FIGURE 1

CCAAGAATTCGGCACGAGGGTGACCCGGACAGCTGTCCTCTCTGACACCACCCCGGCCTGC

CTCTTTGTTGCC**ATG**AGAGCTGCCTACCTCTTCCTGCTATTCCTGCCTGCAGGCTTGCTGG

CTCAGGGCCAGTATGACCTGGACCCGCTGCCGCCGTTCCCTGACCACGTCCAGTACACCCA

CTATAGCGACCAGATCGACAACCCAGACTACTATGATTATCAAGAGGTGACTCCTCGGCCC

TCCGAGGAACAGTTCCAGTTCCAGTCCCAGCAGCAAGTCCAACAGGAAGTCATCCCAGCCC

CAACCCCAGAACCAGGAAATGCAGAGCTGGAGCCCACAGAGCCTGGGCCTCTTGACTGCCG

TGAGGAACAGTACCCGTGCACCCGCCTCTACTCCATACACAGGCCTTGCAAACAGTGTCTC

AACGAGGTCTGCTTCTACAGCCTCCGCCGTGTGTACGTCATTAACAAGGAGATCTGTGTTC

GTACAGTGTGTGCCCACGAGGAGCTCCTCCGAGCTGACCTCTGTCGGGACAAGTTCTCCAA

ATGTGGCGTGATGGCCAGCAGCGGCCTGTGCCAATCCGTGGCGGCCTCCTGTGCCAGGAGC

TGTGGGAGCTGC**TAG**GGTGGTGCTGGCATCCTGAGTCCTGGCCCTCCTGGGATCTGGGGCC

CTCGGGCCCTGCCTGACCTGGTGCTTTTTTCCCCATCCCCATGTTCCTTTTATTCTGTAAA

AAGTTAGTGGACTGCAGCCCTGGGGGTTGCAGGCTGCGGTGCCTCAGGCCCCTCCTTCAGC

CTGTGGCCACCTCTGGGGCACGATGGGGGCTCCCCACTGCCCAGTCTGCCCCTCGGGTTGG

GGGAGTATCCCAGGCCTCTCTGTGGGACCTGGGCCCCTGACGGGCCTTCTCAGCCCGTTTT

GAGGACAGACAGTCCCCCGAGGTAGGCTACATCCCCCCACCCCAGCTGGTCTGCTTGGATT

TCCTACAGCCCCCGTGGGCATGGACCACCTTTATTTTATACAAAATTAAAAACAAGTTTTT

AC

FIGURE 2A

CACTAACGCTCTTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGT
CAAGGCTGGCTTGTGCCAGAACGGCGCGCGCGCGACGCACGCACACACACGGGGGGAAACT
TTTTTAAAAATGAAAGGCTAGAAGAGCTCAGCGGCGGCGCGGGCCGTGCGCGAGGGCTCCG
GAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGACGCCCGGCCCCGCTCGGCGCCC
GCGTGGGATGGTGCAGCGCTCGCCGCCGGGCCCGAGAGCTGCTGCACTGAAGGCCGGCGAC
GATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGT
GCTCTGCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACGAAGGAAGAGCTGATG
AAGTTGTCAGTGCCTCTGTTCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTC
CAAGAATCATCCAGAAGTGCTGAATATTCGACTACAACGGGAAAGCAAAGAACTGATCATA
AATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAG
ACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGTAATTCTGGGTCACTGTTACTACCA
TGGACATGTACGGGGATATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGG
GGACTTATTGTGTTTGAAAATGAAAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACA
GATACAAACTCTTCCCAGCGAAGAAGCTGAAAAGCGTCCGGGGATCATGTGGATCACATCA
CAACACACCAAACCTCGCTGCAAAGAATGTGTTTCCACCACCCTCTCAGACATGGGCAAGA
AGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACA
ACCGAGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAATAGAGAT
TGCTAATCACGTTGACAAGTTTTACAGACCACTGAACATTCGGATCGTGTTGGTAGGCGTG
GAAGTGTGGAATGACATGGACAAATGCTCTGTAAGTCAGGACCCATTCACCAGCCTCCATG
AATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCATGACAATGCGCAGCT
TGTCAGTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTGC
ACGGCAGACCAGTCTGGGGGAATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCG
TGACCCTGGCACATGAGCTGGGCCACAATTTCGGGATGAATCATGACACACTGGACAGGGG
CTGTAGCTGTCAAATGGCGGTTGAGAAAGGAGGCTGCATCATGAACGCTTCCACCGGGTAC
CCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTTGGAGACCAGCCTGGAGAAAG
GAATGGGGGTGTGCCTGTTTAACCTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCAGAAGTG
TGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATG
AATCGCTGCTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGC
TGTGCTGTGAAGACTGCCAGCTGAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTC
CTGTGACCTCCCAGAGTTCTGCACAGGGGCCAGCCCTCACTGCCCAGCCAACGTGTACCTG
CACGATGGGCACTCATGTCAGGATGTGGACGGCTACTGCTACAATGGCATCTGCCAGACTC
ACGAGCAGCAGTGTGTCACACTCTGGGGACCAGGTGCTAAACCTGCCCCTGGGATCTGCTT
TGAGAGAGTCAATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCC
TTTGCCAAATGCGAGATGAGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCA

GCCGGCCAGTCATTGGTACCAATGCCGTTTCCATAGAAACAAACATCCCCCTGCAGCAAGG
AGGCCGGATTCTGTGCCGGGGGACCCACGTGTACTTGGGCGATGACATGCCGGACCCAGGG
CTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCAAA
ATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAA
CAACAGGAAGAACTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGC
TTTGGAGGAAGCACAGACAGCGGCCCCATCCGGCAAGCAGATAACCAAGGTTTAACCATAG
GAATTCTGGTGACCATCCTGTGTCTTCTTGCTGCCGGATTTGTGGTTTATCTCAAAAGGAA
GACCTTGATACGACTGCTGTTTACAAATAAGAAGACCACCATTGAAAAACTAAGGTGTGTG
CGCCCTTCCCGGCCACCCCGTGGCTTCCAACCCTGTCAGGCTCACCTCGGCCACCTTGGAA
AAGGCCTGATGAGGAAGCCGCCAGATTCCTACCCACCGAAGGACAATCCCAGGAGATTGCT
GCAGTGTCAGAATGTTGACATCAGCAGACCCCTCAACGGCCTGAATGTCCCTCAGCCCCAG
TCAACTCAGCGAGTGCTTCCTCCCCTCCACCGGGCCCCACGTGCACCTAGCGTCCCTGCCA
GACCCCTGCCAGCCAAGCCTGCACTTAGGCAGGCCCAGGGGACCTGTAAGCCAAACCCCCC
TCAGAAGCCTCTGCCTGCAGATCCTCTGGCCAGAACAACTCGGCTCACTCATGCCTTGGCC
AGGACCCCAGGACAATGGGAGACTGGGCTCCG

FIGURE 2B

CCTGGCACCCCTCAGACCTGCTCCACAATATCCACACCAAGTGCCCAGATCCACCCACACC
GCCTATATTAAGTGAGAAGCCGACACCTTTTTTCAACAGTGAAGACAGAAGTTTGCACTAT
CTTTCAGCTCCAGTTGGAGTTTTTTGTACCAACTTTTAGGATTTTTTTTAATGTTTAAAAC
ATCATTACTATAAGAACTTTGAGCTACTGCCGTCAGTGCTGTGCTGTGCTATGGTGCTCTG
TCTACTTGCACAGGTACTTGTAAATTATTAATTTATGCAGAATGTTGATTACAGTGCAGTG
CGCTGTAGTAGGCATTTTTACCATCACTGAGTTTTCCATGGCAGGAAGGCTTGTTGTGCTT
TTAGTATTTTAGTGAACTTGAAATATCCTGCTTGATGGGATTCTGGACAGGATGTGTTTGC
TTTCTGATCAAGGCCTTATTGGAAAGCAGTCCCCCAACTACCCCCAGCTGTGCTTATGGTA
CCAGATGCAGCTCAAGAGATCCCAAGTAGAATCTCAGTTGATTTTCTGGATTCCCCATCTC
AGGCCAGAGCCAAGGGGCTTCAGGTCCAGGCTGTGTTTGGCTTTCAGGGAGGCCCTGTGCC
CCTTGACAACTGGCAGGCAGGCTCCCAGGGACACCTGGGAGAAATCTGGCTTCTGGCCAGG
AAGCTTTGGTGAGAACCTGGGTTGCAGACAGGAATCTTAAGGTGTAGCCACACCAGGATAG
AGACTGGAACACTAGACAAGCCAGAACTTGACCCTGAGCTGACCAGCCGTGAGCATGTTTG
GAAGGGGTCTGTAGTGTCACTCAAGGCGGTGCTTGATAGAAATGCCAAGCACTTCTTTTTC
TCGCTGTCCTTTCTAGAGCACTGCCACCAGTAGGTTATTTAGCTTGGGAAAGGTGGTGTTT
CTGTAAGAAACCTACTGCCCAGGCACTGCAAACCGCCACCTCCCTATACTGCTTGGAGCTG
AGCAAATCACCACAAACTGTAATACAATGATCCTGTATTCAGACAGATGAGGACTTTCCAT
GGGACCACAACTATTTTCAGATGTGAACCATTAACCAGATCTAGTCAATCAAGTCTGTTTA
CTGCAAGGTTCAACTTATTAACAATTAGGCAGACTCTTTATGCTTGCAAAAACTACAACCA
ATGGAATGTGATGTTCATGGGTATAGTTCATGTCTGCTATCATTATTCGTAGATATTGGAC
AAAGAACCTTCTCTATGGGGCATCCTCTTTTTCCAACTTGGCTGCAGGAATCTTTAAAAGA
TGCTTTTAACAGAGTCTGAACCTATTTCTTAAACACTTGCAACCTACCTGTTGAGCATCAC
AGAATGTGATAAGGAAATCAACTTGCTTATCAACTTCCTAAATATTATGAGATGTGGCTTG
GGCAGCATCCCCTTGAACTCTTCACTCTTCAAATGCCTGACTAGGGAGCCATGTTTCACAA
GGTCTTTAAAGTGACTAATGGCATGAGAAATACAAAAATACTCAGATAAGGTAAAATGCCA
TGATGCCTCTGTCTTCTGGACTGGTTTTCACATTAGAAGACAATTGACAACAGTTACATAA
TTCACTCTGAGTGTTTTATGAGAAAGCCTTCTTTTGGGGTCAACAGTTTTCCTATGCTTTG
AAACAGAAAAATATGTACCAAGAATCTTGGTTTGCCTTCCAGAAAACAAAACTGCATTTCA
CTTTCCCGGTGTTCCCCACTGTATCTAGGCAACATAGTATTCATGACTATGGATAAACTAA
ACACGTGACACAAACACACACAAAAGGGAACCCAGCTCTAATACATTCCAACTCGTATAGC
ATGCATCTGTTTATTCTATAGTTATTAAGTTCTTTAAAATGTAAAGCCATGCTGGAAAATA
ATACTGCTGAGATACATACAGAATTACTGTAACTGATTACACTTGGTAATTGTACTAAAGC
CAAACATATATATACTATTAAAAAGGTTTACAGAATTTTATGGTGCATTACGTGGGCATTG
TCTTTTTAGATGCCCAAATCCTTAGATCTGGCATGTTAGCCCTTCCTCCAATTATAAGAGG
ATATGAACCAAAAAAAAAAAAAAAAAA

FIGURE 3

GGAGTTTCAAAGAAGCGCTAGTAAGGTCTCTGAGATCCTTGCACTAGCTACATCCTCAGGG

TAGGAGGAAG**ATG**GCTTCCAGAAGCATGCGGCTGCTCCTATTGCTGAGCTGCCTGGCCAAA

ACAGGAGTCCTGGGTGATATCATCATGAGACCCAGCTGTGCTCCTGGATGGTTTTACCACA

AGTCCAATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTG

TCAGTCTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACC

ATAGCAGAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCCTGCACGACC

CACAGAAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTC

TGGCAAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATTAGCTCCAATAACAACTTTTTA

ACTTGGAGCAGCAACGAATGCAACAAGCGCCAACACTTCCTGTGCAAGTACCGACCA**TAG**A

GCAAGAATCAAGATTCTGCTAACTCCTGCACAGCCCCGTCCTCTTCCTTTCTGCTAGCCTG

GCTAAATCTGCTCATTATTTCAGAGGGGAAACCTAGCAAACTAAGAGTGATAAGGGCCCTA

CTACACTGGCTTTTTTAGGCTTAGAGACAGAAACTTTAGCATTGGCCCAGTAGTGGCTTCT

AGCTCTAAATGTTTGCCCCGCCATCCCTTTCCACAGTATCCTTCTTCCCTCCTCCCCTGTC

TCTGGCTGTCTCGAGCAGTCTAGAAGAGTGCATCTCCAGCCTATGAAACAGCTGGGTCTTT

GGCCATAAGAAGTAAAGATTTGAAGACAGAAGGAAGAAACTCAGGAGTAAGCTTCTAGACC

CCTTCAGCTTCTACACCCTTCTGCCCTCTCTCCATTGCCTGCACCCCACCCCAGCCACTCA

ACTCCTGCTTGTTTTTCCTTTGGCCATAGGAAGGTTTACCAGTAGAATCCTTGCTAGGTTG

ATGTGGGCCATACATTCCTTTAATAAACCATTGTGTACATAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

GGGGAGCAGAGAGGAGGCA**ATG**GCCACCATGGAGAACAAGGTGATCTGCGCCCTGGTCCTG
GTGTCCATGCTGGCCCTCGGCACCCTGGCCGAGGCCCAGACAGAGACGTGTACAGTGGCCC
CCCGTGAAAGACAGAATTGTGGTTTTCCTGGTGTCACGCCCTCCCAGTGTGCAAATAAGGG
CTGCTGTTTCGACGACACCGTTCGTGGGGTCCCCTGGTGCTTCTATCCTAATACCATCGAC
GTCCCTCCAGAAGAGGAGTGTGAATTT**TAG**ACACTTCTGCAGGGATCTGCCTGCATCCTGA
CGCGGTGCCATCCCCAGCACGGTGATTAGTCCCAGAGCTCGGCTGCCACCTCCACCGGACA
CCTCAGACACGCTTCTGCAGCTGTGCCTCGGCTCACAACACAGATTGACTGCTCTGACTTT
GACTACTCAAAATTGGCCTAAAAATTAAAAGAGCTCGATATTAAAAA

FIGURE 5


AATCCATCTGAGAATATGCTGCCACAAATACCCTTTTTGCTGCTAGTATCCTTGAACTTGGTTCATGGAGTGTTTTACGCT
GAACGATACCAAACGCCCACAGGCATAAAAGGCCCACTACCCAACACCAAGACACAGTTCTTCATTCCCTACACCATAAAGA
GTAAAGGTATAGCAGTAAGAGGAGAGCAAGGTACTCCTGGTCCACCAGGCCCTGCTGGACCTCGAGGGCACCCAGGTCCTTC
TGGACCACCAGGAAAACCAGGCTACGGAAGTCCTGGACTCCAAGGAGAGCCAGGGTTGCCAGGACCACCGGGACCATCAGCT
GTAGGGAAACCAGGTGTGCCAGGACTCCCAGGAAAACCAGGAGAGAGAGGACCATATGGACCAAAAGGAGATGTTGGACCAG
CTGGCCTACCAGGACCCCGGGGCCCACCAGGACCACCTGGAATCCCTGGACCGGCTGGAATTTCTGTGCCAGGAAAACCTGG
ACAACAGGGACCCACAGGAGCCCCAGGACCCAGGGGCTTTCCTGGAGAAAAGGGTGCACCAGGAGTCCCTGGTATGAATGGA
CAGAAAGGGGAAATGGGATATGGTGCTCCTGGTCGTCCAGGTGAGAGGGGTCTTCCAGGCCCTCAGGGTCCCACAGGACCAT
CTGGCCCTCCTGGAGTGGGAAAAAGAGGTGAAAATGGGGTTCCAGGACAGCCAGGCATCAAAGGTGATAGAGGTTTTCCGGG
AGAAATGGGACCAATTGGCCCACCAGGTCCCCAAGGCCCTCCTGGGGAACGAGGGCCAGAAGGCATTGGAAAGCCAGGAGCT
GCTGGAGCCCCAGGCCAGCCAGGGATTCCAGGAACAAAAGGTCTCCCTGGGGCTCCAGGAATAGCTGGGCCCCCAGGGCCTC
CTGGCTTTGGGAAACCAGGCTTGCCAGGCCTGAAGGGAGAAAGAGGACCTGCTGGCCTTCCTGGGGGTCCAGGTGCCAAAGG
GGAACAAGGGCCAGCAGGTCTTCCTGGGAAGCCAGGTCTGACTGGACCCCCTGGGAATATGGGACCCCAAGGACCAAAAGGC
ATCCCGGGTAGCCATGGTCTCCCAGGCCCTAAAGGTGAGACAGGGCCAGCTGGGCCTGCAGGATACCCTGGGGCTAAGGGTG
AAAGGGGTTCCCCTGGGTCAGATGGAAAACCAGGGTACCCAGGAAAACCAGGTCTCGATGGTCCTAAGGGTAACCCAGGGTT
ACCAGGTCCAAAAGGTGATCCTGGAGTTGGAGGACCTCCTGGTCTCCCAGGCCCTGTGGGCCCAGCAGGAGCAAAGGGAATG
CCCGGACACAATGGAGAGGCTGGCCCAAGAGGTGCCCCTGGAATACCAGGTACTAGAGGCCCTATTGGGCCACCAGGCATTC
CAGGATTCCCTGGGTCTAAAGGGGATCCAGGAAGTCCCGGTCCTCCTGGCCCAGCTGGCATAGCAACTAAGGGCCTCAATGG
ACCCACCGGGCCACCAGGGCCTCCAGGTCCAAGAGGCCCCTCTGGAGAGCCTGGTCTTCCAGGGCCCCCTGGGCCTCCAGGC
CCACCAGGTCAAGCAGTCATGCCTGAGGGTTTTATAAAGGCAGGCCAAAGGCCCAGTCTTTCTGGGACCCCTCTTGTTAGTG
CCAACCAGGGGGTAACAGGAATGCCTGTGTCTGCTTTTACTGTTATTCTCTCCAAAGCTTACCCAGCAATAGGAACTCCCAT
ACCATTTGATAAAATTTTGTATAACAGGCAACAGCATTATGACCCAAGGACTGGAATCTTTACTTGTCAGATACCAGGAATA
TACTATTTTTCATACCACGTGCATGTGAAAGGGACTCATGTTTGGGTAGGCCTGTATAAGAATGGCACCCCTGTAATGTACA
CCTATGATGAATACACCAAAGGCTACCTGGATCAGGCTTCAGGGAGTGCCATCATCGATCTCACAGAAAATGACCAGGTGTG
GCTCCAGCTTCCCAATGCCGAGTCAAATGGCCTATACTCCTCTGAGTATGTCCACTCCTCTTTCTCAGGATTCCTAGTGGCT
CCAATGTGAGTACACCCCACAGAGCTAATCTAAATCTTGTGCTAGAAAAAGCATTCTCTAACTCTACCCCACCCTACAAAAT
GCATATGGAGGTAGGCTGAAAAGAATGTAATTTTTATTTTCTGAAATACAGATTTGAGCTATCAGACCAACAAACCTTCCCC
CTGAAAAGTGAGCAGCAACGTAAAAACGTATGTGAAGCCTCTCTTGAATTTCTAGTTAGCAATCTTAAGGCTCTTTAAGGTT
TTCTCCAATATTAAAAAATATCACCAAAGAAGTCCTGCTATGTTAAAAACAAACAACAAAAAACAAAGCAACAAAAAAAAAA
ATTAAAAAAAAAAAACAGAAATAGAGCTCTAAGTTATGTGAAATTTGATTTGAGAAACTCGGCATTTCCTTTTTAAAAAAGCC
TGTTTCTAACTATGAATATGAGAACTTCTAGGAAACATCCAGGAGGTATCATATAACTTTGTAGAACTTAAATACTTGAATA
TTCAAATTTAAAAGACACTGTATCCCCTAAAATATTTCTGATGGTGCACTACTCTGAGGCCTGTATGGCCCCTTTCATCAAT
ATCTATTCAAATATACAGGTGCATATATACTTGTTAAAGCTCTTATATAAAAAAGCCCCAAAATATTGAAGTTCATCTGAAA
TGCAAGGTGCTTTCATCAATGAACCTTTTCAAAACTTTTCTATGATTGCAGAGAAGCTTTTTATATACCCAGCATAACTTGG
AAACAGGTATCTGACCTATTCTTATTTAGTTAACACAAGTGTGATTAATTTGATTTCTTTAATTCCTTATTGAATCTTATGT
GATATGATTTTCTGGATTTACAGAACATTAGCACATGTACCTTGTGCCTCCCATTCAAGTGAAGTTATAATTTACACTGAGG
GTTTCAAAATTCGACTAGAAGTGGAGATATATTATTTATTTATGCACTGTACTGTATTTTTATATTGCTGTTTAAAACTTTT
AAGCTGTGCCTCACTTATTAAAGCACAAAATGTTTTACCTACTCCTTATTTACGACACAATAAAATAACATCAATAGATTTT
TAGGCTGAATTAATTTGAAAGCAGCAATTTGCTGTTCTCAACCATTCTTTCAAGGCTTTTCATTCGACACAATAAAATAACA
TCAATAGATTTTTAGGGATGGGTGGCTTT

FIGURE 6

CATGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCC
TGTGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA
ATGTCCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCA
GTATTTCATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGA
ACCTACTGTGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTA
AGGCCATTTGGGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTT
AGAAACACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTT
GTCCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAAC
AAATATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTG
GTCCCAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTAC
TGCGTACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGC
AGTGTGCCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTA
TGTTTTGCCCATATCTATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGA
TATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAAT
TTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTAACTTTATCACCCTCAATAT
CTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTA
TCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGG
TGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACAC
GGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACA
GTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGG
AGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTT
GGCAGTCTTGGGCCCGCAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGAC
CCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCC
TGGTCGACTGGGATCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCA
GGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCT
AGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGC
AATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAACTGATGCCAACACTTCCTTTT
GCCTTTTGTTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTG
GGATGAAAGAAGTTTTTTCCAGTTTGTCAGTGTCTGTGAGAA

FIGURE 7

AGCGGGCGTGGGACTGAGCAGTCTGCTGCCCCCCGACATGTGACCCAGCCCCGCCGCCCATG
CGGGCTCCCGGCCGCCCGGCCCTGCGGCCGCTGCCGCTGCCGCCGCTGCTGCTGTTGCTC
CTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTCTCTGGTGGTTTGCCTAAACCTGCAA
ACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACAATGGACTCCACCAGAGGGTCT
TCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATATATGGGCAAAAGAAATGGCTG
AATAAATCAGAATGCAGAAATATCAATAGAACCTACTGTGATCTTTCTGCTGAAACTTCTG
ACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTGGGGAACAAAGTGTTCCAAATG
GGCTGAAAGTGGACGGTTCTATCCTTTTTTAGAAACACAAATTGGCCCACCAGAGGTGGCA
CTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCCAGAGAAGTGGAAGAGAA
ATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGAAGTATAACGTGTC
TGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCACACGCTGGTG
CTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCCAGGGC
CCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCATC
AGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTT
TTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAG
CAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAA
AATCGTGATTAACTTTATCACCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGAT
ATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGA
ACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGAT
GGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCC
CTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCA
CTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACA
AGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCGCAAACGTTACAGTAC
TCATACACCCCTCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGG
AGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAAACTGGCAGGCT
GTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGG
GATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGC
CACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCA
GATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAACAAGTGAGTC
ACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTCAGT
GTCTGTGAGAA

FIGURE 8

GCTTGGTCTAGTGCCCATTTACTCTGGACTCCGG**ATG**GCTGCCGCACGCCTCTGCCTCTCC
CTGCTGCTCCTGTCCACCTGCGTGGCTCTGTTACTACAGCCACTGCTGGGTGCCCAGGGAG
CCCCACTGGAGCCAGTGTACCCAGGGGACAATGCCACACCAGAGCAGATGGCCCAGTATGC
AGCTGATCTCCGTAGATACATCAACATGCTGACCAGGCCTAGGTATGGGAAAAGACACAAA
GAGGACACGCTGGCCTTCTCGGAGTGGGGGTCCCCGCATGCTGCTGTCCCCAGGGAGCTCA
GCCCGCTGGACTTA**TAA**TGCCACCTTCTGTCTCCTACGACTCCATGAGCAGCGCCAGCCCA
GCTCTCCCCTCTGCACCCTTGGCTCTGGCCAAAGCTTGCTCCCTGCTCCCACACAGGCTCA
ATAAAGCAAGTCAAAGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 9

GAAGGACAGCAGGGCCAACAGTCACAGCAGCCCTGACCAGAGCATTCCTGGAGCTCAAGCT
CCTCTACAAAGAGGTGGACAGAGAAGACAGCAGAGACCATGGGACCCCCCTCAGCCCCTCC
CTGCAGATTGCATGTCCCCTGGAAGGAGGTCCTGCTCACAGCCTCACTTCTAACCTTCTGG
AACCCACCCACCACTGCCAAGCTCACTATTGAATCCACGCCGTTCAATGTCGCAGAGGGGA
AGGAGGTTCTTCTACTCGCCCACAACCTGCCCCAGAATCGTATTGGTTACAGCTGGTACAA
AGGCGAAAGAGTGGATGGCAACAGTCTAATTGTAGGATATGTAATAGGAACTCAACAAGCT
ACCCCAGGGCCCGCATACAGTGGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCC
AGAACGTCACCCAGAATGACACAGGATTCTATACCCTACAAGTCATAAAGTCAGATCTTGT
GAATGAAGAAGCAACCGGACAGTTCCATGTATACCCGGAGCTGCCCAAGCCCTCCATCTCC
AGCAACAACTCCAACCCCGTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGG
TTCAGAACACAACCTACCTGTGGTGGGTAAATGGTCAGAGCCTCCCGGTCAGTCCCAGGCT
GCAGCTGTCCAATGGCAACATGACCCTCACTCTACTCAGCGTCAAAAGGAACGATGCAGGA
TCCTATGAATGTGAAATACAGAACCCAGCGAGTGCCAACCGCAGTGACCCAGTCACCCTGA
ATGTCCTCTATGGCCCAGATGGCCCCACCATTTCCCCCTCAAAGGCCAATTACCGTCCAGG
GGAAAATCTGAACCTCTCCTGCCACGCAGCCTCTAACCCACCTGCACAGTACTCTTGGTTT
ATCAATGGGACGTTCCAGCAATCCACACAAGAGCTCTTTATCCCCAACATCACTGTGAATA
ATAGCGGATCCTATATGTGCCAAGCCCATAACTCAGCCACTGGCCTCAATAGGACCACAGT
CACGATGATCACAGTCTCTGGAAGTGCTCCTGTCCTCTCAGCTGTGGCCACCGTCGGCATC
ACGATTGGAGTGCTGGCCAGGGTGGCTCTGATATAGCAGCCCTGGTGTATTTTCGATATTT
CAGGAAGACTGGCAGATTGGACCAGACCCTGAATTCTTCTAGCTCCTCCAATCCCATTTTA
TCCCATGGAACCACTAAAAACAAGGTCTGCTCTGCTCCTGAAGCCCTATATGCTGGAGATG
GACAACTCAATGAAAATTTAAAGGGAAAACCCTCAGGCCTGAGGTGTGTGCCACTCAGAGA
CTTCACCTAACTAGAGACAGGCAAACTGCAAACCATGGTGAGAAATTGACGACTTCACACT
ATGGACAGCTTTTCCCAAGATGTCAAAACAAGACTCCTCATCATGATAAGGCTCTTACCCC
CTTTTAATTTGTCCTTGCTTATGCCTGCCTCTTTCGCTTGGCAGGATGATGCTGTCATTAG
TATTTCACAAGAAGTAGCTTCAGAGGGTAACTTAACAGAGTATCAGATCTATCTTGTCAAT
CCCAACGTTTTACATAAAATAAGAGATCCTTTAGTGCACCCAGTGACTGACATTAGCAGCA
TCTTTAACACAGCCGTGTGTTCAAATGTACAGTGGTCCTTTTCAGAGTTGGACTTCTAGAC
TCACCTGTTCTCACTCCCTGTTTTAATTCAACCCAGCCATGCAATGCCAAATAATAGAATT
GCTCCCTACCAGCTGAACAGGGAGGAGTCTGTGCAGTTTCTGACACTTGTTGTTGAACATG
GCTAAATACAATGGGTATCGCTGAGACTAAGTTGTAGAAATTAACAAATGTGCTGCTTGGT
TAAAATGGCTACACTCATCTGACTCATTCTTTATTCTATTTTAGTTGGNTTGTATCTTGCC
TAAGGTGCGTAGTCCAACTCTTGGTATTACCCTCCTAATAGTCATACTAGTAGTCATACTC
CCTGGTGTAGTGTATTCTCTAAAAGCTTTAAATGTCTGCATGCAGCCAGCCATCAAATAGT
GAATGGTCTCTCTTTGGCTGGAATTACAAAACTCAGAGAAATGTGTCATCAGGAGAACATC

ATAACCCATGAAGGATAAAAGCCCCAAATGGTGGTAACTGATAATAGCACTAATGCTTTAA
GATTTGGTCACACTCTCACCTAGGTGAGCGCATTGAGCCAGTGGTGCATAAATGCTACATA
CTCCAACTGAAATGTTAAGGAAGAAGATAGATCCAATTAAAAAAAATTAAAACCAATTTAA
AAAAAAAAAGAACACAGGAGATTCCAGTCTACTTGAGTTAGCATAATACAGAAGTCCCCTC
TACTTTAACTTTTACAAAAAAGTAACCTGAACTAATCTGATGTTAACCAATGTATTTATTT
CTGTGGTTCTGTTTCCTTGTTCCAATTTGACAAAACCCACTGTTCTTGTATTGTATTGCCC
AGGGGGAGCTATCACTGTACTTGTAGAGTGGTGCTGCTTTAATTCATAAATCACAAATAAA
AGCCAATTAGCTCTAAAAAAAAAAAA

FIGURE 10

GACCTGGGGGGGCGAGGGCCCCCCAAACTCAGTTCGGATCCTACCCGAGTGAGGCGGCGCC

ATGGAGCTCCGGGTGCTGCTCTGCTGGGCTTCGTTGGCCGCAGCTTTGGAAGAGACCCTGC

TGAACACAAAATTGGAAACTGCTGATCTGAAGTGGGTGACATTCCCTCAGGTGGACGGGCA

GTGGGAGGAACTGAGCGGCCTGGATGAGGAACAGCACAGCGTGCGCACCTACGAAGTGTGT

GACGTGCAGCGTGCCCCGGGCCAGGCCCACTGGCTTCGCACAGGTTGGGTCCCACGGCGGG

GCGCCGTCCACGTGTACGCCACGCTGCGCTTCACCATGCTCGAGTGCCTGTCCCTGCCTCG

GGCTGGGCGCTCCTGCAAGGAGACCTTCACCGTCTTCTACTATGAGAGCGATGCGGACACG

GCCACGGCCCTCACGCCAGCCTGGATGGAGAACCCCTACATCAAGGTGGACACGGTGGCCG

CGGAGCATCTCACCCGGAAGCGCCCTGGGGCCGAGGCCACCGGGAAGGTGAATGTCAAGAC

GCTGCGTCTGGGACCGCTCAGCAAGGCTGGCTTCTACCTGGCCTTCCAGGACCAGGGTGCC

TGCATGGCCCTGCTATCCCTGCACCTCTTCTACAAAAAGTGCGCCCAGCTGACTGTGAACC

TGACTCGATTCCCGGAGACTGTGCCTCGGGAGCTGGTTGTGCCCGTGGCCGGTAGCTGCGT

GGTGGATGCCGTCCCCGCCCCTGGCCCCAGCCCCAGCCTCTACTGCCGTGAGGATGGCCAG

TGGGCCGAACAGCCGGTCACGGGCTGCAGCTGTGCTCCGGGGTTCGAGGCAGCTGAGGGGA

ACACCAAGTGCCGAGCCTGTGCCCAGGGCACCTTCAAGCCCCTGTCAGGAGAAGGGTCCTG

CCAGCCATGCCCAGCCAATAGCCACTCTAACACCATTGGATCAGCCGTCTGCCAGTGCCGC

GTCGGGTACTTCCGGGCACGCACAGACCCCCGGGGTGCACCCTGCACCACCCCTCCTTCGG

CTCCGCGGAGCGTGGTTTCCCGCCTGAACGGCTCCTCCCTGCACCTGGAATGGAGTGCCCC

CCTGGAGTCTGGTGGCCGAGAGGACCTCACCTACGCCCTCCGCTGCCGGGAGTGCCGACCC

GGAGGCTCCTGTGCGCCCTGCGGGGGAGACCTGACTTTTGACCCCGGCCCCCGGGACCTGG

TGGAGCCCTGGGTGGTGGTTCGAGGGCTACGTCCTGACTTCACCTATACCTTTGAGGTCAC

TGCATTGAACGGGGTATCCTCCTTAGCCACGGGGCCCGTCCCATTTGAGCCTGTCAATGTC

ACCACTGACCGAGAGGTACCTCCTGCAGTGTCTGACATCCGGGTGACGCGGTCCTCACCCA

GCAGCTTGAGCCTGGCCTGGGCTGTTCCCCGGGCACCCAGTGGGGCTGTGCTGGACTACGA

GGTCAAATACCATGAGAAGGGCGCCGAGGGTCCCAGCAGCGTGCGGTTCCTGAAGACGTCA

GAAAACCGGGCAGAGCTGCGGGGGCTGAAGCGGGGAGCCAGCTACCTGGTGCAGGTACGGG

CGCGCTCTGAGGCCGGCTACGGGCCCTTCGGCCAGGAACATCACAGCCAGACCCAACTGGA

TGAGAGCGAGGGCTGGCGGGAGCAGCTGGCCCTGATTGCGGGCACGGCAGTCGTGGGTGTG

GTCCTGGTCCTGGTGGTCATTGTGGTCGCAGTTCTCTGCCTCAGGAAGCAGAGCAATGGGA

GAGAAGCAGAATATTCGGACAAACACGGACAGTATCTCATCGGACATGGTACTAAGGTCTA

CATCGACCCCTTCACTTATGAAGACCCTAATGAGGCTGTGAGGGAATTTGCAAAAGAGATC

GATGTCTCCTACGTCAAGATTGAAGAGGTGATTGGTGCAGGTGAGTTTGGCGAGGTGTGTC

GGGGGCGGCTCAAGGCCCCAGGGAAGAAGGAGAGCTGTGTGGCAATCAAGACCCTGAAGGG

TGGCTACACGGAGCGGCAGCGGCGTGAGTTTCTGAGCGAGGCCTCCATCATGGGCCAGTTC

GAGCACCCCAATATCATCCGCCTGGAGGGCGTGGTCACCAACAGCATGCCCGTCATGATTC

TCACAGAGTTCATGGAGAACGGCGCCCTGGACTCCTTCCTGCGGCTAAACGACGGACAGTT
CACAGTCATCCAGCTCGTGGGCATGCTGCGGGGCATCGCCTCGGGCATGCGGTACCTTGCC
GAGATGAGCTACGTCCACCGAGACCTGGCTGCTCGCAACATCCTAGTCAACAGCAACCTCG
TCTGCAAAGTGTCTGACTTTGGCCTTTCCCGATTCCTGGAGGAGAACTCTTCCGATCCCAC
CTACACGAGCTCCCTGGGAGGAAAGATTCCCATCCGATGGACTGCCCCGGAGGCCATTGCC
TTCCGGAAGTTCACTTCCGCCAGTGATGCCTGGAGTTACGGGATTGTGATGTGGGAGGTGA
TGTCATTTGGGGAGAGGCCGTACTGGGACATGAGCAATCAGGACGTGATCAATGCCATTGA
ACAGGACTACCGGCTGCCCCCGCCCCCAGACTGTCCCACCTCCCTCCACCAGCTCATGCTG
GACTGTTGGCAGAAAGACCGGAATGCCCGGCCCCGCTTCCCCCAGGTGGTCAGCGCCCTGG
ACAAGATGATCCGGAACCCCGCCAGCCTCAAAATCGTGGCCCGGGAGAATGGCGGGGCCTC
ACACCCTCTCCTGGACCAGCGGCAGCCTCACTACTCAGCTTTTGGCTCTGTGGGCGAGTGG
CTTCGGGCCATCAAAATGGGAAGATACGAAGAAAGTTTCGCAGCCGCTGGCTTTGGCTCCT
TCGAGCTGGTCAGCCAGATCTCTGCTGAGGACCTGCTCCGAATCGGAGTCACTCTGGCGGG
ACACCAGAAGAAAATCTTGGCCAGTGTCCAGCACATGAAGTCCCAGGCCAAGCCGGGAACC
CCGGGTGGGACAGGAGGACCGGCCCCGCAGTAC**TGA**CCTGCAGGAACTCCCCACCCCAGGG
ACACCGCCTCCCCATTTTCCGGGGCAGAGTGGGGACTCACAGAGGCCCCCAGCCCTGTGCC
CCGCTGGATTGCACTTTGAGCCCGTGGGGTGAGGAGTTGGCAATTTGGAGAGACAGGATTT
GGGGGTTCTGCCATAATAGGAGGGGAAAATCACCCCCCAGCCACCTCGGGGAACTCCAGAC
CAAGGGTGAGGGCGCCTTTCCCTCAGGACTGGGTGTGACCAGAGGAAAAGGAAGTGCCCAA
CATCTCCCAGCCTCCCCAGGTGCCCCCCTCACCTTGATGGGTGCGTTCCCGCAGACCAAAG
AGAGTGTGACTCCCTTGCCAGCTCCAGAGTGGGGGGGCTGTCCCAGGGGGCAAGAAGGGGT
GTCAGGGCCCAGTGACAAAATCATTGGGGTTTGTAGTCCCAACTTGCTGCTGTCACCACCA
AACTCAATCATTTTTTTCCCTTGTAAATGCCCCTCCCCCAGCTGCTGCCTTCATATTGAAG
GTTTTTGAGTTTTGTTTTTGGTCTTAATTTTTCTCCCCGTTCCCTTTTTGTTTCTTCGTTT
TGTTTTTCTACCGTCCTTGTCATAACTTTGTGTTGGAGGGAACCTGTTTCACTATGGCCTC
CTTTGCCCAAGTTGAAACAGGGGCCCATCATCATGTCTGTTTCCAGAACAGTGCCTTGGTC
ATCCCACATCCCCGGACCCCGCCTGGGACCCCCAAGCTGTGTCCTATGAAGGGGTGTGGGG
TGAGGTAGTGAAAAGGGCGGTAGTTGGTGGTGGAACCCAGAAACGGACGCCGGTGCTTGGA
GGGGTTCTTAAATTATATTTAAAAAAGTAACTTTTTGTATAAATAAAAGAAAATGGGACGT
GTCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 11

GCTCTATGCCGCCTACCTTGCTCTCGCCGCTGCTGCCGGAGCCGAAGCAGAGAAGGCAGCG
GGTCCCGTGACCGTCCCGAGAGCCCCGCGCTCCCGACCAGGGGGCGGGGGCGGCCCCGGGG
AGGGCGGGGCAGGGGCGGGGGGAAGAAAGGGGGTTTTGTGCTGCGCCGGGAGGGCCGGCGC
CCTCTTCCGAATGTCCTGCGGCCCCAGCCTCTCCTCACGCTCGCGCAGTCTCCGCCGCAGT
CTCAGCTGCAGCTGCAGGACTGAGCCGTGCACCCGGAGGAGACCCCCGGAGGAGGCGACAA
ACTTCGCAGTGCCGCGACCCAACCCCAGCCCTGGGTAGCCTGCAGC**ATG**GCCCAGCTGTTC
CTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTGG
AAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGCA
GGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCCG
AGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGGG
AGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCCG
CGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCTG
CGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGCG
ACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCTA
TGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCACC
CCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTGT
CGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGGA
TGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTAC
TGTTATGCTGAAGACCTAAATGGAGAACTGTTCCTGGGTGACCCTCCAGAGAAGCTGACAT
TGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAACT
GTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAGT
GTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAGA
CTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCTA
CTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAAC
CCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAGC
TGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGGA
GGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCTA
GAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCAT
TGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGGA
GGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGCC
TCTCTCCCCACTGAGCCAGCAGCCCAGGAGAAGTCACTCTCCCAGGCGCCAGCAAGGGCAG
TCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAAG
GGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCCA
TCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTAT
CTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGCT

TCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTCT
GGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCCACTGACAGCG
CCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTGACTGTGTCCCCAGCCCCTGCCA
CAATGGTGGGACATGCTTGGAGGAGGAGGAAGGGGTCCGCTGCCTATGTCTGCCTGGCTAT
GGGGGGGACCTGTGCGATGTTGGCCTCCGCTTCTGCAACCCCGGCTGGGACGCCTTCCAGG
GCGCCTGCTACAAGCACTTTTCCACACGAAGGAGCTGGGAGGAGGCAGAGACCCAGTGCCG
GATGTACGGCGCGCATCTGGCCAGCATCAGCACACCCGAGGAACAGGACTTCATCAACAAC
CGGTACCGGGAGTACCAGTGGATCGGACTCAACGACAGGACCATCGAAGGCGACTTCTTGT
GGTCGGATGGCGTCCCCCTGCTCTATGAGAACTGGAACCCTGGGCAGCCTGACAGCTACTT
CCTGTCTGGAGAGAACTGCGTGGTCATGGTGTGGCATGATCAGGGACAATGGAGTGACGTG
CCCTGCAACTACCACCTGTCCTACACCTGCAAGATGGGGCTGGTGTCCTGTGGGCCGCCAC
CGGAGCTGCCCCTGGCTCAAGTGTTCGGCCGCCCACGGCTGCGCTATGAGGTGGACACTGT
GCTTCGCTACCGGTGCCGGGAAGGACTGGCCCAGCGCAATCTGCCGCTGATCCGATGCCAA
GAGAACGGTCGTTGGGAGGCCCCCCAGATCTCCTGTGTGCCCAGAAGACCTGCCCGAGCTC
TGCACCCAGAGGAGGACCCAGAAGGACGTCAGGGGAGGCTACTGGGACGCTGGAAGGCGCT
GTTGATCCCCCCTTCCAGCCCCATGCCAGGTCCC**TAG**GGGGCAAGGCCTTGAACACTGCCG
GCCACAGCACTGCCCTGTCACCCAAATTTTCCCTCACACCTTGCGCTCCCGCCACCACAGG
AAGTGACAACATGACGAGGGGTGGTGCTGGAGTCCAGGTGACAGTTCCTGAAGGGGCTTCT
GGGAAATACCTAGGAGGCTCCAGCCCAGCCCAGGCCCTCTCCCCCTACCCTGGGCACCAGA
TCTTCCATCAGGGCCGGAGTAAATCCCTAAGTGCCTCAACTGCCCTCTCCCTGGCAGCCAT
CTTGTCCCCTCTATTCCTCTAGGGAGCACTGTGCCCACTCTTTCTGGGTTTTCCAAGGGAA
TGGGCTTGCAGGATGGAGTGTCTGTAAAATCAACAGGAAATAAAACTGTGTATGAGCCCA

FIGURE 12

TGTGGCACTGCCTGCGTACCCAACCCCAGCCCTGGGTAGCCTGCAGC**ATG**GCCCAGCTGTT
CCTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTG
GAAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGC
AGGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCC
GAGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGG
GAGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCC
GCGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCT
GCGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGC
GACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCT
ATGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCAC
CCCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTG
TCGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGG
ATGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTA
CTGTTATGCTGAAGACCTAAATGGAGAATTGTTCCTGGGTGACCCTCCAGAGAAGCTGACA
TTGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAAC
TGTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAG
TGTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAG
ACTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCT
ACTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAA
CCCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAG
CTGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGG
AGGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCT
AGAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCA
TTGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGG
AGGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGC
CTCTCTCCCCACTGAGCCAGCAGCCCAGGAGGAGTCACTCTCCCAGGCGCCAGCAAGGGCA
GTCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAA
GGGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCC
ATCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTA
TCTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGC
TTCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTC
TGGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCCACTGACAGC
GCCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTAATTCTGCCCAAGGCTCAACTG
CCCTCTCTATCCTACTCCTTTTCTTCCCCCTGCAGCTCTGGGTCACC**TGA**CCTGTAGTCCT
TTAACCCACCATCATCCCAAACTCTCCTGTCCTTTGCCTTCATTCTCTTACCCACCTCTAC

CTATGGGTCTCCAATCTCGGATATCCACCTTGTGGGTATCTCAGCTCTCCGCGTCTTTACC
CTGTGATCCCAGCCCCGCCACTGACCATCTGTGACCCTTCCCTGCCATTGGGCCCTCCACC
TGTGGCTCACATCTCGCCAGCCCCACAGAGCATCCTCAGGCCTCTCCAAGGGTCCTCATCA
CCTATTGCAGCCTTCAGGGCTCGGCCTATTTTCCACTACTCCCTTCATCCGCCTGTGTGCC
GTCCCCTTTAGCTGCCTCCTATTGATCTCAGGGAAGCCTGGGAGTCCCTTCTCACCCCTCA
ACCTCCGGAGTCCAGGAGAACCCGTACCCCCACAGAGCCTTAAGCAACTACTTCTGTGAAG
TATTTTTTGACTGTTTCATGGAAAACAAGCCTTGGAAATAAATCTCTATTAAACCGC

FIGURE 13

AGCCAGCTCAGGCTACACTATCCCAGGATCAGC**ATG**GCCGTCCGCCAGTGGGTAATCGCCC
TGGCCTTGGCTGCCCTCCTTGTTGTGGACAGGGAAGTGCCAGTGGCAGCAGGAAAGCTCCC
TTTCTCAAGAATGCCCATCTGTGAACACATGGTAGAGTCTCCAACCTGTTCCCAGATGTCC
AACCTGGTCTGCGGCACTGATGGGCTCACATATACGAATGAATGCCAGCTCTGCTTGGCCC
GGATAAAAACCAAACAGGACATCCAGATCATGAAAGATGGCAAATGC**TGA**TCCCACAGGAG
CACCTCAAGCCATGAAGTGTCAGCTGGAGAACAGTGGTGGGCATGGAGAGGATATGACATG
AAATAAAAGATCCAGCCCAAAAAAAAAAGAAAAAA

FIGURE 14

CGTTTCCAATGCACGTACAGCCCGTACACACCGTGTGCTGGGACACCCCACAGTCAGCCGC

ATGGCTCCCCTGTGCCCCAGCCCCTGGCTCCCTCTGTTGATCCCGGCCCCTGCTCCAGGCC

TCACTGTGCAACTGCTGCTGTCACTGCTGCTTCTGGTGCCTGTCCATCCCCAGAGGTTGCC

CCGGATGCAGGAGGATTCCCCCCTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTGGGC

GAGGAGGATCTGCCCAGTGAAGAGGATTCACCCAGAGAGGAGGATCCACCCGGAGAGGAGG

ATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCTAAATC

AGAAGAAGAGGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCCTGGAGATCCT

CAAGAACCCCAGAATAATGCCCACAGGGACAAAGAAGGGGATGACCAGAGTCATTGGCGCT

ATGGAGGCGACCCGCCCTGGCCCCGGGTGTCCCCAGCCTGCGCGGGCCGCTTCCAGTCCCC

GGTGGATATCCGCCCCCAGCTCGCCGCCTTCTGCCCGGCCCTGCGCCCCCTGGAACTCCTG

GGCTTCCAGCTCCCGCCGCTCCCAGAACTGCGCCTGCGCAACAATGGCCACAGTGTGCAAC

TGACCCTGCCTCCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTACCGGGCTCTGCA

GCTGCATCTGCACTGGGGGGCTGCAGGTCGTCCGGGCTCGGAGCACACTGTGGAAGGCCAC

CGTTTCCCTGCCGAGATCCACGTGGTTCACCTCAGCACCGCCTTTGCCAGAGTTGACGAGG

CCTTGGGGCGCCCGGGAGGCCTGGCCGTGTTGGCCGCCTTTCTGGAGGAGGGCCCGGAAGA

AAACAGTGCCTATGAGCAGTTGCTGTCTCGCTTGGAAGAAATCGCTGAGGAAGGCTCAGAG

ACTCAGGTCCCAGGACTGGACATATCTGCACTCCTGCCCTCTGACTTCAGCCGCTACTTCC

GATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCAGGGTGTCATCTGGACTGTGTTTAA

CCAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCCTCTCTGACACCCTGTGGGGACCT

GGTGACTCTCGGCTACAGCTGAACTTCCGAGCGACGCAGCCTTTGAATGGGCGAGTGATTG

AGGCCTCCTTCCCTGCTGGAGTGGACAGCAGTCCTCGGGCTGCTGAGCCAGTCCAGCTGAA

TTCCTGCCTGGCTGCTGGTGACATCCTAGCCCTGGTTTTTGGCCTCCTTTTTGCTGTCACC

AGCGTCGCGTTCCTTGTGCAGATGAGAAGGCAGCACAGAAGGGGAACCAAAGGGGGTGTGA

GCTACCGCCCAGCAGAGGTAGCCGAGACTGGAGCC**TAG**AGGCTGGATCTTGGAGAATGTGA

GAAGCCAGCCAGAGGCATCTGAGGGGGAGCCGGTAACTGTCCTGTCCTGCTCATTATGCCA

CTTCCTTTTAACTGCCAAGAAATTTTTTAAAATAAATATTTATAATA

FIGURE 15

TCCACATCCAGGTCTTTGTGCTCCTCGCTTGCCTGTTCCTTTTCCACGCATTTTCCAGGAT
AACTGTGACTCCAGGCCCGCA**ATG**GATGCCCTGCAACTAGCAAATTCGGCTTTTGCCGTTG
ATCTGTTCAAACAACTATGTGAAAAGGAGCCACTGGGCAATGTCCTCTTCTCTCCAATCTG
TCTCTCCACCTCTCTGTCACTTGCTCAAGTGGGTGCTAAAGGTGACACTGCAAATGAAATT
GGACAGGTTCTTCATTTTGAAAATGTCAAAGATGTACCCTTTGGATTTCAAACAGTAACAT
CGGATGTAAACAAACTTAGTTCCTTTTACTCACTGAAACTAATCAAGCGGCTCTACGTAGA
CAAATCTCTGAATCTTTCTACAGAGTTCATCAGCTCTACGAAGAGACCGTATGCAAAGGAA
TTGGAAACTGTTGACTTCAAAGATAAATTGGAAGAAACGAAAGGTCAGATCAACAACTCAA
TTAAGGATCTCACAGATGGCCACTTTGAGAACATTTTAGCTGACAACAGTGTGAACGACCA
GACCAAAATCCTTGTGGTTAATGCTGCCTACTTTGTTGGCAAGTGGATGAAGAAATTTTCT
GAATCAGAAACAAAAGAATGTCCTTTCAGAGTCAACAAGACAGACACCAAACCAGTGCAGA
TGATGAACATGGAGGCCACGTTCTGTATGGGAAACATTGACAGTATCAATTGTAAGATCAT
AGAGCTTCCTTTTCAAAATAAGCATCTCAGCATGTTCATCCTACTACCCAAGGATGTGGAG
GATGAGTCCACAGGCTTGGAGAAGATTGAAAAACAACTCAACTCAGAGTCACTGTCACAGT
GGACTAATCCCAGCACCATGGCCAATGCCAAGGTCAAACTCTCCATTCCAAAATTTAAGGT
GGAAAAGATGATTGATCCCAAGGCTTGTCTGGAAAATCTAGGGCTGAAACATATCTTCAGT
GAAGACACATCTGATTTCTCTGGAATGTCAGAGACCAAGGGAGTGGCCCTATCAAATGTTA
TCCACAAAGTGTGCTTAGAAATAACTGAAGATGGTGGGGATTCCATAGAGGTGCCAGGAGC
ACGGATCCTGCAGCACAAGGATGAATTGAATGCTGACCATCCCTTTATTTACATCATCAGG
CACAACAAAACTCGAAACATCATTTTCTTTGGCAAATTCTGTTCTCCT**TAA**GTGGCATAGC
CCATGTTAAGTCCTCCCTGACTTTTCTGTGGATGCCGATTTCTGTAAACTCTGCATCCAGA
GATTCATTTTCTAGATACAATAAATTGCTAATGTTGCTGGATCAGGAAGCCGCCAGTACTT
GTCATATGTAGCCTTCACACAGATAGACCTTTTTTTTTTTTCCAATTCTATCTTTTGTTTC
CTTTTTTCCCATAAGACAATGACATACGCTTTTAATGAAAAGGAATCACGTTAGAGGAAAA
ATATTTATTCATTATTTGTCAAATTGTCCGGGTTAGTTGGCAGAAATACAGTCTTCCACAA
AGAAAATTCCTATAAGGAAGATTTGGAAGCTCTTCTTCCCAGCACTATGCTTTCCTTCTTT
GGGATAGAGAATGTTCCAGACATTCTCGCTTCCCTGAAAGACTGAAGAAAGTGTAGTGCAT
GGGACCCACGAAACTGCCCTGGCTCCAGTGAAACTTGGGCACATGCTCAGGCTACTATAGG
TCCAGAAGTCCTTATGTTAAGCCCTGGCAGGCAGGTGTTTATTAAAATTCTGAATTTTGGG
GATTTTCAAAAGATAATATTTTACATACACTGTATGTTATAGAACTTCATGGATCAGATCT
GGGGCAGCACCCTATAAATCAACACCTTAATATGCTGCAACAAAATGTAGAATATTCAGAC
AAAATGGATACATAAAGACTAAGTAGCCCATAAGGGGTCAAAATTTGCTGCCAAATGCGTA
TGCCACCAACTTACAAAAACACTTCGTTCGCAGAGCTTTTCAGATTGTGGAATGTTGGATA
AGGAATTATAGACCTCTAGTAGCTGAAATGCAAGACCCCAAGAGGAAGTTCAGATCTTAAT
ATAAATTCACTTTCATTTTTGATAGCTGTCCCATCTGGTCATTTGGTTGGCACTAGACTGG

TGGCAGGGGCTTCTAGCTGACTTGCACAGGGATTCTCACAATAGCCGATATCAGAATTTGT
GTTGAAGGAACTTGTCTCTTCATCTAATATGATAGCGGGAAAAGGAGAGGAAACTACTGCC
TTTAGAAAATATAAGTAAAGTGATTAAAGTGCTCACGTTACCTTGACACATAGTTTTTCAG
TCTATGGGTTTAGTTACTTTAGATGGCAAGCATGTAACTTATATTAATAGTAATTTGTAAA
GTTGGTTGGATAAGCTATCCGTGTTGCAGGTTCATGGATTACTTCTCTATAAAAAATATGT
ATTTACCAAAAATTTTGTGACATTCCTTCTCCCATCTCTTCCTTGACCTGCATTGTAAATA
GGTTCTTCTTGTTCTGAGATTCAATATTGAATTTTTCCTATGCTATTGACAATAAAATATT
ATTGAACTACA

