(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 September 2010 (10.09.2010)

PCT

(10) International Publication Number
WO 2010/100632 A3

(51) International Patent Classification:
A61K 39/118 (2006.01)

(21) International Application Number:
PCT/IB2010/050988

(22) International Filing Date:
8 March 2010 (08.03.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/157,921 6 March 2009 (06.03.2009) US

(71) Applicant *(for all designated States except US)*: **NOVARTIS AG** [CH/CH]; Lichstrasse 35, CH-4056 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **GRANDI, Guido** [IT/IT]; Novartis Vaccines and Diagnostics SRL, 1 Via Fiorentina, Siena (IT). **GRIFANTINI, Renata Maria** [IT/IT]; Novartis Vaccines and Diagnostics SRL, 1 Via Fiorentina, Siena (IT). **FINCO, Oretta** [IT/IT]; Novartis Vaccines and Diagnostics SRL, 1 Via Fiorentina, Siena (IT).

(74) Agent: **MARSHALL, Cameron John**; Carpmaels & Ransford, One Southampton Row, London WC1B 5HA (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

(88) Date of publication of the international search report:
20 January 2011

WO 2010/100632 A3

(54) Title: CHLAMYDIA ANTIGENS

(57) Abstract: The invention provides identifies Chlamydia antigens for use in the treatment, prevention and/or diagnosis of Chlamydia infection. In particular, the invention provides antigens CT733, CTl 53, CT601, CT279, CT443, CT372, CT456, CT381, CT255, CT341, CT716, CT745, CT387, CT812, CT869, CT166, CT175, CT163, CT214, CT721, CT127, CT043, CT823 and/or CT600 from *C. trachomatis for the treatment, prevention or diagnosis of Chlamydia infection.*

CHLAMYDIA ANTIGENS

TECHNICAL FIELD

This invention is in the field of *Chlamydia trahomatis* proteins and their uses.

BACKGROUND ART

Vaccine development has been identified as essential to controlling infection with *C. trachomatis*. Vaccines against *C. trachomatis* appear to elicit protective T-cell and/or B-cell immunity in the genital tract mucosa.

Protective immunity to *C. trachomatis* seems to depend on a Th1-polarized cell-mediated immune response, in particular on CD4+ lymphocytes secreting IFNγ. For example, depletion of CD4+ T cells in mice results in loss of protective immunity, and adoptive transfer of Chlamydia-specific CD4+ T cells confers protection against challenge with *C. trachomatis*. Furthermore, recent studies report that *C. trachomatis* infection in mice induces a CD4-Th1 protective immune response, indicating that critical Chlamydia antigens are processed and presented via the MHC class II pathway (Brunham RC and Rey-Ladino J (2005), Nat Rev Immunol 5: 149-1611; Su H and Caldwell HD (1995), Infect Immun 63: 3302-3308).

Although B-cells and antibodies do not have a decisive role in resolution of primary infection, they are likely to be important for enhancing the protective effector T-cell response and to be required to control re-infection with various mechanisms such as antibody-mediated neutralization and opsonization.

Because immune protection against infection with *C. trachomatis* is likely to be mediated by immunization with *C. trachomatis* proteins that are targets of CD4+ T cells and that are capable of inducing B-cell responses, identification of such proteins is particularly important. It is therefore an object of the invention to provide further antigens for use in Chlamydia vaccines.

DISCLOSURE OF THE INVENTION

The invention provides identifies Chlamydia antigens for use in the treatment, prevention and/or diagnosis of Chlamydia infection. In particular, the invention provides one or more of the following antigens (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30) from *C. trachomatis* for the treatment, prevention or diagnosis of Chlamydia infection (and, in particular, *C. trachomatis* infection): CT733, CT153, CT601, CT279, CT443, CT372, CT456, CT381, CT255, CT341, CT716, CT745, CT812, CT869, CT387, CT166, CT175, CT163, CT214, CT721, CT127, CT043, CT823, CT600, CT711, CT114, CT480, CT089, CT734 and CT016 for example, one or more of CT733, CT153, CT601, CT279, CT443, CT372, CT456, CT381, CT255, CT341, CT716 and CT745.

In particular, the invention provides proteins for use in the treatment, prevention and/or diagnosis of Chlamydia infection (and, in particular, *C. trachomatis* infection). Immunisation with the proteins is preferably able to induce a specific CD4+ Th1 cell mediated response against Chlamydia.

In one embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:1 and SEQ ID NO:2 respectively. This protein is also known as "CT733" and is annotated as a hypothetical protein from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:3 and SEQ ID NO:4 respectively. This protein is also known as "CT153" and is annotated as MACPF/ membrane-attack complex (MAC)/ perforin from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:5 and SEQ ID NO:6 respectively. This protein is also known as "CT601" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:7 and SEQ ID NO:8 respectively. This protein is also known as "CT279" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:9 and SEQ ID NO:10 respectively. This protein is also known as "CT443" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:11 and SEQ ID NO:12 respectively. This protein is also known as "CT372" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:13 and SEQ ID NO:14 respectively. This protein is also known as "CT456" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:15 and SEQ ID NO:16 respectively. This protein is also known as "CT381" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:39 and SEQ ID NO:40 respectively. This protein is also known as "CT255" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:41 and SEQ ID NO:42 respectively. This protein is also known as "CT341" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:43 and SEQ ID NO:44 respectively. This protein is also known as "CT716" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:45 and SEQ ID NO:46 respectively. This protein is also known as "CT745" from *C. trachomatis*. In another embodiment, the nucleic acid sequence and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:47 and SEQ ID NO:48, respectively. This protein is also known as "CT387" from *C. trachomatis* and is annotated as a hypothetical protein. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:49 and

SEQ ID NO:50, respectively. This protein is also known as "CT812" from *C. trachomatis* and is annotated as a polymorphic outer membrane protein. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:51 and SEQ ID NO:52, respectively. This protein is also known as "CT869" from *C. trachomatis* and is annotated as a polymorphic outer membrane protein. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:53 and SEQ ID NO:54, respectively. This protein is also known as "CT166" from *C. trachomatis*. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:55 and SEQ ID NO:56, respectively. This protein is also known as "CT175" from *C. trachomatis*. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:155 and SEQ ID NO:156, respectively. This protein is also known as "CT163" from *C. trachomatis*. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:159 and SEQ ID NO:160, respectively. This protein is also known as "CT214" from *C. trachomatis*. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:163 and SEQ ID NO:164, respectively. This protein is also known as "CT721" from *C. trachomatis*. In another embodiment, the nucleic acid and/or amino acid sequence of the protein comprises the sequence presented in SEQ ID NO:167 and SEQ ID NO:168, respectively. This protein is also known as "CT127" from *C. trachomatis*.

In some embodiments, the protein is a variant of a protein as described above. For example, the protein may comprise one or more mutations (for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more mutations) in the sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 40, 42, 44, 46, 48, 50, 52, 54, 56, 136, 140, 156, 160, 164 or 168, for example, in the sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 40, 42, 44, or 46. Preferred mutations are those which do not cause a significant conformational change in the protein such that the protein of the invention retains the ability to elicit an immune response against the wild-type Chlamydia protein. The proteins having the sequences presented in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 40, 42, 44, 46, 48, 50, 52, 54 and 56 are the wild-type proteins.

In some embodiments, the one or more mutations are present in the N-terminal portion of the protein, for example, between residues 1 and 20 of the protein, between residues 21 and 40, between residues 41 and 60, between residues 1 and 60 or between residues 1 and 40 of the protein. In some embodiments, the one or more mutations are present in the C-terminal portion of the protein, for example, between the C-terminal 20 residues of the protein, between residues 21 and 40 from the C-terminus, between residues 41 and 60 from the C-terminus; between residues 1 and 60 from the C-terminus or between residues 1 and 40 from the C-terminus of the protein.

Preferably, the amino acid sequences contain fewer than twenty mutations (e.g. 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1). Each mutation preferably involves a single amino acid

and is preferably a point mutation. The mutations may each independently be a substitution, an insertion or a deletion. Preferred mutations are single amino acid substitutions. The proteins may also include one or more (*e.g.* 1, 2, 3, 4, 5, 6, 7, 8, 9, *etc.*) single amino acid deletions relative to the Chlamydia sequences. The proteins may also include one or more (*e.g.* 1, 2, 3, 4, 5, 6, 7, 8, 9, *etc.*) insertions (*e.g.* each of 1, 2, 3, 4 or 5 or more amino acids) relative to the Chlamydia sequences. Deletions, substitutions or insertions may be at the N-terminus and/or C-terminus, or may be between the two termini. Thus a truncation is an example of a deletion. Truncations may involve deletion of up to 40 (or more) amino acids at the N-terminus and/or C-terminus (for example, 1-10, 11-40, 41-70, 71-100 or more amino acids).

Amino acid substitutions may be to any one of the other nineteen naturally occurring amino acids. Preferably, a substitution mutation is a conservative substitution. Alternatively, a substitution mutation is a non-conservative substitution. A conservative substitution is commonly defined as a substitution introducing an amino acid having sufficiently similar chemical properties, e.g. having a related side chain (*e.g.* a basic, positively charged amino acid should be replaced by another basic, positively charged amino acid), in order to preserve the structure and the biological function of the molecule. Genetically-encoded amino acids are generally divided into four families: (1) acidic *i.e.* aspartate, glutamate; (2) basic *i.e.* lysine, arginine, histidine; (3) non-polar *i.e.* alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar *i.e.* glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In general, substitution of single amino acids within these families does not have a major effect on the biological activity. Further examples of conversative substitutions that may be used in the invention are presented in Table 1.

TABLE 1

Amino Acid	Synonymous Groups	More Preferred Synonymous Groups
Ser	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Arg	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Leu	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Pro	Gly, Ala, Ser, Thr, Pro	Pro
Thr	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Ala	Gly, Thr, Pro, Ala, Ser	Gly, Ala
Val	Met, Phe, Ile, Leu, Val	Met, Ile, Val, Leu
Gly	Ala, Thr, Pro, Ser, Gly	Gly, Ala
Ile	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Phe	Trp, Phe, Tyr	Tyr, Phe
Tyr	Trp, Phe, Tyr	Phe, Tyr
Cys	Ser, Thr, Cys	Cys
His	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Gln	Glu, Asn, Asp, Gln	Asn, Gln
Asn	Glu, Asn, Asp, Gln	Asn, Gln
Lys	Asn, Lys, Gln, Arg, His	Arg, Lys, His

Asp	Glu, Asn, Asp, Gln	Asp, Glu
Glu	Glu, Asn, Asp, Gln	Asp, Glu
Met	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Trp	Trp, Phe, Tyr	Trp

Examples of non-conservative substitutions that may be used in the invention include the substitution of an uncharged polar amino acid with a nonpolar amino acid, the substitution of a nonpolar amino acid with an uncharged polar amino acid, the substitution of an acidic amino acid with a basic amino acid and the substitution of a basic amino acid with an acidic amino acid.

Mutations may also be introduced to improve stability, *e.g.*, the insertion of disulphide bonds (van den Akker et al. Protein Sci., 1997, 6:2644-2649). For example, the protein may comprise an amino acid sequence having sequence identity to the amino acid sequence of any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 40, 42, 44, 46, 48, 50, 52, 54, 56, 136, 140, 156, 160, 164 and 168, for example, of any one of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 40, 42, 44 and 46. The degree of sequence identity is preferably greater than 50% (e.g. 60%, 70%, 80%, 90%, 95%, 97%, 98%, 99% or more). These proteins include homologs, orthologs, allelic variants and functional mutants. Identity between proteins is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters gap open penalty = 12 and gap extension penalty = 1.

The Chlamydia protein of the invention may comprise one or more amino acid derivatives. By "amino acid derivative" is intended an amino acid or amino acid-like chemical entity other than one of the 20 genetically encoded naturally occurring amino acids. In particular, the amino acid derivative may contain substituted or non-substituted, linear, branched, or cyclic alkyl moieties, and may include one or more heteroatoms. The amino acid derivatives can be made de novo or obtained from commercial sources (Calbiochem-Novabiochem AG; Bachem).

In some embodiments, the variant protein is a homologous protein from *C. pneumoniae*, *C. psittaci*, *C. pecorum*, *C. muridarum* or *C. suis*.

The invention further provides a protein comprising or consisting of a fragment of a protein comprising or consisting of the amino acid sequence of any of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 40, 42, 44, 46, 48, 50, 52, 54, 56, 136, 140, 156, 160, 164 or 168, for example, of any one of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 40, 42, 44 or 46, or a fragment of a variant thereof. The fragment should comprise at least *n* consecutive amino acids from the protein and, depending on the particular sequence, *n* is 6 or more (*e.g.* 8, 11, 16, 31, 51, 76, 121, 181, 231, 281, 331, 381, 431, 440, 445, 446, 481, 531, 581, 631, 681, 731, 781, 801, 806, 808 or more). The fragment is n-1 amino acids or less in length, wherein n = the number of amino acids in the full length protein (e.g. n-5, n-20, n-50, n-110, n-180, n-240, n-310, n-380, n-445, n-515, n-595, n-675, n-745, n-785, n-800 amino acids or less in length). Preferably the fragment comprises one or more

epitopes from the protein. Preferably, one or more of the epitopes is an MHC class II epitope, for example, a CD4+ T cell epitope. In some embodiments, the fragment comprises or consists of the amino acid sequence of any of SEQ ID NOs 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 138, 142, 146, 150, 154, 158, 162, 166 and 170. In some embodiments, the invention provides a protein comprising or consisting of a fragment of a protein comprising or consisting of the amino acid sequence recited in SEQ ID NO: 122. Table 3 below shows which fragments correspond to which full length sequences.

TABLE 3

Annotation	SEQ ID NO. for full length sequence	SEQ ID NO. for fragment
CT733	1	63
CT733	2	64
CT153	3	65
CT153	4	66
CT601	5	67
CT601	6	68
CT279	7	69
CT279	8	70
CT443	9	71
CT443	10	72
CT372	11	73
CT372	12	74
CT456	13	75
CT456	14	76
CT381	15	77
CT381	16	78
CT043	17	79
CT043	18	80
CT711	19	81 (nucleotide); 82 (protein)
CT114	20	83 (nucleotide); 84 (protein)
CT480	21	85 (nucleotide); 86 (protein)
CT089	22	87 (nucleotide); 88 (protein)
CT734	23	89 (nucleotide); 90 (protein)
CT016	24	91 (nucleotide); 92 (protein)
TC0551 (CT279)	25	93

TC0551 (CT279)	26	94
TC0651 (CT372)	27	95
TC0651 (CT372)	28	96
TC0727 (CT443)	29	97
TC0727 (CT443)	30	98
TC0313 (CT043)	31	99
TC0313 (CT043)	32	100
TC0890 (CT601)	33	101
TC0890 (CT601)	34	102
TC0741 (CT456)	35	103
TC0741 (CT456)	36	104
TC0660 (CT381)	37	105
TC0660 (CT381)	38	106
CT255	39	107
CT255	40	108
CT341	41	109
CT341	42	110
CT716	43	111
CT716	44	112
CT745	45	113
CT745	46	114
CT387	47	115
CT387	48	116
CT812	49	117 (mature full length); 119 (N-terminal fragment); 121 (C-terminal fragment)
CT812	50	118 (mature full length) 120 (N-terminal fragment) 122 (C-terminal fragment)
CT869	51	123
CT869	52	124
CT166	53	125
CT166	54	126
CT175	55	127
CT175	56	128
TC0666 (CT387)	57	129
TC0666 (CT387)	58	130

TC0197	59	131
TC0197	60	132
TC0261	61	133
TC0261	62	134
CT600	135	137
CT600	136	138
CT823	139	141
CT823	140	142
TC0106	143	145
TC0106	144	146
TC0431	147	149
TC0431	148	150
TC0210	151	153
TC0210	152	154
CT163	155	157
CT163	156	158
CT214	159	161
CT214	160	162
CT721	163	165
CT721	164	166
CT127	167	169
CT127	168	170

The protein of the invention, for example the variant protein or the fragment, is preferably immunogenic.

The term "immunogenic" in the context of "an immunogenic variant" and "immunogenic fragment", is used to mean that the protein is capable of eliciting an immune response, such as a cell-mediated and/or an antibody response, against the wild-type Chlamydia protein from which it is derived, for example, when used to immunise a subject (preferably a mammal, more preferably a human or a mouse). For example, the protein of the invention (for example, the variant or fragment) is preferably capable of stimulating *in vitro* CD4+ IFNγ+ cells in splenocytes purified from mice infected with live *C. trachomatis* to a level comparable with the wild-type Chlamydia protein. The protein of the invention preferably retains the ability to elicit antibodies that recognise the wild-type protein. For example, the protein of the invention preferably elicits antibodies that can bind to, and preferably neutralise the activity of, the wild-type protein. In a further embodiment, the protein of the invention is capable of eliciting antibodies that are capable of neutralising Chlamydia infectivity and/or virulence. In some embodiments, the antibodies are able to cross-react with the protein of the

invention and the wild-type protein, but with no other homologous protein (e.g. from another Chlamydia species). In other embodiments, the antibodies are cross-reactive with the wild-type protein and with homologous proteins from other Chlamydia species. In some embodiments, the antibodies are cross-reactive with the wild-type protein and with homologous protein from other organisms (for example from *E.coli* or *H.influenzae*). Mice immunized with the protein of the invention and the wild-type Chlamydia protein preferably show similar antigen-specific antibody titers. Antibody titres and specificities can be measured using standard methods available in the art. Other methods of testing the immunogenicity of proteins are also well known in the art.

For example, the variant or fragment is preferably capable of eliciting an immune response, such as a cell-mediated and/or an antibody response, against the wild-type Chlamydia protein. In one embodiment the fragment is capable of stimulating *in vitro* CD4+ IFNγ+ cells in splenocytes purified from mice infected with live *C. trachomatis* to a level comparable with the wild-type Chlamydia protein and/or retains the ability to elicit antibodies that recognise the wild-type protein.

Preferably, the variant or the fragment is capable of inducing a specific CD4-Th1 cell mediated response against the wild type Chlamydia protein.

The proteins of the invention can, of course, be prepared by various means (*e.g.* recombinant expression, purification from native host, purification from cell culture, chemical synthesis *etc.*) and in various forms (*e.g.* native, fusions, glycosylated, non-glycosylated, lipidated, non-lipidated, phosphorylated, non-phosphorylated, myristoylated, non-myristoylated, monomeric, multimeric, particulate, denatured, *etc.*). Generally, the recombinant fusion proteins of the present invention are prepared as a GST-fusion protein and/or a His-tagged fusion protein.

The proteins of the invention are preferably prepared in purified or substantially pure form (*i.e.* substantially free from host cell proteins and/or other *Chlamydia* proteins), and are generally at least about 50% pure (by weight), and usually at least about 90% pure, *i.e.* less than about 50%, and more preferably less than about 10% (*e.g.* 5%) of a composition is made up of other expressed polypeptides. Thus the antigens in the compositions are separated from the whole organism with which the molecule is expressed.

Whilst expression of the proteins of the invention may take place in *Chlamydia*, the invention preferably utilises a heterologous host. The heterologous host may be prokaryotic (*e.g.* a bacterium) or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis, Vibrio cholerae, Salmonella typhi, Salmonella typhimurium, Neisseria lactamica, Neisseria cinerea, Mycobacteria* (*e.g. M.tuberculosis*), yeasts, *etc.*

The term "polypeptide" or "protein" refers to amino acid polymers of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation,

phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, *etc.*), as well as other modifications known in the art. Polypeptides can occur as single chains or associated chains.

The invention provides polypeptides comprising a sequence -P-Q- or -Q-P-, wherein: -P- is an amino acid sequence as defined above and -Q- is not a sequence as defined above *i.e.* the invention provides fusion proteins. Where the N-terminus codon of -P- is not ATG, but this codon is not present at the N-terminus of a polypeptide, it will be translated as the standard amino acid for that codon rather than as a Met. Where this codon is at the N-terminus of a polypeptide, however, it will be translated as Met. Examples of -Q- moieties include, but are not limited to, histidine tags (*i.e.* His_n where $n = 3$, 4, 5, 6, 7, 8, 9, 10 or more), maltose-binding protein, or glutathione-S-transferase (GST).

Proteins of the invention may be attached to a solid support. They may comprise a detectable label (*e.g.* a radioactive or fluorescent label, or a biotin label).

Antibodies

The proteins of the invention induce antibodies that may be used as a vaccine capable of neutralising the activity of infectious EB. The antibodies may alternatively be used for the diagnosis of Chlamydia infection. Thus, the invention provides antibodies for use in the treatment, prevention or diagnosis of Chlamydia infection. Preferably, the infection is by *C. trachomatis*, but may alternatively be by *C. psittaci, C. pecorum, C. muridarum* or *C. suis*.

The term "antibody" includes intact immunoglobulin molecules, as well as fragments thereof which are capable of binding an antigen. These include hybrid (chimeric) antibody molecules (Winter *et al.*, (1991) *Nature* 349:293-99; US 4,816,567); F(ab')2 and F(ab) fragments and Fv molecules; non-covalent heterodimers (Inbar *et al.*, (1972) *Proc. Natl. Acad. Sci.* U.S.A. 69:2659-62; Ehrlich *et al.*, (1980) *Biochem* 19:4091-96); single-chain Fv molecules (sFv) (Huston *et al.*, (1988) Proc. Natl. Acad. Sci. U.S.A. 85:5897-83); dimeric and trimeric antibody fragment constructs; minibodies Pack *et al.*, (1992) *Biochem* 31, 1579-84; Cumber *et al.*, (1992) *J. Immunology* 149B, 120-26); humanized antibody molecules (Riechmann *et al.*, (1988) *Nature* 332, 323-27; Verhoeyan *et al.*, (1988) *Science* 239, 1534-36; and GB 2,276,169); and any functional fragments obtained from such molecules, as well as antibodies obtained through non-conventional processes such as phage display. Preferably, the antibodies are monoclonal antibodies. Methods of obtaining monoclonal antibodies are well known in the art. Humanised or fully-human antibodies are preferred.

The antibodies may be polyclonal or monoclonal and may be produced by any suitable means. The antibody may include a detectable label.

Also provided is a method for preparing antibodies comprising immunising a mammal (such as a mouse or a rabbit) with a protein of the invention and obtainining polyclonal antibodies or monoclonal antibodies by conventional techniques. For example, polyclonal antisera may be

obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (*eg.* 1,000*g* for 10 minutes). Monoclonal antibodies may be prepared using the standard method of Kohler & Milstein [*Nature* (1975) 256:495-96], or a modification thereof, or by any other suitable method.

Nucleic acids

According to a further aspect, the invention provides a nucleic acid encoding a protein or antibody of the invention. In some embodiments, the nucleic acid sequence encoding a protein of the invention preferably comprises or consists of any one of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 39, 41, 43, 45, 47, 49, 51, 53, 55, 135, 139, 155, 159, 163 or 167, for example, of any one of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 39, 41, 43 or 45. In some embodiments, the nucleic acid sequence encoding a protein of the invention comprises or consists of any one of SEQ ID NOs: 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131 and 133.

The invention also provides nucleic acid comprising nucleotide sequences having sequence identity to such nucleotide sequences. Identity between sequences is preferably determined by the Smith-Waterman homology search algorithm as described above. Such nucleic acids include those using alternative codons to encode the same amino acid.

The invention also provides nucleic acid which can hybridize to these nucleic acids. Hybridization reactions can be performed under conditions of different "stringency". Conditions that increase stringency of a hybridization reaction of widely known and published in the art (*e.g.* page 7.52 of Kaplitt, *Nature Genetics* (1994) *6*:148). Examples of relevant conditions include (in order of increasing stringency): incubation temperatures of 25°C, 37°C, 50°C, 55°C and 68°C; buffer concentrations of 10 x SSC, 6 x SSC, 1 x SSC, 0.1 x SSC (where SSC is 0.15 M NaCl and 15 mM citrate buffer) and their equivalents using other buffer systems; formamide concentrations of 0%, 25%, 50%, and 75%; incubation times from 5 minutes to 24 hours; 1, 2, or more washing steps; wash incubation times of 1, 2, or 15 minutes; and wash solutions of 6 x SSC, 1 x SSC, 0.1 x SSC, or de-ionized water. Hybridization techniques and their optimization are well known in the art (*e.g.* see US patent 5,707,829, *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.* eds., 1987) Supplement 30, Kaplitt, *Nature Genetics* (1994) *6*:148, and WO 94/03622, *etc.*).

The nucleic acid may be used in hybridisation reactions (*e.g.* Northern or Southern blots, or in nucleic acid microarrays or 'gene chips') or in amplification reactions (*e.g.* PCR, SDA, SSSR, LCR, NASBA, TMA) *etc.*

The invention also provides a nucleic acid comprising sequences complementary to those described above (*e.g.* for antisense or probing, or for use as primers). In one embodiment, the nucleic acid is complementary to the full length of the nucleic acid described above.

Nucleic acid according to the invention may be labelled *e.g.* with a radioactive or fluorescent label. This is particularly useful where the nucleic acid is to be used as a primer or probe *e.g.* in PCR, LCR or TMA.

The term "nucleic acid" includes in general means a polymeric form of nucleotides of any length, which contain deoxyribonucleotides, ribonucleotides, and/or their analogs. It includes DNA, RNA, DNA/RNA hybrids. It also includes DNA or RNA analogs, such as those containing modified backbones (*e.g.* peptide nucleic acids (PNAs) or phosphorothioates) or modified bases. Thus the invention includes mRNA, ribozymes, DNA, cDNA, recombinant nucleic acids, branched nucleic acids, plasmids, vectors, probes, primers, *etc.*. Where nucleic acid of the invention takes the form of RNA, it may or may not have a 5' cap.

Nucleic acids of the invention can take various forms (*e.g.* single stranded, double stranded, vectors, primers, probes *etc.*). Unless otherwise specified or required, any embodiment of the invention that utilizes a nucleic acid may utilize both the double-stranded form and each of two complementary single-stranded forms which make up the double-stranded form. Primers and probes are generally single-stranded, as are antisense nucleic acids.

Nucleic acids of the invention are preferably prepared in substantially pure form (*i.e.* substantially free from naturally-occuring nucleic acids, particularly from chlamydial or other host cell nucleic acids), generally being at least about 50% pure (by weight), and usually at least about 90% pure.

Nucleic acids of the invention may be prepared in many ways *e.g.* by chemical synthesis (*e.g.* phosphoramidite synthesis of DNA) in whole or in part, by digesting longer nucleic acids using nucleases (*e.g.* restriction enzymes), by joining shorter nucleic acids or nucleotides (*e.g.* using ligases or polymerases), from genomic or cDNA libraries, *etc.*

The invention provides vectors comprising nucleotide sequences of the invention (*e.g.* cloning or expression vectors) and host cells transformed with such vectors. Nucleic acids of the invention may be part of a vector *i.e.* part of a nucleic acid construct designed for transduction/transfection of one or more cell types. Vectors may be, for example, "cloning vectors" which are designed for isolation, propagation and replication of inserted nucleotides, "expression vectors" which are designed for expression of a nucleotide sequence in a host cell, "viral vectors" which are designed to result in the production of a recombinant virus or virus-like particle, or "shuttle vectors", which comprise the attributes of more than one type of vector. Preferred vectors are plasmids.

Also provided is a host cell comprising a nucleic acid of the invention. A "host cell" includes an individual cell or cell culture which can be or has been a recipient of exogenous nucleic acid. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or change. Host cells include cells transfected or infected *in vivo* or *in vitro* with nucleic acid of the invention, for example, with a vector of the invention.

Where a nucleic acid is DNA, it will be appreciated that "U" in a RNA sequence will be replaced by "T" in the DNA. Similarly, where a nucleic acid is RNA, it will be appreciated that "T" in a DNA sequence will be replaced by "U" in the RNA.

The term "complement" or "complementary" when used in relation to nucleic acids refers to Watson-Crick base pairing. Thus the complement of C is G, the complement of G is C, the complement of A is T (or U), and the complement of T (or U) is A. It is also possible to use bases such as I (the purine inosine) *e.g.* to complement pyrimidines (C or T).

Nucleic acids of the invention can be used, for example: to produce polypeptides; as hybridization probes for the detection of nucleic acid in biological samples; to generate additional copies of the nucleic acids; to generate ribozymes or antisense oligonucleotides; as single-stranded DNA primers or probes; or as triple-strand forming oligonucleotides.

The invention provides a process for producing nucleic acid of the invention, wherein the nucleic acid is synthesised in part or in whole using chemical means.

For certain embodiments of the invention, nucleic acids are preferably at least 24 nucleotides in length (*e.g.* 60, 120, 240, 390, 540, 720, 900, 1200, 1320, 1500, 1800, 2100, 2400, 2415 nucleotides or longer).

For certain embodiments of the invention, nucleic acids are preferably at most 2430 nucleotides in length (*e.g.* 2427, 2394, 2250, 2034, 1450, 1300, 1150, 1000, 850, 700, 500 nucleotides or shorter).

Primers and probes of the invention, and other nucleic acids used for hybridization, are preferably between 10 and 30 nucleotides in length (*e.g.* 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides).

Immunogenic compositions and medicaments

The protein, antibody, and/or nucleic acid or medicament may be in the form of a composition. These compositions may be suitable as immunogenic compositions (*e.g.* vaccines), or as diagnostic reagents.

Preferably, the composition is an immunogenic composition. It is particularly advantageous to use a protein of the invention in an immunogenic composition such as a vaccine. It is also envisaged that the immunogenic composition may comprise a nucleic acid which encodes a protein of the invention such that the protein is generated *in vivo*.

An immunogenic composition of the invention comprises a protein, antibody, nucleic acid, vector and/or host cell according to the invention. Immunogenic compositions according to the invention may either be prophylactic (*i.e.* to prevent infection) or therapeutic (*i.e.* to treat infection), but will typically be prophylactic. Where the immunogenic composition is for prophylactic use, the human is preferably a child (*e.g.* a toddler or infant) or a teenager; where the immunogenic composition is for

therapeutic use, the human is preferably a teenager or an adult. An immunogenic composition intended for children may also be administered to adults *e.g.* to assess safety, dosage, immunogenicity, *etc.*

In some embodiments, the immunogenic composition is for treatment or prevention of Chlamydia infection or an associated condition (*e.g.* trachoma, blindness, cervicitis, pelvic inflammatory disease, infertility, ectopic pregnancy, chronic pelvic pain, salpingitis, urethritis, epididymitis, infant pneumonia, patients infected with cervical squamous cell carcinoma, and/or HIV infection, *etc.*), preferably, *C. trachomatis* infection. The immunogenic composition may be effective against *C. pneumoniae*.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of the protein of the invention, as well as any other components, as needed. By 'immunologically effective amount', it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, age, the taxonomic group of the individual to be treated (*e.g.* non-human primate, primate, *etc.*), the capacity of the individual's immune system to synthesise antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Antigens in the composition will typically be present at a concentration of at least 1μg/ml each.

In general, the concentration of any given antigen will be sufficient to elicit an immune response against that antigen.

Dosage treatment can be a single dose schedule or a multiple dose schedule. Multiple doses may be used in a primary immunisation schedule and/or in a booster immunisation schedule. In a multiple dose schedule the various doses may be given by the same or different routes *e.g.* a parenteral prime and mucosal boost, a mucosal prime and parenteral boost, *etc.* Multiple doses will typically be administered at least 1 week apart (*e.g.* about 2 weeks, about 3 weeks, about 4 weeks, about 6 weeks, about 8 weeks, about 10 weeks, about 12 weeks, about 16 weeks, *etc.*). In some embodiments, three or more doses are provided (for example, three, four or five) doses. In some embodiments, three doses are given intramuscularly at 2 week-intervals, for example, three doses of 10-20 μg of each protein, at 2 week-intervals, given intramuscularly.

The pH of an immunogenic composition is preferably between 6 and 8, preferably about 7. pH may be maintained by the use of a buffer. The composition may be sterile and/or pyrogen-free. The composition may be isotonic with respect to humans.

Immunogenic compositions of the invention will generally be administered directly to a patient. Direct delivery may be accomplished by parenteral injection (*e.g.* subcutaneously, intraperitoneally, intravenously, intramuscularly, or to the interstitial space of a tissue), or mucosally, such as by rectal, oral (*e.g.* tablet, spray), vaginal, topical, transdermal (*See e.g.* WO99/27961) or transcutaneous (See *e.g.* WO02/074244 and WO02/064162), intranasal (See *e.g.* WO03/028760), ocular, aural, pulmonary or other mucosal administration.

Chlamydia infections affect various areas of the body and so the immunogenic compositions of the invention may be prepared in various forms. For example, the compositions may be prepared as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared (*e.g.* a lyophilised composition). The composition may be prepared for topical administration *e.g.* as an ointment, cream or powder. The composition may be prepared for oral administration *e.g.* as a tablet or capsule, or as a syrup (optionally flavoured). The composition may be prepared for pulmonary administration *e.g.* as an inhaler, using a fine powder or a spray. The composition may be prepared as a suppository or pessary. The composition may be prepared for nasal, aural or ocular administration *e.g.* as drops.

The invention also provides a delivery device pre-filled with an immunogenic composition of the invention.

The invention also provides a kit comprising a first component and a second component wherein neither the first component nor the second component is a composition of the invention as described herein, but wherein the first component and the second component can be combined to provide a composition of the invention as described herein. The kit may further include a third component comprising one or more of the following: instructions, syringe or other delivery device, adjuvant, or pharmaceutically acceptable formulating solution.

A composition as described above may alternatively and/or additionally be used for diagnosis of chlamydia infection.

Combinations with other antigens

The therapeutic or diagnostic efficiency of a Chlamydia antigen may be improved by combination with a different Chlamydia antigen. For example, the immunogenicity of a protein of the invention may be improved by comaintion with antoher protein of the invention or with another known Chlamydia antigen. The invention thus includes an immunogenic composition comprising a combination of *Chlamydia* antigens, said combination comprising a protein of the invention in combination with one or more additional Chlamydia antigens. The one or more additional Chlamydia antigens that are present in the composition may be in the form of a protein or nucleic acid or any other suitable form. A protein of the invention may be combined with one or more (e.g. 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more) different proteins of the invention and/or with one or more (e.g. 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more) other known Chlamydia antigens. For example, an

immunogenic composition is provided comprising two or more (e.g. 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 or more) proteins of the invention. The proteins of the invention may alternatively and/or additionally be provided in the composition in the form of their corresponding nucleic acids, vectors, host cells, etc. Also provided is a protein or nucleic acid of the invention for a use as described above, wherein the protein or nucleic acid is for use in combination with one or more additional Chlamydia antigens (or their encoding nucleic acids). The one or more additional antigens (e.g. 2, 3, 4, 5, 6, 7 or more additional antigens) may be administered simultaneously, separately or sequentially with the protein or nucleic acid of the invention, for example as a combined preparation.

Likewise, the antibodies of the invention may be used in combination with one or more antibodies specific for one or more additional Chlamydia antigens for use in diagnosis of Chlamydia infections.

In one embodiment, one or more of the additional Chlamydia antigens is selected from the antigens presented in Table 2, or their variants. For example, one or more (for example, all) of the additional antigens are selected from the *Chlamydia trachomatis* antigens listed in Table 2, but may alternatively or additionally be selected from the *Chlamydia pneumoniae* antigens listed in Table 2. In some embodiments, the one or more (for example, all) of the additional antigens are selected from the *Chlamydia trachomatis* antigens and/or *Chlamydia pneumoniae* antigens listed in Table 2 and CT387, CT812, CT869, CT166, CT175, CT163, CT214, CT721 and CT127. In one embodiment, one or more of the one or more additional antigens are selected from CT372, CT443, CT043, CT153, CT279, CT601, CT711, CT114, CT480, CT456, CT381, CT089, CT734, CT016, CT600, CT823, CT387, CT812, CT869, CT166, CT175, CT163, CT214, CT721 and CT127 (or their variants), for example, from CT372, CT443, CT043, CT153, CT279, CT601, CT711, CT114, CT480, CT456, CT381, CT089, CT734, CT016, CT600 and CT823. These additional antigens are listed in Table 2 and their sequences are set out in the "Sequences" section that follows Table 2. In one embodiment, one or more proteins of the invention is combined with CT089. In another embodiment, one or more proteins of the invention is combined with CT089 and CT381 (or their variants). In some embodiments, the C-terminal fragment of CT812 "CT812C" (for example, a protein comprising or consisting of the amino acid sequence set out in SEQ ID NO:122 or a fragment or variant thereof) is used instead of full length CT812.

In some embodiments, the following combinations of antigens (or their variants) are used: CT733+CT601, CT733+CT279, CT733+CT443, CT733+CT372, CT733+CT456, CT733+CT381, CT153+CT601, CT153+CT279, CT153+CT443, CT153+CT372, CT153+CT456, CT153+CT381, CT601+CT443, CT601+CT372, CT601+CT456, CT601+CT381, CT279+CT443, CT279+CT372, CT279+CT456, CT279+CT381, CT443+CT372, CT443+CT456, CT443+CT381, CT372+CT456, CT372+CT381, CT387+CT812+CT869, CT387+CT812C+CT869. These combinations may be used in the absence of any other chlamydia antigens or in the presence of one or more additional chlamydia antigens. Particularly preferred combinations are: (i) CT279+CT601; (ii) CT372+CT443; (iii) CT733+CT153; (iv) CT456+CT381; (v) CT279+ CT601+CT733+CT153; (vi)

CT279+CT601+CT372+CT443; (vii) CT823+CT733+CT043+ CT456; (viii) CT387+CT812+CT869; and (ix) CT387+CT812C+CT869 (or their variants).

The human serovariants ("serovars") of *C. trachomatis* are divided into two biovariants ("biovars"). Serovars A-K elicit epithelial infections primarily in the ocular tissue (A-C) or urogenital tract (D-K). Serovars L1, L2 and L3 are the agents of invasive lymphogranuloma venereum (LGV). In some embodiments, one or more of the additional Chlamydial antigens may, for example, be of any of Serovars A-K or L1, L2 or L3. One or more of the additional Chlamydia antigens is preferably from *C. trachomatis* serovar D, or from another epidemiologically prevalent serotype.

In some embodiments, one or more of the additional Chlamydia antigens is a homologous antigen from *C. pneumoniae*, *C. psittaci, C. pecorum, C. muridarum* or *C. suis.*

In some embodiments, TC0551 (the *C. muridarum* homologue of CT279) is used in place of the *C. trachomatis* protein. *C. muridarum* is the mouse adapted strain of *Chlamydia trachomatis.* Although *C. muridarum* is not a human pathogen, infection of mice with *C. muridarum* phenotypically mimics many aspects of *C. trachomatis* infection in humans and is frequently used to measure immunoprotective responses against *C. trachomatis.* In some embodiments, TC0890 (the *C. muridarum* homologue of CT601) is used in place of the *C. trachomatis* protein. In some embodiments, TC0651 (the *C. muridarum* homologue of CT372) is used in place of the *C. trachomatis* protein. In some embodiments, TC0727 (the *C. muridarum* homologue of CT443) is used in place of the *C. trachomatis* protein. In some embodiments, TC0106 (the *C. muridarum* homologue of CT733) is used in place of the *C. trachomatis* protein. In some embodiments, TC0431 (the *C. muridarum* homologue of CT153) is used in place of the *C. trachomatis* protein. In some embodiments, TC0660 (the *C. muridarum* homologue of CT381) is used in place of the *C. trachomatis* protein. In some embodiments, TC0741 (the *C. muridarum* homologue of CT456) is used in place of the *C. trachomatis* protein. In some embodiments, TC0210 (the *C. muridarum* homologue of CT823) is used in place of the *C. trachomatis* protein. In some embodiments, TC0666 (the *C. muridarum* homologue of CT387) is used in place of the *C. trachomatis* protein. TC0666 is annotated as a hypothetical protein. In some embodiments, TC0197 (the *C. muridarum* homologue of CT812) is used in place of the *C. trachomatis* protein. TC0197 is annotated as polymorphic membrane protein D family protein. In some embodiments, TC0261 (the *C. muridarum* homologue of CT869) is used in place of the *C. trachomatis* protein. TC0261 is annotated as polymorphic membrane protein E/F family protein. In some embodiments, TC0313 (the *C. muridarum* homologue of CT043) is used in place of the *C. trachomatis* protein. In some embodiments, TC0889 (the *C. muridarum* homologue of CT600) is used in place of the *C. trachomatis* protein. In some embodiments, TC0210 (the *C. muridarum* homologue of CT823) is used in place of the *C. trachomatis* protein. In some embodiments in which the composition comprises a single Chlamydia antigen, the *C. muridarum* homologue is used in place of the single *C. trachomatis* antigen. In some embodiments in which the composition comprises a combination of Chlamydia antigens, the *C.*

muridarum homologue is used in place of one or more (for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) or all *C. trachomatis* antigens.

Advantageous combinations of the invention are those in which two or more antigens (for example, two, three or four antigens) act synergistically. Thus, the protection against Chlamydia achieved by their combined administration exceeds that expected by mere addition of their individual protective efficacy.

In some embodiments, the one or more additional Chlamydia antigens may comprise an amino acid sequence: (a) which is a variant of a Table 2 antigen (i.e. has 50% or more identity (*e.g.* 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more) to a sequence presented in Table 2); and/or (b) comprising a fragment of at least 'n' consecutive amino acids of a sequence presented in Table 2 or of a variant of a Table 2 antigen, wherein 'n' is 7 or more (*e.g.* 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 350, 450, 550, 650, 750, 780, 800 or more). Preferred fragments of (b) comprise an epitope from a sequence presented in Table 2. Preferably, the epitope is a MHC class II epitope, for example, a CD4+ T cell epitope. Other preferred fragments lack one or more amino acids (*e.g.* 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or more) from the C-terminus and/or one or more amino acids (*e.g.* 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 or more) from the N-terminus of a sequence presented in Table 2, while retaining at least one epitope of a sequence presented in Table 2. Other fragments omit one or more protein domains. When an additional Chlamydia antigen comprises a sequence that is not identical to a complete sequence from Table 2 (e.g. when it comprises a sequence with less than 100% sequence identity thereto, or when it comprises a fragment thereof), it is preferred in each individual instance that the additional Chlamydia antigen can elicit an antibody that recognises a protein having the complete sequence from the Table 2 antigen from which it is derived.

In some embodiments, the combination of two or more chlamydia antigens is provided as a combined preparation for simultaneous, separate or sequential administration. The invention also provides a kit comprising a protein of the invention and one or more additional antigens for simultaneous, separate or sequential administration.

The *Chlamydia* antigens used in the invention may be present in the composition as individual separate polypeptides. Alternatively, the combination may be present as a hybrid polypeptide in which two or more (*i.e.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more) of the antigens are expressed as a single polypeptide chain. Hybrid polypeptides offer two principal advantages: first, a polypeptide that may be unstable or poorly expressed on its own can be assisted by adding a suitable hybrid partner that overcomes the problem; second, commercial manufacture is simplified as only one expression and purification need be employed in order to produce two polypeptides which are both antigenically useful. Different hybrid polypeptides may be mixed together in a single formulation. Within such combinations, a *Chlamydia trachomatis* antigen may

be present in more than one hybrid polypeptide and/or as a non-hybrid polypeptide. It is preferred, however, that an antigen is present either as a hybrid or as a non-hybrid, but not as both.

Hybrid polypeptides can be represented by the formula NH_2-A-{-X-L-}$_n$-B-COOH, wherein: at least one X is an amino acid sequence of a Chlamydia protein according to the invention as described above; L is an optional linker amino acid sequence; A is an optional N-terminal amino acid sequence; B is an optional C-terminal amino acid sequence; *n* is an integer of 2 or more (*e.g.* 2, 3, 4, 5, 6, *etc.*). Usually *n* is 2 or 3.

If a -X- moiety has a leader peptide sequence in its wild-type form, this may be included or omitted in the hybrid protein. In some embodiments, the leader peptides will be deleted except for that of the -X- moiety located at the N-terminus of the hybrid protein *i.e.* the leader peptide of X_1 will be retained, but the leader peptides of X_2 ... X_n will be omitted. This is equivalent to deleting all leader peptides and using the leader peptide of X_1 as moiety -A-.

For each *n* instances of {-X-L-}, linker amino acid sequence -L- may be present or absent. For instance, when *n*=2 the hybrid may be NH_2-X_1-L_1-X_2-L_2-COOH, NH_2-X_1-X_2-COOH, NH_2-X_1-L_1-X_2-COOH, NH_2-X_1-X_2-L_2-COOH, *etc.* Linker amino acid sequence(s) -L- will typically be short (*e.g.* 20 or fewer amino acids *i.e.* 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples comprise short peptide sequences which facilitate cloning, poly-glycine linkers (*i.e.* comprising Gly_n where *n* = 2, 3, 4, 5, 6, 7, 8, 9, 10 or more), and histidine tags (*i.e.* His_n where *n* = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable linker amino acid sequences will be apparent to those skilled in the art. A useful linker is GSGGGG, with the Gly-Ser dipeptide being formed from a *Bam*HI restriction site, thus aiding cloning and manipulation, and the $(Gly)_4$ tetrapeptide being a typical poly-glycine linker.

-A- is an optional N-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include leader sequences to direct protein trafficking, or short peptide sequences which facilitate cloning or purification (*e.g.* histidine tags *i.e.* His_n where *n* = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable N-terminal amino acid sequences will be apparent to those skilled in the art. If X_1 lacks its own N-terminus methionine, -A- is preferably an oligopeptide (*e.g.* with 1, 2, 3, 4, 5, 6, 7 or 8 amino acids) which provides a N-terminus methionine.

-B- is an optional C-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include sequences to direct protein trafficking, short peptide sequences which facilitate cloning or purification (*e.g.* comprising histidine tags *i.e.* His_n where *n* = 3, 4, 5, 6, 7, 8, 9, 10 or more), or sequences which enhance protein stability. Other suitable C-terminal amino acid sequences will be apparent to those skilled in the art.

Where hybrid polypeptides are used, the individual antigens within the hybrid (*i.e.* individual -X- moieties) may be from one or more strains. Where n=2, for instance, X_2 may be from the same strain as X_1 or from a different strain. Where n=3, the strains might be (i) $X_1=X_2=X_3$ (ii) $X_1=X_2\neq X_3$ (iii) $X_1\neq X_2=X_3$ (iv) $X_1\neq X_2\neq X_3$ or (v) $X_1=X_3\neq X_2$, *etc.*

The invention also provides a nucleic acid encoding a hybrid polypeptide of the invention. Furthermore, the invention provides a nucleic acid which can hybridise to this nucleic acid, preferably under "high stringency" conditions (*e.g.* 65°C in a 0.1 x SSC, 0.5% SDS solution).

Further components of the composition

Compositions may thus be pharmaceutically acceptable. They will usually include components in addition to the antigens *e.g.* they typically include one or more pharmaceutical carrier(s) and/or excipient(s). A thorough discussion of such components is available in *Remington The Science and Practice of Pharmacy.*

Compositions will generally be administered to a mammal in aqueous form. Prior to administration, however, the composition may have been in a non-aqueous form. For instance, although some vaccines are manufactured in aqueous form, then filled and distributed and administered also in aqueous form, other vaccines are lyophilised during manufacture and are reconstituted into an aqueous form at the time of use. Thus a composition of the invention may be dried, such as a lyophilised formulation.

The composition may include preservatives such as thiomersal or 2-phenoxyethanol. It is preferred, however, that the vaccine should be substantially free from (*i.e.* less than 5μg/ml) mercurial material *e.g.* thiomersal-free. Vaccines containing no mercury are more preferred. Preservative-free vaccines are particularly preferred.

To control tonicity, it is preferred to include a physiological salt, such as a sodium salt. Sodium chloride (NaCl) is preferred, which may be present at between 1 and 20 mg/ml *e.g.* about 10±2mg/ml NaCl. Other salts that may be present include potassium chloride, potassium dihydrogen phosphate, disodium phosphate dehydrate, magnesium chloride, calcium chloride, *etc.*

Compositions will generally have an osmolality of between 200 mOsm/kg and 400 mOsm/kg, preferably between 240-360 mOsm/kg, and will more preferably fall within the range of 290-310 mOsm/kg.

Compositions may include one or more buffers. Typical buffers include: a phosphate buffer; a Tris buffer; a borate buffer; a succinate buffer; a histidine buffer (particularly with an aluminum hydroxide adjuvant); or a citrate buffer. Buffers will typically be included in the 5-20mM range.

The pH of a composition will generally be between 5.0 and 8.1, and more typically between 6.0 and 8.0 *e.g.* 6.5 and 7.5, or between 7.0 and 7.8.

The composition is preferably sterile. The composition is preferably non-pyrogenic *e.g.* containing <1 EU (endotoxin unit, a standard measure) per dose, and preferably <0.1 EU per dose. The composition is preferably gluten free.

The composition may include material for a single immunisation, or may include material for multiple immunisations (*i.e.* a 'multidose' kit). The inclusion of a preservative is preferred in multidose arrangements. As an alternative (or in addition) to including a preservative in multidose compositions, the compositions may be contained in a container having an aseptic adaptor for removal of material.

Human vaccines are typically administered in a dosage volume of about 0.5ml, although a half dose (*i.e.* about 0.25ml) may be administered to children.

Immunogenic compositions of the invention may also comprise one or more immunoregulatory agents. Preferably, one or more of the immunoregulatory agents include one or more adjuvants. The adjuvants may include a TH1 adjuvant and/or a TH2 adjuvant, further discussed below.

Adjuvants which may be used in compositions of the invention include, but are not limited to:

A. Mineral-containing compositions

Mineral containing compositions suitable for use as adjuvants in the invention include mineral salts, such as aluminium salts and calcium salts (or mixtures thereof). Calcium salts include calcium phosphate (*e.g.* the "CAP" particles disclosed in US patent 6355271). Aluminum salts include hydroxides, phosphates, sulfates, *etc.*, with the salts taking any suitable form (*e.g.* gel, crystalline, amorphous, *etc.*). Adsorption to these salts is preferred. The mineral containing compositions may also be formulated as a particle of metal salt [WO00/23105].

The adjuvants known as aluminum hydroxide and aluminum phosphate may be used. These names are conventional, but are used for convenience only, as neither is a precise description of the actual chemical compound which is present (*e.g.* see chapter 9 of *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum). The invention can use any of the "hydroxide" or "phosphate" adjuvants that are in general use as adjuvants. The adjuvants known as "aluminium hydroxide" are typically aluminium oxyhydroxide salts, which are usually at least partially crystalline. The adjuvants known as "aluminium phosphate" are typically aluminium hydroxyphosphates, often also containing a small amount of sulfate (*i.e.* aluminium hydroxyphosphate sulfate). They may be obtained by precipitation, and the reaction conditions and concentrations during precipitation influence the degree of substitution of phosphate for hydroxyl in the salt.

A fibrous morphology (*e.g.* as seen in transmission electron micrographs) is typical for aluminium hydroxide adjuvants. The pI of aluminium hydroxide adjuvants is typically about 11 *i.e.* the adjuvant itself has a positive surface charge at physiological pH. Adsorptive capacities of between 1.8-2.6 mg protein per mg Al^{+++} at pH 7.4 have been reported for aluminium hydroxide adjuvants.

Aluminium phosphate adjuvants generally have a PO_4/Al molar ratio between 0.3 and 1.2, preferably between 0.8 and 1.2, and more preferably 0.95±0.1. The aluminium phosphate will generally be amorphous, particularly for hydroxyphosphate salts. A typical adjuvant is amorphous aluminium hydroxyphosphate with PO_4/Al molar ratio between 0.84 and 0.92, included at 0.6mg Al^{3+}/ml. The aluminium phosphate will generally be particulate (*e.g.* plate-like morphology as seen in transmission electron micrographs). Typical diameters of the particles are in the range 0.5-20μm (*e.g.* about 5-10μm) after any antigen adsorption. Adsorptive capacities of between 0.7-1.5 mg protein per mg Al^{+++} at pH 7.4 have been reported for aluminium phosphate adjuvants.

The point of zero charge (PZC) of aluminium phosphate is inversely related to the degree of substitution of phosphate for hydroxyl, and this degree of substitution can vary depending on reaction conditions and concentration of reactants used for preparing the salt by precipitation. PZC is also altered by changing the concentration of free phosphate ions in solution (more phosphate = more acidic PZC) or by adding a buffer such as a histidine buffer (makes PZC more basic). Aluminium phosphates used according to the invention will generally have a PZC of between 4.0 and 7.0, more preferably between 5.0 and 6.5 *e.g.* about 5.7.

Suspensions of aluminium salts used to prepare compositions of the invention may contain a buffer (*e.g.* a phosphate or a histidine or a Tris buffer), but this is not always necessary. The suspensions are preferably sterile and pyrogen-free. A suspension may include free aqueous phosphate ions *e.g.* present at a concentration between 1.0 and 20 mM, preferably between 5 and 15 mM, and more preferably about 10 mM. The suspensions may also comprise sodium chloride.

The invention can use a mixture of both an aluminium hydroxide and an aluminium phosphate. In this case there may be more aluminium phosphate than hydroxide *e.g.* a weight ratio of at least 2:1 *e.g.* ≥5:1, ≥6:1, ≥7:1, ≥8:1, ≥9:1, *etc.*

The concentration of Al^{+++} in a composition for administration to a patient is preferably less than 10mg/ml *e.g.* ≤5 mg/ml, ≤4 mg/ml, ≤3 mg/ml, ≤2 mg/ml, ≤1 mg/ml, *etc.* A preferred range is between 0.3 and 1mg/ml. A maximum of 0.85mg/dose is preferred.

Aluminium phosphates are particularly preferred, particularly in compositions which include a *H.influenzae* saccharide antigen, and a typical adjuvant is amorphous aluminium hydroxyphosphate with PO_4/Al molar ratio between 0.84 and 0.92, included at 0.6mg Al^{3+}/ml. Adsorption with a low dose of aluminium phosphate may be used *e.g.* between 50 and 100μg Al^{3+} per conjugate per dose. Where there is more than one conjugate in a composition, not all conjugates need to be adsorbed.

B. Oil Emulsions

Oil emulsion compositions suitable for use as adjuvants in the invention include squalene-water emulsions, such as MF59 [Chapter 10 of *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum; see also WO90/14837] (5% Squalene, 0.5% Tween 80, and 0.5% Span 85,

formulated into submicron particles using a microfluidizer). Complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA) may also be used.

Various oil-in-water emulsion adjuvants are known, and they typically include at least one oil and at least one surfactant, with the oil(s) and surfactant(s) being biodegradable (metabolisable) and biocompatible. The oil droplets in the emulsion are generally less than 5μm in diameter, and ideally have a sub-micron diameter, with these small sizes being achieved with a microfluidiser to provide stable emulsions. Droplets with a size less than 220nm are preferred as they can be subjected to filter sterilization.

The emulsion can comprise oils such as those from an animal (such as fish) or vegetable source. Sources for vegetable oils include nuts, seeds and grains. Peanut oil, soybean oil, coconut oil, and olive oil, the most commonly available, exemplify the nut oils. Jojoba oil can be used *e.g.* obtained from the jojoba bean. Seed oils include safflower oil, cottonseed oil, sunflower seed oil, sesame seed oil and the like. In the grain group, corn oil is the most readily available, but the oil of other cereal grains such as wheat, oats, rye, rice, teff, triticale and the like may also be used. 6-10 carbon fatty acid esters of glycerol and 1,2-propanediol, while not occurring naturally in seed oils, may be prepared by hydrolysis, separation and esterification of the appropriate materials starting from the nut and seed oils. Fats and oils from mammalian milk are metabolizable and may therefore be used in the practice of this invention. The procedures for separation, purification, saponification and other means necessary for obtaining pure oils from animal sources are well known in the art. Most fish contain metabolizable oils which may be readily recovered. For example, cod liver oil, shark liver oils, and whale oil such as spermaceti exemplify several of the fish oils which may be used herein. A number of branched chain oils are synthesized biochemically in 5-carbon isoprene units and are generally referred to as terpenoids. Shark liver oil contains a branched, unsaturated terpenoids known as squalene, 2,6,10,15,19,23-hexamethyl-2,6,10,14,18,22-tetracosahexaene, which is particularly preferred herein. Squalane, the saturated analog to squalene, is also a preferred oil. Fish oils, including squalene and squalane, are readily available from commercial sources or may be obtained by methods known in the art. Other preferred oils are the tocopherols (see below). Mixtures of oils can be used.

Surfactants can be classified by their 'HLB' (hydrophile/lipophile balance). Preferred surfactants of the invention have a HLB of at least 10, preferably at least 15, and more preferably at least 16. The invention can be used with surfactants including, but not limited to: the polyoxyethylene sorbitan esters surfactants (commonly referred to as the Tweens), especially polysorbate 20 and polysorbate 80; copolymers of ethylene oxide (EO), propylene oxide (PO), and/or butylene oxide (BO), sold under the DOWFAX™ tradename, such as linear EO/PO block copolymers; octoxynols, which can vary in the number of repeating ethoxy (oxy-1,2-ethanediyl) groups, with octoxynol-9 (Triton X-100, or t-octylphenoxypolyethoxyethanol) being of particular interest; (octylphenoxy)polyethoxyethanol (IGEPAL CA-630/NP-40); phospholipids such as phosphatidylcholine (lecithin); nonylphenol

ethoxylates, such as the Tergitol™ NP series; polyoxyethylene fatty ethers derived from lauryl, cetyl, stearyl and oleyl alcohols (known as Brij surfactants), such as triethyleneglycol monolauryl ether (Brij 30); and sorbitan esters (commonly known as the SPANs), such as sorbitan trioleate (Span 85) and sorbitan monolaurate. Non-ionic surfactants are preferred. Preferred surfactants for including in the emulsion are Tween 80 (polyoxyethylene sorbitan monooleate), Span 85 (sorbitan trioleate), lecithin and Triton X-100.

Mixtures of surfactants can be used *e.g.* Tween 80/Span 85 mixtures. A combination of a polyoxyethylene sorbitan ester such as polyoxyethylene sorbitan monooleate (Tween 80) and an octoxynol such as t-octylphenoxypolyethoxyethanol (Triton X-100) is also suitable. Another useful combination comprises laureth 9 plus a polyoxyethylene sorbitan ester and/or an octoxynol.

Preferred amounts of surfactants (% by weight) are: polyoxyethylene sorbitan esters (such as Tween 80) 0.01 to 1%, in particular about 0.1 %; octyl- or nonylphenoxy polyoxyethanols (such as Triton X-100, or other detergents in the Triton series) 0.001 to 0.1 %, in particular 0.005 to 0.02%; polyoxyethylene ethers (such as laureth 9) 0.1 to 20 %, preferably 0.1 to 10 % and in particular 0.1 to 1 % or about 0.5%.

Preferred emulsion adjuvants have an average droplets size of <1μm *e.g.* ≤750nm, ≤500nm, ≤400nm, ≤300nm, ≤250nm, ≤220nm, ≤200nm, or smaller. These droplet sizes can conveniently be achieved by techniques such as microfluidisation.

Specific oil-in-water emulsion adjuvants useful with the invention include, but are not limited to:

- A submicron emulsion of squalene, Tween 80, and Span 85. The composition of the emulsion by volume can be about 5% squalene, about 0.5% polysorbate 80 and about 0.5% Span 85. In weight terms, these ratios become 4.3% squalene, 0.5% polysorbate 80 and 0.48% Span 85. This adjuvant is known as 'MF59' (WO90/14837, Podda & Del Giudice (2003) *Expert Rev Vaccines* 2:197-203, Podda (2001) *Vaccine* 19: 2673-2680; as described in more detail in Chapter 10 of *Vaccine Design: The Subunit and Adjuvant Approach* (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X) and chapter 12 of *Vaccine Adjuvants: Preparation Methods and Research Protocols* (Volume 42 of *Methods in Molecular Medicine* series). ISBN: 1-59259-083-7. Ed. O'Hagan). The MF59 emulsion advantageously includes citrate ions *e.g.* 10mM sodium citrate buffer.

- An emulsion of squalene, a tocopherol, and Tween 80. The emulsion may include phosphate buffered saline. It may also include Span 85 (*e.g.* at 1%) and/or lecithin. These emulsions may have from 2 to 10% squalene, from 2 to 10% tocopherol and from 0.3 to 3% Tween 80, and the weight ratio of squalene:tocopherol is preferably ≤1 as this provides a more stable emulsion. Squalene and Tween 80 may be present volume ratio of about 5:2. One such emulsion can be made by dissolving Tween 80 in PBS to give a 2% solution, then mixing 90ml of this solution with a mixture of (5g of DL-α-tocopherol and 5ml squalene), then microfluidising the mixture.

The resulting emulsion may have submicron oil droplets *e.g.* with an average diameter of between 100 and 250nm, preferably about 180nm.

- An emulsion of squalene, a tocopherol, and a Triton detergent (*e.g.* Triton X-100). The emulsion may also include a 3d-MPL (see below). The emulsion may contain a phosphate buffer.

- An emulsion comprising a polysorbate (*e.g.* polysorbate 80), a Triton detergent (*e.g.* Triton X-100) and a tocopherol (*e.g.* an α-tocopherol succinate). The emulsion may include these three components at a mass ratio of about 75:11:10 (*e.g.* 750μg/ml polysorbate 80, 110μg/ml Triton X-100 and 100μg/ml α-tocopherol succinate), and these concentrations should include any contribution of these components from antigens. The emulsion may also include squalene. The emulsion may also include a 3d-MPL (see below). The aqueous phase may contain a phosphate buffer.

- An emulsion of squalane, polysorbate 80 and poloxamer 401 ("Pluronic™ L121"). The emulsion can be formulated in phosphate buffered saline, pH 7.4. This emulsion is a useful delivery vehicle for muramyl dipeptides, and has been used with threonyl-MDP in the "SAF-1" adjuvant (Allison & Byars (1992) *Res Immunol* 143:519-25) (0.05-1% Thr-MDP, 5% squalane, 2.5% Pluronic L121 and 0.2% polysorbate 80). It can also be used without the Thr-MDP, as in the "AF" adjuvant (Hariharan *et al.* (1995) *Cancer Res* 55:3486-9) (5% squalane, 1.25% Pluronic L121 and 0.2% polysorbate 80). Microfluidisation is preferred.

- An emulsion comprising squalene, an aqueous solvent, a polyoxyethylene alkyl ether hydrophilic nonionic surfactant (*e.g.* polyoxyethylene (12) cetostearyl ether) and a hydrophobic nonionic surfactant (*e.g.* a sorbitan ester or mannide ester, such as sorbitan monoleate or 'Span 80'). The emulsion is preferably thermoreversible and/or has at least 90% of the oil droplets (by volume) with a size less than 200 nm (US-2007/014805.). The emulsion may also include one or more of: alditol; a cryoprotective agent (*e.g.* a sugar, such as dodecylmaltoside and/or sucrose); and/or an alkylpolyglycoside. Such emulsions may be lyophilized.

- An emulsion o US-2007/014805.f squalene, poloxamer 105 and Abil-Care (Suli *et al.* (2004) *Vaccine* 22(25-26):3464-9). The final concentration (weight) of these components in adjuvanted vaccines are 5% squalene, 4% poloxamer 105 (pluronic polyol) and 2% Abil-Care 85 (Bis-PEG/PPG-16/16 PEG/PPG-16/16 dimethicone; caprylic/capric triglyceride).

- An emulsion having from 0.5-50% of an oil, 0.1-10% of a phospholipid, and 0.05-5% of a non-ionic surfactant. As described in WO95/11700, preferred phospholipid components are phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, phosphatidylglycerol, phosphatidic acid, sphingomyelin and cardiolipin. Submicron droplet sizes are advantageous.

- A submicron oil-in-water emulsion of a non-metabolisable oil (such as light mineral oil) and at least one surfactant (such as lecithin, Tween 80 or Span 80). Additives may be included, such as QuilA saponin, cholesterol, a saponin-lipophile conjugate (such as GPI-0100, described in US patent 6,080,725, produced by addition of aliphatic amine to desacylsaponin via the carboxyl group of glucuronic acid), dimethyidioctadecylammonium bromide and/or N,N-dioctadecyl-N,N-bis (2-hydroxyethyl)propanediamine.

- An emulsion in which a saponin (*e.g.* QuilA or QS21) and a sterol (*e.g.* a cholesterol) are associated as helical micelles (WO2005/097181).

- An emulsion comprising a mineral oil, a non-ionic lipophilic ethoxylated fatty alcohol, and a non-ionic hydrophilic surfactant (*e.g.* an ethoxylated fatty alcohol and/or polyoxyethylene-polyoxypropylene block copolymer) (WO2006/113373).

- An emulsion comprising a mineral oil, a non-ionic hydrophilic ethoxylated fatty alcohol, and a non-ionic lipophilic surfactant (*e.g.* an ethoxylated fatty alcohol and/or polyoxyethylene-polyoxypropylene block copolymer) (Wu *et al.* (2004) *Antiviral Res.* 64(2):79-83).

In some embodiments an emulsion may be mixed with antigen extemporaneously, at the time of delivery, and thus the adjuvant and antigen may be kept separately in a packaged or distributed vaccine, ready for final formulation at the time of use. In other embodiments an emulsion is mixed with antigen during manufacture, and thus the composition is packaged in a liquid adjuvanted form,. The antigen will generally be in an aqueous form, such that the vaccine is finally prepared by mixing two liquids. The volume ratio of the two liquids for mixing can vary (*e.g.* between 5:1 and 1:5) but is generally about 1:1. Where concentrations of components are given in the above descriptions of specific emulsions, these concentrations are typically for an undiluted composition, and the concentration after mixing with an antigen solution will thus decrease. Where a composition is to be prepared extemporaneously prior to use (*e.g.* where a component is presented in lyophilised form) and is presented as a kit, the kit may comprise two vials, or it may comprise one ready-filled syringe and one vial, with the contents of the syringe being used to reactivate the contents of the vial prior to injection.

Where a composition includes a tocopherol, any of the α, β, γ, δ, ε or ξ tocopherols can be used, but α-tocopherols are preferred. The tocopherol can take several forms *e.g.* different salts and/or isomers. Salts include organic salts, such as succinate, acetate, nicotinate, *etc.* D-α-tocopherol and DL-α-tocopherol can both be used. Tocopherols are advantageously included in vaccines for use in elderly patients (*e.g.* aged 60 years or older) because vitamin E has been reported to have a positive effect on the immune response in this patient group (Han *et al.* (2005) *Impact of Vitamin E on Immune Function and Infectious Diseases in the Aged* at *Nutrition, Immune functions and Health* EuroConference, Paris, 9-10 June 2005). They also have antioxidant properties that may help to stabilize the emulsions (US-6630161). A preferred α-tocopherol is DL-α-tocopherol, and the

preferred salt of this tocopherol is the succinate. The succinate salt has been found to cooperate with TNF-related ligands *in vivo*.

C. Saponin formulations (chapter 22 of Vaccine Design... (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum)

Saponin formulations may also be used as adjuvants in the invention. Saponins are a heterogeneous group of sterol glycosides and triterpenoid glycosides that are found in the bark, leaves, stems, roots and even flowers of a wide range of plant species. Saponin from the bark of the *Quillaia saponaria* Molina tree have been widely studied as adjuvants. Saponin can also be commercially obtained from *Smilax ornata* (sarsaprilla), *Gypsophilla paniculata* (brides veil), and *Saponaria officinalis* (soap root). Saponin adjuvant formulations include purified formulations, such as QS21, as well as lipid formulations, such as ISCOMs. QS21 is marketed as Stimulon™.

Saponin compositions have been purified using HPLC and RP-HPLC. Specific purified fractions using these techniques have been identified, including QS7, QS17, QS18, QS21, QH-A, QH-B and QH-C. Preferably, the saponin is QS21. A method of production of QS21 is disclosed in US 5,057,540. Saponin formulations may also comprise a sterol, such as cholesterol (WO96/33739).

Combinations of saponins and cholesterols can be used to form unique particles called immunostimulating complexs (ISCOMs) (chapter 23 of *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum). ISCOMs typically also include a phospholipid such as phosphatidylethanolamine or phosphatidylcholine. Any known saponin can be used in ISCOMs. Preferably, the ISCOM includes one or more of QuilA, QHA & QHC. ISCOMs are further described in Podda & Del Giudice (2003) *Expert Rev Vaccines* 2:197-203; Podda (2001) *Vaccine* 19: 2673-2680; *Vaccine Design: The Subunit and Adjuvant Approach* (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X); *Vaccine Adjuvants: Preparation Methods and Research Protocols* (Volume 42 of *Methods in Molecular Medicine* series). ISBN: 1-59259-083-7. Ed. O'Hagan; Allison & Byars (1992) *Res Immunol* 143:519-25; Hariharan *et al.* (1995) *Cancer Res* 55:3486-9; US-2007/014805; Suli *et al.* (2004) *Vaccine* 22(25-26):3464-9; WO95/11700; US patent 6,080,725; WO2005/097181; WO2006/113373; Han *et al.* (2005) *Impact of Vitamin E on Immune Function and Infectious Diseases in the Aged* at *Nutrition, Immune functions and Health* EuroConference, Paris, 9-10 June 2005; US- 6630161; US 5,057,540; WO96/33739; EP-A-0109942; and WO96/11711. Optionally, the ISCOMS may be devoid of additional detergent (WO00/07621).

A review of the development of saponin based adjuvants can be found in Barr *et al.* (1998) *Advanced Drug Delivery Reviews* 32:247-271 and Sjolanderet *et al.* (1998) *Advanced Drug Delivery Reviews* 32:321-338.

D. Virosomes and virus-like particles

Virosomes and virus-like particles (VLPs) can also be used as adjuvants in the invention. These structures generally contain one or more proteins from a virus optionally combined or formulated with a phospholipid. They are generally non-pathogenic, non-replicating and generally do not contain

any of the native viral genome. The viral proteins may be recombinantly produced or isolated from whole viruses. These viral proteins suitable for use in virosomes or VLPs include proteins derived from influenza virus (such as HA or NA), Hepatitis B virus (such as core or capsid proteins), Hepatitis E virus, measles virus, Sindbis virus, Rotavirus, Foot-and-Mouth Disease virus, Retrovirus, Norwalk virus, human Papilloma virus, HIV, RNA-phages, Qß-phage (such as coat proteins), GA-phage, fr-phage, AP205 phage, and Ty (such as retrotransposon Ty protein p1). VLPs are discussed further in Niikura *et al.* (2002) *Virology* 293:273-280; Lenz *et al.* (2001) *J Immunol* 166:5346-5355; Pinto *et al.* (2003) *J Infect Dis* 188:327-338; Gerber *et al.* (2001) *J Virol* 75:4752-4760; WO03/024480 and WO03/024481. Virosomes are discussed further in, for example, Gluck *et al.* (2002) *Vaccine* 20:B10-B16.

E. Bacterial or microbial derivatives

Adjuvants suitable for use in the invention include bacterial or microbial derivatives such as non-toxic derivatives of enterobacterial lipopolysaccharide (LPS), Lipid A derivatives, immunostimulatory oligonucleotides and ADP-ribosylating toxins and detoxified derivatives thereof.

Non-toxic derivatives of LPS include monophosphoryl lipid A (MPL) and 3-O-deacylated MPL (3dMPL). 3dMPL is a mixture of 3 de-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains. A preferred "small particle" form of 3 De-O-acylated monophosphoryl lipid A is disclosed in EP-A-0689454. Such "small particles" of 3dMPL are small enough to be sterile filtered through a 0.22μm membrane (US- 6630161). Other non-toxic LPS derivatives include monophosphoryl lipid A mimics, such as aminoalkyl glucosaminide phosphate derivatives *e.g.* RC-529 (Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278; and Evans *et al.* (2003) *Expert Rev Vaccines* 2:219-229).

Lipid A derivatives include derivatives of lipid A from *Escherichia coli* such as OM-174. OM-174 is described for example in Meraldi *et al.* (2003) *Vaccine* 21:2485-2491 and Pajak *et al.* (2003) *Vaccine* 21:836-842.

Immunostimulatory oligonucleotides suitable for use as adjuvants in the invention include nucleotide sequences containing a CpG motif (a dinucleotide sequence containing an unmethylated cytosine linked by a phosphate bond to a guanosine). Double-stranded RNAs and oligonucleotides containing palindromic or poly(dG) sequences have also been shown to be immunostimulatory.

The CpG's can include nucleotide modifications/analogs such as phosphorothioate modifications and can be double-stranded or single-stranded. Kandimalla *et al.* (2003) *Nucleic Acids Research* 31:2393-2400, WO02/26757 and WO99/62923 disclose possible analog substitutions *e.g.* replacement of guanosine with 2'-deoxy-7-deazaguanosine. The adjuvant effect of CpG oligonucleotides is further discussed in Krieg (2003) *Nature Medicine* 9:831-835; McCluskie *et al.* (2002) *FEMS Immunology and Medical Microbiology* 32:179-185; WO98/40100; US 6,207,646; US 6,239,116 and US 6,429,199.

The CpG sequence may be directed to TLR9, such as the motif GTCGTT or TTCGTT (Kandimalla *et al.* (2003) *Biochemical Society Transactions* 31 (part 3):654-658). The CpG sequence may be

specific for inducing a Th1 immune response, such as a CpG-A ODN, or it may be more specific for inducing a B cell response, such a CpG-B ODN. CpG-A and CpG-B ODNs are discussed in Blackwell *et al.* (2003) *J Immunol* 170:4061-4068; Krieg (2002) *Trends Immunol* 23:64-65; and WO01/95935. Preferably, the CpG is a CpG-A ODN.

Preferably, the CpG oligonucleotide is constructed so that the 5' end is accessible for receptor recognition. Optionally, two CpG oligonucleotide sequences may be attached at their 3' ends to form "immunomers". See, for example, Gluck *et al.* (2002) *Vaccine* 20:B10-B16; Kandimalla *et al.* (2003) *BBRC* 306:948-953; Bhagat *et al.* (2003) *BBRC* 300:853-861; and WO03/035836.

A useful CpG adjuvant is CpG7909, also known as ProMune™ (Coley Pharmaceutical Group, Inc.). Another is CpG1826. As an alternative, or in addition, to using CpG sequences, TpG sequences can be used (WO01/22972), and these oligonucleotides may be free from unmethylated CpG motifs. The immunostimulatory oligonucleotide may be pyrimidine-rich. For example, it may comprise more than one consecutive thymidine nucleotide (*e.g.* TTTT, as disclosed in Pajak *et al.* (2003) *Vaccine* 21:836-842), and/or it may have a nucleotide composition with >25% thymidine (*e.g.* >35%, >40%, >50%, >60%, >80%, *etc.*). For example, it may comprise more than one consecutive cytosine nucleotide (*e.g.* CCCC, as disclosed in Pajak *et al.* (2003) *Vaccine* 21:836-842), and/or it may have a nucleotide composition with >25% cytosine (*e.g.* >35%, >40%, >50%, >60%, >80%, *etc.*). These oligonucleotides may be free from unmethylated CpG motifs. Immunostimulatory oligonucleotides will typically comprise at least 20 nucleotides. They may comprise fewer than 100 nucleotides.

A particularly useful adjuvant based around immunostimulatory oligonucleotides is known as IC-31™ (Schellack *et al.* (2006) *Vaccine* 24:5461-72). Thus an adjuvant used with the invention may comprise a mixture of (i) an oligonucleotide (*e.g.* between 15-40 nucleotides) including at least one (and preferably multiple) CpI motifs (*i.e.* a cytosine linked to an inosine to form a dinucleotide), and (ii) a polycationic polymer, such as an oligopeptide (*e.g.* between 5-20 amino acids) including at least one (and preferably multiple) Lys-Arg-Lys tripeptide sequence(s). The oligonucleotide may be a deoxynucleotide comprising 26-mer sequence 5'-$(IC)_{13}$-3'. The polycationic polymer may be a peptide comprising 11-mer amino acid sequence KLKLLLLLKLK.

Bacterial ADP-ribosylating toxins and detoxified derivatives thereof may be used as adjuvants in the invention. Preferably, the protein is derived from *E.coli* (*E.coli* heat labile enterotoxin "LT"), cholera ("CT"), or pertussis ("PT"). The use of detoxified ADP-ribosylating toxins as mucosal adjuvants is described in WO95/17211 and as parenteral adjuvants in WO98/42375. The toxin or toxoid is preferably in the form of a holotoxin, comprising both A and B subunits. Preferably, the A subunit contains a detoxifying mutation; preferably the B subunit is not mutated. Preferably, the adjuvant is a detoxified LT mutant such as LT-K63, LT-R72, and LT-G192. The use of ADP-ribosylating toxins and detoxified derivatives thereof, particularly LT-K63 and LT-R72, as adjuvants can be found in Beignon *et al.* (2002) *Infect Immun* 70:3012-3019; Pizza *et al.* (2001) *Vaccine* 19:2534-2541; Pizza *et al.* (2000) *Int J Med Microbiol* 290:455-461; Scharton-Kersten *et al.* (2000) *Infect Immun*

68:5306-5313; Ryan *et al.* (1999) *Infect Immun* 67:6270-6280; Partidos *et al.* (1999) *Immunol Lett* 67:209-216; Peppoloni *et al.* (2003) *Expert Rev Vaccines* 2:285-293; and Pine *et al.* (2002) *J Control Release* 85:263-270.

