

The terms "polynucleotide" includes DNA, cDNA, RNA and mRNA.

As used herein, the term "comprising" in context of the present specification should be interpreted as "including".

The expressions "cell", "cell line", "cell culture" and "strain" are used interchangeably.

Host cell extract (extract therefrom) as employed herein is used to refer to part or all of the host cell contents i.e.

- material obtained from a partial or total cell lysis to extract some or all chemical matter from the cells, or
- material obtained from subjecting the cells to processing such that one or more of the proteins expressed from the cells are released into the liquid phase.

Chemical matter and material that form the contents of the host cell as employed herein refer to liquids, proteins, lipids, sugars, lipopolysaccharides, peptidoglycans, plasmid and chromosomal DNA's, RNA's, small molecules such as amino acids, metal ions, redox active molecules, tRNA's and fragments of all of the above from within the cell.

Methods for releasing one or more proteins from the cells include heat treatment and/or subjecting the cells non-lysing pressure and/or use of chemicals including: buffers, chelating agents, detergents, physical disruption and/or use sonic energy and/or use of mechanical shear.

Gram-negative cell sample as employed herein refers to a population of gram-negative cells, for example at least two gram-negative cells but more specifically a batch employed in a fermentation process, in particular a commercial fermentation process.

Purified recombinant antibody or binding fragment thereof as employed herein is intended to refer to a form of the antibody or fragment wherein there are less impurities and contaminates than in the corresponding unprocessed form. It is not necessarily intended to be an absolute term because the antibody or fragment may still require further purification steps.

In one embodiment a purified recombinant antibody or binding fragment thereof provided by the method according to the present invention is substantially pure.

Substantially pure as employed herein refers to the where the antibody or binding fragment thereof is great than 90% pure, for example 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% pure, in particular 95% pure or greater.

This invention described herein is based on the surprising and unexpected observation that after a host cell sample transformed with an expression vector encoding a recombinant protein molecule has been cultured a step of adjusting the pH to 5 or less may be used to precipitate recombinant disulphide isomerase and that this has a significant beneficial impact on the purification of the protein.

Step a) of the present invention requires the adjustment of the host cell sample or extract thereof to a pH of 5 or less, a pH of less than 5, a pH of 4.5 or less or a pH of 3 or less. In a preferred embodiment, the pH of the host cell sample or extract thereof is adjusted to a pH of 4.5 or less, more preferably a pH of approximately 4.5. The inventors have found that a pH of 4.5 is particularly advantageous to precipitate and allow separation of recombinant DsbC.

In one embodiment the pH of the host cell sample or extract thereof is not adjusted to a pH which is less than pH 3 or alternatively less than pH 3.5. Accordingly, the pH of the host cell sample or extract thereof is preferably adjusted to a pH of 3 to 4.9, a pH of 3.5 to 4.9, a pH of 3 to 4.5 or a pH of 3.5 to 4.5.

In one embodiment the pH of the host cell sample or extract thereof is adjusted to a pH of 3 or less, more preferably a pH of approximately 3. The inventors have found that a pH of 3 of less is advantageous to precipitate and allow separation of recombinant DsbC and also host cell dipeptide binding protein.

However, whilst pHs of 3 or less are suitable for precipitating the DsbC protein from the host cell or extract low pHs may alter the folded conformation of the antibody or binding fragment thereof and thus may not be optimal when all parameters of the process are considered.

Thus the optimal pH for precipitating the host cell proteins will depend on the specific antibody or antibody fragment expressed by the cell and in particular the stability of the same in the relevant pH range.

At least one disulfide isomerase, for example at least DsbC is precipitated employing the process. In one embodiment one or more other host cell proteins are also precipitated employing the process.

In step a) of the method of the present invention the host cell solution or extract thereof may be held at the lower pH, for example a pH of 5 or less for any suitable period of time before carrying out any further purification steps, such as centrifugation and/or filtation, and/or before a further pH adjustment typically to raise pH of the solution. Generally the host cell solution or extract thereof will be held for 24 hours or less, 12 hours or less, 6 hours or less, 5 hours or less, 2 hours or less, 1 hour or less, 30 minutes or less, 10 minutes or less or 7 minutes or less.

After employing the method of the invention the pH may, for example be raised to pH 6, 6.5 or 7 before the further i.e. subsequence processing such as further purification, for example purification employing anion exchange chromatography is performed.

Step a) of the method of the present invention may be performed at any suitable temperature, for example room temperature, such as 20-23°C or reduced temperature such 1-10°C. Thus suitable temperature ranges for performing the method include at least 1 to 30°C.

The adjustment of the pH of the host cell sample or extract thereof may be performed using any suitable agent capable of changing the pH. Examples of suitable agents are glacial acetic acid, sodium hydroxide, sodium acetate or tris base, and combination thereof. The agent may be at any suitable concentration such as 30 or 60%

(v/v) glacial acetic acid, 1M sodium hydroxide, 1M sodium acetate, and 2M or 3M tris base.

One or more of these agents are particularly advantageous for use in the method according to the present disclosure and, for example enabling precipitation and removal of the recombinant disulphide isomerase in the purification of a therapeutic protein because they are non-toxic.

Disulphide isomerase is an enzyme that catalyzes the formation and breakage of disulphide bonds between cysteine residues within proteins as they fold. Protein disulphide isomerase is released from the host cell upon extraction of the recombinant protein. The host cell used in the method of the present invention produces recombinant disulphide isomerase and, therefore, the recombinant disulphide isomerase constitutes a significant proportion of contaminating host cell protein. The present invention has provided a time and cost-saving means to remove the recombinant disulphide isomerase from the recombinant protein. The recombinant disulphide isomerase, such as DsbC, may comprise a histidine tag (his-tag) at the N- and/or C-terminus.

The histidine tag is advantageous because it allows effective monitoring to show that DsbC contamination has been removed from the antibody or binding fragment thereof. This is important because the presence of DsbC can be difficult to detect when using polyclonal sera raised against the cellular contents from wild type *E. coli* because standard polyclonal sera tend to be poorly reactive or non-reactive to DsbC, and so are unable to detect DsbC even when it is overexpressed at high levels. The presence of a detection tag such as poly-His ensures that the removal of abundantly overexpressed DsbC can be monitored/confirmed using sensitive immunodetection methods.

In a preferred embodiment, the recombinant disulphide isomerase is DsbC. DsbC is a prokaryotic protein found in the periplasm of *E.coli* which catalyzes the formation of disulphide bonds in *E.coli*. DsbC has an amino acid sequence length of 236 (including signal peptide) and a molecular weight of 25.6 KDa (UniProt No. P0AEG6). DsbC was first identified in 1994 (Missiakas et al. The Escherichia coli dsbC (xprA) gene encodes a periplasmic protein involved in disulfide bond formation, The EMBO Journal vol 13, no 8, p2013-2020, 1994 and Shevchik et al. Characterization of DsbC, a periplasmic protein of Erwinia chrysanthemi and *Escherichia coli* with disulfide isomerase activity, The EMBO Jounral vol 13, no 8, p2007-2012, 1994). The DsbC protein may comprise a histidine tag

(his-tag) at the N- and/or C-terminus.

The predicted pI of DsbC is 5.7 and the predicted pI of his-tagged DsbC is 6.3. The pI of a protein can be predicted using a variety of commercially available and free to public softwares and on-line pI prediction facilities such as ExPASy ProtParam.

Proteins may be expected to precipitate if the pH of the solution passes through the pI of the protein. However, it can be seen from Figure 4 that a solution of DsbC (his-tagged) free of host cell or host cell extract when adjusted to pH 3 does not precipitate because the quantity of DsbC is not reduced after the pH adjustment. Surprisingly, it has been found that DsbC does precipitate when a host cell sample or extract thereof comprising recombinant DsbC is subjected to a pH adjustment to less than pH 5 a significant amount of the DsbC is precipitated and can be separated from a recombinant protein of interest, as shown in Figure 3. Accordingly, it is not possible to predict from the predicted pI of a protein, such as DsbC, what pH is required to allow precipitation from a host cell sample or extract thereof.

As used herein, a "recombinant polypeptide" refers to a protein that is constructed or produced using recombinant DNA technology. The polynucleotide sequence encoding disulphide isomerase may be identical to an endogenous sequence encoding disulphide isomerase found in bacterial cells. Alternatively, the recombinant polynucleotide sequence encoding the disulphide isomerase is a mutated version of the wild-type disulphide isomerase sequence, for example having a restriction site removed. In the embodiment wherein the disulphide isomerase is DsbC, the restriction site EcoRI may be removed and/or a sequence encoding a his-tag added. An example modified DsbC nucleotide sequence for use in the present invention is shown in SEQ ID NO: 11, which encodes the his-tagged DsbC amino acid sequence shown in SEQ ID NO: 12.

In one embodiment the mutant DsbC consists of a DsbC protein with a mutated active site represented by –CXXC- wherein XX is TF, GF, HH, NY, SF, MF, VH, SH, RF, FA, GA, MA, GI, AV, PS, QA, SV, PR, PP, AL, PL, FL, TR, LL, VL, QL, LQ.

Prior to step a) of the method of the present invention, the method may comprise culturing a host cell sample transformed with one or more expression vectors encoding a recombinant protein and a recombinant disulphide isomerase. The sample may be at any suitable scale from small-scale production of protein to large-scale manufacture of protein for commercial purposes.

In one embodiment, wherein the protein is an antibody, the recombinant antibodies produced from the host cell may be a mixture of functional and non-functional antibodies.

The host cell comprises a recombinant polynucleotide encoding a disulphide isomerase which may be present on a suitable expression vector transformed into the cell and/or integrated into the host cell's genome. Preferably the polynucleotide encoding the disulphide isomerase is in an expression vector in the cell thereby causing minimal disruption to the host cell's genome. The recombinant protein and the recombinant disulphide isomerase may be present on the same expression vector or on separate expression vectors.

The cells used in the present invention are gram-negative bacteria. Most preferably, the cells are *E. coli*. The cells are recombinant cells which have been genetically engineered. *E. coli* host cells are mutated strains capable of producing recombinant proteins. The recombinant *E. coli* host cells may be derived from any suitable *E. coli* strains include MC4100, TG1, TG2, DHB4, DH5α, DH1, BL21, K12, XL1Blue and JM109. One example is *E. coli* W3110 (ATCC 27,325) a commonly used host strain for recombinant protein fermentations. Examples also include modified *E. coli* strains, for example metabolic mutants and protease deficient strains.

The host cell is genetically engineered to produce recombinant disulphide isomerase. Host cells which produce recombinant disulphide isomerase are particularly advantageous because the presence of the recombinant disulphide isomerase may reduce cell lysis and may assist in the processing of the recombinant protein.

In a preferred embodiment the host cell comprises a polynucleotide sequence encoding recombinant DsbC, including his-tagged DsbC.

The host cell according to the present invention may comprise one or more further genetic modifications.

In one embodiment, the host cell may have reduced protease activity, wherein the cell comprises a mutated Tsp gene encoding a Tsp protein having reduced protease activity or is a knockout mutated Tsp gene.

