

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

(A) NAME: PEPTIDE THERAPEUTICS LIMITED

(B) STREET: 100 Fulbourn Road

(C) CITY: Cambridge

10

(D) STATE: not applicable

(E) COUNTRY: United Kingdom

(F) POSTAL CODE (ZIP): CB1 9PT

(ii) TITLE OF INVENTION: ATTENUATED BACTERIA USEFUL IN VACCINES

15

(iii) NUMBER OF SEQUENCES: 6

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

20

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

25

APPLICATION NUMBER:

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 1690 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: aroC of E.coli

(ix) FEATURE:

(A) NAME/KEY: CDS

5 (B) LOCATION:492..1562

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

GTCGACGCGG TGGATATCTC TCCAGACGCG CTGGCGGTTG CTGAACAGAA CATCGAAGAA      60
10 CACGGTCTGA TCCACAACGT CATTCCGATT CGTTCCGATC TGTTCGCGA CTTGCCGAAA      120
GTGCAGTACG ACCTGATTGT CACTAACCCG CCGTATGTCG ATGCGAAGAT ATGTCCGACC      180
TGCCAAACAA TACCGCCACG AGCCGGAACT GGGCCTGGCA TCTGGCACTG ACGGCCTGAA      240
ACTGACGCGT CGCATTCTCG GTAACGCGGC AGATTACCTT GCTGATGATG GCGTGTTGAT      300
TTGTGAAGTC GGCAACAGCA TGGTACATCT TATGGAACAA TATCCGGATG TTCCGTTTAC      360
15 CTGGCTGGAG TTTGATAACG GCGGCGATGG TGTGTTTATG CTCACCAAAG AGCAGCTTAT      420
TGCCGCACGA GAACATTTCTG CGATTATATA AGATTAAGTA AACACGCAAA CACAACAATA      480
ACGGAGCCGT G ATG GCT GGA AAC ACA ATT GGA CAA CTC TTT CGC GTA ACC      530
      Met Ala Gly Asn Thr Ile Gly Gln Leu Phe Arg Val Thr
              1             5             10

20 ACC TTC GGC GAA TCG CAC GGG CTG GCG CTC GGC TGC ATC GTC GAT GGT      578
Thr Phe Gly Glu Ser His Gly Leu Ala Leu Gly Cys Ile Val Asp Gly
      15             20             25

25 GTT CCG CCA GGC ATT CCG CTG ACG GAA GCG GAC CTG CAA CAT GAC CTC      626
Val Pro Pro Gly Ile Pro Leu Thr Glu Ala Asp Leu Gln His Asp Leu
      30             35             40             45

GAC CGT CGT CGC CCT GGG ACA TCG CGC TAT ACC ACC CAG CGC CGC GAG      674
30 Asp Arg Arg Arg Pro Gly Thr Ser Arg Tyr Thr Thr Gln Arg Arg Glu
      50             55             60

CCG GAT CAG GTC AAA ATT CTC TCC GGT GTT TTT GAA GGC GTT ACT ACC      722
Pro Asp Gln Val Lys Ile Leu Ser Gly Val Phe Glu Gly Val Thr Thr
35             65             70             75

