Purification, some properties and nucleotide sequence of 5-carboxymethyl-2-hydroxymuconate isomerase of *Escherichia coli* C

David I. Roper and Ronald A. Cooper

Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK

Received 2 March 1990; revised version received 17 April 1990

As part of an investigation into the evolution of catabolic pathway enzymes a cloned gene encoding the *Escherichia coli* C 5-carboxymethyl-2-hydroxymuconate (CHM) isomerase, an enzyme of the homoprotocatechuate catabolic pathway, was used to produce large amounts of the protein. The isomerase was purified to homogeneity and some of its properties determined. The reaction occurred optimally at pH 7.6 and the specificity constant was 5.8×10^5 $M^{-1} \cdot s^{-1}$ with CHM and 6.0×10^2 $M^{-1} \cdot s^{-1}$ with 2-hydroxyhepta-2,4-diene-1,7-dioate, the substrate of a second isomerase in the pathway. The pure protein showed one type of subunit of M_r 14000 whilst the molecular mass of the native enzyme was 30000, suggesting that it was a dimer of identical subunits. The first 19 N-terminal amino acids were sequenced and the data used to confirm that the open reading frame of 378 bp, identified from the nucleotide sequence, encoded the CHM isomerase.

5-Carboxymethyl-2-hydroxymuconate isomerase; Protein purification; N-terminal sequence; Gene sequence; Derived primary structure; *Escherichia coli* C

1. INTRODUCTION

How the enzymes that catalyse successive reactions in catabolic pathways have arisen is unclear. They may have evolved from a single ancestral gene, by a stepwise process of tandem gene duplication followed by mutation of the extra gene copy, and would thus be structurally related. Alternatively, they may have arisen by recruitment, from a variety of sources, of proteins with the appropriate catalytic properties and would therefore be structurally unrelated. It is also possible that, within a single pathway, both of these processes have occurred.

The homoprotocatechuate (HPC) pathway is the bacterial route whereby 4-hydroxyphenylacetate, a product of phenylalanine and tyrosine breakdown and one of several compounds released during the degradation of lignin, is catabolized. HPC is converted to succinic semialdehyde and pyruvate by a sequence of reactions initiated by *meta*-cleavage with dioxygen to give 5-carboxymethyl-2-hydroxymuconate semialdehyde (CHMS) which in turn is oxidized in an NAD-dependent reaction to give 5-carboxymethyl-2-hydroxymuconate (CHM). The subsequent reactions of CHM involve, sequentially, isomerization, decarboxylation, isomerization, hydration and aldol cleavage (Fig. 1) [1–3], all of which depend on the presence of an enolizable oxo-group in the molecule [4]. So once a protein capable of isomerizing CHM was available mutations of duplicated genes might have given rise fairly easily to altered proteins that could bring about the similar electronic shifts needed for decarboxylation, isomerization, hydration and aldol cleavage [4].

Analysis of the cloned HPC pathway genes and proteins therefore offers an opportunity to see whether enzymes catalyzing chemically related reactions have arisen by modification of a single ancestral gene. This paper describes the nucleotide sequence, the derived amino acid sequence and some properties of CHM isomerase to serve as a basis for comparison when information on the other pathway enzymes becomes available.

Correspondence address: R.A. Cooper, Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK

2. MATERIALS AND METHODS

2.1. *Bacteria, plasmids and growth conditions*

Cells were grown aerobically at 37°C in Luria broth [5] containing 100 $\mu g \cdot ml^{-1}$ ampicillin. Liquid media were solidified as required by the incorporation of 1.6% (w/v) Bacto Agar (Difco Laboratories). The *E. coli* strain DH5α (Gibco-BRL) was used as host for the pUC18- and pUC19-based plasmids used in this study. The 3.0 kbp *Bam*HI-*Sal*I fragment of pJJ801 [3] was incorporated into pUC18 and pUC19 to give pDR1830 and pDR1930, respectively. pDR1930 was used to prepare a nested set of deletions that included pDR9304 and pDR9313-9317 (Fig. 3) using a Pharmacia exonuclease III deletion kit. pDR1830 was digested with *Bam*HI and *Eco*RV to delete 0.75 kbp of cloned DNA, the cohesive ends removed by digestion with S1 nuclease and the blunt ends ligated to form pDR821 (Fig.3).

