

Glyoxysomal malate dehydrogenase from watermelon is synthesized with an amino-terminal transit peptide

(isoenzymes/organelle/*Citrullus vulgaris*/polymerase chain reaction)

CHRISTINE GIETL*

Institute of Botany, Technical University of Munich, Arcisstrasse 16, D-8000 Munich 2, Federal Republic of Germany; and Department of Physiology, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

Communicated by Diter von Wettstein, May 11, 1990 (received for review March 5, 1990)

ABSTRACT The isolation and sequence of a cDNA clone encoding the complete glyoxysomal malate dehydrogenase [gMDH; (S)-malate:NAD⁺ oxidoreductase, EC 1.1.1.37] of watermelon cotyledons are presented. Partial cDNA clones were synthesized in a three part strategy, taking advantage of the polymerase chain reaction technology with oligonucleotides based on directly determined amino acid sequences. Subsequently, the complete clone for gMDH was synthesized with a sense primer corresponding to the nucleotide sequence of the N-terminal end of pre-gMDH and an antisense primer corresponding to the adenylation site found in the mRNA. The amino acids for substrate and cofactor binding identified by x-ray crystallography for pig heart cytoplasmic malate dehydrogenase are conserved in the 319-amino-acid-long mature plant enzyme. The pre-gMDH contains an N-terminal transit peptide of 37 residues. It has a net positive charge, lacks a long stretch of hydrophobic residues, and contains besides acidic amino acids a cluster of serine residues. This N-terminal extension is cleaved off upon association with or import into glyoxysomes. It contains a putative AHL topogenic signal for microbody import and has no sequence similarity to the 27-residue-long presequence of the watermelon mitochondrial malate dehydrogenase precursor.

Microbodies, discovered by DeDuve (1), belong to the basic set of membrane-bounded organelles in cells of the plant and animal kingdom. Depending on their enzymatic make-up, different forms of microbodies are distinguished. Glyoxysomes are unique to plant cells and house, in addition to the enzymes for β -oxidation of fatty acids, the glyoxylate cycle, a variant of the tricarboxylic acid cycle. Whereas in the latter acetyl CoA from β -oxidation is fully oxidized to two molecules of CO₂, the glyoxylate cycle converts in each turn two molecules of acetyl CoA into one molecule of succinate and one molecule of oxaloacetate. Formation of oxaloacetate from malate is catalyzed by the glyoxysomal malate dehydrogenase [gMDH; (S)-malate:NAD⁺ oxidoreductase, EC 1.1.1.37]. In germinating watermelon seedlings, the glyoxysomes process the fat stored in the cotyledons into the glyoxylate cycle products; succinate is converted into malate within the mitochondria and is used for gluconeogenesis. As the cotyledons turn into green leaves, another member of the microbody family develops—the peroxisome. It is specialized to import glycolate produced by photorespiration in the chloroplast and to convert it via glyoxylate into glycine. This is shuttled into the mitochondria, where two molecules of glycine condense to serine. Serine returns into the peroxisome and is converted into glycerate, which serves the carbon reduction cycle in chloroplasts. These interactions are visualized in electron micrographs, where glyoxysomes are seen wedged between mitochondria and lipid droplets,

and peroxisomes are seen between mitochondria and chloroplasts (2). As in plants, mammalian peroxisomes contain enzymes involved in the production and degradation of H₂O₂; in trypanosomes, glycolysis is sequestered into microbodies called glycosomes (3). All microbodies studied contain enzymes for β -oxidation and a specific spectrum of other enzymes. It is further characteristic that microbody enzyme activities are also present in other cell compartments. Glyoxysomes as well as mitochondria contain MDH, citrate synthase, and enzymes for β -oxidation (4). The organelle-specific isoenzymes are encoded in the nucleus, translated on cytoplasmic ribosomes, and imported into the genome-less microbodies or the genome-containing mitochondria. A majority of enzymes destined for mitochondrial import are synthesized with a cleavable presequence at the N terminus (5), while most microbody enzymes are synthesized in mature size and apparently targeted with an internal or C-terminal recognition sequence (6). There are important exceptions: the peroxisomal 3-ketoacyl-CoA thiolase of rat liver catalyzing the final step in fatty acid β -oxidation is synthesized with a 26-amino acid N-terminal presequence, whereas the mitochondrial isoenzyme is not made as a larger precursor (7, 8).

