

Mitochondrial malate dehydrogenase from watermelon: sequence of cDNA clones and primary structure of the higher-plant precursor protein

Christine Gietl¹, Michael Lehnerer¹ and Ole Olsen²

¹Lehrstuhl für Botanik, Technische Universität München, Arcisstr. 21, D-8000 München 2, FRG (author for correspondence); ²Department of Physiology, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

Received 28 December 1989; accepted in revised form 20 February 1990

Key words: isoenzymes, cotyledons, *Citrullus vulgaris*, polymerase chain reaction

Abstract

The isolation and sequence of a cDNA clone encoding the complete mitochondrial malate dehydrogenase (mMDH) of watermelon cotyledons is presented. Taking advantage of the polymerase chain reaction technology partial cDNA clones from the central part, the 3' part and the 5' part of the mRNA were obtained with oligonucleotides based on directly determined amino acid sequences. Subsequently, two complete cDNA clones for mMDH were synthesized with a sense primer corresponding to the nucleotide sequence of the amino terminal end of pre-mMDH and two antisense primers corresponding to the major alternative adenylation sites found in the mRNA.

The amino acid residues for substrate and cofactor binding identified by X-ray crystallography for pig heart cytoplasmic MDH are conserved in the 320 amino acid long mature higher-plant mMDH. A presequence of 27 amino acids is present at the amino terminal end of the precursor protein.

Introduction

Plant tissues contain multiple molecular forms of malate dehydrogenase (MDH, EC 1.1.1.37), and these are located in different cell compartments [7, 16]. The three cytosolic MDH isozymes (I, II and IV) are involved in the malate shuttle and gluconeogenesis. In guard cells of stomates cytosolic MDH (probably isozyme I) provides the malate which acts as counterion to the imported

potassium in the osmotically effected opening of stomates [5, 25, 30].

The mitochondrial malate dehydrogenase (mMDH) oxidizes malate to oxaloacetate in the tricarboxylic acid cycle. A fifth form of malate dehydrogenase functions in the glyoxylate cycle of the glyoxysomes (gMDH). Fat stored in the cotyledons of seeds such as those of watermelon (*Citrullus vulgaris* Schard.) is mobilized upon ger-

mination and the fatty acids degraded by β -oxidation to acetyl-CoA which is condensed with glyoxylate to malate by malate synthase. Malate is oxidised to oxaloacetate by gMDH. This cycle is required to convert fatty acids into carbohydrates during germination. Upon development of the cotyledons into leaves the glyoxysomes are converted into peroxisomes which cooperate with the chloroplast and the mitochondrion in photorespiration, permitting the recycling of glycolate into glycerate-3-phosphate and CO_2 . In the peroxisomes MDH provides NADH for the reduction of β -hydroxypyruvate to glycerate.

These various isoenzymes are encoded in genes of the nucleus and synthesized on cytoplasmic ribosomes. The mitochondrial and glyoxysomal isoenzymes are made as higher molecular weight precursor proteins with extra sequences which have been estimated by gel electrophoretic methods to 3 kDa and 8 kDa in size. During import of the pre-proteins into their respective organelle the putative targeting peptides are cleaved off and subsequently the holoenzyme comprising two identical subunits is assembled [7].

Molecular cloning of cDNA to the messenger RNAs encoding the higher plant mitochondrial and glyoxysomal MDH isoenzymes and determination of their nucleotide sequences is carried out with the aim of identifying the size and location of the organelle targeting sequences within the precursor polypeptide. Subsequently, the amino acid sequences responsible for the differential import of the isozyme precursors into glyoxysomes and mitochondria will be studied with *in vitro* translation from the cDNA clones.

In the present paper we report on the isolation and sequencing of a cDNA clone encoding the complete watermelon mitochondrial MDH. Taking advantage of the polymerase chain reaction (PCR) technology [28] partial cDNA clones from the central part, the 3' part and the 5' part of the mRNA were obtained. Thereafter, two complete cDNA clones for mMDH were synthesized with a sense primer corresponding to the nucleotide sequence of the amino terminal end of the pre-mMDH and two antisense primers correspond-

ing to the major alternative adenylation sites found in the mRNA.

The catalytic pocket and cofactor binding domains of the cytoplasmic malate dehydrogenase of the pig heart have been determined by X-ray crystallography [1, 2, 3] and critical residues for substrate binding identified by site-directed mutagenesis of the functionally related lactate dehydrogenase from *Bacillus stearothermophilus* [38]. These amino acid residues are entirely conserved in the 320 amino acid long mature higher plant mMDH. A presequence of 27 amino acids is present at the amino terminal end of the precursor protein.

Materials and methods

Plant material

Watermelon seeds (*Citrullus vulgaris* Schard. cv. Sugar Baby, harvest 1987) were supplied by Bayerische Futtersaat, Munich, FRG and germinated under sterile conditions in the dark at 30 °C as described previously [13].

