

Total amino acid number: 2217, MW=246060

Max ORF starts at AA pos 1(may be DNA pos 1) for 2217 AA (6651 bases)

1 ATGTGGAGTACCCTGATTATTTGTCCTTGCTGGTGCAGCGACGCCCAATGCCCAATACCCGCTGG
1 M W T P V L L C L L V A A A N A Q Y G W

61 AAAAAATGGAAACCTCTACAAAATACGAAATCAACGGACGCACCCTGACCGCCCTGAACCAA
21 K N G N L Y K Y E I N G R T L T A L N Q

121 GTCGCCGACCAATACGCCGGAGTCTTATTTCAGAGCCAATTCTACGTCCAGCCCTTCTCC
41 V A D Q Y A G V L F R A N F Y V Q P F S

181 AGTGACAGACTGTCTGCCTACATCCAAAATGCCGAGACCGCTCAAGTTCACGCTGAGCTT
61 S D R L S A Y I Q N A E T A Q V H A E L

241 CCAAGTGGATATGAATCTCACATCCCCCTCAGCCAGTTGAACTACAAGAGTATGCCCCCTC
81 P S G Y E S H I P S S Q L N Y K S M P L

301 AGCCACGAGCCATTTCGAAATTTACCTGAAGAAGGGAGTCGTCTCCAACCTTCGCGTCAAC
101 S H E P F E I Y L K K G V V S N L R V N

361 AAGAATGTCTCCGACTGGGAACTTAACATCATCAAGGCTGTTGTGAGCCAAATCCAAGTT
121 K N V S D W E L N I I K A V V S Q I Q V

421 GACACCCAAGGTCAAAACTTGAAGAAATCCAGCCACAACCAACTCCCCAAGGAAAAACAAG
141 D T Q G Q N L K K S S H N Q L P K E N K

481 CCCTACGGTGTTTACAAGACCATGGAAGACTCCGTCACCGGTGAATGTGAAACTCTCTAC
161 P Y G V Y K T M E D S V T G E C E T L Y

541 GATGTCTCACCTCTGCCAGAAATCACCCCTCCAAACCAAACCCCTGGTTGGTTCCTTTCCCC
181 D V S P L P E I T L Q T K P W L V P F P

601 AACTTCCGTGAAAACGACAATTTCATCGACATCGTCAAGACCACCAACTACAGCAAATGT
201 N F R E N G Q F I D I V K T T N Y S K C

661 GAAGAACGTTCTGCTTACCACTTCGGTATCACTGGTCTGACCAACTGGAAACCCGCCAGC
221 E E R S A Y H F G I T G L T N W K P A S

721 AACCAAATGGGACAATTCCCTCTCCCGCTCCAACATCAACCGTGTGTCATTCTTGAAAC
241 N Q M G Q F L S R S N I N R V V I S G N

781 GTGAAATACTACACCATCCAATCCTCTGTTTTCCACCAACAAAATCGTCATCAGTCCACAG
261 V K Y Y T I Q S S V S T N K I V I S P Q

841 ATGTACGAATCACAGAAGGGAATGGTCATCAGTGTATGAACATGACCCTGGCTTCCTTC
281 M Y E S Q K G M V I S V M N M T I A S F

901 CACCAGGCCAATGGATCTCCCCGCAACGGACACAACGGACACAACGAAAACAACGGACAC
301 H Q A N G S E R N G H N G H N G N N G H

961 AACGGAAACAACGGACACAATGGACACAACGGAACAATGGACAAAAACAACAGCAATGGC
321 N G N N G H N G H N G N N G Q N N S N G

1021 AGCTCAAGCAGCAGCTCCAGCTCTAGCAGCTCTAGCTCTGGCTCCAGCAGCTCTGAAGAA
341 S S S S S S S S S S S S S S G S S S S E E

1081 AACAACAACAACAACGACAACGAAGACGAAAAGAATAACAGTGGCGCATACGCTCAACAC
361 N N N N N D N E D E K N N S G A Y A Q H

1141 TACAACAACAATGGTGCATCTAGCAGCAGCTCCTCCAGCTCCAGCTCCAGCTCTGACTCA
381 Y N N N G A S S S S S S S S S S S S S D S

1201 TCTAGTTCCAGCTCCAGCTCTAGCAGCAGCAGCTCCTCCTCCTCCTCTTCCAGCTCAAGC