GGC ACC AGC ATT GGC TTG TTG ATC GAA AAC ACT GAC CAG CGC TCT CAG      770

```

	Gly Thr Ser Ile Gly Leu Leu Ile Glu Asn Thr Asp Gln Arg Ser Gln	
	80 85 90	
	GAT TAC AGT GCG ATT AAG GAC GTT TTC CGT CCA GGC CAT GCC GAT TAC	818
5	Asp Tyr Ser Ala Ile Lys Asp Val Phe Arg Pro Gly His Ala Asp Tyr	
	95 100 105	
	ACC TAC GAA CAA AAA TAC GGT CTG CGC GAT TAT CGC GGC GGT GGA CGT	866
	Thr Tyr Glu Gln Lys Tyr Gly Leu Arg Asp Tyr Arg Gly Gly Gly Arg	
10	110 115 120 125	
	TCT TCC GCC CGC GAA ACC GCC ATG CGC GTG GCG GCA GGA GCT ATT GCC	914
	Ser Ser Ala Arg Glu Thr Ala Met Arg Val Ala Ala Gly Ala Ile Ala	
	130 135 140	
15	AAA AAA TAT CTC GCC GAG AAA TTT GGT ATT GAA ATC CGT GGC TGC CTG	962
	Lys Lys Tyr Leu Ala Glu Lys Phe Gly Ile Glu Ile Arg Gly Cys Leu	
	145 150 155	
20	ACC CAG ATG GGC GAC ATT CCG CTG GAT ATC AAA GAC TGG TCG CAG GTC	1010
	Thr Gln Met Gly Asp Ile Pro Leu Asp Ile Lys Asp Trp Ser Gln Val	
	160 165 170	
	GAG CAA AAT CCG TTT TTT TGC CCG GAC CCC GAC AAA ATC GAC GCG TTA	1058
25	Glu Gln Asn Pro Phe Phe Cys Pro Asp Pro Asp Lys Ile Asp Ala Leu	
	175 180 185	
	GAC GAG TTG ATG CGT GCG CTG AAA AAA GAG GGC GAC TCC ATC GGC GCT	1106
	Asp Glu Leu Met Arg Ala Leu Lys Lys Glu Gly Asp Ser Ile Gly Ala	
30	190 195 200 205	
	AAA GTC ACC GTT GTT GCC AGT GGC GTT CCT GCC GGA CTT GGC GAG CCG	1154
	Lys Val Thr Val Val Ala Ser Gly Val Pro Ala Gly Leu Gly Glu Pro	
	210 215 220	
35	GTC TTT GAC CGC CTG GAT GCT GAC ATC GCC CAT GCG CTG ATG AGC ATC	1202
	Val Phe Asp Arg Leu Asp Ala Asp Ile Ala His Ala Leu Met Ser Ile	

	225	230	235	
	AAC GCG GTG AAA GGC GTG GAA ATT GGC GAC GGC TTT GAC GTG GTG GCG			1250
	Asn Ala Val Lys Gly Val Glu Ile Gly Asp Gly Phe Asp Val Val Ala			
5	240	245	250	
	CTG CGC GGC AGC CAG AAC CGC GAT GAA ATC ACC AAA GAC GGT TTC CAG			1298
	Leu Arg Gly Ser Gln Asn Arg Asp Glu Ile Thr Lys Asp Gly Phe Gln			
	255	260	265	
10				
	AGC AAC CAT GCG GGC GGC ATT CTC GGC GGT ATC AGC AGC GGG CAG CAA			1346
	Ser Asn His Ala Gly Gly Ile Leu Gly Gly Ile Ser Ser Gly Gln Gln			
	270	275	280	285
15				
	ATC ATT GCC CAT ATG GCG CTG AAA CCG ACC TCC AGC ATT ACC GTG CCG			1394
	Ile Ile Ala His Met Ala Leu Lys Pro Thr Ser Ser Ile Thr Val Pro			
	290	295	300	
	GGT CGT ACC ATT AAC CGC TTT GGC GAA GAA GTT GAG ATG ATC ACC AAA			1442
20	Gly Arg Thr Ile Asn Arg Phe Gly Glu Glu Val Glu Met Ile Thr Lys			
	305	310	315	
	GGC CGT CAC GAT CCC TGT GTC GGG ATC CGC GCA GTG CCG ATC GCA GAA			1490
	Gly Arg His Asp Pro Cys Val Gly Ile Arg Ala Val Pro Ile Ala Glu			
25	320	325	330	
	GCG AAT GCT GGC GAT CGT TTT AAT GGA TCA CCT GTT ACG GCA ACG GGC			1538