A useful CT mutant is or CT-E29H (Tebbey *et al.* (2000) *Vaccine* 18:2723-34). Numerical reference for amino acid substitutions is preferably based on the alignments of the A and B subunits of ADP-ribosylating toxins set forth in Domenighini *et al.* (1995) *Mol Microbiol* 15:1165-1167, specifically incorporated herein by reference in its entirety.

F. Human immunomodulators

Human immunomodulators suitable for use as adjuvants in the invention include cytokines, such as interleukins (*e.g.* IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12 (WO99/40936), *etc.*) (WO99/44636), interferons (*e.g.* interferon-γ), macrophage colony stimulating factor, and tumor necrosis factor. A preferred immunomodulator is IL-12.

G. Bioadhesives and Mucoadhesives

Bioadhesives and mucoadhesives may also be used as adjuvants in the invention. Suitable bioadhesives include esterified hyaluronic acid microspheres (Singh *et al.* (2001) *J Cont Release* 70:267-276) or mucoadhesives such as cross-linked derivatives of poly(acrylic acid), polyvinyl alcohol, polyvinyl pyrollidone, polysaccharides and carboxymethylcellulose. Chitosan and derivatives thereof may also be used as adjuvants in the invention (WO99/27960).

H. Microparticles

Microparticles may also be used as adjuvants in the invention. Microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, *etc.*), with poly(lactide-co-glycolide) are preferred, optionally treated to have a negatively-charged surface (*e.g.* with SDS) or a positively-charged surface (*e.g.* with a cationic detergent, such as CTAB).

I. Liposomes (Chapters 13 & 14 of Vaccine Design... (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.)

Examples of liposome formulations suitable for use as adjuvants are described in US 6,090,406; US 5,916,588; and EP-A-0626169.

J. Polyoxyethylene ether and polyoxyethylene ester formulations

Adjuvants suitable for use in the invention include polyoxyethylene ethers and polyoxyethylene esters (WO99/52549). Such formulations further include polyoxyethylene sorbitan ester surfactants in combination with an octoxynol (WO01/21207) as well as polyoxyethylene alkyl ethers or ester surfactants in combination with at least one additional non-ionic surfactant such as an octoxynol (WO01/21152). Preferred polyoxyethylene ethers are selected from the following group:

polyoxyethylene-9-lauryl ether (laureth 9), polyoxyethylene-9-steoryl ether, polyoxytheylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

K. Phosphazenes

A phosphazene, such as poly[di(carboxylatophenoxy)phosphazene] ("PCPP") as described, for example, in Andrianov *et al.* (1998) *Biomaterials* 19:109-115 and Payne *et al.* (1998) *Adv Drug Delivery Review* 31:185-196, may be used.

L. Muramyl peptides

Examples of muramyl peptides suitable for use as adjuvants in the invention include N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), and N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE).

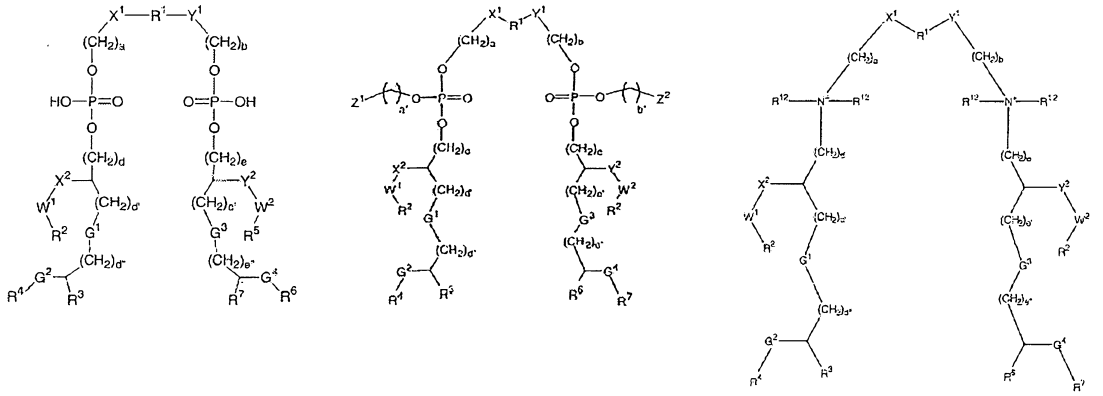
M. Imidazoquinolone Compounds.

Examples of imidazoquinolone compounds suitable for use adjuvants in the invention include Imiquimod ("R-837") (US 4,680,338; US 4,988,815), Resiquimod ("R-848") (WO92/15582), and their analogs; and salts thereof (*e.g.* the hydrochloride salts). Further details about immunostimulatory imidazoquinolines can be found in Stanley (2002) *Clin Exp Dermatol* 27:571-577; Wu *et al.* (2004) *Antiviral Res.* 64(2):79-83; Vasilakos *et al.* (2000) *Cell Immunol.* 204(1):64-74; US patents 4689338, 4929624, 5238944, 5266575, 5268376, 5346905, 5352784, 5389640, 5395937, 5482936, 5494916, 5525612, 6083505, 6440992, 6627640, 6656938, 6660735, 6660747, 6664260, 6664264, 6664265, 6667312, 6670372, 6677347, 6677348, 6677349, 6683088, 6703402, 6743920, 6800624, 6809203, 6888000 and 6924293; and Jones (2003) *Curr Opin Investig Drugs* 4:214-218.

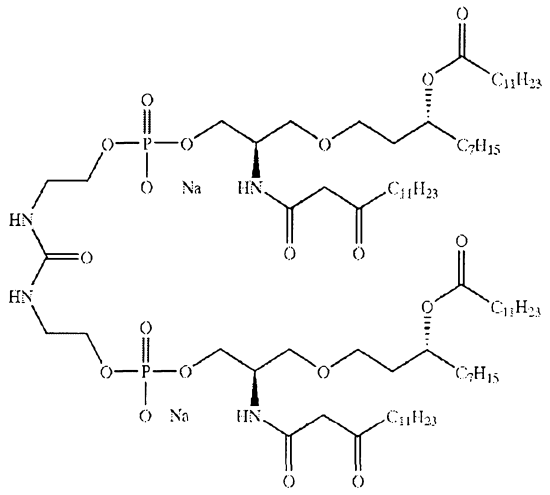
N. Substituted ureas

Substituted ureas useful as adjuvants include compounds of formula I, II or III, or salts thereof:

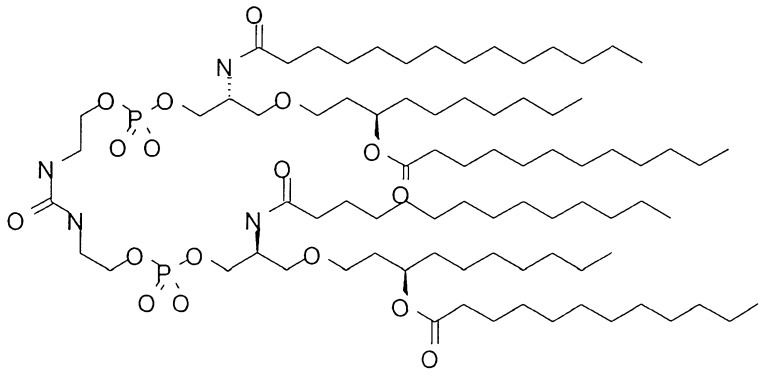
I II III



as defined in WO03/011223, such as 'ER 803058', 'ER 803732', 'ER 804053', ER 804058', 'ER 804059', 'ER 804442', 'ER 804680', 'ER 804764', ER 803022 or 'ER 804057' *e.g.*:



ER804057



ER-803022:

O. Further adjuvants

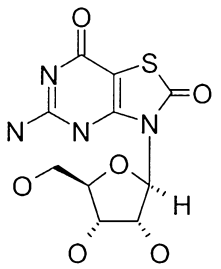
Further adjuvants that may be used with the invention include:

- An aminoalkyl glucosaminide phosphate derivative, such as RC-529 (Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278; Evans *et al.* (2003) *Expert Rev Vaccines* 2:219-229).

- A thiosemicarbazone compound, such as those disclosed in WO2004/060308. Methods of formulating, manufacturing, and screening for active compounds are also described in Bhagat *et al.* (2003) *BBRC* 300:853-861. The thiosemicarbazones are particularly effective in the stimulation of human peripheral blood mononuclear cells for the production of cytokines, such as TNF-α.

- A tryptanthrin compound, such as those disclosed in WO2004/064759. Methods of formulating, manufacturing, and screening for active compounds are also described in WO03/035836. The thiosemicarbazones are particularly effective in the stimulation of human peripheral blood mononuclear cells for the production of cytokines, such as TNF-α.

- A nucleoside analog, such as: (a) Isatorabine (ANA-245; 7-thia-8-oxoguanosine):



and prodrugs thereof; (b) ANA975; (c) ANA-025-1; (d) ANA380; (e) the compounds disclosed in US 6,924,271, US2005/0070556 and US 5,658,731, oxoribine (7-allyl-8-oxoguanosine) (US patent 5,011,828).

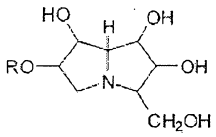
- Compounds disclosed in WO2004/87153, including: Acylpiperazine compounds, Indoledione compounds, Tetrahydraisoquinoline (THIQ) compounds, Benzocyclodione compounds, Aminoazavinyl compounds, Aminobenzimidazole quinolinone (ABIQ) compounds (US 6,605,617, WO02/18383), Hydrapthalamide compounds, Benzophenone compounds, Isoxazole compounds, Sterol compounds, Quinazilinone compounds, Pyrrole compounds (WO2004/018455), Anthraquinone compounds, Quinoxaline compounds, Triazine compounds, Pyrazalopyrimidine compounds, and Benzazole compounds (WO03/082272).

- Compounds containing lipids linked to a phosphate-containing acyclic backbone, such as the TLR4 antagonist E5564 (Wong *et al.* (2003) *J Clin Pharmacol* 43(7):735-42 ; US2005/0215517).

- A polyoxidonium polymer (Dyakonova et al. (2004) Int Immunopharmacol 4(13):1615-23 ; FR-2859633) or other N-oxidized polyethylene-piperazine derivative.

- Methyl inosine 5'-monophosphate ("MIMP") (Signorelli & Hadden (2003) *Int Immunopharmacol* 3(8):1177-86).

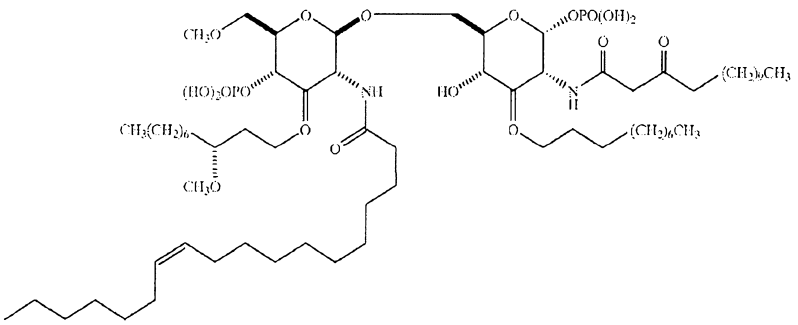
- A polyhydroxlated pyrrolizidine compound (WO2004/064715), such as one having formula:



where R is selected from the group comprising hydrogen, straight or branched, unsubstituted or substituted, saturated or unsaturated acyl, alkyl (*e.g.* cycloalkyl), alkenyl, alkynyl and aryl groups, or a pharmaceutically acceptable salt or derivative thereof. Examples include, but are not limited to: casuarine, casuarine-6-α-D-glucopyranose, 3-*epi*-casuarine, 7-*epi*-casuarine, 3,7-di*epi*-casuarine, *etc.*

- A CD1d ligand, such as an α-glycosylceramide (De Libero *et al*, *Nature Reviews Immunology,* 2005, 5: 485-496; US patent 5,936,076 ; Oki *et al*, *J. Clin. Investig.*, 113: 1631-1640 ; US2005/0192248; Yang *et al*, *Angew. Chem. Int. Ed.*, 2004, 43: 3818-3822; WO2005/102049; Goff *et al, J. Am. Chem., Soc.,* 2004, 126: 13602-13603; WO03/105769) *e.g.* α-galactosylceramide), phytosphingosine-containing α-glycosylceramides, OCH, KRN7000 [(2S,3S,4R)-1-O-(α-D-galactopyranosyl)-2-(N-hexacosanoylamino)-1,3,4-octadecanetriol], CRONY-101, 3"-O-sulfo-galactosylceramide, *etc.*

- A gamma inulin (Cooper (1995) *Pharm Biotechnol* 6:559-80) or derivative thereof, such as algammulin.



Adjuvant combinations

The invention may also comprise combinations of aspects of one or more of the adjuvants identified above. For example, the following adjuvant compositions may be used in the invention: (1) a saponin and an oil-in-water emulsion (WO99/11241); (2) a saponin (*e.g.* QS21) + a non-toxic LPS derivative (*e.g.* 3dMPL) (WO94/00153); (3) a saponin (*e.g.* QS21) + a non-toxic LPS derivative (*e.g.* 3dMPL) + a cholesterol; (4) a saponin (*e.g.* QS21) + 3dMPL + IL-12 (optionally + a sterol) (WO98/57659); (5) combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions (European patent applications 0835318, 0735898 and 0761231); (6) SAF, containing 10% squalane, 0.4% Tween 80™, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion. (7) Ribi™ adjuvant system (RAS), (Ribi Immunochem) containing 2% squalene, 0.2% Tween 80, and one or more bacterial cell wall

components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); and (8) one or more mineral salts (such as an aluminum salt) + a non-toxic derivative of LPS (such as 3dMPL). In some embodiments a combination of a toxin (e.g. LTK63) and an immunostimulatory oligonucleotide (e.g. CpG) is used. In some embodiments, a combination of an emulsion (e.g. montanide) and an immunostimulatory oligonucleotide (e.g. CpG) is used.

Other substances that act as immunostimulating agents are disclosed in chapter 7 of *Vaccine Design,* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.

The use of an aluminium hydroxide and/or aluminium phosphate adjuvant is particularly preferred, and antigens are generally adsorbed to these salts. Calcium phosphate is another preferred adjuvant. Other preferred adjuvant combinations include combinations of Th1 and Th2 adjuvants such as CpG & alum or resiquimod & alum. A combination of aluminium phosphate and 3dMPL may be used.

To improve thermal stability, a composition may include a temperature protective agent. This component may be particularly useful in adjuvanted compositions (particularly those containing a mineral adjuvant, such as an aluminium salt). As described in WO2006/110603, a liquid temperature protective agent may be added to an aqueous vaccine composition to lower its freezing point *e.g.* to reduce the freezing point to below 0°C. Thus the composition can be stored below 0°C, but above its freezing point, to inhibit thermal breakdown. The temperature protective agent also permits freezing of the composition while protecting mineral salt adjuvants against agglomeration or sedimentation after freezing and thawing, and may also protect the composition at elevated temperatures *e.g.* above 40°C. A starting aqueous vaccine and the liquid temperature protective agent may be mixed such that the liquid temperature protective agent forms from 1-80% by volume of the final mixture. Suitable temperature protective agents should be safe for human administration, readily miscible/soluble in water, and should not damage other components (*e.g.* antigen and adjuvant) in the composition. Examples include glycerin, propylene glycol, and/or polyethylene glycol (PEG). Suitable PEGs may have an average molecular weight ranging from 200-20,000 Da. In a preferred embodiment, the polyethylene glycol can have an average molecular weight of about 300 Da ('PEG-300').

The invention provides an immunogenic composition comprising: (i) one or more proteins of the invention; and (ii) a temperature protective agent. This composition may be formed by mixing (i) an aqueous composition comprising one or more proteins of the invention, with (ii) a temperature protective agent. The mixture may then be stored *e.g.* below 0°C, from 0-20°C, from 20-35°C, from 35-55°C, or higher. It may be stored in liquid or frozen form. The mixture may be lyophilised. The composition may alternatively be formed by mixing (i) a dried composition comprising one or more proteins of the invention, with (ii) a liquid composition comprising the temperature protective agent. Thus component (ii) can be used to reconstitute component (i).

The compositions of the invention may elicit either or both of a cell mediated immune response and a humoral immune response. This immune response will preferably induce long lasting (*e.g.* neutralising) antibodies and a cell mediated immunity that can quickly respond upon exposure to chlamydia.

Two types of T cells, CD4 and CD8 cells, are generally thought necessary to initiate and/or enhance cell mediated immunity and humoral immunity. CD8 T cells can express a CD8 co-receptor and are commonly referred to as Cytotoxic T lymphocytes (CTLs). CD8 T cells are able to recognized or interact with antigens displayed on MHC Class I molecules.

CD4 T cells can express a CD4 co-receptor and are commonly referred to as T helper cells. CD4 T cells are able to recognize antigenic peptides bound to MHC class II molecules. Upon interaction with a MHC class II molecule, the CD4 cells can secrete factors such as cytokines. These secreted cytokines can activate B cells, cytotoxic T cells, macrophages, and other cells that participate in an immune response. Helper T cells or CD4+ cells can be further divided into two functionally distinct subsets: TH1 phenotype and TH2 phenotypes which differ in their cytokine and effector function.

Activated TH1 cells enhance cellular immunity (including an increase in antigen-specific CTL production) and are therefore of particular value in responding to intracellular infections. Activated TH1 cells may secrete one or more of IL-2, IFNγ, and TNF-β. A TH1 immune response may result in local inflammatory reactions by activating macrophages, NK (natural killer) cells, and CD8 cytotoxic T cells (CTLs). A TH1 immune response may also act to expand the immune response by stimulating growth of B and T cells with IL-12. TH1 stimulated B cells may secrete IgG2a.

Activated TH2 cells enhance antibody production and are therefore of value in responding to extracellular infections. Activated TH2 cells may secrete one or more of IL-4, IL-5, IL-6, and IL-10. A TH2 immune response may result in the production of IgG1, IgE, IgA and memory B cells for future protection.

An enhanced immune response may include one or more of an enhanced TH1 immune response and a TH2 immune response.

A TH1 immune response may include one or more of an increase in CTLs, an increase in one or more of the cytokines associated with a TH1 immune response (such as IL-2, IFNγ, and TNF-β), an increase in activated macrophages, an increase in NK activity, or an increase in the production of IgG2a. Preferably, the enhanced TH1 immune response will include an increase in IgG2a production.

A TH1 immune response may be elicited using a TH1 adjuvant. A TH1 adjuvant will generally elicit increased levels of IgG2a production relative to immunization of the antigen without adjuvant. TH1 adjuvants suitable for use in the invention may include for example saponin formulations, virosomes and virus like particles, non-toxic derivatives of enterobacterial lipopolysaccharide (LPS),

immunostimulatory oligonucleotides. Immunostimulatory oligonucleotides, such as oligonucleotides containing a CpG motif, are preferred TH1 adjuvants for use in the invention.

A TH2 immune response may include one or more of an increase in one or more of the cytokines associated with a TH2 immune response (such as IL-4, IL-5, IL-6 and IL-10), or an increase in the production of IgG1, IgE, IgA and memory B cells. Preferably, the enhanced TH2 immune resonse will include an increase in IgG1 production.

A TH2 immune response may be elicited using a TH2 adjuvant. A TH2 adjuvant will generally elicit increased levels of IgG1 production relative to immunization of the antigen without adjuvant. TH2 adjuvants suitable for use in the invention include, for example, mineral containing compositions, oil-emulsions, and ADP-ribosylating toxins and detoxified derivatives thereof. Mineral containing compositions, such as aluminium salts are preferred TH2 adjuvants for use in the invention.

Preferably, the invention includes a composition comprising a combination of a TH1 adjuvant and a TH2 adjuvant. Preferably, such a composition elicits an enhanced TH1 and an enhanced TH2 response, i.e., an increase in the production of both IgG1 and IgG2a production relative to immunization without an adjuvant. Still more preferably, the composition comprising a combination of a TH1 and a TH2 adjuvant elicits an increased TH1 and/or an increased TH2 immune response relative to immunization with a single adjuvant (*i.e.*, relative to immunization with a TH1 adjuvant alone or immunization with a TH2 adjuvant alone).

The immune response may be one or both of a TH1 immune response and a TH2 response. Preferably, immune response provides for one or both of an enhanced TH1 response and an enhanced TH2 response. Preferably, the immune response includes an increase in the production of IgG1 and/or IgG2 and/or IgGA.

The invention is preferably used to elicit systemic and/or mucosal immunity. The enhanced immune response may be one or both of a systemic and a mucosal immune response. Preferably, the immune response provides for one or both of an enhanced systemic and an enhanced mucosal immune response. Preferably the mucosal immune response is a TH2 immune response. Preferably, the mucosal immune response includes an increase in the production of IgA.

Methods of treatment, and administration of the vaccine

The invention also provides a method for raising an immune response in a mammal comprising the step of administering an effective amount of a protein, antibody, nucleic acid, vector, host cell or composition of the invention. The immune response is preferably protective and preferably involves antibodies and/or cell-mediated immunity. The method may raise a booster response.

The invention also provides a protein or combination, as defiend above, for use as a medicament *e.g.* for use in raising an immune response in a mammal.

The invention also provides the use of a protein or combination of the invention in the manufacture of a medicament for raising an immune response in a mammal. By raising an immune response in the mammal by these uses and methods, the mammal can be protected against Chlamydia infection. More particularly, the mammal may be protected against *Chlamydia trachomatis*. The invention is effective against Chlamydia of various different serotypes, but can be particularly useful in protecting against disease resulting from Chlamydia infection by strains in serovar D.

Thus, according to a further aspect, the invention also provides a nucleic acid, protein, antibody, vector or host cell according to the invention for use as a medicament (*e.g.* a vaccine) or a diagnostic reagent. In one embodiment, the protein, nucleic acid or antibody is used for treatment, prevention or diagnosis of Chlamydia infection (preferably *C. trachomatis*) in a mammal. The invention also provides a method of treating, preventing of diagnosing Chlamydia infection (preferably, *C. trachomatis* infection) in a patient (preferably a mammal), comprising administering a therapeutically effective amount of a nucleic acid, protein or antibody of the invention.

Preferably, the nucleic acid, protein or antibody according to the invention is for treatment or prevention of Chlamydia infection or an associated condition (*e.g.* trachoma, blindness, cervicitis, pelvic inflammatory disease, infertility, ectopic pregnancy, chronic pelvic pain, salpingitis, urethritis, epididymitis, infant pneumonia, cervical squamous cell carcinoma, *etc.*), preferably, *C. trachomatis* infection. The immunogenic composition may additionally or alternatively be effective against *C. pneumoniae*.

The mammal is preferably a human. Where the vaccine is for prophylactic use, the human is preferably a child (*e.g.* a toddler or infant) or a teenager; where the vaccine is for therapeutic use, the human is preferably a teenager or an adult. A vaccine intended for children may also be administered to adults *e.g.* to assess safety, dosage, immunogenicity, *etc.* Thus a human patient may be less than 1 year old, 1-5 years old, 5-15 years old, 15-55 years old, or at least 55 years old. Preferred patients for receiving the vaccines are people going through purberty, teenagers, sexually active people, the elderly (*e.g.* ≥50 years old, ≥60 years old, and preferably ≥65 years), the young (*e.g.* ≤5 years old), hospitalised patients, healthcare workers, armed service and military personnel, pregnant women, the chronically ill, or immunodeficient patients. The vaccines are not suitable solely for these groups, however, and may be used more generally in a population.

Vaccines produced by the invention may be administered to patients at substantially the same time as (*e.g.* during the same medical consultation or visit to a healthcare professional or vaccination centre) other vaccines *e.g.* at substantially the same time as a human papillomavirus vaccine such as Cervarix™ or Gardasil™; a tetanus, diphtheria and acellular pertussis vaccine such as TDaP, DTaP or Boostrix™; a rubella vaccine such as MMR; or a tubercolosis vaccine such as the BCG. Examples of other vaccines that the vaccine produced by the invention may be administered at substantially the same time as are a measles vaccine, a mumps vaccine, a varicella vaccine, a MMRV vaccine, a diphtheria vaccine, a tetanus vaccine, a pertussis vaccine, a DTP vaccine, a conjugated *H.influenzae*

type b vaccine, an inactivated poliovirus vaccine, a hepatitis B virus vaccine, a meningococcal conjugate vaccine (such as a tetravalent A-C-W135-Y vaccine), a respiratory syncytial virus vaccine, *etc.*

In a preferred embodiment, the protein of the invention is used to elicit antibodies that are capable of neutralising the activity of the wild type Chlamydia protein, for example, of one or more of wild-type Chlamydia CT733, CT153, CT601, CT279, CT443, CT372, CT456, CT381, CT255, CT341, CT716, CT745, CT387, CT812, CT869, CT166, CT175, CT163, CT214, CT721, CT127, CT043, CT600 and/or CT823 for example, of one or more of wild-type Chlamydia CT733, CT153, CT601, CT279, CT443, CT372, CT456 and/or CT381. Neutralizing antibodies may be used as a vaccine capable of neutralising the activity of a native Chlamydia protein expressed by infectious EB. In one embodiment, the protein of the invention is used to elicit antibodies that are capable of neutralising Chlamydia infectivity and/or virulence. Thus, the invention also provides the antibodies of the invention for neutralising wild-type Chlamydia proteins and/or Chlamydia infectivity and/or virulence.

The invention also provides the use of a nucleic acid, protein, or antibody of the invention in the manufacture of: (i) a medicament for treating or preventing bacterial infection; (ii) a diagnostic reagent for detecting the presence of bacteria or of antibodies raised against bacteria; and/or (iii) a reagent which can raise antibodies against bacteria. Said bacteria is preferably a *Chlamydia*, e.g. *Chlamydia trachomatis* or *Chlamydia pneumoniae*, but is preferably *Chlamydia trachomatis*.

Also provided is a method for diagnosing Chlamydia infection, comprising:

(a) raising an antibody against a protein of the invention;

(b) contacting the antibody of step (a) with a biological sample suspected of being infected with Chlamydia under conditions suitable for the formation of antibody-antigen complexes; and

(c) detecting said complexes, wherein detection of said complex is indicative of Chlamydia infection.

Also provided is a method for diagnosing Chlamydia infection, comprising: (a) contacting an antibody which was raised against a protein of the invention with a biological sample suspected of being infected with Chlamydia under conditions suitable for the formation of antibody-antigen complexes; and (b) detecting said complexes, wherein detection of said complex is indicative of Chlamydia infection.

Proteins of the invention can be used in immunoassays to detect antibody levels (or, conversely, antibodies of the invention can be used to detect protein levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to proteins within biological samples, including for example, blood or serum samples, can be

detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

Testing efficacy of compositions

The efficacy of the immunogenic compositions of the present invention can be evaluated in *in vitro* and *in vivo* animal models prior to host, e.g., human, administration. For example, *in vitro* neutralization by Peterson et al (1988) is suitable for testing vaccine compositions directed toward *Chlamydia trachomatis*.

One way of checking efficacy of therapeutic treatment involves monitoring *C. trachomatis* infection after administration of the compositions of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses both systemically (such as monitoring the level of IgG1 and IgG2a production) and mucosally (such as monitoring the level of IgA production) against the *Chlamydia trachomatis* antigens in the compositions of the invention after administration of the composition. Typically, serum *Chlamydia* specific antibody responses are determined post-immunisation but pre-challenge whereas mucosal *Chlamydia* specific antibody body responses are determined post-immunisation and post-challenge.

One example of such an *in vitro* test is described as follows. Hyper-immune antisera is diluted in PBS containing 5% guinea pig serum, as a complement source. *Chlamydia trachomatis* (10^4 IFU; inclusion forming units) are added to the antisera dilutions. The antigen-antibody mixtures are incubated at 37°C for 45 minutes and inoculated into duplicate confluent Hep-2 or HeLa cell monolayers contained in glass vials (e.g., 15 by 45 mm), which have been washed twice with PBS prior to inoculation. The monolayer cells are infected by centrifugation at 1000X g for 1 hour followed by stationary incubation at 37°C for 1 hour. Infected monolayers are incubated for 48 or 72 hours, fixed and stained with Chlamydia specific antibody, such as anti-MOMP. Inclusion-bearing cells are counted in ten fields at a magnification of 200X. Neutralization titer is assigned on the dilution that gives 50% inhibition as compared to control monolayers/IFU.

Another way of assessing the immunogenicity of the compositions of the present invention is to express the proteins recombinantly for screening patient sera or mucosal secretions by immunoblot

and/or microarrays. A positive reaction between the protein and the patient sample indicates that the patient has mounted an immune response to the protein in question. This method may also be used to identify immunodominant antigens and/or epitopes within antigens.

The efficacy of vaccine compositions can also be determined *in vivo* by challenging animal models of *Chlamydia trachomatis* infection, *e.g.*, guinea pigs or mice, with the vaccine compositions. For example, *in vivo* vaccine composition challenge studies in the guinea pig model of *Chlamydia trachomatis* infection can be performed. A description of one example of this type of approach follows. Female guinea pigs weighing 450 – 500 g are housed in an environmentally controlled room with a 12 hour light-dark cycle and immunized with vaccine compositions via a variety of immunization routes. Post-vaccination, guinea pigs are infected in the genital tract with the agent of guinea pig inclusion conjunctivitis (GPIC), which has been grown in HeLa or McCoy cells (Rank et al. (1988)). Each animal receives approximately $1.4x10^7$ inclusion forming units (IFU) contained in 0.05 ml of sucrose-phosphate-glutamate buffer, pH 7.4 (Schacter, 1980). The course of infection monitored by determining the percentage of inclusion-bearing cells by indirect immunofluorescence with GPIC specific antisera, or by Giemsa-stained smear from a scraping from the genital tract (Rank et al 1988). Antibody titers in the serum is determined by an enzyme-linked immunosorbent assay.

Alternatively, *in vivo* vaccine compositions challenge studies can be performed in the murine model of *Chlamydia trachomatis* (Morrison et al 1995). A description of one example of this type of approach is as follows. Female mice 7 to 12 weeks of age receive 2.5 mg of depo-provera subcutaneously at 10 and 3 days before vaginal infection. Post-vaccination, mice are infected in the genital tract with 1,500 inclusion-forming units of *Chlamydia trachomatis* contained in 5ml of sucrose-phosphate-glutamate buffer, pH 7.4. The course of infection is monitored by determining the percentage of inclusion-bearing cells by indirect immunofluorescence with *Chlamydia trachomatis* specific antisera, or by a Giemsa-stained smear from a scraping from the genital tract of an infected mouse. The presence of antibody titers in the serum of a mouse is determined by an enzyme-linked immunosorbent assay.

Nucleic acid immunisation

The immunogenic compositions described above include Chlamydia antigens. In all cases, however, the polypeptide antigens can be replaced by nucleic acids (typically DNA) encoding those polypeptides, to give compositions, methods and uses based on nucleic acid immunisation. Nucleic acid immunisation is now a developed field (*e.g.* see Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648; Strugnell *et al.* (1997) *Immunol Cell Biol* 75(4):364-369; Cui (2005) *Adv Genet* 54:257-89; Robinson & Torres (1997) *Seminars in Immunol* 9:271-283; Brunham *et al.* (2000) *J Infect Dis* 181 Suppl 3:S538-43; Svanholm *et al.* (2000) *Scand J Immunol* 51(4):345-53; *DNA Vaccination - Genetic Vaccination* (1998) eds. Koprowski *et al.* (ISBN 3540633928); *Gene Vaccination : Theory and Practice* (1998) ed. Raz (ISBN 3540644288), *etc.*).

The nucleic acid encoding the immunogen is expressed *in vivo* after delivery to a patient and the expressed immunogen then stimulates the immune system. The active ingredient will typically take the form of a nucleic acid vector comprising: (i) a promoter; (ii) a sequence encoding the immunogen, operably linked to the promoter; and optionally (iii) a selectable marker. Preferred vectors may further comprise (iv) an origin of replication; and (v) a transcription terminator downstream of and operably linked to (ii). In general, (i) & (v) will be eukaryotic and (iii) & (iv) will be prokaryotic.

Preferred promoters are viral promoters *e.g.* from cytomegalovirus (CMV). The vector may also include transcriptional regulatory sequences (*e.g.* enhancers) in addition to the promoter and which interact functionally with the promoter. Preferred vectors include the immediate-early CMV enhancer/promoter, and more preferred vectors also include CMV intron A. The promoter is operably linked to a downstream sequence encoding an immunogen, such that expression of the immunogen-encoding sequence is under the promoter's control.

Where a marker is used, it preferably functions in a microbial host (*e.g.* in a prokaryote, in a bacteria, in a yeast). The marker is preferably a prokaryotic selectable marker (*e.g.* transcribed under the control of a prokaryotic promoter). For convenience, typical markers are antibiotic resistance genes.

The vector of the invention is preferably an autonomously replicating episomal or extrachromosomal vector, such as a plasmid.

The vector of the invention preferably comprises an origin of replication. It is preferred that the origin of replication is active in prokaryotes but not in eukaryotes.

Preferred vectors thus include a prokaryotic marker for selection of the vector, a prokaryotic origin of replication, but a *eukaryotic* promoter for driving transcription of the immunogen-encoding sequence. The vectors will therefore (a) be amplified and selected in prokaryotic hosts without polypeptide expression, but (b) be expressed in eukaryotic hosts without being amplified. This arrangement is ideal for nucleic acid immunization vectors.

The vector of the invention may comprise a eukaryotic transcriptional terminator sequence downstream of the coding sequence. This can enhance transcription levels. Where the coding sequence does not have its own, the vector of the invention preferably comprises a polyadenylation sequence. A preferred polyadenylation sequence is from bovine growth hormone.

The vector of the invention may comprise a multiple cloning site.

In addition to sequences encoding the immunogen and a marker, the vector may comprise a second eukaryotic coding sequence. The vector may also comprise an IRES upstream of said second sequence in order to permit translation of a second eukaryotic polypeptide from the same transcript as the immunogen. Alternatively, the immunogen-coding sequence may be downstream of an IRES.

The vector of the invention may comprise unmethylated CpG motifs *e.g.* unmethylated DNA sequences which have in common a cytosine preceding a guanosine, flanked by two 5' purines and two 3' pyrimidines. In their unmethylated form these DNA motifs have been demonstrated to be potent stimulators of several types of immune cell.

Vectors may be delivered in a targeted way. Receptor-mediated DNA delivery techniques are described in, for example, Findeis *et al.*, *Trends Biotechnol.* (1993) 11:202; Chiou *et al.* (1994) *Gene Therapeutics: Methods And Applications Of Direct Gene Transfer.* ed. Wolff; Wu *et al.*, *J. Biol. Chem.* (1988) *263*:621; Wu *et al.*, *J. Biol. Chem.* (1994) *269*:542; Zenke *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1990) *87*:3655; and Wu *et al.*, *J. Biol. Chem.* (1991) *266*:338.

Therapeutic compositions containing a nucleic acid are administered in a range of about 100ng to about 200mg of DNA for local administration in a gene therapy protocol. Concentration ranges of about 500 ng to about 50 mg, about 1μg to about 2 mg, about 5μg to about 500μg, and about 20μg to about 100μg of DNA can also be used during a gene therapy protocol. Factors such as method of action (*e.g.* for enhancing or inhibiting levels of the encoded gene product) and efficacy of transformation and expression are considerations which will affect the dosage required for ultimate efficacy. Where greater expression is desired over a larger area of tissue, larger amounts of vector or the same amounts re-administered in a successive protocol of administrations, or several administrations to different adjacent or close tissue portions may be required to effect a positive therapeutic outcome. In all cases, routine experimentation in clinical trials will determine specific ranges for optimal therapeutic effect.

Vectors can be delivered using gene delivery vehicles. The gene delivery vehicle can be of viral or non-viral origin (see generally Jolly, *Cancer Gene Therapy* (1994) *1*:51; Kimura, *Human Gene Therapy* (1994) *5*:845; Connelly, *Human Gene Therapy* (1995) *1*:185; and Kaplitt, *Nature Genetics* (1994) *6*:148).

Viral-based vectors for delivery of a desired nucleic acid and expression in a desired cell are well known in the art. Exemplary viral-based vehicles include, but are not limited to, recombinant retroviruses (*e.g.* WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; US patent 5,219,740; WO 93/11230; WO 93/10218; US patent 4,777,127; GB Patent No. 2,200,651; EP-A-0345242; and WO 91/02805), alphavirus-based vectors (*e.g.* Sindbis virus vectors, Semliki forest virus (ATCC VR-67; ATCC VR-1247), Ross River virus (ATCC VR-373; ATCC VR-1246) and Venezuelan equine encephalitis virus (ATCC VR-923; ATCC VR-1250; ATCC VR 1249; ATCC VR-532); hybrids or chimeras of these viruses may also be used), poxvirus vectors (*e.g.* vaccinia, fowlpox, canarypox, modified vaccinia Ankara, *etc.*), adenovirus vectors, and adeno-associated virus (AAV) vectors (*e.g.* see WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; US patent 5,219,740; WO 93/11230; WO 93/10218; US patent 4,777,127; GB Patent No. 2,200,651; EP-A-0345242; WO 91/02805; WO 94/12649; WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984; and WO

95/00655). Administration of DNA linked to killed adenovirus (Curiel, *Hum. Gene Ther.* (1992) *3*:147) can also be employed.

Non-viral delivery vehicles and methods can also be employed, including, but not limited to, polycationic condensed DNA linked or unlinked to killed adenovirus alone (e.g. De Libero *et al*, *Nature Reviews Immunology,* 2005, 5: 485-496), ligand-linked DNA (Wu, *J. Biol. Chem.* (1989) *264*:16985), eukaryotic cell delivery vehicles cells (US patent 5,814,482; WO 95/07994; WO 96/17072; WO 95/30763; and WO 97/42338) and nucleic charge neutralization or fusion with cell membranes. Naked DNA can also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and US patent 5,580,859. Liposomes (*e.g.* immunoliposomes) that can act as gene delivery vehicles are described in US patent 5,422,120; WO 95/13796; WO 94/23697; WO 91/14445; and EP-0524968. Additional approaches are described in Philip, *Mol. Cell Biol.* (1994) *14*:2411 and Woffendin, *Proc. Natl. Acad. Sci.* (1994) *91*:11581.

Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials or use of ionizing radiation (e.g. US patent 5,206,152 and WO 92/11033). Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun (US patent 5,149,655) or use of ionizing radiation for activating transferred genes (Strugnell *et al.* (1997) *Immunol Cell Biol* 75(4):364-369 and Cui (2005) *Adv Genet* 54:257-89).

Delivery DNA using PLG {poly(lactide-co-glycolide)} microparticles is a particularly preferred method *e.g.* by adsorption to the microparticles, which are optionally treated to have a negatively-charged surface (*e.g.* treated with SDS) or a positively-charged surface (*e.g.* treated with a cationic detergent, such as CTAB).

Antibody immunisation

The antibodies of the invention may be used, for example, for neutralising the activity of the wild-type Chlamydia protein. Antibodies against Chlamydia antigens can be used for passive immunisation (Brandt *et al.* (2006) *J Antimicrob Chemother.* 58(6):1291-4. Epub 2006 Oct 26). Thus the invention provides the use of antibodies of the invention in therapy. The invention also provides the use of such antibodies in the manufacture of a medicament. The invention also provides a method for treating a mammal comprising the step of administering an effective amount of an antibody of the invention. As described above for immunogenic compositions, these methods and uses allow a mammal to be protected against Chlamydia infection.

Processes

According to further aspects, the invention provides various processes.

A process for producing a protein of the invention is provided, comprising the step of culturing a host cell of the invention under conditions which induce protein expression.

A process for producing protein or nucleic acid of the invention is provided, wherein the protein or nucleic acid is synthesised in part or in whole using chemical means.

A process for detecting Chlamydia (preferably *C. trachomatis*) in a biological sample is also provided, comprising the step of contacting a nucleic acid according to the invention with the biological sample under hybridising conditions. The process may involve nucleic acid amplification (*e.g.* PCR, SDA, SSSR, LCR, TMA *etc.*) or hybridisation (*e.g.* microarrays, blots, hybridisation with probe in solution *etc.*).

A process for detecting wild-type Chlamydia (preferably, *C. trachomatis*) is provided, comprising the steps of: (a) contacting an antibody of the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complex(es); and (b) detecting said complex(es). This process may advantageously be used to diagnose Chlamydia infection.

General

The practice of the present invention will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, immunology and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th edition, ISBN: 0683306472; *Methods In Enzymology* (S. Colowick and N. Kaplan, eds., Academic Press, Inc.); *Handbook of Experimental Immunology*, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds, 1986, Blackwell Scientific Publications); Sambrook *et al.* (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edition (Cold Spring Harbor Laboratory Press); *Handbook of Surface and Colloidal Chemistry* (Birdi, K.S. ed., CRC Press, 1997); Ausubel *et al.* (eds) (2002) *Short protocols in molecular biology*, 5th edition (Current Protocols); *Molecular Biology Techniques: An Intensive Laboratory Course*, (Ream *et al.*, eds., 1998, Academic Press); and *PCR (Introduction to Biotechniques Series)*, 2nd ed. (Newton & Graham eds., 1997, Springer Verlag) *etc.*

"GI" numbering is used herein. A GI number, or "GenInfo Identifier", is a series of digits assigned consecutively to each sequence record processed by NCBI when sequences are added to its databases. The GI number bears no resemblance to the accession number of the sequence record. When a sequence is updated (*e.g.* for correction, or to add more annotation or information) then it receives a new GI number. Thus the sequence associated with a given GI number is never changed.

Where the invention concerns an "epitope", this epitope may be a B-cell epitope and/or a T-cell epitope. Such epitopes can be identified empirically (*e.g.* using PEPSCAN (Geysen *et al.* (1984) *PNAS USA* 81:3998-4002; Carter (1994) *Methods Mol Biol* 36:207-23) or similar methods), or they can be predicted (*e.g.* using the Jameson-Wolf antigenic index (Jameson, BA *et al.* 1988, *CABIOS* 4(1):181-186), matrix-based approaches (Raddrizzani & Hammer (2000) *Brief Bioinform* 1(2):179-

89), MAPITOPE (Bublil *et al.* (2007) *Proteins* 68(1):294-304), TEPITOPE (De Lalla *et al.* (1999) *J. Immunol.* 163:1725-29; Kwok *et al.* (2001) *Trends Immunol* 22:583-88), neural networks (Brusic *et al.* (1998) *Bioinformatics* 14(2):121-30), OptiMer & EpiMer (Meister *et al.* (1995) *Vaccine* 13(6):581-91; Roberts *et al.* (1996) *AIDS Res Hum Retroviruses* 12(7):593-610), ADEPT (Maksyutov & Zagrebelnaya (1993) *Comput Appl Biosci* 9(3):291-7), Tsites (Feller & de la Cruz (1991) *Nature* 349(6311):720-1), hydrophilicity (Hopp (1993) *Peptide Research* 6:183-190), antigenic index (Welling *et al.* (1985) *FEBS Lett.* 188:215-218) or the methods disclosed in Davenport *et al.* (1995) *Immunogenetics* 42:392-297; Tsurui & Takahashi (2007) *J Pharmacol Sci.* 105(4):299-316; Tong *et al.* (2007) *Brief Bioinform.* 8(2):96-108 ; Schirle *et al.* (2001) *J Immunol Methods.* 257(1-2):1-16; and Chen *et al.* (2007) *Amino Acids* 33(3):423-8, *etc.*). Epitopes are the parts of an antigen that are recognised by and bind to the antigen binding sites of antibodies or T-cell receptors, and they may also be referred to as "antigenic determinants".

Where an antigen "domain" is omitted, this may involve omission of a signal peptide, of a cytoplasmic domain, of a transmembrane domain, of an extracellular domain, *etc.*

The term "comprising" encompasses "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

The term "about" in relation to a numerical value x is optional and means, for example, $x \pm 10\%$.

References to a percentage sequence identity between two amino acid sequences means that, when aligned, that percentage of amino acids are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in section 7.7.18 of *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.*, eds., 1987) Supplement 30. A preferred alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is disclosed in Smith & Waterman (1981) *Adv. Appl. Math.* 2: 482-489.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 is a graph which shows the ability of 20 selected *C. trachomatis* antigens to induce IFNγ production by CD4+ T cells.

Figure 2a shows the bacterial shedding (IFUs recovered from lungs) after Chlamydia challenge in mice to whom EB-CM CD4+ T cells had been adoptively transferred. Figure 2b shows the ability of various *C. muridarum* antigens to stimulate the protective EB-CD4+ T cell line to produce IFNγ.

Figure 3 is a histogram which shows the number of CD4+ T cells that produce IFNγ, upon specific stimulation with *C. trachomatis* recombinant antigens CT153 and CT733.

Figure 4 shows the protective activity of TC0106 (*C. muridarum* homologue of CT733) and TC0431 (*C. muridarum* homologue of CT153) as single antigens. The graph shows mean IFU/ml in BALB/C

mice immunised with the two antigens and then challendged with *C.muridarum*. The three bars are, from left to right: adjuvant alone; TC0106 as immunogen; and TC0431 as immunogen.

Figure 5 shows the protective activity of the combination TC0106+TC0431. The graph shows mean IFU per lung (Log10) recovered from infected lungs of mice immunised with the combination. The three bars are, from left to right: 10^3 live Ebs; adjuvant alone; antigen combination.

Figure 6 shows CD4 T cells producing IFNγ in PBMC of mice immunized with TC0106+TC0431, TC0106, TC0431 and LTK 63+CpG. From left to right, the bars represent stimulation with 1) LTK 63, TC0106+TC0431, TC0106, TC0431 (all EB-immunized mice); 2) LTK 63, TC0106+TC0431, TC0106 (all TC 0106-immunized mice); 3) LTK63, TC0106+TC0431. TC0431 (all TC0431-immunized mice); and 4) LTK63 and TC0106+TC0431 (both TC0106+TC0431-immunized mice). It shows that immunization with TC0106 (*C. muridarum* homologue of CT733) and TC0431 (*C. muridarum* homologue of CT153) elicits a significant frequency of specific CD4+/IFNγ+ cells. The Y axis shows frequency on 10^6 CD4.

Figure 7 is a summary of protection results for various combinations and single antigens in the mouse model of *C. muridarum* intransal challenge. It shows the mean IFU/lung of mice immunised imtramuscularly with single antigens, or antigen combinations, adjuvanted with LTK63 and CpG, then challenged intranasally with 10^3 *C.muridarum* IFU.

Figure 8 is a summary of protection results for various combinations of antigens in the mouse model of *C. muridarum* intransal challenge. It shows mean IFU/lung (log 10) of *C.muridarum* recovered from infected lungs of immunised mice.

Figure 9 shows the results of the combination TC0551+TC0651+TC0727+TC0890 in the mouse model of ovarian bursa challenge with *C. muridarum.* The Y axis shows IFU/swab (log10). The three groups, from left to right, are for different immunizing antigens: ovalbiumin; the combination; and nMOMP.

Figure 10A shows the protection results achieved with various antigens combinations in the mouse model of *C. muridarum* intranasal challenge. Figure 10B shows the frequency of IFNg-producing CD4+ T cells induced by vaccination with the antigen combination TC0890+TC0551. From left to right, the bars represent stimutaion with 1) TC0551, TC0890, TC0551+TC0890 (for adjuvant-immunized mice) and 2) TC0551, TC0890, TC0551+TC0890 (for MIX TC0890+TC0551-immunized mice). Figure 10C shows CD4 T cells producing TFNg and IL2/TNF in PBMC of mice immunized with TC0106+TC0431 with Ltk63 + CpG. From left to right, the bars represent stimulation with a) TC0106. TC0431, TC0106+TC0431, CT153+CT733 (all adjuvant-immunized mice); 2) TC0106, TC0431, TC0106+TC0431, CT153+CT733 (all MIX TC0106+TC0431-immunized mice).

Figure 11 shows an immunoblot analysis of CT601, CT279, CT153 and CT733 in Ct-EBs and *C.trachomatis*-infected HeLa cells using their specific mouse immune antisera.

Figure 12 shows protective activity of antigens TC0313, TC0741, TC0106 and TC0210 given singly or in combination. In 12A to 12D the bars show mean IFU/lung (Log10), with the left-hand bar being adjuvant alone (LTK61+CpG) and the right-hand bar being the TC antigen. Figure 12E shows the IFU reduction over time (Mean IFU/lung against days post-challenge) using the combination (squares) or adjuvant alone (diamonds).

Figure 13A and 13B are histograms showing the antigen specific CD4 Th1 response in BALB/c mice after a primary *C. trachomatis* (CT) infection. Results are the mean of 4 independent experiments. Two results are shown for each experiment: non-infected mice (left hand bar) and primary infected mice (right hand bar). From left to right in Figure 13A, the results relate to stimulation with CT812C, CT387, CT869, CT166 and CT175. From left to right in Figure 13B, the results relate to stimulation with MOMP, CT163, CT812, CT812C, CT166, CT869, CT163, CT812, CT214, CT387, CT721, CT127 and CT175. The frequency on 10^5 CD4 T cells is shown on the Y axis.

Figure 14 is a histogram showing *C. muridarum* IFUs recovered from infected lungs of immunised mice (Day 12 post I.N. challenge with 10^3 IFUs). The immunisation group is shown on the X axis: the left hand bar relates to mice immuised with LTK63+CpG; the right hand bar relates to mice immunised with TC0197+TC0261+TC0666+LTK63+CpG. Mean IFU/lung (Log10) is shown on the Y axis.

Figure 15 is a histogram showing *C. muridarum* IFUs recovered from infected lungs of immunised mice (Day 12 post I.N. challenge with 10^3 IFUs). The immunisation group is shown on the X axis: from left to right, the results relate to mice immunised with i) LTK63+CpG, ii) TC0261+LTK63+CpG, iii) TC0197+LTK63+CpG, and iv) TC0666+LTK63+CpG. Mean IFU/lung (Log10) is shown on the Y axis.

MODES FOR CARRYING OUT THE INVENTION

Example 1: Induction of population of CD4+ T cells to produce IFNγ

20 antigens have been found which induce a population of CD4+ T cells to produce IFNγ (see Figure 1). 17 of these are newly discovered (CT016, CT043, CT114, CT153, CT255, CT279, CT341, CT372, CT480, CT600, CT601, CT711, CT716, CT733, CT734. CT745, CT823), while three antigens (CT681-MOMP, CT396-Hsp60 and CT587-Enolase) have already been described as targets of CD4+ T cells (Goodall JC et al. 2001; Hassell AB et al. 1993). Significantly, some antigens were able to induce a frequency of antigen-specific CD4+ responding T cells at least comparable to what observed with the positive control antigen MOMP.

The 17 new antigens are as follows:

Antigen	Annotation	Gene name
CT016	Hypothetical protein	
CT043	Hypothetical protein	
CT114	Hypothetical protein	
CT153	Hypothetical protein	
CT255	Hypothetical protein	
CT279	Na(+)-translocating NADH-quinone reductase subunit C	nqr3
CT341	Heat shock protein J (Hsp-J)	dnaJ
CT372	Hypothetical protein	
CT480	Oligopeptide Binding Lipoprotein	oppA_4
CT600		
CT601	Invasin repeat family phosphatase	papQ
CT711	Hypothetical protein	
CT716	Hypothetical protein	
CT733	Hypothetical protein	
CT734	Hypothetical protein	
CT745	protoporphyrinogen oxidase	hemG
CT823	DO serine protease	htrA

Of these 17 new antigens, CT341 may be the least suitable for use in immunization because it is a heat shock protein.

Example 2: Characterization of the antigen-specifity of protective Chlamydia specific CD4+ Th1 cell lines

The relevance of the newly discovered antigens for protective immunity to Chlamydia was further corroborated by showing that they were recognized by T cells belonging to a Chlamydia-specific CD4+/IFNγ+ cell line, conferring protection when adoptively transferred to naïve recipient mice. To this aim we have derived a short-term CD4+ T cell line, produced against the extracellular EB form of *C. muridarum* that showed a high capacity to protect adoptively transferred naïve mice from *C. muridarum* challenge. The protective CD4+ cell line, which had undergone only a few cycles of expansion, maintained a polyclonal cell population with broad specificity that should correlate more closely to the *in vivo* protective response than long-term lines or clones. The polyclonal cell line was analysed for its antigen recognition profile versus the *C. muridarum* antigens, homologous to the *C. trachomatis* CD4-Th1 inducing proteins. The dissection of the antigen specificity of the protective CD4+ T cell polyclonal population demonstrated that the *Chlamydia* CD4+/IFNγ+ inducing-antigens identified during an infection are also targets of CD4+ T cells that play a part in the rapid clearance of the bacterium in a protective response to the infection, in the absence of antibodies.

Chlamydia T cell lines were derived from Balb/c infected mice and their protective activity was verified in naïve mice against *C. muridarum* challenge. Subsequently, the antigen recognition profile of the *C. muridarum* CD4+ T cell line was characterized to define the possible contribution of each *C. muridarum* antigen in inducing protective CD4+ T cells. For the preparation of Chlamydia – specific CD4+ T cells, splenic CD4+ T lymphocytes were purified from donor Balb/c mice that had previously been infected intranasally with 10^3 viable Elementary Bodies (EBs) of *C. muridarum.* An EB-responding CD4+ T cell line was derived (referred as EB-CD4+ cell line) and expanded *in vitro* with a short term stimulation with heat inactivated EBs. The line showed the capacity to respond to *C. muridarum* EBs by producing IFNγ with a high frequency (data not shown). To determine the efficacy of the EB-CD4+ cell line in resolving an infection, 10^7 CD4+ T cells were adoptively transferred into 4 Balb/c recipient naïve mice. Mice were challenged intranasally 24 hours after i.v. infusion of CD4+ T cells with 10^3 IFUs of *C. muridarum*. The protective effect of adoptive immunization was evaluated by quantitating the number of IFUs recovered from lungs taken 10 days after Chlamydia challenge. As shown in Figure 2a, naïve mice adoptively transferred with EB-CM CD4+ T cells shed 3 Log_{10} fewer IFUs in the lungs 10 days after intranasal challenge with 10^3 IFUs of *C. muridarum*, as compared to either non treated mice (p value: 0.008) or mice receiving an unrelated CD4+ T cell line. Similarly, splenic CD4+ T cells isolated from mice that resolved an intravaginal primary infection with 10^5 IFUs of *C. trachomatis* conferred significant IFU reduction in adoptively transferred mice (data not shown).

To characterize the antigen recognition profile of the *C. muridarum* CD4+ T cells, most of the *C. muridarum* proteins, homolog of the proteins identified as CD4+ Th1 inducers during *C. trachomatis* infection (Figure 1), were obtained in recombinant form and tested for their ability to stimulate the protective EB-CD4+ T cell line to produce IFNγ. In this analysis we excluded both the proteins which after purification did not reach the purity /endoxin level required for the cytokine stimulation assay, or those which, due to their homology with human bacterial proteins were not suitable for developing a vaccine (e.g. heat shock proteins, enolase). The protective EB-CD4+ T cell line was stimulated *in vitro* with a panel of 19 *C. muridarum* recombinant proteins, including MOMP. Fourteen of them were homologs of *C. trachomatis* CD4+ Th1 inducing antigens identified in the primary screening in infected mice, and 5 were negative controls. As shown in Figure 2b, all the 14 CD4+-inducing antigens tested were found also to be targets of the protective EB-CM CD4+ T cell line, and able to induce IFNγ production in a percentage of CD4+ T cells at least 3 times higher than the frequency of negative control antigens. Therefore the pattern of T cell antigens recognized by the protective Chlamydia EB-CM T cell line is comparable to the recognition profile of T cells identified in the *C. trachomatis* infected mice.

Example 3: CT733 and CT153 specific CD4+ Th1 response in BALB/c mice after a primary C. trachomatis infection.

Splenocytes of primary infected BALB/c mice and non infected controls were collected 10 days after infection and stimulated with LPS-free recombinant antigens CT733 and CT153 (20mg/ml). After 4 hours of stimulation, 5mg/ml of Brefeldin A were added to the cells for the following 12 hrs to block cytokine secretion. Afterwards, cells were fixed, permeabilized and stained. Intracellular IFNγ and IL-5 expression were analyzed versus CD4 surface expression of the gated viable cells and assessed by flow cytometry.

The histogram in Figure 3 shows the number of CD4+ T cells per 10^5 CD4+ T splenocytes of primary infected (dark gery bars) and non-infected (light grey bars) mice that produce IFNγ upon specific stimulation with the *C. trachomatis* recombinant antigens CT153 and CT733. The data were confirmed in several further experiments using the same protcol.

The results indicate that CT733 and CT153 are able to induce significant frequencies of specific CD4+/IFNγ+ cells in splenocytes from Balb/c mice that were infected intravaginally with *C. trachomatis*, suggesting a potential role as antigen candidates for these proteins.

Example 4: Protective activity of single antigens TC0106 and TC0431 against C. muridarum challenge.

CT733 and CT153 were tested in a mouse model of chlamydial infection to evaluate their protective properties. This was done by adopting the mouse model of lung infection with the species *Chlamydia muridarum*.

The *C. muridarum* proteins TC0106 and TC0431, homologous to CT733 and CT153, respectively, were cloned and purified, and used for the mouse model.

Groups of BALB/c mice were immunized with either TC0106 or TC0431 recombinant antigens formulated with LTK63+CpG adjuvant (3 doses of 15 ug protein, at 2 week interval, given intramuscularly). As negative control, mice were immunized with the adjuvant only. Four weeks after the last immunization animals were infected intranasally with 10^3 IFU of infectious *C. muridarum*. After 10 days, the protective activity conferred by the two antigens was measured by counting the infectious IFU in the lung of challenge animals.

As shown in Figure 4, each of the two antigens (middle and right hand columns of the histogram) was able to reduce significantly the number of IFU/lung in challenged mice as compared to adjuvant immunized mice (left hand column of the histogram), indicating that both TC0106 and TC0431 (and therefore CT733 and CT153) confer protective immunity to Chlamydia infection

Example 5: Protective activity of the combination of TC0106+TC0431 against C. muridarum challenge.

Groups of BALB/c mice (10 to 15 mice) were immunized with the combination of TC0106+TC0431 recombinant antigens formulated with LTK63+CpG adjuvant (3 doses of 10 ug of each protein at 2 week-interval, given intramuscularly). As negative control, mice were immunized with the adjuvant only. Four weeks after the last immunization, animals were infected intranasally with 10^3 IFU of infectious *C. muridarum*. After 10 days, the protective activity conferred by the two antigens was measured by counting the infectious IFU in the lung of challenge animals. As positive control, a group of mice receiving a primary and a secondary *C. muridarum* infection was also included (left column in the histogram of Figure 5).

As shown in Figure 5, the antigen combination (right hand column of histogram) was able to significantly reduce the number of IFU/lung in challenged mice as compared to adjuvant immunized mice (middle column of histogram).

Thus, immunization with the CT733 and CT153, either alone or in combination, was able to significantly reduce the bacterial load in the lungs of challenged mice (see Figures 4 and 5).

Example 6: Elicitation of CD4+ Th1 cells in BALB/c mice after immunization with TC0431 and TC0106 recombinant antigens, alone or in combination.

Groups of BALB/c mice (10 to 15 mice) were immunized with the recombinant antigens TC0431 and TC0106 as single antigens or in combination (i.m., 10-15 micrograms/dose, 3 doses at 2 week-intervals) using LTK63+CpG adjuvant. Ten days after the third immunization dose, splenocytes were collected and stimulated with LPS-free recombinant antigens (20mg/ml). As negative control, splenocytes of adjuvant immunized mice were included. After 4 hours of stimulation, 5mg/ml of Brefeldin A was added to the cells for the following 12 hrs to block cytokine secretion. Afterwards, cells were fixed, permeabilized and stained. The intracellular IFNγ was analyzed versus CD4 surface expression of the gated viable cells and assessed by flow cytometry. The histogram in Figure 6 shows the number of CD4+ T cells per 10^5 CD4+ T splenocytes that produce IFNγ upon specific stimulation with the recombinant antigens in mice immunized with TC0106, TC0431, the combination of TC0106+TC0431 and adjuvant immunized mice.

The results indicate that immunization with these antigens elicits a high frequency of CD4+ Th1 cells.

Example 7: Evaluation of the protective effect of the chlamydial antigen(s) against C. muridarum challenge.

The protective effect of combinations of two antigens selected from *C. trachomatis* CT279, CT601, CT372, CT443, CT733, CT153, CT456 and CT381 was tested in the *C. muridarum* mouse model using their *C. muridarum* homologues TC0551 (CT279), TC0651 (CT372), TC0727 (CT443),

TC0890 (CT601), TC0106 (CT733), TC0431 (CT153), TC0660 (CT381) and TC0741 (CT456). The protective effect of CT733 and CT153 individually was also tested.

BALB/c mice were immunized three times intramuscularly with a combination of two antigens or single antigens with LTK63+CpG as adjuvant. Twenty-four days post last immunization mice were challenged intranasally with 10^3 IFU *C. muridarum*. After 10 days, lungs were collected, homogenized and the number of viable chlamydiae (IFU/lung) was measured. The data in Figure 7 shows the mean IFU/lung counts in antigen-immunized mice and adjuvant-immunized control. From left to right, the lanes relate to (a) adjuvant only; (b) TC0551+TC0890 (CT279+CT601); (c) TC0651+TC0727 (CT372+CT443); (d) TC0106+TC0431 (CT733+CT153); (e) TC0660+TC0741 (CT456+CT381); (f) TC0106 (CT733); (g) TC0431 (CT153). For each antigen formulation, the numbers of infected mice out of the total immunized are reported in the form "Inf X/Y", wherein X is the number of infected mice and Y is the total number of mice challenged. The statistical significance of immunizing antigen/s (P), was determined by Student t-test.

Four combinations of two antigens have been identified as capable of conferring protection against *C. muridarum* intranasal challenge. For three of them (TC0431+TC0106; TC0727+TC0651; TC0551+TC0890; homologs of CT733+CT153; CT443+CT372; CT279+CT601) protection has been confirmed in a high number of mice using LTK63+CpG adjuvant (Figure 7). Immunization experiments with TC0431 and TC0106 (CT153 and CT733) as single antigens indicate that the two antigens are both immunogenic individually and that either of the two antigens contributes to protection of the CT153+CT733 combination (Figure 7). A fourth antigen combination has been recently identified (TC660+TC0741; homologs of CT456 and CT381) showing protection in an immunization experiment (15 mice) (Figure 7).

The experiments were repeated where the protocol differed from that described above in that the mice were challenged intranasally with 10^3 IFUs of *C. muridarum* three weeks after the last immunization. Since differences in the duration of infections in the animals may occur, the presence of infectious *Chlamydiae* in the lungs was determined in each mouse at days 10 and 12 after challenge. Immunization experiments were repeated at least three times so as to generate data from a statistically significant number of mice. Figure 8 reports the mean number of infectious chlamydiae recovered from lungs of mice immunized with each antigen formulation, in which data obtained at days 10 and 12 were averaged. As shown in Figure 8, two of the four combinations tested in the mouse model, namely TC0551 (CT279 homolog, 82.6 % identity) + TC0890 (CT0601 homolog, 87.6% identity) and TC0106 (CT733 homolog, 84.8% identity) +TC0431 (CT153 homolog, 64.6% identity), showed a statistically significant protective effect in the immunized groups with an IFU reduction of more than 1 Log as compared to the adjuvant-injected mice (P:<0.001). Moreover, 20-25% of the animals immunized with either of the two combinations resolved completely the infection by days 10-12, as compared to 9% of the adjuvant group.

Example 8: Evaluation of the protective activity of the combination TC0551+TC0890+TC0106+TC0431 against challenge with C. muridarum.

On the basis of the result discussed in the preceeding Example, groups of mice were immunized with a combination of four antigens TC0551+TC0890+TC0106+TC0431 using the same immunization regimen as in the Example above. As shown in Figure 8, the 4-antigen combination appeared to have an additive protective effect over the 2-antigen combinations, showing 2.2 Logs reduction of bacterial shedding in the lung (P:0.0003). Moreover, 39% of animals totally resolved the infection, indicating a higher efficacy of this antigen combination in accelerating the bacterial clearance.

The remarkable reduction observed in the number of viable Chlamydiae recovered from the lungs of immunized mice is the first demonstration of a high level of protection induced by systemic immunization with recombinant Chlamydia proteins. It has also to be pointed out that, since denatured forms of the recombinant antigens were used, further optimization of antigen conformation could maximize their protective activity.

Preliminary data aimed at defining whether any of the 4 recombinant antigens were protective when given as single antigens, indicated that a lower level of IFU reduction was observed (less than 1 log) was obtained with any of them (data not shown). This is in agreement with the notion that, in general, combinatorial vaccination approaches are more effective in conferring protective immunity against a given pathogen than single vaccine approaches, since elicited immune responses target different aspects of the bacterial developmental cycle.

Example 9: Evaluation of the protective activity of the combination TC0551+TC0651+TC0727+TC0890 against intraovarian bursa challenge with C. muridarum.

The protective effect of the combination TC0551+TC0651+TC0727+TC0890 (homologs of *C. trachomatis* CT279+CT372+CT443+CT601) was tested in the mouse model of ovarian bursa challenge with *C. muridarum* using the Montanide+CpG adjuvant. This model has previously been described to assess the protective activity of native MOMP (nMOMP), the chlamydial major outer membrane protein (Pal S et al, Infect Immun., 73:8153, 2005). In this model, the protective activity of the antigens is assessed against progression of infection by counting the chlamydia shedding in vaginals swabs.

BALB/c mice were immunized three times intranasally with a combination of the four antigens or with MOMP, with LTK63+CpG as adjuvant. As negative control, a group of mice immunized with ovalbumin was also included. Four weeks after the last immunization, mice received a *C. muridarum* challenge in the ovarian bursa and chlamydial shedding was measured by counting the IFU in the vaginal swabs of infected animals.

The results shown in Figure 9 represent the number of IFU/vaginal swab at two weeks post challenge. As shown in Figure 9, mice receiving the combination of all four antigens show a reduced bacterial shedding as compared to the negative control group (Ovalbumin). Thus, the combination

reduced the progression of infection. Interestingly, the protection level obtained with the combination does not differ significantly from that obtained with nMOMP, which is the most protective antigen that has been identified so far. Thus, this combination of four antigens is a particularly immunogenic combination.

Example 10: Antigen-specific cytokine profiles of protective CD4+ T cells

Given the importance of the CD4-Th1 response in mediating protection from Chlamydia infection, the type of immune response induced by vaccination with two antigen combinations that elicited protection in mice was analysed (TC0551+TC0890 and TC0106+TC0431). In particular, we measured the simultaneous production from antigen-specific CD4+ T cells of IFNγ, TNF-α and IL-2, considering this as an indication of optimal effector functions of CD4+ T cells, possibly improving protection for vaccines aimed at targeting T-cell responses. The assessment of the cytokine profile induced in a single antigen specific CD4+ T cell by vaccination was performed by multiparametric flow cytometric analysis (Perfetto SP et al., Nat.Rev.Immunol. 4, 648-655, 2004) in immunized mice. Peripheral blood was collected 12 days after the last immunization with antigen combinations TC0551+TC0890 and TC0106+TC0431. PBMC were prepared and the frequency of CD3+, CD4+ antigen-specific IFNγ, IL-2 and TNF-producing cells was assessed by intracellular cytokine staining and flow cytometric determination. As shown in Figure 10B, vaccination with the antigen combination TC0551-TC0890 induced a high frequency of TC0551-responding CD4+ T cells producing IFNγ (93 TC0551 specific CD4+ T cells on 10^5 CD4+ cells), while the response to TC0890 was very low, with a frequency of 16 IFNγ+ responding T cells on 10^5 CD4+ cells. The response to the antigen combination used for immunization showed an increased response compared to single antigens, with 132 IFNγ producing T cells on 10^5 CD4+ cells. Furthermore, there was a predominant frequency of multifunctional CD4+ T cells, producing either IFNγ and TNF-α or IFNγ/TNF-α/IL-2 simultaneously. In the control group of mock immunized mice there was no cytokine secretion in response to any recombinant antigen used for stimulation, indicating the specificity of the response observed in the vaccinated mice. As far as the CD4+ response to the antigen combination TC0106-TC0431 is concerned (Figure 10C) both antigens, TC0106 and TC0431 induced a similar response with a frequency respectively of 120 and 98 IFNγ antigen-specific T cells on 10^5 CD4+, while the antigen combination showed a frequency of 145 IFNγ+ responding T cells on 10^5 CD4+ cells. The further analysis of cytokines produced simultaneously with IFNγ showed that about 50% of IFNγ+ cells produced also TNF-α and IL-2, while about 30% of them produced TNF-α. Overall these data underline that the Th1 cytokines produced by antigen-specific CD4+ T cells induced by vaccination showed a functional difference that could reflect differences in the capacity to clear the infection.

Example: 11. Expression analysis of CD4+ inducing Chlamydia antigens.

We then investigated the expression of CT279 (subunit C of Na(+)-translocating NADH-quinone reductase), CT601 (Invasin repeat family phosphatase), CT733 (-Hypothetical protein) and CT153 (MAC-Perforin Protein) by immunoblot analysis both in Ct-EBs and within *C. trachomatis* infected HeLa cells, using their specific mouse immune antisera (Figure 11A). Total protein lysates of renografin-purified EBs (corresponding to approximately 10^7 EBs per lane) showed that each tested antiserum was able to react with a protein band of the expected molecular weight in both EB samples, showing in general a higher reactivity against CM EBs. For analysis of antigen expression in Chlamydia-infected cells, total protein extracts were prepared from Hela 229 cells at different time points after infection (24 – 48 - 72 h) and tested by immunoblot.

The amount of Chlamydial proteins loaded on the gel was normalized on the basis of MOMP expression as described. As shown in Figure 11B, the four antigens appeared to be expressed at different phases of the Chlamydia development.

Finally, we also investigated antigen cellular localization within infected HeLa cells by confocal microscopy in infected Hela cells at 6, 24, 48 and 72 h post infection. As shown in Figure 11B, expression of all antigens was clearly detected within the inclusions at 24h post infection and was still visible at 72h. Interestingly, CT153 staining appeared to accumulate at the inclusion membrane while the other proteins were homogeneously distributed. Since CT153 encodes a MAC-Perforin protein, belonging to a protein family capable of disrupting the cell membrane, the ammassing of this protein at the inclusion membrane might anticipate its involvement in the Chlamydia exit from infected cells.

The analysis of the immune response after vaccination with the combinations has shown that all the recombinant antigens induced a robust humoral response, with the production of IgG2a antibody titers higher than IgG1, as expected for a Th1 driven immune response. Since the resolution of a Chlamydia infection requires a Th1 type of cellular immune response, the regulation of CD4+ Th1 effector and memory cells after vaccination has also been investigated. Differences in the type of cytokines produced by individual cells have important implications for their capacity to mediate effector functions, be sustained as memory T cells or both. CD4+ T cells that secrete only IFNγ have limited capacity to develop into memory T cells as compared with IL-2-IFNγ double positive cells (Hayashi N. et al. 2002). Therefore vaccines eliciting high frequency of single-positive IFNγ producing cells may be limited in their ability to provide long-lasting protection. Furthermore the majority of CD4+ T cells that produce IL-2, IFNγ and TNF are classified as effector memory cells, playing an essential role for mediating protection against intracellular pathogens (Darrah PA et al. 2007). We demonstrated that antigen-specific CD4+ T cells induced by immunization with the protective combinations were predominantly multifunctional, being differentiated to ensure a population of Th1 cells that included either effectors and memory cells. An appropriate balance of

Th1 lineage cells that can be maintained and those with immediate protective functions might be the successful formula for an effective vaccine.

Example 12: Combination of CT823+CT733+CT043+CT456

To evaluate the protective activity of antigens TC0106, TC0313, TC0210, TC0741 and their combination, groups of mice were immunized with the 4 antigens either as single or in a 4 antigen-combination, using the same immunization regimen described in Example 7. The protective activity of the single antigens was assessed by measuring the IFU/Lung at day 12 post infection. The protective activity of the 4-ag combination was measured at days 10, 12, 14 post infection, to evaluate the kinetics of the infection clearance. As shown in Figure 12, the single antigens conferred approximately 0.5-1log IFU reduction in the lung of infected animals.

The four antigens combination showed a highest protective property, indicating a synergic activity of the four antigens in conferring protection, eliciting approximately 4 logs reduction of bacterial shedding in the lung ($P<0.0001$) at day 12 and showing the tendency to resolve the infection at day 12. Moreover a high number of mice (42 %) totally resolved the infection, indicating the efficacy of the antigen combination in accelerating the bacterial clearance.

Example 13: Evaluation of antigenicity of CT812, CT387, CT869, CT166 and CT175

Antigen specific CD4 Tg1 response in BALB/c mice after a primary C.trachomatis (CT infection

The antigen specific CD4 Th1 response in BALB/c mice after a primary *C. trachomatis* (CT) infection was evaluated. *C. trachomatis* antigens identified by the proteomic characterization of the membrane fraction of CT infected HeLa cells were tested for their capability to induce specific CD4+ Th1 response in mice that received an experimental CT infection. Splenocytes of primary infected BALB/c mice and non infected controls were collected 10 days after infection and stimulated with LPS-free recombinant antigens (20μg/ml). After 4 hours of stimulation, 5μg/ml of Brefeldin A was added to the cells for the following 12 hrs, to block cytokine secretion. Afterwards, cells were fixed, permeabilized and stained. The intracellular IFN-γ expression was analyzed versus CD4 surface expression of the gated viable cells, and assessed by flow cytometry. The histogram in Figure 13A and figure 13B show the number of CD4+ T cells that produce IFNγ, upon specific stimulation with CT recombinant antigens per 10^5 CD4+ T splenocytes of primary infected (right hand bars) and not-infected (left hand bars) mice. Data are representative of 4 different experiments. As shown in Figure 13A, CT812C, CT387, CT869 and CT166 induced a significant frequency of $CD4^+$-IFNγ+ cells in splenocytes of infected animals (Pval <0.05). As shown in Figure 13B, CT812C (a C-terminal fragment of CT812) surprisingly induced a higher frequency of $CD4^+$-IFNγ+ cells in splenocytes of infected animals than did the full length CT812 sequence.

Protective activity of the combination of TC0197+TC0261+TC0666 against C. muridarum challenge

The protective effect of the combination of the three *C. trachomatis* antigens CT387+CT812+CT869 was tested in the *C. muridarum* mouse model using their *C. muridarum* orthologues TC0666, TC0197 and TC0261, respectively. TC0197, TC0261 and TC0666 were cloned and purified for protection studies in the mouse model of intranasal infection with *C. muridarum*. Groups of BALB/c mice (16 mice per group) were immunized with the combination of the three recombinant antigens TC0197+TC0261+TC0666 formulated with LTK63+CpG adjuvant (3 doses of 10 μg of each protein, at 2 week-interval, given intramuscularly). As a negative control, mice were immunized with the adjuvant only. Four weeks after the last immunization, animals were infected intranasally with 10^3 IFU of infectious *C. muridarum*. After 12 days, the protective activity conferred by the two antigens was measured by counting the infectious IFU in the lung of challenge animals. As shown in Figure 14, the antigen combination TC0197+TC0261+TC0666 was able to reduce significantly the number of IFU/lung in challenged mice as compared to adjuvant immunized mice (1.4 log IFU reduction with Pval <0.05). The finding that the combination of TC0197+TC0261+TC0666 is able to protect mice against *C. muridarum* challenge (Figure 14) provides evidence that the combinations CT812+CT869+CT387 and CT812C+CT869+CT387 from *C. trachomatis* are protective against infection by *C. trachomatis*.

Protective activity of TC0197, TC0261 and TC0666 as single antigens against C. muridarum challenge

The protective activity of TC0197, TC0261 and TC0666 as single antigens against *C. muridarum* challenge was assessed. 3 Groups of BALB/c mice (16 mice per group) were immunized with the three recombinant antigens individually formulated with LTK63+CpG adjuvant (3 doses of 20 ug of each protein, at 2 week-interval, given intramuscularly). As a negative control, mice were immunized with the adjuvant only. Four weeks after the last immunization, animals were infected intranasally (I.N.) with 10^3 IFU of infectious *C. muridarum*. After 12 days, the protective activity conferred by the three single antigens was measured by counting the infectious IFU in the lung of challenge animals. As shown in Figure 15, none of the 3 antigens individually were able to reduce significantly the number of IFU/lung in challenged mice as compared to adjuvant immunized mice.

Thus, the combination of TC0197+TC0261+TC0666 is able to protect mice against *C. muridarum* challenge (Figure 14). In particular, Figure 14 shows protection in terms of reduction in the mean number of IFUs recovered from lungs of immunized mice versus adjuvant immunized controls [p=0.0024]. In contrast, the three antigens are not protective when administered individually (Figure 15).

Example 14: Materials and methods

The experimental protocols used in Examples 1, 2, 7 (repeated experiments), 8, 10 and 11 are described in further detail in this Example.