FIGURE 16

GCCGAGCCGGGGCTCCGGAAGCCGGCGGGGGCGCCGCGGCCGTGCGGGGCGTCA**ATG**GATC
GCCACTCCAGCTACATCTTCATCTGGCTGCAGCTGGAGCTCTGCGCCATGGCCGTGCTGCT
CACCAAAGGTGAAATTCGATGCTACTGTGATGCTGCCCACTGTGTAGCCACTGGTTATATG
TGTAAATCTGAGCTCAGCGCCTGCTTCTCTAGACTTCTTGATCCTCAGAACTCAAATTCCC
CACTCACCCATGGCTGCCTGGACTCTCTTGCAAGCACGACAGACATCTGCCAAGCCAAACA
GGCCCGAAACCACTCTGGCACCACCATACCCACATTGGAATGCTGTCATGAAGACATGTGC
AATTACAGAGGGCTGCACGATGTTCTCTCTCCTCCCAGGGGTGAGGCCTCAGGACAAGGAA
ACAGGTATCAGCATGATGGTAGCAGAAACCTTATCACCAAGGTGCAGGAGCTGACTTCTTC
CAAAGAGTTGTGGTTCCGGGCAGCGGTCATTGCCGTGCCCATTGCTGGAGGGCTGATTTTA
GTGTTGCTTATTATGTTGGCCCTGAGGATGCTTCGAAGTGAAAATAAGAGGCTGCAGGATC
AGCGGCAACAGATGCTCTCCCGTTTGCACTACAGCTTTCACGGACACCATTCCAAAAAGGG
GCAGGTTGCAAAGTTAGACTTGGAATGCATGGTGCCGGTCAGTGGGCACGAGAACTGCTGT
CTGACCTGTGATAAAATGAGACAAGCAGACCTCAGCAACGATAAGATCCTCTCGCTTGTTC
ACTGGGGCATGTACAGTGGGCACGGGAAGCTGGAATTCGTA**TGA**CGGAGTCTTATCTGAAC
TACACTTACTGAACAGCTTGAAGGCCTTTTGAGTTCTGCTGGACAGGAGCACTTTATCTGA
AGACAAACTCATTTAATCATCTTTGACAGACAAAATGACCTCTGCAAACAGAATCTTGGAT
ATTTCTTCTGAAGGATTATTTGCACAGACTTAAATACAGTTAAATGTGTTATTTGCTTTTA
AAATTATAAAAAGCAAAGAGAAGACTTTGTACACACTGTCACCAGGGTTATTTGCATCCAA
GGGAGCTGGAATTGAGTACCTAAATAAACAAAAATGTGCCCTAAAAA

FIGURE 17A

GTCACTTAGGAAAAGGTGTCCTTTCGGGCAGCCGGGCTCAGCATGAGGAACAGAAGGAATG
ACACTCTGGACAGCACCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTA
CAGTGAAAGCGACTTGGTGAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTC
TTTACCAAAGATTCCAAGGCCACGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGC
ACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGA
ATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTAT
ATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGC
ACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAA
GCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGG
ATTCTCACGGGAGGCACCCATTATGGCCTGACGAAGTACATCGGGGAGGTGGTGAGAGATA
ACACCATCAGCAGGAGTTCAGAGGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCAT
GGTCTCCAACCGGGACACCCTCATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAG
TACCTTATGGATGACTTCACAAGGGATCCACTGTATATCCTGGACAACAACCACACACATT
TGCTGCTCGTGGACAATGGCTGTCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCA
GCTAGAGAAGCATATCTCTGAGCGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCC
ATTGTGTGTTTTGCCCAAGGAGGTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCA
AAAATAAAATTCCTTGTGTGGTGGTGGAAGGCTCGGGCCGGATCGCTGATGTGATCGCTAG
CCTGGTGGAGGTGGAGGATGCCCCGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTT
TTACCCCGCACGGTGTCCCGGCTGTCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCA
AAGAAATTCTCGAATGTTCTCACCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGA
AATTGTGAGCAATGCCATCTCCTACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGAC
AAGGATAACTGGAATGGGCAGCTGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCA
ATGATGAGATTTTCACCAATGACCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTT
TACGGCTCTCATAAAGGACAGACCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAAC
CTACGGAAGTTTCTCACCCATGATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGC
TTGTGTACCGGAATCTGCAGATCGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGT
CTGGAAACTGGTTGCGAACTTCCGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGAC
GAGATGGACATAGAACTCCACGACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCT
TCATCTGGGCCATTCTTCAGAATAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAG
GGGCTGCACTCTGGCAGCCCTGGGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAG
AACGACATCAATGCTGCTGGGGAGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTG
TTGAGCTGTTCACTGAGTGTTACAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTA
TTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAG
CATTTCACCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTT
CCCGAGACACCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATACCCTTGGTGGGCTG

TGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTAT
GTGGCGTTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCT
TCCTCCTGCTGTTTGCCTACGTGCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGA
GCTGGTCCTGTACTCGCTGGTCTTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTA
AATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACT
TCATAGCAGGAATTGTATTTCGGCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACG
AGTCATTTTCTGTCTGGACTACATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTA
AGCAGAAACTTAGGACCCAAGATTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCT
TCCTGTTCCTCTTTGCGGTGTGGATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAG
GCAGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCC
ATGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCT
TCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTT
CCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTG
GTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACC
AGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCA

FIGURE 17B

GCCGCCTCAATATCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAA
GTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTTTCAAAAAT
GAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAAGGAAAACTACCTTGTCAAGATCA
ACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCATCGATTTAGACAACTGGATACAAA
GCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTAATAAAATCAAATAAAACTGTATG
AAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCATATTAAGGAATGCTGATGAACAA
TTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCCTGGGTACATGGTGGATGATTTT
AAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGTGTGATTGGTTTCATACTTGAAG
ACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTCCAGAATGGTGCCTGTTTCTCTC
TGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAGTGTGTTCTTACCGCCTCCTTTT
TCCTTTAATCTTATTTTTGATGAACACATATATAGGAGAACATCTATCCTATGAATAAGAA
CCTGGTCATGCTTTACTCCTGTATTGTTATTTTGTTCATTTCCAATTGATTCTCTACTTTT
CCCTTTTTTTGTATTATGTGACTAATTAGTTGGCATATTGTTAAAAGTCTCTCAAATTAGGC
CAGATTCTAAAACATGCTGCAGCAAGAGGACCCCGCTCTCTTCAGGAAAAGTGTTTTCATT
TCTCAGGATGCTTCTTACCTGTCAGAGGAGGTGACAAGGCAGTCTCTTGCTCTCTTGGACT
CACCAGGCTCCTATTGAAGGAACCACCCCCATTCCTAAATATGTGAAAAGTCGCCCAAAAT
GCAACCTTGAAAGGCACTACTGACTTTGTTCTTATTGGATACTCCTCTTATTTATTATTTT
TCCATTAAAAATAATAGCTGGCTATTATAGAAAATTTAGACCATACAGAGATGTAGAAAGA
ACATAAATTGTCCCCATTACCTTAAGGTAATCACTGCTAACAATTTCTGGATGGTTTTTCA
AGTCTATTTTTTTTCTATGTATGTCTCAATTCTCTTTCAAAATTTTACAGAATGTTATCAT
ACTACATATATACTTTTTATGTAAGCTTTTTTCACTTAGTATTTTATCAAATATGTTTTTAT
TATATTCATAGCCTTCTTAAACATTATATCAATAATTGCATAATAGGCAACCTCTAGCGAT
TACCATAATTTTGCTCATTGAAGGCTATCTCCAGTTGATCATTGGGATGAGCATCTTTGTG
CATGAATCCTATTGCTGTATTTGGGAAAATTTTCCAAGGTTAGATTCCAATAAATATCTAT
TTATTATTAAATATTAAAATATCGATTTATTATTAAAACCATTTATAAGGCTTTTTCATAA
ATGTATAGCAAATAGGAATTATTAACTTGAGCATAAGATATGAGATACATGAACCTGAACT
ATTAAAATAAAATATTATATTTAACCCTAGTTTAAGAAGAAGTCAATATGCTTATTTAAAT
ATTATGGATGGTGGGCAGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGG
CAAAACCACATCTCTACTAAAAATAAAAAAATTAGCTGGGTGTGGTGGTGCACTCCTGTAA
TCCCAGCTACTCAGAAGGCTGAGGTACAAGAATTGCTGGAACCTGGGAGGCGGAGGTTGCA
GTGAACCAAGATTGCACCACTGCACTCCAGCCGGGGTGACAGAGTGAGACTCCGACTGAAA
ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATATTATGGATGGTGAAGGGAATGG
TATAGAATTGGAGAGATTATCTTACTGAACACCTGTAGTCCCAGCTTTCTCTGGAAGTGGT
GGTATTTGAGCAGGATGTGCACAAGGCAATTGAAATGCCCATAATTAGTTTCTCAGCTTTG
AATACACTATAAACTCAGTGGCTGAAGGAGGAAATTTTAGAAGGAAGCTACTAAAAGATCT

AATTTGAAAAACTACAAAAGCATTAACTAAAAAAGTTTATTTTCCTTTTGTCTGGGCAGTA
GTGAAAATAACTACTCACAACATTCACTATGTTTGCAAGGAATTAACACAAATAAAAGATG
CCTTTTTACTTAAACGCCAAGACAGAAAACTTGCCCAATACTGAGAAGCAACTTGCATTAG
AGAGGGAACTGTTAAATGTTTTCAACCCAGTTCATCTGGTGGATGTTTTTGCAGGTTACTC
TGAGAATTTTGCTTATGAAAAATCATTATTTTTAGTGTAGTTCACAATAATGTATTGAACA
TACTTCTAATCAAAGGTGCTATGTCCTTGTGTATGGTACTAAATGTGTCCTGTGTACTTTT
GCACAACTGAGAATCCTGCGGCTTGGTTTAATGAGTGTGTTCATGAAATAAATAATGGAGG
AATTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

FIGURE 18

GAATTCGCCCTTGGCTCAGCATGAGGAACAGAAGGAATGACACTCTGGACAGCACCCGGAC
CCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTACAGTGAAAGCGACTTGGTGAAT
TTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTCTTTACCAAAGATTCCAAGGCCA
CGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGCACATGGAAGGCACCCAGATCAA
CCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGAATTTCCTACCGACGCCTTTGGG
GATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTATATACGTCTGTCCTGCGACACGG
ACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGCACCTGAAAACACCCAACCTGGT
CATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAAGCCGCGCATGCGCAAGATCTTC
AGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGGAGGCACCCATT
ATGGCCTGATGAAGTACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCAGA
GGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAACCGGGACACCCTC
ATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAGTACCTTATGGATGACTTCACAA
GAGATCCACTGTATATCCTGGACAACAACCACACACATTTGCTGCTCGTGGACAATGGCTG
TCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAG
CGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAG
GTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGT
GGTGGAAGGCTCGGGCCAGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCC
CTGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGC
TGCCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCA
CCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCC
TACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGC
TGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCAATGATGAGATTTTCACCAATGA
CCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTTTACGGCTCTCATAAAGGACAGA
CCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGGAAGTTTCTCACCCATG
ATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATCTGCAGAT
CGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGTCTGGAAACTGGTTGCGAACTTC
CGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGACGAGATGGACATAGAACTCCACG
ACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAA
TAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAGGGGCTGCACTCTGGCAGCCCTG
GGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAGAACGACATCAATGCTGCTGGGG
AGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTGTTGAGCTGTTCACTGAGTGTTA
CAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTATTCCTGTGAAGCTTGGGGTGGA
AGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAGCATTTCATCGCCCAGCCTGGGG
TCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTTCCCGAGACACCAAGAACTGGAA
GATTATCCTGTGTCTGTTTATTATACCCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAG

AAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTATGTGGCGTTCTTCACCTCCCCCT
TCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCTTCCTCCTGCTGTTTGCCTACGT
GCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGAGCTGGTCCTGTACTCGCTGGTC
TTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTAAATGGGGTGAATTATTTTACTG
ACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACTTCATAGCAGGAATTGTATTTCG
GCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACGAGTCATTTTCTGTCTGGACTAC
ATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTAAGCAGAAACTTAGGACCCAAGA
TTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCTTCCTGTTCCTCTTTGCGGTGTG
GATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAGGCAGAATGAGCAGCGCTGGAGG
TGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCCATGTTCGGCCAGGTGCCCAGTG
ACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCTTCACTGGGAATGAGTCCAAGCC
ACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTTCCCCGAGTGGATCACCATCCCC
CTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTGGTCAACCTGCTGGTCGCCATGT
TTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACCAGGTCTGGAAGTTCCAGAGGTA
CTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAATATCCCCTTCCCCTTCATTGTCTTCGCT
TACTTCTACATGGTGGTGAAGAAGTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGT
CTTCTGTCTGCTGTTTCAAAAATGAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAA
GGAAAACTACCTTGTCAAGATCAACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCAT
CGATTTAGACAACTGGATACAAAGCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTA
ATAAAATCAAATAAAACTGTATGAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCAT
ATTAAGGAATGCTGATGAACAATTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCC
TGGGTACATGGTGGATGATTTTAAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGT
GTGATTGGTTTCATACTTGAAGACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTC
CAGAATGGTGCCTGTTTCTCTCTGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAG
TGTGTTCTTACCGCCTCCTTTTTCCTTTAATCTTATTTTTGATGAACACATATATAGGAGA
ACATCTATCCTATGAATAAGAACCTGGTCATGCTTTAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 19

GCGGCGGGGGCGAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGG
ACGCGGAGCGGATGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGA
GTGGGCGGGAGGGAGAGGGGGTGTCCGAGGGCCACAAGAGTATGACGGGGCTGTACGAGCT
GGTGTGGCGGGTGCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGC
GTTCGGTTCGGCACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGG
TCCTAGCGCCGCTCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCA
CCACCGGCACCCGCGCGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGC
GCGGACGGTCGTTCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGG
AGCAGGAACCCAGCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCAT
CTCCTACATTAGCGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATG
GATGAAATTTTAAAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAG
AATTTGCAAATAGTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGT
GCTGTCTCCGGAAGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTA
GTAGCCCAGAAGCAAAGGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTA
GTTCAAATGGTTGTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATT
AGGCTTTCTTCCCTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAAC
ATCAGTTATGAGGACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGG
GAAAG**TAG**TGGTCATTGGTTGCATAATTTGATTTGAGGCTTGTGGAGGAAAGGAACCAAGT
GACTCTGATGTTTACAAAGCACCTATGAAACCCTGTACACACCTATGAAACCCTGTACACA
CCTAGTTCATAATCCTCATAATTTATCAACAAACACAAAGAAGTGTCTTACTTGAGAGTGA
GTGTGTGCGTGTGTGCGTGCACACATGTGCACGTTTGTATGTGTGGAAATAAACATAAATG
GGGACGTGTTGGAGAAGGAAATACATAGACCTACAACTTTGAGCATATAGCAGTGATGTTT
TAGGAACTGAAATGTCACACTTAATAAAGTCTTCAGCCCAGCTACTTCCCTGTTTTCGTGG
GGAGAAGGGGGCCTGATTAGAACTGTTCTGGTTGTGTTTGGCGGGAGGGGAATAAATTTTG
TTCAGTCCTTCTTAGTGACCAAACTTTAATTTTTAGGAATAACATATTGACTTACTGAACT
GAAGCATTCTGAGTTGAAAGGAGCCCCAGAGGAAAGGAGTTCTGTGTTGCTCACATGTTAA
AAGCTTGCTCACCTTCAGAGCAGAGGGAATACCTATCTTCAGATATCCGCCCATTTTCATC
TCTTCATTATAGTCAAACAGTGTGACTTGAGAGTGTTGCTCTGGTGTCTGTATTCTGGCTT
ATGAAGATTATTTGAAAAAGAACTCTTACTACATTGAAATGCAGACTTTTAAAAATTTAAA
TATTGGATTAGGCAGTCAAAAAACCAAACAAGCATAAAAGGTCAGTAAGTTGTAATCTTAA
AAGTAAAGGTGGAAAACTCATTATAAATGGAAGAAAAGTTTTGATTTCCTTTTTTGTTTGA
TGGGCAGTATGCCATATTATATCCAAAGTTGGTTTAAAAAATACTTCCATCAACTATTTTT
ATTTAAAATAAACATTTGAGGGAAGTTACCAAGGCAGCTTTTTTCCTCAAAAGTAACCTGT
TCCTCTTTGGAACAGCACATTTTAGGGGCATGGTTAATACCTGAGATTTTTACTCAGTAAA
CCTGATGGTTACCGTGTGTAAAATATCTTTAAGTAGGATTGAAGGCCTCTGTGGGGAATAT

AATGTTACCAAAGCCTATAAAAATAAATTTTACATGTTCTTTTTTATGACAGAGAGCAGCA
CTGGTTCTGTTATTTTTTAAAATGAATAACTGATTTCTTGACAGGTGTTTAGTATTTCTTC
CCTCACTGCTGATTCTTGGATAGAAACCATTCTTTATATTTGATAGACTGTTTTCAGAAAA
CTCTTATCAACAAGTGTACAATAGTTATCTAAAACTGTACATTTAGAATGGAGCAGTTTAA
TACTAGATCTCAGAAATTTGAAAAATAGCAAAGAAGACTGGATTTGGAAAGCATGGTCTAC
AATTGCTTGTCAATTCTGAAGCTATGAAGAATAAATGTTTCAGCTTTGGATTATGAAACCC
CGTTTATGATTTTTAAAACACACTTGAAATAAAAATGATTGAACTAAATTTTGGTCCGGTG
ACATCATTTTGCACTGCATAGCCCATTATACATTGTAAGTCTTTTTGTTTTAATTTATTAC
TGAGAGTTTTCTGTGAAGCTACAGCATATCTAATCAGAGAATTTCTGATTTGTTTCCTTAT
ACTGTGATTATGTTATATTAAGGCATTTGTAGTGCAGCTTAAGACTGAATTTATGCTTTTT
GTAAACATGATCGGTGTAAATGTCTTATAAACATTCTGGAGTATGTATAACTTTAATGAAT
GAAATTTAATGGACCTGATTAAAATAAAGGGATTTAATCTTTGTTAAAGTGAAGTTAGTGG
AATAAATTACCTACTGGAATATAGTCCAAGTCACTAAAGGTTTAATATTTGCATTTTTGTG
CTTTTATTTTCTCCTTCCATTCATAATTATATACTTGAAAGTACATCTGTAGCCTATGACT
TGAGTCTCTTGAACTTCTAGGAAGAGGCAAACTACAAACTACTAGGATTCTGATTTCAGAT
ATAGGCATTCCAGAATCTTCTCTTTACGAGTTCACCTGCTAGTATAATCTCCACAACTTGA
ATGGCCTTGGTTGTTCTGTAATTGCTGCCAAAATCATCACAAGCTGTACGTCATCAAGGCT
CCCTTTGCACTCCCAAGAAGAACTGTTCATTTTAAACAAAAGTGTATGTCTTTATTTGTAT
TGGAAAATATTGTCTTTAAATTGTTTCTTGTTGACACTTCCCACAATGGAAAAATTACCAA
ATTAAACCTGTTTTATGGATGGCAGCTTGGAGCATAGCAAGAATTTGGAGGATTTGAATTC
CATTCCCAGTTCTCATTGCGTTTTGTTTCTTAAAACTATAATAATTGGTTACTGTTATAAA
GTTTAAAAGGTGGTTTTAATATGAATAGCAAATTCTGGTATATTGTGACTAACGCTTAAGA
ATGCCTGTCTTTGAGAGGAAGGTGTTATAATATTAATCAACAGTGCCAAATACACTGTGCA
TATCTACAATTTAATCTTTGAATGTTTGTTACTGGATTAGCTCCCTCCTCCTTCTGTGTGA
TGGTACCATGCATAGAGTCAATCAAATCCTTGTGATGTTTTGTATGGACTTTGACAACATG
TAACTAATGTGTAAAGCAAGTTTTTATGATTAAGGAATCAAATTTATTGAATTTTATTATT
GAAAGTTGAAACTTAACATGTATAAACAAAAAACAATAAAATAATAAACTATTTTCATTGA
CTATAAAAAAAAAGCCCACGCGTCCGCC

FIGURE 20

GAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGGACGCGGAGCGA
TGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGAGTGGGCGGGAGG
GAGAGGGGGTGTCCGAGGGCCACAAGAGT**ATG**ACGGGGCTGTACGAGCTGGTGTGGCGGGT
GCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGCGTTCGGTTCGGC
ACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGGTCCTAGCGCCGC
TCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCACCACCGGCACCC
GCGCGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGCGCGGACGGTCGT
TCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGGAGCAGGAACCCA
GCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCATCTCCTACATTAG
CGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATGGATGAAATTTTA
AAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAGAATTTGCAAATA
GTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGTGCTGTCTCCGGA
AGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTAGTAGCCCAGAAG
CAAAAGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTAGTTCAAATGGTT
GTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATTAGGCTTTCTTCC
CTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAACATCAGTTATGAG
GACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGGGAAAG**TAG**TGGT
CATTGGTTGCA

FIGURE 21

GGGGGGCGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGA
TCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGC
TGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCA
GTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC
TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACT**ATG**GTCCAGAGGCTGT
GGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT
TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG
GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT
GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC
CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGC
TGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGG
GCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTC
TGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC
AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGG
CCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCAC
CTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCA
GAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTT
TCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCAC
CCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTG
TTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCA
CCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCA
GTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT
CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT
CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCT
GCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC
AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC
CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCT
GCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTG
GTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCA
TCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGT
CCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCC
ATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG**TAG**A
AAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGT
TAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTC
TCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT

CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCT
GGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGTGGGGACTCTGCAGGTGG
ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTT
TTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCT
TCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGA
ACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCC
CTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG
GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTA
TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGT
AGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTG
CCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAAT
TCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAG
GTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAG
CCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACT
GCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA
ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCC
ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCC
CCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGTTTAAGTGCCGTTTGCAATAATG
TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAAT
GTTTATGGTGACAAAATTAAAGCCTTTCTTATATGTTT

FIGURE 22

AACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGCCCCTGGTTCGGCCCACCTCT
GAAGGTTCCAGAATCGATAGTGAATTCGTGGGGAAGCGCAGCC**ATG**GCTCTGCGGAGGCTG
GGGGCCGCGCTGCTGCTGCTGCCGCTGCTCGCCGCCGTGGAAGAAACGCTAATGGACTCCA
CTACAGCGACTGCTGAGCTGGGCTGGATGGTGCATCCTCCATCAGGGTGGGAAGAGGTGAG
TGGCTACGATGAGAACATGAACACGATCCGCACGTACCAGGTGTGCAACGTGTTTGAGTCA
AGCCAGAACAACTGGCTACGGACCAAGTTTATCCGGCGCCGTGGCGCCCACCGCATCCACG
TGGAGATGAAGTTTTCGGTGCGTGACTGCAGCAGCATCCCCAGCGTGCCTGGCTCCTGCAA
GGAGACCTTCAACCTCTATTACTATGAGGCTGACTTTGACTCGGCCACCAAGACCTTCCCC
AACTGGATGGAGAATCCATGGGTGAAGGTGGATACCATTGCAGCCGACGAGAGCTTCTCCC
AGGTGGACCTGGGTGGCCGCGTCATGAAAATCAACACCGAGGTGCGGAGCTTCGGACCTGT
GTCCCGCAGCGGCTTCTACCTGGCCTTCCAGGACTATGGCGGCTGCATGTCCCTCATCGCC
GTGCGTGTCTTCTACCGCAAGTGCCCCCGCATCATCCAGAATGGCGCCATCTTCCAGGAAA
CCCTGTCGGGGGCTGAGAGCACATCGCTGGTGGCTGCCCGGGGCAGCTGCATCGCCAATGC
GGAAGAGGTGGATGTACCCATCAAGCTCTACTGTAACGGGGACGGCGAGTGGCTGGTGCCC
ATCGGGCGCTGCATGTGCAAAGCAGGCTTCGAGGCCGTTGAGAATGGCACCGTCTGCCGAG
GTTGTCCATCTGGGACTTTCAAGGCCAACCAAGGGGATGAGGCCTGTACCCACTGTCCCAT
CAACAGCCGGACCACTTCTGAAGGGGCCACCAACTGTGTCTGCCGCAATGGCTACTACAGA
GCAGACCTGGACCCCCTGGACATGCCCTGCACAACCATCCCCTCCGCGCCCCAGGCTGTGA
TTTCCAGTGTCAATGAGACCTCCCTCATGCTGGAGTGGACCCCTCCCCGCGACTCCGGAGG
CCGAGAGGACCTCGTCTACAACATCATCTGCAAGAGCTGTGGCTCGGGCCGGGGTGCCTGC
ACCCGCTGCGGGGACAATGTACAGTACGCACCACGCCAGCTAGGCCTGACCGAGCCACGCA
TTTACATCAGTGACCTGCTGGCCCACACCCAGTACACCTTCGAGATCCAGGCTGTGAACGG
CGTTACTGACCAGAGCCCCTTCTCGCCTCAGTTCGCCTCTGTGAACATCACCACCAACCAG
GCAGCTCCATCGGCAGTGTCCATCATGCATCAGGTGAGCCGCACCGTGGACAGCATTACCC
TGTCGTGGTCCCAGCCAGACCAGCCCAATGGCGTGATCCTGGACTATGAGCTGCAGTACTA
TGAGAAGGAGCTCAGTGAGTACAACGCCACAGCCATAAAAAGCCCCACCAACACGGTCACC
GTGCAGGGCCTCAAAGCCGGCGCCATCTATGTCTTCCAGGTGCGGGCACGCACCGTGGCAG
GCTACGGGCGCTACAGCGGCAAGATGTACTTCCAGACCATGACAGAAGCCGAGTACCAGAC
AAGCATCCAGGAGAAGTTGCCACTCATCATCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTC
ATTGCTGTGGTTGTCATCGCCATCGTGTGTAACAGAAGACGGGGGTTTGAGCGTGCTGACT
CGGAGTACACGGACAAGCTGCAACACTACACCAGTGGCCACATGACCCCAGGCATGAAGAT
CTACATCGATCCTTTCACCTACGAGGACCCCAACGAGGCAGTGCGGGAGTTTGCCAAGGAA
ATTGACATCTCCTGTGTCAAAATTGAGCAGGTGATCGGAGCAGGGGAGTTTGGCGAGGTCT
GCAGTGGCCACCTGAAGCTGCCAGGCAAGAGAGAGATCTTTGTGGCCATCAAGACGCTCAA
GTCGGGCTACACGGAGAAGCAGCGCCGGGACTTCCTGAGCGAAGCCTCCATCATGGGCCAG

TTCGACCATCCCAACGTCATCCACCTGGAGGGTGTCGTGACCAAGAGCACACCTGTGATGA
TCATCACCGAGTTCATGGAGAATGGCTCCCTGGACTCCTTTCTCCGGCAAAACGATGGGCA
GTTCACAGTCATCCAGCTGGTGGGCATGCTTCGGGGCATCGCAGCTGGCATGAAGTACCTG
GCAGACATGAACTATGTTCACCGTGACCTGGCTGCCCGCAACATCCTCGTCAACAGCAACC
TGGTCTGCAAGGTGTCGGACTTTGGGCTCTCACGCTTTCTAGAGGACGATACCTCAGACCC
CACCTACACCAGTGCCCTGGGCGGAAAGATCCCCATCCGCTGGACAGCCCCGGAAGCCATC
CAGTACCGGAAGTTCACCTCGGCCAGTGATGTGTGGAGCTACGGCATTGTCATGTGGGAGG
TGATGTCCTATGGGGAGCGGCCCTACTGGGACATGACCAACCAGGATGTAATCAATGCCAT
TGAGCAGGACTATCGGCTGCCACCGCCCATGGACTGCCCGAGCGCCCTGCACCAACTCATG
CTGGACTGTTGGCAGAAGGACCGCAACCACCGGCCCAAGTTCGGCCAAATTGTCAACACGC
TAGACAAGATGATCCGCAATCCCAACAGCCTCAAAGCCATGGCGCCCCTCTCCTCTGGCAT
CAACCTGCCGCTGCTGGACCGCACGATCCCCGACTACACCAGCTTTAACACGGTGGACGAG
TGGCTGGAGGCCATCAAGATGGGGCAGTACAAGGAGAGCTTCGCCAATGCCGGCTTCACCT
CCTTTGACGTCGTGTCTCAGATGATGATGGAGGACATTCTCCGGGTTGGGCTCACTTTGGC
TGGCCACCAGAAAAAAATCCTGAACAGTATCCAGGTGATGCGGGCGCAGATGAACCAGATT
CAGTCTGTGGAGGTT**TGA**CATTCACCTGCCTCGGCTCACCTCTTCCTCCAAGCCCCGCCCC
CTCTGCCCCACGTGCCGGCCCTCCTGGTGCTCTATCCACTGCAGGGCCAGCCACTCGCCAG
GAGGCCACGGGCCACGGGAAGAACCAAGCGGTGCCAGCCACGAGACGTCACCAAGAAAACA
TGCAACTCAAACGACGGAAAAAAAAAGGGAATGGGAAAAAAGAAAACAGATCCTGGGAGGG
GGCGGGAAATACAAGGAATATTTTTTAAAGAGGATTCTCATAAGGAAAGCAATGACTGTTC
TTGCGGGGGATAAAAAAGGGCTTGGGAGATTCATGCGATGTGTCCAATCGGAGACAAAAGC
AGTTTCTCTCCAACTCCCTCTGGGAAGGTGACCTGGCCAGAGCCAAGAAACACTTTCAGAA
AAACAAATGTGAAGGGGAGAGACAGGGGCCGCCCTTGGCTCCTGTCCCTGCTGCTCCTCTA
GGCCTCACTCAACAACCAAGCGCCTGGAGGACGGGACAGATGGACAGACAGCCACCCTGAG
AACCCCTCTGGGAAAATCTATTCCTGCCACCACTGGGCAAACAGAAGAATTTTTCTGTCTT
TGGAGAGTATTTTAGAAACTCCAATGAAAGACACTGTTTCTCCTGTTGGCTCACAGGGCTG
AAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGAACGCGGGGACCCCAG

FIGURE 23

CGGGAAGCGCAGCCATGGCTCTGCGGAGGCTGGGGGCCGCGCTGCTGCTGCTGCCGCTGCT
CGCCGCCGTGGAAGAAACGCTAATGGACTCCACTACAGCGACTGCTGAGCTGGGCTGGATG
GTGCATCCTCCATCAGGGTGGGAAGAGGTGAGTGGCTACGATGAGAACATGAACACGATCC
GCACGTACCAGGTGTGCAACGTGTTTGAGTCAAGCCAGAACAACTGGCTACGGACCAAGTT
TATCCGGCGCCGTGGCGCCCACCGCATCCACGTGGAGATGAAGTTTTCGGTGCGTGACTGC
AGCAGCATCCCCAGCGTGCCTGGCTCCTGCAAGGAGACCTTCAACCTCTATTACTATGAGG
CTGACTTTGACTCGGCCACCAAGACCTTCCCCAACTGGATGGAGAATCCATGGGTGAAGGT
GGATACCATTGCAGCCGACGAGAGCTTCTCCCAGGTGGACCTGGGTGGCCGCGTCATGAAA
ATCAACACCGAGGTGCGGAGCTTCGGACCTGTGTCCCGCAGCGGCTTCTACCTGGCCTTCC
AGGACTATGGCGGCTGCATGTCCCTCATCGCCGTGCGTGTCTTCTACCGCAAGTGCCCCCG
CATCATCCAGAATGGCGCCATCTTCCAGGAAACCCTGTCGGGGGCTGAGAGCACATCGCTG
GTGGCTGCCCGGGGCAGCTGCATCGCCAATGCGGAAGAGGTGGATGTACCCATCAAGCTCT
ACTGTAACGGGGACGGCGAGTGGCTGGTGCCCATCGGGCGCTGCATGTGCAAAGCAGGCTT
CGAGGCCGTTGAGAATGGCACCGTCTGCCGAGGTTGTCCATCTGGGACTTTCAAGGCCAAC
CAAGGGGATGAGGCCTGTACCCACTGTCCCATCAACAGCCGGACCACTTCTGAAGGGGCCA
CCAACTGTGTCTGCCGCAATGGCTACTACAGAGCAGACCTGGACCCCCTGGACATGCCCTG
CACAACCATCCCCTCCGCGCCCCAGGCTGTGATTTCCAGTGTCAATGAGACCTCCCTCATG
CTGGAGTGGACCCCTCCCCGCGACTCCGGAGGCCGAGAGGACCTCGTCTACAACATCATCT
GCAAGAGCTGTGGCTCGGGCCGGGGTGCCTGCACCCGCTGCGGGGACAATGTACAGTACGC
ACCACGCCAGCTAGGCCTGACCGAGCCACGCATTTACATCAGTGACCTGCTGGCCCACACC
CAGTACACCTTCGAGATCCAGGCTGTGAACGGCGTTACTGACCAGAGCCCCTTCTCGCCTC
AGTTCGCCTCTGTGAACATCACCACCAACCAGGCAGCTCCATCGGCAGTGTCCATCATGCA
TCAGGTGAGCCGCACCGTGGACAGCATTACCCTGTCGTGGTCCCAGCCGGACCAGCCCAAT
GGCGTGATCCTGGACTATGAGCTGCAGTACTATGAGAAGGAGCTCAGTGAGTACAACGCCA
CAGCCATAAAAAGCCCCACCAACACGGTCACCGTGCAGGGCCTCAAAGCCGGCGCCATCTA
TGTCTTCCAGGTGCGGGCACGCACCGTGGCAGGCTACGGGCGCTACAGCGGCAAGATGTAC
TTCCAGACCATGACAGAAGCCGAGTACCAGACAAGCATCCAGGAGAAGTTGCCACTCATCA
TCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTCATTGCTGTGGTTGTCATCGCCATCGTGTG
TAACAGACGGGGGTTTGAGCGTGCTGACTCGGAGTACACGGACAAGCTGCAACACTACACC
AGTGGCCACATGACCCCAGGCATGAAGATCTACATCGATCCTTTCACCTACGAGGACCCCA
ACGAGGCAGTGCGGGAGTTTGCCAAGGAAATTGACATCTCCTGTGTCAAAATTGAGCAGGT
GATCGGAGCAGGGGAGTTTGGCGAGGTCTGCAGTGGCCACCTGAAGCTGCCAGGCAAGAGA
GAGATCTTTGTGGCCATCAAGACGCTCAAGTCGGGCTACACGGAGAAGCAGCGCCGGGACT
TCCTGAGCGAAGCCTCCATCATGGGCCAGTTCGACCATCCCAACGTCATCCACCTGGAGGG
TGTCGTGACCAAGAGCACACCTGTGATGATCATCACCGAGTTCATGGAGAATGGCTCCCTG

GACTCCTTTCTCCGGCAAAACGATGGGCAGTTCACAGTCATCCAGCTGGTGGGCATGCTTC
GGGGCATCGCAGCTGGCATGAAGTACCTGGCAGACATGAACTATGTTCACCGTGACCTGGC
TGCCCGCAACATCCTCGTCAACAGCAACCTGGTCTGCAAGGTGTCGGACTTTGGGCTCTCA
CGCTTTCTAGAGGACGATACCTCAGACCCCACCTACACCAGTGCCCTGGGCGGAAAGATCC
CCATCCGCTGGACAGCCCCGGAAGCCATCCAGTACCGGAAGTTCACCTCGGCCAGTGATGT
GTGGAGCTACGGCATTGTCATGTGGGAGGTGATGTCCTATGGGGAGCGGCCCTACTGGGAC
ATGACCAACCAGGATGTAATCAATGCCATTGAGCAGGACTATCGGCTGCCACCGCCCATGG
ACTGCCCGAGCGCCCTGCACCAACTCATGCTGGACTGTTGGCAGAAGGACCGCAACCACCG
GCCCAAGTTCGGCCAAATTGTCAACACGCTAGACAAGATGATCCGCAATCCCAACAGCCTC
AAAGCCATGGCGCCCCTCTCCTCTGGCATCAACCTGCCGCTGCTGGACCGCACGATCCCCG
ACTACACCAGCTTTAACACGGTGGACGAGTGGCTGGAGGCCATCAAGATGGGGCAGTACAA
GGAGAGCTTCGCCAATGCCGGCTTCACCTCCTTTGACGTCGTGTCTCAGATGATGATGGAG
GACATTCTCCGGGTTGGGGTCACTTTGGCTGGCCACCAGAAAAAAATCCTGAACAGTATCC
AGGTGATGCGGGCGCAGATGAACCAGATTCAGTCTGTGGAGGTT**TGA**CATTCACCTGCCTC
GGCTCACCTCTTCCTCCAAGCCCCGCCCCCTCTGCCCCACGTGCCGGCCCTCCTGGTGCTC
TATCCACTGCAGGGCCAGCCACTCGCCAGGAGGCCACGGGCCACGGGAAGAACCAAGCGGT
GCCAGCCACGAGACGTCACCAAGAAAACATGCAACTCAAACGACGGAAAAAAAAAGGGAAT
GGGAAAAAAGAAAACAGATCCTGGGAGGGGGCGGGAAATACAAGGAATATTTTTTAAAGAG
GATTCTCATAAGGAAAGCAATGACTGTTCTTGCGGGGGATAAAAAAGGGCTTGGGAGATTC
ATGCGATGTGTCCAATCGGAGACAAAAGCAGTTTCTCTCCAACTCCCTCTGGGAAGGTGAC
CTGGCCAGAGCCAAGAAACACTTTCAGAAAAACAAATGTGAAGGGGAGAGACAGGGGCCGC
CCTTGGCTCCTGTCCCTGCTGCTCCTCTAGGCCTCACTCAACAACCAAGCGCCTGGAGGAC
GGGACAGATGGACAGACAGCCACCCTGAGAACCCCTCTGGGAAAATCTATTCCTGCCACCA
CTGGGCAAACAGAAGAATTTTTCTGTCTTTGGAGAGTATTTTAGAAACTCCAATGAAAGAC
ACTGTTTCTCCTGTTGGCTCACAGGGCTGAAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGA
ACGCGGGGACCCCAGAAAGGTCAGCCTTCCTGAGGATGGGCAACCCCCAGGTCTGCAGCTC
CAGGTACATATCACGCGCACAGCCTGGCAGCCTGGCCCTCCTGGTGCCCACTCCCGCCAGC
CCCTGCCTCGAGGTCGAC

FIGURE 24

CCGCCAGCTCCTTCGGCA**ATG**AACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGG
AGTTCCTCCGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGT
ATGGACGCTTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTT
GTAGCTGTATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCT
ACACCAGGATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTT
CGTCTTGTACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGT
CACAACTTCAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCAACATCTGCTACG
CTGGAAATACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAG**TGA**TTTACCAT
TTTGATAATTAAAAGGAAAAAAAAAGGAAGACTCTCACTGTAAAAACAGCTGTAGGTATAA
TGTATATTCCCAGAGAATTGTATTTAACTAATTAATGTTTTTTATATTCTTAAATTTGCTC
ACAAATTGTGGTTTGTTACAATTAAACTGGATACTTATTTGCAAAGTGTTGTAGCTTATAA
TGAACTCTTAAGTATCTTATTAATGTATTAATGTCTTCATAGATCATATTTTCTTAGACAA
TGTTTAAATAGATAAATTGCTAATATTGAGAATGTGTCAAGTTTGTAAACCTAACTTTTAA
GATGCCAGATTCTTTTTTGATTAAATGTTGCAAAATCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 25

GCCCTTGGCA**ATG**AACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGGAGTTCCTC

CGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGTATGGACGC

TTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTTGTAGCTGT

ATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCTACACCAGG

ATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTTCGTCTTGT

ACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGTCACAACTT

CAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCACCATCTGCTACGCTGGAAAT

ACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAG**TGA**TTTACCAAAGGGC

FIGURE 26

GGCCCTGCCCTGAAGGCTGGTCACTTGCAGAGGTAAACTCCCCTCTTTGACTTCTGGCCAG
GGTTTGTGCTGAGCTGGCTGCAGCCGCTCTCAGCCTCGCTCCGGGCACGTCGGGCAGCCTC
GGGCCCTCCTGCCTGCAGGATC**ATG**CCCACCACCGTGGACGATGTCCTGGAGCATGGAGGG
GAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTCG
CGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCCC
CGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACACG
GTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTGG
ACTGGAACCAGAGCACCTTTGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGAG
CCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCATC
GTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCAG
TGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCCG
TAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCATT
TCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCAG
GCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAGT
GGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTAC
GCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTGCTCTGCCCAACTTCTTCTTCTTGC
TCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTGA
AGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCTT
CAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTGG
TCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCTC
TGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGAT
TTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGACC
GCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGGC
CTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAGA
ATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACAT
TCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCAC
GCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGGC
GTACTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCTT
TGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGAT
GATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAAC**TAA**GAAGAGAGACCGTTGCTG
CTGTCATGACCTAGCTTTGATGGCAGCAAGACCAAAAGTAGAAATCCCTGCACTCATCACA
AAGCCCATACAACTCAACCAAACTTACCCCTGAGCCCTATCAACCTAGGTCTACAGCCAGT
GGAGTCTATTGTACACTGTGGAAAAATACCCATGGGACCAGATCCTGCCAAATTCTTCCAG
CTCACTTTATTCTCAGCATTCCTAGGACATTGGACATTGGTTTTCTGGAGGGTTTTTTTTC
CGATCTTTGTATTTTTTTAAATTTGATTCTTTTCTTTGCAATGCTAGCAACCAGAATACAT

AGGGGAACTGTGGGCTAGGCAAANAAAATAGAAAAAGTGTGAAAAACAGTAAAGTTGGGAG

AGGAGCATCTATTTTCTTAAAGAAATAAAACACCNAAAACAAAAAAAAAAAAAAAAAAAAA

FIGURE 27

GCCCTTCTCCTGCCTGCAGGATC**ATG**CCCACCACCGTGGACGATGTCCTGGAGCATGGAGG
GGAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTC
GCGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCC
CCGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACAC
GGTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTG
GACTGGAACCAGAGCACCTTCGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGA
GCCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCAT
CGTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCA
GTGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCC
GTAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCAT
TTCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCA
GGCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAG
TGGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTA
CGCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTTCTCTGCCCAACTTCTTCTTCTTG
CTCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTG
AAGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCT
TCAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTG
GTCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCT
CTGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGA
TTTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGAC
CGCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGG
CCTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAG
AATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACA
TTCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCA
CGCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGG
CGTGCTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCT
TTGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGA
TGATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAAC**TAA**GAAGAGAGACCGTTGCT
GCTGTCATGACAAGGGC

FIGURE 28

CGGGCCAGGTTTCCAGGCTCGGCCGCCGCCTCCATCCCAGCACCTGCGGAGGGAGCGCTGA
CC**ATG**GCTCCCTGGCCTGAATTGGGAGATGCCCAGCCCAACCCCGATAAGTACCTCGAAGG
GGCCGCAGGTCAGCAGCCCACTGCCCCTGATAAAAGCAAAGAGACCAACAAAACAGATAAC
ACTGAGGCACCTGTAACCAAGATTGAACTTCTGCCGTCCTACTCCACGGCTACACTGATAG
ATGAGCCCACTGAGGTGGATGACCCCTGGAACCTACCCACTCTTCAGGACTCGGGGATCAA
GTGGTCAGAGAGAGACACCAAAGGGAAGATTCTCTGTTTCTTCCAAGGGATTGGGAGATTG
ATTTTACTTCTCGGATTTCTCTACTTTTTCGTGTGCTCCCTGGATATTCTTAGTAGCGCCT
TCCAGCTGGTTGGAGGAAAAATGGCAGGACAGTTCTTCAGCAACAGCTCTATTATGTCCAA
CCCTTTGTTGGGGCTGGTGATCGGGGTGCTGGTGACCGTCTTGGTGCAGAGCTCCAGCACC
TCAACGTCCATCGTTGTCAGCATGGTGTCCTCTTCATTGCTCACTGTTCGGGCTGCCATCC
CCATTATCATGGGGGCCAACATTGGAACGTCAATCACCAACACTATTGTTGCGCTCATGCA
GGTGGGAGATCGGAGTGAGTTCAGAAGAGCTTTTGCAGGAGCCACTGTCCATGACTTCTTC
AACTGGCTGTCCGTGTTGGTGCTCTTGCCCGTGGAGGTGGCCACCCATTACCTCGAGATCA
TAACCCAGCTTATAGTGGAGAGCTTCCACTTCAAGAATGGAGAAGATGCCCCAGATCTTCT
GAAAGTCATCACTAAGCCCTTCACAAAGCTCATTGTCCAGCTGGATAAAAAAGTTATCAGC
CAAATTGCAATGAACGATGAAAAAGCGAAAAACAAGAGTCTTGTCAAGATTTGGTGCAAAA
CTTTTACCAACAAGACCCAGATTAACGTCACTGTTCCCTCGACTGCTAACTGCACCTCCCC
TTCCCTCTGTTGGACGGATGGCATCCAAAACTGGACCATGAAGAATGTGACCTACAAGGAG
AACATCGCCAAATGCCAGCATATCTTTGTGAATTTCCACCTCCCGGATCTTGCTGTGGGCA
CCATCTTGCTCATACTCTCCCTGCTGGTCCTCTGTGGTTGCCTGATCATGATTGTCAAGAT
CCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCACTGTCATCAAGAAGACCATCAACACTGAT
TTCCCCTTTCCCTTTGCATGGTTGACTGGCTACCTGGCCATCCTCGTCGGGGCAGGCATGA
CCTTCATCGTACAGAGCAGCTCTGTGTTCACGTCGGCCTTGACCCCCCTGATTGGAATCGG
CGTGATAACCATTGAGAGGGCTTATCCACTCACGCTGGGCTCCAACATCGGCACCACCACC
ACCGCCATCCTGGCCGCCTTAGCCAGCCCTGGCAATGCATTGAGGAGTTCACTCCAGATCG
CCCTGTGCCACTTTTTCTTCAACATCTCCGGCATCTTGCTGTGGTACCCGATCCCGTTCAC
TCGCCTGCCCATCCGCATGGCCAAGGGGCTGGGCAACATCTCTGCCAAGTATCGCTGGTTC
GCCGTCTTCTACCTGATCATCTTCTTCTTCCTGATCCCGCTGACGGTGTTTGGCCTCTCGC
TGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGGGTTCCCGTCGTCTTCATCATCATCCTGGT
ACTGTGCCTCCGACTCCTGCAGTCTCGCTGCCCACGCGTCCTGCCGAAGAAACTCCAGAAC
TGGAACTTCCTGCCGCTGTGGATGCGCTCGCTGAAGCCCTGGGATGCCGTCGTCTCCAAGT
TCACCGGCTGCTTCCAGATGCGCTGCTGCTACTGCTGCCGCGTGTGCTGCCGCGCGTGCTG
CTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCTGCAGCAAGTGCTGCGAGGACTTGGAGGAG
GCGCAGGAGGGGCAGGATGTCCCTGTCAAGGCTCCTGAGACCTTTGATAACATAACCATTA
GCAGAGAGGCTCAGGGTGAGGTCCCTGCCTCGGACTCAAAGACCGAATGCACGGCCTTG**TA**

GGGGACGCCCCAGATTGTCAGGGATGGGGGGATGGTCCTTGAGTTTTGCATGCTCTCCTCC
CTCCCACTTCTGCACCCTTTCACCACCTCGAGGAGATTTGCTCCCCATTAGCGAATGAAAT
TGATGCAGTCCTAAAAAAAAAAA

FIGURE 29

CAGCCCAGCACCTGCGGAGGGAGCGCTGACC**ATG**GCTCCCTGGCCTGAATTGGGAGATGCC
CAGCCCAACCCCGATAAGTACCTCGAAGGGGCCGCAGGTCAGCAGCCCACTGCCCCTGATA
AAAGCAAAGAGACCAACAAAACAGATAACACTGAGGCACCTGTAACCAAGATTGAACTTCT
GCCGTCCTACTCCACGGCTACACTGATAGATGAGCCCACTGAGGTGGATGACCCCTGGAAC
CTACCCACTCTTCAGGACTCGGGGATCAAGTGGTCAGAGAGAGACACCAAAGGGAAGATTC
TCTGTTTCTTCCAAGGGATTGGGAGATTGATTTTACTTCTCGGATTTCTCTACTTTTTCGT
GTGCTCCCTGGATATTCTTAGTAGCGCCTTCCAGCTGGTTGGAGGAAAAATGGCAGGACAG
TTCTTCAGCAACAGCTCTATTATGTCCAACCCTTTGTTGGGGCTGGTGATCGGGGTGCTGG
TGACCGTCTTGGTGCAGAGCTCCAGCACCTCAACGTCCATCGTTGTCAGCATGGTGTCCTC
TTCATTGCTCACTGTTCGGGCTGCCATCCCCATTATCATGGGGGCCAACATTGGAACGTCA
ATCACCAACACTATTGTTGCGCTCATGCAGGTGGGAGATCGGAGTGAGTTCAGAAGAGCTT
TTGCAGGAGCCACTGTCCATGACTTCTTCAACTGGCTGTCCGTGTTGGTGCTCTTGCCCGT
GGAGGTGGCCACCCATTACCTCGAGATCATAACCCAGCTTATAGTGGAGAGCTTCCACTTC
AAGAATGGAGAAGATGCCCCAGATCTTCTGAAAGTCATCACTAAGCCCTTCACAAAGCTCA
TTGTCCAGCTGGATAAAAAAGTTATCAGCCAAATTGCAATGAACGATGAAAAAGCGAAAAA
CAAGAGTCTTGTCAAGATTTGGTGCAAAACTTTTACCAACAAGACCCAGATTAACGTCACT
GTTCCCTCGACTGCTAACTGCACCTCCCCTTCCCTCTGTTGGACGGATGGCATCCAAAACT
GGACCATGAAGAATGTGACCTACAAGGAGAACATCGCCAAATGCCAGCATATCTTTGTGAA
TTTCCACCTCCCGGATCTTGCTGTGGGCACCATCTTGCTCATACTCTCCCTGCTGGTCCTC
TGTGGTTGCCTGATCATGATTGTCAAGATCCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCA
CTGTCATCAAGAAGACCATCAACACTGATTTCCCCTTTCCCTTTGCATGGTTGACTGGCTA
CCTGGCCATCCTCGTCGGGGCAGGCATGACCTTCATCGTACAGAGCAGCTCTGTGTTCACG
TCGGCCTTGACCCCCCTGATTGGAATCGGCGTGATAACCATTGAGAGGGCTTATCCACTCA
CGCTGGGCTCCAACATCGGCACCACCACCACCGCCATCCTGGCCGCCTTAGCCAGCCCTGG
CAATGCATTGAGGAGTTCACTCCAGATCGCCCTGTGCCACTTTTTCTTCAACATCTCCGGC
ATCTTGCTGTGGTACCCGATCCCGTTCACTCGCCTGCCCATCCGCATGGCCAAGGGGCTGG
GCAACATCTCTGCCAAGTATCGCTGGTTCGCCGTCTTCTACCTGATCATCTTCTTCTTCCT
GATCCCGCTGACGGTGTTTGGCCTCTCGCTGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGG
GTTCCCGTCGTCTTCATCATCATCCTGGTACTGTGCCTCCGACTCCTGCAGTCTCGCTGCC
CACGCGTCCTGCCGAAGAAACTCCAGAACTGGAACTTCCTGCCGCTGTGGATGCGCTCGCT
GAAGCCCTGGGATGCCGTCGTCTCCAAGTTCACCGGCTGCTTCCAGATGCGCTGCTGCTGC
TGCTGCCGCGTGTGCTGCCGCGCGTGCTGCTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCT
GCAGCAAGTGCTGCGAGGACTTGGAGGAGGCGCAGGAGGGGCAGGATGTCCCTGTCAAGGC
TCCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGGTGAGGTCCCTGCCTCG
GACTCAAAGACCGAATGCACGGCCTTG**TAG**GGGACGCCCCAGATTGTCAGGGATGGGGGGA

TGGTCCTTGAGTTTTGCATGCTCTCCTCCCTCCCACTTCTGCACCCTTTCACCACCTCGAG
GAGATTTGCTCCCCATTAGCGAATGAAATTGATGCAGTCCTACCTAACTCGATTCCCTTTG
GCTTGGTGGGTAGGCCTGCAGGGCACTTTTATTCCAACCCATGGCCTCCATGACTTTTTCA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

GGCGTGGGAGCTGCCTAGGCCGGGCCCTGCCAGGGAGCAAGTCTGCATCGCAGAGGCCAGG
GCAGAGTGGCTCCTCACAGCCTGAAGCTCATCCTTCTGCACGGGCCAGCCAGGCCAGCACA
GAGGCACCAGGGCAGCAGTGCACACAGGTCCCCGGGGACCCCACC**ATG**TGGAGCGGATGGT
GGCTGTGGCCCCTTGTGGCCGTCTGCACTGCAGACTTCTTTCGGGACGAGGCAGAGAGGAT
CATGAGGGACTCCCCTGTCATTGATGGGCACAATGACCTCCCCTGGCAGCTGCTGGATATG
TTCAACAACCGGCTGCAGGACGAGAGGGCCAACCTGACCACCTTGGCCGGCACACACACCA
ACATCCCCAAGCTGAGGGCCGGCTTTGTGGGAGGCCAGTTCTGGTCCGTGTACACGCCCTG
CGACACCCAGAACAAAGACGCCGTGCGGAGGACGCTGGAGCAGATGGACGTGGTCCACCGC
ATGTGCCGGATGTACCCGGAGACCTTCCTGTATGTCACCAGCAGTGCAGGCATTCGGCAGG
CCTTCCGGGAAGGGAAGGTGGCCAGCCTGATCGGCGTGGAGGGCGGCCACTCCATTGACAG
CAGTTTGGGCGTCCTGCGGGCACTCTATCAGCTGGGCATGCGGTACCTGACCCTCACCCAC
AGCTGCAACACGCCCTGGGCTGACAACTGGCTGGTGGACACGGGAGACAGCGAGCCCCAGA
GCCAAGGCTTGTCACCCTTTGGGCAGCGTGTGGTGAAGGAGCTGAACCGTCTGGGGGTCCT
CATCGACTTGGCTCACGTGTCTGTGGCCACCATGAAGGCCACCCTGCAGCTGTCCAGAGCC
CCGGTCATCTTCAGCCACTCCTCGGCCTACAGCGTGTGCGCAAGCCGGCGCAACGTGCCTG
ACGACGTCCTGAGGCTGGTGAAACAGACAGACAGCCTGGTGATGGTGAACTTCTACAACAA
TTACATTTCCTGCACCAACAAGGCCAACCTGTCCCAAGTGGCCGACCATCTGGATCACATC
AAGGAGGTGGCAGGAGCCAGAGCCGTGGGTTTTGGTGGGGACTTTGATGGTGTTCCAAGGG
TCCCTGAGGGGCTGGAGGACGTCTCCAAGTATCCAGACCTGATCGCTGAGCTGCTCAGGAG
GAACTGGACGGAGGCGGAGGTCAAGGGCGCACTGGCTGACAACCTGCTGAGGGTCTTCCAG
GCTGTGGAACAGGCCAGCAACCTCACACAGGCTCCCGAGGAGGAGCCCATCCCGCTGGACC
AGCTGGGTGGCTCCTGCAGGACCCATTACGGCTACTCCTCTGGGGCTTCCAGCCTCCATCG
CCACTGGGGGCTCCTGCTGGCCTCCCTCGCTCCCCTGGTCCTCTGTCTGTCTCTCCTG**TGA**
AACCTGGGAGACCAGAGTCCCCTTTAGGGTTCCCGGAGCTCCGGGAAGACCCGCCCATCCC
AGGACTCCAGATGCCAGGAGCCCTGCTGCCCACATGCAAGGACCAGCATCTCCTGAGAGGA
CGCCTGGGCTTACCTGGGGGGCAGGATGCCTGGGGACAGTTCAGGACACACACACAGTAGG
CCCGCAATAAAAGCAACACCCCTTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 31