	Ala Asn Ala Gly Asp Arg Phe Asn Gly Ser Pro Val Thr Ala Thr Gly			
	335	340	345	
30				
	GCA AAA TGC CGA TGT GAA GAC TGA TATTCACGC TGGTAAAAA TGAATAAAAC			1592
	Ala Lys Cys Arg Cys Glu Asp *			
	350	355		
35				
	CGCGATTGCG CTGCTGGCTC TGCTTGCCAG TAGCGCCAGC CTGGCAGCGA CGCCGTGGCA			1652
	AAAAATAACC CAACCTGTGC CGGGTAGCGC CAAATCGA			1690

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 356 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Gly Asn Thr Ile Gly Gln Leu Phe Arg Val Thr Thr Phe Gly
 1 5 10 15

Glu Ser His Gly Leu Ala Leu Gly Cys Ile Val Asp Gly Val Pro Pro
 20 25 30

Gly Ile Pro Leu Thr Glu Ala Asp Leu Gln His Asp Leu Asp Arg Arg
 35 40 45

20 Arg Pro Gly Thr Ser Arg Tyr Thr Thr Gln Arg Arg Glu Pro Asp Gln
 50 55 60

Val Lys Ile Leu Ser Gly Val Phe Glu Gly Val Thr Thr Gly Thr Ser
 25 65 70 75 80

Ile Gly Leu Leu Ile Glu Asn Thr Asp Gln Arg Ser Gln Asp Tyr Ser
 85 90 95

30 Ala Ile Lys Asp Val Phe Arg Pro Gly His Ala Asp Tyr Thr Tyr Glu
 100 105 110

Gln Lys Tyr Gly Leu Arg Asp Tyr Arg Gly Gly Gly Arg Ser Ser Ala
 115 120 125

35 Arg Glu Thr Ala Met Arg Val Ala Ala Gly Ala Ile Ala Lys Lys Tyr
 130 135 140

Leu Ala Glu Lys Phe Gly Ile Glu Ile Arg Gly Cys Leu Thr Gln Met
 145 150 155 160

Gly Asp Ile Pro Leu Asp Ile Lys Asp Trp Ser Gln Val Glu Gln Asn
 5 165 170 175

Pro Phe Phe Cys Pro Asp Pro Asp Lys Ile Asp Ala Leu Asp Glu Leu
 180 185 190

10 Met Arg Ala Leu Lys Lys Glu Gly Asp Ser Ile Gly Ala Lys Val Thr
 195 200 205

Val Val Ala Ser Gly Val Pro Ala Gly Leu Gly Glu Pro Val Phe Asp
 210 215 220

15 Arg Leu Asp Ala Asp Ile Ala His Ala Leu Met Ser Ile Asn Ala Val
 225 230 235 240

Lys Gly Val Glu Ile Gly Asp Gly Phe Asp Val Val Ala Leu Arg Gly
 20 245 250 255

Ser Gln Asn Arg Asp Glu Ile Thr Lys Asp Gly Phe Gln Ser Asn His
 260 265 270

25 Ala Gly Gly Ile Leu Gly Gly Ile Ser Ser Gly Gln Gln Ile Ile Ala
 275 280 285

His Met Ala Leu Lys Pro Thr Ser Ser Ile Thr Val Pro Gly Arg Thr
 290 295 300

30 Ile Asn Arg Phe Gly Glu Glu Val Glu Met Ile Thr Lys Gly Arg His
 305 310 315 320

35 Asp Pro Cys Val Gly Ile Arg Ala Val Pro Ile Ala Glu Ala Asn Ala
 325 330 335

Gly Asp Arg Phe Asn Gly Ser Pro Val Thr Ala Thr Gly Ala Lys Cys
 340 345 350

Arg Cys Glu Asp *
 5 355

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 1713 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: ompC of E.coli

20 (ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 491..1594

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GTTAACAAGC GTTATAGTTT TTCTGTGGTA GCACAGAATA ATGAAAAGTG TGTAAGAAG 60