The MDH isoenzymes in germinating watermelon seedlings, which have to be sorted into glyoxysomes (gMDH) and mitochondria (mMDH), have the unique property of being synthesized as higher molecular weight precursor proteins (9). From the nucleotide sequence of a cDNA clone encoding the mMDH precursor, an N-terminal presequence of 27 amino acids containing the basic and hydroxylated residues characteristic for mitochondrial envelope transit peptides has been identified (10). Here the isolation and sequence of a cDNA clone encoding the complete gMDH precursor is reported.[†] It has an N-terminal transit peptide of 37 amino acids, which is distinctly different from the presequence of the mitochondrial precursor and no C-terminal extension. The mature polypeptide of gMDH is 319 residues and that of mMDH is 320 residues long and thus essentially of the same length as MDHs from pig, rat, mouse, yeast, and *Escherichia coli*. The catalytic pocket and cofactor binding domains of pig heart cytoplasmic MDH and the functionally related lactate dehydrogenase (LDH) from *Bacillus stearothermophilus* are entirely conserved in the higher plant enzymes.

MATERIALS AND METHODS

Plant Material. Watermelon seeds (*Citrullus vulgaris* Schrad., var. Sugar Baby, harvest 1987) were supplied by

Abbreviations: gMDH, glyoxysomal malate dehydrogenase; mMDH, mitochondrial malate dehydrogenase; PCR, polymerase chain reaction; PTS, peroxisome targeting signal; ss, single stranded; LDH, lactate dehydrogenase.

*Reprint requests should be sent to the address in Munich.

[†]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M33148).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Bayrische Futtersaat (Munich) and germinated under sterile conditions in the dark at 30°C as described (11).

Synthetic Oligonucleotides. DNA fragments were prepared on an Applied Biosystems 380 A DNA synthesizer. They were desalted on Pharmacia nick columns containing Sephadex G-50 and were phosphorylated with ATP or [γ -³²P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq; DuPont/New England Nuclear) using T4 polynucleotide kinase.

Synthesis of Single-Stranded cDNA (ss cDNA). The ss cDNA was synthesized by transcribing polyadenylated mRNA from 2-day-old watermelon cotyledons with avian myoblastosis virus reverse transcriptase (Life Sciences, St. Petersburg, FL) and the combined use of random hexamer deoxynucleotides and dT₁₅ as primers and T4 DNA ligase treatment of cDNA-RNA hybrids as described (10). For the production of clones with the polymerase chain reaction (PCR) corresponding to the 5' end of mRNA, a poly(dA)-tail sequence was attached at the 3' end of the cDNA with terminal deoxynucleotidyltransferase (Pharmacia) (10). The ss cDNA-RNA hybrids were used as templates for the PCR without further purification.

PCR. The reaction mixture (100 μ l) contained 20 ng of template, 1 nM each primer, 0.2 mM each dNTP, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, gelatin (100 μ g/ml), and 2.5 units of *Thermus aquaticus* DNA polymerase (*Taq* polymerase) (Pharmacia). Amplification comprised 25 cycles of 1 min at 94°C for denaturation, 2 min at 55°C for annealing, and 3 min at 72°C for primer extension.

Analysis and Cloning of PCR Amplified Products. PCR products were purified by electrophoresis in a 1% agarose gel. Regions of the gel containing specific products were isolated and the DNA was recovered by using the GeneClean kit (Bio 101, La Jolla, CA). Double-stranded products were either blunt-ended with T4 DNA polymerase for cloning into the *Sma* I site vector pUC13 (12) or digested with restriction endonucleases and ligated into a pGEMEX-1 vector (Promega) digested with identical enzymes. *E. coli* DH5 α (Bethesda Research Laboratories), transformed by a standard protocol (13), was used as a host. Plasmids with gMDH inserts were identified by restriction analysis or DNA sequence determination of mini-prep DNA (14, 15). DNA and protein sequence analysis procedures were performed by using the MicroGenie program (16).

RESULTS

Cloning of the cDNA That Encodes the Precursor for gMDH (pre-gMDH). A full-length cDNA clone for pre-gMDH was obtained by a strategy adapted from refs. 17 and 18 and was successfully used in cloning the cDNA for the precursor of mMDH (pre-mMDH) (10). Sequence amplification of watermelon ss cDNA encoding pre-gMDH by the PCR was carried out in three steps: the central part was synthesized with a sense primer of mixed oligonucleotides encoding the 8 N-terminal amino acids of the mature gMDH subunit (AKG-GAPGF) that are known from direct sequencing (19) of the 31 N-terminal residues (cf. Fig. 1). As antisense primer, the mixed oligonucleotide for a highly conserved sequence (MAYAGA) corresponding to amino acid positions 234–239 in the mature mMDH (cf. Fig. 3) was used. This antisense primer had been successfully used for the isolation of the mMDH cDNA clone and it covers a sequence conserved in mMDHs from watermelon, pig, rat, mouse, and yeast (10). A single PCR fragment was generated and cloned into the *Sma* I site of pUC13. Sequence determination of the resulting plasmid showed a 720-base-pair (bp) insert comprising the first 31 amino acids known from the gMDH N terminus, including the primers in the same reading frame. In particular, Gly-Phe in positions 7 and 8 and the two methionines in

positions 25 and 29 are diagnostic for gMDH and distinguish the cloned piece from the mitochondrial counterpart.