CATCTAAAGCCTCCTCAGCCTTCTGAGTCAGCCTGAAAGGAACAGGCCGAACTGCTGTATG
GGCTCTACTGCCAGTGTGACCTCACCCTCTCCAGTCACCCCTCCTCAGTTCCAGCTATGAG
TTCCTGCAACTTCACACATGCCACCTTTGTGCTTATTGGTATCCCAGGATTAGAGAAAGCC
CATTTCTGGGTTGGCTTCCCCCTCCTTTCCATGTATGTAGTGGCAATGTTTGGAAACTGCA
TCGTGGTCTTCATCGTAAGGACGGAACGCAGCCTGCACGCTCCGATGTACCTCTTTCTCTG
CATGCTTGCAGCCATTGACCTGGCCTTATCCACATCCACCATGCCTAAGATCCTTGCCCTT
TTCTGGTTTGATTCCCGAGAGATTAGCTTTGAGGCCTGTCTTACCCAGATGTTCTTTATTC
ATGCCCTCTCAGCCATTGAATCCACCATCCTGCTGGCCATGGCCTTTGACCGTTATGTGGC
CATCTGCCACCCACTGCGCCATGCTGCAGTGCTCAACAATACAGTAACAGCCCAGATTGGC
ATCGTGGCTGTGGTCCGCGGATCCCTCTTTTTTTTCCCACTGCCTCTGCTGATCAAGCGGC
TGGCCTTCTGCCACTCCAATGTCCTCTCGCACTCCTATTGTGTCCACCAGGATGTAATGAA
GTTGGCCTATGCAGACACTTTGCCCAATGTGGTATATGGTCTTACTGCCATTCTGCTGGTC
ATGGGCGTGGACGTAATGTTCATCTCCTTGTCCTATTTTCTGATAATACGAACGGTTCTGC
AACTGCCTTCCAAGTCAGAGCGGGCCAAGGCCTTTGGAACCTGTGTGTCACACATTGGTGT
GGTACTCGCCTTCTATGTGCCACTTATTGGCCTCTCAGTTGTACACCGCTTTGGAAACAGC
CTTCATCCCATTGTGCGTGTTGTCATGGGTGACATCTACCTGCTGCTGCCTCCTGTCATCA
ATCCCATCATCTATGGTGCCAAAACCAAACAGATCAGAACACGGGTGCTGGCTATGTTCAA
GATCAGCTGTGACAAGGACTTGCAGGCTGTGGGAGGCAAGTGACCCTTAACACTACACTTC
TCCTTATCTTTATTGGCTTGATAAACATAATTATTTCTAACACTAGCTTATTTCCAGTTGC
CCATAAGCACATCAGTACTTTTCTCTGGCTGGAATAGTAAACTAAAGTATGGTACATCTAC
CTAAAGGACTATTATGTGGAATAATACATACTAATGAAGTATTACATGATTTAAAGACTAC
AATAAAACCAAACATGCTTATAACATTAAGAAAAACAATAAAGATACATGATTGAAACCAA
GTTGAAAAATAGCATATGCCTTGGAGGAAATGTGCTCAAATTACTAATGATTTAGTGTTGT
CCCTACTTTCTCTCTCTTTTTTCTTTCTTTTTTTTTTATTATGGTTAGCTGTCACATACAA
CTTTTTTTTTTTTTGAGATGGGGTCTCCAGCCTGGGCAACAGAGCAAGACCCTGTCTCAAA
GCATAAAATGGAATAACATATCAAATGAAACAGGGAAAATGAAGCTGACAATTTATGGGAG
CCA

FIGURE 32A

CACACATACGCACGCACGATCTCACTTCGATCTATACACTGGAGGATTAAAACAAACAAAC
AAAAAAAACATTTCCTTCGCTCCCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCCGCACG
GCGAGGGGCCGCAGACCGTCTGGAAATGCGAATCCTAAAGCGTTTCCTCGCTTGCATTCAG
CTCCTCTGTGTTTGCCGCCTGGATTGGGCTAATGGATACTACAGACAACAGAGAAAACTTG
TTGAAGAGATTGGCTGGTCCTATACAGGAGCACTGAATCAAAAAAATTGGGGAAAGAAATA
TCCAACATGTAATAGCCCAAAACAATCTCCTATCAATATTGATGAAGATCTTACACAAGTA
AATGTGAATCTTAAGAAACTTAAATTTCAGGGTTGGGATAAAACATCATTGGAAAACACAT
TCATTCATAACACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTCAGCGG
AGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAACTTTTCACTGGGGAAAATGCAAT
ATGTCATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAA
TCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTGAGGAAGCAGTCAAAGGAAAAGGGAA
GTTAAGAGCTTTATCCATTTTGTTTGAGGTTGGGACAGAAGAAAATTTGGATTTCAAAGCG
ATTATTGATGGAGTCGAAAGTGTTAGTCGTTTTGGGAAGCAGGCTGCTTTAGATCCATTCA
TACTGTTGAACCTTCTGCCAAACTCAACTGACAAGTATTACATTTACAATGGCTCATTGAC
ATCTCCTCCCTGCACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCT
GAAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAATGCAACAATCTGGTTATGTCATGC
TGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTT
TTCCTCATACACTGGAAAGGAAGAGATTCATGAAGCAGTTTGTAGTTCAGAACCAGAAAAT
GTTCAGGCTGACCCAGAGAATTATACCAGCCTTCTTGTTACATGGGAAAGACCTCGAGTCG
TTTATGATACCATGATTGAGAAGTTTGCAGTTTTGTACCAGCAGTTGGATGGAGAGGACCA
AACCAAGCATGAATTTTTGACAGATGGCTATCAAGACTTGGGTGCTATTCTCAATAATTTG
CTACCCAATATGAGTTATGTTCTTCAGATAGTAGCCATATGCACTAATGGCTTATATGGAA
AATACAGCGACCAACTGATTGTCGACATGCCTACTGATAATCCTGAACTTGATCTTTTCCC
TGAATTAATTGGAACTGAAGAAATAATCAAGGAGGAGGAAGAGGGAAAAGACATTGAAGAA
GGCGCTATTGTGAATCCTGGTAGAGACAGTGCTACAAACCAAATCAGGAAAAAGGAACCCC
AGATTTCTACCACAACACACTACAATCGCATAGGGACGAAATACAATGAAGCCAAGACTAA
CCGATCCCCAACAAGAGGAAGTGAATTCTCTGGAAAGGGTGATGTTCCCAATACATCTTTA
AATTCCACTTCCCAACCAGTCACTAAATTAGCCACAGAAAAAGATATTTCCTTGACTTCTC
AGACTGTGACTGAACTGCCACCTCACACTGTGGAAGGTACTTCAGCCTCTTTAAATGATGG
CTCTAAAACTGTTCTTAGATCTCCACATATGAACTTGTCGGGGACTGCAGAATCCTTAAAT
ACAGTTTCTATAACAGAATATGAGGAGGAGAGTTTATTGACCAGTTTCAAGCTTGATACTG
GAGCTGAAGATTCTTCAGGCTCCAGTCCCGCAACTTCTGCTATCCCATTCATCTCTGAGAA
CATATCCCAAGGGTATATATTTTCCTCCGAAAACCCAGAGACAATAACATATGATGTCCTT
ATACCAGAATCTGCTAGAAATGCTTCCGAAGATTCAACTTCATCAGGTTCAGAAGAATCAC
TAAAGGATCCTTCTATGGAGGGAAATGTGTGGTTTCCTAGCTCTACAGACATAACAGCACA

GCCCGATGTTGGATCAGGCAGAGAGAGCTTTCTCCAGACTAATTACACTGAGATACGTGTT
GATGAATCTGAGAAGACAACCAAGTCCTTTTCTGCAGGCCCAGTGATGTCACAGGGTCCCT
CAGTTACAGATCTGGAAATGCCACATTATTCTACCTTTGCCTACTTCCCAACTGAGGTAAC
ACCTCATGCTTTTACCCCATCCTCCAGACAACAGGATTTGGTCTCCACGGTCAACGTGGTA
TACTCGCAGACAACCCAACCGGTATACAATGGTGAGACACCTCTTCAACCTTCCTACAGTA
GTGAAGTCTTTCCTCTAGTCACCCCTTTGTTGCTTGACAATCAGATCCTCAACACTACCCC
TGCTGCTTCAAGTAGTGATTCGGCCTTGCATGCTACGCCTGTATTTCCCAGTGTCGATGTG
TCATTTGAATCCATCCTGTCTTCCTATGATGGTGCACCTTTGCTTCCATTTTCCTCTGCTT
CCTTCAGTAGTGAATTGTTTCGCCATCTGCATACAGTTTCTCAAATCCTTCCACAAGTTAC
TTCAGCTACCGAGAGTGATAAGGTGCCCTTGCATGCTTCTCTGCCAGTGGCTGGGGGTGAT
TTGCTATTAGAGCCCAGCCTTGCTCAGTATTCTGATGTGCTGTCCACTACTCATGCTGCTT
CAGAGACGCTGGAATTTGGTAGTGAATCTGGTGTTCTTTATAAAACGCTTATGTTTTCTCA
AGTTGAACCACCCAGCAGTGATGCCATGATGCATGCACGTTCTTCAGGGCCTGAACCTTCT
TATGCCTTGTCTGATAATGAGGGCTCCCAACACATCTTCACTGTTTCTTACAGTTCTGCAA
TACCTGTGCATGATTCTGTGGGTGTAACTTATCAGGGTTCCTTATTTAGCGGCCCTAGCCA
TATACCAATACCTAAGTCTTCGTTAATAACCCCAACTGCATCATTACTGCAGCCTACTCAT
GCCCTCTCTGGTGATGGGGAATGGTCTGGAGCCTCTTCTGATAGTGAATTTCTTTTACCTG
ACACAGATGGGCTGACAGCCCTTAACATTTCTTCACCTGTTTCTGTAGCTGAATTTACATA
TACAACATCTGTGTTTGGTGATGATAATAAGGCGCTTTCTAAAAGTGAAATAATATATGGA
AATGAGACTGAACTGCAAATTCCTTCTTTCAATGAGATGGTTTACCCTTCTGAAAGCACAG
TCATGCCCAACATGTATGATAATGTAAATAAGTTGAATGCGTCTTTACAAGAAACCTCTGT
TTCCATTTCTAGCACCAAGGGCATGTTTCCAGGGTCCCTTGCTCATACCACCACTAAGGTT
TTTGATCATGAGATTAGTCAAGTTCCAGAAAATAACTTTTCAGTTCAACCTACACATACTG
TCTCTCAAGCATCTGGTGACACTTCGCTTAAACCTGTGCTTAGTGCAAACTCAGAGCCAGC
ATCCTCTGACCCTGCTTCTAGTGAAATGTTATCTCCTTCAACTCAGCTCTTATTTTATGAG
ACCTCAGCTTCTTTTAGTACTGAAGTATTGCTACAACCTTCCTTTCAGGCTTCTGATGTTG
ACACCTTGCTTAAAACTGTTCTTCCAGCTGTGCCCAGTGATCCAATATTGGTTGAAACCCC
CAAAGTTGATAAAATTAGTTCTACAATGTTGCATCTCATTGTATCAAATTCTGCTTCAAGT
GAAAACATGCTGCACTCTACATCTGTACCAGTTTTTGATGTGTCGCCTACTTCTCATATGC
ACTCTGCTTCACTTCAAGGTTTGACCATTTCCTATGCAAGTGAGAAATATGAACCAGTTTT
GTTAAAAAGTGAAAGTTCCCACCAAGTGGTACCTTCTTTGTACAGTAATGATGAGTTGTTC
CAAACGGCCAATTTGGAGATTAACCAGGCCCATCCCCCAAAAGGAAGGCATGTATTTGCTA
CACCTGTTTTATCAATTGATGAACCATTAAATACACTAATAAATAAGCTTATACATTCCGA
TGAAATTTTAACCTCCACCAAAAGTTCTGTTACTGGTAAGGTATTTGCTGGTATTCCAACA
GTTGCTTCTGATACATTTGTATCTACTGATCATTCTGTTCCTATAGGAAATGGGCATGTTG
CCATTACAGCTGTTTCTCCCCACAGAGATGGTTCTGTAACCTCAACAAAGTTGCTGTTTCC
TTCTAAGGCAACTTCTGAGCTGAGTCATAGTGCCAAATCTGATGCCGGTTTAGTGGGTGGT

GGTGAAGATGGTGACACTGATGATGATGGTGATGATGATGATGACAGAGATAGTGATGGCT
TATCCATTCATAAGTGTATGTCATGCTCATCCTATAGAGAATCACAGGAAAAGGTAATGAA

FIGURE 32B

TGATTCAGACACCCACGAAAACAGTCTTATGGATCAGAATAATCCAATCTCATACTCACTA
TCTGAGAATTCTGAAGAAGATAATAGAGTCACAAGTGTATCCTCAGACAGTCAAACTGGTA
TGGACAGAAGTCCTGGTAAATCACCATCAGCAAATGGGCTATCCCAAAAGCACAATGATGG
AAAAGAGGAAAATGACATTCAGACTGGTAGTGCTCTGCTTCCTCTCAGCCCTGAATCTAAA
GCATGGGCAGTTCTGACAAGTGATGAAGAAAGTGGATCAGGGCAAGGTACCTCAGATAGCC
TTAATGAGAATGAGACTTCCACAGATTTCAGTTTTGCAGACACTAATGAAAAAGATGCTGA
TGGGATCCTGGCAGCAGGTGACTCAGAAATAACTCCTGGATTCCCACAGTCCCCAACATCA
TCTGTTACTAGCGAGAACTCAGAAGTGTTCCACGTTTCAGAGGCAGAGGCCAGTAATAGTA
GCCATGAGTCTCGTATTGGTCTAGCTGAGGGGTTGGAATCCGAGAAGAAGGCAGTTATACC
CCTTGTGATCGTGTCAGCCCTGACTTTTATCTGTCTAGTGGTTCTTGTGGGTATTCTCATC
TACTGGAGGAAATGCTTCCAGACTGCACACTTTTACTTAGAGGACAGTACATCCCCTAGAG
TTATATCCACACCTCCAACACCTATCTTTCCAATTTCAGATGATGTCGGAGCAATTCCAAT
AAAGCACTTTCCAAAGCATGTTGCAGATTTACATGCAAGTAGTGGGTTTACTGAAGAATTT
GAGACACTGAAAGAGTTTTACCAGGAAGTGCAGAGCTGTACTGTTGACTTAGGTATTACAG
CAGACAGCTCCAACCACCCAGACAACAAGCACAAGAATCGATACATAAATATCGTTGCCTA
TGATCATAGCAGGGTTAAGCTAGCACAGCTTGCTGAAAAGGATGGCAAACTGACTGATTAT
ATCAATGCCAATTATGTTGATGGCTACAACAGACCAAAAGCTTATATTGCTGCCCAAGGCC
CACTGAAATCCACAGCTGAAGATTTCTGGAGAATGATATGGGAACATAATGTGGAAGTTAT
TGTCATGATAACAAACCTCGTGGAGAAAGGAAGGAGAAAATGTGATCAGTACTGGCCTGCC
GATGGGAGTGAGGAGTACGGGAACTTTCTGGTCACTCAGAAGAGTGTGCAAGTGCTTGCCT
ATTATACTGTGAGGAATTTTACTCTAAGAAACACAAAAATAAAAAAGGGCTCCCAGAAAGG
AAGACCCAGTGGACGTGTGGTCACACAGTATCACTACACGCAGTGGCCTGACATGGGAGTA
CCAGAGTACTCCCTGCCAGTGCTGACCTTTGTGAGAAAGGCAGCCTATGCCAAGCGCCATG
CAGTGGGGCCTGTTGTCGTCCACTGCAGTGCTGGAGTTGGAAGAACAGGCACATATATTGT
GCTAGACAGTATGTTGCAGCAGATTCAACACGAAGGAACTGTCAACATATTTGGCTTCTTA
AAACACATCCGTTCACAAAGAAATTATTTGGTACAAACTGAGGAGCAATATGTCTTCATTC
ATGATACACTGGTTGAGGCCATACTTAGTAAAGAAACTGAGGTGCTGGACAGTCATATTCA
TGCCTATGTTAATGCACTCCTCATTCCTGGACCAGCAGGCAAAACAAAGCTAGAGAAACAA
TTCCAGCTCCTGAGCCAGTCAAATATACAGCAGAGTGACTATTCTGCAGCCCTAAAGCAAT
GCAACAGGGAAAAGAATCGAACTTCTTCTATCATCCCTGTGGAAAGATCAAGGGTTGGCAT
TTCATCCCTGAGTGGAGAAGGCACAGACTACATCAATGCCTCCTATATCATGGGCTATTAC
CAGAGCAATGAATTCATCATTACCCAGCACCCTCTCCTTCATACCATCAAGGATTTCTGGA
GGATGATATGGGACCATAATGCCCAACTGGTGGTTATGATTCCTGATGGCCAAAACATGGC
AGAAGATGAATTTGTTTACTGGCCAAATAAAGATGAGCCTATAAATTGTGAGAGCTTTAAG

GTCACTCTTATGGCTGAAGAACACAAATGTCTATCTAATGAGGAAAAACTTATAATTCAGG
ACTTTATCTTAGAAGCTACACAGGATGATTATGTACTTGAAGTGAGGCACTTTCAGTGTCC
TAAATGGCCAAATCCAGATAGCCCCATTAGTAAAACTTTTGAACTTATAAGTGTTATAAAA
GAAGAAGCTGCCAATAGGGATGGGCCTATGATTGTTCATGATGAGCATGGAGGAGTGACGG
CAGGAACTTTCTGTGCTCTGACAACCCTTATGCACCAACTAGAAAAAGAAAATTCCGTGGA
TGTTTACCAGGTAGCCAAGATGATCAATCTGATGAGGCCAGGAGTCTTTGCTGACATTGAG
CAGTATCAGTTTCTCTACAAAGTGATCCTCAGCCTTGTGAGCACAAGGCAGGAAGAGAATC
CATCCACCTCTCTGGACAGTAATGGTGCAGCATTGCCTGATGGAAATATAGCTGAGAGCTT
AGAGTCTTTAGTT**TAA**CACAGAAAGGGGTGGGGGGACTCACATCTGAGCATTGTTTTCCTC
TTCCTAAAATTAGGCAGGAAAATCAGTCTAGTTCTGTTATCTGTTGATTTCCCATCACCTG
ACAGTAACTTTCATGACATAGGATTCTGCCGCCAAATTTATATCATTAACAATGTGTGCCT
TTTTGCAAGACTTGTAATTTACTTATTATGTTTGAACTAAAATGATTGAATTTTACAGTAT
TTCTAAGAATGGAATTGTGGTATTTTTTTCTGTATTGATTTTAACAGAAAATTTCAATTTA
TAGAGGTTAGGAATTCCAAACTACAGAAAATGTTTGTTTTTAGTGTCAAATTTTTAGCTGT
ATTTGTAGCAATTATCAGGTTTGCTAGAAATATAACTTTTAATACAGTAGCCTGTAAATAA
AACACTCTTCCATATGATATTCAACATTTTACAACTGCAGTATTCACCTAAAGTAGAAATA
ATCTGTTACTTATTGTAAATACTGCCCTAGTGTCTCCATGGACCAAATTTATATTTATAAT
TGTAGATTTTTATATTTTACTACTGAGTCAAGTTTTCTAGTTCTGTGTAATTGTTTAGTTT
AATGACGTAGTTCATTAGCTGGTCTTACTCTACCAGTTTTCTGACATTGTATTGTGTTACC
TAAGTCATTAACTTTGTTTCAGCATGTAATTTTAACTTTTGTGGAAAATAGAAATACCTTC
ATTTTGAAAGAAGTTTTTATGAGAATAACACCTTACCAAACATTGTTCAAATGGTTTTTAT
CCAAGGAATTGCAAAAATAAATATAAATATTGCCATTAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

FIGURE 33

ATTGTCTGGGAATTGCAGCCGCGGGGCGGGCGGCGGCGGCGGCGGCGGCCGGGACCCA
GCGGGCCAGGTGGGGACGGCGCGGAGCGGGTGCGGGAGATGCCGTGCGGGACTGGGGCCAC
CTGAGCCGCCCGCCTCGTCCCCGCCTTCTGTGGGAAGG**ATG**TGCGCGCGGATGGCCGGTCG
CACAAGAGCGGCCCCTCGGGGGCCCTACGGCCCCTGGCTCTGCCTCCTGGTGGCCCTCGCC
CTGGACGTCGTGAGAGTGGACTGTGGCCAGGCTCCCCTGGACCCTGTCTACCTGCATGTGA
CAGCCGCCCGCCCAGCCCAGCCCACACTCTGGACTGCCAAGCTAGACCGCTTCAAGGGCTC
CAGGCACCACACCACCCTCATCACCTGCCACCGTGCTGGGCTCACAGAGCCAGATTCCAGC
AGTCCCCTTGAACTGTCTGAGTTCCTATGGGTGGACTTTGTGGTGGAGAATAGCACTGGTG
GGGGCGTAGCGGTCACTCGCCCCGTCACGTGGCAGCTGGAGTACCCAGGCCAGGCCCCTGA
AGCAGAGAAGGACAAAATGGTGTGGGAAATCCTGGTGTCTGAGCGGGACATCAGAGCCCTT
ATCCCACTGGCCAAGGCTGAGGAGCTGGTGAATACAGCACCACTGACTGGAGTGCCCCAGC
ATGTCCCCGTGCGCCTTGTCACTGTGGACGGCGGGGGGGCCTTGGTGGAGGTGACAGAGCA
TGTCGGCTGCGAGTCTGCCAACACACAGGTCCTGCAGGTGTCTGAGGCCTGTGATGCCGTG
TTCGTGGCTGGCAAGGAGAGCCGGGGCGCCCGGGGGGTGCGAGTGGACTTCTGGTGGCGCC
GGCTCCGCGCCTCGCTGCGGCTGACCGTGTGGGCCCCGCTGCTACCGCTGCGTATCGAGCT
CACCGACACCACCCTCGAGCAGGTCCGCGGCTGGAGGGTACCTGGCCCTGCTGAAGGGCCT
GCGGAACCCGCTGCAGAGGCGTCAGATGAGGCCGAGCGGCGCGCCCGTGGCTGCCACCTGC
AGTACCAGCGGGCCGGTGTGCGCTTCCTCGCCCCCTTCGCGGCCCACCCGCTGGACGGCGG
CCGCCGCCTCACGCACCTGCTTGGCCCCGACTGGCTGCTAGACGTGTCCCACCTCGTGGCG
CCACACGCCCGCGTGCTGGACTCGCGTGTAGCCTCTCTGGAGGGTGGCCGTGTCGTGGTGG
GCCGGGAGCCCGGTGTCACCTCCATTGAGGTGCGTTCCCCACTGTCTGACTCCATCCTGGG
GGAGCAGGCGCTGGCTGTGACGGACGACAAGGTCTCAGTGCTGGAGCTGAGGGTGCAGCCA
GTGATGGGCATCTCGCTGACCTTGAGCCGGGGCACTGCCCACCCCGGGGAGGTCACAGCTA
CGTGCTGGGCACAGTCAGCCCTTCCCGCCCCAAAGCAGGAGGTGGCCCTCTCCCTATGGCT
GTCCTTCTCTGATCACACTGTGGCCCCAGCTGAGCTCTACGACCGCCGTGACCTGGGACTG
TCCGTCTCAGCCGAGGAGCCTGGTGCCATCCTGCCAGCTGAGGAGCAGGGTGCCCAGCTCG
GGGTGGTGGTGAGTGGGGCAGGCGCCGAGGGGCTGCCGCTGCATGTGGCTCTGCACCCGCC
CGAGCCCTGCCGCCGGGGCCGCCACCGTGTGCCTCTGGCCTCTGGCACCGCCTGGCTGGGG
CTGCCCCCTGCCTCCACTCCAGCCCCTGCTCTCCCATCCAGCCCTGCTTGGAGCCCACCAG
CCACAGAAGCCACCATGGGTGGTAAACGGCAGGTGGCAGGCAGTGTCGGGGGCAACACAGG
TGTGAGGGGCAAGTTTGAGCGGGCAGAGGAGGAGGCCAGGAAGGAGGAGACCGAAGCCAGG
GAGGAGGAGGAGGAAGAGGAGGAGGAGATGGTCCCTGCCCCTCAGCATGTCACTGAGCTAG
AGCTGGGCATGTACGCCCTGCTGGGAGTCTTCTGCGTGGCCATCTTCATCTTCTTGGTCAA
TGGTGTGGTCTTCGTCCTGCGCTATCAGCGCAAAGAACCTCCCGACAGTGCCACTGACCCC
ACCTCCCCCCAGCCCCACAACTGGGTCTGGCTGGGCACTGACCAGGAGGAACTGAGCCGCC

AGCTGGACCGGCAGTCCCCTGGCCCGCCCAAGGGGGAGGGGAGCTGCCCCTGTGAGAGTGG
GGGAGGAGGGGAGGCCCCTACCCTGGCCCCTGGCCCTCCTGGGGGCACCACCAGCTCCTCA
AGCACCCTGGCCCGAAAGGAGGCTGGGGGGCGGCGGAAGCGAGTAGAGTTTGTGACATTTG
TGCCAGCCCCTCCAGCCCAGTCACCTGAGGAGCCTGTAGGGGCCCCTGCTGTGCAGTCCAT
CCTTGTGGCAGGCGAGGAGGACATCCGCTGGGTGTGTGAGGACATGGGGCTGAAGGACCCT
GAGGAGCTTCGCAACTACATGGAGAGGATCCGGGGCAGCTCCTGACCCTCCACAGCCACCT
GGTCAGCCACCAGCTGGGGCAACGAGGGTGGAGGTCCCACTGAGCCTCTCGCCTGCCCCCG
CCACTCGTCTGGTGCTTGTTGATCCAAGTCCCCTGCCTGGTCCCCCACAAGGACTCCCATC
CAGGCCCCCTCTGCCCTGCCCCTTGTCATGGACCATGGTCGTGAGGAAGGGCTCATGCCCC
TTATTTATGGGAACCATCTCATTCTAACAGAATAAACCGAGAAGGAAACCAGAAAAAAAAA
AAAAAAAAAA

FIGURE 34

ACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGGTCGGGGCCAGAGCGCAGGTGTACCTGG
CGGCCGTGCTGGAGCACCTGACCGCCGAGATCCTGGAGCTGGCTGGCAACCCGGCCCGCGA
CAAGAAGACCCGCATCATCCTGCGCCACCTGTAGCTGGCCATTCGCAACGGCGAGGAGCTT
AACAAGCTGCTGGGCGAAGTCACCATCGCGCAGGGCGGTGTCCTGCCCAACATTCAGGGCG
TGCTTCTGCCCCAGAAGACCAAGAGCCACCACAAGGCCAAGGGTGAAAACCATTCACTAGG
AGAGGAGAAACACA**ATG**GCCACCAAGACAGAGTTGAGTCCCACAGCAAGGGAGAGCAAGAA
CGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAAAGGTCCAAGTTTATGTTCT
GCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTCACAACGATAGCACAAAATG
TCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAAGCCCTCAATCCCAGCTCAA
TGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCTAAGTAAAGCCCCAAAGAGT
CTTCCTGCAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATTTGGGAAAAGTGGGGCCCTC
CACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAATGTTACTGGGATGCTTTAC
TGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTGGCCCTTTGTCTTCTATATC
TTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTTGTGATTTATGATGACCCCT
TTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCATATCCTCCTTGAAACAACA
GGTCGGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCTCAGATCTCTACCCATTTGG
TCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGCACAATGGTTGTATACATAC
CAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACAATGGACTTCTATCTGCCCT
TCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTATCTGGCAGATTTCCTTCTA
ACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACAATTTTAGGAAGTCTCCCCT
CTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCTATATCACAGCAACTGCCTT
GCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGGGATTTATATCAATGTCTTA
GATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCAAGAGGATTTTCGAGCATAG
CACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTCAGGACCCTGAGTTTGGGTG
GAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGGACTACTCTTCTACCTCATA
TTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAACTCACTCGTTTA**TGA**AGTT
ATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCATAAAATAGAAGGCTTCCGTG
ATGAAAATACCAGTGAAAAGATTTTTTTTTCCTGTGGCTCTTTTCAATTATGAGATCAGTT
CATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATGAATAAAAATTCAAATAAAA
TGATAACTAAGAAAAAAAAAAAAAA

FIGURE 35

GTGAAAACCATTCACTAGGAGAGGAGAAACACAATGGCCACCAAGACAGAGTTGAGTCCCA
CAGCAAGGGAGAGCAAGAACGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAA
AGTTCCAAGTTTATGTTCTGCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTC
ACAACGATAGCACAAAATGTCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAA
GCCCTCAATCCCAGCTCAATGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCT
AAGTAAAGCCCCAAAGAGTCTTCCTACAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATT
TGGGAAAAGTGGGGCCCTCCACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAA
TGTTACTGGGATGCTTTACTGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTG
GCCCTTTGTCTTCTATATCTTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTT
GTGATTTATGATGACCCCGTTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCA
TATCCTCCTTGAAACAACAGGTCAGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCT
CAGATCTCTACCCATTTGGTCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGC
ACAATGGTTGTATACATACCAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACA
ATGGACTTCTATCTGCCCTTCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTA
TCTGGCAGATTTCCTTCTAACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACA
ATTTTAGGAAGTCTCCCCTCTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCT
ATATCACAGCAACTGCCTTGCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGG
GATTTATATCAATGTCTTAGATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCA
AGAGGATTTTCGAGCATAGCACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTC
AGGACCCTGAGTTTGGGTGGAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGG
ACTACTCTTCTACCTCATATTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAA
CTCACTCGTTTATGAAGTTATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCAT
AAAATAGAAGGCTTCCGTGATGAAAATACCAGTGAAAAGATTTTTTTTCCTGTGGCTCTT
TTCAATTATGAGATCAGTTCATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATG
AATAAAAATTCAAATAAAATGATAACTAAGAAAAAAAAAAAAAA

FIGURE 36

GCTCGCTGGGCCGCTGCTCCCGGGTGTCCCAGGCCCGGCCGGTGCGCAGAGCATGGCGGGTGCGGGCCCGAAGCGGCGCGC
GCTAGCGGCGCCGGCGGCCGAGGAGAAGGAAGAGGCGCGGGAGAAGATGCTGGCCGCCAAGAGCGCGGACGGCTCGGCGCCG
GCAGGCGAGGGCGAGGGCGTGACCCTGCAGCGGAACATCACGCTGCTCAACGGCGTGGCCATCATCGTGGGGACCATTATCG
GCTCGGGCATCTTCGTGACGCCCACGGGCGTGCTCAAGGAGGCAGGCTCGCCGGGGCTGGCGCTGGTGGTGTGGGCCGCGTG
CGGCGTCTTCTCCATCGTGGGCGCGCTCTGCTACGCGGAGCTCGGCACCACCATCTCCAAATCGGGCGGCGACTACGCCTAC
ATGCTGGAGGTCTACGGCTCGCTGCCCGCCTTCCTCAAGCTCTGGATCGAGCTGCTCATCATCCGGCCTTCATCGCAGTACA
TCGTGGCCCTGGTCTTCGCCACCTACCTGCTCAAGCCGCTCTTCCCCACCTGCCCGGTGCCCGAGGAGGCAGCCAAGCTCGT
GGCCTGCCTGTGCGTGCTGCTGCTCACGGCCGTGAACTGCTACAGCGTGAAGGCCGCCACCCGGGTCCAGGATGCCTTTGCC
GCCGCCAAGCTCCTGGCCCTGGCCCTGATCATCCTGCTGGGCTTCGTCCAGATCGGAAAGGGTGATGTGTCCAATCTAGATC
CCAAGTTCTCATTTGAAGGCACCAAACTGGATGTGGGGAACATTGTGCTGGCATTATACAGCGGCCTCTTTGCCTATGGAGG
ATGGAATTACTTGAATTTCGTCACAGAGGAAATGATCAACCCCTACAGAAACCTGCCCCTGGCCATCATCATCTCCCTGCCC
ATCGTGACGCTGGTGTACGTGCTGACCAACCTGGCCTACTTCACCACCCTGTCCACCGAGCAGATGCTGTCGTCCGAGGCCG
TGGCCGTGGACTTCGGGAACTATCACCTGGGCGTCATGTCCTGGATCATCCCCGTCTTCGTGGGCCTGTCCTGCTTTGGCTC
CGTCAATGGGTCCCTGTTCACATCCTCCAGGCTCTTCTTCGTGGGGTCCCGGGAAGGCCACCTGCCCTCCATCCTCTCCATG
ATCCACCCACAGCTCCTCACCCCCGTGCCGTCCCTCGTGTTCACGTGTGTGATGACGCTGCTCTACGCCTTCTCCAAGGACA
TCTTCTCCGTCATCAACTTCTTCAGCTTCTTCAACTGGCTCTGCGTGGCCCTGGCCATCATCGGCATGATCTGGCTGCGCCA
CAGAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGAACCTGGCCCTGCCTGTGTTCTTCATCCTGGCCTGCCTCTTCCTGATC
GCCGTCTCCTTCTGGAAGACACCCGTGGAGTGTGGCATCGGCTTCACCATCATCCTCAGCGGGCTGCCCGTCTACTTCTTCG
GGGTCTGGTGGAAAAACAAGCCCAAGTGGCTCCTCCAGGGCATCTTCTCCACGACCGTCCTGTGTCAGAAGCTCATGCAGGT
GGTCCCCCAGGAGACATAGCCAGGAGGCCGAGTGGCTGCCGGAGGAGCATGCGCAGAGGCCAGTTAAAGTAGATCACCTCCT
CGAACCCACTCCGGTTCCCCGCAACCCACAGCTCAGCTGCCCATCCCAGTCCTCGCCGTCCCTCCCAGGTCGGGCAGTGGAG
GCTGCTGTGAAAACTCTGGTACGAATCTCATCCCTCAACTGAGGGCCAGGGACCCAGGTGTGCCTGTGCTCCTGCCCAGGAG
CAGCTTTTGGTCTCCTTGGGCCCTTTTTCCCTTCCCTCCTTTGTTTACTTATATATATATTTTTTTTAAACTTAAATTTTGG
GTCAACTTGACACCACTAAGATGATTTTTTAAGGAGCTGGGGGAAGGCAGGAGCCTTCCTTTCTCCTGCCCCAAGGGCCCAG
ACCCTGGGCAAACAGAGCTACTGAGACTTGGAACCTCATTGCTACCACAGACTTGCACTGAAGCCAGACAGCTGCCCAGACA
CATGGGCTTGTGACATTCGTGAAAACCAACCCTGTGGGCTTATGTCTCTGCCTTAGGGTTTGCAGAGTGGAAACTCAGCCGT
AGGGTGGCACTGGGAGGGGGTGGGGGGATCTGGGCAAGGTGGGTGATTCCTCCCAGGAGGTGCTTGAGGCCCCGATGGACTCC
TGACCATAATCCTAGCCCCGAGACACCATCCTGAGCCAGGGAACAGCCCCAGGGTTGGGGGGTGCCGGCATCTCCCCTAGCT
CACCAGGCCTGGCCTCTGGGCAGTGTGGCCTCTTGGCTATTTCTGTTCCAGTTTTGGAGGCTGAGTTCTGGTTCATGCAGAC
AAAGCCCTGTCCTTCAGTCTTCTAGAAACAGAGACAAGAAAGGCAGACACACCGCGGCCAGGCACCCATGTGGGCGCCCACC
CTGGGCTCCACACAGCAGTGTCCCCTGCCCCAGAGGTCGCAGCTACCCTCAGCCTCCAATGCATTGGCCTCTGTACCGCCCG
GCAGCCCCTTCTGGCCGGTGCTGGGTTCCCACTCCCGGCCTAGGCACCTCCCCGCTCTCCCTGTCACGCTCATGTCCTGTCC
TGGTCCTGATGCCCGTTGTCTAGGAGACAGAGCCAAGCACTGCTCACGTCTCTGCCGCCTGCGTTTGGAGGCCCCTGGGCTC
TCACCCAGTCCCCACCCGCCTGCAGAGAGGGAACTAGGGCACCCCTTGTTTCTGTTGTTCCCGTGAATTTTTTTCGCTATGG
GAGGCAGCCGAGGCCTGGCCAATGCGGCCCACTTTCCTGAGCTGTCGCTGCCTCCATGGCAGCAGCCAAGGACCCCCAGAAC
AAGAAGACCCCCCCGCAGGATCCCTCCTGAGCTCGGGGGGCTCTGCCTTCTCAGGCCCCGGGCTTCCCTTCTCCCCAGCCAG
AGGTGGAGCCAAGTGGTCCAGCGTCACTCCAGTGCTCAGCTGTGGCTGGAGGAGCTGGCCTGTGGCACAGCCCTGAGTGTCC
CAAGCCGGGAGCCAACGAAGCCGGACACGGCTTCACTGACCAGCGGCTGCTCAAGCCGCAAGCTCTCAGCAAGTGCCCAGTG
GAGCCTGCCGCCCCCACCTGGGCACCGGGACCCCCTCACCATCCAGTGGGCCCGGAGAAACCTGATGAACAGTTTGGGGACT
CAGGACCAGATGTCCGTCTCTCTTGCTTGAGGAATGAAGACCTTTATTCACCCCTGCCCCGTTGCTTCCCGCTGCACATGGA
CAGACTTCACAGCGTCTGCTCATAGGACCTGCATCCTTCCTGGGGACGAATTCCACTCGTCCAAGGGACAGCCCACGGTCTG
GAGGCCGAGGACCACCAGCAGGCAGGTGGACTGACTGTGTTGGGCAAGACCTCTTCCCTCTGGGCCTGTTCTCTTGGCTGCA
AATAAGGACAGCAGCTGGTGCCCCACCTGCCTGGTGCATTGCTGTGTGAATCCAGGAGGCAGTGGACATCGTAGGCAGCCAC
GGCCCCAGGTCCAGGAGAAGTGCTCCCTGGAGGCACGGACCACTGCTTCCCACTGGGGCCGGCGGGGCCCACGCACGACGTC
AGCCTCTTACCTTCCCGCCTCGGCTAGGGGTCCTCGGGATGCCGTTCTGTTCCAACCTCCTGTTCTGGGAGGTGGACATGCC
TCAAGGATACAGGGAGCCGGCGGCCTCTCGACGGCACGCACTTCCTGTTGGCTGCTGCGGCTGTGGGCGAGCATGGGGGCTG
CCAGCGTCTGTTGTGGAAAGTAGCTGCTAGTGAAATGGCTGGGGCCGCTGGGGTCCGTCTTCACACTGCGCAGGTCTCTTCT

GGGCGTCTGAGCTGGGGTGGGAGCTCCTCCGCAGAAGGTTGGTGGGGGGTCCAGTCTGTGATCCTTGGTGCTGTGTGCCCCA
CTCCAGCCTGGGGACCCCACTTCAGAAGGTAGGGGCCGTGTCCCGCGGTGCTGACTGAGGCCTGCTTCCCCCTCCCCCTCCT
GCTGTGCTGGAATTCCACAGGGACCAGGGCCACCGCAGGGGACTGTCTCAGAAGACTTGATTTTTCCGTCCCTTTTTCTCCA
CACTCCACTGACAAACGTCCCCAGCGGTTTCCACTTGTGGGCTTCAGGTGTTTTCAAGCACAACCCACCACAACAAGCAAGT
GCATTTTCAGTCGTTGTGCTTTTTTGTTTTGTGCTAACGTCTTACTAATTTAAAGATGCTGTCGGCACCATGTTTATTTATT
TCCAGTGGTCATGCTCAGCCTTGCTGCTCTGCGTGGCGCAGGTGCCATGCCTGCTCCCTGTCTGTGTCCCAGCCACGCAGGG
CCATCCACTGTGACGTCGGCCGACCAGGCTGGACACCCTCTGCCGAGTAATGACGTGTGTGGCTGGGACCTTCTTTATTCTG
TGTTAATGGCTAACCTGTTACACTGGGCTGGGTTGGGTAGGGTGTTCTGGCTTTTTTGTGGGGTTTTTATTTTTAAAGAAAC
ACTCAATCATCCTAG

FIGURE 37

ATTCGCCCTTTGCTCCCGGGTGTCCCAGGCCCGGCCGGTGCGCAGAGC**ATG**GCGGGTGCGG
GCCCGAAGCGGCGCGCGCTAGCGGCGCCGGCGGCCGAGGAGAAGGAAGAGGCGCGGGAGAA
GATGCTGGCCGCCAAGAGCGCGGACGGCTCGGCGCCGGCAGGCGAGGGCGAGGGCGTGACC
CTGCAGCGGAACATCACGCTGCTCAACGGCGTGGCCATCATCGTGGGGACCATTATCGGCT
CGGGCATCTTCGTGACGCCCACGGGCGTGCTCAAGGAGGCAGGCTCGCCGGGGCTGGCGCT
GGTGGTGTGGGCCGCGTGCGGCGTCTTCTCCATCGTGGGCGCGCTCTGCTACGCGGAGCTC
GGCACCACCATCTCCAAATCGGGCGGCGACTACGCCTACATGCTGGAGGTCTACGGCTCGC
TGCCCGCCTTCCTCAAGCTCTGGATCGAGCTGCTCATCATCCGGCCTTCATCGCAGTACAT
CGTGGCCCTGGTCTTCGCCACCTACCTGCTCAAGCCGCTCTTCCCCACCTGCCCGGTGCCC
GAGGAGGCAGCCAAGCTCGTGGCCTGCCTCTGCGTGCTGCTGCTCACGGCCGTGAACTGCT
ACAGCGTGAAGGCCGCCACCCGGGTCCAGGATGCCTTTGCCGCCGCCAAGCTCCTGGCCCT
GGCCCTGATCATCCTGCTGGGCTTCGTCCAGATCGGGAAGGGTGATGTGTCCAATCTAGAT
CCCAACTTCTCATTTGAAGGCACCAAACTGGATGTGGGGAACATTGTGCTGGCATTATACA
GCGGCCTCTTTGCCTATGGAGGATGGAATTACTTGAATTTCGTCACAGAGGAAATGATCAA
CCCCTACAGAAACCTGCCCCTGGCCATCATCATCTCCCTGCCCATCGTGACGCTGGTGTAC
GTGCTGACCAACCTGGCCTACTTCACCACCCTGTCCACCGAGCAGATGCTGTCGTCCGAGG
CCGTGGCCGTGGACTTCGGGAACTATCACCTGGGCGTCATGTCCTGGATCATCCCCGTCTT
CGTGGGCCTGTCCTGCTTCGGCTCCGTCAATGGGTCCCTGTTCACATCCTCCAGGCTCTTC
TTCGTGGGGTCCCGGGAAGGCCACCTGCCCTCCATCCTCTCCATGATCCACCCACAGCTCC
TCACCCCCGTGCCGTCCCTCGTGTTCACGTGTGTGATGACGCTGCTCTACGCCTTCTCCAA
GGACATCTTCTCCGTCATCAACTTCTTCAGCTTCTTCAACTGGCTCTGCGTGGCCCTGGCC
ATCATCGGCATGATCTGGCTGCGCCACAGAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGA
ACCTGGCCCTGCCTGTGTTCTTCATCCTGGCCTGCCTCTTCCTGATCGCCGTCTCCTTCTG
GAAGACACCCGTGGAGTGTGGCATCGGCTTCACCATCATCCTCAGCGGGCTGCCCGTCTAC
TTCTTCGGGGTCTGGTGGAAAAACAAGCCCAAGTGGCTCCTCCAGGGCATCTTCTCCACGA
CCGTCCTGTGTCAGAAGCTCATGCAGGTGGTCCCCCAGGAGACA**TAG**CCAGGAGGCCGAAG
GGCG

FIGURE 38

CCAGGCTCTCCACCCCCACTTCCCAATTGAGGAAACCGAGGCAGAGGAGGCTCAGGTGTGG
CCAATCACCCTGCACATCAGAGTTACCCTGGGCAGGGCCCACTGAGACCTGGGAGGGGCCA
CTCGGGACCTGGAGGGCTGGGGGCTGCCCGGGCGTTAGGGGTAAAGCTCCCTACCCAACTG
CGCAGAAGGCCTCAGAGGCCTGGGGGCTGGGCTTCCCCTTTCACATCGCCCTTTAGAGGCC
CACGTGTGGGCATTGGCCCGCGATCTGAAAGGGGCTGTCCTGTTCCTCATGGGCGCTGCCA
GCGCCACGCACTCCTCTTTCTGCCTGGCCGGCCACTCCCGTCTGCTGTGACGCGCGGACAG
AGAGCTACCGGTGGACCCACGGTGCCTCCCTCCCTGGGATCTACACAGACC**ATG**GCCTTGC
CAACGGCTCGACCCCTGTTGGGGTCCTGTGGGACCCCCGCCCTCGGCAGCCTCCTGTTCCT
GCTCTTCAGCGTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGACAGGGCAGGAG
GCTGCACCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCTCTCCCCTCGCC
AACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGTGTCCGGGAGCT
GGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGCGCTGTCTGGCT
CACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCTGCTGCTATTCC
TCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCCCGCATCACGAA
GGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGCTGCCTGCGGCT
CTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCGGGCTCTGGGAG
GCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTGCTGCTACCCCG
GCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCAGGGCGGCTCTG
CAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCACGATGGACGCTC
TGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCGCAGGGCATCGT
GGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTGAACGGACCATC
CTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGGCAAGAAGGCCC
GCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCCTGCGTGGATGC
GGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCTACGAGCAGCTG
GACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGAGTCTGTGATCC
AGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAGTGGAATGTGAC
GTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAATGAGTCCTCAG
GTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAAAGACACCCTAG
ACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAGGAGCTGAGCTC
CGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGTGTGACCCAAGG
CAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAACGGGTCCGAAT
ACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTGAAGGCGCTCAG
TCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGGATGCGGTGCTG
CCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGGCCTGAAGGCGG
AGGAGCGGCACCGCCCGGTGCGGGACTGGATCCTACGGCAGCGGCAGGACGACCTGGACAC

GCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAGACCTCAGCGTG
CAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCTCACCGTCCTGG
CACTGCTCCTAGCCTCCACCCTGGCC**TGA**GGGCCCCACTCCCTTGCTGGCCCCAGCCCTGC
TGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACCCCAAGAGAACT
CGCGCTCAGTAAACGGGAACATGCCCCCTGCAGACACGTAAAAAAAAAAAAAAAAAA

FIGURE 39

TGGGATCTACACAGACC**ATG**GCCTTGCCAACGGCTCGACCCCTGTTGGGGTCCTGTGGGAC
CCCCGCCCTCGGCAGC
CTCCTGTTCCTGCTCTTCAGCCTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGA
CAGGGCAGGAGGCTGCGCCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCT
CTCCCCTCGCCAACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGT
GTCCGGGAGCTGGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGC
GCTGTCTGGCTCACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCT
GCTGCTATTCCTCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCC
CGCATCACGAAGGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGC
TGCCTGCGGCTCTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCG
GGCTCTGGGAGGCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTG
CTGCTACCCCGGCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCA
GGGCGGCTCTGCAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCAC
GATGGACGCTCTGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCG
CAGGGCATCGTGGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTG
AACGGACCATCCTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGG
CAAGAAGGCCCGCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCC
TGCGTGGATGCGGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCT
ACGAGCAGCTGGACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGA
GTCTGTGATCCAGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAG
TGGAATGTGACGTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAA
TGAGTCCTCAGGTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAA
AGACACCCTAGACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAG
GAGCTGAGCTCCGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGT
GTGACCCAAGGCAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAA
CGGGTCCGAATACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTG
AAGGCGCTCAGTCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGG
ATGCGGTGCTGCCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGG
CCTGAAGGCGGAGGAGCGGCACCGCCCGGTGCGGGACTGGATCCTACGGCAGCGGCAGGAC
GACCTGGACACGCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAG
ACCTCAGCGTGCAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCT
CACCGTCCTGGCACTGCTCCTAGCCTCCACCCTGGCC**TGA**GGGCCCCACTCCCTTGCTGGC
CCCAGCCCTGCTGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACC
CCAAGAGAACT

FIGURE 40

ATAACAGCATGAAGTGCCGTGGAACTGGAATAGGCGTGTCCTCTCCCTCGACCCTCCCCCT
CCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCCTCCCTCCGGCGAGGGCCGCCTTTATAAC
AACTGCTCAGAGTGCGAGGGCGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTC
GCCTGCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTA
GGGTCCAGA**ATG**GCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACT
ACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGG
GGTTGTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGAC
TCCAACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCT
TTGGCCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTT
CCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTG
ACCAAGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGG
GCTTCAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGAC
CAACGTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTG
CTCACCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTG
GTTCCTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCAT
TGCCATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGAT
GACACCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTA
GTCCCGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGC
TTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAA
GAGGAAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACAC
ATTTTCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTG
GCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGC**TAA**CTCTGTCCTGAAGAG
TGGGACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCT
TGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTC
AACCACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCG
CTGTAGTATTTTTTTTTTTTTGTCTCATCCTTTGGATACTTCTTTTAAGTGGGAGTCTCAG
GCAACTCAAGTTTAGACCCTTACTCTTTTTGTTTGTTTTTTGAAACAGGATCTTGCTCTGT
CACCCAGGCTTGAGTGCAGTGGTGCGATCACAGCCCAGTGCAGCCTCGACCACCTGTGCTC
AAGCAATCCTCCCATCTCCATCTCCCAAAGTGCTGGGATGACAGGCGTGAGCCACAGCTCC
CAGCCTAGGCCCTTAATCTTGCTGTTATTTTCCATGGACTAAAGGTCTGGTCATCTGAGCT
CACGCTGGCTCACACAGCTCTAGGGGCCTGCTCCTCTAACTCACAGTGGGTTTTGTGAGGC
TCTGTGGCCCAGAGCAGACCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCA
CTGGCCTGAATCTACACTGGAAGCCAACTTGCTGGCACCCCCGCTCCCCAACCCTTCTTGC
CTGGGTAGGAGAGGCTAAAGATCACCCTAAATTTACTCATCTCTCTAGTGCTGCCTCACAT
TGGGCCTCAGCAGCTCCCCAGCACCAATTCACAGGTCACCCCTCTCTTCTTGCACTGTCCC

CAAACTTGCTGTCAATTCCGAGATCTAATCTCCCCCTACGCTCTGCCAGGAATTCTTTCAG
ACCTCACTAGCACAAGCCCGGTTGCTCCTTGTCAGGAGAATTTGTAGATCATTCTCACTTC
AAATTCCTGGGGCTGATACTTCTCTCATCTTGCACCCCAACCTCTGTAAATAGATTTACCG
CATTTACGGCTGCATTCTGTAAGTGGGCATGGTCTCCTAATGGAGGAGTGTTCATTGTATA
ATAAGTTATTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTCATGGTGGTGGCA
GCAAAAAAAAAAAAAA

FIGURE 41

AGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTGTTAA
CCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGGTCGG
GACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATACTTCA
GCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTTTTATATGA
ATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCCCATT
GACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCTGACG
CCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGAGTAA
TGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAATGTT
GAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGACTTAA
AAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATTATCC
TAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAACAAAT
GGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGACTTCA
TTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATATTGGA
GGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGAGAAA
CTTCCACGAGGTGTCCTAGAAAGGTTCATGGGAGACAAAGTGGCTTCCGAAGCTCTTATGA
AGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTTTGA
GACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAATGGA
ATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATTGATC
AGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAACCAC
CTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGAATAC
ACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGTTCAGCGCC
TCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGCTTTA
CAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCGTACA
GCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGTGCGA
TTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAAAACA
AGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCTCCTG
ACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATGACTA
GTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTTATCA
CCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTTAAAG
ACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAATGAAT
TGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATAAACT
CCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTA
ATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTG
TCTATACAACTAAAATTATAACCAAAGTTGTGGAACCAAAAATTAAAGTGATTGAAGGCAG
TCTTCAGCCTATTATCAAAACTGAAGGACCCACACTAACAAAAGTCAAAATTGAAGGTGAA

CCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAA
TTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGAC
ACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGT
GGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAAT
TCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGG
AGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGTTCTAGAAGACGATTA
AGGGAAGGTCGTTCTCAG**TGA**AAATCCAAAAACCAGAAAAAAATGTTTATACAACCCTAAG
TCAATAACCTGACCTTAGAAAATTGTGAGAGCCAAGTTGACTTCAGGAACTGAAACATCAG
CACAAAGAAGCAATCATCAAATAATTCTGAACACAAATTTAATATTTTTTTTTCTGAATGA
GAAACATGAGGGAAATTGTGGAGTTAGCCTCCTGTGGTAAAGGAATTGAAGAAAATATAAC
ACCTTACACCCTTTTTCATCTTGACATTAAAAGTTCTGGCTAACTTTGGAATCCATTAGAG
AAAAATCCTTGTCACCAGATTCATTACAATTCAAATCGAAGAGTTGTGAACTGTTATCCCA
TTGAAAAGACCGAGCCTTGTATGTATGTTATGGATACATAAAATGCACGCAAGCCATTATC
TCTCCATGGGAAGCTAAGTTATAAAAATAGGTGCTTGGTGTACAAAACTTTTTATATCAAA
AGGCTTTGCACATTTCTATATGAGTGGGTTTACTGGTAAATTATGTTATTTTTTACAACTA
ATTTTGTACTCTCAGAATGTTTGTCATATGCTTCTTGCAATGCATATTTTTTAATCTCAAA
CGTTTCAATAAAACCATTTTTCAGATATAAAGAGAATTACTTCAAATTGAGTAATTCAGAA
AAACTCAAGATTTAAGTTAAAAAGTGGTTTGGACTTGGGAA

FIGURE 42

CCCTTGAGACTCAAG**ATG**ATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA
TTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACAC

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGA
TGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCCGTGAGG
AAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTC
AG**TGA**AGGGC

FIGURE 43

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACA
CCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAG
GTCGTTCTCAG**TGA**AGGGC

FIGURE 44

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA
TTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACAC

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGT
TCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURE 45

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTC
ACCAAGGTCACCAAATTCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTA
AAAGACTGCTTCAGGGAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGG
ATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURE 46

