Volume 310, number 2, 167–170 FEBS 11572 September 1992

Cloning and sequencing of glutamate mutase component S from *Clostridium tetanomorphum*

Homologies with other cobalamin-dependent enzymes

E. Neil G. Marsh and Daniel E. Holloway

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Received 16 July 1992; revised version received 10 August 1992

The gene encoding component S, the small subunit, of glutamate mutase, an adenosylcobalamin (coenzyme B_{12})-dependent enzyme from *Clostridium tetanomorphum* has been cloned and its nucleotide sequence determined. The *mutS* gene encodes a protein of 137 amino acid residues, with M_r 14,748. The deduced amino acid sequence showed homology with the C-terminal portion of adenosylcobalamin-dependent methylmalonyl-CoA mutase [1989, Biochem. J. 260, 345–352] and a region of cobalamin-dependent methionine synthase which has been shown to bind cobalamin [1989, J. Biol. Chem 264, 13888–13895].

Glutamate mutase; Cobalamin-dependent enzyme; Sequence homology

1. INTRODUCTION

Glutamate mutase catalyses the first step in the fermentation of glutamate by *Clostridium tetanomorphum*. This is an unusual isomerization in which L-glutamate is converted to *threo*-β-methyl-L-aspartate [1]. It was the first example to be discovered of a group of adenosylcobalamin (AdoCbl)-dependent enzymes which catalyze the general reaction:

where X is an electron withdrawing group which may be OH, NH_2 or, as in this case, a carbon-containing fragment, so that a skeletal rearrangement is effected [2]. The role of AdoCbl as the intermediate hydrogen carrier in these rearrangements is well established [3]; however, very little is known of the proteins' role in catalyzing these mechanistically complex reactions.

Glutamate mutase consists of two readily separable components, designated E and S [4,5]. Component S is a small, monomeric protein of $M_r \approx 15{,}000$, whilst component E is a homodimer of subunit $M_r \approx 50{,}000$ ([6], E.N.G. Marsh, unpublished results). Component E has been shown to bind both substrate and coenzyme, but is absolutely dependent upon component S for activity [4]. E and S interact weakly with one another in a manner which exhibits saturation kinetics. The function of S, which binds neither AdoCbl nor substrate on its own, is unclear; it may interact with AdoCbl bound on component E, as increasing the molar ratio of S to E decreases the apparent K_m for AdoCbl [5].

The sequences of several AdoCbl-dependent enzymes [7–10] and cobalamin-dependent methionine synthase [11] have recently been determined, together with various cobalamin-binding and transport proteins [12–16]. Until now, though, no obvious homologies have emerged, which might, for instance, identify residues involved in cobalamin binding. The crystallization of methylmalonyl-CoA mutase from *Propionibacterium shermanii* and a cobalamin-binding proteolytic fragment of *Escherichia coli* methionine synthase have been reported [17,18], but at present no three-dimensional structures are known for any of these proteins.

Here we report the cloning and sequencing of *mutS*, the gene encoding component S of glutamate mutase from *C. tetanomorphum*. The sequence has allowed us to identify residues which are conserved between component S, methylmalonyl-CoA mutase from several species, and methionine synthase.

2. MATERIALS AND METHODS

2.1. *Construction of a probe for the* mutS *gene*

Both component S and component E were purified from *C. tetanomorphum* ([4,5] and E.N.G. Marsh, unpublished results) and their N-terminal sequences determined by automated methods. The protein

Correspondence address: E.N.G. Marsh, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK. Fax: (44) (223)333 345.

sequences were used to design two oligonucleotides for use as primers in a PCR experiment. The amplified DNA fragment produced by the PCR was subcloned into pUC118 to give the construct, pNM1. The DNA sequence at each end of the insert of pNM1 was determined to confirm that the PCR had amplified the correct region of genomic DNA. The insert of pNM1 was excised from the vector by restriction with *Eco*RI endonuclease and purified by agarose gel electrophoresis. The purified insert was then used as a template to produce a ^{32}P-labeled random-primed probe [19].

2.2. *Construction and screening of genomic DNA library*

Genomic DNA was purified from *C. tetanomorphum* (NCIMB No. 11547) by standard methods and used to construct a library in the λ vector, EMBL3 [19]. Approximately 60,000 plaques were screened using the random-primed probe according to published protocols [19].