Synthesis of single-stranded cDNA (ss-cDNA)

Polydenylated mRNA was isolated from 2-day-old cotyledons by extraction of total RNA with phenol/chloroform and by affinity-binding it twice on an oligo-(dT)-cellulose column (Boehringer Mannheim, FRG) for removal of most of the rRNA [37]. The ss-cDNA was synthesized with avian myoblastosis virus reverse transcriptase (Life Sciences, USA) according to Rutledge *et al.* [26] with the combined use of random hexamer deoxynucleotides and (dT)₁₅ as primers and T4-DNA ligase treatment of cDNA/RNA hybrids, a procedure promoting the production of full length ss-cDNA molecules. These ss-cDNA/RNA-hybrids were used as templates for the polymerase chain reaction (PCR) without further purification (20 ng/asay). For the production of PCR clones corresponding to the 5' end of mRNA a poly(dA) tail sequence was attached at

the 3' end of the cDNA with terminal deoxynucleotidyl transferase (Pharmacia, Sweden). In the reaction 200 ng DNA was treated with 40 units transferase in a buffer containing 40 mM cacodylate pH 6.8, 10 mM MgCl₂, 0.3 mM ZnCl₂, 1 mM dATP for 1 hour at 37 °C. The ss-cDNA/RNA-hybrids were used as templates for the polymerase chain reaction without further purification (20 ng/assay).

Primer synthesis and processing

Synthetic oligonucleotide primers were prepared on an Applied Biosystems 380A DNA synthesizer according to the procedure described by the manufacturer. Oligonucleotides were desalted on Pharmacia Nick columns containing Sephadex G-50 and phosphorylated with ATP or [γ -³²P]ATP (3000 Ci/mmol, DuPont-New England Nuclear, USA) using T4 polynucleotide kinase.

Analysis and cloning of PCR amplified products

PCR products were separated by electrophoresis in agarose gels. Normal 1% (w/v) SeaKem GTG agarose (FMC BioProducts, USA) was used when amplified fragments were expected to be longer than approximately 800 bp whereas 4% (w/v) NieSieve agarose gel (FMC BioProducts, USA) was used for electrophoresis of smaller DNA fragments. Regions of the gel containing specific products were isolated, and the DNA was recovered by using the GeneClean kit (Bio101, USA) according to the manufacturer's recommendations. Double-stranded products were either blunt-ended with T4-DNA polymerase for cloning into the *Sma* I cloning site of vector pUC13 [36] or digested with restriction endonucleases and ligated into a pGEMEX-1 vector (Promega, USA) digested with identical enzymes. *Escherichia coli* DH5 α (*F*, *endA1*, *hsdR17* (*r_k⁻ m_k⁺*), *supE44*, *thi-1*, λ - , *recA1*, *gyrA96*, *relA1*, Φ 80*dlacZ* Δ *m15*) was supplied by BRL (Bethesda Research Laboratories, USA) and transformed by

a standard protocol [18]. Plasmids with mMDH gene inserts were identified by colony hybridizations [18], restriction analysis of mini-prep DNA [11] or DNA sequence determination of mini-prep DNA [11, 29] with Sequenase (United State Biochemicals, USA). DNA and protein sequence analysis procedures were performed using the MicroGenie program [24]

Results

Cloning of the cDNA that encodes the precursor for mitochondrial malate dehydrogenase (pre-mMDH)

N-terminal amino acid sequences were obtained from purified watermelon mMDH [8] and from two fragments after cleavage with cyanogen bromide (CNBr). Written in the single-letter amino acid code the three polypeptides had the following sequences:

mMDH: NH₂-ATESVPERKVAVLGAAGGIGQPLALLMKL-COOH;
CNBr-I: NH₂-[M]KLNPLVSKLALYDIAGXPGVAAD-COOH;
CNBr-II: NH₂-[M]AYAGALFADACKLGLNGVPDVVECSFV-COOH.

Alignment with the amino acid sequence deduced from a cDNA clone of rat pre-mMDH [9] suggested that CNBr-I is attached directly to the N-terminal peptide with an overlap of 3 amino acids, whereas the methionin of CNBr-II is in a position corresponding to residue 234 of the mature rat mMDH protein. Sequence amplification of watermelon sscDNA encoding pre-mMDH by the polymerase chain reaction was carried out in three steps according to a strategy adapted from [17, 28] and outlined in Fig. 1.

The central part was synthesized with mixed oligonucleotides as primers: the sense primer was based on the 6 N-terminal amino acids ATESVP of the mature subunit (P1: 256-fold degeneration); the antisense primer was based on the first 6 amino acids MAYAGA of CNBr-II (P2: 64-fold degeneration). The first nucleotides in each oligonucleotide were included to generate restriction endonuclease cleavage sites, although these were not used in the subsequent steps. The reaction mixture (100 μ l) contained 20 ng template DNA, 10 μ M of each primer, 0.2 mM of

