

Sequence of the *Escherichia coli* fructose-1,6-bisphosphatase gene

William D.O.Hamilton, Deborah A.Harrison and Tristan A.Dyer

Institute for Plant Science Research, Cambridge Laboratory, Maris Lane, Trumpington, Cambridge, CB2 2JB, UK
 Submitted August 10, 1988

Accession no. X12545

Fructose-1,6-bisphosphatase (FBPase) is an enzyme found in many different types of organism. Although it has basically the same catalytic function in each, its precise role is not the same as it may be a component of several different metabolic pathways. In *E. coli*, for example, it is necessary for growth on substances such as glycerol succinate and acetate. We have sequenced the *E. coli* gene for this protein (1,2) so that we could compare the predicted amino acid sequence of its product with other known FBPase sequences. Amino acids conserved in chloroplasts (3) mammals (4,5), yeasts (6), and *E. coli* FBPases are marked with (*). The putative coding sequence contains 332 codons giving a protein of predicted molecular weight 36,834. As amino acids in several regions are conserved in all the FBPases studied, it would seem probable that all the genes had a common progenitor. However, as discussed elsewhere (3), the ancestral sequence has evolved so that the activity of the enzyme can be regulated to suit the particular subcellular environment in which it must function.

```

1   CTCGAGCTGGCTGGCATCGTACCCCTGAAATCAGCTCAATGCCCTTCTCAAAGTAGGTGCTCATCGGCGGATACACATGGCGTCCGAACCCCTTACT
101  TCATGGCTAACTGGCGCGCCAGCATCGCCGACGCGCCATAAATGCGCCACAAATTCCTAAAATATGAATGCGCATACGTCATCTCTTTTAAATCT
201  GCCGCTCATTTTACGCATATGTCGGCGAAGTGAGAAACGCATTTCAAGSATAATCCGTAATTTTTCGCGCGCAATTCACCTGAGCGCAACATAGAGATTA
                                     *
301  TGTAAAGATTGTTGGCGTGCCTTTACTCCATAAACATTGCAGGGAAGATTTATGAAAACGTTAGGTGAATTTATGTCGAAAAGCAGCAGCAGATTTCT
                                     *
   H A T G E L T A L L S A I K L G A E I I N R D I N E A G L V D I L G
401  CATGCTACCGGTGAGCTCACTGCTTTGCTGTCGCGCAATAAAAACCTGGGCGCCAGAGATTATCCATCCGCGATATCAACAAGCAGACTGGTTGATATCC
                                     *
   A S G A E N V Q G E V Q Q K L D L F A N E K L K A A L K A R D I V
501  GTGCCAGCGGTCTGAGAAACCTGCGAGGGCGAGGTTCCAGCAGAAAACCTGCACTGTTGCTGTAATGAAAACCTGAAAAGCCGCACTGAAAAGCCCGCATGCT
                                     *
   A G I A S E E E D E I V V F E G C E H A K Y V V L M D P L D G S S
601  TGGCGGCTATGCTCTGAAAAGAAAGATGAGATTGCTGCTTTGAAAGGCTGTGAAACAGCAAAAATACGTTGGTGTGATGGACCCCGGATGGCTCGCTCC
                                     *
   H I D V N V S V G I F S I Y R R V T F V G T V E E D F L Q P G
701  AACATCGATGTAACTGCTCTGTCGGTACCATTTTCTCCATCTACC CGCGGTACGCTGTTGGCAGCGCGTAAACGGAGAAAGATTTCCTCCAGCTCG
                                     *
   H K Q V A A G Y V V Y G E S T N L V Y T T G C G V N A F T Y I D F S
801  GTAAACAAACAGGTGGCGCAGGTTACGCGGTATACGGCTCCTCTACCATGCTGGTTTTACACCACCGGATCGGTTTTCACGCTTTACTTACGATCTTCC
                                     *
   L G V F C L C Q E R N R F P E K G T Y S I N E G N Y I K F P F G
901  GCTCGGCTTTCTGCTGTGTCAGGAACGATGCGCTTCCCGGAGAAAGCAAAACCTACTCCATCAACGAGGAAACTCATTAAAGTTCCGAAACCGG
                                     *
   V K E Y I K F C Q E E D E K S T N R P Y T S R I Y G S L V A D F E R N
1001  GTGAAGAATACATTAATCTCGCGAAGAAATAAATCCACCAACCGCCCTTATACCTCAGCTGATATCGGTTCACTGGTCGCGGATTTCCACCGTA
                                     *
   L L E G G I Y L Y P S T A S E H P D G K L R L L Y E C H P N A F L A
1101  ACTCTGTAAAAGCGGTATTTATCTCTACCCAGACACCGCCAGCCCGGACGCAAACTGCGTTGCTGTATGAGTGCACCCGATGGCATCTCTCGGC
                                     *
   E Q A G G K A S D G K E R I L D I I F E T L E Q R S F P V G E D
1201  GGAACAAGCGGCGGTAAAAGCGAGCATGGGCAAGAGCGTATTCTGGATATCATCTCCGGAACCTGCGACCAAGCGCGGTTCACTCTTGTGCGCAACCA
                                     *
   H N V E D V E R F I R E F P D A
1301  CATATGGTGAAGATGTGCGAAGCCTTATCCGAGGTTCCCGGACGCGTAATCTCTGCGCCTCCACACGGAGAGCGTTTTTATGCTGCGCAATGT
1401  TGACTCTTGTCTGCGCAACTCTGACGATTCCTCCACAGCACCCTTAATCCAGCTGTAAATGCTATAAAAATAAACACAAAATACCGATGGCGATTTT
1501  CGTCCACCAAGAACTCAGCGTCCATCAAAATTTATGTAAGTCTGAATCACTCCCTGAATCCGCGACGCCAAAAGAGCGTCCCTAATACCGTTCCCAACGCCA
1601  CCCTCAAAAC
    
```

We are grateful to Prof. Dan G. Fraenkel for providing clones pJS33 and pJS35 for sequencing and Jacqueline C. Timans for technical assistance. W.D.O.H., D.A.H. and J.C.T. are employees of the Agricultural Genetics Company.

References: (1) Sedivy JM, Daldal F & Fraenkel DG (1984) *J Bacteriol* 158, 1048-1053; (2) Sedivy JM, Babul J & Fraenkel DG (1986) *Proc Natl Acad Sci USA* 83, 1656-1659; (3) Raines CA, Lloyd JC, Longstaff M, Bradley D & Dyer TA (1988) *Nucl Acids Res*, in press; (4) Fisher WK & Thompson EOP (1983) *Aust J Biol Sci* 36, 235-250; (5) Marcus F, Edelstein I, Reardon I & Heinrikson RL (1982) *Proc Natl Acad Sci USA* 79, 7161-7165; (6) Rogers DT, Hiller E, Mitsok L & Orr E (1988) *J Biol Chem* 263, 6051-6057.