

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

Carolyn Byrne, Harold W. Stokes¹ and Kevin A. Ward

Division of Animal Production, CSIRO, PO Box 239, Blacktown, NSW 2148 and ¹School of Biological Sciences, Macquarie University, North Ryde, NSW 2113, Australia
 Submitted October 20, 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
GATCTGGAAATTCGGCAGACGGGAAGAGGGGATGCATATCTGTTGCCAGTAAAGATAACGGCATTGCCTTTGCAOOGGCCATCCGGAATGATGCGCAAAAGCTGGCGCAGCA
121
PhePheArgAspValGluAlaGlyLeuAspProAspValProTyrAsnTyrPheProHisAsnAspProGlnAsnThrProArgAlaSerTrpArgSerHisGlyAsnLeuPheThr
TTTTTCGGAGTGTGAAGCOGGACTAGACCOGGATGACCGGTATAAATTTTCCGCAACAATGATCCGCAAAATACACCGGAGCGAGCTGGCGTAGTCCGGTAATTACTGTTTACC
241
AsnTrpLeuAsnTyrTyrValTyrGlnIleThrProTyrAspLeuArgHisMetAsnProThrLeuAsp***
AACCTGGCTCACTATTAGCTTACCAAGCTATACGATACGATCTACCGCATGATCCACCGTGATTAACTctctgtgatcgatcgttaagcgattcagacccttacctcaggc
361
acctctgggtggcctttttattcttcgcgaactcctcagcagggtgaatatttattctatattgttatccacaagttatccagttatcttaacttaaatgganctgtttgatcttg
481 3 -10 4 * 5 4 5 MetThrGluGlnAlaThrTh
CATTTTaaatgagtagcttctgttctgctgaacgaasagcaccacagatcctctgttcacagtggggaagtttctggatccatgacggagctgcagcATGACTGAACAGCGCAACAA
601
rThrAspGluLeuAlaPheThrArgProTyrGlyGluGlnGluLysGlnIleLeuThrAlaGluAlaValGluPheLeuThrGluLeuValThrHisPheThrProGlnArgAsnLysLe
ACCGATGAACCTGGCTTTCACAGCCGGTATGGCGAGCAGGAGAACAAATCTTACTGCGAAGCGGTAGAATTTCTGACTGAGCTGCGAGCATTTTAAGCCACAAAGCGCAATAAAT
721
uLeuAlaAlaArgIleGlnGlnGlnAspIleAspAsnGlyThrLeuProAspPheIleSerGluThrAlaSerIleArgAspAlaAspTrpLysIleArgGlyIleProAlaAspLe
TCTGGCAGCGCAGTACAGCAGCAGCAAGATATTGATAACGGAAGCTGCTGCTGATTTTATTGCGAAGCAGCTCCATTCGGAGCTGCTGATGCAAAATTCGCGGATCTCTCGGACT
841
uGluAspArgArgValGluIleThrGlyProValGluArgLysMetValIleAsnAlaLeuAsnAlaAsnValLysValPheMetAlaAspPheGluAspSerLeuAlaAspTrpAs
AGAACCCGCGCGCTAGAGATACTGGCCGCTAGAGCGCAAGATGCTGATCAACGCCCTCAACGCCAATGCGAAAGCTTTTATGCGGATTTGCAAGATCTCACTGGCAGCAGCTGGAA
961
nLysValIleAspGlyGlnIleAsnLeuArgAspAlaValAsnGlyThrIleSerTyrThrAsnGluAlaGlyLysThrIleSerGlnLeuAsnProAlaValLeuLysCysArgV
CAAAATGATGACCGGCAAAATTAACCTGGCTGATGGCTTAAAGCGACCACTTACACCAATGAAGCAGGCAAAATTAACGCAAGCCAAATTCGAGGCTTTGATTTGCTGGGGT
1081
lArgGlyLeuHisLeuProGluLysHisValThrTrpArgGlyGluAlaIleProGlySerLeuPheAspPheAlaLeuTyrPhePheHisAsnTyrGlnAlaLeuAlaLysGlySe