2.3. *Subcloning and sequence analysis*

Positively hybridizing λ clones were analyzed by restriction mapping, and an appropriate restriction fragment subcloned into pUC119. The DNA sequence was determined by 'shotgun' cloning of sonicated fragments into M13mp18 [20] and by double-stranded sequencing [21] of the parent pUC119 plasmid using specifically synthesized oligonucleotide primers. The sequence data was compiled and analyzed using the Staden [22] and University of Wisconsin Genetics Computer Group programs [23]. Secondary structure predictions used the Garnier–Osguthorpe–Robson algorithm [24].

3. RESULTS AND DISCUSSION

The N-terminal protein sequences obtained from component S and component E by automated methods are shown in Fig. 1. From these, two oligonucleotides were designed to function as primers in a PCR experiment (Fig. 1), the expectation being that both components would be transcribed as part of the same operon. Each oligonucleotide incorporated an *Eco*RI site at the 5′ end to facilitate cloning into a pUC vector. The oligonucleotide based on component E was reverse complemented so that PCR would amplify all of the *mutS* gene, assuming *mutS* to be upstream of *mutE*. Fortunately this assumption proved correct and PCR amplified a 1,800 bp fragment which was subcloned into pUC118 to give pNM1. The PCR product was considerably longer than required to encode *mutS* alone (≈ 450 bp), implying that it was separated from *mutE* by over 1,300 bp on the chromosome. Double-stranded sequencing of pNM1 using 'universal' and 'reverse' primers confirmed that the insert encoded the expected protein sequences and hence that the intended region of DNA had been amplified.

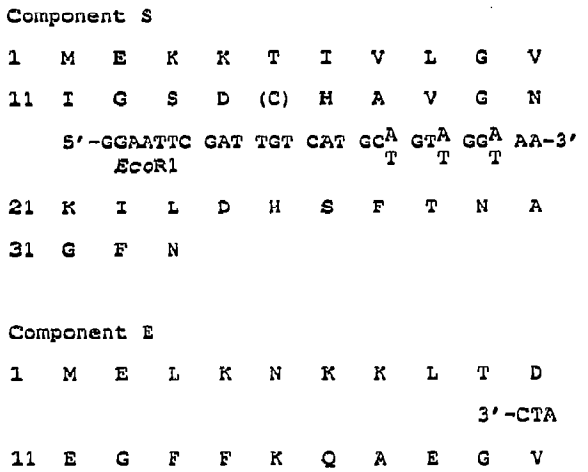

Fig. 1. N-Terminal protein sequences obtained for components S and E of glutamate mutase. The oligonucleotide PCR primers which were designed from the protein sequences are shown aligned with the corresponding amino acid residues. The high A–T content of clostridial DNA allowed the degeneracy of the oligonucleotides to be minimized by using only A and/or T at positions of ambiguity. The cysteinyl residue shown in parenthesis denotes that this assignment was only tentative. Note that the oligonucleotide designed from the component E sequence was synthesized as the non-coding strand and its sequence is shown here running 3′ to 5′.

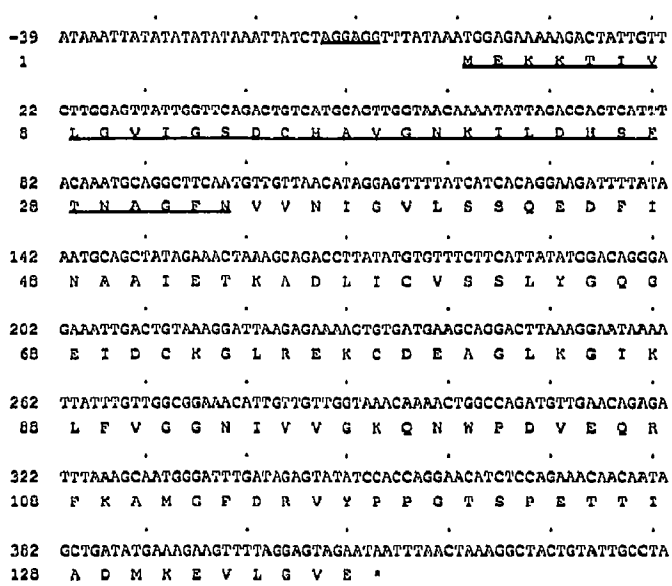

Fig. 2. The nucleotide and deduced protein sequence of the *mutS* gene. The numbering is from the A of the ATG start codon for the DNA sequence and from the N-terminal methionine of the protein sequence. A potential ribosome binding site is underlined in the DNA sequence. The underlined protein sequence was previously determined by automated sequencing of component S. The TAA stop codon is denoted by an asterisk.