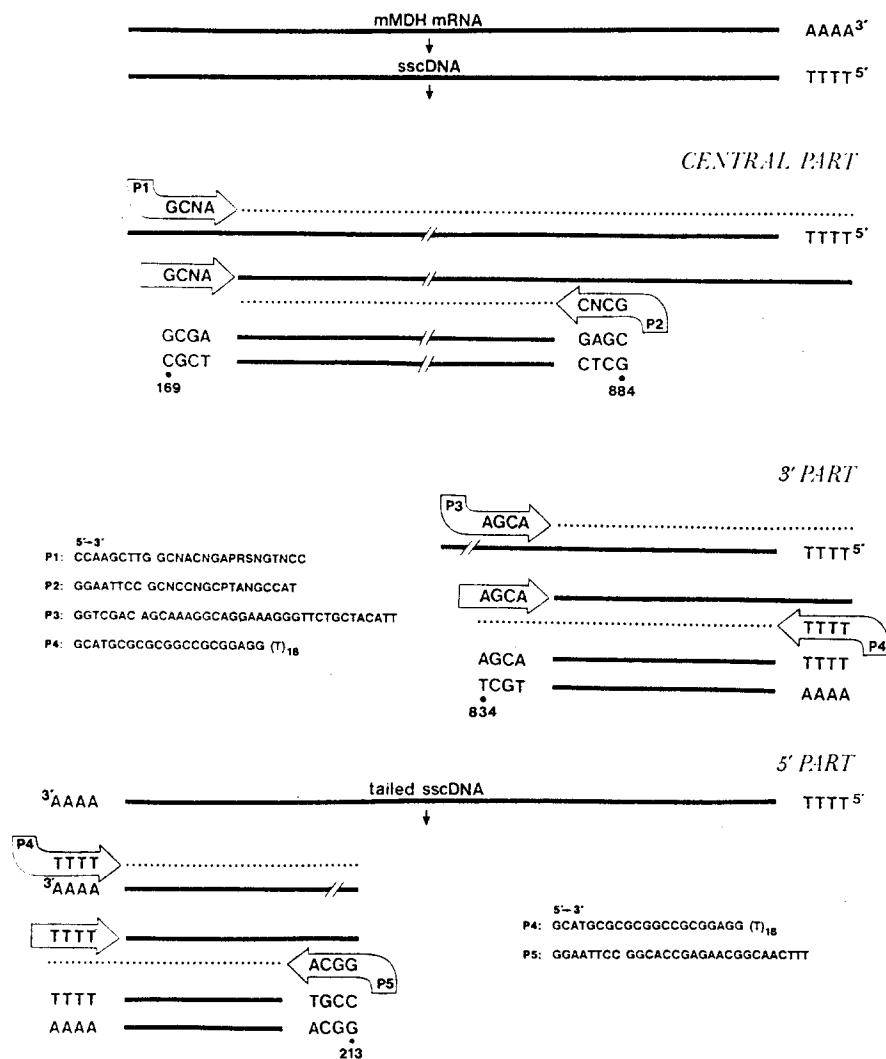


Fig. 1. Strategy for amplification of three overlapping cDNA parts by the polymerase chain reaction. mRNA encoding the mitochondrial pre-malate dehydrogenase (pre-mMDH) was reverse-transcribed into single stranded cDNA (ss-cDNA). The central part of the coding region was amplified with sense primer 1 (P1) based on the 6 N-terminal amino acids of the mature protein and antisense primer 2 (P2) based on the 6 N-terminal amino acids of cyanogen bromide (CNBr) fragment II. This yielded the double-stranded cDNA fragment spanning nucleotide 169 to 884 in the final sequence depicted in Fig. 3. The 3' part of the cDNA clone was generated with sense primer 3 (P3) starting at nucleotide 834 and a poly(dT) antisense primer (P4). In the third step P4 was used as sense primer and primer P5, starting at nucleotide 213, as antisense primer. N = degeneration in primer with all four nucleotides; P = degeneration with the two purine dA,dG nucleotides; R = degeneration with dA,dT; S = degeneration with dC,dG.

each dNTP, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 100 µg/ml gelatin and 2.5 units *Taq* polymerase (Perkin Elmer Cetus, USA). 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. A single fragment was generated and purified on a 4% NuSieve agarose gel and cloned into the *Sma* I site of pUC13. Sequence determination [11, 29] of the resulting plasmid showed a 732 bp insert with the mMDH-N-terminal and CNBr-I codons in the same reading frame.

The 3' part of the cDNA clone was generated as described [17] with P3 as sense primer. It is based on the nucleotide sequence for the 10 residues found one amino acid upstream of CNBr-II (AKAGKGSATL) and the codon usage in the cloned central part. Primer 4 (P4) which includes a poly(dT) stretch of 18 nucleotides was employed as antisense primer. Reaction conditions were as above. The PCR products were separated on a 4% NuSieve agarose gel. It is to be noted that the sense primer had to synthesize the specific sense strand from the sscDNA representing the antisense strand before the unspecific antisense primer could use it as a template. Two sharp bands of about 550 bp and 630 bp were generated, cloned into the *Sma* I site of pUC13 and transformed into *E. coli* cells. Colony hybridization [18] of 1000 clones with P3 as a probe revealed that virtually all the clones contained the sequence of the specific sense primer. Sequence analysis of 26 different clones with the long and short inserts showed that 23 clones contained the P3-primer sequence followed by the nucleotides for the 28 amino acids of CNBr-II in the same reading frame. The difference in length was due to an insert in the 3' non-coding region directly in front of the poly(A) tail. Two clones (530 bp insert) contained the P3 primer sequence followed by a so far unknown sequence, which neither was homologous to the rat pre-mMDH [9]. One clone (630 bp insert) had stop codons very close to the primer sequence in the same reading frame. These clones were discarded.