2.2. *Assay of CHM isomerase*

The enzyme was assayed as described previously [2].

2.3. *Purification of CHM isomerase*

CHM isomerase was purified from *E. coli* DH5α(pDR9304). Cells grown overnight at 37°C in 800 ml medium were harvested by cen-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

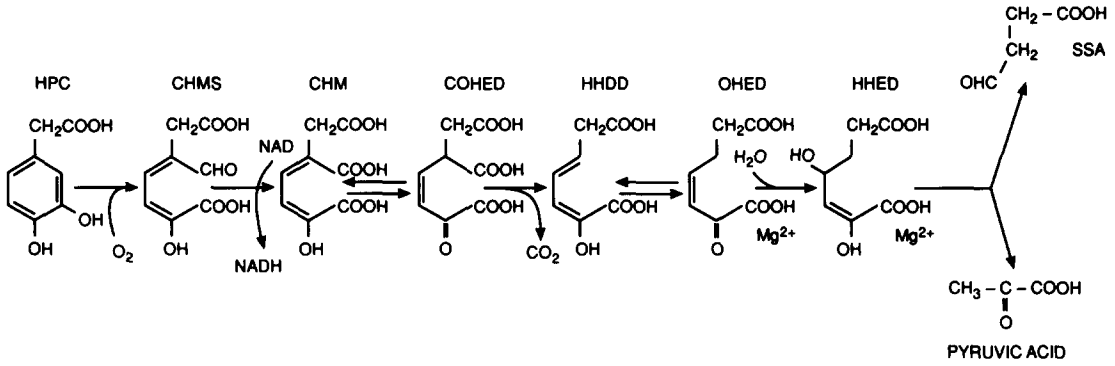

Fig. 1. Homoprotocatechuate (HPC) catabolic pathway in *E. coli*. The pathway intermediates are: 5-carboxymethyl-2-hydroxymuconate semialdehyde (CHMS); 5-carboxymethyl-2-hydroxymuconate (CHM); 5-carboxymethyl-2-oxo-hex-3-ene-1,6-dioate (COHED); 2-hydroxyhepta-2,4-diene-1,7-dioate (HHDD); 2-oxo-hept-3-ene-1,7-dioate (OHED); 2,4-dihydroxy-hept-2-ene-1,7-dioate (HHED) and succinic semialdehyde (SSA). CHM isomerase converts CHM to COHED and HHDD isomerase converts HHDD to OHED.

trifugation at 10000 × *g* for 10 min at 4°C. The pellet was washed in 160 ml 20 mM Tris-HCl buffer, pH 7.5, and finally resuspended in 16 ml of the same buffer. Cells (4 ml portions) were ultrasonicated as described previously [6]. The crude extract was ultracentrifuged at 180000 × *g* for 90 min at 4°C. The supernatant was treated with protamine sulphate (40 mg·ml^{-1} in 20 mM Tris-HCl buffer, pH 7.5) in the proportion 1 mg protamine sulphate to 20 mg bacterial protein. The mixture was stirred gently at room temperature for 10 min then centrifuged at 15000 × *g* for 15 min at 4°C. The supernatant was applied to an HR 10/10 Mono-Q anion exchange column and chromatographed using a Pharmacia fast protein liquid chromatography (FPLC) system. A 70 ml gradient of 0.0–0.5 M NaCl in 20 mM Tris-HCl buffer, pH 7.5, with a flow rate of 4 ml·min^{-1} was used and the fractions with the highest specific activities, eluted at approx. 0.25 M NaCl, were pooled. The volume of the pooled fractions was reduced to 2 ml by ultrafiltration, 2 ml of 1 M ammonium sulphate was added and the solution applied to an HR 5/5 Phenyl Superose column. A 10 ml gradient of 1.7–0.0 M ammonium sulphate in 0.1 M sodium phosphate buffer, pH 7.5, was applied at a flow rate of 0.5 ml·min^{-1}. The CHM isomerase was eluted at the end of the gradient. Fractions with the highest specific activities were pooled and the volume reduced to 2 ml by ultrafiltration. Two Pharmacia HR 10/30 Superose 12 gel filtration columns connected in series were equilibrated with 50 mM sodium phosphate/0.15 M NaCl buffer, pH 7.5. The concentrated fraction was applied to this system and eluted at a flow rate of 0.4 ml·min^{-1}. The CHM isomerase obtained was greater than 99% pure as judged by SDS-PAGE. Details of the purification are given in Table I. Protein in crude extracts was measured by the biuret method [7] and in purified extracts by the Folin method [8] using bovine serum albumin as standard. The concentration of the pure protein was measured spectrophotometrically using a value of 0.8 A_{280} per mg CHM isomerase calculated from the tryptophan and tyrosine content of the deduced amino acid sequence.