The 3' part of the cDNA clone was generated as described (10, 18): the sense primer was based on the precise nucleotide sequence for the 8 residues (VEAKAGAG) beginning at amino acid position 221 (see Figs. 1 and 3)—i.e., upstream of but close to the end of the cloned central part. The antisense primer isolated included a poly(dT) stretch of 18 nucleotides. The PCR product was cloned into the *Sma* I site of pUC13 and sequenced, revealing a 520-bp insert. The homology with the analogous region of mMDH was 60% on a DNA level and 65% on a protein level; the stop codon was found in nearly the same position, indicating that the gMDH shows no C-terminal extension relative to the mMDH. In spite of the absence of direct protein sequence information, the cloned insert was considered to be the C-terminal part of a MDH isoenzyme different from mMDH. The synthesis of the total clone gave proof that the sequence-faithful sense primer indeed recognized the 3' part of gMDH and not of a cytoplasmic MDH.

To generate the 5' part of the clone, the poly(dA)-tailed ss cDNA was used as a template with an antisense primer based on the nucleotide sequence for amino acids 9–16 (KVAIL-GAA) of the mature subunit (see Figs. 1 and 3), with the precise nucleotide sequence found in the cloned central part; the sense primer included an oligo(dT) tail (10). After cloning of the product into the *Sma* I site of pUC13, a 275-bp insert could be sequenced: it had the gMDH-specific antisense primer and the adjacent code for the first 8 amino acids of the mature gMDH subunit, which are diagnostic for gMDH, in the same reading frame. In addition, an N-terminal transit peptide of 37 amino acids was found.

In a final step, the complete cDNA clone for the pre-gMDH was synthesized by PCR: the sense primer was based on the first 21 bases of the coding region plus the T preceding the initiation codon (Fig. 1, bases 69–90); a *Not* I restriction endonuclease cleavage site was added at the 5' end for cloning purposes: 5'-GGCGGCCGCC-TATGCAGCCGAT-TCCGGATGTT-3'. The antisense primer was complementary to bases 1310–1334 (Fig. 1) in front of the poly(dA) tail and included a *Sal* I site at the 5' end: 5'-GGTCGAC-CAT-AAAAGGTACTCTAATCCAACCC-3'. The PCR product was cut with *Not* I and *Sal* I and was cloned into the appropriate sites of the transcription vector pGEMEX-1 (Promega).

Sequence Analysis of the cDNA Encoding gMDH. A 1334-bp nucleotide sequence (Fig. 1) was determined by the sequencing strategy given in Fig. 2. One open reading frame encoding a polypeptide with 356 amino acids was identified. The position of the N-terminal alanine residue of the mature enzyme shows that the gene encodes a 37-amino-acid-long N-terminal transit peptide. The first 31 residues known from sequencing of the mature enzyme are found in-frame. In particular, the octapeptide AKGGAPGF is in front of the classical MDH protein sequence, which is unique to the N terminus of the mature watermelon enzyme and at the same time distinguishes the gMDH from the mitochondrial counterpart (19). A stop codon at base 1138 revealed a total size for the mature subunit of 319 amino acids, which was only 1 amino acid shorter than mMDH from watermelon and 2 amino acids shorter than MDHs from pig, rat, mouse, yeast, and *E. coli* (10). The amino acids glutamic acid, leucine, serine, isoleucine, lysine, glycine, and phenylalanine (positions 305, 306, 309, 310, 312, 313, and 316 in Fig. 3) are highly conserved among all organelle-bound MDHs and provide convincing evidence for a genuine C terminus of MDH. No inconsistencies were encountered between the sequences determined for the total clone and for the separately synthesized PCR products of the central part, 3' part, and 5' part of the cDNA. It is therefore concluded that gMDH does not