The random-primed probe made from the insert of pNM1 was used to screen a λ library of *C. tetanomorphum* genomic DNA. Out of ≈60,000 plaques screened, two positively hybridizing clones were found. Restriction analysis of these identified a 5 kbp *Sal*I fragment that hybridized to the probe and which mapped to the end of the insert adjacent to the left arm of EMBL3. This was subcloned into pUC119 to give the construct, pGM1. Partial sequencing of pGM1 determined the entire nucleotide sequence of *mutS*; it also established that the *mutE* gene was not present as a full-length copy in either pGM1 or the parent λ clone. The nucleotide sequence and deduced protein sequence of *mutS* is

```
          1      +++ + +  * *   *  ++          +++   + ++   ++ 50
mutB_Hum  REGRRPRLLV AKMGQDGHDR GAKVIATGFA DLGFDVDIGP LFQTPREVAQ
mutB_Mou  REGRRLGLLV AKMGKDGHDR GAKVIATGFA DLGFDVDIGP LFQTPREVAH
mutB_Psh  AEGRRPRILL AKMGQDGHDR GQKVIATAYA DLGFDVDVGP LFQTPEETAR
   mutS   ..MEKKTIVL GVIGSDCHAV GNKILDHSFT NAGFNVVNIG VLSSQEDFIN
   metH   QGKTNGKMVI ATVKGDVHDI GKNIVGVVLQ CNNYEIVDLG VMVPAEKILR

         51* +     + ++*+  +         + ++     *   *++ + **   100
mutB_Hum  QAVDADVHAV GVSTLAAGHK TLVPELIKEL NSLGRPDILV MCGGVIPPQD
mutB_Mou  DAVDADVHAV GVSTHAAGHK TLVPELIKEL TALGRPDILV MCGGVIPPQD
mutB_Psh  QAVEADVHVV GVSSLAGGHL TLVPALRKEL DKLGRPDILI TVGGVIPEQD
   mutS   AAIETKADLI CVSSLYGQGE IDCKGLREKC DEAGLKGIKL FVGGNIVVGK
   metH   TAKEVNADLI GLSGLITPSL DEMVHVAKEM ERQGF.TIPL LIGGATTSKA

        101       +        ++         ++   +              150
mutB_Hum  YEF......L FEVGVSNVFG PGTRIPKAAV QVLDDIEKCL EKKQQSV*..
mutB_Mou  YEF......L YEVGVSNVFG PGTRIPRAAV QVLDDIEKCL AEKQQSV*..
mutB_Psh  FDE......L RKDGAVEIYT PGTVIPESAI SLVKKLRASL DA*.......
   mutS   QNWPDVEQRF KAMGFDRVYP PGTS.PETTI ADMKEVLGVE *.........
   metH   HTAVKIEQNY SGPT...VYV QNAS.R..TV GVVAALLSDT QRDDFVARTR
```

Fig. 3. Alignment of the deduced amino acid sequence of mutS with other cobalamin-dependent enzymes. MutB_Hum, mutB_Mou and mutB_Psh represent residues 602–742 of human, 608–748 of mouse and 593–728 of *P. shermanii* methylmalonyl-CoA mutase, respectively. MutS represents residues 1–137 of glutamate mutase component S, and metH residues 741–883 of *E. coli* methionine synthase. Invariant residues are shown in bold type and are marked by *. Conservatively substituted positions are marked by +. The numbering above the sequences is arbitrary.

