



1 GGATCTGCATCGCTCCGGTCCCGTCACTGGGCGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACCGGTGCTA

101 GAGAAGGTGGCGGGGTAACCTGGAAAGTGATGTCGTACTGGCTCCGCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

201 GTGAACGTTCTTTTTCGAACGGGTTTCCGCCAGAACACAGCTGGTGGGTAGGGATGAGGGAGGGGAGGGGCATTGTGATGTACAGGGCTGCTCTGTGAG

301 ATCAAGGGTCTCTTAAGGGTGGGAGCTGGGGCAGGGACTACGAGAGCAGCCAGATGGGCTGAAAGTGAACCTCAAGGGGTTTCTGGCACCTACCTACCTG

401 CTTCCCGCTGGGGGTGGGGAGTTGGCCAGAGTCTTAAGATTGGGGCAGGGTGGAGAGGTGGGCTCTTCCTGCTTCCACTCATCTTATAGCTTTCTTT

501 **EcoRI (511)** CCCCAGATCCGAATTCGAGATCCAAACCAAGGAGGAAAGGATATCACAGAGGAGACCATGGAGAGCAAGGTGCTGTGGCCGTCGCCCTGTGGCTCTGGC
NcoI (556) 1 MetGluSerLysValLeuLeuAlaValAlaLeuTrpLeuCysV

601 TGGAGACCCGGCCGCTCTGTGGTTCGCTAGTGTCTCTTGTATCTGCCAGGCTCAGCATACAAAAAGACATACTTACAATTAAGGCTAATACAAC

15 alGluThrArgAlaAlaSerValGlyLeuProSerValSerLeuAspLeuProArgLeuSerI leGlnLysAspI leLeuThrI leLysAlaAsnThrTh

701 TCTTCAAATTACTTGCAGGGACAGAGGGACTTGGACTGGCTTTGGCCCAATAACAGAGTGGCAGTGAGCAAAGGGTGGAGGTGACTGAGTGCAGCGAT

48 rLeuGlnI leThrCysArgGlyGlnArgAspLeuAspTrpLeuTrpProAsnGlnSerGlySerGluGlnArgValGluValThrValValI I

801 GGCCTCTTCTGAAGACACTACAATTCAAAAGTATCGGAAATGACACTGGAGCCTACAAGTGTCTTACCAGGAAACTGACTTGGCCCTCGGTCATT

82 GlyLeuPheCysLysThrLeuThrI leProLysVal I leGlyAsnAspThrGlyAlaTyrLysCysPheTyrArgGluThrAspLeuAlaSerVal I leT

901 ATGCTATGTTCAAGATTACAGATCTCCATTTATTGCTTCTGTTAGTGACCAACATGGAGTCTGTACATTACTGAGAACAAAAACAAAACCTGTGGTGTAT

115 yrValTyrValGlnAspTyrArgSerProPheI leAlaSerValSerAspGlnHisGlyValValTyrI leThrGluAsnLysAsnLysThrValVal I I

1001 TCCATGTCCTGGTCCATTTCAAATCTCAACGTGTCACTTTGTGCAAGATACCCAGAAAAGAGATTTGTTCTGTATGGTAACAGAATTTCTCGGGACAGC

148 eProCysLeuGlySerI leSerAsnLeuAsnValSerLeuCysAlaArgTyrProGluLysArgPheValProAspGlyAsnArgI leSerTrpAspSer

1101 AAGAAGGGCTTACTATTCCAGCTACATGATCAGCTATGCTGGCATGGTCTTCTGTGAAGCAAAAATAATGATGAAAGTTACCAGTCTATTATGTACA

182 LysLysGlyPheThrI leProSerTyrMetI leSerTyrAlaGlyMetValPheCysGluAlaLysI leAsnAspGluSerTyrGlnSerI leMetTyrI

1201 TAGTTGTCGTTGAGGATAGGATTTATGATGTGGTTCGAGTCGCTCATGGAAATGAACTATCTGTTGGAGAAAAGCTTGTCTTAAATTTGTACAGC

215 leValValValValGlyTyrArgI leTyrAspValValLeuSerProSerHisGlyI leGluLeuSerValGlyGluLysLeuValLeuAsnCysThrAl

1301 AAGAACTGAACATAATGTGGGATTGACTTCAACTGGGAATACCCCTTCTCGAAGCATCAGCATAAGAACTTGTAAACCGAGACCTAAAAACCCAGTCT

248 aArgThrGluLeuAsnValGlyI leAspPheAsnTrpGluTyrProSerSerLysHisGlnHisLysLysLeuValAsnArgAspLeuLysThrGlnSer