2.4. *Molecular mass estimation*

Native molecular masses were estimated by gel filtration as described under purification of CHM isomerase. The system was calibrated using IgG (160000); bovine serum albumin (66000); β-lactoglobulin (35000) and cytochrome *c* (12400) as standards. Subunit molecular masses were estimated by SDS-polyacrylamide gel electrophoresis (PAGE) [9] using gradient gels of 7.5% to 20% acrylamide. The proteins used to calibrate the gels were bovine serum albumin (66000); ovalbumin (45000); glyceraldehyde 3-phosphate dehydrogenase (36000); carbonic anhydrase (29000); trypsinogen (24000); trypsin inhibitor (20100) and bovine lactalbumin (14200).

2.5. *Amino acid sequencing*

The pure enzyme was run on a 7.5%–20% SDS-polyacrylamide gradient gel and electroblotted onto a polyvinylidene difluoride (PVDF) membrane [10] using 50 mM glycine–50 mM Tris, pH 10, transfer buffer and stained with Coomassie blue R-250. The stained protein band was excised and loaded into an Applied Biosystems 470A gas-phase sequencer without polybrene.

2.6. *DNA manipulations*

Small and large scale plasmid preparations were carried out by standard procedures [11]. Restriction endonucleases, T4 DNA ligase and exonuclease III deletion kit were used according to the manufacturer's instructions. Isolation of DNA fragments from low melting point agarose gels was as described [12]. Transformations were carried out using the morpholinepropane sulphonic acid (MOPS)-RbCl method [13].

Table I

A summary of the purification of *E. coli* C CHM isomerase

Purification step	Volume (ml)	Total protein (mg)	Total units (μmol·min^{-1})	Specific activity (μmol·min^{-1}·mg protein^{-1})	Recovery (%)	Purification (fold)
Ultracentrifuged extract	17.5	446.0	2292	5.1	100.0	0.0
Protamine sulphate treated ultracentrifuged extract	17.0	395.0	2924	7.4	127.5	1.5
Pooled Mono-Q fractions	6.0	25.2	1656	65.9	72.2	12.9
Pooled Phenyl Superose fractions	6.0	11.4	1445	127.6	63.0	25.0
Pooled Superose 12 fractions	5.0	9.0	1335	148.3	58.2	29.1

2.7. *Nucleotide sequencing*

Nucleotide sequencing was carried out with T7 DNA polymerase according to Tabor and Richardson [14]. Plasmid DNA for sequencing was prepared by the method of Kraft et al. [15].

2.8. *Chemicals*

CHM was prepared as described [3]. Restriction endonucleases were from Gibco-BRL and Pharmacia, kits for DNA sequencing and for preparing deletions were exonuclease III were from Pharmacia. All other chemicals were of the highest grade commercially available.