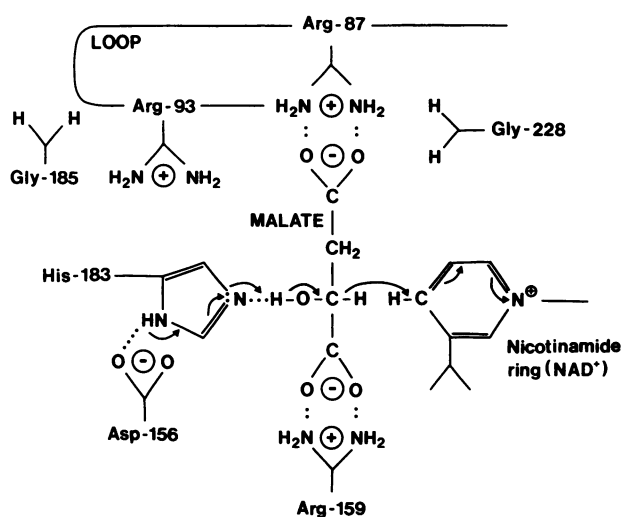


FIG. 4. Model for the active site pocket and catalytic mechanism of watermelon gMDH adapted from ref. 41.

change into a site with malate substrate specificity and high MDH activity. Using the coordinates of the M4 isoenzyme for the pig LDH-NADH-oxamate complex and the catalytic mechanism of LDH, Wilks *et al.* have presented a model for the active site environment of pyruvate in the native enzyme-NADH-substrate complex (figure 1 A and B in ref. 41). By site-directed mutagenesis of a glutamine into an arginine residue, the substrate binding pocket could be enlarged to fit oxaloacetate instead of pyruvate. Indeed, the amino acid chains predicted from the model of Wilks *et al.* are found in the primary structure of the watermelon gMDH and mMDH and are marked in Fig. 3. By adapting the drawing in ref. 41, the active site environment and catalytic mechanism of the higher plant MDHs are given in Fig. 4. Arg-87 in the loop (residues 83-95) that closes over the active site upon binding of the coenzyme and substrate forms an ion pair with the 4-carboxyl group of the malate substrate. The latter is oriented in the pocket with its 1-carboxyl group by ion pairing to Arg-159. The third arginine involved, Arg-93, polarizes the carbonyl bond formed when His-183 receives the proton in the reduction of NAD⁺ and oxidation of malate to oxaloacetate. Asp-156 is considered to stabilize the protonated His-183 by ion pair formation. In MDH enzymes Gly-185 and Gly-228 are rigorously conserved, while in LDH enzymes an equally strong conservation is observed for aspartate and threonine at these positions, respectively. This conceivably reduces the size of the catalytic pocket in the LDH enzymes.

In the electron pair movement during malate oxidation, as portrayed in Fig. 4, nicotinamide accepts a hydride ion from malate and His-183 receives a proton from the 2-hydroxy group of malate. Wilks *et al.* (41) have shown that the exchange of Gln-102 into an arginine in the loop of the LDH enzyme of *B. stearothermophilus* changes the K_m value for oxaloacetate from 1.5 mM to 0.06 mM and the kinetic constant for catalysis k_{cat} changes from 6.0 to 250 s⁻¹, which corresponds to that for pyruvate of the native enzyme. The K_m values for the watermelon gMDH and mMDH toward oxaloacetate have been determined to be 0.18 and 0.15 mM, respectively (9). In conclusion, the higher plant MDHs use the same amino acid chains in catalytic pocket as MDH enzymes in mammals, yeast, and bacteria. Most likely, Arg-87 is decisive for selecting malate as substrate.

I thank Prof. D. von Wettstein, Dr. K. K. Thomsen, and candidate scientist O. Olsen for support and encouragement. I am indebted to Gerda Krämer, Nina Rasmussen, and Ann-Sofi Steinholtz for tech-

nical assistance and graphic work. This work was supported by a grant from the Deutsche Forschungsgemeinschaft.