 GGTAACAAAA ACCGAATGCG AGGCATCCGG TTGAAATAGG GGTAACAGA CATTGAGAAA 120
 30 TGAATGACGG TAATAAATAA AGTTAATGAT GATAGCGGGA GTTATTCTAG TTGCGAGTGA 180
 AGGTTTTGTT TTGACATTCA GTGCTGTCAA ATACTTAAGA ATAAGTTATT GATTTTAACC 240
 35 TTGAATTATT ATTGCTTGAT GTTAGGTGCT TATTTGCCA TTCCGCAATA ATCTTAAAAA 300
 GTTCCCTTGC ATTTACATTT TGAACATCT ATAGCGATAA ATGAAACATC TAAAAGTTT 360

	TAGTATCATA TTCGTGTTGG ATTATTCTGC ATTTTGGGG AGAATGGACT TGCCGACTGA	420
	TTAATGAGGG TTAATCAGTA TGCAGTGGCA TAAAAAGCA AATAAAGGCA TATAACAGAG	480
5	GGTTAATAAC ATG AAA GTT AAA GTA CTG TCC CTC CTG GTC CCA GCT CTG Met Lys Val Lys Val Leu Ser Leu Leu Val Pro Ala Leu 360 365 370	529
10	CTG GTA GCA GGC GCA GCA AAC GCT GCT GAA GTT TAC AAC AAA GAC GGC Leu Val Ala Gly Ala Ala Asn Ala Ala Glu Val Tyr Asn Lys Asp Gly 375 380 385	577
15	AAC AAA TTA GAT CTG TAC GGT AAA GTA GAC GGC CTG CAC TAT TTC TCT Asn Lys Leu Asp Leu Tyr Gly Lys Val Asp Gly Leu His Tyr Phe Ser 390 395 400	625
20	GAC AAC AAA GAT GTA GAT GGC GAC CAG ACC TAC ATG CGT CTT GGC TTC Asp Asn Lys Asp Val Asp Gly Asp Gln Thr Tyr Met Arg Leu Gly Phe 405 410 415	673
25	AAA GGT GAA ACT CAG GTT ACT GAC CAG CTG ACC GGT TAC GGC CAG TGG Lys Gly Glu Thr Gln Val Thr Asp Gln Leu Thr Gly Tyr Gly Gln Trp 420 425 430	721
	GAA TAT CAG ATC CAG GGC AAC AGC GCT GAA AAC GAA AAC AAC TCC TGG Glu Tyr Gln Ile Gln Gly Asn Ser Ala Glu Asn Glu Asn Asn Ser Trp 435 440 445 450	769
30	ACC CGT GTG GCA TTC GCA GGT CTG AAA TTC CAG GAT GTG GGT TCT TTC Thr Arg Val Ala Phe Ala Gly Leu Lys Phe Gln Asp Val Gly Ser Phe 455 460 465	817
35	GAC TAC GGT CGT AAC TAC GGC GTT GTT TAT GAC GTA ACT TCC TGG ACC Asp Tyr Gly Arg Asn Tyr Gly Val Val Tyr Asp Val Thr Ser Trp Thr 470 475 480	865

	GAC GTA CTG CCA GAA TTC GGT GGT GAC ACC TAC GGT TCT GAC AAC TTC	913
	Asp Val Leu Pro Glu Phe Gly Gly Asp Thr Tyr Gly Ser Asp Asn Phe	
	485 490 495	
5	ATG CAG CAG CGT GGT AAC GGC TTC GCG ACC TAC CGT AAC ACT GAC TTC	961
	Met Gln Gln Arg Gly Asn Gly Phe Ala Thr Tyr Arg Asn Thr Asp Phe	
	500 505 510	
	TTC GGT CTG GTT GAC GGC CTG AAC TTT GCT GTT CAG TAC CAG GGT AAA	1009
10	Phe Gly Leu Val Asp Gly Leu Asn Phe Ala Val Gln Tyr Gln Gly Lys	
	515 520 525 530	
	AAC GGC AAC CCA TCT GGT GAA GGC TTT ACT AGT GGC GTA ACT AAC AAC	1057
	Asn Gly Asn Pro Ser Gly Glu Gly Phe Thr Ser Gly Val Thr Asn Asn	
15	535 540 545	
	GGT CGT GAC GCA CTG CGT CAA AAC GGC GAC GGC GTC GGC GGT TCT ATC	1105
	Gly Arg Asp Ala Leu Arg Gln Asn Gly Asp Gly Val Gly Gly Ser Ile	
	550 555 560	
20	ACT TAT GAT TAC GAA GGT TTC GGT ATC GGT GGT GCG ATC TCC AGC TCC	1153