Bacterial strains, cultures and reagents

Chlamydia muridarum Nigg and *Chlamydia trachomatis* serovar D strain D/UW-3/CX were grown on confluent monolayers of LLCMK2 (ATCC CCL7) or HeLa 229 cells (ATCC CCL 2.1) in Earle minimal essential medium (EMEM) as described (Caldwell *et al.* (1981) *Infect Immun* 31: 1161-1176). Purification of *C. trachomatis* and *C. muridarum* EBs was carried out by Renografin density gradient centrifugation as described (Montigiani *et al.* (2002) *Infect Immun* 70: 368-379.). Bacteria were aliquoted and stored at -70°C in sucrose-phosphate-glutamine buffer (SPG) until use. When indicated, EBs were heat inactivated at 56°C for 3 hours.

E. coli DH5α or BL21 (DE3) was grown aerobically in Luria Broth (LB) medium (Difco) at 37°C. When appropriate, ampicillin (100 μg/ml) and isopropyl-beta-D-galactopyranoside (IPTG, 0.5 mM) were added to the medium.

Unless specified, all chemicals were purchased from Sigma. Restriction enzymes and DNA modification enzymes were from New England Biolabs. Unless differently stated, all reagents and antibody for intracellular cytokine staining were from BD Biosciences Pharmingen. Confocal microscopy reagents were from Molecular Probes.

Gene cloning, protein expression and preparation of antisera

To produce *C. trachomatis* recombinant proteins and their *C. muridarum* homologs, genes were PCR-amplified from *C. trachomatis* and *C. muridarum* chromosomal DNA using specific primers annealing at the 5' and 3' ends of either gene. The genes were cloned into plasmid pET21b$^+$ (Invitrogen) or pGEXKG (Amersham) in order to express them both as a C-terminal His-tag fusion and as a double fusion protein with an N-terminal Glutathione transferase-encoding sequence and a C-terminal His-tag.

Cloning and purification of His- and GST fusions were performed as already described (Montigiani *et al.*, 2002). CT0681 and TC0052, encoding for *C.trachomatis* and *C. muridarum* MOMP respectively (Ct MOMP and Cm MOMP, respectively) were expressed as His fusions and purified from the insoluble protein fraction. With the exception of TC0313 and TC0210, all the *C. muridarum* proteins used in this work were purified only from the insoluble protein fraction in a denatured form.

For T cell in vitro stimulation assays, LPS-free proteins were prepared by washing of column–immobilized proteins with buffer Tris-HCl 10mM, pH 8, containing 1% Triton X114 (35 ml) at 4°C. The amount of residual endotoxin was determined using a *Limulus* Amebocyte Lysate Analysis Kit (QCL-100, BioWhittaker, Walkerville, MD).

Mouse antisera were generated and treated as described (Montigiani *et al.*, 2002). Where specified, sera from mice immunized with 20 μg of *E. coli* contaminant proteins (IMAC-purified proteins from *E. coli* bacteria containing pET21b+ empty vector) were used as negative control. Western blot, ELISA and Flow cytometry of *C. trachomatis* EBs were performed as described (Finco *et al.* (2005) *Vaccine* 23: 1178-1188.).

Screening of antigen specific CD4-Th1 response in splenocytes from infected mice

Groups of 6 week-old female BALB/c mice purchased from Charles River Laboratories (3 mice/ group) received a subcutaneus hormonal treatment with 2.5 mg of Depo-provera (Medroxyprogesterone acetate) and after five days mice were inoculated intravaginally with 15 μl of SPG buffer containing 10^6 of *C. trachomatis* IFU. The level of infection was analyzed 7 days post-challenge, by collecting vaginal swabs and counting chlamydial inclusions 48h later stained with FITC-conjugated anti Chlamydia antibody (Merifluor) using a UV microscope.

The swabs were collected in 400 μl of SPG and were inoculated on LLCMK2 cell monolayers seeded on 96w flat bottom plates. After 48 hours incubation the number of infectious chlamydiae was determined by counting chlamydial inclusions.

Ten days post challenge mice were sacrificed and their spleens were taken. Splenocytes were prepared by homogenization through a nylon filter (BD) and the erythrocytes were removed by hypotonic lysis in Ack lysis buffer (NH_4Cl 0,155 M, $KHCO_3$ 10 mM, Na_2EDTA 0,1 mM) for 3 minutes at RT, then the cells were plated in 96 wells plates at $2x10^6$ cells per well and stimulated with 20 μg/ml of endotoxin-free specific antigen or with 4 μg/ml of purified EBS in presence of 1 μg/ml anti-CD28 antibody (BD Biosciences Pharmingen) for 4 h at 37 °C. Brefeldin A (BFA; Sigma-Aldrich) was then added at a final concentration of 2.5 μg/ml and cells were incubated for an additional 16 h before intracellular cytokine staining. Cells were stained for viability with LIVE/DEAD® (Molecular Probes) dye according to the manufacturer's instructions. Cells were then fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences Pharmingen) and stained with fluorochrome-labelled monoclonal antibodies for the detection of cells expressing CD3, CD4 on the surface and intracellular IFNγ and IL-4. Finally, cells were resuspended in PBS 1% BSA. All antibodies for intracellular cytokine staining were purchased from BD Pharmingen. Acquisition of the samples was performed using a BD Canto flow cytometer and data were analyzed using FlowJo software (Tree Star Inc., Ashland, USA). The intracellular expression of IFNγ and IL-4 was analysed in CD4 expressing singlet cells, previously gated for, morphology, CD3 expression and viability. Cells were then harvested and stained for CD4 surface expression and IFNγ, or IL-4 intracellular production, to investigate whether the observed responses were of the Th1 (IFNγ) or Th2 (IL-4) type. As negative control, spleens from not infected mice were harvested and analyzed in parallel.

Preparation of CD4+ Th1 cell lines and of antigen presenting cells (APCs)

Splenocytes were prepared by homogenization from spleens from donor Balb/c mice that had previously been infected intranasally with 10^3 viable Elementary Bodies (EBs) of *Chlamydia muridarum* (*C. muridarum*) as decribed above. Following centrifugation at 1200 rpm and suspension in Macs Buffer (PBS PH 7,2 0,5% BSA and 2mM EDTA), the cells were incubated with CD4 (L3T4) microbeads (Milteny Biotec) for 15 minutes and then loaded on a LS columns. The CD4 cells bound to the magnet were recovered, washed and suspended in RPMI 1640 supplemented with 2,5% fetal bovine serum (Hyclone), antibiotics, L-Glutammine 2mM, Sodium Piruvate 1mM, MEM Not essential amino Acids,MEM Vitamins (Gibco) and Beta-mercaptoethanol 0.5 µM. Then the cells were plated in 6 multiwell plates, 10^7 cells/wells. After the first stimulation, the purified CD4 were washed twice and then plated with APCs as described below.

Also a CD4+ cell line with *C. trachomatis* was obtained by spleens from donor Balb/c mice that had previously been infected intravaginally with 10^6 viable Elementary Bodies (EBs) of *Chlamydia trachomatis* and it was performed as described above for *Chlamydia muridarum.*

The CD4 cells were plated ($6x10^6$/well) with APCs ($2x10^7$/well) prepared by naive mice spleens. Splenocytes were prepared as described above, then were washed twice with the medium, gamma irradiated for 7 minutes washed again and suspended in medium.

Cultures were then incubated at 37°C in a humidified atmosphere containing 5%CO_2. After 24 h, Aldesleukin Proleukin (IL2) was added at a concentration of 20U/ml.

C. muridarum and C. trachomatis-mouse model of adoptive transfer

Groups of 6 week-old female BALB/c mice purchased from Charles River Laboratories (4 mice/ group), were adoptively transferred by intravenous administration of 10^7 CD4+ T cells in 100µl of RPMI-1640 medium (Sigma). Mice were challenged intranasally 24 hours after with 10^3 IFUs of *C. muridarum* or 10^5 IFUs of *C. trachomatis*. The effect of adoptive immunization was evaluated by quantitating the number of IFUs recovered from lungs taken 10 days after *C. muridarum* challenge or 6 days after *C. trachomatis* challenge, as described above.

Characterization of the C. muridarum CD4+ T cell line

The same day of the adoptive transfer, an aliquot of purified CD4+ T cells were taken to assess the capability of *C. muridarum* antigens identified in the previous CD4+ Th1+ screening to stimulate them in vitro. 250000 cells/w were plated in 96 multiwell plates with 10^6 mouse splenocytes CD4 depleted as APC and stimulated with 20 µg/ml of *C. muridarum* proteins, homologous to the *C. trachomatis* proteins identified as CD4+ Th1 inducers, in presence of 1 µg/ml anti-CD28 antibody (BD Biosciences Pharmingen) for 3 h at 37 °C. Then BFA was added and intracellular staining was carried out as described for the splenocytes.

Mouse protection model

Groups of 6 week-old female BALB/c mice (10-15 mice/group), were immunized intramuscularly (i.m.) with 3 doses of the antigen combinations TC0551-TC890 (15 μg/dose) and TC0106-TC0431 (containing 10 μg of each protein/dose) at days 1, 15, and 28 formulated with 5μg of LTK63 (Ryan et al., 2000) + 10 μg of CpG (ODN 1826) adjuvant dissolved in 50 μl PBS. As negative control, groups of mice that received the adjuvant alone were included and treated in parallel.

Three weeks after the last immunization mice were inoculated intranasally (i.n.) with 40 μl of SPG buffer containing 10^3 IFU of *C. muridarum*. The Chlamydia challenge dose given to each mouse was confirmed by culturing in triplicate serial dilutions of the inoculating dose on LLCMK2 cell monolayers seeded on 96 wells flat bottom plates. After 24 hours incubation the number of infectious chlamydiae was determined by counting chlamydial inclusions. In the time period between 10- and 12 days post challenge mice were sacrificed, lungs were isolated and their homogenates were used to assess chlamydia growth.

Analysis of antigen specific CD4-Th1 response in PBMC of mice

PBMC from mouse were isolated from up to 2 ml of heparinized blood, diluted 1/5 in HBSS (Hanks' Balanced Salt Solution) and separated by density gradient centrifugation over Lympholite-M (Cedarlane). 10^6 PBMC were plated in duplicate in 96 multiwell plates with 10^6 mouse splenocytes CD4 depleted as APC and stimulated and stained as described above for mouse splenocytes for 16 h. In this staining was analyzed the expression of IFNγ, TNFα and IL-2.

Confocal microscopy

To examine cellular localization of *C.trachomatis* proteins after infection, HeLa cells (20000) were plated on onto glass coverslides (∅ 13 mm) and after 24 hours were infected with CT EBs in 1:1 ratio as described above. At 6, 24, 48 and 72 hours post infection the cells were fixed in 2% paraformaldehyde in PBS buffer for 20 minutes at room temperature. After 2 washes with PBS the cells were permeabilized with a solution of 1%/saponin-0.1% Triton in PBS for 20 minutes.

After washing twice and blocking with PBS containing 1% BSA (PBS-BSA), the cell samples were subjected to antibody and chemical staining. The samples were incubated for 1h at RT (standard dilution 1:5000 in PBS-BSA) with polyclonal antisera obtained from mice immunized with TC601, TC279, TC733 and TC153, previously pre-adsorbed overnight at 4°C onto nitrocellulose strips containing *E. coli* BL21 cell total proteins. Goat anti-mouse Alexa Fluor (Molecular Probes) conjugated antibodies (excitation at 488) were used to visualize the localization of each antigen. Propidium Iodide and Phalloidin conjugated with Alexa Fluor dye A620 (Molecular Probes) were used to visualize respectively DNA and actin.

After extensive washes in PBS, cells were mounted with Anti-Fade reagent (Molecular Probes) and observed under a laser scanning confocal microscope (Bio-Rad) with 100X oil immersion objective lens.

It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention.

TABLE 2

C. pneumoniae accession number & annotation	*C. trachomatis* accession number & annotation	CT No.
	Hypothetical protein (AAC67968)	CT372
	omcB (AAC68042)	CT443
	Hypothetical protein (AAC67634)	CT043
	Hypothetical protein (AAC67744)	CT153
	Nqr3 (AAC67872)	CT279
	papQ (AAC68203)	CT601
	Hypothetical protein (AAC68306)	CT711
	Hypothetical protein (AAC67705)	CT114
	oppA_4 (AAC68080)	CT480
	Hypothetical protein (AAC68056)	CT456
	ArtJ (AAC67977)	CT381
	lcrE (AAC67680)	CT089
	Hypothetical protein (AAC68329)	CT734
	Hypothetical protein (AAC67606)	CT016
gi|4376729|gb|AAD18590.1| Polymorphic Outer Membrane Protein G Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376729|gb|AAD18590.1| Polymorphic Outer Membrane Protein G Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376731|gb|AAD18591.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376731|gb|AAD18591.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329350|gb|AAC68472.1| Putative Outer Membrane Protein I	
gi|4376731|gb|AAD18591.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376733|gb|AAD18593.1| Polymorphic Outer Membrane Protein G Family	gi|3328840|gb|AAC68009.1| Putative outer membrane protein A	
gi|4376731|gb|AAD18591.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376754|gb|AAD18611.1| Polymorphic Outer Membrane Protein (Frame-shift with C	gi|3329344|gb|AAC68467.1| Putative Outer Membrane Protein E	
gi|4376260|gb|AAD18163.1| Polymorphic Outer Membrane Protein G Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376262|gb|AAD18165.1| hypothetical protein	gi|3328765|gb|AAC67940.1| hypothetical protein	
gi|4376269|gb|AAD18171.1| hypothetical protein	gi|3328825|gb|AAC67995.1| hypothetical protein	
gi|4376270|gb|AAD18172.1| Polymorphic Outer Membrane Protein G Family	gi|3329350|gb|AAC68472.1| Putative Outer Membrane Protein I	
gi|4376272|gb|AAD18173.1| Predicted OMP {leader peptide: outer membrane}	gi|3328772|gb|AAC67946.1| hypothetical protein	CT351
gi|4376273|gb|AAD18174.1| Predicted OMP {leader peptide}	gi|3328771|gb|AAC67945.1| hypothetical protein	CT350
gi|4376296|gb|AAD18195.1| hypothetical protein	gi|3328520|gb|AAC67712.1| Ribulose-P Epimerase	
gi|4376362|gb|AAD18254.1| YbbP family hypothetical protein	gi|3328401|gb|AAC67602.1| hypothetical protein	
gi|4376372|gb|AAD18263.1| Signal Peptidase I	gi|3328410|gb|AAC67610.1| Signal Peptidase I	
gi|4376397|gb|AAD18286.1| CHLPS hypothetical protein	gi|3328506|gb|AAC67700.1| CHLPS hypothetical protein	
gi|4376402|gb|AAD18290.1| ACR family	gi|3328505|gb|AAC67699.1| ACR family	
gi|4376419|gb|AAD18305.1| CT149 hypothetical protein	gi|3328551|gb|AAC67740.1| possible hydrolase	
gi|4376446|gb|AAD18330.1| hypothetical protein	gi|3329261|gb|AAC68390.1| hypothetical protein	
gi|4376466|gb|AAD18348.1| Oligopeptide Binding Protein	gi|3328604|gb|AAC67790.1| Oligopeptide Binding Protein	CT198
gi|4376467|gb|AAD18349.1| Oligopeptide Binding Protein	gi|3328604|gb|AAC67790.1| Oligopeptide Binding Protein	
gi|4376468|gb|AAD18350.1| Oligopeptide Binding Protein	gi|3328539|gb|AAC67730.1| Oligopeptide Binding Protein	

gi|4376469|gb|AAD18351.1| Oligopeptide Binding Protein	gi|3328579|gb|AAC67766.1| Oligopeptide binding protein permease	
gi|4376520|gb|AAD18398.1| Polysaccharide Hydrolase-Invasin Repeat Family	gi|3328526|gb|AAC67718.1| predicted polysaccharide hydrolase-invasin repeat family	
gi|4376567|gb|AAD18441.1| Inclusion Membrane Protein C	gi|3328642|gb|AAC67825.1| Inclusion Membrane Protein C	
gi|4376576|gb|AAD18449.1| Omp85 Analog	gi|3328651|gb|AAC67834.1| Omp85 Analog	**CT241**
gi|4376577|gb|AAD18450.1| (OmpH-Like Outer Membrane Protein)	gi|3328652|gb|AAC67835.1| (OmpH-Like Outer Membrane Protein)	**CT242**
gi|4376601|gb|AAD18472.1| Low Calcium Response D	gi|3328486|gb|AAC67681.1| Low Calcium Response D	
gi|4376602|gb|AAD18473.1| Low Calcium Response E	gi|3328485|gb|AAC67680.1| Low Calcium Response E	**CT089**
gi|4376607|gb|AAD18478.1| Phopholipase D Superfamily	gi|3328479|gb|AAC67675.1| Phopholipase D Superfamily {leader (33) peptide}	
gi|4376615|gb|AAD18485.1| YojL hypothetical protein	gi|3328472|gb|AAC67668.1| hypothetical protein	**CT077**
gi|4376624|gb|AAD18493.1| Solute Protein Binding Family	gi|3328461|gb|AAC67658.1| Solute Protein Binding Family	
gi|4376639|gb|AAD18507.1| Flagellar Secretion Protein	gi|3328453|gb|AAC67651.1| Flagellar Secretion Protein	
gi|4376664|gb|AAD18529.1| Leucyl Aminopeptidase A	gi|3328437|gb|AAC67636.1| Leucyl Aminopeptidase A	**CT045**
gi|4376672|gb|AAD18537.1| CBS Domain protein (Hemolysin Homolog)	gi|3328667|gb|AAC67849.1| Hypothetical protein containing CBS domains	
gi|4376679|gb|AAD18543.1| CT253 hypothetical protein	gi|3328664|gb|AAC67846.1| hypothetical protein	
gi|4376696|gb|AAD18559.1| CT266 hypothetical protein	gi|3328678|gb|AAC67859.1| hypothetical protein	**CT266**
gi|4376717|gb|AAD18579.1| Phospholipase D superfamily	gi|3328698|gb|AAC67877.1| Phospholipase D superfamily	
gi|4376727|gb|AAD18588.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376728|gb|AAD18589.1| Polymorphic Outer Membrane Protein G Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376729|gb|AAD18590.1| Polymorphic Outer Membrane Protein G Family	gi|3329350|gb|AAC68472.1| Putative Outer Membrane Protein I	
gi|4376731|gb|AAD18591.1| Polymorphic Outer Membrane Protein G/I Family	gi|3329350|gb|AAC68472.1| Putative Outer Membrane Protein I	
gi|4376733|gb|AAD18593.1| Polymorphic Outer Membrane Protein G Family	gi|3328840|gb|AAC68009.1| Putative outer membrane protein A	
gi|4376735|gb|AAD18594.1| Polymorphic Outer Membrane Protein (truncated) A/I Fam	gi|3328840|gb|AAC68009.1| Putative outer membrane protein A	
gi|4376736|gb|AAD18595.1| Polymorphic Outer Membrane Protein G Family	gi|3329346|gb|AAC68469.1| Putative Outer Membrane Protein G	
gi|4376737|gb|AAD18596.1| Polymorphic Outer Membrane Protein H Family	gi|3329347|gb|AAC68470.1| Putative Outer Membrane Protein H	
gi|4376751|gb|AAD18608.1| Polymorphic Outer Membrane Protein E Family	gi|3329344|gb|AAC68467.1| Putative Outer Membrane Protein E	
gi|4376752|gb|AAD18609.1| Polymorphic Outer Membrane Protein E Family	gi|3329344|gb|AAC68467.1| Putative Outer Membrane Protein E	
gi|4376753|gb|AAD18610.1| Polymorphic Outer Membrane Protein E/F Family	gi|3329344|gb|AAC68467.1| Putative Outer Membrane Protein E	
gi|4376757|gb|AAD18613.1| hypothetical protein	gi|3328701|gb|AAC67880.1| PP-loop superfamily ATPase	
gi|4376767|gb|AAD18622.1| Arginine Periplasmic Binding Protein	gi|3328806|gb|AAC67977.1| Arginine Binding Protein	**CT381**
gi|4376790|gb|AAD18643.1| Heat Shock Protein-70	gi|3328822|gb|AAC67993.1| HSP-70	**CT396**
gi|4376802|gb|AAD18654.1| CT427 hypothetical protein	gi|3328857|gb|AAC68024.1| hypothetical protein	
gi|4376814|gb|AAD18665.1| CT398 hypothetical protein	gi|3328825|gb|AAC67995.1| hypothetical protein	**CT398**
gi|4376829|gb|AAD18679.1| polymorphic membrane protein A Family	gi|3328840|gb|AAC68009.1| Putative outer membrane protein A	
gi|4376830|gb|AAD18680.1| polymorphic membrane protein B Family	gi|3328841|gb|AAC68010.1| Putative outer membrane protein B	
gi|4376832|gb|AAD18681.1| Solute binding protein	gi|3328844|gb|AAC68012.1| Solute-binding protein	**CT415**
gi|4376834|gb|AAD18683.1| (Metal Transport Protein)	gi|3328846|gb|AAC68014.1| (Metal Transport Protein)	
gi|4376847|gb|AAD18695.1| Tail-Specific Protease	gi|3328872|gb|AAC68040.1| Tail-Specific Protease	
gi|4376848|gb|AAD18696.1| 15 kDa Cysteine-Rich Protein	gi|3328873|gb|AAC68041.1| 15kDa Cysteine-Rich Protein	
gi|4376849|gb|AAD18697.1| 60 kDa Cysteine-Rich OMP	gi|3328874|gb|AAC68042.1| 60kDa Cysteine-Rich OMP	**CT443**
gi|4376850|gb|AAD18698.1| 9 kDa-Cysteine-Rich Lipoprotein	gi|3328876|gb|AAC68043.1| 9kDa-Cysteine-Rich Lipoprotein	**CT444**

gi|4376878|gb|AAD18723.1| 2-Component Sensor	gi|3328901|gb|AAC68067.1| 2-component regulatory system-sensor histidine kinase	CT467
gi|4376879|gb|AAD18724.1| similarity to CHLPS IncA	gi|3328451|gb|AAC67649.1| hypothetical protein	
gi|4376884|gb|AAD18729.1| CT471 hypothetical protein	gi|3328905|gb|AAC68071.1| hypothetical protein	
gi|4376886|gb|AAD18731.1| YidD family	gi|3328908|gb|AAC68073.1| hypothetical protein	
gi|4376890|gb|AAD18734.1| CT476 hypothetical protein	gi|3328911|gb|AAC68076.1| hypothetical protein	
gi|4376892|gb|AAD18736.1| Oligopeptide Permease	gi|3328913|gb|AAC68078.1| Oligopeptide Permease	
gi|4376894|gb|AAD18738.1| Oligopeptide Binding Lipoprotein	gi|3328915|gb|AAC68080.1| oligopeptide Binding Lipoprotein	
gi|4376900|gb|AAD18743.1| Glutamine Binding Protein	gi|3328922|gb|AAC68086.1| Glutamine Binding Protein	
gi|4376909|gb|AAD18752.1| Protease	gi|6578107|gb|AAC68094.2| Protease	
gi|4376952|gb|AAD18792.1| Apolipoprotein N-Acetyltransferase	gi|3328972|gb|AAC68136.1| Apolipoprotein N-Acetyltransferase	
gi|4376960|gb|AAD18800.1| FKBP-type peptidyl-prolyl cis-trans isomerise	gi|3328979|gb|AAC68143.1| FKBP-type peptidyl-prolyl cis-trans isomerise	CT541
gi|4376968|gb|AAD18807.1| CT547 hypothetical protein	gi|3328986|gb|AAC68149.1| hypothetical protein	CT547
gi|4376969|gb|AAD18808.1| CT548 hypothetical protein	gi|3328987|gb|AAC68150.1| hypothetical protein	
gi|4376998|gb|AAD18834.1| Major Outer Membrane Protein	gi|3329133|gb|AAC68276.1| Major Outer Membrane Protein	CT681
gi|4377005|gb|AAD18841.1| YopC/Gen Secretion Protein D	gi|3329125|gb|AAC68269.1| probable Yop proteins translocation protein	
gi|4377015|gb|AAD18851.1| FHA domain; (homology to adenylate cyclase)	gi|3329115|gb|AAC68259.1| (FHA domain; homology to adenylate cyclase)	
gi|4377033|gb|AAD18867.1| CHLPN 76 kDa Homolog_1 (CT622)	gi|3329069|gb|AAC68226.1| CHLPN 76kDa Homolog	CT622
gi|4377034|gb|AAD18868.1| CHLPN 76 kDa Homolog_2 (CT623)	gi|6578109|gb|AAC68227.2| CHLPN 76kDa Homolog	CT623
gi|4377035|gb|AAD18869.1| Integral Membrane Protein	gi|3329071|gb|AAC68228.1| Integral Membrane Protein	
gi|4377072|gb|AAD18902.1| CT648 hypothetical protein	gi|3329097|gb|AAC68825.1| hypothetical protein	
gi|4377073|gb|AAD18903.1| CT647 hypothetical protein	gi|3329096|gb|AAC68824.1| hypothetical protein	CT647
gi|4377085|gb|AAD18914.1| CT605 hypothetical protein	gi|3329050|gb|AAC68208.1| hypothetical protein	
gi|4377090|gb|AAD18919.1| Peptidoglycan-Associated Lipoprotein	gi|3329044|gb|AAC68202.1| Peptidoglycan-Associated Lipoprotein	CT600
gi|4377091|gb|AAD18920.1| macromolecule transporter	gi|3329043|gb|AAC68201.1| component of a macromolecule transport system	
gi|4377092|gb|AAD18921.1| CT598 hypothetical protein	gi|3329042|gb|AAC68200.1| hypothetical protein	
gi|4377093|gb|AAD18922.1| Biopolymer Transport Protein	gi|3329041|gb|AAC68199.1| Biopolymer Transport Protein	CT597
gi|4377094|gb|AAD18923.1| Macromolecule transporter	gi|3329040|gb|AAC68198.1| polysaccharide transporter	
gi|4377101|gb|AAD18929.1| CT590 hypothetical protein	gi|3329033|gb|AAC68192.1| hypothetical protein	
gi|4377102|gb|AAD18930.1| CT589 hypothetical protein	gi|3329032|gb|AAC68191.1| hypothetical protein	CT589
gi|4377106|gb|AAD18933.1| hypothetical protein	gi|3328796|gb|AAC67968.1| hypothetical protein	
gi|4377111|gb|AAD18938.1| Enolase	gi|3329030|gb|AAC68189.1| Enolase	CT587
gi|4377127|gb|AAD18953.1| General Secretion Protein D	gi|3329013|gb|AAC68174.1| Gen. Secretion Protein D	
gi|4377130|gb|AAD18956.1| predicted OMP {leader peptide}	gi|3329010|gb|AAC68171.1| predicted OMP	CT569
gi|4377132|gb|AAD18958.1| CT567 hypothetical protein	gi|3329008|gb|AAC68169.1| hypothetical protein	CT567
gi|4377133|gb|AAD18959.1| CT566 hypothetical protein	gi|3329007|gb|AAC68168.1| hypothetical protein	
gi|4377140|gb|AAD18965.1| Yop Translocation J	gi|3329000|gb|AAC68161.1| Yop proteins translocation lipoprotein J	CT559
gi|4377170|gb|AAD18992.1| Outer Membrane Protein B	gi|3329169|gb|AAC68308.1| Outer Membrane Protein Analog	CT713
gi|4377177|gb|AAD18998.1| Flagellar M-Ring Protein	gi|3329175|gb|AAC68314.1| Flagellar M-Ring Protein	
gi|4377182|gb|AAD19003.1| CT724 hypothetical protein	gi|3329181|gb|AAC68319.1| hypothetical protein	
gi|4377184|gb|AAD19005.1| Rod Shape Protein	gi|3329183|gb|AAC68321.1| Rod Shape Protein	
gi|4377193|gb|AAD19013.1| CT734 hypothetical protein	gi|3329192|gb|AAC68329.1| hypothetical protein	
gi|4377206|gb|AAD19025.1| CHLTR possible phosphoprotein	gi|3329204|gb|AAC68339.1| CHLTR possible phosphoprotein	
gi|4377222|gb|AAD19040.1| Muramidase (invasin repeat family)	gi|3329221|gb|AAC68354.1| Muramidase (invasin repeat family)	CT759
gi|4377223|gb|AAD19041.1| Cell Division Protein FtsW	gi|3329222|gb|AAC68355.1| Cell Division Protein FtsW	
gi|4377224|gb|AAD19042.1| Peptidoglycan Transferase	gi|3329223|gb|AAC68356.1| Peptidoglycan Transferase	CT761
gi|4377225|gb|AAD19043.1| Muramate-Ala Ligase & D-Ala-D-Ala Ligase	gi|3329224|gb|AAC68357.1| UDP-N-acetylmuramate-alanine ligase	
gi|4377248|gb|AAD19064.1| Thioredoxin Disulfide Isomerase	gi|3329244|gb|AAC68375.1| Thioredoxin Disulfide Isomerase	

gi|4377261|gb|AAD19076.1| CT788 hypothetical protein - {leader peptide-periplasmi	gi|3329253|gb|AAC68383.1| {leader (60) peptide-periplasmic}	
gi|4377280|gb|AAD19093.1| Insulinase family/Protease III	gi|3329273|gb|AAC68402.1| Insulinase family/Protease III	
gi|4377287|gb|AAD19099.1| Putative Outer Membrane Protein D Family	gi|3329279|gb|AAC68408.1| Putative Outer Membrane Protein D	
gi|4377306|gb|AAD19116.1| DO Serine Protease	gi|3329293|gb|AAC68420.1| DO Serine Protease	**CT823**
gi|4377342|gb|AAD19149.1| ABC transporter permease	gi|3329327|gb|AAC68451.1| ABC transporter permease — pyrimidine biosynthesis protein	
gi|4377347|gb|AAD19153.1| CT858 hypothetical protein	gi|6578118|gb|AAC68456.2| predicted Protease containing IRBP and DHR domains	
gi|4377353|gb|AAD19159.1| CT863 hypothetical protein	gi|3329337|gb|AAC68461.1| hypothetical protein	
gi|4377367|gb|AAD19171.1| Predicted OMP	gi|3328795|gb|AAC67967.1| hypothetical protein	
gi|4377408|gb|AAD19209.1| hypothetical protein	gi|3328795|gb|AAC67967.1| hypothetical protein	
gi|4377409|gb|AAD19210.1| Predicted Outer Membrane Protein (CT371)	gi|3328795|gb|AAC67967.1| hypothetical protein	
gi|4376411|gb|	gi|3328512|gb|AAC67705.1| hypothetical protein	**CT114**
gi|4376508|gb|	gi|3328585|gb|AAC67772.1| hypothetical protein	**CT181**
gi|4376710|gb|	gi|3328692|gb|AAC67872.1| NADH (Ubiquinone) Oxidoreductase, Gamma	**CT279**
gi|4376777|gb|	gi|3328815|gb|AAC67986.1| hypothetical protein	**CT389**
gi|4376782|gb|	gi|3328817|gb|AAC67988.1| hypothetical protein	**CT391**
gi|4376863|gb|	gi|3328887|gb|AAC68054.1| Arginyl tRNA transferase	**CT454**
gi|4376866|gb|	gi|3328889|gb|AAC68056.1| hypothetical protein	**CT456**
gi|4376972|gb|	gi|3328991|gb|AAC68153.1| D-Ala-D-Ala Carboxypeptidase	**CT551**
gi|4377139|gb|	gi|3329001|gb|AAC68162.1| hypothetical protein	**CT560**
gi|4377154|gb|	gi|3329154|gb|AAC68295.1| hypothetical protein	**CT700**

SEQUENCE LISTING

SEQ ID NO: 1- CT733 nucleotide sequence

ATGTTAATAAACTTTACCTTTCGCAACTGTCTTTTGTTCCTTGTCACACTGTCTAGTGTCCCTGTTTTCTCAGCACC
TCAACCTCGCGGAACGCTTCCTAGCTCGACCACAAAAATTGGATCAGAAGTTTGGATTGAACAAAAAGTCCGCCAAT
ATCCAGAGCTTTTATGGTTAGTAGAGCCGTCCTCTACGGGAGCCTCTTTAAAATCTCCTTCAGGAGCCATCTTTTCT
CCAACATTATTCCAAAAAAAGGTCCCTGCTTTCGATATCGCAGTGCGCAGTTTGATTCACTTACATTTATTAATCCA
GGGTTCCCGCCAAGCCTATGCTCAACTGATCCAACTACAGACCAGCGAATCCCCTCTAACATTTAAGCAATTCCTTG
CATTGCATAAGCAATTAACTCTATTTTTAAATTCCCCTAAGGAATTTTATGACTCTGTTAAAGTGTTAGAGACAGCT
ATCGTCTTACGTCACTTAGGCTGTTCAACTAAGGCTGTTGCTGCGTTTAAACCTTATTTCTCAGAAATGCAAAGAGA
GGCTTTTTACACTAAGGCTCTGCATGTACTACACACCTTCCCAGAGCTAAGCCCATCATTTGCTCGCCTCTCTCCGG
AGCAGAAAACTCTCTTCTTCTCCTTGAGAAAATTGGCGAATTACGATGAGTTACTCTCGCTGACGAACACCCCAAGT
TTTCAGCTTCTGTCTGCTGGGCGCTCGCAACGAGCTCTTTTAGCTCTGGACTTGTACCTCTATGCTTTGGATTCCTG
TGGAGAACAGGGGATGTCCTCTCAATTCCACACAAACTTCGCACCTCTACAGTCCATGTTGCAACAATACGCTACTG
TAGAAGAGGCCTTTTCTCGTTATTTTACTTACCGAGCTAATCGATTAGGATTTGATGGCTCTTCTCGATCCGAGATG
GCTTTAGTAAGAATGGCCACCTTGATGAACTTGTCTCCTTCCGAAGCTGCGATTTTAACCACAAGCTTCAAAACCCT
TCCTACAGAAGAAGCGGATACTTTGATCAATAGTTTCTATACCAATAAGGGCGATTCGTTGGCTCTTTCTCTGCGAG
GGTTGCCTACACTTGTATCCGAACTGACGCGAACTGCCCATGGCAATACCAATGCAGAAGCTCGATCTCAGCAAATT
TATGCAACTACCCTATCGCTAGTAGTAAAGAGTCTGAAAGCGCACAAAGAAATGCTAAACAAGCAAATTCTTTCTAA
GGAAATTGTTTTAGATTTCTCAGAAACTGCAGCTTCTTGCCAAGGATTGGATATCTTTTCCGAGAATGTCGCTGTTC
AAATTCACTTAAATGGAACCGTTAGTATCCATTTATAA

SEQ ID NO: 2 - CT733 protein sequence

MLINFTFRNCLLFLVTLSSVPVFSAPQPRGTLPSSTTKIGSEVWIEQKVRQYPELLWLVEPSSTGASLKSPSGAIFS
PTLFQKKVPAFDIAVRSLIHLHLLIQGSRQAYAQLIQLQTSESPLTFKQFLALHKQLTLFLNSPKEFYDSVKVLETA
IVLRHLGCSTKAVAAFKPYFSEMQREAFYTKALHVLHTFPELSPSFARLSPEQKTLFFSLRKLANYDELLSLTNTPS
FQLLSAGRSQRALLALDLYLYALDSCGEQGMSSQFHTNFAPLQSMLQQYATVEEAFSRYFTYRANRLGFDGSSRSEM
ALVRMATLMNLSPSEAAILTTSFKTLPTEEADTLINSFYTNKGDSLALSLRGLPTLVSELTRTAHGNTNAEARSQQI
YATTLSLVVKSLKAHKEMLNKQILSKEIVLDFSETAASCQGLDIFSENVAVQIHLNGTVSIHL

SEQ ID NO:3 – CT153 nucleotide sequence

ATGACTAAGCCTTCTTTCTTATACGTTATTCAACCTTTTTCCGTATTTAATCCACGATTAGGACGTTTCTCTACAGA
CTCAGATACTTATATCGAAGAAGAAAACCGCCTAGCATCGTTCATTGAGAGTTTGCCACTGGAGATCTTCGATATAC
CTTCTTTCATGGAAACCGCGATTTCCAATAGCCCCTATATTTTATCTTGGGAGACAACTAAAGACGGCGCTCTGTTC
ACTATTCTTGAACCCAAACTCTCAGCTTGCGCAGCCACTTGCCTGGTAGCCCCTTCTATACAAATGAAATCCGATGC
GGAGCTCCTAGAAGAAATTAAGCAAGCGTTATTACGCAGCTCTCATGACGGTGTGAAATATCGCATCACCAGAGAAT
CCTTCTCTCCAGAAAAGAAAACTCCTAAGGTTGCTCTAGTCGATGACGATATTGAATTGATTCGCAATGTCGACTTT
TTGGGTAGAGCTGTTGACATTGTCAAATTAGACCCTATTAATATTCTGAATACCGTAAGCGAAGAGAATATTCTAGA
TTACTCTTTTACAAGAGAAACGGCTCAGCTGAGCGCGGATGGTCGTTTTGGTATTCCTCCAGGGACTAAGCTATTCC
CTAAACCTTCTTTTGATGTAGAAATCAGTACCTCCATTTTCGAAGAAACAACTTCATTTACTCGAAGTTTTTCTGCA
TCGGTTACTTTTAGTGTACCAGACCTCGCGGCGACTATGCCTCTTCAAAGCCCTCCCATGGTAGAAAATGGTCAAAA
AGAAATTTGTGTCATTCAAAAACACTTATTCCCAAGCTACTCTCCTAAACTAGTCGATATTGTTAAACGATACAAAA
GAGAGGCTAAGATCTTGATTAACAAGCTTGCCTTTGGAATGTTATGGCGACATCGGGCTAAAAGCCAAATCCTCACC
GAGGGAAGCGTACGTCTAGACTTACAAGGATTCACAGAATCGAAGTACAATTACCAGATTCAAGTAGGATCCCATAC
GATTGCAGCTGTATTAATCGATATGGATATTTCCAAGATTCAATCCAAATCAGAACAAGCTTATGCAATTAGGAAAA
TCAAATCAGGCTTTCAACGTAGCTTGGATGACTATCATATTTATCAAATTGAAAGAAAACAAACCTTTTCTTTTTCT
CCGAAGCATCGCAGCCTCTCATCCACATCCCATTCCGAAGATTCTGATTTGGATCTTTCTGAAGCAGCCGCCTTTTC
AGGAAGTCTTACCTGCGAGTTTGTAAAAAAAAGCACTCAACATGCCAAGAATACCGTCACATGTTCCACAGCCGCTC
ATTCCCTATACACACTCAAAGAAGATGACAGCTCGAACCCCTCTGAAAAACGATTAGATAGTTGTTTCCGCAATTGG
ATTGAAAACAAACTAAGCGCCAATTCTCCAGATTCCTGGTCAGCGTTTATTCAAAAATTCGGAACACACTATATTGC
ATCAGCAACTTTTGGAGGGATAGGTTTCCAAGTGCTCAAACTATCTTTTGAACAGGTGGAGGATCTACATAGCAAAA
AGATCTCCTTAGAAACCGCAGCAGCCAACTCTCTATTAAAAGGTTCTGTATCCAGCAGCACAGAATCTGGATACTCC
AGCTATAGCTCCACGTCTTCTTCTCATACGGTATTTTTAGGAGGAACGGTCTTACCTTCGGTTCATGATGAACGTTT
AGACTTTAAAGATTGGTCGGAAAGTGTGCACCTGGAACCTGTTCCTATCCAGGTTTCTTTACAACCTATAACGAATT
TACTAGTTCCTCTCCATTTTCCTAATATCGGTGCTGCAGAGCTCTCTAATAAACGAGAATCTCTTCAACAAGCGATT
CGAGTCTATCTCAAAGAACATAAAGTAGATGAGCAAGGAGAACGTACTACATTTACATCAGGAATCGATAATCCTTC
TTCCTGGTTTACCTTAGAAGCTGCCCACTCTCCTCTTATAGTCAGTACTCCTTACATTGCTTCGTGGTCTACGCTTC
CTTATTTGTTCCCAACATTAAGAGAACGTTCTTCGGCAACCCCTATCGTTTTCTATTTTTGTGTAGATAATAATGAA
CATGCTTCGCAAAAAATATTAAACCAATCGTATTGCTTCCTCGGGTCCTTGCCTATTCGACAAAAAATTTTTGGTAG
CGAATTTGCTAGTTTCCCCTATCTATCTTTCTATGGAAATGCAAAAGAGGCGTACTTTGATAACACGTACTACCCAA
CGCGTTGTGGGTGGATTGTTGAAAAGTTAAATACTACACAAGATCAATTCCTCCGGGATGGAGACGAGGTGCGACTA
AAACATGTTTCCAGCGGAAAGTATCTAGCAACAACTCCTCTTAAGGATACCCATGGTACACTCACGCGTACAACGAA
CTGTGAAGATGCTATCTTTATTATTAAAAAATCTTCAGGTTATTGA

SEQ ID NO:4 - CT153 protein sequence

MTKPSFLYVIQPFSVFNPRLGRFSTDSDTYIEEENRLASFIESLPLEIFDIPSFMETAISNSPYILSWETTKDGALF
TILEPKLSACAATCLVAPSIQMKSDAELLEEIKQALLRSSHDGVKYRITRESFSPEKKTPKVALVDDDIELIRNVDF
LGRAVDIVKLDPINILNTVSEENILDYSFTRETAQLSADGRFGIPPGTKLFPKPSFDVEISTSIFEETTSFTRSFSA
SVTFSVPDLAATMPLQSPPMVENGQKEICVIQKHLFPSYSPKLVDIVKRYKREAKILINKLAFGMLWRHRAKSQILT
EGSVRLDLQGFTESKYNYQIQVGSHTIAAVLIDMDISKIQSKSEQAYAIRKIKSGFQRSLDDYHIYQIERKQTFSFS
PKHRSLSSTSHSEDSDLDLSEAAAFSGSLTCEFVKKSTQHAKNTVTCSTAAHSLYTLKEDDSSNPSEKRLDSCFRNW
IENKLSANSPDSWSAFIQKFGTHYIASATFGGIGFQVLKLSFEQVEDLHSKKISLETAAANSLLKGSVSSSTESGYS
SYSSTSSSHTVFLGGTVLPSVHDERLDFKDWSESVHLEPVPIQVSLQPITNLLVPLHFPNIGAAELSNKRESLQQAI
RVYLKEHKVDEQGERTTFTSGIDNPSSWFTLEAAHSPLIVSTPYIASWSTLPYLFPTLRERSSATPIVFYFCVDNNE
HASQKILNQSYCFLGSLPIRQKIFGSEFASFPYLSFYGNAKEAYFDNTYYPTRCGWIVEKLNTTQDQFLRDGDEVRL
KHVSSGKYLATTPLKDTHGTLTRTTNCEDAIFIIKKSSGY

SEQ ID NO:5 - CT601 nucleotide sequence

ATGCTCGCTAATCGCTTATTCTTAATAACCCTTTTAGGGTTAAGTTCGTCTGTTTACGGCGCAGGTAAAGCACCGTC
TTTGCAGGCTATTCTAGCCGAAGTCGAAGACACCTCCTCTCGTCTACACGCTCATCACAATGAGCTTGCTATGATCT
CTGAACGCCTCGATGAGCAAGACACGAAACTACAGCAACTTTCGTCAACACAAGATCATAACCTACCTCGACAAGTT
CAGCGACTAGAAACGGACCAAAAAGCTTTGGCAAAAACACTGGCGATTCTTTCGCAATCCGTCCAAGATATTCGGTC
TTCTGTACAAAATAAATTACAAGAAATCCAACAAGAACAAAAAAAAATTAGCACAAAATTTGCGAGCGCTTCGTAACT
CTTTACAAGCTCTCGTTGATGGCTCTTCTCCAGAAAATTATATTGATTTCCTAACTGGTGAAACCCCGGAACATATT
CATATTGTTAAACAAGGAGAGACCCTGAGCAAGATCGCGAGTAAATATAACATCCCCGTCGTAGAATTAAAAAAACT
TAATAAACTAAATTCGGATACTATTTTTACAGATCAAAGAATTCGCCTTCCGAAAAAGAAATAG

SEQ ID NO:6 - CT601 protein sequence

MLANRLFLITLLGLSSSVYGAGKAPSLQAILAEVEDTSSRLHAHHNELAMISERLDEQDTKLQQLSSTQDHNLPRQV
QRLETDQKALAKTLAILSQSVQDIRSSVQNKLQEIQQEQKKLAQNLRALRNSLQALVDGSSPENYIDFLTGETPEHI
HIVKQGETLSKIASKYNIPVVELKKLNKLNSDTIFTDQRIRLPKKK

SEQ ID NO:7 - CT279 nucleotide sequence

ATGGCATCCAAGTCTCGCCATTATCTTAATCAGCCTTGGTACATTATCTTATTCATCTTTGTTCTTAGTTTAATTGC
TGGTACCCTCCTGTCTTCTGTGTATTATGTCCTTGCACCTATCCAACAGCAAGCTGCGGAATTCGATCGCAATCAAC
AAATGCTAATGGCTGCACAAGTAATTTCTTCCGATAACACATTCCAAGTCTATGAAAAGGGAGATTGGCACCCAGCC
CTATATAATACTAAAAAGCAGTTGCTAGAGATCTCCTCTACTCCTCCTAAAGTAACCGTGACAACTTTAAGCTCATA
TTTTCAAAACTTTGTTAGAGTCTTGCTTACAGATACACAAGGAAATCTTTCTTCATTCGAAGACCATAATCTCAATC
TAGAAGAATTTTTATCTCAACCAACTCCTGTAATACATGGTCTTGCCCTTTATGTGGTCTACGCTATCCTACACAAC
GATGCAGCTTCCTCTAAATTATCTGCTTCCCAAGTAGCGAAAAATCCAACAGCTATAGAATCTATAGTTCTTCCTAT
AGAAGGTTTTGGTTTGTGGGGACCTATCTATGGATTCCTTGCTCTAGAAAAAGACGGGAATACTGTTCTTGGTACTT
CTTGGTATCAACATGGCGAGACTCCTGGATTAGGAGCAAATATCGCTAACCCTCAATGGCAAAAAAATTTCAGAGGC
AAAAAAGTATTTCTAGTCTCAGCTTCTGGAGAAACAGATTTTGCTAAGACAACCCTAGGACTGGAAGTTATAAAAGG
ATCTGTATCTGCAGCATTAGGAGACTCACCTAAAGCTGCTTCTTCCATCGACGGAATTTCAGGAGCTACTTTGACTT
GTAATGGTGTTACCGAATCCTTCTCTCATTCTCTAGCTCCCTACCGCGCTTTGTTGACTTTCTTCGCCAACTCTAAA
CCTAGTGGAGAGTCTCATGACCACTAA

SEQ ID NO:8 - CT279 protein sequence

MASKSRHYLNQPWYIILFIFVLSLIAGTLLSSVYYVLAPIQQQAAEFDRNQQMLMAAQVISSDNTFQVYEKGDWHPA
LYNTKKQLLEISSTPPKVTVTTLSSYFQNFVRVLLTDTQGNLSSFEDHNLNLEEFLSQPTPVIHGLALYVVYAILHN
DAASSKLSASQVAKNPTAIESIVLPIEGFGLWGPIYGFLALEKDGNTVLGTSWYQHGETPGLGANIANPQWQKNFRG
KKVFLVSASGETDFAKTTLGLEVIKGSVSAALGDSPKAASSIDGISGATLTCNGVTESFSHSLAPYRALLTFFANSK
PSGESHDH

SEQ ID NO:9 - CT443 nucleotide sequence

ATGCGAATAGGAGATCCTATGAACAAACTCATCAGACGAGCAGTGACGATCTTCGCGGTGACTAGTGTGGCGAGTTT
ATTTGCTAGCGGGGTGTTAGAGACCTCTATGGCAGAGTCTCTCTCTACAAACGTTATTAGCTTAGCTGACACCAAAG
CGAAAGACAACACTTCTCATAAAAGCAAAAAAGCAAGAAAAAACCACAGCAAAGAGACTCCCGTAGACCGTAAAGAG
GTTGCTCCGGTTCATGAGTCTAAAGCTACAGGACCTAAACAGGATTCTTGCTTTGGCAGAATGTATACAGTCAAAGT
TAATGATGATCGCAATGTTGAAATCACACAAGCTGTTCCTGAATATGCTACGGTAGGATCTCCCTATCCTATTGAAA
TTACTGCTACAGGTAAAAGGGATTGTGTTGATGTTATCATTACTCAGCAATTACCATGTGAAGCAGAGTTCGTACGC
AGTGATCCAGCGACAACTCCTACTGCTGATGGTAAGCTAGTTTGGAAAATTGACCGCTTAGGACAAGGCGAAAAGAG
TAAAATTACTGTATGGGTAAAACCTCTTAAAGAAGGTTGCTGCTTTACAGCTGCAACAGTATGCGCTTGTCCAGAGA
TCCGTTCGGTTACAAAATGTGGACAACCTGCTATCTGTGTTAAACAAGAAGGCCCAGAGAATGCTTGTTTGCGTTGC
CCAGTAGTTTACAAAATTAATATAGTGAACCAAGGAACAGCAACAGCTCGTAACGTTGTTGTTGAAAATCCTGTTCC
AGATGGTTACGCTCATTCTTCTGGACAGCGTGTACTGACGTTTACTCTTGGAGATATGCAACCTGGAGAGCACAGAA
CAATTACTGTAGAGTTTTGTCCGCTTAAACGTGGTCGTGCTACCAATATAGCAACGGTTTCTTACTGTGGAGGACAT
AAAAATACAGCAAGCGTAACAACTGTGATCAACGAGCCTTGCGTACAAGTAAGTATTGCAGGAGCAGATTGGTCTTA
TGTTTGTAAGCCTGTAGAATATGTGATCTCCGTTTCCAATCCTGGAGATCTTGTGTTGCGAGATGTCGTCGTTGAAG
ACACTCTTTCTCCCGGAGTCACAGTTCTTGAAGCTGCAGGAGCTCAAATTTCTTGTAATAAAGTAGTTTGGACTGTG

AAAGAACTGAATCCTGGAGAGTCTCTACAGTATAAAGTTCTAGTAAGAGCACAAACTCCTGGACAATTCACAAATAA
TGTTGTTGTGAAGAGCTGCTCTGACTGTGGTACTTGTACTTCTTGCGCAGAAGCGACAACTTACTGGAAAGGAGTTG
CTGCTACTCATATGTGCGTAGTAGATACTTGTGACCCTGTTTGTGTAGGAGAAAATACTGTTTACCGTATTTGTGTC
ACCAACAGAGGTTCTGCAGAAGATACAAATGTTTCTTTAATGCTTAAATTCTCTAAAGAACTGCAACCTGTATCCTT
CTCTGGACCAACTAAAGGAACGATTACAGGCAATACAGTAGTATTCGATTCGTTACCTAGATTAGGTTCTAAAGAAA
CTGTAGAGTTTTCTGTAACATTGAAAGCAGTATCAGCTGGAGATGCTCGTGGGGAAGCGATTCTTTCTTCCGATACA
TTGACTGTTCCAGTTTCTGATACAGAGAATACACACATCTATTAA

SEQ ID NO:10 - CT443 protein sequence

MRIGDPMNKLIRRAVTIFAVTSVASLFASGVLETSMAESLSTNVISLADTKAKDNTSHKSKKARKNHSKETPVDRKE
VAPVHESKATGPK@DSCFGRMYTVKVNDDRNVEIT@AVPEYATVGSPYPIEITATGKRDCVDVIITQ@LPCEAEFVR
SDPATTPTADGKLVWKIDRLG@GEKSKITVWVKPLKEGCCFTAATVCACPEIRSVTKCG@PAICVK@EGPENACLRC
PVVYKINIVN@GTATARNVVVENPVPDGYAHSSG@RVLTFTLGDM@PGEHRTITVEFCPLKRGRATNIATVSYCGGH
KNTASVTTVINEPCV@VSIAGADWSYVCKPVEYVISVSNPGDLVLRDVVVEDTLSPGVTVLEAAGA@ISCNKVVWTV
KELNPGESL@YKVLVRA@TPG@FTNNVVVKSCSDCGTCTSCAEATTYWKGVAATHMCVVDTCDPVCVGENTVYRICV
TNRGSAEDTNVSLMLKFSKEL@PVSFSGPTKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAILSSDT
LTVPVSDTENTHIY

SEQ ID NO:11 - CT372 nucleotide sequence

ATGCAGGCTGCACACCATCACTATCACCGCTACACAGATAAACTGCACAGACAAAACCATAAAAAAGATCTCATCTC
TCCCAAACCTACCGAACAAGAGGCGTGCAATACTTCTTCCCTTAGTAAGGAATTAATCCCTCTATCAGAACAAAGAG
GCCTTTTATCCCCCATCTGTGACTTTATTTCGGAACGCCCTTGCTTACACGGAGTTTCTGTTAGAAATCTCAAGCAA
GCGCTAAAAAATTCTGCAGGAACCCAAATTGCACTGGATTGGTCTATTCTCCCTCAATGGTTCAATCCTCGGGTCTC
TCATGCCCCTAAGCTTTCTATCCGAGACTTTGGGTATAGCGCACACCAAACTGTTACCGAAGCCACTCCTCCTTGCT
GGCAAAACTGCTTTAATCCATCTGCGGCCGTTACTATCTATGATTCCTCATATGGGAAAGGGGTCTTTCAAATATCC
TATACCCTTGTCCGCTATTGGAGAGAGAATGCTGCGACTGCTGGCGATGCTATGATGCTCGCAGGGAGTATCAATGA
TTATCCCTCTCGTCAGAACATTTTCTCTCAGTTTACTTTCTCCCAAAACTTCCCAAATGAACGGGTGAGTCTGACAA
TTGGTCAGTACTCACTCTATGCAATAGACGGAACATTATACAATAACGATCAACAACTTGGATTCATTAGTTACGCA
TTATCACAAAATCCAACAGCAACTTATTCCTCTGGAAGTCTTGGAGCTTACCTACAAGTCGCTCCTACCGCAAGCAC
AAGTCTTCAAATAGGATTTCAAGACGCTTATAATATCTCCGGATCCTCTATCAAATGGAGTAACCTTACAAAAAATA
GATACAATTTTCACGGTTTTGCTTCCTGGGCTCCCCGCTGTTGCTTAGGATCTGGCCAGTACTCCGTGCTTCTTTAT
GTGACTAGACAAGTTCCAGAACAGATGGAACAAACAATGGGATGGTCAGTCAATGCGAGTCAACACATATCTTCTAA
ACTGTATGTGTTTGGAAGATACAGCGGTGTTACAGGACATGTGTTCCCGATTAACCGCACGTATTCATTCGGTATGG
CCTCTGCAAATTTATTTAACCGTAACCCACAAGATTTATTTGGAATTGCTTGCGCATTCAATAATGTACACCTCTCT
GCTTCTCCAAATACTAAAAGAAAATACGAAACTGTAATCGAAGGGTTTGCAACTATCGGTTGCGGCCCCTATCTTTC
TTTCGCTCCAGACTTCCAACTCTACCTCTACCCAGCTCTTCGTCCAAACAAACAATCTGCCCGTGTTTATAGCGTGC
GAGCTAATTTAGCTATCTAA

SEQ ID NO:12 - CT372 protein sequence

M@AAHHHYHRYTDKLHR@NHKKDLISPKPTE@EACNTSSLSKELIPLSE@RGLLSPICDFISERPCLHGVSVRNLK@
ALKNSAGT@IALDWSILP@WFNPRVSHAPKLSIRDFGYSAH@TVTEATPPCW@NCFNPSAAVTIYDSSYGKGVF@IS
YTLVRYWRENAATAGDAMMLAGSINDYPSR@NIFS@FTFS@NFPNERVSLTIG@YSLYAIDGTLYNND@@LGFISYA
LS@NPTATYSSGSLGAYL@VAPTASTSL@IGF@DAYNISGSSIKWSNLTKNRYNFHGFASWAPRCCLGSG@YSVLLY
VTR@VPE@ME@TMGWSVNAS@HISSKLYVFGRYSGVTGHVFPINRTYSFGMASANLFNRNP@DLFGIACAFNNVHLS
ASPNTKRKYETVIEGFATIGCGPYLSFAPDF@LYLYPALRPNK@SARVYSVRANLAI

SEQ ID NO:13 - CT456 nucleotide sequence

ATGACGAATTCTATATCAGGTTATCAACCTACTGTTACAACTTCTACATCATCAACCACTTCGGCATCAGGTGCTTC
CGGATCTCTGGGAGCTTCTTCTGTATCTACTACCGCAAACGCTACAGTTACACAAACAGCAAACGCAACAAATTCAG
CGGCTACATCTTCTATCCAAACGACTGGAGAGACTGTAGTAAACTATACGAATTCAGCCTCCGCCCCCAATGTAACT
GTATCGACCTCCTCTTCTTCCACACAAGCCACAGCCACTTCGAATAAAACTTCCCAAGCCGTTGCTGGAAAAATCAC
TTCTCCAGATACTTCAGAAAGCTCAGAAACTAGCTCTACCTCATCAAGCGATCATATCCCTAGCGATTACGATGACG
TTGGTAGCAATAGTGGAGATATTAGCAACAACTACGATGACGTAGGTAGTAACAACGGAGATATCAGTAGCAATTAT
GACGATGCTGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCTCTAGAACAAG
TGGCCCAGAAAATACAAGTGGTGGTGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACAGCAATTATGACG
ATGCTGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCTCTAGAACAAGTGGC
CCAGAAAATACGAGTGGTGGTGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACAGCAATTATGACGATGC
TGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCTCTAGAACAAGTGGCCCAG
AAAATACGAGTGATGGTGCAGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACACAACAGGGCCTCGTAAC
GAGGGTGTATTCGGCCCTGGACCGGAAGGACTACCAGACATGTCTCTTCCTTCATACGATCCTACAAATAAAACCTC
GTTATTGACTTTCCTCTCCAACCCTCATGTAAAGTCGAAAATGCTTGAAAACTCGGGGCATTTCGTCTTCATTGATA
CAGATAGAAGTAGTTTCATTCTTGTTCCTAACGGAAATTGGGACCAAGTCTGTTCAATTAAAGTTCAAAATGGAAAG
ACCAAAGAAGATCTCGACATCAAAGACTTGGAAAACATGTGTGCAAAATTCTGTACAGGGTTTAGCAAATTCTCTGG
TGACTGGGACAGTCTTGTAGAACCTATGGTGTCAGCCAAAGCTGGAGTGGCCAGCGGAGGCAATCTTCCCAATACAG
TGATTATCAATAATAAATTCAAAACTTGCGTTGCTTATGGTCCTTGGAATAGCCAGGAAGCAAGTTCTGGTTATACA
CCTTCTGCTTGGAGACGTGGTCATCGAGTAGATTTTGGAGGAATTTTTGAGAAAGCCAACGACTTTAATAAAATCAA
CTGGGGAACTCAAGCCGGGCCTAGTAGCGAAGACGATGGCATTTCCTTCTCCAATGAAACTCCTGGAGCTGGTCCTG

CAGCTGCTCCATCACCAACGCCATCCTCTATTCCTATCATCAATGTCAATGTCAATGTTGGCGGAACTAATGTGAAT
ATTGGAGATACGAATGTCAACACGACTAACACCACACCAACAACTCAATCTACAGACGCCTCTACAGATACAAGCGA
TATCGATGACATAAATACCAACAACCAAACTGATGATATCAATACGACAGACAAAGACTCTGACGGAGCTGGTGGAG
TCAATGGCGATATATCCGAAACAGAATCCTCTTCTGGAGATGATTCAGGAAGTGTCTCTTCCTCAGAATCAGACAAG
AATGCCTCTGTCGGAAATGACGGACCTGCTATGAAAGATATCCTTTCTGCCGTGCGTAAACACCTAGACGTCGTTTA
CCCTGGCGAAAATGGCGGTTCTACAGAAGGGCCTCTCCCAGCTAACCAAACTCTCGGAGACGTAATCTCTGATGTAG
AGAATAAAGGCTCCGCTCAGGATACAAAATTGTCAGGAAATACAGGAGCTGGGGATGACGATCCAACAACCACAGCT
GCTGTAGGTAATGGAGCGGAAGAGATCACTCTTTCCGACACAGATTCTGGTATCGGAGATGATGTATCCGATACAGC
GTCTTCATCTGGGGATGAATCCGGAGGAGTCTCCTCTCCCTCTTCAGAATCCAATAAAAATACTGCCGTTGGAAATG
ACGGACCTTCTGGACTAGATATCCTCGCTGCCGTACGTAAACATTTAGATAAGGTTTACCCTGGCGACAATGGTGGT
TCTACAGAAGGGCCTCTCCAAGCTAACCAAACTCTTGGAGATATCGTCCAGGATATGGAAACAACAGGGACATCCCA
AGAAACCGTTGTATCCCCATGGAAAGGAAGCACTTCTTCAACGGAATCAGCAGGAGGAAGTGGTAGCGTACAAACAC
TACTGCCTTCACCACCTCCAACCCCGTCAACTACAACATTAAGAACGGGCACAGGAGCTACCACCACATCCTTGATG
ATGGGAGGACCAATCAAAGCTGACATAATAACAACTGGTGGCGGAGGACGAATTCCTGGAGGAGGAACGTTAGAAAA
GCTGCTCCCTCGTATACGTGCGCACTTAGACATATCCTTTGATGCGCAAGGCGATCTCGTAAGTACTGAAGAGCCTC
AGCTTGGCTCGATTGTAAACAAATTCCGCCAAGAAACTGGTTCAAGAGGAATCTTAGCTTTCGTTGAGAGTGCTCCA
GGCAAGCCGGGATCTGCACAGGTCTTAACGGGTACAGGGGGAGATAAAGGCAACCTATTCCAAGCAGCTGCCGCAGT
CACCCAAGCCTTAGGAAATGTTGCAGGGAAAGTCAACCTTGCGATACAAGGCCAAAAACTATCATCCCTAGTCAATG
ACGACGGGAAGGGGTCTGTTGGAAGAGATTTATTCCAAGCAGCAGCCCAAACAACTCAAGTGCTAAGCGCACTGATT
GATACCGTAGGATAA

SEQ ID NO:14 - CT456 protein sequence

MTNSISGYQPTVTTSTSSTTSASGASGSLGASSVSTTANATVTQTANATNSAATSSIQTTGETVVNYTNSASAPNVT
VSTSSSSTQATATSNKTSQAVAGKITSPDTSESSETSSTSSSDHIPSDYDDVGSNSGDISSNYDDVGSNNGDISSNY
DDAAADYEPIRTTENIYESIGGSRTSGPENTSGGAAAALNSLRGSSYSNYDDAAADYEPIRTTENIYESIGGSRTSG
PENTSGGAAAALNSLRGSSYSNYDDAAADYEPIRTTENIYESIGGSRTSGPENTSDGAAAAALNSLRGSSYTTGPRN
EGVFGPGPEGLPDMSLPSYDPTNKTSLLTFLSNPHVKSKMLENSGHFVFIDTDRSSFILVPNGNWDQVCSIKVQNGK
TKEDLDIKDLENMCAKFCTGFSKFSGDWDSLVEPMVSAKAGVASGGNLPNTVIINNKFKTCVAYGPWNSQEASSGYT
PSAWRRGHRVDFGGIFEKANDFNKINWGTQAGPSSEDDGISFSNETPGAGPAAAPSPTPSSIPIINVNVNVGGTNVN
IGDTNVNTTNTTPTTQSTDASTDTSDIDDINTNNQTDDINTTDKDSDGAGGVNGDISETESSSGDDSGSVSSSESDK
NASVGNDGPAMKDILSAVRKHLDVVYPGENGGSTEGPLPANQTLGDVISDVENKGSAQDTKLSGNTGAGDDDPTTTA
AVGNGAEEITLSDTDSGIGDDVSDTASSSGDESGGVSSPSSESNKNTAVGNDGPSGLDILAAVRKHLDKVYPGDNGG
STEGPLQANQTLGDIVQDMETTGTSQETVVSPWKGSTSSTESAGGSGSVQTLLPSPPPTPSTTTLRTGTGATTTSLM
MGGPIKADIITTGGGGRIPGGGTLEKLLPRIRAHLDISFDAQGDLVSTEEPQLGSIVNKFRQETGSRGILAFVESAP
GKPGSAQVLTGTGGDKGNLFQAAAAVTQALGNVAGKVNLAIQGGKLSSLVNDDGKGSVGRDLFQAAAQTTQVLSALI
DTVG

SEQ ID NO:15: CT381 nucleotide sequence

ATGTGCATAAAAAGAAAAAAAACATGGATAGCTTTTTTAGCAGTTGTCTGTAGTTTTTGTTTGACGGGTTGTTTAAA
AGAAGGGGGAGACTCCAATAGTGAAAAATTTATTGTAGGGACTAATGCAACCTACCCTCCTTTTGAGTTTGTTGATA
AGCGAGGAGAGGTTGTAGGCTTCGATATAGACTTGGCTAGAGAGATTAGTAACAAGCTGGGGAAAACGCTGGACGTT
CGGGAGTTTTCCTTTGATGCACTCATTCTAAACCTAAAACAGCATCGGATTGATGCGGTTATAACAGGGATGTCCAT
TACTCCTTCTAGATTGAAGGAAATTCTTATGATTCCCTATTATGGGGAGGAAATAAAACACTTGGTTTTAGTGTTTA
AAGGAGAGAATAAGCATCCATTGCCACTCACTCAATATCGTTCTGTAGCTGTTCAAACAGGAACCTATCAAGAGGCC
TATTTACAGTCTCTTTCTGAAGTTCATATTCGCTCTTTTGATAGCACTCTAGAAGTACTCATGGAAGTCATGCATGG
TAAATCTCCCGTCGCTGTTTTAGAGCCATCTATCGCTCAAGTTGTCTTGAAAGATTTCCCGGCTCTTTCTACAGCAA
CCATAGATCTCCCTGAAGATCAGTGGGTTTTAGGATACGGGATTGGCGTTGCTTCAGATCGCCCAGCTTTAGCCTTG
AAAATCGAGGCAGCTGTGCAAGAGATCCGAAAAGAAGGAGTGCTAGCAGAGTTGGAACAGAAGTGGGGTTTGAACAA
CTAA

SEQ ID NO:16: CT381 protein sequence

MCIKRKKTWIAFLAVVCSFCLTGCLKEGGDSNSEKFIVGTNATYPPFEFVDKRGEVVGFDIDLAREISNKLGKTLDV
REFSFDALILNLKQHRIDAVITGMSITPSRLKEILMIPYYGEEIKHLVLVFKGENKHPLPLTQYRSVAVQTGTYQEA
YLQSLSEVHIRSFDSTLEVLMEVMHGKSPVAVLEPSIAQVVLKDFPALSTATIDLPEDQWVLGYGIGVASDRPALAL
KIEAAVQEIRKEGVLAELEQKWGLNN

SEQ ID NO:17: CT043 nucleotide sequence

ATGTCCAGGCAGAATGCTGAGGAAAATCTAAAAAATTTTGCTAAAGAGCTTAAACTCCCCGACGTGGCCTTCGATCA
GAATAATACGTGCATTTTGTTTGTTGATGGAGAGTTTTCTCTTCACCTGACCTACGAAGAACACTCTGATCGCCTTT
ATGTTTACGCACCTCTTCTTGACGGACTGCCAGACAATCCGCAAAGAAGGTTAGCTCTATATGAGAAGTTGTTAGAA
GGCTCTATGCTCGGAGGCCAAATGGCTGGTGGAGGGGTAGGAGTCGCTACTAAGGAACAGTTGATCTTAATGCACTG
CGTGTTAGACATGAAGTATGCAGAGACCAACCTACTCAAAGCTTTTGCACAGCTTTTTATTGAAACCGTTGTGAAAT
GGCGAACTGTTTGTTCTGATATCAGCGCTGGACGAGAACCCACTGTTGATACCATGCCACAAATGCCTCAAGGGGGT
GGCGGAGGAATTCAACCTCCTCCAGCAGGAATCCGTGCATAA

SEQ ID NO:18: CT043 protein sequence

MSR@NAEENLKNFAKELKLPDVAFD@NNTCILFVDGEFSLHLTYEEHSDRLYVYAPLLDGLPDNP@RRLALYEKLLE
GSMLGG@MAGGGVGVATKE@LILMHCVLDMKYAETNLLKAFA@LFIETVVKWRTVCSDISAGREPTVDTMP@MP@GG
GGGI@PPPAGIRA

SEQ ID NO:19: CT711 / hypothetical protein (AAC68306)

MSI@PTSISLTKNITAALAGE@VDAAAVYMP@AVFFF@@LDEKSKGLK@ALGLLEEVDLEKFIPSLEKSPTPITTGT
TSKISADGIEIVGELSSETILADPNKAAA@VFGEGLADSFDDWLRLSENGGI@DPTAIEEEEIVTKY@TELNTLRNKL
K@@SLTDDEYTKLYAIP@NFVKEIESLKNENNVRLIPKSKVTNFW@NIMLTYNSVTSLSEPVTDAMNTTMAEYSLYI
ERATEAAKLIREITNTIKDIFNPVWDVRE@TGIFGLKGAEYNALEGNMI@SLLSFAGLFR@LMSRTATVDEIGALYP
KNDKNEDVIHTAIDDYVNSLADLKANE@VKLNGLLSLVYAYYASTLGFAKKDVFNNA@ASFTDYTNFLN@EI@YWTP
RETSSFNISN@AL@TFKNKPSADYNGVYLFDNKGLETNLFNPTFFFDVVSLMTADPTKTMSR@DYNKVITASESSI@
KIN@AITAWELAIAECGTKKAKLEPSSLNYFNAMVEAKKTFVETSPI@MVYSSLMLDKYLPN@@YILETLGS@MTFS
NKAARYLNDIIAYAVSF@TADVYYSLGMYLR@MN@@EFPEVISRANDTVKKEIDRSRADLFHCKKAIEKIKELVTSV
NADTELTSS@RAELLETLASYAFEFENLYHNLSNVYVMVSKV@ISGVSKPDEVDEAFTAKIGSKEFDTWI@@LTTFE
SAVIEGGRNGVMPGGE@@VL@SLESK@@DYTSFN@N@@LAL@MESAAI@@EWTMVAAALALMN@IFAKLIRRFK

SEQ ID NO:20: CT114 / hypothetical protein (AAC67705)

MCFIGIGSLLLPTALRATERMRKEPIPLLDK@@SFWNVDPYCLESICACFVAHRDPLSAK@LMYLFP@LSEEDVSVF
ARCILSSKRPEYLFSKSEEELFAKLILPRVSLGVHRDDDLARVLVLAEPSAEE@KARYYSLYLDVLALRAYVERERL
ASAAHGDPERIDLATIEAINTILF@EEGWRYPSK@EMFENRFSELAAVTDSKFGVCLGTVVLY@AVA@RLDLSLDPV
TPPGHIYLRYKDKVNIETTSGGRHLPTERYCECIKES@LKVRS@MELIGLTFMNRGAFFL@KGEFL@ASLAYE@A@S
YLSDE@ISDLLGITYVLLGKKAAGEALLKKSAEKTRRGSSIYDYF@GYISPEILGVLFADSGVTY@ETLEYRKKLVM
LSKKYPKSGSLRLRLATTALELGLVKEGV@LLEESVKDAPEDLSLRL@FCKILCNRHDYVRAKYHFD@A@ALLIKEG
LFSEKTSYTLLKTIGKKLSLFAPS

SEQ ID NO:21: CT480 / oppA_4 (AAC68080)

MIDKIIRTILVLSLFLLYWSSDLLEKDVKSIKRELKALHEDVLELVRISH@@KNWV@STDFSVSPEISVLKDCGDPA
FPNLLCEDPYVEKVVPSLLKEGFVPKGILRTA@VGRPDNLSPFNGFVNIVRFYELCVPNLAVEHVGKYEEFAPSLAL
KIEEHYVEDGSGDKEFHIYLRPNMFWEPIDPTLFPKNITLADSFLRPHPVTAHDVKFYYDVVMNPYVAEMRAVAMRS
YFEDMVSVRVENDLKLIVRWRAHTVRNE@GEEEKKVLYSAFANTLAL@PLPCFVY@HFANGEKIVPEDSDPDTYRKD
SVWA@NFSSHWAYNYIVSCGAFRFAGMDDEKITLVRNPNYHNPFAALVEKRYIYMKDSTDSLF@DFKAGKVDIAYFP
PNHVDNLASFM@TSAYKE@AARGEAILEKNSSDRSYSYIGWNCLSLFFNNRSVR@AMNMLIDRDRIIE@CLDGRGVS
VSGPFSLCSPSYNRDVEGW@YSPEEAARKLEEEGWIDADGDGIREKVIDGVVVPFRFRLCYYVKSVTARTIAEYVAT
VCKEVGIECCLLGLDMADYS@ALEEKNFDAILSGWCLGTPPEDPRALWHSEGALEKGSANAVGFCNEEADRIIE@LS
YEYDSNKR@ALYHRFHEVIHEESPYAFLYSR@YSLVYKEFVKNIFVPTEH@DLIPGA@DETVNLSMLWVDKEEGRCS
AIS

SEQ ID NO:22: CT089 / lcrE (AAC67680)

MTASGGAGGLGST@TVDVARA@AAAAT@DA@EVIGS@EASEASMLKGCEDLINPAAATRIKKKGEKFESLEARRKPT
ADKAEKKSESTEEKGDTPLEDRFTEDLSEVSGEDFRGLKNSFDDDSSPDEILDALTSKFSDPTIKDLALDYLI@TAP
SDGKLKSTLI@AKH@LMS@NP@AIVGGRNVLLASETFASRANTSPSSLRSLYF@VTSSPSNCANLH@MLASYLPSEK
TAVMEFLVNGMVADLKSEGPSIPPAKL@VYMTELSNL@ALHSVNSFFDRNIGNLENSLKHEGHAPIPSLTTGNLTKT
FL@LVEDKFPSSSKA@KALNELVGPDTGP@TEVLNLFFRALNGCSPRIFSGAEKK@@LASVITNTLDAINADNEDYP
KPGDFPRSSFSSTPPHAPVP@SEIPTSPTST@PPSP

SEQ ID NO:23: CT734 / hypothetical protein (AAC68329)

MKKFIYKYSFGALLLLSGLSGLSSCCANSYGSTLAKNTAEIKEESVTLREKPDAGCKKKSSCYLRKFFSRKKPKEKT
EPVLPNFKSYADPMTDSERKDLSFVVSAAADKSSIALAMA@GEIKGALSRIREIHPLALL@ALAEDPALIAGMKKM@
GRDWVWNIFITELSKVFS@AASLGAFSVADVAAFASTLGLDSGTVTSIVDGERWAELIDVVI@NPAI

SEQ ID NO:24: CT016 / hypothetical protein (AAC67606)

MKVKIND@FICISPYISARWN@IAFIESCDGGTEGGITLKLHLIDGETVSIPNLG@AIVDEVF@EHLLYLESTAP@K
NKEEEKISSLLGAV@@MAKGCEV@VFS@KGLVSMLLGGAGSINVLL@HSPEHKDHPDLPTDLLERIA@MMRSLSIGP
TSILAKPEPHCNCLHC@IGRATVEEEDAGVSDEDLTFRSWDIS@SGEKMYTVTDPLNPEE@FNVYLGTPIGCTCG@P
YCEHVKAVLYT

SEQ ID NO:25: CM homolog of CT279 = TC_0551

ATGGCATCCAAGTCTCGTCATTATCTTAACCAGCCTTGGTACATTATCTTATTCATCTTTGTTCTTAGTCTGGTTGC
TGGTACCCTTCTTTCTTCAGTTTCCTATGTTCTATCTCCAATCCAAAAACAAGCTGCAGAATTTGATCGTAATCAGC
AAATGTTGATGGCCGCACAAATTATTTCCTATGACAATAAATTCCAAATATATGCTGAAGGGGATTGGCAACCTGCT
GTCTATAATACAAAAAAACAGATACTAGAAAAAAGCTCTTCCACTCCACCACAAGTGACTGTGGCGACTCTATGCTC
TTATTTTCAAAATTTTGTTAGAGTTTTGCTTACAGACTCCCAAGGGAATCTTTCTTCTTTTGAAGATCACAATCTTA
ACCTAGAAGAGTTCTTATCCCACCCCACATCTTCAGTACAAGATCACTCTCTGCATGTAATTTATGCTATTCTAGCA
AACGATGAATCCTCTAAAAAGTTATCATCCTCCCAAGTAGCAAAAAATCCGGTATCCATAGAGTCTATTATTCTTCC

TATAAAAGGATTTGGTTTATGGGGACCAATCTATGGATTTCTTGCTTTAGAAAAGGACGGTAATACGGTTCTAGGGA
CATGCTGGTATCAACATGGTGAGACTCCAGGATTAGGAGCAAATATAACTAATCCCCAATGGCAACAAAATTTCAGA
GGAAAAAAAGTATTTCTCGCTTCCTCTTCCGGAGAAACCGATTTTGCTAAAACAACTCTAGGACTAGAAGTTATAAA
AGGATCTGTTTCTGCATTATTAGGGGACTCTCCCAAAGCTAATTCCGCTGTTGATGGAATTTCAGGAGCTACACTGA
CCTGTAATGGAGTTACTGAAGCTTTTGCTAATTCGCTAGCTCCTTACCGCCCCTTATTGACTTTCTTCGCCAATCTT
AACTCTAGTGGAGAATCTCATGACAACCAATAA