1. DeDuve, C. (1969) *Proc. R. Soc. London Ser. B.* **173**, 71-83.
2. DeDuve, C. (1983) *Sci. Am.* **248**, 52-62.
3. Douglas, M. G., McCammon, M. T. & Vassarotti, A. (1986) *Microbiol. Rev.* **50**, 166-178.
4. Beevers, H. (1979) *Annu. Rev. Plant Physiol.* **30**, 159-193.
5. Hartl, F.-V., Pfanner, N., Nicholson, D. W. & Neupert, W. (1989) *Biochim. Biophys. Acta* **988**, 1-45.
6. Borst, P. (1989) *Biochim. Biophys. Acta* **1008**, 1-13.
7. Hijikata, M., Ishii, N., Kagamiyama, H., Osumi, T. & Hashimoto, T. (1987) *J. Biol. Chem.* **262**, 8151-8158.
8. Arakawa, H., Takiguchi, M., Amaya, Y., Nagata, S., Hayashi, H. & Mori, M. (1987) *EMBO J.* **6**, 1361-1366.
9. Gietl, C. & Hock, B. (1987) in *Isoenzymes: Current Topics in Biological and Medical Research*, eds. Ratazzi, M. C., Scandalios, I. G. & Whitt, G. S. (Liss, New York), Vol. 16, pp. 175-192.
10. Gietl, C., Lehnerer, M. & Olsen, O. (1990) *Plant Mol. Biol.*, in press.
11. Hock, B. (1969) *Planta* **85**, 340-350.
12. Vieira, J. & Messing, Y. (1982) *Gene* **19**, 259-268.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152**, 232-238.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
16. Queen, C. & Korn, L. J. (1984) *Nucleic Acids Res.* **12**, 581-599.
17. Saiki, R. K., Gelford, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Ehrlich, H. A. (1988) *Science* **239**, 487-490.
18. Loh, E. Y., Elliot, J. F., Cwirla, S., Lanier, L. L. & Davis, M. M. (1989) *Science* **243**, 217-243.
19. Gietl, C., Lottspeich, F. & Hock, B. (1986) *Planta* **559**, 555-558.
20. Walk, R.-A. & Hock, B. (1978) *Biochem. Biophys. Res. Commun.* **81**, 636-643.
21. Riezman, H., Weir, E. M., Leaver, C. J., Titus, D. E. & Becker, W. M. (1980) *Plant Physiol.* **65**, 40-46.
22. Gietl, C. & Hock, B. (1982) *Plant Physiol.* **70**, 483-487.
23. Dingwall, C. & Laskey, R. A. (1986) *Annu. Rev. Cell Biol.* **2**, 367-390.
24. Haeuptle, M.-T., Flint, N., Gough, N. M. & Dobberstein, B. (1989) *J. Cell Biol.* **108**, 1227-1236.
25. van Figura, K. & Haslik, A. (1986) *Annu. Rev. Biochem.* **55**, 167-193.
26. von Heijne, G., Steppuhn, J. & Herrmann, R. G. (1989) *Eur. J. Biochem.* **180**, 535-545.
27. Schmidt, G. W. & Mishkind, M. L. (1986) *Annu. Rev. Biochem.* **55**, 879-912.
28. Osumi, T. & Hashimoto, T. (1984) *Trends Biochem. Sci.* **9**, 317-319.
29. Opperdoes, F. R., Baudhuin, P., Coppens, J., DeRos, C., Edwards, S. W., Weijers, P. J. & Misset, O. (1984) *J. Cell Biol.* **98**, 1178-1184.
30. Keller, G.-A., Gould, S., Deluca, M. & Subramani, S. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3264-3268.
31. Wierenga, R. K., Swinkels, B., Michels, P. A. M., Osinga, K., Misset, O., Van Beeumen, J., Gibson, W. C., Postma, J. P. M., Borst, P., Opperdoes, F. R. & Hol, W. G. J. (1987) *EMBO J.* **6**, 215-221.
32. Lake, J. A., Delacruz, V. F., Ferreira, P. C. G., Morel, C. & Simpson, L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4779-4783.
33. Misset, O., Bos, O. J. M. & Opperdoes, F. R. (1986) *Eur. J. Biochem.* **157**, 441-453.
34. Swinkels, B. W., Evers, R. & Borst, P. (1988) *EMBO J.* **7**, 1159-1165.
35. Gould, S. J., Keller, G.-A., Hosken, N., Wilkinson, J. & Subramani, S. (1989) *J. Cell Biol.* **108**, 1657-1664.
36. Okada, H., Ueda, M., Sugaya, T., Atomi, H., Mozaffar, S., Hishida, T., Teranishi, Y., Okazaki, K., Takechi, T., Kamiryo, T. & Tanaka, A. (1987) *Eur. J. Biochem.* **170**, 105-110.
37. Comai, L., Baden, S. C. & Harada, J. J. (1989) *J. Biol. Chem.* **264**, 2778-2782.
38. Greenler, J. McC., Sloan, J. S., Schwartz, B. W. & Becker, W. M. (1989) *Plant Mol. Biol.* **13**, 139-150.
39. Schram, A. W., Strijland, A., Hashimoto, T., Wanders, R. J. A., Schutgens, R. B. H., VandenBosch, H. & Tager, I. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6156-6158.
40. Birktoft, J. J. & Banaszak, J. (1983) *J. Biol. Chem.* **258**, 472-482.
41. Wilks, H. M., Hart, K. N., Feeny, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R. & Holbrook, J. J. (1988) *Science* **242**, 1541-1544.