3. RESULTS AND DISCUSSION

3.1. *Purification of the enzyme*

Contrary to a previous report [3] the 3.0 kbp *Bam*HI-*Sal*I fragment of pJJ801 carries the gene for CHM isomerase (Roper and Cooper, unpublished results) as well as the HPC dioxygenase gene. When pDR1930 that carries this *Bam*HI-*Sal*I fragment was shortened by exonuclease III digestion one of the resulting plasmids (pDR9304) (Fig. 3), expressed much increased amounts of CHM isomerase and HPC dioxygenase and was therefore used as a source of CHM isomerase.

3.2. *Some properties of the purified enzyme*

The purified enzyme was stable for at least 4 weeks at 4°C. The effect of pH on the rate of reaction was measured in 0.1 M sodium phosphate buffer over the pH range of 6.6–8.6 and the optimum was at pH 7.6. At this pH the specificity constant (k_{cat}/K_m) [16] was 5.8×10^5 $M^{-1} \cdot s^{-1}$ with CHM and 6.0×10^2 $M^{-1} \cdot s^{-1}$ with 2-hydroxyhepta-2,4-diene-1,7-dioate, the substrate of the second isomerase in the pathway, indicating a strong preference for CHM. On SDS-PAGE the enzyme showed a single subunit of M_r 14000. The M_r of the native enzyme estimated by gel filtration was 30000 ± 2000 suggesting that the enzyme is a dimer of identical subunits.

3.3. *N-terminal amino acid sequence*

The first analysis indicated the presence of proline at position 1 so a second sample was subjected to an extended trifluoroacetic acid cleavage step during the first cycle. An unambiguous sequence for the first 19 residues was obtained with the exception of residue 7 where no assignment was possible. A sample of the blotted protein was analysed by the procedure of Amons [17] and this resulted in the identification of cysteine at position 7. The first 19 residues of CHM isomerase are shown in Fig. 2.

```
1                                   10
Pro His Phe Ile Val Glu Cys Ser Asp Asn

                                    19
Ile Arg Glu Glu Ala Asp Leu Pro Gly
```

Fig. 2. N-terminal amino acid sequence of *E. coli* C CHM isomerase.

3.4. *Nucleotide sequence of the CHM isomerase gene*

Unidirectional deletions were made from the *Sal*I site in pDR1930 to produce a series of subclones. The enzyme activity expressed by individual subclones was tested using crude extracts and CHM isomerase activity was found for subclones that retained greater than 1.1 kbp of DNA from the *Bam*HI site of the construct. The smallest deletion subclone to retain CHM isomerase activity was pDR9317. Restriction maps and the nucleotide sequencing strategy employed are shown in Fig. 3.

Fig. 3. Restriction map and sequencing strategy for the *E. coli* C CHM isomerase gene. The 3.0 kbp *Bam*HI-*Sal*I fragment (heavy line) cloned into pUC19 to give pDR1930 is shown with major restriction sites. The restriction sites shown in the polylinker are not drawn to the same scale. Unidirectional deletion subclones were created from the *Sal*I end of the insert DNA by the exonuclease III digestion procedure. The coding strand encompassing CHM isomerase was sequenced using the deletion subclones pDR9317, pDR9316, pDR9314 and pDR9313 as shown. The arrows indicate the start, direction and extent of dideoxynucleotide sequencing carried out using the M13 reverse sequencing primer. The complementary strand was sequenced from the polylinker *Eco*RI site of pDR821 using the M13 reverse sequencing primer and continued (broad arrow) using a synthetic oligonucleotide based on the determined DNA sequence at positions 256–276 of the coding strand. The 378 bp region encoding the isomerase polypeptide is indicated by the hatched box. Restriction sites are: E, *Eco*RI; Ss, *Sst*I; K, *Kpn*I; Sm, *Sma*I; B, *Bam*HI; Ev, *Eco*RV; N, *Nde*I; S, *Sal*I; P, *Pst*I; Sp, *Sph*I; H, *Hin*dIII.

Fig. 4. Nucleotide sequence and deduced amino acid sequence of the *E. coli* C CHM isomerase gene. The sequence is written in the 5′ ⟶ 3′ direction of the coding strand with the deduced amino acid sequence below. The *Eco*RV site downstream of the carboxyl terminus and the internal *Nde*I site are indicated. The amino terminus of the gene was identified by matching the predicted sequence to the sequence obtained from the purified protein. The ribosome binding site (RBS) is shown. The numbers above the lines refer to the nucleotide positions and numbering begins with the ATG initiation codon.