	Thr Tyr Asp Tyr Glu Gly Phe Gly Ile Gly Gly Ala Ile Ser Ser Ser	
	565 570 575	
25	AAA CGT ACT GAT GCT CAG AAC ACC GCT GCT TAC ATC GGT AAC GGC GAC	1201
	Lys Arg Thr Asp Ala Gln Asn Thr Ala Ala Tyr Ile Gly Asn Gly Asp	
	580 585 590	
	CGT GCT GAA ACC TAC ACT GGT GGT CTG AAA TAC GAC GCT AAC AAC ATC	1249
30	Arg Ala Glu Thr Tyr Thr Gly Gly Leu Lys Tyr Asp Ala Asn Asn Ile	
	595 600 605 610	
	TAC CTG GCT GCT CAG TAC ACC CAG ACC TAC AAC GCA ACT CGC GTA GGT	1297
	Tyr Leu Ala Ala Gln Tyr Thr Gln Thr Tyr Asn Ala Thr Arg Val Gly	
35	615 620 625	
	TCC CTG GGT TGG GCG AAC AAA GCA CAG AAC TTC GAA GCT GTT GCT CAG	1345

	Ser Leu Gly Trp Ala Asn Lys Ala Gln Asn Phe Glu Ala Val Ala Gln			
	630	635	640	
	TAC CAG TTC GAC TTC GGT CTG CGT CCG TCC CTG GCT TAC CTG CAG TCT	1393		
5	Tyr Gln Phe Asp Phe Gly Leu Arg Pro Ser Leu Ala Tyr Leu Gln Ser			
	645	650	655	
	AAA GGT AAA AAC CTG GGT CGT GGC TAC GAC GAC GAA GAT ATC CTG AAA	1441		
	Lys Gly Lys Asn Leu Gly Arg Gly Tyr Asp Asp Glu Asp Ile Leu Lys			
10	660	665	670	
	TAT GTT GAT GTT GGT GCT ACC TAC TAC TTC AAC AAA AAC ATG TCC ACC	1489		
	Tyr Val Asp Val Gly Ala Thr Tyr Tyr Phe Asn Lys Asn Met Ser Thr			
	675	680	685	690
15				
	TAC GTT GAC TAC AAA ATC AAC CTG CTG GAC GAC AAC CAG TTC ACT CGT	1537		
	Tyr Val Asp Tyr Lys Ile Asn Leu Leu Asp Asp Asn Gln Phe Thr Arg			
	695	700	705	
20				
	GAC GCT GGC ATC AAC ACT GAT AAC ATC GTA GCT CTG GGT CTG GTT TAC	1585		
	Asp Ala Gly Ile Asn Thr Asp Asn Ile Val Ala Leu Gly Leu Val Tyr			
	710	715	720	
	CAG TTC TAA TCTCGATTGA TATCGAACAA GGGCCTGCGG GCCCTTTTTT	1634		
25	Gln Phe *			
	725			
	CATTGTTTTT AGCGTACAAA CTCAGTTTTT TGGTGTACTC TTGCGACCGT TCGCATGAGG	1694		
30	ATAATCACGT ACGGAAATA			
			1713	

(2) INFORMATION FOR SEQ ID NO: 4:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

5 Met Lys Val Lys Val Leu Ser Leu Leu Val Pro Ala Leu Leu Val Ala
 1 5 10 15
 Gly Ala Ala Asn Ala Ala Glu Val Tyr Asn Lys Asp Gly Asn Lys Leu
 10 20 25 30
 Asp Leu Tyr Gly Lys Val Asp Gly Leu His Tyr Phe Ser Asp Asn Lys
 35 40 45
 15 Asp Val Asp Gly Asp Gln Thr Tyr Met Arg Leu Gly Phe Lys Gly Glu
 50 55 60
 Thr Gln Val Thr Asp Gln Leu Thr Gly Tyr Gly Gln Trp Glu Tyr Gln
 65 70 75 80
 20 Ile Gln Gly Asn Ser Ala Glu Asn Glu Asn Asn Ser Trp Thr Arg Val
 85 90 95
 Ala Phe Ala Gly Leu Lys Phe Gln Asp Val Gly Ser Phe Asp Tyr Gly
 25 100 105 110
 Arg Asn Tyr Gly Val Val Tyr Asp Val Thr Ser Trp Thr Asp Val Leu
 115 120 125
 30 Pro Glu Phe Gly Gly Asp Thr Tyr Gly Ser Asp Asn Phe Met Gln Gln
 130 135 140
 Arg Gly Asn Gly Phe Ala Thr Tyr Arg Asn Thr Asp Phe Phe Gly Leu
 145 150 155 160