AGCGGCTGTCAGCTTCGGGAAAAACATCTCACCTGGCTGGTAGGCAATCCCGCGGACGGCTTTGATTTGCGCTCTATTCTTCACAACATACAGGCTGCTGGCAAAGCGCAG
1201
rGlyProTyrPheThrLeuProLysThrGlnSerTrpGlnGluAlaIleTrpTrpSerGluValPheSerTyrAlaGluAspArgPheAsnLeuProArgGlyThrIlePheAlaThrLe
TGCTCCATTTCTATCTGCGAAACCACTCTGCTGGCAGGAGCGCCTGCTGGAGCGAAGCTTTCAGCTATGCGAAGATGCTCTTAATCTGCGCGCGGCACATCAGGCGAAGCTT
1321
uLeuIleGluThrLeuProAlaValPheGlnMetAspGluIleLeuHisAlaLeuArgAspHisIleValGlyLeuAsnCysGlyArgTrpAspTyrIlePheSerTyrIleLysThrLe
CGTATGAAACGCTGCGCGCGCTTCCAGATGGATGAAATCTTCAOOGCGCTGACCATATGTTGCTCTGAAGCTGCGGCTGTTGGGATACATCTTCAAGCTATATAAAAGCTT
1441
uLysAsnTyrProAspArgValLeuProAspArgGlnAlaValThrMetAspLysProPheLeuAsnAlaTyrSerArgLeuLeuIleLysThrCysHisLysArgGlyAlaPheAlaMe
GAAAAACTATCCGATGCGCTCCTGCCAGACAGCAGGCAAGTGAAGATGGATAAACCATCTCTGAATGCTTACTCAGCGCTGTTGATTAACACTGCCATAAAGCGGGCTTTTGGGAT
1561
rGlyGlyMetAlaAlaPheIleProSerLysAspGluGlnHisAsnGlnValLeuAsnLysValLysAlaAspLysSerLeuGluAlaAsnAsnGlyHisAspGlyThrTrpIleAl
GGCGGCTGCGCGCTTATTCGAGCAAAAGATGAAGACCAATAACCAAGCTGCTCAAAAAGTAAAGCGGATAAATGCTGCGAAGCCAATAACGGCTCAGCATGCCATGGATGCGC
1681
hisProGlyLeuAlaAspThrAlaMetAlaValPheAsnAspIleLeuGlySerArgLysAsnGlnLeuGluValMetArgGluGlnAspAlaProIleThrAlaAspGlnLeuAla
TCACCAGGCGCTTCGGCAGCAATGCGGATTTCAACGACATCTCCGCTAAATAATCAAGCTTGAAGCTGATGCGGCAAGAACGCGGATTAAGCTGATGCGGATTTATGCGCTGATGGAAGA
1801
aProCysAspGlyGluArgThrGluGluGlyMetArgAlaAsnIleArgValAlaValGlnTyrIleGluAlaTrpIleSerGlyAsnGlyCysValProIleTyrGlyLeuMetGluAs
ACGTTGTGATGGTGAACGCACAGGAAGGATATGCGGCCAACATTTGCGTGGCTGTCAGTACATGCAAGCGTGGATCTGCGAAGCGGCTGTCGCGGATTTATGCGCTGATGGAAGA
1921
pAlaAlaThrAlaGlnIleSerArgThrSerIleTrpGlnTrpIleHisGlnLysThrLeuSerAsnGlyLysProValThrAlaAspLeuHisLeuPheGlnMetGluGlyGluLe
TGGCGGACCGCTGAATTTCCGATCTGCGATGCGCTGATGATCATCAAAAAGCTGCAAAAGCTGCAAAAGCTGCAAAAGCTGCTGCGCGAGATGCTGGCGAAGAGAT
2041
tLysValIleAlaSerLeuLeuGlyGluIleUArgPheSerGlnGlyArgPheAspAlaAlaArGlyMetGluGlnIleThrThrSerAspGluLeuIleAspPheLeuThrLeuTrp
GAAAGCTATGCGCAGCAATGCGGGAAGCTTCTCCAGGGCGTTTCCAGATGCGGACGCTGATGGAACAGATCACCATCTCGATGAGTAAATTAATGATTTCTTCCAGGCTGCC
2161
oGlyTyrArgLeuLeuAla*** MetLysThrArgThrGlnGlnIleGluGluLeuGlnLysGluTrpThrGlnProArgTrpGlyGlyIle
AGGCTACCGGCTTGTAGGCTAAaccaccacatsectatgggctctcgacATGAAAACCGCTACACACAAATTAAGAAATACAGAAAGAGTGGACTCAACCGGCTGGGAGGCAAT
    
```

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.