

**APPENDIX 4**

**USHA CODING SEQUENCE FROM *ESCHERICHIA COLI***

This is the complete *E. coli ushA* coding sequence as reported by Burns and Beacham (1986). The numbers at the end of each row indicate the position of the first nucleotide in each row, with the first nucleotide in the initiation codon being position 1. The locations of primers used in this study are indicated, with any variation from the *ushA* sequence being noted by the shaded bases. The arrows indicate 5'→3' direction of the primer.

```

-93          CGGACGCAGA CTTGTCTATG ATGTCGCGTC ATACTATTTT
          Lyle01
          TCAACACGTT GAAATCAGGT CAGGGAGAGA AGT
-33
1
ATGAAATTAT TGCAGCGGGG CGTGGCGTTA GCGCTGTTAA CCACATTTAC ACTGGCGAGT
61
GAAACTGCTC TGGCGTATGA GCAGGATAAA ACCTACAAAA TTACAGTTCT GCATACCAAT
121 D41N H43N
GATCATCATG GGCATTTTTG GCGCAATGAA TATGGCGAAT ATGGTCTGGC GCGCAGAAAA
181
ACGCTGGTGG ATGGTATCCG CAAAGAGGTT GCGGCTGAAG GCGGTAGCGT GCTGCTACTT
241 D84N
TCCGGTGGCG ACATTAACAC TGGCGTGCCC GAGTCTGACT TACAGGATGC CGAACCTGAT
301 H117N E118Q
TTTCGCGGTA TGAATCTGGT GGGCTATGAC GCGATGGCGA TCGGTAATCA TGAATTTGAT
361
AATCCGCTCA CCGTATTACG CCAGCAGGAA AAGTGGGCCA AGTCCCCTT GCTTCCGCG
421
AATATCTACC AGAAAAGTAC TGGCGAGCGC CTGTTTAAAC CGTGGGCGCT GTTTAAGCGT
481
CAGGATCTGA AAATTGCCGT TATTGGGCTG ACAACCGATG ACACAGCAAA AATTGGTAAC
541
CCGAATACT TCACTGATAT CGAATTCGT AAGCCCGCCG ATGAAGCGAA GCTGGTGATT
601 H217N
CAGGAGCTGC AACAGACAGA AAAGCCAGAC ATTATTATCG CGGCGACCCA TATGGGGCAT
661
TACGATAATG GTGAGCACGG CTCTAACGCA CCGGGCGATG TGGAGATGGC ACGCGCGCTG

```

721 CCTGCCGGAT CGCTGGCGAT GATCGTCCGT GGTCACTCGC AAGATACGGT CTGCATGGCG  
■ H252N →

781 GCAGAAAACA AAAACAGGT CGATTACGTG CCGGGTACGC CATGCAAACC AGATCAACAA

841 AACGGCATCT GGATTGTGCA GGCGCATGAG TGGGGCAAAT ACGTGGGACG GGCTGATTTT

901 GAGTTTCGTA ATGGCGAAAT GAAAATGGTT AACTACCAGC TGATTCCGGT GAACCTGAAG

961 AAGAAAGTGA CCTGGGAAGA CGGGAAAAGC GAGCGCGTGC TTTACACTCC TGAAATCGCT

1021 GAAAACCAGC AAATGATCTC GCTGTTATCA CCGTTCCAGA ACAAAGGCAA AGCGCAGCTG

1081 GAAGTGAAAA TAGGCGAAAC CAATGGTCGT CTGGAAGGCG ATCGTGACAA AGTGCGTTTT

1141 GTACAGACCA ATATGGGGCG GTTGATTCTG GCAGCCCAA TGGATCGCAC TGGTGCCGAC

1201 TTTGCGGTGA TGAGCGGAGG CGGAATTCGT GATTCTATCG AAGCAGGCGA TATCAGCTAT

1261 AAAAACGTGC TGAAAGTGCA GCCATTCGGC AATGTGGTGG TGTATGCCGA CATGACCGGT

1321 AAAGAGGTGA TTGATTACCT GACCGCCGTC GCGCAGATGA AGCCAGATTC AGGTGCCTAC

1381 CCGCAATTTG CCAACGTTAG CTTTGTGGCG AAAGACGGCA AACTGAACGA CCTTAAATC

1441 AAAGGCGAAC CGGTCGATCC GCGGAAAAC TACCGTATGG CGACAT'AAA CTTCATGCC

1501 ACCGGCGGTG ATGGATATCC GCGCCTTGAT AACAAACCGG GCTATGTGAA TACCGGCTTT

1561 ATTGATGCCG AAGTGCTGAA AGCGTATATC CAGAAAAGCT CGCCGCTGGA TGTGAGTGT

1621 TATGAACCGA AAGGTGAGGT GAGCTGGCAG TAA

1654 TCCGAAAGTG CCGGATGTTT GCATCCGGCA CAATGCTTAA TCGCGGCGGG CGATATCAGC

1714 AAATTTGGCA TCG

***E. COLI* 5'-NUCLEOTIDASE SEQUENCE**

This is the primary structure of the *E. coli* 5'-nucleotidase, with changes to the amino acid sequence induced by oligonucleotide-directed mutagenesis indicated above the wild type sequence. The arrow between residues 25 and 26 indicates the end of the leader peptide, with the mature protein starting at residue 26.

```

1
  MKLLQRGVAL ALLTTFTLAS ETALAYEQDK TYKITVLHTN  █ █ DHHGHFWRNE YGEYGLAAQK
                                     ↓
61
  TLVDGIRKEV AAEGGSVLLL SGGDINTGVP ESDLQDAEPD FRGMNLVGYD AMAIGNHEFD
                                     █
121
  NPLTVLRQQE KWAKFPLLSA NIYQKSTGER LFKPWALFKR QDLKIAVIGL TTDDTAKIGN
181
  PEYFTDIEFR KPADEAKLVI QELQQTEKPD IIIAATHMGH  █ YDNGEHGSNA PGDVEMARAL
241
  PAGSLAMIVG  █ GHSQDTVCMA AENKKQVDYV PGTPCKPDQQ NGIWIVQAHE WGKYVGRADF
301
  EFRNGEMKMN YQLIPVNLKK KVTWEDGKSE RVLYTPEIAE NQQMISLLSP FQNKGKAQLE
361
  VKIGETNGRL EGDRDKVRFV QTNMGRILILA AQMRTGADFA VMSGGGIRDS IEAGDISYKN
421
  VLKVQPFQNV VVYADMTGKE VIDYLTAVAQ MKPDSGAYPQ FANVSFVAKD GKLNDLKIKG
481
  EPVDPAKTYR MATLNFNATG GDGYPRLDNK PGYVNTGFID AEVLKAYIQK SSPLDVSVYE
541
  PKGEVSWQ

```