 35 Val Asp Gly Leu Asn Phe Ala Val Gln Tyr Gln Gly Lys Asn Gly Asn
 165 170 175

	Pro Ser Gly Glu Gly Phe Thr Ser Gly Val Thr Asn Asn Gly Arg Asp	
	180	185 190
5	Ala Leu Arg Gln Asn Gly Asp Gly Val Gly Gly Ser Ile Thr Tyr Asp	
	195	200 205
	Tyr Glu Gly Phe Gly Ile Gly Gly Ala Ile Ser Ser Ser Lys Arg Thr	
	210	215 220
10	Asp Ala Gln Asn Thr Ala Ala Tyr Ile Gly Asn Gly Asp Arg Ala Glu	
	225	230 235 240
	Thr Tyr Thr Gly Gly Leu Lys Tyr Asp Ala Asn Asn Ile Tyr Leu Ala	
	245	250 255
15	Ala Gln Tyr Thr Gln Thr Tyr Asn Ala Thr Arg Val Gly Ser Leu Gly	
	260	265 270
	Trp Ala Asn Lys Ala Gln Asn Phe Glu Ala Val Ala Gln Tyr Gln Phe	
20	275	280 285
	Asp Phe Gly Leu Arg Pro Ser Leu Ala Tyr Leu Gln Ser Lys Gly Lys	
	290	295 300
25	Asn Leu Gly Arg Gly Tyr Asp Asp Glu Asp Ile Leu Lys Tyr Val Asp	
	305	310 315 320
	Val Gly Ala Thr Tyr Tyr Phe Asn Lys Asn Met Ser Thr Tyr Val Asp	
	325	330 335
30	Tyr Lys Ile Asn Leu Leu Asp Asp Asn Gln Phe Thr Arg Asp Ala Gly	
	340	345 350
	Ile Asn Thr Asp Asn Ile Val Ala Leu Gly Leu Val Tyr Gln Phe *	
35	355	360 365

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1808 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: ompF of E.coli

(ix) FEATURE:

- 15 (A) NAME/KEY: CDS
 (B) LOCATION: 457..1545

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

20

AAACTAATC CGCATTCTTA TTGCGGATTA GTTTTTCTT AGCTAATAGC ACAATTTTCA 60

TACTATTTTT TGGCATTCTG GATGTCTGAA AGAAGATTTT GTGCCAGGTC GATAAAGTTT 120

25 CCATCAGAAA CAAAATTTCC GTTTAGTTAA TTAAATATA AGGAAATCAT ATAAATAGAT 180

TAAAATTGCT GTAAATATCA TCACGTCTCT ATGGAAATAT GACGGTGTTC ACAAAGTTCC 240

TTAAATTTTA CTTTGGTTA CATATTTTTT CTTTTGAAA CCAAATCTTT ATCTTTGTAG 300

30

CACTTTCACG GTAGCGAAAC GTTAGTTTGA ATGGAAAGAT GCCTGCAGAC ACATAAAGAC 360

ACCAAATCT CATCAATAGT TCCGTAAATT TTTATTGACA GAACTTATTG ACGGCAGTGG 420

35 CAGGTGTCAT AAAAAAACC ATGAGGGTAA TAAATA ATG ATG AAG CGC AAT ATT 474

Met Met Lys Arg Asn Ile

1

5

	CTG GCA GTG ATC GTC CCT GCT CTG TTA GTA GCA GGT ACT GCA AAC GCT	522
	Leu Ala Val Ile Val Pro Ala Leu Leu Val Ala Gly Thr Ala Asn Ala	
	10 15 20	
5	GCA GAA ATC TAT AAC AAA GAT GGC AAC AAA GTA GAT CTG TAC GGT AAA	570
	Ala Glu Ile Tyr Asn Lys Asp Gly Asn Lys Val Asp Leu Tyr Gly Lys	
	25 30 35	
	GCT GTT GGT CTG CAT TAT TTT TCC AAG GGT AAC GGT GAA AAC AGT TAC	618
10	Ala Val Gly Leu His Tyr Phe Ser Lys Gly Asn Gly Glu Asn Ser Tyr	
	40 45 50	
	GGT GGC AAT GGC GAC ATG ACC TAT GCC CGT CTT GGT TTT AAA GGG GAA	666
	Gly Gly Asn Gly Asp Met Thr Tyr Ala Arg Leu Gly Phe Lys Gly Glu	
15	55 60 65 70	