401 S
1261 AGCAGCTCCGAAGAAGACCAATACTACAGAAGACACAACAACAACCGCAACAACAAC
421 S S S E E D Q Y Y R R H N N N N N R N N N
1321 AACACAACAACAACAACAACAACCGCAACAACAACCACAACAACAACAATAACGCC
441 N N N N N N N N N R N N N H N N N N N N A
1381 AACGACAACAACAATTCCGCCAACAATAACAACAACAACAATTCCGCCAACAACAACAAC
461 N D N N N S A N N N N N N N S A N N N N
1441 AACCGCAACAATGACGAGAACGTCTCTCGCAGTCGCTCCCGCAGAGATATCTCCCAATAC
481 N R N N D E N V S R S [R S R R] D I S Q Y
1501 AAATACAACAGTTTTTGAAAAACAACAACAACGACAACGAAGACGAAAAGAATAACAGT
501 K Y N S F E N N N N N D N E D E K N N S
1561 GGCCGCAACGGACACAACGGACACAACGGAAAACAACGGACACAACGGAAAACAACGGACAC
521 G R N G H N G H N G N N G H N G N N G H
1621 AATGGACACAACGGAAAACAATGGACAAAAACAACGCAATGGCAGCTCAAGCAGCAGCTCC
541 N G H N G N N G Q N N S N G S S S S S S
1681 AGCTCTAGCAGCTCTAGCTCTGGCTCCAGCAGCTCTGAAGAAAACAACCTCCCGTTACAAC
561 S S S S S S S S G S S S S E E N N S R Y N
1741 AACGGTAAATTTCGCCAGCTTTGCCCCGCCACAACGGATCTGGATCCTCCTCCTCCAGCAGC
581 N G K F A S F A R H N G S G S S S S S S S
1801 TCCCCGACTCTTCTGACTCCTCCAGCTCCTCCAGCTCATCCAGTTCTCCAGCTCCTCT
601 S P D S S D S S S S S S S S S S S S S S S
1861 AGCTCCTCCAGCTCTTCCAGCTCCTCATCCAGCTCTGAAGACAACAGCTCCTTTGGATCT
621 S S S S S S S S S S S S S E D N S S F G S
1921 TCCGTCTCCAGCAGCAGCGAAGAGGACTATGAACCACGTCCAAGCATGTACAAGGCTCCT
641 S V S S S S E E D Y E P R P S M Y K A P
1981 CAAACTCCTTTCTTCCCTACTTTCATCGGAACTACGGTAACAGTATCCAATCCGCTAAA
661 Q T P F F P Y F I G N Y G N S I Q S A K
2041 CAAGTTAACGGAGTTGCTCTTGCCCGTAAGCTCGCCCAAGAAATTGCTGAGGAATTGAAC
681 Q V N G V A L A R K L A Q E I A E E L N
2101 GACCCACGTCAAATTACTCAAAAAAGCACTTTGGCTAAATTCAACATGTTGGTTGAGGAA
701 D P R Q I T Q K S T L A K F N M L V E E
2161 CTCAGAACCCTGGACGCGAAACAAATGGAACAAGCTTCCCAAGAGCTTCATTTCAACTCC
721 L R T L D A K Q M E Q A S Q E L H F N S
2221 GCCCAGGCCTCCAGCCACAGCCGTCAAGATGCTCTGAAATCTCTCGCCTGGAAATCCTTC
741 A Q A S S H S R Q D A L K S L A W K S F
2281 TGTGATGCCTTAGTTGAAGCCGGTACCGGACCTGCCTTTTTTGCAAATCCAAAAGATCATT
761 C D A L V E A G T G P A F L Q I Q K I I
2341 GAACACCAACAAGTCTCCGACGCCGAAGCCGCTCGCATGATCAGAATCGAATTCCTCGCGG
781 E H Q Q V S D A E A A [R M I R] I E F P R
2401 CCGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATCCCTATACAACAATGTTCAAC
801 P P W R P G A C D V G P N P Y T T M F [N]
2461 ACCAGAACCAACTCAACGAAACTGCTCTCCTTGCTTTCGCTGAGCTTTGCCGTAAAGCCG
821 [T R T] [N S T K] L L S L L S L S F A V K P