ACACGGACCAAGGAGTCTAACACGTGCGCGAGTCGGGGGCTCGCACGAAAGCCGCCGTGGC
GCAATGAAGGTGAAGGCCGGCGCGCTCGCCGGCCGAGGTGGGATCCCGAGGCCTCTCCAGT
CCGCCGAGGGCGCACCACCGGCCCGTCTCGCCCGCCGCGCCGGGGAGGTGGAGCACGAGCG
CACGTGTTAGGACCCGAAAG**ATG**GTGAACTATGCCTGGGCAGGGCGAAGCCAGAGGAAACT
CTGGTGGAGGTCCGTAGCGGTCCTGACGTGCAAATCGGTCGTCCGACCTGGGTATAGGGGC
GGGCTCCAGGCGAGGCGGTCGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGC
CCGGAGCGATGAAGATGGTCGCGCCCTGGACGCGGTTCTACTCCAACAGCTGCTGCTTGTG
CTGCCATGTCCGCACCGGCACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTG
GTACTGTTGATTTTATTGAGTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTG
AACTGGGAGGTGACTTTGAGTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTC
TCTTCTCATGATCCTGATATGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCC
TGGATCATCCCATTCTTCTGTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAA
TCACTGTGCTTATTTATCCAAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTT
TCCCTACAGAGATGATGTCATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTG
TTTATTAGCATTATCTTGACTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACC
GATACATCAATGGTAGGAACTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTAC
GGTGCTGCTACCCCCGTATGATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCA
CCTTACGTGTCTGCC**TAA**GCCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCA
GACATCTGAGCAATAGTTCTGTTATTTCACTTTTGCCATGAGCCTCTCTGAGCTTGTTTGT
TGCTGAAATGCTACTTTTTAAAATTTAGATGTTAGATTGAAAACTGTAGTTTTCAACATAT
GCTTTGCTAGAACACTGTGATAGATTAACTGTAGAATTCTTCCTGTACGATTGGGGATATA
ATGGGCTTCACTAACCTTCCCTAGGCATTGAAACTTCCCCCAAATCTGATGGACCTAGAAG
TCTGCTTTTGTACCTGCTGGGCCCCAAAGTTGGGCATTTTTCTCTCTGTTCCCTCTCTTTT
GAAAATGTAAAATAAAACCAAAAATAGACAACTTTTTCTTCAGCCATTCCAGCATAGAGAA
CAAAACCTTATGGAAACAGGAATGTCAATTGTGTAATCATTGTTCTAATTAGGTAAATAGA
AGTCCTTATGTATGTGTTACAAGAATTTCCCCCACAACATCCTTTATGACTGAAGTTCAAT
GACAGTTTGTGTTTGGGTGGTAAAGGATTTTCTCCATGGCCTGAATTAAGACCATTAGAAA
GCACCAGGCCGTGGGAGCAGTGACCATCTGCTGACTGTTCTTGTGGATCTTGTGTCCAGGG
ACATGGGGTGACATGCCTCGTATGTGTTAGAGGGTGGAATGGATGTGTTTGGCGCTGCATG
GGATCTGGTGCCCCTCTTCTCCTGGATTCACATCCCCACCCAGGGCCCGCTTTTACTAAGT
GTTCTGCCCTAGATTGGTTCAAGGAGGTCATCCAACTGACTTTATCAAGTGGAATTGGGAT
ATATTTGATATACTTCTGCCTAACAACATGGAAAAGGGTTTTCTTTTCCCTGCAAGCTACA
TCCTACTGCTTTGAACTTCCAAGTATGTCTAGTCACCTTTTAAAATGTAAACATTTTCAGA
AAAATGAGGATTGCCTTCCTTGTATGCGCTTTTTACCTTGACTACCTGAATTGCAAGGGAT
TTTTATATATTCATATGTTACAAAGTCAGCAACTCTCCTGTTGGTTCATTATTGAATGTGC

TGTAAATTAAGTTGTTTGCAATTAAAACAAGGTTTGCCCACAAAAAAAAAA

FIGURE 47

ATTCGCCCTTCCGAAAGATGGTGAACT**ATG**CCTGGGCAGGGCGGGCTCCAGGCGAGGCGGT
CGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGCCCGGAGCGATGAAGATGGT
CGCGCCCTGGACTCGGTTCTACTCCAACAGCTGCTGCTTGTGCTGCCATGTCCGCACCGGC
ACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTGGTACTGTTGATTTTATTGA
GTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTGAACTGGGAGGTGACTTTGA
GTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTCTCTTCTCATGATCCTGATA
TGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCCTGGATCATCCCATTCTTCT
GTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAATCACTGTGCTTATTTATCC
AAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTTTCCCTACAGAGATGATGTC
ATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTGTTTATTAGCATTATCTTGA
CTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACCGATACATCAATGGTAGGAA
CTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTACGGTGCTGCTACCCCCGTAT
GATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCACCTTACGTGTCTGCC**TAA**G
CCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCAGACATCTGAGAAGGGC

FIGURE 48

AGCGCC**ATG**GGGCGAGCCGGGGGCGGGGGCCCGGGCCGGGGGCCGCCGCCACTGCTGCTGT
TTCTGGGGGCCGCGCTGGTCCTGGCCTCTGGGGCCGTGCCGGCGCGTGAGGCGGGCAGCGC
GGTTGAGGCCGAAGAGCTGGTGAAGGGCAGCCCGGCGTGGGAGCCGCCTGCCAACGACACG
CGGGAAGAAGCCGGCCCACCAGCGGCTGGGGAAGATGAGGCGTCGTGGACGGCGCCCGGTG
GCGAGCTGGCCGGGCCAGAAGAGGTGCTGCAGGAGTCGGCTGCGGTGACCGGCACCGCCTG
GCTGGAAGCTGACAGCCCAGGCCTGGGAGGAGTGACCGCAGAGGCGGGCAGCGGCGATGCC
CAGGCCCTTCCAGCTACGCTCCAGGCTCCCCACGAGGTCCTCGGGCAGTCAATCATGCCCC
CTGCCATTCCTGAGGCTACAGAGGCCAGCGGGCCACCCTCCCCCACCCCCGGCGACAAGCT
GAGCCCAGCTTCTGAACTCCCCAAGGAGAGCCCCTTGGAGGTTTGGCTGAACCTGGGGGGC
AGCACACCCGACCCTCAAGTGCCAGAGCTGACTTACCCATTTCAGGGCACCCTGGAGCCCC
AACCGGCATCAGATATCATTGACATCGACTACTTCGAAGGACTGGATGGTGAGGGTCGTGG
CGCAGATCTGGGGAGCTTCCCAGGGTCACCAGGAACCTCAGAGAACCACCCTGATACTGAG
GGAGAGACCCCTTCCTGGAGCCTGCTTGACTTATACGATGATTTCACCCCCTTCGATGAAT
CTGATTTCTACCCCACCACATCCTTTTATGATGACTTGGATGAAGAGGAGGAGGAAGAGGA
GGATGACAAAGATGCAGTAGGAGGTGGAGACCTAGAAGATGAAAATGAGCTTCTAGTGCCC
ACTGGGAAGCCTGGTCTGGGGCCCGGGACAGGCCAGCCCACCAGTCGGTGGCATGCTGTCC
CTCCACAGCACACTCTGGGGTCGGTCCCCGGCAGCAGCATCGCCCTCAGGCCCCGCCCAGG
AGAGCCAGGCAGGGACTTGGCCTCCAGTGAAAATGGCACTGAGTGCCGCAGTGGCTTTGTG
CGGCATAACGGCTCCTGCCGGTCAGTGTGCGACCTCTTCCCAAGTTACTGTCACAATGGCG
GCCAGTGCTACCTGGTGGAGAACATAGGGGCCTTCTGCAGGTGCAACACGCAGGACTACAT
CTGGCACAAGGGGATGCGCTGCGAGTCCATCATCACCGACTTCCAGGTGATGTGCGTGGCC
GTGGGCTCGGCTGCCCTCGTCCTGCTCCTGCTCTTCATGATGACGGTGTTCTTTGCCAAGA
AGCTCTACCTGCTCAAGACGGAGAATACCAAGCTGCGTAGGACCAACAAATTCCGGACCCC
ATCTGAGCTCCACAATGATAACTTCTCCCTCTCCACCATTGCCGAGGGCTCTCACCCAAAT
GATGATCCTAGTGCTCCCCACAAAATCCAGGAGGTTCTCAAGTCCTGCCTGAAAGAGGAGG
AGTCATTTAACATCCAGAACTCCATGTCGCCCAAACTTGAGGGTGGCAAAGGTGACCAGGC
TGACTTGGATGTGAACTGTCTTCAGAATAATTTAACC**TAA**AGCAGAGCAAGAAGAGAGGAA
GCGGGGTAGTGGGTGGGGGTAGGGAAGAAACATTATCTCCTCTTGTACAGAGTCTATTTCT
TGTAACCATTTGTTAAACTC

FIGURE 49

GGACAGGCGTGGCGGCCGGAGCCCCAGCATCCCTGCTTGAGGTCCAGGAGCGGAGCCCGCG
GCCACCGCCGCCTGATCAGCGCGACCCCGGCCCGCGCCCGCCCCGCCCGGCAAGATGCTGC
CCGTGTACCAGGAGGTGAAGCCCAACCCGCTGCAGGACGCGAACATCTGCTCACGCGTGTT
CTTCTGGTGGCTCAATCCCTTGTTTAAAATTGGCCATAAACGGAGATTAGAGGAAGATGAT
ATGTATTCAGTGCTGCCAGAAGACCGCTCACAGCACCTTGGAGAGGAGTTGCAAGGGTTCT
GGGATAAAGAAGTTTTAAGAGCTGAGAATGACGCACAGAAGCCTTCTTTAACAAGAGCAAT
CATAAAGTGTTACTGGAAATCTTATTTAGTTTTGGGAATTTTTACGTTAATTGAGGAAAGT
GCCAAAGTAATCCAGCCCATATTTTTGGGAAAAATTATTAATTATTTTGAAAATTATGATC
CCATGGATTCTGTGGCTTTGAACACAGCGTACGCCTATGCCACGGTGCTGACTTTTTGCAC
GCTCATTTTGGCTATACTGCATCACTTATATTTTTATCACGTTCAGTGTGCTGGGATGAGG
TTACGAGTAGCCATGTGCCATATGATTTATCGGAAGGCACTTCGTCTTAGTAACATGGCCA
TGGGGAAGACAACCACAGGCCAGATAGTCAATCTGCTGTCCAATGATGTGAACAAGTTTGA
TCAGGTGACAGTGTTCTTACACTTCCTGTGGGCAGGACCACTGCAGGCGATCGCAGTGACT
GCCCTACTCTGGATGGAGATAGGAATATCGTGCCTTGCTGGGATGGCAGTTCTAATCATTC
TCCTGCCCTTGCAAAGCTGTTTTGGGAAGTTGTTCTCATCACTGAGGAGTAAAACTGCAAC
TTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATAACTGGTATAAGGATAATAAAA
ATGTACGCCTGGGAAAAGTCATTTTCAAATCTTATTACCAATTTGAGAAAGAAGGAGATTT
CCAAGATTCTGAGAAGTTCCTGCCTCAGGGGGATGAATTTGGCTTCGTTTTTCAGTGCAAG
CAAAATCATCGTGTTTGTGACCTTCACCACCTACGTGCTCCTCGGCAGTGTGATCACAGCC
AGCCGCGTGTTCGTGGCAGTGACGCTGTATGGGGCTGTGCGGCTGACGGTTACCCTCTTCT
TCCCCTCAGCCATTGAGAGGGTGTCAGAGGCAATCGTCAGCATCCGAAGAATCCAGACCTT
TTTGCTACTTGATGAGATATCACAGCGCAACCGTCAGCTGCCGTCAGATGGTAAAAAGATG
GTGCATGTGCAGGATTTTACTGCTTTTTGGGATAAGGCATCAGAGACCCCAACTCTACAAG
GCCTTTCCTTTACTGTCAGACCTGGCGAATTGTTAGCTGTGGTCGGCCCCGTGGGAGCAGG
GAAGTCATCACTGTTAAGTGCCGTGCTCGGGGAATTGGCCCCAAGTCACGGGCTGGTCAGC
GTGCATGGAAGAATTGCCTATGTGTCTCAGCAGCCCTGGGTGTTCTCGGGAACTCTGAGGA
GTAATATTTTATTTGGGAAGAAATATGAAAAGGAACGATATGAAAAAGTCATAAAGGCTTG
TGCTCTGAAAAAGGATTTACAGCTGTTGGAGGATGGTGATCTGACTGTGATAGGAGATCGG
GGAACCACGCTGAGTGGAGGGCAGAAAGCACGGGTAAACCTTGCAAGAGCAGTGTATCAAG
ATGCTGACATCTATCTCCTGGACGATCCTCTCAGTGCAGTAGATGCGGAAGTTAGCAGACA
CTTGTTCGAACTGTGTATTTGTCAAATTTTGCATGAGAAGATCACAATTTTAGTGACTCAT
CAGTTGCAGTACCTCAAAGCTGCAAGTCAGATTCTGATATTGAAAGATGGTAAAATGGTGC
AGAAGGGGACTTACACTGAGTTCCTAAAATCTGGTATAGATTTTGGCTCCCTTTTAAAGAA
GGATAATGAGGAAAGTGAACAACCTCCAGTTCCAGGAACTCCCACACTAAGGAATCGTACC
TTCTCAGAGTCTTCGGTTTGGTCTCAACAATCTTCTAGACCCTCCTTGAAAGATGGTGCTC

TGGAGAGCCAAGATACAGAGAATGTCCCAGTTACACTATCAGAGGAGAACCGTTCTGAAGG
AAAAGTTGGTTTTCAGGCCTATAAGAATTACTTCAGAGCTGGTGCTCACTGGATTGTCTTC
ATTTTCCTTATTCTCCTAAACACTGCAGCTCAGGTTGCCTATGTGCTTCAAGATTGGTGGC
TTTCATACTGGGCAAACAAACAAAGTATGCTAAATGTCACTGTAAATGGAGGAGGAAATGT
AACCGAGAAGCTAGATCTTAACTGGTACTTAGGAATTTATTCAGGTTTAACTGTAGCTACC
GTTCTTTTTGGCATAGCAAGATCTCTATTGGTATTCTACGTCCTTGTTAACTCTTCACAAA
CTTTGCACAACAAAATGTTTGAGTCAATTCTGAAAGCTCCGGTATTATTCTTTGATAGAAA
TCCAATAGGAAGAATTTTAAATCGTTTCTCCAAAGACATTGGACACTTGGATGATTTGCTG
CCGCTGACGTTTTTAGATTTCATCCAGACATTGCTACAAGTGGTTGGTGTGGTCTCTGTGG
CTGTGGCCGTGATTCCTTGGATCGCAATACCCTTGGTTCCCCTTGGAATCATTTTCATTTT
TCTTCGGCGATATTTTTTGGAAACGTCAAGAGATGTGAAGCGCCTGGAATCTACAACTCGG
AGTCCAGTGTTTTCCCACTTGTCATCTTCTCTCCAGGGGCTCTGGACCATCCGGGCATACA
AAGCAGAAGAGAGGTGTCAGGAACTGTTTGATGCACACCAGGATTTACATTCAGAGGCTTG
GTTCTTGTTTTTGACAACGTCCCGCTGGTTCGCCGTCCGTCTGGATGCCATCTGTGCCATG
TTTGTCATCATCGTTGCCTTTGGGTCCCTGATTCTGGCAAAAACTCTGGATGCCGGGCAGG
TTGGTTTGGCACTGTCCTATGCCCTCACGCTCATGGGGATGTTTCAGTGGTGTGTTCGACA
AAGTGCTGAAGTTGAGAATATGATGATCTCAGTAGAAAGGGTCATTGAATACACAGACCTT
GAAAAAGAAGCACCTTGGGAATATCAGAAACGCCCACCACCAGCCTGGCCCCATGAAGGAG
TGATAATCTTTGACAATGTGAACTTCATGTACAGTCCAGGTGGGCCTCTGGTACTGAAGCA
TCTGACAGCACTCATTAAATCACAAGAAAAGGTTGGCATTGTGGGAAGAACCGGAGCTGGA
AAAAGTTCCCTCATCTCAGCCCTTTTTAGATTGTCAGAACCCGAAGGTAAAATTTGGATTG
ATAAGATCTTGACAACTGAAATTGGACTTCACGATTTAAGGAAGAAAATGTCAATCATACC
TCAGGAACCTGTTTTGTTCACTGGAACAATGAGGAAAAACCTGGATCCCTTTAAGGAGCAC
ACGGATGAGGAACTGTGGAATGCCTTACAAGAGGTACAACTTAAAGAAACCATTGAAGATC
TTCCTGGTAAAATGGATACTGAATTAGCAGAATCAGGATCCAATTTTAGTGTTGGACAAAG
ACAACTGGTGTGCCTTGCCAGGGCAATTCTCAGGAAAAATCAGATATTGATTATTGATGAA
GCGACGGCAAATGTGGATCCAAGAACTGATGAGTTAATACAAAAAAAAATCCGGGAGAAAT
TTGCCCACTGCACCGTGCTAACCATTGCACACAGATTGAACACCATTATTGACAGCGACAA
GATAATGGTTTTAGATTCAGGAAGACTGAAAGAATATGATGAGCCGTATGTTTTGCTGCAA
AATAAAGAGAGCCTATTTTACAAGATGGTGCAACAACTGGGCAAGGCAGAAGCCGCTGCCC
TCACTGAAACAGCAAAACAGGTATACTTCAAAAGAAATTATCCACATATTGGTCACACTGA
CCACATGGTTACAAACACTTCCAATGGACAGCCCTCGACCTTAACTATTTTCGAGACAGCA
CTG<u>TGA</u>ATCCAACCAAAATGTCAAGTCCGTTCCGAAGGCATTTTCCACTAGTTTTTGGACT
ATGTAAACCACATTGTACTTTTTTTTTACTTTGGCAACAAATATTTATACATACAAGATGCT
AGTTCATTTGAATATTTCTCCC

FIGURE 50

CTCCGTCCCAGGGGAGAAGGAGAGGCGTCTGCAGGGGGCAGAGACCGCAGCTACCTGCCGG
GTGCGCCCCCCACCCAGGAGCGCTCGCTTCGCCCCCTTTCCTCCCCCCGCCCCCACCTCCT
TATTGGTGCTAGTTTGCAGCGCCCAGCTCCTGCGCCTTCGCTTCGCGTTTGAATCTGGCTC
GCCCCTTCGTATT**ATG**TCTGCACTCCGAAGGAAATTTGGGGACGATTATCAGGTAGTGACC
ACATCGTCCAGCGGCTCGGGCTTGCAGCCCCAGGGGCCAGGCCAGGACCCTCAGCAGCAGC
TTGTGCCCAAGAAGAAGCGGCAGCGGTTCGTGGACAAGAACGGCCGGTGCAATGTACAGCA
CGGCAACCTGGGCAGCGAGACAAGCCGCTACCTCTCGGACCTCTTCACCACGCTGGTGGAC
CTCAAGTGGCGCTGGAACCTCTTCATCTTCATTCTCACCTACACCGTGGCCTGGCTTTTCA
TGGCGTCCATGTGGTGGGTGATCGCCTACACTCGGGGCGACCTGAACAAAGCCCACGTCGG
TAACTACACGCCTTGCGTGGCCAATGTCTATAACTTCCCTTCTGCCTTCCTCTTCTTCATC
GAGACGGAGGCCACCATCGGCTATGGCTACCGATACATCACAGACAAGTGCCCCGAGGGCA
TCATCCTCTTCCTCTTCCAGTCCATCCTGGGCTCCATCGTGGACGCCTTCCTCATCGGCTG
CATGTTCATCAAGATGTCCCAGCCCAAGAAGCGCGCCGAGACCCTCATGTTCAGCGAGCAC
GCGGTGATCTCCATGAGGGACGGAAAACTCACGCTTATGTTCCGGGTGGGCAACCTGCGCA
ACAGCCACATGGTCTCCGCGCAGATTCGCTGCAAGCTGCTCAAATCTCGGCAGACACCTGA
GGGTGAGTTCCTTCCCCTTGACCAACTTGAACTGGATGTAGGTTTTAGTACAGGGGCAGAT
CAACTTTTTCTTGTGTCCCCCCTCACAATTTGCCACGTGATCGATGCCAAAAGCCCCTTTT
ATGACCTATCCCAGCGAAGCATGCAAACTGAACAGTTCGAGATTGTCGTCATCCTAGAAGG
CATTGTGGAAACAACTGGGATGACTTGTCAAGCTCGAACATCATATACTGAAGATGAAGTT
CTTTGGGGTCATCGTTTTTTTCCTGTAATTTCCTTAGAAGAGGGATTCTTTAAAGTTGATT
ACTCCCAGTTCCACGCAACATTTGAAGTCCCCACCCCACCTTACAGTGTGAAAGAGCAGGA
GGAAATGCTTCTCATGTCGTCCCCTTTAATAGCACCAGCCATAACTAACAGCAAAGAAAGA
CATAATTCTGTGGAATGCTTAGATGGACTAGATGATATTACTACAAAACTACCATCTAAGC
TGCAGAAAATTACTGGAAGAGAAGACTTTCCCAAAAAACTCTTGAGGATGAGTTCTACAAC
TTCAGAAAAAGCCTACAGCTTGGGAGACTTGCCCATGAAACTTCAACGAATAAGTTCAGTT
CCGGGCAACTCAGAAGAAAAACTGGTATCTAAAACCACCAAGATGTTATCTGATCCCATGA
GCCAGTCTGTGGCTGATTTGCCACCAAAGCTTCAAAAGATGGCTGGAGGAGCAGCTAGGAT
GGAAGGGAACCTTCCAGCCAAATTAAGAAAAATGAACTCTGATCGCTTCACA**TAA**CAAAGC
ACTCCCTTAGGCATTATTTAATGTTTGATTTAGTAATAGTCCAATATTTGGCGATGAGGTA
ATTCTCCCTAAGGAATCTGAAAGTATATTTTCCTCCCAGTTCTACAGCATATTTCGAGAAC
CCTTCCTTTCCCAAGTATTGCGAATGTGCAGAAAGCAACAGTTACGGAGGGAGGACATCAT
AAGGAAGTTATTAACGGGCATGTATTATCACATCAAGCATGCAATAATGTGCAAATTTTGC
ATTTAGTTTTATGGCATGATTTATATATGGCATATTTATATTGTATATTCTGGAAAAAAAA
ATATATATATATATTTAAAGGGGAGATACTCTCCCTGACATTTCTAACATATGTATTAAGC
CAAACATGAGTGAATAGCTTTCAGGGCGATAAAACTAAATATATGTCTGTGTGTGTGTGTG

TATGTATACACACATATACATATATATATACACATACATACACATACATACATACATACAT
ATATATCTGATAAAATTGTGATGTTTTGTTCAAAGTTGTAGTTCTTGTGCATGTTTACTTT
ATTAGAGTAGGAAGGCTACTGGCATTAATTATTAATACCAAATATTTTAGCCTTAAATTTT
TGTCATTTTAAAATCTGATTTAATGTTTTCTGCTGTTTAAGGTCTTGGGAGGCTTTCAATT
GTATTTTATATGAGAGAATCACACAAGTTTGTGCTATCTATGGCCCTGCAAAAATATAACC
ATTACATGTTTAAATTGTAAATTTTAGAGCATACCAGTACTCAGTATAGCATTGAACATTT
CTTATGATTTTTAAAAGTTGCTAGTACTGGGGAGAAATAATTGTTGATTAATTTGAGAATT
ATTCCTTTCCTAGACTAATTAAAATCTGGAAATCTGTTTTGTATATGATCTAATACAAAGA
TGAGCTCTGAACAAACACTGAATCATGTTAATAGACAGTAGCCAAGTTATATTGAATATAT
CAGAATCTGTGTGAAGTTACACAATTAATTGTCCCTGTTTCAAACTGAGTAAATTGGAAAC
ATTTTCTTTCTTTTTCTGGAAATTTTGTCCATTTTAAAAACCAATCATTTTAAGAAGACAT
GACAATGCAATGAAACAGATGATAAATATTTATGCTTAAAATAAAAAAAAAAAAA

FIGURE 51

GAGAGGGTCCTTCAGGGTCTGCTTATGCCCTTGTTCAAGAACACCAGTGTCAGCTCTCTGT
ACTCTGGTTGCAGACTGACCTTGCTCAGGCCTGAGAAGGATGGGGCAGCCACCAGAGTGGA
TGCTGTCTGCACCCATCGTCCTGACCCCAAAAGCCCTGGACTGGACAGAGAGCGGCTGTAC
TGGAAGCTGAGCCAGCTGACCCACGGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGC
ACAGTCTCTATGTCAATGGTTTCACCCATCAGAGCTCTATGACGACCACCAGAACTCCTGA
TACCTCCACAATGCACCTGGCAACCTCGAGAACTCCAGCCTCCCTGTCTGGACCTACGACC
GCCAGCCCTCTCCTGGTGCTATTCACAATTAACTTCACCATCACTAACCTGCGGTATGAGG
AGAACATGCATCACCCTGGCTCTAGAAAGTTTAACACCACGGAGAGAGTCCTTCAGGGTCT
GCTCAGGCCTGTGTTCAAGAACACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACC
TTGCTCAGGCCCAAGAAGGATGGGGCAGCCACCAAAGTGGATGCCATCTGCACCTACCGCC
CTGATCCCAAAAGCCCTGGACTGGACAGAGAGCAGCTATACTGGGAGCTGAGCCAGCTAAC
CCACAGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGGACAGTCTCTATGTCAATGGT
TTCACACAGCGGAGCTCTGTGCCCACCACTAGCATTCCTGGGACCCCCACAGTGGACCTGG
GAACATCTGGGACTCCAGTTTCTAAACCTGGTCCCTCGGCTGCCAGCCCTCTCCTGGTGCT
ATTCACTCTCAACTTCACCATCACCAACCTGCGGTATGAGGAGAACATGCAGCACCCTGGC
TCCAGGAAGTTCAACACCACGGAGAGGGTCCTTCAGGGCCTGCTCAGGTCCCTGTTCAAGA
GCACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACTTTGCTCAGGCCTGAAAAGGA
TGGGACAGCCACTGGAGTGGATGCCATCTGCACCCACCACCCTGACCCCAAAAGCCCTAGG
CTGGACAGAGAGCAGCTGTATTGGGAGCTGAGCCAGCTGACCCACAATATCACTGAGCTGG
GCCACTATGCCCTGGACAACGACAGCCTCTTTGTCAATGGTTTCACTCATCGGAGCTCTGT
GTCCACCACCAGCACTCCTGGGACCCCCACAGTGTATCTGGGAGCATCTAAGACTCCAGCC
TCGATATTTGGCCCTTCAGCTGCCAGCCATCTCCTGATACTATTCACCCTCAACTTCACCA
TCACTAACCTGCGGTATGAGGAGAACATGTGGCCTGGCTCCAGGAAGTTCAACACTACAGA
GAGGGTCCTTCAGGGCCTGCTAAGGCCCTTGTTCAAGAACACCAGTGTTGGCCCTCTGTAC
TCTGGCTCCAGGCTGACCTTGCTCAGGCCAGAGAAAGATGGGGAAGCCACCGGAGTGGATG
CCATCTGCACCCACCGCCCTGACCCCACAGGCCCTGGGCTGGACAGAGAGCAGCTGTATTT
GGAGCTGAGCCAGCTGACCCACAGCATCACTGAGCTGGGCCCCTACACACTGGACAGGGAC
AGTCTCTATGTCAATGGTTTCACCCATCGGAGCTCTGTACCCACCACCAGCACCGGGGTGG
TCAGCGAGGAGCCATTCACACTGAACTTCACCATCAACAACCTGCGCTACATGGCGGACAT
GGGCCAACCCGGCTCCCTCAAGTTCAACATCACAGACAACGTCATGAAGCACCTGCTCAGT
CCTTTGTTCCAGAGGAGCAGCCTGGGTGCACGGTACACAGGCTGCAGGGTCATCGCACTAA
GGTCTGTGAAGAACGGTGCTGAGACACGGGTGGACCTCCTCTGCACCTACCTGCAGCCCCT
CAGCGGCCCAGGTCTGCCTATCAAGCAGGTGTTCCATGAGCTGAGCCAGCAGACCCATGGC
ATCACCCGGCTGGGCCCCTACTCTCTGGACAAAGACAGCCTCTACCTTAACGGTTACAATG
AACCTGGTCTAGATGAGCCTCCTACAACTCCCAAGCCAGCCACCACATTCCTGCCTCCTCT

GTCAGAAGCCACAACAGCCATGGGGTACCACCTGAAGACCCTCACACTCAACTTCACCATC
TCCAATCTCCAGTATTCACCAGATATGGGCAAGGGCTCAGCTACATTCAACTCCACCGAGG
GGGTCCTTCAGCACCTGCTCAGACCCTTGTTCCAGAAGAGCAGCATGGGCCCCTTCTACTT
GGGTTGCCAACTGATCTCCCTCAGGCCTGAGAAGGATGGGGCAGCCACTGGTGTGGACACC
ACCTGCACCTACCACCCTGACCCTGTGGGCCCCGGGCTGGACATACAGCAGCTTTACTGGG
AGCTGAGTCAGCTGACCCATGGTGTCACCCAACTGGGCTTCTATGTCCTGGACAGGGATAG
CCTCTTCATCAATGGCTATGCACCCCAGAATTTATCAATCCGGGGCGAGTACCAGATAAAT
TTCCACATTGTCAACTGGAACCTCAGTAATCCAGACCCCACATCCTCAGAGTACATCACCC
TGCTGAGGGACATCCAGGACAAGGTCACCACACTCTACAAAGGCAGTCAACTACATGACAC
ATTCCGCTTCTGCCTGGTCACCAACTTGACGATGGACTCCGTGTTGGTCACTGTCAAGGCA
TTGTTCTCCTCCAATTTGGACCCCAGCCTGGTGGAGCAAGTCTTTCTAGATAAGACCCTGA
ATGCCTCATTCCATTGGCTGGGCTCCACCTACCAGTTGGTGGACATCCATGTGACAGAAAT
GGAGTCATCAGTTTATCAACCAACAAGCAGCTCCAGCACCCAGCACTTCTACCCGAATTTC
ACCATCACCAACCTACCATATTCCCAGGACAAAGCCCAGCCAGGCACCACCAATTACCAGA
GGAACAAAAGGAATATTGAGGATGCGCTCAACCAACTCTTCCGAAACAGCAGCATCAAGAG
TTATTTTTCTGACTGTCAAGTTTCAACATTCAGGTCTGTCCCCAACAGGCACCACACCGGG
GTGGACTCCCTGTGTAACTTCTCGCCACTGGCTCGGAGAGTAGACAGAGTTGCCATCTATG
AGGAATTTCTGCGGATGACCCGGAATGGTACCCAGCTGCAGAACTTCACCCTGGACAGGAG
CAGTGTCCTTGTGGATGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAATTCTGAC
CTTCCCTTCTGGGCTGTCATCTTCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCC
TGATCTGCGGTGTCCTGGTGACCACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGTCCA
GCAACAGTGCCCAGGCTACTACCAGTCACACCTAGACCTGGAGGATCTGCAATGACTGGAA
CTTGCCGGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCAAAGAAGCTTGGCTGGGGCAG
AAATAAACCATATTGGTCG

FIGURE 52

GCCCGGGCAGGTGCTGATAGCACAGTTCTGTCCAGAGAAGGAAGGCGGAATAAACTTATTC
ATTCCCAGGAACTCTTGGGGTAGGTGTGTGTTTTTCACATCTTAAAGGCTCACAGACCCTG
CGCTGGACAAATGTTCCATTCCTGAAGGACCTCTCCAGAATCCGGATTGCTGAATCTTCCC
TGTTGCCTAGAAGGGCTCCAAACCACCTCTTGACA**ATG**GGAAACTGGGTGGTTAACCACTG
GTTTTCAGTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCC
TTCCTGAAATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACAT
TGGCCTGTGCCCGAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCC
TGTGTGTCGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTG
AGAAAGCAATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTAC
ATACAGCTATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACA
GGCCACAGATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGG
GGTTCTTGGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCA
GCGTTGCTGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGC
TACTGAGTTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATC
TTCTATATCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGG
AGAGCATGAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCG
TGACTCCCACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATC
CTTGCACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGA
AGGTTGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAA
GCGTGGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTC
CTGGAATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATA
TCCGAGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACC
AATTCCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTAT
GAAGTGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAAT
CCATCTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTA
CTGGATCTGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAA
CAGGAGATGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGAT
GGGACAGCAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGAC
AGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATA
GCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAA
AGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTA
CTTCAACAAAGAAAATTTT**TGA**GTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTT
TGTATCTTCAGTAATTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTA
TCTTCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGATAAGAATGTGCC
TCTCAAGCCTTGACTCCCTGGTATTCTTTTTTTGATTGCATTCAACTTCGTTACTTGAGCT

TCAGCAACTTAAGAACTTCTGAAGTTCTTAAAGTTCTGAAGTTCTTAAAGCCCATGGATCC
TTTCTCAGAAAAATAACTGTAAATCTTTCTGGACAGCCATGACTGTAGCAAGGCTTGATAG
CAGAGGTTTGGTGGTTCAGAGTTATACAACTAATCCCAGGTGATTTTATCAATTCCAGTGT
TACCATCTCCTGAGTTTTGGTTTGTAATCTTTTGTCCCTCCCACCCCCACAGAAGATTTCT
AAGTAGGGTGACTTTTTAAATAAAAATTTATTGAATAATTAATGATAAAACATAATAATAA
ACATAAATAATAAACAAAATTACCGAGAACCCCATCCCCATATAACACCAACAGTGTACAT
GTTTACTGTCACTTTTGATATGGTCTTATCCAGTGTGAACAGCAATTTATTATTTTTGCTC
ATCAAAAAATAAAGGATTTTCTTCTTCACTTGATGAATGAGTCTTTTGTTACTTCTTTTTG
GGCTCTCCATTTGGGAAAGAAAATCTAGAAAGTATGCTCATTGTAATTGAAATTGTATTAT
AATTGCAAGTCAGCTATACTGTCATCTTGTATTATTCTGTTAAAAGCGATTGTGAAAACTT
GAACTCCATGTTAGGATGACTTATTTAGTCAAGAGGAGCTCTTTATCACTATTAAGGGAGT
TTATTAATGTCTAGTTTAAAAAATAACAATTTATAATTTTCATAGATTTTTGTGGACTCCT
ACCTATATCGCTAAGGAATCATTCAGCTAATAACATAAGCAAAATGCCTTCATTCTCTATC
AGAATGCTTTTTATTTCTGCCATCTACCCAACCTCGGCCTACAAGGACCACACTGGCTCTG
GGATGTTAGTCGTCACAGTGAATCCCTAGAAGGAGAGAAGCCCAGGGCTCTTAAGAGAGAA
ACCTCACACTTCATTGTGCCCAGACAATTCAACTCATTATAGATTAGAAGAATGGTTGTAC
TGAATTAGCATTATATTATCTATTTTTTCTACCATTAAGTGACCACGAATGAACAAATGGT
TGTTTCTAATAGTTTAGGAACCTCTACGGCTTGAGGGATAAAACTGTTTTTGGTTTTTTTT
TGTCACTAGTCTGAGCACAATAAAGGTAGGTCATGAACTCATGGTGGCTAATCATGTTTCA
TTTATTACAACAGTAAAACAATACTGATGGTTGTTTACATAGGGAGGGATACCACTGCTGT
TGCCAATTTGTTCACCCGGATTTCAGCTAATTCAGGTTTAATAAGGCTTTGGGAAAATGCA
CAATGAAACCTTGAAGACATGATTCTTATTAGACAACTACTGACCAAATGACAACTTCATC
TTCAATGACCTTGATTTGGAATCTTGAAACTTACAACGAGTTCCTAATGAGAAAGGCATCA
CTCTTGGTCTCCATGTCTAAATATATTGACACACTCACTTCAAGATATGATGGCTGCCAAA
ATCAATATCCCAAGGCTTGAATATTTATGGGAGTAGGCATAAATTCGTGATTATATTCATG
GTAATGAATATAATAAATCCTATGCATAGCCAACATCCAAGTTGAACTGGGGATGAGCAGG
ATACACAGAAAGCTCCTCACCTGCTTTGAGCAAGTTTTTCTTGCAAAGTTGTAATATGAAA
TACCTTAATCAACAAATTTCTTATAAATGTTTATATGAGTAGAAAATCCCTCAATTTCATT
GTCACTCAAAAAGTTGATTATCT

FIGURE 53

CCCTTGGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCA
GTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGA
AATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTG
TGCCCGAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGT
CGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGC
AATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGC
TATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACA
GATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTT
GGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGC
TGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAG
TTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATA
TCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCAT
GAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCC
CACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCAC
CGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGT
GATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGC
TTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAAT
GGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGC
AGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCC
AGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGG
CTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTG
GTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTACTGGATC
TGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAACAGGAGA
TGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGATGGGACAG
CAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTG
AAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCT
CCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCT
GCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAAC
AAAGAAAATTTTTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCT
TCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGAAG

FIGURE 54

GGGCTCCAAACCACCTCTTGACA**ATG**GGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCTTGGAAGTGGATCCTTG
CACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGT
TGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGT
GGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGG
AATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCG
AGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATT
CCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAG
TGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCAT
CTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCA
TTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGA
GACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGG
AGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATAT
TCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTT**TGA**GTTATAG
GAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGT
CAGGTTTGAGCAGTCACTTTAGGA

FIGURE 55

GGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTGTGCCC
GAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGTCGCAA
TCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGCAATTG
GATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGCTATTC
ACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACAGATGG
CTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTTGGCTA
AATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGCTGGTC
TCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAGTTCAT
CCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATATCCTT
GGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCATGAATG
AGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCCCACTG
TAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCACCGGTC
ATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGTGATTA
CCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGCTTCAG
CATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAATGGCAT
CCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGCAGCAG
GGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCCAGGAT
TGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGGCTGTG
CTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTGGTACA
AATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCATTAAACTT
TGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATG
TGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCT
TATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCT
GGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTTTGAGTTATAGGAATAAGG
ACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGTCAGGTTTG
AGCAGTCACTTTAGGAAG

FIGURE 56

GGACTGCTTTGTAACTGCTAAGATTGCAGACAGAAATAGCACACAACCACTGTGAGCTGTA
TGCGATTCAGAAACCAAGACCAAATTTTGCTCACTTTCATTAATCAGTTGCTCAGATAGAA
GGAA**ATG**ACATCTGGTTCTGTCTTCTTCTACATCTTAATTTTTGGAAAATATTTTTCTCAT
GGGGGTGGACAGGATGTCAAGTGCTCCCTTGGCTATTTCCCCTGTGGGAACATCACAAAGT
GCTTGCCTCAGCTCCTGCACTGTAACGGTGTGGACGACTGCGGGAATCAGGCCGATGAGGA
CAACTGTGGAGACAACAATGGATGGTCCATGCAATTTGACAAATATTTTGCCAGTTACTAC
AAAATGACTTCCCAATATCCTTTTGAGGCAGAAACACCTGAATGTTTGGTCGGTTCTGTGC
CAGTGCAATGTCTTTGCCAAGGTCTGGAGCTTGACTGTGATGAAACCAATTTACGAGCTGT
TCCATCGGTTTCTTCAAATGTGACTGCAATGTCACTTCAGTGGAACTTAATAAGAAAGCTT
CCTCCTGATTGCTTCAAGAATTATCATGATCTTCAGAAGCTGTACCTGCAAAACAATAAGA
TTACATCCATCTCCATCTATGCTTTCAGAGGACTGAATAGCCTTACTAAACTGTATCTCAG
TCATAACAGAATAACCTTCCTGAAGCCGGGTGTTTTTGAAGATCTTCACAGACTAGAATGG
CTGATAATTGAAGATAATCACCTCAGTCGAATTTCCCCACCAACATTTTATGGACTAAATT
CTCTTATTCTCTTAGTCCTGATGAATAACGTCCTCACCCGTTTACCTGATAAACCTCTCTG
TCAACACATGCCAAGACTACATTGGCTGGACCTTGAAGGCAACCATATCCATAATTTAAGA
AATTTGACTTTTATTTCCTGCAGTAATTTAACTGTTTTAGTGATGAGGAAAAACAAAATTA
ATCACTTAAATGAAAATACTTTTGCACCTCTCCAGAAACTGGATGAATTGGATTTAGGAAG
TAATAAGATTGAAAATCTTCCACCGCTTATATTCAAGGACCTGAAGGAGCTGTCACAATTG
AATCTTTCCTATAATCCAATCCAGAAAATTCAAGCAAACCAATTTGATTATCTTGTCAAAC
TCAAGTCTCTCAGCCTAGAAGGGATTGAAATTTCAAATATCCAACAAAGGATGTTTAGACC
TCTTATGAATCTCTCTCACATATATTTTAAGAAATTCCAGTACTGTGGGTATGCACCACAT
GTTCGCAGCTGTAAACCAAACACTGATGGAATTTCATCTCTAGAGAATCTCTTGGCAAGCA
TTATTCAGAGAGTATTTGTCTGGGTTGTATCTGCAGTTACCTGCTTTGGAAACATTTTTGT
CATTTGCATGCGACCTTATATCAGGTCTGAGAACAAGCTGTATGCCATGTCAATCATTTCT
CTCTGCTGTGCCGACTGCTTAATGGGAATATATTTATTCGTGATCGGAGGCTTTGACCTAA
AGTTTCGTGGAGAATACAATAAGCATGCGCAGCTGTGGATGGAGAGTACTCATTGTCAGCT
TGTAGGATCTTTGGCCATTCTGTCCACAGAAGTATCAGTTTTACTGTTAACATTTCTGACA
TTGGAAAAATACATCTGCATTGTCTATCCTTTTAGATGTGTGAGACCTGGAAAATGCAGAA
CAATTACAGTTCTGATTCTCATTTGGATTACTGGTTTTATAGTGGCTTTCATTCCATTGAG
CAATAAGGAATTTTTCAAAAACTACTATGGCACCAATGGAGTATGCTTCCCTCTTCATTCA
GAAGATACAGAAAGTATTGGAGCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATT
TGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAG
TGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGT
TTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTC
TTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCT

GCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAA
ATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTCAGA
AAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGA
GTTAATGAAGCCGGACCTTTTCACATACCCCTGTGAAATGTCACTGATTTCTCAATCAACG
AGACTCAATTCCTATTCA**TGA**

FIGURE 57

TAGAGATCCCTCGACCTCGACCCACGCGTCCGAGGAAAGAAAAAAAGAGGAATGGAAAGAG
ACAGAGAAAGGAAATGGGAGTGGAAGGAGGGAGGACTGCTTTGTAACTGCTAAGATTGCAG
ACAGAAATAGCACACAACCACTGTGAGCTGTATGCGATTCAGAAACCAAGACCAAATTTTG
CTCACTTTCATTAATCAGTTGCTCAGATAGAAGGAAATGACATCTGGTTCTGTCTTCTTCT
ACATCTTAATTTTTGGAAAATATTTTTCTCATGGGGGTGGACAGGATGTCAAGTGCTCCCT
TGGCTATTTCCCCTGTGGGAACATCACAAAGTGCTTGCCTCAGCTCCTGCACTGTAACGGT
GTGGACGACTGCGGGAATCAGGCCGATGAGGACAACTGTGGAGACAACAATGGATGGTCTC
TGCAATTTGACAAATATTTTGCCAGTTACTACAAAATGACTTCCCAATATCCTTTTGAGGC
AGAAACACCTGAATGTTTGGTCGGTTCTGTGCCAGTGCAATGTCTTTGCCAAGGTCTGGAG
CTTGACTGTGATGAAACCAATTTACGAGCTGTTCCATCGGTTTCTTCAAATGTGACTGCAA
TGTCACTTCAGTGGAACTTAATAAGAAAGCTTCCTCCTGATTGCTTCAAGAATTATCATGA
TCTTCAGAAGCTGTACCTGCAAAACAATAAGATTACATCCATCTCCATCTATGCTTTCAGA
GGACTGAATAGCCTTACTAAACTGTATCTCAGTCATAACAGAATAACCTTCCTGAAGCCGG
GTGTTTTTGAAGATCTTCACAGACTAGAATGGCTGATAATTGAAGATAATCACCTCAGTCG
AATTTCCCCACCAACATTTTATGGACTAAATTCTCTTATTCTCTTAGTCCTGATGAATAAC
GTCCTCACCCGTTTACCTGATAAACCTCTCTGTCAACACATGCCAAGACTACATTGGCTGG
ACCTTGAAGGCAACCATATCCATAATTTAAGAAATTTGACTTTTATTTCCTGCAGTAATTT
AACTGTTTTAGTGATGAGGAAAAACAAAATTAATCACTTAAATGAAAATACTTTTGCACCT
CTCCAGAAACTGGATGAATTGGATTTAGGAAGTAATAAGATTGAAAATCTTCCACCGCTTA
TATTCAAGGACCTGAAGGAGCTGTCACAATTGAATCTTTCCTATAATCCAATCCAGAAAAT
TCAAGCAAACCAATTTGATTATCTTGTCAAACTCAAGTCTCTCAGCCTAGAAGGGATTGAA
ATTTCAAATATCCAACAAAGGATGTTTAGACCTCTTATGAATCTCTCTCACATATATTTTA
AGAAATTCCAGTACTGTGGGTATGCACCACATGTTCGCAGCTGTAAACCAAACACTGATGG
AATTTCATCTCTAGAGAATCTCTTGGCAAGCATTATTCAGAGAGTATTTGTCTGGGTTGTA
TCTGCAGTTACCTGCTTTGGAAACATTTTTGTCATTTGCATGCGACCTTATATCAGGTCTG
AGAACAAGCTGTATGCCATGTCAATCATTTCTCTCTGCTGTGCCGACTGCTTAATGGGAAT
ATATTTATTCGTGATCGGAGGCTTTGACCTAAAGTTTCGTGGAGAATACAATAAGCATGCG
CAGCTGTGGATGGAGAGTACTCATTGTCAGCTTGTAGGATCTTTGGCCATTCTGTCCACAG
AAGTATCAGTTTTACTGTTAACATTTCTGACATTGGAAAAATACATCTGCATTGTCTATCC
TTTTAGATGTGTGAGACCTGGAAAATGCAGAACAATTACAGTTCTGATTCTCATTTGGATT
ACTGGTTTTATAGTGGCTTTCATTCCATTGAGCAATAAGGAATTTTTCAAAAACTACTATG
GCACCAATGGAGTATGCTTCCCTCTTCATTCAGAAGATACAGAAAGTATTGGAGCCCAGAT
TTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC
TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATC
AAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATT

ATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGT
ACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTC
TCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGTATAACTACAG
ACAAAGAAAATCTATGGACAGCAAAGGTCAGAAAACATATGCTCCATCATTCATCTGGGTG
GAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAATGAAGCCGGACCTTTTCACATACC
CCTGTGAAATGTCACTGATTTCTCAATCAACGAGACTCAATTCCTATTCATGACTGACTCT
GAAATTCATTTCTTCGCAGAGAATACTGTGGGGGTGCTTCATGAGGGATTTACTGGTATGA
AATGAATACCACAAAATTAATTTATAATAATAGCTAAGATAAATATTTTACAAGGACATGA
GGAAAAATAAAAATGACTAATGCTCTTACAAAGGGAAGTAATTATATCAATAATGTATATA
TATTAGTAGACATTTTGCATAAGAAATTAAGAGAAATCTACTTCAGTAACATTCATTCATT
TTTCTAACATGCATTTATTGAGTACCCACTACTATGTGCATAGCATTGCAATATAGTCCTG
GAAGTAGACAGTGCAGAACCTTTCAATCTGTAGATGGTGTTTAATGACAAAAGACTATACA
AAGTCCATCTGCAGTTCCTAGTTTAAAGTAGAGCTTTACCTGTCATGTGCATCAGCAAGAA
TCATAGGCACTTTTAAATAAAGGTTTAAAGTTTTGGAAT

FIGURE 58

GATCCATTGTGCTGGCGCGGATTCTTTATCACTGGAAGGAGGAGGAGGGGGAGAGTCGCTC

CCGCCGGGCGAGC**ATG**GGGCGCCTGGCCTCGAGGCCGCTGCTGCTGGCGCTCCTGTCGTTG

GCTCTTTGCCGAGGGCGTGTGGTGAGAGTCCCCACAGCGACCCTGGTTCGAGTGGTGGGCA

CTGAGCTGGTCATCCCCTGCAACGTCAGTGACTATGATGGCCCCAGCGAGCAAAACTTTGA

CTGGAGCTTCTCATCTTTGGGGAGCAGCTTTGTGGAGCTTGCAAGCACCTGGGAGGTGGGG

TTCCCAGCCCAGCTGTACCAGGAGCGGCTGCAGAGGGGCGAGATCCTGTTAAGGCGGACTG

CCAACGACGCCGTGGAGCTCCACATAAAGAACGTCCAGCCTTCAGACCAAGGCCACTACAA

ATGTTCAACCCCCAGCACAGATGCCACTGTCCAGGGAAACTATGAGGACACAGTGCAGGTT

AAAGTGCTGGCCGACTCCCTGCACGTGGGCCCCAGCGCGCGGCCCCCGCCGAGCCTGAGCC

TGCGGGAGGGGGAGCCCTTCGAGCTGCGCTGCACCGCCGCCTCCGCCTCGCCGCTGCACAC

GCACCTGGCGCTGCTGTGGGAGGTGCACCGCGGCCCGGCCAGGCGGAGCGTCCTCGCCCTG

ACCCACGAGGGCAGGTTCCACCCGGGCCTGGGGTACGAGCAGCGCTACCACAGTGGGGACG

TGCGCCTCGACACCGTGGGCAGCGACGCCTACCGCCTCTCAGTGTCCCGGGCTCTGTCTGC

CGACCAGGGCTCCTACAGGTGTATCGTCAGCGAGTGGATCGCCGAGCAGGGCAACTGGCAG

GAAATCCAAGAAAAGGCCGTGGAAGTTGCCACCGTGGTGATCCAGCCGACAGTTCTGCGAG

CAGCTGTGCCCAAGAATGTGTCTGTGGCTGAAGGAAAGGAACTGGACCTGACCTGTAACAT

CACAACAGACCGAGCCGATGACGTCCGGCCCGAGGTGACGTGGTCCTTCAGCAGGATGCCT

GACAGCACCCTACCTGGCTCCCGCGTGTTGGCGCGGCTTGACCGTGATTCCCTGGTGCACA

GCTCGCCTCATGTTGCTTTGAGTCATGTGGATGCACGCTCCTACCATTTACTGGTTCGGGA

TGTTAGCAAAGAAAACTCTGGCTACTATTACTGCCACGTGTCCCTGTGGGCACCCGGACAC

AACAGGAGCTGGCACAAAGTGGCAGAGGCCGTGTCTTCCCCAGCTGGTGTGGGTGTGACCT

GGCTAGAACCAGACTACCAGGTGTACCTGAATGCTTCCAAGGTCCCCGGGTTTGCGGATGA

CCCCACAGAGCTGGCATGCCGGGTGGTGGACACGAAGAGTGGGGAGGCGAATGTCCGATTC

ACGGTTTCGTGGTACTACAGGATGAACCGGCGCAGCGACAATGTGGTGACCAGCGAGCTGC

TTGCAGTCATGGACGGGGACTGGACGCTAAAATATGGAGAGAGGAGCAAGCAGCGGGCCCA

GGATGGAGACTTTATTTTTTCTAAGGAACATACAGACACGTTCAATTTCCGGATCCAAAGG

ACTACAGAGGAAGACAGAGGCAATTATTACTGTGTTGTGTCTGCCTGGACCAAACAGCGGA

ACAACAGCTGGGTGAAAAGCAAGGATGTCTTCTCCAAGCCTGTTAACATATTTTGGGCATT

AGAAGATTCCGTGCTTGTGGTGAAGGCGAGGCAGCCAAAGCCTTTCTTTGCTGCCGGAAAT

ACATTTGAGATGACTTGCAAAGTATCTTCCAAGAATATTAAGTCGCCACGCTACTCTGTTC

TCATCATGGCTGAGAAGCCTGTCGGCGACCTCTCCAGTCCCAATGAAACGAAGTACATCAT

CTCTCTGGACCAGGATTCTGTGGTGAAGCTGGAGAATTGGACAGATGCATCACGGGTGGAT

GGCGTTGTTTTAGAAAAAGTGCAGGAGGATGAGTTCCGCTATCGAATGTACCAGACTCAGG

TCTCAGACGCAGGGCTGTACCGCTGCATGGTGACAGCCTGGTCTCCTGTCAGGGGCAGCCT

TTGGCGAGAAGCAGCAACCAGTCTCTCCAATCCTATTGAGATAGACTTCCAAACCTCAGGT

CCTATATTTAATGCTTCTGTGCATTCAGACACACCATCAGTAATTCGGGGAGATCTGATCA
AATTGTTCTGTATCATCACTGTCGAGGGAGCAGCACTGGATCCAGATGACATGGCCTTTGA
TGTGTCCTGGTTTGCGGTGCACTCTTTTGGCCTGGACAAGGCTCCTGTGCTCCTGTCTTCC
CTGGATCGGAAGGGCATCGTGACCACCTCCCGGAGGGACTGGAAGAGCGACCTCAGCCTGG
AGCGCGTGAGTGTGCTGGAATTCTTGCTGCAAGTGCATGGCTCCGAGGACCAGGACTTTGG
CAACTACTACTGTTCCGTGACTCCATGGGTGAAGTCACCAACAGGTTCCTGGCAGAAGGAG
GCAGAGATCCACTCCAAGCCCGTTTTTATAACTGTGAAGATGGATGTGCTGAACGCCTTCA
AGTATCCCTTGCTGATCGGCATCGGTCTGTCCACGGTCATCGGGCTCCTGTCCTGTCTCAT
CGGGTACTGCAGCTCCCACTGGTGTTGTAAGAAGGAGGTTCAGGAGACACGGCGCGAGCGC
CGCAGGCTCATGTCGATGGAGATGGACTAGGCTGGCCCGGGAGGGGAGTGACAGAGGGACG
TTCTAGGAGCAATTGGGGCAAGAAGAGGACAGTGATATTTTAAAACAAAGTGTGTTACACT
AAAAACCAGTCCTCTCTAATCTCAGGTGGGACTTGGCGCTCTCTCTTTTCTGCATGTCAAG
TTCTGAGCGCGGACATGTTTACCAGCACACGGCTCTTCTTCCCACGGCACTTTCTGATGTA
ACAATCGAGTGTGTGTTTTCCCAACTGCCAGTGATAAAGAATCCGCGCCAGCACAATGGAT
CTCGAGGGATCTTCCATACCTACCAGTTCTGCGCCTGCAGGTC