In order to generate the 5' part of the clone poly(dA)-tailed sscDNA was used as a template (20 ng/assay) with P5 as antisense primer which is based on the nucleotide sequence for amino acids 9–15 (KVAVLGA) of the mature subunit and with the oligo(dT)-tailed P4 as sense primer. After 25 PCR cycles the products were separated on a 4% NuSieve agarose gel: in contrast to the procedure for generation of the 3' part of the clone the 'unspecific' P4-primer had to assemble first the template strand for the 'specific' P5 primer. This resulted in a smear of linearly amplified products concealing the 5' part of the cDNA. The agarose containing the DNA located above

the 75 bp marker (which corresponds to the length of the primers plus the code for the 8 known N-terminal amino acids) was cut from the gel, and approximately 5% of the DNA was directly reamplified [17]. After 25 PCR cycles the products were separated on a 4% NuSieve gel: three discrete fragments approximately 270 bp, 180 bp and 150 bp in length were obtained as ethidium bromide stained faint bands. 5% of each band was again cut from the gel and directly reamplified in order to achieve higher yield. After cloning the re-amplified DNA into the *Sma* I site of pUC13 the resulting plasmids were sequenced: only the 270 bp insert contained the 'specific' P5 primer sequence and the adjacent code for the first 8 amino acids of the mature mMDH subunit in the same reading frame. In the 180 bp and 150 bp inserts the sequences of the PCR primers were connected by several nucleotide sequences, lacking the expected eight codons for the N-terminal peptide. These clones were discarded.

In the final step the two complete cDNA clones for pre-mMDH were synthesized by the polymerase chain reaction: the sense primer was based on the first 19 bases of the coding region plus the A preceding the initiation codon (Fig. 3: base 87–106):

5'-GGCGGCCGCAATGAAGGCCTCAATTCTCA-3'

A *Not* I-restriction endonuclease cleavage site (underlined) was added at the 5' end for cloning purposes. Two different antisense primers were synthesized corresponding to the two different native 3'-non-coding regions identified among the PCR-products. One primer was complementary to bases 1316–1337 (Fig. 3) in front of the poly(dA)-tail of the short mRNA:

5'-GGTCGACAATCCCTCAGAAAACGTGCCAT-3'

The other was complementary to bases 1392–1413 for the long mRNA:

5'-GGTCGACCCAGTTCAAATGGTCGGACGAT-3'

A *Sal* I restriction endonuclease cleavage site (underlined) was again added for cloning purposes. The reaction was carried out with cycles of 1 min at 94 °C, 2 min at 55 °C and

4 min at 72 °C. After 25 cycles an aliquot of the PCR product was checked on a 1% agarose gel. The rest was cut with *Not* I and *Sal* I and cloned into the transcription vector pGEMEX-1.

Sequence analysis of the cDNA encoding mitochondrial malate dehydrogenase

A 1413 bp nucleotide sequence (Fig. 3) was determined using the sequencing strategy given in Fig. 2. One open reading frame encoding a polypeptide with 347 amino acids was identified. The position of the N-terminal alanine residue of the mature enzyme shows that the gene encodes a 27 amino acid long presequence, which is usual for enzymes targeted into the mitochondrial matrix. In front of the classical MDH protein sequence is found the octapeptide ATESVPER which is unique to the N-terminus of the mature watermelon enzyme and also distinguishes the mitochondrial MDH from the glyoxysomal counterpart [8]. Four internal methionine residues are present in the deduced amino acid sequence, which predicts the purified enzyme to yield five fragments upon cyanogen bromide cleavage. Partial amino acid sequences of all five fragments were determined by automated Edman degradation and could be aligned to the deduced amino acid sequence as presented in Fig. 3. One CNBr fragment is identical in sequence to the N-terminal amino acid sequence of the mature enzyme. For the adjoining CNBr fragment (designated as CNBr-I) the sequencer provided 23 residues which were identical to the residues deduced from the cDNA clone. This was also the case for the

28 residues determined in the C terminally located CNBr fragment (CNBr-II). The other two CNBr fragments could not be separated electrophoretically. As a consequence, the sequencer delivered two amino acid residues at each of the 17 steps, which however, could be assigned to the two peptides underlined in Fig. 3. One starts at nucleotide position 439 and the other at position 532. An analysis of codon usage did not reveal a significant bias.

As described above, the sequence covering nucleotide positions 87 to 1337 was determined from two clones produced independently by the polymerase chain reaction. Identical sense primers were used, but one clone was generated with an antisense primer corresponding to positions 1392 to 1413 immediately preceding the mRNA poly(A)⁺ tail of the long transcript and the other with the antisense primer complementary to positions 1316 to 1337 which adjoin the mRNA poly(A)⁺ tail of the major short transcript. Identical nucleotide sequences were obtained up to position 1337 in both clones and no inconsistencies were encountered to the sequences determined for the separately synthesised PCR products of the central part, 3' part and 5' part (Fig. 1) of the cDNA. It is therefore concluded that two different sites for attaching the polyadenyl tail are used for the transcript of one gene encoding the mMDH.

