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