For example, the host cell may have reduced Tsp protein activity compared to a wild-type cell. Tsp (also known as Prc) is a 60kDa periplasmic protease. The first known substrate of Tsp was Penicillin-binding protein-3 (PBP3) (Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3 of *Escherichia coli*;

Nagasawa H, Sakagami Y, Suzuki A, Suzuki H, Hara H, Hirota Y. J Bacteriol. 1989 Nov;171(11):5890-3 and Cloning, mapping and characterization of the Escherichia coli Tsp gene which is involved in C-terminal processing of penicillin-binding protein 3; Hara H, Yamamoto Y, Higashitani A, Suzuki H, Nishimura Y. J Bacteriol. 1991 Aug;173 (15):4799-813) but it was later discovered that the Tsp was also able to cleave phage tail proteins and, therefore, it was renamed as Tail Specific Protease (Tsp) (Silber et al., Proc. Natl. Acad. Sci. USA, 89: 295-299 (1992)).

In this embodiment the cell has reduced Tsp protein activity compared to a wild-type cell. The expression "reduced Tsp protein activity compared to a wild-type cell" means that the Tsp activity of the cell is reduced compared to the Tsp activity of a wild-type cell. The cell may be modified by any suitable means to reduce the activity of Tsp. In one embodiment the reduced Tsp activity is from modification of the endogenous polynucleotide encoding Tsp and/or associated regulatory expression sequences. The modification may reduce or stop Tsp gene transcription and translation or may provide an expressed Tsp protein having reduced protease activity compared to the wild-type Tsp protein. In one embodiment an associated regulatory expression sequence is modified to reduce Tsp expression. For example, the promoter for the Tsp gene may be mutated to prevent expression of the gene. In a preferred embodiment the cell according to the present invention carries a mutated Tsp gene encoding a Tsp protein having reduced protease activity or a knockout mutated Tsp gene. Preferably the chromosomal Tsp gene is mutated.

The reduction of Tsp (prc) activity is desirable to reduce the proteolysis of proteins of interest. However, it was found that cells lacking protease prc show thermosensitive growth at low osmolarity. Hara *et al* isolated thermoresistant revertants containing extragenic suppressor (spr) mutations (Hara et al., Microbial Drug Resistance, 2: 63-72 (1996)). Spr is an 18kDa membrane bound periplasmic protease and the substrates of spr are Tsp and peptidoglycans in the outer membrane involved in cell wall hydrolysis during cell division. The spr gene is designated as UniProtKB/Swiss-Prot P0AFV4 (SPR_ECOLI). Improved protease deficient strains comprising mutant spr gene have been described. Chen et al (Chen C, Snedecor B, Nishihara JC, Joly JC, McFarland N, Andersen DC, Battersby JE, Champion KM. Biotechnol Bioeng. 2004 Mar 5;85(5):463-74) describes the construction of *E. coli* strains carrying different combinations of

mutations in prc (Tsp) and another protease, DegP, created by amplifying the upstream and downstream regions of the gene and ligating these together on a vector comprising selection markers and a sprW174R mutation (High-level accumulation of a recombinant antibody fragment in the periplasm of *Escherichia coli* requires a triple-mutant (ΔDegP Δprc sprW174R) host strain.

The wild-type amino acid sequence of the spr protein is shown in SEQ ID NO:13 with the signal sequence at the N-terminus and in SEQ ID NO:14 without the signal sequence of 26 amino acids (according to UniProt Accession Number P0AFV4). The amino acid numbering of the spr protein sequence in the present invention includes the signal sequence. Accordingly, the amino acid 1 of the spr protein is the first amino acid (Met) shown in SEQ ID NO: 13.

In a further embodiment the cell comprises a mutated spr gene. The mutant spr gene encoding a spr protein may have a mutation at one or more amino acids selected from C94, S95, V98, Y115, D133, V135, H145, G147, H157 and W174. Preferably the mutant spr gene encodes a spr protein having a mutation at one or more amino acids selected from C94, S95, V98, Y115, D133, V135, H145, G147 and H157. Preferably, the mutant spr gene encodes a spr protein having a mutation at one or more amino acids selected from S95, V98, Y115, D133, V135 and G147. The spr mutations are capable of suppressing the growth phenotype of a cell comprising a mutated Tsp gene.

One or more of the amino acids C94, S95, V98, Y115, D133, V135, H145, G147, H157 and W174 may be mutated to any suitable amino acid which results in a spr protein capable of suppressing the phenotype of a cell comprising a mutated Tsp gene. For example, one or more of S95, V98, Y115, D133 and V135 may be mutated to a small amino acid such as Gly or Ala. In a preferred embodiment the spr protein comprises one or more of the following mutations: C94A, S95F, V98E, Y115F, D133A, V135D or G, H145A, G147C and H157A.

In a further embodiment the mutated spr gene encodes a spr protein having the mutation W174R. In an alternative embodiment the spr protein does not have the mutation W174R.

The designation for a substitution mutant herein consists of a letter followed by a number followed by a letter. The first letter designates the amino acid in the wild-type protein. The number refers to the amino acid position where the amino acid substitution is

being made, and the second letter designates the amino acid that is used to replace the wild-type amino acid.

Accordingly, in a preferred embodiment the cell used in the present invention comprises a recombinant polynucleotide encoding DsbC and a mutated spr gene, as defined above.

In a further preferred embodiment the cell used in the present invention has reduced Tsp protein activity compared to a wild-type cell and comprises a recombinant polynucleotide encoding DsbC and a mutated spr gene, as defined above.

In one embodiment the host cell comprises a mutated DegP gene encoding a DegP protein having chaperone activity and reduced protease activity and/or a mutated ptr gene, wherein the mutated ptr gene encodes a Protease III protein having reduced protease activity or is a knockout mutated ptr gene and/or a mutated OmpT gene, wherein the mutated OmpT gene encodes an OmpT protein having reduced protease activity or is a knockout mutated OmpT gene.

In one embodiment the host cell expresses one or more proteins as follows:

- one or more proteins capable of facilitating protein folding, such as FkpA, Skp, SurA, PPiA and PPiD; and/or
- one or more protein capable of facilitating protein secretion or translocation, such as SecY, SecE, SecG, SecYEG, SecA, SecB, FtsY and Lep; and/or
- one or more proteins capable of facilitating disulphide bond formation, such as DsbA, DsbB, DsbD, DsbG.

One of more of the above proteins may be integrated into the cell's genome and/or inserted in an expression vector.

In one embodiment the host cell does not comprise recombinant polynucleotide encoding one or more of the following proteins:

- one or more proteins capable of facilitating protein folding, such as FkpA, Skp, SurA, PPiA and PPiD;
- one or more protein capable of facilitating protein secretion or translocation, such as SecY, SecE, SecG, SecYEG, SecA, SecB, FtsY and Lep; and
- one or more proteins capable of facilitating disulphide bond formation, such as DsbA, DsbB, DsbD, DsbG.

In one embodiment the cell used in the present invention has reduced Tsp protein activity compared to a wild-type cell and comprises a recombinant polynucleotide encoding DsbC and a mutated spr gene, as defined above comprises a recombinant polynucleotide encoding DsbC and a mutated spr gene, as defined above, and FkpA and/or Skp.

The recombinant protein produced using the methods of the present invention is typically expressed in either the periplasm of the *E. coli* host cell or in the host cell culture supernatant, depending on the nature of the protein and the scale of production. The methods for targeting proteins to these compartments are well known in the art, for a review see Makrides, Microbiological Reviews, 1996, 60, 512-538. Examples of suitable signal sequences to direct proteins to the periplasm of *E. coli* include the *E. coli* PhoA, OmpA, OmpT, LamB and OmpF signal sequences. Proteins may be targeted to the supernatant by relying on the natural secretory pathways or by the induction of limited leakage of the outer membrane to cause protein secretion examples of which are the use of the *pelB* leader, the protein A leader, the coexpression of bacteriocin release protein, the mitomycin-induced bacteriocin release protein along with the addition of glycine to the culture medium and the coexpression of the *kil* gene for membrane permeabilization. Most preferably, in the methods of the invention, the recombinant protein is expressed in the periplasm of the host *E. coli*.

Expression of the recombinant protein in the *E. coli* host cells may also be under the control of an inducible system, whereby the expression of the recombinant protein in *E. coli* is under the control of an inducible promoter. Many inducible promoters suitable for use in *E. coli* are well known in the art and depending on the promoter, expression of the recombinant protein can be induced by varying factors such as temperature or the concentration of a particular substance in the growth medium (Baneyx, Current Opinion in Biotechnology, 1999, 10:411-421; Goldstein and Doi, 1995, Biotechnol.Annu.Rev, 105-128). Examples of inducible promoters include the *E.coli lac*, *tac*, and *trc* promoters which are inducible with lactose or the non-hydrolyzable lactose analog, isopropyl-β-D-1-thiogalactopyranoside (IPTG) and the *pho*A, *trp* and *ara*BAD promoters which are induced by phosphate, tryptophan and L-arabinose respectively. Expression may be induced by, for example, the addition of an inducer or a change in temperature where induction is temperature dependent. Where induction of recombinant protein expression is

achieved by the addition of an inducer to the culture the inducer may be added by any suitable method depending on the fermentation system and the inducer, for example, by single or multiple shot additions or by a gradual addition of inducer through a feed. It will be appreciated that there may be a delay between the addition of the inducer and the actual induction of protein expression for example where the inducer is lactose there may be a delay before induction of protein expression occurs while any pre-existing carbon source is utilized before lactose.

E. coli host cell cultures (fermentations) may be cultured in any medium that will support the growth of *E. coli* and expression of the recombinant protein. The medium may be any chemically defined medium, such as those provided in Pirt S.J. (1975) Principles of Microbe and Cell Cultivation, Blackwell Scientific Publications, with modifications where appropriate to control growth rate as described herein. An example of a suitable medium is 'SM6E' as described by Humphreys *et al.*, 2002, Protein Expression and Purification, 26:309-320.

Culturing of the *E. coli* host cells can take place in any suitable container such as a shake flask or a fermenter depending on the scale of production required. Various large scale fermenters are available with a capacity of more than 1,000 litres up to about 100,000 litres. Preferably, fermenters of 1,000 to 50,000 litres are used, more preferably 1,000 to 10,000 or 12,000 litres. Smaller scale fermenters may also be used with a capacity of between 0.5 and 1,000 litres.

Fermentation of *E. coli* may be performed in any suitable system, for example continuous, batch or fed-batch mode (Thiry & Cingolani, 2002, Trends in Biotechnology, 20:103-105) depending on the protein and the yields required. Batch mode may be used with shot additions of nutrients or inducers where required. Alternatively, a fed-batch culture may be used and the cultures grown in batch mode pre-induction at the maximum specific growth rate that can be sustained using the nutrients initially present in the fermenter and one or more nutrient feed regimes used to control the growth rate until fermentation is complete. Fed-batch mode may also be used pre-induction to control the metabolism of the *E. coli* host cells and to allow higher cell densities to be reached (Lee, 1996, Tibtech, 14:98-105).

If desired, the host cells may be subject to collection from the fermentation medium, *e.g.* host cells may be collected from the sample by centrifugation, filtration or

by concentration. In particular, the methods of the invention are suitable for the large-scale industrial manufacture of antibodies of therapeutic quality.

In one embodiment the method according to the present invention comprises a step of centrifugation after the step of culturing, followed by suspension of the cells by addition of the extraction buffer.

Following the step of culturing, the method may comprise a step of extraction to release the recombinant protein from the host cell. The extraction may be performed by any suitable method such as cell lysis by mechanical or pressure treatment, freeze-thaw treatment, osmotic shock, extraction agents or heat treatment. Such extraction methods are well-known in the art.