Among the sequenced clones with inserts of the short PCR product of the 3' part of the transcript five additional sites for the attachment of the poly(A)⁺ tail were found at positions 1313, 1315, 1320, 1330 and 1334 (marked with short arrows in Fig. 3). The 3' untranslated region does not

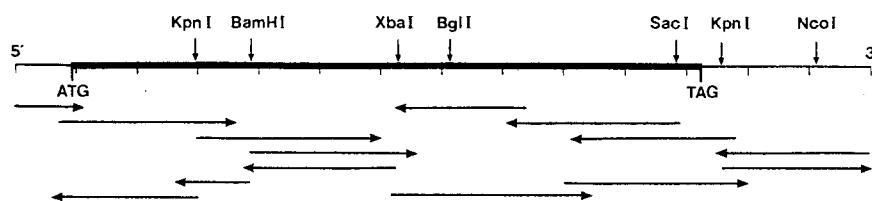


Fig. 2. Restriction endonuclease site map and sequencing strategy for cDNA encoding pre-mMDH. The open reading frame is indicated with a thick line, the untranslated 5' and 3' sequences with a thin line. Direction and extent of (nucleotide) sequencing is given by thin arrows. The vertical marks in the restriction map are placed at a distance of 100 bp.

Discussion

Comparison of the mitochondrial malate dehydrogenase from the higher plant and other organisms

The deduced molecular weight of the preform of the subunit of watermelon mitochondrial malate dehydrogenase is 36200 and that of the mature form 33200. Apparent molecular weights deter-

mined by gel electrophoresis gave values of 41000 and 38000, respectively [6].

Alignment of the amino acid sequence of the watermelon mMDH with porcine mMDH [3] and mMDH from rat [9], mouse [33], yeast [35] as well as the MDH from *E. coli* [20, 22] is presented in Fig. 4. Only those residues not identical with corresponding residues in the watermelon enzyme are shown. For optimal alignment of the

		20		40		60
	NNNN	OO	Q	NNNN	N	OO
W-mMDH	ATESVPERKVAVLGAAGGIGQPLALLMKLN-PLVSKLALYDIA-GTPGVAADVGHVNTRS					
P-mMDH	-----A	S	S L - S	R T	-H	LS IE A
R-mMDH	-----A	S	S L - S	R T	-H	LS IE A
M-mMDH	-----A	S	S L - S	R T	-H	LS IE A
Y-mMDH	-----Y T	G	S L	HK- TD R	LK- AK	T LS IP N
E- MDH	-----M		A L TQL SG E S		PV	V LS IP AV
		80		100		120
		NNNN	NC	C	N	N
W-mMDH	EVTGYVGEEQLGK--ALEGSDVVIIPAGVPRKPGMTRDDLFNINAGIVKSLCTAIKYP					
P-mMDH	T K L P --LPDC K C V			T T	AT TA C QH	
R-mMDH	N K L P --LPDC K C V			T T	AT TA C QH	
M-mMDH	N K L P --LPDC K C V			T T	AT TA C QH	
Y-mMDH	V K FTP PD LNN KDT M L			A S	RD AA T ESA	
E- MDH	KIK FS DA--TP-	A	L S A	D S	V	N VQV T
		140		160		180
	NNN		N	NC	OO	OO
W-mMDH	NALINMISNPFVNSTVPIAAEVFKAGTYDEKKLFGVTTLDVRAKTFYA-GK-ANVPVAE					
P-mMDH	D M CI I T H V NPN I		I	NA V EL GLD-	- R	
R-mMDH	E M CI I T H V NPN I		I	N V EL GLD-	- R	
M-mMDH	E MCI I T H V NPN I		I	N V EL GLD-	- R	
Y-mMDH	A LV V Q L NK V NP		SI	AR ISEVENTD-	TQ	
E- MDH	K C GI T T A L V KN		II SN	V EL GKQ-	-G	
		200		220		240
	C C			OO	OO	NN N OO
W-mMDH	VNVPIVGGHAGITILPLFSQATPRANLSDDTIVALTQDGGTEVVEAKAGKGSATLSM					
P-mMDH	S K I I C KVDFFQ QLST G I EA		K A			
R-mMDH	K I I C KVDFFQ QLST G I EA		K A			
M-mMDH	K I I C KVDFFQ QLST G I EA		K A			
Y-mMDH	R- N S I I - NHKLM KRHE IH I F D		K N A			
E- MDH	E S V L -V GVSFTEQEVAD T I NA		G			
		260		280		300
	Q					
W-mMDH	AYAGALFADACLKGLNGVDPVVECSFVQS--TVTEL-PPFASKVKLGKNGVESVLDLGLPL					
P-mMDH	R VFSLVDAM KEG I K --QE DC- Y STPLL		K I KN GI KI			
R-mMDH	R VFSLVDAM KEG --KE C-TY STPLL		K L KN GI KI			
M-mMDH	R VFSLVDAM KEG --KE C-TY STPLL		K L KN GI KI			
Y-mMDH	H K N V SGFK ER I P D PLFKS GIE P T PD I KIHPI E					
E- MDH	GQ A R GLSLVRALQGEQG AYVEG--DGQYA-R SQPLL		ERKSI T			
		320				
W-mMDH	SDFEKEGIEKLPKASIEKGIQFANAN		Watermelon			
P-mMDH	P EKM AEA I K EE VKNMK		Pig			
R-mMDH	TP EKM AEA I K ED VK MK		Rat			
M-mMDH	TP EKM AEA I K ED VK MK		Mouse			
Y-mMDH	TSE E MLQKC ET KN VN V- SK		Yeast			
E- MDH	A QNAL GMLDTL KD AL EE V --K		E.coli			

Fig. 4. Comparison of the primary structure of the watermelon mitochondrial malate dehydrogenase with the corresponding enzymes from pig [3], rat [9], mouse [33] and yeast [35]. Also given is the *E. coli* MDH amino acid sequence [20]. N = amino acids attributed to NAD binding [3, 35]; C = amino acids participating in the substrate binding pocket [1, 38]; Q = amino acids located at the interphase of the subunits. Only amino acid residues which differ from those found in the higher plant enzyme are written for the other five sequences, an empty space signifying identity. Gaps introduced in the sequences for maximal alignment are indicated by a hyphen.