SEQ ID NO:26: CM homologue of CT279 protein sequence = TC_0551 protein sequence

MASKSRHYLNQPWYIILFIFVLSLVAGTLLSSVSYVLSPIQKQAAEFDRNQQMLMAAQIISYDNKFQIYAEGDWQPA
VYNTKKQILEKSSSTPPQVTVATLCSYFQNFVRVLLTDSQGNLSSFEDHNLNLEEFLSHPTSSVQDHSLHVIYAILA
NDESSKKLSSSQVAKNPVSIESIILPIKGFGLWGPIYGFLALEKDGNTVLGTCWYQHGETPGLGANITNPQWQQNFR
GKKVFLASSSGETDFAKTTLGLEVIKGSVSALLGDSPKANSAVDGISGATLTCNGVTEAFANSLAPYRPLLTFFANL
NSSGESHDNQ

SEQ ID NO:27: CM homologue of CT372 = TC_0651 nucleotide sequence

ATGAATGGAAAAGTTCTGTGTGAGGTTTCTGTGTCCTTCCGTTCGATTCTGCTGACGGCTCTGCTTTCACTTTCTTT
TACAAACACTATGCAGGCTGCACACCATCATTATCACCGTTATGATGATAAACTACGCAGACAATACCATAAAAAGG
ACTTGCCCACTCAAGAGAATGTTCGGAAAGAGTTTTGTAATCCCTACTCTCATAGTAGTGATCCTATCCCTTTGTCA
CAACAACGAGGAGTCCTATCTCCTATCTGTGATTTAGTCTCAGAGTGCTCGTTTTTGAACGGGATTTCCGTTAGGAG
TCTTAAACAAACACTGAAAAATTCTGCTGGGACTCAAGTTGCTTTAGACTGGTCTATCCTTCCTCAATGGTTCAATC
CTAGATCCTCTTGGGCTCCTAAGCTCTCTATTCGAGATCTTGGATATGGTAAACCCCAGTCCCTTATTGAAGCAGAT
TCCCCTTGTTGTCAAACCTGCTTCAACCCATCTGCTGCTATTACGATTTACGATTCTTCATGTGGGAAGGGTGTTGT
CCAAGTGTCATACACCCTTGTTCGTTATTGGAGAGAAACGGCTGCACTTGCAGGGCAAACTATGATGCTTGCAGGAA
GTATTAATGATTATCCTGCTCGCCAAAACATATTCTCTCAACTTACATTTTCCCAAACTTTCCCTAATGAGAGAGTA
AATCTAACTGTTGGTCAATACTCTCTTTACTCGATAGACGGAACGCTGTACAACAATGATCAGCAGCTAGGATTTAT
TAGTTATGCGTTGTCGCAAAATCCAACAGCGACTTATTCCTCTGGAAGCCTTGGCGCCTATCTACAAGTCGCTCCAA
CAGAAAGCACCTGTCTTCAAGTTGGGTTCCAAGATGCCTATAATATTTCAGGTTCCTCGATCAAATGGAATAATCTT
ACAAAAAATAAGTATAACTTCCATGGCTATGCATCTTGGGCTCCACACTGTTGCTTAGGACCTGGACAATACTCTGT
TCTTCTTTATGTAACCAGAAAGGTTCCTGAGCAAATGATGCAGACAATGGGCTGGTCTGTGAATGCAAGTCAATACA
TCTCTTCTAAACTTTATGTATTTGGAAGATACAGCGGAGTCACAGGCCAATTGTCTCCTATTAACCGAACCTATTCA
TTTGGCTTAGTCTCTCCTAATTTATTGAACCGTAACCCACAAGACTTATTTGGAGTAGCTTGCGCATTCAATAATAT
ACACGCCTCCGCCTTTCAAAATGCTCAAAGAAAATATGAAACTGTGATCGAGGGATTTGCAACTATTGGTTGCGGAC
CTTACATCTCCTTTGCTCCAGATTTCCAACTTTACCTCTATCCTGCTCTGCGTCCAAATAAACAAAGCGCCCGAGTC
TATAGCGTTCGCGCAAACCTAGCTATTTAG

SEQ ID NO:28: CM homologue of CT372 = TC_0651 protein sequence

MNGKVLCEVSVSFRSILLTALLSLSFTNTMQAAHHHYHRYDDKLRRQYHKKDLPTQENVRKEFCNPYSHSSDPIPLS
QQRGVLSPICDLVSECSFLNGISVRSLKQTLKNSAGTQVALDWSILPQWFNPRSSWAPKLSIRDLGYGKPQSLIEAD
SPCCQTCFNPSAAITIYDSSCGKGVVQVSYTLVRYWRETAALAGQTMMLAGSINDYPARQNIFSQLTFSQTFPNERV
NLTVGQYSLYSIDGTLYNNDQQLGFISYALSQNPTATYSSGSLGAYLQVAPTESTCLQVGFQDAYNISGSSIKWNNL
TKNKYNFHGYASWAPHCCLGPGQYSVLLYVTRKVPEQMMQTMGWSVNASQYISSKLYVFGRYSGVTGQLSPINRTYS
FGLVSPNLLNRNPQDLFGVACAFNNIHASAFQNAQRKYETVIEGFATIGCGPYISFAPDFQLYLYPALRPNKQSARV
YSVRANLAI

SEQ ID NO:29: CM homologue of CT443 = TC_0727

ATGCGAATAGGAGATCCTATGAACAAACTCATCAGACGAGCTGTGACGATCTTCGCGGTGACTAGTGTGGCGAGTTT
ATTTGCTAGCGGGGTGTTAGAGACCTCTATGGCAGAGTCTCTCTCTACCAACGTTATTAGCTTAGCTGACACCAAAG
CGAAAGAGACCACTTCTCATCAAAAAGACAGAAAAGCAAGAAAAAATCATCAAAATAGGACTTCCGTAGTCCGTAAA
GAGGTTACTGCAGTTCGTGATACTAAAGCTGTAGAGCCTAGACAGGATTCTTGCTTTGGCAAAATGTATACAGTCAA
AGTTAATGATGATCGTAATGTAGAAATCGTGCAGTCCGTTCCTGAATATGCTACGGTAGGATCTCCATATCCTATTG
AGATTACTGCTATAGGGAAAAGAGACTGTGTTGATGTAATCATTACACAGCAATTACCATGCGAAGCAGAGTTTGTT
AGCAGTGATCCAGCTACTACTCCTACTGCTGATGGTAAGCTAGTTTGGAAAATTGATCGGTTAGGACAGGGCGAAAA
GAGTAAAATTACTGTATGGGTAAAACCTCTTAAAGAAGGTTGCTGCTTTACAGCTGCAACGGTTTGTGCTTGTCCAG
AGATCCGTTCGGTTACGAAATGTGGCCAGCCTGCTATCTGTGTTAAACAGGAAGGTCCAGAAAGCGCATGTTTGCGT
TGCCCAGTAACTTATAGAATTAATGTAGTCAACCAAGGAACAGCAACAGCACGTAATGTTGTTGTGGAAAATCCTGT
TCCAGATGGCTATGCTCATGCATCCGGACAGCGTGTATTGACATATACTCTTGGGGATATGCAACCTGGAGAACAGA
GAACAATCACCGTGGAGTTTTGTCCGCTTAAACGTGGTCGAGTCACAAATATTGCTACAGTTTCTTACTGTGGTGGA
CACAAAAATACTGCTAGCGTAACAACAGTGATCAATGAGCCTTGCGTGCAAGTTAACATCGAGGGAGCAGATTGGTC
TTATGTTTGTAAGCCTGTAGAATATGTTATCTCTGTTTCTAACCCTGGTGACTTAGTTTTACGAGACGTTGTAATTG
AAGATACGCTTTCTCCTGGAATAACTGTTGTTGAAGCAGCTGGAGCTCAGATTTCTTGTAATAAATTGGTTTGGACT
TTGAAGGAACTCAATCCTGGAGAGTCTTTACAATATAAGGTTCTAGTAAGAGCTCAAACTCCAGGGCAATTCACAAA
CAACGTTGTTGTGAAAAGTTGCTCTGATTGCGGTATTTGTACTTCTTGCGCAGAAGCAACAACTTACTGGAAAGGAG
TTGCTGCTACTCATATGTGCGTAGTAGATACTTGTGATCCTATTTGCGTAGGAGAGAACACTGTTTATCGTATCTGT
GTGACAAACAGAGGTTCTGCTGAAGATACAAATGTGTCCTTAATTTTGAAATTCTCTAAAGAATTACAACCTATATC
TTTCTCTGGACCAACTAAAGGAACCATTACAGGAAACACGGTAGTGTTTGATTCGTTACCTAGATTAGGTTCTAAAG
AAACTGTAGAGTTTTCTGTAACGTTGAAAGCAGTATCCGCTGGAGATGCTCGTGGGGAAGCTATTCTTTCTTCCGAT
ACATTGACAGTTCCTGTATCTGATACGGAGAATACACATATCTATTAA

SEQ ID NO:30: CM homologue of CT443 = TC_0727

MRIGDPMNKLIRRAVTIFAVTSVASLFASGVLETSMAESLSTNVISLADTKAKETTSHQKDRKARKNHQNRTSVVRK
EVTAVRDTKAVEPRQDSCFGKMYTVKVNDDRNVEIVQSVPEYATVGSPYPIEITAIGKRDCVDVIITQQLPCEAEFV
SSDPATTPTADGKLVWKIDRLGQGEKSKITVWVKPLKEGCCFTAATVCACPEIRSVTKCGQPAICVKQEGPESACLR
CPVTYRINVVNQGTATARNVVVENPVPDGYAHASGQRVLTYTLGDMQPGEQRTITVEFCPLKRGRVTNIATVSYCGG
HKNTASVTTVINEPCVQVNIEGADWSYVCKPVEYVISVSNPGDLVLRDVVIEDTLSPGITVVEAAGAQISCNKLVWT
LKELNPGESLQYKVLVRAQTPGQFTNNVVVKSCSDCGICTSCAEATTYWKGVAATHMCVVDTCDPICVGENTVYRIC
VTNRGSAEDTNVSLILKFSKELQPISFSGPTKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAILSSD
TLTVPVSDTENTHIY

SEQ ID NO:31: CM homologue of CT043 = TC_0313 nucleotide sequence

ATGTCCAGACAGAATGCTGAGGAAAATCTAAAAAATTTTGCTAAAGAGCTCAAGCTCCCCGACGTGGCCTTCGATCA
GAATAATACGTGCATTTTGTTTGTTGATGGAGAGTTTTCTCTTCACCTGACCTACGAAGAGCACTCTGATCGCCTTT
ATGTTTACGCACCTCTCCTTGACGGACTCCCAGATAATCCGCAAAGAAAGTTGGCTCTGTATGAGAAATTGTTGGAA
GGCTCTATGCTCGGAGGCCAAATGGCTGGTGGAGGAGTAGGAGTTGCTACTAAAGAACAGTTGATCCTAATGCATTG
CGTGTTAGATATGAAATATGCAGAGACTAATCTATTGAAAGCTTTTGCACAGCTTTTCATTGAAACTGTTGTGAAAT
GGCGAACGGTCTGTTCTGATATCAGCGCTGGACGAGAACCTTCCGTTGACACTATGCCTCAAATGCCTCAAGGAGGC
AGCGGAGGAATTCAACCTCCTCCAACAGGAATTCGTGCGTAG

SEQ ID NO:32: CM homologue of CT043 = TC_0313 protein sequence

MSRQNAEENLKNFAKELKLPDVAFDQNNTCILFVDGEFSLHLTYEEHSDRLYVYAPLLDGLPDNPQRKLALYEKLLE
GSMLGGQMAGGGVGVATKEQLILMHCVLDMKYAETNLLKAFAQLFIETVVKWRTVCSDISAGREPSVDTMPQMPQGG
SGGIQPPPTGIRA

SEQ ID NO:33: CM homologue of CT601 = TC_0890 nucleotide sequence

ATGCTCGCTAATCGGTTATTTCTAATCACCCTTATAGGTTTTGGCTATTCTGCTTACGGTGCCAGCACAGGGAAATC
ACCTTCTTTACAGGTTATTTTAGCTGAAGTCGAGGATACATCTTCGCGCTTACAAGCTCATCAGAATGAGCTTGTTA
TGCTCTCGGAACGTTTAGATGAGCAAGACACAAAACTTCAACAACTCTCGTCAACTCAGGCCCGTAATCTTCCTCAA
CAAGTTCAACGGCTTGAGATTGATCTGAGAGCTCTGGCTAAAACAGCTGCTGTGCTCTCGCAATCTGTTCAGGATAT
CCGATCATCCGTGCAAAATAAATTACAAGAAATCCAACAAGAACAAAAAAATTTAGCTCAAAATTTACGAGCGCTTC
GCAACTCCTTACAAGCACTAGTTGATGGCTCTTCCCCAGAAAATTATATTGATTTTTTGGCCGGGGAGACACCTGAA
CATATTCACGTTGTTAAACAAGGAGAAACCCTGAGTAAAATCGCTAGTAAGTACAATATCCCTGTCGCAGAATTGAA
AAAACTTAATAAATTAAATTCCGATACTATTTTTACTGATCAAAGAATCCGACTTCCAAAAAAGAAATAA

SEQ ID NO:34: CM homologue of CT601 = TC_0890 protein sequence

MLANRLFLITLIGFGYSAYGASTGKSPSLQVILAEVEDTSSRLQAHQNELVMLSERLDEQDTKLQQLSSTQARNLPQ
QVQRLEIDLRALAKTAAVLSQSVQDIRSSVQNKLQEIQQEQKNLAQNLRALRNSLQALVDGSSPENYIDFLAGETPE
HIHVVKQGETLSKIASKYNIPVAELKKLNKLNSDTIFTDQRIRLPKKK

SEQ ID NO:35: CM homologue of CT456 = TC_0741

ATGACGACTCCAATAAGTAATTCTCCATCTTCTATTCCAACTGTTACAGTATCAACTACTACAGCATCTTCTGGATC
TCTCGGAACTTCTACTGTATCATCAACGACTACAAGTACTTCAGTCGCACAAACAGCAACAACAACATCTTCTGCTT
CTACATCTATAATTCAGTCTAGTGGAGAAAACATCCAATCCACTACAGGTACCCCTTCTCCTATTACGTCTAGTGTT
TCAACATCCGCTCCATCTCCTAAAGCCTCCGCCACTGCAAACAAAACTTCAAGCGCTGTTTCTGGGAAAATTACCTC
ACAAGAAACTTCTGAGGAATCCGAAACCCAAGCCACTACATCTGATGGAGAAGTTAGTAGTAATTACGATGATGTTG
ATACCCCGACCAATTCGTCCGATTCGACAGTTGATAGTGATTACCAAGATGTTGAGACTCAGTACAAAACAATTAGC
AACAATGGTGAAAACACTTATGAAACAATCGGAAGTCATGGTGAGAAAAACACACACGTCCAGGAAAGCCATGCATC
CGGAACAGGAAATCCCATAAATAATCAGCAAGAAGCTATTAGACAGCTCCGATCATCTACCTATACAACCAGCCCTC
GTAATGAGAATATATTTAGTCCAGGACCGGAAGGTCTACCTAATATGTCTCTTCCTAGTTACAGCCCTACAGATAAA
AGTTCTCTACTAGCTTTCCTATCTAATCCCAATACAAAAGCAAAAATGCTCGAACACTCCGGGCATTTAGTCTTTAT
AGACACAACTAGAAGTAGCTTTATCTTTGTTCCGAATGGAAATTGGGATCAAGTCTGTTCCATGAAGGTTCAGAATG
GGAAAACTAAAGAAGACCTTGGCTTAAAGGACTTAGAAGATATGTGTGCAAAGTTTTGCACAGGATACAATAAATTC
TCCTCTGATTGGGGAAATCGAGTTGACCCCTTGGTCTCTTCTAAGGCCGGGATAGAAAGTGGGGGGCACCTCCCAAG
CTCAGTTATCATCAACAACAAATTTAGAACCTGTGTTGCCTATGGGCCGTGGAACCCCAAAGAAAACGGCCCCAATT
ATACTCCTTCAGCCTGGAGACGTGGGCATCGAGTAGATTTTGGAAAGATCTTTGATGGAACAGCGCCGTTTAATAAA
ATCAACTGGGGCTCTTCCCCTACCCCTGGTGATGACGGCATCTCCTTCTCTAATGAAACTATTGGGTCTGAACCATT
CGCGACACCTCCCTCATCCCCATCGCAAACCCCCGTTATCAACGTCAATGTTAATGTCGGTGGAACCAATGTTAATA
TTGGGGATACAAACGTATCTAAAGGATCCGGCACACCAACATCTTCTCAATCTGTGGACATGTCTACAGATACTAGC
GATTTAGATACCAGTGATATTGATACAAACAACCAAACTAACGGCGATATCAACACGAATGACAACTCCAATAATGT
CGATGGAAGTTTATCTGACGTTGATTCAAGGGTGGAAGACGATGACGGTGTATCGGATACAGAGTCCACTAATGGCA
ATGACTCTGGTAAAACTACTTCCACAGAAGAAAATGGTGACCCAAGCGGACCAGACATCCTGGCTGCTGTACGTAAA
CACCTAGACACTGTCTATCCAGGAGAAAATGGCGGATCTACAGAAGGACCTCTCCCTGCTAATCAAAATCTGGGGAA
CGTTATCCATGATGTGGAGCAGAATGGATCTGCTAAAGAAACTATTATCACTCCAGGAGATACAGGGCCTACAGACT
CAAGCTCCTCTGTAGATGCTGATGCAGACGTTGAAGATACTTCTGATACTGACTCTGGAATCGGAGACGACGACGGT
GTATCGGATACAGAGTCCACTAATGGTAATAACTCTGGTAAAACTACTTCCACAGAAGAAAATGGTGACCCAAGCGG
ACCAGACATCCTGGCTGCTGTACGTAAACACCTAGACACTGTCTATCCAGGAGAAAATGGCGGATCTACAGAAGGAC

CTCTCCCTGCTAATCAAAATCTGGGGAACGTTATCCATGATGTAGAACAAAACGGAGCCGCTCAAGAAACTATTATC
ACTCCAGGAGATACGGAATCTACAGACACAAGCTCTAGTGTAAATGCTAATGCAGACTTAGAAGATGTTTCTGATGC
TGATTCAGGATTCGGGGATGATGACGGTATATCGGATACAGAGTCCACTAATGGTAACGACTCTGGAAAAAATACTC
CTGTAGGGGATGGTGGTACACCAAGCGGACCAGATATCCTAGCTGCTGTACGCAAACATCTAGACACTGTCTATCCA
GGAGAAAATGGTGGATCTACAGAGAGACCTTTACCCGCTAATCAAAATTTAGGAGATATCATTCATGATGTAGAACA
AAACGGAAGCGCTAAAGAAACTGTAGTATCGCCTTATCGAGGAGGAGGAGGAAATACATCTTCCCCAATTGGATTAG
CCTCCCTGCTTCCAGCAACACCATCCACACCTTTGATGACAACACCTAGAACAAATGGGAAAGCTGCAGCTTCTTCT
TTGATGATAAAAGGAGGAGAAACTCAAGCCAAGCTAGTTAAGAATGGCGGCAATATCCCTGGAGAAACCACATTAGC
AGAATTACTCCCTCGTTTAAGAGGACACCTTGACAAAGTCTTTACTTCAGACGGGAAGTTTACAAATCTTAATGGAC
CTCAACTTGGAGCCATCATAGACCAATTCCGCAAAGAAACGGGTTCCGGAGGAATCATAGCTCATACAGATAGTGTT
CCAGGAGAGAACGGAACAGCCTCTCCTCTCACAGGAAGTTCAGGGGAAAAAGTCTCTCTCTATGATGCAGCGAAAAA
CGTCACTCAAGCTTTAACAAGTGTTACGAACAAAGTAACCCTAGCAATGCAAGGACAAAAACTGGAAGGAATTATAA
ACAACAACAATACCCCCTCTTCTATTGGACAAAATCTTTTCGCAGCAGCGAGGGCAACGACACAATCCCTCAGTTCA
TTAATTGGAACCGTACAATAA

SEQ ID NO:36: CM homologue of CT456 = TC_0741 protein sequence

MTTPISNSPSSIPTVTVSTTTASSGSLGTSTVSSTTTSTSVAQTATTTSSASTSIIQSSGENIQSTTGTPSPITSSV
STSAPSPKASATANKTSSAVSGKITSQETSEESETQATTSDGEVSSNYDDVDTPTNSSDSTVDSDYQDVETQYKTIS
NNGENTYETIGSHGEKNTHVQESHASGTGNPINNQQEAIRQLRSSTYTTSPRNENIFSPGPEGLPNMSLPSYSPTDK
SSLLAFLSNPNTKAKMLEHSGHLVFIDTTRSSFIFVPNGNWDQVCSMKVQNGKTKEDLGLKDLEDMCAKFCTGYNKF
SSDWGNRVDPLVSSKAGIESGGHLPSSVIINNKFRTCVAYGPWNPKENGPNYTPSAWRRGHRVDFGKIFDGTAPFNK
INWGSSPTPGDDGISFSNETIGSEPFATPPSSPSQTPVINVNVNVGGTNVNIGDTNVSKGSGTPTSSQSVDMSTDTS
DLDTSDIDTNNQTNGDINTNDNSNNVDGSLSDVDSRVEDDDGVSDTESTNGNDSGKTTSTEENGDPSGPDILAAVRK
HLDTVYPGENGGSTEGPLPANQNLGNVIHDVEQNGSAKETIITPGDTGPTDSSSSVDADADVEDTSDTDSGIGDDDG
VSDTESTNGNNSGKTTSTEENGDPSGPDILAAVRKHLDTVYPGENGGSTEGPLPANQNLGNVIHDVEQNGAAQETII
TPGDTESTDTSSSVNANADLEDVSDADSGFGDDDGISDTESTNGNDSGKNTPVGDGGTPSGPDILAAVRKHLDTVYP
GENGGSTERPLPANQNLGDIIHDVEQNGSAKETVVSPYRGGGGNTSSPIGLASLLPATPSTPLMTTPRTNGKAAASS
LMIKGGETQAKLVKNGGNIPGETTLAELLPRLRGHLDKVFTSDGKFTNLNGPQLGAIIDQFRKETGSGGIIAHTDSV
PGENGTASPLTGSSGEKVSLYDAAKNVTQALTSVTNKVTLAMQGQKLEGIINNNNTPSSIGQNLFAAARATTQSLSS
LIGTVQ

SEQ ID NO:37: CM homologue of CT381 = TC_0660

GTGAGTATGTATATAAAAAGAAAGAAAGCTTGGATGACTTTCTTAGCAATTGTCTGTAGTTTCTGTTTGGCGGGCTG
TTCAAAAGAGAGCAAAGACTCTGTTAGTGAAAAATTTATTGTAGGAACTAACGCAACGTATCCTCCTTTTGAGTTTG
TTGATGAAAGAGGTGAGACGGTTGGCTTTGATATTGATTTAGCTAGGGAGATTAGTAAAAAGCTAGGGAAAAAATTA
GAAGTCCGAGAATTTGCTTTTGATGCACTCGTTCTCAATTTAAAACAGCATCGTATTGATGCAATTATGGCAGGGGT
GTCCATTACGTCTTCTCGATTGAAAGAAATTTTGATGATTCCCTACTATGGCGAAGAAATAAAGAGTTTGGTTTTAG
TGTTTAAGGATGGAGACTCAAAGTCTTTACCACTAGATCAGTATAATTCTGTTGCTGTTCAAACTGGCACGTACCAA
GAGGAATATTTACAGTCTCTTCCAGGGGTGCGTATTCGCTCTTTTGATAGTACTTTAGAAGTGCTTATGGAAGTTTT
GCATAGCAAGTCTCCTATAGCTGTTTTAGAACCGTCTATTGCGCAGGTCGTTTTAAAAGATTTTCCGACGCTCACTA
CTGAAACGATAGATCTTCCTGAAGATAAATGGGTTTTAGGGTATGGAATTGGAGTTGCTTCTGATCGACCATCTCTA
GCTTCTGATATAGAAGCTGCTGTACAAGAGATCAAGAAAGAAGGAGTGTTAGCAGAGTTAGAGCAAAAATGGGGTTT
GAACGGCTAA

SEQ ID NO:38: CM homologue of CT381 = TC_0660

MSMYIKRKKAWMTFLAIVCSFCLAGCSKESKDSVSEKFIVGTNATYPPFEFVDERGETVGFDIDLAREISKKLGKKL
EVREFAFDALVLNLKQHRIDAIMAGVSITSSRLKEILMIPYYGEEIKSLVLVFKDGDSKSLPLDQYNSVAVQTGTYQ
EEYLQSLPGVRIRSFDSTLEVLMEVLHSKSPIAVLEPSIAQVVLKDFPTLTTETIDLPEDKWVLGYGIGVASDRPSL
ASDIEAAVQEIKKEGVLAELEQKWGLNG

SEQ ID NO:39 – CT255 nucleotide sequence

ATGGAAGAAAAAGGCATCTTACAATTGGTTGAAATTTCGCGAGCAATGGCTTTACAGGGAGTTTGTCCTTGGACTAA
TTTACAGAGTGTGGAGTCTATGTTGCAGTATATAGCAGGGGAGTGTCAGGAGTTGGCTGATGCTGTACAAGAAAATA
AAGCTTCGTTGGAAATCGCTTCGGAAGCCGGAGACGTACTTACTTTAGTATTGACCTTGTGTTTCTTGCTAGAAAGA
GAAGGAAAGCTTAAAGCTGAAGAAGTATTTGTAGAAGCTTTGGCTAAGTTGCGTCGTCGATCTCCTCATGTTTTTGA
TCCTCATAATCAAATTTCTTTAGAACAGGCTGAAGAATACTGGGCTCGTATGAAACAGCAAGAAAAAATTTCTTAA

SEQ ID NO:40 – CT255 protein sequence

MEEKGILQLVEISRAMALQGVCPWTNLQSVESMLQYIAGECQELADAVQENKASLEIASEAGDVLTLVLTLCFLLER
EGKLKAEEVFVEALAKLRRRSPHVFDPHNQISLEQAEEYWARMKQQEKIS

SEQ ID NO:41 – CT341 nucleotide sequence

ATGGATTACTACACGATATTGGGTGTAGCGAAGACTGCTACTCCTGAAGAAATAAAGAAAGCTTACCGTAAGCTCGC
TGTAAAGTACCATCCAGATAAGAATCCTGGGGATGCTGAAGCGGAGCGACGCTTTAAAGAAGTTTCTGAAGCCTATG
AAGTATTAGGTGATGCGCAGAAGCGGGAGTCATATGATCGTTACGGCAAAGACGGTCCATTTGCTGGTGCTGGAGGA
TTCGGTGGCGCTGGCATGGGGAATATGGAAGACGCTTTGCGAACATTTATGGGAGCTTTTGGCGGCGATTTCGGTGG

TAATGGAGGCGGTTTCTTTGAAGGGCTTTTTGGAGGACTTGGAGAAGCTTTCGGAATGCGTGGAGGCTCAGAAAGTT
CTCGACAAGGAGCTAGTAAGAAGGTGCATATTACGCTGTCCTTCGAGGAGGCGGCAAAAGGTGTTGAAAAAGAACTT
CTTGTTTCAGGCTATAAATCTTGTGATGCTTGTTCTGGTAGTGGAGCCAATACTGCTAAAGGTGTAAAAGTTTGTGA
TCGATGCAAGGGCTCTGGTCAGGTAGTGCAAAGCCGAGGCTTTTTCTCCATGGCTTCTACTTGCCCTGATTGTAGTG
GTGAAGGTCGGGTTATCACAGATCCTTGTTCAGTTTGTCGTGGGCAGGGACGTATCAAGGATAAACGTAGCGTCCAT
GTTAATATCCCAGCTGGAGTCGATTCTGGGATGAGATTAAAGATGGAAGGCTATGGAGATGCTGGCCAAAATGGAGC
GCCTGCAGGGGATCTGTATGTTTTTATTGATGTAGAGCCTCATCCTGTTTTCGAGCGCCATGGGGATGATTTAGTTT
TAGAGCTTCCTATTGGATTTGTTGATGCGGCTTTAGGGATCAAGAAGGAAATCCCTACACTCTTAAAAGAAGGTACT
TGCCGTTTGAGTATCCCAGAAGGGATTCAGAGCGGAACAGTTCTTAAAGTTAGAGGGCAGGGATTCCCTAATGTGCA
TGGGAAATCCAGAGGAGATCTTTTAGTAAGAGTATCTGTGGAGACTCCCCAGCACCTATCTAATGAACAAAAAGATT
TATTGAGACAGTTTGCTGCTACGGAGAAGGCTGAAAATTTCCCTAAGAAACGGAGTTTCTTAGACAAAATCAAAGGT
TTTTTTTCTGACTTTGCTGTATAG

SEQ ID NO:42 – CT341 protein sequence

MDYYTILGVAKTATPEEIKKAYRKLAVKYHPDKNPGDAEAERRFKEVSEAYEVLGDAQKRESYDRYGKDGPFAGAGG
FGGAGMGNMEDALRTFMGAFGGDFGGNGGGFFEGLFGGLGEAFGMRGGSESSRQGASKKVHITLSFEEAAKGVEKEL
LVSGYKSCDACSGSGANTAKGVKVCDRCKGSGQVVQSRGFFSMASTCPDCSGEGRVITDPCSVCRGQGRIKDKRSVH
VNIPAGVDSGMRLKMEGYGDAGQNGAPAGDLYVFIDVEPHPVFERHGDDLVLELPIGFVDAALGIKKEIPTLLKEGT
CRLSIPEGIQSGTVLKVRGQGFPNVHGKSRGDLLVRVSVETPQHLSNEQKDLLRQFAATEKAENFPKKRSFLDKIKG
FFSDFAV

SEQ ID NO:43 – CT716 nucleotide sequence

ATGAATAAAAAACTCCAAGATCTGTCTAAACTGCTCACTATTGAGCTTTTCAAGAAACGTACACGGTTGGAAACAGT
AAAAAAAGCGCTCTCCACAATAGAACATCGCTTACAACAAATACAGGAGCACATCGCGAAAATTTCCTTAACAAGGC
ACAAACAATTCCTATGTCGGTCATATACCCATGAATATGACCAACATTTAGAACATTTACAAAGAGAGCAAACTTCT
CTATATAAACAGCATCAGACCCTGAAAACGTCTTTGAAAGATGCTTATGGCGACATACAAAAACAACTAGACCAAAG
AAAAATTATCGAAAAGATCCATGACAGTAAATATCCTATAAAGAGCGCGAATAACTAA

SEQ ID NO:44 – CT716 protein sequence

MNKKLQDLSKLLTIELFKKRTRLETVKKALSTIEHRLQQIQEHIAKISLTRHKQFLCRSYTHEYDQHLEHLQREQTS
LYKQHQTLKTSLKDAYGDIQKQLDQRKIIEKIHDSKYPIKSANN

SEQ ID NO:45 – CT745 nucleotide sequence

ATGAAACATGCTCTCATTGTTGGCTCAGGTATTGCCGGCCTTTCTGCCGCGTGGTGGCTACACAAACGATTCCCTCA
TGTGCAGCTGTCTATTCTAGAAAAAGAGTCTCGATCTGGAGGGCTAATTGTCACAGAGAAACAACAAGGGTTTTCCC
TCAATATGGGCCCTAAAGGTTTTGTTTTAGCTCATGATGGGCAACACACCCTTCACCTCATTCAGTCTTTAGGCCTA
GCAGACGAGCTATTATATAGCTCTCCAGAGGCTAAAAACCGCTTTATCCACTATAATAATAAAACCCGAAAAGTCTC
GCCTTGGACTATTTTCAAACAAAATCTCCCTCTCTCTTTTGCTAAGGATTTCTTTGCGCGTCCTTACAAACAAGACA
GCTCCGTGGAAGCCTTCTTTAAAAGACACAGTTCTTCCAAGCTTAGAAGAAATCTTTTAAATCCCATTAGCATTGCT
ATTCGTGCAGGACATAGTCATATATTGTCTGCACAGATGGCTTACCCAGAATTAACACGAAGAGAAGCTCAAACAGG
ATCGTTGTTACGTAGTTATCTCAAAGATTTTCCTAAAGAGAAACGCACAGGCCCTTATTTAGCTACCTTGCGGTCTG
GGATGGGAATGCTAACCCAGGCTTTGCATGATAAATTGCCTGCTACCTGGTATTTTTCTGCACCCGTCAGCAAAATC
CGTCAGTTGGCGAATGGGAAAATTTCTCTTTCATCTCCTCAAGGAGAAATAACGGGAGATATGCTCATTTATGCTGG
GTCCGTGCACGATCTCCCTTCCTGTCTAGAAGGGATCCCTGAAACCAAGCTTATCAAGCAAACGACTTCATCTTGGG
ATCTCTCTTGTGTATCTTTAGGATGGCATGCATCCTTCCCTATCCCTCATGGATATGGCATGCTTTTCGCTGATACG
CCTCCCTTATTAGGGATCGTGTTTAATACGGAAGTGTTCCCTCAACCCGAGCGGCCTAATACAATAGTCTCTCTTCT
TTTAGAAGGTCGATGGCACCAAGAAGAAGCGTATGCTTTCTCACTAGCAGCTATTTCTGAGTACCTGCAAATTTACA
CTCCTCCCCAAGCTTTCTCACTATTCTCTCCTCGAGAGGGACTTCCCCAACACCATGTTGGATTTATCCAATCCCGC
CAACGCCTTCTATCTAAACTTCCTCACAATATAAAAATTGTAGGGCAGAATTTTGCAGGTCCAGGTCTCAACCGCGC
TACAGCGTCTGCTTATAAAGCTATAGCTTCTTTACTATCATGA

SEQ ID NO:46 – CT745 protein sequence

MKHALIVGSGIAGLSAAWWLHKRFPHVQLSILEKESRSGGLIVTEKQQGFSLNMGPKGFVLAHDGQHTLHLIQSLGL
ADELLYSSPEAKNRFIHYNNKTRKVSPWTIFKQNLPLSFAKDFFARPYKQDSSVEAFFKRHSSSKLRRNLLNPISIA
IRAGHSHILSAQMAYPELTRREAQTGSLLRSYLKDFPKEKRTGPYLATLRSGMGMLTQALHDKLPATWYFSAPVSKI
RQLANGKISLSSPQGEITGDMLIYAGSVHDLPSCLEGIPETKLIKQTTSSWDLSCVSLGWHASFPIPHGYGMLFADT
PPLLGIVFNTEVFPQPERPNTIVSLLLEGRWHQEEAYAFSLAAISEYLQIYTPPQAFSLFSPREGLPQHHVGFIQSR
QRLLSKLPHNIKIVGQNFAGPGLNRATASAYKAIASLLS

SEQ ID NO: 47 – CT387 nucleotide sequence

ATGACGCTCTTTCATTCTCATCATGATGCCGTCTCTCCAGACAGCTACCTATGTTCTTCCCTTCAGTTAGTTGGTAC
TGGCGTATACGAAGGAGAAATCGAGATTCAAAATATCCCCTCTTATTTCCTTGGATTCCAATTACCCTCTCATTGCA
TACACCTTAATTTAAAGAGCTCTCTAGCTCAATTAGGAATAGATGCCTCCCTTCTTCACTGCGAATTGAGCAAAAAT
CAACATCGAGCACATATACATGCTCAATTTACCGGTCATGGCCCCATTGCTGAATCTATGCTAGCCCTTCTCCAACC
AGGAGATCGTGTAGCAAAACTATTTGCTGCAGACGATCGCAGACTGGTCCGATCTCCAGATTACCTCGAAAGCATGC
TGAAAAATACAGATAAAGCTGGCCATCCTTTGCTCTGTTTTGGGAAAAAATTAGAACACTTGATTTCTTTTGATGTG

GTAGATGATCGCCTTGTCGTCTCCCTTCCTACCCTGCCGGGAGTTGTTCGTTATGATTCGGATATTTATGGACTCCT
TCCTCTTATTCAAAAATCACTCAGTAATCCCAAACTCAGCATTCGTCACTTTTTAGCTCTGTACCAACAGATTGTGG
AAGGGCAACATGTCTCTTGCGGAAACCATATTCTTCTGATCAAAACAGAACCGCTGCACATCCGCACTGTATTTGCT
CGCGTGGTAAATCAACTCCTCCCTCAAGGTCTCTCCCACACTTCTGCCAATATTTTGGAACCAACCACTCGAGAATC
CGGGGATATCTTTGAATTTTTTGGGAACCCTTCTGCACAGATAGAAAGAATTCCTTTAGAATTTTTCACTATCGAAC
CCTATAAAGAACATTCTTACTTCTGTAATCGGGATTTATTACAAACCATCTTACAATCAGAAAGCGAAATCAAAAAA
ATATTCGAAACAGCGCCCAAAGAACCTGTCAAAGCTGCCACCTATTTATCAAAAGGCAGTGAAATCTCTTCCCTGCA
CACAGACTCTTGGCTCACAGGATCCGCAGCTGCCTATCAATATAGTGAGCAAGCAGATAAAAACGAGTACACTCATG
CTCAACCTTGCTATCCTTTCTTAGAAGCAATGGAAATGGGCCTGATCAATAGCGAAGGAGCCTTACTCACTCGTTAT
TTCCCTTCAGCTAGCTTAAAAGGAATGTTGATTTCCTACCATGTGCGCCACTATCTCAAACAAATCTACTTTCAAGT
TCCCTCTTATACACATGGAAACTATTTCTCTCATAATGACAGAGGTTTGCTATTAGATCTGCAGCAAGCAGATATTG
ATGTTTTCTGGGCAGATGAAGAAAGCGGCCGTGTGTTGCAATATACAAAACGACGCGATAAGAATAGCGGTATGTTC
GTGATCAAAAATCGTGTTGAAGAGTTTCGATCAGCTTATTTTATTGCTATTTATGGCTCTCGTCTCCTTGAGAATAA
TTTCTCTGCTCAGCTCCATACCCTCCTAGCGGGCTTACAGCAAGCAGCACATACTCTCGGCATTCCTGGATTCTCAA
AGCCTACCCCACTTGCAGTCATCACCGGAGGCGGCACTGGAGTTATGGCCACAGGAAATCGTGTAGCTAAAGAACTA
GGAATCCTATCTTGTGGAACCGTTCTTGATTTAGAAGCTTCTCCAGCACAAATCGACCAACCTACCAATGAATTCTT
AGATGCTAAAATGACATACCGCCTACCTCAACTTATAGAAAGGCAAGAACACTTTTATGCAGACCTTCCTATCCTTG
TAGTTGGCGGTGTAGGAACCGATTTCGAACTCTACCTAGAACTTGTCTATCTCAAAACAGGAGCTAAACCACCGACT
CCCATTTTCCTAATTGGACCTATTGAATACTGGAAAGAAAAAGTGGCCCACGCCTACGAGATCAACCTCAAAGCAGG
AACCATCCGTGGATCCGAATGGATCAGCAACTGCCTATATTGTATCACTTCTCCGGAAGCTGGAATTGCCGTATTCG
AACAATTCCTAGCTGGAGAACTCCCTATAGGATACGACTATCCTCCAGCTCCAGATGGATTAGTGATCGTCTAA

SEQ ID NO:48 – CT387 protein sequence

MTLFHSHHDAVSPDSYLCSSLQLVGTGVYEGEIEIQNIPSYFLGFQLPSHCIHLNLKSSLAQLGIDASLLHCELSKN
QHRAHIHAQFTGHGPIAESMLALLQPGDRVAKLFAADDRRLVRSPDYLESMLKNTDKAGHPLLCFGKKLEHLISFDV
VDDRLVVSLPTLPGVVRYDSDIYGLLPLIQKSLSNPKLSIRHFLALYQQIVEGQHVSCGNHILLIKTEPLHIRTVFA
RVVNQLLPQGLSHTSANILEPTTRESGDIFEFFGNPSAQIERIPLEFFTIEPYKEHSYFCNRDLLQTILQSESEIKK
IFETAPKEPVKAATYLSKGSEISSLHTDSWLTGSAAAYQYSEQADKNEYTHAQPCYPFLEAMEMGLINSEGALLTRY
FPSASLKGMLISYHVRHYLKQIYFQVPSYTHGNYFSHNDRGLLLDLQQADIDVFWADEESGRVLQYTKRRDKNSGMF
VIKNRVEEFRSAYFIAIYGSRLLENNFSAQLHTLLAGLQQAAHTLGIPGFSKPTPLAVITGGGTGVMATGNRVAKEL
GILSCGTVLDLEASPAQIDQPTNEFLDAKMTYRLPQLIERQEHFYADLPILVVGGVGTDFELYLELVYLKTGAKPPT
PIFLIGPIEYWKEKVAHAYEINLKAGTIRGSEWISNCLYCITSPEAGIAVFEQFLAGELPIGYDYPPAPDGLVIV

SEQ ID NO:49 – CT812 nucleotide sequence

ATGAGTTCCGAGAAAGATATAAAAAGCACCTGTTCTAAGTTTTCTTTGTCTGTAGTAGCAGCTATCCTTGCCTCTGT
TAGCGGGTTAGCTAGTTGCGTAGATCTTCATGCTGGAGGACAGTCTGTAAATGAGCTGGTATATGTAGGCCCTCAAG
CGGTTTTATTGTTAGACCAAATTCGAGATCTATTCGTTGGGTCTAAAGATAGTCAGGCTGAAGGACAGTATAGGTTA
ATTGTAGGAGATCCAAGTTCTTTCCAAGAGAAAGATGCGGATACTCTTCCCGGGAAGGTAGAGCAAAGTACTTTGTT
CTCAGTAACCAATCCCGTGGTTTTCCAAGGTGTGGACCAACAGGATCAAGTCTCTTCCCAAGGGTTAATTTGTAGTT
TTACGAGCAGCAACCTTGATTCTCCTCGTGACGGAGAATCTTTTTTAGGTATTGCTTTTGTTGGGGATAGTAGTAAG
GCTGGAATCACATTAACTGACGTGAAAGCTTCTTTGTCTGGAGCGGCTTTATATTCTACAGAAGATCTTATCTTTGA
AAAGATTAAGGGTGGATTGGAATTTGCATCATGTTCTTCTCTAGAACAGGGGGGAGCTTGTGCAGCTCAAAGTATTT
TGATTCATGATTGTCAAGGATTGCAGGTTAAACACTGTACTACAGCCGTGAATGCTGAGGGGTCTAGTGCGAATGAT
CATCTTGGATTTGGAGGAGGCGCTTTCTTTGTTACGGGTTCTCTTTCTGGAGAGAAAAAGTCTCTATATGCCTGCAGG
AGATATGGTAGTTGCGAATTGTGATGGGGCTATATCTTTTGAAGGAAACAGCGCGAACTTTGCTAATGGAGGAGCGA
TTGCTGCCTCTGGGAAAGTGCTTTTTGTCGCTAATGATAAAAAGACTTCTTTTATAGAGAACCGAGCTTTGTCTGGA
GGAGCGATTGCAGCCTCTTCTGATATTGCCTTTCAAAACTGCGCAGAACTAGTTTTCAAAGGCAATTGTGCAATTGG
AACAGAGGATAAAGGTTCTTTAGGTGGAGGGGCTATATCTTCTCTAGGCACCGTTCTTTTGCAAGGGAATCACGGGA
TAACTTGTGATAAGAATGAGTCTGCTTCGCAAGGAGGCGCCATTTTTGGCAAAAATTGTCAGATTTCTGACAACGAG
GGGCCAGTGGTTTTCAGAGATAGTACAGCTTGCTTAGGAGGAGGCGCTATTGCAGCTCAAGAAATTGTTTCTATTCA
GAACAATCAGGCTGGGATTTCCTTCGAGGGAGGTAAGGCTAGTTTCGGAGGAGGTATTGCGTGTGGATCTTTTTCTT
CCGCAGGTGGTGCTTCTGTTTTAGGGACCATTGATATTTCGAAGAATTTAGGCGCGATTTCGTTCTCTCGTACTTTA
TGTACGACCTCAGATTTAGGACAAATGGAGTACCAGGGAGGAGGAGCTCTATTTGGTGAAAATATTTCTCTTTCTGA
GAATGCTGGTGTGCTCACCTTTAAAGACAACATTGTGAAGACTTTTGCTTCGAATGGGAAAATTCTGGGAGGAGGAG
CGATTTTAGCTACTGGTAAGGTGGAAATTACTAATAATTCCGAAGGAATTTCTTTTACAGGAAATGCGAGAGCTCCA
CAAGCTCTTCCAACTCAAGAGGAGTTTCCTTTATTCAGCAAAAAAGAAGGGCGACCACTCTCTTCAGGATATTCTGG
GGGAGGAGCGATTTTAGGAAGAGAAGTAGCTATTCTCCACAACGCTGCAGTAGTATTTGAGCAAAATCGTTTGCAGT
GCAGCGAAGAAGAAGCGACATTATTAGGTTGTTGTGGAGGAGGCGCTGTTCATGGGATGGATAGCACTTCGATTGTT
GGCAACTCTTCAGTAAGATTTGGTAATAATTACGCAATGGGACAAGGAGTCTCAGGAGGAGCTCTTTTATCTAAAAC
AGTGCAGTTAGCTGGGAATGGAAGCGTCGATTTTTCTCGAAATATTGCTAGTTTGGGAGGAGGAGCTCTTCAAGCTT
CTGAAGGAAATTGTGAGCTAGTTGATAACGGCTATGTGCTATTCAGAGATAATCGAGGGAGGGTTTATGGGGGTGCT
ATTTCTTGCTTACGTGGAGATGTAGTCATTTCTGGAAACAAGGGTAGAGTTGAATTTAAAGACAACATAGCAACACG
TCTTTATGTGGAAGAAACTGTAGAAAAGGTTGAAGAGGTAGAGCCAGCTCCTGAGCAAAAAGACAATAATGAGCTTT
CTTTCTTAGGGAGAGCAGAACAGAGTTTTATTACTGCAGCTAATCAAGCTCTTTTCGCATCTGAAGATGGGGATTTA
TCACCTGAGTCATCCATTTCTTCTGAAGAACTTGCGAAAAGAAGAGAGTGTGCTGGAGGAGCTATTTTTGCAAAACG
GGTTCGTATTGTAGATAACCAAGAGGCCGTTGTATTCTCGAATAACTTCTCTGATATTTATGGCGGCGCCATTTTTA
CAGGTTCTCTTCGAGAAGAGGATAAGTTAGATGGGCAAATCCCTGAAGTCTTGATCTCAGGCAATGCAGGGGATGTT
GTTTTTTCCGGAAATTCCTCGAAGCGTGATGAGCATCTTCCTCATACAGGTGGGGGAGCCATTTGTACTCAAAATTT

GACGATTTCTCAGAATACAGGGAATGTTCTGTTTTATAACAACGTGGCCTGTTCGGGAGGAGCTGTTCGTATAGAGG
ATCATGGTAATGTTCTTTTAGAAGCTTTTGGAGGAGATATTGTTTTTAAAGGAAATTCTTCTTTCAGAGCACAAGGA
TCCGATGCTATCTATTTTGCAGGTAAAGAATCGCATATTACAGCCCTGAATGCTACGGAAGGACATGCTATTGTTTT
CCACGACGCATTAGTTTTTGAAAATCTAGAAGAAAGGAAATCTGCTGAAGTATTGTTAATCAATAGTCGAGAAAATC
CAGGTTACACTGGATCTATTCGATTTTTAGAAGCAGAAAGTAAAGTTCCTCAATGTATTCATGTACAACAAGGAAGC
CTTGAGTTGCTAAATGGAGCCACATTATGTAGTTATGGTTTTAAACAAGATGCTGGAGCTAAGTTGGTATTGGCTGC
TGGAGCTAAACTGAAGATTTTAGATTCAGGAACTCCTGTACAACAAGGGCATGCTATCAGTAAACCTGAAGCAGAAA
TCGAGTCATCTTCTGAACCAGAGGGTGCACATTCTCTTTGGATTGCGAAGAATGCTCAAACAACAGTTCCTATGGTT
GATATCCATACTATTTCTGTAGATTTAGCCTCCTTCTCTTCTAGTCAACAGGAGGGGACAGTAGAAGCTCCTCAGGT
TATTGTTCCTGGAGGAAGTTATGTTCGATCTGGAGAGCTTAATTTGGAGTTAGTTAACACAACAGGTACTGGTTATG
AAAATCATGCTTTATTGAAGAATGAGGCTAAAGTTCCATTGATGTCTTTCGTTGCTTCTGGTGATGAAGCTTCAGCC
GAAATCAGTAACTTGTCGGTTTCTGATTTACAGATTCATGTAGTAACTCCAGAGATTGAAGAAGACACATACGGCCA
TATGGGAGATTGGTCTGAGGCTAAAATTCAAGATGGAACTCTTGTCATTAGTTGGAATCCTACTGGATATCGATTAG
ATCCTCAAAAAGCAGGGGCTTTAGTATTTAATGCATTATGGGAAGAAGGGGCTGTCTTGTCTGCTCTGAAAAATGCA
CGCTTTGCTCATAATCTCACTGCTCAGCGTATGGAATTCGATTATTCTACAAATGTGTGGGGATTCGCCTTTGGTGG
TTTCCGAACTCTATCTGCAGAGAATCTGGTTGCTATTGATGGATACAAAGGAGCTTATGGTGGTGCTTCTGCTGGAG
TCGATATTCAATTGATGGAAGATTTTGTTCTAGGAGTTAGTGGAGCTGCTTTCCTAGGTAAAATGGATAGTCAGAAG
TTTGATGCGGAGGTTTCTCGGAAGGGAGTTGTTGGTTCTGTATATACAGGATTTTTAGCTGGATCCTGGTTCTTCAA
AGGACAATATAGCCTTGGAGAAACACAGAACGATATGAAAACGCGTTATGGAGTACTAGGAGAGTCGAGTGCTTCTT
GGACATCTCGAGGAGTACTGGCAGATGCTTTAGTTGAATACCGAAGTTTAGTTGGTCCTGTGAGACCTACTTTTTAT
GCTTTGCATTTCAATCCTTATGTCGAAGTATCTTATGCTTCTATGAAATTCCCTGGCTTTACAGAACAAGGAAGAGA
AGCGCGTTCTTTTGAAGACGCTTCCCTTACCAATATCACCATTCCTTTAGGGATGAAGTTTGAATTGGCGTTCATAA
AAGGACAGTTTTCAGAGGTGAACTCTTTGGGAATAAGTTATGCATGGGAAGCTTATCGAAAAGTAGAAGGAGGCGCG
GTGCAGCTTTTAGAAGCTGGGTTTGATTGGGAGGGAGCTCCAATGGATCTTCCTAGACAGGAGCTGCGTGTCGCTCT
GGAAAATAATACGGAATGGAGTTCTTACTTCAGCACAGTCTTAGGATTAACAGCTTTTTGTGGAGGATTTACTTCTA
CAGATAGTAAACTAGGATATGAGGCGAATACTGGATTGCGATTGATCTTTTAA

SEQ ID NO:50 – CT812 protein sequence

MSSEKDIKSTCSKFSLSVVAAILASVSGLASCVDLHAGGQSVNELVYVGPQAVLLLDQIRDLFVGSKDSQAEGQYRL
IVGDPSSFQEKDADTLPGKVEQSTLFSVTNPVVFQGVDQQDQVSSQGLICSFTSSNLDSPRDGESFLGIAFVGDSSK
AGITLTDVKASLSGAALYSTEDLIFEKIKGGLEFASCSSLEQGGACAAQSILIHDCQGLQVKHCTTAVNAEGSSAND
HLGFGGGAFFVTGSLSGEKSLYMPAGDMVVANCDGAISFEGNSANFANGGAIAASGKVLFVANDKKTSFIENRALSG
GAIAASSDIAFQNCAELVFKGNCAIGTEDKGSLGGGAISSLGTVLLQGNHGITCDKNESASQGGAIFGKNCQISDNE
GPVVFRDSTACLGGGAIAAQEIVSIQNNQAGISFEGGKASFGGGIACGSFSSAGGASVLGTIDISKNLGAISFSRTL
CTTSDLGQMEYQGGGALFGENISLSENAGVLTFKDNIVKTFASNGKILGGGAILATGKVEITNNSEGISFTGNARAP
QALPTQEEFPLFSKKEGRPLSSGYSGGGAILGREVAILHNAAVVFEQNRLQCSEEEATLLGCCGGGAVHGMDSTSIV
GNSSVRFGNNYAMGQGVSGGALLSKTVQLAGNGSVDFSRNIASLGGGALQASEGNCELVDNGYVLFRDNRGRVYGGA
ISCLRGDVVISGNKGRVEFKDNIATRLYVEETVEKVEEVEPAPEQKDNNELSFLGRAEQSFITAANQALFASEDGDL
SPESSISSEELAKRRECAGGAIFAKRVRIVDNQEAVVFSNNFSDIYGGAIFTGSLREEDKLDGQIPEVLISGNAGDV
VFSGNSSKRDEHLPHTGGGAICTQNLTISQNTGNVLFYNNVACSGGAVRIEDHGNVLLEAFGGDIVFKGNSSFRAQG
SDAIYFAGKESHITALNATEGHAIVFHDALVFENLEERKSAEVLLINSRENPGYTGSIRFLEAESKVPQCIHVQQGS
LELLNGATLCSYGFKQDAGAKLVLAAGAKLKILDSGTPVQQGHAISKPEAEIESSSEPEGAHSLWIAKNAQTTVPMV
DIHTISVDLASFSSSQQEGTVEAPQVIVPGGSYVRSGELNLELVNTTGTGYENHALLKNEAKVPLMSFVASGDEASA
EISNLSVSDLQIHVVTPEIEEDTYGHMGDWSEAKIQDGTLVISWNPTGYRLDPQKAGALVFNALWEEGAVLSALKNA
RFAHNLTAQRMEFDYSTNVWGFAFGGFRTLSAENLVAIDGYKGAYGGASAGVDIQLMEDFVLGVSGAAFLGKMDSQK
FDAEVSRKGVVGSVYTGFLAGSWFFKGQYSLGETQNDMKTRYGVLGESSASWTSRGVLADALVEYRSLVGPVRPTFY
ALHFNPYVEVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFELAFIKGQFSEVNSLGISYAWEAYRKVEGGA
VQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYFSTVLGLTAFCGGFTSTDSKLGYEANTGLRLIF

SEQ ID NO:51 – CT869 nucleotide sequence

ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAGGTTCCTTCTAGAATCTT
TCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACA
ATCTCACTAACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCT
GTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGA
AAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCT
TTGAAAATAATACTTGTTGCAGACTATTTACATGGAGAAATCCTTATGCTGCTGATAAAATAAGAGAAGGCGGAGCC
ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAAGG
AGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTCTTTATGGACAACATCTGTA
TTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGAT
CTCTTCTTCATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGG
TAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAA
TTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTAT
GGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAA
TAAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATG
AAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACA
GTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGT
TAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTG
TTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAATGGTTTTCTA

TGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGC
AGTTCTGAGTTGCTATAAAAATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGA
ATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT
ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTACGGGAACTCTCCTTA
TGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTAC
AATCAGAAAATATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCA
AAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCATAGAACCTT
ACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGG
GGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATT
ACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGG
ATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATG
AGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGT
TTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAA
TCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAACCCTTTG
GATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTG
GGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTT
TATGAATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAG
GGATCGCAGCCCAGCTCCTAGCCAGTAAGGGTATTTGGTTCGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCC
TATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTC
AACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTAG

SEQ ID NO:52 – CT869 protein sequence

MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAA
VTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRLFTWRNPYAADKIREGGA
IHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD
LFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNY
GGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAIT
VASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFL
CIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY
TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWGWA
KTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSGHPFWGI
TGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEG
FLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEV
GAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIAAQLLASKGIWFGSGSPSSRHAMS
YKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF

SEQ ID NO:53 – CT166 nucleotide sequence

GTGAACGTTCGTACGTACTCTGTTCAGAGGGGGGGGGTAAAAACGATTTCTGCTAGTGCAGTTCCTCCTACAGCAGC
TGTTTTATCGAGAAAAAAGCGTGCTATAGAAGAGAAGAAGGAGGAAGCTTCTTCTGGAAAGATAGAAAATCTTGATG
CTAGCAAATACGATCTTACTCCCAAGAACATAGAAGAAAAACTAGGAATTACTCCTGAACAGAAATCTACTGTTAAA
GACCTATTAAATAAACTGAAAAAGGTCATTAGTGCTTACAACTCTATGCCAGATAAAAATTCGGAAGCGGGACAGAA
TTCCTTGATTCAACAAGGAAAATACGTCGATGCCATTCAGAAGAAGCTTCCAGCATCATCGCAGGCTCAGCCTAAAC
AGGCAAAAGCTAAGGAACAGAAAGCCGAAGAAAAACCTAAGACGACTCCGATTGAAGGTGTTCTTGAAACCATCAAA
ACAGAATTTAAAGGCCATCGTGTACCTGTTGAGAAAATCATCCATGGAATATGGATCGCAGGAGCGCCTCCGGATGG
TATCGAAGATTATATGCGAGTCTTTTTAGATACTTATGAAGGTTTTGACTTCTACTTCTGGGTAGATGAGAATGCTT
ATGCAGCAGCTAAATTTTCTAGCATTTTGAAGAAGGTCGCTTTCGATGCGGCTATTCAAGATCTACGATCTGCCACA
GATGAGTCTACGAAGGCCTTTGTTAAAGACTACGATGAATTAAAACAGAAATATGAAAAGAAAGTTGCGGAGACGAC
TTCTCAAGCAGAAAAAGACCAATATCTCAAAGATCTAAAGGATCTTTTAGAGAAATTTACAAAAATCAGTGATGAGA
TTCGTGGAAAATTTGATCGGCTGTTTCTTAAGAATGTGATTGTTGCTCAGAACGGATTCTTTAATTTCTGCTTGCTG
AAAGGCCTCGGCAATATCAATGACGAAACGCGTGCAGAGTATTTAGAGAAAGAACTCAAACTTCCTACTGAGGAGAT
CGAACAGTATAAAAAGCTTAAAGAGACGAACAAAGAGAAGATAGCCGCTATTGTAAAACAACTAAACGAGAAACTTG
GATCGGATCGGGTAAAAATCAAAGACATTAAAGAGCTGCAATCTATGAAGCAAGCTCGAAATGTCTACAATTATGAA
CAGGAAATGTTTCTGCGCTGGAACTATGCAGCCGCAACAGATCAGATTCGTATGTATATGTTGGAGGAACTTGGAGG
TCTTTATACTGATCTGGATATGATGCCTTCATACTCTCAGGAAGTATTGGAGCTTATCAAAAAGCACAGTGATGGAA
ACCGAATGTTTGAGGATATGAGCTCTAGACGGGCGATTTCTGATGCGGTTTTAAAGATGGCTGTAGGTAAGGCGACA
ACAGTTTCCATGGAAGAGGTAGCAAAGGATATCGATGTTTCTCGCTTAACAGAAGAGGATAAGACAAAATTAAATGC
TCTATTTAAGGATCTAGAGCCATTTGCAAAACCGGATTCTAAAGGAGCTGAAGCAGAAGGGGGTGAAGGAGCAAAAG
GTATGAAAAAGAGCTTTTTCCAGCCCATAGATCTGAATATTGTCAGAAATACCATGCCTATCTTGAGACGCTATCAT
CACTATCCTGAGTTAGGATGGTTTATTCGAGGATTGAACGGATTGATGGTCTCTCATAAGGGAAGCACTGCGGTTTC
TGCTGTCATTGTAGGGCAACAGGCTGCCTACCAGGAACTAGCAGCACTTAGACAAGATGTCCTTTCAGGGGAGTTTT
TCCATTCTTTAGAAAATTTGACACATAGAAACCATAAGGAGCGTATTGGAAATCATCTCGTCGCTAATTATTTGGCT
AAAAGTCTCTTTTTTGATTACTGCCAAGATTCAGTGATGCCGGAGGCTGTAAGTACCTTAGGTATTAGATGA

SEQ ID NO: 54 – CT166 protein sequence

MNVRTYSVQRGGVKTISASAVPPTAAVLSRKKRAIEEKKEEASSGKIENLDASKYDLTPKNIEEKLGITPEQKSTVK
DLLNKLKKVISAYNSMPDKNSEAGQNSLIQQGKYVDAIQKKLPASSQAQPKQAKAKEQKAEEKPKTTPIEGVLETIK
TEFKGHRVPVEKIIHGIWIAGAPPDGIEDYMRVFLDTYEGFDFYFWVDENAYAAAKFSSILKKVAFDAAIQDLRSAT
DESTKAFVKDYDELKQKYEKKVAETTSQAEKDQYLKDLKDLLEKFTKISDEIRGKFDRLFLKNVIVAQNGFFNFCLL

KGLGNINDETRAEYLEKELKLPTEEIEQYKKLKETNKEKIAAIVKQLNEKLGSDRVKIKDIKELQSMKQARNVYNYE
QEMFLRWNYAAATDQIRMYMLEELGGLYTDLDMMPSYSQEVLELIKKHSDGNRMFEDMSSRRAISDAVLKMAVGKAT
TVSMEEVAKDIDVSRLTEEDKTKLNALFKDLEPFAKPDSKGAEAEGGEGAKGMKKSFFQPIDLNIVRNTMPILRRYH
HYPELGWFIRGLNGLMVSHKGSTAVSAVIVGQQAAYQELAALRQDVLSGEFFHSLENLTHRNHKERIGNHLVANYLA
KSLFFDYCQDSVMPEAVSTLGIR

SEQ ID NO:55 – CT175 nucleotide sequence

ATGCATCACAGGAAGTTTTTAGCAGTTTCCATTGCTTTCGTAAGTTTAGCTTTTGGGCTAACATCTTGTTATCATAA
AAAAGAAGAACCAAAAGATGTTTTGCGGATTGCGATCTGTCATGATCCAATGTCTTTAGATCCGCGTCAGGTTTTTT
TAAGCAAAGATGTTTCTATTGTAAAAGCTCTCTATGAAGGGTTAGTCCGGGAAAAAGAAGCTGCGTTCCAGCTAGCT
TTGGCAGAAAGATATCATCAATCTGATGATGGTTGTGTTTATACTTTTTTTCTAAAAAATACATTCTGGAGCAACGG
AGATGTTGTAACAGCATATGATTTTGAAGAGTCTATTAAACAAATTTATTTCCGAGAAATTGATAACCCTTCGTTAC
GCTCTCTTGCATTAATTAAAAATTCTCATGCTGTTTTAACAGGAGCTCTCCCTGTTGAAGATTTAGGTGTTAGAGCT
TTGAATGCGAAAACTCTAGAAATTGTTTTAGAAAACCCGTTTCCTTATTTTCTAGAGATATTGGCGCACCCGGTTTT
TTATCCGGTGCACACCTCTTTACGAGAATATTACAAAGATAAGCGTAACAAACGCGTTTTCCCGATAATTTCTAATG
GTCCTTTTGCGATTCAATGTTATGAGCCGCAAAGATATTTACTAATCAACAAAAACCCTCTGTATCATGCCAAGCAC
GATGTTCTGTTAAATTCGGTATGTTTGCAGATAGTTCCTGATATCCATACAGCTATGCAGTTATTCCAAAAAAATCA
TATCGATTTAGTTGGGTTACCCTGGAGCTCCTCCTTTTCTTTAGAAGAACAAAGAAATCTCCCTAGAGAAAAATTAT
TTGATTATCCTGTATTGAGTTGCTCTGTTTTATTCTGTAACATTCATCAAACACCTTTAAATAATCCCTCGCTGAGA
ACAGCCCTCTCTTTAGCAATCAATCGAGAAACTTTATTAAAACTAGCAGGTAAAGGCTGTAGCGCTACGAGCTTTGT
TCACCCACAATTATCTCAGATACCTGCTACTACTTTGTCTCAAGATGAGCGGATTGCTTTAGCAAAAGGCTACTTGA
CCGAAGCTTTAAAGACTTTATCTCAAGAAGATTTAGAAAAAATTACATTAATTTATCCTATAGAATCTGTTTGCTTA
CGAGCCGTTGTTCAAGAAATTCGCCAACAATTATTTGATGTACTGGGATTTAAAATTTCTACATTAGGATTAGAATA
TCATTGTTTTTTAGACAAACGTTCCAGAGGAGAATTCTCCTTAGCAACTGGTAATTGGATTGCAGACTATCATCAAG
CTAGTGCTTTCCTGTCTGTCCTAGGTAATGGGACAAGATATAAAGACTTTCAATTGATTAACTGGCAGAACCAAAAG
TACACAAATATAGTTGCTCAACTTCTGATTCAAGAATCAAGCGACCTACAGCTTATGGCAGAGCAGTTGTTGCTTAA
AGAAAGTCCTCTTATTCCTCTATACCACCTCGATTATGTGTATGCGAAACAGCCTCGGGTGTCTGATCTCCAAACCT
CTTCTCGTGGAGAAATTGATTTAAAAAGAGTTTCATTAGCTGAAGGATAG

SEQ ID NO:56 – CT175 protein sequence

MHHRKFLAVSIAFVSLAFGLTSCYHKKEEPKDVLRIAICHDPMSLDPRQVFLSKDVSIVKALYEGLVREKEAAFQLA
LAERYHQSDDGCVYTFFLKNTFWSNGDVVTAYDFEESIKQIYFREIDNPSLRSLALIKNSHAVLTGALPVEDLGVRA
LNAKTLEIVLENPFPYFLEILAHPVFYPVHTSLREYYKDKRNKRVFPIISNGPFAIQCYEPQRYLLINKNPLYHAKH
DVLLNSVCLQIVPDIHTAMQLFQKNHIDLVGLPWSSSFSLEEQRNLPREKLFDYPVLSCSVLFCNIHQTPLNNPSLR
TALSLAINRETLLKLAGKGCSATSFVHPQLSQIPATTLSQDERIALAKGYLTEALKTLSQEDLEKITLIYPIESVCL
RAVVQEIRQQLFDVLGFKISTLGLEYHCFLDKRSRGEFSLATGNWIADYHQASAFLSVLGNGTRYKDFQLINWQNQK
YTNIVAQLLIQESSDLQLMAEQLLLKESPLIPLYHLDYVYAKQPRVSDLQTSSRGEIDLKRVSLAEG

SEQ ID NO: 57 – TC0666 nucleotide sequence (homologue of CT387)

ATGAGGATTCCAATGACACTCTTTCACACTCATCACGATGCCGTCTCTCCGGACGGCTACTTATGTTCTTCCCTTCA
GTTAGTTGGCTCTGGCACATATGAAGGAGAAATCGAAATCCAAAATATTCCTTCTTATTTCCTTGGATTCCGATTAC
CCACCCATTGCGTTCATCTTAATTTGAAGAGTTCTCTAGCCCAGTTAGGAGTAGATGCATCTCTTCTTCACTGCGAA
CTAAGCAAAAATCAACAACGTGCACATATGCACGTGCAGTTCACCGGCTATGGCCCTATCGCTGAGTCCATGCTATC
TCTTCTCAAACCCGGAGATCGAGTAGCCAAACTGTTTGCTGCAGATGATCGTAGACTAGTCCGCTCCCCTGATTATC
TTGAAAGCATGCTAAAAAATACTGATAAGACAGGACATCCTCTGCTCCGATTTGGAAAAAAACTCGAGCATCTTATC
TCTTTTGATGTGGTGGACGATCGCCTCGTTGTATCACTCCCCACCTTGCCAGGCATAGTCAATTATGACCCAGACAT
CTATGGACTTCTTCCCTTAATTCAAAAATCACTAAGCAATCCTAAATTGAGTATTCGCCACTTCTTGTCTCTCTATC
AGAAGATCGTAGAAGGACCACACATCCCTTATGAAGGAAACATTTTGTTAATCAAAACAGAGCCTCTTCATATCCGC
ACAGTATTTGCTCGCGTGGTCGATCAAATGCTCCCTCAAGGTCTATTTCACACTTCTGCCAACATTTTAGAACCCAC
AACGCGAGAGTCTGGAGATATTTTTGAATTTTTTGGAAATCCCTCCACTCTTGTAGAAAGAATCCCTCTAGAATTCT
TCACTATCGAACCCTACAAAGAACACTCTTACTTCTGTAATCGAGATCTATTGCAAACTACCTTGCAATCGGAAAGT
GAAATCAAAAAAATATTCGATACAGCTCCTCAAGAGCCTGTAAAAGCCGCCACTTATTTATCAAAAGGAAGTGAAAT
TTCTTCTCTTGATGCAGATTCTTGGCTTACGGGATCCGCAGCTGCATACCAATGTAGCGAAAAACAGGCAGCTAAAG
ACGAATACATCCACGCTCAACCCTGTTATCCATTTTTGGAAGCAATGGAAACGGGACTCATCAATAGCGAAGGAGCT
TTACTCACTCGGTTTTTCCCCTCTTCCAGCTTAAAAGGGATGTTGATCTCCTATCATGTACGCCACTATCTTAAGCA
AATTTACTTTCAAGTTCCTTCTTATACATATGGAGACTACTTCTCTCATAATGACCGAGGATTACTGTTAGATCTAT
ATCAGGCGAACATTGATGTGTTCTGGGCTGATGAAGAGAGCGGCCGTGTATTGCAATATACAAAACGGCGCGACAAA
AATAGTGGAATGTTCGTCGTTAAAAATCGAGTAGAAGAGTTCCAATCAGCATATTTCGTAGCGATTTATGGATCACG
TCTCCTGGAAAATAATTTCTCGGCCCAACTAAACACGCTTCTTGCAGGGTTACAAAAAGCTGCACACACTCTAGGCA
TTCCAGGCTTCTCAAAACCCACTCCTCTTGCCGTAATCACAGGAGGAGGGACTGGCGTTATGGCTACAGGAAATCGT
GTTGCAAAAGAGTTGGGAATTCTTTCTTGCGGGACCGTTCTCGATTTGGAAGCTTCACCTGCACAAATAGATCAGCC
TGCAAACGAATTTTTAGATGCCAAAATGACATACCGTCTACCGCAACTTATAGAAAGACAAGAACATTTTTATTCAG
ACCTTGCCATTTTAGTTGTTGGTGGTGTTGGAACAGATTTCGAACTTTACCTAGAACTCGTCTACTTGAAAACAGGC
GCCAAACCTCCTACTCCAATTTTCCTTATTGGGCCTGTTGAATACTGGAAAGAGAAAGTTGCTCATGCCTATGAGAT
TAATCTTAAAGCAGGAACTATTCGTGGTTCTGAGTGGATCAGCAACTGCTTATTCTGCATTACATCTCCTGAAGCAG
GAATTGCTGTATTCGAACAGTTCCTCGCTGGAGAACTTCCCATAGGATATGATTATCCTCCAGCTCCAGACGGATTA
GTTATCGTCTAA

SEQ ID NO: 58 - TC0666 protein sequence (homologue of CT387)

MRIPMTLFHTHHDAVSPDGYLCSSLQLVGSGTYEGEIEIQNIPSYFLGFRLPTHCVHLNLKSSLAQLGVDASLLHCE
LSKNQQRAHMHVQFTGYGPIAESMLSLLKPGDRVAKLFAADDRRLVRSPDYLESMLKNTDKTGHPLLRFGKKLEHLI
SFDVVDDRLVVSLPTLPGIVNYDPDIYGLLPLIQKSLSNPKLSIRHFLSLYQKIVEGPHIPYEGNILLIKTEPLHIR
TVFARVVDQMLPQGLFHTSANILEPTTRESGDIFEFFGNPSTLVERIPLEFFTIEPYKEHSYFCNRDLLQTTLQSES
EIKKIFDTAPQEPVKAATYLSKGSEISSLDADSWLTGSAAAYQCSEKQAAKDEYIHAQPCYPFLEAMETGLINSEGA
LLTRFFPSSSLKGMLISYHVRHYLKQIYFQVPSYTYGDYFSHNDRGLLLDLYQANIDVFWADEESGRVLQYTKRRDK
NSGMFVVKNRVEEFQSAYFVAIYGSRLLENNFSAQLNTLLAGLQKAAHTLGIPGFSKPTPLAVITGGGTGVMATGNR
VAKELGILSCGTVLDLEASPAQIDQPANEFLDAKMTYRLPQLIERQEHFYSDLAILVVGGVGTDFELYLELVYLKTG
AKPPTPIFLIGPVEYWKEKVAHAYEINLKAGTIRGSEWISNCLFCITSPEAGIAVFEQFLAGELPIGYDYPPAPDGL
VIV