In a preferred embodiment an extraction buffer is added to the sample and the sample is then subjected to a heat treatment step. The heat treatment step is preferably as described in detail in US 5,665,866 (the contents of which are incorporated herein by reference). The heat treatment step makes it possible to obtain a sample of soluble, correctly folded and assembled antibody by facilitating the removal of other antibody-related material. Antibody which is "correctly folded and assembled" is shown by the presence of a single band corresponding to the expected molecular weight for assembled heavy and light chains on non-reducing SDS PAGE. Other antibody related material will typically be free heavy and light chain or part thereof, partially degraded fragments of correctly folded and assembled antibody.

The heat treatment step is performed by subjecting the sample to a desired elevated temperature. Most preferably, heat treatment step is performed within the range of 30°C to 70°C. The temperature can be selected as desired and may depend on the stability of the protein for purification. In another embodiment, the temperature is within the range 40°C to 65°C, or preferably within the range 40°C to 60°C, more preferably within the range 45°C to 60°C, even more preferably within the range 50°C to 60°Cand most preferably at 55°C to 60°C, 58°C to 60°C or 59°C. Thus, the minimum temperatures are 30°C, 35°C or 40°C and the maximum temperatures 60°C, 65°C or 70°C.

The heat treatment step is preferably carried out for a prolonged period of time. The length of heat treatment is preferably between 1 and 24 hours, more preferably between 4 and 18 hours, even more preferably between 6 and 16 hours and most preferably between 10 and 14 hours or between 10 and 12 hours, for example 12 hours.

Thus, the minimum time for heat treatment is 1, 2 or 3 hours and the maximum is 20, 22 or 24 hours.

In a particular embodiment, the heat treatment is performed at 50°C to 60°C for 10 to 16 hours, and more preferably at 59°C for 10 to 12 hours. One skilled in the art will understand that temperatures and time can be selected as suits the sample in question and the characteristics of the protein being produced.

Following the step of extraction the sample may be subjected to a step of centrifugation and/or filtration prior to the step of adjusting the pH.

The sample which is subjected to step a) of the present invention is a host cell sample or extract thereof. Accordingly, step a) may, for example be carried out on:

- a solution comprising the population of host cells expressing the recombinant protein and recombinant disulphide isomerase;
- the supernatant of the population host cells expressing the recombinant protein and recombinant disulphide isomerase following centrifugation and/or filtration to remove the host cells;
- the extract of the population of host cells expressing the recombinant protein and recombinant disulphide isomerase following a step of extraction, such as heat treatment; or
- the extract of the population of host cells expressing the recombinant protein and recombinant disulphide isomerase following a step of extraction, such as heat treatment, and one or more subsequent steps of centrifugation and/or filtration.

Prior to step a), the host cell solution or extract thereof is at a suitable pH, typically the pH of the host cell solution or extract thereof is pH 5 to 10, preferably pH 6 to 8 or approximately pH 7. The pH of the host cell solution or extract thereof prior to step a) may depend on the pI of the recombinant protein of interest. Proteins may have a tendency to precipitate if the pH is adjusted to or through the pI of the protein. Therefore, it is preferable that the pH of the solution prior to the pH adjustment step is lower than the pI of the recombinant protein to ensure that the pH adjustment step does not bring the pH of the solution to or through the pI of the recombinant protein. For example, if the pI of the recombinant protein is around 8, the pH of the solution is preferably less than pH 8 in

order to minimize precipitation of the recombinant protein. The step of lowering the pH to less than pH 5 will also preferably avoid the pI of the recombinant protein.

The pI of a protein will depend upon the ionic strength of the solution it is in. Accordingly, the calculated pI of a protein in water or very low ionic strength buffer may be different to the pI of a protein in a host cell sample or extract thereof comprising buffer and HCPs. The pI of a protein may be 1, 1.5 or 2 pH units lower than the measured pI of the protein in water or very low ionic strength buffer. Therefore, the pH of the host cell solution or extract thereof is preferably kept at a pH of 1, 1.5 or 2 pH units below the pI of the recombinant protein of interest measured in water or very low ionic strength buffer.

In step b) of the method according to the present invention, the precipitated recombinant disulphide isomerase is separated from the recombinant protein to produce a recombinant protein sample. Step b) typically comprises centrifugation and/or filtration in order to separate the precipitated recombinant disulphide isomerase.

The resulting recombinant protein sample has a reduced quantity of recombinant disulphide isomerase compared to the quantity of recombinant disulphide isomerase prior to step a). Preferably the recombinant protein sample after step b) is substantially free of recombinant disulphide isomerase. However, it is possible that some recombinant disulphide isomerase will remain in the sample, which may be removed by later purification steps. The skilled person can determine if recombinant disulphide isomerase is still present in the recombinant protein sample using methods well known in the art. For example, an ELISA analysis may be used to detect the recombinant disulphide isomerase, such as DsbC. An antibody specific for the recombinant disulphide isomerase, such as DsbC, may be used or if the recombinant disulphide isomerase is his-tagged, such as DsbC-his-tag, an anti-His-tag antibody may be used for detection of the recombinant disulphide isomerase.

Substantially free of, as employed herein is intended to refer to containing 5%w/w or less, such as 4, 3, 2, 1 or 0.5% w/w or less.

In one embodiment, where the pH of the host cell sample or extract thereof is adjusted to a pH of 3 or less (such as pH 3), the host cell protein dipeptide binding protein (also referred to as "DBP") is precipitated and separated from the recombinant protein in step b). In this embodiment the resulting recombinant protein sample also has reduced

quantity of host cell dipeptide binding protein (DBP). Preferably the resulting recombinant protein sample is substantially free of DBP.

Preferably the resulting recombinant protein sample also has reduced quantity of other host cell proteins which may also precipitate during step a). Preferably the resulting recombinant protein sample is substantially free of other host cell proteins.

The yield of the recombinant protein of interest following steps a) and b) of the present invention is typically 75% or more, 80% or more, 85% or more or 90% or more.

After step b), the pH of the recombinant protein sample may be adjusted to a suitable pH for either storing the recombinant protein or for performing a downstream purification step, such as chromatography. As discussed previously, the pH of the solution can be adjusted to minimize precipitation of the recombinant protein of interest by ensuring that the pH is not at the pI of the recombinant protein. Preferably the pH of the recombinant protein sample is adjusted to a pH of 5 to 7, more preferably a pH of 5 to 6. The exact pH required will depend upon the properties of the recombinant protein, including the pI of the protein, and which downstream purification steps are to be carried out. Any suitable agent may be used to adjust the pH of the recombinant protein sample, such as a base selected from NaOH, Na Acetate or Tris base.

Alternatively, the method does not comprise a step of pH adjustment after step b). In this embodiment, the pH of the recombinant protein sample may be suitable for storage of the recombinant protein and/or for one or more subsequent downstream purification steps. In one embodiment, in step a) the pH of the host cell sample or extract thereof is adjusted to a pH of 4.5, step b) comprises centrifugation and filtration and the resulting recombinant protein solution after step b) is then subjected to a cation exchange chromatography step at pH 4.5.

Alternatively a chromatography step may be performed at higher pH, for example pH or above, such as 5 to 8, in particular 6 or 7.

Following step b), the recombinant protein sample may be subjected to one or more further purification steps in order to remove contaminants and/or undesired protein fragments such as antibody fragments. Typically the recombinant protein sample is subjected to one or more steps of chromatography, wherein each step of chromatography may be followed by a step of filtration, such as microfiltration, diafiltration or ultrafiltration. The one or more steps of chromatograph may be affinity or non-affinity

chromatography steps. In one embodiment the chromatograph is a non-affinity chromatography step such as cation or anion ion-exchange chromatography.

In one embodiment, in step a) the pH of the host cell sample or extract thereof is adjusted to a pH of 4.5, step b) comprises centrifugation and filtration and the resulting recombinant protein solution after step b) is then subjected to a cation exchange chromatography step at pH 4.5.

The method of the present invention is particularly advantageous when non-affinity chromatography steps are used in the downstream purification. This is because a large quantity of host cell proteins, including recombinant disulphide isomerase, may bind to the chromatography columns if they have not been previously separated. Accordingly, the capacity of the chromatography column to bind the desired recombinant protein is reduced. The separation of host cell proteins, including recombinant disulphide isomerase increases the capacity of subsequent non-affinity chromatography columns which provides a faster and cheaper means for purification of the recombinant protein.

During step a) of the present invention, the recombinant protein may partially unfold whilst the solution is held at the pH of 5 or less. However, partial unfolding of the protein is reversible and following the separation of the precipitated disulphide isomerase, the pH of the protein solution may be adjusted to a pH above 5, preferably a pH of 5 to 7, or pH of 5 to 6, and the protein may then adopt it's original folded conformation. Thus overall the low pH holding step provides relatively gentle conditions for removing impurities, with minimal long term consequence. The recombinant protein sample provided by the method of the present invention comprises a significant proportion of functional protein.

The recombinant protein of interest expressed by the host cells, preferably has a pI of 6 or higher, more preferably pI of 7 or higher. As discussed previously, if the protein has a pI of 6 or higher is it easier to minimize precipitation of the protein by keeping the pH of the sample below the pI of the protein, preferably 1, 1.5 or 2 pH units below the pI of the protein in water or very low ionic strength buffer, and the step of adjusting the pH of the sample to pH 5 or less does not adjust the pH of the solution to the pI of the protein.

The recombinant protein of interest is preferably a recombinant antibody. The recombinant antibody preferably has a pI of 6 to 10, 7 to 10, 6 to 9, 7 to 9, 8 to 9, 6, 7, 8 or 9. The pH of the sample is preferably kept below the pI of the recombinant antibody,

more preferably 1, 1.5 or 2 pH units below the pI of the recombinant antibody, before, during and after step a) in order to ensure minimal precipitation of the antibody.

In a preferred embodiment, the recombinant antibody has a pI of 7-9 and the pH of the sample is adjusted to pH 3 to 4.5, more preferably pH 4.5, in step a).

A specific example of a recombinant antibody is an anti-TNF Fab' CDP870, as described herein and in WO01/094585, which has a pI of 8.

In one embodiment the recombinant antibody or binding fragment thereof is anti-TNF Fab' CDP870.

As used herein, 'functional antibody' includes antibody molecules that retain the ability to specifically recognise or bind to the antigen against which they were raised (cognate antigen) i.e. the antigen to which it is specific. The production of a functional antibody is shown by the presence of a single band on non-reducing SDS-PAGE corresponding to the expected molecular weight of the antibody, or by direct binding assay using BIAcore or other methods known to the person skilled in the art, for example but not limited to, ELISA. Non-functional antibodies include fragments which do not recognise their cognate antigen, and include incorrectly-folded or incorrectly-assembled antibodies, free heavy and light chains, and fragments thereof, including partially degraded fragments of antibodies which do not recognise or bind to their cognate antigen.

A binding fragment of an antibody is a fragment that is able to bind the antigen to which the antibody is specific.

In a preferred example, the recombinant antibody molecule is at least part of an antibody light chain and at least part of an antibody heavy chain, such that at least some of the expressed light and heavy chain antibody molecules are able to combine to form functional antibody.

As used herein, 'antibodies' include antibodies having full length heavy and light chains; functionally active fragments, binding fragments, derivatives or analogues thereof and may be, but are not limited to VH, VL, VHH, Fab, modified Fab, an altered hinge Fab, Fab', $F(ab')_2$ or Fv fragment; a light chain or heavy chain monomer or dimer; a single chain antibody, e.g. a single chain Fv in which the heavy and light chain variable domains are joined by a peptide linker, Fab-Fv, or a dual specificity antibody, such as a Fab-dAb, as described in PCT/GB2008/003331.