320 amino acids 9 gaps had to be introduced which gives a numbering of 329 positions. These are used in the comparisons to be described in the present paper.

As is evident from the figure the three mammalian enzymes differ among each other only by a few residues, identity being greater than 98%. Comparison of the watermelon mMDH with the porcine, rat and mouse mMDH gives 59–60% overall identity of residues. The yeast mMDH shares 55% identical residues with watermelon mMDH and sequence identity of the latter with the *E. coli* enzyme amounts also to 55%. The overall identity drops to 20% when one compares the sequence of MDH from the thermophilic bacterium *Thermus flavus* [23]. It has been pointed out that the MDH from this bacterium is more closely related to the mammalian cytoplasmic MDH with 52% overall identity of residues [14]. It will be interesting from an evolutionary point of view to find out if this is also the case for the higher plant cytosolic MDH. The close relationship between the mammalian mitochondrial MDH with *E. coli* MDH on the one hand and the similarity between the cytosolic MDH with the *Thermus flavus* on the other hand is taken as evidence that the two prokaryotic primordial genes have arisen by duplication prior to the origin of mammals [19].

The mature form of the watermelon enzyme is distinguished from the other enzymes by an amino terminal extension of 7 amino acids. Like in the mammalian enzyme the N-terminal amino acid is an alanine. Amino acid conservation among all enzymes is considerably higher in the amino terminal half of the polypeptide chain than in the carboxy terminal half. Epitopes with very strong conservation are found in the following positions: 9–26, 85–101, 127–134, 139–144, 154–160, 183–191 and 223–247.

The functional epitopes of the plant enzymes

An instigation to determine the primary structure of the mitochondrial malate dihydrogenase (mMDH) in a higher plant such as the watermelon

is provided by the prominent conservation of the amino acids of the nicotinamide-adenine-dinucleotide (NAD) binding sites and the residues involved in catalysis among the mammalian and yeast mMDH enzymes. X-ray crystallography of the porcine heart cytoplasmic MDH [2, 3] assigned four β sheets and one α helix as the core structure to which the cofactor NAD(H) binds. This core structure is highly conserved in the mitochondrial MDH of pig, other mammals and yeast [3, 35]. Furthermore, consistent similarities are also evident if other 2-hydroxy acid dehydrogenases are studied. Especially a comparison with the most extensively explored lactate dehydrogenase NAD binding domain and catalytic pocket [3, 19] reveals residues of critical importance. An analysis of the primary structure of the watermelon mMDH reveals that also the higher plant enzyme employs these residues.

In Fig. 4 the amino acid residues attributed to the binding of NAD(H) are labelled with N. The substrate binding pocket of the cytoplasmic porcine MDH has been determined by Birktoft and Banaszak [1]. Amino acid side chains predicted to be involved in the catalytic mechanism are found in the primary structure of the watermelon mMDH and marked C in Fig. 4. Also indicated in Fig. 4 are the residues attributed to the helices forming the subunit interface. These are designated Q and seem to fall frequently into conserved epitopes.

The cDNA of the watermelon mMDH identifies a mitochondrial presequence of 27 amino acids (Fig. 3). Mitochondrial target sequences for matrix enzymes have recently been compiled by Hartl *et al.* [10] and by von Heijne *et al.* [12] and it is concluded that no general significant sequence homologies can be deduced. However, several common features of the primary structures of mitochondrial presequences can be listed: (1) they are rich in positively charged amino acids (mainly arginines); (2) they lack to a large extent acidic residues, and (3) they have a relatively high content of hydroxylated amino acid side chains.

A comparison of the watermelon presequence with that of mMDH of rat, mouse and yeast corroborates these features:

				5			10			15			20		25		
				+	+	.	+	.	+	.	.	.	.	+	.	+	↓
mMDH watermelon	M	K	A	S	I	L	R	S	V	R	S	A	V	S	R	S	S
	+	.	.	.	.	.	+	.	+	+	.	.	.	.	.	.	.
rat [9]	M	L	S	A	L	A	R	P	V	G	A	A	L	R	R	S	F
	+	.	.	.	.	.	+	.	+	+	.	.	.	.	.	.	.
mouse [15]	M	L	S	A	L	A	R	P	A	G	A	A	L	R	R	S	F
	+	.	.	.	.	.	+	.	+	+	.	.	.	.	.	.	.
yeast [35]							+	.	+	+	+	.	.	.	.	.	.
							M	L	S	R	V	A	K	R	A	F	S