	ACT CAA ATC AAT TCC GAT CTG ACC GGT TAT GGT CAG TGG GAA TAT AAC	714
	Thr Gln Ile Asn Ser Asp Leu Thr Gly Tyr Gly Gln Trp Glu Tyr Asn	
	75 80 85	
20	TTC CAG GGT AAC AAC TCT GAA GGC GCT GAC GCT CAA ACT GGT AAC AAA	762
	Phe Gln Gly Asn Asn Ser Glu Gly Ala Asp Ala Gln Thr Gly Asn Lys	
	90 95 100	
25	ACG CGT CTG GCA TTC GCG GGT CTT AAA TAC GCT GAC GTT GGT TCT TTC	810
	Thr Arg Leu Ala Phe Ala Gly Leu Lys Tyr Ala Asp Val Gly Ser Phe	
	105 110 115	
	GAT TAC GGC CGT AAC TAC GGT GTG GTT TAT GAT GCA CTG GGT TAC ACC	858
30	Asp Tyr Gly Arg Asn Tyr Gly Val Val Tyr Asp Ala Leu Gly Tyr Thr	
	120 125 130	
	GAT ATG CTG CCA GAA TTT GGT GGT GAT ACT GCA TAC AGC GAT GAC TTC	906
	Asp Met Leu Pro Glu Phe Gly Gly Asp Thr Ala Tyr Ser Asp Asp Phe	
35	135 140 145 150	
	TTC GTT GGT CGT GTT GGC GGC GTT GCT ACC TAT CGT AAC TCC AAC TTC	954

	Phe Val Gly Arg Val Gly Gly Val Ala Thr Tyr Arg Asn Ser Asn Phe	
	155 160 165	
	TTT GGT CTG GTT GAT GGC CTG AAC TTC GCT GTT CAG TAC CTG GGT AAA	1002
5	Phe Gly Leu Val Asp Gly Leu Asn Phe Ala Val Gln Tyr Leu Gly Lys	
	170 175 180	
	AAC GAG CGT GAC ACT GCA CGC CGT TCT AAC GGC GAC GGT GTT GGC GGT	1050
	Asn Glu Arg Asp Thr Ala Arg Arg Ser Asn Gly Asp Gly Val Gly Gly	
10	185 190 195	
	TCT ATC AGC TAC GAA TAC GAA GGC TTT GGT ATC GTT GGT GCT TAT GGT	1098
	Ser Ile Ser Tyr Glu Tyr Glu Gly Phe Gly Ile Val Gly Ala Tyr Gly	
	200 205 210	
15	GCA GCT GAC CGT ACC AAC CTG CAA GAA GCT CAA CCT CTT GGC AAC GGT	1146
	Ala Ala Asp Arg Thr Asn Leu Gln Glu Ala Gln Pro Leu Gly Asn Gly	
	215 220 225 230	
20	AAA AAA GCT GAA CAG TGG GCT ACT GGT CTG AAG TAC GAC GCG AAC AAC	1194
	Lys Lys Ala Glu Gln Trp Ala Thr Gly Leu Lys Tyr Asp Ala Asn Asn	
	235 240 245	
	ATC TAC CTG GCA GCG AAC TAC GGT GAA ACC CGT AAC GCT ACG CCG ATC	1242
25	Ile Tyr Leu Ala Ala Asn Tyr Gly Glu Thr Arg Asn Ala Thr Pro Ile	
	250 255 260	
	ACT AAT AAA TTT ACA AAC ACC AGC GGC TTC GCC AAC AAA ACG CAA GAC	1290
	Thr Asn Lys Phe Thr Asn Thr Ser Gly Phe Ala Asn Lys Thr Gln Asp	
30	265 270 275	
	GTT CTG TTA GTT GCG CAA TAC CAG TTC GAT TTC GGT CTG CGT CCG TCC	1338
	Val Leu Leu Val Ala Gln Tyr Gln Phe Asp Phe Gly Leu Arg Pro Ser	
	280 285 290	
35	ATC GCT TAC ACC AAA TCT AAA GCG AAA GAC GTA GAA GGT ATC GGT GAT	1386
	Ile Ala Tyr Thr Lys Ser Lys Ala Lys Asp Val Glu Gly Ile Gly Asp	

	295	300	305	310	
	GTT GAT CTG GTG AAC TAC TTT GAA GTG GGC GCA ACC TAC TAC TTC AAC				1434
	Val Asp Leu Val Asn Tyr Phe Glu Val Gly Ala Thr Tyr Tyr Phe Asn				
5		315	320	325	
	AAA AAC ATG TCC ACC TAT GTT GAC TAC ATC ATC AAC CAG ATC GAT TCT				1482
	Lys Asn Met Ser Thr Tyr Val Asp Tyr Ile Ile Asn Gln Ile Asp Ser				