SEQ ID NO: 59 - TC0197 nucleotide sequence

ATGAGTTCCGAGAAAGATAAAAAAAAACTCCTGTTCTAAGTTTTCCTTATCGGTAGTAGCAGCTATTCTCGCTTCTAT
GAGTGGTTTATCGAATTGTTCCGATCTTTATGCCGTAGGAAGTTCTGCAGACCATCCTGCCTACTTGATTCCTCAAG
CGGGGTTATTATTGGATCATATTAAGGATATATTCATTGGCCCTAAAGATAGTCAGGATAAGGGGCAGTATAAGTTG
ATTATTGGTGAGGCTGGCTCTTTCCAAGATAGTAATGCAGAGACTCTTCCTCAAAAGGTAGAGCACAGCACTTTGTT
TTCAGTTACAACACCTATAATTGTGCAAGGAATAGATCAACAAGATCAGGTCTCTTCGCAGGGATTGGTCTGTAATT
TTTCAGGAGATCATTCAGAGGAGATTTTTGAGAGAGAATCCTTTTTAGGGATCGCTTTCCTAGGGAATGGTAGCAAG
GATGGAATCACGTTAACAGATATAAAATCTTCGTTATCTGGTGCTGCCTTGTATTCTTCAGATGATCTTATTTTTGA
AAGAATTAAGGGAGATATAGAGCTTTCTTCTTGTTCATCTTTAGAAAGAGGAGGAGCTTGTTCAGCTCAAAGTATTT
TAATTCATGATTGTCAAGGATTAACGGTAAAACATTGTGCCGCAGGGGTGAATGTTGAAGGAGTTAGTGCTAGCGAC
CATCTCGGATTTGGGGGCGGGGCCTTCTCTACTACAAGTTCTCTTTCTGGAGAGAAGAGTTTGTATATGCCTGCAGG
CGATATTGTGGTGGCTACCTGCGATGGTCCTGTGTGTTTCGAAGGAAATAGTGCTCAGTTAGCAAATGGTGGCGCTA
TTGCCGCTTCTGGTAAAGTTCTTTTTGTAGCTAACGAAAAAAAGATTTCCTTTACAGACAACCAAGCTTTGTCTGGA
GGAGCTATTTCTGCATCTTCTAGTATTTCTTTCCAAAATTGTGCTGAGCTTGTGTTCAAGAGTAATCTTGCAAAAGG
AGTTAAAGATAAATGTTCTTTGGGAGGAGGTGCTTTAGCCTCTTTAGAATCCGTAGTTTTGAAAGATAATCTCGGTA
TTACTTATGAAAAAAATCAGTCCTATTCGGAAGGAGGGGCTATTTTTGGGAAGGATTGTGAGATTTTTGAAAACAGG
GGGCCTGTTGTATTCAGAGATAATACAGCTGCTTTAGGAGGCGGAGCTATTTTGGCGCAACAAACTGTGGCGATTTG
TGGTAATAAGTCTGGAATATCTTTTGAAGGAAGTAAGTCTAGTTTTGGAGGGGCCATTGCTTGTGGAAATTTCTCTT
CTGAGAATAATTCTTCAGCTTTGGGATCAATTGATATCTCTAACAATCTAGGAGATATCTCTTTTCTTCGGACTCTG
TGTACTACTTCGGATTTAGGGCAAACGGATTACCAAGGGGGAGGGGCCTTATTCGCTGAAAATATTTCTCTTTCTGA
GAATGCTGGTGCAATTACTTTCAAAGACAATATTGTGAAGACATTTGCCTCAAATGGAAAAATGTTGGGTGGAGGGG
CAATTTTAGCTTCAGGAAATGTTTTGATTAGCAAAAACTCTGGAGAGATTTCTTTTGTAGGGAATGCTCGAGCTCCT
CAGGCTATTCCGACTCGTTCATCTGACGAATTGTCTTTTGGCGCACAATTAACTCAAACTACTTCAGGATGTTCTGG
AGGAGGAGCTCTTTTTGGTAAAGAGGTTGCCATTGTTCAAAATGCCACTGTTGTATTCGAGCAAAATCGCTTACAGT
GTGGCGAGCAGGAAACACATGGTGGAGGCGGTGCTGTTTATGGTATGGAGAGTGCCTCTATTATTGGAAACTCTTTT
GTGAGATTCGGAAATAATTACGCTGTAGGGAATCAGATTTCTGGAGGAGCTCTTTTATCCAAGAAGGTCCGTTTAGC
TGAAAATACAAGGGTAGATTTTTCTCGAAATATCGCTACTTTCTGCGGCGGGGCTGTTCAAGTTTCTGATGGAAGTT
GCGAATTGATCAACAATGGGTATGTGCTATTCAGAGATAACCGAGGGCAGACATTTGGTGGGGCTATTTCTTGCTTG
AAAGGAGATGTGATCATTTCCGGAAATAAAGATAGGGTTGAGTTTAGAGATAACATTGTGACGCGGCCTTATTTTGA
AGAAAATGAAGAAAAAGTTGAGACAGCAGATATTAATTCAGATAAGCAAGAAGCAGAAGAGCGCTCTTTATTAGAGA
ACATTGAGCAGAGCTTTATTACTGCAACTAATCAGACCTTTTTCTTAGAGGAAGAGAAACTCCCATCAGAAGCTTTT
ATCTCTGCTGAAGAACTTTCAAAGAGAAGAGAATGTGCTGGTGGGGCGATTTTTGCAAAACGGGTCTACATTACGGA
TAATAAAGAACCTATCTTGTTTTCGCATAATTTTTCTGATGTTTATGGGGGAGCTATTTTTACGGGTTCTCTACAGG
AAACTGATAAACAAGATGTTGTAACTCCTGAAGTTGTGATATCAGGCAACGATGGGGATGTCATTTTTTCTGGAAAT
GCAGCTAAACATGATAAGCATTTACCTGATACAGGTGGTGGAGCCATTTGTACACAGAATTTGACGATTTCCCAAAA
CAATGGGAATGTCTTGTTCTTGAACAATTTTGCTTGTTCTGGTGGAGCAGTTCGCATAGAGGATCATGGAGAAGTTC
TTTTAGAGGCTTTTGGGGGAGATATTATTTTCAATGGAAACTCTTCTTTCAGAGCTCAAGGATCGGATGCGATCTAT
TTTGCTGGTAAGGACTCTAGAATTAAAGCTTTAAATGCTACTGAAGGACATGCGATTGTGTTCCAAGATGCATTGGT
GTTTGAAAATATAGAAGAAAGAAAGTCTTCGGGACTATTGGTGATTAACTCTCAGGAAAATGAGGGTTATACGGGAT
CCGTCCGATTTTTAGGATCTGAAAGTAAGGTTCCTCAATGGATTCATGTGCAACAGGGAGGTCTTGAGTTGCTACAT
GGAGCTATTTTATGTAGTTATGGGGTTAAACAAGATCCTAGAGCTAAAATAGTATTATCTGCTGGATCTAAATTGAA
GATTCTAGATTCAGAGCAAGAAAATAACGCAGAAATTGGAGATCTTGAAGATTCTGTTAATTCAGAAAAAACACCAT
CTCTTTGGATTGGGAAGAACGCTCAAGCAAAAGTCCCTCTGGTTGATATCCATACTATTTCTATTGATTTAGCATCA
TTTTCTTCTAAAGCTCAGGAAACCCCTGAGGAAGCTCCACAAGTCATCGTCCCTAAGGGAAGTTGTGTCCACTCGGG
AGAGTTAAGTTTGGAGTTGGTTAATACAACAGGAAAAGGTTATGAGAATCATGCGTTGTTAAAAAATGATACTCAGG
TTTCTCTCATGTCTTTCAAAGAGGAAAATGATGGATCTTTAGAAGATTTGAGTAAGTTGTCTGTTTCGGATTTACGC
ATTAAAGTTTCTACTCCAGATATTGTAGAAGAAACTTATGGCCATATGGGGGATTGGTCTGAAGCTACAATTCAAGA
TGGGGCTCTTGTCATTAATTGGCATCCTACTGGATATAAATTAGATCCGCAAAAAGCTGGTTCTTTGGTATTCAATG
CATTATGGGAGGAAGAGGCTGTATTGTCTACTCTAAAAAATGCTCGGATTGCCCATAACCTTACCATTCAGAGAATG
GAATTTGATTATTCTACAAATGCTTGGGGATTAGCTTTTAGTAGCTTTAGAGAGCTATCTTCAGAGAAGCTTGTTTC
TGTTGATGGATATAGAGGCTCTTATATAGGGGCTTCTGCAGGCATTGATACTCAGTTGATGGAAGATTTTGTTTTGG
GAATCAGCACGGCTTCCTTCTTCGGGAAAATGCATAGTCAGAATTTTGATGCAGAGATTTCTCGACATGGTTTTGTT
GGTTCGGTCTATACAGGCTTCCTAGCTGGGGCCTGGTTCTTCAAGGGGCAGTACAGTCTTGGCGAAACACATAACGA
TATGACAACTCGTTACGGGGTTTTGGGAGAATCTAATGCTACTTGGAAGTCTCGAGGAGTACTAGCAGATGCTTTAG
TTGAATATCGTAGTTTAGTCGGTCCAGCACGACCTAAATTTTATGCTTTGCATTTTAATCCTTATGTCGAGGTATCT

TATGCATCTGCGAAGTTCCCTAGTTTTGTAGAACAAGGAGGAGAAGCTCGTGCTTTTGAAGAAACCTCTTTAACAAA
CATTACCGTTCCCTTTGGTATGAAATTTGAACTATCTTTTACAAAAGGACAGTTTTCAGAGACTAATTCTCTTGGAA
TAGGTTGTGCATGGGAAATGTATCGGAAAGTCGAAGGAAGATCTGTAGAGCTACTAGAAGCTGGTTTTGATTGGGAA
GGATCTCCTATAGATCTCCCTAAACAAGAGCTGAGAGTGGCTTTAGAAAACAATACGGAATGGAGTTCGTATTTTAG
TACAGCTCTAGGAGTAACAGCATTTTGTGGAGGATTTTCTTCTATGGATAATAAACTAGGATACGAAGCGAATGCTG
GAATGCGTTTGATTTTCTAG

SEQ ID NO: 60 - TC0197 protein sequence

MSSEKDKKNSCSKFSLSVVAAILASMSGLSNCSDLYAVGSSADHPAYLIPQAGLLLDHIKDIFIGPKDSQDKGQYKL
IIGEAGSFQDSNAETLPQKVEHSTLFSVTTPIIVQGIDQQDQVSSQGLVCNFSGDHSEEIFERESFLGIAFLGNGSK
DGITLTDIKSSLSGAALYSSDDLIFERIKGDIELSSCSSLERGGACSAQSILIHDCQGLTVKHCAAGVNVEGVSASD
HLGFGGGAFSTTSSLSGEKSLYMPAGDIVVATCDGPVCFEGNSAQLANGGAIAASGKVLFVANEKKISFTDNQALSG
GAISASSSISFQNCAELVFKSNLAKGVKDKCSLGGGALASLESVVLKDNLGITYEKNQSYSEGGAIFGKDCEIFENR
GPVVFRDNTAALGGGAILAQQTVAICGNKSGISFEGSKSSFGGAIACGNFSSENNSSALGSIDISNNLGDISFLRTL
CTTSDLGQTDYQGGGALFAENISLSENAGAITFKDNIVKTFASNGKMLGGGAILASGNVLISKNSGEISFVGNARAP
QAIPTRSSDELSFGAQLTQTTSGCSGGGALFGKEVAIVQNATVVFEQNRLQCGEQETHGGGGAVYGMESASIIGNSF
VRFGNNYAVGNQISGGALLSKKVRLAENTRVDFSRNIATFCGGAVQVSDGSCELINNGYVLFRDNRGQTFGGAISCL
KGDVIISGNKDRVEFRDNIVTRPYFEENEEKVETADINSDKQEAEERSLLENIEQSFITATNQTFFLEEEKLPSEAF
ISAEELSKRRECAGGAIFAKRVYITDNKEPILFSHNFSDVYGGAIFTGSLQETDKQDVVTPEVVISGNDGDVIFSGN
AAKHDKHLPDTGGGAICTQNLTISQNNGNVLFLNNFACSGGAVRIEDHGEVLLEAFGGDIIFNGNSSFRAQGSDAIY
FAGKDSRIKALNATEGHAIVFQDALVFENIEERKSSGLLVINSQENEGYTGSVRFLGSESKVPQWIHVQQGGLELLH
GAILCSYGVKQDPRAKIVLSAGSKLKILDSEQENNAEIGDLEDSVNSEKTPSLWIGKNAQAKVPLVDIHTISIDLAS
FSSKAQETPEEAPQVIVPKGSCVHSGELSLELVNTTGKGYENHALLKNDTQVSLMSFKEENDGSLEDLSKLSVSDLR
IKVSTPDIVEETYGHMGDWSEATIQDGALVINWHPTGYKLDPQKAGSLVFNALWEEEAVLSTLKNARIAHNLTIQRM
EFDYSTNAWGLAFSSFRELSSEKLVSVDGYRGSYIGASAGIDTQLMEDFVLGISTASFFGKMHSQNFDAEISRHGFV
GSVYTGFLAGAWFFKGQYSLGETHNDMTTRYGVLGESNATWKSRGVLADALVEYRSLVGPARPKFYALHFNPYVEVS
YASAKFPSFVEQGGEARAFEETSLTNITVPFGMKFELSFTKGQFSETNSLGIGCAWEMYRKVEGRSVELLEAGFDWE
GSPIDLPKQELRVALENNTEWSSYFSTALGVTAFCGGFSSMDNKLGYEANAGMRLIF

SEQ ID NO: 61 - TC0261 nucleotide sequence

ATGAAAAAACTGTTCTTTTTTGTCCTTATTGGAAGCTCTATACTGGGATTTACTCGAGAAGTCCCTCCTTCGATTCT
TTTAAAGCCTATACTAAATCCATACCATATGACCGGGTTATTTTTTCCCAAGGTTAATTTGCTTGGAGACACACATA
ATCTCACTGATTACCATTTGGATAATCTAAAATGCATTCTGGCTTGCCTACAAAGAACTCCTTATGAAGGAGCTGCT
TTCACAGTAACCGATTACTTAGGTTTTTCAGATACACAAAAGGATGGTATTTTTTGTTTTAAAAATCTTACTCCAGA
GAGTGGAGGGGTTATTGGTTCCCCAACTCAAAACACTCCTACTATAAAAATTCATAATACAATCGGCCCCGTTCTTT
TCGAAAATAATACCTGTCATAGACTGTGGACACAGACCGATCCCGAAAATGAAGGAAACAAAGCACGCGAAGGCGGG
GCAATTCATGCTGGGGACGTTTACATAAGCAATAACCAGAACCTTGTCGGATTCATAAAGAACTTTGCTTATGTTCA
AGGTGGAGCTATTAGTGCTAATACTTTTGCCTATAAAGAAAATAAATCGAGCTTTCTTTGCCTAAATAACTCTTGTA
TACAAACTAAGACGGGAGGGAAAGGTGGTGCTATTTACGTTAGTACGAGCTGCTCTTTCGAGAACAATAACAAGGAT
CTGCTTTTCATCCAAAACTCCGGCTGTGCAGGAGGAGCTATCTTCTCTCCAACCTGTTCTCTAATAGGAAACCAAGG
AGATATTGTTTTTTACAGCAACCACGGTTTTAAAAATGTTGATAATGCAACTAACGAATCTGGGGATGGAGGAGCTA
TTAAAGTAACTACCCGCTTGGACATCACCAATAATGGTAGTCAAATCTTTTTTTCTGATAATATCTCAAGAAATTTT
GGAGGAGCTATTCATGCTCCTTGTCTTCATCTTGTTGGTAATGGGCCAACCTATTTTACAAACAATATAGCTAATCA
CACAGGTGGGGCTATTTATATAACAGGAACAGAAACCTCAAAGATTTCTGCAGATCACCATGCTATTATTTTTGATA
ATAACATTTCTGCAAACGCCACCAATGCGGACGGATCTAGCAGCAACACTAATCCTCCTCACAGAAATGCGATCACT
ATGGACAATTCCGCTGGAGGAATAGAACTTGGTGCAGGGAAGAGCCAGAATCTTATTTTCTATGATCCTATTCAAGT
GACGAATGCTGGAGTTACCGTAGACTTCAATAAGGATGCCTCCCAAACCGGATGTGTAGTTTTCTCTGGAGCGACTG
TCCTTTCTGCAGATATTTCTCAGGCTAATTTGCAAACTAAAACACCTGCAACGCTTACTCTCAGTCACGGTCTTCTG
TGTATCGAAGATCGTGCTCAGCTCACAGTGAACAATTTTACACAAACAGGAGGGATTGTAGCCTTAGGAAATGGAGC
AGTTTTAAGCAGCTACCAACACAGCACTACAGACGCCACTCAAACTCCCCCTACAACCACCACTACAGATGCTTCCG
TAACTCTTAATCACATTGGATTAAATCTCCCCTCTATTCTTAAGGATGGAGCAGAGATGCCTCTATTATGGGTAGAA
CCTATAAGCACAACTCAAGGTAACACTACAACATATACGTCAGATACCGCGGCTTCCTTCTCATTAAATGGAGCCAC
ACTCTCTCTCATTGATGAAGATGGAAATTCTCCCTATGAAAACACGGACCTCTCTCGTGCATTGTACGCTCAACCTA
TGCTAGCAATTTCTGAGGCCAGTGATAACCAATTGCAATCCGAAAGCATGGACTTTTCTAAAGTTAATGTTCCTCAC
TATGGATGGCAAGGACTTTGGACCTGGGGGTGGGCAAAAACTGAAAATCCAACAACAACTCCTCCAGCAACAATTAC
TGATCCGAAAAAAGCTAATCAGTTTCATAGAACTTTATTATTAACGTGGCTCCCTGCTGGTTATATCCCCAGCCCTA
AACATAAAAGCCCTTTAATAGCTAATACCTTGTGGGGGAATATACTTTTTGCAACGGAAAACTTAAAAAATAGCTCA
GGGCAAGAACTTCTTGATCGTCCTTTCTGGGGAATTACAGGAGGGGGCTTGGGGATGATGGTCTATCAAGAACCTAG
AAAAGACCATCCTGGATTCCACATGCATACCTCCGGATATTCAGCAGGAATGATTACAGGAAACACACATACCTTCT
CATTACGATTCAGCCAGTCCTATACAAAACTCAATGAACGTTATGCCAAGAACTATGTGTCTTCTAAAAATTACTCT
TGCCAAGGGGAAATGCTTTTGTCCTTACAAGAAGGACTCATGCTGACTAAACTAATTGGTCTCTATAGTTATGGGAA
TCACAACAGCCACCATTTCTATACCCAAGGAGAAGACCTATCGTCTCAAGGGGAGTTCCATAGTCAGACTTTTGGAG
GGGCTGTCTTTTTTGATCTACCTCTGAAACCTTTTGGAAGAACACACATACTTACAGCTCCTTTCTTAGGTGCCATT
GGTATGTATTCTAAGCTGTCTAGCTTTACAGAAGTAGGAGCCTATCCAAGAACCTTTATTACAGAAACGCCTTTAAT
CAATGTCCTGATTCCTATCGGAGTAAAAGGTAGCTTCATGAATGCCACCCATAGACCTCAGGCCTGGACTGTAGAGC
TTGCTTACCAACCTGTTCTTTACAGACAAGAACCTAGTATCTCTACCCAATTACTCGCTGGTAAAGGTATGTGGTTT
GGGCATGGAAGTCCTGCATCTCGCCACGCTCTAGCTTATAAAATTTCACAGAAAACACAGCTTTTGCGATTTGCAAC

ACTTCAACTCCAGTATCACGGATACTATTCGTCTTCCACTTTCTGTAATTATCTGAATGGAGAGGTATCTTTACGTT
TCTAA

SEQ ID NO: 62 - TC0261 protein sequence

MKKLFFFVLIGSSILGFTREVPPSILLKPILNPYHMTGLFFPKVNLLGDTHNLTDYHLDNLKCILACLQRTPYEGAA
FTVTDYLGFSDTQKDGIFCFKNLTPESGGVIGSPTQNTPTIKIHNTIGPVLFENNTCHRLWTQTDPENEGNKAREGG
AIHAGDVYISNNQNLVGFIKNFAYVQGGAISANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIYVSTSCSFENNNKD
LLFIQNSGCAGGAIFSPTCSLIGNQGDIVFYSNHGFKNVDNATNESGDGGAIKVTTRLDITNNGSQIFFSDNISRNF
GGAIHAPCLHLVGNGPTYFTNNIANHTGGAIYITGTETSKISADHHAIIFDNNISANATNADGSSSNTNPPHRNAIT
MDNSAGGIELGAGKSQNLIFYDPIQVTNAGVTVDFNKDASQTGCVVFSGATVLSADISQANLQTKTPATLTLSHGLL
CIEDRAQLTVNNFTQTGGIVALGNGAVLSSYQHSTTDATQTPPTTTTTDASVTLNHIGLNLPSILKDGAEMPLLWVE
PISTTQGNTTTYTSDTAASFSLNGATLSLIDEDGNSPYENTDLSRALYAQPMLAISEASDNQLQSESMDFSKVNVPH
YGWQGLWTWGWAKTENPTTTPPATITDPKKANQFHRTLLLTWLPAGYIPSPKHKSPLIANTLWGNILFATENLKNSS
GQELLDRPFWGITGGGLGMMVYQEPRKDHPGFHMHTSGYSAGMITGNTHTFSLRFSQSYTKLNERYAKNYVSSKNYS
CQGEMLLSLQEGLMLTKLIGLYSYGNHNSHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHILTAPFLGAI
GMYSKLSSFTEVGAYPRTFITETPLINVLIPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPSISTQLLAGKGMWF
GHGSPASRHALAYKISQKTQLLRFATLQLQYHGYYSSSTFCNYLNGEVSLRF

SEQ ID NO: 63 - CT733 fragment nucleotide sequence

GCACCTCAACCTCGCGGAACGCTTCCTAGCTCGACCACAAAAATTGGATCAGAAGTTTGGATTGAACAAAAAGTCCG
CCAATATCCAGAGCTTTTATGGTTAGTAGAGCCGTCCTCTACGGGAGCCTCTTTAAAATCTCCTTCAGGAGCCATCT
TTTCTCCAACATTATTCCAAAAAAAGGTCCCTGCTTTCGATATCGCAGTGCGCAGTTTGATTCACTTACATTTATTA
ATCCAGGGTTCCCGCCAAGCCTATGCTCAACTGATCCAACTACAGACCAGCGAATCCCCTCTAACATTTAAGCAATT
CCTTGCATTGCATAAGCAATTAACTCTATTTTTAAATTCCCCTAAGGAATTTTATGACTCTGTTAAAGTGTTAGAGA
CAGCTATCGTCTTACGTCACTTAGGCTGTTCAACTAAGGCTGTTGCTGCGTTTAAACCTTATTTCTCAGAAATGCAA
AGAGAGGCTTTTTACACTAAGGCTCTGCATGTACTACACACCTTCCCAGAGCTAAGCCCATCATTTGCTCGCCTCTC
TCCGGAGCAGAAAACTCTCTTCTTCTCCTTGAGAAAATTGGCGAATTACGATGAGTTACTCTCGCTGACGAACACCC
CAAGTTTTCAGCTTCTGTCTGCTGGGCGCTCGCAACGAGCTCTTTTAGCTCTGGACTTGTACCTCTATGCTTTGGAT
TCCTGTGGAGAACAGGGGATGTCCTCTCAATTCCACACAAACTTCGCACCTCTACAGTCCATGTTGCAACAATACGC
TACTGTAGAAGAGGCCTTTTCTCGTTATTTTACTTACCGAGCTAATCGATTAGGATTTGATGGCTCTTCTCGATCCG
AGATGGCTTTAGTAAGAATGGCCACCTTGATGAACTTGTCTCCTTCCGAAGCTGCGATTTTAACCACAAGCTTCAAA
ACCCTTCCTACAGAAGAAGCGGATACTTTGATCAATAGTTTCTATACCAATAAGGGCGATTCGTTGGCTCTTTCTCT
GCGAGGGTTGCCTACACTTGTATCCGAACTGACGCGAACTGCCCATGGCAATACCAATGCAGAAGCTCGATCTCAGC
AAATTTATGCAACTACCCTATCGCTAGTAGTAAAGAGTCTGAAAGCGCACAAAGAAATGCTAAACAAGCAAATTCTT
TCTAAGGAAATTGTTTTAGATTTCTCAGAAACTGCAGCTTCTTGCCAAGGATTGGATATCTTTTCCGAGAATGTCGC
TGTTCAAATTCACTTAAATGGAACCGTTAGTATCCATTTA

SEQ ID NO: 64 - CT733 fragment protein sequence

APQPRGTLPSSTTKIGSEVWIEQKVRQYPELLWLVEPSSTGASLKSPSGAIFSPTLFQKKVPAFDIAVRSLIHLHLL
IQGSRQAYAQLIQLQTSESPLTFKQFLALHKQLTLFLNSPKEFYDSVKVLETAIVLRHLGCSTKAVAAFKPYFSEMQ
REAFYTKALHVLHTFPELSPSFARLSPEQKTLFFSLRKLANYDELLSLTNTPSFQLLSAGRSQRALLALDLYLYALD
SCGEQGMSSQFHTNFAPLQSMLQQYATVEEAFSRYFTYRANRLGFDGSSRSEMALVRMATLMNLSPSEAAILTTSFK
TLPTEEADTLINSFYTNKGDSLALSLRGLPTLVSELTRTAHGNTNAEARSQQIYATTLSLVVKSLKAHKEMLNKQIL
SKEIVLDFSETAASCQGLDIFSENVAVQIHLNGTVSIHL

SEQ ID NO:65 – CT153 fragment nucleotide sequence

ACTAAGCCTTCTTTCTTATACGTTATTCAACCTTTTTCCGTATTTAATCCACGATTAGGACGTTTCTCTACAGACTC
AGATACTTATATCGAAGAAGAAAACCGCCTAGCATCGTTCATTGAGAGTTTGCCACTGGAGATCTTCGATATACCTT
CTTTCATGGAAACCGCGATTTCCAATAGCCCCTATATTTTATCTTGGGAGACAACTAAAGACGGCGCTCTGTTCACT
ATTCTTGAACCCAAACTCTCAGCTTGCGCAGCCACTTGCCTGGTAGCCCCTTCTATACAAATGAAATCCGATGCGGA
GCTCCTAGAAGAAATTAAGCAAGCGTTATTACGCAGCTCTCATGACGGTGTGAAATATCGCATCACCAGAGAATCCT
TCTCTCCAGAAAAGAAAACTCCTAAGGTTGCTCTAGTCGATGACGATATTGAATTGATTCGCAATGTCGACTTTTTG
GGTAGAGCTGTTGACATTGTCAAATTAGACCCTATTAATATTCTGAATACCGTAAGCGAAGAGAATATTCTAGATTA
CTCTTTTACAAGAGAAACGGCTCAGCTGAGCGCGGATGGTCGTTTTGGTATTCCTCCAGGGACTAAGCTATTCCCTA
AACCTTCTTTTGATGTAGAAATCAGTACCTCCATTTTCGAAGAAACAACTTCATTTACTCGAAGTTTTTCTGCATCG
GTTACTTTTAGTGTACCAGACCTCGCGGCGACTATGCCTCTTCAAAGCCCTCCCATGGTAGAAAATGGTCAAAAAGA
AATTTGTGTCATTCAAAAACACTTATTCCCAAGCTACTCTCCTAAACTAGTCGATATTGTTAAACGATACAAAAGAG
AGGCTAAGATCTTGATTAACAAGCTTGCCTTTGGAATGTTATGGCGACATCGGGCTAAAAGCCAAATCCTCACCGAG
GGAAGCGTACGTCTAGACTTACAAGGATTCACAGAATCGAAGTACAATTACCAGATTCAAGTAGGATCCCATACGAT
TGCAGCTGTATTAATCGATATGGATATTTCCAAGATTCAATCCAAATCAGAACAAGCTTATGCAATTAGGAAAATCA
AATCAGGCTTTCAACGTAGCTTGGATGACTATCATATTTATCAAATTGAAAGAAAACAAACCTTTTCTTTTTCTCCG
AAGCATCGCAGCCTCTCATCCACATCCCATTCCGAAGATTCTGATTTGGATCTTTCTGAAGCAGCCGCCTTTTCAGG
AAGTCTTACCTGCGAGTTTGTAAAAAAAAGCACTCAACATGCCAAGAATACCGTCACATGTTCCACAGCCGCTCATT
CCCTATACACACTCAAAGAAGATGACAGCTCGAACCCCTCTGAAAAACGATTAGATAGTTGTTTCCGCAATTGGATT
GAAAACAAACTAAGCGCCAATTCTCCAGATTCCTGGTCAGCGTTTATTCAAAAATTCGGAACACACTATATTGCATC
AGCAACTTTTGGAGGGATAGGTTTCCAAGTGCTCAAACTATCTTTTGAACAGGTGGAGGATCTACATAGCAAAAAGA

TCTCCTTAGAAACCGCAGCAGCCAACTCTCTATTAAAAGGTTCTGTATCCAGCAGCACAGAATCTGGATACTCCAGC
TATAGCTCCACGTCTTCTTCTCATACGGTATTTTTAGGAGGAACGGTCTTACCTTCGGTTCATGATGAACGTTTAGA
CTTTAAAGATTGGTCGGAAAGTGTGCACCTGGAACCTGTTCCTATCCAGGTTTCTTTACAACCTATAACGAATTTAC
TAGTTCCTCTCCATTTTCCTAATATCGGTGCTGCAGAGCTCTCTAATAAACGAGAATCTCTTCAACAAGCGATTCGA
GTCTATCTCAAAGAACATAAAGTAGATGAGCAAGGAGAACGTACTACATTTACATCAGGAATCGATAATCCTTCTTC
CTGGTTTACCTTAGAAGCTGCCCACTCTCCTCTTATAGTCAGTACTCCTTACATTGCTTCGTGGTCTACGCTTCCTT
ATTTGTTCCCAACATTAAGAGAACGTTCTTCGGCAACCCCTATCGTTTTCTATTTTTGTGTAGATAATAATGAACAT
GCTTCGCAAAAAATATTAAACCAATCGTATTGCTTCCTCGGGTCCTTGCCTATTCGACAAAAAATTTTTGGTAGCGA
ATTTGCTAGTTTCCCCTATCTATCTTTCTATGGAAATGCAAAAGAGGCGTACTTTGATAACACGTACTACCCAACGC
GTTGTGGGTGGATTGTTGAAAAGTTAAATACTACACAAGATCAATTCCTCCGGGATGGAGACGAGGTGCGACTAAAA
CATGTTTCCAGCGGAAAGTATCTAGCAACAACTCCTCTTAAGGATACCCATGGTACACTCACGCGTACAACGAACTG
TGAAGATGCTATCTTTATTATTAAAAAATCTTCAGGTTAT

SEQ ID NO:66 - CT153 fragment protein sequence

TKPSFLYVIQPFSVFNPRLGRFSTDSDTYIEEENRLASFIESLPLEIFDIPSFMETAISNSPYILSWETTKDGALFT
ILEPKLSACAATCLVAPSIQMKSDAELLEEIKQALLRSSHDGVKYRITRESFSPEKKTPKVALVDDDIELIRNVDFL
GRAVDIVKLDPINILNTVSEENILDYSFTRETAQLSADGRFGIPPGTKLFPKPSFDVEISTSIFEETTSFTRSFSAS
VTFSVPDLAATMPLQSPPMVENGQKEICVIQKHLFPSYSPKLVDIVKRYKREAKILINKLAFGMLWRHRAKSQILTE
GSVRLDLQGFTESKYNYQIQVGSHTIAAVLIDMDISKIQSKSEQAYAIRKIKSGFQRSLDDYHIYQIERKQTFSFSP
KHRSLSSTSHSEDSDLDLSEAAAFSGSLTCEFVKKSTQHAKNTVTCSTAAHSLYTLKEDDSSNPSEKRLDSCFRNWI
ENKLSANSPDSWSAFIQKFGTHYIASATFGGIGFQVLKLSFEQVEDLHSKKISLETAAANSLLKGSVSSSTESGYSS
YSSTSSSHTVFLGGTVLPSVHDERLDFKDWSESVHLEPVPIQVSLQPITNLLVPLHFPNIGAAELSNKRESLQQAIR
VYLKEHKVDEQGERTTFTSGIDNPSSWFTLEAAHSPLIVSTPYIASWSTLPYLFPTLRERSSATPIVFYFCVDNNEH
ASQKILNQSYCFLGSLPIRQKIFGSEFASFPYLSFYGNAKEAYFDNTYYPTRCGWIVEKLNTTQDQFLRDGDEVRLK
HVSSGKYLATTPLKDTHGTLTRTTNCEDAIFIIKKSSGY

SEQ ID NO:67 - CT601 fragment nucleotide sequence

GGTAAAGCACCGTCTTTGCAGGCTATTCTAGCCGAAGTCGAAGACACCTCCTCTCGTCTACACGCTCATCACAATGA
GCTTGCTATGATCTCTGAACGCCTCGATGAGCAAGACACGAAACTACAGCAACTTTCGTCAACACAAGATCATAACC
TACCTCGACAAGTTCAGCGACTAGAAACGGACCAAAAAGCTTTGGCAAAAACACTGGCGATTCTTTCGCAATCCGTC
CAAGATATTCGGTCTTCTGTACAAAATAAATTACAAGAAATCCAACAAGAACAAAAAAAAATTAGCACAAAATTTGCG
AGCGCTTCGTAACTCTTTACAAGCTCTCGTTGATGGCTCTTCTCCAGAAAATTATATTGATTTCCTAACTGGTGAAA
CCCCGGAACATATTCATATTGTTAAACAAGGAGAGACCCTGAGCAAGATCGCGAGTAAATATAACATCCCCGTCGTA
GAATTAAAAAAACTTAATAAACTAAATTCGGATACTATTTTTACAGATCAAAGAATTCGCCTTCCGAAAAAGAAA

SEQ ID NO:68 - CT601 fragment protein sequence

GKAPSLQAILAEVEDTSSRLHAHHNELAMISERLDEQDTKLQQLSSTQDHNLPRQVQRLETDQKALAKTLAILSQSV
QDIRSSVQNKLQEIQQEQKKLAQNLRALRNSLQALVDGSSPENYIDFLTGETPEHIHIVKQGETLSKIASKYNIPVV
ELKKLNKLNSDTIFTDQRIRLPKKK

SEQ ID NO:69 - CT279 fragment nucleotide sequence

GCACAAGTAATTTCTTCCGATAACACATTCCAAGTCTATGAAAAGGGAGATTGGCACCCAGCCCTATATAATACTAA
AAAGCAGTTGCTAGAGATCTCCTCTACTCCTCCTAAAGTAACCGTGACAACTTTAAGCTCATATTTTCAAAACTTTG
TTAGAGTCTTGCTTACAGATACACAAGGAAATCTTTCTTCATTCGAAGACCATAATCTCAATCTAGAAGAATTTTTA
TCTCAACCAACTCCTGTAATACATGGTCTTGCCCTTTATGTGGTCTACGCTATCCTACACAACGATGCAGCTTCCTC
TAAATTATCTGCTTCCCAAGTAGCGAAAAATCCAACAGCTATAGAATCTATAGTTCTTCCTATAGAAGGTTTTGGTT
TGTGGGGACCTATCTATGGATTCCTTGCTCTAGAAAAAGACGGGAATACTGTTCTTGGTACTTCTTGGTATCAACAT
GGCGAGACTCCTGGATTAGGAGCAAATATCGCTAACCCTCAATGGCAAAAAAATTTCAGAGGCAAAAAAGTATTTCT
AGTCTCAGCTTCTGGAGAAACAGATTTTGCTAAGACAACCCTAGGACTGGAAGTTATAAAAGGATCTGTATCTGCAG
CATTAGGAGACTCACCTAAAGCTGCTTCTTCCATCGACGGAATTTCAGGAGCTACTTTGACTTGTAATGGTGTTACC
GAATCCTTCTCTCATTCTCTAGCTCCCTACCGCGCTTTGTTGACTTTCTTCGCCAACTCTAAACCTAGTGGAGAGTC
TCATGACCAC

SEQ ID NO:70 - CT279 fragment protein sequence

AQVISSDNTFQVYEKGDWHPALYNTKKQLLEISSTPPKVTVTTLSSYFQNFVRVLLTDTQGNLSSFEDHNLNLEEFL
SQPTPVIHGLALYVVYAILHNDAASSKLSASQVAKNPTAIESIVLPIEGFGLWGPIYGFLALEKDGNTVLGTSWYQH
GETPGLGANIANPQWQKNFRGKKVFLVSASGETDFAKTTLGLEVIKGSVSAALGDSPKAASSIDGISGATLTCNGVT
ESFSHSLAPYRALLTFFANSKPSGESHDH

SEQ ID NO:71 - CT443 fragment nucleotide sequence

GGGGTGTTAGAGACCTCTATGGCAGAGTCTCTCTCTACAAACGTTATTAGCTTAGCTGACACCAAAGCGAAAGACAA
CACTTCTCATAAAAGCAAAAAAGCAAGAAAAAACCACAGCAAAGAGACTCCCGTAGACCGTAAAGAGGTTGCTCCGG
TTCATGAGTCTAAAGCTACAGGACCTAAACAGGATTCTTGCTTTGGCAGAATGTATACAGTCAAAGTTAATGATGAT
CGCAATGTTGAAATCACACAAGCTGTTCCTGAATATGCTACGGTAGGATCTCCCTATCCTATTGAAATTACTGCTAC
AGGTAAAAGGGATTGTGTTGATGTTATCATTACTCAGCAATTACCATGTGAAGCAGAGTTCGTACGCAGTGATCCAG
CGACAACTCCTACTGCTGATGGTAAGCTAGTTTGGAAAATTGACCGCTTAGGACAAGGCGAAAAGAGTAAAATTACT

GTATGGGTAAAACCTCTTAAAGAAGGTTGCTGCTTTACAGCTGCAACAGTATGCGCTTGTCCAGAGATCCGTTCGGT
TACAAAATGTGGACAACCTGCTATCTGTGTTAAACAAGAAGGCCCAGAGAATGCTTGTTTGCGTTGCCCAGTAGTTT
ACAAAATTAATATAGTGAACCAAGGAACAGCAACAGCTCGTAACGTTGTTGTTGAAAATCCTGTTCCAGATGGTTAC
GCTCATTCTTCTGGACAGCGTGTACTGACGTTTACTCTTGGAGATATGCAACCTGGAGAGCACAGAACAATTACTGT
AGAGTTTTGTCCGCTTAAACGTGGTCGTGCTACCAATATAGCAACGGTTTCTTACTGTGGAGGACATAAAAATACAG
CAAGCGTAACAACTGTGATCAACGAGCCTTGCGTACAAGTAAGTATTGCAGGAGCAGATTGGTCTTATGTTTGTAAG
CCTGTAGAATATGTGATCTCCGTTTCCAATCCTGGAGATCTTGTGTTGCGAGATGTCGTCGTTGAAGACACTCTTTC
TCCCGGAGTCACAGTTCTTGAAGCTGCAGGAGCTCAAATTTCTTGTAATAAAGTAGTTTGGACTGTGAAAGAACTGA
ATCCTGGAGAGTCTCTACAGTATAAAGTTCTAGTAAGAGCACAAACTCCTGGACAATTCACAAATAATGTTGTTGTG
AAGAGCTGCTCTGACTGTGGTACTTGTACTTCTTGCGCAGAAGCGACAACTTACTGGAAAGGAGTTGCTGCTACTCA
TATGTGCGTAGTAGATACTTGTGACCCTGTTTGTGTAGGAGAAAATACTGTTTACCGTATTTGTGTCACCAACAGAG
GTTCTGCAGAAGATACAAATGTTTCTTTAATGCTTAAATTCTCTAAAGAACTGCAACCTGTATCCTTCTCTGGACCA
ACTAAAGGAACGATTACAGGCAATACAGTAGTATTCGATTCGTTACCTAGATTAGGTTCTAAAGAAACTGTAGAGTT
TTCTGTAACATTGAAAGCAGTATCAGCTGGAGATGCTCGTGGGGAAGCGATTCTTTCTTCCGATACATTGACTGTTC
CAGTTTCTGATACAGAGAATACACACATCTAT

SEQ ID NO:72 - CT443 fragment protein sequence

GVLETSMAESLSTNVISLADTKAKDNTSHKSKKARKNHSKETPVDRKEVAPVHESKATGPKQDSCFGRMYTVKVNDD
RNVEITQAVPEYATVGSPYPIEITATGKRDCVDVIITQQLPCEAEFVRSDPATTPTADGKLVWKIDRLGQGEKSKIT
VWVKPLKEGCCFTAATVCACPEIRSVTKCGQPAICVKQEGPENACLRCPVVYKINIVNQGTATARNVVVENPVPDGY
AHSSGQRVLTFTLGDMQPGEHRTITVEFCPLKRGRATNIATVSYCGGHKNTASVTTVINEPCVQVSIAGADWSYVCK
PVEYVISVSNPGDLVLRDVVVEDTLSPGVTVLEAAGAQISCNKVVWTVKELNPGESLQYKVLVRAQTPGQFTNNVVV
KSCSDCGTCTSCAEATTYWKGVAATHMCVVDTCDPVCVGENTVYRICVTNRGSAEDTNVSLMLKFSKELQPVSFSGP
TKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAILSSDTLTVPVSDTENTHIY

SEQ ID NO:73 - CT372 fragment nucleotide sequence

CAGGCTGCACACCATCACTATCACCGCTACACAGATAAACTGCACAGACAAAACCATAAAAAAGATCTCATCTCTCC
CAAACCTACCGAACAAGAGGCGTGCAATACTTCTTCCCTTAGTAAGGAATTAATCCCTCTATCAGAACAAAGAGGCC
TTTTATCCCCCATCTGTGACTTTATTTCGGAACGCCCTTGCTTACACGGAGTTTCTGTTAGAAATCTCAAGCAAGCG
CTAAAAAATTCTGCAGGAACCCAAATTGCACTGGATTGGTCTATTCTCCCTCAATGGTTCAATCCTCGGGTCTCTCA
TGCCCCTAAGCTTTCTATCCGAGACTTTGGGTATAGCGCACACCAAACTGTTACCGAAGCCACTCCTCCTTGCTGGC
AAAACTGCTTTAATCCATCTGCGGCCGTTACTATCTATGATTCCTCATATGGGAAAGGGGTCTTTCAAATATCCTAT
ACCCTTGTCCGCTATTGGAGAGAGAATGCTGCGACTGCTGGCGATGCTATGATGCTCGCAGGGAGTATCAATGATTA
TCCCTCTCGTCAGAACATTTTCTCTCAGTTTACTTTCTCCCAAAACTTCCCAAATGAACGGGTGAGTCTGACAATTG
GTCAGTACTCACTCTATGCAATAGACGGAACATTATACAATAACGATCAACAACTTGGATTCATTAGTTACGCATTA
TCACAAAATCCAACAGCAACTTATTCCTCTGGAAGTCTTGGAGCTTACCTACAAGTCGCTCCTACCGCAAGCACAAG
TCTTCAAATAGGATTTCAAGACGCTTATAATATCTCCGGATCCTCTATCAAATGGAGTAACCTTACAAAAAATAGAT
ACAATTTTCACGGTTTTGCTTCCTGGGCTCCCCGCTGTTGCTTAGGATCTGGCCAGTACTCCGTGCTTCTTTATGTG
ACTAGACAAGTTCCAGAACAGATGGAACAAACAATGGGATGGTCAGTCAATGCGAGTCAACACATATCTTCTAAACT
GTATGTGTTTGGAAGATACAGCGGTGTTACAGGACATGTGTTCCCGATTAACCGCACGTATTCATTCGGTATGGCCT
CTGCAAATTTATTTAACCGTAACCCACAAGATTTATTTGGAATTGCTTGCGCATTCAATAATGTACACCTCTCTGCT
TCTCCAAATACTAAAAGAAAATACGAAACTGTAATCGAAGGGTTTGCAACTATCGGTTGCGGCCCCTATCTTTCTTT
CGCTCCAGACTTCCAACTCTACCTCTACCCAGCTCTTCGTCCAAACAAACAATCTGCCCGTGTTTATAGCGTGCGAG
CTAATTTAGCTATC

SEQ ID NO:74 - CT372 fragment protein sequence

QAAHHHYHRYTDKLHRQNHKKDLISPKPTEQEACNTSSLSKELIPLSEQRGLLSPICDFISERPCLHGVSVRNLKQA
LKNSAGTQIALDWSILPQWFNPRVSHAPKLSIRDFGYSAHQTVTEATPPCWQNCFNPSAAVTIYDSSYGKGVFQISY
TLVRYWRENAATAGDAMMLAGSINDYPSRQNIFSQFTFSQNFPNERVSLTIGQYSLYAIDGTLYNNDQQLGFISYAL
SQNPTATYSSGSLGAYLQVAPTASTSLQIGFQDAYNISGSSIKWSNLTKNRYNFHGFASWAPRCCLGSGQYSVLLYV
TRQVPEQMEQTMGWSVNASQHISSKLYVFGRYSGVTGHVFPINRTYSFGMASANLFNRNPQDLFGIACAFNNVHLSA
SPNTKRKYETVIEGFATIGCGPYLSFAPDFQLYLYPALRPNKQSARVYSVRANLAI

SEQ ID NO:75 - CT456 fragment nucleotide sequence

ACAAATTCAGCGGCTACATCTTCTATCCAAACGACTGGAGAGACTGTAGTAAACTATACGAATTCAGCCTCCGCCCC
CAATGTAACTGTATCGACCTCCTCTTCTTCCACACAAGCCACAGCCACTTCGAATAAAACTTCCCAAGCCGTTGCTG
GAAAAATCACTTCTCCAGATACTTCAGAAAGCTCAGAAACTAGCTCTACCTCATCAAGCGATCATATCCCTAGCGAT
TACGATGACGTTGGTAGCAATAGTGGAGATATTAGCAACAACTACGATGACGTAGGTAGTAACAACGGAGATATCAG
TAGCAATTATGACGATGCTGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCT
CTAGAACAAGTGGCCCAGAAAATACAAGTGGTGGTGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACAGC
AATTATGACGATGCTGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCTCTAG
AACAAGTGGCCCAGAAAATACGAGTGGTGGTGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACAGCAATT
ATGACGATGCTGCTGCTGATTACGAGCCGATAAGAACTACTGAAAATATTTATGAGAGTATTGGTGGCTCTAGAACA
AGTGGCCCAGAAAATACGAGTGATGGTGCAGCAGCAGCAGCACTCAATTCTCTAAGAGGCTCCTCCTACACAACAGG
GCCTCGTAACGAGGGTGTATTCGGCCCTGGACCGGAAGGACTACCAGACATGTCTCTTCCTTCATACGATCCTACAA
ATAAAACCTCGTTATTGACTTTCCTCTCCAACCCTCATGTAAAGTCGAAAATGCTTGAAAACTCGGGGCATTTCGTC
TTCATTGATACAGATAGAAGTAGTTTCATTCTTGTTCCTAACGGAAATTGGGACCAAGTCTGTTCAATTAAAGTTCA

AAATGGAAAGACCAAAGAAGATCTCGACATCAAAGACTTGGAAAACATGTGTGCAAAATTCTGTACAGGGTTTAGCA
AATTCTCTGGTGACTGGGACAGTCTTGTAGAACCTATGGTGTCAGCCAAAGCTGGAGTGGCCAGCGGAGGCAATCTT
CCCAATACAGTGATTATCAATAATAAATTCAAAACTTGCGTTGCTTATGGTCCTTGGAATAGCCAGGAAGCAAGTTC
TGGTTATACACCTTCTGCTTGGAGACGTGGTCATCGAGTAGATTTTGGAGGAATTTTTGAGAAAGCCAACGACTTTA
ATAAAATCAACTGGGGAACTCAAGCCGGGCCTAGTAGCGAAGACGATGGCATTTCCTTCTCCAATGAAACTCCTGGA
GCTGGTCCTGCAGCTGCTCCATCACCAACGCCATCCTCTATTCCTATCATCAATGTCAATGTCAATGTTGGCGGAAC
TAATGTGAATATTGGAGATACGAATGTCAACACGACTAACACCACACCAACAACTCAATCTACAGACGCCTCTACAG
ATACAAGCGATATCGATGACATAAATACCAACAACCAAACTGATGATATCAATACGACAGACAAAGACTCTGACGGA
GCTGGTGGAGTCAATGGCGATATATCCGAAACAGAATCCTCTTCTGGAGATGATTCAGGAAGTGTCTCTTCCTCAGA
ATCAGACAAGAATGCCTCTGTCGGAAATGACGGACCTGCTATGAAAGATATCCTTTCTGCCGTGCGTAAACACCTAG
ACGTCGTTTACCCTGGCGAAAATGGCGGTTCTACAGAAGGGCCTCTCCCAGCTAACCAAACTCTCGGAGACGTAATC
TCTGATGTAGAGAATAAAGGCTCCGCTCAGGATACAAAATTGTCAGGAAATACAGGAGCTGGGGATGACGATCCAAC
AACCACAGCTGCTGTAGGTAATGGAGCGGAAGAGATCACTCTTTCCGACACAGATTCTGGTATCGGAGATGATGTAT
CCGATACAGCGTCTTCATCTGGGGATGAATCCGGAGGAGTCTCCTCTCCCTCTTCAGAATCCAATAAAAATACTGCC
GTTGGAAATGACGGACCTTCTGGACTAGATATCCTCGCTGCCGTACGTAAACATTTAGATAAGGTTTACCCTGGCGA
CAATGGTGGTTCTACAGAAGGGCCTCTCCAAGCTAACCAAACTCTTGGAGATATCGTCCAGGATATGGAAACAACAG
GGACATCCCAAGAAACCGTTGTATCCCCATGGAAAGGAAGCACTTCTTCAACGGAATCAGCAGGAGGAAGTGGTAGC
GTACAAACACTACTGCCTTCACCACCTCCAACCCCGTCAACTACAACATTAAGAACGGGCACAGGAGCTACCACCAC
ATCCTTGATGATGGGAGGACCAATCAAAGCTGACATAATAACAACTGGTGGCGGAGGACGAATTCCTGGAGGAGGAA
CGTTAGAAAAGCTGCTCCCTCGTATACGTGCGCACTTAGACATATCCTTTGATGCGCAAGGCGATCTCGTAAGTACT
GAAGAGCCTCAGCTTGGCTCGATTGTAAACAAATTCCGCCAAGAAACTGGTTCAAGAGGAATCTTAGCTTTCGTTGA
GAGTGCTCCAGGCAAGCCGGGATCTGCACAGGTCTTAACGGGTACAGGGGGAGATAAAGGCAACCTATTCCAAGCAG
CTGCCGCAGTCACCCAAGCCTTAGGAAATGTTGCAGGGAAAGTCAACCTTGCGATACAAGGCCAAAAACTATCATCC
CTAGTCAATGACGACGGGAAGGGGTCTGTTGGAAGAGATTTATTCCAAGCAGCAGCCCAAACAACTCAAGTGCTAAG
CGCACTGATTGATACCGTAGGA

SEQ ID NO:76 - CT456 fragment protein sequence

TNSAATSSIQTTGETVVNYTNSASAPNVTVSTSSSSTQATATSNKTSQAVAGKITSPDTSESSETSSTSSSDHIPSD
YDDVGSNSGDISNNYDDVGSNNGDISSNYDDAAADYEPIRTTENIYESIGGSRTSGPENTSGGAAAALNSLRGSSYS
NYDDAAADYEPIRTTENIYESIGGSRTSGPENTSGGAAAALNSLRGSSYSNYDDAAADYEPIRTTENIYESIGGSRT
SGPENTSDGAAAAALNSLRGSSYTTGPRNEGVFGPGPEGLPDMSLPSYDPTNKTSLLTFLSNPHVKSKMLENSGHFV
FIDTDRSSFILVPNGNWDQVCSIKVQNGKTKEDLDIKDLENMCAKFCTGFSKFSGDWDSLVEPMVSAKAGVASGGNL
PNTVIINNKFKTCVAYGPWNSQEASSGYTPSAWRRGHRVDFGGIFEKANDFNKINWGTQAGPSSEDDGISFSNETPG
AGPAAAPSPTPSSIPIINVNVNVGGTNVNIGDTNVNTTNTTPTTQSTDASTDTSDIDDINTNNQTDDINTTDKDSDG
AGGVNGDISETESSSGDDSGSVSSSESDKNASVGNDGPAMKDILSAVRKHLDVVYPGENGGSTEGPLPANQTLGDVI
SDVENKGSAQDTKLSGNTGAGDDDPTTTAAVGNGAEEITLSDTDSGIGDDVSDTASSSGDESGGVSSPSSESNKNTA
VGNDGPSGLDILAAVRKHLDKVYPGDNGGSTEGPLQANQTLGDIVQDMETTGTSQETVVSPWKGSTSSTESAGGSGS
VQTLLPSPPPTPSTTTLRTGTGATTTSLMMGGPIKADIITTGGGGRIPGGGTLEKLLPRIRAHLDISFDAQGDLVST
EEPQLGSIVNKFRQETGSRGILAFVESAPGKPGSAQVLTGTGGDKGNLFQAAAAVTQALGNVAGKVNLAIQGQKLSS
LVNDDGKGSVGRDLFQAAAQTTQVLSALIDTVG

SEQ ID NO:77: CT381 fragment nucleotide sequence

TGTTTAAAAGAAGGGGGAGACTCCAATAGTGAAAAATTTATTGTAGGGACTAATGCAACCTACCCTCCTTTTGAGTT
TGTTGATAAGCGAGGAGAGGTTGTAGGCTTCGATATAGACTTGGCTAGAGAGATTAGTAACAAGCTGGGGAAAACGC
TGGACGTTCGGGAGTTTTCCTTTGATGCACTCATTCTAAACCTAAAACAGCATCGGATTGATGCGGTTATAACAGGG
ATGTCCATTACTCCTTCTAGATTGAAGGAAATTCTTATGATTCCCTATTATGGGGAGGAAATAAAACACTTGGTTTT
AGTGTTTAAAGGAGAGAATAAGCATCCATTGCCACTCACTCAATATCGTTCTGTAGCTGTTCAAACAGGAACCTATC
AAGAGGCCTATTTACAGTCTCTTTCTGAAGTTCATATTCGCTCTTTTGATAGCACTCTAGAAGTACTCATGGAAGTC
ATGCATGGTAAATCTCCCGTCGCTGTTTTAGAGCCATCTATCGCTCAAGTTGTCTTGAAAGATTTCCCGGCTCTTTC
TACAGCAACCATAGATCTCCCTGAAGATCAGTGGGTTTTAGGATACGGGATTGGCGTTGCTTCAGATCGCCCAGCTT
TAGCCTTGAAAATCGAGGCAGCTGTGCAAGAGATCCGAAAAGAAGGAGTGCTAGCAGAGTTGGAACAGAAGTGGGGT
TTGAACAAC

SEQ ID NO:78: CT381 fragment protein sequence

CLKEGGDSNSEKFIVGTNATYPPFEFVDKRGEVVGFDIDLAREISNKLGKTLDVREFSFDALILNLKQHRIDAVITG
MSITPSRLKEILMIPYYGEEIKHLVLVFKGENKHPLPLTQYRSVAVQTGTYQEAYLQSLSEVHIRSFDSTLEVLMEV
MHGKSPVAVLEPSIAQVVLKDFPALSTATIDLPEDQWVLGYGIGVASDRPALALKIEAAVQEIRKEGVLAELEQKWG
LNN

SEQ ID NO:79: CT043 fragment nucleotide sequence

TCCAGGCAGAATGCTGAGGAAAATCTAAAAAATTTTGCTAAAGAGCTTAAACTCCCCGACGTGGCCTTCGATCAGAA
TAATACGTGCATTTTGTTTGTTGATGGAGAGTTTTCTCTTCACCTGACCTACGAAGAACACTCTGATCGCCTTTATG
TTTACGCACCTCTTCTTGACGGACTGCCAGACAATCCGCAAAGAAGGTTAGCTCTATATGAGAAGTTGTTAGAAGGC
TCTATGCTCGGAGGCCAAATGGCTGGTGGAGGGGTAGGAGTCGCTACTAAGGAACAGTTGATCTTAATGCACTGCGT
GTTAGACATGAAGTATGCAGAGACCAACCTACTCAAAGCTTTTGCACAGCTTTTTATTGAAACCGTTGTGAAATGGC
GAACTGTTTGTTCTGATATCAGCGCTGGACGAGAACCCACTGTTGATACCATGCCACAAATGCCTCAAGGGGGTGGC
GGAGGAATTCAACCTCCTCCAGCAGGAATCCGTGCA

SEQ ID NO:80: CT043 fragment protein sequence

SRQNAEENLKNFAKELKLPDVAFDQNNTCILFVDGEFSLHLTYEEHSDRLYVYAPLLDGLPDNPQRRLALYEKLLEG
SMLGGQMAGGGVGVATKEQLILMHCVLDMKYAETNLLKAFAQLFIETVVKWRTVCSDISAGREPTVDTMPQMPQGGG
GGIQPPPAGIRA

SEQ ID NO:81: CT711 fragment nucleotide.seq Length: 2298

TCAATACAACCTACATCCATTTCTTTAACTAAGAATATAACGGCAGCTTTAGCCGGAGAGCAGGTCGATGCTGCTGC
AGTGTATATGCCGCAGGCTGTTTTTTTCTTTCAGCAACTGGATGAAAAAAGCAAGGGGCTGAAACAGGCTTTAGGAT
TGCTCGAAGAGGTTGATCTAGAAAAATTTATACCGTCTTTAGAAAAATCACCTACACCTATCACTACGGGAACAACG
AGTAAAATTTCCGCTGATGGGATTGAGATTGTTGGAGAGCTTTCTTCAGAAACAATTTTGGCAGATCCTAATAAAGC
TGCAGCTCAGGTTTTTGGAGAGGGGCTTGCAGATAGTTTTGATGATTGGCTCAGATTATCTGAAAATGGGGGGATTC
AAGATCCTACAGCAATAGAAGAAGAGATTGTTACTAAGTATCAAACAGAACTCAATACTCTGCGCAATAAACTCAAG
CAACAATCTTTAACAGACGATGAGTATACGAAGCTTTATGCTATTCCTCAAAACTTTGTTAAAGAGATAGAAAGCTT
AAAGAATGAAAATAATGTGAGGTTAATTCCCAAAAGTAAAGTCACTAACTTTTGGCAGAATATCATGCTCACTTACA
ACTCGGTAACCTCGTTATCAGAACCTGTTACCGATGCGATGAATACGACTATGGCGGAGTACTCTCTTTATATTGAG
AGAGCTACAGAGGCTGCCAAGTTGATACGGGAGATAACCAACACGATCAAAGACATTTTCAATCCAGTTTGGGATGT
GCGTGAACAAACAGGAATTTTTGGGTTAAAAGGAGCTGAGTATAACGCTTTAGAAGGCAATATGATTCAAAGCTTGC
TTAGCTTTGCGGGTCTATTCCGGCAGTTAATGAGTCGTACTGCAACAGTTGATGAGATAGGCGCACTTTATCCTAAA
AATGATAAAAACGAAGACGTCATTCATACTGCTATTGATGATTATGTGAATTCTTTAGCTGATTTGAAAGCCAATGA
ACAGGTCAAACTCAACGGTCTGTTGAGTTTAGTATATGCTTATTATGCTAGTACTTTAGGTTTTGCTAAGAAGGATG
TATTCAATAATGCACAAGCTTCTTTTACAGATTATACTAATTTTCTAAACCAAGAGATCCAATATTGGACGCCTAGA
GAGACTTCAAGTTTTAATATCTCCAATCAAGCATTGCAAACCTTTAAAAATAAGCCTTCGGCTGATTATAACGGCGT
ATATCTTTTTGATAATAAAGGATTAGAGACTAATCTCTTTAATCCTACGTTCTTCTTTGATGTTGTGAGTCTCATGA
CAGCTGATCCTACGAAGACTATGTCTCGACAGGATTACAATAAGGTGATTACAGCCTCGGAATCCAGTATTCAGAAG
ATTAATCAGGCTATTACCGCTTGGGAACTAGCTATTGCAGAATGTGGGACTAAAAAAGCGAAGCTCGAACCATCCAG
TTTAAATTATTTTAATGCTATGGTCGAAGCGAAGAAGACCTTCGTAGAGACCTCTCCAATACAGATGGTCTATTCAT
CTTTGATGTTGGATAAGTATCTTCCGAATCAGCAGTACATATTAGAGACATTAGGAAGTCAGATGACTTTCTCTAAC
AAGGCTGCTCGGTATTTAAATGATATCATTGCGTATGCAGTTAGCTTCCAAACAGCTGACGTCTATTATTCTTTAGG
GATGTATCTTCGACAAATGAACCAGCAGGAATTTCCTGAGGTGATTTCTCGTGCTAACGATACTGTGAAAAAAGAGA
TAGATCGGAGTCGTGCGGATCTCTTTCACTGTAAAAAAGCTATCGAAAAGATTAAAGAATTAGTGACTTCTGTAAAT
GCGGATACTGAATTGACCTCATCTCAGCGTGCAGAGTTATTAGAGACGTTAGCTAGTTATGCTTTTGAATTTGAGAA
TCTCTATCACAACCTCTCTAATGTTTACGTCATGGTTTCTAAGGTACAGATTTCTGGCGTAAGCAAGCCTGATGAAG
TGGATGAGGCTTTTACTGCTAAGATTGGATCGAAGGAATTCGATACTTGGATTCAGCAGCTTACAACATTTGAAAGT
GCTGTGATTGAAGGTGGGCGTAATGGTGTGATGCCTGGGGGAGAGCAGCAGGTTTTACAGAGTTTAGAGAGCAAGCA
GCAAGATTACACGTCGTTCAACCAGAATCAGCAATTAGCTCTACAAATGGAGTCCGCAGCGATTCAACAAGAGTGGA
CTATGGTAGCAGCAGCCTTAGCATTAATGAATCAGATTTTTGCTAAGTTGATCCGTAGATTTAAA

SEQ ID NO:82: CT711 fragment protein sequence (AAC68306)

SIQPTSISLTKNITAALAGEQVDAAAVYMPQAVFFFQQLDEKSKGLKQALGLLEEVDLEKFIPSLEKSPTPITTGTT
SKISADGIEIVGELSSETILADPNKAAAQVFGEGLADSFDDWLRLSENGGIQDPTAIEEEIVTKYQTELNTLRNKLK
QQSLTDDEYTKLYAIPQNFVKEIESLKNENNVRLIPKSKVTNFWQNIMLTYNSVTSLSEPVTDAMNTTMAEYSLYIE
RATEAAKLIREITNTIKDIFNPVWDVREQTGIFGLKGAEYNALEGNMIQSLLSFAGLFRQLMSRTATVDEIGALYPK
NDKNEDVIHTAIDDYVNSLADLKANEQVKLNGLLSLVYAYYASTLGFAKKDVFNNAQASFTDYTNFLNQEIQYWTPR
ETSSFNISNQALQTFKNKPSADYNGVYLFDNKGLETNLFNPTFFFDVVSLMTADPTKTMSRQDYNKVITASESSIQK
INQAITAWELAIAECGTKKAKLEPSSLNYFNAMVEAKKTFVETSPIQMVYSSLMLDKYLPNQQYILETLGSQMTFSN
KAARYLNDIIAYAVSFQTADVYYSLGMYLRQMNQQEFPEVISRANDTVKKEIDRSRADLFHCKKAIEKIKELVTSVN
ADTELTSSQRAELLETLASYAFEFENLYHNLSNVYVMVSKVQISGVSKPDEVDEAFTAKIGSKEFDTWIQQLTTFES
AVIEGGRNGVMPGGEQQVLQSLESKQQDYTSFNQNQQLALQMESAAIQQEWTMVAAALALMNQIFAKLIRRFK

SEQ ID NO:83: CT114 fragment nucleotide sequence - Length: 1296

GATCCTTTGAGTGCAAAACAGTTAATGTATCTGTTTCCTCAGCTCTCAGAAGAGGATGTATCTGTTTTTGCTCGATG
CATTTTGTCTTCAAAGCGTCCAGAATACCTCTTTTCAAAATCGGAGGAAGAGCTCTTTGCAAAATTGATTTTGCCAA
GGGTTTCTCTAGGTGTTCATCGGGACGATGATTTAGCGAGAGTGTTGGTGTTAGCGGAGCCTTCTGCAGAAGAGCAG
AAGGCTCGATACTATTCATTGTATCTGGATGTTTTAGCTTTGCGTGCATACGTTGAAAGAGAGCGTTTGGCGAGTGC
TGCACACGGAGATCCTGAGCGGATAGATTTGGCAACCATAGAAGCTATTAATACCATCCTTTTTCAGGAAGAAGGAT
GGAGGTATCCTTCAAAACAAGAGATGTTTGAAAACAGGTTTTCTGAGTTAGCTGCTGTTACAGATAGTAAGTTTGGA
GTTTGCTTGGGAACTGTAGTGCTTTATCAAGCTGTCGCCCAGCGGCTTGATTTGTCTCTGGACCCTGTCACCCCTCC
TGGACATATTTACTTACGCTATAAGGACAAGGTGAATATTGAAACCACTTCTGGAGGAAGGCATCTTCCTACTGAAA
GGTATTGTGAATGCATAAAAGAGTCGCAGTTAAAGGTGCGTTCGCAGATGGAGCTTATAGGGTTAACTTTTATGAAT
AGAGGAGCTTTCTTTTTGCAAAAAGGAGAGTTTCTTCAGGCGTCCTTAGCTTATGAGCAAGCTCAATCATATTTATC
AGACGAGCAGATTTCTGATTTGTTAGGGATTACTTATGTTCTTTTAGGAAAGAAGGCGGCGGGAGAGGCTCTTTTAA
AGAAATCTGCAGAAAAGACTCGGCGAGGGTCATCTATCTATGACTATTTCCAAGGATATATTTCCCCCGAAATCCTA
GGGGTGTTGTTTGCCGATTCAGGGGTGACCTATCAAGAAACTTTGGAGTATCGAAAAAAAACTAGTGATGCTTTCCAA
GAAGTATCCAAAAAGTGGATCTCTTAGGTTGAGGTTGGCGACAACAGCATTGGAGCTAGGGCTGGTCAAGGAGGGGG
TGCAGTTGTTAGAAGAGAGTGTTAAGGATGCCCCAGAGGACCTCTCTTTACGTCTGCAGTTTTGTAAAATTCTTTGC

AATCGACATGATTATGTCCGAGCAAAATATCATTTTGATCAAGCGCAAGCTCTTCTCATTAAAGAAGGGTTGTTTTC
CGAAAAAACTTCCTATACTCTCTTAAAAACTATCGGGAAAAAGCTATCTCTTTTTGCTCCGAGT

SEQ ID NO:84: CT114 fragment protein sequence (AAC67705)

DPLSAKQLMYLFPQLSEEDVSVFARCILSSKRPEYLFSKSEEELFAKLILPRVSLGVHRDDDLARVLVLAEPSAEEQ
KARYYSLYLDVLALRAYVERERLASAAHGDPERIDLATIEAINTILFQEEGWRYPSKQEMFENRFSELAAVTDSKFG
VCLGTVVLYQAVAQRLDLSLDPVTPPGHIYLRYKDKVNIETTSGGRHLPTERYCECIKESQLKVRSQMELIGLTFMN
RGAFFLQKGEFLQASLAYEQAQSYLSDEQISDLLGITYVLLGKKAAGEALLKKSAEKTRRGSSIYDYFQGYISPEIL
GVLFADSGVTYQETLEYRKKLVMLSKKYPKSGSLRLRLATTALELGLVKEGVQLLEESVKDAPEDLSLRLQFCKILC
NRHDYVRAKYHFDQAQALLIKEGLFSEKTSYTLLKTIGKKLSLFAPS

SEQ ID NO:85: CT480 fragment nucleotide sequence

TCTTCAGATCTACTTGAAAAAGATGTGAAATCGATCAAAAGAGAACTCAAGGCTTTACATGAAGATGTTCTTGAGTT
AGTCCGGATCTCGCATCAGCAAAAAAATTGGGTCCAGTCTACAGATTTTTCTGTTTCTCCAGAGATCAGTGTATTGA
AGGATTGCGGAGATCCTGCGTTCCCTAATTTATTATGCGAAGACCCTTATGTTGAAAAAGTGGTCCCTTCGTTGTTA
AAGGAAGGTTTTGTTCCGAAAGGTATTTTGCGTACAGCTCAAGTAGGAAGGCCTGATAACCTAAGTCCGTTTAATGG
CTTTGTTAATATCGTTCGATTTTATGAATTGTGCGTTCCTAATTTGGCTGTTGAGCATGTTGGTAAATACGAGGAGT
TTGCGCCTAGTTTAGCCTTAAAGATAGAAGAGCATTATGTAGAGGATGGGTCTGGGGATAAAGAATTTCATATTTAT
TTGCGTCCTAATATGTTTTGGGAGCCGATAGATCCTACGCTGTTCCCTAAAAATATAACTTTAGCAGACAGCTTCTT
AAGACCACATCCTGTCACCGCTCATGATGTGAAGTTCTATTACGATGTAGTCATGAATCCCTATGTTGCAGAAATGC
GTGCAGTGGCTATGAGATCTTATTTTGAGGATATGGTTTCGGTTCGGGTAGAAAACGATTTGAAATTAATCGTTCGT
TGGAGAGCTCATACTGTACGTAATGAACAGGGAGAGGAAGAGAAAAAAGTGCTCTATTCTGCTTTCGCGAATACATT
GGCACTCCAACCGTTACCTTGTTTCGTGTATCAGCATTTCGCAAATGGAGAGAAGATCGTTCCAGAAGATTCTGATC
CCGATACGTATCGCAAAGATTCGGTATGGGCGCAAAACTTTTCTTCACATTGGGCGTATAATTACATAGTGAGCTGT
GGAGCATTCCGATTTGCAGGGATGGATGATGAGAAAATTACTTTAGTTCGTAATCCTAATTATCATAATCCGTTTGC
GGCTCTTGTGGAGAAGCGCTATATCTATATGAAAGATAGTACAGATTCTCTCTTCCAAGATTTCAAAGCTGGGAAGG
TGGATATTGCGTATTTCCCTCCTAACCATGTCGATAATCTAGCGAGCTTCATGCAAACCTCTGCTTATAAGGAACAA
GCTGCTAGAGGAGAGGCAATTTTAGAAAAAAATTCATCAGACCGGTCCTATTCTTACATCGGATGGAATTGTCTTTC
TCTTTTCTTTAACAATCGTTCGGTACGACAAGCCATGAATATGTTGATCGATCGGGATCGCATTATTGAGCAGTGCT
TGGATGGTCGTGGAGTCTCTGTGAGTGGGCCTTTTTCTCTCTGCTCTCCATCATACAACAGAGATGTAGAGGGATGG
CAATACTCTCCGGAAGAGGCCGCACGTAAATTAGAGGAAGAGGGCTGGATCGATGCTGATGGAGATGGTATTCGTGA
GAAAGTAATCGATGGAGTTGTAGTGCCTTTCCGTTTCCGGTTATGCTACTATGTGAAAAGTGTAACAGCACGAACGA
TTGCCGAATATGTAGCTACGGTATGTAAAGAGGTGGGTATCGAGTGTTGCTTACTCGGGTTAGATATGGCGGATTAT
TCACAAGCCCTCGAGGAGAAAAATTTCGATGCTATTCTTTCCGGATGGTGTTTAGGAACCCCTCCAGAAGATCCTCG
TGCTCTATGGCATTCGGAAGGAGCTTTGGAGAAAGGATCTGCCAATGCTGTTGGATTTTGTAATGAGGAAGCAGACC
GTATCATCGAACAGCTCAGTTACGAGTATGATTCTAATAAGCGCCAAGCCTTGTATCACCGTTTTCACGAGGTGATT
CATGAGGAATCTCCTTACGCGTTTCTCTATTCAAGACAGTACTCCCTTGTCTATAAGGAGTTTGTAAAAAATATTTT
TGTGCCAACAGAACATCAGGATTTGATTCCTGGAGCTCAAGATGAGACAGTGAATTTATCCATGTTGTGGGTAGATA
AAGAGGAGGGTCGATGCTCCGCTATATCT

SEQ ID NO:86: CT480 / oppA_4 fragment protein sequence (AAC68080)

SSDLLEKDVKSIKRELKALHEDVLELVRISHQQKNWVQSTDFSVSPEISVLKDCGDPAFPNLLCEDPYVEKVVPSLL
KEGFVPKGILRTAQVGRPDNLSPFNGFVNIVRFYELCVPNLAVEHVGKYEEFAPSLALKIEEHYVEDGSGDKEFHIY
LRPNMFWEPIDPTLFPKNITLADSFLRPHPVTAHDVKFYYDVVMNPYVAEMRAVAMRSYFEDMVSVRVENDLKLIVR
WRAHTVRNEQGEEEKKVLYSAFANTLALQPLPCFVYQHFANGEKIVPEDSDPDTYRKDSVWAQNFSSHWAYNYIVSC
GAFRFAGMDDEKITLVRNPNYHNPFAALVEKRYIYMKDSTDSLFQDFKAGKVDIAYFPPNHVDNLASFMQTSAYKEQ
AARGEAILEKNSSDRSYSYIGWNCLSLFFNNRSVRQAMNMLIDRDRIIEQCLDGRGVSVSGPFSLCSPSYNRDVEGW
QYSPEEAARKLEEEGWIDADGDGIREKVIDGVVVPFRFRLCYYVKSVTARTIAEYVATVCKEVGIECCLLGLDMADY
SQALEEKNFDAILSGWCLGTPPEDPRALWHSEGALEKGSANAVGFCNEEADRIIEQLSYEYDSNKRQALYHRFHEVI
HEESPYAFLYSRQYSLVYKEFVKNIFVPTEHQDLIPGAQDETVNLSMLWVDKEEGRCSAIS

SEQ ID NO:87: CT089 fragment nucleotide.sequence - Length: 1194

GCTGCAGCTACTCAAGATGCACAAGAGGTTATCGGCTCTCAGGAAGCTTCTGAGGCAAGTATGCTCAAAGGATGTGA
GGATCTCATAAATCCTGCAGCTGCAACCCGAATCAAAAAAAAAAGGAGAGAAGTTTGAATCATTAGAAGCTCGTCGCA
AACCAACAGCGGATAAAGCAGAAAAGAAATCCGAGAGCACAGAGGAAAAAGGCGATACTCCTCTTGAAGATCGTTTC
ACAGAAGATCTTTCCGAAGTCTCCGGAGAAGATTTTCGAGGATTGAAAAATTCGTTCGATGATGATTCTTCTCCTGA
CGAAATTCTCGATGCGCTCACAAGTAAATTTTCTGATCCCACAATAAAGGATCTAGCTCTTGATTATCTAATTCAAA
CAGCTCCCTCTGATGGGAAACTTAAGTCCACTCTCATTCAGGCAAAGCATCAACTGATGAGCCAGAATCCTCAGGCG
ATTGTTGGAGGACGCAATGTTCTGTTAGCTTCAGAAACCTTTGCTTCCAGAGCAAATACATCTCCTTCATCGCTTCG
CTCCTTATATTTCCAAGTAACCTCATCCCCCTCTAATTGCGCTAATTTACATCAAATGCTTGCTTCTTACTTGCCAT
CAGAGAAAACCGCTGTTATGGAGTTTCTAGTAAATGGCATGGTAGCAGATTTAAAATCGGAGGGCCCTTCCATTCCT
CCTGCAAAATTGCAAGTATATATGACGGAACTAAGCAATCTCCAAGCCTTACACTCTGTAAATAGCTTTTTTGATAG
AAATATTGGGAACTTGGAAAATAGCTTAAAGCATGAAGGACATGCCCCTATTCCATCCTTAACGACAGGAAATTTAA
CTAAAACCTTCTTACAATTAGTAGAAGATAAATTCCCTTCCTCTTCCAAAGCTCAAAAGGCATTAAATGAACTGGTA
GGCCCAGATACTGGTCCTCAAACTGAAGTTTTAAACTTATTCTTCCGCGCTCTTAATGGCTGTTCGCCTAGAATATT
CTCTGGAGCTGAAAAAAAACAGCAGCTGGCATCGGTTATCACAAATACGCTAGATGCGATAAATGCGGATAATGAGG

ATTATCCTAAACCAGGTGACTTCCCACGATCTTCCTTCTCTAGTACGCCTCCTCATGCTCCAGTACCTCAATCTGAG
ATTCCAACGTCACCTACCTCAACACAGCCTCCATCACCC

SEQ ID NO:88: CT089 / lcrE fragment protein sequence (AAC67680)

AAATQDAQEVIGSQEASEASMLKGCEDLINPAAATRIKKKGEKFESLEARRKPTADKAEKKSESTEEKGDTPLEDRF
TEDLSEVSGEDFRGLKNSFDDDSSPDEILDALTSKFSDPTIKDLALDYLIQTAPSDGKLKSTLIQAKHQLMSQNPQA
IVGGRNVLLASETFASRANTSPSSLRSLYFQVTSSPSNCANLHQMLASYLPSEKTAVMEFLVNGMVADLKSEGPSIP
PAKLQVYMTELSNLQALHSVNSFFDRNIGNLENSLKHEGHAPIPSLTTGNLTKTFLQLVEDKFPSSSKAQKALNELV
GPDTGPQTEVLNLFFRALNGCSPRIFSGAEKKQQLASVITNTLDAINADNEDYPKPGDFPRSSFSSTPPHAPVPQSE
IPTSPTSTQPPSP

SEQ ID NO:89: CT734 fragment nucleotide sequence - Length: 591

TGTTGCGCCAACTCTTATGGATCGACTCTTGCAAAAAATACAGCCGAGATAAAAGAAGAATCTGTTACACTTCGCGA
GAAGCCGGATGCCGGCTGTAAAAAGAAATCTTCTTGTTACTTGAGAAAATTTTTCTCGCGCAAGAAACCTAAAGAGA
AGACAGAGCCTGTGTTGCCGAACTTTAAGTCTTACGCAGATCCAATGACAGATTCCGAAAGAAAAGACCTTTCTTTC
GTAGTATCTGCTGCTGCTGATAAGTCTTCTATTGCTTTGGCTATGGCTCAGGGGGAAATTAAAGGCGCATTATCGCG
TATTAGAGAGATCCATCCTCTTGCATTGTTACAAGCTCTTGCAGAAGATCCTGCTTTAATTGCTGGAATGAAAAAGA
TGCAAGGACGGGATTGGGTCTGGAATATCTTTATCACAGAATTAAGCAAAGTTTTTTCTCAAGCAGCATCTTTAGGG
GCTTTCAGCGTTGCAGACGTTGCCGCGTTCGCGTCGACCTTAGGATTAGACTCGGGGACCGTTACCTCAATTGTTGA
TGGGGAAAGGTGGGCTGAGCTGATCGATGTCGTGATTCAGAACCCTGCTATA

SEQ ID NO:90: CT734 fragment protein sequence (AAC68329)

CCANSYGSTLAKNTAEIKEESVTLREKPDAGCKKKSSCYLRKFFSRKKPKEKTEPVLPNFKSYADPMTDSERKDLSF
VVSAAADKSSIALAMAQGEIKGALSRIREIHPLALLQALAEDPALIAGMKKMQGRDWVWNIFITELSKVFSQAASLG
AFSVADVAAFASTLGLDSGTVTSIVDGERWAELIDVVIQNPAI

SEQ ID NO:91:CT016 fragment nucleotide sequence

AAAGTTAAAATTAATGATCAGTTCATTTGTATTTCCCCATACATTTCTGCTCGATGGAATCAGATAGCTTTCATAGA
GTCTTGTGATGGAGGGACGGAAGGGGGTATTACTTTGAAACTCCATTTAATTGATGGAGAGACAGTCTCTATACCTA
ATCTAGGACAAGCGATTGTTGATGAGGTGTTCCAAGAGCACTTGCTATATTTAGAGTCCACAGCTCCTCAGAAAAAC
AAGGAAGAGGAAAAAATTAGCTCTTTGTTAGGAGCTGTTCAACAAATGGCTAAAGGATGCGAAGTACAGGTTTTTTC
TCAAAAGGGCTTGGTTTCTATGTTACTAGGAGGAGCTGGTTCGATTAATGTGTTGTTGCAACATTCTCCAGAACATA
AGGATCATCCTGATCTTCCTACCGATTTACTGGAGAGGATAGCGCAAATGATGCGTTCATTATCTATAGGACCAACT
TCTATTTTAGCTAAGCCAGAGCCTCATTGCAACTGTTTGCATTGTCAAATTGGACGAGCTACAGTGGAAGAAGAGGA
TGCCGGAGTATCGGATGAGGATCTTACTTTTCGTTCATGGGATATCTCTCAAAGTGGAGAAAAGATGTACACTGTTA
CAGATCCTTTGAATCCAGAAGAGCAGTTTAATGTGTATTTAGGAACGCCGATTGGATGCACATGTGGGCAGCCATAC
TGTGAACACGTGAAAGCTGTTCTTTATACT

SEQ ID NO:92: CT016 fragment protein sequence (AAC67606)