1401 GGGAGTGAGATGAAGAAATTTTGTAGCACCTAATATAGATGGTGTAAACCGGAGTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCTGATGA

282 rArgGluMetLysCysPheLeuSerThrLeuThrI leAspGlyValThrArgSerAspGlnGlyLeuTyrThrCysAlaAlaSerGlyLeuMetT

1501 CCAAGAAGAACAGCATTGTGTCAGGGTCCATGAAAAACCTTTTGTGCTTTTGGAAAGTGGCATGGAATCTCTGGTGAAGCCACGGTGGGGAGCGTGT

315 hrLysLysAsnSerThrPheValArgValHisGluLysProPheValAlaPheGlySerGlyMetGluSerLeuValGluAlaThrValGlyGluArgVa

1601 CAGAATCCCTGCGAAGTACCTTGGTTACCCACCCCGAGAAATAAAATGGTATAAAATGGAATACCCCTTGAGTCCAATCACACAATTAAGCGGGGCAT

348 IArgI leProAlaLysTyrLeuGlyTyrProProGluI leLysTrpTyrLysAsnGlyI leProLeuGluSerAsnHisThrI leLysAlaGlyHis

1701 GTACTGACGATTATGGAAGTGAAGAGACACAGGAAATFACACTGTCATCTTACCATTCCCAATTTCAAAGGAGAAGCAGAGCCATTTGGTCTCTC

382 ValLeuThrI leMetGluValSerGluArgAspThrGlyAsnTyrThrVal I leLeuThrAsnProI leSerLysGluLysGlnSerHisValValSerL

1801 TGGTTGTGTATGTCCACCCAGATTGGTGAAGAACTCTAATCTCTCTGTTGATTCTACCAGTACGGCACCCTCAAACGCTGACATGACGGTCTA

415 euValValTyrValProProGlnI leGlyGluLysSerLeuI leSerProValAspSerTyrGlnTyrGlyThrThrGlnThrLeuThrCysThrValTy

1901 TGCCATTCTCCCGCATCACATCCACTGGTATTGGCAGTTGGAGGAAGAGTGGCCAAACGAGCCAGCCAGCTGTCTCAGTGACAAACCCATACCT

448 rAlAl leProProHisHisI leHisTrpTyrTrpGlnGluCysAlaAsnGluProSerGlnAlaValThrAsnProTyrPro

2001 TGTGAAGAATGGAGAAGTGTGGAGGACTCCAGGAGGAAATAAAATTAAGTTAAATAAAATCAATTTGCTCTAATTTGAAGGAAAAACAAAACGTAA

482 CysGluGluTrpArgSerValGluAspPheGlnGlyGlyAsnLysI leGluValAsnLysAsnGlnPheAlaLeuI leGluGlyLysAsnLysThrValS

2101 GTACCCTTGTATCCAAGCGCAATGTGTGAGCTTTGTACAAATGTGAAGCGGTCAACAAAGTGGGAGAGGAGAGAGGGTGTCTCTTCCACGTGAC

515 erThrLeuVal I leGlnAlaAlaAsnValSerAlaLeuTyrLysCysGluAlaValAsnLysValGlyArgGlyGluArgVal I leSerPheHisValTh

2201 CAGGGTCTGAAATTACTTGCACCTGACAGCCACTGAGCAGGAGAGGCTTGTGTTGGTGCAGTGCAGACAGATCTACGTTTGGAGACCTC

548 rArgGlyProGluI leThrLeuGlnProAspMetGlnProThrGluGlnGluSerValSerLeuTrpCysThrAlaAspArgSerThrPheGluAsnLeu

2301 ACATGGTACAAGCTTGGCCACAGCCTCTGCCAATCCATGTGGGAGAGTGGCCACACCTGTTGCAAGAAGTGGATACTCTTGGAAATGAATGCCA

582 ThrTrpTyrLysLeuGlyProGlnProLeuI leHisValGlyGluLeuProThrProValCysLysAsnLeuAspThrLeuTrpLysLeuAsnAlaT

2401 CCATGTTCTCTAATAGCACAATGACATTTGATCATGGAGCTTAAGAATGCATCTTGCAGGACCAAGGAGACTATGTCTGCCTGTCTCAAGACAGGAA

615 hrMetPheSerAsnSerThrAsnAspI leLeuI leMetMetLeuLysAsnAlaSerGlnGlyAspTyrValCysLeuAlaGlnAspArgLy

