Volume 16 Number 19 1988 Nucleic Acids Research

Nucleotide sequence of the *aceB* gene encoding malate synthase A in *Escherichia coli*

Carolyn Byrne, Harold W.Stokes[1] and Kevin A.Ward

Division of Animal Production, CSIRO, PO Box 239, Blacktown, NSW 2148 and [1]School of Biological Sciences, Macquarie University, North Ryde, NSW 2113, Australia

Submitted September 5, 1988

Accession no. X12431

The DNA sequence of Escherichia coli aceB encoding malate synthase A has been determined(1). Shown is the coding region of aceB flanked by the C-terminal region of an upstream open reading frame and the N-terminal coding region of the adjacent aceA gene (2). The transcription start site is shown by an asterisk and the -10 and -35 site of the promoter are marked and underlined (3). Sequences similar to ribosome binding sites for aceB and aceA (2) are underlined. Regions of dyad symmetry in or near the promoter are marked by arrows and numbered. 1 is probably the terminator for the upstream gene, 2, 3, 4 and 5 are possible regulatory sites. DNA from the aceB gene hybridizes to a 10 kb RNA transcript.

```
AspLeuGluIleLeuAlaGluThrGluGluGlyAspAlaTyrLeuPheAlaSerLysAspLysArgIleAlaPheValThrGlyHisProGluTyrAspAlaGlnThrLeuAlaGlnGlu
GATCTGGAAATTCTGGCAGAGACGGAAGAAGGGGATGCATATCTGTTTGCCAGTAAAGATAAGCGCATTGCCTTTGTGACGGGCCATCCCGAATATGATGCGCAAACGCTGGCGCAGGAA
121
PhePheArgAspValGluAlaGlyLeuAspProAspValProTyrAsnTyrPheProHisAsnAspProGlnAsnThrProArgAlaSerTrpArgSerHisGlyAsnLeuLeuPheThr
TTTTTCCGCGATGTGGAAGCCGGACTAGACCCGGATGTACCGTATAACTATTTCCCGCACAATGATCCGCAAAATACACCGCGAGCGAGCTGGCGTAGTCACGGTAATTTACTGTTTACC
241
AsnTrpLeuAsnTyrTyrValTyrGlnIleThrProTyrAspLeuArgHisMetAsnProThrLeuAsp***                                               1
AACTGGCTCAACTATTACGTCTACCAGATCACGCCATACGATCTACGGCACATGAATCCAACGCTGGATTAAtcttctgtgatagtcgatcgttaagcgattcagcaccttacctcaggc
361       1                                   2        2                                        3                -35
accttcgggtgcctttttatttccgaaacgtacctcagcaggtgaataaatttattcatattgttatcaacaagttatcaagtatttttaattaaaatggaaattgtttttgatttttg

481 3          -10      4 *    5           4           5                                            MetThrGluGlnAlaThrTh
cattttaaatgagtagtcttagttgtgctgaacgaaaagagcacaacgatccttcgttcacagtggggaagttttcggatccatgacgaggagctgcacgATGACTGAACAGGCAACAAC
601
rThrAspGluLeuAlaPheThrArgProTyrGlyGluGlnGluLysGlnIleLeuThrAlaGluAlaValGluPheLeuThrGluLeuValThrHisPheThrProGlnArgAsnLysLe
AACCGATGAACTGGCTTTCACAAGGCCGTATGGCGAGCAGGAGAAGCAAATTCTTACTGCCGAAGCGGTAGAATTTCTGACTGAGCTGGTGACGCATTTTACGCCACAACGCAATAAACT
721
uLeuAlaAlaArgIleGlnGlnGlnGlnAspIleAspAsnGlyThrLeuProAspPheIleSerGluThrAlaSerIleArgAspAlaAspTrpLysIleArgGlyIleProAlaAspLe
TCTGGCAGCGCGCATTCAGCAGCAGCAAGATATTGATAACGGAACGTTGCCTGATTTTATTTCGGAAACAGCTTCCATTCGCGATGCTGATTGGAAAATTCGCGGGATTCCTGCGGACTT
841
uGluAspArgArgValGluIleThrGlyProValGluArgLysMetValIleAsnAlaLeuAsnAlaAsnValLysValPheMetAlaAspPheGluAspSerLeuAlaProAspTrpAs
AGAAGACCGCCGCGTAGAGATAACTGGCCCGGTAGAGCGCAAGATGGTGATCAACGCGCTCAACGCCAATGTGAAAGTCTTTATGGCCGATTTCGAAGATTCACTGGCACCAGACTGGAA