		330	335	340	
10	GAC AAC AAA CTG GGC GTA GGT TCA GAC GAC ACC GTT GCT GTG GGT ATC				1530
	Asp Asn Lys Leu Gly Val Gly Ser Asp Asp Thr Val Ala Val Gly Ile				
		345	350	355	
15	GTT TAC CAG TTC TAA TAGCACCT CTTTGTTAAA TGCCGAAAAA ACAGGACTTT				1585
	Val Tyr Gln Phe *				
		360			
	GGTCCTGTTT TTTTATACC TTCCAGAGCA ATCTCAGTC TTGCAAAAAC AGCCTGCGTT				1645
20	TTCATCAGTA ATAGTTGGAA TTTTGTAAT CTCCGTTAC CCTGATAGCG GACTTCCCTT				1705
	CTGTAACCAT AATGGAACCT CGTCATGTTT GAGAACATTA CCGCCGCTCC TGCCGACCCG				1765
25	ATTCTGGGCC TGGCCGATCT GTTCGTGCC GATGAACGTC CCG				1808

(2) INFORMATION FOR SEQ ID NO: 6:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 362 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Met Lys Arg Asn Ile Leu Ala Val Ile Val Pro Ala Leu Leu Val
 1 5 10 15
 Ala Gly Thr Ala Asn Ala Ala Glu Ile Tyr Asn Lys Asp Gly Asn Lys
 5 20 25 30
 Val Asp Leu Tyr Gly Lys Ala Val Gly Leu His Tyr Phe Ser Lys Gly
 35 40 45
 10 Asn Gly Glu Asn Ser Tyr Gly Gly Asn Gly Asp Met Thr Tyr Ala Arg
 50 55 60
 Leu Gly Phe Lys Gly Glu Thr Gln Ile Asn Ser Asp Leu Thr Gly Tyr
 65 70 75 80
 15 Gly Gln Trp Glu Tyr Asn Phe Gln Gly Asn Asn Ser Glu Gly Ala Asp
 85 90 95
 Ala Gln Thr Gly Asn Lys Thr Arg Leu Ala Phe Ala Gly Leu Lys Tyr
 20 100 105 110
 Ala Asp Val Gly Ser Phe Asp Tyr Gly Arg Asn Tyr Gly Val Val Tyr
 115 120 125
 25 Asp Ala Leu Gly Tyr Thr Asp Met Leu Pro Glu Phe Gly Gly Asp Thr
 130 135 140
 Ala Tyr Ser Asp Asp Phe Phe Val Gly Arg Val Gly Gly Val Ala Thr
 145 150 155 160
 30 Tyr Arg Asn Ser Asn Phe Phe Gly Leu Val Asp Gly Leu Asn Phe Ala
 165 170 175
 Val Gln Tyr Leu Gly Lys Asn Glu Arg Asp Thr Ala Arg Arg Ser Asn
 35 180 185 190
 Gly Asp Gly Val Gly Gly Ser Ile Ser Tyr Glu Tyr Glu Gly Phe Gly

	195	200	205
	Ile Val Gly Ala Tyr Gly Ala Ala Asp Arg Thr Asn Leu Gln Glu Ala		
	210	215	220
5	Gln Pro Leu Gly Asn Gly Lys Lys Ala Glu Gln Trp Ala Thr Gly Leu		
	225	230	235 240
	Lys Tyr Asp Ala Asn Asn Ile Tyr Leu Ala Ala Asn Tyr Gly Glu Thr		
10	245	250	255
	Arg Asn Ala Thr Pro Ile Thr Asn Lys Phe Thr Asn Thr Ser Gly Phe		
	260	265	270
15	Ala Asn Lys Thr Gln Asp Val Leu Leu Val Ala Gln Tyr Gln Phe Asp		
	275	280	285
	Phe Gly Leu Arg Pro Ser Ile Ala Tyr Thr Lys Ser Lys Ala Lys Asp		
	290	295	300
20	Val Glu Gly Ile Gly Asp Val Asp Leu Val Asn Tyr Phe Glu Val Gly		
	305	310	315 320
	Ala Thr Tyr Tyr Phe Asn Lys Asn Met Ser Thr Tyr Val Asp Tyr Ile		
25	325	330	335
	Ile Asn Gln Ile Asp Ser Asp Asn Lys Leu Gly Val Gly Ser Asp Asp		
	340	345	350
30	Thr Val Ala Val Gly Ile Val Tyr Gln Phe *		
	355	360	