FIGURE 59

CCCACGCGTCCGCGGTAAATCACTTCTTGGAGGTGCCCTGCACGCCGGTCCTGGGAGCAGG
CGGCCTCCCGGGGGTGCGGGAGCCCCACTCCTCCGTGGTGTGTTCCATTTGCTTCCCACAT
CTGGAGGAGCTGACGTGCCAGCCTCCCCCAGCACCACCCAGGGACGGGAGGC**ATG**AGCCGG
TCAAGGCACCTGGGCAAAATCCGGAAGCGTCTGGAAGATGTCAAGAGCCAGTGGGTCCGGC
CAGCCAGGGCTGACTTTAGTGACAACGAGAGTGCCCGGCTGGCCACGGACGCCCTCTTGGA
TGGGGGTTCTGAAGCCTACTGGCGGGTGCTCAGCCAGGAAGGCGAGGTGGACTTCTTGTCC
TCGGTGGAGGCCCAGTACATCCAGGCCCAGGCCAGGGAGCCCCCGTGTCCCCCAGACACCC
TGGGAGGGGCGGAAGCAGGCCCTAAGGGACTGGACTCCAGCTCCCTACAGTCCGGCACCTA
CTTCCCTGTGGCCTCAGAGGGCAGCGAGCCGGCCCTACTGCACAGCTGGGCCTCAGCTGAG
AAGCCCTACCTGAAGGAAAAATCCAGCGCCACTGTGTACTTCCAGACCGTCAAGCACAACA
ACATCAGAGACCTCGTCCGCCGCTGCATCACCCGGACTAGCCAGGTCCTGGTCATCCTGAT
GGATGTGTTCACGGATGTGGAGATCTTCTGTGACATTCTAGAGGCAGCCAACAAGCGTGGG
GTGTTCGTTTGTGTGCTCCTGGACCAGGGAGGTGTGAAGCTCTTCCAGGAGATGTGTGACA
AAGTCCAGATCTCTGACAGTCACCTCAAGAACATTTCCATCCGGAGTGTGGAAGGAGAGAT
ATACTGTGCCAAGTCAGGCAGGAAATTCGCTGGCCAAATCCGGGAGAAGTTCATCATCTCG
GACTGGAGATTTGTCCTGTCTGGATCTTACAGCTTCACCTGGCTCTGCGGACACGTGCACC
GGAACATCCTCTCCAAGTTCACAGGCCAGGCGGTGGAGCTGTTTGACGAGGAGTTCCGCCA
CCTCTACGCCTCCTCCAAGCCTGTGATGGGCCTGAAGTCCCCGCGGCTGGTCGCCCCCGTC
CCGCCCGGAGCAGCCCCGGCCAATGGCCGCCTTAGCAGCAGCAGTGGCTCCGCCAGTGACC
GCACGTCCTCCAACCCCTTCAGCGGCCGCTCGGCAGGCAGCCACCCCGGTACCCGAAGTGT
GTCCGCGTCTTCAGGGCCCTGTAGCCCCGCGGCCCCACACCCGCCTCCACCGCCCCGGTTC
CAGCCCCACCAAGGCCCTTGGGGAGCCCCGAGTCCCCAGGCCCACCTCTCCCCGCGGCCCC
ACGACGGCCCGCCCGCCGCTGTCTACAGCAACCTGGGGGCCTACAGGCCCACGCGGCTGCA
GCTGGAGCAGCTGGGCCTGGTGCCGAGGCTGACTCCAACCTGGAGGCCCTTCCTGCAGGCC
TCCCCTCACTTC**TGA**AGGTCCCATCCCCTGCTGCCCTCCGCAGGCCCAGGGCTGGGCACTC
CCTGAGACCCAAAGACCCACCTCAACGACGAGTGGCGTTGAGCCACTTCCCTTTGAAAAGA
CACTCAAAATCACTGCCATGGTTCAATGTTCCCAGGCCCCAGGCCATCCACTTGCCGGCCC
CCACCAGTTCTTGGGTTCCCCGCTCTAGTTTGACCTGTGCAGCACATTCCAGAAGGTTCCA
GGGGAGGTTGTGGGGCAGCTAGAGGACAAAATCATGAAAACAGAGTCCCTGTCTTCCAGAG
ATCATCCGGGGCTTTAATATTAATGGCCCCCAAAACTCCGTAAGAAGCAGGAAATGCAGCC
CAAGTTTTACAAATGGGTAAACAGAGGCACTGAGAGATAGATGGTAGTTTGGTACTTCTGG
TTCCCAGTGCCCAGGAATGGTCCACTCCCAAGAAATTCAGGAAAGAAAGACTGAGGAGAAG
GTGTGGGAACATTCTGGATGTTTCGGGAGAGTTGGGGAAACTCCTCCTCTTAGGAAAGGCT
AATACTAGGGTATCCTTGGGCCCAATGAATTAGGGGTGAGGCCCCAGAACCCGTTATCTAT
GAGTTGTATGGGGGAGCCATCTGAAGCTGTAGCCACCAGGGATGCAGCTAGCTGAGGAGTT

TGGGGTGTTGGGTTGGACAAGGCAGGTTAGTAGACTCAGATTCTTGCTTCAAAGAGCCTTG
GGCTGGCCTGGAGGTCCCTGGAGTCTAGACTGGACCTAGGAGCTTGAGTTGTCAGGGGCCA
GGACTGGCCCCACTGCAGTGCCCAGGCCAGTCTTGAGCAGCAGGGAGGGCTCAGCTGTCCC
CAGATCCAGGTGCCTCTGACCAGCCTGGTCACCTCCTGAGGAATAAATGCTGAACCTCACA
AGCCCCATCATTCATTTCTTCTCAATTCACAGTGCCCCTCTTTGTTTCTGGGGTGGAACTA
GGTCCTGAGGGCACAGCCTAGCTGAGTGCAAAGAAATATAGGATGCTTAGAAAGCATACAG
GAGGGGCCAGGCGTGGTGGCTCATGCCTGTAATCCCAGAACTTTGGGATGCCAAGGTGGTT
GGATTACCTGAGATCAGGTGGATTACCTGGTCTCGAGACCAGCCTGACCAATATGGTGAAA
CCCCGTCTCTACTAAAAATACAAAAATTAGGCTGAGACAGGAGAATTGCTTGAACCCAGGA
AGCAGAGGTTGCAATGAGCTGAGATTGCATCACTGCACTCCAGCATGGGCAACAAAGCAAG
ACTCCGTCACAGAAAAAAAAAAAAA

FIGURE 60

ATGGCTCTGCCGCAGCCGCCGCCGCCCGCCGGGGCGCGTGACCCGGTGACCCCGGAGCACT
GGGCATCCGGACCCCCCTCGCAGGCGCAGCCGCTATTAAGGCAGGAGGCTAAAGAGGAGGA
GGAAGGGGAGGAGACAGGCGTCCAGGGCGCCTGGGGAACCGGCACGGCGGAGCAGCGGCGG
CGGGGCTGGGGGGAGGCCGCCGAGTCGGCCGCGGCCGAGGAGGGGCAGGCGGAGGTCGGGG
GCGCCGCGGCCGCGGGGTCAGGTTCCCCGGCGGGAGGCGCAGGTGGCGGCCTGGGGAGCTG
GCGGCCGCTCCTCGCCTGGCTGCAGCGGCGGCAGCCCCAGTGCTGCCCCTGTGCGGCGCCC
CTTTCCCGCTCCGCCGCGCACTGTTGTCATGGAGGAACCAAGATGGCGGCTCTGGCCTACA
ACCTGGGCAAGCGGGAGATCAACCACTACTTCAGCGTGAGGAGCGCCAAGGTGCTGGCGCT
GGTGGCCGTGCTGCTGCTCGCAGCGTGCCACCTCGCCTCCCGCCGCTACCGAGGCAATGAT
TCGTGTGAATACCTTCTCTCAAGTGGCAGATTTCTTGGAGAGAAAGTTTGGCAACCTCACA
GTTGTATGATGCATAAATACAAAATCAGTGAAGCAAAGAACTGCCTTGTAGATAAACATAT
TGCATTTATTGGAGATTCCAGAATTCGTCAATTGTTTTATTCTTTTGTAAAAATAATTAAT
CCCCAATTCAAAGAAGAAGGAAATAAGCATGAAAACATTCCTTTTGAAGACAAGACTGCAT
CAGTTAAAGTGGATTTTCTGTGGCATCCTGAAGTTAATGGTTCTATGAAACAGTGTATCAA
AGTGTGGACTGAGGATTCCATTGCAAAGCCACATGTGATTGTAGCAGGAGCTGCCACATGG
TCCATCAAGATTCACAATGGTAGCAGTGAAGCGCTTTCTCAATATAAAATGAACATCACCT
CCATAGCACCACTTTTAGAAAAATTGGCAAAGACTAGTGATGTTTATTGGGTCTTACAAGA
TCCTGTTTATGAAGATCTATTAAGTGAAAATAGGAAGATGATCACTAATGAGAAGATAGAT
GCTTACAATGAAGCTGCAGTCAGTATTTTGAATAGTAGCACCAGAAATTCTAAATCAAATG
TTAAGATGTTCAGTGTTTCCAAATTAATTGCTCAAGAAACCATCATGGAATCTTTGGATGG
CTTACATCTTCCTGAATCGAGCAGAGAAACTACTGCAATGATTCTTATGAATGTGTATTGC
AATAAGATTTTGAAGCCTGTAGATGGGTCCTGTTGTCAACCTCGGCCTCCTGTTACTCTCA
TACAGAAGCTAGCTGCTTGTTTTTTCACTTTATCTATTATCGGATATTTAATTTTTTACAT
AATTCATCGTAATGCTCATCGGAAGAATAAGCCGTGTACTGATTTGGAAAGTGGAGAGGAA
AAGAAAAATATTATCAATACCCCTGTGTCTTCATTAGAAATACTTTTACAATCTTTCTGCA
AACTTGGCCTGATTATGGCATATTTCTATATGTGTGACCGTGCAAATCTGTTCATGAAGGA
AAACAAATTTTATACACATTCATCTTTCTTTATTCCAATTATCTACATTTTGGTTTTGGGA
GTATTTTATAATGAAAATACTAAAGAGACTAAAGTATTAAATAGAGAACAAACAGACGAAT
GGAAAGGCTGGATGCAACTTGTGATTTTGATTTATCACATTTCTGGAGCAAGTACATTTTT
GCCTGTATACATGCACATTCGAGTTCTGGTTGCTGCATATTTATTTCAGACAGGGTATGGG
CATTTCTCATACTTTTGGATAAAAGGAGATTTTGGAATCTATAGAGTATGTCAGGTTTTAT
TTCGTCTCAATTTCCTGGTAGTGGTGTTATGTATAGTAATGGATCGACCTTATCAATTCTA
TTACTTTGTCCCCTTGGTCACTGTATGGTTCATGGTCATATATGTTACTTTAGCACTATGG
CCACAAATAATCCAAAAAAAAGCAAACGGAAATTGTTTCTGGCATTTTGGCTTACTGTTGA
AACTAGGCTTTTTGCTGTTATTCATATGTTTTTTGGCATATTCTCAGGGTGCATTTGAGAA

GATCTTTTCTCTTTGGCCATTGTCCAAGTGTTTTGAACTGAAAGGGAATGTATATGAATGG
TGGTTCAGATGGAGGTTAGACCGTTATGTAGTTTTCCACGGAATGCTGTTTGCTTTTATTT
ATCTGGCTTTGCAGAAGCGTCAAATACTTTCTGAAGGAAAGGGTGAACCTCTTTTTTCAAA
CAAAATTTCAAATTTTCTGTTGTTTATTTCAGTAGTTTCTTTCTTGACCTATTCCATCTGG
GCTAGCAGTTGTAAAAACAAAGCAGAGTGCAATGAACTCCATCCGTCTGTTTCTGTGGTAC
AGATTTTAGCCTTCATCCTAATAAGAAACATCCCTGGATATGCCCGTTCAGTTTACAGTTC
ATTTTTTGCTTGGTTTGGAAAAATTTCATTAGAGCTATTTATTTGCCAGTATCACATATGG
CTGGCAGCGGACACAAGGGGTATCTTGGTACTGATACCTGGAAACCCTATGCTCAACATCA
TTGTCAGCACTTTCATATTTGTTTGTGTGGCACATGAAATTTCTCAGATCACTAATGATCT
TGCACAGATTATTATTCCTAAAGATAACTCATCTCTCTTGAAAAGGTTGGCATGTATAGCT
GCATTTTTTTGTGGACTCCTCATCTTATCATCCATTCAAGATAAATCAAAACATTAGGTTC
CAAAAATTCTAAAAAACCTAAACTCTTCAGGCTACCTTTGTGTGTCTCTAGAAGAGAAAAG
CATCTATCTGGAGATATAAATGTGTWTGTAAATATAAACGTTTGTGGCAAGAGGACAGTTC
TGTGACATCTGTTGAACATATGTGGTTGTATATATTGGAAATGTACATATCCAATATGAAA
TACTAAAACAAACAAACAAACAAAAAACCAGAATGCATTGTATAGGATTGCATGTGAAGTC
TTTTCTACTGAATCTATATTTCCATTTGTAAGTGATTTTAAGTTAACATATGAAGGCAGGG
AAATGATTACCTTTCCAGTAAAAAGTATAGATAATTTAATTAACTTAGTGACACCA:CCAA
GTGTTTTGATATAACTAAATTTGTGGTAATAAGACTGTCTGCACCTGTATTCATTGTGGAA
CTTCCTCTTTCATTGGAAACTTTCTTACTCAAGAATGACGGCAGTATTGTTTTCTTATATG
TGCAATGAAGTGGAATGATAAACAGTATGCCTTTAATTTATATGTGTTCTTGTTCTGATGT
TGTTTCCTGAAATGATTTTTCTTCCTAACTGTGGTTTTCGGGTATGCAAGCCTAAATCTTT
GTACACTTTGTCTCACAGAATAGTTCTGAGGCTCCATGACAGGGTTTTGTCATTGTTGATG
TTATTGTTGCTTCGTTTTATAAAAAAGCCAAAATTTTTTTTCCAATCCAAACGTTCACCTG
TTTCCTTTCCTCAAGCTATACCAGTGTAATACCAGTTACCCTGTGGATCCATTTAATATGT
TATCCCCACTAATTAATTTTCGTATATTATTTCCAATATTTGGAAAGCTCTTTATAGCCAT
TTGGTATTTCCTATTACCCACCTCCTATTTTAAATATTTATCAGTCTAAACTTGTGCAGTG
TAGTAAACATGCAAGTTGTTACGATTGAGCTGTATTACCATAAGTAGAATTTTAAGTAAAC
TGGTGAATTTGGGCAATAAATGTTTTTGCTTTTTGTTTGATTTTTTTTTACAAGCTAACTG
TTAGAGGTATACATTTATTTATCTGTTGTACAGATTTGATTATGATTTTAATGTTTGAAAG
ATTGCACTTGTTTGCTTTTACTATATGTGGGGTAAAATATATTTTCTGTTCACAGTATATG
AAAATATGGAGTAATTTAAACAGTAAATAAACATTCTGTGGATGCTTATTTTTGTATTGGC
AAAGTATCAATTAAACTATATGTGTTCTTTTTCAA

FIGURE 61A

GAAAGACATACACACTTCATGTAATGCTACCTGCAAGTCTCCCTAGAAAAGCAGTTTTTGT
AGGTGAAAACAATGAAGCCAGGTAATATTGCAAGGAGGCTGTAATTTTAGCAGACCTACCA
ACAACACTGATGTAGGAAGCTCATTATTTTAATTTCTGGAGCCTTTTAATTTTTTCTTTAG
AAAGTGTATAAATAATTGCAGTGCTGCTTTGCTTCCAAAACTGGGCAGTGAGTTCAACAAC
AACGACAACAACAGCCGCAGCTCATCCTGGCCGTCATGGAGTTTCTTGAAAGAACGTATCT
TGTGAATGATAAAGCTGCCAAGATGTATGCTTTCACACTAGAAAGAAGGAGCTGCAAATGA
ACACTTCATAGCAATGTGGAACTCCAACAGAAACCGGTGAATAAAGATCAGTGTCCCAGAG
AGAGACCAGAGGAGCTGGAGTCAGGAGGC**ATG**TACCACTGCCACAGTGGCTCCAAGCCCAC
AGAAAAGGGGGCGAATGAGTACGCCTATGCCAAGTGGAAACTCTGTTCTGCTTCAGCAATA
TGCTTCATTTTCATGATTGCAGAGGTCGTGGGTGGGCACATTGCTGGGAGTCTTGCTGTTG
TCACAGATGCTGCCCACCTCTTAATTGACCTGACCAGTTTCCTGCTCAGTCTCTTCTCCCT
GTGGTTGTCATCGAAGCCTCCCTCTAAGCGGCTGACATTTGGATGGCACCGAGCAGAGATC
CTTGGTGCCCTGCTCTCCATCCTGTGCATCTGGGTGGTGACTGGCGTGCTAGTGTACCTGG
CATGTGAGCGCCTGCTGTATCCTGATTACCAGATCCAGGCGACTGTGATGATCATCGTTTC
CAGCTGCGCAGTGGCGGCCAACATTGTACTAACTGTGGTTTTGCACCAGAGATGCCTTGGC
CACAATCACAAGGAAGTACAAGCCAATGCCAGCGTCAGAGCTGCTTTTGTGCATGCCCTTG
GAGATCTATTTCAGAGTATCAGTGTGCTAATTAGTGCACTTATTATCTACTTTAAGCCAGA
GTATAAAATAGCCGACCCAATCTGCACATTCATCTTTTCCATCCTGGTCTTGGCCAGCACC
ATCACTATCTTAAAGGACTTCTCCATCTTACTCATGGAAGGTGTGCCAAAGAGCCTGAATT
ACAGTGGTGTGAAAGAGCTTATTTTAGCAGTCGACGGGGTGCTGTCTGTGCACAGCCTGCA
CATCTGGTCTCTAACAATGAATCAAGTAATTCTCTCAGCTCATGTTGCTACAGCAGCCAGC
CGGGACAGCCAAGTGGTTCGGAGAGAAATTGCTAAAGCCCTTAGCAAAAGCTTTACGATGC
ACTCACTCACCATTCAGATGGAATCTCCAGTTGACCAGGACCCCGACTGCCTTTTCTGTGA
AGACCCCTGTGAC**TAG**CTCAGTCACACCGTCAGTTTCCCAAATTTGACAGGCCACCTTCAA
ACATGCTGCTATGCAGTTTCTGCATCATAGAAAATAAGGAACCAAAGGAAGAAATTCATGT
CATGGTGCAATGCACATTTTATCTATTTATTTAGTTCCATTCACCATGAAGGAAGAGGCAC
TGAGATCCATCAATCAATTGGATTATATACTGATCAGTAGCTGTGTTCAATTGCAGGAATG
TGTATATAGATTATTCCTGAGTGGAGCCGAAGTAACAGCTGTTTGTAACTATCGGCAATAC
CAAATTCATCTCCCTTCCAATAATGCATCTTGAGAACACATAGGTAAATTTGAACTCAGGA
AAGTCTTACTAGAAATCAGTGGAAGGGACAAATAGTCACAAAATTTTACCAAAACATTAGA
AACAAAAAATAAGGAGAGCCAAGTCAGGAATAAAAGTGACTCTGTATGCTAACGCCACATT
AGAACTTGGTTCTCTCACCAAGCTGTAATGTGATTTTTTTTTCTACTCTGAATTGGAAATA
TGTATGAATATACAGAGAAGTGCTTACAACTAATTTTTATTTACTTGTCACATTTTGGCAA
TAAATCCCTCTTATTTCTAAATTCTAACTTGTTTATTTCAAAACTTTATATAATCACTGTT
CAAAAGGAAATATTTTCACCTACCAGAGTGCTTAAACACTGGCACCAGCCAAAGAATGTGG

TTGTAGAGACCCAGAAGTCTTCAAGAACAGCCGACAAAAACATTCGAGTTGACCCCACCAA
GTTGTTGCCACAGATAATTTAGATATTTACCTGCAAGAAGGAATAAAGCAGATGCAACCAA
TTCATTCAGTCCACGAGCATGATGTGAGCACTGCTTTGTGCTAGACATTGGGCTTAGCATT
GAAACTATAAAGAGGAATCAGACGCAGCAAGTGCTTCTGTGTTCTGGTAGCAACTCAACAC
TATCTGTGGAGAGTAAACTGAAGATGTGCAGGCCAACATTCTGGAAATCCTATGTCACTGG
GTTTGGTTTGGAAACCTGGACTTCTGCATTTTTAAAAGTTACCCAGAGATGCTTCTAAAGA
TGAGCCATAGTCTAGAAGATTGTCAACCACAGGAGTTCATTGAGTGGGACAGCTAGAYACA
TACATTGGCAGCTACAATAGTATCATGAATTGCAATGATGTAGTGGGGTATAAAAGGAAAG
CGATGGATATTGCCGGATGGGCATGGCCAGTGATGTTTCACGTCATTGAGGTGACAGCTCT
GCTGGACTTTGAATTACATATGGAGGCTCTCCAGGAAGACGAAGAAGAGAAGGACATTCTA
GGCAAAAAGAAGACTAGGCACAAGGCACACTTATGTTTGTCTGTTAGCTTTTAGTTGAAAA
AGCAAAATACATGATGCAAAGAAACCTCTCCACGCTGTGATTTTTAAAACTACATACTTTT
TGCAACTTTATGGTTATGAGTATTGTAGAGAACAGGAGATAGGTCTTAGATGATTTTTATG
TTGTTGTCAGACTCTAGCAAGGTACTAGAAACCTAGCAGGCATTAATAATTGTTGAGGCAA
TGACTCTGAGGCTATATCTGGGCCTTGTCATTATTTATCATTTATATTTGTATTTTTTTCT
GAAATTTGAGGGCCAAGAAAACATTGACTTTGACTGAGGAGGTCACATCTGTGCCATCTCT
GCAAATCAATCAGCACCACTGAAATAACTACTTAGCATTCTGCTGAGCTTTCCCTGCTCAG
TAGAGACAAATATACTCATCCCCCACCTCAGTGAGCTTGTTTAGGCAACCAGGATTAGAGC
TGCTCAGGTTCCCAACGTCTCCTGCCACATCGGGTTCTCAAAATGGAAAGAATGGTTTATG
CCAAATCACTTTTCCTGTCTGAAGGACCACTGAATGGTTTTGTTTTTCCATATTTTGCATA
GGACGCCCTAAAGACTAGGTGACTTGGCAAACACACAAGTGTTAGTATAATTCTTTGCTTC
TGCTTCTTTTTGAAAATCATGTTTAGATTTGATTTTAAGTCAGAAATTCACTGAATGTCAG
GTAATCATTATGGAGGGAGATTTGTGTGTCAACCAAAGTAATTGTCCCATGGCCCCAGGGT
ATTTCTGTTGTTTCCCTGAAATTCTGCTTTTTTAGTCAGCTAGATTGAAAACTCTGAACAG
TAGATGTTTATATGGCAAAATGCAAGACAATCTATAAGGGAGATTTTAAGGATTTTGAGAT
GAAAAAACAGATGCTACTCAGGGGCTTTATGGACCATCCATCAATTCTGAAGTTCTGACTC
TCCCATTACCCTTTCCCTGGTGTGGTCAGAACTCCAGGTCACTGGAAGTTAGTGGAATCAT
GTAGTTGAATTCTTTACTTCAAGACATTGTATTCTCTCCAGCTATCAAAACATTAATGATC
TTTTATGTCTTTTTTTTGTTATTGTTATACTTTAAGTTCTGGGGTACATGTGCGGAACATG
TAGGTTTGTTACATAGGTATACATGTGCCATGGTGGTTTGCTGCACTCATCAACCTGTCAT
CTACATTCTTTTATGTCTGTCTTTCAAAGCAACACTCTGTTCTTCTGAGTAGTGAAATCAG
GTCAACTTTACCACCAGCCTCCATTTTTAATATGCTTCACCATCATCCAGCACCTACTTAA
GATTTATCTAGGGCTCTGTGGTGATGTTAGGACCCATAAAAGAAATTTATGCCTTCCATAT
GTTTGGTTACAGATGGGAAATGGGAATGTTGAAGGACATGAAAGAAAGGATGTTTACACAT
TAAGCATCAGTTCTGAAGCTAGATTGTCTGAGTTTGAATCTTAGCTCTTCCCTTTATTAGC
TCTGTGACCTCGAGCTAGTTACTTAAATGCTCTGATCCTCTATTTCCTGATCAGTGAAACC
TCCCTATTCAAATGTGTGAGAGTTTAATAAATTAGGACACTTAAAAATGTTGGAGCAGTGC

ATAGCATGTAGTGTTCAGTACATGTTAAATGTTGTTTTTTATTATGTACAAACATGTGTGG
GCACAGAATTTTAAATCATCTCAACTTTTGAGAAATTTTGAGTTATCAACACCGTTCCCAC

FIGURE 61B

AAGACAGTGGCAAAATTATTGGTGAGAATTAAACAGCTGTTTCTCAGAGGAAGCAATGGAG
GCTTGCTGGGATAAAGGCATTTACTGAGAGGCTGTTACCTAGTGAGAGTGATGAATTAATT
AAAATAGTCGAATCCCTTTCTGACTGTCTCTGAAAGCTTCCGCTTTTATCTTTGAAGAGCA
GAATTGTCACCCCAAGGACATTTATTAATAAAAAGAACAACTGTCCAGTGCAATGAAGGCA
AAGTCATAGGTCTCCCAAGTCTTACCCCATTCCTGTGAAATATCAAGTTCTTGGCTTTTCT
CTGTCATGTAGCCTCAACTTTCTCCGACCGGGTGCATTTCTTTCTCTGGTTTCTAAATTGC
CAGTGGCAAATTTGGATCACTTACTTAATATCTGTTAAATTTTGTGACCCAACAAAGTCTT
TTAGCACTGTGGTGTCAAAAAGAAAAACACCTCCCAGGCATATACATTTTATAGATTCCTG
GAGAATGTTGCTCTCCAGCTCCATCCCCACCCAATGAAATATGATCCAGAGAGTCTTGCAA
AGAGACAAGCCTCATTTTCCACAATTAGCTCTAAAGTGCCTCCAGGAAATGATTTTCTCAG
CTCATCTCTCTGTATTCCCTGTTTTGGATCACAGGGCAATCTGTTTAAATGACTAATTACA
GAAATCATTAAAGGCACCAAGCAAATGTCATCTCTGAATACACACATCCCAAGCTTTACAA
ATCCTGCCTGGCTTGACAGTGATGAGGCCACTTAACAGTCCAGCGCAGGCGGATGTTAAAA
AAAATAAAAAGGTGACCATCTGCGGTTTAGTTTTTTAACTTTCTGATTTCACACTTAACGT
CTGTCATTCTGTTACTGGGCACCTGTTTAAATTCTATTTTAAAATGTTAATGAGTGTTGTT
TAAAATAAAATCAGGAAAGAGAGAAAAAAAAAAAAAAAAAAA

FIGURE 62A

GCGGCGGGAGAGCTGCTGGCTCGCCCGGATCCCGGGAGCTGCCTGGAGGCGGGCCCGGCCC
GGGGAAGGTGAGCGGCTGCGGGACCCAGCCCCTCGCCGGGAGCGGGCACC**ATG**GTGCTGTC
GGTGCCTGTGATCGCGCTGGGCGCCACGCTGGGCACAGCCACCAGCATCCTCGCGTTGTGC
GGGGTCACCTGCCTGTGTCGGCACATGCACCCCAAGAAGGGGCTGCTGCCGCGGGACCAGG
ACCCCGACCTGGAGAAGGCGAAGCCCAGCTTGCTCGGGTCTGCACAACAGTTCAATGTTAA
AAAGTCCACGGAACCTGTTCAGCCCCGTGCCCTCCTCAAGTTCCCAGACATCTATGGACCC
AGGCCAGCTGTGACGGCTCCAGAGGTCATCAACTATGCAGACTATTCACTGAGGTCTACGG
AGGAGCCCACTGCACCTGCCAGCCCCCAACCCCCGAATGACAGTCGCCTCAAGAGGCAGGT
CACAGAGGAGCTGTTCATCCTCCCTCAGAATGGTGTGGTGGAGGATGTCTGTGTCATGGAG
ACCTGGAACCCAGAGAAGGCTGCCAGTTGGAACCAGGCCCCCAAACTCCACTACTGCCTGG
ACTATGACTGTCAGAAGGCAGAATTGTTTGTGACTCGCCTGGAAGCTGTGACCAGCAACCA
CGACGGAGGCTGTGACTGCTACGTCCAAGGGAGTGTGGCCAATAGGACCGGCTCTGTGGAG
GCTCAGACAGCCCTAAAGAAGCGGCAGCTGCACACCACCTGGGAGGAGGGCCTGGTGCTCC
CCCTGGCGGAGGAGGAGCTCCCCACAGCCACCCTGACGCTGACCTTGAGGACCTGCGACCG
CTTCTCCCGTCACAGCGTGGCCGGGGAGCTCCGCCTGGGCCTGGACGGGACATCTGTGCCT
CTAGGGGCTGCCCAGTGGGGCGAGCTGAAGACTTCAGCGAAGGAGCCATCTGCAGGAGCTG
GAGAGGTCCTACTATCCATCAGCTACCTCCCGGCTGCCAACCGCCTCCTGGTGGTGCTGAT
TAAAGCCAAGAACCTCCACTCTAACCAGTCCAAGGAGCTCCTGGGGAAGGATGTCTCTGTC
AAGGTGACCTTGAAGCACCAGGCTCGGAAGCTGAAGAAGAAGCAGACTAAACGAGCTAAGC
ACAAGATCAACCCCGTGTGGAACGAGATGATCATGTTTGAGCTGCCTGACGACCTGCTGCA
GGCCTCCAGTGTGGAGCTGGAAGTGCTGGGCCAGGACGATTCAGGGCAGAGCTGTGCGCTT
GGCCACTGCAGCCTGGGCCTGCACACCTCGGGCTCTGAGCGCAGCCACTGGGAGGAGATGC
TCAAAAACCCTCGCCGGCAGATTGCCATGTGGCACCAGCTGCACCTG**TAA**CCAGCTGCCCA
GCTGCCTCCCTTCTTGGACAGCCCTGACCCGTCCTCTGCAACCTCCTTTCTGTGCCCCTTT
CCTCATTCTGACACCCAGAAGACAGTGACAGATGTGTTTGCAAGGCTGGGATGGCTCTCTC
ATCATACTCTTGTTTCTTAGAAATAAGCAAGACAGAGCAGGAAATGGAATATGCGGGTCAC
ACTGAGGAATGCATTTTGCTCATCTGTGTTATTGAAGGAGGTGCTTATTAAATACAGTTCC
TATGCCTGTTTTATAGGTGGGGTTAGGCCAGATGCAGAGAAAGCTAAATGTGGGAATCATG
GATGCAAAGAAGAATTTGGCTTTTTGAAAAACAAGCATTTCAAAAATGATGAAGGAAGTGA
AAGTATCCTGGATCAACTCCTAGAGTTAGAGATTGCCCAGGTGGAAAGAAACCTTAGCCAG
CGTTCAATCAAGCTCACCATGCAGGGCAGTCACCCGGCAGTTCTCAAACTTTAGCATGTGA
AGAGTCACCAGCAGATTCCTGGGCTCGCCTGGAGACATTCCTAGTCGGTATTCCTGGTCGA
AGCCCAGGAGCCTTCCTTTTTAACAAGCTGATGTAGAGGGTGGAGCACTGTATGTGGAGAA
ATTCCTTCTACAATATTCCACACAGGTTTTTGGCCACAGTCCTTGATGGAGTCCCAAAACC
ATGGTGCAGCCAGTTCCAATGCTGGACACCTCAACCATCAGGGTGAAATCTGGGGCCTCAG

CTTTTTAATTTAATTATTTTAATTCTTAATACTTTAATTTGTGCATTTCATAAGCCCCCTG
CTCTTGGACTGAATTTTGTGCTTTTTATTGAAGAATTTTATTGTTTTTATCTTAAAATCAG
TTTCTATTATCCTTGGGGAGACCATCCCTAACAAAGTACAGGTGGGATCTCCTGTGAGTCA
TTGGCTGGGTTCTGATTGCTAGATGTCACACCCACCAGCATCACCAAAGTGACTCTGAGAT
AGACCGGTCCCTTCTCAGCGTTCCAGTCACTTCAGGAGGAATTTAGTTATTGACTTAGTCT
ATGACATCTGGCTACATGTAGGTAGAGAAGAAAGACAATTTTAAAAAGGAAATCAGGTCTT
TTGCAACTGTGCCTCCCTCTGTCTGTTTTCACTTGAATGGGTAAATAACCAGCAGCTAGGT
TTTGAATTCCTACCTTGTTATTCTAAACAGATGTCCACATTGTTAATTAAATCTAAATTAT
GAGCCTTGCTGAGTGGATACGGTACTTACACCTGAACCAGGATTCCTGGGTTCTGTTGTTG
ACATTGCCCTTCAGCACCTGTTTGGCCAGCTGTATAAGATAGGACTAATGACTAGGAAGCC
TACCCCAATGAATGATATACTAGATGAAATAGTGTTCAAAACCTGTAGGCACTCTCTGGCT
AAAAACAAACTCTGAGGCCACCAGCAGATCATCTTTAAGCTAAGTTACTATTTTTCACCTT
TTTTTTTAGACGGAGTTTTGCTCTTTGTTGCCCAGGCTGGAGTGCAGTGGCACGATCTCGG
CTCACTGCAACCTCCGCCTCCCAAGTTCAAGCGATTCTCCTGTCTCAGCCTCCTGGGTAGC
TGGGATTACAGGTGCCCACCAACATGCCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGG
TTTCACCATGTTGGCCAGGCTGGTCTTCAACTCCAGATCTCAGGTGATCTACCCTCCTCGG
CCTCCCAAAGTACTGGGATTACAGGCCTGAGCCACCGCGCCCGGCCTATTTTTCACTTTAA
TTTGGCAGCTGAGAATGCCCAAAAAGTGCCAGAAGCATCGTGGCATTTCCAGAACCATGGA
TTCTGCCTTTGGACCCCTCTCTATTAATATTAAAACTCTGGGCCTTCAGATGTCACCCTAA
TCCACTGCCCTAAGACAGAATTTCTGGACAAGATGGGTAAGGGCTTCATTCCTTCAACAAG
TCAAGTCATACTTGGCCTCTCCCTGAGAATCTGAGCAGGAGCCTTATAACCTGTGGTCATT
ATTTTTTCTTTCTGTACAGAAATAGAAAAGCATTAGAAATAACTTCTAACCATCCTCTGAA
AAAACAGAAAAAATATCGAATCCCTCTTTCATGAGAAGTCTTTTGGATAATTGGAAACCTT
CATCACTGAGGTTGGCCAGCCCCTGCCAAGTGTTGTGTAGGCAAAGCACTTGTTAGTGGCT
TCCTATGAAATGTTTTAGAGATCTCTTCACCATACTGGTTTCTTCTCTTTGGTTGGTGTGG
GTAAAAGAAAACAAAACATTTCCTATAAGCTGAAAGCTGACCAGCATTCTCTTCTTGGTAA
CATCTACTACTCCAACCTAGAAAATTTGGATTCTAGACCAAAAATCAGGAAACATGGCTCC
TTATAAATCTGTGCAGCTGCCTTATAGTACCATCAAAGGAATTTCAGGTGGGCTGGGCGGG
GCCCCGATCCCAGAATTATCAACTCCACCCATCATCATTTGGTCATGAAGCATCCTTTCAT
TCTTCTTCTTCTTTTTTTTGGGGGGGGCGGGGCGGGGGAGGGATCTCAAAGTTTTAGTCTT
CCAGAATCCAAATTAAAGGTTGCCCCTGATGGGGGCCAGGTTCCGCCACAGAACATCTTAG
ATGTCAGCCTTGACCTCACTTAGCAGGGATTACAGAAATGAGATACATTTTGAAGGAGAGT
TGTCTGTTATGTTCACTGTATTCTAAGTGCCTGGGATAAAGCTGTCTCATGGGTGCTCCAT
ATATATTCATATATATTTGTTGAGTGAATTAATGAATTAAGAGTGGCTGGCAGAGTAGGCA
GAAAAAGACACTGCAAATGGCATAAAAATTAAAGTCCTAGCTGAGTTCTCAATGGTAAAGG
CATCAGATGTCTTAGCAGTCAAGCTAGAAATTCATGACAATGAGTATTACTATTTGCCTAA
TGACAACTCATTGCTCTCCATGTAAATGTAATCAACAGATGAAGAGAATATAATTGCTCTG

CTTTTCCACTAAAACTCCATCTTAGTGAATTTTAAATTATCCAGAGATGTCAAACTGCCAA
ATAAAAATATTTCAGTAGTCTTTGCATCAGCTTACCTTGTACCAGAAACATTTCCAATTTA

FIGURE 62B

CTATCAAATTATAGTAACTGAGCCTGTGTGAAGTATCTCATCATTTTCGAAAGGAACACCT
TGTGTGATGCCAGTGAGCATTTCTAAAAAGGGTGTGAGGTAGAGGTAAAAATAAGGTGAGA
GACCATTTCAGAATGCACTGTTGCTCAAAAAGGTGATCTGGTTCTTTCTTCAGAGATTTCT
ACGGGGATAGAAAATCGGGAGTCTGCCCTCATTAATCTGTGACTCCACCTCTTGCATCAAA
TCAATATCTATTTGTTGAGCACTTATTGATTAAGACCTTGCATATGTCTGTCCATTTTGAT
TTGAGATACAACTTTTTGTGTGGGTTGAATGACAAATCACTCCAAACAAAACTGGGCACAG
AGAATCAGCTAGGAGACCAGTTATTCAGGGTCCATTTCTCTTGGATGTAAAGGAGTCCTGG
GTAAAATGTGGCTGTAACCTAAACCAACTAGTCCTTGTGATTTGTTTCTGCCCTCTGTGTT
TCCTGTTGTCAAATGCTAAGTGTGTGTTTTGCAGTCATGAACTAAAGCACAAAAAGATGCA
TGAGACATTGTAGTCATATGTCTGGTGTGACACTTTGGAGCAAAAACCTTGCAGTGGTAAA
TAAAAAATTTCCAACAGGGAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63A

AGCGGCAGCTCCGCTGAAAACTGCATTCAGCCAGTCCTCCGGACTTCTGGAGCGGGGACAG
GGCGCAGGGCATCAGCAGCCACCAGCAGGACCTGGGAAATAGGGATTCTTCTGCCTCCACT
TCAGGTTTTAGCAGCTTGGTGCTAAATTGCTGTCTCAAAATGCAGAGGATCTAATTTGCAG
AGGAAAACAGCCAAAGAAGGAAGAGGAGGAAAAGGAAAAAAAAAGGGGTATATTGTGGATG
CTCTACTTTTCTTGGAA**ATG**CAAAAGATTATGCATATTTCTGTCCTCCTTTCTCCTGTTTT
ATGGGGACTGATTTTTGGTGTCTCTTCTAACAGCATACAGATAGGGGGGCTATTTCCTAGG
GGCGCCGATCAAGAATACAGTGCATTTCGAGTAGGGATGGTTCAGTTTTCCACTTCGGAGT
TCAGACTGACACCCCACATCGACAATTTGGAGGTGGCAAACAGCTTCGCAGTCACTAATGC
TTTCTGCTCCCAGTTTTCGAGAGGAGTCTATGCTATTTTTGGATTTTATGACAAGAAGTCT
GTAAATACCATCACATCATTTTGCGGAACACTCCACGTCTCCTTCATCACTCCCAGCTTCC
CAACAGATGGCACACATCCATTTGTCATTCAGATGAGACCCGACCTCAAAGGAGCTCTCCT
TAGCTTGATTGAATACTATCAATGGGACAAGTTTGCATACCTCTATGACAGTGACAGAGGC
TTATCAACACTGCAAGCTGTGCTGGATTCTGCTGCTGAAAAGAAATGGCAAGTGACTGCTA
TCAATGTGGGAAACATTAACAATGACAAGAAAGATGAGATGTACCGATCACTTTTTCAAGA
TCTGGAGTTAAAAAAGGAACGGCGTGTAATTCTGGACTGTGAAAGGGATAAAGTAAACGAC
ATTGTAGACCAGGTTATTACCATTGGAAAACACGTTAAAGGGTACCACTACATCATTGCAA
ATCTGGAATTTACTGATGGAGACCTATTAAAAATCCAGTTTGGAGGTGCAAATGTCTCTGG
ATTTCAGATAGTGGACTATGATGATTCGTTGGTATCTAAATTTATAGAAAGATGGTCAACA
CTGGAAGAAAAAGAATACCCTGGAGCTCACACAACAACAATTAAGTATACTTCTGCTCTGA
CCTATGATGCCGTTCAAGTGATGACTGAAGCCTTCCGCAACCTAAGGAAGCAAAGAATTGA
AATCTCCCGAAGGGGGAATGCAGGAGACTGTCTGGCAAACCCAGCAGTGCCCTGGGGACAA
GGTGTAGAAATAGAAAGGGCCCTCAAACAGGTTCAGGTTGAAGGTCTCTCAGGAAATATAA
AGTTTGACCAGAATGGAAAAAGAATAAACTATACAATTAACATCATGGAGCTCAAAACTAA
TGGGCCCCGGAAGATTGGCTACTGGAGTGAAGTGGACAAAATGGTTGTTACCCTTACTGAG
CTCCCTTCTGGAAATGACACCTCTGGGCTTGAGAATAAGACTGTTGTTGTCACCACAATTT
TGGAATCTCCGTATGTTATGATGAAGAAAAATCATGAAATGCTTGAAGGCAATGAGCGCTA
TGAGGGCTACTGTGTTGACCTGGCTGCAGAAATCGCCAAACATTGTGGGTTCAAGTACAAG
TTGACAATTGTTGGTGATGGCAAGTATGGGGCCAGGGATGCAGACACGAAAATTTGGAATG
GGATGGTTGGAGAACTTGTATATGGGAAAGCTGATATTGCAATTGCTCCATTAACTATTAC
CCTTGTGAGAGAAGAGGTGATTGACTTCTCAAAGCCCTTCATGAGCCTCGGGATATCTATC
ATGATCAAGAAGCCTCAGAAGTCCAAACCAGGAGTGTTTTCCTTTCTTGATCCTTTAGCCT
ATGAGATCTGGATGTGCATTGTTTTTGCCTACATTGGGGTCAGTGTAGTTTTATTCCTGGT
CAGCAGATTTAGCCCCTACGAGTGGCACACTGAGGAGTTTGAAGATGGAAGAGAAACACAA
AGTAGTGAATCAACTAATGAATTTGGGATTTTTAATAGTCTCTGGTTTTCCTTGGGTGCCT
TTATGCGGCAAGGATGCGATATTTCGCCAAGATCCCTCTCTGGGCGCATTGTTGGAGGTGT

GTGGTGGTTCTTTACCCTGATCATAATCTCCTCCTACACGGCTAACTTAGCTGCCTTCCTG
ACTGTAGAGAGGATGGTGTCTCCCATCGAAAGTGCTGAGGATCTTTCTAAGCAAACAGAAA
TTGCTTATGGAACATTAGACTCTGGCTCCACTAAAGAGTTTTTCAGGAGATCTAAAATTGC
AGTGTTTGATAAAATGTGGACCTACATGCGGAGTGCGGAGCCCTCTGTGTTTGTGAGGACT
ACGGCCGAAGGGGTGGCTAGAGTGCGGAAGTCCAAAGGGAAATATGCCTACTTGTTGGAGT
CCACGATGAACGAGTACATTGAGCAAAGGAAGCCTTGCGACACCATGAAAGTTGGTGGAAA
CCTGGATTCCAAAGGCTATGGCATCGCAACACCTAAAGGATCCTCATTAGGAACCCCAGTA
AATCTTGCAGTATTGAAACTCAGTGAGCAAGGCGTCTTAGACAAGCTGAAAAACAAATGGT
GGTACGATAAAGGTGAATGTGGAGCCAAGGACTCTGGAAGTAAGGAAAAGACCAGTGCCCT
CAGTCTGAGCAACGTTGCTGGAGTATTCTACATCCTTGTCGGGGGCCTTGGTTTGGCAATG
CTGGTGGCTTTGATTGAGTTCTGTTACAAGTCAAGGGCCGAGGCGAAACGAATGAAGGTGG
CAAAGAATGCACAGAATATTAACCCATCTTCCTCGCAGAATTCACAGAATTTTGCAACTTA
TAAGGAAGGTTACAACGTATATGGCATCGAAAGTGTTAAAATT**TAG**GGGATGACCTTGAAT
GATGCCATGAGGAACAAGGCAAGGCTGTCAATTACAGGAAGTACTGGAGAAAATGGACGTG
TTATGACTCCAGAATTTCCCAAAGCAGTGCATGCTGTCCCTTACGTGAGTCCTGGCATGGG
AATGAATGTCAGTGTGACTGATCTCTCGTGATTGATAAGAACCTTTTGAGTGCCTTACACA
ATGGTTTTCTTGTGTGTTTATTGTCAAAGTGGTGAGAGGCATCCAGTATCTTGAAGACTTT
TCTTTCAGCCAAGAATTCTTAAATATGTGGAGTTCATCTTGAATTGTAAGGAATGATTAAT
TAAAACACAACATCTTTTTCTACTCGAGTTACAGACAAAGCGTGGTGGACATGCACAGCTA
ACATGGAAGTACTATAATTTACCTGAAGTCTTTGTACAGACAACAAACCTGTTTCTGCAGC
CACTATTGTTAGTCTCTTGATTCATAATGACTTAAGCACACTTGACATCAACTGCATCAAG
ATGTGACATGTTTTATAAAAAAAGGAAAAAAAACATTTAAAACTAAAAAATATTTTTAGGT
ATTTTCACAAACAAACTGGCTTTTAAATAAATTTGCTTCCATATTGGTTGAATAAGACAAA
AACAATTAAACTGAGTGGGAAGTGAATAAAAAAAGGCTTTAGGTATCGATTCCATATTTTT
CAAAGCCAAATATGTAAATGCTAAGGAAAGTAAACAAAGAGGAGATTCCAATCTTGTAATT
TAATATTGTTATTAAAACTTTAATGTATCCTATTCTTTAACATTTGGTGTTAATATAAAAT
TACTTGGCAATGCTTGACATTTGAAATAAACATTTTTCTATTGTTTTATTGCAAGTGGTCC
AATTAATTTTGCTTAGCTACAGTTTGGTCATAAATCAAGTGAGTTTAAAGACACTACCAAG
TTGTTAGGTGCCCAGAGAAAATTTCTCCCTTTTAAAAAGGCCAGGTGATTTTTCAAATGTA
ATCTTGCCCCCAAAGTAATATCCGAATATCTTTTTGACATGTCTAAATATATATATATATA
AAGAAATATTTGTTAACACAAAAGCATTTGATCTATGTAGATAAATGCTAATAGATTTAAA
AAGCTAATATTAACAAATACCAGAATACGTGAAGTTCCATTTTTAAAGTGTTTGAGCTTAC
AGAAGAGAAACATTCATTTTAAATGAAGTAAAAAATGCCTTGAAAGTAATTCTTTAGATAG
TTGCCCATTGATTAAATTCCAAAAACTAAATATGTTTTTAGCTTTAAAATTATAAAAGCTG
TCATAAACTTTATATATTATGAATTTTAAAATATGTTTGAGTCTCCTGCAATATAGTTTCA
TCCCATTGACATCAATTAAAAATAACCCTAATATATTATTTTTATATTTATTCCTCAGGTG
GAATGGCTATTTTAATATGCCCAGTGTGGATAAAATGTCACATTTCTGTAACTTTTGACTA

AAGAGCCTATATTTATCTAGTTAATGAATTTAAAGGATCTATCTTTCCCTTCATAAAATAC
CTCTTATTTCCATTAAAGCCCCCCAAGTTTAATTAATTTAGGATTTTGAATGATTATTGAC

FIGURE 63B

ATCCAATAGTTATTTTTAATATTTGTATTCTTGTTATTTCTGGAAGAAAGCCTTTGTGTAG
CACTTGGTATTTTGCAAAGTGCTTTTAAAACATTCTTACTTACCGTATTTCATAGAAGGGA
AGGAAAAATGTAAGGTTTAACAGTAAGCACTTGCATTGAACATGGAGGCATGTGGTATCAT
GATATTCTTCACTAAATTTAGCTGTCCCTAATCACAGATCCTAAGGTAATATAATATAATT
TTAGTGCATTTCTCCTCATCAGGAATGCTGGAGGTGCATTTTAAGTTTTAATAATAAGTGC
TAGAATGACCAAATTGCAGACTAATTGTTTCCATATTGTACTTAAAATGAGTTTTTAAAAG
TGAAAAAGAAATGACTATATACAATCAATGCTATTTATTGTACCTCTGGGCCTACTCTTCT
AAAAATTGTAGCTTATCGATTTTTCTCTGTCAAGCTTGAACTAATGTAAATAATTGAAATA
ATGTAAAGTTATATTTTCATGTTTTTATAGATACAACATGACAAGAATACATAATGTAAGA
GTATTTCAACTATGGATAATGTTGATTGGATAATGCACATCTCAGTTACAAGCAGTACTCA
TAGTTTAATATCCATGTAACGGTGCATCAATATATTGCTATATAAATATGTCTGTGTGCAT
ATAAGTGAAAAGTGGTCAAACAAGAGTGATGACAGCTGTCTAAAGGTTTTTTTATTCATTT
TATATAAAAACTGTTATGGAAAGACCAAAATGTTTATGAACTATTCTTATGTAAATTTACA
ATTGTCCTTTACTGTACTTTTTTGTTTACAGTATAGTACCTTATTTTCTGCTGTGTTAAGT
GGGTGTCAAACTCCAAGAAGACATACACTTTCTATAACTTCTATTGAAGATATTGGAATTT
CCAATTTTTCATGTGTACTATGTCAGAAAATGCTTTCGATTTTATTTTTAAATCTAACATC
GGATGGCTTTTCCGGAGTGTTGTAAAAACTTCAATCATACATAAAACATGTTCTTACAAAA
GGCAAAGATCTT

FIGURE 64

GCCAACACTGGCCAAAAGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTCGCCT
GCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTAGGGT
CCAGA**ATG**GCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACTACAG
ACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGGGGTT
GTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGACTCCA
ACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCTTTGG
CCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTTCCTC
TTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTGACCA
AGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGGGCTT
CAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGACCAAC
GTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTGCTCA
CCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTGGTTC
CTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCATTGCC
ATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGATGACA
CCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTAGTCC
CGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGCTTTC
TGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAAGAGG
AAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACACATTT
TCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTGGCCG
AGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGC**TAA**CTCTGTCCTGAAGAGTGGG
ACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCTTGAG
TCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTCAACC
ACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCGCTGT
AGTATTTTTTTTTTTTTGTCTCATCCTTTAGATACTTCTTTTAAGTGGGAGTCTCAG

FIGURE 65

TCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAATAATAATT
TAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCGGTAGGAAT
AAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTTAGTTGGCT
GACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTCCCAGGCTT
CCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAATATTTAATA
TTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGTACTTAATG
GAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCTTTTATTCA
TGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGGATAACACA
AGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCACATTCCTA
TTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTATGTTTCAAA
TGGCTAAAAGCTTATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGAGATGTGGC
TTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTAGTCATTCT
GAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCTACTTACTA
AATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTACCTGTCTA
CTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACTACAAATTC
CATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACTTCCTATCC
ACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGCATACCACA
GTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACACATGTAAAT
CACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTANNAAAAAAACAAAAAA
GGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGTAATTTTTT
TAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAAAAATAAAA
ATAAAAATAAGATTTTGTTACTT

FIGURE 66

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCG**ATG**CTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAA**TGA**GCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
GTAGGAATAAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 67

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGGAGA**ATG**CCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAA**TGA**GCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
GTAGGAATAAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 68

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGA**ATG**CCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAA**TGA**GCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
GTAGGAATAAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 69

CAGGCCTCTGAGGCTCCCTTGCCGAGGGCCCCGAGCTGCAGGGACAGTGAGCAGTGAGTCC
CTTGGGCATCCCGCTCCTGGGCAGGTCACCAATAGGTCCCCGCAGTTCCCAATGGAACTGT
TCCAGTCCTCCCCGAGGCCTCCACTTCAACCTGTCTGTGTCTGCCCAGGCCTGGAGTTGTG
TGACCCTCCCCACCGCCTGGCCTTCTCCATGGGGGCTGGCCTTTTCTCGGTGGTGGGCACC
CTGCTGCTGCCCGGCCTGGCTGCGCTTGTGCAGGACTGGCGTCTTCTGCAGGGGCTGGGTG
CCCTGATGAGTGGACTCTTGCTGCTCTTTTGGGGGTTCCCGGCCCTGTTCCCCGAGTCTCC
CTGCTGGCTGCTGGCCACAGGTCAGGTAGCTCGAGCCAGGAAGATCCTGTGGCGCTTTGCA
GAAGCCAGTGGCGTGGGCCCCGGGGACAGTTCCTTGGAGGAGAACTCCCTGGCTACAGAGC
TGACCATGCTGTCTGCACGGAGCCCCCAGCCCCGGTACCACTCCCCACTGGGGCTTCTGCG
TACCCGAGTCACCTGGAGAAACGGGCTTATCTTGGGCTTCAGCTCGCTGGTTGGTGGAGGC
ATCAGAGCTAGCTTCCGCCGCAGCCTGGCACCTCAGGTGCCGACCTTCTACCTGCCCTACT
TCCTGGAGGCCGGCCTGGAGGCGGCAGCCTTGGTCTTCCTGCTCCTGACGGCAGATTGCTG
TGGACGCCGCCCCGTGCTGCTGCTGGGCACCATGGTCACAGGCCTGGCATCCCTGCTGCTC
CTCGCTGGGGCCCAGTATCTGCCAGGCTGGACTGTGCTGTTCCTCTCTGTCCTGGGGCTCC
TGGCCTCCCGGGCTGTGTCCGCACTCAGCAGCCTCTTCGCGGCCGAGGTCTTCCCCACGGT
GATCAGGGGGGCCGGGCTGGGCCTGGTGCTGGGGGCCGGGTTCCTGGGCCAGGCAGCCGGC
CCCCTGGACACCCTGCACGGCCGGCAGGGCTTCTTCCTGCAACAAGTCGTCTTCGCCTCCC
TTGCTGTCCTTGCCCTGCTGTGTGTCCTGCTGCTGCCTGAGAGCCGAAGCCGGGGGCTGCC
CCAGTCACTGCAGGACGCCGACCGCCTGCGCCGCTCCCCACTCCTGCGGGGCCGCCCCCGC
CAGGACCACCTGCCTCTGCTGCCGCCCTCCAACTCCTACTGGGCCGGCCACACCCCCGAGC
AGCAC**TAG**TCCTGCCTGGTGGCCCTGGGAGCCAGGATGGGACCAAAGTCAAGGCCTGGGGC
ATGGCTGAGTACCCCAGACGTCTGGTCCAGGGCAGACACATTCCTCTCAGAAGCCCGTGTC
TCAGTGCAGGTGGAGCCGTGGGGACAGCGTGAAGGTGTCTCCAGCCAGGCCCCAGGCACTG
GGAGGCCCTGGGTCTCCCCCCAGCCACACCCAGTAGGTGTGGAGGATAAAGGCTTCTGTGG
AACTTGGCGTCCTCCCTGGTTCTTCCCTCAGGCCACACCTGGGGCCTGGAAGGGTCTCACC
TGTGCCTGCTCCCCACCTCTGGGCTCACCAGCAAGGAGCCAGGGTGTCAGCGGCTGCTCCA
TAGCCCCTGTCCAGGTAGTGGCAGCCCCCGGCCGCCTCCTGCCCCAGCCAAGCAGGGTGGT
ATGGGGGCCTGGCCCTCAGGTCCCTGCCAGGCCAGGCCTCACTGGCCCCATCCCCCTCATG
GTCAGTGTGGCAGTGGGCGCCCGTCCTGCATGAAAGAGACTAGCACG