2501 GACCAAGAAAAAGACATTTGCGTGGTCAGGCAGCTCACAGTCTTAGAGCGTGTGGCACCCACGATCACAGGAAACCTGGAGAATCAGACGACAAGTATTGG

648 sThrLysLysArgHisCysValValArgGlnLeuThrValLeuGluArgValAlaProThrI leThrGlyAsnLeuGluAsnGlnThrThrSerI leGly

2601 GAAAGCATCGAAGTCTCATGCACGGCATCTGGGAATCCCCCTCCACAGATCATGTGGTTAAAGATAATGAGACCTTGTAGAAGACTCAGGCATTGTAT

682 GluSerI leGluValSerCysThrAlaSerGlyAsnProProGlnI leMetTrpPheLysAspAsnGluThrLeuValGluAspSerGlyI leValI

2701 TGAAGGATGGGAACCGAACCTCACTATCCGAGAGTGAGGAAGGAGGACGAAGGCCTCTACACCTGCCAGGCATGCAAGTCTTGGCTGTGCAAAAGT

715 euLysAspGlyAsnArgAsnLeuThrI leArgArgValArgLysGluAspGluGlyLeuTyrThrCysGlnAlaCysSerValLeuGlyCysAlaLysVa

NheI (2832)

2801 GGAGGCATTTTTATAATAGAAGGTGCCTAAGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAACCCACAACCTAGAATGCAGTAAAAAAAT

748 I GLeuAlaPhePheI leI leGluGlyAla●●●

2901 GCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAGTTAAC

3001 AACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAGATCATTAAAT

3101 GTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAG

3201 CATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTCG

3301 TTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGT

3401 GGTCGTTGCTCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCAACCCGGTAAGA

3501 CACGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACG

3601 GCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCCGC

3701 TGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAG
 3801 TGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAAGCTGTACTGTGGAATGTGTGTCAGTTAGGGTGTGAAAAGTCCCCAGGCT
 3901 CCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCAT
 4001 GCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTA **NcoI (4088)**
 4101 ATTTTTTTTATTTATGCAGAGGCCGAGGCCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAAGC
 4201 TCCCGGGAGCTTGTATATCCATTTTCGGATCTGATcagCACGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAAGGTGAGG
 4301 **BspHI (4308)** AACTAAATCATGAAGACCTTCAACATCTCTCAGCAGGATCTGGAGCTGGTGGAGGTCGCCACTGAGAAGATCACCATGCTCTATGAGGACAACAAGCACC
 1▶ MetLysThrPheAsnI leSerGlnGlnAspLeuGluLeuValGluValAlaThrGluLysI leThrMetLeuTyrGluAspAsnLysHisH
 4401 ATGTCGGGGCGGCCATCAGGACCAAGACTGGGGAGATCATCTCTGTGTCCACATTGAGGCCTACATTGGCAGGGTCACTGTCTGTGCTGAAGCCATTGC
 31▶ isValGlyAlaAlaI leArgThrLysThrGlyGluI leI leSerAlaValHisI leGluAlaTyrI leGlyArgValThrValCysAlaGluAlaI leAl
 4501 CATTGGGTCTGCTGTGAGCAACGGGCAGAAGGACTTTGACACCATTGTGGCTGTCAGGCACCCCTACTCTGATGAGGTGGACAGATCCATCAGGGTGGTC
 64▶ aI leGlySerAlaValSerAsnGlyGlnLysAspPheAspThrI leValAlaValArgHisProTyrSerAspGluValAspArgSerI leArgValVal
 4601 AGCCCTGTGGCATGTGCAGAGAGCTCATCTCTGACTATGCTCCTGACTGCTTTGTGCTCATTGAGATGAATGGCAAGCTGGTCAAAACCACCATTGAGG
 98▶ SerProCysGlyMetCysArgGluLeuI leSerAspTyrAlaProAspCysPheValLeuI leGluMetAsnGlyLysLeuValLysThrThrI leGluG
 4701 AACTCATCCCCCTCAAGTACACCAGGAACCTAAACCTGAATTCGCTAGAGGGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCG
 131▶ luLeuI leProLeuLysTyrThrArgAsn••• **EcoRI (4737)**
 4801 ACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCTAATAAAATG
 4901 AGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCA
 5001 TGCGCAGGGCCCAATTGCTCGAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTC
 5101 TCCATCAAAACAAAACGAAAACAAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTCTCTATCGAA **NotI (5023)**