+ = basic; . = hydroxylated

It has been repeatedly found that the spacing of the arginines and the hydroxylated residues are of significance for the mitochondrial import of a specific protein. In this respect the similarity of the spacing between the arginines in the targeting sequences of watermelon, rat and mouse is remarkable as is the heaping of serine and threonine residues C-terminal to the conserved arginine in position 15. The presequence of yeast shows this arginine and the aliphatic residues. Thompson and McAlister-Henn [34] have recently shown that the presequence in yeast is dispensable for the import of mMDH into the mitochondrial matrix. The MDH 1 gene of a yeast strain was inactivated by gene disruption and the strain then transformed with a MDH 1 gene in which the nucleotides encoding the presequence had been deleted. Normal amounts of the mMDH were restored inside the mitochondria in the cells transformed with the truncated gene. This is in contrast to studies with the transit peptide of rat mMDH. Chu *et al.* [4] found that replacement of Leu-13 with a charged residue such as Glu or Asp abolishes import of the precursor into the mitochondria. It is further shown by site-directed mutants that this residue is important for binding of the precursor to the outer mitochondrial membrane. In a subsequent paper [32] it is demonstrated that the presequence is processed in two steps, the first cleavage site being after Arg-15. With the clone for the higher plant mMDH available it will be of interest to study the importance of its presequence for import into higher plant and yeast mitochondria. Schmitz and Lonsdale [31] demonstrated that the putative presequence of the tryptophanyl-

tRNA-synthetase of yeast mitochondria comprising 35 amino acids [21] can target the bacterial β -glucuronidase equally well into tobacco plant and yeast cell mitochondria. Processing of the imported precursor occurred at the appropriate cleavage site.

Sequence analysis of a number of PCR products from the 3' non-translated region of the mMDH messenger RNA in the cotyledons of watermelon revealed two major adenylation sites at nucleotide 1337 and 1413 and 5 minor ones between nucleotides 1313 and 1337 (Fig. 3). This is comparable to the polymorphisms for poly(A)⁺ additions encountered in the mRNA encoding alcohol dehydrogenase 1 in maize [27]. By S1 mapping four major alternative poly(A)⁺ addition sites were identified in the *Adh1-1S* allele and seven in the *Adh1-1F* allele. In the *Adh1* gene the canonical adenylation signal AATAAA does not occur within a reasonable distance upstream of the adenylation sites, a situation parallel to the one encountered with the watermelon mMDH gene.

Acknowledgement

The authors would like to thank Prof. F. Lottspeich, Max Planck Institut für Biochemie, München-Martinsried for the amino acid sequence analysis of cyanogen bromide fragments and Prof. D. von Wettstein and Dr K.K. Thomsen for support and encouragement. The work was supported by grant Gi 154/2-1 of the Deutsche Forschungsgemeinschaft to C.G.