FIGURE 70

AGCGGGGCGGGTCCCCGAAGCCCAGTACCAGACCTCCTGTGCGAAGAGTAACCTGCTCTTC
GGACTGATAACTGGATGGAATGGCCTTTAGAAAGGGAGCATCTTCAGAATCTGCAGCTGGG
ATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCCGAGCCAGCATCTTGCTAAGGGTT
ACAAAATAA**ATG**CCATTTGAACAGTGCCATCTGTCATGGAAAAACGTGAGACGTTTGTACA
ATGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAATTTGTTCAACTTGATAAAGAGGC
CTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATTATCAAGGAAACAGAAAGAAGAA
TTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTGTTAGAAAGCCCAGTGGATGGGT
GGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAACCGAACATGGCTCAAAAATGAG
AAAAAGCATAGAAATAATTTATGCAATTACATCTGTGATTCTTGCTTCTGTGTCTGTAATA
AATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTTATTATATTAAATGGTATTTTAT
ATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTATTCAGGATTTGTGTGTTACCTG
GTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAAGACAGCCTTTGTCATGTTACTA
AGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGTCGGCTTTGGCGTATCCATCAAG
CTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAGAAATTAAAGATATGTTACTTGA
GTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGGAAGAAGATTTCTTAGTTGTCTC
TTCAACTGGAATATCAACTTCATCAAAATGATCCACGGGACCATTAAAAACCAGTTACAGG
GATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTTATTTCAGAGCTTGGAAAAAGGC
TTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCATCCAGGACTTCATGTTCCACGGG
ATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTGCGGGAGGTTCTGAGTTACTTTC
ACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGCTTTATAGATTATATAAGCCCAT
CCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCGATCAAATGCTGCATTGTTGTTT
GTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCTATTGAAATGGATAGTRAAATCC
AGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATCCTTACCCGATGGTCCGTTCCAC
AGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTGGGAAATGATGCCCCCGACCATT
CTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCATTTGACACGAGCTCAGCTGATG
TTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGGACAACAAACTGAGCCACCCATT
GTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCACGACAATTCGGAGAAAGTGAGG
GTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTGAGGGCTGCTAAGTTTTGGAAAA
TATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTGATTCTCGACCTGTGTCTCGGCG
CCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAATCAGCCGGAGGAGGTCTGGTGC
GAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCTGCCAGGAGGTTCTATCAGTACG
CCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGATTCACGTTATTCGTCATTGCTT
AAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGAGGACGAGGAGGAAGAGGACGGA
AGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTGTCAGTAAACGATGTTGCATGCA
TGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAAGTATTGACAGAAGTATGGAAAA

TAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGCCTCTGTGCTTCCAGAGTATCTG
AAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTCATGCTAATGTCCTTTATGCCGG
CCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCACGCTGAGAAGCCGGGAGGAGGG
CGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCTCTGCTCCTGGGGGCAGGTGGGG
CACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAGCATGCCCAGGCCAAGAGCAACA
CAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCCCAGTCAAACCTGAATTGGCATT
GGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCGCGAGTGCTTGCTCTCTGCTCCT
CGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACGTCAAAGGCAGATCTGGAGTCAC
TTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTGAAGCAGCTGCCCCGCGAGCCTT
TGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAAGTTCTGCTCAGAAGGAAAGGTG
TATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAAAGCAAAATTTTATCTTTTATTC
AAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCATTTATCAGCAAATTATCCAGAC
CTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCTTGGTGACCATCAGTTTCAGATG
CAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTGAAGGGATTTTTTTATGTTTCAT
TACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCTTGATTCAGAAAACAGATTCAGA
TGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGTATTTCAGAAAATGTTGGAATGT
ATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTGCGGCTGCTTTATTCTGTTCAGA
GGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGCACACAGACACCCCTGTGCACCG
GGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGAGATAAGTCACCAGCTACGGAAG
GTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTTTCTGATCTTCCACCATTTTCAA
GGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGGTCAGGTCATTTTTGGATGAATT
AAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGTGTGCCTCACGGCTGCTGTGCAT
ATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCAAAAGTGAGGGAGGTTGCAGCCA
CTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTTTGGAAGAGGATAGCATTGAAAG
ATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACTTTTGAATTCA**TAA**CCAAGCCAA
CATCTCCAGACATGTAAAAATAGGGAAAAGTGATTCAAATTGAAATGCCTGTGTATTTTCC
TATTGTTTTTAATGTTAATAACCCATATAATAGGGAAAGGGTGGGATTTTTTTGTGGGAAT
GTGGGAAGGTGGGGGTTATGGAGGAGATAACTCAAAACTTCTTCAATTTTGCCTAGTGCCT
GCGTAAATAATATATTTAATATAAAGGACTCCAGGTATGAATGGTGTAGAAATCCATGATT
CCAAGAAAAAACACTTTTCTAGCAAACCTGGTTGTTTTTAAAATGACTTTTATATATGTAA
TATTGCTTGGAAACTATGAGTAATAAAGCAATGACAACAAMAAAAAAAAAAAAAAA

FIGURE 71

GCCCTTCAGAATCTGCAGCTGGGATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCC
GAGCCAGCATCTTGCTAAGGGTTACAAAATAAATGCCATTTGAACAGTGCCATCTGTC**ATG**
GAAAAACGTGAGACGTTTGTACAAGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAAT
TTGTTCAACTTGATAAAGAGGCCTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATT
ATCAAGGAAACAGAAAGAAGAATTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTG
TTAGAAAGCCCAGTGGATGGGTGGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAA
CCGAACATGGCTCAAAAATGAGAAAAAGCATAGAAATAATTTATGCAATTACATCTGTGAT
TCTTGCTTCTGTGTCTGTAATAAATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTT
ATTATATTAAATGGTATTTTATATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTA
TTCAGGATTTGTGTGTTACCTGGTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAA
GACAGCCTTTGTCATGTTACTAAGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGT
CGGCTTTGGCGTATCCATCAAGCTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAG
AAATTAAAGATATGTTACTTGAGTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGG
AAGAAGATTTCTTAGTTGTCTCTTCAACTGGAATATCAACTTCATCAAAATGATCCACGGG
ACCATTAAAAACCAGTTACAGGGATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTT
ATTTCAGAGCTTGGAAAAAGGCTTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCAT
CCAGGACTTCATGTTCCACGGGATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTG
CGGGAGGTTCTGAGTTACTTTCACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGC
TTTATAGATTATATAAGCCCATCCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCG
ATCAAATGCTGCATTGTTGTTTGTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCT
ATTGAAATGGATAGTGAAATCCAGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATC
CTTACCCGATGGTCCGTTCCACAGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTG
GGAAATGATGCCCCCGACCATTCTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCA
TTTGACACGAGCTCAGCTGATGTTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGG
ACAACAAACTGAGCCACCCATTGTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCA
CGACAATTCGGAGAAAGTGAGGGTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTG
AGGGCTGCTAAGTTTTGGAAAATATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTG
ATTCTCGACCTGTGTCTCGGCGCCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAA
TCAGCCGGAGGAGGTCTGGTGCGAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCT
GCCAGGAGGTTCTATCAGTACGCCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGA
TTCACGTTATTCGTCATTGCTTAAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGA
GGACGAGGAGGAAGAGGACGGAAGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTG
TCAGTAAACGATGTTGCATGCATGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAA
GTATTGACAGAAGTATGGAAAATAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGC
CTCTGTGCTTCCAGAGTATCTGAAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTC

ATGCTAATGTCCTTTATGCCGGCCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCA
CGCTGAGAAGCCGGGAGGAGGGCGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCT
CTGCTCCTGGGGGCAGGTGGGGCACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAG
CATGCCCAGGCCAAGAGCAACACAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCC
CAGTCAAACCTGAATTGGCATTGGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCG
CGAGTGCTTGCTCTCTGCTCCTCGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACG
TCAAAGGCAGATCTGGAGTCACTTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTG
AAGCAGCTGCCCCGCGAGCCTTTGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAA
GTTCTGCTCAGAAGGAAAGGTGTATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAA
AGCAAAATTTTATCTTTTATTCAAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCA
TTTATCAGCAAATTATCCAGACCTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCT
TGGTGACCATCAGTTTCAGATGCAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTG
AAGGGATTTTTTTATGTTTCATTACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCT
TGATTCAGAAAACAGATTCAGATGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGT
ATTTCAGAAAATGTTGGAATGTATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTG
CGGCTGCTTTATTCTGTTCAGAGGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGC
ACACAGACACCCCTGTGCACCGGGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGA
GATAAGTCACCAGCTACGGAAGGTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTT
TCTGATCTTCCACCATTTTCAAGGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGG
TCAGGTCATTTTTGGATGAATTAAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGT
GTGCCTCACGGCTGCTGTGCATATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCA
AAAGTGAGGGAGGTTGCAGCCACTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTT
TGGAAGAGGATAGCATTGAAAGATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACT
TTTGAATTCA**TAA**CCAAGCCAACATCTCCAGACATGAAGGGC

FIGURE 72

TCCAACCAGATTTTCACCGCTATGCCTCCTTTCACTCTGGGAGTCTTCCAGAGGTCTTGCA
CTCGGGAGAGCATGCTCAGGTTTCCCCAGCTCTACAAAATCACCCAGAATGCCAAAGACTT
CAACGTGACCAGTGGTCATGCCACTGCCTGTTGATTTGTTGAAAATATTGTTTACACGTAT
GTTCTTGTTACTGATTGTCAGAAAGCTGGTTTTGAGACTGCAGCTTGGACTAAATTCAGTC
ATCTGGCTGTCTGGGGAAGCATGCTGACCAGTCTGGTGTTCTTTGGCATCTACTCAGCCAT
CTGGTCCACCATTCTCATTGCCCCAAATATGAGAGGACAGAAGAATGGTACCGGTACTGCC
AATGGAGATGGAGGAAGGAGACAGAAAGAAACAGAGCCCAGACCCTAGGGACCACCAGCAT
TTGCAGAATGGATAAACAGCCTTCTTCCTAACAAAGGAAGCACAGCAACTGTGATCCTGAG
CTGTGCACACTTCTGGTTGGGATTATTTCTGGTTTCTACTTCCTGTTTGAAGATGTGGCAT
GGAGAGTGAACAAGCTGCTGCCCACCACCTGGCATCACAGCCCCAGAACTCAGCTATTTCC
ATGGGACCACAGCATCTCATCTCTGGGCTGAGCCAGAAAGACCCCTACTGAAGTCCAGAGG
CACTTTTCTGAAAGGCTCTGCTTTGACCTGAAGTATTTTATCTATCCTCAGTCTCAGGACA
CTGTTGATGGAATTAAGGCCAAGCACATCTGCAAAAAAGACATTGCTGGAGGAGGTGCAAA
GAGCTGGAAACCAAGTCTCCAGTCCTGGGAAAAGCAGTGGTATGGAAAAGCAATGGAAAGA
GCATTTTGAAAATGCCATTCCACTGTTTTCTGGCCTTTATGATTTCTGCTGAGAAATCCAC
TGTTAGTCTGATGGGGTCTCCTTCATAGCACCAATGACCTGAAGAGCCTTGTTGAAGGAAG
ACTCCATCTGATGACTCAGAGCAAGTATTTTTTAGTGTGTTATTGTTATTAGCAGAAAGAG
GGCCATAAAATACATGGGGCAAGCTGAATATATCTTAGGCAAAAGAAGAAAATATTCAAAT
TCTTATGTTATTTTATCTAATTATTTTATCTCTTTTTGTGTGTGACTTATAATGTGTGTAT
TGTATTAATAAAAGTATATAAACATGTAGTTT

FIGURE 73

CATTCCCACCCTTCCTTCTTTAATAAGCAGGAGCGAAAAAGACAAATTCCAAAGAGGATTG
TTCAGTTCAAGGGAATGAAGAATTCAGAATAATTTTGGTAAATGGATTCCAATATCGGGAA
TAAGAATAAGCTGAACAGTTGACCTGCTTTGAAGAAACATACTGTCCATTTGTCTAAAATA
ATCTATAACAACCAAACCAATCAAA**ATG**AATTCAACATTATTTTCCCAGGTTGAAAATCAT
TCAGTCCACTCTAATTTCTCAGAGAAGAATGCCCAGCTTCTGGCTTTTGAAAATGATGATT
GTCATCTGCCCTTGGCCATGATATTTACCTTAGCTCTTGCTTATGGAGCTGTGATCATTCT
TGGTGTCTCTGGAAACCTGGCCTTGATCATAATCATCTTGAAACAAAAGGAGATGAGAAAT
GTTACCAACATCCTGATTGTGAACCTTTCCTTCTCAGACTTGCTTGTTGCCATCATGTGTC
TCCCCTTTACATTTGTCTACACATTAATGGACCACTGGGTCTTTGGTGAGGCGATGTGTAA
GTTGAATCCTTTTGTGCAATGTGTTTCAATCACTGTGTCCATTTTCTCTCTGGTTCTCATT
GCTGTGGAACGACATCAGCTGATAATCAACCCTCGAGGGTGGAGACCAAATAATAGACATG
CTTATGTAGGTATTGCTGTGATTTGGGTCCTTGCTGTGGCTTCTTCTTTGCCTTTCCTGAT
CTACCAAGTAATGACTGATGAGCCGTTCCAAAATGTAACACTTGATGCGTACAAAGACAAA
TACGTGTGCTTTGATCAATTTCCATCGGACTCTCATAGGTTGTCTTATACCACTCTCCTCT
TGGTGCTGCAGTATTTTGGTCCACTTTGTTTTATATTTATTTGCTACTTCAAGATATATAT
ACGCCTAAAAAGGAGAAACAACATGATGGACAAGATGAGAGACAATAAGTACAGGTCCAGT
GAAACCAAAAGAATCAATATCATGCTGCTCTCCATTGTGGTAGCATTTGCAGTCTGCTGGC
TCCCTCTTACCATCTTTAACACTGTGTTTGATTGGAATCATCAGATCATTGCTACCTGCAA
CCACAATCTGTTATTCCTGCTCTGCCACCTCACAGCAATGATATCCACTTGTGTCAACCCC
ATATTTTATGGGTTCCTGAACAAAAACTTCCAGAGAGACTTGCAGTTCTTCTTCAACTTTT
GTGATTTCCGGTCTCGGGATGATGATTATGAAACAATAGCCATGTCCACGATGCACACAGA
TGTTTCCAAAACTTCTTTGAAGCAAGCAAGCCCAGTCGCATTTAAAAAAATCAACAACAAT
GATGATAATGAAAAAATC**TGA**AACTACTTATAGCCTATGGTCCCGGATGACATCTGTTTAA
AAACAAGCACAACCTGCAACATACTTTGATTACCTGTTCTCCCAAGGAATGGGGTTGAAAT
CATTTGAAAATGACTAAGATTTTCTTGTCTTGCTTTTTACTGCTTTTGTTGTAGTTGTCAT
AATTACATTTGGAACAAAAGGTGTGGGCTTTGGGGTCTTCTGGAAATAGTTTTGACCAGAC
ATCTTTGAAGTGCTTTTTGTGAATTTATGCATATAATATAAAGACTTTTATACTGTACTTA
TTGGAATGAAATTTCTTTAAAGTATTACGATGCGCTGACTTCAGAAGTACCTGCCATCCAA
TACGGTCATTAGATTGGGTCATCTTGATTAGATTAGATTAGATTAGATTGTCAACAGATTG
GGCCATCCTTACTTTATGATAGGCATCATTTTAGTGTGTTACAATAGTAACAGTATGCAAA
AGCAGCATTCAGGAGCCGAAAGATAGTCTTGAAGTCATTCAGAAGTGGTTTGAGGTTTCTG
TTTTTTGGTGGTTTTTGTTTGTTTTTTTTTTTTTTCACCTTAAGGGAGGCTTTCATTTCCT
CCCGACTGATTGTCACTTAAATCAAAATTTAAAAATGAATAAAAAGACATACTTCTCAGCT
GCAAATATTATGGAGAATTGGGCACCCACAGGAATGAAGAGAGAAAGCAGCTCCCCAACTT
CAAAACCATTTTGGTACCTGACAACAAGAGCATTTTAGAGTAATTAATTTAATAAAGTAAA

TTAGTATTGCTGCAAATAGCTAAATTATATTTATTTGAATTGATGGTCAAGAGATTTTCCA
TTTTTTTTACAGACTGTTCAGTGTTTGTCAAGCTTCTGGTCTAATATGTACTCGAAAGACT
TTCCGCTTACAATTTGTAGAAACACAAATATCGTTTTCCATACAGCAGTGCCTATATAGTG
ACTGATTTTAACTTTCAATGTCCATCTTTCAAAGGAAGTAACACCAAGGTACAATGTTAAA
GGAATATTCACTTTACCTAGCAGGGAAAAATACACAAAAACTGCAGATACTTCATATAGCC
CATTTTAACTTGTATAAACTGTGTGACTTGTGGCGTCTTATAAATAATGCACTGTAAAGAT
TACTGAATAGTTGTGTCATGTTAATGTGCCTAATTTCATGTATCTTGTAATCATGATTGAG
CCTCAGAATCATTTGGAGAAACTATATTTTAAAGAACAAGACATACTTCAATGTATTATAC
AGATAAAGTATTACATGTGTTTGATTTTAAAAGGGCGGACATTTTATTAAAATCAATATTG
TTTTTGCTTTTTTCTGAGGAGTCTCTTTCAGTTTCATTTTTTCTCATCCCATGACTTCCCTC
CGATGGT

FIGURE 74

AGTATGAGGCGCCCACC**ATG**GTGCGATGCGACCGCGGGCTGCAGATGCTGCTGACCACGGC
CGGAGCCTTCGCCGCCTTCTCGCTCATGGCCATCGCCATCGGCACCGACTACTGGCTGTAC
TCCAGCGCGCACATCTGCAACGGCACCAACCTGACCATGGACGACGGGCCCCCGCCCCGCC
GCGCCCGCGGCGACCTCACCCACTCTGGTCTGTGGCGGGTGTGCTGCATCGAAGGGATCTA
TAAAGGGCACTGCTTCCGGATCAATCACTTCCCAGAGGACAATGACTACGACCACGACAGC
TCGGAGTACCTCCTCCGCATCGTGCGAGCCTCCAGCGTCTTCCCCATCCTCAGCACCATCC
TGCTCCTGCTGGGTGGCCTGTGCATCGGTGCTGGCAGGATCTACAGCCGCAAGAACAACAT
CGTCCTCAGTGCCGGCATCCTCTTCGTGGCTGCAGGCCTCAGTAACATCATCGGTATCATC
GTCTACATTTCCAGCAACACAGGTGACCCGAGTGACAAGCGGGACGAAGACAAAAAGAACC
ATTACAACTACGGCTGGTCTTTTTACTTTGGAGCTCTGTCTTTCATTGTGGCTGAGACCGT
GGGCGTCCTGGCTGTAAACATTTACATTGAGAAAAATAAAGAGTTGAGGTTTAAGACCAAA
CGGGAATTCCTTAAGGCGTCTTCCTCTTCTCCTTATGCCAGGATGCCGAGCTACAGGTACC
GGCGACGGCGCTCGAGGTCCAGCTCAAGGTCCACCGAGGCCTCGCCCTCCAGGGACGTGTC
GCCCATGGGCCTGAAGATCACAGGGGCCATCCCCATGGGGGGAGCTGTCCATGTACACGCTG
TCCAGGGAGCCCCTCAAGGTGACCACCGCAGCCAGCTACAGCCCCGACCAGGAGGCCAGCT
TCCTGCAGGTGCATGACTTTTTCCAGCAGGACCTGAAGGAAGGTTTCCACGTCAGCATGCT
GAACCGACGGACGACCCCTGTG**TGA**

FIGURE 75

CGGGTCTGATAGTCCCTACCTGTCAGGACTGGTGTTAGGATGAGATAATGTTTGTGAACTG
TAAACATATATAAACGTGTGCTACTGTGAGAACTGGAACAAAGAAGAGAGGGAGTGAGAGA
AATCAAGGGAGGGCTGGGGCTGGGAAAGAACGAAAAGGGAGTCGCGTATAGAGGAGAGGCG
ACAGTCGCGAGCCACACTTTGCAATGAAACTCTTTAGACTTTCTGCCGGGAGAGCGGCCCA
GACGCGCCAGGTCTGTAGCAGGAGGCCGCGAGGGCGGGTCCCCAGAAGCCTACAGGTGAGT
ATCGGTTCTCCCCTTCCCGGCTTTCGGTCCGGAGGAGGCGGGAGCAGCTTCCCTGTTCTGA
TCCTATCGCGGGCGGCGCAGGGCCGGCTTGGCCTTCCGTGGGACGGGGAGGGGGGCGGGAT
GTGTCACCCAAATACCAGTGGGGACGGTCGGTGGTGGAACCAGCCGGGCAGGTCGGGTAGA
GTATAAGAGCCGGAGGGAGCGGCCGGGGCGCAGACGCCTGCAGACCATCCCAGACGCCGGA
GCCCGAGCCCCGACGAGTCCCCGCGCCTCATCCGCCCGCGTCCGGTCCGCGTTCCTCCGCC
CCACC**ATG**GCTCGGGGCCCCGGCCTCGCGCCGCCACCGCTGCGGCTGCCGCTGCTGCTGCT
GGTGCTGGCGGCGGTGACCGGCCACACGGCCGCGCAGGACAACTGCACGTGTCCCACCAAC
AAGATGACCGTGTGCAGCCCCGACGGCCCCGGCGGCCGCTGCCAGTGCCGCGCGCTGGGCT
CGGGCATGGCGGTCGACTGCTCCACGCTGACCTCCAAGTGTCTGCTGCTCAAGGCGCGCAT
GAGCGCCCCCAAGAACGCCCGCACGCTGGTGCGGCCGAGTGAGCACGCGCTCGTGGACAAC
GATGGCCTCTACGACCCCGACTGCGACCCCGAGGGCCGCTTCAAGGCGCGCCAGTGCAACC
AGACGTCGGTGTGCTGGTGCGTGAACTCGGTGGGCGTGCGCCGCACGGACAAGGGCGACCT
GAGCCTACGCTGCGATGAGCTGGTGCGCACCCACCACATCCTCATTGACCTGCGCCACCGC
CCCACCGCCGGCGCCTTCAACCACTCAGACCTGGACGCCGAGCTGAGGCGGCTCTTCCGCG
AGCGCTATCGGCTGCACCCCAAGTTCGTGGCGGCCGTGCACTACGAGCAGCCCACCATCCA
GATCGAGCTGCGGCAGAACACGTCTCAGAAGGCCGCCGGTGAAGTGGATATCGGCGATGCC
GCCTACTACTTCGAGAGGGACATCAAGGGCGAGTCTCTATTCCAGGGCCGCGGCGGCCTGG
ACTTGCGCGTGCGCGGAGAACCCCTGCAGGTGGAGCGCACGCTCATCTATTACCTGGACGA
GATTCCCCCGAAGTTCTCCATGAAGCGCCTCACCGCCGGCCTCATCGCCGTCATCGTGGTG
GTCGTGGTGGCCCTCGTCGCCGGCATGGCCGTCCTGGTGATCACCAACCGGAGAAAGTCGG
GGAAGTACAAGAAGGTGGAGATCAAGGAACTGGGGGAGTTGAGAAAGGAACCGAGCTTG**TA**
GGTACCCGGCGGGGCAGGGGATGGGGTGGGGTACCGGATTTCGGTATCGTCCCAGACCCAA
GTGAGTCACGCTTCCTGATTCCTCGGCGCAAAGGAGACGTTTATCCTTTCAAATTCCTGCC
TTCCCCCTCCCTTTTGCGCACACACCAGGTTTAATAGATCCTGGCCTCAGGGTCTCCTTTC
TTTCTCACTTCTGTCTTGAGGGAAGCATTTCTAAAATGTATCCCCTTTCGGTCCAACAACA
GGAAACCTGACTGGGGCAGTGAAGGAAGGGATGGCACAGCGTTATGTGTAAAAAACAAGTA
TCTGTATGACAACCCGGGATCGTTTGCAAGTAACTGAATCCATTGCGACATTGTGAAGGCT
TAAATGAGTTTAGATGGGAAATAGCGTTGTTATCGCCTTGGGTTTAAATTATTTGATGAGT
TCCACTTGTATCATGGCCTACCCGAGGAGAAGAGGAGTTTGTTAACTGGGCCTATGTAGTA
GCCTCATTTACCATCGTTTGTATTACTGACCACATATGCTTGTCACTGGGAAAGAAGCCTG

TTTCAGCTGCCTGAACGCAGTTTGGATGTCTTTGAGGACAGACATTGCCCGGAAACTCAGT
CTATTTATTCTTCAGCTTGCCCTTACTGCCACTGATATTGGTAATGTTCTTTTTTGTAAAA
TGTTTGTACATATGTTGTCTTTGATAATGTTGCTGTAATTTTTTAAAATAAAACACGAATT
TAATAAAATATGGGAAAGGCACAAACCAGAAGTCGGCATTTGTGAAAAGTCCCTCCAGATT
TCTATCACTTTGGTCTCTAATTTCCCAAGACTTGTATTTTTTTTTTATTTCAAATTATAAC
ACTTTTTTTTCCCCCAGAAGTGGGTGTTTCATGTTGCTACTCTGGTGTGTCCCAAGATATC
CTAACTGGCCAGTGTAAATGCTATTCTTTCTAAATAAGATTATTTGGAAACTTCCTTCAAA
CTGCAGGAGGGCGAGCTCTGAGGGCACGAGAAGCTAAAACTAGCTGCTTTTGATGAAAAAG
AGTGCCAGTCTTTGGTCATCTCTAAACAAGGCTTATCACCAATGGAGACAGAAAACTCTAG
TTCAAGAGCTGTACCTCCTTTGAATCCCAGCCCTACTCGAAATAAGTGGTACTATTTCCAT
TTAGCCTTTGAGCAAATCACTTAACTCAAAGGCGTTGTGGCTCTAAGATTAAACGACTTT

FIGURE 76

GCGGCCGCGGAGCAAGAAGGGCGCCGCGTCGTGCGGCCCGCGCAGCCCCCGGAGCCATGGG
CAAGTGCAGCGGGCGCTGCACGCTGGTCGCCTTCTGCTGCCTGCAGCTGGTGGCTGCGCTG
GAGCGGCAGATCTTTGACTTCCTGGGCTACCAGTGGGCTCCCATCCTAGCCAACTTCCTGC
ACATCATGGCAGTCATCCTGGGCATCTTTGGCACCGTGCAGTACCGCTCCCGGTACCTCAT
CCTGTATGCAGCCTGGCTGGTGCTCTGGGTTGGCTGGAATGCATTTATCATCTGCTTCTAC
TTGGAGGTTGGACAGCTGTCCCAGGACCGGGACTTCATCATGACCTTCAACACATCCCTGC
ACCGCTCCTGGTGGATGGAGAATGGGCCAGGCTGCCTGGTGACACCTGTTCTGAACTCCCG
CCTGGCTCTGGAGGACCACCATGTCATCTCTGTCACTGGCTGCCTGCTTGACTACCCCTAC
ATTGAAGCCCTCAGCAGCGCCCTGCAGATCTTCCTGGCACTGTTCGGCTTCGTGTTCGCCT
GCTACGTGAGCAAAGTGTTCCTGGAGGAGGAGGACAGCTTTGACTTCATCGGCGGCTTTGA
CTCCTACGGATACCAGGCGCCCCAGAAGACGTCGCATTTACAGCTGCAGCCTCTGTACACG
TCGGGGTAGCCTCTGCCCCGCGCCCACCCCGGCGCCTCGCCCTGGGCTGACCGCAGCTGCC
GCGAGCTCGGGCCAAGGCGCAGGCGTGTCCCCCTGGTGGCCCGCGCTCTCACTGCAGCCTG
TGCCCAACCCCGCGTCTGCATCTGGAGATGCGGACTTGGACGTGGACTTGGACTTGGACTT
GGATTTGAGCTTGGCTCTTCGCAGCCCGGACTTCGGAGGAGTGGGGCGGGGCGGGGGAGGG
GCACCACGGGTTTTTTGTTTTTTGTTTGTTTGTTTTTAATCTCAGCCTTGGCGTGAGCTGG
GGCCTTCCTCTCTTCTCCAGCCTCTCCCTTTCACTCTTCGCCCAGCATCCTGCCCCCCTGT
CCAAAAACAGCAGGACATCAGGCCCATCCCATCCCACCACACTCACTCACCAGCTCTGGGG
AAAGCTACTGTGAACTAGGAGCAGGATTCCTGGGTTCTAATCGCAGGTCCATCACTGACTG
TGACGTCTAGCAAAGCCCTTGCCCTCTCTGAGCCTCGGTTTCCGCACCTCAAGTAATTAAT
CCCTTAGCAAATGGACTCTTTTAGACTTCTCATTTAACTCAATTCCCTGAGCTAGACTGGG
ATTAAAATTCTCATTTTGCAGTACATTAAAACTGATGCCCAGAGATGTGATTTGCTTGAGG
CCACACAGCTAGATTTTTGGTGGAAGTGGGCCTTGAACACAGTGTACTTTCTGCAGTTTCT
GACTGTAAAACCCAGTGTCTGCTCTCTGAGTTCCATTTCCAAGCCCCCCTCCATCTTGGAC
CTATGTGGTCTCCACCATATTCACACACCACCACCACCACTTGCCAATGCCTCTCTTAAAG
CAATATACCCATTCGTTCTCTTATTGGGAACTGGATGGATGAAGCCCCAAATTCAGCCCCA
CCCACAGAGAAGCCTTCCTACACTCAGCCTCTGTCCACCCTTGGCAAATCTTTCAAGCTCT
CTCCTCCAGGAAAGTGGGGCCCCAACTCAGTCACTCCACCCCCTTCCAGGTCCCTGAGGCT
GGTTCTACTGTATCCCCATCACCTCCACAACTCCACTCACCCCTGACGGCTCCATCCACCT
CACCAGTTGGAAGGCTTGTGGTTTCAGAGAGGAGCAATGCTGGTCAGCGCTGCCCAGACTC
CAGTGTTTACAGATCACCAGCATTTACAACCAATCCAATGGCCAGAAGCCTCCTCTAACAA
GCCCAGAAGGAGTTCTGAAGGGGCAGATGGGGGTGTGAGTAGTCGGGGAGTCGGGATTGCC
AGCACCCTCACCCTTCCTTGGGGGCAAGTAGAGGTGAGAACACTTTCCCCACCTCCCTCCA
CAGACACTCCTGAGGACGCTGCATCCCACGCACTGCCTGGTGCGTCCATAGAGAGAGGATC
AGGTCTCAGCATTTCATCTGTGAAAGAGGCATGGCCCTGGGTTAGAAAGGAGGGCAGGAGA

CATGGAGGAACTGGGGGGCACCCAGATGGTGCAGATGGTTTGCACACCTGAGCCTGTCTGT
GGTGACCATTCCGCTCCTCTCCCACTACCCTCCAATCTATCATTCCCTACTCTCTAAGGCC
AAAATATCCTGAGCAAGGCTGGCAACCCCACCCCACCATCCCAAATGCAAGCAGCCAGGCC
CAGGAGTTCCTCTGGCCCCCACAGGCATGGAGCTCCCAGCTGGTGGGTACAGCTTGAGAGG
GGGGCAGCTCCCTCAGGCTAAGCTACTGCCCTTCACTGGGCCAGCCCTGCCTCCAGCCCTC
ACCTCTCTCACCCCAACTCTCCCCCAAGCCCCTTTCTACTCAACGGGTGTAGCCACTGGTG
CTTTGAAGCCTTTTGTTTTTATAAGATGGTTTTTGCAAGGGGACCAGGTTCTCTTTTCACT
GGGACCTTGCAAGGAGGGGAGTGCTCTCCTGGTTTCTGTGCAGGCGGGTTGATTAAAGATG
GTGTTTTCTTCTCT

FIGURE 77

GGCGGGGAGAGGAGAGGAGAAGAGCCGCGGGGGGCCCAGCCCGGAGCCAGG**ATG**CCCGC
GCCGCGCGCCCGGGAGCAGCCCCGCGTGCCCGGGGAGCGCCAGCCGCTGCTGCCTCGCGGT
GCGCGGGGCCCTCGACGGTGGCGGCGGGCGGCGGGCGCGGCCGTGCTGCTGGTGGAGATGC
TGGAGCGCGCCGCCTTCTTCGGCGTCACCGCCAACCTCGTGCTGTACCTCAACAGCACCAA
CTTCAACTGGACCGGCGAGCAGGCGACGCGCGCCGCGCTGGTATTCCTGGGCGCCTCCTAC
CTGCTGGCGCCCGTGGGCGGCTGGCTGGCCGACGTGTACCTGGGCCGCTACCGCGCGGTCG
CGCTCAGCCTGCTGCTCTACCTGGCCGCCTCGGGCCTGCTGCCCGCCACCGCCTTCCCCGA
CGGCCGCAGCTCCTTCTGCGGAGAGATGCCCGCGTCGCCGCTGGGACCTGCCTGCCCCTCG
GCCGGCTGCCCGCGCTCCTCGCCCAGCCCCTACTGCGCGCCCGTCCTCTACGCGGGCCTGC
TGCTACTCGGCCTGGCCGCCAGCTCCGTCCGGAGCAACCTCACCTCCTTCGGTGCCGACCA
GGTGATGGATCTCGGCCGCGACGCCACCCGCCGCTTCTTCAACTGGTTTTACTGGAGCATC
AACCTGGGTGCTGTGCTGTCGCTGCTGGTGGTGGCGTTTATTCAGCAGAACATCAGCTTCC
TGCTGGGCTACAGCATCCCTGTGGGCTGTGTGGGCCTGGCATTTTTCATCTTCCTCTTTGC
CACCCCCGTCTTCATCACCAAGCCCCCGATGGGCAGCCAAGTGTCCTCTATGCTTAAGCTC
GCTCTCCAAAACTGCTGCCCCCAGCTGTGGCAACGACACTCGGCCAGAGACCGTCAATGTG
CCCGCGTGCTGGCCGACGAGAGGTCTCCCCAGCCAGGGGCTTCCCCGCAAGAGGACATCGC
CAACTTCCAGGTGCTGGTGAAGATCTTGCCCGTCATGGTGACCCTGGTGCCCTACTGGATG
GTCTACTTCCAGATGCAGTCCACCTATGTCCTGCAGGGTCTTCACCTCCACATCCCAAACA
TTTTCCCAGCCAACCCGGCCAACATCTCTGTGGCCCTGAGAGCCCAGGGCAGCAGCTACAC
GATCCCGGAAGCCTGGCTCCTCCTGGCCAATGTTGTGGTGGTGCTGATTCTGGTCCCTCTG
AAGGACCGCTTGATCGACCCTTTACTGCTGCGGTGCAAGCTGCTTCCCTCTGCTCTGCAGA
AGATGGCGCTGGGGATGTTCTTTGGTTTTACCTCCGTCATTGTGGCAGGAGTCCTGGAGAT
GGAGCGCTTACACTACATCCACCACAACGAGACCGTGTCCCAGCAGATTGGGGAGGTCCTG
TACAACGCGGCACCACTGTCCATCTGGTGGCAGATCCCTCAGTACCTGCTCATTGGGATCA
GTGAGATCTTTGCCAGCATCCCAGGCCTGGAGTTTGCCTACTCAGAGGCCCCGCGCTCCAT
GCAGGGCGCCATCATGGGCATCTTCTTCTGCCTGTCGGGGGTGGGCTCACTGTTGGGCTCC
AGCCTAGTGGCACTGCTGTCCTTGCCCGGGGGCTGGCTGCACTGCCCCAAGGACTTTGGGA
ACATCAACAATTGCCGGATGGACCTCTACTTCTTCCTGCTGGCTGGCATTCAGGCCGTCAC
GGCTCTCCTATTTGTCTGGATCGCTGGACGCTATGAGAGGGCGTCCCAGGGCCCAGCCTCC
CACAGCCGTTTCAGCAGGGACAGGGGC**TGA**ACAGGCCCTATTCCAGCCCCCTTGCTTCAC

FIGURE 78A

GGCCGCGGGCTCTCGCGGGGCGGCGACGCCGCGGGGAGGATGCTGCTTGCCGCGCCCGCGT
CCTCACCGTCCTCCCGGGCCGCCTGCTGGGGCTTTGTTGTGGCCCGGACGCCGCGGGCCAC
CCCCTGAAGTCGCCTGCCGCCGCCGCCGCCGCACCTAGCGGACGGGCGGGCGGGCGCGCGT
GTGCCCAGGAGTGCGCGCCTGTCGCGGTGGTGGGTGCAGGACTGGACCCACGGGCCCATTG
TGCGCCCGCCCGCGGCAGCCAGGACCATGTGGGTGAACCCGGAGGAGGTGTTGCTGGCCAA
CGCGCTGTGGATCACCGAGAGGGCCAACCCATACTTCATCCTGCAGCGGAGGAAGGGCCAC
GCCGGCGATGGAGGCGGCGGCGGCGGACTGGCGGGCCTGCTGGTGGGTACCCTTGATGTTG
TGTTGGACTCCAGCGCCCGGGTCGCTCCTTACCGAATCTTGTACCAGACTCCAGACTCCCT
GGTCTACTGGACCATCGCCTGTGGTGGTTCCAGGAAAGAAATCACTGAACACTGGGAATGG
CTTGAGCAAAATCTCTTGCAGACACTCTCCATCTTTGAAAATGAGAATGATATCACCACAT
TTGTGAGAGGAAAAATACAGGGCATCATTGCAGAATACAACAAAATCAATGATGTAAAGGA
AGATGATGACACGGAGAAGTTTAAAGAAGCCATTGTGAAATTTCATAGGCTGTTTGGGATG
CCAGAGGAAGAGAAACTCGTCAACTATTACTCTTGCAGCTATTGGAAGGGGAAGGTCCCCC
GTCAGGGTTGGATGTACCTCAGCATTAACCACCTTTGCTTTTATTCTTTTCTTATGGGAAG
GGAAGCGAAACTGGTCATCCGGTGGGTAGACATCACTCAGCTTGAGAAGAATGCCACCCTG
CTTCTGCCTGATGTGATCAAAGTGAGCACACGGTCCAGTGAGCATTTCTTCTCTGTATTCC
TCAACATCAACGAGACCTTCAAGTTAATGGAGCAGCTTGCCAACATAGCCATGAGGCAACT
CTTAGACAATGAGGGATTTGAACAAGATCGATCCCTGCCCAAACTCAAAAGGAAATCTCCT
AAAAAAGTGTCTGCTCTAAAACGTGATCTTGATGCCAGGGCAAAGAGTGAGAGATACCGTG
CACTTTTCCGGCTGCCCAAAGATGAAAAATTAGATGGCCACACAGACTGCACTCTCTGGAC
TCCATTTAACAAAATGCACATTTTGGGGCAGATGTTTGTGTCCACAAATTACATCTGTTTT
ACCAGCAAGGAGGAGAACTTATGTAGCCTCATTATCCCGCTCCGTGAGGTGACAATTGTGG
AAAAGGCAGACAGCTCCAGTGTGCTCCCCAGTCCCTTATCCATCAGCACCCGAAACAGGAT
GACCTTCCTATTTGCCAACTTGAAAGATAGAGACTTTCTAGTGCAGAGGATCTCAGATTTC
CTGCAACAGACTACTTCCAAAATATATTCTGACAAGGAGTTTGCAGGAAGTTACAACAGTT
CAGATGATGAGGTGTACTCTCGACCCAGCAGCCTCGTCTCCTCCAGCCCCCAGAGAAGCAC
GAGCTCTGATGCTGATGGAGAGCGCCAGTTTAACCTAAATGGCAACAGCGTCCCCACAGCC
ACACAGACCCTGATGACCATGTATCGGCGGCGGTCTCCCGAGGAGTTCAACCCGAAATTGG
CCAAAGAGTTTCTGAAAGAGCAAGCCTGGAAGATTCACTTTGCTGAGTATGGGCAAGGGAT
CTGCATGTACCGCACAGAGAAAACGCGGGAGCTGGTGTTGAAGGGCATCCCGGAGAGCATG
CGTGGGGAGCTCTGGCTGCTGCTGTCAGGTGCCATCAATGAGAAGGCCACACATCCTGGGT
ACTATGAAGACCTAGTGGAGAAGTCCATGGGGAAGTATAATCTCGCCACGGAGGAGATTGA
GAGGGATTTACACCGCTCCCTTCCAGAACACCCAGCTTTTCAGAATGAAATGGGCATTGCT
GCACTAAGGAGAGTCTTAACAGCTTATGCTTTTCGAAATCCCAACATAGGGTATTGCCAGG
CCATGAATATTGTCACTTCAGTGCTGCTGCTTTATGCCAAAGAGGAGGAAGCTTTCTGGCT

GCTTGTGGCTTTGTGTGAGCGCATGCTCCCAGATTACTACAACACCAGAGTTGTGGGTGCA
CTGGTGGACCAAGGTGTCTTTGAGGAGCTAGCACGAGACTACGTCCCACAGCTGTACGACT
GCATGCAAGACCTGGGCGTGATTTCCACCATCTCCCTGTCTTGGTTCCTCACACTATTTCT
CAGTGTGATGCCTTTTGAGAGTGCAGTTGTGGTTGTTGACTGTTTCTTCTATGAAGGAATT
AAAGTGATATTCCAGTTGGCCCTAGCTGTGCTGGATGCAAATGTGGACAAACTGTTGAACT
GCAAGGATGATGGGGAGGCCATGACCGTTTTGGGAAGGTATTTAGACAGTGTGACCAATAA
AGACAGCACACTGCCTCCCATTCCTCACCTCCACTCCTTGCTCAGCGATGATGTGGAACCT
TACCCTGAGGTAGACATCTTTAGACTCATCAGAACTTCCTACGAGAAATTCGGAACTATCC
GGGCAGATTTGATTGAACAGATGAGATTCAAACAGAGACTGAAAGTGATCCAGACGCTGGA
GGATACTACGAAACGCAACGTGGTACGAACCATTGTGACAGAAACTTCCTTTACCATTGAT
GAGCTGGAAGAACTTTATGCTCTTTTCAAGGCAGAACATCTCACCAGCTGCTACTGGGGCG
GGAGCAGCAACGCGCTGGACCGGCATGACCCCAGCCTGCCCTACCTGGAACAGTATCGCAT
TGACTTCGAGCAGTTCAAGGGAATGTTTGCTCTTCTCTTTCCTTGGGCATGTGGAACTCAC
TCTGACGTTCTGGCCTCCCGCTTGTTCCAGTTATTAGATGAAAATGGAGACTCTTTGATTA
ACTTCCGGGAGTTTGTCTCTGGGCTAAGTGCTGCATGCCATGGGGACCTCACAGAGAAGCT
CAAACTCCTGTACAAAATGCACGTCTTGCCTGAGCCATCCTCTGATCAAGATGAACCAGAT
TCTGCTTTTGAAGCAACTCAGTACTTCTTTGAAGATATTACCCCAGAATGTACACATGTTG
TTGGATTGGATAGCAGAAGCAAACAGGGTGCAGATGATGGCTTTGTTACGGTGAGCCTAAA
GCCAGACAAAGGGAAGAGAGCAAATTCCCAAGAAAATCGTAATTATTTGAGACTGTGGACT
CCAGAAAATAAATCTAAGTCAAAGAATGCAAAGGATTTACCCAAATTAAATCAGGGGCAGT
TCATTGAACTGTGTAAGACAATGTATAACATGTTCAGCGAAGACCCCAATGAGCAGGAGCT
GTACCATGCCACGGCAGCAGTGACCAGCCTCCTGCTGGAGATTGGGGAGGTCGGCAAGTTG
TTCGTGGCCCAGCCTGCAAAGGAGGGCGGGAGCGGAGGCAGTGGGCCGTCCTGCCACCAGG
GCATCCCAGGCGTGCTCTTCCCCAAGAAAGGGCCAGGCCAGCCTTACGTGGTGGAGTCTGT
TGAGCCCCTGCCGGCCAGCCTGGCCCCCGACAGCGAGGAACACTCCCTTGGAGGACAAATG
GAGGACATCAAGCTGGAGGACTCCTCGCCCCGGGACAACGGGGCCTGCTCCTCCATGCTGA
TCTCTGACGACGACACCAAGGACGACAGCTCCATGTCCTCATACTCGGTGCTGAGTGCCGG
CTCCCACGAGGAGGACAAGCTGCACTGCGAGGACATCGGAGAGGACACGGTCCTGGTGCGG
AGCGGCCAGGGCACGGCGGCACTGCCCCGGAGCACCAGCCTGGACCGGGACTGGGCCATCA
CCTTCGAGCAGTTCCTGGCCTCCCTCTTAACTGAGCCTGCCCTGGTCAAGTACTTTGACAA
GCCCGTGTGCATGATGGCCAGGATTACCAGTGCAAAAAACATCCGGATGATGGGCAAGCCC
CTCACCTCGGCCAGTGACTATGAAATCTCGGCCATGTCCGGC**TGA**CACGGGCGCCTTCCCG
GGGGAGTGGGAGGAGAGGGAGGGGAGGGATTTTTTATGTTCTTCTGTGTTGAGTTTTTTCT
TTCTTTCTTTTAAATTAAATATTTATTAGTACCTGGCTTGAAGCCTAGTGTTTTCATAATG
TAATTCAATGAAAACTGTTGGAGAAATATTTAAACACCTCAATGTAGGTACATTACACTCT
TGTTGCGGGGAGGGGATTTACCAGAATACAGTTTATTTCGTGAATTCTAAAAAACAAAAAG
ATGAATCTGTCAGTGATATGTGTGTATTATAACTTATTAATCTTGCTGTTGAGCTGTATAC

ATGGTTTAAAAAATAGTACTGTTTAATGCTAAGTAAGGCAGCAGTCATTTGTGTATTCAGG
CTTTTTAAATAAAATTAGAGCTGTAAGGAAAATGAAAAGCCACAAATGCAAGACTGTTCTT

FIGURE 78B

AAATGGAAGGCATAGTCAGCGAGGGTAAATCCTATACCACTTTAGGAAGTATTAAAAATAT
TTTTAAGATTTGAAATATATTTCATAGAAGTCCTCTATTCAAAATCATATTCCACAGATGT
TCCCCTTCAAAGGGAAAACATTTGGGGTTCTAAACAGTTATGAAAGTAAGTGATTTTTACA
TGATTCCAGAATAACACTTGTATTGACCAATTTAGACAGATACCAGACCAATTTTGCATTT
AAGAAATTGTTCTGATTATTTACGTCAACTCATTAGAATTCAGTGAAAAGTAACAGTCTTT
TGTCACAGAGAATCTGAAAGTAGCAGCAAAGACAGAGGGCTCATGACAGGTTTTTGCTTTT
GCTTTGCTTTTGTTTTTGAAAGAGTAAAAGTACTGATGCTTCTGATACTGGATGTTTAGCT
TCTTACTGCAAAAACATAAGTAAAACAGTCAACTTTACCATTTCCGTATTCTCCATAGATT
GAAGAAATTTATACCACATATCGCATATGACCATCTTTCCATCAAATCAATGTAGAGATAA
TGTAAACTGAAAAAAAATCTGCAAGATAATGTAACTGAATGTTTTAAAAACAGAACTTGTC
ACTTTATATAAAAGAATAGTATGCTCTATTTCCTGAATGGATGTGGAAATGAAAGCTAGCG
CACCTGCACTTTGAATTCTTGCTTCTTTTTTATTACTGTTATGATTTTGCTTTTTACAGAT
GTTGGACGATTTTTTCTTCTGATTGTTGAATTCATAATCATGGTCTCATTTCCTTTGCTTC
TTTGGAATATTTCTTTCAACACATTCCTTTATTTTATTATACATTGTGTCCTTTTTTTAGC
TATTGCTGCTGTTGTTTTTTATTCTATTTACAGGATGATTTTTAAACTGTCAAATGAAGTA
GTGTTAACCTCAAATAGGCTAAATGTGAACAAATAAAATACAGCAAATACTCAGATACAGC
TTTTTATCTTTGTGCTTGAGTTCCTGCCTAAGGAATAACATTATTCTTTTGACAACTTTTG
CAGGGGAAATTATATCAGGCAACCATTTTGATTAAGTAAATAAATTTTATAGGCAAACATA
TAGAGAGANATACAATTTGTAGTATATCAATGACTATATTTAAAATAAGGANTATAATTGT
TATCAGTTATCTAACTTAAAATGCTTATCCATAATGATCAGTGATATTCAGCTTTTTAAAA
TATGCTTGTTGG

FIGURE 79

MRAAYLFLLFLPAGLLAQGQYDLDPLPPFPDHVQYTHYSDQIDNPDYYDYQEVTPRPSEEQ
FQFQSQQQVQQEVIPAPTPEPGNAELEPTEPGPLDCREEQYPCTRLYSIHRPCKQCLNEVC
FYSLRRVYVINKEICVRTVCAHEELLRADLCRDKFSKCGVMASSGLCQSVAASCARSCGSC

Signal sequence.
amino acids 1-19

N-myristoylation sites.
amino acids 161-166, 167-172

FIGURE 80

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNEGRADEVVSASVRSGDLWIPVKSFDS
KNHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTVILGHCYYH
GHVRGYSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHH
NTPNLAAKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLEKVKQRLIEI
ANHVDKFYRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQL
VSGVYFQGTTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRG
CSCQMAVEKGGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKC
GNRFVEEGEECDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNS
CDLPEFCTGASPHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICF
ERVNSAGDPYGNCGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQG
GRILCRGTHVYLGDDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCN
NRKNCHCEAHWAPPFCDKFGFGGSTDSGPIRQADNQGLTIGILVTILCLLAAGFVVYLKRK
TLIRLLFTNKKTTIEKLRCVRPSRPPRGFQPCQAHLGHLGKGLMRKPPDSYPPKDNPRRLL
QCQNVDISRPLNGLNVPQPQSTQRVLPPLHRAPRAPSVPARPLPAKPALRQAQGTCKPNPP
QKPLPADPLARTTRLTHALARTPGQWETGLRLAPLRPAPQYPHQVPRSTHTAYIK

Signal sequence.
amino acids 1-28

Transmembrane domain.
amino acids 709-729

N-glycosylation sites.
amino acids 111-114, 149-152, 381-384, 452-455, 651-654

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 206-209, 730-733, 742-745

N-myristoylation sites.
amino acids 88-93, 103-108, 177-182, 180-185, 313-318, 343-348, 377-382, 407-412, 423-428, 424-429, 524-529, 630-635, 635-640, 656-661, 668-673, 694-699, 708-713, 712-717, 878-883

Neutral zinc metallopeptidases, zinc-binding region

signature.

amino acids 347-356

Disintegrins proteins.

amino acids 454-504, 410-460, 486-536

FIGURE 81

MASRSMRLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQSY
GNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPQKRQQWQWIDGAMYLYRSWSGKS
MGGNKHCAEISSNNNFLTWSSNECNKRQHFLCKYRP

Signal sequence.
amino acids 1-22

N-glycosylation site.
amino acids 50-53

N-myristoylation sites.
amino acids 64-69, 125-130

C-type lectin domain signature.
amino acids 129-154

Lectin C-type domain.
amino acids 47-156

FIGURE 82

MATMENKVICALVLVSMLALGTLAEAQTETCTVAPRERQNCGFPGVTPSQCANKGCCFDDT
VRGVPWCFYPNTIDVPPEEECEF

Signal sequence.
amino acids 1-24

N-myristoylation sites.
amino acids 45-50, 64-69

Trefoil (P-type) domain.
amino acids 30-71

FIGURE 83

MLPQIPFLLLVSLNLVHGVFYAERYQTPTGIKGPLPNTKTQFFIPYTIKSKGIAVRGEQGT
PGPPGPAGPRGHPGPSGPPGKPGYGSPGLQGEPGLPGPPGPSAVGKPGVPGLPGKPGERGP
YGPKGDVGPAGLPGPRGPPGPPGIPGPAGISVPGKPGQQGPTGAPGPRGFPGEKGAPGVPG
MNGQKGEMGYGAPGRPGERGLPGPQGPTGPSGPPGVGKRGENGVPGQPGIKGDRGFPGEMG
PIGPPGPQGPPGERGPEGIGKPGAAGAPGQPGIPGTKGLPGAPGIAGPPGPPGFGKPGLPG
LKGERGPAGLPGGPGAKGEQGPAGLPGKPGLTGPPGNMGPQGPKGIPGSHGLPGPKGETGP
AGPAGYPGAKGERGSPGSDGKPGYPGKPGLDGPKGNPGLPGPKGDPGVGGPPGLPGPVGPA
GAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIPGFPGSKGDPGSPGPPGPAGIATKGLNGP
TGPPGPPGPRGPSGEPGLPGPPGPPGPPGPPGQAVMPEGFIKAGQRPSLSGTPLVSANQGVTGM
PVSAFTVILSKAYPAIGTPIPFDKILYNRQQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHV
WVGLYKNGTPVMYTYDEYTKGYLDQASGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSS
FSGFLVAPM

Signal sequence.
amino acids 1-18

Tyrosine kinase phosphorylation site.
amino acids 116-123

N-myristoylation sites.
amino acids 18-23, 276-281, 317-322, 350-355, 380-385,
415-420, 446-451, 529-534, 548-553, 587-592, 613-618, 661-666

Amidation site.
amino acids 219-222

C1q domain signature.
amino acids 571-601

C1q domain.
amino acids 553-677

Collagen triple helix repeat (20 copies).
amino acids 92-150, 156-214, 223-281, 282-340, 344-403,

146247

404-463, 464-522

FIGURE 84

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQ
YFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFL
ETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTW
SQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQSSEFKAKIIFWY
VLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAEKIVINFITLNI
SDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENT
EGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQAAL
AVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQ
DSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMEN

Signal sequence.
amino acids 1-18

Transmembrane domain.
amino acids 239-259

N-glycosylation sites.
amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192, 304-307, 523-526

Tyrosine kinase phosphorylation sites.
amino acids 385-392, 518-526

N-myristoylation sites.
amino acids 53-58, 106-111, 368-373, 492-497

Phosphatidylinositol 3- and 4-kinases proteins.
amino acids 253-279

Tissue factor.
amino acids 1-278

FIGURE 85

MRAPGRPALRPLPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEG
LQGVKVTYTVQYFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSK
WAESGRFYPFLETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNV
SVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQS
SEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAE
KIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHL
MEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVST
QGTLLESQAALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGR
LCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYV
QMEN