The antibodies may be polyclonal, monoclonal, bi-, tri- or tetra-valent antibodies, humanized or chimeric antibodies. These antibodies and their fragments may be naturally occurring, humanized, chimeric or CDR grafted antibodies and standard molecular biology techniques may be used to modify, add or delete amino acids or domains as desired. Humanized antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule (see, for example, US 5,585,089). The antibody molecules purified using the methods of the invention can be of any class (*e.g.* IgG, IgE, IgM, IgD and IgA) or subclass of immunoglobulin molecule.

The methods for creating these antibody molecules are well known in the art (see for example, Shrader et al., WO 92/02551; Ward et al., 1989, Nature, 341:544; Orlandi et al., 1989, Proc.Natl.Acad.Sci. USA, 86:3833; Riechmann *et al.*, 1988, Nature, 322:323; Bird et al, 1988, Science, 242:423; Queen et al., US 5,585,089; Adair, WO91/09967; Mountain and Adair, 1992, Biotechnol. Genet. Eng. Rev, 10:1-142; Verma *et al.*, 1998, Journal of Immunological Methods, 216:165-181).

Monoclonal antibodies may be prepared by any method known in the art such as the hybridoma technique (Kohler & Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, 1983, Immunology Today, 4:72) and the EBV-hybridoma technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, pp77-96, Alan R Liss, Inc., 1985).

Chimeric antibodies are those antibodies encoded by immunoglobulin genes that have been genetically engineered so that the light and heavy chain genes are composed of immunoglobulin gene segments belonging to different species. These chimeric antibodies are likely to be less antigenic. Bivalent antibodies may be made by methods known in the art (Milstein *et al.*, 1983, Nature 305:537-539; WO 93/08829, Traunecker *et al.*, 1991, EMBO J. 10:3655-3659). Bi-, tri- and tetra-valent antibodies may comprise multiple specificities or may be monospecific (see for example WO 92/22853).

Antibody sequences may also be generated using single lymphocyte antibody methods based on the molecular cloning and expression of immunoglobulin variable region cDNAs generated from single lymphocytes that were selected for the production of specific antibodies such as described by Babcook, J. *et al.*, 1996, Proc. Natl. Acad. Sci. USA 93(15):7843-7848 and in WO 92/02551. The latter methods rely on the isolation of

individual antibody producing cells which are then clonally expanded followed by screening for those clones which are producing an antibody which recognises its cognate antigen, and, if desired, the subsequent identification of the sequence of their variable heavy (V_H) and light (V_L) chain genes. Alternatively, the cells producing antibody that recognises its cognate antigen may be cultured together followed by screening.

The antibody may be specific for any target antigen. The antigen may be a cell-associated protein, for example a cell surface protein on cells such as bacterial cells, yeast cells, T-cells, endothelial cells or tumour cells, or it may be a soluble protein. Antigens of interest may also be any medically relevant protein such as those proteins upregulated during disease or infection, for example receptors and/or their corresponding ligands. Particular examples of cell surface proteins include adhesion molecules, for example integrins such as β1 integrins e.g. VLA-4, E-selectin, P selectin or L-selectin, CD2, CD3, CD4, CD5, CD7, CD8, CD11a, CD11b, CD18, CD19, CD20, CD23, CD25, CD33, CD38, CD40, CD40L, CD45, CDW52, CD69, CD134 (OX40), ICOS, BCMP7, CD137, CD27L, CDCP1, CSF1 or CSF1-Receptor, DPCR1, DPCR1, dudulin2, FLJ20584, FLJ40787, HEK2, KIAA0634, KIAA0659, KIAA1246, KIAA1455, LTBP2, LTK, MAL2, MRP2, nectin-like2, NKCC1, PTK7, RAIG1, TCAM1, SC6, BCMP101, BCMP84, BCMP11, DTD, carcinoembryonic antigen (CEA), human milk fat globulin (HMFG1 and 2), MHC Class I and MHC Class II antigens, KDR and VEGF, PD-1, DC-SIGN, TL1A, IL-7 receptor A and where appropriate, receptors thereof.

Soluble antigens include interleukins such as IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-8, IL-12, IL-13, IL-14, IL-16 or IL-17, such as IL17A and/or IL17F, viral antigens for example respiratory syncytial virus or cytomegalovirus antigens, immunoglobulins, such as IgE, interferons such as interferon α, interferon β or interferon γ, tumour necrosis factor TNF (formerly known as tumour necrosis factor-α and referred to herein as TNF or TNFα), tumor necrosis factor-β, colony stimulating factors such as G-CSF or GM-CSF, and platelet derived growth factors such as PDGF-α, and PDGF-β, WISP-1 and where appropriate receptors thereof. Other antigens include bacterial cell surface antigens, bacterial toxins, viruses such as influenza, EBV, HepA, B and C, bioterrorism agents, radionuclides and heavy metals, and snake and spider venoms and toxins.

In one embodiment, the antibody may be used to functionally alter the activity of the antigen of interest. For example, the antibody may neutralize, antagonize or agonise the activity of said antigen, directly or indirectly.

In a preferred embodiment the antibody is an anti-TNF antibody, more preferably an anti-TNF Fab' CDP870, as described in WO01/094585 (the contents of which are incorporated herein by reference).

In a one embodiment the antibody having specificity for human TNFα, comprises a heavy chain wherein the variable domain comprises a CDR having the sequence shown in SEQ ID NO:1 for CDRH1, the sequence shown in SEQ ID NO:2 for CDRH2 or the sequence shown in SEQ ID NO:3 for CDRH3.

In one embodiment the antibody comprises a light chain wherein the variable domain comprises a CDR having the sequence shown in SEQ ID NO:4 for CDRL1, the sequence shown in SEQ ID NO:5 for CDRL2 or the sequence shown in SEQ ID NO:6 for CDRL3.

In one embodiment the antibody comprises a heavy chain wherein the variable domain comprises a CDR having the sequence shown in SEQ ID NO:1 for CDRH1, the sequence shown in SEQ ID NO:2 for CDRH2 or the sequence shown in SEQ ID NO:3 for CDRH3 and a light chain wherein the variable domain comprises a CDR having the sequence shown in SEQ ID NO:4 for CDRL1, the sequence shown in SEQ ID NO:5 for CDRL2 or the sequence shown in SEQ ID NO:6 for CDRL3.

In one embodiment the antibody comprises SEQ ID NO:1 for CDRH1, SEQ ID NO: 2 for CDRH2, SEQ ID NO:3 for CDRH3, SEQ ID NO:4 for CDRL1, SEQ ID NO:5 for CDRL2 and SEQ ID NO:6 for CDRL3.

The antibody is preferably a CDR-grafted antibody molecule and typically the variable domain comprises human acceptor framework regions and non-human donor CDRs.

Preferably, the antibody comprises the light chain variable domain CDP870 (SEQ ID NO:7) and the heavy chain variable domain CDP870 (SEQ ID NO:8).

It is preferred that the antibody is a modified Fab fragment wherein the modification is the addition to the C-terminal end of its heavy chain one or more amino acids to allow the attachment of an effector or reporter molecule. Preferably, the additional amino acids form a modified hinge region containing one or two cysteine

residue to which the effector or reporter molecule may be attached. Such a modified Fab fragment preferably has a heavy chain comprising or consisting of the sequence given as SEQ ID NO:10 and the light chain comprising or consisting of the sequence given as SEQ ID NO:9.

The host cell used in the present invention comprises DNA sequence encoding the antibody. Preferably, the DNA sequence encodes the heavy and the light chain of the antibody.

In one preferred embodiment, the DNA sequence encodes a light chain and comprises the sequence shown in SEQ ID NO:7 or SEQ ID NO:9 (CDP870) or a degenerate equivalent thereof.

In an alternatively preferred embodiment, the DNA sequence encodes a heavy chain and comprises the sequence shown in SEQ ID NO:8 or SEQ ID NO:10 (CDP870) or a degenerate equivalent thereof.

The DNA sequence of the may comprise synthetic DNA, for instance produced by chemical processing, cDNA, genomic DNA or any combination thereof.

The constant region domains of the antibody, if present, may be selected having regard to the proposed function of the antibody molecule, and in particular the effector functions which may be required. For example, the constant region domains may be human IgA, IgD, IgE, IgG or IgM domains. In particular, human IgG constant region domains may be used, especially of the IgG1 and IgG3 isotypes when the antibody molecule is intended for therapeutic uses and antibody effector functions are required. Alternatively, IgG2 and IgG4 isotypes may be used when the antibody molecule is intended for therapeutic purposes and antibody effector functions are not required, e.g. for simply blocking TNFα activity.

Methods for the expression of recombinant proteins are well known in the art. Suitable examples of host cells for the expression of recombinant antibody molecules include bacteria such as gram positive or gram negative bacteria, *e.g. E. coli*, or yeast cells, e.g. *S. cerevisiae*, or mammalian cells, *e.g.* CHO cells and myeloma or hybridoma cell lines, *e.g.* NSO cells. Most preferably, in the methods of the invention, a recombinant antibody is produced in bacteria, *e.g. E. coli* (see Verma *et al.*, 1988, J. Immunol. Methods 216:165-181; Simmons *et al.*, 2002, J. Immunol. Methods 263:133-147).

After expression, the antibody may be further processed, for example by conjugation to another entity such as an effector molecule. Accordingly, the method according to the present invention may comprise a further step of attaching an effector molecule to the antibody.

The term effector molecule as used herein includes, for example, antineoplastic agents, drugs, toxins (such as enzymatically active toxins of bacterial or plant origin and fragments thereof e.g. ricin and fragments thereof) biologically active proteins, for example enzymes, other antibody or antibody fragments, synthetic or naturally occurring polymers, nucleic acids and fragments thereof e.g. DNA, RNA and fragments thereof, radionuclides, particularly radioiodide, radioisotopes, chelated metals, nanoparticles and reporter groups such as fluorescent compounds or compounds which may be detected by NMR or ESR spectroscopy. Effector molecular may be attached to the antibody or fragment thereof by any suitable method, for example an antibody fragment may be modified to attach at least one effector molecule as described in WO05/003171 or WO05/003170 (the contents of which are incorporated herein by reference). WO05/003171 or WO05/003170 also describe suitable effector molecules.

The antibody may have a macrocycle, for chelating a heavy metal atom, or a toxin, such as ricin, attached to it by a covalent bridging structure. Alternatively, procedures of recombinant DNA technology may be used to produce an antibody molecule in which the Fc fragment (CH2, CH3 and hinge domains), the CH2 and CH3 domains or the CH3 domain of a complete immunoglobulin molecule has (have) been replaced by, or has attached thereto by peptide linkage, a functional non-immunoglobulin protein, such as an enzyme or toxin molecule. In the embodiment wherein the antibody is a modified Fab fragment having at the C-terminal end of its heavy chain one or more amino acids to allow attachment of an effector or reporter molecule, the additional amino acids preferably form a modified hinge region containing one or two cysteine residues to which the effector or reporter molecule may be attached.

A preferred effector group is a polymer molecule, which may be attached to the modified Fab fragment to increase its half-life *in vivo*.

The polymer molecule may, in general, be a synthetic or a naturally occurring polymer, for example an optionally substituted straight or branched chain polyalkylene,

polyalkenylene or polyoxyalkylene polymer or a branched or unbranched polysaccharide, e.g. a homo- or hetero- polysaccharide.