2521 ATGTCAACGCCAGAAACGCCCACAACACTACTACCCTGTCCACGTTTATGGACGTGTCCTCC
 841 M S T P E T P T T T T L S T F M D V S S
 2581 CAGAGCATGCCAAGGCTGTTGCTCACCAATACCTTCCCTACTATGAACAAAACCTCAAGA
 861 Q S M P R L L L T N T F P T M N K T S R
 2641 GAGCCGTTGCCAACGGTGACAGCCGCAAGATCCAAGCTTACATCCGTGCCATTGGAACT
 881 E P L P T V T A A R S K L T S V P L E T
 2701 TTGCTCACCCCAAGATCCTCGAAGTTTTTCGAGCCCTACCTTGAAGGAAAGGTCCCAATCT
 901 L L T P R S S K F S S P T L K E R S Q S
 2761 CCAACTTCCAACGCACTGTCATGGTCCTCTCCCTTAACGAACTTGCACGCGTCTACCCCA
 921 P T S N A L S W S S P L T N L H A S T P
 2821 ACCTTGCCCGCAATGTCTCTTCAAGATCTACCAAAACACCCAAGAAAACCAAGAAGTCC
 941 T L P A M S S S R S T K T P K K T K K S
 2881 GTGTTGCCGCTGTCTTCTTAATCTTTCGGAACCAACCCATCTGCCAGACCCCTCCAACGTA
 961 V L P L S S S S S E P T H L P R P S N V
 2941 TGGCTCAATTACCAATGAAGACCAAGACCAGCAAGTCAACGCTGCCGTCAAATTCCGCC
 981 W L N S P M K T K T S K S T L P S N S A
 3001 CTTGAAAACGCTGCCAAGGCCCCACTCCGAGTCTCGTCAAGAACTCGCCCAAGCTGCTCAA
 1001 L E N A A K A H S E S R Q E L A Q A A Q
 3061 TCCGCTATTGCTCTCCTCAGCCCTAAGGCTTACGGTCTTCAATACTCCAAGAAATGGCTC
 1021 S A I A L L S P K A Y G L Q Y S K K W L
 3121 CGCGACTACATCGTCAAGGAAGAAAACCTTGCCTACAGAGTGCTTGCCGATATGATCCAG
 1041 R D Y I V K E E N L A Y R V S A D M I Q
 3181 AGTGAGGACTCCTTAATCCCCAACCAAGTTTACGTTGCCCTTACCCGTTACCTCGGTGGA
 1061 S E D S L I P N Q V Y V A L H R Y L G G
 3241 TTCGCTCAACGTGTTGCCAGCTTCAGAGCCATGACCTCCAGTGCCAGTGACCTCGTTGAA
 1081 F A Q R V A S F R A M T S S A S D L V E
 3301 AAGATCCAAGAACAGTTACCAATGGAGAAGAATACCAACAACAATCCGAAATGAACCAG
 1101 K I Q E Q F T N G E E Y Q Q Q S E M N Q
 3361 CAATTCTCTGCCGAACAAATCTTCAGACAGTTCAACATTAAGCCTGACTACCCACAAGAA
 1121 Q F S A E Q I F R Q F N I K P D Y P Q E
 3421 GTTGAAGCCCTCCTCCAATACACCGTTTTTCGAGCCAAGCGCTGGGCTTCTTCGACGAA
 1141 V E A L L Q Y T V F G A K R W A F F D E
 3481 GAATTCTTCAACCAATCCCAAGACGTCTGAACGATGCTCTCTCTAAGGTTCAAAATGGA
 1161 E F F N Q I P R R L N D A L S K V Q N G
 3541 CAATCTTTCAACTCCACCAAATTCTACAACGACATCTCTCTCTCCCTTGCTTTCCCCACT
 1181 Q S F N S T K F Y N D I S L S L A F P T
 3601 GCCACTGGTCTGCCATTCTCTTACACTCTCAAGGTCCCCACTCTCGTCCAAGCCGGTGA
 1201 A T G L P F S Y T L K V P T L V Q A G G
 3661 GAAGTTCAAGCCAGAGTTCAAGGTCACAACTCCAACAACAACAACCTCTTCCGTATC
 1221 E V Q A R V Q G H N S N N N N N L F R I
 3721 CCTGAAGCTGTTAACGTAACCGCCGAAATTGAAATTGTTTACGCCACTGAATTGAAATCT
 1241 P E A V N V T A E I E I V Y A T E L K S
 3781 GAACTTGGATTGCTCACTCCATTCAACCACGAACGTTACGTTGCTGGTCTTGCCAAGAAC
 1261 E L G F V T P F N H E R Y V A G L A K N