DNA sequencing was begun from the *Hin*dIII site of pDR9317 towards the *Bam*HI site utilizing the reverse primer site of pUC19. Because the deletion subclones had an average size difference of 100 bp and sequence for at least 300 residues was obtained from each construct there was very significant repetition of sequence measurement. The single-strand sequence was thus obtained with a high degree of accuracy. Sequence for the opposite strand was obtained by analysis of pDR821. The reverse primer site of pUC18 was used to obtain the sequence of the first 280 bases. A synthetic oligonucleotide corresponding to the sequence of bases 256–276 was then constructed and used as a primer to obtain the sequence for a further 260 bases. Computer analysis of the nucleotide sequence revealed an open reading frame (ORF) of 378 bp with associated ribosome binding site [18] 186 bp upstream of the *Nde*I site. This ORF was confirmed as that encoding CHM isomerase since the predicted amino acid sequence for residues 2–20 corresponded exactly with the amino terminal amino acid sequence obtained from the pure protein. The methionine at position 1 of the ORF is absent from the purified protein. The molecular mass calculated from the predicted amino acid sequence was 14044 which agrees very well with that of the purified enzyme measured by SDS-PAGE. The nucleotide sequence and predicted amino acid sequence for CHM isomerase are shown in Fig. 4.

When various data bases were searched for other proteins with amino acid sequences similar to that of CHM isomerase no obvious matches were found. However, it will be possible to see whether CHM isomerase has any obvious sequence similarities to the other enzymes of the HPC catabolic pathway once the nucleotide and predicted amino acid sequences for those enzymes becomes available. In turn this might shed some light on how catabolic pathway enzymes have evolved.

Acknowledgements: We are indebted to Dr K.S. Lilley for the amino acid sequencing, Mr John Keyte for oligonucleotide synthesis, Mrs A. Taylor and Ms Z. Master for typing the manuscript. D.I.R. thanks the SERC for a research studentship.

REFERENCES

[1] Barbour, M.G. and Bayly, R.C. (1980) J. Bacteriol. 142, 480–485.
[2] Garrido-Pertierra, A. and Cooper, R.A. (1981) Eur. J. Biochem. 117, 581–584.
[3] Jenkins, J.R. and Cooper, R.A. (1988) J. Bacteriol. 170, 5317–5324.
[4] Dagley, S. (1975) in: Essays in Biochemistry, vol. 11 (Campbell, P.N. and Aldridge, W.N. eds) pp. 81–138, Academic Press, London.
[5] Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, NY.
[6] Ferrer, E. and Cooper, R.A. (1988) FEBS Lett. 52, 155–160.
[7] Gornall, A.C., Bardawill, C.S. and David, M.M. (1949) J. Biol. Chem. 177, 751–756.
[8] Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265–275.
[9] Hames, B.D. (1981) in: Gel Electrophoresis of Proteins – a Practical Approach (Hames, B.D. and Rickwood, D. eds) pp. 1–91, IRL Press, Oxford.
[10] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[11] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[12] Perbal, B. (1988) A Practical Guide to Molecular Cloning, pp. 368–374, Wiley, New York.
[13] Kushner, S.R. (1978) in: Genetic Engineering (Boyer, H.B. and Nicosia, S. eds) pp. 17–23, Elsevier, Amsterdam.
[14] Tabor, S. and Richardson, C.C. (1987) Proc. Natl. Acad. Sci. USA 84, 4767–4771.
[15] Kraft, R., Tardiff, J., Krauter, K.S. and Leinwand, L.A. (1988) BioTechniques 6, 544–547.
[16] Fersht, A. (1985) Enzyme Structure and Mechanism, 2nd edn, Freeman, New York.
[17] Amons, R. (1987) FEBS Lett. 212, 68–71.
[18] Shine, J. and Dalgarno, L. (1975) Nature 254, 34–38.