Particular optional substituents which may be present on the above-mentioned synthetic polymers include one or more hydroxy, methyl or methoxy groups. Particular examples of synthetic polymers include optionally substituted straight or branched chain poly(ethyleneglycol), poly(propyleneglycol) poly(vinylalcohol) or derivatives thereof, especially optionally substituted poly(ethyleneglycol) such as methoxypoly(ethyleneglycol) or derivatives thereof. Particular naturally occurring polymers include lactose, amylose, dextran, glycogen or derivatives thereof. "Derivatives" as used herein is intended to include reactive derivatives, for example thiol-selective reactive groups such as maleimides and the like. The reactive group may be linked directly or through a linker segment to the polymer. It will be appreciated that the residue of such a group will in some instances form part of the product as the linking group between the antibody fragment and the polymer.

The size of the polymer may be varied as desired, but will generally be in an average molecular weight range from 500Da to 50,000Da, preferably from 5000 to 40,000Da and more preferably from 25,000 to 40,000Da. The polymer size may in particular be selected on the basis of the intended use of the product. Thus, for example, where the product is intended to leave the circulation and penetrate tissue, for example for use in the treatment of a tumour, it may be advantageous to use a small molecular weight polymer, for example with a molecular weight of around 5,000Da. For applications where the product remains in the circulation, it may be advantageous to use a higher molecular weight polymer, for example having a molecular weight in the range from 25,000Da to 40,000Da.

Particularly preferred polymers include a polyalkylene polymer, such as a poly(ethyleneglycol) or, especially, a methoxypoly(ethyleneglycol) or a derivative thereof, and especially with a molecular weight in the range from about 25,000Da to about 40,000Da.

Each polymer molecule attached to the modified antibody fragment may be covalently linked to the sulphur atom of a cysteine residue located in the fragment. The covalent linkage will generally be a disulphide bond or, in particular, a sulphur-carbon bond.

Where desired, the antibody fragment may have one or more effector or reporter molecules attached to it. The effector or reporter molecules may be attached to the antibody fragment through any available amino acid side-chain or terminal amino acid functional group located in the fragment, for example any free amino, imino, hydroxyl or carboxyl group. One or more effector or reporter molecules may be attached to an amino acid at or towards the C-terminal end of the heavy chain and/or the light chain of the antibody.

An activated polymer may be used as the starting material in the preparation of polymer-modified antibody fragments as described above. The activated polymer may be any polymer containing a thiol reactive group such as an α-halocarboxylic acid or ester, e.g. iodoacetamide, an imide, e.g. maleimide, a vinyl sulphone or a disulphide. Such starting materials may be obtained commercially (for example from Shearwater Polymers Inc., Huntsville, AL, USA) or may be prepared from commercially available starting materials using conventional chemical procedures.

As regards attaching poly(ethyleneglycol) (PEG) moieties, reference is made to "Poly(ethyleneglycol) Chemistry, Biotechnical and Biomedical Applications", 1992, J. Milton Harris (ed), Plenum Press, New York, "Poly(ethyleneglycol) Chemistry and Biological Applications", 1997, J. Milton Harris and S. Zalipsky (eds), American Chemical Society, Washington DC and "Bioconjugation Protein Coupling Techniques for the Biomedical Sciences", 1998, M. Aslam and A. Dent, Grove Publishers, New York.

Where it is desired to obtain an antibody fragment linked to an effector or reporter molecule, this may be prepared by standard chemical or recombinant DNA procedures in which the antibody fragment is linked either directly or via a coupling agent to the effector or reporter molecule either before or after reaction with the activated polymer as appropriate. Particular chemical procedures include, for example, those described in WO 93/62331, WO 92/22583, WO 90,195 and WO 89/1476. Alternatively, where the effector or reporter molecule is a protein or polypeptide the linkage may be achieved using recombinant DNA procedures, for example as described in WO 86/01533 and EP-A-0392745.

Preferably, the modified Fab fragment provided by the method of the present invention is PEGylated (i.e. has PEG (poly(ethyleneglycol)) covalently attached thereto) according to the method disclosed in EP-A-0948544. Preferably the antibody is a

PEGylated modified Fab fragment as shown in Figure 9. As shown in Figure 2, the modified Fab fragment has attached to one of the cysteine residues at the C-terminal end of the modified hinge region of the heavy chain a lysyl-maleimide-derived group wherein each of the two amino groups of the lysyl residue has covalently linked to it a methoxypoly(ethyleneglycol) residue having a molecular weight of about 20,000 Da, such that the total average molecular weight of the methoxypoly(ethyleneglycol) residues is about 40,000Da, more preferably the lysyl-maleimide-derived group is [1-[[[2-[[3-(2,5-dioxo-1-pyrrolidinyl)-1-oxopropyl]amino]ethyl]amino]-carbonyl]-1,5-pentanediyl]bis (iminocarbonyl). A lysine residue is covalently linked to the maleimide group. To each of the amine groups on the lysine residue is attached a methoxypoly(ethyleneglycol) polymer having a molecular weight of approximately 20,000 Da. The total molecular weight of the entire effector molecule is therefore approximately 40,000 Da.

Accordingly, the method according to the present invention preferably comprises attaching to one of the cysteine residues at the C-terminal end of the heavy chain a lysyl-maleimide group wherein each amino group of the lysyl residue has covalently linked to it a methoxypoly(ethyleneglycol) residue having a molecular weight of about 20,000 Da.

Preferably, in the compound shown in Figure 9, the heavy chain of the antibody part has the sequence given as SEQ ID NO:10 and the light chain has the sequence given in SEQ ID NO:9. This compound is referred to herein as CDP870.

Preferred features of each embodiment of the invention are as for each of the other embodiments *mutatis mutandis*. All publications, including but not limited to patents and patent applications cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

The present disclosure explicitly discloses embodiment comprising certain combinations of integers. The present disclosure also extends to embodiments consisting or consisting essentially of the said combinations of integers.

Preferences and/or embodiments may be combined as technically feasible.

The invention will now be described with reference to the following examples, which are merely illustrative and should not in any way be construed as limiting the scope of the present invention.

Examples

Example 1: Generation of Host Cell Line Expressing Recombinant DsbC and Recombinant Antibody

For all experiments the *E.coli* cell line W3110 was used as the parental wild-type cell line. Cell lines were created carrying the following mutations:

a. a mutated Tsp gene;
b. a mutated Tsp gene and carrying recombinant DsbC;
c. a mutated Tsp gene and a mutated spr gene;
d. a mutated Tsp gene and a mutated spr gene and carrying recombinant DsbC.

Generation of Cell Line carrying mutated Tsp gene MXE001 (ΔTsp)

The MXE001 strain was generated as follows:

The Tsp cassette was moved as Sal I, Not I restriction fragments into similarly restricted pKO3 plasmids. The pKO3 plasmid uses the temperature sensitive mutant of the pSC101 origin of replication (*RepA*) along with a chloramphenicol marker to force and select for chromosomal integration events. The *sacB* gene which encodes for levansucrase is lethal to *E . coli* grown on sucrose and hence (along with the chloramphenicol marker and pSC101 origin) is used to force and select for de-integration and plasmid curing events. This methodology had been described previously (Hamilton *et al.*, 1989, *Journal of Bacteriology*, 171, 4617-4622 and Blomfield *et al.*, 1991, *Molecular Microbiology*, 5, 1447-1457). The pKO3 system removes all selective markers from the host genome except for the inserted gene.

The following plasmids were constructed.

pMXE191 comprising the knockout mutated Tsp gene as shown in the SEQ ID NO: 15 comprising *EcoR I* and *Ase I* restriction markers.

The plasmid was then transformed into electro-competent competent *E. coli* W3110 cells prepared using the method found in Miller, E.M. and Nickoloff, J.A., "Escherichia coli electrotransformation," in Methods in Molecular Biology, vol. 47, Nickoloff, J.A. (ed.), Humana Press, Totowa, NJ, 105 (1995).

Day 1 40µl of *E.coli* cells were mixed with (10pg) **1**µl of pKO3 DNA in a chilled BioRad **0.2cm** electroporation cuvette before electroporation at 2500V, 25µF and 200Ω. **1000**µl of

2xPY was added immediately, the cells recovered by shaking at 250rpm in an incubator at 30°C for 1 hour. Cells were serially 1/10 diluted in 2xPY before 100µl aliquots were plated out onto 2xPY agar plates containing chloramphenicol at 20µg/ml prewarmed at 30°C and 43°C. Plates were incubated overnight at 30°C and 43°C.

Day 2 The number of colonies grown at 30°C gave an estimate of the efficiency of electroporation whilst colonies that survive growth at 43°C represent potential integration events. Single colonies from the 43°C plate were picked and resuspended in 10ml of 2xPY. 100µl of this was plated out onto 2xPY agar plates containing 5% (w/v) sucrose pre-warmed to 30°C to generate single colonies. Plates were incubated overnight at 30°C.

Day 3 Colonies here represent potential simultaneous de-integration and plasmid curing events. If the de-integration and curing events happened early on in the growth, then the bulk of the colony mass will be clonal. Single colonies were picked and replica plated onto 2xPY agar that contained either chloramphenicol at 20µg/ml or 5% (w/v) sucrose. Plates were incubated overnight at 30°C.

Day 4 Colonies that both grow on sucrose and die on chloramphenicol represent potential chromosomal replacement and plasmid curing events. These were picked and screened by PCR with a mutation specific oligonucleotide. Colonies that generated a positive PCR band of the correct size were struck out to produce single colonies on 2xPY agar containing 5% (w/v) sucrose and the plates were incubated overnight at 30°C.

Day 5 Single colonies of PCR positive, chloramphenicol sensitive and sucrose resistant *E. coli* were used to make glycerol stocks, chemically competent cells and act as PCR templates for a PCR reaction with 5' and 3' flanking oligos to generate PCR product for direct DNA sequencing using **Taq** polymerase.

Cell strain MXE001 was tested to confirm successful modification of genomic DNA carrying the mutated Tsp gene by PCR amplification of the region of the Tsp gene comprising a non-naturally occurring Ase I restriction site using oligonucleotides primers. The amplified regions of the DNA were then analyzed by gel electrophoresis before and after incubation with Ase I restriction enzyme to confirm the presence of the non-naturally occurring Ase I restriction site in the mutated genes.

The lysates were prepared by heating a single colony of cells for 10 minutes at 95 ^{O}C in 20ul of 1x PCR buffer. The mixture was allowed to cool to room temperature then

centrifugation at 13,200rpm for 10 minutes. The supernatant was removed and labeled as 'cell lysate'.

Each strain was amplified using Tsp oligos pair.

The DNA was amplified using a standard PCR procedure.

5ul	Buffer x10 (Roche)
1ul	dNTP mix (Roche, 10mM mix)
1.5ul	5' oligo (5 pmol)
1.5ul	3' oligo (5 pmol)
2ul	Cell lysate
0.5ul	Taq DNA polymerase (Roche 5U/ul)
38.5ul	H2O

PCR cycle.

94 ^{o}C	1 minute
94 ^{o}C	1 minute)
55 ^{o}C	1 minute) repeated for 30 cycles
72 ^{o}C	1 minute)
72 ^{o}C	10 minutes

Once the reactions were complete 25ul was removed to a new microfuge tube for digestion with Ase I. To the 25ul of PCR reaction 19ul of H_2O, 5ul of buffer 3 (NEB), 1ul of Ase I (NEB) was added, mixed and incubated at 37 ^{o}C for 2 hours.

To the remaining PCR reaction 5ul of loading buffer (x6) was added and 20ul was loaded onto a 0.8% TAE 200ml agarose gel (Invitrogen) plus Ethidium Bromide (5ul of 10mg/ml stock) and run at 100 volts for 1 hour. 10ul of size marker (Perfect DNA marker 0.1-12Kb, Novagen) was loaded in the final lane.