Signal sequence.
amino acids 1-29

Transmembrane domain.
amino acids 240-270

FIGURE 86

MAAARLCLSLLLLSTCVALLLQPLLGAQGAPLEPVYPGDNATPEQMAQYAADLRRYINMLT
RPRYGKRHKEDTLAFSEWGSPHAAVPRELSPLDL

Signal sequence.
amino acids 1-29

N-myristoylation site.
amino acids 80-85

Amidation site.
amino acids 65-68

Pancreatic hormone family signature.
amino acids 49-65

Pancreatic hormone peptide.
amino acids 30-65

FIGURE 87

MGPPSAPPCRLHVPWKEVLLTASLLTFWNPPTTAKLTIESTPFNVAEGKEVLLLAHNLPQN
RIGYSWYKGERVDGNSLIVGYVIGTQQATPGPAYSGRETIYPNASLLIQNVTQNDTGFYTL
QVIKSDLVNEEATGQFHVYPELPKPSISSNNSNPVEDKDAVAFTCEPEVQNTTYLWWVNGQ
SLPVSPRLQLSNGNMTLTLLSVKRNDAGSYECEIQNPASANRSDPVTLNVLYGPDGPTISP
SKANYRPGENLNLSCHAASNPPAQYSWFINGTFQQSTQELFIPNITVNNSGSYMCQAHNSA
TGLNRTTVTMITVSGSAPVLSAVATVGITIGVLARVALI

Signal sequence.
amino acids 1-34

Transmembrane domain.
amino acids 322-342

N-glycosylation sites.
amino acids 104-107, 111-114, 115-119, 152-155, 173-176, 197-200, 224-227, 256-259, 274-277, 288-291, 292-295, 309-312

Tyrosine kinase phosphorylation site.
amino acids 206-213

N-myristoylation sites.
amino acids 85-90, 211-216, 295-300, 307-312, 332-337

Immunoglobulin domain.
amino acids 160-217, 252-301

FIGURE 88

MELRVLLCWASLAAALEETLLNTKLETADLKWVTFPQVDGQWEELSGLDEEQHSVRTYEVC
DVQRAPGQAHWLRTGWVPRRGAVHVYATLRFTMLECLSLPRAGRSCKETFTVFYYESDADT
ATALTPAWMENPYIKVDTVAAEHLTRKRPGAEATGKVNVKTLRLGPLSKAGFYLAFQDQGA
CMALLSLHLFYKKCAQLTVNLTRFPETVPRELVVPVAGSCVVDAVPAPGPSPSLYCREDGQ
WAEQPVTGCSCAPGFEAAEGNTKCRACAQGTFKPLSGEGSCQPCPANSHSNTIGSAVCQCR
VGYFRARTDPRGAPCTTPPSAPRSVVSRLNGSSLHLEWSAPLESGGREDLTYALRCRECRP
GGSCAPCGGDLTFDPGPRDLVEPWVVVRGLRPDFTYTFEVTALNGVSSLATGPVPFEPVNV
TTDREVPPAVSDIRVTRSSPSSLSLAWAVPRAPSGAVLDYEVKYHEKGAEGPSSVRFLKTS
ENRAELRGLKRGASYLVQVRARSEAGYGPFGQEHHSQTQLDESEGWREQLALIAGTAVVGV
VLVLVVIVVAVLCLRKQSNGREAEYSDKHGQYLIGHGTKVYIDPFTYEDPNEAVREFAKEI
DVSYVKIEEVIGAGEFGEVCRGRLKAPGKKESCVAIKTLKGGYTERQRREFLSEASIMGQF
EHPNIIRLEGVVTNSMPVMILTEFMENGALDSFLRLNDGQFTVIQLVGMLRGIASGMRYLA
EMSYVHRDLAARNILVNSNLVCKVSDFGLSRFLEENSSDPTYTSSLGGKIPIRWTAPEAIA
FRKFTSASDAWSYGIVMWEVMSFGERPYWDMSNQDVINAIEQDYRLPPPPDCPTSLHQLML
DCWQKDRNARPRFPQVVSALDKMIRNPASLKIVARENGGASHPLLDQRQPHYSAFGSVGEW
LRAIKMGRYEESFAAAGFGSFELVSQISAEDLLRIGVTLAGHQKKILASVQHMKSQAKPGT
PGGTGGPAPQY

Signal sequence.
amino acids 1-15
Transmembrane domain.
amino acids 539-559
N-glycosylation sites.
amino acids 203-206, 335-338, 426-429, 768-771
Glycosaminoglycan attachment site.
amino acids 280-283
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 564-567, 639-642, 795-798
Tyrosine kinase phosphorylation sites.
amino acids 508-515, 588-596, 608-614, 729-736
N-myristoylation sites.
amino acids 152-157, 182-187, 264-269, 298-303, 317-322, 374-379, 496-501, 681-686, 699-704, 719-724, 723-728,

951-956, 975-980

Amidation site.

amino acids 637-640

Tyrosine protein kinases specific active-site signature.

amino acids 736-748

FIGURE 89

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALWAWPSELSS
PGPEASLPTEPAAQEKSLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGDCVPSPCHNGGTCLEEEEGVRCL
CLPGYGGDLCDVGLRFCNPGWDAFQGACYKHFSTRRSWEEAETQCRMYGAHLASISTPEEQ
DFINNRYREYQWIGLNDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG
QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAQVFGRPRLRYEVDTVLRYRCREGLAQRNLP
LIRCQENGRWEAPQISCVPRRPARALHPEEDPEGRQGRLLGRWKALLIPPSSPMPGP

Signal peptide.
amino acids 1-15
N-glycosylation sites.
amino acids 130-133, 337-340
Tyrosine kinase phosphorylation sites.
amino acids 128-135, 451-459
N-myristoylation sites.
amino acids 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591, 677-682, 684-689, 720-725, 772-777, 811-816
EGF-like domain cysteine pattern signature.
amino acids 670-682
C-type lectin domain signature.
amino acids 784-809
Immunoglobulins and major histocompatibility complex proteins signature.
amino acids 135-142

Link domain proteins.

amino acids 166-216, 264-314

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 655-676

FIGURE 90

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALWAWPSELSS
PGPEASLPTEPAAQEESLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGNSAQGSTALSILLLFFPLQLWVT

Signal sequence.
amino acids 1-15

Transmembrane domain.
amino acids 652-670

N-glycosylation sites.
amino acids 130-133, 337-340

Tyrosine kinase phosphorylation sites.
amino acids 128-135, 451-459

N-myristoylation sites.
amino acids 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591

Immunoglobulins and major histocompatibility complex.
amino acids 135-141

Extracellular link domain.
amino acids 156-251, 257-353

Immunoglobulin domain.

amino acids 50-139

FIGURE 91

MAVRQWVIALALAALLVVDREVPVAAGKLPFSRMPICEHMVESPTCSQMSNLVCGTDGLTY
TNECQLCLARIKTKQDIQIMKDGKC

Signal sequence.
amino acids 1-20

N-myristoylation site.
amino acids 58-63

Kazal-type serine protease inhibitor domain.
amino acids 37-86

FIGURE 92

MAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLVPVHPQRLPRMQEDSPLGGGSSGEDDPLG
EEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPEVKPKSEEEGSLKLEDLPTVEAPGDP
QEPQNNAHRDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELL
GFQLPPLPELRLRNNGHSVQLTLPPGLEMALGPGREYRALQLHLHWGAAGRPGSEHTVEGH
RFPAEIHVVHLSTAFARVDEALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSE
TQVPGLDISALLPSDFSRYFRYEGSLTTPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGP
GDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVFGLLFAVT
SVAFLVQMRRQHRRGTKGGVSYRPAEVAETGA

Signal sequence.
amino acids 1-37

Transmembrane domain.
amino acids 409-429

N-glycosylation site.
amino acids 346-349

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 440-443

N-myristoylation sites.
amino acids 50-55, 51-56, 209-214, 236-241, 310-315, 339-344, 421-426, 442-447

Eukaryotic-type carbonic anhydrases signature.
amino acids 237-253

Eukaryotic-type carbonic anhydrase.
amino acids 141-390

FIGURE 93

MDALQLANSAFAVDLFKQLCEKEPLGNVLFSPICLSTSLSLAQVGAKGDTANEIGQVLHFE
NVKDVPFGFQTVTSDVNKLSSFYSLKLIKRLYVDKSLNLSTEFISSTKRPYAKELETVDFK
DKLEETKGQINNSIKDLTDGHFENILADNSVNDQTKILVVNAAYFVGKWMKKFSESETKEC
PFRVNKTDTKPVQMMNMEATFCMGNIDSINCKIIELPFQNKHLSMFILLPKDVEDESTGLE
KIEKQLNSESLSQWTNPSTMANAKVKLSIPKFKVEKMIDPKACLENLGLKHIFSEDTSDFS
GMSETKGVALSNVIHKVCLEITEDGGDSIEVPGARILQHKDELNADHPFIYIIRHNKTRNI
IFFGKFCSP

Signal sequence.
amino acids 1-42

N-glycosylation sites.
amino acids 99-102, 133-136, 188-191, 361-364

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 173-176

N-myristoylation sites.
amino acids 130-135, 207-212, 306-311, 312-317

Serpin (serine protease inhibitor).
amino acids 3-375

Fibrillarin proteins.
amino acids 150-169

FIGURE 94

MDRHSSYIFIWLQLELCAMAVLLTKGEIRCYCDAAHCVATGYMCKSELSACFSRLLDPQNS
NSPLTHGCLDSLASTTDICQAKQARNHSGTTIPTLECCHEDMCNYRGLHDVLSPPRGEASG
QGNRYQHDGSRNLITKVQELTSSKELWFRAAVIAVPIAGGLILVLLIMLALRMLRSENKRL
QDQRQQMLSRLHYSFHGHHSKKGQVAKLDLECMVPVSGHENCCLTCDKMRQADLSNDKILS
LVHWGMYSGHGKLEFV

Signal sequence.
amino acids 1-26

Transmembrane domain.
amino acids 152-172

N-glycosylation site.
amino acids 87-90

Glycosaminoglycan attachment sites.
amino acids 121-124, 252-255

N-myristoylation sites.
amino acids 68-73, 249-254

FIGURE 95

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLTKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKHISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGRIADVIASLVEVEDAPTSSAVKEKLVRFLPRTVSRLSEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFTAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIANKIK

Transmembrane domains.
amino acids 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

N-glycosylation sites.
amino acids 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 83-86

Tyrosine kinase phosphorylation sites.
amino acids 21-27, 219-226, 279-286

N-myristoylation sites.
amino acids 71-76, 141-146, 174-179, 209-214, 299-304,

577-582, 643-648, 904-909

Amidation site.
amino acids 102-105

FIGURE 96

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLPEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFIAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIANKIK

Transmembrane domains.
amino acids 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

N-glycosylation sites.
amino acids 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 83-86

Tyrosine kinase phosphorylation sites.
amino acids 21-27, 219-226, 279-286

N-myristoylation sites.
amino acids 71-76, 141-146, 174-179, 209-214, 299-304,

577-582, 643-648, 904-909

Amidation site.

amino acids 102-105

FIGURE 97

RRGRGEVLAVERGSGSGGGGTRSGWPAPAAGADKKPSRCGSGREGEGVSEGHKSMTGLYEL
VWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPAVGRNRRH
HRHPRGGSCLAAAHHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVVWCMAVGI
SYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLNCHLAVKV
LSPEDGKADIVRAAQDFCQLVAQKQRRPTDLDVDTLASLLSSNGCPDPDLVLKFGPVDSTL
GFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Transmembrane domain.
amino acids 166-186

N-glycosylation sites.
amino acids 198-201, 325-328

Glycosaminoglycan attachment sites.
amino acids 14-17, 16-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 34-37, 270-273

N-myristoylation sites.
amino acids 13-18, 15-20, 17-22, 19-24, 20-25, 47-52, 129-134, 156-161, 215-220

FIGURE 98

MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPA
VGRNRRHHRHPRGGSCLAAAHHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVV
WCMAVGISYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLN
CHLAVKVLSPEDGKADIVRAAQDFCQLVAQKQKRPTDLDVDTLASLLSSNGCPDPDLVLKF
GPVDSTLGFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Signal sequence.
amino acids 1-23

Transmembrane domain.
amino acids 115-132

N-glycosylation sites.
amino acids 144-147, 271-274

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 216-219

N-myristoylation sites.
amino acids 75-80, 102-107, 161-166

FIGURE 99

MVQRLWVSRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLLEVGVEEKFMTMVLGIGP
VLGLVCVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPLEL
ALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPAIDW
DTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSLSPHCCPC
RARLAFRNLGALLPRLHQLCCRMPRTLRRLFVAELCSWMALMTFTLFYTDFVGEGLYQGVP
RAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFGTRAVYLASVAAFPVAA
GATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDSL
MTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGI
CLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLA
KYSA

Signal sequence.
amino acids 1-35

Transmembrane domains.
amino acids 15-35, 52-72, 87-107, 122-142, 160-180, 240-260, 277-297, 321-341, 486-506, 521-541

N-myristoylation sites.
amino acids 27-32, 95-100, 171-176, 299-304, 303-308, 320-325, 419-424, 447-452, 462-467

Amidation site.
amino acids 82-85

Cell attachment sequence.
amino acids 415-417

FIGURE 100

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAV
REFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSE
ASIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIA
AGMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRW
TAPEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPS
ALHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTS
FNTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGLTLAGHQKKILNSIQVMR
AQMNQIQSVEV

Signal sequence.

amino acids 1-18

Transmembrane domain.

amino acids 542-562

N-glycosylation sites.

amino acids 265-268, 336-339, 428-431, 482-485, 705-708

Glycosaminoglycan attachment site.

amino acids 367-370

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 802-805

Tyrosine kinase phosphorylation sites.

amino acids 374-381, 595-603, 736-743

N-myristoylation sites.

amino acids 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 706-711, 726-731, 730-735, 875-880, 958-963

Amidation site - amino acids 644-647

Tyrosine protein kinases specific active-site signature - amino acids 743-755

Receptor tyrosine kinase class V signature 1 - amino acids 182-198

Receptor tyrosine kinase class V signature 2 - amino acids 241-261

Ephrin receptor ligand binding domain - amino acids 20-197

Protein kinase domain - amino acids 622-881

Fibronectin type III domain - amino acids 325-421, 436-520

SAM domain (Sterile alpha motif) - amino acids 912-976

FIGURE 101

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAVR
EFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSEA
SIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIAA
GMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRWT
APEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPSA
LHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTSF
NTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGVTLAGHQKKILNSIQVMRA
QMNQIQSVEV

Signal sequence.
amino acids 1-18
Transmembrane domain.
amino acids 543-563
N-glycosylation sites.
amino acids 265-268, 336-339, 428-431, 482-485, 704-707
Glycosaminoglycan attachment site.
amino acids 367-370
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 801-804
Tyrosine kinase phosphorylation sites.
amino acids 374-381, 594-602, 735-742
N-myristoylation sites.
amino acids 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 705-710, 725-730, 729-734, 874-879, 957-962

Amidation site.

amino acids 643-646

Tyrosine protein kinases specific active-site signature.

amino acids 742-754

Receptor tyrosine kinase class V signature 1.

amino acids 182-198

Receptor tyrosine kinase class V signature 2.

amino acids 241-264

Ephrin receptor ligand binding domain.

amino acids 20-197

Protein kinase domain.

amino acids 621-880

Fibronectin type III domains.

amino acids 325-421, 436-520

SAM domain (Sterile alpha motif).

amino acids 911-975

Src homology 2 (SH2) domains.

amino acids 733-769, 787-797, 806-819

FIGURE 102

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVNICYAGNTYFSFIAWRSRTIQ

Signal sequence.
amino acids 1-41

Transmembrane domains.
amino acids 53-73, 86-106, 122-142

N-glycosylation sites.
amino acids 2-5, 92-95

N-myristoylation site.
amino acids 88-93

Leucine zipper pattern.
amino acids 17-38

FIGURE 103

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVTICYAGNTYFSFIAWRSRTIQ

Transmembrane domains.
amino acids 19-39, 53-73, 86-106, 122-142

N-glycosylation sites.
amino acids 2-5, 92-95

N-myristoylation site.
amino acids 88-93

Leucine zipper pattern.
amino acids 17-38

FIGURE 104

MPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVALPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MHMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Transmembrane domains.
amino acids 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

N-glycosylation sites.
amino acids 72-75, 97-100, 113-116

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 345-348

Tyrosine kinase phosphorylation site.
amino acids 536-553

N-myristoylation sites.
amino acids 191-196, 254-259, 442-447, 466-471, 504-509

Amidation site.
amino acids 174-177, 229-232, 327-330, 401-404

Sugar (and other) transporter.
amino acids 23-529

FIGURE 105

MPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVSLPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MHMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Transmembrane domains.
amino acids 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

N-glycosylation sites.
amino acids 72-75, 97-100, 113-116

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 345-348

Tyrosine kinase phosphorylation site.
amino acids 536-544

N-myristoylation sites.
amino acids 191-196, 254-259, 442-447, 466-471, 504-509

Amidation sites.
amino acids 174-177, 229-232, 327-330, 401-404

Sugar (and other) transporter.
amino acids 23-529

FIGURE 106

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCYCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Transmembrane domains.
amino acids 96-116, 136-156, 178-198, 219-239, 356-376, 372-392, 406-426, 445-465, 488-508, 523-543, 549-569, 565-585, 592-612

N-glycosylation sites.
amino acids 36-39, 136-139, 295-298, 308-311, 313-316, 321-324, 335-338, 340-343, 495-498, 520-532, 667-670

N-myristoylation sites.
amino acids 23-28, 79-84, 126-131, 131-136, 146-151, 150-155, 187-192, 191-196, 393-398, 423-428, 460-465, 464-469, 519-524, 546-551, 634-639

4Fe-4S ferredoxins, iron-sulfur binding region signature.
amino acids 635-646

Insulin family signature.
amino acids 621-635

Heat-stable enterotoxins.

amino acids 617-633, 625-641, 613-629

Na+/Pi-cotransporter.

amino acids 118-549

FIGURE 107

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCCCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Transmembrane domains.
amino acids 96-116, 136-156, 178-198, 219-239, 356-376, 372-392, 406-426, 445-465, 488-508, 523-543, 547-567, 563-583, 592-612

N-glycosylation sites.
amino acids 36-39, 136-139, 295-298, 308-311, 313-316, 321-324, 335-338, 340-343, 495-498, 520-523, 667-670

N-myristoylation sites.
amino acids 23-28, 79-84, 126-131, 131-136, 146-151, 150-155, 187-192, 191-196, 393-398, 423-428, 460-465, 464-469, 519-524, 546-551, 634-639

4Fe-4S ferredoxins, iron-sulfur binding region signature.
amino acids 635-645

Insulin family signature.
amino acids 621-635

Na+/Pi-cotransporter.
amino acids 118-549

FIGURE 108

MWSGWWLWPLVAVCTADFFRDEAERIMRDSPVIDGHNDLPWQLLDMFNNRLQDERANLTTL
AGTHTNIPKLRAGFVGGQFWSVYTPCDTQNKDAVRRTLEQMDVVHRMCRMYPETFLYVTSS
AGIRQAFREGKVASLIGVEGGHSIDSSLGVLRALYQLGMRYLTLTHSCNTPWADNWLVDTG
DSEPQSQGLSPFGQRVVKELNRLGVLIDLAHVSVATMKATLQLSRAPVIFSHSSAYSVCAS
RRNVPDDVLRLVKQTDSLVMVNFYNNYISCTNKANLSQVADHLDHIKEVAGARAVGFGGDF
DGVPRVPEGLEDVSKYPDLIAELLRRNWTEAEVKGALADNLLRVFQAVEQASNLTQAPEEE
PIPLDQLGGSCRTHYGYSSGASSLHRHWGLLLASLAPLVLCLSLL

Signal sequence.
amino acids 1-16

Transmembrane domain.
amino acids 392-410

N-glycosylation sites.
amino acids 57-60, 279-282, 332-335, 358-361

Tyrosine kinase phosphorylation site.
amino acids 110-118

N-myristoylation sites.
amino acids 63-68, 78-83, 124-129, 139-144, 151-156, 375-380, 395-400

Leucine zipper pattern.
amino acids 390-411

Renal dipeptidase active site.
amino acids 140-162

Renal dipeptidase.
amino acids 1-411

FIGURE 109

MSSCNFTHATFVLIGIPGLEKAHFWVGFPLLSMYVVAMFGNCIVVFIVRTERSLHAPMYLF
LCMLAAIDLALSTSTMPKILALFWFDSREISFEACLTQMFFIHALSAIESTILLAMAFDRY
VAICHPLRHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLIKRLAFCHSNVLSHSYCVHQDV
MKLAYADTLPNVVYGLTAILLVMGVDVMFISLSYFLIIRTVLQLPSKSERAKAFGTCVSHI
GVVLAFYVPLIGLSVVHRFGNSLHPIVRVVMGDIYLLLPPVINPIIYGAKTKQIRTRVLAM
FKISCDKDLQAVGGK

Transmembrane domains.
amino acids 26-46, 57-77, 96-116, 144-164, 197-217, 238-258, 272-292

N-glycosylation sites.
amino acids 5-8, 136-139

N-myristoylation site.
amino acids 18-23, 238-243, 245-250

7 transmembrane receptor (rhodopsin family).
amino acids 40-131, 212-291

FIGURE 110A

MRILKRFLACIQLLCVCRLDWANGYYRQQRKLVEEIGWSYTGALNQKNWGKKYPTCNSPKQ
SPINIDEDLTQVNVNLKKLKFQGWDKTSLENTFIHNTGKTVEINLTNDYRVSGGVSEMVFK
ASKITFHWGKCNMSSDGSEHSLEGQKFPLEMQIYCFDADRFSSFEEAVKGKGKLRALSILF
EVGTEENLDFKAIIDGVESVSRFGKQAALDPFILLNLLPNSTDKYYIYNGSLTSPPCTDTV
DWIVFKDTVSISESQLAVFCEVLTMQQSGYVMLMDYLQNNFREQQYKFSRQVFSSYTGKEE
IHEAVCSSEPENVQADPENYTSLLVTWERPRVVYDTMIEKFAVLYQQLDGEDQTKHEFLTD
GYQDLGAILNNLLPNMSYVLQIVAICTNGLYGKYSDQLIVDMPTDNPELDLFPELIGTEEI
IKEEEEGKDIEEGAIVNPGRDSATNQIRKKEPQISTTTHYNRIGTKYNEAKTNRSPTRGSE
FSGKGDVPNTSLNSTSQPVTKLATEKDISLTSQTVTELPPHTVEGTSASLNDGSKTVLRSP
HMNLSGTAESLNTVSITEYEEESLLTSFKLDTGAEDSSGSSPATSAIPFISENISQGYIFS
SENPETITYDVLIPESARNASEDSTSSGSEESLKDPSMEGNVWFPSSTDITAQPDVGSGRE
SFLQTNYTEIRVDESEKTTKSFSAGPVMSQGPSVTDLEMPHYSTFAYFPTEVTPHAFTPSS
RQQDLVSTVNVVYSQTTQPVYNGETPLQPSYSSEVFPLVTPLLLDNQILNTTPAASSSDSA
LHATPVFPSVDVSFESILSSYDGAPLLPFSSASFSSELFRHLHTVSQILPQVTSATESDKV
PLHASLPVAGGDLLLEPSLAQYSDVLSTTHAASETLEFGSESGVLYKTLMFSQVEPPSSDA
MMHARSSGPEPSYALSDNEGSQHIFTVSYSSAIPVHDSVGVTYQGSLFSGPSHIPIPKSSL
ITPTASLLQPTHALSGDGEWSGASSDSEFLLPDTDGLTALNISSPVSVAEFTYTTSVFGDD
NKALSKSEIIYGNETELQIPSFNEMVYPSESTVMPNMYDNVNKLNASLQETSVSISSTKGM
FPGSLAHTTTKVFDHEISQVPENNFSVQPTHTVSQASGDTSLKPVLSANSEPASSDPASSE
MLSPSTQLLFYETSASFSTEVLLQPSFQASDVDTLLKTVLPAVPSDPILVETPKVDKISST
MLHLIVSNSASSENMLHSTSVPVFDVSPTSHMHSASLQGLTISYASEKYEPVLLKSESSHQ
VVPSLYSNDELFQTANLEINQAHPPKGRHVFATPVLSIDEPLNTLINKLIHSDEILTSTKS
SVTGKVFAGIPTVASDTFVSTDHSVPIGNGHVAITAVSPHRDGSVTSTKLLFPSKATSELS
HSAKSDAGLVGGGEDGDTDDDGDDDDDRDSDGLSIHKCMSCSSYRESQEKVMNDSDTHENS
LMDQNNPISYSLSENSEEDNRVTSVSSDSQTGMDRSPGKSPSANGLSQKHNDGKEENDIQT
GSALLPLSPESKAWAVLTSDEESGSGQGTSDSLNENETSTDFSFADTNEKDADGILAAGDS
EITPGFPQSPTSSVTSENSEVFHVSEAEASNSSHESRIGLAEGLESEKKAVIPLVIVSALT
FICLVVLVGILIYWRKCFQTAHFYLEDSTSPRVISTPPTPIFPISDDVGAIPIKHFPKHVA
DLHASSGFTEEFETLKEFYQEVQSCTVDLGITADSSNHPDNKHKNRYINIVAYDHSRVKLA
QLAEKDGKLTDYINANYVDGYNRPKAYIAAQGPLKSTAEDFWRMIWEHNVEVIVMITNLVE
KGRRKCDQYWPADGSEEYGNFLVTQKSVQVLAYYTVRNFTLRNTKIKKGSQKGRPSGRVVT
QYHYTQWPDMGVPEYSLPVLTFVRKAAYAKRHAVGPVVVHCSAGVGRTGTYIVLDSMLQQI
QHEGTVNIFGFLKHIRSQRNYLVQTEEQYVFIHDTLVEAILSKETEVLDSHIHAYVNALLI
PGPAGKTKLEKQFQLLSQSNIQQSDYSAALKQCNREKNRTSSIIPVERSRVGISSLSGEGT
DYINASYIMGYYQSNEFIITQHPLLHTIKDFWRMIWDHNAQLVVMIPDGQNMAEDEFVYWP

NKDEPINCESFKVTLMAEEHKCLSNEEKLIIQDFILEATQDDYVLEVRHFQCPKWPNPDSP
ISKTFELISVIKEEAANRDGPMIVHDEHGGVTAGTFCALTTLMHQLEKENSVDVYQVAKMI
NLMRPGVFADIEQYQFLYKVILSLVSTRQEENPSTSLDSNGAALPDGNIAESLESLV

Signal sequence.
amino acids 1-19
Transmembrane domain.
amino acids 1638-1658
N-glycosylation sites.
amino acids 105-108, 134-137, 223-226, 232-235, 324-327, 381-384, 497-500, 501-504, 552-555, 602-605, 629-632, 677-680, 1017-1020, 1050-1053, 1082-1085, 1122-1125, 1456-1459, 1561-1564, 1617-1620, 1868-1871, 2051-2054, 2078-2081
Glycosaminoglycan attachment sites.
amino acids 490-493, 991-994, 1548-1551, 1550-1553, 2070-2073
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 1877-1880
Tyrosine kinase phosphorylation site.
amino acids 148-156

FIGURE 110B

N-myristoylation sites.

amino acids 139-144, 186-191, 372-377, 471-476, 486-491, 533-538, 555-560, 582-587, 588-593, 638-643, 893-898, 960-965, 1097-1102, 1259-1264, 1385-1390, 1411-1416, 1415-1420, 1549-1554, 1551-1556, 1553-1558, 1579-1584, 1625-1630, 1879-1884, 1935-1940, 2123-2128, 2225-2230, 2226-2231, 2230-2235

Amidation sites.

amino acids 49-52, 1831-1834

Tyrosine specific protein phosphatases active site.

amino acids 1930-1942

Protein-tyrosine phosphatase.

amino acids 1749-1990, 2047-2280

Eukaryotic-type carbonic anhydrase.

amino acids 38-300

Fibronectin type III domain.

amino acids 312-401

MAM domain proteins.

amino acids 1757-1785, 1060-1088, 1813-1853, 2047-2079

FIGURE 111

MCARMAGRTRAAPRGPYGPWLCLLVALALDVVRVDCGQAPLDPVYLHVTAARPAQPTLWTA
KLDRFKGSRHHTTLITCHRAGLTEPDSSSPLELSEFLWVDFVVENSTGGGVAVTRPVTWQL
EYPGQAPEAEKDKMVWEILVSERDIRALIPLAKAEELVNTAPLTGVPQHVPVRLVTVDGGG
ALVEVTEHVGCESANTQVLQVSEACDAVFVAGKESRGARGVRVDFWWRRLRASLRLTVWAP
LLPLRIELTDTTLEQVRGWRVPGPAEGPAEPAAEASDEAERRARGCHLQYQRAGVRFLAPF
AAHPLDGGRRLTHLLGPDWLLDVSHLVAPHARVLDSRVASLEGGRVVVGREPGVTSIEVRS
PLSDSILGEQALAVTDDKVSVLELRVQPVMGISLTLSRGTAHPGEVTATCWAQSALPAPKQ
EVALSLWLSFSDHTVAPAELYDRRDLGLSVSAEEPGAILPAEEQGAQLGVVVSGAGAEGLP
LHVALHPPEPCRRGRHRVPLASGTAWLGLPPASTPAPALPSSPAWSPPATEATMGGKRQVA
GSVGGNTGVRGKFERAEEEARKEETEAREEEEEEEEEMVPAPQHVTELELGMYALLGVFCV
AIFIFLVNGVVFVLRYQRKEPPDSATDPTSPQPHNWVWLGTDQEELSRQLDRQSPGPPKGE
GSCPCESGGGGEAPTLAPGPPGGTTSSSSTLARKEAGGRRKRVEFVTFVPAPPAQSPEEPV
GAPAVQSILVAGEEDIRWVCEDMGLKDPEELRNYMERIRGSS

N-glycosylation site.
amino acids 106-109

Glycosaminoglycan attachment sites.
amino acids 480-483, 678-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 314-317

Tyrosine kinase phosphorylation site.
amino acids 758-766

N-myristoylation sites.
amino acids 109-114, 111-116, 193-198, 397-402, 454-459, 472-477, 476-481, 516-521, 550-555, 553-558, 672-677, 693-698, 694-694

Amidation sites.
amino acids 312-315, 543-546, 708-711

FIGURE 112

MATKTELSPTARESKNAQDMQVDETLIPRKGPSLCSARYGIALVLHFCNFTTIAQNVIMNI
TMVAMVNSTSPQSQLNDSSEVLPVDSFGGLSKAPKSLPAKSSILGGQFAIWEKWGPPQERS
RLCSIALSGMLLGCFTAILIGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPFSYPW
ISTSEKEYIISSLKQQVGSSKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYIS
SVYHVNIRDNGLLSALPFIVAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSLPSSALI
VSLPYLNSGYITATALLTLSCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIV
PTVSGFLLSQDPEFGWRNVFFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Transmembrane domains.

amino acids 124-144, 154-174, 255-275, 287-307, 319-339, 350-370, 385-405

N-glycosylation sites.

amino acids 49-52, 60-63, 68-71, 77-80

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-418

N-myristoylation sites.

amino acids 106-111, 131-136, 135-140, 143-148, 160-165, 163-168, 255-260, 269-274, 272-277, 297-302, 335-340, 353-358

FIGURE 113

MQVDETLIPRKVPSLCSARYGIALVLHFCNFTTIAQNVIMNITMVAMVNSTSPQSQLNDSS
EVLPVDSFGGLSKAPKSLPTKSSILGGQFAIWEKWGPPQERSRLCSIALSGMLLGCFTAIL
IGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPVSYPWISTSEKEYIISSLKQQVRS
SKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYISSVYHVNIRDNGLLSALPFI
VAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSLPSSALIVSLPYLNSGYITATALLTL
SCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIVPTVSGFLLSQDPEFGWRNV
FFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Transmembrane domains.
amino acids 22-42, 105-125, 137-157, 236-256, 268-288, 300-320, 331-351, 366-386

N-glycosylation sites.
amino acids 30-33, 41-44, 49-52, 58-61

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 396-399

N-myristoylation sites.
amino acids 87-92, 112-117, 116-121, 124-129, 141-146, 144-149, 236-241, 250-255, 253-258, 278-283, 316-321, 334-339

FIGURE 114

MAGAGPKRRALAAPAAEEKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPKFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Transmembrane domains.
amino acids 48-68, 83-103, 120-140, 137-157, 165-185, .
200-220, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

N-glycosylation sites.
amino acids 49-52, 340-343

Tyrosine kinase phosphorylation site.
amino acids 112-119

N-myristoylation sites.
amino acids 34-39, 61-66, 107-112, 114-119, 125-130, 337-342,
341-346, 353-358, 459-464, 488-493

Amino acid permease.
amino acids 46-481

FIGURE 115

MAGAGPKRRALAAPAAEEKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPNFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Transmembrane domains.
amino acids 48-68, 83-103, 120-140, 137-157, 165-185,
199-219, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

N-glycosylation sites.
amino acids 49-52, 230-233, 340-343

Tyrosine kinase phosphorylation site.
amino acids 112-119

N-myristoylation sites.
amino acids 34-39, 61-66, 107-112, 114-119, 125-130, 337-342,
341-346, 353-358, 459-464, 488-493

Amino acid permeases.
amino acids 46-481

FIGURE 116

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA

Signal sequence.
amino acids 1-33

Transmembrane domain.
amino acids 603-621

N-glycosylation sites.
amino acids 57-60, 388-391, 488-491, 515-518

Tyrosine kinase phosphorylation site.
amino acids 353-360, 471-477

N-myristoylation sites.
amino acids 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Amidation site.
amino acids 304-307

Leucine zipper pattern.
amino acids 101-122

FIGURE 117

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA

Signal sequence.
amino acids 1-33

Transmembrane domain.
amino acids 603-621

N-glycosylation sites.
amino acids 57-60, 388-391, 488-491, 515-518

Tyrosine kinase phosphorylation sites.
amino acids 353-360, 471-477

N-myristoylation sites.
amino acids 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Amidation site.
amino acids 304-307

Leucine zipper pattern.
amino acids 101-122

FIGURE 118

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAHIYLTMLLSIAIWVAWITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Transmembrane domains.
amino acids 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

N-glycosylation site.
amino acids 158-161

N-myristoylation sites.
amino acids 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Amidation site.
amino acids 124-127

7 transmembrane receptor.
amino acids 27-273

FIGURE 119

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERFMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMVQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYTT
KIITKVVEPKIKVIEGSLQPIIKTEGPTLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKK
YTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTKFIEG
GDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 507-510, 731-734

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 768-773

Fasciclin domain.
amino acids 94-232, 496-630

FIGURE 120

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEEVTKVTKFIEGGDGHLFEDEEIKRLLQGDTPVRKLQAN
KKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 507-510, 704-707

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 121

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEDTPVRKL
QANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 674-677

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 122

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 704-707

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 123

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTK
FIEGGDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 674-677

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 124

MVNYAWAGRSQRKLWWRSVAVLTCKSVVRPGYRGGLQARRSTLLKTCARARATAPGAMKMV
APWTRFYSNSCCLCCHVRTGTILLGVWYLIINAVVLLILLSALADPDQYNFSSSELGGDFE
FMDDANMCIAIAISLLMILICAMATYGAYKQRAAWIIPFFCYQIFDFALNMLVAITVLIYP
NSIQEYIRQLPPNFPYRDDVMSVNPTCLVLIILLFISIILTFKGYLISCVWNCYRYINGRN
SSDVLVYVTSNDTTVLLPPYDDATVNGAAKEPPPPYVSA

Transmembrane domains.
amino acids 83-103, 124-144, 162-182, 205-225

N-glycosylation sites.
amino acids 111-114, 244-227, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 39-42

Tyrosine kinase phosphorylation site.
amino acids 243-251

N-myristoylation site.
amino acids 34-39

FIGURE 125

MPGQGGLQARRSTLLKTCARARATAPGAMKMVAPWTRFYSNSCCLCCHVRTGTILLGVWYL
IINAVVLLILLSALADPDQYNFSSSELGGDFEFMDDANMCIAIAISLLMILICAMATYGAY
KQRAAWIIPFFCYQIFDFALNMLVAITVLIYPNSIQEYIRQLPPNFPYRDDVMSVNPTCLV
LIILLFISIILTFKGYLISCVWNCYRYINGRNSSDVLVYVTSNDTTVLLPPYDDATVNGAA
KEPPPPYVSA

Transmembrane domains.
amino acids 54-74, 95-115, 133-153, 176-196

N-glycosylation sites.
amino acids 82-85, 215-218, 226-229

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 10-13

Tyrosine kinase phosphorylation site.
amino acids 214-222

N-myristoylation site.
amino acids 5-10

FIGURE 126

MGRAGGGGPGRGPPPLLLFLGAALVLASGAVPAREAGSAVEAEELVKGSPAWEPPANDTRE
EAGPPAAGEDEASWTAPGGELAGPEEVLQESAAVTGTAWLEADSPGLGGVTAEAGSGDAQA
LPATLQAPHEVLGQSIMPPAIPEATEASGPPSPTPGDKLSPASELPKESPLEVWLNLGGST
PDPQVPELTYPFQGTLEPQPASDIIDIDYFEGLDGEGRGADLGSFPGSPGTSENHPDTEGE
TPSWSLLDLYDDFTPFDESDFYPTTSFYDDLDEEEEEEEDDKDAVGGGDLEDENELLVPTG
KPGLGPGTGQPTSRWHAVPPQHTLGSVPGSSIALRPRPGEPGRDLASSENGTECRSGFVRH
NGSCRSVCDLFPSYCHNGGQCYLVENIGAFCRCNTQDYIWHKGMRCESIITDFQVMCVAVG
SAALVLLLLFMMTVFFAKKLYLLKTENTKLRRTNKFRTPSELHNDNFSLSTIAEGSHPNDD
PSAPHKIQEVLKSCLKEEESFNIQNSMSPKLEGGKGDQADLDVNCLQNNLT

Signal sequence.
amino acids 1-30

Transmembrane domain.
amino acids 423-443

N-glycosylation sites.
amino acids 57-60, 355-358, 367-370, 473-476

N-myristoylation sites.
amino acids 6-11, 29-34, 79-84, 109-114, 116-121, 222-227, 226-231, 230-235, 233-238, 308-313, 314-319, 330-335, 334-339, 368-373, 482-487

FIGURE 127

MLPVYQEVKPNPLQDANICSRVFFWWLNPLFKIGHKRRLEEDDMYSVLPEDRSQHLGEELQ
GFWDKEVLRAENDAQKPSLTRAIIKCYWKSYLVLGIFTLIEESAKVIQPIFLGKIINYFEN
YDPMDSVALNTAYAYATVLTFCTLILAILHHLYFYHVQCAGMRLRVAMCHMIYRKALRLSN
MAMGKTTTGQIVNLLSNDVNKFDQVTVFLHFLWAGPLQAIAVTALLWMEIGISCLAGMAVL
IILLPLQSCFGKLFSSLRSKTATFTDARIRTMNEVITGIRIIKMYAWEKSFSNLITNLRKK
EISKILRSSCLRGMNLASFFSASKIIVFVTFTTYVLLGSVITASRVFVAVTLYGAVRLTVT
LFFPSAIERVSEAIVSIRRIQTFLLLDEISQRNRQLPSDGKKMVHVQDFTAFWDKASETPT
LQGLSFTVRPGELLAVVGPVGAGKSSLLSAVLGELAPSHGLVSVHGRIAYVSQQPWVFSGT
LRSNILFGKKYEKERYEKVIKACALKKDLQLLEDGDLTVIGDRGTTLSGGQKARVNLARAV
YQDADIYLLDDPLSAVDAEVSRHLFELCICQILHEKITILVTHQLQYLKAASQILILKDGK
MVQKGTYTEFLKSGIDFGSLLKKDNEESEQPPVPGTPTLRNRTFSESSVWSQQSSRPSLKD
GALESQDTENVPVTLSEENRSEGKVGFQAYKNYFRAGAHWIVFIFLILLNTAAQVAYVLQD
WWLSYWANKQSMLNVTVNGGGNVTEKLDLNWYLGIYSGLTVATVLFGIARSLLVFYVLVNS
SQTLHNKMFESILKAPVLFFDRNPIGRILNRFSKDIGHLDDLLPLTFLDFIQTLLQVVGVV
SVAVAVIPWIAIPLVPLGIIFIFLRRYFLETSRDVKRLESTTRSPVFSHLSSSLQGLWTIR
AYKAEERCQELFDAHQDLHSEAWFLFLTTSRWFAVRLDAICAMFVIIVAFGSLILAKTLDA
GQVGLALSYALTLMGMFQWCVRQSAEVENMMISVERVIEYTDLEKEAPWEYQKRPPPAWPH
EGVIIFDNVNFMYSPGGPLVLKHLTALIKSQEKVGIVGRTGAGKSSLISALFRLSEPEGKI
WIDKILTTEIGLHDLRKKMSIIPQEPVLFTGTMRKNLDPFKEHTDEELWNALQEVQLKETI
EDLPGKMDTELAESGSNFSVGQRQLVCLARAILRKNQILIIDEATANVDPRTDELIQKKIR
EKFAHCTVLTIAHRLNTIIDSDKIMVLDSGRLKEYDEPYVLLQNKESLFYKMVQQLGKAEA
AALTETAKQVYFKRNYPHIGHTDHMVTNTSNGQPSTLTIFETAL

Transmembrane domains:

amino acids 134-154, 233-253, 346-366, 708-728, 767-787, 834-854, 857-877, 948-968, 967-987

N-glycosylation sites.

amino acids 651-654, 690-693, 746-749, 754-757, 792-795, 1176-1179, 1309-1312

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 1115-1118

Tyrosine kinase phosphorylation sites.

amino acids 37-45, 497-504, 1253-1259

N-myristoylation sites.

amino acids 192-197, 318-323, 343-348, 430-435, 448-453, 487-492, 532-537, 537-542, 624-629, 672-677, 752-575,

766-771, 770-775, 779-784, 980-985, 1078-1083, 1174-1179, 1313-1318

Amidation sites.

amino acids 405-410, 495-500

ATP/GTP-binding site motif A (P-loop).

amino acids 445-452, 1075-1082

ABC transporter.

amino acids 438-609, 1068-1250

ABC transporter transmembrane region.

amino acids 92-365, 711-997

FIGURE 128

MSALRRKFGDDYQVVTTSSSGSGLQPQGPGQDPQQQLVPKKKRQRFVDKNGRCNVQHGNLG
SETSRYLSDLFTTLVDLKWRWNLFIFILTYTVAWLFMASMWWVIAYTRGDLNKAHVGNYTP
CVANVYNFPSAFLFFIETEATIGYGYRYITDKCPEGIILFLFQSILGSIVDAFLIGCMFIK
MSQPKKRAETLMFSEHAVISMRDGKLTLMFRVGNLRNSHMVSAQIRCKLLKSRQTPEGEFL
PLDQLELDVGFSTGADQLFLVSPLTICHVIDAKSPFYDLSQRSMQTEQFEIVVILEGIVET
TGMTCQARTSYTEDEVLWGHRFFPVISLEEGFFKVDYSQFHATFEVPTPPYSVKEQEEMLL
MSSPLIAPAITNSKERHNSVECLDGLDDITTKLPSKLQKITGREDFPKKLLRMSSTTSEKA
YSLGDLPMKLQRISSVPGNSEEKLVSKTTKMLSDPMSQSVADLPPKLQKMAGGAARMEGNL
PAKLRKMNSDRFT

Transmembrane domains.
amino acids 83-103, 120-140, 158-178

Glycosaminoglycan attachment site.
amino acids 20-23

N-myristoylation sites.
amino acids 58-63, 61-66, 216-221, 301-306, 486-491

Cell attachment sequence.
amino acids 109-114

Inward rectifier potassium channel.
amino acids 47-393

FIGURE 129

MPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTH
GITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPTTASPLLVLF
TINFTITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDG
AATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRDSLYVNGFTQRSSVP
TTSIPGTPTVDLGTSGTPVSKPGPSAASPLLVLFTLNFTITNLRYEENMQHPGSRKFNTTE
RVLQGLLRSLFKSTSVGPLYSGCRLTLLRPEKDGTATGVDAICTHHPDPKSPRLDREQLYW
ELSQLTHNITELGHYALDNDSLFVNGFTHRSSVSTTSTPGTPTVYLGASKTPASIFGPSAA
SHLLILFTLNFTITNLRYEENMWPGSRKFNTTERVLQGLLRPLFKNTSVGPLYSGSRLTLL
RPEKDGEATGVDAICTHRPDPTGPGLDREQLYLELSQLTHSITELGPYTLDRDSLYVNGFT
HRSSVPTTSTGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNITDNVMKHLLSPLFQRSSL
GARYTGCRVIALRSVKNGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQQTHGITRLGPYS
LDKDSLYLNGYNEPGLDEPPTTPKPATTFLPPLSEATTAMGYHLKTLTLNFTISNLQYSPD
MGKGSATFNSTEGVLQHLLRPLFQKSSMGPFYLGCQLISLRPEKDGAATGVDTTCTYHPDP
VGPGLDIQQLYWELSQLTHGVTQLGFYVLDRDSLFINGYAPQNLSIRGEYQINFHIVNWNL
SNPDPTSSEYITLLRDIQDKVTTLYKGSQLHDTFRFCLVTNLTMDSVLVTVKALFSSNLDP
SLVEQVFLDKTLNASFHWLGSTYQLVDIHVTEMESSVYQPTSSSSTQHFYPNFTITNLPYS
QDKAQPGTTNYQRNKRNIEDALNQLFRNSSIKSYFSDCQVSTFRSVPNRHHTGVDSLCNFS
PLARRVDRVAIYEEFLRMTRNGTQLQNFTLDRSSVLVDGYSPNRNEPLTGNSDLPFWAVIF
IGLAGLLGLITCLICGVLVTTRRRKKEGEYNVQQQCPGYYQSHLDLEDLQ

Transmembrane domains.

amino acids 107-127, 418-438, 1093-1113

N-glycosylation sites.

amino acids 6-9, 125-128, 146-149, 162-165, 281-284, 302-305, 374-377, 385-388, 437-440, 457-460, 473-476, 570-573, 591-594, 721-724, 741-744, 836-839, 853-856, 895-898, 928-931, 967-970, 1004-1007, 1058-1061, 1064-1067

Glycosaminoglycan attachment site.

amino acids 644-647

Tyrosine kinase phosphorylation sites.

amino acids 358-365, 1121-1128

N-myristoylation sites.

amino acids 15-20, 171-176, 260-265, 310-315, 327-332, 339-344, 482-487, 583-588, 611-616, 778-783, 782-787, 1106-1111, 1114-1119

SEA domain.

amino acids 112-245, 270-391, 661-684

FIGURE 130

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSVAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Signal sequence.
amino acids 1-29

Transmembrane domain.
amino acids 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

Ferric reductase like.
amino acids 46-563

FIGURE 131

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Transmembrane domains.
amino acids 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

Ferric reductase like transmembrane.
amino acids 46-563

FIGURE 132

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSWKWILAPVILYIC
ERILRFYRSQQKVVITKVVMHPSKVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLT
SAPEEDFFSIHIRAAGDWTENLIRAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAG
IGVTPFASILKSIWYKFQCADHNLKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEF
STIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSSLDPRKVQFYFNKENFO

Transmembrane domains.
amino acids 6-26, 44-64

Tyrosine kinase phosphorylation site.
amino acids 146-152

N-myristoylation sites.
amino acids 181-186, 257-262

Ferric reductase like transmembrane.
amino acids 1-293

FIGURE 133

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLA
KSLRKCCHRYSSLDPRKVQFYFNKENF

Transmembrane domains.
amino acids 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 478-483

Ferric reductase like transmembrane.
amino acids 46-514

FIGURE 134

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Transmembrane domains.
amino acids 212-232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-649, 665-685

N-glycosylation sites.
amino acids 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

N-myristoylation sites.
amino acids 49-54, 352-357, 586-591, 659-664

Leucine zipper pattern.
amino acids 487-508

LDL-receptor class A.
amino acids 40-59

G-protein coupled receptor.
amino acids 159-189, 413-432, 422-452, 436-466, 478-518, 528-547, 625-644, 673-689

Leucine rich repeats.
amino acids 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

7 transmembrane receptor (rhodopsin family proteins).
amino acids 485-681

Low-density lipoprotein receptor domain.
amino acids 25-64

FIGURE 135

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSLQFDKYFASYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Transmembrane domains.
amino acids 212-232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-649, 665-685

N-glycosylation sites.
amino acids 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

N-myristoylation sites.
amino acids 49-54, 352-357, 586-591, 659-664

Leucine zipper pattern.
amino acids 487-508

Leucine rich repeats.
amino acids 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

7 transmembrane receptor (rhodopsin family).
amino acids 59-259

Low-density lipoprotein receptor domain.

amino acids 25-64

FIGURE 136

MGRLASRPLLLALLSLALCRGRVVRVPTATLVRVVGTELVIPCNVSDYDGPSEQNFDWSFS
SLGSSFVELASTWEVGFPAQLYQERLQRGEILLRRTANDAVELHIKNVQPSDQGHYKCSTP
STDATVQGNYEDTVQVKVLADSLHVGPSARPPPSLSLREGEPFELRCTAASASPLHTHLAL
LWEVHRGPARRSVLALTHEGRFHPGLGYEQRYHSGDVRLDTVGSDAYRLSVSRALSADQGS
YRCIVSEWIAEQGNWQEIQEKAVEVATVVIQPTVLRAAVPKNVSVAEGKELDLTCNITTDR
ADDVRPEVTWSFSRMPDSTLPGSRVLARLDRDSLVHSSPHVALSHVDARSYHLLVRDVSKE
NSGYYYCHVSLWAPGHNRSWHKVAEAVSSPAGVGVTWLEPDYQVYLNASKVPGFADDPTEL
ACRVVDTKSGEANVRFTVSWYYRMNRRSDNVVTSELLAVMDGDWTLKYGERSKQRAQDGDF
IFSKEHTDTFNFRIQRTTEEDRGNYYCVVSAWTKQRNNSWVKSKDVFSKPVNIFWALEDSV
LVVKARQPKPFFAAGNTFEMTCKVSSKNIKSPRYSVLIMAEKPVGDLSSPNETKYIISLDQ
DSVVKLENWTDASRVDGVVLEKVQEDEFRYRMYQTQVSDAGLYRCMVTAWSPVRGSLWREA
ATSLSNPIEIDFQTSGPIFNASVHSDTPSVIRGDLIKLFCIITVEGAALDPDDMAFDVSWF
AVHSFGLDKAPVLLSSLDRKGIVTTSRRDWKSDLSLERVSVLEFLLQVHGSEDQDFGNYYC
SVTPWVKSPTGSWQKEAEIHSKPVFITVKMDVLNAFKYPLLIGIGLSTVIGLLSCLIGYCS
SHWCCKKEVQETRRERRRLMSMEMD

Signal sequence.
amino acids 1-21

Transmembrane domain.
amino acids 833-853

N-glycosylation sites.
amino acids 44-47, 286-289, 300-303, 383-386, 413-416, 525-528, 600-603, 618-621, 691-694

Tyrosine kinase phosphorylation site.
amino acids 632-640

N-myristoylation sites.
amino acids 243-248, 398-403, 511-516, 651-656, 753-758, 789-794, 836-841, 844-849

Cell attachment sequence.
amino acids 703-705

Immunoglobulin domain.

amino acids 36-121, 162-249, 292-375, 422-517, 564-657, 704-795

FIGURE 137

MSRSRHLGKIRKRLEDVKSQWVRPARADFSDNESARLATDALLDGGSEAYWRVLSQEGEVD
FLSSVEAQYIQAQAREPPCPPDTLGGAEAGPKGLDSSSLQSGTYFPVASEGSEPALLHSWA
SAEKPYLKEKSSATVYFQTVKHNNIRDLVRRCITRTSQVLVILMDVFTDVEIFCDILEAAN
KRGVFVCVLLDQGGVKLFQEMCDKVQISDSHLKNISIRSVEGEIYCAKSGRKFAGQIREKF
IISDWRFVLSGSYSFTWLCGHVHRNILSKFTGQAVELFDEEFRHLYASSKPVMGLKSPRLV
APVPPGAAPANGRLSSSSGSASDRTSSNPFSGRSAGSHPGTRSVSASSGPCSPAAPHPPPP
PRFQPHQGPWGAPSPQAHLSPRPHDGPPAAVYSNLGAYRPTRLQLEQLGLVPRLTPTWRPF
LQASPHF

Transmembrane domain.
amino acids 154-175

N-glycosylation sites.
amino acids 32-35, 217-220

Tyrosine kinase phosphorylation site.
amino acids 221-228

N-myristoylation sites.
amino acids 45-50, 86-91, 94-99, 112-117, 186-191, 311-316, 341-346

Amidation site.
amino acids 232-235

FIGURE 138

MALPQPPPPAGARDPVTPEHWASGPPSQAQPLLRQEAKEEEEGEETGVQGAWGTGTAEQRR
RGWGEAAESAAAEEGQAEVGGAAAAGSGSPAGGAGGGLGSWRPLLAWLQRRQPQCCPCAAP
LSRSAAHCCHGGTKMAALAYNLGKREINHYFSVRSAKVLALVAVLLLAACHLASRRYRGND
SCEYLLSSGRFLGEKVWQPHSCMMHKYKISEAKNCLVDKHIAFIGDSRIRQLFYSFVKIIN
PQFKEEGNKHENIPFEDKTASVKVDFLWHPEVNGSMKQCIKVWTEDSIAKPHVIVAGAATW
SIKIHNGSSEALSQYKMNITSIAPLLEKLAKTSDVYWVLQDPVYEDLLSENRKMITNEKID
AYNEAAVSILNSSTRNSKSNVKMFSVSKLIAQETIMESLDGLHLPESSRETTAMILMNVYC
NKILKPVDGSCCQPRPPVTLIQKLAACFFTLSIIGYLIFYIIHRNAHRKNKPCTDLESGEE
KKNIINTPVSSLEILLQSFCKLGLIMAYFYMCDRANLFMKENKFYTHSSFFIPIIYILVLG
VFYNENTKETKVLNREQTDEWKGWMQLVILIYHISGASTFLPVYMHIRVLVAAYLFQTGYG
HFSYFWIKGDFGIYRVCQVLFRLNFLVVVLCIVMDRPYQFYYFVPLVTVWFMVIYVTLALW
PQIIQKKANGNCFWHFGLLLKLGFLLLFICFLAYSQGAFEKIFSLWPLSKCFELKGNVYEW
WFRWRLDRYVVFHGMLFAFIYLALQKRQILSEGKGEPLFSNKISNFLLFISVVSFLTYSIW
ASSCKNKAECNELHPSVSVVQILAFILIRNIPGYARSVYSSFFAWFGKISLELFICQYHIW
LAADTRGILVLIPGNPMLNIIVSTFIFVCVAHEISQITNDLAQIIIPKDNSSLLKRLACIA
AFFCGLLILSSIQDKSKH

Transmembrane domains.

amino acids 153-173, 447-467, 495-515, 532-552, 577-597, 624-644, 651-671, 677-697, 693-713, 733-753, 771-791, 808-828, 866-886, 904-924

N-glycosylation sites.

amino acids 182-185, 277-280, 311-314, 323-326, 377-380, 904-907

Tyrosine kinase phosphorylation site.

amino acids 180-187

N-myristoylation sites.

amino acids 47-52, 50-55, 53-58, 81-86, 82-87, 89-94, 93-98, 94-99, 96-101, 97-102, 134-139, 181-186, 312-317, 511-516, 549-554, 746-751

Amidation site.