References

1. Birktoft JJ, Banaszak J: The presence of a histidine-aspartic acid pair in the action site of 2-hydroxyacid dehydrogenases. X-ray refinement of cytoplasmic malate dehydrogenase. *J Biol Chem* 258: 472–482 (1983).
2. Birktoft JJ, Bradshaw RA, Banaszak LJ: Structure of porcine heart cytoplasmic malate dehydrogenase: Combining x-ray diffraction and chemical sequence data in structural studies. *Biochemistry* 26: 2722–2734 (1987).
3. Birktoft JJ, Fernley RT, Bradshaw RA, Banaszak LJ: Amino acid sequence homology among the 2-hydroxy acid dehydrogenases: Mitochondrial and cytoplasmic malate dehydrogenases form a homologous system with lactate dehydrogenase. *Proc Natl Acad Sci USA* 79: 5166–5170 (1982).
4. Chu TW, Grant PM, Strauss A: Mutation of a neutral amino acid in the transit peptide of rat mitochondrial malate dehydrogenase abolishes binding and import. *J Biol Chem* 262: 15759–15764 (1987).
5. Dittrich P, Raschke K: Malate metabolism in isolated epidermis of *Commelina communis* L. in relation to stomatal functioning. *Planta* 134: 77–81 (1977).
6. Gietl C, Hock B: Organell-bound malate dehydrogenase isoenzymes are synthesized as higher molecular weight precursors. *Plant Physiol* 70: 483–487 (1982).
7. Gietl C, Hock B: Glyoxysomal and mitochondrial malate dehydrogenase in watermelon cotyledons. In: Rattazzi MC, Scandalios IG, Whitt GS: *Isozymes: Current Topics in Biological and Medical Research*, vol. 16, pp. 175–192. Alan R. Liss, Inc., New York (1987).
8. Gietl C, Lottspeich F, Hock B: Sequence homologies between glyoxysomal and mitochondrial malate dehydrogenase. *Planta* 169: 555–558 (1986).
9. Grant PM, Tellam J, May VL, Strauss AW: Isolation and nucleotide sequence of a cDNA clone encoding rat mitochondrial malate dehydrogenase. *Nucl Acids Res* 14: 6053–6066 (1986).
10. Hartl F-V, Pfanner N, Nicholson DW, Neupert W: Mitochondrial protein import. *Biochim Biophys Acta* 988: 1–45 (1989).
11. Hattori M, Sakaki Y: Dideoxy sequencing method using denatured plasmid templates. *Anal Biochem* 152: 232–238 (1986).
12. Heijne G von, Steppuhn J, Herrmann RG: Domain structure of mitochondrial and chloroplast targeting peptides. *Eur J Biochem* 180: 535–545 (1989).
13. Hock B: Die Hemmung der Isocitrat-Lyase bei Wassermelonenkeimlingen durch Weisslicht. *Planta* 85: 340–350 (1969).
14. Joh T, Takeshima H, Tsuzuki T, Setoyama C, Shimada K, Tanase S, Kuramitsu S, Kagamiyama H, Morino Y: Cloning and sequence analysis of cDNAs encoding mammalian cytosolic malate dehydrogenase. *J Biol Chem* 262: 15127–15131 (1987).
15. Joh T, Takeshima H, Tsuzuki T, Shimada K, Tanase S, Morino Y: Cloning and sequence analysis of cDNA encoding mammalian mitochondrial malate dehydrogenase. *Biochemistry* 26: 2515–2520 (1987).
16. Kindl H: *Biochemie der Pflanzen*. Springer-Verlag Berlin/Heidelberg/New York (1987).
17. Loh EY, Elliott JF, Cwirla S, Lanier LL, Davis MM: Polymerase chain reaction with single sided specificity: analysis of T cell receptor Δ chain. *Science* 243: 217–243 (1989).
18. Maniatis T, Fritsch EF, Sambrook J: *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1983).
19. McAlister-Henn L: Evolutionary relationships among the malate dehydrogenases. *TIBS* 13: 178–181 (1988).
20. McAlister-Henn L, Blaber M, Bradshaw RA, Nisco SJ: Complete nucleotide sequence of *Escherichia coli* gene encoding malate dehydrogenase. *Nucleic Acids Res* 15: 4993 (1987).
21. Myers AM, Tzagoloff A: *MSW*, a yeast gene coding for mitochondrial tryptophanyl-tRNA synthetase. *J Biol Chem* 260: 15371–15377 (1985).
22. Nicholls DJ, Minton NP, Alkinson T, Sundaram TK: Cloning sequencing and over-expression of *Escherichia coli* malate dehydrogenase. *Appl Microbiol Biotechnol* 31: 376–382 (1989).
23. Nishiyama M, Matsubara N, Yamamoto K, Iijima S, Vozumi T, Beppu T: Nucleotide sequence of the malate dehydrogenase gene of *Thermus flavus* and its mutation directing an increase in enzyme activity. *J Biol Chem* 261: 14178–14183 (1986).
24. Queen C, Korn LJ: A comprehensive sequence analysis program for the IBM personal computer. *Nucleic Acids Res* 12: 581–599 (1984).
25. Raschke K: Stomatal action. *Ann Rev Plant Physiol* 26: 309–340 (1975).
26. Rutledge RG, Seligy VL, Côté M-J, Dimack K, Lewin LL, Tenniswood MP: Rapid synthesis and cloning of complementary DNA from any RNA molecule into plasmid and phage M13 vectors. *Gene* 68: 1251–158 (1985).
27. Sachs MM, Dennis ES, Gerlach WL, Peacock WJ: Two alleles of maize *alcohol dehydrogenase 1* have 3' structural and poly(A) addition polyporphisms. *Genetics* 113: 449–467 (1986).
28. Saiki RK, Gelford DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA: Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–490 (1988).
29. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463–5467 (1977).
30. Sautter C, Hock B: Fluorescence immunohistochemical localization of malate dehydrogenase isoenzymes in watermelon cotyledons. *Plant Physiol* 70: 1162–1168 (1982).
31. Schmitz UK, Lonsdale DM: A yeast mitochondrial pre-sequence functions as a signal for targeting to plant mitochondria in vivo. *Plant Cell* 1: 783–791 (1989).
32. Sztu ES, Chu TW, Strauss A, Rosenberg LE: Import of

- the malate dehydrogenase precursor by mitochondria. *J Biol Chem* 263: 12085–12091 (1988).
33. Takeshima H, Joh T, Tsuzuki T, Shimada K, Matsukado Y: Structural organization of the mouse malate dehydrogenase gene. *J Mol Biol* 200: 1–11 (1988).
34. Thompson LM, McAlister-Henn L: Dispensable presequence for cellular localization and function of mitochondrial malate dehydrogenase from *Saccharomyces cerevisiae*. *J Biol Chem* 264: 12091–12096 (1989).
35. Thompson LM, Sutherland P, Steffan JS, McAlister-Henn L: Gene sequence and primary structure of mitochondrial malate dehydrogenase from *Saccharomyces cerevisiae*. *Biochemistry* 27: 8393–8400 (1988).
36. Vieira J, Messing Y: The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259–268 (1982).
37. Walk R-A, Hock B: Cell free synthesis of glyoxysomal malate dehydrogenase. *Biochem Biophys Res Commun* 81: 636–643 (1977).
38. Wilks HM, Hart KN, Feeney R, Dunn CR, Muirhead H, Chia WN, Barstow DA, Atkinson T, Clarke AR, Holbrook JJ: A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. *Science* 242: 1541–1544 (1988).
39. Zhang H, Scholl R, Browse J, Sommerville C: Double stranded sequencing as a choice for DNA sequencing. *Nucleic Acids* 16: 1220 (1988).