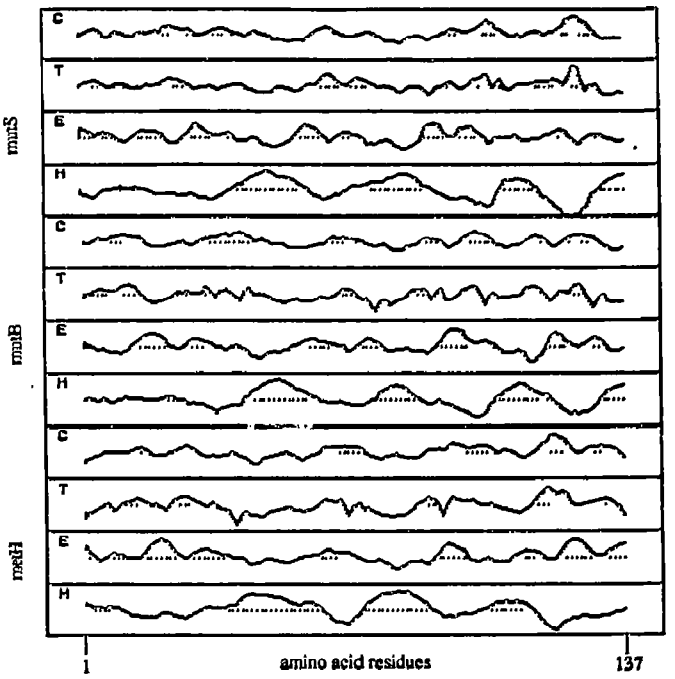

Fig. 4. Secondary structure predictions based on the Garnier–Osguthorpe–Robson algorithm [24]. Top, mutS: residues 1–137 of glutamate mutase component S. Middle, mutB: residues 592–728 of *P. shermanii* methylmalonyl-CoA mutase. Bottom, metH: residues 743–880 of *E. coli* methionine synthase. C, random coil; T, β-turn; E, β-sheet (extended); H, α-helix.

shown in Fig. 2. The gene is flanked on both sides by unidentified open reading frames which are transcribed in the same direction as *mutS*. *MutS* encodes a protein of 137 amino acid residues with M_r 14,748, a value which is in good agreement with that previously determined by SDS-PAGE and gel filtration ([6], E.N.G. Marsh, unpublished results).

A search of the SWISSPROT protein sequence database identified only one protein as significantly similar to mutS which was the large subunit of adenosylcobalamin-dependent methylmalonyl-CoA mutase (mutB) from *P. shermanii* [7]. MutS aligned with the C-terminal portion of mutB and was homologous throughout its length (Fig 3.). This result was surprising since mutS on its own does not bind cobalamin [5]. The sequences of other cobalamin-binding proteins were extracted from the database [8–16] and compared with mutS. Similar homologies were found between mutS and methylmalonyl-CoA mutases from human and mouse [8,9], and also methionine synthase (metH) from *E. coli* [11] (Fig. 3). Significantly, the homology with metH is in a region which is encompassed by a 28 kDa tryptic peptide that has been shown to bind cobalamin [25]. Two short regions of highly conserved sequence stood out in the alignment and were identified by the motifs, DXHXXG (from D14 of mutS), and GXXXIXXXXGG (from G82 of mutS), which were invariant in all the proteins.

When secondary structure predictions [24] were computed for mutS and the regions of metH and mutB which aligned with it, very similar patterns emerged for all three proteins (Fig. 4). The analysis predicted an overall pattern of strands of β-sheet alternating with α-helices. In particular, the invariant residues, DXHXXG, lie at a predicted β-turn between two strands of β-sheet whilst the glycyl residues of GXXXIXXXXGG mark each end of a region of predicted β-sheet between two helices. This pattern of alternating sheet and helix is found in the Rossman fold of nucleotide-binding proteins, and it has been proposed that such a fold may be involved in binding the dimethylbenzimidazole ribofuranosyl portion of cobalamin [12]. Thus, one attractive hypothesis is that mutS may activate AdoCbl towards cleavage of the axial cobalt–carbon bond by interacting with the dimethylbenzimidazole ligand on the opposite face of the corrin ring. Caution is necessary in making such predictions, though, as a benzimidazole ribofuranosyl-binding Rossman fold structure has previously been suggested for regions of metH and mutB, based solely on secondary structure homology [25]. However, this homology was between *different* regions of mutB and metH from the primary structure homology with mutS we have reported here.