Once the Ase I digestions were complete 10ul of of loading buffer (x6) was added and 20ul was loaded onto a 0.8% TAE agarose gel (Invitrogen) plus Ethidium Bromide (5ul of 10mg/ml stock) and run at 100 volts for 1 hour. 10ul of size marker (Perfect DNA marker 0.1-12Kb, Novagen) was loaded in the final lane. Both gels were visualized using UV transluminator.

The genomic fragment amplified showed the correct sized band of 2.8Kb for Tsp. Following digestion with Ase I this confirmed the presence of the introduced Ase I sites in the Tsp deficient strain MXE001 but not in the W3110 control.

MXE001: genomic DNA amplified using the Tsp primer set and the resulting DNA was digested with Ase I to produce 2.2 and 0.6 Kbps bands.

W3110 PCR amplified DNA was not digested by Ase I restriction enzyme.

Generation of Cell Lines carrying mutated spr gene and Cell Lines carrying mutated Tsp gene and mutated spr gene

The spr mutations were generated and selected for using a complementation assay.

The spr gene was mutated using the Clontech® random mutagenisis diversity PCR kit which introduced 1 to 2 mutations per 1000bp. The mutated spr PCR DNA was cloned into an inducible expression vector [pTTO CDP870] which expresses CDP870 Fab' along with the spr mutant. This ligation was then electro-transformed into an *E.coli* strain MXE001 (ΔTsp) prepared using the method found in Miller, E.M. and Nickoloff, J.A., "*Escherichia coli* electrotransformation," in Methods in Molecular Biology, vol. 47, Nickoloff, J.A. (ed.), Humana Press, Totowa, NJ, 105 (1995). The following protocol was used, 40ul of electro competent MXE001, 2.5ul of the ligation (100pg of DNA) was added to a 0.2cm electroporation cuvette, electro-transformation was performed using as BioRad Genepulser Xcell with the following conditions, 2500V, 25μF and 200Ω. After the electro-transformation 1ml of SOC (Invitrogen) (pre-warmed to 37 ^{o}C) was added and the cells left to recover at 37 ^{o}C for 1 hour with gentle agitation.

The cells where plated onto Hypotonic agar (5g/L Yeast extract, 2.5g/L Tryptone, 15g/L Agar (all Difco)) and incubated at 40 ^{o}C. Cells which formed colonies were re-plated onto HLB at 43 ^{o}C to confirm restoration of the ability to grow under low osmotic conditions at high temperature to the MXE001 strain. Plasmid DNA was prepared from the selected clones and sequenced to identify spr mutations.

Using this method five single and one double mutation in the spr protein were isolated which complemented the ΔTsp phenotype as follows:

1. V98E
2. D133A
3. V135D
4. V135G
5. G147C
6. S95F and Y115F

The individual mutations identified above and three catalytic triad mutations of spr (C94A, H145A, H157A) and W174R were used to generate new strains using either the wild-type W3110 E. coli strain (genotype: F- LAM- IN (rrnD-rrnE)1 rph1 (ATCC no. 27325)) to create spr mutated strains or MXE001 strain from Example 1 to make combined ΔTsp/mutant spr strains.

The following mutant *E. coli* cell strains were generated using a gene replacement vector system using the pKO3 homologous recombination/replacement plasmid (Link *et al.*, 1997, *Journal of Bacteriology*, 179, 6228-6237), as described in Example 1 for the generation of MXE001.

Table 1

Mutant *E. coli* Cell Strain	**Genotype**	**Spr Vectors**
MXE001	ΔTsp	-
MXE008	ΔTsp, spr D133A	pMXE339, pK03 spr D133A (-SalI)
MXE009	ΔTsp, spr H157A	pMXE345, pK03 spr H157A (-SalI)
MXE010	spr G147C	pMXE338, pK03 spr G147C (-SalI)
MXE011	spr C94A	pMXE343, pK03 spr C94A (-SalI)
MXE012	spr H145A	pMXE344, pK03 spr H145A (-SalI)
MXE013	spr W174R	pMXE346, pK03 spr W174R (-SalI)
MXE014	ΔTsp, spr V135D	pMXE340, pK03 spr V135D (-SalI)
MXE015	ΔTsp, spr V98E	pMXE342, pK03 spr V98E (-SalI)
MXE016	ΔTsp, spr C94A	pMXE343, pK03 spr C94A (-SalI)
MXE017	ΔTsp, spr H145A	pMXE344, pK03 spr H145A (-SalI)
MXE018	ΔTsp, spr V135G	pMXE341, pK03 spr V135G (-SalI)

The mutant spr integration cassettes were moved as Sal I, Not I restriction fragments into similarly restricted pKO3 plasmids.

The plasmid uses the temperature sensitive mutant of the pSC101 origin of replication (*RepA*) along with a chloramphenicol marker to force and select for chromosomal

integration events. The *sacB* gene which encodes for levansucrase is lethal to *E . coli* grown on sucrose and hence (along with the chloramphenicol marker and pSC101 origin) is used to force and select for de-integration and plasmid curing events. This methodology had been described previously (Hamilton *et al.*, 1989, *Journal of Bacteriology*, 171, 4617-4622 and Blomfield *et al.*, 1991, *Molecular Microbiology*, 5, 1447-1457). The pKO3 system removes all selective markers from the host genome except for the inserted gene.

The pK03 vectors listed below were constructed, comprising the mutated spr genes including a silent mutation within the spr sequence which removes a SalI restriction site for clone identification.

pMXE336, pK03 spr S95F (-SalI)
pMXE337, pK03 spr Y115F (-SalI)
pMXE338, pK03 spr G147C (-SalI)
pMXE339, pK03 spr D133A (-SalI)
pMXE340, pK03 spr V135D (-SalI)
pMXE341, pK03 spr V135G (-SalI)
pMXE342, pK03 spr V98E (-SalI)
pMXE343, pK03 spr C94A (-SalI)
pMXE344, pK03 spr H145A (-SalI)
pMXE345, pK03 spr H157A (-SalI)
pMXE346, pK03 spr W174R (-SalI)

These plasmids were then transformed into chemically competent *E. coli* W3110 cells prepared using the method found in Miller, E.M. and Nickoloff, J.A., "*Escherichia coli* electrotransformation," in Methods in Molecular Biology, vol. 47, Nickoloff, J.A. (ed.), Humana Press, Totowa, NJ, 105 (1995) or into MXE001 strain from Example 1 to make combined ΔTsp/mutant spr strains, as shown in Table 1.

Day 1 40μl of electro-compentent *E.coli* cells or MXE001 cells were mixed with (10pg) 1μl of pKO3 DNA in a chilled BioRad **0.2cm** electroporation cuvette before electroporation at 2500V, 25μF and 200Ω. **1000**μl of 2xPY was added immediately, the cells recovered by shaking at 250rpm in an incubator at 30°C for 1 hour. Cells were serially 1/10 diluted in 2xPY before 100μl aliquots were plated out onto 2xPY agar plates containing chloramphenicol at 20μg/ml prewarmed at 30°C and 43°C. Plates were incubated overnight at 30°C and 43°C.

Day 2 The number of colonies grown at 30°C gave an estimate of the efficiency of electroporation whilst colonies that survive growth at 43°C represent potential integration events. Single colonies from the 43°C plate were picked and resuspended in 10ml of 2xPY. 100µl of this was plated out onto 2xPY agar plates containing 5% (w/v) sucrose pre-warmed to 30°C to generate single colonies. Plates were incubated overnight at 30°C.

Day 3 Colonies here represent potential simultaneous de-integration and plasmid curing events. If the de-integration and curing events happened early on in the growth, then the bulk of the colony mass will be clonal. Single colonies were picked and replica plated onto 2xPY agar that contained either chloramphenicol at 20µg/ml or 5% (w/v) sucrose. Plates were incubated overnight at 30°C.

Day 4 Colonies that both grow on sucrose and die on chloramphenicol represent potential chromosomal replacement and plasmid curing events. These were picked and screened by PCR plus restriction digest for the loss of a SalI site. Colonies that generated a positive PCR band of the correct size and resistance to digestion by SalI were struck out to produce single colonies on 2xPY agar containing 5% (w/v) sucrose and the plates were incubated overnight at 30°C.

Day 5 Single colonies of PCR positive, chloramphenicol sensitive and sucrose resistant *E. coli* were used to make glycerol stocks, chemically competent cells and act as PCR templates for a PCR reaction with 5' and 3' flanking oligos to generate PCR product for direct DNA sequencing using **Taq** polymerase to confirm the correct mutation.

Generation of plasmid for Fab' and DsbC expression

A plasmid was constructed containing both the heavy and light chain sequences of an anti-TNF Fab' and the sequence encoding DsbC.

A dicistronic message was created of the anti-TNFα Fab' fragment (referred to as CDP870) described in WO01/94585. The upstream cistron encoded the light chain of the antibody (SEQ ID NO: 9) whilst the downstream cistron encoded the heavy chain of the antibody (SEQ ID NO: 10). A DNA sequence encoding the OmpA signal peptide was fused to the 5' end of the DNA coding for each of the light chain and the heavy chain to allow efficient secretion to the periplasm. The intergenic sequence (IGS) 2 was used as shown in SEQ ID NO: 16.

Plasmid pDPH358 (pTTO 40.4 CDP870 IGS2), an expression vector for the CDP870 Fab' (an anti-TNF Fab') and DsbC (a periplasmic polypeptide), was constructed using

conventional restriction cloning methodologies which can be found in Sambrook *et al* 1989, Molecular cloning: a laboratory manual. CSHL press, N.Y. The plasmid pDPH358 contained the following features; a strong *tac* promoter and *lac* operator sequence. The plasmid contained a unique EcoRI restriction site after the coding region of the Fab' heavy chain, followed by a non-coding sequence and then a unique NdeI restriction site. The DsbC gene was PCR cloned using W3110 crude chromosomal DNA as a template such that the PCR product encoded for a 5' EcoRI site followed by a strong ribosome binding, followed by the native start codon, signal sequence and mature sequence of DsbC, terminating in a C-terminal His tag and finally a non-coding NdeI site. The EcoRI-NdeI PCR fragment was restricted and ligated into the expression vector such that all three polypeptides: Fab' light chain, Fab' heavy chain and DsbC were encoded on a single polycistronic mRNA.

The Fab light chain, heavy chain genes and DcbC gene were transcribed as a single polycistronic message. DNA encoding the signal peptide from the *E. coli* OmpA protein was fused to the 5' end of both light and heavy chain gene sequences, which directed the translocation of the polypeptides to the *E. coli* periplasm. Transcription was terminated using a dual transcription terminator *rrnB t1t2*. The *lacIq* gene encoded the constitutively expressed Lac I repressor protein. This repressed transcription from the tac promoter until de-repression was induced by the presence of allolactose or IPTG. The origin of replication used was p15A, which maintained a low copy number. The plasmid contained a tetracycline resistance gene for antibiotic selection.

Expression of anti-TNF Fab' and DsbC

The MXE008 strain was transformed with the plasmid encoding Fab' light chain, Fab' heavy chain and DsbC.

The transformation of the strain was carried out using the method found in Chung C.T *et al* Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989).