961
nLysValIleAspGlyGlnIleAsnLeuArgAspAlaValAsnGlyThrIleSerTyrThrAsnGluAlaGlyLysIleTyrGlnLeuLysProAsnProAlaValLeuIleCysArgVa
CAAAGTGATCGACGGGCAAATTAACCTGCGTGATGCGGTTAACGGCACCATCAGTTACACCAATGAAGCAGGCAAAATTTACCAGCTCAAGCCCAATCCAGCGGTTTTGATTTGTCGGGT
1081
lArgGlyLeuHisLeuProGluLysHisValThrTrpArgGlyGluAlaIleProGlySerLeuPheAspPheAlaLeuTyrPhePheHisAsnTyrGlnAlaLeuLeuAlaLysGlySe
ACGCGGTCTGCACTTGCCGGAAAAACATGTCACCTGGCGTGGTGAGGCAATCCCCGGCAGCCTGTTTGATTTTGCGCTCTATTTCTTCCACAACTATCAGGCACTGTTGGCAAAGGGCAG
1201
rGlyProTyrPheTyrLeuProLysThrGlnSerTrpGlnGluAlaAlaTrpTrpSerGluValPheSerTyrAlaGluAspArgPheAsnLeuProArgGlyThrIleLysAlaThrLe
TGGTCCCTATTTCTATCTGCCGAAAACCCAGTCCTGGCAGGAAGCGGCCTGGTGGAGCGAAGTCTTCAGCTATGCAGAAGATCGCTTTAATCTGCCGCGCGGCACCATCAAGGCGACGTT
1321
uLeuIleGluThrLeuProAlaValPheGlnMetAspGluIleLeuHisAlaLeuArgAspHisIleValGlyLeuAsnCysGlyArgTrpAspTyrIlePheSerTyrIleLysThrLe
GCTGATTGAAACGCTGCCCGCCGTGTTCCAGATGGATGAAATCCTTCACGCGCTGCGTGACCATATTGTTGGTCTGAACTGCGGTCGTTGGGATTACATCTTCAGCTATATCAAAACGTT
1441
uLysAsnTyrProAspArgValLeuProAspArgGlnAlaValThrMetAspLysProPheLeuAsnAlaTyrSerArgLeuLeuIleLysThrCysHisLysArgGlyAlaPheAlaMe
GAAAAACTATCCCGATCGCGTCCTGCCAGACAGACAGGCAGTGACGATGGATAAACCATTCCTGAATGCTTACTCACGCCTGTTGATTAAAACCTGCCATAAACGCGGTGCTTTTGCGAT
1561
tGlyGlyMetAlaAlaPheIleProSerLysAspGluGluHisAsnAsnGlnValLeuAsnLysValLysAlaAspLysSerLeuGluAlaAsnAsnGlyHisAspGlyThrTrpIleAl
GGGCGGCATGGCGGCGTTTATTCCGAGCAAAGATGAAGAGCACAATAACCAGGTGCTCAACAAAGTAAAAGCGGATAAATCGCTGGAAGCCAATAACGGTCACGATGGCACATGGATCGC
1681
aHisProGlyLeuAlaAspThrAlaMetAlaValPheAsnAspIleLeuGlySerArgLysAsnGlnLeuGluValMetArgGluGlnAspAlaProIleThrAlaAspGlnLeuLeuAl
TCACCCAGGCCTTGCGGACACGGCAATGGCGGTATTCAACGACATTCTCGGCTCCCGTAAAAATCAGCTTGAAGTGATGCGCGAACAAGACGCGCCGATTACTGCCGATCAGCTGCTGGC
1801
aProCysAspGlyGluArgThrGluGluGlyMetArgAlaAsnIleArgValAlaValGlnTyrIleGluAlaTrpIleSerGlyAsnGlyCysValProIleTyrGlyLeuMetGluAs
ACCTTGTGATGGTGAACGCACCGAAGAAGGTATGCGCGCCAACATTCGCGTGGCTGTGCAGTACATCGAAGCGTGGATCTCTGGCAACGGCTGTGTGCCGATTTATGGCCTGATGGAAGA
1921
pAlaAlaThrAlaGluIleSerArgThrSerIleTrpGlnTrpIleHisHisGlnLysThrLeuSerAsnGlyLysProValThrLysAlaLeuPheArgGlnMetLeuGlyGluGluMe
TGCGGCGACGGCTGAAATTTCCCGTACCTCGATCTGGCAGTGGATCCATCATCAAAAAACGTTGAGCAATGGCAAACCGGTGACCAAAGCCTTGTTCCGCCAGATGCTGGGCGAAGAGAT
2041
tLysValIleAlaSerGluLeuGlyGluGluArgPheSerGlnGlyArgPheAspAspAlaAlaArgLeuMetGluGlnIleThrThrSerAspGluLeuIleAspPheLeuThrLeuPr
GAAAGTCATTGCCAGCGAACTGGGCGAAGAACGTTTCTCCCAGGGGCGTTTTGACGATGCCGCACGCTTGATGGAACAGATCACCACTTCCGATGAGTTAATTGATTTCCTGACCCTGCC
2161
oGlyTyrArgLeuLeuAla***                             MetLysThrArgThrGlnGlnIleGluGluLeuGlnLysGluTrpThrGlnProArgTrpGluGlyIle
AGGCTACCGCCTGTTAGCGTAAaccaccacataactatggagcatctgcacATGAAAACCCGTACACAACAAATTGAAGAATTACAGAAAGAGTGGACTCAACCGCGTTGGGAAGGCATT
```

References

(1) Nunn, W.D. (1986) Microbiol. Rev. 50, 179-192. (2) Rieul, C., Bleicher, F., Duclos, B., Cortay, J.C., and Cozzone, A.J. (1988) Nucleic Acids Res. 16, 568. (3) Chung, T., Klumpp, D.J. and LaPorte, D.C. (1988) J. Bacteriol. 170, 386-392.