In conclusion, the sequence of mutS represents the first detailed structural information on glutamate mutase. Using this information it has been possible to iden-

tify clear sequence homologies relating three cobalamin-dependent enzymes which tentatively suggest a function for mutS.

Acknowledgements: We thank Dr. Jésus Cortes for his generous gifts of vectors and strains and especially for his help and advice throughout. Automated protein sequence determinations were carried out by Dr. James McCormick and oligonucleotide synthesis undertaken by Mr. Mike Weldon, both of the SERC/Wellcome-funded Protein and Nucleic Acid Chemistry Facility (Department of Biochemistry, University of Cambridge). Financial support was provided by SERC and the Royal Society. E.N.G.M. is a member of the Cambridge Centre for Molecular Recognition.

REFERENCES

[1] Barker, H.A., Roove, V., Suzuki, F. and Iodice, A.A. (1964) J. Biol. Chem. 239, 3260–3266.
[2] Babior, B.M. and Krouwer, J. (1979) C.R.C. Crit. Rev. Biochem. 6, 35–102.
[3] Golding, B.T. and Rao, D.N.R. (1986) in: Enzyme Mechanisms (Page. M.I. and Wiliams, A. eds.) Adenosylcobalamin-dependent Enzymic Reactions, pp. 404–428, Royal Society of Chemistry, London.
[4] Suzuki, F. and Barker, H.A. (1966) J. Biol. Chem. 241, 878–888.
[5] Switzer, R.L. and Barker, H.A. (1967) J. Biol. Chem. 242, 2658–2674.
[6] Leutbecher, U., Bocher, R., Dietman, L. and Buckel, W. (1992) Eur. J. Biochem. 205, 759–765.
[7] Marsh, E.N., McKie, N., Davis, N.K. and Leadlay, P.F. (1989) Biochem. J. 260, 345–352.
[8] Jansen, R., Kalousek, F., Fenton, W.A., Rosenberg, L.E. and Ledley, F.D. (1989) Genomics 4, 198–205.
[9] Wilkemeyer, M.F., Crane, A.M. and Ledley, F.D. (1990) Biochem. J. 271, 449–455.
[10] Faust, L.R.P., Connor, J.A., Roof, D.M., Hoch, J.A. and Babior, B.M. (1990) J. Biol. Chem. 265, 12462–12466.
[11] Old, I.G., Margarita, D., Glass, R.E. and Saint Girons, I. (1990) Gene 87, 15–21.
[12] Dieckgraefe, B.K., Seetharam, B., Banaszak, L., Leykam, J.F. and Alpers, D.H. (1988) Proc. Natl. Acad. Sci. USA 85, 46–50.
[13] Lundrigan, M.D. and Kadner, R.J. (1989) J. Bacteriol. 171, 154–161.
[14] Heller, K. and Kadner, R.J. (1985) J. Bacteriol. 161, 904–908.
[15] Johnson, J., Bollenkens, J., Allen, R.H. and Berliner, N. (1989) J. Biol. Chem. 264, 15754–15757.
[16] Platica, O., Janeczko, R., Quadros, E.V., Regec, A., Romain, R. and Rothenburg, S.P. (1991) J. Biol. Chem. 266, 7860–7863.
[17] Marsh, N., Leadlay, P.F. and Evans, P.R. (1988) J. Mol. Biol. 200, 421–422.
[18] Luschinsky, C.L., Drummond, J.T., Matthews, R.G. and Ludwig, M.L. (1992) J. Mol. Biol. 225, 557–560.
[19] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[20] Bankier, A.T., Weston, K.M. and Barrell, B.G. (1987) Methods Enzymol. 155, 51–93.
[21] Murphy, G. and Ward, E.S. (1989) in: Nucleic Acids Sequencing: A Practical Approach (Howe, C.J. and Ward, E.S. eds.) pp. 99–115, IRL Press, Oxford.
[22] Staden, R. (1986) Nucleic Acids Res. 14, 217–238.
[23] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[24] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97–120.
[25] Banerjee, R.V., Johnston, N.L., Sobeski, J.K., Dutton, P. and Matthews, R.G. (1989) J. Biol. Chem. 264, 13888–13895.