KVKINDQFICISPYISARWNQIAFIESCDGGTEGGITLKLHLIDGETVSIPNLGQAIVDEVFQEHLLYLESTAPQKN
KEEEKISSLLGAVQQMAKGCEVQVFSQKGLVSMLLGGAGSINVLLQHSPEHKDHPDLPTDLLERIAQMMRSLSIGPT
SILAKPEPHCNCLHCQIGRATVEEEDAGVSDEDLTFRSWDISQSGEKMYTVTDPLNPEEQFNVYLGTPIGCTCGQPY
CEHVKAVLYT

SEQ ID NO:93: CM homolog of CT279 = TC_0551 fragment nucleotide sequence

GCATCCAAGTCTCGTCATTATCTTAACCAGCCTTGGTACATTATCTTATTCATCTTTGTTCTTAGTCTGGTTGCTGG
TACCCTTCTTTCTTCAGTTTCCTATGTTCTATCTCCAATCCAAAAACAAGCTGCAGAATTTGATCGTAATCAGCAAA
TGTTGATGGCCGCACAAATTATTTCCTATGACAATAAATTCCAAATATATGCTGAAGGGGATTGGCAACCTGCTGTC
TATAATACAAAAAAACAGATACTAGAAAAAAGCTCTTCCACTCCACCACAAGTGACTGTGGCGACTCTATGCTCTTA
TTTTCAAAATTTTGTTAGAGTTTTGCTTACAGACTCCCAAGGGAATCTTTCTTCTTTTGAAGATCACAATCTTAACC
TAGAAGAGTTCTTATCCCACCCCACATCTTCAGTACAAGATCACTCTCTGCATGTAATTTATGCTATTCTAGCAAAC
GATGAATCCTCTAAAAAGTTATCATCCTCCCAAGTAGCAAAAAATCCGGTATCCATAGAGTCTATTATTCTTCCTAT
AAAAGGATTTGGTTTATGGGGACCAATCTATGGATTTCTTGCTTTAGAAAAGGACGGTAATACGGTTCTAGGGACAT
GCTGGTATCAACATGGTGAGACTCCAGGATTAGGAGCAAATATAACTAATCCCCAATGGCAACAAAATTTCAGAGGA
AAAAAAGTATTTCTCGCTTCCTCTTCCGGAGAAACCGATTTTGCTAAAACAACTCTAGGACTAGAAGTTATAAAAGG
ATCTGTTTCTGCATTATTAGGGGACTCTCCCAAAGCTAATTCCGCTGTTGATGGAATTTCAGGAGCTACACTGACCT
GTAATGGAGTTACTGAAGCTTTTGCTAATTCGCTAGCTCCTTACCGCCCCTTATTGACTTTCTTCGCCAATCTTAAC
TCTAGTGGAGAATCTCATGACAACCAA

SEQ ID NO:94: CM homologue of CT279 = TC_0551 fragment protein sequence

ASKSRHYLNQPWYIILFIFVLSLVAGTLLSSVSYVLSPIQKQAAEFDRNQQMLMAAQIISYDNKFQIYAEGDWQPAV
YNTKKQILEKSSSTPPQVTVATLCSYFQNFVRVLLTDSQGNLSSFEDHNLNLEEFLSHPTSSVQDHSLHVIYAILAN
DESSKKLSSSQVAKNPVSIESIILPIKGFGLWGPIYGFLALEKDGNTVLGTCWYQHGETPGLGANITNPQWQQNFRG
KKVFLASSSGETDFAKTTLGLEVIKGSVSALLGDSPKANSAVDGISGATLTCNGVTEAFANSLAPYRPLLTFFANLN
SSGESHDNQ

SEQ ID NO:95: CM homologue of CT372 = TC_0651 fragment nucleotide sequence

AATGGAAAAGTTCTGTGTGAGGTTTCTGTGTCCTTCCGTTCGATTCTGCTGACGGCTCTGCTTTCACTTTCTTTTAC
AAACACTATGCAGGCTGCACACCATCATTATCACCGTTATGATGATAAACTACGCAGACAATACCATAAAAAGGACT
TGCCCACTCAAGAGAATGTTCGGAAAGAGTTTTGTAATCCCTACTCTCATAGTAGTGATCCTATCCCTTTGTCACAA
CAACGAGGAGTCCTATCTCCTATCTGTGATTTAGTCTCAGAGTGCTCGTTTTTGAACGGGATTTCCGTTAGGAGTCT
TAAACAAACACTGAAAAATTCTGCTGGGACTCAAGTTGCTTTAGACTGGTCTATCCTTCCTCAATGGTTCAATCCTA
GATCCTCTTGGGCTCCTAAGCTCTCTATTCGAGATCTTGGATATGGTAAACCCCAGTCCCTTATTGAAGCAGATTCC
CCTTGTTGTCAAACCTGCTTCAACCCATCTGCTGCTATTACGATTTACGATTCTTCATGTGGGAAGGGTGTTGTCCA
AGTGTCATACACCCTTGTTCGTTATTGGAGAGAAACGGCTGCACTTGCAGGGCAAACTATGATGCTTGCAGGAAGTA
TTAATGATTATCCTGCTCGCCAAAACATATTCTCTCAACTTACATTTTCCCAAACTTTCCCTAATGAGAGAGTAAAT
CTAACTGTTGGTCAATACTCTCTTTACTCGATAGACGGAACGCTGTACAACAATGATCAGCAGCTAGGATTTATTAG
TTATGCGTTGTCGCAAAATCCAACAGCGACTTATTCCTCTGGAAGCCTTGGCGCCTATCTACAAGTCGCTCCAACAG
AAAGCACCTGTCTTCAAGTTGGGTTCCAAGATGCCTATAATATTTCAGGTTCCTCGATCAAATGGAATAATCTTACA
AAAAATAAGTATAACTTCCATGGCTATGCATCTTGGGCTCCACACTGTTGCTTAGGACCTGGACAATACTCTGTTCT
TCTTTATGTAACCAGAAAGGTTCCTGAGCAAATGATGCAGACAATGGGCTGGTCTGTGAATGCAAGTCAATACATCT
CTTCTAAACTTTATGTATTTGGAAGATACAGCGGAGTCACAGGCCAATTGTCTCCTATTAACCGAACCTATTCATTT
GGCTTAGTCTCTCCTAATTTATTGAACCGTAACCCACAAGACTTATTTGGAGTAGCTTGCGCATTCAATAATATACA
CGCCTCCGCCTTTCAAAATGCTCAAAGAAAATATGAAACTGTGATCGAGGGATTTGCAACTATTGGTTGCGGACCTT
ACATCTCCTTTGCTCCAGATTTCCAACTTTACCTCTATCCTGCTCTGCGTCCAAATAAACAAAGCGCCCGAGTCTAT
AGCGTTCGCGCAAACCTAGCTATT

SEQ ID NO:96: CM homologue of CT372 = TC_0651 fragment protein sequence

NGKVLCEVSVSFRSILLTALLSLSFTNTM@AAHHHYHRYDDKLRR@YHKKDLPT@ENVRKEFCNPYSHSSDPIPLS@
@RGVLSPICDLVSECSFLNGISVRSLK@TLKNSAGT@VALDWSILP@WFNPRSSWAPKLSIRDLGYGKP@SLIEADS
PCC@TCFNPSAAITIYDSSCGKGVV@VSYTLVRYWRETAALAG@TMMLAGSINDYPAR@NIFS@LTFS@TFPNERVN
LTVG@YSLYSIDGTLYNND@@LGFISYALS@NPTATYSSGSLGAYL@VAPTESTCL@VGF@DAYNISGSSIKWNNLT
KNKYNFHGYASWAPHCCLGPG@YSVLLYVTRKVPE@MM@TMGWSVNAS@YISSKLYVFGRYSGVTG@LSPINRTYSF
GLVSPNLLNRNP@DLFGVACAFNNIHASAF@NA@RKYETVIEGFATIGCGPYISFAPDF@LYLYPALRPNK@SARVY
SVRANLAI

SEQ ID NO:97: CM homologue of CT443 = TC_0727 fragment nucleotide sequence

AGCGGGGTGTTAGAGACCTCTATGGCAGAGTCTCTCTCTACCAACGTTATTAGCTTAGCTGACACCAAAGCGAAAGA
GACCACTTCTCATCAAAAAGACAGAAAAGCAAGAAAAAATCATCAAAATAGGACTTCCGTAGTCCGTAAAGAGGTTA
CTGCAGTTCGTGATACTAAAGCTGTAGAGCCTAGACAGGATTCTTGCTTTGGCAAAATGTATACAGTCAAAGTTAAT
GATGATCGTAATGTAGAAATCGTGCAGTCCGTTCCTGAATATGCTACGGTAGGATCTCCATATCCTATTGAGATTAC
TGCTATAGGGAAAAGAGACTGTGTTGATGTAATCATTACACAGCAATTACCATGCGAAGCAGAGTTTGTTAGCAGTG
ATCCAGCTACTACTCCTACTGCTGATGGTAAGCTAGTTTGGAAAATTGATCGGTTAGGACAGGGCGAAAAGAGTAAA
ATTACTGTATGGGTAAAACCTCTTAAAGAAGGTTGCTGCTTTACAGCTGCAACGGTTTGTGCTTGTCCAGAGATCCG
TTCGGTTACGAAATGTGGCCAGCCTGCTATCTGTGTTAAACAGGAAGGTCCAGAAAGCGCATGTTTGCGTTGCCCAG
TAACTTATAGAATTAATGTAGTCAACCAAGGAACAGCAACAGCACGTAATGTTGTTGTGGAAAATCCTGTTCCAGAT
GGCTATGCTCATGCATCCGGACAGCGTGTATTGACATATACTCTTGGGGATATGCAACCTGGAGAACAGAGAACAAT
CACCGTGGAGTTTTGTCCGCTTAAACGTGGTCGAGTCACAAATATTGCTACAGTTTCTTACTGTGGTGGACACAAAA
ATACTGCTAGCGTAACAACAGTGATCAATGAGCCTTGCGTGCAAGTTAACATCGAGGGAGCAGATTGGTCTTATGTT
TGTAAGCCTGTAGAATATGTTATCTCTGTTTCTAACCCTGGTGACTTAGTTTTACGAGACGTTGTAATTGAAGATAC
GCTTTCTCCTGGAATAACTGTTGTTGAAGCAGCTGGAGCTCAGATTTCTTGTAATAAATTGGTTTGGACTTTGAAGG
AACTCAATCCTGGAGAGTCTTTACAATATAAGGTTCTAGTAAGAGCTCAAACTCCAGGGCAATTCACAAACAACGTT
GTTGTGAAAAGTTGCTCTGATTGCGGTATTTGTACTTCTTGCGCAGAAGCAACAACTTACTGGAAAGGAGTTGCTGC
TACTCATATGTGCGTAGTAGATACTTGTGATCCTATTTGCGTAGGAGAGAACACTGTTTATCGTATCTGTGTGACAA
ACAGAGGTTCTGCTGAAGATACAAATGTGTCCTTAATTTTGAAATTCTCTAAAGAATTACAACCTATATCTTTCTCT
GGACCAACTAAAGGAACCATTACAGGAAACACGGTAGTGTTTGATTCGTTACCTAGATTAGGTTCTAAAGAAACTGT
AGAGTTTTCTGTAACGTTGAAAGCAGTATCCGCTGGAGATGCTCGTGGGGAAGCTATTCTTTCTTCCGATACATTGA
CAGTTCCTGTATCTGATACGGAGAATACACATATCTAT

SEQ ID NO:98: CM homologue of CT443 = TC_0727 fragment protein sequence

SGVLETSMAESLSTNVISLADTKAKETTSH@KDRKARKNH@NRTSVVRKEVTAVRDTKAVEPR@DSCFGKMYTVKVN
DDRNVEIV@SVPEYATVGSPYPIEITAIGKRDCVDVIITQ@LPCEAEFVSSDPATTPTADGKLVWKIDRLG@GEKSK
ITVWVKPLKEGCCFTAATVCACPEIRSVTKCG@PAICVK@EGPESACLRCPVTYRINVVN@GTATARNVVVENPVPD
GYAHASG@RVLTYTLGDM@PGE@RTITVEFCPLKRGRVTNIATVSYCGGHKNTASVTTVINEPCV@VNIEGADWSYV
CKPVEYVISVSNPGDLVLRDVVIEDTLSPGITVVEAAGA@ISCNKLVWTLKELNPGESL@YKVLVRA@TPG@FTNNV
VVKSCSDCGICTSCAEATTYWKGVAATHMCVVDTCDPICVGENTVYRICVTNRGSAEDTNVSLILKFSKEL@PISFS
GPTKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAILSSDTLTVPVSDTENTHIY

SEQ ID NO:99: CM homologue of CT043 = TC_0313 fragment nucleotide sequence

TCCAGACAGAATGCTGAGGAAAATCTAAAAAATTTTGCTAAAGAGCTCAAGCTCCCCGACGTGGCCTTCGATCAGAA
TAATACGTGCATTTTGTTTGTTGATGGAGAGTTTTCTCTTCACCTGACCTACGAAGAGCACTCTGATCGCCTTTATG
TTTACGCACCTCTCCTTGACGGACTCCCAGATAATCCGCAAAGAAAGTTGGCTCTGTATGAGAAATTGTTGGAAGGC

TCTATGCTCGGAGGCCAAATGGCTGGTGGAGGAGTAGGAGTTGCTACTAAAGAACAGTTGATCCTAATGCATTGCGT
GTTAGATATGAAATATGCAGAGACTAATCTATTGAAAGCTTTTGCACAGCTTTTCATTGAAACTGTTGTGAAATGGC
GAACGGTCTGTTCTGATATCAGCGCTGGACGAGAACCTTCCGTTGACACTATGCCTCAAATGCCTCAAGGAGGCAGC
GGAGGAATTCAACCTCCTCCAACAGGAATTCGTGCG

SEQ ID NO:100: CM homologue of CT043 = TC_0313 fragment protein sequence

SRQNAEENLKNFAKELKLPDVAFDQNNTCILFVDGEFSLHLTYEEHSDRLYVYAPLLDGLPDNPQRKLALYEKLLEG
SMLGGQMAGGGVGVATKEQLILMHCVLDMKYAETNLLKAFAQLFIETVVKWRTVCSDISAGREPSVDTMPQMPQGGS
GGIQPPPTGIRA

SEQ ID NO:101: CM homologue of CT601 = TC_0890 fragment nucleotide sequence

CTCGCTAATCGGTTATTTCTAATCACCCTTATAGGTTTTGGCTATTCTGCTTACGGTGCCAGCACAGGGAAATCACC
TTCTTTACAGGTTATTTTAGCTGAAGTCGAGGATACATCTTCGCGCTTACAAGCTCATCAGAATGAGCTTGTTATGC
TCTCGGAACGTTTAGATGAGCAAGACACAAAACTTCAACAACTCTCGTCAACTCAGGCCCGTAATCTTCCTCAACAA
GTTCAACGGCTTGAGATTGATCTGAGAGCTCTGGCTAAAACAGCTGCTGTGCTCTCGCAATCTGTTCAGGATATCCG
ATCATCCGTGCAAAATAAATTACAAGAAATCCAACAAGAACAAAAAAATTTAGCTCAAAATTTACGAGCGCTTCGCA
ACTCCTTACAAGCACTAGTTGATGGCTCTTCCCCAGAAAATTATATTGATTTTTTGGCCGGGGAGACACCTGAACAT
ATTCACGTTGTTAAACAAGGAGAAACCCTGAGTAAAATCGCTAGTAAGTACAATATCCCTGTCGCAGAATTGAAAAA
ACTTAATAAATTAAATTCCGATACTATTTTTACTGATCAAAGAATCCGACTTCCAAAAAAGAAA

SEQ ID NO:102: CM homologue of CT601 = TC_0890 fragment protein sequence

LANRLFLITLIGFGYSAYGASTGKSPSLQVILAEVEDTSSRLQAHQNELVMLSERLDEQDTKLQQLSSTQARNLPQQ
VQRLEIDLRALAKTAAVLSQSVQDIRSSVQNKLQEIQQEQKNLAQNLRALRNSLQALVDGSSPENYIDFLAGETPEH
IHVVKQGETLSKIASKYNIPVAELKKLNKLNSDTIFTDQRIRLPKKK

SEQ ID NO:103: CM homologue of CT456 = TC_0741 fragment nucleotide sequence

ACGACTCCAATAAGTAATTCTCCATCTTCTATTCCAACTGTTACAGTATCAACTACTACAGCATCTTCTGGATCTCT
CGGAACTTCTACTGTATCATCAACGACTACAAGTACTTCAGTCGCACAAACAGCAACAACAACATCTTCTGCTTCTA
CATCTATAATTCAGTCTAGTGGAGAAAACATCCAATCCACTACAGGTACCCCTTCTCCTATTACGTCTAGTGTTTCA
ACATCCGCTCCATCTCCTAAAGCCTCCGCCACTGCAAACAAAACTTCAAGCGCTGTTTCTGGGAAAATTACCTCACA
AGAAACTTCTGAGGAATCCGAAACCCAAGCCACTACATCTGATGGAGAAGTTAGTAGTAATTACGATGATGTTGATA
CCCCGACCAATTCGTCCGATTCGACAGTTGATAGTGATTACCAAGATGTTGAGACTCAGTACAAAACAATTAGCAAC
AATGGTGAAAACACTTATGAAACAATCGGAAGTCATGGTGAGAAAAACACACACGTCCAGGAAAGCCATGCATCCGG
AACAGGAAATCCCATAAATAATCAGCAAGAAGCTATTAGACAGCTCCGATCATCTACCTATACAACCAGCCCTCGTA
ATGAGAATATATTTAGTCCAGGACCGGAAGGTCTACCTAATATGTCTCTTCCTAGTTACAGCCCTACAGATAAAAGT
TCTCTACTAGCTTTCCTATCTAATCCCAATACAAAAGCAAAAATGCTCGAACACTCCGGGCATTTAGTCTTTATAGA
CACAACTAGAAGTAGCTTTATCTTTGTTCCGAATGGAAATTGGGATCAAGTCTGTTCCATGAAGGTTCAGAATGGGA
AAACTAAAGAAGACCTTGGCTTAAAGGACTTAGAAGATATGTGTGCAAAGTTTTGCACAGGATACAATAAATTCTCC
TCTGATTGGGGAAATCGAGTTGACCCCTTGGTCTCTTCTAAGGCCGGGATAGAAAGTGGGGGGCACCTCCCAAGCTC
AGTTATCATCAACAACAAATTTAGAACCTGTGTTGCCTATGGGCCGTGGAACCCCAAAGAAAACGGCCCCAATTATA
CTCCTTCAGCCTGGAGACGTGGGCATCGAGTAGATTTTGGAAAGATCTTTGATGGAACAGCGCCGTTTAATAAAATC
AACTGGGGCTCTTCCCCTACCCCTGGTGATGACGGCATCTCCTTCTCTAATGAAACTATTGGGTCTGAACCATTCGC
GACACCTCCCTCATCCCCATCGCAAACCCCCGTTATCAACGTCAATGTTAATGTCGGTGGAACCAATGTTAATATTG
GGGATACAAACGTATCTAAAGGATCCGGCACACCAACATCTTCTCAATCTGTGGACATGTCTACAGATACTAGCGAT
TTAGATACCAGTGATATTGATACAAACAACCAAACTAACGGCGATATCAACACGAATGACAACTCCAATAATGTCGA
TGGAAGTTTATCTGACGTTGATTCAAGGGTGGAAGACGATGACGGTGTATCGGATACAGAGTCCACTAATGGCAATG
ACTCTGGTAAAACTACTTCCACAGAAGAAAATGGTGACCCAAGCGGACCAGACATCCTGGCTGCTGTACGTAAACAC
CTAGACACTGTCTATCCAGGAGAAAATGGCGGATCTACAGAAGGACCTCTCCCTGCTAATCAAAATCTGGGGAACGT
TATCCATGATGTGGAGCAGAATGGATCTGCTAAAGAAACTATTATCACTCCAGGAGATACAGGGCCTACAGACTCAA
GCTCCTCTGTAGATGCTGATGCAGACGTTGAAGATACTTCTGATACTGACTCTGGAATCGGAGACGACGACGGTGTA
TCGGATACAGAGTCCACTAATGGTAATAACTCTGGTAAAACTACTTCCACAGAAGAAAATGGTGACCCAAGCGGACC
AGACATCCTGGCTGCTGTACGTAAACACCTAGACACTGTCTATCCAGGAGAAAATGGCGGATCTACAGAAGGACCTC
TCCCTGCTAATCAAAATCTGGGGAACGTTATCCATGATGTAGAACAAAACGGAGCCGCTCAAGAAACTATTATCACT
CCAGGAGATACGGAATCTACAGACACAAGCTCTAGTGTAAATGCTAATGCAGACTTAGAAGATGTTTCTGATGCTGA
TTCAGGATTCGGGGATGATGACGGTATATCGGATACAGAGTCCACTAATGGTAACGACTCTGGAAAAAATACTCCTG
TAGGGGATGGTGGTACACCAAGCGGACCAGATATCCTAGCTGCTGTACGCAAACATCTAGACACTGTCTATCCAGGA
GAAAATGGTGGATCTACAGAGAGACCTTTACCCGCTAATCAAAATTTAGGAGATATCATTCATGATGTAGAACAAAA
CGGAAGCGCTAAAGAAACTGTAGTATCGCCTTATCGAGGAGGAGGAGGAAATACATCTTCCCCAATTGGATTAGCCT
CCCTGCTTCCAGCAACACCATCCACACCTTTGATGACAACACCTAGAACAAATGGGAAAGCTGCAGCTTCTTCTTTG
ATGATAAAAGGAGGAGAAACTCAAGCCAAGCTAGTTAAGAATGGCGGCAATATCCCTGGAGAAACCACATTAGCAGA
ATTACTCCCTCGTTTAAGAGGACACCTTGACAAAGTCTTTACTTCAGACGGGAAGTTTACAAATCTTAATGGACCTC
AACTTGGAGCCATCATAGACCAATTCCGCAAAGAAACGGGTTCCGGAGGAATCATAGCTCATACAGATAGTGTTCCA
GGAGAGAACGGAACAGCCTCTCCTCTCACAGGAAGTTCAGGGGAAAAAGTCTCTCTCTATGATGCAGCGAAAAACGT
CACTCAAGCTTTAACAAGTGTTACGAACAAAGTAACCCTAGCAATGCAAGGACAAAAACTGGAAGGAATTATAAACA
ACAACAATACCCCCTCTTCTATTGGACAAAATCTTTTCGCAGCAGCGAGGGCAACGACACAATCCCTCAGTTCATTA
ATTGGAACCGTACAA

SEQ ID NO:104: CM homologue of CT456 = TC_0741 fragment protein sequence

TTPISNSPSSIPTVTVSTTTASSGSLGTSTVSSTTTSTSVAQTATTTSSASTSIIQSSGENIQSTTGTPSPITSSVS
TSAPSPKASATANKTSSAVSGKITSQETSEESETQATTSDGEVSSNYDDVDTPTNSSDSTVDSDYQDVETQYKTISN
NGENTYETIGSHGEKNTHVQESHASGTGNPINNQQEAIRQLRSSTYTTSPRNENIFSPGPEGLPNMSLPSYSPTDKS
SLLAFLSNPNTKAKMLEHSGHLVFIDTTRSSFIFVPNGNWDQVCSMKVQNGKTKEDLGLKDLEDMCAKFCTGYNKFS
SDWGNRVDPLVSSKAGIESGGHLPSSVIINNKFRTCVAYGPWNPKENGPNYTPSAWRRGHRVDFGKIFDGTAPFNKI
NWGSSPTPGDDGISFSNETIGSEPFATPPSSPSQTPVINVNVNVGGTNVNIGDTNVSKGSGTPTSSQSVDMSTDTSD
LDTSDIDTNNQTNGDINTNDNSNNVDGSLSDVDSRVEDDDGVSDTESTNGNDSGKTTSTEENGDPSGPDILAAVRKH
LDTVYPGENGGSTEGPLPANQNLGNVIHDVEQNGSAKETIITPGDTGPTDSSSSVDADADVEDTSDTDSGIGDDDGV
SDTESTNGNNSGKTTSTEENGDPSGPDILAAVRKHLDTVYPGENGGSTEGPLPANQNLGNVIHDVEQNGAAQETIIT
PGDTESTDTSSSVNANADLEDVSDADSGFGDDDGISDTESTNGNDSGKNTPVGDGGTPSGPDILAAVRKHLDTVYPG
ENGGSTERPLPANQNLGDIIHDVEQNGSAKETVVSPYRGGGGNTSSPIGLASLLPATPSTPLMTTPRTNGKAAASSL
MIKGGETQAKLVKNGGNIPGETTLAELLPRLRGHLDKVFTSDGKFTNLNGPQLGAIIDQFRKETGSGGIIAHTDSVP
GENGTASPLTGSSGEKVSLYDAAKNVTQALTSVTNKVTLAMQGQKLEGIINNNNTPSSIGQNLFAAARATTQSLSSL
IGTVQ

SEQ ID NO:105: CM homologue of CT381 = TC_0660 fragment nucleotide sequence

TGTTCAAAAGAGAGCAAAGACTCTGTTAGTGAAAAATTTATTGTAGGAACTAACGCAACGTATCCTCCTTTTGAGTT
TGTTGATGAAAGAGGTGAGACGGTTGGCTTTGATATTGATTTAGCTAGGGAGATTAGTAAAAAGCTAGGGAAAAAAT
TAGAAGTCCGAGAATTTGCTTTTGATGCACTCGTTCTCAATTTAAAACAGCATCGTATTGATGCAATTATGGCAGGG
GTGTCCATTACGTCTTCTCGATTGAAAGAAATTTTGATGATTCCCTACTATGGCGAAGAAATAAAGAGTTTGGTTTT
AGTGTTTAAGGATGGAGACTCAAAGTCTTTACCACTAGATCAGTATAATTCTGTTGCTGTTCAAACTGGCACGTACC
AAGAGGAATATTTACAGTCTCTTCCAGGGGTGCGTATTCGCTCTTTTGATAGTACTTTAGAAGTGCTTATGGAAGTT
TTGCATAGCAAGTCTCCTATAGCTGTTTTAGAACCGTCTATTGCGCAGGTCGTTTTAAAAGATTTTCCGACGCTCAC
TACTGAAACGATAGATCTTCCTGAAGATAAATGGGTTTTAGGGTATGGAATTGGAGTTGCTTCTGATCGACCATCTC
TAGCTTCTGATATAGAAGCTGCTGTACAAGAGATCAAGAAAGAAGGAGTGTTAGCAGAGTTAGAGCAAAAATGGGGT
TTGAACGGC

SEQ ID NO:106: CM homologue of CT381 = TC_0660 fragment protein sequence

CSKESKDSVSEKFIVGTNATYPPFEFVDERGETVGFDIDLAREISKKLGKKLEVREFAFDALVLNLKQHRIDAIMAG
VSITSSRLKEILMIPYYGEEIKSLVLVFKDGDSKSLPLDQYNSVAVQTGTYQEEYLQSLPGVRIRSFDSTLEVLMEV
LHSKSPIAVLEPSIAQVVLKDFPTLTTETIDLPEDKWVLGYGIGVASDRPSLASDIEAAVQEIKKEGVLAELEQKWG
LNG

SEQ ID NO:107 – CT255 fragment nucleotide sequence

GAAGAAAAAGGCATCTTACAATTGGTTGAAATTTCGCGAGCAATGGCTTTACAGGGAGTTTGTCCTTGGACTAATTT
ACAGAGTGTGGAGTCTATGTTGCAGTATATAGCAGGGGAGTGTCAGGAGTTGGCTGATGCTGTACAAGAAAATAAAG
CTTCGTTGGAAATCGCTTCGGAAGCCGGAGACGTACTTACTTTAGTATTGACCTTGTGTTTCTTGCTAGAAAGAGAA
GGAAAGCTTAAAGCTGAAGAAGTATTTGTAGAAGCTTTGGCTAAGTTGCGTCGTCGATCTCCTCATGTTTTTGATCC
TCATAATCAAATTTCTTTAGAACAGGCTGAAGAATACTGGGCTCGTATGAAACAGCAAGAAAAAATTTCT

SEQ ID NO:108 – CT255 fragment protein sequence

EEKGILQLVEISRAMALQGVCPWTNLQSVESMLQYIAGECQELADAVQENKASLEIASEAGDVLTLVLTLCFLLERE
GKLKAEEVFVEALAKLRRRSPHVFDPHNQISLEQAEEYWARMKQQEKIS

SEQ ID NO:109 – CT341 fragment nucleotide sequence

GATTACTACACGATATTGGGTGTAGCGAAGACTGCTACTCCTGAAGAAATAAAGAAAGCTTACCGTAAGCTCGCTGT
AAAGTACCATCCAGATAAGAATCCTGGGGATGCTGAAGCGGAGCGACGCTTTAAAGAAGTTTCTGAAGCCTATGAAG
TATTAGGTGATGCGCAGAAGCGGGAGTCATATGATCGTTACGGCAAAGACGGTCCATTTGCTGGTGCTGGAGGATTC
GGTGGCGCTGGCATGGGGAATATGGAAGACGCTTTGCGAACATTTATGGGAGCTTTTGGCGGCGATTTCGGTGGTAA
TGGAGGCGGTTTCTTTGAAGGGCTTTTTGGAGGACTTGGAGAAGCTTTCGGAATGCGTGGAGGCTCAGAAAGTTCTC
GACAAGGAGCTAGTAAGAAGGTGCATATTACGCTGTCCTTCGAGGAGGCGGCAAAAGGTGTTGAAAAAGAACTTCTT
GTTTCAGGCTATAAATCTTGTGATGCTTGTTCTGGTAGTGGAGCCAATACTGCTAAAGGTGTAAAAGTTTGTGATCG
ATGCAAGGGCTCTGGTCAGGTAGTGCAAAGCCGAGGCTTTTTCTCCATGGCTTCTACTTGCCCTGATTGTAGTGGTG
AAGGTCGGGTTATCACAGATCCTTGTTCAGTTTGTCGTGGGCAGGGACGTATCAAGGATAAACGTAGCGTCCATGTT
AATATCCCAGCTGGAGTCGATTCTGGGATGAGATTAAAGATGGAAGGCTATGGAGATGCTGGCCAAAATGGAGCGCC
TGCAGGGGATCTGTATGTTTTTATTGATGTAGAGCCTCATCCTGTTTTCGAGCGCCATGGGGATGATTTAGTTTTAG
AGCTTCCTATTGGATTTGTTGATGCGGCTTTAGGGATCAAGAAGGAAATCCCTACACTCTTAAAAGAAGGTACTTGC
CGTTTGAGTATCCCAGAAGGGATTCAGAGCGGAACAGTTCTTAAAGTTAGAGGGCAGGGATTCCCTAATGTGCATGG
GAAATCCAGAGGAGATCTTTTAGTAAGAGTATCTGTGGAGACTCCCCAGCACCTATCTAATGAACAAAAAGATTTAT
TGAGACAGTTTGCTGCTACGGAGAAGGCTGAAAATTTCCCTAAGAAACGGAGTTTCTTAGACAAAATCAAAGGTTTT
TTTTCTGACTTTGCTGTA

SEQ ID NO:110 – CT341 fragment protein sequence

DYYTILGVAKTATPEEIKKAYRKLAVKYHPDKNPGDAEAERRFKEVSEAYEVLGDAQKRESYDRYGKDGPFAGAGGF
GGAGMGNMEDALRTFMGAFGGDFGGNGGGFFEGLFGGLGEAFGMRGGSESSRQGASKKVHITLSFEEAAKGVEKELL
VSGYKSCDACSGSGANTAKGVKVCDRCKGSGQVVQSRGFFSMASTCPDCSGEGRVITDPCSVCRGQGRIKDKRSVHV
NIPAGVDSGMRLKMEGYGDAGQNGAPAGDLYVFIDVEPHPVFERHGDDLVLELPIGFVDAALGIKKEIPTLLKEGTC
RLSIPEGIQSGTVLKVRGQGFPNVHGKSRGDLLVRVSVETPQHLSNEQKDLLRQFAATEKAENFPKKRSFLDKIKGF
FSDFAV

SEQ ID NO:111 – CT716 fragment nucleotide sequence

AATAAAAAACTCCAAGATCTGTCTAAACTGCTCACTATTGAGCTTTTCAAGAAACGTACACGGTTGGAAACAGTAAA
AAAAGCGCTCTCCACAATAGAACATCGCTTACAACAAATACAGGAGCACATCGCGAAAATTTCCTTAACAAGGCACA
AACAATTCCTATGTCGGTCATATACCCATGAATATGACCAACATTTAGAACATTTACAAAGAGAGCAAACTTCTCTA
TATAAACAGCATCAGACCCTGAAAACGTCTTTGAAAGATGCTTATGGCGACATACAAAAACAACTAGACCAAAGAAA
AATTATCGAAAAGATCCATGACAGTAAATATCCTATAAAGAGCGCGAATAAC

SEQ ID NO:112 – CT716 fragment protein sequence

NKKLQDLSKLLTIELFKKRTRLETVKKALSTIEHRLQQIQEHIAKISLTRHKQFLCRSYTHEYDQHLEHLQREQTSL
YKQHQTLKTSLKDAYGDIQKQLDQRKIIEKIHDSKYPIKSANN

SEQ ID NO:113 – CT745 fragment nucleotide sequence

GCGTGGTGGCTACACAAACGATTCCCTCATGTGCAGCTGTCTATTCTAGAAAAAGAGTCTCGATCTGGAGGGCTAAT
TGTCACAGAGAAACAACAAGGGTTTTCCCTCAATATGGGCCCTAAAGGTTTTGTTTTAGCTCATGATGGGCAACACA
CCCTTCACCTCATTCAGTCTTTAGGCCTAGCAGACGAGCTATTATATAGCTCTCCAGAGGCTAAAAACCGCTTTATC
CACTATAATAATAAAACCCGAAAAGTCTCGCCTTGGACTATTTTCAAACAAAATCTCCCTCTCTCTTTTGCTAAGGA
TTTCTTTGCGCGTCCTTACAAACAAGACAGCTCCGTGGAAGCCTTCTTTAAAAGACACAGTTCTTCCAAGCTTAGAA
GAAATCTTTTAAATCCCATTAGCATTGCTATTCGTGCAGGACATAGTCATATATTGTCTGCACAGATGGCTTACCCA
GAATTAACACGAAGAGAAGCTCAAACAGGATCGTTGTTACGTAGTTATCTCAAAGATTTTCCTAAAGAGAAACGCAC
AGGCCCTTATTTAGCTACCTTGCGGTCTGGGATGGGAATGCTAACCCAGGCTTTGCATGATAAATTGCCTGCTACCT
GGTATTTTTCTGCACCCGTCAGCAAAATCCGTCAGTTGGCGAATGGGAAAATTTCTCTTTCATCTCCTCAAGGAGAA
ATAACGGGAGATATGCTCATTTATGCTGGGTCCGTGCACGATCTCCCTTCCTGTCTAGAAGGGATCCCTGAAACCAA
GCTTATCAAGCAAACGACTTCATCTTGGGATCTCTCTTGTGTATCTTTAGGATGGCATGCATCCTTCCCTATCCCTC
ATGGATATGGCATGCTTTTCGCTGATACGCCTCCCTTATTAGGGATCGTGTTTAATACGGAAGTGTTCCCTCAACCC
GAGCGGCCTAATACAATAGTCTCTCTTCTTTTAGAAGGTCGATGGCACCAAGAAGAAGCGTATGCTTTCTCACTAGC
AGCTATTTCTGAGTACCTGCAAATTTACACTCCTCCCCAAGCTTTCTCACTATTCTCTCCTCGAGAGGGACTTCCCC
AACACCATGTTGGATTTATCCAATCCCGCCAACGCCTTCTATCTAAACTTCCTCACAATATAAAAATTGTAGGGCAG
AATTTTGCAGGTCCAGGTCTCAACCGCGCTACAGCGTCTGCTTATAAAGCTATAGCTTCTTTACTATCA

SEQ ID NO:114 – CT745 fragment protein sequence

AWWLHKRFPHVQLSILEKESRSGGLIVTEKQQGFSLNMGPKGFVLAHDGQHTLHLIQSLGLADELLYSSPEAKNRFI
HYNNKTRKVSPWTIFKQNLPLSFAKDFFARPYKQDSSVEAFFKRHSSSKLRRNLLNPISIAIRAGHSHILSAQMAYP
ELTRREAQTGSLLRSYLKDFPKEKRTGPYLATLRSGMGMLTQALHDKLPATWYFSAPVSKIRQLANGKISLSSPQGE
ITGDMLIYAGSVHDLPSCLEGIPETKLIKQTTSSWDLSCVSLGWHASFPIPHGYGMLFADTPPLLGIVFNTEVFPQP
ERPNTIVSLLLEGRWHQEEAYAFSLAAISEYLQIYTPPQAFSLFSPREGLPQHHVGFIQSRQRLLSKLPHNIKIVGQ
NFAGPGLNRATASAYKAIASLLS

SEQ ID NO: 115 – CT387 fragment nucleotide sequence

ACGCTCTTTCATTCTCATCATGATGCCGTCTCTCCAGACAGCTACCTATGTTCTTCCCTTCAGTTAGTTGGTACTGG
CGTATACGAAGGAGAAATCGAGATTCAAAATATCCCCTCTTATTTCCTTGGATTCCAATTACCCTCTCATTGCATAC
ACCTTAATTTAAAGAGCTCTCTAGCTCAATTAGGAATAGATGCCTCCCTTCTTCACTGCGAATTGAGCAAAAATCAA
CATCGAGCACATATACATGCTCAATTTACCGGTCATGGCCCCATTGCTGAATCTATGCTAGCCCTTCTCCAACCAGG
AGATCGTGTAGCAAAACTATTTGCTGCAGACGATCGCAGACTGGTCCGATCTCCAGATTACCTCGAAAGCATGCTGA
AAAATACAGATAAAGCTGGCCATCCTTTGCTCTGTTTTGGGAAAAAATTAGAACACTTGATTTCTTTTGATGTGGTA
GATGATCGCCTTGTCGTCTCCCTTCCTACCCTGCCGGGAGTTGTTCGTTATGATTCGGATATTTATGGACTCCTTCC
TCTTATTCAAAAATCACTCAGTAATCCCAAACTCAGCATTCGTCACTTTTTAGCTCTGTACCAACAGATTGTGGAAG
GGCAACATGTCTCTTGCGGAAACCATATTCTTCTGATCAAAACAGAACCGCTGCACATCCGCACTGTATTTGCTCGC
GTGGTAAATCAACTCCTCCCTCAAGGTCTCTCCCACACTTCTGCCAATATTTTGGAACCAACCACTCGAGAATCCGG
GGATATCTTTGAATTTTTTGGGAACCCTTCTGCACAGATAGAAAGAATTCCTTTAGAATTTTTCACTATCGAACCCT
ATAAAGAACATTCTTACTTCTGTAATCGGGATTTATTACAAACCATCTTACAATCAGAAAGCGAAATCAAAAAAATA
TTCGAAACAGCGCCCAAAGAACCTGTCAAAGCTGCCACCTATTTATCAAAAGGCAGTGAAATCTCTTCCCTGCACAC
AGACTCTTGGCTCACAGGATCCGCAGCTGCCTATCAATATAGTGAGCAAGCAGATAAAAACGAGTACACTCATGCTC
AACCTTGCTATCCTTTCTTAGAAGCAATGGAAATGGGCCTGATCAATAGCGAAGGAGCCTTACTCACTCGTTATTTC
CCTTCAGCTAGCTTAAAAGGAATGTTGATTTCCTACCATGTGCGCCACTATCTCAAACAAATCTACTTTCAAGTTCC
CTCTTATACACATGGAAACTATTTCTCTCATAATGACAGAGGTTTGCTATTAGATCTGCAGCAAGCAGATATTGATG
TTTTCTGGGCAGATGAAGAAAGCGGCCGTGTGTTGCAATATACAAAACGACGCGATAAGAATAGCGGTATGTTCGTG
ATCAAAAATCGTGTTGAAGAGTTTCGATCAGCTTATTTTATTGCTATTTATGGCTCTCGTCTCCTTGAGAATAATTT
CTCTGCTCAGCTCCATACCCTCCTAGCGGGCTTACAGCAAGCAGCACATACTCTCGGCATTCCTGGATTCTCAAAGC

CTACCCCACTTGCAGTCATCACCGGAGGCGGCACTGGAGTTATGGCCACAGGAAATCGTGTAGCTAAAGAACTAGGA
ATCCTATCTTGTGGAACCGTTCTTGATTTAGAAGCTTCTCCAGCACAAATCGACCAACCTACCAATGAATTCTTAGA
TGCTAAAATGACATACCGCCTACCTCAACTTATAGAAAGGCAAGAACACTTTTATGCAGACCTTCCTATCCTTGTAG
TTGGCGGTGTAGGAACCGATTTCGAACTCTACCTAGAACTTGTCTATCTCAAAACAGGAGCTAAACCACCGACTCCC
ATTTTCCTAATTGGACCTATTGAATACTGGAAAGAAAAAGTGGCCCACGCCTACGAGATCAACCTCAAAGCAGGAAC
CATCCGTGGATCCGAATGGATCAGCAACTGCCTATATTGTATCACTTCTCCGGAAGCTGGAATTGCCGTATTCGAAC
AATTCCTAGCTGGAGAACTCCCTATAGGATACGACTATCCTCCAGCTCCAGATGGATTAGTGATCGTC

SEQ ID NO:116 – CT387 fragment protein sequence

TLFHSHHDAVSPDSYLCSSLQLVGTGVYEGEIEIQNIPSYFLGFQLPSHCIHLNLKSSLAQLGIDASLLHCELSKNQ
HRAHIHAQFTGHGPIAESMLALLQPGDRVAKLFAADDRRLVRSPDYLESMLKNTDKAGHPLLCFGKKLEHLISFDVV
DDRLVVSLPTLPGVVRYDSDIYGLLPLIQKSLSNPKLSIRHFLALYQQIVEGQHVSCGNHILLIKTEPLHIRTVFAR
VVNQLLPQGLSHTSANILEPTTRESGDIFEFFGNPSAQIERIPLEFFTIEPYKEHSYFCNRDLLQTILQSESEIKKI
FETAPKEPVKAATYLSKGSEISSLHTDSWLTGSAAAYQYSEQADKNEYTHAQPCYPFLEAMEMGLINSEGALLTRYF
PSASLKGMLISYHVRHYLKQIYFQVPSYTHGNYFSHNDRGLLLDLQQADIDVFWADEESGRVLQYTKRRDKNSGMFV
IKNRVEEFRSAYFIAIYGSRLLENNFSAQLHTLLAGLQQAAHTLGIPGFSKPTPLAVITGGGTGVMATGNRVAKELG
ILSCGTVLDLEASPAQIDQPTNEFLDAKMTYRLPQLIERQEHFYADLPILVVGGVGTDFELYLELVYLKTGAKPPTP
IFLIGPIEYWKEKVAHAYEINLKAGTIRGSEWISNCLYCITSPEAGIAVFEQFLAGELPIGYDYPPAPDGLVIV

SEQ ID NO:117 – CT812 fragment nucleotide sequence

TGCGTAGATCTTCATGCTGGAGGACAGTCTGTAAATGAGCTGGTATATGTAGGCCCTCAAGCGGTTTTATTGTTAGA
CCAAATTCGAGATCTATTCGTTGGGTCTAAAGATAGTCAGGCTGAAGGACAGTATAGGTTAATTGTAGGAGATCCAA
GTTCTTTCCAAGAGAAAGATGCGGATACTCTTCCCGGGAAGGTAGAGCAAAGTACTTTGTTCTCAGTAACCAATCCC
GTGGTTTTCCAAGGTGTGGACCAACAGGATCAAGTCTCTTCCCAAGGGTTAATTTGTAGTTTTACGAGCAGCAACCT
TGATTCTCCTCGTGACGGAGAATCTTTTTTAGGTATTGCTTTTGTTGGGGATAGTAGTAAGGCTGGAATCACATTAA
CTGACGTGAAAGCTTCTTTGTCTGGAGCGGCTTTATATTCTACAGAAGATCTTATCTTTGAAAAGATTAAGGGTGGA
TTGGAATTTGCATCATGTTCTTCTCTAGAACAGGGGGGAGCTTGTGCAGCTCAAAGTATTTTGATTCATGATTGTCA
AGGATTGCAGGTTAAACACTGTACTACAGCCGTGAATGCTGAGGGGTCTAGTGCGAATGATCATCTTGGATTTGGAG
GAGGCGCTTTCTTTGTTACGGGTTCTCTTTCTGGAGAGAAAAGTCTCTATATGCCTGCAGGAGATATGGTAGTTGCG
AATTGTGATGGGGCTATATCTTTTGAAGGAAACAGCGCGAACTTTGCTAATGGAGGAGCGATTGCTGCCTCTGGGAA
AGTGCTTTTTGTCGCTAATGATAAAAAGACTTCTTTTATAGAGAACCGAGCTTTGTCTGGAGGAGCGATTGCAGCCT
CTTCTGATATTGCCTTTCAAAACTGCGCAGAACTAGTTTTCAAAGGCAATTGTGCAATTGGAACAGAGGATAAAGGT
TCTTTAGGTGGAGGGGCTATATCTTCTCTAGGCACCGTTCTTTTGCAAGGGAATCACGGGATAACTTGTGATAAGAA
TGAGTCTGCTTCGCAAGGAGGCGCCATTTTTGGCAAAAATTGTCAGATTTCTGACAACGAGGGGCCAGTGGTTTTCA
GAGATAGTACAGCTTGCTTAGGAGGAGGCGCTATTGCAGCTCAAGAAATTGTTTCTATTCAGAACAATCAGGCTGGG
ATTTCCTTCGAGGGAGGTAAGGCTAGTTTCGGAGGAGGTATTGCGTGTGGATCTTTTTCTTCCGCAGGTGGTGCTTC
TGTTTTAGGGACCATTGATATTTCGAAGAATTTAGGCGCGATTTCGTTCTCTCGTACTTTATGTACGACCTCAGATT
TAGGACAAATGGAGTACCAGGGAGGAGGAGCTCTATTTGGTGAAAATATTTCTCTTTCTGAGAATGCTGGTGTGCTC
ACCTTTAAAGACAACATTGTGAAGACTTTTGCTTCGAATGGGAAAATTCTGGGAGGAGGAGCGATTTTAGCTACTGG
TAAGGTGGAAATTACTAATAATTCCGAAGGAATTTCTTTTACAGGAAATGCGAGAGCTCCACAAGCTCTTCCAACTC
AAGAGGAGTTTCCTTTATTCAGCAAAAAAGAAGGGCGACCACTCTCTTCAGGATATTCTGGGGGAGGAGCGATTTTA
GGAAGAGAAGTAGCTATTCTCCACAACGCTGCAGTAGTATTTGAGCAAAATCGTTTGCAGTGCAGCGAAGAAGAAGC
GACATTATTAGGTTGTTGTGGAGGAGGCGCTGTTCATGGGATGGATAGCACTTCGATTGTTGGCAACTCTTCAGTAA
GATTTGGTAATAATTACGCAATGGGACAAGGAGTCTCAGGAGGAGCTCTTTTATCTAAAACAGTGCAGTTAGCTGGG
AATGGAAGCGTCGATTTTTCTCGAAATATTGCTAGTTTGGGAGGAGGAGCTCTTCAAGCTTCTGAAGGAAATTGTGA
GCTAGTTGATAACGGCTATGTGCTATTCAGAGATAATCGAGGGAGGGTTTATGGGGGTGCTATTTCTTGCTTACGTG
GAGATGTAGTCATTTCTGGAAACAAGGGTAGAGTTGAATTTAAAGACAACATAGCAACACGTCTTTATGTGGAAGAA
ACTGTAGAAAAGGTTGAAGAGGTAGAGCCAGCTCCTGAGCAAAAAGACAATAATGAGCTTTCTTTCTTAGGGAGAGC
AGAACAGAGTTTTATTACTGCAGCTAATCAAGCTCTTTTCGCATCTGAAGATGGGGATTTATCACCTGAGTCATCCA
TTTCTTCTGAAGAACTTGCGAAAAGAAGAGAGTGTGCTGGAGGAGCTATTTTTGCAAAACGGGTTCGTATTGTAGAT
AACCAAGAGGCCGTTGTATTCTCGAATAACTTCTCTGATATTTATGGCGGCGCCATTTTTACAGGTTCTCTTCGAGA
AGAGGATAAGTTAGATGGGCAAATCCCTGAAGTCTTGATCTCAGGCAATGCAGGGGATGTTGTTTTTTCCGGAAATT
CCTCGAAGCGTGATGAGCATCTTCCTCATACAGGTGGGGGAGCCATTTGTACTCAAAATTTGACGATTTCTCAGAAT
ACAGGGAATGTTCTGTTTTATAACAACGTGGCCTGTTCGGGAGGAGCTGTTCGTATAGAGGATCATGGTAATGTTCT
TTTAGAAGCTTTTGGAGGAGATATTGTTTTTAAAGGAAATTCTTCTTTCAGAGCACAAGGATCCGATGCTATCTATT
TTGCAGGTAAAGAATCGCATATTACAGCCCTGAATGCTACGGAAGGACATGCTATTGTTTTCCACGACGCATTAGTT
TTTGAAAATCTAGAAGAAAGGAAATCTGCTGAAGTATTGTTAATCAATAGTCGAGAAAATCCAGGTTACACTGGATC
TATTCGATTTTTAGAAGCAGAAAGTAAAGTTCCTCAATGTATTCATGTACAACAAGGAAGCCTTGAGTTGCTAAATG
GAGCCACATTATGTAGTTATGGTTTTAAACAAGATGCTGGAGCTAAGTTGGTATTGGCTGCTGGAGCTAAACTGAAG
ATTTTAGATTCAGGAACTCCTGTACAACAAGGGCATGCTATCAGTAAACCTGAAGCAGAAATCGAGTCATCTTCTGA
ACCAGAGGGTGCACATTCTCTTTGGATTGCGAAGAATGCTCAAACAACAGTTCCTATGGTTGATATCCATACTATTT
CTGTAGATTTAGCCTCCTTCTCTTCTAGTCAACAGGAGGGGACAGTAGAAGCTCCTCAGGTTATTGTTCCTGGAGGA
AGTTATGTTCGATCTGGAGAGCTTAATTTGGAGTTAGTTAACACAACAGGTACTGGTTATGAAAATCATGCTTTATT
GAAGAATGAGGCTAAAGTTCCATTGATGTCTTTCGTTGCTTCTGGTGATGAAGCTTCAGCCGAAATCAGTAACTTGT
CGGTTTCTGATTTACAGATTCATGTAGTAACTCCAGAGATTGAAGAAGACACATACGGCCATATGGGAGATTGGTCT
GAGGCTAAAATTCAAGATGGAACTCTTGTCATTAGTTGGAATCCTACTGGATATCGATTAGATCCTCAAAAAGCAGG
GGCTTTAGTATTTAATGCATTATGGGAAGAAGGGGCTGTCTTGTCTGCTCTGAAAAATGCACGCTTTGCTCATAATC
TCACTGCTCAGCGTATGGAATTCGATTATTCTACAAATGTGTGGGGATTCGCCTTTGGTGGTTTCCGAACTCTATCT

GCAGAGAATCTGGTTGCTATTGATGGATACAAAGGAGCTTATGGTGGTGCTTCTGCTGGAGTCGATATTCAATTGAT
GGAAGATTTTGTTCTAGGAGTTAGTGGAGCTGCTTTCCTAGGTAAAATGGATAGTCAGAAGTTTGATGCGGAGGTTT
CTCGGAAGGGAGTTGTTGGTTCTGTATATACAGGATTTTTAGCTGGATCCTGGTTCTTCAAAGGACAATATAGCCTT
GGAGAAACACAGAACGATATGAAAACGCGTTATGGAGTACTAGGAGAGTCGAGTGCTTCTTGGACATCTCGAGGAGT
ACTGGCAGATGCTTTAGTTGAATACCGAAGTTTAGTTGGTCCTGTGAGACCTACTTTTTATGCTTTGCATTTCAATC
CTTATGTCGAAGTATCTTATGCTTCTATGAAATTCCCTGGCTTTACAGAACAAGGAAGAGAAGCGCGTTCTTTTGAA
GACGCTTCCCTTACCAATATCACCATTCCTTTAGGGATGAAGTTTGAATTGGCGTTCATAAAAGGACAGTTTTCAGA
GGTGAACTCTTTGGGAATAAGTTATGCATGGGAAGCTTATCGAAAAGTAGAAGGAGGCGCGGTGCAGCTTTTAGAAG
CTGGGTTTGATTGGGAGGGAGCTCCAATGGATCTTCCTAGACAGGAGCTGCGTGTCGCTCTGGAAAATAATACGGAA
TGGAGTTCTTACTTCAGCACAGTCTTAGGATTAACAGCTTTTTGTGGAGGATTTACTTCTACAGATAGTAAACTAGG
ATATGAGGCGAATACTGGATTGCGATTGATCTTT

SEQ ID NO:118 – CT812 fragment protein sequence

CVDLHAGGQSVNELVYVGPQAVLLLDQIRDLFVGSKDSQAEGQYRLIVGDPSSFQEKDADTLPGKVEQSTLFSVTNP
VVFQGVDQQDQVSSQGLICSFTSSNLDSPRDGESFLGIAFVGDSSKAGITLTDVKASLSGAALYSTEDLIFEKIKGG
LEFASCSSLEQGGACAAQSILIHDCQGLQVKHCTTAVNAEGSSANDHLGFGGGAFFVTGSLSGEKSLYMPAGDMVVA
NCDGAISFEGNSANFANGGAIAASGKVLFVANDKKTSFIENRALSGGAIAASSDIAFQNCAELVFKGNCAIGTEDKG
SLGGGAISSLGTVLLQGNHGITCDKNESASQGGAIFGKNCQISDNEGPVVFRDSTACLGGGAIAAQEIVSIQNNQAG
ISFEGGKASFGGGIACGSFSSAGGASVLGTIDISKNLGAISFSRTLCTTSDLGQMEYQGGGALFGENISLSENAGVL
TFKDNIVKTFASNGKILGGGAILATGKVEITNNSEGISFTGNARAPQALPTQEEFPLFSKKEGRPLSSGYSGGGAIL
GREVAILHNAAVVFEQNRLQCSEEEATLLGCCGGGAVHGMDSTSIVGNSSVRFGNNYAMGQGVSGGALLSKTVQLAG
NGSVDFSRNIASLGGGALQASEGNCELVDNGYVLFRDNRGRVYGGAISCLRGDVVISGNKGRVEFKDNIATRLYVEE
TVEKVEEVEPAPEQKDNNELSFLGRAEQSFITAANQALFASEDGDLSPESSISSEELAKRRECAGGAIFAKRVRIVD
NQEAVVFSNNFSDIYGGAIFTGSLREEDKLDGQIPEVLISGNAGDVVFSGNSSKRDEHLPHTGGGAICTQNLTISQN
TGNVLFYNNVACSGGAVRIEDHGNVLLEAFGGDIVFKGNSSFRAQGSDAIYFAGKESHITALNATEGHAIVFHDALV
FENLEERKSAEVLLINSRENPGYTGSIRFLEAESKVPQCIHVQQGSLELLNGATLCSYGFKQDAGAKLVLAAGAKLK
ILDSGTPVQQGHAISKPEAEIESSSEPEGAHSLWIAKNAQTTVPMVDIHTISVDLASFSSSQQEGTVEAPQVIVPGG
SYVRSGELNLELVNTTGTGYENHALLKNEAKVPLMSFVASGDEASAEISNLSVSDLQIHVVTPEIEEDTYGHMGDWS
EAKIQDGTLVISWNPTGYRLDPQKAGALVFNALWEEGAVLSALKNARFAHNLTAQRMEFDYSTNVWGFAFGGFRTLS
AENLVAIDGYKGAYGGASAGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRKGVVGSVYTGFLAGSWFFKGQYSL
GETQNDMKTRYGVLGESSASWTSRGVLADALVEYRSLVGPVRPTFYALHFNPYVEVSYASMKFPGFTEQGREARSFE
DASLTNITIPLGMKFELAFIKGQFSEVNSLGISYAWEAYRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTE
WSSYFSTVLGLTAFCGGFTSTDSKLGYEANTGLRLIF

SEQ ID NO:119 - CT812N nucleotide sequence

TGCGTAGATCTTCATGCTGGAGGACAGTCTGTAAATGAGCTGGTATATGTAGGCCCTCAAGCGGTTTTATTGTTAGA
CCAAATTCGAGATCTATTCGTTGGGTCTAAAGATAGTCAGGCTGAAGGACAGTATAGGTTAATTGTAGGAGATCCAA
GTTCTTTCCAAGAGAAAGATGCGGATACTCTTCCCGGGAAGGTAGAGCAAAGTACTTTGTTCTCAGTAACCAATCCC
GTGGTTTTCCAAGGTGTGGACCAACAGGATCAAGTCTCTTCCCAAGGGTTAATTTGTAGTTTTACGAGCAGCAACCT
TGATTCTCCTCGTGACGGAGAATCTTTTTTAGGTATTGCTTTTGTTGGGGATAGTAGTAAGGCTGGAATCACATTAA
CTGACGTGAAAGCTTCTTTGTCTGGAGCGGCTTTATATTCTACAGAAGATCTTATCTTTGAAAAGATTAAGGGTGGA
TTGGAATTTGCATCATGTTCTTCTCTAGAACAGGGGGGAGCTTGTGCAGCTCAAAGTATTTTGATTCATGATTGTCA
AGGATTGCAGGTTAAACACTGTACTACAGCCGTGAATGCTGAGGGGTCTAGTGCGAATGATCATCTTGGATTTGGAG
GAGGCGCTTTCTTTGTTACGGGTTCTCTTTCTGGAGAGAAAAGTCTCTATATGCCTGCAGGAGATATGGTAGTTGCG
AATTGTGATGGGGCTATATCTTTTGAAGGAAACAGCGCGAACTTTGCTAATGGAGGAGCGATTGCTGCCTCTGGGAA
AGTGCTTTTTGTCGCTAATGATAAAAAGACTTCTTTTATAGAGAACCGAGCTTTGTCTGGAGGAGCGATTGCAGCCT
CTTCTGATATTGCCTTTCAAAACTGCGCAGAACTAGTTTTCAAAGGCAATTGTGCAATTGGAACAGAGGATAAAGGT
TCTTTAGGTGGAGGGGCTATATCTTCTCTAGGCACCGTTCTTTTGCAAGGGAATCACGGGATAACTTGTGATAAGAA
TGAGTCTGCTTCGCAAGGAGGCGCCATTTTTGGCAAAAATTGTCAGATTTCTGACAACGAGGGGCCAGTGGTTTTCA
GAGATAGTACAGCTTGCTTAGGAGGAGGCGCTATTGCAGCTCAAGAAATTGTTTCTATTCAGAACAATCAGGCTGGG
ATTTCCTTCGAGGGAGGTAAGGCTAGTTTCGGAGGAGGTATTGCGTGTGGATCTTTTTCTTCCGCAGGTGGTGCTTC
TGTTTTAGGGACCATTGATATTTCGAAGAATTTAGGCGCGATTTCGTTCTCTCGTACTTTATGTACGACCTCAGATT
TAGGACAAATGGAGTACCAGGGAGGAGGAGCTCTATTTGGTGAAAATATTTCTCTTTCTGAGAATGCTGGTGTGCTC
ACCTTTAAAGACAACATTGTGAAGACTTTTGCTTCGAATGGGAAAATTCTGGGAGGAGGAGCGATTTTAGCTACTGG
TAAGGTGGAAATTACTAATAATTCCGAAGGAATTTCTTTTACAGGAAATGCGAGAGCTCCACAAGCTCTTCCAACTC
AAGAGGAGTTTCCTTTATTCAGCAAAAAAGAAGGGCGACCACTCTCTTCAGGATATTCTGGGGGAGGAGCGATTTTA
GGAAGAGAAGTAGCTATTCTCCACAACGCTGCAGTAGTATTTGAGCAAAATCGTTTGCAGTGCAGCGAAGAAGAAGC
GACATTATTAGGTTGTTGTGGAGGAGGCGCTGTTCATGGGATGGATAGCACTTCGATTGTTGGCAACTCTTCAGTAA
GATTTGGTAATAATTACGCAATGGGACAAGGAGTCTCAGGAGGAGCTCTTTTATCTAAAACAGTGCAGTTAGCTGGG
AATGGAAGCGTCGATTTTTCTCGAAATATTGCTAGTTTGGGAGGAGGAGCTCTTCAAGCTTCTGAAGGAAATTGTGA
GCTAGTTGATAACGGCTATGTGCTATTCAGAGATAATCGAGGGAGGGTTTATGGGGGTGCTATTTCTTGCTTACGTG
GAGATGTAGTCATTTCTGGAAACAAGGGTAGAGTTGAATTTAAAGACAACATAGCAACACGTCTTTATGTGGAAGAA
ACTGTAGAAAAGGTTGAAGAGGTAGAGCCAGCTCCTGAGCAAAAAGACAATAATGAGCTTTCTTTCTTAGGGAGAGC
AGAACAGAGTTTTATTACTGCAGCTAATCAAGCTCTTTTCGCATCTGAAGATGGGGATTTATCACCTGAGTCATCCA
TTTCTTCTGAAGAA

SEQ ID NO:120: CT812N protein sequence

CVDLHAGGQSVNELVYVGPQAVLLLDQIRDLFVGSKDSQAEGQYRLIVGDPSSFQEKDADTLPGKVEQSTLFSVTNP
VVFQGVDQQDQVSSQGLICSFTSSNLDSPRDGESFLGIAFVGDSSKAGITLTDVKASLSGAALYSTEDLIFEKIKGG
LEFASCSSLEQGGACAAQSILIHDCQGLQVKHCTTAVNAEGSSANDHLGFGGGAFFVTGSLSGEKSLYMPAGDMVVA
NCDGAISFEGNSANFANGGAIAASGKVLFVANDKKTSFIENRALSGGAIAASSDIAFQNCAELVFKGNCAIGTEDKG
SLGGGAISSLGTVLLQGNHGITCDKNESASQGGAIFGKNCQISDNEGPVVFRDSTACLGGGAIAAQEIVSIQNNQAG
ISFEGGKASFGGGIACGSFSSAGGASVLGTIDISKNLGAISFSRTLCTTSDLGQMEYQGGGALFGENISLSENAGVL
TFKDNIVKTFASNGKILGGGAILATGKVEITNNSEGISFTGNARAPQALPTQEEFPLFSKKEGRPLSSGYSGGGAIL
GREVAILHNAAVVFEQNRLQCSEEEATLLGCCGGGAVHGMDSTSIVGNSSVRFGNNYAMGQGVSGGALLSKTVQLAG
NGSVDFSRNIASLGGGALQASEGNCELVDNGYVLFRDNRGRVYGGAISCLRGDVVISGNKGRVEFKDNIATRLYVEE
TVEKVEEVEPAPEQKDNNELSFLGRAEQSFITAANQALFASEDGDLSPESSISSEE

SEQ ID NO:121: CT812C nucleotide sequence

GAAGAACTTGCGAAAAGAAGAGAGTGTGCTGGAGGAGCTATTTTTGCAAAACGGGTTCGTATTGTAGATAACCAAGA
GGCCGTTGTATTCTCGAATAACTTCTCTGATATTTATGGCGGCGCCATTTTTACAGGTTCTCTTCGAGAAGAGGATA
AGTTAGATGGGCAAATCCCTGAAGTCTTGATCTCAGGCAATGCAGGGGATGTTGTTTTTTCCGGAAATTCCTCGAAG
CGTGATGAGCATCTTCCTCATACAGGTGGGGGAGCCATTTGTACTCAAAATTTGACGATTTCTCAGAATACAGGGAA
TGTTCTGTTTTATAACAACGTGGCCTGTTCGGGAGGAGCTGTTCGTATAGAGGATCATGGTAATGTTCTTTTAGAAG
CTTTTGGAGGAGATATTGTTTTTAAAGGAAATTCTTCTTTCAGAGCACAAGGATCCGATGCTATCTATTTTGCAGGT
AAAGAATCGCATATTACAGCCCTGAATGCTACGGAAGGACATGCTATTGTTTTCCACGACGCATTAGTTTTTGAAAA
TCTAGAAGAAAGGAAATCTGCTGAAGTATTGTTAATCAATAGTCGAGAAAATCCAGGTTACACTGGATCTATTCGAT
TTTTAGAAGCAGAAAGTAAAGTTCCTCAATGTATTCATGTACAACAAGGAAGCCTTGAGTTGCTAAATGGAGCCACA
TTATGTAGTTATGGTTTTAAACAAGATGCTGGAGCTAAGTTGGTATTGGCTGCTGGAGCTAAACTGAAGATTTTAGA
TTCAGGAACTCCTGTACAACAAGGGCATGCTATCAGTAAACCTGAAGCAGAAATCGAGTCATCTTCTGAACCAGAGG
GTGCACATTCTCTTTGGATTGCGAAGAATGCTCAAACAACAGTTCCTATGGTTGATATCCATACTATTTCTGTAGAT
TTAGCCTCCTTCTCTTCTAGTCAACAGGAGGGGACAGTAGAAGCTCCTCAGGTTATTGTTCCTGGAGGAAGTTATGT
TCGATCTGGAGAGCTTAATTTGGAGTTAGTTAACACAACAGGTACTGGTTATGAAAATCATGCTTTATTGAAGAATG
AGGCTAAAGTTCCATTGATGTCTTTCGTTGCTTCTGGTGATGAAGCTTCAGCCGAAATCAGTAACTTGTCGGTTTCT
GATTTACAGATTCATGTAGTAACTCCAGAGATTGAAGAAGACACATACGGCCATATGGGAGATTGGTCTGAGGCTAA
AATTCAAGATGGAACTCTTGTCATTAGTTGGAATCCTACTGGATATCGATTAGATCCTCAAAAAGCAGGGGCTTTAG
TATTTAATGCATTATGGGAAGAAGGGGCTGTCTTGTCTGCTCTGAAAAATGCACGCTTTGCTCATAATCTCACTGCT
CAGCGTATGGAATTCGATTATTCTACAAATGTGTGGGGATTCGCCTTTGGTGGTTTCCGAACTCTATCTGCAGAGAA
TCTGGTTGCTATTGATGGATACAAAGGAGCTTATGGTGGTGCTTCTGCTGGAGTCGATATTCAATTGATGGAAGATT
TTGTTCTAGGAGTTAGTGGAGCTGCTTTCCTAGGTAAAATGGATAGTCAGAAGTTTGATGCGGAGGTTTCTCGGAAG
GGAGTTGTTGGTTCTGTATATACAGGATTTTTAGCTGGATCCTGGTTCTTCAAAGGACAATATAGCCTTGGAGAAAC
ACAGAACGATATGAAAACGCGTTATGGAGTACTAGGAGAGTCGAGTGCTTCTTGGACATCTCGAGGAGTACTGGCAG
ATGCTTTAGTTGAATACCGAAGTTTAGTTGGTCCTGTGAGACCTACTTTTTATGCTTTGCATTTCAATCCTTATGTC
GAAGTATCTTATGCTTCTATGAAATTCCCTGGCTTTACAGAACAAGGAAGAGAAGCGCGTTCTTTTGAAGACGCTTC
CCTTACCAATATCACCATTCCTTTAGGGATGAAGTTTGAATTGGCGTTCATAAAAGGACAGTTTTCAGAGGTGAACT
CTTTGGGAATAAGTTATGCATGGGAAGCTTATCGAAAAGTAGAAGGAGGCGCGGTGCAGCTTTTAGAAGCTGGGTTT
GATTGGGAGGGAGCTCCAATGGATCTTCCTAGACAGGAGCTGCGTGTCGCTCTGGAAAATAATACGGAATGGAGTTC
TTACTTCAGCACAGTCTTAGGATTAACAGCTTTTTGTGGAGGATTTACTTCTACAGATAGTAAACTAGGATATGAGG
CGAATACTGGATTGCGATTGATCTTT

SEQ ID NO:122: CT812C protein sequence

EELAKRRECAGGAIFAKRVRIVDNQEAVVFSNNFSDIYGGAIFTGSLREEDKLDGQIPEVLISGNAGDVVFSGNSSK
RDEHLPHTGGGAICTQNLTISQNTGNVLFYNNVACSGGAVRIEDHGNVLLEAFGGDIVFKGNSSFRAQGSDAIYFAG
KESHITALNATEGHAIVFHDALVFENLEERKSAEVLLINSRENPGYTGSIRFLEAESKVPQCIHVQQGSLELLNGAT
LCSYGFKQDAGAKLVLAAGAKLKILDSGTPVQQGHAISKPEAEIESSSEPEGAHSLWIAKNAQTTVPMVDIHTISVD
LASFSSSQQEGTVEAPQVIVPGGSYVRSGELNLELVNTTGTGYENHALLKNEAKVPLMSFVASGDEASAEISNLSVS
DLQIHVVTPEIEEDTYGHMGDWSEAKIQDGTLVISWNPTGYRLDPQKAGALVFNALWEEGAVLSALKNARFAHNLTA
QRMEFDYSTNVWGFAFGGFRTLSAENLVAIDGYKGAYGGASAGVDIQLMEDFVLGVSGAAFLGKMDSQKFDAEVSRK
GVVGSVYTGFLAGSWFFKGQYSLGETQNDMKTRYGVLGESSASWTSRGVLADALVEYRSLVGPVRPTFYALHFNPYV
EVSYASMKFPGFTEQGREARSFEDASLTNITIPLGMKFELAFIKGQFSEVNSLGISYAWEAYRKVEGGAVQLLEAGF
DWEGAPMDLPRQELRVALENNTEWSSYFSTVLGLTAFCGGFTSTDSKLGYEANTGLRLIF

SEQ ID NO:123: CT869 fragment nucleotide sequence

AGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTATCAAATAAAAT
TAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAA
AAACTCCCAATGAAGGAGCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT
TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCG
TGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGGAGAAATCCTTATGCTGCTG
ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATG
AAGAACTTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTT
TCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATT
CTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATC

TGTTCTCTAACAGGAAATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTC
AGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTA
GTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTAC
TTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGA
CCGCCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATC
CTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGTAGCCAAAATTTA
ATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTC
TGTAGTATTTTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCC
TTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGGGGT
GTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGATTCTGCTAGCAATGCCTCTAT
AACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGC
CTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATT
GATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTC
TGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAG
GACTTTGGACTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA
GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGTCC
CCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAGCTGACACCTA
GTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCT
GGATTCCATATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAG
TCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAA
TGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCAC
CATTTCTATACTCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTT
TGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCTA
GCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTC
CCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACC
CGTTCTGTATAGACAAGAACCAGGGATCGCAGCCCAGCTCCTAGCCAGTAAGGGTATTTGGTTCGGTAGTGGAAGCC
CCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAG
TATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTC

SEQ ID NO: 124: CT869 fragment protein sequence

REVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIY
FAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRLFTWRNPYAADKIREGGAIHAQNLYINHNHDVVGFM
KNFSYVQGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPI
CSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTY
FINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL
IFYDPIEVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGG
VVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLI
DDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQK
ANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHP
GFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCH
HFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLV
PIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIAAQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQ
YHGFYSSSTFCNYLNGEIALRF

SEQ ID NO:125: CT166 fragment nucleotide sequence

AACGTTCGTACGTACTCTGTTCAGAGGGGGGGGGTAAAAACGATTTCTGCTAGTGCAGTTCCTCCTACAGCAGCTGT
TTTATCGAGAAAAAAGCGTGCTATAGAAGAGAAGAAGGAGGAAGCTTCTTCTGGAAAGATAGAAAATCTTGATGCTA
GCAAATACGATCTTACTCCCAAGAACATAGAAGAAAAACTAGGAATTACTCCTGAACAGAAATCTACTGTTAAAGAC
CTATTAAATAAACTGAAAAAGGTCATTAGTGCTTACAACTCTATGCCAGATAAAAATTCGGAAGCGGGACAGAATTC
CTTGATTCAACAAGGAAAATACGTCGATGCCATTCAGAAGAAGCTTCCAGCATCATCGCAGGCTCAGCCTAAACAGG
CAAAAGCTAAGGAACAGAAAGCCGAAGAAAAACCTAAGACGACTCCGATTGAAGGTGTTCTTGAAACCATCAAAACA
GAATTTAAAGGCCATCGTGTACCTGTTGAGAAAATCATCCATGGAATATGGATCGCAGGAGCGCCTCCGGATGGTAT
CGAAGATTATATGCGAGTCTTTTTAGATACTTATGAAGGTTTTGACTTCTACTTCTGGGTAGATGAGAATGCTTATG
CAGCAGCTAAATTTTCTAGCATTTTGAAGAAGGTCGCTTTCGATGCGGCTATTCAAGATCTACGATCTGCCACAGAT
GAGTCTACGAAGGCCTTTGTTAAAGACTACGATGAATTAAAACAGAAATATGAAAAGAAAGTTGCGGAGACGACTTC
TCAAGCAGAAAAAGACCAATATCTCAAAGATCTAAAGGATCTTTTAGAGAAATTTACAAAAATCAGTGATGAGATTC
GTGGAAAATTTGATCGGCTGTTTCTTAAGAATGTGATTGTTGCTCAGAACGGATTCTTTAATTTCTGCTTGCTGAAA
GGCCTCGGCAATATCAATGACGAAACGCGTGCAGAGTATTTAGAGAAAGAACTCAAACTTCCTACTGAGGAGATCGA
ACAGTATAAAAAGCTTAAAGAGACGAACAAAGAGAAGATAGCCGCTATTGTAAAACAACTAAACGAGAAACTTGGAT
CGGATCGGGTAAAAATCAAAGACATTAAAGAGCTGCAATCTATGAAGCAAGCTCGAAATGTCTACAATTATGAACAG
GAAATGTTTCTGCGCTGGAACTATGCAGCCGCAACAGATCAGATTCGTATGTATATGTTGGAGGAACTTGGAGGTCT
TTATACTGATCTGGATATGATGCCTTCATACTCTCAGGAAGTATTGGAGCTTATCAAAAAGCACAGTGATGGAAACC
GAATGTTTGAGGATATGAGCTCTAGACGGGCGATTTCTGATGCGGTTTTAAAGATGGCTGTAGGTAAGGCGACAACA
GTTTCCATGGAAGAGGTAGCAAAGGATATCGATGTTTCTCGCTTAACAGAAGAGGATAAGACAAAATTAAATGCTCT
ATTTAAGGATCTAGAGCCATTTGCAAAACCGGATTCTAAAGGAGCTGAAGCAGAAGGGGGTGAAGGAGCAAAAGGTA
TGAAAAAGAGCTTTTTCCAGCCCATAGATCTGAATATTGTCAGAAATACCATGCCTATCTTGAGACGCTATCATCAC
TATCCTGAGTTAGGATGGTTTATTCGAGGATTGAACGGATTGATGGTCTCTCATAAGGGAAGCACTGCGGTTTCTGC
TGTCATTGTAGGGCAACAGGCTGCCTACCAGGAACTAGCAGCACTTAGACAAGATGTCCTTTCAGGGGAGTTTTTCC

ATTCTTTAGAAAATTTGACACATAGAAACCATAAGGAGCGTATTGGAAATCATCTCGTCGCTAATTATTTGGCTAAA
AGTCTCTTTTTTGATTACTGCCAAGATTCAGTGATGCCGGAGGCTGTAAGTACCTTAGGTATTAGA

SEQ ID NO: 126 – CT166 fragment protein sequence

NVRTYSVQRGGVKTISASAVPPTAAVLSRKKRAIEEKKEEASSGKIENLDASKYDLTPKNIEEKLGITPEQKSTVKD
LLNKLKKVISAYNSMPDKNSEAGQNSLIQQGKYVDAIQKKLPASSQAQPKQAKAKEQKAEEKPKTTPIEGVLETIKT
EFKGHRVPVEKIIHGIWIAGAPPDGIEDYMRVFLDTYEGFDFYFWVDENAYAAAKFSSILKKVAFDAAIQDLRSATD
ESTKAFVKDYDELKQKYEKKVAETTSQAEKDQYLKDLKDLLEKFTKISDEIRGKFDRLFLKNVIVAQNGFFNFCLLK
GLGNINDETRAEYLEKELKLPTEEIEQYKKLKETNKEKIAAIVKQLNEKLGSDRVKIKDIKELQSMKQARNVYNYEQ
EMFLRWNYAAATDQIRMYMLEELGGLYTDLDMMPSYSQEVLELIKKHSDGNRMFEDMSSRRAISDAVLKMAVGKATT
VSMEEVAKDIDVSRLTEEDKTKLNALFKDLEPFAKPDSKGAEAEGGEGAKGMKKSFFQPIDLNIVRNTMPILRRYHH
YPELGWFIRGLNGLMVSHKGSTAVSAVIVGQQAAYQELAALRQDVLSGEFFHSLENLTHRNHKERIGNHLVANYLAK
SLFFDYCQDSVMPEAVSTLGIR