Expression of anti-TNF Fab'

The strain MXE008 was transformed with plasmid pMXE117 (pTTO CDP870 or 40.4 IGS2), an expression vector for the CDP870 Fab' (an anti-TNF Fab' having a light chain sequence shown in SEQ ID NO: 9 and a heavy chain sequence shown in SEQ ID NO: 10), which was constructed using conventional restriction cloning methodologies which can be

found in Sambrook *et al* 1989, Molecular cloning: a laboratory manual. CSHL press, N.Y. The plasmid pMXE117 (pTTO CDP870 or 40.4 IGS2) contained the following features; a strong *tac* promoter and *lac* operator sequence. The Fab light and heavy chain genes were transcribed as a single dicistronic message. DNA encoding the signal peptide from the *E. coli* OmpA protein was fused to the 5' end of both light and heavy chain gene sequences, which directed the translocation of the polypeptides to the *E. coli* periplasm. Transcription was terminated using a dual transcription terminator *rrnB t1t2*. The *lacIq* gene encoded the constitutively expressed Lac I repressor protein. This repressed transcription from the tac promoter until de-repression was induced by the presence of allolactose or IPTG. The origin of replication used was p15A, which maintained a low copy number. The plasmid contained a tetracycline resistance gene for antibiotic selection.

The transformation of the strains was carried out using the method found in Chung C.T *et al* Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989).

Example 2 – Growth of Host Cell Expressing Recombinant Antibody anti-TNF Fab'

The MXE008 strain produced in Example 1 was used in a fermentation experiments to express the anti-TNFα Fab'

MXE008 expressing anti-TNF Fab' and DsbC was fermented using standard fermentation conditions and medium. Following fermentation the resulting host cell sample was subjected to centrifugation and the cell pellet was frozen for storage.

Example 3 – extraction of anti-TNFα Fab' from *E. coli* strain & pH adjustment

The frozen pellet cell pellet from Example 2 was then thawed and re-suspended in extraction buffer (0.1M Tris/10mM EDTA pH 7.4). Extraction was carried out overnight at 60°C.

Following the extraction step the extraction sample was centrifuged for 1hour at 4200rpm at 4°C. Post centrifugation the sample was clarified via a 0.45um+0.22um filtration.

After filtration, the resulting host cell sample extract having a pH of 6.9 was divided into 3 aliquots of 10mL and was adjusted to pH 5.0, 4.5 or 3.0 using 30% (v/v) glacial acetic acid. A separate aliquot was retained at pH 6.9 as a control sample.

As can be seen in Figure 1, increased levels of precipitation were observed as the pH decreased. The most precipitation of host cell protein was seen at pH3.0, followed by pH 4.5 and then pH 5.0.

Example 4 – Analysis of host cell sample extracts following pH adjustment by reverse phase HPLC

Samples of each host cell sample extract following pH adjustment to pH 5.0, 4.5 or 3.0 or the control (no-pH adjustment, pH 7) were taken across the time course at time (T)=0, T=1 hour, T=2 hours, T=4.5 hours, T=8 hours and T=24 hours, and real time analysis was carried out via reverse phase HPLC. The reverse phase HPLC allows separation of proteins on the basis of hydrophobicity and was carried out using a suitable HPLC column (Poroshell ™ 300SB-C8, 5micron, 2.1x75mm from Agilent Technologies). Samples were filtered (0.22um) prior to analysis.

The results are shown in Figures 2, 3, 5 and 6. Figure 2 shows the overlaid chromatograms for the reverse phase HPLC analysis at T=0 (T=0 means as soon as was possible to conduct the analysis after pH adjustment, typically around 7 minutes after pH adjustment) for the host cell sample extract after pH adjustment. Figure 3 shows the individual chromatograms for the reverse phase HPLC analysis at T=0 for the host cell sample extract after pH adjustment. The peaks corresponding to the light chain (LC) of the Fab', the heavy chain (HC) of the Fab', the whole Fab', the DsbC-his (his-tagged DsbC) and dipeptide binding protein (DBP) are shown for pH 7 and for pH adjustment to pH 3.0, pH 4.5 and pH 5.0. It can be seen from Figures 2 and 3 that the quantity of DsbC-his is lower after pH adjustment to pH 5 compared to the control at pH 7, the quantity of DsbC-his is significantly lowered further after pH adjustment to pH 4.5 or to pH 3.0. It can also be seen that the quantity of DBP is significantly lowered after pH adjustment to pH 3. The pI of DBP is 5.7. Therefore, in the host cell extract the pH had to be significantly lowered in order cause precipitation of DBP. The quantity of Fab' is not significantly affected by the pH adjustment.

Figure 5 shows the total peak area for DsbC-his from the reverse phase HPLC chromatogram after pH adjustment to pH 3.0, 4.5 and 5.0. As seen in Figures 2 and 3, the amount of DsbC is reduced after adjustment to pH 5.0 and further reduced after adjustment to pH 4.5 or 3.0.

Figure 6 shows the total peak area for DBP from the reverse phase HPLC chromatogram after pH adjustment to pH 3.0, 4,5 and 5.0. As seen in Figures 2 and 3, the amount of DBP is reduced after adjustment to pH 3.0.

A purified solution of DsbC-his was obtained from the host cell extract of strain MXE008 by mixed-mode cation exchange chromatography Analysis of this purified solution comprising DsbC-his was perfomed at pH 7 and at pH 3.0 after T=0, 1, 8 and 24 hours. The results are shown in Figure 4 which shows the chromatograms for reverse phase HPLC analysis of the purified solution comprising DsbC-his at pH 7 and at pH 3.0 after T=0, 1, 8 and 24 hours. It can be seen from Figure 4 that the pH adjustment had no effect on the precipitation and removal of DsbC-his from the solution.

Example 5 – SDS-PAGE Analysis of Samples

Non-reducing SDS-PAGE analysis was carried after an extended storage at pH 7, pH 5.0, pH 4.5 and pH 3.0. Figure 7 shows the SDS-PAGE analysis gel. It can be seen from Figure 7 that the quantity of DBP is reduced after pH adjustment to pH 3. The DsbC and the heavy chain (HC) of the Fab' show at the same molecular weight on the gel but it can be seen that the band for DsbC and HC is reduced following pH adjstument to pH 5, pH 4.5 and pH 3 due to the reduction in DsbC confirms that the quantity of DsbC-his and DBP is reduced after pH adjustment.

The anti-TNF Fab' CDP870 tested in the above examples after pH adjustment to pH 4.5 has been tested by Biacore analysis and found to have retained affinity to TNF. (data not shown).

Claims:

1. A method for purifying a recombinant protein from a gram-negative bacterial host cell sample or extract thereof wherein said host cell expresses a recombinant protein and a recombinant disulphide isomerase DsbC; wherein the method comprises:
   a. adjusting the pH of the host cell sample or extract thereof to a pH of 5 or less to precipitate the recombinant disulphide isomerase; and
   b. separating precipitated recombinant disulphide isomerase from the recombinant protein to produce a recombinant protein sample.
2. The method according to claim 1, wherein the pH of the host cell sample or extract thereof is adjusted to a pH of 4.5 or less.
3. The method according to claim 2, wherein the pH of the host cell sample or extract thereof is adjusted to a pH of 4.5 to 3.0.
4. The method according to claim 2, wherein the pH of the host cell sample or extract thereof is adjusted to a pH of 3 or less.
5. The method according to claim 4, wherein in step a) host cell dipeptide binding protein is further precipitated and in step b) host cell dipeptide binding protein is further separated from the recombinant protein.
6. The method according to any preceding claim, wherein the disulphide isomerase comprises a histidine tag.
7. The method according any preceding claim, wherein the gram-negative bacterial host cell is is *E. coli*.
8. The method according to any preceding claim, wherein the host cell comprises Fkpa and/r Skp.
9. The method according to any preceding claim, wherein in step a) the pH of the host cell sample or extract thereof is adjusted to a pH of less than 5 using glacial acetic acid.
10. The method according to any preceding claim, wherein the host cell sample or extract thereof is held at a pH of 5 or less for one hour or less.
11. The method according to any preceding claim, wherein the separation in step b) comprises centrifugation and/or filtration.

12. The method according to any preceding claim, wherein after step b) the method comprises a further purification step of subjecting the recombinant protein sample to chromatography.

13. The method according to claim 11, wherein the chromatography is ion-exchange chromatography,

14. The method according to any preceding claim, wherein after step b) the pH of the recombinant protein sample is adjusted to a pH of 5 to 7.

15. The method according to any preceding claim, wherein prior to step a) an extraction buffer is added to the host cell sample or extract thereof and the host cell sample or extract thereof is subjected to a heat treatment step.

16. The method according to any preceding claim, wherein the recombinant protein has a pI of 6 to 9.

17. The method according to claim 16, wherein the recombinant protein has a pI of 8 to 9.

18. The method according to any preceding claim, wherein the recombinant protein is a recombinant antibody.

19. The method according to claim 18, wherein the recombinant antibody is a monoclonal, humanised or chimeric antibody, or a fragment thereof, such as a binding fragment thereof.

20. The method according to claim 18 or claim 19, wherein the recombinant antibody is a Fab or Fab' fragment.

21. The method according to any of claims 18 to 20, wherein the recombinant antibody is specific for human TNFα.

22. The method according to claim 21, wherein the antibody a heavy chain wherein the variable domain comprises a CDR having the sequence given in SEQ ID NO:1 for CDRH1, SEQ ID NO:2, for CDRH2 and SEQ ID NO:3 for CDRH3 and a light chain wherein the variable domain comprises a CDR having the sequence given in SEQ ID NO:4 for CDRL1, SEQ ID NO:5 for CDRL2 and SEQ ID NO:6 for CDRL3.

23. The method according to claim 22, wherein the antibody comprises the light chain variable region sequence given in SEQ ID NO:7 and the heavy chain variable region given in SEQ ID NO:8.

24. The method according to claim 23, wherein the antibody is a Fab fragment and comprises a heavy chain having the sequence given in SEQ ID NO:10 and a light chain having the sequence given in SEQ ID NO:9.

25. Use of a step of adjusting the pH of a gram-negative bacterial host cell sample or extract thereof transformed with an expression vector encoding a recombinant protein to a pH of 5 or less to precipitate host cell recombinant disulphide isomerase, separate the precipitated host cell recombinant disulphide isomerase from the recombinant protein and produce a purified recombinant protein sample.