3841 ATCTTTGTAAACATCCCAGTTAAGGTCGCCGCAATGTGGACATCGCCAACACCAAGGTC
1281 I F V N I P V K V A A N V D I A N T K V
3901 GAATTCTACATGAAGCCCATGAACAACGAGAACGAAAAAATCTTCCACTACGGATCT
1301 E F Y M K P M N N Q N E Q K I F H Y G S
3961 TACCCTTACACCGCAATCCAAAACATCTTCGACTTCCGCCCCCTCCAAGAAAACGAAAAC
1321 Y P Y T A I Q N I F D F R P L Q E N E N
4021 ACTAAGTACATCTTCGCCAACGAGAACAAGAACAAGTTCGAGAAAGTCTACGGTGAAGAA
1341 T K Y I F A N E N K N K F E K V Y G E E
4081 AAGACCGGTTTTGCCTTCCGTTGCCAATACAAGGGTGACCAACAAAGCTTCCAATTTCGCT
1361 K T G F A F R C Q Y K G D Q Q S F Q F A
4141 GACTTCTACAACCTTCGCTAAGCGTAACGATTTCTTCTCCGCTGCTTTCTTCCCATGGGCT
1381 D F Y N F A K R N D F F S A A F F P W A
4201 GAAAAGACCATCCAATACAACAACCTTCGATGGCTACTATGACCCAGCTAGATCTGCCGCC
1401 E K T I Q Y N N F D G Y Y D P A R S A A
4261 AAATCCGCCAAGTTCGCCTTGAACCTATGCCAACAAATACGCCAACAAAGGAAAACAACAAC
1421 K S A K F A L N Y A N K Y A N K E N N N
4321 GAAGGTAACCTCTAACAACCACAACAACCAACGATGCTGTTCCATCTTCTTACCAA
1441 E G N S N N H N N N H N D A V P S S Y Q
4381 CCTGACAACGAACAGAGAATGAACCAATTCCGCCAGCCGCGCTCAATCTGGAGTTCAAAGT
1461 P D N E Q R M N Q F A S R A Q S G V Q S
4441 GCCAACATCGATGTTATCGACATCTCTGCTCAATTCTTCGGACAAAAGAAGCTGATTAC
1481 A N I D V I D I S A Q F F G Q K N A D Y
4501 GTCGCCACTCTTGCTTACGCTCGTAGCCCGAGTCGCTGAAAAGGCTCGTTTCTCTTCTAC
1501 V A T L A Y A R S P V A E K A R F L F Y
4561 GCCGGTGCTAACAATGCCAACAAACAACAAAAACAAGGTTGCTGTTGACGCCACCGCTTAC
1521 A G A N N A N N N K N K V A V D A T A Y
4621 ATGCCCAATGTTCTCTCGTCAACGCCGCTCAAGCTTTCAACGCCGACGCCAACTCCCGC
1541 M P N V P L V N A A Q A F N A D A N S R
4681 CTCTACGTCAACGTTAAATCTGGAGAAAACCTCGACAATGGTGCTCAATTCCAATTCCAA
1561 L Y V N V K S G E N L D N G A Q F Q F Q
4741 GCTAACATGAAACAAAGCCAAGAGTTCCGTGACTACTTCCGTCAAAGCCAAATGTACAAA
1581 A N M K Q S Q E F R D Y F R Q S Q M Y K
4801 CAGTGCTCTCAACAAATGGAACAAGGAGAATATATGATGCCTGCTTGCCGTAACGCTACC
1601 Q C S Q Q M E Q G E Y M M P A C R N A T
4861 GTTGCCGCTAACAGATTAAACGAAGCCCACTTCTCCATCAACTTCGACAAGGTTTCTGAT
1621 V A A N R L N E A H F S I N F D K V S D
4921 GCCGTCAAGAACTACACCTACCAAGCTTTTCGCCTACGCCCGTCATTTAGGATACCAATAC
1641 A V K N Y T Y Q A F A Y A R H L G Y Q Y
4981 CAGTCTGAAAACCTTCGGAAACCCCAATGGCCAACACAACAAGATTGATGGTTACTTCAAG
1661 Q S E N F G N P N G Q H N K I D G Y F K
5041 TTCTCCCTAAGTTCGACTTCGCCCAGTTCTACTTCAACGCTCCTTCCGTAGCCGCTTCC
1681 F S P K F D F A Q F Y F N A P S V A A S
5101 TTCAAGAACGTTCCAGTTCACCAATACGTCGCTGACTTTTTTCGCCCCCACCCTGCTAC
1701 F K N V P V H Q Y V A D F F A P H P V Y