SEQ ID NO:127 – CT175 fragment nucleotide sequence

TGTTATCATAAAAAAGAAGAACCAAAAGATGTTTTGCGGATTGCGATCTGTCATGATCCAATGTCTTTAGATCCGCG
TCAGGTTTTTTTAAGCAAAGATGTTTCTATTGTAAAAGCTCTCTATGAAGGGTTAGTCCGGGAAAAAGAAGCTGCGT
TCCAGCTAGCTTTGGCAGAAAGATATCATCAATCTGATGATGGTTGTGTTTATACTTTTTTTCTAAAAAATACATTC
TGGAGCAACGGAGATGTTGTAACAGCATATGATTTTGAAGAGTCTATTAAACAAATTTATTTCCGAGAAATTGATAA
CCCTTCGTTACGCTCTCTTGCATTAATTAAAAATTCTCATGCTGTTTTAACAGGAGCTCTCCCTGTTGAAGATTTAG
GTGTTAGAGCTTTGAATGCGAAAACTCTAGAAATTGTTTTAGAAAACCCGTTTCCTTATTTTCTAGAGATATTGGCG
CACCCGGTTTTTTATCCGGTGCACACCTCTTTACGAGAATATTACAAAGATAAGCGTAACAAACGCGTTTTCCCGAT
AATTTCTAATGGTCCTTTTGCGATTCAATGTTATGAGCCGCAAAGATATTTACTAATCAACAAAAACCCTCTGTATC
ATGCCAAGCACGATGTTCTGTTAAATTCGGTATGTTTGCAGATAGTTCCTGATATCCATACAGCTATGCAGTTATTC
CAAAAAAATCATATCGATTTAGTTGGGTTACCCTGGAGCTCCTCCTTTTCTTTAGAAGAACAAAGAAATCTCCCTAG
AGAAAAATTATTTGATTATCCTGTATTGAGTTGCTCTGTTTTATTCTGTAACATTCATCAAACACCTTTAAATAATC
CCTCGCTGAGAACAGCCCTCTCTTTAGCAATCAATCGAGAAACTTTATTAAAACTAGCAGGTAAAGGCTGTAGCGCT
ACGAGCTTTGTTCACCCACAATTATCTCAGATACCTGCTACTACTTTGTCTCAAGATGAGCGGATTGCTTTAGCAAA
AGGCTACTTGACCGAAGCTTTAAAGACTTTATCTCAAGAAGATTTAGAAAAAATTACATTAATTTATCCTATAGAAT
CTGTTTGCTTACGAGCCGTTGTTCAAGAAATTCGCCAACAATTATTTGATGTACTGGGATTTAAAATTTCTACATTA
GGATTAGAATATCATTGTTTTTTAGACAAACGTTCCAGAGGAGAATTCTCCTTAGCAACTGGTAATTGGATTGCAGA
CTATCATCAAGCTAGTGCTTTCCTGTCTGTCCTAGGTAATGGGACAAGATATAAAGACTTTCAATTGATTAACTGGC
AGAACCAAAAGTACACAAATATAGTTGCTCAACTTCTGATTCAAGAATCAAGCGACCTACAGCTTATGGCAGAGCAG
TTGTTGCTTAAAGAAAGTCCTCTTATTCCTCTATACCACCTCGATTATGTGTATGCGAAACAGCCTCGGGTGTCTGA
TCTCCAAACCTCTTCTCGTGGAGAAATTGATTTAAAAAGAGTTTCATTAGCTGAAGGATAG

SEQ ID NO:128 – CT175 fragment protein sequence

CYHKKEEPKDVLRIAICHDPMSLDPRQVFLSKDVSIVKALYEGLVREKEAAFQLALAERYHQSDDGCVYTFFLKNTF
WSNGDVVTAYDFEESIKQIYFREIDNPSLRSLALIKNSHAVLTGALPVEDLGVRALNAKTLEIVLENPFPYFLEILA
HPVFYPVHTSLREYYKDKRNKRVFPIISNGPFAIQCYEPQRYLLINKNPLYHAKHDVLLNSVCLQIVPDIHTAMQLF
QKNHIDLVGLPWSSSFSLEEQRNLPREKLFDYPVLSCSVLFCNIHQTPLNNPSLRTALSLAINRETLLKLAGKGCSA
TSFVHPQLSQIPATTLSQDERIALAKGYLTEALKTLSQEDLEKITLIYPIESVCLRAVVQEIRQQLFDVLGFKISTL
GLEYHCFLDKRSRGEFSLATGNWIADYHQASAFLSVLGNGTRYKDFQLINWQNQKYTNIVAQLLIQESSDLQLMAEQ
LLLKESPLIPLYHLDYVYAKQPRVSDLQTSSRGEIDLKRVSLAEG

SEQ ID NO: 129 – TC0666 fragment nucleotide sequence (homologue of CT387)

ATGACACTCTTTCACACTCATCACGATGCCGTCTCTCCGGACGGCTACTTATGTTCTTCCCTTCAGTTAGTTGGCTC
TGGCACATATGAAGGAGAAATCGAAATCCAAAATATTCCTTCTTATTTCCTTGGATTCCGATTACCCACCCATTGCG
TTCATCTTAATTTGAAGAGTTCTCTAGCCCAGTTAGGAGTAGATGCATCTCTTCTTCACTGCGAACTAAGCAAAAAT
CAACAACGTGCACATATGCACGTGCAGTTCACCGGCTATGGCCCTATCGCTGAGTCCATGCTATCTCTTCTCAAACC
CGGAGATCGAGTAGCCAAACTGTTTGCTGCAGATGATCGTAGACTAGTCCGCTCCCCTGATTATCTTGAAAGCATGC
TAAAAAATACTGATAAGACAGGACATCCTCTGCTCCGATTTGGAAAAAAACTCGAGCATCTTATCTCTTTTGATGTG
GTGGACGATCGCCTCGTTGTATCACTCCCCACCTTGCCAGGCATAGTCAATTATGACCCAGACATCTATGGACTTCT
TCCCTTAATTCAAAAATCACTAAGCAATCCTAAATTGAGTATTCGCCACTTCTTGTCTCTCTATCAGAAGATCGTAG
AAGGACCACACATCCCTTATGAAGGAAACATTTTGTTAATCAAAACAGAGCCTCTTCATATCCGCACAGTATTTGCT
CGCGTGGTCGATCAAATGCTCCCTCAAGGTCTATTTCACACTTCTGCCAACATTTTAGAACCCACAACGCGAGAGTC
TGGAGATATTTTTGAATTTTTTGGAAATCCCTCCACTCTTGTAGAAAGAATCCCTCTAGAATTCTTCACTATCGAAC
CCTACAAAGAACACTCTTACTTCTGTAATCGAGATCTATTGCAAACTACCTTGCAATCGGAAAGTGAAATCAAAAAA
ATATTCGATACAGCTCCTCAAGAGCCTGTAAAAGCCGCCACTTATTTATCAAAAGGAAGTGAAATTTCTTCTCTTGA
TGCAGATTCTTGGCTTACGGGATCCGCAGCTGCATACCAATGTAGCGAAAAACAGGCAGCTAAAGACGAATACATCC
ACGCTCAACCCTGTTATCCATTTTTGGAAGCAATGGAAACGGGACTCATCAATAGCGAAGGAGCTTTACTCACTCGG
TTTTTCCCCTCTTCCAGCTTAAAAGGGATGTTGATCTCCTATCATGTACGCCACTATCTTAAGCAAATTTACTTTCA
AGTTCCTTCTTATACATATGGAGACTACTTCTCTCATAATGACCGAGGATTACTGTTAGATCTATATCAGGCGAACA
TTGATGTGTTCTGGGCTGATGAAGAGAGCGGCCGTGTATTGCAATATACAAAACGGCGCGACAAAAATAGTGGAATG
TTCGTCGTTAAAAATCGAGTAGAAGAGTTCCAATCAGCATATTTCGTAGCGATTTATGGATCACGTCTCCTGGAAAA
TAATTTCTCGGCCCAACTAAACACGCTTCTTGCAGGGTTACAAAAAGCTGCACACACTCTAGGCATTCCAGGCTTCT
CAAAACCCACTCCTCTTGCCGTAATCACAGGAGGAGGGACTGGCGTTATGGCTACAGGAAATCGTGTTGCAAAAGAG

TTGGGAATTCTTTCTTGCGGGACCGTTCTCGATTTGGAAGCTTCACCTGCACAAATAGATCAGCCTGCAAACGAATT
TTTAGATGCCAAAATGACATACCGTCTACCGCAACTTATAGAAAGACAAGAACATTTTTATTCAGACCTTGCCATTT
TAGTTGTTGGTGGTGTTGGAACAGATTTCGAACTTTACCTAGAACTCGTCTACTTGAAAACAGGCGCCAAACCTCCT
ACTCCAATTTTCCTTATTGGGCCTGTTGAATACTGGAAAGAGAAAGTTGCTCATGCCTATGAGATTAATCTTAAAGC
AGGAACTATTCGTGGTTCTGAGTGGATCAGCAACTGCTTATTCTGCATTACATCTCCTGAAGCAGGAATTGCTGTAT
TCGAACAGTTCCTCGCTGGAGAACTTCCCATAGGATATGATTATCCTCCAGCTCCAGACGGATTAGTTATCGTC

SEQ ID NO: 130 - TC0666 fragment protein sequence (homologue of CT387)

MTLFHTHHDAVSPDGYLCSSLQLVGSGTYEGEIEIQNIPSYFLGFRLPTHCVHLNLKSSLAQLGVDASLLHCELSKN
QQRAHMHVQFTGYGPIAESMLSLLKPGDRVAKLFAADDRRLVRSPDYLESMLKNTDKTGHPLLRFGKKLEHLISFDV
VDDRLVVSLPTLPGIVNYDPDIYGLLPLIQKSLSNPKLSIRHFLSLYQKIVEGPHIPYEGNILLIKTEPLHIRTVFA
RVVDQMLPQGLFHTSANILEPTTRESGDIFEFFGNPSTLVERIPLEFFTIEPYKEHSYFCNRDLLQTTLQSESEIKK
IFDTAPQEPVKAATYLSKGSEISSLDADSWLTGSAAAYQCSEKQAAKDEYIHAQPCYPFLEAMETGLINSEGALLTR
FFPSSSLKGMLISYHVRHYLKQIYFQVPSYTYGDYFSHNDRGLLLDLYQANIDVFWADEESGRVLQYTKRRDKNSGM
FVVKNRVEEFQSAYFVAIYGSRLLENNFSAQLNTLLAGLQKAAHTLGIPGFSKPTPLAVITGGGTGVMATGNRVAKE
LGILSCGTVLDLEASPAQIDQPANEFLDAKMTYRLPQLIERQEHFYSDLAILVVGGVGTDFELYLELVYLKTGAKPP
TPIFLIGPVEYWKEKVAHAYEINLKAGTIRGSEWISNCLFCITSPEAGIAVFEQFLAGELPIGYDYPPAPDGLVIV

SEQ ID NO: 131 - TC0197 fragment nucleotide sequence

AATTGTTCCGATCTTTATGCCGTAGGAAGTTCTGCAGACCATCCTGCCTACTTGATTCCTCAAGCGGGGTTATTATT
GGATCATATTAAGGATATATTCATTGGCCCTAAAGATAGTCAGGATAAGGGGCAGTATAAGTTGATTATTGGTGAGG
CTGGCTCTTTCCAAGATAGTAATGCAGAGACTCTTCCTCAAAAGGTAGAGCACAGCACTTTGTTTTCAGTTACAACA
CCTATAATTGTGCAAGGAATAGATCAACAAGATCAGGTCTCTTCGCAGGGATTGGTCTGTAATTTTTCAGGAGATCA
TTCAGAGGAGATTTTTGAGAGAGAATCCTTTTTAGGGATCGCTTTCCTAGGGAATGGTAGCAAGGATGGAATCACGT
TAACAGATATAAAATCTTCGTTATCTGGTGCTGCCTTGTATTCTTCAGATGATCTTATTTTTGAAAGAATTAAGGGA
GATATAGAGCTTTCTTCTTGTTCATCTTTAGAAAGAGGAGGAGCTTGTTCAGCTCAAAGTATTTTAATTCATGATTG
TCAAGGATTAACGGTAAAACATTGTGCCGCAGGGGTGAATGTTGAAGGAGTTAGTGCTAGCGACCATCTCGGATTTG
GGGGCGGGGCCTTCTCTACTACAAGTTCTCTTTCTGGAGAGAAGAGTTTGTATATGCCTGCAGGCGATATTGTGGTG
GCTACCTGCGATGGTCCTGTGTGTTTCGAAGGAAATAGTGCTCAGTTAGCAAATGGTGGCGCTATTGCCGCTTCTGG
TAAAGTTCTTTTTGTAGCTAACGAAAAAAAGATTTCCTTTACAGACAACCAAGCTTTGTCTGGAGGAGCTATTTCTG
CATCTTCTAGTATTTCTTTCCAAAATTGTGCTGAGCTTGTGTTCAAGAGTAATCTTGCAAAAGGAGTTAAAGATAAA
TGTTCTTTGGGAGGAGGTGCTTTAGCCTCTTTAGAATCCGTAGTTTTGAAAGATAATCTCGGTATTACTTATGAAAA
AAATCAGTCCTATTCGGAAGGAGGGGCTATTTTTGGGAAGGATTGTGAGATTTTTGAAAACAGGGGGCCTGTTGTAT
TCAGAGATAATACAGCTGCTTTAGGAGGCGGAGCTATTTTGGCGCAACAAACTGTGGCGATTTGTGGTAATAAGTCT
GGAATATCTTTTGAAGGAAGTAAGTCTAGTTTTGGAGGGGCCATTGCTTGTGGAAATTTCTCTTCTGAGAATAATTC
TTCAGCTTTGGGATCAATTGATATCTCTAACAATCTAGGAGATATCTCTTTTCTTCGGACTCTGTGTACTACTTCGG
ATTTAGGGCAAACGGATTACCAAGGGGGAGGGGCCTTATTCGCTGAAAATATTTCTCTTTCTGAGAATGCTGGTGCA
ATTACTTTCAAAGACAATATTGTGAAGACATTTGCCTCAAATGGAAAAATGTTGGGTGGAGGGGCAATTTTAGCTTC
AGGAAATGTTTTGATTAGCAAAAACTCTGGAGAGATTTCTTTTGTAGGGAATGCTCGAGCTCCTCAGGCTATTCCGA
CTCGTTCATCTGACGAATTGTCTTTTGGCGCACAATTAACTCAAACTACTTCAGGATGTTCTGGAGGAGGAGCTCTT
TTTGGTAAAGAGGTTGCCATTGTTCAAAATGCCACTGTTGTATTCGAGCAAAATCGCTTACAGTGTGGCGAGCAGGA
AACACATGGTGGAGGCGGTGCTGTTTATGGTATGGAGAGTGCCTCTATTATTGGAAACTCTTTTGTGAGATTCGGAA
ATAATTACGCTGTAGGGAATCAGATTTCTGGAGGAGCTCTTTTATCCAAGAAGGTCCGTTTAGCTGAAAATACAAGG
GTAGATTTTTCTCGAAATATCGCTACTTTCTGCGGCGGGGCTGTTCAAGTTTCTGATGGAAGTTGCGAATTGATCAA
CAATGGGTATGTGCTATTCAGAGATAACCGAGGGCAGACATTTGGTGGGGCTATTTCTTGCTTGAAAGGAGATGTGA
TCATTTCCGGAAATAAAGATAGGGTTGAGTTTAGAGATAACATTGTGACGCGGCCTTATTTTGAAGAAAATGAAGAA
AAAGTTGAGACAGCAGATATTAATTCAGATAAGCAAGAAGCAGAAGAGCGCTCTTTATTAGAGAACATTGAGCAGAG
CTTTATTACTGCAACTAATCAGACCTTTTTCTTAGAGGAAGAGAAACTCCCATCAGAAGCTTTTATCTCTGCTGAAG
AACTTTCAAAGAGAAGAGAATGTGCTGGTGGGGCGATTTTTGCAAAACGGGTCTACATTACGGATAATAAAGAACCT
ATCTTGTTTTCGCATAATTTTTCTGATGTTTATGGGGGAGCTATTTTTACGGGTTCTCTACAGGAAACTGATAAACA
AGATGTTGTAACTCCTGAAGTTGTGATATCAGGCAACGATGGGGATGTCATTTTTTCTGGAAATGCAGCTAAACATG
ATAAGCATTTACCTGATACAGGTGGTGGAGCCATTTGTACACAGAATTTGACGATTTCCCAAAACAATGGGAATGTC
TTGTTCTTGAACAATTTTGCTTGTTCTGGTGGAGCAGTTCGCATAGAGGATCATGGAGAAGTTCTTTTAGAGGCTTT
TGGGGGAGATATTATTTTCAATGGAAACTCTTCTTTCAGAGCTCAAGGATCGGATGCGATCTATTTTGCTGGTAAGG
ACTCTAGAATTAAAGCTTTAAATGCTACTGAAGGACATGCGATTGTGTTCCAAGATGCATTGGTGTTTGAAAATATA
GAAGAAAGAAAGTCTTCGGGACTATTGGTGATTAACTCTCAGGAAAATGAGGGTTATACGGGATCCGTCCGATTTTT
AGGATCTGAAAGTAAGGTTCCTCAATGGATTCATGTGCAACAGGGAGGTCTTGAGTTGCTACATGGAGCTATTTTAT
GTAGTTATGGGGTTAAACAAGATCCTAGAGCTAAAATAGTATTATCTGCTGGATCTAAATTGAAGATTCTAGATTCA
GAGCAAGAAAATAACGCAGAAATTGGAGATCTTGAAGATTCTGTTAATTCAGAAAAAACACCATCTCTTTGGATTGG
GAAGAACGCTCAAGCAAAAGTCCCTCTGGTTGATATCCATACTATTTCTATTGATTTAGCATCATTTTCTTCTAAAG
CTCAGGAAACCCCTGAGGAAGCTCCACAAGTCATCGTCCCTAAGGGAAGTTGTGTCCACTCGGGAGAGTTAAGTTTG
GAGTTGGTTAATACAACAGGAAAAGGTTATGAGAATCATGCGTTGTTAAAAAATGATACTCAGGTTTCTCTCATGTC
TTTCAAAGAGGAAAATGATGGATCTTTAGAAGATTTGAGTAAGTTGTCTGTTTCGGATTTACGCATTAAAGTTTCTA
CTCCAGATATTGTAGAAGAAACTTATGGCCATATGGGGGATTGGTCTGAAGCTACAATTCAAGATGGGGCTCTTGTC
ATTAATTGGCATCCTACTGGATATAAATTAGATCCGCAAAAAGCTGGTTCTTTGGTATTCAATGCATTATGGGAGGA
AGAGGCTGTATTGTCTACTCTAAAAAATGCTCGGATTGCCCATAACCTTACCATTCAGAGAATGGAATTTGATTATT
CTACAAATGCTTGGGGATTAGCTTTTAGTAGCTTTAGAGAGCTATCTTCAGAGAAGCTTGTTTCTGTTGATGGATAT
AGAGGCTCTTATATAGGGGCTTCTGCAGGCATTGATACTCAGTTGATGGAAGATTTTGTTTTGGGAATCAGCACGGC

TTCCTTCTTCGGGAAAATGCATAGTCAGAATTTTGATGCAGAGATTTCTCGACATGGTTTTGTTGGTTCGGTCTATA
CAGGCTTCCTAGCTGGGGCCTGGTTCTTCAAGGGGCAGTACAGTCTTGGCGAAACACATAACGATATGACAACTCGT
TACGGGGTTTTGGGAGAATCTAATGCTACTTGGAAGTCTCGAGGAGTACTAGCAGATGCTTTAGTTGAATATCGTAG
TTTAGTCGGTCCAGCACGACCTAAATTTTATGCTTTGCATTTTAATCCTTATGTCGAGGTATCTTATGCATCTGCGA
AGTTCCCTAGTTTTGTAGAACAAGGAGGAGAAGCTCGTGCTTTTGAAGAAACCTCTTTAACAAACATTACCGTTCCC
TTTGGTATGAAATTTGAACTATCTTTTACAAAAGGACAGTTTTCAGAGACTAATTCTCTTGGAATAGGTTGTGCATG
GGAAATGTATCGGAAAGTCGAAGGAAGATCTGTAGAGCTACTAGAAGCTGGTTTTGATTGGGAAGGATCTCCTATAG
ATCTCCCTAAACAAGAGCTGAGAGTGGCTTTAGAAAACAATACGGAATGGAGTTCGTATTTTAGTACAGCTCTAGGA
GTAACAGCATTTTGTGGAGGATTTTCTTCTATGGATAATAAACTAGGATACGAAGCGAATGCTGGAATGCGTTTGAT
TTTCTAG

SEQ ID NO: 132 - TC0197 fragment protein sequence

NCSDLYAVGSSADHPAYLIPQAGLLLDHIKDIFIGPKDSQDKGQYKLIIGEAGSFQDSNAETLPQKVEHSTLFSVTT
PIIVQGIDQQDQVSSQGLVCNFSGDHSEEIFERESFLGIAFLGNGSKDGITLTDIKSSLSGAALYSSDDLIFERIKG
DIELSSCSSLERGGACSAQSILIHDCQGLTVKHCAAGVNVEGVSASDHLGFGGGAFSTTSSLSGEKSLYMPAGDIVV
ATCDGPVCFEGNSAQLANGGAIAASGKVLFVANEKKISFTDNQALSGGAISASSSISFQNCAELVFKSNLAKGVKDK
CSLGGGALASLESVVLKDNLGITYEKNQSYSEGGAIFGKDCEIFENRGPVVFRDNTAALGGGAILAQQTVAICGNKS
GISFEGSKSSFGGAIACGNFSSENNSSALGSIDISNNLGDISFLRTLCTTSDLGQTDYQGGGALFAENISLSENAGA
ITFKDNIVKTFASNGKMLGGGAILASGNVLISKNSGEISFVGNARAPQAIPTRSSDELSFGAQLTQTTSGCSGGGAL
FGKEVAIVQNATVVFEQNRLQCGEQETHGGGGAVYGMESASIIGNSFVRFGNNYAVGNQISGGALLSKKVRLAENTR
VDFSRNIATFCGGAVQVSDGSCELINNGYVLFRDNRGQTFGGAISCLKGDVIISGNKDRVEFRDNIVTRPYFEENEE
KVETADINSDKQEAEERSLLENIEQSFITATNQTFFLEEEKLPSEAFISAEELSKRRECAGGAIFAKRVYITDNKEP
ILFSHNFSDVYGGAIFTGSLQETDKQDVVTPEVVISGNDGDVIFSGNAAKHDKHLPDTGGGAICTQNLTISQNNGNV
LFLNNFACSGGAVRIEDHGEVLLEAFGGDIIFNGNSSFRAQGSDAIYFAGKDSRIKALNATEGHAIVFQDALVFENI
EERKSSGLLVINSQENEGYTGSVRFLGSESKVPQWIHVQQGGLELLHGAILCSYGVKQDPRAKIVLSAGSKLKILDS
EQENNAEIGDLEDSVNSEKTPSLWIGKNAQAKVPLVDIHTISIDLASFSSKAQETPEEAPQVIVPKGSCVHSGELSL
ELVNTTGKGYENHALLKNDTQVSLMSFKEENDGSLEDLSKLSVSDLRIKVSTPDIVEETYGHMGDWSEATIQDGALV
INWHPTGYKLDPQKAGSLVFNALWEEEAVLSTLKNARIAHNLTIQRMEFDYSTNAWGLAFSSFRELSSEKLVSVDGY
RGSYIGASAGIDTQLMEDFVLGISTASFFGKMHSQNFDAEISRHGFVGSVYTGFLAGAWFFKGQYSLGETHNDMTTR
YGVLGESNATWKSRGVLADALVEYRSLVGPARPKFYALHFNPYVEVSYASAKFPSFVEQGGEARAFEETSLTNITVP
FGMKFELSFTKGQFSETNSLGIGCAWEMYRKVEGRSVELLEAGFDWEGSPIDLPKQELRVALENNTEWSSYFSTALG
VTAFCGGFSSMDNKLGYEANAGMRLIF

SEQ ID NO: 133 - TC0261 fragment nucleotide sequence

ACTCGAGAAGTCCCTCCTTCGATTCTTTTAAAGCCTATACTAAATCCATACCATATGACCGGGTTATTTTTTCCCAA
GGTTAATTTGCTTGGAGACACACATAATCTCACTGATTACCATTTGGATAATCTAAAATGCATTCTGGCTTGCCTAC
AAAGAACTCCTTATGAAGGAGCTGCTTTCACAGTAACCGATTACTTAGGTTTTTCAGATACACAAAAGGATGGTATT
TTTTGTTTTAAAAATCTTACTCCAGAGAGTGGAGGGGTTATTGGTTCCCCAACTCAAAACACTCCTACTATAAAAAT
TCATAATACAATCGGCCCCGTTCTTTTCGAAAATAATACCTGTCATAGACTGTGGACACAGACCGATCCCGAAAATG
AAGGAAACAAAGCACGCGAAGGCGGGGCAATTCATGCTGGGGACGTTTACATAAGCAATAACCAGAACCTTGTCGGA
TTCATAAAGAACTTTGCTTATGTTCAAGGTGGAGCTATTAGTGCTAATACTTTTGCCTATAAAGAAAATAAATCGAG
CTTTCTTTGCCTAAATAACTCTTGTATACAAACTAAGACGGGAGGGAAAGGTGGTGCTATTTACGTTAGTACGAGCT
GCTCTTTCGAGAACAATAACAAGGATCTGCTTTTCATCCAAAACTCCGGCTGTGCAGGAGGAGCTATCTTCTCTCCA
ACCTGTTCTCTAATAGGAAACCAAGGAGATATTGTTTTTTACAGCAACCACGGTTTTAAAAATGTTGATAATGCAAC
TAACGAATCTGGGGATGGAGGAGCTATTAAAGTAACTACCCGCTTGGACATCACCAATAATGGTAGTCAAATCTTTT
TTTCTGATAATATCTCAAGAAATTTTGGAGGAGCTATTCATGCTCCTTGTCTTCATCTTGTTGGTAATGGGCCAACC
TATTTTACAAACAATATAGCTAATCACACAGGTGGGGCTATTTATATAACAGGAACAGAAACCTCAAAGATTTCTGC
AGATCACCATGCTATTATTTTTGATAATAACATTTCTGCAAACGCCACCAATGCGGACGGATCTAGCAGCAACACTA
ATCCTCCTCACAGAAATGCGATCACTATGGACAATTCCGCTGGAGGAATAGAACTTGGTGCAGGGAAGAGCCAGAAT
CTTATTTTCTATGATCCTATTCAAGTGACGAATGCTGGAGTTACCGTAGACTTCAATAAGGATGCCTCCCAAACCGG
ATGTGTAGTTTTCTCTGGAGCGACTGTCCTTTCTGCAGATATTTCTCAGGCTAATTTGCAAACTAAAACACCTGCAA
CGCTTACTCTCAGTCACGGTCTTCTGTGTATCGAAGATCGTGCTCAGCTCACAGTGAACAATTTTACACAAACAGGA
GGGATTGTAGCCTTAGGAAATGGAGCAGTTTTAAGCAGCTACCAACACAGCACTACAGACGCCACTCAAACTCCCCC
TACAACCACCACTACAGATGCTTCCGTAACTCTTAATCACATTGGATTAAATCTCCCCTCTATTCTTAAGGATGGAG
CAGAGATGCCTCTATTATGGGTAGAACCTATAAGCACAACTCAAGGTAACACTACAACATATACGTCAGATACCGCG
GCTTCCTTCTCATTAAATGGAGCCACACTCTCTCTCATTGATGAAGATGGAAATTCTCCCTATGAAAACACGGACCT
CTCTCGTGCATTGTACGCTCAACCTATGCTAGCAATTTCTGAGGCCAGTGATAACCAATTGCAATCCGAAAGCATGG
ACTTTTCTAAAGTTAATGTTCCTCACTATGGATGGCAAGGACTTTGGACCTGGGGGTGGGCAAAAACTGAAAATCCA
ACAACAACTCCTCCAGCAACAATTACTGATCCGAAAAAAGCTAATCAGTTTCATAGAACTTTATTATTAACGTGGCT
CCCTGCTGGTTATATCCCCAGCCCTAAACATAAAAGCCCTTTAATAGCTAATACCTTGTGGGGGAATATACTTTTTG
CAACGGAAAACTTAAAAAATAGCTCAGGGCAAGAACTTCTTGATCGTCCTTTCTGGGGAATTACAGGAGGGGGCTTG
GGGATGATGGTCTATCAAGAACCTAGAAAAGACCATCCTGGATTCCACATGCATACCTCCGGATATTCAGCAGGAAT
GATTACAGGAAACACACATACCTTCTCATTACGATTCAGCCAGTCCTATACAAAACTCAATGAACGTTATGCCAAGA
ACTATGTGTCTTCTAAAAATTACTCTTGCCAAGGGGAAATGCTTTTGTCCTTACAAGAAGGACTCATGCTGACTAAA
CTAATTGGTCTCTATAGTTATGGGAATCACAACAGCCACCATTTCTATACCCAAGGAGAAGACCTATCGTCTCAAGG
GGAGTTCCATAGTCAGACTTTTGGAGGGGCTGTCTTTTTTGATCTACCTCTGAAACCTTTTGGAAGAACACACATAC
TTACAGCTCCTTTCTTAGGTGCCATTGGTATGTATTCTAAGCTGTCTAGCTTTACAGAAGTAGGAGCCTATCCAAGA
ACCTTTATTACAGAAACGCCTTTAATCAATGTCCTGATTCCTATCGGAGTAAAAGGTAGCTTCATGAATGCCACCCA

TAGACCTCAGGCCTGGACTGTAGAGCTTGCTTACCAACCTGTTCTTTACAGACAAGAACCTAGTATCTCTACCCAAT
TACTCGCTGGTAAAGGTATGTGGTTTGGGCATGGAAGTCCTGCATCTCGCCACGCTCTAGCTTATAAAATTTCACAG
AAAACACAGCTTTTGCGATTTGCAACACTTCAACTCCAGTATCACGGATACTATTCGTCTTCCACTTTCTGTAATTA
TCTGAATGGAGAGGTATCTTTACGTTTC

SEQ ID NO: 134 - TC0261 fragment protein sequence

TREVPPSILLKPILNPYHMTGLFFPKVNLLGDTHNLTDYHLDNLKCILACLQRTPYEGAAFTVTDYLGFSDTQKDGI
FCFKNLTPESGGVIGSPTQNTPTIKIHNTIGPVLFENNTCHRLWTQTDPENEGNKAREGGAIHAGDVYISNNQNLVG
FIKNFAYVQGGAISANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIYVSTSCSFENNNKDLLFIQNSGCAGGAIFSP
TCSLIGNQGDIVFYSNHGFKNVDNATNESGDGGAIKVTTRLDITNNGSQIFFSDNISRNFGGAIHAPCLHLVGNGPT
YFTNNIANHTGGAIYITGTETSKISADHHAIIFDNNISANATNADGSSSNTNPPHRNAITMDNSAGGIELGAGKSQN
LIFYDPIQVTNAGVTVDFNKDASQTGCVVFSGATVLSADISQANLQTKTPATLTLSHGLLCIEDRAQLTVNNFTQTG
GIVALGNGAVLSSYQHSTTDATQTPPTTTTTDASVTLNHIGLNLPSILKDGAEMPLLWVEPISTTQGNTTTYTSDTA
ASFSLNGATLSLIDEDGNSPYENTDLSRALYAQPMLAISEASDNQLQSESMDFSKVNVPHYGWQGLWTWGWAKTENP
TTTPPATITDPKKANQFHRTLLLTWLPAGYIPSPKHKSPLIANTLWGNILFATENLKNSSGQELLDRPFWGITGGGL
GMMVYQEPRKDHPGFHMHTSGYSAGMITGNTHTFSLRFSQSYTKLNERYAKNYVSSKNYSCQGEMLLSLQEGLMLTK
LIGLYSYGNHNSHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHILTAPFLGAIGMYSKLSSFTEVGAYPR
TFITETPLINVLIPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPSISTQLLAGKGMWFGHGSPASRHALAYKISQ
KTQLLRFATLQLQYHGYYSSSTFCNYLNGEVSLRF

SEQ ID NO: 135 - CT600 nucleotide sequence

ATGAGAAAGACTATTTTTAAAGCGTTTAATTTATTATTCTCCCTTCTTTTTCTTTCTTCATGCTCTTATCCTTGCAG
AGATTGGGAATGCCATGGTTGCGACTCCGCAAGACCTCGTAAATCCTCTTTTGGATTCGTACCTTTCTACTCCGATG
AAGAAATTCAACAAGCTTTTGTTGAAGATTTTGATTCCAAAGAAGAGCAGCTGTACAAAACGAGCGCACAGAGTACC
TCTTTCCGAAATATCACTTTCGCTACAGATAGTTATTCTATTAAAGGAGAGGATAACCTCACGATTCTTGCAAGCTT
AGTTCGTCATTTGCATAAATCTCCTAAAGCTACGCTATATATAGAGGGCCATACAGATGAACGTGGAGCTGCAGCTT
ATAACCTAGCTTTAGGAGCTCGTCGTGCGAATGCTGTAAAACAATACCTCATCAAACAGGGAATCGCTGCAGACCGC
TTATTCACTATTTCTTACGGAAAAGAACATCCTGTTCATCCAGGCCATAATGAATTAGCTTGGCAACAAAATCGTCG
TACTGAATTTAAGATCCATGCTCGCTAA

SEQ ID NO: 136 - CT600 protein sequence

MRKTIFKAFNLLFSLLFLSSCSYPCRDWECHGCDSARPRKSSFGFVPFYSDEEIQQAFVEDFDSKEEQLYKTSAQST
SFRNITFATDSYSIKGEDNLTILASLVRHLHKSPKATLYIEGHTDERGAAAYNLALGARRANAVKQYLIKQGIAADR
LFTISYGKEHPVHPGHNELAWQQNRRTEFKIHAR

SEQ ID NO: 137 - CT600 fragment nucleotide sequence

TGCTCTTATCCTTGCAGAGATTGGGAATGCCATGGTTGCGACTCCGCAAGACCTCGTAAATCCTCTTTTGGATTCGT
ACCTTTCTACTCCGATGAAGAAATTCAACAAGCTTTTGTTGAAGATTTTGATTCCAAAGAAGAGCAGCTGTACAAAA
CGAGCGCACAGAGTACCTCTTTCCGAAATATCACTTTCGCTACAGATAGTTATTCTATTAAAGGAGAGGATAACCTC
ACGATTCTTGCAAGCTTAGTTCGTCATTTGCATAAATCTCCTAAAGCTACGCTATATATAGAGGGCCATACAGATGA
ACGTGGAGCTGCAGCTTATAACCTAGCTTTAGGAGCTCGTCGTGCGAATGCTGTAAAACAATACCTCATCAAACAGG
GAATCGCTGCAGACCGCTTATTCACTATTTCTTACGGAAAAGAACATCCTGTTCATCCAGGCCATAATGAATTAGCT
TGGCAACAAAATCGTCGTACTGAATTTAAGATCCATGCTCGC

SEQ ID NO: 138 - CT600 fragment protein sequence

CSYPCRDWECHGCDSARPRKSSFGFVPFYSDEEIQQAFVEDFDSKEEQLYKTSAQSTSFRNITFATDSYSIKGEDNL
TILASLVRHLHKSPKATLYIEGHTDERGAAAYNLALGARRANAVKQYLIKQGIAADRLFTISYGKEHPVHPGHNELA
WQQNRRTEFKIHAR

SEQ ID NO: 139 - CT823 nucleotide sequence

ATGATGAAAAGATTATTATGTGTGTTGCTATCGACATCAGTTTTCTCTTCGCCAATGCTAGGCTATAGTGCGTCAAA
GAAAGATTCTAAGGCTGATATTTGTCTTGCAGTATCCTCAGGAGATCAAGAGGTTTCACAAGAAGATCTGCTCAAAG
AAGTATCCCGAGGATTTTCTCGGGTCGCTGCTAAGGCAACGCCTGGAGTTGTATATATAGAAAATTTTCCTAAAACA
GGGAACCAGGCTATTGCTTCTCCAGGAAACAAAAGAGGCTTTCAAGAGAACCCTTTTGATTATTTTAATGACGAATT
TTTTAATCGATTTTTTGGATTGCCTTCGCATAGAGAGCAGCAGCGTCCGCAGCAGCGTGATGCTGTAAGAGGAACTG
GGTTCATTGTTTCTGAAGATGGTTATGTTGTTACTAACCATCATGTAGTCGAGGATGCAGGAAAAATTCATGTTACT
CTCCACGACGGACAAAAATACACAGCTAAGATCGTGGGGTTAGATCCAAAAACAGATCTTGCTGTGATCAAAATTCA
AGCGGAGAAATTACCATTTTTGACTTTTGGGAATTCTGATCAGCTGCAGATAGGTGACTGGGCTATTGCTATTGGAA
ATCCTTTTGGATTGCAAGCAACGGTCACTGTCGGGGTCATTAGTGCTAAAGGAAGAAATCAGCTACATATTGTAGAT
TTCGAAGACTTTATTCAAACAGATGCTGCCATTAATCCTGGGAATTCAGGCGGTCCATTGTTAAACATCAATGGTCA
AGTTATCGGGGTTAATACTGCCATTGTCAGTGGTAGCGGGGGATATATTGGAATAGGGTTTGCTATTCCTAGCTTGA
TGGCTAAACGAGTCATTGATCAATTGATTAGTGATGGGCAGGTAACAAGAGGCTTTTTGGGAGTTACCTTGCAACCG
ATAGATTCTGAATTGGCTACTTGTTACAAATTGGAAAAAGTGTACGGAGCTTTGGTGACGGATGTTGTTAAAGGTTC
TCCAGCAGAAAAAGCAGGGCTGCGCCAAGAAGATGTCATTGTGGCTTACAATGGAAAAGAAGTAGAGTCTTTGAGTG
CGTTGCGTAATGCCATTTCCCTAATGATGCCAGGGACTCGTGTTGTTTTAAAAATCGTTCGTGAAGGGAAAACAATC
GAGATACCTGTGACGGTTACACAGATCCCAACAGAGGATGGCGTTTCAGCGTTGCAGAAGATGGGAGTCCGTGTTCA

GAACATTACTCCAGAAATTTGTAAGAAACTCGGATTGGCAGCAGATACCCGAGGGATTCTGGTAGTTGCTGTGGAGG
CAGGCTCGCCTGCAGCTTCTGCAGGCGTCGCTCCTGGACAGCTTATCTTAGCGGTGAATAGGCAGCGAGTCGCTTCC
GTTGAAGAGTTAAATCAGGTTTTGAAAAACTCGAAAGGAGAGAATGTTCTCCTTATGGTTTCTCAAGGAGATGTGGT
GCGATTCATCGTCTTGAAATCAGACGAGTAG

SEQ ID NO: 140 - CT823 protein sequence

MMKRLLCVLLSTSVFSSPMLGYSASKKDSKADICLAVSSGDQEVSQEDLLKEVSRGFSRVAAKATPGVVYIENFPKT
GNQAIASPGNKRGFQENPFDYFNDEFFNRFFGLPSHREQQRPQQRDAVRGTGFIVSEDGYVVTNHHVVEDAGKIHVT
LHDGQKYTAKIVGLDPKTDLAVIKIQAEKLPFLTFGNSDQLQIGDWAIAIGNPFGLQATVTVGVISAKGRNQLHIVD
FEDFIQTDAAINPGNSGGPLLNINGQVIGVNTAIVSGSGGYIGIGFAIPSLMAKRVIDQLISDGQVTRGFLGVTLQP
IDSELATCYKLEKVYGALVTDVVKGSPAEKAGLRQEDVIVAYNGKEVESLSALRNAISLMMPGTRVVLKIVREGKTI
EIPVTVTQIPTEDGVSALQKMGVRVQNITPEICKKLGLAADTRGILVVAVEAGSPAASAGVAPGQLILAVNRQRVAS
VEELNQVLKNSKGENVLLMVSQGDVVRFIVLKSDE

SEQ ID NO:141 - CT823 fragment nucleotide sequence

TCGCCAATGCTAGGCTATAGTGCGTCAAAGAAAGATTCTAAGGCTGATATTTGTCTTGCAGTATCCTCAGGAGATCA
AGAGGTTTCACAAGAAGATCTGCTCAAAGAAGTATCCCGAGGATTTTCTCGGGTCGCTGCTAAGGCAACGCCTGGAG
TTGTATATATAGAAAATTTTCCTAAAACAGGGAACCAGGCTATTGCTTCTCCAGGAAACAAAAGAGGCTTTCAAGAG
AACCCTTTTGATTATTTTAATGACGAATTTTTTAATCGATTTTTTGGATTGCCTTCGCATAGAGAGCAGCAGCGTCC
GCAGCAGCGTGATGCTGTAAGAGGAACTGGGTTCATTGTTTCTGAAGATGGTTATGTTGTTACTAACCATCATGTAG
TCGAGGATGCAGGAAAAATTCATGTTACTCTCCACGACGGACAAAAATACACAGCTAAGATCGTGGGGTTAGATCCA
AAAACAGATCTTGCTGTGATCAAAATTCAAGCGGAGAAATTACCATTTTTGACTTTTGGGAATTCTGATCAGCTGCA
GATAGGTGACTGGGCTATTGCTATTGGAAATCCTTTTGGATTGCAAGCAACGGTCACTGTCGGGGTCATTAGTGCTA
AAGGAAGAAATCAGCTACATATTGTAGATTTCGAAGACTTTATTCAAACAGATGCTGCCATTAATCCTGGGAATTCA
GGCGGTCCATTGTTAAACATCAATGGTCAAGTTATCGGGGTTAATACTGCCATTGTCAGTGGTAGCGGGGGATATAT
TGGAATAGGGTTTGCTATTCCTAGCTTGATGGCTAAACGAGTCATTGATCAATTGATTAGTGATGGGCAGGTAACAA
GAGGCTTTTTGGGAGTTACCTTGCAACCGATAGATTCTGAATTGGCTACTTGTTACAAATTGGAAAAAGTGTACGGA
GCTTTGGTGACGGATGTTGTTAAAGGTTCTCCAGCAGAAAAAGCAGGGCTGCGCCAAGAAGATGTCATTGTGGCTTA
CAATGGAAAAGAAGTAGAGTCTTTGAGTGCGTTGCGTAATGCCATTTCCCTAATGATGCCAGGGACTCGTGTTGTTT
TAAAAATCGTTCGTGAAGGGAAAACAATCGAGATACCTGTGACGGTTACACAGATCCCAACAGAGGATGGCGTTTCA
GCGTTGCAGAAGATGGGAGTCCGTGTTCAGAACATTACTCCAGAAATTTGTAAGAAACTCGGATTGGCAGCAGATAC
CCGAGGGATTCTGGTAGTTGCTGTGGAGGCAGGCTCGCCTGCAGCTTCTGCAGGCGTCGCTCCTGGACAGCTTATCT
TAGCGGTGAATAGGCAGCGAGTCGCTTCCGTTGAAGAGTTAAATCAGGTTTTGAAAAACTCGAAAGGAGAGAATGTT
CTCCTTATGGTTTCTCAAGGAGATGTGGTGCGATTCATCGTCTTGAAATCAGACGAG

SEQ ID NO:142 - CT823 fragment protein sequence

SPMLGYSASKKDSKADICLAVSSGDQEVSQEDLLKEVSRGFSRVAAKATPGVVYIENFPKTGNQAIASPGNKRGFQE
NPFDYFNDEFFNRFFGLPSHREQQRPQQRDAVRGTGFIVSEDGYVVTNHHVVEDAGKIHVTLHDGQKYTAKIVGLDP
KTDLAVIKIQAEKLPFLTFGNSDQLQIGDWAIAIGNPFGLQATVTVGVISAKGRNQLHIVDFEDFIQTDAAINPGNS
GGPLLNINGQVIGVNTAIVSGSGGYIGIGFAIPSLMAKRVIDQLISDGQVTRGFLGVTLQPIDSELATCYKLEKVYG
ALVTDVVKGSPAEKAGLRQEDVIVAYNGKEVESLSALRNAISLMMPGTRVVLKIVREGKTIEIPVTVTQIPTEDGVS
ALQKMGVRVQNITPEICKKLGLAADTRGILVVAVEAGSPAASAGVAPGQLILAVNRQRVASVEELNQVLKNSKGENV
LLMVSQGDVVRFIVLKSDE

SEQ ID NO:143 – TC0106 nucleotide sequence

ATGCTAACTAACTTTACCTTTCGCAACTGTCTTTTGTTTTTCGTCACATTGTCCAGTGTCCCTGTTTTCTCGGCACC
CCAACCTCGCGTAACGCTTCCTAGTGGAGCCAATAAAATCGGATCAGAAGCTTGGATAGAGCAAAAAGTCCGTCAAT
ATCCAGAACTTTTGTGGTTAGTTGAACCTTCTCCTGCAGGAACTTCTTTAAACGCTCCTTCGGGGATGATCTTTTCT
CCCCTATTGTTCCAAAAGAAAGTCCCTGCTTTTGATATCGCAGTACGCAGTCTGATTCACCTACACCTGCTTATCCA
GGGCTCCCGCCAAGCTTATGCTCAGCTTGTCCAGCTGCAGGCTAATGAATCCCCTATGACATTTAAACAGTTCCTTA
CCCTACATAAGCAGCTCTCCTTATTCCTAAATTCTCCTAAAGAGTTTTATGATTCCGTCAAAATTTTAGAAACTGCT
ATCATCCTACGCCACTTAGGATGTTCAACAAAAGCTGTTGCCACATTTAAGCCTTATTTTTCAGAAACGCAAAAAGA
GGTCTTCTATACAAAAGCTTTGCATGTTCTGCATACTTTCCCAGAATTGAGCCCTTCGTTTGCTAGACTCTCTCCAG
AACAAAAAACGCTCTTCTTCTCATTGAGAAAGCTCGCTAATTATGATGAGTTACTTTCCCTGACAAATGCCCCTAGT
TTACAACTACTATCTGCTGTACGCTCGCGACGCGCGCTTTTGGCTCTAGACTTGTATCTCTATGCTTTAGATTTTTG
TGGAGAACAGGGGATATCCTCTCAGTTTCATATGGACTTTTCTCCTTTACAGTCCATGTTGCAACAATATGCTACGG
TTGAAGAAGCCTTCTCCCGCTACTTTACTTACCGAGCTAATCGCCTAGGATTTGCGGGTTCTTCTCGAACTGAAATG
GCCTTAGTTAGAATAGCTACTTTAATGAACCTATCCCCTTCAGAAGCTGCTATTTTAACAACAAGCTTTAAGTCTCT
TTCCTTGGAAGATGCTGAAAGCTTAGTGAATAGCTTTTATACAAATAAGGGAGACTCTTTAGCTCTTTCTTTACGAG
GACTACCAACTCTTATATCTGAACTAACACGCGCTGCGCATGGAAATACGAATGCGGAAGCTCGAGCTCAGCAAATT
TACGCCACAACGTTATCATTGGTAGCAAAAAGCTTGAAAGCTCACAAAGAGATGCAAAACAAACAAATTCTTCCCGA
AGAAGTCGTTTTAGATTTCTCTGAAACTGCTTCTTCCTGTCAAGGATTGGACATCTTCTCTGAGAACGTTGCTGTTC
AAATCCACTTGAATGGATCTGTCAGCATCCATCTATAA

SEQ ID NO:144 - TC0106 protein sequence

MLTNFTFRNCLLFFVTLSSVPVFSAPQPRVTLPSGANKIGSEAWIEQKVRQYPELLWLVEPSPAGTSLNAPSGMIFS
PLLFQKKVPAFDIAVRSLIHLHLLIQGSRQAYAQLVQLQANESPMTFKQFLTLHKQLSLFLNSPKEFYDSVKILETA

IILRHLGCSTKAVATFKPYFSETQKEVFYTKALHVLHTFPELSPSFARLSPEQKTLFFSLRKLANYDELLSLTNAPS
LQLLSAVRSRRALLALDLYLYALDFCGEQGISSQFHMDFSPLQSMLQQYATVEEAFSRYFTYRANRLGFAGSSRTEM
ALVRIATLMNLSPSEAAILTTSFKSLSLEDAESLVNSFYTNKGDSLALSLRGLPTLISELTRAAHGNTNAEARAQQI
YATTLSLVAKSLKAHKEMQNKQILPEEVVLDFSETASSCQGLDIFSENVAVQIHLNGSVSIHL

SEQ ID NO:145 - TC0106 fragment nucleotide sequence

TCAGAAGCTTGGATAGAGCAAAAAGTCCGTCAATATCCAGAACTTTTGTGGTTAGTTGAACCTTCTCCTGCAGGAAC
TTCTTTAAACGCTCCTTCGGGGATGATCTTTTCTCCCCTATTGTTCCAAAAGAAAGTCCCTGCTTTTGATATCGCAG
TACGCAGTCTGATTCACCTACACCTGCTTATCCAGGGCTCCCGCCAAGCTTATGCTCAGCTTGTCCAGCTGCAGGCT
AATGAATCCCCTATGACATTTAAACAGTTCCTTACCCTACATAAGCAGCTCTCCTTATTCCTAAATTCTCCTAAAGA
GTTTTATGATTCCGTCAAAATTTTAGAAACTGCTATCATCCTACGCCACTTAGGATGTTCAACAAAAGCTGTTGCCA
CATTTAAGCCTTATTTTTCAGAAACGCAAAAAGAGGTCTTCTATACAAAAGCTTTGCATGTTCTGCATACTTTCCCA
GAATTGAGCCCTTCGTTTGCTAGACTCTCTCCAGAACAAAAAACGCTCTTCTTCTCATTGAGAAAGCTCGCTAATTA
TGATGAGTTACTTTCCCTGACAAATGCCCCTAGTTTACAACTACTATCTGCTGTACGCTCGCGACGCGCGCTTTTGG
CTCTAGACTTGTATCTCTATGCTTTAGATTTTTGTGGAGAACAGGGGATATCCTCTCAGTTTCATATGGACTTTTCT
CCTTTACAGTCCATGTTGCAACAATATGCTACGGTTGAAGAAGCCTTCTCCCGCTACTTTACTTACCGAGCTAATCG
CCTAGGATTTGCGGGTTCTTCTCGAACTGAAATGGCCTTAGTTAGAATAGCTACTTTAATGAACCTATCCCCTTCAG
AAGCTGCTATTTTAACAACAAGCTTTAAGTCTCTTTCCTTGGAAGATGCTGAAAGCTTAGTGAATAGCTTTTATACA
AATAAGGGAGACTCTTTAGCTCTTTCTTTACGAGGACTACCAACTCTTATATCTGAACTAACACGCGCTGCGCATGG
AAATACGAATGCGGAAGCTCGAGCTCAGCAAATTTACGCCACAACGTTATCATTGGTAGCAAAAAGCTTGAAAGCTC
ACAAAGAGATGCAAAACAAACAAATTCTTCCCGAAGAAGTCGTTTTAGATTTCTCTGAAACTGCTTCTTCCTGTCAA
GGATTGGACATCTTCTCTGAGAACGTTGCTGTTCAAATCCACTTGAATGGATCTGTCAGCATCCATCTA

SEQ ID NO:146 - TC0106 fragment protein sequence

SEAWIEQKVRQYPELLWLVEPSPAGTSLNAPSGMIFSPLLFQKKVPAFDIAVRSLIHLHLLIQGSRQAYAQLVQLQA
NESPMTFKQFLTLHKQLSLFLNSPKEFYDSVKILETAIILRHLGCSTKAVATFKPYFSETQKEVFYTKALHVLHTFP
ELSPSFARLSPEQKTLFFSLRKLANYDELLSLTNAPSLQLLSAVRSRRALLALDLYLYALDFCGEQGISSQFHMDFS
PLQSMLQQYATVEEAFSRYFTYRANRLGFAGSSRTEMALVRIATLMNLSPSEAAILTTSFKSLSLEDAESLVNSFYT
NKGDSLALSLRGLPTLISELTRAAHGNTNAEARAQQIYATTLSLVAKSLKAHKEMQNKQILPEEVVLDFSETASSCQ
GLDIFSENVAVQIHLNGSVSIHL

SEQ ID NO:147 - TC0431 nucleotide sequence

ATGCCCCACTCTCCTTTTTTATATGTTGTTCAACCGCATTCTGTTTTTAATCCTAGATTGGGAGAGCGGCACCCTAT
TACTTTAGATTTCATCAAAGAAAAGAATCGATTAGCTGATTTTATTGAAAACCTACCTTTAGAAATTTTTGGAGCCC
CTTCTTTCTTGGAAAATGCTTCTTTAGAAGCCTCTTATGTCTTGTCTAGGGAATCCACAAAAGATGGCACTCTTTTT
ACCGTTCTAGAACCCAAACTATCTGCCTGCGTAGCTACTTGCCTTGTGGATTCTTCTATTCCTATGGAGCCCGATAA
CGAGCTCTTAGAAGAAATTAAACACACTTTGTTGAAAAGCTCTTGTGATGGCGTACAATATCGTGTAACCCGAGAGA
CTCTCCAAAACAAAGATGAAGCCCCCAGAGTCTCTTTAGTTGCTGATGATATCGAACTTATCCGCAATGTAGATTTT
TTAGGACGTTCCGTTGATATTGTAAAATTGGATCCCTTGAATATTCCTAATACCGTAAGCGAGGAGAATGCTCTCGA
TTACTCTTTCACAAGGGAAACCGCCAAACTTAGCCCTGACGGACGAGTTGGCATCCCTCAAGGGACAAAAATTTTGC
CAGCTCCCTCTCTTGAAGTTGAAATTAGCACCTCTATTTTTGAGGAAACCTCTTCTTTTGAACAAAACTTTTCTTCC
TCTATTACTTTTTGTGTACCACCTCTTACCTCTTTTTCTCCTTTGCAAGAACCTCCTCTAGTGGGAGCTGGACAGCA
GGAAATTCTTGTGACTAAAAAGCACTTATTCCCTAGCTATACCCCTAAACTTATTGATATTGTCAAACGACACAAAA
GAGACGCAAAGATTCTAGTAAACAAGATCCAGTTCGAGAAACTATGGAGAAGTCATGCCAAAAGTCAAATCTTAAAA
GAAGGCTCTGTTCGCTTGGATTTACAAGGATTTACAGGGGAGCTGTTTAACTACCAACTTCAAGTAGGATCTCATAC
AATTGCAGCCGTGTTAATTGATCCGGAAATTGCTAACGTCAAATCCCTCCCCGAACAAACTTACGCTGTAAGAAAAA
TTAAATCAGGGTTCCAATGTAGTTTGGATGACCAACACATTTATCAAGTCGCAGTAAAAAAACATCTTTCTCTGTCT
TCACAACCTCCGAAGATATCTCCGTTATCTCAATCCGAAAGCTCCGATTTAAGTCTCTTTGAAGCAGCAGCGTTTTC
AGCAAGCCTAACTTACGAGTTCGTAAAGAAAAATACATATCATGCTAAGAATACTGTAACTTGCTCCACGGTATCGC
ACTCTCTGTATATTCTCAAAGAAGATGACGGGGCTAATGCTGCAGAAAAACGCTTAGACAACAGTTTCCGAAACTGG
GTCGAAAATAAGTTGAACGCAAATTCTCCAGATTCTTGTACTGCATTTATTCAAAAATTCGGCACACATTACATCAC
ATCGGCAACTTTTGGAGGATCTGGGTTCCAAGTTCTTAAATTATCCTTTGAACAGGTAGAAGGCCTCCGTAGTAAGA
AGATCTCCCTAGAAGCAGCAGCAGCAAATTCCTTATTAAAAAGCTCTGTGTCAAACAGCACGGAATCTGGCTACTCT
ACTTACGATTCCTCTTCTTCTTCTCATACAGTATTCCTAGGGGGCACTGTATTACCCTCTGTTCATGATGGACAGTT
AGATTTTAAAGATTGGTCTGAAAGTGTCTGTTTAGAACCTGTTCCCATTCACATTTCTTTACTCCCCTTAACAGACT
TGCTCACCCCTCTTTATTTTCCTGAAACGGATACAACCGAACTATCTAATAAACGTAATGCTCTCCAACAAGCGGTT
CGAGTTTACCTTAAAGACCATCGTTCAGCTAAACAAAGCGAACGCTCCGTATTCACAGCGGGGATCAATAGTCCTTC
TTCCTGGTTCACATTAGAATCTGCTAATTCACCTCTTGTTGTGAGTTCTCCTTACATGACGTATTGGTCTACTCTCC
CCTATCTCTTCCCCACATTAAAAGAGCGTTCTTCAGCAGCTCCCATCGTTTTTTATTTTTGTGTGGATAATAATGAA
CACGCCTCCCAAAAAATTTTAAACCAAACATATTGCTTCATAGGTTCTTTACCTATTCGACAAAAGATTTTTGGCAG
AGAATTTGCTGAGAATCCTTATTTATCTTTCTATGGAAGGTTTGGAGAAGCTTATTTTGATGGCGGTTATCCAGAAC
GTTGTGGATGGATTGTTGAAAAGTTAAATACTACTAAAGATCAAATTCTCCGCGATGAGGATGAAGTGCAACTAAAG
CATGTTTATAGCGGAGAGTATCTGTCTACAATTCCTATTAAGGATTCCCATTGCACACTCTCGCGTACATGCACCGA
ATCGAATGCTGTTTTTATTATCAAAAAACCTTCGAGCTATTGA

SEQ ID NO:148 - TC0431 protein sequence

MPHSPFLYVVQPHSVFNPRLGERHPITLDFIKEKNRLADFIENLPLEIFGAPSFLENASLEASYVLSRESTKDGTLF
TVLEPKLSACVATCLVDSSIPMEPDNELLEEIKHTLLKSSCDGVQYRVTRETLQNKDEAPRVSLVADDIELIRNVDF
LGRSVDIVKLDPLNIPNTVSEENALDYSFTRETAKLSPDGRVGIPQGTKILPAPSLEVEISTSIFEETSSFEQNFSS
SITFCVPPLTSFSPLQEPPLVGAGQQEILVTKKHLFPSYTPKLIDIVKRHKRDAKILVNKIQFEKLWRSHAKSQILK
EGSVRLDLQGFTGELFNYQLQVGSHTIAAVLIDPEIANVKSLPEQTYAVRKIKSGFQCSLDDQHIYQVAVKKHLSLS
SQPPKISPLSQSESSDLSLFEAAAFSASLTYEFVKKNTYHAKNTVTCSTVSHSLYILKEDDGANAAEKRLDNSFRNW
VENKLNANSPDSCTAFIQKFGTHYITSATFGGSGFQVLKLSFEQVEGLRSKKISLEAAAANSLLKSSVSNSTESGYS
TYDSSSSSHTVFLGGTVLPSVHDGQLDFKDWSESVCLEPVPIHISLLPLTDLLTPLYFPETDTTELSNKRNALQQAV
RVYLKDHRSAKQSERSVFTAGINSPSSWFTLESANSPLVVSSPYMTYWSTLPYLFPTLKERSSAAPIVFYFCVDNNE
HASQKILNQTYCFIGSLPIRQKIFGREFAENPYLSFYGRFGEAYFDGGYPERCGWIVEKLNTTKDQILRDEDEVQLK
HVYSGEYLSTIPIKDSHCTLSRTCTESNAVFIIKKPSSY

SEQ ID NO:149 - TC0431 fragment nucleotide sequence

CCCCACTCTCCTTTTTTATATGTTGTTCAACCGCATTCTGTTTTTAATCCTAGATTGGGAGAGCGGCACCCTATTAC
TTTAGATTTCATCAAAGAAAAGAATCGATTAGCTGATTTTATTGAAAACCTACCTTTAGAAATTTTTGGAGCCCCTT
CTTTCTTGGAAAATGCTTCTTTAGAAGCCTCTTATGTCTTGTCTAGGGAATCCACAAAAGATGGCACTCTTTTTACC
GTTCTAGAACCCAAACTATCTGCCTGCGTAGCTACTTGCCTTGTGGATTCTTCTATTCCTATGGAGCCCGATAACGA
GCTCTTAGAAGAAATTAAACACACTTTGTTGAAAAGCTCTTGTGATGGCGTACAATATCGTGTAACCCGAGAGACTC
TCCAAAACAAAGATGAAGCCCCCAGAGTCTCTTTAGTTGCTGATGATATCGAACTTATCCGCAATGTAGATTTTTTA
GGACGTTCCGTTGATATTGTAAAATTGGATCCCTTGAATATTCCTAATACCGTAAGCGAGGAGAATGCTCTCGATTA
CTCTTTCACAAGGGAAACCGCCAAACTTAGCCCTGACGGACGAGTTGGCATCCCTCAAGGGACAAAAATTTTGCCAG
CTCCCTCTCTTGAAGTTGAAATTAGCACCTCTATTTTTGAGGAAACCTCTTCTTTTGAACAAAACTTTTCTTCCTCT
ATTACTTTTTGTGTACCACCTCTTACCTCTTTTTCTCCTTTGCAAGAACCTCCTCTAGTGGGAGCTGGACAGCAGGA
AATTCTTGTGACTAAAAAGCACTTATTCCCTAGCTATACCCCTAAACTTATTGATATTGTCAAACGACACAAAAGAG
ACGCAAAGATTCTAGTAAACAAGATCCAGTTCGAGAAACTATGGAGAAGTCATGCCAAAAGTCAAATCTTAAAAGAA
GGCTCTGTTCGCTTGGATTTACAAGGATTTACAGGGGAGCTGTTTAACTACCAACTTCAAGTAGGATCTCATACAAT
TGCAGCCGTGTTAATTGATCCGGAAATTGCTAACGTCAAATCCCTCCCCGAACAAACTTACGCTGTAAGAAAAATTA
AATCAGGGTTCCAATGTAGTTTGGATGACCAACACATTTATCAAGTCGCAGTAAAAAAACATCTTTCTCTGTCTTCA
CAACCTCCGAAGATATCTCCGTTATCTCAATCCGAAAGCTCCGATTTAAGTCTCTTTGAAGCAGCAGCGTTTTCAGC
AAGCCTAACTTACGAGTTCGTAAAGAAAAATACATATCATGCTAAGAATACTGTAACTTGCTCCACGGTATCGCACT
CTCTGTATATTCTCAAAGAAGATGACGGGGCTAATGCTGCAGAAAAACGCTTAGACAACAGTTTCCGAAACTGGGTC
GAAAATAAGTTGAACGCAAATTCTCCAGATTCTTGTACTGCATTTATTCAAAAATTCGGCACACATTACATCACATC
GGCAACTTTTGGAGGATCTGGGTTCCAAGTTCTTAAATTATCCTTTGAACAGGTAGAAGGCCTCCGTAGTAAGAAGA
TCTCCCTAGAAGCAGCAGCAGCAAATTCCTTATTAAAAAGCTCTGTGTCAAACAGCACGGAATCTGGCTACTCTACT
TACGATTCCTCTTCTTCTTCTCATACAGTATTCCTAGGGGGCACTGTATTACCCTCTGTTCATGATGGACAGTTAGA
TTTTAAAGATTGGTCTGAAAGTGTCTGTTTAGAACCTGTTCCCATTCACATTTCTTTACTCCCCTTAACAGACTTGC
TCACCCCTCTTTATTTTCCTGAAACGGATACAACCGAACTATCTAATAAACGTAATGCTCTCCAACAAGCGGTTCGA
GTTTACCTTAAAGACCATCGTTCAGCTAAACAAAGCGAACGCTCCGTATTCACAGCGGGGATCAATAGTCCTTCTTC
CTGGTTCACATTAGAATCTGCTAATTCACCTCTTGTTGTGAGTTCTCCTTACATGACGTATTGGTCTACTCTCCCCT
ATCTCTTCCCCACATTAAAAGAGCGTTCTTCAGCAGCTCCCATCGTTTTTTATTTTTGTGTGGATAATAATGAACAC
GCCTCCCAAAAAATTTTAAACCAAACATATTGCTTCATAGGTTCTTTACCTATTCGACAAAAGATTTTTGGCAGAGA
ATTTGCTGAGAATCCTTATTTATCTTTCTATGGAAGGTTTGGAGAAGCTTATTTTGATGGCGGTTATCCAGAACGTT
GTGGATGGATTGTTGAAAAGTTAAATACTACTAAAGATCAAATTCTCCGCGATGAGGATGAAGTGCAACTAAAGCAT
GTTTATAGCGGAGAGTATCTGTCTACAATTCCTATTAAGGATTCCCATTGCACACTCTCGCGTACATGCACCGAATC
GAATGCTGTTTTTATTATCAAAAAACCTTCGAGCTAT

SEQ ID NO:150 - TC0431 fragment protein sequence

PHSPFLYVVQPHSVFNPRLGERHPITLDFIKEKNRLADFIENLPLEIFGAPSFLENASLEASYVLSRESTKDGTLFT
VLEPKLSACVATCLVDSSIPMEPDNELLEEIKHTLLKSSCDGVQYRVTRETLQNKDEAPRVSLVADDIELIRNVDFL
GRSVDIVKLDPLNIPNTVSEENALDYSFTRETAKLSPDGRVGIPQGTKILPAPSLEVEISTSIFEETSSFEQNFSSS
ITFCVPPLTSFSPLQEPPLVGAGQQEILVTKKHLFPSYTPKLIDIVKRHKRDAKILVNKIQFEKLWRSHAKSQILKE
GSVRLDLQGFTGELFNYQLQVGSHTIAAVLIDPEIANVKSLPEQTYAVRKIKSGFQCSLDDQHIYQVAVKKHLSLSS
QPPKISPLSQSESSDLSLFEAAAFSASLTYEFVKKNTYHAKNTVTCSTVSHSLYILKEDDGANAAEKRLDNSFRNWV
ENKLNANSPDSCTAFIQKFGTHYITSATFGGSGFQVLKLSFEQVEGLRSKKISLEAAAANSLLKSSVSNSTESGYST
YDSSSSSHTVFLGGTVLPSVHDGQLDFKDWSESVCLEPVPIHISLLPLTDLLTPLYFPETDTTELSNKRNALQQAVR
VYLKDHRSAKQSERSVFTAGINSPSSWFTLESANSPLVVSSPYMTYWSTLPYLFPTLKERSSAAPIVFYFCVDNNEH
ASQKILNQTYCFIGSLPIRQKIFGREFAENPYLSFYGRFGEAYFDGGYPERCGWIVEKLNTTKDQILRDEDEVQLKH
VYSGEYLSTIPIKDSHCTLSRTCTESNAVFIIKKPSSY

SEQ ID NO:151 - TC0210 nucleotide sequence

ATGATGAAAAGATTATTATGTGTGTTGCTATCGACATCAGTTTTCTCTTCGCCCATGTTGGGCTATAGTGCGCCAAA
GAAAGATTCCAGTACTGGCATTTGTCTTGCAGCATCTCAAAGTGATCGGGAACTTTCCCAAGAAGATTTGCTAAAAG
AAGTGTCTAGAGGATTTTCCAAAGTCGCTGCTCAGGCAACTCCAGGAGTTGTGTATATAGAAAATTTTCCTAAAACT
GGGAGTCAAGCTATTGCTTCTCCTGGGAATAAAAGGGGTTTTCAAGAGAATCCCTTTGATTATTTCAATGATGAGTT
TTTCAATCGATTTTTTGGTTTACCCTCGCATAGAGAGCAGCCTCGTCCCCAACAGCGTGATGCTGTAAGAGGAACAG

GTTTTATTGTGTCAGAAGATGGGTACGTTGTGACCAACCATCACGTAGTGGAAGATGCGGGGAAAATTCATGTTACT
TTACACGATGGACAAAAATACACCGCAAAAATCATAGGATTAGATCCTAAAACGGATCTCGCTGTGATTAAGATCCA
AGCAAAAAATCTCCCTTTTTTAACTTTTGGAAACTCTGATCAGCTTCAGATAGGGGATTGGTCAATAGCCATTGGAA
ATCCTTTCGGATTACAAGCCACAGTAACCGTTGGCGTGATTAGTGCTAAGGGAAGAAACCAATTACATATTGTTGAT
TTTGAAGATTTTATTCAGACGGATGCAGCAATTAATCCCGGGAATTCAGGTGGTCCATTATTGAACATTGATGGACA
GGTTATTGGAGTGAATACAGCAATCGTTAGCGGTAGCGGGGGATACATTGGAATAGGATTTGCCATTCCTAGCTTAA
TGGCTAAACGAGTTATTGACCAACTCATTAGCGATGGACAGGTGACGAGAGGATTTTTAGGAGTAACCTTACAGCCT
ATTGATTCGGAGCTTGCCGCTTGTTACAAATTAGAAAAGGTGTACGGAGCCTTGATTACGGATGTTGTTAAGGGATC
TCCTGCAGAAAAAGCAGGTTTGCGCCAGGAAGATGTCATTGTTGCTTACAATGGGAAAGAAGTGGAGTCTTTGAGTG
CTTTACGTAATGCGATTTCTTTGATGATGCCAGGGACTCGTGTTGTCTTAAAAGTTGTGCGTGAAGGGAAATTCATT
GAAATACCTGTCACTGTTACACAAATTCCTGCGGAGGATGGGGTATCTGCTCTTCAAAAAATGGGAGTTCGGGTACA
GAATCTTACTCCAGAGATATGCAAGAAACTAGGATTAGCGTCTGATACTCGAGGGATTTTTGTAGTGTCCGTAGAAG
CTGGTTCTCCTGCAGCTTCTGCAGGAGTGGTTCCAGGACAACTTATTCTGGCTGTAAACAGACAGAGAGTTTCTTCT
GTTGAAGAATTGAATCAGGTCTTGAAGAATGCAAAAGGAGAGAAATGTTCTCCTTATGGTTTCTCAAGGAGAAGTCAT
TCGATTCGTTGTTTTAAAGTCTGATGAATAG

SEQ ID NO:152 - TC0210 protein sequence

MMKRLLCVLLSTSVFSSPMLGYSAPKKDSSTGICLAASQSDRELSQEDLLKEVSRGFSKVAAQATPGVVYIENFPKT
GSQAIASPGNKRGFQENPFDYFNDEFFNRFFGLPSHREQPRPQQRDAVRGTGFIVSEDGYVVTNHHVVEDAGKIHVT
LHDGQKYTAKIIGLDPKTDLAVIKIQAKNLPFLTFGNSDQLQIGDWSIAIGNPFGLQATVTVGVISAKGRNQLHIVD
FEDFIQTDAAINPGNSGGPLLNIDGQVIGVNTAIVSGSGGYIGIGFAIPSLMAKRVIDQLISDGQVTRGFLGVTLQP
IDSELAACYKLEKVYGALITDVVKGSPAEKAGLRQEDVIVAYNGKEVESLSALRNAISLMMPGTRVVLKVVREGKFI
EIPVTVTQIPAEDGVSALQKMGVRVQNLTPEICKKLGLASDTRGIFVVSVEAGSPAASAGVVPGQLILAVNRQRVSS
VEELNQVLKNAKGENVLLMVSQGEVIRFVVLKSDE

SEQ ID NO:153 - TC0210 fragment nucleotide sequence

TCGCCCATGTTGGGCTATAGTGCGCCAAAGAAAGATTCCAGTACTGGCATTTGTCTTGCAGCATCTCAAAGTGATCG
GGAACTTTCCCAAGAAGATTTGCTAAAAGAAGTGTCTAGAGGATTTTCCAAAGTCGCTGCTCAGGCAACTCCAGGAG
TTGTGTATATAGAAAATTTTCCTAAAACTGGGAGTCAAGCTATTGCTTCTCCTGGGAATAAAAGGGGTTTTCAAGAG
AATCCCTTTGATTATTTCAATGATGAGTTTTTCAATCGATTTTTTGGTTTACCCTCGCATAGAGAGCAGCCTCGTCC
CCAACAGCGTGATGCTGTAAGAGGAACAGGTTTTATTGTGTCAGAAGATGGGTACGTTGTGACCAACCATCACGTAG
TGGAAGATGCGGGGAAAATTCATGTTACTTTACACGATGGACAAAAATACACCGCAAAAATCATAGGATTAGATCCT
AAAACGGATCTCGCTGTGATTAAGATCCAAGCAAAAAATCTCCCTTTTTTAACTTTTGGAAACTCTGATCAGCTTCA
GATAGGGGATTGGTCAATAGCCATTGGAAATCCTTTCGGATTACAAGCCACAGTAACCGTTGGCGTGATTAGTGCTA
AGGGAAGAAACCAATTACATATTGTTGATTTTGAAGATTTTATTCAGACGGATGCAGCAATTAATCCCGGGAATTCA
GGTGGTCCATTATTGAACATTGATGGACAGGTTATTGGAGTGAATACAGCAATCGTTAGCGGTAGCGGGGGATACAT
TGGAATAGGATTTGCCATTCCTAGCTTAATGGCTAAACGAGTTATTGACCAACTCATTAGCGATGGACAGGTGACGA
GAGGATTTTTAGGAGTAACCTTACAGCCTATTGATTCGGAGCTTGCCGCTTGTTACAAATTAGAAAAGGTGTACGGA
GCCTTGATTACGGATGTTGTTAAGGGATCTCCTGCAGAAAAAGCAGGTTTGCGCCAGGAAGATGTCATTGTTGCTTA
CAATGGGAAAGAAGTGGAGTCTTTGAGTGCTTTACGTAATGCGATTTCTTTGATGATGCCAGGGACTCGTGTTGTCT
TAAAAGTTGTGCGTGAAGGGAAATTCATTGAAATACCTGTCACTGTTACACAAATTCCTGCGGAGGATGGGGTATCT
GCTCTTCAAAAAATGGGAGTTCGGGTACAGAATCTTACTCCAGAGATATGCAAGAAACTAGGATTAGCGTCTGATAC
TCGAGGGATTTTTGTAGTGTCCGTAGAAGCTGGTTCTCCTGCAGCTTCTGCAGGAGTGGTTCCAGGACAACTTATTC
TGGCTGTAAACAGACAGAGAGTTTCTTCTGTTGAAGAATTGAATCAGGTCTTGAAGAATGCAAAAGGAGAGAATGTT
CTCCTTATGGTTTCTCAAGGAGAAGTCATTCGATTCGTTGTTTTAAAGTCTGATGAA

SEQ ID NO:154 - TC0210 fragment protein sequence

SPMLGYSAPKKDSSTGICLAASQSDRELSQEDLLKEVSRGFSKVAAQATPGVVYIENFPKTGSQAIASPGNKRGFQE
NPFDYFNDEFFNRFFGLPSHREQPRPQQRDAVRGTGFIVSEDGYVVTNHHVVEDAGKIHVTLHDGQKYTAKIIGLDP
KTDLAVIKIQAKNLPFLTFGNSDQLQIGDWSIAIGNPFGLQATVTVGVISAKGRNQLHIVDFEDFIQTDAAINPGNS
GGPLLNIDGQVIGVNTAIVSGSGGYIGIGFAIPSLMAKRVIDQLISDGQVTRGFLGVTLQPIDSELAACYKLEKVYG
ALITDVVKGSPAEKAGLRQEDVIVAYNGKEVESLSALRNAISLMMPGTRVVLKVVREGKFIEIPVTVTQIPAEDGVS
ALQKMGVRVQNLTPEICKKLGLASDTRGIFVVSVEAGSPAASAGVVPGQLILAVNRQRVSSVEELNQVLKNAKGENV
LLMVSQGEVIRFVVLKSDE

SEQ ID NO:155 – CT163 nucleotide sequence

ATGTTTGTGTCGTTCGATAAATCCCGTTGCAGAGCGGATGTCCCCGATTTTTTTGAAAGGACAGGAAACTTTCTTCT
CCATTGTGTGGCAAGAGGGATCAATGTTTTATATCGTGTGAAACAAATCTCTAACTATCCTTCATGCTATTTCTCAC
ATAAAGAGATTTCGTGTTGTCGTCGTATTGCAAACATTGTGATCTGTATTCTCACAGGGCCTCTGATGTTATTGGCC
ACTGTGTTAGGATTATTAGCGTATAGGTTTTCTTCTACTTACCAGACTTCTTTACAAGAACGCTTTCGTTATAAATA
TGAACAAAAGCAAGCTTTAGATGAATACCGTGATAGGGAAGAAAAAGTCATTACGCTTCAGAAGTTTTGTAGAGGAT
TTCTAGTTAGAAATCATTTGCTCAACCAAGAAACTTTAACAACGTGTAAGCAATGGGGGCAAAAACTATTAGAAGGA
GAAAAATTCCCAAGGGTCCCAGAAGGACGGTCTCTTGTATATATTTCAAAACAGTTTCCTTCTTTAGTAGCAAAACA
CGTTGGGGCTCAAGATGCCAGGTCTCGTTGGCATCATATTTTTTCTATGCGCAAAGCGCTTGCTTATTTAGATATTA
AGCGCATACGAGCACCACGCGCTAGAGTTTATCAAAACTTTATATTCGAAGAAAAACTTCCTGTTTCACGAATTTCT
GTAGATTCAATGTGTCTCTATAAAGAAAATCCACAAGCTTTCGATGAGGCGATCAAAGAACTCTTATTTCTATTTAA
AGAAGTGCATTTCAGGGATTTTGTTGTAGAAACAGAGTCTCCAACAGACGATTTCCCCTTAGCCGTGAAAGTACACA