SEQUENCE LISTING

<110> UCB CELLTECH

<120> PROCESS FOR PURIFYING PROTEINS

<130> G0126-WO01

<140> PCT/GB2011/001129
<141> 2011-07-27

<150> GB1012599.5
<151> 2010-07-27

<160> 16

<170> PatentIn version 3.5

<210> 1
<211> 5
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRH1 of CDP870

<400> 1

Asp Tyr Gly Met Asn
1 5

<210> 2
<211> 17
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRH2 of CDP870

<400> 2

Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> 3
<211> 9
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRH3 of CDP870

<400> 3

Gly Tyr Arg Ser Tyr Ala Met Asp Tyr
1 5

<210> 4
<211> 11
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRL1 of CDP870

<400> 4

Lys Ala Ser Gln Asn Val Gly Thr Asn Val Ala
1 5 10

<210> 5
<211> 7
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRL2 of CDP870

<400> 5

Ser Ala Ser Phe Leu Tyr Ser
1 5

<210> 6
<211> 9
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of CDRL3 of CDP870

<400> 6

Gln Gln Tyr Asn Ile Tyr Pro Leu Thr
1 5

<210> 7
<211> 107
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of the light chain variable region CDP870

<400> 7

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val Gly Thr Asn
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Ala Leu Ile
35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Tyr Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
100 105

<210> 8
<211> 118
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of the heavy chain variable region CDP870

<400> 8

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr
100 105 110

Leu Val Thr Val Ser Ser
115

<210> 9
<211> 214
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of a grafted anti-TNF? Fab CDP870 light chain

<400> 9

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Gln Asn Val Gly Thr Asn
20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Ala Leu Ile
35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Tyr Arg Phe Ser Gly
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205

Phe Asn Arg Gly Glu Cys
210

<210> 10
<211> 229
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of a grafted anti-TNF? Fab CDP870 heavy chain

<400> 10

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met
35 40 45

Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr
100 105 110

Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro
115 120 125

Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
130 135 140

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn
145 150 155 160

Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln
165 170 175

Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser
180 185 190

Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser
195 200 205

Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr
210 215 220

His Thr Cys Ala Ala
225

<210> 11
<211> 720
<212> DNA
<213> Artificial

<220>
<223> Nucleotide sequence of his-tagged DsbC

<400> 11
atgaagaaag gttttatgtt gtttactttg ttagcggcgt tttcaggctt tgctcaggct 60

gatgacgcgg caattcaaca aacgttagcc aaaatgggca tcaaaagcag cgatattcag 120

cccgcgcctg tagctggcat gaagacagtt ctgactaaca gcggcgtgtt gtacatcacc 180

gatgatggta aacatatcat tcaggggcca atgtatgacg ttagtggcac ggctccggtc 240

aatgtcacca ataagatgct gttaaagcag ttgaatgcgc ttgaaaaaga gatgatcgtt 300

tataaagcgc cgcaggaaaa acacgtcatc accgtgttta ctgatattac ctgtggttac 360

tgccacaaac tgcatgagca aatggcagac tacaacgcgc tggggatcac cgtgcgttat 420

cttgctttcc cgcgccaggg gctggacagc gatgcagaga aagaaatgaa agctatctgg 480

tgtgcgaaag ataaaaacaa agcgtttgat gatgtgatgg caggtaaaag cgtcgcacca 540

gccagttgcg acgtggatat tgccgaccat tacgcacttg gcgtccagct tggcgttagc 600

ggtactccgg cagttgtgct gagcaatggc acacttgttc cgggttacca gccgccgaaa 660

gagatgaaag aatttctcga cgaacaccaa aaaatgacca gcggtaaaca ccatcaccat 720

<210> 12
<211> 242
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of his-tagged DsbC

<400> 12

Met Lys Lys Gly Phe Met Leu Phe Thr Leu Leu Ala Ala Phe Ser Gly
1 5 10 15

Phe Ala Gln Ala Asp Asp Ala Ala Ile Gln Gln Thr Leu Ala Lys Met
20 25 30

Gly Ile Lys Ser Ser Asp Ile Gln Pro Ala Pro Val Ala Gly Met Lys
35 40 45

Thr Val Leu Thr Asn Ser Gly Val Leu Tyr Ile Thr Asp Asp Gly Lys
50 55 60

His Ile Ile Gln Gly Pro Met Tyr Asp Val Ser Gly Thr Ala Pro Val
65 70 75 80

Asn Val Thr Asn Lys Met Leu Leu Lys Gln Leu Asn Ala Leu Glu Lys
85 90 95

Glu Met Ile Val Tyr Lys Ala Pro Gln Glu Lys His Val Ile Thr Val
100 105 110

Phe Thr Asp Ile Thr Cys Gly Tyr Cys His Lys Leu His Glu Gln Met
115 120 125

Ala Asp Tyr Asn Ala Leu Gly Ile Thr Val Arg Tyr Leu Ala Phe Pro
130 135 140

Arg Gln Gly Leu Asp Ser Asp Ala Glu Lys Glu Met Lys Ala Ile Trp
145 150 155 160

Cys Ala Lys Asp Lys Asn Lys Ala Phe Asp Asp Val Met Ala Gly Lys
165 170 175

```
Ser Val Ala Pro Ala Ser Cys Asp Val Asp Ile Ala Asp His Tyr Ala
            180                 185                 190

Leu Gly Val Gln Leu Gly Val Ser Gly Thr Pro Ala Val Val Leu Ser
        195                 200                 205

Asn Gly Thr Leu Val Pro Gly Tyr Gln Pro Pro Lys Glu Met Lys Glu
    210                 215                 220

Phe Leu Asp Glu His Gln Lys Met Thr Ser Gly Lys His His His His
225                 230                 235                 240

His His
```

<210> 13
<211> 188
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of wild-type spr gene including the signal sequence which is the first 26 amino acid residues

<400> 13

```
Met Val Lys Ser Gln Pro Ile Leu Arg Tyr Ile Leu Arg Gly Ile Pro
1               5                   10                  15

Ala Ile Ala Val Ala Val Leu Leu Ser Ala Cys Ser Ala Asn Asn Thr
            20                  25                  30

Ala Lys Asn Met His Pro Glu Thr Arg Ala Val Gly Ser Glu Thr Ser
        35                  40                  45

Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu Asn Leu Val Arg Asn Val
    50                  55                  60

Asp Val Lys Ser Arg Ile Met Asp Gln Tyr Ala Asp Trp Lys Gly Val
65                  70                  75                  80

Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys Gly Ile Asp Cys Ser Gly
                85                  90                  95

Phe Val Gln Arg Thr Phe Arg Glu Gln Phe Gly Leu Glu Leu Pro Arg
            100                 105                 110

Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys Ser Val Ser Arg Ser Asn
        115                 120                 125

Leu Arg Thr Gly Asp Leu Val Leu Phe Arg Ala Gly Ser Thr Gly Arg
    130                 135                 140
```

```
His Val Gly Ile Tyr Ile Gly Asn Asn Gln Phe Val His Ala Ser Thr
145                 150                 155                 160

Ser Ser Gly Val Ile Ile Ser Ser Met Asn Glu Pro Tyr Trp Lys Lys
                165                 170                 175

Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser Arg Ser
            180                 185
```

<210> 14
<211> 162
<212> PRT
<213> Artificial

<220>
<223> Amino acid sequence of the wild-type spr gene without the signal sequence

<400> 14

```
Cys Ser Ala Asn Asn Thr Ala Lys Asn Met His Pro Glu Thr Arg Ala
1               5                   10                  15

Val Gly Ser Glu Thr Ser Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu
            20                  25                  30

Asn Leu Val Arg Asn Val Asp Val Lys Ser Arg Ile Met Asp Gln Tyr
        35                  40                  45

Ala Asp Trp Lys Gly Val Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys
    50                  55                  60

Gly Ile Asp Cys Ser Gly Phe Val Gln Arg Thr Phe Arg Glu Gln Phe
65                  70                  75                  80

Gly Leu Glu Leu Pro Arg Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys
                85                  90                  95

Ser Val Ser Arg Ser Asn Leu Arg Thr Gly Asp Leu Val Leu Phe Arg
            100                 105                 110

Ala Gly Ser Thr Gly Arg His Val Gly Ile Tyr Ile Gly Asn Asn Gln
        115                 120                 125

Phe Val His Ala Ser Thr Ser Ser Gly Val Ile Ile Ser Ser Met Asn
    130                 135                 140

Glu Pro Tyr Trp Lys Lys Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser
145                 150                 155                 160

Arg Ser
```

<210> 15
<211> 2048
<212> DNA
<213> Artificial

<220>
<223> DNA sequence of a mutated knockout Tsp gene including the 6 nucleotides ATGAAT upstream of the start codon

<400> 15
atgaattcgt ttttaggctt accgcgttag ctggcctgct tgcaatagca ggccagacat 60

taattgtaga agatatcacg cgtgctgatc aaattccggt attaaaggaa gagacgcagc 120

atgcgacggt aagtgagcgc gtaacgtcgc gcttcacccg ttctcattat cgccagttcg 180

acctcgatca ggcattttcg gccaaaatct ttgaccgcta cctgaatctg ctcgattaca 240

gccacaacgt gctgctggca agcgatgttg aacagttcgc gaaaaagaaa accgagttag 300

gcgatgaact gcgttcaggc aaactcgacg ttttctacga tctctacaat ctggcgcaaa 360

agcgccgttt tgagcgttac cagtacgctt tgtcggtact ggaaaagccg atggatttca 420

ccggcaacga cacttataac cttgaccgca gcaaagcgcc ctggccgaaa aacgaggctg 480

agttgaacgc gctgtgggac agtaaagtca aattcgacga gttaagcctg aagctgacag 540

gaaaaacgga taaagaaatt cgtgaaaccc tgactcgccg ctacaaattt gccattcgtc 600

gtctggcgca aaccaacagc gaagatgttt tctcgctggc aatgacggcg tttgcgcgtg 660

aaatcgaccc gcataccaac tatctttccc cgcgtaatac cgaacagttc aacactgaaa 720

tgagtttgtc gctggaaggt attggcgcag tgctgcaaat ggatgatgac tacaccgtta 780

tcaattcgat ggtggcaggt ggtccggcag cgaagagtaa agctatcagc gttggtgaca 840

aaattgtcgg tgttggtcaa acaggcaagc cgatggttga cgtgattggc tggcgtcttg 900

atgatgtggt tgccttaatt aaagggccga agggcagtaa agttcgtctg gaaattttac 960

ctgctggtaa agggaccaag acccgtactg taacgttgac ccgtgaacgt attcgtctcg 1020

aagaccgcgc ggttaaaatg tcggtgaaga ccgtcggtaa agagaaagtc ggcgtgctgg 1080

atattccggg cttctatgtg ggtttgacag acgatgtcaa agtgcaactg cagaaactgg 1140

aaaaacagaa tgtcagcagc gtcatcatcg acctgcgtag caatggcggt ggggcgttaa 1200

ctgaagccgt atcgctctcc ggtctgttta ttcctgcggg tcccattgtt caggtccgcg 1260

ataacaacgg caaggttcgt gaagatagcg ataccgacgg acaggttttc tataaaggcc 1320

cgctggtggt gctggttgac cgcttcagtg cttcggcttc agaaatcttt gccgcggcaa 1380

tgcaggatta cggtcgtgcg ctggttgtgg gtgaaccgac gtttggtaaa ggcaccgttc 1440

agcaataccg ttcattgaac cgtatttacg atcagatgtt acgtcctgaa tggccagcgc 1500

tgggttctgt gcagtacacg atccagaaat tctatcgcgt taacggcggc agtacgcaac 1560

gtaaaggcgt aacgccagac atcatcatgc cgacgggtaa tgaagaaacg gaaacgggtg 1620

agaaattcga agataacgcg ctgccgtggg atagcattga tgccgcgact tatgtgaaat 1680

caggagattt aacggccttt gaaccggagc tgctgaagga acataatgcg cgtatcgcga 1740

aagatcctga gttccagaac atcatgaagg atatcgcgcg cttcaacgct atgaaggaca 1800

agcgcaatat cgtttctctg aattacgctg tgcgtgagaa agagaataat gaagatgatg 1860

cgacgcgtct ggcgcgtttg aacgaacgct ttaaacgcga aggtaaaccg gagttgaaga 1920

aactggatga tctaccgaaa gattaccagg agccggatcc ttatctggat gagacggtga 1980

atatcgcact cgatctggcg aagcttgaaa aagccagacc cgcggaacaa cccgctcccg 2040

tcaagtaa 2048

<210> 16
<211> 69
<212> DNA
<213> Artificial

<220>
<223> Oligonucleotide cassette encoding intergenic sequence 2 (IGS2) for E. coli Fab expression

<400> 16
gagctcacca gtaacaaaaa gttttaatag aggggagtgt taaaatgaag aagactgcta 60

tagcaattg 69