5161 TCTGGCTTCGACCGTCTCATGCAGGACACTTTCCAAGCCAAATACCAAGCTGCTTGCGTT
1721 S G F D R L M Q D T F Q A K Y Q A A C V
5221 GCTGACAAGATGCACGCTACCACTTTTGACAACAAAACCTTACCCCTTCATATGCAACAA
1741 A D K M H A T T F D N K T Y P L H M Q Q
5281 AACAACTGGTATGTCTCATGGCTACGTCAACAGGAACAACCTACTACAACAACCAATAC
1761 N N W Y V L M A Y V N R N N Y Y N N Q Y
5341 AACAGCTACTTTCAAGGAAACAACAAGAACCAATACAGCTACAGAGACTACAACCAAAAG
1781 N S Y F Q G N N K N Q Y S Y R D Y N Q K
5401 AGATTCTACACTACCGACTATGCTAGACAGACTAGCAACGGCCAAAAGGAATTGAAAATT
1801 R F Y T T D Y A R Q T S N G Q K E L K I
5461 GTCCTGAACAACGGAGAATACGAAATCTTCATGCAACCAGCTTCCTCTCAAGCCGGTCTC
1821 V L N N G E Y E I F M Q P A S S Q A G L
5521 CACAGCTCTAACTCCGGAAGAACAACGCTGCCATCAAGGTCTTCATCAACAAACAAGAA
1841 H S S N S G K N N A A I K V F I N K Q E
5581 CAACAATTCAACGACAAGCAATTCCTGACTTCCACGGACACAATGGAAAGATCTACGCT
1861 Q Q F N D K Q F T D F H G H N G K I Y A
5641 CAATTCTACGCTCTCCCTGATGGTGCTATCCGCTTCTTCGCTCCCCAATCTGGACTTCAA
1881 Q F Y A L P D G A I R F F A P Q S G L Q
5701 GCCATCTATGACGGAGCTCGTATCAAGATTCAAGCTGCCAACCAATACAGAGGTGCTGTC
1901 A I Y D...G...A...R I K I Q A A N Q Y R G A V
5761 CGTGGTATGTGCGGAACCTACTCCAACCAATACGCTGACACTTCACCGCCCCCTCAAACT
1921 R G...M...C...G T Y S N Q Y A D T S P P L K T
5821 GTGTCTACAAGAACCCAGAAAGACTTCGCTGCCAGCTACGCTGTCAATTGACTCTTCTTCC
1941 V S T R T Q K D F A A S Y A V I D S S S
5881 CCCTCTCAAGTGAAATCCCAGAAGGAACGCGCTCAGCAAAACCTTCTGTGCCCGCAAGAAC
1961 P S Q V K S Q K E R A Q Q N F C A R K N
5941 AACCAATTCGGAAACTACGTTAGCCGCGAGCGACGCCGGATACGGATACAAATACAACAAC
1981 N Q F G N Y V S R S D A G Y G Y K Y N N
6001 AACGACAAATACTACGAATCCGCTTACAAGAACACCAAAATACTACGACCCCTCCAACCTCC
2001 N D K Y Y E S A Y K N T K Y Y D P S N S
6061 CAGTACAACAAGTACTACAAGAACAACAAGTACGCCAAGGACCAGGACTCCTCTACTCT
2021 Q Y N K Y Y K N N K Y A K D Q D S S Y S
6121 AGCAGCAGCTCTTCATCTAGCGACTCTAGCAGCAGCTCATCCAGTGACTCCAGCAGCAGC
2041 S S S S S S S S D S S S S S S S D S S S S
6181 TCCTCTTCTTCTCTCCCCAGCTCATCCAGCTCTAGCAGCGACTCCAGCTCCTCTTCCAGC
2061 S S S S S P S S S S S S S D S S S S S S
6241 TCTTCCCAGTCCAACGACAACAACCGCAACAACAACACCGCAACAACAGCAACAACAAC
2081 S S Q S N D N N R N N N N R N N S N N N ,
6301 AACAAACCGCAACAACAACCGCAACAACAACAACAACAACAGCAGCAGCAGCCAAGAA
2101 N N R N N N R N N N N N N N S S S S Q E
6361 GCCAGTTACGAACAGAGAAACCAAAACGGACCCTCTATCCGCAAACGTACCGTGCCATC
2121 A S Y E Q R N Q N G P S I R K L Y R A I
6421 AACCAAGGAGACGACATGTGCTTCAACCATCAACGCTATCCCAACTTGACAGATACCCAGCC
2141 N Q G D D M C F T I N A I P T C R Y P A

6481 AAGCCCGTTGGAAGCGCCAAGAAAAATGGTTGACTTCTACTGTGCGCCCAAATCCAGCTCT
2161 K P V G S A K K M V D F Y C A P K S S S
6541 GAAGCTCAACACTTCAGCAAATTGATTGCCAAGGGTGCTGCTCCCAGCCAATTGTCTCTC
2181 E A Q H F S K L I A K G A A P S Q L S L
6601 AAGAAACCAAACCAAAAGTTCGAAGTGAACATTCCCGAATACTGTGTTGCCTAA
2201 K K P N Q K F E V N I P E Y C V A *