ACTATTGGGTATGCCCACGATACGATAATTTACCTTTATTTATTCAAGAAGGAAAAGATGGCTCTCCAGAAGGGCGT
ATAGGACTGGTCGATCTAGAAACTTTTTCTTGGTCTCCACATCCATACCCCGTAGAAGAACTAGCTGTGATGTTTCC
TATGCATAAAGAGCTTCTTATGACAGAGGCGAAAAAACTACAAATCCCTTTCTCTACAAAGGAGGTCGAGCGCTCTG
TAGAGAAAGGGCTTGCTTTTTTTGAACATATGCTAGGGCATCAAGATTTTTGTTCCCAAAAAAGCGTAACGCCATTG
CGTAATTGTGCCCCTTATATTCATCTAGAAGTATGGAGATTCTCACTGAAAATTTTTGATATTTTAAAAGCTGCTAT
TCAACTAAATGGAGCACTCAATGTTCTGTTATCTCCAGATATTCGAGAGCGGTTGAGTGCTATTTCGGATAAGCAAT
GGTTGGCTATTAGCTCCCAGGTTACGTCATCGTTACTCGAGCAAGTTTCTACAAACATCTATCAGTCTCATACTGAA
GAGGCTAAACGAGTAAATTCTTCAGGGACTTTTATCATGTGTCGATCTCCTATCTTCCGGAAAAGCATCTTCATTAA
AAATCTCCCACAATTCTTAAACAAGAAATTGCAGTTGCTTCCAGAGGAGAAAGCAATCAGCGAGGCGCTTGCTTCTC
TATGTTTACGTGCAGTAATGGAAGAGCTAGTAGCAACAGGAAATATTTATTCTTATGATTCTATGGATGATTTTTTT
GAAGGGCAGTATTGTCGCATTCGTTATTAG

SEQ ID NO:156 – CT163 protein sequence

MFVSFDKSRCRADVPDFFERTGNFLLHCVARGINVLYRVKQISNYPSCYFSHKEISCCRRIANIVICILTGPLMLLA
TVLGLLAYRFSSTYQTSLQERFRYKYEQKQALDEYRDREEKVITLQKFCRGFLVRNHLLNQETLTTCKQWGQKLLEG
EKFPRVPEGRSLVYISKQFPSLVAKHVGAQDARSRWHHIFSMRKALAYLDIKRIRAPRARVYQNFIFEEKLPVSRIS
VDSMCLYKENPQAFDEAIKELLFLFKEVHFRDFVVETESPTDDFPLAVKVHNYWVCPRYDNLPLFIQEGKDGSPEGR
IGLVDLETFSWSPHPYPVEELAVMFPMHKELLMTEAKKLQIPFSTKEVERSVEKGLAFFEHMLGHQDFCSQKSVTPL
RNCAPYIHLEVWRFSLKIFDILKAAIQLNGALNVLLSPDIRERLSAISDKQWLAISSQVTSSLLEQVSTNIYQSHTE
EAKRVNSSGTFIMCRSPIFRKSIFIKNLPQFLNKKLQLLPEEKAISEALASLCLRAVMEELVATGNIYSYDSMDDFF
EGQYCRIRY

SEQ ID NO:157 – CT163 fragment nucleotide sequence

TTTGTGTCGTTCGATAAATCCCGTTGCAGAGCGGATGTCCCCGATTTTTTTGAAAGGACAGGAAACTTTCTTCTCCA
TTGTGTGGCAAGAGGGATCAATGTTTTATATCGTGTGAAACAAATCTCTAACTATCCTTCATGCTATTTCTCACATA
AAGAGATTTCGTGTTGTCGTCGTATTGCAAACATTGTGATCTGTATTCTCACAGGGCCTCTGATGTTATTGGCCACT
GTGTTAGGATTATTAGCGTATAGGTTTTCTTCTACTTACCAGACTTCTTTACAAGAACGCTTTCGTTATAAATATGA
ACAAAAGCAAGCTTTAGATGAATACCGTGATAGGGAAGAAAAAGTCATTACGCTTCAGAAGTTTTGTAGAGGATTTC
TAGTTAGAAATCATTTGCTCAACCAAGAAACTTTAACAACGTGTAAGCAATGGGGGCAAAAACTATTAGAAGGAGAA
AAATTCCCAAGGGTCCCAGAAGGACGGTCTCTTGTATATATTTCAAAACAGTTTCCTTCTTTAGTAGCAAAACACGT
TGGGGCTCAAGATGCCAGGTCTCGTTGGCATCATATTTTTTCTATGCGCAAAGCGCTTGCTTATTTAGATATTAAGC
GCATACGAGCACCACGCGCTAGAGTTTATCAAAACTTTATATTCGAAGAAAAACTTCCTGTTTCACGAATTTCTGTA
GATTCAATGTGTCTCTATAAAGAAAATCCACAAGCTTTCGATGAGGCGATCAAAGAACTCTTATTTCTATTTAAAGA
AGTGCATTTCAGGGATTTTGTTGTAGAAACAGAGTCTCCAACAGACGATTTCCCCTTAGCCGTGAAAGTACACAACT
ATTGGGTATGCCCACGATACGATAATTTACCTTTATTTATTCAAGAAGGAAAAGATGGCTCTCCAGAAGGGCGTATA
GGACTGGTCGATCTAGAAACTTTTTCTTGGTCTCCACATCCATACCCCGTAGAAGAACTAGCTGTGATGTTTCCTAT
GCATAAAGAGCTTCTTATGACAGAGGCGAAAAAACTACAAATCCCTTTCTCTACAAAGGAGGTCGAGCGCTCTGTAG
AGAAAGGGCTTGCTTTTTTTGAACATATGCTAGGGCATCAAGATTTTTGTTCCCAAAAAAGCGTAACGCCATTGCGT
AATTGTGCCCCTTATATTCATCTAGAAGTATGGAGATTCTCACTGAAAATTTTTGATATTTTAAAAGCTGCTATTCA
ACTAAATGGAGCACTCAATGTTCTGTTATCTCCAGATATTCGAGAGCGGTTGAGTGCTATTTCGGATAAGCAATGGT
TGGCTATTAGCTCCCAGGTTACGTCATCGTTACTCGAGCAAGTTTCTACAAACATCTATCAGTCTCATACTGAAGAG
GCTAAACGAGTAAATTCTTCAGGGACTTTTATCATGTGTCGATCTCCTATCTTCCGGAAAAGCATCTTCATTAAAAA
TCTCCCACAATTCTTAAACAAGAAATTGCAGTTGCTTCCAGAGGAGAAAGCAATCAGCGAGGCGCTTGCTTCTCTAT
GTTTACGTGCAGTAATGGAAGAGCTAGTAGCAACAGGAAATATTTATTCTTATGATTCTATGGATGATTTTTTTGAA
GGGCAGTATTGTCGCATTCGTTAT

SEQ ID NO:158 – CT163 fragment protein sequence

FVSFDKSRCRADVPDFFERTGNFLLHCVARGINVLYRVKQISNYPSCYFSHKEISCCRRIANIVICILTGPLMLLAT
VLGLLAYRFSSTYQTSLQERFRYKYEQKQALDEYRDREEKVITLQKFCRGFLVRNHLLNQETLTTCKQWGQKLLEGE
KFPRVPEGRSLVYISKQFPSLVAKHVGAQDARSRWHHIFSMRKALAYLDIKRIRAPRARVYQNFIFEEKLPVSRISV
DSMCLYKENPQAFDEAIKELLFLFKEVHFRDFVVETESPTDDFPLAVKVHNYWVCPRYDNLPLFIQEGKDGSPEGRI
GLVDLETFSWSPHPYPVEELAVMFPMHKELLMTEAKKLQIPFSTKEVERSVEKGLAFFEHMLGHQDFCSQKSVTPLR
NCAPYIHLEVWRFSLKIFDILKAAIQLNGALNVLLSPDIRERLSAISDKQWLAISSQVTSSLLEQVSTNIYQSHTEE
AKRVNSSGTFIMCRSPIFRKSIFIKNLPQFLNKKLQLLPEEKAISEALASLCLRAVMEELVATGNIYSYDSMDDFFE
GQYCRIRY

SEQ ID NO:159 - CT214 nucleotide sequence

ATGCGAACAGACTCTCTTTTCAATCCTCCCGACTCTACTAGAGGAGTTTTTCAGTTTTTAGAGACTCAGTGTGATCG
AGCCGTGGCTCGGTCCAGACAAAGCCAATTTATAGGGTTAGTCTCTGCTGTAGCAGCTGCAGCATTATTATTGTTGC
TTGTGGTCGCTCTATCTGTTCCAGGATTCCCAGTTGCAGCTTCAATTGTTGTAGGGGTTCTCTTTGCTTTATCGATC
GTAGCATTAACAGCTTCGTTTTTGGTATATATAGCTAATGCTAAGCTTGTTGCAATAAGAATTAAATTCTTGAGTAG
TGGTCTGCAAGATCACTTTTCGGAGTCATCTATTTTAGGGACTCTCCGTAAAGGACGTGGTGCTAGTATTCCGCTTA
TTTCCGGACAAGCAGATGATCCTCTCCCTAATCGGATTGGGATCAAAAAAAGCACTGAAATGCGTGTTCTTCAAAAA
GGAATTGGGACAGATTATAAAAAATATAAGCAGCATCTTGATAGAGTGAATAATGATTTCACTTTTGTCTGTGAGGG
GATTAGCGCTTTAATTCCTACAGAAAAAGATGCTCCATTCCCTATAGAACCTTCTCATTTAGCAGGTGTTTTTTTAG
TATCATTTTCACCAGACAAGAATCCGATTCTAAAGATTACGCGTCATGCTGAGAAGATGTTACAGCCTCCTCAAGGC
GGATTCCCTAACGGGCTGGTTTGGTTGTGTGGAGCTCTTTCTGATCCTAAGAAATTTGCAGCTCCCTTTCTATCTTT

GATTGAGAAGACTCACCAAGGGATTTTGGTGAGTAAAGACTTGAAAGACAATAAGGAAAGAAAGCTAGCTTTAGAGG
CTTCCCTTCTTTCATTGAATATTTTCTTTTCCGGTTGGTGTTTGGGGAATCCGGAGTACAATCAGTATATCACAACT
GCTGTAGCTGAGAAATATAGGGATGTCTCTGTAAGAAATTGTATTTATGATTTCCTGGATACAGGGAATGTGATTTC
AGCTCTTGCTTTAGCAAGTAGTTATTCACAAGATTCCGCTTGGGCTGCAGGGTTGCAGAAAGTTTTACGTGAAGAAG
ATAAAAAGACTAAGAAAAAGTCACGTGAAGAAGTCTCTTGTTTGTATCGTGATATAGATCCAGGCTGTTGTTTAAGA
GCCCTTCCTAAGCGATTTGAATCCAAGTCTTCAGGTAGTCAAGGTAGTCCTAAAGAGCAGTTAAGCTCTTTGTTGAA
AGCTTTAGACCAGAAAATTCCTTCAGGGATTTTAGGATTGATTGCAAAAGCTTCTTCTGCAGATCTCAAGGCTGATT
TTGCAGGTATGCTTGAAGTTATTAAGCAATTACAAGCTTTATTCGATTCTTACCCACCTTTATGCGAAGACAATATT
CTCTTGTGGTTAAGCGCTTCTTTAGAACAAGTAGGCTTGCAGAAGAAATTGAGAACCTTTTTACCTTCATCAGAAAA
AAAACTCTTAGAAAGAGTTCTCTCTACATTTTTATTAGGTTTGTATACTCGAGGAGTCTTTTCTGTAGGGCAAGTGA
ATCAGCTAGCTACTATTTGTAATACTCAGGACTCTACAGAATTCTGCCAGAGAGTAAGTGACCTTTCGTTAATTAAA
CGAGCTCTACCTGCATTATTTGGTTAA

SEQ ID NO:160 - CT214 protein sequence

MRTDSLFNPPDSTRGVFQFLETQCDRAVARSRQSQFIGLVSAVAAAALLLLLVVALSVPGFPVAASIVVGVLFALSI
VALTASFLVYIANAKLVAIRIKFLSSGLQDHFSESSILGTLRKGRGASIPLISGQADDPLPNRIGIKKSTEMRVLQK
GIGTDYKKYKQHLDRVNNDFTFVCEGISALIPTEKDAPFPIEPSHLAGVFLVSFSPDKNPILKITRHAEKMLQPPQG
GFPNGLVWLCGALSDPKKFAAPFLSLIEKTHQGILVSKDLKDNKERKLALEASLLSLNIFFSGWCLGNPEYNQYITT
AVAEKYRDVSVRNCIYDFLDTGNVISALALASSYSQDSAWAAGLQKVLREEDKKTKKKSREEVSCLYRDIDPGCCLR
ALPKRFESKSSGSQGSPKEQLSSLLKALDQKIPSGILGLIAKASSADLKADFAGMLEVIKQLQALFDSYPPLCEDNI
LLWLSASLEQVGLQKKLRTFLPSSEKKLLERVLSTFLLGLYTRGVFSVGQVNQLATICNTQDSTEFCQRVSDLSLIK
RALPALFG

SEQ ID NO:161 - CT214 fragment nucleotide sequence

CGAACAGACTCTCTTTTCAATCCTCCCGACTCTACTAGAGGAGTTTTTCAGTTTTTAGAGACTCAGTGTGATCGAGC
CGTGGCTCGGTCCAGACAAAGCCAATTTATAGGGTTAGTCTCTGCTGTAGCAGCTGCAGCATTATTATTGTTGCTTG
TGGTCGCTCTATCTGTTCCAGGATTCCCAGTTGCAGCTTCAATTGTTGTAGGGGTTCTCTTTGCTTTATCGATCGTA
GCATTAACAGCTTCGTTTTTGGTATATATAGCTAATGCTAAGCTTGTTGCAATAAGAATTAAATTCTTGAGTAGTGG
TCTGCAAGATCACTTTTCGGAGTCATCTATTTTAGGGACTCTCCGTAAAGGACGTGGTGCTAGTATTCCGCTTATTT
CCGGACAAGCAGATGATCCTCTCCCTAATCGGATTGGGATCAAAAAAAGCACTGAAATGCGTGTTCTTCAAAAAGGA
ATTGGGACAGATTATAAAAAATATAAGCAGCATCTTGATAGAGTGAATAATGATTTCACTTTTGTCTGTGAGGGGAT
TAGCGCTTTAATTCCTACAGAAAAAGATGCTCCATTCCCTATAGAACCTTCTCATTTAGCAGGTGTTTTTTTAGTAT
CATTTTCACCAGACAAGAATCCGATTCTAAAGATTACGCGTCATGCTGAGAAGATGTTACAGCCTCCTCAAGGCGGA
TTCCCTAACGGGCTGGTTTGGTTGTGTGGAGCTCTTTCTGATCCTAAGAAATTTGCAGCTCCCTTTCTATCTTTGAT
TGAGAAGACTCACCAAGGGATTTTGGTGAGTAAAGACTTGAAAGACAATAAGGAAAGAAAGCTAGCTTTAGAGGCTT
CCCTTCTTTCATTGAATATTTTCTTTTCCGGTTGGTGTTTGGGGAATCCGGAGTACAATCAGTATATCACAACTGCT
GTAGCTGAGAAATATAGGGATGTCTCTGTAAGAAATTGTATTTATGATTTCCTGGATACAGGGAATGTGATTTCAGC
TCTTGCTTTAGCAAGTAGTTATTCACAAGATTCCGCTTGGGCTGCAGGGTTGCAGAAAGTTTTACGTGAAGAAGATA
AAAAGACTAAGAAAAAGTCACGTGAAGAAGTCTCTTGTTTGTATCGTGATATAGATCCAGGCTGTTGTTTAAGAGCC
CTTCCTAAGCGATTTGAATCCAAGTCTTCAGGTAGTCAAGGTAGTCCTAAAGAGCAGTTAAGCTCTTTGTTGAAAGC
TTTAGACCAGAAAATTCCTTCAGGGATTTTAGGATTGATTGCAAAAGCTTCTTCTGCAGATCTCAAGGCTGATTTTG
CAGGTATGCTTGAAGTTATTAAGCAATTACAAGCTTTATTCGATTCTTACCCACCTTTATGCGAAGACAATATTCTC
TTGTGGTTAAGCGCTTCTTTAGAACAAGTAGGCTTGCAGAAGAAATTGAGAACCTTTTTACCTTCATCAGAAAAAAA
ACTCTTAGAAAGAGTTCTCTCTACATTTTTATTAGGTTTGTATACTCGAGGAGTCTTTTCTGTAGGGCAAGTGAATC
AGCTAGCTACTATTTGTAATACTCAGGACTCTACAGAATTCTGCCAGAGAGTAAGTGACCTTTCGTTAATTAAACGA
GCTCTACCTGCATTATTTGGT

SEQ ID NO:162 - CT214 fragment protein sequence

RTDSLFNPPDSTRGVFQFLETQCDRAVARSRQSQFIGLVSAVAAAALLLLLVVALSVPGFPVAASIVVGVLFALSIV
ALTASFLVYIANAKLVAIRIKFLSSGLQDHFSESSILGTLRKGRGASIPLISGQADDPLPNRIGIKKSTEMRVLQKG
IGTDYKKYKQHLDRVNNDFTFVCEGISALIPTEKDAPFPIEPSHLAGVFLVSFSPDKNPILKITRHAEKMLQPPQGG
FPNGLVWLCGALSDPKKFAAPFLSLIEKTHQGILVSKDLKDNKERKLALEASLLSLNIFFSGWCLGNPEYNQYITTA
VAEKYRDVSVRNCIYDFLDTGNVISALALASSYSQDSAWAAGLQKVLREEDKKTKKKSREEVSCLYRDIDPGCCLRA
LPKRFESKSSGSQGSPKEQLSSLLKALDQKIPSGILGLIAKASSADLKADFAGMLEVIKQLQALFDSYPPLCEDNIL
LWLSASLEQVGLQKKLRTFLPSSEKKLLERVLSTFLLGLYTRGVFSVGQVNQLATICNTQDSTEFCQRVSDLSLIKR
ALPALFG

SEQ ID NO:163 – CT721 nucleotide sequence

ATGGACGGGACAAAAATTCACGAAACACGCTCCTTCTCTTGGTTAAACAACCAACAAGCCATCCCTCCTTCCGAAAT
GGTGAAGGAGGCTTTTCAACGTTACGCAGACGTATTTTCGTACAGCGCAAATACCTCCATTCTGACTTTACAAGCAG
AAGCTGAAGCTTCTGCCCGCAAACTCACAGGGTGTCAGGAGAAGGCTTTTACCTTTCATTTTATTCTTCATTACCCG
AATGTCACGGCCATTATCGTGGCCGCTCTTCTGGAAAACCAAAATGCCTTCCAGGGGCGTAATCACCTTCTTGTTCC
TTCTTGCGAGCAACAATTTATCATTAATGCTCTCTGCCGTCGGCAAAACTTAGGGACAACCTATGATTGGGTAACCA
GCAAAAACGGCCGCGTAAAAGAATCCGATCTAGCAGAAGCTCTTTCCCCGCGGACCTTGCTGTTTTCCATATCTGCT
GCGAATGGTATGACAGGATTTCTGGAAGCGATCCCTGAGCTTGCTGCGTTATGTAAAGAACGCGGGGTAATTTTCCA
CATAGACCTGAGTGATATCTTAGGAAGATGCGCGCTACCCGCAGAACTCTATCAAGCAGATATCCTTACTTTTTCTT
CACAGTCTCTTGGTGGGATTGGTCCCTCAGGAGCGATGTTTATTTCTCCCGCTTTAACAAAATATTTTTCCTTATGG

CTTCCTAGTAATCCACAAGTCCCTACCTGCCTGAGTTCTCTTGCAGCTTTTTCTCTTGCCTGTCAGGAACGTACAAC
CGCTTTCTCCTCTCTTGTGCTTTCTGCTATTTCTTCTCGAGCAGCTCTTAAACAGGCTCTTTCCGCTATTCCTCAAG
TCGAATTCCTTTTGGAAGACAGTGCCCCTCGTCTCCCTAATGTCGCTGTCTTTGCTATTCCTGGTATCCCTGCAGAG
TCCTTAGGATTTTTCCTTTCCCAGAAAAATATTTTTGTAGGGTTAGGCTATGAACGCTTCCAGCCTCTATCGCAGAT
TTTACAAAGTTCGGGCATCTCTCCCTTCTTATGCCACAGCGCTTTACACGTATCTTTTACTGAACGTACTCCTACTA
CACACTTCTCTGCATTAGCAACCGCCTTACAAGAAGGGATCTCTCACCTACAACCACTGGTTACTCAATCCTTATGA

SEQ ID NO:164 – CT721 protein sequence

MDGTKIHETRSFSWLNNQQAIPPSEMVKEAFQRYADVFSYSANTSILTLQAEAEASARKLTGCQEKAFTFHFILHYP
NVTAIIVAALLENQNAFQGRNHLLVPSCEQQFIINALCRRQNLGTTYDWVTSKNGRVKESDLAEALSPRTLLFSISA
ANGMTGFLEAIPELAALCKERGVIFHIDLSDILGRCALPAELYQADILTFSSQSLGGIGPSGAMFISPALTKYFSLW
LPSNPQVPTCLSSLAAFSLACQERTTAFSSLVLSAISSRAALKQALSAIPQVEFLLEDSAPRLPNVAVFAIPGIPAE
SLGFFLSQKNIFVGLGYERFQPLSQILQSSGISPFLCHSALHVSFTERTPTTHFSALATALQEGISHLQPLVTQSL

SEQ ID NO:165 – CT721 fragment nucleotide sequence

GACGGGACAAAAATTCACGAAACACGCTCCTTCTCTTGGTTAAACAACCAACAAGCCATCCCTCCTTCCGAAATGGT
GAAGGAGGCTTTTCAACGTTACGCAGACGTATTTTCGTACAGCGCAAATACCTCCATTCTGACTTTACAAGCAGAAG
CTGAAGCTTCTGCCCGCAAACTCACAGGGTGTCAGGAGAAGGCTTTTACCTTTCATTTTATTCTTCATTACCCGAAT
GTCACGGCCATTATCGTGGCCGCTCTTCTGGAAAACCAAAATGCCTTCCAGGGGCGTAATCACCTTCTTGTTCCTTC
TTGCGAGCAACAATTTATCATTAATGCTCTCTGCCGTCGGCAAAACTTAGGGACAACCTATGATTGGGTAACCAGCA
AAAACGGCCGCGTAAAAGAATCCGATCTAGCAGAAGCTCTTTCCCCGCGGACCTTGCTGTTTTCCATATCTGCTGCG
AATGGTATGACAGGATTTCTGGAAGCGATCCCTGAGCTTGCTGCGTTATGTAAAGAACGCGGGGTAATTTTCCACAT
AGACCTGAGTGATATCTTAGGAAGATGCGCGCTACCCGCAGAACTCTATCAAGCAGATATCCTTACTTTTTCTTCAC
AGTCTCTTGGTGGGATTGGTCCCTCAGGAGCGATGTTTATTTCTCCCGCTTTAACAAAATATTTTTCCTTATGGCTT
CCTAGTAATCCACAAGTCCCTACCTGCCTGAGTTCTCTTGCAGCTTTTTCTCTTGCCTGTCAGGAACGTACAACCGC
TTTCTCCTCTCTTGTGCTTTCTGCTATTTCTTCTCGAGCAGCTCTTAAACAGGCTCTTTCCGCTATTCCTCAAGTCG
AATTCCTTTTGGAAGACAGTGCCCCTCGTCTCCCTAATGTCGCTGTCTTTGCTATTCCTGGTATCCCTGCAGAGTCC
TTAGGATTTTTCCTTTCCCAGAAAAATATTTTTGTAGGGTTAGGCTATGAACGCTTCCAGCCTCTATCGCAGATTTT
ACAAAGTTCGGGCATCTCTCCCTTCTTATGCCACAGCGCTTTACACGTATCTTTTACTGAACGTACTCCTACTACAC
ACTTCTCTGCATTAGCAACCGCCTTACAAGAAGGGATCTCTCACCTACAACCACTGGTTACTCAATCCTTA

SEQ ID NO:166 – CT721 fragment protein sequence

DGTKIHETRSFSWLNNQQAIPPSEMVKEAFQRYADVFSYSANTSILTLQAEAEASARKLTGCQEKAFTFHFILHYPN
VTAIIVAALLENQNAFQGRNHLLVPSCEQQFIINALCRRQNLGTTYDWVTSKNGRVKESDLAEALSPRTLLFSISAA
NGMTGFLEAIPELAALCKERGVIFHIDLSDILGRCALPAELYQADILTFSSQSLGGIGPSGAMFISPALTKYFSLWL
PSNPQVPTCLSSLAAFSLACQERTTAFSSLVLSAISSRAALKQALSAIPQVEFLLEDSAPRLPNVAVFAIPGIPAES
LGFFLSQKNIFVGLGYERFQPLSQILQSSGISPFLCHSALHVSFTERTPTTHFSALATALQEGISHLQPLVTQSL

SEQ ID NO:167 – CT127 nucleotide sequence

ATGCCGCACCAAGTCTTATTGTCTCCTGTTTGCGATCTTTTATCGAATGCTGAAGGTATAGAGACGCAAGTACTGTT
TGGAGAAAGGATATGCAACCATAACCATCGACACTATGCCTATTCTCAACTAGTCTTTTCTTCTATATGGAAGCCAT
ACCCTGGCGACTCTCTACAGAATATTCCTCTATTCTCTTCCCAACTGCAGCCTCCTAATGCTGTTGTCTGCTCTCAA
GAAGCTTTTTTAGATCCTTGGCATATCCCCTTACCTTTTGCCGCTCCGCTCCACATAGATAACCAAAATCAAGTGTC
CCTATCTCCTGCTAGCATAGCATTATTAAATTCCAATTCCAGAAGTAACTATGCAAAAGCTTTCTGCTCTACCAAAG
AGATTCGTTTTTTAAATTCTTCATTCTCTCCAAGAGATTTAGTTTCTTTCGCAGAACAATTGATAGATACTCCGTAC
GTTTGGGGTGGCCGGTGCATTCATAAACAGCTTCCTCGTAATGGTGTAGATTGTTCGGGGTATATTCAACTACTTTA
CCAAGTCACAGGAAGAAATATCCCTCGCAATGCTAGAGATCAATACAGAGACTGTTCTCCAGTAAAAGATTTCTCGT
CTCTACCTATAGGAGGACTTATCTTCCTCAAGAAAGCAAGCACGGGACAAATCAACCATGTTATGATGAAAATCTCG
GAGCATGAATTCATTCATGCTGCGGAAAAAATAGGGAAAGTAGAAAAAGTAATCCTAGGAAATAGGGCTTTCTTTAA
AGGGAATCTATTCTGCTCATTAGGTGAACCGCCTATAGAAGCTGTTTTTGGCGTTCCTAAAAATAGAAAAGCCTTCT
TTTGA

SEQ ID NO:168 – CT127 protein sequence

MPHQVLLSPVCDLLSNAEGIETQVLFGERICNHNHRHYAYSQLVFSSIWKPYPGDSLQNIPLFSSQLQPPNAVVCSQ
EAFLDPWHIPLPFAAPLHIDNQNQVSLSPASIALLNSNSRSNYAKAFCSTKEIRFLNSSFSPRDLVSFAEQLIDTPY
VWGGRCIHKQLPRNGVDCSGYIQLLYQVTGRNIPRNARDQYRDCSPVKDFSSLPIGGLIFLKKASTGQINHVMMKIS
EHEFIHAAEKIGKVEKVILGNRAFFKGNLFCSLGEPPIEAVFGVPKNRKAFF

SEQ ID NO:169 – CT127 fragment nucleotide sequence

CCGCACCAAGTCTTATTGTCTCCTGTTTGCGATCTTTTATCGAATGCTGAAGGTATAGAGACGCAAGTACTGTTTGG
AGAAAGGATATGCAACCATAACCATCGACACTATGCCTATTCTCAACTAGTCTTTTCTTCTATATGGAAGCCATACC
CTGGCGACTCTCTACAGAATATTCCTCTATTCTCTTCCCAACTGCAGCCTCCTAATGCTGTTGTCTGCTCTCAAGAA
GCTTTTTTAGATCCTTGGCATATCCCCTTACCTTTTGCCGCTCCGCTCCACATAGATAACCAAAATCAAGTGTCCCT
ATCTCCTGCTAGCATAGCATTATTAAATTCCAATTCCAGAAGTAACTATGCAAAAGCTTTCTGCTCTACCAAAGAGA
TTCGTTTTTTAAATTCTTCATTCTCTCCAAGAGATTTAGTTTCTTTCGCAGAACAATTGATAGATACTCCGTACGTT
TGGGGTGGCCGGTGCATTCATAAACAGCTTCCTCGTAATGGTGTAGATTGTTCGGGGTATATTCAACTACTTTACCA

AGTCACAGGAAGAAATATCCCTCGCAATGCTAGAGATCAATACAGAGACTGTTCTCCAGTAAAAGATTTCTCGTCTC
TACCTATAGGAGGACTTATCTTCCTCAAGAAAGCAAGCACGGGACAAATCAACCATGTTATGATGAAAATCTCGGAG
CATGAATTCATTCATGCTGCGGAAAAAATAGGGAAAGTAGAAAAAGTAATCCTAGGAAATAGGGCTTTCTTTAAAGG
GAATCTATTCTGCTCATTAGGTGAACCGCCTATAGAAGCTGTTTTTGGCGTTCCTAAAAATAGAAAAGCCTTCTTT

SEQ ID NO:170 – CT127 fragment protein sequence

PHQVLLSPVCDLLSNAEGIETQVLFGERICNHNHRHYAYSQLVFSSIWKPYPGDSLQNIPLFSSQLQPPNAVVCSQE
AFLDPWHIPLPFAAPLHIDNQNQVSLSPASIALLNSNSRSNYAKAFCSTKEIRFLNSSFSPRDLVSFAEQLIDTPYV
WGGRCIHKQLPRNGVDCSGYIQLLYQVTGRNIPRNARDQYRDCSPVKDFSSLPIGGLIFLKKASTGQINHVMMKISE
HEFIHAAEKIGKVEKVILGNRAFFKGNLFCSLGEPPIEAVFGVPKNRKAFF

CLAIMS

1. A protein comprising the amino acid sequence of any one of SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO: 18, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO: 48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:136 or SEQ ID NO:140 for use in therapy or diagnosis.
2. A protein having 50% or greater sequence identity to a protein according to claim 1.
3. A protein comprising a fragment of the amino acid sequence of claim 1 or claim 2.
4. A protein according to claim 3, wherein the fragment comprises at least 8 consecutive amino acids of the amino acid sequence of claim 1 or claim 2.
5. An antibody which binds to a protein according to any one of claims 1 to 4 for use in therapy or diagnosis.
6. A nucleic acid molecule which encodes a protein or antibody according to any one of claims 1 to 5 for use in therapy or diagnosis.
7. A nucleic acid molecule according to claim 6, comprising the amino acid sequence of any one of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:135 or SEQ ID NO:139.
8. A nucleic acid molecule comprising of a fragment of the nucleotide sequence of any one of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43 or SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:135 or SEQ ID NO:139.
9. A nucleic acid molecule comprising a nucleotide sequence complementary to a nucleic acid molecule according to any one of claims 6 to 8.
10. A nucleic acid molecule comprising a nucleotide sequences having 50% or greater sequence identity to a nucleic acid molecule according to any one of claims 6 to 9.
11. A nucleic acid molecule which can hybridise to a nucleic acid molecule according to any one of claims 6 to 10 under high stringency conditions.
12. A vector comprising a nucleic acid according to any one of claims 6 to 11 for use in therapy or diagnosis.
13. A host cell comprising a nucleic acid or vector according to any one of claims 6 to 12 for use in therapy or diagnosis.

14. A composition comprising a protein, antibody, nucleic acid molecule, vector or host cell according to any preceding claim for use in therapy or diagnosis.

15. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim for use as a vaccine composition.

16. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim for use as a pharmaceutical.

17. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim, for use in the treatment, prevention or diagnosis of Chlamydia.

18. A protein, antibody, nucleic acid, vector, host cell or composition according claim 17, for use in the treatment, prevention or diagnosis of *Chlamydia trachomatis*.

19. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim, for raising an immune response in a mammal.

20. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim, for eliciting antibodies that are capable of neutralising Chlamydia infection.

21. A protein or nucleic acid according to any one of claims 1 to 19, wherein the immune response is a CD4+ Th1 cell-mediated response.

22. The use of a protein, antibody, nucleic acid, vector or host cell according to any one of claims 1 to 13 in the manufacture of a medicament for the treatment or prevention of infection due to *Chlamydia* bacteria, particularly *Chlamydia trachomatis*.

23. A method of treating, preventing or diagnosing Chlamydia in a patient, comprising administering a therapeutically effective amount of a protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim.

24. A protein, antibody, nucleic acid, vector, host cell or composition according to any preceding claim, for use as a medicament in combination with one or more additional Chlamydia antigens or their encoding nucleic acids.

25. A protein, antibody, nucleic acid, vector, host cell or composition according to claim 24, wherein the combination comprises CT279+CT601, CT372+CT443, CT733+CT153, CT456+CT381, CT279+CT153+CT733+CT601, CT279+CT601+CT372+CT443, CT823+CT733+CT043+CT456, CT387+CT812+CT869, CT387+CT812C+CT869 (or variants thereof).

26. A protein, antibody, nucleic acid, vector, host cell or composition according to claim 24 or claim 25, wherein the protein, antibody, nucleic acid, vector, host cell or composition and the one or more additional Chlamydia antigens or their encoding nucleic acids are a combined preparation for simultaneous, separate or sequential administration.

27. A method for diagnosing Chlamydia infection, comprising:

(a) raising an antibody against a protein according to any one of claims 1 to 4;

(b) contacting the antibody of step (a) with a biological sample suspected of being infected with Chlamydia under conditions suitable for the formation of antibody-antigen complexes; and

(c) detecting said complexes, wherein detection of said complex is indicative of Chlamydia infection.

FIGURE 1

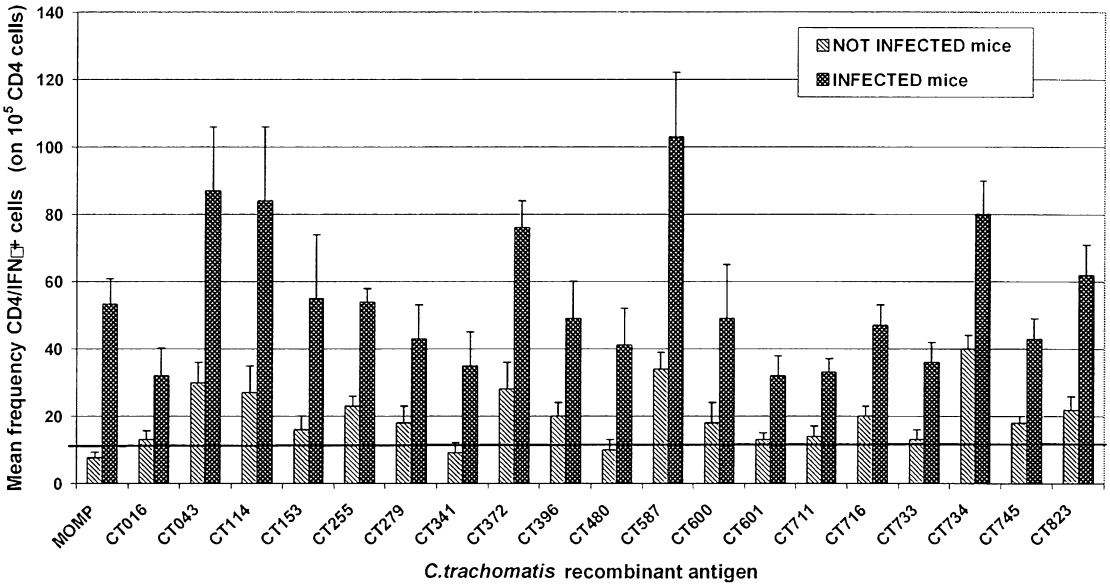


FIGURE 2

FIGURE 2A

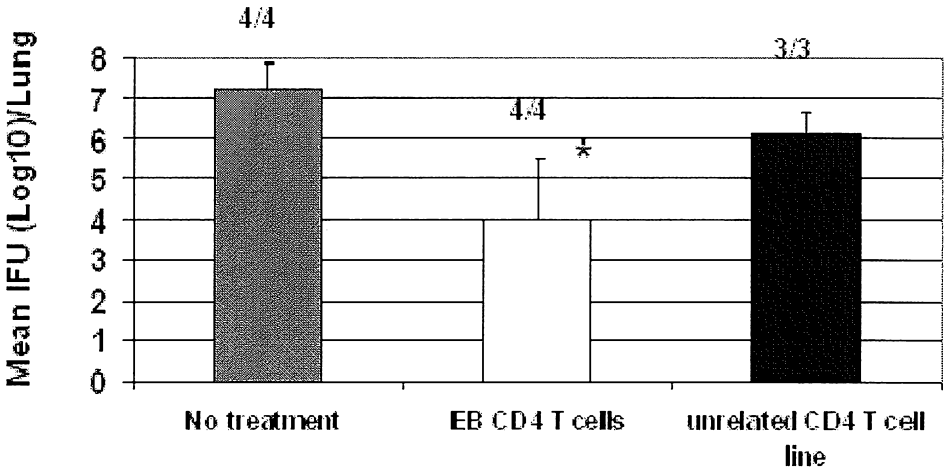


FIGURE 2B

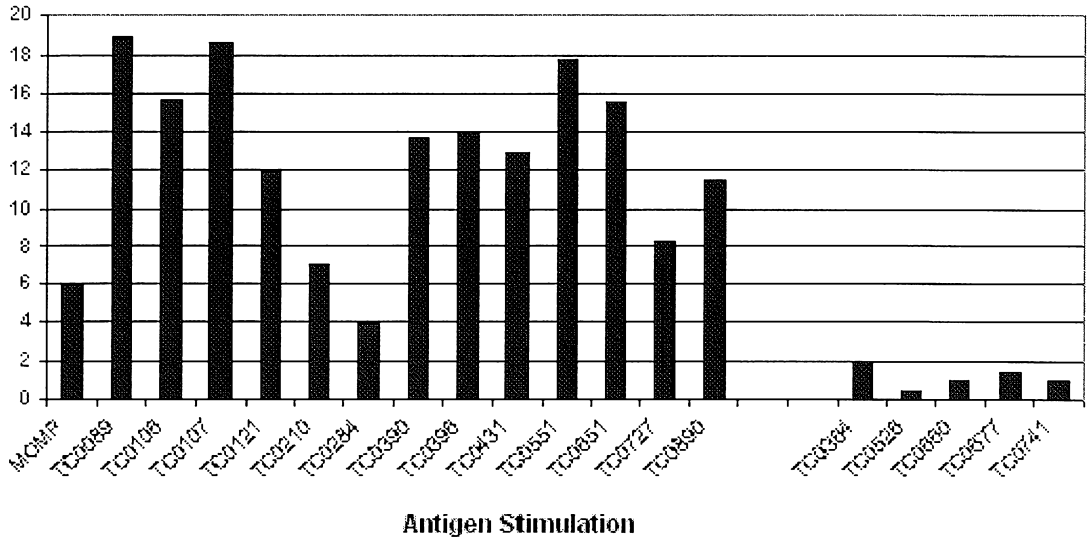


FIGURE 3

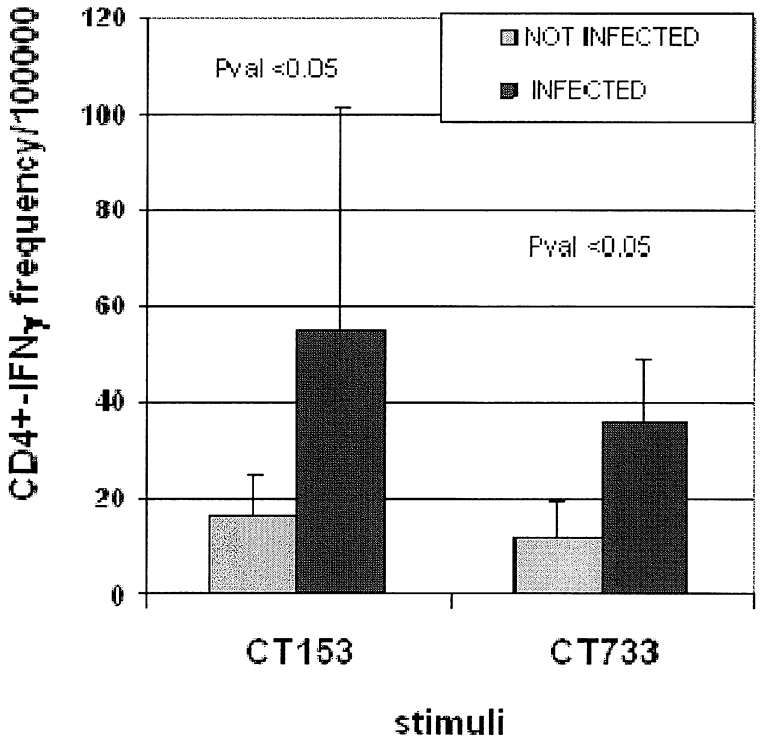


FIGURE 4

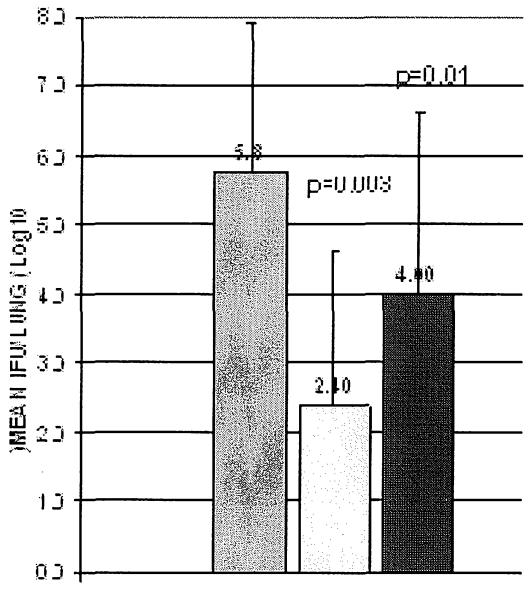


FIGURE 5

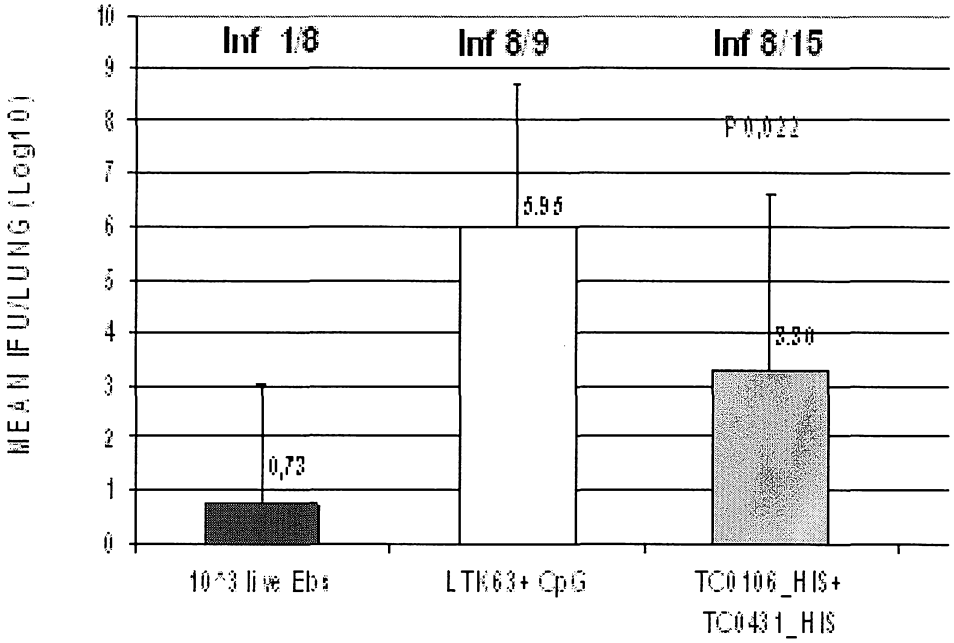


FIGURE 6

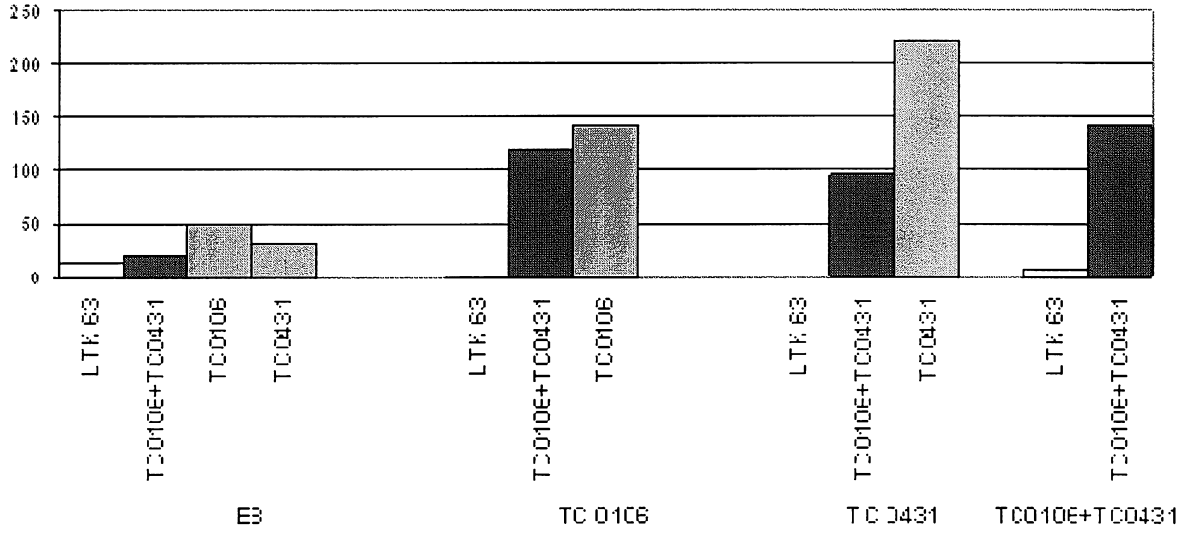


FIGURE 7

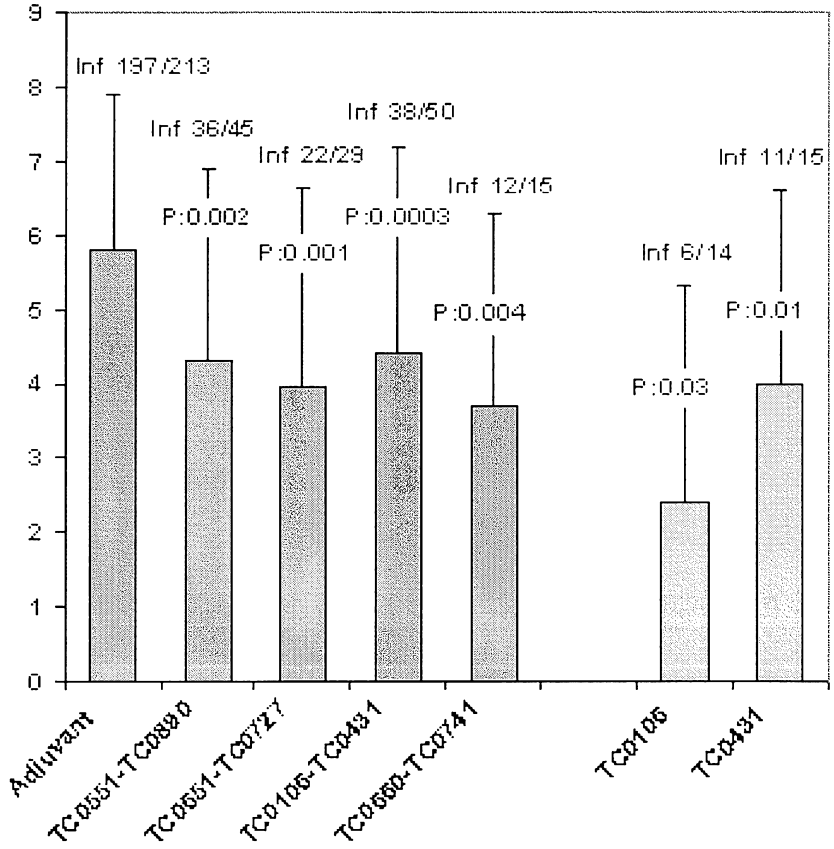


FIGURE 8

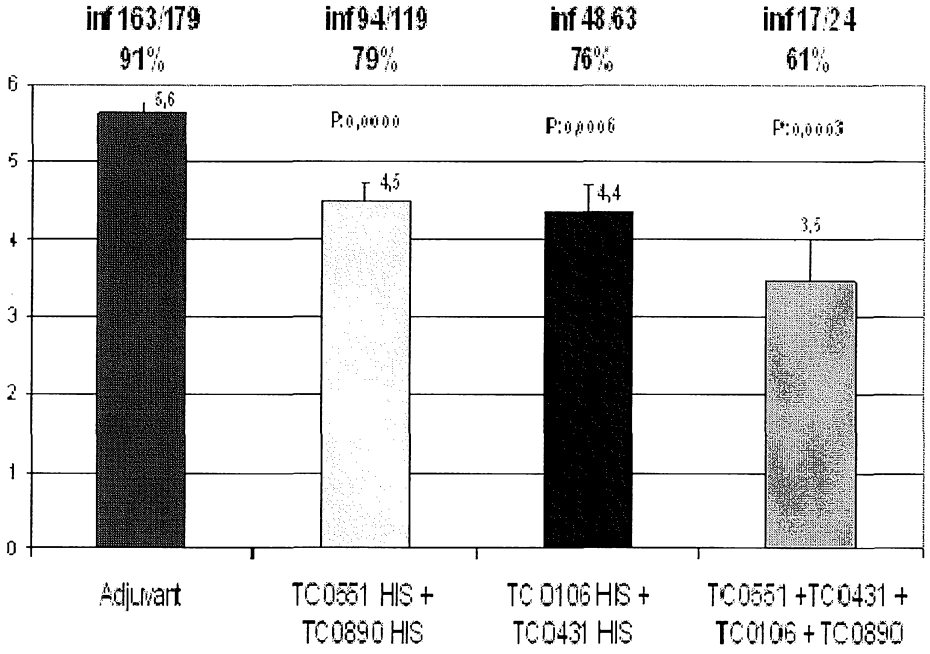


FIGURE 9

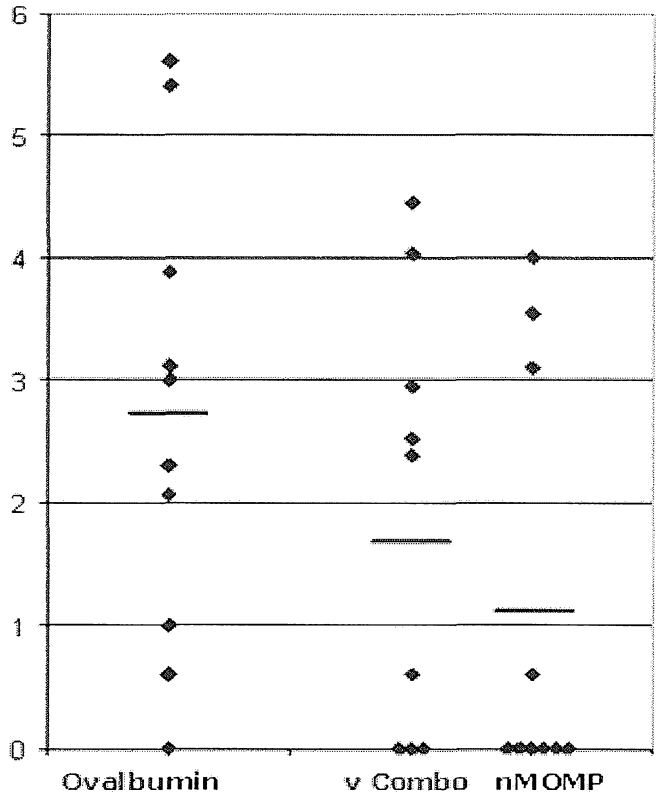


FIGURE 10

FIGURE 10A

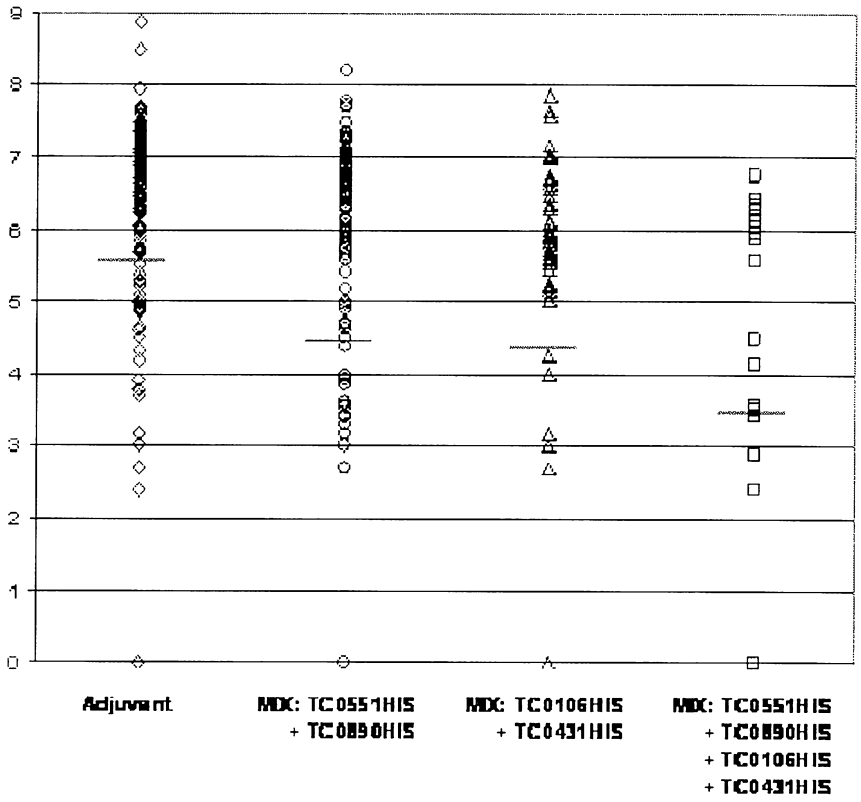


FIGURE 10B

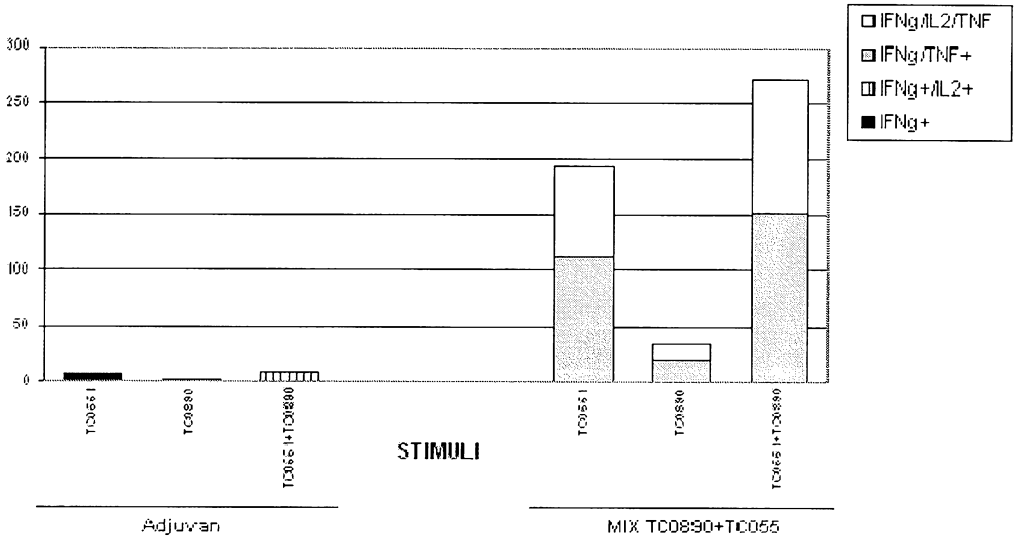


FIGURE 10C

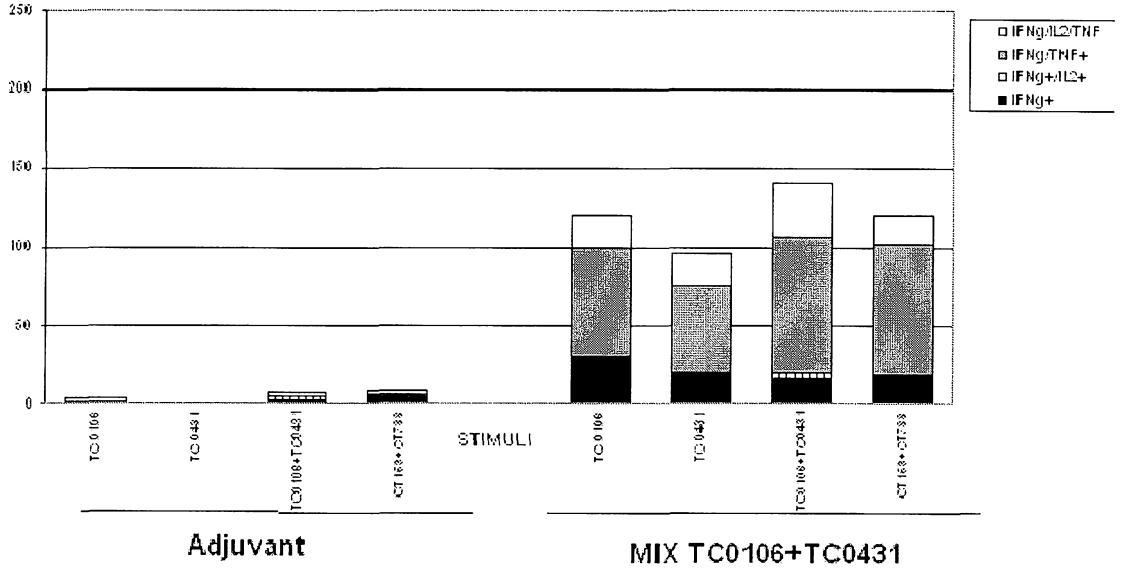

250
200
150
100
50
0
IFNg+/IL2+/TNF
IFNg+/TNF+
IFNg+/IL2+
IFNg+
TC 0 106
TC 0431
TC0 106+TC0431
CT 1688+CTF38
STIMULI
TC 0 106
TC 0431
TC0 106+TC0431
CT 1688+CTF38
Adjuvant
MIX TC0106+TC0431

FIGURE 11

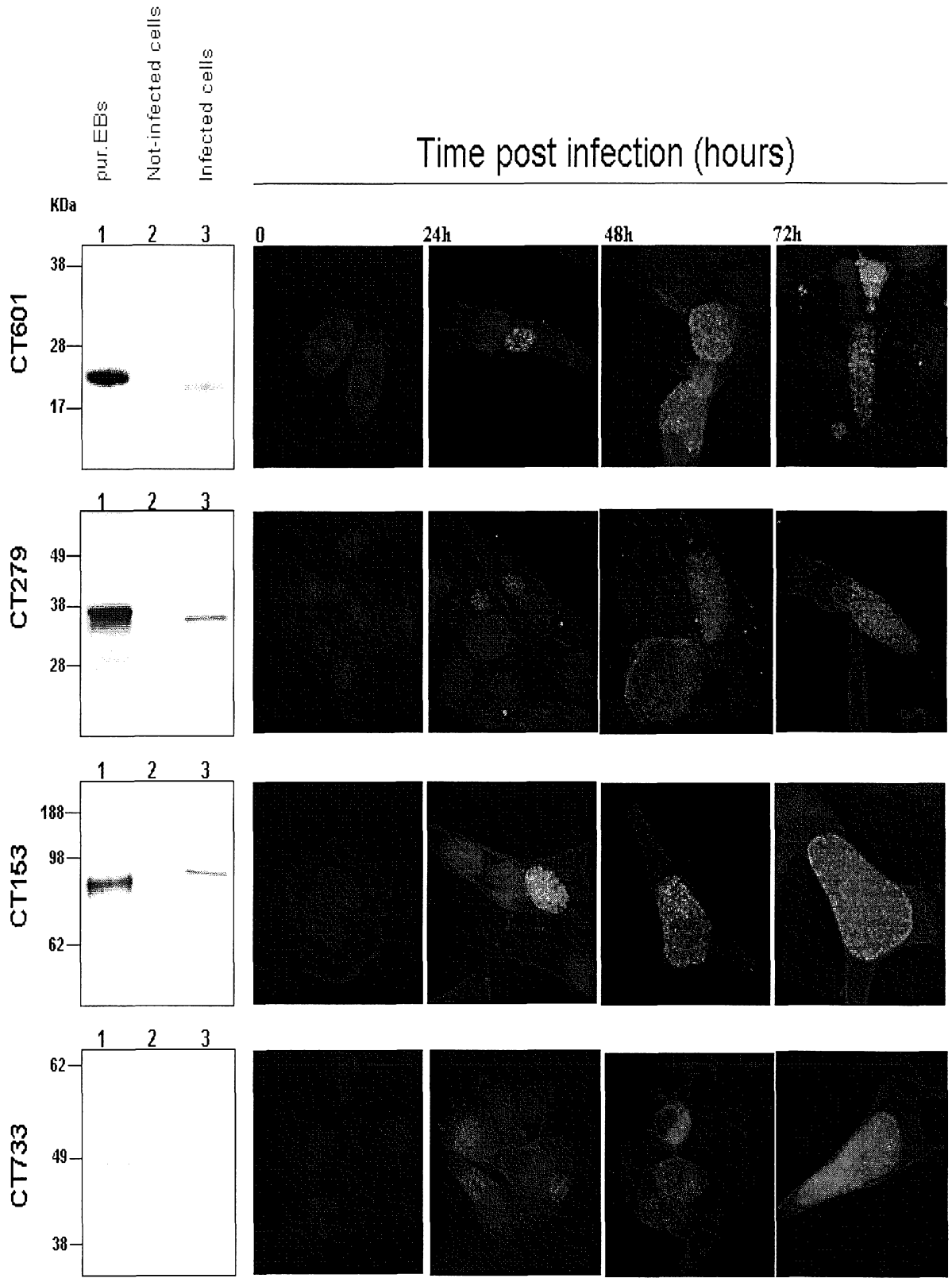

pur.EBs
Not-infected cells
Infected cells
Time post infection (hours)
KDa
1 2 3
0
24h
48h
72h
CT601
38
28
17
CT279
49
38
28
CT153
188
98
62
CT733
62
49
38

FIGURE 12

FIGURE 12A

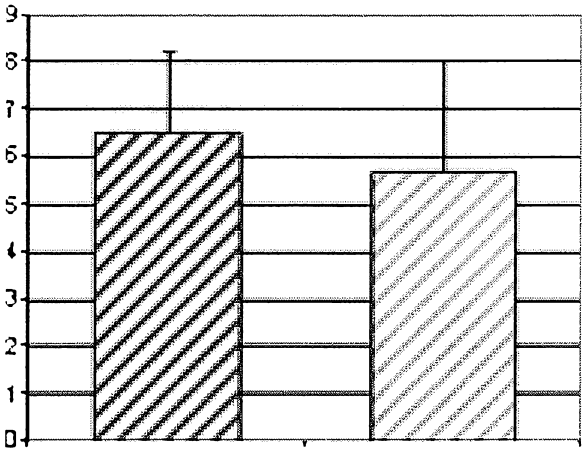


FIGURE 12B

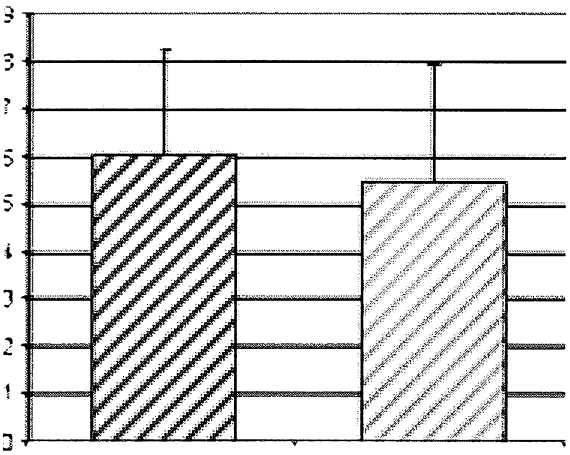


FIGURE 12C

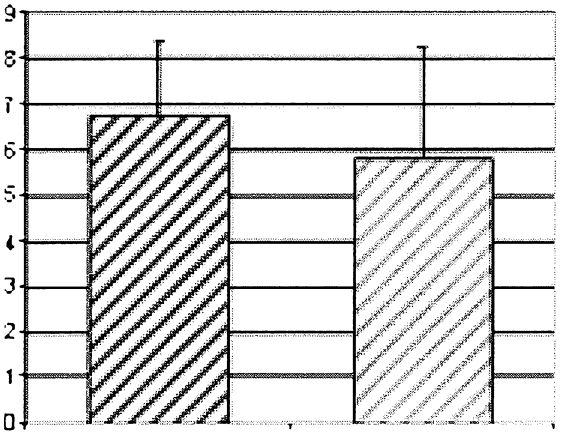


FIGURE 12D

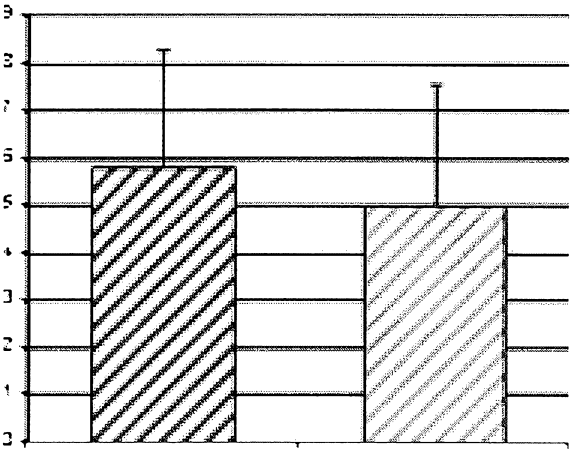
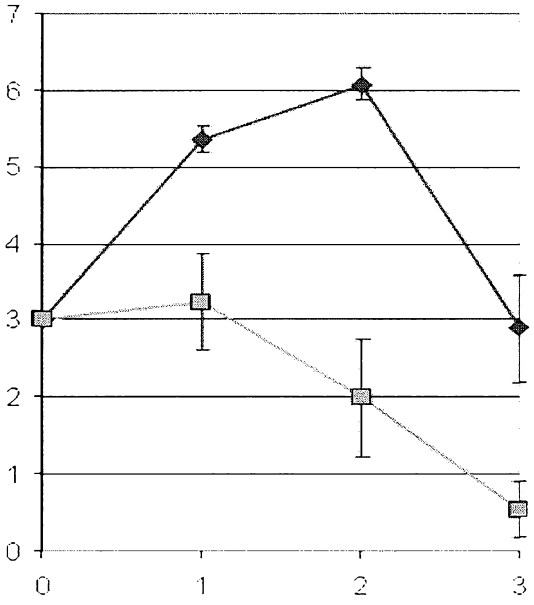


FIGURE 12E



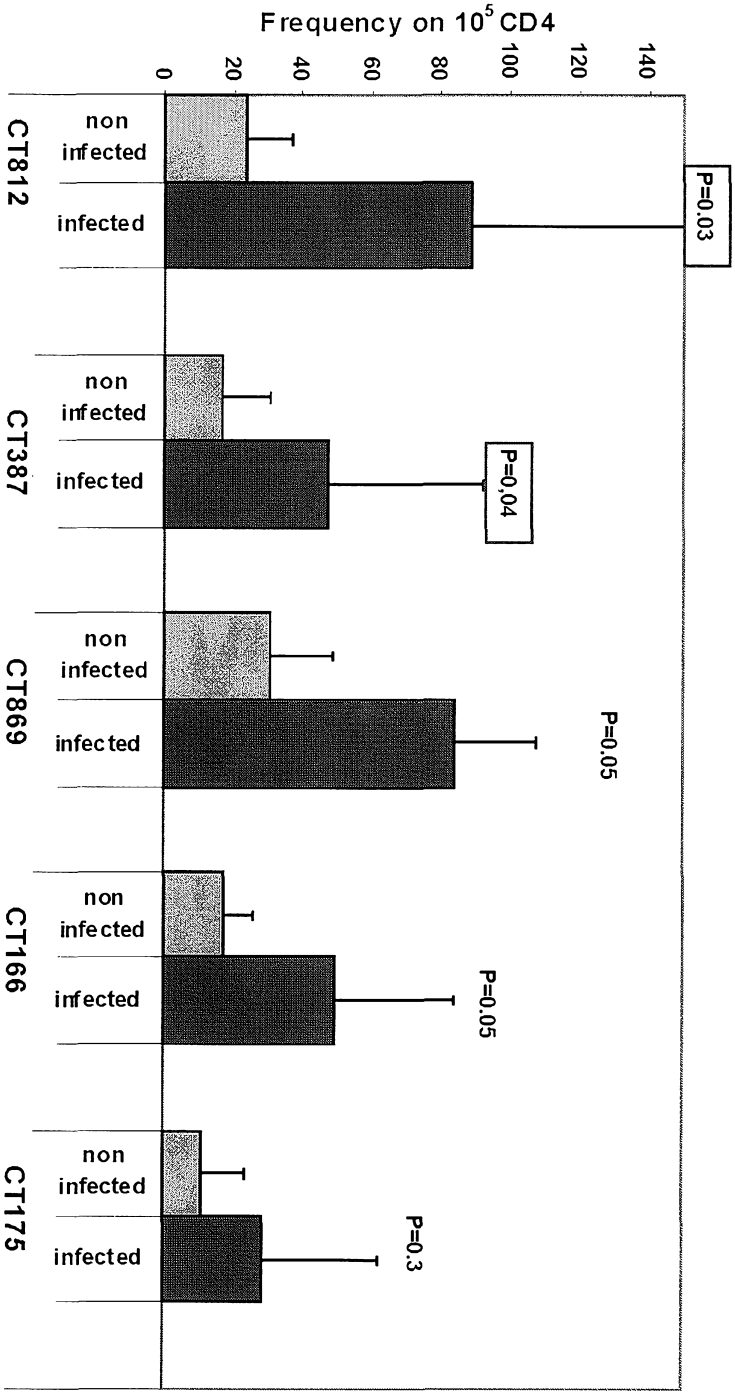

Frequency on 10^5 CD4
0
20
40
60
80
100
120
140
CT812
non infected
infected
P=0.03
CT387
non infected
infected
P=0,04
CT869
non infected
infected
P=0.05
CT166
non infected
infected
P=0.05
CT175
non infected
infected
P=0.3

FIGURE 13A

FIGURE 13B

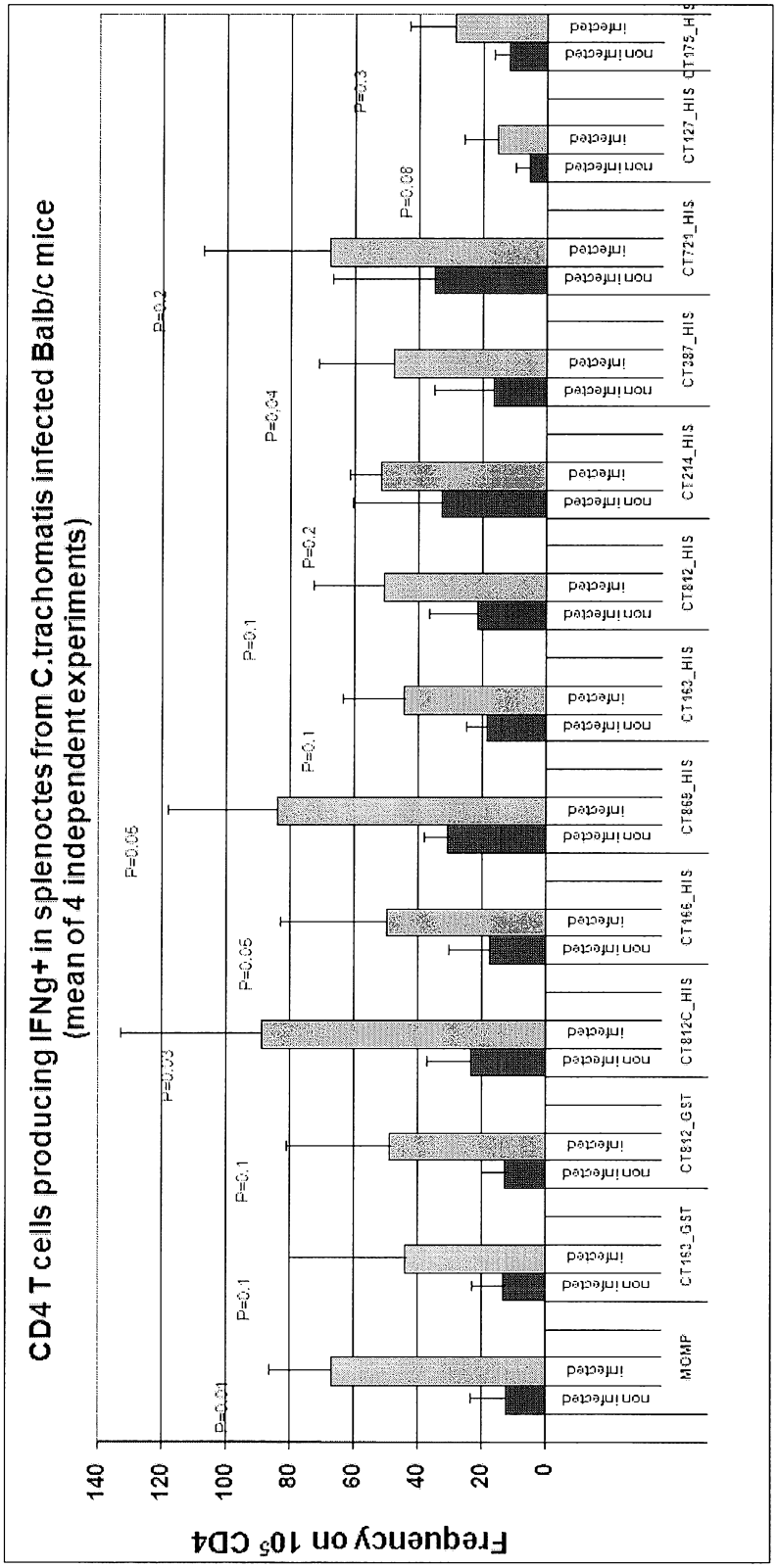

CD4 T cells producing IFNg+ in splenoctes from C.trachomatis infected Balb/c mice
(mean of 4 independent experiments)
Frequency on 10⁵ CD4
0
20
40
60
80
100
120
140
P=0.01
P=0.1
P=0.03
P=0.05
P=0.1
P=0.2
P=0.04
P=0.1
P=0.2
P=0.08
P=0.3
MOMP
CT153_GST
CT842_GST
CT812C_HIS
CT166_HIS
CT869_HIS
CT153_HIS
CT842_HIS
CT244_HIS
CT387_HIS
CT729_HIS
CT127_HIS
CT875_HIS
non infected
infected

FIGURE 14

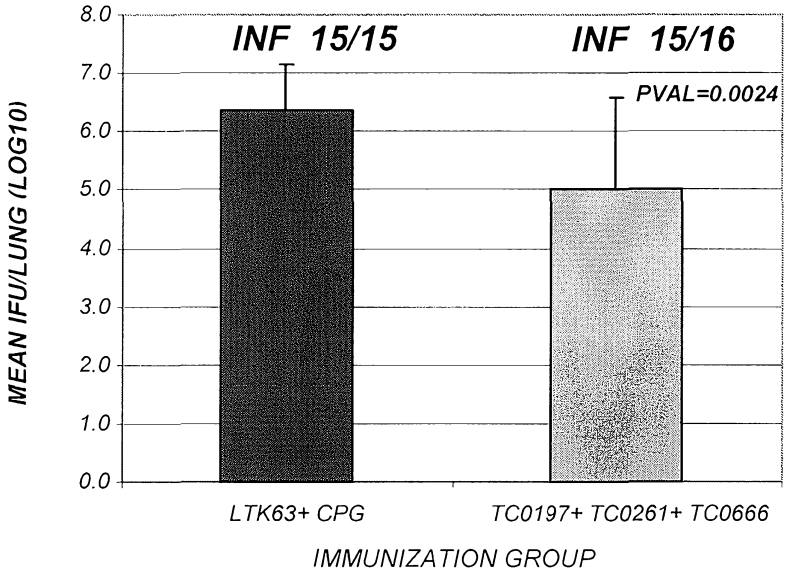


FIGURE 15