amino acids 144-147

FIGURE 139

MYHCHSGSKPTEKGANEYAYAKWKLCSASAICFIFMIAEVVGGHIAGSLAVVTDAAHLLID
LTSFLLSLFSLWLSSKPPSKRLTFGWHRAEILGALLSILCIWVVTGVLVYLACERLLYPDY
QIQATVMIIVSSCAVAANIVLTVVLHQRCLGHNHKEVQANASVRAAFVHALGDLFQSISVL
ISALIIYFKPEYKIADPICTFIFSILVLASTITILKDFSILLMEGVPKSLNYSGVKELILA
VDGVLSVHSLHIWSLTMNQVILSAHVATAASRDSQVVRREIAKALSKSFTMHSLTIQMESP
VDQDPDCLFCEDPCD

Transmembrane domains.
amino acids 53-73, 91-111, 127-147, 166-186, 202-222

N-glycosylation sites.
amino acids 162-165, 234-237

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 81-84

Tyrosine kinase phosphorylation site.
amino acids 13-20

N-myristoylation sites.
amino acids 7-12, 42-47, 94-99, 228-233

Cation efflux family.
amino acids 78-306

FIGURE 140

MVLSVPVIALGATLGTATSILALCGVTCLCRHMHPKKGLLPRDQDPDLEKAKPSLLGSAQQ
FNVKKSTEPVQPRALLKFPDIYGPRPAVTAPEVINYADYSLRSTEEPTAPASPQPPNDSRL
KRQVTEELFILPQNGVVEDVCVMETWNPEKAASWNQAPKLHYCLDYDCQKAELFVTRLEAV
TSNHDGGCDCYVQGSVANRTGSVEAQTALKKRQLHTTWEEGLVLPLAEEELPTATLTLTLR
TCDRFSRHSVAGELRLGLDGTSVPLGAAQWGELKTSAKEPSAGAGEVLLSISYLPAANRLL
VVLIKAKNLHSNQSKELLGKDVSVKVTLKHQARKLKKKQTKRAKHKINPVWNEMIMFELPD
DLLQASSVELEVLGQDDSGQSCALGHCSLGLHTSGSERSHWEEMLKNPRRQIAMWHQLHL

Signal sequence.
amino acids 1-24

Transmembrane domain.
amino acids 4-24

N-glycosylation sites.
amino acids 118-121, 201-204, 317-320

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 65-68, 342-345

N-myristoylation sites.
amino acids 11-16, 15-20, 189-194, 197-202, 204-209, 261-266, 380-385, 385-390, 396-401, 401-406

C2 domain.
amino acids 304-394

FIGURE 141

MQKIMHISVLLSPVLWGLIFGVSSNSIQIGGLFPRGADQEYSAFRVGMVQFSTSEFRLTPH
IDNLEVANSFAVTNAFCSQFSRGVYAIFGFYDKKSVNTITSFCGTLHVSFITPSFPTDGTH
PFVIQMRPDLKGALLSLIEYYQWDKFAYLYDSDRGLSTLQAVLDSAAEKKWQVTAINVGNI
NNDKKDEMYRSLFQDLELKKERRVILDCERDKVNDIVDQVITIGKHVKGYHYIIANLEFTD
GDLLKIQFGGANVSGFQIVDYDDSLVSKFIERWSTLEEKEYPGAHTTTIKYTSALTYDAVQ
VMTEAFRNLRKQRIEISRRGNAGDCLANPAVPWGQGVEIERALKQVQVEGLSGNIKFDQNG
KRINYTINIMELKTNGPRKIGYWSEVDKMVVTLTELPSGNDTSGLENKTVVVTTILESPYV
MMKKNHEMLEGNERYEGYCVDLAAEIAKHCGFKYKLTIVGDGKYGARDADTKIWNGMVGEL
VYGKADIAIAPLTITLVREEVIDFSKPFMSLGISIMIKKPQKSKPGVFSFLDPLAYEIWMC
IVFAYIGVSVVLFLVSRFSPYEWHTEEFEDGRETQSSESTNEFGIFNSLWFSLGAFMRQGC
DISPRSLSGRIVGGVWWFFTLIIISSYTANLAAFLTVERMVSPIESAEDLSKQTEIAYGTL
DSGSTKEFFRRSKIAVFDKMWTYMRSAEPSVFVRTTAEGVARVRKSKGKYAYLLESTMNEY
IEQRKPCDTMKVGGNLDSKGYGIATPKGSSLGTPVNLAVLKLSEQGVLDKLKNKWWYDKGE
CGAKDSGSKEKTSALSLSNVAGVFYILVGGLGLAMLVALIEFCYKSRAEAKRMKVAKNAQN
INPSSSQNSQNFATYKEGYNVYGIESVKI

Signal sequence.
amino acids 1-21

Transmembrane domains.
amino acids 4-24, 543-563, 624-644, 811-831

N-glycosylation sites.
amino acids 256-259, 370-373, 406-409, 413-416

Tyrosine kinase phosphorylation sites.
amino acids 35-41, 662-667

N-myristoylation sites.
amino acids 17-22, 21-26, 134-140, 181-186, 254-259, 287-292, 355-360, 405-410, 472-477, 593-598, 669-674, 746-751, 760-765, 764-769, 795-800, 823-828

Amidation site.
amino acids 365-368

Ligand-gated ion channel.
amino acids 543-834

Receptor family ligand binding region.
amino acids 56-398

FIGURE 142

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAHIYLTMLLSIAIWVAWITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Transmembrane domains.
amino acids 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

N-glycosylation site.
amino acids 158-161

N-myristoylation sites.
amino acids 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Amidation site.
amino acids 124-127

7 transmembrane receptor.
amino acids 27-273

FIGURE 143

MLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWIGDGVDTPQAGYFGL
FHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIACIICFTLFFFCNTAT
VYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGACSVRWAYILAIIGI
LDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.

amino acids 21-41, 93-113, 127-147, 173-193

N-myristoylation sites.

amino acids 23-28, 67-72, 89-94, 182-187

FIGURE 144

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.

amino acids 1-15, 34-54, 107-127, 141-161, 187-207

N-myristoylation sites.

amino acids 3-8, 37-42, 81-86, 103-108, 196-201

FIGURE 145

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.

amino acids 1-15, 34-54, 107-127, 141-161, 187-207

N-myristoylation sites.

amino acids 3-8, 37-42, 81-86, 103-108, 196-201

FIGURE 146

RPLRLPCRGPRAAGTVSSESLGHPAPGQVTNRSPQFPMELFQSSPRPPLQPVCVCPGLELC
DPPHRLAFSMGAGLFSVVGTLLLPGLAALVQDWRLLQGLGALMSGLLLLFWGFPALFPESP
CWLLATGQVARARKILWRFAEASGVGPGDSSLEENSLATELTMLSARSPQPRYHSPLGLLR
TRVTWRNGLILGFSSLVGGGIRASFRRSLAPQVPTFYLPYFLEAGLEAAALVFLLLTADCC
GRRPVLLLGTMVTGLASLLLLAGAQYLPGWTVLFLSVLGLLASRAVSALSSLFAAEVFPTV
IRGAGLGLVLGAGFLGQAAGPLDTLHGRQGFFLQQVVFASLAVLALLCVLLLPESRSRGLP
QSLQDADRLRRSPLLRGRPRQDHLPLLPPSNSYWAGHTPEQH

Transmembrane domains.
amino acids 67-87, 102-122, 220-240, 249-269, 265-285, 301-321, 338-358

Glycosaminoglycan attachment site.
amino acids 145-148

N-myristoylation sites.
amino acids 14-19, 27-32, 57-62, 101-106, 146-151, 180-185, 191-196, 202-207, 203-208, 228-233, 253-258, 283-288, 308-313, 312-317, 364-369

Amidation site.
amino acids 244-247

Leucine zipper pattern.
amino acids 96-117

Sugar transport proteins.
amino acids 117-126, 241-253, 276-321

FIGURE 147

MPFEQCHLSWKNVRRLYNAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQ
RLKNLLTDVLLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINES
ENYEALLECVIILNGILYALPESERKLQSSIQDLCVTWWEKGLPAKEDTGKTAFVMLLRRS
LETKTGADVCRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNW
NINFIKMIHGTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHL
PRRSPVHSKVREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEA
FPIRDPNLHAIEMDSXIQKQFEELYSLLEDPYPMVRSTGILGVCKITSKYWEMMPPTILID
LLKKVTGELAFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAF
VDMLLKIKAVRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERC
VTLVQMNHAAARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREK
ENVTVLDKTLSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVF
KDDRCKIPLFMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHIL
ELVDNWLPTEHAQAKSNTASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKK
LNHLLKALETSKADLESLLQTPGGKPRGFSEAAAPRAFGLHCRLSIHLQHKFCSEGKVYLS
MLEDTGFWLESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLL
QRSLGIMQTVKGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIAR
SFRKQPEEGLRLLYSVQRPLHEFITAVQSRHTDTPVHRGVLSTLIAGPVVEISHQLRKVSD
VEELTPPEHLSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVHIIL
VINAGKHKSSKVREVAATVHRKLKTFMEITLEEDSIERFLYESSSRTLGELLNS

Transmembrane domain.
amino acids 1081-1101
N-glycosylation sites.
amino acids 120-123, 612-615
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 280-283, 430-433, 1033-1036
Tyrosine kinase phosphorylation sites.
amino acids 99-106, 314-321, 329-337, 385-391, 844-852
N-myristoylation sites.
amino acids 137-142, 189-194, 254-259, 262-267, 698-703, 920-925, 1015-1020
Amidation site.
amino acids 233-236

ATP/GTP-binding site motif A (P-loop).

amino acids 167-174

Leucine zipper pattern.

amino acids 128-149, 135-156, 787-808

Mitochondrial energy transfer proteins signature.

amino acids 982-990

FIGURE 148

MEKRETFVQAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQRLKNLLTDV
LLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINESENYEALLEC
VIILNGILYALPESERKLQSSIQDLCVTWWEKGLPAKEDTGKTAFVMLLRRSLETKTGADV
CRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNWNINFIKMIH
GTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHLPRRSPVHSK
VREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEAFPIRDPNLH
AIEMDSEIQKQFEELYSLLEDPYPMVRSTGILGVCKITSKYWEMMPPTILIDLLKKVTGEL
AFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAFVDMLLKIKA
VRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERCVTLVQMNHA
AARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREKENVTVLDKT
LSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVFKDDRCKIPL
FMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHILELVDNWLPT
EHAQAKSNTASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKKLNHLLKALE
TSKADLESLLQTPGGKPRGFSEAAAPRAFGLHCRLSIHLQHKFCSEGKVYLSMLEDTGFWL
ESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLLQRSLGIMQT
VKGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIARSFRKQPEEG
LRLLYSVQRPLHEFITAVQSRHTDTPVHRGVLSTLIAGPVVEISHQLRKVSDVEELTPPEH
LSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVHIILVINAGKHKS
SKVREVAATVHRKLKTFMEITLEEDSIERFLYESSSRTLGELLNS

Transmembrane domain.
amino acids 1072-1092
N-glycosylation sites.
amino acids 111-114, 603-606
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 3-6, 271-274, 421-424, 1024-1027
Tyrosine kinase phosphorylation site.
amino acids 91-97, 305-312, 320-328, 376-382, 835-843
N-myristoylation sites.
amino acids 128-133, 180-185, 245-250, 253-258, 689-694, 911-916, 1006-1011
Amidation site.
amino acids 224-227

ATP/GTP-binding site motif A (P-loop).

amino acids 158-165

Leucine zipper pattern.

amino acids 119-140, 126-147, 778-799

Mitochondrial energy transfer proteins signature.

amino acids 973-981

FIGURE 149

MNSTLFSQVENHSVHSNFSEKNAQLLAFENDDCHLPLAMIFTLALAYGAVIILGVSGNLAL
IIIILKQKEMRNVTNILIVNLSFSDLLVAIMCLPFTFVYTLMDHWVFGEAMCKLNPFVQCV
SITVSIFSLVLIAVERHQLIINPRGWRPNNRHAYVGIAVIWVLAVASSLPFLIYQVMTDEP
FQNVTLDAYKDKYVCFDQFPSDSHRLSYTTLLLVLQYFGPLCFIFICYFKIYIRLKRRNNM
MDKMRDNKYRSSETKRINIMLLSIVVAFAVCWLPLTIFNTVFDWNHQIIATCNHNLLFLLC
HLTAMISTCVNPIFYGFLNKNFQRDLQFFFNFCDFRSRDDDYETIAMSTMHTDVSKTSLKQ
ASPVAFKKINNNDDNEKI

Transmembrane domains.
amino acids 38-58, 78-98, 117-137, 155-175, 208-228, 260-280, 296-316

N-glycosylation sites.
amino acids 2-5, 11-14, 17-20, 73-76, 81-84, 186-189

Tyrosine kinase phosphorylation sites.
amino acids 247-253, 341-347

N-myristoylation site.
amino acids 54-59

7 transmembrane receptor (rhodopsin family).
amino acids 57-320

FIGURE 150

MVRCDRGLQMLLTTAGAFAAFSLMAIAIGTDYWLYSSAHICNGTNLTMDDGPPPRRARGDL
THSGLWRVCCIEGIYKGHCFRINHFPEDNDYDHDSSEYLLRIVRASSVFPILSTILLLLGG
LCIGAGRIYSRKNNIVLSAGILFVAAGLSNIIGIIVYISSNTGDPSDKRDEDKKNHYNYGW
SFYFGALSFIVAETVGVLAVNIYIEKNKELRFKTKREFLKASSSSPYARMPSYRYRRRRSR
SSSRSTEASPSRDVSPMGLKITGAIPMGELSMYTLSREPLKVTTAASYSPDQEASFLQVHD
FFQQDLKEGFHVSMLNRRTTPV

Signal sequence.
amino acids 1-19

Transmembrane domains.
amino acids 103-123, 140-160, 182-202

N-glycosylation sites.
amino acids 42-45, 45-48

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 240-243, 322-325

N-myristoylation sites.
amino acids 16-21, 43-48, 74-79, 122-127, 262-267

Cell attachment sequence.
amino acids 58-60

PMP-22/EMP/MP20/Claudin family.
amino acids 6-202

FIGURE 151

MARGPGLAPPPLRLPLLLLVLAAVTGHTAAQDNCTCPTNKMTVCSPDGPGGRCQCRALGSG
MAVDCSTLTSKCLLLKARMSAPKNARTLVRPSEHALVDNDGLYDPDCDPEGRFKARQCNQT
SVCWCVNSVGVRRTDKGDLSLRCDELVRTHHILIDLRHRPTAGAFNHSDLDAELRRLFRER
YRLHPKFVAAVHYEQPTIQIELRQNTSQKAAGEVDIGDAAYYFERDIKGESLFQGRGGLDL
RVRGEPLQVERTLIYYLDEIPPKFSMKRLTAGLIAVIVVVVVALVAGMAVLVITNRRKSGK
YKKVEIKELGELRKEPSL

Signal sequence.
amino acids 1-26

Transmembrane domain.
amino acids 274-294

N-glycosylation sites.
amino acids 33-36, 120-123, 168-171, 208-211

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 271-274, 300-303

N-myristoylation sites.
amino acids 59-64, 132-137

Thyroglobulin type-1 repeat.
amino acids 73-145

FIGURE 152

MGKCSGRCTLVAFCCLQLVAALERQIFDFLGYQWAPILANFLHIMAVILGIFGTVQYRSRY
LILYAAWLVLWVGWNAFIICFYLEVGQLSQDRDFIMTFNTSLHRSWWMENGPGCLVTPVLN
SRLALEDHHVISVTGCLLDYPYIEALSSALQIFLALFGFVFACYVSKVFLEEEDSFDFIGG
FDSYGYQAPQKTSHLQLQPLYTSG

Signal sequence.
amino acids 1-21

Transmembrane domains.
amino acids 33-53, 62-82, 147-167

N-glycosylation site.
amino acids 100-103

Tyrosine kinase phosphorylation site.
amino acids 24-32

N-myristoylation sites.
amino acids 50-55, 182-187

FIGURE 153

MPAPRAREQPRVPGERQPLLPRGARGPRRWRRAAGAAVLLVEMLERAAFFGVTANLVLYLN
STNFNWTGEQATRAALVFLGASYLLAPVGGWLADVYLGRYRAVALSLLLYLAASGLLPATA
FPDGRSSFCGEMPASPLGPACPSAGCPRSSPSPYCAPVLYAGLLLLGLAASSVRSNLTSFG
ADQVMDLGRDATRRFFNWFYWSINLGAVLSLLVVAFIQQNISFLLGYSIPVGCVGLAFFIF
LFATPVFITKPPMGSQVSSMLKLALQNCCPQLWQRHSARDRQCARVLADERSPQPGASPQE
DIANFQVLVKILPVMVTLVPYWMVYFQMQSTYVLQGLHLHIPNIFPANPANISVALRAQGS
SYTIPEAWLLLANVVVVLILVPLKDRLIDPLLLRCKLLPSALQKMALGMFFGFTSVIVAGV
LEMERLHYIHHNETVSQQIGEVLYNAAPLSIWWQIPQYLLIGISEIFASIPGLEFAYSEAP
RSMQGAIMGIFFCLSGVGSLLGSSLVALLSLPGGWLHCPKDFGNINNCRMDLYFFLLAGIQ
AVTALLFVWIAGRYERASQGPASHSRFSRDRG

Transmembrane domains.
amino acids 35-55, 73-93, 98-118, 157-177, 199-219, 250-270, 341-361, 369-389, 408-428, 457-477, 499-519, 537-557

N-glycosylation sites.
amino acids 61-64, 66-69, 178-181, 223-226, 356-359, 439-442

Glycosaminoglycan attachment site.
amino acids 503-506

N-myristoylation sites.
amino acids 51-56, 90-95, 116-121, 147-152, 169-174, 209-214, 258-263, 365-370, 414-419, 479-484, 493-498, 497-502, 506-511, 531-536

POT family.
amino acids 101-503

FIGURE 154

MWVNPEEVLLANALWITERANPYFILQRRKGHAGDGGGGGGLAGLLVGTLDVVLDSSARVA
PYRILYQTPDSLVYWTIACGGSRKEITEHWEWLEQNLLQTLSIFENENDITTFVRGKIQGI
IAEYNKINDVKEDDDTEKFKEAIVKFHRLFGMPEEEKLVNYYSCSYWKGKVPRQGWMYLSI
NHLCFYSFLMGREAKLVIRWVDITQLEKNATLLLPDVIKVSTRSSEHFFSVFLNINETFKL
MEQLANIAMRQLLDNEGFEQDRSLPKLKRKSPKKVSALKRDLDARAKSERYRALFRLPKDE
KLDGHTDCTLWTPFNKMHILGQMFVSTNYICFTSKEENLCSLIIPLREVTIVEKADSSSVL
PSPLSISTRNRMTFLFANLKDRDFLVQRISDFLQQTTSKIYSDKEFAGSYNSSDDEVYSRP
SSLVSSSPQRSTSSDADGERQFNLNGNSVPTATQTLMTMYRRRSPEEFNPKLAKEFLKEQA
WKIHFAEYGQGICMYRTEKTRELVLKGIPESMRGELWLLLSGAINEKATHPGYYEDLVEKS
MGKYNLATEEIERDLHRSLPEHPAFQNEMGIAALRRVLTAYAFRNPNIGYCQAMNIVTSVL
LLYAKEEEAFWLLVALCERMLPDYYNTRVVGALVDQGVFEELARDYVPQLYDCMQDLGVIS
TISLSWFLTLFLSVMPFESAVVVVDCFFYEGIKVIFQLALAVLDANVDKLLNCKDDGEAMT
VLGRYLDSVTNKDSTLPPIPHLHSLLSDDVEPYPEVDIFRLIRTSYEKFGTIRADLIEQMR
FKQRLKVIQTLEDTTKRNVVRTIVTETSFTIDELEELYALFKAEHLTSCYWGGSSNALDRH
DPSLPYLEQYRIDFEQFKGMFALLFPWACGTHSDVLASRLFQLLDENGDSLINFREFVSGL
SAACHGDLTEKLKLLYKMHVLPEPSSDQDEPDSAFEATQYFFEDITPECTHVVGLDSRSKQ
GADDGFVTVSLKPDKGKRANSQENRNYLRLWTPENKSKSKNAKDLPKLNQGQFIELCKTMY
NMFSEDPNEQELYHATAAVTSLLLEIGEVGKLFVAQPAKEGGSGGSGPSCHQGIPGVLFPK
KGPGQPYVVESVEPLPASLAPDSEEHSLGGQMEDIKLEDSSPRDNGACSSMLISDDDTKDD
SSMSSYSVLSAGSHEEDKLHCEDIGEDTVLVRSGQGTAALPRSTSLDRDWAITFEQFLASL
LTEPALVKYFDKPVCMMARITSAKNIRMMGKPLTSASDYEISAMSG

Transmembrane domains.

amino acids 35-55, 669-689

N-glycosylation sites.

amino acids 212-215, 239-242, 417-420, 1011-1014

Glycosaminoglycan attachment site.

amino acids 1192-1195

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 272-275, 277-280, 468-471

N-myristoylation sites.

amino acids 36-41, 37-42, 39-44, 40-45, 44-49, 414-419, 497-502, 515-520, 668-673, 782-787, 845-850, 846-851, 914-919, 977-982, 1078-1083, 1079-1084, 1082-1087, 1144-1149,

1193-1198

Amidation site.

amino acids 991-994

TBC domain.

amino acids 512-724

S-100/ICaBP type calcium.

amino acids 891-911

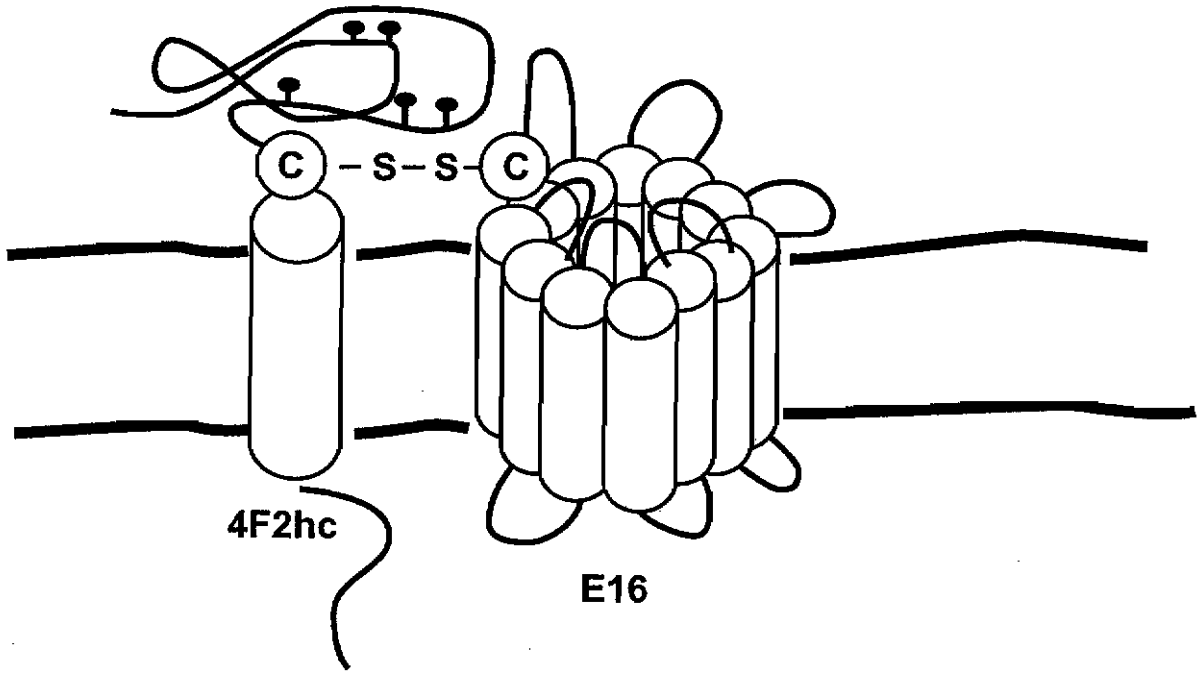


FIG. 155

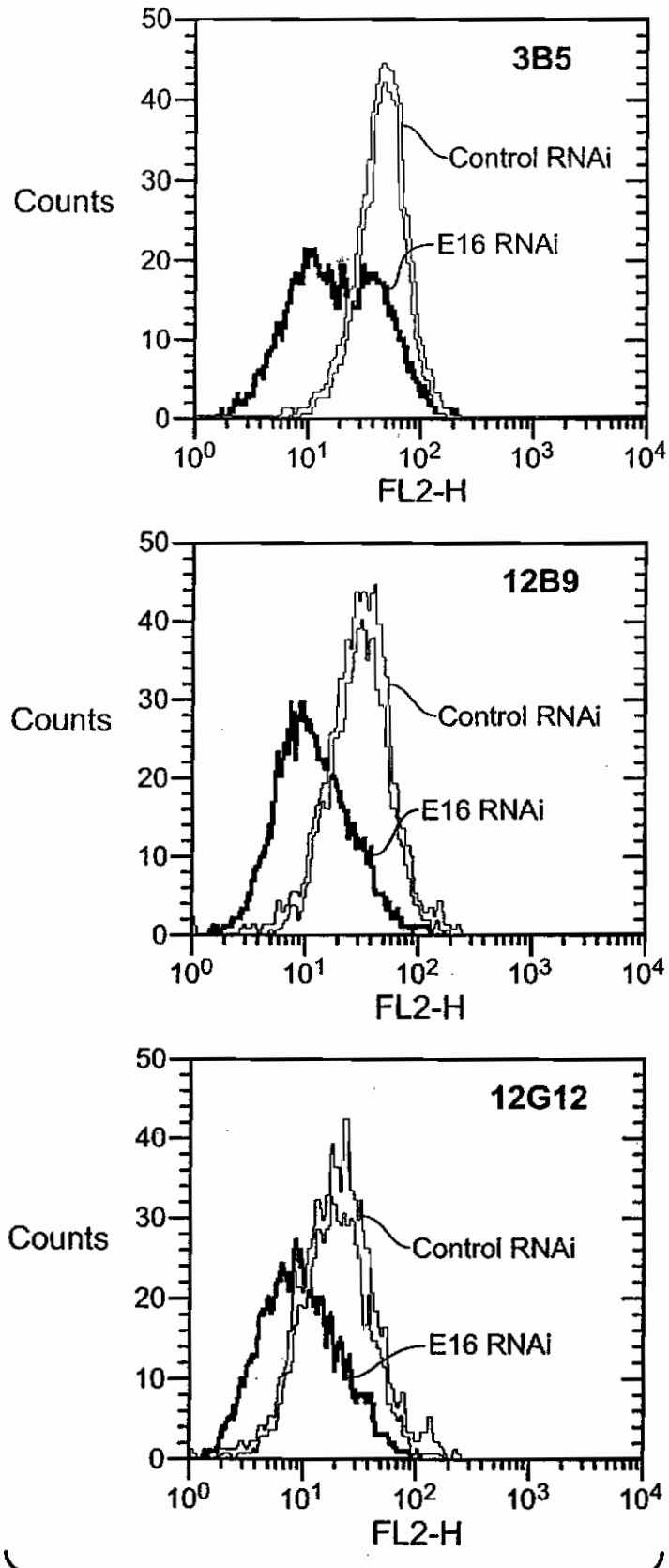
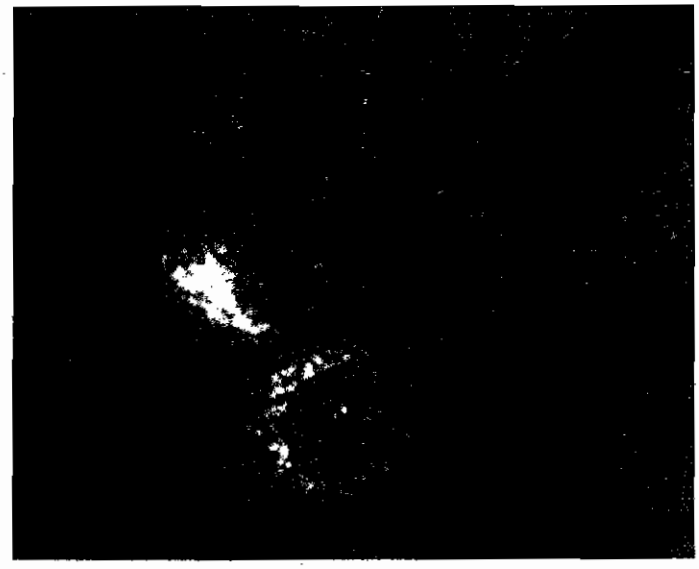

3B5
Counts
Control RNAi
E16 RNAi
FL2-H
12B9
Counts
Control RNAi
E16 RNAi
FL2-H
12G12
Counts
Control RNAi
E16 RNAi
FL2-H

FIG. 156



PC3 Cells Treated with E16 mAb (3B5) for Overnight.

FIG. 157

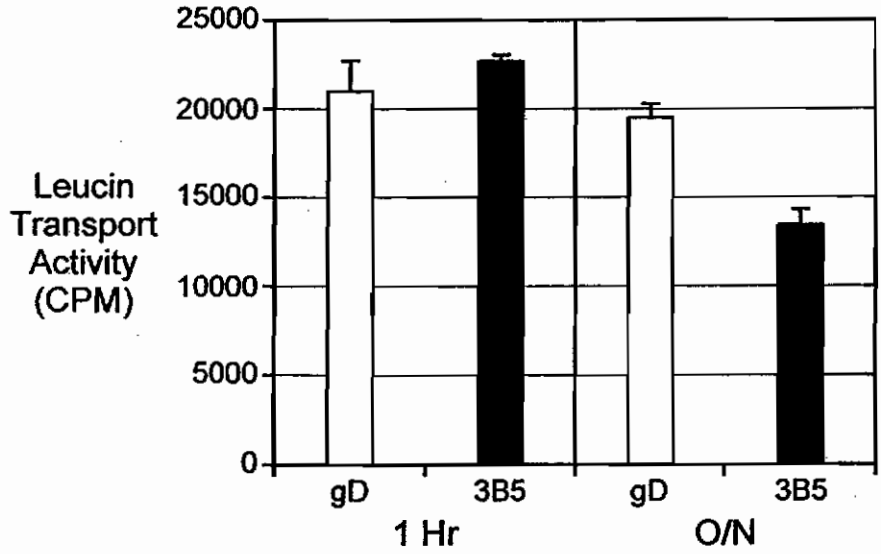


FIG. 158

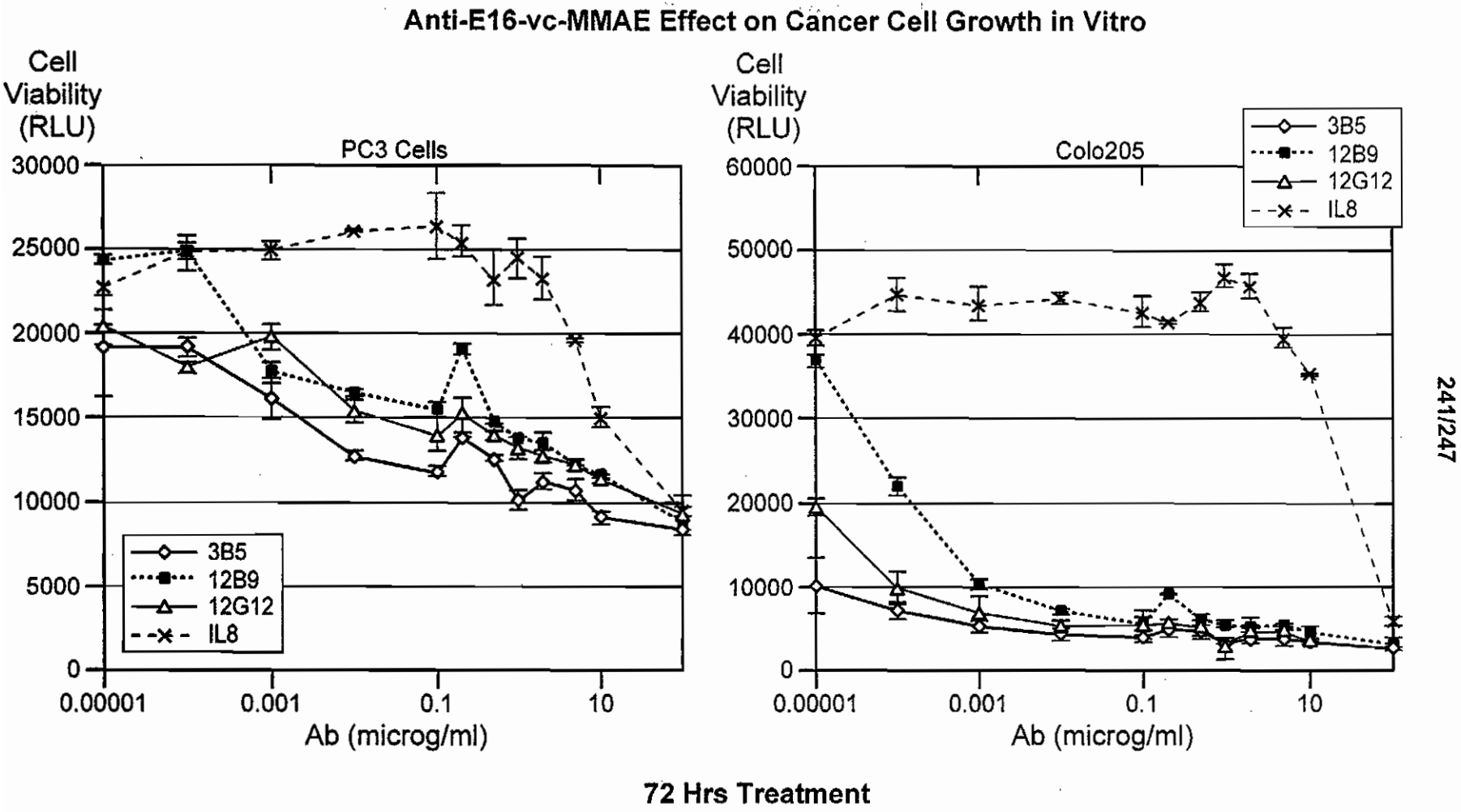

Anti-E16-vc-MMAE Effect on Cancer Cell Growth in Vitro
Cell Viability (RLU)
PC3 Cells
30000
25000
20000
15000
10000
5000
0
0.00001
0.001
0.1
10
Ab (microg/ml)
3B5
12B9
12G12
IL8
Cell Viability (RLU)
Colo205
60000
50000
40000
30000
20000
10000
0
0.00001
0.001
0.1
10
Ab (microg/ml)
3B5
12B9
12G12
IL8
72 Hrs Treatment

FIG. 159

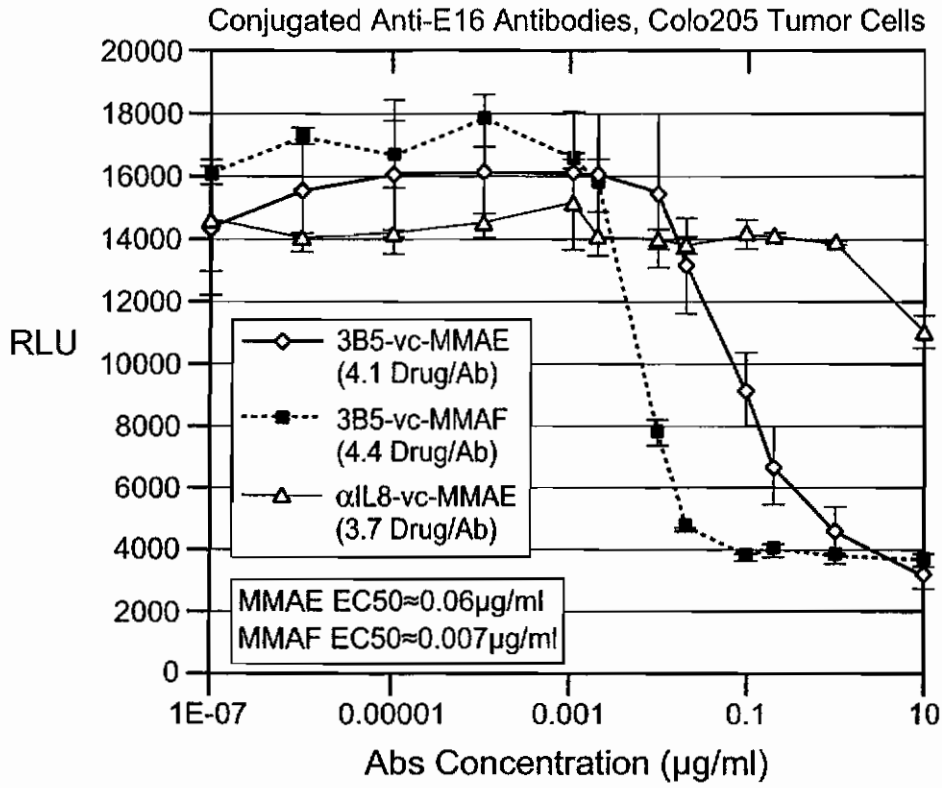
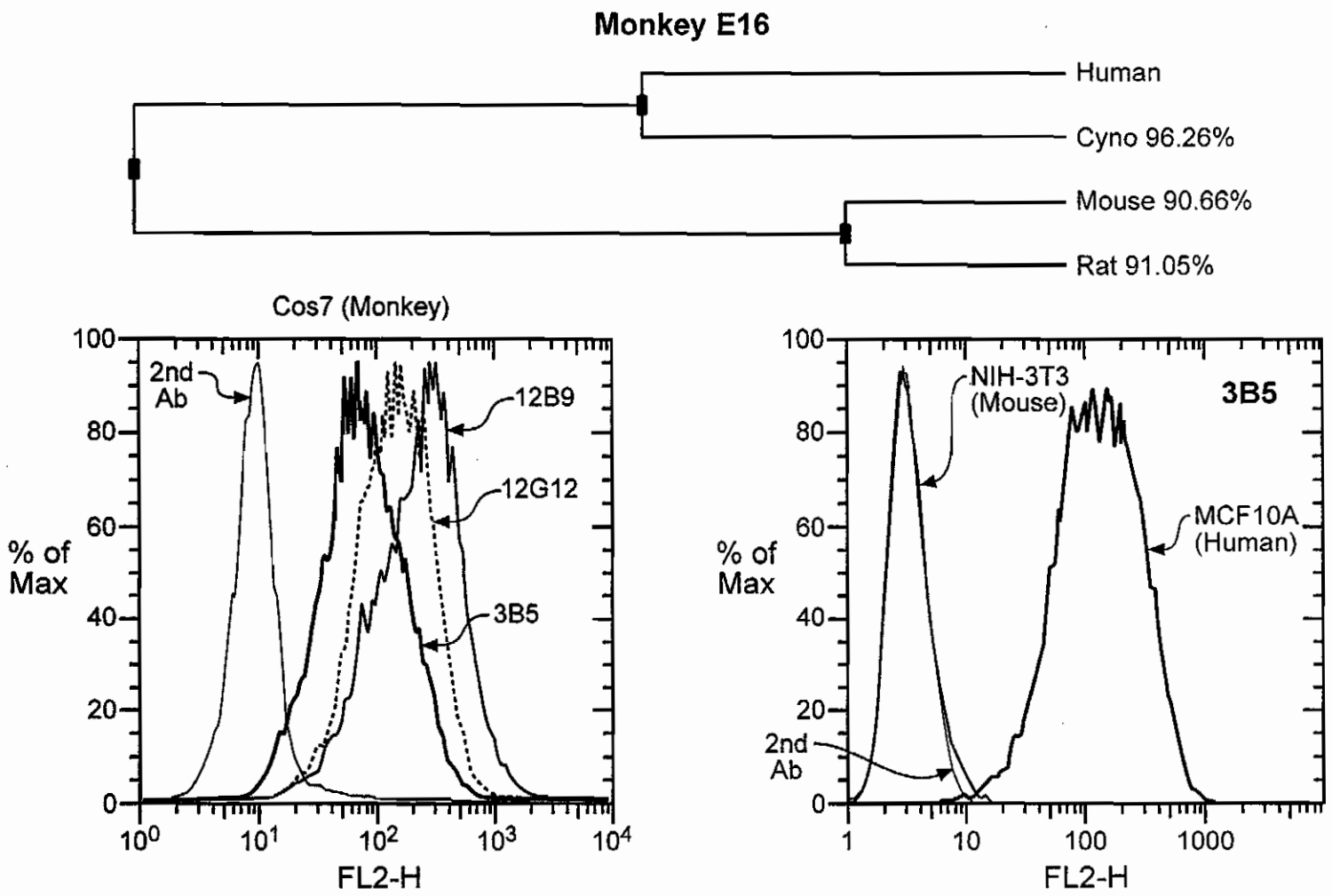

Conjugated Anti-E16 Antibodies, Colo205 Tumor Cells
RLU
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
1E-07
0.00001
0.001
0.1
10
Abs Concentration (µg/ml)
3B5-vc-MMAE
(4.1 Drug/Ab)
3B5-vc-MMAF
(4.4 Drug/Ab)
αIL8-vc-MMAE
(3.7 Drug/Ab)
MMAE EC50≈0.06µg/ml
MMAF EC50≈0.007µg/ml

FIG. 160



Anti-E16 antibody do not react with mouse E16, but cross react with monkey E16

FIG. 161

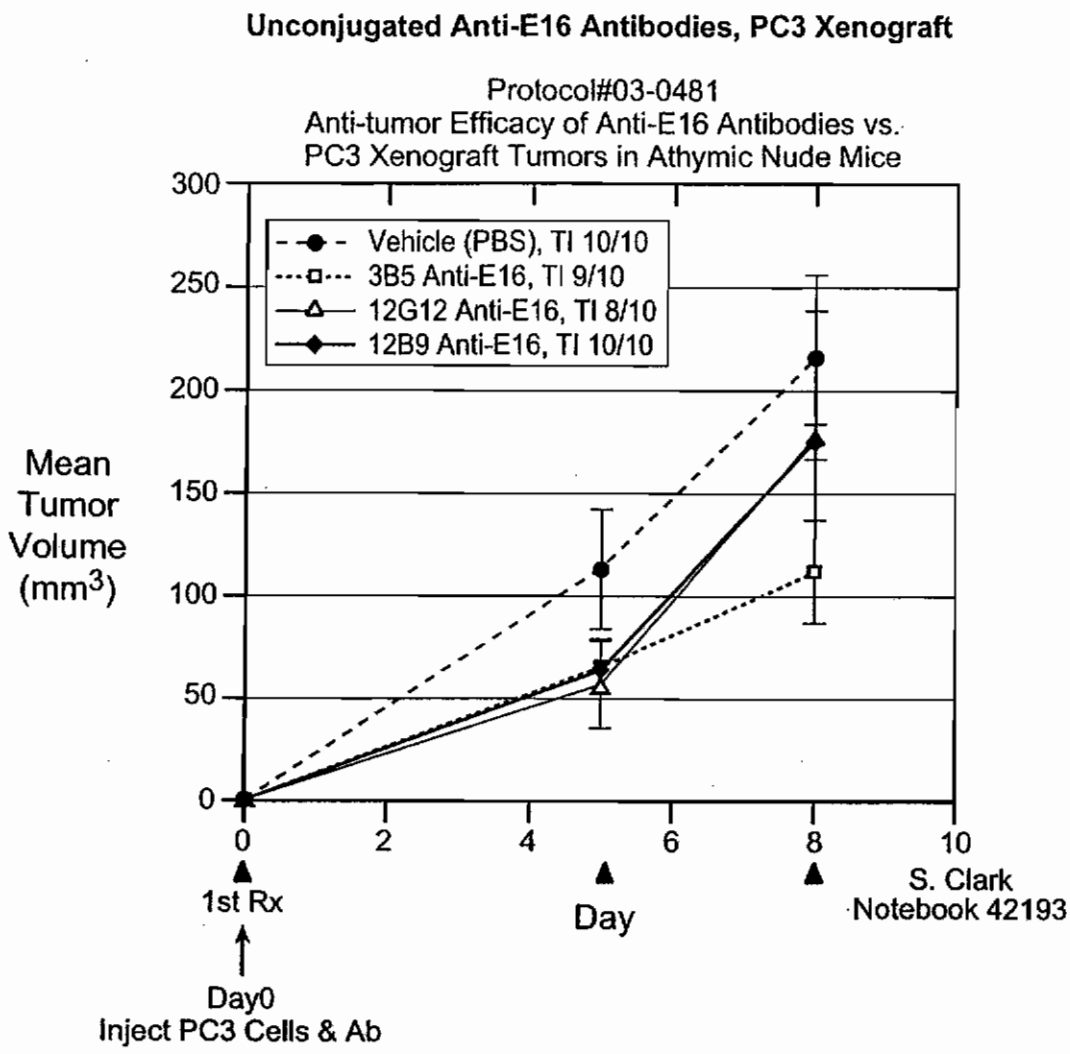

Unconjugated Anti-E16 Antibodies, PC3 Xenograft
Protocol#03-0481
Anti-tumor Efficacy of Anti-E16 Antibodies vs.
PC3 Xenograft Tumors in Athymic Nude Mice
Vehicle (PBS), TI 10/10
3B5 Anti-E16, TI 9/10
12G12 Anti-E16, TI 8/10
12B9 Anti-E16, TI 10/10
Mean Tumor Volume (mm3)
300
250
200
150
100
50
0
0
2
4
6
8
10
1st Rx
Day
S. Clark Notebook 42193
Day0
Inject PC3 Cells & Ab

FIG. 162

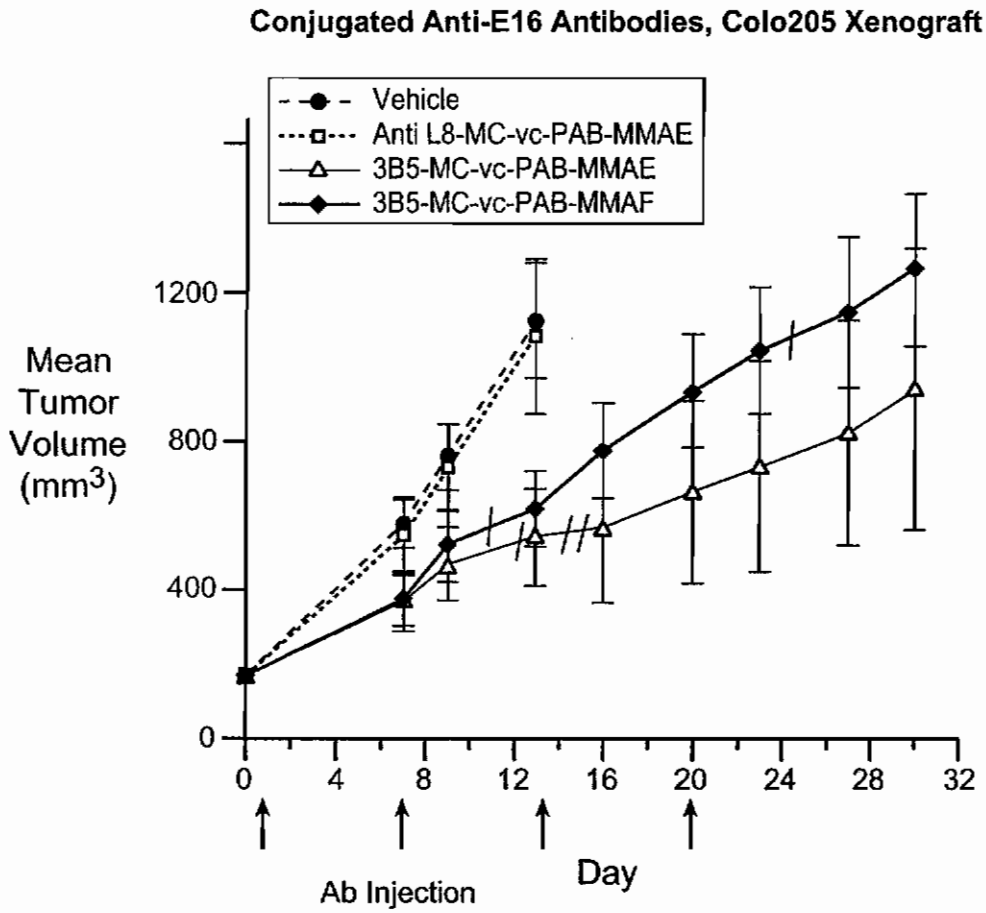

Conjugated Anti-E16 Antibodies, Colo205 Xenograft
Vehicle
Anti L8-MC-vc-PAB-MMAE
3B5-MC-vc-PAB-MMAE
3B5-MC-vc-PAB-MMAF
Mean Tumor Volume (mm3)
0
400
800
1200
0
4
8
12
16
20
24
28
32
Day
Ab Injection

FIG. 163

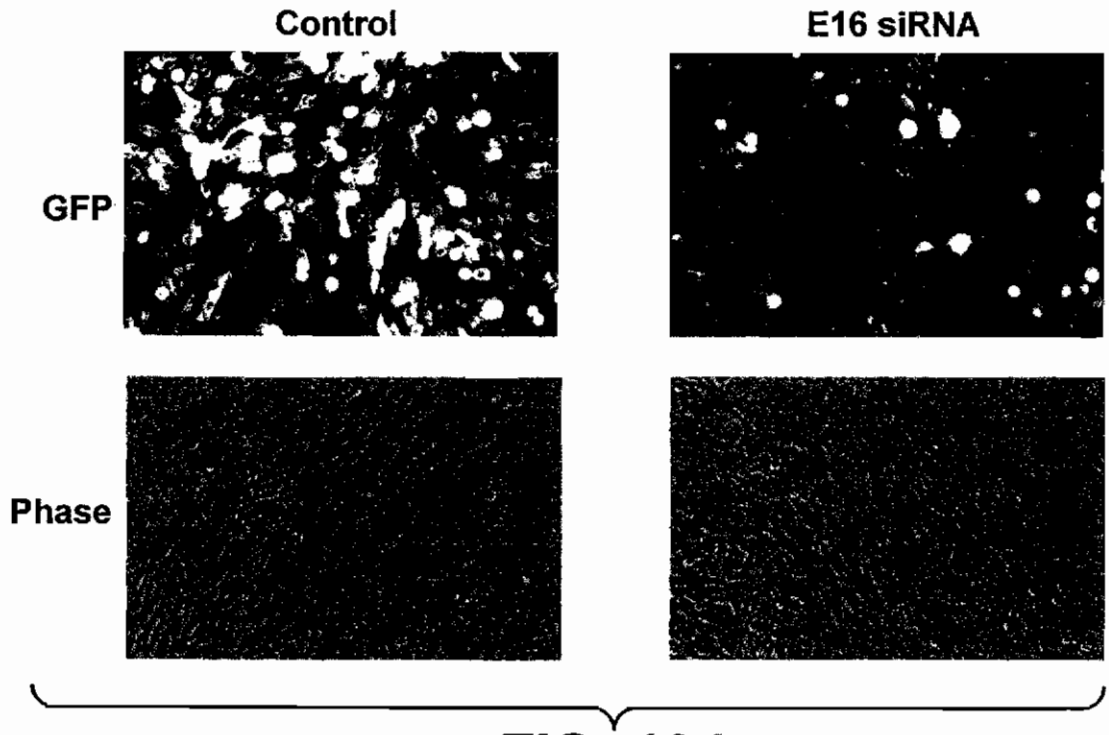


FIG. 164

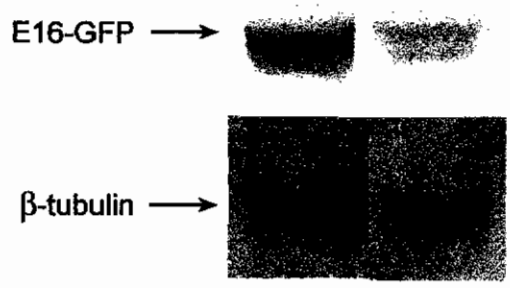


FIG. 165

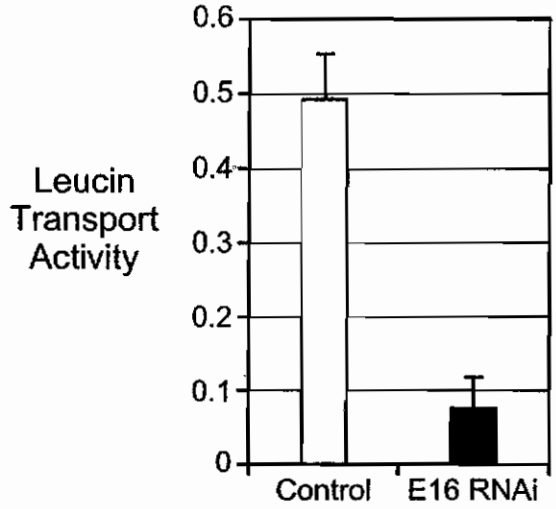


FIG. 166

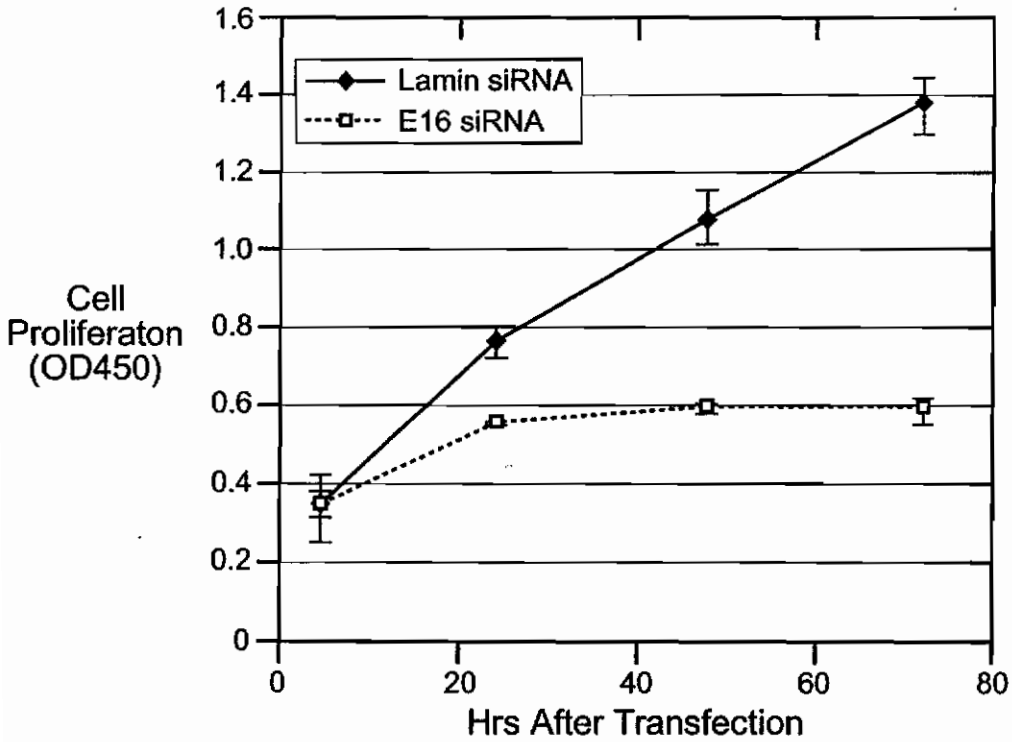


FIG. 167